今野　(2009 年 12 月 08 日)

情報統計学の問題 (その５)

$\boxed{\text{問題 1}}$　$X_1, X_2, \ldots, X_n \,(n \geq 2)$ を正規母集団 $N(\mu, 1)$ からの大きさ n の無作為標本とし，検定問題

$$H_0:\, \mu = 0,\ \mathrm{vs.} H_1:\, \mu > 0$$

を考える．
(1) $\bar{x}_n = (1/n)\sum_{i=1}^n x_i$ としたとき，

$$\sum_{i=1}^n (x_i - \mu)^2 - \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

を求めよ．
(2) 観測 $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ を得たとき，μ に関する尤度関数 $L(\mu)$ を書け．ただし，$L(\mu)$ は $\exp\left(-\frac{n}{2}(\bar{x}_n - \mu)^2.\right)$ の乗数倍という形でかけるが，その乗数も含めて書くこと．
(3) 検定問題 $H_0:\, \mu = 0,\ \mathrm{vs.} H_1:\, \mu' \,(\mu' > 0)$ を考える．定数 $k\,(0 < k < 1)$ に対して，

$$L(0) < kL(\mu') \Longleftrightarrow \bar{x}_n > k'$$

となるが，k' を k, μ', n 等必要な記号を用いて表現せよ．
(4) 帰無仮説 H_0 のもと，$\bar{X}_n$ の分布を求め，$\bar{X}_n$ を一次変換することにより，標準正規分布に従う確率変数を構成せよ．
(5) 検定関数を

$$\varphi(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & (\bar{x}_n > k') \\ 0 & (\bar{x}_n \leq k') \end{cases}$$

と定めたとき，有意水準 $\alpha\,(0 < \alpha < 1)$ となるように k' を

$$\int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt = \alpha$$

となる u の値 $u(\alpha)$ を用いて定めよ．

$\boxed{\textbf{Solution}}$
(1)

$$\begin{aligned}
\sum_{i=1}^n (x_i - \mu)^2 &= \sum_{i=1}^n (x_i - \bar{x}_n + \bar{x}_n - \mu)^2 \\
&= \sum_{i=1}^n \left\{(x_i - \bar{x}_n)^2 + (\bar{x}_n - \mu)^2 + 2(x_i - \bar{x}_n)(\bar{x}_n - \mu)\right\} \\
&= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + \sum_{i=1}^n (\bar{x}_n - \mu)^2 + 2(\bar{x}_n - \mu)\sum_{i=1}^n (x_i - \bar{x}_n)\} \\
&= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + n(\bar{x}_n - \mu)^2.
\end{aligned}$$

最後の等号は，

$$\sum_{i=1}^n (x_i - \bar{x}_n) \;=\; \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{x}_n = n\bar{x}_n - n\bar{x}_n = 0$$

からわかる．よって，

$$\sum_{i=1}^n (x_i - \mu)^2 - \sum_{i=1}^n (x_i - \bar{x}_n)^2 = n(\bar{x}_n - \mu)^2$$

(2)

$$\begin{aligned}
L(\mu) &= \Pi_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i-\mu)^2\right) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2\right) \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i-\bar{x}_n)^2 - \frac{n}{2}(\bar{x}_n-\mu)^2.\right) \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i-\bar{x}_n)^2\right) \exp\left(-\frac{n}{2}(\bar{x}_n-\mu)^2.\right)
\end{aligned}$$

(3)

$$\begin{aligned}
L(0) < kL(\mu') &\iff \frac{L(0)}{L(\mu')} < k \\
&\iff \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i-\bar{x}_n)^2\right)\exp\left(-\frac{n}{2}\bar{x}_n^2\right)}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i-\bar{x}_n)^2\right)\exp\left(-\frac{n}{2}(\bar{x}_n-\mu')^2\right)} < k \\
&\iff \frac{\exp\left(-\frac{n}{2}\bar{x}_n^2\right)}{\exp\left(-\frac{n}{2}(\bar{x}_n-\mu')^2\right)} < k \\
&\iff \exp\left(-\frac{n}{2}\bar{x}_n^2 + \frac{n}{2}(\bar{x}_n-\mu')^2\right) < k \\
&\iff -\frac{n}{2}\bar{x}_n^2 + \frac{n}{2}(\bar{x}_n-\mu')^2 < \log k \\
&\iff \bar{x}_n^2 - (\bar{x}_n-\mu')^2 > -\frac{2}{n}\log k \\
&\iff \bar{x}_n^2 - \bar{x}_n^2 + 2\mu'\bar{x}_n - (\mu')^2 > -\frac{2}{n}\log k \\
&\iff 2\mu'\bar{x}_n > (\mu')^2 - \frac{2}{n}\log k \\
&\iff \bar{x}_n > \frac{\mu'}{2} - \frac{1}{n\mu'}\log k
\end{aligned}$$

よって，

$$k' = \frac{\mu'}{2} - \frac{1}{n\mu'}\log k$$

(4) $\bar{X}_n$ は正規分布に従う確率変数 $X_1, \ldots, X_n$ の一次結合なので，その分布は正規分布になるので，$\bar{X}_n$ の期待値と分散を計算する：H_0 のもと，$\mathbb{E}[X_i]=0$ となることに注意すれば，

$$\begin{aligned}
\mathbb{E}[\bar{X}_n] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[X_i\right] = 0.
\end{aligned}$$

また，$X_1,\,\ldots,\,X_n$ は互いに独立であるので，$i \neq j$ のとき，$\mathbb{COV}[X_i,\,X_j] = 0,\,\mathbb{VAR}[X_i] = 1$ となることと分散の性質より

$$\begin{aligned}
\mathbb{VAR}[\bar{X}_n] &= \mathbb{VAR}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \\
&= \left(\frac{1}{n}\right)^2 \mathbb{VAR}\left[\sum_{i=1}^{n} X_i\right] \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathbb{VAR}\left[X_i\right] + \sum_{i,\,j:\,i\neq j} \mathbb{COV}[X_i,\,X_j] \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathbb{VAR}\left[X_i\right] = \frac{1}{n}\mathbb{VAR}[X_1] = \frac{1}{n}
\end{aligned}$$

よって，

$$\frac{\bar{X}_n - \mathbb{E}[\bar{X}_n]}{\sqrt{\mathbb{VAR}[\bar{X}_n]}} = \frac{\bar{X}_n -}{\sqrt{1/n}} = \sqrt{n}\bar{X}_n$$

は標準正規分布に従う．

$$\begin{aligned}
\mathbb{E}_{\mu=0}[\varphi(X_1,\,\ldots,\,X_n)] &= \mathbb{E}_{\mu=0}[\bar{X}_n > k'] \\
\mathbb{E}_{\mu=0}[\sqrt{n}\bar{X}_n > \sqrt{n}k'] & \\
&= \int_{\sqrt{n}k'}^{i} nfty \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,dt
\end{aligned}$$

より

$$\sqrt{n}k' = u(\alpha) \Longleftrightarrow k' = \frac{u(\alpha)}{\sqrt{n}}$$