①

Chapter 3

Conditional Probability and Conditional Expectation

Term Final → At least 1 set question

- Calculating prob. and expectation when some partial information is available.

The Discreet Case

$P(E|F) \rightarrow$ Cond. prob. of E, given F.

$$\boxed{P(E|F) = \frac{P(EF)}{P(F)}}$$

$P_{X|Y}(x|y) = P\{X = x \mid Y = y\}$

↓

(Prob. mass func. of X, given Y = y)

So, $p_{X|Y}(x|y) = P\{X=x \mid Y=y\}$

$$= \frac{P\{X=x,\ Y=y\}}{P\{Y=y\}}$$

pmf $$\boxed{p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}}$$

Now, conditional Probability distribution function of X, given that $Y=y$,

$$F_{X|Y}(x|y) = P\{X \le x \mid Y=y\}$$

pdf $$\boxed{F_{X|Y}(x|y) = \sum_{a \le x} P_{X|Y}(a|y)}$$

③

Conditional expectation of X, given Y = y,

$$E[X \mid Y=y] = \sum_x x P\{X=x \mid Y=y\}$$

$$= \sum_x x P_{X|Y}(x|y)$$

$$\therefore \boxed{E[X \mid Y=y] = \sum_x x P\{X=x \mid Y=y\} = \sum_x x P_{X|Y}(x|y)}$$

If X and Y are independent,

$$P_{X|Y}(x|y) = P\{X=x \mid Y=y\}$$

$$= \frac{P\{X=x, Y=y\}}{P\{Y=y\}}$$

$$= \frac{P\{X=x\} P\{Y=y\}}{P\{Y=y\}}$$

$$= P\{X=x\}$$

$P_{X|Y}(x|y) = P_x(x)$; X & Y independent

$$\boxed{P\{X=x \mid Y=y\} = P\{X=x\}}$$

$\hookrightarrow$ X and Y independent

(3-3)

Example: $X \to \text{poisson}(\lambda_1)$, $Y \to \text{poisson}(\lambda_2)$
$X, Y \to$ Independent, $(X+Y=n)$

Let, $p\{X=k \mid X+Y=n\}$

$$= \frac{p\{X=k,\ X+Y=n\}}{p\{X+Y=n\}}$$

$$= \frac{p\{X=k,\ Y=n-k\}}{p\{X+Y=n\}}$$

$$= \frac{p\{X=k\}\ p\{Y=n-k\}}{p\{X+Y=n\}}$$

$$= \frac{e^{-\lambda_1}\lambda_1^{k}}{k!}\ \frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!} \cdot \frac{n!}{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^{n}}$$

// $X+Y \to \text{poisson}(\lambda_1+\lambda_2)$
if X, Y independent

(5)

$$= \binom{n}{k} \frac{(\lambda_1)^k (\lambda_2)^{n-k}}{(\lambda_1+\lambda_2)^n}$$

$$= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}$$

So, $P\{X=k \mid X+Y=n\}$ is the binomial distribution with parameters

$$= \left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$$

$E[X=k \mid X+Y=n]$

$$= n \frac{\lambda_1}{\lambda_1+\lambda_2}$$

// Expectation of
// Binomial (n, p)
// is np

Example: (3.1)

$p(x,y) \rightarrow$ joint prob. mass func.

$p(1,1) = 0.5$, $p(1,2) = 0.1$

$p(2,1) = 0.1$, $p(2,2) = 0.3$

So, $P\{X = \cancel{x} 1 \mid Y = 1\} = ?$

Solve:

$P_Y(1) = \sum_x p(x,1) = p(1,1) + p(2,1)$

$= 0.5 + 0.1$

$= 0.6$

Now,

$P_{X|Y}(\cancel{x} 1 \mid 1) = P\{X = 1 \mid Y = 1\}$

$= \dfrac{P\{X = 1, Y = 1\}}{P\{Y = 1\}}$

$= \dfrac{p(1,1)}{P_Y(1)} = \dfrac{0.5}{0.6}$

$= \dfrac{5}{6}$ (Ans)

7

Ex - 3.2

X_1, X_2 binomial RV with parameters (n_1, p) and (n_2, p)

