Probability theory

Prof.dr.hab. Viorel Bostan

Technical University of Moldova

viorel.bostan@adm.utm.md

Lecture 11

Continuous Probability

Consider now an experiment with a random outcome, such that the number of possible outcomes is not finite or infinitely countable, but rather uncountable, in other words, we have as many outcomes as real numbers.

Continuous Probability

Consider now an experiment with a random outcome, such that the number of possible outcomes is not finite or infinitely countable, but rather uncountable, in other words, we have as many outcomes as real numbers.

For example, consider the experiment of picking randomly a real number from interval $[0, 1]$, or from a unit square $[0, 1] \times [0, 1]$.

Continuous Probability

Consider now an experiment with a random outcome, such that the number of possible outcomes is not finite or infinitely countable, but rather uncountable, in other words, we have as many outcomes as real numbers.

For example, consider the experiment of picking randomly a real number from interval $[0, 1]$, or from a unit square $[0, 1] \times [0, 1]$.

That is our sample space is either $\Omega = [0, 1]$ or $\Omega = [0, 1] \times [0, 1]$.

Next, we need distribution function, function that assigns probability to all outcomes from the sample space.

Continuous Probability

Consider now an experiment with a random outcome, such that the number of possible outcomes is not finite or infinitely countable, but rather uncountable, in other words, we have as many outcomes as real numbers.

For example, consider the experiment of picking randomly a real number from interval $[0, 1]$, or from a unit square $[0, 1] \times [0, 1]$.

That is our sample space is either $\Omega = [0, 1]$ or $\Omega = [0, 1] \times [0, 1]$.

Next, we need distribution function, function that assigns probability to all outcomes from the sample space.

If we proceed further as we did in the discrete case, then we must assign the probability 0 to each outcome, since otherwise, the sum of the probabilities over all of the possible outcomes, would not equal 1.

Continuous Probability

Consider now an experiment with a random outcome, such that the number of possible outcomes is not finite or infinitely countable, but rather uncountable, in other words, we have as many outcomes as real numbers.

For example, consider the experiment of picking randomly a real number from interval $[0, 1]$, or from a unit square $[0, 1] \times [0, 1]$.

That is our sample space is either $\Omega = [0, 1]$ or $\Omega = [0, 1] \times [0, 1]$.

Next, we need distribution function, function that assigns probability to all outcomes from the sample space.

If we proceed further as we did in the discrete case, then we must assign the probability 0 to each outcome, since otherwise, the sum of the probabilities over all of the possible outcomes, would not equal 1.

However, if all of the assigned probabilities are 0, then the sum is 0, not 1, as it should be.

Continuous Probability

Consider now an experiment with a random outcome, such that the number of possible outcomes is not finite or infinitely countable, but rather uncountable, in other words, we have as many outcomes as real numbers.

For example, consider the experiment of picking randomly a real number from interval $[0, 1]$, or from a unit square $[0, 1] \times [0, 1]$.

That is our sample space is either $\Omega = [0, 1]$ or $\Omega = [0, 1] \times [0, 1]$.

Next, we need distribution function, function that assigns probability to all outcomes from the sample space.

If we proceed further as we did in the discrete case, then we must assign the probability 0 to each outcome, since otherwise, the sum of the probabilities over all of the possible outcomes, would not equal 1.

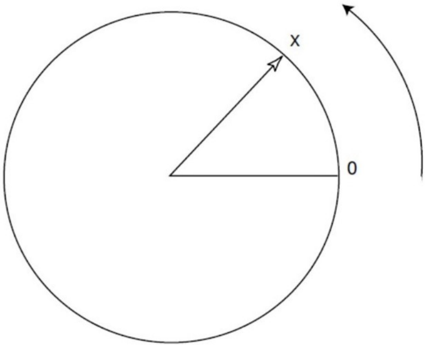
However, if all of the assigned probabilities are 0, then the sum is 0, not 1, as it should be.

Attention!

In continuous probability, there is no meaning to consider probability of individual outcome!

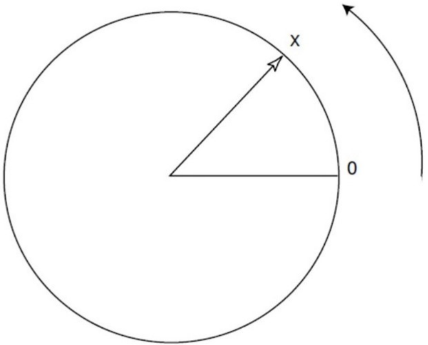
Continuous Probability. Spinner example

Consider a circle of unit length (i.e with radius $\frac{1}{2\pi}$):



Continuous Probability. Spinner example

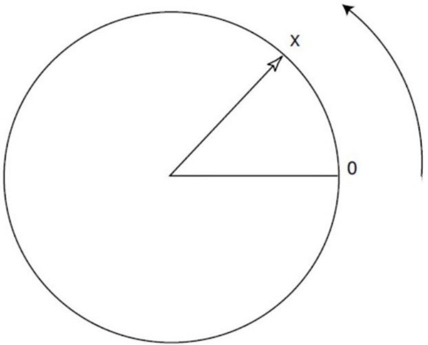
Consider a circle of unit length (i.e with radius $\frac{1}{2\pi}$):



What is probability $P(X < 0.5)$?

Continuous Probability. Spinner example

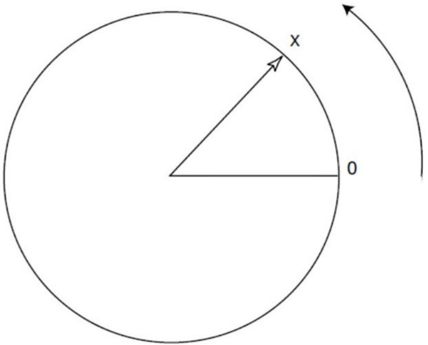
Consider a circle of unit length (i.e with radius $\frac{1}{2\pi}$):



What is probability $P(X < 0.5)$? Or $P(1/3 < X < 2/3)$?

Continuous Probability. Spinner example

Consider a circle of unit length (i.e with radius $\frac{1}{2\pi}$):

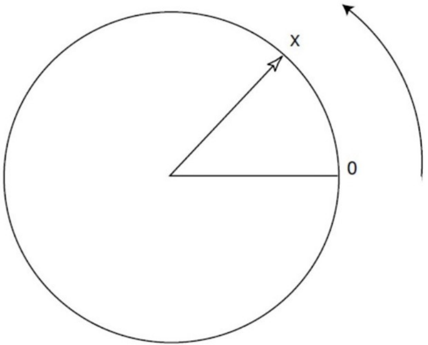


What is probability $P(X < 0.5)$? Or $P(1/3 < X < 2/3)$?

Obviously, $P(X = 0.5)$ does not make sense.

Continuous Probability. Spinner example

Consider a circle of unit length (i.e with radius $\frac{1}{2\pi}$):



What is probability $P(X < 0.5)$? Or $P(1/3 < X < 2/3)$?

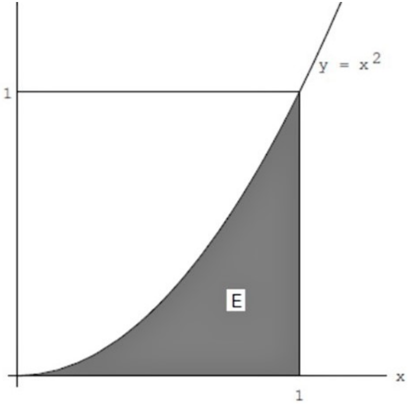
Obviously, $P(X = 0.5)$ does not make sense.

We rather ask what is $P(0.5 - 0.01 < X < 0.5 + 0.01)$?

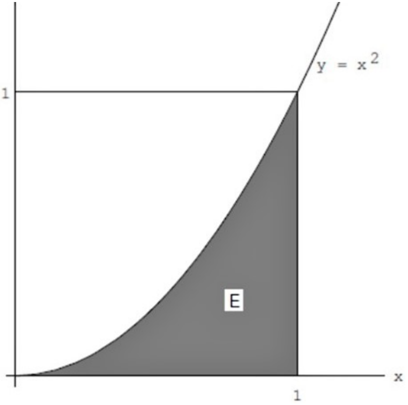
Monte Carlo Method

Suppose we generate a pair (x, y), s.t. $0 \leqslant x, y \leqslant 1$.

Suppose we generate a pair (x, y), s.t. $0 \leqslant x, y \leqslant 1$.

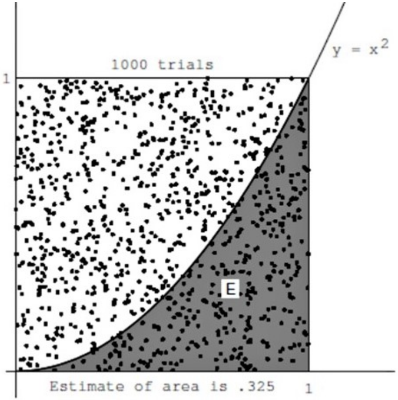


Suppose we generate a pair (x, y), s.t. $0 \leqslant x, y \leqslant 1$.

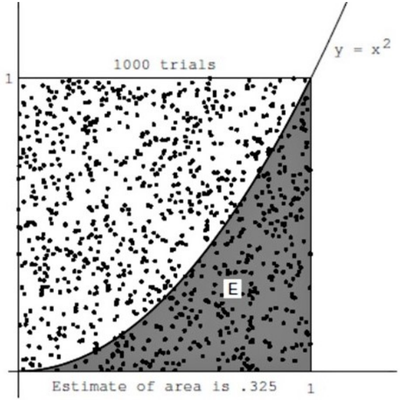


And ask what is the area of the region E.

Simulate 1000 points (x, y) in unit square and get, say, 325 points in E.



Simulate 1000 points (x, y) in unit square and get, say, 325 points in E.



$$\frac{325}{1000} = 0.325 \approx \text{Area of } E = \int_0^1 x^2\,dx = \frac{1}{3}.$$

Sum example (2 variables)

Let us choose two random real numbers in $[0,1]$ and add them together.

Sum example (2 variables)

Let us choose two random real numbers in $[0, 1]$ and add them together.

$X \in [0, 1]$ and $Y \in [0, 1]$ are two **continuous random variables** (CRV).

Sum example (2 variables)

Let us choose two random real numbers in $[0, 1]$ and add them together.

$X \in [0, 1]$ and $Y \in [0, 1]$ are two **continuous random variables** (CRV).

Let Z be the sum: Z=X+Y. It is another CRV. Clearly Z can take values from 0 to 2.

Sum example (2 variables)

Let us choose two random real numbers in $[0, 1]$ and add them together.

$X \in [0, 1]$ and $Y \in [0, 1]$ are two **continuous random variables** (CRV).

Let Z be the sum: Z=X+Y. It is another CRV. Clearly Z can take values from 0 to 2.

How is Z distributed over $[0, 2]$?

Sum example (2 variables)

Let us choose two random real numbers in $[0, 1]$ and add them together.

$X \in [0, 1]$ and $Y \in [0, 1]$ are two **continuous random variables** (CRV).

Let Z be the sum: Z=X+Y. It is another CRV. Clearly Z can take values from 0 to 2.

How is Z distributed over $[0, 2]$?

Write a program that simulates this experiment many times (for example, 10000 times).

Sum example (2 variables)

Let us choose two random real numbers in $[0, 1]$ and add them together.

$X \in [0, 1]$ and $Y \in [0, 1]$ are two **continuous random variables** (CRV).

Let Z be the sum: Z=X+Y. It is another CRV. Clearly Z can take values from 0 to 2.

How is Z distributed over $[0, 2]$?

Write a program that simulates this experiment many times (for example, 10000 times).

