Class No. 5
Date: August 12, 13,14 …
MA 2302: Introduction to Probability and Statistics

RANDOM VARIABLES

Instructor
Prof. Gopal Krishna Panda
Department of Mathematics
NIT Rourkela

Random Variables

Consider the following problem done in the last class:

Three capacitors are chosen at random and with replacement from a lot consisting of 8 capacitors of capacity 500 μF*. and 16 capacitors of capacity 1000* μF*. Find the probability that, out of the three chosen capacitors, (a) there is at least one capacitors of capacity 500* μF*. and (b) not more that two capacitors of capacity 1000* μF*?*

As usual A is the event of choosing a 500 μF capacitor and B, the event of choosing a 1000 μF capacitor from the lot Observe that $B = A^c$. The sample space for this problem is

$$S = \{AAA, AAB, ABA, BAA, ABB, BAB, BBA, BBB\}.$$

Let X be the number of 500 μF capacitor in the sample. Observe that X depends on the outcome of the experiment, that the capacity of the three capacitors drawn. In particular,

$X(AAA) = 3, X(AAB) = 2, (BAB) = 1, X(BBB) = 0$ and so on. Thus, the possible values of X are $0, 1, 2, 3$ and we have seen that $Pr\{X = 0\} = \frac{8}{27}$, $Pr\{X = 1\} = \frac{12}{27}$, $Pr\{X = 2\} = \frac{6}{27}$ and $Pr\{X = 3\} = \frac{1}{27}$. Let us write $f(x) = Pr\{X = x\}$.

Random Variables

Then we can describe these probabilities as:

x:	0	1	2	3
$f(x)$:	$8/27$	$12/27$	$6/27$	$1/27$

In this example, we notice that X is a mapping (function) from the sample space to the real line and the range of X is the set $R = \{0, 1, 2, 3\}$. Observe that $f(a) = Pr\{X = a\} = 0$ if $a \notin R$.

Definition: A random variable is a function from a sample space to the real line. It is called a discrete random variable if its range R is either a finite or a countable set. If R is a finite or infinite interval, then R is called a continuous random variable.

Definition: Given a discrete random X, A function $f: R \rightarrow [0,1]$ as $f(x) = \Pr\{X = x\}$ is known as a probability mass function or simply probability function.

Random variables

For a probability mass function f,

$$f(x) \geq 0 \text{ and } \sum_x f(x) = 1.$$

The distribution of the total probability 1 to the different values of the random variable X is known as a probability distribution of X.

Example 1: Two players A and B go on tossing a fair coin with the agreement that if head appears B has to pay \$1 to A and if tail appears A has to pay \$1 to B. Let X be the gain of A after 3 tosses. Then the sample space is

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}.$$

Observe that the probability of each element of the sample space is $1/8$ and $X(HHH) = 3$, $X(HHT) = 1, \dots, X(TTH) = -1$ and $X(TTT) = -3$.

Random variables

Thus, the range of X is $R = \{-3, -1, 1, 3\}$ and the probability distribution of X is

x	$f(x)$
-3	1/8
-1	3/8
1	3/8
3	1/8

Observe that $f(-1) = \Pr\{X = -1\} = P(\{HTT, THT, TTH\})$, $\quad f(-1) = \Pr\{X = 3\} = P(\{HHH\}) = 1/8$ and so on.

