MATEMATYKA UBEZPIECZEŃ MAJĄTKOWYCH I OSOBOWYCH

RYSZARD SZEKLI

Uniwersytet Wrocławski
2018

Spis treści

Wstęp

Skrypt jest przeznaczony dla studentów kierunku matematyka na Wydziale Matematyki i Informatyki UWr.

Dla wygody wiele używanych faktów z teorii prawdopodobieństwa znajduje się w Dodatku.

Rozdziały oznaczone * wymagają znajomości bardziej zaawansowanych narzędzi rachunku prawdopodobieństwa spoza kursu rachunku prawdopodobieństwa A.

Kursywą podana jest terminologia angielska.

Kurs zawiera matematyczne podstawy i klasyczne metody używane w zawodzie *aktuariusza*.

Specjalistą w zakresie oszacowania ryzyka jest *aktuariusz*. Miejscem pracy aktuariusza mogą być wszystkie instytucje finansowe, w których zarządza się ryzykiem. W Polsce istnieje wciąż zapotrzebowanie na aktuariuszy.

Aktuariusz to specjalista ubezpieczeniowy, który oszacowuje za pomocą metod matematyki aktuarialnej, wysokość składki, świadczeń, odszkodowań, rezerw ubezpieczeniowych. Aktuariusze w oparciu o dane historyczne, regulacje prawne i prognozy dokonują kalkulacji prawdopodobieństw zdarzeń losowych. Oszacowują również ryzyko powstania szkód majątkowych. Aktuariusz przypisuje finansową wartość przyszłym zdarzeniom.

Korzenie zawodu aktuariusza sięgają przełomu XVII i XVIII w. i były powiązane przede wszystkim z rozwojem ubezpieczeń na życie, ale głównego znaczenia profesja ta nabrała dopiero w XIX w. Matematykę aktuarialną zapoczątkowały pod koniec XVII w. prace angielskiego astronoma E. Halleya dotyczące wymieralności w wybranej populacji, a w 1948 r. w Londynie powstał Instytut Aktuariuszy - pierwsza placówka naukowa prowadząca prace z zakresu matematyki aktuarialnej.

W Polsce za początek zawodu aktuariusza można uznać rok 1920, w którym działalność rozpoczął Polski Instytut Aktuariuszy. Środowisko aktuariuszy w 1991 r. powołało Polskie Stowarzyszenie Aktuariuszy. Zadaniem Stowarzyszenia jest wspieranie tej grupy zawodowej, a także uczestnictwo w pracach legislacyjnych w zakresie ubezpieczeń. Stowarzyszenie jest członkiem Międzynarodowego Stowarzyszenia Aktuariuszy.

Sektor towarzystw ubezpieczeniowych, zarówno na życie jak i majątkowo-osobowych, nie może funkcjonować bez aktuariuszy, którzy w większości właśnie tam pracują. Zgodnie z Ustawą o działalności ubezpieczeniowej z 11 września 2015, (zob. szczegóły na $www.knf.gov.pl/dla_rynku/egzaminy$ w zakładce egzamin na aktuariusza)

do zadań aktuariusza w Polsce należy:

- ustalanie wartości rezerw techniczno-ubezpieczeniowych,

- kontrolowanie aktywów stanowiących pokrycie rezerw techniczno-ubezpieczeniowych,

- wyliczanie marginesu wypłacalności,

- sporządzanie rocznego raportu o stanie portfela ubezpieczeń,

- ustalanie wartości składników zaliczanych do środków własnych.

Aktuariusze mogą pracować we wszystkich instytucjach finansowych zarządzających ryzykiem. Mogą pracować w firmach konsultingowych, udzielając porad w zakresie podejmowania decyzji finansowych. W szczególności pomagają zaprojektować programy emerytalne, a w trakcie ich działania wyceniają ich aktywa i zobowiązania. Aktuariusze mogą również oszacowywać koszt różnego rodzaju ryzyk w działalności przedsiębiorstw. Mogą pracować również w instytucjach państwowych związanych np. z systemem ubezpieczeń społecznych czy zdrowotnych.

Ponadto aktuariusze mogą znaleźć zatrudnienie wszędzie tam, gdzie konieczne jest rozwiązywanie problemów finansowych i statystycznych - banki i firmy inwestycyjne, duże korporacje, związki zawodowe.

Zgodnie z ustawą o działalności ubezpieczeniowej, aktuariuszem może zostać osoba fizyczna, która:

- ukończyła studia wyższe,

- przez okres co najmniej 2 lat wykonywała czynności z zakresu matematyki ubezpieczeniowej, finansowej i statystyki, pod kierunkiem aktuariusza,

- złożyła z pozytywnym wynikiem egzamin aktuarialny,

- posiada pełną zdolność do czynności prawnych,

- korzysta z pełni praw publicznych,

- nie była prawomocnie skazana za umyślne przestępstwo przeciwko wiarygodności dokumentów, przestępstwo przeciwko mieniu lub za przestępstwo skarbowe.

Jednym z powyższych wymogów dla uzyskania licencji aktuariusza jest zdanie egzaminu aktuarialnego. Zgodnie z rozporządzeniem Ministra Finansów z 23 kwietnia 2015 r. w sprawie zakresu obowiązujących tematów egzaminów aktuarialnych oraz trybu przeprowadzania tych egzaminów zakres tego egzaminu obejmuje cztery działy:

- matematykę finansową,

- matematykę ubezpieczeń na życie,

- matematykę pozostałych ubezpieczeń osobowych i majątkowych,

- prawdopodobieństwo i statystykę.

Egzaminy są organizowane co najmniej 2 razy w roku kalendarzowym. Każda część egzaminu składa się z 10 pytań. Każde pytanie oceniane jest według następującej skali:

- dobra odpowiedź: 3 punkty, - błędna odpowiedź: -2 punkty, - brak odpowiedzi: 0 punktów.

Egzamin uważa się za zaliczony po uzyskaniu 13 punktów z jednej części. Zaliczenie wszystkich działów nie może trwać dłużej niż 2 lata.

Aktuariuszem najczęściej mogą zostać osoby w wykształceniem matematycznym lub ekonomicznym.

Jednym z głównych zadań w działalności firm ubezpieczeniowych jest dbałość o wypłacalność. Na firmy ubezpieczeniowe nałożone jest wiele wymogów zapewniających bezpieczeństwo działalności ubezpieczeniowej. Działalność ubezpieczeniowa ze względu na swoje społeczne i gospodarcze znaczenie została poddana nadzorowi wyspecjalizowanego organu administracji państwowej.

Wypłacalność to zdolność firmy do spłaty zobowiązań w terminie. Jest podstawowym kryterium oceny kondycji finansowej zakładu ubezpieczeń.

Jeden z podstawowych wymogów działalności ubezpieczeniowej dotyczy marginesu wypłacalności. Margines wypłacalności jest to określona przepisami prawa wielkość środków własnych zakładu ubezpieczeń, która ma na celu zapewnienie wypłacalności i nie może być niższa od minimalnej wysokości kapitału gwarancyjnego.

Wymogi dotyczące marginesu wypłacalności dla zakładów ubezpieczeń zostały wprowadzone w 1973 roku.

Wraz z rozwojem rynku ubezpieczeniowego, pojawieniem się nowych produktów oraz ryzyk istniejące wymogi przestały w pełni odzwierciedlać wszystkie ryzyka, na które były narażone firmy ubezpieczeniowe. Dotyczyło to głównie ryzyk finansowych np. ryzyka zmiany stóp procentowych. Pomimo spełniania istniejących wymogów wypłacalności przez firmy ubezpieczeniowe, kondycja finansowa tych firm pogarszała się. Obowiązujące wymogi wypłacalności nie spełniały już oczekiwań związanych z zapewnieniem bezpieczeństwa działalności ubezpieczeniowej. Nie bez znaczenia był również fakt coraz większego skupienia działalności ubezpieczeniowej wokół międzynarodowych grup kapitałowych.

Pierwszym krokiem w kierunku poprawienia systemu badania wypłacalności było wprowadzenie Solvency I. W prawie polskim Solvency I zwiększyło wysokość minimalnego kapitału gwarancyjnego dla spółek akcyjnych z grupy I (ubezpieczenia na życie) z 800 tys. euro do 3 mln euro, dla działu II (ubezpieczenia majątkowe) grup 1-9 oraz 16-18 z 300 tys. euro i 200 tys. euro do 2 mln euro. Wprowadzono również coroczną indeksację minimalnego kapitału gwarancyjnego.

Zmieniająca się rzeczywistość finansowa i gospodarcza wymusiła debatę nad zmianami w nowym systemie wypłacalności zakładów ubezpieczeń. Wykonano szereg analiz ryzyk działalności ubezpieczeniowej, analiz bankructw, analiz istniejących modeli wypłacalności wdrożonych w innych krajach. Wynikiem tych działań miało być powstanie nowego systemu badania wypłacalności Solvency II. Został on zapoczątkowany w 2001 roku przez Komisje Europejską w ramach Komitetu Europejskiego.

U podstaw dyskusji nad koniecznością wprowadzenia Solvency II leży szereg niedoskonałości w istniejących regulacjach dotyczących wypłacalności. Spośród nich należy tu chociażby wymienić metody bazujące na składce, które nie uwzględniają istotnych ryzyk; brak

uwzględnienia kompletnych form transferu ryzyka, brak uwzględnienia zależności pomiędzy aktywami i pasywami oraz zakresem prowadzonej działalności.

Nowo powstający system Solvency II ma być uniwersalny i ma objąć wszystkie firmy ubezpieczeniowe prowadzące działalność na terenie UE. Jest on wzorowany na Bazylei II, która określa zasady wypłacalności dla banków.

Nowy system oceny wypłacalności zgodny z Solvency II ma być dopasowany do rzeczywistych ryzyk, na jakie narażony jest zakład ubezpieczeń.

W przypadku instytucji ubezpieczeniowej potencjalne ryzyka są specyficzne dla typów zawieranych umów ubezpieczenia w zakresie ubezpieczeń na życie lub ubezpieczeń majątkowych. Umiejętność skutecznej identyfikacji, oceny i monitorowania ryzyk może uchronić przed znacznymi stratami. Kluczową rolę odgrywają tu przyjęte metodologie zarządzania ryzykiem, służące eliminacji ich negatywnego wpływu na wyniki finansowe.

Ryzyka, na które jest narażony zakład ubezpieczeń można podzielić na *ryzyka aktuarialne* związane z przyszłymi wynikami technicznymi zależnymi od czynników losowych częstości, intensywności szkód, kosztów operacyjnych, zmian w składzie portfela wypowiedzeń bądź konwersji umów ubezpieczenia oraz *ryzyka finansowe* ryzyka, na które jest narażona każda instytucja finansowa, (np. bank), do tej grupy zaliczają się ryzyka takie jak: ryzyko zmian stopy procentowej, ryzyko kredytowe, ryzyko rynkowe, ryzyko walutowe.

Większa uwaga nadzoru ubezpieczeniowego ma skupić się na kontroli sposobów zarządzania ryzykiem przez firmy ubezpieczeniowe, jak również na poprawności przyjętych w tym zakresie założeń. Idea Solvency II polega na ściślejszym uzależnieniu wysokości kapitału od wielkości ryzyka podejmowanego przed firmy ubezpieczeniowe. Ujednoliceniu mają być poddane sposoby raportowania firm ubezpieczeniowych w różnych krajach. Solvency II ma mieć o wiele większy zakres od Solvency I, ma uwzględnić, bowiem wpływ nowych tendencji z zakresu metodologii zarządzania ryzykiem w ubezpieczeniach, szeroko pojętej inżynierii finansowej oraz standardów sprawozdawczości zgodnych z wymogami IASB (International Accounting Standard Board). Pierwszorzędnymi zamierzeniami projektu jest znalezienie wymogu marginesu wypłacalności oraz osiągnięcie większej synchronizacji w ustalaniu poziomu rezerw technicznych.

Znaczącą rolę techniczną w ramach Solvency II odgrywają miary ryzyka takie jak VaR, TVaR, CVaR itp. oraz kopuły (copulas), które będą omówione w obecnym skrypcie.

Rozdział 1

Wprowadzenie

Zawód aktuariusza jest jednym z najstarszych w świecie finansów. Historia tego zawodu rozpoczyna się w połowie dziewietnastego wieku wraz z ubezpieczeniami na życie i aż do lat sześćdziesiątych dwudziestego wieku matematyczne metody aktuariusza związane były z wyceną kontraktów ubezpieczeniowych, tworzeniem tablic przeżycia na podstawie danych statystycznych oraz z wyliczniem rezerw pieniężnych firmy. W latach sześćdziesiątych rozpoczęto stosowanie matematycznych metod do stworzenia teorii ryzyka na użytek ubezpieczeń majątkowych i osobowych. Punktem wyjścia był standardowy złożony proces Poissona, którego pomysł pochodzi od Filipa Lundberga z 1903 roku, a który matematycznie został opracowany przez Haralda Cramera w latach trzydziestych. Do lat dziewięćdziesiątych był on rozwijany na różne sposoby. Proces Poissona został zastąpiony przez proces odnowy oraz przez proces Coxa, następnie użyto procesów Markowa kawałkami deterministycznych, wreszcie wprowadzono losowe otoczenie pozwalające na modelowanie losowych zmian w intensywności zgłoszeń szkód i wielkości szkód. Pojawia się wiele książek z teorii ryzyka (zob. listę referencji). Jednym z najbardziej *matematycznie* interesujących zagadnień w teorii ryzyka jest zagadnienie ruiny, gdzie czasy pierwszego przekroczenia wysokiego poziomu rezerwy kapitałowej są w centrum uwagi. Stare i nowe rezultaty na tym polu mogą być wytłumaczone przez teorię martyngałów i użyte do pokazania nierówności Lundberga dla bardzo ogólnych modeli dowodząc, iż dla małych szkód prawdopodobieństwo ruiny dąży do zera wykładniczo szybko wraz z rezerwą początkową. Specjalna teoria pojawia się dla szkód potencjalnie dużych. Warunkowe twierdzenia graniczne pozwalają zrozumieć trajektorie prowadzące do ruiny. Interesujący rozkwit metod matematycznych w latach dziewięćdziesiątych dokonał się głównie z dwóch przyczyn: wzrostu szkód związanych z katastrofami oraz z gwałtownego rozwoju rynków finansowych.

Wielkie katastrofy i szkody lat siedemdziesiątych i osiemdziesiątych spowodowały przekroczenia rezerw na rynku ubezpieczeń pierwotnych i wtórnych. Szybko rosnący rynek finansowy w tym czasie poszukiwał nowych możliwości inwestycyjnych również w zakresie przyjmowania zakładów w zakresie naturalnych katastrof takich jak trzęsienia ziemi i huragany. Częstość występowania i rozmiary wielkich szkód stworzyły potrzebę wprowadzenia wyszukanych modeli statystycznych do badania procesu szkód. **Teoria wartości**

ekstremalnych dostarcza niezbędnych matematycznych narzędzi do wprowadzenia nowych metod. Pojawiają się książki w zakresie teorii wartości ekstremalnych w kontekście problematyki ubezpieczeniowej.

W latach osiemdziesiątych banki inwestycyjne dostrzegają, iż zabezpieczanie się przed ryzykiem finansowym nie jest wystarczające ze względu na dodatkowe ryzyka rynkowe. Tak zwany traktat z Bazylei z roku 1988 z poprawkami z lat 1994-1996, wprowadza tradycyjne metody ubezpieczeniowe budowania rezerw do sfery ryzyka bankowego. Rezerwy muszą być tworzone na pokrycie tzw. earning at risk, to znaczy różnicy między wartością średnią, a kwantylem jednoprocentowym rozkładu zysku/straty (profit/loss). Wyznaczenie tak małego kwantyla wymaga bardzo specjalnych metod statystycznych. Metody aktuarialne stosowane są również do modelowania ryzyka kredytowego. Portfele kredytowe są porównywalne z portfelami ryzyk ubezpieczeniowych. Przyszły rozwój metod ubezpieczeniowych związany jest z powstawaniem złożonych rynków ubezpieczeniowych, firmy ubezpieczeniowe oczekują elastycznych rozwiązań zapewniających pomoc w całościowym podejściu do zarządzania ryzykiem.

Całkiem naturalnie na tym tle wprowadzane są metody pochodzące z teorii stochastycznej optymalizacji. Wiele zmiennych kontrolnych takich jak wielkość reasekuracji, dywidendy, inwestycje są badane łącznie w sposób dynamiczny prowadząc do równań Hamiltona-Jakobiego-Bellmana, rozwiązywanych numerycznie.

Po tym krótkim nakreśleniu historii rozwoju metod matematycznych w ubezpieczeniach wracamy do podstawowego modelu. Pomyślmy o konkretnej sytuacji. Przeglądając wszystkie polisy ubezpieczeniowe, zakupione w jednej firmie ubezpieczeniowej, które ubezpieczają skutki pożaru mieszkań w pewnej dzielnicy dużego miasta, najprawdopodobniej natkniemy się na porównywalną wartość ubezpieczanych dóbr oraz możemy przyjąć, iż szanse na pożar w poszczególnych budynkach są podobne. Taki zbiór polis tworzy **jednorodny portfel** ubezpieczeniowy. Większość firm ubezpieczeniowych używa tego rodzaju portfeli jako podstawowych cegiełek swej działalności. Cegiełki takie, odpowiednio ułożone, tworzą większe bloki działalności takie jak ubezpieczenia od ognia, ubezpieczenia ruchu drogowego, ubezpieczenia przed kradzieżami, ubezpieczenia majątkowe itd. Blok ubezpieczeń od ognia zawiera wtedy wiele portfeli różniących się rodzajami ryzyka, na przykład dla: wolno stojących domów, domów szeregowych, budynków wielomieszkaniowych, sklepów, marketów itd., które wymagają osobnego określenia **ryzyka ubezpieczeniowego** dla każdego rodzaju i wyliczenia innej składki ubezpieczeniowej, choćby z tego tylko powodu, iż rozmiar szkody w poszczególnych portfelach może być nieporównywalny. W dalszym ciągu skupiać będziemy naszą uwagę na analizie pojedynczych portfeli, które składać się będą z wielu elementów natury losowej lub deterministycznej. Podstawowym parametrem portfela jest czasokres w którym ubezpieczone ryzyka mogą generować szkody. Zwykle dane odnoszące się do danego portfela obejmują okres jedengo roku. Kluczowym parametrem jest rezerwa początkowa (kapitał początkowy), wyznaczany na początku czasokresu w celu pokrycia kosztów wynikających ze zgłoszonych szkód w portfelu. Same zgłoszenia wyznaczone są przez chwile zgłoszeń $T_1 < T_2 < T_3 < ...$, przy czym wygodnie jest przyjąć iż $T_0 = 0 < T_1$. Liczbę zgłoszeń do chwili $t > 0$ definiujemy przez $N(t) = \max\{n : T_n \leqslant t\}$. Każde zgłoszenie związane jest z wielkością zgłaszanej szkody oznaczanej przez X_n , dla

$n-$tego zgłoszenia. Przy tych oznaczeniach całkowita wartość szkód zgłoszonych do chwili t równa się $S(t) = \sum_{i=1}^{N(t)} X_i$. (Przyjmujemy $S(t) = 0$, gdy $N(t) = 0$). Oznaczmy przez $H(t)$ wartość składek zebranych w portfelu do chwili t. Zwykle przyjmujemy, że $H(t) = ct$, dla pewnej stałej wartości $c > 0$. Wtedy rezerwa kapitału w portfelu, przy założeniu, że kapitał początkowy wynosi u, wyraża się wzorem $R(t) = u + H(t) - S(t)$. Zakładając, że momenty zgłoszeń oraz wielkości szkód są zmiennymi losowymi, możemy interpretować kolekcję zmiennych $(R(t), t > 0)$ jako proces stochastyczny. (Jest to tak zwany **proces ryzyka**). Badanie procesu ryzyka jest centralnym zagadnieniem tak zwanej teorii ryzyka, która z kolei stanowi niewątpliwie jądro matematyki ubezpieczeniowej poświęconej ubezpieczeniom majatkowym i osobowym.

Nakreślimy teraz bliżej zestawy założeń przyjmowanych o zmiennych losowych tego modelu, które umożliwiają dokładniejszą analizę portfeli.

Rozpoczniemy od podania detali dotyczących ciągu zgłoszeń. O zmiennych losowych $T_1, T_2, ...$można przyjąć wiele różnych założeń. W pewnych szczególnych przypadkach użytecznym i odpowiednim założeniem jest to, iż ciąg ten tworzy **proces odnowy**, tzn. ciąg zmiennych losowych odstępów między zgłoszeniami $W_i = T_i - T_{i-1}, i = 1, 2, ...$, jest ciągiem niezależnych zmiennych losowych o jednakowych rozkładach. Taki proces zgłoszeń jest elementem składowym modelu *Sparre Andersena*, który będzie opisany detalicznie później. Klasycznym przykładem procesu odnowy jest **proces Poissona,** w którym odstępy między zgłoszeniami mają rozkład wykładniczy. Ponieważ rozkład wykładniczy jako jedyny ma własność **braku pamięci**, proces Poissona ma wiele strukturalnych własności odróżniających go od innych procesów. (Własność braku pamięci rozkładu wykładniczego jest zdefiniowana przez równość $P(W > x + y \mid W > y) = P(W > x)$, dla $x, y > 0$ lub równoważnie $P(W > x + y) = P(W > x)P(W > y)$). Na przykład, dla procesu Poissona $P(N(t) = k) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}, k = 0, 1, ...$, gdzie $0 < \lambda = (EW)^{-1}$, przy tym, $EN(t) = \lambda t = VarN(t)$. Ponadto liczby zgłoszeń w rozłącznych przedziałach czasowych w procesie Poissona tworzą kolekcję niezależnych zmiennych losowych.

W praktyce aktuarialnej zauważono już dawno, iż stosunek wartości oczekiwanej do wariancji w procesach zgłoszeń $(N(t), t > 0)$ bardzo często nie jest równy jeden (tak jest w procesie Poissona). Można to wytłumaczyć tym, że indywidualne szkody w portfelu są zgłaszane zgodnie z procesem Poissona o pewnej wartości średniej, lecz wartość średnia ilości indywidualnych zgłoszeń może być różna dla każdej z polis w portfelu. Takie założenie prowadzi do procesu zgłoszeń dla którego $P(N(t) = k) = \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^k}{k!}dF(\lambda)$, gdzie F jest pewną dystrybuantą określającą rozkład parametru λ w zbiorze możliwych wartości w danym portfelu (zakładamy zawsze, że $\lambda > 0$). Wygodnie jest przyjąć, że istnieje zmienna losowa Λ określajaca losową wartość parametru λ, spełniająca $P(\Lambda \leqslant \lambda) = F(\lambda)$. Zakładamy przy tym, że Λ jest zmienną losową niezależną od indywidualnych procesów Poissona. Proces $(N(t), t > 0)$ spełniający te założenia jest tak zwanym **mieszanym Procesem Poissona**. Szczególny przypadek, gdy Λ ma rozkład gamma, odpowiada tak zwanemu **procesowi Polya.**

Inna użyteczna klasa procesów zgłoszeń jest wyznaczona związkiem rekurencyjnym postaci $P(N(t) = k) = (a + \frac{b}{k})P(N(t) = k - 1)$,dla $k = 1, 2, ...$ oraz pewnych stałych a, b (być

może zależnych jedynie od t). Rozkład geometryczny, dwumianowy i Poissona znajdują się w tej klasie, przy odpowiedniej specyfikacji stałych a, b. Dla takich procesów Panjer pokazał użyteczną rekurencję pozwalającą wyznaczyć rozkład całkowitej wartości szkód w portfelu.

Wspomniana wcześniej własność procesu Poissona, iż liczby zgłoszeń w rozłącznych przedziałach czasowych tworzą kolekcję niezależnych zmiennych losowych stanowi punkt wyjścia do teorii procesów o niezależnych przyrostach. Procesy zgłoszeń posiadające tę własność są procesami, dla których $P(N(t) = k) = \sum_{i=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^i}{i!} p_k^{*i}$, gdzie p_k^{*i} oznacza $i-$krotny splot funkcji prawdopodobieństwa $(p_k, k = 0, 1, ...)$. Oznacza to, że liczbę zgłoszeń można zapisać w postaci $N(t) = \sum_{i=1}^{K(t)} Y_i$, gdzie $(K(t), t > 0)$ jest Procesem Poissona niezależnym od ciągu zmiennych $(Y_i, i = 1, 2, ...)$, które są z kolei wzajemnie niezależne o jednakowym rozkładzie $(p_k, k = 0, 1, ...)$ Takie procesy są **złożonymi procesami Poissona.**

Podstawowym założeniem o wielkościach zgłaszanych szkód w portfelu jest to, iż tworzą one ciąg $X_1, X_2, ...$ niezależnych zmiennych losowych o jednakowych rozkładach. W zasadzie każda dystrybuanta skoncentrowana na $[0, \infty)$ może być użyta do określenia rozkładu wielkości szkód, jednakże często odróżnia się dystrubuanty o **lekkich i cieżkich ogonach**. Dystrybuanty o lekkich ogonach są asymptotycznie równoważne rozkładowi wykładniczemu. Dystrybuanty o ciężkich ogonach służą do modelowania szkód, które mogą osiągać wartości relatywnie bardzo duże z istotnymi prawdopodobieństwami (tak jak się zdarza w przypadku portfeli ubezpieczeń od pożarów). Typowym rozkładem ciężkoogonowym używanym w praktyce jest rozkład Pareto.

Łatwo wyobrazić sobie sytuacje, w których proces zgłoszeń $(N(t), t > 0)$ i ciąg wielkości zgłaszanych szkód $(X_n, n = 1, 2, ...)$ są zależne, jak na przykład w przypadku szkód wynikających z wypadków drogowych, kiedy to intensywność zgłoszeń jak również rozmiar szkód zależą od warunków drogowych związanych z porą roku. Obliczenie rozkładu całkowitej wartości szkód jest w tym przypadku możliwe jedynie w bardzo specjalnych przypadkach. Dlatego przyjmuje się bardzo często, że $(N(t), t > 0)$ oraz $(X_n, n = 1, 2, ...)$ są niezależne. Nawet przy tym założeniu wyliczenie rozkładu $S(t)$ nie jest łatwym zadaniem. Podstawowym wzorem w tym przypadku jest $P(S(t) \leqslant x) = \sum_{i=0}^{\infty} P(N(t) = i) F_X^{*i}(x)$, gdzie $F_X(x) = P(X_1 \leqslant x)$. Jak widzimy potrzebne są sploty F_X^{*i}, dla których proste wzory są znane jedynie w nielicznych przypadkach. Z tego powodu musimy zdać się często na aproksymacje. W przypadku, gdy liczba zgłoszeń jest duża a rozkłady mają skończone wariancje można będzie zastosować Centralne Twierdzenie Graniczne (CTG) i wtedy $P(S(t) \leqslant x) \approx \Phi(\frac{x - ES(t)}{(VarS(t))^{1/2}})$. Aproksymacja tego rodzaju jest bardzo niedokładna, gdy tylko niewielka ilość szkód wyznacza wartość całego portfela (tak jak w przypadku szkód o ciężkich ogonach). Wyznaczenie dobrych aproksymacji w takich przypadkach jest bardzo trudne.

Użyliśmy oznaczenia $H(t)$ dla oznaczenia wielkości składek zebranych w portfelu do chwili t. Zwykle składki pobierane są raz do roku od indywidualnych posiadaczy polis, jednakże wygodniej jest założyc, iż napływ składek odbywa się jednorodnie w ciagu całego roku. Wyznaczenie wielkości $H(t)$ jest jedną z niewielu rzeczy na jakie może wpłynąć ubezpieczający

i musi być dokonane w taki sposób, aby pokryć zobowiązania w portfelu wynikające ze zgłaszanych szkód. Z drugiej strony zawyżanie wysokości składek jest ograniczane konkurencją na rynku ubezpieczeń. Najbardziej popularną formą składki jest $H(t) = (1+\theta)EN(t)EX$, dla pewnej stałej θ odzwierciedlajacej narzut gwarantujący bezpieczeństwo działania (safety loading). Taki sposób naliczania składki nie odzwieciedla losowej zmienności portfela, dlatego alternatywnie używa się wzorów uwzględniających wariancje składowych zmiennych losowych. Jeszcze innym aspektem w trakcie naliczania składek jest fakt, że nie wszyscy indywidualni posiadacze polis w danym portfelu powinni płacić składki w tej samej wysokości oraz składki powinny zależeć od historii indywidualnej polisy.

Rezerwa kapitału $R(t) = u + H(t) - S(t)$ przybiera szczególnie prostą postać, gdy przyjmiemy iż parametr czasu przebiega zbiór liczb naturalnych. Oznaczając wtedy przez H_n składki zebrane w n jednostkach czasu oraz przez S_n sumaryczne szkody zgłoszone w n jednostkach czasu otrzymujemy rezerwę w n tej chwili $R_n = u + H_n - S_n$ (przyjmujemy $S_0 = 0, H_0 = 0$). Przy dodatkowym założeniu, że przyrosty $H_n - H_{n-1}$ oraz $S_n - S_{n-1}$ są wzajemnie niezależne dla $n = 2, 3, ...$, otrzymujemy ciąg $(R_n, n = 0, 1, 2, ...)$ zwany błądzeniem losowym (*random walk*). Ogólnie trajektorie przebiegu w czasie wartości $R(t)$ obrazują zachowanie się losowego procesu, w którym trend dodatni reprezentuje $H(t)$,a trend ujemny $S(t)$. Przedmiotem intensywnych badań teoretycznych jest tak zwane prawdopodobieństwo ruiny w procesie $(R(t), t > 0)$. Jesli przez $\tau = \inf\{t > 0 : R(t) < 0\}$ oznaczymy pierwszą chwilę, gdy rezerwa przyjmie wartość ujemną (tak zwana chwila ruiny), to prawdopodobieństwem ruiny jest $\psi(u) = P(\tau < \infty)$. W przypadku, gdy wielkości szkód mają rozkład lekkoogonowy, można podać aproksymacje i ograniczenia na $\psi(u)$ (będą to wzory oparte o funkcję wykładniczą). W przypadku ciężkich ogonów aproksymacje istnieją dla tak zwanych rozkładów podwykładniczych (*subexponential*).

Rozdział 2

Rozkłady wielkości portfela

Portfelem nazywamy zbiór ryzyk $X = \{X_1, \ldots, X_N\}$ określonego typu, które są zmiennymi losowymi. Podstawową wielkością związaną z portfelem jest **wielkość portfela**, czyli suma zmiennych losowych składających się na portfel $S_N = X_1 + \cdots + X_N$. Mówimy o modelu **prostym**, gdy N jest ustaloną liczbą. Gdy $S = X_1 + \cdots + X_N$, gdzie N jest zmienną losową całkowitoliczbową, to mówimy o modelu **złożonym**. Podstawowym założeniem jest to, że zmienne losowe $(X_i)_{1 \leqslant i \leqslant N}$ są niezależne oraz N jest niezależne od $(X_i)_{i \geqslant 1}$.

2.1 Rozkład wielkości portfela w modelu prostym

Dla prostoty przyjmijmy najpierw $N = 2$ oraz $X_1 = X$, $X_2 = Y$, wtedy $S := S_2 = X + Y$ gdzie X, Y są niezależnymi indywidualnymi szkodami.

Przypadek I. Rozkłady kratowe.

Przyjmijmy na chwilę założenie, że X, Y przyjmują jedynie wartości ze zbioru liczb naturalnych $\mathbb{N} = \{0, 1, \ldots\}$ z prawdopodobieństwami $P(X = i) = p_X(i) \in [0, 1]$, $P(Y = i) = p_Y(i) \in [0, 1]$, $i \in \mathbb{N}$. Przyjmujemy $p_X(s) = p_Y(s) = 0$ dla $s \notin \mathbb{N}$. Stosując wzór na prawdopodobieństwo całkowite, dla $s \in \mathbb{R}$ otrzymujemy

$$\begin{aligned} F_S(s) := P(S \leqslant s) &= \sum_{i=1}^{\infty} P(X + Y \leqslant s | Y = i) P(Y = i) \\ &= \sum_{i=0}^{\infty} P(X \leqslant s - i | Y = i) P(Y = i). \end{aligned}$$

Korzystając z niezależności X i Y otrzymujemy

$$F_S(s) = \sum_{i=0}^{\infty} F_X(s - i) p_Y(i) \tag{2.1.1}$$

oraz

$$p_S(s) = \sum_{i=1}^{\infty} p_X(s-i)p_Y(i). \tag{2.1.2}$$

Zauważmy, że wartości $p_S(s)$ mogą być dodatnie jedynie dla $s \in \mathbb{N}$, dla s spoza zbioru $\mathbb{N}$ są równe 0. Tak samo możemy argumentować w celu otrzymania wzorów w przypadku, gdy zmienne losowe przyjmują wartości w dowolnym przeliczalnym zbiorze kratowym $\{d \cdot i : i \in \mathbb{Z}\}$, gdzie $d > 0$ (zmienne losowe o rozkładach kratowych). Ustawiając dopuszczalne wartości zmiennych w ciąg, załóżmy, że X, Y przyjmują przeliczalną ilość wartości $y_1, y_2, \ldots$, ze zbioru $\{d \cdot i : i \in \mathbb{Z}\}$ z dodatnimi prawdopodobieństwami $p_X(y_i)$ i $p_Y(y_i)$, odpowiednio. Otrzymujemy z niezależności, dla $s \in \mathbb{R}$

$$F_S(s) = \sum_{i=1}^{\infty} F_X(s-y_i)p_Y(y_i) \tag{2.1.3}$$

oraz

$$p_S(s) = \sum_{i=1}^{\infty} p_X(s-y_i)p_Y(y_i). \tag{2.1.4}$$

Mówimy, że dystrybuanta F_S jest **splotem** F_X i F_Y i oznaczamy $F_S(s) = F_X * F_Y(s)$. Podobnie dla funkcji prawdopodobieństwa oznaczamy $p_S(s) = p_X * p_Y(s)$ jeśli zachodzi (2.1.4). Wygodnie jest wprowadzić oznaczenia na potęgi splotowe. $p_X^{*2} = p_X * p_X$ oraz $p_X^{*n} = p_X^{*(n-1)} * p_X$, dla $n \geqslant 1$. Dla $n = 0$, $p_X^{*0}(s) = \mathbb{I}_{\{0\}}(s)$, $F_X^{*0}(s) = \mathbb{I}_{[0,\infty)}(s)$.

Przypadek II. Rozkłady absolutnie ciągłe względem miary Lebesgua.

Dla zmiennych X, Y typu absolutnie ciągłego, czyli dla zmiennych o dystrybuantach postaci $F_X(s) = \int_{-\infty}^{s} f_X(x)dx$, $F_Y(s) = \int_{-\infty}^{s} f_Y(x)dx$, dla $s \in \mathbb{R}$ można zastosować analogiczne rozumowanie używając prawdopodobieństw warunkowych w celu otrzymania analogicznych wzorów . Można też zastosować inne metody.

Metoda I. *Przejście graniczne.* Niech dla $n \geqslant 1$, $Y^{(n)}$ będzie zmienną losową przyjmującą wartości w zbiorze $\{\frac{1}{2^n} \cdot i : i \in \mathbb{Z}\}$, zdefiniowaną przez $Y^{(n)} = \sum_{i\in\mathbb{Z}} \frac{i}{2^n} \mathbb{I}_{\{\frac{i}{2^n} \leqslant Y < \frac{i+1}{2^n}\}}$. Funkcja prawdopodobieństwa tej zmiennej jest określona przez $P(Y^{(n)} = \frac{i}{2^n}) = F_Y(\frac{i+1}{2^n}) - F_Y(\frac{i}{2^n})$. Podobnie jak w (2.1.3) otrzymujemy

$$\begin{aligned} F_{X+Y^{(n)}}(s) &= \sum_{i\in\mathbb{Z}}^{\infty} F_X(s - \frac{i}{2^n}) p_{Y^{(n)}}(\frac{i}{2^n}) \\ &= \sum_{i\in\mathbb{Z}}^{\infty} F_X(s - \frac{i}{2^n})(F_Y(\frac{i+1}{2^n}) - F_Y(\frac{i}{2^n})) \\ &= \sum_{i\in\mathbb{Z}}^{\infty} F_X(s - \frac{i}{2^n}) f_Y(\xi_{i,n}) \frac{1}{2^n}, \end{aligned}$$

dla pewnych $\xi_{i,n} \in [\frac{i}{2^n}, \frac{i+1}{2^n})$ wybranych na podstawie twierdzenia o wartości średniej. Przechodząc w ostatniej równości z $n \to \infty$, z lewej strony mamy $F_{X+Y^{(n)}}(s) \to_{n\to\infty} F_{X+Y}(s)$, bo zbieżność zmiennych losowych (prawie wszędzie) pociąga zbieżność dystrybuant (tutaj dystrybuanta graniczna jest ciągła). Z prawej strony ostatniej równości mamy sumę aproksymacyjną całki Riemanna, więc otrzymujemy w granicy

$$F_S(s) = \int_{-\infty}^{\infty} F_X(s-y) f_Y(y) dy = F_X * F_Y(s) \tag{2.1.5}$$

oraz różniczkując

$$f_S(s) = \int_{-\infty}^{\infty} f_X(s-y) f_Y(y) dy = f_X * f_Y(s). \tag{2.1.6}$$

Metoda II. *Wartość oczekiwana.* Dla pary zmiennych losowych X, Y o rozkładach absolutnie ciągłych możemy użyć następującego lematu (który natychmiast można uogólnić na większą liczbę zmiennych).

Lemat 2.1.1 *Niech X, Y będą zmiennymi losowymi o łącznej dystrybuancie*

$$F_{(X,Y)}(x,y) = P(X \leqslant x, Y \leqslant y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(x,y) dy dx,$$

wtedy

$$E(\psi(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x,y) f_{(X,Y)}(x,y) dy dx,$$

gdzie $\psi : \mathbb{R}^2 \to \mathbb{R}$ jest dowolną mierzalną funkcją.

Dowód. Dla $\psi(x,y) = \mathbb{I}_{(-\infty,x]\times(-\infty,y]}(x,y)$, teza wynika natychmiast z równości

$$E(\mathbb{I}_{(-\infty,x]\times(-\infty,y]}(X,Y)) = P(X \leqslant x, Y \leqslant y)$$

i z założenia lematu. Ponieważ dowolna funkcja ψ może być przybliżona kombinacjami liniowymi indykatorów takiej postaci, teza jest natychmiastowa.

Przyjmując teraz $\psi(x,y) = \mathbb{I}_{\{x+y\leqslant s\}}(x,y)$ otrzymujemy z powyższego lematu

$$\begin{aligned} P(X+Y \leqslant s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{I}_{\{x+y\leqslant s\}}(x,y) f_{(X,Y)}(x,y) dy dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{I}_{\{x+y\leqslant s\}}(x,y) f_X(x) f_Y(y) dx dy, \end{aligned}$$

gdzie ostatnia równość wynika z niezależności zmiennych X, Y. Ponieważ $\mathbb{I}_{\{x+y\leqslant s\}}(x,y) = \mathbb{I}_{\{x\leqslant s-y\}}(x)$ otrzymujemy (2.1.5).

Przykład 2.1.2 Niech X ma gęstość $f_X(x) = \frac{1}{2}I_{(0,2)}(x)$ oraz niezależnie, Y ma gęstość $f_Y(x) = \frac{1}{3}I_{(0,3)}(x)$. Wtedy ze wzoru (2.1.5)

$$F_S(s) = \begin{cases} 1 & \text{dla} \quad s \geqslant 5 \\ 1 - \frac{(5-s)^2}{12} & \text{dla} \quad 3 \leqslant s < 5 \\ \frac{s-1}{3} & \text{dla} \quad 2 \leqslant s < 3 \\ \frac{s^2}{12} & \text{dla} \quad 0 \leqslant s < 2 \\ 0 & \text{dla} \quad s < 0 \end{cases}.$$

Rzeczywiście, mamy dla $0 < x < 2$, $F_X(x) = \int_0^x f_X(u)du = 0.5\int_0^x du = 0.5x$, a stąd

$$F_X(x) = \begin{cases} 0 & \text{dla} \quad x \leqslant 0 \\ 0.5x & \text{dla} \quad 0 < x < 2 \\ 1 & \text{dla} \quad x \geqslant 2 \end{cases}.$$

Dla $0 \leqslant s < 2$ dostajemy

$$F_S(s) = \int_0^s \frac{1}{2}(s-y)\frac{1}{3}dy = \frac{1}{12}s^2.$$

Dla $2 \leqslant s < 3$ mamy: jeżeli $s - y > 2$ (czyli $0 < y < s - 2$), to $F_X(s - y) = 1$. Jeżeli $0 < s - y < 2$ (czyli $s - 2 < y < s$), to $F_X(s - y) = \frac{1}{2}(s - y)$, stąd

$$\begin{aligned} F_S(s) &= \int_0^{s-2} 1 \cdot \frac{1}{3}I_{(0,3)}(y)dy + \int_{s-2}^{s} \frac{1}{2}(s-y)\frac{1}{3}I_{(0,3)}(y)dy \\ &= \frac{1}{3}\int_0^{s-2} dy + \int_{s-2}^{s} \frac{1}{2}(s-y)\frac{1}{3}dy \\ &= \frac{s-1}{3}. \end{aligned}$$

Dla $3 \leqslant s < 5$, podobnie jak wyżej,

$$\begin{aligned} F_S(s) &= \int_0^{s-2} 1 \cdot \frac{1}{3}I_{(0,3)}(y)dy + \int I_{(s-2,s)}(y)\frac{1}{2}(s-y) \cdot \frac{1}{3}I_{(0,3)}(y)dy \\ &= \int_0^{s-2} \frac{1}{3}dy + \int \frac{1}{2}(s-y) \cdot \frac{1}{3}I_{(0,3)\cap(s-2,s)}(y)dy \\ &= \int_0^{s-2} \frac{1}{3}dy + \int_{s-2}^{3} \frac{1}{2}(s-y) \cdot \frac{1}{3}dy \\ &= 1 - \frac{(5-s)^2}{12}. \end{aligned}$$

Ten sam wynik otrzymamy licząc wielkości odpowiednich pól na rysunku przedstawiającym łączną gęstość (tak jak na wykładzie).

□

Niech X będzie zmienną o rozkładzie mieszanym, tzn. $F_X(s) = \alpha F_X^d(s) + (1-\alpha)F_X^c(s)$, dla pewnego $\alpha \in (0,1)$, gdzie F_X^d jest częścią dyskretną dystrybuanty F_X, a F_X^c jest częścią absolutnie ciagłą dystrybuanty F_X. Niech Y będzie zmienną o rozkładzie mieszanym, tzn. $F_Y(s) = \beta F_Y^d(s) + (1-\beta)F_Y^c(s)$, dla pewnego $\beta \in (0,1)$, gdzie $F_Y^d(s) = \sum_i P(Y = y_i) I_{[y_i,\infty)}(s)$ jest częścią dyskretną dystrybuanty F_Y, a $F_Y^c(s) = \int_{-\infty}^s f_Y^c(y)dy$ jest częścią absolutnie ciagłą dystrybuanty F_Y. Wygodnie jest wprowadzić ogólne oznaczenie na splot dystrybuant następująco,

$$F_X * F_Y(s) = \int_{-\infty}^{\infty} F_X(s-y)dF_Y(y),$$

gdzie $\int_{-\infty}^{\infty} h(y)dF_Y(y) = \beta \sum_i h(y_i)P(Y = y_i)/\beta + (1-\beta)\int_{-\infty}^{\infty} h(y)f_Y^c(y)dy$, dla dowolnej funkcji całkowalnej $h(s)$.

Możemy więc bezpośrednio określić

$$F_S(s) = F_X * F_Y(s) = \sum_i F_X(s-y_i)P(Y = y_i) + (1-\beta)\int_{-\infty}^{\infty} F_X(s-y)f_Y^c(y)dy,$$

gdzie $\beta = \sum_i P(Y = y_i)$.

Przykład 2.1.3 Niech X ma rozkład z atomami $P(X = 0) = 0.2$, $P(X = 1) = 0.7$ i gęstością $f_X(x) = 0.1$ dla $x \in (0,1)$. Zmienna losowa Y ma rozkład z atomami $P(Y = 0) = 0.3$, $P(Y = 1) = 0.2$ i gęstością $f_Y(x) = 0.5$, $x \in (0,1)$. Zakładając, że X i Y są niezależne obliczymy $P(X + Y \in [1, 1.5))$.

Metoda I (wyliczenie bezpośrednie poprzez analizę zdarzeń sprzyjających). Mamy

$$\begin{aligned}
&\{X + Y \in [1, 1.5)\} = \\
&= \{X = 1, Y = 0\} \cup \{X = 0, Y = 1\} \cup \{X \in (0,1), Y \in (0,1), X + Y \in [1, 1.5)\} \\
&\cup \{X = 1, Y \in (0,1), X + Y \in [1, 1.5)\} \cup \{Y = 1, X \in (0,1), X + Y \in [1, 1.5)\}
\end{aligned}$$

Stąd

$$\begin{aligned}
P(X+Y\in[1,1.5)) &= P(X=1,Y=0)+P(X=0,Y=1)+\\
&\quad +P(X\in(0,1),Y\in(0,1),X+Y\in[1,1.5))\\
&\quad +P(X=1,Y\in(0,1),X+Y\in[1,1.5))\\
&\quad +P(Y=1,X\in(0,1),X+Y\in[1,1.5))\\
&= 0.7\cdot 0.3+0.2\cdot 0.2+\int_0^1 P(Y\in(0,1),X+Y\in(1,1.5)|X=x)f_X(x)dx+\\
&\quad +0.7\int_0^{0.5} f_Y(x)dx+0.2\int_0^{0.5} f_X(x)dx\\
&= 0.7\cdot 0.3+0.2\cdot 0.2+0.1\int_0^1 P(Y\in(1-x,1.5-x)\cap(0,1))dx\\
&\quad +0.7\cdot 0.25+0.2\cdot 0.05\\
&= 0.7\cdot 0.3+0.2\cdot 0.2+0.7\cdot 0.25+0.2\cdot 0.05\\
&\quad +0.1\left(\int_0^{0.5} P(Y\in(1-x,1))dx+\int_{0.5}^1 P(Y\in(1-x,1.5-x))dx\right)\\
&= 0.7\cdot 0.3+0.2\cdot 0.2+0.7\cdot 0.25+0.2\cdot 0.05\\
&\quad +0.1*0.5\left(\int_0^{0.5} xdx+\int_{0.5}^1 0.5dx\right)=0.45375.
\end{aligned}$$

Metoda II (graficzna)

Podobnie jak w poprzednim przykładzie, alternatywną metodą rozwiązania tego zagadnienia jest geometryczne przedstawienie masy łącznego rozkładu (ćwiczenia).

Metoda III. (Sploty dla rozkładów mieszanych).

Dla zmiennej Y mamy, $F_Y^d(s)=(0.3I_{[0,\infty)}(s)+0.2I_{[1,\infty)}(s))/0.5$, $F_Y^c(s)=\int_{-\infty}^s I_{(0,1)}(y)dy$, $\beta=0.5$ oraz

$P(X+Y\in[1,1.5))=P(X+Y<1.5)-P(X+Y<1)$. $P(X+Y<1.5)=P(X+Y\leqslant 1.5)=F_S(1.5)$, bo $P(X+Y=1.5)=0$. $P(X+Y<1)=P(X+Y\leqslant 1)-P(X+Y=1)=F_S(1)-P(X+Y=1)$. Czyli $P(S\in[1,1.5))=F_S(1.5)-F_S(1)+P(X+Y=1)$. Używając definicji splotu

$$F_S(s)=F_X(s-0)P(Y=0)+F_X(s-1)P(Y=1)+0.5\int_{-\infty}^{\infty} F_X(s-y)I_{(0,1)}(y)dy.$$

Wstawiając znane wartości, otrzymujemy

$$F_S(s)=F_X(s)0.3+F_X(s-1)0.2+0.5\int_0^1 F_X(s-y)dy,$$

oraz

$$\begin{aligned}
F_S(1.5) &= 0.3+0.25\cdot 0.2+0.5\int_0^1 F_X(1.5-y)dy\\
&= 0.35+0.5\int_{0.5}^{1.5} F_X(y)dy\\
&= 0.35+0.5\cdot 0.6375=0.66875
\end{aligned}$$

Podobnie otrzymujemy

$$\begin{aligned} F_S(1) &= 0.3 + 0.2 \cdot 0.2 + 0.5 \int_0^1 F_X(1-y)dy \\ &= 0.34 + 0.5 \int_0^1 F_X(y)dy \\ &= 0.34 + 0.5 \cdot 0.25 = 0.465. \end{aligned}$$

Ostatecznie $P(X+Y \in [1, 1.5)) = 0.66875 - 0.465 + 0.21 + 0.04 = 0.45375$.

□

Niech teraz $S = S_n = X_1 + \cdots + X_n$, gdzie $(X_i)_{i \geqslant 1}$ są niezależnymi zmiennymi losowymi. Policzenie rozkładu sumy S bezpośrednio ze wzorów (2.1.3)-(2.1.4) jest zazwyczaj pracochłonnym zadaniem. Rzeczywiście, w przypadku, gdy wartości X_i są naturalne, mając $P(S_1 = k) = P(X_1 = k)$, liczymy $P(S_n = k)$ w sposób rekurencyjny:

$$P(S_n = k) = \sum_{m=0}^{k} P(S_{n-1} = k - m)P(S_1 = m). \tag{2.1.7}$$

W przypadku rozkładów dyskretnych o skończonym nośniku wzory te moga być użyte, jest to jednak (oprócz przypadku, gdy n jest bardzo małe, np. $n = 1, 2, 3$) bardzo nieefektywne. Aby obliczyć rozkład np. sumy trzech zmiennych losowych niezależnych $X_1 + X_2 + X_3$ najpierw ze wzoru (2.1.4) obliczamy rozkład f_{S_2} sumy $S_2 = X_1 + X_2$, a następnie zastosujemy powyższy wzór do obliczenia rozkładu $S_3 = S_2 + X_3$. Widać, że w przypadku dowolnego n w celu obliczenia rozkładu S_n będziemy musieli zastosować takie postępowanie rekurencyjne $n - 1$ razy.

Przykład 2.1.4 Trzy niezależne ryzyka mają rozmiary szkód jak w tabeli:

i	0	1	2	3
$P(X_1 = i)$	0.3	0.2	0.4	0.1
$P(X_2 = i)$	0.6	0.1	0.3	0
$P(X_3 = i)$	0.4	0.2	0	0.4

Policz rozkład zmiennej $S = S_3 = X_1 + X_2 + X_3$.

Najpierw obliczymy rozkład p_{S_2} dla $S_2 = X_1 + X_2$. Ze wzoru (2.1.4) otrzymujemy

$$\begin{aligned} p_{S_2}(0) &= p_{X_1}(0)p_{X_2}(0) = 0.18, \\ p_{S_2}(1) &= p_{X_1}(0)p_{X_2}(1) + p_{X_1}(1)p_{X_2}(0) = 0.15, \\ p_{S_2}(2) &= p_{X_1}(0)p_{X_2}(2) + p_{X_1}(1)p_{X_2}(1) + p_{X_1}(2)p_{X_2}(0) = 0.35, \\ &\cdots \\ p_{S_2}(5) &= p_{X_1}(3)p_{X_2}(2) = 0.03. \end{aligned}$$

Następnie w ten sam sposób obliczymy rozkład S_3

$$p_{S_3}(0) = p_{S_2}(0)p_{X_3}(0) = 0.072,$$
$$p_{S_3}(1) = p_{S_2}(0)p_{X_3}(1) + p_{S_2}(1)p_{X_3}(0) = 0.096,$$
$$\ldots.$$

Wyniki te przedstawimy w tabeli.

x	$p_{X_1}(x)$	$p_{X_2}(x)$	$p_{X_3}(x)$	$p_{S_2}(x)$	$p_S(x)$
0	0.3	0.6	0.4	0..18	0.072
1	0.2	0.1	0.2	0.15	0.096
2	0.6	0.3	0	0.35	0.170
3	0.4	0	0.4	0.16	0.206
4	-	-	-	0.13	0.144
5	-	-	-	0.03	0.178
6	-	-	-	-	0.070
7	-	-	-	-	0.052
8	-	-	-	-	0.012

□

Dla zmiennych całkowitoliczbowych (oraz dla zmiennych o wartościach w zbiorze wielokrotności $\{dn : n \in N\}$ dla $d > 0$) o wiele efektywniejsze jest użycie funkcji tworzących, a dla dowolnych zmiennych, funkcji tworzących momenty.

Definicja 2.1.5 *Dla dowolnego ciągu liczb rzeczywistych* $(a_n)_{n\geqslant 0}$ *funkcję*

$$A(t) = \sum_{n=0}^{\infty} a_n t^n,$$

nazywamy ***funkcją tworzącą*** *tego ciągu.*

Jeśli $(a_n)_{n\geqslant 0}$ jest ograniczony, to funkcja tworząca przyjmuje wartości skończone dla $\mid t \mid < 1$.

Dla zmiennej losowej X przyjmującej wartości ze zbioru liczb naturalnych definiujemy funkcję

$$P_X(t) = \mathrm{E}\left[t^X\right],$$

jest to **funkcja tworząca prawdopodobieństwa**.

Jeśli oznaczymy przez $p_n := P(X = n)$, to funkcja tworząca ciągu $(p_n)_{n\geqslant 0}$ równa się $P_X(\cdot)$. Zauważmy, że $P_X(t)$ przyjmuje wartości skończone przynajmniej dla $\mid t \mid \leqslant 1$.

Funkcję ogona rozkładu określamy przez $q_n := p_{n+1} + p_{n+2} + \cdots$. Funkcja tworząca ciągu $(q_n)_{n\geqslant 0}$, $Q_X(t) = \sum_{n=0}^{\infty} q_n t^n$ jest skończona przynajmniej dla $\mid t \mid < 1$. Łatwo zauważyć, że

$$Q_X(t) = \frac{1 - P_X(t)}{1 - t}.$$

Ponadto

$$P'_X(t) = \sum_{k=1}^{\infty} k p_k t^{k-1},$$

i funkcja ta jest skończona przynajmniej dla $\mid t \mid < 1$. Ponadto, jeśli wartość oczekiwana zmiennej X jest skończona, to

$$EX = P'_X(1).$$

Zauważmy, że funkcję Q_X możemy zapisać jako iloraz różnicowy $Q_X(t) = \frac{P_X(1) - P_X(1-h)}{h}$, dla $h := 1 - t$. Gdy $t \to 1$, to $h \to 0$ i mamy $Q_X(1) = P'_X(1) = EX$, co daje

$$EX = q_0 + q_1 + q_2 + \cdots .$$

Podobnie możemy otrzymać

$$P''_X(1) = 2Q'_X(1) = E(X(X-1)).$$

Dla wariancji zmiennej X zachodzi więc równość

$$VarX = P''_X(1) + P'_X(1) - (P'_X(1))^2.$$

Podobne rozumowania możemy powtórzyć dla wyższych momentów zmiennej losowej X.

Funkcje tworzące prawdopodobieństwa, oprócz przydatnosci do liczenia momentów, przydają się do liczenia rozkładów sum zmiennych losowych. Funkcja tworząca sumy niezależnych zmiennych losowych jest równa iloczynowi funkcji tworzących składników.

Lemat 2.1.6 *Niech X, Y będą niezależnymi zmiennymi losowymi o wartościach w zbiorze liczb naturalnych, o funkcjach tworzących prawdopodobieństwa, odpowiednio P_X, P_Y. Wtedy zmienna losowa $X + Y$ ma funkcję prawdopodobieństwa daną splotem (2.1.4) oraz funkcję tworzącą prawdopodobieństwa P_{X+Y} równą iloczynowi $P_X P_Y$.*

Wynika to bezpośrednio z porównania współczynników w szeregach potęgowych - po wymnożeniu i z równości (2.1.4).

Funkcje tworzące prawdopodobieństwa są przydatne również do badania zbieżności ciągu rozkładów (zob. Feller (1981), rozdz.).

Lemat 2.1.7 *Niech $(X_n)_{n\geqslant 1}$ będzie ciagiem zmiennych losowych przyjmujących wartości naturalne, o funkcjach prawdopodobieństwa p_{X_n} i funkcjach tworzących prawdopodobieństwa P_{X_n}. Wtedy następujące warunki są równoważne*

1. *$p_{X_n}(k) \to_{n\to\infty} p_Y(k)$, dla każdego naturalnego k i dla pewnej zmiennej losowej Y,*
2. *$P_{X_n}(t) \to_{n\to\infty} P_Y(t)$, dla każdego $t \in [0,1)$ i dla pewnej zmiennej losowej Y.*

Przykład 2.1.8 Funkcja tworząca prawdopodobieństwa rozkładu dwumianowego (binomialnego). Rozkład dwumianowy jest rozkładem liczby sukcesów w próbach Bernoulliego, tzn. rozkładem sumy $S_n = X_1 + \cdots + X_n$, dla n prób, gdzie $\{X_i, i = 1, \ldots, n\}$ są niezależne o rozkładzie (funkcji prawdopodobieństwa) $p_{X_i}(0) = 1 - p_{X_i}(1) = 1 - p = q$, gdzie $p \in (0,1)$ jest prawdopodobieństwem sukcesu. Ponieważ $P_{X_i}(t) = q + pt$, z definicji, więc

$$P_{S_n}(t) = (q + pt)^n.$$

□

Przykład 2.1.9 Zbieżność ciągu rozkładów dwumianowych do rozkładu Poissona. Rozkład Poissona jest dany przez $p_Y(k) = e^{-\lambda}\frac{\lambda^k}{k!}$, dla $k = 0, 1, 2, \ldots$. Z definicji liczymy natychmiast

$$P_Y(t) = e^{-\lambda(1-t)}.$$

Rozważmy ciąg zmiennych losowych o rozkładach dwumianowych $p_{S_n}(k) = P(S_n = k) = \binom{n}{k} p_n^k q_n^{n-k}$, gdzie prawdopodobieństwo sukcesu jest zależne od n, w taki sposób, że $np_n \to_{n\to\infty} \lambda > 0$. Dla funkcji tworzących prawdopodobieństwa mamy

$$P_{S_n}(t) = (q_n + p_n t)^n = (1 - \frac{np_n(1-t)}{n})^n,$$

stąd $P_{S_n}(t) \to_{n\to\infty} e^{-\lambda(1-t)} = P_Y(t)$. Z lematu 2.1.7 otrzymujemy $p_{S_n}(k) \to p_Y(k)$, tzn. dla dużych n prawdopodobieństwa dwumianowe możemy przybliżać rozkładem Poissona, o ile zachodzi $np \approx \lambda$.

□

Inne transformacje zmiennej losowej X zdefiniowane są jako:

$$M_X(t) = \mathrm{E}\left[e^{tX}\right],$$

funkcja tworząca momenty,

$$C_X(t) = \log \mathrm{E}\left[e^{tX}\right],$$

funkcja tworząca kumulanty.

Mamy wtedy ogólnie dla niezależnych zmiennych $(X_i)_{i\geqslant 1}$

$$\begin{aligned} M_{S_n}(t) &= \prod_{i=1}^{n} M_{X_i}(t), \\ P_{S_n}(t) &= \prod_{i=1}^{n} P_{X_i}(t), \\ C_{S_n}(t) &= \sum_{i=1}^{n} C_{X_i}(t). \end{aligned} \tag{2.1.8}$$

Rzeczywiście, z niezależności

$$\begin{aligned} M_{S_n}(t) &= \mathrm{E}\left[e^{tS_n}\right] = \mathrm{E}\left[e^{t(X_1+\ldots+X_n)}\right] \\ &= \mathrm{E}\left[e^{tX_1}\right]\cdots\mathrm{E}\left[e^{tX_n}\right] = M_{X_1}(t)\cdots M_{X_n}(t) \end{aligned}$$

i analogicznie dla pozostałych funkcji.

Przykład 2.1.10 Liczba porażek przed uzyskaniem pierwszego sukcesu w kolejnych próbach Bernoulliego jest zmienną losową o rozkładzie geometrycznym $P(X = k) = p_X(k) = q^k p$, $k = 0, 1, 2, \ldots$. Z definicji (szereg geometryczny) otrzymujemy

$$P_X(t) = \frac{p}{1 - qt}.$$

Liczba porażek przy oczekiwaniu na r-ty sukces jest więc sumą r niezależnych zmiennych losowych o rozkładach geometrycznych $S_r = X_1 + \cdots + X_r$. Zmienna ta ma funkcję tworzącą prawdopodobieństwa

$$P_{S_r}(t) = (\frac{p}{1 - qt})^r.$$

Rozkład ten nazywamy rozkładem Pascala (szczególny przypadek ujemnego rozkładu dwumianowego).

□

Twierdzenie 2.1.11 *Załóżmy, że $X_1, \ldots, X_n$ są niezależne. Wtedy*

1. *Jeżeli $X_i \sim Poi(\lambda_i)$ to $S_n \sim Poi(\lambda)$, gdzie $\lambda = \sum_{i=1}^n \lambda_i$.*
2. *Jeżeli $X_i \sim Bin^-(r_i, q)$ to $S_n \sim Bin^-(r, q)$,gdzie $r = \sum_{i=1}^n r_i$.*
3. *Jeżeli $X_i \sim Bin(m_i, p)$ to $S_n \sim Bin(m, p)$,gdzie $m = \sum_{i=1}^n m_i$*

4. *Jeżeli* $X_i \sim \Gamma(\alpha_i, \beta)$ *to* $S_n \sim \Gamma(\alpha, \beta)$, *gdzie* $\alpha = \sum_{i=1}^n \alpha_i$.

5. *Jeżeli* $X_i \sim N(\mu_i, \sigma_i^2)$ *to* $S_n \sim N(\mu, \sigma^2)$, *gdzie* $\mu = \sum_{i=1}^n \mu_i$, $\sigma^2 = \sum_{i=1}^n \sigma_i^2$.

Korzystając z transformat możemy wyliczyć momenty zmiennych losowych.

Oznaczmy

$$\mu_k(X) = \mathrm{E}\left[X^k\right],$$

$$m_k(X) = \mathrm{E}\left[(X - \mathrm{E}[X])^k\right], \ k > 0.$$

W przypadku, gdy wiadomo o jaką zmienną losową chodzi piszemy m_k i μ_k. Parametr μ_k nazywany jest k-tym **momentem zwykłym**, m_k - k-tym **momentem centralnym**. W szczególności $\mu_1(X) =: \mu_X$ jest średnią, $m_2(X) =: \sigma_X^2$ jest wariancją, a σ_X jest odchyleniem standardowym. Pomijamy indeks X w powyższych oznaczeniach, jeśli z kontekstu jasno wynika jakich zmiennych losowych dotyczą rozważania.

Parametr

$$\gamma_3 := \frac{m_3}{\sigma^3}$$

jest nazywamy **skośnością**, a

$$\gamma_4 := \frac{m_4}{\sigma^4} - 3$$

nazywamy **kurtozą**. Iloraz

$$\gamma_1 := \frac{\sigma^2}{\mu}$$

nazywamy **indeksem dyspersji**, a

$$\gamma_2 = \frac{\sigma}{\mu}$$

współczynnikiem zmienności.

Przy założeniu niezależności zmiennych $M_{S_n}^{(1)}(0) = C_{S_n}^{(1)}(0) = \mathrm{E}[S_n]$, $C_{S_n}^{(2)}(0) = \mathrm{Var}[S_n]$, $C_{S_n}^{(3)}(0) = m_3(S_n)$, a stąd na przykład

$$\mu_1(S_n) = \sum_{i=1}^n \mu_1(X_i),$$
$$\mu_2(S_n) = \sum_{i=1}^n \mu_2(X_i).$$

Dla niezależnych zmiennych losowych o tym samym rozkładzie dostajemy

$$\mu_1(S_n) = n\mu_1(X),$$
$$\mu_2(S_n) = n\mu_2(X).$$

Rozwijając w szereg Taylora funkcję C_X otrzymujemy

$$C_X(t) = \sum_{n=1}^{\infty} k_n(X) t^n/n!,$$

gdzie współczynniki $k_n(X) = C_X^{(n)}(0)$ nazywamy kumulantami zmiennej losowej X. Łatwo sprawdzamy, że $k_1(X) = \mu_X$, $k_2(X) = \sigma_X^2$, $k_3(X) = \gamma_3\sigma_X^3 = m_3(X)$, $k_4(X) = \gamma_4(X)\sigma_X^4 = m_4(X) - 3\sigma_X^4$.

Zachodzą następujące własności ogólne $k_n(X+c) = k_n(X)$, dla $n \geqslant 2$, , $k_n(cX) = c^n k_n(X)$, $c \in \mathbb{R}$. Dla niezależnych zmiennych losowych X, Y, $k_n(X+Y) = k_n(X) + k_n(Y)$.

Jako ilustrację metody liczenia rozkładu przy użyciu funkcji tworzących przedstawimy jeszcze raz wyliczenia z przykładu 2.1.4.

