Probability & Statistics 1 2012–13

Progress Test

17 January 2013

Answer all 5 questions.
The total number of marks available is 100.

QUESTION 1 *[25 marks]*
A function F is defined by

$$F(x) = \frac{e^{\theta x}}{e^{\theta x} + e^{-\theta x}}, \qquad x \in \mathbb{R}$$

where $\theta > 0$ is an unknown parameter.

(i) (a) State the properties possessed by the (cumulative) distribution function of a continuous random variable.

(b) Show that F is a distribution function and calculate the associated density function, f. **[15 marks]**

(ii) (a) Calculate a value for θ which results in

$$\mathbb{P}(-1.96 \leq X \leq +1.96) = 0.95,$$

where X is a random variable with distribution function F.

(b) If θ has this value, does X have a Normal distribution? Give a reason for your answer. **[10 marks]**

QUESTION 2 *[30 marks]*
An area contains three volcanoes, A, B and C. Volcano A erupts on average once every 5 years, volcano B on average once every 10 years, and volcano C on average once every 20 years. No volcano erupts more than once in a given year. A and C will not both erupt in the same year. The eruptions of A and B occur independently of one another. In addition, the probability that both B and C erupt is one quarter the probability that either B or C (or both) erupts.

(i) Draw a Venn diagram and mark in all the probabilities. **[11 marks]**

(ii) Let A represent the event that volcano A erupts, and similarly for B and C.

(a) Give an interpretation in words of the event $(A \cup C)^c \cap B$.

(b) Write down in set-theoretic notation the event that two volcanoes erupt in the same year. **[4 marks]**

(iii) Let N denote the number of volcanic eruptions in one year. Write down the probability function of N and calculate its expectation and variance. **[8 marks]**

(iv) Use a Poisson approximation to calculate the probability that there are 3 or more eruptions in 10 years. Comment on whether the Poisson distribution is likely to give a good approximation in this case. **[7 marks]**

QUESTION 3 *[15 marks]*
If one day is rainy, the next day will be rainy with probability 0.5, sunny with probability 0.2 or cloudy with probability 0.3. If a day is sunny, the next day will be cloudy with probability 0.3, sunny with probability 0.4, or rainy otherwise. If a day is cloudy, the next day will be cloudy with probability 0.5, sunny with probability 0.2 or rainy otherwise.
Today is Sunday and it is sunny.

(i) Calculate the probability that Tuesday will be sunny. **[4 marks]**

(ii) If Tuesday is sunny, what is the probability that Monday was rainy? **[4 marks]**

(iii) Calculate the probability that the first two cloudy days will be Wednesday and Thursday. **[7 marks]**

QUESTION 4 *[10 marks]*
A statistician records, for 100 Test Matches, the durations of the opening partnerships in terms of the number of balls faced (which must be an integer). The sample mean is 74.93 and the results are tabulated as follows:

Balls faced	1–10	11–20	21–30	31–50	51–75	76–100	101–150	>150
Frequency	12	12	15	16	10	8	14	13

Source: cricinfo.com

(i) Calculate the median and quartiles of the observations. **[6 marks]**

(ii) Use **two** different methods to assess symmetry or skewness of the data. **[4 marks]**

QUESTION 5 *[20 marks]*
Tim Henman played in six grand-slam tennis semi-finals; up until Auguts 2011 Andy Murray had played in 10. The number of games scored by the players on each occasion are shown below.

Tim Henman	17	17	27	11	15	11				
Andy Murray	23	22	21	14	24	13	15	14	22	22

(i) Draw a box-and-whisker plot to compare the performances of the two players. Ensure that you check for the presence of outliers. **[14 marks]**

(ii) Comment on the data sets in terms of location, spread and symmetry. **[6 marks]**

Formulae

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$
$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Distribution	Notation	Mean	Variance	$p(x)$	Domain
Binomial	$\mathrm{Bin}(k, \theta)$	$k\theta$	$k\theta(1-\theta)$	$\binom{k}{x}\theta^x (1-\theta)^{k-x}$	$x = 0, 1, \ldots, k$
Poisson	$\mathrm{Pois}(\lambda)$	λ	λ	$e^{-\lambda}\dfrac{\lambda^x}{x!}$	$x = 0, 1, \ldots$
Geometric	$\mathrm{Geom}(\theta)$	θ^{-1}	$\theta^{-2}(1-\theta)$	$\theta(1-\theta)^{x-1}$	$x = 1, 2, \ldots$