確率及び統計／レポート７

学籍番号： ０９５７０７Ｂ
氏名： 大 城 佳 明
提出日： 2010 年 7 月 30 日

1 問題

2 つの正規母集団からそれぞれ次の標本を得た。

標本 A： 1.5 2.1 2.5 1.8 3.0 4.1 1.8 2.9 1.6 2.3

標本 B： 2.3 1.8 1.9 2.4 2.0 2.9 3.2 2.7

2 つの母平均の違いについて、検定と情報量規準の双方の視点から考察せよ。

2 解答

2.1 検定

始めに、母平均の差を検定する。帰無仮説 $H_0 : \mu x = \mu y$ を、有意水準$\alpha = 0.05$ で検定する。$\sigma x, \sigma y$ は未知であり、$\sigma^2 x = \sigma^2 y$ とする。標本 A の標本平均、標本 B の標本平均をそれぞれ$\overline{X}, \overline{Y}$とすると、

$$\begin{aligned} \overline{X} &= \frac{1}{10}(1.5 + 2.1 + 2.5 + 1.8 + 3.0 + 4.1 + 1.8 + 2.9 + 1.6 + 2.3) \\ &= 2.36 \qquad (1) \\ \overline{Y} &= \frac{1}{8}(2.3 + 1.8 + 1.9 + 2.4 + 2.0 + 2.9 + 3.2 + 2.7) \\ &= 2.4 \qquad (2) \end{aligned}$$

となる。$S^2_{\overline{X}-\overline{Y}}$は$\overline{X} - \overline{Y}$ の標本不偏分散であり、

$$\begin{aligned} S^2_{\overline{X}-\overline{Y}} &= \frac{1}{N+M-2}\left\{\sum_{i=1}^{N}(X^{(i)} - \overline{X})^2 + \sum_{i=1}^{M}(Y^{(i)} - \overline{Y})^2\right\} \\ &= \frac{1}{10+8-2}(5.764 + 1.76) \\ &= 0.470 \qquad (3) \end{aligned}$$

となる。これらの母分散が未知な場合の検定量である次式に (1),(2),(3) を代入する

$$\begin{aligned} K &= \frac{\overline{X} - \overline{Y} - (\mu x - \mu y)}{\sqrt{(\frac{1}{N} - \frac{1}{M})S^2_{\overline{X}-\overline{Y}}}} \\ &= \frac{2.36 - 2.4}{\sqrt{(\frac{1}{10} - \frac{1}{8}) \cdot 0.470}} \end{aligned}$$

$$
\begin{aligned}
&= \frac{-0.04}{0.3252} \\
&= -0.123 \qquad (4)
\end{aligned}
$$

自由度１６の t 分布における$\alpha = 0.05$ の境界値は 2.120 である。したがって、棄却されない

2.2　情報量規準

　次に情報量規準を行う
$Model1.A\sim N(\mu,\sigma^2)\quad B\sim N(\mu,\sigma^2)$
$Model2.A\sim N(\mu,\sigma^2)\quad B\sim N(\mu,\sigma^2)\quad \mu_A \neq \mu_B$
$Model3.A\sim N(\mu,\sigma^2)\quad B\sim N(\mu,\sigma^2)\quad \sigma_A \neq \sigma_B$
$Model4.A\sim N(\mu,\sigma^2)\quad B\sim N(\mu,\sigma^2)\quad \mu_A \neq \mu_B \quad \sigma_A \neq \sigma_B$
にわけることが出来る。

2.2.1　Model 1

$$
\begin{aligned}
\overline{\mu} &= \frac{1}{18} \cdot 42.8 \\
&= 2.378 \qquad (5) \\
\overline{\sigma^2} &= \frac{1}{18} \cdot 7.531 \\
&= 0.418 \qquad (6)
\end{aligned}
$$

より、対数尤度関数は、

$$
l(\theta) = -\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x^{(i)} - \mu)^2 - \frac{N}{2}\log 2\pi\sigma^2 \qquad (7)
$$

