Subject CT3 — Probability and Mathematical Statistics Core Technical

EXAMINERS' REPORT

April 2009

Introduction

The attached subject report has been written by the Principal Examiner with the aim of helping candidates. The questions and comments are based around Core Reading as the interpretation of the syllabus to which the examiners are working. They have however given credit for any alternative approach or interpretation which they consider to be reasonable.

R D Muckart
Chairman of the Board of Examiners

June 2009

Comments

The paper was answered quite well overall. Some questions were answered less well (or by noticeably fewer candidates) than others – they were:

Question 7(i) – confidence intervals based on small samples
Question 9(ii) – differences between pairs of treatment means in an ANOVA context
Question 11(i) – deciding whether or not certain variables derived from Poisson variables are themselves Poisson variables
Question 12 – writing down the correct likelihood function, in which the stated probabilities are *raised to powers* given by the observed frequencies of occurrence (not *multiplied* by the frequencies)

There were no other misunderstandings widely evident, and no particular errors were made so repeatedly as to be worthy of comment.

1 Revised mean = (216.9 – 3.1 – 46.2)/10 = 16.76 i.e. £16,760

Revised median = original median = £16,900

Revised $\Sigma x^2 = 5052.13 - 3.1^2 - 46.2^2 = 2908.08$

so revised standard deviation = $\left\{\frac{1}{9}\left(2908.08 - \frac{167.6^2}{10}\right)\right\}^{1/2} = 3.31837$ i.e. £3,318.37

2 $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
$= 0.3 + 0.5 + 0.2 - 0.1 - (0.3 \times 0.2) = 0.84$

(OR via a Venn diagram)

So P(none occur) = 1 – 0.84 = 0.16

3 (i) $\int_0^1 f(x)dx = 1 \Rightarrow \int_0^1 k(1-x^2)dx = 1 \Rightarrow k\left[x - \frac{x^3}{3}\right]_0^1 = 1 \Rightarrow k\left(1 - \frac{1}{3}\right) = 1 \Rightarrow k = 1.5$

(ii) $P(X > 0.25) = \int_{0.25}^1 f(x)dx$

$= \int_{0.25}^1 1.5(1-x^2)dx = 1.5\left[x - \frac{x^3}{3}\right]_{0.25}^1 = 1.5 \times 0.422 = 0.633.$

4 (i) $E[Y_i] = 2$, $Var[Y_i] = 4$

Therefore $E[Y_1 + Y_2] = 4$ and (since Y_1, Y_2 independent) $Var[Y_1 + Y_2] = 8$.

So, for total loss, mean = £4,000 and variance = 8×10^6 (£^{2}).
(OR from $Y_1 + Y_2 \sim \chi_4^2$)

(ii) For the total loss we have $Y_1 + Y_2 \sim \chi_4^2$, so we want a constant k such that $P\left(\chi_4^2 > k\right) = 0.95$.

From tables of the χ_4^2 distribution $P\left(\chi_4^2 > 0.7107\right) = 0.95$.

$\therefore$ The total losses will exceed £710.7 with probability 0.95.

5 Let X be the number still in employment after one year.

$\therefore X \sim bin(280, 0.82) \approx N(229.6, 6.43^2)$

$P(X \geq 240) = P(X > 239.5)$ applying a continuity correction

$$= P(Z > \frac{239.5 - 229.6}{6.43}) = P(Z > 1.54) = 1 - 0.93822 = 0.062$$

6 (i) $E[X_i] = 4/0.5 = 8$ and $Var[X_i] = 4/0.5^2 = 16$ (or by noting $X_i \sim \chi_8^2$).

Using the CLT: $\bar{X} = \dfrac{\sum_{i=1}^{40} X_i}{40} \approx N\left(E[X_i], \dfrac{Var[X_i]}{40}\right)$, i.e. $N(8, 0.4)$ approximately.

[Note: The exact distribution of $\bar{X}$ is Gamma(160,20)]

(ii) The symmetry of the distribution gives: median$\left[\bar{X}\right]$ = mean$\left[\bar{X}\right]$ = 8, i.e. £800.

7 (i) For a single observation x from Poisson(λ) a 95% confidence interval for λ is (λ_1, λ_2) where

$$\sum_{r=x}^{\infty} p(r;\lambda_1) = 0.025 \quad \text{and} \quad \sum_{r=0}^{x} p(r;\lambda_2) = 0.025$$

So for $x = 2$

$$\lambda_1 \text{ is s.t. } \sum_{r=2}^{\infty} p(r;\lambda_1) = 0.025 \qquad \text{i.e. } \sum_{r=0}^{1} p(r;\lambda_1) = 0.975$$

From tables λ_1 is between 0.20 and 0.30 being about 0.24.

$$\lambda_2 \text{ is s.t. } \sum_{r=0}^{2} p(r;\lambda_2) = 0.025$$

From tables λ_2 is between 7.00 and 7.25 being about 7.23.

