STATISTICS (C) UNIT 2 TEST PAPER 4

1. A fair die is rolled 300 times. Use an appropriate normal distribution to estimate the probability that the score is 3 on less than 40 of the 300 occasions. [5]

2. A discus thrower achieves a range that is normally distributed, with mean 32 m and variance 49 m^2. Anyone throwing further than 45 m is eligible for a National Championship event.
   (i) Find the probability that the thrower achieves this on his first throw. [2]
   Given that he has three throws,
   (ii) find the probability that the mean distance thrown is between 40 m and 44 m. [4]

3. A television company claimed that 55% of all viewers watched a certain event on its channel. However, in a poll of 160 viewers, only 76 had watched that particular channel.
   (i) Carry out a hypothesis test at the 5% significance level to decide whether this is evidence against the company's claim. [5]
   (ii) Write down the probability of making a Type I error, and explain what is meant, in this context, by such an error. [2]

4. A random sample of A-level results is to be taken from the marks obtained by 6th Formers in a school.
   (i) Would it be advisable simply to use the results of all those doing A-level Maths? Explain your answer. [2]
   The individual UCAS scores have a standard deviation of 3·2 points. Last year, the school's average was 19·6 points. This year, the thirty A-level Maths students achieved a mean of 21·3 points.
   (ii) Working at the 2% significance level, decide whether this supports the hypothesis that the school's results overall are better this year. [5]

5. A random variable X has a Poisson distribution with a mean, λ, which is assumed to equal 5.
   (i) Find $P(X = 0)$. [2]
   (ii) In 100 measurements, the value 0 occurs three times. Find the highest significance level at which you should reject the original hypothesis in favour of $\lambda < 5$. [7]

6. In World War II, the number of V2 missiles that landed on each square mile of London was, on average, 3·5. Assuming that the hits were randomly distributed throughout London,
   (i) suggest a suitable model for the number of hits on each square mile, giving a suitable value for any parameters. [1]
   (ii) calculate the probability that a particular square mile received

(a) no hits, [1]

(b) more than 7 hits. [2]

(iii) State, with a reason, whether the model is likely to be accurate. [1]

In contrast, the number of bombs weighing more than 1 ton landing on each square mile was 45.

(iv) Use a suitable approximation to find the probability that a randomly selected square mile received more than 60 such bombs. Explain what adjustment must be made when using this approximation. [6]

7. A continuous random variable X has probability density function

$$f(x) = kx \qquad 1 < x < 4,$$
$$f(x) = 0 \qquad \text{otherwise.}$$

(i) Sketch a graph of $f(x)$, and hence find the value of k. [4]

(ii) Calculate the mean and the variance of X. [5]

(iii) Show that the interquartile range of X is 1·321, correct to 3 decimal places. [6]

STATISTICS 2 (C) TEST PAPER 4 : ANSWERS AND MARK SCHEME

1. $X \sim B(300, 1/6)$ $X \sim N(50, 41{\cdot}666)$, so B1 M1 A1

$P(X < 40) = P(X < 39{\cdot}5) = P(Z < -10{\cdot}5/\sqrt{41{\cdot}6666}) = P(Z < -1{\cdot}627) = 0{\cdot}0519$ M1 A1 5

2. (i) $P(X > 45) = P(Z > (45 - 32)/7) = P(Z > 1{\cdot}857) = 0{\cdot}0317$ M1 A1

(ii) The mean X is distributed $N(32, 49/3)$, so $P(40 < X < 44)$ B1

$= P(8/\sqrt{16{\cdot}333} < Z < 12/\sqrt{16{\cdot}333}) = P(1{\cdot}979 < Z < 2{\cdot}969)$ M1 A1

$= 0{\cdot}9985 - 0{\cdot}9761 = 0{\cdot}0224$ A1 6

3. (i) H0 : $p = 0{\cdot}55$ $X \sim Bin(160, 0.55)$ $X \sim N(88, 39.6)$ B1

Test statistic is $z = (76.5 - 88)/\sqrt{39.6} = -1{\cdot}827$ M1 A1

This is less than the critical value $-1{\cdot}645$, so reject H0 at 5% level M1 A1

(ii) 0·005; assuming that $p < 0{\cdot}55$ when it is not B1 B1 7

4. (i) No - they are not likely to be a representative sample of the year B1 B1

(ii) H0 : $\mu = 19{\cdot}6$, H1 : $\mu > 19{\cdot}6$ B1

Assuming H0, probability of 30 getting mean of 21·3 is

$P(X>29.3)$ $Z = (21{\cdot}3 - 19{\cdot}6)/(3{\cdot}2/\sqrt{30}) = 2{\cdot}910$

$P(X>29.3) = 0{\cdot}0018 < 2\%$ M1 A1

so this is a significant result at 2% level, and we can reject H0 A1

and accept that the school's results have improved A1 7

5. (i) If mean = 5, $X \sim Po(5)$ $P(X = 0) = 0{\cdot}0067$ M1 A1

(ii) $X \sim Po(\lambda)$ H0 : $\lambda = 5$ H1 : $\lambda < 5$ B1

Under H0, no. of '0's in 100 measurements ~ Po(0·67) M1 A1

$P(X \geq 3) = 1 - e^{-0{\cdot}67}(1 + 0{\cdot}67 + 0.67^2/2!) = 0{\cdot}031 = 3{\cdot}1\%$ M1 A1 A1

e.g. reject H0 at the 5% significance level, but not at the 1% level. A1 9

6. (i) Poisson, Po(3·5) B1

(ii) (a) $P(X = 0) = 0{\cdot}0302$ (from tables) B1

(b) $P(X > 7) = 1 - P(X \leq 7) = 1 - 0{\cdot}9733 = 0{\cdot}0267$ M1 A1

(iii) Might not be random – possibly aimed at specific targets B1

(iv) X~Po(45) X~ N (45, 45) M1 A1

$P(X > 60) = P(X > 60{\cdot}5) = P(Z > 15{\cdot}5/6{\cdot}71) = P(Z > 2{\cdot}31)$ M1 A1

$= 1 - 0{\cdot}9896 = 0{\cdot}0104$ A1

Continuity correction, to convert from discrete to continuous B1 11

7. (i) Graph : straight line from $(1, k)$ to $(4, 4k)$; on x-axis elsewhere B2

Area of trapezium = $\frac{1}{2} \times 3 \times (k + 4k) = 1$, so $k = {}^2/_{15}$ M1 A1

(ii) $E(X) = \int_1^4 \frac{2x^2 dx}{15} = 2{\cdot}8$ Var= $\int_1^4 \frac{2x^3 dx}{15}$ -(mean)$^2 = 0{\cdot}66$ M1 A1 M1 A1 A1

(iii) $k \int_1^n x dx = 0{\cdot}75$ ${}^2/_{15}[u^2 - 1] = 1{\cdot}5$ $u = \sqrt{12{\cdot}25} = 3{\cdot}5$ M1 A1

$k \int_1^n x dx = 0{\cdot}25$ ${}^2/_{15}[l^2 - 1] = 0{\cdot}5$ $l = \sqrt{4{\cdot}75} = 2{\cdot}18$ M1 A1

IQR = $u - l = 1{\cdot}321$ M1 A1 15