4767 Statistics 2

Question 1

(i)	**EITHER:** $S_{xy} = \Sigma xy - \frac{1}{n}\Sigma x\Sigma y = 316345 - \frac{1}{50} \times 2331.3 \times 6724.3$ $= 2817.8$ $S_{xx} = \Sigma x^2 - \frac{1}{n}(\Sigma x)^2 = 111984 - \frac{1}{50} \times 2331.3^2 = 3284.8$ $S_{yy} = \Sigma y^2 - \frac{1}{n}(\Sigma y)^2 = 921361 - \frac{1}{50} \times 6724.3^2 = 17036.8$ $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{2817.8}{\sqrt{3284.8 \times 17036.8}} = 0.377$ **OR:** $\text{cov}(x,y) = \frac{\sum xy}{n} - \bar{x}\bar{y} = 316345/50 - 46.626 \times 134.486$ $= 56.356$ $\text{rmsd}(x) = \sqrt{\frac{S_{xx}}{n}} = \sqrt{(3284.8/50)} = \sqrt{65.696} = 8.105$ $\text{rmsd}(y) = \sqrt{\frac{S_{yy}}{n}} = \sqrt{(17036.8/50)} = \sqrt{340.736} = 18.459$ $r = \frac{\text{cov(x,y)}}{rmsd(x)rmsd(y)} = \frac{56.356}{8.105 \times 18.459} = 0.377$	M1 for method for S_{xy} M1 for method for at least one of S_{xx} or S_{yy} A1 for at least one of S_{xy}, S_{xx} or S_{yy} correct M1 for structure of *r* A1 (AWRT 0.38) M1 for method for cov (*x*,*y*) M1 for method for at least one msd A1 for at least on of cov(*x*,*y*), rmsd(*x*) or rmsd(*y*) correct M1 for structure of *r* A1 (AWRT 0.38)	**5**
(ii)	H_0: $\rho = 0$ H_1: $\rho \neq 0$ (two-tailed test) where ρ is the population correlation coefficient For $n = 50$, 5% critical value = 0.2787 Since 0.377 > 0.2787 we can reject H_0: There is sufficient evidence at the 5% level to suggest that there is correlation between oil price and share cost	B1 for H_0, H_1 in symbols B1 for defining ρ B1FT for critical value M1 for sensible comparison leading to a conclusion A1 for result B1 FT for conclusion in context	**6**
(iii)	Population The scatter diagram has a roughly elliptical shape, hence the assumption is justified.	B1 B1 elliptical shape E1 conclusion	**3**
(iv)	Because the alternative hypothesis should be decided without referring to the sample data and there is no suggestion that the correlation should be positive rather than negative.	E1 E1	**2**
		TOTAL	**16**

Question 2

(i)	Meteors are seen randomly and independently There is a uniform (mean) rate of occurrence of meteor sightings	B1 B1	**2**
(ii)	(*A*) *Either* P(X = 1) = 0.6268 – 0.2725 = 0.3543 *Or* P(X = 1) = $e^{-1.3}\dfrac{1.3^1}{1!}$ = 0.3543 (*B*) Using tables: P($X \ge 4$) = 1 – P($X \le 3$) = 1 – 0.9569 = 0.0431	M1 for appropriate use of tables or calculation A1 M1 for appropriate probability calculation A1	**4**
(iii)	$\lambda = 10 \times 1.3 = 13$ P(X = 10) = $e^{-13}\dfrac{13^{10}}{10!}$ = 0.0859	B1 for mean M1 for calculation A1 CAO	**3**
(iv)	Mean no. per hour = 60 × 1.3 = 78 Normal approx. to the Poisson, X ~ N(78, 78) P($X \ge 100$) = $P\left(Z > \dfrac{99.5-78}{\sqrt{78}}\right)$ = P(Z >2.434) = 1 – Φ(2.434) = 1 – 0.9926 = 0.0074	B1 for Normal approx. B1 for correct parameters (SOI) B1 for continuity corr. M1 for correct Normal probability calculation using correct tail A1 CAO**, (**but FT wrong or omitted CC)	**5**
(v)	*Either* P(At least one) = 1 - $e^{-\lambda}\dfrac{\lambda^0}{0!}$ = 1 - $e^{-\lambda} \ge 0.99$ $e^{-\lambda} \le 0.01$ $-\lambda \le \ln 0.01$, so $\lambda \ge 4.605$ 1.3 $t \ge 4.605$, so $t \ge 3.54$ Answer t = 4 *Or* t = 1, λ = 1.3, P(At least one) = 1 - $e^{-1.3}$ = 0.7275 t = 2, λ = 2.6, P(At least one) = 1 - $e^{-2.6}$ = 0.9257 t = 3, λ = 3.9, P(At least one) = 1 - $e^{-3.9}$ = 0.9798 t = 4, λ = 5.2, P(At least one) = 1 - $e^{-5.2}$ = 0.9944 Answer t = 4	M1 formation of equation/inequality using P($X \ge 1$) = 1 – P(X = 0) with Poisson distribution. A1 for correct equation/inequality M1 for logs A1 for 3.54 A1 for t (correctly justified) M1 at least one trial with any value of t A1 correct probability. M1 trial with either t = 3 or t = 4 A1 correct probability of t = 3 <u>and</u> t = 4 A1 for t	**5**
		TOTAL	**19**