Divide interval $[0, 2]$ into, for example, 20 equal sub-intervals:

$$0.0 < 0.1 < 0.2 < 0.3 < \ldots < 1.8 < 1.9 < 2.0$$

Sum example (2 variables)

Let us choose two random real numbers in $[0, 1]$ and add them together.

$X \in [0, 1]$ and $Y \in [0, 1]$ are two **continuous random variables** (CRV).

Let Z be the sum: Z=X+Y. It is another CRV. Clearly Z can take values from 0 to 2.

How is Z distributed over $[0, 2]$?

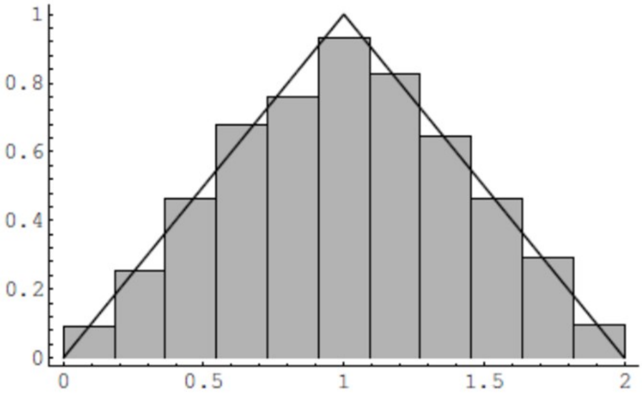
Write a program that simulates this experiment many times (for example, 10000 times).

Divide interval $[0, 2]$ into, for example, 20 equal sub-intervals:

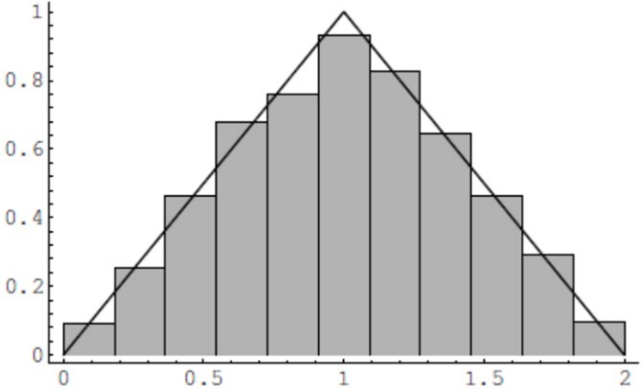
$$0.0 < 0.1 < 0.2 < 0.3 < \ldots < 1.8 < 1.9 < 2.0$$

Produce a bar graph with the property that on each interval, the area of the bar (rectangle) is equal to the fraction of outcomes that fell in the corresponding interval.

Sum example (2 variables)



Sum example (2 variables)

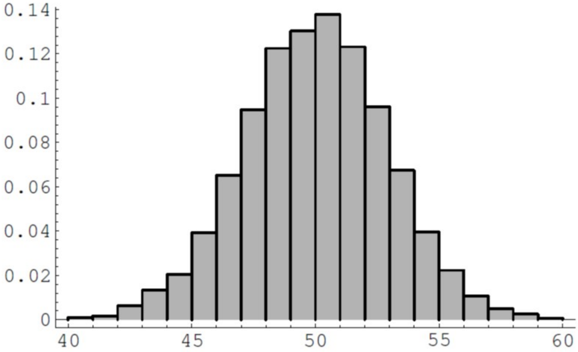


The bar diagram approximates the function:

$$f(t) = \begin{cases} t, & \text{if } 0 \leqslant t < 1, \\ 2 - t, & \text{if } 1 \leqslant t < 2. \end{cases}$$

Repeat the experiment with summing 100 variables:

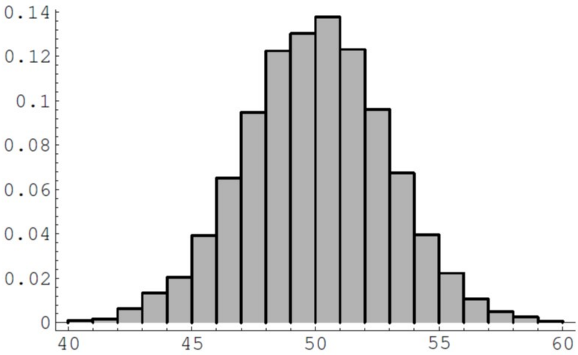
$$Z = X_1 + X_2 + \ldots + X_{99} + X_{100}, \quad \text{with } X_i \in [0, 1].$$



Sum example (100 variables)

Repeat the experiment with summing 100 variables:

$$Z = X_1 + X_2 + \ldots + X_{99} + X_{100}, \quad \text{with } X_i \in [0,1].$$



Bar diagram approximates "bell shape" function, like in binomial distribution from discrete probability case (we will call it later, the normal distribution).

Continuous density functions

Definition

Let X be a real-valued continuous random variable (CRV).
A **density function** for X is a real-valued function f such that

$$P(a \leqslant X < b) = \int_a^b f(t)\, dt.$$

If $E \subset \mathbb{R}$ then

$$P(X \in E) = \int_E f(t)\, dt.$$

Continuous density functions

Definition

Let X be a real-valued continuous random variable (CRV).
A **density function** for X is a real-valued function f such that

$$P(a \leqslant X < b) = \int_a^b f(t)\,dt.$$

If $E \subset \mathbb{R}$ then

$$P(X \in E) = \int_E f(t)\,dt.$$

For example:

$$P(2 < X < 5) = \int_2^5 f(t)\,dt, \qquad P(X \geqslant 7) = \int_7^{+\infty} f(t)\,dt.$$

Continuous density functions

Definition

Let X be a real-valued continuous random variable (CRV).
A **density function** for X is a real-valued function f such that

$$P(a \leqslant X < b) = \int_a^b f(t)\, dt.$$

If $E \subset \mathbb{R}$ then

$$P(X \in E) = \int_E f(t)\, dt.$$

For example:

$$P(2 < X < 5) = \int_2^5 f(t)\, dt, \qquad P(X \geqslant 7) = \int_7^{+\infty} f(t)\, dt.$$

Obviously, a density function $f(t)$ for random variable X must satisfy:

$$\int_{-\infty}^{+\infty} f(t)\, dt = 1.$$

Cumulative distribution function of CRV

Definition

Let X be a CRV. Then, the **cumulative distribution function** of X is defined by

$$F_X(t) = P(X \leqslant t) = \int_{-\infty}^{t} f(u)\, du.$$

Definition

Let X be a CRV. Then, the **cumulative distribution function** of X is defined by

$$F_X(t) = P(X \leqslant t) = \int_{-\infty}^{t} f(u)\, du.$$

If X is a real-valued CRV which possesses a density function $f(t)$, then it also has a cumulative distribution function $F_X(t)$, and the following theorem shows that the two functions are related:

Cumulative distribution function of CRV

Definition

Let X be a CRV. Then, the **cumulative distribution function** of X is defined by

$$F_X(t) = P(X \leqslant t) = \int_{-\infty}^{t} f(u)\, du.$$

If X is a real-valued CRV which possesses a density function $f(t)$, then it also has a cumulative distribution function $F_X(t)$, and the following theorem shows that the two functions are related:

Theorem

Let X be a real-valued continuous random variable (CRV) with density function $f(t)$. Then

$$\frac{d}{dt} F_X(t) = f(t).$$

In other words, derivative of cumulative distribution function is the density function.

Example

A real number U is chosen at random from $[0,1)$ and then it is squared.
Let $X = U^2$ represents the result.

Example

A real number U is chosen at random from $[0, 1)$ and then it is squared.
Let $X = U^2$ represents the result.

If $0 \leqslant t < 1$, then

$$\begin{aligned} F_X(t) &= P(X \leqslant t) \\ &= P(U^2 \leqslant t) \\ &= P(U \leqslant \sqrt{t}) \\ &= \sqrt{t}. \end{aligned}$$

Example

A real number U is chosen at random from $[0, 1)$ and then it is squared.
Let $X = U^2$ represents the result.

If $0 \leqslant t < 1$, then

$$\begin{aligned} F_X(t) &= P(X \leqslant t) \\ &= P(U^2 \leqslant t) \\ &= P(U \leqslant \sqrt{t}) \\ &= \sqrt{t}. \end{aligned}$$

If $t < 0$, then $P(X < t) = 0$ (since X is always non-negative).

Example

A real number U is chosen at random from $[0,1)$ and then it is squared.
Let $X = U^2$ represents the result.

If $0 \leqslant t < 1$, then

$$\begin{aligned} F_X(t) &= P(X \leqslant t) \\ &= P(U^2 \leqslant t) \\ &= P(U \leqslant \sqrt{t}) \\ &= \sqrt{t}. \end{aligned}$$

If $t < 0$, then $P(X < t) = 0$ (since X is always non-negative).
If $t \geqslant 1$, then $P(X < t) = 1$ (since X is always less or equal to 1).

Example

A real number U is chosen at random from $[0, 1)$ and then it is squared.
Let $X = U^2$ represents the result.

If $0 \leqslant t < 1$, then

$$\begin{aligned} F_X(t) &= P(X \leqslant t) \\ &= P(U^2 \leqslant t) \\ &= P(U \leqslant \sqrt{t}) \\ &= \sqrt{t}. \end{aligned}$$

If $t < 0$, then $P(X < t) = 0$ (since X is always non-negative).
If $t \geqslant 1$, then $P(X < t) = 1$ (since X is always less or equal to 1).
Thus, cumulative distribution of CRV X is

$$F_X(t) = \begin{cases} 0, & \text{if } t < 0, \\ \sqrt{t}, & \text{if } 0 \leqslant t < 1, \\ 1, & \text{if } t \geqslant 1. \end{cases}$$

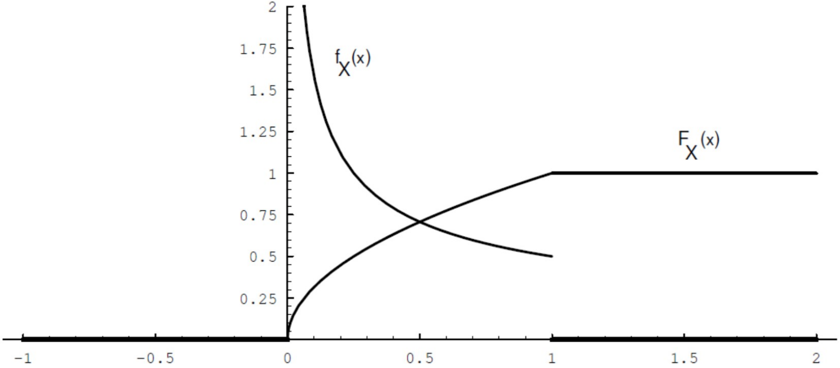
Example

Differentiate $F(t)$ to obtain density function $f(t)$:

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

Differentiate $F(t)$ to obtain density function $f(t)$:

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



Example

Let X and Y be CRV from $[0, 1)$ and let $Z = X + Y$. Clearly $Z \in [0, 2]$.

Example

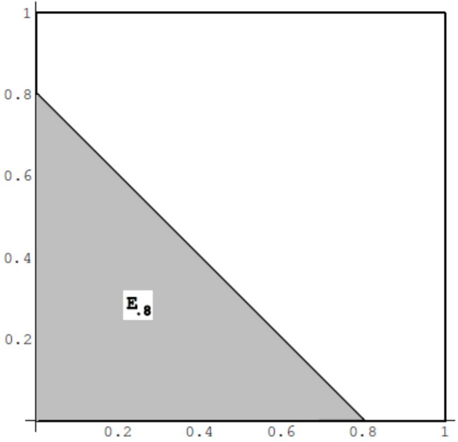
Let X and Y be CRV from $[0, 1)$ and let $Z = X + Y$. Clearly $Z \in [0, 2]$.
Sample space Ω will be unit square in $\mathbb{R}^2$.

