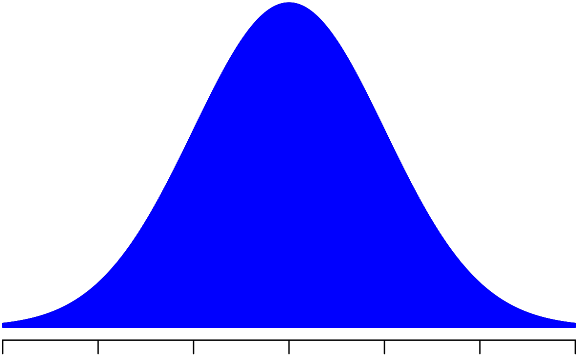
Ch. 4 – Continuous Random Variables



Continuous Random Variables

So far, we have only discussed discrete random variables, which have only a sequence of possible values (usually whole numbers):

- The number of defective widgets in a batch.
- The number of widgets inspected before finding one defective.
- The number of customers who visit a store in an hour.

However, many quantities in real life vary continuously:

- The length of a metal rod.
- The strength of a specimen of concrete.
- The weight of a bottled drink.
- The amount of time until the next customer arrives.

We will need different techniques to deal with continuous random variables.

Uniform Random Variable

Suppose we choose a random number X "uniformly" from the interval $[0, 1]$. What is the probability that we get a number between 0.2 and 0.6?

The interval $[0.2, 0.6]$ has length $0.6 - 0.2 = 0.4$, which is 40% of the total length of the interval $[0, 1]$. Therefore, intuitively the probability that X would be in the interval $[0.2, 0.6]$ should be

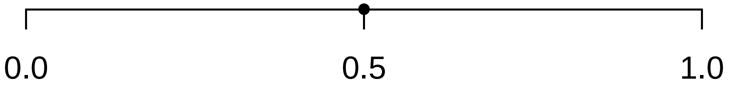
$$P(0.2 \leq X \leq 0.6) = 0.6 - 0.2 = 0.4$$

In general, for an interval $[a, b]$ inside $[0, 1]$ we should have

$$P(a \leq X \leq b) = b - a$$

Uniform Random Variable

Suppose we choose a random number from the interval $[0, 1]$. What is the probability that we get exactly the number 0.5?



The probability is $P(0.5 \leq X \leq 0.5) = 0.5 - 0.5 = 0$. In fact, for any x in $[0, 1]$ the probability that X is exactly x is 0. And yet, X will always be some number in $[0, 1]$.

- Even if an event has probability 0, that doesn't mean it is impossible for it to occur.
- For a continuous random variable, the concept of a probability mass function is useless: every probability $P(X = x)$ is zero.
- We *cannot* find the probability $P(a \leq X \leq b)$ by simply adding up all the probabilities $P(X = x)$ over all x in $[a, b]$.

Continuous Random Variable

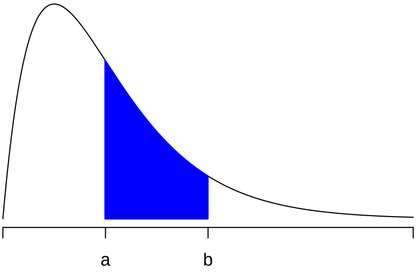
We say that a random variable X is **continuous** if $P(X = x) = 0$ for every x. If there is a function $f(x)$ such that for all $a \leq b$,

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

then we call $f(x)$ a **probability density function** (pdf) of X.

To be a valid pdf, we must have

1. $f(x) \geq 0$ for all x.
2. $\int_{-\infty}^{\infty} f(x) = 1$.



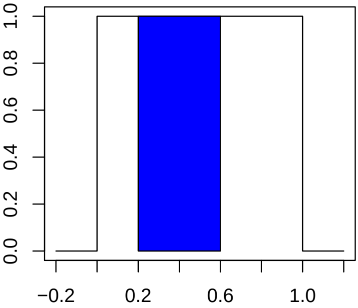
Define a pdf by

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

The continuous random variable X with this pdf is called a **standard uniform** random variable; it takes values uniformly on the interval [0,1].

