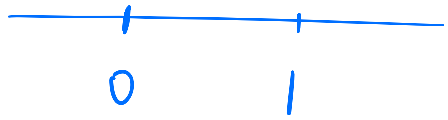
Section 2.2 The Distribution Function of a Random Variable

School of Sciences, BUPT

Attandance:

1\. Tossing a fair coin. Let

$$X(\omega)=\begin{cases}1, & \omega=H\\ 0, & \omega=T.\end{cases}$$



Find $F_X(x)$.

2\. Selecting a point from $[0,1]$ at random.

Let $Y(\omega)=\omega,\quad \omega\in[0,1]$

Find $F_Y(y)$.

1\. $$\begin{cases}P(X=1)=\frac{1}{2}=P(\{\omega:\ X(\omega)=1\})\\ P(X=0)=\frac{1}{2}\end{cases}$$

X	0	1
$p(x)$	$\frac{1}{2}$	$\frac{1}{2}$

$p(x) \Longleftrightarrow F(x)$

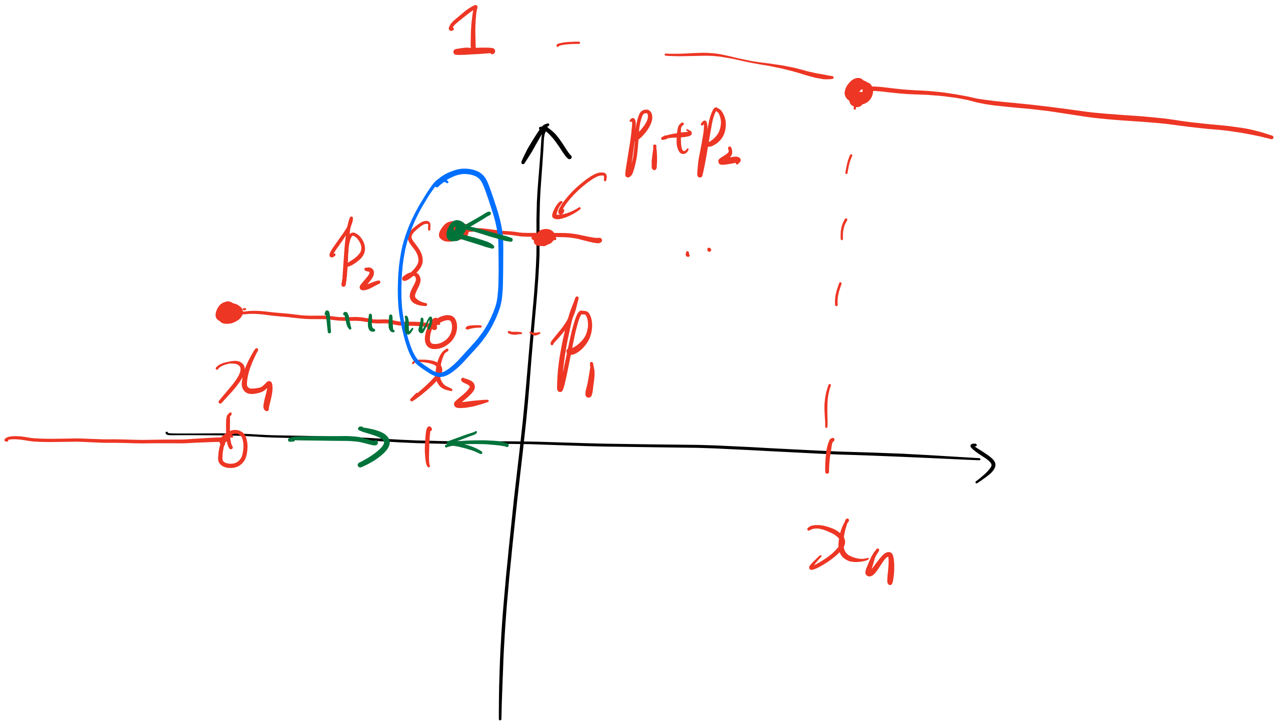
X	x_1	x_2	$\cdots$	x_n	$(\cdots\cdots)$
$p(x)$	p_1	p_2	$\cdots$	p_n	$(\cdots)$

p.f. ① $p(x) \Rightarrow F(x)$

$$F(x) = P(X \le x) = \begin{cases} 0, & x < x_1 \\ p_1, & x \in [x_1, x_2) \\ p_1 + p_2, & x \in [x_2, x_3) \\ \cdots\cdots & \\ 1, & x \ge x_n \end{cases}$$