Przykład 2.1.12 Funkcje tworzące prawdopodobieństwa zmiennych X_1, X_2, X_3 mają postać

$$P_{X_1}(t) = 0.3 + 0.2t + 0.4t^2 + 0.1t^3,$$
$$P_{X_2}(t) = 0.6 + 0.1t + 0.3t^2,$$
$$P_{X_3}(t) = 0.4 + 0.2t + 0.4t^3,$$

i po wymnożeniu otrzymujemy funkcję tworzącą rozkładu sumy

$$P_S(t) = 0.072 + 0.096t + 0.170t^2 + 0.206t^3 + \\ + 0.144t^4 + 0.178t^5 + 0.070t^6 + 0.052t^7 + 0.012t^8,$$

a stąd odczytując współczynniki przy t^k, $k = 0, 1, \ldots, 8$, odczytujemy rozkład:

i	0	1	2	3	4	5	6	7	8
$P(S=i)$	0.072	0.096	0.170	0.20	0.144	0.178	0.070	0.052	0.012

.

Metoda ta jest bardzo efektywna przy użyciu komputera, bo łatwo można funkcje tworzące rozwinąć w szereg potęgowy Taylora, automatycznie otrzymując rozkład prawdopodobieństwa z funkcji tworzącej prawdopodobieństwa danej w postaci szeregu (por. zadania na ćwiczeniach).

□

2.2 Rozkłady w modelu złożonym

2.2.1 Zmienne losowe liczące ilość szkód

Sposób wyboru zmiennej liczącej w modelowaniu portfela złożonego zależy od modelowanego portfela. Pewne podstawowe cechy dobieranych rozkładów można rozpoznać z próbki

używając (próbkowej) średniej i wariancji. Ponieważ dla zmiennej losowej N o rozkładzie dwumianowym $Bin(n,p)$ mamy $\mathrm{E}[N] = np > \mathrm{Var}[N] = np(1-p)$, więc rozkłady dwumianowe można stosować wtedy, gdy średnia próbkowa jest dużo większa niż wariancja próbkowa.

Ponieważ dla zmiennej losowej N o rozkładzie Poissona $Poi(\lambda)$, mamy $\mathrm{E}[N] = \lambda = \mathrm{Var}[N]$, więc rozkład ten jest odpowiedni, gdy średnia próbkowa ilości szkód jest w przybliżeniu równa wariancji próbkowej. Założenie Poissonowskości ilości szkód jest zazwyczaj bardziej realistyczne niż założenie o dwumianowości rozkładu z innych względów, lecz sytuacja równości średniej i wariancji występuje dość rzadko.

Ilość szkód modeluje się często mieszanymi rozkładami Poissona.

Rozważmy portfel ubezpieczeń składający się z polis dla których liczba roszczeń jest zmienną losową N o rozkładzie Poissona z parametrem Θ. Jeżeli przyjmiemy, że Θ jest zmienną losową, to rozkład zmiennej N ma parametr, który też jest zmienną losową Θ przyjmującą wartości dodatnie i posiadającą dystrubuantę U. Taka modyfikacja prowadzi do tzw. **mieszanego rozkładu Poissona,** dla którego

$$P(N=n) = \int_0^\infty P(N=n|\Theta=\theta)dF_\Theta(\theta) = \int_0^\infty \frac{e^{-\theta}\theta^n}{n!}dF_\Theta(\theta).$$

Mieszany rozkład Poissona będziemy oznaczać przez $MPoi(\Theta)$.

Uwaga 2.2.1 W obliczeniach wygodnie jest używać zmiennych losowych $E[X \mid Y]$ oraz $Var[X \mid Y]$ (warunkową wartość oczekiwaną i warunkową wariancję), które są zdefiniowane przy użyciu warunkowego rozkładu. Niech

$$F_{X|Y=y}(t) := P(X \leqslant t|Y=y)$$

oznacza warunkową dystrybuantę X względem $Y=y$ oraz

$$E[X|Y=y] := \int_{-\infty}^{\infty} x dF_{X|Y=y}(x)$$
$$Var[X|Y=y] := \int_{-\infty}^{\infty} x^2 dF_{X|Y=y}(x) - (E[X|Y=y])^2$$

oznaczają warunkową wartość oczekiwaną i warunkową wariancję (które są funkcjami zmiennej y). Definiujemy zmienne losowe

$$E[X \mid Y] := \varphi(Y),$$
$$Var[X \mid Y] := \psi(Y),$$

dla rzeczywistych funkcji φ, ψ takich, że dla prawie każdego (wzgl. rozkładu zmiennej Y) y

$$E[X \mid Y=y] = \varphi(y),$$
$$Var[X \mid Y=y] = \psi(y).$$

Lemat 2.2.2 *Dla dowolnych zmiennych losowych X, Y zachodzi następujący związek*

$$E[X] = E[E[X \mid Y]]. \tag{2.2.1}$$

Podamy uzasadnienie powyższego wzoru dla zmiennej losowej o rozkładzie dyskretnym (w przypadku zmiennej losowej ciągłej dowód przebiega analogicznie, tylko sumy należy zamienić na całki, alternatywnie, dowolny rozkład możemy przybliżyć rozkładami dyskretnymi monotonicznie)

$$\begin{aligned} E[E[X \mid Y]] &= \sum_k E[X \mid Y = y_k] \Pr(Y = y_k) \\ &= \sum_k \sum_i x_i \Pr(X = x_i \mid Y = y_k) \Pr(Y = y_k) \\ &= \sum_k \sum_i x_i \frac{\Pr(X = x_i, Y = y_k)}{\Pr(Y = y_k)} \Pr(Y = y_k) \\ &= \sum_i x_i \sum_k \Pr(X = x_i, Y = y_k) = \sum_i x_i \Pr(X = x_i) = E[X] \end{aligned}$$

Bezpośrednio z definicji mamy

$$Var[X|Y] = E[X^2|Y] - (E[X|Y])^2.$$

Nie należy mylić zmiennych $Var[X|Y]$ i $E[(X - E[X])^2|Y]$.

Lemat 2.2.3 *Dla dowolnych zmiennych X, Y, dla których odpowiednie momenty istnieją, mamy*

$$Var[X] = E[Var[X|Y]] + Var[E[X|Y]] \tag{2.2.2}$$

Rzeczywiście,

$$\begin{aligned} Var[X] = E[X^2] - (E[X])^2 = E[E[X^2|Y]] - (E[E[X|Y]])^2 = \\ = E[E[X^2|Y] - (E[X|Y])^2] + E[(E[X|Y])^2] - (E[E[X|Y]])^2 = \\ = E[Var[X|Y]] + Var[E[X|Y]]. \end{aligned}$$

Korzystając z (2.2.1),

$$P_N(t) = \mathrm{E}\left[\mathrm{E}\left[t^N \mid \Theta\right]\right] = \mathrm{E}\left[e^{\Theta(t-1)}\right] = M_\Theta(t-1).$$

Ponadto

$$C_N(t) = \log M_N(t) = \log P_N(e^t) = \log M_\Theta(e^t - 1)$$

oraz

$$\begin{aligned} \mathrm{E}[N] &= \mathrm{E}[\Theta] \\ \mathrm{Var}[N] &= \mathrm{E}[\Theta] + \mathrm{Var}[\Theta] = \mathrm{E}[N] + \mathrm{Var}[\Theta], \\ \mathrm{E}\left[(N - \mathrm{E}[N])^3\right] &= \mathrm{E}\left[(\Theta - \mathrm{E}[\Theta])^3\right] + 3\mathrm{Var}[\Theta] + \mathrm{E}[\Theta]. \end{aligned}$$

Z powyższych wzorów wynika, że model taki będziemy stosowali wtedy, gdy dla próbki danych średnia próbkowa ilości szkód jest mniejsza niż wariancja próbkowa.

Przykład 2.2.4 Załóżmy, że zmienna losowa N ma mieszany rozkład Poissona, a Θ ma rozkład $\Gamma(\alpha, \beta)$. Ponieważ funkcja tworząca momenty dla rozkładu $\Gamma(\alpha, \beta)$ dana jest wzorem

$$M_\Theta(t) = \left(\frac{\beta}{\beta - t}\right)^\alpha \quad \text{dla } t < \beta,$$

więc podstawiając

$$r = \alpha, \qquad p = \frac{\beta}{\beta + 1}, \qquad q = 1 - p,$$

dostajemy

$$\begin{aligned} M_N(t) &= M_\Theta(e^t - 1) = \left(\frac{\beta}{\beta - (e^t - 1)}\right)^\alpha = \left(\frac{\frac{\beta}{\beta+1}}{1 - \left(1 - \frac{\beta}{1+\beta}\right)e^t}\right)^\alpha \\ &= \left(\frac{p}{1 - qe^t}\right)^r. \end{aligned}$$

Jest to funkcja tworząca rozkładu **ujemnego dwumianowego** $Bin^-(r, p)$. Mamy więc $MPoi(Gamma(\alpha, \beta)) = Bin^-(\alpha, \frac{\beta}{\beta+1})$. Funkcja prawdopodobieństwa tego rozkładu zadana jest wzorem

$$P(N = n) = \binom{r + n - 1}{n} p^r q^n, n \in \mathbb{N} \tag{2.2.3}$$

Jeżeli $r = 1$, to otrzymujemy **rozkład geometryczny**, $N \sim Geo(p)$, co oznacza, że randomizacja rozkładem wykładniczym parametru wartości średniej w rozkładzie Poissona daje w rezultacie rozkład geometryczny, $MPoi(Exp(\beta)) = Geo(\frac{\beta}{\beta+1})$.

□

Przekształcając gęstość (2.2.3) możemy ją zapisać w postaci

$$P(N=n)=\begin{cases}(-1)^n\binom{-r}{n}p^r q^n & \text{dla} \quad n=0,1,2,\ldots,\\ 0 & \text{dla} \quad \text{pozostałych wartości } n,\end{cases} \tag{2.2.4}$$

gdzie

$$\binom{-r}{n}=\frac{(-r)(-r-1)...(-r-n+1)}{n!}.$$

Dla tego rozkładu

$$E[N]=\frac{rq}{p}, \tag{2.2.5}$$

$$Var[N]=\frac{rq}{p^2}. \tag{2.2.6}$$

$$\tag{2.2.7}$$

Oznaczmy skrótowo funkcję prawdopodobieństwa zmiennej N przez $p_k=f_N(k)=P(N=k), k\in\mathbb{N}$. Załóżmy, że

$$p_k=\left(a+\frac{b}{k}\right)p_{k-1},\ k\geqslant 1, \tag{2.2.8}$$

dla pewnego doboru parametrów a i b. Zapisując to inaczej dostajemy

$$k\frac{p_k}{p_{k-1}}=ka+b=:l(k). \tag{2.2.9}$$

W szczególności, dla rozkładu dwumianowego $Bin(n,p)$

$$a:=\frac{-p}{1-p},\ b:=\frac{p(n+1)}{1-p},$$

Poissona $Poi(\lambda)$,

$$a:=0,\ b:=\lambda$$

oraz ujemnie dwumianowego $Bin^-(r,p)$,

$$a:=q,\ b:=(r-1)q.$$

Jeżeli więc dla próbki $N_1,\ldots,N_n$, z rozkładu zmiennej N zdefiniujemy licznik $n_k:=\#\{i:N_i=k\}$, to wykres funkcji $\hat{l}:k\to k\frac{n_k}{n_{k-1}}$ powinien być w przybliżeniu liniowy

na podstawie (2.2.9). Punkt przecięcia linii $\hat{l}(k)$ z osią OY jest przybliżeniem parametru b, natomiast z osią OX, ilorazu $\frac{-b}{a}$. Metodę tą nazwiemy metodą Panjera. Jeśli wykres nie jest w przybliżeniu liniowy, to rozkład nie należy do klasy rozkładów spełniających rekurencję (2.2.8).

Okazuje się, że tylko te trzy wymienione rozkłady spełniają tę rekurencję.

Twierdzenie 2.2.5 *Przypuśćmy, że rozkład $(p_k)_{k\geqslant 0}$ spełnia rekurencję*

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1},\ k \geqslant 1.$$

Wtedy $(p_k)_{k\geqslant 0}$ jest rozkładem Poissona, dwumianowym lub ujemnym dwumianowym.

Dowód:

Gdy $a = 0$ wtedy, aby rekurencja miała sens, przyjmujemy $b > 0$. Z zależności rekurencyjnej mamy

$$p_k = p_0 \frac{b^k}{k!},$$

co natychmiast implikuje, że (p_k) jest rozkładem Poissona z parametrem $\lambda = b$.

Załóżmy, że $a \neq 0$. Zauważmy, że

$$p_k = \frac{a^k}{k!}(\Delta + k - 1)(\Delta + k - 2)\cdots(\Delta + 1)\Delta p_0, \qquad k \in \mathbb{N},$$

gdzie $\Delta = (1 + \frac{b}{a})$.
Sumując obie strony względem k mamy

$$\begin{aligned} 1 &= p_0 \sum_{k=0}^{\infty} (\Delta + k - 1)(\Delta + k - 2)\ldots\Delta \frac{a^k}{k!} \\ &= p_0 \sum_{k=0}^{\infty} \binom{-\Delta}{k} (-a)^k = p_0 (1 - a)^{-\Delta}. \end{aligned}$$

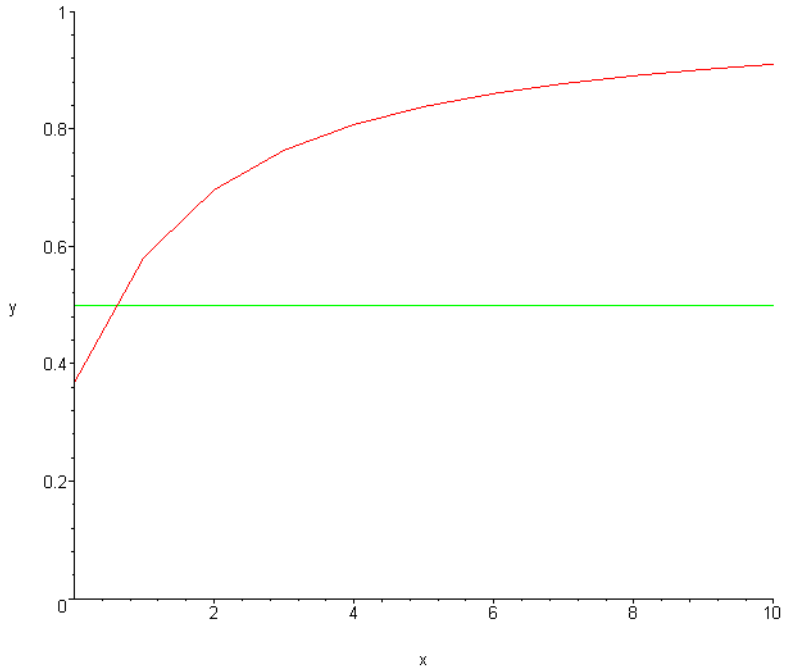
a stąd $p_0 = (1 - a)^{\Delta}$ oraz

$$p_k = \binom{-\Delta}{k} (-a)^k (1 - a)^{\Delta} = \binom{\Delta + k - 1}{k} a^k (1 - a)^{\Delta}, \qquad k \in \mathbb{N}. \tag{2.2.10}$$

Zauważmy, że dla $0 < a < 1$ prawa strona wyrażenia (2.2.10) jest dodatnia dla każdego $k \in \mathbb{N}$, gdy $\Delta > 0$ (tzn. $b > -a$). W tym przypadku (p_k) ma rozkład ujemny dwumianowy z parametrem $p = 1 - a$ oraz $r = \Delta$.

Ponieważ $a > 1$ wykluczamy, rozpatrujemy w końcu $a < 0$. W tym przypadku liczby p_k są nieujemne, gdy $-\Delta \in \mathbb{N}$. W tym przypadku otrzymujemy rozkład dwumianowy. □

Inna metoda graficzna będzie oparta na **funkcji hazardowej** zdefiniowanej dla n należących do nośnika rozkładu zmiennej N przyjmującej wartości ze zbioru liczb naturalnych,



Rysunek 2.2.1: Funkcje hazardowe: *Poi(1), Geo(0.5).*

$$r_N(n) = \frac{P(N = n)}{P(N \geqslant n)}.$$

W szczególności, dla rozkładu

- Poissona $Poi(\lambda)$ jest ona rosnąca dla $\lambda > 0$, (rys. 2.2.1);
- ujemnie dwumianowego $Bin^-(r, p)$ jest ona malejąca dla $r < 1$, rosnąca dla $r > 1$ i stała dla $r = 1$, tzn. dla rozkładu geometrycznego (rys 2.2.1).

Przybliżeniem funkcji $r_N(n)$ jest

$$\frac{\#\{i : N_i = n\}}{\#\{i : N_i \geqslant n\}}.$$

Nanosząc powyższe wartości na wykres dostaniemy przybliżoną funkcję hazardową, na podstawie której możemy stawiać hipotezy dotyczące typu rozkładu.

Przykład 2.2.6 Rozważmy portfel składający się z $n = 421240$ polis samochodowych. W tabeli, w drugiej kolumnie przedstawiono ilość polis n_k, które wygenerowały k szkód. Chcemy znaleźć rozkład szkód najlepiej opisujący nasze dane.

k	Obserwowane n_k	$Poi(0.131)$ r_k	Bin^-	$MixedPoi$
0	370412	369247	370460	370409
1	46545	48644	46413	46558
2	3935	3204	4044	3916
3	317	141	301	328
4	28	5	20	27
5	3	0	1	2

Wykres dla rekurencji Panjera nie jest liniowy, funkcja hazardowa nie jest monotoniczna. Sugeruje to, że rozkład ilości szkód nie będzie ani Poissona, ani dwumianowy ani ujemny dwumianowy. Liczymy teraz średnią, wariancję i skośność próbkową ilości szkód i dostajemy:

$$\begin{aligned} \overline{N} &= \frac{1}{n}\sum kn_k = 0.131 \\ S^2 &= \frac{1}{n}\sum (k-\overline{N})^2 n_k = 0.138 \\ A &:= \frac{1}{n}\sum (k-\overline{N})^3 n_k = 0.153 \end{aligned}$$

Średnia jest więc mniejsza od wariancji. Odrzuca to ponownie możliwość dopasowania rozkładu dwumianowego. Spróbujmy dopasować rozkład mieszany Poissona, gdzie zmienna mieszająca Θ przyjmuje dwie wartości: $P(\Theta = \theta_1) = p = 1 - P(\Theta = \theta_2)$. Średnia, wariancja i trzeci centralny moment Θ liczymy więc ze wzorów:

$$\begin{aligned} \mathrm{E}[\Theta] &= p\theta_1 + (1-p)\theta_2; \\ \mathrm{Var}[\Theta] &= p(\theta_1 - \mathrm{E}[\Theta])^2 + (1-p)(\theta_2 - \mathrm{E}[\Theta])^2; \\ m_3(\Theta) &= p(\theta_1 - \mathrm{E}[\Theta])^3 + (1-p)(\theta_2 - \mathrm{E}[\Theta])^3. \end{aligned}$$

Korzystając teraz ze wzoru (2.2.3) otrzymujemy, estymując $\mathrm{E}[N] = \overline{N}$, $\mathrm{Var}[N] = S^2$, $m_3(N) = A$, układ trzech równań z trzema niewiadomymi, który po rozwiązaniu daje: $p = 0.4633$, $\theta_1 = 0.2243$, $\theta_2 = 0.537$. Można przyjąć, że nasze dane pochodzą właśnie z takiego mieszanego rozkładu Poissona.

□

2.2.2 Własności ogólne

Rozważmy model złożony $S_N = \sum_{i=1}^{N} X_i$ dla niezależnych $(X_i)_{i\geqslant 1}$ o jednakowych rozkładach F_X i wartościach naturalnych oraz dla niezależnej od nich zmiennej liczącej N.

Rozkład sumy $S = S_N$ można przedstawić następującym wzorem:

$$F_S(x) = P(S \leqslant x) = \sum_{n=0}^{\infty} F_X^{*n}(x) P(N = n),$$

a stąd

$$f_S(x) = \sum_{n=0}^{\infty} f_X^{*n}(x) P(N = n).$$

Znalezienie rozkładu sumy losowej bezpośrednio z powyższych wzorów, podobnie jak w przypadku sumy deterministycznej ilości składników, jest zazwyczaj trudnym zadaniem. Dlatego też będziemy używali funkcji tworzących.

Zaczniemy od przypomnienia podstawowych wzorów dla sum losowej długości. Wzór na wartość oczekiwaną łącznych roszczeń w złożonym modelu, gdzie zmienne losowe $X_1, X_2, ...,$ mają wartość oczekiwaną $E[X]$ i wariancję $Var[X]$ ma postać

$$E[S] = E[X] \cdot E[N]. \tag{2.2.11}$$

Mówi on, że przeciętna wysokość łącznych roszczeń jest równa iloczynowi przeciętnej liczby roszczeń i przeciętnej wysokości pojedyńczego roszczenia.

Korzystając z lematu 2.2.2

$$\begin{aligned} E[S] &= E[X_1 + ... + X_N] = E[E[X_1 + X_2 + ... + X_N \mid N]] \\ &= \sum_{n=1}^{\infty} E[X_1 + ... + X_n | N = n] \Pr[N = n] \\ &= \sum_{n=1}^{\infty} E[X_1 + ... + X_n] \Pr[N = n] = \sum_{n=1}^{\infty} n E[X] \Pr[N = n] \\ &= E[X] \sum_{n=1}^{\infty} n \Pr[N = n] = E[X] E[N] \end{aligned}$$

Wzór na wariancję łącznych roszczeń jest następujący

$$Var[S] = E[N] \cdot Var[X] + (E[X])^2 \cdot Var[N]. \tag{2.2.12}$$

Z powyższego wzoru wynika, że wariancja składa się z dwóch składników; pierwszy z nich odnosi się do zmienności liczby roszczeń, drugi zaś do zmienności wysokości pojedynczego roszczenia.

Rzeczywiście, bezpośrednio z definicji:

$$\begin{aligned}
Var[S] &= E[S^2] - (E[S])^2 = E[E[(X_1 + ...X_N)^2|N] - (E[X]E[N])^2 = \\
&= \sum_{n=1}^{\infty} E[(X_1 + ...X_N)^2|N = n] \Pr(N = n) - (E[X]E[N])^2 = \\
&= \sum_{n=1}^{\infty} E[(X_1 + ... + X_n)^2] \Pr(N = n) - (E[X]E[N])^2 = \\
&= \sum_{n=1}^{\infty} E[\sum_{i=1}^{n} X_i \sum_{j=1}^{n} X_j] \Pr(N = n) - (E[X]E[N])^2 = \\
&= \sum_{n=1}^{\infty} E[\sum_{i=j=1}^{n} (X_i)^2 + \sum_{i \neq j=1}^{n} X_i X_j] \Pr(N = n) - (E[X]E[N])^2 = \\
&= \sum_{n=1}^{\infty} E[\sum_{i=j=1}^{n} (X_i)^2] \Pr(N = n) + \sum_{n=1}^{\infty} E[\sum_{i \neq j=1}^{n} X_i X_j] \Pr(N = n) - (E[X]E[N])^2 = \\
&= \sum_{n=1}^{\infty} nE[X^2] \Pr(N = n) + \sum_{n=1}^{\infty} (n^2 - n)(E[X])^2 \Pr(N = n) - (E[X]E[N])^2 = \\
&= E[X^2]E[N] + (E[X])^2 E[N^2] - (E[X])^2 E[N] - (E[X]E[N])^2 = \\
&= E[N]Var[X] + (E[X])^2 Var[N].
\end{aligned}$$

Używając funkcji tworzących, dostajemy szybciej związek pomiędzy rozkładem S_N a N i X, a także między momentami.

Fakt 2.2.7 *Zachodzą następujące wzory*

$$\begin{aligned}
&M_{S_N}(t) = M_N(\log M_X(t)), \\
&C_{S_N}(t) = C_N(C_X(t)), \\
&P_{S_N}(t) = P_N(P_X(t))
\end{aligned}$$

a stąd, dla momentów

$$\mathrm{E}\left[S_N\right] = \mathrm{E}\left[N\right] \mathrm{E}\left[X\right], \tag{2.2.13}$$

$$\mathrm{Var}\left[S_N\right] = \mathrm{E}\left[N\right] \mathrm{Var}\left[X\right] + \mathrm{Var}\left[N\right] \left(\mathrm{E}\left[X\right]\right)^2, \tag{2.2.14}$$

$$\begin{aligned}
\mathrm{E}\left[\left(S_N - \mathrm{E}\left[S_N\right]\right)^3\right] &= \mathrm{E}\left[\left(N - \mathrm{E}\left[N\right]\right)^3\right] \left(\mathrm{E}\left[X\right]\right)^3 \\
&\quad + 3\mathrm{Var}\left[N\right] \mathrm{E}\left[X\right] \mathrm{Var}\left[X\right] + \mathrm{E}\left[N\right] \mathrm{E}\left[\left(X - \mathrm{E}\left[X\right]\right)^3\right].
\end{aligned} \tag{2.2.15}$$

Dowód: Udowodnimy najpierw tożsamość dla funkcji tworzących momenty. Mamy

$$\mathrm{E}\left[e^{tS_N}\right] = \mathrm{E}\left[\mathrm{E}\left[e^{tS_N}|N\right]\right] = \sum_n \mathrm{E}\left[e^{tS_n}|N=n\right] P(N=n).$$

Korzystając z niezależności zmiennych X_i od S otrzymujemy

$$\begin{aligned}\mathrm{E}\left[e^{tS_N}\right] &= \sum_n \mathrm{E}\left[e^{tS_n}\right] P(N=n) = \sum_n \mathrm{E}\left[e^{t(X_1+\cdots+X_n)}\right] P(N=n)\\ &= \sum_n \mathrm{E}\left[\prod_{i=1}^n \exp(tX_i)\right] P(N=n)\\ &= \sum_n \prod_{i=1}^n \mathrm{E}\left[\exp(tX_i)\right] P(N=n) = \sum_n \prod_{i=1}^n M_{X_i}(t) P(N=n),\end{aligned}$$

gdzie w przedostatniej równości skorzystaliśmy z niezależności zmiennych losowych X_i. Ponieważ X_i mają ten sam rozkład więc $M_{X_i}(t) = M_X(t)$, a stąd

$$\mathrm{E}\left[e^{tS_N}\right] = \sum_n (M_X(t))^n P(N=n) = \sum_n e^{un} P(N=n),$$

gdzie $u = \log M_X(t)$. Ostatnia suma jest równa z definicji $M_N(u)$, a stąd otrzymujemy tezę.

Tożsamość dla funkcji tworzących kumulanty i funkcji tworzących prawdopodobieństwa wynika bezpośrednio z definicji. Dla wartości oczekiwanej mamy

$$C^{(1)}_{S_N}(t) = C^{(1)}_N(C_X(t)) C^{(1)}_X(t)$$

i wstawiając $t = 0$ otrzymujemy

$$E[S_N] = C^{(1)}_{S_N}(0) = C^{(1)}_N(0) C^{(1)}_X(0) = \mathrm{E}\left[N\right]\mathrm{E}\left[X\right].$$

Dla drugiej pochodnej mamy

$$C^{(2)}_{S_N}(t) = C^{(2)}_N(C_X(t))[C^{(1)}_X(t)]^2 + C^{(1)}_N(C_X(t)) C^{(2)}_X(t)$$

i wstawiając $t = 0$ otrzymujemy

$$Var[S_N] = C^{(2)}_{S_N}(0) = C^{(2)}_N(0)[C^{(1)}_X(0)]^2 + C^{(1)}_N(0) C^{(2)}_X(0),$$

co daje tezę. Obliczając trzecią pochodną dostajemy

$$\begin{aligned}C^{(3)}_{S_N}(t) &= C^{(3)}_N(C_X(t))[C^{(1)}_X(t)]^3 + 2C^{(2)}_N(C_X(t)) C^{(1)}_X(t) C^{(2)}_X(t)\\ &\quad + C^{(2)}_N(C_X(t)) C^{(1)}_X(t) C^{(2)}_X(t) + C^{(1)}_N(C_X(t)) C^{(3)}_X(t)\\ &= C^{(3)}_N(C_X(t))[C^{(1)}_X(t)]^3 + 3C^{(2)}_N(C_X(t)) C^{(1)}_X(t) C^{(2)}_X(t)\\ &\quad + C^{(1)}_N(C_X(t)) C^{(3)}_X(t)\end{aligned}$$

i biorąc $t = 0$ otrzymujemy wynik. □

Przykład 2.2.8 Rozważmy portfel ubezpieczeń, w którym ilość pojawiających się roszczeń opisuje i rozkład wysokości pojedynczego roszczenia opisują tabele poniżej. Szukamy rozkładu zmiennej losowej $S_N = X_1 + \ldots + X_N$.

Rozkład zmiennej N.

n	0	1	2	3	4	5	6	7	8
p_n	0.05	0.10	0.15	0.20	0.25	0.15	0.06	0.03	0.01

Rozkład wysokości pojedynczego roszczenia.

x	1	2	3	4	5	6	7	8	9	10
$f_X(x)$	0.15	0.2	0.25	0.125	0.075	0.05	0.05	0.05	0.025	0.025

Obliczymy funkcję tworzącą zmiennej losowej N

$$P_N(t) = 0.05 + 0.10t + 0.15t^2 + 0.2t^3 + 0.25t^4 + 0.15t^5 + 0.06t^6 + 0.03t^7 + 0.01t^8,$$

więc

$$\begin{aligned} P_S(t) = {} & 0.05 + 0.10P_X(t) + 0.15(P_X(t))^2 + 0.2(P_X(t))^3 + 0.25(P_X(t))^4 \\ & + 0.15(P_X(t))^5 + 0.06(P_X(t))^6 + 0.03(P_X(t))^7 + 0.01(P_X(t))^8. \end{aligned}$$

Funkcja tworząca zmiennej losowej X ma postać

$$\begin{aligned} P_X(t) = {} & 0.15t + 0.2t^2 + 0.25t^3 + 0.125t^4 + 0.075t^5 \\ & + 0.05t^6 + 0.05t^7 + 0.05t^8 + 0.025t^9 + 0.025t^{10}. \end{aligned}$$

Łącząc oba wzory otrzymujemy

$$\begin{aligned} P_S(t) = {} & 0.05 + 0.015t + 0.02338t^2 + 0.03468t^3 + 0.03258t^4 \\ & + 0.03579t^5 + 0.03981t^6 + 0.04356t^7 + \ldots \end{aligned}$$

Teraz, wartości $P(S = k)$ odczytujemy jako współczynniki przy t^k, $k = 0, \ldots, 80$.

□

2.2.3 Złożony rozkład dwumianowy

Załóżmy, że portfel składa się z niezależnych ryzyk $Y_1, \ldots, Y_n$. Każde roszczenie pojawia się z prawdopodobieństwem p, niezależnie od jego wielkości X_i. Roszczenie ma więc strukturę: $Y_i = I_i X_i$, gdzie $(X_i)_{i\geqslant 1}$, $(I_i)_{i\geqslant 1}$ są niezależnymi od siebie ciągami niezależnych zmiennych losowych o tych samych rozkładach z $P(I_i = 1) = p = 1 - P(I_i = 0)$. Łączna szkoda ma wtedy postać $S = \sum_{i=1}^{N} X_i$, gdzie $P(N = k) = \binom{n}{k} p^k q^{n-k}$, gdzie $p \in (0, 1)$, $q = 1 - p$, $k = 0, 1, \ldots, n$, tzn. N ma rozkład dwumianowy. Mówimy wtedy, że S_N ma złożony rozkład dwumianowy i zapisujemy wtedy $S_N \sim CBin(n, p, F_X)$. Mamy

$$M_N(t) = (q + pe^t)^n$$
$$C_N(t) = n \log(q + pe^t),$$
$$M_{S_N}(t) = (q + pM_X(t))^n,$$
$$C_{S_N}(t) = n \log(q + pM_X(t)),$$

a stąd

$$\mathrm{E}\left[S_N\right] = np\mathrm{E}\left[X\right],$$
$$\mathrm{Var}\left[S_N\right] = np\mathrm{Var}\left[X\right] + npq(\mathrm{E}\left[X\right])^2,$$
$$\mathrm{E}\left[\left(S_N - \mathrm{E}\left[S_N\right]\right)^3\right] = np\mathrm{E}\left[X^3\right] - 3np^2\mathrm{E}\left[X^2\right]\mathrm{E}\left[X\right] + 2np^3(\mathrm{E}\left[X\right])^3.$$

W szczególności, dla deterministycznej szkody o wielkości x_0 dostajemy:

$$\mathrm{E}\left[\left(S_N - \mathrm{E}\left[S_N\right]\right)^3\right] \begin{cases} > 0 & \text{dla} \quad p < \frac{1}{2} \\ = 0 & \text{dla} \quad p = \frac{1}{2} \\ < 0 & \text{dla} \quad p > \frac{1}{2} \end{cases}.$$

Złożone rozkłady dwumianowe mogą więc służyć do modelowania portfeli o dowolnym znaku skośności.

Złożony rozkład dwumianowy ma bardzo ważną cechę: suma niezależnych $CBin(m^{(i)}, p, F)$ jest znowu złożonym rozkładem dwumianowym.

Twierdzenie 2.2.9 *Jeśli $S^{(1)}, S^{(2)}, \ldots, S^{(n)}$ są niezależnymi zmiennymi losowymi i $S^{(i)}$ ma rozkład złożony $CBin(m^{(i)}, p, F)$ to*

$$S = \sum_{i=1}^{n} S^{(i)}$$

ma rozkład złożony ujemny dwumianowy $CBin(m, p, F)$ dla

$$m = \sum_{i=1}^{n} m^{(i)}.$$

2.2.4 Złożony rozkład Poissona

Niech $P(N = n) = \frac{\lambda^n}{n!}e^{-\lambda}$, $\lambda > 0$, $n = 0, 1, \ldots$. Jeżeli $S_N = \sum_{i=1}^{N} X_i$, dla ciągu niezależnych zmiennych losowych $(X_i)_{i \geqslant 0}$ o dystrybuancie F_X, niezależnego od N, to mówimy, że S_N ma **złożony rozkład Poissona** i zapisujemy $S_N \sim CPoi(\lambda, F_X)$.

Dla złożonego rozkładu Poissona mamy

$$M_N(t) = e^{\lambda(e^t - 1)},\ C_N(t) = \lambda(e^t - 1),$$

$$M_{S_N}(t) = e^{\lambda(M_X(t)-1)}$$
$$C_{S_N}(t) = \lambda(M_X(t)-1),$$

co daje $C^{(k)}_{S_N}(0) = \lambda M^{(k)}_X(0)$ i dalej

$$\mathrm{E}\left[S_N\right] = \lambda \mathrm{E}\left[X\right],$$
$$\mathrm{Var}\left[S_N\right] = \lambda \mathrm{E}\left[X^2\right],$$
$$\mathrm{E}\left[\left(S_N - \mathrm{E}\left[S_N\right]\right)^3\right] = \lambda \mathrm{E}\left[X^3\right].$$

Przykład 2.2.10 Rozważmy portfel ubezpieczeń, w którym pojawiające się roszczenia mają rozkład Poissona z parametrem $\lambda = 10$, natomiast rozkład wysokości pojedyńczego roszczenia jest jak w tabeli. Policzymy rozkład zmiennej losowej $S = X_1 + \ldots, X_N$.

Rozkład wysokości pojedynczego roszczenia.

i	1	2	12	13	18
$P(X=i)$	0.1	0.35	0.05	0.2	0.3

Znając funkcję tworzącą zmiennej losowej N, $P_N(t) = e^{10(t-1)}$ dostajemy

$$P_S(t) = e^{10(P_X(t)-1)}.$$

Funkcja tworząca zmiennej losowej X ma postać

$$P_X(t) = 0.1t + 0.35t^2 + 0.05t^{12} + 0.2t^{13} + 0.3t^{18}. \tag{2.2.16}$$

Wstawiając (2.2.16) do wzoru na P_S otrzymujemy funkcję tworzącą rozkładu złożonego

$$P_S(t) = e^{10(0.1t+0.35t^2+0.05t^{12}+0.2t^{13}+0.3t^{18}-1)}.$$

Rozwijając powyższe równanie w szereg Taylora, możemy odczytać rozkład zmiennej S

$$P_S(t) = 0.0000454 + 0.0000454t + 0.00018t^2 + 0.00017t^3 + \cdots$$

□

Przykład 2.2.11 Niech $S = \sum_{i=1}^{N} X_i$, gdzie $X_i \sim Poi(\beta)$, a $N \sim Poi(\lambda)$. Funkcja tworząca prawdopodobieństwa jest postaci

$$P_S(t) = \exp\left(\lambda\left(\exp(\beta(t-1)) - 1\right)\right)$$

Rozkład taki nazywamy *rozkładem Neymanna typu A*. W szczególnych przypadkach rozkład ten aproksymujemy:

- Jeżeli λ jest duże i $\lambda\beta > 0$, to $\frac{S-\lambda\beta}{\sqrt{\lambda\beta(1+\beta)}} \sim N(0,1)$;
- Jeżeli $\lambda \in (0,1)$, to S ma przesunięty rozkład Poissona, tzn. $P(S=0) = (1-\lambda) + \lambda e^{-\beta}$, $P(S=k) = \lambda e^{-\beta}\beta^k/k!$, $k \geqslant 1$.
- Jeżeli β jest *małe*, to S przybliżamy za pomocą rozkładu $Poi(\lambda\beta)$.

□

Przykład 2.2.12 Niech $S = \sum_{i=1}^{N} X_i$, gdzie $X_i \sim Exp(\alpha)$, a $N \sim Poi(\lambda)$. Stosując wzór na prawdopodobieństwo całkowite,

$$F_{S_N}(x) = \exp(-\lambda)\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}P(\sum_{i=1}^{n} X_i \leqslant x).$$

Wiemy jednak z Twierdzenia 2.1.11, że $S_n = \sum_{i=1}^{n} X_i$ ma rozkład Gamma o gęstości $f_{S_n}(x) = \frac{\alpha^n}{(n-1)!}\exp(-\alpha x)x^{n-1}$. Stąd, różniczkując $F_{S_N}(x)$,

$$f_{S_N}(x) = \exp(-(\lambda+\alpha x))2\sqrt{\frac{\lambda\alpha}{x}}B_1(2\sqrt{\lambda\alpha x}),$$

gdzie $B_1(x) = \sum_{k=0}^{\infty}\frac{(x/2)^{2k+1}}{k!(k+1)!}$ jest zmodyfikowaną funkcją Bessela.

□

Przykład 2.2.13 Niech $S = S_N = X_1 + \cdots + X_N$, gdzie $N \sim Poi(\lambda)$. Załóżmy, że X_i mają rozkład logarytmiczny

$$P(X=k) = \frac{p^k}{-k\ln(1-p)}, \quad p \in (0,1), k \geqslant 1.$$

Wtedy S_N ma rozkład ujemny dwumianowy. Policzymy w tym celu najpierw funkcję tworzącą dla X. Dla $t < -\ln(1-p)$ mamy

$$\begin{aligned} M_X(t) &= -\frac{1}{\ln(1-p)} \sum_{k=1}^{\infty} \exp(tk) \frac{p^k}{k} \\ &= -\frac{1}{\ln(1-p)} \sum_{k=1}^{\infty} \int_0^{p\exp(t)} u^{k-1} du \\ &= -\frac{1}{\ln(1-p)} \int_0^{p\exp(t)} \sum_{k=1}^{\infty} u^{k-1} du \\ &= -\frac{1}{\ln(1-p)} \int_0^{p\exp(t)} \frac{1}{1-u} du \\ &= \frac{\ln(1-p\exp(t))}{\ln(1-p)}, \end{aligned}$$

a stąd (patrz wzór (2.2.4))

$$\begin{aligned} M_S(t) &= \exp\left(\lambda\left(\tfrac{\ln(1-p\exp(t))}{\ln(1-p)} - 1\right)\right) \\ &= \exp\left(\tfrac{\lambda}{\ln(1-p)} \ln\left(\tfrac{1-pe^t}{1-p}\right)\right) \\ &= \left(\tfrac{1-p}{1-pe^t}\right)^r, \end{aligned}$$

gdzie $r = -\frac{\lambda}{\ln(1-p)}$. Jest to funkcja tworząca dla rozkładu $Bin^-(-\frac{\lambda}{\ln(1-p)}, 1-p)$.

□

Sumy niezależnych zmiennych o rozkładach złożonych Poissona są zawsze zmiennymi o złożonym rozkładzie Poissona.

Twierdzenie 2.2.14 *Jeśli $S^{(1)}, S^{(2)}, \ldots, S^{(n)}$ są niezależnymi zmiennymi losowymi i $S^{(i)}$ ma rozkład złożony Poissona z parametrem $\lambda^{(i)}$ i dystrybuancie składników $F^{(i)}$ ($CPoi(\lambda^{(i)}, F^{(i)})$, dla $i = 1, \ldots, n$), to*

$$S = \sum_{i=1}^{n} S^{(i)}$$

ma rozkład złożony Poissona $CPoi(\lambda, F)$ dla

$$\lambda = \sum_{i=1}^{n} \lambda^{(i)},$$

$$F(x) = \sum_{i=1}^{n} \frac{\lambda^{(i)}}{\lambda} F^{(i)}(x).$$

Z założenia

$$M_{S^{(i)}}(t) = \exp(\lambda^{(i)}(M_{X^{(i)}}(t) - 1)),$$

gdzie zmienna losowa $X^{(i)}$ ma rozkład $F^{(i)}$. Z niezależności dostajemy

$$\begin{aligned} M_S(t) &= \prod_{i=1}^{n} M_{S^{(i)}}(t) = \exp\left(\sum_{i=1}^{n} \lambda^{(i)}(M_{X^{(i)}}(t) - 1)\right) \\ &= \exp\left(\lambda\left(\sum_{i=1}^{n} \frac{\lambda^{(i)}}{\lambda} M_{X^{(i)}}(t) - 1\right)\right) \end{aligned}$$

co jest funkcją tworzącą żądanego rozkładu złożonego. □

Przykład 2.2.15 Niech $x_1, \ldots, x_n$ będą wartościami wypłat w K portfelach, których wielkości są losowe, niezależne $N_1, \ldots, N_K$ o rozkładach Poissona z parametrami $\lambda_1, \ldots, \lambda_K$ odpowiednio. Wtedy z powyższego twierdzenia

$$S = x_1 N_1 + \ldots + x_K N_K = \sum_{i=1}^{N_1} x_1 + \ldots + \sum_{i=1}^{N_n} x_n$$

ma rozkład złożony Poissona $CPoi(\lambda, F)$ dla

$$\lambda = \lambda_1 + \ldots + \lambda_K,$$

$$F(x) = \sum_{i=1}^{K} \frac{\lambda_i}{\lambda} I_{([x_i,\infty)}(x).$$

□

Przykład 2.2.16 Przypuśćmy, że $S^{(1)}$ ma złożony rozkład Poissona z parametrem $\lambda^{(1)} = 0.5$ i wielkościami szkód $1, 2, 3, 4, 5$ z prawdopodobieństwami $0.15, 0.3, 0.3, 0.05, 0.2$, odpowiednio. $S^{(2)}$ ma rozkład złożony Poissona z $\lambda^{(2)} = 0.8$ oraz rozkładem szkód o wielkościach $1, 2, 3$ z prawdopodobieństwami 0.25, 0.5, 0.25, odpowiednio. Ponadto $S^{(3)}$ ma rozkład złożony Poissona z $\lambda^{(3)} = 1.2$ oraz rozkładem szkód o wielkościach $3, 4, 5$ z prawdopodobieństwami 0.15, 0.5, 0.35, odpowiednio. Jeśli $S^{(1)}, S^{(2)}, S^{(3)}$ są niezależne, policzymy rozkład $S = S^{(1)} + S^{(2)} + S^{(3)}$.

Z Twierdzenia 2.2.14 mamy

$$\begin{aligned} \lambda &= \lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)} = 0.5 + 0.8 + 1.2 = 2.5, \\ f_X(x) &= \frac{\lambda^{(1)}}{\lambda} f_{X^{(1)}}(x) + \frac{\lambda^{(2)}}{\lambda} f_{X^{(2)}}(x) + \frac{\lambda^{(3)}}{\lambda} f_{X^{(3)}}(x), \\ f_X(x) &= 0.2 f_{X^{(1)}}(x) + 0.32 f_{X^{(2)}}(x) + 0.48 f_{X^{(3)}}(x). \end{aligned}$$

zatem

$$f_X(x) = \begin{cases} 0.110 & \text{dla} \quad x = 1 \\ 0.220 & \text{dla} \quad x = 2 \\ 0.212 & \text{dla} \quad x = 3 \\ 0.250 & \text{dla} \quad x = 4 \\ 0.208 & \text{dla} \quad x = 5 \end{cases}.$$

Otrzymujemy złożony rozkład Poissona z parametrem $\lambda = 2.5$ i rozmiarem szkód podanym powyżej.

□

Twierdzenie 2.2.17 *Jeżeli zmienna $S = X_1 + \cdots + X_N$ ma złożony rozkład Poissona $CPoi(\lambda, F)$, z dyskretnym rozkładem indywidualnych roszczeń o dystrybuancie F i funkcji prawdopodobieństwa $\pi_i = P(X = x_i), \quad i = 1, \ldots, K$, to*

(i) zmienne losowe $N_1, N_2, \ldots, N_K$, zdefiniowane przez

$$N_i = card\{k : X_k = x_i\},$$

$i = 1, \ldots, K$ są wzajemnie niezależne, (wtedy $S = x_1 N_1 + \cdots + x_K N_K$)

(ii) zmienne losowe N_i mają rozkłady Poissona z parametrami $\lambda^{(i)} = \lambda\pi_i$, $i = 1, 2, \ldots, K$.

Dowód. Załóżmy, że $N = \sum_{i=1}^{K} N_i = k$. Wtedy

$$\begin{aligned} \mathrm{E}\left[\exp\left(\sum_{i=1}^{K} t_i N_i\right)\right] &= \sum_{k=0}^{\infty} \mathrm{E}\left[\exp\left(\sum_{i=1}^{K} t_i N_i\right) \mid N = k\right] \Pr(N = k) \\ &= \sum_{k=0}^{\infty} (\pi_1 e^{t_1} + \pi_2 e^{t_2} + \ldots + \pi_K e^{t_K})^k \frac{e^{-\lambda}\lambda^k}{k!} \qquad (2.2.17) \\ &= \exp(-\lambda) \exp\left(\lambda \sum_{i=1}^{K} \pi_i e^{t_i}\right) \\ &= \prod_{i=1}^{K} \exp[\lambda\pi_i(e^{t_i} - 1)]. \qquad (2.2.18) \end{aligned}$$

Rzeczywiście, rozkład wektora $(N_1, \ldots, N_K)$ pod warunkiem $N = k$ jest wielomianowy, tzn.

$$P(N_1 = k_1, \ldots, N_K = k_K | N = k) = \frac{k!}{k_1! \ldots k_K!} \pi_1^{k_1} \ldots \pi_K^{k_K},$$

a stąd

$$\mathrm{E}\left[\exp\left(\sum_{i=1}^{K} t_i N_i\right) | N = k\right] = \left[\pi_1 e^{t_1} + \ldots + \pi_K e^{t_K}\right]^k,$$

co zostało użyte w (2.2.17).

W równaniu (2.2.18) otrzymaliśmy więc iloczyn K funkcji, z których każda jest innej zmiennej t_i. Powyższa formuła pokazuje wzajemną niezależność zmiennych losowych N_i. Dalej, podstawiając $t_i = t$ i $t_j = 0$ dla $j \neq i$ otrzymujemy

$$\mathrm{E}\left[\exp(tN_i)\right] = \exp[\lambda\pi_i(e^t - 1)]. \tag{2.2.19}$$

Wzór (2.2.19) jest funkcją tworzącą rozkładu Poissona z parametrem $\lambda\pi_i$. □

2.2.5 Złożony rozkład ujemny dwumianowy

Rozkład łącznych roszczeń jest rozkładem złożonym ujemnym dwumianowym, jeśli jego dystrybuanta jest postaci

$$F_S(x) = \Pr(S \leqslant x) = \sum_{n=0}^{\infty} \binom{r+n-1}{n} p^r q^n G^{*n}(x). \tag{2.2.20}$$

Wartość oczekiwana, wariancja i skośność zmiennej losowej S w tym przypadku dane są przez

$$\begin{aligned} E[S] &= \frac{rq}{p}E[X], \\ Var[S] &= \frac{rq}{p}E[X^2] + \frac{rq^2}{p^2}\left(E[X]\right)^2 \\ \mathrm{E}\left[\left(S_N - \mathrm{E}\left[S_N\right]\right)^3\right] &= \frac{rq}{p}\mathrm{E}\left[X^3\right] + 3\frac{rq^2}{p^2}\mathrm{E}\left[X^2\right]\mathrm{E}\left[X\right] + 2\frac{rq^3}{p^3}\left(\mathrm{E}\left[X\right]\right)^3. \end{aligned}$$

Ponadto

$$M_S(t) = \left(\frac{p}{1 - qM_X(t)}\right)^r. \tag{2.2.21}$$

Przykład 2.2.18 Załóżmy, że szkoda S ma złożony rozkład ujemny dwumianowy, z parametrami $r = 5$ i $p = 0.6$ i rozkładem szkód jak w poniższej tabeli. Wyliczymy $\Pr(S = x)$.

Rozkład pojedynczego roszczenia.

x	1	2	3	4	5	6
$f_X(x)$	0.05	0.10	0.15	0.20	0.25	0.25

Mamy więc

$$P_S(t) = \left(\frac{0.6}{1 - 0.4P_X(t)}\right)^5, \tag{2.2.22}$$

gdzie funkcja tworząca zmiennej losowej X jest równa

$$P_X(t) = 0.05t + 0.1t^2 + 0.15t^3 + 0.2t^4 + 0.25t^5 + 0.25t^6. \tag{2.2.23}$$

Wstawiając (2.2.23) do (2.2.22) otrzymujemy

$$P_S(t) = \left(\frac{0.6}{1 - 0.4(0.05t + 0.1t^2 + 0.15t^3 + 0.2t^4 + 0.25t^5 + 0.25t^6)}\right)^5. \tag{2.2.24}$$

Rozwijając powyższe równanie w szereg Taylora, możemy odczytać rozkład zmiennej S

$$\begin{aligned} P_S(t) = {} & 0.0777 + 0.0077t + 0.0160t^2 + 0.0252t^3 + 0.0359t^4 + \\ & + 0.0486t^5 + 0.0564t^6 + 0.0280t^7 + 0.0365t^8 + 0.0427t^9 + ... \end{aligned}$$

□

W szczególności, gdy $r = 1$ otrzymujemy **złożony rozkład geometryczny** z atomem w zerze i piszemy $S_N \sim CGeo(p, F)$. Jeśli dla złożonego rozkładu geometrycznego X ma rozkład standardowy wykładniczy $F_X(x) = 1 - e^{-x}, x > 0$, to $M_X(t) = \frac{1}{1-t}$ oraz

$$M_{S_N}(t) = p + q\frac{p}{p-t}, \tag{2.2.25}$$

co oznacza, że rozkład sumy jest mieszanką atomu w zerze wielkości p i rozkładu wykładniczego z parametrem p, tzn.

$$F_{S_N}(x) = pI_{(0,\infty)}(x) + q(1 - e^{-px}).$$

Powyższy fakt można zobaczyć w bardziej ogólnym kontekście.

Przykład 2.2.19 Rozkłady złożone ujemne dwumianowe można czasem utożsamić z rozkładami złożonymi dwumianowymi. Niech $S = S_N = X_1 + \cdots + X_N$, gdzie $N \sim Bin^-(r, p)$, $r \in \mathbb{N}$, X_i mają rozkład $Exp(\lambda)$, tzn. $S_N \sim CBin^-(r, p, Exp(\lambda))$. Porównanie funkcji tworzących momenty daje jednak również, że

$$S_N \sim CBin(r, 1 - p, Exp(\lambda p)).$$

Jeżeli natomiast X_i mają rozkład $Geo(\lambda)$, tzn. $S_N \sim CBin^-(r, p, Geo(\lambda))$ to zachodzi również

$$S_N \sim CBin(r, 1 - p, Geo(\lambda)).$$

Warto tu zauważyć, że wyjściowy rozkład ujemny dwumianowy ma nośnik nieskończony, a rozkład dwumianowy ma nośnik $\{0, \ldots, r\}$.

□

Używając funkcji tworzących momenty, jak w Twierdzeniu 2.2.14, otrzymujemy zamkniętość na operację splotu klasy rozkładów złożonych ujemnych dwumianowych.

Twierdzenie 2.2.20 *Jeśli $S^{(1)}, S^{(2)}, \ldots, S^{(n)}$ są niezależnymi zmiennymi losowymi i $S^{(i)}$ ma rozkład złożony $CBin^{-}(r^{(i)}, p, F)$ to*

$$S = \sum_{i=1}^{n} S^{(i)}$$

ma rozkład złożony ujemny dwumianowy $CBin^{-}(r, p, F)$ dla

$$r = \sum_{i=1}^{n} r^{(i)}.$$

2.2.6 Wzory rekurencyjne Panjera

Przypomnijmy wzór rekurencyjny dla rozkładu licznika ilości szkód wprowadzony w równaniu (2.2.8). Oznaczmy znowu funkcję prawdopodobieństwa zmiennej N przez $p_k = f_N(k) = P(N = k), k \in \mathbb{N}$. Załóżmy, że

$$p_k = \left(a + \frac{b}{k}\right) p_{k-1}, \ k \geqslant 1,$$

dla pewnego doboru parametrów a i b. Zapisując to inaczej dostajemy

$$k\frac{p_k}{p_{k-1}} = ka + b =: l(k).$$

W szczególności, dla rozkładu dwumianowego $Bin(n, p)$

$$a := \frac{-p}{1-p}, \ b := \frac{p(n+1)}{1-p},$$

Poissona $Poi(\lambda)$,

$$a := 0, \ b := \lambda$$

oraz ujemnie dwumianowego $Bin^{-}(r, p)$,

$$a := q, \ b := (r-1)q.$$

Twierdzenie 2.2.21 *Jeśli $S = X_1 + \cdots + X_N$ jest sumarycznym roszczeniem w portfelu, $g_k :== p_X(k) = P(X_i = k)$, $k \in \mathbb{N}$ oznacza rozkład pojedynczego roszczenia, to*

$$P(S=k) = \begin{cases} \sum_{i=0}^{\infty} g_0^i \cdot p_i & dla \quad k=0 \\ \frac{1}{1-ag_0} \sum_{i=1}^{k} (a + \frac{b \cdot i}{k}) g_i P(S=k-i) & dla \quad k \geqslant 1 \end{cases}$$

oraz

$$\mathrm{E}\left[S^n\right] = \frac{1}{1-a} \sum_{k=1}^{n} n \binom{n}{k} (a + b\frac{k}{n}) \mathrm{E}\left[S^{n-k}\right] \mathrm{E}\left[X^k\right], \ n \geqslant 1.$$

Dowód: Niech $f_k := P(S = k)$. Rezultat dla $k = 0$ wynika bezpośrednio ze wzoru na prawdopodobieństwo całkowite. Dla pozostałych k zauważmy najpierw, że dla jednakowo rozłożonych $X_1, X_2, \ldots$ mamy, dla $S_n = X_1 + \cdots + X_n$,

$$nE(X_1 \mid S_n = k) = \sum_{m=1}^{n} E(X_m \mid S_n = k) = E(S_n \mid S_n = k) = k,$$

Korzystając z tego, że $\frac{1}{n} = E(\frac{X_1}{k} | S_n = k)$ mamy

$$\begin{aligned} f_k &= \sum_{n=1}^{\infty} P(S=k|N=n)p_n = \\ &= \sum_{n=1}^{\infty} p_n g_k^{*n} = \sum_{n=1}^{\infty} (a + \frac{b}{n}) p_{n-1} g_k^{*n} = \\ &= \sum_{n=1}^{\infty} p_{n-1} E[a + b\frac{X_1}{k} \mid \sum_{i=1}^{n} X_i = k] g_k^{*n}. \end{aligned}$$

Licząc wprost z definicji, otrzymujemy dla $n > 1$

$$E[a+b\frac{X_1}{k} \mid \sum_{i=1}^{n} X_i = k] = (a+b\cdot 0/k)\frac{g_0 g_k^{*n-1}}{g_k^{*n}} + (a+b\cdot 1/k)\frac{g_1 g_{k-1}^{*n-1}}{g_k^{*n}} + \cdots + (a+b\cdot k/k)\frac{g_k g_0^{*n-1}}{g_k^{*n}},$$

oraz dla $n = 1$

$$E[a + b\frac{X_1}{k} \mid \sum_{i=1}^{1} X_i = k] = a + b.$$

Wstawiając te wyrażenia do sumy względem n, skracając g_k^{*n} i wyodrębniając wyraz $n = 1$

mamy

$$\begin{aligned}
f_k &= p_0(a+b)g_k + \sum_{n=2}^{\infty}\sum_{i=0}^{k} p_{n-1}(a+b\frac{i}{k})g_i g_{k-i}^{*(n-1)} = \\
&= p_0(a+b)g_k + \sum_{i=0}^{k}(a+b\frac{i}{k})g_i \sum_{n=1}^{\infty} p_n g_{k-i}^{*(n)} = \\
&= p_0(a+b)g_k + ag_o\sum_{n=1}^{\infty} p_n g_k^{*n} + \sum_{i=1}^{k-1}(a+b\frac{i}{k})g_i \sum_{n=1}^{\infty} p_n g_{k-i}^{*n} + (a+b)g_k \sum_{n=1}^{\infty} p_n g_0^{*n} \\
&= ag_0 f_k + \sum_{i=1}^{k-1}(a+b\frac{i}{k})g_i f_{k-i} + (a+b)g_k f_0 \\
&= ag_0 f_k + \sum_{i=1}^{k}(a+b\frac{i}{k})g_i f_{k-i}.
\end{aligned}$$

Wyliczając f_k z tego równania otrzymujemy tezę. □

Uwaga 2.2.22 Przewaga rekurencji Panjera nad postępowaniem rekurencyjnym bezpośrednio z definicji (wzór (2.1.7)) polega na różnicy w złożoności obliczeniowej. W celu obliczenia $P(S_n = j)$ potrzebujemy w pierwszym przypadku $O(j^2)$ operacji, podczas gdy w drugim przypadku algorytm wymaga $O(j^3)$ obliczeń.

Przykład 2.2.23 Niech $N \sim Poi(\lambda)$ i $X > 0$. Wtedy

$$P(S=k) = \begin{cases} \exp(-\lambda) & \text{dla} \quad k=0 \\ \frac{\lambda}{k}\sum_{i=1}^{k} i g_i P(S=k-i) & \text{dla} \quad k \geqslant 1 \end{cases}$$

□

Przykład 2.2.24 Niech $N \sim Geo(p)$ i $X > 0$. Wtedy

$$P(S=k) = p\sum_{i=1}^{k} g_i P(S=k-i), \qquad k \geqslant 1.$$

□

2.3 Twierdzenia graniczne-aproksymacje

2.3.1 Aproksymacja rozkładem dwumianowym i Poissona.

Przykład 2.3.1 Niech $X_i \sim Bin(1, p_i)$, $i = 1, \ldots, n$, $S_n = X_1 + \cdots + X_n$. Wtedy nie można użyć Twierdzenia 2.1.11. Ponieważ $\mathrm{E}[S_n] \geqslant \mathrm{Var}[S_n]$, aproksymacja zmienną lo-

sową N o rozkładzie dwumianowym (dla której $\mathrm{E}[N] \geqslant \mathrm{Var}[N]$) jest dopuszczalna. Niech $N \sim Bin(n,p)$, gdzie $p = \frac{\sum_{i=1}^{n} p_i}{n}$. Ponieważ $\mathrm{E}[N] = np$, więc $\mathrm{E}[N] = \mathrm{E}[S_n]$. Mamy jednak $\mathrm{Var}[N] - \mathrm{Var}[S_n] = \sum_{i=1}^{n} p_i^2(1-\frac{1}{n})$.

□

Przykład 2.3.2 Niech $X_i \sim Bin(1,p_i)$, $i = 1,\ldots,n$. Można przybliżyć rozkład S_n rozkładem Poissona tak, aby została zachowana średnia. Mamy $\mathrm{E}[S_n] = \sum p_i$, $\mathrm{Var}[S_n] = \sum p_i(1-p_i)$.

Niech $N \sim Poi(\lambda)$, $\lambda = \sum p_i$. Wtedy $\mathrm{E}[N] = \mathrm{E}[S_n]$. Następuje jednak przeszacowanie wariancji: mamy $\mathrm{Var}[N] - \mathrm{Var}[S_n] = \sum p_i^2$. Stosowanie aproksymacji Poissonowskiej ma sens w przypadku, gdy $\mathrm{E}[S_n] \approx \mathrm{Var}[S_n]$, tzn., gdy $\sum p_i^2$ jest małe. W przeciwnym przypadku aproksymacja ta jest gorsza niż w przypadku aproksymacji dwumianowej. Przykładowo, niech $n = 19$, $p_i = 0.05i$, $i = 1,\ldots,19$, wtedy $\sum p_i^2 = 6.175$.

□

Przykład 2.3.3 Jeżeli S_n ma rozkład $Bin(n,p_n)$ oraz p_n jest małe, to rozkład S_n przybliża się w praktyce rozkładem Poissona. Sytuacja taka nastąpi np. wtedy, gdy w portfelu mamy n ryzyk, każde generujące szkodę X_i, $i = 1,\ldots,n$ i interesują nas te szkody, które przekroczą poziom u. Wtedy $K = \sum_{i=1}^{n} I(X_i > u)$ ma rozkład $Bin(n,p)$, $p = P(X_i > u)$. Jeżeli u jest duże, to p jest małe i aproksymacja ma sens.

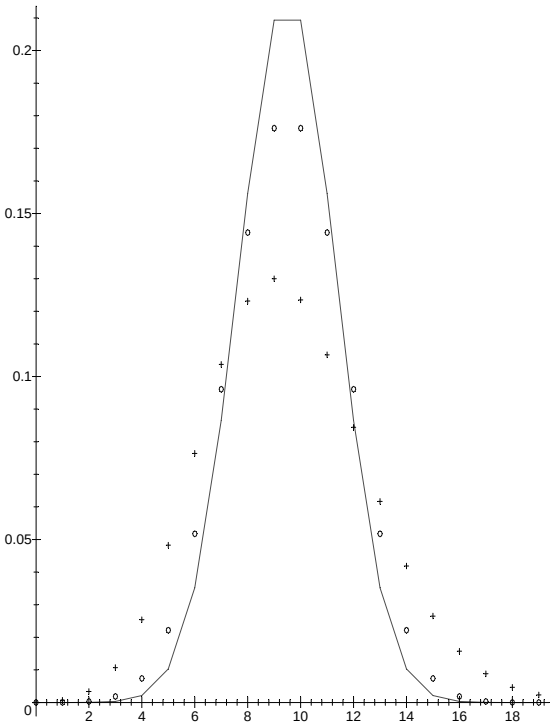
□

2.3.2 Aproksymacja rozkładem normalnym.

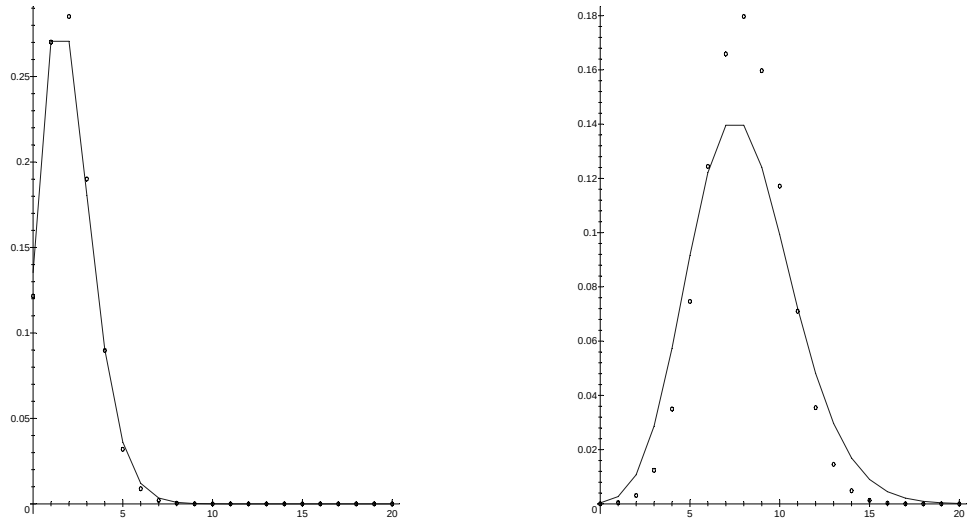
Jeśli $X_1, X_2, \ldots$ są zmiennymi losowymi niezależnymi o jednakowym rozkładzie, skończonej wartości średniej $\mu = \mathrm{E}[X] < \infty$ i wariancji $\sigma^2 = \mathrm{Var}[X] < \infty$, to z Centralnego Twierdzenia Granicznego, dla $S_n = X_1 + \ldots + X_n$ mamy

$$P\left(\frac{S_n - \mathrm{E}[S_n]}{\sqrt{\mathrm{Var}[S_n]}} \leqslant x\right) \to_{n\to\infty} \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$$

Wartości $\Phi(x)$ są zawarte w tablicach rozkładu normalnego i można je wykorzystać do przybliżenia rozkładu S_n. Mamy mianowicie (*dla dużych wartości* n)



Rysunek 2.3.1: Przybliżanie sumy niejednorodnych rozkładów dwumianowych rozkładami Poissona i dwumianowym: wartości dokładne (linia ciągła), aproksymacja Poissona (krzyżyki), aproksymacja dwumianowa (kółka)



Rysunek 2.3.2: Aproksymacja Poissonowska dla rozkładu dwumianowego $Bin(n, p_n)$: Rysunek lewy - $n = 20, p_n = 0.1$; rysunek prawy - $n = 20, p_n = 0.4$. Linia ciągła - wartości dokładne.

$$P(S_n \leqslant y) \approx \Phi\left(\frac{y - \mathrm{E}\left[S_n\right]}{\sqrt{\mathrm{Var}\left[S_n\right]}}\right),$$

przy czym $\mathrm{E}\left[S_n\right] = n\mu$ oraz $\mathrm{Var}\left[S_n\right] = n\sigma^2$. Tego rodzaju przybliżenie stosuje się również w przypadku, gdy (X_i) są niezależne, ale być może o różnych rozkładach. Wtedy kładziemy $\mathrm{E}\left[S\right] = \sum_{i=1}^{n} \mu_i$, $\mathrm{Var}\left[S\right] = \sum_{i=1}^{n} \sigma_i^2$, gdzie $\mu_i = \mathrm{E}\left[X_i\right]$, $\sigma_i^2 = \mathrm{Var}\left[X_i\right]$.

