

# Statistinės duomenų analizės praktinės užduotys

2017

## 5. Pasiklovimo intervalai.

```
library(knitr)
set.seed(42)
n <- 300
```

(a) Sukonstruoti 0.95 (0.90) - pasikliautinį intervalą *normalaus skirstinio* vidurkiui, naudojant imtį `norm1`

```
norm1 <- rnorm(n, mean=65, sd=11) # Dėstytojo sprendiniuose irgi imta 11, o ne sqrt(11)
```

Tarkime, kad  $X_1, \dots, X_n$  yra imtis,  $X_i \sim N(\mu, \sigma^2)$ . Vidurkio lygio  $1 - \alpha$  pasiklovimo intervalą

$$(\underline{\mu}, \bar{\mu}) = \bar{X} - t_{\alpha/2}(n-1) \frac{s}{\sqrt{n}}, \bar{X} + t_{\alpha/2}(n-1) \frac{s}{\sqrt{n}}$$

```
mu_u95 <- mean(norm1) + qt(1 - .05/2, n - 1) * sd(norm1) / sqrt(n)
mu_l95 <- mean(norm1) - qt(1 - .05/2, n - 1) * sd(norm1) / sqrt(n)
mu_u90 <- mean(norm1) + qt(1 - 0.1/2, n - 1) * sd(norm1) / sqrt(n)
mu_l90 <- mean(norm1) - qt(1 - 0.1/2, n - 1) * sd(norm1) / sqrt(n)

CI <- rbind(c(mu_l90, mu_u90), c(mu_l95, mu_u95))
colnames(CI) <- c("$\\underline{\\mu}$", "$\\overline{\\mu}$")
rownames(CI) <- c("0.90", "0.95")
kable(CI)
```

|      | $\underline{\mu}$ | $\bar{\mu}$ |
|------|-------------------|-------------|
| 0.90 | 63.72591          | 65.79495    |
| 0.95 | 63.52655          | 65.99431    |

(b) Sukonstruoti 0.95 (0.90) - pasikliautinį intervalą *normalaus skirstinio* dispersijai, naudojant tą pačią imtį `norm1`.

$1 - \alpha$  pasiklovimo intervalas *normalaus skirstinio* dispersijai yra

$$\underline{\sigma^2} = \frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)},$$

$$\overline{\sigma^2} = \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)}.$$

```
sigma_l95 <- (n - 1) * var(norm1) / qchisq(1 - .05/2, n - 1)
sigma_u95 <- (n - 1) * var(norm1) / qchisq(1 - (1 - .05/2), n - 1)
sigma_l90 <- (n - 1) * var(norm1) / qchisq(1 - .1/2, n - 1)
sigma_u90 <- (n - 1) * var(norm1) / qchisq(1 - (1 - .1/2), n - 1)

CI <- rbind(c(sigma_l90, sigma_u90), c(sigma_l95, sigma_u95))
colnames(CI) <- c("$\\underline{\\sigma^2}$", "$\\overline{\\sigma^2}$")
rownames(CI) <- c("0.90", "0.95")
kable(CI)
```

|      | $\sigma^2$ | $\overline{\sigma^2}$ |
|------|------------|-----------------------|
| 0.90 | 103.6149   | 135.6556              |
| 0.95 | 101.0998   | 139.3838              |

(c) Bandant sportinį lėktuvą, gautos šios jo maksimalaus greičio (m/s) reikšmės:

422.2; 418.7; 425.6; 420.3; 425.8; 423.1; 431.5; 428.2; 438.3; 434.0; 411.3; 417.2; 413.5; 441.3; 423.0.

Tarę, kad buvo stebimas *normalus a. d.*, raskite maksimalaus greičio vidurkio ir vidutinio kvadratinio nuokrypio taškinius įvertinius ir 0.95 pasiklovimo intervalus.

```
dat <- c(422.2, 418.7, 425.6, 420.3, 425.8,
        423.1, 431.5, 428.2, 438.3, 434.0,
        411.3, 417.2, 413.5, 441.3, 423.0)

estimates <- c(mean(dat), var(dat))
names(estimates) <- c("$\\hat{\\mu}$", "$\\hat{\\sigma}^2$")
kable(estimates)
```

