CV2018 - Probability & Statistics

Semester I 2012 - 2013

1. (a) $P(A) = 0.3 \quad P(B) = 0.5 \quad P(A \cap B) = 0.2$

(i) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.5} = 0.4 \qquad \underline{\underline{P(A|B) = 0.4}}$

(ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.5 - 0.2 = 0.6 \qquad \underline{\underline{P(A \cup B) = 0.6}}$

(iii) $P(\bar{A} \cup \bar{B}) = 1 - P(\overline{\bar{A} \cup \bar{B}}) = 1 - P(A \cap B) = 1 - 0.2 = 0.8 \qquad \underline{\underline{P(\bar{A} \cup \bar{B}) = 0.8}}$

(b) The probability that box C is chosen is $P(C) = 0.5$ and hence $P(D) = 0.5$.

Probability that a red bar is chosen given that box C is chosen $P(R|C) = 2/5$

———— " ———— box D is chosen $P(R|D) = 4/9$

Probability that a white bar is chosen given that box C is chosen $P(W/C) = 3/5$

———— " ———— box D is chosen $P(W|D) = 5/9$

(i) $P(R) = P(R|C)P(C) + P(R|D)P(D) = \left(\frac{2}{5} \cdot \frac{1}{2}\right) + \left(\frac{4}{9} \cdot \frac{1}{2}\right) = \frac{19}{45} \Rightarrow \underline{\underline{P(R) = 19/45}}$

(ii) $P(C|R) = \ldots\ldots$ We know that $P(C \cap R) = P(C|R)P(R) = P(R|C)P(C) \rightarrow$ Bayes' Theorem

$P(C|R) = \frac{P(R|C)P(C)}{P(R)} = \frac{(2/5)(1/2)}{19/45} = 9/19 \qquad \underline{\underline{P(C|R) = 9/19}}$

(iii) $P(4R) = P(2R \text{ from } C) \times P(2R \text{ from } D)$

$P(4R) = \left(\frac{2}{5} \times \frac{1}{4}\right) \times \left(\frac{4}{9} \times \frac{3}{8}\right) = \frac{1}{60} \qquad \underline{P(4R) = 1/60}$

(iv) $P(4R \text{ or } 4W) = P(4R) + P(4W)$

$P(4W) = P(2W \text{ from } C) \times P(2W \text{ from } D)$

$P(4W) = \left(\frac{3}{5} \times \frac{2}{4}\right) \times \left(\frac{5}{9} \times \frac{4}{8}\right) = \frac{1}{12}$

$P(4R \text{ or } 4W) = \frac{1}{60} + \frac{1}{12} = \frac{6}{60} = \frac{1}{10} \qquad \underline{P(4R \text{ or } 4W) = 1/10}$

CV2018 - Probability and Statistics

Semester 1 2012 - 2013

2. (a) (i) $E[X] = 1(0.2) + 2(0.1) + 3(0.4) + 4(0.3)$

$\underline{E[X] = 2.8}$

$E[X^2] = 1^2(0.2) + 2^2(0.1) + 3^2(0.4) + 4^2(0.3) = 9$

$Var[X] = E[X^2] - E[X]^2 = 9 - (2.8)^2 = 1.16$

$\underline{Var[X] = 1.16}$

(ii) $Y = aX + b$

$Var[Y] = Var[aX+b] = a^2\,Var[X] = 10.44$

$a^2 = 9 \rightarrow \underline{a = 3} \vee \underline{a = -3}$

$E[Y] = E[aX+b] = aE[X] + b = 13.4$

for $a = 3$, $b = 13.4 - 3(2.8) = 5$, $\underline{b = 5}$

$a = -3$, $b = 13.4 + 3(2.8) = 21.8$, $\underline{b = 21.8}$

(b) X: total number of containers acceptable

(i) $P(X \geq 4) = P(X=4) + P(X=5) = \binom{5}{4}(0.9)^4(0.1) + \binom{5}{5}(0.9)^5 = \underline{0.91854}$

