経済統計：前期第 3 回中間試験

村澤 康友

2011 年 7 月 11 日

注意：3 問とも解答すること．

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) 母数（パラメーター）
   (b) 不偏推定量
   (c) 漸近正規推定量
   (d) 信頼係数
2. （30 点）サッカーの試合の得（失）点数はポアソン分布に従うと考えられる．ポアソン分布の確率関数は $x = 0, 1, \ldots$ について
$$p(x) := \frac{\mathrm{e}^{-\lambda}\lambda^x}{x!}.$$
   n 試合分の得点データ $(X_1, \ldots, X_n)$ から母数 λ を推定したい．
   (a) $X_1, \ldots, X_n$ を独立と仮定して $(X_1, \ldots, X_n)$ の同時確率関数を書きなさい．
   (b) $(X_1, \ldots, X_n) = (x_1, \ldots, x_n)$ を観測したときの λ の対数尤度関数を書きなさい．
   (c) λ の最尤推定量が標本平均 $\bar{X}$ となることを示しなさい．
3. （50 点）真夏の電力需要は日中最高気温に左右される．したがって「でんき予報」のためには日中最高気温の確率分布を知る必要がある．例として 8 月 15 日の大阪の日中最高気温を考える．過去 5 年間のデータは次表の通り．

年	日中最高気温
2006	37.9
2007	36.9
2008	35.7
2009	31.5
2010	34.5

   母集団分布は $\mathrm{N}\left(\mu, \sigma^2\right)$ と仮定する．（※解答に際しては分布表を適宜参照すること．）
   (a) 標本平均 $\bar{X}$ と標本分散 s^2 の値を計算しなさい．
   (b) 日中最高気温が 35 度以上の日を「猛暑日」という．$\bar{X}, s^2$ を母平均・母分散とみなして 8 月 15 日に大阪が「猛暑日」となる確率を求めなさい．
   (c) $\bar{X}, s^2$ の確率分布を書きなさい（既知の数値は代入すること）．
   (d) μ の 95 %信頼区間を求めなさい（計算が面倒なので $\bar{X}, s^2$ の数値は代入しなくてよい）．
   (e) σ^2 の 95 %信頼区間を求めなさい（同上）．

解答例

1. 統計学の基本用語
   (a) 確率分布の特性を表す定数.
   (b) 期待値が母数と等しい推定量.
   (c) 漸近分布が正規分布である推定量.
   (d) 信頼域が母数を含む確率.
2. 最尤推定
   (a) $(X_1, \dots, X_n)$ の同時確率関数は，任意の $(x_1, \dots, x_n) \in \mathbb{N}^n$ について

$$\begin{aligned} p_{X_1,\dots,X_n}(x_1, \dots, x_n) &= p(x_1) \cdots p(x_n) \\ &= \frac{\mathrm{e}^{-\lambda}\lambda^{x_1}}{x_1!} \cdots \frac{\mathrm{e}^{-\lambda}\lambda^{x_n}}{x_n!} \\ &= \frac{\mathrm{e}^{-n\lambda}\lambda^{x_1+\cdots+x_n}}{x_1! \cdots x_n!}. \end{aligned}$$

   (b) 対数尤度関数は

$$\ln L(\lambda; x_1, \dots, x_n) = -n\lambda + (x_1 + \cdots + x_n) \ln \lambda - \ln(x_1! \cdots x_n!).$$

   (c) 1 階の条件は

$$-n + \frac{x_1 + \cdots + x_n}{\lambda^*} = 0.$$

   したがって

$$\lambda^* = \frac{x_1 + \cdots + x_n}{n}.$$

3. 点推定と区間推定
   (a) 標本平均は

$$\begin{aligned} \bar{X} &= \frac{37.9 + 36.9 + 35.7 + 31.5 + 34.5}{5} \\ &= 35.3. \end{aligned}$$

   標本分散は

$$\begin{aligned} s^2 &= \frac{(37.9 - 35.3)^2 + (36.9 - 35.3)^2 + (35.7 - 35.3)^2 + (31.5 - 35.3)^2 + (34.5 - 35.3)^2}{4} \\ &= \frac{2.6^2 + 1.6^2 + 0.4^2 + (-3.8)^2 + (-0.8)^2}{4} \\ &= \frac{6.76 + 2.56 + 0.16 + 14.44 + 0.64}{4} \\ &= 6.14. \end{aligned}$$

   - 各 5 点.
   - 計算ミスは 1 点減.

   (b) 8 月 15 日の大阪の日中最高気温を X とすると

$$X \sim \mathrm{N}(35.3, 6.14).$$

したがって

$$\begin{aligned}\Pr[X \geq 35.0] &= \Pr\left[\frac{X-35.3}{\sqrt{6.14}} \geq \frac{35.0-35.3}{\sqrt{6.14}}\right] \\ &\approx \Pr\left[\frac{X-35.3}{\sqrt{6.14}} \geq -0.12\right] \\ &= 1-\Pr\left[\frac{X-35.3}{\sqrt{6.14}} \geq 0.12\right] \\ &= 1-0.45224 \\ &= 0.54776.\end{aligned}$$

(c) 標本の大きさ $n=5$ なので

$$\begin{aligned}\bar{X} &\sim \mathrm{N}\left(\mu, \frac{\sigma^2}{5}\right), \\ \frac{4s^2}{\sigma^2} &\sim \chi^2(4).\end{aligned}$$

- 各 5 点.
- 左辺の形を示さず $\chi^2(4)$ のみは 2 点.
- $\bar{X}, s^2$ の値を代入したらダメ（確率変数でなくなる）.

(d) 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

$$\Pr\left[-2.776 \leq \frac{\bar{X}-\mu}{\sqrt{s^2/5}} \leq 2.776\right] = .95,$$

すなわち

$$\Pr\left[-2.776\sqrt{\frac{s^2}{5}} \leq \bar{X}-\mu \leq 2.776\sqrt{\frac{s^2}{5}}\right] = .95,$$

または

$$\Pr\left[\bar{X}-2.776\sqrt{\frac{s^2}{5}} \leq \mu \leq \bar{X}+2.776\sqrt{\frac{s^2}{5}}\right] = .95.$$

したがって μ の 95 %信頼区間は

$$\left[\bar{X}-2.776\sqrt{\frac{s^2}{5}}, \bar{X}+2.776\sqrt{\frac{s^2}{5}}\right] \approx [32.2, 38.4].$$

- t 分布を使わなければ 0 点.
- 自由度の間違いは 0 点.

(e) χ^2 分布表より

$$\Pr\left[0.484419 \le \frac{4s^2}{\sigma^2} \le 11.1433\right] = .95,$$

すなわち

$$\Pr\left[\frac{1}{11.1433} \le \frac{\sigma^2}{4s^2} \le \frac{1}{0.484419}\right] = .95,$$

または

$$\Pr\left[\frac{4s^2}{11.1433} \le \sigma^2 \le \frac{4s^2}{0.484419}\right] = .95.$$

したがって σ^2 の 95 %信頼区間は

$$\left[\frac{4s^2}{11.1433}, \frac{4s^2}{0.484419}\right] \approx [2.20, 50.70].$$

- χ^2 分布表の読み間違いは 5 点減.
- 自由度の間違いは 0 点.

※答案は返却します．採点や成績に関する質問にも応じます．**オフィスアワーの時間（月水木金の昼休み）に**研究室まで来てください.