Continuous Probability Distributions

Sections 5.1 - 5.4

Cathy Poliak, Ph.D.
cathy@math.uh.edu
Office in 11c Flemming

Department of Mathematics
University of Houston

Lecture 7 - 3339

You Try Questions

For each random variable, determine if it is:

a. Discrete   b. Continuous

2. The number of cars passing a busy intersection between 4:30 PM and 6:30 PM. X = number of cars a)
3. The weight of a fire fighter. $= X$ b)
4. The amount of soda in a can of Pepsi.
   Continuous

Types of Random Variables

- A random variable that may assume either a finite number of values or an infinite sequence of values such as $0, 1, \ldots$ is referred to as a **discrete random variable.**

- A random variable that may assume any numerical value in an interval or collection of intervals is called a **continuous random variable.**

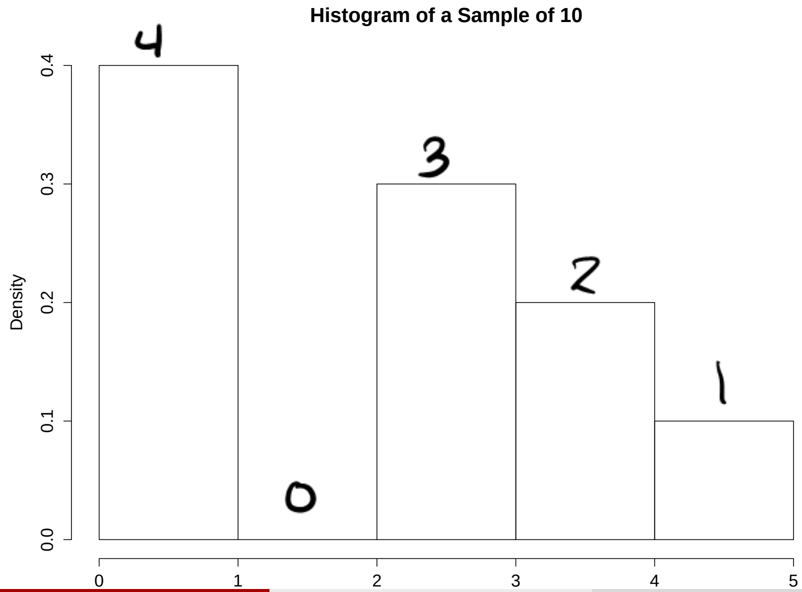
Example

- Suppose we want to determine the probability of waiting for an elevator where the longest waiting time is 5 minutes.
- What type of variable do we have?
- Suppose we take a sample of 10, 50, 1000, and 10,000 people to see how long they wait for the elevator. The following are histograms for the waiting times of each sample.

