経済統計 II：中間試験

村澤 康友

2020 年 11 月 30 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) 単純無作為抽出
   (b) 標本比率
   (c) 有限標本（小標本）特性
   (d) 信頼域
2. （50 点）F 大生の男子と女子の（試験前の）1 日当たり勉強時間の母集団分布を，それぞれ $\mathrm{N}(4,3)$，$\mathrm{N}(5,3)$ とする．無作為に選んだ F 大生の男子 16 人と女子 8 人の勉強時間を，それぞれ $(X_1,\dots,X_{16})$，$(Y_1,\dots,Y_8)$ とする．
   (a) 標本平均 $\bar{X},\bar{Y}$ と，その差 $\bar{X}-\bar{Y}$ の分布を求めなさい．
   (b) $\bar{X}>\bar{Y}$ となる確率を求めなさい．
   (c) プールした標本分散 s^2 を定義し，その分布を求めなさい．
   (d) 標本分散 s_X^2,s_Y^2 と，その比 s_X^2/s_Y^2 の分布を求めなさい．
   (e) $\Pr\left[a\le s_X^2/s_Y^2\le b\right]=.95$ となる a,b を求めなさい．ただし $0<a<b<\infty$ とする．
3. （30 点）F 大生の通学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 F 大生 5 人に通学時間を尋ねたところ，20 分・40 分・50 分・60 分・80 分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) μ の 90 %信頼区間を求めなさい．
   (c) σ^2 の 90 %信頼区間を求めなさい．

解答例

1. 統計学の基本用語
   (a) どの個体の組合せも等確率で取り出される抽出.
      - 「どの個体も等確率で取り出される抽出」は不十分なので 2 点.
   (b) ベルヌーイ母集団からの標本における成功（$=1$）の比率.
   (c) 推定量の厳密な分布に関する性質.
      - 漸近特性と対比されるので「厳密な」が必要.
   (d) ある確率で母数を含む確率的な領域.
2. 統計量の標本分布
   (a)

$$\begin{aligned}\bar{X} &\sim \mathrm{N}\left(4, \frac{3}{16}\right)\\ \bar{Y} &\sim \mathrm{N}\left(5, \frac{3}{8}\right)\\ \bar{X}-\bar{Y} &\sim \mathrm{N}\left(-1, \frac{9}{16}\right)\end{aligned}$$

   - 標本平均の分布で各 3 点，その差の分布で 4 点.

   (b) $Z \sim \mathrm{N}(0,1)$ とすると

$$\begin{aligned}\Pr\left[\bar{X}-\bar{Y}>0\right] &= \Pr\left[\frac{\bar{X}-\bar{Y}-(-1)}{3/4} > \frac{0-(-1)}{3/4}\right]\\ &= \Pr\left[Z > \frac{4}{3}\right]\\ &= \Pr[Z>1.33]\\ &= .091759\end{aligned}$$

   (c) プールした標本分散は

$$s^2 := \frac{1}{22}\left[\sum_{i=1}^{16}\left(X_i-\bar{X}\right)^2 + \sum_{j=1}^{8}\left(Y_j-\bar{Y}\right)^2\right]$$

   s^2 の分布は

$$\frac{22s^2}{3} \sim \chi^2(22)$$

   - 定義で 5 点，分布で 5 点.
   - m, n, σ^2 を問題から特定しなければ 0 点.

   (d)

$$\begin{aligned}\frac{15s_X^2}{3} &\sim \chi^2(15)\\ \frac{7s_Y^2}{3} &\sim \chi^2(7)\\ \frac{s_X^2}{s_Y^2} &\sim \mathrm{F}(15,7)\end{aligned}$$

- 標本分散の分布で各 3 点，その比の分布で 4 点.
- $m, n, \sigma_X^2, \sigma_Y^2$ を問題から特定しなければ 0 点.

(e) $F \sim \mathrm{F}(15,7)$ とすると

$$\Pr[a \le F \le b] = .95$$

これを満たす a, b の例は

$$\begin{aligned}\Pr[F < a] &= \Pr\left[\frac{1}{F} > \frac{1}{a}\right] \\ &= .025 \\ \Pr[F > b] &= .025\end{aligned}$$

$1/F \sim \mathrm{F}(7,15)$ より $1/a = 3.293$ すなわち $a = 1/3.293$．また $b = 4.568$.

3. 母平均・母分散の区間推定

(a) 標本平均は

$$\begin{aligned}\bar{X} &:= \frac{20+40+50+60+80}{5} \\ &= 50\end{aligned}$$

標本分散は

$$\begin{aligned}s^2 &:= \frac{(20-50)^2+(40-50)^2+(50-50)^2+(60-50)^2+(80-50)^2}{4} \\ &= \frac{900+100+0+100+900}{4} \\ &= 500\end{aligned}$$

(b) 標本平均の分布は

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

標準化すると

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

σ^2 を s^2 に置き換えると

$$\frac{\bar{X}-\mu}{\sqrt{s^2/n}} \sim \mathrm{t}(n-1)$$

t 分布表より $n = 5$ なら

$$\Pr\left[-2.132 \le \frac{\bar{X}-\mu}{\sqrt{s^2/n}} \le 2.132\right] = .9$$

または

$$\Pr\left[-2.132\sqrt{\frac{s^2}{n}} \le \bar{X}-\mu \le 2.132\sqrt{\frac{s^2}{n}}\right] = .9$$

または

$$\Pr\left[\bar{X}-2.132\sqrt{\frac{s^2}{n}} \le \mu \le \bar{X}+2.132\sqrt{\frac{s^2}{n}}\right] = .9$$

$n = 5$，$\bar{X} = 50$，$s^2 = 500$ を代入すると，90 %信頼区間は $[28.68, 71.32]$.

（c）標本分散の分布は

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

χ^2 分布表より $n=5$ なら

$$\Pr\left[.710723 \leq \frac{(n-1)s^2}{\sigma^2} \leq 9.48773\right] = .9$$

または

$$\Pr\left[\frac{1}{9.48773} \leq \frac{\sigma^2}{(n-1)s^2} \leq \frac{1}{.710723}\right] = .9$$

または

$$\Pr\left[\frac{(n-1)s^2}{9.48773} \leq \sigma^2 \leq \frac{(n-1)s^2}{.710723}\right] = .9$$

$n=5$，$s^2=500$ を代入すると，90 %信頼区間は $[42.16, 562.81]$.