Now, $p\{X_1 = k \mid X_1 + X_2 = m\} = ?$

Solve:

$$p\{X_1 = k \mid X_1 + X_2 = m\} = \frac{p\{X_1 = k,\ X_1 + X_2 = m\}}{p\{X_1 + X_2 = m\}}$$

$$= \frac{p\{X_1 = k,\ X_2 = m-k\}}{p\{X_1 + X_2 = m\}}$$

$$= \frac{p\{X_1 = k\}\ p\{X_2 = m-k\}}{p\{X_1 + X_2 = m\}}$$

$$= \frac{\binom{n_1}{k} p^k (1-p)^{n_1-k} \binom{n_2}{m-k} p^{m-k} (1-p)^{n_2-m+k}}{\binom{n_1+n_2}{m} p^m (1-p)^{n_1+n_2-m}}$$

// $\underset{(n_1,p)}{X} + \underset{(n_2,p)}{Y} \rightarrow$ Binomial (n_1+n_2, p)

$$= \frac{\binom{n_1}{k}\binom{n_2}{m-k} p^{k+m-k} (1-p)^{n_1-k+n_2-m+k}}{\binom{n_1+n_2}{m} p^m (1-p)^{n_1+n_2-m}}$$

$$= \frac{\binom{n_1}{k}\binom{n_2}{m-k}}{\binom{n_1+n_2}{m}}$$ [hypergeometric distribution]

Ex - 3.4

Rainy day $\rightarrow$ p_i

Non " " $\rightarrow$ q_i

$P(\text{rain tomorrow}) = \alpha$

E [expected # components function | it rains] = ?

Let, $X_i = \begin{cases} 1, & i \text{ functions} \\ 0, & i \text{ doesn't function} \end{cases}$

$$E\left[\sum_{i=1}^{n} X_i \mid Y=1\right] = \sum_{i=1}^{n} E[X_i \mid Y=1]$$

$$= \sum_{i=1}^{n} p_i$$

Continuous RV

$f_{X|Y}(x|y) \rightarrow$ Conditional probability density function, given $Y = y$.

$$\boxed{f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}}$$

$$\boxed{E[X|Y=y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x|y)\, dx}$$

#Ex 3.7

$$f(x,y) = \begin{cases} \frac{1}{2} y e^{-xy} & ; 0<x<\infty,\ 0<y<2 \\ 0 & ; \text{otherwise} \end{cases}$$

$$E\left[e^{x/2} \mid Y=1\right] = ?$$

Solve: $E\left[e^{x/2} \mid Y=1\right] = \int e^{x/2} f_{X|Y}(x|1)\, dx$

$f(x,y) \longrightarrow$ given

$$f_{X|Y}(x|\cancel{y}\ 1) = \frac{f(x,1)}{f_Y(1)}$$

$$\# \; f_Y(y) = \int_0^\infty \frac{1}{2} y e^{-xy} \, dx$$

$$\therefore f_Y(1) = \int_0^\infty \frac{1}{2} e^{-x} \, dx$$

$$= \frac{1}{2} \left[\frac{e^{-x}}{-1} \right]_0^\infty$$

$$\therefore f_Y(1) = \frac{1}{2}$$

Now, $f(x,1) = \frac{1}{2} e^{-x}$

So, $f_{X|Y}(x \mid 1) = \dfrac{\frac{1}{2} e^{-x}}{\frac{1}{2}} = e^{-x}$

$$\therefore E\left[e^{x/2} \mid Y = 1\right] = \int_0^\infty e^{x/2} e^{-x} \, dx$$

$$= \left[\frac{e^{-x/2}}{-\frac{1}{2}} \right]_0^\infty$$

$$= 2 \quad \text{(Ans)}$$

(11)

