

# 1 確率空間

## 1.1 もっと数学的な定義

標本点：指向によって生じる個々の結果)

根元事象：標本点 1 つの事象

標本空間：標本点の集合 集合

事象の集合 F：標本空間  $\Omega$  の部分集合族  $2^{\Omega}$

部分集合族：部分集合の集合、べき集合

$F < 2^{\Omega}$  である

有限集合のときは  $F = 2^{\Omega}$

事象：標本空間  $\Omega$  の部分集合 A、または、事象の集合 F の各要素  $A \in F$

空事象：標本点を持たない事象

確率測度：各事象に対して、0 以上の数に対応させる実関数 P

確率空間：三つ組  $(\Omega, F, P)$  を確率空間と呼ぶ ← 確率の世界的なもの

### 1.1.1 例

コイン投げで、表を O とし、裏を 1 とする。

根元事象：「表が出る」 → 0 「裏が出る」 → 1

標本空間： $\Omega = \{0, 1\}$

事象の集合： $F = 2^{\Omega} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\} \rightarrow$  できる集合全て、今回だと「何も取らない」「0 または 1 だけ」「両方」

確率測度：例えば、 $P(\{0\}) = 2/3$ ,  $P(\{1\}) = 1/3$ ,  $P(\emptyset) = 0$ ,  $P(\{0, 1\}) = 1 \rightarrow$  どっちかは必ず出る

### 1.1.2 同様に確からしい

n 個の標本点からなる標本空間  $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$  において、標本点  $\omega_i$  のそれぞれが生起する確率が等しい時、すなわち

$P(\omega_1) = P(\omega_2) = \dots = P(\omega_n)$

が成り立つ時、「同様に確からしい」という。

### 1.1.3 例題 2

コインを 2 枚投げた時、結果を (a, b) であらわす。

根元事象： $\{(0, 0), \{(0, 1)\}, \{(1, 0)\}, \{(1, 1)\}\}$

$\Omega = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$

$F = 2^{\Omega}$  個ある

$F = \{\emptyset, \{(0, 0)\}, \{(0, 1)\}, \{(1, 0)\}, \{(1, 1)\}, \{(0, 0), (0, 1)\}, \{(0, 0), (1, 0)\}, \{(0, 0), (1, 1)\}, \{(0, 1), (1, 0)\}, \{(0, 1), (1, 1)\}, \{(1, 0), (1, 1)\}, \{(0, 0), (0, 1), (1, 0)\}, \{(0, 0), (0, 1), (1, 1)\}, \{(0, 0), (1, 0), (1, 1)\}, \{(0, 0), (0, 1), (1, 0), (1, 1)\}\}$

## 1.2 確率の公理的定義

コルモゴロフの定義

公理 1：全ての要素  $E_i$  に対して  $0 \leq P(E_i)$  である

公理 2 :  $P(\Omega) = 1$

公理 3 : お互いに排反な事象  $E_i$  について、 $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$

※ 「互いに排反な事象  $E_i$  とは

すべての  $n \neq m$  に対して

$E_n \cap E_m = \emptyset$

が成り立つこと

公理 1 3 を満たさなければ、確率とは言えない。

公理から次が導き出される

1.  $P(\emptyset) = 0$

2. 単調性 :  $A \subseteq B$  ならば、 $P(A) \leq P(B)$

3. 上界 :  $0 \leq P(A_i) \leq 1$

4. 加法定理  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

5. 余事象の法則  $P(A^c) = 1 - P(A)$

\*\*\*\*\*宿題 公理 1 3 のみを用いて、1~5 の関係が成り立つことを示せ。A/B はつかう ??

## 2 確率変数と確率分布

離散変数の場合を考える

### 2.1 確率変数

定義 ( いくつかある )

1. 確率変数とは、試行する実験結果として取り得る値を表す変数 サイコロ 1~6

2. 事象 X が確率を伴った出現する時、X を確率変数と呼ぶ

3. 標本空間  $\Omega$  の各要素に対して、実数の集合 R の 1 つの値を割り当てる関数として、確率変数 X を定義する

$X : \Omega \rightarrow R$

$y = f(x)$   $x, y$  は変数  $f : R \rightarrow R$

※ 「事象を数値で表したもの」を確率変数とする

確率変数は大文字の X で表し、その実現値は小文字で表す

サイコロ

確率変数を X とする。

$P(X=1), P(X=2), \dots, (P(X=6))$

$P(X)=x$

## 3 10/19 :

前回、三つ組  $(\Omega, F, P)$  ですべての確率を表すことができる

### 3.1 離散確率分布

確率分布とは . . .

