# STAC62F : 2016 Assignment 3 - Solutions

1. $E(\underset{\sim}{a} + B\underset{\sim}{x}) = \left( E\left(a_i + \sum_{j=1}^{k} b_{ij} X_j\right) \right)$

$= \left(a_i + \sum_{j=1}^{k} b_{ij} E(x_j)\right)$ since $E(a+bX) = a + bE(X)$ for r.v.'s

$= \underset{\sim}{a} + B E(\underset{\sim}{x})$

2. $E(AXB + C) = \left( E\left( \sum_{i_1=1}^{m} a_{i i_1} \sum_{i_2=1}^{n} x_{i_1 i_2} b_{i_2 j} + c_{ij} \right) \right)$

$$= \left( \sum_{i_1=1}^{m} \sum_{i_2=1}^{n} a_{i i_1} E(x_{i_1 i_2}) b_{i_2 j} + c_{ij} \right)$$

$$= A E(X) B + C.$$

3. Put $X = (\underset{\sim}{x}_1 \cdots \underset{\sim}{x}_n)'$ $\underset{\sim}{1} = (1, \ldots, 1)' \in R^n$

(a) $E(\bar{\underset{\sim}{x}}) = E\left(\frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{x}_i\right) = E\left(\left(\frac{1}{n}\right) X' \underset{\sim}{1}'\right)$

$\overset{②}{=} \frac{1}{n} E(X') \underset{\sim}{1}' = \frac{1}{n} (\underset{\sim}{\mu}\ \underset{\sim}{\mu} \cdots \underset{\sim}{\mu}) \underset{\sim}{1}' = \frac{1}{n} \sum_{i=1}^{n} \underset{\sim}{\mu} = \underset{\sim}{\mu}.$

(b) $\mathrm{Var}(\bar{\underset{\sim}{x}}) = E\left( (\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})(\bar{\underset{\sim}{x}} - \underset{\sim}{\mu})' \right)$

$$= E\left( \left(\frac{1}{n} \sum_{i=1}^{n} (\underset{\sim}{x}_i - \underset{\sim}{\mu})\right) \left(\frac{1}{n} \sum_{j=1}^{n} (\underset{\sim}{x}_j - \underset{\sim}{\mu})\right)' \right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} E\left( (\underset{\sim}{x}_i - \underset{\sim}{\mu})(\underset{\sim}{x}_j - \underset{\sim}{\mu})' \right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} E\left( (\underset{\sim}{x}_i - \underset{\sim}{\mu})(\underset{\sim}{x}_i - \underset{\sim}{\mu})' \right)$$

since $E\left((\underset{\sim}{x}_i - \underset{\sim}{\mu})(\underset{\sim}{x}_j - \underset{\sim}{\mu})'\right)$ when $i \neq j$

$= E(\underset{\sim}{x}_i - \underset{\sim}{\mu}) E(\underset{\sim}{x}_j - \underset{\sim}{\mu})'$ $\underset{\sim}{x}_i, \underset{\sim}{x}_j$ are ind.

$= 0$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}(\underset{\sim}{x}_i) = \frac{1}{n} \Sigma.$$

(c) $E(S) = \frac{1}{n-1} E\left( (X - \underset{\sim}{1}_n \bar{\underset{\sim}{x}}')' (X - \underset{\sim}{1}_n \bar{\underset{\sim}{x}}') \right)$

$= \frac{1}{n-1}\left[ E(X'X) - E(X' \underset{\sim}{1}_n \bar{\underset{\sim}{x}}') - E(\bar{\underset{\sim}{x}} \underset{\sim}{1}' X) + E(\bar{\underset{\sim}{x}} \underset{\sim}{1}_n' \underset{\sim}{1}_n \bar{\underset{\sim}{x}}') \right]$

$= \frac{1}{n-1}\left[ E(X'X) - nE(\bar{\underset{\sim}{x}} \bar{\underset{\sim}{x}}') - nE(\bar{\underset{\sim}{x}} \bar{\underset{\sim}{x}}') + nE(\bar{\underset{\sim}{x}} \bar{\underset{\sim}{x}}') \right]$

$= \frac{1}{n-1}\left[ \sum_{i=1}^{n} E(\underset{\sim}{x}_i \underset{\sim}{x}_i') - nE(\bar{\underset{\sim}{x}} \bar{\underset{\sim}{x}}') \right]$

and note $\mathrm{Var}(\underset{\sim}{x}) = E\left( (\underset{\sim}{x} - E(\underset{\sim}{x}))(\underset{\sim}{x} - E(\underset{\sim}{x}))' \right)$

$= E(\underset{\sim}{x} \underset{\sim}{x}') - E(\underset{\sim}{x}) E(\underset{\sim}{x}') = E(\underset{\sim}{x} \underset{\sim}{x}') - E(\underset{\sim}{x})(E(\underset{\sim}{x}))'$

