Lecture note 5: Continuous Random variables

Foundation of Probability Theory/STA 203

Zhuosong ZHANG

Department of Statistics and Data Science, SUSTech

Fall, 2023

Continuous random variables

Introduction

There are random variables other than discrete random variables, that is, it take continuous values:

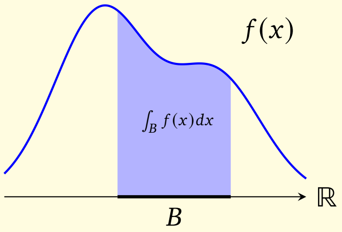
- the waiting time for the next bus;
- the height of a randomly selected SUSTech student;
- the delay time of a flight;
- and so on...

Definition 1

We say that X is a continuous random variable if there exists a nonnegative function f, defined for all real $x \in \mathbb{R}$, having the property

$$\mathbb{P}\{X \in B\} = \int_B f(x)dx$$

for any $B \in \mathscr{B}(\mathbb{R})$.



Remark

- This definition is also known as “absolutely continuous random variables”.
- The support of X is defined as $\mathbb{S} : \{x : f(x) > 0\}$.

Probability density function

Definition 2 (pdf)

The function f defined as in the last page is called the probability density function (pdf) of the random variable X.

Proposition 3

Any pdf f satisfies the following properties:

(i) $f(x) \geqslant 0$ for all $x \in \mathbb{R}$.

(ii) $\int_{-\infty}^{\infty} f(x)dx = \int_{\mathbb{S}} f(x)dx = 1$.

Example 4

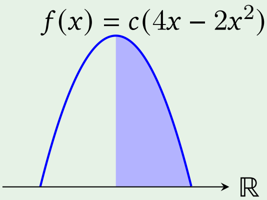
Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} c(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) What is the value of c?

(b) Find $\mathbb{P}\{X > 1\}$.

(c) Find the distribution function of X.



Solution.

(a) Since f is a pdf, we have $\int f(x)dx = 1$, implying that

$$c\int_0^2 (4x - 2x^2)dx = 1,$$

which further gives $c = \frac{3}{8}$.

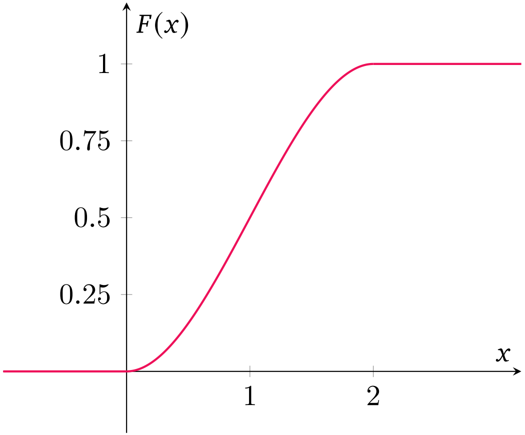
(b) We have

$$\mathbb{P}\{X > 1\} = \int_1^\infty f(x)dx = \frac{3}{8}\int_1^2 (4x - 2x^2)dx = \frac{1}{2}.$$

(c) For $x \leqslant 0$, $F(x) = \mathbb{P}(X \leqslant x) = 0$. For $x \in (0, 2)$,

$$F(x) = \mathbb{P}(X \leqslant x) = \int_0^x f(x) = \int_0^x \frac{3}{8}(4t - 2t^2)dt = \frac{3}{4}x^2 - \frac{1}{4}x^3.$$

For $x \geqslant 2$, $F(x) = \mathbb{P}(X \leqslant x) = 1$. ■


F(x)
1
0.75
0.5
0.25
x
1
2

Example 5

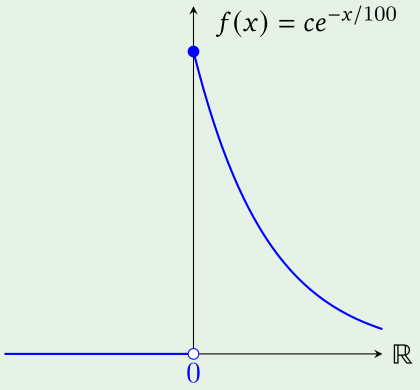
The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} ce^{-x/100} & x \geqslant 0 \\ 0 & x < 0. \end{cases}$$

What is the probability that

(a) a computer will function between 50 and 150 hours before breaking down?

(b) it will function for fewer than 100 hours?



Solution.

(a) Since

$$1 = \int_{-\infty}^{\infty} f(x)dx = c\int_{-\infty}^{\infty} e^{-x/100}dx,$$

we obtain

$$1 = 100c \implies c = \frac{1}{100}.$$

Let X be the random variable representing the function time (in hours) of a computer. Hence, the probability that a computer will function between 50 and 150 hours before breaking down is given by

$$\begin{aligned}\mathbb{P}\{50 < X < 150\} &= \int_{50}^{150} f(x)dx \\ &= \frac{1}{100}\int_{50}^{150} e^{-x/100}dx \\ &= e^{-1/2} - e^{-3/2} \approx 0.384.\end{aligned}$$

(b) Similarly,

$$\begin{aligned}\mathbb{P}\{X < 100\} &= \int_{0}^{100} f(x)dx \\ &= \int_{0}^{100} \frac{1}{100}e^{-x/100}dx \\ &= 1 - e^{-1} \approx 0.633.\end{aligned}$$

■

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Proposition 6

Let $F(x) = \mathbb{P}(X \leqslant x)$ be the distribution function of X with support (a, b). Here, a may be $-\infty$ and b may be ∞.

