Лабораторна робота №3
з асимптотичної статистики
Варіант №4

Горбунова Даніела Денисовича
4 курс бакалаврату
група "комп'ютерна статистика"

18 травня 2021 р.

1 Вступ.

У даній роботі побудовано тест відношення вірогідностей для перевірки простих гіпотез про параметри розподілу спостережень з кратної вибірки. Оцінювання порогу тесту, його потужності проведено за допомогою нормального наближення у випадку альтернатив, що зближуються, та з використанням імітаційного моделювання. Додатково визначили мінімальний обсяг вибірки, при якому ймовірності помилки першого та другого роду не перевищуть 0.05.

2 Хід роботи.

2.1 Інформація за Фішером.

Для подальших підрахунків, нам необхідно обчислити інформацію за Фішером за одним спостереженням $\xi_1 \sim Bin(m, \theta)$, де $\theta \in (0, 1)$ - невідома ймовірність успіху, а $m \in \mathbb{N}$ - відома кількість випробувань. Нагадаємо, що розподіл ξ_1 описується за формулою:

$$p(k; \theta) := \mathbb{P}(\xi_1 = k) = C_m^k \theta^k (1-\theta)^{m-k},\ k = \overline{0, m}$$

Обчислимо логарифм від $p(k; \theta)$, а потім візьмемо похідну за невідомим параметром θ:

$$\ln(p(k;\theta)) = \ln C_m^k + k(\ln\theta - \ln(1-\theta)) + m\ln(1-\theta)$$
$$\frac{\partial}{\partial\theta}\ln(p(k;\theta)) = k\left(\frac{1}{\theta} - \frac{1}{1-\theta}\right) - \frac{m}{1-\theta} = \frac{k}{\theta(1-\theta)} - \frac{m}{1-\theta} = \frac{k - m\theta}{\theta(1-\theta)}$$

Обчислимо за означенням інформацію за Фішером $I_{\xi_1}(\theta)$:

$$I_{\xi_1}(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln(p(\xi_1;\theta))\right)^2\right] = \frac{1}{(\theta(1-\theta))^2}\sum_{k=0}^{m} p(k;\theta)\left(k^2 - 2km\theta + (m\theta)^2\right) =$$
$$= \frac{1}{(\theta(1-\theta))^2}\left(\mathbb{E}\left[\xi_1^2\right] - 2m\theta\mathbb{E}\left[\xi_1\right] + (m\theta)^2\right) = \frac{m\theta(1-\theta) + (m\theta)^2 - 2(m\theta)^2 + (m\theta)^2}{(\theta(1-\theta))^2} =$$
$$= \frac{m\theta(1-\theta)}{(\theta(1-\theta))^2} = \frac{m}{\theta(1-\theta)}$$

Інформацію за Фішером порахували, тепер можна сміливо приступати до основної частини завдання.

2.2 Побудова тесту відношення вірогідностей.

2.2.1 Статистика тесту. Рішуче правило.

Розглядається задача перевірка двох простих гіпотез про ймовірність успіху в біноміальному розподілі $Bin(2, \theta_*)$ незалежних спостережень з вибірки $\overline{\xi} = \{\xi_j\}_{j=1}^n$, $n := 65$:

$$\begin{aligned} H_0 &: \ \theta_* = \theta_0 = 0.5 \\ H_1 &: \ \theta_* = \theta_1 = 0.6 \end{aligned} \tag{1}$$

Вводимо статистику відношення вірогідності:

$$\mathrm{LR}(\overline{\xi}) = \frac{L(\overline{\xi};\theta_1)}{L(\overline{\xi};\theta_0)},\ L(\overline{\xi};\theta_i) = \prod_{j=1}^{n} p(\xi_j;\theta_i),\ i = 0, 1$$

де $p(k;\theta)$ - та сама функція вірогідності одного спостереження, що наведена в попередньому розділі (2.1). Запишемо тест π з фіксованим рівнем значущості $\alpha \in (0, 1)$ у вигляді:

$$\pi(\overline{\xi}) = \mathbb{1}_{\mathrm{LR}(\overline{\xi}) \geq C_\alpha},\ C_\alpha : \ \mathbb{P}_{H_0}\left(\pi(\overline{\xi}) = 1\right) = \alpha \tag{2}$$

Тест відношення вірогідностей (2) можна записати в еквівалентній формі, де розглядається логарифмована LR-статистика, тобто:

$$\pi(\overline{\xi}) = \mathbb{1}_{\ln \mathrm{LR}(\overline{\xi}) \geq \tilde{C}_\alpha},\ \tilde{C}_\alpha = \ln C_\alpha,\ \ln \mathrm{LR}(\overline{\xi}) = \sum_{j=1}^{n} \ln\left(\frac{p(\xi_j;\theta_1)}{p(\xi_j;\theta_0)}\right) \tag{3}$$

У даній задачі знаходити явний вигляд порогу тесту ми не будемо, але спробуємо його наближено обчислити. Для цього скористаємося нормальною апроксимацією відношення вірогідності у випадку альтернатив, що зближуються.