For example, the probability that X is between .2 and .6 is

$$\begin{aligned} &P(.2 \leq X \leq .6) \\ &= \int_{.2}^{.6} 1 \; dx \\ &= x|_{.2}^{.6} \\ &= .6 - .2 \\ &= .4 \end{aligned}$$



Uniform Random Variable

We say that X is a **uniform** random variable on the interval $[a, b]$ if X has pdf

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

Example: Suppose that the time we have to wait at a bus stop is a uniform random variable X between 0 and 15 minutes. What is the probability that we will have to wait more than 10 minutes?

$$\begin{aligned} P(X \geq 10) &= \int_{10}^{\infty} f(x)\ dx \\ &= \int_{10}^{15} \frac{1}{15-0}\ dx \\ &= \frac{1}{15} x \Big|_{10}^{15} \\ &= \frac{15-10}{15} = 1/3 \end{aligned}$$

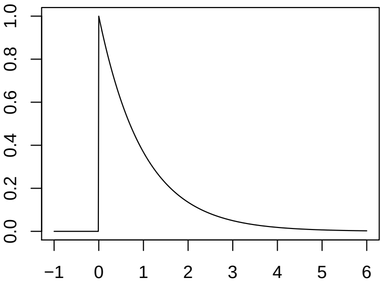
Exponential Random Variable

We say that X is an **exponential** random variable with rate $\lambda > 0$ if X has pdf

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

We can check that this is a valid pdf:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\ dx &= \int_{0}^{\infty} \lambda e^{-\lambda x}\ dx \\ &= \lambda \frac{-1}{\lambda} e^{-\lambda x} \Big|_{x=0}^{\infty} \\ &= 0 - (-1) = 1 \end{aligned}$$



CDF of Continuous Random Variable

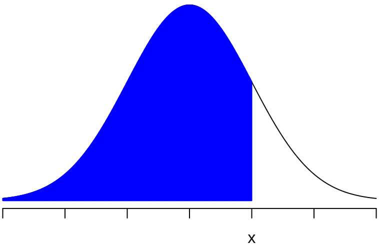
The **cumulative distribution function** (cdf), $F(x)$, of a continuous random variable X is defined the same as in the discrete case:

$$F(x) = P(X \leq x)$$

If X has pdf $f(x)$, then this becomes

$$F(x) = \int_{-\infty}^{x} f(t)\ dt$$

By the Fundamental Theorem of Calculus, $F'(x) = f(x)$, if f is continuous at x .



CDF of Standard Uniform Random Variable

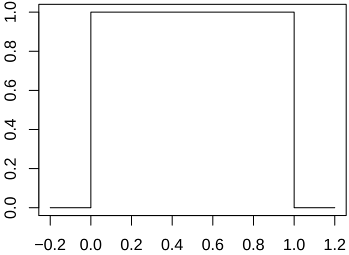
Recall that a standard uniform random variable X has pdf

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

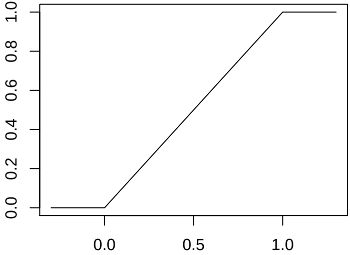
Let $F(x)$ be the cdf of X. For $0 \le x \le 1$, we have

$$F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{0}^{x} 1\ dt = x$$

For $x \le 0$, clearly $F(x) = 0$, while for $x \ge 1$, $F(x) = 1$.



$f(x) =$ pdf of X



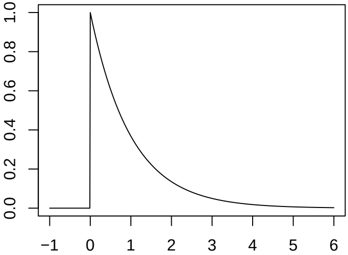
$F(x) =$ cdf of X

Recall the pdf of an exponential random variable x with rate λ is

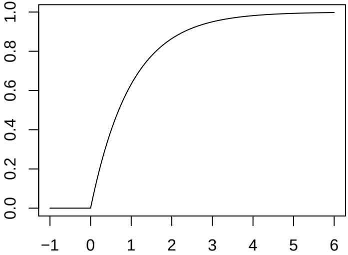
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

The cdf $F(x)$ is therefore, for $x \geq 0$,

$$F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_{0}^{x} \lambda e^{-\lambda t}\ dt$$
$$= -e^{-\lambda t}\Big|_{t=0}^{x} = 1 - e^{-\lambda x}$$



$f(x) =$ pdf of X



$F(x) =$ cdf of X

Expected Value

The **expected value** or **mean** of a continuous random variable X with pdf $f(x)$ is

$$E(X) = \int_{-\infty}^{\infty} xf(x)\ dx$$

For example, the mean of a uniform random variable X on $[a, b]$ is

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xf(x)\ dx \\
&= \int_{a}^{b} \frac{x}{b-a}\ dx \\
&= \frac{x^2}{2(b-a)}\Big|_a^b \\
&= \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}
\end{aligned}$$

Mean of Exponential

Using integration by parts, we can find the mean of an exponential random variable X of rate λ:

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x)\ dx \\
&= \int_{0}^{\infty} x \lambda e^{-\lambda x}\ dx \\
&= -x e^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\ dx \\
&= 0 - 0 - \frac{1}{\lambda} e^{-\lambda x}\Big|_0^{\infty} = \frac{1}{\lambda}
\end{aligned}
$$

Example

Suppose that the lifetime X of a lightbulb follows an exponential distribution with mean $\mu = 100$ days. What is the probability that the lifetime is at least 50 days?

