概率论与数理统计第二次习题课题目

题 1 求区间 $[a,b]$ 上取值的随机变量可能达到的最大方差？何时达到？

题 2 设 X 为一连续型随机变量，求实数 c，使得 $E|X-c|$ 达到最小.

题 3 设随机变量 X,Y 的联合概率密度函数为

$$f_{X,Y}(x,y)=\frac{1+xy(x^2-y^2)}{4},\qquad |x|\le 1,|y|\le 1.$$

求 $X+Y$ 的概率分布函数 F_{X+Y}。

题 4 设随机变量 X，Y 独立，都服从标准正态分布 $N(0,1)$，(X,Y) 的联合密度函数记为 $f(x,y)$；

1. 证明；函数 $g(x,y)=\begin{cases} f(x,y)+\dfrac{xy}{100}, & when\ x^2+y^2\le 1 \\ f(x,y), & when\ x^2+y^2>1 \end{cases}$，是二维概率密度函数。
2. 若随机变量 (U,V) 的密度函数为 $g(x,y)$，证明：U，V 都服从标准正态分布 $N(0,1)$，但 (U,V) 不服从二维正态分布。

注：本例说明各分量为正态推不出联合分布为正态。

题 5 投掷一枚公平的硬币，记正面为 H、反面为 T，

1. 直至首次出现HH时停止，请计算投掷次数的期望与方差；
2. 如果停止准则变为首次出现HT，此时投掷次数的期望和方差分别是多少；
3. 假设甲、乙进行一场比赛，投掷一个硬币直至首次出现HH或TH停止。如果以HH结束，则甲胜，TH结束为乙胜，请问甲、乙的获胜概率; 若改为HH先出现甲胜，HT先出现乙胜，结果怎样。

题 6 设随机变量$X\sim N(0,1)$, 随机变量W服从分布律$\begin{pmatrix} W & 1 & -1 \\ P & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$，$X$与$W$相互独立。令$Y=XW$。
证明: (1). $Y\sim N(0,1)$, (2). $Cov(X,Y)=0$, (3). X与Y不独立。
（这个例子表明两个正态分布随机变量相互不相关并不能推出这两个随机变量独立。只有当两个随机变量是二元正态分布时，不相关才与独立等价。）

题 7 设随机变量$X\sim N(0,1)$, 随机变量

$$Y=\begin{cases} X, & 若|X|\ge c \\ -X, & 若|X|<c \end{cases}.$$

(1). 求常数c使得X与Y不相关，
(2). 证明: $Y\sim N(0,1)$,
(3). 说明对任意$c>0$, 如上定义的随机变量X与Y不独立。

概率论与数理统计第二次习题课题目解答

题 1 求区间 $[a,b]$ 上取值的随机变量可能达到的最大方差？何时达到？

解：$E((X-E(X))^2 \le E((X-\frac{a+b}{2})^2) \le E((b-\frac{a+b}{2})^2) = \frac{(b-a)^2}{4}$

第一个不等号等式成立的条件是 $E(X)=\frac{a+b}{2}$，

第二个不等号成立的条件是 $P(X=a)+P(X=b)=1$,

所以达到最大方差 $\frac{(b-a)^2}{4}$ 的条件是 $P(X=a)=P(X=b)=\frac{1}{2}$。

题 2 设 X 为一连续型随机变量，求实数 c，使得 $E|X-c|$ 达到最小.

解：X 有概率密度函数 $f(x)$，并且 $f(x)$ 处处连续。

$$h(c) := E|X-c| = \int_{-\infty}^{+\infty} |x-c| f(x)dx = \int_c^{+\infty} (x-c)f(x)dx + \int_{-\infty}^{c} (c-x)f(x)dx.$$

于是 h 关于 c 可微，

$$\begin{aligned} h'(c) &= \frac{d}{dc}\left(\int_c^{+\infty}(x-c)f(x)dx\right) + \frac{d}{dc}\left(\int_{-\infty}^{c}(c-x)f(x)dx\right) \\ &= -(c-c)f(c) - \int_c^{+\infty} f(x)dx + (c-c)f(c) + \int_{-\infty}^{c} f(x)dx \\ &= F(c) - \big[1-F(c)\big] = 2F(c)-1. \end{aligned}$$