2.2.2 Оцінювання порогу тесту.

Для цього зауважимо, що гіпотези з (1) можемо представити через фіксовані альтернативи, що зближуються:

$$\begin{aligned} H_0 &: \theta_* = \theta_0 = 0.5 \\ H_1 &: \theta_* = \theta_1 = 0.6 = \theta_0 + \frac{v}{\sqrt{n}}, \end{aligned} \tag{4}$$

де швидкість збіжності v неважко визначити з рівняння, знаючи n та θ_i:

$$\theta_1 = \theta_0 + \frac{v}{\sqrt{n}} \Leftrightarrow v = \sqrt{n}(\theta_1 - \theta_0) = 0.8062258$$

Якщо справджується H_0, то справедлива збіжність $\mathrm{LR}(\bar{\xi})$ за розподілом:

$$\mathrm{LR}(\bar{\xi}) \to |v|\zeta_\infty - \frac{v^2}{2} I_{\xi_1}(\theta_0),\ n \to \infty;\ \zeta_\infty \sim N(0, I_{\xi_1}(\theta))$$

Внаслідок цього, помилка першого роду тесту збігається до ймовірності вигляду:

$$\alpha(\pi) \to \mathbb{P}\left(|v|\zeta_\infty - \frac{v^2}{2} I_{\xi_1}(\theta_0) > C\right) = \mathbb{P}\left(\frac{\zeta_\infty}{\sqrt{I_{\xi_1}(\theta_0)}} > \frac{\frac{v^2}{2} I_{\xi_1}(\theta_0) + C}{|v|\sqrt{I_{\xi_1}(\theta_0)}}\right),$$

звідки знаходимо поріг тесту $C = C_\alpha$:

$$C_\alpha = -\frac{v^2}{2} I_{\xi_1}(\theta_0) + Q^{N(0,1)}(1-\alpha)|v|\sqrt{I_{\xi_1}(\theta_0)} \tag{5}$$

Поріг (5) забезпечує асимпт. ймовірність помилки першого роду рівній α. Застосуємо це для наближеного обчислення порогового значення тесту (2):

```
# Обсяг вибірки, що обстежується
n <- 65
# Кількість випробувань
m <- 2
# Ймовірність у випадку нульової гіпотези
theta.0 <- 0.5
# Ймовірність у випадку альтернативи
theta.1 <- 0.6
# Обчислення швидкості збіжності альтернатив
v <- sqrt(n)*(theta.1 - theta.0)
# Рівень значущості та відповідний квантиль
alpha <- 0.05
q.alpha <- qnorm(1 - alpha)
# Інформація за Фішером одного спостереження
fisher.info <- function(m, theta) m/(theta*(1-theta))
fish.0 <- fisher.info(2, theta.0)
# Наближене значення порогу тесту
C.appr <- -v^2/2*fish.0+q.alpha*abs(v)*sqrt(fish.0)
print(C.appr)
# 1.150843
```

Спробуємо порівняти нормальне наближення із тим, що буде отримано за допомогою імітаційного моделювання.

Для цього ми визначимо наближення $\hat{C}_{\alpha,B}$ як квантиль рівня $1-\alpha$ за вибіркою з lnLR-статистик на кожній повторній вибірці $\overline{\xi}_{H_0,b}$, $b=\overline{1,B}$ за виконання нульової гіпотези.

```
# Статистика логарифмічного відношення вірогідностей
log.LR.binom <- function(x, m, p.0, p.1)
{
  l.f <- function(p) {log(dbinom(x, size=m, prob=p))}
  sum(l.f(p.1)-l.f(p.0))
}
set.seed(1)
# Обчислення бустрепованого порогу тесту
C.boot <- function(B)
{
  B <- 10^3
  log.LR.boot <- replicate(B,
                           {
                             x.boot <- rbinom(n, m, theta.0)
                             log.LR.binom(x.boot, m, theta.0, theta.1)
                           })
  C.quan <- quantile(log.LR.boot, 1 - alpha)
  C.quan
}
# B = 1000
print(C.boot(10^3))
# 1.401221
# B = 10000
print(C.boot(10^4))
# 0.9957563
```

Як бачимо із результатів, уточнення порогу тесту за допомогою імітаційного моделювання дає менші значення, порівняно з нормальним наближенням:

$$\hat{C}_{\alpha,1000} = 1.401221,\ \hat{C}_{\alpha,10000} = 0.9957563$$

Можливо це можна пояснити тим, що апроксимація не зовсім адекватно спрацювала на вибірці відносно малого обсягу. Далі переходимо до наближеного обчислення потужності тесту.

2.2.3 Оцінювання потужності тесту.

