
積率母関数

$$E[e^{tX}] = \int_{-\infty}^{\infty} e^{tX} f(X) \mathrm{d}X \tag{1}$$

X が正規分布 $N(\mu, \sigma^2)$ に従う場合

$$E[e^{tX}] = e^{\mu t + \frac{\sigma^2 t^2}{2}} = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \tag{2}$$

··

記号

1. $\bar{X}$　確率変数 X の算術平均
2. $E[X]$　確率変数 X の期待値で $\bar{X}$ と同じ
3. $V[X]$　確率変数 X の分散
4. $\mathrm{Cov}(X, Y)$　確率変数 X, Y の共分散
5. S_{xy}　偏差の積 $(x - \bar{x})(y - \bar{y})$ の総和
6. S_{xx}　偏差の平方 $(x - \bar{x})^2$ の総和
7. R　相関係数
8. R^2　決定係数 $R^2 = 1 - S_e/S_T = S_{xy}^2/(S_{xx}S_{yy})$
9. S_T　全平方和 $S_T = S_{yy}$
10. S_R　回帰平方和 $S_R = S_{xy}^2/S_{xx}$
11. S_e　残差平方和 $S_e = S_T - S_R$
12. V_R　回帰分散、回帰平均平方 $V_R = S_R/\phi_R$
13. V_e　残差分散、残差平均平方 $V_e = S_e/\phi_e$
14. ϕ_T　自由度
15. ϕ_R　回帰自由度
16. ϕ_e　残差自由度 $\phi_e = \phi_T - \phi_R$
17. $y = \alpha + \beta x + \varepsilon$　一次線形回帰モデル
18. α, β　回帰係数
19. $\hat{\beta}$　β の推定値 $\hat{\beta} = S_{xy}/S_{xx}$
20. $\hat{\alpha}$　α の推定値 $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

··

性質

- $E[k] = k \quad (k : const)$
- $E[X+Y] = E[X] + E[Y]$
- $E[kX] = kE[X]$
- $V[k] = 0 \quad (k : const)$
- $V[X+Y] = V[X] + V[Y] + 2\mathrm{Cov}(X,Y)$
- $V[kX] = k^2 V[X]$
- $V[X] = E[X^2] - (E[X])^2$
- $\mathrm{Cov}(X,X) = V[X]$
- $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
- $\mathrm{Cov}(X,k) = 0 \quad (k : const)$
- $\mathrm{Cov}(X,k) = E[XY] - E[X]E[Y]$
- $\mathrm{Cov}(X+k,Y) = \mathrm{Cov}(X,Y) \quad (k : const)$
- $\mathrm{Cov}(kX,Y) = k\mathrm{Cov}(X,Y) \quad (k : const)$
- $\mathrm{Cov}(X+Z,Y) = \mathrm{Cov}(X,Y) + \mathrm{Cov}(Z,Y)$

....

分散共分散行列

確率変数 X_i に対し次の行列 Σ を分散共分散行列という。

$$\Sigma = \begin{pmatrix} \mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Cov}(X_n,X_n) \end{pmatrix} \tag{3}$$

数値は小数第 3 位で解答する。

---

1. $m \in \mathbb{R}$, $v > 0$ を定数とし、実数値確率変数 X は正規分布 $N(m,v)$ に従うとする。
   $Y = pX + q$ $(p, q \in \mathbb{R},\ p \neq 0)$ とおく時、Y の積率母関数 $E[e^{tY}]$ $(t \in \mathbb{R})$ を求めよ。

   ....

