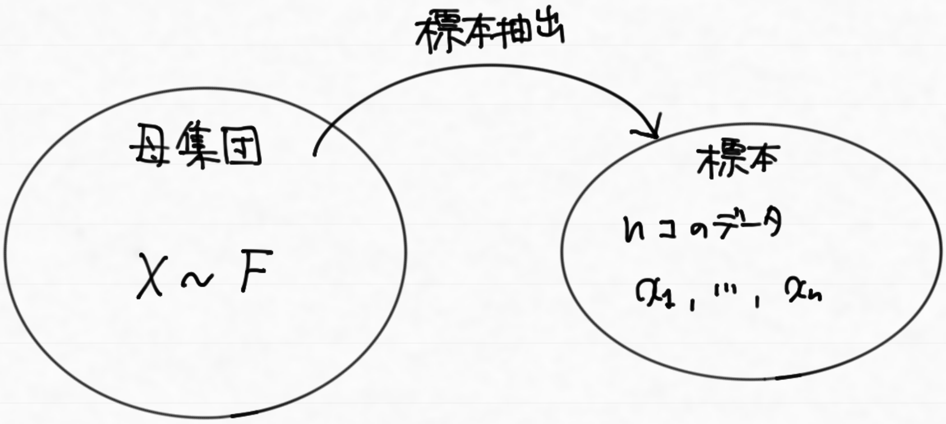
Ch.5 標本分布とその近似

§5.1 統計量と標本分布



Def.（ランダム・サンプル）

$X_1, \cdots, X_n$: ランダム・サンプル

$\overset{\text{def.}}{\iff} \quad X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$ $\square$

・母数：母集団の特性値（パラメータ）

・母平均 μ ：母集団の平均

・母分散 σ^2 ：母集団の分散

・μの推定： $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ （標本平均）

σ^2の推定： $S^2 := \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2$ （標本分散）

・実際には、$\bar{X}, S^2$ は X_i の実現値 x_i で推定される。

→ 推定誤差や信頼区間など、その推定値の信頼性を与えるために、$\bar{X}, S^2$ の分布を求める必要がある。

Def.（統計量，標本分布）

$\{X_i\}_{i=1}^n$ の関数 $t(X_1, \cdots, X_n)$ を**統計量**といい，

その確率分布を**標本分布**という．　$\square$

・母平均 μ，母分散 σ^2 の母集団からのランダム・サンプリング：

$X_1, \cdots, X_n \overset{i.i.d.}{\sim} (\mu, \sigma^2)$　とかく．

・$\mathbb{E}[X_i] = \mu$，$\mathbb{V}[X_i] = \sigma^2$．

・$\mathbb{E}[\overline{X}] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \frac{1}{n}\sum_{i=1}^n \mu = \mu$．

$\mathbb{V}[\overline{X}] = \mathbb{V}\left[\sum_{i=1}^n \frac{1}{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}[X_i] = \frac{\sigma^2}{n}$

。$\mathbb{V}[\overline{X}] = \mathbb{E}[(\overline{X}-\mu)]$ なので，$\mathbb{V}[\overline{X}]$ は $\overline{X}$ が μ から平均的にどの程度離れているかを表す．$n \to \infty$ で $\mathbb{V}[\overline{X}] \to 0$．

・$\mathbb{E}[S^2] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2\right]$

$$= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n (X_i - \mu + \mu - \overline{X})^2\right]$$

$$= \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n (X_i-\mu)^2\right] + \frac{2}{n}\mathbb{E}\left[\sum_{i=1}^n (X_i-\mu)(\mu-\overline{X})\right] + \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^n (\mu-\overline{X})^2\right]$$

$$= \frac{1}{n}\sum_{i=1}^n \mathbb{V}[X_i] - \frac{2}{n}\sum_{i=1}^n \mathrm{Cov}(X_i, \overline{X}) + \frac{1}{n}\sum_{i=1}^n \mathbb{V}[\overline{X}]$$

$$= \frac{1}{n}\sum_{i=1}^n \sigma^2 - \frac{2}{n}\sum_{i=1}^n (\mathbb{E}[X_i\overline{X}] - \mu^2) + \frac{1}{n}\sum_{i=1}^n \frac{\sigma^2}{n}$$

$$= \frac{n+1}{n}\sigma^2 - \frac{2}{n}\sum_{i=1}^{n}\left(\mathbb{E}[X_i\bar{X}] - \mu^2\right)$$

