Actuarial Society of India

Examinations

May 2006

CT3 – Probability and Mathematical Statistics

Indicative Solutions

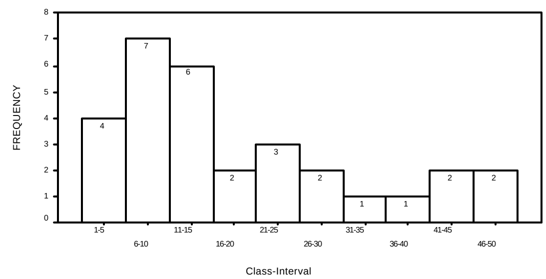
1) a) Stem Leaf

Stem	Leaf
0	7, 6, 2, 8, 1, 3, 8, 4, 7
1	3, 2, 4, 3, 0, 4, 0, 3, 6, 9
2	5, 6, 1, 8, 4
3	4, 6
4	7, 5, 1
5	0

b)

Class	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	Total
Frequency	4	7	6	2	3	2	1	1	2	2	30

Histogram



c) $Mean: \dfrac{\Sigma fx}{\Sigma f} = \dfrac{560}{30} = 18.67$

$$Variance: \frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx}{N}\right)^2$$

$$= \frac{16140}{30} - \left(\frac{560}{30}\right)^2 = 189.56$$

[6]

2) $P(A) = \dfrac{26}{52} = \dfrac{1}{2}; P(B) = \dfrac{4}{52} = \dfrac{1}{13}; P(A \cap B) = \dfrac{2}{52} = \dfrac{1}{26}$

Computed values lead to *P(A Ç B) = P(A) . P(B)*. Hence *A* and *B* are independent.

So the color of the card does not affect its probability of being a king.

[2]

3) M : The event that the student passed in Mathematics test.

S : The event that the student passed in Statistics test.

a) $P(M^c \cap S^c) = P((M \cup S)^c) = 1 - P(M \cup S)$

Now *P(MÈS) = P(M) + P(S) – P(MÇ S)*

$$= \frac{65}{120} + \frac{75}{120} - \frac{35}{120} = \frac{105}{120}$$

$$\therefore P(M^c \cap S^c) = 1 - \frac{105}{120} = \frac{15}{120} = \frac{1}{8}$$

b) B : The event that student passes at least one test

: $M \subset S$

$$P(M/B) = \frac{P(M \cap B)}{P(B)}; P(B) > 0$$

$$P(M \cap B) = P(M) = \frac{65}{120}; P(B) = P(M \cup S) = \frac{105}{120}$$

Hence $P(M/B) = \frac{65}{120} / \frac{105}{120} = \frac{65}{105} = \frac{13}{21}$

[5]

4) a) Joint *pdf* of (X,Y) = $f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$

$$= \begin{cases} 6e^{-(2x+3y)} & ; x, y > 0 \\ 0 \ elsewhere \end{cases}$$

b) Marginal *pdf* of X : $f_X(x) = \begin{cases} 2e^{-2x} & ; x > 0 \\ 0 \ elsewhere \end{cases}$

Marginal *pdf* of Y : $f_Y(x) = \begin{cases} 3e^{-3y} & ; y > 0 \\ 0 \ elsewhere \end{cases}$

c) $P((X \le 1) \cap (Y \le 1)) = 1 - e^{-2} - e^{-3} + e^{-5}$

d) $P[(1 < X < 3) \cap (1 < Y < 2)] = F(3,2) - F(3,1) - F(1,2) + F(1,1)$

$= e^{-5} - e^{-8} - e^{-9} + e^{-12}$

[6]

5) $M_X(t) = \left[\frac{1}{2a} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{1}{a}[-(x-a)]} dx + \int_{a}^{\infty} e^{tx} e^{-\frac{1}{a}(x-a)} dx \right]$

