四．(12 分)随机变量(X_1,X_2)的密度函数为 $p(x,y)=\frac{1}{2\sqrt{3}\,\pi}e^{\frac{-2}{3}\left[(x-1)^2-\frac{(x-1)(y-2)}{2}+\frac{(y-2)^2}{4}\right]}$，$Y=2X_1+X_2$。

（1）求Y的分布；（2）计算X_1和Y的相关系数$Corr(X_1,Y)$；（3）计算$E(X_1|Y=1)$。

解：（1）(X_1,X_2)服从二维正态分布$N(1,2,1,4,-0.5)$，$Y=2X_1+X_2$服从正态分布

$$E(Y)=E(2X_1+X_2)=2E(X_1)+E(X_2)=4$$

$$Cov(X_1,X_2)=\rho\cdot\sigma_1\cdot\sigma_2=-0.5\cdot 2=-1$$

$$Var(Y)=Var(2X_1+X_2)=4Var(X_1)+Var(X_2)+4Cov(X_1,X_2)$$

$$=4\cdot 1+4-4\cdot 1=4$$

$Y\sim N(4,2^2)$。

（2）$Cov(X_1,Y)=Cov(X_1,2X_1+X_2)=2Var(X_1)+Cov(X_1,X_2)=2-1=1$

$$Corr(X_1,Y)=\frac{Cov(X_1,Y)}{\sqrt{Var(X_1)\cdot Var(Y)}}=\frac{1}{1\cdot 2}=0.5$$

（3）因为 $Cov(X_1,Y)=Cov(X_1,2X_1+X_2)=1$，$Var(Y)=4$

所以 $Cov\left(X_1-\frac{Y}{4},Y\right)=0$，$\left(X_1-\frac{Y}{4},Y\right)$服从二位正态分布，所以相互独立

$$E(X_1|Y=1)=E\left(X_1-\frac{Y}{4}+\frac{Y}{4}\middle|Y=1\right)=E\left(X_1-\frac{Y}{4}\right)+\frac{1}{4}=\frac{1}{4}。$$

7. 二维随机变量$(X,Y)\sim N(0,0,1,1,0.5)$，$E(X^2\,|\,X+Y=0)=$____________。

【答案】 $\frac{1}{4}$

解： $U=X-Y$和$V=X+Y$，相互独立，

$$E(X^2\,|\,X+Y=0)=E\left(\left(\frac{U+V}{2}\right)^2\middle|V=0\right)=\frac{1}{4}E(U^2+2UV+V^2\,|\,V=0)$$

5. 随机变量的联合密度函数 $p(x,y)=\begin{cases}\dfrac{e^{-(x+y)/2}}{4}, & x,y>0\\ 0, & \text{其他}\end{cases}$，$Z=\begin{cases}2X+Y, & \text{若 } X\ge Y\\ 3X, & \text{若 } X<Y\end{cases}$，

则 $E(Z|X<Y)=$______________。

【答案】 3

$$E(Z|X<Y)=E(3X|X<Y)=\frac{3\int_0^{+\infty}dx\int_x^{+\infty}x\frac{e^{-x/2}}{2}\frac{e^{-y/2}}{2}dy}{P(X<Y|)}=6\int_0^{+\infty}x\frac{e^{-x/2}}{2}dx\int_x^{+\infty}\frac{e^{-y/2}}{2}dy$$

$$=6\int_0^{+\infty}x\frac{e^{-x/2}}{2}e^{-x/2}dx=3\int_0^{+\infty}xe^{-x}dx=3$$

注意公式：$E(X|A)=E(XI_A)/p(A)$

习题 2 设 X 与 Y 独立同分布于标准正态分布 $\mathcal{N}(0,1)$，求 $\mathbb{E}[\max\{X,Y\}]$.

一个结论 $max(x,y)=x-y+|x-y|$

X，Y$\sim\mathcal{N}(0,0,1,1,\rho)$. 求X-Y与XY的协方差及相关系数

$Cov(X-Y,XY)=Cov(X,XY)-Cov(Y,XY)=0$（对称性）

E|X|=$\sigma\sqrt{\frac{2}{\pi}}$，$if\ X\sim N(0,\sigma)$

习题 5 设随机变量 $X_1,X_2,\ldots,X_n$ 中任意两个的相关系数都是 ρ，试证：$\rho\ge-\frac{1}{n-1}$.

一个结论 $E[(\frac{X-\mu_x}{\sigma_x})(\frac{Y-\mu_y}{\sigma_y})]=r(X,Y)$

习题 9 设随机变量 $X\sim\mathcal{N}(\mu,1)$，$Y\sim\mathcal{N}(0,1)$，且 X 与 Y 相互独立。令

$$I=\begin{cases}1, & Y<X\\ 0, & Y\ge X\end{cases}$$

试证明：

1) $\mathbb{E}[I|X=x]=\Phi(x)$;
2) $\mathbb{E}[\Phi(X)]=P(Y<X)$;
3) $\mathbb{E}[\Phi(X)]=\Phi(\mu/\sqrt{2})$;.