$$F(x) = \sum_{x_k \le x} p_k$$

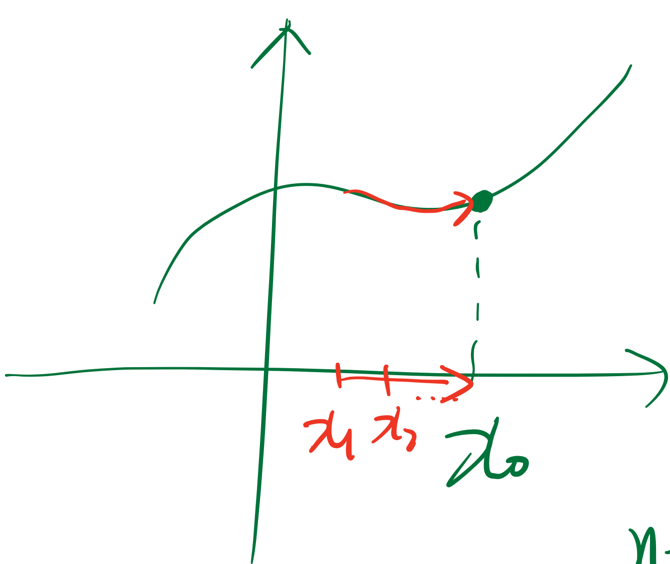
$$P(X \in A) = \sum_{x_k \in A} p_k$$



② $F(x) \Longrightarrow p(x)$

$$p_k = P(X = x_k) = F(x_k) - F(x_k-)$$

$$k = 1, 2, \ldots (n)$$



$g(x)$

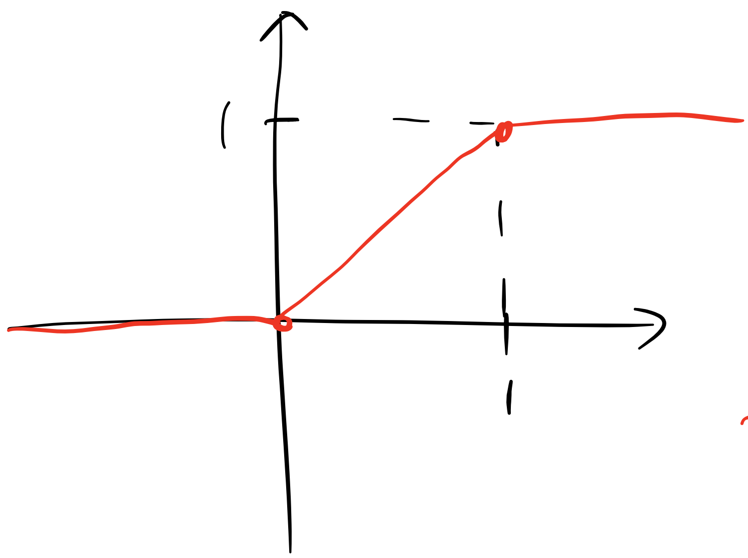
$\{x_n\} \uparrow x$

$$\lim_{n\to\infty} g(x_n) = g(x-)$$

2.

$$F_Y(y) = \begin{cases} 0 & y \le 0 \\ y & 0 \le y < 1 \\ 1 & y \ge 1 \end{cases}$$

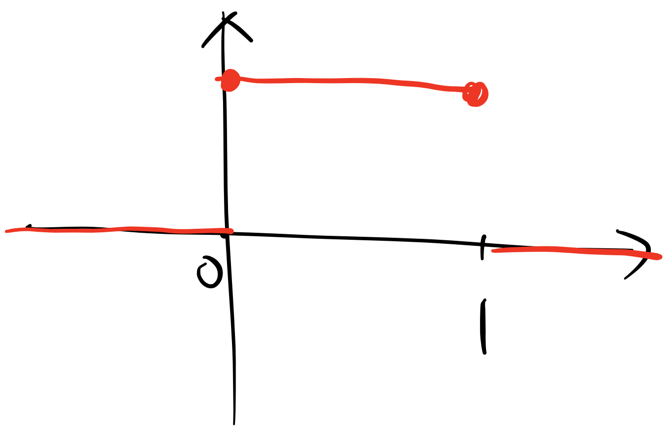
$$P(Y \le 0.3) = 0.3$$



$F_Y(y)$

$$f_Y(y) = F_Y'(y)$$

$$= \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

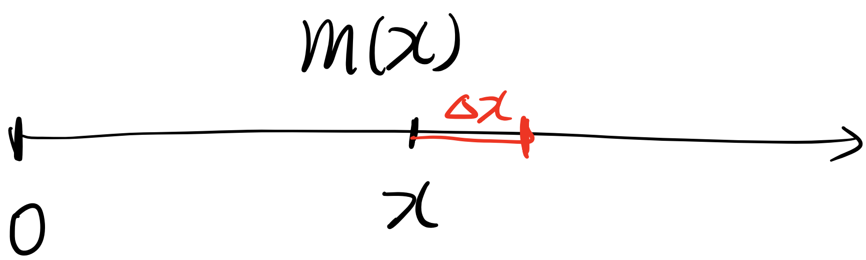


$$\Omega \xrightarrow{X(\omega)} \mathbb{R} \xrightarrow{F(x)} [0,1]$$

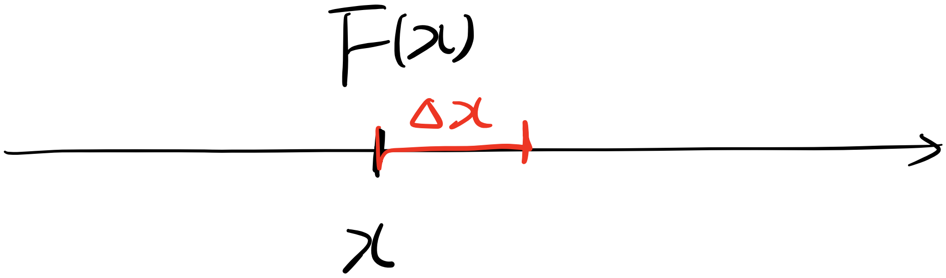
$F(x)$. d.f. p.f.

$\begin{cases} \text{discrete: } p(x) & \text{probability (mass) function} \\ \text{continuous: } f(x). & \text{prob. density function} \end{cases}$

p.d.f.



$$d(x) = \lim_{\Delta x \to 0} \frac{m(x+\Delta x) - m(x)}{\Delta x} = m'(x)$$

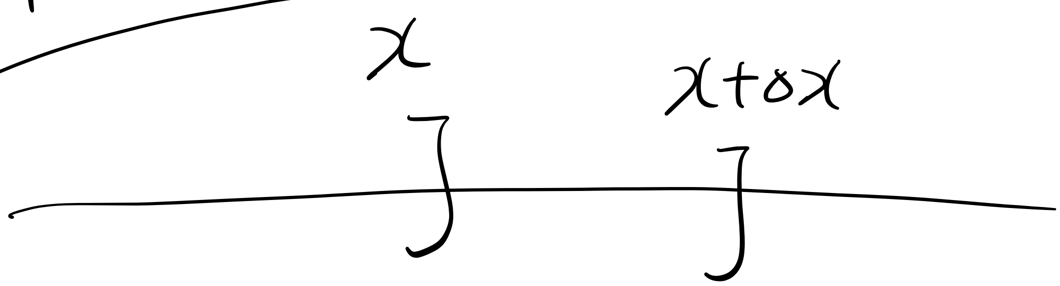


$$f(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = F'(x)$$

$$F(x+\Delta x) - F(x) = P(X \le x+\Delta x) - P(X \le x)$$

$A \supset B$

$P(A) - P(B) = P(A-B)$



$$= P(X \in (x, x+\Delta x])$$

Contents

$Y = g(X)$

$F_Y(y) = ?$

The Definition of Distribution Function

Definition

The function $F(x)$ that associates with each real number x the probability $P(X \leqslant x)$ that the random variable X takes on a value smaller than or equal to this number is called the ***distribution function*** of X. *That is*

$$F(x) = P(X \leqslant x), \qquad \forall\, x \in \mathbb{R}. \tag{1}$$

The abbreviation for distribution function is d.f.. Some authors use the term ***cumulative distribution function***, instead of distribution function, and use the abbreviation c.d.f..

