

確率分布のまとめ

◆ 正規分布

定義 1 x が平均 μ 、分散 σ^2 の正規分布 $N(\mu, \sigma^2)$ に従うとき、その確率密度は

$$\text{関数 } f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \text{で表される. } \Rightarrow \text{p.22 解説}$$

定理 1 正規分布 $N(\mu, \sigma^2)$ に従う x を標準化した $y = \frac{x-\mu}{\sigma}$ は、標準正規分布 $N(0, 1)$ に従う.

$\Rightarrow$ p.28 例題 2.5(5)

定理 2 平均 μ 、分散 σ^2 の任意の分布の母集団からランダムに $x_1, x_2, \dots, x_n$ を選んだとき、その平均 $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$ の分布は、 n が大きくなると正規分布 $N(\mu, \sigma^2/n)$ に近づく.

$\Rightarrow$ p.21 例題 2.3 (中心極限定理). $E(\bar{x}) = \mu$ と $V(\bar{x}) = \sigma^2/n$ の導出については $\Rightarrow$ p.18 解説

定理 3 偶然誤差 ε は、平均 0 の正規分布 $N(0, \sigma^2)$ に従う.

証明: $\Rightarrow$ p.126 参考

◆ 二項分布

定義 2 事象 A の確率が p (したがって余事象 $\bar{A}$ の確率が $1-p$) である試行を独立に n 回反復したとき、 A が起こる回数 r は二項分布 $B(n, p)$ に従い、その確率は ${}_nC_rp^r(1-p)^{n-r}$ で表される. $\Rightarrow$ p.36 解説

定理 4 二項分布 $B(n, p)$ の平均は np 、分散は $np(1-p)$ である. $\Rightarrow$ p.36 解説

定理 5 二項分布 $B(n, p)$ は n が十分大きいとき、正規分布 $N(np, np(1-p))$ に近づく. $\Rightarrow$ p.34 例題 2.8

◆ χ^2 分布

定義 3 標準正規分布 $N(0, 1)$ に従う母集団からランダムに $x_1, x_2, \dots, x_n$ を選んだとき、

$y = x_1^2 + x_2^2 + \dots + x_n^2$ が従う分布を自由度 n の χ^2 分布といい、その確率密度は $y > 0$ として

$$\text{関数 } f(y) = \frac{1}{2\Gamma(\frac{n}{2})} \left(\frac{y}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y}{2}} \text{で表される. } \Rightarrow \text{p.69 解説}$$

注意: ガンマ関数は正の実数 s に対して $\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt$ として定義され、特に $s = n$ (自然数) のとき $\Gamma(n) = (n-1)!$ が成り立つ (ガンマ関数は階乗の定義を正の実数全体に拡張したもの).

$\Rightarrow$ p.32 解説

定理 6 正規分布 $N(\mu, \sigma^2)$ に従う母集団からランダムに $x_1, x_2, \dots, x_n$ を選んだとき、

$$y = \frac{(x_1-\mu)^2}{\sigma^2} + \frac{(x_2-\mu)^2}{\sigma^2} + \dots + \frac{(x_n-\mu)^2}{\sigma^2} \text{は自由度 } n \text{ の } \chi^2 \text{ 分布に従う.}$$

証明：定理 1 より， $t_1 = \frac{x_1 - \mu}{\sigma}$ ， $t_2 = \frac{x_2 - \mu}{\sigma}$ ， $\dots$ ， $t_n = \frac{x_n - \mu}{\sigma}$ は標準正規分布 $N(0, 1)$ に従うから，定義 3 より $y = t_1^2 + t_2^2 + \dots + t_n^2$ は自由度 n の χ^2 分布に従う（終）。

定理 7 正規分布 $N(\mu, \sigma^2)$ に従う母集団からランダムに $x_1, x_2, \dots, x_n$ を選んだとき，その平均を

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \text{ とすると， } y = \frac{(\bar{x} - \mu)^2}{\sigma^2/n} \text{ は自由度 } 1 \text{ の } \chi^2 \text{ 分布に従う。}$$

証明：定理 2 より $\bar{x}$ は正規分布 $N(\mu, \sigma^2/n)$ に従うから， y は定理 6 で $n = 1$ の場合に相当する（終）。

定理 8 正規分布 $N(\mu, \sigma^2)$ に従う母集団からランダムに $x_1, x_2, \dots, x_n$ を選んだとき，その平均を

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}, \text{ 分散を } s^2 \text{（不偏分散を } u^2 \text{）とすると，}$$

$$y = \frac{(x_1 - \bar{x})^2}{\sigma^2} + \frac{(x_2 - \bar{x})^2}{\sigma^2} + \dots + \frac{(x_n - \bar{x})^2}{\sigma^2} = \frac{ns^2}{\sigma^2} = \frac{(n-1)u^2}{\sigma^2} \text{ は自由度 } n-1 \text{ の } \chi^2 \text{ 分布に従う。}$$

