クレジット:
UTokyo Online Education 統計データ解析Ⅰ 2017 小池祐太

統計データ解析 (I) 6 章 参考資料

1　正規分布の平均・分散の計算

確率密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

で与えられる連続分布を平均 μ, 分散 σ^2 の正規分布または Gauss 分布と呼ぶ.

X を平均 μ, 分散 σ^2 の正規分布に従う確率変数とする. 積分の線形性より,

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} (x-\mu)f(x)dx + \mu \int_{-\infty}^{\infty} f(x)dx$$

が成り立つ. ここで,

$$\begin{aligned}(\text{右辺第 1 項}) &= \int_{-\infty}^{\infty} yf(y+\mu)dy \ (y = x-\mu \text{ と置換}) \\ &= 0 \ (\text{奇関数の性質})\end{aligned}$$

であり, また $\int_{-\infty}^{\infty} f(x)dx = 1$ より (右辺第 2 項)$= \mu$ である. 以上より,

$$E[X] = \mu.$$

一方で,

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy \ (y = (x-\mu)/\sigma \text{と置換})$$

であり, 部分積分公式より

$$\int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy = \int_{-\infty}^{\infty} y \cdot ye^{-y^2/2} dy = \left[-ye^{-y^2/2}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}$$

であるから,

$$\mathrm{Var}[X] = \sigma^2.$$

2　ガンマ分布の平均・分散の計算

ν, α を正の実数とする. 確率密度関数が

$$f(x) = \frac{1}{\Gamma(\nu)} \alpha^\nu x^{\nu-1} e^{-\alpha x} \quad (x > 0), \qquad f(x) = 0 \quad (x \leq 0)$$

で与えられる連続分布をパラメータ ν, α のガンマ分布と呼ぶ. ただし, $\Gamma(\nu)$ はガンマ関数

$$\Gamma(\nu) = \int_0^\infty x^{\nu-1} e^{-x} dx$$

を表す.

X をパラメータ ν, α のガンマ分布に従う確率変数とする.

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\Gamma(\nu)} \int_0^\infty \alpha^\nu x^\nu e^{-\alpha x} dx = \frac{1}{\alpha \Gamma(\nu)} \int_0^\infty y^\nu e^{-y} dy \ (y = \alpha x \text{ と置換}) \\ &= \frac{\Gamma(\nu+1)}{\alpha \Gamma(\nu)} \end{aligned}$$

であるから, 公式 $\Gamma(\nu+1) = \nu \Gamma(\nu)$ (「解析入門 I」, IV 章定理 12.2) より,

$$E[X] = \nu / \alpha.$$

一方で,

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx = \frac{1}{\Gamma(\nu)} \int_0^\infty \alpha^\nu x^{\nu+1} e^{-\alpha x} dx = \frac{1}{\alpha^2 \Gamma(\nu)} \int_0^\infty y^{\nu+1} e^{-y} dy \ (y = \alpha x \text{ と置換}) \\ &= \frac{\Gamma(\nu+2)}{\alpha^2 \Gamma(\nu)} \end{aligned}$$

であるから, 公式 $\Gamma(\nu+2) = (\nu+1)\nu\Gamma(\nu)$ (「解析入門 I」, IV 章定理 12.2) より

$$E[X^2] = (\nu+1)\nu / \alpha^2.$$

従って,

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2 = \nu / \alpha^2.$$

3 ベータ分布の平均・分散の計算

α, β を正の実数とする. 確率密度関数が

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad (0 < x < 1), \qquad f(x) = 0 \quad (x \notin (0,1))$$

で与えられる連続分布を, パラメーター α, β のベータ分布と呼ぶ. ただし, $B(\alpha, \beta)$ はベータ関数

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$$

を表す.

X をパラメーター α, β のベータ分布に従う確率変数とする.