累积

The Properties of Distribution Function

Proposition

The d.f. $F(x)$ of every random variable X have the following three properties.
(i) *It is nondecreasing as x increases; that is, if $x_1 < x_2$, then $F(x_1) \leqslant F(x_2)$.*
(ii) $\lim_{x \to -\infty} F(x) = 0$ *and* $\lim_{x \to +\infty} F(x) = 1$.
(iii) *It is always right-continuous; that is, $F(x) = F(x^+)$ at every point x.*

Proof. (i) If $x_1 < x_2$, then $\{X \leqslant x_1\} \subseteq \{X \leqslant x_2\}$. Hence, $P(X \leqslant x_1) \leqslant P(X \leqslant x_2)$.
(ii) These limiting values follow directly from the fact that $P(X \leqslant x)$ must approach 0 as $x \to -\infty$ and $P(X \leqslant x)$ must approach 1 as $x \to +\infty$. These relations can in turn be rigorously established by using the continuity of probability.

The Properties of Distribution Function

(iii) It is always right-continuous; that is, $F(x) = F(x^+)$ at every point x.

$\{y_n\} \downarrow x$

Proof. (iii) Let $y_1 > y_2 > \cdots$ be a sequence of numbers that are decreasing such that $\lim_{n\to\infty} y_n = x$. Then the event $\{X \leqslant x\}$ is the intersection of all the events $\{X \leqslant y_n\}$ for $n = 1, 2, \cdots$. Hence, by the continuity of probability,

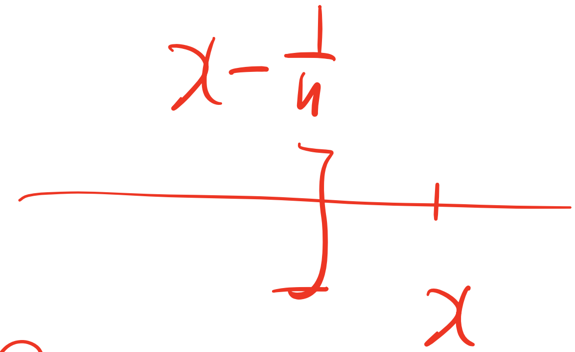
$$F(x) = P(X \leqslant x) = \lim_{n\to\infty} P(X \leqslant y_n) = F(x^+).$$

$$\{X \leq x\} = \bigcap_{n=1}^{\infty} \{X \leq x + \frac{1}{n}\}$$

□

$\{y_n\} \uparrow x$

$\{x - \frac{1}{n}\}$



$$\{X \quad x\} = \bigcup_{n=1}^{\infty} \{X \le x - \frac{1}{n}\}$$

The Properties of Distribution Function

Let I be a subset of the real line and

$$P(X \in I) = P(\omega : X(\omega) \in I).$$

That is to say, by means of random variable, we could use $\{X \in I\}$ denote random events which is different from the way we used in Chapter 1.

The form of the subset I can be various. For any $x_1, x_2 \in \mathbb{R}$, A could be any of the following subsets:

$$(-\infty, x_1),\ (-\infty, x_1],\ \{x_1\},\ (x_1, x_2),\ (x_1, x_2],$$

$$[x_1, x_2],\ [x_1, x_2),\ [x_2, +\infty),\ (x_2, +\infty).$$

The Properties of Distribution Function

Theorem

(i) *For every value* x, $P(X > x) = 1 - F(x)$.

(ii) *For all values* x_1 *and* x_2 *such that* $x_1 < x_2$,

$(x_1, x_2]$

$$P(x_1 < X \leqslant x_2) = F(x_2) - F(x_1).$$

$\Rightarrow f(x)$

(iii) *For each value* x, $P(X < x) = F(x^-)$.

$P(X = x) = F(x) - F(x-)$

Proof. (i) $P(X > x) = 1 - P(X \leqslant x) = 1 - F(x)$.

(ii)

$$\begin{aligned} P(x_1 < X \leqslant x_2) &= P(\{X \leqslant x_2\} - \{X \leqslant x_1\}) \\ &= P(X \leqslant x_2) - P(X \leqslant x_1) \\ &= F(x_2) - F(x_1). \end{aligned}$$

(iii) For each value x, $P(X < x) = F(x^-)$.

Proof. (iii) Let $y_1 < y_2 < \cdots$ be an increasing sequence of numbers such that $\lim_{n\to\infty} y_n = x$. Then

$$\{X < x\} = \bigcup_{n=1}^{\infty} \{X \leqslant y_n\}.$$

Thus

$$P(X < x) = \lim_{n\to\infty} P(X \leqslant y_n) = \lim_{n\to\infty} F(y_n) = F(x^-).$$

$\square$

The Properties of Distribution Function

Corralary

For every value x, $\Rightarrow p(x)$

$$P(X = x) = F(x) - F(x^-) \quad \text{and} \quad P(X \geqslant x) = 1 - F(x^-).$$

For all values x_1 *and* x_2 *such that* $x_1 < x_2$,

$$P(x_1 < X < x_2) = F(x_2^-) - F(x_1),$$

$$P(x_1 \leqslant X < x_2) = F(x_2^-) - F(x_1^-),$$

and

$$P(x_1 \leqslant X \leqslant x_2) = F(x_2) - F(x_1^-).$$

The Definition and Properties of Distribution Function

Example

Suppose that the d.f. F of a random variable X is as following. Find each of the following probabilities:

