4769 Statistics 4

Q1				
(i)	$\text{L} = \frac{e^{-\theta}\theta^{x_1}}{x_1!} \cdots \frac{e^{-\theta}\theta^{x_n}}{x_n!} \left[= \frac{e^{-n\theta}\theta^{\sum x_i}}{x_1! x_2! \cdots x_n!} \right]$	M1 A1	product form fully correct	
	ln L = const $- n\theta + \sum x_i \ln\theta$	M1 A1		
	$\frac{d \ln \text{L}}{d\theta} = -n + \frac{\sum x_i}{\theta} = 0$	M1 A1		
	$\Rightarrow \hat{\theta} = \frac{\sum x_i}{n} (= \bar{x})$	A1	CAO	
	Check this is a maximum	M1		
	e.g. $\frac{d^2 \ln \text{L}}{\text{d}\theta^2} = -\frac{\sum x_i}{\theta^2} < 0$	A1		9
(ii)	$\lambda = \text{P}(X = 0) = e^{-\theta}$	B1		1
(iii)	We have $R \sim \text{B}(n, e^{-\theta})$,	M1		
	so $\text{E}(R) = ne^{-\theta}$	B1		
	$\text{Var}(R) = ne^{-\theta}(1 - e^{-\theta})$	B1		
	$\tilde{\lambda} = \frac{R}{n}$	M1		
	$\therefore \text{E}(\tilde{\lambda}) = e^{-\theta}$ i.e. unbiased	A1 A1		
	$\text{Var}(\tilde{\lambda}) = \frac{e^{-\theta}(1 - e^{-\theta})}{n}$	A1	BEWARE PRINTED ANSWER	7

(iv)	Relative efficiency of $\tilde{\lambda}$ wrt ML est $= \dfrac{\text{Var(ML Est)}}{\text{Var}(\tilde{\lambda})}$ $= \dfrac{\theta e^{-2\theta}}{n} \cdot \dfrac{n}{e^{-\theta}(1-e^{-\theta})} = \dfrac{\theta}{e^{\theta}-1}$ Eg:- Expression is $\dfrac{\theta}{\theta + \dfrac{\theta^2}{2!} + \ldots}$ always < 1 and this is $\approx$ 1 if θ is small $\approx$ 0 if θ is large	M1 M1 A1 M1 E1 E1 E1	any attempt to compare variances if correct BEWARE PRINTED ANSWER Allow statement that $\dfrac{\theta}{e^{\theta}-1} \to 0$ as $\theta \to \infty$	7

Q2				
(i)	$\mathrm{P}(X=x)=q^{x-1}p$	B1	FT into pgf only	
	Pgf $\mathrm{G}(t)=\mathrm{E}(t^X)=\sum_{x=1}^{\infty} pt^x q^{x-1}$	M1		
	$= pt(1+qt+q^2t^2+\ldots)$	A1		
	$= \underline{\underline{pt(1-qt)^{-1}}}$	A1	BEWARE PRINTED ANSWER [consideration of |qt| < 1 not required]	
	$\mu=\mathrm{G}'(1) \quad \sigma^2=\mathrm{G}''(1)+\mu-\mu^2$	M1	for attempt to find G'(t) and/or G"(t)	
	$G'(t)= pt(-1)(1-qt)^{-2}(-q)+p(1-qt)^{-1}$ $= pqt(1-qt)^{-2}+p(1-qt)^{-1}$	A1		
	$\therefore \mathrm{G}'(1)= pq(1-q)^{-2}+p(1-q)^{-1}=\frac{q}{p}+1=\underline{\underline{\frac{1}{p}}}$	A1	BEWARE PRINTED ANSWER	
	$\mathrm{G}''(t)= pqt(-2)(1-qt)^{-3}(-q)+pq(1-qt)^{-2}+$ $p(-1)(1-qt)^{-2}(-q)$	A1		
	$\therefore \mathrm{G}''(1)=2pq^2(1-q)^{-3}+pq(1-q)^{-2}+pq(1-q)^{-2}$ $=\frac{2q^2}{p^2}+\frac{2q}{p}$	A1		
	$\therefore \sigma^2=\frac{2q^2}{p^2}+\frac{2q}{p}+\frac{1}{p}-\frac{1}{p^2}=\frac{2q^2+2pq+p-1}{p^2}$	M1	For inserting their values	
	$=\frac{q}{p^2}(2q+2p-1)=\underline{\underline{\frac{q}{p^2}}}$	A1	BEWARE PRINTED ANSWER	
				11

