Actuarial Society of India

Examinations

November 2005

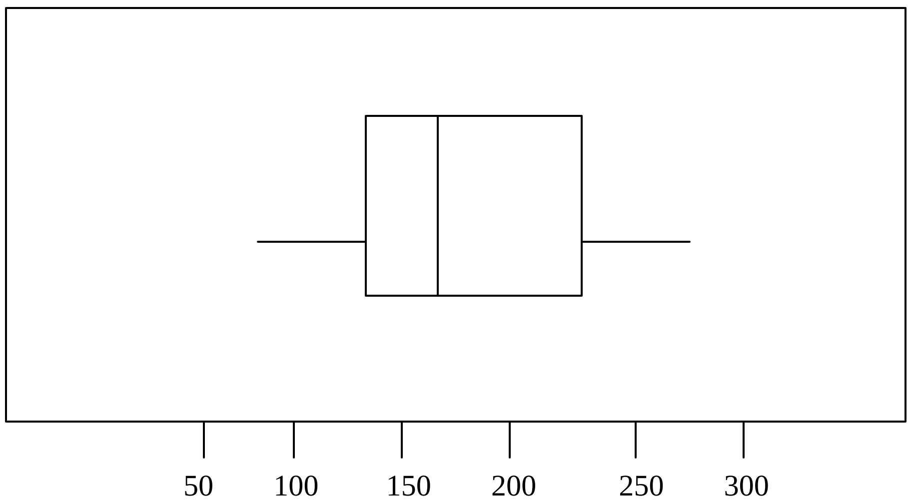
CT3 – Probability and Mathematical Statistics

Indicative Solutions

Q.1) Median = 176

$Q_1 = 4.25\text{th item} = 136.5$
$Q_3 = 11.75\text{th item} = 245$
Alternatively
$Q_1 = 136$ and $Q_3 = 253$

[1]



[1]
Total [2]

Q.2) Given $\bar{x} = 15$, $s^2 = 9$ n = 5

$\Sigma x = n\bar{x} = 75$ and

corrected $\Sigma x = 75 - 3 + 10 = 82$

New Mean = $\frac{82}{7} = 11.71$ **[1]**

Variance $s^2 = \dfrac{n\Sigma x_i^2 - (\Sigma x_i)^2}{n(n-1)}$

$9 = \dfrac{5\Sigma x_i^2 - (75)^2}{5 \times 4}$

$180 = 5\Sigma x_i^2 - (75)^2$

$\therefore \Sigma x_i^2 = 1161$

$corrected\ \Sigma x_i^2 = 1161 + (-3)^2 + (10)^2$
$= 1270$ **[1]**

New $s^2 = \dfrac{7(1270) - (82)^2}{7 \times 6}$

$= \dfrac{8890 - 6724}{42} = \dfrac{2166}{42}$

= 51.5 (approximately) **[1]**

Total [3]

Q.3) Let A and B denote the event that the actuary will be selected in X and will be rejected in Y respectively.

Given

$P(A) = 0.7, \quad P(\overline{A}) = 0.3$

$P(B) = 0.5 \quad P(\overline{B}) = 0.5$

$P(\overline{A} \cup \overline{B}) = 0.6$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ **[1]**

But $P(A \cap B) = 1 - P((\overline{A} \cup \overline{B}) = 0.4$ **[1]**

$\therefore P(A \cup B) = 0.7 + 0.5 - 0.4$

$= 0.8$ **[1]**

Total [3]

Q.4) Let P(A), P(B) and P(C) denote the probabilities that the loan application was processed by the Actuaries A,B and C respectively.

Given that $P(A) = 0.40, \quad P(B) = 0.35, \quad P(C) = 0.25$ **[1]**

Let E be the event that the loan application containing error.

Further, it is given that $P(E/A) = 0.04, \quad P(E/B) = 0.06, \quad P(E/C) = 0.03$ **[1]**