各事象 J に対して、それが起こる確率を割り当てたもの。つまり、確率測度を定義したものである。

確率分布表やグラフ、数式で与えることができる。

## 3.2 例題

### コイントス

例題	
事象	裏 表
確率変数 X	0 1
確率 P	$\frac{1}{2} \quad \frac{1}{2}$
	↑ $\frac{1}{3} \quad \frac{2}{3}$ 今までも
	確率分布表

### 2つのサイコロ

2つのサイコロを投げた時、出た目の数を X と Y で表す

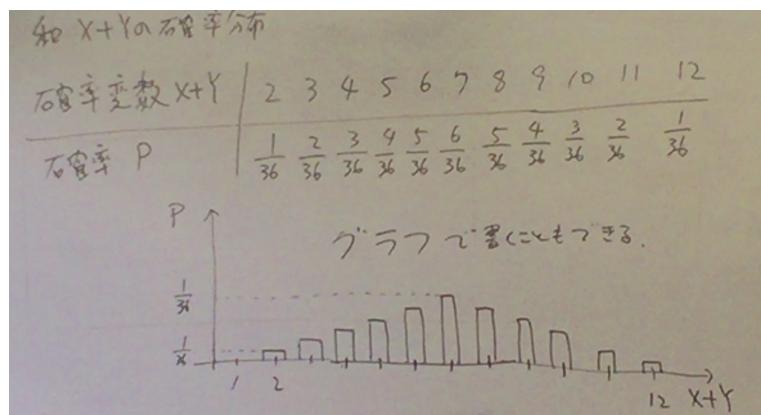
※ X と Y はそれぞれ確率変数

「出た目の和」を考えたい

和 X+Y も確率変数である

事象	1	2	3	4	5	6
確率変数 X	1	2	3	4	5	6
確率 P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

### 和 X+Y の確率分布



確率分布が数式で表せる場合

### 離散一様分布

n 個の事象が等しく起こる確率分布

$$P(X = r) = \frac{1}{n}, \quad r = 1, 2, 3, \dots, n$$

$P(X = r)$  は  $P(r), P_x(r)$  と書くこともある

↑ 「確率変数 X が r となる確率」

#### 3.2.1 例：サイコロ

サイコロの目が r となる確率

$$P(X = r) = \frac{1}{6}$$

### 3.3 二項分布

確率  $p$  で起こる事象が、 $n$  回の試行中  $r$  回起こる確率分布

$$P(X = r) = P_x(r) = {}_n C_r P^r (1 - P)^{n-r}$$

#### 3.3.1 例： $n$ 回のコイン投げ

コインを 10 回投げた時表が 3 回出る  $P(X = 3) = {}_{10} C_3 (\frac{1}{2})^3 (1 - (\frac{1}{2})^{10-3})$

