Chapter 4 Continuous Random Variables

One big difference that we notice here as opposed to discrete random variables is that the **distribution function** of a continuous random variable is does not have any jumps. Remember that DF of a discrete random variable is a step function. The jumps in the DF correspond to points x for which $P(X = x) > 0$. Thus, the fact that the DF does not have jumps is consistent with the fact that $P(X = x) = 0$ for all x. Indeed, we have the following definition for continuous random variables.

Definition 4.1

A random variable X with DF $F_X(x) = P(X \le x)$ is said to be **continuous** if there exists a nonnegative function f_X, defined for all real $x \in \boldsymbol{R}$, having the property that for any (reasonable) set A of real numbers,

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

where the function f_X is called the **probability density function** (pdf) **of the random variable** X.

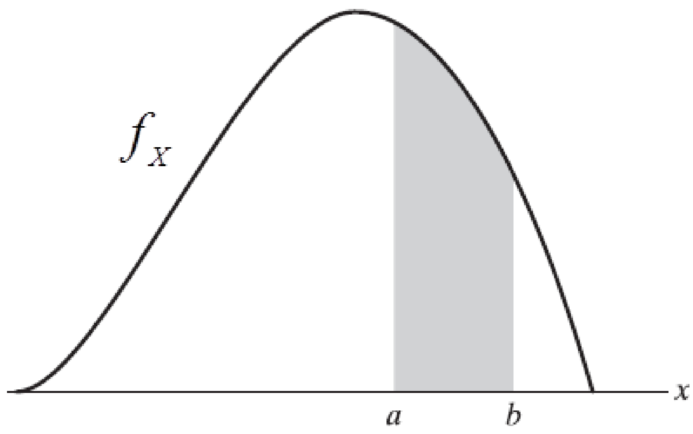
Letting $A = [a, b]$, we obtain

$$P(a \le X \le b) = \int_a^b f_X(x)dx.$$

If we let $a = b$, we get

$$P(X = a) = \int_a^a f_X(x)dx = 0.$$

In words, this equation states that the probability that a continuous random variable will assume any fixed value is zero.



$P(a \le X \le b)$ = area of shaded region

It states that the probability that X will be in A may be obtained by integrating the probability density function over the set A. Since X must take some value in $\boldsymbol{R}$, f_X must satisfy

$$1 = P(X \in \boldsymbol{R}) = \int_{-\infty}^{\infty} f_X(x)dx.$$

The **distribution function of a continuous random variable** X is

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)dt, \quad \text{for all } x \in \boldsymbol{R}.$$

Differentiating both sides of the preceding equation yields

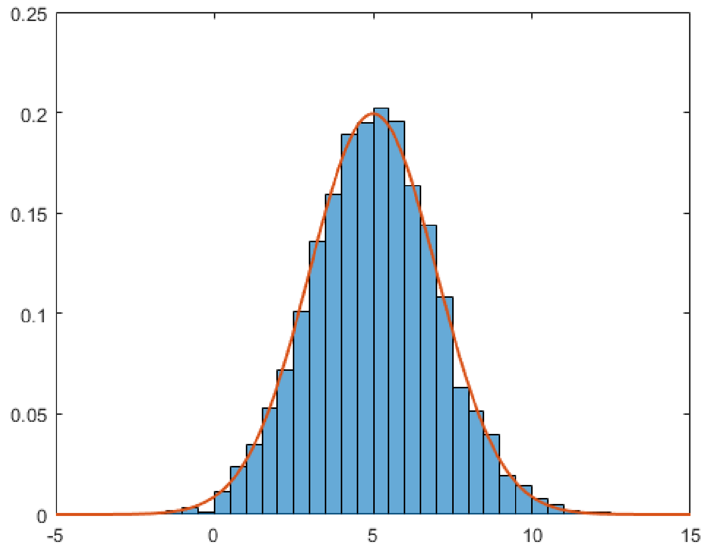
$$\frac{d}{dx} F_X(x) = f_X(x).$$

That is, the density is the derivative of the cumulative distribution function.

A somewhat more intuitive interpretation of the density function may be obtained from as follows:

$$P\left(a - \frac{\varepsilon}{2} \le X \le a + \frac{\varepsilon}{2}\right) = \int_{a+\varepsilon/2}^{a-\varepsilon/2} f_X(x)dx \approx \varepsilon f_X(a)$$

when ε is small and when f_X is continuous at $x = a$. In other words, the probability that X will be contained in an interval of length ε around the point a is approximately $\varepsilon f_X(a)$. From this result, we see that $f_X(a)$ is a measure of how likely it is that the random variable will be near a.