$= e^{at} / (1 - a^2 t^2)$

$K_X(t) = \log M_X(t) = at - \log(1 - a^2 t^2)$

Mean: $K'_X(t)|_{t=0} = a$; variance : $K''_X(t)|_{t=0} = 2a^2$

[5]

6) $f(x) = \begin{cases} \frac{1}{16} x^2 e^{-x/2} & ; \quad x > 0 \\ 0 & otherwise \end{cases}$

$P(X > 12) = \int_{12}^{\infty} f(x) dx$

$= \int_{12}^{\infty} \frac{1}{16} x^2 e^{-x/2} dx$ $\quad = 13 e^{-6}$

$= 0.0322$ **[3]**

7) This problem is an application of Binomial distribution

Here $n = 7$; $p = 1/10$ and $q = 9/10$

a) $P(X \geq 1) = 1 - P(X = 0) = 0.5217$

b) $P(X = 0) = 0.4783$

c) $P(X = 7) \approx 0$

[3]

8) $P(X(t) = n)\ e^{-lt}\ (l\,t)^n / n! \quad ; \quad n = 0,1,2,\ldots$

i) *P (6 calls received in 2 hours)* is evaluated by taking $n = 6,\ t = 2$ so that $lt = 16$

$P(X(t) = 6) = e^{-16}\ 16^6/6! = 0.0026$

ii) *P (atleast 2 calls in the next 20 minutes)* is evaluated by taking t = 1/3, so that $l\,t = 8/3$

$P(X(t) \geq 2) = 1 - [P(X(t) = 0) + P(X(t) = 1)]$

$= 1 - [\ e^{-8/3} + e^{-8/3}\,(8/3)]$

$= 0.7452$

[3]

9) a) Statement of central limit theorem

b) Here $n = 950,\ p = 0.57,\ q = 0.43$

Mean : $m = np = 950 \times 0.57 = 541.5$

SD : $s = \sqrt{npq} = 15.2593$

i) $P(X \geq 565) = P(X \geq 565.5)$

$$= P\left(\frac{X - m}{s} \geq \frac{565.5 - 541.5}{15.2593}\right)$$

$= P\,(Z > 1.573) = 0.0583$

ii) $P(535 \leq X \leq 575) = P(534.5 \leq X \leq 575.5)$

$$= P\left(\frac{534.5 - 541.5}{15.2593} \leq Z \leq \frac{575.5 - 541.5}{15.2593}\right)$$

$= P(-0.46 \leq Z \leq 2.23)$

$= P(0 \leq Z \leq 0.46) + P(0 \leq Z \leq 2.23)$

$= 0.6643$

[5]

10) $X \sim f(x) = m\,e^{-mx} \quad ; x > 0, m > 0$

$H_0 : m = 20 \qquad H_1 : m = 30$

$$P(\text{Type I error}) = \int_{28}^{\infty} 20\, e^{-20x} dx = e^{-560} = 0$$

$$P(\text{Type II error}) = \int_{0}^{28} 30\, e^{-30x} dx = 1 - e^{-840} = 1$$ **[2]**

11) a) H_0: There is no association between gender and color blindness

Observed	Expected
904	946.25
998	955.75
91	48.75
7	49.75

$$c^2 = \sum \frac{(0-E)^2}{E}$$

$$c^2 = 76.598$$

Critical value of $c^2_{0.05}$ (1) : 3.8414

Inference : Reject H_0

b) $L(q/\text{data}) = \text{Constant}\left(\frac{1-q}{2}\right)^{904}\left(\frac{1-q^2}{2}\right)^{998}\left(\frac{q}{2}\right)^{91}\left(\frac{q^2}{2}\right)^{7}$

$$\log L = \text{Constant} + 904 \log\left(\frac{1-q}{2}\right) + 998 \log\left(\frac{1-q^2}{2}\right)$$
$$+ 91 \log\left(\frac{q}{2}\right) + 7 \log\left(\frac{q^2}{2}\right)$$

$$\frac{\partial \log L}{\partial q} = 904\left(\frac{2}{1-q}\right)\left(-\frac{1}{2}\right) + 998\left(\frac{2}{1-q^2}\right)(-q) + 91\left(\frac{2}{q}\right) + 7\left(\frac{2}{q^2}\right)2q$$

$$\frac{\partial \log L}{\partial q} = 0 \Rightarrow q^2(3005) + q(904) - 105 = 0$$

solution of this quadratic equation gives an admissible $q = 0.089$

[15]