$= \frac{1}{n-1}\left[ \sum_{i=1}^{n} (\Sigma + \underset{\sim}{\mu} \underset{\sim}{\mu}') - n\left( \frac{1}{n}\Sigma + \underset{\sim}{\mu} \underset{\sim}{\mu}' \right) \right]$

$= \frac{1}{n-1}\left[ n\Sigma + n\underset{\sim}{\mu} \underset{\sim}{\mu}' - \Sigma - n\underset{\sim}{\mu} \underset{\sim}{\mu}' \right] = \Sigma.$

4. First way: $E(X^2) = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

(put $u = x$, $dv = x e^{-\frac{x^2}{2}}$ so $du = 1$, $v = -e^{-x^2/2}$)

$= \frac{-x e^{-x^2/2}}{\sqrt{2\pi}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

$= 0 + 1 = 1$

Second way $F_{X^2}(x) = P(X^2 \le x) = 0$ when $x < 0$ and so suppose $x > 0$. Then $F_{X^2}(x) =$ $P(-\sqrt{x} \le X \le \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$ where $\Phi$ is the $N(0,1)$ cdf. Therefore

$F_{X^2}(x) = 2\Phi(\sqrt{x}) - 1$ and

$$f_X(x) = \frac{dF_{X^2}(x)}{dx} = 2\varphi(\sqrt{x})\,\frac{1}{2}x^{-\frac{1}{2}} \quad x>0$$

$$= \frac{1}{\sqrt{2\pi}}\, x^{-\frac{1}{2}} e^{-x/2} = \frac{1}{\Gamma(1/2)}\left(\frac{x}{2}\right)^{\frac{1}{2}-1} e^{-\frac{x}{2}}\, \frac{1}{2}.$$

which we recognize as the density of a chi-squared (1) distribution.

Then $E(X^2) = \int_0^\infty x\, f_X(x)\,dx$

$$= \frac{2}{\Gamma(\frac{1}{2})}\int_0^\infty \left(\frac{x}{2}\right)\left(\frac{x}{2}\right)^{\frac{1}{2}-1} e^{-\frac{x}{2}}\,\frac{1}{2}\,dx$$

put $y = \frac{x}{2}$, $dy = \frac{1}{2}dx$

$$= \frac{2}{\Gamma(\frac{1}{2})}\int_0^\infty y^{3/2-1} e^{-y}\,dy$$

$$= 2\,\Gamma(3/2)/\Gamma(\tfrac{1}{2}) = 2\,\tfrac{1}{2}\,\Gamma(\tfrac{1}{2})/\Gamma(\tfrac{1}{2}) = 1$$

using $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.

(5) We first need to show that each joint distribution $(X_{t_1}, \ldots, X_{t_k})$ is properly defined and this entails showing that $\Sigma(t_1, \ldots, t_k)$ is a variance and, from results in class we know that this means that $\Sigma(t_1, \ldots, t_k)$ is positive semidefinite.

Let $\underset{\sim}{a} \in \mathbb{R}^k$ then

$$\underset{\sim}{a}' \, \Sigma(t_1, \ldots, t_k) \, \underset{\sim}{a} = \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j \, \sigma(t_i, t_j)$$

$$= v^2 \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j \exp\left\{ -\frac{(t_i - t_j)^2}{2\ell^2} \right\}$$

$$= v^2 \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j \, E\left( \exp\left\{ \sqrt{-1} \, \frac{(t_i - t_j)}{\ell} Z \right) \right.$$

where $Z \sim N(0,1)$

$$= v^2 E\left( \sum_{i=1}^{k} \sum_{j=1}^{k} a_i a_j \exp\left\{ \sqrt{-1} \, \frac{(t_i - t_j)}{\ell} Z \right\} \right)$$

$$= v^2 E\left( \sum_{i=1}^{k} a_i \exp\left\{ \sqrt{-1} \, \frac{t_i}{\ell} Z \right\} a_j \exp\left\{ -\sqrt{-1} \, \frac{t_j}{\ell} Z \right\} \right)$$

$$= v^2 E\left( \left( \sum_{i=1}^{k} a_i \exp\left\{ \sqrt{-1} \, \frac{t_i}{\ell} Z \right\} \right) \overline{\left( \sum_{i=1}^{k} a_i \exp\left\{ \sqrt{-1} \, \frac{t_i}{\ell} Z \right\} \right)} \right)$$

$$= v^2 E\left( \left| \sum_{i=1}^{k} a_i \exp\left\{ \sqrt{-1} \, \frac{t_i}{\ell} Z \right\} \right|^2 \right) \geq 0$$

since $| \cdot |^2 \geq 0$.

Therefore $\Sigma(t_1, \ldots, t_k)$ is a variance matrix.