$$= \frac{n+1}{n}\sigma^2 - \frac{2}{n}\mathbb{E}\left[\sum_{i=1}^{n}X_i\bar{X}\right] + 2\mu^2$$

$$= \frac{n+1}{n}\sigma^2 - 2\mathbb{E}[\bar{X}^2] + 2\mu^2$$

$$= \frac{n+1}{n}\sigma^2 - 2\left(\mathbb{E}[\bar{X}^2] - \mu^2\right)$$

$$= \frac{n+1}{n}\sigma^2 - 2\mathbb{V}[\bar{X}]$$

$$= \frac{n-1}{n}\sigma^2 .$$

→ S^2の期待値はσ^2にならない。

期待値をσ^2にするには，nで割らずに$n-1$で割る必要がある。

<u>Def.</u>（不偏分散）

$V^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ を**不偏分散**という。 $\square$

§5.2 正規母集団からの代表的な標本分布.

5.2.1. 標本平均と標本分散の独立性

Th. $X_1, \cdots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$.

(1) $\overline{X} \perp\!\!\!\perp V^2$.

(2) $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

(3) $\frac{n-1}{\sigma^2} V^2 \sim \chi^2_{n-1}$. $\square$

pf $Z_i = \frac{X_i - \mu}{\sigma}$ とすると, $Z_i \overset{i.i.d.}{\sim} N(0,1)$

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

$\sum_{i=1}^{n}(Z_i - \overline{Z})^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{n-1}{\sigma^2}V^2$.

上記(1)〜(3)は

(1)' $\overline{Z} \perp\!\!\!\perp \sum_{i=1}^{n}(Z_i - \overline{Z})^2$

(2)' $\sqrt{n}\,\overline{Z} \sim N(0,1)$

(3)' $\sum_{i=1}^{n}(Z_i - \overline{Z})^2 \sim \chi^2_{n-1}$

とかける. これらを示す.

Helmert行列を利用して示す:

$$H := \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \cdots & 0 \\ -\frac{1}{\sqrt{2\cdot 3}} & -\frac{1}{\sqrt{2\cdot 3}} & \frac{2}{\sqrt{2\cdot 3}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{\sqrt{(n-1)n}} & -\frac{1}{\sqrt{(n-1)n}} & -\frac{1}{\sqrt{(n-1)n}} & \cdots & \frac{n-1}{\sqrt{(n-1)n}} \end{pmatrix}$$

[Claim] H：直交行列．

pf. 略．

$\mathbb{Z} = (Z_1, \cdots, Z_n)^T$ の $\mathbb{z} = (z_1, \cdots, z_n)^T$ における joint pdf. は

$$f_{\mathbb{Z}}(\mathbb{z}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_i^2}{2}\right)$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{i=1}^{n} z_i^2\right)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\mathbb{z}^T\mathbb{z}\right)$$

変数変換 $\mathbb{Y} = H\mathbb{Z}$ を考える．$\mathbb{Z} = H^T\mathbb{Y}$．

$$f_{\mathbb{Y}}(\mathbb{y}) = f_{\mathbb{Z}}(H^T\mathbb{y})\,|J(\mathbb{y})|$$

$$J(\mathbb{y}) = \det(H^T).$$

$H^TH = I$ より $\det(H^T)\det(H) = 1$

$\det(H^T) = \det(H)$ なので，$\det(H^T) = \pm 1$．

$\therefore |J(\mathbb{y})| = 1$．

$$f_Y(y) = f_Z(H^T y)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2} y^T H^T H y\right)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2} y^T y\right)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_i^2}{2}\right)$$

$\therefore Y_1, \cdots, Y_n \overset{i.i.d.}{\sim} N(0,1)$.

$$Y = HZ = \begin{pmatrix} \sum_{i=1}^{n} \frac{Z_i}{\sqrt{n}} \\ \\ * \\ \\ \end{pmatrix} = \begin{pmatrix} \sqrt{n}\,\overline{Z} \\ \\ * \\ \\ \end{pmatrix}.$$

$\therefore Y_1 = \sqrt{n}\,\overline{Z} \sim N(0,1)$. （(2)'が示せた）

また，

$$\sum_{i=1}^{n} (Z_i - \overline{Z})^2 = \sum_{i=1}^{n} Z_i^2 - 2\sum_{i=1}^{n} Z_i \overline{Z} + n\overline{Z}^2$$

$$= \sum_{i=1}^{n} Z_i^2 - n\overline{Z}^2$$

$$= Z^T Z - Y_1^2$$

$$= Y^T Y - Y_1^2$$

$$= Y_2^2 + \cdots + Y_n^2 \sim \chi^2_{n-1}$$ （(3)'が示せた）

$\therefore \overline{Z} \perp\!\!\!\perp \sum_{i=1}^{n} (Z_i - \overline{Z})^2$. （(1)'が示せた）

$\boxed{/\!/\!/}$

5.2.2 t-分布.