(ii) $P(X \leq 2) = 1 - P(X=3) = 1 - \binom{3}{3}(0.9)^3 = \underline{0.271}$

(iii) $P(\mu - 1.5 < H \leq \mu + 1.5) \geq 0.99$

$P\left(\frac{\mu - 1.5 - \mu}{\sigma} < \frac{H - \mu}{\sigma} \leq \frac{\mu + 1.5 - \mu}{\sigma}\right) \geq 0.99$

$P\left(\frac{-1.5}{\sigma} < \frac{H-\mu}{\sigma} \leq \frac{1.5}{\sigma}\right) \geq 0.99$

$\Phi\left(\frac{1.5}{\sigma}\right) - \Phi\left(-\frac{1.5}{\sigma}\right) \geq 0.99$

$\Phi\left(\frac{1.5}{\sigma}\right) - \left(1 - \Phi\left(\frac{1.5}{\sigma}\right)\right) \geq 0.99$

$2\Phi\left(\frac{1.5}{\sigma}\right) - 1 \geq 0.99$

$\Phi\left(\frac{1.5}{\sigma}\right) \geq 0.995$

$\Phi\left(\frac{1.5}{\sigma}\right) \geq \Phi(2.576)$

$\frac{1.5}{\sigma} \geq 2.576$

$\sigma \leq 0.582$

Maximum value of σ is $\underline{\underline{0.582}}$

CV2018 - Probability and Statistics

Semester I - 2012-2013

3. (d) $E[X] = \mu \quad E[Z] = \beta \qquad cov(X_i, Z_j) = \gamma$ if $i = j$ and $cov(X_i, Z_j) = 0$ if $i \neq j$

(i) $E[X_i Z_j] = \mu\beta$ for $i \neq j$

(ii) $cov(X_i Z_j) = \gamma = E[X_i Z_j] - E[X_i]E[Z_j] = \gamma$

$E[X_i Z_j] = \gamma + \mu\beta$ for $i = j$

$$E[W] = \alpha \left[E\left[n \sum_{k=1}^{n} X_k Z_k - \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{j=1}^{n} Z_j\right)\right]\right]$$

linearity of expectation

$$= \alpha \left[n \sum_{k=1}^{n} E[X_k Z_k] - E\left[\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{j=1}^{n} Z_j\right)\right]\right]$$

$$\begin{matrix} X_1Z_1 & X_1Z_2 & X_1Z_3 & \cdots & X_1Z_n \\ X_2Z_1 & X_2Z_2 & X_2Z_3 & \cdots & X_2Z_n \\ \vdots & \vdots & \vdots & & \vdots \\ X_nZ_1 & X_nZ_2 & X_nZ_3 & \cdots & X_nZ_n \end{matrix}$$

$$= \alpha \left[n(n(\gamma + \mu\beta)) - E[\underbrace{(X_1 + X_2 + X_3 + \cdots + X_n)(Z_1 + Z_2 + Z_3 + \cdots + Z_n)}]\right]$$

There will be n^2 terms: n terms with $i = j$; $n^2 - n$ terms with $i \neq j$

$$= \alpha \left[n^2(\gamma + \mu\beta) - \left(n(\gamma + \mu\beta) + (n^2 - n)(\mu\beta)\right)\right]$$

$$= \alpha \left[n^2\gamma + \cancel{n^2\mu\beta} - n\gamma - \cancel{n\mu\beta} - \cancel{n^2\mu\beta} + \cancel{n\mu\beta}\right]$$

$$E[W] = \alpha(n^2\gamma - n\gamma) = \gamma$$

$$\boxed{\alpha = \frac{1}{n(n-1)}}$$

For these steps, I will elaborate more clearly.

$E[(X_1 + X_2 + X_3 + \cdots + X_n)(Z_1 + Z_2 + Z_3 + Z_4 + \cdots + Z_n)]$

$= E[\underbrace{X_1Z_1 + X_1Z_2 + X_1Z_3 + \cdots + X_1Z_n}_{n \text{ terms}} + \underbrace{X_2Z_1 + X_2Z_2 + X_2Z_3 + \cdots X_2Z_n}_{n \text{ terms}} + \underbrace{\cdots\cdots}_{n \text{ terms}} + \underbrace{\cdots}_{n \text{ terms}} + X_nZ_1 + X_nZ_2 + \cdots + X_nZ_n$

So, for each value of i for X, we have $j = 1$ to n for Z i.e. for X_1, we have n terms of Z, for X_2, we have n terms of Z as well, and so on. Because we have n terms of X, then we have $n \times n$ terms altogether. Out of n^2 terms, we have $\underbrace{X_1Z_1, X_2Z_2, \ldots, X_nZ_n}_{n \text{ terms}}$

n terms for $i = j$ and $n^2 - n$ for $i \neq j$.