Solution: The rate of failure is $\lambda = 1/\mu = 1/100$ per day. Therefore,

$$\begin{aligned} P(X \geq 50) &= \int_{50}^{\infty} \lambda e^{-\lambda x} \, dx \\ &= -e^{-\lambda x} \Big|_{50}^{\infty} \\ &= e^{-50\lambda} \\ &= e^{-50/100} = e^{-1/2} \approx .607 \end{aligned}$$

In general, if X is an exponential random variable with rate λ,

$$P(X \geq t) = e^{-\lambda t}$$

Example

Again suppose that the lifetime X of a lightbulb follows an exponential distribution with mean $\mu = 100$ days. Given that the bulb has survived for 30 days, what is the probability that it will last for at least 50 more?

$$
\begin{aligned}
P(X \geq 80 \mid X \geq 30) &= \frac{P(X \geq 80 \cap X \geq 30)}{P(X \geq 30)} \\
&= \frac{P(X \geq 80)}{P(X \geq 30)} \\
&= \frac{e^{-80\lambda}}{e^{-30\lambda}} \\
&= e^{-50\lambda} = e^{-50/100} = e^{-1/2} \approx .607
\end{aligned}
$$

This is the same as the probability of a new bulb lasting 50 days, as we calculated on the previous slide.

Memoryless Property of Exponential

The exponential distribution is useful for modeling the lifetime of components where breakdowns are the result of sudden, random failures rather than gradual deterioration.

In general, an exponential random has the "memoryless" property:

$$P(X \geq s + t \mid X \geq s) = P(X \geq t)$$

Proof:
$$\begin{aligned} P(X \geq s + t \mid X \geq s) &= \frac{P(X \geq s + t \cap X \geq s)}{P(X \geq s)} \\ &= \frac{P(X \geq s + t)}{P(X \geq s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = P(X \geq t) \end{aligned}$$

Exponential Waiting Times

Consider a Poisson process with rate λ.

- Let Y_t be the number of events occurring in the interval $[0, t]$.
- So Y_t is a Poisson random variable with mean λt.
- Let X be the time of the first event.

$$\begin{aligned}
P(X \leq t) &= P(\text{first event occurs by time } t) \\
&= P(\text{at least one event occurs in the interval } [0, t]) \\
&= P(Y_t \geq 1) \\
&= 1 - P(Y_t = 0) \\
&= 1 - \frac{e^{-\lambda t}(-\lambda t)^0}{0!} \\
&= 1 - e^{-\lambda t}
\end{aligned}$$

This is the cdf of an exponential random variable of rate λ. Therefore, in a Poisson process, the waiting time for the first event is an exponential random variable with rate λ.

Variance and Standard Deviation

The **variance** of a continuous random variable X with pdf $f(x)$ is

$$V(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x^2 - \mu) f(x)\, dx$$

As before, the **standard deviation** is $\sigma = \sqrt{V(X)}$.

The shortcut formula for the variance also works for continuous random variables:

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

Variance of Uniform Distribution

$$E(X^2) = \int_a^b \frac{x^2}{b-a}\,dx = \frac{x^3}{3(b-a)}\bigg|_a^b = \frac{b^3-a^3}{3(b-a)}$$
$$= \frac{(b-a)(a^2+ab+b^2)}{3(b-a)} = \frac{a^2+ab+b^2}{3}$$

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} \\ &= \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12} \end{aligned}$$

Variance of Exponential Distribution

We can find the variance of an exponential random variable X by using integration by parts twice:

$$
\begin{aligned}
E(X^2) &= \int_0^\infty x^2 \cdot \lambda e^{-\lambda x}\ dx \\
&= -x^2 e^{-\lambda x}\Big|_0^\infty - \int_0^\infty -2xe^{-\lambda x}\ dx \\
&= 2\int_0^\infty xe^{-\lambda x}\ dx \\
&= 0 - 0 + 2\left[-\frac{1}{\lambda}xe^{-\lambda x}\Big|_0^\infty - \int_0^\infty -\frac{1}{\lambda}e^{-\lambda x}\ dx\right] \\
&= 2\left[0 - 0 - \frac{1}{\lambda^2}e^{-\lambda x}\Big|_0^\infty\right] = \frac{2}{\lambda^2}
\end{aligned}
$$

So $V(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$. In other words, the standard deviation is $\sigma = 1/\lambda = \mu$.