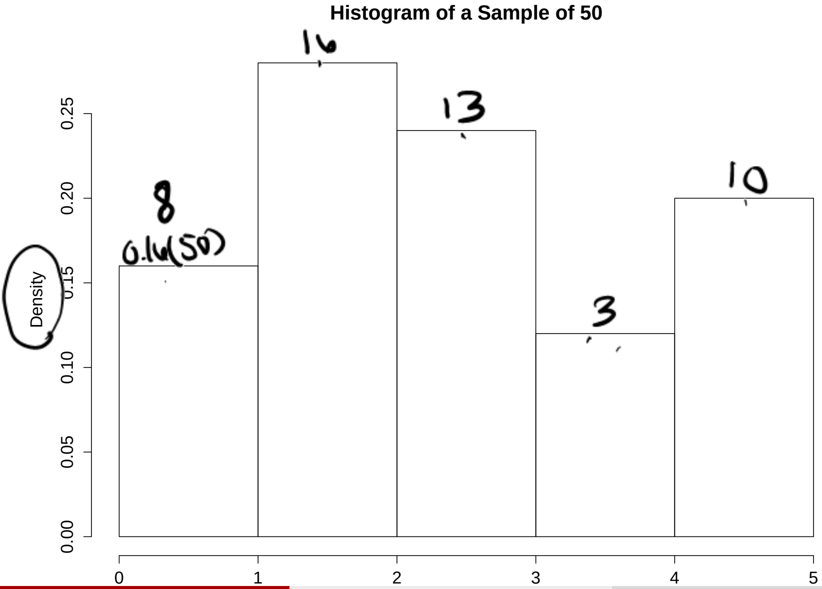
X = waiting time for an elevator

$\{x \mid$ all real numbers between 0 and 5$\}$

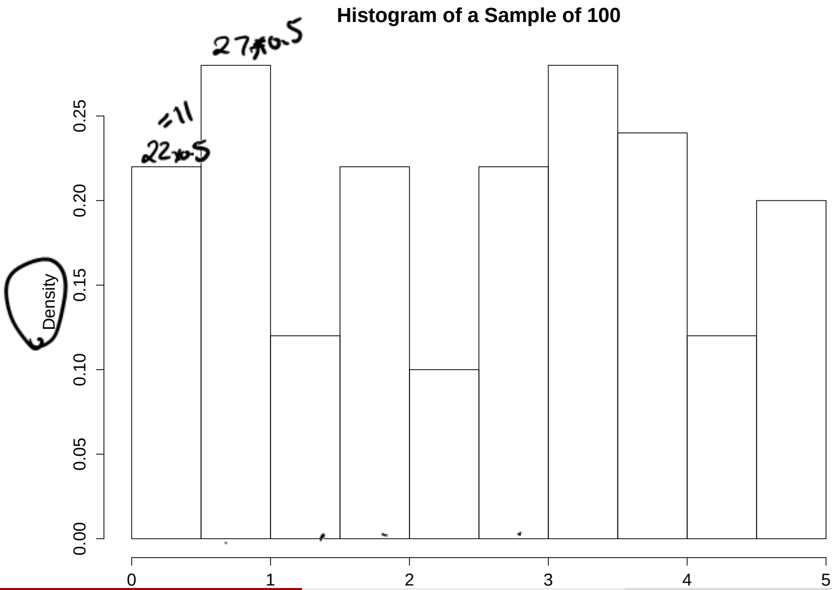
Sample of 10 People Waiting for the Elevator



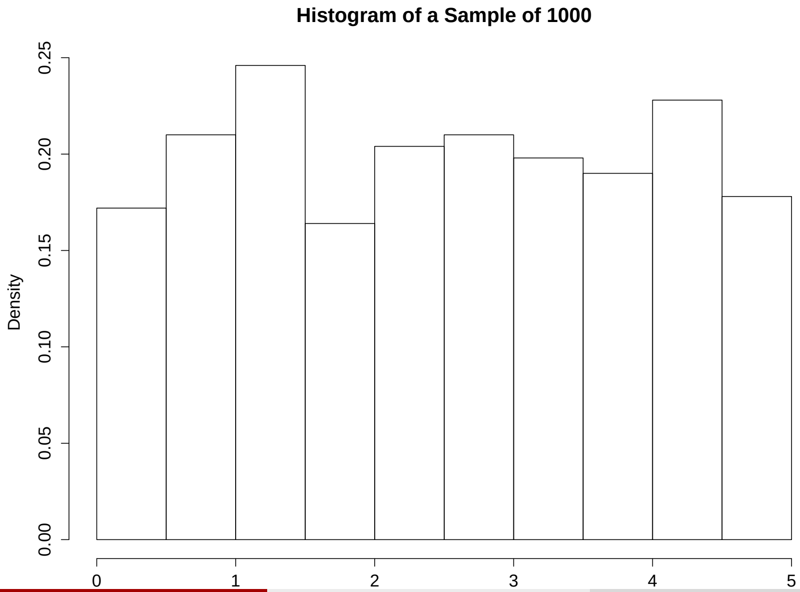
Sample of 50 People Waiting for the Elevator



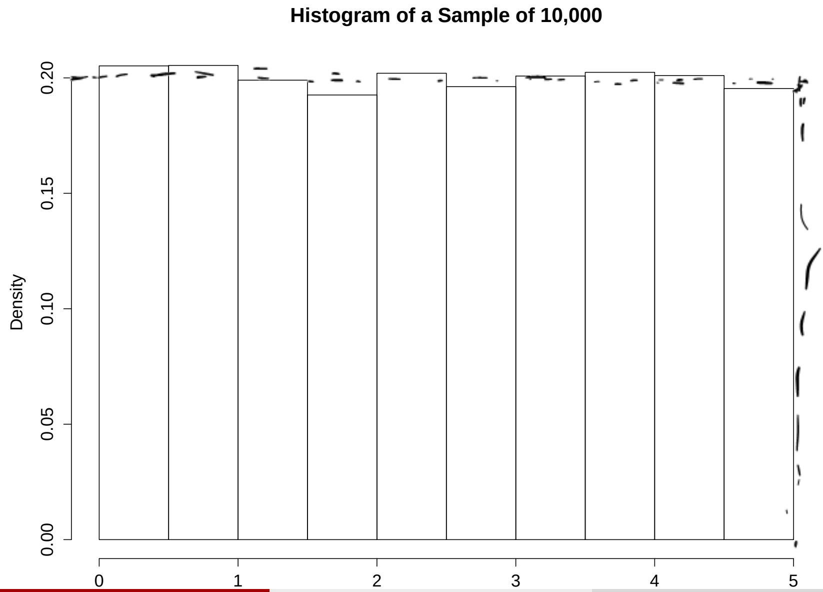
Sample of 100 People Waiting for the Elevator