Using Bayes theorem, we have

$$P(A/E) = \frac{P(A).P(E/A)}{P(A).P(E/A) + P(B).P(E/B) + P(C).P(E/C)}$$ **[1]**

$$= \frac{0.4 x 0.04}{0.4(0.04) + 0.35(0.06) + 0.25(0.03)}$$ **[1]**

$= 0.36$ **Total [4]**

Q.5) $$E(X) = \int_{-a}^{a} \frac{2a}{p} \frac{x}{(a^2 + x^2)} dx$$

$= 0$ (the integrand is an odd function) **[1]**

$$EX^2 = \frac{2a}{p} \int_{-a}^{a} \frac{x^2}{a^2 + x^2} dx$$

$$= \frac{4a}{p} \int_{0}^{a} \frac{x^2}{a^2 + x^2} dx$$

$$= \frac{4a}{p} \int_{0}^{a} \left[1 - \frac{a^2}{x^2 + a^2} \right] dx$$

$$= \frac{4a}{p} \left[x - a^2 \tan^{-1} \frac{x}{a} \right]_0^a$$

$$= \frac{a^2}{p} [4 - p]$$ **[1]**

$$\sigma^2 = EX^2 - (EX)^2 = \frac{a^2}{p} [4 - p]$$ **[1]**

Total [3]

Q.6) Given $X_i \sim G(\theta)$; i = 1,2,…, k

Then each X_i has pgf

$$G(t) = \frac{q}{[1 - (1 - q)t]}$$ **[1]**

Let Y = $X_1 + X_2 + \dots + X_k$

Then pgf of Y is $\left[\dfrac{q}{1-(1-q\,t)}\right]^k$ **[1]**

which is the pgf of negative Binomial distribution with parameters(k,θ)

Total [2]

Q.7) Invoking Hyper geometric distribution with

x = 2, n = 5, N = 120 and M = 80,

$$p(x,n,N,M) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{x}}$$ **[1]**

$$= \frac{\binom{80}{2}\binom{40}{3}}{\binom{120}{5}} = 0.164$$ **[1]**

Total [2]

Q.8)

(a) Postulates : Let X(t) denotes number of occurrences in (0,t)

1) Probability of one occurrence in the interval of length $\Delta\, t$ is $\lambda\;\Delta\; t + o(\Delta t)$
2) Probability of no occurrence in $\Delta\, t$ is 1-$\lambda\;\Delta\; t + o(\Delta\; t)$
3) Probability of more than one occurrences in $\Delta\, t$ is $o(\Delta\; t)$
4) X(t) is independent of the number of occurrences of the event in any interval prior and after (0,t)

[1]

(b) Since $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$,

P[N_1(t) = k/N_1(t) + N_2 (t) = n]

$$= \frac{P\left[\{N_1(t) = k\} \cap \{N_1(t) + N_2(t) = n\}\right]}{P\left[N_1(t) + N_2(t) = n\right]}$$

$$= \frac{P\left[\{N_1(t) = k\} \cap \{N_2(t) = n-k\}\right]}{P\left[N_1(t) + N_2(t) = n\right]}$$ **[2]**

$$= \left[\frac{e^{-l_1 t}(l_1 t)^k}{k!} \quad \frac{e^{-l_2 t}(l_2 t)^{n-k}}{(n-k)!}\right] \quad \left[\frac{e^{-(l_1+l_2)t} \qquad \{(l_1 + l_2)t\}^n}{n!}\right]$$

$$= \frac{n!}{k!(n-k)!} \; \frac{(l_1 t)^k (l_2 t)^{n-k}}{\{(l_1 + l_2)t\}^n}$$

$$= \binom{n}{k} \left(\frac{l_1}{l_1 + l_2} \right)^k \left(\frac{l_2}{l_1 + l_2} \right)^{n-k} ; \; k = 0,1,2,\ldots,n$$ **[2]**

$$= \binom{n}{k} \; p^k \quad q^{n-k}$$

where $p = \left(\frac{l_1}{l_1 + l_2} \right)$ *and* $q = \left(\frac{l_2}{l_1 + l_2} \right)$

Total [5]

Q.9)

i.) $E(X) = 0.\frac{5}{12} + 1\left(\frac{1}{2}\right) + 2\left(\frac{1}{12}\right) = 2/3$

$$E(Y) = 0.\frac{7}{12} + 1\left(\frac{7}{18}\right) + 2\left(\frac{1}{36}\right) = 4/9$$ **[2]**

E(XY) = 1/6

ii.) Cov(X,Y) = E(XY) – (EX) (E(Y))
= 1/6 – (2/3) (4/9) = -7/54 **[1]**

iii.) Conditional distribution of given Y = 1

$$\left.\begin{array}{c} P(X = 0/Y = 1) = (2/9) \div (7/18) = 4/7 \\ P(X = 1/Y = 1) = (1/6) \div (7/18) = 3/7 \\ P[X = 2/Y = 1] = 0 \div (7/18) = 0 \end{array}\right\}$$ **[2]**