(i) For any $x < y$,

$$\mathbb{P}\{x < X \leqslant y\} = F(y) - F(x) = \int_x^y f(u)du.$$

(ii) F is continuous on (a, b).

(iii) For any $x \in (a, b)$, $\mathbb{P}\{X = x\} = 0$.

Proof.

(i) The first equality follows directly from the definition of cdf. For the second one,

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

(ii) For any $x \in (a, b)$ and $\delta > 0$,

$$F(x + \delta) - F(x - \delta) = \int_{x-\delta}^{x+\delta} f(u)du.$$

Because f is integrable, then it is bounded[a], and thus, F is continuous.

(iii) For any $x \in (a, b)$,

$$\mathbb{P}(X = x) = F(x) - F(x-) = 0.$$

∎

[a] We have assumed f is Riemann integrable.

Distribution function

The relationship between the cumulative distribution function F and the probability density function f is expressed by

$$F(x) = \mathbb{P}\{X \leqslant x\} = \int_{-\infty}^{x} f(t)dt.$$

Therefore, if f is continuous,

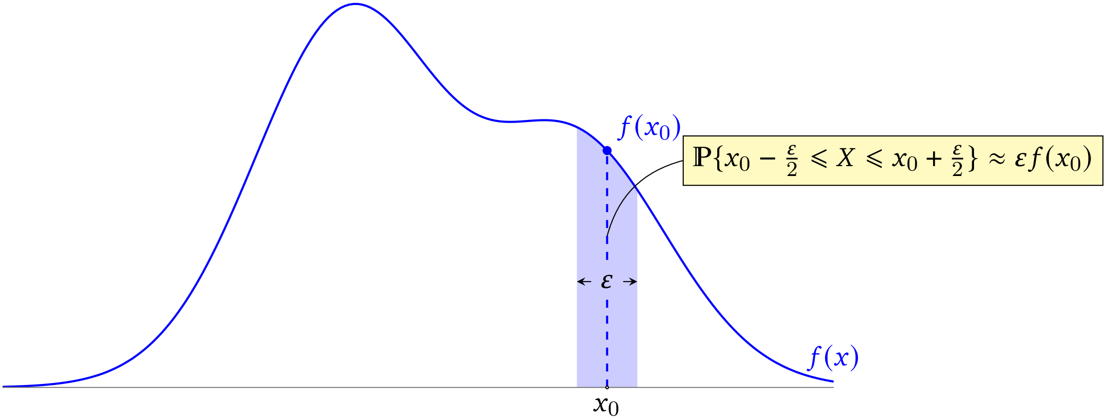
$$\frac{d}{dx}F(x) = f(x).$$

Moreover,

$$\mathbb{P}(x < X \leqslant y) = P(x \leqslant X \leqslant y) = \ldots = F(y) - F(x).$$

A more intuitive interpretation

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

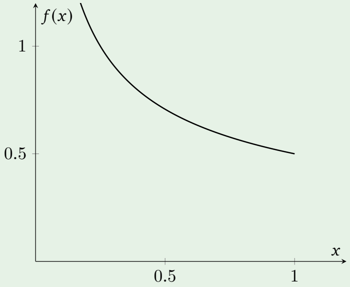


A pdf can take arbitrarily large value

Example 7

Consider a random variable X with pdf

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{if } 0 < x \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$



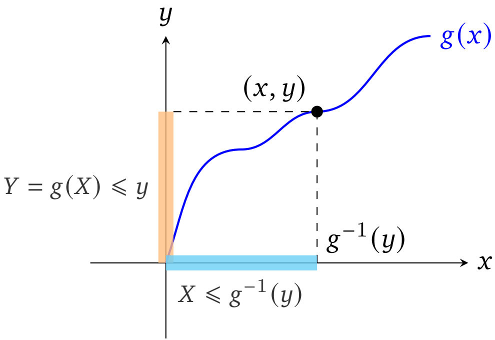
Functions of a random variable

- Let X be a continuous random variable with support $\mathbb{S}$.
- Let $g : \mathbb{S} \to \mathbb{R}$ be a function.
- Suppose that we want to know the distribution of $g(X)$.
- The distribution function of $Y = g(X)$ is given by

$$\begin{aligned} F_Y(y) &= \mathbb{P}\{Y \leqslant y\} \\ &= \mathbb{P}\{g(X) \leqslant y\} \\ &= \mathbb{P}\{X \in g^{-1}(-\infty, y]\}. \end{aligned}$$

- If g is increasing, then

$$\{X \in g^{-1}(-\infty, y]\} = \{X \leqslant g^{-1}(y)\}.$$



Example 8

If X is continuous with distribution function F_X and density function f_X, find the density function of $Y = 2X$.