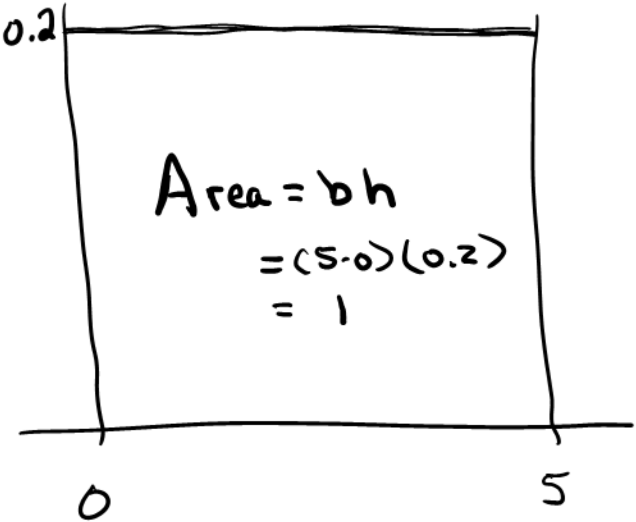


Sample of 1000 People Waiting for the Elevator



Sample of 10,000 People Waiting for the Elevator




0.2
Area = bh
= (5-0)(0.2)
= 1
0
5

Probability distributions

- A **probability distribution** for random variables describes how probabilities are distributed over the values of the random variable.
- For a discrete random variable X, the probability distribution is defined by **probability mass function**, denoted by $f(X)$. This provides the probability for each value of the random variable.
- For a continuous random variable, this is called the **probability density function** $f(x)$.The probability density function (pdf) $f(x)$ is a graph of an equation. The area under the graph of $f(x)$ corresponding to a given interval provides the probability that the random variable x assumes a value in that interval.

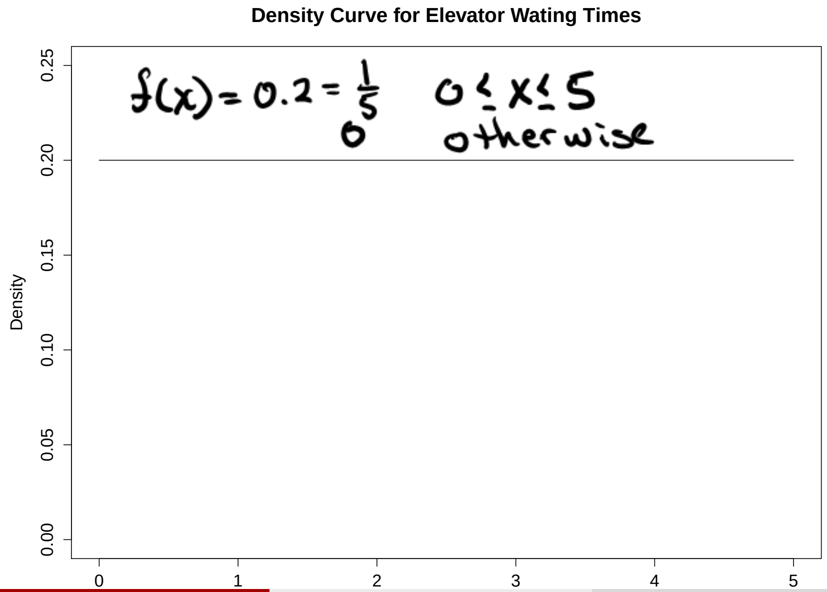
Probability Density Function

For $f(x)$ to be a legitimate pdf, it must satisfy the following two conditions:

1. $f(x) \geq 0$ for all x.

2. The area under the entire graph of $f(x)$ must equal 1.

Probability Density Function of Elevator Waiting Times



Uniform Distribution

A continuous random variable X is said to have a **uniform distribution** on the interval $[A, B]$ if the pdf of X is:

$$f(x) = \begin{cases} \frac{1}{B-A}, & A \leq x \leq B \\ 0, & \text{otherwise} \end{cases}$$

This is denoted as $X \sim U(a, b)$

Elevator example X = min. waiting for an elevator

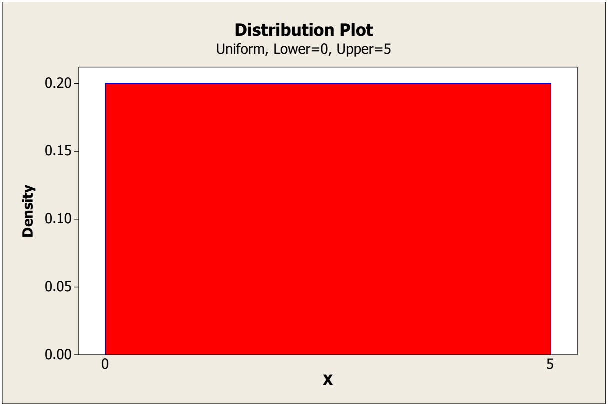
$X \sim U(0, 5)$

pdf

$$f(x) = \begin{cases} \frac{1}{5} & 0 \leq x \leq 5 \\ 0, & \text{otherwise} \end{cases}$$

Density curve for waiting time

The rectangle ranges between 0 and 5. The height of the rectangle is: $\frac{1}{\text{highest value} - \text{lowest value}} = \frac{1}{5-0} = 0.2$.