   確率変数 X が正規分布 $N(m,v)$ に従う為、$Y = pX + q$ は $N(pm+q, p^2 v)$

に従う。これにより、Y の積率母関数は次のようになる。

$$E[\exp(tY)] = \exp\left((pm+q)t + \frac{p^2vt^2}{2}\right) \tag{4}$$

---

2. n 次対角行列 V とし、$\mathbb{R}^n$ 値確率変数 $\boldsymbol{X} : \Omega \to \mathbb{R}^n$ は n 次元正規分布 $N_n(\mathbf{0}, V)$ に従うものとする。

$$V = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & n \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \tag{5}$$

(a) $1 \leq i \leq n$ に対して、X_i の平均 $E[X_i]$ を求めよ。

(b) $1 \leq i \leq n, 1 \leq j \leq n$ に対して、共分散 $\mathrm{Cov}(X_i, X_j)$ を求めよ。

....

(a) $N_n(\mathbf{0}, V)$ より $E[\boldsymbol{X}] = \mathbf{0}$。つまり、$E[X_i] = 0$ $(i = 1, \ldots, n)$ である。

(b) V は分散共分散行列である。この為、各々の確率変数 X_i の共分散には次のようになる。

$$\mathrm{Cov}(X_i, X_j) = \begin{cases} i & i = j \\ 0 & i \neq j \end{cases} \tag{6}$$

---

3. 次のデータについて、以下の問いに答えよ。ただし、$t(8, 0.05) = 2.306$ とする。

説明変数	2.2	4.1	5.5	1.9	3.4	2.6	4.2	3.7	4.9	3.2
目的変数	71	81	86	72	77	73	80	81	85	74

(a) $\bar{x}$ 及び $\bar{y}$ を求めよ。

(b) S_{xy}、S_{xx}、S_{yy} を求めよ。

(c) 最小二乗法によって $\hat{\beta}_0$ 及び $\hat{\beta}_1$ を求めよ。

(d) S_R、R^2、S_e、V_e を求めよ。

(e) 回帰係数 β_1 が 0 であるかどうかを、有意水準 0.05 として検定せよ。

(f) 回帰係数 β_1 の信頼率 95% の信頼区間を求めよ。

....

(a) 説明変数

$$\bar{x} = (2.2 + 4.1 + 5.5 + 1.9 + 3.4 + 2.6 + 4.2 + 3.7 + 4.9 + 3.2) \div 10 \tag{7}$$

$$= 3.57 \tag{8}$$

目的変数

$$\bar{y} = (71 + 81 + 86 + 72 + 77 + 73 + 80 + 81 + 85 + 74) \div 10 \tag{9}$$

$$= 78 \tag{10}$$

(b) まず、説明変数と目的変数の偏差を求める。

$x - \bar{x}$	-1.37	0.53	1.93	-1.67	-0.17	-0.97	0.63	0.13	1.33	-0.37
$y - \bar{y}$	-7	3	8	-6	-1	-5	2	3	7	-4
偏差の積	9.59	1.59	15.44	10.02	0.17	4.85	1.26	0.39	9.31	1.48

$$S_{xy} = 9.59 + 1.59 + 15.44 + 10.02 + 0.17 \tag{11}$$

$$+ 4.85 + 1.26 + 0.39 + 9.31 + 1.48 \tag{12}$$

$$= 54.1 \tag{13}$$

$$S_{xx} = (-1.37)^2 + (0.53)^2 + (1.93)^2 + (-1.67)^2 + (-0.17)^2 \tag{14}$$

$$+ (-0.97)^2 + (0.63)^2 + (0.13)^2 + (1.33)^2 + (-0.37)^2 \tag{15}$$

$$= 11.961 \tag{16}$$

$$S_{yy} = (-7)^2 + 3^2 + 8^2 + (-6)^2 + (-1)^2 \tag{17}$$

$$+ (-5)^2 + 2^2 + 3^2 + 7^2 + (-4)^2) \tag{18}$$

$$= 262 \tag{19}$$

(c) 説明変数 x と目的変数 y の間には $y = \beta_0 + \beta_1 x$ という関係があると考え、ここに誤差 ε を含んだ $y = \beta_0 + \beta_1 x + \varepsilon$ という回帰モデルを考える。最小 2 乗法における β_1 の推定値 $\hat{\beta}_1$ は $\hat{\beta}_1 = S_{xy}/S_{xx}$ により求まる。

