分布論 (Distribution Theory)

(July 31, 2015)

広島大学 大学院理学研究科 柳原宏和

e-mail: *yanagi@math.sci.hiroshima-u.ac.jp*

0. 序論

何故分布論が必要か?

分布論は統計的推測のために必要!

今, データ $x_1, \ldots, x_n$ に対して統計モデルを想定する. つまり, データを確率変数 $X_1, \ldots, X_n$ の実現値と仮定する. ただし n は標本数とする. ここでは, $(X_1, \ldots, X_n)$ に確率分布を想定 (正規分布などの分布のタイプを仮定) して以下のようなパラメトリックモデルを考える.

$$\mathrm{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \int_{-\infty}^{x_1} \cdots \int_{\infty}^{x_n} f(t_1, \ldots, t_n; \theta) dt_1 \cdots dt_n, \ \theta \in \Omega \ (f \text{ は既知}, \theta \text{ は未知}).$$

以下のような統計的推測を行うためには f を仮定し使用する必要がある.

- 仮説検定: 以下の検定問題

$$\text{帰無仮説 } H: \ \theta \in \Omega_0 \text{ vs 対立仮説 } K: \ \theta \in \Omega_1 = \Omega - \Omega_0,$$

  を考える. 一般的な仮説検定の手法では, ある閾値 c に対して, 検定統計量 $T(X_1, \ldots, X_n)$ の実現値 $T(x_1, \ldots, x_n)$ が $T(x_1, \ldots, x_n) \geq c$ であれば H を棄却する. ここでは $T = T(X_1, \ldots, X_n)$ と書くことにする. 仮説検定を行うためには, 具体的な c を決定する必要がある. これは検定の有意水準によって変化するので, f により有意水準を計算する必要がある. また使用する検定統計量の良し悪しを測るためには検出力を考える必要がある.

  1. 有意水準 α: $\mathrm{P}(T \geq c|H) \leq \alpha$ (通常は $\alpha = 0.05$ or 0.01). この確率は第一種の過誤 (H が正しいにもかかわらず H を棄却する間違い) を起こす確率に等しい. また H が正しい下での T の分布のことを帰無分布 (null distribution) と呼ぶ.
  2. 検出力 β: $\beta(\theta) = \mathrm{P}(T \geq c|K)$. この確率は 1 から第二種の過誤 ($K$ が正しいにもかかわらず H を採択する) を起こす確率を引いたものに等しい. また K が正しい下での T の分布のことを対立分布 (non-null distribution) と呼ぶ.

- 推定: θ の推定量 $\hat{\theta}(X_1, \ldots, X_n)$ とし (推定値は $\hat{\theta}(x_1, \ldots, x_n)$) $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ と書くこととする. 推定量の良さを測る指数, 平均 (バイアス) $\mathrm{E}[\hat{\theta}]$, 分散 $\mathrm{Var}[\hat{\theta}]$, 平均二乗誤差 (Mean Squared Error: MSE) $\mathrm{MSE}[\hat{\theta}] = \mathrm{E}[(\hat{\theta} - \theta)^2]$ を計算するためには $\hat{\theta}$ の分布を知る必要がある. また $\hat{\theta}$ の分布が分かれば θ の区間推定 (例えば $a \leq \theta \leq b$ と言った形で推定) を行うことができる.

注意

1. 検定統計量, 推定量などを統計量と呼び, 一般的には T と書く. このとき統計量 T の分布とは T の分布関数 (Distribution Function; df) や確率密度関数 (Probability Density Function; pdf) を指す.

$$\begin{aligned} \mathrm{P}(T \leq x) &= F(x) \quad (F(x)\text{: 分布関数}) \\ &= \int_{-\infty}^{x} f(t) dt \quad (f(x)\text{: 確率密度関数}). \end{aligned}$$

2. 確率変数 X の確率密度関数が確率分布 G の確率密度関数であるとき, X は G に従うといい, $X \sim G$ とかく. 確率変数 $X_1, \ldots, X_n$ が互いに独立で同じ確率分布 G に従うとき, $X_1, \ldots, X_n \sim i.i.d.\ G$ と書く (i.i.d. は independently and identically distributed の略). G の代わりに, G の分布関数や確率密度関数を書くこともある.

3. **期待値** (Expectation) と特性量:

   定義

   確率変数 X の確率密度関数を $f(x)$ とする. ある関数 $g(\cdot)$ に対して, 以下を $g(X)$ の期待値という.

   $$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

   ここで, $\mathrm{E}[X^r]$ を X の原点回りの r 次**モーメント** (r**th Moment**) と呼び, μ_r で表す. さらに, $\mathrm{E}[(X-\mu)^r]$ を X の平均周りの r 次モーメント (または単に r 次モーメント) と呼び α_r で表す. ここで, $\mu = \mathrm{E}[X]$ である.

   以下は確率分布の形状を表す特性量である.

   (a) **平均** (**Mean**): $\mu = \mu_1$. 確率分布の重心, または中心.
   (b) **分散** (**Variance**): $\mathrm{Var}[X] = \sigma^2 = \alpha_2$. 確率分布の散らばり. $\sigma = \sqrt{\sigma^2}$ を**標準偏差** (**Standard Deviation**) と呼ぶ.
   (c) **歪度** (**Skewness**): $\kappa_3 = \alpha_3/\sigma^3$. 確率分布の歪み. 左右対称な分布であれば $\kappa_3 = 0$. $\kappa_3 > 0$ のとき, 右に歪んだ分布 (分布の裾が右側に広がっている), $\kappa_3 < 0$ のとき, 左に歪んだ分布 (分布の裾が左側に広がっている) という.
   (d) **尖度** (**Kurtosis**): $\kappa_4 = \alpha_4/\sigma^3 - 3$. 分散 1 に固定したときの分布の裾の重さ, または尖り具合. 尖度が大きくなるほど裾の重い分布. $\kappa_4 \geq -2$ であり, $\kappa_4 = -2$ のときは一点分布 (退化した分布).

   正規分布の場合は $\kappa_3 = 0$, $\kappa_4 = 0$ となることから, 正規性からの逸脱度を表す指標としても用いられる.

4. **特性関数** (**Characteristic Function**; cf):

   定義

   確率変数 X の確率密度関数を $f(x)$ とする. ${}^{\forall}t \in \mathbb{R}$ に対して, $\mathrm{E}[\exp(itX)]$ を X の特性関数といい, $C(t)$ で表す (X の特性関数ということを明記するために $C_X(t)$ と書くこともある).

   特性関数は以下の様な特性を持つ.

   (a) 一意性定理: 確率変数 X_1, X_2 の特性関数を $C_1(t)$, $C_2(t)$ とすると,

   $$C_1(t) = C_2(t) \Leftrightarrow X_1 \text{ と } X_2 \text{ の分布は同じ.}$$

   (b) モーメント公式:

   $$\mathrm{E}[|X|^n] < \infty \Rightarrow \left.\frac{d^n}{dt^n}C(t)\right|_{t=0} = i^n\mathrm{E}[X^n].$$

   ここでの条件 $\mathrm{E}[|X|^n] < \infty$ は微分と積分の順序を入れ替えるために必要である.

計算機を使った分布の近似法

$U \sim \mathrm{U}(0,1)$ ($\mathrm{U}(a,b)$ は区間 (a,b) 上の一様分布を表す) とする. このとき U の確率密度関数は,

$$f(x) = \begin{cases} 1 & (0 \le x \le 1) \\ 0 & (\text{その他}) \end{cases},$$

であり, 分布関数は

$$\mathrm{P}(U \le x) = \begin{cases} 0 & (0 < x) \\ x & (0 \le x \le 1) \\ 1 & (x > 1) \end{cases},$$

である. また $X \sim \mathrm{N}(0,1)$ ($\mathrm{N}(\mu, \sigma^2)$ は平均 μ, 分散 σ^2 の正規分布を表す) とすると,

$$\mathrm{P}(X \le x) = \int_{\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x).$$

$\Phi(x)$ は狭義単調増加関数であるので, Φ^{-1} が存在する. いま, $Y = \Phi^{-1}(U)$ ($U \sim \mathrm{U}(0,1)$) とすると,

$$\mathrm{P}(Y \le y) = \mathrm{P}(\Phi^{-1}(U) \le y) = \mathrm{P}(U \le \Phi(y)) = \Phi(y).$$

よって Y は標準正規分布に従う確率変数であることがわかる. 一般に, $U \sim \mathrm{U}(0,1)$ で, X の分布関数 $F(x)$ を狭義単調増加関数としたとき, $Y = F^{-1}(U)$ の分布は X の分布に等しい. つまり,

$$\mathrm{P}(Y \le x) = \mathrm{P}(F^{-1}(U) \le x) = \mathrm{P}(U \le F(x)) = F(x),$$

が成り立つ.

(*Example*) $U_1, \ldots, U_n \sim i.i.d.\ \mathrm{U}(0,1)$ とするとし, F を狭義単調増加なある分布関数とする. このとき,

$$X_1 = F^{-1}(U_1), \ldots, X_n = F^{-1}(U_n),$$

とすれば

$$X_1, \ldots, X_n \sim i.i.d.\ F,$$

である.

上記の方法を用いて $T(X_1, \ldots, X_n)$ の分布関数の近似関数を作る. 今,

$$\begin{array}{ccc} (x_1^{(1)}, \ldots, x_n^{(1)}) & \to & T^{(1)} = T(X_1^{(1)}, \ldots, X_n^{(1)}) = F^{-1}(U^{(1)}), \\ \vdots & \vdots & \vdots \\ (x_1^{(m)}, \ldots, x_n^{(m)}) & \to & T^{(m)} = T(X_1^{(m)}, \ldots, X_n^{(m)}) = F^{-1}(U^{(m)}), \end{array}$$

とすると, $T^{(1)}, \ldots, T^{(m)}$ と $U^{(1)}, \ldots, U^{(m)}$ の順序は同じである (F は狭義単調増加関数なので). ここで, $U^{(1)} = F(T^{(1)}), \ldots, U^{(m)} = F(T^{(m)})$ を小さい順に並べて

$$U^{<1>} = F(T^{<1>}), \ldots, U^{<m>} = F(T^{<m>}),$$

とおくと,

$$\mathrm{P}(T \le x) = \frac{\alpha}{m} \quad (\alpha = 1, \ldots, m),$$

であるとき, $x = T^{<\alpha>} = F^{-1}(U^{<\alpha>})$ である.

しかしながら, F の具体的な形が分かっていなければ, 上記のような方法で分布を近似することはできない.

1. 1次元正規分布とその拡張分布

1.1. 1次元正規分布 (Univariate Normal Distributuon)

定義

$X \sim \mathrm{N}(\mu, \sigma^2)$ であるとは, 確率変数 X が従う分布の確率密度関数が

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \quad (\mu \in \mathbb{R}, \ \sigma > 0), \tag{1.1}$$

であるときのことをいう. また, この確率分布を平均 μ, 分散 σ^2 の1次元正規分布という. 特に, $\mu = 0$, $\sigma^2 = 1$ のときの1次元正規分布を**標準正規分布** (**Standard Normal Distribution**) と呼び, その確率密度関数を $\phi(x)$, 分布関数を $\Phi(x)$ と書く.

正規分布は統計学において最も使用頻度が高い分布である. 特に $X_1, \ldots, X_n \sim i.i.d.$ $\mathrm{E}[X_i] = \mu$, $\mathrm{Var}[X_i] = \sigma^2$ $(< \infty)$ のとき, **中心極限定理** (**Central Limit Theory**) により

$$\frac{\sqrt{n}}{\sigma}\left\{\frac{1}{n}(X_1 + \cdots + X_n) - \mu\right\} \xrightarrow{d} \mathrm{N}(0,1),$$

であることが知られている. また同様の中心極限定理よりほとんどの推定量が漸近的に正規分布に従うことも知られている.

特性

1. 正規分布は μ を中心に左右対称の分布であり, μ は中心の位置, σ^2 は裾の広がりを表している.

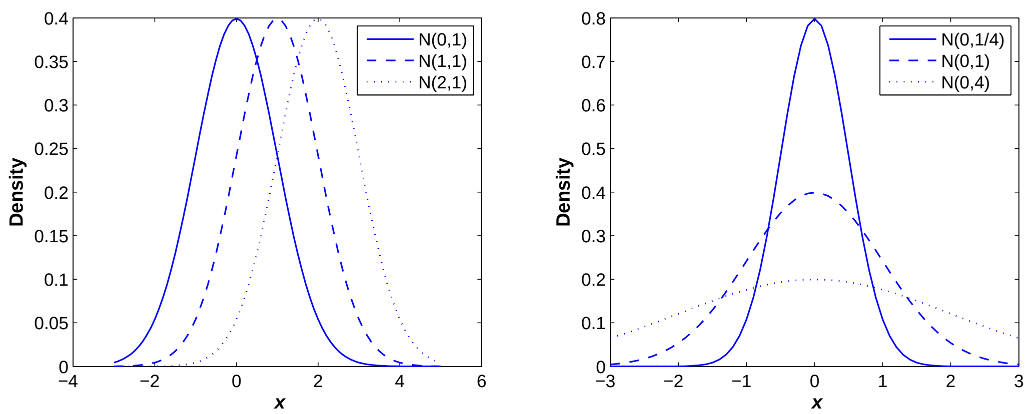

Figure 1.1. 正規分布の形

2. $X \sim \mathrm{N}(\mu, \sigma^2)$ のとき, $Z = (X - \eta)/\psi$ $(\psi > 0)$ とおくと, $Z \sim \mathrm{N}((\mu - \eta)/\psi, \sigma^2/\psi^2)$ となる.

   (*Proof*) Z の分布関数が $\mathrm{N}((\mu - \eta)/\psi, \sigma^2/\psi^2)$ の分布関数になることを言えばよい. Z の分布関数は,

$$\mathrm{P}(Z \leq z) = \mathrm{P}((X - \eta)/\psi) = \mathrm{P}(X \leq \psi z + \eta) = \int_{-\infty}^{\psi z + \eta} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right\} dt.$$

   ここで, $u = (t - \eta)/\psi$ とおくと, $du/dt = \psi^{-1}$ で, $t = \psi u + \eta$ である. このとき, 積分範囲は, $-\infty \leq t \leq$

$\psi z+\eta$ より, $-\infty \le u \le z$ となる. よって,

$$\begin{aligned}\mathrm{P}(Z \le z) &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\left(\frac{\psi u+\eta-\mu}{\sigma}\right)^2\right\}\psi^{-1}du \\ &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma^2/\psi^2}}\exp\left\{-\frac{1}{2}\left(\frac{u-(\mu-\eta)/\psi}{\sigma/\psi}\right)^2\right\}du.\end{aligned}$$

これは, $\mathrm{N}((\mu-\eta)/\psi, \sigma^2/\psi^2)$ の分布関数である. □

この特性より, $X \sim N(\mu,\sigma^2)$ のとき, $Z=(X-\mu)/\sigma \sim \mathrm{N}(0,1)$, $X \sim \mathrm{N}(0,1)$ のとき, $Z=\sigma X+\mu \sim \mathrm{N}(\mu,\sigma^2)$ が言える.

3. $X \sim \mathrm{N}(\mu,\sigma^2)$ のとき $\mathrm{E}[X]=\mu$, $\mathrm{Var}[X]=\sigma^2$ である. つまり

$$\mathrm{E}[X]=\int_{-\infty}^{\infty} xf(x;\mu,\sigma^2)dx=\mu, \quad \mathrm{Var}[X]=\mathrm{E}[(X-\mu)^2]=\int_{-\infty}^{\infty}(x-\mu)^2 f(x;\mu,\sigma^2)dx=\sigma^2.$$

また, r 次モーメントは

$$\mathrm{E}[(X-\mu)^r]=\begin{cases} 0 & (r \text{ が奇数}) \\ 1\cdot 3\times\cdots\times(r-1)\sigma^r & (r \text{ が偶数})\end{cases}. \tag{1.2}$$

(*Proof*) $X \sim \mathrm{N}(\mu,\sigma^2)$ のとき, r 次モーメントは

$$\mathrm{E}[(X-\mu)^r]=\int_{-\infty}^{\infty}(x-\mu)^r\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{1}{2\sigma^2}(x-\mu)^2\right\}dx,$$

である. 今, $z=(x-\mu)/\sigma$ とおくと, $dx=\sigma dz$ であるので,

$$\mathrm{E}[(X-\mu)^r]=\int_{-\infty}^{\infty}\sigma^r z^r\frac{1}{\sqrt{2\pi}}e^{z^2/2}dz=\sigma^r\int_{-\infty}^{\infty}z^r\phi(z)dz,$$

となる. ここで $\phi(z)$ は偶関数であることに注意すれば, r が奇数のとき $\mathrm{E}[(X-\mu)^r]=0$ であることがわかる. また r が偶数のときは,

$$\mathrm{E}[(X-\mu)^r]=2\sigma^r\int_{0}^{\infty}z^r\phi(z)dz,$$

となり, ここで $t=z^2/2$ とおくと, $dz=(2t)^{-1/2}dt$ であるので,

$$\begin{aligned}\mathrm{E}[(X-\mu)^r] &= 2\sigma^r\int_0^{\infty}(2t)^{(r-1)/2}\frac{1}{\sqrt{2\pi}}e^{-t}dt \\ &= \frac{2\cdot 2^{(r-1)/2}\sigma^r}{\sqrt{2\pi}}\int_0^{\infty}t^{(r+1)/2-1}e^{-t}dt=\frac{2^{r/2}\sigma^r}{\sqrt{\pi}}\Gamma((r+1)/2),\end{aligned} \tag{1.3}$$

となる. ここで $\Gamma(x)$ はガンマ関数である. 今, $\Gamma(x)=(x-1)\Gamma(x-1)$ と $\Gamma(1/2)=\sqrt{\pi}$ であることに注意すれば,

$$\begin{aligned}\Gamma((r+1)/2) &= \left\{\frac{1}{2}(r+1)-1\right\}\left\{\frac{1}{2}(r+1)-2\right\}\times\cdots\times\frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma(1/2) \\ &= 2^{-r/2}\sqrt{\pi}\cdot 1\cdot 3\times\cdots\times(r-1),\end{aligned}$$

となる. この式を (1.3) 式に代入すれば (1.2) 式を得ることができる. □

この特性より, $X \sim \mathrm{N}(\mu,\sigma^2)$ のとき, $\kappa_3=0$, $\kappa_4=0$ が言える.

4. $X \sim \mathrm{N}(\mu, \sigma^2)$ のとき, X の特性関数は

$$C(t) = \exp\left(it\mu - \frac{1}{2}t^2\sigma^2\right), \quad (t \in \mathbb{R}). \tag{1.4}$$

(*Proof*) $X \sim \mathrm{N}(\mu, \sigma^2)$ のとき,

$$C(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{1}{2\sigma^2}(x-\mu)^2\right\} dx,$$

である. 今 $z = (x-\mu)/\sigma$ とおくと, $dx = \sigma dz$ であるので,

$$\begin{aligned} C(t) &= \int_{-\infty}^{\infty} e^{it(\mu+\sigma z)} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = e^{it\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(z^2 - 2it\sigma z)\right\} dz \\ &= \exp\left(it\mu - \frac{1}{2}t^2\sigma^2\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(z - it\sigma)^2\right\} dz. \end{aligned}$$

上記の式の被積分関数は $\mathrm{N}(it\sigma, 1)$ の確率密度関数と見ることができるので, 確率密度関数の定義より,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(z - it\sigma)^2\right\} dz = 1.$$

よって (1.4) 式を導出することができる. □

5. 再生性: 互いに独立な確率変数 X_1, X_2 を, $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$ とする. このとき, $X_1 + X_2 \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ となる.

(*Proof*) $Z = X_1 + X_2$ の特性関数は,

$$C_Z(t) = \mathrm{E}\left[\exp(itZ)\right] = \mathrm{E}\left[\exp\{it(X_1 + X_2)\}\right] = \mathrm{E}\left[\exp(itX_1)\right]\mathrm{E}\left[\exp(itX_2)\right] = C_{X_1}(t)C_{X_2}(t).$$

ここで, 特性 4 より, $C_{X_i}(t) = \exp(it\mu_i - t^2\sigma_i^2/2)$ であるので,

$$C_Z(t) = \exp\left(it\mu_1 - \frac{1}{2}t^2\sigma_1^2\right)\exp\left(it\mu_2 - \frac{1}{2}t^2\sigma_2^2\right) = \exp\left\{it(\mu_1 + \mu_2) - \frac{1}{2}t^2(\sigma_1^2 + \sigma_2^2)\right\}.$$

これは, $\mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ の特性関数である. □

6. 乱数発生方法 (ボックス＝ミュラー法): $U_1, U_2 \sim i.i.d.\ \mathrm{U}(0,1)$ とすると,

$$Z_1 = \sqrt{-2\log U_1}\cos(2\pi U_2), \quad Z_2 = \sqrt{-2\log U_1}\sin(2\pi U_2),$$

は, $Z_1, Z_2 \sim i.i.d.\ \mathrm{N}(0,1)$ となる.

