Imperial College London

MATH50011

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May-June 2021

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

Statistical Modelling 1

Date: Friday, 21 May 2021

Time: 09:00 to 11:00

Time Allowed: 2 hours

Upload Time Allowed: 30 minutes

This paper has 4 Questions.

Candidates should start their solutions to each question on a new sheet of paper.

Each sheet of paper should have your CID, Question Number and Page Number on the top.

Only use 1 side of the paper.

Allow margins for marking.

Any required additional material(s) will be provided.

Credit will be given for all questions attempted.

Each question carries equal weight.

SUBMIT YOUR ANSWERS AS SEPARATE PDFs TO THE RELEVANT DROPBOXES ON BLACKBOARD (ONE FOR EACH QUESTION) WITH COMPLETED COVERSHEETS WITH YOUR CID NUMBER, QUESTION NUMBERS ANSWERED AND PAGE NUMBERS PER QUESTION.

1. Let $Y_1, \ldots, Y_n$ be independent and identically distributed (i.i.d.) $N(\mu, \sigma^2)$ with $\sigma^2 > 0$ known. Suppose that interest lies in estimating $\theta = P(Y_i \leq 0)$.

(a) State the maximum likelihood estimator (MLE) $\hat{\mu}$ for μ based on $Y_1, \ldots, Y_n$ and state the distribution for Z_a such that $\sqrt{n}(\hat{\mu} - \mu) \to_d Z_a$. (2 marks)

(b) Find the MLE $\hat{\theta}_Y$ for θ based on $Y_1, \ldots, Y_n$ and derive the distribution for Z_b such that $\sqrt{n}(\hat{\theta}_Y - \theta) \to_d Z_b$.

(6 marks)

(c) For each $i = 1, \ldots, n$, define

$$W_i := I(Y_i \leq 0) = \begin{cases} 1 & \text{if } Y_i \leq 0, \\ 0 & \text{else.} \end{cases}$$

What is the distribution of W_i? (2 marks)

(d) State the MLE $\hat{\theta}_W$ for θ based on $W_1, \ldots, W_n$ and state the distribution for Z_d such that $\sqrt{n}(\hat{\theta}_W - \theta) \to_d Z_d$.

(4 marks)

(e) Show that $Var(Z_b) \leq Var(Z_d)$, where Z_b and Z_d follow the distributions identified in parts (b) and (d). (6 marks)

(Total: 20 marks)

2. The distribution of the concentration S (in milligrams per litre) of an enzyme in a certain biological system is assumed to be adequately described by the density function

$$f_S(s;\theta) = 3\theta^{-3}s^2, \qquad 0 < s < \theta$$

where $\theta > 0$ is an unknown parameter. A biologist is interested in making statistical inferences about θ.

Suppose we have available n i.i.d. random variables S_i, $i = 1, \dots, n$.

(a) (i) Show that $S_{(n)} := \max\{S_1, \dots, S_n\}$ has density function

$$f_{S_{(n)}}(s;\theta) = 3n\theta^{-3n}s^{3n-1}, \quad 0 < s < \theta.$$

(4 marks)

(ii) Using $S_1, S_2, \dots, S_n$, construct an exact $100(1-\alpha)\%$ upper one-sided confidence interval $(0, U)$ for the unknown parameter θ, where U is a function of $S_{(n)} := \max\{S_1, \dots, S_n\}$.
(4 marks)

(b) (i) Use the central limit theorem to show that

$$\sqrt{n}\left(\bar{S} - \frac{3}{4}\theta\right) \to_d N\left(0, \frac{3}{80}\theta^2\right).$$

(4 marks)

(ii) Using $S_1, S_2, \dots, S_n$, construct a large-sample $100(1-\alpha)\%$ two-sided confidence interval for the unknown parameter θ based on $\bar{S} = \frac{1}{n}\sum_{i=1}^{n} S_i$. (4 marks)