Next we have to show that these joint distributions satisfy the consistency requirements so we can apply KCT. In class we proved the following result: if $\underset{\sim}{X} \sim N_n(\underline{\mu}, \Sigma)$ and $\underset{\sim}{Y} = (X_{i_1}, \ldots, X_{i_s})$ where $s \le t$ and $i_1, \ldots, i_s$ are distinct then $\underset{\sim}{Y} \sim N((\mu_{i_1}, \ldots, \mu_{i_s})', (\sigma_{i_1 i_2})_{1 \le i_1 \le s, 1 \le i_2 \le s})$.

Now observe that $(X_{t_1}, \ldots, X_{t_n}) \overset{\text{def}}{\sim} N_n((m(t_1), \ldots, m(t_n)), \Sigma(t_1, \ldots, t_n))$ and the above implies that $(X_{t_{i_1}}, \ldots, X_{t_{i_s}}) \sim N((m(t_{i_1}), \ldots, m(t_{i_s}))', \Sigma(t_{i_1}, \ldots, t_{i_s}))$ which corresponds to the definition. Therefore KCT applies and the stochastic process is properly defined.

6. 3.2.1 $P(X=1, Z=1) = P(X=1, Y=1)$ $= P(X=1)P(Y=1) = P(X=1)P(Z=1)$ since $P(Z=1) = P(X=1, Y=1) + P(X=-1, Y=-1)$ $= 1/4 + 1/4 = 1/2 = P(Y=1)$. Similarly $P(X=a, Z=b) = P(X=a)P(Z=b)$, $P(Y=a, Z=b) = P(Y=a)P(Z=b)$ for all choices of $a$ and $b$. Since $X$ and $Y$ are defined to be independent we have that $X$, $Y$ and $Z$ are pairwise independent.

But $0 = P(X=1, Y=1, Z=-1) \neq 1/8$ $= P(X=1)P(Y=1)P(Z=-1)$ and so $X$, $Y$ and $Z$ are not mutually statistically independent.

3.5.4 $P(X=n) = P\left(\begin{array}{c}\text{last of } m \text{ marked} \\ \text{animals was } n\text{th sampled}\end{array}\right)$

$$= P\left(\begin{array}{c}\text{1st } n-1 \text{ animals} \\ \text{contains exactly} \\ m-1 \text{ marked}\end{array}\right) P\left(n \mid \ \right)$$

$$= \frac{\binom{a}{m-1}\binom{b-a}{n-m}}{\binom{b}{n-1}} \frac{a-m+1}{b-n+1}$$

$$= \frac{\frac{a!}{(a-m+1)!\,(m-1)!}\binom{b-a}{n-m}}{\frac{b!}{(b-n+1)!\,(n-1)!}} \frac{a-m+1}{b-n+1}$$

$$= \frac{a}{b}\binom{a-1}{m-1}\binom{b-a}{n-m} \Big/ \binom{b-1}{n-1}$$

Now $P(X=n) = 0$ when $n < m$ and $P(X=n) = 0$ when $n > b-a+m$. So

$$E(X) = \sum_{i=m}^{b-a+m} i \,\frac{a}{b}\binom{a-1}{m-1}\binom{b-a}{i-m} \Big/ \binom{b-1}{i-1}$$

$$= \frac{a}{b}\binom{a-1}{m-1}\sum_{i=m}^{b-a+m} i \binom{b-a}{i-m} \Big/ \binom{b-1}{i-1}$$

$$= m \sum_{i=m}^{b-a+m} \binom{a}{m}\binom{b-a}{i-m} \Big/ \binom{b}{i}$$

It is "difficult" to sum this so we try another approach let $X_i$ = # of unmarked animals sampled between the $(i-1)$st and the $i$-th marked animal. Then $X = X_1 + \cdots + X_m + m$ and $E(X) = \sum_{i=1}^{m} E(X_i) + m$ and since the $X_i$ are symmetrically distributed $E(X_i) = E(X_1)$ $\forall i$ so $E(X) = m(E(X_1)+1)$ and we only need $E(X_1)$

Now consider the case when $m = a$ and let $X_{a+1}$ be the remaining unmarked animals after all $a$ have been obtained. Again all the $X_i$ have the same distribution and clearly $X_1 + \cdots + X_a + X_{a+1} = b - a$ so $(a+1)E(X_1) = b - a$ and $E(X_1) = (b-a)/(a+1)$. This implies $E(X) = m\left(\frac{b-a}{a+1} + 1\right) = m(b+1)/(a+1)$

8. 3.6.5 (a) Note that log is a concave function and so $-\log$ is convex. Therefore, whenever $X \geq 0$ a.s. and $E(\log X)$ exists we have $E(-\log X) \geq -\log E(X)$ by Jensen.