(*Proof*)

$$\begin{aligned} Z_1^2 + Z_2^2 &= -2\log U_1 \cos^2(2\pi U_2) - 2\log U_1 \sin^2(2\pi U_2) = -2\log U_1, \\ \frac{Z_2}{Z_1} &= \frac{\sqrt{-2\log U_1}\sin(2\pi U_2)}{\sqrt{-2\log U_1}\cos(2\pi U_2)} = \frac{\sin(2\pi U_2)}{\cos(2\pi U_2)} = \tan(2\pi U_2), \end{aligned}$$

より, $(U_1, U_2) \to (Z_1, Z_2)$ の逆変換は,

$$U_1 = \exp\left\{-\frac{1}{2}(Z_1^2 + Z_2^2)\right\}, \quad U_2 = \frac{1}{2\pi}\tan^{-1}(Z_2/Z_1),$$

となる. この変換に関するヤコビアンは,

$$
\begin{aligned}
J &= \begin{vmatrix} \partial u_1/\partial z_1 & \partial u_1/\partial z_2 \\ \partial u_2/\partial z_1 & \partial u_2/\partial z_2 \end{vmatrix} = \begin{vmatrix} -z_1\exp\{-(z_1^2+z_2^2)/2\} & -z_2\exp\{-(z_1^2+z_2^2)/2\} \\ -z_2(z_1^2+z_2^2)^{-1}/(2\pi) & z_1(z_1^2+z_2^2)^{-1}/(2\pi) \end{vmatrix} \\
&= -\frac{z_1^2}{2\pi(z_1^2+z_2^2)}\exp\left\{-\frac{1}{2}(z_1^2+z_2^2)\right\} - \frac{z_2^2}{2\pi(z_1^2+z_2^2)}\exp\left\{-\frac{1}{2}(z_1^2+z_2^2)\right\} \\
&= -\frac{1}{2\pi}\exp\left\{-\frac{1}{2}(z_1^2+z_2^2)\right\}.
\end{aligned}
$$

一方, U_i の確率密度関数は,

$$
g(u) = \begin{cases} 1 & (0 \le u \le 1) \\ 0 & (\text{その他}) \end{cases},
$$

であるので, Z_1, Z_2 の同時確率密度関数は,

$$
\begin{aligned}
f(z_1, z_2) &= g(\exp\{-(z_1^2+z_2^2)/2\})g(\tan^{-1}(z_2/z_2)/(2\pi))|J| = |J| \\
&= \frac{1}{2\pi}\exp\left\{-\frac{1}{2}(z_1^2+z_2^2)\right\} = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z_1^2}{2}\right) \times \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z_2^2}{2}\right).
\end{aligned}
$$

これは互いに独立に標準正規分布に従う確率変数の同時確率密度関数である. □

実際のデータが従う分布は正規分布のように綺麗な分布であるとは限らない!

$\rightarrow$ 歪度や尖度を自由に変えることができないか?

$\rightarrow$ 混合正規分布や歪正規分布へ.

1.2. 混合正規分布 (Contaminated Normal Distribution)

<u>定義</u>

$X \sim \mathrm{CN}(\lambda, \alpha)$ であるとは, 確率変数 X が従う分布の確率密度関数が

$$f(x;\lambda,\alpha) = (1-\alpha)\phi(x) + \frac{\alpha}{\lambda}\phi(x/\lambda),\ (\lambda > 0, \alpha \in [0,1]), \tag{1.5}$$

であるときのことをいう. これは分散の異なる正規分布を混合比 α で混ぜた分布となっているため, この確率分布を分散パラメータ λ, 混合比 α の混合正規分布という.

<u>特性</u>

1. 混合正規分布は 0 を中心とした左右対称な分布であり, $\alpha = 0, 1$ か $\lambda = 1$ のときは正規分布となる.

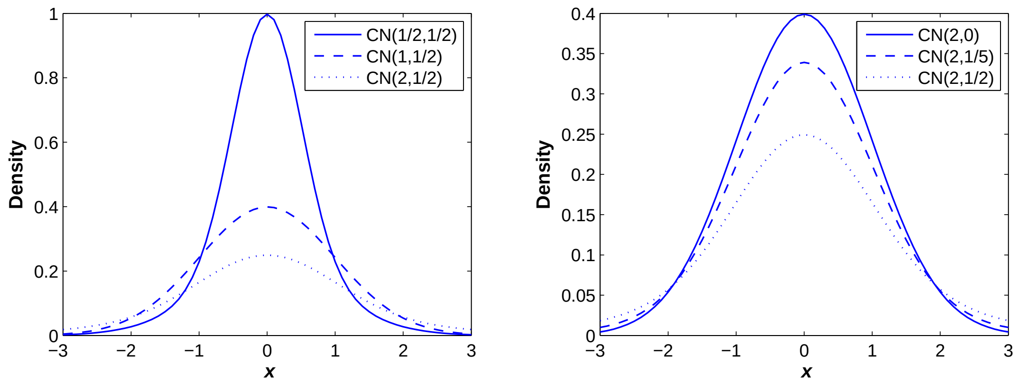

Figure 1.2. 混合正規分布の形

2. $X \sim \mathrm{CN}(\lambda, \alpha)$ のとき

$$\mathrm{E}[X] = 0,\ \mathrm{Var}[X] = 1 - \alpha + \alpha\lambda^2,\ \kappa_3 = 0,\ \kappa_4 = \frac{3\alpha(1-\alpha)(\lambda^2-1)^2}{(1-\alpha+\alpha\lambda^2)^2}.$$

r 次モーメントは

$$\mathrm{E}[(X-\mu)^r] = \begin{cases} 0 & (r \text{ が奇数}) \\ 1 \cdot 3 \times \cdots \times (r-1)(1-\alpha+\alpha\lambda^r) & (r \text{ が偶数}) \end{cases}. \tag{1.6}$$

(*Proof*) $X \sim \mathrm{CN}(\lambda, \alpha)$ のとき, r 次モーメントは

$$\mathrm{E}[X^r] = (1-\alpha)\int_{-\infty}^{\infty} x^r \phi(x)dx + \alpha \int_{-\infty}^{\infty} x^r \frac{1}{\lambda}\phi(x/\lambda)dx = (1-\alpha)\mathrm{E}[X_1^r] + \alpha\mathrm{E}[X_2^r],$$

となる. ただし $X_1 \sim \mathrm{N}(0,1)$, $X_2 \sim \mathrm{N}(0,\lambda^2)$ である. 上記の式と正規分布の r 次モーメント (1.2) より (1.6) 式を得ることができる. □

特に $\alpha = 1/(1+\lambda^2)$ とおくと, (1.5) 式は λ にだけ依存する形になり,

$$f(x;\lambda) = \frac{\lambda^2}{1+\lambda^2}\phi(x) + \frac{1}{\lambda(1+\lambda^2)}\phi(x/\lambda),\ (\lambda > 0),$$

と簡単な形となる. このとき $\lambda = 1$ であれば正規分布である. また, $\kappa_4 = 3(\lambda - 1/\lambda)^2/4$ となり, $\lambda \to \infty$ とすれば $\kappa_4 \to \infty$ となることがわかる.

3. $X \sim \mathrm{CN}(\lambda, \alpha)$ のとき, X の特性関数は

$$C(t) = (1-\alpha)\mathrm{exp}\left(-\frac{1}{2}t^2\right) + \alpha\mathrm{exp}\left(-\frac{1}{2}t^2\lambda^2\right), \quad (t \in \mathbb{R}). \tag{1.7}$$

(*Proof*) $X \sim \mathrm{CN}(\lambda, \alpha)$ のとき,

$$C(t) = (1-\alpha)\int_{-\infty}^{\infty} e^{itx}\phi(x)dx + \alpha\int_{-\infty}^{\infty} e^{itx}\frac{1}{\lambda}\phi(x/\lambda)dx = \alpha\mathrm{E}[\exp(itX_1)] + (1-\alpha)\mathrm{E}[\exp(itX_2)],$$

である. ただし, $X_1 \sim \mathrm{N}(0,1)$, $X_2 \sim \mathrm{N}(0,\lambda^2)$. ここで, 正規分布の特性 4 より,

$$\mathrm{E}[\exp(itX_1)] = \exp\left(-\frac{1}{2}t^2\right), \quad \mathrm{E}[\exp(itX_2)] = \exp\left(-\frac{1}{2}t^2\lambda^2\right).$$

よって, (1.7) を得る. □

4. 乱数発生方法: それぞれ独立な確率変数 Z_1, Z_2, U を $Z_1 \sim \mathrm{N}(0,1)$, $Z_2 \sim \mathrm{N}(0,\lambda^2)$, $U \sim \mathrm{B}(1,\alpha)$ とすると $X \sim \mathrm{CN}(\lambda,\alpha)$ となる X は以下のように発生させることができる.

$$X = (1-U)Z_1 + UZ_2.$$

ただし, $\mathrm{B}(n,p)$ は繰り返し回数 n, 成功確率 p $(0 \leq p \leq 1)$ の二項分布を表す. 二項分布の確率関数は,

$$\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}, \ (k = 0, 1, \ldots, n),$$

である.

(*Proof*) X の分布関数が $\mathrm{CN}(\lambda,\alpha)$ の分布関数になることを言えばよい. X の分布関数は,

$$\mathrm{P}(X \leq x) = \mathrm{P}((1-U)Z_1 + UZ_2 \leq x) = \mathrm{P}(\{\{Z_1 \leq x\} \cap \{U=0\}\} \cup \{\{Z_2 \leq x\} \cap \{U=1\}\}).$$

ここで, $\{Z_1 \leq x\} \cap \{U=0\}\}$ と $\{Z_2 \leq x\} \cap \{U=1\}\}$ は背反であり, Z_1, Z_2, U は互いに独立なので,

$$\begin{aligned}
\mathrm{P}(X \leq x) &= \mathrm{P}(\{Z_1 \leq x\} \cap \{U=0\}) + \mathrm{P}(\{Z_2 \leq x\} \cap \{U=1\}) \\
&= \mathrm{P}(Z_1 \leq x)\mathrm{P}(U=0) + \mathrm{P}(Z_2 \leq x)\mathrm{P}(U=1) \\
&= (1-\alpha)\int_{-\infty}^{x}\phi(t)dt + \alpha\int_{-\infty}^{x}\frac{1}{\lambda}\phi(t/\lambda)dt \\
&= \int_{-\infty}^{x}\left\{(1-\alpha)\phi(t) + \frac{\alpha}{\lambda}\phi(t/\lambda)\right\}dt.
\end{aligned}$$

これは, $\mathrm{CN}(\lambda,\alpha)$ の分布関数である. □

1.3. 歪正規分布 (Skew Normal Distribution)

定義

$X \sim \mathrm{SN}(\lambda)$ であるとは, 確率変数 X が従う分布の確率密度関数が

$$f(x;\lambda) = 2\phi(x)\Phi(\lambda x), \ (\lambda \in \mathbb{R}), \tag{1.8}$$

であるときのことをいう. また, この確率分布を歪度パラメータ λ の歪正規分布という.

(1.8) が確率密度関数になっていることは, 以下の定理により示すことができる.

定理 1.3.1: $f(x)$ を確率密度関数, $F(x)$ をその分布関数とすると, f が 0 を中心に対称であれば $2f(x)F(\lambda x)$ は確率密度関数となる.

(*Proof*) $2f(x)F(\lambda x) \geq 0$ は明らか. $X, Y \sim i.i.d.\ f(x)$ とすると,

$$\frac{1}{2} = \mathrm{P}(X - \lambda Y \leq 0) = \mathrm{E}_Y[\mathrm{P}(X \leq \lambda y)|Y = y] = \int_{-\infty}^{\infty} F(\lambda y) f(y) dy.$$

よって $\int_{-\infty}^{\infty} 2f(x)F(\lambda x)f(x)dx = 1$ が成り立つ. □

以下の定理は, 歪正規分布の特性の証明のために必要となる.

定理 1.3.2: X と Y を独立な連続型の確率変数とし, X と Y の確率密度関数をぞれぞれ $f(x)$, $g(y)$ とする. このとき, $Z = X + Y$ の確率密度関数 $h(z)$ は,

$$h(z) = \int_{-\infty}^{\infty} f(z - y) g(y) dy, \tag{1.9}$$

となる.

(*Proof*) $V = Y$ とおくと, $(X, Y) \to (Z, V)$ の逆変換は, $X = Z - V$, $Y = V$ となり, この変換に関するヤコビアンは,

$$J = \begin{vmatrix} \partial x/\partial z & \partial x/\partial v \\ \partial y/\partial z & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

よって, Z と V の同時確率密度関数は, $f(z - v)g(v) \cdot 1 = f(z - v)g(v)$ であり, Z の確率密度関数は,

$$h(z) = \int_{-\infty}^{\infty} f(z - v) g(v) dv.$$

これは (1.9) に一致する. □

特性

1. 歪正規分布は非対称な分布であり, $\lambda = 0$ のときは正規分布となる. また, $\lambda \to \pm\infty$ のときは半正規分布となる.

   注意

   半正規分布 (half-normal distribution) とは, 平均 μ の正規分布を μ から折り返した分布である. つまり, $X \sim \mathrm{N}(\mu, \sigma^2)$ のとき, $|X - \mu|$ や $-|X - \mu|$ が従う分布であり, $\mathrm{N}^+(\mu, \sigma^2)$, $\mathrm{N}^-(\mu, \sigma^2)$ と書く. $\mathrm{N}^+(\mu, \sigma^2)$ の確率密度関数は,

$$f(x;\mu,\sigma^2) = \begin{cases} \sqrt{\dfrac{2}{\pi\sigma^2}} \exp\left\{-\dfrac{1}{2\sigma^2}(x-\mu)^2\right\} & (x \geq 0) \\ 0 & (x < 0) \end{cases},$$

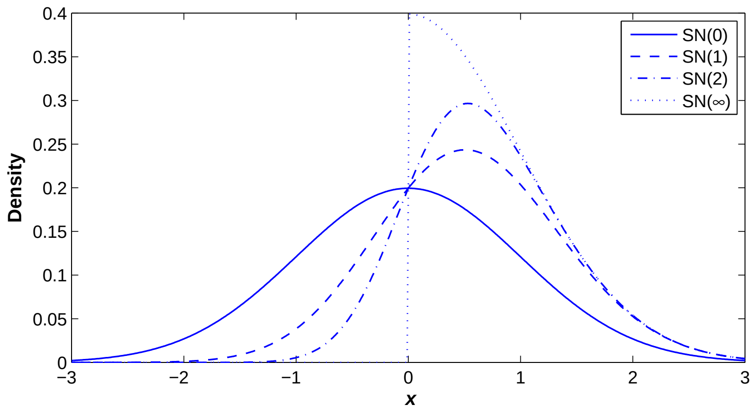

Figure 1.3. 歪正規分布の形

であり, $\mathrm{N}^-(\mu,\sigma^2)$ の確率密度関数は,

$$f(x;\mu,\sigma^2)=\begin{cases}0 & (x>0)\\ \sqrt{\dfrac{2}{\pi\sigma^2}}\exp\left\{-\dfrac{1}{2\sigma^2}(x-\mu)^2\right\} & (x\le 0)\end{cases}.$$

ここで,

$$\lim_{\lambda\to\infty}\Phi(\lambda x)=\begin{cases}1 & (x>0)\\ 1/2 & (x=0)\\ 0 & (x<0)\end{cases},\quad \lim_{\lambda\to-\infty}\Phi(\lambda x)=\begin{cases}0 & (x>0)\\ 1/2 & (x=0)\\ 1 & (x<0)\end{cases},$$

となるので, $\lambda\to\pm\infty$ のとき, 歪正規分布が半正規分布になることがわかる.

2. $X\sim\mathrm{SN}(\lambda)$ のとき $\delta=\lambda/\sqrt{1+\lambda^2}$ とすると,

$$\mathrm{E}[X]=\sqrt{\frac{2}{\pi}}\delta,\quad \mathrm{Var}[X]=1-\frac{2}{\pi}\delta^2,$$
$$\kappa_3=\frac{1}{2}(4-\pi)\mathrm{sign}(\lambda)\left\{\frac{\lambda^2}{\pi/2+(\pi/2-1)\lambda^2}\right\}^{3/2},$$
$$\kappa_4=2(\pi-3)\left\{\frac{\lambda^2}{\pi/2+(\pi/2-1)\lambda^2}\right\}^2,$$

である. ここで $\lambda\to\infty$ とすれば,

$$\lim_{\lambda\to\infty}\kappa_3=\frac{\sqrt{2}(4-\pi)}{(\pi-2)^{3/2}},\quad \lim_{\lambda\to\infty}\kappa_4=\frac{8(\pi-3)}{(\pi-2)^2},$$

となり, λ を大きくしても歪度と尖度はそこまで大きくはならないことがわかる.

3. 乱数発生方法: それぞれ独立な確率変数 Z_1,Z_2 を $Z_1\sim\mathrm{N}(0,1/(1+\lambda^2))$, $Z_2\sim\mathrm{N}(0,\lambda^2/(1+\lambda^2))$ とすると $X\sim\mathrm{SN}(\lambda)$ となる X は以下のように発生させることができる.

$$X=\mathrm{sign}(\lambda)(Z_1+|Z_2|).$$

(*Proof*) $Y=|Z_2|$ とおくと, Y は半正規分布に従い, その確率密度関数は,

$$g(y)=\sqrt{\frac{2}{\pi\tau^2}}\exp\left(-\frac{y^2}{2\tau^2}\right),$$

となる．ただし $\tau^2 = \lambda^2/(1+\lambda^2)$．まず $\lambda > 0$ のときを考える．このとき, 定理 1.3.2 より, X の確率密度関数は,

$$
\begin{aligned}
h(x) &= \int_0^\infty \frac{1}{\sqrt{1-\tau^2}} \phi\left((x-y)/\sqrt{1-\tau^2}\right) g(y)dy \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi(1-\tau^2)}} \exp\left\{-\frac{1}{2(1-\tau^2)}(x-y)^2\right\} \sqrt{\frac{2}{\pi\tau^2}} \exp\left(-\frac{y^2}{2\tau^2}\right) dy.
\end{aligned}
$$

ここで,

$$
\begin{aligned}
&-\frac{1}{2(1-\tau^2)}(x-y)^2 - \frac{y^2}{2\tau^2} = -\frac{1}{2\tau^2(1-\tau^2)}\left\{\tau^2(x-y)^2 + (1-\tau^2)y^2\right\} \\
&= -\frac{1}{2\tau^2(1-\tau^2)}\left\{(y-\tau^2 x)^2 + \tau^2(1-\tau^2)x^2\right\} = -\frac{1}{2\tau^2(1-\tau^2)}(y-\tau^2 x)^2 - \frac{1}{2}x^2,
\end{aligned}
$$

となるので,

$$
h(x) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2}x^2\right) \int_0^\infty \frac{1}{\sqrt{2\pi\tau^2(1-\tau^2)}} \exp\left\{-\frac{1}{2\tau^2(1-\tau^2)}(y-\tau^2 x)^2\right\} dy.
$$

今, $u = (y-\tau^2 x)/\sqrt{\tau^2(1-\tau^2)}$ とおくと, $du/dy = 1/\sqrt{\tau^2(1-\tau^2)}$ であり,

$$
\frac{\tau^2 x}{\sqrt{\tau^2(1-\tau^2)}} = -\sqrt{\frac{\tau^2}{1-\tau^2}}x = -\sqrt{\frac{\lambda^2/(1+\lambda^2)}{1/(1+\lambda^2)}}x = -\lambda x,
$$

となるので, u の積分範囲は $-\lambda x$ から ∞ となる．よって,

$$
h(x) = 2\phi(x) \int_{-\lambda x}^\infty \phi(u)du = 2\phi(x) \int_{-\infty}^{\lambda x} \phi(u)du = 2\phi(x)\Phi(\lambda x).
$$

また, $\lambda < 0$ のとき, $-X \sim \mathrm{SN}(-\lambda)$ であり, また $\phi(-u) = \phi(u)$ であるので,

$$
\mathrm{P}(X \le x) = \mathrm{P}(-X \ge -x) = \int_{-x}^\infty 2\phi(u)\Phi(-\lambda u)du = \int_{-x}^\infty 2\phi(-u)\Phi(-\lambda u)du
$$

ここで, $-u = t$ とおくと,

$$
\mathrm{P}(X \le x) = \int_{-\infty}^x 2\phi(t)\Phi(\lambda t)dt.
$$

これは $\mathrm{SN}(\lambda)$ の分布関数である．以上により X の確率密度関数が $2\phi(x)\Phi(\lambda x)$ であることが言える．
□

2. 多次元確率分布と多次元正規分布

2.1. 多次元確率分布

- **分布関数**

<u>**定義**</u>

p 次元確率変数ベクトル $\boldsymbol{X} = (X_1, \ldots, X_p)'$ の同時確率密度関数 (単に確率密度関数と呼ぶこともある) を $f(x_1, \ldots, x_p)$ とする. このとき, $\boldsymbol{X}$ の分布関数は以下のように定義される.

$$\mathrm{P}(X_1 \leq x_1, \ldots, X_p \leq x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) du_1 \cdots du_p.$$

- **期待値と分散**

<u>**定義**</u>

$f(x_1, \ldots, x_p)$ を確率変数ベクトル $\boldsymbol{X} = (X_1, \ldots, X_p)'$ の確率密度関数とすると, $\boldsymbol{X}$ の期待値は以下のように定義される.