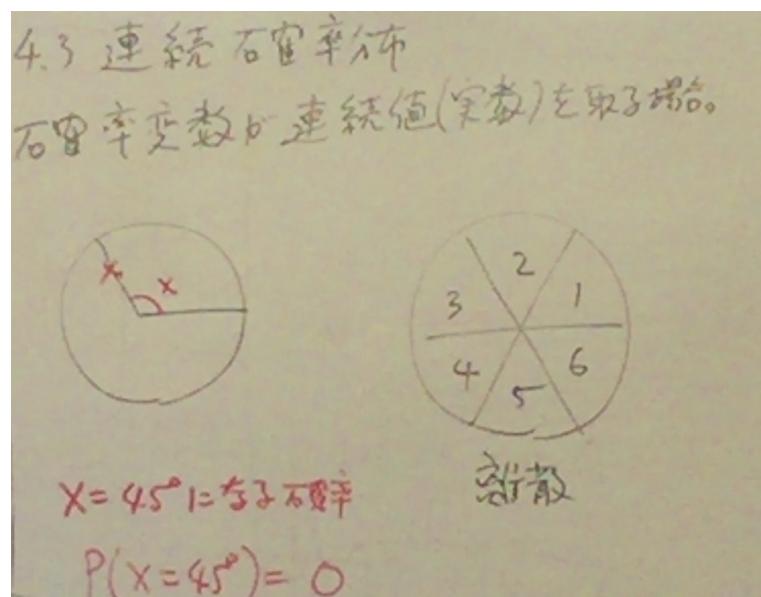
### 3.4 ポアソン分布

単位時間中に平均で  $\lambda$  回起こる事象が  $k$  回起こる確率分布

$$P(X = r) = \frac{\lambda^k}{k!} e^{-\lambda}$$

### 3.5 連続確率分布

確率変数が連続値（実数）をとる場合



連続確率分布のときは、ある値  $r$  を取り確率は 0 であることが多い。

↓

「 $X = r$  となる確率」には意味はない

「 $a \leq x \leq b$  となる確率」を考えることができる

#### 3.5.1 定義

確率変数  $X$  に対して、 $a \leq x \leq b$  となる確率が

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

で与えられる時、 $f(x)$  を  
確率密度関数

と定義する。

ただし、 $f(x) \geq 0$  とする

全区間で積分すると

$$\int_a^b f(x)dx = 1$$

である

※  $f(x)$  の値は 1 を超えることもある

連続系の場合は、確率密度関数を考える

※確率密度関数であることの必須条件

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

$f(x)$  が確率密度関数ならば、1 と 2 を満たす

### 3.5.2 例題

区間  $[0:1]$  において、 $f(x) = x$  は確率密度関数の必要条件を満たすか？

1.  $f(x) \geq 0$  は満たす ( $0 \leq x \leq 1$ )
2.  $\int_{-\infty}^{\infty} f(x)dx = \int_0^1 xdx = 1/2 \neq 1$

これより、 $f(x) = x$  は必要条件を満たしていない

### 3.5.3 連続確率分布

連続一様分布

区間  $[a:b]$  で一様である時

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{その他} \end{cases}$$

### 3.5.4 正規分布 (ガウス分布)

平均  $\mu$ , 分散  $\sigma^2$  の正規分布

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

### 3.5.5 指数分布

単位時間当たりに  $\lambda$  回起こる事象が怒った後つぎに発生するまでの生起間隔が  $x$  単位時間である確率

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{その他} \end{cases}$$

### 3.5.6 ガンマ分布

$$f(x) = x^{k-1} \frac{e^{-\frac{x}{\theta}}}{\Gamma(k) \theta^k} (x \geq 0)$$

$\Gamma(k)$  はガンマ関数

$$\Gamma(k) = (k-1)!$$

カーネル分布

$$f(x) = x^{k-1} \frac{e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k}, \quad x > 0$$

$\Gamma(k)$  は カーネル関数

$$\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx, \quad (k > 0)$$

- $\Gamma(k+1) = k \Gamma(k)$
- $\Gamma(n) = (n-1)!$   $n$  は整数
- $\Gamma(1) = 1$

\*\*\*\*\*宿題。。。画像

4 10/26

#### 4.1 累積分布関数

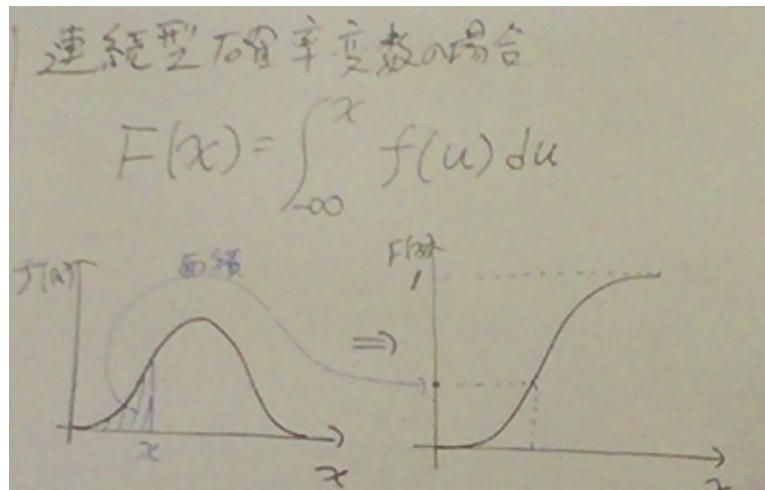
累積分布関数  $F(x)$  ・・・ ある値が  $x$  以下である確率  $P(X \leq x)$

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

と定義する

連続型確率変数の場合

$$F(x) = \int_{-\infty}^x f(u) du$$



#### 4.2 期待値と分散

確率分布の特徴を考えてみる

代表値

- 平均値 (average) : 標本平均。全てのデータの総和をデータ数で割ったもの

- ・中央値 (median) : データを大きさ順に並べた時、真ん中に来る値のこと
- ・最頻値 (mode) : 最もデータ数が多い値のこと

#### 4.2.1 例

データ (1,1,2,4,5,8,9,10,11) 平均値=5.66

中央値=5 番目が真ん中=5

最頻値=1

### 4.3 期待値

期待値は、ある確率分布の平均的な値を知る尺度である

#### 4.3.1 理論的期待値

離散確率空間  $(\Omega, F, P)$  において、標本空間  $\Omega$  が

$$\Omega = \{x_1, x_2, \dots, x_n\}$$

であり、この確率空間上の確率変数を  $X$  とする

このとき、確率変数  $X$  の期待値  $E$ (エクスペクテーション) は

$$E[x] = \sum_{i=1}^n x_i P(X = x_i)$$

で定義する

定義だから覚える

※  $P(X = x_1)$  を  $P_i$  と略すことがある

連続確率空間に対しては、確率変数  $X$  の期待値は、確率密度関数  $f(x)$  を用いて

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

で定義する

#### 4.3.2 例題

10 本のくじがある。A 賞が当たると 1000 円、B 賞は 500 円、C 賞は 100 円もらえるとする。

ただし、各賞は A 賞 2 本、B 賞は 3 本、C 賞は 5 本入っている。

くじを 1 回引いた時、いくら貰えると期待できるか。

とりあえず確率分布表を書こう

石雀率分布表			
確率変数 X	1000	500	100
確率 P	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{5}{10}$

