STA 431 Assignment One

① (a) F (b) T (c) T

② (a and b)

	$x=1$	$x=2$	$x=3$	
$y=1$	$3/12$	$\frac{1}{12}$	$3/12$	$7/12$
$y=2$	$\frac{1}{12}$	$3/12$	$\frac{1}{12}$	$5/12$
	$\frac{4}{12}$	$\frac{4}{12}$	$\frac{4}{12}$	1

(c) $E(X) = \sum_x x\,P(x) = (1)(\frac{1}{3}) + (2)(\frac{1}{3}) + 3(\frac{1}{3}) = \frac{6}{3} = 2$

(d) $E(Y) = \sum_y y\,P(y) = (1)\frac{7}{12} + (2)\frac{5}{12} = \frac{17}{12}$

(e)

	$x=1$	$x=2$	$x=3$
$y=1$	$xy=1$ $(3/12)$	$xy=2$ $(1/12)$	$xy=3$ $(3/12)$
$y=2$	$xy=2$ $(1/12)$	$xy=4$ $(3/12)$	$xy=6$ $(1/12)$

z	1	2	3	4	6
$P(z)$	$3/12$	$2/12$	$3/12$	$3/12$	$1/12$

(f) $E(Z) = \sum_z z\,P(z) = (1)\frac{3}{12} + (2)\frac{2}{12} + (3)\frac{3}{12} + (4)\frac{3}{12} + (6)\frac{1}{12}$

$= (3+4+9+12+6)/12 = \frac{34}{12} = \frac{17}{6}$

(2g) Yes $E(XY) = \frac{17}{6} = (2)(\frac{17}{12}) = E(X)E(Y)$

(h) $Cov(X, Y) = 0$

(i) No. $P(X=1, Y=1) = \frac{1}{4} \neq P(X=1)P(Y=1) = (\frac{4}{12})(\frac{7}{12})$

$0.25 = \frac{28}{144} = \frac{14}{72} = \frac{7}{36} = 0.1944$

This is where I use independence.

(3) $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_X(x) f_Y(y)\,dx\,dy$

$= \int_{-\infty}^{\infty} y f_Y(y) \left(\int_{-\infty}^{\infty} x\, f_X(x)\,dx\right) dy = \int_{-\infty}^{\infty} y f_Y(y) E(X)\,dy$

$= E(X)\int_{-\infty}^{\infty} y\, f_Y(y)\,dy = E(X)E(Y)$

Yes, if X and Y are discrete, replace the integral signs with summation signs.

This shows that independence implies zero co-variance. Problem one shows that you can have zero covariance without independence.

4 (a) $E(X) = \sum_x x\, p(x) = (a)(1) = a$

$Var(X) = \sum_x (x-\mu_x)^2 p(x) = (a-a)^2(1)$
$= 0$

(b) $E(Y) = \int_{-\infty}^{\infty} a\, f_X(x)\, dx = a\int_{-\infty}^{\infty} f_X(x)\, dx = a \cdot 1 = a$

$Var(Y) = \int_{-\infty}^{\infty} (a-a)^2 f_X(x)\, dx = 0$

5 (a) $Var(Y) = E\{(Y-\mu_y)^2\}$

$= E\{Y^2 - 2\mu_y Y + \mu_y^2\}$

$= E(Y^2) - 2\mu_y E(Y) + \mu_y^2$

$= E(Y^2) - 2\mu_y^2 + \mu_y^2$

$= E(Y^2) - \mu_y^2$

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

(5b) $Cov(X,Y) = E\{(X-\mu_x)(Y-\mu_y)\}$

$= E\{XY - X\mu_y - Y\mu_x + \mu_x\mu_y)$

$= E(XY) - E(X)\mu_y - E(Y)\mu_x + \mu_x\mu_y$

$= E(XY) - 2E(X)E(Y) + E(X)E(Y)$

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

(c) By problem 3, $E(XY) = E(X)E(Y)$ so by (5b), $Cov(X,Y) = 0$

6 (a) First, $E(aX) = aE(X)$, so

$$Var(aX) = E\{(aX - aE(X))^2\} = E\{a^2(X-E(X))^2\}$$
$$= a^2 E\{(X-E(X))^2\} = a^2 Var(X)$$

(b) First, $E(aX+b) = aE(X) + b$, so

$$Var(aX+b) = E\{(aX + \cancel{b} - aE(X) \cancel{-b})^2\}$$
$$= E\{(aX - aE(X))^2\} = Var(aX) \overset{\text{by (a)}}{=} a^2 Var(X)$$

(c) $Var(a) = 0$ by Problem 4.

