Midterm Exam, Mathematical Statistics I, 2013 Fall

Name: Huang Siao-Han

18/18

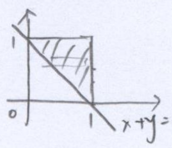
- Note only the answer but also the derivation
- You may use the notation $\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}u^2)du$.

1. [+2]

+4/4

1) A bivariate random vector (X,Y) has the pdf

$$f(x,y)=\begin{cases}6xy^2 & 0<x<1,\ 0<y<1\\ 0 & \text{otherwise}\end{cases}$$. Calculate $P(X+Y>1)$.



$$P(X+Y>1)=\int_0^1\int_{1-y}^1 6xy^2\,dx\,dy$$
$$=\int_0^1 y^2\left[6\left(\frac{x^2}{2}\right)\Big|_{1-y}^1\right]dy$$
$$=\int_0^1 y^2\left[3(1-(1-y)^2)\right]dy$$
$$=\int_0^1 y^2(6y-3y^2)\,dy$$
$$=\int_0^1 6y^3-3y^4\,dy$$
$$=\left(6\frac{y^4}{4}-3\frac{y^5}{5}\right)\Big|_0^1=\frac{3}{2}-\frac{3}{5}=\frac{9}{10}$$ #

2) Let $(\mu_X,\ \mu_Y,\ \sigma_X^2,\ \sigma_Y^2,\ \rho)$ be the parameters of the bivariate normal random variable (X,Y). Let $\begin{bmatrix}U\\V\end{bmatrix}=\begin{bmatrix}a & b\\ b & c\end{bmatrix}\begin{bmatrix}X\\Y\end{bmatrix}$, where $a,\ b,\ c$ are constant.

Derive $Cov(U,V)$ in terms of $(\mu_X,\ \mu_Y,\ \sigma_X^2,\ \sigma_Y^2,\ \rho)$.

$$\begin{bmatrix}U\\V\end{bmatrix}=\begin{bmatrix}a & b\\ b & c\end{bmatrix}\begin{bmatrix}X\\Y\end{bmatrix}=\begin{bmatrix}aX+bY\\ bX+cY\end{bmatrix}$$

$$Cov(U,V)=Cov(aX+bY,\ bX+cY)$$
$$=ab\,Cov(X,X)+ac\,Cov(X,Y)+b^2\,Cov(Y,X)+bc\,Cov(Y,Y)$$
$$=ab\,Var(X)+bc\,Var(Y)+(ac+b^2)\,Cov(X,Y)$$
$$=ab\,\sigma_Y^2+bc\,\sigma_Y^2+(ac+b^2)\,\rho\,\sigma_X\sigma_Y$$ #

$$\rho=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}$$
$$\Rightarrow Cov(X,Y)=\rho\,\sigma_X\sigma_Y$$

+2/2

2. [+2] An *iid* sequence of bivariate random vectors (X_i, Y_i), $i = 1, \ldots, n$ follows a bivariate normal distribution with $EX_i = \mu_X$, $EY_i = \mu_Y$, $Var(X_i) = \sigma_X^2$, $Var(Y_i) = \sigma_Y^2$ and the correlation ρ_{XY}. Calculate $P(X_1 \le Y_1, \cdots, X_n \le Y_n)$.

$$P(X_1 \le Y_1, \ldots, X_n \le Y_n) \overset{\text{independent}}{=} P(X_1 \le Y_1) P(X_2 \le Y_2) \cdots P(X_n \le Y_n)$$

$$\overset{\text{iid}}{=} [P(X_1 \le Y_1)]^n \qquad \bullet\ X_1 - Y_1 \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

$$= [P(X_1 - Y_1 \le 0)]^n$$

$$= \Bigg[P\Bigg(\underbrace{\boxed{\frac{(X_1-Y_1)-(\mu_X-\mu_Y)}{\sqrt{\sigma_X^2+\sigma_Y^2}}}}_{\sim N(0,1)} \le \frac{-(\mu_X-\mu_Y)}{\sqrt{\sigma_X^2+\sigma_Y^2}}\Bigg)\Bigg]^n$$

$$= \Bigg[\Phi\Bigg(\frac{-(\mu_X-\mu_Y)}{\sqrt{\sigma_X^2+\sigma_Y^2}}\Bigg)\Bigg]^n \quad \#$$

+2/2.

3. [+2] Let $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{N}(\mu, \sigma^2)$.

1) Calculate $P(x < X_1 \le x+h, \ldots, x \le X_n \le x+h)$ for $h > 0$.

2) Calculate $\lim_{h \to 0} \dfrac{P(x < X_1 \le x+h, \ldots, x \le X_n \le x+h)}{h^n}$.

