

入門 統計学

第2章 確率分布入門

著 橋本智雄

発表者 野口耕二朗

目次

- 確率の概念
- 確率変数と確率分布

確率の概念

基本的な用語

- 試行・・・一定の条件の下での実験や調査, 観測
- 標本点・・・ある試行の結果を ω_i などの記号で表したもの
- 標本空間・・・標本点全体の集合 全体集合 $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$
- 事象・・・試行の結果のある集合 事象 $A = \{\omega | \omega < 10\} \subset \Omega$

統計的確率

● 定義

- ある試行を n 回繰り返し行なったとき、事象 A の生起回数が n_A であり、 n を大きくしていくとき相対度数 n_A/n が一定の値 p に近づくならば、この値 p を事象 A の確率といい、 $P(A) = p$ で表す。

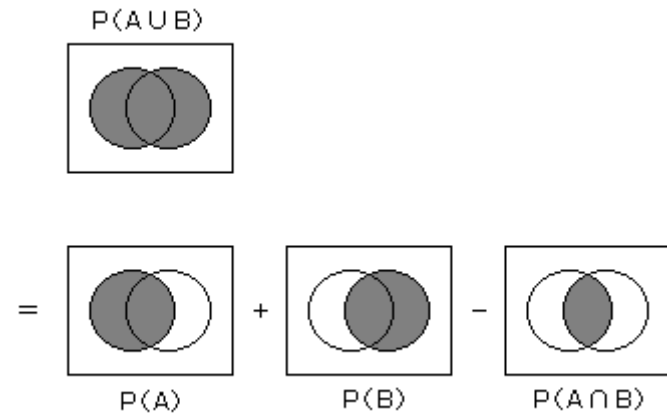
● 性質

- I. $0 \leq P(A) \leq 1$
- II. $P(\phi) = 0, P(\Omega) = 1$
- III. A と B が排反ならば、 $P(A \cup B) = P(A) + P(B)$

基本的な定理

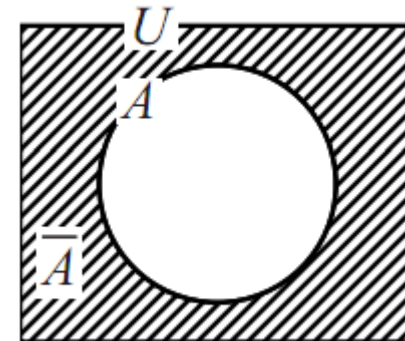
● 加法定理

➤ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



● 余事象の確率

➤ $P(\bar{A}) = 1 - P(A)$



条件付き確率

- 2つの事象 A と B があつて、事象 B が起こつたという条件の下での事象 A の確率を $P(A|B)$ で表し、以下のように定義する

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$

- 乗法公式

- $P(A \cap B) = P(B)P(A|B)$

- $P(A|B) = P(A)$ のとき、2つの事象 A と B は独立

- $P(A \cap B) = P(A)P(B)$ が成り立つ

確率変数と確率分布

確率変数

● 離散型

- 飛び飛びの値をとる
- 確率関数は $P(X = x)$ ($x = 1, 2, 3, \dots$) と表される

● 連続型

- とり得る値が連続的である
- 確率関数は $P(a \leq X \leq b) = \int_a^b f(x) dx$ と表される
- $f(x)$ は確率密度関数といい

$$\int_{-\infty}^{\infty} f(x) dx = 1, f(x) \geq 0$$

という性質を持つ

分布関数

- 離散型

- $F(x) = P(X \leq x) = \sum_{k \leq x} p(k)$

- 連続型

- $F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$

- 基本的性質

- I. $0 \leq F(x) \leq 1$

- II. $F(x)$ は x について単調増加関数

- III. $\lim_{x \rightarrow \infty} F(x) = 1, \lim_{x \rightarrow -\infty} F(x) = 0$

- IV. $F(x)$ は右側連続である. すなわち, $\lim_{x \rightarrow a+0} F(x) = F(a)$

離散型の重要な確率分布

- 2点分布 $B(1, p)$

- X のとり値が0と1で, $P(X = 1) = p, P(X = 0) = 1 - p$ すなわち

$$p(x) = P(X = x) = p^x(1 - p)^{1-x} \quad (x = 0, 1)$$

で与えられる分布

- 2項分布 $B(n, p)$

- X のとり値が0, 1, ..., n のいずれかで, 確率関数 $p(x)$ が

$$p(x) = P(X = x) = {}_nC_x p^x (1 - p)^{n-x}$$

で与えられる分布

離散型の重要な確率分布

●ポアソン分布 $P_0(\lambda)$

- X のとり得る値が $0, 1, \dots, n$ のいずれかで、確率関数 $p(x)$ が

$$p(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!} \quad (\lambda > 0; x = 0, 1, 2, \dots)$$