$EX^2 = 0^2 (5/12) + 1^2 (1/2) + 2^2 (1/12) = 5/6$

$$\therefore s_x^2 = EX^2 - (EX)^2 = \frac{5}{6} - (2/3)^2 = 7/18$$ **[1]**

$EY^2 = 0^2 (7/12) + 1^2 (7/18) + 2^2 (1/36)$
$= 7/18 + 1/9 = ½$

$$s_y^2 = EY^2 - (EY)^2 = \frac{1}{2} - \frac{16}{81} = \frac{49}{162}$$ **[1]**

$$r(X,Y) = \frac{(-7/54)}{\sqrt{\frac{7}{18} X \frac{49}{162}}} = -\frac{\sqrt{6}}{9\sqrt{7}}$$ **[1]**

Total [8]

Q.10) X ~ B(n,p)
n = 225
p – probability of refusing loan application p = 0.2
For a Binomial distribution
Mean = np = 225 x 0.2 = 45 = μ (say)
variance = npq = 225 x 0.2 x 0.8 = 36 = σ^2 (say) **[1]**

$$Z = \frac{X - np}{\sqrt{npq}} = \frac{X - 45}{6} \sim N(0,1)$$ **[1]**

Applying continuity correction, one has to find P(X < 40.5).

$$Z = \frac{40.5 - 45}{6} = -0.75$$

$\therefore$ P(Z < -0.75) = 0.5 – 0.2734 = 0.2266 **[2]**

Total [4]

Q.11) n = 5, $\sigma^2 = 25$

we know that $\frac{(n-1)s^2}{s^2} \sim c^2(n-1)$ **[1]**

$$\therefore \frac{(5-1)s^2}{25} = \frac{4}{25}s^2 \sim c^2(4)$$

Let Y = $\frac{4}{25}s^2$

$\therefore$ f(y) = $\frac{1}{4\Gamma 2} y\, e^{-y/2} = \frac{1}{4}\, y\, e^{-y/2}$ $(\because \Gamma 2 = 1)$ **[1]**

when $s^2 = 20$, Y = $\frac{80}{25} = 3.2$ and

when $s^2 = 30$, Y = $\frac{120}{25} = 4.8$

$\therefore$ It is required to compute

P[20 < s^2 < 30] = P[3.2 < y < 4.8]

$$= \frac{1}{4}\int_{3.2}^{4.8} y\, e^{-Y/2}\, dy$$

$$= \left[-e^{-Y/2}\left(\frac{y}{2} + 1\right) \right]_{3.2}^{4.8}$$

$$= -3.4\, e^{-2.4} + 2.6\, e^{-1.6}$$ **[1]**

$$= -3.4(0.091) + 2.6((0.202) = 0.216$$

Total [3]

Q.12) P(no claims) = $e^{-\lambda}$ **[1]**

Observed proportion of “no claims” = 0.78 **[1]**

$\therefore$ MLE of e^{λ} is 0.78 so MLE of λ is –log(0.78) = 0.248 **[1]**

Total [3]

Q.13) Total # claims $\sum_{i=1}^{200} X_i \sim P(200\, l)$ **[1]**

$$\therefore \hat{l} = \overline{X} \sim N\left(l, \tfrac{l}{200}\right)$$ **[1]**

$$\therefore P\left(\frac{\hat{l} - l}{\sqrt{l/200}} < 1.645 \right) = 0.95$$

$$\therefore P\left[\hat{l} < l + 1.645\sqrt{\frac{l}{200}}\right] = 0.95$$

$$\therefore P\left[\hat{l} - 1.645\sqrt{\frac{\hat{l}}{200}} < l\right] \cong 0.95$$ **[2]**

observed $\hat{l} = \frac{52}{200} = 0.26$

so lower limit is $0.26 - 1645\sqrt{0.0013}$

$= 0.26 - 0.059 = 0.201$ **[1]**

so $(0.201, \infty)$

Total [5]

Q.14) $P(Y < y \mid H_0) = P\left(Z = \frac{y-10}{3} < -1.645\right) = 0.95$

determines the critical region, which is y < 10 – 4.935 = 5.065 **[2]**

we want

$$1 - P(Y < 5.065 \mid H_1) = 1 - P\left(Z < \frac{5.065 - 4}{\sqrt{4x.96}}\right) = 1 - \Phi(0.543)$$

= 1 – 0.71 = 0.29 **[2]**

OR in practice, we reject for Y ≤ 5, so we want

$$1 - P(Y \le 5 \mid H_1) = 1 - P\left(Z < \frac{5.5 - 4}{\sqrt{4x.96}}\right) = 1 - \Phi(0.765)$$