Properties of Expected Value and Variance

The same properties of expected value and variance which we used for discrete random variables also work for continuous random variables:

1. $E(c) = c$
2. $E(cX) = cE(X)$
3. $E(X + Y) = E(X) + E(Y)$

1. $V(c) = 0$
2. $V(cX) = c^2 V(X)$
3. $V(X + c) = V(X)$
4. If X and Y are independent, $V(X + Y) = V(X) + V(Y)$.

Shifting and Scaling a Uniform Distribution

Starting with a uniform random variable X on $[a, b]$, if we add or multiply by a constant c, then we obtain a new uniform random variable:

Namely, $X + c$ is a uniform random variable on $[a + c, b + c]$, while cX is a uniform random variable on $[ca, cb]$.

Starting from a standard uniform random variable U on $[0, 1]$, a uniform random variable X on $[a, b]$ may be obtained by scaling and shifting:

$$X = (b - a)U + a$$

- First notice that $(b - a)U$ is uniform on $[0, b - a]$.
- Therefore $(b - a)U + a$ is uniform on $[0 + a, (b - a) + a] = [a, b]$.

Variance of Uniform Distribution

We can use properties of variance to more easily find the variance of a uniform random variable X on $[a, b]$:

First we find the variance of a standard uniform random variable U:

$$E(U^2) = \int_{-\infty}^{\infty} x^2 f(x)\ dx = \int_0^1 x^2\ dx = 1/3$$
$$V(U) = E(U^2) - [E(U)]^2 = 1/3 - (1/2)^2 = 1/12$$

Now any uniform random variable X on $[a, b]$ may be written in terms of a standard uniform random variable as $X = (b-a)U + a$. Therefore,

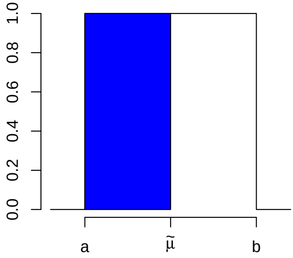
$$\begin{aligned} V(X) &= V((b-a)U + a) \\ &= V((b-a)U) \\ &= (b-a)^2 V(U) = \frac{(b-a)^2}{12} \end{aligned}$$

Median of a Continuous Random Variable

Given a continuous random variable X, the **median** of X is the value $\tilde{\mu}$ such that $P(X \leq \tilde{\mu}) = \frac{1}{2}$.

Example: The median of a uniform random variable X on $[a, b]$ is $\frac{a+b}{2}$, since

$$
\begin{aligned}
P(X \leq \frac{a+b}{2}) &= \int_a^{\frac{a+b}{2}} f(x)\ dx \\
&= \int_a^{\frac{a+b}{2}} \frac{1}{b-a}\ dx \\
&= \frac{1}{b-a}\left(\frac{a+b}{2} - a\right) \\
&= \frac{1}{b-a} \cdot \frac{b-a}{2} = \frac{1}{2}
\end{aligned}
$$



In this case, the median $\tilde{\mu}$ is the same as the mean μ.

Median of a Exponential Random Variable

Find the median of an exponential random variable X with rate λ.

Solution: We need to find x such that $P(X \leq x) = 1/2$. Recall

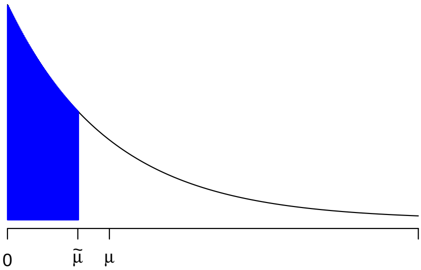
$$P(X \leq x) = F(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x}$$

We just need to solve

$$\frac{1}{2} = 1 - e^{-\lambda x}$$

for x. This gives

$$x = \frac{\ln 2}{\lambda} = \mu \ln 2$$



So the median is $\tilde{\mu} = \mu \ln 2 \approx .693\mu$.

Percentiles of a Continuous Random Variable

Given a continuous random variable X and $0 \le p \le 1$, the $100p$th **percentile** of X is the value x such that $P(X \le x) = p$.

Note: The *median* is the same thing as the 50th percentile.