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{B(\alpha, \beta)} \int_0^1 x^\alpha (1-x)^{\beta-1} dx = \frac{B(\alpha+1, \beta)}{B(\alpha, \beta)}$$

であるから, 公式 $B(\alpha+1, \beta) = \frac{\alpha}{\alpha+\beta} B(\alpha, \beta)$ (「解析入門 I」, IV 章定理 12.2–12.3) より,

$$E[X] = \frac{\alpha}{\alpha + \beta}.$$

一方で,

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{1}{B(\alpha,\beta)}\int_0^1 x^{\alpha+1}(1-x)^{\beta-1}dx = \frac{B(\alpha+2,\beta)}{B(\alpha,\beta)}$$

であるから, 公式 $B(\alpha+2,\beta) = \frac{\alpha+1}{\alpha+\beta+1}\frac{\alpha}{\alpha+\beta}B(\alpha,\beta)$ (「解析入門 I」, IV 章定理 12.2–12.3) より,

$$E[X^2] = \frac{\alpha+1}{\alpha+\beta+1}\frac{\alpha}{\alpha+\beta}.$$

従って,

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

4 t 分布の平均・分散の計算

ν を正の実数とする. 確率密度関数が

$$f(x) = \frac{1}{\sqrt{\pi\nu}}\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}$$

で与えられる連続分布を, 自由度 ν の (Student の) t 分布と呼ぶ.

$xf(x) = O(|x|^{-\nu})$ $(|x|\infty)$ であるから, $\nu > 1$ のとき積分 $\int_{-\infty}^{\infty} xf(x)dx$ は絶対収束する (「解析入門 I」, IV 章定理 11.3). 従って X は平均をもち, 奇関数の性質から,

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = 0$$

となる. 一方で, $x^2 f(x) = O(|x|^{-(\nu-1)})$ $(|x|\infty)$ であるから, $\nu > 2$ のとき積分 $\int_{-\infty}^{\infty} x^2 f(x)dx$ は絶対収束する (「解析入門 I」, IV 章定理 11.3). 従って X は分散をもち,

$$\begin{aligned}\mathrm{Var}[X] &= \int_{-\infty}^{\infty} x^2 f(x)dx = 2\int_0^{\infty} x^2 f(x)dx \ (\text{偶関数の性質}) \\ &= \frac{\nu}{\sqrt{\pi}}\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)}\int_0^{\infty}\frac{y^{\frac{1}{2}}}{(1+y)^{(\nu+1)/2}}dy \ (y = x^2/\nu\text{と置換})\end{aligned}$$

であるから, 公式

$$\int_0^{\infty}\frac{y^{\frac{1}{2}}}{(1+y)^{(\nu+1)/2}}dy = B\left(\frac{\nu}{2}-1,\frac{3}{2}\right)$$

(「解析入門 I」, IV 章 (12.6) 式) および

$$B\left(\frac{\nu}{2}-1,\frac{3}{2}\right) = \frac{\Gamma(\frac{\nu}{2}-1)\Gamma(\frac{3}{2})}{\Gamma(\frac{\nu+1}{2})}$$

(「解析入門 I」, IV 章定理 12.3) より

$$\mathrm{Var}[X] = \frac{\nu}{\sqrt{\pi}}\frac{\Gamma(\frac{\nu}{2}-1)\Gamma(\frac{3}{2})}{\Gamma(\nu/2)}$$

を得る. さらに, 公式

$$\Gamma(\nu/2) = (\nu/2-1)\Gamma(\nu/2-1), \qquad \Gamma(3/2) = \Gamma(1/2)/2$$

(「解析入門 I」, IV 章定理 12.2) および

$$\Gamma(1/2) = \sqrt{\pi}$$

(「解析入門 I」, IV 章定理 12.4) より,

$$\mathrm{Var}[X] = \frac{\nu}{\nu-2}.$$

5　F 分布の平均・分散の計算

ν_1, ν_2 を正の実数とする. 確率密度関数が

$$f(x) = \frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2, \nu_2/2)} \frac{x^{\nu_1/2-1}}{(1+\nu_1 x/\nu_2)^{(\nu_1+\nu_2)/2}} \quad (x > 0), \qquad f(x) = 0 \quad (x \leq 0)$$

で与えられる連続分布を, 自由度 ν_1, ν_2 の F 分布と呼ぶ.