標本空間  $\Omega$

$$\Omega = \{x_1 = 1000, x_2 = 500, x_3 = 100\}$$

確率測度  $\mathbf{P}$

$$P_1 = P(X = x_1) = \frac{2}{10}, P_2 = P(X = x_2) = \frac{3}{10}, P_3 = P(X = x_3) = \frac{5}{10}$$

期待値

$$E[x] = \sum_{i=1}^3 x_i P(X = x_i) = 1000 \times 2/10 + 500 \times 3/10 + 100 \times 5/10 = 400$$

1回当たり 400 円貰えると期待できる

#### 4.4 分散

\*\*\*\*\*もしかしたら  $x$  と  $X$  がごちゃごちゃになってる??

ある確率分布の値のばらつきは分散または標準偏差で測ることができる。

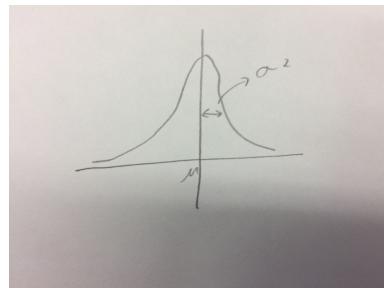
分散と標準偏差

確率変数  $X$  の期待値を  $\mu = E[X]$  とするこ  $X$  の分散  $\sigma^2$  は

$$\sigma^2 = V[X] = E[(x - \mu)^2]$$

で定義されるまた、 $\sigma$  を標準偏差と呼ぶ

$$\sigma = \sqrt{V[X]}$$



離散確率空間  $(\Omega, \mathcal{F}, \mathbf{P})$  において

$$\Omega = x_1, x_2, \dots, x_n$$

であり、この確率空間上の確率変数を  $X$  とする

このとき、確率変数  $X$  の分散は

$$V[X] = \sum_{i=1}^n (x_i - \mu)^2 P(X = x_i)$$

同様に連続確率空間では

$$V[X] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

で定義する

分散の公式

$$V[X] = E[X^2] - (E[X])^2$$

\*\*\*\*\*宿題この式を示せ

#### 4.5 計算公式

確率変数  $X, Y$  と、定数  $a, b, c$  において

E1 :  $E[c] = c$

E2 :  $E[aX + b] = aE[X] + b$

**E3** :  $E[X + Y] = E[X] + E[Y]$

**V1** :  $V[c] = 0$

**V2** :  $V[X + c] = V[X]$

**V3** :  $V[cX] = c^2V[X]$

**V4** :  $V[X \pm Y] = V[X] + V[Y] \rightarrow$ ただし、**X** と **Y** は互いに独立であるとする

分散にマイナスはない

独立 . . .

お互いの結果が、影響し合わないこと

$$P(X \cap Y) = P(X)P(Y)$$

#### 4.5.1 練習問題

1. サイコロを1回降った時、出る目の期待値  $\mu$  と分散  $\sigma^2$  を求めなさい

標本空間  $\Omega$

$$\Omega = \{x_1 = 1, x_2 = 2, \dots, x_6 = 6\}$$

確率測度

$$P_1 = P_2 = \dots = P_6 = 1/6$$

出る目の値を確率変数 **X** とおく

$$\mu = E[X] = \sum_{i=1}^6 x_i P_i = 1 * 1/6 + 2 * 1/6 + \dots + 6 * 1/6 = 3.5 = 7/2$$