Ex 3.5

$$f(x,y) = \begin{cases} 6xy(2-x-y) & ; \quad 0<x<1,\ 0<y<1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$\boxed{E[X \mid Y=y]} = ?$$

$$\boxed{f_{X|Y}(x|y)} = ?$$

Solve

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$= \frac{6xy(2-x-y)}{\int_0^1 6xy(2-x-y)\,dx}$$

$$= \frac{6xy(2-x-y)}{y\int_0^1 (12x-6x^2-6xy)\,dx}$$

$$= \frac{6xy(2-x-y)}{y\left[6x^2-2x^3-3x^2y\right]_0^1}$$

$$= \frac{6xy(2-x-y)}{y(4-3y)}$$

$$= \frac{6x(2-x-y)}{4-3y}$$

$$E[X \mid Y=y] = \int_0^1 x \, f_{X|Y}(x|y)\, dx$$

$$= \int_0^1 x \, \frac{6x(2-x-y)}{4-3y}\, dx$$

$$= \frac{1}{4-3y} \int_0^1 12x^2 - 6x^3 - 6x^2 y \, dx$$

$$= \frac{1}{4-3y} \left[4x^3 - \frac{3}{2}x^4 - 2x^3 y \right]_0^1$$

$$= \frac{1}{4-3y} \left(\frac{5}{2} - 2y \right)$$

$$= \frac{5-4y}{8-6y} \quad \text{(Ans)}$$

Computing Expectation by Conditioning

$E[X|Y] \rightarrow$ function of the RV, Y whose value at $Y=y$ is $E[X|Y=y]$

$E[X|Y]$ is itself a RV.

$$\boxed{E[X] = E\left[E[X|Y]\right]}$$

#Discreet

$$\boxed{E[X] = \sum_{y} E[X|Y=y]\, p\{Y=y\}}$$

Continuous:

$$\boxed{E[X] = \int_{-\infty}^{\infty} E[X|Y=y]\, f_Y(y)\, dy}$$

Prove that,

$$E[X] = \sum_y E[X \mid Y=y]\, P\{Y=y\}$$

$$R.H.S = \sum_y E[X \mid Y=y]\, P\{Y=y\}$$

$$= \sum_y \sum_x x\, P\{X=x \mid Y=y\}\, P\{Y=y\}$$

$$= \sum_y \sum_x \frac{x P\{X=x, Y=y\}}{P\{Y=y\}}\, P\{Y=y\}$$

$$= \sum_x x \sum_y P\{X=x, Y=y\}$$

$$= \sum_x x\, P_X\{X=x\}$$

$$= E[X] \quad \text{[proved]}$$

(18)

Like (Ex - 3.9)

$X \rightarrow$ # access to resources

$$Y = \begin{cases} 1, & \text{if application is CPU intensive} \\ 2, & \text{" " Memory "} \end{cases}$$

$E[X|Y=1] = 5$

$E[X|Y=2] = 3$

$P\{Y=1\} = P\{Y=2\} = \frac{1}{2}$

(equally likely)

