Sta532 Hw3

Xu Chen 2/2/2020

Exercise 1

(a) Function $f(x) = x^\alpha$ is convex for any $\alpha \geqslant 1$ on $\mathbb{R}^+$ because $f''(x) = \alpha(\alpha-1)x^{\alpha-2} \geqslant 0$. By Jensen's inequality, we have

$$\mathsf{E}[Y^p] = \mathsf{E}[(Y^q)^{p/q}] \geqslant (\mathsf{E}[Y^q])^{p/q}$$

Therefore, $\mathsf{E}[Y^p]^{1/p} \geqslant \mathsf{E}[Y^q]^{1/q}$.

(b) Function $f(x) = 1/x$ is convex on $\mathbb{R}^+$ because $f''(x) = 2/x^3 \geqslant 0$. Therefore, we have $\mathsf{E}[1/Y] \geqslant 1/\mathsf{E}[Y]$.

(c) Function $f(x) = -\log x$ is convex on $\mathbb{R}^+$ because $f''(x) = 1/x^2 \geqslant 0$. Therefore, we have $\mathsf{E}[-\log Y] \geqslant -\log \mathsf{E}[Y]$ and hence $\log \mathsf{E}[Y] \geqslant \mathsf{E}[\log Y]$.

Exercise 2

(a) Denote the three dimensional simplex as $\Delta_3 = \{w = (w_1, w_2, w_3) \mid w_1 + w_2 + w_3 = 1, w_i \in (0,1), i = 1,2,3\}$. Let $\alpha = \sum_{i=1}^3 \alpha_i$. Therefore, taking w_1 as an example, we have

$$\begin{aligned}\mathsf{E}[w_1] &= \int_{\omega\in\Delta_3} \omega_1 p_w(\omega_1,\omega_2,\omega_3)\mathrm{d}\omega = \int_{\omega\in\Delta_3} \frac{1}{\mathrm{B}(\alpha_1,\alpha_2,\alpha_3)}\omega_1^{\alpha_1}\omega_2^{\alpha_2-1}\omega_3^{\alpha_3-1}\mathrm{d}\omega \\ &= \frac{\mathrm{B}(\alpha_1+1,\alpha_2,\alpha_3)}{\mathrm{B}(\alpha_1,\alpha_2,\alpha_3)} = \frac{\Gamma(\alpha_1+1)}{\Gamma(\alpha+1)}\frac{\Gamma(\alpha)}{\Gamma(\alpha_1)} = \frac{\alpha_1}{\alpha}\end{aligned}$$

Similarly,

$$\mathsf{E}[w_1^2] = \int_{\omega\in\Delta_3} \omega_1^2 p_w(\omega_1,\omega_2,\omega_3)\mathrm{d}\omega = \frac{\mathrm{B}(\alpha_1+2,\alpha_2,\alpha_3)}{\mathrm{B}(\alpha_1,\alpha_2,\alpha_3)} = \frac{\Gamma(\alpha_1+2)}{\Gamma(\alpha+2)}\frac{\Gamma(\alpha)}{\Gamma(\alpha_1)} = \frac{\alpha_1(\alpha_1+1)}{\alpha(\alpha+1)}$$

Therefore,

$$\mathsf{Var}[w_1] = \frac{\alpha_1(\alpha_1+1)}{\alpha(\alpha+1)} - \frac{\alpha_1^2}{\alpha^2} = \frac{\alpha_1(\alpha_2+\alpha_3)}{\alpha^2(\alpha+1)}$$

The expectations and variances for w_2 and w_3 can be derived similarly by symmetry.

(b)

$$\begin{aligned}\mathsf{E}[w_1w_2] &= \int_{\omega\in\Delta_3} \omega_1\omega_2 p_w(\omega_1,\omega_2,\omega_3)\mathrm{d}\omega = \frac{\mathrm{B}(\alpha_1+1,\alpha_2+1,\alpha_3)}{\mathrm{B}(\alpha_1,\alpha_2,\alpha_3)} = \frac{\Gamma(\alpha_1+1)\Gamma(\alpha_2+1)}{\Gamma(\alpha+2)}\frac{\Gamma(\alpha)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \\ &= \frac{\alpha_1\alpha_2}{\alpha(\alpha+1)}\end{aligned}$$

Therefore,

$$\begin{aligned}\mathsf{Cov}(w_1, w_2) &= \frac{\alpha_1\alpha_2}{\alpha(\alpha+1)} - \frac{\alpha_1\alpha_2}{\alpha^2} \\ &= -\frac{\alpha_1\alpha_2}{\alpha^2(\alpha+1)}\end{aligned}$$

The covariance between w_1 and w_2 is negative. It makes sense because $\sum_{i=1}^{3} w_i = 1$ and when w_1 becomes larger, w_2 is likely to become smaller.