(ii)	X_1=number of trials to first success X_2= " " " " " next " . . . . X_n= " " " " " nth " $\therefore Y=X_1+X_2+\ldots..X_n$ = total no of trials to the nth success $\therefore$ pgf of $Y = (\text{pgf of } X)^n = \underline{\underline{p^n t^n (1-qt)^{-n}}}$ $\mu_Y = n\mu_X = \underline{\underline{\dfrac{n}{p}}}$ $\sigma_Y^2 = n\sigma_X^2 = \underline{\underline{\dfrac{nq}{p^2}}}$	E1 E1 1 1 1		5
(iii)	N(candidate's μ_Y , candidate's σ_Y^2)	1		1
(iv)	Y = no of tickets to be sold ~ random variable as in (ii) with n = 140 and p = 0.8 ~ Approx N($\frac{140}{0.8} = 175, \frac{140\times 0.2}{(0.8)^2} = 43.75$) $P(Y \geq 160) \approx P(N(175,43.75) > 159\frac{1}{2})$ = P(N(0,1)>-2.343) = 0.9905 For any sensible discussion <u>in context</u> (eg groups of passengers ⇒ not indep.)	E1 1 M1 A1 A1 E1 E1	Do not award if cty corr absent or wrong, but FT if 160 used → -2.268, 0.9884 CAO	7
Q3	X = amount of salt ~ N($\mu[750], \sigma^2[20^2]$) Sample of n=9			
(i)	Type I error: rejecting null hypothesis … … when it is true. Type II error: accepting null hypothesis … … when it is false. OC: P (accepting null hypothesis … … as a function of the parameter under investigation)	B1 B1 B1 B1 B1 B1	Allow B1 for P(rej H_0 when true) Allow B1 for P(acc H_0 when false) [P(type II error | the true value of the parameter) scores B1+B1]	6
(ii)	Reject if $\bar{x} < 735$ *or* $\bar{x} > 765$ $\alpha = P(\bar{X} < 735 \text{ or } \bar{X} > 765 \mid \bar{X} \sim N(750, \frac{20^2}{9}))$ $= P(Z < \frac{(735-750)3}{20} = -2.25$ or $Z > \frac{(765-750)3}{20} = 2.25)$ = 2(1-0.9878) = 2 × 0.0122 = 0.0244 This is the probability of rejecting good output and unnecessarily re-calibrating the machine – seems small [but not very small?]	M1 A1 A1 A1 E1 E1	Might be implicit CAO Accept any sensible comments	6

(iii)	Accept if $735 < \bar{x} < 765$, and now $\mu = 725$. $\beta = P(735 < \overline{X} < 765 \mid \overline{X} \sim N(725, 20^2/9))$ = P(1.5 < Z< 6) = 1 – 0.9332 = 0.0668 This is the probability of accepting output and carrying on when in fact μ has slipped to 725 – small[-ish?]	M1 A1 A1 A1 E1 E1	might be implicit CAO If upper limit 765 not considered, maximum 2 of these 4 marks. If Φ(6) not considered, maximum 3 out of 4. accept sensible comments	6
(iv)	OC = $P(735 < \overline{X} < 765 \mid \overline{X} \sim N(\mu, 20^2/9))$ $= \Phi\left(\frac{(765-\mu)3}{20}\right) - \Phi\left(\frac{(735-\mu)3}{20}\right)$ "Φ − Φ" μ=720:Φ(6.75) – Φ(2.25)=1– 0.9878 =0.0122 730: 5.25 0.75 =1– 0.7734 =0.2266 740: 3.75 −0.75 =1– (1–0.7734)= 0.7734 750: similarly or by write-down from part (ii) [FT] : 0.9756 760, 770, 780 by symmetry [FT]: 0.7734, 0.2266, 0.0122	M1 M1 A1 1 1 1	 both correct if any two correct	6
Q4				
(i)	$x_{ij} = \mu + \alpha_i + e_{ij}$ μ = population … .. grand mean for whole experiment α_i = population … .. mean by which *i* th treatment differs from μ e_{ij} are experimental errors… ~ ind N $(0, \sigma^2)$	1 1 1 1 1 1 3	Allow "uncorrelated" 1 for ind N; 1 for 0; 1 for σ^2.	9
(ii)	Totals are 240, 246, 254, 264, 196 each from sample of size 5 Grand total 936 "Correction factor" CF = $\frac{936^2}{20}$ = 43804.8 Total SS = 44544 - CF = 739.2			

	Between contractors SS = $\frac{240^2}{5}+\ldots+\frac{196^2}{5}-\mathrm{CF}$ = 44209.6 – CF = 404.8	M1 M1	For correct methods for any two,	
	Residual SS (by subtraction) = 739.2 – 404.8 = 334.4	A1	if each calculated SS is correct.	
	(ANOVA table below)	M1 M1 1 A1 1	CAO	

Source of Variation	SS	df	MS	MS ratio
Between Contractors	404.8	3	$134.9\dot{3}$	6.456
Residual	334.4	16	20.9	
Total	739.2	19		

	Refer to $F_{3,16}$	1	NO FT IF WRONG	
	Upper 5% point is 3.24	1	NO FT IF WRONG	
	Significant	1		
	Seems performances of contractors are not all the same	1		12
(iii)	Randomised blocks	B1		
	Description	E1 E1	Take the subject areas as "blocks", ensure each contractor is used at least once in each block	3