(c) Construct a level α hypothesis test of $H_0 : \theta \le \theta_0$ against $H_1 : \theta > \theta_0$ based on part (a).
(4 marks)

(Total: 20 marks)

3. For $i = 1, \ldots, n$, consider the linear models

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{1}$$
$$x_i = \delta_0 + \delta_1 z_i + \eta_i. \tag{2}$$

In matrix form we have $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ and $\mathbf{x} = \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\eta}$. Suppose that $\mathbf{Z}^T\mathbf{X}$ is non-singular, that $\mathbf{Z}$ and $\mathbf{X}$ have full column rank, and that the errors in each model are i.i.d. with $Var(\epsilon_i) = \sigma^2$ and $Var(\eta_i) = \tau^2$.

Unless requested otherwise, express your responses in vector/matrix notation.

(a) Derive expressions for the hat matrix, $\mathbf{P}$, and the vector of fitted values, $\hat{\boldsymbol{x}}$, from least squares regression based on model (2). (3 marks)

(b) Replace x_i by $\hat{x}_i$ in model (1) and show that the least squares estimator using the $\hat{x}_i$ is

$$\tilde{\boldsymbol{\beta}} = (\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{Y}.$$

(4 marks)

(c) Find $E(\tilde{\boldsymbol{\beta}})$ and $Cov(\tilde{\boldsymbol{\beta}})$, treating the x_i and z_i as known constants. (6 marks)

(d) Show that the estimator of β_1 obtained in part (c) is

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^n (z_i - \bar{z})(Y_i - \bar{Y})}{\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})},$$

where $\bar{Y}, \bar{x}$, and $\bar{z}$ are the usual sample means. (3 marks)

(e) Let $\hat{\boldsymbol{\beta}}$ denote the usual least squares estimator of $\boldsymbol{\beta}$ based on model (1). Show that

$$Var(\hat{\beta}_1) \leq Var(\tilde{\beta}_1),$$

treating the x_i and z_i as known constants. (4 marks)

(Total: 20 marks)

4. Suppose we wish to investigate the relationship between the daily cases of Salmonella and season (autumn, winter, spring, summer) in a particular city based on $n = 365$ consecutive days. Suppose further that we perform our analysis modelling the logarithm of the number of Salmonella cases as a function of season, rainfall (yes/no), and a possible season-rainfall interaction.

(a) Describe a linear model you could use to address this problem, clearly defining the predictor variables and their corresponding model parameters. Assume that the error terms are i.i.d. $N(0, \sigma^2)$ for some σ^2. (6 marks)

(b) Suppose we wish to test for the existence of a seasonal effect on Salmonella cases. State the hypotheses you would test in terms of your model parameters, the form of the test statistic you would use to test those hypotheses, and the methods you would use to determine a rejection region for the test. Provide explicit formulas wherever possible (matrix notation is allowed, provided the matrices are adequately defined). (5 marks)

(c) Suppose instead that we merely want to determine whether a season-rainfall interaction on Salmonella cases exists. State the hypotheses you would test in terms of your model parameters, the form of the test statistic you would use to test those hypotheses, and the methods you would use to determine a rejection region for the test. Provide explicit formulas wherever possible (matrix notation is allowed, provided the matrices are adequately defined). (5 marks)

(d) A more realistic distribution for the errors is proposed to account for the correlation between days. Let $\epsilon_i = \rho\epsilon_{i-1} + \delta_i$ where the δ_i are i.i.d. $N(0, \sigma^2)$ random variables and $\epsilon_0 := 0$. Derive the joint distribution of the first three errors $(\epsilon_1, \epsilon_2, \epsilon_3)$ for this model. (4 marks)

(Total: 20 marks)

