Probability and Statistics 1 2013–14

Solutions to Progress Test

Question 1

(i) $1 = \int_{0.5}^{\infty} a(2+x)^{-3}\,dx = \frac{a}{2}(2.5)^{-2} = \frac{2a}{25}$. Hence $a = 12.5$.

(ii) $F(x) = \int_{0.5}^{x} 12.5(2+y)^{-3}\,dy = 6.25((2.5)^{-2} - (2+x)^{-2}) = 1 - \frac{25}{4(2+x)^2}$.
If $F(x) = \frac{1}{2}$ then $(2+x)^2 = \frac{25}{2}$, so that $x = \frac{5}{\sqrt{2}} - 2 = 1.536$.

(iii)

$$\begin{aligned}\mathbb{E}(X) &= 12.5\int_{0.5}^{\infty} x(2+x)^{-3}\,dx \\ &= \left[-6.25x(2+x)^{-2}\right]_{0.5}^{\infty} + \int_{0.5}^{\infty} 6.25(2+x)^{-2}\,dx \\ &= \frac{1}{2} + 2.5 = 3.\end{aligned}$$

The expectation is noticeably larger than the median, so the distribution is right-skewed.

(iv) $\mathbb{P}(X > 10) = 1 - F(10) = \frac{25}{576} = 0.0434$.
Let Y be the number of claims greater than £10,000 in one day, so that $Y \sim \text{Bin}(10, 0.0434)$.
Then $\mathbb{P}(Y > 1) = 1 - p_Y(0) - p_Y(1) = 1 - 0.6416 - 0.2911 = 0.0672$. *[Total: 24 marks]*

Question 2

(i) Let L denote the number of languages surveyed. Then $L \sim \text{Geom}(\frac{2}{3})$.
Therefore $\mathbb{P}(3 \le L \le 5) = \frac{2}{3}\left(\frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4}\right) = \frac{26}{243} = 0.1070$.

(ii) (a) Let T be the number of languages which use 'th', so that $T \sim \text{Bin}(120, 0.03)$. We wnat $\mathbb{P}(T \ge 5)$. Using the Binomial distribution we get $1 - 0.7076 = 0.2924$, or using the Poisson approximation (mean 3.6) we get $1 - 0.7064 = 0.2936$.

(b) For any given group, the probability that at least one language in the group uses 'th' is $1 - 0.97^{10} = 0.2626$, or $1 - e^{-0.3} = 0.2592$ if the Poisson approximation is used.
The required expectation is therefore $12 \times 0.2626 = 3.15$ or 3.11 under the Poisson approximation.

[Total: 16 marks]

Question 3

(i) $0.22 \times 0.75 = 0.165$.

(ii) $\mathbb{P}(\text{Refuse}) = 0.22\times0.25 + 0.3\times0.2 + 0.48\times0.05 = 0.139$. Therefore we want $0.165/(1-0.139) = 0.192$.

(iii) $(0.22 \times 0.25 + 0.3 \times 0.2 + 0.14 \times 0.05)/0.139 = 0.122/0.139 = 0.878$.

[Total: 14 marks]

Question 4

(i) unordered outcomes: bb, bg, br, gg, gr, rr; ordered similarly. $A = \{bb, bg, br\}$, $B = \{bb, bg, gg\}$, $C = \{gg, gr\}$.

(ii) Venn diagram: this should show that A and C are mutually exclusive and should list all the outcomes, including rr.

(iii) (a) $\mathbb{P}(A) = \mathbb{P}(bb) + \mathbb{P}(bg) + \mathbb{P}(br) = \frac{0+2(9-r)+2r}{90} = 0.2$.
$\mathbb{P}(B) = \mathbb{P}(bb) + \mathbb{P}(bg) + \mathbb{P}(gg) = \frac{0+2(9-r)+(9-r)(8-r)}{90} = \frac{(9-r)(10-r)}{90}$.
$\mathbb{P}(C) = \mathbb{P}(gg) + \mathbb{P}(gr) = \frac{(9-r)(8-r)+2r(9-r)}{90} = \frac{(8+r)(9-r)}{90}$.
$\mathbb{P}(A \cup B \cup C) = 1 - \mathbb{P}(rr) = 1 - \frac{r(r-1)}{90}$.

(b) $\mathbb{P}(B \cap C) = \mathbb{P}(gg) = \frac{(9-r)(8-r)}{90}$. For independence we require that $\frac{(9-r)(10-r)}{90} \times \frac{(8+r)(9-r)}{90} = \frac{(9-r)(8-r)}{90}$, so that $(8+r)(9-r)(10-r) = 90(8-r)$. Solution: $r = 4$.

[Total: 26 marks]

Question 5

(i) Quartiles for type A are 378, 407 and 418.5 (IQR 40.5); for type B 331, 361, 374 (IQR 43); for type C 322, 378, 391 (IQR 69).

(ii) For type A 295 is a possible outlier; there are no others. But the intervals should be calculated properly.

(iii) Box-and-whisker. There should be a single horizontal axis with a linear scale, the three data sets should be clearly labelled and any outliers should be indicated appropriately.

(iv) The IQR is much larger for type C than for A or B, although A has the largest range (partly due to the outlier).

Type A has the highest quartiles, so can be assumed to be the largest animals; there do not appear to be systematic differences between the sizes of B and C, to judge by the quartiles.

[Total: 20 marks]