Example

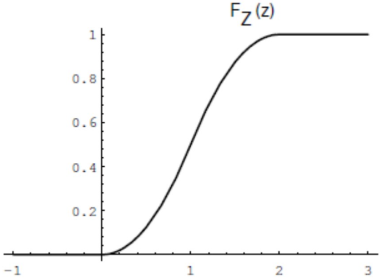
Let X and Y be CRV from $[0, 1)$ and let $Z = X + Y$. Clearly $Z \in [0, 2]$.
Sample space Ω will be unit square in $\mathbb{R}^2$.
Let E_t denote the event that $Z \leqslant t$. For example, below is $E_{0.8}$

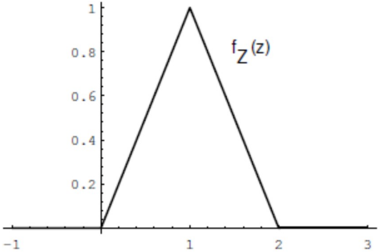
Example

Let X and Y be CRV from $[0, 1)$ and let $Z = X + Y$. Clearly $Z \in [0, 2]$.
Sample space Ω will be unit square in $\mathbb{R}^2$.
Let E_t denote the event that $Z \leqslant t$. For example, below is $E_{0.8}$



$$F_Z(t) = P(Z \leqslant t) = \text{Area of } E_t.$$





$$F_Z(t) = \begin{cases} 0, & t < 0, \\ \frac{t^2}{2}, & 0 \leqslant t < 1, \\ 1 - \frac{(2-t)^2}{2}, & 1 \leqslant t < 2, \\ 1, & t \geqslant 2. \end{cases} \qquad f_Z(t) = \begin{cases} 0, & t < 0, \\ t, & 0 \leqslant t < 1, \\ 2-t, & 1 \leqslant t < 2, \\ 1, & t \geqslant 2. \end{cases}$$

Continuous Uniform Density

The simplest density function corresponds to the random variable X whose value represents the outcome of the experiment consisting of choosing a real number at random from the interval $[a, b]$.

Continuous Uniform Density

The simplest density function corresponds to the random variable X whose value represents the outcome of the experiment consisting of choosing a real number at random from the interval $[a, b]$.

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

$$F_X(t) = P(X < t) = \begin{cases} 0, & t < a \\ \frac{t-a}{b-a}, & a \leqslant t < b \\ 1, & t \geqslant b. \end{cases}$$

It is easy to simulate on a computer a random variable with this density:

$$X = (b-a) \cdot rnd + a.$$

Exponential Density

The exponential density function is defined by

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{if} \quad 0 \leqslant t < +\infty \\ 0, & \text{otherwise.} \end{cases}$$

Exponential Density

The exponential density function is defined by

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{if} \quad 0 \leqslant t < +\infty \\ 0, & \text{otherwise.} \end{cases}$$

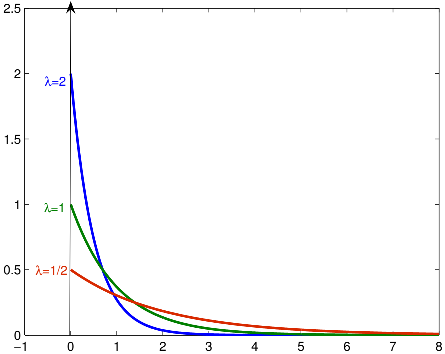
Here λ is a positive constant depending on the experiment.

Exponential Density

The exponential density function is defined by

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{if} \quad 0 \leqslant t < +\infty \\ 0, & \text{otherwise.} \end{cases}$$

Here λ is a positive constant depending on the experiment.
Graphs of exponential densities with different choices of λ:



Exponential Density

- The exponential density is often used to describe experiments involving a question of the form: How long until something happens?

Exponential Density

- The exponential density is often used to describe experiments involving a question of the form: How long until something happens?
- For example, the exponential density is often used to study the time between emissions of particles from a radioactive source.

- The exponential density is often used to describe experiments involving a question of the form: How long until something happens?
- For example, the exponential density is often used to study the time between emissions of particles from a radioactive source.
- The cumulative distribution function of the exponential density is easy to compute. Let T be an exponentially distributed random variable with parameter. If $t \geqslant 0$, then we have

$$F_T(t) = P(T \leqslant t) = \int_{-\infty}^{t} \lambda e^{-\lambda u}\, du = 1 - e^{-\lambda t}.$$

Therefore,

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

Exponential Density

- The exponential density is often used to describe experiments involving a question of the form: How long until something happens?
- For example, the exponential density is often used to study the time between emissions of particles from a radioactive source.
- The cumulative distribution function of the exponential density is easy to compute. Let T be an exponentially distributed random variable with parameter. If $t \geqslant 0$, then we have

$$F_T(t) = P(T \leqslant t) = \int_{-\infty}^{t} \lambda e^{-\lambda u}\, du = 1 - e^{-\lambda t}.$$

Therefore,

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

- Both the exponential density and the geometric distribution share a “memory-less” property, that will be discussed further.

Continuous Conditional Probability

- When the sample space Ω is not countable, we define the conditional probability in a similar manner.

Continuous Conditional Probability

- When the sample space Ω is not countable, we define the conditional probability in a similar manner.
- For example, if X is a continuous random variable with density function $f(t)$, and if E is an event with positive probability, we define a **conditional density function** by the formula:

$$f(t \mid E) = \begin{cases} \dfrac{f(t)}{P(E)}, & \textit{if} \quad t \in E, \\ 0, & \textit{if} \quad t \notin E. \end{cases}$$

Continuous Conditional Probability

- When the sample space Ω is not countable, we define the conditional probability in a similar manner.
- For example, if X is a continuous random variable with density function $f(t)$, and if E is an event with positive probability, we define a **conditional density function** by the formula:

$$f(t \mid E) = \begin{cases} \dfrac{f(t)}{P(E)}, & \textit{if} \quad t \in E, \\ 0, & \textit{if} \quad t \notin E. \end{cases}$$

- Then, for any event F, we have

$$P(F \mid E) = \int_F f(t \mid E)\, dt.$$

Continuous Conditional Probability

- When the sample space Ω is not countable, we define the conditional probability in a similar manner.
- For example, if X is a continuous random variable with density function $f(t)$, and if E is an event with positive probability, we define a **conditional density function** by the formula:

$$f(t \mid E) = \begin{cases} \dfrac{f(t)}{P(E)}, & \textit{if} \quad t \in E, \\ 0, & \textit{if} \quad t \notin E. \end{cases}$$

- Then, for any event F, we have

$$P(F \mid E) = \int_F f(t \mid E)\, dt.$$

- The expression $P(F \mid E)$ is called the **conditional probability** of event F given event E.

Continuous Conditional Probability

- It is easy to obtain an alternative expression for this probability:

$$
\begin{aligned}
P(F \mid E) &= \int_F f(t \mid E)\, dt \\
&= \int_{F \cap E} \frac{f(t)}{P(E)}\, dt \\
&= \frac{1}{P(E)} \int_{F \cap E} f(t)\, dt \\
&= \frac{P(F \cap E)}{P(E)}.
\end{aligned}
$$

Continuous Conditional Probability

- It is easy to obtain an alternative expression for this probability:

$$\begin{aligned} P(F \mid E) &= \int_F f(t \mid E)\, dt \\ &= \int_{F \cap E} \frac{f(t)}{P(E)}\, dt \\ &= \frac{1}{P(E)} \int_{F \cap E} f(t)\, dt \\ &= \frac{P(F \cap E)}{P(E)}. \end{aligned}$$

- So, we have the same rule as in discrete case:

$$P(F \mid E) = \frac{P(F \cap E)}{P(E)}.$$

Continuous Conditional Probability

Example. Pick at random a real number in $[2,9)$. Let X be its value.
Compute $P(X < 3 \mid X \leqslant 6)$.
Solution. Recall the density function for X is

$$f(t) = \begin{cases} \frac{1}{7}, & \text{if} \quad t \in [2,9), \\ 0, & \text{otherwise.} \end{cases}$$

Then, if $E = "X \leqslant 6"$, we have conditional density function:

$$f(t \mid E) = \begin{cases} \frac{\frac{1}{7}}{\frac{4}{7}} = \frac{1}{4}, & \text{if} \quad t \in [2,6], \\ 0, & \text{otherwise.} \end{cases}$$

Also, conditional probability can be computed direct form the formula:

$$P(X < 3 \mid X \leqslant 6) = \frac{P((X < 3) \cap (X \leqslant 6))}{P(X \leqslant 6)} = \frac{P(X < 3)}{P(X \leqslant 6)} = \frac{\frac{1}{7}}{\frac{4}{7}} = \frac{1}{4}.$$

Radioactive example

- Suppose that we are observing a lump of plutonium-239. Experiment consists of waiting for an emission, then starting a clock, and recording the length of time X that passes until the next emission.

Radioactive example

- Suppose that we are observing a lump of plutonium-239. Experiment consists of waiting for an emission, then starting a clock, and recording the length of time X that passes until the next emission.
- Experience has shown that X has an exponential density with some parameter, which depends upon the size of the lump.

Radioactive example

- Suppose that we are observing a lump of plutonium-239. Experiment consists of waiting for an emission, then starting a clock, and recording the length of time X that passes until the next emission.
- Experience has shown that X has an exponential density with some parameter, which depends upon the size of the lump.
- Suppose that when we perform this experiment, we notice that the clock reads r seconds, and is still running.

Radioactive example

- Suppose that we are observing a lump of plutonium-239. Experiment consists of waiting for an emission, then starting a clock, and recording the length of time X that passes until the next emission.
- Experience has shown that X has an exponential density with some parameter, which depends upon the size of the lump.
- Suppose that when we perform this experiment, we notice that the clock reads r seconds, and is still running.
- What is the probability that there is no emission in a further s seconds?

Radioactive example

- Suppose that we are observing a lump of plutonium-239. Experiment consists of waiting for an emission, then starting a clock, and recording the length of time X that passes until the next emission.
- Experience has shown that X has an exponential density with some parameter, which depends upon the size of the lump.
- Suppose that when we perform this experiment, we notice that the clock reads r seconds, and is still running.
- What is the probability that there is no emission in a further s seconds?
- Let $G(t)$ be the probability that the next particle is emitted after time t. Then,

$$\begin{aligned} G(t) &= P(X \geqslant t) = P(t \geqslant X < +\infty) \\ &= \int_t^{+\infty} f(u)\,du = \int_t^{+\infty} \lambda e^{-\lambda u}\,du = e^{-\lambda t}. \end{aligned}$$

Radioactive example

Let E be the event **"the next particle is emitted after time r"** and F the event **"the next particle is emitted after time $r+s$ "**. Then,

$$\begin{aligned} P(F \mid E) &= \frac{P(F \cap E)}{P(E)} \\ &= \frac{G(r+s)}{G(r)} \\ &= \frac{e^{-\lambda(r+s)}}{e^{-\lambda r}} \\ &= e^{-\lambda s}. \end{aligned}$$

Let E be the event **"the next particle is emitted after time r"** and F the event **"the next particle is emitted after time $r + s$ "**. Then,

$$\begin{aligned} P(F \mid E) &= \frac{P(F \cap E)}{P(E)} \\ &= \frac{G(r+s)}{G(r)} \\ &= \frac{e^{-\lambda(r+s)}}{e^{-\lambda r}} \\ &= e^{-\lambda s}. \end{aligned}$$

In other words, we just proved that

$$P(T \geqslant r + s \mid T \geqslant r) = P(T \geqslant s).$$

- Therefore, the probability that we have to wait s seconds more for an emission, given that there has been no emission in r seconds, is independent of the time r.