Przykład 2.3.4 (cd. Przykładu 2.1.4) Zastosujemy najpierw aproksymację normalną w przypadku, gdy n jest małe ($n = 3$). Jeżeli prawdziwy rozkład S jest typu dyskretnego wtedy poprawiamy aproksymację poprzez dodanie 0.5 do wartości S. Zatem dla $S = X_1 + X_2 + X_3$ mamy $\mathrm{E}\left[S\right] = 3.4, \mathrm{Var}\left[S\right] = 3.66$ oraz

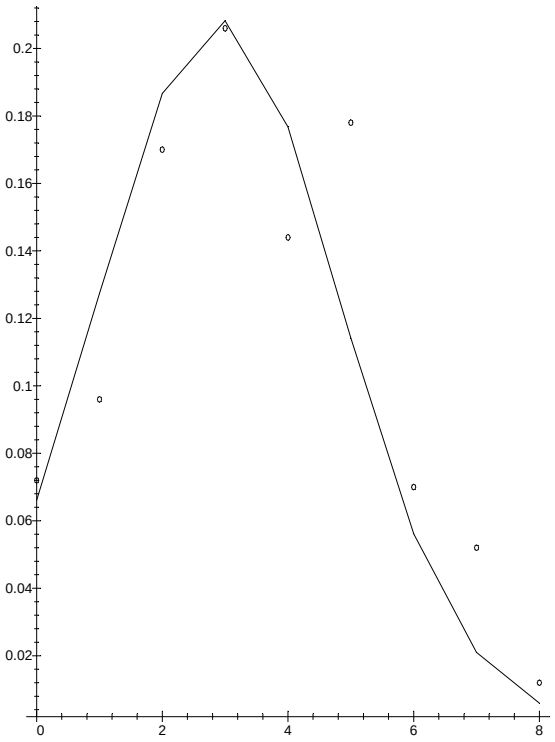
$$\begin{aligned} P(S \leqslant x) &= P\left(\frac{S + 0.5 - \mathrm{E}\left[S\right]}{\sqrt{\mathrm{Var}\left[S\right]}} \leqslant \frac{x + 0.5 - \mathrm{E}\left[S\right]}{\sqrt{\mathrm{Var}\left[S\right]}}\right) \\ &= P\left(\frac{S + 0.5 - 3.4}{\sqrt{3.66}} \leqslant \frac{x + 0.5 - 3.4}{\sqrt{3.66}}\right) \approx \Phi\left(\frac{x + 0.5 - 3.4}{\sqrt{3.66}}\right). \end{aligned}$$

Dla poszczególnych x odczytujemy rozkład z tablic. Poniżej w tabeli oraz na rysunku został przedstawiony rozkład zmiennej S obliczony przy pomocy wzoru dokładnego (lub równoważnie, funkcji tworzącej) i aproksymacji normalnej.

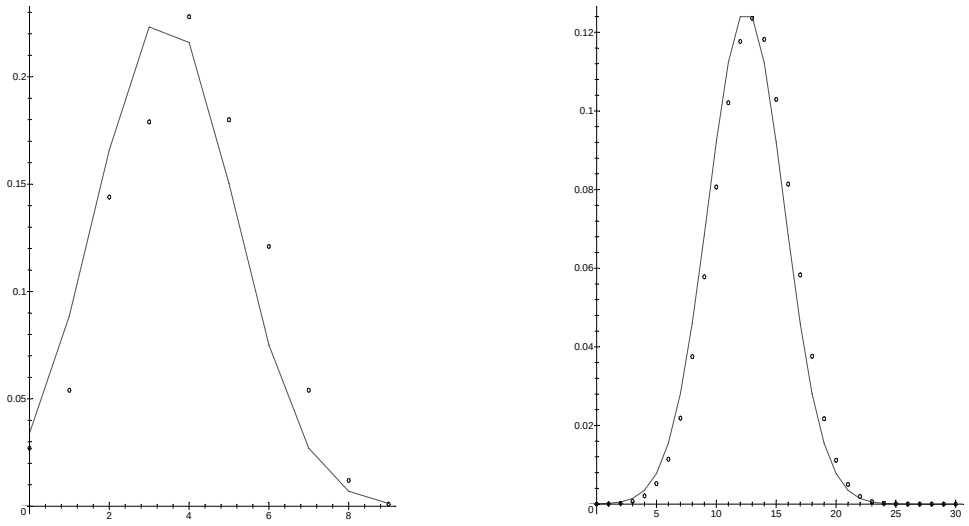
x	Dokładne wyliczenie $\mathrm{f}_S(\mathrm{x})$	$\mathrm{F}_S(\mathrm{x})$	Aproksymacja $\mathrm{f}_S(\mathrm{x})$	rozkł. normalnym $\Phi\left(\frac{x+0.5-3.4}{\sqrt{3.66}}\right)$
0	0.072	0.072	0.0661	0.0648
1	0.096	0.168	0.1273	0.1603
2	0.170	0.338	0.1867	0.3190
3	0.206	0.544	0.2082	0.5208
4	0.144	0.688	0.1768	0.7173
5	0.178	0.866	0.1142	0.8638
6	0.070	0.936	0.0561	0.9474
7	0.052	0.988	0.0210	0.9839
8	0.012	1.00	0.0059	0.9962

□

Obliczenia powyższe pokazują, że dla małego rozmiaru portfela przybliżenie rozkładem normalnym nie jest dobre.



Rysunek 2.3.3: Wartości dokladne (kółka) i aproksymacja normalna (linia ciągła) dla sumy trzech zmiennych losowych.



Rysunek 2.3.4: Przybliżenie rozkładem normalnym: wartości dokładne (kółka) i aproksymacja (linia ciągła) dla $n = 3$ i $n = 10$.

Przykład 2.3.5 Chcemy policzyć rozkład zmiennej $S_n = X_1 + \cdots + X_n$, gdzie $(X_i)_{i\geqslant 1}$ mają ten sam rozkład

i	0	1	2	3
$P(X = i)$	0.3	0.2	0.4	0.1

Funkcja tworząca prawdopodobieństwa ma postać $P_X(t) = 0.3 + 0.2t + 0.4t^2 + 0.1t^3$, a stąd $P_S(t) = (0.3+0.2t+0.4t^2+0.1t^3)^n$. Otrzymujemy więc szereg potęgowy i odczytując współczynniki przy t^m, $m \in \mathbb{N}$, dostajemy rozkład S. Analogicznie jak w poprzednim przykładzie stosujemy aproksymację normalną ($\mathrm{E}[S] = n \cdot 1.3$, $\mathrm{Var}[S] = n \cdot 1.01$) i dla $n = 3$ oraz $n = 10$ sporządzamy przybliżenia (rysunek)

□

Widzimy, że dla $n = 10$ mamy lepszą zgodność z rozkładem normalnym.

Przykład 2.3.6 Jezeli zmienna losowa X ma rozkład $Bin(n, p)$ oraz $n \to \infty$, to z Twier-

dzenia Moivre'a-Laplace'a mamy

$$\frac{X - np}{\sqrt{npq}} \to^{d}_{n\to\infty} Z \sim N(0,1).$$

Lepsze wyniki daje zastosowanie czynnika korygującego 0.5, tzn. przybliżamy

$$P\left(X \leqslant x\right) \approx \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right).$$

□

Przykład 2.3.7 Załóżmy, że zmienna losowa N ma rozkład $Poi(\lambda)$. Zauważmy, że jeśli λ jest całkowite, to N można przedstawić (co do rozkładu) jako $N_1 + \cdots + N_\lambda$, gdzie składniki są niezależnymi zmiennymi losowymi o rozkładzie $Poi(1)$. Centralne Twierdzenie Graniczne można więc stosować w ogólnym przypadku i otrzymać przybliżenie (*dla dużych* λ)

$$P(N \leqslant k) \approx \Phi\left(\frac{k - \lambda}{\sqrt{\lambda}}\right).$$

Przybliżenie to jest jednak niedokładne dla małych wartości λ. Intuicyjnie jest to jasne: rozkład normalny będący symetrycznym nie może dawać dobrych przybliżeń dla Poissona z małym λ, a więc z dużą skośnością $\frac{1}{\sqrt{\lambda}}$.

□

Przykład 2.3.8 Załóżmy, że $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi o rozkładzie wykładniczym ze średnią 1. Wtedy zmienna losowa $S_n = \sum_{i=1}^{n} X_i$ ma rozkład $\Gamma(n,1)$ (patrz Twierdzenie 2.1.11). Rozkład *standaryzowanej zmiennej losowej* $S_n^* = (S_n - n\mu_X)/(\sigma_X\sqrt{n})$, $k = 1, \ldots, n$, ma wtedy postać

$$\begin{aligned} F_{S_n^*}(u) &= P(S_n^* \leqslant u) = P\left(\frac{S_n - n\mu_X}{\sigma_X\sqrt{n}} \leqslant u\right) \\ &= P\left(S_n \leqslant \sigma_X\sqrt{n}u + n\mu_X\right) \\ &= F_{S_n}(\sigma_X\sqrt{n}u + n\mu_X), \end{aligned}$$

a stąd (uwzględniając $\mu_X = \sigma_X = 1$), $f_{S_n^*}(u) = \sqrt{n} f_{S_n}(\sqrt{n}u + n)$. Ponieważ dla rozkładu $\Gamma(n,1)$ mamy $f_{S_n}(x) = \frac{x^{n-1}\exp(-x)}{(n-1)!}$, więc kładąc $x = \sqrt{n}u + n$ dostajemy

$$f_{S_n^*}(u) = \sqrt{n}\frac{(\sqrt{n}u + n)^{n-1}\exp(-\sqrt{n}u)\exp(-n)}{(n-1)!}.$$

Zauważmy, że zmienne losowe S_n^* przyjmują wartości w przedziale $(-\sqrt{n}\mu_X/\sigma_X, \infty)$. Z Centralnego Twierdzenia Granicznego zmienna losowa S_n^* ma dla dużych n rozkład w przybliżeniu normalny. Na jednym rysunku przedstawimy wykresy gęstości zmiennej S_n^* dla różnych wartości n oraz gęstość rozkładu $N(0,1)$. Dobre dopasowanie mamy już od $n = 30$.

□

Przykład 2.3.9 Towarzystwo ubezpieczeniowe oferuje roczne kontrakty w wysokości 1 i 2, w grupach osób, gdzie w pierwszej grupie prawdopodobieństwo śmierci równa sie 0.02, a w drugiej 0.10:

k	q_k	b_k	n_k
1	0.02	1	500
2	0.02	2	500
3	0.10	1	300
4	0.10	2	500

gdzie n_k oznaczają liczebności konkraktów. Towarzystwo to zamierza za 1800 powyżej opisanych kontraktów zebrać tyle składek, aby z prawdopodobieństwem 0.95 sumaryczna szkoda $S = X_1 + \ldots + X_{1800}$ była mniejsza od zebranej sumy. Przy tym wymaga się, aby składka każdego osobnika była proporcjonalna do wartości oczekiwanej jego szkody, tzn. ma być postaci $(1+\theta)E[X_j]$, dla pewnej stałej $\theta > 0$ (składka wartości oczekiwanej). Zmienne X_i są niezależne o rozkładach $P(X_i = b_k) = q_k$, $P(X_i = 0) = 1 - q_k$, przy czym rozkłady zależą od typu kontraktu (np. dla $i = 1, \ldots, 500$, $k = 1$). Należy więc znaleźć θ takie, że

$$P(S \leqslant (1+\theta)\mathrm{E}\,[S]) = 0.95,$$

co jest równoważne

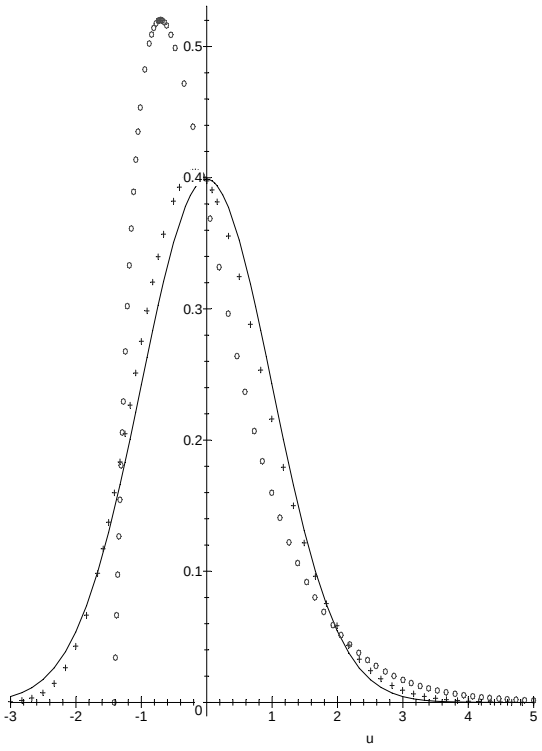
$$P\left(\frac{S - \mathrm{E}\,[S]}{\sqrt{\mathrm{Var}\,[S]}} \leqslant \frac{\theta \mathrm{E}\,[S]}{\sqrt{\mathrm{Var}\,[S]}}\right) = 0.95,$$

i korzystając z przybliżenia rozkładem normalnym

$$\Phi\left(\frac{\theta \mathrm{E}\,[S]}{\sqrt{\mathrm{Var}\,[S]}}\right) = 0.95.$$

Z tablic odczytujemy

$$\frac{\theta \mathrm{E}\,[S]}{\sqrt{\mathrm{Var}\,[S]}} = 1.645.$$



Rysunek 2.3.5: Aproksymacja rozkładu Gamma rozkładem normalnym. Rozkład normalny (linia ciągła) i przesunięte rozkłady Gamma:
$k = 2$ (kółka) i $k = 30$ (krzyżyki).

Pozostaje więc policzyć $\mathrm{E}[S]$ i $\mathrm{Var}[S]$.

k	q_k	b_k	$E[X_k]$	$Var[X_k]$	n_k	$n_k E[X_k]$
1	0.02	1	0.02	0.0196	500	10
2	0.02	2	0.04	0.0784	500	20
3	0.1	1	0.1	0.09	300	30
4	0.1	2	0.2	0.36	500	100

Sumując odpowiednio otrzymujemy $\mathrm{E}[S] = 160$ oraz $\mathrm{Var}[S] = 256$. Wstawiając te wartości wyżej otrzymujemy $\theta = 0.1645$ (czyli narzut na skladkę netto powinien wynosić 16.5%, aby zapewnić 95% pewność, że składki pokryją szkody).

□

Przykład 2.3.10 Portfel $X_1, \ldots, X_n$ jest dwuwarstwowy, dla $n = n_1 + n_2$ ryzyk postaci $X_i = I_i \cdot B_i$, gdzie $P(I_i = 1) = q_i$. Inaczej mówiąc, ubezpieczenia wypadkowe są dwojakiego rodzaju ($k = 1, 2$) i mają taka samą strukturę. Rozkład wypłaty (pod warunkiem, że szkoda następuje) tzn. $X_k \mid I_k = 1$ jest rozkładem obciętym wykładniczym. Rozkład obcięty wykładniczy zadany jest dystrybuantą

$$F_B(x) = \begin{cases} 0 & \text{dla} \quad x < 0 \\ 1 - e^{-\lambda x} & \text{dla} \quad 0 \leqslant x < L \\ 1 & \text{dla} \quad x \geqslant L \end{cases},$$

dla pewnego $L > 0$. Specyfikując różne wartości parametrów skali i poziomu obcięcia, zakładamy następujące dane

k	n_k	q_k	λ_k	L_k
1	500	0.10	1	2.5
2	2000	0.05	2	5.0

Wyliczając momenty otrzymujemy

$$\mathrm{E}[X_k | I_k = 1] = \frac{1 - e^{-\lambda_k L_k}}{\lambda_k},$$

$$\mathrm{E}\left[X_k^2 | I_k = 1\right] = \frac{2}{\lambda_k^2}(1 - e^{-\lambda_k L_k}) - \frac{2L_k}{\lambda_k} e^{-\lambda_k L_k},$$

$$\mathrm{Var}[X_k | I_k = 1] = \frac{1 - 2\lambda_k L_k e^{-\lambda_k L_k} - e^{-2\lambda_k L_k}}{\lambda_k^2}.$$

Zbierając wyniki dla poszczególnych typów kontraktów mamy

k	q_k	μ_k	σ_k^2	$E[X_k]$	$Var[X_k]$	n_k
1	0.10	0.9179	0.5828	0.09179	0.13411	500
2	0.05	0.5000	0.2498	0.0250	0.02436	2000

Ostatecznie

$$\mathrm{E}[S] = 500 \cdot 0.09179 + 2000 \cdot 0.025 = 95.89,$$

$$\mathrm{Var}[S] = 500 \cdot 0.13411 + 2000 \cdot 0.02436 = 115.78$$

i szukamy θ takiego, że

$$P(S \leqslant (1+\theta)\mathrm{E}[S]) = 0.95$$

Stosując tablice rozkładu normalnego i przybliżenie rozkładem normalnym, znajdujemy dla $S = X_1 + \cdots + X_n$,

$$\frac{\theta \mathrm{E}[S]}{\sqrt{\mathrm{Var}[S]}} = 1.645$$

Stąd $\theta = 0.1846$, to znaczy, że należy przyjąć *narzut* ponad osiemnastoprocentowy na wartość średnią, aby z prawdopodobieństwem 0.95 składki (z narzutem) pobierane jako wartość średnia z narzutem pokryły wartość zgłoszonych szkód.

□

2.3.3 Aproksymacja rozkładów złożonych rozkładem normalnym

Centralne Twierdzenie Graniczne jest prawdziwe również dla sum losowych.

Twierdzenie 2.3.11 *Jeśli $S = X_1 + \ldots + X_N$ ma rozkład złożony Poissona $CPoi(\lambda, F)$, to*

$$\frac{S - \lambda \mathrm{E}[X]}{\sqrt{\lambda \mathrm{E}[X^2]}} \to^d_{\lambda \to \infty} N(0,1)$$

Dowód: Niech

$$Y = \frac{S - \lambda \mathrm{E}[X]}{\sqrt{\lambda \mathrm{E}[X^2]}}. \tag{2.3.1}$$

Pokażemy, że

$$\lim_{\lambda \to \infty} M_Y(t) = \exp(t^2/2), \qquad t \in \mathbb{R}.$$

Mamy

$$\begin{aligned} M_Y(t) = \mathrm{E}\left[e^{tY}\right] &= \mathrm{E}\left[\exp\left(\frac{S}{\sqrt{\lambda \mathrm{E}[X^2]}} t\right)\right] \exp\left(-\frac{\lambda t \mathrm{E}[X]}{\sqrt{\lambda \mathrm{E}[X^2]}}\right) \\ &= M_S\left(\frac{t}{\sqrt{\lambda E(X^2)}}\right) \exp\left(-\frac{\lambda t \mathrm{E}[X]}{\sqrt{\lambda \mathrm{E}[X^2]}}\right). \end{aligned}$$

Używając wzoru (2.2.4) otrzymujemy

$$M_Y(t) = \exp\left(\lambda\left[M_X\left(\frac{t}{\sqrt{\lambda \mathrm{E}\left[X^2\right]}}\right) - 1\right] - \frac{\lambda t \mathrm{E}\left[X\right]}{\sqrt{\lambda \mathrm{E}\left[X^2\right]}}\right).$$

Korzystając z rozwinięcia funkcji tworzącej momenty w szereg Taylora

$$M_X(t) = 1 + \frac{t\mathrm{E}\left[X\right]}{1!} + \frac{t^2\mathrm{E}\left[X^2\right]}{2!} + \ldots$$

oraz podstawiając $t/\sqrt{\lambda \mathrm{E}\left[X\right]}$ w miejsce t otrzymujemy

$$M_Y(t) = \exp\left(\frac{1}{2}t^2 + \frac{1}{6\sqrt{\lambda}}\frac{\mathrm{E}\left[X^3\right]}{\mathrm{E}\left[X^{3/2}\right]}t^3 + \ldots\right) \underset{\lambda\to\infty}{\to} \exp(t^2/2).$$

□

Przykład 2.3.12 Rozważmy złożony rozkład Poissona dla parametrów $\lambda = 5$ i $\lambda = 15$. Szkoda przyjmuje pięć wartości jak w Przykładzie 2.2.10. Dla parametru $\lambda = 15$ mamy dobrą zgodność z rozkładem normalnym.

□

Podobnie jak dla rozkładów $CPoi$ zachodzi Centralne Twierdzenie Graniczne dla rozkładów złożonych ujemnych dwumianowych.

Twierdzenie 2.3.13 *Jeśli $S = X_1 + \ldots + X_N$ ma rozkład złożony ujemny dwumianowy $Bin^-(r,p)$, to*

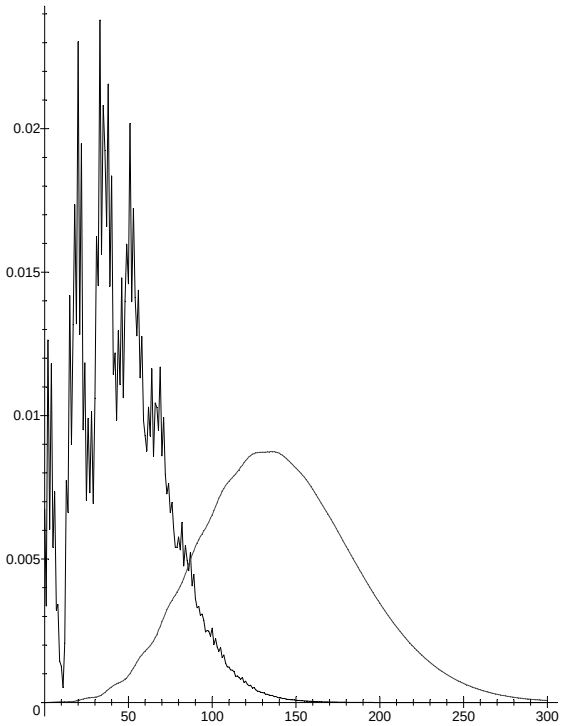
$$\frac{S - rq\mathrm{E}\left[X\right]/p}{\sqrt{(rq\mathrm{Var}\left[X\right])/p + rq(\mathrm{E}\left[X\right])^2/p^2}} \to^d_{r\to\infty} N(0,1).$$

2.3.4 Aproksymacja przesuniętym rozkładem Gamma

Niech teraz

$$G(x;\alpha,\beta) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1}e^{-\beta t}dt,$$

$$H(x;\alpha,\beta,x_0) = G(x - x_0;\alpha,\beta).$$



Rysunek 2.3.6:

Funkcja $G(\cdot;\alpha,\beta)$ jest dystrybuantą rozkładu gamma $Gamma(\alpha,\beta)$ o parametrze kształtu α i parametrze skali β, .

Sumaryczną szkodę S można przybliżyć rozkładem gamma zachowując zgodność trzech pierwszych centralnych momentów, tzn.

$$\mathrm{E}[S] = x_0 + \frac{\alpha}{\beta},$$

$$\mathrm{Var}[S] = \frac{\alpha}{\beta^2},$$

$$\mathrm{E}\left[(S - \mathrm{E}[S])^3\right] = \frac{2\alpha}{\beta^3}.$$

Zauważmy, że aproksymacja gamma wymaga, by skośność była dodatnia.

Stąd należy dobrać parametry x_0, α, β w następujący sposób

$$\alpha = 4\frac{(\mathrm{Var}[S])^3}{(\mathrm{E}[(S - \mathrm{E}[S])^3])^2},$$

$$\beta = 2\frac{\mathrm{Var}[S]}{\mathrm{E}[(S - \mathrm{E}[S])^3]},$$

$$x_0 = \mathrm{E}[S] - 2\frac{(\mathrm{Var}[S])^2}{\mathrm{E}[(S - \mathrm{E}[S])^3]}.$$

W szczególności dla rozkładu złożonego Poissona

$$\alpha = 4\lambda\frac{(\mathrm{E}[X^2])^3}{(\mathrm{E}[X^3])^2}, \tag{2.3.2}$$

$$\beta = 2\frac{\mathrm{E}[X^2]}{\mathrm{E}[X^3]}, \tag{2.3.3}$$

$$x_0 = \lambda\mathrm{E}[X] - 2\lambda\frac{(\mathrm{E}[X^2])^2}{\mathrm{E}[X^3]}. \tag{2.3.4}$$

Uwaga 2.3.14 Powstaje naturalne pytanie: kiedy stosować aproksymację normalną, a kiedy Gamma? Jeżeli S współczynnik skośności jest mały, to należy stosować aproksymację normalną, w przeciwnym razie - Gamma. Regułę tę należy stosować ostrożnie.

Przykład 2.3.15 (cd. Przykładu 2.2.10) Rozważmy złożony rozkład Poissona z parametrem $\lambda = 10$ i szkodą o rozkładzie

i	1	2	12	13	18
$P(X = i)$	0.1	0.35	0.05	0.2	0.3

Zastosujemy najpierw aproksymację normalną. W złożonym rozkładzie mamy

$$\begin{aligned}\mathrm{E}[S] &= \mathrm{E}[N]\,\mathrm{E}[X] = 94,\\ \mathrm{Var}[S] &= \lambda(\mathrm{E}\left[X^2\right]) = 1397.\end{aligned}$$

Otrzymujemy więc

$$\Pr(S \leqslant x) \approx \Phi(\frac{x - 94 + 0.5}{\sqrt{1397}}).$$

Aby skorzystać przybliżenia Gamma trzeba dla danych w powyższym przykładzie obliczyć parametry x_0, α, β. Dla złożonego rozkładu Poissona mamy ze wzorów (2.3.2)-(2.3.4)

$$\begin{aligned}\alpha &= 4 \cdot 10 \cdot \frac{(139.7)^3}{(2278.3)^2} = 21.01,\\ \beta &= 2 \cdot \frac{139.7}{2278.3} = 0.12,\\ x_0 &= 10 \cdot 9.4 - 2 \cdot 10 \cdot \frac{(139.7)^2}{2278.3} = -77.3.\end{aligned}$$

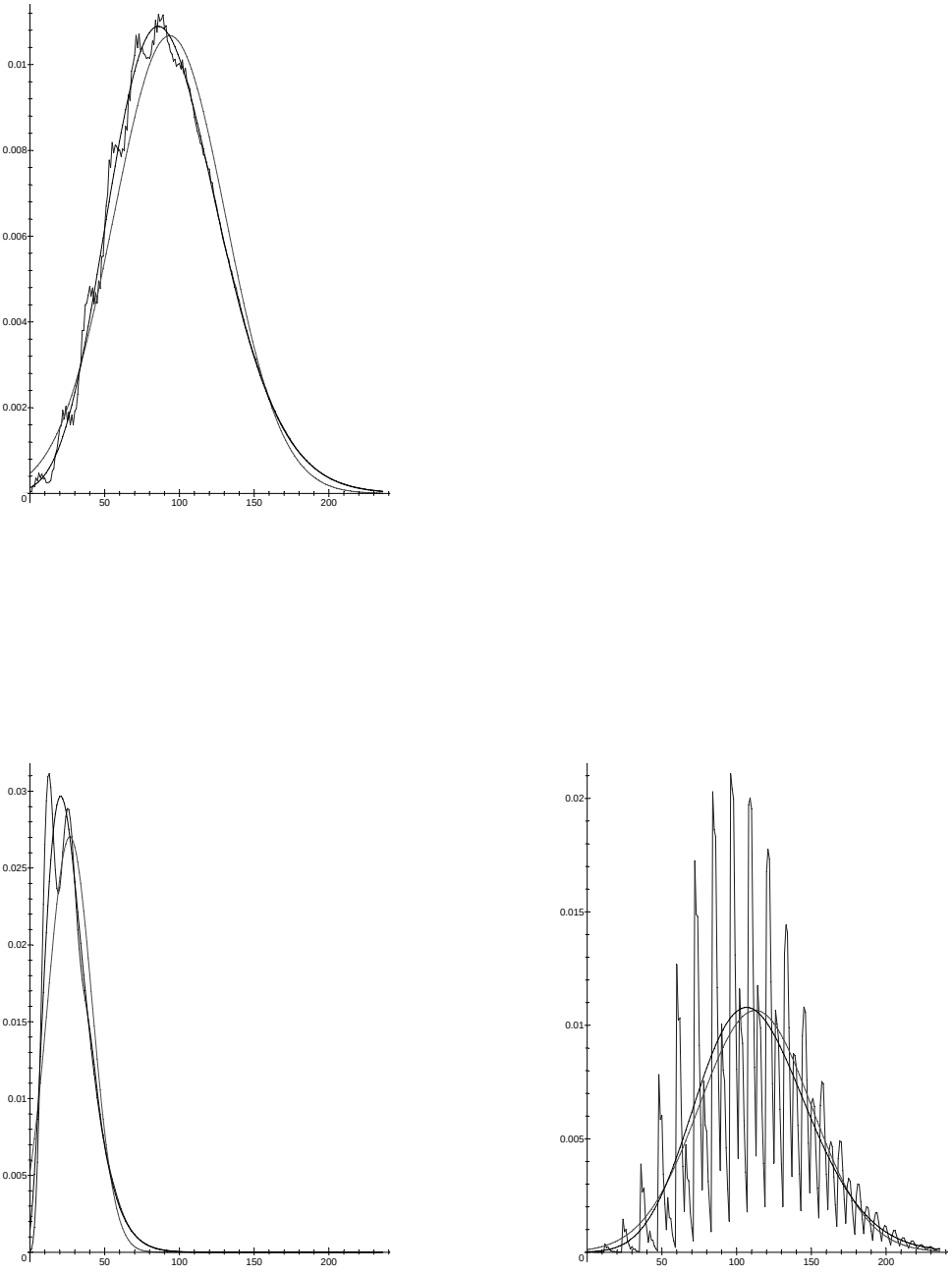
W naszym przypadku skośność wynosi 0.44. Widzimy lepsze dopasowanie rozkładem Gamma niż normalnym.

Rozpatrzmy jednak inne dane:

i	1	2	12	13	18	γ	ξ
$P(X=i)$	0.45	0.45	0.05	0.025	0.025	0.9	0.3
$P(X=i)$	0.5	0.5	0	0	0	0.36	0.06
$P(X=i)$	0.05	0.05	0.8	0.05	0.05	0.34	0.055

W pierwszym przypadku aproksymacja normalna, ze względu na dużą skośność (0.44) nie jest dopuszczalna, przybliżenie Gamma jest dobre tylko w ogonie rozkładu. W drugim , skośność wynosi 0.9, w trzecim - 0.34.

□



Rysunek 2.3.7: Aproksymacja Gamma (linia gruba) i normalna (linia cienka) wraz z wartościami dokładnymi (linia łamana)

Rozdział 3

Składki

3.1 Składka netto

Składka jest opłatą, którą podmiot narażony na ryzyko płaci ubezpieczycielowi za przejęcie części ryzyka związanego ze swoją działalnością. Składkę wylicza ubezpieczyciel bazując na zasadzie, że losowe, sumaryczne roszczenia ubezpieczonych klientów w danej grupie działalności (w danym **portfelu**) powinny być skompensowane przez ustalone z góry opłaty (składki).

Wartość składki H powinna zależeć od dystrybuanty F_S zmiennej S, która oznacza ubezpieczaną wielkość. Tak więc $H = H(F_S)$ (piszemy również $H = H(S)$).

Składka netto. Podstawową składką jest $H = E[S]$. Składka taka nazywana jest składką netto. Wycena szkody jedynie poprzez wartość oczekiwaną nie gwarantuje jednak bezpiecznego pokrycia wielkości szkody.

Rzeczywiście, załóżmy bowiem, że portfel składa się z n niezależnych, jednakowo rozłożonych ryzyk $X_1, ..., X_n$. Mamy $S_n = X_1 + ... + X_n$. Oczywiście, składka dla całego portfela powinna być równa sumie pojedynczych składek, tzn. $H(S_n) = H(X_1) + ... + H(X_n) = nH(X_1)$, gdyż ryzyka mają ten sam rozkład. Z Prawa Wielkich Liczb, dla każdego $\varepsilon > 0$ i dostatecznie dużych wartości n, $P\left(\left|\frac{S_n}{n} - \mathrm{E}\left[X_1\right]\right| > \varepsilon\right) = 0$, a stąd otrzymujemy

$$\lim_{n\to\infty} P\left(\sum_{i=1}^{n} X_i > nH(X_1)\right) = \begin{cases} 1 & \text{dla} \quad H(X_1) < \mathrm{E}\left[X_1\right] \\ 0 & \text{dla} \quad H(X_1) \geqslant \mathrm{E}\left[X_1\right] \end{cases} . \tag{3.1.1}$$

W przypadku, gdy $H(X_1) < \mathrm{E}\left[X_1\right]$ prawdopodobienstwo, że suma wypłaconych odszkodowań przekroczy zebrane składki dąży do 1. Dlatego dla dowolnego portfela S pożądaną własnością składki jest by

$$H(S) > \mathrm{E}\left[S\right],$$

czyli

$$H(S) = (1+\rho)\mathrm{E}\left[S\right] \quad \rho > 0$$

Wartość ρ nazywamy **narzutem bezpieczeństwa**. Wyliczenie wartości ρ w konkretnym portfelu jest związane z zagadnieniem bezpieczeństwa pokrycia szkody przez składkę.

3.2 Składka z ustalonym poziomem bezpieczeństwa

Równość (3.1.1) wskazuje, że dla dużego rozmiaru portfela i składki przewyższającej $\mathrm{E}[X_1]$, prawdopodobieństwo niewypłacalności dąży do zera. W praktyce jednak mamy do czynienia z portfelami skończonego rozmiaru, dlatego stawiamy następujące pytanie: jak duża powinna być składka (lub jak duży powinien być narzut bezpieczeństwa), aby prawdopodobieństwo niewypłacalności było równe z góry zadanej (małej) liczbie $\alpha \in (0,1)$? Inaczej mówiąc: szukamy takiego $H(S)$, by

$$P(S > H(S)) = \alpha.$$

Wartość ta nazywana jest kwantylem rzędu $1-\alpha$ rozkładu zmiennej S (wartość H(S) jest nazywana *Value-at-Risk* (VaR) i jest jedną ze standardowych miar oceny wypłacalności instytucji finansowych). W przypadku $S_n = X_1 + ... + X_n$, gdy nie znamy rozkładu pojedynczej szkody X_i, a rozmiar portfela jest dostatecznie duży ($n \geqslant 30$) możemy posłużyć się Centralnym Twierdzeniem Granicznym, aby otrzymać

$$\begin{aligned} P(S_n > H(S_n)) &= P\left(\frac{S_n - \mathrm{E}[S_n]}{\sigma(S_n)} > \frac{H(S_n) - \mathrm{E}[S_n]}{\sigma(S_n)}\right) \\ &\approx P\left(Z > \frac{H(S_n) - \mathrm{E}[S_n]}{\sigma(S_n)}\right), \end{aligned}$$

gdzie zmienna losowa Z ma rozkład normalny $N(0,1)$ ze średnią 0 i wariancją 1. Przyjmując $P\left(Z > \frac{H(S_n)-\mathrm{E}[S_n]}{\sigma(S_n)}\right) = \alpha$ otrzymujemy

$$\frac{H(S_n) - \mathrm{E}[S_n]}{\sigma(S_n)} = z_{1-\alpha},$$

gdzie $z_\alpha = F_Z^{-1}(\alpha)$, $\alpha \in (0,1)$. Mamy więc $H(S_n) = (1+\rho)\mathrm{E}[S_n]$ z

$$\rho = \frac{z_{1-\alpha}\sigma(S_n)}{\mathrm{E}[S_n]}.$$

Powyższy wzór można zapisać inaczej w postaci

$$H(S_n) = \mathrm{E}[S_n] + z_{1-\alpha}\sigma(S_n) \tag{3.2.1}$$

lub ogólniej

$$H(S) = \mathrm{E}[S] + \delta\sigma(S), \qquad \delta > 0.$$

Powyższy wzór definiuje tak zwaną **składkę odchylenia standardowego**.

Powyżej podaliśmy sposób wyliczenia składki dla portfela S_n złożonego z niezależnych ryzyk $X_1, ..., X_n$ o tym samym rozkładzie. Z formuły na składkę $H(S_n)$ dla całego portfela łatwo można uzyskać składkę dla pojedynczego ryzyka:

$$H(X_i) = \mathrm{E}[X_i] + z_{1-\alpha}\sigma(S_n)/n, \qquad i = 1, ..., n. \tag{3.2.2}$$

Zauważmy, że składka ta jest identyczna dla każdego i. Ogólniej, jeżeli $X_1, ..., X_n$ mają inne rozkłady, ale są niezależne i mają równe wariancje, to składkę indywidualną można skalkulować według wzoru (3.2.2).

Często postuluje się, by dla dowolnych dwóch ryzyk S i S^* zachodziła nierówność

$$H(S + S^*) \leqslant H(S) + H(S^*).$$

Powyższa nierówność może być ostra jak pokazuje poniższy przykład.

Przykład 3.2.1 [EA:26.10.1998(8)] Łączenie ryzyk dwóch różnych portfeli może mieć wpływ na wielkość składki. Rozważmy w tym celu dwie grupy ryzyk: pierwsza składa się z 900 niezależnych ryzyk $X_1, ...X_{900}$ z tą samą średnią 1 i wariancją 4, druga składa się z 800 niezależnych (nawzajem i od ryzyk z pierwszej grupy) ryzyk $Y_1, ..., Y_{800}$ ze średnią 2 i wariancją 8. W obu grupach ustalono składki netto z narzutem H_1 i H_2 tak, by prawdopodobieństwo, iż suma szkód przekroczy sumę składek wyniosło 0.0013. W tym celu przyjęto, że ryzyka łączne w obu portfelach można przybliżyć rozkładem normalnym. Powstaje pytanie o ustalenie nowych składek z narzutem H_1^* i H_2^* tak, by po połączeniu portfeli został zachowany standard bezpieczeństwa stosowany w pojedynczych portfelach. Jaki jest dopuszczalny zbiór wartości dla H_1^* i H_2^*?

Ustalmy najpierw początkową wielkość składek. Z warunków zadania

$$P\left(\sum_{i=1}^{900} X_i > 900H_1\right) = 0.0013$$

oraz

$$P\left(\sum_{i=1}^{800} Y_i > 800H_2\right) = 0.0013$$

i z aproksymacji normalnej dostajemy $(900H_1 - 900)/\sqrt{900 \cdot 4} = z_{1-0.0013} \approx 3$ oraz $(800H_2 - 1600)/\sqrt{800 \cdot 8} = z_{1-0.0013} \approx 3$, gdzie $z_{1-0.0013}$ jest kwantylem rzędu $1 - 0.0013$ standardowego rozkładu normalnego. Stąd $H_1 \approx \frac{6}{5}$, $H_2 \approx 2.3$. Chcemy teraz znaleźć takie H_1^* i H_2^*, by

$$P\left(\sum_{i=1}^{900} X_i + \sum_{i=1}^{800} Y_i > 900H_1^* + 800H_2^*\right) = 0.0013,$$

czyli $(900H_1^* + 800H_2^* - 1600 - 900)/\sqrt{900 \cdot 4 + 800 \cdot 8} \approx 3$ oraz by zmiana stawek była korzystna, tzn. $H_1^* \leqslant H_1$, $H_2^* \leqslant H_2$. Rozwiązujemy więc, dla składek z narzutem, zagadnienie

$$\begin{cases} 900H_1^* + 800H_2^* \approx 2800 \\ H_1^* \leqslant H_1 \\ H_2^* \leqslant H_2 \end{cases}.$$

Otrzymujemy stąd dopuszczalny przedział dla H_2^* : $(2.15, 2.30)$. Zauważmy, że przed zmianą taryfy łączna zebrana składka wyniosła $900H_1 + 800H_2 = 2920.0$, a więc była wyższa niż po zmianie stawek. Łączenie niezależnych portfeli ryzyk może obniżyć wielkośc składek.

□

3.3 Składki oparte o funkcję użyteczności

Funkcją użyteczności będziemy nazywać dowolną niemalejącą funkcję, która mierzy użyteczność kwoty (losowej) X. Jeśli kwota X przyjmuje wartości w przedziale $[m, M]$, to wygodnie jest unormować taką funkcję, przyjmując $u(m) = 0$ i $u(M) = 1$.

Jeżeli kontrahent posiada kapitał w i posługuje się funkcją użyteczności u to wysokość składki, którą gotowy jest zapłacić za zabezpieczenie ryzyka zależy od funkcji u. Graniczną wysokością składki jest kwota H wyznaczona równaniem balansującym oczekiwany efekt ryzyka mierzony wartością oczekiwaną względem u z efektem deterministycznym zabezpieczenia po opłaceniu H,

$$\mathrm{E}\,[u(w - S)] = u(w - H)$$

Przyjęcie warunku $w = H$ w powyższym wzorze prowadzi do równania

$$\mathrm{E}\,[u(H - S)] = u(0).$$

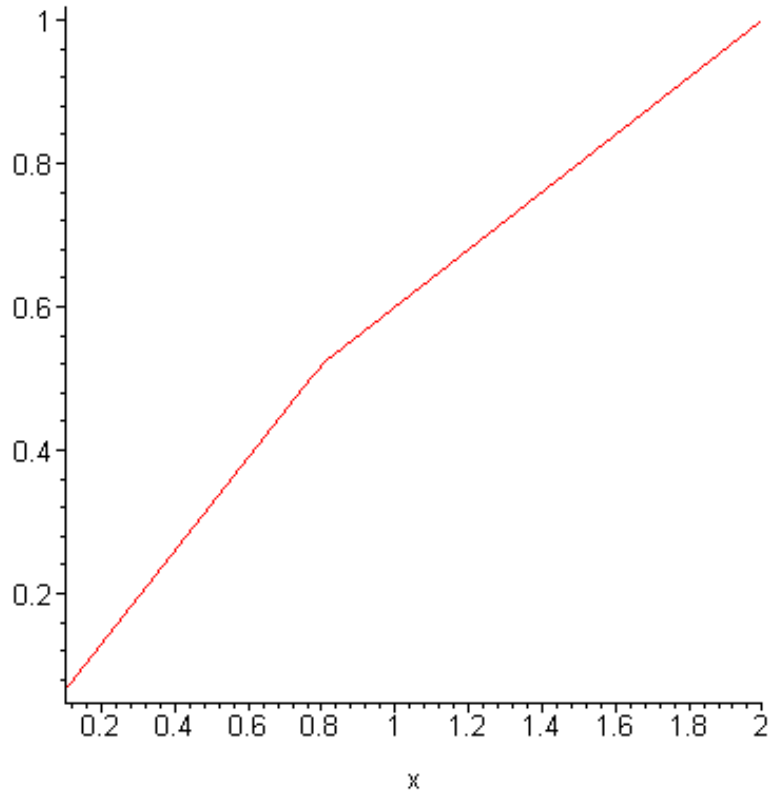
Wyliczona z tego równania składka H nazywana jest **składką funkcji użyteczności** (*zero utility principle*).

Rozpoznanie kształtu funkcji użyteczności jest możliwe w wyniku analizy potencjalnych decyzji dotyczących zabezpieczania się i gotowości płacenia składki.

Przykład 3.3.1 Załóżmy, że $M = w = 20000$ oraz $P(X = 0) = 1/2 = P(X = M)$. To znaczy posiadając kapitał w wysokości 20000 jesteśmy narażeni na ryzyko X polegające na stracie całego kapitału z prawdopodobieństwem $1/2$. Określając teraz maksymalną składkę, powiedzmy $H = 12000$, za ubezpieczenie takiego ryzyka, automatycznie wyznaczamy wartość $u(8000) = 1/2$, bo $E(u(M - X)) = 1/2$ (zakładając, że wyznaczamy unormowane u). Zob. rys. 3.3.1

□

Intuicyjne zrozumienie roli funkcji użyteczności możliwe jest również poprzez interpretację w języku prowadzenia gry.



Rysunek 3.3.1: Pierwsze przybliżenie funkcji użyteczności. Oś OX razy 10000.

Przykład 3.3.2 Załóżmy, że dla dowolnego ustalonego w z przedziału $[m, M]$, i funkcji użyteczności u takiej, że $u(m) = 0$, $u(M) = 1$ mamy następujący wybór: akceptujemy kwotę w lub otrzymujemy kwotę wynikajacą z losowania, gdzie otrzymujemy M z prawdopodobieństwem p (sukces) lub m z prawdopodobieństwem $1 - p$ (porażka). Każdy się zgodzi, że wyboru nie możemy dokonać nie znając wartości p. Oznaczmy przez X zmienną losową taką, że $P(X = M) = p = 1 - P(X = m)$. Wtedy $Eu(X) = p$. Jesli p^* jest tą wartością prawdopodobieństwa sukcesu, która spowoduje, że będziemy chcieli otrzymać kwotę wynikającą z losowania X zamiast pewnej kwoty w, to możemy przyjąć, że $u(w) = p^*$. Kształt funkcji u będzie zależał od indywidualnej skłonności do podjęcia ryzyka (wybrania opcji losowania).

□

Mówimy, że decydent ma *awersję do ryzyka* jeśli jego funkcja użyteczności jest wklęsła. Przypomnijmy, że u jest **wklęsła** jeśli

$$u(w + h) - u(w) \geqslant u(w' + h) - u(w'),$$

dla każdego $h > 0$ oraz $w \leqslant w'$, tzn. gdy u ma przyrosty malejące. Dla gładkich u, jest to równoważne z warunkiem, że pochodna u jest malejąca (słabo). Funkcja u o przyrostach rosnących (równoważnie o pochodnej rosnącej) nazywana jest funkcją **wypukłą**. Własności wklęsłości i wypukłości odnoszą się zawsze do konkretnego zbioru argumentów (zwykle rozważa się przedziały otwarte, niekoniecznie skończone).

Ponieważ dla funkcji wklęsłej u styczna w dowolnym ustalonym punkcie w_0 leży ponad wykresem u, zob. rys. 3.3.2, to

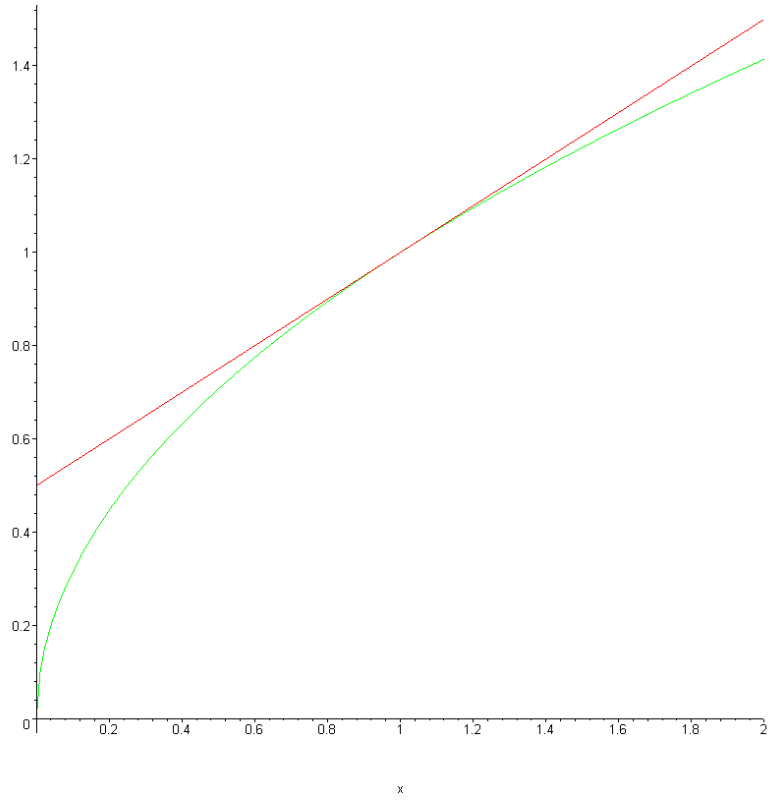
$$u(w) \leqslant u'(w_0)(w - w_0) + u(w_0).$$

Dla dowolnej zmiennej losowej X o wartości oczekiwanej EX, podstawiając $w_0 := EX$, $w := X$ i biorąc wartość oczekiwaną obustronnie w powyższej nierówności otrzymujemy natychmiast

$$\mathrm{E}\,[u(X)] \leqslant u(\mathrm{E}\,[X]).$$

Nierówność ta jest tak zwaną **nierównością Jensena** dla funkcji wklęsłych (dla u wypukłych zachodzi nierówność odwrotna).

Podstawiając w nierówności Jensena dla u wklęsłej, $X := w - S$, otrzymujemy $\mathrm{E}\,[u(w - S)] \leqslant u(w - \mathrm{E}\,[S])$. Jeśli teraz H jest taka, że spełnia równowagę $\mathrm{E}\,[u(w - S)] = u(w - H)$, to $u(w - H) \leqslant u(w - \mathrm{E}\,[S])$, z czego wynika ($u$ jest niemalejąca), że $H \geqslant \mathrm{E}\,[S]$. Oznacza to, że decydent mający awersję do ryzyka jest gotowy płacić składkę większą od wartości średniej ryzyka. Wielkość tej składki zależy jednak wyraźnie od rodzaju u i może zależeć od kapitału w, który posiada decydent .



Rysunek 3.3.2: Ilustracja nierówności Jensena. Linia prosta styczna w punkcie (1,1) leży nad wykresem krzywej wklesłej.

Przykład 3.3.3 Niech $u(w) = -e^{-\alpha w}$, $\alpha > 0$. Nie przyjmując dokładniejszych założeń o X, jedynie istnienie skończonej wartości $M_X(\alpha)$, z równania $\mathrm{E}\,[u(w-S)] = u(w-H)$ otrzymujemy natychmiast

$$H = C_X(\alpha)/\alpha.$$

Przy takim u, składka wyliczona na podstawie u nie zależy od w!

□

Przykład 3.3.4 Niech $u(w) = w^\gamma$, $\gamma \in (0,1)$. Jeśli $w = 20$, S ma rozkład jednostajny na $[0,20]$ i $\gamma = 1/2$, to $\mathrm{E}\,[u(w-S)] = u(w-H)$ sprowadza się do równości $\sqrt{20-H} = \int \sqrt{20-x} f_S(x)dx$ dla $f_S(x) = I_{[0,20]}(x)/20$. Wyliczając składkę otrzymujemy $H = 11,11...$ Widzimy, że $H > ES = 10$.

□

Przykład 3.3.5 Niech $u(w) = w - \alpha w^2$, $w < 1/2\alpha$, $\alpha > 0$. Załóżmy, że $w = 10$, $P(S = 10) = 0.5 = 1 - P(S = 0)$, $\alpha = 0.01$, wtedy $H = 5.28$. Gdy $w = 20$, to dla tej zmiennej S, $H = 5.37$, co pokazuje, że dla kwadratowej funkcji użyteczności składka zależy od posiadanego kaptału w.

□

3.4 Reasekuracja, podział ryzyka

Reasekuracja jest przekazaniem przez ubezpieczyciela (**cedenta**) części ryzyka do ubezpieczenia w innym towarzystwie ubezpieczeniowym. Ogólnie biorąc dla ryzyka X, wartość $X_r = X - h(X)$ przekazuje się do ponownego ubezpieczenia, gdzie $h(x)$ jest pewną funkcją (funkcja retencyjna) o własnościach: $h(x)$ jest niemalejąca, $x - h(x)$ jest niemalejąca, $0 \leqslant h(x) \leqslant x$ oraz $h(0) = 0$. Funkcja kompensacji $k(x)$ wyznaczona jest przez równość $k(x) = x - h(x)$. Udział własny oznaczamy przez $X_c = h(X)$.

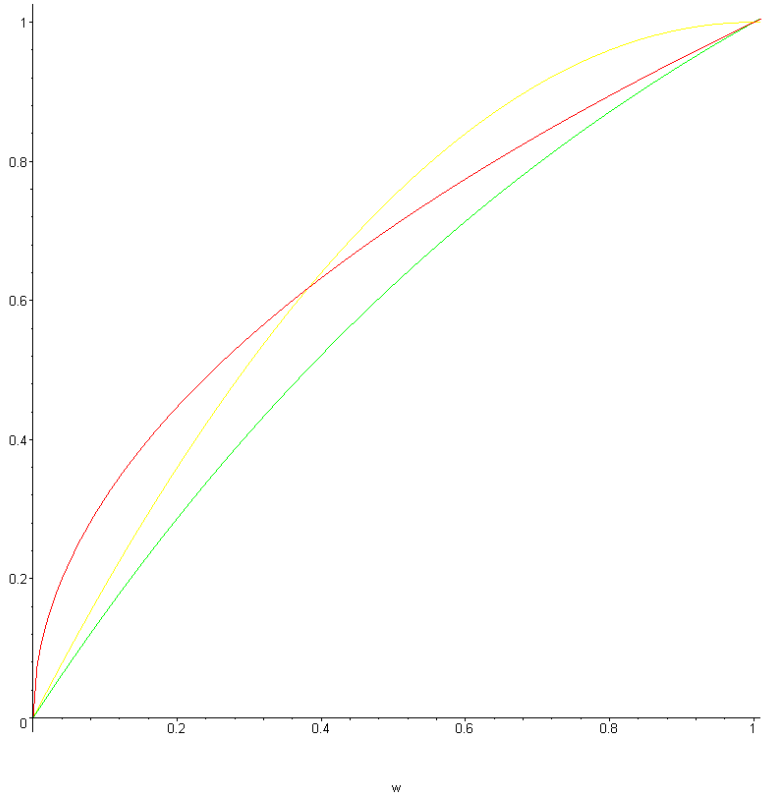
Typowymi kontraktami reasekuracyjnymi są:

1. Reasekuracja proporcjonalna

   a) reasekuracja z udziałem procentowym (*quota-share*),

   $$\begin{aligned} X_r &= \alpha X, \quad \alpha \in (0,1), \\ X_c &= (1-\alpha)X. \end{aligned}$$

   Wady kontraktu proporcjonalnego:



Rysunek 3.3.3: Wykres unormowanych funkcji użyteczności $(-exp(-w)+1)/(-exp(-1)+1)$- zielona, $\sqrt{(w)}$- czerwona, $(w-0.5*w^2)/0.5$- żółta

- małe szkody dzielone są pomiędzy cedenta i reasekuratora. Dominują wtedy koszty administracyjne likwidacji szkody.

b) reasekuracja nadwyżkowa (*surplus*)

$$X_r = \begin{cases} (1-\frac{s}{I})X & , \quad I > s \\ 0 & , \quad I \leqslant s \end{cases},$$

$$X_c = \begin{cases} \frac{s}{I}X & , \quad I > s \\ X & , \quad I \leqslant s \end{cases},$$

gdzie s jest poziomem retencji, czyli górnym limitem odpowiedzialności towarzystwa ubezpieczeniowego, a I górną wartością szkody.

2. Reasekuracja nieproporcjonalna

a) reasekuracja nadwyżkowa (*kontrakt stop-loss*)

$$\begin{aligned} X_r &= (X-d)_+, \\ X_c &= \min(X,d) \end{aligned}$$

Uwaga: Jeżeli reasekurujemy szkody kolektywne, tzn. łączna szkoda ma postać $X := S_N = \sum_{i=1}^N X_i$, to rozróżniamy dwa typy reasekuracji nadwyżkowej: (Z oznacza wartość przekazaną do reasekuracji)

$$Z = \sum_{i=1}^{N}(X_i - b)_+ \quad \text{(excess-of-loss)}, \quad b > 0,$$

$$Z = (X-b)_+ \quad \text{(stop-loss)}, \quad b > 0,$$

b) reasekuracja r największych wypłat:

Niech $X_{(1)} \leqslant \cdots \leqslant X_{(n)}$ będą uporządkowanymi stratami w portfelu. Umowa reasekuracyjna mówi, że r największych wypłat pokrywa reasekurator, tzn.

$$\begin{aligned} X_r &= \sum_{i=1}^{r} X_{(n-i+1)}, \\ X_c &= \sum_{i=1}^{n-r} X_{(i)}. \end{aligned}$$

c) reasekuracja ECOMOR

Niech $X_{(1)} \leqslant \cdots \leqslant X_{(n)}$ będą uporządkowanymi stratami w portfelu. Reasekurator pokrywa nadwyżkę ponad poziom $X_{(n-r)}$ dla ustalonego r, a więc

$$\begin{aligned} X_r &= \sum_{i=1}^{r}(X_{(n-i+1)} - X_{(n-r)}), \\ X_c &= \sum_{i=1}^{n-r} X_{(i)} + rX_{(n-r)}. \end{aligned}$$

3.4.1 Wycena kontraktu stop-loss

Zajmiemy się teraz bliżej własnościami składki netto kontraktu stop-loss. Przyjmujemy oznaczenie $I_d(s) = (s-d)_+$ (tutaj udział własny jest wyznaczony funkcją $h(s) = \min(s, d)$). Natychmiast możemy wyliczyć wartość netto ryzyka części ryzyka S przekazywanego do reasekuracji:

$$\mathrm{E}\left[I_d(S)\right] = \int_d^\infty (s-d)dF_S(s) = \int_d^\infty (1 - F_S(s))ds.$$

Zauważmy, że jest to pole pod ogonem dystrybuanty ryzyka na przedziale $[d, \infty)$. Przy założeniu, że dystrybuanta F_S jest różniczkowalna, różniczkując $E[I_d(S)]$ względem d otrzymujemy

$$E[I_d(S)]' = -(1 - F_S(d)),$$

różniczkując ponownie otrzymujemy

$$E[I_d(S)]'' = f_S(d).$$

Ponieważ $-(1 - F_S(d)) \leqslant 0$ oraz $f_S(d) \geqslant 0$, wnioskujemy, że $\mathrm{E}\left[I_d(S)\right]$ jest funkcją nierosnącą i wypukłą zmiennej d. W szczególności, dla ryzyka wykładniczego $S \sim Exp(\beta)$ mamy

$$\mathrm{E}\left[I_d(S)\right] = \int_d^\infty (\exp(-\beta s))ds = \frac{e^{-d\beta}}{\beta}.$$

Niech teraz $L(d) = \mathrm{E}\left[\min(S, d)\right]$. Funkcja ta jest wypukła i rosnąca zmiennej d, bo

$$L(d) = \int_0^d 1 - F_S(s)ds,$$

gdyż jak wiadomo $\mathrm{E}\left[I_d(S)\right] + L(d) = \mathrm{E}\left[S\right] = \int_0^\infty 1 - F_S(s)ds$.

Przykład 3.4.1 [EA:2.06.2001(3)] Wielkość szkody S jest zmienną losową o rozkładzie $U[0, 10]$. Poza pytaniem o wartość netto warto wiedzieć ile wynosi wariancja $VarI_5(S)$. Zauważmy, że $\mathrm{E}\left[(S-5)_+^k\right] = \frac{1}{10}\int_0^{10}(x-5)_+^k dx = \frac{1}{10}\int_5^{10}(x-5)^k dx = \frac{1}{10(k+1)}5^{k+1}$. W szczególności, dla $k = 1$ i $k = 2$ mamy, odpowiednio, $\frac{5}{4}$ i $\frac{25}{6}$, a stąd $\mathrm{Var}\left[(S-5)_+\right] = \frac{125}{48}$.

Z drugiej strony mamy $\mathrm{E}\left[\min(S, 5)^k\right] = 5^k P(S > 5) + \frac{1}{10}\int_0^5 x^k dx = \frac{1}{2}5^k + \frac{1}{10}\frac{1}{k+1}5^{k+1}$ i w szczególności

$$\begin{aligned} \mathrm{E}\left[\min(S,5)\right] &= \frac{15}{4}, \\ \mathrm{Var}\left[\min(S,5)\right] &= \frac{100}{6} - \left(\frac{15}{4}\right)^2 = \frac{125}{48}. \end{aligned}$$

Obliczając współczynnik zmienności $\gamma_2(Y) = \frac{\sigma_Y}{\mu_Y}$ dla zmiennych losowych $Y := S$, $(S-5)_+$,

$\min(S,5)$ dostajemy odpowiednio $\frac{\sqrt{3}}{3}$, $\frac{\sqrt{15}}{3}$, $\frac{\sqrt{15}}{9}$. Kontrakt stop-loss zmniejsza więc współczynnik zmienności posiadacza ryzyka $\min(S,d)$, podczas, gdy analogiczna wielkość dla $(S-d)_+$ jest większa niż dla wyjściowej zmiennej losowej S.

□

Przykład 3.4.2 [EA:23.10.1999(6)] O rozkładzie pewnego ryzyka S wiemy, że:

- $E[(S-20)_+] = 8$
- $E[S|10 < S \leqslant 20] = 13$
- $P(S \leqslant 20) = \frac{3}{4}$
- $P(S \leqslant 10) = \frac{1}{4}$.

Obliczymy $\mathrm{E}\left[(S-10)_+\right]$. Mamy

$$\begin{aligned}
\mathrm{E}\left[(S-10)_+\right] &= \mathrm{E}\left[(S-10)I(S>10)\right] = \mathrm{E}\left[SI(S>10)\right] - 10P(S>10) \\
&= \mathrm{E}\left[SI(20 \geqslant S > 10)\right] + \mathrm{E}\left[SI(S>20)\right] - 10P(S>10) \\
&= \mathrm{E}\left[SI(20 \geqslant S > 10)\right] + \mathrm{E}\left[(S-20)I(S>20)\right] \\
&\quad +20P(S>20) - 10P(S>10) \\
&= \mathrm{E}\left[S|10 < S \leqslant 20\right] P(10 < S \leqslant 20) + \mathrm{E}\left[(S-20)_+\right] \\
&\quad +20P(S>20) - 10P(S>10) \\
&= 13 \cdot 0.5 + 8 + 5 - 7.5 = 12.
\end{aligned}$$

□

Przykład 3.4.3 [EA:15.01.2000(4)] Zmienna losowa S przyjmuje wartości nieujemne. Mamy następujące dane

i	d_i	$F(d_i)$	$\mathrm{E}\left[(S-d_i)_+\right]$
1	5	0.45	20
2	7	0.65	19

Obliczymy $\mathrm{E}\left[S|5 < S \leqslant 7\right]$. Mamy

$$\begin{aligned}
\mathrm{E}\left[S|5 < S \leqslant 7\right] &= \frac{\mathrm{E}\left[SI(5 < S \leqslant 7)\right]}{P(5 < S \leqslant 7)} \\
&= \frac{\mathrm{E}\left[(S-5)_+\right] - \mathrm{E}\left[(S-7)_+\right] + 5P(S>5) - 7P(S>7)}{P(5 < S \leqslant 7)} \\
&= 6.5
\end{aligned}$$

□

3.4.2 Własności kontraktu stop-loss

Dwa ryzyka X, Y możemy porównać również poprzez odpowiadające im wartości netto wielkości przekazywanej do reasekuracji w kontraktach stop-loss. Wprowadza się nastepującą relację

$$X <_{icx} Y \equiv \mathrm{E}\left[(X-d)_+\right] \leqslant \mathrm{E}\left[(Y-d)_+\right], \ \forall d \in \mathbb{R}$$

Natychmiast otrzymujemy charakteryzacje.

Twierdzenie 3.4.4 *Następujące warunki są równoważne*

(i) $X <_{icx} Y$,

(ii) *dla wszystkich wypukłych niemalejących funkcji* g, $\mathrm{E}\left[g(X)\right] \leqslant \mathrm{E}\left[g(Y)\right]$,

(iii) $\int_t^\infty 1 - F_X(s)ds \leqslant \int_t^\infty 1 - F_Y(s)ds, \ \forall t \in \mathbb{R}$.

Gdy $EX = EY$, to piszemy wtedy $X <_{cx} Y$. W tym przypadku zachodzi natychmiast $VarX \leqslant VarY$.

Nastepujące przykłady są wzięte z książki Muellera i Stoyana (2002).

Przykład 3.4.5 Dla zmiennych X, Y o rozkładach $Gamma(\alpha_1, \beta_1)$, $Gamma(\alpha_2, \beta_2)$ odpowiednio, $X <_{icx} Y$ jeśli $\alpha_1 \geqslant \alpha_2$ i $\alpha_1/\beta_1 \leqslant \alpha_2/\beta_2$.

