

# Probability and Statistics 1 2011–12

## Solutions to Progress Test

1. (i) The sketch should demonstrate that the density is symmetric about 0.
- (ii)  $1 = -c \int_{-1}^0 x(1-x^2) dx + c \int_0^1 x(1-x^2) dx = 2c \left(\frac{1}{2} - \frac{1}{4}\right) = \frac{c}{2}$ . Hence  $c = 2$ .  
By symmetry,  $\mathbb{E}(X) = 0$ .  
Therefore  $\text{Var}(X) = \mathbb{E}(X^2) = 2 \int_0^1 (2x^3 - 2x^5) dx = 4 \left(\frac{1}{4} - \frac{1}{6}\right) = \frac{1}{3}$ .
- (iii)  $\mathbb{P}(0 < X \leq x) = \int_0^x cy(1-y^2) dy = x^2 - \frac{1}{2}x^4$  for  $0 \leq x \leq 1$ .  $\mathbb{P}(-x < X \leq 0)$  is the same, by symmetry.
- (iv)  $\mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = 2y - y^2$  for  $0 \leq y \leq 1$ .  
Therefore  $f_Y(y) = 2(1-y)$  for  $0 < y < 1$ .
2. (i) The probability that there are two or more claims on an individual policy is  $\theta = 1 - e^{-0.2}(1 + 0.2) = 1 - 1.2e^{-0.2} = 0.0175$ .
- (ii) The number of policies,  $X$ , which have to be considered before finding one with 2 or more claims has a geometric distribution with parameter  $\theta$ .  $\mathbb{P}(X \geq 50) = \sum_{50}^{\infty} \theta(1-\theta)^{x-1} = (1-\theta)^{49} = 0.4205$ .
- (iii) a) The distribution is Binomial with parameters 150 and  $\theta$ .  
The expectation is  $150\theta = 2.63$ , the variance  $150\theta(1-\theta) = 2.58$ .
- b) Using binomial probabilities, we get
- $$\mathbb{P}(Y \geq 3) = 1 - [p_Y(0) + p_Y(1) + p_Y(2)] = 1 - [0.0705 + 0.1887 + 0.2507] = 0.4901.$$
- Instead, we could use the Poisson approximation, since  $\theta$  is small and  $k\theta < 5$ .  
This gives us
- $$\mathbb{P}(Y \geq 3) \approx 1 - [p_Y(0) + p_Y(1) + p_Y(2)] = 1 - [0.0722 + 0.1897 + 0.2494] = 0.4887.$$
3. (i) This is  $\frac{1}{2} \times 0.4 + \frac{1}{2} \times 0.3 = \frac{7}{20}$  or 0.35.
- (ii) By independence, the answer is  $0.35^2 = 0.1225$  or  $\frac{49}{400}$ .
- (iii) We already have the probability of a 2–0 scoreline in the series. We need to add the probability of a 1–0 scoreline, i.e. the probability that team A wins one and the other is drawn. This is  $2 \times 0.35 \times 0.3 = 0.21$ . Adding this to the previous answer gives 0.3325 or  $\frac{133}{400}$ .
- (iv) The probability that team A won the toss both times and won both matches is  $(\frac{1}{2} \times 0.4)^2 = 0.04$ .  
Therefore the answer is  $\frac{0.04}{0.1225} = 0.3265 = \frac{16}{49}$ .
4. (i) a) Diagram.
- b) By independence,  $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$ .
- $$\begin{aligned} \mathbb{P}(A^c \cap B^c) &= 1 - \mathbb{P}(A \cup B) = 1 - [\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)] \\ &= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A) \mathbb{P}(B) = [1 - \mathbb{P}(A)][1 - \mathbb{P}(B)] \\ &= \mathbb{P}(A^c) \mathbb{P}(B^c). \end{aligned}$$
- (ii) a) The event that (c) is correct and either the two correct answers are adjacent or (d) is also correct.
- b) No.  $B \cap A^c$  includes the possibility that (b) and (c) are the correct answers, which is not an element of  $C$ .

c)  $A \cap B$  or  $A \cap B \cap C^c$  would do here.

5. (i) Philosophy: 38, 46, 58; Zoology: 49, 56.5, 65.5.  
Philosophy: IQR is 20; 92 is a possible outlier.  
Zoology: IQR is 16.5; no outliers.
- (ii) Diagram.
- (iii) Location: The quartiles for Z are all higher than those for P, as is the sample mean.  
Spread: P has a higher IQR as well as a larger range, so it is more spread out.  
Skewness: P seems positively skewed, based on inter-quartile distances and on comparison of mean with median. Z is more symmetric, based on the same data, but slightly positively skewed.
6. (i) A histogram would be most appropriate, though marks would also be awarded for a correctly drawn cumulative frequency diagram.
- (ii)  $\bar{x} = \frac{\sum xf_x}{n} = \frac{4217}{71} = 59.39$ .  
 $s_x^2 = \frac{\sum x^2 f_x - n\bar{x}^2}{n-1} = \frac{14000}{70} = 200$ .
- (iii) It is intended that linear interpolation be used here. There are 71 observations.  
LQ: in the 50–59 range. We solve  $\frac{x-49.5}{10} = \frac{17.75-16}{20}$ , giving a LQ equal to  $49.5 + 0.88 = 50.4$  to 1dp.  
Median: in the 50–59 range. We solve  $\frac{x-49.5}{10} = \frac{35.5-16}{20}$ , giving a LQ equal to  $49.5 + 9.75 = 59.3$  to 1dp.  
UQ: in the 60–69 range. We solve  $\frac{x-59.5}{10} = \frac{53.25-36}{21}$ , giving a LQ equal to  $59.5 + 8.21 = 67.7$  to 1dp.