Example 4.2

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

$$f_X(x) = \begin{cases} cx^2 & -1 \le x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find a suitable value of c.

(b) Find the cumulative function F_X.

(c) What is $P(0 \le X \le 1.5)$?

Solution

(a) $1 = P(X \in \boldsymbol{R}) = \int_{-\infty}^{\infty} f_X(x)dx = \int_{-1}^{2} cx^2 dx = c\left.\dfrac{x^3}{3}\right|_{-1}^{2} = 3c.$ Then $c = \dfrac{1}{3}.$

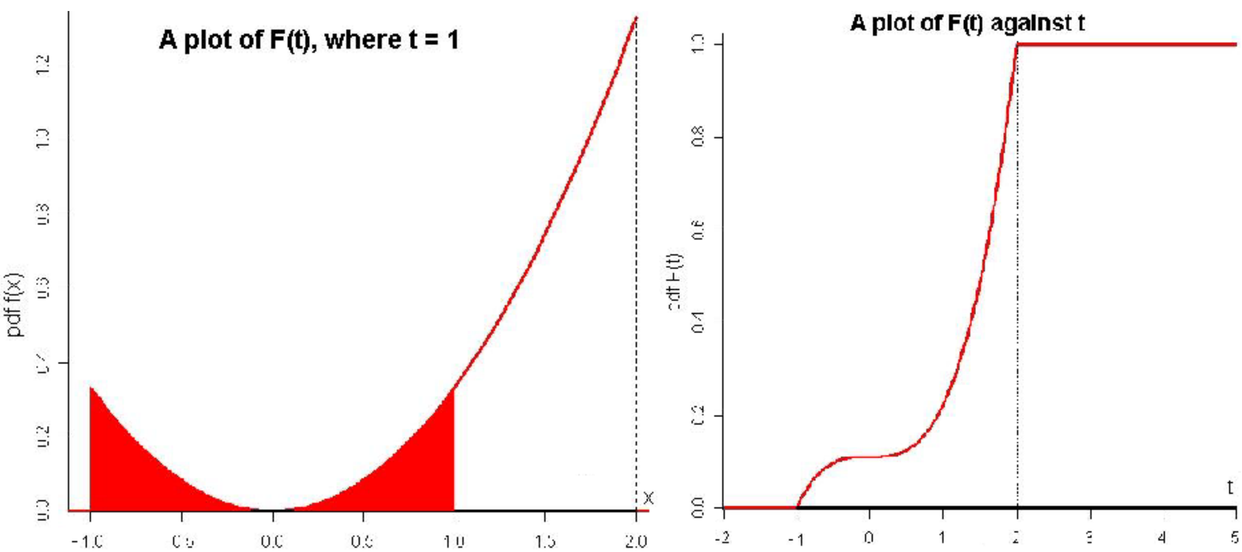
Unlike the probability mass function of a discrete random variable, the probability density function is NOT a probability. Note that we only require that the density function $f_X(x) \geq 0$. So, it can be greater than 1!!! For instance, $f_X(2) = \frac{4}{3} > 1$.

(b) Clearly, $F_X(x) = 0$ if $x < -1$, and $F_X(x) = 1$ if $x > 2$. For $-1 \leq x \leq 2$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{-1}^{x} \frac{1}{3}t^2 dt = \frac{1}{9}t^3\Big|_{-1}^{x} = \frac{1}{9}(x^3+1).$$

In conclusion, we have

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \frac{1}{9}(x^3+1) & -1 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$



(c) $P(0 \leq X \leq 1.5) = \int_{0}^{1.5} \frac{1}{3}x^2 dx = \frac{1}{9}x^3\Big|_{0}^{1.5} = \frac{3}{8}$. On the other hand,

$$P(0 \leq X \leq 1.5) = F_X(1.5) - F_X(0) = \frac{1}{9}(1.5^3+1) - \frac{1}{9} = \frac{3}{8}.$$

Definition 4.3

Let X be a continuous random variable with probability density function $f_X(x)$, then, for any real-valued function g,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)dx.$$