・ $X_i \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma^2)$

$$\to \bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\therefore Z = \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim \mathcal{N}(0,1).$$

・ σ のかわりに V を代入したt統計量 $T = \dfrac{\sqrt{n}(\bar{X}-\mu)}{V}$ の分布は？

・ $U := \dfrac{n-1}{\sigma^2}V^2$ とおくと. $U \sim \chi^2_{n-1}$.

$$V^2 = \frac{\sigma^2}{n-1}U, \quad V = \sigma\sqrt{\frac{1}{n-1}U}$$

$$\therefore T = \frac{\sigma Z}{\sigma\sqrt{\frac{U}{n-1}}} = \frac{Z}{\sqrt{\frac{U}{n-1}}}$$

Def. (t-分布)

$Z \sim \mathcal{N}(0,1)$, $U \sim \chi^2_m$, $Z \perp\!\!\!\perp U$.

このとき, $T = \dfrac{Z}{\sqrt{\frac{U}{m}}}$

は, 自由度 m の Studentのt-分布 に従うといい,

$T \sim t_m$ とかく. $\square$

Prop. $T \sim t_m$ のとき、T の pdf は

$$f_T(t|m) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} \frac{1}{\sqrt{\pi m}} \frac{1}{(1+t^2/m)^{\frac{m+1}{2}}}. \qquad \square$$

- $t_1 = C(0,1)$.

Prop. $t_m \xrightarrow{d} N(0,1) \quad (\text{as } m \to \infty)$. $\square$

5.2.3 F-分布

Def. (F-分布)

$S \sim \chi_m^2$, $T \sim \chi_n^2$, $S \perp\!\!\!\perp T$.

$$Y = \frac{\frac{S}{m}}{\frac{T}{n}}$$

は自由度 (m,n) の Snedecor の F-分布に従うといい、

$Y \sim F_{m,n}$ とかく。 $\square$

Prop. $Y \sim F_{m,n}$. Y の pdf は

$$f_Y(y|m,n) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}} y^{\frac{m}{2}-1}}{\left(1+\left(\frac{m}{n}\right)y\right)^{\frac{m+n}{2}}}, \quad y>0.$$

$$W = \frac{S}{S+T} \sim \mathrm{Beta}\left(\frac{m}{2}, \frac{n}{2}\right). \qquad \square$$

§5.3 確率変数と確率分布の収束.

5.3.1 確率収束と大数の弱法則

Def.（確率収束）

$\{U_n\}_{n=1}^{\infty}$: r.v.列 : U : r.v. に確率収束

$$\overset{\text{def.}}{\iff} \ {}^{\forall}\varepsilon>0, \ \lim_{n\to\infty} P(|U_n-U| \geq \varepsilon) = 0.$$

これを $U_n \xrightarrow{P} U$ とかく. $\square$

・確率収束を示すのに，次の不等式が便利：

Th.（Markovの不等式）

$Y \geq 0$: r.v., $\mathbb{E}[Y] < \infty$.

$$\Rightarrow {}^{\forall}\varepsilon>0, \quad P(Y \geq \varepsilon) \leq \frac{1}{\varepsilon}\mathbb{E}[Y]. \quad \square$$

pf. $I(Y \geq c) := \begin{cases} 1 & (Y \geq c) \\ 0 & (\text{otherwise}) \end{cases}$

とすると,

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}[YI(Y\geq c) + YI(Y<c)] \\
&\geq \mathbb{E}[YI(Y \geq c)] \\
&\geq \mathbb{E}[cI(Y \geq c)] \\
&= cP(Y \geq c). \quad \blacksquare
\end{aligned}$$

Th. (Chebyshevの不等式)

X : r.v. $\mu = \mathbb{E}[X]$, $\sigma^2 = \mathbb{V}[X]$ は存在.

$\Rightarrow \forall k \neq 0$, $P(|X-\mu| \geq k) \leq \dfrac{\sigma^2}{k^2}$. □

pf. Markov's ineq. で $Y = (X-\mu)^2$, $\varepsilon = k^2$ とする. ▨

Th. (大数の弱法則)

$X_1, X_2, \cdots \overset{i.i.d.}{\sim} (\mu, \sigma^2)$.

