情報数学 追加5
確率論

第1章

離散確率変数

離散確率変数

確率p_iで値$x_i (i = 0, 1, 2 \ldots n)$を取る変数$X$

$$\sum_{i=1}^{n} p_i = p_1 + p_2 + \cdots + p_n = 1 \quad (0 \le p_i \le 1)$$

例1.　確率変数Xの確率分布が次のようになる

X	1	2	3	4
P	0.1	0.3	0.1	0.5

離散確率変数

例1.　確率変数Xの確率分布が次のようになる

X	1	2	3	4
P	0.1	0.3	0.1	0.5

期待値あるいは平均　$E[X] = \sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n$

分散　$Var[X] = E[(X - E[X])^2] = \boxed{E[X^2] - (E[X])^2}$　$E[X^2 - 2E[X] X + E[X]^2]$

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

標準偏差　$\sigma[X] = \sqrt{Var[X]}$

例1の $E[X] = 3$　　$Var[X] = 1.2$

$E[X^2] = 0.1 + 1.2 + 0.9 + 8 = 10.2$

$(E(X))^2 = 9$

離散確率変数

例1.　確率変数Xの確率分布が次のようになる

X	1	2	3	4
P	0.1	0.3	0.1	0.5

積率母関数　$M_X(t) = E[e^{tX}] = \sum_{i=1}^{n} e^{tx_i} p_i = e^{tx_1}p_1 + e^{tx_2}p_2 + \cdots + e^{tx_n}p_n$

性質　$M_X(t)$をXの積率母関数とする。このとき、正の整数nに対して、

$$[M_X(t)]_{t=0}^{(n)} = E[X^n]$$

例1の $M_X(t) = E[e^{tX}] = E[e^t \cdot 0.1 + e^{2t} \cdot 0.3 + e^{3t} \cdot 0.1 + e^{4t} \cdot 0.5]$

第2章

2つの離散確率変数

結合確率分布と周辺確率分布

例2.　確率変数XとYの**結合確率分布**が次のようになる

X\Y	0	1
0	0.6	0.2
1	0.1	0.1

Xの**周辺確率分布**が次のようになる

X	0	1
P	0.8	0.2

$E[X] = 0.2$
$Var[X] = 0.16$

$\sigma^2 = E[X^2] - (E[X])^2 = 0.2 - 0.04$

Yの**周辺確率分布**が次のようになる

Y	0	1
P	0.7	0.3

$E[Y] = 0.3$
$Var[Y] = 0.21$

$0.3 - 0.09$

2つの離散確率変数の共分散

例2.　確率変数XとYの**結合確率分布**が次のようになる

X\Y	0	1
0	0.6	0.2
1	0.1	0.1

XとYの**共分散**: 协方差

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = \boxed{E[XY] - E[X]E[Y]}$$

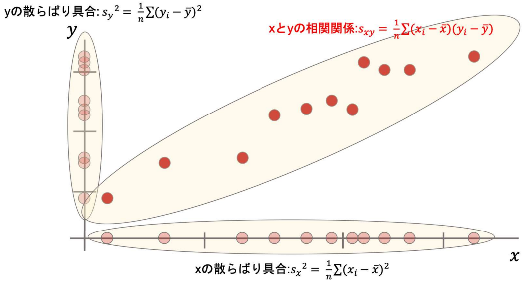
例2の$Cov(X,Y) = [0 \cdot 0 \cdot 0.6 + 0 \cdot 1 \cdot 0.2 + 1 \cdot 0 \cdot 0.1 + 1 \cdot 1 \cdot 0.1] - E[X]E[Y]$

$0.2 \cdot 0.3$　　　$0.1 - 0.06 = 0.04$

2つの離散確率変数の共分散

XとYの<u>共分散:</u>

$$Cov(X,Y) = E[(X-E[X])(Y-E[Y])] = E[XY] - E[X]E[Y]$$



共分散仅评价正负相关性（正负有意义）

<u>不适于评价相关的强弱</u>

2つの離散確率変数の相関係数

XとYの<u>相関係数:</u>

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

例2の$\rho(X,Y) \approx 0.22$

$-1 \le \rho(X,Y) \le 1$，$\rho(X,Y)$ 代表X与Y的相关程度

2つの離散確率変数性質

性質

$$E[aX+bY+c] = aE[X] + bE[Y] + c$$
$$Var[aX+bY+c] = a^2Var[X] + b^2Var[Y] + 2abCov(X,Y)$$

