HERIOT-WATT UNIVERSITY

F71SB2 Statistics II

Wednesday, 14 March 2007, 16.30 – 18.30

Attempt ALL 8 questions.
A total of 90 marks is available.

Approved electronic calculators may be used.

1. (a) A pack of 52 cards consists of 13 Hearts, 13 Spades, 13 Clubs and 13 Diamonds. Suppose that one card is drawn randomly and then replaced in the pack. This process is repeated until a card that is *not* a Diamond is selected. The draws are assumed to be independent from each other. Let X be the random variable giving the number of draws until a card that is *not* a Diamond is drawn.

   i) Identify fully the distribution of X. [2]

   ii) Calculate the probability $P(X \geq 4)$. [3]

   (b) Consider the discrete random variable Y with range $S_y = \{1, 2, 3, 4, ...\}$. The following table gives the cumulative distribution function of Y for its first 3 values:

$$F_Y(y) = \begin{cases} \frac{3}{4}, & \text{for } y = 1 \\ \frac{15}{16}, & \text{for } y = 2 \\ \frac{63}{64}, & \text{for } y = 3 \\ \dots & \dots \end{cases}$$

   i) Calculate the probabilities $P(Y > 2)$ and $P(Y > 3 \mid Y > 1)$. [6]

   ii) State the property of the geometric distribution that your answers in part i) illustrate. [1]

2. Suppose that the lifetime, T (measured in days), of a specific type of electronic component, follows an Exp(0.005) distribution.

   (a) Calculate the probability that the lifetime of a randomly selected component of this type will be between 150 and 200 days. [4]

   (b) Identify fully the distribution of a new random variable Y, which gives the lifetime of the same type of component in weeks. You should explain your answer. [4]

   (c) Hence, or otherwise, calculate the probability that the lifetime of randomly chosen component of this type will exceed 32 weeks. [3]

CONTINUED/

3. Let X be a continuous random variable whose probability density function (*pdf*) is

$$f(x) = \begin{cases} \frac{2(\beta - x)}{\beta^2}, & \text{for } 0 < x < \beta; \\ 0, & \text{elsewhere} \end{cases}$$

where β is a parameter of the distribution.

(a) Show that $E(X) = \frac{\beta}{3}$ by evaluating an appropriate integral. [4]

(b) Suppose that a random sample is taken from the distribution of X and gives a sample mean $\bar{x} = 0.65$. Calculate $\hat{\beta}$, the method-of-moments estimate of β, using this information. [3]

4. Let X be a continuous random variable whose probability density function (*pdf*) is

$$f_X(x) = \begin{cases} \frac{1}{x^2}, & \text{for } 1 < x < \infty; \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify that $f_X(x)$ is a proper *pdf*. [3]

(b) Consider now a transformation of X given by $Y = g(X) = X^{-1}$.

i) Determine the range of the random variable Y. [2]

ii) Derive the *pdf* of Y. [5]

ii) Identify fully the distribution of Y using your answers to i) and ii). [2]

5. A very large group of students obtains test scores that are normally distributed with mean $\mu = 60$ and standard deviation $\sigma = 15$.

(a) Calculate the probability that the score of a randomly selected student from this group will be between 85 and 95. [5]

(b) What proportion of the students obtained a score between 85 and 95? [1]

(c) Find the score which was exceeded by the top 10% of all students. [5]

CONTINUED/

6. Let $X_1, \ldots, X_n$ be independently and identically distributed random variables whose common distribution has mean μ and variance σ^2.

   (a) Show that the mean and variance of $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ are given respectively by μ and $\frac{\sigma^2}{n}$. [4]

   (b) Use the Central Limit Theorem to identify approximately the distribution of $\bar{X}$. [2]

   (c) A fair coin is tossed independently 100 times. Use the Central Limit Theorem to calculate the approximate probability that the number of times that a 'head' is achieved will exceed 40. [6]

7. The height (in cm) of 30 randomly selected plants of a certain kind are given below (in ascending order):

25.1	25.5	27.2	27.9	28.8	28.9	29.1	29.7	29.9	30.1
30.3	30.6	30.6	31.1	31.3	31.7	31.7	31.9	32.2	32.4
32.6	32.9	33.3	33.5	33.8	34.2	34.5	34.9	35.3	36.3

   (a) Calculate the 5-point summary for these data. [5]

   (b) Calculate the interquartile range (IQR) and explain briefly its meaning. [2]

   (c) Construct a stem-and-leaf diagram for the data. Does the plot support the suggestion that the distribution of heights is Normal? Give reasons for your answer. [5]

   (d) Assume now that the variance of the distribution of plant heights from which the above sample is selected, is known and equal to $\sigma^2 = 8.1$. Using this information, calculate a 90% confidence interval for the unknown mean, μ, of the population. (For the above data $\sum x_i = 937.3$) [4]

CONTINUED/

8. For a random sample of 7 vehicles travelling over a particular stretch of motorway, radar monitoring recorded the following speeds, in miles per hour (*mph*):

$$79 \quad 73 \quad 68 \quad 77 \quad 86 \quad 71 \quad 69.$$

For these data $\sum x_i = 523$ and $\sum x_i^2 = 39321$.

(a) Stating clearly any assumptions you need to make, calculate a 95% confidence interval for the unknown mean speed, μ, of all vehicles travelling over this stretch of motorway. [6]

(b) Suppose now that the speed of 86 *mph* in the above sample was wrongly recorded, and that the correct recording is instead 76 *mph*. Without doing any calculations, state whether a 95% confidence interval for μ will now be wider or narrower than the one found in (a). Give reasons for your answer. [3]

[Grand total: 90]

END OF PAPER