であるので、(5),(6) を代入する。

$$
\begin{aligned}
l(\theta_1) &= -\frac{1}{2\cdot 0.418}\cdot 7.531 - \frac{18}{2}\log 2\pi \cdot 0.418 \\
&= -9 - 8.707 \\
&= -17.707
\end{aligned}
$$

$$
AIC(\theta) = -2\{l(\hat{\theta}) - k\} \qquad (8)
$$

より、母数の数は２なので、

$$
AIC(\theta_1) = 39.414 \qquad (9)
$$

2.2.2 Model 2

$$
\begin{aligned}
\overline{\mu_A} &= \frac{1}{10} \cdot 23.599998 \\
&= 2.36 && (10) \\
\overline{\mu_B} &= \frac{1}{8} \cdot 19.2 \\
&= 2.4 && (11) \\
\overline{\sigma^2} &= \frac{1}{18} \cdot (1.76 + 5.768) \\
&= 0.418 && (12)
\end{aligned}
$$

より、(7) に代入すると

$$
\begin{aligned}
l(\theta_2) &= -\frac{1}{2 \cdot 0.418} \cdot (5.764 + 1.76) - \frac{18}{2} \log 2\pi \cdot 0.418 \\
&= -9 - 8.707 \\
&= -17.686
\end{aligned}
$$

よって、母数の数は 3 なので

$$
AIC(\theta_2) = 41.388 \tag{13}
$$

2.2.3 Model 3

$$
\begin{aligned}
\overline{\mu} &= \frac{\frac{\sum_{i=1}^{N_A} x_i}{\sigma_A^2} + \frac{\sum_{i=1}^{N_B} x_j}{\sigma_B^2}}{\frac{N_A}{\sigma_A^2} + \frac{N_B}{\sigma_B^2}} \\
\overline{\sigma_A^2} &= \frac{1}{N_A} \sum_{i=1}^{N_A} (x_i - \overline{\mu})^2 = \frac{1}{N_A} \sum_{i=1}^{N_A} x_i^2 - \overline{\mu}^2 \\
\overline{\sigma_B^2} &= \frac{1}{N_B} \sum_{j=1}^{N_B} (x_i - \overline{\mu})^2 = \frac{1}{N_B} \sum_{j=1}^{N_B} x_i^2 - \overline{\mu}^2
\end{aligned}
$$

$$
\begin{aligned}
\overline{\mu} &= 2.382 && (14) \\
\overline{\sigma_A^2} &= 0.472 && (15) \\
\overline{\sigma_B^2} &= 0.306 && (16)
\end{aligned}
$$

より、(7) に代入すると

$$
\begin{aligned}
l(\theta_3) &= \left\{ -5 - \frac{10}{2} \log 2\pi \cdot 0.472 \right\} - \left\{ -4 - \frac{8}{2} \log 2\pi \cdot 0.306 \right\} \\
&= (-5 - 5.4329) + (-4 - 2.6127) \\
&= 17.05 && (17)
\end{aligned}
$$

よって、母数の数は 3 なので、

$$
AIC(\theta_3) = 40.10 \tag{18}
$$

2.2.4 Model 4

$$\begin{aligned}
\overline{\mu_A} &= \frac{1}{10} \cdot 23.599998 \\
&= 2.36 && (19) \\
\overline{\mu_B} &= \frac{1}{8} \cdot 19.2 \\
&= 2.4 && (20) \\
\overline{\sigma_A^2} &= \frac{1}{10} \cdot 5.764 \\
&= 0.576 && (21) \\
\overline{\sigma_B^2} &= \frac{1}{8} \cdot 1.76 \\
&= 0.220 && (22)
\end{aligned}$$

より、(7) に代入すると

$$\begin{aligned}
l(\theta_4) &= \left\{ -\frac{1}{2 \cdot 0.576} \cdot 5.77 - \frac{10}{2} \log 2\pi \cdot 0.576 \right\} - \left\{ \frac{1}{2 \cdot 0.220} \cdot 1.764 - \frac{8}{2} \log 2\pi \cdot 0.220 \right\} \\
&= (-5 - 6.438) + (-4 - 1.29) \\
&= -16.728 && (23)
\end{aligned}$$

よって、母数の数は 4 なので、

$$AIC(\theta_4) = 41.456 \tag{24}$$

2.2.5 結論

(9),(13),(18),(24) の結果より、

$$AIC(\theta_1) < AIC(\theta_3) < AIC(\theta_2) < AIC(\theta_4) \tag{25}$$

したがって、Model 1 が一番尤もらしい