$$\mathrm{E}[\boldsymbol{X}] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \boldsymbol{x} f(x_1, \ldots, x_p) dx_1 \cdots dx_p.$$

ただし $\boldsymbol{x} = (x_1, \ldots, x_p)'$. また, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)' = \mathrm{E}[\boldsymbol{X}]$ とすると, $\boldsymbol{X}$ の分散は以下のように定義される.

$$\mathrm{Var}[\boldsymbol{X}] = \mathrm{E}[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})' f(x_1, \ldots, x_p) dx_1 \cdots dx_p.$$

$\boldsymbol{X}$ の期待値は, $\boldsymbol{X}$ の成分ごとの期待値を並べたもの, 分散も行列 $(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'$ のそれぞれの成分ごとの期待値を並べたものであることがわかる. 行列 $\mathrm{Var}[\boldsymbol{X}]$ の第 (j, j) 成分は X_j の分散, 第 (i, j) 成分は X_i と X_j の**共分散** (**Covariance**) となっているので, $\mathrm{Var}[\boldsymbol{X}]$ は**分散共分散行列** (**Variance Covariance Matrix**) と呼ばれる (単に**共分散行列** (**Covariance Matrix**) と呼ばれることもある). $\mathrm{Var}[\boldsymbol{X}]$ は多くの場合, Σ と表記される. また特に, X_i と X_j の**相関** (**Correlation**),

$$\rho_{ij} = \frac{\mathrm{E}[(X_i - \mu_i)(X_j - \mu_j)]}{\sqrt{\mathrm{Var}[X_i]\mathrm{Var}[X_j]}},$$

を成分に持つ $p \times p$ 行列 Φ は**相関行列** (**Correlation Matrix**) と呼ばれる. ここで, $\mathrm{Var}[X_j] = \sigma_j^2$ とおくと, 分散共分散行列 $\mathrm{Var}[X] = \Sigma$ と相関行列 Φ の間には以下のような関係がある.

$$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p)\Phi\mathrm{diag}(\sigma_1, \ldots, \sigma_p).$$

- **特性関数**

<u>**定義**</u>

p 次元確率変数ベクトル $\boldsymbol{X}$ の確率密度関数を $f(\boldsymbol{x})$ とする. ${}^\forall \boldsymbol{t} \in \mathbb{R}^p$ に対して, $\mathrm{E}[\exp(i\boldsymbol{t}'\boldsymbol{X})]$ を $\boldsymbol{X}$ の特性関数といい, $C(\boldsymbol{t})$ で表わす. つまり

$$C(\boldsymbol{t}) = \mathrm{E}[\exp(i\boldsymbol{t}'\boldsymbol{X})] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(i\boldsymbol{t}'\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}, \ (\boldsymbol{t} \in \mathbb{R}^p).$$

${}^\forall \boldsymbol{\ell} \in \mathbb{R}^p$ に関して $Y = \boldsymbol{\ell}'\boldsymbol{X}$ とし, Y の特性関数を $C_Y(t)$ とおくと,

$$C_Y(t) = \mathrm{E}[\exp(itY)] = \mathrm{E}[\exp(it\boldsymbol{\ell}'\boldsymbol{X})], \tag{2.1}$$

となる. (2.1) 式において, $t\boldsymbol{\ell} = \boldsymbol{t}$ と置き換えると, $\boldsymbol{X}$ の特性関数となることがわかる. 以上により, 特性関数の一意性により Y の分布から $\boldsymbol{X}$ の分布が分かることになる.

特性関数は以下の様な特性を持つ.

(a) 一意性定理: 確率変数ベクトル $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ の特性関数を $C_1(\boldsymbol{t})$, $C_2(\boldsymbol{t})$ とすると,

$$C_1(\boldsymbol{t}) = C_2(\boldsymbol{t}) \Leftrightarrow \boldsymbol{X}_1 \text{ と } \boldsymbol{X}_2 \text{ の分布は同じ.}$$

(b) モーメント公式: $n = n_1 + \cdots + n_p$ とすると,

$$\mathrm{E}[\|\boldsymbol{X}\|^n] < \infty \Rightarrow \left.\frac{\partial^n}{\partial t_1^{n_1} \cdots \partial t_p^{n_p}}\right|_{\boldsymbol{t}=\boldsymbol{0}} C(\boldsymbol{t}) = i^n \mathrm{E}[X_1^{n_1} \cdots X_p^{n_p}].$$

ここでの条件 $\mathrm{E}[\|\boldsymbol{X}\|^n] < \infty$ は微分と積分の順序を入れ替えるために必要である.

- **行列について**

1. **正定値行列 (Positive Definite Matrix)**:

   定義

   $p \times p$ 行列 $M = (m_{ij})$ が正定値行列 ($M > 0$ と書く) とは以下の条件を満たすことである.

$$M > 0 \Leftrightarrow \begin{cases} M' = M \\ \sum_{i,j}^{p} m_{ij} x_i x_j = \boldsymbol{x}' M \boldsymbol{x} > 0 \ ({}^{\forall}\boldsymbol{x} \in \mathbb{R}^p \backslash \{\boldsymbol{0}_p\}) \end{cases}.$$

   ただし $\boldsymbol{x} = (x_1, \ldots, x_p)'$ である.

2. **半正定値行列 (Positive Semi-Definite Matrix)**:

   定義

   $p \times p$ 行列 $M = (m_{ij})$ が半正定値行列 ($M \geq 0$ と書く) とは以下の条件を満たすことである.

$$M \geq 0 \Leftrightarrow \begin{cases} M' = M \\ \sum_{i,j}^{p} m_{ij} x_i x_j = \boldsymbol{x}' M \boldsymbol{x} \geq 0 \ ({}^{\forall}\boldsymbol{x} \in \mathbb{R}^p \backslash \{\boldsymbol{0}_p\}) \end{cases}.$$

   ただし $\boldsymbol{x} = (x_1, \ldots, x_p)'$ である.

   注意

   一般に分散共分散行列 $\Sigma = \mathrm{Var}[\boldsymbol{X}]$ は半正定値行列である. なぜならば,

$$\Sigma' = \{\mathrm{E}[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})']\}' = \mathrm{E}[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'] = \Sigma,$$
$$\boldsymbol{x}'\Sigma\boldsymbol{x} = \mathrm{E}[\boldsymbol{x}'(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'\boldsymbol{x}] = \mathrm{E}[\{\boldsymbol{x}'(\boldsymbol{X} - \boldsymbol{\mu})\}^2] \geq 0.$$

   しかしながら, 特に断らない限り, Σ は正定値行列であるとする. 同様に, 相関行列も正定値行列であるとする.

3. **スペクトル分解 (Spectral Decomposition)**:

   p 次対称行列 M はある p 次直交行列 $H = (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_p)$ $(H'H = HH' = I_p)$ を用いて

$$M = H\Lambda H', \tag{2.2}$$

   と分解できる. ただし

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p),$$

であり $\lambda_1 \geq \cdots \geq \lambda_p$ である. 今 (2.2) 式の右から H をかけると

$$MH = H\Lambda \Leftrightarrow M\boldsymbol{h}_j = \lambda_j \boldsymbol{h}_j \ (j = 1, \ldots, p), \tag{2.3}$$

となり, λ_j は M の固有値, $\boldsymbol{h}_j$ は固有値 λ_j に対する固有ベクトルであることがわかる. ここで, $M > 0$ とすると $\lambda_p > 0$ となり, 正則行列 $A = H\Lambda^{1/2}$ を用いて M は $M = AA'$ と分解できる. また, (2.3) 式より,

$$M = H\Lambda H' = (\lambda_1 \boldsymbol{h}_1, \ldots, \lambda \boldsymbol{h}_p)(\boldsymbol{h}_1, \ldots, \boldsymbol{h}_p)' = \sum_{j=1}^{p} \lambda_j \boldsymbol{h}_j \boldsymbol{h}_j'. \tag{2.4}$$

4. 正定値行列の行列式と逆行列:

定理 2.1.1. p 次正定値行列 M を

$$M = \begin{array}{c c} & \begin{array}{cc} p_1 & p_2 \end{array} \\ \begin{array}{c} p_1 \\ p_2 \end{array} & \begin{pmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{pmatrix} \end{array},$$

と分割したとき, M の行列式と逆行列は以下となる.

$$|M| = |M_{22}||M_{11\cdot 2}|, \quad M^{-1} = \begin{pmatrix} O & O \\ O & M_{22}^{-1} \end{pmatrix} + \begin{pmatrix} I_{p_1} \\ -M_{22}^{-1}M_{12}' \end{pmatrix} M_{11\cdot 2}^{-1} \begin{pmatrix} I_{p_1} \\ -M_{22}^{-1}M_{12}' \end{pmatrix}'. \tag{2.5}$$

ただし, $M_{11\cdot 2} = M_{11} - M_{12}M_{22}^{-1}M_{12}'$ である.

(*Proof*) 今, 行列 B を,

$$B = \begin{pmatrix} I_{p_1} & -M_{12}M_{22}^{-1} \\ O & I_{p_2} \end{pmatrix},$$

とおくと,

$$BMB' = \begin{pmatrix} M_{11\cdot 2} & O \\ O & M_{22} \end{pmatrix},$$

である. ここで, ブロック三角行列の行列式の公式から, $|B| = |I_{p_1}||I_{p_2}| = 1$, $|BMB'| = |M_{11\cdot 2}||M_{22}|$ がわかり, $|BMB'| = |B||M||B'| = |B||M||B| = |M|$ であるので, $|M| = |M_{11\cdot 2}||M_{22}|$ が成り立つ. また,

$$\begin{pmatrix} M_{11\cdot 2} & O \\ O & M_{22} \end{pmatrix}^{-1} = \begin{pmatrix} M_{11\cdot 2}^{-1} & O \\ O & M_{22}^{-1} \end{pmatrix},$$

であり, $(BMB')^{-1} = (B')^{-1}M^{-1}B^{-1}$ であるので,

$$\begin{aligned} M^{-1} = B'(BMB')^{-1}B &= \begin{pmatrix} I_{p_1} & O \\ -M_{22}^{-1}M_{12}' & I_{p_2} \end{pmatrix} \begin{pmatrix} M_{11\cdot 2}^{-1} & O \\ O & M_{22}^{-1} \end{pmatrix} \begin{pmatrix} I_{p_1} & -M_{12}M_{22}^{-1} \\ O & I_{p_2} \end{pmatrix} \\ &= \begin{pmatrix} I_{p_1} & O \\ -M_{22}^{-1}M_{12}' & I_{p_2} \end{pmatrix} \left\{ \begin{pmatrix} O & O \\ O & M_{22}^{-1} \end{pmatrix} + \begin{pmatrix} M_{11\cdot 2}^{-1} & O \\ O & O \end{pmatrix} \right\} \begin{pmatrix} I_{p_1} & -M_{12}M_{22}^{-1} \\ O & I_{p_2} \end{pmatrix} \\ &= \begin{pmatrix} I_{p_1} & O \\ -M_{22}^{-1}M_{12}' & I_{p_2} \end{pmatrix} \left\{ \begin{pmatrix} O & O \\ O & M_{22}^{-1} \end{pmatrix} + \begin{pmatrix} I_{p_1} \\ O \end{pmatrix} M_{11\cdot 2}^{-1} \begin{pmatrix} I_{p_1} \\ O \end{pmatrix}' \right\} \begin{pmatrix} I_{p_1} & -M_{12}M_{22}^{-1} \\ O & I_{p_2} \end{pmatrix} \\ &= \begin{pmatrix} O & O \\ O & M_{22}^{-1} \end{pmatrix} + \begin{pmatrix} I_{p_1} \\ -M_{22}^{-1}M_{12}' \end{pmatrix} M_{11\cdot 2}^{-1} \begin{pmatrix} I_{p_1} \\ -M_{22}^{-1}M_{12}' \end{pmatrix}'. \end{aligned}$$

以上により (2.5) 式を得ることができる. □

2.2. 多次元正規分布 (Multivariate Normal Distribution)

定義 1

p 次元確率変数ベクトル $\boldsymbol{X} = (X_1, \ldots, X_p)'$ に対し, $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ であるとは, $\boldsymbol{X}$ の従う分布の確率密度関数が

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma) = \left(\frac{1}{2\pi}\right)^{p/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \quad (\boldsymbol{\mu} \in \mathbb{R}^p, \Sigma > 0), \tag{2.6}$$

であるときのことをいう. ただし, $\boldsymbol{x} = (x_1, \ldots, x_p)'$. また, この確率分布を平均ベクトル $\boldsymbol{\mu}$, 分散共分散行列 Σ の p 次元正規分布という.

$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$, $\Sigma = (\sigma_{ij})$ とおく. このとき, $p = 1$ のとき, $\Sigma = \sigma_{11} = \sigma^2$, $\mu_1 = \mu$, $x_1 = x$ とおけば, $|\Sigma|^{-1/2} = 1/\sigma$, $(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = (x-\mu)^2/\sigma^2$ となるので, (2.6) 式は 1 次元正規分布の確率密度関数となる. よって, 多次元正規分布は 1 次元正規分布を拡張した分布であることがわかる.

(2.6) 式が確率密度関数であることを確かめるためには, $f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma) \geq 0$ でかつ $f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma)$ を全区間で積分して 1 となることを確かめればよい. $f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma) \geq 0$ は明らかであるので, 全区間で積分して 1 になることを示せばよいことになる. 今, $\Sigma > 0$ なので, スペクトル分解により $\Sigma = AA'$ と分解でき, $\Sigma^{-1} = (A')^{-1}A^{-1}$ である. ここで, 変数変換 $\boldsymbol{z} = A^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ 行うと $\boldsymbol{x} = \boldsymbol{\mu} + A\boldsymbol{z}$ となる. このとき, $\boldsymbol{x} \in \mathbb{R}^p \Leftrightarrow \boldsymbol{z} \in \mathbb{R}^p$ であり, 変数変換に関するヤコビアンは

$$J = \begin{vmatrix} \partial x_1/\partial z_1 & \cdots & \partial x_1/\partial z_p \\ \vdots & \ddots & \vdots \\ \partial x_p/\partial z_1 & \cdots & \partial x_p/\partial z_p \end{vmatrix} = |A|,$$

である. また,

$$|\Sigma| = |AA'| = |A||A'| = |A|^2,$$

となるので $|A| = |\Sigma|^{1/2}$ となる. よって求めたい積分は変数変換を使って,

$$\begin{aligned}
&\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\right)^{p/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} dx_1 \cdots dx_p \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\right)^{p/2} |\Sigma|^{-1/2} e^{-\boldsymbol{z}'\boldsymbol{z}/2} |\Sigma|^{1/2} dz_1 \cdots dz_p \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\right)^{p/2} e^{-(z_1^2+\cdots+z_p^2)/2} dz_1 \cdots dz_p \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\right)^{p/2} \prod_{j=1}^{p} e^{-z_j^2/2} dz_1 \cdots dz_p = \prod_{j=1}^{p} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z_j^2/2} dz_j = 1.
\end{aligned}$$

以上により (2.6) 式が確率密度関数であることをわかる.

ここで, $\boldsymbol{Z} = A^{-1}(\boldsymbol{X}-\boldsymbol{\mu})$ という変換を考えると, $\boldsymbol{Z}$ の確率密度関数は

$$g(\boldsymbol{z}) = f(\boldsymbol{\mu} + A\boldsymbol{z}; \boldsymbol{\mu}, \Sigma)|\Sigma|^{1/2} = \left(\frac{1}{2\pi}\right)^{p/2} e^{-\boldsymbol{z}'\boldsymbol{z}/2} = \prod_{j=1}^{p} \frac{1}{\sqrt{2\pi}} e^{-z_j^2/2},$$

となる. これは

$$\boldsymbol{Z} \sim \mathrm{N}_p(\boldsymbol{0}, I_p) \Leftrightarrow Z_1, \ldots, Z_p \sim i.i.d.\ \mathrm{N}(0, 1), \tag{2.7}$$

であることを示し, また $\boldsymbol{Z} \sim \mathrm{N}_p(\boldsymbol{0}, I_p)$ である $\boldsymbol{Z}$ に対して,

$$\boldsymbol{X} = \boldsymbol{\mu} + A\boldsymbol{Z} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)\ (\Sigma = AA'),$$

であることも示している．つまり $\boldsymbol{X}$ は $\boldsymbol{Z}$ の線形変換によって表すことができる．

先の定義 1 以外に以下の様な多次元正規分布の定義もある．

定義 2

p 次元確率変数ベクトル $\boldsymbol{X} = (X_1, \ldots, X_p)'$ に対し, $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ であるとは, $\mathrm{E}[\boldsymbol{X}] = \mu$, $\mathrm{Var}[\boldsymbol{X}] = \Sigma$ で, ${}^{\forall}\boldsymbol{\ell} \in \mathbb{R}^p$ に対し, $\boldsymbol{\ell}'\boldsymbol{X}$ が 1 次元正規分布に従うときのことをいう．

今, $\boldsymbol{X}$ を定義 1 で定義された多次元正規分布に従う確率変数ベクトルとする．このとき $\boldsymbol{Z} \sim \mathrm{N}_p(\boldsymbol{0}, I_p)$ となる $\boldsymbol{Z}$ を用いて $\boldsymbol{X} = \boldsymbol{\mu} + A\boldsymbol{Z}$ と書ける．この式の両辺に $\boldsymbol{\ell}'$ を左からかけると

$$\boldsymbol{\ell}'\boldsymbol{X} = \boldsymbol{\ell}'\boldsymbol{\mu} + \boldsymbol{\ell}'A\boldsymbol{Z} = c_0 + \boldsymbol{c}'\boldsymbol{Z}, \ (c_0 = \boldsymbol{\ell}'\boldsymbol{\mu}, \boldsymbol{c} = (c_1, \ldots, c_p)' = A'\boldsymbol{\ell}),$$

となり, 正規分布の再生性により, $\boldsymbol{\ell}'\boldsymbol{X} \sim \mathrm{N}(c_0, \boldsymbol{c}'\boldsymbol{c})$ であることがわかる．よって, 定義 1$\Rightarrow$ 定義 2 であることがわかる．また, 上記に式において A が非正則行列であれば, 定義 1 を満たさないが定義 2 を満たすことになる．よって定義 1$\subset$ 定義 2 であるので, 定義 2 で定義した多次元正規分布の方が定義 1 で定義した多次元正規分布よりも広いクラスとなることがわかる．

特性

1. 多次元正規分布は $\boldsymbol{\mu}$ を中心に左右対称の分布であり, $\boldsymbol{\mu}$ は中心の位置, Σ はばらつきと変量ごとの関係の強さを表している．

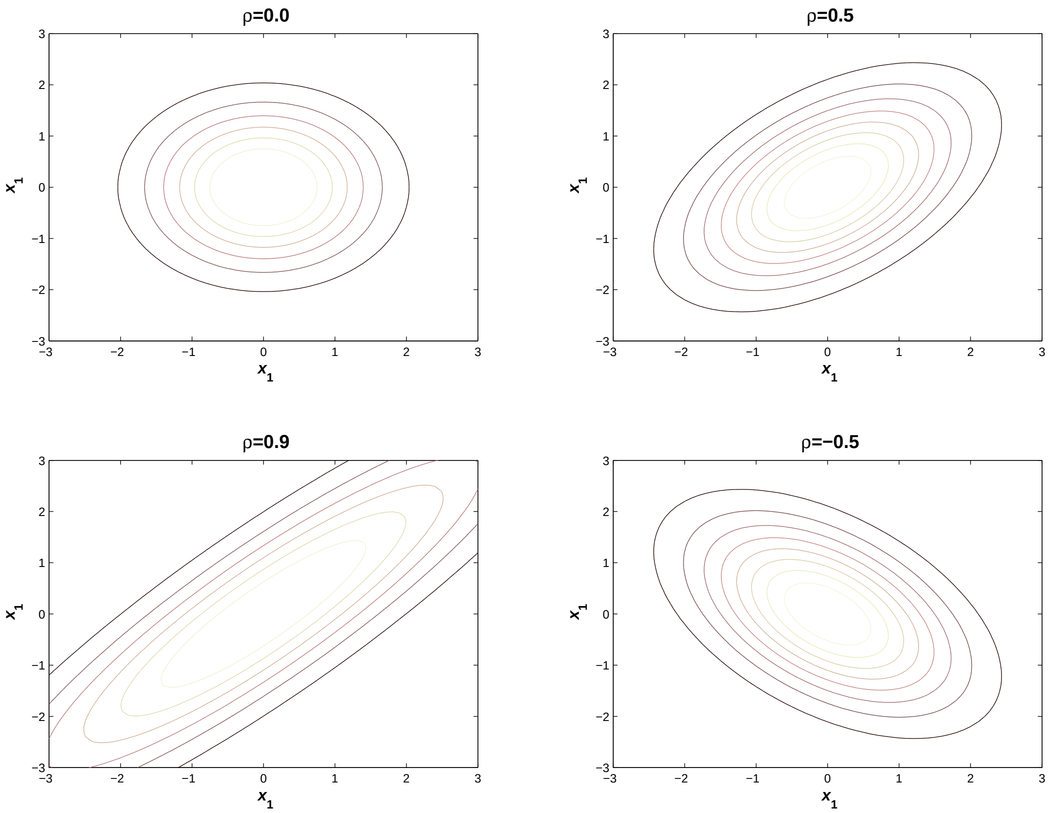

Figure 2.1. 2 次元正規分布を上から見た図
(ただしそれぞれの変量の分散は 1. また図中の線は等高線を表す)

2. $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ のとき, $\mathrm{E}[\boldsymbol{X}] = \boldsymbol{\mu}$, $\mathrm{Var}[\boldsymbol{X}] = \Sigma$.