(c) The variance of w_1 has been derived in part **(a)**.

$$\mathsf{Var}[w_1 + w_2] = \mathsf{Var}[1 - w_3] = \mathsf{Var}[w_3] = \frac{\alpha_3(\alpha_1+\alpha_2)}{\alpha^2(\alpha+1)}$$

(d) The density function of w_1 is given by

$$\begin{aligned}p_{w_1}(\omega_1) &= \int_{\omega\in\Delta_3} p_w(\omega_1, \omega_2, \omega_3)\mathrm{d}\omega_3 \\ &\propto \int_0^{1-\omega_1} \omega_1^{\alpha_1-1}(1-\omega_1-\omega_3)^{\alpha_2-1}\omega_3^{\alpha_3-1}\mathrm{d}\omega_3 \\ &\propto \omega_1^{\alpha_1-1}(1-\omega_1)^{\alpha_2+\alpha_3-2}\int_0^{1-\omega_1}(1-\omega_3/(1-\omega_1))^{\alpha_2-1}(\omega_3/(1-\omega_1))^{\alpha_3-1}\mathrm{d}\omega_3 \\ &\propto \omega_1^{\alpha_1-1}(1-\omega_1)^{\alpha_2+\alpha_3-1}\int_0^1 (1-u)^{\alpha_2-1}u^{\alpha_3-1}\mathrm{d}u \\ &\propto \omega_1^{\alpha_1-1}(1-\omega_1)^{\alpha_2+\alpha_3-1}\end{aligned}$$

Therefore, $w_1 \sim \mathsf{Beta}(\alpha_1, \alpha_2+\alpha_3)$. Similarly, $w_3 \sim \mathsf{Beta}(\alpha_3, \alpha_1+\alpha_2)$ and hence

$$p_{w_3}(\omega_3) = \frac{1}{\mathrm{B}(\alpha_3, \alpha_1+\alpha_2)}\omega_3^{\alpha_3-1}(1-\omega_3)^{\alpha_1+\alpha_2-1}$$

Let $u = w_1 + w_2 = 1 - w_3$. Then by change of variable, the density function of u is

$$p_u(u) = \frac{1}{\mathrm{B}(\alpha_3, \alpha_1+\alpha_2)}(1-u)^{\alpha_3-1}u^{\alpha_1+\alpha_2-1}$$

Therefore, $w_1 + w_2 \sim \mathsf{Beta}(\alpha_1+\alpha_2, \alpha_3)$.

Exercise 3

(a) Let us assume that X and Y are continuous random variable with marginal density functions $p_X(x)$ and $p_Y(y)$. Since X and Y are independent, we have $p_{X,Y}(x,y) = p_X(x)p_Y(y)$. Therefore,

$$\begin{aligned}\mathsf{E}[XY] &= \int xyp_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = \int xyp_X(x)p_Y(y)\mathrm{d}x\mathrm{d}y \\ &= \int xp_X(x)\mathrm{d}x\int yp_Y(y)\mathrm{d}y = \mathsf{E}[X]\mathsf{E}[Y]\end{aligned}$$

Therefore, $\mathsf{Cov}(X,Y) = \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] = 0$. The proof for discrete random variables is similar.

(b) Since $X = a + bY$, we have

$$\begin{aligned}\mathsf{E}[X] &= \mathsf{E}[a + bY] = a + b\mathsf{E}[Y] \\ \mathsf{Var}[X] &= \mathsf{Var}[a + bY] = b^2\mathsf{Var}[Y]\end{aligned}$$

Therefore,

$$\begin{aligned}\mathsf{Cor}(X, Y) &= \frac{\mathsf{E}[(X - \mathsf{E}[X])(Y - \mathsf{E}[Y])]}{\sqrt{\mathsf{Var}[X]\mathsf{Var}[Y]}} \\ &= \frac{b\mathsf{E}[(Y - \mathsf{E}[Y])^2]}{\sqrt{b^2\mathsf{Var}[Y]^2}} = \frac{b}{|b|}\end{aligned}$$

Therefore, $\mathsf{Cor}(X, Y) = 1$ if $b > 0$ or -1 if $b < 0$.

