

1.1 Wykres *poissonness*

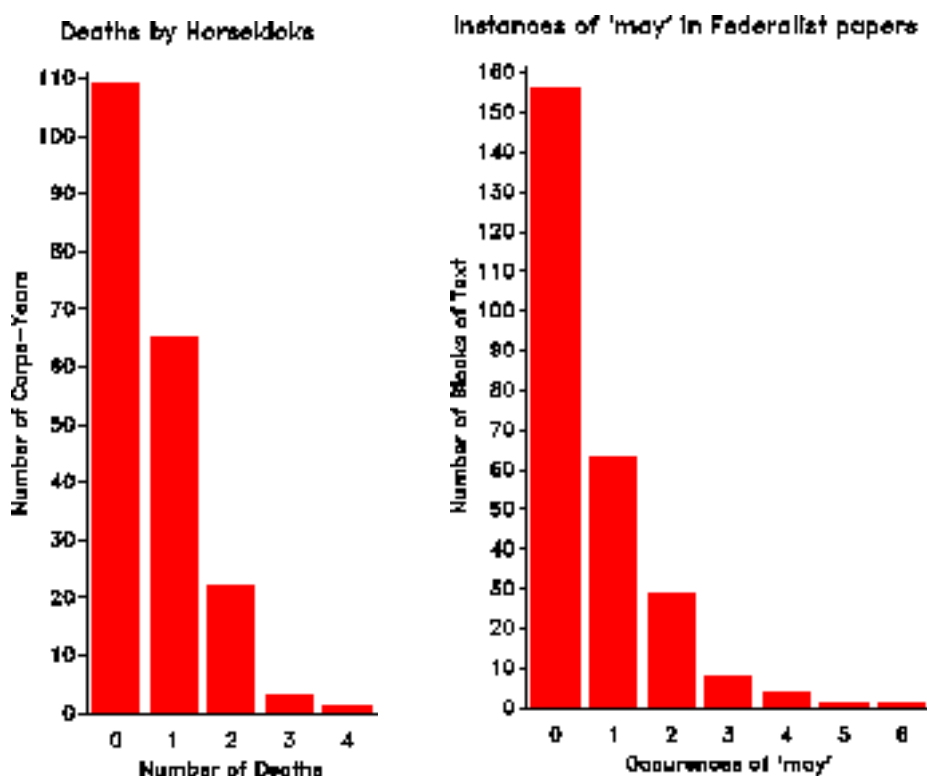
Dane:

*Dane von Bortkiewicza* (1898). Liczba wypadków śmiertelnych w 10 korpach armii pruskiej w ciągu 20 lat:

liczba wypadków	0	1	2	3	4
liczba obserwacji (korpasy x lata)	109	65	22	3	1

*Listy Federalistów*. Występowanie słowa *may* w 262 blokach po 200 słów.

liczba wystąpień	0	1	2	3	4	5	6
liczba bloków	156	63	29	8	4	1	1



*Metoda.*

1.1.1 Pokaż, że gdy w  $n_k$  próbach wystąpiło  $k$  sukcesów i gdy rozkład liczby sukcesów jest rozkładem Poissona z parametrem  $\lambda$  to zachodzi równość

$$u_k \stackrel{df}{=} \ln \left( \frac{k! n_k}{n} \right) = -\lambda + (\ln \lambda) k$$

Wielkość  $u_k$  nazywamy *pseudolicznikiem* (ang. *count metameter*)

1.1.2. Napisz za pomocą najwygodniejszego dla ciebie narzędzia procedurę, która rysuje wykres punktowy  $\{(k, u_k) : k = 0, 1, \dots\}$  oraz wpisuje w ten układ prostą regresji, oblicza jej równanie i drukuje wartość współczynnika determinacji  $R^2$ .

1.1.3. Oceń wizualnie, na podstawie sporządzonych wykresów czy można przyjąć, że *Dane von Bortkiewicza* pochodzą z rozkładu Poissona.

1.1.4. Zrób zadanie 1.1.3. Dla *Listów Federalistów*.

1.2. **Wykresy Orda.**

Metoda (Ord,1967)

1.2.1 Dla niektórych rozkładów dyskretnych o rozkładzie prawdopodobieństwa  $(k, p_k)$  zachodzi równanie:

$$v_k \stackrel{df}{=} \frac{k p_k}{p_{k-1}} = a + bk$$

Dla rzeczywistych danych  $p_k$  zastępuje się przez  $\frac{n_k}{n}$  co daje przybliżoną relację

$$v_k = \frac{k n_k}{n_{k-1}} = a + bk$$

Sprawdź, że równanie Orda zachodzi dla poniższych rozkładów. Wyznacz parametry  $a$  i  $b$

a)Poissona:  $p_k = e^{-\lambda} \frac{\lambda^k}{k!}, k = 0, 1, 2, \dots,$

b)dwumianowego:  $p_k = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \dots, n,$

c)ujemnego dwumianowego:  $p_k = (1-p)^k p, k = 0, 1, 2, \dots,$

d)logarytmicznego:  $p_k = -\frac{\vartheta^k}{k \ln(1-\vartheta)}, k = 1, 2, \dots$

Sformułuj proste kryteria graficzne identyfikacji rozkładów a)-d)

1.2.2. Oceń metodą Orda rozkład *Danych von Bortkiewicza i Listów Federalistów*

1.2.3. Oceń z jakiego rozkładu pochodzą dane

a) Liczba cząsteczek złota w cienkiej warstwie roztworu złota ( $k$  jest liczbą cząsteczek,  $n_k$  - liczbą okresów, w których obserwowano  $k$  cząsteczek):

k	0	2	3	4	5	6	7	8
$n_k$	112	168	130	68	32	5	1	1

b)Liczba organizacji społecznych, do których należą Amerykanie ( $k$  jest liczbą organizacji,  $n_k$  liczba ankietowanych, którzy należą do  $k$  organizacji)

$k$	0	1	2	3	4
$n_k$	1523	476	214	95	71