所以，在区间 $I_1 := \{c : F(c) < 1/2\}$ 上，$h'(c)<0$，h 严格减，在区间 $I_3 := \{c : F(c) > 1/2\}$ 上，$h'(c)>0$，h 严格增，在区间

$$I_2 := \{c : F(c) = 1/2\}$$

上，h 为常数，所以区间 I_2 中的点都是 h 的最小值点。当区间 I_2 是单点集时，这个唯一的 c 值恰是 X 的中位数。

题 3 设随机变量 X,Y 的联合概率密度函数为

$$f_{X,Y}(x,y) = \frac{1+xy(x^2-y^2)}{4}, \qquad |x|\le 1, |y|\le 1.$$

求 $X+Y$ 的概率分布函数 F_{X+Y}。

<u>解法 1</u>**:**

$$
\begin{aligned}
f_{X+Y}(u) &= \int_{-\infty}^{+\infty} f_{X,Y}(u-v,v)dv \\
&= \int_{-\infty}^{+\infty} \frac{1+(u-v)v[(u-v)^2-v^2]}{4} I_{|u-v|\le 1,|v|\le 1} dv \\
&= \int_{-\infty}^{+\infty} \frac{1+(u-v)v[u^2-2uv]}{4} I_{u-1\le v\le 1+u,-1\le v\le 1} dv \\
&= I_{\max\{u-1,-1\}\le\max\{u+1,1\}} \int_{\max\{u-1,-1\}}^{\min\{u+1,1\}} \frac{1+(u-v)v[u^2-2uv]}{4} dv \\
&= I_{-2\le u\le 2} \int_{\frac{u}{2}-1+\left|\frac{u}{2}\right|}^{\frac{u}{2}+1-\left|\frac{u}{2}\right|} \frac{1+(u-v)v[u^2-2uv]}{4} dv \\
&= I_{-2\le u\le 2} \int_{-1+\left|\frac{u}{2}\right|}^{1-\left|\frac{u}{2}\right|} \frac{1-2u\left[\frac{u^2}{4}-w^2\right]w}{4} dw \qquad \left(\text{注意积分的对称性，令} w=v-\frac{u}{2}\right) \\
&= I_{-2\le u\le 2} \int_{-1+\left|\frac{u}{2}\right|}^{1-\left|\frac{u}{2}\right|} \frac{1}{4} dw \\
&= I_{-2\le u\le 2} \frac{2-|u|}{4}.
\end{aligned}
$$

于是

$$
\begin{aligned}
F_{X+Y}(z) = \int_{-\infty}^{z} f_{X+Y}(u)du &= \int_{-2}^{\min\{2,z\}} \frac{2-|u|}{4} du \\
&= I_{z\ge 2} + I_{-2\le z<2} \left(\frac{1}{2} + \int_0^z \frac{2-u\cdot \mathrm{sgn}(z)}{4} du\right) \\
&= I_{z\ge 2} + I_{-2\le z<2} \left(\frac{1}{2} + \frac{z}{2} - \frac{z^2\cdot \mathrm{sgn}(z)}{8}\right) \\
&= I_{z\ge 2} + I_{-2\le z<2} \left(\frac{1}{2} + \frac{z}{2} - \frac{z|z|}{8}\right).
\end{aligned}
$$

<u>解法 2</u>**:** 由于 X,Y 联合概率密度函数关于自变量 (x,y) 具有一定对称性，所以考虑对称的变量替换

$$
\begin{cases} u &= x+y, \\ v &= x-y, \end{cases}
$$

由此解得

$$
\begin{cases} x &= (u+v)/2, \\ y &= (u-v)/2, \end{cases}
$$

于是得到 $U=X+Y$ 和 $V=X-Y$ 的联合概率密度函数

$$
\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}\left(\frac{u+v}{2},\frac{u-v}{2}\right) \frac{1}{\left|\det\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\right|} \\
&= \frac{1+\frac{u^2-v^2}{4}uv}{8} I_{\left|\frac{u+v}{2}\right|\le 1,\left|\frac{u-v}{2}\right|\le 1}.
\end{aligned}
$$