Exponential Density

- Therefore, the probability that we have to wait s seconds more for an emission, given that there has been no emission in r seconds, is independent of the time r.
- This property is called the memory-less property .

Exponential Density

- Therefore, the probability that we have to wait s seconds more for an emission, given that there has been no emission in r seconds, is independent of the time r.
- This property is called the memory-less property .
- When trying to model various phenomena, this property is helpful in deciding whether the exponential density is appropriate.

Exponential Density

- Therefore, the probability that we have to wait s seconds more for an emission, given that there has been no emission in r seconds, is independent of the time r.
- This property is called the memory-less property .
- When trying to model various phenomena, this property is helpful in deciding whether the exponential density is appropriate.
- The fact that the exponential density is memory-less means that it is reasonable to assume if one comes upon a lump of a radioactive isotope at some random time, then the amount of time until the next emission has an exponential density with the same parameter as the time between emissions.

Exponential Density

- A well-known example, known as the **“bus paradox”** replaces the emissions by buses:

Exponential Density

- A well-known example, known as the **“bus paradox”** replaces the emissions by buses:
- ”If you know that, on the average, the buses come by every 30 minutes and the buses arrival times are being modeled by the exponential density, then no matter when you arrive, you will have to wait, on the average, for 30 minutes for a bus.

How to simulate a random variable with exponential density?

Exponential Density

How to simulate a random variable with exponential density?
The following random variable does the job:

$$Y = -\frac{1}{\lambda}\ln(rnd).$$

Exponential Density

How to simulate a random variable with exponential density?
The following random variable does the job:

$$Y = -\frac{1}{\lambda}\ln(rnd).$$

The following arguments show that the random variable Y has the required property:

Exponential Density

How to simulate a random variable with exponential density?
The following random variable does the job:

$$Y = -\frac{1}{\lambda}\ln(rnd).$$

The following arguments show that the random variable Y has the required property:
We have

$$\begin{aligned} P(Y \leqslant t) =& P(-\frac{1}{\lambda}\ln(rnd) \leqslant t) \\ =& P(\ln(rnd) \geqslant -\lambda t) \\ =& P(rnd \geqslant e^{-\lambda t}) \\ =& 1 - e^{-\lambda t}. \end{aligned}$$

Exponential Density

How to simulate a random variable with exponential density?
The following random variable does the job:

$$Y = -\frac{1}{\lambda}\ln(rnd).$$

The following arguments show that the random variable Y has the required property:
We have

$$\begin{aligned} P(Y \leqslant t) =& P(-\frac{1}{\lambda}\ln(rnd) \leqslant t) \\ =& P(\ln(rnd) \geqslant -\lambda t) \\ =& P(rnd \geqslant e^{-\lambda t}) \\ =& 1 - e^{-\lambda t}. \end{aligned}$$

This last expression is seen to be the cumulative distribution function of an exponentially distributed random variable with parameter λ.

Independent Events

If E and F are two events with positive probability in a continuous sample space, then, as in the case of discrete sample spaces we have:

Definition

Events E and F are said to be independent if

$$P(E \mid F) = P(E) \quad \textit{and} \quad P(F \mid E) = P(F).$$

Independent Events

If E and F are two events with positive probability in a continuous sample space, then, as in the case of discrete sample spaces we have:

Definition

Events E and F are said to be independent if

$$P(E \mid F) = P(E) \quad \textit{and} \quad P(F \mid E) = P(F).$$

As in discrete case, each of the above equations imply the other
(i.e. no need to check both!)

Independent Events

If E and F are two events with positive probability in a continuous sample space, then, as in the case of discrete sample spaces we have:

Definition

Events E and F are said to be independent if

$$P(E \mid F) = P(E) \quad \textit{and} \quad P(F \mid E) = P(F).$$

As in discrete case, each of the above equations imply the other
(i.e. no need to check both!)

Also, if E and F are independent, then

$$P(E \cap F) = P(E)P(F).$$

Definition

Let $X_1, X_2, \ldots, X_n$ be continuous random variables associated with an experiment, and let $X = (X_1, X_2, \ldots, X_n)$. Then the joint cumulative distribution function of X is defined by

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n).$$

If $f(x_1, x_2, \ldots, x_n)$ is the joint density function then it can be shown that

$$F(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(t_1, t_2, \ldots, t_n)\, dt_1 dt_2 \ldots dt_n.$$

and

$$f(x_1, x_2, \ldots, x_n) = \frac{\partial^n F(x_1, x_2, \ldots, x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}.$$

Definition

Let $X_1, X_2, \ldots, X_n$ be continuous random variables associated with an experiment, and let $F_1(x), F_2(x), \ldots, F_n(x)$ be their cumulative distribution functions. Then these random variables are mutually independent if

$$F(x_1, x_2, \ldots, x_n) = F_1(x_1)F_2(x_2)\ldots F_n(x_n).$$

for any choice of $x_1, x_2, \ldots, x_n$.

In other words, if $X_1, X_2, \ldots, X_n$ are mutually independent, then the cumulative distribution function of the joint random variable $X = (X_1, X_2, \ldots, X_n)$ is the product of individual cumulative distribution functions.

Theorem

Let $X_1, X_2, \ldots, X_n$ be continuous random variables with density functions $f_1(x), f_2(x), \ldots, , f_n(x)$, respectively. Then, these random variables are mutually independent if and only if

$$f(x_1, x_2, ..., x_n) = f_1(x_1) f_2(x_2) \ldots \; f_n(x_n),$$

for any choice of $x_1, x_2, \ldots, x_n$.

Example.
Choose a point $\omega = (\omega_1, \omega_2)$ at random from the unit square.
Set $X_1 = \omega_1^2$, $X_2 = \omega_2^2$, and $X_3 = \omega_1 + \omega_2$.

Find the joint cumulative distribution $F_{12}(r_1, r_2)$ of $X_{12} = (X_1, X_2)$ and joint cumulative distribution $F_{23}(r_2, r_3)$ of $X_{23} = (X_2, X_3)$.

Solution.
First, need the cumulative distribution functions of X_1, X_2, and X_3.

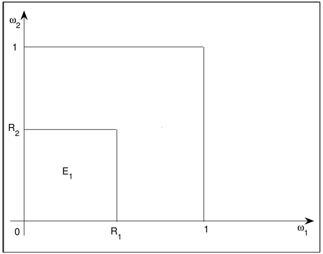
$$\begin{aligned} F_1(r_1) &= P(X_1 \leqslant r_1) = P(\omega_1^2 \leqslant r_1) \\ &= P(\omega_1 \leqslant \sqrt{r_1}) \\ &= \sqrt{r_1}, \quad \text{if } 0 \leqslant r_1 < 1. \end{aligned}$$

Similarly,

$$F_2(r_2) = \sqrt{r_2}, \quad \text{if } 0 \leqslant r_2 < 1.$$

Previously, we have shown that

$$F_3(r_3) = \begin{cases} 0, & \textit{if} \quad r_3 \leqslant 0, \\ \frac{1}{2} r_3^2 & \textit{if} \quad 0 \leqslant r_3 < 1, \\ 1 - \frac{1}{2}(2 - r_3)^2 & \textit{if} \quad 1 \leqslant r_3 < 2, \\ 1, & \textit{if} \quad 2 \leqslant r_3. \end{cases}$$



Let $R_1 = \sqrt{r_1}$ and $R_2 = \sqrt{r_2}$. Then,

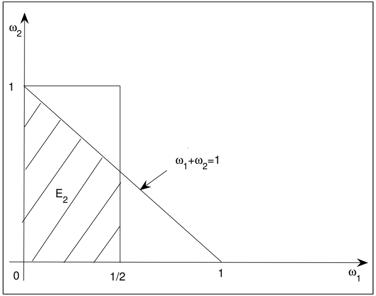
$$\begin{aligned} F_{12}(r_1, r_2) &= P(X_1 \leqslant r_1 \ \text{ and } \ X_2 \leqslant r_2) \\ &= P(\omega_1 \leqslant \sqrt{r_1} \ \text{ and } \ \omega_2 \leqslant \sqrt{r_2}) \\ &= P(\omega_1 \leqslant R_1 \ \text{ and } \ \omega_2 \leqslant R_2) \\ &= Area(E_1) = R_1 R_2 = \sqrt{r_1}\sqrt{r_2} = F_1(r_1)F_2(r_2). \end{aligned}$$

So, X_1 and X_2 are independent.

Example (contd)

On the other hand, if $r_1 = 1/4$ and $r_3 = 1$, then

$$\begin{aligned} F_{13}(\tfrac{1}{4}, 1) &= P(X_1 \leqslant \tfrac{1}{4} \ \text{ and } \ X_3 \leqslant 1) = P(\omega_1 \leqslant \tfrac{1}{2} \ \text{ and } \ \omega_1 + \omega_2 \leqslant 1) \\ &= Area(E_2) = \tfrac{1}{2} - \tfrac{1}{8} = \tfrac{3}{8}. \end{aligned}$$



So, $F_{13}(\frac{1}{4}, 1) = \frac{3}{8} \neq F_1(\frac{1}{4}) F_3(1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.
Therefore, the random variables X_1 and X_3 are not independent.

Independent Random Variables

Theorem

Let $X_1, X_2, \ldots, X_n$ be mutually independent continuous random variables and let $\varphi_1(x), \varphi_2(x), \ldots, , \varphi_n(x)$ be continuous functions. Then $\varphi_1(X_1)$, $\varphi_2(X_2)$, $\ldots$, $\varphi_n(X_n)$, are mutually independent variables.

For example, if X_1, X_2, X_3 are mutually independent continuous random variables, then so are Y_1, Y_2, Y_3, where

$$\begin{aligned} Y_1 &= X_1^2, \\ Y_2 &= \sin X_2, \\ Y_3 &= X_3^4 - 5X_3^3 + 9X_3 - 7. \end{aligned}$$

Queues Example

Suppose that customers arrive at random times at a service station with one server and that each customer is served immediately, if no one is ahead of him, but must wait his turn in line otherwise.

Queues Example

Suppose that customers arrive at random times at a service station with one server and that each customer is served immediately, if no one is ahead of him, but must wait his turn in line otherwise.

How long should each customer expect to wait?

Queues Example

Suppose that customers arrive at random times at a service station with one server and that each customer is served immediately, if no one is ahead of him, but must wait his turn in line otherwise.

How long should each customer expect to wait?

The waiting time of a customer is the length of time between the time that he arrives and the time that he begins to be served.

Queues Example

Suppose that customers arrive at random times at a service station with one server and that each customer is served immediately, if no one is ahead of him, but must wait his turn in line otherwise.

How long should each customer expect to wait?

The waiting time of a customer is the length of time between the time that he arrives and the time that he begins to be served.

Assume that inter-arrival times between successive customers are given by random variables $X_1, X_2, \ldots, X_n$ that are **mutually independent** and exponentially distributed with cumulative distribution function:

$$F_X(t) = 1 - e^{-\lambda t}.$$

Queues Example

Suppose that customers arrive at random times at a service station with one server and that each customer is served immediately, if no one is ahead of him, but must wait his turn in line otherwise.

How long should each customer expect to wait?

The waiting time of a customer is the length of time between the time that he arrives and the time that he begins to be served.

Assume that inter-arrival times between successive customers are given by random variables $X_1, X_2, \ldots, X_n$ that are **mutually independent** and exponentially distributed with cumulative distribution function:

$$F_X(t) = 1 - e^{-\lambda t}.$$

Also, assume that the service times for successive customers are given by **mutually independent** random variables $Y_1, Y_2, \ldots, Y_n$ distributed with another exponential cumulative distribution function:

$$F_Y(t) = 1 - e^{-\mu t}.$$

Queues Example

The parameters λ and μ represent, respectively, the reciprocals of the average time between arrivals of customers and the average service time of the customers.