Replacing the sum and the probability mass function for a discrete random variable by an integral and the probability density function for a continuous random variable, we also have

$$E[aX+b] = \int_{-\infty}^{\infty} (ax+b) f_X(x)dx = a\int_{-\infty}^{\infty} xf_X(x)dx + b\int_{-\infty}^{\infty} f_X(x)dx = aE[X]+b.$$

Definition 4.4

The **variance of a continuous random variable** X is defined exactly as it is for a discrete random variable, namely, if X is a random variable with expected value μ, then the variance of X is defined (for any type of random variable) by

$$Var(X) = E\left[(X-\mu)^2\right].$$

The alternative formula,

$$Var(X) = E[X^2] - E[X]^2$$

is established in a manner similar to its counterpart in the discrete case. We still have

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

Definition 4.5

Insurance policies are often sold with a per-loss **deductible of** d. When the loss, x, is at or below d, the insurance pays nothing. When the loss is above d, the insurance pays $x-d$. Let X be the random variable of the loss from policy holder. The deductible of d that paid by insurance company is

$$(X-d)^+ = \max\{X-d, 0\} = \begin{cases} X-d & \text{if } X \ge d \\ 0 & \text{if } X < d. \end{cases}$$

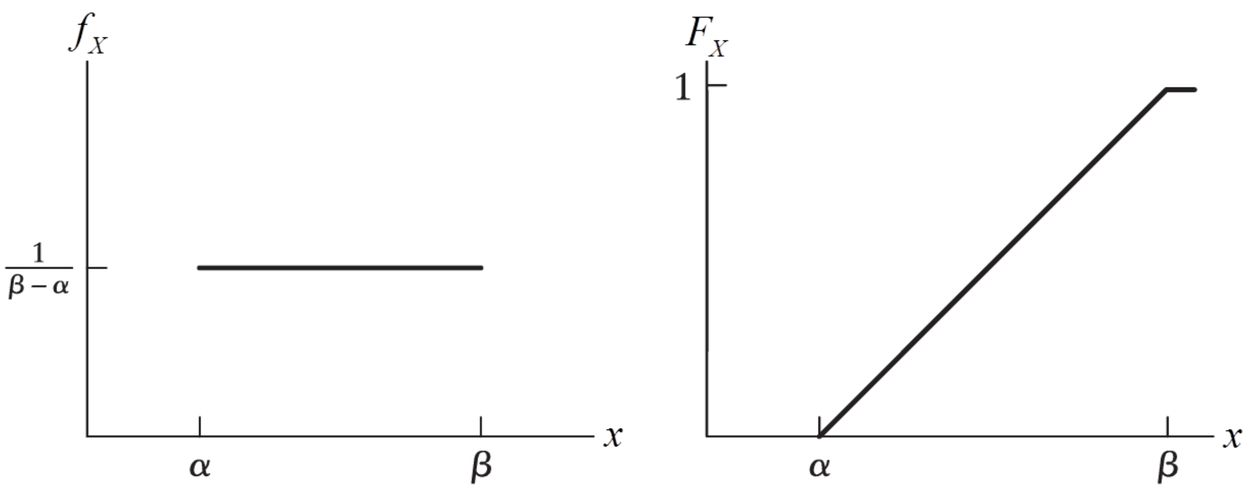
Definition 4.6 (Uniform (α, β))

A random variable is said to be **uniformly distributed over the interval** (α, β) if its probability density function is given by

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

Equivalently, its cumulative distribution function is given by

$$F_X(x) = \begin{cases} 0 & x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \alpha \le x \le \beta \\ 1 & x > \beta \end{cases}$$



Theorem 4.7

Let X be uniformly distributed over (a, b). Then

$$E[X] = \frac{a+b}{2} \quad \text{and} \quad Var[X] = \frac{(b-a)^2}{12}.$$

In words, the expected value of a random variable that is uniformly distributed over some interval is equal to the midpoint of that interval.

Definition 4.8 (Exponential (λ))

A continuous random variable whose probability density function is given, for some $\lambda > 0$, by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is said to be an **exponential random variable** (or, more simply, is said to be **exponentially distributed**) **with parameter** λ. Equivalently, its cumulative density function is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

The exponential distribution is used for instance in physics to represent the lifetime of a particle, the parameter λ representing the exponential decay rate at which the particle ages.

Theorem 4.9

Let X be an exponential random variable with parameters λ. Then

$$E[X] = \frac{1}{\lambda} \quad \text{and} \quad Var[X] = \frac{1}{\lambda^2}.$$

Example 4.10

An auto insurance policy has a deductible of 1 and a maximum claim payment of 5. Auto loss amounts follow an exponential distribution with mean 2. Calculate the expected claim payment paid by insurance company for an auto loss.