3. $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ のとき, $\boldsymbol{X}$ の特性関数は,

$$C(\boldsymbol{t}) = \exp\left(i\boldsymbol{t}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'\Sigma\boldsymbol{t}\right),\ (\boldsymbol{t} \in \mathbb{R}^p). \tag{2.8}$$

(*Proof*) $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ のとき, ${}^\forall\boldsymbol{\ell} \in \mathbb{R}^p$ に対して $Y = \boldsymbol{\ell}'\boldsymbol{X}$ とおくと, $Y \sim \mathrm{N}(\boldsymbol{\ell}'\boldsymbol{\mu}, \boldsymbol{\ell}'\Sigma\boldsymbol{\ell})$ である. よって, Y の特性関数は, (1.4) 式より,

$$C_Y(t) = \exp\left(it\boldsymbol{\ell}'\boldsymbol{\mu} - \frac{1}{2}t^2\boldsymbol{\ell}'\Sigma\boldsymbol{\ell}\right),$$

となる. ここで (2.1) 式から, $t\boldsymbol{\ell} = \boldsymbol{t}$ とおけば $\boldsymbol{X}$ の特性関数になり, その式は (2.8) となる. □

4. $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ のとき, $\boldsymbol{b}$ を q $(\leq p)$ 次元ベクトル, B を $q \times p$ 行列で $\mathrm{rank}(B) = q$ ($\Leftrightarrow$ BB' が正則) とし, $\boldsymbol{Y} = \boldsymbol{b} + B\boldsymbol{X}$ とする. このとき,

$$\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma) \Rightarrow \boldsymbol{Y} \sim \mathrm{N}_q(\boldsymbol{b} + B\boldsymbol{\mu}, B\Sigma B'). \tag{2.9}$$

(*Proof*) $\boldsymbol{Y}$ の特性関数を $C_{\boldsymbol{Y}}(\boldsymbol{t})$ $(\boldsymbol{t} \in \mathbb{R}^q)$, $\boldsymbol{X}$ の特性関数を $C_{\boldsymbol{X}}(\boldsymbol{s})$ $(\boldsymbol{s} \in \mathbb{R}^p)$ とすると,

$$C_{\boldsymbol{Y}}(\boldsymbol{t}) = \mathrm{E}[\exp(i\boldsymbol{t}'\boldsymbol{Y})] = \mathrm{E}[\exp\{i\boldsymbol{t}'(\boldsymbol{b} + B\boldsymbol{X})\}] = \exp(i\boldsymbol{t}'\boldsymbol{b})\mathrm{E}[\exp(i\boldsymbol{t}'B\boldsymbol{X})] = \exp(i\boldsymbol{t}'\boldsymbol{b})C_{\boldsymbol{X}}(B'\boldsymbol{t}).$$

ここで, $C_{\boldsymbol{X}}(\boldsymbol{s})$ は, (2.8) 式より, $C_{\boldsymbol{X}}(\boldsymbol{s}) = \exp(i\boldsymbol{s}'\boldsymbol{\mu} - \boldsymbol{s}'\Sigma\boldsymbol{s}/2)$ なので, $\boldsymbol{s} = B'\boldsymbol{t}$ を代入すれば,

$$C_{\boldsymbol{X}}(B'\boldsymbol{t}) = \exp\left(i\boldsymbol{t}'B\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}'B\Sigma B'\boldsymbol{t}\right),$$

が得られる. これを用いると,

$$C_{\boldsymbol{Y}}(\boldsymbol{t}) = \exp\left\{i\boldsymbol{t}'(\boldsymbol{b} + B\boldsymbol{\mu}) - \frac{1}{2}\boldsymbol{t}'B\Sigma B'\boldsymbol{t}\right\},$$

となり, これは $\mathrm{N}_q(\boldsymbol{b} + B\boldsymbol{\mu}, B\Sigma B')$ の特性関数である. よって, 特性関数の一意性により, $\boldsymbol{Y} \sim \mathrm{N}_q(\boldsymbol{b} + B\boldsymbol{\mu}, B\Sigma B')$ となることがわかる. □

注意

この特性により, 以下が言える.

$X_i \sim \mathrm{N}(\mu_i, \sigma^2)$ $(i = 1, \ldots, p)$ である互いに独立な確率変数 $X_1, \ldots, X_p$ に対し, p 次直交行列 $H = (h_{ij})$ を用いて, $(Y_1, \ldots, Y_p)' = H(X_1, \ldots, X_p)'$ と定める. このとき, $Y_1, \ldots, Y_p$ も互いに独立で, $Y_i \sim \mathrm{N}(\eta_i, \sigma^2)$ となる. ただし, $\eta_i = \sum_{j=1}^p h_{ij}\mu_j$.

5. 再生性: 互いに独立な p 次元確率変数ベクトル $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ を, $\boldsymbol{X}_1 \sim \mathrm{N}_p(\boldsymbol{\mu}_1, \Sigma_1)$, $\boldsymbol{X}_2 \sim \mathrm{N}_p(\boldsymbol{\mu}_2, \Sigma_2)$ とする. このとき, $\boldsymbol{X}_1 + \boldsymbol{X}_2 \sim \mathrm{N}_p(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2, \Sigma_1 + \Sigma_2)$ となる.

   $\boldsymbol{Z} = \boldsymbol{X}_1 + \boldsymbol{X}_2$ の特性関数は,

$$C_{\boldsymbol{Z}}(\boldsymbol{t}) = \mathrm{E}\left[\exp(i\boldsymbol{t}'\boldsymbol{Z})\right] = \mathrm{E}\left[\exp\{i\boldsymbol{t}'(\boldsymbol{X}_1 + \boldsymbol{X}_2)\}\right] = \mathrm{E}\left[\exp(i\boldsymbol{t}'\boldsymbol{X}_1)\right]\mathrm{E}\left[\exp(i\boldsymbol{t}'\boldsymbol{X}_2)\right] = C_{\boldsymbol{X}_1}(\boldsymbol{t})C_{\boldsymbol{X}_2}(\boldsymbol{t}).$$

   ここで, 特性 3 より, $C_{\boldsymbol{X}_i}(\boldsymbol{t}) = \exp(i\boldsymbol{t}'\boldsymbol{\mu}_i - \boldsymbol{t}'\Sigma_i\boldsymbol{t}/2)$ であるので,

$$C_{\boldsymbol{Z}}(\boldsymbol{t}) = \exp\left(i\boldsymbol{t}'\boldsymbol{\mu}_1 - \frac{1}{2}\boldsymbol{t}'\Sigma\boldsymbol{t}\right)\exp\left(i\boldsymbol{t}'\boldsymbol{\mu}_2 - \frac{1}{2}\boldsymbol{t}'\Sigma\boldsymbol{t}\right) = \exp\left\{i\boldsymbol{t}'(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) - \frac{1}{2}\boldsymbol{t}'(\Sigma_1 + \Sigma_2)\boldsymbol{t}\right\}.$$

   これは, $\mathrm{N}_p(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2, \Sigma_1 + \Sigma_2)$ の特性関数である. □

6. 条件付き分布: $\boldsymbol{X} \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ のとき, $\boldsymbol{X}$, $\boldsymbol{\mu}$, Σ を,

$$\boldsymbol{X} = \begin{array}{c} p_1 \\ p_2 \end{array}\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix},\ \boldsymbol{\mu} = \begin{array}{c} p_1 \\ p_2 \end{array}\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix},\ \Sigma = \begin{array}{cc} & \begin{array}{cc} p_1 & p_2 \end{array} \\ \begin{array}{c} p_1 \\ p_2 \end{array} & \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix} \end{array},$$

と分割する. このとき, $\boldsymbol{X}_2 = \boldsymbol{x}_2$ を与えたときの $\boldsymbol{X}_1$ の条件付き分布は,

$$\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{x}_2 \sim \mathrm{N}_{p_1}(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11\cdot 2}),\ (\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'),$$

となり, $\boldsymbol{X}_2$ の分布は $\boldsymbol{X}_2 \sim \mathrm{N}_{p_2}(\boldsymbol{\mu}_2, \Sigma_{22})$ となる.

(*Proof*) $\boldsymbol{X}_2$ の分布は (2.9) 式において, $B = (O, I_{p_2})$, $\boldsymbol{b} = \boldsymbol{0}$ とおけば $B\boldsymbol{\mu} = \boldsymbol{\mu}_2$, $B\Sigma B' = \Sigma_{22}$ であるので, $\boldsymbol{X}_2 \sim \mathrm{N}_{p_2}(\boldsymbol{\mu}_2, \Sigma_{22})$ となることがわかる. また, $\boldsymbol{X}_1$ の条件付き分布は

$$\begin{aligned} f_{\boldsymbol{X}_1|\boldsymbol{X}_2=\boldsymbol{x}_2}(\boldsymbol{x}_1) &= \frac{f_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma)}{f_{\boldsymbol{X}_2}(\boldsymbol{x}_2; \boldsymbol{\mu}_2, \Sigma_{22})} = \frac{(2\pi)^{-p/2}|\Sigma|^{-1/2}\exp\{-(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})/2\}}{(2\pi)^{-p_2/2}|\Sigma_{22}|^{-1/2}\exp\{-(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)/2\}} \\ &= \left(\frac{1}{2\pi}\right)^{p_1/2} \frac{|\Sigma|^{-1/2}\exp\{-(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})/2\}}{|\Sigma_{22}|^{-1/2}\exp\{-(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)/2\}}, \end{aligned}$$

である. ここで, 定理 2.1.1 より,

$$|\Sigma| = |\Sigma_{22}||\Sigma_{11\cdot 2}|, \quad \Sigma^{-1} = \begin{pmatrix} O & O \\ O & \Sigma_{22}^{-1} \end{pmatrix} + \begin{pmatrix} I_{p_1} \\ -\Sigma_{22}^{-1}\Sigma_{12}' \end{pmatrix} \Sigma_{11\cdot 2}^{-1} \begin{pmatrix} I_{p_1} \\ -\Sigma_{22}^{-1}\Sigma_{12}' \end{pmatrix}', \tag{2.10}$$

であり, また,

$$\begin{aligned} &(\boldsymbol{x} - \boldsymbol{\mu})' \begin{pmatrix} O & O \\ O & \Sigma_{22}^{-1} \end{pmatrix} (\boldsymbol{x} - \boldsymbol{\mu}) = (\boldsymbol{x}_2 - \boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \\ &\begin{pmatrix} I_{p_1} \\ -\Sigma_{12}\Sigma_{22}^{-1} \end{pmatrix}' (\boldsymbol{x} - \boldsymbol{\mu}) = \boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \end{aligned}$$

なので, (2.10) を用いると,

$$\begin{aligned} &|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right\} \\ &= |\Sigma_{22}|^{-1/2}|\Sigma_{11\cdot 2}|^{-1/2}\exp\left\{-\frac{1}{2}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)\right\} \\ &\times \exp\left(-\frac{1}{2}\{\boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)\}'\Sigma_{22}^{-1}\{\boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)\}\right), \end{aligned}$$

となることがわかる. 上記の式を代入すれば,

$$\begin{aligned} f_{\boldsymbol{X}_1|\boldsymbol{X}_2=\boldsymbol{x}_2}(\boldsymbol{x}_1) = &\left(\frac{1}{2\pi}\right)^{p_1/2} |\Sigma_{11\cdot 2}|^{-1/2} \\ &\times \exp\left(-\frac{1}{2}\{\boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)\}'\Sigma_{22}^{-1}\{\boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)\}\right), \end{aligned}$$

を得ることができる. これは $\mathrm{N}_{p_1}(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11\cdot 2})$ の確率密度関数であるので, $\boldsymbol{X}_1$ の条件付き分布が $\mathrm{N}_{p_1}(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11\cdot 2})$ となることがわかる. $\square$

3. χ^2 分布とその関連分布

3.1. χ^2 分布 (Chi-Square Distribution)

定義 1

非負値をとる確率変数 Y に対し, $Y \sim \chi^2_n$ であるとは, Y の従う分布の確率密度関数が,

$$f(y;n) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1} & (y > 0) \\ 0 & (y \leq 0) \end{cases} \qquad (n \in \mathbb{N}), \tag{3.1}$$

であるときのことをいう. また, この確率分布を自由度 n の χ^2 分布という.

(3.1) 式が確率密度関数であることを確かめるためには, $f(y;n) \geq 0$ でかつ $f(y;n)$ を全区間で積分して 1 となることを確かめればよい. $f(y;n) \geq 0$ は明らかであるので, 全区間で積分して 1 になることを示せばよいことになる. 今, (3.1) の右部分の積分を考えると,

$$\int_0^\infty e^{-y/2} y^{n/2-1} dy = 2^{n/2-1} \int_0^\infty e^{-y/2} (y/2)^{n/2-1} dy,$$

である. ここで, $y/2 = x$ とおけば, $dx/dy = 1/2$ であるので, ガンマ関数の定義 $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ を用いると,

$$\int_0^\infty e^{-y/2} (y/2)^{n/2-1} dy = 2 \int_0^\infty e^{-x} x^{n/2-1} dx = 2\Gamma(n/2).$$

以上のことから,

$$\int_0^\infty f(y;n) dy = \int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1} dy = \frac{1}{2^{n/2}\Gamma(n/2)} (2^{n/2-1}) \cdot 2 \cdot \Gamma(n/2) = 1,$$

となり, (3.1) 式が確率密度関数となることがわかる.

χ^2 分布の定義として, 以下のような定義を用いているテキストも多い.

定義 2

$Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(0,1)$ とし, $Y = \sum_{j=1}^n Z_j^2$ とおく. このとき, Y の従う分布を自由度 n の χ^2 分布といい, $Y \sim \chi^2_n$ と書く.

注意

1. 定義 2 での Y の確率密度関数は, もちろん (3.1) となる.

   (*Proof*) Y の確率密度関数を $f(y)$, 分布関数を $F(y)$ とすると,

   $$f(y) = \frac{d}{dy} F(y) = \lim_{h \to 0} \frac{F(y+h) - F(y)}{h} = \lim_{h \to 0} \frac{1}{h} \mathrm{P}(y < Y \leq y+h),$$

   である. ここで,

   $$\mathrm{P}(y < Y \leq y+h) = \mathrm{P}(y < Z_1^2 + \cdots + Z_n^2 \leq y+h),$$

   であり,

   $$\mathrm{P}((Z_1, \ldots, Z_n) \in D) = \int_D f(z_1, \ldots, z_n) dz_1 \cdots dz_n,$$

   であるので,

   $$\mathrm{P}(y < Y \leq y+h) = \int_{y < z_1^2 + \cdots + z_n^2 < y+h} \left(\frac{1}{2\pi}\right)^{n/2} \exp\left\{ -\frac{1}{2}(z_1^2 + \cdots + z_n^2) \right\} dz_1 \cdots dz_n,$$

となる. よって,

$$\begin{aligned}
&\left(\frac{1}{2\pi}\right)^{n/2} e^{-(y+h)/2} \int_{y<z_1^2+\cdots+z_n^2<y+h} dz_1 \cdots dz_n \\
&\leq \mathrm{P}(y < Y \leq y+h) \\
&\leq \left(\frac{1}{2\pi}\right)^{n/2} e^{-y/2} \int_{y<z_1^2+\cdots+z_n^2<y+h} dz_1 \cdots dz_n,
\end{aligned}$$

が成り立つ. ここで, 積分,

$$\int_{y<z_1^2+\cdots+z_n^2<y+h} dz_1 \cdots dz_n,$$

は二つの同心球面に囲まれた部分の体積であり, 半径 r の球の体積は $\pi^{n/2}r^n/\Gamma(n/2+1)$ であるので,

$$\begin{aligned}
\int_{y<z_1^2+\cdots+z_n^2<y+h} dz_1 \cdots dz_n &= \frac{\pi^{n/2}}{\Gamma(n/2+1)} \left\{(y+h)^{n/2} - y^{n/2}\right\} \\
&= \frac{1}{\Gamma(n/2)2^{n/2}} \left(\frac{2}{n}\right) (2\pi)^{n/2} \left\{(y+h)^{n/2} - y^{n/2}\right\},
\end{aligned}$$

となる. これを用いると,

$$\frac{e^{-(y+h)/2}}{\Gamma(n/2)2^{n/2}} \left(\frac{2}{n}\right) \left\{(y+h)^{n/2} - y^{n/2}\right\} \leq \mathrm{P}(y < Y \leq y+h) \leq \frac{e^{-y/2}}{\Gamma(n/2)2^{n/2}} \left(\frac{2}{n}\right) \left\{(y+h)^{n/2} - y^{n/2}\right\},$$

となる. 今, $h=0$ 周りのテーラー展開より, $(y+h)^{n/2} = y^{n/2} + (n/2)y^{n/2-1}h + O(h^2)$ であるので,

$$\lim_{h\to 0} \frac{1}{h} \left\{(y+h)^{n/2} - y^{n/2}\right\} = \lim_{h\to 0} \frac{1}{h} \left\{y^{n/2} + (n/2)y^{n/2-1}h + O(h^2) - y^{n/2}\right\} = \frac{n}{2} y^{n/2-1},$$

となる. 以上により

$$\lim_{h\to 0} \frac{1}{h} \mathrm{P}(y < Y \leq y+h) = \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1},$$

となることがわかり, Y の確率密度関数が (3.1) となることがわかる. □

2. 定義 2 より, $X_1, \ldots, X_n \sim i.i.d.\ \mathrm{N}(\mu, \sigma^2)$ のとき, X_j を基準化したものを 2 乗して足したものも χ^2 分布に従う. つまり, $Y = \sum_{j=1}^n \{(X_j - \mu)/\sigma\}^2$ とすれば, $Y \sim \chi_n^2$ となる.

3. 定義 2 では, n は自然数でなくてはならないが, 定義 1 では, n は自然数でなく実数であっても問題ない.

特性

1. 自由度 n の χ^2 分布は右に歪んだ分布であり, n は位置とばらつきを表す母数である. また n が大きくなれば正規分布に近づく.

2. $Y \sim \chi_n^2$ のとき, Y の特性関数は

$$C(t) = (1-2it)^{-n/2}, \ (t \in \mathbb{R}), \tag{3.2}$$

となる.

(*Proof*) $Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(0,1)$ とすると, 定義 2 から $Y = Z_1^2 + \cdots + Z_n^2$ と置くことができる. よって

$$C_Y(t) = \mathrm{E}[\exp(itY)] = \mathrm{E}[\exp\{it(Z_1^2 + \cdots + Z_n^2)\}] = \prod_{j=1}^n \mathrm{E}[\exp(itZ_j^2)], \tag{3.3}$$

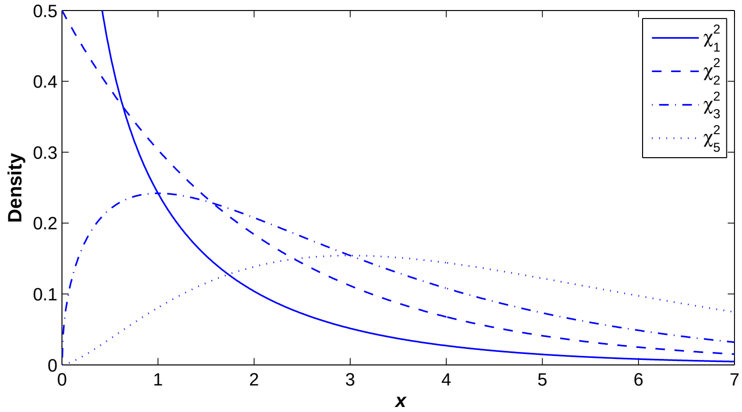

Figure 3.1. χ^2 分布の形

となる．今 $Z \sim \mathrm{N}(0,1)$ とすると,

$$\mathrm{E}[\exp(itZ^2)] = \int_{-\infty}^{\infty} e^{itz^2} \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz = (1-2it)^{-1/2} \int_{-\infty}^{\infty} \frac{1}{\{2\pi/(1-2it)\}^{1/2}} e^{-(1-2it)z^2/2} dz,$$

となる．ここで積分の中身は $\mathrm{N}(0, 1/(1-2it))$ の密度関数であるので, その全区間での積分値は 1 となる．よって $\mathrm{E}[\exp(itZ^2)] = (1-2it)^{-1/2}$．これを (3.3) に代入すると (3.2) となる． □

3. $Y \sim \chi^2_n$ のとき, $\mathrm{E}[Y] = n$, $\mathrm{Var}[Y] = 2n$ が成り立つ．また一般的には,

$$\mathrm{E}[Y^r] = n(n+2)\cdots\{n+2(r-1)\}, \tag{3.4}$$

が成り立つ.