95% confidence interval for λ is (0.24, 7.23).

(ii) For a sample of n with observed mean $\bar{x}$

$$\frac{\bar{X}-\lambda}{\sqrt{\frac{\hat{\lambda}}{n}}} \approx N(0,1) \quad \text{where } \hat{\lambda} = \bar{X} \qquad \left[\text{OR: could use } \frac{\Sigma X - n\lambda}{\sqrt{n\hat{\lambda}}} \approx N(0,1)\right]$$

giving an approximate 95% confidence interval as $\bar{X} \pm 1.96\sqrt{\frac{\bar{X}}{n}}$

So for $n = 30$, $\bar{x} = 2.4$

95% CI is $2.4 \pm 1.96\sqrt{\frac{2.4}{30}} \Rightarrow 2.4 \pm 0.55 \Rightarrow (1.85, 2.95)$

(iii) The CI in (ii) is much narrower due to having more data.

(It is also centred higher due to the larger estimate.)

8 Let σ^2 be the population variance.

$$\frac{24S^2}{\sigma^2} \sim \chi^2_{24} \Rightarrow P\left(\frac{24S^2}{\sigma^2} > 13.85\right) = 0.95$$

$$\Rightarrow P\left(\sigma^2 < \frac{24S^2}{13.85}\right) = 0.95 \Rightarrow k^2 = 24 \times 2.105^2 / 13.85 = 7.678$$

$\Rightarrow k = 2.771$ so upper confidence limit for σ is 2.771 i.e. £2771
(OR: CI is (0, 2.771) i.e. (0, £2771)).

9 (i) $F = \frac{2239}{120} = 18.66$ on 3,28 d.f.

from tables $F_{3,28}(1\%) = 4.568$

$\therefore$ P-value < 0.01 or $0 < P$-v < 0.01

So there is overwhelming evidence of a difference between the underlying treatment means.

(ii) (a) LSD = $t_{28}(2.5\%)\sqrt{\hat{\sigma}^2(\frac{1}{8}+\frac{1}{8})} = 2.048\sqrt{120(\frac{1}{8}+\frac{1}{8})} = 11.2$

(b) means in order $\underline{\bar{y}_{3.} < \bar{y}_{2.}} < \underline{\bar{y}_{1.} < \bar{y}_{4.}}$
underlined thus:

Treatments 2 & 3 are separate from treatments 1 & 4 which have significantly higher means.

10 $E[N] = 0.7(0) + 0.15(1) + 0.1(2) + 0.05(3) = 0.5$

$E[N^2] = 0.15(1) + 0.1(4) + 0.05(9) = 1.00 \quad \therefore Var[N] = 1.00 - 0.5^2 = 0.75$

$E[X] = \frac{2}{0.1} = 20, \; Var[X] = \frac{2}{0.1^2} = 200$

Let S = total of the claim amounts.

$E[S] = E[N]E[X] = (0.5)(20) = 10$, i.e. £10,000.

$V[S] = E[N]Var[X] + Var[N][E(X)]^2 = (0.5)(200) + (0.75)(20)^2 = 400$

$\therefore sd(S) = 20$, i.e. £20,000.

11 (i) (a) Let $S = \sum_{i=1}^{n} X_i$

$M_X(t) = \exp\{\lambda(e^t - 1)\}$

$\Rightarrow M_S(t) = \{M_X(t)\}^n = \left[\exp\{\lambda(e^t-1)\}\right]^n = \exp\{n\lambda(e^t-1)\}$

$\Rightarrow S \sim$ Poisson($n\lambda$)

(b) No

One reason is that $E[2X_1 + 5] = 2\lambda + 5$, which is not equal to $V[2X_1 + 5] = 4\lambda$

[Note: another obvious reason is that $2X_1 + 5$ can only takes values 5, 7, 9, … , not 0, 1, 2, 3,…]

(c) No

One reason is that $E\left[\bar{X}\right]=\lambda$, which is not equal to $V\left[\bar{X}\right]=\lambda/2$

[Note: another obvious reason is that $\bar{X}$ can take values 0.5, 1.5, 2.5, … , which a Poisson variable cannot.]