(c)

$$\begin{aligned}\mathsf{Cov}(a_1 + b_1X_1, a_2 + b_2X_2) &= \mathsf{E}[(a_1 + b_1X_1 - \mathsf{E}[a_1 + b_1X_1])(a_2 + b_2X_2 - \mathsf{E}[a_2 + b_2X_2])] \\ &= \mathsf{E}[b_1b_2(X_1 - \mathsf{E}[X_1])(X_2 - \mathsf{E}[X_2])] = b_1b_2\mathsf{Cov}(X_1, X_2)\end{aligned}$$

Let us assume the joint density function of $X = (X_1, X_2, X_3)$ is $p_X(x_1, x_2, x_3)$. We have

$$\begin{aligned}\mathsf{E}[X_1 + X_2 + X_3] &= \int\int\int(x_1 + x_2 + x_3)p_X(x_1, x_2, x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &= \sum_{i=1}^{3}\int\int\int x_ip_X(x_1, x_2, x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 \\ &= \sum_{i=1}^{3}\int x_ip_{X_i}(x_i)\mathrm{d}x_i = \sum_{i=1}^{3}\mathsf{E}[X_i]\end{aligned}$$

Using the result above, we have

$$\begin{aligned}\mathsf{Var}[X_1 + X_2 + X_3] &= \mathsf{E}[((X_1 + X_2 + X_3) - \mathsf{E}(X_1 + X_2 + X_3))^2] \\ &= \mathsf{E}[((X_1 - \mathsf{E}[X_1]) + (X_2 - \mathsf{E}[X_2]) + (X_3 - \mathsf{E}[X_3]))^2] \\ &= \mathsf{E}[(X_1 - \mathsf{E}[X_1])^2 + (X_2 - \mathsf{E}[X_2])^2 + (X_3 - \mathsf{E}[X_3])^2 + (X_1 - \mathsf{E}[X_1])(X_2 - \mathsf{E}[X_2])] \\ &+ \mathsf{E}[(X_1 - \mathsf{E}[X_1])(X_3 - \mathsf{E}[X_3]) + (X_2 - \mathsf{E}[X_2])(X_3 - \mathsf{E}[X_3])] \\ &= \mathsf{Var}[X_1] + \mathsf{Var}[X_2] + \mathsf{Var}[X_3] + \mathsf{Cov}(X_1, X_2) + \mathsf{Cov}(X_1, X_3) + \mathsf{Cov}(X_2, X_3)\end{aligned}$$

Exercise 4

(a) $\mathsf{E}[Y_1] = \mathsf{E}[Z] + \mathsf{E}[X_1] = 0, \quad \mathsf{E}[Y_2] = \mathsf{E}[Z] + \mathsf{E}[X_2] = 0, \quad \mathsf{E}[Y_3] = \mathsf{E}[Z^2] + \mathsf{E}[X_3] = 1$

Since Z, X_1, X_2, X_3 are independent standard normal random variables, we have

$$\begin{aligned}\mathsf{Var}[Y_1] &= \mathsf{Var}[Z] + \mathsf{Var}[X_1] = 1 + 1 = 2 \\ \mathsf{Var}[Y_2] &= \mathsf{Var}[Z] + \mathsf{Var}[X_2] = 1 + 1 = 2 \\ \mathsf{Var}[Y_3] &= \mathsf{Var}[Z^2] + \mathsf{Var}[X_3] = \mathsf{E}[Z^4] - (\mathsf{E}[Z^2])^2 + 1 = 3 - 1 + 1 = 3\end{aligned}$$

(b)

$$\begin{aligned}
\mathsf{Cov}(Y_1, Y_2) &= \mathsf{E}[(Z + X_1)(Z + X_2)] = \mathsf{E}[Z^2 + ZX_1 + ZX_2 + X_1X_2] = 1 \\
\mathsf{Cov}(Y_1, Y_3) &= \mathsf{E}[(Z + X_1)(Z^2 + X_3 - 1)] = \mathsf{E}[Z^3 + ZX_3 - Z + X_1Z^2 + X_1X_3 - X_1] = 0 \\
\mathsf{Cov}(Y_2, Y_3) &= \mathsf{E}[(Z + X_2)(Z^2 + X_3 - 1)] = \mathsf{E}[Z^3 + ZX_3 - Z + X_2Z^2 + X_2X_3 - X_2] = 0
\end{aligned}$$