Solution.

We first calculate the distribution function F_Y of Y:

$$F_Y(y) = \mathbb{P}\{Y \leqslant y\} = \mathbb{P}\{2X \leqslant y\} = \mathbb{P}\{X \leqslant \frac{y}{2}\} = F_X(\frac{y}{2}).$$

Therefore,

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{2}F'_X(\frac{y}{2}) = \frac{1}{2}f_X(\frac{y}{2}).$$

■

Example 9

If X is a continuous random variable with probability density f_X, then the distribution of $Y = X^2$ is obtained as follows:

$$
\begin{aligned}
F_Y(y) &= \mathbb{P}\{Y \leqslant y\} \\
&= \mathbb{P}\{X^2 \leqslant y\} \\
&= \mathbb{P}\{-\sqrt{y} \leqslant X \leqslant \sqrt{y}\} \\
&= F_X(\sqrt{y}) - F_X(-\sqrt{y}),
\end{aligned}
$$

and thus

$$
f_Y(y) = F_Y'(y) = \frac{1}{2\sqrt{y}}\{f_Y(\sqrt{y}) - f_Y(-\sqrt{y})\}.
$$

Expectation of continuous random variables

Definition 10

Let X be a continuous random variable with probability density f. Then, the expectation of X is defined as

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$$

provided that $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$.

Remark

If X is supported on $\mathbb{S}$, that is, $f(x) = 0$ when $x \notin \mathbb{S}$, then

$$\mathbb{E}[X] = \int_{\mathbb{S}} x f(x) dx.$$

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Example 11

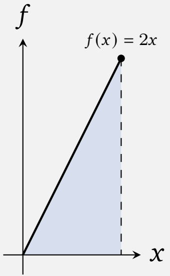
Find $\mathbb{E}[X]$ when the density function of X is

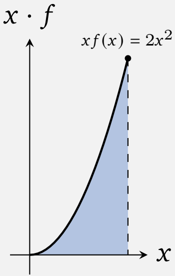
$$f(x) = \begin{cases} 2x & 0 \leqslant x \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

Solution.

Note that the support of X is $\mathbb{S} = [0, 1]$. Therefore,

$$\mathbb{E}[X] = \int_0^1 x f(x) dx = \int_0^1 2x^2 dx$$
$$= \frac{2}{3}.$$





Expectation of $g(X)$

Proposition 12

If X is a continuous random variable with probability density function $f(x)$ on a support $\mathbb{S}$, then, for any function $g : \mathbb{S} \to \mathbb{R}$,

$$\mathbb{E}[g(X)] = \int_{\mathbb{S}} g(x) f(x) dx.$$

Proof of the preceding Proposition

We need to prove the following lemma:

Lemma 13

For a nonnegative continuous random variable Y,

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}\{Y > y\} dy.$$

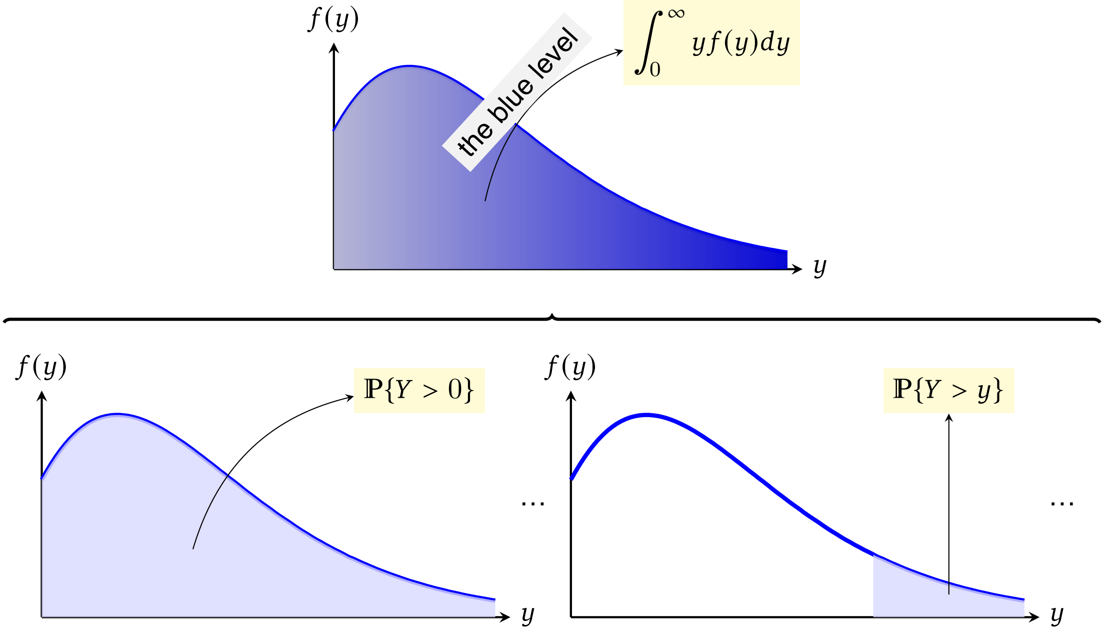
Proof of the preceding Proposition

Proof.