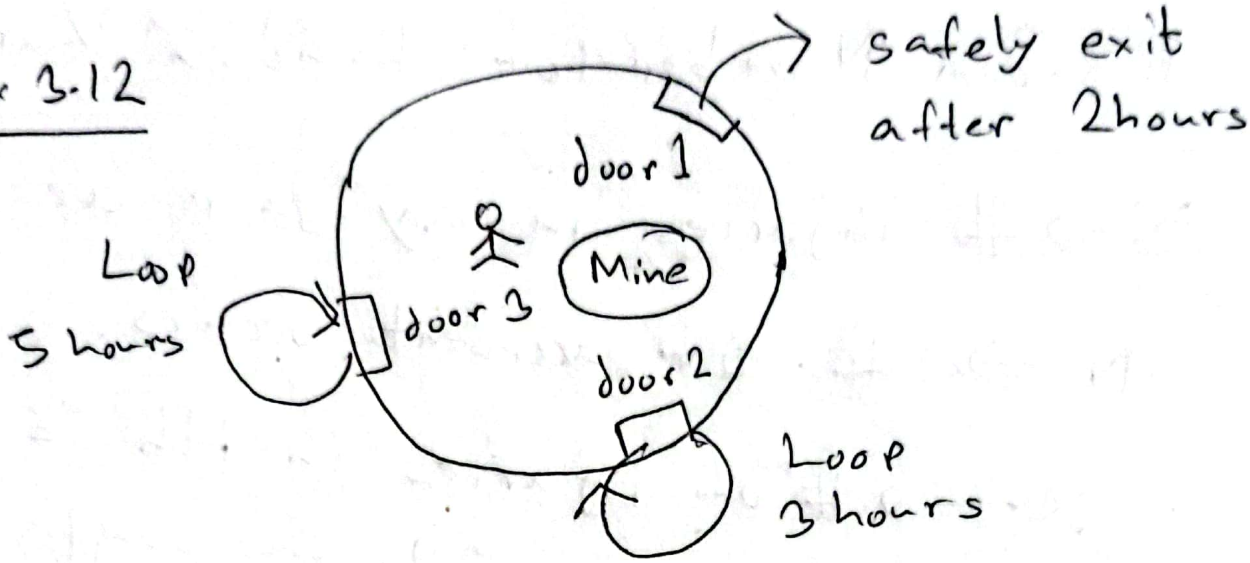
$$E[X] = \sum_{y} E[X|Y=y]\, P\{Y=y\}$$

$$= E[X|Y=1]\, P\{Y=1\} + E[X|Y=2]\, P\{Y=2\}$$

$$= 5 \times \frac{1}{2} + 3 \times \frac{1}{2}$$

$$= 4$$

#Ex 3.12



X : Time until the person reaches safely.

Y : The door he chooses (equally likely) [Memoryless] [He forgets each time what door he chosen last]

$$E[X] = E[X \mid Y=1]\, P\{Y=1\} + E[X \mid Y=2]\, P\{Y=2\} + E[X \mid Y=3]\, P\{Y=3\}$$

$$\Rightarrow E[X] = 2 \times \frac{1}{3} + \left(E[X] + 3\right)\frac{1}{3} + \left(E[X]+5\right)\frac{1}{3}$$

$$\Rightarrow 3E[X] = 2+3+5+2E[X]$$

$$\therefore E[X] = 10 \text{ hours.}$$

18

Random Number of Random Variables:

$X \to$ # injuries in any car accident

$N \to$ # of accident occurs.

$X_i \to$ # of injuries in ith accident

$$E\left[\sum_{i=1}^{N} X_i\right] = E\left[E\left[\sum_{i=1}^{N} X_i \mid N\right]\right]$$

$\to$ Total # injuries

Now, $$E\left[\sum_{i=1}^{N} X_i \mid N=n\right] = E\left[\sum_{i=1}^{n} X_i \mid N=n\right]$$

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

$$= E[X_1]+E[X_2]+\cdots+E[X_n] = n\,E[X]$$

Again, $$E\left[E\left[\sum_{i=1}^{N} X_i \mid N\right]\right]$$

$$= E\left[N\,E[X]\right]$$

$$= E[N]\,E[X]$$

Variance of Random number of RV

$$Var\left(\sum_{i=1}^{N} X_i\right) = E\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] - \left(E\left[\sum_{i=1}^{N} X_i\right]\right)^2$$

Now,

$$E\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] = E\left[E\left[\left(\sum_{i=1}^{N} X_i\right)^2 \Big| N\right]\right]$$

$$E\left[\left(\sum_{i=1}^{N} X_i\right)^2 \Big| N=n\right] = Var\left(\sum_{i=1}^{n} X_i\right) + \left(E\left[\sum_{i=1}^{n} X_i\right]\right)^2$$

$$= n\,Var(X) + \left(n\,E[X]\right)^2$$

Again, $E\left[\left(\sum_{i=1}^{N} X_i\right)^2 \Big| N\right] = N\,Var(X) + \left(N\,E[X]\right)^2$

taking $E[\,]$ on both sides,

$$E\left[E\left[\left(\sum_{i=1}^{N} X_i\right)^2 \Big| N\right]\right] = E[N]\,Var(X) + E[N^2]\left(E[X]\right)^2$$

$$\therefore Var\left(\sum_{i=1}^{N} X_i\right) = E[N]\,Var(X) + E[N^2]\left(E[X]\right)^2 - \left(E[N]\,E[X]\right)^2$$

$$= E[N]\,Var(X) + \left(E[X]\right)^2\left\{E[N^2] - \left(E[N]\right)^2\right\}$$

$$= E[N]\,Var(X) + \left(E[X]\right)^2 Var(N)$$

Computing Probability by Conditioning

$E[X] = E[E[X|Y]]$

R.V. $X \longrightarrow E$ event

$$X = \begin{cases} 1 & ; \text{ if } E \text{ occurs} \\ 0 & ; \text{ if } E \text{ doesn't occur} \end{cases}$$