Example: Find the 90th percentile of a uniform random variable X on [10,40]:

$$.9 = P(X \le x) = \int_{10}^{x} \frac{1}{40-10}\, dt = \frac{x-10}{30}$$

Solving for x,

$$x = (.9)(30) + 10 = 37$$

So the 90th percentile of X is 37.

Problem

Given a Poisson process with rate λ, find the pdf of the amount of time X we must wait until k events have occurred.

Let Y_t be the number of events which have occurred by time t, so Y_t is a Poisson random variable with mean $\mu = \lambda t$. The cdf of X is

$$\begin{aligned}
F(t) = P(X \leq t) &= P(Y_t \geq k) \\
&= 1 - P(Y_t < k) \\
&= 1 - \sum_{j=0}^{k-1} \frac{e^{-\mu}\mu^j}{j!} \\
&= 1 - \sum_{j=0}^{k-1} \frac{e^{-\lambda t}(\lambda t)^j}{j!}
\end{aligned}$$

Differentiating the cdf, we find the pdf of X:

$$f(t) = F'(t) = \frac{d}{dt}\left[1 - \sum_{j=0}^{k-1} \frac{e^{-\lambda t}(\lambda t)^j}{j!}\right] = -\sum_{j=0}^{k-1} \frac{\lambda^j}{j!}\frac{d}{dt}(e^{-\lambda t}t^j)$$

$$= -\sum_{j=0}^{k-1} \frac{\lambda^j}{j!}(-\lambda e^{-\lambda t}t^j + e^{-\lambda t}jt^{j-1})$$

$$= \sum_{j=0}^{k-1} \frac{\lambda^{j+1}e^{-\lambda t}t^j}{j!} - \sum_{j=1}^{k-1} \frac{\lambda^j e^{-\lambda t}t^{j-1}}{(j-1)!}$$

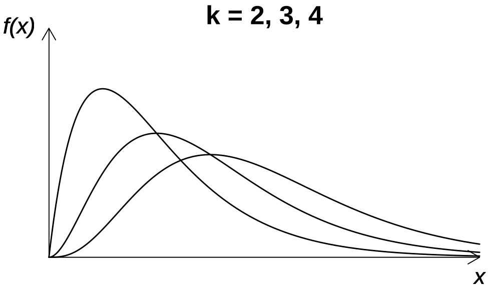
$$= \sum_{j=0}^{k-1} \frac{\lambda^{j+1}e^{-\lambda t}t^j}{j!} - \sum_{j=0}^{k-2} \frac{\lambda^{j+1}e^{-\lambda t}t^j}{j!} = \frac{\lambda^k e^{-\lambda t}t^{k-1}}{(k-1)!}$$

Gamma Distribution

Given a Poisson process with rate λ, the waiting time for k events has a **gamma distribution**, with pdf

$$f(x) = \begin{cases} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

When $k = 1$ this is an exponential distribution: $f(x) = \lambda e^{-\lambda x}$



Gamma as Sum of Exponentials

Given a Poisson process with rate λ, the waiting time for the first event is an exponential random variable Y_1 with rate λ.

After the first event, the waiting time for the next event is an independent exponential random variable Y_2.

Continuing, we see that the waiting time X for k events is the sum of k independent exponential random variables with rate λ:

$$X = Y_1 + \cdots + Y_k$$

In other words, a gamma random variable with parameters k and λ may be expressed as a sum of k independent exponential random variables with rate λ.

Mean and Variance of Gamma Distribution

Expressing a gamma random variable X as sum of independent exponential random variables allows us to easily calculate the mean and variance of X:

$$\begin{aligned}
E(X) &= E(Y_1 + \cdots + Y_k) \\
&= E(Y_1) + \cdots + E(Y_k) \\
&= 1/\lambda + \cdots + 1/\lambda \\
&= k/\lambda
\end{aligned}$$

$$\begin{aligned}
V(X) &= V(Y_1 + \cdots + Y_k) \\
&= V(Y_1) + \cdots + V(Y_k) \\
&= 1/\lambda^2 + \cdots + 1/\lambda^2 \\
&= k/\lambda^2
\end{aligned}$$

Example

Cars pass a certain point on a road according to a Poisson process with rate $\lambda = 20$ per hour. If we wait until 100 cars have passed, what are the mean and standard deviation of the amount of time we will have to wait?

