確率変数と確率分布

確率変数：関数

事象 → 数値

確率的挙動を

確率分布関数として扱う

2.1 確率変数と確率分布

例: サイコロ投げ

サイコロの出目の確率を記述する

確率変数 (r.v.と略)

X: サイコロの取りうる値として

$$\underbrace{P(X=x) = 1/6}_{} \quad \underbrace{(x=1,\cdots,6)}_{\text{離散型}}$$

確率変数　値

(discrete) 離散型 r.v.

r.v. Xが可算個の離散値を取りうる

離散型r.v.の確率を表現する (by 関数)

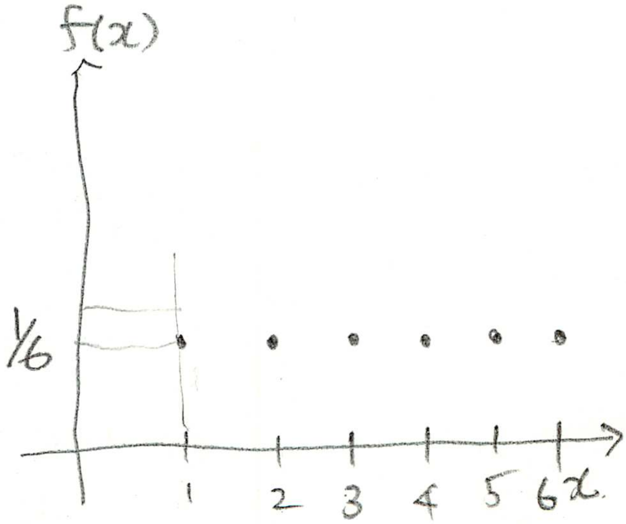
確率関数

$$f(x) \equiv \begin{cases} P(X=x) & (x=x_1, x_2, \cdots) \\ 0 & (x \neq \ 〃 \ ) \end{cases}$$

def
定ギ

$$f(x_i) > 0$$

$$\sum_{i=1}^{\infty} f(x_i) = 1$$



連続型
(contineous)

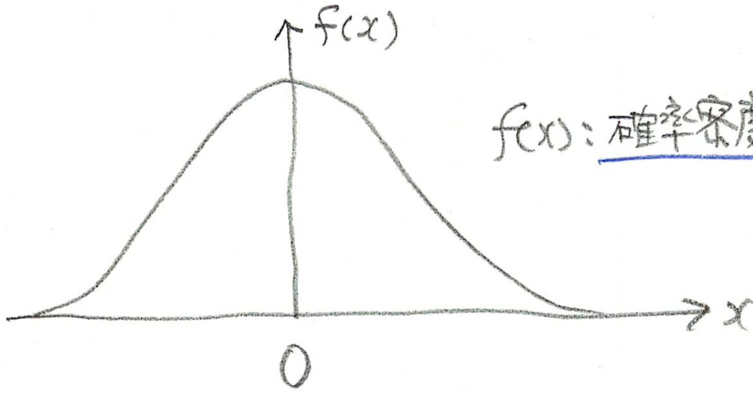
r.v. Xの取りえる値が連続のとき。

例：計測誤差を表わす r.v. X.

Xの値は $x=0$ の近傍の実数値を取ると考えられる

0のまわりが数多く現われて、

0から離れると出現数が少なくなる



$f(x)$：確率密度関数

Xの値が区間 $[a,b]$ にある確率を $f(x)$ を用いて表わす：

確率密度関数 $f(x)$：

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \quad \longleftrightarrow \quad P(-\infty < X < \infty) = 1$$

$[a,b]$ の巾が十分に小さい時

近似式

$$P(a < X \leq b) \simeq \underbrace{f(\xi)}_{\text{密度}}(b-a) \quad (a < \xi < b)$$