(*Proof*) $Y \sim \chi^2_n$ のとき, Y の特性関数は, (3.2) より, $C(t) = (1-2it)^{-n/2}$ である．よって,

$$\begin{aligned}\frac{d^r}{dt^r} C(t) &= (-2it)^r \left(-\frac{n}{2}\right)\left(-\frac{n}{2}-1\right)\cdots\left\{-\frac{n}{2}-(r-1)\right\}(1-2it)^{-n/2-r} \\ &= i^r n(n+2)\cdots\{n+2(r-1)\}(1-2it)^{-n/2-r},\end{aligned}$$

となる．特性関数の性質から $\mathrm{E}[Y^r] = i^{-r} d^r C(t)/dt^r|_{t=0}$ であるので, 上記に式に $t=0$ を代入し, i^r で割ると (3.4) を得ることができる.

(3.4) より $r=1$, $r=2$ のときを計算すれば, $\mathrm{E}[Y] = n$, $\mathrm{E}[Y^2] = n(n+2)$ となる．ここで, $\mathrm{Var}[Y] = \mathrm{E}[Y^2] - (\mathrm{E}[Y])^2$ より, $\mathrm{Var}[Y] = n(n+2) - n^2 = 2n$ となる． □

4. 逆数のモーメント: $Y \sim \chi^2_n$ $(n - 2r > 0)$ のとき,

$$\mathrm{E}[Y^{-r}] = \{(n-2)\cdots(n-2r)\}^{-1}, \tag{3.5}$$

となる.

(*Proof*) $Y \sim \chi^2_n$ のとき, $\mathrm{E}[Y^{-r}]$ は,

$$\mathrm{E}[Y^{-r}] = \int_0^{\infty} y^{-r} \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1} dy = \int_0^{\infty} \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{(n-2r)/2-1} dy,$$

である．ここで, $2^{n/2} = 2^{(n-2r)/2} 2^r$ であり,

$$\Gamma(n/2) = \left(\frac{n}{2}-1\right)\cdots\left(\frac{n}{2}-r\right)\Gamma(n/2-r),$$

であるので,

$$\mathrm{E}[Y^{-r}] = 2^{-r}\left(\frac{n}{2}-1\right)^{-1}\cdots\left(\frac{n}{2}-r\right)^{-1}\int_0^\infty \frac{1}{2^{(n-2r)/2}\Gamma((n-2r)/2)}e^{-y/2}y^{(n-2r)/2-1}dy,$$

となる．上記の式の積分の中身は自由度 $n-2r$ の χ^2 分布の確率密度関数であるので, 全区間で積分すると 1 になる．よって,

$$\mathrm{E}[Y^{-r}] = 2^{-r}\left(\frac{n}{2}-1\right)^{-1}\cdots\left(\frac{n}{2}-r\right)^{-1} = \{(n-2)\cdots(n-2r)\}^{-1},$$

となり, (3.5) を得ることができる． □

5. 再生性: 互いに独立な確率変数 Y_1, Y_2 が, $Y_1 \sim \chi^2_{n_1}$, $Y_2 \sim \chi^2_{n_2}$ のとき, $Y_1+Y_2 \sim \chi^2_{n_1+n_2}$ となる．

(*Proof*) $W = Y_1+Y_2$ とし, Y_1 と Y_2 の特性関数をそれぞれ $C_{Y_1}(t)$, $C_{Y_2}(t)$ とおくと, W の特性関数 $C_W(t)$ は,

$$C_W(t) = \mathrm{E}[\exp(itW)] = \mathrm{E}[\exp\{it(Y_1+Y_2)\}] = C_{Y_1}(t)C_{Y_2}(t),$$

となる．ここで, (3.2) より, $C_{Y_j}(t) = (1-2it)^{-n_j/2}$ $(j=1,2)$ であるので, これを代入すると,

$$C_W(t) = (1-2it)^{-n_1/2}(1-2it)^{-n_2/2} = (1-2it)^{-(n_1+n_2)/2},$$

となる．これは自由度 n_1+n_2 の χ^2 分布の特性関数であり, 特性関数の一意性により $W \sim \chi^2_{n_1+n_2}$ であることがわかる． □

6. 歪正規分布と χ^2 分布: $X_1,\ldots,X_n \sim i.i.d.\ \mathrm{SN}(\lambda)$ のとき, $Y = \sum_{j=1}^n X_j^2 \sim \chi^2_n$ となる．

(*Proof*) $X \sim \mathrm{SN}(\lambda)$ のとき $X^2 \sim \chi^2_1$ であることが示せれば, あとは χ^2 分布の再生性から $Y = \sum_{j=1}^n X_j^2 \sim \chi^2_n$ を示すことができる．よって, 以下, $X \sim \mathrm{SN}(\lambda)$ のとき $X^2 \sim \chi^2_1$ を示す．X の分布関数を $F_\lambda(x)$ とすると,

$$\mathrm{P}(X^2 \le y) = \mathrm{P}(-\sqrt{y} \le X \le \sqrt{y}) = F_\lambda(\sqrt{y}) - F_\lambda(-\sqrt{y}),$$

となるので, $Y = X^2$ の密度関数 $f_Y(y)$ は,

$$f_Y(y) = \frac{d}{dy}\mathrm{P}(X^2 \le y) = y^{-1/2}\phi(\sqrt{y})\left\{\Phi(\lambda\sqrt{y}) + \Phi(-\lambda\sqrt{y})\right\}$$

となる．ここで $\Phi(y)+\Phi(-y) = 1$ であることに注意すれば,

$$f_Y(y) = y^{-1/2}\phi(\sqrt{y}) = \frac{1}{\sqrt{2\pi}}e^{-y/2}y^{-1/2},$$

となり, Y は自由度 1 の χ^2 分布に従うことがわかる． □

7. 対数 χ^2 分布の平均と分散: $Y \sim \chi^2_n$ のとき, $W = \log Y$ が従う分布を対数 χ^2 分布という．このとき,

$$\mathrm{E}[W] = \log 2 + \psi(n/2), \quad \mathrm{Var}[W] = \psi^{(1)}(n/2),$$

である．ただし, $\psi(x)$, $\psi^{(1)}(x)$ は,

$$\psi(x) = \frac{d}{dx}\log\Gamma(x), \quad \psi^{(1)}(x) = \frac{d}{dx}\psi(x),$$

で定義されるディガンマ関数, トリガンマ関数である．

(*Proof*) まず, 定数 α に対して $\mathrm{E}[Y^\alpha]$ を計算すると,

$$\begin{aligned}\mathrm{E}[Y^\alpha] &= \int_0^\infty y^\alpha \frac{1}{2^{n/2}\Gamma(n/2)}e^{-y/2}y^{n/2-1}dy = \int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)}e^{-y/2}y^{(n+2\alpha)/2-1}dy \\ &= \frac{2^{(n+2\alpha)/2}\Gamma(n/2+\alpha)}{2^{n/2}\Gamma(n/2)}\int_0^\infty \frac{1}{2^{(n+2\alpha)/2}\Gamma(n/2+\alpha)}e^{-y/2}y^{(n+2\alpha)/2-1}dy = \frac{2^\alpha\Gamma(n/2+\alpha)}{\Gamma(n/2)}.\end{aligned}$$

ここで, W の特性関数を $C_W(t)$ $(t \in \mathbb{R})$ とすると, 先の結果より,

$$C_W(t) = \mathrm{E}\left[\exp(itW)\right] = \mathrm{E}\left[\exp(it\log Y)\right] = \mathrm{E}\left[\{\exp(\log Y)\}^{it}\right] = \mathrm{E}\left[Y^{it}\right] = \frac{2^{it}\Gamma(n/2+it)}{\Gamma(n/2)}.$$

このとき, $\Gamma^{(k)}(x) = d^k\Gamma(x)/dx^k$ とすると,

$$\begin{aligned}
\frac{d}{dt}C_W(t) &= i\log 2 \cdot C_W(t) + i2^{it}\frac{\Gamma^{(1)}(n/2+it)}{\Gamma(n/2)}, \\
\frac{d^2}{dt^2}C_W(t) &= i\log 2\frac{d}{dt}C_W(t) + i^2\log 2\cdot 2^{it}\frac{\Gamma^{(1)}(n/2+it)}{\Gamma(n/2)} + i^2 2^{it}\frac{\Gamma^{(2)}(n/2+it)}{\Gamma(n/2)} \\
&= i\log 2\left\{i\log 2C_W(t) + i2^{it}\frac{\Gamma^{(1)}(n/2+it)}{\Gamma(n/2)}\right\} \\
&\quad + i^2\log 2\cdot 2^{it}\frac{\Gamma^{(1)}(n/2+it)}{\Gamma(n/2)} + i^2 2^{it}\frac{\Gamma^{(2)}(n/2+it)}{\Gamma(n/2)} \\
&= i^2(\log 2)^2 C_W(t) + 2i^2\log 2\cdot 2^{it}\frac{\Gamma^{(1)}(n/2+it)}{\Gamma(n/2)} + i^2 2^{it}\frac{\Gamma^{(2)}(n/2+it)}{\Gamma(n/2)}.
\end{aligned}$$

ここで, $C_W(0) = 1$ であり,

$$\begin{aligned}
\psi(x) &= \frac{d}{dx}\log\Gamma(x) = \frac{\Gamma^{(1)}(x)}{\Gamma(x)}, \\
\psi^{(1)}(x) &= \frac{d}{dx}\frac{\Gamma^{(1)}(x)}{\Gamma(x)} = \frac{\Gamma^{(2)}(x)\Gamma(x) - \{\Gamma^{(1)}(x)\}^2}{\{\Gamma(x)\}^2} = \frac{\Gamma^{(2)}(x)}{\Gamma(x)} - \left\{\frac{\Gamma^{(1)}(x)}{\Gamma(x)}\right\}^2 = \frac{\Gamma^{(2)}(x)}{\Gamma(x)} - \{\psi(x)\}^2,
\end{aligned}$$

であるので, これらを用いると,

$$\begin{aligned}
\left.\frac{d}{dt}C_W(t)\right|_{t=0} &= i\left\{\log 2 + \psi(n/2)\right\}, \\
\left.\frac{d^2}{dt^2}C_W(t)\right|_{t=0} &= -\left[(\log 2)^2 + 2\log 2\cdot\psi(n/2) + \psi^{(1)}(n/2) + \{\psi(n/2)\}^2\right] \\
&= -\left\{\log 2 + \psi(n/2)\right\}^2 - \psi^{(1)}(n/2).
\end{aligned}$$

ここで特性関数のモーメント公式を用いると,

$$\mathrm{E}[W] = \log 2 + \psi(n/2), \quad \mathrm{E}[W^2] = \{\log 2 + \psi(n/2)\}^2 + \psi^{(1)}(n/2),$$

であり, $\mathrm{Var}[W] = \mathrm{E}[W^2] - (\mathrm{E}[W])^2 = \psi^{(1)}(n/2)$ となるので, 題意は示された. □

3.2. 2 次形式 (Quadratic Form) 統計量の分布

標本分散 (Sample Variance)

$X_1, \dots, X_n$ を $\mathrm{E}[X_i] = \mu$, $\mathrm{Var}[X_i] = \sigma^2$ を持つ分布に従う独立な確率変数とし, 標本平均 $\bar{X}$ を $\bar{X} = n^{-1}\sum_{i=1}^n X_i$ とする. このとき, 標本の散らばりを測る指数である標本分散は

$$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2, \tag{3.6}$$

で定義される.

(3.6) 式の S^2 は,

$$\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{i=1}^n (X_i - \mu + \mu - \bar{X})^2 \\
&= \frac{1}{n-1}\sum_{i=1}^n \left\{(X_i-\mu)^2 - 2(X_i-\mu)(\bar{X}-\mu) + (\bar{X}-\mu)^2\right\}, \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^n (X_i-\mu)^2 - 2n(\bar{X}-\mu)(\bar{X}-\mu) + n(\bar{X}-\mu)^2\right\} \\
&= \frac{1}{n-1}\sum_{i=1}^n (X_i-\mu)^2 - \frac{n}{n-1}(\bar{X}-\mu)^2,
\end{aligned} \tag{3.7}$$

と分解できる. $\mathrm{E}[\bar{X}] = \mu$ なので, 上記の式を使うと, S^2 の期待値は,

$$\mathrm{E}\left[S^2\right] = \frac{1}{n-1}\sum_{i=1}^n \mathrm{E}\left[(X_i-\mu)^2\right] - \frac{n}{n-1}\mathrm{E}\left[(\bar{X}-\mu)^2\right] = \frac{n}{n-1}\left(\mathrm{Var}[X_i] - \mathrm{Var}[\bar{X}]\right), \tag{3.8}$$

となる. ここで, 定義から $\mathrm{Var}[X_i] = \sigma^2$ であり, 独立な確率変数の和の分散の公式から,

$$\mathrm{Var}[\bar{X}] = \frac{1}{n^2}\mathrm{Var}\left[\textstyle\sum_{i=1}^n X_i\right] = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}[X_i] = \frac{1}{n}\sigma^2,$$

となるから, これらを (3.8) に代入すれば,

$$\mathrm{E}[S^2] = \frac{n}{n-1}\left(\sigma^2 - \frac{1}{n}\sigma^2\right) = \sigma^2,$$

となることがわかる. これは S^2 が分散 σ^2 の**不偏推定量** (**Unbiased Estimator**) となっていることを示している.

ベキ等行列に関する補題

補題 3.2.1:A を n 次**ベキ等行列** (**Idempotent Matrix**), つまり $A^2 = A$ であるとする. このとき, A の固有値は 1 か 0 だけとなり, $\mathrm{rank}(A) = \mathrm{tr}(A)$ が成り立つ.

(*Proof*) λ を A の固有値, $\boldsymbol{x}$ を固有値 λ に対する固有ベクトルとする. このとき, 固有値の定義から, $A\boldsymbol{x} = \lambda\boldsymbol{x}$ が成り立つ. この式に左から A をかけると $A^2\boldsymbol{x} = \lambda A\boldsymbol{x}$ となり, $A^2 = A$ と $A\boldsymbol{x} = \lambda\boldsymbol{x}$ を代入すると, $A\boldsymbol{x} = \lambda^2\boldsymbol{x}$ を得る. 以上の 2 式から, $\lambda\boldsymbol{x} = \lambda^2\boldsymbol{x}$ を得, これは λ が 1 か 0 かのどちらかとなることを示している. また行列のランクは 0 ではない固有値の数に一致し, 固有値の和は行列のトレースに一致するので, 固有値が 1 か 0 の場合, 固有値の和は 0 ではない固有値の数に一致する. よって $\mathrm{rank}(A) = \mathrm{tr}(A)$. □

対称行列に関する補題

補題 3.2.2:$O_{n,p}$ はすべての成分が 0 である $n \times p$ 零行列とし, A, B を n 次対称行列とする. このとき, $AB = O_{n,n}$ であれば, A, B とも共通の直交行列で対角化可能である. 特に, Λ_1 を A の非零固有値を対角成分にもつ対角行列, Λ_2 を B の非零固有値を対角成分にもつ対角行列とすると,

$$A = H \begin{pmatrix} \Lambda_1 & O_{a,n-a} \\ O_{n-a,a} & O_{n-a,n-a} \end{pmatrix} H', \quad B = H \begin{pmatrix} O_{a,a} & O_{a,b} & O_{a,n-a-b} \\ O_{b,a} & \Lambda_2 & O_{b,n-a-b} \\ O_{n-a-b,a} & O_{n-a-b,b} & O_{n-a-b,n-a-b} \end{pmatrix} H',$$

となる直交行列 H が存在する. ただし, $a = \mathrm{rank}(A)$, $b = \mathrm{rank}(B)$ である.

(*Proof*) 階数は非零固有値の数に一致するので, A の非零固有値の数は a 個であり, B の非零固有値の数は b 個となる. ここで, $AB = O_{n,n}$ なので, 固有値の定義より, λ を B の固有値とし, $\boldsymbol{h}$ をその固有ベクトルとすると,

$$AB = O_{n,n} \Rightarrow AB\boldsymbol{h} = \boldsymbol{0} \Leftrightarrow \lambda A\boldsymbol{h} = \boldsymbol{0} \Leftrightarrow \lambda \cdot A\boldsymbol{h} = 0 \cdot \boldsymbol{h}.$$

よって, B の非零固有値に対する固有ベクトルは A の零固有値に対する固有ベクトルとなる. また, $O_{n,n} = O'_{n,n} = (AB)' = B'A' = BA$ であるので, 同様な方法を用いると, A の非零固有値に対する固有ベクトルは B の零固有値に対する固有ベクトルとなる. 以上により, A の固有ベクトル, $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n$ は,

$\boldsymbol{h}_1, \ldots, \boldsymbol{h}_a$: A の非零固有値に対する固有ベクトルで B の零固有値に対する固有ベクトル
$\boldsymbol{h}_{a+1}, \ldots, \boldsymbol{h}_{a+b}$: A の零固有値に対する固有ベクトルで B の非零固有値に対する固有ベクトル , (3.9)
$\boldsymbol{h}_{a+b+1}, \ldots, \boldsymbol{h}_n$: B の零固有値に対する固有ベクトル

と分割することがでる. B の非零固有値を $\lambda_{a+1}, \ldots, \lambda_{a+b}$ とおくと, 上記の式と (2.4) 式のスペクトル分解により,

$$B = \sum_{j=a+1}^{a+b} \lambda_j \boldsymbol{h}_j \boldsymbol{h}'_j. \tag{3.10}$$

ここで, $H = (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n)$ とおくと, $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n$ は A の固有ベクトルなので, H により A は直交可能で, $\lambda_1, \ldots, \lambda_a$ を A の非零固有値とし, $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_a)$ とおくと,

$$H'AH = \begin{pmatrix} \Lambda_1 & O_{a,n-a} \\ O_{n-a,a} & O_{n-a,n-a} \end{pmatrix}.$$

また, $\Lambda_2 = \mathrm{diag}(\lambda_{a+1}, \ldots, \lambda_{a+b})$ とおくと, (3.9), (3.10) 式より, $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n$ が互いに直交する単位ベクトルであることに注意すれば,

$$H'BH = \sum_{j=a+1}^{a+b} \lambda_j \begin{pmatrix} \boldsymbol{h}'_1 \\ \vdots \\ \boldsymbol{h}'_n \end{pmatrix} \boldsymbol{h}_j \boldsymbol{h}'_j (\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n) = \begin{pmatrix} O_{a,a} & O_{a,b} & O_{a,n-a-b} \\ O_{b,a} & \Lambda_2 & O_{b,n-a-b} \\ O_{n-a-b,a} & O_{n-a-b,b} & O_{n-a-b,n-a-b} \end{pmatrix}.$$

よって, B は H により対角化可能. 以上により題意が示せた. □

S^2 の標本分布

定理 3.2.1: $X_1, \ldots, X_n \sim i.i.d.\ \mathrm{N}(\mu, \sigma^2)$ とし, $W = (n-1)S^2/\sigma^2$ とする. このとき $W \sim \chi^2_{n-1}$ である.