Solution: Let X be the amount of time until 100 cars have passed. X is a gamma random variable with parameters $k = 100$ and $\lambda = 20$. We find the mean and standard deviation of X (in hours) using the formulas on the previous slide:

$$\mu = E(X) = k/\lambda = 5$$
$$\sigma = \sqrt{V(X)} = \sqrt{k/\lambda^2} = \sqrt{1/4} = 1/2$$

Bernoulli Process vs. Poisson Process

A sequence of independent Bernoulli random variables $Y_1, Y_2, \ldots$ each with parameter p is called a **Bernoulli process** with rate p. We interpret this as a process where events occur only at discrete times, 1, 2, 3, $\ldots$, as opposed to a Poisson process where the time of occurence of an event may be any positive real number.

	Bernoulli process	Poisson process
# of events in a unit time period	Bernoulli(p)	Poisson(λ)
# of events in a period of length n	Binomial(n, p)	Poisson($n\lambda$)
Waiting time for first event	Geometric(p)	Exponential(λ)
Waiting time for r events	Negative Binomial (r, p)	Gamma(r, λ)

Standard Gamma Distribution

Recall: the pdf of a gamma random variable with parameters k and λ is

$$f(x) = \begin{cases} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}, & x \geq 0 \\ 0 & x < 0 \end{cases}$$

A gamma random variable with $\lambda = 1$ is called a **standard gamma** random variable. In this case, the pdf becomes

$$f(x) = \begin{cases} \frac{1}{(k-1)!} x^{k-1} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Since $f(x)$ is a valid pdf, $\int_0^\infty \frac{1}{(k-1)!} x^{k-1} e^{-x} \, dx = 1$. In other words,

$$\int_0^\infty x^{k-1} e^{-x} \, dx = (k-1)!$$

Gamma Function

We showed that for any integer $k \geq 1$,

$$\int_0^\infty x^{k-1} e^{-x} \, dx = (k-1)!$$

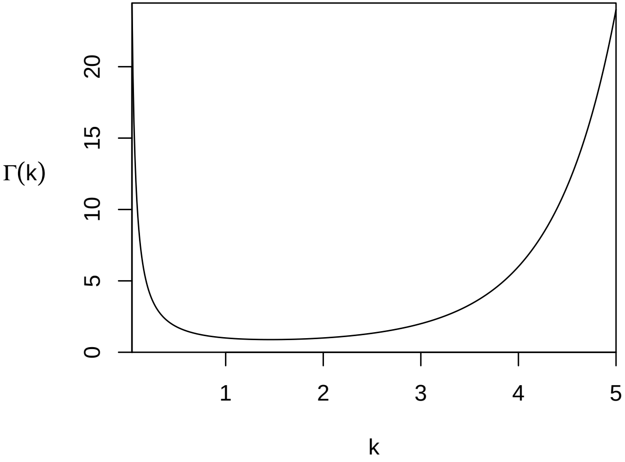
In general, this integral exists for any real number $k > 0$ and is the definition of the **gamma function** $\Gamma(k)$:

$$\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} \, dx$$

For integers $k \geq 1$, then, $\Gamma(k) = (k-1)!$.

Gamma Function

$$\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} \, dx$$



Gamma Distribution

Recall: for integer $k \geq 1$, the pdf of a gamma random variable is

$$f(x) = \begin{cases} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Using the gamma function, we may extend the definition of gamma random variables to include cases where k may not be an integer:

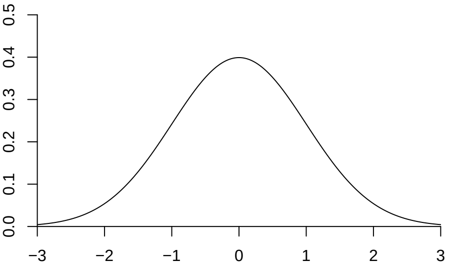
A **gamma** random variable X with parameters $k, \lambda > 0$ is given by the pdf

$$f(x) = \begin{cases} \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Standard Normal Distribution

We say a random variable X has the **standard normal distribution** if it has pdf

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

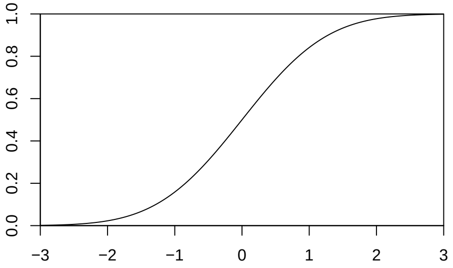


CDF of Standard Normal Distribution

The cdf of the standard normal distribution is

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

There is no simple formula for evaluating this integral. However, it can easily be evaluated numerically by a computer.



Standard Normal Distribution

We want to show that the standard normal pdf is a valid pdf.

We will use the special integral $\int_{-\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi}$, which we will prove later.

