22B1053 - 22B1205 - 221025

PAGE NO
DATE / /

Q1

(a) X_1: denotes number of additional books such we move from having picked 0 distinct to 1 distinct colors.

hence X_1: denotes picking any book

$\therefore X_1 = 1$.

When books with $i-1$ distinct colors have been collected, then number option for next book to be of different color

$= n+1-i$

$\therefore P(\text{picking different color book after } i-1 \text{ already}) = \frac{n+1-i}{n}$

(b) We calculate PMF of X_i i.e. $P(X_i = r)$. $X_i = r$, means that after we had $i-1$ different color books, it takes r pickups to have i different color i.e. at r^{th} pick we get different color, and at remaining $r-1$ pickups we get any color picked already in those $i-1$ color ($\therefore$ $i-1$ option for each option).

$$\therefore P(X_i = r) = \frac{\text{considered options}}{\text{Total options}} = \frac{(i-1)^{r-1}(n+1-i)}{n^r}$$

$$= \left(\frac{n+1-i}{n}\right)\left(1 - \frac{n+1-i}{n}\right)^{r-1}$$

$\therefore$ parameter $P = \frac{n+1-i}{n}$

(c) Let Z be Geometric R.V i.e. $P(Z=k) = P(1-P)^{k-1}$, $k = 1, 2, \ldots$

$$\therefore E(Z) = \sum_{k=1}^{\infty} (1-P)^{k-1} P k = P + \sum_{k=2}^{\infty} (1-P)^{k-1} P k \quad -(1)$$

$$(1-P)E(Z) = \sum_{k=2}^{\infty} (1-P)^{k-1} P(k-1) \quad -(2)$$

Subtracting (2) from (1),

$$P\,E(Z) = P + \sum_{k=2}^{\infty} (1-P)^{k-1} P$$

$$E(Z) = 1 + \sum_{k=2}^{\infty} (1-P)^{k-1} = 1 + \frac{1-P}{1-(1-P)} = 1 + \frac{1-P}{P} = 1/P$$

$$E(Z) = 1/P.$$

PAGE No.
DATE / /

$$E(Z^2) = \sum_{k=1}^{\infty} p(1-p)^{k-1}k^2 = p + \sum_{k=2}^{\infty} p(1-p)^{k-1}k^2 \quad - (4)$$

$$(1-p)E(Z^2) = \sum_{k=2}^{\infty} p(1-p)^{k-1}(k-1)^2 \quad - (3)$$

Subtracting (3) from (4)

$$pE(Z^2) = p + \sum_{k=2}^{\infty} p(1-p)^{k-1}(2k-1)$$

$$E(Z^2) = 1 + \sum_{k=2}^{\infty} \cancel{p}(1-p)^{k-1}(2k-1) = 1 + 3\cancel{p}(1-p) + \sum_{k=3}^{\infty} \cancel{p}(1-p)^{k-1}(2k-1) \quad - (5)$$

$$(1-p)E(Z^2) = 1-p + \sum_{k=3}^{\infty} \cancel{p}(1-p)^{k-1}(2k-3) \quad - (6)$$

Subtracting (6) from (5)

$$pE(Z^2) = p + 3\cancel{p}(1-p) + \sum_{k=3}^{\infty} 2\cancel{p}(1-p)^{k-1}$$

$$E(Z^2) = \frac{1+3(1-p)}{p} + \sum_{k=3}^{\infty} \frac{2(1-p)^{k-1}}{p} = \frac{1+3(1-p)}{p} + \frac{2(1-p)^2}{p^2}$$

$$= \frac{1+3(1-p)}{p} + \frac{2(1+p^2-2p)}{p^2}$$

$$= \frac{2}{p^2} + 2 + 1 - 3 + \frac{3}{p} - \frac{4}{p}$$

$$= \frac{2}{p^2} - \frac{1}{p}$$

$$\therefore Var(Z) = E(Z^2) - E^2(Z)$$

$$Var(Z) = \frac{1}{p^2} - \frac{1}{p}.$$

(d) As $X^{(n)} = \sum_{i=1}^{n} X_i$

$$\therefore E(X^{(n)}) = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

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

PAGE No.
DATE / /

(e) $\because$ X_i deals with things after X_{i-1} and before X_{i+1} had happened $\therefore$ There X_{i-1}, X_i, X_{i+1} are indepent, and by induction over i, we claim ~~X_i, X_j~~ X_i's are independent R.V.

$$\therefore \quad Var(X^{(n)}) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

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

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

$$Var(X^{(n)}) \leq n^2\left(1 + 1/2^2 + 1/3^2 \cdots\right) - n\left(1 + 1/2 \cdots + 1/n\right)$$

$$= \frac{n^2\pi^2}{6} - n\left(1 + 1/2 \cdots + 1/n\right)$$

$$Var(X^{(n)}) \leq \frac{n^2\pi^2}{6}$$

(f)

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

$$\leq 1 + 1 + 1 + 1 \cdots + 1$$

$$= m$$

$$\leq m+1$$

where $2^m > n \geq 2^{m-1}$ or $m > \log_2 n \geq m-1$

$$\therefore \quad 1 + \frac{1}{2} + \frac{1}{3} \cdots \frac{1}{n} \leq \log_2 n + 2$$

using same m,

$$1 + \frac{1}{2} \cdots \frac{1}{n} \geq 1 + \frac{1}{2} + \frac{1}{3} \cdots \frac{1}{2^{m-1}} > \left(\frac{1}{2} + \frac{1}{2}\right) + \left(\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}\right) \cdots + \left(\frac{1}{2^{m-1}} + \frac{1}{2^{m-1}} \cdots \frac{1}{2^{m-1}}\right)$$

$$= 1 + 1 + \cdots 1$$

$$= m-1 \geq \log_2 n - 1$$

$$\therefore \log_2 n - 1 < 1 + 1/2 \cdots \frac{1}{n} \leq \log_2 n + 2$$

PAGE No.
DATE / /

$\therefore$ for $n > 2$ we have

$$2\log_2 n > \log_2 n + 2$$

or

$$\log_2 n - 1 > \frac{\log_2 n}{2}$$

$\therefore$

$$\frac{\log_2 n}{2} < 1 + \frac{1}{2} + \frac{1}{3} \cdots \frac{1}{n} < 2\log_2 n \quad \text{for } n > 2$$

$\therefore$ as $n > 0$

$$\frac{n\log_2 n}{2} < \left(1 + \frac{1}{2} \cdots \frac{1}{n}\right) n < 2(\log_2 n) \times n \quad \text{for } n > 2$$

$\therefore$ $E(x^{(n)}) \in \Theta(n\log_2 n)$.

$\therefore$ $f(n) = n\log_2 n$