2.1 Przypuśćmy, że $P(Y_i = 1) = 1 - P(Y_i = 0) = \pi, i = 1, 2, \ldots, n$ oraz że $\{Y_i\}$ są niezależne. Niech $Y = \sum_{i=1}^{n} Y_i$.

a) Podaj rozkład zmiennej Y i oblicz wariancję $V(Y)$ zmiennej Y.

b) Niech teraz współczynnik korelacji między dowolną parą zmiennych Y_i oraz zmienną $Y_j$jest taki sam i wynosi $\rho > 0$. Pokaż, że $V(Y) > n\pi(1-\pi)$.

c) Załóżmy, że prawdopodobieństwo sukcesu π nie jest znane. Niech rozkład a priori parametru π ma gęstość $g(\pi)$[1], czyli $P(Y_i = 1 \mid \pi) = \pi$ i π jest zmienną losową o rozkładzie, skupionym na odcinku $(0, 1)$, o wartości oczekiwanej μ i wariancji $\sigma^2 > 0$. Pokaż, że wtedy $V(Y) > n\mu(1-\mu)$.[2]

2.2 Niech dla ciągu niezależnych prób Bernouilliego z prawdopodobieństwem sunkcesu π zmienna Y_k oznacza liczbę sukcesów poprzedzających k-tą porażkę.

a) Pokaż, że mają one *uogólniony rozkład dwumianowy*

$$P(Y_k = y) = \binom{y+k-1}{y} \pi^y (1-\pi)^k, \; y = 0, 1, \ldots$$

b) Sprawdź, że rozkład ten spełnia równanie Orda. Opisz jak z wykresu Orda można odczytać parametry π i k.

c) Pokaż, że funkcja tworząca[3] dla tego rozkładu ma postać:

$$F(z) = \left(\frac{1-\pi}{1-\pi z}\right)^k$$

d) Posługując się funkcją tworzącą[4] pokaż, że

$$E(Y_k) = k\frac{\pi}{1-\pi}, \; V(Y_k) = k\frac{\pi}{(1-\pi)^2}$$

Zwróć uwagę, że dla tych danych wariancja jest bardzo duża: $V(Y_k) > E(Y_k)$

e) Udowodnij prawo małych liczb dla uogólnionego rozkładu dwumianowego: z założeń $k \to \infty$, $\pi \to 0$, $k\pi \to \mu$ wynika, że $Y = \lim Y_k$ ma rozkład Poissona.

---

[1] gdy π ma rozkład dyskretny, to g jest rozkładem prawdopodobieństwa

[2] Zjawisko opisane w punktach b) i c) w którym wariancja po wprowadzeniu pewnych zaburzeń w założeniach jest większa niż w przypadku a) nazywa się nadzmiennością (*overdispersion*)

[3] Funkcja tworząca dla rozkładu dyskretnego $\{p_k\}, k = 0, 1, \ldots$ jest opisana wzorem $F(z) = \sum_k z^k p_k$

[4] gdy nie znasz wzorów, możesz je znaleźć np w Wikipedii