Note that $Y = \int_0^Y dy = \int_0^\infty \mathbb{1}(y < Y)dy$, and taking expectations on both sides yields

$$
\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}\left[\int_0^\infty \mathbb{1}(y < Y)dy\right] \\
&= \int_0^\infty \mathbb{E}[\mathbb{1}(Y > y)]dy && \text{exchange the order of } \mathbb{E} \text{ and } \int \\
&= \int_0^\infty \mathbb{P}\{Y > y\}dy. && (\text{because } \mathbb{P}(A) = \mathbb{E}[\mathbb{1}(A)])
\end{aligned}
$$

∎



$f(y)$
the blue level
$\int_0^\infty y f(y) dy$
y
$f(y)$
$\mathbb{P}\{Y > 0\}$
y
...
$f(y)$
$\mathbb{P}\{Y > y\}$
y
...

Proof of Proposition 12.

We only prove for the case where $g \geqslant 0$. By the above lemma with $Y = g(X)$,

$$\begin{aligned}
\mathbb{E}[g(X)] &= \int_0^\infty \mathbb{P}\{g(X) > y\}dy \\
&= \int_0^\infty \left(\int_{x \in \mathbb{S}: g(x) > y} f(x)dx \right) dy \\
&= \int_{\mathbb{S}} \left(\int_{y: 0 < y < g(x)} 1 dy \right) f(x)dx \\
&= \int_{\mathbb{S}} g(x) f(x) dx.
\end{aligned}$$

■

Example 14

The density function of X is given by

$$f(x) = \begin{cases} 1 & \text{if } 0 \leqslant x \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find $\mathbb{E}[e^X]$.

Solution.

$$\mathbb{E}[e^X] = \int_0^1 e^x dx = e - 1.$$

Properties of expectation

Proposition 15

Let X be a continuous random variable supported on $\mathbb{S}$. For any $a, b \in \mathbb{R}$ and $g, h : \mathbb{S} \to \mathbb{R}$,

$$\mathbb{E}[ag(X) + bh(X)] = a\,\mathbb{E}[g(X)] + b\,\mathbb{E}[h(X)].$$

Linearity property of expectation

Proposition 16

More generally, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and let X and Y be two random variables. Then, for any a, b,

$$\mathbb{E}[aX + bY] = a\,\mathbb{E}[X] + b\,\mathbb{E}[Y].$$

Variance

Definition 17

The variance of a continuous random variable is defined exactly as it is for a discrete random variable. If X is a random variable with expected value μ, then the variance of X is defined as

$$\mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Proposition 18

For any $a, b \in \mathbb{R}$,

$$\mathrm{Var}(aX+b) = a^2 \,\mathrm{Var}(X).$$

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 19

Suppose that the density function of X is

$$f(x) = \begin{cases} 2x & 0 \leqslant x \leqslant 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $\mathrm{Var}(X)$.

Solution.

We first compute $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 2x^3 dx = \frac{1}{2}.$$

Hence,

$$\mathrm{Var}(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}.$$

■

Commonly used continuous random variables

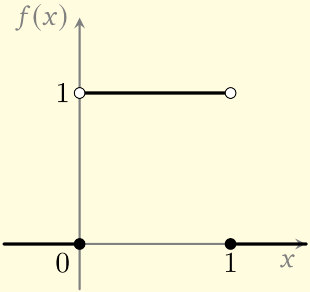
Uniform random variables

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 20 (Uniform distribution)

A random variable X is said to be uniformly distributed over the interval $(0, 1)$ if its probability density function is given by

$$f(x) = \begin{cases} 1 & 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$



Note that for any $0 < a < b < 1$,

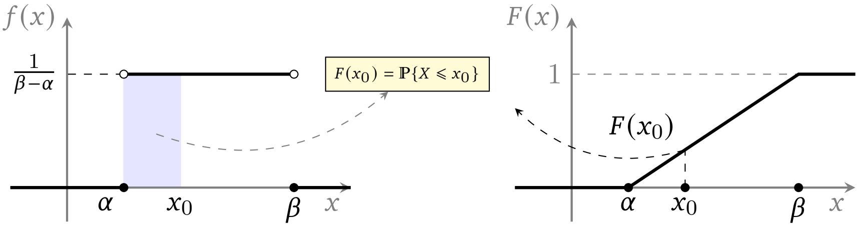
$$\mathbb{P}\{a \leqslant X \leqslant b\} = \int_a^b f(x)dx = b - a,$$

which is the length of the interval $[a, b]$.

Definition 21 (General uniform distribution)

We say X is a uniform random variable on the interval (α, β), denoted by $X \sim$ Uniform(α, β), if the pdf of X is given by

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



Proposition 22

Let $X \sim \text{Uniform}(\alpha, \beta)$. Then

$$\mathbb{E}[X] = \frac{\alpha+\beta}{2}, \quad \text{Var}(X) = \frac{(\beta-\alpha)^2}{12}.$$

Proof.

