第8回 数理統計 レポート

小森 一輝

2020 年 11 月 20 日

定義 3.9. 期待値

i) 確率変数 X が離散型の場合, X の確率関数 f_X が $\boxed{\sum_{i=1}^{\infty} |x_i| f_X(x_i) < \infty}$ $(D = \{x_i | i = 1, 2, \ldots\})$ を満たすとき, X の**期待値 (expected value)** μ を次式で定義する.

$$\mu = E(X) = \sum_{i=1}^{\infty} x_i f_X(x_i)$$

ii) 確率変数 X が連続型の場合, X の確率変数 f_X が $\boxed{\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty}$ を 満たすとき, X の**期待値** μ を次式で定義する.

$$\mu = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

確率変数 X の期待値は**平均値 (mean)** と呼ばれることもある.

注

常に期待値は存在するとは限らない.
$E_{\text{分布を表す}}$(確率変数)
$E_X(X) = \int x f_X(t) dt$

定理 3.18. X, Y を確率変数とし, $Y = h(X)$ とする.ただし, $h : \mathbb{R} \to \mathbb{R}$ は微分可能な単調関数とする. このとき, $E(Y) < \infty$ であれば, 以下が成り立つ.

$$E(Y) = E(h(X))$$

命題 3.19. 期待値の性質
$\phi, \psi : \mathbb{R} \to \mathbb{R}$ をボレル関数とし, $a, b \in \mathbb{R}$ を定数とする. このとき, $E(\phi(X)) < \infty, E(\psi(X)) < \infty$ であれば, 以下が成り立つ.

$$\begin{aligned} &E(a\phi(X) + b\psi(X)) = aE(\phi(X)) + bE(\psi(X)) \\ &|E(\phi(X)| \leq E(|\phi(X)|) \end{aligned}$$

Proof. 離散型

$$\begin{aligned} E_X(a\phi(X) + b\psi(X)) &= \sum_{i=1}^{\infty} (a\phi(X) + b\psi(X)) f_X(x_i) \\ &= a \sum \phi(X) f_X(x_i) + b \sum \psi(X) f_X(x_i) \\ &= aE(\phi(X)) + bE(\psi(X)) \end{aligned}$$

□

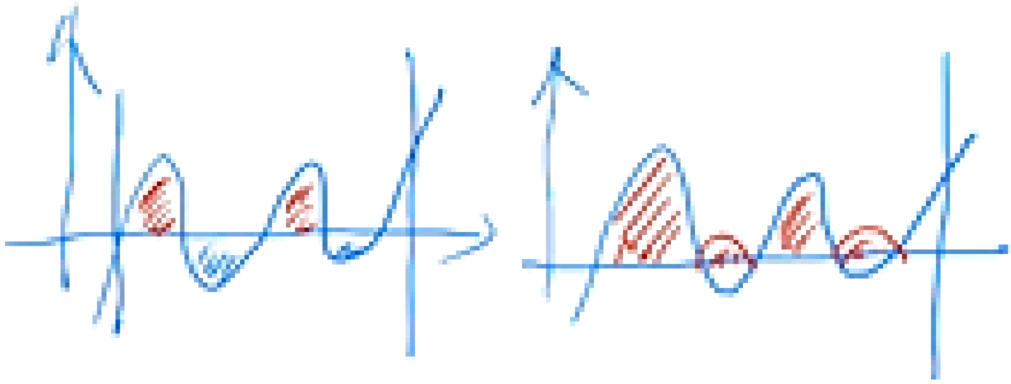
Proof. 連続型

$$\begin{aligned} E_X(a\phi(X) + b\psi(X)) &= \int (a\phi(X) + b\psi(X)) f_X(x_i) \\ &= a \int \phi(X) f_X(x_i) dx + b \int \psi(X) f_X(x_i) dx \\ &= aE(\phi(X)) + bE(\psi(X)) \end{aligned}$$

□

Proof.

$$\begin{aligned} &|\sum f(x_i)| \leq \sum |f(x_i)| \\ &|\int f(x_i)| \leq \int |f(x_i)| \end{aligned}$$



□

定義 3.10. 分散, 標準偏差

i) 確率変数 X が離散型の場合, X の確率関数 f_Xが $\sum_{i=1}^{\infty} x_i^2 f_X(x_i) < \infty (D = \{x_i \mid i = 1, 2, \dots\})$ を満たすとき, X の**分散 (varaiance)** を次式で定義し, σ^2と記す.

