4.2 正規分布からのランダム標本

定理 4.4 $X_1, X_2, \ldots, X_n$ を正規分布 $N(\mu, \sigma^2)$ からの標本の大きさが n のランダム標本とし，$\bar{X}_n = (1/n)\sum_{i=1}^n X_i$ と $S_n^2 = (1/(n-1))\sum_{i=1}^n (X_i - \bar{X}_n)^2$ とおく．このとき，以下が成立する：

(1) $\bar{X}_n$ と S_n^2 は独立である．

(2) $\bar{X}_n$ は正規分布 $N(\mu, \sigma^2/n)$ に従う．

(3) $(n-1)S_n^2/\sigma^2$ は自由度 $n-1$ のカイ自乗分布に従う．

証明 (2) は例 4.2 からわかる．次に，(1) を示す．各 i に対して，$Y_i = (X_i - \mu)/\sigma$ とすれば，Y_i は 正規分布 $N(0, 1)$ に従い，$(\bar{X}_n - \mu)/\sigma = \bar{Y}_n/\sigma = (1/n)\sum_{i=1}^n Y_i$ と $S_n^2/\sigma^2 = (1/(n-1))\sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ となるので，一般性を失わず，X_i は正規分布 $N(0, 1)$ に従うとして，$\bar{X}_n$ と S_n^2 の独立性を示せばよいことがわかる．

$X_i - \bar{X}_n$ は正規分布 $N(0, (1-1/n))$ に従うことがわかる．なぜならば，

$$
\begin{aligned}
M_{X_i - \bar{X}_n}(t) &= \mathbb{E}\left[\exp\left(t(1-1/n)X_i - \sum_{j=\neq i}(t/n)X_j\right)\right] \\
&= \mathbb{E}[e^{t(1-1/n)X_i}]\prod_{j\neq i}\mathbb{E}[e^{(t/n)X_j}] \\
&= \exp\left(\frac{t^2(1-(1/n))^2}{2}\right)\prod_{j\neq i}\exp\left(\frac{((t/n)^2}{2}\right) \\
&= \exp\left(\frac{(1-1/n)t^2}{2}\right)
\end{aligned}
$$

からわかる．$\bar{X}_n$ と $X_i - \bar{X}_n$ はともに正規分布に従うので，$\bar{X}_n$ と $X_i - \bar{X}_n$ が独立であることをいうためには，$\mathbb{COV}[\bar{X}_n, X_i - \bar{X}_n] = 0$ を示せばよい．しかし，

$$
\begin{aligned}
\mathbb{COV}[\bar{X}_n, X_i - \bar{X}_n] &= \mathbb{E}[\bar{X}_n(X_i - \bar{X}_n)] \\
&= \mathbb{E}[(1/n)\sum_{j=1}^n X_j X_i] - \mathbb{E}[\bar{X}_n^2] = \frac{1}{n}\sum_{j\neq i}\mathbb{E}[X_i X_j] + \frac{1}{n}\mathbb{E}[X_i^2] - \mathbb{E}[\bar{X}_n^2] \\
&= \frac{1}{n}\mathbb{VAR}[X_i] - \mathbb{VAR}[\bar{X}_n] = \frac{1}{n} - \frac{1}{n} = 0
\end{aligned}
$$

となる．よって，$\bar{X}_n$ と $X_i - \bar{X}_n$ が独立である．これから $\bar{X}_n$ と $(X_1 - \bar{X}_n, X_2 - \bar{X}_n, \ldots, X_n - \bar{X}_n)$ は独立[2]となり，$\bar{X}_n$ と S_n^2 は独立であることがわかる．

(3) を示すために，

$$
\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu)^2
$$

[2] ここの議論は正規分布に従っていることから保障される．

に注意する．$Y_i = (X_i - \mu)/\sigma,\ i = 1,\ 2,\ \ldots,\ n,$ とおけば，$\bar{Y}_n = (1/n)\sum_{i=1}^n Y_i = (\bar{X}_n - \mu)/\sigma$ となり，Y_i と $\bar{Y}_n$ は $N(0,\ 1)$ と $N(0, 1/n)$ に従う．しがたって，$U = \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ が自由度 $n-1$ のカイ自乗分布に従うことを示せばよい．いま，$W = \sum_{i=1}^n Y_i^2,\ V = n\bar{Y}_n^2$ とおけば，$t < 1/2$ に対して，W, V の積率母関数は

$$\begin{aligned} M_W(t) &= \mathbb{E}[\exp(tW)] = \left(\frac{1}{1-2t}\right)^n, \\ M_V(t) &= \mathbb{E}[\exp(tV)] = \left(\frac{1}{1-2t}\right) \end{aligned}$$

となる．さらに，$\bar{Y}_n$ と $(Y_1 - \bar{Y}_n,\ Y_2 - \bar{Y}_n,\ \ldots,\ Y_n - \bar{Y}_n)$ は独立であることに注意して，$U = \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ の積率母関数を求める：

$$\begin{aligned} M_W(t) &= \mathbb{E}[\exp(tW)] = \mathbb{E}[\exp\{t\sum_{i=1}^n (Y_i - \bar{Y}_n)^2\} + tn\bar{Y}_n^2] \\ &= \mathbb{E}[\exp\{tU + tV\}] = \mathbb{E}[\exp\{tU\}]\mathbb{E}[\exp(tV)] = M_U(t)M_V(t) \end{aligned}$$