Random variables

Example: There are 25 girls and 35 boys in a class room. 10 questions are to be asked to the students of this class. For each question, a student of the class is randomly chosen without bothering for repetition of the students. Let X of questions asked to the girl students. Then the possible values of X are $0,1,2,\dots,10$. Furthermore, the probability of choosing a girl student is $\frac{25}{60}=\frac{5}{12}$ and that for a boy student is $\frac{7}{12}$. Observe that $X=3$ if exactly 3 questions are asked to girl students and 7 to boy students. However, choosing 3 question out of 10 can be done in $\binom{10}{3}$ ways and probability of choosing 3 girls and 7 boys is equal to $\left(\frac{5}{12}\right)^3\left(\frac{7}{12}\right)^7$. Hence

$$Pr\{X=3\}=\binom{10}{3}\left(\frac{5}{12}\right)^3\left(\frac{7}{12}\right)^7.$$

If there are n questions, proportion of girls is p and that of boys is q, then observe that

$$f(k)=Pr\{X=k\}=\binom{n}{k}p^k q^{n-k}, k=0,1,2,\dots,n.$$

Random variables

Given a random variable X, a function $F: \mathbb{R} \to [0,1]$ defined as $F(x) = \Pr\{X \leq x\}$ is called a *probability distribution function* (or *cumulative distribution function* or simply *distribution function*). Look at the following table:

