STATISTICS (C) UNIT 2 TEST PAPER 5

1. (i) Explain briefly why it is often useful to take a sample from a population. [2]
   (ii) A school has 1240 pupils. Describe how random numbers could be used to select a random sample of 40 of these pupils. [2]

2. A certain type of lettuce seed has a 12% failure rate for germination. In a new sample of 50 genetically modified seeds, only 3 did not germinate.
   Clearly stating your hypotheses, test, at the 1% significance level, whether the GM seeds are better. [5]

3. When a large number of candidates take a particular exam, it is found that marks are distributed normally, with mean 63·4% and standard deviation 7·8%.
   (i) Find the probability that a candidate scores more than 80% [2]
   (ii) The top 10% of candidates are awarded A grades, and the next 15% are awarded B grades. Find the upper and lower boundary marks for a B grade [6]

4. A car ferry has lanes 27 m long. It is thought that car lengths are normally distributed, with standard deviation 0·8 m and a mean of 4·7 m.
   (i) Using these assumptions, find the probability that a random sample of 6 cars will fit in one lane of the ferry. [4]
   (ii) If, in fact, 7 cars can fit, should you reject, at the 1% significance level, the hypothesis that the mean length is 4·7 m? [4]

5. The waiting time, in minutes, at a dentist is modelled by the continuous random variable T with probability density function

$$f(t) = k(10 - t) \qquad 0 \le t \le 10,$$
$$f(t) = 0 \qquad \text{otherwise.}$$

   (i) Sketch the graph of f(t) and find the value of k. [3]
   (ii) Find the mean value of T. [2]
   (iii) Find the median value of T. [3]
   (iv) Write down the modal value of T. [1]
   It is sometimes suggested that, for most distributions,

$$3 \times (\text{mean} - \text{median}) \approx \text{mean} - \text{mode}.$$

   (v) Show that this result is not satisfied in this case, and suggest a reason why. [2]

6. In a fruit packing plant, apples are packed on to trays of 100, and then checked for blemishes. Th chance of any particular apple having a blemish is 7%. A tray is selected at random.
   (i) Use a suitable normal distribution to estimate the probability that the tray contains less than five blemished apples [5]

   Trays are rejected if they contain more than 11 blemished apples; otherwise they are accepted.
   (ii) Find the probability that a tray is accepted even if the proportion of blemished apples is higher than originally thought and is, in fact, 0·09. [5]
   (iii) State whether this is a Type I or Type II error. [2]

7. A textbook contains, on average, 1·2 misprints per page. Assuming that the misprints are randomly distributed throughout the book,
   (i) specify a suitable model for X, the random variable representing the number of misprints on a given page. [1]
   (ii) Find the probability that a particular page has more than 2 misprints. [2]
   (iii) Find the probability that Chapter 1, with 8 pages, has no misprints at all. [2]

   Chapter 2 is longer, at 20 pages.
   (iv) Use a suitable approximation to find the probability that Chapter 2 has less than ten misprints altogether. Explain what adjustment is necessary when using this approximation. [7]

STATISTICS 2 (C) TEST PAPER 5 : ANSWERS AND MARK SCHEME

1.	(i) Quicker and cheaper than surveying the whole population	B2	
	(ii) Allocate each pupil a number, from 0000 to 1239 inclusive. Then select successive groups of 4 figures from random number table, and use the first 40 which are in the range 0000 to 1239	B2	4
2.	$X \sim B(50, 0.12)$ $X \sim N(6, 5.28)$ $H0 : p = 0{\cdot}12$ $H1 : p < 0{\cdot}12$	B1	
	$P(X \le 3)$ $z = (3 - 6) / \sqrt{5.28} = -1{\cdot}3056$	M1 A1	
	The critical value at 1% level, one-tailed, is $-2{\cdot}326$, therefore the	B1	
	result is not significant – GM seeds are not proven to be better	A1	5
3.	(i) $P(X > 80) = P(Z > (80 - 63{\cdot}4)/7{\cdot}8) = P(Z > 2{\cdot}128) = 0{\cdot}0167$	M1 A1	
	(ii) Top 10% is $Z = 1{\cdot}282$, top 25% is $Z = 0{\cdot}674$	B2	
	Therefore, upper boundary is $63{\cdot}4 + 1{\cdot}282 \times 7{\cdot}8 = 73{\cdot}39$, and	M1 A1	
	lower boundary is $63{\cdot}4 + 0{\cdot}674 \times 7{\cdot}8 = 68{\cdot}67$	A1	
	i.e. $68.7 < X < 73.4$	A1	8