Solution

The expected claim payment is

$$\begin{aligned} \int_1^6 (x-1)0.5e^{-0.5x}dx + \int_6^{\infty} 5 \times 0.5e^{-0.5x}dx &= -(x-1)e^{-0.5x}\Big|_1^6 + \int_1^6 e^{-0.5x}dx - \left[5e^{-0.5x}\right]_6^{\infty} \\ &= -5e^{-3} + 0 - \left[2e^{-0.5x}\right]_1^6 + 5e^{-3} \\ &= -2e^{-3} + 2e^{-1/2} \end{aligned}$$

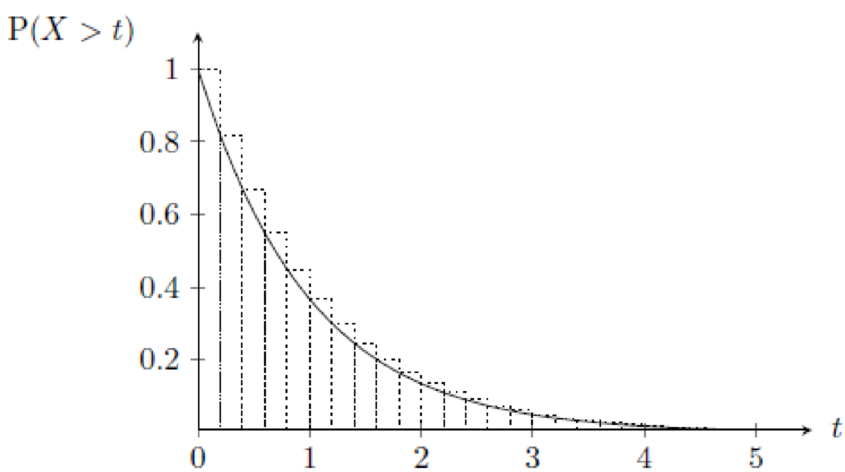
Theorem 4.11

Exponential random variable X satisfies memoryless property, that is

$$P(X > s+t \mid X > t) = P(X > s) \quad \text{for all } s, t \geq 0.$$

If we think of *X* as being the lifetime of some instrument, the equation states that the probability that the instrument survives for at least $s + t$ hours, given that it has survived *t* hours, is the same as the initial probability that it survives for at least *s* hours. In other words, if the instrument is alive at age *t*, the distribution of the remaining amount of time that it survives is the same as the original lifetime distribution. (That is, it is as if the instrument does not "remember" that it has already been in use for a time *t*.)

The exponential distribution is in continuous time but the geometric distribution is in discrete time. Both of them satisfy memoryless property.



Approximation of the exponential distribution by the geometric distribution.

Example 4.12

The time *T* required to repair a machine is an exponentially distributed random variable with mean $\frac{1}{2}$ (hours). What is the probability that a repair takes at least $12\frac{1}{2}$ hours given that its duration exceeds 12 hours?

Solution

Clearly $\lambda = 2$. $P(T > 12\frac{1}{2} \mid T > 12) = P(T > \frac{1}{2}) = e^{-1}$.

Definition 4.13 (Gamma (α, λ))