Sketch of $f(x)$

x	$f(x)$
-3	1/8
-1	3/8
1	3/8
3	1/8

$F(-\infty) = F(-7) = F(-12) = F(-3.1) = 0.$

$F(-3) = F(-2) = F(-1.1) = 1/8.$

$F(-1) = F(0) = F(0.9999) = 4/8.$

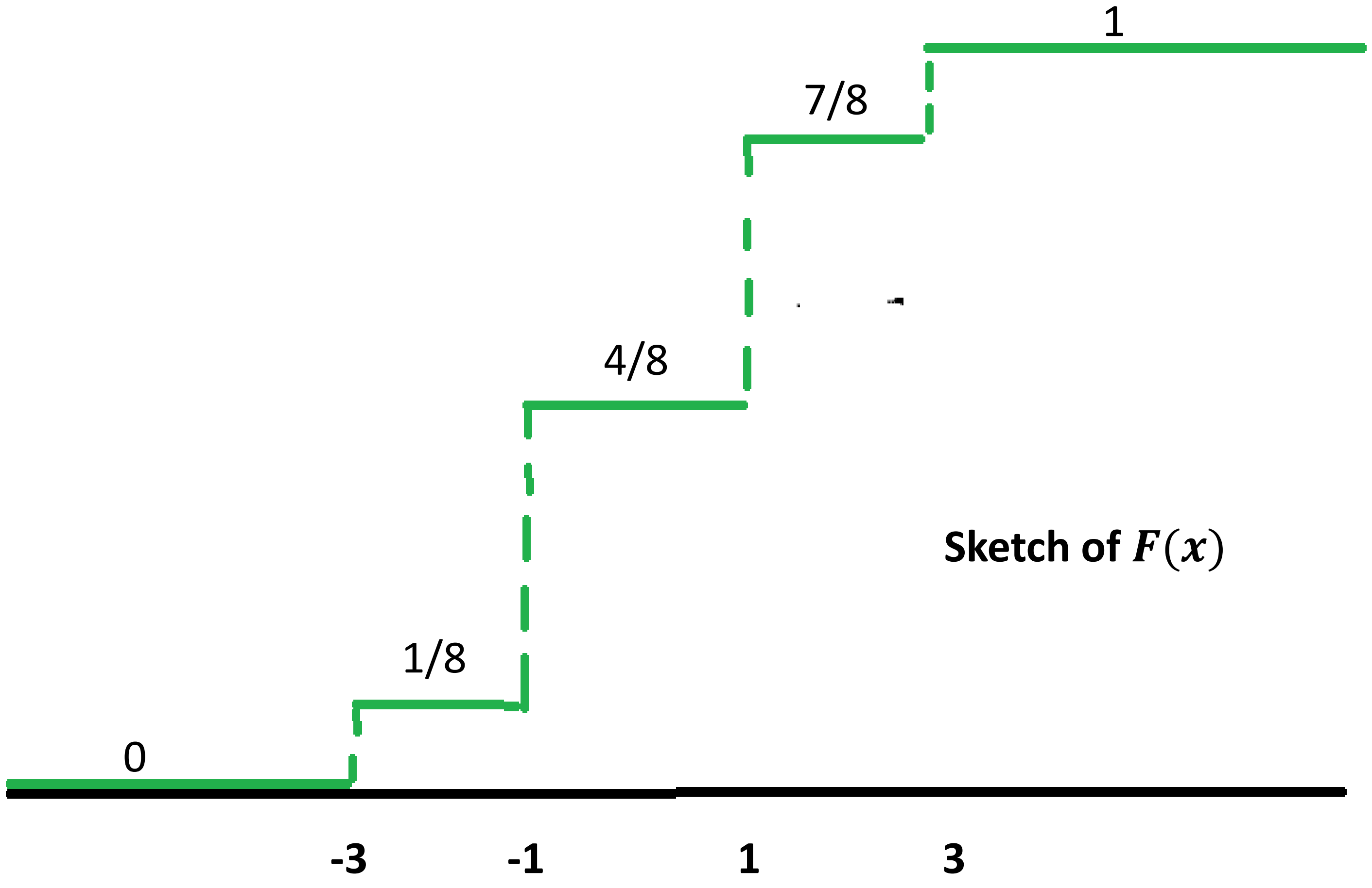
$F(1) = F(2) = F(2.5) = 7/8.$

$F(3) = F(79.67) = F(10^{10}) = F(\infty) = 1.$

The value of the probability distribution function $F(x)$ evaluated at the range of X is given in the following table:

Thus,

$$\boldsymbol{F(x) = \begin{cases} 0 & if\ x < -3, \\ \frac{1}{8} & if - 3 \leq x < -1, \\ \frac{4}{8} & if - 1 \leq x < 1, \\ \frac{7}{8} & if\ 1 \leq x < 3, \\ 1 & if\ x \geq 3. \end{cases}}$$



Sketch of $F(x)$

Observe that the function $F(x)$ is right continuous, i.e. $\lim_{x \to a^+} F(x) = F(a)$ for each $a \in \mathbb{R}$.

Random variables

Example: Two fair dice are thrown. Let X be the sum of points arising out of one throw. Then, the possible values of X are $x = 2,3,4,5,6,7,8,9,10,11,12$. Observe that, in this case the sample space is

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

$\Pr\{X = 2\} = P\{11\} = \frac{1}{36}$, $Pr\{X = 3\} = P\{12,21\} = \frac{2}{36}$, $\Pr\{X = 4\} = P\{13,22,31\} = \frac{3}{36}$, ..., $\Pr\{X = 11\} = P\{56,65\} = 2/36$, $\Pr\{X = 12\} = P\{66\} = 1/36$. The following table gives the values of $f(x)$ for different values of x.

Hence,

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$$\boldsymbol{F}(x) = \begin{cases} 0 & \text{if } x < 2, \\ \frac{1}{36} & \text{if } 2 \le x < 3, \\ \frac{3}{36} & \text{if } 3 \le x < 4, \\ \frac{6}{36} & \text{if } 4 \le x < 5, \\ \frac{10}{36} & \text{if } 5 \le x < 6,. \end{cases}$$ and $$F(x) = \begin{cases} \frac{15}{36} & \text{if } 6 \le x < 7, \\ \frac{21}{36} & \text{if } 7 \le x < 8, \\ \frac{26}{36} & \text{if } 8 \le x < 9, \\ \frac{30}{36} & \text{if } 9 \le x < 10, \\ \frac{33}{36} & \text{if } 10 \le x < 11, \\ \frac{35}{36} & \text{if } 11 \le x < 12, \\ 1 & \text{if } x \ge 12, \end{cases}$$ and one can sketch $f(x)$ and $F(x)$ as usual.