1. (a) *[Seen]* (2A marks) The MLE is $\hat{\mu} = \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.
Its distribution is $\hat{\mu} \sim N(\mu, \sigma^2/n)$ for each $n = 1, 2, 3, \ldots$. Hence its asymptotic distribution is

$$\sqrt{n}(\hat{\mu} - \mu) \rightarrow_d N(0, \sigma^2).$$

(b) *[Seen Similar]* (6B marks) By functional invariance of the MLE, we have $\hat{\theta}_Y = \Phi(-\hat{\mu}/\sigma)$. To find its asymptotic distribution, we apply the delta method with $g(\mu) = \Phi(-\mu/\sigma)$ and $g'(\mu) = -\phi(-\mu/\sigma)/\sigma$. Hence,

$$\sqrt{n}(\hat{\theta}_Y - \theta) = \sqrt{n}(g(\hat{\mu}) - g(\mu)) \rightarrow_d g'(\mu)N(0, \sigma^2) = N(0, \phi(-\mu/\sigma)^2).$$

(c) *[Seen Similar]* (2A marks) Since $W_i = 1$ with probability θ and $W_i = 0$ with probability $1 - \theta$, we have $W_i \sim Bernoulli(\theta)$.

(d) *[Seen]* (4A marks) We have previously seen that the MLE is $\hat{\theta}_W = \bar{W}$.
The asymptotic distribution of $\hat{\theta}_W$ follows by either the central limit theorem or results about MLEs in regular models and is

$$\sqrt{n}(\hat{\theta}_W - \theta) \rightarrow_d N(0, \theta(1 - \theta)).$$

(e) *[Unseen]* (6D marks) We know that $Var(Z_b)$ is the Cramèr-Rao lower bound (CRLB) for a sample with $n = 1$. Hence, $n^{-1}Var(Z_b)$ is the CRLB for a random sample of size n.

Next, we note that, for fixed sample size n, the MLE $\hat{\theta}_W$ is an unbiased estimator of θ with variance $Var(\hat{\theta}_W) = n^{-1}Var(Z_d)$.

From the above observation, we have that $n^{-1}Var(Z_b) \leq n^{-1}Var(Z_d)$ by the CRLB. This is equivalent to $Var(Z_b) \leq Var(Z_d)$, which completes the proof.

2. (a) *[Seen Similar]*

(i) (4A marks) The cdf of $S_{(n)}$ is $F_{S_{(n)}}(s;\theta) = F_S(s;\theta)^n = (s^3/\theta^3)^n$. Differentiating this, the density of $S_{(n)}$ is

$$f_{S_{(n)}}(s;\theta) = n\left[\frac{s^3}{\theta^3}\right]^{n-1} 3\theta^{-3}s^2 = 3n\theta^{-3n}s^{3n-1}, 0 < s < \theta.$$

(ii) (4C marks) Now, since $S_{(n)} < \theta$, let U have the structure $cS_{(n)}$, where $c > 1$. Then,

$$\begin{aligned} P(cS_{(n)} > \theta) &= P\left(S_{(n)} > \frac{\theta}{c}\right) \\ &= \int_{\theta/c}^{\theta} 3n\theta^{-3n}s^{3n-1}ds = \theta^{-3n}[s^{3n}]_{\theta/c}^{\theta} \\ &= 1 - c^{-3n} = (1-\alpha) \end{aligned}$$

so that $c = \alpha^{-1/3n}$. So, the exact $100(1-\alpha)\%$ confidence interval for θ is

$$(0, \alpha^{-1/3n}S_{(n)}).$$

(b) *[Seen Similar]*

(i) (4A marks) We have

$$\begin{aligned} E(S) &= \int_0^{\theta} 3\theta^{-3}s^3 ds = \frac{3}{4}\theta; \\ E(S^2) &= \int_0^{\theta} 3\theta^{-3}s^4 ds = \frac{3}{5}\theta^2; \\ Var(S) &= \frac{3}{5}\theta^2 - \frac{9}{16}\theta^2 = \frac{3}{80}\theta^2. \end{aligned}$$

