Assignment 1

Deadline: Wednesday, 15th January, 2014

There are five questions on this assignment. Total marks: [40]. Marks are shown in boxes [ ].

1. Consider the partially completed ANOVA for a completely randomized one-way design presented below. The missing elements of the table are labelled A, B, C, D and E.

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>5</td>
<td>37.9</td>
<td>B</td>
<td>4.27</td>
</tr>
<tr>
<td>Residuals</td>
<td>19</td>
<td>D</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>A</td>
<td>E</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) How many treatments are being compared in this experiment? [1]
(b) Determine the numerical values for A, B, C, D and E. [4]

2. Suppose \( Y \) is a random variable with probability mass function given by the values in the following table:

\[
\begin{array}{c|cccc}
  k & 1 & 2 & 3 & 4 \\
  \hline
  p_Y(k) & \frac{4}{10} & \frac{3}{10} & \frac{2}{10} & \frac{1}{10}
\end{array}
\]

(a) Find \( E[Y] \) and \( \text{var}(Y) \). [3]
(b) Using the law of the unconscious statistician and not otherwise, find \( E[1/Y] \). [2]
3. Suppose that a firm manufactures fuses such that each fuse has a 5% chance of being
defective, independently of all other fuses.

Let the random variable $Y$ represent the number of defective fuses contained within a
randomly selected batch of six fuses, where $Y$ has a probability mass function given by

$$p_Y(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in R_Y = \{0, 1, \ldots, n\}$$

and $n \in \mathbb{Z}^+$ and $0 < p < 1$ are parameters.

(a) State the values of $n$ and $p$. \hspace{1cm} [1]

(b) Find the expected number of defective fuses in the batch, $E[Y]$. \hspace{1cm} [2]

(c) Is there more than a 5% chance that at least two of the fuses in the batch are
defective? Justify your answer. \hspace{1cm} [2]

4. Let the joint probability density function (p.d.f.) of $X$ and $Y$ be given by

$$f_{(X,Y)}(x,y) = \begin{cases} \frac{1}{8}xe^{-(x+y)/2} & x > 0, \ y > 0 \\ 0 & \text{otherwise} \end{cases}.$$ 

(a) Show that the marginal p.d.f. of $X$ is given by

$$f_X(x) = \frac{1}{4}xe^{-\frac{x}{2}} \quad x > 0.$$ \hspace{1cm} [2]

(b) Show that the marginal p.d.f. of $Y$ is given by

$$f_Y(y) = \frac{1}{2}e^{-\frac{y}{2}} \quad y > 0.$$ \hspace{1cm} [2]

(c) Are $X$ and $Y$ independent? \hspace{1cm} [1]
5. A driving school is in the process of drawing up a new performance related pay scale for its driving instructors. To ensure that this is fair, the school would like to assess the variability of the number of hours of tuition required by pupils prior to taking the practical car test.

A random sample of 11 students who recently completed their practical driving tests were drawn from the school’s database and the number of hours of tuition for each student (which also includes time spent accessing online self-study material) was calculated. The data are presented below.

\[
\begin{aligned}
28.89 & & 30.97 & & 29.47 & & 28.84 & & 26.89 \\
33.71 & & 32.36 & & 31.60 & & 32.91 & & 25.68 & & 29.66
\end{aligned}
\]

It is assumed that the data are normally distributed, and that the population from which the data were drawn has an unknown mean \( \mu \) and unknown variance \( \sigma^2 \).

(a) With justification, suggest an appropriate point estimator for \( \sigma^2 \), and calculate its observed value for the data above. [4]

(b) Suppose the sample size is \( n \), and that the sample variance is denoted by \( S^2 \). What is the distribution of \((n - 1)S^2/\sigma^2\)? [2]

(c) Calculate a 90% confidence interval for \( \sigma^2 \). [4]

(d) Show that the expectation of the length of a 100(1 - \( \alpha \))% confidence interval for \( \sigma^2 \), based on a sample of size \( n \), is given by the formula

\[
(n - 1)\sigma^2 \left[ \frac{1}{\chi^2_{1-\alpha/2,n-1}} - \frac{1}{\chi^2_{\alpha/2,n-1}} \right].
\]

(e) Determine the smallest sample size \( n \) in the range 13 to 16 inclusive, i.e. for \( n \in \{13, 14, 15, 16\} \), that would ensure that the expected length of a 95% confidence interval for \( \sigma^2 \) is less than or equal to 2\( \sigma^2 \). [5]

Important Note:
- Please read the current version of the Mathematics & Statistics Coursework Policy. Copies can be obtained from the course website, or in hardcopy from the programme administrator.