**HERIOT-WATT UNIVERSITY**

---

F71SB2 Statistics 2

Friday, 14 March 2008, 09:30 – 11:30

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

Approved electronic calculators may be used

1. A group of 10 students consists of 5 sets of twins. Suppose that 3 students are selected at random without replacement.

   (a) Calculate the probability that the selection does *not* contain any pairs of twins. [4]

   (b) Hence, or otherwise, find the probability that the selection contains a pair of twins. [4]

2. Let $X$ be a random variable representing the number shown when we roll a balanced 6-sided die (i.e. a die with all 6 possible outcomes having equal probability).
   Find the expected value of $g(X) = 2X^2 + 1$. [4]

3. A study is conducted on the number of accidents occurring on a busy street per day. A random sample of the number of accidents for 100 days is collected, and is given in the following table.

| | | | | | | |
|---|---|---|---|---|---|---|
| Number of accidents per day ($x$) : | 0 | 1 | 2 | 3 | 4 | $\geq 5$ |
| Number of days ($f_x$) : | 23 | 38 | 27 | 11 | 1 | 0 |

   It is believed that the number of accidents per day, $X$, follows a Poisson($\lambda$) distribution. (*Recall that the probability mass function of the Poisson*($\lambda$) *distribution is given by* $f_X(x) = P(X = x) = e^{-\lambda}\lambda^x/x!$, *and its mean is given by* $\lambda$.)

   (a) Apply the method of moments to estimate the parameter of the Poisson distribution $\lambda$. [4]

   (b) Calculate the expected frequencies for the number of accidents per day under the assumed distribution and using your estimate of $\lambda$ from above. [6]

   (c) Compare the expected numbers in each category with the observed given in the table and comment briefly. [2]

CONTINUED/

4. A contractor's profit (expressed in £1000s) on a certain job can be regarded as a *continuous* random variable with probability density function (p.d.f.) given by

$$f(x) = \begin{cases} \frac{1}{c}\,(x+1), & \text{for } -1 < x < 5; \\ 0, & \text{otherwise} \end{cases}$$

where $c$ is a constant.

(a) Show that for $f(x)$ above to be a proper p.d.f. we must have $c = 18$. [4]

(b) Calculate $E(X)$, $E(X^2)$ and hence $\text{var}(X)$ by evaluating the appropriate integrals. (*You will need to use the value* $c = 18$ *verified in (a).*) [6]

(c) Calculate the probability that the contractor's profit will be at least £2000. (*Again, you will need to use the value* $c = 18$ *verified in (a).*) [4]

5. (a) Let $X$ be a random variable whose distribution is $N(\mu, \sigma^2)$. State the distribution of the quantity $Z = \frac{X-\mu}{\sigma}$. [2]

(b) Suppose that $X \sim N(80, 16)$. Use statistical tables to calculate the probability $P(77 < X < 81)$. [6]

(c) Consider the random variable $T \sim \text{Exp}(0.5)$. Find the probability $P(1 \leq T \leq 2)$. [5]

6. The concentration (mg/litre) of a certain blood protein in the blood was measured on a random sample of 30 healthy males. The results were (in ascending order):

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 13 | 14 | 15 | 16 | 17 | 18 | 18 | 19 | 21 | 22 | 24 |
| 26 | 27 | 29 | 29 | 36 | 40 | 40 | 40 | 41 | 42 | 43 | 44 |
| 44 | 45 | 46 | 46 | 49 | 50 | | | | | | |

(a) Calculate the median, the quartiles ($Q_1$, $Q_3$) and the interquartile range ($IQR$) for these data. [5]

(b) Construct a stem-and-leaf diagram for the data. Does the plot support the suggestion that the distribution of concentrations is Normal? Give brief reasons for your answer. [6]

CONTINUED/

7. Let $X_1, \ldots, X_n$ be independently identically distributed random variables whose common distribution has mean $\mu$ and variance $\sigma^2$.

   (a) Show that the mean and variance of $Y = \sum_{i=1}^{n} X_i$ are given by $n\mu$ and $n\sigma^2$ respectively, and use the Central Limit Theorem to identify approximately the distribution of $Y$. [6]

   (b) A fair coin is tossed at random 100 times. Use the Central Limit Theorem to calculate the approximate probability that the number of heads will exceed 60. (*You will also need to apply a continuity correction.*) [7]

8. The population distribution of the length (measured in microns) of a certain kind of bacteria is considered to be Normal with *known* variance 0.25 and unknown mean $\mu$. Eight bacteria are selected at random and their lengths measured. The resulting data are:

   6.3 7.3 6.6 6.8 8.0 7.6 7.1 7.0

   For these data $\sum x_i = 56.7$ and $\sum x_i^2 = 403.95$.

   (a) Evaluate the sample mean for these data and calculate a 90% confidence interval for the population mean $\mu$. [6]

   (b) Evaluate the sample variance and sample standard deviation for these data. [3]

   (c) Suppose now that the population variance is *unknown*. Use the sample variance, evaluated in (b), to calculate a 90% confidence interval for $\mu$ in this case. [6]

[Grand total: 90]

END OF PAPER