Queues Example

The parameters λ and μ represent, respectively, the reciprocals of the average time between arrivals of customers and the average service time of the customers.

Thus, for example, the larger the value of λ, the smaller the average time between arrivals of customers.

Queues Example

The parameters λ and μ represent, respectively, the reciprocals of the average time between arrivals of customers and the average service time of the customers.

Thus, for example, the larger the value of λ, the smaller the average time between arrivals of customers.

We can guess that the length of time a customer will spend in the queue depends on the relative sizes of the average inter-arrival time and the average service time.

Queues Example

The parameters λ and μ represent, respectively, the reciprocals of the average time between arrivals of customers and the average service time of the customers.

Thus, for example, the larger the value of λ, the smaller the average time between arrivals of customers.

We can guess that the length of time a customer will spend in the queue depends on the relative sizes of the average inter-arrival time and the average service time.

You can simulate this queuing process on computer.

Queues Example

The parameters λ and μ represent, respectively, the reciprocals of the average time between arrivals of customers and the average service time of the customers.

Thus, for example, the larger the value of λ, the smaller the average time between arrivals of customers.

We can guess that the length of time a customer will spend in the queue depends on the relative sizes of the average inter-arrival time and the average service time.

You can simulate this queuing process on computer.

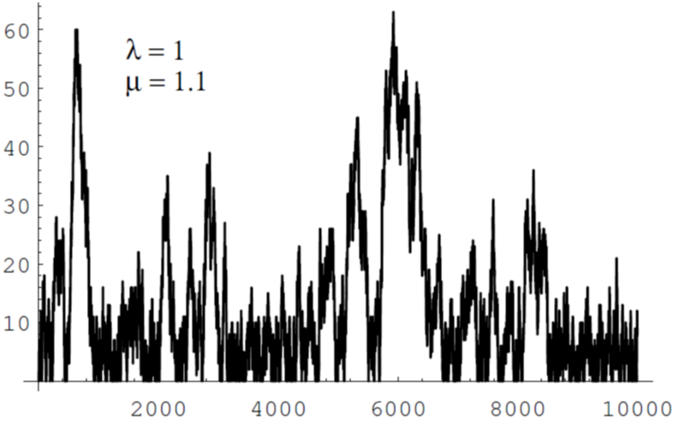
Generate random variables $X_1, X_2, \ldots$ with exponential density with parameter λ and random variables $Y_1, Y_2, \ldots$ with exponential density with parameter μ.

Queues Example

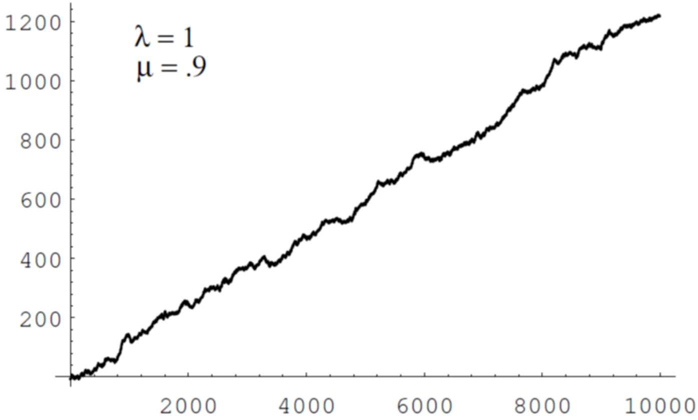
Let $N(t)$ be the number of customers in the queue at time t.

Queues Example

Let $N(t)$ be the number of customers in the queue at time t.
If $\lambda < \mu$, then the number of customers in the line will be finitely bounded:



If $\lambda > \mu$, then the number of customers in the line will be infinitely increasing:



Queues Example

How long will a customer have to wait in the queue for service?

Queues Example

How long will a customer have to wait in the queue for service?

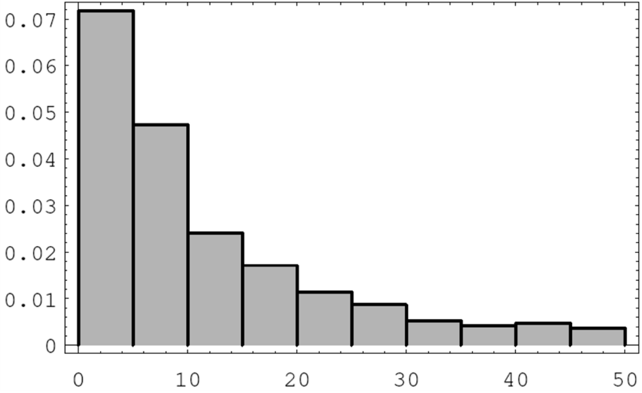
Let W_i be the length of time that the $i-$th customer has to wait in line.

Queues Example

How long will a customer have to wait in the queue for service?

Let W_i be the length of time that the $i-$th customer has to wait in line.

The bar diagram shows how W_i are distributed with $\lambda = 1$ and $\mu = 1.1$.



Functions of a Random Variable

Consider random variables which are functions of other random variables.

Functions of a Random Variable

Consider random variables which are functions of other random variables.

Theorem

Let X be a continuous random variable, and suppose that $\phi(t)$ is a strictly increasing function on the range of X. Define $Y = \phi(X)$. Suppose that random variables X and Y have cumulative distribution functions F_X and F_Y, respectively. Then, these functions are related by

$$F_Y(t) = F_X(\phi^{-1}(t)).$$

If $\phi(t)$ is strictly decreasing on the range of X, then

$$F_Y(t) = 1 - F_X(\phi^{-1}(t)).$$

Functions of a Random Variable

Proof.

Since ϕ is a strictly increasing function on the range of X,

Proof.

Since ϕ is a strictly increasing function on the range of X, the events $(X \leqslant \phi^{-1}(t))$ and $(\phi(X) \leqslant t)$ are equal.

Proof.

Since ϕ is a strictly increasing function on the range of X, the events $(X \leqslant \phi^{-1}(t))$ and $(\phi(X) \leqslant t)$ are equal. Thus, we have

$$\begin{aligned} F_Y(t) &= P(Y \leqslant t) = P(\phi(X) \leqslant t) \\ &= P(X \leqslant \phi^{-1}(t)) = F_X(\phi^{-1}(t)). \end{aligned}$$

Proof.

Since ϕ is a strictly increasing function on the range of X, the events $(X \leqslant \phi^{-1}(t))$ and $(\phi(X) \leqslant t)$ are equal. Thus, we have

$$\begin{aligned} F_Y(t) &= P(Y \leqslant t) = P(\phi(X) \leqslant t) \\ &= P(X \leqslant \phi^{-1}(t)) = F_X(\phi^{-1}(t)). \end{aligned}$$

If ϕ is strictly decreasing on the range of X, then we have

$$\begin{aligned} F_Y(t) &= P(Y \leqslant t) = P(\phi(X) \leqslant t) \\ &= P(X \geqslant \phi^{-1}(t)) \\ &= 1 - P(X \leqslant \phi^{-1}(t)) = 1 - F_X(\phi^{-1}(t)). \end{aligned}$$

$\square$

Corollary

Let X be a continuous random variable, and suppose that $\phi(t)$ is a strictly increasing function on the range of X. Define $Y = \phi(X)$. Suppose that the density functions of random variables X and Y are f_X and f_Y, respectively.

Corollary

Let X be a continuous random variable, and suppose that $\phi(t)$ is a strictly increasing function on the range of X. Define $Y = \phi(X)$. Suppose that the density functions of random variables X and Y are f_X and f_Y, respectively. Then, these functions are related by

$$f_Y(t) = f_X(\phi^{-1}(t))\tfrac{d}{dt}\phi^{-1}(t)$$

and if $\phi(t)$ is strictly decreasing then

$$f_Y(t) = -f_X(\phi^{-1}(t))\tfrac{d}{dt}\phi^{-1}(t).$$

Proof follows from Chain Rule (differentiation of composite).

If the function ϕ is neither strictly increasing nor strictly decreasing, then the situation is somewhat more complicated, but can be treated by the same methods.

If the function ϕ is neither strictly increasing nor strictly decreasing, then the situation is somewhat more complicated, but can be treated by the same methods. For example, suppose that $Y = X^2$, Then $\phi(t) = t^2$, and

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

If the function ϕ is neither strictly increasing nor strictly decreasing, then the situation is somewhat more complicated, but can be treated by the same methods. For example, suppose that $Y = X^2$, Then $\phi(t) = t^2$, and

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

Moreover,

$$\begin{aligned} f_Y(t) &= \tfrac{d}{dt} F_Y(t) = \tfrac{d}{dt}\Big(F_X(\sqrt{t}) - F_X(-\sqrt{t})\Big) \\ &= \Big(f_X(\sqrt{t}) + f_X(-\sqrt{t})\Big)\frac{1}{2\sqrt{t}}. \end{aligned}$$

Corollary

If $F(t)$ is a given cumulative distribution function that is strictly increasing when $0 < F(t) < 1$ and if U is a random variable with uniform distribution $[0, 1)$, then

$$Y = F^{-1}(U)$$

has the cumulative distribution $F(t)$.

Corollary

If $F(t)$ is a given cumulative distribution function that is strictly increasing when $0 < F(t) < 1$ and if U is a random variable with uniform distribution $[0, 1)$, then

$$Y = F^{-1}(U)$$

has the cumulative distribution $F(t)$.

Thus, to simulate a random variable with a given cumulative distribution F, we need only to set

$$Y = F^{-1}(rnd).$$

Normal Density

- One of the most important density function is the **normal density** function.

Normal Density

- One of the most important density function is the **normal density** function.
- We have seen that the binomial distribution functions are bell-shaped, even for moderate size values of n.

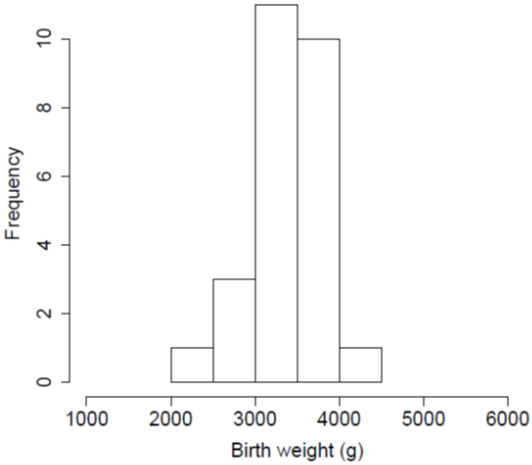
Normal Density

- One of the most important density function is the **normal density** function.
- We have seen that the binomial distribution functions are bell-shaped, even for moderate size values of n.
- We recall that a binomial-distributed random variable with parameters n and p can be considered to be the sum of n mutually independent 0-1 random variables.

- One of the most important density function is the **normal density** function.
- We have seen that the binomial distribution functions are bell-shaped, even for moderate size values of n.
- We recall that a binomial-distributed random variable with parameters n and p can be considered to be the sum of n mutually independent 0-1 random variables.
- A very important theorem in probability theory, called the **Central Limit Theorem**, states that under very general conditions, if we sum a large number of mutually independent random variables, then the distribution of the sum can be closely approximated by a certain specific continuous density, called the **normal density**.