XとYが独立 → $E[XY]=0$, $E[X]\cdot E[Y]=0$

$\leftrightarrow$すべてのi,jについて、<u>$P(X=i,Y=j) = P(X=i)P(Y=j)$</u>

例2のXとYが独立であるか？

考试技巧：考试若出题，则一定不独立

$E(X) = \frac{3}{8} + 2\cdot\frac{3}{8} + 3\cdot\frac{1}{8} = \frac{3}{2}$

$E(Y) = \frac{2}{8} + \frac{8}{8} + \frac{6}{8} = 2$

$Var[X] = \frac{3}{8} + \frac{12}{8} + \frac{9}{8} - \frac{6}{4} = \frac{12}{8} = \frac{3}{2}$

$Var[Y] = \frac{2}{8} + \frac{16}{8} + \frac{18}{8} - \frac{32}{8} = \frac{4}{8} = \frac{1}{2}$

$Cov[X,Y] = E[XY] - E[X]\cdot E[Y]$

2つの離散確率変数の性質

注意事项

XとYが独立 $\Rightarrow Cov(X,Y) = 0$

$Cov(X,Y) = 0 \nRightarrow X$と$Y$が独立

考试技巧：考试若出题，则一定不独立

$1-\frac{1}{20}$

問題1

X\Y	1	2	3
0	1/8	0	0
1	0	2/8	1/8
2	0	2/8	1/8
3	1/8	0	0

1. XとYの期待値と分散をそれぞれ求めよ。
2. XとYの相関係数を求めよ。
3. XとYは独立かどうか、理由とともに答えよ。

問題2

離散確率変数$X_1, X_2, \dots, X_n$は独立な確率分布

$$\Pr[X_i = k] = (\frac{1}{2})^k \quad (k = 1, 2, 3, \dots)$$

に従うものとする。

1. $\Pr[X_i \le k]$を求めよ。
2. X_iの期待値$E[X_i]$を求めよ。
3. $\sum_{i=1}^{n} X_i$の期待値$E[\sum_{i=1}^{n} X_i]$を求めよ。
4. $Y = \max X_i$とするとき、$\Pr[Y \le k]$を求めよ。
5. $\lim_{n\to\infty} \Pr[Y \le \log_2 n - 1]$を求めよ。

第3章

条件つき確率

条件つき確率

公式 事象Sのもとで、事象Tの条件つき確率

$$P(T|S) = \frac{P(S \cap T)}{P(S)}$$

$P(S \cap T)$代表事象S与事象T同时发生的概率

例3. サイコロを2回投げる。出る目の和が3で、1回目が1である条件付き確率を求めよ。

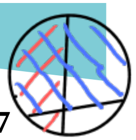
S:出る目の和が3で、$P(S) = \frac{2}{36} = \frac{1}{18}$ T:1回目が1である

$S \cap T$:1回目が1でかつ出る目の和が3である$P(S \cap T) = \frac{1}{36}$

$$P(T|S) = \frac{P(S \cap T)}{P(S)} = \frac{1}{2}$$

阪大，与东工一样　　7月底出成绩

条件つき確率



例4.　がんになる確率$P_C = 0.003$, がんにならない確率$P_{NC} = 0.997$ある検診で、かんである人を陽性と判定する確率は0.85、がんでない人を陽性と誤判定する確率は0.15となっている。
ある患者がこの検診で陽性と判定されたとき、この患者が実際にがんである確率を求めよ。

S　　　　T

S:~~検診の陽性判定~~、$P(S) = 0.003 \times 0.85 + 0.997 \times 0.15 = 0.1521$
T:この患者ががんである
$S \cap T$:この患者がんでかつ陽性と判定され

$$P(S \cap T) = 0.003 \times 0.85 = 0.00255$$

結論:

$$P(T|S) = \frac{P(S \cap T)}{P(S)} \approx 0.0168$$

問題3　同じ形の2つの箱AとBがある。箱Aに赤球2個、青球3個があり、箱Bに赤球4個、青球2個がある。無作為に１つの箱から1個の球を取り出したら、球の色が青であった。このとき、その球が箱Bの球である確率を求めよ。