$\hookrightarrow$ Indicator Random Variable

$$E[X] = \sum_x x\, P(x) = 1 \times P\{X=1\} + 0 \times P\{X=0\}$$

$$= P\{X=1\}$$

$$\therefore \boxed{E[X] = P(E)}$$

$$E[X|Y=y] = P\{X|Y=y\}$$

$$P(E) = E[X] = E[E[X|Y]]$$

$$= \sum_y E[X|Y=y]\, P\{Y=y\}$$

$$= \sum_y P\{E|Y=y\}\, P\{Y=y\}$$

Continuous,

$$P(E) = \int_{-\infty}^{\infty} P\{E|Y=y\}\, f_Y(y)\, dy$$

2) Computing Probabilities by Conditioning

Customers entering a webstore buys product with probability p.

Customer arrival → poisson distribution with parameter λ.

$X \rightarrow$ # of products sold

$N \rightarrow$ # of customers entering the website

$$P\{X=0\} = \sum_{n=0}^{\infty} P\{X=0 \mid N=n\} P\{N=n\}$$

$$= \sum_{n=0}^{\infty} (1-p)^n \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= e^{-\lambda} \sum \frac{\{\lambda(1-p)\}^n}{n!}$$

$$= e^{-\lambda} e^{\lambda(1-p)}$$

$$= e^{-\lambda p}$$

Assuming, each customer buy 1 product each.

$$P\{X=k\} = \sum_{n=k}^{\infty} P\{X=k \mid N=n\}\, P\{N=n\}$$

$$= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{e^{-\lambda}\lambda^n}{n!}$$

$$= e^{-\lambda} p^k \sum_{n=k}^{\infty} \frac{\cancel{n!}}{k!\,(n-k)!} \frac{(1-p)^{n-k} \lambda^n}{\cancel{n!}}$$

$$= \frac{e^{-\lambda}(\lambda p)^k}{k!} \sum_{n=k}^{\infty} \frac{\{\lambda(1-p)\}^{n-k}}{(n-k)!}$$

$$= \frac{e^{-\lambda}(\lambda p)^k}{k!} e^{\lambda(1-p)}$$

$$= \frac{e^{-\lambda p}(\lambda p)^k}{k!}$$

$\therefore$ X has a poisson distribution with parameter λp.

The HAT problem:

n men take off their hat, put in a box and take them again.

Match $\longrightarrow$ if a man gets his own hat.

Lets $E \rightarrow$ Event that no matches occur.

$M \longrightarrow$ Event that 1st man gets match.

$M^c \longrightarrow$ Event " " " doesn't get his match.

$$P_n = P(E)$$

$$= P(E|M)\, P(M) + P(E|M^c) \cdot P(M^c)$$

$$= 0 \times \frac{1}{n} + P(E|M^c)\, \frac{n-1}{n}$$

// if 1st man get match $E=0$

$$P(E) = \frac{n-1}{n} P(E|M^c)$$

$$\therefore \boxed{P_n = \frac{n-1}{n} P(E|M^c)}$$

(24)

Now, 2 cases:, if 1st man gets ith man's hat

(i) ith man takes 1st man's hat (swap)

$$\Rightarrow \frac{1}{n-1} P_{n-2}$$

(ii) ith man doesn't take 1st man's hat

$$\Rightarrow P_{n-1}$$

$$P(E|M^c) = \frac{1}{n-1} P_{n-2} + P_{n-1}$$

$$\therefore P_n = \frac{n-1}{n} P_{n-1} + \frac{1}{n} P_{n-2}$$

$$\therefore P_n - P_{n-1} = -\frac{1}{n}\left(P_{n-1} - P_{n-2}\right)$$

Base case. $P_1 = 0$ // only 1 man surely gets match

$$P_2 = \frac{1}{2}$$

$$P_3 - P_2 = -\frac{1}{3}(P_2 - P_1) \cancel{= \frac{1}{2!} - \frac{1}{3!}}$$

$$\Rightarrow P_3 = \frac{1}{2!} - \frac{1}{3!}$$

$$P_4 - P_3 = -\frac{1}{4}(P_3 - P_2)$$

$$\therefore P_4 = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}$$

$$P_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots$$

$$\approx e^{-1} \quad \text{as } n \to \infty$$

<u>Exactly k matches:</u> $P\{\text{Exactly } k \text{ matches}\}$

$$= \binom{n}{k} \frac{1}{n} \frac{1}{n-1} \frac{1}{n-2} \cdots \frac{1}{n-k+1} P_{n-k}$$

