Random Variables and Probability Distribution

Definitions:

- Random Variable: A random variable is a numerically valued function defined over a sample space S.
- Discrete Random Variable: It is one that can assume a countable number of values.

  Examples:

  (1) # of defective items drawn from a sample of size 10.

  (2) # of people in the waiting line in a doctor's office from a sample of size 100.
- Continuous Random Variable: It is one that can assume infinitely large number of values.

  Examples:

  (1) The failure times of electric bulbs

  (2) The weights of children entering the doctor's office

Probability Distribution for Discrete R.V.'s:

$$0 \leq p(x) \leq 1$$
$$\sum_x p(x) = 1$$

Example: Tossing a coin twice. Let X= number of heads=0,1,2.

$$P(X=0) = \frac{1}{4}$$
$$P(X=1) = \frac{1}{2}$$
$$P(X=2) = \frac{1}{4}.$$

X	0	1	2
$p(x)$	1/4	$1/2$	$1/4$

Example: Out of 5 items, 3 items are defectives. Two items are drawn at random. Let X = number of defective items chosen. That is, $X = 0, 1, 2$.

$$P(X=0) = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}} = 0.1$$

$$P(X=1) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = 0.6$$

$$P(X=2) = \frac{\binom{3}{2}\binom{2}{0}}{\binom{5}{2}} = 0.3$$

X	0	1	2
$p(x)$	0.1	0.6	0.3

Example: Assume $p = P(S)$=(success)=0.4, $q = P(F)$=P(Failure), and X= number of drillings until the first success occurs. So $X = 1, 2, 3, ...$

$$P(X = 1) = P(S) = 0.4$$

$$P(X = 2) = P(FS) = (0.6)(0.4) = 0.24$$

$$P(X = 3) = P(FFS) = (0.6)(0.6)(0.4) = 0.144$$

X	$P(X)$
1	0.4
2	0.24
3	0.144
...	

Expectation and Its Properties:

X	$P(x)$	$xP(x)$	$x^2P(x)$
0	0.1	0	0
1	0.6	0.6	0.6
2	0.3	0.6	1.2
		1.2	1.8

$$EX = \sum_x x\ P(x) = 1.2, \text{ and } EX^2 = \sum_x x^2\ P(x) = 1.8$$

$$\begin{aligned} \sigma^2 = Var(X) = E(X-\mu)^2 &= EX^2 - \mu^2 \\ &= 1.8 - (1.2)^2 = 0.36. \end{aligned}$$

and

$$\sigma = \sqrt{0.36} = 0.6$$

Properties on Expectation:

(i) $E(X + a) = E(X) + a$

(ii)$E(a + bX) = a + bE(X)$

(iii) $E(a + bX + cX^2) = a + bE(X) + cE(X^2)$

(iv) $Var(X + a) = Var(X)$

(v) $Var(aX + b) = a^2\ Var(X)$

(vi) $Std(aX + b) = |a|Std(X)$.

Example: A sales person receives orders for $0, 1, 2, 3$ or 4 units of products each day with probability $0.4, 0.1, 0.3, 0.1$, and 0.1, respectively. His salary is 50 JD per day plus 10 JD for each unit ordered.

X	$P(x)$	$xP(x)$	$x^2P(x)$
0	0.4	0	0
1	0.1	0.1	0.1
2	0.3	0.6	1.2
3	0.1	0.3	0.9
4	0.1	0.4	1.6
		1.4	3.8

His expected salary is $S = 50 + 10E(X) = 50 + 14 = 64$.

Joint Distribution of 2 R.V.'s:

Let

$$X = x_1, x_2, ..., x_k$$

$$Y = y_1, y_2, ..., y_l,$$

with

$$f(x_i, y_j) = P(X = x_i, Y = y_j)$$

Suppose there are 10 cars in a lot of which 5 are in good condition, 2 have defective transmission (DT), and the other 3 have defective steering (DS). 2 cars are chosen at random.

Define

$X = \#$ of cars with DT=0,1,2 and $Y = \#$ of cars with DS=0,1,2.