From Waiting Time Example Determine the Following

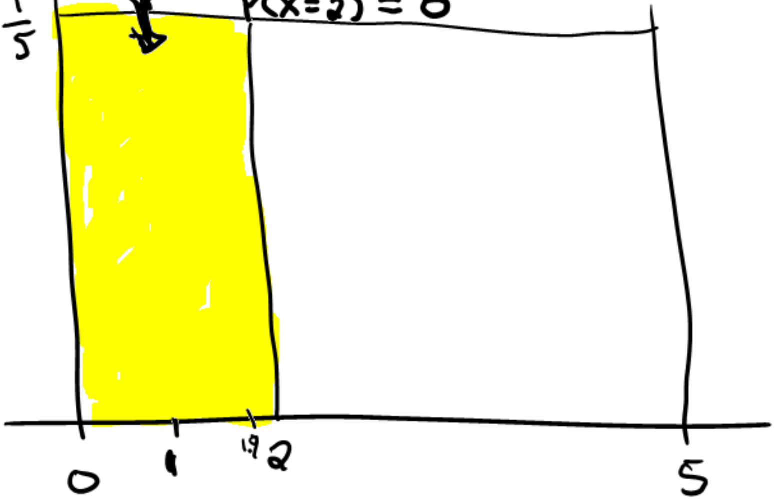
Let X = the waiting time for the elevator. With $X \sim U(0, 5)$.

1. $P(X \leq 2)$
2. $P(X < 2)$
3. $P(X = 2)$
4. $P(2 < X < 4)$ $= 2(\frac{1}{5}) = 0.4$
5. $P(X > 4)$ $= 1(\frac{1}{5}) = 0.2$
6. Find X_0 such that $P(X \leq x_0) = 0.25$.

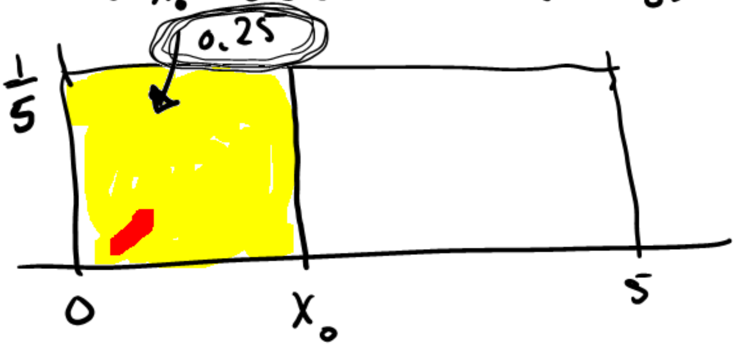
$P(X \le 2) = 2(0.2) = 0.4$

$P(X < 2) = 0.4 = P(X \le 2)$

$P(X = 2) = 0$



6. Find x_0 such that $P(X \le x_0) = 0.25$



$$X_0(0.2) = 0.25$$

$$X_0 = \frac{0.25}{0.2}$$

$$x_0 = 1.25 \leftarrow \text{quantile}$$

Q1

Your Turn

Old Faithful erupts every 91 minutes. Let X = the time you wait for Old Faithful to erupt. Assume a uniform distribution

1. What is the pdf of the time waiting? $X \sim U(0, 91)$

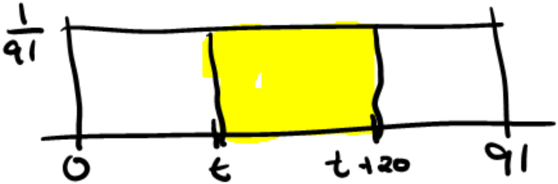
$$f(x) = \begin{cases} \frac{1}{91-0} = \frac{1}{91} & 0 \le x \le 91 \\ 0 & \text{otherwise} \end{cases}$$

2. You arrive there at random and wait for 20 minutes ... what is the probability you will see it erupt? $P(t < X < t+20)$

$$= (t+20-t)\left(\frac{1}{91}\right)$$

$$= \frac{20}{91}$$

$$= 0.2198$$



Example of a density function

Let the random variable X = a dealer's profit, in units of \$5000, on a new automobile with a density function:

$$f(x) = \begin{cases} 2(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

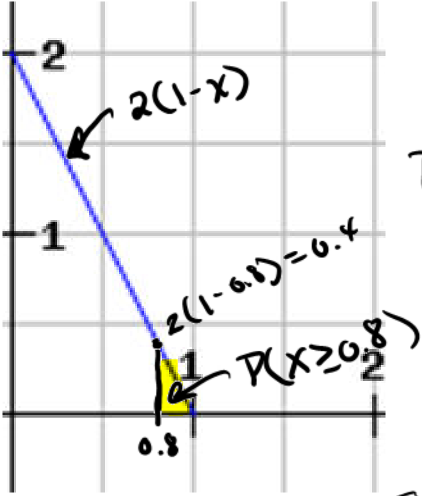
What is the probability that the dealer's profit is at least \$4000 for a new automobile. That is $P(X \geq \frac{4000}{5000}) = P(X \geq 0.8)$.

Finding Probability

To find the probability of the profit at least \$4000, we need to find the area under the curve between 0.8 and 1.

