STA 532 HW3 Due Wed 2/5 Yi Mi, ym154

1. (a) let $f(y) = y^{p/q}$, $1 \le q < p$

$$\frac{\partial^2 f(y)}{\partial y^2} = \frac{p}{q} \times \left(\frac{p}{q} - 1\right) \times y^{\frac{p}{q}-2} > 0 \quad \text{for } y \in (0, +\infty)$$

so $f(y)$ is strictly convex

Jensen: $E[f(y)] \geq f[E(y)]$ for convex $f(y)$

$$\Rightarrow E[(Y^q)^{p/q}] \geq [E(Y^q)]^{p/q}$$

$$E[Y^p] \geq [E(Y^q)]^{p/q}$$

$$E(Y^p)^{1/p} \geq E(Y^q)^{1/q}$$

(b) since $g(Y) = 1/Y$ is strictly convex in $Y \in (0, +\infty)$

$$E[1/Y] \geq 1/E[Y]$$

(c) since $h(Y) = \log Y$ is strictly concave in $Y \in (0, +\infty)$

$$E[\log Y] \leq \log E[Y]$$

2. (a) $(w_1, w_2, w_3) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3)$

The pdf for $\underset{\sim}{w} = (w_1, w_2, w_3)$ is

$$p(\underset{\sim}{w}) = \frac{1}{B(\alpha)} \prod_{i=1}^{3} w_i^{\alpha_i - 1},$$

where $B(\alpha) = \dfrac{\prod_{i=1}^{3} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{3} \alpha_i)}$

Let $\alpha_0 = \alpha_1 + \alpha_2 + \alpha_3$,

$$p(\underset{\sim}{w}) = \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} \prod_{i=1}^{3} w_i^{\alpha_i - 1}$$

Mean: $E(w_1) = \iiint w_1 \, p(\underset{\sim}{w}) \, dw_1 dw_2 dw_3$

$$= \iiint w_1 \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} w_1^{\alpha_1 - 1} w_2^{\alpha_2 - 1} w_3^{\alpha_3 - 1} \, dw_1 dw_2 dw_3$$

with kernel: $\iiint \dfrac{\Gamma(\alpha_0 + 1)}{\Gamma(\alpha_1 + 1)\Gamma(\alpha_2)\Gamma(\alpha_3)} w_1^{\alpha_1} w_2^{\alpha_2 - 1} w_3^{\alpha_3 - 1} \, dw_1 dw_2 dw_3 = 1$

we have:

$$E(w_1) = \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} \times \frac{\Gamma(\alpha_1 + 1)\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_0 + 1)}$$

$$= \frac{\alpha_1}{\alpha_0}$$

Therefore: $E(w_i) = \dfrac{\alpha_i}{\alpha_0}$ for $i \in \{1, 2, 3\}$

Variance: $Var(w_i) = E(w_i^2) - E(w_i)^2$

Similarly, $E(w_1^2) = \iiint w_1^2 \, p(\underset{\sim}{w}) \, dw_1 dw_2 dw_3$

$$= \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} \times \frac{\Gamma(\alpha_1+2)\Gamma(\alpha_2)\Gamma(\alpha_3)}{\Gamma(\alpha_0+2)}$$

$$= \frac{(\alpha_1+1)\alpha_1}{(\alpha_0+1)\alpha_0}$$

Thus, $E(w_i^2) = \dfrac{(\alpha_i+1)\alpha_i}{(\alpha_0+1)\alpha_0}$

And, $Var(w_i) = \dfrac{(\alpha_i+1)\alpha_i}{(\alpha_0+1)\alpha_0} - \left(\dfrac{\alpha_i}{\alpha_0}\right)^2$

$$= \frac{(\alpha_0-\alpha_i)\alpha_i}{(\alpha_0+1)\alpha_0^2}$$

(b). $cov(w_1, w_2) = E[(w_1 - E(w_1))(w_2 - E(w_2))]$

$= E(w_1 w_2) - E(w_1)E(w_2) - E(w_1)E(w_2) + E(w_1)E(w_2)$

$= E(w_1 w_2) - E(w_1)E(w_2)$

Similarly, $E(w_1 w_2) = \iiint w_1 w_2 \, p(\underset{\sim}{w}) \, dw_1 dw_2 dw_3$