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{54.1}{11.961} = 4.523033 \fallingdotseq 4.523 \tag{20}$$

推定値 $\hat{\beta}_0$ は $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$ を満たす為、

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \bar{y} - \frac{S_{xy}}{S_{xx}} \bar{x} \tag{21}$$

$$= 78 - \frac{54.1}{11.961} \times 3.57 = 61.85277 \fallingdotseq 61.853 \tag{22}$$

（d）回帰平方和 S_R

$$S_R = \frac{S_{xy}^2}{S_{xx}} = \frac{54.1^2}{11.961} = 244.69609\cdots \fallingdotseq 244.696 \tag{23}$$

決定係数 R^2

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{54.1^2}{11.961 \times 262} = 0.9339545\cdots \fallingdotseq 0.934 \tag{24}$$

残差平方和 S_e

$$S_e = S_{yy} - \frac{S_{xy}^2}{S_{xx}} = 262 - \frac{54.1^2}{11.961} = 17.303904\cdots \fallingdotseq 17.304 \tag{25}$$

残差分散 V_e

$$V_e = \frac{1}{10-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right) = \frac{1}{8}\left(262 - \frac{54.1^2}{11.961}\right) = 2.162988\cdots \fallingdotseq 2.163 \tag{26}$$

（e）帰無仮説を $\beta_1 = 0$ とする。

$t(8, 0.05) = 2.306$ より有意水準 0.05 において次の式を満たせば仮説を棄却する。

$$\frac{|\hat{\beta}_1 - 0|}{\sqrt{S_e}8S_{xx}} \geq 2.306 \tag{27}$$

左辺を計算すると次の値となる。

$$\frac{|\hat{\beta}_1 - 0|}{\sqrt{\frac{S_e}{8S_{xx}}}} = \frac{S_{xy}}{S_{xx}}\sqrt{\frac{8S_{xx}^2}{S_{xx}S_{yy} - S_{xy}^2}} = \sqrt{\frac{8S_{xy}^2}{S_{xx}S_{yy} - S_{xy}^2}} \tag{28}$$

$$= \sqrt{\frac{8 \times 54.1^2}{11.961 \times 262 - 54.1^2}} = 10.6362 \fallingdotseq 10.636 \tag{29}$$

これにより、帰無仮説は有意水準 0.05 にて棄却された。

（f）$\hat{\beta}_1$ の期待値と分散は次の通りである。

$$E[\hat{\beta}_1] = \beta_1, \quad V[\hat{\beta}_1] = \frac{\sigma^2}{S_{xx}} \tag{30}$$

この為、$\hat{\beta}_1$ を正規化した $\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{S_{xx}}}}$ は正規分布 $N(0,1)$ に従う。この時、σ^2 は母数であり未知の値である。この為、σ^2 の代わりに残差分散 V_e を用いる。

$$t = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{V_e}{S_{xx}}}} = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{S_e}{8S_{xx}}}} \tag{31}$$

この t は自由度 8 の t 分布に従う。

$t(8, 0.05) = 2.306$ より信頼率 $95\%(= 1 - 0.05)$ で次の区間が求まる。

$$-2.306 < \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{S_e}{8S_{xx}}}} < 2.306 \tag{32}$$

式を変形すると β_1 の区間が求まる。

$$\hat{\beta}_1 - 2.306\sqrt{\frac{S_e}{8S_{xx}}} < \beta_1 < \hat{\beta}_1 + 2.306\sqrt{\frac{S_e}{8S_{xx}}} \tag{33}$$

