S.2 June 2010

1 (a) Population – a collection of all individuals or items.

(b) Statistic – a random variable which is a function of the sample values and involves no unknown values.

(c) Population – all people in the town eligible to vote.

Statistic – sample proportion 35%.

(d) Sample proportions calculated from all possible samples of size 100 will form a probability distribution – called the sampling distribution.

2 (a) X = no. of games Bhim loses out of 9

$X \sim B(9, 0.2)$ $\quad P(X=3) = {}^{9}C_{3}(0.2)^3(0.8)^6$

$= \underline{0.176}$

(b) fewer than half of the games

$\Rightarrow P(X \leq 4) = \underline{0.9804}$

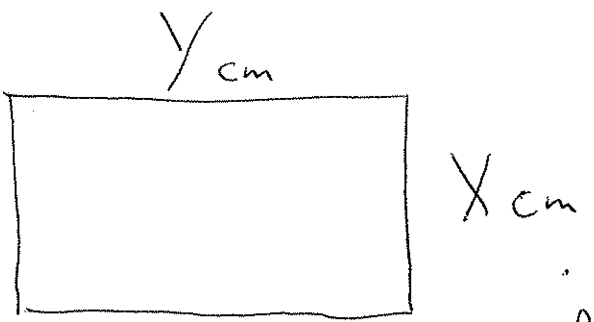
(c) $X \sim B(60, 0.05) \Rightarrow$ Mean $= 60 \times 0.05 = \underline{3}$

Variance $= 60 \times 0.05 \times 0.95 = \underline{2.85}$

(d) n large, p small $\Rightarrow X \approx Po(60 \times 0.05)$ i.e. $Po(3)$

$P(X > 4) = 1 - P(X \leq 4) = 1 - 0.8153 = \underline{0.1847}$

③



$X \sim U(1, 7)$

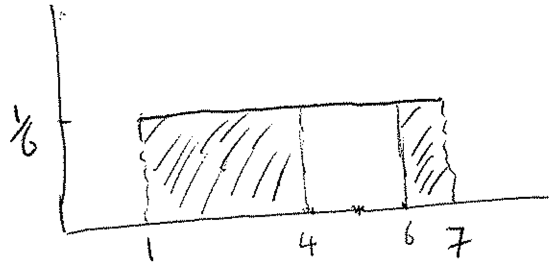
Perimeter 20 cm

X	Y
1	9
2	8
3	7
4	6
5	5
6	4
7	3

Prob.(longer side > 6)

$= P(X < 4) + P(X > 6)$

$=$



$= \frac{3}{6} + \frac{1}{6} = \underline{\underline{2/3}}$

④

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{4}{9}(x^2 + 2x - 3), & 1 \leq x \leq 1.5 \\ 1, & x > 1.5 \end{cases}$$

(a) $F(m) = 0.5 \Rightarrow \frac{4}{9}(m^2 + 2m - 3) = \frac{1}{2}$

$\Rightarrow m^2 + 2m - 3 = \frac{9}{8} \Rightarrow (m+1)^2 - 4 = 9/8$

$\Rightarrow (m+1)^2 = \frac{41}{8} \Rightarrow m = -1 \pm \sqrt{\frac{41}{8}}$

$\Rightarrow$ median is $\underline{1.26}$ (tens of hours)

(b) $f(x) = \begin{cases} \frac{4}{9}(2x+2), & 1 \leq x \leq 1.5 \\ 0, & \text{otherwise.} \end{cases}$

(c) $P(X \geq 1.2) = 1 - F(1.2)$

$= 1 - \frac{28}{75} = \frac{47}{75}$ or $\underline{0.627}$

(d) Prob.(working after 12 hours)

$= \left[P(X > 1.2)\right]^4 = \underline{0.154}$

(5) (a) The connections to the website may occur independently, singly and randomly $\Rightarrow$ Poisson suitable.

(b) X = no. of users who fail to connect in 2 hour period (at first attempt)

$X \sim Po(8)$.

(i) P(all users connect at first attempt)

$= P(X=0) = e^{-8} = \underline{0.000335}$

(ii) $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.0424$

$= \underline{0.9576}$

(c) $H_0: \lambda = 4 \quad H_1: \lambda > 4$

X = no. of users who fail to connect at their first attempt over a 12 hour period.

Under H_0, $X \sim Po(48) \approx N(48, 48)$

$P(X \geq 60) \approx P(X \geq 59.5)$

$= P\left(Z \geq \frac{59.5 - 48}{\sqrt{48}}\right) = P(Z \geq 1.66)$

$= 1 - \Phi(1.66) = 0.0485 < 0.05$

$\Rightarrow$ Test is significant, reject H_0. There is evidence that more users are failing to connect first time.

6 (a) We can assume that the probability that a bolt is faulty, $\frac{1}{4}$, is constant and that the bolts have been selected independently.