S　　T

$$P(T|S) = \frac{P(S\cap T)}{P(S)} = \frac{\frac{2}{6}\ \frac{1}{3}}{\frac{1}{2}\cdot\frac{3}{5}+\frac{1}{2}\cdot\cancel{\frac{2}{6}}\frac{1}{3}} = \frac{5}{9+5} = \frac{5}{14}$$

問題4　ある製品を生産する企業では、3つの機械設備A,B,Cが設置されている。各機械設備はそれぞれ全製品の20% 50% 30%を生産している。各機械設備の不良品の発生する確率はそれぞれ1% 4% 4%である。
ある製品に対して検査を行い、不良品であることがわかった。このとき、この不良品が機械設備Bの製品である確率を求めよ。

S　T

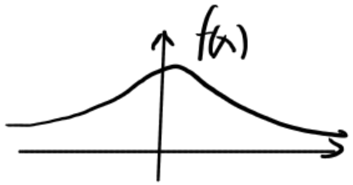
$$P(T|S) = \frac{P(S\cap T)}{P(S)} = \frac{50\% \times 4\%}{20\% \times 1\% + 50\% \times 4\% + 30\% \times 4\%}$$

$$= \frac{20}{2+20+12} = \frac{10}{17}$$

第4章

連続確率変数

連続確率変数と確率密度関数



確率変数Xのとる値は$\mathbb{R}$(実数)である
Xの確率密度関数　$f(x)$　　(または$f_X(x)$)

1. $f(x)$代表$X=x$的可能性； $f(x)\geq 0$
2. $P(c\leq X\leq d)=\int_c^d f(x)dx$
3. $\int_{-\infty}^{\infty} f(x)dx=1$

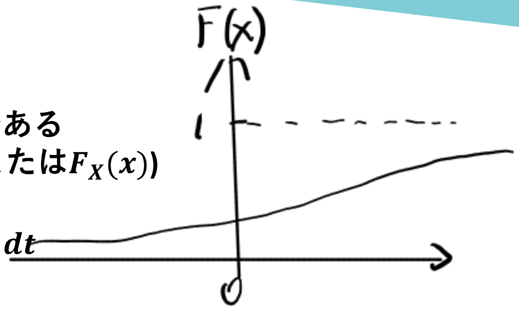
連続確率変数と確率密度関数

確率変数Xのとる値は$\mathbb{R}$(実数)である
Xの確率密度関数　$f(x)$　　(または$f_X(x)$)

1. $f(x)$代表$X=x$的可能性； $f(x)\geq 0$
2. $P(c\leq X\leq d)=\int_c^d f(x)dx$
3. $\int_{-\infty}^{\infty} f(x)dx=1$

Xの期待値: $E[X]=\int_{-\infty}^{\infty} xf(x)dx$
Xの分散: $Var[X]=E[X^2]-\{E[X]\}^2=\int_{-\infty}^{\infty} x^2 f(x)dx-\{\int_{-\infty}^{\infty} xf(x)dx\}^2$
Xの積率母関数: $M_X(t)=E[e^{tX}]=\int_{-\infty}^{\infty} e^{tx}f(x)dx$

確率分布関数

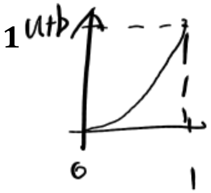


確率変数Xのとる値は$\mathbb{R}$(実数)である
Xの確率分布関数　$F(x)$　　(または$F_X(x)$)

1. $F(x)=P(X\leq x)=\int_{-\infty}^{x} f(t)dt$
2. $F'(x)=f(x)$
3. $F(x)$は単調増加関数　$x\to\infty$

問題5　確率変数Xの確率密度関数$f(x)$が以下のように与えられている。また、Xの期待値$E[X]=\frac{1}{2}$である。

$$f(x)=\begin{cases} ax^2+bx & 0\leq x\leq 1 \\ 0 & その他 \end{cases}$$



1. a,bの値を求めよ。
2. Xの分散$Var[X]$を求めよ。
3. Xの確率分布関数$F(x)$を求めよ。

① $E[X]=\int_0^1 x\cdot(ax^2+bx)dx=\frac{1}{2}$
$\frac{a}{4}x^4\Big|_0^1+\frac{b}{3}x^3\Big|_0^1=\frac{1}{2}$ $\Big\}$ $\frac{a}{4}+\frac{b}{3}=\frac{1}{2}$

