Cover Sheet for Submission of Maths Examinations Summer 2020

We would advise preparing your coversheets with your CID, Module Name and Code and Date, before the exams are due to take place.

CID: 01738166

Module Name: Probability and Statistics

Module Code: MATH40005

Date: 14/05/2020

Questions Answered (in the file):

Please tick next to the question or questions you have answered in this file.

Q1	
Q2	
Q3	
Q4	
Q5	✓
Q6	

(Note: this is a coversheet for all students - not all students will have exams with 6 questions. Please tick the boxes which are appropriate for your exam and/or the file you are submitting).

(Optional) Page Numbers for each question;

Page Number	Question Answered

If handwritten, please complete in CAPITAL Letters, in Blue or Black Ink, ensuring the cover sheet is legible.

(a) For any given a

$$\sum_{i=1}^{n}(x_i-a)^2=\sum_{i=1}^{n}\left[(x_i-\bar{x})+(\bar{x}-a)\right]^2=\sum_{i=1}^{n}\left[(x_i-\bar{x})^2+2(x_i-\bar{x})(\bar{x}-a)+(\bar{x}-a)^2\right]$$

$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+2(\bar{x}-a)\sum_{i=1}^{n}(x_i-\bar{x})+\sum_{i=1}^{n}(\bar{x}-a)^2$$

$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+2(\bar{x}-a)\cdot 0+n(\bar{x}-a)^2$$

$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-a)^2$$, where we have used

$$\sum_{i=1}^{n}(x_i-\bar{x})=\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}\bar{x}=n\bar{x}-n\bar{x}=0$$

Since $n(\bar{x}-a)^2\geq 0$, we get $\sum_{i=1}^{n}(x_i-a)^2\geq\sum_{i=1}^{n}(x_i-\bar{x})^2$, with equality only when $\bar{x}=a$.

(b) Denote $RSS(b_0,b_1)=\sum_{i=1}^{n}\left[(y_i-b_1x_i)-b_0\right]^2$ and write $z_i=y_i-b_1x_i$

$\Rightarrow$ $RSS(b_0,b_1)=\sum_{i=1}^{n}(z_i-b_0)^2$.

From (a) we get that $RSS(b_0,b_1)\geq\sum_{i=1}^{n}\left[z_i-\bar{z}\right]^2$

$\Rightarrow$ $\hat{b}_0=\frac{1}{n}\sum_{i=1}^{n}(y_i-b_1x_i)=\underline{\bar{y}-b_1\bar{x}}$. $RSS(\hat{b}_0,b_1)\leq RSS(b_0,b_1)$.

Now find $\hat{b}_1$:

$$RSS(\hat{b}_0,b_1)=\sum_{i=1}^{n}\left[(y_i-b_1x_i)-(\bar{y}-b_1\bar{x})\right]^2=\sum_{i=1}^{n}\left[(y_i-\bar{y})-b_1(x_i-\bar{x})\right]^2$$

$$=\sum_{i=1}^{n}\left[(y_i-\bar{y})^2-2b_1(x_i-\bar{x})(y_i-\bar{y})+b_1^2(x_i-\bar{x})^2\right]=S_{yy}-2b_1S_{xy}+b_1^2S_{xx}$$

$\Rightarrow RSS(\hat{b}_0, b_1) = \cancel{S_{yy}}\ S_{xx}\left(b_1 - \frac{S_{xy}}{S_{xx}}\right)^2 + \left(S_{yy} - \frac{(S_{xy})^2}{S_{xx}}\right)^2$

$\Rightarrow RSS(\hat{b}_0, \hat{b}_1) \leq RSS(\hat{b}_0, b_1)$ where $\underline{\hat{b}_1 = \frac{S_{xy}}{S_{xx}}}$

$\Rightarrow \underline{\hat{b}_1 = \frac{S_{xy}}{S_{xx}}}$, $\hat{b}_0 = \bar{y} - \hat{b}_1\bar{x} = \underline{\bar{y} - \left(\frac{S_{xy}}{S_{xx}}\right)\bar{x}}$

(c)

(i) $H_0 : \mu_x = \mu_y$

(ii) ~~Assumptions : $x_1, x_2, \dots x_n$ are observations of independent random variables~~

Assume the observations $(x_1, \dots x_n)$ are drom independent r.v. following a normal distribution with unknown mean μ_x and unknown variance σ_x^2. Same for $(y_1, \dots y_m)$ with μ_y, σ_y^2. Also assume the unknown variences are equal : $\sigma_y^2 = \sigma_x^2 = \sigma^2$.

(d) $L(\theta | \mathbf{x}) = f(\vec{x}|\theta) = \prod_{i=1}^{n} f(x_i | \theta)$ (since x_i are independent)

$\Rightarrow L(\theta|\mathbf{x}) = \prod_{i=1}^{n}\left(\frac{1}{\theta}\right) = \begin{cases} \frac{1}{\theta^n} & \text{for } 0 \leq x_i \leq \theta \\ 0 & \text{otherwise} \end{cases}$

$f_x(x) = \begin{cases} \frac{1}{b-a} = \frac{1}{\theta} & , x \in [0,\theta] \\ 0 & , \text{otherwise} \end{cases}$

$\frac{d\ L(\theta|\mathbf{x})}{d\theta} = \frac{d\ \theta^{-n}}{d\theta} = -n\ \theta^{-n-1} = \frac{-n}{\theta^{n+1}}$, which is always negative (as $n > 0, \theta > 0$)

Now find maximum.

$\frac{dL(\theta|\mathbf{x})}{d\theta} < 0$, $\forall \theta \Rightarrow L(\theta|\mathbf{x})$ is decreasing,

$L(\theta|X)$ is decreasing $\Rightarrow$ the MLE will be the smallest value of θ such that $\theta \geq x_i$. This value is $\theta = \max(x_1, \dots, x_n)$

$\Rightarrow \underline{\hat{\theta} = \max(X_1, \dots, X_n)}$