(b) $H_0: p = \frac{1}{4} \quad H_1: p \neq \frac{1}{4}$

X = no. of faulty bolts out of 50.

Under H_0, $X \sim B(50, \frac{1}{4})$.

$P(X \leq c) \approx 0.025$

$\Rightarrow c = 6$

$P(X \geq c) = 1 - P(X \leq c-1) \approx 0.025$

$\Rightarrow P(X \leq c-1) \approx 0.975$

$\Rightarrow c-1 = 18, \; c = 19$

So critical region is $X \leq 6$ or $X \geq 19$.

(c) Actual significance level $= P(X \leq 6) + P(X \geq 19)$

$= 0.0194 + (1 - 0.9713) = \underline{0.0481}$

(d) $X = 8$ faulty bolts $\Rightarrow$ not in critical region $\therefore$ do not reject H_0. There is insufficient evidence that the proportion of faulty bolts has changed.

(e) $H_0: p = \frac{1}{4}$ $\quad H_1: p < \frac{1}{4}$

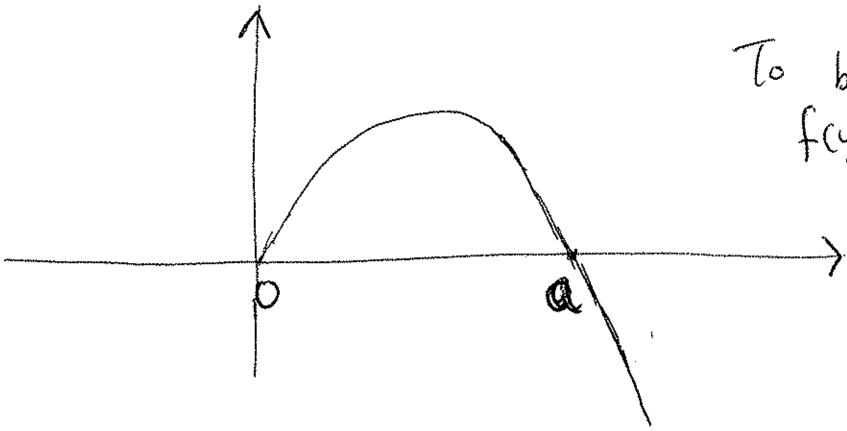
Under H_0, $X \sim B(50, \frac{1}{4})$.

$P(X \leq 5) = 0.007 < 0.01$

$\Rightarrow$ Test is significant, reject H_0. There is evidence that the proportion of faulty bolts has decreased.

(7) $f(y) = \begin{cases} ky(a-y), & 0 \leq y \leq 3 \\ 0, & \text{otherwise} \end{cases}$ $\quad k, a > 0.$

(a)(i)



To be a valid pdf $f(y) \geq 0 \Rightarrow a \geq 3$

(ii) $\int_0^3 ky(a-y)\,dy = 1$

$\Rightarrow \left[k\left(\frac{ay^2}{2} - \frac{y^3}{3}\right)\right]_0^3 = 1$

$\Rightarrow k\left(\frac{9a}{2} - 9\right) = 1 \Rightarrow k(9a-18) = 2$

$\Rightarrow k = \dfrac{2}{9(a-2)}$

(b) $E(Y) = 1.75 \Rightarrow \int_0^3 ky^2(a-y)\,dy = 1.75$

$\Rightarrow \left[k\left(\frac{ay^3}{3} - \frac{y^4}{4}\right)\right]_0^3 = 1.75$

$\Rightarrow k\left(9a - \frac{81}{4}\right) = 1.75$

$\Rightarrow k(36a - 81) = 7$

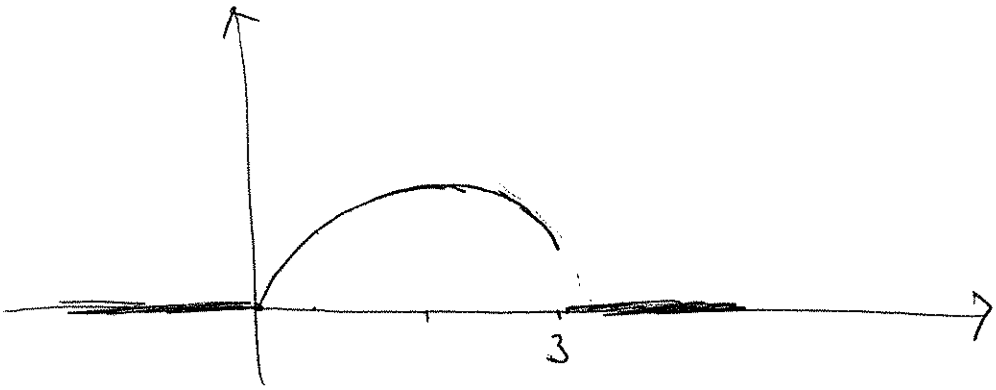
So $\dfrac{2}{9(a-2)}(36a-81) = 7 \Rightarrow 72a - 162 = 63(a-2)$

$9a = 36 \Rightarrow a = 4$

So $k = \dfrac{2}{9(4-2)} = \dfrac{1}{9}$

(c) $f(y) = \begin{cases} \frac{1}{9} y(4-y), & 0 \leq y \leq 3 \\ 0, & \text{otherwise} \end{cases}$

$f(3) = \frac{1}{3}$



(d) Mode of Y is $\underline{\underline{2}}$.