分散  $\sigma^2$

$$\sigma^2 = V[X] = \sum_{i=1}^6 (x_i - \mu)^2 P_i$$

もしくは

$$E[X^2] = \sum_{i=1}^6 x_i^2 P_i = 91/6$$

これより

$$V[X] = E[X^2] - (E[X])^2 = 91/6 - (7/2)^2 = 35/12$$

## 5 11/21

### 5.1 高次の統計量

#### 5.1.1 m 次モーメント

期待値  $\mu = E[X]$  のまわりの m 次モーメント

$\mu_m, m = 1, 2, 3, \dots$  を

$$\mu_m = E[(X - \mu)^m] = \sum_{i=1}^m (x_i - \mu)^m P(X = x_i)$$

と定義する

分散  $\sigma^2$  は 2 次モーメントに等しい

$$\sigma^2 = \mu_2$$

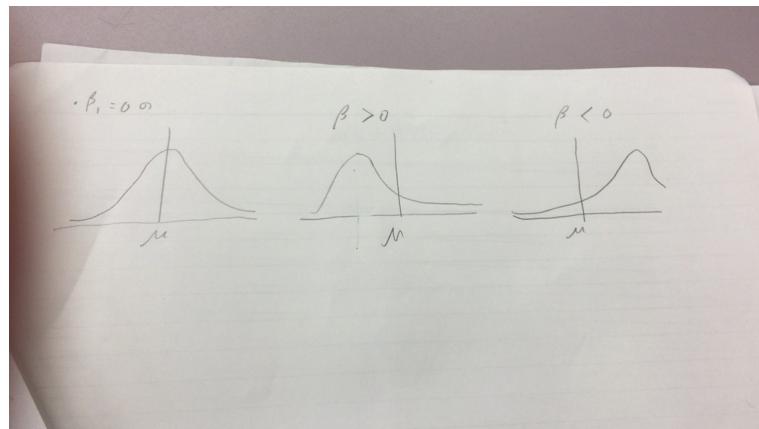
### 5.1.2 歪度

分布の歪み度をあらわす尺度

$$\beta_1 = \frac{\mu_3}{\sigma^3} = \frac{E[(X - \mu)^3]}{\sigma^3}$$

$\sigma$  : 標準偏差

- ・  $\beta_1 = 0$  のとき、期待値  $\mu$  を中心に左右対称になる
- ・  $\beta_1 > 0$  のとき、中心よりも左に歪んだ分布になる
- ・  $\beta_1 < 0$  のとき、中心よりも右に歪んだ分布になる



### 5.1.3 尖度

分布の尖り度を表す尺度である

$$\beta_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{E[(X - \mu)^4]}{\sigma^4} - 3$$

・  $\beta_2$  が大きい時

鋭いピークと太く長い裾を持つ

・  $\beta_2$  がちいさいとき  
丸みがかったピークで、短く細い裾を持つ

## 5.2 様々な確率分布

### 5.2.1 ベルヌーイ分布

・ ベルヌーイ試行 . . . .

1回の試行で、2つの結果 A または B が起こる実験をベルヌーイ試行と呼ぶ

例

・ コインを1枚1回投げる試行

・ ある病気である、または、ある病気でない

事象 A と B に対して、確率変数 X を用いてそれぞれ 1,0 で表す

1が出る確率を  $p$  で表す

ベルヌーイ分布

$$\begin{cases} P(X = 1) = p \\ P(X = 0) = 1 - p \end{cases} \quad (1)$$

(2)

確率変数 X はパラメータ  $p$  のベルヌーイ分布  $Be(p)$  に従うという

$$X \sim Be(p)$$

期待値 :  $E[X] = p$  分散 :  $V[X] = p(1 - p)$

### 5.2.2 二項分布

ベルヌーイ試行を  $n$  回行った時、結果  $A$  が怒る回数  $X$  は確率変数となる

二項分布  $B(n,p)$  : 離散分布

確率  $p$  で結果  $A$  が怒る試行において、 $n$  回行った時、ちょうど  $k$  回怒る確率は

$$P(X = k) = {}_n C_k p^k (1-p)^{n-k}$$

$$k = 0, 1, 2, 3, \dots$$

で与えられる

確率変数  $X$  はパラメータ  $n, p$  の二項分布  $B(n, p)$  に従うという

$$X \sim B(n, p)$$

### 5.2.3 例題

表が出る確率が  $p = \frac{3}{5}$  である曲がったコインを 5 回投げた時、表が 3 回出る確率はいくらか

$$P(X = 3) = {}_5 C_3 p^3 (1-p)^{5-3} = 10 * (3/5)^3 * (2/5)^2 = \frac{216}{625}$$

$k=0,1,2,3,4,5$  の場合について計算すると、確率分布表がかける

石窓確率変数	0	1	2	3	4	5
確率 $P(X=k)$	それを書いなして					

$B(n,p)$  の特徴

期待値  $E[X] = np$

分散  $V[X] = np(1-p)$

※二項定理

### 5.2.4 ポアソン分布 $P_o(\lambda)$ : 離散分布

単位時間あたりに入会起こる事象が  $k$  回起こる確率は

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

で与えられる。 $\lambda > 0$  とする

確率変数  $X$  はパラメータ  $\lambda$  のポアソン分布  $P_o(\lambda)$  に従うという

$$X \sim P_o(\lambda)$$

### 5.2.5 例題

1 時間あたりに  $\lambda = 3$  回メールが届く人がいる。→単位時間あたりだからポアソン  
この人が  $k=5$  回メールを受信する確率はいくらか

$$P(X = 5) = \frac{3^5}{5!} e^{-3} = 0.1008$$

ポアソン分布の特徴

期待値  $E[X] = \lambda$

分散  $V[X] = \lambda$

### 5.2.6 指数分布：連續分布

パラメータ  $\lambda > 0$  を持つ確率密度関数

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3)$$

(4)