4 (i) The mean X must be < 27/6 = 4·5 m. B1

$X \sim N(4{\cdot}7, 0{\cdot}8^2/6)$,

$P(X < 4{\cdot}5)$ $Z = (4{\cdot}5 - 4{\cdot}7)/(0{\cdot}8/\sqrt{6}) = -0.6124$ M1 A1

$P(X < 4{\cdot}5) = 1 - 0{\cdot}7297 = 0{\cdot}270$ A1

(ii) For 7 cars, $X \sim N(4{\cdot}7, 0{\cdot}8^2/7)$, so B1

$P(X < {}^{27}/_{7})$ $Z = ({}^{27}/_{7} - 4{\cdot}7) / (0{\cdot}8/\sqrt{7}) = -2{\cdot}787$

$P(Z < {}^{27}/_{7}) = 0{\cdot}0026 < 1\%$ M1 A1

so at 1% level, reject hypothesis that mean is 4·7 m A1 8

5. (i) Graph : straight line from (0, 10k) to (10, 0); on x-axis otherwise B1

$\frac{1}{2} \times 10 \times 10k = 1$ $k = {}^{1}/_{50}$ M1 A1

(ii) $E(T) = \int_0^{10} t \times f(t)dt = \frac{1}{50}\int_0^{10} 10t - t^2 dt = {}^{1}/_{50}(5t^2 - \frac{1}{3}t^3) = 3\frac{1}{3}$ M1 A1

(iii) Need ${}^{1}/_{50}\int_0^{m} 10 - t\,dt = 0.5$ $10m - \frac{1}{2}m^2 = 25$ M1 A1

$m^2 - 20m + 50 = 0$ $m = 2{\cdot}93$ (iv) Mode = 0, from graph A1; B1

(iv) 3(mean – median) = 1·213, mean – mode = 3·3333, so not similar; this is because the mode is not centrally located B1 B1 11

6. (i) No. of blemished apples is $X \sim B(100, 0{\cdot}07)$ $X \sim N(7, 6{\cdot}51)$ B1 B1

$P(X < 5) = P(X < 4{\cdot}5)$

$Z = (4.5 - 7)/\sqrt{6{\cdot}51}) = -0{\cdot}980$ $P(X < 4{\cdot}5) = 0{\cdot}164$ M1 A1 A1

(ii) Under H1, $X \sim B(100, 0{\cdot}09)$ $X \sim N(9, 8{\cdot}19)$, so M1 A1

$P(X \le 11) = P(X < 11{\cdot}5)$

$Z = (11.5 - 9)2{\cdot}5 / \sqrt{8{\cdot}19}) = 0{\cdot}8736)$ $P(X < 11{\cdot}5) = 0{\cdot}809$ M1 A1 A1

(iii) Type II error : accepting original hypothesis, even though wrong B2 12

7. (i) Poisson : $X \sim Po(1{\cdot}2)$ B1

(ii) $P(X > 2) = 1 - e^{-1{\cdot}2}(1 + 1{\cdot}2 + 1{\cdot}2^2/2) = 0{\cdot}121$ M1 A1

(iii) $P(X = 0) = e^{-1{\cdot}2} = 0{\cdot}301$ P(X= 0 in Ch. 1) = 0·3018 = 0·0000677 M1 A1

(iv) Total for Ch. 2 $X \sim Po(24)$ $X \sim N(24, 24)$ M1 A1

Then $P(X < 10) = P(X < 9{\cdot}5) = P(Z < -14{\cdot}5/4{\cdot}90) = P(Z < -2{\cdot}96) = 0{\cdot}0015$ M1 A1 M1 A1

Continuity correction needed, from discrete to continuous B1 12