(*Proof*) (3.7) 式を用いて W を分解すると

$$W = \frac{n-1}{\sigma^2} S^2 = \frac{1}{\sigma^2} \left\{ \sum_{i=1}^{n} (X_i - \mu)^2 - n(\bar{X} - \mu)^2 \right\} = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 - n \left(\frac{\bar{X} - \mu}{\sigma} \right)^2, \tag{3.11}$$

となる. ここで $Z_i = (X_i - \mu)/\sigma$ とおくと, $X_1, \ldots, X_n \sim i.i.d.\ \mathrm{N}(\mu, \sigma^2)$ なので, $Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(0,1)$ となる. また,

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma} = \frac{\bar{X} - \mu}{\sigma},$$

となるので, Z_i を使って (3.11) 式を書き換えると,

$$W = \sum_{i=1}^{n} Z_i^2 - n\bar{Z}^2 = \sum_{i=1}^{n} Z_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Z_i\right)^2, \tag{3.12}$$

となる. ここで, $\boldsymbol{Z} = (Z_1, \ldots, Z_n)'$ とし, $\mathbf{1}_n$ をすべての成分が 1 である n 次元ベクトルとする. このとき, $\sum_{i=1}^{n} Z_i = \mathbf{1}_n'\boldsymbol{Z}$, $\sum_{i=1}^{n} Z_i^2 = \boldsymbol{Z}'\boldsymbol{Z}$ であるので, $\mathbf{1}_n'\boldsymbol{Z} = \boldsymbol{Z}'\mathbf{1}_n$ であることに注意すると, (3.12) 式は,

$$W = \boldsymbol{Z}'\boldsymbol{Z} - \frac{1}{n}(\mathbf{1}_n'\boldsymbol{Z})^2 = \boldsymbol{Z}'\boldsymbol{Z} - \frac{1}{n}\boldsymbol{Z}'\mathbf{1}_n\mathbf{1}_n'\boldsymbol{Z} = \boldsymbol{Z}'\left(I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\boldsymbol{Z} = \boldsymbol{Z}'A\boldsymbol{Z}, \tag{3.13}$$

となる. ここで $A' = A$ であり, かつ,

$$A^2 = I_n - \frac{2}{n}\mathbf{1}_n\mathbf{1}_n' + \frac{1}{n^2}\mathbf{1}_n\mathbf{1}_n'\mathbf{1}_n\mathbf{1}_n' = I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n' = A,$$

であるので, A は対称ベキ等行列であることがわかる. 対称行列は直交行列によって対角化可能であるので, H を $H'AH = \Lambda$ とする n 次直交行列 ($H'H = HH' = I_n$) とする. このときの Λ は, A の固有値を対角成分に持つ対角行列であり, A のベキ等性から, 補題 3.2.1 により, A の固有値は 1 か 0 で, $\mathrm{tr}(A) = \mathrm{tr}(I_n) - n^{-1}\mathrm{tr}(\mathbf{1}_n\mathbf{1}_n') = n - n^{-1}n = n - 1$ であるので, 1 の個数は $n-1$ 個となる. よって, Λ は,

$$\Lambda = \begin{pmatrix} I_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1}' & 0 \end{pmatrix},$$

となる. ただし $\mathbf{0}_n$ は全ての成分が 0 である n 次元ベクトルとする. $\boldsymbol{U} = (U_1, \ldots, U_n)' = H'\boldsymbol{Z}$ とし, $A = H\Lambda H'$ であることを利用すれば, (3.13) 式は,

$$W = \boldsymbol{Z}'H\Lambda H'\boldsymbol{Z} = \boldsymbol{U}'\Lambda\boldsymbol{U} = \begin{pmatrix} U_1 & \ldots & U_n \end{pmatrix}\begin{pmatrix} I_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1}' & 0 \end{pmatrix}\begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix} = \sum_{i=1}^{n-1} U_i^2. \tag{3.14}$$

ここで, $\boldsymbol{Z} \sim \mathrm{N}_n(\mathbf{0}_n, I_n)$ であり, $\mathrm{E}[\boldsymbol{U}] = \mathrm{E}[H'\boldsymbol{Z}] = \mathbf{0}$, $\mathrm{Var}[\boldsymbol{U}] = H'\mathrm{Var}[\boldsymbol{X}]H = H'H = I_n$ であるので, 第 2.2 節の多次元正規分布の特性 4 より, $\boldsymbol{U} \sim \mathrm{N}_n(\mathbf{0}_n, I_n)$. これは $U_1, \ldots, U_n \sim i.i.d.\ N(0,1)$ を示すものであり, (3.14) 式と χ^2 分布の定義 2 より $W \sim \chi^2_{n-1}$ であることがわかる. □

<u>2 次形式統計量の分布と独立性</u>

定理 3.2.2: $Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(0,1)$ とし, $\boldsymbol{Z} = (Z_1, \ldots, Z_n)'$ とおき, A を対称な n 次ベキ等行列とする. このとき, $\boldsymbol{Z}'A\boldsymbol{Z} \sim \chi^2_f$ である. ただし, $f = \mathrm{tr}(A)\ (\leq n)$ である.

(*Proof*) $Y = \boldsymbol{Z}'\boldsymbol{A}\boldsymbol{Z}$ とおく. 今, A は対称ベキ等行列なので, n 次直交行列 H によって, $H'AH = \Lambda$ と対角化可能である. このときの Λ は $\mathrm{tr}(A) = f$ なので,

$$\Lambda = \begin{pmatrix} I_f & O_{f,n-k} \\ O_{n-f,f} & O_{n-f,n-f} \end{pmatrix}, \tag{3.15}$$

となる. 定理 3.2.1 の証明と同様に, $\boldsymbol{U} = (U_1, \ldots, U_n)' = H'\boldsymbol{Z}$ とすれば, $U_1, \ldots, U_n \sim i.i.d.\ \mathrm{N}(0,1)$ であり, (3.14) 式の Λ を (3.15) 式の Λ に変えることにより,

$$Y = \sum_{i=1}^{f} U_i^2,$$

を得ることができる. ここで, χ^2 分布の定義 2 より, $Y \sim \chi^2_f$ となることがわかる. □

定理 3.2.3: $Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(0,1)$ とし, $\boldsymbol{Z} = (Z_1, \ldots, Z_n)'$ とおき, A と B を対称な n 次ベキ等行列で $\mathrm{tr}(A) = a$, $\mathrm{tr}(B) = b$, $AB = O_{n,n}$ とする (ただし $a + b \leq n$). このとき, $\boldsymbol{Z}'A\boldsymbol{Z}$ と $\boldsymbol{Z}'B\boldsymbol{Z}$ は独立となる.

(*Proof*) $X = \boldsymbol{Z}'A\boldsymbol{Z}$, $Y = \boldsymbol{Z}'B\boldsymbol{Z}$ とおく．$AB = O_{n,n}$ より, 補題 3.2.2 より, A, B を分解する共通の直交行列 H が存在し, A, B はベキ等行列であることを注意すれば,

$$A = H\begin{pmatrix} I_a & O_{a,n-a} \\ O_{n-a,a} & O_{n-a,n-a} \end{pmatrix}H', \quad B = H\begin{pmatrix} O_{a,a} & O_{a,b} & O_{a,n-a-b} \\ O_{b,a} & I_b & O_{b,n-a-b} \\ O_{n-a-b,a} & O_{n-a-b,b} & O_{n-a-b,n-a-b} \end{pmatrix}H',$$

となる．ここで, $\boldsymbol{U} = (U_1, \ldots, U_n)' = H'\boldsymbol{Z}$ とおけば, $U_1, \ldots, U_n \sim i.i.d.\mathrm{N}(0,1)$ であり,

$$X = \boldsymbol{Z}'A\boldsymbol{Z} = \boldsymbol{Z}'HH'AHH'\boldsymbol{Z} = \boldsymbol{U}'\begin{pmatrix} I_a & O_{a,n-a} \\ O_{n-a,a} & O_{n-a,n-a} \end{pmatrix}\boldsymbol{U} = \sum_{i=1}^{a} U_i^2,$$

$$Y = \boldsymbol{Z}'B\boldsymbol{Z} = \boldsymbol{Z}'HH'BHH'\boldsymbol{Z} = \boldsymbol{U}'\begin{pmatrix} O_{a,a} & O_{a,b} & O_{a,n-a-b} \\ O_{b,a} & I_b & O_{b,n-a-b} \\ O_{n-a-b,a} & O_{n-a-b,b} & O_{n-a-b,n-a-b} \end{pmatrix}\boldsymbol{U} = \sum_{i=a+1}^{a+b} U_i^2.$$

$U_1, \ldots, U_a, U_{a+1}, \ldots, U_{a+b}$ は互いに独立なので, X と Y は独立. □

定理 3.2.4: $Z_1, \ldots, Z_n \sim i.i.d.$ $\mathrm{N}(0,1)$ とし, $\boldsymbol{Z} = (Z_1, \ldots, Z_n)'$ とおき, A を対称な n 次ベキ等行列, $\boldsymbol{b}$ を $A\boldsymbol{b} = \boldsymbol{0}_n$ を満たす n 次元ベクトルとする．このとき, $\boldsymbol{Z}'A\boldsymbol{Z}$ と $\boldsymbol{b}'\boldsymbol{Z}$ は独立となる．

(*Proof*) $Y = \boldsymbol{Z}'\boldsymbol{A}\boldsymbol{Z}$, $X = \boldsymbol{b}'\boldsymbol{Z}$ とおく．今,

$$W = \frac{1}{\boldsymbol{b}'\boldsymbol{b}}\boldsymbol{Z}'\boldsymbol{b}\boldsymbol{b}'\boldsymbol{Z} = \boldsymbol{Z}'B\boldsymbol{Z},$$

とおく．このとき, B は対称ベキ等行列であり, $AB = O_{n,n}$ となる．よって, 定理 3.2.3 の証明により, A と B を対角化する直交行列を H とし, $\boldsymbol{U} = (U_1, \ldots, U_n)' = H'\boldsymbol{Z}$ とおけば, $U_1, \ldots, U_n$ は互いに独立であり, $\mathrm{rank}(B) = 1$ であることに注意すれば, $\mathrm{rank}(A) = a$ とすると,

$$Y = \sum_{i=1}^{n} U_i^2, \quad W = U_{a+1}^2.$$

また, B の定義から, B の非零固有値に対する固有ベクトルは $\boldsymbol{b}/\|\boldsymbol{b}\|$ であり, $U_{a+1} = \boldsymbol{b}'\boldsymbol{Z}/\|\boldsymbol{b}\| = X/\|\boldsymbol{b}\|$. $\|\boldsymbol{b}\|$ が定数であることに注意すれば, Y と X も独立となることがわかる． □

定理 3.2.4 から, 以下の定理を得ることができる．

定理 3.2.5: $X_1, \ldots, X_n \sim i.i.d.$ $\mathrm{N}(\mu, \sigma^2)$ のとき, S^2 と $\bar{X}$ は独立である．

(*Proof*) 今, $Z_i = (X_i - \mu)/\sigma$ とすると, $Z_1, \ldots, Z_n \sim i.i.d.$ $\mathrm{N}(0,1)$ であり, $X_i = \mu + \sigma Z_i$ となるので,

$$\bar{X} = \mu + \sigma\bar{Z},$$

となる．ここで, $W = (n-1)S^2/\sigma^2$ とすると, $n-1$, σ, μ は定数なので, S^2 と $\bar{X}$ の独立性を調べるためには, W と $\bar{Z}$ の独立性を調べればよい．定理 3.2.1 の証明より, $A = I_n - n^{-1}\boldsymbol{1}_n\boldsymbol{1}_n'$ とおけば, $W = \boldsymbol{Z}'A\boldsymbol{Z}$ であり, $\bar{Z} = n^{-1}\boldsymbol{1}_n'/\boldsymbol{Z} = \boldsymbol{b}'\boldsymbol{Z}$ となる．ここで, A はベキ等行列であり,

$$A\boldsymbol{b} = \left(I_n - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n'\right)\left(\frac{1}{n}\boldsymbol{1}_n\right) = \frac{1}{n}\left(I_n - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n'\right)\boldsymbol{1}_n = \frac{1}{n}\left(\boldsymbol{1}_n - \boldsymbol{1}_n\right) = \boldsymbol{0}_n,$$

となる．よって定理 3.2.4 から, W と $\bar{Z}$ は独立となることがわかる．以上により, S^2 と $\bar{X}$ は独立となる．
□

3.3. 非心 χ^2 分布 (Non-central Chi-Square Distribution)

3.1 節で, $Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(0,1)$ のとき, $Y = \sum_{j=1}^{n} Z_j^2 \sim \chi_n^2$ であることを示した. では, $Z_1, \ldots, Z_n \sim i.i.d.\ \mathrm{N}(\mu, 1)$ のとき, $Y = \sum_{j=1}^{n} X_j^2$ の分布はどうなるだろうか? 非心 χ^2 分布は, その問に対する回答となる分布である.

定義 1

$Y \sim \chi_n^2(\delta)$ であるとは, Y が従う分布の確率密度関数が,

$$f(y; n, \delta) = \begin{cases} \dfrac{1}{2^{n/2}} e^{-(y+\delta)/2} \displaystyle\sum_{j=0}^{\infty} \left(\frac{\delta}{4}\right)^j \frac{y^{n/2+j-1}}{j!\Gamma(n/2+j)} & (y > 0) \\ 0 & (y \leq 0) \end{cases} \quad (n \in \mathbb{N}, \delta \geq 0), \tag{3.16}$$

であるときのことをいう. また, この確率分布を自由度 n, 非心度 δ の非心 χ^2 分布という.

非心度が $\delta = 0$ のとき, $0^0 = 1$ であることに注意すれば,

$$f(y; n, 0) = \frac{1}{2^{n/2}} e^{-(y+0)/2} \sum_{j=0}^{\infty} \left(\frac{0}{4}\right)^j \frac{y^{n/2+j-1}}{j!\Gamma(n/2+j)} = \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1},$$

となり, (3.1) 式から, 自由度 n の χ^2 分布の確率密度関数に一致していることがわかる. そのことから, 非心 χ^2 分布は特別な場合 ($\delta = 0$) として χ^2 分布を含む, χ^2 分布よりも広いクラスの分布であることがわかる.

χ^2 分布と同様に, 非心 χ^2 分布の定義として, 以下のような定義を用いているテキストも多い.

定義 2

$Z_1, \ldots, Z_n$ は互いに独立に, $Z_j \sim \mathrm{N}(\mu_j, 1)$ $(j = 1, \ldots, n)$ に従うとする. このとき, $Y = \sum_{j=1}^{n} Z_j^2$ が従う分布が $\chi_n^2(\delta)$ である. ただし, $\delta = \sum_{j=1}^{n} \mu_j^2$ である.

特性

1. 自由度 n, 非心度 δ を持つ非心 χ^2 分布は右に歪んだ分布であり, n と δ はともに位置とばらつきを表す母数である. また, n または δ が大きくなれば正規分布に近づく.

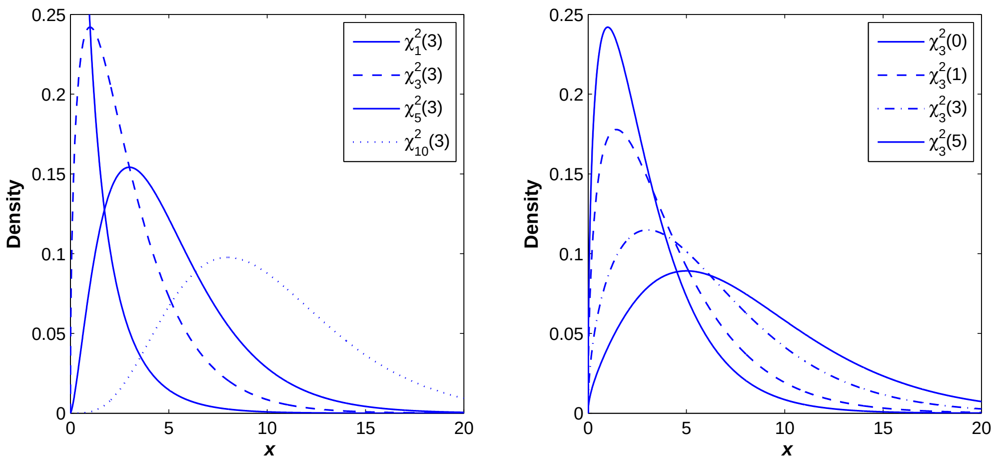

Figure 3.2. 非心 χ^2 分布の形

2. $Y \sim \chi_n^2(\delta)$ のとき, Y の特性関数は

$$C(t) = (1-2it)^{-n/2} \exp\left(\frac{it\delta}{1-2it}\right), \quad (t \in \mathbb{R}), \tag{3.17}$$

となる.

(*Proof*) $Z_1, \ldots, Z_n$ を独立な確率変数として, $Z_j \sim i.i.d.\ \mathrm{N}(\mu_j, 1)$ $(j = 1, \ldots, n)$ とすると, 定義 2 から $Y = Z_1^2 + \cdots + Z_n^2$ と置くことができる. よって

$$C_Y(t) = \mathrm{E}[\exp(itY)] = \mathrm{E}[\exp\{it(Z_1^2 + \cdots + Z_n^2)\}] = \prod_{j=1}^n \mathrm{E}[\exp(itZ_j^2)], \tag{3.18}$$

となる. 今 $Z \sim \mathrm{N}(\mu, 1)$ とすると,

$$\begin{aligned}
\mathrm{E}[\exp(itZ^2)] &= \int_{-\infty}^{\infty} e^{itz^2} \frac{1}{(2\pi)^{1/2}} e^{-(z-\mu)^2/2} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{1}{2}(1-2it)z^2 + \mu z - \frac{1}{2}\mu^2\right\} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{1}{2}(1-2it)\left(z - \frac{\mu}{1-2it}\right)^2\right\} \exp\left(\frac{it\mu^2}{1-2it}\right) dz \\
&= (1-2it)^{-1/2} \exp\left(\frac{it\mu^2}{1-2it}\right) \\
&\quad \times \int_{-\infty}^{\infty} \frac{1}{\{2\pi/(1-2it)\}^{1/2}} \exp\left\{-\frac{1}{2}(1-2it)\left(z - \frac{\mu}{1-2it}\right)^2\right\} dz,
\end{aligned}$$

となる. ここで積分の中身は $\mathrm{N}(\mu/(1-2it), 1/(1-2it))$ の密度関数であるので, その全区間での積分値は 1 となる. よって $\mathrm{E}[\exp(itZ^2)] = (1-2it)^{-1/2}\exp(\mu^2 it/(1-2it))$. これを (3.18) 式に代入し, $\delta = \mu_1^2 + \cdots + \mu_n^2$ であることに注意すれば,

$$C_Y(t) = \prod_{j=1}^n (1-2it)^{-1/2} \exp\left(\frac{\mu_j^2 it}{1-2it}\right) = (1-2it)^{-n/2} \exp\left(\frac{\delta it}{1-2it}\right),$$

となり, (3.17) 式を導くことができる. □

3. $Y \sim \chi_n^2(\delta)$ のとき, $\mathrm{E}[Y] = n + \delta$, $\mathrm{Var}[Y] = 2(n + 2\delta)$ が成り立つ.

   (*Proof*) Y の特性関数の t に関する 1・2 回微分にそれぞれに $t = 0$ を代入して, i と i^2 で割れば $\mathrm{E}[Y]$ と $\mathrm{E}[Y^2]$ を得ることができる. 分散に関しては, これらを用いて, $\mathrm{E}[Y^2] - (\mathrm{E}[Y])^2$ を計算すればよい. □

4. 再生性:互いに独立な確率変数 Y_1, Y_2 が, $Y_1 \sim \chi_{n_1}^2(\delta_1)$, $Y_2 \sim \chi_{n_2}^2(\delta_2)$ のとき, $Y_1 + Y_2 \sim \chi_{n_1+n_2}^2(\delta_1 + \delta_2)$ となる.

   (*Proof*) $W = Y_1 + Y_2$ とし, Y_1 と Y_2 の特性関数をそれぞれ $C_{Y_1}(t)$, $C_{Y_2}(t)$ とおくと, W の特性関数 $C_W(t)$ は,

$$C_W(t) = \mathrm{E}[\exp(itW)] = \mathrm{E}[\exp\{it(Y_1 + Y_2)\}] = C_{Y_1}(t)C_{Y_2}(t),$$

   となる. ここで, (3.17) より, $C_{Y_j}(t) = (1-2it)^{-n_j/2}\exp\{it\delta_j/(1-2it)\}$ $(j = 1, 2)$ であるので, これを代入すると,

$$\begin{aligned}
C_W(t) &= (1-2it)^{-n_1/2}(1-2it)^{-n_2/2} \exp\left(\frac{it\delta_1}{1-2it}\right) \exp\left(\frac{it\delta_2}{1-2it}\right) \\
&= (1-2it)^{-(n_1+n_2)/2} \exp\left\{\frac{it(\delta_1 + \delta_2)}{1-2it}\right\},
\end{aligned}$$

   となる. これは自由度 $n_1 + n_2$, 非心度 $\delta_1 + \delta_2$ の非心 χ^2 分布の特性関数であり, 特性関数の一意性により $W \sim \chi_{n_1+n_2}^2(\delta_1 + \delta_2)$ であることがわかる. □

3.4. ウィッシャート分布 (Wishart Distribution)

2.1 節では 1 次元正規分布を多次元に拡張した多次元正規分布を紹介した. では, χ^2 分布を多次元に拡張することができないのであろうか? ウィッシャート分布は, その問に対する回答となる分布である.

定義 1

V を対称で正則な $p \times p$ 確率変数行列とする. $V \sim \mathrm{W}_p(n, \Sigma)$ であるとは, S が従う分布の確率密度関数が,

$$f(V; n, \Sigma) = \frac{|V|^{(n-p-1)/2} \exp\{-\mathrm{tr}(V\Sigma^{-1})/2\}}{2^{pn/2}\pi^{p(p-1)/4}|\Sigma|^{n/2}\prod_{j=1}^{p}\Gamma((n-i+1)/2)} \qquad (n \in \mathbb{N}, \Sigma > 0), \tag{3.19}$$

であるときのことをいう. また, この確率分布を自由度 n, 平均パラメータ Σ のウィッシャート分布という.

$p = 1$ のとき, $|V| = v$, $V^{-1} = 1/v$ であることに注意すれば,

$$\frac{v^{(n-2)/2}\exp\{-v/(2\sigma)\}}{2^{n/2}\sigma^{n/2}\Gamma(n/2)} = \frac{1}{2^{n/2}\sigma\Gamma(n/2)} e^{-(v/\sigma)/2} (v/\sigma)^{n/2-1},$$

となり, (3.1) 式から, $\sigma = 1$ のとき, 自由度 n の χ^2 分布に一致することがわかる. そのことから, ウィッシャート分布は特別な場合 (Σ が単位行列で $p = 1$) として χ^2 分布を含む, χ^2 分布よりも広いクラスの分布であることがわかる.

χ^2 分布と同様に, ウィッシャート分布の定義として, 以下のような定義を用いているテキストも多い.

定義 2

$\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \sim i.i.d.\ \mathrm{N}_p(\boldsymbol{0}, \Sigma)$ とし, $V = \sum_{j=1}^{n} \boldsymbol{X}_j \boldsymbol{X}_j'$ とおく. このとき, V が従う分布が $\mathrm{W}_p(n, \Sigma)$ である.