$\Rightarrow \bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu$ (as $n \to \infty$). □

pf. $\mathbb{E}[\bar{X}] = \mu$, $\mathbb{V}[\bar{X}] = \dfrac{\sigma^2}{n}$.

Chebyshev's ineq. より. $\forall \varepsilon > 0$,

$0 \leq P(|\bar{X}-\mu| \geq \varepsilon) \leq \dfrac{\sigma^2}{n\varepsilon} \xrightarrow[n\to\infty]{} 0$ ▨

Def. (平均2乗収束)

$\{U_n\}_{n=1}^{\infty}$: r.v.列 : U : r.v. に平均2乗収束

$\overset{def.}{\Longleftrightarrow} \lim_{n\to\infty} \mathbb{E}[|U_n - U|^2] = 0$.

これを $U_n \xrightarrow{L^2} U$ とかく. □

Prop. $U_n \xrightarrow{L^2} U \Rightarrow U_n \xrightarrow{p} U$. □

pf. Chebyshev's ineq. を用いる.

$^\forall \varepsilon > 0,$

$$0 \leq P(|U_n - U| \geq \varepsilon) \leq \frac{\mathbb{E}[|U_n - U|^2]}{\varepsilon^2} \xrightarrow[n\to\infty]{} 0.$$

5.3.2 法則収束と中心極限定理

Def. (法則収束)

$\{U_n\}_{n=1}^{\infty}$: r.v.列 : U : r.v. に法則収束

$\overset{\text{def.}}{\Longleftrightarrow}$ F_n : U_n の cdf. F_U : U の cdf.. として,

$$\lim_{n\to\infty} F_n(x) = F_U(x) \quad \text{for } ^\forall x : F_U(x) \text{ の連続点}.$$

これを $U_n \xrightarrow{d} U$ とかく。

Th. (中心極限定理)

$$X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} (\mu, \sigma^2). \qquad Z \sim N(0,1)$$

$$\Rightarrow \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} Z.$$

pf. $Z_i = \dfrac{X_i - \mu}{\sigma}$ とおくと, $\mathbb{E}[Z_i] = 0$, $\mathbb{V}[Z_i] = 1$.

$\bar{Z} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} Z_i$ であり, $\mathbb{E}[\bar{Z}] = 0$, $\mathbb{V}[\bar{Z}] = \dfrac{1}{n}$.

$$\bar{Z} = \frac{1}{n}\cdot\frac{1}{\sigma}\left(\sum_{i=1}^{n} X_i - n\mu\right) = \frac{\bar{X} - \mu}{\sigma} \text{ なので},$$

示すことは. $\sqrt{n}\,\bar{Z} \xrightarrow{d} Z \sim N(0,1)$.

$$\phi_{\sqrt{n}\bar{Z}}(t) = \mathbb{E}\left[e^{i\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i\right)t}\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{i\frac{t}{\sqrt{n}}Z_i}\right] = \left(\mathbb{E}\left[e^{i\frac{t}{\sqrt{n}}Z_1}\right]\right)^n.$$

$\phi_{Z_1}(\theta) := \mathbb{E}[e^{i\theta Z_1}]$ とおくと． $\phi_{\sqrt{n}\bar{Z}}(t) = \left(\phi_{Z_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$.

n：十分大のとき，$\frac{t}{\sqrt{n}}$ は微小量．Maclaurin展開して，

$$\phi_{Z_1}\left(\frac{t}{\sqrt{n}}\right) = \phi_{Z_1}(0) + \phi_{Z_1}'(0)\frac{t}{\sqrt{n}} + \phi_{Z_1}''(0)\frac{t^2}{2n} + o(n^{-1})$$

$$\phi_{Z_1}(0) = \mathbb{E}[e^0] = 1.$$

$$\frac{1}{i}\phi_{Z_1}'(0) = \mathbb{E}[Z_1] = 0,$$

$$\frac{1}{i^2}\phi_{Z_1}''(0) = \mathbb{E}[Z_1^2] = 1.$$

$$\therefore \phi_{Z_1}\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + o(n^{-1})$$

$$\therefore \phi_{\sqrt{n}\bar{Z}}(t) = \left(1 - \frac{t^2}{2n} + o(n^{-1})\right)^n \xrightarrow[n\to\infty]{} e^{-\frac{t^2}{2}} = \phi_Z(t).$$ ▨

5.3.3 収束に関する諸結果

<u>Prop.</u> $\{U_n\}_{n=1}^{\infty}$: r.v.列． U : r.v.

