計量経済 II：中間試験

村澤 康友

2016 年 11 月 15 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) 統計的推測
   (b) 最尤法
   (c) （k 次の）標本積率
   (d) 漸近有効推定量
2. （30 点）$\mathrm{N}\left(\mu,\sigma^2\right)$ から抽出した大きさ n の無作為標本の標本分散を s^2 とする．μ は未知とする．
   (a) $\mathrm{E}\left(s^2\right)=\sigma^2$ となることを示しなさい．
   (b) s^2 はどのような分布をもつか？（証明不要）
   (c) $\sigma^2=1$ とする．$n=17$ のとき $s^2>2$ の確率を χ^2 分布表を利用して求めなさい．
3. （50 点）ある将棋の対局の中終盤 50 手の将棋ソフトとの一致率は X 棋士が 88%（22/25），Y 棋士が 60%（15/25）であった．両棋士の（無限仮説母集団における）真の一致率を p_X,p_Y，この対局の中終盤各 25 手における一致率（＝標本平均）を $\hat{p}_X,\hat{p}_Y$ とする．各対局者と将棋ソフトの指し手の一致／不一致は独立かつ同一なベルヌーイ試行であり，また互いの指し手についても独立と仮定する．
   (a) 2 項母集団 $\mathrm{Bin}(1,p_X)$, $\mathrm{Bin}(1,p_Y)$ の平均と分散を求めなさい．
   (b) $\hat{p}_X$, $\hat{p}_Y$ の漸近分布を求めなさい．
   (c) $\hat{p}_X-\hat{p}_Y$ の漸近分布を求めなさい．
   (d) p_X-p_Y の 99% 信頼区間を近似的に求めなさい．
   (e) $n_X=n_Y=25$ として 99% 信頼区間を近似的に計算し，それが $p_X-p_Y=0$ を含むかどうか調べなさい．

   ※数値例はフィクションです．また分析の目的は両棋士の棋風の差異の検証であり，本データのみで不正行為の有無の検証は不可能です．

解答例

1. 統計学の基本用語
   (a) 一部の観察から全体について推測すること.
   (b) (対数) 尤度関数を最大にする解を母数の推定値とする手法.
   (c) $\hat{\mu}_k := (1/n)\sum_{i=1}^{n} X_i^k$.
   (d) 漸近正規推定量の中で漸近分散が最小となる推定量.
      - 「漸近分散が最小」のみは 2 点.
2. 標本分散の標本分布
   (a) 標本分散は
   $$s^2 := \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$
   次式を示せばよい.
   $$\mathrm{E}\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right) = (n-1)\sigma^2$$
   ここで
   $$\begin{aligned}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 &= \sum_{i=1}^{n}\left[(X_i-\mu) - \left(\bar{X}-\mu\right)\right]^2 \\ &= \sum_{i=1}^{n}\left[(X_i-\mu)^2 - 2(X_i-\mu)\left(\bar{X}-\mu\right) + \left(\bar{X}-\mu\right)^2\right] \\ &= \sum_{i=1}^{n}(X_i-\mu)^2 - 2\sum_{i=1}^{n}(X_i-\mu)\left(\bar{X}-\mu\right) + n\left(\bar{X}-\mu\right)^2 \\ &= \sum_{i=1}^{n}(X_i-\mu)^2 - n\left(\bar{X}-\mu\right)^2\end{aligned}$$
   期待値をとると
   $$\begin{aligned}\mathrm{E}\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right) &= \sum_{i=1}^{n}\mathrm{E}\left((X_i-\mu)^2\right) - n\,\mathrm{E}\left(\left(\bar{X}-\mu\right)^2\right) \\ &= \sum_{i=1}^{n}\mathrm{var}(X_i) - n\,\mathrm{var}\left(\bar{X}\right) \\ &= n\sigma^2 - n\frac{\sigma^2}{n} \\ &= (n-1)\sigma^2\end{aligned}$$
      - 標本分散の定義で 5 点.

   (b)
   $$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$
      - $\chi^2(n-1)$ のみは $s^2 \sim \chi^2(n-1)$ の意味に取れるので 0 点.

(c)

$$\begin{aligned}\Pr\left[s^2 > 2\right] &= \Pr\left[\frac{(n-1)s^2}{\sigma^2} > \frac{2(n-1)}{\sigma^2}\right] \\ &= \Pr\left[\chi^2(16) > 32\right] \\ &\approx .01\end{aligned}$$

- $\chi^2(16)$ で 5 点.

3. 母比率の差の信頼区間

(a) $\mathrm{Bin}(1, p_X)$ の平均は

$$1 \cdot p_X + 0 \cdot (1 - p_X) = p_X$$

分散は

$$\begin{aligned}(1-p_X)^2 \cdot p_X + (0-p_X)^2 \cdot (1-p_X) &= (1-p_X)^2 p_X + p_X^2(1-p_X) \\ &= p_X(1-p_X)\end{aligned}$$

$\mathrm{Bin}(1, p_Y)$ についても同様.

- 平均で 5 点，分散で 5 点.

(b)

$$\hat{p}_X \overset{a}{\sim} \mathrm{N}\left(p_X, \frac{p_X(1-p_X)}{n_X}\right)$$
$$\hat{p}_Y \overset{a}{\sim} \mathrm{N}\left(p_Y, \frac{p_Y(1-p_Y)}{n_Y}\right)$$

- 前問の解答と整合的なら OK.

(c)

$$\hat{p}_X - \hat{p}_Y \overset{a}{\sim} \mathrm{N}\left(p_X - p_Y, \frac{p_X(1-p_X)}{n_X} + \frac{p_Y(1-p_Y)}{n_Y}\right)$$

- 前問の解答と整合的なら OK.

(d)

$$\frac{\hat{p}_X - \hat{p}_Y - (p_X - p_Y)}{\sqrt{p_X(1-p_X)/n_X + p_Y(1-p_Y)/n_Y}} \overset{a}{\sim} \mathrm{N}(0,1)$$

または

$$\frac{\hat{p}_X - \hat{p}_Y - (p_X - p_Y)}{\sqrt{\hat{p}_X(1-\hat{p}_X)/n_X + \hat{p}_Y(1-\hat{p}_Y)/n_Y}} \overset{a}{\sim} \mathrm{N}(0,1)$$

したがって

$$\Pr\left[-2.58 \le \frac{\hat{p}_X - \hat{p}_Y - (p_X - p_Y)}{\sqrt{\hat{p}_X(1-\hat{p}_X)/n_X + \hat{p}_Y(1-\hat{p}_Y)/n_Y}} \le 2.58\right] \approx .99$$

99% 信頼区間の上限・下限は

$$\hat{p}_X - \hat{p}_Y \pm 2.58\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}}$$

- 前問の解答と整合的なら OK.

(e)

$$\begin{aligned}\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} &= \frac{.88(1-.88)}{25} \\ &= \frac{.1056}{25} \\ \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y} &= \frac{.6(1-.4)}{25} \\ &= \frac{.24}{25}\end{aligned}$$

したがって

$$\frac{\hat{p}_X(1-\hat{p}_X)}{n_X}+\frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}=\frac{.3456}{25}$$

すなわち

$$\begin{aligned}\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X}+\frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}} &\approx \frac{.5879}{5} \\ &= .1176\end{aligned}$$

99% 信頼区間は

$$[.28-2.58\cdot .1176, .28+2.58\cdot .1176]\approx[-.023, .583]$$

したがって 0 を含む.