$$= \frac{(n-k)!}{n!} \times \frac{n!}{k!\,(n-k)!} \times P_{n-k}$$

$$= \frac{P_{n-k}}{k!}$$

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

$$\approx \frac{e^{-1}}{k!} \quad \text{as } n \to \infty$$

(28)

The Ballot Problem:

(A) $\longrightarrow$ n votes
(B) $\longrightarrow$ m votes

total votes: $(n+m)$
$n > m$
A wins.

Let,

$P_{n,m}$ = Probability that A is always ahead of B.

Conditioning on last vote

$$P_{n,m} = P\{A \text{ always ahead} \mid A \text{ receives last vote}\} \times P\{A \text{ rec. last vote}\} + P\{A \text{ always ahead} \mid B \text{ rec. last vote}\} \times P\{B \text{ rec. last vote}\}$$

$$= P_{n-1,m} \times \frac{n}{n+m} + P_{n,m-1}\frac{m}{n+m}$$

Prove that, $P_{n,m} = \dfrac{n-m}{n+m}$

Proof:

induction on $(n+m)$.

Base Case:

$n+m = 1$; $n > m$ $\Rightarrow n = 1, m = 0$

$$P_{1,0} = \frac{1-0}{1+0} = 1 \text{, true.}$$

Induction step:

Assuming true for, $n+m = k$

$$P_{n,m} = P_{n-1,m} \frac{n}{n+m} \times P_{n,m-1} \frac{m}{n+m}$$

$$= \frac{n-1-m}{n-1+m} \frac{n}{n+m} \times \frac{n-m+1}{n+m-1} \frac{m}{n+m}$$

$$= \frac{n^2 - mn - n + mn - m^2 + m}{(m+n)(m+n-1)}$$

$$= \frac{n^2 - m^2 + m - n}{(m+n)(m+n-1)}$$

$$= \frac{(n-m)(m+n-1)}{(m+n)(m+n-1)}$$

$$= \frac{n-m}{n+m} \quad \text{(proved)}$$

(24)

A List Model

Self organizing file system.

e_1, e_2, e_3, e_4, e_5 (disk)

e_4 $\quad e_4, e_1, e_2, e_3, e_5$

Front of the line rule

E[position of an element to be retrieved]

$P\{e_i \text{ is selected}\} = P_i$

So, $\sum_{i=1}^{n} P_i = 1$

Now, E[Pos. of an elem. to be retr.]

$= \sum_{i=1}^{n} E[\text{Pos of an elem...} \mid e_i \text{ is selected}] \times P\{e_i \text{ is selected}\}$

$= \sum_{i=1}^{n} E[\text{Position of } e_i]\, P_i$

Now, position of $e_i = 1 + \sum_{i \neq j} I_j$

where, $I_j = \begin{cases} 1 ; & \text{If } e_j \text{ preceeds } e_i \\ 0 ; & \text{Otherwise} \end{cases}$

$$E[I_j] = P\{e_j \text{ preceeds } e_i\}$$

$$= P\{e_j \mid e_j \text{ or } e_i\}$$

$$= \frac{P_j}{P_i + P_j}$$

$$E[\text{Position of } e_i] = 1 + \sum_{i \neq j} E[I_j]$$

$$= 1 + \sum_{j \neq i} \frac{P_j}{P_i + P_j}$$

$\therefore E[\text{Pos of elem. to be ret.}]$

$$= \sum_{i=1}^{n} \left(1 + \sum_{i \neq j} \frac{P_j}{P_i + P_j}\right) P_i$$

$$= \sum_{i=1}^{n} P_i + \sum_{i=1}^{N} \sum_{i \neq j} \left(\frac{P_j}{P_i + P_j} \times P_i\right)$$

$$= 1 + \sum_{i=1}^{n} P_i \sum_{j \neq i} \frac{P_j}{P_i + P_j}$$