(1). $P(x < X_1 \le x+h, \ldots, x \le X_n \le x+h)$

$$\overset{\text{indep}}{=} P(x < X_1 \le x+h) \cdots P(x \le X_n \le x+h)$$

$$\overset{\text{iid}}{=} [P(x < X_1 \le x+h)]^n$$

$$= [P(X_1 \le x+h) - P(X_1 \le x)]^n$$

$$= \Big[P\Big(\underbrace{\frac{X_1-\mu}{\sigma}}_{\sim N(0,1)} \le \frac{(x+h)-\mu}{\sigma}\Big) - P\Big(\underbrace{\frac{X_1-\mu}{\sigma}}_{\sim N(0,1)} \le \frac{x-\mu}{\sigma}\Big)\Big]^n$$

$$= \Big[\Phi\Big(\frac{(x+h)-\mu}{\sigma}\Big) - \Phi\Big(\frac{x-\mu}{\sigma}\Big)\Big]^n \quad \#$$

(2) $\lim_{h \to 0} \dfrac{P(x < X_1 \le x+h, \ldots, x \le X_n \le x+h)}{h^n}$

$$= \lim_{h \to 0} \left[\frac{\Phi\left(\frac{(x+h)-\mu}{\sigma}\right) - \Phi\left(\frac{x-\mu}{\sigma}\right)}{h}\right]^n \qquad \left(\Phi\left(\frac{x-\mu}{\sigma}\right)\right)' = \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \frac{1}{\sigma}$$

$$= \lim_{h \to 0} \left[\frac{\phi\left(\frac{(x+h)-\mu}{\sigma}\right) \cdot \frac{1}{\sigma}}{1}\right]^n$$

$$= \left[\frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right)\right]^n \quad \#$$

+4/4

4. [+4] Let $X_1,\ldots,X_n \overset{iid}{\sim} \mathrm{N}(\mu,\sigma^2)$.

1) Derive the distribution of $(n-1)S^2/\sigma^2$, where S^2 is the sample variance. You may use the independence between the sample mean and sample variance.

$\bar{X} \perp S^2$

2) Compute $E(S^2-\sigma^2)^2$, which is the mean squared error.

$$\frac{-n(\mu-\bar{X})^2}{\sigma^2} = \frac{-2n(\mu-\bar{X})^2}{\sigma^2} + n\frac{(\mu-\bar{X})^2}{\sigma^2} = \frac{2n(\bar{X}-\mu)(\mu-\bar{X})}{\sigma^2} + \frac{n(\mu-\bar{X})^2}{\sigma^2}$$

(1). $\left(\frac{X_i-\mu}{\sigma}\right) \sim N(0,1) \Rightarrow \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_1^2 \Rightarrow \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_n^2$

$$E\left(\exp\left(t\cdot\frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2}\right)\right) = \underbrace{(1-2t)^{-\frac{n}{2}}}_{\text{mgf of } \chi_n^2}$$

$$\frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} + \boxed{2\frac{\sum_{i=1}^n (X_i-\mu)(\mu-\bar{X})}{\sigma^2} + \frac{n(\mu-\bar{X})^2}{\sigma^2}}$$

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i-\bar{X})^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i-\mu+\mu-\bar{X})^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} - \frac{n(\mu-\bar{X})^2}{\sigma^2}$$

$$\Rightarrow \frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} + \frac{n(\mu-\bar{X})^2}{\sigma^2}$$

$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

$$\Rightarrow E\left(\exp\left(t\cdot\frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2}\right)\right) \overset{\because \bar{X}\perp S^2}{=} E\left(\exp\left(t\cdot\frac{(n-1)S^2}{\sigma^2}\right)\right) + E\left(\exp\left(t\cdot\underbrace{\boxed{\frac{(\bar{X}-\mu)^2}{\sigma^2/n}}}_{\sim\chi_1^2}\right)\right)$$

$$\Rightarrow (1-2t)^{-\frac{n}{2}} = E\left(\exp\left(t\cdot\frac{(n-1)S^2}{\sigma^2}\right)\right) + (1-2t)^{-\frac{1}{2}}$$

$$\Rightarrow E\left(\exp\left(t\cdot\frac{(n-1)S^2}{\sigma^2}\right)\right) = \frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}} = (1-2t)^{-\frac{(n-1)}{2}} \longleftarrow \text{mgf of } \chi_{n-1}^2$$

$$\therefore \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

(2) $E\left(\frac{(n-1)S^2}{\sigma^2}\right) = n-1 \Rightarrow \frac{(n-1)}{\sigma^2}E(S^2) = (n-1) \Rightarrow E(S^2) = \sigma^2$