Density Function

This is the graph of the density function.



$P(X \geq 0.8)$

$= \frac{1}{2}(1-0.8)(0.4)$

$= 0.04$

$P(X \geq 0.8)$

$= \int_{0.8}^{1} 2(1-x)\,dx$

$= \int_{0.8}^{1} 2-2x\,dx$

$= 2x - x^2 \Big|_{0.8}^{1}$

$= [2(1)-1^2] - [2(0.8)-0.8^2]$

$= 0.04$

Definition of a Density Function

- A **density function** is a non-negative function f defined of the set of real numbers such that:

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

- If f is a density function, then its integral $F(x) = \int_{-\infty}^{x} f(u)du$ is a continuous cumulative distribution function (cdf), that is $P(X \leq x) = F(x)$.

- If X is a random variable with this density function, then for any two real numbers, a and b

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx.$$

Integration

1. Evaluate $\int_0^5 (\frac{1}{5}) dx$. $= \frac{1}{5} x \Big|_0^5 = \frac{1}{5}(5-0) = 1$

$t \sim exp(\lambda = 0.3)$

2. Evaluate $\int_0^1 0.3e^{-0.3t} dt$. $= 1 - e^{-0.3}$

$u = -0.3t$

$du = -0.3dt$

$= -\int e^u du$

$= -e^{-03t} \Big|_0^1 = -e^{-0.3} - (-1)$

$= 1 - e^{-0.3}$

$= 0.2592$

3. Evaluate $\int_0^\infty 2xe^{-2x} dx$. ibp

$\int u dv = uv - \int v du$

$u = 2x$

$du = 2$

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

$dv = e^{-2x} dx$

$$\int_0^\infty 2xe^{-2x}\,dx = -xe^{-2x}\Big|_0^\infty - \int_0^\infty -e^{-2x}\,dx$$

$$= [0-0] + -\frac{1}{2}e^{-2x}\Big|_0^\infty$$

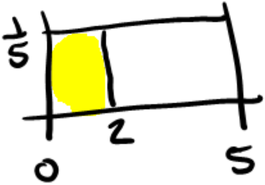
$$= 0 - \left(-\frac{1}{2}\right)$$

$$= \frac{1}{2}$$

```
> f=function(x) 2*x*exp(-2*x)
> integrate(f,0,Inf)
0.5 with absolute error < 8.6e-06
```

Let $X \sim U(0, 5)$ such that the pdf of X is:

$$f(x) = \begin{cases} \frac{1}{5} & 0 \le x \le 5 \\ 0 & otherwise \end{cases}$$



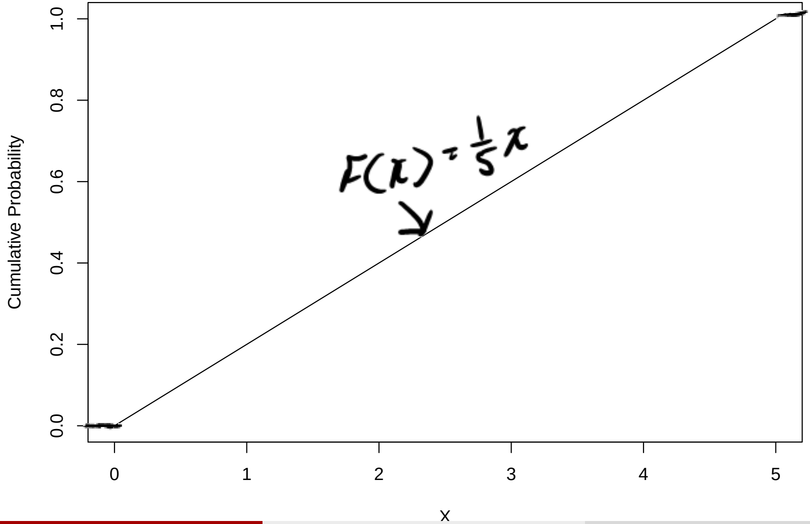
Find the cdf $F(x)$for X.

$$F(x) = P(X \le x)$$
$$= P(X \le 2)$$
$$= \frac{1}{5}(2) = 0.4$$

$$F(x) = \int_0^x \frac{1}{5}\,du = \frac{1}{5}x \quad 0 \le x \le 5$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{5}x & 0 \le x \le 5 \\ 1 & x > 5 \end{cases} \qquad P(X \le 5) = 1$$

Cumulative Density Function



Using the cdf $F(X)$ to Compute Probabilities

Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. Then for any number a,

$$= 1 - P(X \leq a)$$

$$P(X > a) = 1 - F(a)$$

and for any two numbers a and b with $a < b$,

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

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

Example Using CDF

Suppose we have a cdf;

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \frac{x^3+1}{9}, & -1 \leq x < 2 \\ 1, & x \geq 2. \end{cases}$$

1. Determine $P(X \leq 0) = F(0) = \frac{0^3+1}{9} = \frac{1}{9}$
2. Determine $P(0 < X \leq 1) = F(1) - F(0)$
   $= \frac{1^3+1}{9} - \frac{1}{9} = \frac{1}{9}$
3. Determine $P(X \geq 0.5) = 1 - P(X \leq 0.5)$
   $= 1 - \frac{0.5^3+1}{9} = 1 - 0.125$
   $= 0.875$
4. Given this CDF determine the pdf $f(x)$.