Note that

$$\mathbb{E}[X] = \int_\alpha^\beta \frac{x}{\beta-\alpha}dx = \frac{\alpha+\beta}{2},$$

and

$$\mathbb{E}[X^2] = \int_\alpha^\beta \frac{x^2}{\beta-\alpha}dx = \frac{\alpha^2+\alpha\beta+\beta^2}{3},$$

and it follows that

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{(\beta-\alpha)^2}{12},$$

as desired. ∎

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Example 23

If $X \sim \text{Uniform}(0, 10)$, find the probability that

(a) $X < 3$,

(b) $X > 6$,

(c) $3 < X < 8$.

Solution.

(a) $\mathbb{P}\{X < 3\} = \int_{-\infty}^{3} f(x)dx = \frac{1}{10}\int_{0}^{3} 1dx = 0.3.$

(b) $\mathbb{P}\{X > 6\} = \int_{6}^{\infty} f(x)dx = \frac{1}{10}\int_{6}^{10} 1dx = 0.4.$

(c) $\mathbb{P}\{3 < X < 8\} = \int_{3}^{8} f(x)dx = \frac{1}{10}\int_{3}^{8} 1dx = 0.5.$ ∎

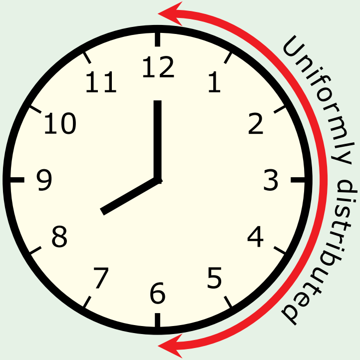
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Example 24

Buses arrive at a specified stop at 15-minute intervals starting at 8 A.M. That is, they arrive at 8, 8:15, 8:30, 8:45, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 8 and 8:30, find the probability that he waits

(a) less than 5 minutes for a bus;

(b) more than 10 minutes for a bus.



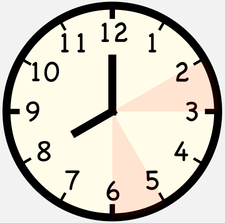
Solution.

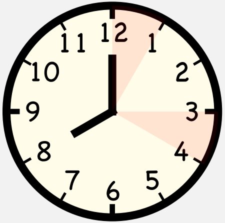
Let X denote the number of minutes past 8 that the passenger arrives at the stop. Then, $X \sim \text{Uniform}(0, 30)$.
The passenger will have to wait less than 5 minutes if he arrives between $8:10$ and $8:15$ or between $8:25$ and $8:30$. Hence, the desired probability for part (a) is

$$\mathbb{P}\{10 < X < 15\} + \mathbb{P}\{25 < X < 30\} = \frac{1}{3}.$$

Similarly, the probability for part (b) is

$$\mathbb{P}\{0 < X < 5\} + \mathbb{P}\{15 < X < 20\} = \frac{1}{3}. \quad \blacksquare$$





Definition 25 (Normal distribution)

A random variable X is said to have a normal distribution, or a Gaussian distribution, with parameters μ and σ^2, denoted by

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

if the pdf of X is given by

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

Specially, if $X \sim N(0,1)$, then X is said to be a standard normal random variable.

Figure: Gauss (高斯, 1777–1855)

The normal curve

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

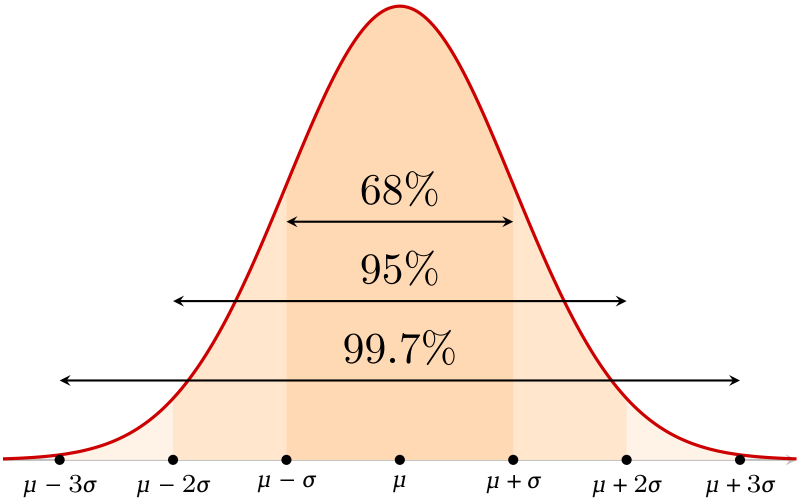
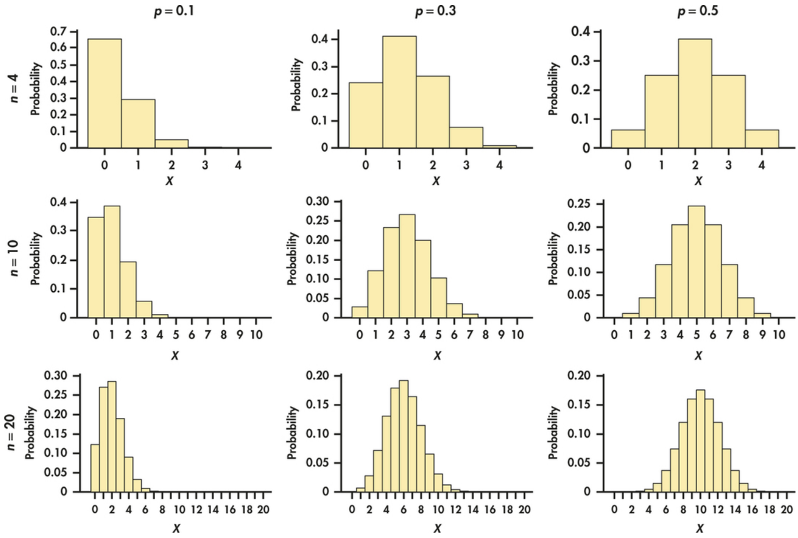