$$= \frac{\Gamma(\alpha_0)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} \times \frac{\Gamma(\alpha_1+1)\Gamma(\alpha_2+1)\Gamma(\alpha_3)}{\Gamma(\alpha_0+2)}$$

$$= \frac{\alpha_1\alpha_2}{(\alpha_0+1)\alpha_0}$$

Thus, $E(w_i w_j) = \dfrac{\alpha_i\alpha_j}{(\alpha_0+1)\alpha_0}$

$cov(w_i, w_j) = \dfrac{\alpha_i\alpha_j}{(\alpha_0+1)\alpha_0} - \dfrac{\alpha_i}{\alpha_0}\dfrac{\alpha_j}{\alpha_0} = -\dfrac{\alpha_i\alpha_j}{\alpha_0^2(\alpha_0+1)} \quad i \neq j$

explain the sign: $\sum_{i=1}^{n} w_i = 1$, when w_i increases, w_j may decrease, so their relationship is negative.

(c) $Var(w_1+w_2) = E(w_1+w_2)^2 - [E(w_1+w_2)]^2$

$= E(w_1^2 + 2w_1w_2 + w_2^2) - [(Ew_1)^2 + 2Ew_1Ew_2 + (Ew_2)^2]$

$= \frac{(\alpha_1+1)\alpha_1}{(\alpha_0+1)\alpha_0} + 2\frac{\alpha_1\alpha_2}{(\alpha_0+1)\alpha_0} + \frac{(\alpha_2+1)\alpha_2}{(\alpha_0+1)\alpha_0} - \left[\frac{\alpha_1^2}{\alpha_0^2} + 2\frac{\alpha_1\alpha_2}{\alpha_0\alpha_0} + \frac{\alpha_2^2}{\alpha_0^2}\right]$

$= \frac{(\alpha_1+\alpha_2)^2 + \alpha_1 + \alpha_2}{(\alpha_0+1)\alpha_0} - \frac{(\alpha_1+\alpha_2)^2}{\alpha_0^2}$

$Cov(w_1, w_1+w_2) = E[(w_1 - E(w_1))(w_1+w_2 - E(w_1+w_2))]$

$= E[w_1(w_1+w_2)] - E(w_1+w_2)E(w_1) - E(w_1)E(w_1+w_2) + E(w_1)E(w_1+w_2)$

$= E(w_1^2) + E(w_1w_2) - E(w_1)[E(w_1) + E(w_2)]$

$= \frac{(\alpha_1+1)\alpha_1}{(\alpha_0+1)\alpha_0} + \frac{\alpha_1\alpha_2}{(\alpha_0+1)\alpha_0} - \frac{\alpha_1}{\alpha_0}\left(\frac{\alpha_1}{\alpha_0} + \frac{\alpha_2}{\alpha_0}\right)$

$= \frac{\alpha_1(\alpha_0 + \alpha_1 - \alpha_2)}{(\alpha_0+1)\alpha_0^2}$

(d) $p(\underset{\sim}{w}) = \dfrac{\Gamma(\alpha_0)}{\prod_{i=1}^{3}\Gamma(\alpha_i)} \prod_{i=1}^{3} w_i^{\alpha_i - 1}$

let $W = \dfrac{w_2}{1-w_1}$

We have $w_3 = 1 - w_1 - w_2 = (1-w_1)(1-W)$

thus $p(w_1, W) = \dfrac{\Gamma(\alpha_0)}{\prod_{i=1}^{3}\Gamma(\alpha_i)} w_1^{\alpha_1-1}\left[W(1-w_1)\right]^{\alpha_2-1}\left[(1-w_1)(1-W)\right]^{\alpha_3-1}$

and $p(w_1) = \int p(w_1, W)\, dW$

$= \dfrac{\Gamma(\alpha_0)}{\prod_{i=1}^{3}\Gamma(\alpha_i)} w_1^{\alpha_1-1}(1-w_1)^{\alpha_2+\alpha_3-2} \int W^{\alpha_2-1}(1-W)^{\alpha_3-1}\, dW$