Dla zmiennych X, Y o rozkładach Weibulla $Wei(r_1, c_1)$, $Wei(r_2, c_2)$ (rozkłady zadane przez ogon dystrybuanty $1 - F(x) = \exp(-cx^r)$, dla $r, c > 0$), $X <_{icx} Y$ jeśli $r_1 \geqslant r_2$ i $\mu_1(X) \leqslant \mu_1(Y)$.

Dla zmiennych X, Y o rozkładach log-normalnych $LN(\mu_1, \sigma_1)$, $LN(\mu_2, \sigma_2)$, $X <_{icx} Y$ jeśli $\sigma_1 \leqslant \sigma_2$ i $\mu_1(X) \leqslant \mu_1(Y)$.

□

Przykład 3.4.6 W klasie wszystkich zmiennych losowych, które mają ustaloną wartość średnią, powiedzmy μ, najmniejszą zmienną względem relacji $<_{icx}$ jest zmienna $X \equiv \mu$. Nie istnieje element maksymalny w tej klasie. Jeśli jednak zawężymy klasę zmiennych o ustalonej wartości średniej μ przez dodatkowe wymaganie, aby wartości zmiennych leżały w skończonym przedziale $[a, b]$, to rozkładem maksymalnym względem $<_{icx}$ w takiej klasie jest rozkład o dystrybuancie

$$F_{max}(x) = \frac{b-\mu}{b-a} I_{[a,\infty)}(x) + \frac{\mu-a}{b-a} I_{[b,\infty)}(x).$$

□

Przykład 3.4.7 W klasie wszystkich zmiennych losowych, które mają ustaloną wartość średnią, powiedzmy μ oraz ustalona wariancję σ^2 nie istnieje element maksymalny względem $<_{icx}$. Istnieje jednak dystrybuanta, która jest ograniczeniem górnym dla tej klasy, najmniejszym z możliwych (tzn. jest ona kresem górnym tej klasy, lecz do niej nie należy). Jest to dystrybuanta

$$F_{sup}(x) = F_0(\frac{x-\mu}{\sigma}),\ F_0(x) = 1/2 + \frac{x}{2(x^2+1)^{1/2}}.$$

□

Skutecznym kryterium do otrzymywania relacji $<_{icx}$ jest tak zwane **kryterium Karlina-Novikowa** .

Twierdzenie 3.4.8 *Jeśli* $\mathrm{E}[X] \leqslant \mathrm{E}[Y]$ *oraz istnieje* x_0 *takie, że*

$$F_Y(x) \geqslant F_X(x),\ \forall x < x_0$$
$$F_Y(x) \leqslant F_X(x),\ \forall x > x_0,$$

to $X <_{icx} Y$.

Dowód. Korzystajc z faktu, że

$$\mathrm{E}[X] = \int_0^\infty 1 - F_X(u)du - \int_{-\infty}^0 F_X(u)du,$$

otrzymujemy

$$\mathrm{E}[Y] - \mathrm{E}[X] = \int_{-\infty}^\infty F_X(u) - F_Y(u)du. \tag{3.4.1}$$

Wystarczy pokazać, korzystając z twierdzenia 3.4.4, (iii), że

$$\int_x^\infty F_X(u) - F_Y(u)du \geqslant 0,\ \forall x \in \mathbb{R}.$$

Gdy $x > x_0$ wynika to założenia pierwszego. Gdy $x < x_0$, zauważmy, korzystając z (3.4.1), że

$$\int_x^\infty F_X(u) - F_Y(u)du = \mathrm{E}[Y] - \mathrm{E}[X] + \int_{-\infty}^x F_Y(u) - F_X(u)du \geqslant 0,$$

gdzie nierówność wynika z założenia o uporządkowaniu wartości oczekiwanych oraz z warunku drugiego na liście założeń.

□

Zajmiemy się teraz wyborem najodpowiedniejszego (dla cedenta) sposobu reasekuracji. Poniższe twierdzenie pokazuje, że sposród wszystkich kontraktów o ustalonej składce netto $\mathrm{E}\left[h(S)\right]$ wariancję minimalizuje kontrakt stop-loss.

Twierdzenie 3.4.9 *Niech $S \geqslant 0$ będzie ustalonym ryzykiem. Jeśli d^* jest wyznaczone równaniem*

$$\mathrm{E}\left[\min(d^*, S)\right] = C,$$

dla ustalonej wartosci netto C udziału własnego, to

$$\min(d^*, S) <_{icx} h(S),$$

dla każdego kontraktu $0 \leqslant h(S) \leqslant S$, takiego, że $\mathrm{E}\left[h(S)\right] = C$.

Dowód. Pokażemy, że dla każdego d

$$\int_d^\infty \bar{F}_{\min(d^*,S)}(u)du \leqslant \int_d^\infty \bar{F}_{h(S)}(u)du.$$

Niech $d \leqslant d^*$. Wtedy (z definicji obcięcia)

$$\int_0^d \bar{F}_{\min(d^*,S)}(u)du = \int_0^d \bar{F}_S(u)du \geqslant \int_0^d \bar{F}_{h(S)}(u)du,$$

bo $h(S) \leqslant S$ założenia. Ponieważ $\int_0^\infty \bar{F}_{\min(d^*,S)}(u)du = \int_0^\infty \bar{F}_{h(S)}(u)du = C$ (równość wartości oczekiwanych), to

$$\int_d^\infty \bar{F}_{\min(d^*,S)}(u)du \leqslant \int_0^d \bar{F}_{h(S)}(u)du.$$

Gdy $d \geqslant d^*$, to $\int_d^\infty \bar{F}_{\min(d^*,S)}(u)du = 0$.

□

Jako natychmiastowy wniosek z powyższego twierdzenia otrzymujemy następujący wynik (który zaopatrzymy w niezależny dowód).

Twierdzenie 3.4.10 *Niech S będzie ustalonym ryzykiem oraz C ustaloną wartością netto udziału własnego, wtedy*

$$\min_{\{h:\mathrm{E}[h(S)]=C\}} \mathrm{Var}\left[h(S)\right] = \mathrm{Var}\left[\min(d^*, S)\right],$$

przy czym d^* wyznaczone jest przez $\mathrm{E}\left[\min(d^*, S)\right] = C$.

Dowód. Licząc wariancję udziału własnego widzimy, że dla dowolnego ryzyka X i dowolnego udziału własnego $h(X)$ takiego, że $\mathrm{E}\left[h(X)\right] = C$ i dla każdego $d \geqslant 0$ zachodzi

$$\begin{aligned}\mathrm{Var}\left[h(X)\right] &= \int_0^\infty (h(x) - C)^2 dF_X(x) \\ &= \int_0^\infty (h(x) - d + d - C)^2 dF_X(x) \\ &= \int_0^\infty (h(x) - d)^2 - 2\int_0^\infty (d - h(x))(d - C) dF_X(x) + (d - C)^2 \\ &= \int_0^\infty (h(x) - d)^2 dF_X(x) - (d - C)^2.\end{aligned}$$

Wstawiając w szczególności $h(X) := min(d^*, S)$, mamy

$$\begin{aligned}\mathrm{Var}\left[\min(S, d^*)\right] &= \int_0^\infty (\min(x, d^*) - d^*)^2 dF_S(x) - (C - d^*)^2 \\ &= \int_0^{d^*} (x - d^*)^2 dF_S(x) - (C - d^*)^2.\end{aligned}$$

Dla $h(x) \leqslant x \leqslant d^*$ zachodzi $(x - d^*)^2 \leqslant (h(x) - d^*)^2$, stąd

$$\int_0^{d^*} (x - d^*)^2 dF_S(x) \leqslant \int_0^{d^*} (h(x) - d^*)^2 dF_S(x) \leqslant \int_0^\infty (h(x) - d^*)^2 dF_S(x),$$

a więc wykorzystując ogólny wzór na wariancję z $d = d^*$ i $X = S$ widzimy, że

$$\mathrm{Var}\left[h(S)\right] \leqslant \mathrm{Var}\left[\min(S, d^*)\right],$$

dla każdego h takiego, że $\mathrm{E}\left[h(X)\right] = C$, co kończy dowód. □

Niech $X = h(X) + k(X)$ będzie podziałem ryzyka X, takim, że jak zwykle $0 \leqslant h(x) \leqslant x$, $0 \leqslant k(x) \leqslant x$ oraz $\mathrm{E}\left[k(X)\right] = P \leqslant \mathrm{E}\left[X\right]$ jest ustalone. Niech d^* będzie poziomem odpowiedzialności w kontrakcie stop-loss takim, że $\mathrm{E}\left[I_{d^*}(X)\right] = P$. Zauważmy, że jest ta sama wartość d^*, która wcześniej spełniała warunek $\mathrm{E}\left[\min(X, d^*)\right] = C$, dla $C = \mathrm{E}\left[X\right] - P$.

Twierdzenie 3.4.11 *Dla ustalonego kapitału początkowego w, wklęsłej funkcji użyteczności u oraz $0 \leqslant P \leqslant \mathrm{E}\left[X\right]$,*

$$\max_{k(x):\mathrm{E}[k(X)]=P} \mathrm{E}\left[u(w + k(X) - X - P)\right] = \mathrm{E}\left[u(w + I_{d^*}(X) - X - P)\right],$$

gdzie d^* spełnia

$$\int_{d^*}^\infty (x - d^*) dF_X(x) = P.$$

Dowód 1. Ponieważ u jest wklęsła, niemalejąca, to funkcja $g(x) := -u(-x+c)$ jest wypukła niemalejąca. Wystarczy pokazać, że

$$\mathrm{E}\left[u(w + k(X) - X - P)\right] \leqslant \mathrm{E}\left[u(w + I_{d^*}(X) - X - P)\right],$$

co jest równoważne

$$\mathrm{E}\left[u(w-h(X)-P)\right] \leqslant \mathrm{E}\left[u(w-\min(d^*,X)-P)\right],$$

dla każdego podziału ryzyka $X = h(X)+k(X)$ spełniającego założenia twierdzenia. Biorąc $c := w - P$, ostatnią nierówność można zapisać w postaci

$$\mathrm{E}\left[g(h(X)\right] \geqslant \mathrm{E}\left[g(\min(d^*,X)),\right]$$

co wynika z faktu pokazanego wcześniej, że $\min(d^*, X) <_{icx} h(X)$. □

Dowód 2. Zauważmy, że dla dowolnej różniczkowalnej funkcji wklęsłej i $a < b$ mamy

$$u'(b) \leqslant \frac{u(b)-u(a)}{b-a} \leqslant u'(a). \tag{3.4.2}$$

Z (3.4.2) natychmiast otrzymujemy dla każdego x

$$u(w-x-P+k(x))-u(w-x-P+I_{d^*}(x)) \leqslant u'(w-x-P+I_{d^*}(x))\cdot(k(x)-I_{d^*}(x)).$$

Korzystając z faktu, że $I_{d^*}(x) - x \geqslant -d^*$ i z monotoniczności u' mamy

$$u(w-x-P+k(x))-u(w-x-P+I_{d^*}(x)) \leqslant u'(w-P-d^*)\cdot(k(x)-I_{d^*}(x)).$$

Ponieważ nierówność ta jest prawdziwa dla każdego x, więc jest również prawdziwa, gdy wstawimy zamiast zmiennej x zmienną losową X, a następnie po obu stronach nierówności nałożymy wartość oczekiwaną. W rezultacie otrzymujemy dla każdego $k(x)$ takiego, że $\mathrm{E}\left[k(X)\right] = \mathrm{E}\left[I_{d^*}(X)\right] = P$

$$\mathrm{E}\left[u(w-X-P+k(X))-u(w-X-P+I_{d^*}(X))\right] \leqslant 0,$$

co dowodzi tezy twierdzenia.

□

3.5 Stochastyczne porównywanie ryzyk

Dwa ryzyka X, Y możemy porównać poprzez odpowiadające im wartości netto udziału własnego w kontraktach stop-loss. Wprowadza się nastepującą relację

$$X <_{icv} Y \equiv \mathrm{E}\left[\min(X,d)\right] \leqslant \mathrm{E}\left[\min(Y,d)\right], \forall d \in \mathbb{R}$$

Proponowany sposób porównywania ryzyk można opisać następująco

Twierdzenie 3.5.1 *Następujące warunki są równoważne*

(i) $X <_{icv} Y$,

(ii) dla wszystkich wklęsłych niemalejących funkcji g, $\mathrm{E}\left[g(X)\right] \leqslant \mathrm{E}\left[g(Y)\right]$,

(iii) $\int_0^{\max(0,t)} 1 - F_X(s)ds - \int_{-\infty}^{\min(0,t)} F_X(s)ds$
$\leqslant \int_0^{\max(0,t)} 1 - F_Y(s)ds - \int_{-\infty}^{\min(0,t)} F_Y(s)ds,\ \forall t \in \mathbb{R}.$

Dowód. Z definicji $X <_{icv} Y$, nierówność $\mathrm{E}\left[g(X)\right] \leqslant \mathrm{E}\left[g(Y)\right]$ zachodzi dla funkcji $g(x) = \min(d, x)$, gdzie $d \in \mathbb{R}$, które oczywiście są niemalejące wklęsłe. Łatwo zauważyć, że dowolną funkcję niemalejącą wklęsłą możemy punktowo w sposób monotoniczny przybliżyć kombinacjami liniowymi funkcji postaci $\min(d, x)$. Stąd po nałożeniu wartości oczekiwanych i przejściu do granicy w aproksymacji dostajemy $\mathrm{E}\left[g(X)\right] \leqslant \mathrm{E}\left[g(Y)\right]$ dla dowolnych wklęsłych niemalejących funkcji g.

Ostatni warunek jest bezpośrednią konsekwencją użycia wzoru na wartość oczekiwaną dowolnej zmiennej losowej Z

$$\mathrm{E}\left[Z\right] = \int_0^\infty 1 - F_Z(x)dx - \int_{-\infty}^0 F_Z(x)dx$$

oraz postaci ogona dystrybuanty zmiennej $\min(t, X)$,

$$1 - F_{\min(t,X)}(x) = (1 - F_X(x))I_{(-\infty,t)}(x).$$

□

Wprowadzone stochastyczne metody porządkowania ryzyk są ogólniejszym spojrzeniem na klasyczny sposób porządkowania ryzyk poprzez reguły decyzyjne. Klasyczna **reguła decyzyjna Markowitza** jest określona przez porównywanie wartości

$$U(X) := \mathrm{E}\left[X\right] - \alpha \mathrm{Var}\left[X\right],\ \alpha > 0.$$

Jeśli ryzyka X, Y mają rozkłady normalne, to $X <_{icv} Y$ wtedy i tylko wtedy, gdy $\mathrm{E}\left[X\right] \leqslant \mathrm{E}\left[Y\right]$ i $\mathrm{Var}\left[X\right] \geqslant \mathrm{Var}\left[Y\right]$, z czego wynika, że zgodnie z regułą Markowitza $U(X) \leqslant U(Y)$ wtedy i tylko wtedy, gdy $X <_{icv} Y$, dla ryzyk o rozkładzie normalnym. Powstaje naturalne pytanie o inne rozkłady i ich zgodność z regułami podobnymi do reguły Markowitza. Zachodzi następująca zgodność, która nie wymaga założeń o typie rozkładu, ale dotyczy reguły mierzącej rozrzut inaczej niż w regule Markowitza:

Twierdzenie 3.5.2 *Jeśli*

$$U_1(X) := \mathrm{E}\left[X\right] - \alpha\delta^{(1)}(X),$$

gdzie $\alpha \in (0, 1)$, $\delta^{(1)}(X) = \mathrm{E}\left[\left|X - \mathrm{E}\left[X\right]\right|\right]/2$, *to*

$$X <_{icv} Y \Rightarrow U_1(X) \leqslant U_1(Y).$$

Dowód. Zauważmy najpierw, że dla dowolnej liczby rzeczywistej a, możemy dokonać rozkładu na dwie części $a = a_+ - a_-$, gdzie $a_+ = \max(0, a)$, $a_- = -\min(0, a)$. Wtedy

$|a| = a_+ + a_-$. Stąd, jeśli $\mathrm{E}[X] = 0$, to $\mathrm{E}[X_+] = \mathrm{E}[X_-]$ oraz $\mathrm{E}[(X - \mathrm{E}[X])_+] = \mathrm{E}[(X - \mathrm{E}[X])_-]$, co daje

$$\mathrm{E}[|X - \mathrm{E}[X]|] = 2\mathrm{E}[(X - \mathrm{E}[X])_-].$$

Wystarczy więc pokazać, że

$$\mathrm{E}[X] - \mathrm{E}[(X - \mathrm{E}[X])_-] \leqslant \mathrm{E}[Y] - \mathrm{E}[(Y - \mathrm{E}[Y])_-].$$

Nierówność ta wynika z nastepujących nierówności

$$\begin{aligned}\mathrm{E}[(Y - \mathrm{E}[Y])_-] - \mathrm{E}[(X - \mathrm{E}[X])_-] &\leqslant \mathrm{E}[(X - \mathrm{E}[Y])_-] - \mathrm{E}[(X - \mathrm{E}[X])_-] \\ &\leqslant \mathrm{E}[Y] - \mathrm{E}[X].\end{aligned}$$

Pierwsza z nich zachodzi ponieważ z warunku $X <_{icv} Y$ wynika, że

$$-\mathrm{E}[(X - t)_-] \leqslant -\mathrm{E}[(Y - t)_-], \forall t \in \mathbb{R},$$

ponieważ funkcja określona przez $x \to -(x - t)_-$ jest niemalejąca i wklęsła. Wstawiając $t := \mathrm{E}[Y]$ i monożąc obustronnie przez -1 otrzymujemy pierwszą nierówność. Druga nierówność wynika z nastepującej relacji

$$\mathrm{E}[(X - (t + \Delta))_-] - \mathrm{E}[(X - t)_-] \leqslant \Delta, \ \forall \Delta \geqslant 0,$$

gdy wstawimy $t := \mathrm{E}[X]$, $\Delta := \mathrm{E}[Y] - \mathrm{E}[X]$. Ostatnia relacja jest oczywista w świetle tożsamości

$$\begin{aligned}\mathrm{E}[(X - (t + \Delta))_-] &= -\mathrm{E}[\min(X - t - \Delta, 0)] \\ \mathrm{E}[(X - t)_-] &= -\mathrm{E}[\min(X - t, 0)].\end{aligned}$$

□

Oprócz prostej zawartości portfela $S_n = X_1 + \cdots + X_n$ interesujące są kombinacje liniowe portfela $S_n(\mathbf{a}) = a_1X_1 + \cdots + a_nX_n$, dla wekotora $\mathbf{a} = (a_1, \ldots, a_n)$ o nieujemnych współrzędnych. Klasycznym zagadnieniem jest wycena wartości $S_n(\mathbf{a})$ oraz wyznaczenie optymalnego wyboru $\mathbf{a}$ o ustalonej sumie współrzędnych $\sum_{i=1}^n a_i = m > 0$. Wyznaczenie

$$\max_{\{\mathbf{a}:\sum_{i=1}^n a_i = m\}} \mathrm{E}[u(S_n(\mathbf{a}))],$$

można zinterpretować jako zagadnienie optymalnej alokacji środków wielkości m względem funkcji użyteczności u. Jeśli u jest wklęsła, to zagadnienie to sprowadza się do poszukiwania maksimum względem porządku $<_{icv}$ w klasie zmiennych losowych $\{S_n(\mathbf{a}) : \sum_{i=1}^n a_i = m\}$. Okazuje się, że jeśli rozkład łączny portfela $(X_1, \ldots, X_n)$ jest niezmienniczy na permutacje, to optymalną alokacją jest rozkład równomierny (zob. Tw. 8.2.3, Mueller, Stoyan (2002)).

Twierdzenie 3.5.3 *Jeśli* $(X_{\pi(1)}, \ldots, X_{\pi(n)})$ *ma taki sam rozkład łączny jak* $(X_1, \ldots, X_n)$, *dla dowolnej permutacji* $\pi(n)$ *indeksów portfela, to dla u wklęsłej*

$$\max_{\{\mathbf{a}:\sum_{i=1}^{n} a_i = m\}} \mathrm{E}\left[u(S_n(\mathbf{a}))\right]$$

jest osiagnięte w punkcie $\mathbf{a}^* = (m/n, \ldots, m/n)$.

Szczególnym przypadkiem portfela spełniającego warunek permutowalności jest portfel prosty niezależnych ryzyk o jednakowym rozkładzie.

Kolejną relację stochastycznego uporządkowania między ryzykami, z której relacje $<_{icx}$ oraz $<_{icv}$ wynikają, jest

$$X <_{st} Y \equiv \mathrm{E}\left[g(X)\right] \leqslant \mathrm{E}\left[g(Y)\right],$$

dla wszystkich niemalejących funkcji g, dla których wartości oczekiwane istnieją.

Wstawiając indykatory półprostych w miejsce g natychmiast widzimy, że $X <_{st} Y$ implikuje, że $F_X \geqslant F_Y$. Implikacja w drugą stronę jest również prawdziwa, co widać przy użyciu odpowiedniej aproksymacji funkcjami prostymi (tzn. kombinacjami liniowymi indykatorów). Ta metoda porównywania zwykle stosowana jest do zawartości portfela S_n. Zamiast porównywania dystrybuant, można porównywać odpowiednie funkcje kwantylowe, które w tym kontekście oznaczamy symbolem

$$VaR(S_n, p) := F_{S_n}^{-1}(p) = \inf\{t : F_{S_n}(t) \geqslant p\},\ p \in (0,1).$$

Symbol VaR pochodzi od określenia tej wielkości w języku angielskim - Value at Risk. Mamy więc równoważność

$$VaR(S_n, p) \leqslant VaR(S'_n, p)\ \forall p \in (0,1) \Leftrightarrow S_n <_{st} S'_n$$

Porównywanie wartości VaR różnych ryzyk zwykle jest interesujące dla wartości parametru p bliskiej 1, a do tego celu nie musi zachodzić tak mocna relacja między ryzykami jak $<_{st}$. Zauważmy, że kryterium Karlina-Novikowa implikuje nierówność dla VaR dla dostatecznie dużych p (bliskich 1).

Stochastyczny wzrost wielkości portfela może nastąpić oczywiście wtedy, gdy zmienią się parametry rozkładów składowych w portfelu (zwiększając np. rozkład wzgledem relacji $<_{icx}$ lub nawet $<_{st}$ i prowadząc do zwiększenia VaR). Wartą podkreślenia jest obserwacja, że wzrost zawartości portfela mierzony relacją $<_{icx}$ może nastąpić w wyniku zmiany typu zależności składowych portfela przy ustalonych indywidulanych rozkładach składowych. Jest to dość ogólna prawidłowość związana z konkretnymi typami zależności stochastycznej, którą najpierw zilustrujemy prostym przykładem.

Przykład 3.5.4 Portfel ubezpieczeniowy pewnej firmy składa się z $n = 10^6$ kontraktów ubezpieczających dom na następny rok. Przyjmuje się, że szansa zniszczenia domu wynosi 10^{-4}. Przyjmujemy $P(X_i = 1) = 10^{-4} = 1 - P(X_i = 0)$. Wtedy $\mathrm{E}\left[S_n\right] = 100$. Niech

$d = 150$ i załóżmy najpierw, że $X_1, \ldots, X_n$ są niezależne. Wtedy $\mathrm{Var}\,[S_n] = 100(1-10^{-4}) \approx 100$. Z CTG można użyć aproksymacji $S_n \approx N(100, \sigma^2 = 100)$. Wartość netto nadwyżki nad d wynosi

$$\mathrm{E}\,[(S_n - d)_+] = \int_{150}^{\infty} 1 - \Phi_{N(100,100)}(u)du \approx 3 \cdot 10^{-8} \approx 0.$$

Co oznacza, że składka za reasekurację powinna być bardzo mała. Założenie o niezależności ryzyk przy ubezpieczaniu domów jest jednak bardzo mało realistyczne. Szczególnie w rejonie huraganów, gdzie wiele domów jednocześnie jest niszczonych (zdarzenia te powodowane są jedną przyczyną). Można więc założyć, że istnieje jedna losowa przyczyna, którą będziemy modelować zakładając istnienie zmiennej losowej Θ indykującej zajście huraganu. Niech $P(\Theta = 1) = 1/100 = 1 - P(\Theta = 0)$. Załóżmy teraz, że zmienne $X_i', i = 1, \ldots, n$ nie są niezależne, lecz, że są niezależne warunkowo, tzn.

$$P(X_1' = x_1, \ldots, X_n' = x_n|\Theta = \theta) = P(X_1' = x_1|\Theta = \theta) \cdots P(X_n' = x_n|\Theta = \theta).$$

Przyjmijmy, że dla $i = 1, \ldots, n$

$$P(X_i' = 1|\Theta = 1) = 10^{-3} P(X_i' = 1|\Theta = 0) = 1/11000.$$

Wtedy $P(X_i' = 1) = 10^{-4}$, czyli rozkłady pojedynczych ryzyk są takie same jak poprzednio, przy założeniu niezależności. Łatwo znajdujemy, że dla $i \neq j$

$$Corr(X_i', X_j') = \mathrm{E}\,[(X_i' - \mathrm{E}\,[X_i'])(X_i' - \mathrm{E}\,[X_i'])]\,/\sigma_{X_i'}\sigma_{X_j'} = 10^{-4},$$

czyli sądząc jedynie po wartości korelacji można by przypuszczać, że nowe zmienne nie są mocno zależne. Warunkując względem Θ i pod warunkiem ustalonej wartości stosując przybliżenie CTG, mamy

$$S_n' \sim 0.01N(1000, 1000) + 0.99N(1000/11, 1000/11).$$

Licząc dla tej mieszanki wartość netto nadwyżki nad $d = 150$, otrzymujemy

$$\mathrm{E}\,[(S_n' - d)_+] \approx 8.5.$$

Widać więc wyraźny wzrost wartości netto w kontrakcie stop-loss, w wyniku wprowadzenia zależności między ryzykami. Powtarzając rachunki dla innych wartości d widzimy, że taki wzrost następuje dla każdego d, a stąd S_n' w przypadku zależności jest większe w relacji $<_{icx}$ od wartości portfela S_n w przypadku niezależności.

□

Istnieje wiele możliwości zdefiniowania zależności między zmiennymi losowymi. Przypomnijmy kilka z nich.

(A). Mówimy, że rozkład wektora $(X_1, \ldots, X_n)$ jest MTP_2 jeśli istnieje gęstość tego rozkładu $f_{\mathbf{X}} = f_{(X_1,\ldots,X_n)}$ spełniająca

$$f_{\mathbf{X}}(\min(\mathbf{x},\mathbf{y}))f_{\mathbf{X}}(\max(\mathbf{x},\mathbf{y})) \geqslant f_{\mathbf{X}}(\mathbf{x})f_{\mathbf{X}}(\mathbf{y}),$$

dla każdego $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$.

(B). Mówimy, że rozkład wektora $(X_1,\ldots,X_n)$ jest CIS jeśli dla każdego $i = 2,\ldots,n$, funkcja argumentów $x_1,...,x_{i-1}$ zdefiniowana przez

$$\mathrm{E}\left[g(X_i)|X_1 = x_1,\ldots,X_{i-1} = x_{i-1}\right]$$

jest niemalejąca dla każdej niemalejącej funkcji g.

(C). Mówimy, że rozkład wektora $(X_1,\ldots,X_n)$ jest stowarzyszony (lub wektor jest stowarzyszony) jeśli

$$Cov(g(\mathbf{X}),h(\mathbf{X})) \geqslant 0$$

dla wszystkich funkcji $g,h : \mathbb{R}^n \to \mathbb{R}$ niemalejących po współrzędnych.

(D). Mówimy, że rozkład wektora $(X_1,\ldots,X_n)$ jest WAS jeśli dla każdego $i = 2,\ldots,n$

$$Cov(I_{(t,\infty)}(X_i),g(X_{i+1},\ldots,X_n)) \geqslant 0 \ \forall t \in \mathbb{R}$$

(E). Mówimy, że rozkład wektora $(X_1,\ldots,X_n)$ jest $PSMD$ jeśli

$$\mathrm{E}\left[g(\mathbf{X}^*)\right] \leqslant \mathrm{E}\left[g(\mathbf{X})\right]$$

dla wszystkich supermodularnych funkcji $g : \mathbb{R}^n \to \mathbb{R}$, tzn. funkcji dla których $\frac{\partial^2 g}{\partial x_i \partial x_j} \geqslant 0$, $\forall i < j$, gdzie $\mathbf{X}^*$ jest wektorem o współrzędnych niezależnych, takim, że X_i ma identyczny rozkład z rozkładem X_i^*.

Powyższa relacja jest szczególnym przypadkiem tak zwanego porządku supermodularnego $\mathbf{X}^* <_{sm} \mathbf{X}$, który odzwierciedla porządkowanie parametrów zależności między współrzędnymi wektorów posiadających współrzędne o tych samych rozkładach.

Powyższe definicje typów zależności są ustawione od najmocniejszej do najsłabszej w tym sensie, że jeśli wektor $\mathbf{X}$ spełnia (A), to spełnia (B) itd. tzn. symbolicznie (A)⇒(B)⇒(C)⇒(D)⇒(

Każdy z tych typów zależności składowych portfela implikuje, że zawartość portfela z zależnościami jest większa w sensie relacji $<_{icx}$ od zawartości w portfelu o składowych niezależnych.

Twierdzenie 3.5.5 *Jeśli* $\mathbf{X}$ *spełnia (E), to* $S_n^* <_{icx} S_n$.

Ryzyka w portfelu $(X_1,\ldots,X_n)$ wzajemnie się wyłączają, gdy

$$P(X_i > 0, X_j > 0) = 0 \ , \ \forall i < j.$$

Ryzyka wzajemnie wyłączające się generują najmniejszą w pewnym sensie zawartość portfela (zob. Mueller, Stoyan [17], §8.3).

Twierdzenie 3.5.6 *Jeśli* $(X_1, \ldots, X_n)$ *wzajemnie się wyłączają, to ich suma jest ograniczeniem dolnym w relacji* $<_{icx}$ *w klasie wszystkich portfeli o tych samych rozkładach pojedynczych ryzyk, tzn.*

$$X_1 + \cdots + X_n <_{icx} Y_1 + \cdots + Y_n,$$

dla dowolnego wektora $(Y_1, \ldots, Y_n)$*, dla którego* $F_{X_i} = F_{Y_i}$*,* $i = 1, \ldots, n$*.*

3.6 Miary ryzyka

Ogólna idea miary ryzyka opiera się o założenie, że miarą ryzyka powinna być funkcja, powiedzmy ϱ, która każdemu ryzyku X przyporządkowuje wartość $\varrho(X) \in (0, \infty]$, która odpowiada wartości kwoty, którą trzeba dodać do X, aby ryzyko $X + \varrho(X)$ było *akceptowalne*. Ponieważ *akceptowalność* można określić na różne sposoby istnieje wiele możliwości precyzyjnego określenia takiej miary, która w pewnym sensie *zabezpiecza* ryzyko. Często wymagania dotyczące *zabezpieczania* wynikają z konkretnych regulacji prawnych, takich jak na przykład zawartych w układzie Solvency II, wspomnianym we wstępie skryptu.

Na pojęcie *zabezpieczania* można spojrzeć w następujący sposób. Ryzyko X jest zabezpieczone przez $\varrho(X)$, gdy zdarzenie $\{X > \varrho(X)\}$ zachodzi z ustalonym, małym, prawdopodobieństwem. Inną miarą zabezpieczenia może być wartość $\mathrm{E}\left[(X - \varrho(X))_+\right]$, która powinna być dostatecznie mała.

Istnieje cały szereg porządanych własności, które powinna miara ryzyka posiadać.

1. (no-ripoff) $\varrho(X) \leqslant \max_{\omega \in \Omega}(X(\omega))$,
2. (non-negative loading) $\varrho(X) \geqslant \mathrm{E}\left[X\right]$,
3. tranzytywność $\varrho(X + c) = \varrho(X) + c, \forall c \in \mathbb{R}$,
4. (neutral position) $\varrho(X - \varrho(X)) = 0$,
5. brak narzutu $\varrho(c) = c\ , \forall c \in \mathbb{R}$,
6. podaddytywność $\varrho(X + Y) \leqslant \varrho(X) + \varrho(Y)$,
7. jednorodność $\varrho(cX) = c\varrho(X),\ \forall c \in \mathbb{R}$,
8. monotoniczność $P(X \leqslant Y) = 1\ \Rightarrow \varrho(X) \leqslant \varrho(Y)$.

Ryzyko X może być mierzone nie tylko w oparciu o jedną, wyjściową miarę prawdopodobieństwa, na danej przestrzeni próbkowej. Można przyjąć,

$$\varrho(X) = \sup\{\int_{\mathbb{R}} x dF(x), F \in \mathcal{P}_{\mathcal{X}}\},$$

gdzie klasa rozkładów $\mathcal{P}_{\mathcal{X}}$ jest w pewnej relacji do rozkładu wyjściowego F_X zmiennej X, na przykład będąc klasą rozkładów spełniających pewne ograniczenia. Klasę $\mathcal{P}_{\mathcal{X}}$ można interpretować jako klasę opisującą różne secenariusze realizujące ryzyko, a miara ryzyka jest wtedy wyborem wartości oczekiwanej *najgorszego* scenariusza.

Dla ryzyka X o dystrybuancie F_X często używaną miarą ryzyka jest wspomniana już miara

$$\varrho(X) = VaR[X, p] := F_X^{-1}(p),$$

dla $p \in (0, 1)$.

Używanie tej miary motywowane jest częściowo przez następującą własność

Lemat 3.6.1 *Dla każdego X i $\epsilon > 0$ oraz dla każdej funkcji $\varrho(X)$*

$$\mathrm{E}\left[(X - VaR[X, 1-\epsilon])_+ + VaR[X, 1-\epsilon]\epsilon\right] \leqslant \mathrm{E}\left[(X - \varrho(X))_+ + \varrho(X)\epsilon\right]$$

Dowód. Niech $g(d) := \mathrm{E}\left[(X-d)_+\right] + d\epsilon$, będzie funkcją zmiennej d na $\mathbb{R}$. Wtedy, ponieważ $g(d) = \int_d^\infty 1 - F_X(u)du + d\epsilon$, istnieje $g'(d) = -(1 - F_X(d)) + \epsilon$. Przyrównując pochodną do zera i zauważając, że funkcja osiąga minimum otrzymujemy, że punktem minimum jest d^* taki, że $F_X(d^*) = 1 - \epsilon$, co oznacza, że $d^* = VaR[X, 1-\epsilon]$. □

Wielkość $\mathrm{E}\left[(X - \varrho(X))_+ + \varrho(X)\epsilon\right]$ reprezentuje średnią nadwyżkę nad wielkość zabezpieczenia (wielkość rezerwy) plus koszt utrzymania rezerwy $\varrho(X)$. Jak widać wielkość ta jest najmniejsza jeśli rezerwa jest wyliczona na podstawie VaR, która jako miara ryzyka ma większość pożądanych własności, ale niestety nie wszystkie.

VaR posiada własność no-ripoff, bo

$$VaR[X, p] = F_X^{-1}(p) \leqslant F_X^{-1}(1) = \max_{\omega \in \Omega}(X(\omega)).$$

Dla $p > F_X(\mathrm{E}\left[X\right])$ zachodzi $VaR[X, p] \geqslant \mathrm{E}\left[X\right]$, czyli własność non-negative loading zachodzi dla dostatecznie dużych p.

Tranzytywność $VaR[X + c, p] = VaR[X, p] + c$ jest oczywista. Zachodzi jednorodność

$$VaR[cX, p] = cVaR[X, p].$$

Rzeczywiście $VaR[cX, p] = F_{cX}^{-1}(p) = \inf\{t : F_X(\frac{t}{c}) \geqslant p\} = \inf\{t : \frac{t}{c} \geqslant F_X^{-1}(p)\} = c\inf\{t : t \geqslant F_X^{-1}(p)\} = c\inf\{t : F_X(t) \geqslant p\} = cF_X^{-1}(p) = cVaR[X, p]$.

Zachodzi też własność braku narzutu $VaR[c, p] = c$. Nie zachodzi jednak własność podaddytywności.

Inne częściej używane miary ryzyka:

(Tail Value at Risk)

$$TVaR[X,p] := \frac{1}{1-p}\int_p^1 VaR[X,u]du,$$

(Conditional Tail Expectation)

$$CTE[X,p] := \mathrm{E}\left[X|X > Var[X,p]\right],$$

(Conditional Value at Risk)

$$CVaR[X,p] := CTE[X,p] - VaR[X,p] = \mathrm{E}\left[X - VaR[X,p]|X > VaR[X,p]\right],$$

(Expected Shortfall)

$$ES[X,p] := \mathrm{E}\left[(X - VaR[X,p])_+\right]$$

Zachodzą następujące związki między wprowadzonymi miarami ryzyka.

Twierdzenie 3.6.2 *Dla* $p \in (0,1)$

$$\begin{aligned} TVaR[X,p] &= VaR[X,p] + \frac{ES[X,p]}{1-p}, \\ CTE[X,p] &= VaR[X,p] + \frac{ES[X,p]}{1 - F_X(VaR[X,p])}, \\ CVaR[X,p] &= \frac{ES[X,p]}{1 - F_X(VaR[X,p])}. \end{aligned}$$

Dowód. Licząc odpowiednie pola pod krzywą ogona rozkładu X mamy

$$ES[X,p] = \int_p^1 (F_X^{-1}(u) - VaR[X,p])du = (1-p)TVaR[X,p] - (1-p)VaR[X,p],$$

co daje pierwszą zależność. Ponieważ resztowy rozkład ma ogon dany przez

$$P(X > u + d|X > d) = \frac{1 - F_X(u+d)}{1 - F_X(d)},$$

więc wstawiając $d := VaR[X,p]$ i całkując otrzymujemy

$$\begin{aligned} CVaR[X,p] &= \mathrm{E}\left[X - VaR[X,p]|X > VaR[X,p]\right] \\ &= \frac{\int_0^\infty 1 - F_X(u + VaR[X,p])du}{1 - F_X(VaR[X,p])} \\ &= \frac{ES[X,p]}{1 - F_X(VaR[X,p])}, \end{aligned}$$

co daje ostatnią równość. Kombinując dwie uzyskane równości otrzymujemy trzecią. □

Zauważmy, że gdy F_X jest dystrybuantą ciągłą, to

$$TVaR[X,p] = CTE[X,p].$$

Warto zauważyć, że $TVaR$ ma porządane własności:

$TVaR$ ma własność no-ripoff, gdyż

$$TVaR[X,p] \leqslant \frac{1}{1-p}\int_p^1 \max_{\omega\in\Omega}(X(\omega))du = \max_{\omega\in\Omega}(X(\omega)).$$

Oczywiście

$$TVaR[c,p] = c.$$

Ponieważ

$$TVaR[X,0] = \int_0^1 F_X^{-1}(u)du = \mathrm{E}\left[X\right],$$

wystarczy pokazać, że funkcja $p \to TVaR[X,p]$ jest niemalejąca, aby zobaczyć, że

$$TVaR[X,p] \geqslant \mathrm{E}\left[X\right].$$

Monotoniczność funkcji $TVaR[X,p]$ zmiennej $p \in (0,1)$ sprawdzamy różniczkując ją i korzystając z faktu, że $TVaR[X,p] \geqslant VaR[X,p]$.

Oczywiście

$$TVaR[X+c,p] = TVaR[X,p] + c.$$

Zachodzi również podaddytywność

$$TVaR[X+Y,p] \leqslant TVaR[X,p] + TVaR[Y,p].$$

Związek między oceną ryzyka X poprzez funkcję użyteczności u oraz miarą $VaR[X,p]$ dany jest poprzez

$$\mathrm{E}\left[u(X)\right] = \int_0^1 u(VaR[X,p])dp,$$

bo zmienna losowa $VaR[X,U] = F_X^{-1}(U) = \tilde{X}$ uzyskana przez podstawienie zmiennej losowej U o rozkładzie jednostajnym na $(0,1)$ w miejsce argumentu p jest zmienną o tym samym rozkładzie co zmienna X, stąd $\mathrm{E}\left[u(X)\right] = \mathrm{E}\left[u(\tilde{X})\right]$ i zachodzi powyższa równość.

Istnieje też możliwość oceny ryzyka przez tak zwaną funkcję *zniekształcającą* $g : [0,1] \to [0,1]$ o własnościach $g(0) = 0$, $g(1) = 1$ i g jest niemalejąca. Używa się jej do obliczenia *zniekształconej* (distorted) wartości oczekiwanej

$$H_g[X] = \int_0^\infty g(1 - F_X(x))dx - \int_{-\infty}^0 1 - g(1 - F_X(x))dx.$$

Gdy $g(x) := x$, to $H_g[X] = \mathrm{E}\,[X]$.

Jeśli $X \geqslant 0$, to

$$\begin{aligned} H_g[X] = \int_0^\infty g(1 - F_X(x))dx = \\ &= \int_0^\infty \int_0^{1-F_X(x)} dg(p)dx = \\ &= \int_0^\infty \int_0^1 I_{[0,1-F_X(x)]}(p)dg(p)dx = \\ &= \int_0^1 \int_0^\infty I_{[0,F^{-1}(1-p)]}(x)dxdg(p) = \\ &= \int_0^1 VaR[X, 1-p]dg(p). \end{aligned}$$

Dla $g(x) = I_{[1-p,1]}(x)$ mamy $H_g[X] = VaR[X, p]$.

Jeśli $g(x) = 1-(1-x)^k$, to $H_g[X] = \mathrm{E}\,[\max\{X_1^*, \ldots, X_k^*\}]$ dla niezależnych X_i^* o rozkładzie F_X.

Jeśli $g(x) = \min\{x/(1-p), 1\}$, to $H_g[X] = TVaR[X, p]$.

Wartość $H_g[X]$ dla $X \geqslant 0$ nazywana jest miarą ryzyka Wanga (the Wang risk measure).

Przykład 3.6.3 Jeśli $X \sim N(\mu, \sigma^2)$, to

$$VaR[X, p] = \mu + \sigma\Phi^{-1}(p),$$

$$TVaR[X, p] = \mu + \sigma\frac{\Phi'(\Phi^{-1}(p))}{1-p}.$$

Dla $X \sim LN(\mu, \sigma^2)$

$$VaR[X, p] = \exp(\mu + \sigma\Phi^{-1}(p)),$$

$$TVaR[X, p] = \exp(\mu + \sigma^2/2)\frac{\Phi(\sigma - \Phi^{-1}(p))}{1-p}.$$

□

3.7 Modelowanie zależności przez funkcje copula

Rozpoczniemy od klasycznych rozkładów dwuwymiarowych.

Przykład 3.7.1 (Rozkład Marshalla-Olkina)

Niech $X \sim Exp(\lambda_X)$, $Y \sim Exp(\lambda_Y)$, $Z \sim Exp(\lambda_Z)$ będą niezależnymi zmiennymi losowymi. Niech $X_1 = \min(X, Z)$, $Y_1 = \min(Y, Z)$. Wtedy X_1 i Y_1 nie są niezależnymi zmiennymi losowymi. Łączny rozkład (X_1, Y_1) nazywamy rozkładem Marshalla-Olkina. Natychmiast widać, że $X_1 \sim Exp(\lambda_X + \lambda_Z)$ oraz $Y_1 \sim Exp(\lambda_Y + \lambda_Z)$.

$$\begin{aligned} P(X_1 > x, Y_1 > y) &= P(\min(X, Z) > x, \min(Y, Z) > y) \\ &= P(X > x, Y > y, Z > \max(x, y)) \\ &= \exp(-\lambda_X x) \exp(-\lambda_Y y) \exp(-\lambda_Z \max(x, y)) \\ &= \exp(-\lambda_X x) \exp(-\lambda_Y y) \min(\exp(-\lambda_Z x, \exp(-\lambda_Z y)). \end{aligned}$$

Stąd

$$\begin{aligned} F_{(X_1,Y_1)}(x, y) &= 1 - P(X_1 > x) - P(Y_1 > y) + P(X_1 > x, Y_1 > y) = \\ &= 1 - \exp(-(\lambda_X + \lambda_Z)x) - \exp(-(\lambda_Y + \lambda_Z)y) \\ &+ \exp(-\lambda_X x) \exp(-\lambda_Y y) \min(\exp(-\lambda_Z x), \exp(-\lambda_Z y)). \end{aligned}$$

□

Przykład 3.7.2 (Rozkład Pareto)

Losujemy parametr θ zgodnie z rozkładem $Gamma(\alpha, \lambda)$. Następnie losujemy niezależnie zmienne X, Y o rozkładzie wykładniczym $Exp(\theta)$. Mamy więc

$$P(X > x, Y > y | \Theta = \theta) = \exp(-\theta x) \exp(-\theta y).$$

Licząc rozkład bezwarunkowy otrzymujemy

$$\begin{aligned} P(X > x, Y > y) &= \int_0^\infty \exp(-\theta x) \exp(-\theta y) \exp(-\lambda\theta) \theta^{\alpha-1} \lambda^\alpha / \Gamma(\alpha) d\theta = \\ &= (x/\lambda + y/\lambda + 1)^{-\alpha}. \end{aligned}$$

Stąd

$$F_{(X,Y)}(x, y) = 1 - (1 + x/\lambda)^{-\alpha} - (1 + y/\lambda)^{-\alpha} + (x/\lambda + y/\lambda + 1)^{-\alpha}.$$

□

Dystrybuanty wielowymiarowe o brzegowych rozkładach jednostajnych na $(0, 1)$ nazywamy *funkcjami copula*. Przypomnijmy, że $C(x, y)$ jest dystrybuantą dwuwymiarową jeśli

- $C : \mathbb{R}^2 \to [0,1]$, jest niemalejąca po współrzędnych,
- $C(-\infty,-\infty) = 0$, $C(\infty,\infty) = 1$ (przy iterowanym przejściu do granicy w dowolnej kolejności)
- C jest supermodularna, tzn.

$$C(\min(\mathbf{x},\mathbf{y})) + C(\max(\mathbf{x},\mathbf{y})) \geqslant C(\mathbf{x}) + C(\mathbf{y}),$$

dla dowolnych $\mathbf{x},\mathbf{y} \in \mathbb{R}^2$ (minimum i maksimum brane są po współrzędnych).

W przypadku funkcji copula zakładamy, że $C(x,\infty) = x,\ x \in (0,1)$ i $C(\infty,y) = y,\ y \in (0,1)$, tzn. dystrybuanty brzegowe są jednostajne na $(0,1)$.

Twierdzenie 3.7.3 *(Sklar)*

Jeśli $F_{(X,Y)}$ jest dystrybuantą zmiennych losowych (X,Y) o dystrybuantach brzegowych F_X i F_Y ciągłych, to istnieje wyznaczona jednoznacznie funkcja copula $C_{(X,Y)}$ taka, że

$$F_{(X,Y)}(x,y) = C_{(X,Y)}(F_X(x), F_Y(y)).$$

Dowód. Przyjmijmy

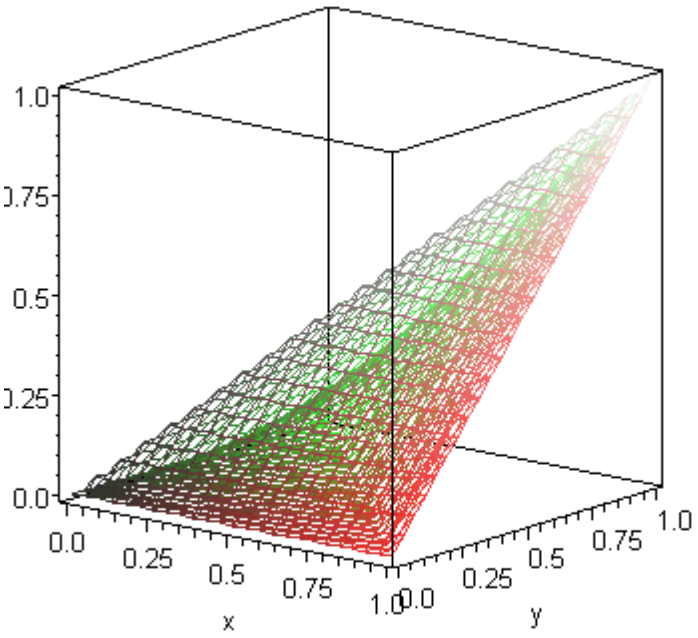
$$C_{(X,Y)}(x,y) := P(F_X(X) \leqslant x, F_Y(Y) \leqslant y).$$

Wtedy $C_{(X,Y)}(x,y) = F_{(X,Y)}(F^{-1}(x), F^{-1}(y))$ i przez odpowiednie podstawienie otrzymujemy tezę. □

Przyjmuje się, że funkcja C jest odzwierciedleniem sposobu w jaki współrzędne X i Y wektora losowego (X,Y) są stochastycznie zależne. Warto podkreślić, że zależność opisywana przez $C_{(X,Y)}$ nie musi odpowiadać intuicjom związanym z wielkością współczynnika korelacji.

Następujące funkcje copula zasługują na szczególną uwagę:

- $C^*(x,y) = \min(1,x)\min(1,y)I_{(0,\infty)\times(0,\infty)}(x,y)$,
  rozkład jednostajny na $(0,1)\times(0,1)$, który odpowiada niezależnym X i Y,
- $C^+(x,y) = \min(\min(1,x),\min(1,y))I_{(0,\infty)\times(0,\infty)}(x,y)$,
  rozkład jednostajny na przekątnej $(0,1)\times(0,1)$, który odpowiada *maksymalnie dodatnio* zależnym zmiennym X i Y,
- $C^-(x,y) = \max(0,\min(1,x)+\min(1,y)-1)I_{(0,\infty)\times(0,\infty)}(x,y)$,
  rozkład jednostajny na przeciwprzekątnej $(0,1)\times(0,1)$, który odpowiada *maksymalnie ujemnie* zależnym zmiennym X i Y.



Rysunek 3.7.1: Wykresy C^*, C^+, C^-.

Znaczenie C^+ oraz C^- wynika z faktu, iż są one ograniczeniami dla dowolnych funkcji copula.

Twierdzenie 3.7.4 *(Frechet)*
Dla dowolnej funkcji copula C zachodzą nierówności

$$C^-(x,y) \leqslant C(x,y) \leqslant C^+(x,y), \ \forall x,y \in \mathbb{R}.$$

Przykład 3.7.5 Funkcja Calytona.

$$C^{Clay(\alpha)}(x,y) = (x^{-\alpha} + y^{-\alpha} - 1)^{-1/\alpha}, \ \alpha > 0, \ x,y \in (0,1).$$

Funkcja Franka.

$$C^{F(\alpha)}(x,y) = -\frac{1}{\alpha}\ln(1 + \frac{(1-\exp(-\alpha x))(1-\exp(-\alpha y))}{\exp(-\alpha)-1}), \alpha \neq 0.$$

Funkcja Archimedesa.

$$C^{A(\phi)}(x,y) = \phi^{-1}(\phi(x)+\phi(y))I_{\{\phi(x)+\phi(y)\leqslant\phi(0)\}}(x,y), \ \phi:[0,1]\to\mathbb{R}_+, \ \phi(1)=0,$$

gdzie ϕ jest funkcją ściśle malejącą i ściśle wypukłą.
Gdy $\phi(t) = (t^{-\alpha}-1)/\alpha$, to $C^{A(\phi)} = C^{Clay(\alpha)}$ jest funkcją Claytona. Gdy $\phi(t) = -\ln(\frac{1-\exp(-\alpha t)}{1-\exp(-\alpha)})$, to $C^{A(\phi)} = C^{F(\alpha)}$ jest funkcją Franka. Gdy $\phi(t) = -\ln(t)$, to $C^{A(\phi)} = C^*$.

□

Przykład 3.7.6 Rozważmy dwa ryzyka o rozkładach logarytmicznie normalnych $X \sim LN(0,\sigma^2)$, $Y \sim LN(0,1)$. Stosując funkcję C^+ możemy zdefiniować łączny rozkład przez

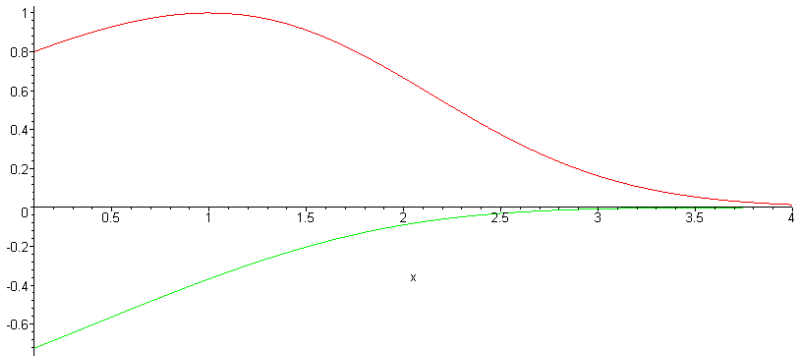
$$F_{(X^+,Y^+)}(x,y) = C^+(F_X(x), F_Y(y)),$$

który odpowiada zmiennym X^+ i Y^+ *współmonotonicznym* (comonotone). Podobnie definiujemy rozkład

$$F_{(X^-,Y^-)}(x,y) = C^-(F_X(x), F_Y(y)),$$

który odpowiada X^-, Y^- *przeciwmonotonicznym* (countermonotone). Współczynnniki korelacji dla tych dystrybuant dane są odpowiednio przez

$$Corr(X^+,Y^+) = \frac{e^\sigma - 1}{\sqrt{(e-1)(e^{\sigma^2}-1)}},$$



Rysunek 3.7.2: Wykresy funkcji korelacji w rozkładach współmonotonicznym i przeciwmonotonicznym o brzegowych log-normalnych $LN(0,1)$ i $LN(0,\sigma^2)$.

$$Corr(X^-,Y^-)=\frac{e^{-\sigma}-1}{\sqrt{(e-1)(e^{\sigma^2}-1)}}.$$

Wykresy funkcji korelacji w zależności od σ są na rysunku 3.7.2. Jak widać w obu przypadkach, przy wartości σ dostatecznie dużej korelacja jest bliska zeru, mimo, że zmienne są strukturalnie maksymalnie zależne.

□

Naturalną relacją między parami ryzyk o jednakowych rozkładach brzegowych, $F_X = F_{X'}$, $F_Y = F_{Y'}$, jest (concordance)

$$(X,Y) <_{cc} (X',Y') \Leftrightarrow F_{(X,Y}(x,y) \leqslant F_{(X',Y'}(x,y)\ \forall x,y \in \mathbb{R}.$$

Ponieważ rozkłady brzegowe są równe, więc relacja ta jest równoważna do porównywania odpowiednich funkcji copula:

$$(X,Y) <_{cc} (X',Y') \Leftrightarrow C_{(X,Y)} \leqslant C_{(X',Y')},$$

a stąd mamy dla dowolnych zmiennych losowych (X, Y)

$$F_{(X^-,Y^-)} \leqslant F_{(X,Y)} \leqslant F_{(X^+,Y^+)}.$$

Ze wzorów

$$F_{(X^+,Y^+)}(x, y) = \min(F_X(x), F_Y(y)),$$
$$F_{(X^-,Y^-)}(x, y) = \max(0, F_X(x) + F_Y(y) - 1),$$

licząc dystrybuanty z definicji, widzimy, że zmienne (X^+, Y^+) możemy zadać przez

$$(X^+, Y^+) := (F_X^{-1}(U), F_Y^{-1}(U)),$$
$$(X^-, Y^-) := (F_X^{-1}(U), F_Y^{-1}(1 - U)),$$

dla danej zmiennej U o rozkładzie jednostajnym na $(0, 1)$.

Relację $<_{cc}$ można opisać przy pomocy wartości oczekiwanych

$$(X, Y) <_{cc} (X', Y') \Leftrightarrow \mathrm{E}\left[f(X)g(Y)\right] \leqslant \mathrm{E}\left[f(X')g(Y')\right],$$

dla wszystkich f, g współmonotonicznych, tzn. dla f, g niemalejących, lub dla f, g nierosnących.

Rzeczywiście, gdy $f(x) = I_{(-\infty,a)}(x)$, $g(y) = I_{(-\infty,b)}(y)$, $a, b \in \mathbb{R}$, otrzymujemy nierówność dla dystrybuant. Każdą niemalejącą funkcję f można aproksymować monotonicznie przez kombinacje liniowe funkcji indykatorowych, stąd nierówność jest prawdziwa dla dowolnych f, g.

Mamy więc w szczególności

$$(X, Y) <_{cc} (X', Y') \Rightarrow Corr(X, Y) \leqslant Corr(X', Y'),$$

a stąd

$$Corr(X^-, Y^-) \leqslant Corr(X, Y) \leqslant Corr(X^+, Y^+).$$

Przykład 3.7.7 Rozważmy znowu dwa dowolne ryzyka o rozkładach logarytmicznie normalnych $X \sim LN(0, \sigma^2)$, $Y \sim LN(0, 1)$ o których nie mamy informacji o łącznej dystrybuancie. Stosując wzory z poprzedniego przykładu i powyższą nierówność dla korelacji otrzymujemy na przyklad dla $\sigma = 4$,

$$-0.00025 \leqslant Corr(X, Y) \leqslant 0.01372.$$

Oznacza to, że takie ryzyka (X, Y), niezależnie od sposobu zależności między nimi będą skorelowane na poziomie bliskim zero. Widać, że **w tym przypadku korelacja nie jest dobrą miarą zależności**.

□

Rozdział 4

Prawdopodobieństwo ruiny: czas dyskretny

Rozważmy następujący proces:

$$R_n = u + cn - S_n,\ n = 0, 1, \ldots, \tag{4.0.1}$$

gdzie $u \geqslant 0$ jest kapitałem początkowym towarzystwa ubezpieczeniowego, c - składką otrzymaną w ciągu jednego okresu, a $S_n = W_1 + \cdots + W_n$ - sumą szkód wypłaconych do chwili n. Zmienne losowe $(W_i)_{i\geqslant 1}$ są wypłatami w i-tym okresie. Proces $(R_n)_{n\geqslant 1}$ nazywamy **procesem nadwyżki ubezpieczyciela** lub **procesem ryzyka**.

Interesować nas będzie zmienna losowa T postaci

$$T = \min\{n : R_n < 0\},$$

(zmienna ta nazywana jest **momentem technicznej ruiny**), jak również

$$\psi(u) = P(T < \infty),$$

czyli **prawdopodobieństwo ruiny**, jeżeli kapitał początkowy wynosił u.

4.1 Proces ryzyka jako błądzenie losowe- prawdopodobieństwo ruiny

Ciąg R_n możemy zapisać jako błądzenie losowe startujące z poziomu u

$$R_n = u + (c - W_1) + \cdots + (c - W_n),$$

oraz

$$\begin{aligned}\psi(u) &= P(\bigcup_{i\geqslant 1}\{R_i < 0\}) \\ &= P(\bigcup_{i\geqslant 1}\{S_i - ci > u\}) \\ &= P(\max(S_i - ci : i \geqslant 1) > u).\end{aligned}$$

Prawdopodobieństwo ruiny jest więc równe prawdopodobieństwu, że maksymalna wartość pewnego błądzenia losowego przekroczy poziom u. Oznaczmy to maksimum przez

$$M = \max(0, W_1 - c, (W_1 - c) + (W_2 - c), \ldots),$$

gdzie błądzenie losowe ma przyrosty postaci $W_i - c$, stąd

$$\psi(u) = P(M > u).$$

Powstaje pytanie, kiedy błądzenie losowe osiąga skończone maksimum? Intuicja wskazuje na błądzenia, które dryfują "do dołu"i w związku z tym osiągają skończony pułap "do góry". Rzeczywiście zachodzi

Lemat 4.1.1 $P(M < \infty) = 1 \Leftrightarrow \mathrm{E}\,[W_i - c] < 0.$

Dowód. Z mocnego prawa wielkich liczb

$$P(\lim_{n\to\infty} \frac{\sum_{i=1}^{n}(W_i - c)}{n} = \mathrm{E}\,[W_i - c]) = 1,$$

stąd, jeśli $\mathrm{E}\,[W_i - c] < 0$, to $P(\lim_{n\to\infty}\sum_{i=1}^{n}(W_i - c) = -\infty) = 1$, czyli trajektorie błądzenia z prawdopodobieństwem 1 dążą do $-\infty$ osiągając w związku z tym $M < \infty$. Z drugiej strony, jeśli $M < \infty$, to $\lim_{n\to\infty}\frac{\sum_{i=1}^{n}(W_i - c)}{n} \leqslant 0$, czyli $\mathrm{E}\,[W_i - c] \leqslant 0$. Ponieważ w symetrycznym błądzeniu losowym trajektorie z prawdopodobieństwem 1 powracają do punktu wyjścia nieskończenie wiele razy, warunek $M < \infty$ nie może zachodzić z prawdopodobieństwem 1, stąd $\mathrm{E}\,[W_i - c] = 0$ jest wykluczone. Oznacza to, że $M < \infty$ z prawdopodobieństwem 1 pociąga $\mathrm{E}\,[W_i - c] < 0$. □

Okazuje się, że rozkład maksimum M nie zmieni się jeśli dodamy jeden *extra* przyrost do M i wyrównamy do zera.

Lemat 4.1.2 *Rozkład zmiennej M jest taki sam jak zmiennej $max(0, M + (W - c))$, gdzie W jest zmienną niezależną od całego błądzenia $\{W_i - c, i \geqslant 1\}$, ale posiadającą rozkład równy rozkładowi W_i, pisząc krótko*

$$M =^{d} (M + W - c)_{+}.$$

Dowód. Możemy M zapisać w postaci

$$M = \max(0, W_1 - c + L),$$

gdzie

$$L = \max(0, W_2 - c, (W_2 - c) + (W_3 - c), (W_2 - c) + (W_3 - c) + (W_4 - c), \ldots).$$

Jeśli teraz w powyższych wzorach zastąpimy ciąg $W_1, W_2, W_3, \ldots$ ciągiem $W, W_1, W_2, \ldots$, to rozkład uzyskanych zmiennych nie zmieni się, bo wyjściowy i *nowy* ciąg są o tym samym rozkładzie. To znaczy, że rozkład M jest taki sam jak rozkład $\max(0, W - c + \max(0, W_1 - c, (W_2 - c) + (W_3 - c), \ldots)) = (W - c + M)_+$. □

Używając powyższy lemat uzyskujemy rozkład M w języku funkcji tworzących w przypadku $c = 1$.

Twierdzenie 4.1.3 *Załóżmy, że $W_i \in \mathbb{N}$ mają funkcję tworzącą P_W. Jeśli $\mathrm{E}[W_i] < 1$ to zmienna M ma funkcję tworzacą*

$$P_M(t) = \frac{1 - \mathrm{E}[W_i]}{1 - \mathrm{E}[W_i] P_{\tilde{W}}(t)},$$

gdzie

$$P_{\tilde{W}}(t) = \frac{1 - P_{W_i}(t)}{\mathrm{E}[W_i](1 - t)}.$$

Dowód. Liczymy funkcję tworzacą

$$\begin{aligned}
P_M(t) &= \mathrm{E}\left[t^M\right] = \mathrm{E}\left[t^{(M+W-1)_+}\right] \\
&= \mathrm{E}\left[t^{(M+W-1)_+} I_{\{M+W-1 \geqslant 0\}}\right] + \mathrm{E}\left[t^{(M+W-1)_+} I_{\{M+W-1<0\}}\right] \\
&= P(M + W = 0) \\
&+ P(M + W - 1 = 0) + P(M + W - 1 = 1)t + P(M + W - 1 = 2)t^2 + \cdots \\
&= P(M + W = 0) + \frac{1}{t}(P(M + W = 1)t + P(M + W = 2)t^2 + \cdots) \\
&= P(M + W = 0) - \frac{1}{t}P(M + W = 0) + \frac{1}{t}P_{M+W}(t) \\
&= P(M + W = 0)(1 - \frac{1}{t}) + \frac{1}{t}P_M(t)P_W(t).
\end{aligned}$$

Wyliczamy stąd

$$P_M(t) = \frac{(t-1)P(M + W = 0)}{t - P_W(t)}.$$

Przechodząc z $t \to 1$ w powyższej równości otrzymujemy

$$1 = \frac{P(M + W = 0)}{1 - \mathrm{E}[W]},$$

co daje $P(M+W=0) = 1 - \mathrm{E}[W]$. Ostatecznie więc

$$P_M(t) = \frac{(1-\mathrm{E}[W])(1-t)}{P_W(t)-t},$$

co po przekształceniu możemy zapisać jako

$$P_M(t) = \frac{1-\mathrm{E}[W]}{1-\mathrm{E}[W]\,P_{\tilde{W}}(t)}.$$

□

Funkcja tworząca $P_{\tilde{W}}$ odpowiada zmiennej losowej o rozkładzie

$$P(\tilde{W}=k) = \frac{P(W>k)}{\mathrm{E}[W]},$$

który nazywamy rozkładem resztowym rozkladu zmiennej W.

Zmienna M ma rozkład złożony geometryczny $CGeo(p = 1-\mathrm{E}[W], F_{\tilde{W}})$. Tak więc widzimy, że **prawdopodobieństwo ruiny, przy wprowadzonych założeniach, jest równe ogonowi dystrybuanty zmiennej losowej o złożonym rozkładzie geometrycznym**. Okaże się, że taka struktura jest również prawdziwa w modelu z czasem ciągłym.

