中級統計学：第 3 回中間試験

村澤 康友

2023 年 12 月 22 日

注意： 3 問とも解答すること．結果より思考過程を重視するので，途中計算等も必ず書くこと（部分点は大いに与えるが，結果のみの解答は 0 点とする）．教科書のみ参照してよい（他の講義資料・ノートは持込不可）．

1. （20 点）以下で定義される統計学の専門用語をそれぞれ書きなさい．
   （a）考察の対象全体
   （b）標本の関数
   （c）期待値が母数と等しい推定量
   （d）ある確率で母数を含む確率的な領域
2. （30 点）分布表を用いて以下の問いに答えなさい．
   （a）$X \sim \chi^2(2)$ とする．$\Pr[a \leq X \leq b] = .99$ となる a, b を求めなさい．
   （b）$Y \sim \mathrm{t}(4)$ とする．$\Pr[|Y| \leq c] = .95$ となる c を求めなさい．
   （c）$Z \sim \mathrm{F}(6, 8)$ とする．$\Pr[d \leq Z \leq e] = .9$ となる d, e を求めなさい．
   なお a〜e はすべて正の実数（$0, \infty$ は含まない）とする．
3. （50 点）岸田内閣の支持率 p を区間推定したい．ベルヌーイ母集団から抽出した大きさ n の無作為標本 $(X_1, \dots, X_n)$ における支持率を $\hat{p}_n$ とする．
   （a）$\mathrm{Bin}(1, p)$ の平均と分散を求めなさい（要証明）．
   （b）$\hat{p}_n$ の平均と分散を求めなさい（要証明）．
   （c）$\hat{p}_n$ の漸近分布を求めなさい．
   （d）p の 95% 信頼区間を近似的に求めなさい．
   （e）$n = 100$，$\hat{p}_n = .2$ として p の 95% 信頼区間を近似的に計算しなさい．

解答例

1. 統計学の基本用語
   (a) 母集団
   (b) 統計量
   (c) 不偏推定量
   (d) 信頼域
      - 「信頼区間」は 1 点減.
2. 分布表の読み方
   (a)

$$\begin{aligned}\Pr[a \le X \le b] &= \Pr[X \ge a] - \Pr[X > b] \\ &= .99\end{aligned}$$

これを満たす例は

$$\begin{aligned}\Pr[X \ge a] &= .995 \\ \Pr[X > b] &= .005\end{aligned}$$

χ^2 分布表より $X \sim \chi^2(2)$ なら $a = .0100251$, $b = 10.5966$.
   - 各 5 点.

   (b) t 分布の対称性より

$$\begin{aligned}\Pr[|Y| \le c] &= \Pr[-c \le Y \le c] \\ &= 1 - 2\Pr[Y > c] \\ &= .95\end{aligned}$$

すなわち

$$\Pr[Y > c] = .025$$

t 分布表より $Y \sim \mathrm{t}(4)$ なら $c = 2.776$.

   (c)

$$\begin{aligned}\Pr[d \le Z \le e] &= 1 - \Pr[Z < d] - \Pr[Z > e] \\ &= .9\end{aligned}$$

これを満たす例は

$$\begin{aligned}\Pr[Z < d] &= \Pr\left[\frac{1}{Z} > \frac{1}{d}\right] \\ &= .05 \\ \Pr[Z > e] &= .05\end{aligned}$$

$Z \sim \mathrm{F}(6,8)$ なら $1/Z \sim \mathrm{F}(8,6)$ なので F 分布表より $1/d = 4.147$, すなわち $d = 1/4.147$. 同じく F 分布表より $Z \sim \mathrm{F}(6,8)$ なら $e = 3.581$.
   - 各 5 点.

3. 母比率の信頼区間

(a) $X \sim \mathrm{Bin}(1, p)$ とすると

$$\begin{aligned}
\mathrm{E}(X) &:= 1 \cdot p + 0 \cdot (1-p) \\
&= p \\
\mathrm{var}(X) &= \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 \\
&= \mathrm{E}(X) - \mathrm{E}(X)^2 \\
&= p - p^2 \\
&= p(1-p)
\end{aligned}$$

- 各 5 点.
- 要証明.

(b) 期待値の線形性より

$$\begin{aligned}
\mathrm{E}(\hat{p}_n) &= \mathrm{E}\left(\frac{X_1 + \cdots + X_n}{n}\right) \\
&= \frac{\mathrm{E}(X_1) + \cdots + \mathrm{E}(X_n)}{n} \\
&= \frac{np}{n} \\
&= p
\end{aligned}$$

$X_1, \ldots, X_n$ は独立なので

$$\begin{aligned}
\mathrm{var}(\hat{p}_n) &= \mathrm{var}\left(\frac{X_1 + \cdots + X_n}{n}\right) \\
&= \frac{\mathrm{var}(X_1) + \cdots + \mathrm{var}(X_n)}{n^2} \\
&= \frac{np(1-p)}{n^2} \\
&= \frac{p(1-p)}{n}
\end{aligned}$$

- 各 5 点.
- 要証明.

(c) $X_1, \ldots, X_n$ は iid なので，中心極限定理より

$$\hat{p}_n \overset{a}{\sim} \mathrm{N}\left(p, \frac{p(1-p)}{n}\right)$$

(d) 標準化すると

$$\frac{\hat{p}_n - p}{\sqrt{p(1-p)/n}} \overset{a}{\sim} \mathrm{N}(0, 1)$$

分母の p を $\hat{p}_n$ に置き換えても

$$\frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)/n}} \overset{a}{\sim} \mathrm{N}(0, 1)$$

標準正規分布表より

$$\Pr\left[-1.96 \leq \frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)/n}} \leq 1.96\right] \approx .95$$

ここで

$$-1.96 \leq \frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)/n}} \leq 1.96 \iff -1.96\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \leq \hat{p}_n - p \leq 1.96\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}$$

$$\iff \hat{p}_n - 1.96\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \leq p \leq \hat{p}_n + 1.96\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}$$

したがって p の 95 %信頼区間は近似的に

$$\left[\hat{p}_n - 1.96\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}, \hat{p}_n + 1.96\sqrt{\frac{\hat{p}_n(1-\hat{p})_n}{n}}\right]$$

（e）$n = 100$，$\hat{p}_n = .2$ を代入すると

$$\begin{aligned}\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} &= \sqrt{\frac{.2(1-.2)}{100}} \\ &= \frac{.4}{10} \\ &= \frac{4}{100}\end{aligned}$$

したがって p の 95 %信頼区間は近似的に

$$\left[.2 - \frac{1.96 \cdot 4}{100}, .2 + \frac{1.96 \cdot 4}{100}\right] \approx [.1216, .2784]$$