

4

Continuous Random Variables and Probability Distributions

Stat 4570/5570

Material from Devore's book (Ed 8) – Chapter 4 - and Cengage

Continuous r.v.

A random variable X is **continuous** if possible values comprise either a single interval on the number line or a union of disjoint intervals.

Example: If in the study of the ecology of a lake, X , the r.v. may be depth measurements at randomly chosen locations.

Then X is a continuous r.v. The range for X is the minimum depth possible to the maximum depth possible.

Continuous r.v.

In principle variables such as height, weight, and temperature are continuous, in practice the limitations of our measuring instruments restrict us to a discrete (though sometimes very finely subdivided) world.

However, continuous models often approximate real-world situations very well, and continuous mathematics (calculus) is frequently easier to work with than mathematics of discrete variables and distributions.

Probability Distributions for Continuous Variables

Suppose the variable X of interest is the depth of a lake at a randomly chosen point on the surface.

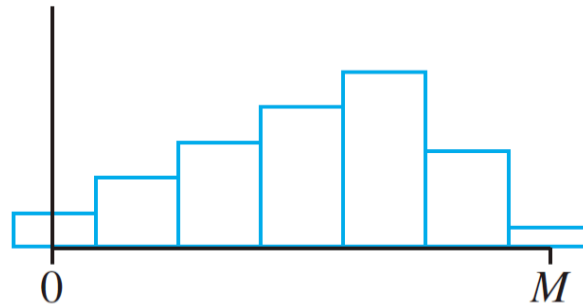
Let M = the maximum depth (in meters), so that any number in the interval $[0, M]$ is a possible value of X .

If we “discretize” X by measuring depth to the nearest meter, then possible values are nonnegative integers less than or equal to M .

The resulting discrete distribution of depth can be pictured using a probability histogram.

Probability Distributions for Continuous Variables

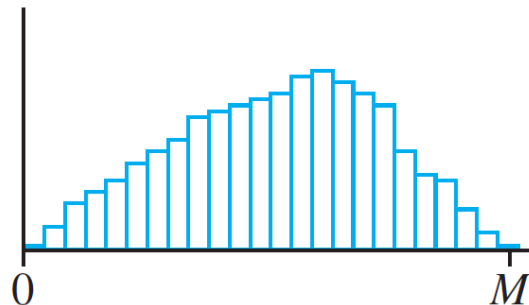
If we draw the histogram so that the area of the rectangle above any possible integer k is the proportion of the lake whose depth is (to the nearest meter) k , then the total area of all rectangles is 1:



Probability histogram of depth measured to the nearest meter

Probability Distributions for Continuous Variables

If depth is measured much more accurately, each rectangle in the resulting probability histogram is much narrower, though the total area of all rectangles is still 1.

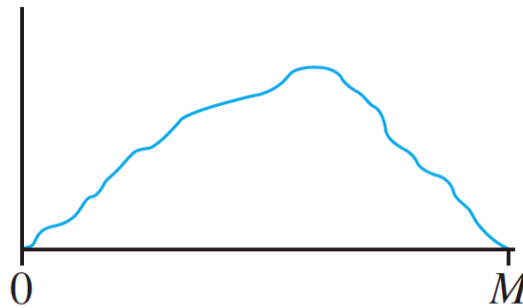


Probability histogram of depth measured to the nearest centimeter

Probability Distributions for Continuous Variables

If we continue in this way to measure depth more and more finely, the resulting sequence of histograms approaches a smooth curve.

Because for each histogram the total area of all rectangles equals 1, the total area under the smooth curve is also 1.