4.1.1 Współczynnik dopasowania

Załóżmy, że zmienne losowe $(W_i)_{i\geqslant 1}$ mają ten sam rozkład i funkcję tworzącą momenty $M_W(t)$ oraz $E[W] < c$. Definiujemy **współczynnik dopasowania** $R(W,c)$ jako dodatnie rozwiązanie równania $M_{W-c}(r) = 1$, co jest równoważne

$$\exp(-cr)M_W(r) = 1. \tag{4.1.1}$$

Zauważmy, że

$$\frac{d}{dr}M_{W-c}(r) = \mathrm{E}\left[(W-c)e^{r(W-c)}\right],$$

$$\frac{d^2}{dr^2}M_{W-c}(r) = \mathrm{E}\left[(W-c)^2e^{r(W-c)}\right] \geqslant 0.$$

Co więcej

$$\frac{d}{dr}M_{W-c}(r) = \mathrm{E}\left[(W-c)e^{r(W-c)}\right] = \mathrm{E}[W]-c < 0$$

i $M_{W-c}(0) = 1$. Oznacza to, że funkcja $M_{W-c}(r)$ maleje w otoczeniu 0 i jest wypukła, co daje istnienie współczynnika dopasowania.

Przykład 4.1.4 Załóżmy, że $W \sim N(\mu, \sigma^2)$. Wtedy $M_W(r) = \exp(\mu r + \sigma^2 r^2/2)$ i stąd $R(N(\mu, \sigma^2), c) = \frac{2(c-\mu)}{\sigma^2}$. Jeżeli składka za jeden okres naliczana jest według zasady wartości oczekiwanej, to $c = (1+\theta)\mathrm{E}[W]$, a stąd $R(N(\mu, \sigma^2), c) = \frac{2\theta\mu}{\sigma^2}$.

□

Zauważmy, że założenie o normalności zmiennej losowej W ma sens tylko wtedy, gdy do łącznych wypłat stosujemy aproksymację normalną (szkody nie mogą być ujemne).

Przykład 4.1.5 Załóżmy, że W przyjmuje dwie wartości: $P(W = a) = p = 1 - P(W = b)$. Wtedy $M_W(r) = p\exp(ra) + (1-p)\exp(rb)$. Współczynnik dopasowania wylicza się więc ze wzoru

$$\exp(cr) = p\exp(ra) + (1-p)\exp(rb)$$

Jeżeli $a = 2$, $b = 0$, $p < \frac{1}{2}$, $c = 1$ powyższe równanie staje się równaniem kwadratowym i otrzymujemy $R(W, c) = -\log\left(\frac{p}{1-p}\right)$.

□

Powyższy przykład pokazuje, że rozwiązanie równania (4.1.1) rzadko da się przedstawić w postaci jawnej. Współczynnik R można jednak przybliżyć stosując podobne rozumowanie jak w przypadku aproksymacji Edgewortha. Rozwijając $\log M_W(r)$ w szereg Taylora i uwzględniając dwa (trzy) pierwsze składniki dostajemy równanie (4.1.1) w następującej postaci:

$$\begin{aligned} &r\mathrm{E}[W] + \frac{1}{2}r^2\mathrm{Var}[W] - cr = 0, \\ &(r\mathrm{E}[W] + \frac{1}{2}r^2\mathrm{Var}[W] + \frac{1}{6}r^3\mathrm{E}\left[(W - \mathrm{E}[W])^3,\right] - cr = 0), \end{aligned} \tag{4.1.2}$$

co w pierwszym przypadku daje przybliżony współczynnik dopasowania

$$R(W, c) \approx \frac{2(c - \mathrm{E}[W])}{\mathrm{Var}[W]}. \tag{4.1.3}$$

Dla W będącego rozkładem złożonym $W = \sum_{i=1}^{N} X_i$ dostajemy więc

$$R(\sum_{i=1}^{N} X_i, c) \approx \frac{2(c - \mathrm{E}[N]\mathrm{E}[X])}{\mathrm{E}[N]\mathrm{Var}[X] + (\mathrm{E}[X])^2\mathrm{Var}[N]}.$$

Przykład 4.1.6 Załóżmy, że $c = (1+\theta)\mathrm{E}[W] = (1+\theta)\mathrm{E}[N]\mathrm{E}[X]$.

1. Jeżeli $W \sim CPoi(\lambda, F_X)$, to

$$R(CPoi(\lambda, F_X), c) \approx \frac{2\theta\mathrm{E}[X]}{\mathrm{E}[X^2]}.$$

2. Jeżeli $W \sim CBin(n, p, F_X)$, to

$$R(CBin(n, p, F_X), c) \approx \frac{2\theta \mathrm{E}[X]}{\mathrm{Var}[X] + q(\mathrm{E}[X])^2}.$$

3. Jeżeli $W \sim CBin^{-}(r, p, F_X)$, to

$$R(CBin^{-}(r, p, F_X), c) \approx \frac{2\theta \mathrm{E}[X]}{\mathrm{E}[X^2] + (\mathrm{E}[X])^2(1/p - 1)}.$$

Zauważmy, że jeżeli $p \to 1$ to współczynnik dopasowania $R(CBin^{-}(r, p, F_X), c)$ jest równy analogicznej wielkości dla rozkładu $CPoi(\lambda, F_X)$. Nie jest to zaskakujący, gdyż rozkład Poissona można w tym przypadku traktować jako graniczny dla rozkładu ujemnego dwumianowego.

□

Podobnie jak w przypadku rozwinięcia Edgewortha wzór (4.1.3) daje złe przybliżenie jeżeli zmienna W ma dużą skośność.

4.1.2 Prawdopodobieństwo ruiny - lekkie ogony

Twierdzenie 4.1.7 *Załóżmy, że w procesie ryzyka (4.0.1)*

- *zmienne losowe $(W_i)_{i \geqslant 1}$ są niezależne i o tym samym rozkładzie,*
- *funkcja tworząca momenty $M_W(t)$ istnieje i jest skończona,*
- $\mathrm{E}[W] < c$.

Wtedy dla $R := R(W, c)$

$$\psi(u) = \exp(-Ru) \frac{1}{\mathrm{E}[\exp(-RR_T | T < \infty)]}. \tag{4.1.4}$$

Dowód: Zauważmy najpierw, że ciąg $(R_n)_{n \geqslant 1}$ ma przyrosty niezależn, tzn. zmienne losowe $R_{n_1} - R_0, R_{n_2} - R_{n_1}, R_{n_3} - R_{n_2}, \ldots$ są niezależne dla dowolnych $0 \leqslant n_1 < n_2 < \cdots$. Poza tym dla dowolnych $n > i$ mamy z niezależności

$$\begin{aligned} \mathrm{E}[\exp(-R(R_n - R_i))] &= \mathrm{E}[\exp(-R(c - W_{i+1}) - \cdots - R(c - W_n))] = \\ &= \mathrm{E}[\exp(-R(c - W))]^{n-i} = 1 \end{aligned}$$

i w szczególności dla $i = 0$ mamy

$$\mathrm{E}\left[\exp(-R(R_n)\right] = \exp(-Ru). \tag{4.1.5}$$

Stąd

$$\begin{aligned}
\exp(-Ru) &= \mathrm{E}\left[\exp(-RR_n)\right] \\
&= \sum_{i=1}^{n} \mathrm{E}\left[\exp(-RR_n|T=i)\right] P(T=i) + \mathrm{E}\left[\exp(-RR_n|T>n)\right] P(T>n) \\
&= \sum_{i=1}^{n} \mathrm{E}\left[\exp(-RR_i - R(R_n - R_i)|T=i)\right] P(T=i) \\
&\quad + \mathrm{E}\left[\exp(-RR_n|T>n)\right] P(T>n) \\
&= \sum_{i=1}^{n} \mathrm{E}\left[\exp(-RR_i|T=i)\right] P(T=i) + \mathrm{E}\left[\exp(-RR_n|T>n)\right] P(T>n).
\end{aligned}$$

Teraz, ostatni składnik w powyższym równaniu dąży do 0 przy $n \to \infty$, a stąd

$$\exp(-Ru) = \sum_{i=1}^{\infty} \mathrm{E}\left[\exp(-RR_i|T=i)\right] P(T=i) = \mathrm{E}\left[\exp(-RU_T|T<\infty)\right].$$

□

Przykład 4.1.8 (cd. Przykładu 5.5.1 z $a = 1$ i $b = 0$) Zauważmy, że $R_T = -1$ z prawdopodobieństwem 1. Stąd

$$\psi(u) = \exp(-R(u+1)) = \left(\frac{p}{1-p}\right)^{u+1}.$$

□

Wyliczenie wartości znajdującej się w mianowniku (4.1.4) jest zazwyczaj trudne i możliwe jedynie tylko w kilku przypadkach. Z Twierdzenia 4.1.7 otrzymujemy jednak górne oszacowanie na prawdopodobieństwo ruiny: **Nierówność Cramera**

$$\psi(u) \leqslant \exp(-Ru). \tag{4.1.6}$$

Rozdział 5

*Prawdopodobieństwo ruiny: czas ciągły

Podobnie do dyskretnego procesu ryzyka (4.0.1) rozważa się model w czasie ciągłym:

$$R(t) = u + ct - \sum_{i=1}^{N(t)} X_i,$$

gdzie $(X_i)_{i\geqslant 1}$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie, a $(N(t), t \geqslant 0)$ jest procesem opisującym ilość szkód zgłoszonych do chwili t. Jeżli więc szkody zgłaszane są w losowych chwilach $0 = T_0 < T_1 < \cdots$, to

$$N(t) := \max\{n : T_n \leqslant t\} = \sum_{n=1}^{\infty} I_{\{T_n \leqslant\}}.$$

zajmiemy się najpierw własnosciami procesu $(N(t), t > 0)$.

5.1 Proces zgłoszeń - teoria odnowy

Jeśli $U_i = T_i - T_{i-1}$ dla $i = 1, 2, \ldots$ są niezależnymi zmiennymi losowymi o jednakowych rozkładach, to proces $(N(t), t > 0)$ jest **procesem odnowy**. Zachodzą podstawowe związki zdarzeń opisujących ten proces.

- $\{N(t) = 0\} = \{T_1 > t\}$,
- $\{N(t) = k\} = \{T_k \leqslant t < T_{k+1}\}$,
- $\{N(t) \geqslant k\} = \{T_k \leqslant t\}$,
- $\{N(t) < k\} = \{T_k > t\}$.

Stąd $P(N(t)=0) = 1 - F_U(t)$, oraz $P(N(t) \geqslant k) = F_U^{*k}(t)$, gdzie F_U jest dystrybuantą odstępów U_i.

Twierdzenie 5.1.1 *Jeśli* $\mathrm{E}\,[U_i] > 0$, *to*

$$P(\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mathrm{E}\,[U]}) = 1.$$

Dowód: Korzystając z podstawowych zależności, mamy

$$\frac{T_{N(t)}}{N(t)} \leqslant \frac{t}{N(t)} \leqslant \frac{T_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)}.$$

z mocnego prawa wielkich liczb otrzymujemy z prawdopodobieństwem 1

$$\lim_{n\to\infty} \frac{T_n}{n} = \mathrm{E}\,[U]\,,$$

stąd teza, gdyż z prawdopodobieństwem 1, $\lim_{t\to\infty} N(t) = \infty$.

□

Wielkość $N(t)$ można postrzegać jako efekt sumaryczny wielu zmiennych losowych, gdy t rośnie do ∞. Należy więc oczekiwać, że zachodzi twierdzenie graniczne podobne do CTG. Rzeczywiście:

Twierdzenie 5.1.2 *Jeśli* $0 < \mathrm{Var}\,[U_i] < \infty$, *to*

$$P(\frac{N(t) - t/\mathrm{E}\,[U]}{\sqrt{(\mathrm{Var}\,[U]\,t/(\mathrm{E}\,[U])^3)}} \leqslant x) \to_{t\to\infty} \Phi(x),$$

gdzie Φ *oznacza dystrybuantę standardowego rozkładu normalnego.*

Dowód:

$$\begin{aligned} P(\frac{N(t) - t/\mathrm{E}\,[U]}{\sqrt{(\mathrm{Var}\,[U]\,t/(\mathrm{E}\,[U])^3)}} \leqslant x) &= P(N(t) \leqslant x\sqrt{(\mathrm{Var}\,[U]\,t/(\mathrm{E}\,[U])^3)} + t/\mathrm{E}\,[U]) \\ &= P(T_{\lfloor x\sqrt{(\mathrm{Var}[U]t/(\mathrm{E}[U])^3)}+t/\mathrm{E}[U]\rfloor+1} > t) \\ &= P(\frac{U_1 + \cdot + U_K - K\mathrm{E}\,[U]}{\sqrt{(K\mathrm{Var}\,[U])}} \geqslant \frac{t - K\mathrm{E}\,[U]}{\sqrt{(K\mathrm{Var}\,[U])}}, \end{aligned}$$

gdzie $K = \lfloor x\sqrt{(\mathrm{Var}\,[U]\,t/(\mathrm{E}\,[U])^3)} + t/\mathrm{E}\,[U]\rfloor + 1$. Ponieważ

$$\frac{t - K\mathrm{E}\,[U]}{\sqrt{(K\mathrm{Var}\,[U])}} \to_{t\to\infty} -x,$$

z CTG dla ciągu (U_i) i równości $1 - \Phi(-x) = \Phi(x)$, otrzymujemy tezę. □

Wartość oczekiwaną liczby zgłoszeń do chwili t nazywamy **funkcją odnowy** i oznaczamy

$$H(t) := \mathrm{E}\left[N(t)\right],\ t > 0.$$

Otrzymujemy natychmiast z podstawowych zależności między zdarzeniami
$H(t) = \sum_{k=0}^{\infty} P(N(t) > k) = \sum_{k=0}^{\infty} P(N(t) \geqslant k+1) =$

$$\sum_{k=0}^{\infty} P(T_{k+1} \leqslant t) = \sum_{k=1}^{\infty} F_U^{*k}(t). \tag{5.1.1}$$

Z twierdzenia 5.1.1 wiemy, że z prawdopodobieństwem 1 uśredniona w czasie liczba zgłoszeń jest zbieżna do odwrotności wartości oczekiwanej odstępów między zgłoszeniami. Można się spodziewać, że wartość oczekiwana liczby zgłoszeń uśredniona w czasie będzie zbieżna do tej samej granicy. Rzeczywiście tak jest, ale wymaga to technicznie dodatkowej uwagi, ponieważ zbieżność z prawdopodobieństwem 1 ciągu zmiennych losowych nie implikuje - ogólnie rzecz biorąc - zbieżności wartości oczekiwanych tego ciągu.

Twierdzenie 5.1.3 *Jeśli* $\mathrm{E}\left[U_1\right] < \infty$, *to*

$$\lim_{t\to\infty} \frac{H(t)}{t} = \frac{1}{\mathrm{E}\left[U\right]}.$$

Dowód: Z lematu Fatou otrzymujemy

$$\frac{1}{\mathrm{E}\left[U\right]} = \mathrm{E}\left[\lim_{t\to\infty} \frac{N(t)}{t}\right] \leqslant \liminf_{t\to\infty} \frac{H(t)}{t}.$$

Wystarczy więc pokazać, że

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leqslant \frac{1}{\mathrm{E}\left[U\right]}.$$

W tym celu pokażemy najpierw tę nierówność dla obciętych odstępów

$$U_n^K := \min(K, U_n),\ K > 0, n \in \mathbb{N}.$$

Ze względu na obcięcie mamy

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leqslant \limsup_{t\to\infty} \frac{H^K(t)}{t},$$

gdzie H^K jest funkcją odnowy dla obciętych odstępów. Aby oszacować wartość $H^K(t)$, pokażemy, że

$$H^K(t) + 1 = \mathrm{E}\left[N^K(t) + 1\right] = \frac{\mathrm{E}\left[U_1 + \cdots + U_{N^K(t)+1}\right]}{\mathrm{E}\left[U_1^K\right]}.$$

Rzeczywiście

$$\begin{aligned}
\mathrm{E}\left[U_1+\cdots+U_{N^K(t)+1}\right] &= \mathrm{E}\left[\sum_{i=1}^{\infty} U_i^K I_{\{i\leqslant N^K(t)+1\}}\right] \\
&= \sum_{i=1}^{\infty} \mathrm{E}\left[U_i^K I_{\{i\leqslant N^K(t)+1\}}\right] \\
&= \sum_{i=1}^{\infty} \mathrm{E}\left[U_i^K I_{\{T_{i-1}^K\leqslant t\}}\right] \\
&= \sum_{i=1}^{\infty} \mathrm{E}\left[U_i^K\right] P(T_{i-1}^K \leqslant t) \\
&= \mathrm{E}\left[U_1^K\right] \sum_{i=0}^{\infty} P(N^K(t)+1>i) \\
&= \mathrm{E}\left[U_1^K\right] \mathrm{E}\left[N^K(t)+1\right],
\end{aligned}$$

gdzie korzystaliśmy z niezależności T_{i-1}^K od U_i^K. Otrzymujemy więc oszacowanie

$$\limsup_{t\to\infty} \frac{H^K(t)}{t} \leqslant \limsup_{t\to\infty} \frac{\mathrm{E}\left[U_1+\cdots+U_{N^K(t)+1}\right]}{t\mathrm{E}\left[U_1^K\right]}.$$

Ponieważ $\mathrm{E}\left[U_1+\cdots+U_{N^K(t)+1}\right] \leqslant t+K$ mamy

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leqslant \limsup_{t\to\infty} \frac{H^K(t)}{t} \leqslant \limsup_{t\to\infty} \frac{t+K}{t\mathrm{E}\left[U_1^K\right]} = \frac{1}{\mathrm{E}\left[U^K\right]}.$$

Przechodząc z $K\to\infty$ otrzymujemy

$$\limsup_{t\to\infty} \frac{H(t)}{t} \leqslant \frac{1}{\mathrm{E}\left[U\right]}.$$

□

Funkcja odnowy spełnia równanie całkowe (Fredholma)

$$H(t) = F_U(t) + H * F_U(t),\ t>0, \tag{5.1.2}$$

gdzie tutaj operacja splotu $*$ jest zdefiniowana na $(0,\infty)$ dla dowolnych funkcji f,g ograniczonych na przedziałach zwartych o wahaniu ograniczonym, tzn.

$$f*g(t) = \int_0^t f(t-y)dg(t).$$

Przypomnijmy, że $f^{*0}(t) = I_{(0,\infty)}(t)$. Rzeczywiście

$$F_U(t) + H*F_U(t) = F_U(t) + F_U^{*2}(t) + F_U^{*3}(t) + \ldots = H(t),$$

z równania (5.1.1).

Okazuje się, że rozważanie równań tego typu jest bardzo owocne w kontekście badania procesu ryzyka. Zachodzi następujący ogólny lemat.

Lemat 5.1.4 *Niech F będzie być może ułomną dystrybuantą nieujemnej zmiennej losowej, tzn. $F(0) = 0$, F jest niemalejąca i prawostronnie ciągła oraz $F(\infty) \leqslant 1$. Jeśli dla pewnej funkcji $z(t) \geqslant 0$, ograniczonej na przedziałach zwartych zachodzi równanie*

$$Z(t) = z(t) + Z * F(t),$$

to rozwiązaniem tego równania jest

$$Z(t) = z * \hat{H}(t), \ \hat{H}(t) = \sum_{n=0}^{\infty} F^{*n}(t).$$

Dowód: Niech

$$Z_k(t) = z * (\sum_{n=0}^{k} F^{*n}(t),$$

wtedy

$$Z_{k+1}(t) = z(t) + Z_k * F(t).$$

Przy założeniu $z \geqslant 0$ mamy monotoniczność ciągu $Z_k(t)$, więc przechodząc z $k \to \infty$ w powyższej równości otrzymujemy tezę. □

Przykład 5.1.5 Niech $F := F_U$, o gęstości $F'_U = f_U$. Niech $z := f_U$. Wtedy $Z(t) = H'(t)$, tzn, zachodzi

$$H'(t) = f_U(t) + H' * F_U(t).$$

Wynika to natychmiast ze wzoru $(f * g)' = f' * g = f * g'$, dla dowolnych różniczkowalnych splatanych funkcji f, g.

□

Przykład 5.1.6 Niech $F := qG$, dla $q \in (0, 1)$ i właściwej dystrybuanty G. Niech $z(t) \equiv 1 - q = p$. Mamy więc równanie

$$Z = p + Z * qG.$$

Wtedy Z jest dystrybuantą złożonego rozkładu geometrycznego $CGeo(p, G)$, tzn.

$$Z(t) = \sum_{n=0}^{\infty} pq^n G^{*n}(t).$$

□

Własności asymptotyczne rozwiązań $Z(t)$ tego typu równań są zależne od całkowalności funkcji $z(t)$. **Kluczowym twierdzeniem odnowy** jest fakt o istnieniu granicy rozwiązania równania typu odnowy.

Twierdzenie 5.1.7 *Jeśli $z \geqslant 0$ jest nierosnącą i całkowalną oraz F jest właściwą dystrybuantą ciągłą, to rozwiązanie równania z lematu 5.1.4 ma granicę*

$$\lim_{t\to\infty} Z(t) = \frac{\int_0^\infty z(x)dx}{\int_0^\infty 1 - F(x)dx}$$

Twierdzenie to można uogólnić na funkcje bezpośrednio całkowalne w sensie Riemanna.

5.2 Prawdopodobieństwo ruiny: proces zgłoszeń Poissona

Mówimy, że $(N(t), t \geqslant 0)$ jest **procesem Poissona** z parametrem λ jeśli ten proces ma następujące własności:

- $N(0) = 0$;
- Dla $t_0 < t_1 < \cdots < t_k$, przyrosty $N(t_1) - N(t_0), N(t_2) - N(t_1), \ldots, N(t_k) - N(t_{k-1})$ są niezależnymi zmiennymi losowymi o rozkładach Poissona

  $$P(N(t_i) - N(t_{i-1}) = m) = \exp(-\lambda(t_i - t_{i-1}))\frac{(\lambda(t_i - t_{i-1}))^m}{m!}, \quad m \geqslant 0, i = 1, \ldots, k;$$

- dla $h > 0$, $P(N(h) = 1) = \lambda h + o(h)$ oraz $P(N(h) > 1) = o(h)$;
- Odstępy między zgłoszeniami są niezależne i mają rozkład wykładniczy: $P(T_i - T_{i-1} \leqslant x) = 1 - \exp(-\lambda x)$, $i \geqslant 1$.

Niech $(N(t), t \geqslant 0)$ będzie więc procesem Poissona z parametrem λ. Ponadto, niech $(X_i)_{i\geqslant 1}$ będzie ciągiem dodatnich, niezależnych zmiennych losowych o tej samej dystrybuancie $F_X(x) = P(X_1 \leqslant x)$, niezależnym od procesu $(N(t), t \geqslant 0)$. Niech $f_X = F_X'$. **Prawdopodobieństwem ruiny** przy kapitale początkowym u jest

$$\psi(u) = P(T < \infty),\ T := \inf(t > 0 : R(t) < 0).$$

Możemy, więc równoważnie napisać

$$\psi(u) = P\left(\sum_{i=1}^{N(t)} X_i - ct > u \quad \text{dla pewnego } t \geqslant 0\right).$$

Wygodnie jest wprowadzić zmienną $M = \sup(t > 0 : \sum_{i=1}^{N(t)} X_i - ct)$, bo wtedy, analogicznie do modelu w czasie dyskretnym, możemy napisać

$$\psi(u) = P(M > u).$$

Zauważmy, że ruina może nastąpić jedynie w chwili jednego ze zgłoszeń, stąd supremum M wystarczy badać w chwilach zgłoszeń:

$$M = \sup(0, X_1 - cU_1, (X_1 - cU_1) + (X_2 - cU_2), \ldots).$$

Widać, że musimy założyć

$$q := \frac{\lambda \mathrm{E}[X]}{c} < 1, \tag{5.2.1}$$

gdzie $\mathrm{E}[X] = \mathrm{E}[X_1] < \infty$. W przeciwnym razie, M nie będzie skończoną zmienną losową i $\psi(u) = 1$ dla każdego $u > 0$. Przy tym warunku, $\mathrm{E}\left[\sum_{i=1}^{N(t)} X_i\right] = \mathrm{E}[N(t)]\,\mathrm{E}[X_1] = \lambda t \mathrm{E}[X]$, stąd (5.2.1) oznacza, że składka c zebrana w jednostce czasu jest większa niż średnia wartość wypłaconych szkód.

Okazuje się, że funkcja ψ oraz funkcja $\bar{\psi} := 1 - \psi$ spełniają pewne równania typu odnowy, z których będziemy mogli wywnioskować ich podstawowe własności.

Lemat 5.2.1 *Jeśli $q \in (0, 1)$, to*

$$\bar{\psi}(u) = p + q\bar{\psi} * \tilde{F}_X(u),$$

gdzie $\tilde{F}_X(u) = \frac{1}{\mathrm{E}[X]} \int_0^u 1 - F_X(x)dx$, $p = 1 - q$.

Zanim podamy dowód tego faktu, zauważmy, że równanie w tym lemacie ma postać taką jak w przykładzie 5.1.6, a stąd natychmiast widzimy, że $\bar{\psi}$ jako rozwiązanie ma postać

Twierdzenie 5.2.2 ***(wzór Pollaczka-Chinczyna)***

Przy założeniu, że $q = \frac{\lambda \mathrm{E}[X]}{c} < 1$ mamy

$$\bar{\psi}(u) = \sum_{n=0}^{\infty} pq^n \tilde{F}_X^{*n}(t),$$

Czyli prawdopodobieństwo nie zajścia ruiny $\bar{\psi}(u)$ jest jako funkcja kapitału początkowego u dystrybuantą (zmiennej M) złożonego rozkładu geometrycznego $CGeo(p, \tilde{F}_X)$. Jest to sytuacja analogiczna do modelu w czasie dyskretnym, ale tutaj nie zakładamy, że mamy w bładzeniu losowym, dla którego szukamy maksimum, do czynienia ze zmiennymi o wartościach naturalnych.

Dowód: lematu 5.2.1.

$$\begin{aligned}
&\bar{\psi}(u) = P(R(t), t > 0) = P(S(t) < u + ct, t > 0) \\
&= P(S(T_1) < u + cT_1, S(t) - S(T_1) < u + ct - S(T_1), t > T_1) \\
&= P(X_1 < u + cT_1, S(z + T_1) - S(T_1) < u + c(z + T_1) - X_1, z > 0) \\
&= \int_0^\infty \int_0^\infty P(X_1 < u + cT_1, S(z + T_1) - S(T_1) < \\
&\qquad u + c(z + T_1) - X_1, z > 0 | X_1 = x, T_1 = y) dF_X(x) \lambda \exp(-\lambda y) dy \\
&= \int_0^\infty \int_0^\infty I_{\{x<u+cy\}}(x, y) P(S(z + y) - S(y) < u + cz + cy - x, z > 0) \\
&\qquad dF_X(x) \lambda \exp(-\lambda y) dy \\
&= \int_0^\infty \int_0^{u+cy} P(S(z) < u + cy - x + cz, z > 0) dF_X(x) \lambda \exp(-\lambda y) dy \\
&= \int_0^\infty \int_0^{u+cy} \bar{\psi}(u + cy - x) dF_X(x) \lambda \exp(-\lambda y) dy \\
&= \frac{\lambda}{c} \int_u^\infty \int_0^s \bar{\psi}(s - x) dF_X(x) \exp(-\lambda \frac{s - u}{c}) ds,
\end{aligned}$$

gdzie ostatnia równość zachodzi po podstawieniu $u + cy := s$.

Przypomnijmy regułę różniczkowania całki oznaczonej. Dla

$$\Psi(u) = \int_{a(u)}^{b(u)} f(u, s) ds,$$

$$\frac{d}{du}\Psi(u) = \int_{a(u)}^{b(u)} \frac{\partial}{\partial u} f(u, s) ds + f(u, b(u)) b'(u) - f(u, a(u)) a'(u).$$

Różniczkując względem u otrzymujemy

$$\begin{aligned}
\bar{\psi}'(u) &= \frac{\lambda}{c}[\int_u^\infty \frac{d}{du}(\int_0^s \bar{\psi}(s - x) dF_X(x) \exp(-\lambda \frac{s - u}{c})) ds - \\
&\int_0^u \bar{\psi}(u - x) dF_X(x)] \\
&= \frac{\lambda}{c}[\frac{\lambda}{c} \int_u^\infty \int_0^s \bar{\psi}(s - x) dF_X(x) \exp(-\lambda \frac{s - u}{c}) ds - \\
&\int_0^u \bar{\psi}(u - x) dF_X(x)] \\
&= \frac{\lambda}{c}[\bar{\psi}(u) - \int_0^u \bar{\psi}(u - x) dF_X(x)].
\end{aligned}$$

Całkując po zmiennej u w zakresie od 0 do t mamy

$$\bar{\psi}(t) - \bar{\psi}(0) = \frac{\lambda}{c} \int_0^t [\bar{\psi}(u) - \int_0^u \bar{\psi}(u - x) dF_X(x)] du.$$

Aby wyliczyć

$$\int_0^t \int_0^u \bar{\psi}(u-x)dF_X(x)du,$$

podstawiamy $s := u - x$ i całkujemy przez części otrzymując

$$\begin{aligned}\int_0^t \int_0^u \bar{\psi}(u-x)dF_X(x)du &= \int_0^\infty \int_0^{t-x} \bar{\psi}(s)dsdF_X(x) \\ &= \int_0^t \bar{\psi}(s)ds - \int_0^\infty \bar{\psi}(t-x)(1-F_X(x))dx,\end{aligned}$$

czyli

$$\begin{aligned}\bar{\psi}(t) - \bar{\psi}(0) = \frac{\lambda}{c}\int_0^t \bar{\psi}(u)du - \\ \frac{\lambda}{c}\int_0^t \bar{\psi}(s)ds + \frac{\lambda}{c}\int_0^\infty \bar{\psi}(t-x)(1-F_X(x))dx.\end{aligned}$$

Gdy $t \to \infty$

$$1 - \bar{\psi}(0) = \frac{\lambda}{c}\int_0^\infty (1-F_X(x))dx = \frac{\lambda}{c}\mathrm{E}\left[X\right],$$

co daje

$$\bar{\psi}(t) = (1 - \frac{\lambda}{c}\mathrm{E}\left[X\right]) + \frac{\lambda}{c}\mathrm{E}\left[X\right]\bar{\psi} * \tilde{F}_X(t).$$

□

Z równania dla $\bar{\psi}$ natychmiast otrzymujemy równanie dla ψ.

Lemat 5.2.3 *Funkcja prawdopodobieństwa ruiny ψ spełnia następujące równanie typu odnowy:*

$$\psi(u) = q(1 - \tilde{F}_X(u)) + q\psi * \tilde{F}_X(u), \tag{5.2.2}$$

Przykład 5.2.4 Gdy wielkości szkód X_i mają rozkład wykładniczy $Exp(1/\mathrm{E}\left[X\right])$, to $\tilde{F}_X$ jest znowu dystrybuantą wykładniczą $Exp(1/\mathrm{E}\left[X\right])$ i wtedy $\bar{\psi}$ jest dystrybuantą $CGeo(p, Exp(1/\mathrm{E}\left[X\right])$, wiemy z (2.2.25), że jest to dystrybuanta wykładnicza wymieszana z atomem w zerze, dokładniej, dla $p = 1 - \lambda\mathrm{E}\left[X\right]/c$:

$$\begin{aligned}\bar{\psi}(u) &= p + q(1 - \exp(-pu/\mathrm{E}\left[X\right])), \\ \psi(u) &= q\exp(-pu/\mathrm{E}\left[X\right]).\end{aligned}$$

Wprowadzając narzut (security loading) $\vartheta > 0$ poprzez równość

$$c = (1+\vartheta)\lambda\mathrm{E}\left[X\right],$$

otrzymujemy alternatywną postać

$$\bar{\psi}(u) = \frac{\vartheta}{1+\vartheta} + \frac{1}{1+\vartheta}(1 - \exp(-\frac{\vartheta}{\mathrm{E}[X](1+\vartheta)}u)),$$
$$\psi(u) = \frac{1}{1+\vartheta}\exp(-\frac{\vartheta}{\mathrm{E}[X](1+\vartheta)}u).$$

□

Rozkłady lekkoogonowe- model Cramera-Lundberga

W modelu z czasem ciągłym przydatny będzie również współczynnik dopasowania, zdefiniowany jednak inaczej niż w modelu z czasem dyskretnym. Dodatnie rozwiązanie równania

$$M_X(r) = \frac{c}{\lambda}r + 1,$$

nazywamy współczynnikiem dopasowania w modelu ciągłym i oznaczamy $\tilde{R} := \tilde{R}(X, \lambda, c)$.

Przykład 5.2.5 Dla $X \sim Exp(1/\mathrm{E}[X])$, równanie definiujące przyjmuje postać

$$\frac{1}{1 - \mathrm{E}[X]\, r} = \frac{c}{\lambda}r + 1,$$

co prowadzi do

$$\tilde{R}(Exp(1/\mathrm{E}[X]), \lambda, c) = p/\mathrm{E}[X],$$

dla $p = (1 - \lambda \mathrm{E}[X]/c)$. W języku narzutu otrzymujemy

$$\tilde{R}(Exp(1/\mathrm{E}[X]), \lambda, c) = \frac{\vartheta}{\mathrm{E}[X](1+\vartheta)}.$$

□

Porównując wzory na $\tilde{R} = \tilde{R}(Exp(1/\mathrm{E}[X]), \lambda, c)$ z wzorami na prawdopodobienstwo ruiny dla szkód o rozkładach wykładniczych z poprzedniego przykładu widać, że możemy napisać krótko, w języku współczynnika dopasowania, że w tym przypadku

$$\psi(u) = (1 - \mathrm{E}[X]\,\tilde{R})\exp(-\tilde{R}u).$$

Widzimy, z tych wzorów, że prawdopodobieństwo ruiny w przypadku szkód o rozkładach wykładniczych zależy od wielkości narzutu i wartości średniej szkody. Prawdopodobieństwo ruiny nie zmieni się w takim modelu przy jednoczesnym proporcjonalnym zwiększeniu intensywności nadchodzenia szkód i intensywności pobierania składek.

Dla rozkładów innych niż wykładniczy, wyliczenie prawdopodobieństwa ruiny jest kłopotliwe. Dla rozkładów lekkoogonowych można jednakże wyznaczyć jego granicę przy $u \to \infty$.

Twierdzenie 5.2.6 *Jeśli istnieje skończony współczynnik dopasowania dla rozkładu szkód w modelu ciągłym $\tilde{R} = \tilde{R}(X, \lambda, c)$, oraz $M'_X(\tilde{R}) < \infty$, to*

$$\lim_{u\to\infty} \psi(u)\exp(\tilde{R}u) = \frac{p}{M'_X(\tilde{R})(\lambda/c) - 1},$$

gdzie $p = 1 - \lambda \mathrm{E}[X]/c$.

Dla dużych wartości kapitału początkowego można prawdopodobieństwo ruiny przybliżyć następująco:

$$\psi(u) \approx \exp(-\tilde{R}u)\frac{p}{M'_X(\tilde{R})(\lambda/c) - 1}.$$

Dowód: Wychodząc od równania

$$\psi(u) = q(1 - \tilde{F}_X(u)) + q\psi * \tilde{F}_X(u),$$

mnożymy je obustronnie przez $\exp(\tilde{R})$, otrzymując

$$\begin{aligned}\psi(u)\exp(\tilde{R}u) = {} & q(1 - \tilde{F}_X(u))\exp(\tilde{R}u) \\ & + \int_0^u \psi(u-x)\exp(\tilde{R}(u-x))\frac{\lambda}{c}(1 - F_X(x))\exp(\tilde{R}x)dx,\end{aligned}$$

czyli traktując $\frac{\lambda}{c}(1 - F_X(x))\exp(\tilde{R}x)$ jako gęstość dystrybuanty powiedzmy F, oraz przyjmując $Z(u) = \psi(u)\exp(\tilde{R}u)$, $z(u) = q(1 - \tilde{F}_X(u))\exp(\tilde{R}u)$, widzimy, że jest to równanie typu odnowy

$$Z(u) = z(u) + Z * F(u). \tag{5.2.3}$$

Z twierdzenia 5.1.7 otrzymujemy więc

$$\lim_{u\to\infty} \psi(u)\exp(\tilde{R}u) = \frac{\int_0^\infty q(1 - \tilde{F}_X(u))\exp(\tilde{R}u)du}{\int_0^\infty 1 - F(x)dx}.$$

Pozostaje wyliczyć wartości całek pojawiajacych się w granicy,

$$\int_0^\infty q(1 - \tilde{F}_X(u))\exp(\tilde{R}u)du = p/\tilde{R},$$

$$\int_0^\infty 1 - F(x)dx = \frac{M'_X(\tilde{R})(\lambda/c) - 1}{\tilde{R}},$$

co wynika natychmiast z całkownia przez częsci. □

Dla dowolnych wartości u można podać ograniczenie górne i ograniczenie dolne na $\psi(u)$ jeśli założymy lekkoogonowość rozkładu wielkości szkód.

Twierdzenie 5.2.7 *Jeśli istnieje skończony współczynnik dopasowania dla rozkładu szkód w modelu ciągłym $\tilde{R} = \tilde{R}(X, \lambda, c) < \infty$, to*

$$a_- \exp(-\tilde{R}u) \leqslant \psi(u) \leqslant a_+ \exp(-\tilde{R}u),$$

gdzie

$$a_+ = \sup_{x \geqslant 0} \frac{z(x)}{1 - F(x)},$$

$$a_- = \inf_{x \geqslant 0} \frac{z(x)}{1 - F(x)},$$

dla funkcji z i F zdefiniowanych w równaniu (5.2.3).

5.3 Prawdopodobieństwo ruiny dla rozkładów fazowych

Przypomnijmy , że zmienna losowa ma rozkład fazowy $PH(\mathbf{p}, \mathbf{Q}, n)$, jeżeli $\bar{F}(x) = P(X > x) = \mathbf{p} \exp(\mathbf{Q}x)\mathbf{e}$, gdzie $\mathbf{p} = (p_1, \ldots, p_n)$ jest wektorem liczb nieujemnych sumujących sią do 1, $\mathbf{Q}$ jest macierzą $n \times n$, w której poza przekątną są liczby dodatnie, a na przekątnej są wartości ujemne takie, że suma w wierszach macierzy jest mniejsza lub równa zero, $\mathbf{e} = (1, \ldots, 1)'$, $\exp(\mathbf{Q})$ oznacza tzw. eksponens macierzowy, $\exp(\mathbf{Q}) := \mathbf{I} + \mathbf{Q} + \mathbf{Q}^2/2 + ... + \mathbf{Q}^n/n! + ...$, gdzie $\mathbf{I}$ jest macierzą jednostkową.

Dla rozkładu fazowego dystrybuanta resztowa F_e jest też fazowa i ma reprezentację: $PH(\mathbf{r}, \mathbf{Q}, n)$, gdzie $\mathbf{r} = -\mathbf{p}\mathbf{Q}^{-1}/\mathrm{E}[X]$ i $\mathrm{E}[X]$ wylicza się ze wzoru

$$\mathrm{E}[X] = -n!\mathbf{Q}^{-1}\mathbf{e}.$$

Jeżeli teraz dla zadanego ciągu szkód $(X_i)_{i \geqslant 1}$ o dystrybuancie fazowej F_X, rozważymy ciąg $(Y_i)_{i \geqslant 1}$ o dystrybuancie resztowej $\tilde{F}_X$, to złożony rozkład geometryczny $\sum_{i=1}^{T} Y_i$ jest znowu fazowy z parametrami $(\rho\mathbf{r}, \mathbf{Q} + \mathbf{t_0}\mathbf{r}\rho, n)$, gdzie $\mathbf{t_0} = -\mathbf{Q}\mathbf{e}$. Funkcja ψ ma więc postać

$$\psi(u) = \rho\mathbf{r} \exp\left((\mathbf{Q} + \mathbf{t_0}\mathbf{r}\rho)x\right)\mathbf{e}. \tag{5.3.1}$$

Przykład 5.3.1 Rozkład wykładniczego $Exp(\mu)$ jest fazowy z parametrami $\mathbf{p} = 1$, $\mathbf{Q} = [-1/\mu]$, $n = 1$. Wtedy $\mathbf{r} = 1$ i rozkład resztowy jest też wykładniczy. Wyliczając $\psi(u)$ ze wzoru (5.3.1) dostajemy $\psi(u) = \rho \exp(-x/(1-\rho)/\mu)$.

□

W ogólnym przypadku rozkładów fazowych, dla zadanych wartości x powyższy wzór wyliczamy numerycznie.

Przykład 5.3.2 Niech $\mathbf{p} = (1/4, 1/8, 1/8, 1/4)$ i

$$\mathbf{Q} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}$$

Wtedy

$$\bar{F}_X(x) = 1/4\exp(-x) + 1/8\exp(-2x) + 1/8\exp(-3x) + 1/4\exp(-4x)$$

jest mieszanką rozkładów wykładniczych. Rozkład resztowy dany jest wzorem

$$\tilde{F}_X(x) = 3/5\exp(-x) + 3/20\exp(-2x) + 1/10\exp(-3x) + 3/20\exp(-4x).$$

Funkcja ψ nie ma już tak ładnej postaci, można ją jednak łatwo wyrysować lub podać konkretne wartości. Na przykład $\psi(10) = 0.095$ [1].

□

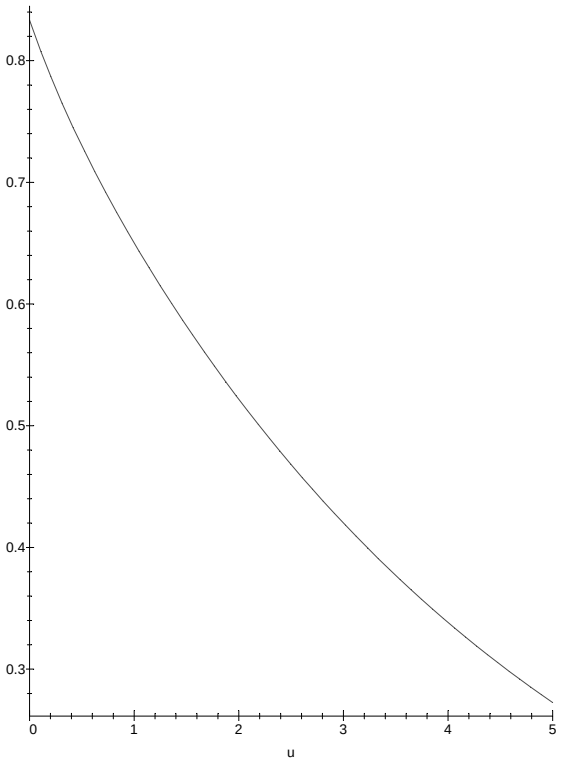
5.4 Prawdopodobieństwo ruiny dla rozkładów ciężkoogonowych

Ze wzoru (5.3.1) wynika, że w przypadku rozkładów fazowych funkcja prawdopodobieństwa ruiny maleje wykładniczo. Podobnie wygląda to w przypadku wszystkich rozkładów, dla których istnieje współczynnik dopasowania, zdefinowany w Rozdziale 4.1.1. Przypomnijmy jednak, że dla rozkładów ciężkoogonowych nie istnieje funkcja tworząca momenty, nie istnieje więc współczynnik dopasowania. W takich przypadkach jesteśmy w stanie jednak otrzymać asymptotykę prawdopodobieństwa ruiny.

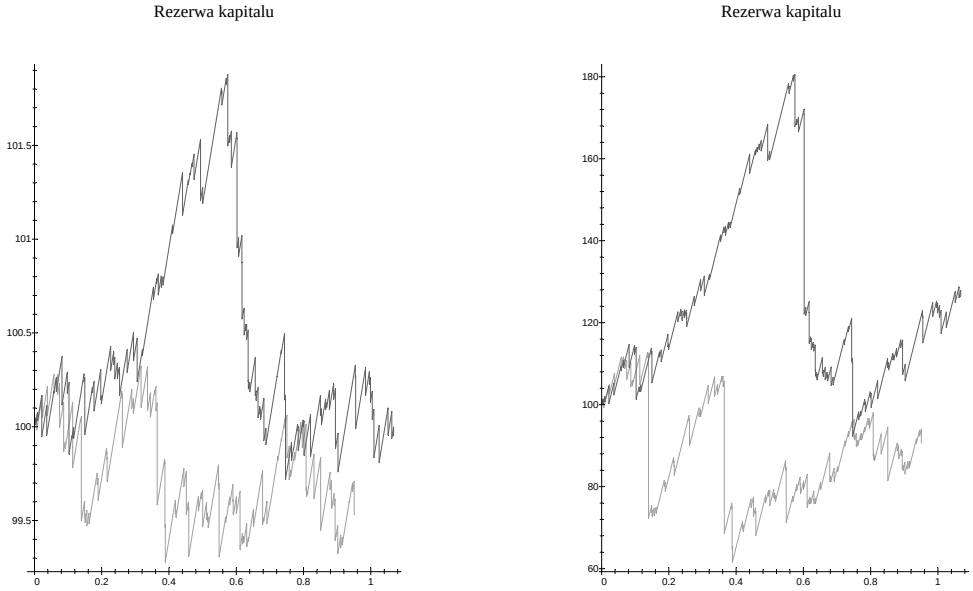
Przed podaniem twierdzenia, przyjrzyjmy się najpierw trajektoriom procesu ryzyka dla dwóch modeli.

Lewy rysunek przedstawia proces ryzyka, gdzie szkody mają rozkład wykładniczy, prawy rysunek - szkody Pareto z $\alpha = 2$, a więc z nieskończoną wariancją. Widzimy ogólną tendencję 'do góry', co jest zagwarantowane przez $\rho < 1$. Na lewym rysunku trajektorie mają lokalnie tendencję w dół poprzez nagromadzenie małych szkód, na prawym - poprzez jedną

[1] Plik ruina-ciagla-1.mws



Rysunek 5.3.1:



Rysunek 5.4.1: Trajektorie procesu ryzyka

dużą szkodę[2].

Jak należy się więc spodziewać, funkcja prawdopodobieństwa ruiny dla szkód ciężkoogonowych będzie zachowywała się inaczej niż dla fazowych. W następującym twierdzeniu użyjemy klasę rozkładów $\mathcal{S}^*$, wprowadzoną w Rozdziale 6.3.1.

Twierdzenie 5.4.1 *Jeżeli $F \in \mathcal{S}^*$, to*

$$\lim_{x\to\infty} \frac{\psi(u)}{\overline{F}_e(u)} = \frac{\rho}{1-\rho}.$$

Dowód: Podamy najpierw bez dowodu dwa lematy techniczne.

Lemat 5.4.2 *Jeżeli $F_e \in \mathcal{S}$, to dla każdego $\varepsilon > 0$ istnieje $D > 0$ takie, że dla każdego $n \geqslant 2$,*

$$\frac{\overline{F_e^{*n}}(u)}{\overline{F}_e(u)} \leqslant D(1+\varepsilon)^n, \qquad x \geqslant 0. \tag{5.4.1}$$

Lemat 5.4.3 *Niech $\overline{G}(u) = \sum_{n=0}^{\infty} p_k \overline{H^{*n}}(u)$, (p_n) jest rozkładem prawdopodobieństwa, $H \in \mathcal{S}$. Jeżeli $\sum_{n=0}^{\infty} p_n(1+\varepsilon)^n < \infty$ dla pewnego $\varepsilon > 0$, to*

$$\lim_{x\to\infty} \frac{\overline{G}(u)}{\overline{H}(u)} = \sum_{n=0}^{\infty} np_n.$$

Ze wzoru (**??**) dostajemy

$$\psi(u) = (1-\rho)\sum_{n=0}^{\infty} \rho^n P\left(\sum_{i=1}^{n} Y_i > x\right) = (1-\rho)\sum_{n=0}^{\infty} \rho^n \overline{F^{*n}}(u).$$

Połóżmy $p_n = \rho^n$ i weźmy $\varepsilon > 0$ tak, by $\rho(1+\varepsilon) < 1$. Stąd $\sum_{n=0}^{\infty} p_n(1+\varepsilon)^n < \infty$.
Z faktu, że $F \in \mathcal{S}^*$ mamy $F_e \in \mathcal{S}$, a więc korzystając z Lematu 5.4.2, istnieje $D > 0$ takie, że (5.4.1) jest prawdziwe. Z Lematu 5.4.3 mamy

$$\lim_{u\to\infty} \frac{\psi(u)}{\overline{F_e}(u)} = (1-\rho)\sum_{n=0}^{\infty} n\rho^n = (1-\rho)\frac{\rho}{(1-\rho)^2} = \frac{\rho}{1-\rho}.$$

□

5.5 Funkcje Copula

Bieżący rozdział jest oparty głównie na publikacji T. Schmidt'a [**?**] oraz R.Nelsena [**?**]. Przedstawione poniżej funkcje copula zostaną w kolejnych rozdziałach wykorzystane do opisu struktury zależności w modelach z zależnymi roszczeniami lub odstępami czasowymi.

[2] Plik ruina-ciagla-2.mws

5.5.1 Definicja i własności funkcji copula

Definicja 5.5.1 *Niech $d \geqslant 2$. Funkcja C: $R^d \to R$ nazywana jest* **funkcją copula**, *gdy jest d-wymiarową dystrybuantą o jednowymiarowych, brzegowych rozkładach jednostajnych na (0,1).*

Twierdzenie 5.5.2 *(Sklar)*
Dla d wymiarowej dystrybuanty o ciągłych dystrybuantach brzegowych $F_1, ..., F_d$ istnieje wyznaczona jednoznacznie funkcja copula C taka, że:

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)) \tag{5.5.1}$$

Odwrotnie, dla danej funkcji copula C i dystrybuant brzegowych $F_1, ..., F_d$ wzór (5.5.1) definiuje d-wymiarową dystrybuantę F.

Podstawowe własności funkcji Copula $C(u) = C(u_1, ..., u_d)$ to:
1) $C : R^d \to [0, 1]$, jest rosnąca po współrzędnych
2) Rozkład brzegowy u_i otrzymuje się przez podstawienie $u_j = 1$ dla wszystkich $j \neq i$

$$C(1, ..., 1, u_i, 1, ..., 1) = u_i, \quad u_i \in [0, 1]$$

3) Dla $a_i < b_i$ prawdopodobieństwo $P(U_1 \in [a_1, b_1], ..., U_d \in [a_d, b_d])$ jest nieujemne, co prowadzi do nierówności:

$$\sum_{i_1=1}^{2} ... \sum_{i_d=1}^{2} (-1)^{i_1+...+i_d} C(u_{1,i_1}, ..., u_{d,i_d}) \geqslant 0,$$

gdzie $u_{j,1} = a_j$ oraz $u_{j,2} = b_j$.

Z drugiej strony każda funkcja $C : [0, 1]^d - > [0, 1]$, która spełnia powyższe własności jest funkcją copula. Ponadto, mając d-wymiarową funkcję copula $C(u_1, ..., u_d)$, $C(1, u_1, ..., u_{d-1})$ jest również funkcją copula.

Hoeffding i Frechet niezależnie dowiedli, że funkcja copula zawsze leży pomiędzy określonymi granicami. Powodem tego jest istnienie skrajnych przypadków zależności. Pierwszy przypadek takiej skrajnej zależności to doskonała dodatnia zależność zmiennych (comonotonic). Wówczas funkcję copula podaje wzór:

$$C(u_1, ..., u_d) = \min(u_1, ..., u_d).$$

W przypadku niezależnych zmiennych losowych funkcja copula to:

$$C(u_1, ..., u_d) = u_1 \cdot ... \cdot u_d. \tag{5.5.2}$$

Bezpośrednio z twierdzenia Sklar'a otrzymujemy, że zmienne losowe są niezależne wtedy i tylko wtedy, gdy ich funkcja copula, to opisana wzorem (5.5.2), funkcja copula niezależności.
Jednakże, niezależność jest tylko etapem pośrednim od doskonałej dodatniej zależności do doskonałej ujemnej zależności (countermonotonic), dla której funkcję copula dla dwóch zmiennych losowych, tj. w przypadku $U_1 = 1 - U_2$, przedstawia wzór:

$$C(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\} \tag{5.5.3}$$

W całej ogólności, nie istnieje taka funkcja copula dla trzech zmiennych, ponieważ równość $U_1 = 1 - U_2$ nakłada pewne ograniczenia na zmienną losową U_3 tzn. jeśli $U_3 = 1 - U_2$, to $U_3 = U_1$, a jeśli $U_3 = 1 - U_1$, to $U_3 = U_2$. Z drugiej strony, nawet jeśli taka funkcja copula nie istnieje to wynikające z niej ograniczenia nadal obowiązują, mianowicie:

Twierdzenie 5.5.3 *(ograniczenia Frechet-Hoeffding)*
Rozważmy funkcje copula $C(u_1, ..., u_d)$. Wtedy:

$$\max\{\sum_{i_1=1}^{d} u_i + 1 - d, 0\} \leqslant C(u_1, ..., u_d) \leqslant \min\{u_1, ..., u_d\}. \tag{5.5.4}$$

W analogiczny sposób, w jaki funkcja copula wiąże dystrybuantę łączną z jej dystrybuantami brzegowymi, można opisać związek pomiędzy funkcjami przeżycia:

$$\bar{F}(x_1, ..., x_d) = \hat{C}(\bar{F}_1(x_1), ..., \bar{F}_d(x_d)) \tag{5.5.5}$$

gdzie $\bar{F}(x_1, ..., x_d)$ to łączna funkcja przeżycia opisana wzorem:

$$\bar{F}(x_1, ..., x_d) = P(X_1 > x_1, ..., X_d > x_d)$$

natomiast $\bar{F}_i(x_i)$, dla $i = 1, ..., d$, to funkcje brzegowe:

$$\bar{F}_i(x_i) = \bar{F}(-\infty, ..., x_i, ..., -\infty)$$

Funkcja $\hat{C}$ również jest funkcją copula, zwaną funkcją copula przeżycia (ang. survival copula). Funkcje C i $\hat{C}$ wiąże równość:

$$\hat{C}(u_1, ..., u_d) = u_1 + ... + u_d - 1 + C(1 - u_1, ..., 1 - u_d). \tag{5.5.6}$$

5.5.2 Funkcje copula Archimedesa

Bardzo ważną klasą kopuł są, tak zwane, kopuły Archimedesa. Charakteryzują się one dużą różnorodnością oraz łatwością, z jaką mogą być konstruowane.

Definicja 5.5.4 *Niech $\varphi : [0, 1] \to [0, \infty)$ będzie ciągłą, ściśle malejącą funkcją taką, że $\varphi(1) = 0$. Funkcją pseudo-odwrotną do φ jest funkcja $\varphi^{[-1]}$, zdefiniowana wzorem:*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leqslant t \leqslant \varphi(0) \\ 0\,, & \varphi(0) \leqslant t \leqslant \infty \end{cases} \tag{5.5.7}$$

Funkcja $\varphi^{[-1]}$ jest ciągła i nierosnąca na przedziale $[0,\infty]$, oraz ściśle malejąca na przedziale $[0,\varphi(0)]$. Poza tym, $\varphi^{[-1]}(\varphi(u)) = u$ na przedziale $[0,1]$, oraz:

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t, & 0 \leqslant t \leqslant \varphi(0), \\ \varphi(0), & \varphi(0) \leqslant t \leqslant \infty, \end{cases}$$

$$= min(t, \varphi(0)).$$

W końcu, jeśli $\varphi(0) = \infty$, to $\varphi^{[-1]} = \varphi^{-1}$.

Celem prezentacji kolejnego lematu wprowadzamy definicję funkcji całkowicie monotonicznych:

Definicja 5.5.5 *Funkcja f jest całkowicie monotoniczna na przedziale J, jeżeli jest na tym przedziale ciągła i posiada pochodne $f^{(n)}$ wszystkich rzędów spełniające nierówność:*

$$(-1)^n f^{(n)}(\lambda) \geqslant 0, \tag{5.5.8}$$

dla każdego λ należącego do wnętrza przedziału J oraz n=0,1,2,... .

Lemat 5.5.1 *Niech $\varphi : [0,1] \to [0,\infty)$ będzie ciągłą, ściśle malejącą funkcją taką, że $\varphi(1) = 0$ oraz niech $\varphi^{[-1]}$ będzie funkcją pseudo-odwrotną do φ zdefiniowaną przez (5.5.7). Funkcja $C : [0,1]^d \to [0,1]$ postaci:*

$$C(u) = \varphi^{[-1]}\left(\sum_{i=1}^{d} \varphi(u_i)\right) \tag{5.5.9}$$

jest funkcją copula wtedy i tylko wtedy, gdy funkcja $\varphi^{[-1]}$ jest całkowicie monotoniczna na przedziale $[0,\infty)$.

Konsekwencją (5.5.8) jest, że jeśli $\varphi^{[-1]}$ jest funkcją całkowicie monotoniczną oraz $\varphi^{[-1]}(c) = 0$ dla skończonego $c > 0$, wtedy $\varphi^{[-1]}$ jest tożsamościowo równa 0. Zatem funkcja $\varphi^{[-1]}$ musi być dodatnia na przedziale $[0,\infty)$, co z kolei oznacza, że zachodzi równość $\varphi^{[-1]} = \varphi^{-1}$.

Funkcje copula postaci (5.5.9) są nazywane funkcjami copula Archimedesa. Natomiast funkcja φ jest zwana generatorem funkcji copula.

Przykład 5.5.1 Funkcja Copula Gumbell'a jest szczególną funkcją Archimedesa, uzyskaną dla generatora $\varphi(t) = (-\ln t)^\theta$:

$$C_\theta{}^{Gumbel}(u_1, ..., u_d) = exp[-((-\ln u_1)^\theta + ... + (-\ln u_d)^\theta)^{1/\theta}] \tag{5.5.10}$$

gdzie $\theta \in [1,\infty)$. Dla $\theta = 1$ otrzymujemy funkcję copula niezależności. Natomiast dla $\theta \to \infty$ szukana funkcja jest zbieżna do funkcji copula doskonale dodatnio zależnej, tak

więc funkcja Copula Gumbel'a jest funkcją copula albo niezależności, albo doskonałej dodatniej zależności.

Zgodnie z wzorem (5.5.6) funkcja copula przeżycia w tym przypadku to:

$$\hat{C}(u_1, ..., u_d) = u_1 + ... + u_d - 1 + exp[-((-\ln(1-u_1))^\theta + ... + (-\ln(1-u_d))^\theta)^{1/\theta}].$$

Przykład 5.5.2 Funkcja Claytona również należy do rodzin funkcji copula Archimedesa, otrzymujemy ją przez podstawienie $\varphi(t) = (t^{-\theta} - 1)/\theta$:

$$C_\theta{}^{Clayton}(u_1, ..., u_d) = (max\{{u_1}^{-\theta} + ... + {u_d}^{-\theta} - 1, 0\})^{-1/\theta} \tag{5.5.11}$$

gdzie $\theta \in (0, \infty)$. Dla $\theta \to 0$, otrzymujemy funkcję copula niezależności, a dla $\theta \to \infty$ copula Clayton'a zmierza do funkcji copula doskonale dodatnio skorelowanej. Dla $\theta = -1$ otrzymujemy ograniczenie z dołu Frechet-Hoeffding'a. Podsumowując, tak jak funkcja copula Gumbel'a, copula Clayton'a interpoluje wśród struktur zależności takich jak: niezależność oraz doskonała dodatnia zależność.
Funkcja copula przeżycia wyliczona dla omawianego przykładu z wzoru (5.5.6) to:

$$\hat{C}(u_1, ..., u_d) = u_1 + ... + u_d - 1 + (max\{(1-u_1)^{-\theta} + ... + (1-u_d)^{-\theta} - 1, 0\})^{-1/\theta}.$$

5.6 Model zrandomizowany

Rozważmy teraz klasyczny model ryzyka (**??**) z wykładniczymi wielkościami roszczeń, dla których, dla każdego n, spełniona jest równość:

$$P(X_1 > x_1, ..., X_n > x_n | \Theta = \theta) = \prod_{k=1}^{n} e^{-\theta x_k}, \tag{5.6.1}$$

gdzie Θ to dodatnia zmienna losowa. Wyliczając rozkład brzegowy X_k okazuje się, że nie jest on wykładniczy, a wielkości roszczeń są zależne. Pokażmy to na przykładzie.

Przykład 5.6.1 Weźmy dwuwymiarowy wektor losowy $X = (X_1, X_2)$ oraz zmienną losową Θ o rozkładzie

$$F_\Theta(\theta) = \frac{1}{2}\ \mathbf{1}_{[1,\infty)}(\theta) + \frac{1}{2}\ \mathbf{1}_{[2,\infty)}(\theta)$$

Z wcześniejszych rozważań mamy równość:

$$P(X_1 > x_1, X_2 > x_2 | \Theta = \theta) = \prod_{k=1}^{n} e^{-\theta x_k} = e^{-\theta(x_1 + x_2)}$$

$$P(X_1 > x_1, X_2 > x_2) = \int_0^\infty \exp^{-\theta(x_1+x_2)}\, dF_\Theta(\theta)$$

Liczymy rozkład brzegowy X_1. Najprościej będzie skorzystać z rozkładu łącznego wykluczając z niego wpływ zmiennej X_2, co osiągniemy przez podstawienie $x_2 = 0$:

$$P(X_1 > x_1, X_2 > 0) = P(X_1 > x_1) = \int_0^\infty \exp^{-\theta(x_1+0)}\, dF_\Theta(\theta) = \frac{1}{2}\exp^{-x_1} + \frac{1}{2}\exp^{-2x_1}$$

Analogicznie dla X_2:

$$P(X_1 > 0, X_2 > x_2) = P(X_2 > x_2) = \int_0^\infty \exp^{-\theta(0+x_2)}\, dF_\Theta(\theta) = \frac{1}{2}\exp^{-x_2} + \frac{1}{2}\exp^{-2x_2}$$

Sprawdzamy niezależność:

$$P(X_1 > x_1)P(X_2 > x_2) = (\frac{1}{2}\exp^{-x_1} + \frac{1}{2}\exp^{-2x_1})(\frac{1}{2}\exp^{-x_2} + \frac{1}{2}\exp^{-2x_2}) =$$

$$\frac{1}{4}\exp^{-x_1-x_2} + \frac{1}{4}\exp^{-x_1-2x_2} + \frac{1}{4}\exp^{-2x_1-x_2} + \frac{1}{4}\exp^{-2x_1-2x_2}$$

$$P(X_1 > x_1, X_2 > x_2) = \int_0^\infty \exp^{-\theta(x_1+x_2)}\, dF_\Theta(\theta) = \frac{1}{2}\exp^{-x_1-x_2} + \frac{1}{2}\exp^{-2x_1-2x_2}$$

Iloczyn brzegowych rozkładów i rozkład łączny są różne, co świadczy o tym, że zmienne X_1 i X_2 nie są niezależne. Rozkłady brzegowe nie są wykładnicze, ale są mieszankami rozkładów wykładniczych.

Kolejnym krokiem jest przedstawienie wzoru na prawdopodobieństwo ruiny dla modelu z zależnymi ryzykami. W tym celu przypomnę wzór na prawdopodobieństwo ruiny (**??**) w klasycznym modelu z wykładniczymi wielkościami roszczeń $Exp(\theta)$, który został wyprowadzony w przykładzie (1.1):

$$\psi_\theta(u) = min\{\frac{\lambda}{\theta c} exp\{-(\theta - \frac{\lambda}{c})u\}, 1\}, u \geqslant 0. \tag{5.6.2}$$

Wówczas, po uwzględnieniu zmiennej θ, prawdopodobieństwo ruiny dla modelu zależności dane jest wzorem:

$$\psi(u) = \int_0^\infty \psi_\theta(u) dF_\Theta(\theta). \tag{5.6.3}$$

Dla $\theta \leqslant \theta_0 = \frac{\lambda}{c}$, złamany jest wspomniany warunek zysku netto, mówiący, że składka c zebrana w jednostce czasu musi być większa niż średnia wypłaconych szkód. Oznacza to,

że dla $\theta \leqslant \theta_0 = \frac{\lambda}{c}$ pojawienie się ruiny jest pewne tzn. $\psi_\theta(u) = 1$, dla wszystkich $u \geqslant 0$, dlatego wzór ogólny na prawdopodobieństwo ruiny można zastąpić przez:

$$\psi(u) = F_\Theta(\theta_0) + \int_{\theta_0}^{\infty} \psi_\theta(u) dF_\Theta(\theta). \tag{5.6.4}$$

Bezpośrednią konsekwencją jest:

$$\lim_{u\to\infty} \psi(u) = F_\Theta(\theta_0) \tag{5.6.5}$$

co jest dodatnie, gdy zmienna losowa Θ ma dodatnią masę prawdopodobieństwa na poziomie $\theta = \frac{\lambda}{c}$ lub poniżej.