所以，

$$
\begin{aligned}
f_{X+Y}(u) = f_U(u) &= \int_{-\infty}^{+\infty} f_{U,V}(u,v)dv \\
&= \int_{-\infty}^{+\infty} \frac{1+\frac{u^2-v^2}{4}uv}{8} I_{|u+v|\le 2, |u-v|\le 2} dv \\
&= \int_{-\infty}^{+\infty} \frac{1}{8} I_{|u+v|\le 2, |u-v|\le 2} dv \qquad \text{(利用了积分的对称性)} \\
&= \frac{1}{8}\int_{-\infty}^{+\infty} I_{-2+|u|\le v\le 2-|u|} dv \\
&= \frac{2-|u|}{4} I_{|u|\le 2}.
\end{aligned}
$$

然后，再象解法 1 中那样求出 F_{X+Y}。

<u>**解法 3**</u>：直接计算 $X+Y$ 的概率分布函数

$$
\begin{aligned}
F_{X+Y}(z) &= P(X+Y\le z) \\
&= \iint\limits_{x+y\le z} f_{X,Y}(x,y)dxdy \\
&= \iint\limits_{\mathbb{R}^2} \frac{1+xy(x^2-y^2)}{4} I_{|x|\le 1, |y|\le 1, x+y\le z} dxdy \\
&= \frac{1}{4}\iint\limits_{\mathbb{R}^2} I_{|x|\le 1, |y|\le 1, x+y\le z} dxdy + \iint\limits_{\mathbb{R}^2} \frac{xy(x^2-y^2)}{4} I_{|x|\le 1, |y|\le 1, x+y\le z} dxdy \\
&= \frac{1}{4}\iint\limits_{\mathbb{R}^2} I_{|x|\le 1, |y|\le 1, x+y\le z} dxdy \qquad \text{(上式第 2 个积分关于 } (x,y)\mapsto(y,x) \text{ 具有对称性)} \\
&= I_{-2\le z<0}\frac{(z+2)^2}{8} + I_{0\le z<2}\left(1-\frac{(2-z)^2}{8}\right) + I_{z\ge 2}.
\end{aligned}
$$

这与前两个解法求得的 F_{X+Y} 是相同的。

题 4 设随机变量 X，Y 独立，都服从标准正态分布 $N(0,1)$，(X,Y) 的联合密度函数记为 $f(x,y)$；

1. 证明；函数 $g(x,y) = \begin{cases} f(x,y) + \dfrac{xy}{100}, & when\ x^2+y^2\le 1 \\ f(x,y), & when\ x^2+y^2>1 \end{cases}$，是二维概率密度函数。
2. 若随机变量 (U,V) 的密度函数为 $g(x,y)$，证明：U，V 都服从标准正态分布 $N(0,1)$，但 (U,V) 不服从二维正态分布。

注：本例说明各分量为正态推不出联合分布为正态。
<u>**解**</u>：

1. 证明：$f(x,y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$,当$x^2+y^2\le 1$时，$f(x,y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}} \ge \frac{1}{2\pi}e^{-\frac{1}{2}} > \frac{1}{100}$,
   所以当$x^2+y^2\le 1$时，$g(x,y)\ge 0$,

又因为

$$
\begin{aligned}
&\iint\limits_{R^2} g(x,y)dxdy \\
&\qquad = \iint\limits_{x^2+y^2\le 1} (f(x,y)+\frac{xy}{100})dxdy + \iint\limits_{x^2+y^2>1} f(x,y)dxdy \\
&\qquad = \iint\limits_{R^2} f(x,y)dxdy + \iint\limits_{x^2+y^2\le 1} \frac{xy}{100})dxdy \\
&\qquad = 1.
\end{aligned}
$$