$$P(a < X < a+dx) = f(x)\,dx$$

教科書での記法について

これを採用

連続型 r.v. の区間 $\underline{P(a < X \leq b)} = P(a \leq X \leq b)$

$\because P(X = a) = 0$

離散型とは異なる

<u>密度関数</u>という総称を用いる

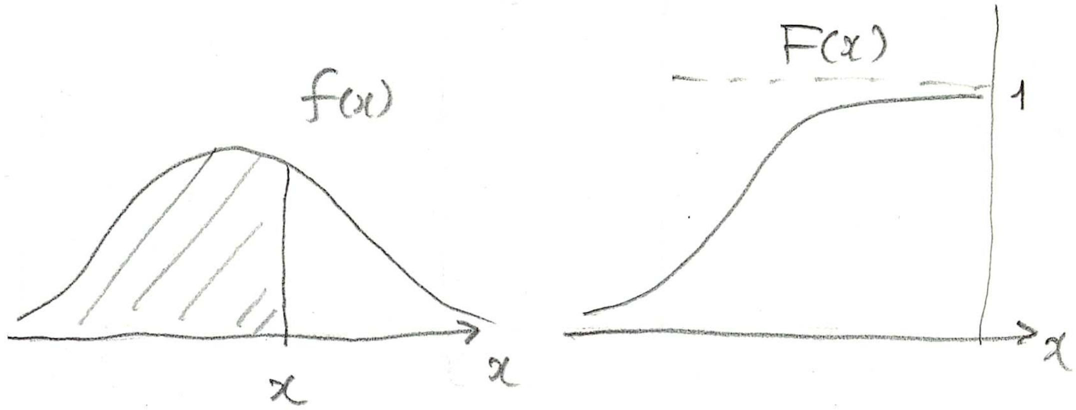
d.f.

(時に p.d.f)

$$\begin{cases} \text{離散型 r.v. の確率関数 (p.f.)} \\ \text{連続型 r.v. の確率密度関数 (p.d.f)} \end{cases}$$

(累積)
分布関数
(c.d.f)

$$F(x) \equiv P(X \leq x)$$



F(x)の性質

(1) $F(a) \leq F(b) \quad (a < b)$

(2) $0 \leq F(x) \leq 1$

(3) $F(x)$は連続

離散型
分布関数

$$\begin{cases} F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) \\ f(x) = F(x) - F(x_-) \end{cases}$$

xの1つ前のx

差分

連続型
〃

$$\begin{cases} F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt \\ f(x) = \dfrac{d}{dx} F(x) \end{cases}$$

確率分布

確率変数のとり得る値,
その値となる確率,
確率を $\left\{\begin{array}{l}\text{密度関数} \\ \text{分布 〃}\end{array}\right\}$ で記述する。

2.2 期待値, 平均, 分散

例: サイコロ投げ

サイコロを n 回振って,
出た目 (x_i) の平均はいくらか

① 目 i が出る確率は 1/6 なので $\left(\sum_1^6 i\right)/6 = 3.5$

② $\sum_{i=1}^{6} P(X=i)\cdot i = 3.5$

平均の求め方の一般化

$\begin{cases} \text{r.v. } X \\ \text{d.f. } f(x) \end{cases}$

離散型

離散型を想定.
X が取りえる値の集合を $\mathcal{X} = \{x_1, x_2, \cdots\}$ とする

平均の定義

$$\text{平均 } \mu \equiv \sum_{i=1,2,\cdots} x_i \cdot f(x_i)$$

$$= \sum_{x \in \mathcal{X}} x \cdot f(x)$$

総和表現

よく似ている

リーマン和

積分表現

連続型

X が取りえる値の領域 $\mathcal{X}$ で表わす.

$$\text{平均 } \mu \equiv \int_{\mathcal{X}} x f(x)\, dx$$

期待値

E[X]　r.v. X の期待値 E[X]

$$\mu = E[X] = \int x f(x)\,dx$$

E[g(x)]　統計値 g(X) の期待値

$$E[g(X)] = \int g(x)\cdot f(x)\,dx$$

E[ ]の線型性

$$E[a\cdot g(X) + b\cdot h(X)]$$

$$= \int \{a g(x) + b h(x)\} f(x)\,dx$$

$$= a\int g(x)\cdot f(x)\,dx + b\int h(x)\cdot f(x)\,dx$$

$$= a E[g(X)] + b E[h(X)]$$

スカラー倍

作用前の＋が
〃 後の＋になる．

分散 (σ^2)　$V[X] = E[(X-\mu)^2]$

$$= \int (x-\mu)^2 f(x)\,dx$$

variance σ^2

μからの距離の2乗の平均

データの散りばり度, 広がり
慣性モーメント（回転体の回転エネルギー）

二種の普通でないサイコロによる考察、

1種目: D_1　⚀ ⚁ ⚂ ⚀ ⚁ ⚂ ,　$\mu = 2$

2種目: D_2　⚀ ⚁ ⚂ ⚁ ⚁ ⚁ ,　$\mu = 2$

$$\sigma_{D_1}^2 = (1-2)^2 \cdot \frac{1}{3} + (2-2)^2 \cdot \frac{1}{3} + (3-2)^2 \cdot \frac{1}{3} = \frac{2}{3}$$

$$\sigma_{D_2}^2 = (1-2)^2 \frac{1}{6} + (2-2)^2 \cdot \frac{4}{6} + (3-2)^2 \cdot \frac{1}{3} = \frac{1}{3}$$