(a) $P(X = -1)$	(b) $P(X < 0)$	(c) $P(X \leqslant 0)$
(d) $P(X = 1)$	(e) $P(0 < X \leqslant 3)$	(f) $P(0 < X < 3)$
(g) $P(0 \leqslant X \leqslant 3)$	(h) $P(1 < X \leqslant 2)$	(i) $P(1 \leqslant X \leqslant 2)$
(j)$P(X > 5)$	(k) $P(X \geqslant 5)$	(l) $P(3 \leqslant X \leqslant 4)$

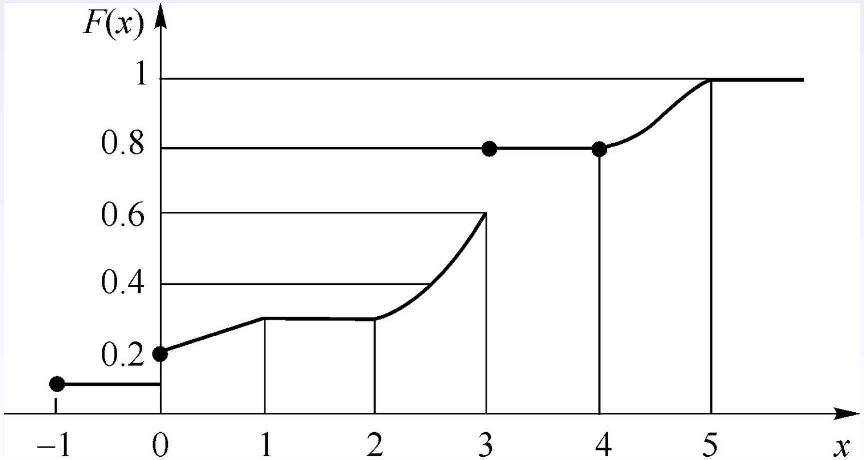


Figure: d.f. of Example 1

Solution.

(a) $P(X = -1) = F(-1) - F(-1^-) = 0.1 - 0 = 0.1,$

(b) $P(X < 0) = F(0^-) = 0.1,$

(c) $P(X \leqslant 0) = F(0) = 0.2,$

(d) $P(X = 1) = F(1) - F(1^-) = 0.3 - 0.3 = 0,$

(e) $P(0 < X \leqslant 3) = F(3) - F(0) = 0.8 - 0.2 = 0.6,$

(f) $P(0 < X < 3) = F(3^-) - F(0) = 0.6 - 0.2 = 0.4,$

(g) $P(0 \leqslant X \leqslant 3) = F(3) - F(0^-) = 0.8 - 0.1 = 0.7,$

(h) $P(1 < X \leqslant 2) = F(2) - F(1) = 0.3 - 0.3 = 0,$

(i) $P(1 \leqslant X \leqslant 2) = F(2) - F(1^-) = 0.3 - 0.3 = 0,$

(j) $P(X > 5) = 1 - P(X \leqslant 5) = 1 - F(5) = 1 - 1 = 0,$

(k) $P(X \geqslant 5) = 1 - F(5^-) = 1 - 1 = 0,$

(l) $P(3 \leqslant X \leqslant 4) = F(4) - F(3^-) = 0.8 - 0.6 = 0.2.$ $\square$

The Properties of Distribution Function

Example

Are the following function d.f. of some random variables?

$$F_1(x) = \begin{cases} 0 & \text{for } x < -2, \\ 1/2 & \text{for } \quad -2 \leqslant x < 0, \\ \cancel{2} & \text{for } \quad x \geqslant 0, \end{cases}$$

Solution. Since $\lim\limits_{x \to +\infty} F_1(x) = 2 > 1$, $F_1(x)$ can not be a d.f..

The Properties of Distribution Function

Example

Are the following function d.f. of some random variables?

$$F_2(x) = \begin{cases} 0 & \text{for } x < 0, \\ \sin x & \text{for } 0 \leqslant x < \pi/2, \\ 1 & \text{for } x \geqslant \pi, \end{cases}$$

Solution. Since $F_2(x) = \sin x$ is decreasing in the interval $(\pi/2, \pi)$ as x increases, $F_2(x)$ can not be a d.f..

The Properties of Distribution Function

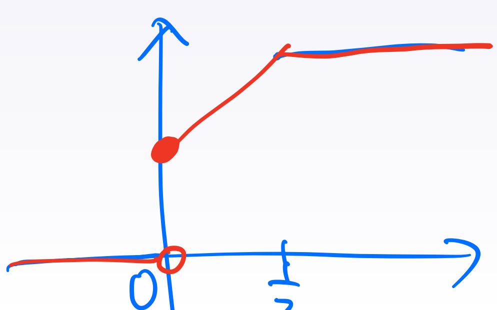
Example

Are the following function d.f. of some random variables?

$$F_3(x) = \begin{cases} 0 & \text{for } x < 0, \\ x + \dfrac{1}{2} & \text{for } 0 \leqslant x < 1/2, \\ 1 & \text{for } x \geqslant 1/2. \end{cases}$$

Solution. The function $F_3(x)$ might be a d.f. of some random variable since it is satisfied the three conditions in Proposition. (In fact, it must be a d.f. of some random variable.)

$P(\frac{1}{4} < X < \frac{1}{3})$



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

$P(X \in (0, \frac{1}{2})) = \frac{1}{2}$

The Properties of Distribution Function

Example

Suppose that the d.f. of a random variable X is

$$F(x) = \begin{cases} 0 & \text{for } x < -1, \\ a & \text{for } -1 \leqslant x < 1, \\ 2/3 - a & \text{for } 1 \leqslant x < 2, \\ a + b & \text{for } x \geqslant 2, \end{cases}$$

and $P(X = 1) = 1/2$. Find the value of a and b.

Solution. Since $\lim_{x \to +\infty} F(x) = 1$, $a + b = 1$. Because $P(X = 1) = F(1) - F(1^-)$, $2/3 - a - a = 1/2$. So we have $a = 1/12, b = 11/12$. □

Contents

Discrete Random Variable

Observing the above examples, we can find the possible values of some random variables are finite, some are countably infinite and some are uncountably infinite. According to this point, we could divide the random variables into different categories.

