STATISTICS (C) UNIT 2 TEST PAPER 1

1. The arrival time of commuters at a station is normally distributed, with 8% of them arriving more than 5 minutes before the train's scheduled departure at 7.35 am, and 3% missing it when it leaves on time.
   (i) Find the mean and standard deviation of their arrival times. [4]
   (ii) If the train leaves one minute late one day, find the percentage of commuters who catch it. [2]

2. The number of copies of *The Statistician* that a newsagent sells each week is modelled by a Poisson distribution. On average, he sells 1·5 copies per week.
   (i) Find the probability that he sells no copies in a particular week. [1]
   (ii) If he stocks 5 copies each week, find the probability he will not have enough copies to meet that week's demand. [2]
   (iii) Find the minimum number of copies that he should stock in order to have at least a 95% probability of being able to satisfy the week's demand. [3]

3. A fair die is rolled 60 times.
   Use an suitable approximation to find the probability of scoring less than five sixes. [6]

4. Briefly describe the difference between a population and a sample. [2]
   A village council has to decide whether or not to build a new village hall. In one road of the village, out of 120 residents, 54 think that the new hall would be a waste of money.
   Working at the 5% significance level, test the hypothesis that the villagers are evenly divided for and against the new building. [5]

5. In a certain field, daisies are randomly distributed, at an average density of 0·8 daisies per cm^2.
   One particular patch, of area 1 cm^2, is selected at random.
   Assuming that the number of daisies per cm^2 has a Poisson distribution,
   (i) find the probability that the chosen patch contains
   (a) no daisies, (b) one daisy. [2]
   Ten such patches are chosen. Using your answers to part (i),
   (ii) find the probability that the total number of daisies is less than two. [4]
   (iii) Use a suitable approximation to find the probability that a patch of area 1 m^2 contains more than 8 100 daisies. [5]

6. Each day on the way to work, a commuter encounters a similar traffic jam. The length of time, in 10-minute units, spent waiting in the traffic jam is modelled by the random variable T with the probability density function:

$$f(t) = k(t^3 - 4t^2 + 4t) \qquad 0 < t < 2,$$
$$f(t) = 0 \qquad \text{otherwise.}$$

   (i) Find the value of k. [3]
   (ii) Find the mean waiting time [2]
   (iii) Show that 0·77 is approximately the median value of T. [3]
   (iv) Given that he has already waited for 12 minutes, find the probability that he will have to wait at least another 3 minutes. [4]

7. A drug currently used to relieve a certain disease has a recovery time which is normally distributed with a mean of 7·2 hours and a standard deviation of 1·4 hours. A new drug, when trialled on 20 patients, has a mean recovery time of 6·3 hours, with the same standard deviation.
   (i) Test, at the 0.1% significance level, whether the new drug is better than the old. [4]

(ii) In this situation, explain what is meant by a Type I error and find the probability of making it. [2]

(iii) If the new drug actually has a mean recovery time of 6·1 hours, find the probability of making a Type II error on the basis of the sample of 20 patients. State any assumptions made in your working. [6]

STATISTICS 2 (C) TEST PAPER 1 : ANSWERS AND MARK SCHEME

1. (i) $1.88=(455-\mu)/\sigma$ $-1.406=(450-\mu)/\sigma$ Work in minutes — B1 B1

$\mu = 452.14$ = 7.32am $\sigma = 1.52$ — M1 A1 (both)

(ii) $z = (456 - 452{\cdot}14) / 1.52 = 2{\cdot}537$, so 99·4% of commuters catch it — M1 A1 6

2. (i) $X \sim \text{Po}(1{\cdot}5)$ $P(X = 0) = e^{-1.5} = 0{\cdot}223$ — B1

(ii) $P(X > 5) = 1 - 0{\cdot}9955 = 0{\cdot}0045$ — M1 A1

(iii) $P(X < 3) = 0{\cdot}9344$ and $P(X < 4) = 0{\cdot}9814$, so he needs 4 copies — M1 M1 A1 6