Assuming this, substituting $u = x/\sqrt{2}$,

$$\begin{aligned}
\int_{-\infty}^{\infty} \phi(x)\ dx &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\ dx \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2}\ du \\
&= \frac{1}{\sqrt{\pi}} \cdot \sqrt{\pi} = 1
\end{aligned}$$

Mean of Standard Normal

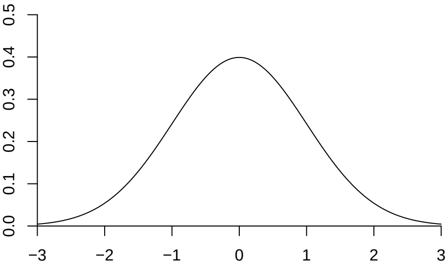
The standard normal distribution is symmetric in the sense that the pdf $\phi(x)$ is an even function, i.e., $\phi(-x) = \phi(x)$:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Therefore the mean of a standard normal random variable is

$$E(X) = \int_{-\infty}^{\infty} x\phi(x)\ dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\ dx = 0$$

since the integrand is an odd function.



Variance of Standard Normal

Substituting $u = x^2/2$, note that

$$\int xe^{-x^2/2}\ dx = \int e^{-u}\ du = -e^{-u} = -e^{-x^2/2}$$

Integrating by parts,

$$\begin{aligned} V(X) = E[(X-\mu)^2] = E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\ dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot xe^{-x^2/2}\ dx \\ &= \frac{1}{\sqrt{2\pi}} \left[x(-e^{-x^2/2})\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-x^2/2}\ dx \right] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\ dx = 1 \end{aligned}$$

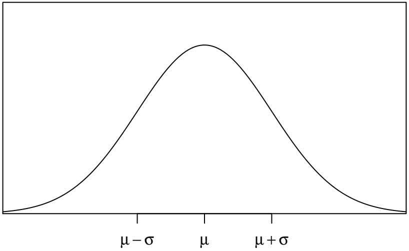
A standard normal random variable has mean 0 and variance 1.

Normal Distributions

If Z is a standard normal random variable, then given numbers μ and $\sigma > 0$, we can define a new scaled, shifted random variable $X = \sigma Z + \mu$:

$$E(X) = E(\sigma Z + \mu) = \sigma E(Z) + E(\mu) = \sigma \cdot 0 + \mu = \mu$$
$$V(X) = V(\sigma Z + \mu) = V(\sigma Z) = \sigma^2 V(Z) = \sigma^2$$

We call X a **normal** random variable with mean μ and standard deviation σ, and we write $X \sim N(\mu, \sigma^2)$.



CDF and PDF of Normal Distributions

The cdf of a normal random variable $X \sim N(\mu, \sigma^2)$ is

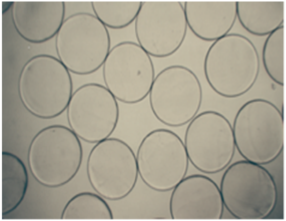
$$\begin{aligned} F(x) &= P(X \leq x) = P(\sigma Z + \mu \leq x) \\ &= P\left(Z \leq \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \end{aligned}$$

By taking the derivative and applying the chain rule, we find the pdf of X:

$$\begin{aligned} f(x) = F'(x) &= \frac{d}{dx}\Phi\left(\frac{x-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma}\Phi'\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{x-\mu}{\sigma})^2/2} \end{aligned}$$

Example

A process produces alginate beads with diameters normally distributed with mean 1 mm and standard deviation .05 mm. Given a random bead, what is the probability that its diameter is between .95 mm and 1.1 mm?

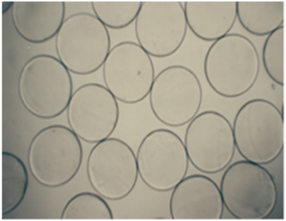


Solution: The random diameter may be written $X = .05Z + 1$, where Z is a standard normal random variable. Then

$$\begin{aligned} P(.95 \leq X \leq 1.1) &= P(.95 \leq .05Z + 1 \leq 1.1) \\ &= P\left(\frac{.95 - 1}{.05} \leq Z \leq \frac{1.1 - 1}{.05}\right) \\ &= P(-1 \leq Z \leq 2) \\ &= \Phi(2) - \Phi(-1) \approx .9772 - .1587 = .8185 \end{aligned}$$

Example – Percentiles

A process produces alginate beads with diameters normally distributed with mean 1 mm and standard deviation .05 mm. Find the 90th percentile of bead diameters.