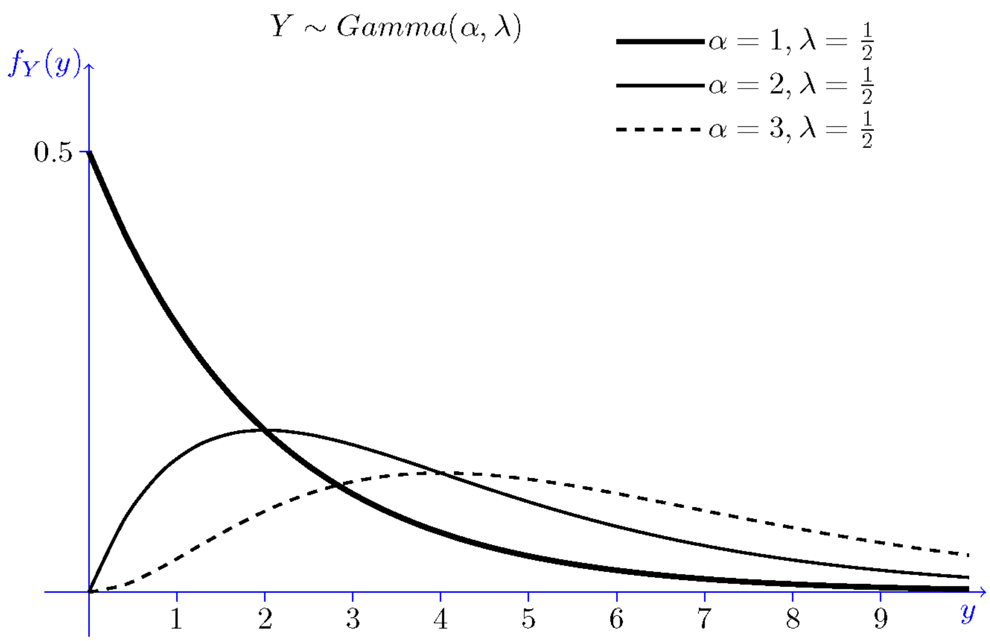
A random variable is said to have a **gamma distribution with parameters** $(\alpha, \lambda), \lambda > 0, \alpha > 0,$ if its density function is given by

$$f_X(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

where $\Gamma(\alpha),$ called the **gamma function**, is defined as

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy.$$

Using integration by parts, $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1).$ In particular, $\Gamma(n) = (n-1)!$ for positive integer n.



Theorem 4.14

Let X be a gamma random variable with parameters $(\alpha, \lambda).$ Then

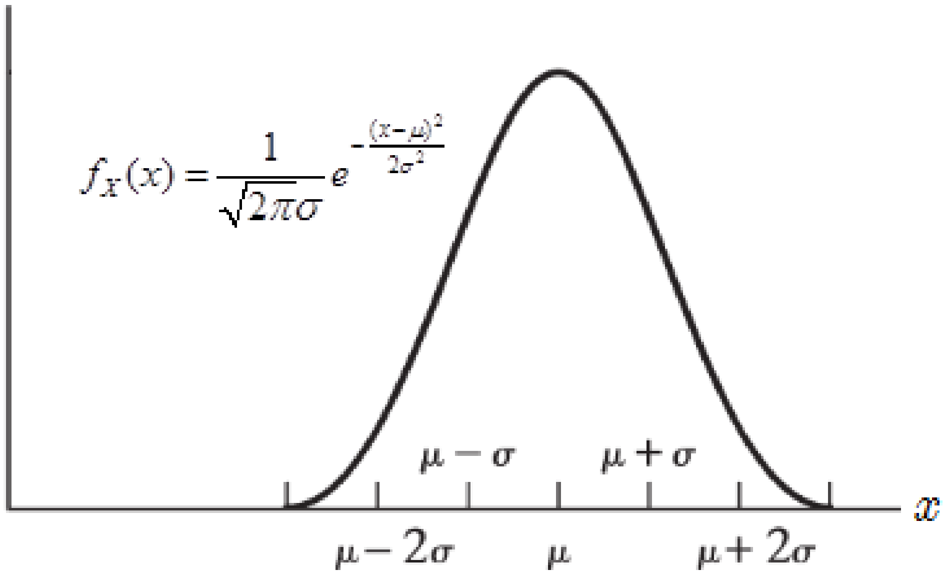
$$E[X] = \frac{\alpha}{\lambda} \quad \text{and} \quad Var[X] = \frac{\alpha}{\lambda^2}.$$

Definition 4.15 (Normal (μ,σ^2))

We say that X is a **normal random variable**, or simply that X is **normally distributed**, **with parameters** μ and σ^2 if the density of X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

This density function is a bell-shaped curve that is symmetric about μ.



Theorem 4.16

If X is normally distributed with parameters μ and σ^2, then $Y = aX + b$ is normally distributed with parameters $a\mu + b$ and $a^2\sigma^2$.

Proof

Suppose that $a > 0$. (The proof when $a < 0$ is similar.) Let F_Y denote the cumulative distribution function of Y. Then

$$
\begin{aligned}
F_Y(x) &= P(Y \le x) \\
&= P(aX + b \le x) \\
&= P\left(X \le \frac{x-b}{a} \right) \\
&= F_X\left(\frac{x-b}{a} \right)
\end{aligned}
$$

where F_X is the cumulative distribution function of *X*. By differentiation, the density function of *Y* is then

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

which shows that *Y* is normal with parameters $a\mu + b$ and $a^2\sigma^2$.