で与えられる確率分布を、指数分布という

確率変数  $X$  はパラメータ  $\lambda$  の指数分布に従うという

$$X \sim Exp(\lambda)$$

### 5.2.7 例題

・地震が起こる間隔

地震が起きてから次に起こるまでの時間

・メールを受信する間隔

※ポアソン分布との違い

単位時間あたり平均  $\lambda$  回起こる現象に対して

・単位時間に事象が起こる回数  $\sim P_o(\lambda)$

・事象の発生間隔は平均  $\frac{1}{\lambda}$  の指数分布に従う

指数分布の特徴

期待値  $E[X] = \frac{1}{\lambda}$  ← 部分積分

分散  $V[X] = \frac{1}{\lambda^2}$

\*\*\*\*\*宿題

二項分布、ポアソン分布、指数分布の期待値と分散を導出しなさい

## 6 11/9

### 6.1 正規分布（連續分布）

平均  $\mu$ , 分散  $\sigma^2$  をもつ確率密度関数

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

で定められる確率分布を正規分布という

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

確率変数  $X$  は、平均  $\mu$ 、分散  $\sigma^2$  の正規分布に従うという

$$X \sim N(\mu, \sigma^2)$$

※ガウス分布ということもある

### 正規分布の特徴

期待値

$$E[X] = \mu$$

分散

$$V[X] = \sigma^2$$

### 標準正規分布 $N(0,1)$

平均  $\mu=0$ 、分散  $\sigma^2 = 1$  であるとき恭順正規分布といふ

$N(\mu, \sigma^2)$  に対して、

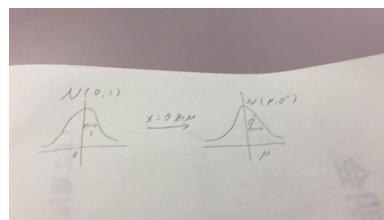
#### 標準化変数

$$z = \frac{x - \mu}{\sigma}$$

を導入する

$z$  は標準正規分布  $N(0,1)$  に従う

$$Z \sim N(0, 1)$$



#### 6.1.1 累積分布関数

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

標準正規分布の場合

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

#### 6.1.2 チェビシェフの不等式

期待値  $\mu$  からの標準偏差  $\sigma$  の単位で、 $k$  倍以内になる確率は

$$P(|X - \mu| \leq k\sigma) > 1 - \frac{1}{k^2}$$

である

ある確率分布

2  $\sigma$  以内  $1 - \frac{1}{2^2} = 75\%$  以内

3  $\sigma$  以内 88% 以内

より大きい

正規分布の場合

2  $\sigma$  以内 95% 以内

3  $\sigma$  以内 99.7% 以内

より大きい (正規分布のは覚えといたほうがいい??)

### 6.1.3 誤差関数と Q 関数

誤差関数  $erf(x)$

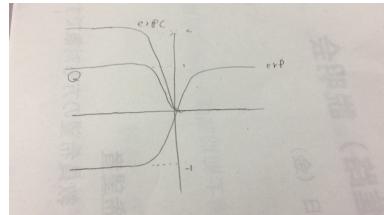
$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

相補誤差関数  $erfc(x)$

$$erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$$

**Q 関数**

$$Q(x) = \frac{1}{2} [1 - erf(\frac{x}{\sqrt{2}})]$$



- $erf(-x) = -erf(x) \rightarrow$  奇関数
- $\Phi(x) = \frac{1}{2}[1 + erf(\frac{x}{\sqrt{2}})] = \frac{1}{2}erfc(-\frac{x}{\sqrt{2}})$
- $erf(x) = 2\Phi(\sqrt{2}x) - 1$
- $erfc(x) = 2[1 - \Phi(\sqrt{2}x)]$