After this, we can then use the linearity of expectation to expand it.

3. (b) X (random variable) = amount of taxi fare.

$X \sim N(\mu, \sigma^2)$ and assume that X is independent and identically distributed.

$\bar{X} = 12.5$, $S = 2.5$ → $n = 10$

(i) $$\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2}), n-1}}$$

$\chi^2_{0.025, 9} = 19.023$

$\chi^2_{0.975, 9} = 2.700$

$$\frac{9(2.5)^2}{\chi^2_{0.025,9}} \le \sigma^2 \le \frac{9(2.5)^2}{\chi^2_{0.975,9}}$$

$$2.957 \le \sigma^2 \le 20.833 \implies 95\% \text{ C.I. for } \sigma^2$$

(ii) $$\bar{X} - t_{\frac{(1-p)}{2}, n-1} \cdot S\sqrt{\frac{n+1}{n}} \le X_{n+1} \le \bar{X} + t_{(\frac{1-p}{2}), n-1} \cdot S\sqrt{\frac{n+1}{n}}$$

with $p = 0.95$, $n = 10$, $S = 2.5$, and $\bar{X} = 12.5$

$$12.5 - 2.262(2.5)\sqrt{\tfrac{11}{10}} \le X_{11} \le 12.5 + 2.262(2.5)\sqrt{\tfrac{11}{10}}$$

$$6.569 \le X_{11} \le 18.431 \implies 95\% \text{ C.I. for } X_{11}$$

(iii) Let r.v. Y to be $\frac{X_{11}+X_{12}+X_{13}+X_{14}+X_{15}}{5}$

so $E[Y] = \frac{1}{5}\left[E[X_{11}] + E[X_{12}] + E[X_{13}] + E[X_{14}] + E[X_{15}]\right] = \frac{1}{5}(5\mu) = \mu$

$Var[Y] = \frac{1}{25}(5\sigma^2) = \frac{1}{5}\sigma^2$

We also know that $E[\bar{X}] = \mu$ and $Var[\bar{X}] = \frac{\sigma^2}{10}$ where $\bar{X}$ is sample mean of first 10 taxis.

Let's construct a new random variable $Z_0 = Y - \bar{X}$. $E[Z_0] = E[Y] - E[\bar{X}] = 0$

Here we assume Y & $\bar{X}$ are independent. $Var[Z_0] = Var[Y] + Var[\bar{X}] = \frac{3}{10}\sigma^2$

$\therefore Z_1 = \frac{Y-\bar{X}}{\sqrt{\frac{3}{10}\sigma^2}} \sim N(0,1) \approx \frac{Y-\bar{X}}{\sqrt{\frac{3}{10}S^2}}$ → Z_1 follow $N(0,1)$ so we can make use of the table in page (ii).

$= P(10 \le Y \le 15)$

$$= P\left(\frac{10-\bar{X}}{\sqrt{\frac{3}{10}S^2}} \le \frac{Y-\bar{X}}{\sqrt{\frac{3}{10}S^2}} \le \frac{15-\bar{X}}{\sqrt{\frac{3}{10}S^2}}\right)$$

$$= P(-1.83 \le Z_1 \le 1.83)$$

$$= \Phi(1.83) - \Phi(-1.83) = 2\Phi(1.83) - 1 = 0.93275$$

CV2018 – Probability & Statistics

Semester I, 2012-2013

3 (b) (IV) $Y = X_{12} - X_{11}$

$E[Y] = E[X_{12} - X_{11}] = E[X_{12}] - E[X_{11}] = \mu - \mu = 0$

$Var[Y] = Var[X_{12} - X_{11}] = Var[X_{12}] + Var[X_{11}] = 2\sigma^2$

$Z = \frac{Y-0}{\sqrt{2}\sigma} \sim N(0,1) \Rightarrow Z = \frac{Y}{\sqrt{2}\sigma} \sim N(0,1)$