Question 3

(i)	$X \sim N(1720, 90^2)$ $P(X < 1700) = P\left(Z < \frac{1700 - 1720}{90}\right)$ $= P(Z < -0.2222)$ $= \Phi(-0.2222) = 1 - \Phi(0.2222)$ $= 1 - 0.5879$ $= 0.4121$	M1 for standardising A1 M1 use of tables (correct tail) A1CAO **NB ANSWER GIVEN**	**4**
(ii)	P(2 of 4 below 1700) $= \binom{4}{2} \times 0.4121^2 \times 0.5879^2 = 0.3522$	M1 for coefficient M1 for $0.4121^2 \times 0.5879^2$ A1 FT (min 2sf)	**3**
(iii)	Normal approx with $\mu = np = 40 \times 0.4121 = 16.48$ $\sigma^2 = npq = 40 \times 0.4121 \times 0.5879 = 9.691$ $P(X \geq 20) = P\left(Z \geq \frac{19.5 - 16.48}{\sqrt{9.691}}\right)$ $= P(Z \geq 0.9701) = 1 - \Phi(0.9701)$ $= 1 - 0.8340 = 0.1660$	B1 B1 B1 for correct continuity corr. M1 for correct Normal probability calculation using correct tail A1 CAO**, (**but FT wrong or omitted CC)	**5**
(iv)	H_0: $\mu = 1720$; H_1 is of this form since the consumer organisation suspects that the mean is below 1720 μ denotes the mean intensity of 25 Watt low energy bulbs made by this manufacturer.	B1 E1 B1 for definition of μ	**3**
(v)	Test statistic $= \frac{1703 - 1720}{90/\sqrt{20}} = \frac{-17}{20.12}$ $= -0.8447$ Lower 5% level 1 tailed critical value of $z = -1.645$ $-0.8447 > -1.645$ so not significant. There is not sufficient evidence to reject H_0 There is insufficient evidence to conclude that the mean intensity of bulbs made by this manufacturer is less than 1720	M1 must include $\sqrt{20}$ A1FT B1 for –1.645 No FT from here if wrong. Must be –1.645 unless it is clear that absolute values are being used. M1 for sensible comparison leading to a conclusion. FT only candidate's test statistic A1 for conclusion in words in context	**5**
		TOTAL	**20**

Question 4

(i)

H_0: no association between type of car and sex;
H_1: some association between type of car and sex;

B1

EXPECTED	Male	Female
Hatchback	83.16	48.84
Saloon	70.56	41.44
People carrier	51.66	30.34
4WD	17.01	9.99
Sports car	29.61	17.39

M1 A2 for expected values (to 2 dp)

(allow A1 for at least one row or column correct)

CONTRIBUTION	Male	Female
Hatchback	1.98	3.38
Saloon	0.59	1.00
People carrier	3.61	6.15
4WD	0.23	0.40
Sports car	1.96	3.33

M1 for valid attempt at $(O-E)^2/E$
A1 for all correct
NB These M1A1 marks cannot be implied by a correct final value of X^2

$X^2 = 22.62$

M1 for summation
A1 for X^2 CAO

Refer to χ_4^2
Critical value at 5% level = 9.488

B1 for 4 deg of f
B1 CAO for cv

22.62 > 9.488
Result is significant
There is evidence to suggest that there is some association between sex and type of car.

M1 sensible comparison leading to a conclusion
A1

NB if H_0 H_1 reversed, or 'correlation' mentioned, do not award first B1or final A1

12

(ii)

- In hatchbacks, male drivers are more frequent than expected. E1
- In saloons, male drivers are slightly more frequent than expected. E1
- In people carriers, female drivers are much more frequent than expected. E1
- In 4WDs the numbers are roughly as expected E1
- In sports cars, female drivers are more frequent than expected. E1

5

TOTAL **17**

Deleted: ¶