Definition

A random variable is said to be of ***discrete type*** *if the number of different values it can take is finite or countably infinite.*

Question

Can we call a random variable continuous if it may take an uncountably infinite number of values? In most cases it is, but sometimes it is not.

Value of X

finite

countable

D.

infinite

un countable

C.

Discrete Random Variable

Example

Suppose that our experiment consists of tossing two fair coins. Let X denote the number of heads appearing. Then X is a random variable taking on one of the values 0, 1, 2 with respective probabilities

$$\begin{aligned} P(X=0) &= P(\omega \mid X(\omega)=0) = P(\{TT\}) = 1/4, \\ P(X=1) &= P(\omega \mid X(\omega)=1) = P(\{TH, HT\}) = 2/4, \\ P(X=2) &= P(\omega \mid X(\omega)=2) = P(\{HH\}) = 1/4. \end{aligned}$$

Find the d.f. of the random variable X.

$$F(x) = P(X \leqslant x) = \sum_{x_k \leqslant x} P(X = x_k)$$

Discrete Random Variable

Solution. If $x < 0$, then $F(x) = P(X \leqslant x) = P(\emptyset) = 0$.
If $0 \leqslant x < 1$, then $F(x) = P(X \leqslant x) = P(X = 0) = 1/4$.
If $1 \leqslant x < 2$, then

$$F(x) = P(X \leqslant x) = P(X = 0 \text{ or } 1)$$

$$= P(X = 0) + P(X = 1) = 1/4 + 2/4 = 3/4.$$

If $x \geqslant 2$, then

$$F(x) = P(X \leqslant x) = P(X = 0,\ 1 \text{ or } 2)$$

$$= P(X = 0) + P(X = 1) + P(X = 2) = 1.$$

That is,

$$F(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1/4 & \text{for } 0 \leqslant x < 1, \\ 3/4 & \text{for } 1 \leqslant x < 2, \\ 1 & \text{for } x \geqslant 2. \end{cases}$$

Discrete Random Variable

For a discrete random variable X, we define the ***probability (mass) function*** $p(x)$ of X by $p(x) = P(X = x)$. The abbreviation for probability function is p.f..

$\cancel{X}$	x_1	x_2	$\cdots$	x_n	$(\cdots)$
$p(x)$	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$(\cdots)$

Table: Tabular form of probability function

$$P(X \in A) = \sum_{x_k \in A} P(X = x_k) = \sum_{x_k \in A} p(x_k). \qquad (*)$$

Let $\{x_1, x_2, \cdots\}$ be the set of possible values of the discrete random variable X. The function p has the following properties:

(i) $p(x_k) \geqslant 0$ for all x_k; $p(x) = 0$ for all other values of x,

(ii) $\sum_{k=1}^{\infty} p(x_k) = 1.$

$$\begin{cases} f(x) \geq 0 \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{cases}$$

Discrete Random Variable

Example

A sample space Ω containing n outcomes $\omega_1, \cdots, \omega_n$ is called a simple sample space if the probability assigned to each of the outcomes $\omega_1, \cdots, \omega_n$ is $1/n$. Let $X(\omega_i) = i/n$, $i = 1, 2, \cdots, n$. The p.f. of X is given by $p(i/n) = 1/n$, where $i = 1, 2, \cdots, n$. $\square$

x	$1/n$	$2/n$	$\cdots$	1
$p(x)$	$1/n$	$1/n$	$\cdots$	$1/n$

Discrete Random Variable

Example

Suppose that the p.f. of random variable X is given by the following table. Find the values of the constant a in the table, $P(X > -1/2)$, and $P(-3 < X < 1)$.

x	-2	0	3	5
$p(x)$	$1/4$	a	$1/2$	$1/12$

Sum $= 1$

Solution. Since $1/4 + a + 1/2 + 1/12 = 1$, $a = 1/6$.
$P(X > -1/2) = P(X = 0) + P(X = 1) + P(X = 3) = 1/6 + 1/2 + 1/12 = 3/4$.
$P(-3 < X < 1) = P(X = -2) + P(X = 0) = 1/4 + 1/6 = 5/12$. □

$$P(X \in A) = \sum_{x_k \in A} p(x_k)$$

Discrete Random Variable

Example

If the d.f. of random variable X is the same as on the blackboard. Find the p.f. of random variable X. $F(x) \Rightarrow p(x)$

Solution. When $x \in (-\infty, 0)$, $(0, 1)$, $(1, 2)$ or $(2, +\infty)$, respectively, we have

p_k

$$P(X = x) = F(x) - F(x^-) = 0.$$

When $x = 0$, $P(X = 0) = F(0) - F(0^-) = 1/4 - 0 = 1/4.$
When $x = 1$, $P(X = 1) = F(1) - F(1^-) = 3/4 - 1/4 = 1/2.$
When $x = 2$, $P(X = 2) = F(2) - F(2^-) = 1 - 3/4 = 1/4.$ $\square$

x_k

For discrete random variable, using p.f. is more straightforward than using d.f. in many cases. Because we can easily know all possible values taken by X and the corresponding probabilities.