特性

1. $V \sim \mathrm{W}_p(n, \Sigma)$ のとき, V の特性関数は

$$C(T) = |I_p - 2i\Sigma T|^{-n/2}, \ (T \text{ は } p \times p \text{ の対称行列でそれぞれの成分は実数}). \tag{3.20}$$

2. $V \sim \mathrm{W}_p(n, \Sigma)$ のとき, $\mathrm{E}[V] = n\Sigma$. $n - p - 1 > 0$ ならば, $\mathrm{E}[V^{-1}] = \Sigma^{-1}/(n - p - 1)$.

3. 再生性: 互いに独立な $p \times p$ 確率変数行列 V_1, V_2 が, $V_1 \sim \mathrm{W}_p(n_1, \Sigma)$, $V_2 \sim \mathrm{W}_p(n_2, \Sigma)$ のとき, $V_1 + V_2 \sim \mathrm{W}_p(n_1 + n_2, \Sigma)$ となる.

4. バートレット分解 (Bartlett Decomposition): $V \sim \mathrm{W}_p(n, I_p)$ のとき, $S = TT'$ を満たす以下のような p 次下側三角行列が存在する.

$$T = \begin{pmatrix} t_{11} & 0 & \cdots & 0 \\ t_{21} & t_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ t_{p1} & t_{p2} & \cdots & t_{pp} \end{pmatrix}, \quad \begin{cases} \text{下三角成分は互いに独立} \\ t_{ii} > 0, \quad t_{ii}^2 \sim \chi^2_{n-i+1} \\ t_{ij} \sim \mathrm{N}(0, 1) \ (i \neq j) \end{cases}.$$

3.5. 行列 2 次形式 (Matrix Quadratic Form) 統計量の分布

標本分散共分散行列 (Sample Variance Covariance Matrix)

$\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ を $\mathrm{E}[\boldsymbol{X}_i] = \boldsymbol{\mu}$, $\mathrm{Var}[\boldsymbol{X}_i] = \Sigma$ を持つ分布に従う独立な p 次元確率変数ベクトルとし, 標本平均 $\bar{\boldsymbol{X}}$ を $\bar{\boldsymbol{X}} = n^{-1}\sum_{i=1}^n \boldsymbol{X}_i$ とする. このとき, 標本の散らばりを測る指数である標本分散共分散行列は

$$S = \frac{1}{n-1}\sum_{i=1}^n (\boldsymbol{X}_i - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \bar{\boldsymbol{X}})', \tag{3.21}$$

で定義される.

(3.21) 式の S は,

$$\begin{aligned} S &= \frac{1}{n-1}\sum_{i=1}^n (\boldsymbol{X}_i - \boldsymbol{\mu} + \boldsymbol{\mu} - \bar{\boldsymbol{X}})(\boldsymbol{X}_i - \boldsymbol{\mu} + \boldsymbol{\mu} - \bar{\boldsymbol{X}})' \\ &= \frac{1}{n-1}\sum_{i=1}^n \left\{(\boldsymbol{X}_i - \boldsymbol{\mu})(\boldsymbol{X}_i - \boldsymbol{\mu})' - (\boldsymbol{X}_i - \boldsymbol{\mu})(\bar{\boldsymbol{X}} - \boldsymbol{\mu})' - (\bar{\boldsymbol{X}} - \boldsymbol{\mu})(\boldsymbol{X}_i - \boldsymbol{\mu})' + (\bar{\boldsymbol{X}} - \boldsymbol{\mu})(\bar{\boldsymbol{X}} - \boldsymbol{\mu})'\right\}, \\ &= \frac{1}{n-1}\left\{\sum_{i=1}^n (\boldsymbol{X}_i - \boldsymbol{\mu})(\boldsymbol{X}_i - \boldsymbol{\mu})' - 2n(\bar{\boldsymbol{X}} - \boldsymbol{\mu})(\bar{\boldsymbol{X}} - \boldsymbol{\mu})' + n(\bar{\boldsymbol{X}} - \boldsymbol{\mu})(\bar{\boldsymbol{X}} - \boldsymbol{\mu})'\right\} \\ &= \frac{1}{n-1}\sum_{i=1}^n (\boldsymbol{X}_i - \boldsymbol{\mu})(\boldsymbol{X}_i - \boldsymbol{\mu})' - \frac{n}{n-1}(\bar{\boldsymbol{X}} - \boldsymbol{\mu})(\bar{\boldsymbol{X}} - \boldsymbol{\mu})', \end{aligned} \tag{3.22}$$

と分解できる. $\mathrm{E}[\bar{\boldsymbol{X}}] = \boldsymbol{\mu}$ なので, 上記の式を使うと, S の期待値は,

$$\begin{aligned} \mathrm{E}[S] &= \frac{1}{n-1}\sum_{i=1}^n \mathrm{E}\left[(\boldsymbol{X}_i - \boldsymbol{\mu})(\boldsymbol{X}_i - \boldsymbol{\mu})'\right] - \frac{n}{n-1}\mathrm{E}\left[(\bar{\boldsymbol{X}} - \boldsymbol{\mu})(\bar{\boldsymbol{X}} - \boldsymbol{\mu})'\right] \\ &= \frac{n}{n-1}\left(\mathrm{Var}[\boldsymbol{X}_i] - \mathrm{Var}[\bar{\boldsymbol{X}}]\right), \end{aligned} \tag{3.23}$$

となる. ここで, 定義から $\mathrm{Var}[\boldsymbol{X}_i] = \Sigma$ であり, 独立な確率変数の和の分散共分散行列の公式から,

$$\mathrm{Var}[\bar{\boldsymbol{X}}] = \frac{1}{n^2}\mathrm{Var}\left[\textstyle\sum_{i=1}^n \boldsymbol{X}_i\right] = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}[\boldsymbol{X}_i] = \frac{1}{n}\Sigma,$$

となるから, これらを (3.23) に代入すれば,

$$\mathrm{E}[S] = \frac{n}{n-1}\left(\Sigma - \frac{1}{n}\Sigma\right) = \Sigma,$$

となることがわかる. これは S が分散共分散行列 Σ の不偏推定量となっていることを示している.

S の標本分布

定理 3.5.1: $X_1, \ldots, X_n \sim i.i.d.\ \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ とし, $V = (n-1)S$ とする. このとき $V \sim \mathrm{W}_p(n-1, \Sigma)$ である.

行列 2 次形式統計量の分布と独立性

定理 3.5.2: $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n \sim i.i.d.\ \mathrm{N}_p(\boldsymbol{0}_p, \Sigma)$ とし, $Z = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n)'$ とおき, A を対称な n 次ベキ等行列とする. このとき, $Z'AZ \sim \mathrm{W}_p(f, \Sigma)$ である. ただし, $f = \mathrm{tr}(A)\ (\leq n)$ である.

定理 3.5.3: $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n \sim i.i.d.\ \mathrm{N}_p(\boldsymbol{0}_p, \Sigma)$ とし, $Z = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n)'$ とおき, A と B を対称な n 次ベキ等行列で $\mathrm{tr}(A) = a$, $\mathrm{tr}(B) = b$, $AB = O_{n,n}$ とする (ただし $a + b \leq n$). このとき, $Z'AZ$ と $Z'BZ$ は独立となる.

定理 3.5.4: $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n \sim i.i.d.\ \mathrm{N}_p(\boldsymbol{0}_p, \Sigma)$ とし, $\boldsymbol{Z} = (Z_1, \ldots, Z_n)'$ とおき, A を対称な n 次ベキ等行列, $\boldsymbol{b}$ を $A\boldsymbol{b} = \boldsymbol{0}_n$ を満たす n 次元ベクトルとする. このとき, $Z'AZ$ と $Z'\boldsymbol{b}$ は独立となる.

定理 3.5.5: $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \sim i.i.d.\ \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ のとき, S と $\bar{\boldsymbol{X}}$ は独立である.

4. t-分布と F-分布

4.1. t-分布 (t-Distribution)

t-分布は 1908 年にギネスビール社で働くゴセット (Gosset, 1876–1936) により発見された. ギネスビール社は社員が学術雑誌に論文を投稿することを禁じていたので, ゴセットはスチューデント (Student) という偽名を使って t-分布に関する論文を発表した. そのため, 現在でも t-分布のことをスチューデントの t-分布と呼ぶことがある.

$X_1, \dots, X_n \sim i.i.d.\ \mathrm{N}(\mu, \sigma^2)$ のとき, μ の**信頼区間** (**Confidence Interval**) を考える.

(*i*) σ^2 が既知の場合: $\bar{X} = n^{-1} \sum_{j=1}^n X_j$ とすると, $\bar{X}$ は平均 μ の推定量である. この場合,

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim \mathrm{N}(0, 1),$$

であるので, 標準正規分布の上側 2.5%点である 1.96 を用いて,

$$\mathrm{P}\left(|\sqrt{n}(\bar{X} - \mu)/\sigma| \le 1.96\right) = 0.95 \ \Rightarrow\ \bar{X} - \frac{\sigma}{\sqrt{n}} 1.96 \le \mu \le \bar{X} + \frac{\sigma}{\sqrt{n}} 1.96,$$

を得る. これは「区間 $[\bar{X} - 1.96\sigma/\sqrt{n},\ \bar{X} + 1.96\sigma/\sqrt{n}]$ が μ を含む確率は 0.95(95%)」ということを示しており, この区間が信頼係数 0.95 の μ の信頼区間である.

(*ii*) σ^2 が未知の場合: σ^2 が未知であるので, σ^2 を推定量 $S^2 = (n-1)^{-1} \sum_{j=1}^n (X_j - \bar{X})^2$ で置き換えて考える. この場合,

$$\bar{X} - \frac{S}{\sqrt{n}} 1.96 \le \mu \le \bar{X} + \frac{S}{\sqrt{n}} 1.96,$$

となる区間を考えたとき, $\sqrt{n}(\bar{X} - \mu)/S$ の従う分布は正規分布ではないので,

$$\mathrm{P}\left(|\sqrt{n}(\bar{X} - \mu)/S| \le 1.96\right) \ne 0.95,$$

となり, 上記の信頼区間の信頼係数は厳密な意味での 0.95 にはならない. しかしながら, n が大きいときは, $S^2 \xrightarrow{p} \sigma^2$ であり, この結果から,

$$\frac{\sqrt{n}(\bar{X} - \mu)}{S} \xrightarrow{d} \mathrm{N}(0, 1),$$

であるので,

$$\mathrm{P}\left(|\sqrt{n}(\bar{X} - \mu)/S| \le 1.96\right) \to 0.95,$$

となり, 漸近的には信頼係数は 0.95 と言ってもよい. 一方, n が小さいときは, 漸近的な近似には問題があるので, 上記の信頼区間を信頼係数を 0.95 とすることには問題がある. 実は $\sqrt{n}(\bar{X} - \mu)/S$ が従う分布が自由度 $n-1$ の t-分布であり, t-分布の上側 2.5%点を用いれば, 信頼係数 0.95 の μ の信頼区間を構成することができる.

定義

互いに独立な確率変数 Z と Y が $Z \sim \mathrm{N}(0, 1)$, $Y \sim \chi_n^2$ であるとき,

$$T = \frac{Z}{\sqrt{Y/n}},$$

の従う分布を自由度 n の t-分布といい, $T \sim t_n$ と書く.

定理 4.1.1: $X_1, \dots, X_n \sim i.i.d.\ \mathrm{N}(\mu, \sigma^2)$ のとき,

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{n-1},$$

である.

(*Proof*) 今, T を

$$T = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \Bigg/ \sqrt{\frac{1}{\sigma^2}\sum_{j=1}^{n}(X_j-\bar{X})^2/(n-1)} = Z/\sqrt{Y/(n-1)},$$

と変形する. ここで, 多次元正規分布の特性 4 と定理 3.2.1 から,

$$Z \sim \mathrm{N}(0,1), \quad Y \sim \chi^2_{n-1},$$

であることがわかる. また, 定理 3.2.5 から $\bar{X}$ と S^2 は独立であるので, Z と Y は独立. よって, T は自由度 $n-1$ の t-分布に従うことがわかる. □

(独立性の別解) 今, $V_j = (X_j-\mu)/\sigma$ とおくと, $V_1,\ldots,V_n \sim i.i.d.\ \mathrm{N}(0,1)$ となる. このとき,

$$Z = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} = \frac{\sqrt{n}}{\sigma}\left\{\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)\right\} = \sqrt{n}\cdot\frac{1}{n}\sum_{j=1}^{n}\frac{X_j-\mu}{\sigma} = \sqrt{n}\bar{V},$$

である. また, 式 (3.7) により

$$(n-1)S^2 = \sum_{j=1}^{n}(X_j-\mu)^2 - n(\bar{X}-\mu)^2,$$

であるので,

$$Y = \frac{1}{\sigma^2}\sum_{j=1}(X_j-\mu)^2 - \frac{n}{\sigma^2}(\bar{X}-\mu)^2 = \sum_{j=1}^{n}V_j^2 - n\bar{V},$$

である. ここで, 第 1 行が $(1/\sqrt{n},\ldots,1/\sqrt{n})$ となる n 次直交行列 $H(H'H = HH' = I_n)$ を用いて, 変換

$$\boldsymbol{U} = (U_1,\ldots,U_n)' = H\boldsymbol{V} = H(V_1,\ldots,V_n)',$$

を考える. (2.7) 式から, $V_1,\ldots,V_n \sim \mathrm{N}(0,1)$ であるので, $\boldsymbol{V} \sim \mathrm{N}_n(\boldsymbol{0}_n, I_n)$ となる. また, 多次元正規分布の特性 4 から,

$$\boldsymbol{U} \sim \mathrm{N}_n(H\boldsymbol{0}_n, HI_nH') = \mathrm{N}_n(\boldsymbol{0}_n, I_n),$$

が成り立ち, (2.7) 式から $U_1,\ldots,U_n \sim i.i.d.\ \mathrm{N}(0,1)$ であることがわかる. ここで,

$$U_1 = \frac{1}{\sqrt{n}}(V_1+\cdots+V_n) = \sqrt{n}\bar{V} = Z,$$

であり,

$$Y = \sum_{j=1}^{n}V_j^2 - n\bar{V}^2 = \boldsymbol{V}'\boldsymbol{V} - Z^2 = \boldsymbol{U}'HH'\boldsymbol{U} - U_1^2 = \sum_{j=2}^{n}U_j^2,$$

である. 今, $U_1,\ldots,U_n$ は互いに独立であるから, Z と Y は独立となる. □

t-分布の確率密度関数

$T \sim t_n$ のとき, T の確率密度関数は,

$$f(t;n) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2} \quad (n \in \mathbb{N}), \tag{4.1}$$

となる.

(*Proof*) 今, Z と Y をそれぞれ独立で $Z \sim \mathrm{N}(0,1)$, $Y \sim \chi_n^2$ となる確率変数とする. このとき, (Z,Y) の同時確率密度関数は, それぞれ独立であるので,

$$h(z,y) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \cdot \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1},$$

となる. ここで,

$$\begin{cases} T = Z/\sqrt{Y/n}, \\ U = Y, \end{cases}$$

といった変換を考える. この変換の逆変換は,

$$\begin{cases} Z = T\sqrt{U/n}, \\ Y = U, \end{cases}$$

であり, 逆変換に関するヤコビアンは,

$$J = \begin{vmatrix} \partial z/\partial t & \partial z/\partial u \\ \partial y/\partial t & \partial y/\partial u \end{vmatrix} = \begin{vmatrix} \sqrt{u/n} & t/(2\sqrt{un}) \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{u}{n}},$$

となるので, (T,U) の同時確率密度関数 $g(t,u)$ は, Z と Y の同時確率密度関数 $h(z,y)$ を用いると,

$$\begin{aligned} g(t,u) &= h\left(t\sqrt{u/n}, u\right)|J| \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\frac{u}{n}\right) \cdot \frac{1}{2^{n/2}\Gamma(n/2)} e^{-u/2} u^{n/2-1} \cdot \frac{1}{\sqrt{n}} u^{1/2} \\ &= \frac{1}{\sqrt{n\pi}2^{(n+1)/2}\Gamma(n/2)} \exp\left\{-\frac{1}{2}\left(1+\frac{t^2}{n}\right)u\right\} u^{(n-1)/2}, \end{aligned}$$

となる. T の確率密度関数は $g(t,u)$ の周辺確率密度関数なので,

$$f(t) = \int_0^\infty g(t,u)du = \int_0^\infty \frac{1}{\sqrt{n\pi}2^{(n+1)/2}\Gamma(n/2)} \exp\left\{-\frac{1}{2}\left(1+\frac{t^2}{n}\right)u\right\} u^{(n-1)/2} du,$$

で求めることができる. ここで, $(1+t^2/n)u = x$ とおくと, $dx/du = (1+t^2/n)$ であるので,

$$\begin{aligned} f(t) &= \int_0^\infty \frac{1}{\sqrt{n\pi}2^{(n+1)/2}\Gamma(n/2)} e^{-x/2} x^{(n-1)/2} \left(1+\frac{t^2}{n}\right)^{-(n-1)/2-1} dx \\ &= \frac{1}{\sqrt{n\pi}2^{(n+1)/2}\Gamma(n/2)} \left(1+\frac{t^2}{n}\right)^{-(n+1)/2} \int_0^\infty e^{-x/2} x^{(n+1)/2-1} dx \\ &= \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} \left(1+\frac{t^2}{n}\right)^{-(n+1)/2} \int_0^\infty \frac{1}{2^{(n+1)/2}\Gamma((n+1)/2)} e^{-x/2} x^{(n+1)/2-1} dx. \end{aligned}$$

上記の被積分関数は自由度 $n+1$ の χ^2 分布の確率密度関数であることに注意すれば, 積分値は 1 となる. よって, 式 (4.1) を得ることができる. □

定理 4.1.2: ${}^\forall t \in \mathbb{R}$ に対して,

$$\lim_{n\to\infty} f(t;n) = \phi(t).$$

(*Proof*) スターリングの公式より,

$$\Gamma(x) = \sqrt{2\pi} e^{-x} x^{x-1/2} \left\{1 + O(x^{-1})\right\},$$

であるので,

$$\begin{aligned} \log\Gamma\left(\frac{n}{2}\right) &= \frac{1}{2}\log 2\pi - \frac{n}{2} + \left(\frac{n}{2} - \frac{1}{2}\right)(\log n - \log 2) + O(n^{-1}), \\ \log\Gamma\left(\frac{n+1}{2}\right) &= \frac{1}{2}\log 2\pi - \frac{n+1}{2} + \left(\frac{n+1}{2} - \frac{1}{2}\right)\{\log(n+1) - \log 2\} + O(n^{-1}) \\ &= \frac{1}{2}\log 2\pi - \frac{n+1}{2} + \frac{n}{2}\{\log(n+1) - \log 2\} + O(n^{-1}). \end{aligned}$$

よって,

$$\log \frac{\Gamma((n+1)/2)}{\sqrt{n}\Gamma(n/2)} = \frac{1}{2}\log 2\pi - \frac{n+1}{2} + \frac{n}{2}\{\log(n+1) - \log 2\}$$
$$- \frac{1}{2}\log 2\pi + \frac{n}{2} - \left(\frac{n}{2} - \frac{1}{2}\right)(\log n - \log 2) - \frac{1}{2}\log n + O(n^{-1}).$$
$$= -\frac{1}{2} + \frac{n}{2}\log\left(1 + \frac{1}{n}\right) + \frac{1}{2}(\log n - \log 2) - \frac{1}{2}\log n + O(n^{-1})$$
$$= -\frac{1}{2} + \frac{1}{2} - \frac{1}{2}\log 2 + O(n^{-1}) = -\frac{1}{2}\log 2 + O(n^{-1}),$$

となるので,

$$\lim_{n\to\infty} \frac{\Gamma((n+1)/2))}{\sqrt{n}\Gamma(n/2)} = \frac{1}{\sqrt{2}}.$$

また,

$$\log\left\{\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}\right\} = \frac{-(n+1)}{2}\log\left(1+\frac{t^2}{n}\right) = \frac{-(n+1)}{2}\left\{1+\frac{t^2}{n}+O(n^{-2})\right\}$$
$$= -\frac{1}{2}t^2 + O(n^{-1}),$$

となるので,

$$\lim_{n\to\infty}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2} = e^{-t^2/2}.$$

以上により,

$$\lim_{n\to\infty} f(t;n) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2} = \phi(t).$$

よって, 題意が示せた. □

特性

1. 自由度 n の t-分布は, 標準正規分布と同様に, 0 を中心に左右対称な分布であるが, 分布の裾は標準正規分布よりも重くなる. また n が大きくなれば, t-分布は標準正規分布に近づく.