(d) Noting $E(aX) = aE(X)$ and $E(bY) = bE(Y)$,

$$Cov(aX, bY) = E\{(aX - aE(X))(bY - bE(Y))\}$$
$$= E\{a(X-E(X))\,b(Y-E(Y))\}$$
$$= ab\,E\{(X-E(X))(Y-E(Y))\}$$
$$= ab\,Cov(X,Y)$$

(e) Noting $E(X+a) = E(X) + a$ and $E(Y+b) = E(Y) + b$,

$$Cov(X+a, Y+b) = E\{(X+\cancel{a}-E(X)-\cancel{a})(Y+\cancel{b}-E(Y)-\cancel{b})\}$$
$$= E\{(X-E(X))(Y-E(Y))\}$$
$$= Cov(X,Y)$$

(6f) Note $E(X+Y) = E(X) + E(Y)$, so

$$\text{Var}(X+Y) = E\{(X+Y-E(X)-E(Y))^2\}$$

$$= E\{(X-E(X) + Y-E(Y))^2\}$$

$$= E\{(X-E(X))^2 + 2(X-E(X))(Y-E(Y)) + (Y-E(Y))^2\}$$

$$= E\{(X-E(X))^2\} + 2E\{(X-E(X))(Y-E(Y))\} + E\{(Y-E(Y))^2\}$$

$$= \text{Var}(X) + \text{Var}(Y) + 2\,\text{cov}(X,Y)$$

(7) $\mathrm{Cov}(X, Y+Z) = E\{(X-\mu_x)(Y+Z-\mu_y-\mu_z)\}$

$= E\{(X-\mu_x)(Y-\mu_y + Z-\mu_z)\}$

$= E\{(X-\mu_x)(Y-\mu_y) + (X-\mu_x)(Z-\mu_z)\}$

$= E\{(X-\mu_x)(Y-\mu_y)\} + E\{(X-\mu_x)(Z-\mu_z)\}$

$= \mathrm{Cov}(X,Y) + \mathrm{Cov}(X,Z)$

(8) In Problem (7), let $X = X_1+X_2$, $Y=Y_1$ and $Z=Y_2$

so, $\mathrm{cov}(X_1+X_2, Y_1+Y_2)$

$= \mathrm{cov}(X_1+X_2, Y_1) + \mathrm{Cov}(X_1+X_2, Y_2)$

$= \mathrm{Cov}(X_1, Y_1) + \mathrm{Cov}(X_2, Y_1) + \mathrm{Cov}(X_1, Y_2) + \mathrm{Cov}(X_2 Y_2)$

The generalization is

$$\mathrm{Cov}\left(\sum_{i=1}^{p} a_i X_i, \sum_{j=1}^{q} b_j Y_j\right) = \sum_{i=1}^{p}\sum_{j=1}^{q} a_i b_j \mathrm{Cov}(X_i, Y_j)$$

(9) $$\mathrm{Corr}(aX, Y) = \frac{\mathrm{Cov}(aX, Y)}{SD(aX)\, SD(Y)}$$

$$= \frac{a\,\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(aX)}\,\sqrt{\mathrm{Var}(Y)}} = \frac{a\,\mathrm{Cov}(X,Y)}{\sqrt{a^2\mathrm{Var}(X)}\,\sqrt{\mathrm{Var}(Y)}}$$

$$= \frac{a\,\mathrm{Cov}(X,Y)}{|a|\sqrt{\mathrm{Var}(X)}\,\sqrt{\mathrm{Var}(Y)}} = \frac{\mathrm{sign}(a)\cancel{|a|}\,\mathrm{Cov}(X,Y)}{\cancel{|a|}\,SD(X)\;SD(Y)}$$