|                  |           |
|------------------|-----------|
| $\hat{\mu}$      | 424.93333 |
| $\hat{\sigma}^2$ | 73.92952  |

```
mu_u <- mean(dat) + qt(1 - 0.05/2, n - 1) * sd(dat) / sqrt(n)
mu_l <- mean(dat) - qt(1 - 0.05/2, n - 1) * sd(dat) / sqrt(n)
sigma_l <- (n - 1) * var(dat) / qchisq(1 - .05/2, n - 1)
sigma_u <- (n - 1) * var(dat) / qchisq(1 - (1 - .05/2), n - 1)

CI <- rbind(c(mu_l, mu_u), c(sigma_l, sigma_u))
colnames(CI) <- c("apat.", "virs.")
rownames(CI) <- c("$\\mu$", "$\\sigma^2$")
kable(CI)
```

|            | apat.     | virs.     |
|------------|-----------|-----------|
| $\mu$      | 423.95642 | 425.91025 |
| $\sigma^2$ | 63.37525  | 87.37386  |

(d) Sukonstruoti 0.95 - pasikliautinį intervalą eksponentinio skirstinio vidurkiui, naudojant imtį *eksp.*

```
eksp <- rexp(n, rate=1 / 65)
```

$1 - \alpha$  pasiklovimo intervalas *eksponentinio skirstinio* vidurkiui yra

$$\underline{\theta} = \frac{2n\bar{X}}{\chi_{\alpha/2}^2(2n)},$$

$$\bar{\theta} = \frac{2n\bar{X}}{\chi_{1-\alpha/2}^2(2n)},$$

```
lambda_l <- 2*n * mean(eksp) / qchisq(1 - .05/2, 2*n)
# [1] 66.70573
lambda_u <- 2*n * mean(eksp) / qchisq(1 - (1 - .05/2), 2*n)
```

```
# [1] 83.66272
```

```
CI <- c(lambda_l, lambda_u)
names(CI) <- c("$\\underline{\\lambda}$", "$\\overline{\\lambda}$")
kable(CI)
```

|                       |          |
|-----------------------|----------|
| $\underline{\lambda}$ | 59.86283 |
| $\overline{\lambda}$  | 75.08031 |

- (e) Laikas nuo užsakymo pateikimo iki jo gavimo (pristatymo laikas) yra pasiskirstęs pagal *eksponentinę skirstinį*. Lentelėje pateikiamos atsitiktinai parinktų užsakymų pristatymo laikas ( $i$  - eilės nr.,  $X_i$  - laikas).

| $i$ | $X_i$ | $i$ | $X_i$ | $i$ | $X_i$ | $i$ | $X_i$ |
|-----|-------|-----|-------|-----|-------|-----|-------|
| 1   | 10    | 6   | 7     | 11  | 10    | 16  | 7     |
| 2   | 10    | 7   | 11    | 12  | 6     | 17  | 6     |
| 3   | 6     | 8   | 12    | 13  | 13    | 18  | 16    |
| 4   | 11    | 9   | 12    | 14  | 8     | 19  | 9     |
| 5   | 8     | 10  | 6     | 15  | 12    | 20  | 5     |

Raskite vidutinio pristatymo laiko taškinę įvertinę ir 0.9 pasiklovimo intervalą.

```
X <- c(10, 10, 6, 11, 8,
      7, 11, 12, 12, 6,
      10, 6, 12, 8, 12,
      7, 6, 16, 9, 5)

mean_hat <- mean(X)
names(mean_hat) <- "$\\hat{\\lambda}$"
kable(mean_hat)
```

|                 |     |
|-----------------|-----|
| $\hat{\lambda}$ | 9.2 |
|-----------------|-----|

```
lambda_l <- 2*length(X) * mean(X) / qchisq(1 - .1/2, 2*length(X))
lambda_u <- 2*length(X) * mean(X) / qchisq(1 - (1 - .1/2), 2*length(X))

CI <- c(lambda_l, lambda_u)
names(CI) <- c("$\\underline{\\lambda}$", "$\\overline{\\lambda}$")
kable(CI)
```

|                       |           |
|-----------------------|-----------|
| $\underline{\lambda}$ | 6.599893  |
| $\overline{\lambda}$  | 13.881919 |

- (f) Sukonstruoti 0.95 pasiklovimo intervalą Bernulio skirstinio parametrui  $p$ , naudojant imtį `ber`.
- ```
ber <- rbinom(n, size=1, prob=0.4)
```