Therefore, the variance-covariance matrix of $Y = [Y_1, Y_2, Y_3]$ is

$$\mathsf{Cov}[Y] = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

(c) No, they are not independent.

$$\begin{aligned}
\mathsf{E}[Y_1^2 Y_3] &= \mathsf{E}[(Z + X_1)^2(Z^2 + X_3)] \\
&= \mathsf{E}[Z^4 + Z^2X_3 + X_1^2Z^2 + X_1^2X_3 + 2Z^3X_1 + 2ZX_1X_3] \\
&= 3 + 1 = 4 \\
\mathsf{E}[Y_1^2]\mathsf{E}[Y_3] &= \mathsf{E}[Z^2 + X_1^2 + 2ZX_1] = 2
\end{aligned}$$

Therefore, $\mathsf{Cov}(Y_1^2, Y_3) \neq 0$ and hence Y_1^2 and Y_3 are not independent. Therefore, Y_1 and Y_3 are not independent (Otherwise, based on homework 2, we must have $f(Y_1)$ and $g(Y_3)$ are independent.).

Remark 1. Some students argue that Y_1 and Y_3 are dependent by showing $\mathsf{E}[Y_3 \mid Y_1] \neq \mathsf{E}[Y_3]$. This reasoning is correct. However, the calculation of $\mathsf{E}[Y_3 \mid Y_1]$ is not very straightforward.

$$\begin{aligned}
\mathsf{E}[Y_3] &= \mathsf{E}[Z^2 + X_3] = \mathsf{E}[Z^2] + \mathsf{E}[X_3] = 1 + 0 = 1 \\
\mathsf{E}[Y_3 \mid Y_1] &= \mathsf{E}[Z^2 + X_3 \mid Y_1] = \mathsf{E}[Z^2 \mid Y_1] + \underbrace{\mathsf{E}[X_3 \mid Y_1]}_{\text{(I)}} = \mathsf{E}[Z^2 \mid Y_1]
\end{aligned}$$

Part (I) is 0 as Z, X_1, X_3 are mutually independent.

$$\mathsf{E}[Z^2 \mid Y_1 = y] = \mathsf{E}[Z^2 \mid Z + X_1 = y]$$

Since $Z, X_1 \overset{\text{iid}}{\sim} \mathsf{N}(0, 1)$, we have (by change of variable)

$$(Z, Z + X_1)^\top \sim \mathsf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\right)$$

Therefore, $Z \mid Z + X_1 = y \sim \mathsf{N}(y/2, 1/2)$ and hence $\mathsf{E}[Z^2 \mid Z + X_1 = y] = y^2/4 + 1/2$. We have $\mathsf{E}[Y_3 \mid Y_1] = Y_1^2/4 + 1/2 \neq \mathsf{E}[Y_3]$.

Remark 2. Some students argue that Y_1 and Y_3 are dependent because both are functions of Z. This is a wrong reasoning. A simple counterexample is as follows. Suppose $Z_1, Z_2 \overset{\text{iid}}{\sim} \mathsf{N}(0,1)$. Let $Y_1 = Z_1 + Z_2$ and $Y_2 = Z_1 - Z_2$. Again by change of variable, one can show that Y_1 and Y_2 are independent although Y_1 and Y_2 are both functions of Z_1 and Z_2.

Exercise 5

(a)

$$\begin{aligned}
\mathsf{E}[\mathsf{E}[f(Y) \mid X]] &= \int \int f(y) p_{Y|X}(y \mid x) \mathrm{d}y p_X(x) \mathrm{d}x \\
&= \int f(y) \int p_{X,Y}(x,y) \mathrm{d}x \mathrm{d}y \\
&= \int f(y) p_Y(y) \mathrm{d}y = \mathsf{E}[f(Y)]
\end{aligned}$$

(b)

$$\begin{aligned}
\mathsf{E}[f(X)g(X,Y) \mid X] &= \int f(x) g(x,y) p_{Y|X}(y \mid x) \mathrm{d}y \\
&= f(x) \int g(x,y) p_{Y|X}(y \mid x) \mathrm{d}y \\
&= f(X) \mathsf{E}[g(X,Y) \mid X]
\end{aligned}$$