$= w_1^{\alpha_1-1}(1-w_1)^{\alpha_2+\alpha_3-1} \dfrac{\Gamma(\alpha_0)}{\Gamma(\alpha_2+\alpha_3)\Gamma(\alpha_1)} \int \dfrac{\Gamma(\alpha_2+\alpha_3)}{\Gamma(\alpha_2)\Gamma(\alpha_3)} W^{\alpha_2-1}(1-W)^{\alpha_3-1}\, dW$

$\Rightarrow w_1 \sim \mathrm{Beta}(\alpha_1, \alpha_2+\alpha_3)$

similarly, $w_3 \sim \mathrm{Beta}(\alpha_3, \alpha_1+\alpha_2)$

and, $w_1 + w_2 = 1 - w_3$

thus, $p(w_1+w_2) = P_{w_3}(1-(w_1+w_2)) \left|\dfrac{d(1-w_3)}{d w_3}\right|$

$= \left[1-(w_1+w_2)\right]^{\alpha_3-1}(w_1+w_2)^{\alpha_1+\alpha_2-1}\dfrac{\Gamma(\alpha_0)}{\Gamma(\alpha_1+\alpha_2)\Gamma(\alpha_3)}$

$= \mathrm{Beta}(\alpha_1+\alpha_2, \alpha_3)$

3. (a) If X and Y are independent,

$$P(X,Y) = P(X)\, P(Y)$$

$$E(XY) = \iint xY\, P(xY)\, dx\, dY.$$

$$= \iint xY\, P(X)\, P(Y)\, dx\, dY$$

$$= \int x\, P(X)\, dx \int Y\, P(Y)\, dY$$

$$= E(X)\, E(Y)$$

$$\mathrm{cov}(X,Y) = E[(X-E(X))(Y-E(Y))]$$

$$= E(XY) - E(X)E(Y)$$

$$= 0$$

(b) $$\mathrm{Cov}(X,Y) = \mathrm{cov}(a+bY, Y)$$

$$= E[(a+bY-E(a+bY))(Y-E(Y))]$$

$$= b\,E[Y-E(Y)]^2$$

$$= b\,\mathrm{Var}(Y)$$

$$\mathrm{Var}(X) = E[a+bY-E(a+bY)]^2$$

$$= b^2 E(Y-E(Y))^2$$

$$= b^2\,\mathrm{Var}(Y)$$

$$\mathrm{cor}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{b\,\mathrm{Var}(Y)}{\sqrt{b^2\,\mathrm{Var}^2(Y)}} = \pm 1$$

(c) i. $\text{cov}(a_1+b_1X_1,\ a_2+b_2X_2)$

$$= E[(a_1+b_1X_1-E(a_1+b_1X_1))(a_2+b_2X_2-E(a_2+b_2X_2))]$$

$$= b_1b_2\ E[(X_1-EX_1)(X_2-EX_2)]$$

$$= b_1b_2\ \text{cov}(X_1,X_2)$$

ii. $E(X_1+X_2+X_3) = \iiint (X_1+X_2+X_3)\,p(X_1,X_2,X_3)\,dX_1\,dX_2\,dX_3$

$$= \sum_{i=1}^{3}\iiint X_i\,p(X_1,X_2,X_3)\,dX_1\,dX_2\,dX_3$$

$$= E(X_1)+E(X_2)+E(X_3)$$

$\text{Var}(X_1+X_2+X_3) = E(X_1+X_2+X_3-E(X_1+X_2+X_3))^2$

$$= E(X_1+X_2+X_3)^2-(E(X_1+X_2+X_3))^2$$

$$= E(X_1^2+X_2^2+X_3^2+2X_1X_2+2X_1X_3+2X_2X_3)$$

$$-(E^2X_1+E^2X_2+E^3X_3+2EX_1EX_2+2EX_1EX_3+2EX_2EX_3)$$

$$= \sum_{i=1}^{3}\text{Var}(X_i)+2\sum_{i\neq j}\text{Cov}(X_i,X_j)$$

$$= \sum_{j}^{3}\sum_{i}^{3}\text{Cov}(X_i,X_j)$$

Yes, it matches the result from class.