$$= \mathrm{sign}(a)\,\mathrm{Corr}(X,Y)$$

(10) (a) $$\sum_{i=1}^{n}(y_i - \bar{y}) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\bar{y} = \sum_{i=1}^{n} y_i - n\bar{y}$$

$$= \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} y_i = 0$$

(b) $$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i^2 - 2\bar{y}y_i + \bar{y}^2)$$

$$= \sum_{i=1}^{n} y_i^2 - 2\bar{y}\sum_{i=1}^{n} y_i + \sum_{i=1}^{n}\bar{y}^2$$

$$= \sum_{i=1}^{n} y_i^2 - 2\bar{y}\,n\bar{y} + n\bar{y}^2 = \sum_{i=1}^{n} y_i^2 - 2n\bar{y}^2 + n\bar{y}^2$$

$$= \sum_{i=1}^{n} y_i^2 - n\bar{y}^2$$

(4) $\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$

$$= \sum_{i=1}^{n}\left(x_i y_i - \bar{y}x_i - \bar{x}y_i + \bar{x}\,\bar{y}\right)$$

$$= \sum_{i=1}^{n} x_i y_i - \bar{y}\sum_{i=1}^{n} x_i + \bar{x}\sum_{i=1}^{n} y_i + \sum_{i=1}^{n}\bar{x}\,\bar{y}$$

$$= \sum_{i=1}^{n} x_i y_i - \bar{y}\,n\bar{x} - \bar{x}\,n\bar{y} + n\bar{x}\,\bar{y}$$

$$= \sum_{i=1}^{n} x_i y_i - 2n\bar{x}\,\bar{y} + n\bar{x}\,\bar{y}$$

$$= \sum_{i=1}^{n} x_i y_i - n\bar{x}\,\bar{y}$$

12 (a) $E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} E(Y_i) = \sum_{i=1}^{n} \mu = n\mu$

(b) $Var\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} Var(Y_i) = \sum_{i=1}^{n} \sigma^2 = n\sigma^2$

using (6f) and (5c)

(c) $Var(\bar{Y}) = Var\left(\frac{1}{n} \sum_{i=1}^{n} Y_i\right) = \frac{1}{n^2} Var \sum_{i=1}^{n} Y_i$

$= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$

(d) $E(\bar{Y}) = E\left(\frac{1}{n} \sum_{i=1}^{n} Y_i\right) = \frac{1}{n} \sum_{i=1}^{n} E(Y_i) = \frac{1}{n} \cancel{n} \cancel{\mu} = \cancel{\mu}$

(e) If L is unbiased,

$\cancel{\mu} = E(L) = E\left(\sum_{i=1}^{n} a_i Y_i\right) = \sum_{i=1}^{n} a_i E(Y_i)$

$= \sum_{i=1}^{n} a_i \mu = \cancel{\mu} \sum_{i=1}^{n} a_i \Rightarrow \sum_{i=1}^{n} a_i = 1$

(f) Yes; $a_i = \frac{1}{n}$ for $i = 1, \ldots, n$

(g) $Var(L) = Var \sum_{i=1}^{n} a_i Y_i = \sum_{i=1}^{n} a_i^2 Var(Y_i)$

$= \sigma^2 \sum_{i=1}^{n} a_i^2$

13 (a) $E\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$

$= E\left(\sum_{i=1}^{n}(X_i-\mu+\mu-\bar{X})^2\right)$

$= E\left(\sum_{i=1}^{n}\left[(X_i-\mu)^2+2(X_i-\mu)(\mu-\bar{X})+(\mu-\bar{X})^2\right]\right)$

$= E\left(\sum_{i=1}^{n}(X_i-\mu)^2+2(\mu-\bar{X})\sum_{i=1}^{n}(X_i-\mu)+n(\bar{X}-\mu)^2\right)$

$= \sum_{i=1}^{n}E(X_i-\mu)^2-2E\left((\bar{X}-\mu)\left(\sum_{i=1}^{n}X_i-n\mu\right)\right)+nE(\bar{X}-\mu)^2$