= 1 – 0.78 = 0.22

Total [4]

Q.15)

e_{ij}

55.3	29.7	11.1
123.2	66.2	24.6
61.6	33.1	12.3

[2]

$o_{ij} - e_{ij}$

7.7	−3.7	−4.1
2.8	−3.2	0.4
−10.6	6.9	3.7

[1]

$$\frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

1.072	0.461	1.514
0.064	0.155	0.007
1.824	1.438	1.113

[1]

Cal χ^2 for 4df : 7.65

Tab : χ^2 for 4df at 5% level 9.488

$\therefore$ No real evidence agains H_0 : operationand side-effect are independent. **[1]**

Combine slight/moderate categories.

o_{ij}		e_{ij}		$o_{ij} - e_{ij}$	
63	33	55.3	40.8	7.7	−7.8
126	88	123.2	90.8	2.8	−2.8
51	56	61.6	45.4	−10.6	10.6

[2]

$\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$

1.072	1.491
0.064	0.086
1.824	2.475

Cal χ^2 for 2df : 7.01
Tab : χ^2 for 2df at 5% 5.991 **[1]**

Some evidence to suggest that presence / absence of side-effect depends on operation.

Total [8]

Q.16) For the data given in the problem,

a) $\overline{X}_1 = 35.22, \quad \overline{X}_2 = 31.56$

$n_1 = n_2 = 9$

Further it is given that

$$\sum_{i=1}^{9}(X_1 - \overline{X}_1)^2 = 195.5556$$

$$\sum_{i=1}^{9}(X_2 - \overline{X}_2)^2 = 160.2222$$

$$s^2 = \frac{\Sigma(X_1 - \overline{X}_1)^2 + \Sigma(X_2 - \overline{X}_2)^2}{n_2 + n_2 - 2}$$

$$= \frac{195.5556 + 160.2222}{9+9-2} = 22.2361$$

$\therefore$ s = 4.716 **[1]**

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{35.22 - 31.56}{4.716\sqrt{\frac{1}{9} + \frac{1}{9}}} = 1.65$$ **[1]**

Table t value = $t_{0.05}$ (16) = 1.746

There is some evidence that the mean time to assemble the device under the both procedure are equal **[1]**

b) 95% confidence interval for the difference in population mean is

$$\left(\overline{X}_1 - \overline{X}_2\right) \pm t_{a/2}\ s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$ **[1]**

substituting the values,

$$(35.22 - 31.56) \pm 2.120\ (4.716)\ \sqrt{\frac{1}{9}+\frac{1}{9}}$$

= 3.66 ± 4.71

= (-1.05, 8.37) **[1]**

c) $F = \frac{s_1^2}{s_2} = \frac{195.5556}{160.2222} = 1.22$ **[2]**

Table $F_{8,8}$ at 5% is 3.44

Accept H_0 the variances are equal. **[1]**

Total [8]

Q.17)

a)

```
   -                                                      *
   -
120+                                                 *
   -
   -
   -                                            *
   -                                  *
 90+
   -
   -                        *    *
   -
   -                   *
 60+
   -         *
   -              *
   -    *
   -
     --------+---------+---------+---------+---------+--------C1
           1.0       2.0       3.0       4.0       5.0
```

Linear model looks reasonable. **[1]**

b) $S_{xx} = 110.5\ \frac{-29^{2}}{10} = 26.4$ **[1]**

$S_{yy} = 73225\frac{-805^{2}}{10} = 8422.5$ **[1]**

$S_{xy} = 2795\ \frac{29.805}{10} = 460.5$ **[1]**

Call-out charge is intercept, hourly rate is slope. **[1]**

c) $\hat{b} = \frac{460.5}{26.4} = Rs.17.443$ (hourly rate) **[1]**

$\hat{a} = 1/10\{805 - 17.443(29)\} = \text{Rs.}29.915$ (call-out charge) **[1]**

Need $s^2 = \frac{1}{8}\left(84225 - \frac{460.5^2}{26.4}\right) = 48.739$ **[1]**