Random Variables

Observe that $F(x) = \Pr\{X \leq x\}$. Hence,

$$\Pr\{\mathrm{X} \leq a\} + \Pr\{\mathrm{a} < \mathrm{X} \leq b\} = \Pr\{\mathrm{X} \leq b\}.$$

Hence,

$$\Pr\{\mathrm{a} < \mathrm{X} \leq b\} = \Pr\{\mathrm{X} \leq b\} - \Pr\{\mathrm{X} \leq a\} = F(b) - F(a)$$

which holds for both discrete random variables. The probability (or cumulative) distribution function of a continuous random variable X is given by

$$F(x) = \int_{-\infty}^{x} f(u)\, du$$

and f is called the *probability density function*, or *density fun*ction or simply a *density* or, in short, *pdf*. Thus, at each point x where F is differentiable, $F'(x) = f(x)$. The pdf of a continuous random variable satisfies

$$f(x) \geq 0 \text{ for all } x \in \mathbb{R} \text{ and } \int_{-\infty}^{\infty} f(x) dx = 1\,.$$

In case of a continuous random variable, the total probability 1 is distributed in some interval and hence $\Pr\{X = a\} = 0$ for each $a \in \mathbb{R}$.

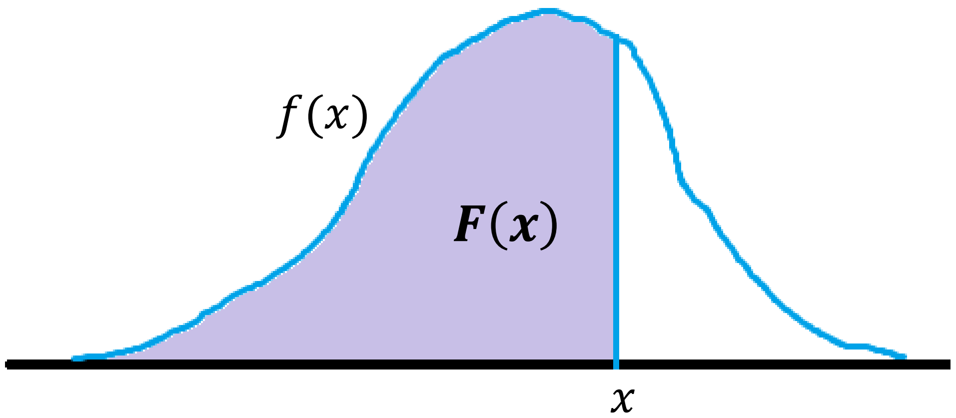
Random Variables

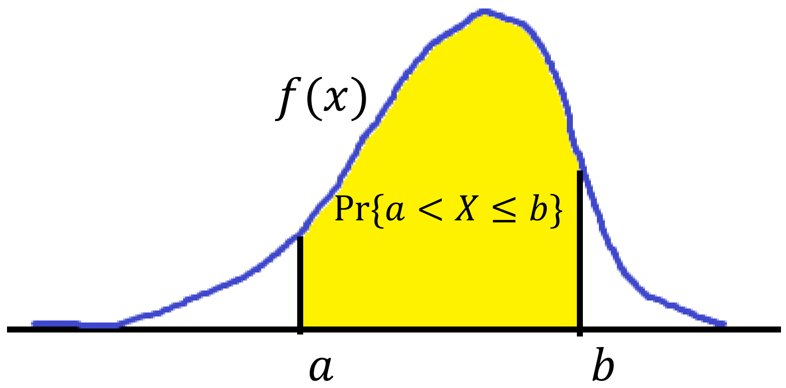
Observe that $F(x) = \Pr\{X \leq x\}$. Hence,

$$\Pr\{X \leq a\} + \Pr\{a < X \leq b\} = \Pr\{X \leq b\}.$$

Hence,

$$\Pr\{a < X \leq b\} = \Pr\{X \leq b\} - \Pr\{X \leq a\} = F(b) - F(a)$$





Example: Let X be a discrete random variable with pmf $f(x) = Kx^2, x = 1,2,3,4$. Find K, and the probability (cumulative) distribution function $F(x)$. Sketch both $f(x)$ and $F(x)$.