$1 - \alpha$  pasiklovimo intervalas *Bernulio skirstinio parametrui*  $p$  yra  $(\underline{p}, \overline{p})$ , kur

$$\underline{p} = 1 - \beta_{\alpha/2}(n - T + 1, T), (T = 1, \dots, n),$$

$$\bar{p} = 1 - \beta_{1-\alpha/2}(n - T, T + 1), (T = 0, \dots, n - 1),$$

čia  $T = n\bar{X}$ ,  $\beta_\alpha(a, b)$  yra *beta skirstinio* kritinės reikšmės. Jei  $T = 0$ , tai  $\underline{p} = 0$ . Jei  $T = n$ , tai  $\bar{p} = 1$ .

```
p_l <- 1 - qbeta(1 - .05/2, n - n*mean(ber) + 1, n*mean(ber))
p_u <- 1 - qbeta(1 - (1 - .05/2), n - n*mean(ber), n*mean(ber) + 1)

CI <- c(p_l, p_u)
names(CI) <- c("$\\underline{p}$", "$\\overline{p}$")
kable(CI)
```

|                 |           |
|-----------------|-----------|
| $\underline{p}$ | 0.3280441 |
| $\bar{p}$       | 0.4409513 |

- (g) Bandat kiekvieną iš 10 prietaisų, nebuvo rasta nė vieno defektinio. Raskite tikimybės, kad prietaisas defektinis, pasiklovimo intervalą, kai pasiklovimo lygmenys yra 0.8, 0.9, 0.99. Išspręskite uždavinį, tarę, kad buvo rasti trys defektingi gaminiai.

```
CI <- c()

for (alpha in c(.2, .1, .01)) {
  p_l <- 1 - qbeta(1 - alpha/2, 10 - 3 + 1, 3)
  p_u <- 1 - qbeta(1 - (1 - alpha/2), 10 - 3, 3 + 1)
  CI <- rbind(CI, c(p_l, p_u))
}

CI <- cbind(CI, c(.8, .9, .99))
colnames(CI) <- c("$\\underline{p}$", "$\\overline{p}$", "$1 - \\alpha$")
kable(CI)
```

| $\underline{p}$ | $\bar{p}$ | $1 - \alpha$ |
|-----------------|-----------|--------------|
| 0.1158253       | 0.5517308 | 0.80         |
| 0.0872644       | 0.6066242 | 0.90         |
| 0.0370072       | 0.7351140 | 0.99         |

- (h) Sukonstruoti 0.95 pasikliautinį intervalą Puasono skirstinio vidurkiui, naudojant imtį `puas`.

```
puas <- rpois(n, lambda=65)
```

$1 - \alpha$  pasiklovimo intervalas *Puasono skirstinio* parametrui  $\lambda$  yra  $(\underline{\lambda}, \bar{\lambda})$ , kur

$$\underline{\lambda} = \frac{1}{2n} \chi_{1-\alpha/2}^2(2T), (T = 1, 2, \dots),$$

$$\bar{\lambda} = \frac{1}{2n} \chi_{\alpha/2}^2(2T + 2),$$

o  $T = \sum_{i=1}^n X_i$ . Jei  $T = 0$ , tai  $\underline{\lambda} = 0$ .

```
lambda_l <- 1 / (2*n) * qchisq(1 - (1 - .05/2), 2 * sum(puas))
lambda_u <- 1 / (2*n) * qchisq(1 - .05/2, 2 * sum(puas) + 2)

CI <- c(lambda_l, lambda_u)
names(CI) <- c("$\\underline{\\lambda}$", "$\\overline{\\lambda}$")
kable(CI)
```

$$\frac{\lambda}{\bar{\lambda}} \begin{array}{l} 64.97131 \\ 66.81169 \end{array}$$

- (i) Lentelėje pateikti skaičiai  $m_i$  tokių vienodo ploto ( $0.25 \text{ km}^2$ ) pietinės Londono dalies rajonų, į kuriuos Antrojo pasaulinio karo metu pataikė po  $i$  lėktuvų, sviedinių.

| $i$   | 0   | 1   | 2  | 3  | 4 | 5 | $\Sigma$ |
|-------|-----|-----|----|----|---|---|----------|
| $m_i$ | 229 | 211 | 93 | 35 | 7 | 1 | 576      |

Tarę, kad buvo stebimas *Puasono a. d.*, raskite parametro  $\lambda$  taškinį įvertinį, ir 0.95 pasikliovimo intervalą.

---

Padaryta su R version 3.4.2 (2017-09-28), x86\_64-pc-linux-gnu.