Lista_3

Analiza przeżycia

Olga Foriasz

2025-10-15

LISTA 3

Zadanie 1a - Oszacowania największej wiarygodności

Dane bierzeny z zadania 3 z litsy 2. Wyznaczymy oszacowania największej wiarogodności średniego czasu do remisji choroby dla pacjentów leczonych lekiem A oraz pacjentów leczonych lekiem B.

```
T_a = sum(grupa_A[1:10]) + t0*(n-R)
T_b = sum(grupa_B[1:10]) + t0*(n-R)

theta_a <- R/T_a
theta_b <- R/T_b
```

Średni czas do remisji dla leku A wynosi: 0.7034414, natomiast dla leku B: 0.7430155

Zadanie 1b - Realizacje przedziału ufności

Wyznaczamy realizację przedziału ufności na poziomie 1-α dla średniego czasu do remisji choroby osobno dla pacjentów leczonych lekiem A oraz lekiem B (dane z listy 2. z zadania 3.) Zakładamy, że są to zmienne z rozkładu wykładniczego. Za α przyjmujemy kolejno α=0.05 oraz α=0.01.

```
library(binom)

## dla alfy = 0.05
alfa=0.05
T.L <- binom.confint(R,n,conf.level=1-alfa, methods="exact")$lower
T.U <- binom.confint(R,n,conf.level=1-alfa, methods = "exact")$upper

TL1= (-log(1-T.L))/t0
TU1= (-log(1-T.U))/t0

TL1 = round(TL1,2)
TU1 = round(TU1,2)
## dla alfy = 0.01
alfa=0.01
T.L <- binom.confint(R,n,conf.level=1-alfa, methods="exact")$lower
T.U <- binom.confint(R,n,conf.level=1-alfa, methods = "exact")$upper
```

```
TL2= (-log(1-T.L))/t0
TU2= (-log(1-T.U))/t0

TL2 = round(TL2,2)
TU2 = round(TU2,2)
```

Realizacją przedziału ufności dla $\alpha = 0.05$ jest przedział $[T_L, T_U]$ =[0.32, 1.3], natomiast dla $\alpha =$ 0.01 przedział $[T_L, T_U]$ =[0.25, 1.52]

Zadanie 2 - Oszacowania największej wiarygodności i realizacja przedziału ufności dla danych z listy 2. zad.3

Przyjmujemy, że obserwacje czasu do remisji choroby były prowadzone do momentu, w którym u dziesięciu pacjentów zostanie ona zaobserwowana. Dla takich danych wyznaczamy najpierw oszacowania największej wiarygodności średniego czasu do remisji choroby dla pacjentów leczonych lekiem A oraz pacjentów leczonych lekiem B.

```
m=10
n=20
theta_a <- round(m/(sum(grupa_A[1:m])+(n-m)*grupa_A[m]),2)
theta_b <- round(m/(sum(grupa_B[1:m])+(n-m)*grupa_B[m]),2)
```

Zatem z wyliczeń wynika, że oszacowania największej wiarygodności średniego czasu do remisji w grupie A wynoszą: 0.73 a w grupie B: 0.96.

Następnie dla danych wyznaczamy realizację przedziału ufności na poziomie 1-α, dla odpowiednio dwóch wartości α: 0.05 i 0.01.

```
m=10
n=20
Ta<-sum(grupa_A[1:m])+(n-m)*grupa_A[m]
Tb<-sum(grupa_B[1:m])+(n-m)*grupa_B[m]
## dla alfy = 0.05
alfa=0.05
qm_a1 = qgamma(alfa/2,m,rate=m)
TLa <- round((m*qm_a1)/Ta,2)
qm_a2 = qgamma((1-alfa/2),m,rate=m)
TUa <-round((m*qm_a2)/Ta,2)

qm_b1 = qgamma(alfa/2,m,rate=m)
TLb <- round((m*qm_b1)/Tb,2)
qm_b2 = qgamma((1-alfa/2),m,rate=m)
TUb <-round((m*qm_b2)/Tb,2)
```

Dla α=0.05 realizacjami przedziałów ufności dla grupy A jest przedział [0.35, 1.25] a dla grupy B: [0.46, 1.65].

Następnie wyznaczamy realizację przedziału ufności dla α=0.01:

```r
m=10
n=20
## dla alfy = 0.01
alfa=0.01
qm_a1 = qgamma(alfa/2,m,m)
TLa <- round((m*qm_a1)/Ta,2)
qm_a2 = qgamma((1-alfa/2),m,m)
TUa <-round((m*qm_a2)/Ta,2)

qm_b1 = qgamma(alfa/2,m,m)
TLb <- round((m*qm_a1)/Tb,2)
qm_b2 = qgamma((1-alfa/2),m,m)
TUb <- round((m*qm_b2)/Tb,2)
```

Stąd otrzymujemy, że realizacjami przedziałów ufności dla grupy A jest przedział [0.27, 1.47] a dla grupy B: [0.36, 1.93]

Realizacją przedziału ufności dla $\alpha = 0.05$ jest przedział $[T_L, T_U]$ =[0.32,1.3], natomiast dla $\alpha = 0.01$ przedział $[T_L, T_U]$ =[0.25,1.52]

Zadanie 3

Przeprowadzono symulację w celu porównania dwóch estymatorów parametru ϑ rozkładu wykładniczego przy danych cenzurowanych I-go typu:

$$\hat{\vartheta} = \frac{R}{T_1} \tag{1}$$

$$\tilde{\vartheta} = -\frac{\log(1 - R/n)}{t_0} \tag{2}$$

gdzie $R = \sum_{i=1}^{n} 1(X_i \leq t_0)$ oraz $T_1 = \sum_{i=1}^{R} X_i + t_0(n - R)$.

Symulację wykonano dla $\vartheta = 1$, liczności prób $n = 10, 30$ oraz czasów obserwacji $t_0 = 0.5, 1, 2$, powtarzając eksperyment $M = 10{,}000$ razy. Dla każdego przypadku obliczono obciążenie (Bias) oraz błąd średniokwadratowy (MSE) obu estymatorów.

Tabela 1: Wyniki symulacji dla różnych wartości n i t0

n	t_0	Bias $\hat{\vartheta}$	Bias $\tilde{\vartheta}$	MSE $\hat{\vartheta}$	MSE $\tilde{\vartheta}$
10	0.5	0.0646	0.0759	0.3189	0.3481
10	1.0	0.0626	0.0854	0.1923	0.2242
10	2.0	-0.0544	-0.0782	0.0789	0.0607
30	0.5	0.0247	0.0267	0.0915	0.0953
30	1.0	0.0229	0.0306	0.0579	0.0667
30	2.0	0.0222	0.0514	0.0416	0.0688

Z przeprowadzonej analizy wynika, że wraz ze wzrostem liczności próby n obciążenie i MSE obu estymatorów maleją, co potwierdza ich zgodność. Większe wartości t_0 (czyli mniejsze cenzurowanie) prowadzą do dokładniejszych oszacowań. W większości przypadków estymator $\hat{\vartheta} = \frac{R}{T_1}$ charakteryzuje się mniejszym błędem średniokwadratowym, co wskazuje na jego większą efektywność w porównaniu z estymatorem $\tilde{\vartheta} = \frac{-\log(1-R/n)}{t_0}$.