Figure: The normal curve

History and examples



History and examples

Examples include:

- the height of a man: In the UK, the mean male height is 1.778 m, and the standard deviation 0.076 m;
- the error made in measuring a physical quantity: length of a piece of string using a ruler;
- exam scores: the SAT Reasoning Test has a distribution that is roughly unimodal and symmetric and is designed to have an overall mean of about 500 and a standard deviation of 100 for all test takers;
- ...

History and examples

de Moivre

棣莫弗

Laplace

拉普拉斯

Gauss

高斯

1733	1782	1809	→ Time

- Abraham de Moivre (1733) used it to approximate probabilities associated with binomial random variables when the binomial parameter n is large.
- Laplace calculated the normalizing constant $\sqrt{2\pi}\sigma$ in 1782, and proved the CLT in 1810.
- Gauss discovered the normal distribution in 1809 as a way to rationalize the method of least squares.

The proof of the normalizing constant

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Proposition 26

We have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}dx = 1.$$

Remark

More generally, for $N(\mu, \sigma^2)$, we have

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-y^2/2}dy = 1$$

if we take $y = \frac{x-\mu}{\sigma}$.

Proof.

Let $I = \int_{-\infty}^{\infty} e^{-x^2/2} dx$. Then,

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2/2} dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2/2} dy\right)$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)} dy dx.$$

Now, let $x = r\cos\theta, y = r\sin\theta$, and it follows that $dydx = rd\theta dr$. Therefore,

$$I^2 = \int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} r d\theta dr$$
$$= 2\pi \int_0^{\infty} r e^{-r^2/2} dr$$
$$= -2\pi e^{-r^2/2}\Big|_0^{\infty}$$
$$= 2\pi.$$

■

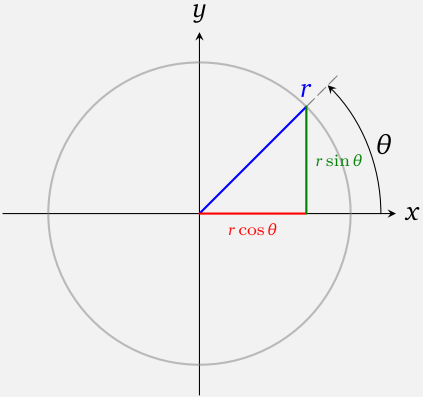


Figure: Change to Polar Coodinates in a Double Integral

Linear transformation of X

Proposition 27

If $X \sim N(\mu, \sigma^2)$, then for any $a, b \in \mathbb{R}$,

$$aX + b \sim N(a\mu + b, a^2\sigma^2).$$

Remark

In particular,

$$\frac{X - \mu}{\sigma} \sim N(0, 1).$$

If $a = 0$, then $aX + b = b$ is a constant. We can consider a constant as a normal random variable with mean b and variance 0.

Linear transformation of X

Proof.

Without loss of generality, we assume that $a > 0$. For $a < 0$, the proof is similar. Let $Y = aX + b$. Note that

$$
\begin{aligned}
F_Y(x) &= \mathbb{P}\{Y \leqslant x\} \\
&= \mathbb{P}\{aX + b \leqslant x\} \\
&= \mathbb{P}\{X \leqslant \frac{x-b}{a}\} \\
&= F_X(\frac{x-b}{a}),
\end{aligned}
$$

where F_X is the cumulative distribution function of X. By differentiation, the pdf of Y is given by

$$
\begin{aligned}
f_Y(x) &= \frac{1}{a} f_X(\frac{x-b}{a}) \\
&= \frac{1}{\sqrt{2\pi}a\sigma} \exp\left\{-\frac{-(x-b-a\mu)^2}{2(a\sigma)^2}\right\},
\end{aligned}
$$

which shows that

$$Y \sim N(a\mu + b, a^2\sigma^2). \quad \blacksquare$$

Expectation and variance of $N(\mu, \sigma^2)$

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Proposition 28

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

$$\mathbb{E}[X] = \mu, \quad \mathrm{Var}(X) = \sigma^2.$$

Proof of the expectation.