12. From the data, we have

New method : $\bar{x} = \frac{608}{8} = 76$

Old method : $\bar{y} = \frac{576}{8} = 72$

$$s_1^2 = \frac{1}{n_1 - 1}\sum (x_i - \bar{x})^2 = \frac{336}{7} = 48.00$$

$$s_2^2 = \frac{1}{n_2 - 1}\sum (y_i - \bar{y})^2 = \frac{288}{7} = 41.14$$

pooled variance : $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

$$= \frac{336 + 228}{14} = 44.57$$

$s = 6.676$

a) H_0 : The mean scores under two methods are equal.

Test statistic : $t = \dfrac{|\bar{x} - \bar{y}|}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

$= \dfrac{4}{6.676\sqrt{2/8}} = 1.198$

Critical value of t at 5% level for 14 df : 2.145

Inference : Do not reject H_0

b) H_0 : The variance of scores under the two methods are equal

Test statistic : F = $\dfrac{s_1^2}{s_2^2} = \dfrac{48.00}{41.14} = 1.1667$

Critical value $F_{0.05}$ (7,7) = 4.99

Inference : Do not reject H_0.

c) $P\left((\bar{x} - \bar{y}) - t_{a/2} s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} < m_1 - m_2 < (\bar{x} - \bar{y}) + t_{a/2} s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} \right) = 1 - a$

The 95% confidence interval is $(76 - 72) \pm 2.145 \text{ X } 6.766\sqrt{\dfrac{1}{8} + \dfrac{1}{8}}$

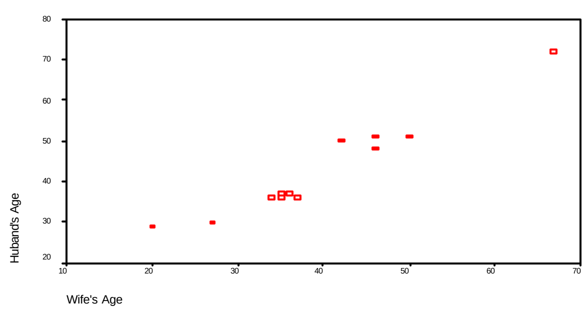
This gives the 95% confidence interval as

(4 ± 2.145 x 6.676 x 0.5)

ie. (-3.16, 11.16)

[12]

13.a)



From the scatter plot we infer that there is a positive association.

b) For the data given

$\bar{x} = 42 \qquad \bar{y} = 39.59 \qquad n = 12$

$\Sigma(x - \bar{x})\ (y - \bar{y}) = 1{,}554.75 \qquad \Sigma(x - \bar{x})^2 = 1626.25 \quad \Sigma(y - \bar{y})^2 = 1582.92$

$r = \dfrac{\Sigma(x - \bar{x})\ (y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2}\sqrt{\Sigma(y - \bar{y})^2}}$

$= \dfrac{1{,}554.75}{\sqrt{1626.25}\sqrt{1582.92}} = 0.969$ (High positive correlation)

c) $H_0 : \rho = 0$; $H_1 : \rho \neq 0$

Test statistic $t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \dfrac{0.969\times\sqrt{10}}{\sqrt{1-0.969^2}} = -\dfrac{0.969\times 3.16}{0.247}$

$$= \frac{3.06}{.245} = 12.498$$

Critical value of $t_{0.05}$ at 10 df : 2.228. Therefore, Reject H_0

[8]