If *X* is normally distributed with parameters μ and σ^2, then $Z = \frac{X-\mu}{\sigma}$ is normally distributed with parameters 0 and 1. Such a random variable is called a **standard normal random variable**.

Theorem 4.17

If *X* is normally distributed with parameters μ and σ^2, then

$$E[X] = \mu \text{ and } Var[X] = \sigma^2.$$

Theorem 4.18

Let *X* be a continuous random variable having probability density function 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 probability density function given by

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

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

Proof

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

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

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. If *g*(*x*) is a decreasing function,

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

In conclusion, $f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\dfrac{d}{dy} g^{-1}(y)\right|.$

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

Example 4.19

Let *X* be normal distribution with parameters $(\mu, \sigma^2),$ and let $Y = e^X$. Find $f_Y(y).$

Solution

Method 1:

The probability density function of *X* is $f_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$

$$F_Y(y) = P(Y \le y) = P\left(e^X \le y\right) = P\left(X \le \ln y\right) = F_X(\ln y) \quad \Rightarrow \quad f_Y(y) = \frac{1}{y} f_X(\ln y)$$

$$f_Y(y) = \begin{cases} \dfrac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}} & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

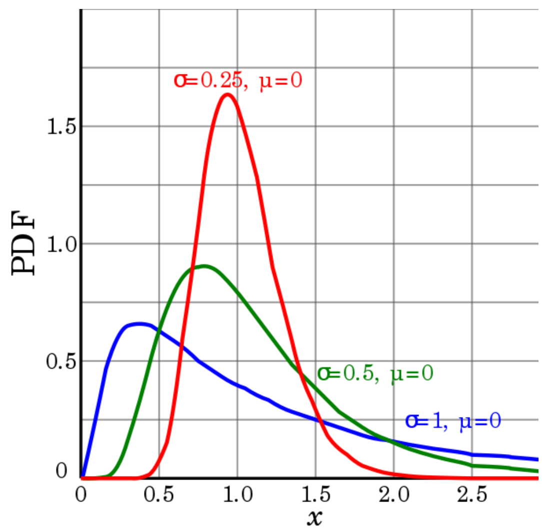
Method 2:

Let $y = g(x) = e^x$. Then $g^{-1}(y) = \ln y$ and $\dfrac{d}{dy}\left(g^{-1}(y)\right) = \dfrac{1}{y}$. Furthermore $y > 0$ if $-\infty < x < \infty$.

$$f_Y(y) = \begin{cases} f_X(\ln y)\left|\dfrac{1}{y}\right| = \dfrac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}} & \text{if } y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Definition 4.20 (Lognormal (μ, σ^2))

The random variable Y in Example 3.25 is said to be a lognormal random variable with parameters μ and σ^2.

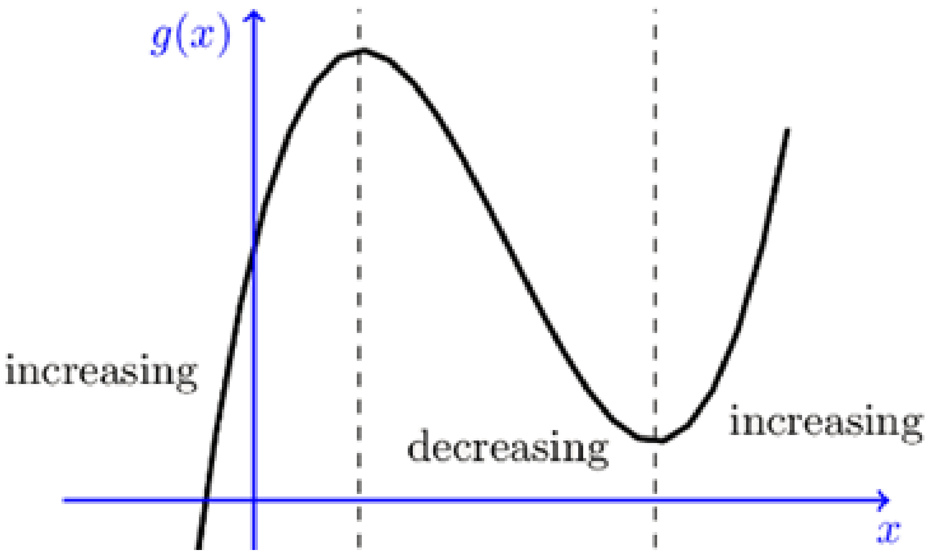


Theorem 4.21

Let Y be lognormal random variable with parameters μ and σ^2. Then

$$y_{0.5} = e^{\mu}, \quad E[Y] = e^{\mu+\frac{1}{2}\sigma^2} \quad \text{and} \quad Var[Y] = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right).$$

Theorem 4.18 can be extended to a more general case. In particular, if g is not monotonic, we can usually divide it into a finite number of monotonic differentiable functions. The below figure shows a function g that has been divided into monotonic parts.