注意Monte-Carol算法的平均值法要讨论期望的存在性

24. 为确定某城市成年男子中吸烟者的比例 p，任意调查 n 个成年男子，记其中的吸烟人数为 m，问 n 至少为多大才能保证 m/n 与 p 的差异小于 0.01 的概率大于 95%.

25. 设 $X \sim Ga(n,1)$，试问 n 应该多大，才能满足

$$P\left(\left|\frac{X}{n}-1\right|>0.1\right)<0.01.$$

$P(|X-EX|\geq\varepsilon)\leq\frac{VarX}{\varepsilon^2}$

26. 设 $\{X_n\}$ 为一独立同分布的随机变量序列，已知 $E(X_i^k)=\alpha_k, k=1,2,3,4$. 试证明：当 n 充分大时，$Y_n=\frac{1}{n}\sum_{i=1}^{n}X_i^2$ 近似服从正态分布，并指出此正态分布的参数.

$P\left(\left|\frac{m}{n}-p\right|\leq 0.01\right)\geq 0.95$

$\frac{m-np}{n}$

27. 用概率论的方法证明：

$$\lim_{n\to\infty}\left(1+n+\frac{n^2}{2!}+\cdots+\frac{n^n}{n!}\right)e^{-n}=\frac{1}{2}.$$

24. 要对p使用不等式 27. Poisson分布可加性

1. 设 $\{X_n\}$ 为独立同分布的随机变量序列，其共同分布为

$$P(X_n=\frac{2^k}{k^2})=\frac{1}{2^k},\qquad k=1,2,\ldots$$

试问 $\{X_n\}$ 是否服从大数定律？

由辛钦大数定律，只要数学期望存在则服从大数定律

回顾马尔科夫大数定律：$\frac{1}{n^2}Var(\sum X_i)\to 0(n\to\infty)$

Monte-Carol算法（随机投点+平均值法）

$$J=\int_{-1}^{1}e^x dx$$

首先把积分区间修改为[0,1]，变为$J=\int_0^1 2e^{2t-1}dt$，

再将被积函数区间修改为[0,1]，变为$J=2(e-e^{-1})\int_0^1\frac{2e^{2t-1}-e^{-1}}{e-e^{-1}}dt+2e^{-1}$

（注意先修改区间下限，再改上限）

即对任意的 $\varepsilon>0$，有 $P\{|X-Y|\geq\varepsilon\}=0$. 于是有

$$P(X\neq Y)=P\left(\bigcup_{k=1}^{+\infty}\left\{|X-Y|\geq\frac{1}{k}\right\}\right)\leq\sum_{k=1}^{+\infty}P\left\{|X-Y|\geq\frac{1}{k}\right\}=0,$$

从而 $P(X=Y)=1$ 成立，结论得证.

9. 设随机变量 X 服从(1,2)上的均匀分布，在 $X=x$ 的条件下，随机变量 Y 的条件分布是参数为 x 的指数分布，证明：XY 服从参数为 1 的指数分布.

做法1：XY双射变换

做法2：条件概率$F_T(t)=P(XY\leq t)=\int P(XY\leq t|X=x)P_X(x)dx=\int_1^2 P(Y\leq t/x|X=x)P_X(x)dx$

16. 设随机变量 X 与 Y 独立同分布，都服从参数为 λ 的指数分布. 令

$$Z=\begin{cases}3X+1, & X\geq Y,\\ 6Y, & X<Y.\end{cases}$$

求 $E(Z)$.

方法1：$E(Z)=E(Z|X>Y)P(X>Y)+E(Z|X<Y)P(X<Y)$

方法2：$E(Z)=\int E(Z|X=x)P_X(x),\int E(Z|X=x)=\int_0^x(3x+1)P_Y(y)dy+\int()P_Y(y)dy$（要注意只有独立才能这么做）

方法3：
$E(Z)=\int E(Z|X=x)P_X(x)=\int\int_{(y)}E(Z|X=x,Y=y)P_{Y|X}(y|x)P_X(x)dydx=\int\int E(Z|X=x,Y=y)P_{X,Y}(x,y)dydx$

Ans: $1/2+27/(4\lambda)$

18. 设 $X_1,X_2,\cdots$ 为独立同分布的随机变量序列，且方差存在. 随机变量 N 只取正整数值，$\mathrm{Var}(N)$ 存在，且 N 与 $\{X_n\}$ 独立. 证明

$$\mathrm{Var}\left(\sum_{i=1}^{N}X_i\right)=\mathrm{Var}(N)[E(X_1)]^2+E(N)\mathrm{Var}(X_1).$$

$E(\sum_{i=1}^{N}X_i))=\sum E(\sum_{i=1}^{N}X_i|N=n)P(N=n)=\sum nP(N=n)EX_i=ENEX_i$

7. 随机变量 (X,Y) 服从以点 $(0,1),(1,0),(1,1)$ 为顶点的三角形区域上的均匀分布，试求 $E(X+Y)$ 和 $\mathrm{Var}(X+Y)$.