When $g$ is a probability function then $0 \leq g(x) \leq 1$ and so $-\infty \leq \log g(x) \leq 0$ which implies $E(-\log g(X))$ exists.

Therefore, when $X$ is discrete and $g$ is a probability function we have.

$E(\log(f_X(X)/g(X))) = E(-\log g(X)/f(X))$
$\geq -\log E(g(X)/f_X(X)) = -\log 1 = 0$.
and so $E(\log f_X(X)) \geq E \log g(X)$ and the result follows by taking $g = f_Y$.

Also we get equality (by Jensen) iff $\log g(X)/f_X(X) = a + b(g(X)/f(X))$ a.s. for some constants $a$ and $b$ and this occurs iff $g(X) = c f_X(X)$ for some constant $c$. But then $1 = \sum_x g(x) = c \sum_x f_X(x) = c$ and so equality occurs iff $g = f_X$.

**Note** (1) the result $E(\log(f_X(X)/g(X))) \geq 0$ with equality iff $f_X = g$ holds quite generally for densities $f_X$ and $g$ and $KL(f_X \| g) = E(\log(f_X(X)/g(X)))$ is called the Kullback-Liebler divergence. It is not necessary either that $X$ be a random variable.

(b) Take $g(x,y) = f_X(x) f_Y(y)$ and note that

$$I = E\left( \log\left( \frac{f_{X,Y}(X,Y)}{f_X(X) f_Y(Y)} \right) \right)$$

$$= KL(f_{X,Y} \| g) \geq 0$$

with equality iff $f_{X,Y} = g$ or equivalently iff $X$ and $Y$ are statistically independent.

9) 3.7.4 $\mathrm{Var}(Y|X) \overset{\text{def}}{=} E((Y-E(Y|X))^2|X)$
and so $\mathrm{Var}(Y|X) : (\mathbb{R}^1, \mathcal{B}^1) \to (\mathbb{R}^1, \mathcal{B}^1)$
is the r.v. satisfying $E(h(X)\,\mathrm{Var}(Y|X)(X))$
$= E(h(X)(Y-E(Y|X))^2)$ for every
$h : (\mathbb{R}^1, \mathcal{B}^1) \to (\mathbb{R}^1, \mathcal{B}^1)$ for which both
expectations exist.

Now $\mathrm{Var}(Y) = E((Y-E(Y))^2)$

$$= E((Y-E(Y|X)+E(Y|X)-E(Y))^2)$$

$$= E((Y-E(Y|X))^2) + 2E((Y-E(Y|X))(E(Y|X)-E(Y))) + E((E(Y|X)-E(Y))^2)$$

$$\overset{\text{TTE}}{=} E(E((Y-E(Y|X))^2|X))$$

$$+ 2E(E((Y-E(Y|X))(E(Y|X)-E(Y))|X))$$

$$+ \mathrm{Var}(E(Y|X)) \quad (\text{as } E(E(Y|X)) \overset{\text{TTE}}{=} E(Y))$$

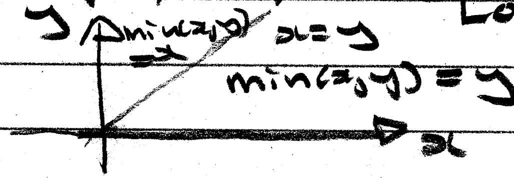
$$= E(\mathrm{Var}(Y|X)) + \mathrm{Var}(E(Y|X))$$

since $= E(E(Y-E(Y|X)|X)(E(Y|X)-E(Y)))$
by result in class

$= 0$ since $E(Y) - E(E(Y|X)) = 0$

10 4.5.4 $X, Y \overset{iid}{\sim} U(0,1)$. Then

$$E(U) = E(\min(X,Y)) = \int_{[0,1]^2} \min(x,y)\,dx\,dy$$




$$= \int_0^1 \int_0^x y\,dy\,dx + \int_0^1 \int_0^y x\,dx\,dy.$$

$$= \int_0^1 \frac{x^2}{2}\,dx + \int_0^1 \frac{y^2}{2}\,dy$$

$$= \frac{x^3}{6}\Big|_0^1 + \frac{y^3}{6}\Big|_0^1 = \frac{1}{3}$$

Since $X+Y = U+V$ then $E(V) = E(X) + E(Y) - E(U) = \frac{1}{2} + \frac{1}{2} - \frac{1}{3} = \frac{2}{3}$, $cov(U,V) = E(UV) - E(U)E(V) = E(UV) - \frac{1}{3}\frac{2}{3}$ and

$$E(UV) = E(\min(X,Y)\max(X,Y))$$

$$= E(XY) \overset{ind}{=} E(X)E(Y) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

since $E(X) = E(Y) = \frac{1}{2}$.

Therefore $cov(U,V) = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}$