## 6.2 多次元確率変数

### 6.2.1 同時確率分布

2変数の場合

2つの離散型確率変数  $X, Y$  があるとする。 $X = x$  であり、かつ、 $Y = y$  である確率

$$P(X = x, Y = y)$$

を2次元確率変数  $(X, Y)$  の同時確率分布という

※大文字  $X, Y$  は確率変数。小文字  $x, y$  は、その実現値である

※  $P(X = x, Y = y) = P_{XY}(x, y) = P(x, y)$  と書くことがある

## 6.3 同時確率分布表

例・・・確率変数  $X, Y$  がどちらも 0,1 をとる場合

		$Y = 0$	$Y = 1$
$X = 0$	$P(X=0, Y=0)$	$P(X=0, Y=1)$	
	$P(X=1, Y=0)$	$P(X=1, Y=1)$	

ここで確率  $P$  を全部足すと 1 となる

**n** 変数の場合 **n** 個の確率変数  $X_1, X_2 \dots X_n$  があるとする  
各確率変数が

$$X_i = x_i$$

である

確率分布

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

を **n** 次元確率変数の同時確率分布という。

※多変数の場合

$$P(x_1, x_2 \dots x_n)$$

と書くことが多いここで  $x_i$  が取り得るすべての場合について和を求める 1 になる

**例題**

2 枚の大小の歪んだコインを同時に投げたとする。大きいコインの得点を **X**、小さいコインの得点を **Y** で表すとする。

大きいコインが表の時 5 点とし、裏のとき 1 点とする。

小さいコインが表の時 3 点とし、裏の時 0 点とする。

各コインの出方は次の同時確率分布の通りである。

		$Y=0$	$Y=3$	
$X=1$		$\frac{1}{3}$	$\frac{1}{4}$	
		$\frac{1}{4}$	$\frac{1}{4}$	
$X=5$				

この大小のコインの合計得点の期待値を求めなさい

$$E[X + Y] = \sum_x \sum_y (x + y) P(X = x, Y = y) \quad (5)$$

$$= (1 + 0)P(X = 1, Y = 0) + (1 + 3)P(X = 1, Y = 3) \quad (6)$$

$$+ (5 + 0)P(X = 5, Y = 0) + (5 + 3)P(X = 5, Y = 3) \quad (7)$$

$$= \frac{47}{12} \quad (8)$$

7 11/30

## 7.1 同時確率分布

		<u>例題</u>	
		0	3
X	1	$\frac{1}{3}$	$\frac{1}{4}$
	5	$\frac{1}{4}$	$\frac{1}{6}$
		$\frac{2}{12}$	$\frac{5}{12}$

## 7.2 周辺確率分布

同時確率分布が与えられた時、たとえば、 $P(X = 1)$  はいくらか？

$X=1$  となるのは、 $Y=0$  でも、 $Y=3$  でもいいから、2つの確立の和になる

$$P(X = 1) = P(X = 1, Y = 0) + P(X = 1, Y = 3) = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

\*\*\*( $P = 1$ ) を確率変数  $X$  の周辺確率という。

確率変数  $(X, Y)$  の同時確率分布において、確率変数  $X$  の周辺確率分布は

$$P(X = x) = \sum_y P(X = x, Y = y)$$

で求められる。n変数の場合は

$$P(X_1 = x_1) = \sum_{x_2} \sum_{x_3} \dots \sum_{x_n} P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

となる

## 7.3 条件付き確率

例題において、

$Y=3$  であることがわかったとき、 $X=5$  である確率はいくらか

この確率を

$$P(X = 5|Y = 3)$$

と表し、条件付き確率と呼ぶ

$Y=3$  は確定なので、同時確率分布表の  $Y=3$  の列のみを見ればいい

$$P(X = 1) = P(X = 1, Y = 0) + P(X = 1, Y = 3) = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

よって

$$P(X = 1|Y = 3) = \frac{3}{5} \tag{9}$$

$$P(X = 5|Y = 3) = \frac{2}{5} \tag{10}$$

(11)

確率変数  $(X, Y)$  において、ある事象  $Y$  が起こる条件のもとで、事象  $X$  が起こる条件付き確率は

$$P(X|Y) = \frac{P(X, Y)}{P(Y)}$$

で求めることができる。

以下の式から求めることができる。こっちの方が大事

同時確率  $P(X, Y)$  は

$$P(X, Y) = P(X|Y)P(Y) = \underbrace{P(Y|X)P(X)}_{\text{と書くこともできる。}}$$

ある事象  $Y$  が起きたとき  $X$  が起きた条件の下で、 $X$  が起る確率

ある事象  $X$  が起きたとき  $Y$  が起きた条件の下で、 $Y$  が起る確率

これを覚えとけば上の式は導き出せる

### 7.3.1 例題

まず、周囲確率  $P(Y=3)$  を求めよ

$$\begin{aligned} P(Y=3) &= \sum_x P(X=x, Y=3) \\ &= P(X=1, Y=3) + P(X=5, Y=3) \\ &= \frac{1}{4} + \frac{1}{6} = \frac{5}{12} \end{aligned}$$