Contents

Continuous Random Variable

Definition

We call X a ***continuous random variable*** *if there is a function f defined for all $x \in \mathbb{R}$ and having the following properties:*

(i) $f(x) \geqslant 0$ *for any real number* x,

(ii) *if A is any subset of* $\mathbb{R}$, *then*

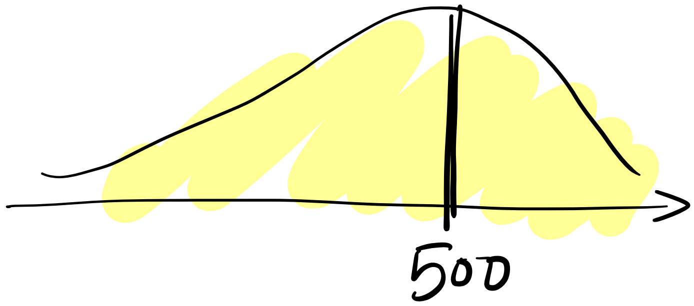
$$P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

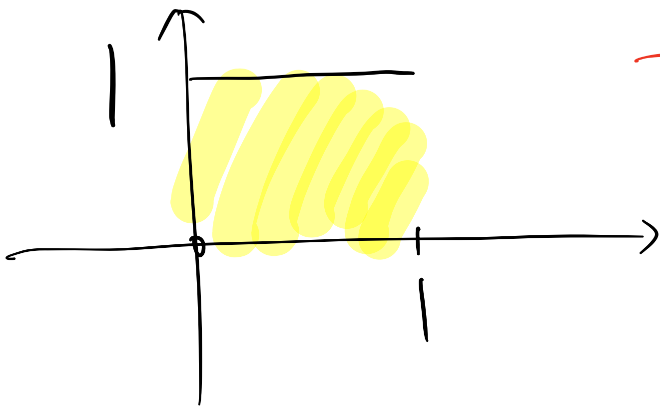
C.

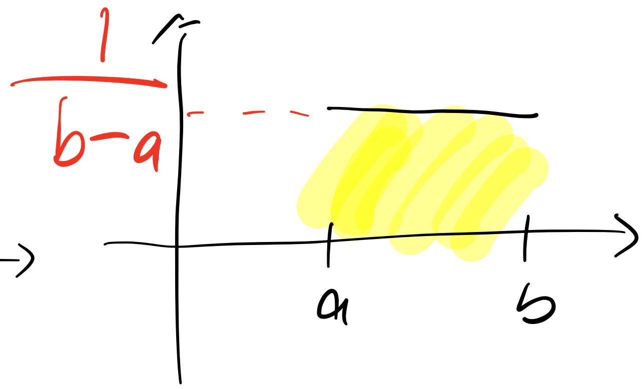
$$P(X \in A) = \int_A f(x)\,dx,$$

where $f(x)$ is called the ***probability density function*** *of r.v. X. (The abbreviation for probability density function is p.d.f.)*

$$P(a < X < b) = \int_a^b f(x)\,dx$$



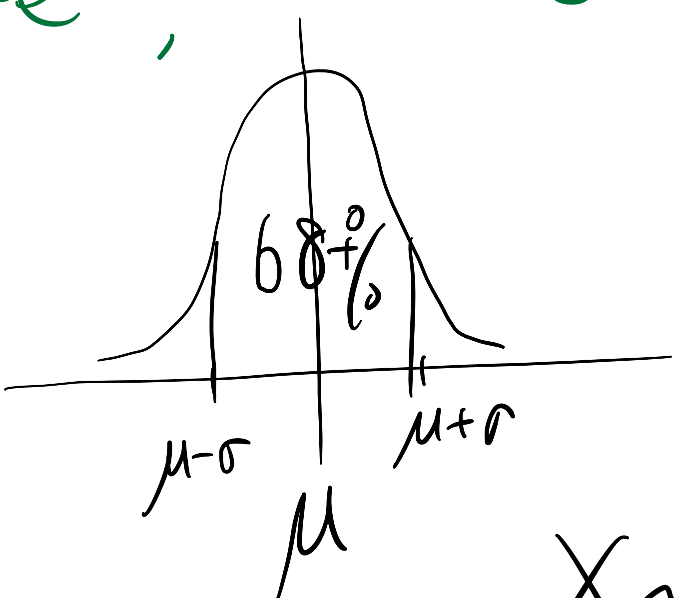




$X \sim U(a,b)$

$$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b \\ 0, & \text{otherwise}\end{cases}$$

$\sim e^{-x^2}$, $\sim x^{-3}$ $\sim e^{-x}$



Bell-shaped.

normal

$X \sim N(\mu, \sigma^2)$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$\sim x^{-3}$

power-law

幂律

Continuous Random Variable

If f is the p.d.f. of the r.v. X, then

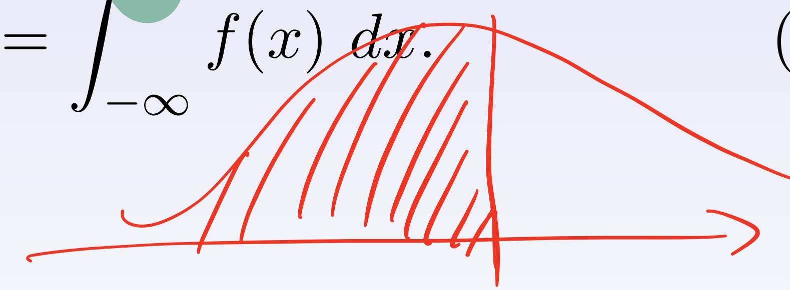
$f(x) \cdot \Delta x$

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)dx.$$

$\simeq P(X \in (x, x+\Delta x])$

The probability is the area, and the area is the integral from a to b.
If $A = (-\infty, x]$, then the distribution function F of X is given by

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(x)\, dx. \tag{2}$$



We also deduce that

$x \rightarrow$

$$\frac{d}{dx}F(x) = f(x) \tag{3}$$

for any x where $F(x)$ is differentiable.

Continuous Random Variable

Remark

(i) *The d.f.* $F(x)$ *of continuous r.v.* X *is continuous since* $P(X = x) = 0$.

(ii) *By using equation (2) and* $F(\infty) = 1$, *we get*

$$\int_{-\infty}^{\infty} f(x)\, dx = F(\infty) = 1. \tag{4}$$

To be a p.d.f. of some continuous random variable, the nonnegative function $f(x)$ *need satisfy equation (4).*

Continuous Random Variable

Remark

(iii) *Moreover, we may have that* $f(x) > 1$. *Actually, we may write that*

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x) = \lim_{\triangle x \to 0} \frac{F(x+\triangle x) - F(x)}{\triangle x}$$

$$= \lim_{\triangle x \to 0} \frac{P(x < X \leqslant x + \triangle x)}{\triangle x}.$$

Thus, $f(x)\triangle x$ *is approximately equal to the probability that* X *takes on a value in an interval of length* $\triangle x$ *about* x. *A probability density acts like an ordinary density (like grams per centimeter). Even though there is no mass exactly at a particular point, the total mass can be found by integrating.*

Continuous Random Variable

x^n, e^{-ax}

Example

Suppose that $f(x) = a(1-x)$ $(0 \leqslant x \leqslant 1)$ is the probability density function of random variable X.
(a) Find the value a.
(b) Find the corresponding d.f. $F(x)$.
(c) Calculate the value of $P(0.3 \leqslant X \leqslant 0.6)$. $= F(0.6) - F(0.3)$

Solution. (a) Since

$$\int_0^1 f(x)dx = \int_0^1 (a - ax)dx = \left(ax - \frac{a}{2}x^2\right)\Big|_0^1 = 1.$$

we have $a = 2$.