3. No. of sixes $X \sim B(60, \frac{1}{6})$ $X \sim N(10, 8{\cdot}333)$ — B1 B1

$P(X < 5) = P(Z < (4{\cdot}5 - 10) / 2{\cdot}887) = P(Z < -1{\cdot}905) = 0{\cdot}0284$ — M1 A1 M1 A1 6

4. Population : **all** items being considered Sample : a selected subset — B1 B1

Let p = proportion in favour of new hall; H0 : $p = 0{\cdot}5$, H1 : $p < 0{\cdot}5$ — B1

$X \sim \text{Bin}(120,0.5)$ $X \sim N(60,30)$

$P(X \le 54) = P(X \le 54.5)$ $z = (54.5-60) / \sqrt{30} = -1{\cdot}004$ — M1 A1

This is $> -1{\cdot}645$, the critical value at 5% level, so do not reject H0 — M1 A1 7

5. (i) (a) $P(X = 0) = e^{-0{\cdot}8} = 0{\cdot}449$ (b) $P(X = 1) = 0{\cdot}8e^{-0{\cdot}8} = 0{\cdot}359$ — B1 B1

(ii) $P(0) + P(1) = 0{\cdot}449^{10} + 10 \times 0{\cdot}449^{9} \times 0{\cdot}359 = 0{\cdot}002996$ — M1 M1 A1 A1

(iii) In 1 m^2, expect 8000 daisies so use $\text{Po}(8000) \approx N(8000, 8000)$ — B1 B1

$P(X > 8100{\cdot}5) = P(Z > 100{\cdot}5/89{\cdot}44) = P(Z > 1{\cdot}12) = 0{\cdot}131$ — M1 A1 A1 11

6. (i) $k\int_0^2 f(t)\,dt = 1$, $k\left[\frac{t^4}{4} - \frac{4t^3}{3} + 2t^2\right]_0^2 = 1$ $\frac{4}{3}k = 1$ $k = \frac{3}{4}$ — M1 A1 A1

(ii) Mean $= \frac{3}{4}\int t \times f(t)\,dt = \frac{3}{4}\left[\frac{t^5}{5} - \frac{4t^4}{4} + \frac{4t^3}{3}\right]_0^2 = 0{\cdot}8$, i.e. 8 minutes — M1 A1

(iii) We require $\frac{3}{4}\left[\frac{t^4}{4} - \frac{4t^3}{3} + 2t^2\right] = \frac{1}{2}$ — M1

Substitute $m = 0{\cdot}77$, get 0·4987, so $m \approx 0{\cdot}77$ — M1 A1

(iv) $P(T > 1{\cdot}2) = 1 - k\int_0^{1.2} f(t)\,dt = \int_0^{1.2} t^3 - 4t^2 + 4t\,dt = 0{\cdot}179$ — B1

$P(T > 1{\cdot}5) = 1 - k\int_0^{1.5} f(t)\,dt = \int_0^{1.5} t^3 - 4t^2 + 4t\,dt = 0{\cdot}051$ — B1

Therefore, $P(T > 1{\cdot}5 \mid T > 1{\cdot}2) = 0{\cdot}051 / 0{\cdot}179 = 0{\cdot}283$ — M1 A1 12

7. (i) H0 : mean = 7·2 Test statistic

$z = (6{\cdot}3 - 7{\cdot}2) / (1{\cdot}4/\sqrt{20})$ — B1 M1

$= -2{\cdot}875 > -3{\cdot}09$ (crit. value at 0.1% level), so do not reject H0 — A1 A1

(ii) Type I error means rejecting old drug in favour of new, when new is actually no better; probability = significance level i.e. 0.1% — B1 B1

(iii) At 0·1% level, critical value is $7{\cdot}2 - 3{\cdot}090 \times 1{\cdot}4 / \sqrt{20} = 6{\cdot}23$ — M1

So $P(X > 6{\cdot}23)$, given mean = 6·1, is $P(Z > (6{\cdot}23 - 6{\cdot}1)/(1{\cdot}4/\sqrt{20}))$ — M1 A1

$= P(Z > 0{\cdot}424) = 0{\cdot}336$. This is probability of Type II error — A1 A1

Assumed standard deviation is the same — B1 12