X を自由度 ν_1, ν_2 の F 分布に従う確率変数とする. $xf(x) = O(x^{-\nu_2/2})$ $(x \to \infty)$ であるから, $\nu_2 > 2$ のとき積分 $\int_{-\infty}^{\infty} xf(x)dx$ は絶対収束する (「解析入門 I」, IV 章定理 11.3). 従って X は平均をもち,

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} xf(x)dx = \frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2, \nu_2/2)} \int_0^{\infty} \frac{x^{\nu_1/2}}{(1+\nu_1 x/\nu_2)^{(\nu_1+\nu_2)/2}} dx \\ &= \frac{\nu_2/\nu_1}{B(\nu_1/2, \nu_2/2)} \int_0^{\infty} \frac{y^{\nu_1/2}}{(1+y)^{(\nu_1+\nu_2)/2}} dy \ (y = (\nu_1/\nu_2)x \text{ と置換}) \end{aligned}$$

が成り立つ. 公式

$$\int_0^{\infty} \frac{y^{\nu_1/2}}{(1+y)^{(\nu_1+\nu_2)/2}} dy = B(\nu_2/2 - 1, \nu_1/2 + 1)$$

(「解析入門 I」, IV 章 (12.6) 式) および

$$B(\nu_2/2 - 1, \nu_1/2 + 1) = \frac{\Gamma(\nu_2/2 - 1)\Gamma(\nu_1/2 + 1)}{\Gamma((\nu_1+\nu_2)/2)}, \qquad B(\nu_1/2, \nu_2/2) = \frac{\Gamma(\nu_2/2)\Gamma(\nu_1/2)}{\Gamma((\nu_1+\nu_2)/2)}$$

(「解析入門 I」, IV 章定理 12.3) および

$$\Gamma(\nu_2/2) = (\nu_2/2 - 1)\Gamma(\nu_2/2 - 1), \qquad \Gamma(\nu_1/2 + 1) = (\nu_1/2)\Gamma(\nu_1/2)$$

(「解析入門 I」, IV 章定理 12.2) より

$$E[X] = \frac{\nu_2}{\nu_1} \frac{1}{\nu_2/2 - 1} \frac{\nu_1}{2} = \frac{\nu_2}{\nu_2 - 2}.$$

一方で, $x^2 f(x) = O(x^{-(\nu_2/2-1)})$ $(x \to \infty)$ であるから, $\nu_2 > 4$ のとき積分 $\int_{-\infty}^{\infty} xf(x)dx$ は絶対収束する (「解析入門 I」, IV 章定理 11.3). 従って X は 2 次のモーメントをもち,

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2, \nu_2/2)} \int_0^{\infty} \frac{x^{\nu_1/2+1}}{(1+\nu_1 x/\nu_2)^{(\nu_1+\nu_2)/2}} dx \\ &= \frac{(\nu_2/\nu_1)^2}{B(\nu_1/2, \nu_2/2)} \int_0^{\infty} \frac{y^{\nu_1/2+1}}{(1+y)^{(\nu_1+\nu_2)/2}} dy \ (y = (\nu_1/\nu_2)x \text{ と置換}) \end{aligned}$$

が成り立つ. 公式

$$\int_0^{\infty} \frac{y^{\nu_1/2+1}}{(1+y)^{(\nu_1+\nu_2)/2}} dy = B(\nu_2/2 - 2, \nu_1/2 + 2)$$

(「解析入門 I」, IV 章 (12.6) 式) および

$$B(\nu_2/2 - 2, \nu_1/2 + 2) = \frac{\Gamma(\nu_2/2 - 2)\Gamma(\nu_1/2 + 2)}{\Gamma((\nu_1+\nu_2)/2)}, \qquad B(\nu_1/2, \nu_2/2) = \frac{\Gamma(\nu_2/2)\Gamma(\nu_1/2)}{\Gamma((\nu_1+\nu_2)/2)}$$

(「解析入門 I」, IV 章定理 12.3) および

$$\Gamma(\nu_2/2) = (\nu_2/2 - 1)(\nu_2/2 - 2)\Gamma(\nu_2/2 - 2), \qquad \Gamma(\nu_1/2 + 2) = (\nu_1/2 + 1)(\nu_1/2)\Gamma(\nu_1/2)$$

(「解析入門 I」, IV 章定理 12.2) より

$$E[X^2] = \left(\frac{\nu_2}{\nu_1}\right)^2 \frac{(\nu_1/2+1)(\nu_1/2)}{(\nu_2/2-1)(\nu_2/2-2)} = \frac{\nu_2^2(\nu_1+2)}{\nu_1(\nu_2-2)(\nu_2-4)}.$$

従って,

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2 = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}.$$

6 参考文献

1. 杉浦光夫著「解析入門 I」, 東京大学出版会 (1980 年).