$\sigma_{D_1}^2 > \sigma_{D_2}^2$: 何故なら 平均である2が離れた
1や3の出る確率が高いから

標準偏差 σ　$\sigma = \sqrt{\sigma^2}$　r.v. Xと次元を合せた量,

r.v.の k次モーメント

$E[X^k]$

平均は 一次モーメント

k次の中心モーメント

$E[(X-\mu)^k]$

分散は 二次の 中心モーメント

モーメント間の関係、

$$
\begin{aligned}
\sigma^2 &= E[(X-\mu)^2] \\
&= E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - 2\mu E[X] + E[\mu^2] \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2 \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

r.v.の
変数変換
(一次変換)
(=線形変換)

r.v. X に対し r.v. Y を

$Y = aX + b$,　　　a, b は数.

$E[X] \equiv \mu_X$, $E[Y] \equiv \mu_Y$ と記して,
$V[X] \equiv \sigma_X^2$　$V[Y] = \sigma_Y$

μ_Y は $a \cdot \mu_X + b$ であろう

σ_Y^2 は 平行移動に関して不変なので

$\sigma_Y^2 = a^2 \sigma_X^2$ であろう

$$\mu_Y = E[Y] = E[aX+b]$$
$$= aE[X] + bE[1]$$
$$= a\mu_x + b$$

$$\sigma_Y^2 = V[Y] = E[(Y-\mu_Y)^2]$$
$$= E[\{(aX+b)-(a\mu_x+b)\}^2]$$
$$= E[\{a(X-\mu_x)\}^2]$$
$$= a^2 E[(X-\mu_x)^2]$$
$$= a^2\sigma_X^2$$

標準化

r.v. X の $\left\{\begin{array}{l}\text{平均が}\ \mu \\ \text{分散が}\ \sigma^2\end{array}\right\}$ の時,

Xの変数変換（標準化）

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

は

$$E[Z] = \frac{1}{\sigma}E[X-\mu] = 0$$
$$V[Z] = E[(Z-0)^2]$$
$$= E\left[\frac{(X-\mu)^2}{\sigma^2}\right]$$
$$= \frac{\sigma^2}{\sigma^2} = 1$$

逆標準化

$z = \dfrac{X-\mu}{\sigma}$ の逆変換.

$X = \sigma z + \mu$ は,

$\mu_z = 0,\ \sigma_z^2 = 1$ の時.

$\mu_X = \mu,\ \sigma_X^2 = \sigma^2$ になる。

2.9 確率変数と確率分布と確率空間

サイコロ投げ　$(\Omega, \overset{\Sigma}{\mathcal{A}}, P)$ の構成

確率空間

標本空間　$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$

ω_x：サイコロの目が x であるという標本点とした。

Σ-集合体　$\mathcal{A} = 2^{\Omega}$（Ω の全部分集合）

確率変数　$X(\omega_x) = x \quad (x = 1, 2, 3, 4, 5, 6)$

点→値へ

$X = x$ と記述してきた意味は

一点からなる事象　$A_x = \{\omega_x\} = \{\omega \in \Omega \mid X(\omega) = x\}$

$P(X=x)$：　$P(A_x) = P(X=x) = 1/6$

偶数の目の事象

$$A_{\{2,4,6\}} = A_2 \cup A_4 \cup A_6$$

$$\to P(A_{\{2,4,6\}}) = P(A_2 \cup A_4 \cup A_6)$$

$$= P(A_2) + P(A_4) + P(A_6)$$

確率変数と確率分布と確率空間
X　P　(modelとも言う)

確率空間 $(\Omega, \mathcal{A}, P)$ があるとする.

Xは連続値をとり得る

$X : \mathcal{A} \to R$

$$X^{-1}((-\infty, x]) = \{\omega \mid X(\omega) \leq \underset{\in R}{x}\} \in \mathcal{A}$$

$$\therefore P(\underbrace{X^{-1}((-\infty, x])}_{\text{標本空間上の事象}}) = P(\{\omega \mid \underbrace{X(\omega) \leq x}_{\text{実数上の区間}}\})$$