Leave blank

1. Explain what you understand by

(a) a population,

(1)

(b) a statistic.

(1)

A researcher took a sample of 100 voters from a certain town and asked them who they would vote for in an election. The proportion who said they would vote for Dr Smith was 35%.

(c) State the population and the statistic in this case.

(2)

(d) Explain what you understand by the sampling distribution of this statistic.

(1)

Leave blank

2. Bhim and Joe play each other at badminton and for each game, independently of all others, the probability that Bhim loses is 0.2

Find the probability that, in 9 games, Bhim loses

(a) exactly 3 of the games,

(3)

(b) fewer than half of the games.

(2)

Bhim attends coaching sessions for 2 months. After completing the coaching, the probability that he loses each game, independently of all others, is 0.05

Bhim and Joe agree to play a further 60 games.

(c) Calculate the mean and variance for the number of these 60 games that Bhim loses.

(2)

(d) Using a suitable approximation calculate the probability that Bhim loses more than 4 games.

(3)

Leave blank

3. A rectangle has a perimeter of 20 cm. The length, X cm, of one side of this rectangle is uniformly distributed between 1 cm and 7 cm.

Find the probability that the length of the longer side of the rectangle is more than 6 cm long.

(5)

Leave blank

4. The lifetime, X, in tens of hours, of a battery has a cumulative distribution function $\mathrm{F}(x)$ given by

$$\mathrm{F}(x)=\begin{cases} 0 & x<1 \\ \dfrac{4}{9}(x^2+2x-3) & 1\leqslant x\leqslant 1.5 \\ 1 & x>1.5 \end{cases}$$

(a) Find the median of X, giving your answer to 3 significant figures.

(3)

(b) Find, in full, the probability density function of the random variable X.

(3)

(c) Find $\mathrm{P}(X\geqslant 1.2)$

(2)

A camping lantern runs on 4 batteries, all of which must be working. Four new batteries are put into the lantern.

(d) Find the probability that the lantern will still be working after 12 hours.

(2)

Leave blank

5. A company has a large number of regular users logging onto its website. On average 4 users every hour fail to connect to the company's website at their first attempt.

(a) Explain why the Poisson distribution may be a suitable model in this case. **(1)**

Find the probability that, in a randomly chosen **2 hour** period,

(b) (i) all users connect at their first attempt,

(ii) at least 4 users fail to connect at their first attempt. **(5)**

The company suffered from a virus infecting its computer system. During this infection it was found that the number of users failing to connect at their first attempt, over a 12 hour period, was 60.

(c) Using a suitable approximation, test whether or not the mean number of users per hour who failed to connect at their first attempt had increased. Use a 5% level of significance and state your hypotheses clearly. **(9)**

Leave blank

6. A company claims that a quarter of the bolts sent to them are faulty. To test this claim the number of faulty bolts in a random sample of 50 is recorded.

(a) Give two reasons why a binomial distribution may be a suitable model for the number of faulty bolts in the sample.

(2)

(b) Using a 5% significance level, find the critical region for a two-tailed test of the hypothesis that the probability of a bolt being faulty is $\frac{1}{4}$. The probability of rejection in either tail should be as close as possible to 0.025

(3)

(c) Find the actual significance level of this test.

(2)

In the sample of 50 the actual number of faulty bolts was 8.

(d) Comment on the company's claim in the light of this value. Justify your answer.

(2)

The machine making the bolts was reset and another sample of 50 bolts was taken. Only 5 were found to be faulty.

(e) Test at the 1% level of significance whether or not the probability of a faulty bolt has decreased. State your hypotheses clearly.

(6)

Leave blank

7. The random variable Y has probability density function f(y) given by

$$\mathrm{f}(y) = \begin{cases} ky(a-y) & 0 \leqslant y \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

where k and a are positive constants.

(a) (i) Explain why $a \geqslant 3$

(ii) Show that $k = \dfrac{2}{9(a-2)}$

(6)

Given that $\mathrm{E}(Y) = 1.75$

(b) show that $a = 4$ and write down the value of k.

(6)

For these values of a and k,

(c) sketch the probability density function,

(2)

(d) write down the mode of Y.

(1)