By the central limit theorem, we have

$$\sqrt{n}(\bar{S} - \frac{3}{4}\theta) \to_d N(0, \frac{3}{80}\theta^2).$$

(ii) *[Seen Method]* (4C marks) By the weak law of large numbers, $\frac{4}{3}\bar{S} \to_p \theta$. Using this with Slutsky's lemma, we conclude

$$\frac{\sqrt{n}(\bar{S} - \frac{3}{4}\theta)}{\frac{4}{3}\bar{S}\sqrt{\frac{3}{80}}} = \frac{\sqrt{n}(\frac{4}{3}\bar{S} - \theta)}{\frac{16}{9}\bar{S}\sqrt{\frac{3}{80}}} = \frac{\sqrt{n}(\frac{4}{3}\bar{S} - \theta)}{\frac{4}{9}\sqrt{\frac{3}{5}}\bar{S}} \to_d N(0,1).$$

Define c to be the value so that $P(N(0,1) > c) = \alpha/2$. Then, using the approximately pivotal distribution we have

$$\left(\frac{4}{3}\bar{S} - \frac{c}{\sqrt{n}}\frac{4}{9}\sqrt{\frac{3}{5}}\bar{S}, \frac{4}{3}\bar{S} + \frac{c}{\sqrt{n}}\frac{4}{9}\sqrt{\frac{3}{5}}\bar{S}\right)$$

is an asymptotically valid $100(1-\alpha)\%$ two-sided confidence interval for θ.

(c) *[Seen]* (4A marks) Let $(0, U)$ be the $(1-\alpha)\times 100\%$ confidence interval from part (a). By results in the lecture notes, a test that rejects H_0 when $\theta_0 \notin (0, U)$ will have level α.

3. (a) *[Seen Method]* (3A marks) For this model, the hat matrix is $\mathbf{P} = \mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T$. The fitted values are then $\hat{\boldsymbol{x}} = \mathbf{P}\mathbf{x}$.

(b) *[Seen Method]* (4C marks) First, we note that $\tilde{\boldsymbol{X}} = (\mathbf{1} \quad \hat{\boldsymbol{x}}) = \mathbf{P}\mathbf{X}$. Then,

$$\begin{aligned}
\tilde{\boldsymbol{\beta}} &= (\tilde{\boldsymbol{X}}^T\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^T\mathbf{Y} \\
&= ((\mathbf{P}\mathbf{X})^T\mathbf{P}\mathbf{X})^{-1}(\mathbf{P}\mathbf{X})^T\mathbf{Y} \\
&= (\mathbf{X}^T\mathbf{P}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{P}\mathbf{Y} \\
&= (\mathbf{X}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\
&= (\mathbf{Z}^T\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1})^{-1}\mathbf{X}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\
&= (\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{Y}
\end{aligned}$$

Note: the penultimate step makes use of the identity $(AB)^{-1} = B^{-1}A^{-1}$.

(c) *[Seen Method]* (3B marks) By linearity of expectation we have

$$E(\tilde{\boldsymbol{\beta}}) = E[(\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{Y}] = (\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^TE(\mathbf{Y}) = (\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}.$$

(3B marks) Using properties of covariance and the matrix inverse we have

$$Cov(\tilde{\boldsymbol{\beta}}) = Cov[(\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{Y}] = (\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T[(\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T]^T\sigma^2 = (\mathbf{X}^T\mathbf{P}\mathbf{X})\sigma^2.$$

(d) *[Seen Method]* (3A marks) We found in part (b) that $\tilde{\boldsymbol{\beta}} = (\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{Y}$.
Note that

$$\mathbf{Z}^T\mathbf{Y} = (n\bar{y} \quad \mathbf{z}^T\mathbf{Y})^T, \qquad \mathbf{Z}^T\mathbf{X} = \begin{pmatrix} n & n\bar{x} \\ n\bar{z} & \mathbf{x}^T\mathbf{Z} \end{pmatrix},$$

$$(\mathbf{Z}^T\mathbf{X})^{-1} = \frac{1}{n\mathbf{x}^T\mathbf{z} - n^2\bar{x}\bar{z}}\begin{pmatrix} \mathbf{x}^T\mathbf{z} & -n\bar{x} \\ -n\bar{z} & n \end{pmatrix} = \frac{n^{-1}}{\sum(z_i-\bar{z})(x_i-\bar{x})}\begin{pmatrix} \mathbf{x}^T\mathbf{z} & -n\bar{x} \\ -n\bar{z} & n \end{pmatrix}.$$