Let $Z = \frac{X-\mu}{\sigma}$. It suffices to prove $\mathbb{E}[Z] = 0$ and $\mathrm{Var}(Z) = 1$. To this end, note that $Z \sim N(0,1)$, we have

$$\begin{aligned}\mathbb{E}[Z] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} dx \\ &= -\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big|_{-\infty}^{\infty} = 0.\end{aligned}$$

■

Proof of the variance.

For the variance,

$$
\begin{aligned}
\mathbb{E}[Z^2] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx \\
&= -\frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} x d e^{-x^2/2} \right) && \text{(integration by parts)} \\
&= -\frac{1}{\sqrt{2\pi}} \left(x e^{-x^2/2} \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} e^{-x^2/2} dx \right) \\
&= 1.
\end{aligned}
$$

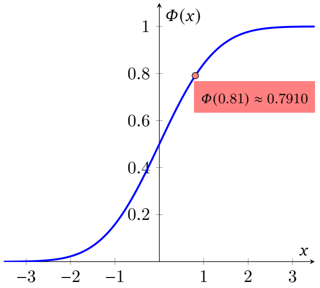
■

The function $\Phi(x)$

- It is customary to denote the cumulative distribution function of $N(0,1)$ by $\Phi(x)$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

z	.00	.01	.02	.03	.04	.05
.0	.5000	.5040	.5080	.5120	.5160	.5199
.1	.5398	.5438	.5478	.5517	.5557	.5596
.2	.5793	.5832	.5874	.5910	.5948	.5987
.3	.6179	.6217	.6255	.6293	.6331	.6368
.4	.6554	.6591	.6628	.6664	.6700	.6736
.5	.6915	.6950	.6985	.7019	.7054	.7088
.6	.7257	.7291	.7324	.7357	.7389	.7422
.7	.7580	.7611	.7642	.7673	.7704	.7734
.8	.7881	**.7910**	.7939	.7967	.7995	.8023
.9	.8159	.8186	.8212	.8238	.8264	.8289



CDF of $N(0,1)$

- For negative x, $\Phi(x)$ can be obtained from the following relationship:

$$\Phi(-x) = 1 - \Phi(x), \quad -\infty < x < \infty.$$

- For example,

$$\Phi(0.21) = 0.5832, \quad \Phi(-0.21) = 1 - \Phi(0.21) = 0.4168.$$

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

$$\mathbb{P}\{X \leqslant a\} = \mathbb{P}\left\{\frac{X-\mu}{\sigma} \leqslant \frac{a-\mu}{\sigma}\right\} = \Phi\left(\frac{a-\mu}{\sigma}\right).$$

R codes

- We can use R codes to calculate the probabilities: for example, if $Z \sim N(0,1)$, $\mathbb{P}\{Z \leqslant 0.81\}$ equals

```
pnorm(0.81)
[1] 0.7910299
```

- $\mathbb{P}\{Z > 0.81\}$ equals

```
pnorm(0.81, lower=FALSE)
[1] 0.2089701
```

- If $X \sim N(3, 2^2)$, then $\mathbb{P}\{X \leqslant 5\}$ equals

```
pnorm(5, mean=3, sd=2)
[1] 0.8413447
```

Examples

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 29

If $X \sim N(\mu = 3, \sigma^2 = 9)$, find

(a) $\mathbb{P}\{2 < X < 5\}$;

(b) $\mathbb{P}\{X > 0\}$;

(c) $\mathbb{P}\{|X - 3| > 6\}$.

Solution.

(a) Note that with $Z = (X - 3)/3 \sim N(0, 1)$,

$$\mathbb{P}\{2 < X < 5\} = \mathbb{P}\left\{\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right\} = \mathbb{P}\left\{-\frac{1}{3} < Z < \frac{2}{3}\right\}$$

$$= \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) \approx 0.3779.$$

■

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 30

An expert witness in a paternity suit testifies that the length (in days) of human gestation is approximately normally distributed with parameters $\mu = 270$ and $\sigma^2 = 100$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had the very long or very short gestation indicated by the testimony?

在一个亲子诉讼案件中，一名专家证人作证说，人类妊娠的长度（以天为单位）近似于正态分布，其参数为 $\mu = 270$，$\sigma^2 = 100$。被告能够证明他在孩子出生前的 *290* 天到 *240* 天期间一直在国外。如果被告实际上是孩子的父亲，那么母亲可能出现在证言中所述的非常长或非常短的妊娠期的概率是多少？

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 31 (Exponential distribution)

A continuous random variable X is said to follow an exponential distribution with parameter $\lambda > 0$, denoted by $X \sim \text{Exp}(\lambda)$, if its pdf is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0, \\ 0 & x < 0. \end{cases}$$

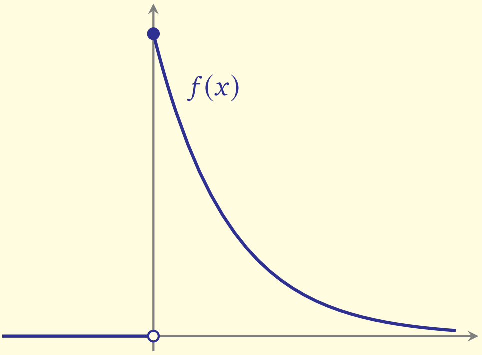


Figure: Exponential density

Proposition 32

The CDF of $X \sim \text{Exp}(\lambda)$ is

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geqslant 0, \\ 0 & x < 0. \end{cases}$$

Proof.

