Асимптотична статистика
4 курс, статистика, Шкляр Ірина Володимирівна

Завдання 3, варіант 7

Побудуємо тест для перевірки простої основної гіпотези Ho проти альтернативи H1 для вибірки обсягу n. Поріг тесту спочатку визначимо використовуючи нормальну апроксимацію відношення вірогідності, а потім, застосовуючи імітаційне моделювання. H0 : Exp(6), H1 : Exp(5), n=55. alpha = 0.05.

1.

```
# експоненційна щільність
dexp_custom <- function(x, rate) {
        rate * exp(-rate * x)   }


# функція відношення вірогідності
likelihood_ratio <- function(x, n, lambda0, lambda1) {
        (lambda1 / lambda0)^n * exp(-(lambda1 - lambda0) * sum(x))    }


lambda0 <- 6  # H0 - нульова гіпотеза
lambda1 <- 5  # H1 - альтернативна
n <- 55
alpha <- 0.05


# генеруємо вибірку з експоненційного розподілу (нульова гіпотеза)
x <- rexp(n, lambda0)


# відношення вірогідності
lr <- likelihood_ratio(x, n, lambda0, lambda1)


# імітаційне моделювання
B <- 10000
lr_sim <- replicate(B, likelihood_ratio(rexp(n, lambda0), n, lambda0, lambda1))
c <- quantile(lr_sim, alpha)
```

```
# результат тесту
decision <- ifelse(lr <= c, "відхилити H0", "прийняти H0")
cat("результат:", decision, "\n")
```

результат: прийняти H0

Для розрахунку порогу тесту за асимптотичною формулою, використовуємо наближення відношення вірогідності логарифмічною функцією: ln(λ1/λ0) + n(λ0 - λ1). Ця статистика при нульовій гіпотезі H0 асимптотично нормально розподілена з математичним сподіванням -(n/2)(λ1 - λ0)^2/λ0 і дисперсією (λ1 - λ0)^2/λ0.

Тому, поріг тесту за асимптотичною формулою можна обчислити як: exp(-(n/2)(λ1 - λ0)^2/λ0 + z_alpha * sqrt((λ1 - λ0)^2/λ0)),

де z_alpha - квантиль стандартного нормального розподілу, який відповідає заданому рівню значущості alpha.

```
# Розрахунок порогу тесту за асимптотичною формулою

z_alpha <- qnorm(1 - alpha)
math_spos <- -(n/2) * (lambda1 - lambda0)^2 / lambda0
disp <- (lambda1 - lambda0)^2 / lambda0
threshold_approx <- exp(math_spos + z_alpha * sqrt(disp)) # Поріг
cat("поріг за асимптотичною формулою:", threshold_approx, "\n")
```

поріг за асимптотичною формулою: 0.02000397

2.

Для розрахунку потужності за асимптотичною формулою використовуємо інформацію за Фішером. У випадку експоненційного розподілу з параметром λ, інформація за Фішером дорівнює 1/λ.

Для альтернативної гіпотези H1: Exp(λ1) і нульової гіпотези H0: Exp(λ0), відстань за Кульбаком-Лейблером між двома розподілами дорівнює: KL = (λ1 - λ0)/λ0.

Асимптотична формула для потужності тесту відношення вірогідності має вигляд:
Потужність ≈ Φ(-sqrt(n*KL)),
де Φ - функція стандартного нормального розподілу, n - обсяг вибірки.

Отже, отримуємо: Φ(-sqrt(n*(λ1 - λ0)/λ0)).

```
# Інформація за Фішером для експоненційного розподілу: 1/lambda
# Відстань Кульбака-Лейблера: KL = (lambda1 - lambda0)/lambda0
# Асимптотична формула для потужності:
# Потужність ≈ Φ(-sqrt(n*KL))
# Φ - функція стандартного нормального розподілу

# потужність для альтернативи, близької до нульової гіпотези
lambda_alt <- 5.5
power_approx <- pnorm((lambda_alt - lambda0) / sqrt(lambda0 / n), lower.tail = FALSE)
cat("наближена потужність для lambda =", lambda_alt, ":", power_approx, "\n")
```

наближена потужність для lambda = 5.5 : 0.9349649

Отже, ймовірність відхилити H0, коли альтернатива $\lambda = 5.5$ є дійсною, дорівнює 0.9349649, тобто майже 1.

```
# перевірка точності наближення через імітаційне моделювання
power_sim <- mean(replicate(B, likelihood_ratio(rexp(n, lambda_alt), n, lambda0, lambda1)
<= c))
cat("потужність для lambda = 5.5 :", power_sim, "\n")
```

потужність для lambda = 5.5 : 0.0114

Точна потужність, обчислена за допомогою імітаційного моделювання для альтернативи $\lambda = 5.5$, становить 0.0114. Бачимо, що результат дуже відрізняється від наближеної оцінки, і є дуже малим.

3.

```
# мінімальний обсяг вибірки для забезпечення потужності 0.95

target_power <- 0.95
n_min <- ceiling((qnorm(target_power) / ((lambda1 - lambda0) / lambda0))^2 * 2)
cat("мінімальний обсяг вибірки:", n_min, "\n")
```

мінімальний обсяг вибірки: 195

Отже, в нашому випадку,щоб вийшли хороші результати має бути мінімальний обсяг вибірки = 195.