THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

UCS654: Predictive Analytics using Statistics

Topic

Probability Distributions

Example

- Consider a the random experiment be that of throwing a die.
- The six faces of the die can be treated as the six sample points in $S=\{s_1\ s_2\ s_3\ s_4\ s_5\ s_6\}$.
- Let $X(s_i)=i$, which transforms sample space to number line.
- This can be helpful in enquiring the probabilities such as

$$P[\{s:X(s)<=a_1\}]$$

$$P[\{s:X(s)<=c\}]$$

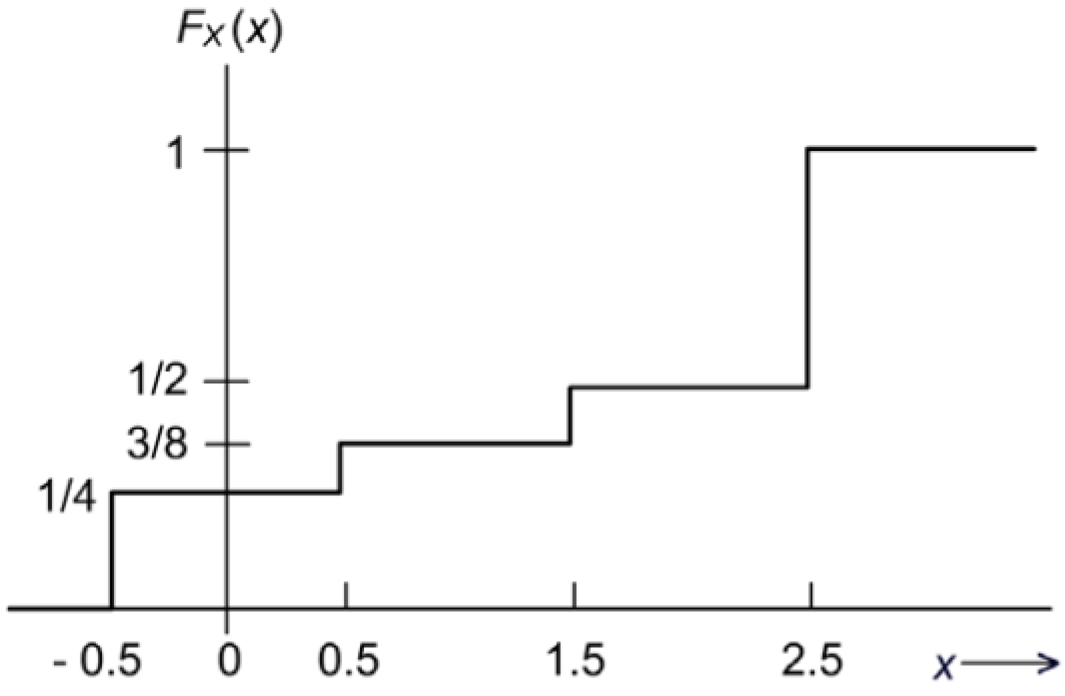
$$P[\{s: b_1 <=X(s)<=b_1\}]$$

If distribution function (cumulative distribution function) of X() is known.

$$F_X(x) = P[\{s: X(s) \leq x\}]$$

Example

- Let $S=\{s_1 \ s_2 \ s_3 \ s_4\}$ with $P(s_1)=1/4$, $P(s_2)=1/8$, $P(s_3)=1/8$, $P(s_4)=1/2$
- Let $X(s_i)=i-1.5$, where i=1,2,3,4, then CDF $F_X(x)$ will as follows.



Properties of CDF

$F_X(\ )$ satisfies the following properties:

i) $F_X(x) \geq 0, \ -\infty < x < \infty$

ii) $F_X(-\infty) = 0$

iii) $F_X(\infty) = 1$

iv) If $a > b$, then $\left[F_X(a) - F_X(b)\right] = P\left[\{s: b < X(s) \leq a\}\right]$

v) If $a > b$, then $F_X(a) \geq F_X(b)$

Exercise-1

Let S be a sample space with six sample points, s_1 to s_6. The events identified on S are the same as above, namely, $A = \{s_1, s_2\}$, $B = \{s_3, s_4, s_5\}$ and $C = \{s_6\}$ with $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{2}$ and $P(C) = \frac{1}{6}$.

Let $Y(\ )$ be the transformation,

$$Y(s_i) = \begin{cases} 1, & i = 1, 2 \\ 2, & i = 3, 4, 5 \\ 3, & i = 6 \end{cases}$$

Show that $Y(\ )$ is a random variable by finding $F_Y(y)$. Sketch $F_Y(y)$.

Probability Density Function

$$f_X(x) = \frac{d F_X(x)}{d x}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(\alpha)\, d\alpha$$

Exercise-2

A random variable X has

$$F_X(x) = \begin{cases} 0 & , \quad x < 0 \\ K x^2 & , \quad 0 \le x \le 10 \\ 100K & , \quad x > 10 \end{cases}$$

(i) Find the constant K

(ii) Evaluate $P[X \le 5]$ and $P[5 < X \le 7]$

(iii) What is $f_x(X)$=?

Exercise-2: Solution

i) $F_X(\infty) = 100K = 1 \Rightarrow K = \dfrac{1}{100}.$

ii) $P(x \le 5) = F_X(5) = \left(\dfrac{1}{100}\right) \times 25 = 0.25$

$P(5 < X \le 7) = F_X(7) - F_X(5) = 0.24$

$$f_X(x) = \frac{d F_X(x)}{d x} = \begin{cases} 0 & , \quad x < 0 \\ 0.02x, & 0 \le x \le 10 \\ 0 & , \quad x > 10 \end{cases}$$