STEP 2020, P2, Q12 - Solution (3 pages; 1/7/21)

(i) **1st part**

$P(\text{Same score is shown on both rolls}) = \sum_{i=1}^{n}\left(\frac{1}{n} + \varepsilon_i\right)^2$

$= \frac{1}{n^2} \, .n + \frac{2}{n}\{\sum_{i=1}^{n} \varepsilon_i\} + \sum_{i=1}^{n} \varepsilon_i{}^2$

Also, $\sum_{i=1}^{n} P(X = i) = 1$, so that $\frac{1}{n} \, .n + \sum_{i=1}^{n} \varepsilon_i = 1$, and hence

$\sum_{i=1}^{n} \varepsilon_i = 0$

Then $P(\text{Same score is shown on both rolls}) = \frac{1}{n} + \sum_{i=1}^{n} \varepsilon_i{}^2$

2nd part

For an unbiased die, each ε_i is zero, and so the corresponding probability is $\frac{1}{n}$, which is less than $\frac{1}{n} + \sum_{i=1}^{n} \varepsilon_i{}^2$ when not all the ε_i are equal to zero. Thus it is more likely for a biased die to show the same score on two successive rolls.

(ii) Consider n lengths x_i laid out next to each other, with

$P(X = i)$ being the probability that a point chosen at random lies within the length x_i.

Then $P(X = i) = \frac{x_i}{L}$, where $L = \sum_{i=1}^{n} x_i$

Also, from (i), we can write $P(X = i) = \frac{1}{n} + \varepsilon_i$,

so that $\frac{x_i}{L} = \frac{1}{n} + \varepsilon_i$ (1)

From (i), $\sum_{i=1}^{n} \varepsilon_i{}^2 \geq 0$ [for want of anything better to investigate]

and so, from (1), $\sum_{i=1}^{n}(\frac{x_i}{L} - \frac{1}{n})^2 \geq 0$

$\Leftrightarrow \left(\frac{1}{L^2}\sum_{i=1}^{n} {x_i}^2\right) - \left(\frac{2}{Ln}\sum_{i=1}^{n} x_i\right) + \frac{n}{n^2} \geq 0$

Writing $A = \sum_{i=1}^{n} {x_i}^2$

$\Leftrightarrow \frac{A}{L^2} - \frac{2}{n} + \frac{1}{n} \geq 0 \Leftrightarrow \frac{A}{L^2} \geq \frac{1}{n}$ (2)

Now consider $(\sum_{i=1}^{n} x_i)^2 = (\sum_{i=1}^{n} {x_i}^2)$

$+2([x_1x_2 + x_1x_3 + \cdots + x_1x_n] + [x_2x_3 + x_2x_4 + \cdots + x_1x_n]$

$+ \cdots + [x_{n-1}x_n])$

ie $L^2 = A + 2B$, where $B = \sum_{i=2}^{n}\sum_{j=1}^{i-1} x_i x_j$ (3)

Result to be proved: $\boldsymbol{B \leq \frac{n-1}{2n} L^2}$

From (2) & (3), $\frac{L^2 - 2B}{L^2} \geq \frac{1}{n}$

$\Leftrightarrow 1 - \frac{1}{n} \geq \frac{2B}{L^2}$

so that $B \leq \frac{L^2}{2}\left(1 - \frac{1}{n}\right) = \frac{n-1}{2n} L^2$; ie the result to be proved.

[The question seems to be a bit misleading, as the comparison of the two probabilities in (i) isn't actually used.]

(iii) P(Same score is shown on 3 successive rolls)

$= \sum_{i=1}^{n}\left(\frac{1}{n} + \varepsilon_i\right)^3 = \sum_{i=1}^{n}(\frac{1}{n^3} + \frac{3\varepsilon_i}{n^2} + \frac{3{\varepsilon_i}^2}{n} + {\varepsilon_i}^3)$

$= \frac{1}{n^2} + \left(\frac{3}{n^2}\sum_{i=1}^{n} \varepsilon_i\right) + \left(\frac{3}{n}\sum_{i=1}^{n} {\varepsilon_i}^2\right) + \sum_{i=1}^{n} {\varepsilon_i}^3$

$= \frac{1}{n^2} + \left(\frac{3}{n}\sum_{i=1}^{n} {\varepsilon_i}^2\right) + \sum_{i=1}^{n} {\varepsilon_i}^3$, as $\sum_{i=1}^{n} \varepsilon_i = 0$

For an unbiased die, each ε_i is zero, and so the probability is $\frac{1}{n^2}$.

We need to establish whether $\left(\frac{3}{n}\sum_{i=1}^{n}{\varepsilon_i}^2\right) + \sum_{i=1}^{n}{\varepsilon_i}^3 \geq 0$ or ≤ 0 in all cases (the implication in the question is that one of these is true).

Now, $\left(\frac{3}{n}\sum_{i=1}^{n}{\varepsilon_i}^2\right) + \sum_{i=1}^{n}{\varepsilon_i}^3 = \sum_{i=1}^{n}{\varepsilon_i}^2\left(\frac{3}{n} + \varepsilon_i\right)$

and $\frac{3}{n} + \varepsilon_i > \frac{1}{n} + \varepsilon_i = P(X = i) \geq 0$

Thus $\sum_{i=1}^{n}{\varepsilon_i}^2\left(\frac{3}{n} + \varepsilon_i\right) > 0$, as not all of the ε_i are zero.

Thus it is more likely for a biased die to show the same score on three successive rolls.