Partitioning a function to monotone parts.

Theorem 4.22

Let X be a continuous random variable with probability density function f_X. Let g be piecewise strictly monotone and continuously differentiable: that is, there exist intervals $I_1, I_2, \cdots, I_n$ which partition range of X such that g is strictly monotone and continuously differentiable on the interior of each I_i. Then the probability density function of $Y = g(X)$ is given by

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} \frac{f_X(x_i)}{|g'(x_i)|} = \sum_{i=1}^{k} f_X(x_i) \left| \frac{dx_i}{dy} \right| & \text{for some solutions } x_1, x_2, \cdots, x_n \text{ for } y = g(x) \\ 0 & \text{if no solution for } y = g(x) \end{cases}$$

Example 4.23

Let X be a uniform random variable over an interval $(-1,1)$ with probability density function

$$f_X(x) = \begin{cases} \frac{1}{2} & x \in (-1,1) \\ 0 & \text{otherwise} \end{cases}$$

and let $Y = X^2$. Find $f_Y(y)$.

Solution

Method 1:

Clearly, $F_Y(y) = P(Y \le y) = P(X^2 \le y) = 0$ if $y \le 0$. Then $f_Y(y) = 0$. Let $y > 0$. The distribution of Y is

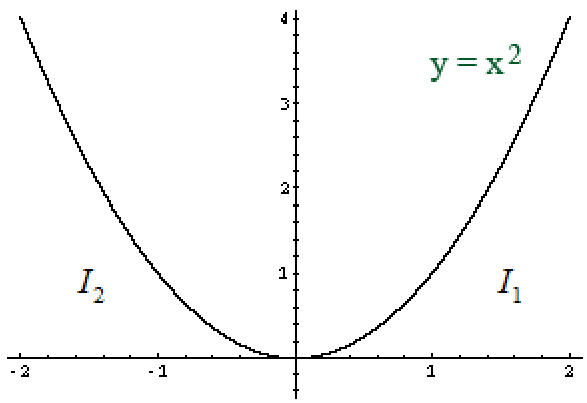
$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P\left(-\sqrt{y} \le X \le \sqrt{y}\right) = F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right) = \begin{cases} \sqrt{y} & 0 < y < 1 \\ 1 & y \ge 1. \end{cases}$$

Its probability density function is

$$f_Y(y) = F_Y'(y) = \begin{cases} \dfrac{1}{2\sqrt{y}} & 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Method 2:

Clearly, $g(x) = x^2$ is not monotone. Partition $R = (-\infty, 0) \cup [0, \infty)$ such that g is strictly monotone on $I_1 = [0, \infty)$ and $I_2 = (-\infty, 0)$.



Suppose $y > 0$. Let $x_1 = \sqrt{y} \in I_1$ and $x_2 = -\sqrt{y} \in I_2$. The probability density function of Y is

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} \frac{f_X(x_i)}{|g'(x_i)|} = \frac{f_X\left(\sqrt{y}\right)}{|g'\left(\sqrt{y}\right)|} + \frac{f_X\left(-\sqrt{y}\right)}{|g'\left(-\sqrt{y}\right)|} = 0 & \text{if } y \geq 1 \\ \sum_{i=1}^{k} \frac{f_X(x_i)}{|g'(x_i)|} = \frac{f_X\left(\sqrt{y}\right)}{|g'\left(\sqrt{y}\right)|} + \frac{f_X\left(-\sqrt{y}\right)}{|g'\left(-\sqrt{y}\right)|} = \frac{\frac{1}{2}}{|2\sqrt{y}|} + \frac{\frac{1}{2}}{|-2\sqrt{y}|} = \frac{1}{2\sqrt{y}} & \text{if } 0 < y < 1 \\ 0 & \text{if } y \leq 0 \end{cases}$$

$$= \begin{cases} \frac{1}{2\sqrt{y}} & \text{if } 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$