$$\sigma^2 = V(X) = E((X-\mu)^2) = \sum_{i=1}^{\infty} (x_i - \mu)^2 f_X(x_i)$$

ii) 確率変数 X が連続型の場合, X の確率密度関数 f_Xが $\int_{-\infty}^{\infty} x^2 f_X(x)dx < \infty$ を満たすとき, X の**分散**を次式で定義し, 同じくσ^2と記す.

$$\sigma^2 = V(X) = E((X-\mu)^2) = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx$$

iii) 確率変数 X の**標準偏差 (standard deviation)** を次式で定義し, σと記す.

$$\sigma = \sqrt{V(X)}$$

すなわち, 標準偏差は分散の正の平方根である.

例 3.4. 確率変数 X の期待値, 分散について, 次式が成り立つ.

$$E(aX+b) = aE(X)+b, \qquad V(aX+b) = a^2V(X)$$
$$V(X) = E(X^2) - (E(X))^2$$

Proof.

連続型

$$\begin{aligned} E(aX+b) &= \int (ax+b)f(x)dx \\ &= a\int xf(x)dx + b\int xf(x)dx \\ &= aE(X)+b \end{aligned}$$

※離散型も同様

□

Proof.

$$\begin{aligned} V(aX+b) &= E((aX+b-E(aX+b))^2) \\ &= E((aX+b-aE(X)+b)^2) \\ &= E(a^2(X-E(X))^2) = a^2E((X-E(X))^2) = a^2V(X) \end{aligned}$$

□

Proof.

$$\begin{aligned} V(X) &= E((X-E(X))^2) \\ &= E(X^2-2XE(X)+(E(X))^2) \\ &= E(X^2-2E(X)E(X)+(E(X))^2) \\ &= E(X^2)-(E(X))^2 \end{aligned}$$

□

例 3.5. 確率変数 X の期待値, 分散について, 次式が成り立つ.

$$E(X-\mu) = 0$$
$$E((X-c)^2) \geq E((X-\mu)^2) = V(X) \qquad (c \in \mathbb{R})$$

Proof.

$$\begin{aligned} E(X-\mu) &= E(X)-\mu \\ &= \mu-\mu=0 \end{aligned}$$

□

Proof.

$$\begin{aligned} \text{左辺}-\text{右辺} &= E((X-c)^2)-E((X-\mu)^2) \\ &= E((X-c)^2-(X-\mu)^2) \\ &= E(-2cX+c^2+2\mu X-\mu^2) \\ &= (-2c+2\mu)E(X)+c^2-\mu^2 \\ &= -2c\mu+2\mu^2-\mu^2+c^2=(c-\mu)^2 \geq 0 \\ &\text{等号成立は } c=\mu \end{aligned}$$

□

補足

データで以下のようにあらわすことができる.

$$\sum(x_i-\overline{x})=0$$
$$\sum(x_i-c)^2 \geq \sum(x_i-\overline{x})^2$$

定義 3.11. 積率
確率変数 X について, ある $n \in \mathbb{N}$ に対し, $E(X^n)<\infty$ であれば,
$E(X^n)$ を**原点まわりの n 次積率 (nth moment about the origin)**, または, **原点積率** **(nth origin moment)** という.
さらに, ある $\mu=E(X)$ が与えられるとき, $E((X-\mu)^n)$ を**平均周りの n 次積率 (nth moment about the mean)**, または, **中心積率** **(nth central moment)** という.以下, $\mu_n=E(X^n)$, $\mu'_n=E((X-\mu)^n)$ と記す
.

補足

原点まわりの n 次積率 : $\mu'_n = E(X^n)$
平均まわりの n 次積率 : $\mu_n = E((X-\mu)^n)$
$\mu'_1 = E(X) = \mu$
$\mu_2 = V(X) = \sigma^2$

定理 3.20. 積率の存在
確率変数 X および, $n(\in \mathbb{N})$ に対して, $E(X^n) < \infty$ が成り立てば, $0 < k \leq n$ に対して, $E(X^k)$ が成り立つ.つまり, n 次の積率が存在すれば, n 次以下の積率も存在する.

$$\int |x^n| f(x)dx < \infty \to \int |x^k| f(x)dx < \infty$$

命題 3.21. 確率変数 X の積率に関して, 次式が成り立つ.
(平均積率を中心積率で表す!!)