Combining these, we find that

$$\begin{aligned}
(\mathbf{Z}^T\mathbf{X})^{-1}\mathbf{Z}^T\mathbf{Y} &= \frac{n^{-1}}{\sum(z_i-\bar{z})(x_i-\bar{x})}\begin{pmatrix} \mathbf{x}^T\mathbf{z} & -n\bar{x} \\ -n\bar{z} & n \end{pmatrix}\begin{pmatrix} n\bar{y} \\ \mathbf{z}^T\mathbf{Y} \end{pmatrix} \\
&= \frac{n^{-1}}{\sum(z_i-\bar{z})(x_i-\bar{x})}\begin{pmatrix} n\bar{y}\mathbf{z}^T\mathbf{x} - n\bar{x}\mathbf{z}^T\mathbf{Y} \\ n\mathbf{z}^T\mathbf{Y} - n^2\bar{z}\bar{y} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\bar{y}\mathbf{z}^T\mathbf{x} - \bar{x}\mathbf{z}^T\mathbf{Y}}{\sum(z_i-\bar{z})(x_i-\bar{x})} \\ \frac{\mathbf{z}^T\mathbf{Y} - n\bar{z}\bar{y}}{\sum(z_i-\bar{z})(x_i-\bar{x})} \end{pmatrix} \\
&= \begin{pmatrix} \bar{y} - \bar{x}\frac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})(x_i-\bar{x})} \\ \frac{\sum(z_i-\bar{z})(y_i-\bar{y})}{\sum(z_i-\bar{z})(x_i-\bar{x})} \end{pmatrix} = \begin{pmatrix} \tilde{\beta}_0 \\ \tilde{\beta}_1 \end{pmatrix}.
\end{aligned}$$

(e) *[Seen Method]* (4D marks) Since $\tilde{\beta}_1 = \boldsymbol{c}^T\tilde{\boldsymbol{\beta}}$ for $\boldsymbol{c}^T = (0, 1)$, we see that $\tilde{\beta}_1$ is a linear unbiased estimator for β_1. Hence, we can apply the Gauss-Markov theorem to reach the desired conclusion. Alternatively, we have $Var(\tilde{\beta}_1) = \sigma^2\frac{\sum(z_i-\bar{z})^2}{[\sum(z_i-\bar{z})(x_i-\bar{x})]^2} =: \sigma^2 S_{zz}/S_{xz}^2$ and $Var(\hat{\beta}_1) = \sigma^2/S_{xx}$. By the Cauchy-Schwarz inequality $S_{xz}^2 \le S_{xx}S_{zz}$ so $\sigma^2/S_{xx} \le \sigma^2 S_{zz}/S_{xz}^2$, which proves the claim.

4. (a) *[Seen Method]* (6A marks) Multiple formulations are possible, providing season was modelled with three indicator variables in the regression model. Let $Y = \log CASES$. Here, we will define indicator variables $SUMMER$, $AUTUMN$, $WINTER$ and $RAIN$ and interaction terms $S.R = SUMMER \times RAIN$, $A.R = AUTUMN \times RAIN$, and $W.R = WINTER \times RAIN$ and use the regression model

$$E[Y] = \beta_0 + \beta_1 \times RAIN + \beta_2 \times SUMMER + \beta_3 \times AUTUMN + \beta_4 \times WINTER + \beta_5 \times S.R + \beta_6 \times A.R + \beta_7 \times W.R$$

Note that the intercept in this model corresponds to the mean cases on a spring day without rain. Various correct reparametrisations are possible.