$\therefore$ s.e. $\hat{b} = \sqrt{\frac{48.739}{26.4}} = 1.359$

$\therefore$ 90% CI : 17.443 ± 1,860(1,359) **[1]**
= 17.44 ± 2.53
= (Rs.14.91, Rs.19.97)

d) $r = \frac{S_{xy}}{\sqrt{S_{xn}S_{xy}}} = 0.97$ **[2]**

$H_0 : \rho = 0$ $H_1 : \rho \neq 0$

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$ **[1]**

$t = \frac{0.97\sqrt{8}}{\sqrt{1-0.94}} \sim 12.9$ **[1]**

t value for 8df at 5% level 2.31 reject $\rho = 0$

Total [13]

Q.18)

a) $\Sigma X = 10216$, $\Sigma X^2 = 2621210$ **[1]**

$SS_T = 2621210 - \frac{10216}{40} = 12043.6$ **[1]**

$SS_R = \frac{2479 + 2619 + 2441 + 2677}{10} - \frac{10216}{40} = 3774.8$ **[1]**

$\therefore SS_E = 8268.8$ by subtraction.

Source	df	SS	MS
Regions	3	3774.8	1258.3
Residual	36	8268.8	229.7
Total	39	12043.6	

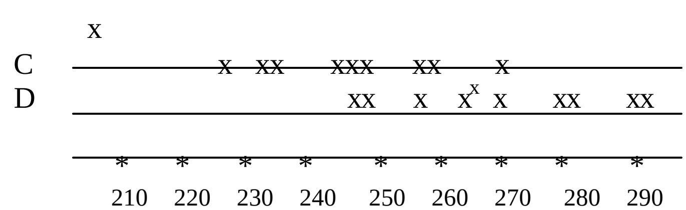
[2]

$F = \frac{1258.3}{229.7} = 5.48$ on (3.36) df **[1]**

$F_{3,36}$ at 5% = 2.9 (app)

reject $H_0 : \mu_A = \mu_B = \mu_C = \mu_D$ at the 5% level **[1]**

b)



[1]

Normality – OK
Equal variances for C and D.

(c) $\dfrac{36\hat{\sigma}^2}{\sigma^2} \sim \chi^2$ where $36\hat{\sigma}^2 = SS_R$

$\therefore$ 95% CI for σ^2 is $\left(\dfrac{SS_R}{\chi^2_{0.975,36}} - \dfrac{SS_R}{\chi^2_{0025.36}}\right)$ **[1]**

$= \left(\dfrac{8268.8}{54.4}, \dfrac{8268.8}{21.37}\right) = (152.0, 386.9)$ (Interpolate in tables)

$\therefore$ 95% CI for σ is (12.3, 19.7) **[1]**

Total [10]

Q.19)

a) $E(NS_N / N = n)$
$= E(N(X_1 + \ldots + X_N) / N = n)$
$= E(n(X_1 + \ldots + X_n) / N = n)$
$= E(n(X_1 + \ldots + X_n))$ (since N & $X_1 + \ldots + X_n$ are independent)
$= n^2 \mu_x$ **[2]**

$$E(NS_N) = \sum_n E(NS_N / N = n)\, P(N = n)$$
$$= \sum_n n^2 \mu_x P(N = n)$$
$$= \mu_x \sum_n n^2 P(N = n)$$
$$= \mu_x EN^2 = \mu_x (Var\, N + (EN)^2)$$
$$\mu_x \left(\sigma_N^2 + \mu_N^2\right)$$
[1]

$EN = \mu_N$. It can be shown that
$ES_N = EN\, EX_1$
$= \mu_N \mu_x$ **[1]**

$Cov(NS_N)$
$= E(NS_N) - EN\, E\, S_N$
$= \mu_x\left(\mu_N^2 + \sigma_N^2\right) - \mu_N . \mu_N \mu_x$
$= \mu_x \sigma_N^2$ **[1]**

Total [5]

Q.20) The conditional density of X given Y =1/2 is

$$f(x/y) = \frac{2x + 4y}{1 + 4y} \quad 0 < x < 1$$
$= 0$ elsewhere **[1]**

so that $f\left(x / \dfrac{1}{2}\right) = \dfrac{2}{3}(x + 1)$; $0 < x < 1$
$= 0$ elsewhere

$$\therefore E\left(X / \frac{1}{2}\right) = \int_0^1 \frac{2}{3} x(x+1)\,dx$$
$= 5/9$ **[1]**

$$E\left(X^2 / \frac{1}{2}\right) = \int_0^1 \frac{2}{3} x^2 (x+1) dx$$

$= 7/18$ **[1]**

$$V\left(X / \frac{1}{2}\right) = \frac{7}{18} - (5/9)^2$$

$= \frac{13}{162}$ **[2]**

Total [5]

****