Probability and Statistics 1 2012–13

Solutions to Progress Test

Question 1

(i) (a) The requirements are that $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to+\infty} F(x) = 1$ and $F'(x) > 0$.

(b) As $x \to -\infty$, $e^{\theta x} \to 0$ and $e^{-\theta x} \to \infty$, so $F(x) \to 0$.
As $x \to +\infty$, $e^{\theta x} \to \infty$ and $e^{-\theta x} \to 0$, so $F(x) \to 1$.
The only additional requirement is that $F'(x) > 0$. Now

$$\frac{d}{dx}F(x) = \frac{\theta e^{\theta x}}{e^{\theta x} + e^{-\theta x}} - \frac{\theta e^{\theta x}(e^{\theta x} - e^{-\theta x})}{(e^{\theta x} + e^{-\theta x})^2} = \frac{2\theta}{(e^{\theta x} + e^{-\theta x})^2} > 0.$$

As F is a distribution function, the derivative $\frac{dF}{dx}$ is the density function, f.

(ii) (a) Since F is the distribution function of X, we want to find θ so that

$$0.95 = F(1.96) - F(-1.96) = \frac{e^{1.96\theta} - e^{-1.96\theta}}{e^{1.96\theta} + e^{-1.96\theta}}.$$

Therefore

$$0.05e^{1.96\theta} = 1.95e^{-1.96\theta},$$

or $\theta = \ln(39)/3.92 = 0.9346$.

(b) No. Despite the similarities, the density function is completely different from the Normal density.

[Total: 25 marks]

Question 2

(i) Let A, B and C denote the eruption events in a given year. Then $\mathbb{P}(A) = 0.2$, $\mathbb{P}(B) = 0.1$ and $\mathbb{P}(C) = 0.05$. In addition, $\mathbb{P}(A \cap C) = 0$, $\mathbb{P}(A \cap B) = 0.02$ and $\mathbb{P}(B \cap C) = 0.25\mathbb{P}(B \cup C)$.

Now $\mathbb{P}(B \cup C) = 0.15 - \mathbb{P}(B \cap C)$, so the final equation above enables us to deduce that $1.25\mathbb{P}(B \cap C) = 0.25 \times 0.15$, or $\mathbb{P}(B \cap C) = 0.03$.

This enables the Venn diagram to be drawn.

(ii) (a) This is the event that the only volcano to erupt is B.

(b) This is $B \cap (A \cup C)$, although other formulae are possible.

(iii) $p_N(0) = 0.7$, $p_N(1) = 0.25$ and $p_N(2) = 0.05$.

$\mathbb{E}(N) = 0.25 \times 1 + 0.05 \times 2 = 0.35$.

$\mathbb{E}(N^2) = 0.25 \times 1 + 0.05 \times 4 = 0.45$.

Hence $\mathrm{Var}\,(N) = 0.45 - 0.35^2 = 0.3275$.

(iv) The expected number of eruptions in 10 years is 3.5.

Using the Poisson approximation, the probabiilty of 3 or more eruptions is

$$1 - e^{-3.5}\left(1 + 3.5 + \frac{3.5^2}{2!}\right) = 1 - 10.625e^{-3.5} = 1 - 0.3208 = 0.6792.$$

Since the variance of N is very close to the expectation, it is not unreasonable to suggest that a Poisson distribution might give a very good approximation.

[Total: 30 marks]

Question 3

(i) We want the probability that Monday is sunny and Tuesday is sunny, the probability that Monday is cloudy and Tuesday is sunny and the probability that Monday is rainy and Tuesday is sunny. These are $0.4 \times 0.4 + 0.3 \times 0.2 + 0.3 \times 0.2 = 0.28$.

(ii) The conditional probability is $\frac{0.3 \times 0.2}{0.28} = \frac{3}{14}$.

(iii) The possibilities for Monday, Tuesday, Wednesday and Thursday are $SSCC$, $SRCC$, $RSCC$ and $RRCC$. The probabilities are $0.4 \times 0.4 \times 0.3 \times 0.5 = 0.024$, $0.4 \times 0.3 \times 0.3 \times 0.5 = 0.018$, $0.3 \times 0.2 \times 0.3 \times 0.5 = 0.009$ and $0.3 \times 0.5 \times 0.3 \times 0.5 = 0.0225$, coming to a total of 0.0735.

[Total: 15 marks]

Question 4

(i) (a) By linear interpolation. We have

$$\frac{LQ - 20.5}{30.5 - 20.5} = \frac{25 - 24}{39 - 24} \Longrightarrow LQ = 20.5 + \frac{10}{15} = 21.2.$$

Similarly, the median is $30.5 + \frac{11}{16} \times 20 = 44.25$ and the upper quartile is $100.5 + \frac{2}{14} \times 50 = 107.6$.

(b) The mean is much larger than the median, suggesting positive skew.

The difference between the UQ and the median is 63.35, whereas the distance between median and lower quartile is 23.05, again suggesting positive skew.

[Total: 10 marks]

Question 5

(i) The quartiles for Tim Henman are 11, 16, 17. The value of 27 is a possible outlier.

The quartiles for Andy Murray are 14, 22 and 22. There are no outliers.

Plot.

(ii) Location: All Murray's quartiles are higher.

Spread: Murray's IQR is larger, but Henman has the wider range.

Skewness: both data sets have the median close to the UQ, suggesting left skew.

[Total: 20 marks]