4. (a) Since $Z, X_1, X_2, X_3 \overset{iid}{\sim} N(0,1)$

$$E(Y_1) = E(Z + X_1) = 0$$

$$E(Y_2) = E(Z + X_2) = 0$$

$$E(Y_3) = E(Z^2 + X_3) = Var(Z) + (EZ)^2 + E(X_3) = 1$$

$$Var(Y_1) = EY_1^2 - (EY_1)^2 = EY_1^2 = E(Z^2 + 2ZX_1 + X_1^2)$$

$$= EZ_1^2 + 2EZ\,EX_1 + EX_1^2$$

$$= 2$$

Similarly, $Var(Y_2) = 2$

$$Var(Y_3) = EY_3^2 - (EY_3)^2 = EY_3^2 - 1 = E(Z^4 + 2X_3Z^2 + X_3^2) - 1$$

$$= EZ^4 + 2EX_3\,EZ^2 + EX_3^2 - 1$$

$$= 3 + 1 - 1$$

$$= 3$$

(b)
$$cov(\underset{\sim}{Y}) = \begin{bmatrix} V(Y_1) & c(Y_1,Y_2) & c(Y_1,Y_3) \\ c(Y_2,Y_1) & V(Y_2) & c(Y_2,Y_3) \\ c(Y_3,Y_1) & c(Y_3,Y_2) & V(Y_3) \end{bmatrix}$$

$$c(Y_1,Y_2) = EY_1Y_2 - EY_1EY_2 = E(Z+X_1)(Z+X_2)$$

$$= EZ^2 + EZ\,EX_1 + EZ\,EX_2 + EX_1EX_2$$

$$= 1$$

$$c(Y_1,Y_3) = EY_1Y_3 - EY_1EY_3 = E(Z+X_1)(Z^2+X_3)$$

$$= EZ^3 + EZ\,EX_3 + EX_1Z^2 + EX_1EX_3$$

since $X_1 \perp Z$, $X_2 \perp Z$

$EX_1Z = EX_1EZ$

$EX_2Z = EX_2EZ$

$EX_1X_2 = EX_1EX_2$

$$E X_1 Z^2 = \iint x_1 z^2 p(x_1, z)\, dx\, dz$$
$$= \int x_1 p(x)\, dx \int z^2 p(z)\, dz$$
$$= E X_1 \, E Z^2$$

thus, $c(Y_1, Y_3) = E Z^3 = 0$

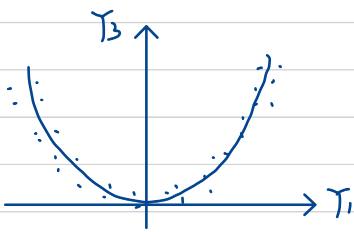
Similarly, $c(Y_2, Y_3) = 0$

$$\mathrm{Cov}(\underset{\sim}{Y}) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

(c) $c(Y_1, Y_3) = 0$, indicating that Y_1 and Y_3 have no linear relationship, however, it does not mean that Y_1 and Y_3 are independent.

$Y_1 = Z + X_1$, $Y_3 = Z^2 + X_3$, they both refer to the same randomvariable Z

We can see X_1 and X_3 as noise, and thus the sketch of the relationship between Y_1 and Y_3 is like the randomly scattered points around the line $f(z) = z^2$, resulting from the noise X_1 and X_3.



Therefore, Y_3 and Y_1 are not independent.

5. (a)

$$
\begin{aligned}
E[f(Y)] &= \int f(y)\, p(y)\, dy \\
&= \int_y f(y) \int_x p(x, y)\, dx\, dy \\
&= \int_y f(y) \int_x p(y \mid x)\, p(x)\, dx\, dy \\
&= \int_x \int_y f(y)\, p(y \mid x)\, dy\, p(x)\, dx \\
&= \int_x E(f(y) \mid X)\, p(x)\, dx \\
&= E[E(f(y) \mid X)]
\end{aligned}
$$

(b)

$$
\begin{aligned}
E[f(X)\, g(X, Y) \mid X] &= \int f(X)\, g(X, Y)\, p(Y \mid X)\, dY \\
&= f(X) \int g(X, Y)\, p(Y \mid X)\, dY \\
&= f(X)\, E(g(X, Y) \mid X)
\end{aligned}
$$