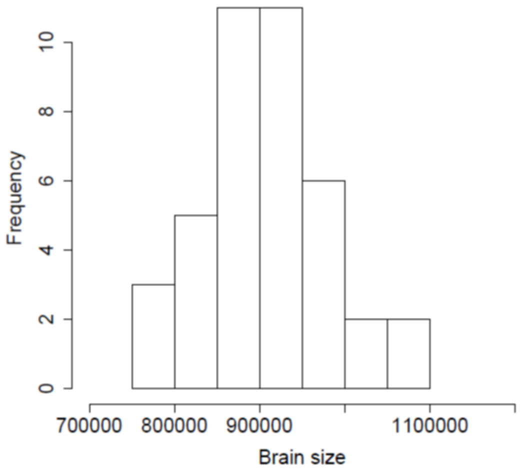
Normal Density Examples from Nature

The birth weights of the babies in USA data set:



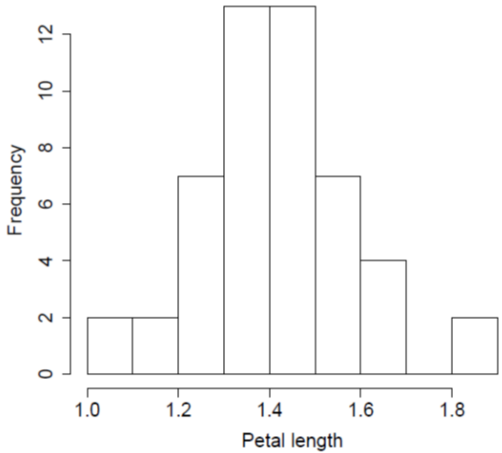
Normal Density Examples from Nature

The brain sizes of 40 students:



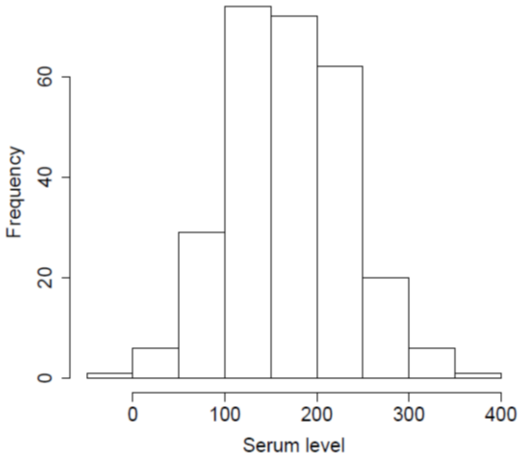
Normal Density Examples from Nature

The petal length of a type of flower:



Normal Density Examples from Nature

Serum level measurements from healthy volunteers:



Definition

The normal density function with parameters μ and σ is defined as follows:

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

Normal Density

Definition

The normal density function with parameters μ and σ is defined as follows:

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

The parameter μ represents the "center" of the density and quite soon, we will show that it is the **average**, or **expected, value** of the density.

Normal Density

Definition

The normal density function with parameters μ and σ is defined as follows:

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

The parameter μ represents the "center" of the density and quite soon, we will show that it is the **average**, or **expected, value** of the density.

The parameter σ is a measure of the "spread" of the density, and thus it is assumed to be positive. Later, we will show that is the **standard deviation** of the density.

Definition

The normal density function with parameters μ and σ is defined as follows:

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

The parameter μ represents the "center" of the density and quite soon, we will show that it is the **average**, or **expected, value** of the density.

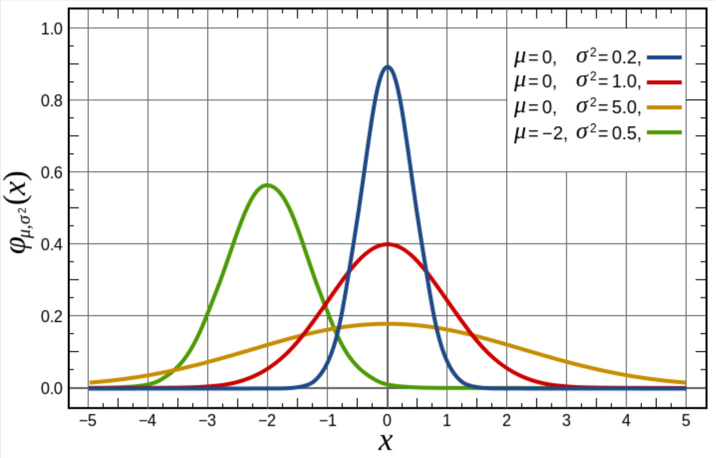
The parameter σ is a measure of the "spread" of the density, and thus it is assumed to be positive. Later, we will show that is the **standard deviation** of the density.

It is not at all obvious that the above function is a density, i.e., that

$$\int_{-\infty}^{+\infty} f_X(t)\, dt = 1.$$

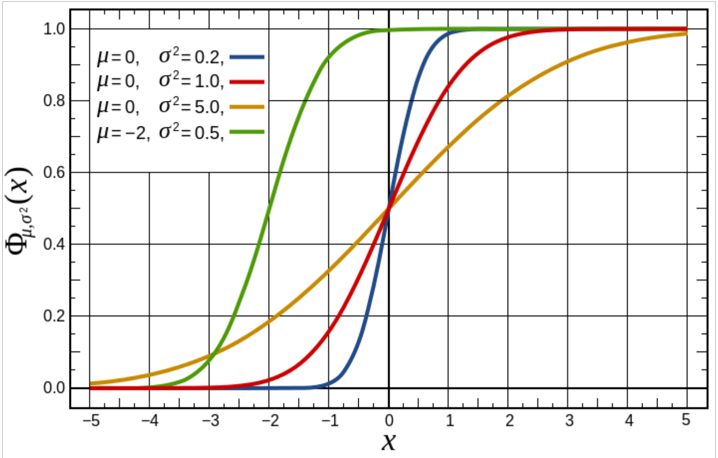
The normal density function:

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



The normal cumulative distribution function:

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt.$$



Normal Density

If real-valued continuous random variable Z is **normally distributed** with mean μ and standard deviation σ, we write

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

and Z is called a **normal random variable** with parameters μ and σ.

Normal Density

If real-valued continuous random variable Z is **normally distributed** with mean μ and standard deviation σ, we write

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

and Z is called a **normal random variable** with parameters μ and σ.

Definition

Let Z be a normal random variable with parameters $\mu = 0$ and $\sigma = 1$. A normal random variable with these parameters is said to be a **standard normal random variable**.

Normal Density

If real-valued continuous random variable Z is **normally distributed** with mean μ and standard deviation σ, we write

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

and Z is called a **normal random variable** with parameters μ and σ.

Definition

Let Z be a normal random variable with parameters $\mu = 0$ and $\sigma = 1$. A normal random variable with these parameters is said to be a **standard normal random variable**.

In order to simulate a normally distributed random variable, generate a uniform distributed variables $U, V \in [0, 1]$ and let

$$X = \sqrt{-2\ln(U)}\cos(2\pi V),$$

$$Y = \sqrt{-2\ln(U)}\sin(2\pi V).$$

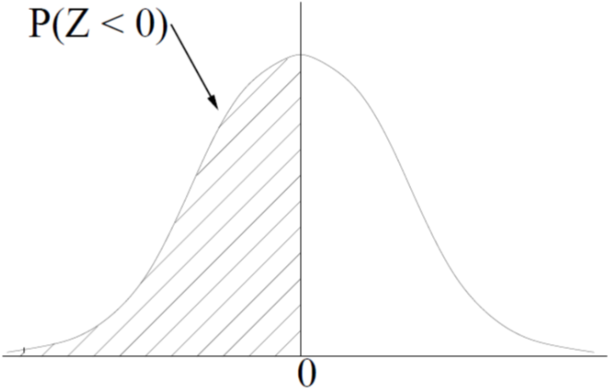
X and Y will be independent random variables with normal density functions with parameters $\mu = 0$ and $\sigma = 1$.

Normal Density

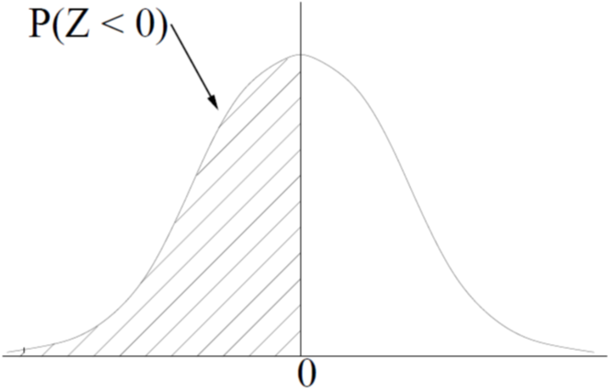
Suppose $Z \sim N(0,1)$. What is $P(Z < 0)$?

Normal Density

Suppose $Z \sim N(0,1)$. What is $P(Z < 0)$?



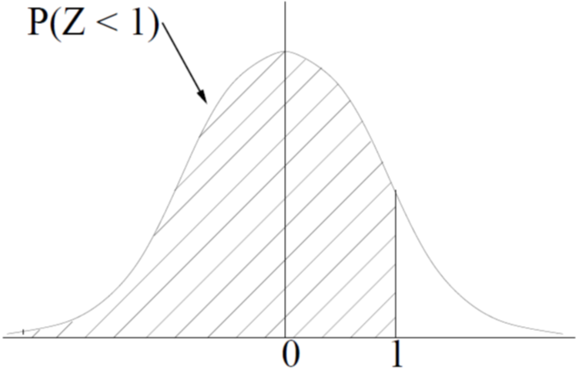
Suppose $Z \sim N(0,1)$. What is $P(Z < 0)$?



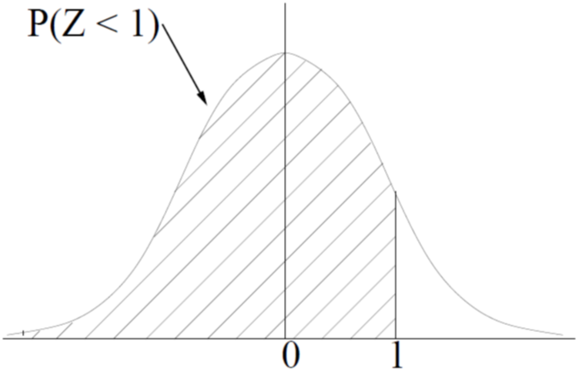
From symmetry results that $P(Z < 0) = 0.5$

How about $P(Z < 1)$?

How about $P(Z < 1)$?

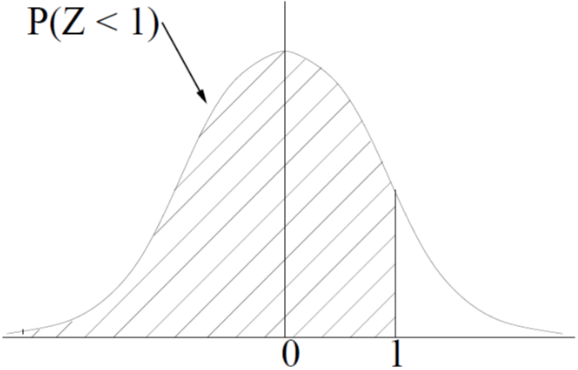


How about $P(Z < 1)$?



Calculating this area is not easy and so we use probability tables.

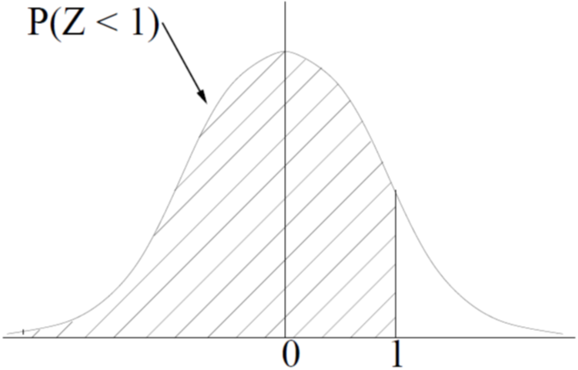
How about $P(Z < 1)$?