Twierdzenie 5.6.1 *Model ryzyka z zależnymi wielkościami szkód, spełniający równanie (5.6.1), można opisać za pomocą wektora szkód $(X_1, X_2, ...)$ o całkowicie monotonicznych brzegowych rozkładach roszczeń zmiennych $X_1, X_2, ...$ takiego, że odpowiednie funkcje copula przeżycia są funkcjami Archimedesa z generatorem $\varphi = (\mathcal{L}(F_\Theta))^{-1}$, gdzie $\mathcal{L}(F_\Theta)$ oznacza transformatę Laplace-Stieltjesa F_Θ, a potęga (-1) funkcję odwrotną.*

Dowód: Powołując się na wzór (5.6.1), łączny rozkład $X_1, .., X_n$ możemy zapisać jako:

$$P(X_1 > x_1, ..., X_n > x_n) = \int_0^{\infty} e^{-\theta(x_1+...+x_n)} dF_\Theta(\theta) = \mathcal{L}(F_\Theta)(x_1 + ... + x_n). \tag{5.6.6}$$

Na podstawie twierdzenia Sklara (tw. 2.2) oraz wzoru dla kopuł przeżycia (5.5.5), dla każdej wielowymiarowej dystrybuanty z rozkładami brzegowymi $F_1, ..., F_n$ istnieje kopuła przeżycia $\hat{C}$ taka, że:

$$P(X_1 > x_1, ..., X_n > x_n) = \hat{C}(\bar{F}_{X_1}(x_1), ..., \bar{F}_{X_n}(x_n))$$

gdzie $\bar{F}_X(x_i) = 1 - F_X(x_i)$ to ogon dystrybuanty rozkładu brzegowego X_i. W omawianym przykładzie X_i, $i = 1, ..., n$ mają takie same rozkłady. Jeśli kopuła jest funkcją Archimedesa z generatorem φ wtedy powyższy wzór możemy wyrazić jako: $\hat{C}(\bar{F}_X(x_1), ..., \bar{F}_X(x_n)) = \varphi^{-1}(\varphi(\bar{F}_X(x_1)) + ... + \varphi(\bar{F}_X(x_n)))$, gdzie

$$\bar{F}_X(x_i) = \int_0^{\infty} e^{-\theta x_i} dF_\Theta(\theta) = \mathcal{L}(F_\Theta)(x_i), \qquad i = 1...n, \tag{5.6.7}$$

co dokładnie pasuje do (5.6.6), gdy generator $\varphi(t) = (\mathcal{L}(F_\Theta))^{-1}(t)$. Zauważmy, że φ jako odwrotna tranformata Laplace'a dystrybuanty jest funkcją ciągłą, ściśle malejącą z $[0, 1]$ do $[0, \infty]$, dla której $\varphi(0) = \infty$ i $\varphi(1) = 0$ i φ^{-1} jest całkowicie monotoniczna, więc zgodnie z lematem 5.5.1 funkcja copula Archimedesa jest dobrze określona dla każdego n. Z równości (5.6.7) widzimy, że zmienne losowe X_i muszą być całkowicie monotoniczne.

Uwaga

Powyższą zmianę konstrukcji można zobrazować jako pobieranie próbek θ z Θ według F_Θ, a następnie poprowadzenie trajektorii dla niezależnego modelu ryzyka z parametrem θ. Tak, więc zależność zostaje wprowadzana poprzez realizację modelu dla wszystkich możliwych wartości θ. Odpowiednio, otrzymana zależność będzie tym silniejsza im bardziej dytrybuanta Θ będzie rozrzucona. Taka zmiana w budowie modelu jest tylko narzędziem do określenia wzoru na prawdopodobieństwo ruiny, natomiast nie jest konieczna do stworzenia modelu zależności. Kolejno można równoważnie pójść w dwie strony: dla każdego modelu ryzyka z całkowicie monotonicznymi wielkościami roszczeń - zmiennymi losowymi X_i, oraz dla dodatniej zmiennej losowej Θ zachodzi równość (5.6.7). Wówczas wzór (5.6.2) odnosi się do struktury zależności opisanej za pomocą kopuły Archimedesa z generatorem $\varphi = (\mathcal{L}(F_\Theta))^{-1}$. Alternatywnie można zacząć od określenia kopuły Archimedesa przez jej generator, wówczas powyższe równania dadzą rozkład brzegowy, dla którego jawny wzór (5.6.2) zachodzi.
Poniżej znajduje się kilka szczegółowych przykładów modeli zależnych.

5.7 Przykłady modeli z zależnymi wielkościami roszczeń

5.7.1 Roszczenia o rozkładzie Pareto z funkcją Copula Claytona

Rozważmy model zrandomizowany z wykładniczymi wielkościami roszczeń (5.6.1) oraz z parametrem Θ o rozkładzie $Gamma(\alpha, \beta)$ z gęstością:

$$f_\Theta(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}, \qquad \theta > 0$$

Otrzymany wówczas rozkład łączny wielkości roszczeń to:

$$P(X_1 > x_n, ..., X_n > x_n) = \int_0^\infty e^{-\theta(x_1+...+x_n)}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta,$$

co wynika bezpośrednio z:

$$P(X_1 > x_n, ..., X_n > x_n|\Theta = \theta) = e^{-\theta(x_1+...+x_n)}.$$

Na podstawie twierdzenia (3.1) strukturę zależności rozważanego modelu można przedstawić za pomocą rozkładów brzegowych oraz funkcji copula przeżycia Archimedesa z generatorem:

$$\varphi(t) = (\mathcal{L}(F_\Theta))^{-1}(t)$$

W tym celu obliczamy rozkład brzegowy wielkości roszczeń X_k, zgodnie ze wzorem (5.6.7):

$$\bar{F}_X(x) = \int_0^\infty e^{-\theta x}f_\Theta(\theta)d\theta = \int_0^\infty e^{-\theta x}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta = (\tfrac{\beta}{x+\beta})^\alpha.$$

$$\int_0^\infty \frac{(x+\beta)^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\theta(x+\beta)}d\theta = (\tfrac{\beta}{x+\beta})^\alpha = (1+\tfrac{x}{\beta})^{-\alpha}, \qquad x \geqslant 0$$

Wyrażenie pod całką $\frac{(x+\beta)^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\theta(x+\beta)}$ jest gęstością zmiennej θ z rozkładem Gamma z parametrami α oraz $x+\beta$, co na przedziale $(0,\infty)$ całkuje się do 1.
Otrzymany rozkład brzegowy X_k dla $k=1,...,n$ to rozkład Pareto(α,β). Mając dany rozkład brzegowy, na podstawie wzoru (5.6.7) z dowodu twierdzenia (3.1), można obliczyć generator funkcji copula Archimedesa:

$$\varphi(t)=(\mathcal{L}(F_\Theta))^{-1}(t)=(\bar{F}_X)^{-1}(t)=t^{-1/\alpha}-1.$$

Kopuła Archimedesa z takim generatorem φ, opisana w przykładzie (2.2), jest kopułą Claytona z parametrem α. Z równości (5.6.4) wynika dla tego modelu, że:

$$\psi(u)=F_\Theta(\theta_0)+\int_{\theta_0}^{\infty}\frac{\lambda}{\theta c}e^{-(\theta-\frac{\lambda}{c})u}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta.$$

Na początek obliczymy $F_\Theta(\theta_0)$ poprzez oszacowanie ogonu dystrybunty zmiennej Θ w punkcie θ_0 i wykorzystanie prawdopodobieństwa przeciwnego :

$$\bar{F}_\Theta(\theta_0)=\int_{\theta_0}^{\infty}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta=\frac{\beta}{\Gamma(\alpha)}\int_{\theta_0}^{\infty}(\beta\theta)^{\alpha-1}e^{-\beta\theta}d\theta=\left[\begin{array}{ccc}t&:=&\beta\theta\\ dt&:=&\beta d\theta\\ \frac{1}{\beta}dt&:=&d\theta\end{array}\right]=$$

$$\frac{1}{\Gamma(\alpha)}\int_{\beta\theta_0}^{\infty}t^{\alpha-1}e^{-t}dt=\frac{\Gamma(\alpha,\beta\theta_0)}{\Gamma(\alpha)}$$

Czyli:

$$F_\Theta(\theta_0)=1-\frac{\Gamma(\alpha,\beta\theta_0)}{\Gamma(\alpha)}$$

gdzie $\Gamma(\alpha,x)=\int_x^{\infty}\omega^{\alpha-1}e^{-\omega}d\omega$ jest niekompletną funkcją Gamma.
Pozostaje obliczyć drugą część tj. całkę ze wzoru na prawdopodobieństwo ruiny, pamiętając z wcześniejszych rozważań, że $\theta_0=\frac{\lambda}{c}$:

$$\int_{\theta_0}^{\infty}\frac{\lambda}{\theta c}e^{-(\theta-\frac{\lambda}{c})u}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta=\frac{\theta_0e^{\theta_0u}\beta}{\Gamma(\alpha)}\int_{\theta_0}^{\infty}\beta^{\alpha-1}\theta^{\alpha-2}e^{-\theta(u+\beta)}d\theta=$$

$$\frac{\theta_0e^{\theta_0u}\beta^\alpha}{\Gamma(\alpha)(u+\beta)^{\alpha-2}}\int_{\theta_0}^{\infty}(\theta(\beta+u))^{\alpha-2}e^{-\theta(u+\beta)}d\theta=\left[\begin{array}{ccc}t&:=&\theta(\beta+u)\\ dt&:=&(\beta+u)d\theta\\ \frac{1}{(\beta+u)}dt&:=&d\theta\end{array}\right]=$$

$$\frac{\theta_0e^{\theta_0u}\beta^\alpha}{\Gamma(\alpha)(u+\beta)^{\alpha-1}}\int_{(\beta+u)\theta_0}^{\infty}t^{\alpha-2}e^{-t}dt=\theta_0e^{\theta_0u}\beta(1+\frac{u}{\beta})^{-(\alpha-1)}\frac{\Gamma(\alpha-1,(\beta+u)\theta_0)}{\Gamma(\alpha)}$$

Ostateczny wzór na prawdopodobieństwo ruiny dla danego przykładu to:

$$\psi(u)=1-\frac{\Gamma(\alpha,\beta\theta_0)}{\Gamma(\alpha)}+\theta_0e^{\theta_0u}\beta(1+\frac{u}{\beta})^{-(\alpha-1)}\frac{\Gamma(\alpha-1,(\beta+u)\theta_0)}{\Gamma(\alpha)}.$$

W szczególności, z (5.6.5) wynika:

$$\lim_{u\to\infty}\psi(u) = 1 - \frac{\Gamma(\alpha, \frac{\beta\lambda}{c})}{\Gamma(\alpha)} \tag{5.7.1}$$

Wreszcie dla kapitału początkowego u=0 otrzymujemy prosty wzór:

$$\psi(0) = 1 - \frac{\Gamma(\alpha, \beta\theta_0)}{\Gamma(\alpha)} + \beta\theta_0\frac{\Gamma(\alpha-1, \beta\theta_0)}{\Gamma(\alpha)}$$

Warto podjąć próbę porównania wyników wzoru na prawdopodobieństwo ruiny dla modelu niezależnego i dla modelu z zależnymi ryzykami. Dlatego w celu porównania, dla obu przypadków załóżmy $\lambda = c = 1$ oraz, że wartości oczekiwane wielkość roszczeń będą równe. Rozkład brzegowy w niezależnym modelu to zwykły rozkład wykładniczy z wartości oczekiwaną $E[X] = 1/\theta$. Z kolei rozkład brzegowy wielkości roszczeń w modelu zależnym to rozkład $Pareto(\alpha, \beta)$. Wartość oczekiwana zmiennej z rozkładu $Pareto(\alpha, \beta)$ to $E[X] = \frac{\alpha\cdot\beta}{\alpha-1}$. Tak, więc szukana równość to: $\frac{1}{\theta} = \frac{\alpha\cdot\beta}{\alpha-1}$. Dodatkowe ograniczenia narzuca warunek składki netto, tj. $E[X] < \frac{c}{\lambda}$, czyli przy omówionych założeniach: $E[X] < 1$. Dla modelu niezależnego warunek jest spełniony, gdy $\theta > 1$, a dla zależnego, gdy $\beta < \frac{\alpha-1}{\alpha}$. Po rozważeniu ograniczeń można zaproponować odpowiednie parametry np.: $\theta = 2$, $\alpha = 3$, $\beta = \frac{1}{3}$.

Kolejno zostaną obliczone wzory na prawdopodobieństwo ruiny dla modelu niezależnego i z zależnościami.

Prawdopodobieństwo ruiny w modelu niezależnym:

$$\psi(u) = min\{\frac{1}{2}exp\{-u\}, 1\}$$

Prawdopodobieństwo ruiny w modelu z zależnymi wielkościami roszczeń:

$$\psi(u) = 1 - \frac{\Gamma(3, \frac{1}{3})}{\Gamma(3)} + e^u\frac{1}{3}(1+3u)^{-2}\frac{\Gamma(2, \frac{1}{3}+u)}{\Gamma(3)}$$

Wyniki obliczeń potrzebnych do ostatecznego wzoru, wykonane w Matlabie:

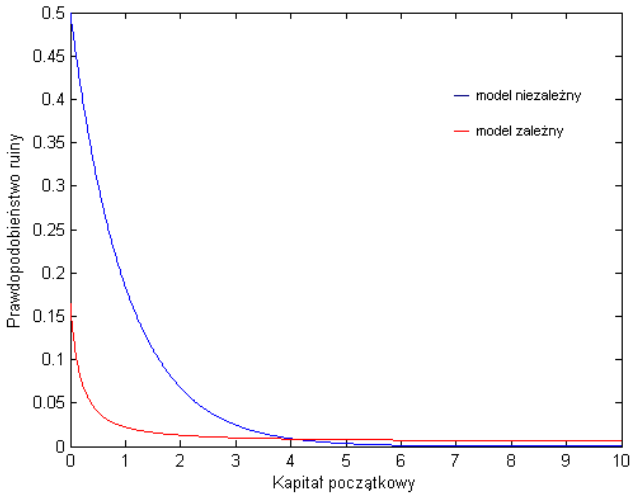
$$\Gamma(3) = \int_0^\infty x^2e^{-x}dx = 2$$

$$\Gamma(3, 1/3) = \int_{1/3}^\infty x^2e^{-x}dx = \tfrac{25}{9}e^{-1/3}$$

$$\Gamma(2, 1/3+u) = (4/3+u)e^{-(1/3+u)}$$

Prowadzi to do ostatecznego wzoru na prawdopodobieństwo ruiny w modelu zależnym:

$$\psi(u) = 1 - \frac{25}{18}e^{-1/3} + \frac{\frac{4}{3}+u}{6(1+3u)^2}e^{-1/3}$$



Z wykresu odczytujemy, że prawdopodobieństwo ruiny dla modelu niezależnego jest z początku dużo wyższe niż prawdopodobieństwo ruiny w modelu z zależnymi ryzykami, ale również szybciej maleje. Dla $u > 4$ prawdopodobieństwo ruiny dla modelu niezależnego jest już mniejsze. Przykładowo obliczę prawodpodobieństwa ruiny obu modeli dla u=0 oraz dla u=5:

Model niezależny:

$$\psi(0) = min\{\frac{1}{2}, 1\} = 0,5$$

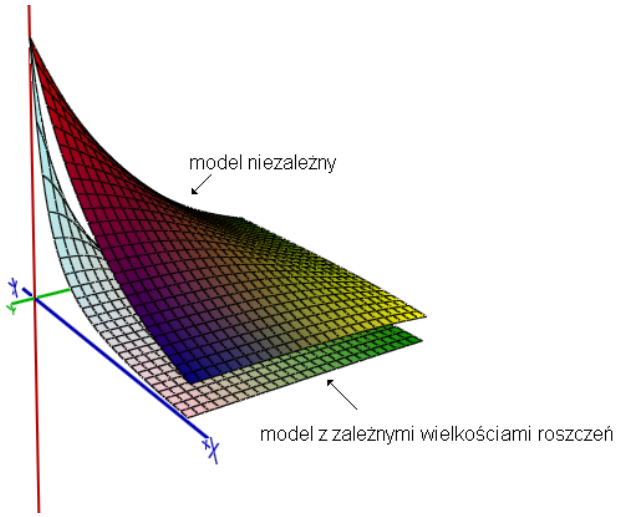
$$\psi(5) = min\{\frac{1}{2}exp\{-5\}, 1\} = 0.5 \cdot e^{-5} \approx 0,0035$$

Model zależny:

$$\psi(0) = 1 - \frac{25}{18}e^{-1/3} + \frac{4}{18}e^{-1/3} = 1 - \frac{21}{18}e^{-1/3} \approx 0,2$$

$$\psi(5) = 1 - \frac{25}{18}e^{-1/3} + \frac{\frac{4}{3}+5}{6(1+15)^2}e^{-1/3} \approx 0,0077$$

Dla zerowego kapitału początkowego prawdopodobieństwo ruiny w modelu niezależnym jest dużo większe, ale dla $u = 5$ jest już nieznacznie mniejsze niż w modelu z zależnymi ryzykami. Poniżej został umieszczony wykres funkcji copula przeżycia dla obu modeli, zgodnie z ustalonymi parametrami, dla dwóch zmiennych X i Y:



5.7.2 Roszczenia o rozkładzie Weibull'a z funkcją Copula Gumbel'a

Rozważmy kolejny model zależności. Tym razem zmienna losowa Θ pochodzi z rozkładu stabilnego (1/2) (zwanego również rozkładem Levy'ego) z gęstością:

$$f_\Theta(\theta) = \frac{\alpha}{2\sqrt{\pi\theta^3}} e^{-\alpha^2/4\theta}, \qquad \theta > 0$$

Otrzymany wówczas rozkład łączny wielkości roszczeń to:

$$P(X_1 > x_1, ..., X_n > x_n) = \int_0^\infty e^{-\theta(x_1+...+x_n)} \frac{\alpha}{2\sqrt{\pi\theta^3}} e^{-\alpha^2/4\theta} d\theta$$

Analogicznie do poprzedniego przykładu zostanie obliczony rozkład brzegowy $\bar{F}_X(x)$ i generator funkcji copula, aby następnie skorzystać z twierdzenia 3.1. W celu otrzymania rozkładu brzegowego zmiennej X_1, należy wykluczyć z rozkładu łącznego wpływ pozostałych zmiennych przez podstawienie $X_2 = ... = X_n = 0$:

$$\bar{F}_X(x) = \int_0^\infty e^{-\theta x} f_\Theta(\theta) d\theta = exp\{-\alpha x^{1/2}\}, \qquad x \geqslant 0.$$

Uzyskany rozkład brzegowy to rozkład Weibull'a z parametrem kształtu 1/2. Ponieważ $\mathcal{L}(F_\Theta)(s) = \bar{F}_\Theta(s) = e^{-\alpha\sqrt{s}}$, to otrzymany generator $\varphi(t)$, jako funkcja odwrotna ogona dystrybuanty brzegowego rozkładu wielkości roszczeń, ma wartość $(\mathcal{L}(F_\Theta))^{-1}(t) = (-\ln t)^2$ (dla wszystkich wartości α). Zgodnie z definicją dana kopuła Archimedesa jest kopułą Gumbel'a (przykład 2.2), którą opisuje wzór:

$$C_2{}^{Gumbel}(F_1(x_1), ..., F_d(x_d)) = exp\left[-\left(\sum_{i=1}^{d}(-ln(1-e^{-\alpha\sqrt{x_i}}))^2\right)^{1/2}\right].$$

Należy pamiętać, że rozkład brzegowy roszczeń zmienia się w zależności od wyboru α, natomiast kopuła pozostaje niezmieniona. Z równości (5.6.5) otrzymujemy:

$$\psi(u) = F_\Theta(\lambda/c) + \int_{\lambda/c}^{\infty} \frac{\lambda}{\theta c} e^{-\theta u} e^{\lambda c/u} \frac{\alpha}{2\sqrt{\pi\theta^3}} e^{-\alpha^2/4\theta} d\theta,$$

co można też wyrazić za pomocą funkcji błędu

$$Erfc(z) = 1 - Erf(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-\omega^2} d\omega = 2\Phi(-z\sqrt{2})$$

jako:

$$\psi(u) = Erfc(\tfrac{\alpha}{2\sqrt{\lambda/c}}) + \tfrac{\lambda}{\alpha^2 c} e^{-\alpha^2 c/4\lambda} (-\tfrac{2\alpha}{\sqrt{\pi\lambda/c}} + e^{(c\alpha - 2\sqrt{u}\lambda)^2/4c\lambda}(1 + \alpha\sqrt{u}) \cdot$$

$$Erfc(\sqrt{u\lambda/c} - \tfrac{\alpha}{2\sqrt{\lambda/c}}) + e^{(c\alpha + 2\sqrt{u}\lambda)^2/4c\lambda}(-1 + \alpha\sqrt{u}) Erfc(\sqrt{u\lambda/c} + \tfrac{\alpha}{2\sqrt{\lambda/c}}))$$

Dla $z > 0$: $Erfc(z) = \Gamma(1/2, z^2)/\sqrt{\pi}$. Dla u=0 otrzymujemy:

$$\psi(0) = Erfc(\frac{\alpha}{2\sqrt{\lambda/c}}) - \frac{2\sqrt{\lambda/c}}{\alpha\sqrt{\pi}} e^{-c\alpha^2/4\lambda} + \frac{2\lambda}{c\alpha^2} Erf(\frac{\alpha}{2\sqrt{\lambda/c}})$$

oraz

$$\lim_{u\to\infty} \psi(u) = Erfc(\frac{\alpha}{2\sqrt{\lambda/c}})$$

Jeśli $F_\Theta(a) = 0$ dla $a > 0$ (tzn. nie ma żadnej masy prawdopodobieństwa w okolicach 0), otrzymany rozkład brzegowy wielkości roszczeń jest lekkoogonowy. Oznacza to, że ogon dystrybuanty da się ograniczyć przez pewną funkcję wykładniczą.

Tak jak w poprzednim przykładzie spróbujmy porównać model niezależny z modelem z zależnymi ryzykami (tym razem Θ ma rozkład stabilny (1/2)). Zakładamy $\lambda = c = 1$ oraz, że wartości oczekiwane wielkość roszczeń w modelach będą równe. Wartość oczekiwana zmiennej z rozkładu Weibulla to: $\frac{1}{\alpha^2}$, czyli szukamy parametrów spełniających równość: $\frac{1}{\theta} = \frac{2}{\alpha^2}$. Dodatkowo, średnia szkoda musi być mniejsza od składki tj.: $E[X] < \frac{c}{\lambda} = 1$. Parametry spełniające wszystkie podane warunki to np.: $\theta = 2$, $\alpha = 2$. Kolejno zostaną obliczone wzory na prawdopodobieństwo ruiny dla modelu niezależnego i z zależnościami.

Prawdopodobieństwo ruiny w modelu niezależnym:

$$\psi(u) = min\{\frac{1}{2} exp\{-u\}, 1\}$$

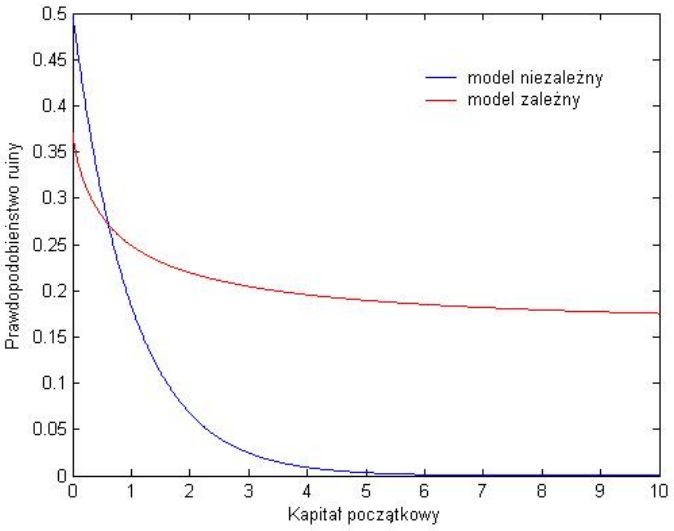
Prawdopodobieństwo ruiny w modelu z zależnymi wielkościami roszczeń:

$$\psi(u) = Erfc(1) + \tfrac{1}{4} e^{-1} (-\tfrac{4}{\sqrt{\pi}} + e^{(1-\sqrt{u})^2}(1 + 2\sqrt{u}) \cdot$$

$$Erfc(\sqrt{u} - 1) + e^{(1+\sqrt{u})^2}(-1 + 2\sqrt{u}) Erfc(\sqrt{u} + 1))$$

Kolejno zostały porównane prawdopodobieństwa ruiny dla kilku wybranych wartości kapitału początkowego:

	Prawdopodobieństwo ruiny	
Kapitał początkowy	Model niezależny	Model zależny
0	0,5	0,3712
1	0,184	0,2492
3	0,0249	0,2039
5	0,0034	0,192188



Tak, jak w poprzednim przykładzie, dla zerowego kapitału początkowego prawdopodobieństwo ruiny w modelu niezależnym jest większe, ale również szybciej maleje. Już dla u=1 prawdopodobieństwo ruiny w modelu niezależnym jest mniejsze niż w modelu z zależnymi ryzykami.

5.7.3 Odstępy między roszczeniami o rozkładzie Pareto z funkcją Copula Claytona

Rozważmy model, gdzie parametr Λ ma rozkład $Gamma(\alpha, \beta)$ z gęstością:

$$f_\Lambda(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}, \qquad \lambda > 0.$$

Wzór na rozkład brzegowy odstępów czasowych jest analogiczny do (5.6.7):

$$\bar{F}_T(t) = \int_0^\infty e^{-\lambda t} f_\Lambda(\lambda) d\lambda = (1 + t/\beta)^{-\alpha}, \ \ t \geqslant 0.$$

Natomiast ich strukturę zależności, zgodnie z twierdzeniem (**??**), opisuje funkcja copula Claytona z generatorem:

$$\varphi(t) = t^{-1/\alpha} - 1.$$

Jeśli rozważamy szczególny przypadek, gdy wielkości roszczeń mają rozkład wykładniczy z parametrem $\frac{1}{EX}$ (tym razem ustalonym), to otrzymujemy wzór (**??**):

$$\psi_\lambda(u) = min\{EX\frac{\lambda}{c}exp\{-(\frac{1}{EX} - \frac{\lambda}{c})u\}, 1\}, \quad u \geqslant 0. \tag{5.7.2}$$

Zgodnie ze wzorem (**??**) oraz warunkiem zysku netto z wartością progową $\lambda_0 = \frac{c}{EX}$ otrzymujemy ostateczny wzór na prawdopodobieństwo ruiny:

$$\psi(u) = EX\frac{\beta^\alpha e^{-\frac{u}{EX}}}{c}(\beta - u/c)^{-1-\alpha}\left(\alpha - \frac{\Gamma(\alpha+1, \frac{1}{EX}(c\beta - u))}{\Gamma(\alpha)}\right) + \frac{\Gamma(\alpha, \frac{\beta c}{EX})}{\Gamma(\alpha)}, \quad u \geqslant 0. \tag{5.7.3}$$

W szczególności dla u=0 mamy:

$$\psi(0) = EX\frac{1}{\beta c}\left(\alpha - \frac{\Gamma(\alpha+1, \frac{c\beta}{EX})}{\Gamma(\alpha)}\right) + \frac{\Gamma(\alpha, \frac{c\beta}{EX})}{\Gamma(\alpha)}$$

i

$$\lim_{u\to\infty}\psi(u) = \frac{\Gamma(\alpha, \frac{c\beta}{EX})}{\Gamma(\alpha)}.$$

W tym przypadku również podejmiemy próbę porównania wyników z modelem niezależnym. Dla obu modeli przyjmujemy $\frac{1}{EX} = \frac{1}{3}$, $c = 1$. Zakładam również, że wartości oczekiwane odstępów czasowych w modelu zależnym i niezależnym będą równe tj.: $\frac{1}{\lambda} = \frac{\alpha\beta}{\alpha-1}$ oraz nie zostanie złamany warunek zysku netto tzn.: $\lambda < \frac{c}{EX} = 3$. Parametry, które spełniają wszystkie powyższe warunki to np.: $\lambda = 2$, $\alpha = 3$, $\beta = \frac{1}{3}$. Poniżej przedstawiam wzory na prawdopodobieństwo ruiny dla modelu zależnego i niezależnego, obliczone dla wybranych parametrów:

Prawdopodobieństwo ruiny w modelu niezależnym:

$$\psi(u) = min\{\frac{2}{3}exp\{-u\}, 1\}$$

Prawdopodobieństwo ruiny w modelu z zależnymi odstępami czasowymi:

$$\psi(u) = \frac{(\frac{1}{3})^3 e^{-3}}{3}(\frac{1}{3} - u)^{-4}\left(3 - \frac{\Gamma(4, 1-3u)}{\Gamma(3)}\right) + \frac{\Gamma(3,1)}{\Gamma(3)}$$

Potrzebne obliczenia zostały wykonane w Matlabie:

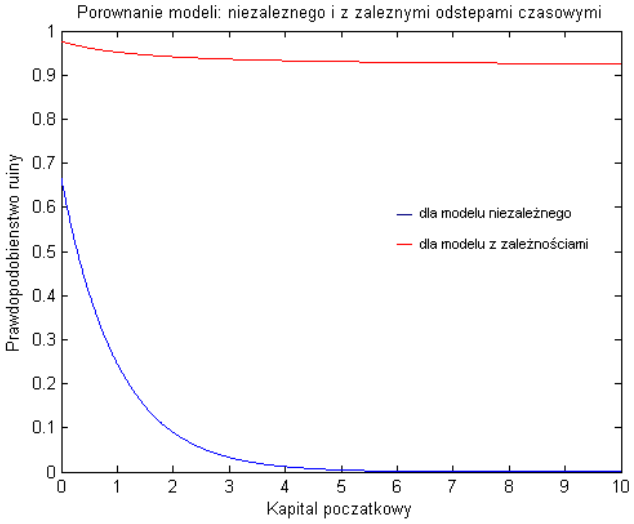
$$\Gamma(3,1) = \int_1^\infty x^2 e^{-x}dx = 5e^{-1}$$

$$\Gamma(4, 1-3u) = \int_{1-3u}^\infty x^3 e^{-x}dx = e^{-(1-3u)}\left((1-3u)^3 + 3(1-3u)^2 + 6(1-3u) + 6\right)$$

Ostateczny wzór dla modelu zależnego to:

$$\psi(u) = \frac{e^{-3u}}{3^4}(\frac{1}{3} - u)^{-4}\left(3 - \frac{e^{-(1-3u)}}{2}((1-3u)^3 + 3(1-3u)^2 + 6(1-3u) + 6)\right) + \frac{5}{2e}$$

Wzory zostały naniesione na wykres.



Otrzymane prawdopodobieństwo ruiny w modelu z zależnościami, dla każdego $u \geqslant 0$, jest dużo wyższe niż w modelu niezależnym. Poniżej obliczę prawdopodobieństwa ruiny w obu modelach dla przykładowych wartości u:

Model niezależny:

$$\psi(0) = min\{\frac{2}{3}, 1\} = \frac{2}{3} = 0,(6)$$

$$\psi(5) = min\{\frac{2}{3}exp\{-5\}, 1\} = \frac{2}{3} \cdot e^{-5} \approx 0,0045$$

Model zależny:

$$\psi(0) = 3 - \frac{\Gamma(4,1)}{\Gamma(3)} + \frac{\Gamma(3,1)}{\Gamma(3)} = \approx 0,9767$$

$$\psi(5) = \frac{e^{-3\cdot 5}}{(3\cdot(\frac{1}{3}-5))^4}\left(3 - \frac{e^{-(1-3\cdot 5)}}{2}((1-3\cdot 5)^3 + 3(1-3\cdot 5)^2 + 6(1-3\cdot 5) + 6)\right) + \frac{5}{2e} \approx 0,9304$$

5.7.4 Odstępy między roszczeniami o rozkładzie Weibulla z funkcją Copula Gumbel'a

Rozważmy model, w którym zmienna losowa Λ (nasz czynnik "frailty") ma rozkład stabilny (1/2) z gęstością:

$$f_\Lambda(\lambda) = \frac{\alpha}{2\sqrt{\pi\lambda^3}}e^{-\alpha^2/4\lambda}, \qquad \lambda > 0$$

Wtedy, analogicznie do przykładu (4.2), otrzymany rozkład brzegowy odstępów czasowych T_k to:

$$\bar{F}_T(t) = \int_0^\infty e^{-\lambda t} f_\Lambda(\lambda) d\lambda = exp\{-\alpha t^{1/2}\}, \qquad t \geqslant 0$$

Oznacza to, że odstępy czasowe mają rozkład Weibulla z parametrem kształu równym 1/2, natomiast ich struktura zależności, na podstawie twierdzenia (**??**), jest opisana funkcją copula Gumbela tj. funkcją copula z rodziny copul Archimedesa z generatorem $\varphi(t) = (-\ln t)^2$. Prawdopodobieństwo ruiny, dla tego modelu, na podstawie wzoru (**??**) to:

$$\psi(u) = \frac{\alpha i e^{-i\alpha\sqrt{u/c}-u\theta}}{4\theta\sqrt{cu}} \left(-1 + Erf\left(\frac{\alpha}{2\sqrt{c\theta}} - i\sqrt{u\theta}\right) + e^{2i\sqrt{u/c}\alpha} Erfc\left(\frac{\alpha}{2\sqrt{c\theta}} + i\sqrt{u\theta}\right)\right) + Erfc\left(\frac{\alpha}{2\sqrt{c\theta}}\right) \tag{5.7.4}$$

gdzie $i = \sqrt{-1}$. Można wykazać, że otrzymana część urojona z prawej strony równania (5.7.4) jest równa zero, co oznacza, że wyrażenie jest w istocie liczbą rzeczywistą. Dla u=0 otrzymujemy:

$$\psi(0) = \left(1 - \frac{\alpha^2}{2c\theta}\right) Erfc\left(\frac{\alpha}{2\sqrt{c\theta}}\right) + \frac{\alpha}{\sqrt{\theta\pi c}} e^{-\alpha^2/4c\theta}$$

dla dużych u: $lim_{u\to\infty}\psi(u) = Erfc(\frac{\alpha}{2\sqrt{c\theta}})$.

W celu porównania prawdopodobieństwa ruiny opisywanego modelu z modelem niezależnym, należy ustalić parametry λ, α, EX, c. Przy czym, wartości oczekiwane odstępów czasowych w obu modelach powinny być równe, tj.: $\frac{1}{\lambda} = \frac{2}{\alpha^2}$ oraz musi być spełniony warunek zysku netto tj.: $\lambda < \frac{c}{EX}$. Wybrane wartości to: $\lambda = \alpha = 2$, $\frac{1}{EX} = 3$, $c = 1$. Poniżej przedstawiam wzory, dla obu modeli, dla wybranych parametrów:

Model niezależny:

$$\psi(u) = min\{\tfrac{2}{3} exp\{-u\}, 1\}$$

$$\psi(0) = min\{\tfrac{2}{3} exp\{0\}, 1\} = \tfrac{2}{3} = 0,(6)$$

$$lim_{u\to\infty}\psi(u) = 0$$

Model zależny:

$$\psi(u) = \frac{2ie^{-i2\sqrt{u}-3u}}{12\sqrt{u}} \left(-1 + Erf\left(\frac{2}{2\sqrt{3}} - i\sqrt{3u}\right) + e^{2i\sqrt{u}2} Erfc\left(\frac{2}{2\sqrt{3}} + i\sqrt{3u}\right)\right) + Erfc\left(\frac{2}{2\sqrt{3}}\right)$$

$$\psi(0) = \left(1 - \tfrac{2}{3}\right) Erfc\left(\tfrac{1}{\sqrt{3}}\right) + \tfrac{2}{\sqrt{3\pi}} e^{-\frac{1}{3}} \approx 0,6042$$

$$lim_{u\to\infty}\psi(u) = Erfc(\tfrac{1}{\sqrt{3}}) \approx 0,4122$$

Dla $u = 0$ prawdopodobieństwo ruiny w modelu niezależnym jest większe, jednak wraz ze wzrostem u, dąży ono do 0. Natomiast dolna granica prawdopodobieństwa ruiny w modelu zależnym to w przybliżeniu 0,4122, co oznacza, że dla dużych u, prawdopodobieństwo ruiny w tym modelu jest większe niż w niezależnym.

Uwaga

Dla ciężko-ogonowych rozkładów odstępów czasowych między roszczeniami zwykłe techniki, takie jak łańcuchy Markowa nie działają, a w literaturze nie jest dostępny żaden jawny wzór na prawdopodobieństwo ruiny. W konsekwencji, wzory (5.7.3) i (5.7.4) mogą być pierwszymi przykładami jawnych wzorów na prawdopodobieństwo ruiny dla identycznych, ciężkoogonowych rozkładów odstępów czasowych między roszczeniami, z zależnościami opisanymi odpowiednio przez funkcję copula Claytona i Gumbela.

5.8 Dalsze rozszerzenia metody mieszania

Pomysł mieszania przedstawiony w obecnej pracy może być kontynuowany na różne sposoby. Oto kilka z nich:

Przykład 5.8.1 Jednym ze sposobów jest jednoczesne mieszanie odstępów czasowych między roszczeniami oraz wielkości roszczeń, niezależnie od siebie. Prowadzi to w obu przypadkach do funkcji copula Archimedesa. Prawdopodobieństwo ruiny wyraża się wówczas wzorem:

$$\psi(u) = \int_0^\infty \int_0^\infty \psi_{(\theta,\lambda)}(u) dF_\Theta(\theta) dF_\Lambda(\lambda)$$

gdzie $\psi_{(\theta,\lambda)}(u)$ to warunkowe prawdopodobieństwo ruiny dla $\Theta = \theta$ i $\Lambda = \lambda$. Jawny wzór na $\psi_{(\theta,\lambda)}(u)$, zawsze doprowadzi nas do otrzymania jawnego wzoru na $\psi(u)$ w modelach odnowy z zależnymi zarówno odstępami czasowymi jak i wielkościami roszczeń.

Przykład 5.8.2 Istnieje również możliwość wprowadzenia zależności Archimedesa pomiędzy wielkości roszczeń, a odstępy czasowe, w tym samym czasie wprowadzając zależności pomiędzy odstępami czasowymi i zależność pomiędzy wielkościami roszczeń. Sposobem na to jest comotonic mixing, gdzie realizacja λ jest funkcją deterministyczną realizacji θ w postaci:

$$\lambda(\theta) = F_\Lambda^{-1}(F_\Theta(\theta)).$$

Prawdopodobieństwo ruiny w takim modelu jest opisane wzorem:

$$\psi(u) = \int_0^\infty \psi_{(\theta,\lambda(\theta))}(u) dF_\Theta(\theta),$$

gdzie $\psi_{(\theta,\lambda)}(u)$ to warunkowe prawdopodobieństwo ruiny dla $\Theta = \theta$ i $\Lambda = \lambda$.

Przykład 5.8.3 Odwołując się do podrozdziału (4.1), w celu uzyskania wzoru na prawdopodobieństwo ruiny dla pewnych, całkowicie monotonicznych rozkładów brzegowych roszczeń, konieczne było ustalenie ich struktury zależności. W rzeczywistości istnieje sposób, aby zmienić strukturę zależności, pozostawiając jednocześnie asymptotyczny ogon rozkładu brzegowego roszczenia niezmieniony i nadal otrzymywać jawne wzory.Niech proces nadwyżki będzie opisany wzorem:

$$R(t) = u + ct - \sum_{k=1}^{N(t)} X_k - \sum_{l=1}^{N'(t)} Y_l$$

gdzie są dwa niezależne procesy Poisson'a: $N'(t)$ z intensywnością λ' generujący niezależne roszczenia $Y_1, Y_2, ...$ o rozkładzie wykładniczym $Exp(\nu)$, gdzie ν jest ustaloną stałą. Z kolei proces N(t) z intensywnością λ generuje strumień zależnych roszczeń $X_1, X_2, ...$, tak jak w (5.6.1), gdzie Θ jest dodatnią zmienną losową. Jeśli rozkład zmiennej Θ jest np. taki jak w podrozdziale (4.1) lub w podrozdziale (4.2), to zależne roszczenia $X_1, X_2, ...$ są ciężkoogonowe. Równoważnie możemy przedstawić powstały proces ryzyka jako:

$$R(t) = u + ct - \sum_{k=1}^{N''(t)} Z_k,$$

gdzie $N''(t)$ jest procesem Poisson'a o intensywności $\lambda+\lambda'$ oraz brzegowe rozkłady wielkości roszczeń $Z_1, Z_2, ...$ są wyrażone jako:

$$f_Z(x) = \frac{\lambda}{\lambda+\lambda'}\theta e^{-\theta x} + \frac{\lambda'}{\lambda+\lambda'}\nu e^{-\nu x}, \quad x \geqslant 0. \tag{5.8.1}$$

Oznacza to, że rozkład wielkości roszczeń jest mieszanką dwóch rozkładów wykładniczych, ale wymieszany jest po wartośćiach parametru θ. Brzegowy rozkład $X_1, X_2, ...$ określa zachowanie ogona, jeśli jest on ciężkoogonowy. Ponieważ dla ustalonego θ, prawdopodobieństwo ruiny $\psi_\theta(u)$ w klasycznym modelu ryzyka z rozkładem wielkości roszczeń z równania (5.8.1) ma jawną formę jako ważona suma dwóch wykładniczych wyrażeń, po raz kolejny otrzymywany jest jawny wzór na prawdopodobieństwo ruiny na mocy (5.6.3). Należy pamiętać, że mieszanie strumienia zależnych roszczeń $X_1, X_2, ...$ z niezależnymi (w tym przypadku wykładniczymi) roszczeniami zredukuje zależności. W szczególności, jeśli tau Kendall'a dwóch dowolnych roszczeń X_i, X_j wynosi τ_0, to wartość tau Kendall'a dla dwóch dowolnych roszczeń Z_i, Z_j w nowym modelu wynosi $\tau_1 = (\frac{\lambda}{\lambda+\lambda'})^2\tau_0$, ponieważ jedyny sposób na dodatnią korelację pomiędzy losowo wybranych wielkościami roszczeń jest wybrać dwa roszczenia pochodzące z procesu $N(t)$ i ponieważ prawdopodobieństwo, że roszczenie wygenerowane przez $N''(t)$ pochodzi z $N(t)$ jest równe $\frac{\lambda}{\lambda+\lambda'}$. W ten sposób określiliśmy metodę, w której zależności pomiędzy wielkościami roszczeń powinny zostać osłabione, zachowanie brzegowego ogona powinno zostać takie jak opisane w procesie $N(t)$, a przez odpowiednie wybranie parametru λ', można wygenerować jawne wzory dla modeli dla każdej wartości τ_1 pomiędzy 0 i τ_0. Należy pamiętać, że odbywa się to kosztem utraty jawnego typu zależności Archimedesa, który był dostępny dla $\lambda' = 0$. Zamiast tego, ogon łącznej dystrybuanty zmiennych $Z_1, Z_2, ...$ otrzymano w następujący sposób:

każdy z zaistniałych roszczeń Z_i jest typu X_i z prawdopodobieństwem $\frac{\lambda}{\lambda+\lambda'}$ oraz typu Y_i w pozostałych przypadkach. Wówczas łączna dystrybuanta jest odpowiednią mieszanką niezależnych typów Y_i oraz struktury zależności Archimedesa typów X_i.

W całkowicie analogiczny sposób można zmieniać strukturę zależność pomiędzy odstępami czasowymi, w dalszym ciągu otrzymując jawne wzory na prawdopodobieństwa ruiny.

Uwaga

Model ryzyka otrzymany w powyższym przykładzie obejmuje dwa rodzaje roszczeń: niezależne, lekko-ogonowe roszczenia oraz zależne, ciężko-ogonowe (na odpowiednich założeniach dla θ). W zakresie praktycznej interpretacji, może to rzeczywiście opisywać realną sytuację portfela: zależność pomiędzy ciężko-ogonowymi wielkościami roszczeń może pochodzić z niepewności parametrów, lub z innych źródeł korelacji jak ryzyko środowiskowe, zmiany klimatu lub ryzyko prawne. Rzeczywiście, wiele modeli wewnętrznych dla Solvency II współpracuje z regularnie zmieniającymi się losowymi stratami agregowanymi przez kopułę przeżycia Claytona.

Uwaga

Analizując przykład (5.8.3), staje się oczywista możliwość wygenerowania wiele więcej przykładów jawnych wzorów dla modeli ryzyka z zależnoscią przez zastąpienie dystrybuanty rozkładu wykładniczego bardziej ogólną dystrybuantą dla której wzór na $\psi_\theta(u)$ jest jawnie określony. Otrzymana struktura zależności oraz brzegowe rozkłady wielkości roszczeń będą wynikiem oddziaływania pomiędzy tym wyborem, a rozkładem zmiennej Θ (analogicznie do równania (5.6.1).

Rozdział 6

Techniki statystyczne dla rozkładów ciągłych

W rozdziale tym zajmiemy się opisem rozkładów ciągłych, które służą do modelowania wielkości szkód oraz metodami statystycznymi pozwalającymi zidentyfikować owe rozkłady na podstawie zgromadzonych danych.

Najpopularniejsze rozkłady ciągłe używane w matematyce ubezpieczeniowej (nie tylko w kontekście wielkości szkód) to:

- rozkład normalny $N(\mu, \sigma^2)$
- rozkłady fazowe, m.in.
  - rozkład wykładniczy $Exp(\lambda)$;
  - rozkład Gamma $\Gamma(\alpha, \beta)$;
- rozkład logarytmiczno-normalny $LN(\mu, \sigma)$;
- rozkład Pareto $Par(\alpha, c)$;
- rozkład Weibulla

Podstawowe wiadomości o rozkładach normalnym, wykładniczym, Gamma, czy logarytmiczno-normalnym, wraz z estymacją parametrów, można znaleźć w Dodatku. Tutaj zajmiemy się głównie rozkładami Pareto i ich modyfikacjami, szczególnie w kontekście dość skomplikowanej estymacji parametru α.

Na początek omówimy pewne techniki statystyczne służące do dopasowanie rozkładu do danych.

6.1 Dopasowanie rozkładu do danych

Następujące funkcje mogą służyç do oceny dopasowania modelu empirycznego do wybranego modelu teoretycznego. Niech $X_1, \ldots, X_n$ będzie skoǫczonym ciągiem zmiennych losowych iid.

- **Dystrybuanta empiryczna**

$$\widehat{F}_n(x) = \frac{1}{n}\#\{i = 1, ..., n : X_i \leqslant x\}.$$

  Tak zdefiniowana funkcja zmiennej x jest funkcją losową, zależącą od zmiennych X_i.

- **Empiryczna funkcja kwantylowa**

$$\widehat{Q}_n(p) = X_i^* \text{jeżeli} p \in \left(\frac{i-1}{n}, \frac{i}{n}\right],$$

  gdzie $X_1^* \leqslant ... \leqslant X_n^*$ są uporządkowanymi wielkościami szkód (statystykami porządkowymi

  Funkcja ta jest empirycznym odpowiednikiem **funkcji kwantylowej**

$$Q(p) := F^{-1}(p),$$

  gdzie $p \in (0, 1]$ oraz $F^{-1}(p) := \inf\{t : F(t) \geqslant p\}$ (uogǫlniona funkcja odwrotna, lewostronnie ciągła). Oczywicie $\widehat{Q}_n(p) = \widehat{F}_n^{-1}(p),\ p \in (0, 1]$.

- **Empiryczna funkcja nadwyżki** $e_{\hat{F}_n}(u) = \frac{\sum_{i=1}^n (X_i - u) I(u < X_i)}{\sum_{i=1}^n I(u < X_i)}$. Wielkość ta będzie przybliżała prawdziwą średnią funkcję nadwyżki.

6.1.1 Dystrybuanta empiryczna

Z twierdzenia Kołmogorowa-Smirnowa wiadomo, że z prawdopodobieństwem 1, losowa funkcja dystrybuanty empirycznej spełnia

$$D_n = \sup_x |\widehat{F}_n(x) - F(x)| \to 0,$$

$n \to \infty$. Możemy się spodziewać, że przy dużej liczbie obserwacji dystrybuanta empiryczna przybliża prawdziwy rozkład szkody F. Odległość D_n jest zmienn losową i przy n dążącym do nieskończoności jej rozkład dąży do znanego rozkładu (twierdzenie Komogorowa). Rozkład ten jest otrzymany w języku tak zwanego mostu Browna i jego maksimum. Stanowi on podstawę testu Smirnowa-Komogorowa, który pozwala statystycznie ocenić, czy dana dystrybuanta empiryczna reprezentuje wybrany rozkład.

Przykład 6.1.1 Rozważamy próbkę dla $n = 100$ liczb $X_1, ..., X_{100}$ z rozkładu o gęstości $f(x) = \exp(-(x - 1))$, $x > 1$.

Do dystrybuanty empirycznej dopasowujemy najpierw rozkład przesunięty wykładniczy, a potem (dla porównania) Pareto z parametrami wyestymowanymi z próbki.

Ponieważ $\mathrm{E}\,[X_1] = \frac{1}{\lambda} + 1$, więc metodą momentów estymujemy parametr λ

$$\widehat{\lambda} = \frac{1}{\sum_{i=1}^{n} X_i/n - 1}.$$

Na jednym rysunku (lewy) przedstawiamy dystrybuantę empiryczną $\widehat{F}_n$ i przesunięty rozkład wykładniczy o gęstości $\widehat{\lambda}\exp(-\widehat{\lambda}(x-1))$.

Rozkad Pareto definujemy tutaj jako rozkład o gęstości

$$f(x) = x^{-(\alpha+1)} I_{\{x>1\}},$$

dla $\alpha > 1$.

W przypadku rozkładu Pareto jest wiele metod estymacji parametru α. Uyjemy estymatora otrzymanego metod najwikszej wiarogodnoc. Estymujemy

$$\widehat{\alpha} = \frac{n}{\sum \log X_i}$$

i przedstawiamy $\widehat{F}_n$ wspólnie z rozkładem $Par(\widehat{\alpha}, 1)$. [1]

□

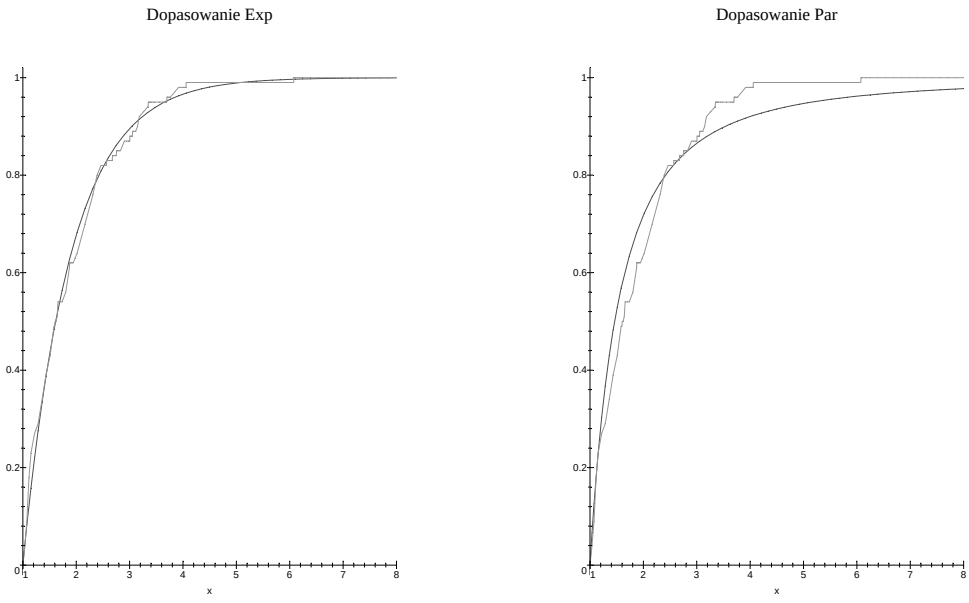
6.1.2 Wykres kwantylowy (Q-Q plot)

Idea wykresu kwantylowego jest następująca: wybierzmy rozkład F. Nanosimy na jednym układzie współrzędnych pary punktów postaci $(Q(p), \widehat{Q}_n(p))$, $p \in (0,1)$. Najbardziej typowym wyborem punktów p jest $\frac{1}{n}, \frac{2}{n}, ..., \frac{n-1}{n}, 1$. Wtedy $\widehat{Q}_n(\frac{i}{n}) = X_i^*$. Z pewnych przyczyn technicznych wybiera się $\frac{1}{n+1}, \frac{2}{n+1}, ..., \frac{n-1}{n+1}, \frac{n}{n+1}$. Pary punktów będą miały więc współrzędne: (kwantyl teoretyczny, kwantyl próbkowy). Jeżeli teraz utworzona wykres będzie w przybliżeniu linią prostą, będzie to oznaczało, że kwantyle próbkowe są bliskie kwantylom teoretycznym, a więc nasze dane pasują do wybranego przez nas modelu teoretycznego.

Przykład 6.1.2 Wykładniczy wykres kwantylowy: Dla rozkładu wykładniczego $Exp(\lambda)$ mamy: $F(x) = 1 - \exp(-\lambda x)$,

$$Q(p) = -\frac{1}{\lambda}\log(1-p).$$

[1] ciagle-1.mws



Wykres kwantylowy będzie miał więc postać:

$$\{(-\log(1-\frac{i}{n+1}), X_i^*), i = 1..., n\}.$$

W przypadku, gdy otrzymalismy wykres w przybliżeniu liniowy, to ma sens dopasowanie "najlepszej" funkcji liniowej $y = ax$ metodą najmniejszych kwadratów. Współczynnik nachylenia a ma wtedy postać

$$\widehat{a} = \frac{\sum_{i=1}^n X_i^* q_i}{\sum_{i=1}^n q_i^2},$$

gdzie $q_i = -\log(1-\frac{i}{n+1})$, i jest oszacowaniem parametru $1/\lambda$.

Weibullowski wykres kwantylowy: Dla rozkładu $Weibull(\lambda, r)$ mamy: $F(x) = 1 - \exp(-\lambda x^r)$,

$$Q(p) = \left(-\frac{1}{\lambda}\log(1-p)\right)^{1/r}$$

lub alternatywnie $\log Q(p) = \frac{1}{r}\log(\frac{1}{\lambda}) + \frac{1}{r}\log(-\log(1-p))$. Wykres kwantylowy będzie miał więc postać:

$$\{(\log(-\log(1-\frac{i}{n+1})), \log(X_i^*)), i = 1..., n\}.$$

W przypadku, gdy otrzymalismy wykres w przybliżeniu liniowy, to nachylenie prostej stanowi aproksymację parametru $1/r$, natomiast minimalna wartość na osi x przybliża

$\log(1/\lambda)/r$.

Log-normalny wykres kwantylowy: Niech Φ oznacza dystrybuantę standardowego rozkładu normalnego. Lognormalny wykres kwantylowy jest postaci

$$\{\Phi^{-1}(\frac{i}{n+1}), \log(X_i^*)), i = 1..., n\}.$$

Nachylenie krzywej daje aproksymację dla σ, a minimalna wartość na osi poziomej przybliża wartość μ.

Wykres kwantylowy Pareto: Dla rozkładu $Par(\alpha)$ mamy:

$$Q(p) = (1-p)^{-1/\alpha} - 1$$

lub alternatywnie

$$\log Q(p) = -\frac{1}{\alpha}\log(1-p).$$

Rysunek kwantylowy będzie miał więc postać: $\{(-\log(1-\frac{i}{n+1}), \log(X_i^*)), i = 1..., n\}$. W przypadku, gdy otrzymalismy wykres w przybliżeniu liniowy, to nachylenie prostej stanowi aproksymację parametru $1/\alpha$.

Następnie, do wygenerowanych liczb z rozkładu wykładniczego, stosujemy Weibullowski wykres kwantylowy, log-normalny wykres kwantylowy, wykres kwantylowy Pareto. Żaden wykres nie jest 'linią prostą'.

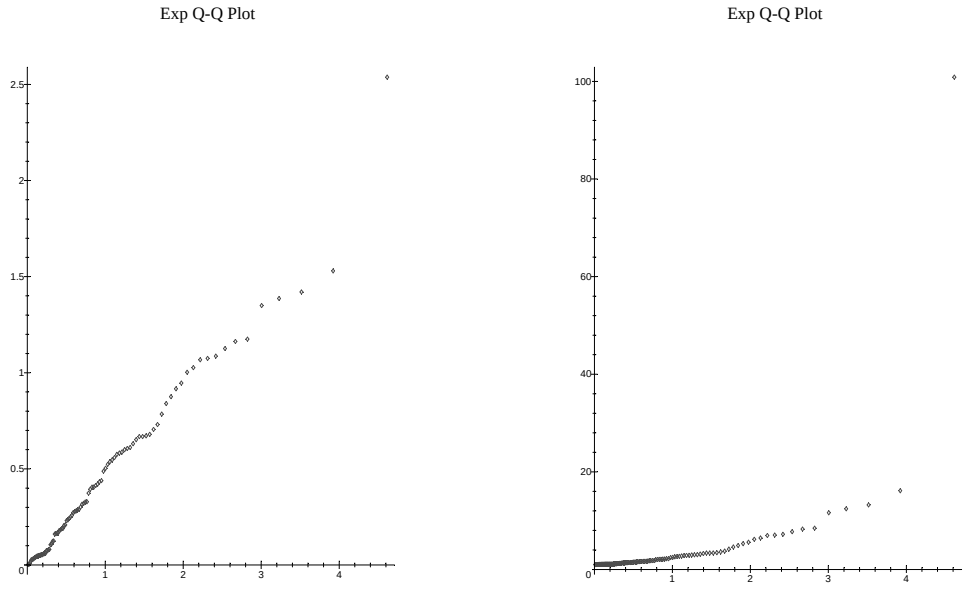
Na podstawie Q-Q plots odrzucimy na pewno rozkład Pareto, raczej też odrzucimy rozkład log-normalny, nie będziemy jednak w stanie odrzucić 'hipotezy' o tym, że rozkład jest Weibulla.[2]

□

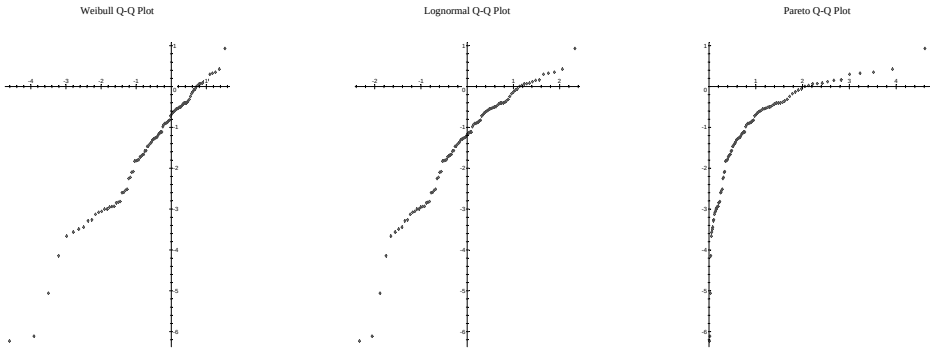
Przykład 6.1.3 Do wygenerowanych liczb z rozkładu Pareto, stosujemy Weibullowski wykres kwantylowy, log-normalny wykres kwantylowy, wykres kwantylowy Pareto. Tylko ostatni rysunek jest 'linią prostą'.[3]

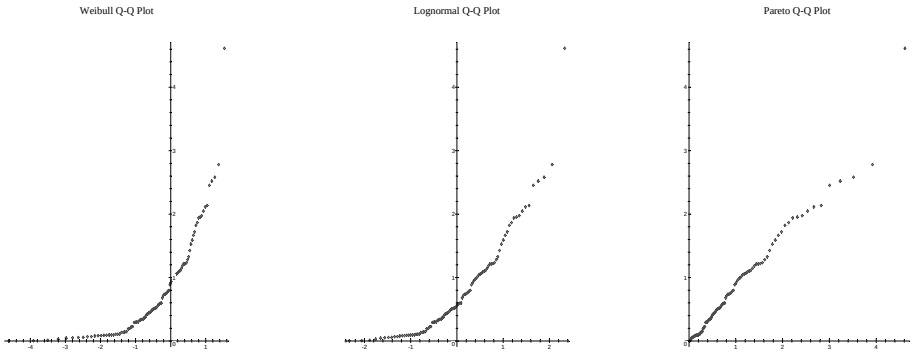
□

[2]Plik ciagle-2a.mws
[3]Plik ciagle-2b.mws



Rysunek 6.1.1: Wykładniczy wykres kwantylowy dla 100 danych z rozkładów $Exp(2)$ i $Par(1.1, 1)$. Lewy wykres pokazuje dobre dopasowanie do rozkładu wykładniczego, podczas gdy prawy wykres pokazuje, że dane zostały z rozkładu 'cięższego' niż wykładniczy.





6.1.3 Średnia funkcja nadwyżki

Innym sposobem diagnostyki jest tzw. średnia funkcja nadwyżki (*mean excess function*))

$$e_F(u) = \mathrm{E}\left[X - u | X > u\right].$$

Funkcja ta jednoznacznie wyznacza dystrybuantę. Bezpośrednio z definicji otrzymujemy

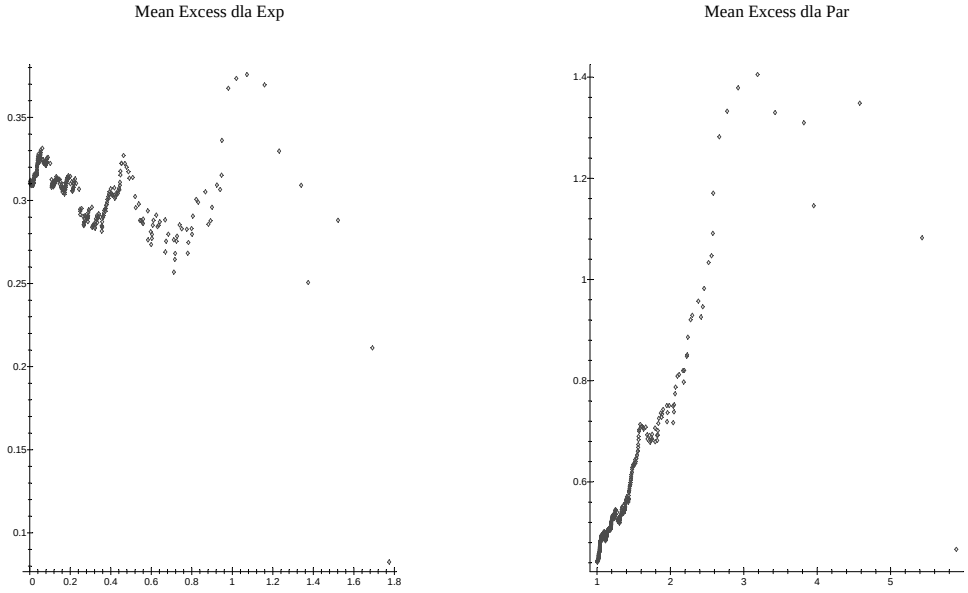
$$\begin{aligned} e_F(u) &= \frac{\mathrm{E}\left[(X-u)I(X>u)\right]}{P(X>u)} \\ &= \frac{\int_u^\infty \overline{F}(s)ds}{\overline{F}(u)} = \frac{\mathrm{E}\left[(X-u)_+\right]}{P(X>u)}. \end{aligned}$$

przy założeniu, że $\mathrm{E}[X] < \infty$. W praktyce estymujemy tą funkcję próbkową funkcją nadwyżki $e_{\hat{F}_n}(u)$. Funkcję $e_{\hat{F}_n}(u)$ rysuje się w punktach $u = X^*_{n-k}$, przyjmuje ona wtedy postać $e_{\hat{F}_n}(X^*_{n-k}) = \frac{1}{k}\sum_{i=1}^k X^*_{n-i+1} - X^*_{n-k}$, $k = 1, ..., n-1$. Dla rozkładu wykładniczego $Exp(\lambda)$ mamy $e_F(u) = \frac{1}{\lambda}$. W przypadku, gdy dane pochodzą z rozkładu cięższego niż wykładniczy, to funkcja nadwyżki jest rosnąca, w przeciwnym razie jest malejąca.

6.2 Rozkład Pareto

Nazwa tej klasy rozkładów wzięła się od nazwiska Vilfredo Pareto (1897), szwajcarskiego ekonomisty, który bogactwo w populacji opisywał za pomocą $M = Ax^{-\alpha}$, gdzie M jest iloscią osób, które mają dochód większy niż x. Jeżeli $\overline{F}(x) = P(X > x) = c^\alpha x^{-\alpha}$, $x > c$, to mówimy, że zmienna losowa X ma **rozkład Pareto** $Par(\alpha, c)$. Jeżeli $c = 1$ to będziemy pisali $Par(\alpha)$. Gęstość zadana jest wzorem $f(x) = c^\alpha x^{-(\alpha+1)}$, $x > c$. Rozklad taki ma średnią

$$\mathrm{E}[X] = \frac{c\alpha}{\alpha - 1}. \tag{6.2.1}$$



Rysunek 6.1.2: Funkcja nadwyżki dla 300 danych z rozkładów $Exp(3)$ i $Par(3,1)$.

Aby powyższa tożsamość miała sens, musimy założyć $\alpha > 1$.

Rozkład Pareto jest tzw. **rozkładem ciężkoogonowym**. W szczególności, dla $\alpha > k$ nie istnieje k-ty moment.