Continuous Random Variable

Solution. (b) The d.f. is if $x < 0$, then $F(x) = 0$.
If $x > 1$, then $F(x) = 1$.
If $0 \leqslant x \leqslant 1$, then
$F(x) = \int_0^x f(y)dy = \int_0^x (2-2y)dy = (2y - y^2)\Big|_0^x = 2x - x^2.$
(c) With the p.d.f. and d.f. above, we can find the probability that x is between 0.3 and 0.6 either by integrating the p.d.f. or by evaluating the d.f. at the endpoints and subtracting. In the first case, we find

$$P(0.3 \leqslant X \leqslant 0.6) = \int_{0.3}^{0.6} (2-2x)dx = (2x - x^2)\Big|_{0.3}^{0.6} = 0.33.$$

Alternatively, $P(0.3 \leqslant X \leqslant 0.6) = F(0.6) - F(0.3) = 0.33.$ $\square$

$F(b) - F(a) = P(X \in [a, b])$ (C.)

Continuous Random Variable

$F(x) = F(x+)$

~~Continuous~~

Example

Suppose that the d.f. of random variable X is

$$F(x) = \begin{cases} A + Be^{-x^2/2} & \text{for } x > 0, \\ 0 & \text{for } x \leqslant 0. \end{cases}$$

Calculate (a) the values of A and B; (b) $f(x)$; (c) $P(1 < X < 2)$.

Solution. (a) Since $F(+\infty) = 1$, $A = 1$. Since $F(0) = F(0^+)$, $A + B = 0$. Thus we can obtain that $B = -1$.

Continuous Random Variable

$$\left(1 - e^{-\frac{x^2}{2}}\right)'$$

Solution. (a) Since $F(+\infty) = 1$, $A = 1$. Since $F(0) = F(0^+)$, $A + B = 0$. Thus we can obtain that $B = -1$.
(b) By using equation (3), we have

$$f(x) = F'(x) = \begin{cases} xe^{-x^2/2} & \text{for } x > 0, \\ 0 & \text{for } x \leqslant 0. \end{cases}$$

(c) $P(1 < X < 2) = F(2) - F(1) = 0.47.$
Alternatively, $P(1 < X < 2) = \int_1^2 xe^{-x^2/2} dx = 0.47.$ $\square$

The Definition and Properties of Distribution Function

For one random variable X, the two things we must pay attention to are **the range** and **the rule** of X. The range of X means all the possible values that X could take. The rule of X means the d.f. of X. If the two things are known, then the probabilities that $\{X \in A\}$, for any $A \subseteq \mathbb{R}$ can be calculated.

Contents

The Distribution Function of Function of a Random Variable

Because a random variable is a real-valued function and the composition of two functions is also a function, we can assert that if X is a random variable, then $Y := g(X) = g(X(\omega))$, where g is a real-valued function defined on the real line, is a random variable as well. In this section, we show how to obtain the probability function or the density function of Y.

Let us see discrete case first. Here we mainly discuss the case X and $Y = g(X)$ are both discrete.

The Distribution Function of Function of a Random Variable

Example

Let X be a discrete random variable whose probability function is given in Table

x	-1	0	1
$p_X(x)$	$1/4$	$1/4$	$1/2$

Let $Y = 2X$ and $W = X^2$. Find the probability functions of Y and W respectively.

The Distribution Function of Function of a Random Variable

Solution. Because the function $g : x \to 2x$ is bijective (i.e., to a given $y = g(x) = 2x$, there corresponds one and only one x and vice versa), the number of possible values of the random variable Y will be the same as the number of possible values of X. We find that

$$\begin{aligned} p_Y(-2) &= P(Y = -2) = P(X = -1) = 1/4, \\ p_Y(0) &= P(Y = 0) = P(X = 0) = 1/4, \\ p_Y(2) &= P(Y = 2) = P(X = 1) = 1/2. \end{aligned}$$

We can describe the p.f. of Y by

x	-1	0	1
y	-2	0	2
$p_Y(y)$	$1/4$	$1/4$	$1/2$

The Distribution Function of Function of a Random Variable

Solution. For $W = X^2$, the possible values of the random variable of W is 0 and 1. And

$$p_W(0) = P(W = 0) = P(X = 0) = 1/4,$$

$$\begin{aligned} p_W(1) &= P(W = 1) = P(\{X = -1\} \cup \{X = 1\}) \\ &= P(X = -1) + P(X = 1) = 3/4. \end{aligned}$$

The p.f. of Y can be described by Table

w	1	0	1
x	-1	0	1
$p_Y(y)$	$1/4$	$1/4$	$1/2$

$\Longrightarrow$

w	0	1
$p_W(x)$	$1/4$	$3/4$

The Distribution Function of Function of a Random Variable

In the case when the random variable X can take a (countably) infinite number of values, we apply the transformation to an arbitrary value of X, and we can find a general formula for the function $p_Y(y)$. That is,

$$p_Y(y) = P(Y = y) = P(g(X) = y) = \sum_{x:g(x)=y} p(x). \tag{5}$$

The Distribution Function of Function of a Random Variable

Example

Suppose that the p.f. of X is

$$p_X(x) = (1-p)^{x-1}p, \qquad x = 1, 2, \cdots$$

where $0 < p < 1$. Let $Y = X^2$. Because the value of X is $1, 2, \cdots$, the quadratic function is (here) a bijective transformation and we easily calculate

$$p_Y(y) = P(Y = y) = P(X^2 = y)$$

$$= p_X(\sqrt{y}) = (1-p)^{\sqrt{y}-1}p \quad \text{for } y = 1, 4, 9, \cdots .$$

$\square$

The Distribution Function of Function of a Random Variable

Next, let us consider continuous case. Here we assume that X and $Y = g(X)$ are both continuous random variables.