$\int_0^1(ax^2+bx)dx=1$
$\left[\frac{a}{3}x^3+\frac{b}{2}x^2\right]_0^1=1$ $\longrightarrow$ $\Big\}$ $\frac{a}{3}+\frac{b}{2}=1$

$\begin{bmatrix}\frac{1}{4} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\\ 1\end{bmatrix}$ → 高斯消元

$\left.\begin{array}{l}3a+4b=6\\ 2a+3b=6\end{array}\right\}\begin{array}{l}a=-6\\ b=6\end{array}$

② $Var[X]=\int_0^1 x^2(ax^2+bx)dx-\left[\int_0^1 x(ax^2+bx)dx\right]^2$
$=6\left[-\frac{1}{5}x^5+\frac{1}{4}x^4\right]_0^1-36\left[-\frac{1}{4}x^4+\frac{1}{3}x^3\right]_0^{1\,2}$

③ $F(x)=\int_{-\infty}^{\infty}-6x^2+6x\,dx$
$=6\left[-\frac{1}{3}x^3+\frac{1}{2}x^2\right]$
$=-2x^3+3x^2$

$= \frac{6}{20} - \left(\frac{1}{2}\right)^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}$

第5章

2つの連続確率変数

2つの連続確率変数と同時確率密度関数

確率変数XとYのとる値はともに$\mathbb{R}$(実数)である
XとYの同時確率密度関数　$f(x,y)$

1. $f(x,y)$代表$X=x,Y=y$的可能性；　$f(x,y)\geq 0$
2. $P(X,Y\in D)=\iint_D f(x,y)dxdy$
3. $\iint_{R^2} f(x,y)dxdy=1$

周辺確率密度関数と共分散

Xの周辺確率密度関数　$f(x)=\int_{-\infty}^{\infty} f(x,y)dy$
Yの周辺確率密度関数　$f(y)=\int_{-\infty}^{\infty} f(x,y)dx$

XとYの共分散

$$Cov(X,Y)=E[XY]-E[X]E[Y]$$
$$=\iint_{R^2} xyf(x,y)dxdy-\int_{-\infty}^{\infty} xf(x)dx\int_{-\infty}^{\infty} yf(y)dy$$

$\boxed{X\text{と}Y\text{が独立}}$ $\leftrightarrow$ 　$f(x,y)=f(x)f(y)$

問題6　確率変数XとYの同時確率密度関数が次のようになる

$$f(x,y)=\begin{cases}\alpha(x+y^3) & (0\leq x\leq 1; 0\leq y\leq 1)\\ 0 & \text{その他}\end{cases}$$

1. αを求めよ。
2. Xの周辺確率分布関数$F(x)$を求めよ。
3. $P(X>Y)$を求めよ。

① $\iint a(x+y^3)dxdx=1$

$a\int_0^1\left[\frac{1}{2}x^2+xy^3\right]_0^1 dy=1$

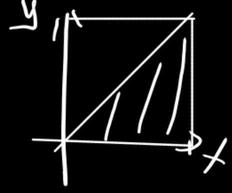
$\frac{1}{2}+y^3$　　$\frac{1}{2}+\frac{1}{4}=\frac{1}{a}$

$a\left[\frac{y}{2}+\frac{y^4}{4}\Big|_0^1\right]=1$　　$\frac{4}{3}$

② $F(x,y)=\frac{4}{3}\left[\frac{1}{2}x^2y+\frac{1}{4}xy^4\right]$ $(x,y\in[0,1])$

$=\frac{2}{3}x^2y+\frac{1}{3}xy^4$

③ $P(X>Y)=\frac{23}{45}$



$$P(X>Y) = \iint \frac{4}{3}(x+y^3)\,dx\,dy$$

$$= \frac{4}{3}\int_0^y \quad dy \quad / \quad \int_x^1 \quad dx$$