A limit of a sequence of discrete histograms

Probability Distributions for Continuous Variables

Definition

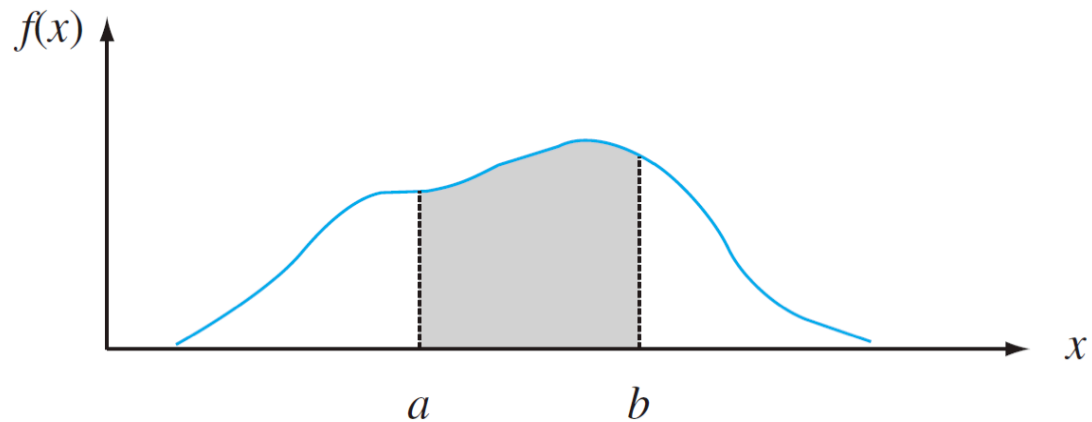
Let X be a continuous r.v. Then a **probability distribution** or **probability density function** (pdf) of X is a function $f(x)$ such that for any two numbers a and b with $a \leq b$, we have

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

The probability that X is in the interval $[a, b]$ can be calculated by integrating the pdf of the r.v. X .

Probability Distributions for Continuous Variables

The probability that X takes on a value in the interval $[a, b]$ is the area above this interval and under the graph of the density function:



$P(a \leq X \leq b)$ = the area under the density curve between a and b

Probability Distributions for Continuous Variables

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

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

2. $\int_{-\infty}^{\infty} f(x) dx = \text{area under the entire graph of } f(x) = 1$

Example

Consider the reference line connecting the valve stem on a tire to the center point.

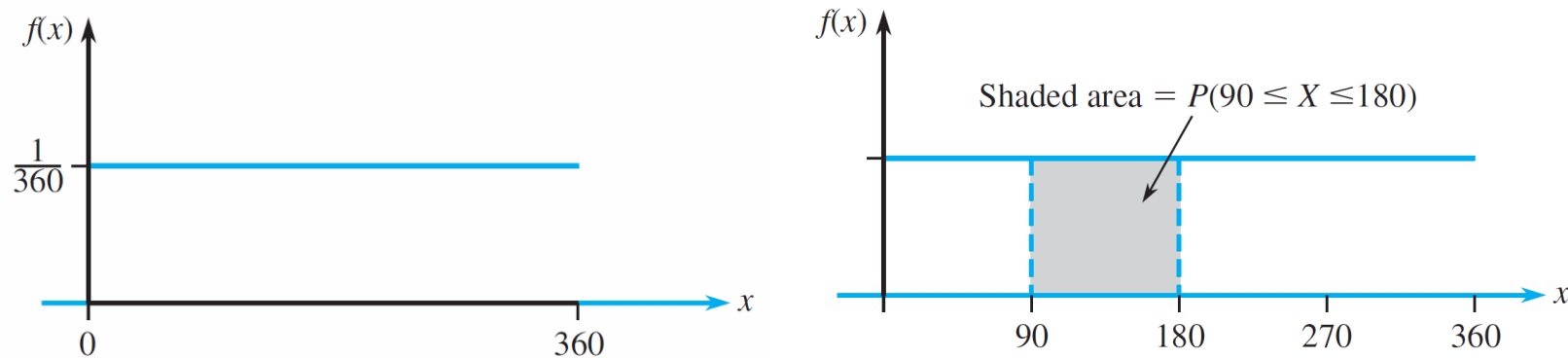
Let X be the angle measured clockwise to the location of an imperfection. One possible pdf for X is

$$f(x) = \begin{cases} \frac{1}{360} & 0 \leq x < 360 \\ 0 & \text{otherwise} \end{cases}$$

Example, cont

cont' d

The pdf is graphed in Figure 4.3.