$P(-Z_{0.025} \leq Z \leq Z_{0.025}) = 0.95$

$P\left(-1.960 \leq \frac{Y}{\sqrt{2}\sigma} \leq 1.960\right) = 0.95$

95% C.I for $X_{12} - X_{11}$ is $-2.77\sigma \leq Y \leq 2.77\sigma$

Since you don't have the value of σ, you may use $\hat{\sigma}$ or s (estimator of σ).

so the 95% C.I for $X_{12} - X_{11}$ is $-6.93 \leq Y \leq 6.93$

4 (a) $G \sim N(\mu_G, \sigma_G)$ $\bar{G} = 0.1$ $S_G = 0.02$ $n = 5$

$J \sim N(\mu_J, \sigma_J)$ $\bar{J} = 0.08$ $S_J = 0.025$ $n = 5$ } 95% confidence level

$H_0: \mu_G - \mu_J \leq 0$ $H_1: \mu_G - \mu_J > 0$

Test Statistics:

$$Z = \frac{(\bar{G} - \bar{J}) - (\mu_G - \mu_J)}{\sqrt{\frac{S_G^2}{5} + \frac{S_J^2}{5}}} \sim N(0,1)$$

$$Z = \frac{(0.1 - 0.08) - (0)}{\sqrt{\frac{0.02^2 + 0.025^2}{5}}} = 1.397$$

$Z_{0.05} = 1.645$

Since $Z < Z_{0.05}$ we can not reject H_0.

In conclusion, we are not confident that the average percentage of profit for company G is more than that for company J (since we can't reject H_0)

4 (b) $L = a t^{1.1} + b \rightarrow$ let's take $L = Y$, $t^{1.1} = X$, $b = a_0$, $a = a_1$

X	1	1.793	3.348	5.615
L(Y)	3190	3334	3705	4162

$\bar{X} = 2.939$ $\bar{Y} = 3597.75$

$S_{xx} = 12.401$ $S_{xy} = 2646.683$

(i) $\hat{a} = \hat{a}_1 = \frac{S_{xy}}{S_{xx}} = 213.425$ $\hat{b} = \hat{a}_0 = 2970.494$

(ii)

Y	3190	3334	3705	4162
$\hat{Y}$	3183.919	3353.165	3685.041	4168.875

$\hat{Y} = \hat{a}_0 + \hat{a}_1 X$

$$SS_E = \sum_{i=1}^{4} (Y - \hat{Y})^2 = 6.081^2 + (-19.165)^2 + 19.959^2 + (-6.875)^2$$

$$SS_E = 849.903$$

$$a_1 = \hat{a}_1 \pm t_{\frac{(1-p)}{2}, n-2} \sqrt{\frac{SS_E}{(n-2)S_{xx}}} \qquad t_{0.025, 2} = 4.303$$

$$a_1 = 213.425 \pm 4.303 \sqrt{\frac{849.903}{(2)(12.401)}}$$

$$a_1 = 213.425 \pm 25.189$$

(iii) $$Y_{n+1} = \hat{Y}_{n+1} \pm t_{\frac{(1-p)}{2}, n-2} \sqrt{\frac{SS_E}{n-2}} \sqrt{1 + \frac{1}{n} + \frac{(X_{n+1} - \bar{X})^2}{S_{xx}}}$$

$$\hat{Y}_{n+1} = \hat{a}_0 + \hat{a}_1 X = 2790.494 + 213.425 (3.5)^{1.1} = 3637.176$$

$$Y_{n+1} = 3637.176 \pm t_{0.005, 2} \sqrt{\frac{849.903}{2}} \sqrt{1 + \frac{1}{4} + \frac{(3.967 - 2.939)^2}{12.401}}$$

$$\cancel{= 3637.176 \pm 9.925 (20.614)(1.156)}$$

$$Y_{n+1} = 3637.176 \pm 236.515$$

I wish the best for your exam. If you want to discuss about this solution, feel free to contact me at 96130369 (Yandi). Good Luck!