経済統計：第3回中間試験

村澤 康友

2018年7月9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）．

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）統計量
   （b）標本分散
   （c）尤度関数
   （d）漸近（大標本）特性
2. （30点）分布表を用いて以下の問いに答えなさい．
   （a）$X \sim \chi^2(1)$ とする．$\Pr[a \leq X \leq b] = .98$ となる a, b を求めなさい．
   （b）$Y \sim \mathrm{t}(2)$ とする．$\Pr[Y \leq c] = .1$ となる c を求めなさい．
   （c）$Z \sim \mathrm{F}(3, 4)$ とする．$\Pr[d \leq Z \leq e] = .95$ となる d, e を求めなさい．
   なお a〜e はすべて実数とする（$\pm\infty$ は除く）．
3. （50点）$\mathrm{N}\left(\mu_X, \sigma^2\right)$，$\mathrm{N}\left(\mu_Y, \sigma^2\right)$ から独立に抽出した無作為標本を $(X_1, \dots, X_m)$，$(Y_1, \dots, Y_n)$ とする．
   （a）標本平均 $\bar{X}, \bar{Y}$ の分布をそれぞれ求めなさい．
   （b）$\bar{X} - \bar{Y}$ の分布を求めなさい．
   （c）プールした標本分散 s^2 を式で定義しなさい．
   （d）s^2 はどのような分布をもつか？
   （e）$m = n = 8$ とする．$\bar{X}, \bar{Y}, s^2$ を用いて $\mu_X - \mu_Y$ の95％信頼区間を作りなさい．

解答例

1. 統計学の基本用語
   (a) 標本の関数.
   (b) $(X_1,\dots,X_n)$ の標本分散は
   $$s^2 := \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$$
   (c) 標本の pmf・pdf を母数の尤度を表す関数とみたもの.
      - 「標本の pmf・pdf」のみは 2 点.
   (d) 推定量（統計量）の漸近分布に関する性質.
2. 分布表の読み方
   (a)
   $$\begin{aligned}\Pr[a \le X \le b] &= \Pr[X \ge a] - \Pr[X > b]\\ &= .98\end{aligned}$$
   これを満たす例は
   $$\begin{aligned}\Pr[X \ge a] &= .99\\ \Pr[X > b] &= .01\end{aligned}$$
   χ^2 分布表より $X \sim \chi^2(1)$ なら $a = .000157088$，$b = 6.63490$.
      - 各 5 点.
   (b) t 分布の対称性より
   $$\begin{aligned}\Pr[Y \le c] &= \Pr[Y \ge -c]\\ &= .1\end{aligned}$$
   t 分布表より $Y \sim \mathrm{t}(2)$ なら $-c = 1.886$，すなわち $c = -1.886$.
      - 符号のミスは 5 点.
   (c)
   $$\begin{aligned}\Pr[d \le Z \le e] &= 1 - \Pr[Z < d] - \Pr[Z > e]\\ &= .95\end{aligned}$$
   これを満たす例は
   $$\begin{aligned}\Pr[Z < d] &= \Pr\left[\frac{1}{Z} > \frac{1}{d}\right]\\ &= .025\\ \Pr[Z > e] &= .025\end{aligned}$$
   $Z \sim \mathrm{F}(3,4)$ なら $1/Z \sim \mathrm{F}(4,3)$ なので F 分布表より $1/d = 15.101$，すなわち $d = 1/15.101$．同じく F 分布表より $Z \sim \mathrm{F}(3,4)$ なら $e = 9.979$.
      - 各 5 点.
      - $d = 0$，$e = 6.591$ も可.

3. 母平均の差の区間推定

(a)

$$\bar{X} \sim \mathrm{N}\left(\mu_X, \frac{\sigma^2}{m}\right)$$
$$\bar{Y} \sim \mathrm{N}\left(\mu_Y, \frac{\sigma^2}{n}\right)$$

- 各 5 点.
- 平均で 1 点，分散で 2 点，正規分布で 2 点.

(b)

$$\bar{X} - \bar{Y} \sim \mathrm{N}\left(\mu_X - \mu_Y, \frac{\sigma^2}{m} + \frac{\sigma^2}{n}\right)$$

- 平均で 2 点，分散で 4 点，正規分布で 4 点.

(c)

$$s^2 := \frac{1}{m+n-2}\left[\sum_{i=1}^{m}\left(X_i - \bar{X}\right)^2 + \sum_{j=1}^{n}\left(Y_j - \bar{Y}\right)^2\right]$$

(d)

$$\frac{(m+n-2)s^2}{\sigma^2} \sim \chi^2(m+n-2)$$

(e)

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{s^2(1/m + 1/n)}} \sim \mathrm{t}(m+n-2)$$

$m = n = 8$ より

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{s/2} \sim \mathrm{t}(14)$$

t 分布表より

$$\Pr\left[-2.145 \leq \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{s/2} \leq 2.145\right] = .95$$

すなわち

$$\Pr\left[-1.0725s \leq \bar{X} - \bar{Y} - (\mu_X - \mu_Y) \leq 1.0725s\right] = .95$$

または

$$\Pr\left[\bar{X} - \bar{Y} - 1.0725s \leq \mu_X - \mu_Y \leq \bar{X} - \bar{Y} + 1.0725s\right] = .95$$

したがって $\mu_X - \mu_Y$ の 95 %信頼区間は

$$\left[\bar{X} - \bar{Y} - 1.0725s, \bar{X} - \bar{Y} + 1.0725s\right]$$

- t(14) で 5 点.