所以 $g(x,y)=\begin{cases} f(x,y)+\dfrac{xy}{100}, & when\ x^2+y^2\le 1 \\ f(x,y), & when\ x^2+y^2>1 \end{cases}$，是二维概率密度函数。

2. 随机变量 (U,V) 的密度函数为 $g(x,y)$，则随机变量U的概率密度函数为：

$$
\begin{aligned}
f_U(x) &= \int_{-\infty}^{+\infty} g(x,y)dy \\
&= \int_{-\infty}^{-\sqrt{1-x^2}} f(x,y)dy + \int_{\sqrt{1-x^2}}^{+\infty} f(x,y)dy + \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (f(x,y)+\frac{xy}{100})dy \\
&= \int_{-\infty}^{+\infty} f(x,y)dy.
\end{aligned}
$$

因为 $f(x,y)$为独立标准正态分布联合密度， U为标准正态分布. 同理可得 V为标准正态分布，但是 U,V的联合密度函数为 $g(x,y)=\begin{cases} f(x,y)+\dfrac{xy}{100}, & when\ x^2+y^2\le 1 \\ f(x,y), & when\ x^2+y^2>1 \end{cases}$， U,V不服从二维正态分布.

题 5 投掷一个公平的硬币，及正面为H、反面为T，

1. 直至首次出现HH时停止，请计算投掷次数的期望与方差；
2. 如果停止准则变为首次出现HT，此时投掷次数的期望和方差分别是多少；
3. 假设甲、乙进行一场比赛，投掷一个硬币直至首次出现HH或TH停止。如果以HH结束，则甲胜，TH结束为乙胜，请问甲、乙的获胜概率; 若改为HH先出现甲胜，HT先出现乙胜，结果怎样。

解:

1. 设X为投掷次数，Y为随机变量，其样本为:首次掷出反面（记为T），前两次掷出正反（记为HT），前两次掷出反正（记为TH）。则

$$
\begin{aligned}
E(X) &= E(E(X|Y)) \\
&= P(Y=T)E(X|Y=T)+P(Y=HT)E(X|Y=HT)+P(Y=HH)E(X|Y=HH) \\
&= \tfrac{1}{2}(1+E(X))+\tfrac{1}{4}(2+E(X))+\tfrac{1}{4}\cdot 2
\end{aligned}
$$

解得$E(X)=6$。

$$\begin{aligned}
E(X^2) &= E(E(X^2|Y)) \\
&= P(Y=T)E(X^2|Y=T)+P(Y=HT)E(X^2|Y=HT)+P(Y=HH)E(X^2|Y=HH) \\
&= \tfrac{1}{2}E((X+1)^2)+\tfrac{1}{4}E((2+X)^2)+\tfrac{1}{4}\cdot 4
\end{aligned}$$

解得$E(X^2)=58$，$Var(X)=22$。

2.

$$\begin{aligned}
E(X|H) &= P(Y=H)E(X|HH)+P(Y=T)E(X|HT) \\
&= \tfrac{1}{2}(1+E(X|H))+\tfrac{1}{2}\cdot 2 \\
&= \tfrac{1}{2}(E(X|H))+\tfrac{3}{2}
\end{aligned}$$

解得$E(X|H)=3$。
又由

$$\begin{aligned}
E(X) &= P(T)E(X|T)+P(H)E(X|H) \\
&= \tfrac{1}{2}E(X)+2
\end{aligned}$$

解得$E(X)=4$。

$$\begin{aligned}
E(X^2|H) &= P(T)E(X^2|HT)+P(H)E(X^2|HH) \\
&= \tfrac{1}{2}E((X+1)^2|H)+\tfrac{1}{2}\cdot 4 \\
&= \tfrac{1}{2}E(X^2|H)+\tfrac{11}{2}
\end{aligned}$$

