STEP II, 2014, Q13 MS

Question 13:

Considering the sequence of events for $X = 4$, the 1st, 2nd and 3rd numbers must all be different and then the 4th must be the same as one of the first three. The probability is therefore

$$P(X = 4) = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\frac{3}{n}$$

The same reasoning applied to $X = r$ gives

$$P(X = r) = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{r-2}{n}\right)\frac{r-1}{n}$$

The result of part (i) is then found by observing that the probabilities of all possible outcomes add up to 1.

Substituting the probabilities into the formula for $E(X)$ gives

$$E(X) = \frac{2}{n} + 3\left(1 - \frac{1}{n}\right)\frac{2}{n} + 4\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\frac{3}{n} + \cdots + (n+1)\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right)$$

For part (iii) observe that any case where $X \geq k$ will have the first $k - 1$ numbers all different from each other. Therefore

$$P(X \geq k) = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-2}{n}\right)$$

The first formula in part (iv) can be shown by considering $kP(Y = k)$ to be equal to the sum of k copies of $P(Y = k)$ and then regrouping the sum for $E(Y)$. Finally this gives two different expressions for $E(Y)$, which must be equal to each other:

$$\frac{2}{n} + 3\left(1 - \frac{1}{n}\right)\frac{2}{n} + 4\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\frac{3}{n} + \cdots + (n+1)\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right)$$
$$= 1 + 1 + \left(1 - \frac{1}{n}\right) + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right)$$

Rearranging and using the result from part (i) then gives the required result.