# 確率・統計

## 1 順列・組合せ

### §1 順列

1.1 n個の異なるものから重複を許さずにr個をとって，これを1列に並べたものを，n個のものの r 順列といい，その数は $_nP_r = \frac{n!}{(n-r)!} = n(n-1)\cdots(n-r+1)$

1.2 $_nP_n = n(n-1)\cdots 2\cdot 1 = n!$，$_nP_0 = 1$，$_0P_0 = 0! = 1$

1.3 n個の異なるものから重複を許してr個をとってできる順列の数は $n^r$

1.4 n個のもののうちaが $n_1$ 個，bが $n_2$ 個，cが $n_3$ 個…であるとき，このn個を1列に並べる順列の数は，$\frac{n!}{n_1!n_2!n_3!\cdots}$

### §2 組合せ

2.1 n個の異なるものから重複を許さずにr個をとった組を，n個のものの r 組合せといい，その数は $_nC_r = \binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{(n-r)!\,r!}$

2.2 $_nC_n = 1$　$_nC_0 = 1$

2.3 n個の異なるものから重複を許してr個をとってできる組合せの数は $_{n+r-1}C_r = \frac{n(n+1)\cdots(n+r-1)}{r!}$

### §3 2項定理と多項定理

3.1 a,bを任意の数，nを正整数とするとき $(a+b)^n = \sum_{r=0}^{n} {_nC_r}\, a^{n-r} b^r = {_nC_0}a^n + {_nC_1}a^{n-1}b$

3.2 $a=1,\ b=x$ とおけば $(1+x)^n = \sum_{r=0}^{n} {_nC_r} x^r = {_nC_0} + {_nC_1}x + \cdots {_nC_n}x^n \cdots + {_nC_n} b^n$

## 2 確率

### §1 事象

$\Omega$：全事象(標本空間)，$\phi$：空事象，A：事象，$A^c$：余事象　$A_1 \cup A_2$：和事象　$A_1 \cap A_2$：積事象

$A_1 \cap A_2 = \phi$：排反事象　　$A \subset B$：AはBの部分集合

### §2 確率の基本定理

(i) $0 \le P(A) \le 1$　(ii) $P(\Omega)=1$，$P(\phi)=0$　(iii) $P(A^c) = 1-P(A)$　(iv) $P(A_1 \cup A_2)$ $= P(A_1)+P(A_2)-P(A_1\cap A_2)$

(v) $A_1, A_2$ が排反するとき　$P(A_1 \cup A_2) = P(A_1)+P(A_2)$

(vi) $A_1, A_2 \cdots A_k$ が排反するとき $P(A_1\cup A_2\cup\cdots\cup A_k) = P(A_1)+P(A_2)+\cdots+P(A_k)$（加法定理）

### §3 古典概型

例：设有N件产品，其中有D件次品，今从中任取n件，问其中恰有k件（k<D）件次品的概率？

$P = \binom{D}{k}\binom{N-D}{n-k} \Big/ \binom{N}{n}$（超幾何分布確率公式）

### §4 条件付き確率と独立性

4.1 $P(A_1) > 0$ のとき，$P(A_2|A_1) = \frac{P(A_1\cap A_2)}{P(A_1)}$ これを，条件 $A_1$ のもとで $A_2$ の起こる条件付き確率という（定義）

4.2 $A_1$ と $A_2$ が独立，$P(A_1)>0$ のとき，$P(A_2|A_1) = P(A_2)$

4.3 $P(A_1)>0$ のとき，$P(A_1\cap A_2) = P(A_1)P(A_2|A_1) = P(A_2)P(A_1|A_2)$（乗法公式）

4.4 $A_1$ と $A_2$ が独立のとき，$P(A_1\cap A_2) = P(A_1)P(A_2)$

4.5 $P(A_1\cap\cdots\cap A_{k-1}) > 0$ のとき，$P(A_1\cap A_2\cdots A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1\cap A_2)\cdots P(A_k|A_1\cap A_2\cap\cdots\cap A_{k-1})$

4.6 $A_1,\cdots A_k$ が独立のとき，$P(A_1\cap A_2\cap\cdots\cap A_k) = P(A_1)\cdots P(A_k)$

4.7 全確率公式 $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)$

Bayes 公式 $P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}$

### §5 確率変数

確率変数X，任意の実数 $x\ (-\infty<x<+\infty)$ に対して，$F(x) = P(X\le x)$ をXの分布関数という。

5.1 離散的な1次元確率分布

Xのとる値が有限または可付番個の値 $x_1, x_2, \cdots$ に限られていて，それぞれのとる確率 $P(X=x_i) = P_i \ge 0$ が定まり，$\sum_i P_i = 1$ のとき，$F(x) = \sum_{x_i\le x} P_i$ を満足する $P_i$ を離散確率変数Xの確率分布という。

(I) (0−1)分布

| X | 0 | 1 |
|---|---|---|
| $P_k$ | $1-p$ | $p$ |

$P\{X=k\} = p^k(1-p)^{1-k}$，$k=0,1$ $(0<p<1)$

(II) n重 Bernoulli実験、2項分布 $B(n,p)$

確率分布：$P_k = \binom{n}{k} p^k q^{n-k}$ $(p,q>0,\ p+q=1,\ k=0,1,\cdots n)$

分布関数 $= \sum_{k\le x}\binom{n}{k} p^k q^{n-k}$

期待値 $= np$　　分散(方差) $= npq$

(III) ポアッソン分布 $P(\lambda)$

確率分布：$P_k = e^{-\lambda}\frac{\lambda^k}{k!}$ $(\lambda>0,\ k=0,1,2\ldots)$

分布関数 $= \sum_{k\le x} e^{-\lambda}\frac{\lambda^k}{k!}$

期待値 $= \lambda$　　分散 $= \lambda$

5.2 連続的な1次元確率分布

$f(x)\ge 0$，$\int_{-\infty}^{\infty} f(x)dx = 1$ のとき，$P\{a<x\le b\} = \int_a^b f(x)dx$ を満足する $f(x)$ を連続確率変数Xの確率密度関数または密度関数といい，

