

経済統計：第3回中間試験

村澤 康友

2016年6月27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（部分点は大きいに与えるが、結果のみの解答は0点とする）。

1. (20点) 以下の用語の定義を式または言葉で書きなさい（各20字程度）。
 - (a) 標本
 - (b) 標本分布
 - (c) 推定量
 - (d) 尤度
2. (50点) F大生の（試験前の）1日当たり勉強時間の母集団分布を $N(5, 4)$ とする。無作為に選んだF大生10人の勉強時間を (X_1, \dots, X_{10}) とする。
 - (a) 標本和 $X_1 + \dots + X_{10}$ の分布を求めなさい。
 - (b) 標本平均 \bar{X} の分布を求めなさい。
 - (c) (母平均を未知とした) 標本分散 s^2 の分布を示しなさい（証明不要）。
 - (d) $\Pr[|\bar{X} - 5| \leq c] = .95$ となる c を求めなさい。
 - (e) $\Pr[a < s^2 \leq b] = .95$ となる a, b を求めなさい。
3. (30点) $\text{Bin}(1, p)$ から抽出した大きさ n の無作為標本を $X := (X_1, \dots, X_n)$ とする。
 - (a) $X = x$ の確率を求めなさい。
 - (b) $X = x$ を観測したときの p の対数尤度関数を書きなさい。
 - (c) p のML推定量を求めなさい（導出の過程を示すこと）。

解答例

1. 統計学の基本用語

(a) 母集団のうち実際に観察される部分.

- 「母集団の一部」がなければ 0 点.

(b) 確率的な標本抽出にともなう統計量の分布.

- 「統計量」がなければ 0 点.

(c) 推定に用いる統計量.

- 「統計量」がなければ 0 点.

(d) ある母数の下で標本の実現値を観測する確率 (密度).

- 「(母数の) 尤もらしさ」は定義でないので 0 点.

2. 標本平均・標本分散の標本分布

(a)

$$\begin{aligned} E(X_1 + \cdots + X_{10}) &= E(X_1) + \cdots + E(X_{10}) \\ &= 10 \cdot 5 \\ &= 50 \\ \text{var}(X_1 + \cdots + X_{10}) &= \text{var}(X_1) + \cdots + \text{var}(X_{10}) \\ &= 10 \cdot 4 \\ &= 40 \end{aligned}$$

したがって $X_1 + \cdots + X_{10} \sim N(50, 40)$.

- 平均・分散のみは 5 点.
- $N(n\mu, n\sigma^2)$ は 5 点.

(b)

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{X_1 + \cdots + X_{10}}{10}\right) \\ &= \frac{E(X_1 + \cdots + X_{10})}{10} \\ &= \frac{50}{10} \\ &= 5 \\ \text{var}(\bar{X}) &= \text{var}\left(\frac{X_1 + \cdots + X_{10}}{10}\right) \\ &= \frac{\text{var}(X_1 + \cdots + X_{10})}{100} \\ &= \frac{40}{100} \\ &= \frac{2}{5} \end{aligned}$$

したがって $\bar{X} \sim N(5, 2/5)$.

- 平均・分散のみは 5 点.
- $N(\mu, \sigma^2/n)$ は 5 点.

(c) 一般に $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$ より

$$\frac{9s^2}{4} \sim \chi^2(9)$$

- $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$ で 5 点.

(d)

$$\begin{aligned} \Pr[|\bar{X} - 5| \leq c] &= \Pr[-c \leq \bar{X} - 5 \leq c] \\ &= \Pr\left[-\frac{c}{\sqrt{2/5}} \leq \frac{\bar{X} - 5}{\sqrt{2/5}} \leq \frac{c}{\sqrt{2/5}}\right] \\ &= .95 \end{aligned}$$

したがって

$$Q\left(\frac{c}{\sqrt{2/5}}\right) = .025$$

標準正規分布表より

$$\frac{c}{\sqrt{2/5}} \approx 1.96$$

したがって

$$\begin{aligned} c &\approx 1.96\sqrt{\frac{2}{5}} \\ &\approx 1.24 \end{aligned}$$

- 標準化で 5 点.

(e) $(9/4)s^2 \sim \chi^2(9)$ なので χ^2 分布表より

$$\Pr\left[2.70039 < \frac{9s^2}{4} \leq 19.0228\right] = .95$$

または

$$\Pr\left[2.70039 \cdot \frac{4}{9} < s^2 \leq 19.0228 \cdot \frac{4}{9}\right] = .95$$

したがって $(a, b) = (1.200, 8.455)$.

- 分布表の読み取りで 5 点.
- 95% 信頼区間を真ん中でなく片側にするのは 1 点減 (真ん中にするより広くなる).

3. 母比率の ML 推定 (復習テスト 18 の問 1 と同じ問題)

(a)

$$\begin{aligned} \Pr[X = x] &= \Pr[(X_1, \dots, X_n) = (x_1, \dots, x_n)] \\ &= \Pr[X_1 = x_1, \dots, X_n = x_n] \\ &= \Pr[X_1 = x_1] \cdots \Pr[X_n = x_n] \\ &= p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \end{aligned}$$

- $X \sim \text{Bin}(n, p)$ は 0 点 ($X \neq X_1 + \cdots + X_n$).

(b)

$$\ell(p; x_1, \dots, x_n) = \sum_{i=1}^n x_i \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1-p)$$

(c) 最大化の1階の条件は

$$\frac{\sum_{i=1}^n x_i}{p^*} - \frac{n - \sum_{i=1}^n x_i}{1 - p^*} = 0$$

すなわち

$$(1 - p^*) \sum_{i=1}^n x_i - p^* \left(n - \sum_{i=1}^n x_i \right) = 0$$

したがって

$$p^* = \frac{\sum_{i=1}^n x_i}{n}$$