14.a) β_0 and β_1 are the parameters of the model

Y : yield of chemical process is dependent variable.
X : temperature is the independent variable

b) From the data $\Sigma x = 0$ $\Sigma y = 92$
$\Sigma x^2 = 110$ $\Sigma xy = 108$

Hence the normal equations give $\hat{b}_0 = 8.36$ and $\hat{b}_1 = 0.98$

The prediction equation is $\hat{y} = 8.36 + 0.98x$

c) Total sum of squares : $\Sigma(y-\bar{y})^2 = 164.545$

Residual sum of squares: $\Sigma(y-\hat{y})^2 = 58.509$

ANOVA				
Source	SS	d.f.	MSS	F
Regression	106.036	1	106.036	
Residual	58.509	9	6.501	$F = \dfrac{106.036}{6.501} = 16.311$
Total	164.545	10		

For testing $H_0 : \beta_1 = 0$ at 5% level, we have the critical value F(1,9) at 5% as : 5.12. Hence we reject H_0

d) 95% confidence interval for β_1

$$\hat{b}_1 - t_{n-2,\alpha/2}\frac{S}{\sqrt{S_{xx}}} < b_1 < \hat{b}_1 + t_{n-2,\alpha/2}\frac{S}{\sqrt{S_{xx}}}$$

where $S_{xx} = \Sigma(x-\bar{x})^2$ and S: Mean residual sum of squares
= 110 = 6.501

This gives 95% CI for β_1 (- 0.423, 2.383)

e) $(1-\alpha)$% CI for true mean value of y when $x = x_0$:

$$\left(\hat{b}_0 + \hat{b}_1 x_0 \mp t_{n-2,\alpha/2}\, S\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}}\right)$$

This gives 95% CI for y when x = 3 as (5.186, 17.414)

f) The residual sum of squares seems to be minimum. This model is O.K

[13]

15a) $\int_0^\infty\int_0^\infty kxy(4x+9y)e^{-(x+y)}dxdy = 1$

$$\Rightarrow \int_0^\infty 4x^2e^{-x}dx\int_0^\infty ye^{-y}dy + \int_0^\infty 9xe^{-x}dx\int_0^\infty y^2e^{-y}dy = \frac{1}{k}$$

$$\Rightarrow 4\Gamma 3\ \Gamma 2 + 9\Gamma 2\ \Gamma 3 = 1/k$$

$$\Rightarrow 4\,x\,2\,x1 + 9\ \ x1\,x\,2 = 1/k \Rightarrow k = \frac{1}{26}$$

Marginal pdf of Y :

$$f_y(y) = \int_0^\infty \frac{1}{26}xy(4x+9y)e^{-(x+y)}dx$$

$$= \frac{1}{26}\left[ye^{-y}(8+9y)\right];\ y>0$$

Marginal pdf of X :

$$f_x(x) = \frac{xe^{-x}(4x+18)}{26}\ ;\ x>0$$

Conditional pdf of X given Y = y

$$f\left(x/Y=y\right) = \frac{xe^{-x}(4x+9y)}{(8+9y)}$$

$$E(X/Y=y) = \int_0^\infty x\frac{xe^{-x}(4x+9y)}{(8+9y)}dx$$

$$= \frac{24+18y}{(8+9y)}$$

$$E(X^2/Y=y) = \int_0^\infty x^2\ \frac{x\,e^{-x}(4x+9y)}{(8+9y)}dx$$

$$= \frac{96+54y}{(8+9y)}$$

$$V(X/Y=y) = \frac{96+54y}{8+9y)} - \left(\frac{24+18y}{8+9y}\right)^2$$

[10]

16.) X_i ~ Poisson with mean 5

S = $\sum_{t=1}^{N} X_i$, $N \sim B(10, p=0.2)$

E(S) = $m_s = m_N\ m_x$ = (10 x 0.2) x 5

V(S) = $s_S^2 = m_N s_x^2 + s_N^2 m_x^2$

= (10 x 0.2) 5 + (10 x 0.2 x 0.8) x 5^2

= 10 + 40 = 50

[2]

*****