$$\begin{aligned}
\mu_2 &= \mu'_2 - (\mu'_1)^2 \\
\mu_3 &= \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3 \\
\mu_4 &= \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4
\end{aligned}$$

Proof.

$$\begin{aligned}
V(X) &= E(X^2) - (E(X))^2 \\
&= \mu'_2 - (\mu'_1)^2
\end{aligned}$$

□

Proof.

$$\begin{aligned}
\mu_3 &= E((X-\mu)^3) \\
&= E(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3) \\
&= E(X^3) - 3\mu E(X^2) + 3\mu^2 E(X) - \mu^3 \\
&= \mu'_3 - 3\mu'_1\mu'_2 + 3\mu'^2_1\mu'_1 - \mu'^3_1 \\
&= \mu'_3 - 3\mu'_1\mu'_2 + 2\mu'^2_1
\end{aligned}$$

□

定理 3.22. 平均まわりの積率と原点まわりの積率の関係
平均まわりの積率μ_kと, 原点まわりの積率μ'_kには以下の関係が成り立つ.

$$\mu_k = \sum_{i=1}^{k} (-1)^{k-i} \begin{pmatrix} i \\ k \end{pmatrix} \mu'_i (\mu'_1)^{k-i}$$
$$\mu'_k = \sum_{i=1}^{k} \begin{pmatrix} i \\ k \end{pmatrix} \mu_i (\mu'_1)^{k-i}$$

ここで, () は二項係数を表し, 形式的に $0! = 1, \mu'_0 = 1$ を仮定している.

Proof.

$$\begin{aligned}\mu_k = E((X-\mu)^k) &= E(\sum_{i=1}^{k} (-1)^{k-i} \begin{pmatrix} i \\ k \end{pmatrix} X^i \mu^{k-i}) \\ &= \sum_{i=1}^{k} (-1)^{k-i} \begin{pmatrix} i \\ k \end{pmatrix} E(X^i)(\mu'_1)^{k-i} \\ &= \sum_{i=1}^{k} (-1)^{k-i} \begin{pmatrix} i \\ k \end{pmatrix} \mu'_i (\mu'_1)^{k-i}\end{aligned}$$

□

Proof.

$$\begin{aligned}\mu'_k &= E(X^k) \\ &= E((X-\mu+\mu)^k) \\ &= \sum_{i=1}^{k} \begin{pmatrix} i \\ k \end{pmatrix} E((X-\mu)^i)\mu^{k-i} \\ &= \sum_{i=1}^{k} \begin{pmatrix} i \\ k \end{pmatrix} E((X-\mu)^i)(\mu'_1)^{k-i} \\ &= \sum_{i=1}^{k} \begin{pmatrix} i \\ k \end{pmatrix} \mu^i (\mu'_1)^{k-i}\end{aligned}$$

□

補足

$$\begin{pmatrix} i \\ k \end{pmatrix} =_k C_i$$

定義 3.12. 歪度と尖度
確率変数 X について, $E(X^4) < \infty$ のとき, X の**歪度 (skewness)**γ_1および, **尖度 (kurtosis)**γ_2を以下の式で定義する.

$$\gamma_1 = E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right) = \frac{\mu_3}{(\mu_2)^{2/3}}, \qquad \gamma_2 = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right) = \frac{\mu_4}{(\mu_2)^2}$$