$F'(x) = \frac{2x}{9}$ $\quad$ $F(x) = \frac{x^2+1}{9}$

pdf

$$f(x) = \begin{cases} \frac{2x}{9} & -1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Example

Suppose we have a pdf of

$$f(x) = \begin{cases} \frac{3}{8}x^2 & 0 \leq X \leq k \\ 0 & otherwise \end{cases}$$

$$\int_0^k \frac{3}{8}x^2 dx = 1$$

$$\frac{x^3}{8} = 1$$

$$k^3 = 8$$

$$k = 2$$

a) Determine k.

$$\int_0^k \frac{3}{8}x^2 dx = \frac{1}{8}x^3 \Big|_0^k = \frac{k^3}{8}$$

b) Give the cdf of this distribution.

$$F(x) = \int_0^x \frac{3}{8}u^2 du = \frac{x^3}{8}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^3}{8} & 0 \leq X \leq 2 \\ 1 & X > 2 \end{cases}$$

c) Determine x_0 such that $P(X \leq x_0) = 0.125$

Find x_0 such that

$$P(X \leq x_0) = 0.125$$

$$F(x_0) = 0.125$$

$$\frac{x_0^3}{8} = 0.125$$

$$x_0^2 = 1$$

$$x_0 = 1$$ "quantile"

Quantiles

Let F be a given cumulative distribution and let p be any real number between 0 and 1. The **(100p)th percentile** of the distribution of a continuous random variable X is defined as

$$F^{-1}(p) = min\{x|F(x) \geq p\}.$$

For continuous distributions, $F^{-1}(p)$ is the smallest number x such that $F(x) = p$.

Determine the Percentiles

Given a cdf,

$$F(x) = \begin{cases} 0 & X < 0 \\ \frac{1}{8}x^3 & 0 \le X \le 2 \\ 1 & X > 2 \end{cases}$$

1. Determine the 90^{th} percentile.

$P(X \le x_0) = 0.9$

$F(x_0) = 0.9$

$\frac{x^3}{8} = 0.9$

$x^3 = 7.2$

$x = 1.931$

```
> 7.2^(1/3)
[1] 1.930979
```

2. Determine the 50^{th} percentile.

$\frac{x^3}{8} = 0.5 \Rightarrow x^3 = 4 \Rightarrow x = 1.5874$ median

3. Find the value of c such that $P(X \le c) = 0.75$.

$\frac{c^3}{8} = 0.75 \Rightarrow c^3 = 6 \Rightarrow c = 1.8171$ Q3

Expected Values for Continuous Random Variables

Discrete r.v. $E(x) = \sum xp$

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

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

More generally, if h is a function defined on the range of X,

$$E(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx.$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$$

$$var(x) = E(x^2) - E(x)^2$$

Example

The following is a pdf of X,

$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq X \leq 1 \\ 0 & \textit{otherwise} \end{cases}$$

1. Determine $E(X)$.

$$E(x) = \int_0^1 x\left[\frac{3}{2}(1-x^2)\right]dx = 0.375$$

```
> f=function(x) 3/2*x*(1-x^2)
> integrate(f,0,1)
0.375 with absolute error <
4.2e-15
```

2. Determine $E(X^2)$

$$E(x^2) = \int_0^1 x^2\left[\frac{3}{2}(1-x^2)\right]dx = 0.2$$

```
> #E(x^2)
> f=function(x) 3/2*x^2*(1-x^2)
> integrate(f,0,1)
0.2 with absolute error < 2.2e-15
```

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

$$= 0.2 - 0.375^2$$

$$= 0.059375$$

$$SD(x) = \sqrt{0.059375} = 0.24367$$

```
> sqrt(.Last.value)
[1] 0.2436699
```

Mean and Variance of the Uniform Distribution

Let $X \sim$ Unif(a, b)

- $E(X) = \frac{a+b}{2}$
- $\text{Var}(X) = \frac{(b-a)^2}{12}$

Elevator example

$f(x) = \frac{1}{5} \quad 0 \leq x \leq 5$

$E(x) = \frac{0+5}{2} = 2.5$

$\text{Var}(x) = \frac{(5-0)^2}{12} = \frac{25}{12}$

$E(x) = \int_0^5 \frac{1}{5} x \, dx$

$E(x^2) = \int_0^5 \frac{1}{5} x^2 \, dx$

$$E(X) = \int_0^5 \frac{1}{5} x\,dx = \frac{1}{10} x^2 \Big|_0^5 = \frac{5}{2}$$

$$E(X^2) = \int_0^5 \frac{1}{5} x^2\,dx = \frac{1}{15} x^3 \Big|_0^5 = \frac{25}{3}$$

$$\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= \frac{25}{3} - \left(\frac{5}{2}\right)^2 \\
&= \frac{25}{3} - \frac{25}{4} \\
&= \frac{100-75}{12} = \frac{25}{12}
\end{aligned}$$

Example From Quiz 7

Let X be the amount of time (in hours) the wait is to get a table at a restaurant. Suppose the cdf is represented by

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

Use the cdf to determine E[X].