Dla próby $X_1, \ldots, X_n$ z rokładu $Par(\alpha, c)$ estymujemy parametry za pomocą metody **największej wiarogodności**. Jeżeli $L(\alpha, c, x_1, \ldots, x_n) = \prod_{i=1}^n f(x_i)$, to

$$\log L(\alpha, c; X_1, \ldots, X_n) = n \log(\alpha) + n\alpha \log c - (\alpha + 1) \sum_{i=1}^n \log X_i.$$

Szukamy maksimum (względem α i c) funkcji L. Różniczkując względem α i c oraz przyrównując do zera, dostajemy układ równań

$$\alpha = \frac{n}{\sum_{i=1}^n \log(X_i/c)}, \tag{6.2.2}$$

$$\alpha = \frac{\alpha + 1}{n} \sum_{i=1}^n (1 + X_i/c)^{-1}, \tag{6.2.3}$$

który należy rozwiązać numerycznie.

Jeżeli chcemy uchronić się od procedur numerycznych możemy postąpić w następujący sposób. Parametr α estymujemy za pomocą wzoru (6.2.2), a następnie wstawiamy estymator c jako

$$\widehat{c} = \min(X_1, \ldots, X_n). \tag{6.2.4}$$

Postępowanie takie jest uzasadnione, gdyż musimy mieć $c \leqslant \min(X_1, \ldots, X_n)$ i w celu maksymalizacji funkcji wiarogodności kładziemy (6.2.4). Należy dodać, że tak otrzymany estymator dla c jest obciążony: $\mathrm{E}\,[\widehat{c}] = \frac{nc\alpha}{n\alpha - 1} \neq c$. Rzeczywiście, licząc rozkład minimum dostajemy

$$P(\min(X_1, \ldots, X_n) > x) = P(X > x)^n = c^{n\alpha} x^{-n\alpha},$$

a więc rozkład Pareto $Par(n\alpha, c)$ (patrz (6.2.1)).

Jeżeli chcemy uchronić się od procedur numerycznych lub dysponujemy tylko obserwacjami przekraczającymi poziom u możemy postąpić jeszcze w inny sposób: Jeżeli $X \sim Par(\alpha, \sigma)$, to rozkład warunkowy $X|X > u$ jest też Pareto, nie zależy od σ i jego dystrybuanta ma postać $1 - u^\alpha x^{-\alpha}$, czyli $Par(\alpha, u)$. Oznaczając przez $Y_i, i = 1, \ldots, n$, realizacje zmiennej losowej Y o powyższym rozkładzie warunkowym, funkcja wiarogodności ma postać

$$\ln L(\alpha; Y_1, \ldots, Y_n) = n \log(\alpha) + n\alpha \log u - (\alpha + 1) \sum_{i=1}^n \log Y_i.$$

Różniczkując względem α i przyrównując do zera dostajemy

$$\widehat{\alpha} = \frac{n}{\sum_{i=1}^n \log(Y_i/u)}. \tag{6.2.5}$$

Jeżeli nasze dane dotyczą nie tych obserwacji, które są większe od u, ale k największych obserwacji, to

$$\widehat{\alpha} = \frac{k}{\sum_{i=1}^{k} \log(X^*_{n+1-i}/X^*_{n-k})}, \tag{6.2.6}$$

gdzie $X^*_1 < X^*_2 < \cdots X^*_n$. W dalszej części będziemy nazywali estymator zadany wzorem (6.2.5) estymatorem Hilla typu I, a w drugim przypadku estymatorem Hilla typu II.

Parametr σ można wyestymować z zależności $\widehat{\sigma} = u/\widehat{\alpha}$. Jeżeli estymacje oparte są na k największych obserwacjach, to σ estymujemy z warunku

$$\widehat{\sigma} = X^*_{n-k} \left(\frac{k}{n}\right)^{1/\widehat{\alpha}}.$$

6.2.1 *Rozkłady typu Pareto

Mówimy, że L jest **funkcją wolno zmieniającą się**, jeżeli $\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1$ dla każdego $t > 0$. Typowe przykłady to $L(x) = 1$, $L(x) = \log x$, $L(x) = \log\log x$.

Rozkłady, dla których $\overline{F}(x) = x^{-\alpha}L(x)$, L jest funkcją wolno zmieniającą się w nieskończoności, nazywamy **rozkładami typu Pareto**. W klasie tej znajdują się m.in. omówione wcześniej rozkłady Pareto oraz rozkłady Burra, log-Gamma, log-logistyczny.

Przykład 6.2.1 Rozkład Burra ma ogon $\bar{F}(x) = \left(\frac{\beta}{\beta+x^\gamma}\right)^\lambda$, $\alpha, \gamma, \lambda, x > 0$. Rozkład o gęstości

$$f(x) = \frac{\beta^\alpha (\log x)^{\alpha-1} x^{-\beta-1}}{\Gamma(\alpha)}, \qquad x > 1,$$

gdzie $\alpha > 0$ nazywamy **rozkładem log-Gamma**. Rozkład o ogonie $\bar{F}(x) = (1+\beta x)^{-\alpha}$, $\alpha > 1$, $\beta, x >$, nazywamy **rozkładem log-logistycznym**.

□

Wykładnik α charakteryzuje tempo zbieżności ogona do 0 i będziemy go nazywali **indeksem Pareto**.

Zauważmy delikatną różnię pomiędzy rozkładami Pareto a typu Pareto. Z uwagi na obecność funkcji L mamy jedynie informację o asymptotycznym zachowaniu ogona $\bar{F}(x) = x^{-\alpha}L(x)$. Funkcja L jest zazwyczaj nieznana i niemżliwa do wyestymowania. Powoduje to problemy związane z estymacją parametru α, w przeciwieństwie do zwykłego rozkładu Pareto.

Parametr α można estymować w sposób następujący: Niech $\gamma = \frac{1}{\alpha}$. Wtedy za estymator γ kładziemy

$$H_{k,n} = \frac{1}{k}\sum_{j=1}^{k} \log(X^*_{n-j+1}) - \log(X^*_{n-k}), \qquad k = 1, \ldots, n.$$

Otrzymujemy wtedy zbiór tzw. estymatorów Hilla.

Problem, który się pojawia, to wybór wartości $k = k_{opt}$ dla której będziemy rozważali estymator Hilla $H_{k,n}$. Procedura wyboru będzie oparta na własnościach wykresu kwantylowego, który został opisany w Rozdziale 6.1.2. Wykres kwantylowy Pareto jest postaci $\{(-\log(1 - \frac{i}{n+1}), \log(X^*_i)), i = 1, ..., n\}$. Jeżeli dopasowanie do rozkładu Pareto jest dobre, to wykres kwantylowy jest w przybliżeniu liniowy i jego nachylenie można traktować jako estymator parametru γ. Estymator Hilla $H_{k,n}$ jest oszacowaniem nachylenia wykresu na prawo od punktu $(-\log(\frac{k+1}{n+1}), \log(X^*_{n-k}))$. Optymalną wartość k_{opt} (a tym samym $\gamma_{opt} := H_{k_{opt},n}$) wyznaczymy na podstawie linii regresji $l_1, \ldots, l_{n-1}$ przechodzących przez punkty, odpowiednio, $(-\log(\frac{k+1}{n+1}), \log(X^*_{n-k}))$, $k = 1, \ldots, n-1$ i mających nachylenie, odpowiednio, $H_{k,n}$ tak, aby średni błąd kwadratowy dopasowania prostej $l_{k_{opt}}$ do punktów $\{(-\log(\frac{j}{n+1}), \log(X^*_{n-j+1})), j = 1, \ldots, k_{opt}\}$ na prawo od $(-\log(\frac{k_{opt}+1}{n+1}), \log(X^*_{n-k_{opt}}))$ był minimalny. Równanie l_k ma postać

$$y = \log X^*_{n-k} + H_{k,n}\left(x + \log\left(\frac{k+1}{n+1}\right)\right). \tag{6.2.7}$$

Licząc średni błąd kwadratowy $MSE(k)$ dostajemy

$$\begin{aligned} MSE(k) \quad &: \quad = \frac{1}{k}\sum_{j=1}^{k}\left(\log(X^*_{n-j+1}) - \log X^*_{n-k} - H_{k,n}\left(-\log\left(\frac{j}{n+1}\right) + \log\left(\frac{k+1}{n+1}\right)\right)\right)^2 \\ &= \quad \frac{1}{k}\sum_{j=1}^{k}\left(\log\left(\frac{X^*_{n-j+1}}{X^*_{n-k}}\right) - H_{k,n}\log\left(\frac{k+1}{j}\right)\right)^2. \end{aligned}$$

Wartość optymalną k_{opt} wyznaczamy z warunku

$$MSE(k_{opt}) = \min_k MSE(k).$$

Ponieważ estymator Hilla jest asymptotycznie normalny, otrzymujemy przedział ufności na poziomie p dla indeksu Pareto:

$$(\gamma^-_{opt}, \gamma^+_{opt}) = \left(\frac{\gamma_{opt}}{1 + \frac{z_{p/2}}{\sqrt{k_{opt}}}}, \frac{\gamma_{opt}}{1 - \frac{z_{p/2}}{\sqrt{k_{opt}}}}\right). \tag{6.2.8}$$

Wykres kwantylowy służy także do estymacji kwantyli. Kładąc w (6.2.7) z $k = k_{opt}$ oraz $x = -\log p$ możemy traktować y jako estymator $\log Q(1-p)$ lub $\exp(y)$ jako estymator $Q(1-p)$, oznaczmy go przez $\widehat{q}_{n,p}$.

$$\log \widehat{q}_{n,p} = \log X^*_{n-k} + H_{k,n}\log\left(\frac{k+1}{(n+1)p}\right).$$

Daje to

$$\widehat{q}_{n,p} = X^*_{n-k} \left(\frac{k+1}{(n+1)p} \right)^{H_{k,n}}, \tag{6.2.9}$$

gdzie $\widehat{q}_{n,p}$ jest kwantylem rzędu p opartym na próbce n elementowej. Używając przedziału ufności dla γ_{opt} dostajemy analogiczne oszacowanie dla $\widehat{q}_{n,p}$:

$$\left(X^*_{n-k} \left(\frac{k+1}{(n+1)p} \right)^{\gamma^-_{opt}}, X^*_{n-k} \left(\frac{k+1}{(n+1)p} \right)^{\gamma^+_{opt}} \right). \tag{6.2.10}$$

Podobnie jak wyżej dokonujemy estymacji $P(X > x)$:

$$\left(\frac{k+1}{n+1} \right) \left(\frac{x}{X^*_{n-k}} \right)^{-1/H_{k,n}}. \tag{6.2.11}$$

Przykład 6.2.2 [4]Wygenerowano 100 liczb z rozkładu $Par(1.5, 1)$. Mamy wtedy $\gamma = 1/\alpha = 0.66$. Przedstawimy najpierw wykres dla estymatora Hilla:

Wyliczamy wartości $MSE(k)$ dla $k = 1, \ldots, n$. Z rysunku odczytujemy, że k_{opt} jest pomiędzy 90 a 100. Wyliczając: $k_{opt} = 95$.

Wartość $H_{k_{opt},n} = \gamma_{opt} = 0.5867$ daje optymalny wybór dla estymatora parametru γ. Ze wzoru (6.2.8) dostajemy dla $p = 0.05$: $[0.48, 0.73]$, a więc prawdziwa wartość 0.66 należży do przedziału ufności.

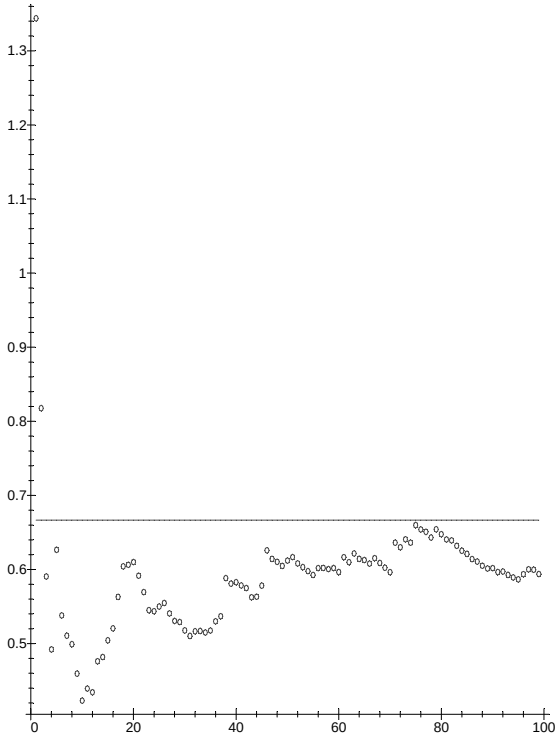
Dalej, ze wzorów (6.2.9), (6.2.10) dostajemy kwantyle wraz z przedziałem ufności:

Następnie estymujemy wartości $P(X > x)$ za pomocą wzoru (6.2.11) (kółka) i $P(X > x) = x^{-\alpha_{opt}}$(linia ciągła). Wartości te pokrywają się.

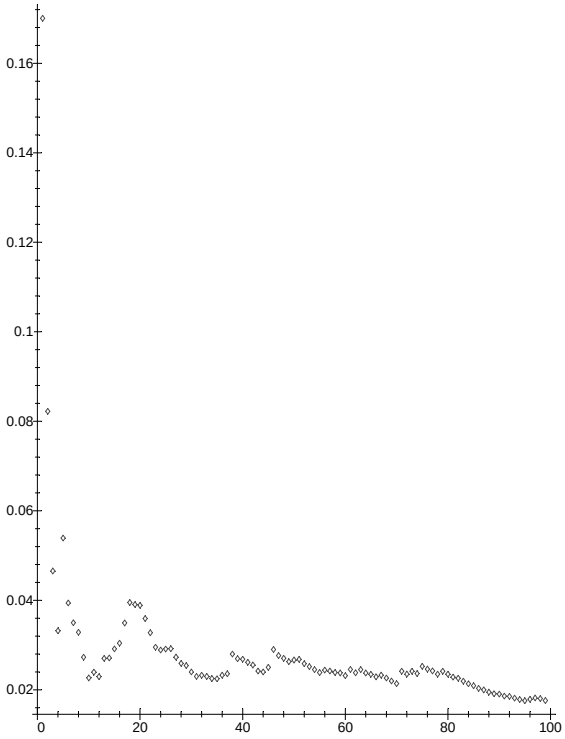
□

[4]ciagle-4.mws

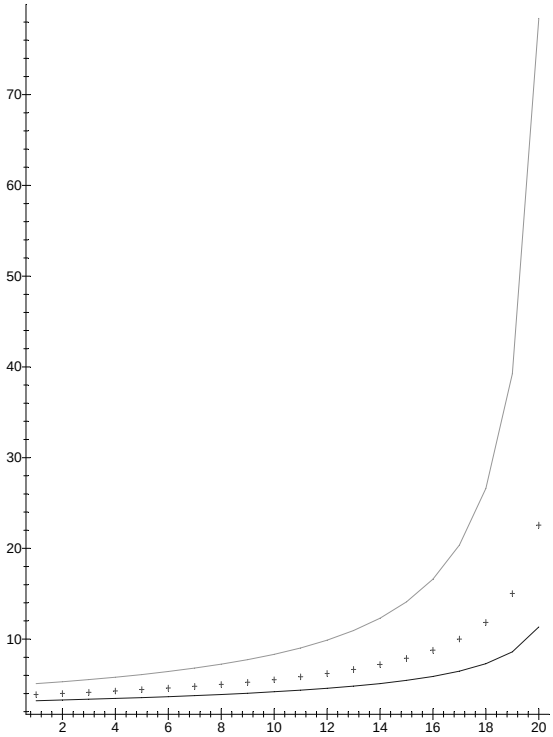
Estymator Hilla



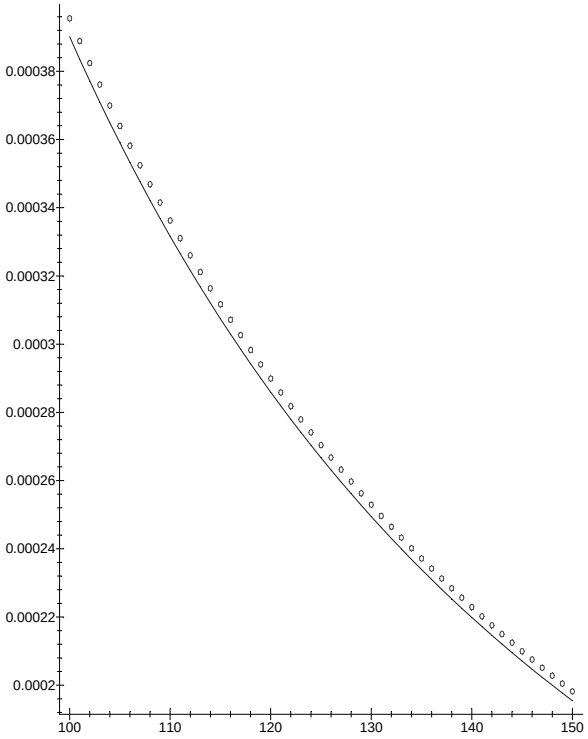
MSE dla indeksu Hilla

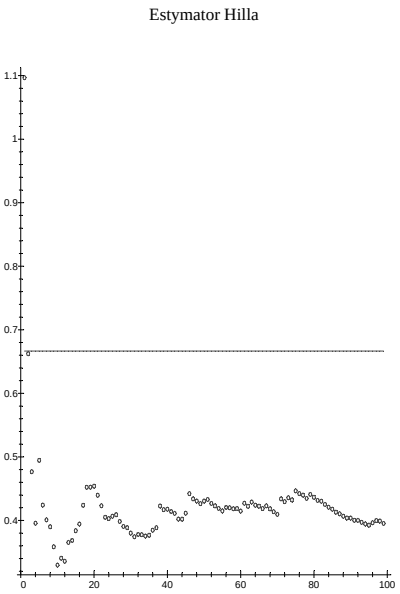


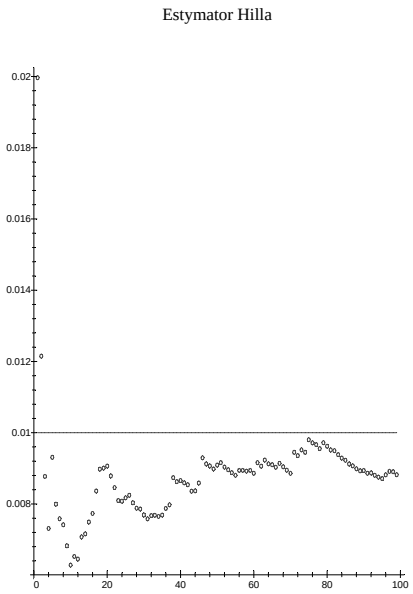
Przedzial ufnosci dla kwantyli



Dopasowanie ogona







Pojawiają się dwa naturalne pytania:

1. Czy warto stosować skomplikowaną procedurę otrzymywania estymatora Hilla zamiast np. używać metodę momentów?

2. Jeżeli pozytywnie odpowiemy na poprzednie pytanie, to po co estymować wartości $P(X > u)$ za pomocą wzoru (6.2.11) zamiast od razu z $P(X > x) = x^{-\alpha_{opt}}$?

W przypadku pierwszego pytania odpowiedź otrzymamy w rozdziale poświęconym ciężkim ogonom. Estymatory MM są obarczone dużym błędem. Po drugie estymatory MM stosują się w przypadku rozkładu Pareto $P(X > x) = x^{-\alpha}$, a nie rozkładu typu Pareto $P(X > x) = x^{-\alpha}L(x)$ z nieznaną funkcją L. Powyższa uwaga daje też odpowiedź na drugie pytanie. Nie wiemy, czy nasze dane pochodzą z rozkładu Pareto czy z rozkładu typu Pareto i przyjęcia założenia, że $L(x) \equiv 1$ może prowadzić do znacznych błędów.

Przykład 6.2.3 [5]Wygenerowano 100 liczb z rozkładu o dystrybuancie $1-x^{-1.5}/\ln(\exp(1)*x)$, $x \geqslant 1$ oraz z o dystrybuancie $1 - x^{-100}/\ln(\exp(1) * x)$, $x \geqslant 1$. W każdym przypadku policzyliśmy estymator Hilla, odpowiednio lewy i rawy rysunek.

[5]ciagle-5.mws

Estymator zachowuje się źle dla $\alpha = 1.5$. Wytłumaczenie może być następujące. Funkcja $\ln(\exp(1) * x)$ powoduje szybsze malenie ogona. W efekcie estymator ma tendencję do pokazywania większej wartości niż prawdziwe α, czyli mniejsze wartości niż prawdziwe γ. W przypadku $\alpha = 100$ ogon maleje na tyle szybko, że dodatkowy efekt jest niezauważalny.

□

6.3 Rozkłady z ciężkimi ogonami

W praktyce stosuje sie często metodę momentów, tzn. estymuje się parametry modelu na bazie momentów próbkowych, np.

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Rozważmy próbę $X_1, \ldots, X_n$ z rozkładu $Par(\alpha, 1)$ i załóżmy na moment, że wiemy iż $X_n^* = \max\{X_1, \ldots, X_n\} \leqslant d$. Obliczamy

$$\begin{aligned}
\mathrm{E}\left[\widehat{\mu}|X_n^* \leqslant d\right] &= \mathrm{E}\left[\widehat{\mu}|X_1 \leqslant d, \ldots, X_n \leqslant d\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[X_i|X_1 \leqslant d, \ldots, X_n \leqslant d\right] \\
&= \mathrm{E}\left[X_1|X_1 \leqslant d, \ldots, X_n \leqslant d\right] \\
&= \mathrm{E}\left[X_1|X_1 \leqslant d\right] \\
&= \frac{\alpha}{\alpha-1}\frac{1-(1+d)^{-(\alpha-1)}}{1-(1+d)^{-\alpha}} - 1.
\end{aligned}$$

Ponieważ prawdziwa średnia wynosi $\frac{1}{\alpha-1}$, więc błąd względny jaki popełniamy estymując średnią za pomocą $\widehat{\mu}$ wynosi:

$$b := \frac{\mathrm{E}\left[\widehat{\mu}|X_n^* \leqslant d\right] - \frac{1}{\alpha-1}}{\frac{1}{\alpha-1}} = \frac{-\alpha(1+d)^{-(\alpha-1)} - (1+d)^{-\alpha}}{1-(1+d)^{-\alpha}}.$$

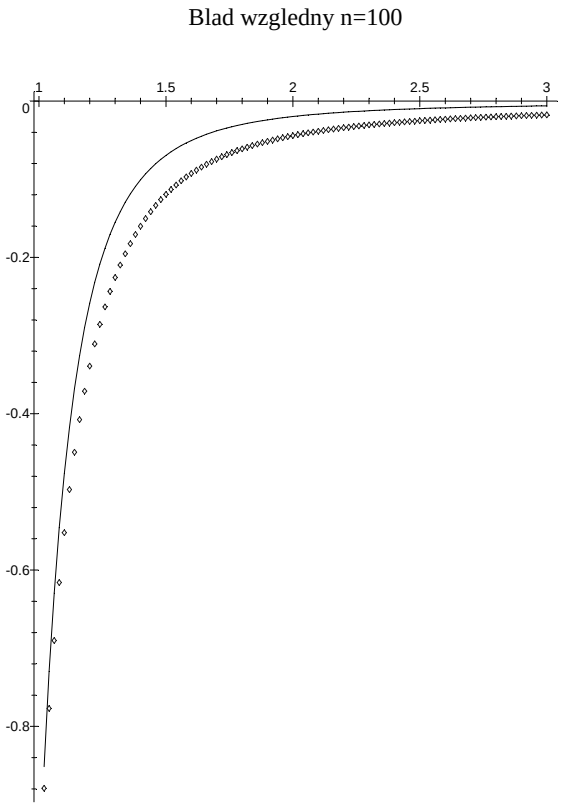
Niech teraz $p \in (0,1)$ i wybierzmy d takie, że $P(X_n^* \leqslant d) = p$, tzn. chcemy aby z (dużym) prawdopodobieństwem p nie popełniać błędu zakładając, że $X_n^* \leqslant d$. Wtedy

$$P(X_n^* \leqslant d) = (1-(1+d)^{-\alpha})^n = p,$$

a stąd

$$b := -\alpha\frac{(1-p)^{(\alpha-1)/\alpha} - (1-p)^{1/n}}{p^{1/n}}.$$

Biorąc standardowe wartości $p = 0.99$ i $p = 0.95$ oraz próbę o liczebności $n = 100$ dostajemy dla różnych parametrów α:



Rysunek 6.3.1: Błąd względny dla $n = 100$ oraz $p = 0.95$ (linia kropkowana) i $p = 0.99$ (linia ciągła)

W szczególności, jeżeli $\alpha = 1.5$ i $p = 0.95$ to błąd wynosi -12%. Jeśli więc prawdziwy parametr wynosi 1.5, to nasza estymacja dla średniej jest aż o 12% za niska 'w 95 na 100 przypadkach'.

Powyższe rozważania pokazują, że używanie średniej próbkowej jako estymatora średniej jest bardzo niebezpieczne w przypadku rozkładu Pareto i ogólniej, w przypadku **rozkładów ciężkoogonowych**.

Dla nieujemnych zmiennych losowych i dla rozkładów skoncentrowanych na półprostej dodatniej istnieje wiele sposobów wyrażenia intuicji: zmienna X przyjmuje "duże" wartości z "istotnym" prawdopodobieństwem.

Definicja 6.3.1 *Niech $F(0) = 0$. Zmienna losowa $X \sim F$ ma ciężki ogon, jeśli jej funkcja tworząca momenty*

$$M(s) = \mathrm{E}\left[e^{sX}\right] = \int_{-\infty}^{\infty} e^{sx} dF(x) = \infty$$

dla wszystkich $s > 0$.

Uwaga 6.3.2 Dla rozkładu F z ciężkim ogonem $\lim_{x\to\infty} e^{sx}\overline{F}(x) = \infty$ dla wszystkich $s > 0$.

6.3.1 Klasy podwykładnicze

Rozkłady podwykładnicze są specjalną klasą rozkładów o ciężkich ogonach. Tego typu rozkłady często stosuje się w modelach ubezpieczeniowych do modelowania wielkości szkód po pożarach, huraganach lub powodziach.

Definicja 6.3.3 *Rozkład F na $(0, \infty)$ nazywamy podwykładniczym jeśli*

$$\lim_{x\to\infty} \frac{1 - F^{*2}(x)}{1 - F(x)} = 2.$$

Klasę podwykładniczych rozkładów oznaczamy przez $\mathcal{S}$.

Niech $(X_i)_{i\geqslant 1}$ będą niezależnymi dodatnimi zmiennymi losowymi z tym samym rozkładem F takim, że $F(x) < 1$ dla wszystkich $x > 0$. Oznaczmy przez

$$\overline{F}^{n*}(x) = 1 - F^{n*}(x) = P(X_1 + \ldots + X_n > x)$$

ogon n-tego splotu F.

Twierdzenie 6.3.4 *Warunek $F \in \mathcal{S}$ jest równoważny każdemu z następujących warunków:*

a) $\lim_{x\to\infty} \frac{\overline{F}^{n*}(x)}{\overline{F}(x)} = n$ *dla pewnego* $n \geqslant 2$,

b) $\lim_{x\to\infty} \frac{P(X_1+\ldots+X_n>x)}{P(\max(X_1,\ldots,X_n)>x)} = 1$ *dla pewnego* $n \geqslant 2$.

Jeżeli F jest rozkładem podwykładniczym, to funkcja tworząca momenty nie istnieje.

Definicja 6.3.5 *Niech F będzie rozkładem na $(0,\infty)$ takim, że $F(x) < 1$ dla wszystkich $x > 0$. Mówimy, że $F \in S^*$, jeśli F ma skończoną średnią $\mu > 0$ i*

$$\lim_{x\to\infty} \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)} \overline{F}(y)dy = 2\mu.$$

Twierdzenie 6.3.6 *Jeśli $F \in S^*$, wtedy $F \in S$ i $\widetilde{F} \in S$, gdzie*

$$\tilde{F}(x) = \frac{1}{\mu} \int_0^x \bar{F}(u)du$$

jest **rozkładem resztowym**.

Powyższe twierdzenia ma ogromne znaczenia dla wyliczania prawdopodobieństwa ruiny dla rozkładów ciężkoogonowych (Rozdział 6.3).

Przykład 6.3.7 Typowe przykłady rozkładów z klasy $\mathcal{S}^*$ to $Wei(r,c)$, $0 < r < 1$, $c > 0$; $Par(\alpha,c)$, $\alpha > 1$; $LN(\mu,\sigma^2)$.

□

Rozdział 7

Modele bayesowskie

7.1 Model portfela niejednorodnego.

Jednym z ważniejszych pytań, na które trzeba odpowiedzieć kalkulując składkę, jest pytanie o **jednorodność** portfela. Aby zbudować dobry model opisujący napływ roszczeń należy dopasować w modelu rozkłady ryzyk indywidualnych tak, aby łączny efekt statystycznie dobrze pasował do zebranych danych. Strukturalnie prosty model powstaje przy założeniu, że ryzyka indywidualne w portfelu mają ten sam parametryczny rozkład, a różnią się jedynie wartością pewnego parametru θ. To znaczy zakładamy, że możemy rozkład szkody indywidualnej indeksować rodziną rozkładów F^θ, $\theta \in \mathbb{R}$.

Przyjmijmy, że nasze dane dotyczą k grup jednorodnych kontraktów ubezpieczeniowych, każda z grup składa się z n_i ryzyk, $i = 1, ..., k$. Załóżmy, że i-ta grupa pochodzi z rozkładu F^{θ_i} i zaobserwowaliśmy następujące szkody:

kontrakt	rozkład	szkody
1	F^{θ_1}	S_{11},...,S_{1n_1}
$\vdots$		
k	F^{θ_k}	S_{k1},...,S_{kn_k}

Możliwe są dwa skrajne podejścia: albo zakładamy, że $\theta_1 = \cdots = \theta_k$ i estymujemy θ na podstawie połączonej próby $S_{11}, ..., S_{1n_1}, ..., S_{k1}$,...,$S_{kn_k}$ albo przyjmujemy, że każda z grup jest inna. Wtedy estymujemy θ_i na podstawie i-tej próby. Oba te podejścia mają swoje wady: obciążanie wszystkich klientów jednakową składką może prowadzić do "negatywnej selekcji ryzyk", tzn. klienci spodziewający się niższych strat będą wybierali inne oferty, z drugiej strony dysponujemy zazwyczaj szczupłą ilością danych dotyczących i-tego typu. Rozsądne jest więc skorzystanie z aparatu statystycznego do testowania hipotezy:

$$H_0 : \theta_1 = \cdots = \theta_k. \tag{7.1.1}$$

Przykład 7.1.1 Przypuśćmy, że S_{ij}, $j = 1, ..., n_i$ są niezależne o wartościach $0, 1$ z prawdopodobieństwem θ_i uzyskania 1, $i = 1, ..., k$. Zmienna S_{ij} jest indykatorem zdarzenia, że

i-ty kierowca miał wypadek w j-tym roku. Rozpatrujemy więc ubezpieczenie dla k kierowców. Wtedy $S_i \equiv \sum_{j=1}^{n_i} S_{ij}$ oznacza liczbę szkód i-tego kierowcy w ciągu n_i lat i ma rozkład Bernoulliego $B(n_i, \theta_i)$. Do testowania hipotezy (7.1.1) możemy użyć testu chi-kwadrat. Rozpatrzmy konkretne dane liczbowe: Mamy 20 ($k = 20$) kierowców, którzy są ubezpieczani przez 10 lat ($n_i = n = 10$). W tabeli wartość 1 oznacza, że kierowca miał w danym roku wypadek. Ostatni wiersz daje ilość lat, w których kierowca miał wypadek.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	0
2	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0
3	0	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
5	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0
8	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1	0
$\sum$	0	0	2	0	0	2	2	0	6	4	3	1	1	1	0	0	5	1	1	0

Kilku kierowców nie miało żadnego wypadku, dlatego nie będą chcieli składki obliczonej na podstawie wartości średniej dla całego portfela. Sprawdzimy więc teraz, czy nasze dane wynikają z niejednorodności populacji kierowców, czy może niektórzy z nich mieli trochę mniej szczęścia. Przetestujemy hipotezę, że średnia ilość lat wypadkowych $\alpha_i = n\theta_i$ dla każdego kierowcy jest taka sama, tzn.

$$H_0 : \alpha_0 = \cdots = \alpha_{20}.$$

Niech $\widehat{\alpha}_i$ będzie ilością lat wypadkowych dla i-tego kierowcy oraz

$$\alpha = \frac{1}{20} \sum_{i=1}^{20} \widehat{\alpha}_i$$

(Jest to estymator średniej ilości lat wypadkowych w populacji). Statystyka do testowania hipotezy H_0 ma postać

$$\frac{n \sum_{i=1}^{20} (\widehat{\alpha}_i - \widehat{\alpha})}{\widehat{\alpha}(1 - \widehat{\alpha})} = 49.1631.$$

Przy założeniu H_0, ma ona rozkład χ^2_{19}. Ponieważ $P(\chi^2_{19} > 49.1631) \approx 0.0002$, więc należy odrzucić hipotezę H_0. Populacja kierowców jest więc niejednorodna i nie powinni oni płacić tej samej składki.

□

Jeżeli przyjęliśmy hipotezę H_0 (czyli przyjęlismy jednorodność portfela) to składkę możemy wyliczyć poznanymi już wcześniej metodami, np. metodą wartości średniej, czy wariancji. W przypadku odrzucenia H_0 musimy postąpić trochę inaczej. Oznaczając wartość oczekiwaną i wariancję dla ryzyka indywidualnego odpowiednio przez

$$\mu(\theta) = \int x dF^{\theta}(x) = \int x f_{\theta}(x) dx, \tag{7.1.2}$$

$$\sigma^2(\theta) = \int (x - \mu(\theta))^2 dF^{\theta}(x), \tag{7.1.3}$$

wyliczamy **składki ryzyka indywidualnego** ze wzorów:

1. dla składki wartości oczekiwanej

$$H(\theta) = (1 + \rho)\mu(\theta), \qquad \rho > 0,$$

2. dla składki odchylenia standardowego

$$H(\theta) = \mu(\theta) + \delta\sigma(\theta), \qquad \delta > 0,$$

Z rodziną parametryczną F^{θ} związujemy rozkład o dystrybuancie F_{Θ}, który opisuje losowy udział poszczególnych parametrów w całym zbiorze wartości parametru θ. Inaczej mówiąc, zakładamy, że parametr (struktury) θ jest losowy i opisujemy jego wartość przez zmienną losową Θ o dystrybuancie F_{Θ} Wtedy *typowa* szkoda S w danym portfelu będzie miała mieszany rozkład

$$F_S(x) = \int F^{\theta}(x) dF_{\Theta}(\theta). \tag{7.1.4}$$

Dokładniej mówiąc, przyjmujemy, że na pewnej przestrzeni probabilistycznej istnieją zmienne losowe S, Θ, takie, że rozkład warunkowy S przy warunku $\Theta = \theta$ jest równy F^{θ} oraz Θ ma rozkład (tzw. a priori) o dystrybuancie F_{Θ}. Zmienne S oraz Θ są zależne. Znajomość wartości $\Theta = \theta$ wyznacza rozkład indywidualnych szkód o parametrze struktury θ.

Znajdziemy teraz, korzystając z równań (7.1.2)-(7.1.4), relacje między wartością średnią μ i wariancją σ^2 powyższego (mieszanego) rozkładu F_S, a odpowiednimi parametrami ryzyk indywidualnych. Bezpośrednio z określenia S i Θ widzimy, że $\mu(\theta) = \mathrm{E}[S|\Theta = \theta]$, $\sigma^2(\theta) = \mathrm{Var}[S|\Theta = \theta]$.

Lemat 7.1.2 *W powyższym modelu*

$$\begin{aligned} \mathrm{E}[S] &:= \mu = \mathrm{E}[\mu(\Theta)] \\ \mathrm{Var}[S] &:= \sigma^2 = \mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{Var}[\mu(\Theta)]. \end{aligned}$$

Dowód: Dla pierwszego wzoru mamy

$$\mu \stackrel{\text{def}}{=} \int x dF(x) \stackrel{(7.1.4)}{=} \int \left(\int x dF^{\theta}(x) \right) dU(\theta) \stackrel{(7.1.2)}{=} \int \mu(\theta) dU(\theta) = \mathrm{E}\left[\mu(\Theta)\right],$$

natomiast dla drugiego mamy z definicji

$$\sigma^2 = \int (x-\mu)^2 dF(x) = \int x^2 dF(x) - \left(\int x dF(x) \right)^2.$$

Korzystając teraz ze wzoru (7.1.4) otrzymujemy

$$\begin{aligned}
\sigma^2 &= \int \left(\int x^2 dF^{\theta}(x) \right) dU(\theta) - \left(\int \left(\int x dF^{\theta}(x) \right) dU(\theta) \right)^2 \\
&= \int \left(\int (x - \mu(\theta) + \mu(\theta))^2 dF^{\theta}(x) \right) dU(\theta) - \left(\int \left(\int x dF^{\theta}(x) \right) dU(\theta) \right)^2 \\
&= \int \left(\int (x-\mu(\theta))^2 dF^{\theta}(x) \right) dU(\theta) + \int \left(\int (\mu(\theta))^2 dF^{\theta}(x) \right) dU(\theta) + \\
&+ 2 \int \left(\int (x-\mu(\theta))\mu(\theta) dF^{\theta}(x) \right) dU(\theta) - \left(\int \left(\int x dF^{\theta}(x) \right) dU(\theta) \right)^2.
\end{aligned}$$

Pierwszy i czwarty składnik są z definicji (7.1.2) i (7.1.3) równe $\int \sigma^2(\theta) dU(\theta)$ i $\left(\int \mu(\theta) dU(\theta)\right)^2$, drugi składnik redukuje się do $\int (\mu(\theta))^2 dU(\theta)$, gdyż $\int dF^{\theta}(x) = 1$. Rozbijając trzeci składnik na dwie części i zamieniając kolejność całkowania dostajemy

$$\begin{aligned}
\sigma^2 &= \int \sigma^2(\theta) dU(\theta) + \int (\mu(\theta))^2 dU(\theta) + 2 \int \mu(\theta) \left(\int (x dF^{\theta}(x) \right) dU(\theta) \\
&- 2 \int (\mu(\theta))^2 \left(\int dF^{\theta}(x) \right) dU(\theta) - \left(\int \mu(\theta) dU(\theta) \right)^2.
\end{aligned}$$

Podobnie jak wyżej, czwarty składnik redukuje się do $2\int (\mu(\theta))^2 dU(\theta)$. Z definicji (7.1.2) trzeci składnik redukuje się do $2\int (\mu(\theta))^2 dU(\theta)$, a stąd

$$\sigma^2 = \int \sigma^2(\theta) dU(\theta) + \int (\mu(\theta))^2 dU(\theta) - \left(\int \mu(\theta) dU(\theta) \right)^2.$$

Pamietając teraz, że U jest dystrybuantą zmiennej losowej Θ otrzymujemy

$$\begin{aligned}
\sigma^2 &= \mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{E}\left[\mu(\Theta)^2\right] - \left(\mathrm{E}\left[\mu(\Theta)\right]\right)^2 \\
&= \mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{Var}\left[\mu(\Theta)\right]
\end{aligned}$$

co kończy dowód dla wariancji. □

Bardziej ogólnym związkiem, określonym dla dowolnych ryzyk X, Y zdefiniowanych na tej samej przestrzeni probabilistycznej, jest następujący wzór na kowariancję

$$\begin{aligned}
\mathrm{Cov}\left[X,Y\right] &\equiv \mathrm{E}\left[XY\right] - \mathrm{E}\left[X\right]\mathrm{E}\left[Y\right] = \mathrm{E}\left[\mathrm{E}\left[XY|\Theta\right]\right] - \mathrm{E}\left[\mathrm{E}\left[X|\Theta\right]\right]\mathrm{E}\left[\mathrm{E}\left[Y|\Theta\right]\right] \\
&= \mathrm{E}\left[\mathrm{Cov}\left[X,Y|\Theta\right]\right] + \mathrm{E}\left[X|\Theta\right]\mathrm{E}\left[Y|\Theta\right] - \mathrm{E}\left[\mathrm{E}\left[X|\Theta\right]\right]\mathrm{E}\left[\mathrm{E}\left[Y|\Theta\right]\right] \\
&= \mathrm{E}\left[\mathrm{Cov}\left[X,Y|\Theta\right]\right] + \mathrm{E}\left[\mathrm{E}\left[X|\Theta\right]\mathrm{E}\left[Y|\Theta\right]\right] - \mathrm{E}\left[\mathrm{E}\left[X|\Theta\right]\right]\mathrm{E}\left[\mathrm{E}\left[Y|\Theta\right]\right] \\
&= \mathrm{E}\left[\mathrm{Cov}\left[X,Y|\Theta\right]\right] + \mathrm{Cov}\left[\mathrm{E}\left[X|\Theta\right], \mathrm{E}\left[Y|\Theta\right]\right].
\end{aligned}$$

Wzór ten zapisaliśmy używając (dla dowolnej zmiennej losowej Z) notacji $\mathrm{E}[Z|\Theta]$ dla określenia zmiennej losowej, którą otrzymujemy z funkcji $\psi(\theta) = \mathrm{E}[Z|\Theta = \theta]$ nakładając ja na zmienną Θ, tzn. $\mathrm{E}[Z|\Theta] = \psi(\Theta)$.

Powyższe wzory można zapisać w sposób następujący:

$$\mu = \mathrm{E}[S] = \mathrm{E}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\mu(\Theta)], \tag{7.1.5}$$

$$\sigma^2 = \mathrm{Var}[S] = \mathrm{E}[\mathrm{Var}[S|\Theta]] + \mathrm{Var}[\mathrm{E}[S|\Theta]]. \tag{7.1.6}$$

Metoda iteracyjna

Do obliczania składek w portfelu ryzyk z parametrem struktury θ służy również inna metoda polegająca na iteracyjnym zastosowaniu metody liczenia składki: raz do policzenia $H(\theta)$, następnie traktując $H(\Theta)$ jako zmienną losową do policzenia $H^* = H(H(\Theta))$, przy tej samej metodzie H liczenia składki. Na przykład przy metodzie wartości oczekiwanej mamy

$$H(\theta) = (1+\rho)\mu(\theta), \quad \rho > 0,$$
$$H^* = (1+\rho)\mathrm{E}[H(\Theta)] = (1+\rho)^2\mathrm{E}[\mu(\Theta)] = (1+\rho)^2\mu.$$

Przy składce odchylenia standardowego natomiast

$$H(\theta) = \mu(\theta) + \delta\sigma(\theta),$$

$$H^* = \mathrm{E}[H(\Theta)] + \delta\sqrt{\mathrm{Var}[H(\Theta)]} = \mu + \delta\mathrm{E}[\sigma(\Theta)] + \delta\sqrt{\mathrm{Var}[\mu(\Theta)] + \delta\sigma(\Theta)}.$$

Zwykle składki H liczone z parametrów mieszanego rozkładu F nie są równe składkom H^* liczonym metodą iteracyjną, na przykład dla składki wartości oczekiwanej mamy $H^* = (1+\rho)H$.

Przykład 7.1.3 Zmienna losowa S ma, pod warunkiem $\Theta = \theta$, rozkład $Poi(\theta)$. Zmienna losowa Θ ma rozkład jednostajny na odcinku $[a, b]$. Znajdziemy średnią i wariancję S.

Przypomnijmy, że dla zmiennej losowej $Z \sim Poi(\theta)$ mamy: $\mathrm{E}[Z] = \mathrm{Var}[Z] = \theta$, czyli $\mu(\theta) = \mathrm{E}[S|\Theta = \theta] = \sigma^2(\theta) = \mathrm{Var}[S|\Theta = \theta] = \theta$. Ze wzorów (7.1.5)-(7.1.6) mamy: $\mathrm{E}[S] = \mathrm{E}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\Theta]$. Analogicznie $\mathrm{Var}[S] = \mathrm{E}[\mathrm{Var}[S|\Theta]] + \mathrm{Var}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\Theta] + \mathrm{Var}[\Theta]$. Stąd $\mathrm{E}[S] = \frac{(a+b)}{2}$ oraz $\mathrm{Var}[S] = \frac{(a+b)}{2} + \frac{(b-a)^2}{12}$.

□

Przykład 7.1.4 [EA:5.10.1996(5)] Rozkład warunkowy dwóch ryzyk X i Y przy danej wartości parametru $\Theta = \theta$ ma następujące charakterystyki:

$$\mathrm{Cov}[X, Y|\Theta = \theta] = \frac{1}{2}\theta$$

$$\mathrm{E}\left[X|\Theta=\theta\right]=\theta$$
$$\mathrm{E}\left[Y|\Theta=\theta\right]=\theta$$

podczas, gdy rozkładem parametru Θ w populacji ryzyk jest rozkład $Gamma(3,6)$. Obliczymy $\mathrm{Cov}\left[X,Y\right]$.

Ze wzoru na kowariancję dostajemy :

$$\begin{aligned}\mathrm{Cov}\left[X,Y\right] &\equiv \mathrm{E}\left[\mathrm{Cov}\left[X,Y|\Theta\right]\right]+\mathrm{E}\left[\mathrm{E}\left[X|\Theta\right]\mathrm{E}\left[Y|\Theta\right]\right]-\mathrm{E}\left[\mathrm{E}\left[X|\Theta\right]\right]\mathrm{E}\left[\mathrm{E}\left[Y|\Theta\right]\right]\\ &=\frac{1}{2}\mathrm{E}\left[\Theta\right]+\mathrm{E}\left[\Theta^2\right]-\left(\mathrm{E}\left[\Theta\right]\right)^2=\frac{1}{2}\mathrm{E}\left[\Theta\right]+\mathrm{Var}\left[\Theta\right].\end{aligned}$$

Dla zmiennej losowej Θ o rozkładzie $Gamma(\alpha,\beta)$ mamy: $\mathrm{E}\left[\Theta\right]=\frac{\alpha}{\beta}$ oraz $\mathrm{Var}\left[\Theta\right]=\frac{\alpha}{\beta^2}$. Stąd $\mathrm{Cov}\left[X,Y\right]=\frac{1}{3}$.

□

W przykładzie 7.1.1 pokazaliśmy niejednorodny portfel ryzyk, w szczególności θ_i, prawdopodobieństwo szkody w danym roku, było tam inne dla różnych kierowców. Typowym zadaniem w takiej sytuacji jest wyestymowanie parametru θ_i dla każdego z kierowców z osobna. Procedura ta nazywana jest metodą **wiarogodności** (*credibility, experience rating*). Podstawowymi założeniami metody wiarogodności są:

- zależność indywidualnych ryzyk $S_{i1},\ldots,S_{in_i},S_i$ od parametru struktury $\Theta=\theta_i$;
- $S_{i1},...,S_{in_i},S_i$ tworzą ciąg niezależnych zmiennych losowych o dystrybuancie F^{θ_i} z gęstością f_{θ_i};
- parametr struktury Θ jest zmienną losową o dystrybuancie F_Θ i gęstości $f_\Theta(\theta)=\pi(\theta)$ (gęstość **a priori**).

Niech $\overline{S}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}S_{ij}$ będzie średnią wielkością szkody dla i-tego typu ryzyka, $\mu(\theta_i)=\mathrm{E}\left[S_{ij}|\Theta=\theta_i\right]$, $\mu=\mathrm{E}\left[\mu(\Theta)\right]$ oraz $\sigma^2(\theta_i)=\mathrm{Var}\left[S_{ij}|\Theta=\theta_i\right]$. Jeżeli rozważamy składkę netto, to w przypadku przyjęcia hipotezy o jednorodności składka dla każdego klienta bedzie wynosiła μ. W przypadku odrzucenia składka netto dla i-tego klienta wynosi $\mu(\theta_i)$. Zauważmy, że wartości Θ nie są obserwowane. Estymatorem $\mu(\theta_i)$ jest $\overline{S}_i$.

Podamy najpierw kryterium, kiedy możemy stosować $\overline{S}_i$ jako dobrego estymatora dla składki indywidualnej. Rozważmy w tym celu *małe* $a\in(0,1)$ i *duże* $p\in(0,1)$. Będziemy używali $\overline{S}_i$ jako estymatora dla $\mu(\theta_i)$, gdy

$$P(|\overline{S}_i-\mu(\Theta)|<a\mu(\Theta)|\Theta=\theta_i)\geqslant p. \tag{7.1.7}$$

Powyższy warunek nazywany jest warunkiem **całkowitej wiarogodności**. Jest on równoważny warunkowi minimalnej wielkości próby dla dużych prób, gdzie można zastosować aproksymację normalną do $\overline{S}_i$. Mamy wtedy

$$P\left(\frac{|\overline{S}_i-\mu(\Theta_i)|}{\sqrt{\sigma^2(\Theta)/n_i}}<\frac{a\mu(\Theta)}{\sqrt{\sigma^2(\Theta)/n_i}}|\Theta=\theta_i\right)\approx 2\Phi\left(\frac{a\sqrt{n_i}\mu(\theta_i)}{\sqrt{\sigma^2(\theta_i)}}\right)-1\geqslant p,$$

a stąd

$$n_i \geqslant \frac{z^2_{(1+p)/2}\sigma^2(\theta_i)}{a^2(\mu(\theta_i))^2}. \tag{7.1.8}$$

Przykład 7.1.5 Przypuśćmy, że $(S_{ij})_{j\geqslant 1}$ mają rozkład Poissona z nieznaną średnią $\Theta_i = \lambda_i$. Wtedy $\mu(\lambda_i) = \lambda_i = \sigma^2(\lambda_i)$ (porównaj Przykład 7.1.3). Wtedy (7.1.8) redukuje się do

$$\lambda_i n_i \geqslant \frac{z^2_{(1+p)/2}}{a^2}$$

i w szczególności, gdy $a = 0.05$, $p = 0.9$ dostajemy

$$\lambda_i n_i \geqslant 1082.28.$$

Ponieważ lewa strona powyższej nierówności może byc interpretowana jako oczekiwana ilość szkód w ciągu n_i lat, dostajemy, że całkowita wiarogodność może być zastosowana, gdy zostanie zgłoszonych 1083 szkód z i-tego ryzyka.

□

Z powyższego przykładu widzimy, że metoda całkowitej wiarogodności może być stosowana w wyjątkowych przypadkach. Alternatywnym podejściem jest tzw. **wiarogodność częściowa**. Polega ona na estymacji parametru $\mu(\theta_i)$ za pomocą

$$b\overline{S}_i + (1-b)\mu. \tag{7.1.9}$$

Wielkość $b \in [0,1]$ powinna być tym większa im więcej danych mamy do dyspozycji. Warunek wiarogodności częściowej ma następującą postać:

$$P(b|\overline{S}_i - \mu(\Theta)| < a\mu(\Theta)|\Theta = \theta_i) \geqslant p.$$

Przykład 7.1.6 (cd. Przykładu 7.1.5). Zastosowanie aproksymacji normalnej daje

$$b = \min\left\{\frac{a\sqrt{n_i}(\mu(\theta_i))}{z_{(1+p)/2}\sqrt{\sigma^2(\theta_i)}}, 1\right\}.$$

Dla S_{ij} o rozkładzie Poissona z nieznanym parametrem $\Theta = \lambda_i$

$$b = \min\left\{\frac{a\sqrt{\lambda_i n_i}}{z_{(1+p)/2}}, 1\right\}.$$

Ponieważ wielkość λ_i jest nieznana, więc zastępujemy $\lambda_i n_i$ odpowiednim estymatorem, czyli ilością zaobserwowanych szkód w ciągu n_i lat.

□

Przedstawione wyżej metody całkowitej i częściowej wiarogodności wymagają dużej wielkości obserwacji. Przedstawimy teraz pewien ogólny sposób wyliczenia składki.

Niech $(S_1(\theta), ..., S_n(\theta), S(\theta))$ będzie ciągiem niezależnych zmiennych losowych o rozkładzie F^θ.

Łączna gęstość wektora $(S_1(\theta), \ldots, S_n(\theta), S(\theta))$ dana jest wzorem

$$\underline{f}_\theta(x_1, \ldots, x_n, x) = f_\theta(x_1) \cdots f_\theta(x_n) f_\theta(x).$$

Jest to warunkowa gęstość wielkości szkód przy warunku $\Theta = \theta$. Bezwarunkowy rozkład wielkości szkód w portfelu $(S_1, \ldots, S_n, S)$ ma więc gęstość

$$\underline{f}(x_1, \ldots, x_n, x) = \int \underline{f}_\theta(x_1, \ldots, x_n, x)\pi(\theta)d\theta \tag{7.1.10}$$

Oczywiście zmienne $(S_1, \ldots, S_n)$ nie muszą być niezależne (są tylko warunkowo niezależne). Rozkładem **a posteriori** zmiennej Θ przy warunku $S_1 = x_1, \ldots, S_n = x_n$ nazywamy rozkład o gęstości

$$\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta) = \underline{f}_\theta(x_1, \ldots, x_n)\pi(\theta)/\underline{f}(x_1, \ldots, x_n). \tag{7.1.11}$$

Rozkład szkody S pod warunkiem $S_1 = x_1, \ldots, S_n = x_n$ dany jest przez:

$$f_S(x|x_1, \ldots, x_n) = \int f_\theta(x)\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta)d\theta. \tag{7.1.12}$$

Wyliczymy składkę netto kompensującą szkodę S (w danym okresie rozliczeniowym) przy użyciu informacji o poprzednich szkodach $S_1 = x_1, \ldots, S_n = x_n$ (w poprzedzających okresach rozliczeniowych). Ponieważ składkę możemy oprzeć jedynie na obserwowanych szkodach (nie znając Θ), szukamy więc "najlepszego" estymatora $\widehat{\mu}(\Theta)$ zmiennej $\mu(\Theta) = \mathrm{E}[S|\Theta]$ opartego na $S_1, \ldots, S_n$. Jako kryterium "dobroci" weźmiemy kwadratową funkcję straty, tzn. w klasie estymatorów $\widehat{\mu}(\Theta)$ zbudowanych na próbie $S_1, \ldots, S_n$ szukamy takiego, który minimalizuje $\mathrm{E}[\mu(\Theta) - \widehat{\mu}(\Theta)]^2$. Niech $H_{n+1} = \mathrm{E}[\mu(\Theta)|S_1, \ldots, S_n]$. Mamy wtedy

$$\begin{aligned}
\mathrm{E}\left[(\mu(\Theta) - \widehat{\mu}(\Theta))^2\right] &= \mathrm{E}\left[(\mu(\Theta) - H_{n+1} + H_{n+1} - \widehat{\mu}(\Theta))^2\right] \\
&= \mathrm{E}\left[(\mu(\Theta) - H_{n+1})^2 + (H_{n+1} - \widehat{\mu}(\Theta))^2\right] + \\
&+ 2\mathrm{E}\left[(\mu(\Theta) - H_{n+1})(H_{n+1} - \widehat{\mu}(\Theta))\right].
\end{aligned}$$

Zauważmy, że

$$\begin{aligned}
\mathrm{E}\left[(\mu(\Theta) - H_{n+1})(H_{n+1} - \widehat{\mu}(\Theta))\right] &= \mathrm{E}\left[\mathrm{E}\left[(\mu(\Theta) - H_{n+1})(H_{n+1} - \widehat{\mu}(\Theta))|S_1, \ldots, S_n\right]\right] \\
&= \mathrm{E}\left[(H_{n+1} - \widehat{\mu}(\Theta))E(\mu(\Theta) - H_{n+1})|S_1, \ldots, S_n\right],
\end{aligned}$$

gdyż $H_{n+1} - \widehat{\mu}(\Theta)$ jest (mierzalną) funkcją obserwacji $S_1, \ldots, S_n$. Ale z definicji H_{n+1}

$$\begin{aligned} \mathrm{E}\left[(\mu(\Theta) - H_{n+1})|S_1, \ldots, S_n\right] &= \mathrm{E}\left[\mu(\Theta)|S_1, \ldots, S_n\right] - \mathrm{E}\left[H_{n+1}|S_1, \ldots, S_n\right] \\ &= H_{n+1} - H_{n+1} = 0. \end{aligned}$$

Mamy więc $\mathrm{E}\left[(\mu(\Theta) - \widehat{\mu}(\Theta))^2\right] = \mathrm{E}\left[(\mu(\Theta) - H_{n+1})^2\right] + \mathrm{E}\left[(H_{n+1} - \widehat{\mu}(\Theta))^2\right]$. Zauważmy teraz, że pierwszy składnik nie zależy od wyboru estymatora, natomiast drugi jest nieujemny i zeruje się wtedy i tylko wtedy, gdy $\widehat{\mu}(\Theta) = H_{n+1}$ prawie wszędzie.

Definicja 7.1.7 *Składką bayesowską wyznaczoną na podstawie obserwacji $S_1 = x_1, \ldots, S_n = x_n$ nazywamy*

$$\begin{aligned} H_{n+1}(x_1, \ldots, x_n) &= \mathrm{E}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] \\ &= \int \mu(\theta)\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta)d\theta, \end{aligned} \tag{7.1.13}$$

przy czym $H_{n+1}(S_1, \ldots, S_n)$ jest najlepszym estymatorem $\mu(\Theta)$ w sensie średniokwadratowym.

Zauważmy, że wyliczenie $H_{n+1}(x_1, \ldots, x_n)$ wymaga znajomości rozkładu a'posteriori. Pokażemy jak można pominąć ten wymóg.

Optymalność $H_{n+1}(S_1, \ldots, S_n)$ zachodzi również w trochę innym sensie. Zauważmy, że

$$\begin{aligned} H_{n+1}(S_1, \ldots, S_n) &= \mathrm{E}\left[\mu(\Theta)|S_1, \ldots, S_n\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[S|\Theta\right]|S_1, \ldots, S_n\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[S|\Theta, S_1, \ldots, S_n\right]|S_1, \ldots, S_n\right] \\ &= \mathrm{E}\left[S|S_1, \ldots, S_n\right]. \end{aligned}$$

Otrzymujemy stąd następujący wynik.

Twierdzenie 7.1.8 *Przy założeniach dotyczących struktury modelu mamy*

$$H_{n+1}(S_1, \ldots, S_n) = \mathrm{E}\left[S|S_1, \ldots, S_n\right],$$

a więc $H_{n+1}(S_1, \ldots, S_n)$ jest również najlepszym estymatorem zmiennej losowej S na podstawie obserwacji $S_1, \ldots, S_n$.

7.2 Model liniowy Bühlmanna (Bayesian credibility)

Jak wspominaliśmy powyżej, bayesowskie podejście do wyliczania składki wiarogodności wymaga znajomości rozkładu parametru Θ. Zazwyczaj pozostaje on jednak nieznany,

często nie wiemy nawet jaka jest przestrzeń stanów dla Θ. Spróbujemy ominąć ten problem. Zamiast, jak w poprzednim rozdziale rozpatrywać klasę wszystkich estymatorów dla $\mu(\Theta) = \mathrm{E}\,[S|\Theta]$ opartych na $S_1, \ldots, S_n$ rozpatrzmy estymatory postaci $a_0 + \sum_{j=1}^{n} a_j S_j$. Z uwagi na symetrię (zmienne losowe $S_1, \ldots, S_n$ mają te same rozkłady brzegowe) musimy przyjąć $a_1 = \cdots = a_n$, a więc rozpatrujemy estymatory postaci $\widehat{\mu}_L(\Theta) = a_0 + a_1\overline{S}$, gdzie $\overline{S}$ jest średnią arytmetyczną. Naszym kryterium wyboru będzie ponownie minimalizacja średniego błędu kwadratowego $\mathrm{E}\left[(\mu(\Theta) - a_0 - a_1\overline{S})^2\right]$. Różniczkując względem a_0 i a_1 otrzymujemy

$$\begin{cases} \mathrm{E}\,[(\mu(\Theta)] - a_0 - a_1\mathrm{E}\left[\overline{S}\right] = 0 \\ \mathrm{E}\left[(\mu(\Theta) - a_0 - a_1\overline{S})\overline{S}\right] = 0 \end{cases}. \tag{7.2.1}$$

Mnożąc pierwsze wyrażenie przez $\mathrm{E}\left[\overline{S}\right]$ i odejmując od drugiego otrzymujemy $\mathrm{Cov}\left[\mu(\Theta), \overline{S}\right] - a_1\mathrm{Var}\left[\overline{S}\right] = 0$, a stąd korzystając z (7.1.6)

$$\begin{aligned} a_1 &= \frac{\mathrm{Cov}\left[\mu(\Theta), \overline{S}\right]}{\mathrm{Var}\left[\overline{S}\right]} = \frac{\mathrm{E}\left[\mu(\Theta)\overline{S}\right] - \mathrm{E}\,[\mu(\Theta)]\,\mathrm{E}\left[\overline{S}\right]}{\mathrm{E}\left[\mathrm{Var}\left[\overline{S}|\Theta\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[\overline{S}|\Theta\right]\right]} \\ &= \frac{\mathrm{E}\left[\mathrm{E}\left[\mu(\Theta)\overline{S}|\Theta\right]\right] - \mathrm{E}\,[\mu(\Theta)]\,\mathrm{E}\left[\mathrm{E}\left[\overline{S}|\Theta\right]\right]}{\mathrm{E}\left[\mathrm{Var}\left[\overline{S}|\Theta\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[\overline{S}|\Theta\right]\right]}. \\ &= \frac{\mathrm{E}\left[\mu(\Theta)\mathrm{E}\left[\overline{S}|\Theta\right]\right] - \mathrm{E}\,[\mu(\Theta)]\,\mathrm{E}\left[\mathrm{E}\left[\overline{S}|\Theta\right]\right]}{\mathrm{E}\left[\mathrm{Var}\left[\overline{S}|\Theta\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[\overline{S}|\Theta\right]\right]}. \end{aligned}$$

Zauważmy teraz, że $\mathrm{E}\left[\overline{S}|\Theta\right] = \mathrm{E}\left[\frac{1}{n}(S_1 + \cdots + S_n)|\Theta\right] = \frac{1}{n}n\mathrm{E}\,[S_1|\Theta] = \mathrm{E}\,[S|\Theta] = \mu(\Theta)$, gdyż $S_1, \ldots, S_n$ mają te same rozkłady brzegowe (warunkowe i bezwarunkowe). Analogicznie, $\mathrm{Var}\left[\overline{S}|\Theta\right] = \frac{1}{n}\mathrm{Var}\,[S|\Theta]$. Otrzymujemy więc

$$a_1 = \mathrm{Var}\,[\mu(\Theta)]\,/\left(\frac{1}{n}\mathrm{E}\,[\mathrm{Var}\,[S|\Theta]] + \mathrm{Var}\,[\mu(\Theta)]\right)$$

i ostatecznie dla

$$k = \mathrm{E}\,[\mathrm{Var}\,[S|\Theta]]\,/\mathrm{Var}\,[\mu(\Theta)] = \mathrm{E}\left[\sigma^2(\Theta)\right]/\mathrm{Var}\,[\mu(\Theta)]$$

mamy

$$\widehat{\mu}_L(\Theta) = \frac{n}{n+k}\overline{S} + \frac{k}{n+k}\mathrm{E}\,[\mu(\Theta)]\,, \tag{7.2.2}$$

lub zapisując inaczej

$$\widehat{\mu}_L(\Theta) = b\overline{S} + (1-b)\mathrm{E}\,[\mu(\Theta)] \tag{7.2.3}$$

z

$$b = \frac{1}{1 + \frac{\mathrm{E}[\sigma^2(\Theta)]}{n\mathrm{Var}[\mu(\Theta)]}}.$$

Zauważmy, że estymator we wzorze (7.2.3) ma analogiczną postać jak wzór w przypadku częściowej wiarogodności (7.1.9). Powyższy wzór definiuje nam **liniową składkę wiarogodności** opartą na $S_1, \ldots, S_n$. Współczynnik przy $\overline{S}$ jest malejący wzgledem $\mathrm{E}\left[\sigma^2(\Theta)\right]$, a więc czym większa zmienność szkód, tym mniej wnoszą one informacji na temat estymowanej wielkości, z drugiej strony współczynnik jest rosnący względem $\mathrm{Var}\left[\mu(\Theta)\right]$ (wariancji składki indywidualnej), a więc czym większa zmienność między indywidualnymi polisami, tym większą wagę przykładamy do wartości szkód.