両側の値を計算する。

$$\hat{\beta}_1 + 2.306\sqrt{\frac{S_e}{8S_{xx}}} = \frac{S_{xy}}{S_{xx}} + 2.306\sqrt{\frac{S_{yy}}{8S_{xx}} - \frac{S_{xy}^2}{8S_{xx}^2}} \tag{34}$$

$$= \frac{54.1}{11.961} + 2.306\sqrt{\frac{262}{8 \times 11.961} - \frac{54.1^2}{8 \times 11.961^2}} \tag{35}$$

$$= 5.503657 \fallingdotseq 5.504 \tag{36}$$

同様にして

$$\hat{\beta}_1 + 2.306\sqrt{\frac{S_e}{8S_{xx}}} = 3.542409 \fallingdotseq 3.542 \tag{37}$$

よって、信頼率 95% の信頼区間は次のようになる。

$$3.542 < \beta_1 < 5.504 \tag{38}$$

---

4. $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{y} \in \mathbb{R}^n$ を次のようなベクトルとする。

$$\boldsymbol{x}_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} x_{12} \\ \vdots \\ x_{n2} \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \tag{39}$$

ある定数 α_0, β_1, $\beta_2 \in \mathbb{R}$ が存在して

$$y_i = \alpha_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \varepsilon_i, \quad 1 \leq i \leq n \tag{40}$$

が成り立つとする。ただし、

$$\bar{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{i1}, \quad \bar{x}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{i2}, \tag{41}$$

とおき、$\sigma > 0$ は定数とし、$\{\varepsilon_1, \ldots, \varepsilon_n\}$ は独立同分布確率変数列で、各 ε_i は正規分布 $N(0, \sigma^2)$ に従うものとする。また、$S_{x_1x_1}S_{x_2x_2} > (S_{x_1x_2})^2$ が成り立っているものとする。

（a）α_0, β_1, β_2 を利用し、残差平方和 S_e を 求めよ。

（b）$\dfrac{\partial S_e}{\partial \alpha_0}$, $\dfrac{\partial S_e}{\partial \beta_1}$, $\dfrac{\partial S_e}{\partial \beta_2}$ を求めよ。

（c）S_e に対するヘッセ行列 H が正定値行列になることを示せ。

$$H = \begin{pmatrix} \dfrac{\partial^2 S_e}{\partial \alpha_0^2} & \dfrac{\partial^2 S_e}{\partial \alpha_0 \partial \beta_1} & \dfrac{\partial^2 S_e}{\partial \alpha_0 \partial \beta_2} \\ \dfrac{\partial^2 S_e}{\partial \alpha_0 \partial \beta_1} & \dfrac{\partial^2 S_e}{\partial \beta_1^2} & \dfrac{\partial^2 S_e}{\partial \beta_1 \partial \beta_2} \\ \dfrac{\partial^2 S_e}{\partial \alpha_0 \partial \beta_2} & \dfrac{\partial^2 S_e}{\partial \beta_1 \partial \beta_2} & \dfrac{\partial^2 S_e}{\partial \beta_2^2} \end{pmatrix} \tag{42}$$

……………………………………………………………………………………

（a）残差平方和 S_e

$$S_e = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n}(y_i - \alpha_0 - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2))^2 \tag{43}$$

（b）$\bar{x}_1$ と $\bar{x}_2$ は平均なので次のように変形できる。

$$\sum_{i=1}^{n} x_{i1} - n\bar{x}_1 = 0, \quad \sum_{i=1}^{n} x_{i2} - n\bar{x}_2 = 0 \tag{44}$$