$E(S^2-\sigma^2)^2 = E(S^2-E(S^2))^2 = Var(S^2)$

$Var\left(\frac{(n-1)S^2}{\sigma^2}\right) = \frac{(n-1)^2}{\sigma^4}Var(S^2) = 2(n-1) \Rightarrow Var(S^2) = \frac{2\sigma^4}{n-1}$, $\therefore E(S^2-\sigma^2)^2 = \frac{2\sigma^4}{n-1}$ #

+2/2

5. [+2] Let $X_1,X_2,X_3 \overset{iid}{\sim} \mathrm{N}(\mu,\sigma^2)$. For $s \le t$, derive $P(X_{(1)} \le s \le X_{(2)} \le t \le X_{(3)})$.

$$\left.\begin{aligned} P(X_{(1)} \le s \le X_{(2)} \le t \le X_{(3)}) &= P(X_1 \le s \le X_2 \le t \le X_3) \\ &+ P(X_2 \le s \le X_3 \le t \le X_1) \\ &\ \ \vdots \\ &+ P(X_3 \le s \le X_2 \le t \le X_1) \end{aligned}\right\} 3!$$

$= 3!\, P(X_1 \le s \le X_2 \le t \le X_3)$

$\overset{iid}{=} 3!\, P(X_1 \le s)\, P(s \le X_2 \le t)\, P(t \le X_3)$

$= 6\, P\left(\frac{X_1-\mu}{\sigma} \le \frac{s-\mu}{\sigma}\right) P\left(\frac{s-\mu}{\sigma} \le \frac{X_2-\mu}{\sigma} \le \frac{t-\mu}{\sigma}\right)\left(1-P\left(\frac{X_3-\mu}{\sigma} \le \frac{t-\mu}{\sigma}\right)\right)$

$= 6\, \Phi\left(\frac{s-\mu}{\sigma}\right)\left[\Phi\left(\frac{t-\mu}{\sigma}\right) - \Phi\left(\frac{s-\mu}{\sigma}\right)\right]\left[1-\Phi\left(\frac{t-\mu}{\sigma}\right)\right]$

$U \sim uniform(0,1)$.

$P(F_Y(y_0) \le U \le F_Y(y_{0+1})) \Rightarrow Y = y_{0+1}$

+2/2

6. [+2] Derive the algorithm for generating data from $binomial(n=3, p=2/3)$ by transforming a uniform random variable. Then, check if your algorithm has the same pmf as $binomial(n=3, p=2/3)$

$U \sim uniform(0,1)$.

$Y \sim bin(n=3, p=\frac{2}{3})$, $f_Y(y) = \binom{3}{y} (\frac{2}{3})^y (\frac{1}{3})^{3-y}$, $y = 0,1,2,3$. $= \begin{cases} \frac{1}{27}, & y=0 \\ \frac{2}{9}, & y=1 \\ \frac{4}{9}, & y=2 \\ \frac{8}{27}, & y=3 \end{cases}$

$F_Y(y) = \begin{cases} \frac{1}{27}, & y=0 \\ \frac{7}{27}, & y=1 \\ \frac{19}{27}, & y=2 \\ 1, & y=3 \end{cases}$ $\Rightarrow Y = \begin{cases} 0, & \cancel{0 \le U < F_Y(0)} = 0 \le U < \frac{1}{27} \\ 1, & \cancel{F_Y(0) \le U < F_Y(1)} = \frac{1}{27} \le U < \frac{7}{27} \\ 2, & \cancel{F_Y(1) \le U < F_Y(2)} = \frac{7}{27} \le U < \frac{19}{27} \\ 3, & \cancel{F_Y(2) \le U < F_Y(3)} = \frac{19}{27} \le U \le 1 \end{cases}$

Define $F_Y(-1) = 0$.

$P(Y=y) = P(F_Y(y-1) \le U \le F_Y(y)) = F_Y(y) - F_Y(y-1) = \begin{cases} \frac{1}{27} - 0 = \frac{1}{27}, & y=0 \\ \frac{7}{27} - \frac{1}{27} = \frac{6}{27} = \frac{2}{9}, & y=1 \\ \frac{19}{27} - \frac{7}{27} = \frac{12}{27} = \frac{4}{9}, & y=2 \\ 1 - \frac{19}{27} = \frac{8}{27}, & y=3 \end{cases} = f_Y(y)$

$Y \sim bin(3, \frac{2}{3})$

#

7. [+2] A bivariate random vector (X,Y) has a bivariate normal distribution with parameters (μ_X, μ_Y, σ_X^2, σ_Y^2, ρ). Calculate $E(X-\mu_X)^2(Y-\mu_Y)$.

+2/2.