Uwaga 7.2.1 Zauważmy, że kładąc $B_i = S_i - \mathrm{E}\left[S|\Theta\right]$, $A = \mathrm{E}\left[S|\Theta\right] - \mu$, zmienne losowe S_i w naszym modelu można zareprezentować w następującej postaci

$$S_i = \mu + A + B_i$$

gdzie μ jest liczbą, natomiast wszystkie zmienne losowe A, B_i są nieskorelowane oraz $\mathrm{E}\left[A\right] = \mathrm{E}\left[B_i\right] = 0$, $\mathrm{Var}\left[A\right] = \mathrm{Var}\left[\mu(\Theta)\right]$, $\mathrm{Var}\left[B_i\right] = \mathrm{E}\left[\sigma^2(\Theta)\right]$. Rzeczywiście,

$$\begin{aligned}
\mathrm{Var}\left[B_i\right] &= \mathrm{E}\left[(S_i - \mu(\Theta))^2\right] = \mathrm{E}\left[S_i^2\right] - 2\mathrm{E}\left[S_i\mu(\Theta)\right] + \mathrm{E}\left[\mu(\Theta)^2\right] \\
&= \mathrm{Var}\left[S_i\right] + \left(\mathrm{E}\left[S_i\right]\right)^2 - 2\mathrm{E}\left[\mathrm{E}\left[S_i\mu(\Theta)|\Theta\right]\right] + \mathrm{E}\left[\mu(\Theta)^2\right] \\
&= \mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{Var}\left[\mu(\Theta)\right] + \left(\mathrm{E}\left[\mu(\Theta)\right]\right)^2 - 2\mathrm{E}\left[\mu(\Theta)\mathrm{E}\left[S_i|\Theta\right]\right] + \mathrm{E}\left[\mu(\Theta)^2\right] \\
&= \mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{Var}\left[\mu(\Theta)\right] + \left(\mathrm{E}\left[\mu(\Theta)\right]\right)^2 - 2\mathrm{E}\left[\mu(\Theta)^2\right] + \mathrm{E}\left[\mu(\Theta)^2\right] \\
&= \mathrm{E}\left[\sigma^2(\Theta)\right]
\end{aligned}$$

oraz

$$\begin{aligned}
\mathrm{Cov}\left[A, B_i\right] &= \mathrm{E}\left[AB_i\right] = \mathrm{E}\left[(S_i - \mathrm{E}\left[S|\Theta\right])(\mathrm{E}\left[S|\Theta\right] - \mu)\right] \\
&= \mathrm{E}\left[S_i\mathrm{E}\left[S|\Theta\right]\right] - \mu\mathrm{E}\left[S_i\right] - \mathrm{E}\left[\left(\mathrm{E}\left[S|\Theta\right]\right)^2\right] + \mu\mathrm{E}\left[\mathrm{E}\left[S|\Theta\right]\right] \\
&= \mathrm{E}\left[S_i\mu(\Theta)\right] - \mu^2 - \mathrm{E}\left[\mu(\Theta)^2\right] + \mu^2 = 0.
\end{aligned}$$

Tutaj: μ jest średnią wysokością szkód dla całej populacji, A jest losowym efektem właściwym dla całej populacji, B_i-indywidualnym losowym efektem dla i-tego ryzyka.

Aby ominąć potrzebę znajomości rozkładu a'posteriori w przypadku liniowym estymujemy współczynniki pojawiające się we wzorze (7.2.2). W tym celu oznaczmy: $\varphi = \mathrm{E}\left[\sigma^2(\Theta)\right]$, $v = \mathrm{Var}\left[\mu(\Theta)\right]$. Załóżmy, że mamy portfel składający się z N jednakowych, niezależnych polis, które obserwujemy przez n lat. Niech S_{ij} oznacza szkodę z i-tej polisy w j-tym roku. Przyjmijmy:

$$\overline{S}_i = \frac{1}{n}\sum_{j=1}^{n} S_{ij},\ \overline{S} = \frac{1}{N}\sum_{i=1}^{N}\overline{S}_i = \frac{1}{nN}\sum_{i=1}^{N}\sum_{j=1}^{n} S_{ij}.$$

Jako estymatory kładziemy:

$$\widehat{\mu} = \overline{S},\ \widehat{\varphi} = \frac{1}{N(n-1)} \sum_{i=1}^{N} \sum_{j=1}^{n} (S_{ij} - \overline{S}_i)^2.$$

Estymatory te są nieobciążone. Naturalnym estymatorem dla v powinno być

$$\frac{1}{N-1} \sum_{i=1}^{N} (\overline{S}_i - \widehat{\mu})^2.$$

Mamy jednak

$$\begin{aligned}
\frac{1}{N-1} \sum_{i=1}^{N} \mathrm{E}\left[(\overline{S}_i - \widehat{\mu})^2\right] &= \frac{N}{N-1} \mathrm{E}\left[\left(\frac{N-1}{N}\overline{S}_1 - \frac{1}{N}\sum_{j=2}^{N}\overline{S}_j\right)^2\right] \\
&= \frac{N}{N-1} \mathrm{E}\left[\left(\frac{N-1}{N}(\overline{S}_1 - \mu) - \frac{1}{N}\sum_{j=2}^{N}(\overline{S}_j - \mu)\right)^2\right] \\
&= \frac{N}{N-1}\left(\left(\frac{N}{N-1}\right)^2 \mathrm{E}\left[\left(\overline{S}_1 - \mu\right)^2\right] + \frac{N-1}{N^2}\mathrm{E}\left[\left(\overline{S}_1 - \mu\right)^2\right]\right) \\
&= \mathrm{E}\left[\left(\overline{S}_1 - \mu\right)^2\right] = \mathrm{E}\left[\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} S_{1i}S_{1j} - \frac{2}{n}\mu\sum_{j=1}^{n} S_{1j} + \mu^2\right] \\
&= \frac{1}{n}\left(\mathrm{E}\left[\mathrm{Var}\left[\Theta\right]\right] + \mathrm{Var}\left[\mu(\Theta)\right] + \mu^2\right) + \frac{n-1}{n}(\mathrm{Var}\left[\mu(\Theta)\right] + \mu^2) \\
&= \mathrm{Var}\left[\mu(\Theta)\right] + \frac{\mathrm{E}\left[\mathrm{Var}\left[\mu(\Theta)\right]\right]}{n}.
\end{aligned}$$

Oznacza to, że proponowany estymator jest obciążony. Możemy jednak położyć

$$\widehat{v} = \frac{1}{N-1} \sum_{i=1}^{N} (\overline{S}_i - \overline{S})^2 - \frac{\widehat{\varphi}}{n}.$$

Ponieważ jednak $\widehat{v}$ może przyjmować ujemne wartości, jako estymator bierzemy $(\widehat{v})_+ = \max\{\widehat{v}, 0\}$. Biorąc $\widehat{k} = \widehat{\varphi}/\ (\widehat{v})_+$ wzór (**??**) przybierze postać

$$\widehat{\mu}_L(\Theta) = \begin{cases} \frac{n}{n+\widehat{k}}\overline{S} + \frac{\widehat{k}}{n+\widehat{k}}\widehat{\mu} & , \quad (\widehat{v})_+ > 0 \\ \widehat{\mu} & , \quad (\widehat{v})_+ = 0 \end{cases}. \tag{7.2.4}$$

Przykład 7.2.2 Niech $S_1, \ldots, S_n$ będą warunkowo (pod warunkiem $\Theta = \theta$) niezależne o tym samym rozkładzie $Poi(\theta)$:

$$\underline{f}_\theta(x_1, \ldots, x_n) = P(S_1 = x_1, \ldots, S_n = x_n | \Theta = \theta) = e^{-n\theta} \frac{\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!},\ x_i \in \mathbb{N}$$

$$\pi(\theta) = \lambda^\alpha \theta^{\alpha-1} e^{-\lambda\theta}/\Gamma(\alpha) \quad \text{(gęstość gamma)}.$$

Wtedy rozkład bezwarunkowy dany jest wzorem

$$\begin{aligned}P(S_1 = x_1, \ldots, S_n = x_n) &= \frac{\lambda^\alpha}{\Gamma(\alpha)\prod_{i=1}^n x_i!}\int_0^\infty \theta^{\alpha-1}e^{-\lambda\theta}e^{-n\theta}\theta^{\sum_{i=1}^n x_i}d\theta \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)\prod_{i=1}^n x_i!}\int_0^\infty \theta^{\sum_{i=1}^n x_i+a-1}e^{-\theta(\lambda+n)}d\theta \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)\prod_{i=1}^n x_i!}\frac{\Gamma(\sum_{i=1}^n x_i+\alpha)}{(\lambda+n)^{\sum_{i=1}^n x_i+\alpha}}.\end{aligned}$$

Biorąc $n = 1$ dostajemy

$$\begin{aligned}P(S_1 = x_1) &= \frac{\lambda^\alpha}{\Gamma(\alpha)x_1}\frac{\Gamma(x_1+\alpha)}{(\lambda+1)^{x+\alpha}} \\ &= \binom{x_1+\alpha-1}{x_1}p^\alpha q^{x_1}\end{aligned}$$

z $p = \left(\frac{\lambda}{\lambda+1}\right) = 1-q$, czyli rozkład ujemny dwumianowy. Rozkład a posteriori na podstawie obserwacji $S_1 = x_1, \ldots, S_n = x_n$ dany jest wzorem (7.1.11), stąd

$$\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta) = \frac{(\lambda+n)^{\sum_{i=1}^n x_i+\alpha}}{\Gamma(\sum_{i=1}^n x_i+\alpha)}\theta^{\sum_{i=1}^n x_i+\alpha-1}e^{-(n+\lambda)\theta}$$

Jest to więc rozkład $Gamma(\sum_{i=1}^n x_i+\alpha, n+\lambda)$. Mamy wtedy też (patrz (7.1.12))

$$\begin{aligned}f_S(x|x_1,\ldots,x_n) &= \int_0^\infty e^{-\theta}\frac{\theta^x}{x!}\frac{(\lambda+n)^{\sum_{i=1}^n x_i+\alpha}}{\Gamma(\sum_{i=1}^n x_i+\alpha)}\theta^{\sum_{i=1}^n x_i+\alpha-1}e^{-(n+\lambda)\theta}d\theta \\ &= \frac{(\lambda+n)^{\sum_{i=1}^n x_i+\alpha}}{\Gamma(\sum_{i=1}^n x_i+\alpha)x!}\int_0^\infty \theta^{x+\sum_{i=1}^n x_i+\alpha-1}e^{-(n+\lambda+1)\theta}d\theta \\ &= \frac{(\lambda+n)^{\sum_{i=1}^n x_i+\alpha}}{\Gamma(\sum_{i=1}^n x_i+\alpha)x!}\frac{\Gamma(x+\sum_{i=1}^n x_i+\alpha)}{(n+\lambda+1)^{x+\sum_{i=1}^n x_i+\alpha}} \\ &= \binom{x+\sum_{i=1}^n x_i+\alpha-1}{x}\left(\frac{\lambda+n}{\lambda+n+1}\right)^{\sum_{i=1}^n x_i+\alpha}\frac{1}{(n+\lambda+1)^x}\end{aligned}$$

Ponieważ S pod warunkiem $\Theta = \theta$ ma rozkład $Poi(\theta)$, więc $\mathrm{E}[S|\Theta=\theta] = \mathrm{Var}[S|\Theta=\theta] = \theta$. Momenty zmiennej losowej S obliczamy następująco: $\mathrm{E}[S] = \mathrm{E}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\Theta] = \frac{\alpha}{\lambda}$, $\mathrm{Var}[S] = \mathrm{E}[\mathrm{Var}[S|\Theta]] + \mathrm{Var}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\Theta] + \mathrm{Var}[\Theta] = \frac{\alpha}{\lambda} + \frac{\alpha}{\lambda^2}$. Mamy także

$$\begin{aligned}\mathrm{Var}\left[\sum_{i=1}^n S_i\right] &= \mathrm{E}\left[\mathrm{Var}\left[\sum_{i=1}^n S_i|\Theta\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[\sum_{i=1}^n S_i|\Theta\right]\right] \\ &= \mathrm{E}[n\Theta] + \mathrm{Var}[n\Theta] = n\mathrm{E}[\Theta] + n^2\mathrm{Var}[\Theta] \\ &= n\frac{\alpha}{\lambda} + n^2\frac{\alpha}{\lambda^2}.\end{aligned}$$

Zauważmy, że wariancja sumy nie jest sumą wariancji, w szczególności zmienne $S_1, \ldots, S_n$ nie są niezależne. Warunkowa średnia i wariancja dla rozkładu a'posteriori dane są więc wzorami (korzystamy z momentów rozkładu Gamma): $\mathrm{E}\left[\Theta | S_1 = x_1, \ldots, S_n = x_n\right] = \frac{n\overline{S}+\alpha}{\lambda+n}$, $\mathrm{Var}\left[\Theta | S_1 = x_1, \ldots, S_n = x_n\right] = \frac{n\overline{S}+\alpha}{(\lambda+n)^2}$. Ponieważ $\mu(\Theta) = \Theta$ więc $H_{n+1} = \mathrm{E}\left[\Theta | S_1, \ldots, S_n\right] = \frac{n\overline{S}+\alpha}{\lambda+n}$. Stąd $\mathrm{E}\left[H_{n+1}\right] = \frac{n\mu+\alpha}{\lambda+n} = \frac{\alpha}{\lambda}$, $\mathrm{Var}\left[H_{n+1}\right] = \mathrm{Var}\left[\frac{n\overline{S}+\alpha}{\lambda+n}\right] = \frac{n\alpha}{\lambda}(1+\frac{n}{\lambda})/\left(n+\lambda\right)^2$. Estymator liniowy jest postaci (por. wzór (7.2.2))

$$\widehat{\mu}_L(\Theta) = \frac{n}{n+\lambda}\overline{S} + \frac{\alpha}{n+\lambda}.$$

Rzeczywiście, $\mu(\Theta) = \sigma^2(\Theta) = \Theta$ (z własności rozkładu Poissona) oraz $\mathrm{Var}\left[\mu(\Theta)\right] = \mathrm{Var}\left[\Theta\right] = \frac{\alpha}{\lambda^2}$, $\mathrm{E}\left[\sigma^2(\Theta)\right] = \mathrm{E}\left[\Theta\right] = \frac{\alpha}{\lambda}$, a więc **estymator liniowy jest równy składce bayesowskiej**.

□

Przykład 7.2.3 Niech $S_1, \ldots, S_n$ będą warunkowo (pod warunkiem $\Theta = \theta$) niezależne o tym samym rozkładzie $Exp(\theta)$:

$$f_\theta(x_1, \ldots, x_n) = \theta^n e^{-\theta\sum_{i=1}^n x_i}, \quad x_1, \ldots, x_n > 0,$$
$$\pi(\theta) = \lambda^\alpha \theta^{\alpha-1} e^{-\lambda\theta}/\Gamma(\alpha) \quad \text{(gęstość gamma)}.$$

Wtedy rozkład bezwarunkowy ma postać

$$f(x_1, \ldots, x_n) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \theta^{n+\alpha-1} e^{-\theta(\lambda+\sum_{i=1}^n x_i)} d\theta = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+n)}{(\lambda+\sum_{i=1}^n x_i)^{\alpha+n}}.$$

W szczególności, dla $n = 1$ jest to przesunięty rozkład Pareto. Rozkład a posteriori uzyskany na podstawie obserwacji $S_1 = x_1, \ldots, S_n = x_n$ jest postaci

$$\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta) = \frac{(\lambda+\sum_{i=1}^n x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \theta^{n+\alpha-1} e^{-\theta(\lambda+\sum_{i=1}^n x_i)},$$

jest to więc rozkład $Gamma(n+\alpha, \lambda+\sum_{i=1}^n x_i)$. Mamy także

$$\begin{aligned} f_S(x|x_1, \ldots, x_n) &= \int_0^\infty \theta e^{-\theta x} \frac{(\lambda+\sum_{i=1}^n x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \theta^{n+\alpha-1} e^{-\theta(\lambda+\sum_{i=1}^n x_i)} d\theta \\ &= \frac{(\lambda+\sum_{i=1}^n x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \int_0^\infty \theta^{n+\alpha} e^{-\theta(\lambda+\sum_{i=1}^n x_i+x)} d\theta \\ &= \frac{(\lambda+\sum_{i=1}^n x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \frac{\Gamma(n+\alpha+1)}{(\lambda+\sum_{i=1}^n x_i+x)^{n+\alpha}} \\ &= (n+a)\left(\frac{\lambda+\sum_{i=1}^n x_i}{\lambda+\sum_{i=1}^n x_i+x}\right)^{n+\alpha} \end{aligned}$$

Ponieważ S pod warunkiem $\Theta = \theta$ ma rozkład $Exp(\theta)$, więc $\mathrm{E}[S|\Theta = \theta] = 1/\theta$, $\mathrm{Var}[S|\Theta = \theta] = 1/\theta^2$. Momenty zmiennej losowej S możemy obliczyć bezpośrednio z własności rozkładu Pareto: $\mu = \mathrm{E}[S] = \frac{\lambda}{\alpha-1}$, $\alpha > 1$, $\mathrm{Var}[S] = \lambda^2 \frac{\alpha}{(\alpha-1)^2(\alpha-2)}$, $\alpha > 2$, lub w następujący sposób: $\mathrm{E}[S] = \mathrm{E}[\mathrm{E}[S|\Theta]] = \mathrm{E}[1/\Theta]$, $\mathrm{Var}[S] = \mathrm{E}[\mathrm{Var}[S|\Theta]] + \mathrm{Var}[\mathrm{E}[S|\Theta]] = \mathrm{E}[1/\Theta^2] + \mathrm{Var}[1/\Theta] = 2\mathrm{E}[1/\Theta^2] - (\mathrm{E}[1/\Theta])^2$. Wartości te obliczamy w sposób następujący:

$$\begin{aligned}\mathrm{E}\left[1/\Theta^k\right] &= \int_0^\infty \frac{1}{\theta^k}\lambda^\alpha \theta^{\alpha-1}e^{-\lambda\theta}/\Gamma(\alpha)d\theta = \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty \theta^{\alpha-1-k}e^{-\lambda\theta}d\theta \\ &= \frac{\lambda^k}{\Gamma(\alpha)}\Gamma(\alpha-k)\end{aligned}$$

dla $k > \alpha$, a stąd $\mathrm{E}[1/\Theta] = \frac{\lambda}{\alpha-1}$, $\mathrm{E}[1/\Theta^2] = \frac{\lambda^2}{(\alpha-1)(\alpha-2)}$. Dla wariancji sumy mamy natomiast:

$$\begin{aligned}\mathrm{Var}\left[\sum_{i=1}^n S_i\right] &= \mathrm{E}\left[\mathrm{Var}\left[\sum_{i=1}^n S_i\right]|\Theta\right] + \mathrm{Var}\left[\mathrm{E}\left[\sum_{i=1}^n S_i\right]|\Theta\right] \\ &= n\mathrm{E}\left[\frac{1}{\Theta^2}\right] + n^2\mathrm{Var}\left[\frac{1}{\Theta}\right] = n\lambda^2\frac{n+\alpha-1}{(\alpha-1)^2(\alpha-2)}.\end{aligned}$$

Warunkowa średnia i wariancja dane są więc wzorami: $\mathrm{E}[\Theta|S_1 = x_1, \ldots, S_n = x_n] = \frac{n+\alpha}{\lambda+n\overline{S}}$, $Var[\Theta|S_1 = x_1, \ldots, S_n = x_n] = \frac{n+\alpha}{(\lambda+n\overline{S})^2}$. Składka bayesowska zadana jest wzorem ($\mu(\Theta) = \frac{1}{\Theta}$):

$$H_{n+1} = \mathrm{E}\left[\frac{1}{\Theta}|S_1, \ldots, S_n\right] = \frac{n}{n+\alpha-1}\overline{S} + \frac{\alpha-1}{n+\alpha-1}\mu. \tag{7.2.5}$$

Rzeczywiście, rozkład a posteriori jest rozkładem $Gamma(n+\alpha, \lambda + \sum x_i)$ a moment rzędu -1 tego rozkładu wynosi $\frac{\lambda}{\Gamma(\alpha)}\Gamma(\alpha-1) = \frac{\lambda}{\alpha-1} = \mu$. Mamy także $\mathrm{E}[H_{n+1}] = \mu$, $\mathrm{Var}[H_{n+1}] = \frac{n}{n+\alpha-1}\frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}$. Estymator liniowy jest postaci (por. wzór (7.2.2))

$$\widehat{\mu}_L(\Theta) = \frac{n}{n+\alpha-1}\overline{S} + \frac{\alpha-1}{n+\alpha-1}\mu.$$

Rzeczywiście, $\mu(\Theta) = 1/\Theta$, $\sigma^2(\Theta) = 1/\Theta^2$, a stąd: $\mathrm{E}[\mu(\Theta)] = \frac{\lambda}{\alpha-1}$, $\mathrm{Var}[\mu(\Theta)] = \mathrm{Var}[1/\Theta] = \frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}$, $\mathrm{E}[\sigma^2(\Theta)] = \mathrm{E}[1/\Theta^2] = \frac{\lambda^2}{(\alpha-1)(\alpha-2)}$, a więc **estymator liniowy jest równy składce bayesowskiej**.

□

Przykład 7.2.4 Niech $X_1, \ldots, X_n$ będą warunkowo (pod warunkiem $\Theta = \theta$) niezależne o tym samym rozkładzie normalnym ze średnią θ i wariancją ξ^2, tzn.

$$f_\theta(x_1,\ldots,x_n) = \frac{1}{(2\pi)^{n/2}\xi^n}\exp(-\sum_{i=1}^{n}(x_i-\theta)^2/2\xi^2).$$

Zmienna losowa Θ ma rozkład $N(\eta,\sigma^2)$,

$$\pi(\theta) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-(\theta-\eta)^2/2\sigma^2).$$

Wtedy rozkład bezwarunkowy jest dany przez

$$f(x_1,\ldots,x_n) = \frac{1}{(2\pi)^{(n+1)/2}\xi^n\sigma}\int_{-\infty}^{\infty}\exp(-\sum_{i=1}^{n}(x_i-\theta)^2/2\xi^2)\exp(-(\theta-\eta)^2/2\sigma^2)d\theta.$$

Rozkład a posteriori uzyskany na podstawie obserwacji $S_1 = x_1,\ldots,S_n = x_n$ jest wtedy rozkładem normalnym ze średnią ze średnią $\frac{n}{n+k}\overline{X} + \frac{k}{n+k}\eta$ i wariancją $\frac{\sigma^2}{n+k}$, gdzie $k = \frac{\sigma^2}{\xi}$. Momenty zmiennej losowej S obliczamy następująco: $\mathrm{E}[S] = \mathrm{E}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\Theta] = \eta$, $\mathrm{Var}[S] = \mathrm{E}[\mathrm{Var}[S|\Theta]] + \mathrm{Var}[\mathrm{E}[S|\Theta]] = \mathrm{E}[\xi^2] + \mathrm{Var}[\Theta] = \xi^2 + \sigma^2$.

□

Zauważmy, że w powyższych przykładach estymatory liniowe były jednocześnie estymatorami bayesowskimi. Okazuje się jednak, że nie zawsze tak jest.

Przykład 7.2.5 Rozpatrzmy ubezpieczenia komunikacyjne. Załóżmy, że wartość szkody $S_i \in \{0,1\}$, $i \geqslant 1$, jest uzależniona od sposobu użytkowania samochodu. Niech Θ opisuje spsób użytkowania samochodu: $\Theta = 1$ oznacza, że ubezpieczony używa samochodu na dojazd do pracy, $\Theta = 0$ w przeciwnym wypadku. Zauważmy, że ta informacja jest najczęściej niedostępna dla towarzystwa ubezpieczeniowego. Przyjmijmy: $\pi(1) = P(\Theta = 1) = p = 1 - P(\Theta = 0) = 1 - \pi(0)$ oraz $f_\theta(1) = P(S_i = 1|\Theta = \theta) = \frac{1}{2}\theta + \frac{1}{4} = 1 - P(S_i = 0|\Theta = \theta) = 1 - f_\theta(0)$. Zakładamy, że ryzyka S_i, $i = 1,\ldots,n$, są warunkowo niezależne. Intuicyjnie, $\Theta = 1$ powinno oznaczać, że klient zapłaci większą składkę. Załóżmy, że w chwili $n+1$ przychodzi nowy klient (oczywiście, nie znamy jego parametru θ), mamy jednak obserwacje jego szkód $S_1 = x_1,\ldots,S_n = x_n$. Mamy ze wzoru (**??**)

$$\mathrm{E}[\mu(\Theta)|S_1 = x_1,\ldots,S_n = x_n] = \mu(0)\pi_{x_1,\ldots,x_n}(0) + \mu(1)\pi_{x_1,\ldots,x_n}(1). \tag{7.2.6}$$

Używając teraz wzoru (7.1.2) otrzymujemy $\mu(0) = f_0(1) = \frac{1}{4}$, $\mu(1) = f_1(1) = \frac{3}{4}$. Ze wzoru (7.1.10) i z warunkowej niezależności dostajemy

$$\begin{aligned} f(x_1,\ldots,x_n) &= \int f_\theta(x_1,\ldots,x_n)\pi(\theta)d\theta \\ &= \int f_\theta(x_1)\cdots f_\theta(x_n)\pi(\theta)d\theta \\ &= f_0(x_1)\cdots f_0(x_n)\pi(0) + f_1(x_1)\cdots f_1(x_n)\pi(1) \\ &= \left(\frac{1}{4}\right)^{\sum x_i}\left(\frac{3}{4}\right)^{n-\sum x_i}(1-p) + \left(\frac{3}{4}\right)^{\sum x_i}\left(\frac{1}{4}\right)^{n-\sum x_i}p. \end{aligned}$$

Z (7.1.11) i warunkowej niezależności dostajemy teraz

$$\begin{aligned}\pi_{S_1,\ldots,S_n}(0) &= f_0(x_1,\ldots,x_n)\pi(0)/f(x_1,\ldots,x_n)\\ &= f_0(x_1)\cdots f_0(x_n)\pi(0)/f(x_1,\ldots,x_n)\\ &= \left(\frac{1}{4}\right)^{\sum x_i}\left(\frac{3}{4}\right)^{n-\sum x_i}(1-p)/f(x_1,\ldots,x_n)\end{aligned}$$

oraz

$$\begin{aligned}\pi_{S_1,\ldots,S_n}(1) &= f_1(x_1,\ldots,x_n)\pi(1)/f(x_1,\ldots,x_n)\\ &= f_1(x_1)\cdots f_1(x_n)\pi(1)/f(x_1,\ldots,x_n)\\ &= \left(\frac{3}{4}\right)^{\sum x_i}\left(\frac{1}{4}\right)^{n-\sum x_i}p/f(x_1,\ldots,x_n).\end{aligned}$$

Wstawiając powyższe wartości do wzoru (7.2.6) otrzymujemy składkę dla $n+1$-klienta:

$$H_{n+1}=\frac{\frac{1}{4}\left(\frac{1}{4}\right)^{\sum x_i}\left(\frac{3}{4}\right)^{n-\sum x_i}(1-p)+\frac{3}{4}\left(\frac{3}{4}\right)^{\sum x_i}\left(\frac{1}{4}\right)^{n-\sum x_i}p}{\left(\frac{1}{4}\right)^{\sum x_i}\left(\frac{3}{4}\right)^{n-\sum x_i}(1-p)+\left(\frac{3}{4}\right)^{\sum x_i}\left(\frac{1}{4}\right)^{n-\sum x_i}p}.$$

W szczególności, **składka H_{n+1} obliczona metodą bayesowską nie jest liniową funkcją sumy szkód** jak w Rozdziale 7.2.

□

7.3 Składka wiarogodności: metoda wariancji

Składka wiarogodności daje się zdefiniować nie tylko dla metody netto wyliczania składki. Prześledzimy metodę wariancji. Zakładamy więc, że

$$H(\theta)=\mu(\theta)+\delta\sigma^2(\theta)$$

oraz θ zostało wybrane losowo zgodnie z dystrybuanta $U(\theta)$ o gęstości $\pi(\theta)$ (jako realizacja zmiennej losowej Θ). Wtedy składka w portfelu ma postać

$$H=\mu+\delta\sigma^2=\mathrm{E}\left[\mu(\Theta)\right]+\delta(\mathrm{E}\left[\sigma^2(\Theta)\right]+\mathrm{Var}\left[\mu(\Theta)\right]).$$

W języku wielkości S można powyższe wartości zapisać jako

$$H=\mathrm{E}\left[S\right]+\delta\mathrm{Var}\left[S\right],$$

$$H(\theta)=\mathrm{E}\left[S|\Theta=\theta\right]+\delta\mathrm{Var}\left[S|\Theta=\theta\right].$$

Przy ogólności przyjętych założeń, można zamiast Θ rozważyć dowolną inną zmienną losową, a nawet wektor zmiennych losowych. W szczególności biorąc $S_1, \ldots, S_n$ rozważamy

$$\mathrm{E}\left[S|S_1 = x_1, \ldots, S_n = x_n\right] + \delta \mathrm{Var}\left[S|S_1 = x_1, \ldots, S_n = x_n\right]$$

oraz

$$\mathrm{E}\left[H(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] + \delta \mathrm{Var}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right].$$

Okazuje się (zobacz Twierdzenie 7.1.8), że powyższe wielkości są sobie równe. Oznaczamy je H_{n+1} i nazywamy składką wiarogodności. Biorąc $H(\Theta) = \mu(\Theta) + \delta\sigma^2(\Theta)$ otrzymujemy z drugiego z powyższych wzorów

$$\begin{aligned} H_{n+1} &= \mathrm{E}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] \\ &\quad + \delta \mathrm{E}\left[\sigma^2(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] + \delta \mathrm{Var}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right]. \end{aligned}$$

Pierwsze dwa składniki są stanowią część przybliżającą składkę, trzeci składnik jest częścią fluktuacyjną.

Podobnie jak dla składki bayesowskiej netto powstaje problem wyliczenia poszczególnych składników znajdujących się w powyższym wzorze. Jest to natychmiastowe, jeżeli znamy rozkłady a'posteriori, w przeciwnym razie składniki te będziemy przybliżali funkcjami liniowymi. Przybliżenie pierwszego składnika znamy już ze wzoru (7.2.2): dla $k = \mathrm{E}\left[\sigma^2(\Theta)\right] / \mathrm{Var}\left[\mu(\Theta)\right]$ mamy

$$\mathrm{E}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] \approx b\overline{S} + (1-b)\mathrm{E}\left[\mu(\Theta)\right],$$

gdzie $b = \frac{n}{n+k}$, czyli

$$b = \frac{\mathrm{Var}\left[\mu(\Theta)\right]}{\frac{1}{n}\mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{Var}\left[\mu(\Theta)\right]}. \tag{7.3.1}$$

Drugi składnik składki aproksymujemy funkcją liniową postaci $d + c\Sigma^2$, gdzie $\Sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(S_i - \overline{S})^2$ Analogicznie otrzymuje się

$$\mathrm{E}\left[\sigma^2(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] \approx c\Sigma^2 + (1-c)\mathrm{E}\left[\sigma^2(\Theta)\right],$$

z

$$c = \frac{\mathrm{Var}\left[\sigma^2(\Theta)\right]}{\mathrm{E}\left[\mathrm{Var}\left[\Sigma^2|\Theta\right]\right] + \mathrm{Var}\left[\sigma^2(\Theta)\right]}. \tag{7.3.2}$$

Z ostatnią częścią postępujemy w sposób następujący. Mamy z definicji wariancji

$$\begin{aligned} &\mathrm{Var}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right] \\ &= \mathrm{E}\left[(\mu(\Theta) - \mathrm{E}\left[\mu(\Theta)|S_1 = x_1, \ldots, S_n = x_n\right])^2|S_1 = x_1, \ldots, S_n = x_n\right]. \end{aligned}$$

Korzystając jednak z przybliżenia pierwszego składnika otrzymujemy oszacowanie na $\mathrm{Var}\left[\mu(\Theta)|S_1 = x_1, \ldots\right.$

$$\mathrm{E}\left[(\mu(\Theta) - b\overline{S} - (1-b)\mathrm{E}\left[\mu(\Theta)\right])^2|S_1 = x_1, \ldots, S_n = x_n\right],$$

co przybliżamy poprzez

$$\begin{aligned}\mathrm{E}\left[(\mu(\Theta) - b\overline{S} - (1-b)E[\mu(\Theta)])^2\right] &= b^2\mathrm{E}\left[\frac{\sigma^2(\Theta)}{n}\right] + (1-b)^2\mathrm{Var}\left[\mu(\Theta)\right] \\ &= (1-b)\mathrm{Var}\left[\mu(\Theta)\right].\end{aligned}$$

Reasumując, estymator liniowy dla H_{n+1} jest postaci

$$b\overline{S} + (1-b)\mathrm{E}\left[\mu(\Theta)\right] + \delta(c\Sigma^2 + (1-c)\mathrm{E}\left[\sigma^2(\Theta)\right]) + \delta(1-b)\mathrm{Var}\left[\mu(\Theta)\right], \quad (7.3.3)$$

gdzie parametry b i c są wyliczone we wzorach (7.3.1) i (7.3.2).

Podobnie jak we wzorze (7.2.4) możemy estymować potrzebne wielkości we wzorze (7.3.3) na podstawie danych. I tak korzystając z tożsamości $\mathrm{Var}\left[\overline{S}\right] = \frac{1}{n}\mathrm{E}\left[\sigma^2(\Theta)\right] + \mathrm{Var}\left[\mu(\Theta)\right]$ otrzymujemy

$$b = \frac{n\mathrm{Var}\left[\overline{S}\right] - \mathrm{Var}\left[S\right]}{(n-1)\mathrm{Var}\left[\overline{S}\right]}, \quad (7.3.4)$$

$$c = \frac{(n-1)\mathrm{Var}\left[\Sigma^2\right] - \mathrm{Var}\left[\Sigma_2^2\right]}{(n-1)\mathrm{Var}\left[\Sigma^2\right]}, \quad (7.3.5)$$

gdzie $\Sigma_2^2 = (S_1 - \frac{S_1+S_2}{2})^2 + (S_2 - \frac{S_1+S_2}{2})^2$,

$$\mathrm{E}\left[\sigma^2(\Theta)\right] = \frac{n}{n-1}(\mathrm{Var}\left[S\right] - \mathrm{Var}\left[\overline{S}\right]), \quad (7.3.6)$$

$$\mathrm{Var}\left[\mu(\Theta)\right] = \frac{n\mathrm{Var}\left[\overline{S}\right] - \mathrm{Var}\left[S\right]}{n-1}. \quad (7.3.7)$$

Przykład 7.3.1 W przykładzie tym nie znamy rozkładu a priori. Rozważmy polisy samochodowe przyjmując dla uproszczenia, że jednorazowa wypłata wynosi 1. Na jedną polisę może przypaść k wypłat, $k \geqslant 0$. Mamy następujące dane dla roku 1:

ilość zgłoszeń	ilość polis	całkowita ilość zgłoszeń	
k	m	$k \cdot m$	k^2m
0	103704	0	0
1	14075	14075	14075
2	1766	3532	7064
3	255	765	2295
4	45	180	720
5	6	30	150
6	2	12	72
7	0	0	0

Mamy: $\sum m = 119853$, $\sum k \cdot m = 18594$, $\sum k^2 m = 24376$. Posiadane dane są danymi kolektywnymi. Niech S_1 oznacza ilość szkód na jedną polisę w pierwszym roku. Wtedy dla $\mathrm{E}[S_1] = \mu$ i $\mathrm{Var}[S_1] = \sigma^2$ mamy: $\widehat{\mu} = \frac{18594}{119853} = 0.155$, $\widehat{\sigma^2} = \frac{24376}{119853} - 0.155^2 = 0.179$. Stąd składka wariancji ma postać $H = 0.155 + 0.79b$. Załóżmy teraz, że ilość szkód na jedną polisę ma rozkład Poissona z parametrem θ (założenie to jest całkiem rozsądne: średnia i wariancja są sobie bliskie). W tym przypadku $\mu(\theta) = \sigma^2(\theta) = \theta$ i składka indywidualna wariancji ma postać: $H(\theta) = (1+\delta)\theta$. Składkę wiarogodności estymujemy w sposób następujący. $\mathrm{E}[\sigma^2(\Theta)] = \mathrm{E}[\Theta] = \mathrm{E}[\mu(\Theta)] = \mathrm{E}[S_1]$, $\mathrm{Var}[S_1] = \mathrm{Var}[\mathrm{E}[S_1|\Theta]] + \mathrm{E}[\mathrm{Var}[S_1|\Theta]] = \mathrm{Var}[\mu(\Theta)] + \mathrm{E}[\sigma^2(\Theta)] = \mathrm{Var}[\mu(\Theta)] + \mathrm{E}[\Theta]$, a stąd $\mathrm{Var}[\mu(\Theta)] = \mathrm{Var}[S_1] - \mathrm{E}[S_1]$. Wartość $\mathrm{E}[\sigma^2(\Theta)]$ estymujemy więc za pomocą $\widehat{\mu}$, a wartość $Var[\mu(\Theta)]$ za pomocą $\widehat{\sigma}^2 - \widehat{\mu}$. Stąd $\widehat{k} = \frac{\widehat{\mu}}{\widehat{\sigma^2} - \widehat{\mu}}$. Jeżeli teraz w następnych latach ilości szkód z pojedynczej polisy bedą wynosiły $S_1, \ldots, S_n$ to oszacowanie komponenty średniej we wzorze (7.3.3) ma postać $\frac{n}{n+\widehat{k}}\overline{S} + \frac{\widehat{k}}{n+\widehat{k}}\widehat{\mu} = \frac{n}{n+6.46}\overline{S} + \frac{6.46}{n+6.46}0.155$. Dla zmiennej losowej Poissona współczynnik c w komponencie wariacyjnej jest taki sam jak b (w obu przypadkach minimalizujemy tę samą zmienną losową Θ). Dla części fluktuacyjnej mamy $(1-b)\mathrm{Var}[\mu(\Theta)] = \frac{6.46}{n+6.46}0.024$. Stąd

$$P(k_1, \ldots, k_n) = (\frac{n}{n+6.46}0.155 + \frac{6.46}{n+6.46}\overline{S})(1+\delta) + \delta\frac{6.46}{n+6.46}0.024.$$

□

7.4 Estymatory największej wiarogodności (NW) dla modeli bayesowskich

Przykład 7.4.1 (cd. Przykładu 7.2.2) Estymatorem największej wiarogodności dla parametru $\mu(\Theta) = \Theta$ na bazie $S_1 = x_1, \ldots, S_n = x_n$ jest

$$\widehat{\Theta}_{NW} = \frac{n\overline{S} + \alpha - 1}{n + \lambda},$$

gdyż gęstość warunkowa $\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta)$ jest gęstością $Gamma\,(\sum_{i=1}^n x_i + \alpha, n + \lambda)$. Momenty estymatorów największej wiarogoności i bayesowskich dane są wzorami: $\mathrm{E}\left[\widehat{\Theta}_{NW}\right] = \frac{n\mu+\alpha-1}{n+\lambda}$, $\mathrm{Var}\left[\widehat{\Theta}_{NW}\right] = \mathrm{Var}\left[\frac{n\overline{S}+\alpha-1}{n+\lambda}\right] = \frac{n\alpha}{\lambda}(1+\frac{n}{\lambda})/\,(n+\lambda)^2$.

□

Przykład 7.4.2 (cd. Przykładu 7.2.3) Estymatorem największej wiarogodności dla parametru Θ na bazie $S_1 = x_1, \ldots, S_n = x_n$ jest (podobnie jak we wcześniejszym przykładzie)

$$\widehat{\Theta}_{NW} = \frac{n + \alpha - 1}{\lambda + n\overline{S}}.$$

Estymatorem największej wiarogodności dla parametru $\mu(\Theta) = \mathrm{E}\,[S|\Theta] = 1/\Theta$ na bazie $S_1 = x_1, \ldots, S_n = x_n$ jest

$$\widehat{\mu}_{NW}(\Theta) = \frac{n}{n+\alpha+1}\overline{S} + \frac{\alpha-1}{n+\alpha+1}\mu. \tag{7.4.1}$$

Rzeczywiście, gęstość zmiennej losowej Z i $1/Z$ są związane zależnością: $f_{1/Z}(z) = \frac{f_Z(1/z)}{z^2}$, stąd gęstość warunkowa zmiennej losowej $1/\Theta$ pod warunkiem $S_1 = x_1, \ldots, S_n = x_n$ dana jest wzorem:

$$\pi_{1/\Theta|S_1=x_1,\ldots,S_n=x_n}(\theta) = \frac{(\lambda + \sum_{i=1}^n x_i)^{\alpha+n}}{\Gamma(\alpha+n)}\theta^{-(n+\alpha+1)}e^{-(\lambda+\sum_{i=1}^n x_i)/\theta}$$

i maksimum jest osiągnięte w punkcie $\theta_0 = \frac{\sum_{i=1}^n x_i+\lambda}{n+\alpha+1}$ ale $\mu = \mathrm{E}\,[S] = \frac{\lambda}{\alpha-1}$. Nakładając wartość oczekiwaną na obie strony wzoru (7.4.1) otrzymujemy $\mathrm{E}\,[\widehat{\mu}_{NW}(\Theta)] = \frac{n+\alpha-1}{n+\alpha+1}\mu$, $\mathrm{Var}\,[\widehat{\mu}_{NW}(\Theta)] = \frac{\mathrm{Var}[\sum S_i]}{n(n+\alpha+1)} = \frac{n+\alpha-1}{n+\alpha+1}\frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}$. Zauważmy, że $\mu(\widehat{\Theta}_{NW}) = 1/\widehat{\Theta}_{NW} = \frac{n+\alpha+1}{\lambda+n\overline{S}}$, a więc policzenie najpierw estymatora NW dla Θ, a następnie wyliczenie składki $\mu(\widehat{\Theta}_{NW})$ daje inny wynik niż policzenie estymatora NW od razu dla $\mu(\Theta) = 1/\Theta$.

□

7.4.1 Porównanie modeli bayesowskich

$f_\theta(x_1, \ldots, x_n)$	$Poi(\theta)$
$\pi(\theta)$	$\Gamma(\alpha, \lambda)$
$f(x_1, \ldots, x_n)$	ujemny wielomianowy
$\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta)$	$\Gamma(\sum x_i + \alpha, n + \lambda)$
$\mathrm{E}[S] = \mu$	$\frac{\alpha}{\lambda}$
$\mathrm{Var}[S] = \sigma^2$	$\frac{\alpha(\alpha+1)}{\lambda^2}$
$\mathrm{E}[\Theta \mid S_1, \ldots, S_n]$	$\frac{n\overline{S}+\alpha}{\lambda+n}$
$\mathrm{Var}[\Theta \mid S_1, \ldots, S_n]$	$\frac{n\overline{S}+\alpha}{(\lambda+n)^2}$
$\mu(\Theta)$, $\sigma^2(\Theta)$	Θ, Θ
$H_{n+1} = \mathrm{E}[\mu(\Theta) \mid S_1, \ldots, S_n]$	$\frac{n\overline{S}+\alpha}{\lambda+n}$
$\mathrm{E}[H_{n+1}] = \mathrm{E}[\mu(\Theta)]$	$\frac{\alpha}{\lambda}$
$\mathrm{Var}[H_{n+1}]$	$\frac{n\alpha}{\lambda}(1 + \frac{n}{\lambda})/(n + \lambda)^2$
$\widehat{\Theta}_{NW}$	$\frac{n\overline{S}+\alpha-1}{\lambda+n}$
$\mathrm{E}\left[\widehat{\Theta}_{NW}\right]$	$\frac{n\mu+\alpha-1}{\lambda+n}$
$\mathrm{Var}\left[\widehat{\Theta}_{NW}\right]$	$\frac{n\alpha}{\lambda}(1 + \frac{n}{\lambda})/(n + \lambda)^2$
$\mathrm{Var}[\sum_{i=1}^{n} S_i]$	$\frac{n\alpha}{\lambda}(1 + \frac{n}{\lambda})$
$\widehat{\mu}_L(\Theta)$	$\frac{n}{n+\lambda}\overline{S} + \frac{\alpha}{n+\lambda}$
$\widehat{\mu}_{NW}(\Theta)$	$\frac{n\overline{S}+\alpha-1}{\lambda+n}$
$\mathrm{E}[\widehat{\mu}_{NW}(\Theta)]$	$\frac{n\mu+\alpha-1}{\lambda+n}$
$\mathrm{Var}[\widehat{\mu}_{NW}(\Theta)]$	$\frac{n\alpha}{\lambda}(1 + \frac{n}{\lambda})/(n + \lambda)^2$

$f_\theta(x_1,\ldots,x_n)$	$Exp(\theta)$
$\pi(\theta)$	$\Gamma(\alpha,\lambda)$
$f(x_1,\ldots,x_n)$	Par
$\pi_{S_1=x_1,\ldots,S_n=x_n}(\theta)$	$\Gamma(n+\alpha+1,\sum x_i+\lambda)$
$\mathrm{E}[S]=\mu$	$\frac{\lambda}{\alpha-1}$
$\mathrm{Var}[S]=\sigma^2$	$\lambda^2\frac{\alpha}{(\alpha-1)^2(\alpha-2)}$
$\mathrm{E}[\Theta\|S_1,\ldots,S_n]$	$\frac{n+\alpha}{\lambda+n\overline{S}}$
$\mathrm{Var}[\Theta\|S_1,\ldots,S_n]$	$\frac{n+\alpha}{(\lambda+n\overline{S})^2}$
$\mu(\Theta)$, $\sigma^2(\Theta)$	$1/\Theta$, $1/\Theta^2$
$H_{n+1}=\mathrm{E}[\mu(\Theta)\|S_1,\ldots,S_n]$	$\frac{n}{n+\alpha-1}\overline{S}+\frac{\alpha-1}{n+\alpha-1}\mu$
$\mathrm{E}[H_{n+1}]=\mathrm{E}[\mu(\Theta)]$	$\frac{\lambda}{\alpha-1}$
$\mathrm{Var}[H_{n+1}]$	$\frac{n}{n+\alpha-1}\frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}$
$\widehat{\Theta}_{NW}$	$\frac{n+\alpha-1}{\lambda+n\overline{S}}$
$\mathrm{E}\left[\widehat{\Theta}_{NW}\right]$	
$\mathrm{Var}\left[\widehat{\Theta}_{NW}\right]$	
$\mathrm{Var}[\sum_{i=1}^n S_i]$	$n\lambda^2\frac{n+\alpha-1}{(\alpha-1)^2(\alpha-2)}$
$\widehat{\Theta}_L$	$\frac{n}{n+\alpha-1}\overline{S}+\frac{\alpha-1}{n+\alpha-1}\mu$
$\widehat{\mu}_{NW}(\Theta)$	$\frac{n}{n+\alpha+1}\overline{S}+\frac{\alpha-1}{n+\alpha+1}\mu.$
$\mathrm{E}[\widehat{\mu}_{NW}(\Theta)]$	$\frac{n+\alpha-1}{n+\alpha+1}\mu$
$\mathrm{Var}[\widehat{\mu}_{NW}(\Theta)]$	$\frac{n+\alpha-1}{n+\alpha+1}\frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}$

Rozdział 8

Dodatek

8.1 Funkcje specjalne

- **Funkcja Gamma**

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0.$$

Zachodzi: $\Gamma(x+1) = x\Gamma(x)$.

- **Niekompletna funkcja Gamma**

$$\Gamma(x, a) = \int_a^\infty t^{x-1} e^{-t} dt, \quad x > 0.$$

- **Funkcja Beta**

$$B(a, b) := \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

Zachodzi: $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

8.2 Parametry i funkcje rozkładów

Niech X będzie zmienną losową. Ze zmiennymi losowymi będziemy utożsamiali następujące funkcje:

- **Dystrybuanta**

$$F(x) \quad = \quad F_X(x) := P(X \leqslant x);$$

- **Funkcja przeżycia (ogon rozkładu)**

$$\overline{F}(x) := 1 - F(x);$$

- **Gęstość**

$$f(x) \quad = \quad f_X(x) = \frac{d}{dx}F(x);$$

- **Funkcja tworząca momenty**

$$M(t) \quad = \quad M_X(t) = \mathrm{E}\left[\exp(tX)\right];$$

- **Funkcja tworząca kumulanty**

$$C(t) = C_X(t) = \log M_X(t).$$

- **Funkcja tworząca prawdopodobieństwa**

$$P(t) = P_X(t) = \mathrm{E}\left[t^X\right] = M_X(\log t).$$

Oznaczmy teraz $\mu_k(X) = \mathrm{E}\left[X^k\right]$, $m_k(X) = \mathrm{E}\left[(X - \mathrm{E}\left[X\right])^k\right]$. W przypadku, gdy wiadomo o jaką zmienną losową chodzi piszemy m_k i μ_k. Parametr μ_k nazywany jest k-tym **momentem zwykłym**, m_k - k-tym **momentem centralnym**. W szczególności $\mu_1 =: \mu$ jest średnią, a m_2 jest wariancją. Parametr $\gamma_3 := \frac{m_3}{m_2^{3/2}}$ jest nazywamy **skośnością**, a $\gamma_4 := \frac{m_4}{m_2^2} - 3$ **kurtozą**. Iloraz $\gamma_1 := \frac{m_2}{\mu}$ nazywamy **indeksem dyspersji**, a $\gamma_2 = \frac{\sqrt{m_2}}{\mu}$ **współczynnikiem zmienności**.

Drugi i trzeci moment centralny można wyrazić za pomocą momentów zwykłych:

$$\begin{aligned} \mathrm{Var}\left[X\right] &= \mathrm{E}\left[(X - \mathrm{E}\left[X\right])^2\right] = \mathrm{E}\left[X^2\right] - (\mathrm{E}\left[X\right])^2, \\ \mathrm{E}\left[(X - \mathrm{E}\left[X\right])^3\right] &= \mathrm{E}\left[X^3\right] - 3\mathrm{E}\left[X\right]\mathrm{E}\left[X^2\right] + 2(\mathrm{E}\left[X\right])^3. \end{aligned}$$

Zachodzą następujące wzory pozwalające wyliczać momenty za pomocą funkcji tworzących:

$$M^{(k)}(0) = \mathrm{E}\left[X^k\right],$$

$$C_X^{(1)}(0) = \mathrm{E}\left[X\right], \quad C_X^{(2)}(0) = \mathrm{Var}\left[X\right], \quad C_X^{(3)}(0) = \mathrm{E}\left[(X - \mathrm{E}\left[X\right])^3\right].$$

Tak więc pochodne funkcji tworzącej momenty pozwalają wyliczać momenty centralne, podczas gdy funkcji tworzącej C_X zwane są **kumulantami**.

8.3 Estymacja momentów

Najpopularniejszą metodą estymacji parametrów μ_k i m_k jest metoda momentów. Zasada jest następująca: estymujemy 'momenty teoretyczne' za pomocą odpowiednich momentów próbkowych dla danych $X_1, \ldots, X_k$. I tak

$$\hat{\mu_1} = \frac{1}{k}\sum_{i=1}^{k} X_i =: \bar{X}; \qquad \mathrm{E}\,[\hat{\mu_1}] = \mu_1, \quad \mathrm{Var}\,[\hat{\mu_1}] = \frac{\sigma^2}{k}$$

$$\hat{m}_2 = \frac{1}{k-1}\sum_{i=1}^{k}(X_i - \mu_1)^2 \quad \mu_1 \text{ znane},$$

$$\hat{m_2} = \frac{1}{k-1}\sum_{i=1}^{k}(X_i - \bar{X})^2 =: S^2 \quad \mu_1 \text{ nieznane}; \qquad \mathrm{E}\,[\hat{m_2}] = m_2; \quad \mathrm{Var}\,[\widehat{m_2}] = \frac{2m_2^2}{k-1}.$$

8.4 Rozkłady dyskretne

8.4.1 Rozkład dwumianowy $Bin(n,p)$

Jeżeli

$$P(X = m) = \binom{n}{m} p^m q^{n-m},$$

gdzie $p \in (0,1)$, $q = 1-p$, $m = 0, 1, \ldots, n$, to X ma rozkład dwumianowy $Bin(n,p)$. Mamy

$P(t)$	$\mathrm{E}\,[X]$	$\mathrm{Var}\,[X]$	γ_1	γ_2	γ_3	γ_4
$(q+pt)^n$	np	npq	q	$\frac{\sqrt{q}}{\sqrt{np}}$	$n\frac{(q-p)}{\sqrt{npq}}$	$3 + \frac{1-6pq}{npq}$

Rozważmy teraz próbę $X_1, \ldots, X_k \sim Bin(n_i, p)$, gdzie $\sum_{i=1}^{k} n_i = n$ jest znane. Parametr p rozkładu estymujemy w sposób następujący:

$$\begin{aligned} \widehat{p} &= \frac{\sum_{i=1}^{k} X_i}{n}, \\ \mathrm{E}\,[\widehat{p}] &= p, \\ \mathrm{Var}\,[\widehat{p}] &= pq\frac{1}{n}. \end{aligned}$$

8.4.2 Rozkład Poissona $Poi(\lambda)$

Jeżeli

$$P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda},$$

gdzie $\lambda > 0$, $m = 1, 2, 3, \ldots$, to X ma rozkład Poissona $Poi(\lambda)$.

$P(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
$\exp(\lambda(t-1))$	λ	λ	1	$\frac{1}{\sqrt{\lambda}}$	$\frac{1}{\sqrt{\lambda}}$	$\frac{1}{\lambda}$

Dla próby $X_1, \ldots, X_k \sim Poi(\lambda)$,

$$\begin{aligned} \widehat{\lambda} &= \frac{\sum_{i=1}^k X_i}{k} \\ \mathrm{E}\left[\widehat{\lambda}\right] &= \lambda \\ \mathrm{Var}\left[\widehat{\lambda}\right] &= \frac{\lambda}{k} \end{aligned}$$

8.4.3 Rozkład ujemny dwumianowy $Bin^-(r, p)$

Jeżeli

$$P(X = m) = \binom{r+m-1}{m} p^r q^m,$$

$r \in \mathbf{R}_+$, $m = 0, 1, \ldots$, tzn. X ma rozkład ujemny dwumianowy $Bin^-(r, p)$.

$P(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
$\left(\frac{p}{1-qt}\right)^r$	$\frac{rq}{p}$	$\frac{rq}{p^2}$	$\frac{1}{p}$	$\frac{1}{\sqrt{rq}}$	$\frac{1+q}{\sqrt{rq}}$	$3 + \frac{p^2+6q}{rq}$

Jeżeli $r \in \mathbf{N}$, to dostajemy rozkład Pascala, jeżeli $r = 1$ - rozkład geometryczny $Geo(p)$. Jeżeli X ma rozkład $Geo(p)$ to zmienna losowa M o rozkładzie warunkowym takim jak X pod warunkiem $X > 0$ ma przesunięty rozkład geometryczny z $P(M = n) = pq^n$, $n \in \{1, 2, \ldots\}$. Oba rozkłady geometryczne różnią się średnią, wariancje są takie same. Inaczej: M ma rozkład postaci $Geo(p) * \delta_1$.

Na podstawie próby $X_1, \ldots, X_k \sim Poi(\lambda)$, estymacja wygląda następująco:

1. r znane:

$$\begin{aligned} \widehat{p} &= \frac{r-1}{r + \sum_{i=1}^k X_i/k - 1}, \\ \mathrm{E}\left[\widehat{\lambda}\right] &= p \end{aligned}$$

2. r i p nieznane:

$$\begin{cases} \widehat{r} = \frac{\overline{X}^2}{S^2 - \overline{X}} \\ \frac{1-\widehat{p}}{\widehat{p}} = \frac{S^2}{\overline{X}} - 1 \end{cases}$$

8.5 Rozkłady ciągłe

8.5.1 Rozkład normalny

Gęstość zmiennej losowej X o rozkładzie normalnym ze średnią μ i wariancją σ^2 jest postaci

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-(x-\mu)^2/2\sigma^2).$$

Piszemy wtedy $X \sim N(\mu, \sigma^2)$. Jeżeli $\mu = 0$ i $\sigma^2 = 1$ to mówimy o standardowym rozkładzie normalnym.

Parametry:

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
$e^{\frac{(t\sigma)^2}{2}+t\mu}$,	μ	σ^2	$\frac{\sigma^2}{\mu}, \mu \neq 0$	$\frac{\sigma}{\mu}, \mu \neq 0$	0	0

Mając dane $X_1, \ldots, X_k$, parametry μ i σ estymujemy metodą momentów.

8.5.2 Rozkład odwrotny normalny $IG(\mu, \sigma^2)$

Niech X ma gęstość zadaną wzorem

$$f_X(x) = \sqrt{\frac{\sigma}{2\pi x^3}} \exp\left\{-\frac{\sigma}{2\mu}\left(\frac{x}{\mu} - 2 + \frac{\mu}{x}\right)\right\},$$

gdzie $\mu \in R$, $\sigma > 0$, $x \in \mathbf{R}$, tzn. X ma rozkład odwrotny normalny $IG(\mu, \sigma^2)$.

Parametry:

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
,	μ	$\frac{\mu^3}{\sigma}$	$\frac{\mu^2}{\sigma}$	$\sqrt{\frac{\mu}{\sigma}}$	$3\sqrt{\frac{\mu}{\sigma}}$	$\frac{15\mu}{\sigma}$

Dla próby $X_1, \ldots, X_k \sim IG(\mu, \sigma^2)$

$$\begin{aligned} \widehat{\mu} &= \overline{X} = \frac{1}{k}\sum_{i=1}^{k} X_i, \\ \widehat{\sigma} &= \left[\frac{1}{k}\sum_{i=1}^{k}\left(X_i^{-1} - \overline{X}^{-1}\right)^2\right]^{-1}. \end{aligned}$$

8.5.3 Rozkład logarytmiczno-normalny $LN(\mu, \sigma)$

Niech X ma gęstość zadaną wzorem

$$f(x) = \frac{1}{(x\sigma\sqrt{2\pi})} \exp\left(\frac{-(\log x - \mu)^2}{2\sigma^2}\right), \quad x > 0,\ \mu \in \mathbf{R},\ \sigma > 0.$$

Wtedy X ma rozkład logarytmiczno-normalny $LN(\mu, \sigma)$.

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
,	$e^{\mu+\frac{1}{2}\sigma^2}$	$\left(e^{\sigma^2}-1\right)e^{2\mu+\sigma^2}$			$\left(e^{\sigma^2}+2\right)\sqrt{e^{\sigma^2}-1}$	$e^{\sigma^4}+2e^{\sigma^3}+3e^{\sigma^2}-3$

Jeśli Y jest $N(\mu, \sigma)$, to $X = e^Y \sim LN(\mu, \sigma)$.
Dla próby $X_1, \ldots, X_k \sim LN(\mu, \sigma^2)$

$$\begin{aligned} \widehat{\mu} &= \frac{1}{k}\sum_{i=1}^{k} \log X_i, \\ \widehat{\sigma} &= \sqrt{\frac{1}{k}\sum_{i=1}^{k} (\log X_i - \widehat{\mu})^2}. \end{aligned}$$

8.5.4 Rozkład wykładniczy $Exp(\lambda)$

Niech X ma gęstość zadaną wzorem

$$f_X(x) = \lambda e^{-\lambda x},$$

gdzie $x > 0$, $\lambda > 0$, tzn. X ma rozkład wykładniczy $Exp(\lambda)$.
Parametry:

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
$\frac{\lambda}{\lambda - t}$,	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{\lambda}$	1	0	

Dla próby $X_1, \ldots, X_k \sim Exp(\lambda)$,

$$\widehat{\lambda} = \frac{1}{\sum_{i=1}^{k} X_i}.$$

8.5.5 Rozkład Gamma $Gamma(\alpha, \beta)$

Niech X ma gęstość zadaną wzorem

$$f_X(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},$$

$\alpha > 0$, $\beta > 0$, $x > 0$, tzn. X ma rozkład Gamma $Gamma(\alpha, \beta)$.

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
$\frac{\beta^\alpha}{(\beta-t)^\alpha}$,	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\frac{1}{\beta}$	$\frac{1}{\sqrt{\alpha}}$		

Jeżeli $X \sim \Gamma(1, \beta)$, to $X \sim Exp(\beta)$.

Dla próby $X_1, \ldots, X_k \sim \Gamma(\alpha, \beta)$,

$$\begin{aligned} \widehat{\alpha} &= \frac{\left(\overline{X}\right)^2}{\frac{1}{k-1}\sum_{i=1}^n (X_i - \overline{X})^2}, \\ \widehat{\beta} &= \frac{\overline{X}}{\frac{1}{k-1}\sum_{i=1}^k (X_i - \overline{X})^2}. \end{aligned}$$

8.5.6 Rozkład Weibulla $Wei(r, c)$

Niech X ma gęstość zadaną wzorem

$$f(x) = rcx^{r-1}\exp(-cx^r), \qquad x > 0,$$

gdzie $0 < r$ jest parametrem kształtu, $c > 0$ jest parametrem skali. Wtedy X ma rozkład Weibulla $Wei(r, c)$.

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
,	$\left(\frac{1}{c}\right)^{\frac{1}{r}} \Gamma\left(\frac{1}{r}+1\right)$	$\left(\frac{1}{c}\right)^{\frac{2}{r}} \left\{\Gamma\left(\frac{2}{r}+1\right) - \left[\Gamma\left(\frac{1}{r}+1\right)\right]^2\right\}$				

Jeśli X jest $Wei(1, c)$, to $X \sim Exp(c)$.

Dla próby $X_1, \ldots, X_n \sim Wei(r, c)$ parametry c i r estymujemy rozwiązują układ równań:

$$\begin{aligned} c\tfrac{1}{k}\textstyle\sum_{i=1}^k X_i^r &= 1 \\ \tfrac{r}{k}\textstyle\sum_{i=1}^k (cX_i^r - 1)\log X_i &= 1 \end{aligned}$$

8.5.7 Rozkład Pareto $Par(\alpha, c)$

Niech X ma gęstość zadaną wzorem

$$f(x) = (\frac{\alpha}{c})(\frac{c}{x})^{\alpha+1}, \qquad x > c,$$

gdzie $\alpha > 0$, a $c > 0$ jest parametrem skali. Wtedy X ma rozkład Pareto $Par(\alpha, c)$.

$M(t)$	$\mathrm{E}[X]$	$\mathrm{Var}[X]$	γ_1	γ_2	γ_3	γ_4
nie istnieje	$c\frac{\alpha}{\alpha-1},\ \alpha > 1$	$c^2\frac{\alpha}{(\alpha-1)^2(\alpha-2)},\ \alpha > 2$			$2\frac{\alpha+1}{\alpha-3}\sqrt{\frac{\alpha-2}{\alpha}},\ \alpha > 3$	

Bibliografia

[1] S. Asmussen (2000). *Ruin probabilities*. World Scientific.

[2] N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones, C.J. Nesbitt (1997). *Actuarial Mathematics*. The Society of Actuaries, Schaumburg, Illinois.

[3] F. Bassi, P. Embrechts, M. Kafetzaki (1997). *A survival kit on quantile estimation.* Switzerland: Departament of Mathematics, ETHZ, CH-8092 Zürich.

[4] J. Beirlant, J. L. Teugels, P. Vynckier (1996). *Practical analysis of extreme values.* Leuven University Press, Belgia.

[5] N. Bingham, C.M. Goldie, J.L. Teugels (1989). *Regular variation*. Cambridge University Press, Cambridge.

[6] H. Bühlmann (1970). *Mathematical Methods in Risk Theory*. Springer Verlag, Nowy Jork.

[7] C.D. Daykin, T. Pentikäinen i M. Pesonen (1994). *Practical Risk Theory for Actuaries.* Chapman & Hall, Londyn.

[8] A. Dąbrowski (2000). *Statystyka matematyczna.* Skrypt, Instytut Matematyczny, Uniwersytet Wrocławski

[9] P. Duszeńko (2001). *Reasekuracja; typy i ich własności*. Praca magisterska, Instytut Matematyczny, Uniwersytet Wrocławski.

[10] P. Embrechts, C. Klüppelberg i T. Mikosch (1997). *Modelling extremal events. For insurance and finance*. Springer-Verlag, Berlin.

[11] W. Feller (1981). *Wstęp do rachunku prawdopodobieństwa*. Wydawnictwo Naukowe PWN, Warszawa.

[12] J. Grandell (1991). *Aspects of risk theory*. Springer-Verlag, New York.

[13] J. Jakubowski i R. Sztencel. (2000) *Wstęp do teorii prawdopodobieństwa*, SCRIPT.

[14] P. Jaworski i J. Micał (2005). *Modelowanie matematyczne w finansach i ubezpieczeniach*, Poltext, Warszawa.

[15] J. Koronacki i J. Mielniczuk. (2001) *Statystyka dla studentów kierunków techniczych i przyrodniczych.* WNT.

[16] M. Mosior (2001). *Komputerowe aproksymacje prawdopodobieństwa ruiny i maksymalizacja zysków przy użyciu optymalnej kontroli dywidendy.* Praca dyplomowa, Instytut Matematyki, Politechnika Wrocławska.

[17] Müller, A. and Stoyan, D. *Comparison Methods for Stochastic Models and Risks.* Wiley, Chichester, 2002.

[18] W. Niemiro (1997). *Teoria wiarygodności.* Podyplomowe Studium Ubezpieczeń WNE UW.

[19] D. Pfeifer (1996-97). *Versicherungsmathematik.* Materiały do wykładów i ćwiczeń, Universität Hamburg.

[20] R.-D. Reiss, M. Thomas (1997). *Statistical Analysis of Extreme Values from Insurance, Finance, Hydrology and other fields.* Birkhäuser Verlag.

[21] T. Rolski, H. Schmidli, V. Schmidt, J. L. Teugels (1998). *Stochastic Processes for Insurance and Finance.* John Willey & Sons.

[22] H-P. Schmidli (2002). *Lecture notes on Risk Theory.* University of Aarhus, Dania.

[23] K. Schmidt (1996). *Lectures on Risk Theory.* Teubner, Stuttgart.

[24] E. Straub (1988). *Non-Life Insuracne Mathematics.* Springer, Association of Swiss Actuaries, Zurich.

[25] B. Sundt. (1993). *An Introduction to Non-Life Insurance Mathematics.* Verlag Versicherungswirtschaft e.V., Karlsruhe.

[26] S. Wieteska (2001). *Zbiór zadań z matematycznej teorii ryzyka ubezpieczeniowego.* Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

[27] A. Wojtala (2002). *Matematyka ubezpieczeń majątkowych - teoria i zadania.* Praca magisterska, Instytut Matematyczny, Uniwersytet Wrocławski.

Spis rysunków