Example

Let X be a continuous nonnegative random variable with density function

$$f_X(x) = \begin{cases} \dfrac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise.} \end{cases}$$

and let $Y = 2X + 1$. Find $f_Y(y)$, the probability density function of Y.

The Distribution Function of Function of a Random Variable

Solution. First we find the d.f. of the random variable Y.

$$\begin{aligned} F_Y(y) =& P(Y \leqslant y) = P(2X+1 \leqslant y) \\ =& P(X \leqslant \frac{y-1}{2}) = \int_{-\infty}^{\frac{y-1}{2}} f_X(x)dx. \end{aligned}$$

If $y \leqslant 1$, $\dfrac{y-1}{2} \leqslant 0$, then $F_Y(y) = 0$.
If $1 < y < 9$, $0 \leqslant \dfrac{y-1}{2} \leqslant 4$, then

$$F_Y(y) = \int_0^{\frac{y-1}{2}} \frac{x}{8}dx = \frac{(y-1)^2}{64}.$$

The Distribution Function of Function of a Random Variable

Solution. If $y \geqslant 9$, then $F_Y(y) = 1$. That is,

$$F_Y(y) = \begin{cases} 0 & \text{for } y \leqslant 1, \\ \dfrac{(y-1)^2}{64} & \text{for } 1 < y < 9, \\ 1 & \text{for } y \geqslant 9. \end{cases}$$

Hence, by differentiating, we obtain

$$f_Y(y) = \begin{cases} \dfrac{y-1}{32} & \text{for } 1 < y < 9, \\ 0 & \text{otherwise.} \end{cases}$$

$\square$

The Distribution Function of Function of a Random Variable

Example

Suppose that X is a continuous random variable with probability density $f_X(x)$. Find the probability density function $f_Y(y)$ of $Y = X^2$.

Solution. For $y < 0$, $F_Y(y) = 0$. So $f_Y(y) = 0$. For $y \geqslant 0$,

$$F_Y(y) = P(Y \leqslant y) = P(X^2 \leqslant y)$$

$$= P(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Differentiation yields

$$f_Y(y) = \frac{1}{2\sqrt{y}}\Big[f_X(\sqrt{y}) + f_X(-\sqrt{y})\Big].$$

□

The Distribution Function of Function of a Random Variable

The method employed above, **the distribution function method**, is the usual way to obtain p.d.f. of a function of a continuous random variable. This method can be divided into four steps:

(i) transform the event $\{Y \leqslant y\}$ to $\{g(X) \leqslant y\}$,

(ii) transform the event $\{g(X) \leqslant y\}$ to $\{X \in I_y\}$, where $I_y = \{x : g(x) \leqslant y\}$,

(iii) calculate the probability $P(X \in I_y)$, which is just the d.f. $F(y) = P(Y \leqslant y))$ of Y,

(iv) obtain p.d.f. $f_Y(y)$ of Y by $f_Y(y) = F_Y^{'}(y)$.

The Distribution Function of Function of a Random Variable

Theorem

Let X be a continuous random variable having p.d.f. f_X. Suppose that $g(x)$ is a strictly monotonic (increasing or decreasing), differentiable (and thus continuous) function of x. Then the random variable Y defined by $Y = g(X)$ has a p.d.f. given by

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\left|\dfrac{d}{dy}g^{-1}(y)\right| & \text{if } y = g(x) \text{ for some } x, \\ 0 & \text{if } y \neq g(x) \text{ for all } x, \end{cases} \tag{6}$$

where $x = g^{-1}(y)$ is defined to equal that value of x such that $g(x) = y$.

The Distribution Function of Function of a Random Variable

Proof. Suppose that $g(x)$ is an increasing function. Suppose that $y = g(x)$ for some x. Then, with $Y = g(X)$,

$$F_Y(y) = P(g(X) \leqslant y) = P(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y)).$$

Differentiation gives

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y),$$

since $g^{-1}(y)$ is nondecreasing, so its derivative is nonnegative.

When $y \neq g(x)$ for any x, then $F_Y(y)$ is either 0 or 1, and in either case $f_Y(y) = 0$. $\square$

The Distribution Function of Function of a Random Variable

Example

Let X be a continuous nonnegative random variable with density function

$$f_X(x) = \begin{cases} \dfrac{x}{8} & \text{for } 0 < x < 4, \\ 0 & \text{otherwise.} \end{cases}$$

and let $Y = 2X + 1$. Find $f_Y(y)$ by using Theorem.

Solution. If $y = g(x) = 2x + 1$, then

$$g^{-1}(y) = \frac{y-1}{2} \quad \text{and} \quad \frac{d}{dy} g^{-1}(y) = \frac{1}{2}.$$

Hence,

$$f_Y(y) = \begin{cases} \frac{1}{2} \cdot \frac{\frac{y-1}{2}}{8} & \text{for } 0 < \frac{y-1}{2} < 4 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{y-1}{32} & \text{for } 1 < y < 9, \\ 0 & \text{otherwise.} \end{cases}$$

$\square$

The Distribution Function of Function of a Random Variable

Example

Let X be a continuous nonnegative random variable with density function f_X, and let $Y = X^n$. Find f_Y, the probability density function of Y.

Solution. If $g(x) = x^n$, then

$$g^{-1}(y) = y^{1/n}$$

and

$$\frac{d}{dy} g^{-1}(y) = \frac{1}{n} y^{1/n-1}.$$

Hence,

$$f_Y(y) = \frac{1}{n} y^{1/n-1} f_X(y^{1/n}).$$

For $n = 2$, this gives $f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y})$. $\square$

The end

Thank you for your patience !