$\boxed{\text{풀이 I}}$

$E(X-\mu_X)^2(Y-\mu_Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)^2(y-\mu_Y) \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}\{(\frac{x-\mu_X}{\sigma_X})^2 - 2\rho(\frac{x-\mu_X}{\sigma_X})(\frac{y-\mu_Y}{\sigma_Y}) + (\frac{y-\mu_Y}{\sigma_Y})^2\}\}dxdy$

Let $u = (\frac{x-\mu_X}{\sigma_X})$, $v = (\frac{y-\mu_Y}{\sigma_Y})$. $du = \frac{dx}{\sigma_X}$, $dv = \frac{dy}{\sigma_Y}$

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\sigma_X u)^2(\sigma_Y v) \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}\{u^2 - 2\rho uv + v^2\}\} dv\, du$

$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} (\sigma_X u)^2 \exp\{-\frac{u^2}{2(1-\rho^2)}\} \int_{-\infty}^{\infty} (\sigma_Y v) \exp\{-\frac{1}{2(1-\rho^2)}\{(v-\rho u)^2 - \rho^2 u^2\}\} dv\, du$

$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} (\sigma_X u)^2 \exp\{-\frac{u^2-\rho^2u^2}{2(1-\rho^2)}\}\left[\int_{-\infty}^{\infty} (\sigma_Y v) \underbrace{\exp\{-\frac{1}{2(1-\rho^2)}(v-\rho u)^2\}}_{\propto N(\rho u, (1-\rho^2))} dv\right] du$

$= \frac{1}{2\pi\cancel{\sqrt{1-\rho^2}}} \int_{-\infty}^{\infty} (\sigma_X u)^2 \exp\{-\frac{u^2}{2}\} (\sigma_Y \rho u) \sqrt{2\pi}\cancel{\sqrt{1-\rho^2}}\, du$

$= \int_{-\infty}^{\infty} (\sigma_X u)^2 (\sigma_Y \rho u) \frac{1}{\sqrt{2\pi}} \exp(-\frac{u^2}{2}) du$

$= E(\sigma_X u)^2(\sigma_Y \rho u)$, $u \sim N(0,1)$. $\Rightarrow E(u) = 0$, $E(u^2) = 1$, $E(u^3) = 0$.

$= E(\sigma_X^2 u^2)(\sigma_Y \rho u)$

$\hookrightarrow M_u(t) = e^{\frac{t^2}{2}}$

$= E(\sigma_X^2 \sigma_Y \rho u^3)$

$E(u) = \frac{d}{dt} M_u(t)|_{t=0} = \frac{2t}{2} e^{\frac{t^2}{2}}|_{t=0} = 0$

$= \sigma_X^2 \sigma_Y \rho E(u^3)$

$E(u^2) = \frac{d^2}{dt^2} M_u(t)|_{t=0} = e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}}|_{t=0} = e^0 = 1$

$= 0$ #.

$E(u^3) = \frac{d^3}{dt^3} M_u(t)|_{t=0} = t e^{\frac{t^2}{2}} + 2t e^{\frac{t^2}{2}} + t^3 e^{\frac{t^2}{2}}|_{t=0} = 0$

$\boxed{\text{풀이 II}}$ $E(X-\mu_X)^2(Y-\mu_Y) = E \frac{\partial^3}{\partial s^2 \partial t} e^{s(X-\mu_X)+t(Y-\mu_Y)}|_{s=t=0} = \frac{\partial^3}{\partial s^2 \partial t} E e^{s(X-\mu_X)+t(Y-\mu_Y)}|_{s=t=0}$

$= \frac{\partial^3}{\partial s^2 \partial t} M_{s(X-\mu_X)+t(Y-\mu_Y)}(1)|_{s=t=0}$.

$s(X-\mu_X) + t(Y-\mu_Y) \sim N(0, s^2\sigma_X^2 + t^2\sigma_Y^2 + 2st\rho)$

$M_{s(X-\mu_X)+t(Y-\mu_Y)}(1) = \exp(\frac{s^2\sigma_X^2 + t^2\sigma_Y^2 + 2st\rho}{2})$.

$\frac{\partial}{\partial s} M_{s(X-\mu_X)+t(Y-\mu_Y)}(1) = (s\sigma_X^2 + t\rho) M_{s(X-\mu_X)+t(Y-\mu_Y)}(1)$

$\frac{\partial^2}{\partial s^2} M_{s(X-\mu_X)+t(Y-\mu_Y)}(1) = (\sigma_X^2) M_{s(X-\mu_X)+t(Y-\mu_Y)}(1) + (s\sigma_X^2 + t\rho)^2 M_{s(X-\mu_X)+t(Y-\mu_Y)}(1)$.