Calculating this area is not easy and so we use probability tables.
Probability tables are tables of probabilities that have been calculated on a computer.

Normal Density

How about $P(Z < 1)$?



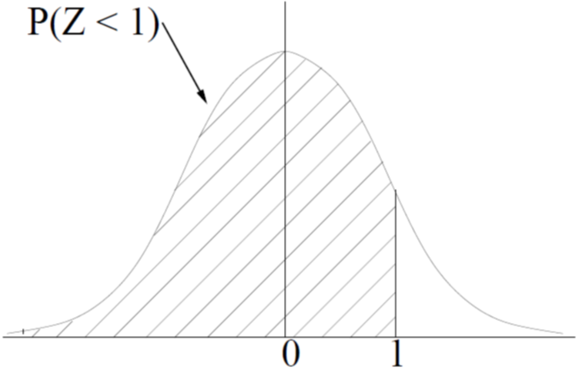
Calculating this area is not easy and so we use probability tables.
Probability tables are tables of probabilities that have been calculated on a computer.
All we have to do is identify the right probability in the table!

Normal Density

The tables allow us to read off probabilities of the form $P(Z < z)$:

z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0.1	0.5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
0.2	0.5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0.3	0.6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0.4	0.6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0.5	0.6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0.6	0.7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0.7	0.7580	7611	7642	7673	7704	7734	7764	7794	7823	7852
0.8	0.7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0.9	0.8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1.0	$\boxed{0.8413}$	8438	8461	8485	8508	8531	8554	8577	8599	8621
1.1	0.8643	8665	8686	8708	8729	8749	8770	8790	8810	8830

Normal Density



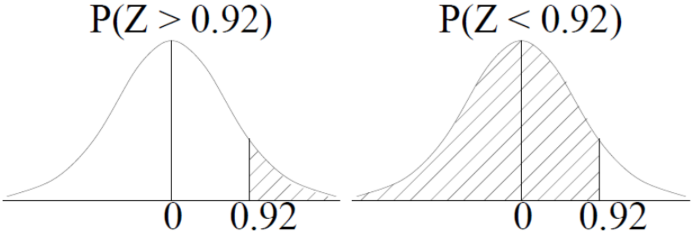
From this table we can identify that $P(Z < 1.0) = 0.8413$.

Normal Density

If $Z \sim N(0,1)$ what is $P(Z > 0.92)$?

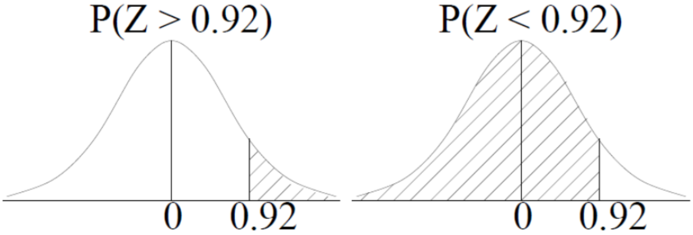
Normal Density

If $Z \sim N(0,1)$ what is $P(Z > 0.92)$?



Normal Density

If $Z \sim N(0,1)$ what is $P(Z > 0.92)$?

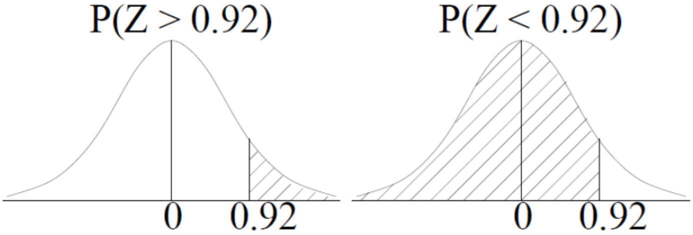


We know that

$$P(Z > 0.92) = 1 - P(Z < 0.92)$$

and we can calculate $P(Z < 0.92)$ from the tables.

If $Z \sim N(0, 1)$ what is $P(Z > 0.92)$?



We know that

$$P(Z > 0.92) = 1 - P(Z < 0.92)$$

and we can calculate $P(Z < 0.92)$ from the tables.

Thus,

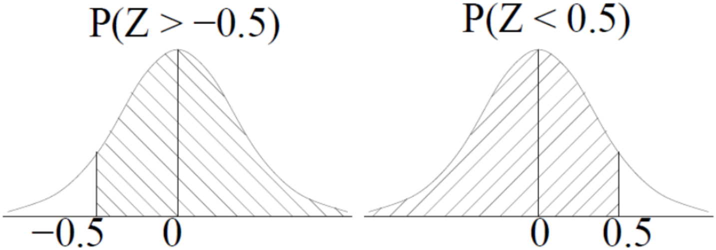
$$P(Z > 0.92) = 1 - 0.8212 = 0.1788.$$

Normal Density

If $Z \sim N(0,1)$, what is $P(Z > -0.5)$?

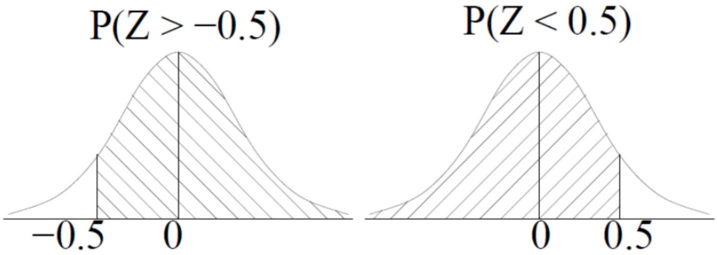
Normal Density

If $Z \sim N(0,1)$, what is $P(Z > -0.5)$?



Normal Density

If $Z \sim N(0,1)$, what is $P(Z > -0.5)$?



The Normal distribution is symmetric, so we know that

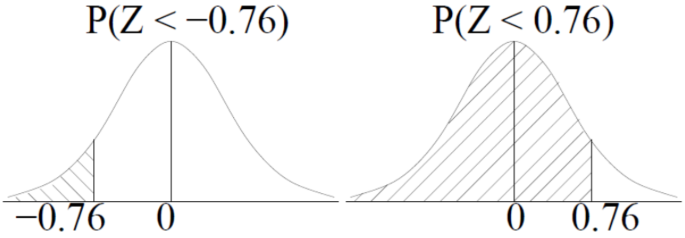
$$P(Z > -0.5) = P(Z < 0.5) = 0.6915.$$

Normal Density

If $Z \sim N(0,1)$, what is $P(Z < -0.76)$?

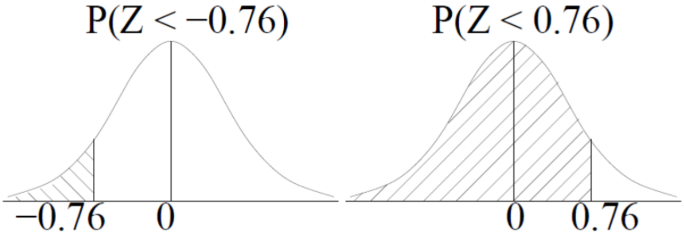
Normal Density

If $Z \sim N(0,1)$, what is $P(Z < -0.76)$?



Normal Density

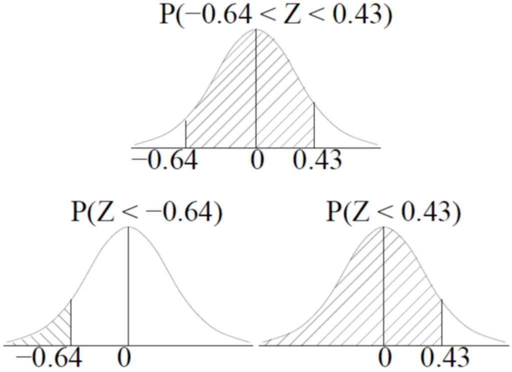
If $Z \sim N(0,1)$, what is $P(Z < -0.76)$?



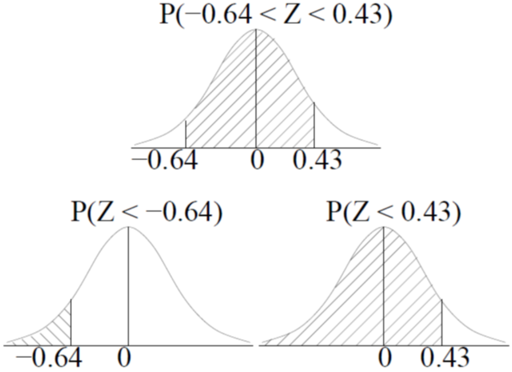
By symmetry we have

$$\begin{aligned} P(Z < -0.76) =& P(Z > 0.76) \\ =& 1 - P(Z < 0.76) \\ =& 1 - 0.7764 = 0.2236. \end{aligned}$$

If $Z \sim N(0, 1)$, what is $P(-0.64 < Z < 0.43)$?



If $Z \sim N(0,1)$, what is $P(-0.64 < Z < 0.43)$?



$$\begin{aligned} P(-0.64 < Z < 0.43) =& P(Z < 0.43) - P(Z < -0.64) \\ =& 0.6664 - (1 - 0.7389) = 0.4053. \end{aligned}$$

Normal Density

Lemma

Let Z be standard normal random variable. Then, random variable $X = \sigma Z + \mu$ is a normal random variable with parameters μ and σ.

Lemma

Let Z be standard normal random variable. Then, random variable $X = \sigma Z + \mu$ is a normal random variable with parameters μ and σ.

Proof.

We have

$$\phi(z) = \sigma z + \mu \quad \text{and} \quad \phi^{-1}(x) = \tfrac{x-\mu}{\sigma}.$$

Z has a standard normal density and thus,

$$f_Z(z) = \tfrac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

On the other hand,

$$F_X(x) = F_Z\left(\tfrac{x-\mu}{\sigma}\right), \quad f_X(x) = f_Z\left(\tfrac{x-\mu}{\sigma}\right)\tfrac{1}{\sigma} = \tfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We have seen above that it is possible to simulate a standard normal random variable Z.
If we wish to simulate a normal random variable X with parameters μ and σ, then we need only transform the simulated values for Z using the equation $X = \sigma Z + \mu$.

Normal Density

We have seen above that it is possible to simulate a standard normal random variable Z.
If we wish to simulate a normal random variable X with parameters μ and σ, then we need only transform the simulated values for Z using the equation $X = \sigma Z + \mu$.

Suppose that we wish to calculate the value of a cumulative distribution function for the normal random variable X, with parameters μ and σ.

We have seen above that it is possible to simulate a standard normal random variable Z.
If we wish to simulate a normal random variable X with parameters μ and σ, then we need only transform the simulated values for Z using the equation $X = \sigma Z + \mu$.

Suppose that we wish to calculate the value of a cumulative distribution function for the normal random variable X, with parameters μ and σ.
We can reduce this calculation to one concerning the standard normal random variable Z as follows:

$$\begin{aligned} F_X(x) = P(X \leqslant x) = P(\sigma Z + \mu \leq x) =& P(Z \leqslant \tfrac{x-\mu}{\sigma}) \\ =& F_Z\left(\tfrac{x-\mu}{\sigma}\right). \end{aligned}$$

Normal Density

We have seen above that it is possible to simulate a standard normal random variable Z.
If we wish to simulate a normal random variable X with parameters μ and σ, then we need only transform the simulated values for Z using the equation $X = \sigma Z + \mu$.

Suppose that we wish to calculate the value of a cumulative distribution function for the normal random variable X, with parameters μ and σ.
We can reduce this calculation to one concerning the standard normal random variable Z as follows:

$$\begin{aligned} F_X(x) = P(X \leqslant x) = P(\sigma Z + \mu \leq x) &= P(Z \leqslant \tfrac{x-\mu}{\sigma}) \\ &= F_Z\left(\tfrac{x-\mu}{\sigma}\right). \end{aligned}$$

This last expression can be found in a table of values of the cumulative distribution function for a standard normal random variable.