(1) $U_n \xrightarrow{p} U \Rightarrow U_n \xrightarrow{d} U$

(2) a : const. $U_n \xrightarrow{d} a \Rightarrow U_n \xrightarrow{p} a$. □

<u>Th.</u> (portmanteauの定理)

$\{U_n\}_{n=1}^{\infty}$: r.v.列． U : r.v.

$$U_n \xrightarrow{d} U$$

$\Leftrightarrow \forall f$: 有界連続，$\lim_{n\to\infty} \mathbb{E}[f(U_n)] = \mathbb{E}[f(U)]$. □

Th. (連続写像定理)

$\{U_n\}_{n=1}^{\infty}$: r.v.列. U : r.v. , h : continuous.

(1) $U_n \xrightarrow{P} U \Rightarrow h(U_n) \xrightarrow{P} h(U)$.

(2) $U_n \xrightarrow{d} U \Rightarrow h(U_n) \xrightarrow{d} h(U)$ $\square$

Th. (Slutskyの定理)

$\{U_n\}_{n=1}^{\infty}, \{V_n\}_{n=1}^{\infty}$: r.v.列. U : r.v. , a : const.

$U_n \xrightarrow{d} U$, $V_n \xrightarrow{P} a$.

(1) $U_n + V_n \xrightarrow{d} U + a$.

(2) $U_n V_n \xrightarrow{d} aU$. $\square$

Th. (デルタ法)

$\{U_n\}_{n=1}^{\infty}$: r.v.列. θ : const. $\{a_n\}_{n=1}^{\infty}$: $a_n \nearrow \infty$.

$a_n(U_n - \theta) \xrightarrow{d} U$

$g \in C^1$: $g'(\theta) \neq 0$.

$\Rightarrow a_n(g(U_n) - g(\theta)) \xrightarrow{d} g'(\theta)U$.

特に, $\sqrt{n}(U_n - \mu) \xrightarrow{d} N(0, \sigma^2)$ のとき

$$\sqrt{n}(g(U_n) - g(\mu)) \xrightarrow{d} N(0, \sigma^2 g'(\mu)^2). \quad \square$$

pf. $g(U_n)$ を $U_n=\theta$ のまわりでTaylor展開.

$$g(U_n) = g(\theta) + g'(\theta^*)(U_n-\theta).$$

ここで, θ^* は $|\theta^*-\theta| < |U_n-\theta|$ となる点.

$a_n(g(U_n)-g(\theta)) = a_n g'(\theta^*)(U_n-\theta)$ なので,

$a_n g'(\theta^*)(U_n-\theta)$ の極限分布を求める.

Slutsky's Th. より,

$$U_n-\theta = \frac{1}{a_n}\cdot a_n(U_n-\theta) \xrightarrow{d} 0\cdot U = 0$$ なので,

$U_n \xrightarrow{P} \theta$. よって, $\theta^* \xrightarrow{P} \theta$.

$g' \in C$ なので, $g'(\theta^*) \xrightarrow{P} g'(\theta)$.

Slutsky's Th. より

$$g'(\theta^*)\, a_n(U_n-\theta) \xrightarrow{d} g'(\theta)U.$$ ▨

・ $g'(\theta)=0$ のときには, 2次のデルタ法が用いられる:

Th. (2次のデルタ法)

$\{U_n\}_{n=1}^{\infty}$: r.v.列. θ : const. $\{a_n\}_{n=1}^{\infty}$: $a_n \nearrow \infty$.

$a_n(U_n-\theta) \xrightarrow{d} U$. $g \in C^2$: $g'(\theta)=0$, $g''(\theta)\neq 0$.

$$\Rightarrow a_n^2(g(U_n)-g(\theta)) \xrightarrow{d} \frac{g''(\theta)}{2}U^2.$$

特に、$\sqrt{n}(U_n-\mu) \xrightarrow{d} N(0,\sigma^2)$ のとき、$Y \sim \chi_1^2$ として

$$n(g(U_n)-g(\mu)) \xrightarrow{d} \frac{g''(\theta)}{2}\sigma^2 Y. \qquad \square$$

・デルタ法の $\sqrt{n}(U_n-\mu) \xrightarrow{d} N(0,\sigma^2)$ について.