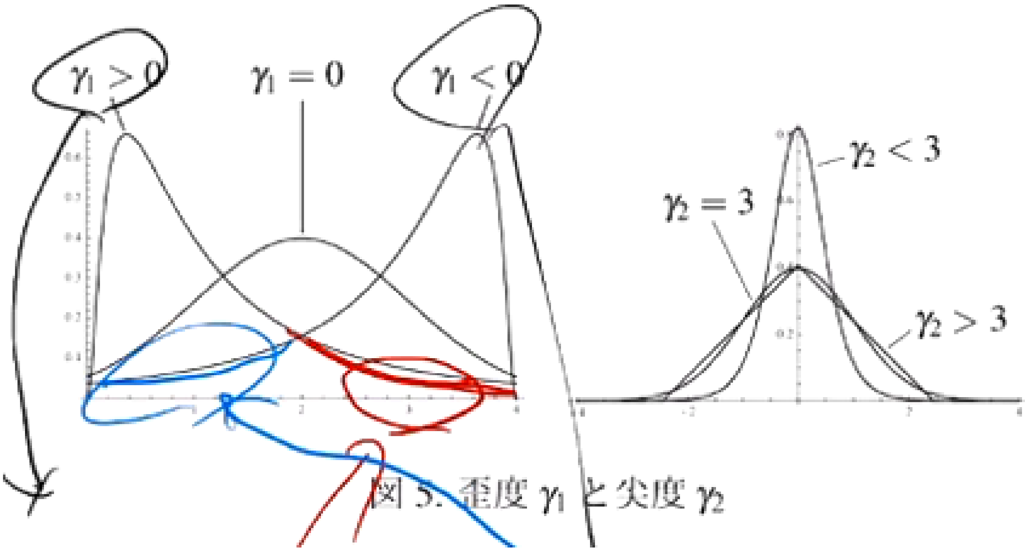


図 5: 歪度 γ_1 と尖度 γ_2

定義 3.15. 標準化
確率変数 X について, $E(X) = \mu, V(X) = \sigma^2 (\sigma > 0)$ とする. このとき, 変換

$$Z = \frac{X-\mu}{\sigma}$$

を X の**標準化 (standardization)** とよぶ.

補足

$$E(X) = 0$$
$$V(X) = 1$$

命題 3.24. 標準化された確率変数の期待値, 分散
標準化された, 確率変数 Z の期待値 $E(Z)$, 分散 $V(Z)$ は以下で与えられる.

$$E(Z) = 0, \qquad V(Z) = 1$$

Proof.

$$\begin{aligned} E(Z) &= E\left(\frac{X-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma}(E(X)-\mu) = 0 \end{aligned}$$

$$\begin{aligned} V(Z) &= E\left(\left(\frac{X-\mu}{\sigma}\right)^2\right) \\ &= \frac{1}{\sigma^2}E((X-\mu)^2) \\ &= \frac{\sigma^2}{\sigma^2} = 1 \end{aligned}$$

□

また, Z を変換した確率変数 $X = \mu + \sigma Z$ の期待値 $E(Z)$, 分散 $V(Z)$ は以下で与えられる.

$$E(Z) = \mu, \qquad V(Z) = \sigma^2$$

Proof.

$$\begin{aligned} &X = \mu + \sigma Z \\ &E(X) = \mu + \sigma E(Z) = \mu \qquad (\because E(Z) = 0) \\ &V(X) = \sigma^2 V(Z) = \sigma^2 \qquad (\because V(Z) = 1) \end{aligned}$$

□

定義 3.16. スチューデント化
確率変数 X について, $E(X)=\mu, V(X)=\sigma^2(\sigma>0)$ とする. このとき, 変換

$$T=\frac{X-\mu}{\sqrt{Y}}$$

を X の**スチューデント化 (studentization)** とよぶ. ただし, Y は, $Y \xrightarrow{P} \sigma^2$ を満たす確率変数である.

定義 3.17. $\boxed{\text{積率母関数}}$ とキュムラント母関数
確率変数 X について, $E(e^{tX})<\infty$ が t の原点を含むある区間 I で成り立つとする. この区間で定義される t の関数

$$m_X(t)=E(e^{tX}) \qquad (t\in I)$$

を X の**積率母関数 (moment generating function)** とよぶ. さらに,

$$\Psi_X(t)=log m_X(t) \qquad (t\in I)$$

を**キュムラント母関数 (cumulant generating function)** とよぶ.
定理 3.25. 確率変数 X について, $E(X^n)<\infty$ ならば, X の積率母関数について以下が成り立つ. (使い方のみ)

$$\left.\frac{d^k}{dt^k}m_X(t)\right|_{t=0}=E(X^k)=\mu'_k \qquad (k=1,2,\dots n)$$

定理 3.26. 積率母関数の一致性
2 つの確率変数 X_1, X_2 に対して, 積率母関数 m_{X_1}, m_{X_2} が存在すると仮定する. このとき, 確率分布 P_{X_1}, P_{X_2} が一致するための必要十分条件は, 積率母関数 m_{X_1}, m_{X_2} が一致することである.