方法1：$E(X+Y)=\int E(X+Y|X=x)P_X(x)=\int\int E(X+Y|X=x,Y=y)P_{X,Y}(x,y)dxdy$

方法2：$E(X+Y)=EX+EY,Var(X+Y)=VarX+VarY+2Cov(X,Y)$

22. 某箱装 100 件产品，其中一、二和三等品分别为 80，10 和 10 件. 现从中随机取一件，定义三个随机变量 X_1,X_2,X_3 如下

$$X_i=\begin{cases}1, & 若抽到 i 等品,\\ 0, & 其他,\end{cases}\quad i=1,2,3.$$

试求随机变量 X_1 和 X_2 的相关系数 $\mathrm{Corr}(X_1,X_2)$.

方法1：
$X_1+X_2+X_3=1\Rightarrow X_1+X_2=1-X_3\Rightarrow Var(1-X_3)=Var(X_3)=Var(X_1+X_2)=VarX_1+VarX_2+2Cov$

方法2：写出联合分布列

X1\X2	0	1
0	0.1	0.1
1	0.8	0

X1X2	0	1
P	1	0

Cov=EX1X2-EX1EX2

期望积分时不要漏掉乘x或y

证明协方差矩阵正定性：$Cov(\sum a_iX_i,\sum a_iX_i)=(a_1\cdots a_n)Cov(X)(a_1\cdots a_n)^T\geq 0\Rightarrow Cov(X)$非负定

随机向量(X_1,X_2,X_3)的相关系数分别为$\rho_{12},\rho_{13},\rho_{23}$，证明：

$$\rho_{12}^2+\rho_{13}^2+\rho_{23}^2\leq 1+\rho_{12}\rho_{13}\rho_{23}$$

pf: 利用$Cov(X)$非负定性质：$|Cov(X)|=\begin{vmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_3\rho_{13}\\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 & \sigma_2\sigma_3\rho_{23}\\ \sigma_1\sigma_3\rho_{13} & \sigma_2\sigma_3\rho_{23} & \sigma_3^2\end{vmatrix}=\sigma_1^2\sigma_2^2\sigma_3^2\begin{vmatrix}1 & \rho_{12} & \rho_{13}\\ \rho_{12} & 1 & \rho_{23}\\ \rho_{13} & \rho_{23} & 1\end{vmatrix}\geq 0$

几种常用的不等式：Chebyshev、Cov内积，$E(\sum(\frac{X_i-\mu_i}{\sigma_i})^2)$、|CovX|、变量求导、$Var\leq E(x-c)^2$

习题3.3.6 注意变量范围，算出以后可以积个分试试

结论：在$P_{X,Y}$中X,Y可分离变量即为独立，但注意，变量的取值范围也不能相互纠缠

对数正态分布$LN(\mu,\sigma^2)$的中位数.

solve: $P(Y\leq y_{0.5})=P(e^X\leq y_{0.5})=P(X\leq\ln y_{0.5})=0.5\Rightarrow\mu=\ln y_{0.5}\Rightarrow y_{0.5}=e^{\mu}$

习题1.5.21

3. 由正态总体 $N(100,4)$ 抽取两个独立样本,样本均值分别为 $\bar{x},\bar{y}$,样本容量分别为 15,20,试求 $P(|\bar{x}-\bar{y}|>0.2)$.

解 由条件得 $\bar{x}\sim N\left(100,\frac{4}{15}\right)$,$\bar{y}\sim N\left(100,\frac{4}{20}\right)$,且 $\bar{x}$ 和 $\bar{y}$ 相互独立,从而

$\bar{x}-\bar{y}\sim N\left(0,\frac{4}{15}+\frac{4}{20}\right)$,即 $\bar{x}-\bar{y}\sim N\left(0,\frac{7}{15}\right)$,于是

$$P(|\bar{x}-\bar{y}|>0.2)=P\left(\frac{|\bar{x}-\bar{y}|}{\sqrt{7/15}}>\frac{0.2}{\sqrt{7/15}}\right)=2(1-\Phi(0.29))=0.7718.$$

记住求最大似然估计时，先连乘、再取ln、再求导，不可错漏步骤

相合性定理:

相合性定理（定理6.2.1） 设$\hat{\theta}_n=\hat{\theta}_n(x_1,\cdots,x_n)$是$\theta$的一个估计量，若

$$\lim_{n\to\infty}E(\hat{\theta}_n)=\theta,\ \lim_{n\to\infty}Var(\hat{\theta}_n)=0$$

则$\hat{\theta}_n$是θ的相合估计.

习题 7 设 $(X_1,X_2,\ldots,X_n)$ 是 n 维随机变量，其协方差矩阵 $\boldsymbol{B}=(b_{ij})$ 存在，其行列式 $\det(\boldsymbol{B})=0$. 证明：各分量之间以概率 1 存在线性关系，即存在一组不全为零的实数 $c_1,c_2,\ldots,c_n$，使得

$$P(c_1X_1+c_2X_2+\cdots+c_nX_n=\text{常数})=1$$

习题 8. 设随机变量 X 在 $[a,b]$ 中取值，证明:

$$\mathrm{Var}(X)\le\frac{(b-a)^2}{4},$$

并说明等号何时成立。

$E(x-c)^2\ge Var(X)$，取$c=(a+b)/2$