STEP Support Programme

STEP 3 Statistics Questions

1 **2007 S3 Q13**

A frog jumps towards a large pond. Each jump takes the frog either 1 m or 2 m nearer to the pond. The probability of a 1 m jump is p and the probability of a 2 m jump is q, where $p + q = 1$, the occurrence of long and short jumps being independent.

(i) Let $p_n(j)$ be the probability that the frog, starting at a point $(n - \frac{1}{2})$ m away from the edge of the pond, lands in the pond for the first time on its jth jump. Show that $p_2(2) = p$.

(ii) Let u_n be the expected number of jumps, starting at a point $(n - \frac{1}{2})$ m away from the edge of the pond, required to land in the pond for the first time. Write down the value of u_1. By finding first the relevant values of $p_n(m)$, calculate u_2 and show that $u_3 = 3 - 2q + q^2$.

(iii) Given that u_n can be expressed in the form $u_n = A(-q)^{n-1} + B + Cn$, where A, B and C are constants (independent of n), show that $C = (1+q)^{-1}$ and find A and B in terms of q. Hence show that, for large n, $u_n \approx \dfrac{n}{p+2q}$ and explain carefully why this result is to be expected.

2 2010 S3 Q12

The infinite series S is given by

$$S = 1 + (1+d)r + (1+2d)r^2 + \cdots + (1+nd)r^n + \cdots \ ,$$

for $|r| < 1$. By considering $S - rS$, or otherwise, prove that

$$S = \frac{1}{1-r} + \frac{rd}{(1-r)^2}\,.$$

Arthur and Boadicea shoot arrows at a target. The probability that an arrow shot by Arthur hits the target is a; the probability that an arrow shot by Boadicea hits the target is b. Each shot is independent of all others. Prove that the expected number of shots it takes Arthur to hit the target is $1/a$.

Arthur and Boadicea now have a contest. They take alternate shots, with Arthur going first. The winner is the one who hits the target first. The probability that Arthur wins the contest is α and the probability that Boadicea wins is β. Show that

$$\alpha = \frac{a}{1-a'b'}\,,$$

where $a' = 1 - a$ and $b' = 1 - b$, and find β.

Show that the expected number of shots in the contest is $\dfrac{\alpha}{a} + \dfrac{\beta}{b}$.

3 **2013 S3 Q12**

A list consists only of letters A and B arranged in a row. In the list, there are a letter As and b letter Bs, where $a \geqslant 2$ and $b \geqslant 2$, and $a + b = n$. Each possible ordering of the letters is equally probable. The random variable X_1 is defined by

$$X_1 = \begin{cases} 1 & \text{if the first letter in the row is } A; \\ 0 & \text{otherwise.} \end{cases}$$

The random variables X_k $(2 \leqslant k \leqslant n)$ are defined by

$$X_k = \begin{cases} 1 & \text{if the } (k-1)\text{th letter is } B \text{ and the } k\text{th is } A; \\ 0 & \text{otherwise.} \end{cases}$$

The random variable S is defined by $S = \sum_{i=1}^{n} X_i$.

(i) Find expressions for $\mathrm{E}(X_i)$, distinguishing between the cases $i = 1$ and $i \neq 1$, and show that $\mathrm{E}(S) = \dfrac{a(b+1)}{n}$.

(ii) Show that:

(a) for $j \geqslant 3$, $\mathrm{E}(X_1 X_j) = \dfrac{a(a-1)b}{n(n-1)(n-2)}$;

(b) $\displaystyle\sum_{i=2}^{n-2}\left(\sum_{j=i+2}^{n} \mathrm{E}(X_i X_j)\right) = \frac{a(a-1)b(b-1)}{2n(n-1)}$;

(c) $\mathrm{Var}\,(S) = \dfrac{a(a-1)b(b+1)}{n^2(n-1)}$.

4 2005 S3 Q14

In this question, you may use the result

$$\int_0^\infty \frac{t^m}{(t+k)^{n+2}}\,\mathrm{d}t = \frac{m!\,(n-m)!}{(n+1)!\,k^{n-m+1}}\,,$$

where m and n are positive integers with $n \geqslant m$, and where $k > 0$.

The random variable V has density function

$$\mathrm{f}(x) = \frac{C\,k^{a+1}\,x^a}{(x+k)^{2a+2}} \qquad (0 \leqslant x < \infty)\,,$$

where a is a positive integer. Show that $C = \dfrac{(2a+1)!}{a!\,a!}$.

Show, by means of a suitable substitution, that

$$\int_0^v \frac{x^a}{(x+k)^{2a+2}}\,\mathrm{d}x = \int_{\frac{k^2}{v}}^\infty \frac{u^a}{(u+k)^{2a+2}}\,\mathrm{d}u$$

and deduce that the median value of V is k. Find the expected value of V.

The random variable V represents the speed of a randomly chosen gas molecule. The time taken for such a particle to travel a fixed distance s is given by the random variable $T = \dfrac{s}{V}$.

Show that

$$\mathrm{P}(T < t) = \int_{\frac{s}{t}}^\infty \frac{C\,k^{a+1}\,x^a}{(x+k)^{2a+2}}\,\mathrm{d}x \qquad (*)$$

and hence find the density function of T. You may find it helpful to make the substitution $u = \dfrac{s}{x}$ in the integral $(*)$.

Hence show that the product of the median time and the median speed is equal to the distance s, but that the product of the expected time and the expected speed is greater than s.