(b) *[Seen Method]* (2B marks) If there is no difference in Salmonella cases by season, then the regression parameter for every term that involves season in some way must be 0. Hence our null hypothesis needs to be $H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$.
(3D marks) Using the least squares estimator $\hat{\beta}$ we can compute the F statistic of the form

$$Q = \frac{(A\hat{\beta})^T (A(X^T X)^{-1} A^T)^{-1} A\hat{\beta}}{\hat{\sigma}^2}$$

where A contains the bottom 6 rows of an 8 dimensional identity matrix. In this case, we can use the F distribution with 6 and $n - 8 = 357$ degrees of freedom, rejecting H_0 if $Q > F_{1-\alpha,6,357}$. The same statistic arises by considering

$$Q = \frac{(RSS_0 - RSS)/6}{RSS/357}$$

where $RSS = 357\hat{\sigma}^2$ is computed from the full model and RSS_0 is computed from a simple linear regression of Y on $RAIN$.

(c) *[Seen Method]* (2B marks) If there is to be no interaction between rain and season, then the regression parameter for every interaction term must be 0. Hence our null hypothesis needs to be $H_0 : \beta_5 = \beta_6 = \beta_7 = 0$.
(3D marks) This time the quadratic form is

$$Q = \frac{(A\hat{\beta})^T (A(X^T X)^{-1} A^T)^{-1} A\hat{\beta}}{\hat{\sigma}^2}$$

where A contains the bottom 3 rows of an 8 dimensional identity matrix. In this case, we can use the F distribution with 3 and $n - 8 = 357$ degrees of freedom, rejecting H_0 if $Q > F_{1-\alpha,3,357}$. The same statistic arises by considering

$$Q = \frac{(RSS_0 - RSS)/3}{RSS/357}$$

where $RSS = 357\hat{\sigma}^2$ is computed from the full model and RSS_0 is computed from a linear regression of Y on $RAIN$, $SUMMER$, $AUTUMN$, and $WINTER$.

(d) *[Unseen]* (4B marks) Note that $\epsilon_1 = \delta_1$ so that $\epsilon_2 = \rho\delta_1 + \delta_2$ and $\epsilon_3 = \rho^2\delta_1 + \rho\delta_2 + \delta_3$.
We rewrite this in matrix form as

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \rho & 1 & 0 \\ \rho^2 & \rho & 1 \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} =: AZ$$

where $Z \sim N_3(0, \Sigma = \sigma^2 I_{3\times3})$. By linearity properties of the multivariate normal distribution, we have

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix} \sim N_3(0, \sigma^2 AA^T).$$

If your module is taught across multiple year levels, you might have received this form for each level of the module. You are only required to fill this out once for each question.
Please record below, some brief but non-trivial comments for students about how well (or otherwise) the questions were answered. For example, you may wish to comment on common errors and misconceptions, or areas where students have done well. These comments should note any errors in and corrections to the paper. These comments will be made available to students via the MathsCentral Blackboard site and should not contain any information which identifies individual candidates. Any comments which should be kept confidential should be included as confidential comments for

ExamModuleCode	Question	Comments for Students
MATH50011	1	Most of you correctly solved the basic material in part (i) - well done. The second part involved deriving a pivotal quantity for the equation P(U > theta)=1-alpha. Please consult the solutions for more detail.
MATH50011	2	Most of you correctly solved the basic material in part (i) - well done. The second part also required the use of Slutsky's Lemma to approximate the unknown variance term. Again please consult the solutions for more detail.
MATH50011	3	Full marks were given when a correct argument on a level alpha test was presented, even if the critical value (which was to be constructed in part a) was incorrect.
MATH50011	4	Overall the question was answered well. In part a there were very few complete and correct answers, most students struggled with dummy variables for categorical and interaction predictors – it was very common to have models with 4 parameters rather than 8 perhaps confusing R syntax with actual mathematical model. Despite this, in most cases still good answers were written for parts b and c, where students overall showed good understanding of model comparisons. Part d when attempted was answered well and many students with tidy linear algebra managed to get full marks.