(d) $\bar{X} \approx N\left(\lambda, \dfrac{\lambda}{n}\right)$

(ii) (a) Under H_0, $\bar{X} \sim N\left(1, \dfrac{1}{100}\right)$ approximately

k is such that $P\left(\bar{X} > k \mid H_0\right) = 0.01$ so $\dfrac{k-1}{0.1} = 2.3263$

$\Rightarrow k = 1.2326$

(b) Power(λ) = P(reject $H_0|\lambda$)

Power(λ = 1.2) $= P\left(\bar{X} > 1.2326\right)$ where $\bar{X} \sim N(1.2, 0.012)$
$= P(Z > 0.298) = 0.383$

Power(λ = 1.5) $= P\left(\bar{X} > 1.2326\right)$ where $\bar{X} \sim N(1.5, 0.015)$
$= P(Z > -2.183) = 0.985$

(c) Power of test increases as the value of λ increases further away from $\lambda = 1$.

12 (i) (a) $L(\theta) = [\frac{1}{4}(2+\theta)]^{1071}[\frac{1}{4}(1-\theta)]^{62}[\frac{1}{4}(1-\theta)]^{68}[\frac{1}{4}\theta]^{299}$ ($\times$ *constant*)

$$\propto (2+\theta)^{1071}(1-\theta)^{130}\theta^{299}$$

$$\log L(\theta) = const + 1071\log(2+\theta) + 130\log(1-\theta) + 299\log\theta$$

(b) $\dfrac{d}{d\theta}\log L(\theta) = \dfrac{1071}{2+\theta} - \dfrac{130}{1-\theta} + \dfrac{299}{\theta}$

$$= \frac{1071\theta(1-\theta) - 130\theta(2+\theta) + 299(2+\theta)(1-\theta)}{(2+\theta)(1-\theta)\theta}$$

numerator = $1071\theta - 1071\theta^2 - 260\theta - 130\theta^2 + 598 - 299\theta - 299\theta^2$

equate to zero: $750\theta^2 - 256\theta - 299 = 0$ 1

$$\therefore \hat{\theta} = \frac{256 \pm \sqrt{256^2 - 4(750)(-299)}}{2(750)} = 0.17067 \pm 0.65406$$

So MLE $\hat{\theta} = 0.82473$ (or 0.825 to 3dp) as other root is negative.

(ii) (a) $$\frac{d^2}{d\theta^2}\log L(\theta) = -\frac{1071}{(2+\theta)^2} - \frac{130}{(1-\theta)^2} - \frac{299}{\theta^2}$$

at $\hat{\theta} = 0.825$, $\frac{d^2}{d\theta^2}\log L(\theta) = -134.20 - 4244.90 - 439.30 = -4818.4$

$$\text{CRlb} = \frac{1}{-E\left[\frac{d^2}{d\theta^2}\log L(\theta)\right]} \approx \frac{1}{4818.4} = 0.0002075$$

(b) $\hat{\theta} \approx N(\theta, CRlb)$ for large samples

and so an approximate 95% CI for θ is $\hat{\theta} \pm 1.96\sqrt{CRlb}$

Here: $0.825 \pm 1.96\sqrt{0.0002075} \Rightarrow 0.825 \pm 0.028$ or (0.797, 0.853)

(iii) (a) With $\theta = 0.775$ the four probabilities are

0.69375, 0.05625, 0.05625, 0.19375 respectively

and the corresponding expected frequencies are

1040.625, 84.375, 84.375, 290.625.

$$\therefore \chi^2 = \sum \frac{(o-e)^2}{e} = 0.887 + 5.934 + 3.178 + 0.241 = 10.24 \quad \text{on 3 df}$$

