STATISTICS (C) UNIT 2 TEST PAPER 3

1. A tour company wishes to assess attitudes to family holidays in Greece. It decides to interview all the passengers on a plane returning from Crete.
   (i) Comment on this way of obtaining a sample, and suggest a better way. [2]
   The plane has seats for 120 passengers, and on average it is 97% full. The random variables X and Y represent the number of passengers and the number of empty seats respectively.
   (ii) State, with a reason, which of X and Y is more suitable to be approximated by a Poisson distribution. [3]

2. A die is rolled 30 times, and the mean score X is calculated.
   Given that, for a very large number of trials with this die, the mean score is 3·5 and the variance is 2·91666,
   (i) write down the distribution of X. [3]
   (ii) Hence find the probability that, in any particular sample of 30 trials, $3 < X < 4$. [3]

3. A group of 74 babies, all suffering from a certain disease, have a mean birth-weight of 2·96 kg. In the population as a whole, the mean birth-weight of a baby is 3·2 kg, with a standard deviation of 0·7 kg.
   Taking the null hypothesis that babies with the disease have the same mean birth-weight as other babies, and the alternative hypothesis that their birth-weight is below average,
   (i) find the rejection region for H0 at the 1% significance level. [4]
   (ii) State the probability of making a Type I error. [1]
   (iii) Describe what a Type II error means in this context. [2]

4. In a certain school, 32% of Year 9 pupils are left-handed. A random sample of 35 Year 9 pupils is chosen.
   Use an appropriate normal distribution to estimate the probability that the sample contains more than 5 but less than 15 left-handed pupils.
   [6]
   Explain what adjustment is necessary when using this approximation. [1]

5. A sample of radioactive material decays randomly, with an approximate mean of 1·5 counts per minute.

(i) Name a distribution that would be suitable for modelling the number of counts per minute. Give any parameters required for the model. [1]

(ii) Find the probability of at least 4 counts in a randomly chosen minute. [2]

(iii) Find the probability of 3 counts or fewer in a random interval lasting 5 minutes. [2]

More careful measurements, over 50 one-minute intervals, give the following data for x, the number of counts per minute:

$$\Sigma x = 84, \quad \Sigma x^2 = 226.$$

(iv) Calculate unbiased estimates of the mean and variance of the number of decays per minute, and explain why this data supports your answer to part (i). [4]

6. A coin is tossed 20 times, giving 14 heads.

Working at the 1% significance level

(i) test whether the coin is fair, or whether it is biased towards Heads, stating your hypotheses clearly. [6]

(ii) Find the acceptance region for the test. [2]

(iii) If the coin is in fact biased, such that the probability of scoring Heads is 0·6, find the probability of making a Type II error. [4]

7. Patients suffering from 'flu are treated with a drug. The number of days, t, that it then takes for them to recover is modelled by the continuous random variable T with the probability density function

$$f(t) = kt^3(4 - t) \qquad 0 < t < 4,$$
$$f(t) = 0 \qquad \text{otherwise.}$$

(i) Find the value of k. [3]

(ii) Find the mean and standard deviation of T. [7]

(iii) Find the probability that a patient takes more than 3 days to recover. [3]

(iv) Comment on the suitability of the model. [1]

STATISTICS 2 (C) TEST PAPER 3 : ANSWERS AND MARK SCHEME

1.	(i) Tourists to Crete might have very different views from tourists to		
	other parts of Greece. A sample of tourists covering all regions	B1	
	of Greece should be interviewed	B1	
	(ii) X is B(120, 0.97) and Y is B(120, 0.03). Y is well approximated	B1 B1	
	by Poisson, because n is large and p is small	B1	5
2	(i) By Central Limit Theorem, mean of X is 3·5, variance is	B1	
	2·91666/30, and distribution is N(3·5, 0·0972222)	B1 B1	

(ii) $P(3 < X < 4) = P(-0{\cdot}5/\sqrt{0{\cdot}097222} < Z < 0{\cdot}5/\sqrt{0{\cdot}097222})$	M1 A1	
$= P(-1{\cdot}604 < Z < 1{\cdot}604) = 0{\cdot}891$	A1	6
3 (i) One-tail test, at 1% level, has critical value $z = -2{\cdot}326$	B1	
so rejection region is $X < 3{\cdot}2 - 2{\cdot}326 \times 0{\cdot}7/\sqrt{74} = 3{\cdot}01$ kg	M1 A1 A1	
(ii) 1%	B1	
(iii) If $X > 3{\cdot}011$ kg, then the null hypothesis will not be rejected. This is an error if the mean birth-weight of all the sick babies is $< 3{\cdot}2$ kg	B2	7
4. No. of left-handed is $X \sim B(35, 0{\cdot}32)$ $X \sim N(11{\cdot}2, 7{\cdot}616)$	M1 A1	
$P(5 < X < 15) = P(5{\cdot}5 < X < 14{\cdot}5) = P(-2{\cdot}07 < Z < 1{\cdot}20)$	M1 A1 A1	
$= 0{\cdot}8842 - 0{\cdot}0195 = 0{\cdot}865$	A1	
Continuity correction, going from discrete to continuous variable	B1	7
5. (i) Poisson : Po(1·5)	B1	
(ii) $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0{\cdot}9344 = 0{\cdot}0656$	M1 A1	
(iii) Counts in 5 minutes are Po(7·5), so $P(X \leq 3) = 0{\cdot}0591$	B1 M1	
(iv) Mean = 84/50 = 1·68 Var. = $226/49 - (50/49) \times 1{\cdot}68^2 = 1{\cdot}732$	B1 M1 A1	
Mean $\approx$ variance so this supports Poisson model	B1	9
6. (i) $X \sim B(20, p)$ H0 : $p =$, H1 : $p > \frac{1}{2}$	B1 B1	
Assuming H0, $P(X \geq 14) = 1 - 0{\cdot}9423 = 0{\cdot}0577$	M1 A1 A1	
> 1%, so do not reject H0 at 1% level.	A1	
(ii) From tables, if $X \leq 15$, accept H0	M1 A1	
(iii) If $p = 0{\cdot}6$, then $P(X \leq 15) = 0{\cdot}949$, and this is the probability of a Type II error	M1 A1 A1 A1	12
7. (i) $k\int_0^4 4t^3 - t^4\,dt = 1$ so $k[4^4 - 4^5/5] = 1$ $k = {}^5/_{256}$	M1 A1 A1	
(ii) Mean $= k\int_0^4 4t^4 - t^5\,dt = {}^5/_{256}[4^6/5 - 4^6/6] = 2{\cdot}67$	M1 A1	
$\text{Var}(T) = k\int_0^4 4t^5 - t^6\,dt - 2{\cdot}67^2 = {}^5/_{256}[4^7/6 - 4^7/7] - 7{\cdot}111$	M1 A1 A1	
$= 0{\cdot}508$ Standard deviation $= \sqrt{0{\cdot}508} = 0{\cdot}713$	M1 A1	
(iii) $P(T > 3) = k\int_0^4 4t^3 - t^4\,dt = 0{\cdot}367$	M1 A1 A1	
(iv) The cut-off at $t = 4$ is unlikely to be so definite in reality	B1	14