Для наближеного обчислення потужності тесту, природньо оцінити лише похибку другого роду тесту, що досліджується. Як і в попередньому випадку, оцінювати будемо двома способами: за допомогою нормальної апроксимації у випадку альтернатив, що зближуться, та з використанням імітаційного моделювання. Декілька слів щодо першого підходу: якщо виконується H_1, то має місце наступна збіжність за розподілом:

$$\mathrm{LR}(\overline{\xi}) \to |v|\zeta_\infty + \frac{v^2}{2} I_{\xi_1}(\theta_0),\ n\to\infty;\ \zeta_\infty \sim N(0, I_{\xi_1}(\theta))$$

Внаслідок цього, помилка другого роду тесту збігається до ймовірності вигляду:

$$\beta(\pi) \to \mathbb{P}\left(|v|\zeta_\infty + \frac{v^2}{2} I_{\xi_1}(\theta_0) < C\right) = \mathbb{P}\left(\frac{\zeta_\infty}{\sqrt{I_{\xi_1}(\theta_0)}} < \frac{C - \frac{v^2}{2} I_{\xi_1}(\theta_0)}{|v|\sqrt{I_{\xi_1}(\theta_0)}}\right) = \Phi\left(\frac{C - \frac{v^2}{2} I_{\xi_1}(\theta_0)}{|v|\sqrt{I_{\xi_1}(\theta_0)}}\right)$$

Обчислимо ймовірність похибки другого роду, використовуючи запропоновану апроксимацію.

```
# Наближене значення помилки другого роду тесту
beta.appr <- pnorm((C.boots - v^2/2*fish.0)/(v*sqrt(fish.0)))
print(beta.appr)
# 0.2408698
# Потужніть тесту
print(1 - beta.appr)
# 0.7591302
```

У випадку імітаційного моделювання, оцінку помилки другого роду визначимо як частоту помилок класифікації, коли вірна альтернатива:

$$\hat{\beta} = \frac{1}{B} \sum_{b=1}^{B} \mathbb{1}_{\ln \mathrm{LR}(\bar{\xi}_{H_1,b}) < C_\alpha}$$

Проводимо обчислення при $B = 1000$:

```
# Бустреповані вибірки з альтернативи
log.LR.boot.H1 <- replicate(1000,
                            {
                              x.boot <- rbinom(n, m, theta.1)
                              log.LR.binom(x.boot, m, theta.0, theta.1)
                            })
# Оцінка помилки другого роду, B = 1000
print(mean(log.LR.boot.H1 < C.boots))
# 0.204
# Потужніть тесту, B = 1000: 0.796
...
# Оцінка помилки другого роду, B = 10000
print(mean(log.LR.boot.H1 < C.boots))
# 0.211
# Потужніть тесту, B = 10000: 0.789
```

Порівнюючи отримані оцінки бачимо несуттєве відхилення значень за бутстрепованою оцінкою від нормальної апроксимації. Але в цілому, значення більш-менш узгоджені. Залишається визначити мінімальний обсяг вибірки, для якого забезпечується такі ймовірності похибок тесту, що не перевищують 0.05.

2.2.4 Визначення мінімального обсягу вибірки.

Ми хочемо знайти мінімальний обсяг вибірки, для якого $\alpha(\pi) < \alpha_0, \beta(\pi) < \beta_0$, де $\alpha_0, \beta_0 \in (0,1)$ - фіксовані. Застосуємо наближену формулу для визначення кількості елементів:

$$n = \frac{(Q^{N(0,1)}(1-\alpha_0) + Q^{N(0,1)}(1-\beta_0))^2}{I_{\xi_1}(\theta_0)(\theta_1 - \theta_0)^2}$$

Для $\alpha_0 = \beta_0 = 0.05$, маємо приблизний мінімальний обсяг вибірки:

```
n.min <- (2*q.alpha)^2/(fish.0*(theta.1 - theta.0)^2)
print(n.min)
# 135.2772 <=> n_min := 136
```

Цікаво, чи справді цього достатньо? Спробуємо розібратися.

2.2.5 Теорія на практиці.

Оцінимо ймовірності помилок першого та другого роду з використанням імітаційного моделювання при $B = 10000$. Оцінки будемо знаходити для кожного $n \in \{136, 150, 175, 200\}$. У тесті будемо використовувати бутстрепований поріг $\hat{C}_{\alpha,10000}$. Маємо такі результати:

n	$\hat{\alpha}$	$\hat{\beta}$
136	0.0222	0.0955
150	0.0226	0.0708
175	0.0141	0.0595
200	0.0125	0.0403

Рис. 1: Оцінки ймовірностей помилок тесту для різних обсягів вибірки n.

Ми можемо припустити, що для забезпечення того, щоб ймовірності помилок тесту не перевищували встановлений поріг 0.05, потрібно мати вибірку як мінімум з 200 елементів. Запропонована апроксимація не дає коректну відповідь на поставлене питання.

3 Висновки.

Побудували тест, порівняли різні методи оцінювання його характеристик. Різниця між значеннями, отриманими за допомогою нормальної апроксимації, та бустрепованими досить суттєва. Можливо це можна пояснити незначним стартових обсягом вибірок, що становить $n = 65$ одиниць. Припускається, що при незначному зільшенні n, апроксимовані результати могли бути кращими, особливо для розв'язання останньої задачі.