For any $x \geqslant 0$,

$$\begin{aligned} F(x) &= \mathbb{P}\{X \leqslant x\} \\ &= \int_0^x \lambda e^{-\lambda t} dt \\ &= 1 - e^{-\lambda x}. \end{aligned}$$

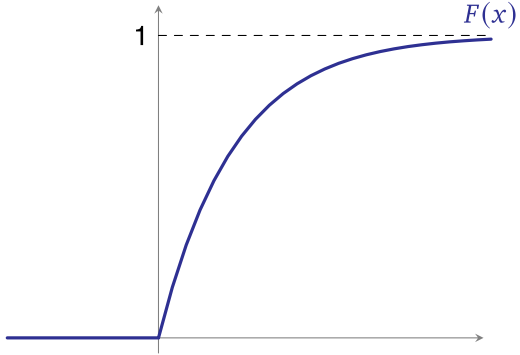


Figure: CDF of $\text{Exp}(\lambda)$

How to use Exp(λ)?

- In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs.
- The amount of time (starting from now) until an earth quake occurs.
- The waiting time until you receive a phone call.
- The time length between mutations on a DNA strand (基因串突变).
- How long it takes for a bank teller etc. to serve a customer.
- ...

Example 33

Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = 0.1$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait

(a) more than 10 minutes;

(b) between 10 and 20 minutes.

Solution.

Let X denote the length of the call made by the person in the booth. Then, $X \sim \text{Exp}(0.1)$.

(a) The probability that you have to wait more than 10 minutes is

$$\mathbb{P}\{X > 10\} = 1 - F(10) = e^{-(0.1)(10)} = e^{-1} \approx 0.368.$$

(b) The probability that you will wait between 10 and 20 minutes is

$$\begin{aligned}\mathbb{P}\{10 < X < 20\} &= F(20) - F(10)\\ &= e^{-(0.1)(10)} - e^{-(0.1)(20)}\\ &= e^{-1} - e^{-2}\\ &\approx 0.233.\end{aligned}$$

■

Expectation and variance

Proposition 34

Let $X \sim \text{Exp}(\lambda)$. Then

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

Remark

For example, in the bank customer example, the parameter λ can be understood as the average number of customers that a banker serves in a unit time interval, and thus the average waiting time is $1/\lambda$.

Proposition 35

If $X \sim \text{Exp}(\lambda)$, then it is memoryless, that is,

$$\mathbb{P}\{X > s + t | X > t\} = \mathbb{P}\{X > s\} \quad \text{for any } s, t > 0.$$

Proof.

For any $s, t > 0$, since $\{X > s + t\} \subset \{X > t\}$, it follows that

$$\mathbb{P}(\{X > s + t\} \cap \{X > t\}) = \mathbb{P}\{X > s + t\} = e^{-\lambda(s+t)} = \mathbb{P}\{X > s\}\,\mathbb{P}\{X > t\}.$$

Then, it follows from the definition of conditional probability that

$$\mathbb{P}\{X > s + t | X > t\} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbb{P}\{X > s\}. \qquad \blacksquare$$

Examples

Example 36

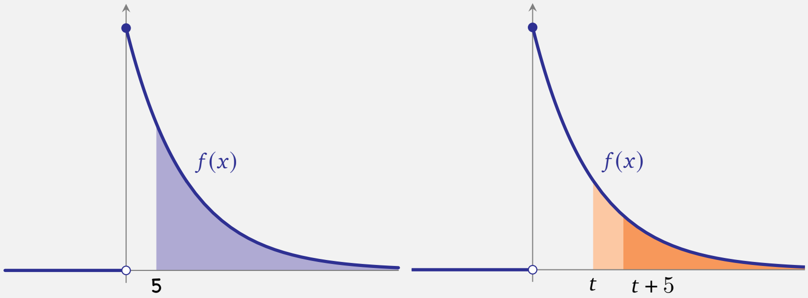
Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with an average value of 10,000 miles. If a person desires to take a 5000-mile trip, what is the probability that he or she will be able to complete the trip without having to replace the car battery?

Solution.

Let X be the lifetime (in 1000 miles) of the car, and it follows that $\mathbb{E}[X] = 10$ which further implies that $\lambda = 0.1$. Let x be the number of miles (in 1000 miles) that the battery had been in use prior to the start of the trip. The desired probability is

$$\mathbb{P}\{X > x + 5 | X > x\} = 1 - F(5) = e^{-(0.1)(5)} = e^{-0.5} \approx 0.604.$$

■



Further reading

[1] Sheldon M. Ross (谢尔登·M. 罗斯).

A first course in probability (概率论基础教程): Chapter 5.

10th edition (原书第十版), 机械工业出版社

[2] 李贤平.

概率论基础: Chapters 3 and 4.

第三版, 高等教育出版社