☞ p.68 例題 4.1 の解で使用した定理

証明の骨子：詳細は ☞ 小針暁宏「確率・統計入門」（岩波書店）命題 7.6

$x_1, x_2, \dots, x_n$ が独立なとき， $x_1 - \mu, x_2 - \mu, \dots, x_n - \mu$ は独立であるが， $x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x}$ は $x_1 + x_2 + \dots + x_n = n\bar{x}$ という関係式のため独立ではない。

そこで，まず母集団が標準正規分布 $N(0, 1)$ に従う場合， $x_1, x_2, \dots, x_n$ に対して直交行列により

$$\begin{cases} y_1 = \frac{1}{\sqrt{2}}(x_1 - x_2) \\ y_2 = \frac{1}{\sqrt{6}}(x_1 + x_2 - 2x_3) \\ y_3 = \frac{1}{\sqrt{12}}(x_1 + x_2 + x_3 - 3x_4) \\ \vdots \\ y_n = \frac{1}{\sqrt{n}}\{x_1 + x_2 + \dots + x_n\} \end{cases} \text{ として } y_1, y_2, \dots, y_n \text{ を定めると， } y_1, y_2, \dots, y_n \text{ も標準正規分布 } N(0, 1) \text{ に従い，}$$

かつ $\begin{cases} y_1^2 + y_2^2 + \dots + y_n^2 = x_1^2 + x_2^2 + \dots + x_n^2 \\ y_n^2 = n\bar{x}^2 \end{cases}$ とすることができる（☞ $x_1, x_2, \dots, x_n$ の代わりに

$y_1, y_2, \dots, y_{n-1}, \bar{x} (= \frac{y_n}{\sqrt{n}})$ を考える）。このとき，

$y = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2 = \dots = y_1^2 + y_2^2 + \dots + y_{n-1}^2$ が成り立つので，定義 3 より y は自由度 $n-1$ の χ^2 分布に従う。…①

母集団が一般の正規分布 $N(\mu, \sigma^2)$ に従う場合には，定理 1 より $t_1 = \frac{x_1 - \mu}{\sigma}$ ， $t_2 = \frac{x_2 - \mu}{\sigma}$ ， $\dots$ ， $t_n = \frac{x_n - \mu}{\sigma}$ が標準正規分布 $N(0, 1)$ に従い，このとき $\bar{t} = \frac{t_1 + t_2 + \dots + t_n}{n} = \frac{(x_1 + x_2 + \dots + x_n) - n\mu}{n\sigma} = \frac{\bar{x} - \mu}{\sigma}$ だから，

$y = \frac{(x_1 - \bar{x})^2}{\sigma^2} + \frac{(x_2 - \bar{x})^2}{\sigma^2} + \dots + \frac{(x_n - \bar{x})^2}{\sigma^2} = \dots = (t_1 - \bar{t})^2 + (t_2 - \bar{t})^2 + \dots + (t_n - \bar{t})^2$ は，①より自由度 $n-1$ の χ^2 分布に従う。（終）

定理 9 n 個のデータについて，カテゴリ C_1 の観測度数が O_1 で期待度数が E_1 ，カテゴリ C_2 の観測度数が O_2 で期待度数が

E_2 とすると， $y = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2}$ は n が十分大きいと

カテゴリ	C_1	C_2	計
観測度数	O_1	O_2	n
期待度数	E_1	E_2	n
確率	p_1	p_2	1

き、自由度1の χ^2 分布に従う。☞ p.71 例題 4.2 の解で使用した定理はこの拡張

証明： $p_1 = \frac{E_1}{n}$, $p_2 = \frac{E_2}{n} = 1 - p_1$ とおくと、 O_1 は n 回の反復試行中事象 C_1 が起こる回数 ($O_2 = n - O_1$ は C_1 が起こらない回数) と見なすことができ、定義 2 より二項分布 $B(n, p_1)$ に従うので、 n が十分大きいときは定理 5 より正規分布 $N(np_1, np_1(1 - p_1))$ に近づく。

$$\text{いま } y = \frac{(O_1 - np_1)^2}{np_1} + \frac{(O_2 - np_2)^2}{np_2} = \frac{(O_1 - np_1)^2}{np_1} + \frac{(n - O_1 - n + np_1)^2}{n(1 - p_1)} = \frac{(O_1 - np_1)^2}{np_1(1 - p_1)} = \left\{ \frac{O_1 - np_1}{\sqrt{np_1(1 - p_1)}} \right\}^2 \text{ において,}$$

$\frac{O_1 - np_1}{\sqrt{np_1(1 - p_1)}}$ は定理 1 より標準正規分布 $N(0, 1)$ に従うので、定義 3 より y は自由度 1 の χ^2 分布に従う。