解得$E(X^2|H)=11$。
又由

$$\begin{aligned}
E(X^2) &= P(T)E(X^2|T)+P(H)E(X^2|H) \\
&= P(T)(E(X)^2+9)+P(H)E(X^2|H) \\
&= \tfrac{1}{2}E(X^2)+10
\end{aligned}$$

解得$E(X^2)=20$，$Var(X)=4$。

3. 设甲胜为事件X，Y为随机变量，其样本为:掷出反面（记为T），与掷出正面（记为H）。

$P(X)=P(Y=H)P(X|Y=H)+P(Y=T)P(X|Y=T)=\frac{1}{2}P(X|Y=H)+\frac{1}{2}P(X|Y=T)$

$P(X|Y=H)=P(Y=H)P(X|HH)+P(Y=T)P(X|HT)=\frac{1}{2}\cdot 1+\frac{1}{2}P(X|T)$

$P(X|Y=T)=P(Y=H)P(X|TH)+P(Y=T)P(X|TT)=\frac{1}{2}\cdot 0+\frac{1}{2}P(X|T)$

$\Rightarrow P(X|Y=T)=0$。

所以$P(X)=1/4$。

同理，若记乙胜为事件X,有$P(X)=3/4$。具体计算如下：

设乙胜为事件X，Y为随机变量，其样本为:掷出反面（记为T），与掷出正面（记为H）。

$P(X)=P(Y=H)P(X|Y=H)+P(Y=T)P(X|Y=T)=\frac{1}{2}P(X|Y=H)+\frac{1}{2}P(X|Y=T)$

$P(X|Y=H)=P(Y=H)P(X|HH)+P(Y=T)P(X|HT)=\frac{1}{2}\cdot 0+\frac{1}{2}P(X|T)$

$P(X|Y=T)=P(Y=H)P(X|TH)+P(Y=T)P(X|TT)=\frac{1}{2}\cdot 1+\frac{1}{2}P(X|T)$

$\Rightarrow P(X|Y=T)=1$，$P(X|Y=H)=\frac{1}{2}$。

所以

$$\begin{aligned} P(X) &= P(Y=H)P(X|Y=H)+P(Y=T)P(X|Y=T) \\ &= \tfrac{1}{2}P(X|Y=H)+\tfrac{1}{2}P(X|Y=T) \\ &= \tfrac{1}{2}\cdot\tfrac{1}{2}+\tfrac{1}{2}\cdot 1=\tfrac{3}{4}. \end{aligned}$$

若HH为甲胜，HT为乙胜，则投掷出现反面对于二者输赢无影响，若出现正面，则游戏一定结束，则H之后出现正面反面的概率各为 1/2，则甲胜的概率为 1/2。（详细证明如下：可设甲胜为事件X，Y为随机变量，其样本为:掷出反面（记为T），与掷出正面（记为H）。$P(X)=P(Y=H)P(X|Y=H)+P(Y=T)P(X|Y=T)=\frac{1}{2}P(X|Y=H)+\frac{1}{2}P(X|Y=T)$

$P(X|Y=H)=P(Y=H)P(X|HH)+P(Y=T)P(X|HT)=\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 0$

$P(X|Y=T)=P(Y=H)P(X|TH)+P(Y=T)P(X|TT)=\frac{1}{2}\P(X|Y=H)+\frac{1}{2}P(X|T)$

$\Rightarrow P(X|Y=T)=P(X|Y=H)=\frac{1}{2}$

所以$P(X)=1/2$。）

思考：如果更一般的情况，有$1\sim n$编号的n个筹码，随机抽取筹码，如果相连的两个11先出现甲胜，相连的两个n1先出现乙胜，求甲的获胜概率。

题 6 设随机变量$X\sim N(0,1)$, 随机变量W服从分布律$\begin{pmatrix} W & 1 & -1 \\ P & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$，$X$与$W$相互独立。令$Y=XW$。