より

$$M_U(t) = \frac{M_W(t)}{M_V(t)} = \left(\frac{1}{1-2t}\right)^{n-1}$$

がわかる． □

4.2.1 t 分布と F 分布

$X_1,\ X_2,\ \ldots,\ X_n$ が正規分布 $N(\mu,\ \sigma^2)$ からの標本の大きさが n のランダム標本としたとき，

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \tag{4.6}$$

は標準正規分布 $N(0,\ 1)$ に従うことが定理 4.4 (2) からわかる． σ が既知であれば，$\bar{X}_n$ を観測したときに，(4.6) は μ の推測に利用できる．しかし，σ が未知のときは，(4.6) の代わりに

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \tag{4.7}$$

を μ の推測に用いる．ただし，$S_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ で S_n は S_n^2 の正の平方根である．(4.7) の標本分布を求めるために

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} = \frac{(\bar{X}_n - \mu)/(\sigma/\sqrt{n})}{\sqrt{S_n^2/\sigma^2}} \tag{4.8}$$

と書き換える．(4.8) の分子は標準正規分布 $N(0,\ 1)$ に従い，分母は $\sqrt{\chi_{n-1}^2/(n-1)}$ と同じ分布で，分母と分子は独立である．ただし，χ_{n-1}^2 は自由度 $(n-1)$ のカイ自乗分布である．したがって，(4.8) の分布は $V/\sqrt{U/(n-1)}$ と同じ分布である．ただし，U と V は独立に自由度 $(n-1)$ のカイ自乗分布と標準正規分布に従うものとする．

定義 4.3 $X_1, X_2, \ldots, X_n, n \geq 2,$ は正規分布 $N(\mu, \sigma^2)$ からの標本の大きさが n のランダム標本とするとき，

$$\frac{\bar{X}_n - \mu}{S/\sqrt{n}}$$

は自由度 $n-1$ の t 分布に従うという．p を自然数としたとき，確率変数 T が自由度 p の t 分布に従うとは，T が確率密度関数

$$f_T(t) = \frac{\Gamma((p+1)/2)}{\Gamma(p/2)} \frac{1}{\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}}, \qquad -\infty < t < \infty$$

を持つときをいう．

<u>t 分布の確率密度関数の導出について</u>

証明

$$U = (n-1)\frac{S_n^2}{\sigma^2}, \qquad V = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}, \qquad p = n-1$$

とおくと U と V は独立で，それぞれは自由度 p のカイ自乗分布 χ_p^2 と標準正規分布 $N(0, 1)$ に従い，

$$T = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} = \frac{V}{\sqrt{U/p}}$$

となる．したがって，U と V から出発して，$\sqrt{p}U/V$ の確率密度関数を求める．まず，

$$f_{U,V}(u, v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2} \frac{1}{\Gamma(p/2)2^{p/2}} u^{(p/2)-1} e^{-u/2}, \qquad -\infty < v < \infty, \quad 0 < u < \infty$$

に注意する．いま

$$t = \frac{v}{\sqrt{u/p}}, \qquad w = u$$

とおくと

$$J = \begin{vmatrix} (\partial u/\partial t) & (\partial u/\partial w) \\ (\partial v/\partial t) & (\partial v/\partial w) \end{vmatrix} = \begin{vmatrix} \sqrt{\frac{w}{p}} & * \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{w}{p}}$$

から

$$\begin{aligned} f_T(t) &= \int_0^\infty f_{U,V}\left(t\sqrt{\frac{w}{p}}, w\right)\sqrt{\frac{w}{p}}\,dw \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(p/2)2^{p/2}} \int_0^\infty e^{-(1/2)t^2 w/p} w^{(p/2)-1} e\ -w/2 \sqrt{\frac{w}{p}}\,dw \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(p/2)2^{p/2}p^{1/2}} \int_0^\infty e^{-(1/2)(1+t^2/p)w} w^{(p+1)/2-1}\,dw \end{aligned}$$

となる．さらに，

$$z = \left(1 + \frac{t^2}{p}\right)w$$

とおけば,

$$\begin{aligned} f_T(t) &= \frac{1}{\Gamma(p/2)\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}} \frac{1}{2^{(p+1)/2}} \int_0^\infty w^{(p+1)/2-1} e^{-w/2}\, dw \\ &= \frac{\Gamma((p+1)/2)}{\Gamma(p/2)} \frac{1}{\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}} \end{aligned}$$

を得る. □

定義 4.4 $X_1, X_2, \ldots, X_n, n \geq 2,$ は正規分布 $N(\mu_X, \sigma_X^2)$ からの標本の大きさが n のランダム標本とし，$Y_1, Y_2, \ldots, Y_m, m \geq 2,$ は $X_i, i = 1, 2, \ldots, n,$ とは独立な正規分布 $N(\mu_Y, \sigma_Y^2)$ からの標本の大きさが m のランダム標本とする．このとき,

$$F = \frac{(S_X^2/\sigma_X^2)}{(S_Y^2/\sigma_Y^2)}$$

は自由度 $n-1$ と $m-1$ の F 分布に従うという．p, q を自然数としたとき，確率変数 F が自由度 p と q の t 分布に従うとは，T が確率密度関数

$$f_F(x) = \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \left(\frac{p}{q}\right)^{p/2} \frac{x^{(p/2)-1}}{(1+(p/q)x)^{(p+q)/2}}, \qquad 0 < x < \infty$$

を持つときをいう.