中級統計学：第 3 回中間試験

村澤 康友

2024 年 12 月 10 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) 母集団から標本を取り出すこと
   (b) 統計量の標本分布から母数について推測すること
   (c) $Z_1, \dots, Z_n \sim \mathrm{N}(0,1)$ が独立のとき，$Z_1^2 + \cdots + Z_n^2$ の分布
   (d) 母数に確率収束する推定量
2. （30 点）分布表を用いて以下の問いに答えなさい．
   (a) $X \sim \chi^2(20)$ とする．$\Pr[a \le X \le b] = 0.98$ となる a, b を求めなさい．
   (b) $Y \sim \mathrm{t}(24)$ とする．$\Pr[|Y| \le c] = 0.98$ となる c を求めなさい．
   (c) $Z \sim \mathrm{F}(12, 10)$ とする．$\Pr[d \le Z \le e] = 0.98$ 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）K 大生の（1 日平均）スマホ使用時間の分布を調べたい．母集団分布を $\mathrm{N}\left(\mu, \sigma^2\right)$ と仮定する（μ, σ^2 は未知）．無作為に選んだ K 大生 5 人に使用時間を尋ねたところ，1 時間・3 時間・4 時間・5 時間・7 時間という回答が得られた．
   (a) 標本平均 $\bar{X}$ と標本分散 s^2 を求めなさい．
   (b) $\bar{X}$ と s^2 はどのような分布をもつか？（証明不要）
   (c) $\bar{X}$ の分散の推定値を求めなさい．
   (d) μ の 95 %信頼区間を求めなさい．
   (e) σ^2 の 95 %信頼区間を求めなさい．

解答例

1. 統計学の基本用語
   (a) 標本抽出
   (b) 統計的推測
      - 「推定」のみは 1 点（統計的推測は「検定」も含む）.
      - 「点推定」「区間推定」は意味を限定しすぎなので 0 点.
   (c) 自由度 n の χ^2 分布
      - 「χ^2 分布」のみは 2 点（自由度なしだと分布が特定されない）.
      - 問題より自由度は n であり，k はダメ.
   (d) 一致推定量
      - 推定量の性質でなく種類の定義なので「一致性」は 1 点減.
2. 分布表の読み方
   (a)

$$\begin{aligned}\Pr[a \le X \le b] &= \Pr[X \ge a] - \Pr[X > b] \\ &= 0.98\end{aligned}$$

これを満たす例は

$$\begin{aligned}\Pr[X \ge a] &= 0.99 \\ \Pr[X > b] &= 0.01\end{aligned}$$

χ^2 分布表より $X \sim \chi^2(20)$ なら $a = 8.26040,\ b = 37.5662.$

- 各 5 点.

(b) t 分布の対称性より

$$\begin{aligned}\Pr[|Y| \le c] &= \Pr[-c \le Y \le c] \\ &= 1 - 2\Pr[Y > c] \\ &= 0.98\end{aligned}$$

すなわち

$$\Pr[Y > c] = 0.01$$

t 分布表より $Y \sim \mathrm{t}(24)$ なら $c = 2.492.$

(c)

$$\begin{aligned}\Pr[d \le Z \le e] &= 1 - \Pr[Z < d] - \Pr[Z > e] \\ &= 0.98\end{aligned}$$

これを満たす例は

$$\begin{aligned}&\Pr[Z < d] = \Pr\left[\frac{1}{Z} > \frac{1}{d}\right] = 0.01 \\ &\Pr[Z > e] = 0.01\end{aligned}$$

$Z \sim \mathrm{F}(12, 10)$ なら $1/Z \sim \mathrm{F}(10, 12)$ なので F 分布表より $1/d = 4.296,$ すなわち $d = 1/4.296.$ 同じく F 分布表より $Z \sim \mathrm{F}(12, 10)$ なら $e = 4.706.$

- 各 5 点.

3. 母平均・母分散の区間推定

(a) 標本平均は

$$\begin{aligned}\bar{X} &= \frac{1+3+4+5+7}{5} \\ &= 4\end{aligned}$$

標本分散は

$$\begin{aligned}s^2 &= \frac{(1-4)^2+(3-4)^2+(4-4)^2+(5-4)^2+(7-4)^2}{5-1} \\ &= \frac{9+1+0+1+9}{4} \\ &= 5\end{aligned}$$

- 各 5 点.

(b)

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{5}\right)$$

$$\frac{4s^2}{\sigma^2} \sim \chi^2(4)$$

- 各 5 点.
- $n = 5$ を代入しなければ 1 点減.
- 左辺の $4s^2/\sigma^2$ がなければダメ.

(c) $\bar{X}$ の分散の推定値は

$$\frac{s^2}{n} = \frac{5}{5} = 1$$

- $n = 5, s^2 = 5$ を代入しなければダメ.
- (a) の s^2 と整合的なら OK.

(d) $\bar{X}$ を標準化すると

$$\frac{\bar{X}-\mu}{\sqrt{\sigma^2/5}} \sim \mathrm{N}(0,1)$$

σ^2 を s^2 に置き換えると

$$\frac{\bar{X}-\mu}{\sqrt{s^2/5}} \sim \mathrm{t}(4)$$

t 分布表より

$$\Pr\left[-2.776 \le \frac{\bar{X}-\mu}{\sqrt{s^2/5}} \le 2.776\right] = 0.95$$

すなわち

$$\Pr\left[\bar{X} - 2.776\sqrt{\frac{s^2}{5}} \le \mu \le \bar{X} + 2.776\sqrt{\frac{s^2}{5}}\right] = 0.95$$

$\bar{X} = 4,\ s^2 = 5$ より μ の 95 %信頼区間は $[1.224, 6.776]$.

- 標準化で 2 点.

- t(4) までは 4 点.
- t 分布表の読み取りまでは 6 点.
- $\bar{X} = 4,\ s^2 = 5$ を代入しなければ 2 点減.

(e) $4s^2/\sigma^2 \sim \chi^2(4)$ なので χ^2 分布表より

$$\Pr\left[0.484419 \leq \frac{4s^2}{\sigma^2} \leq 11.1433\right] = 0.95$$

すなわち

$$\Pr\left[\frac{4s^2}{11.1433} \leq \sigma^2 \leq \frac{4s^2}{0.484419}\right] = 0.95$$

$s^2 = 5$ より σ^2 の 95 %信頼区間は $[20/11.1433, 20/0.484419] \approx [1.79, 41.29]$.

- χ^2 分布表の読み取りまでは 5 点.
- $s^2 = 5$ を代入しなければ 2 点減.