$= \sum_{i=1}^{n}\sigma^2-2E\{(\bar{X}-\mu)(n\bar{X}-n\mu)\}+n\,Var(\bar{X})$

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

$= n\sigma^2-n\,Var(\bar{X}) = n\sigma^2-\cancel{n}\frac{\sigma^2}{\cancel{n}}$

$= \sigma^2(n-1)$, so that

$E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)$

$= \frac{1}{n-1}E\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right) = \frac{1}{\cancel{n-1}}\sigma^2\cancel{(n-1)}$

$= \sigma^2$ <u>Unbiased</u>

(13b) $\mathrm{Cov}(X_j, \bar{X}) = \mathrm{Cov}\left(X_j, \frac{1}{n}\sum_{i=1}^{n} X_i\right)$

$= \frac{1}{n}\mathrm{Cov}\left(X_j, X_j + \sum_{i \neq j} X_i\right)$

$= \frac{1}{n}\left(\mathrm{Var}(X_j) + \sum_{i \neq j}\mathrm{Cov}(X_i, X_j)\right) = \frac{1}{n}\sigma^2 + 0$

$= \frac{\sigma^2}{n}$

(c) $\mathrm{Cov}(\bar{X}, X_j - \bar{X}) = \mathrm{Cov}(\bar{X}, X_j) - \mathrm{Cov}(\bar{X}, \bar{X})$

$= \mathrm{Cov}(X_j, \bar{X}) - \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} - \frac{\sigma^2}{n} = 0$

(14) (a) $\mathrm{Cov}(\bar{X}, \bar{Y}) = \mathrm{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} X_i, \frac{1}{n}\sum_{j=1}^{n} Y_j\right)$

$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(X_i, Y_j)$

$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathrm{Cov}(X_i, Y_i) + \sum\sum_{i \neq j}\mathrm{Cov}(X_i, Y_j)\right)$

$= \frac{1}{n^2}\left(n\,\sigma_{xy} + 0\right) = \sigma_{xy}/n$

(b) $\mathrm{Corr}(\bar{X}, \bar{Y}) = \dfrac{\mathrm{Cov}(\bar{X}, \bar{Y})}{\sqrt{\sigma_x^2/n}\sqrt{\sigma_y^2/n}}$

$= \dfrac{\sigma_{xy}/\cancel{n}}{\cancel{\frac{1}{n}}\sqrt{\sigma_x^2}\sqrt{\sigma_y^2}} = \dfrac{\sigma_{xy}}{\sigma_x\,\sigma_y} = \mathrm{Corr}(X_i, Y_i)$

(15) $Cov(Y_1, Y_2) = Cov(\alpha_1 X_1 + \alpha_2 X_2 + \varepsilon_1, \beta_1 X_1 + \beta_2 X_2 + \varepsilon_2)$

$= Cov(\alpha_1 X_1 + \alpha_2 X_2, \beta_1 X_1 + \beta_2 X_2) + Cov(\varepsilon_1, \varepsilon_2)$

$= \alpha_1 \beta_1 Cov(X_1, X_1) + \alpha_1 \beta_2 Cov(X_1, X_2) + \alpha_2 \beta_1 Cov(X_2, X_1)$
$+ \alpha_2 \beta_2 Cov(X_2, X_2) + \sigma_{12}$

$= \alpha_1 \beta_1 Var(X_1) + \alpha_2 \beta_2 Var(X_2) + (\alpha_1 \beta_2 + \alpha_2 \beta_1) Cov(X_1, X_2)$
$+ \sigma_{12}$

$= \alpha_1 \beta_1 \phi_{11} + \alpha_2 \beta_2 \phi_{22} + (\alpha_1 \beta_2 + \alpha_2 \beta_1) \phi_{12} + \sigma_{12}$

(16) (a) $H_0: \beta_1 = 0$. "Affect" is okay because of random assignment

(b) $H_0: \beta_4 = \beta_5 = 0$

(c) $H_0: \beta_2 = \beta_3 = 0$

(d) $H_0: \beta_3 = 0$

(e) X_1 should be independent of the other explanatory variables because of random assignment. It would be surprising if the other explanatory variables were unrelated to one another.