证明: (1). $Y\sim N(0,1)$, (2). $Cov(X,Y)=0$, (3). X与Y不独立。

（这个例子表明两个正态分布随机变量相互不相关并不能推出这两个随机变量独立。只有当两个随机变量是二元正态分布时，不相关才与独立等价。）

解：（1）利用全概率公式

$$\begin{aligned} F_Y(y) &= P(Y\le y) \\ &= P(Y\le y|W=1)\cdot P(W=1)+P(Y\le y|W=-1)\cdot P(W=-1) \\ &= P(X\le y)\cdot\frac{1}{2}+P(X\ge -y)\cdot\frac{1}{2} \\ &= P(X\le y)\cdot\frac{1}{2}+P(X\le y)\cdot\frac{1}{2} \\ &= \Phi(y). \end{aligned}$$

所以$Y\sim N(0,1)$.

(2) 利用全期望公式

$$\begin{aligned} E(XY) &= E(XY|W=1)\cdot P(W=1)+E(XY|W=-1)\cdot P(W=-1) \\ &= E(X^2)\cdot\frac{1}{2}+E(-X^2)\cdot\frac{1}{2} \\ &= 0. \end{aligned}$$

E(X)=0，E(Y)=0，所以 $Cov(X,Y)=E(XY)-E(X)\cdot E(Y)=0$.

(3) 举一反例，因为$p_{XY}(1,0)=0$, $p_X(1)>0$, $p_Y(0)>0$，所以X与Y不独立。或取定$P(X\in(0,\frac{1}{2}),Y\in(1,2))=0$,但是$P(X\in(0,\frac{1}{2}))>0,P(Y\in(1,2))>0)$，所以不独立。

题 7 设随机变量$X\sim N(0,1)$, 随机变量

$$Y=\begin{cases} X, & 若|X|\ge c \\ -X, & 若|X|<c \end{cases}.$$

(1). 求常数c使得X与Y不相关，
(2). 证明: $Y \sim N(0,1)$,
(3). 说明对任意$c>0$, 如上定义的随机变量X与Y不独立。

解：(1) 利用全期望公式

$$\begin{aligned}
E(XY) &= E(XY \mid |X| \geq c) \cdot P(|X| \geq c) + E(XY \mid |X| < c) \cdot P(|X| < c) \\
&= E(X^2 \mid |X| \geq c) \cdot P(|X| \geq c) + E(-X^2 \mid |X| < c) \cdot P(|X| < c) \\
&= E(X^2 \mid |X| \geq c) \cdot P(|X| \geq c) + E(X^2 \mid |X| < c) \cdot P(|X| < c) \\
&\quad - 2E(X^2 \mid |X| < c) \cdot P(|X| < c) \\
&= E(X^2) - 2E(X^2 \mid |X| < c) \cdot P(|X| < c) \\
&= 1 - 4\int_0^c \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx.
\end{aligned}$$

$E(XY) = 1 - 4\int_0^c \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}} dx = 0$，此方程无法解析求解，用数值方法求得 $c \approx 1.5383$.

(2) 利用全概率公式

$$\begin{aligned}
F_Y(y)) &= P(Y \leq y) \\
&= P(Y \leq y \mid |X| \geq c) \cdot P(|X| \geq c) + P(Y \leq y \mid |X| < c) \cdot P(|X| < c) \\
&= P(X \leq y, |X| \geq c) + P(-X \leq y, |X| < c) \\
&= P(X \leq y, |X| \geq c) + P(X \leq y, |X| < c) \\
&= P(X \leq y) = \Phi(y).
\end{aligned}$$

其中倒数第二个等号，利用了X分布的对称性与$|X|<c$的对称性，即有等式 $P(-X \leq y, |X| < c) = P(X \leq y, |X| < c)$.

(3) 举一反例，因为$p_{XY}(1,0) = 0$, $p_X(1) > 0$, $p_Y(0) > 0$，所以X与Y不独立。或者由$P(X \in (c, c+1), Y \in (c+2, c+3)) = 0$,但$P(X \in (c, c+1)) > 0, P(Y \in (c+2, c+3)) > 0)$，也可得$X, Y$不独立。