Cont./ Example:

$$f(0,0) = P(X=0, Y=0) = \frac{\binom{5}{2}}{\binom{10}{2}} = \frac{10}{45}$$

$$f(0,1) = P(X=0, Y=1) = \frac{\binom{2}{0}\binom{3}{1}\binom{5}{1}}{\binom{10}{2}} = \frac{15}{45}$$

$$f(0,2) = P(X=0, Y=2) = \frac{\binom{2}{0}\binom{3}{2}\binom{5}{0}}{\binom{10}{2}} = \frac{3}{45}$$

$$f(1,2) = P(X=1, Y=2) = 0$$

Cont./ Example:

		Y			
		0	1	2	$P(x)$
	0	$10/45$	$15/45$	$3/45$	$28/45$
X	1	$10/45$	$6/45$	0	$16/45$
	2	$1/45$	0	0	$1/45$
	$P(y)$	$21/45$	$21/45$	$3/45$	1

Cont./ Example:

- $$P(X=0)=\frac{28}{45}$$
  $$P(Y=1)=\frac{21}{45}$$
  $$P(X>Y)=f(1,0)+f(2,0)+f(2,1)=\frac{10}{45}+\frac{1}{45}+0=\frac{11}{45}$$

- Find the distribution of $Z=X+Y$.

  Now $Z=0,1,2,3,4$.

  $$P(Z=0)=P(0,0)=\frac{10}{45},\quad P(Z=1)=P(1,0)+P(0,1)=\frac{25}{45}$$
  $$P(Z=2)=P(0,2)+P(2,0)+P(1,1)=\frac{10}{45}$$
  $$P(Z=3)=P(1,2)+P(2,1)=0,\quad P(Z=4)=P(2,2)=0.$$

Cont./ Example:

- Find the mean and Variance of X

X	$P(x)$	$xP(x)$	$x^2P(x)$
0	$28/45$	0	0
1	$16/45$	$16/45$	$16/45$
2	$1/45$	$2/45$	$4/45$
		18/45	20/45

Therefore $EX = 18/45 = 0.4,\ Var(X) = (20/45) - (18/45)^2 = 0.44 - 0.16 = 0.284\ \sigma_X = 0.53$.

Similarly,
$EY = 27/45 = 0.6, Var(X) = (33/45) - (27/45)^2 = 0.373,\ \sigma_Y = 0.61$.

Cont./ Example:

- Covariance and Correlation Coefficient X and Y. It is a measure of association. It is defined as

$$Cov(X,Y) = E(XY) - E(X).E(Y)$$

where $E(XY) = \sum_{x,y}(x\ y)\ p(x,y)$. The correlation coefficient between X and Y is

$$Corr(X,Y) = \rho = \frac{Cov(X,Y)}{\sigma_X\ \sigma_Y}$$

Then

$$E(X\ Y) = (0)(0)(10/45) + (1)(1)(6/45) + (1)(2)(0) + ... + = 6/45,$$

$$Cov(X,Y) = \frac{6}{45} - \frac{18}{45}\frac{27}{45} = 0.13 - 0.24 = -0.11$$

and

$$\rho = Corr(X,Y) = \frac{-0.11}{(0.53)(0.61)} = -0.33$$

Properties on Covariance and Correlation:

- $-1 \leq \rho \leq 1$
- Corr(aX+b, c Y+d)= Corr(X,Y) if a and c have the same signs and Corr(aX+b, c Y+d)= - Corr(X,Y) if a and c have opposite signs
- $Var(X \mp Y) = Var(X) + Var(Y) \mp 2Cco(X, Y)$
- $Var(aX \mp bY + c) = a^2\, Var(X) + b^2\, Var(Y) \mp 2a\, b\, Cov(X, Y)$
- If X and Y are indep. with $EX = 0, Std(X) = 2, EY = -1, Std(Y) = 4$, then

$$Var(X - Y) = 4 + 16 = 20$$
$$Var(0.5X + 0.5Y) = (1/4)(4) + (1/4)(16) = 1 + 4 = 5$$

- If $Cov(X, Y) = 1$, then
  $Var(0.5X + 0.5Y) = (1/4)(4) + (1/4)(16) + 2(1/2)(1/2) = 5.5$