STEP Support Programme

STEP 3 Statistics: Hints

1 **(i)** Since the pond is **large** the frog cannot jump over it!

Consider what $p_2(2)$ actually is, and then think about the different things that can happen with the first jump.

(ii) u_1, u_2 and u_3 are the expected number of jumps the frog takes to land in the pond if it starts $\frac{1}{2}$ m, $1\frac{1}{2}$ m and $2\frac{1}{2}$ m away from the edge. In each case start by finding the different possible number of jumps and use $q = 1 - p$ to write your final answers in terms of q.

(iii) There are three constants to find and three values of u_i so this looks like being a case of solving simultaneous equations. The constants will all be in terms of q. As n gets large certain parts of u_n can be ignored. For the last part, start by working out the expected distance covered by the frog in one jump.

2 It is helpful to write the series for S and the one for rS on separate lines, making sure the r, r^2, r^3 etc terms line up. You should expect to use this first result later on in the question!

Write out the probability that Arthur hits the target on his first, second, third etc shots (remembering that if he hits it on his second shot, he must have missed it on the first shot). Write out an expression for the expected number of shots, and then use the first result shown in the question to express the expectation in the required form.

If Arthur wins, he either hit the target on the first go, or he missed on his first go followed by Boadicea missing and then Arthur hitting, or ... (you will have an infinite sum of probabilities). To find β, start by writing out the infinite sum in the same way that you found α.

For the last part write out an expression for the expected number of shots. You can then separate this into two different infinite series (one of which corresponds to Arthur winning and the other to Boadicea winning). You can use the first result (twice) and then manipulate your result until it is in the required form.

3 **(i)** Start by working out the probabilities that $X_1 = 1$ and $X_k = 1$. You might need to think about the number of possible arrangements. Since the only possibly values of X_k are 0 and 1, we have $\mathrm{E}(X_k) = \mathrm{P}(X_k = 1)$. Note that $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$ (even if X and Y are not independent).

(ii) **(a)** $X_1X_j = 1$ if and only if **both** $X_1 = 1$ **and** $X_j = 1$. Otherwise $X_1X_j = 0$. Consider what we know about the sequence of letters if $X_1X_j = 1$.

(b) Start by finding $\mathrm{E}(X_iX_j)$ (where $j \geqslant i+2$). Then you can find the inner sum, followed by the outer sum. At some point you will need to find $\displaystyle\sum_{k=2}^{n-2} k$, which can be done in various ways.

(c) Remember that $\mathrm{Var}(S) = \mathrm{E}(S^2) - (\mathrm{E}(S))^2$. You have $\mathrm{E}(S)$ from part **(i)**. To find $\mathrm{E}(S^2)$ start by thinking about $S^2 = (X_1 + X_2 + \cdots + X_n) \times (X_1 + X_2 + \cdots + X_n)$ and what happens when you expand the brackets.

4 This question might look a bit off-putting, but it is basically an integration question and you are given the key result needed. There are a few uses of "hence" in this question, remember that "hence" is an instruction which should be obeyed (unlike "hence or otherwise").

Remember that $\displaystyle\int_{\text{all } x} \mathrm{f}(x)\,\mathrm{d}x = 1$, and using the given result will enable you to find C.

Have a look at the limits of the integrals to help you find "a suitable substitution".

For the median we need $\displaystyle\int_0^m \mathrm{f}(x)\,\mathrm{d}x = \int_m^\infty \mathrm{f}(x)\,\mathrm{d}x$.

The expectation is $\displaystyle\int_{\text{all } x} x \times \mathrm{f}(x)\,\mathrm{d}x$ and the result given at the start of the question will be useful again.

The question asks you to find the *density* function of T — this is another name for the probability distribution function, i.e. you want to find $\mathrm{f}(x)$ such that $\mathrm{P}(T < t) = \displaystyle\int_0^t \mathrm{f}(x)\,\mathrm{d}x$. It is possible to write the density function of T in the same form as the density function of V.