

Statystyczne modelowanie danych biologicznych

Egzamin I termin

Zadanie 1. (10 punktów) Dla dyskretnej zmiennej losowej, mającej rozkład zadany tabelą

X	10	11	12
P(x)	0.1	0.8	0.1

wyznaczyć średnią i wariancję. Przedstaw niezbędne obliczenia!

$$EX = 10 \cdot 0.1 + 11 \cdot 0.8 + 12 \cdot 0.1 = 11$$

$$EX^2 = 10^2 \cdot 0.1 + 11^2 \cdot 0.8 + 12^2 \cdot 0.1 = 121.2$$

$$\text{Var}X = EX^2 - (EX)^2 = 121.2 - 121 = 0.2$$

Zadanie 2. (10 punktów) Wypełnij brakujące dane w tabeli, a następnie odpowiedz na poniższe pytania.

Czas przeżycia (miesiące)	Ilość pacjentów (n_t)	Obserwowana ilość zgonów (d_t)	$p_t = 1 - \frac{d_t}{n_t}$	Funkcja przeżycia $S(t)$
0	100	0	1.000	1.000
3	99	10	0.899	0.899
6	80	25	0.688	0.619
12	50	5	0.900	0.557
18	35	10	0.714	0.397
24	20	5	0.750	0.298
30	10	0	1.000	0.298
36	10	0	1.000	0.298

- a) Ile wynosi kwantyl rzędu 25% dla czasu? **6 miesięcy**
- b) Obliczyć funkcję hazardu dla trzeciego roku i podać jej interpretację.
**Przedział [24, 36]. Ilość zgonów $d = 5$, $f_3 = 5 \cdot 1 = 5$, $F_3 = 5 \cdot 7 = 35$, $G_3 = 10 \cdot 12 = 120$
 $H_3 = 5 / (5 + 35 + 120) = 5 / 160 = 0.03125$ (w skali rocznej 0.375)**
- c) Wyznaczyć przedział ufności dla $S(12)$ przy pomocy metody transformacyjnej.
 $Ss \text{ Se}_{Tr}(12) = 10 / (99 \cdot (99 - 10)) + 25 / (80 \cdot (80 - 25))$

$$Se_{Tr}(12) = \frac{\sqrt{\frac{10}{99 \cdot (99 - 10)} + \frac{25}{80 \cdot (80 - 25)}}}{-\log \frac{99 - 10}{99} - \log \frac{80 - 25}{80}} = 0.172$$

$$CI(95\%) = (S(12)^{\exp(1.96 \cdot 0.172)}, S(12)^{\exp(-1.96 \cdot 0.172)}) = (0.441, 0.659)$$

Zadanie 3. (10 punktów) Na podstawie tabeli wyznacz wartość testu log-rank i odpowiedz na pytanie, czy ilość przeżytych zawałów serca wpływa na śmiertelność ludzi chorych na ostrą niewydolność serca.

Ilość zawałów	Ilość pacjentów	Obserwowana ilość zgonów	Oczekiwana ilość zgonów	$\chi^2_{\log\text{-rank}}$	p-wartość
0	30	15	12	7.639	0.054
1	30	15	10		
2	30	20	15		
>2	30	25	18		
Suma	120	75	55		

$$\chi^2 = \frac{(15 - 12)^2}{12} + \frac{(15 - 10)^2}{10} + \frac{(20 - 15)^2}{15} + \frac{(25 - 18)^2}{18} = 7.639 < \chi^2_3(0.95) = 7.81$$

Nie ma podstaw do odrzucenia H_0 .

Dodatkowo wybrać dwie grupy pacjentów i wyznaczyć dla nich miary: ryzyko względne, iloraz ryzyka i iloraz szans. Podać interpretacje uzyskanych współczynników.

Grupy 0 i 1.

$$HR = \frac{\frac{15}{12}}{\frac{15}{10}} = 0.83(3)$$

$$RR = \frac{\frac{15}{30}}{\frac{15}{30}} = 1.000$$

$$OR = \frac{\frac{15}{15}}{\frac{15}{15}} = 1.000$$

Zadanie 4. (10 punktów) Opisać proces losowania liczb pseudolosowych z rozkładu jednostajnego na odcinku $[0;4]$ przy pomocy dystrybuanty odwrotnej. Funkcja gęstości dla rozkładu jednostajnego na odcinku $[a;b]$ ma postać: $f(x) = \left(\frac{1}{b-a}\right)$. Wyznaczyć zmienną pseudolosową dla tego rozkładu odpowiadającą następującej zmiennej z rozkładu jednostajnego $u_1 = 0.35$. Przedstaw niezbędne obliczenia!!!

Dystrybuanta wynosi:

$$F(x) = \int_0^x f(x)dx = \int_0^x \frac{1}{4}dx = \frac{1}{4}x$$

Dystrybuanta odwrotna:

$$F^{-1}(x) = 4x$$

$$F^{-1}(u_1) = 4 \cdot 0.35 = 1.4$$

Do każdego z zadań przedstaw wykonane obliczenia oraz zinterpretuj uzyskane wyniki!

$$\chi_1^2(0.95) = 3.84, \chi_2^2(0.95) = 5.99, \chi_3^2(0.95) = 7.81$$