補足

P_{X_1} と P_{X_2} が一致 $\leftrightarrow$ m_{X_1} と m_{X_1} が一致

定理 3.27. 確率変数 X について, $E(X^n) < \infty$ 　ならば, X のキュムラント母関数に関して, 以下が成り立つ.

$$\left.\frac{d^k}{dt^k}\psi_X(t)\right|_{t=0} = \mathcal{K}_k \qquad (k = 1, 2, \dots n)$$

命題 3.28. 積率母関数, キュムラント母関数の線形変換
$a, b \in \mathbb{R}$ を定数とし, $Y = aX + b$ とする. このとき, 確率変数 X, Y の積率母関数, および, キュムラント母関数に関して, 以下が成り立つ.

$$\begin{aligned} m_Y(t) &= e^{bt} m_X(at) \\ \psi_Y(t) &= bt + \psi_X(at) \end{aligned}$$

Proof.

$$\begin{aligned} m_Y(t) &= E(e^{ty}) \\ &= E(e^{t(aX+b)}) \\ &= E(e^{bt} \cdot e^{tax}) \\ &= e^{bt} E(e^{tax}) \\ &= e^{bt} m_X(at) \end{aligned}$$

□

命題 3.29. 積率母関数とキュムラント母関数について, 以下の式が成り立つ.

$$\begin{aligned} \psi'_X(0) &= m'_X(0) = E(X) \\ \psi''_X(0) &= m''_X(0) - (m'_X(0))^2 = V(X) \end{aligned}$$

ここでの, $'$ および, $''$ はそれぞれ, 1 階微分, 2 階微分を表している.

定理 3.37. マルコフの不等式 (Markov's inequality)
$h:\mathbb{R}\to\mathbb{R}$ を非負関数とする. 確率変数 X について, $E(h(X))<\infty$ ならば, 任意の$\epsilon>0$ に対して,

$$P(h(X)\geq\epsilon)\leq\frac{E(h(X))}{\epsilon}$$

が成り立つ. この不等式において, X を $X-E(X)$ で置き換え, $h(X)=x^2,\epsilon=a^2$とした

$$P(|X-E(X)|\geq a)\leq\frac{V(X)}{a^2}$$

は**チェビシェフの不等式 (Chebyshev's inequality)** とよばれる.

Proof.

$$\begin{aligned}E(h(X))=\int_{-\infty}^{\infty}h(x)f(x)dx&\geq\int_{h(x)\geq\epsilon}h(x)f_X(x)dx\\&\geq\int_{h(x)\geq\epsilon}\epsilon f_X(x)dx\\&=\epsilon\int_{h(x)\geq\epsilon}f_X(x)dx\\&=\epsilon P(h(x)\geq\epsilon)\end{aligned}$$

□

定理 3.38. ジャンセンの不等式 (Jensen's inequality)
(凸関数と期待値の関係)
$g:\mathbb{R}\to\mathbb{R}$ を凸関数とする. 確率変数 X について, $E(X),E(g(X))<\infty$ であれば,

$$g(E(X))\leq E(g(X))$$

が成り立つ. ここで, 区間 I で定義された実数関数 g が凸関数であるとは, $a,b\in I,t\in(0,1)$ について,

$$g(ta+(1-t)b)\leq t(g(a))+(1-t)g(b)$$

が成り立つことである.

定理 3.39. ヘルダーの不等式 (holder's inequality)
(コーシーシュワルツの一般形)
確率変数 X, Y について, $E(|X|^p), E(|X|^q) < \infty$ であり, $1 < p < \infty, \frac{1}{p} + \frac{1}{q} = 1$ であれば,

$$E(|XY|) \leq (E(|X|^p))^{1/p}(E(|Y|^q))^{1/q}$$

が成り立つ./; この不等式において, $p = q = 2$ とした

$$\underline{(E(|XY|))^2 \leq E(X^2)E(Y^2)}$$

は, **コーシーシュワルツの不等式 (Cauchy-Schwarz inequality)** とよばれる.

定理 3.40. ミンコフスキーの不等式 (Minkowski's inequality)
確率変数 X, Y について, $E(|X|^r), E(|Y|^r) < \infty$　であり, $1 \leq r < \infty$ であれば,

$$(E(|X+Y|^r))^{1/r} \leq (E(|X|^r))^{1/r} + (E(|Y|^r))^{1/r}$$

が成り立つ.

定理 3.41. リアプノフの不等式 (Lyapunov's inequality)
確率変数 X について, $E(|X|^q) < \infty$ であり, $1 < p < q < \infty$ であれば,

$$(E(|X|^p))^{1/p} \leq (E(|X|^q))^{1/q}$$

が成り立つ.