Ans. Since sum of all probabilities is equal to 1, we have

$$K(1^2 + 2^2 + 3^2 + 4^2) = 1.$$

Hence $K = \frac{1}{30}$ and $f(x) = \frac{x^2}{30}, x = 1,2,3,4$. Rest is to be done by you.

Random Variables

Problem 1: Find the density function whose cdf is $F(x) = \begin{cases} 0 & \text{if } x \le 0 \\ 1 - e^{-3x} & \text{if } x > 0. \end{cases}$ Sketch both $f(x)$ and $F(x)$.

Ans. The function $F(x)$ is differentiable except at $x = 0$. Hence,

$$f(x) = F'(x) = \begin{cases} 0 & if\ x < 0, \\ 3e^{-3x} & if\ x > 0. \end{cases}$$

Now sketch $f(x)$ and $F(x)$ as usual.

Problem 2: Let X [millimeters] be the thickness of washers. Assume that X has the density $f(x) = kx$ if $0.9 < x < 1.1$ and $f(x) = 0$ elsewhere. Find k. What is the probability that a washer will have thickness between 0.95 mm and 1.05 mm?

Ans. k: $\int_{-\infty}^{\infty} f(x)\, dx = 1$. Hence $k \int_{0.9}^{1.1} x\, dx = 1.$ Thus, $k = 5$. probability that a washer will have thickness between 0.95 mm and 1.05 mm is equal to

$$\Pr\{0.95 \le X \le 1.05\} = \int_{0.95}^{1.05} 5x\, dx = 0.5\, .$$

Random Variables

Problem 2: Find the probability that none of three bulbs in a traffic signal will have to be replaced during the first 1500 hours of operation if the lifetime X of a bulb is a random variable with density $f(x) = 6\{0.25 - (x - 1.5)^2\}$ if $1 \leq x \leq 2$ and $f(x) = 0$ otherwise, where x is measured in multiples of 1000 hours.

Ans. Probability of no failure of any given bulb in first 1500 hr of operation

$$p = \Pr\{X > 1.5\} = \int_{1.5}^{\infty} f(x)\ dx = \int_{1.5}^{2} 6\{0.25 - (x - 1.5)^2\}\ dx$$

To make the calculation simple use the transformation $y = x - 1.5$. Then

$$p = \int_{0}^{0.5} 6\{0.25 - y^2\}\ dy = 1.5 \times 0.5 - \frac{0.5^3}{3} 0.125 = \frac{2.25 - 0.125}{3} = \frac{17}{24}$$

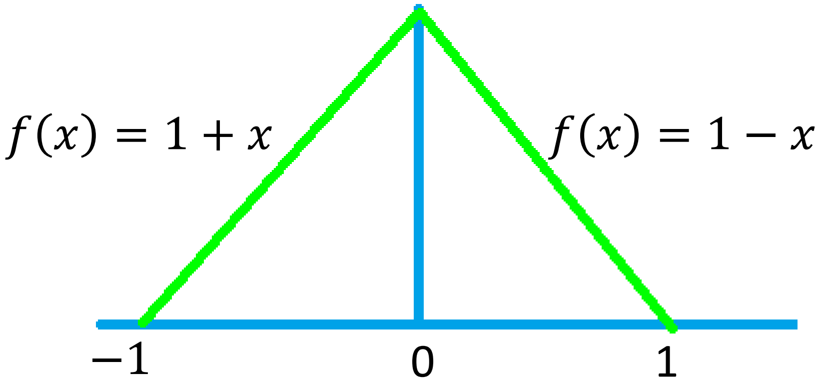
The probability that none of three bulbs in a traffic signal will have to be replaced during the first 1500 hours $= \left(\frac{17}{24}\right)^3 = 0.3554.$

Random Variables

Problem 3: Suppose that in an automatic process of filling oil cans, the content of a can (in gallons) is $Y = 100 + X$, where X is a random variable with density $f(x) = 1 - |x|$ if $|x| \leq 1$ and $f(x) = 0$ when $|x| > 1$. Sketch $f(x)$ and $F(x)$. In a lot of 1000 cans, about how many will contain 100 gallons or more? What is the probability that a can will contain less than 99.5 gallons? Less than 99 gallons?.