これが  $P(Y=3)$ 。

$$\begin{aligned} P(X=1 | Y=3) &= \frac{P(X=1, Y=3)}{P(Y=3)} \\ &= \frac{\frac{1}{4}}{\frac{5}{12}} = \frac{3}{5} \end{aligned}$$

これとおんなじようにしたら  $P(X=5|Y=3) = \frac{2}{5}$  がもとまる

### 7.4 条件付き確率の期待値

$Y=0$  であるときの条件付き確率を求めておく。

$$\begin{aligned} P(Y=0) &= P(X=1, Y=0) + P(X=5, Y=0) \\ &= \frac{1}{3} + \frac{1}{4} = \frac{7}{12} \end{aligned}$$

$$P(X=1 | Y=0) = \frac{P(X=1, Y=0)}{P(Y=0)}$$

$$= \frac{\frac{1}{3}}{\frac{7}{12}} = \frac{4}{7}$$

$$P(X=5 | Y=0) = \frac{P(X=5, Y=0)}{P(Y=0)}$$

$$= \frac{\frac{1}{4}}{\frac{7}{12}} = \frac{3}{7}$$

したがり、 $Y=3$ であることを知ったとき  
得た  $X$  の期待値は

$$E[X | Y=3] = \sum x P(X=x | Y=3)$$

$$E[X | Y=3] = 1 \times \frac{3}{7} + 5 \times \frac{2}{7} = \frac{13}{7}$$

同様に  $Y=0$  であることを知ったときの  $X$  の期待値

$$E[X | Y=0] = 1 \times P(X=1 | Y=0) + 5 \times P(X=5 | Y=0)$$

$$= 1 \times \frac{4}{7} + 5 \times \frac{3}{7} = \frac{19}{7}$$

ここは絶対に試験に出る！！

条件付き確率の意味をしっかり理解しておくこと

## 7.5 共分散

2つの確率変数  $X, Y$  の間に関連があれば、一方の結果は他方に影響する。  
共分散

2つの確率変数  $X, Y$  の期待値を  
それぞれ  $\mu_x = E[X], \mu_y = E[Y]$  とす。  
ここで  $X$  と  $Y$  の共分散は、

$$\text{Cov}[X, Y] = E[(X - \mu_x)(Y - \mu_y)]$$

と定義する。

Cov はコバリアンスとよむ

離散はくつついでないから足し算。連続だとくつついでいわゆる面積になるから積分

離散確率変数について。  
 $\text{Cov}[X, Y] = \sum_x \sum_y (x - \mu_x)(y - \mu_y) P(X=x, Y=y)$

連続確率変数について。  
 $\text{Cov}[X, Y] = \iint_{\Omega} (x - \mu_x)(y - \mu_y) f(x, y) dx dy$   
 ただし、 $f(x, y)$  は、確率密度関数である。  
 $\Omega$  は標本空間とする。

$$\begin{aligned} \text{Q} \quad \text{Cov}[X, X] &= V[X] \\ &= E[(X - \mu)^2] \end{aligned}$$

## 7.6 計算公式

\*\*\*\*\*試験に式の証明の問題出るかも

○計算公式

- $\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$   
 $(V[X] = E[X^2] - (E[X])^2)$
- $V[X+Y] = V[X] + V[Y] + 2\text{Cov}[X, Y]$
- $\text{Cov}[X, Y]^2 \leq V[X]V[Y]$

## 7.7 分散共分散行列

多変数の場合の共分散は？？

確率変数  $X_1, X_2, \dots, X_n$  の  $i$  番目と  $j$  番目の共分散を

$$\sigma_{ij} = \text{Cov}[X_i, X_j]$$

とおく。これをまとめて行列で表したものを見ると分散共分散行列という。これを大文字のシグマ  $\Sigma$  で表す

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \dots & \sigma_{2n} \\ \vdots & & & & \\ \sigma_{n1} & \sigma_{n2} & \dots & & \sigma_{nn} \end{bmatrix}$$

ただし、対角成分。  
 $\sigma_{ii} = V[X_i] = \sigma_i^2$  となる。

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & & & \\ \sigma_{n1} & \dots & & \sigma_n^2 \end{bmatrix}$$

対角成分  
分散

分散

分散の式に等しい

## 7.8 相関係数

共分散を規格化した量

相関係数  $\rho$

$$\rho = \frac{\text{Cov}[X, Y]}{\sqrt{V[X]}\sqrt{V[Y]}}$$

と定義する。

- $-1 \leq \rho \leq 1$  になる。

こう定義すると、 $-1 < \rho < 1$  となり、正規化される。

$\rho = 0$  のとき、 $X$  と  $Y$  は無相関 (関係がない) である