$\sigma^2 = V(\mu)$ となるとき,

$$\sqrt{n}(g(U_n)-g(\mu)) \xrightarrow{d} N(0, V(\mu)g'(\mu)^2).$$

→分散が μ に依存する.

g として, $(g'(\mu))^2 V(\mu) = 1$ or $g'(\mu) = \frac{1}{\sqrt{V(\mu)}}$

となる g をみつければ, 漸近分布が μ に依存しない.

このような g を **分散安定化変換** という.

ex $X_1, X_2, \dots \overset{i.i.d.}{\sim} B(1,\theta)$

$$\mathbb{E}[\overline{X}] = \theta, \quad \mathbb{V}[\overline{X}] = \frac{\theta(1-\theta)}{n}$$

C.L.T. $\sqrt{n}(\overline{X}-\theta) \xrightarrow{d} N(0, \theta(1-\theta)).$

$g'(\theta) = \frac{1}{\sqrt{\theta(1-\theta)}}$ となる g では, デルタ法から

$$\sqrt{n}(g(\overline{X})-g(\theta)) \xrightarrow{d} N(0,1).$$

そのような g は, $g(\theta) = \arcsin(2\theta-1)$. $\square$

ex $X_1, X_2, \cdots \overset{i.i.d.}{\sim} Po(\lambda)$

$$\sqrt{n}(\bar{X}-\lambda) \xrightarrow{d} N(0, \lambda)$$

$$g'(\lambda) = \frac{1}{\sqrt{\lambda}} \qquad \therefore g(\lambda) = 2\sqrt{\lambda}$$

$$\sqrt{n}(2\sqrt{\bar{X}} - 2\sqrt{\lambda}) \xrightarrow{d} N(0, 1).$$

i.e., $\sqrt{n}(\sqrt{\bar{X}} - \sqrt{\lambda}) \xrightarrow{d} N(0, \frac{1}{4}).$ $\square$

5.4 順序統計量

Def.（順序統計量）

$X_1, \cdots, X_n \overset{i.i.d.}{\sim} P$. X_i を小さい順に並び変えたものを

$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ と表し，順序統計量 という. $\square$

- $X_{(1)} = \min_i X_i$: 最小統計量

  $X_{(n)} = \max_i X_i$: 最大統計量

  $R := X_{(n)} - X_{(1)}$: 標本範囲

  $$M := \begin{cases} X_{(\frac{n+1}{2})} & (n : \text{odd}) \\ \dfrac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} & (n : \text{even}) \end{cases}$$: メディアン

Prop. $X_{(j)}$ の cdf. は，X_i の cdf. を F とすると

$$F_{X_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} (F(x))^k (1-F(x))^{n-k}. \quad \square$$

pf. $X_j \leq x$ となる j の個数を Y とおき，

$X_j \leq x$ のとき"成功"，$X_j > x$ のとき"失敗"と考えると，

$Y \sim B(n, F(x))$.

$$P(X_{(j)} \leq x) = P(Y \geq j)$$

$$= \sum_{k=j}^{n} \binom{n}{k} (F(x))^k (1-F(x))^{n-k}.$$ ▨

Prop. $X_{(j)}$ の pdf. は, X_i の pdf を f とすると,

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f(x)\left(F(x)\right)^{j-1}\left(1-F(x)\right)^{n-j}. \qquad \square$$

pf. $F_{X_{(j)}}(x)$ を微分する. $\blacksquare$

・Prop. の解釈.

$X_{(j)} = x$ となる確率は,

$\overset{\text{prob. } F(x)}{\{X_k < x\}}$ が $j-1$ 回, $\overset{\text{prob. } f(x)}{\{X_k = x\}}$ が 1回, $\overset{\text{prob. } 1-F(x)}{\{X_k > x\}}$ が $n-j$ 回

起きる3項分布とみることができるので

Prop. の式が得られる.

・同様に, $X_{(i)}$ と $X_{(j)}$ の joint pdf. は. 5項分布とみることができ,

$$f_{X_{(i)}, X_{(j)}}(x, y)$$

$$= \frac{n!}{(i-1)!\,1!\,(j-i-1)!\,1!\,(n-j)!} F(x)^{i-1} f(x) \left(F(y)-F(x)\right)^{j-i-1} f(y) \left(1-F(y)\right)^{n-j}$$

$$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} f(x) f(y) F(x)^{i-1} \left(F(y)-F(x)\right)^{j-i-1} \left(1-F(y)\right)^{n-j}.$$

・$f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n)$ は, n 項分布とみなして,

$$f_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = n! \prod_{i=1}^{n} f(x_i) \quad (x_1 < \dots < x_n)$$