また、$S_{x_1x_1}$, $S_{x_1x_2}$, $S_{x_2x_2}$ は次のような式である。

$$S_{x_1x_1} = \sum_{i=1}^{n}\left(x_{i1} - \bar{x}_1\right)^2, \qquad S_{x_2x_2} = \sum_{i=1}^{n}\left(x_{i2} - \bar{x}_2\right)^2 \tag{45}$$

$$S_{x_1x_2} = \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \tag{46}$$

$$\begin{aligned}
\frac{\partial S_e}{\partial \alpha_0} =& -2\sum_{i=1}^{n}(y_i - \alpha_0 - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)) && (47)\\
=& -2\sum_{i=1}^{n} y_i + 2n\alpha_0 + 2\beta_1\left(\sum_{i=1}^{n} x_{i1} - n\bar{x}_1\right) + 2\beta_2\left(\sum_{i=1}^{n} x_{i2} - n\bar{x}_2\right) && (48)\\
=& -2\sum_{i=1}^{n} y_i + 2n\alpha_0 && (49)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial S_e}{\partial \beta_1} =& -2\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(y_i - \alpha_0 - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)) && (50)\\
=& -2\sum_{i=1}^{n} y_i(x_{i1} - \bar{x}_1) + 2\alpha_0\left(\sum_{i=1}^{n} x_{i1} - n\bar{x}_1\right) && (51)\\
& + 2\beta_1\sum_{i=1}^{n}\left(x_{i1} - \bar{x}_1\right)^2 + 2\beta_2\sum_{i=1}^{n}\left(x_{i1} - \bar{x}_1\right)\left(x_{i2} - \bar{x}_2\right) && (52)\\
=& -2\sum_{i=1}^{n} y_i(x_{i1} - \bar{x}_1) + 2\beta_1 S_{x_1x_1} + 2\beta_2 S_{x_1x_2} && (53)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial S_e}{\partial \beta_2} =& -2\sum_{i=1}^{n}(x_{i2} - \bar{x}_2)(y_i - \alpha_0 - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)) && (54)\\
=& -2\sum_{i=1}^{n} y_i(x_{i2} - \bar{x}_2) + 2\beta_1 S_{x_1x_2} + 2\beta_2 S_{x_2x_2} && (55)
\end{aligned}$$

（c）ヘッセ行列 H は次のような行列である。

$$H = \begin{pmatrix} 2n & 0 & 0 \\ 0 & 2S_{x_1x_1} & 2S_{x_1x_2} \\ 0 & 2S_{x_1x_2} & 2S_{x_2x_2} \end{pmatrix} \tag{56}$$

ここで、$\boldsymbol{x}_1$ の全ての成分が等しい場合に限り $S_{x_1x_1} = 0$ であり、そうでない場合 $S_{x_1x_1} > 0$ である。これは $\boldsymbol{x}_2$ についても同様である。

$\boldsymbol{x}_1$ の成分は異なる場合を考える。(つまり、$S_{x_1x_1} > 0$ である)

行列が正定値であるためには首座小行列式が全て正であれば良い。ヘッセ行列 H の首座小行列は次の 3 つである。

$$\begin{pmatrix} 2n \end{pmatrix}, \quad \begin{pmatrix} 2n & 0 \\ 0 & 2S_{x_1x_1} \end{pmatrix}, \quad H \tag{57}$$

この 3 つの行列式を考える。

$$\det \begin{pmatrix} 2n \end{pmatrix} = 2n > 0, \quad \det \begin{pmatrix} 2n & 0 \\ 0 & 2S_{x_1x_1} \end{pmatrix} = 4nS_{x_1x_1} > 0 \tag{58}$$

H の行列式は次のようになる。

$$\det H = \begin{vmatrix} 2n & 0 & 0 \\ 0 & 2S_{x_1x_1} & 2S_{x_1x_2} \\ 0 & 2S_{x_1x_2} & 2S_{x_2x_2} \end{vmatrix} = 2n \begin{vmatrix} 2S_{x_1x_1} & 2S_{x_1x_2} \\ 2S_{x_1x_2} & 2S_{x_2x_2} \end{vmatrix} \tag{59}$$

$$= 8n(S_{x_1x_1}S_{x_2x_2} - S_{x_1x_2}^2) > 0 \tag{60}$$

これによりヘッセ行列 H は正定値であることが分かる。