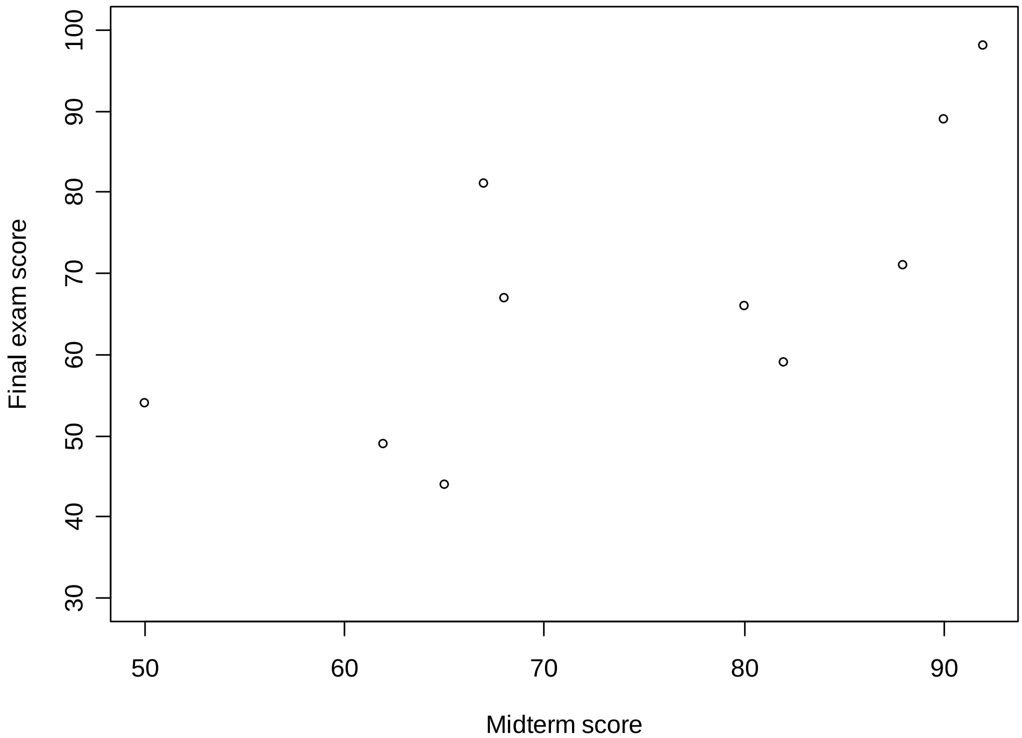
P-value = $P(\chi_3^2 > 10.24) = 1 - 0.983 = 0.017$

These data do not support the model with the value $\theta = 0.775$ in that the probability of observing these data when $\theta = 0.775$ is only 0.017.

[OR: could say "do not support at 5% level, but do support at 1% level"]

(b) This is consistent with the fact that $\theta = 0.775$ is well outside the approximate 95% CI.

13 (i) (a) The plot is given below.



There seems to be a positive relationship between final and midterm score. However it is not clear if this relationship is linear.

[Following comment also valid: the relationship looks linear but with substantial scatter.]

(b) $$\hat{\sigma}^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right) = \frac{1}{8}\left(2737.6 - \frac{(1529.8)^2}{1760.4}\right) = 176.0241$$

$$\text{s.e.}(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\frac{176.0241}{1760.4}} = 0.3162$$

To test $H_0 : \beta = 0$ v $H_1 : \beta > 0$, the test statistic is

$$\frac{\hat{\beta} - 0}{\text{s.e.}(\hat{\beta})} = \frac{0.869}{0.3162} = 2.748,$$

and under the assumption that the errors of the regression are *i.i.d.* $N(0, \sigma^2)$ random variables, it has a *t* distribution with $n - 2 = 8$ df.

From statistical tables we find $t_{8,0.05} = 1.860$, and $t_{8,0.01} = 2.896$.

We reject the hypothesis H_0: β = 0 in favour of H_1: β > 0 at the 5% level (but not at the 1% level).

(c) The mean response, $\hat{y}_{new}$ for $\hat{x}_{new} = 75$ is $\hat{y}_{new} = 3.146 + 0.869 \times 75 = 68.321$.

Its standard error is calculated as

$$\text{s.e.}(\hat{y}_{new}) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_{new} - \bar{x})^2}{S_{xx}}}$$

$$= 13.26741\sqrt{\frac{1}{10} + \frac{(75 - 74.4)^2}{1760.4}} = 13.26741 \times 0.31655$$

$$= 4.1998$$

The 95% CI is given by $\hat{y}_{new} \pm t_{0.025,8} \times \text{s.e.}(\hat{y}_{new})$,

i.e. $68.321 \pm 2.306 \times 4.1998 = 68.321 \pm 9.6847$

or (58.636, 78.006)

(d) The CI for the individual predicted score will be wider than the CI for the mean score in (i)(c), because the variance for the individual predicted value is larger.

(ii) (a) Both statistics follow a t_{n-2} distribution under the null hypothesis.

In addition

$$\frac{\hat{\beta} - 0}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}} = \frac{\dfrac{S_{xy}}{S_{xx}}}{\sqrt{\dfrac{1}{(n-2)S_{xx}}\left(S_{yy} - \dfrac{S_{xy}^2}{S_{xx}}\right)}} = \frac{\sqrt{(n-2)S_{xx}}\dfrac{S_{xy}}{S_{xx}}}{\sqrt{S_{yy}}\sqrt{1 - \dfrac{S_{xy}^2}{S_{xx}S_{yy}}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

(b) $r = \dfrac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \dfrac{1529.8}{\sqrt{1760.4 \times 2737.6}} = 0.6969.$

Then $\dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \dfrac{0.6969 \times \sqrt{8}}{\sqrt{1 - 0.6969^2}} = 2.748$, same as in (i)(b).

END OF EXAMINERS' REPORT