1. Take the derivative.

$$F'(x) = \frac{2x}{9}$$

$$\int_0^3 x\left(\frac{2x}{9}\right)dx = \int_0^3 \frac{2x^2}{9}\,dx = \frac{2x^3}{27}\Big|_0^3 = 2$$

The Exponential Distribution

X is said to have an **exponential distribution** with parameter λ = rate ($\lambda > 0$) if the pdf of X is:

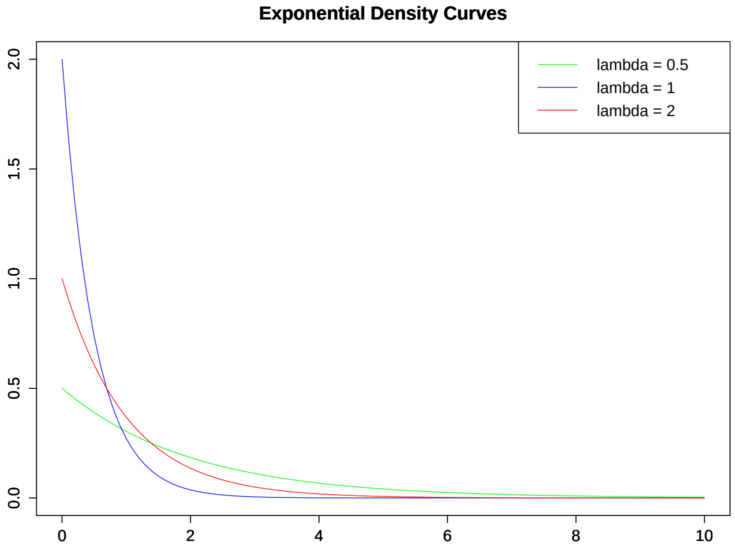
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases}$$

Where λ is a rate parameter, we write $X \sim Exp(\lambda)$. The cdf of a exponential random variable is:

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

The mean of the exponential distribution is $\mu_X = E(X) = \frac{1}{\lambda}$ the standard deviation is also $\frac{1}{\lambda}$.

Exponential Density Curves



Exponential Distribution Related to the Poisson Distribution

- The exponential distribution is frequently used as a model for the distribution of times between the occurrence of successive events until the first arrival.
- Suppose that the number of events occurring in any time of length t has a Poisson distribution with parameter αt.
- Where α, the rate of the event process, is the expected number of events occurring in 1 unit of time.
- The number of occurrences are in non overlapping intervals and are independent of one another.
- Then the distribution of elapsed time between the occurrence of two successive events is exponential with parameter $\lambda = \alpha$.

Example

- Suppose you usually get 3 phone calls per hour.
- 3 phone calls per hour means that we would expect one phone call every $\frac{1}{3}$ hour so $\lambda = \frac{1}{3}$.
- Compute the probability that a phone call will arrive within the next hour.

$$P(X \leq 1) = 1 - e^{-1/3(1)} = 0.2835$$

R code

λ

```
> pexp(1,1/3)
[1] 0.2834687
```

- To find the probability of an exponential distribution in R: pexp(x,λ).
- To find the percentile (quantile) in R: qexp(x,λ).

Examples

Applications of the exponential distribution occurs naturally when describing the waiting time in a homogeneous Poisson process. It can be used in a range of disciplines including queuing theory, physics, reliability theory, and hydrology. Examples of events that may be modeled by exponential distribution include:

- The time until a radioactive particle decays
- The time between clicks of a Geiger counter
- The time until default on payment to company debt holders
- The distance between roadkills on a given road
- The distance between mutations on a DNA strand
- The time it takes for a bank teller to serve a customer
- The height of various molecules in a gas at a fixed temperature and pressure in a uniform gravitational field
- The monthly and annual maximum values of daily rainfall and river discharge volumes

Example from Quiz 7

1. Suppose the time a child spends waiting at for the bus as a school bus stop is exponentially distributed with mean 6 minutes. Determine the probability that the child must wait at least 9 minutes on the bus on a given morning.

$\mu = 6 \quad \lambda = 1/6$

$P(X \geq 9) = 1 - P(X \leq 9)$

```
> 1-pexp(9,1/6)
[1] 0.2231302
```

2. Suppose the time a child spends waiting at for the bus as a school bus stop is exponentially distributed with mean 4 minutes. Determine the probability that the child must wait between 3 and 6 minutes on the bus on a given morning.