で与えられる分布

- 「 $np = \lambda$ (一定)として、 n が十分に大きいとき、2項分布 $B(n, p)$ はポアソン分布 $P_0(\lambda)$ により近似される。」という定理がある

●幾何分布 $G(p)$

- 無制限に試行を繰り返すため、 X のとり得る値が $0, 1, \dots$ であり、確率変数 $p(x)$ は

$$p(x) = P(X = x) = p^x(1 - p)^{1-x} \quad (x = 1, 2, 3, \dots)$$

で与えられる分布

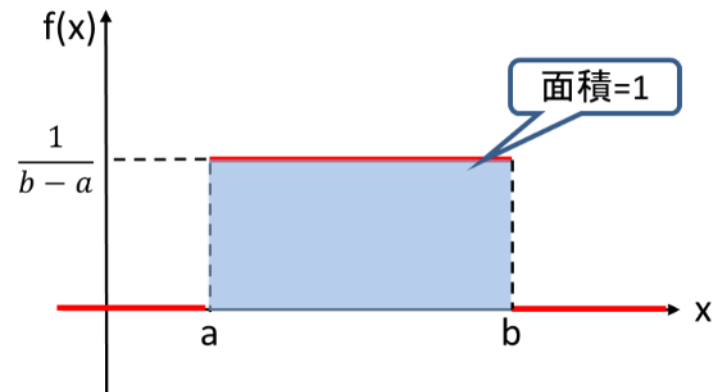
連続型の重要な確率分布

- 一様分布 $U[a, b]$

- X の確率密度関数が

$$\begin{aligned} f(x) &= \frac{1}{b-a} & (a \leq x \leq b) \\ &= 0 & (x < a, b < x) \end{aligned}$$

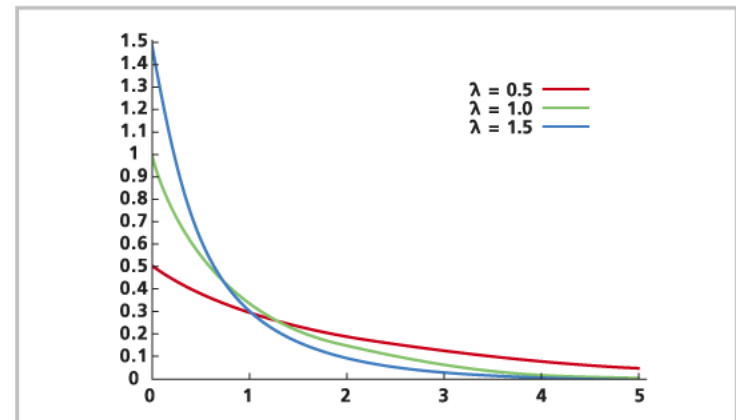
で与えられる分布



- 指数分布 $EXP(\lambda)$

- X の確率密度関数が

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} & (0 \leq x) \\ &= 0 & (x < 0) \end{aligned}$$



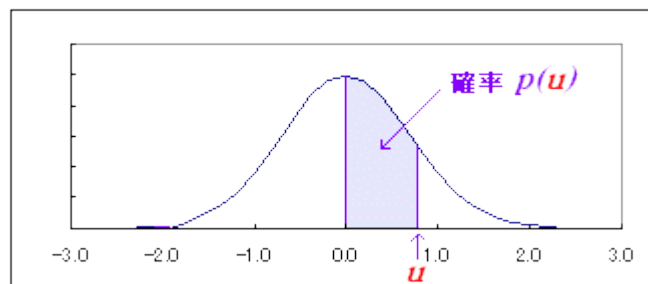
正規分布 $N(\mu, \sigma^2)$

- 密度関数がすべての x に対して

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

で与えられる分布. μ は平均, σ^2 は分散である

- 0から3.09までの標準正規分布の分布関数の積分値が与えられた正規分布表があり, 正規分布を用いた統計処理では重要になる



平均と分散

● 平均

- 離散型 $\cdots E(X) = \sum_i x_i p(x_i)$
- 連続型 $\cdots E(X) = \int_{-\infty}^{\infty} x f(x) dx$

● 分散

- 離散型 $\cdots V(X) = \sum_i (x_i - E(X))^2 p(x_i)$
- 連続型 $\cdots V(X) = \int_{-\infty}^{\infty} (x_i - E(X))^2 f(x) dx$