Ans. Given that

$$f(x) = \begin{cases} 1 + x & \text{if } -1 \leq x \leq 0 \\ 1 - x & \text{if } \quad 0 \leq x \leq 1 \\ 0 & \qquad \text{otherwise.} \end{cases}$$



Sketch of $f(x)$

Thus, for $x < -1$, $f(x) = 0$. If $-1 \leq x \leq 0$, then

$$F(x) = \int_{-\infty}^{x} f(u)\, du = \int_{-1}^{x} (1 + u)\ du = \frac{(1 + x)^2}{2} = \frac{1}{2} + x + \frac{x^2}{2}.$$

Thus, $F(0) = \int_{-\infty}^{0} f(u) du = \frac{1}{2}$. If $0 \leq x \leq 1$, then

$$F(x) = \int_{-\infty}^{x} f(u)\, du = \int_{-\infty}^{0} f(u)\, du + \int_{0}^{x} (1 - u)\ du = 1 - \frac{(1 - x)^2}{2} = \frac{1}{2} + x - \frac{x^2}{2}.$$

If $x \geq 1$, then $F(x) = 1$.

Thus,

$$F(x) = \begin{cases} 0 & if\ x < -1 \\ \dfrac{1}{2} + x + \dfrac{x^2}{2} & if - 1 < x < 0 \\ \dfrac{1}{2} + x - \dfrac{x^2}{2} & if\ \ 0 < x < 1 \\ 1 & if\ x \geq 1. \end{cases}$$

Now, the probability that a can will contain 100 gallons or more is equal to

$$\Pr\{Y \geq 100\} = \Pr\{X \geq 0\} = \int_0^\infty f(x)\,dx = \int_0^1 (1 - x)\,dx = \frac{1}{2}\,.$$

Thus, out of 1000 cans about 500 cans will contain 100 gallons or more.

The probability that a can will contain less than 99.5 gallons is equal to

$$\Pr\{Y < 99.5\} = \Pr\{X < -0.5\} = \int_{-\infty}^{-0.5} f(x)\,dx = \int_{-1}^{-0.5} (1 + x)\,dx = \frac{1}{8}\,.$$

The probability that a can will contain less than 99 gallons is equal to

$$\Pr\{Y < 99\} = \Pr\{X < -1\} = \int_{-\infty}^{-1} f(x)\,dx = 0\,.$$

Random Variables

Problem 4: A packet contains 3 hard drives of capacity 1 TB and 7 hard drives of capacity 2 TB. Hard drives are drawn one by one and with replacement till a hard drive of capacity 2 TB is drawn. Let X be the number of drives chosen till the first 2 TB drive is chose. Find the pmf and cdf of X. Find the probability of drawing a 2 TB hard drive not before the third draw?

Ans. Let A be the event of choosing a 2 TB hard drive and B the event of choosing a 1 TB hard drive in one draw. Then $P(A) = \frac{7}{10}$, $P(B) = \frac{3}{10}$. The first 2 TB hard drive can be chosen as follows

$$A, BA, BBA, BBBA, \ldots$$

Hence the possible values of X are $1, 2, 3, \ldots$ and

$$f(x) = \Pr\{X = x\} = P(BBB \ldots \ldots BA) = \left(\frac{3}{10}\right)^{x-1} \times \frac{7}{10}.$$

Now you prepare tables for $f(x)$ and $F(x)$ and do the rest work.