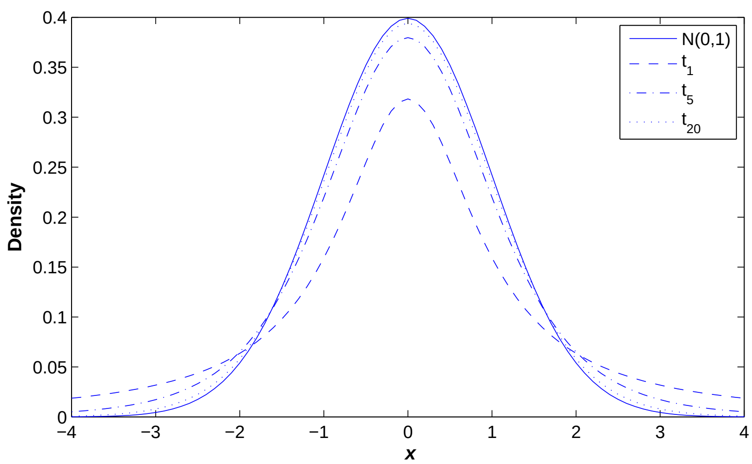

Figure 4.1. t-分布の形

2. $T \sim t_n$ のとき,

$$\mathrm{E}[T] = 0\ (n \geq 2), \quad \mathrm{Var}[T] = n/(n-2)\ (n \geq 3),$$

が成り立つ (自由度が 1 のときは平均と分散, 自由度が 2 のときは分散が存在しない). より一般的には, $n \geq r+1$ であれば,

$$\mathrm{E}[T^r] = \begin{cases} 0 & (r \text{ が奇数}) \\ \dfrac{1 \cdot 3 \times \cdots \times (r-1) n^{r/2}}{(n-2)(n-4) \times \cdots \times (n-r)} & (r \text{ が偶数}) \end{cases} . \tag{4.2}$$

(*Proof*) Z と Y は独立に $Z \sim \mathrm{N}(0,1)$, $Y \sim \chi^2_n$ である確率変数とすると, t-分布の定義より, $T = n^{-1/2} Y^{-1/2} Z$ とかける. それぞれの独立なので, r 次モーメントが存在するならば,

$$\mathrm{E}[T^r] = \mathrm{E}\left[n^{-r/2} Y^{-r/2} Z^r\right] = n^{-r/2} \mathrm{E}\left[Y^{-r/2}\right] \mathrm{E}\left[Z^r\right].$$

正規分布の特性 3 より, r が奇数であれば $\mathrm{E}[Z^r] = 0$ なので, $\mathrm{E}[T^r] = 0$. また, r が偶数の場合, 正規分布の特性 3 と χ^2 分布の特性 4 より,

$$n^{-r/2} \mathrm{E}\left[Y^{-r/2}\right] \mathrm{E}\left[Z^r\right] = \frac{1 \cdot 3 \times \cdots \times (r-1)}{n^{r/2}(n-2) \cdots (n-r)}.$$

以上により, (4.2) 式を得ることができる. また, (4.2) 式より, $\mathrm{E}[T] = 0$ であるので, $\mathrm{Var}[T] = \mathrm{E}[T^2] = n/(n-2)$ を得ることができる. □

注意

自由度が 1 である t-分布を特別に**コーシー分布** (**Cauchy Distribution**) と呼ぶ. コーシー分布はモーメントが一切存在しない分布の代表的なものである.

4.2. F-分布 (F-Distribution)

<u>**定義**</u>

互いに独立な確率変数 X と Y が $X \sim \chi_m^2$, $Y \sim \chi_n^2$ であるとき,

$$U = \frac{X/m}{Y/n},$$

の従う分布を自由度 m, n の F-分布といい, $U \sim F_{m,n}$ と書く.

<u>**F-分布の確率密度関数**</u>

$U \sim F_{m,n}$ のとき, U の確率密度関数は,

$$f(u;m,n) = \begin{cases} \dfrac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} \left(\dfrac{m}{n}\right)^{m/2} u^{m/2-1} \left(1+\dfrac{m}{n}u\right)^{-(m+n)/2} & (u>0) \\ 0 & (u \le 0) \end{cases} \qquad (m,n \in \mathbb{N}), \tag{4.3}$$

となる.

(*Proof*) 今, X と Y をそれぞれ独立で $X \sim \chi_m^2$, $Y \sim \chi_n^2$ となる確率変数とする. このとき, (X,Y) の同時確率密度関数は, それぞれ独立であるので,

$$\begin{aligned} h(x,y) &= \frac{1}{2^{m/2}\Gamma(m/2)} e^{-x/2} x^{m/2-1} \cdot \frac{1}{2^{n/2}\Gamma(n/2)} e^{-y/2} y^{n/2-1} \\ &= \frac{1}{2^{(m+n)/2}\Gamma(m/2)\Gamma(n/2)} e^{-x/2} x^{m/2-1} e^{-y/2} y^{n/2-1} \quad (x>0, y>0), \end{aligned}$$

となる. ここで,

$$\begin{cases} U = (X/m)/(Y/n), \\ V = Y, \end{cases}$$

といった変換を考える. この変換の逆変換は,

$$\begin{cases} X = (m/n)UV, \\ Y = V, \end{cases}$$

であり, 逆変換に関するヤコビアンは,

$$J = \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} (m/n)v & (m/n)u \\ 0 & 1 \end{vmatrix} = \frac{m}{n}v,$$

となるので, (U,V) の同時確率密度関数 $g(u,v)$ は, X と Y の同時確率密度関数 $h(x,y)$ を用いると,

$$\begin{aligned} g(u,v) &= h\left((m/n)uv, v\right)|J| \\ &= \frac{1}{2^{(m+n)/2}\Gamma(m/2)\Gamma(n/2)} e^{-(m/n)uv/2} \left(\frac{m}{n}uv\right)^{m/2-1} e^{-v/2} v^{n/2-1} \cdot \frac{m}{n}v \\ &= \frac{m}{n2^{(m+n)/2}\Gamma(m/2)\Gamma(n/2)} \exp\left\{-\frac{1}{2}\left(1+\frac{m}{n}u\right)v\right\} \left(\frac{m}{n}u\right)^{m/2-1} v^{(m+n)/2-1}, \end{aligned}$$

となる. U の確率密度関数は $g(u,v)$ の周辺確率密度関数なので,

$$\begin{aligned} f(u) &= \int_0^\infty g(u,v)dv \\ &= \frac{m}{n2^{(m+n)/2}\Gamma(m/2)\Gamma(n/2)} \int_0^\infty \exp\left\{-\frac{1}{2}\left(1+\frac{m}{n}u\right)v\right\} \left(\frac{m}{n}u\right)^{m/2-1} v^{(m+n)/2-1} dv \\ &= c(m,n) \int_0^\infty \exp\left\{-\frac{1}{2}\left(1+\frac{m}{n}u\right)v\right\} \left(\frac{m}{n}u\right)^{m/2-1} v^{(m+n)/2-1} dv \end{aligned}$$

で求めることができる. ここで, $(1+mu/n)v = z$ とおくと, $dz/dv = (1+mu/n)$ であるので,

$$\begin{aligned}
f(u) &= c(m,n)\int_0^{\infty} e^{-z/2}\left(\frac{m}{n}u\right)^{m/2-1}\left\{\left(1+\frac{m}{n}u\right)^{-1}z\right\}^{(m+n)/2-1}\left(1+\frac{m}{n}u\right)^{-1}dz \\
&= c(m,n)\int_0^{\infty} e^{-z/2}z^{(m+n)/2-1}\left(\frac{m}{n}u\right)^{m/2-1}\left(1+\frac{m}{n}u\right)^{-(m+n)/2}dz \\
&= c(m,n)2^{(m+n)/2}\Gamma((m+n)/2)\left(\frac{m}{n}\right)^{m/2-1}u^{m/2-1}\left(1+\frac{m}{n}u\right)^{-(m+n)/2} \\
&\quad\times\int_0^{\infty}\frac{1}{2^{(m+n)/2}\Gamma((m+n)/2)}e^{-z/2}z^{(m+n)/2-1}dz \\
&= \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}\left(\frac{m}{n}\right)^{m/2}u^{m/2-1}\left(1+\frac{m}{n}u\right)^{-(m+n)/2} \\
&\quad\times\int_0^{\infty}\frac{1}{2^{(m+n)/2}\Gamma((m+n)/2)}e^{-z/2}z^{(m+n)/2-1}dz.
\end{aligned}$$

上記の被積分関数は自由度 $m+n$ の χ^2 分布の確率密度関数であることに注意すれば, 積分値は 1 となる. よって, 式 (4.3) を得ることができる. □

特性

1. 自由度 m, n の F-分布は, 常に右に歪んだ分布であり, m が大きくなれば対称な分布に, n が大きくなれば自由度 m の χ^2 分布を m で割った分布に近づく.

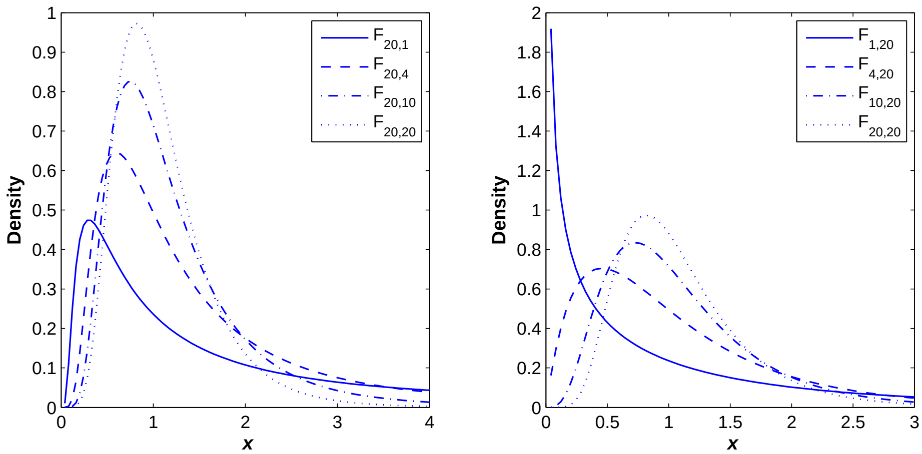

Figure 4.2. F-分布の形

2. $U \sim F_{m,n}$ のとき,

$$\mathrm{E}[U] = \frac{n}{n-2}\ (n \geq 3), \quad \mathrm{Var}[U] = \frac{2n^2(n+m-2)}{m(n-2)^2(n-4)}\ (n \geq 5),$$

が成り立つ. より一般的には, $n \geq 2r+1$ であれば,

$$\mathrm{E}[U^r] = \left(\frac{n}{m}\right)^r \frac{m(m+2)\cdots\{m+2(r-1)\}}{(n-2)\cdots(n-2r)}. \tag{4.4}$$

(*Proof*) 互いに独立な確率変数 X, Y $(X \sim \chi^2_m, Y \sim \chi^2_n)$ を用いて, $U = (X/m)/(Y/n)$ と書ける. それぞれの独立なので, U の r 次モーメントが存在するならば,

$$\mathrm{E}[U^r] = \left(\frac{n}{m}\right)^r \mathrm{E}[X^r]\cdot\mathrm{E}[Y^{-r}],$$

となる．ここで, χ^2 分布の特性 3, 4 より,

$$\mathrm{E}[X^r]\mathrm{E}[Y^{-r}] = \frac{m(m+2)\cdots\{m+2(r-1)\}}{(n-2)\cdots(n-2r)}.$$

よって，(4.4) 式を得ることができる．また, (4.2) 式より,

$$\mathrm{E}[U] = \frac{n}{m}\frac{m}{n-2} = \frac{n}{n-2} \quad (n \geq 3),$$

を得る．同様に,

$$\mathrm{E}[U^2] = \frac{n^2}{m^2}\frac{m(m+2)}{(n-2)(n-4)} = \frac{n^2(m+2)}{m(n-2)(n-4)} \quad (n \geq 5),$$

を得る．よって, U の分散は,

$$\begin{aligned}\mathrm{Var}[U] &= \mathrm{E}[U^2] - (\mathrm{E}[U])^2 = \frac{n^2(m+2)}{m(n-2)(n-4)} - \frac{n^2}{(n-2)^2} \\ &= \frac{n^2}{m(n-2)^2(n-4)}\{(n-2)(m+2) - m(n-4)\} = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad (n \geq 5),\end{aligned}$$

となる． □

3. $T \sim t_n$ のとき, $T^2 \sim F_{1,n}$ である．

(*Proof*) t-分布の定義から, $T = Z/\sqrt{Y/n}$ (Z と Y は独立, $Z \sim \mathrm{N}(0,1)$, $Y \sim \chi_n^2$) と書けるので, $T^2 = Z^2/(Y/n)$. ここで, $Z^2 \sim \chi_1^2$ であるので, $T^2 \sim F_{1,n}$ がわかる． □

4.3. 一元配置計画と分散分析

<u>問題設定</u>

- データ: ある湖の 4 地点で取られた酸素溶解量の無作為標本

地点	酸素溶解量 (%)						平均
A_1	7.8	6.4	8.2	6.9			7.324
A_2	6.7	6.8	7.1	6.9	7.3		6.960
A_3	7.2	7.4	6.9	6.4	6.5		6.880
A_4	6.0	7.4	6.5	6.9	7.2	6.8	6.800

- 問題: 4 地点における酸素融解量に有意な差はあるか?
- モデル化: 要因 A について k 個の水準 $A_1,\dots,A_k$ を考え, それぞれの水準における標本を

$$\begin{array}{ccc} \text{第 1 水準} & : & Y_{11},\dots,Y_{1n_1} \\ \vdots & & \vdots \\ \text{第 } k \text{ 水準} & : & Y_{k1},\dots,Y_{kn_k} \end{array},$$

とし, $Y_{i1},\dots,Y_{in_i}$ に関する統計モデルを,

$$Y_{i1},\dots,Y_{in_i} \sim i.i.d.\ \mathrm{N}(\mu_i,\sigma^2) \quad (i=1,\dots,k),$$

とする. このとき, 考えている問題は,

$$\begin{array}{lcl} H_0 & : & \mu_1=\cdots=\mu_k \ (\text{仮説}) \\ H_1 & : & \text{not } H_0 \ (\mu_1,\dots,\mu_k \text{ のうちどれかは異なる}) \end{array},$$

という検定問題になり, この検定における検定統計量を $T(\boldsymbol{Y})$ ($\boldsymbol{Y}=(Y_{11},\dots,Y_{1n_1},\dots,Y_{k1},\dots,Y_{kn_k})'$) とすると, $T(\boldsymbol{Y})\geq c$ であれば H_0 を棄却 (H_1 を受容) することになる. ここでの c は H_0 が正しいという仮定の下ので $\mathrm{P}(T(\boldsymbol{Y})\geq c)\leq\alpha$ (α はほとんどの場合 0.05 か 0.01) を満たす閾値である.

<u>基本統計量</u>

検定統計量を考えるために, 以下のような統計量を考える.

- 残差平方和: $\bar{Y}_i=n_i^{-1}\sum_{j=1}^{n_i}Y_{ij}$ とおくと,

$$S^2=\min_{(\mu_1,\dots,\mu_k)\in\mathbb{R}^k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\mu_i)^2=\sum_{i=1}^{k}\min_{\mu_i\in\mathbb{R}}\sum_{j=1}^{n_i}(Y_{ij}-\mu_i)^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\bar{Y}_i)^2.$$

- H_0 が正しいという仮定の下での残差平方和: $\bar{Y}=n^{-1}\sum_{i=1}^{k}\sum_{j=1}^{n_i}Y_{ij}$ ($n=n_1+\cdots+n_k$) とおくと,

$$S_0^2=\min_{\mu\in\mathbb{R}}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\mu)^2=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\bar{Y})^2.$$

- 群内平方和: $W=S^2$.
- 群間平方和:

$$\begin{aligned} B=S_0^2-S^2&=\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left\{(Y_{ij}-\bar{Y})^2-(Y_{ij}-\bar{Y}_i)^2\right\} \\ &=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{Y}_i-\bar{Y})(2Y_{ij}-\bar{Y}-\bar{Y}_i)=\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{Y}_i-\bar{Y})^2. \end{aligned}$$

<u>検定統計量と帰無分布</u>

基本統計量により,

$$S_0^2 = B + S^2 = B + W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{Y}_i - \bar{Y})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y})^2,$$

が成り立つ. ここで, H_0 が正しければ, S^2 と S_0^2 の値は近くなり, その結果, B の値は小さくなる. よって, B が W に比べ大きければ H_0 を棄却 ($\mu_1 = \cdots = \mu_k$ ではないと) する検定法を考えることができる. つまり, 以下の検定統計量を用いる.

$$T = \frac{B/(k-1)}{W/(n-k)}.$$

このとき, 以下の定理が成り立つ. この定理により, 検定には, F-分布のパーセント点を閾値に用いればよいことがわかる.

定理 4.3.1: $Y_{11}, \ldots, Y_{1n_1}, \ldots, Y_{k1}, \ldots, Y_{kn_k}$ は互いに独立で, $Y_{i1}, \ldots, Y_{in_i} \sim i.i.d.\ \mathrm{N}(\mu_i, \sigma^2)$ $(i = 1, \ldots, k)$ とする. このとき, $\mu_1 = \cdots = \mu_k$ であれば, B と W は独立で,

$$T = \frac{B/(k-1)}{W/(n-k)} \sim F_{k-1, n-k}.$$

(*Proof*) $Y_{ij} = \mu_i + \varepsilon_{ij}$ とおくと,

$$\varepsilon_{11}, \ldots, \varepsilon_{1n_1}, \ldots, \varepsilon_{k1}, \ldots, \varepsilon_{kn_k} \sim i.i.d.\ \mathrm{N}(0, \sigma^2),$$

となる. $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \ldots, \varepsilon_{1n_1}, \ldots, \varepsilon_{k1}, \ldots, \varepsilon_{kn_k})'$ を用いると, $\bar{Y}_i$ と $\bar{Y}$ は,

$$\bar{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij} = \frac{1}{n_i}\sum_{j=1}^{n_i}(\mu_i + \varepsilon_{ij}) = \mu_i + \bar{\varepsilon}_i,$$
$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\mu_i + \varepsilon_{ij}) = \bar{\mu} + \bar{\varepsilon},$$

となる. よって, W と B は,

$$W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\varepsilon_{ij} - \bar{\varepsilon}_i)^2,$$
$$B = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\bar{Y}_i - \bar{Y})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\mu_i + \bar{\varepsilon}_i - \bar{\mu} - \bar{\varepsilon})^2 = \sum_{i=1}^{k} n_i(\mu_i - \bar{\mu} + \bar{\varepsilon}_i - \bar{\varepsilon})^2,$$

と書き換えることができる. ここで, $n \times k$ 行列 $D = \mathrm{diag}(\mathbf{1}_{n_1}, \ldots, \mathbf{1}_{n_k})$ を用いると, $(\bar{\varepsilon}_1, \ldots, \bar{\varepsilon}_k)' = (D'D)^{-1}D'\boldsymbol{\varepsilon}$ となる. よって,

$$\begin{aligned} W &= \left\{\boldsymbol{\varepsilon} - D(D'D)^{-1}D'\boldsymbol{\varepsilon}\right\}'\left\{\boldsymbol{\varepsilon} - D(D'D)^{-1}D'\boldsymbol{\varepsilon}\right\} \\ &= \boldsymbol{\varepsilon}'\left\{I_n - D(D'D)^{-1}D'\right\}^2\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'\left\{I_n - D(D'D)^{-1}D'\right\}\boldsymbol{\varepsilon}, \end{aligned}$$

となる. $I_n - D(D'D)^{-1}D'$ はベキ等行列で, $\mathrm{tr}\{I_n - D(D'D)^{-1}D'\} = n - k$ となることに注意すれば, $W/\sigma^2 \sim \chi^2_{n-k}$ となる. また, もし H_0 が正しければ, $\mu_1 = \cdots = \mu_k = \mu$ なので, $\bar{\mu} = \mu$ が成り立つ. その結果, B は, $B = \sum_{i=1}^{k} n_i(\bar{\varepsilon}_i - \bar{\varepsilon})^2$ となる. ここで, D と $\mathbf{1}_n$ を用いると,

$$\begin{aligned} B &= \left\{D(D'D)^{-1}D'\boldsymbol{\varepsilon} - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'\boldsymbol{\varepsilon}\right\}'\left\{D(D'D)^{-1}D'\boldsymbol{\varepsilon} - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'\boldsymbol{\varepsilon}\right\} \\ &= \boldsymbol{\varepsilon}'\left\{D(D'D)^{-1}D - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'\right\}^2\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'\left\{D(D'D)^{-1}D - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'\right\}\boldsymbol{\varepsilon}, \end{aligned}$$

となる. $D(D'D)^{-1}D - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'$ はベキ等行列で, $\mathrm{tr}\{D(D'D)^{-1}D - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'\} = k-1$ となることに注意すれば, $B/\sigma^2 \sim \chi^2_{k-1}$ となる. さらに,

$$\{I_n - D(D'D)^{-1}D'\}\{D(D'D)^{-1}D' - \mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'\} = O_{n,n},$$

であるので, B と W は独立となることがわかる. よって, F-分布の定義より, $T \sim F_{k-1,n-k}$ となる. □

検定結果

酸素溶解量のデータで仮説検定を行ってみると, $n = 20$, $k = 4$ で, $F_{3,16}$ の上側 5%点は, 3.2389 である. $T = (0.7180/3)/(4.2675/16) = 0.8973$ となるので, 有意水準 5%で仮説を棄却できない.