$\mu = 4 \quad \lambda = 1/4$

```
> pexp(6,1/4)-pexp(3,1/4)
[1] 0.2492364
```

$P(3 < X < 6) = P(X \leq 6) - P(X \leq 3)$

The "Memoryless" Property

$P(X \geq x) = 1 - [1 - e^{-\lambda x}]$
$= \exp(-\lambda x)$

Another application of the exponential distribution is to model the distribution of component lifetime.

- Suppose component lifetime is exponentially distributed with parameter λ.
- After putting the component into service, we leave for a period of t_0 hours and then return to find the components still working; what now is the probability that it last at least an addition t hours?
- We want to find $P(X \geq t + t_0 | X \geq t_0)$

$$= \frac{P(X \geq t + t_0 \cap X \geq t_0)}{P(X \geq t_0)}$$

$$= \frac{P(X \geq t + t_0)}{P(X \geq t_0)} = P(X \geq t)$$

The Gamma Function

The gamma function $\Gamma(\alpha)$ is defined by:

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

Properties of the Gamma Function

The most important properties of the gamma function are the following:

1. For any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
2. For any positive integer, n, $\Gamma(n) = (n - 1)!$
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

The PDF of a Gamma Distribution

A continuous random variable X is said to have a **gamma distribution** if the pdf of X is

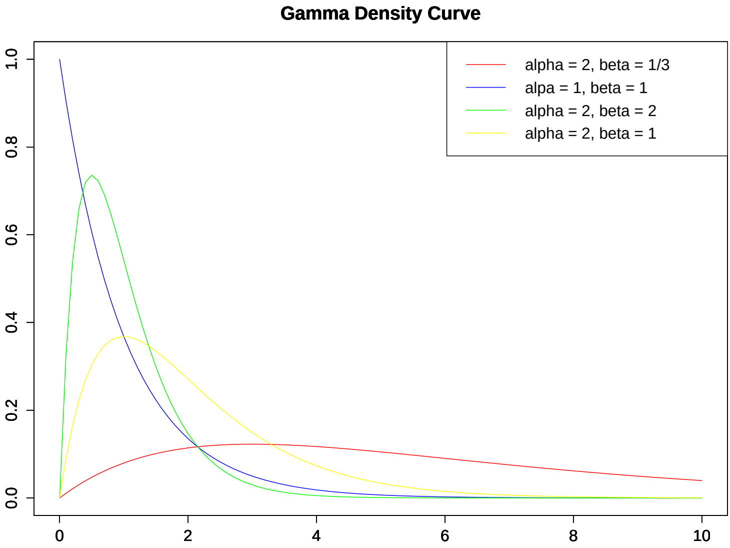
$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \geq 0 \\ 0 & \textit{otherwise} \end{cases}$$

where parameters α and β satisfy $\alpha > 0$, $\beta > 0$.

Gamma Distribution Related to the Poisson

- Gamma distribution is a distribution that arises naturally in processes for which the waiting times between events are relevant.
- It can be thought of as a waiting time between Poisson distributed events, unitl k arrivals.
- Thus the scale parameter can also be thought of as the inverse of the rate parameter (μ), $\frac{1}{\mu}$.
- Then $\alpha = k$ and $\beta = \frac{1}{\mu}$
- In R, $P(X \leq x) = pgamma(x, \alpha, \frac{1}{\beta})$

Gamma Density Curve



Applications of the Gamma Distribution

The gamma distribution can be used a range of disciplines including queuing models, climatology, and financial services. Examples of events that may be modeled by gamma distribution include:

- The amount of rainfall accumulated in a reservoir
- The size of loan defaults or aggregate insurance claims
- The flow of items through manufacturing and distribution processes
- The load on web servers
- The many and varied forms of telecom exchange

Example

Suppose that the telephone calls arriving at a particular switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse until 2 calls have come in to the switchboard?

- Average of 5 calls coming per minute means that $\beta = \frac{1}{5}$.
- Until 2 calls have come into the switchboard means that $\alpha = 2$.

$P(X \leq 1) = \text{pgamma}(1, 2, 1/(1/5))$

```
> pgamma(1,2,1/(1/5))
[1] 0.9595723
```

Mean and Variance of the Gamma Distribution

The mean and variance of a random variable X having the gamma distribution are:

$$E(X) = \mu = \alpha\beta$$
$$Var(X) = \sigma^2 = \alpha\beta^2$$

Example of Gamma Distribution

Suppose that a transistor of a certain type is subjected to an accelerated life test, the lifetime Y (in weeks) has a gamma distribution with a mean of 24 and a standard deviation of 12.

1. Find the values of α and β.

$E(x) = 24 \quad SD(x) = 12 \quad Var(x) = 144$

$E(x) = \alpha\beta$

$24 = \alpha\beta$

$\alpha = 24/\beta$

$Var(x) = \alpha\beta^2$

$144 = \alpha\beta^2$

$144 = \frac{24}{\beta}\beta^2$

$\beta = 6$

$\alpha = 4$

2. Find $P(Y \leq 24)$

$P(Y \leq 24) = \text{pgamma}(24, 4, 1/6)$

```
> pgamma(24,4,1/6)
[1] 0.5665299
```