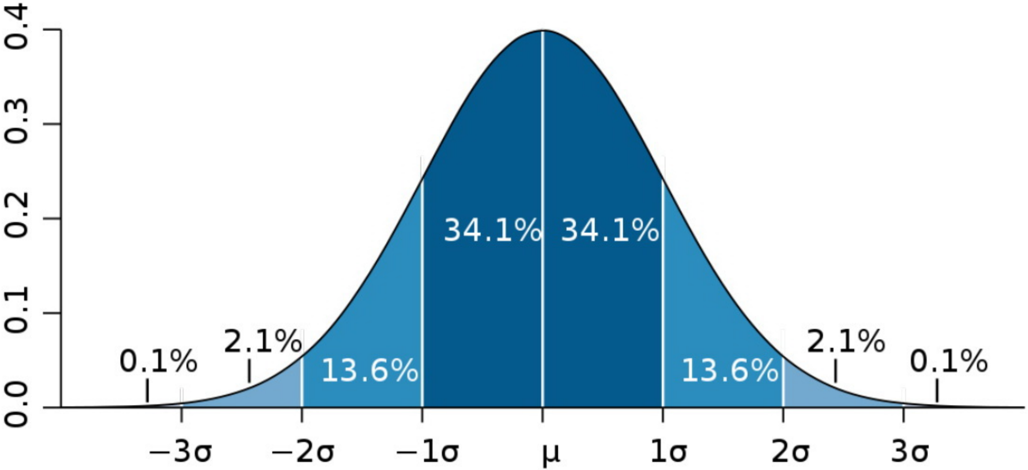
Example.

Suppose that X is a normally distributed random variable with parameters $\mu = 10$ and $\sigma = 3$. Find the probability that X is between 4 and 16. That is $P(4 \leqslant X \leqslant 16)$?

Example.
Suppose that X is a normally distributed random variable with parameters $\mu = 10$ and $\sigma = 3$. Find the probability that X is between 4 and 16. That is $P(4 \leqslant X \leqslant 16)$?
Solution.
To solve this problem, we note that $Z = \frac{X-10}{3}$ is the standardized version of X.
So, we have

$$
\begin{aligned}
P(4 \leqslant X \leqslant 16) =& P(X \leqslant 16) - P(X \leqslant 4) \\
=& F_X(16) - F_X(4) \\
=& F_Z\left(\tfrac{16-10}{3}\right) - F_Z\left(\tfrac{4-10}{3}\right) \\
=& F_Z(2) - F_Z(-2) \\
=& 0.9772 - 0.0228 \\
=& 0.9544.
\end{aligned}
$$

Normal Density



Expected Value of CRV

Definition

Let X be a real-valued continuous random variable with density function $f(t)$. The **expected value** $\mu = E(X)$ is defined by:

$$\mu = E(X) = \int_{-\infty}^{+\infty} t\, f(t)\, dt$$

provided the integral

$$\int_{-\infty}^{+\infty} |t|\, f(t)\, dt$$

is finite (convergent).

Expected Value of CRV

Theorem

If X and Y are real-valued continuous random variables and c is any constant, then

$$\begin{aligned} E(X+Y) &= E(X) + E(Y), \\ E(cX) &= cE(X). \end{aligned}$$

Example. Let X be uniformly distributed on the interval $[0,1)$. Then

$$E(X) = \int_0^1 t\,dt = \frac{1}{2}.$$

It follows that, if we choose a large number N of random numbers from $[0,1)$ and take the average, then we can expect that this average should be close to the expected value of $1/2$.

Example.

Suppose that Maria and Ion agree to meet at Stefan cel Mare statue between 19:00 and 20:00 on Friday.

Suppose each arrives at a time between 19:00 and 20:00 chosen at random with uniform probability.

Let Z be the the length of time that the first to arrive has to wait for the other.

Find the cumulative distribution and density functions of Z.

Also find $E(Z)$.

Example.
Suppose that Maria and Ion agree to meet at Stefan cel Mare statue between 19:00 and 20:00 on Friday.

Suppose each arrives at a time between 19:00 and 20:00 chosen at random with uniform probability.

Let Z be the the length of time that the first to arrive has to wait for the other.

Find the cumulative distribution and density functions of Z.

Also find $E(Z)$.

Solution.
Let $X \in [0, 1)$ and $Y \in [0, 1)$ be the arrival time after 19:00 of Maria and Ion, respectively ($X = 0$ means 19:00 and $X = 1$ corresponds to 20:00).

Example.
Suppose that Maria and Ion agree to meet at Stefan cel Mare statue between 19:00 and 20:00 on Friday.

Suppose each arrives at a time between 19:00 and 20:00 chosen at random with uniform probability.

Let Z be the the length of time that the first to arrive has to wait for the other.

Find the cumulative distribution and density functions of Z.

Also find $E(Z)$.

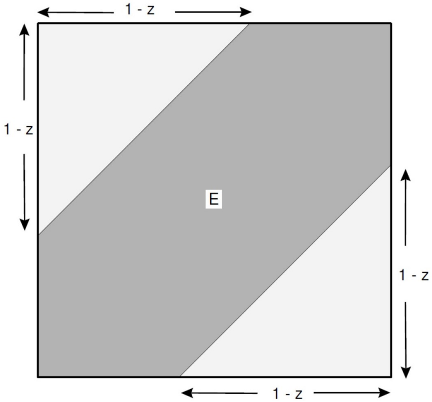
Solution.
Let $X \in [0, 1)$ and $Y \in [0, 1)$ be the arrival time after 19:00 of Maria and Ion, respectively ($X = 0$ means 19:00 and $X = 1$ corresponds to 20:00).

Then waiting time is

$$Z = |X - Y|.$$

Expected Value of CRV

$$
\begin{aligned}
F_Z(z) =& P(Z < z) \\
=& P(|X - Y| < z) \\
=& \text{ Area of } E
\end{aligned}
$$



Expected Value of CRV

Cumulative distribution function is:

$$F_Z(z) = \begin{cases} 0, & z < 0, \\ 1 - (1 - z)^2, & 0 \leqslant z < 1, \\ 1, & z \geqslant 1. \end{cases}$$

Expected Value of CRV

Cumulative distribution function is:

$$F_Z(z) = \begin{cases} 0, & z < 0, \\ 1 - (1-z)^2, & 0 \leqslant z < 1, \\ 1, & z \geqslant 1. \end{cases}$$

Obtain density function by differentiating F_Z with respect to z:

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ 2(1-z), & 0 \leqslant z < 1, \\ 0, & z \geqslant 1. \end{cases}$$

Expected Value of CRV

Cumulative distribution function is:

$$F_Z(z) = \begin{cases} 0, & z < 0, \\ 1 - (1-z)^2, & 0 \leqslant z < 1, \\ 1, & z \geqslant 1. \end{cases}$$

Obtain density function by differentiating F_Z with respect to z:

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ 2(1-z), & 0 \leqslant z < 1, \\ 0, & z \geqslant 1. \end{cases}$$

Compute expected value from definition:

$$E(Z) = \int_{-\infty}^{+\infty} z\, f(z)\, dz = \int_0^1 z \cdot 2(1-z)\, dz = 2\int_0^1 (z - z^2)\, dz = \tfrac{1}{3}.$$

Expected Value of CRV

Cumulative distribution function is:

$$F_Z(z) = \begin{cases} 0, & z < 0, \\ 1 - (1-z)^2, & 0 \leqslant z < 1, \\ 1, & z \geqslant 1. \end{cases}$$

Obtain density function by differentiating F_Z with respect to z:

$$f_Z(z) = \begin{cases} 0, & z < 0, \\ 2(1-z), & 0 \leqslant z < 1, \\ 0, & z \geqslant 1. \end{cases}$$

Compute expected value from definition:

$$E(Z) = \int_{-\infty}^{+\infty} z\, f(z)\, dz = \int_0^1 z \cdot 2(1-z)\, dz = 2\int_0^1 (z - z^2)\, dz = \tfrac{1}{3}.$$

Therefore, the average waiting time will be 20 minutes.

Expected Value of CRV

Theorem

If X is a real-valued random variable and if $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous real-valued function with domain $[a, b]$, then

$$E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(t)\, f(t)\, dt$$

provided the integral exists and is finite.

Expected Value of CRV

Theorem

If X is a real-valued random variable and if $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous real-valued function with domain $[a, b]$, then

$$E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(t)\, f(t)\, dt$$

provided the integral exists and is finite.

Theorem

Let X and Y be independent real-valued continuous random variables with finite expected values. Then, we have

$$E(XY) = E(X) \cdot E(Y).$$

Theorem

If X is a real-valued random variable and if $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous real-valued function with domain $[a, b]$, then

$$E(\phi(X)) = \int_{-\infty}^{+\infty} \phi(t)\, f(t)\, dt$$

provided the integral exists and is finite.

Theorem

Let X and Y be independent real-valued continuous random variables with finite expected values. Then, we have

$$E(XY) = E(X) \cdot E(Y).$$

Observe that these are the same properties as in discrete case.

Example.
Let $Z = (X, Y)$ be a point chosen at random in the unit square.
Let $A = X^2$ and $B = Y^2$. Then, A and B are independent.
Using last theorem, the expectations of A and B are easy to calculate:

$$E(A) = E(B) = \int_0^1 t^2 \, dt = \tfrac{1}{3}.$$

The expectation of CRV AB is just the product of $E(A)$ and $E(B)$, or $1/9$.
The usefulness of this theorem is demonstrated by noting that it is quite a bit more difficult to calculate $E(AB)$ from the definition of expectation.
One finds that the density function of AB is $f_{AB}(t) = \frac{-\ln(t)}{4\sqrt{t}}$, so

$$E(AB) = \int_0^1 t \, f_{AB}(t) \, dt = \tfrac{1}{9}.$$

Variance of CRV

Definition

Let X be a real-valued random variable with density function $f(x)$.
The **variance** $\sigma^2 = V(X)$ is defined by

$$\sigma^2 = V(X) = E((X-\mu)^2).$$

The **standard deviation** $\sigma = D(X)$ is

$$\sigma = D(X) = \sqrt{V(X)}.$$

Variance of CRV

Definition

Let X be a real-valued random variable with density function $f(x)$.
The **variance** $\sigma^2 = V(X)$ is defined by

$$\sigma^2 = V(X) = E((X - \mu)^2).$$

The **standard deviation** $\sigma = D(X)$ is

$$\sigma = D(X) = \sqrt{V(X)}.$$

Theorem

If X is a real-valued random variable with $E(X) = \mu$, then

$$\sigma^2 = V(X) = \int_{-\infty}^{+\infty} (t - \mu^2) f(t)\, dt.$$

Example.
Let X be an exponentially distributed random variable with parameter λ.

Then, the density function of X is

$$f_X(t) = \lambda e^{-\lambda t}, \qquad t \in [0, +\infty).$$

Exponential Density

Example.
Let X be an exponentially distributed random variable with parameter λ.

Then, the density function of X is

$$f_X(t) = \lambda e^{-\lambda t}, \qquad t \in [0, +\infty).$$

From the definition of expectation and variance, and using integration by parts, we have

$$E(X) = \int_{-\infty}^{+\infty} t\,\lambda e^{-\lambda t} dt \int_{0}^{+\infty} t\,\lambda e^{-\lambda t} dt = \frac{1}{\lambda}.$$

Similarly, it can be computed the variance and standard deviation for a CRV with exponential density with parameter λ:

$$V(X) = \frac{1}{\lambda^2}, \qquad D(X) = \frac{1}{\lambda}.$$

The End of Lecture