分布関数 $F(x) = \int_{-\infty}^{x} f(t)dt$

(I) 一様分布（矩形分布）$U(a,b)$

確率密度関数：$f(x) = \begin{cases} 1/(b-a) & (a\le x\le b) \\ 0 & (\text{その他}) \end{cases}$

分布関数 $= \begin{cases} 0 & (-\infty<x<a) \\ (x-a)/(b-a) & (a\le x\le b) \\ 1 & (b<x<+\infty) \end{cases}$

期待値 $= (a+b)/2$　　分散 $= (b-a)^2/12$

(II) 指数分布 $Exp(\lambda)$

確率密度関数 $f(x) = \begin{cases} \lambda e^{-\lambda x} & (0<x) \\ 0 & (x\le x) \end{cases}$

分布関数 $= \begin{cases} 1-e^{-\lambda x} & (0<x) \\ 0 & (x\le 0) \end{cases}$

期待値 $= 1/\lambda$　　分散 $= 1/\lambda^2$

(III) 正規分布 $N(\mu, \sigma^2)$

確率密度関数 $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $(-\infty<\mu<\infty,\ \sigma>0,\ -\infty<x<\infty)$

分布関数 $= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{x} e^{-(t-\mu)^2/2\sigma^2}dt$

期待値 $= \mu$　　分散 $= \sigma^2$

$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$　$N(0,1)$　$\Phi(-x) = 1-\Phi(x)$

5.3 確率変数の関数の分布

如何由已知随机变量X的概率分布求它的函数 $Y=g(x)$ 的概率分布（$g(\cdot)$ 为已知连续函数）.

**定理** 设随机变量X具有概率密度 $f_X(x)$，$x\in(-\infty,\infty)$，又设函数 $g(x)$ 处处可导且恒有 $g'(x)>0$（或 $g'(x)<0$）

则 $Y=g(x)$ 是连续型随机变量，其概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta \\ 0 & \text{其他} \end{cases}$$

$\alpha = \min\{g(-\infty), g(\infty)\}$　$\beta = \max\{g(-\infty), g(\infty)\}$　$h(y)$ 是 $g(x)$ 的反函数

5.4 2次元(多次元)確率分布

結合/同時分布関数：$F(x_1,x_2) = P(-\infty<X_1\le x_1,\ -\infty<X_2\le x_2)$

結合確率密度関数：

離散的な場合：$P(X_1=x_{1i}, X_2=x_{2j}) = P_{ij}\ge 0$，$\sum P_{ij}=1$　$F(x_1,x_2) = \sum_{x_{1i}\le x_1}\sum_{x_{2j}\le x_2} P_{ij}$

連続的な場合：$f(x_1,x_2)\ge 0$，$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1,x_2)dx_1dx_2 = 1$

$F(x_1,x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(t_1,t_2)dt_1dt_2$

5.5 周辺分布

Xの周辺分布関数：$F_X(x) = F(x,\infty) = P(X\le x, Y<\infty)$

Yの周辺分布関数：$F_Y(y) = F(\infty,y) = P(X<\infty, Y\le y)$

離散的な場合：$F_1(x_1) = \sum_{x_{1i}\le x_1}\sum_j P_{ij}$　$F_2(x_2) = \sum_i\sum_{x_{2j}\le x_2} P_{ij}$

連続的な場合：$F_1(x_1) = \int_{-\infty}^{x_1}\int_{-\infty}^{\infty} f(t_1,t_2)dt_2dt_1$

$F_2(x_2) = \int_{-\infty}^{x_2}\int_{-\infty}^{\infty} f(t_1,t_2)dt_1dt_2$

周辺密度関数

離散的な場合：$P(X_1=x_{1i}) = P_{i\cdot} = \sum_j P_{ij}$　$P(X_2=x_{2j}) = P_{\cdot j} = \sum_i P_{ij}$

連続的な場合：$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$　$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$

5.6 条件付き分布

$Y=y$ のもとでのXの条件付き確率密度関数

確率密度関数

離散的な場合：$P_{X|Y}(x|y) = P(X=x|Y=y) = \frac{P(x,y)}{P_Y(y)}$ $(P_Y(y)>0)$

連続的な場合：$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$ $(f_Y(y)>0)$

$Y=y$ のもとでのXの条件付き分布関数

離散的な場合：$F_{X|Y}(x|y) = P(X\le x|Y=y) = \sum_{a\le x} P_{X|Y}(a|y)$

連続的な場合：$F_{X|Y}(a|y) = P(X\le a|Y=y) = \int_{-\infty}^{a} f_{X|Y}(x|y)dx$

5.7 独立性

$P\{X\le x, Y\le y\} = P\{X\le x\}P\{Y\le y\}$

$F(x,y) = F_X(x)F_Y(y)$　　X，Yは独立であるという

5.8 たたみ込み（合成積）

(I) $Z = X+Y$

巻积公式：$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy$

$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_Y(z-x)f_X(x)dx$

(II) $Z = XY$、$Z = \frac{Y}{X}$

$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f(x,xz)dx$

$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x,\frac{z}{x})dx$

X,Yが独立すれば $f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f_X(x) f_Y(xz)dx$

$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_X(x) f_Y(\frac{z}{x})dx$

(III) $M = \max\{X,Y\}$ 及び $N = \min\{X,Y\}$

$F_{max}(z) = F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z)$

$F_{min}(z) = 1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)]$