Solution: As before, the random diameter may be written $X = .05Z + 1$, where Z is a standard normal random variable. If x is the 90th percentile, then

$$\begin{aligned} .9 = P(X \leq x) &= P(.05Z + 1 \leq x) \\ &= P\left(Z \leq \frac{x-1}{.05}\right) = \Phi\left(\frac{x-1}{.05}\right) \end{aligned}$$

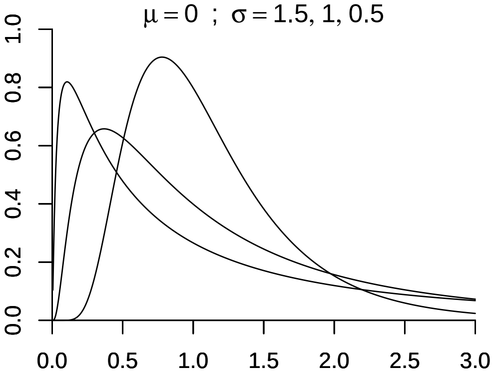
Solving for x,

$$x = .05\Phi^{-1}(.9) + 1 \approx (.05)(1.28) + 1 \approx 1.064 \text{ mm}$$

Log-normal Distribution

A random variable X is said to have a **log-normal** distribution if $\ln(X)$ is a normal random variable.

A log-normal random variable X may be written in the form $X = e^{\sigma Z + \mu}$, where Z is a standard normal random variable.



Example

A certain aerosol spray contains particles whose diameters (in microns) have a log-normal distribution with $\mu = 5$, $\sigma = 0.5$. What proportion of the particles are larger than 300 microns?

Solution: The diameter X of a random particle may be written $X = e^{\sigma Z + \mu}$, so

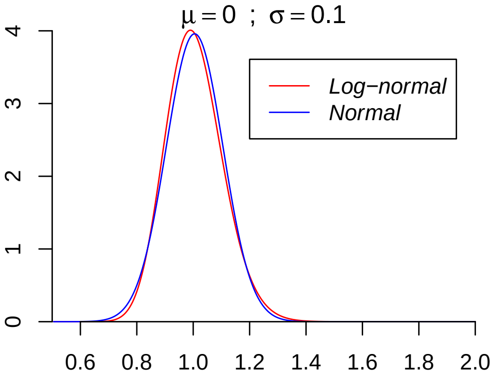
$$\begin{aligned} P(X > 300) &= P(e^{\sigma Z + \mu} > 300) \\ &= P(e^{0.5Z+5} > 300) \\ &= P(Z > (\ln(300) - 5)/0.5) \\ &\approx P(Z > 1.41) \\ &\approx .0793 \end{aligned}$$

Mean and Variance of Log-normal

If X is log-normal with parameters μ and σ, then

$$E(X) = e^{\mu+\sigma^2/2}$$
$$V(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

If σ is small, then X is approximately normal:



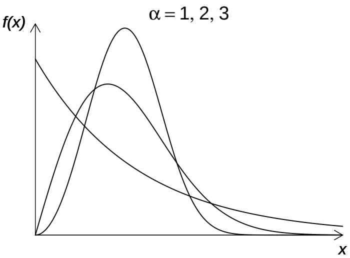
Weibull Distribution

A **Weibull** random variable X with shape $\alpha > 0$ and scale $\beta > 0$ has cdf

$$F(x) = \begin{cases} 1 - e^{-(x/\beta)^\alpha}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Its pdf is

$$f(x) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$



Example

The amount X of NO_x emission (g/gal) from a randomly selected engine of a certain type may be modeled as a Weibull random variable with $\alpha = 2$ and $\beta = 10$. What is the probability that a randomly selected engine has $X \geq 20$?

Solution: The cdf of X is

$$F(x) = P(X \leq x) = 1 - e^{-(x/\beta)^\alpha} = 1 - e^{-(x/10)^2}$$

Therefore,

$$\begin{aligned} P(X \geq 20) &= 1 - P(X \leq 20) = 1 - F(20) \\ &= e^{-(20/10)^2} = e^{-4} \approx .0183 \end{aligned}$$

Summary

Distribution	PDF	Mean	Variance
Uniform $a \leq x \leq b$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential $x \geq 0$	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
Gamma $x \geq 0$	$\frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}$	k/λ	k/λ^2
Normal $-\infty < x < \infty$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{x-\mu}{\sigma})^2/2}$	μ	σ^2
Weibull $x \geq 0$	$\frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha}$	$\beta\Gamma(\frac{\alpha+1}{\alpha})$	$\beta^2\Gamma(\frac{\alpha+2}{\alpha}) - \mu^2$
Log-normal $x \geq 0$	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-(\frac{\ln x-\mu}{\sigma})^2/2}$	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$