◆ F 分布

定義 4 x_A, x_B は独立で、各々自由度 n_A, n_B の χ^2 分布に従うとき、 $y = \frac{x_A/n_A}{x_B/n_B}$ が従う分布を自由度 (n_A, n_B) の

F 分布 といい、その確率密度は関数 $f_{n_A, n_B}(y) = \frac{n_A^{\frac{n_A}{2}} n_B^{\frac{n_B}{2}}}{B(\frac{n_A}{2}, \frac{n_B}{2})} \cdot \frac{x^{\frac{n_A}{2}-1}}{(n_A y + n_B)^{\frac{n_A+n_B}{2}}}$ で表される。☞ p.98 参考

注意：ベータ関数は正の実数 p, q に対して $B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} dt$ として定義され、ガンマ関数

とは $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ の関係がある。☞ p.84 解説

定理 10 正規母集団 $N(\mu_A, \sigma_A^2)$ から抽出した大きさ n_A の標本の不偏分散が u_A^2 、正規母集団 $N(\mu_B, \sigma_B^2)$ から

抽出した大きさ n_B の標本の不偏分散が u_B^2 のとき、 $y = \frac{u_A^2/\sigma_A^2}{u_B^2/\sigma_B^2}$ は自由度 $(n_A - 1, n_B - 1)$ の F 分布

に従う。☞ p.82 例題 5.1 の解で使用した定理

証明：定理 8 より、 $\frac{(n_A-1)u_A^2}{\sigma_A^2}$ は自由度 $n_A - 1$ の χ^2 分布に従い、 $\frac{(n_B-1)u_B^2}{\sigma_B^2}$ は自由度 $n_B - 1$ の χ^2 分布に従う

から、定義 4 より $y = \frac{\frac{(n_A-1)u_A^2}{\sigma_A^2}/(n_A-1)}{\frac{(n_B-1)u_B^2}{\sigma_B^2}/(n_B-1)} = \frac{u_A^2/\sigma_A^2}{u_B^2/\sigma_B^2}$ は自由度 $(n_A - 1, n_B - 1)$ の F 分布に従う (終)。

☞ p.98 参考

定理 11 正規母集団 $N(\mu, \sigma^2)$ から抽出した大きさ n の標本の平均が $\bar{x}$ 、不偏分散が u^2 のとき、 $y = \frac{(\bar{x}-\mu)^2}{u^2/n}$ は

自由度 $(1, n - 1)$ の F 分布に従う。

証明：定理 7 より $\frac{(\bar{x}-\mu)^2}{\sigma^2/n}$ は自由度 1 の χ^2 分布に従い、定理 8 より $\frac{ns^2}{\sigma^2} = \frac{(n-1)u^2}{\sigma^2}$ は自由度 $n - 1$ の χ^2 分布に

従うから、定義 4 より $y = \frac{\frac{(\bar{x}-\mu)^2}{\sigma^2/n}/1}{\frac{(n-1)u^2}{\sigma^2}/(n-1)} = \frac{(\bar{x}-\mu)^2}{u^2/n}$ は自由度 $(1, n - 1)$ の F 分布に従う (終)。

◆ t 分布

定義 5 x の確率密度が関数 $f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$ で表される分布を、**自由度 n の t 分布** という。

☞ p.32 解説

定理 12 n が十分大きいときの t 分布は、**標準正規分布 $N(1, 0)$** に近づく。☞ p.32 解説

定理 13 x が自由度 n の t 分布に従うとき、 x^2 は**自由度 $(1, n)$ の F 分布** に従う。

証明：☞ p.99 参考

定理 14 正規母集団 $N(\mu, \sigma^2)$ から抽出した大きさ n の標本の平均が $\bar{x}$ 、分散が s^2 のとき、

$y = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$ は**自由度 $n - 1$ の t 分布** に従う。☞ p.30 例題 2.7 の解で使用した定理

証明：標本の不偏分散を u^2 とすると、定理 11 より $\frac{(\bar{x} - \mu)^2}{u^2/n} = \frac{(\bar{x} - \mu)^2}{s^2/(n-1)}$ は**自由度 $(1, n - 1)$ の F 分布** に従うから、

定理 13 より $y = \frac{\bar{x} - \mu}{s/\sqrt{n-1}}$ は**自由度 $n - 1$ の t 分布** に従う（終）。

◆ ポアソン分布

定義 6 二項分布 $B(n, p)$ において、 $np = \lambda$ を一定のまま $n \rightarrow \infty$ としたときの分布を**ポアソン分布**といい、

その確率は $\frac{e^{-\lambda}\lambda^r}{r!}$ で表される。☞ p.146 解説

定理 15 ポアソン分布の平均と分散は、ともに λ である。☞ p.146 解説