The pdf and probability from Example 4

Figure 4.3

Example , cont

cont' d

Clearly $f(x) \geq 0$. How can we show that the area of this pdf is equal to 1?

How do we calculate $P(90 \leq X \leq 180)$?

The probability that the angle of occurrence is within 90° of the reference line (reference line is at 0 degrees)?

Probability Distributions for Continuous Variables

Because whenever $0 \leq a \leq b \leq 360$ in Example 4.4 and $P(a \leq X \leq b)$ depends only on the width $b - a$ of the interval, X is said to have a **uniform distribution**.

Definition

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

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

Example

“Time headway” in traffic flow is the elapsed time between the time that one car finishes passing a fixed point and the instant that the next car begins to pass that point.

Let X = the time headway for two randomly chosen consecutive cars on a freeway during a period of heavy flow.

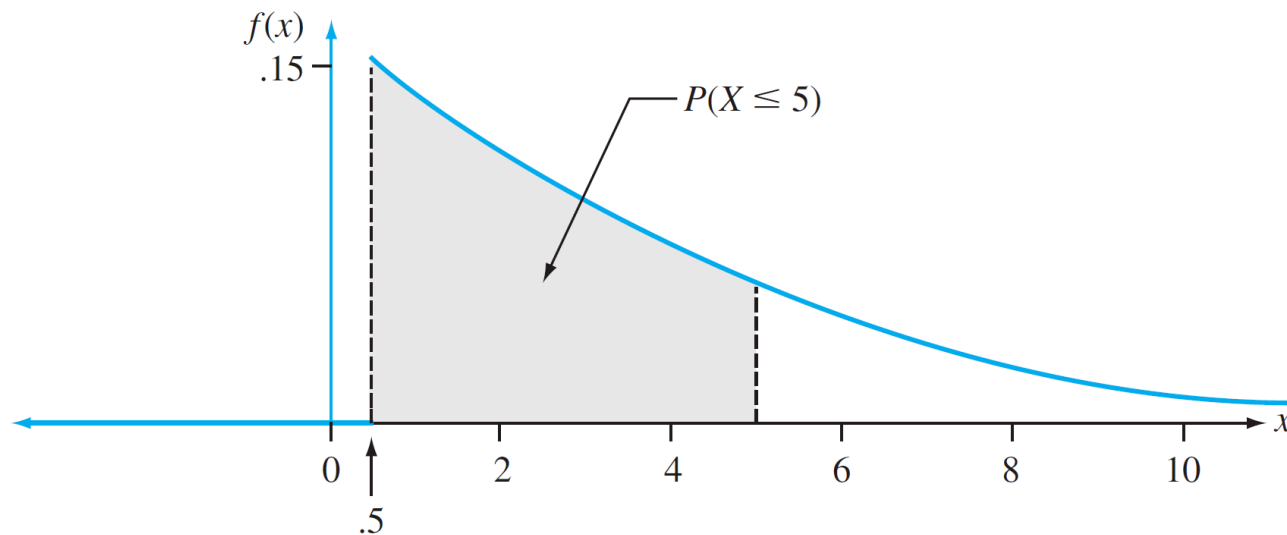
$$f(x) = \begin{cases} .15e^{-.15(x-.5)} & x \geq .5 \\ 0 & \text{otherwise} \end{cases}$$

This pdf of X is essentially the one suggested in “The Statistical Properties of Freeway Traffic” (*Transp. Res.*, vol. 11: 221–228):

Example , cont

cont' d

The graph of $f(x)$ is given in Figure 4.4; there is no density associated with headway times less than .5, and headway density decreases rapidly (exponentially fast) as x increases from .5.



The density curve for time headway in Example 5

Figure 4.4

Example

cont' d

Clearly, $f(x) \geq 0$; and

$$\begin{aligned}\int_{-\infty}^{\infty} f(x) dx &= \int_{-\infty}^{0.5} 0 dx + \int_{0.5}^{\infty} 0.15e^{-0.15(x-0.5)} dx \\ &= 0.15e^{0.075} \int_{0.5}^{\infty} e^{-0.15x} dx = 1\end{aligned}$$

What is the probability that headway time is at most 5 seconds?

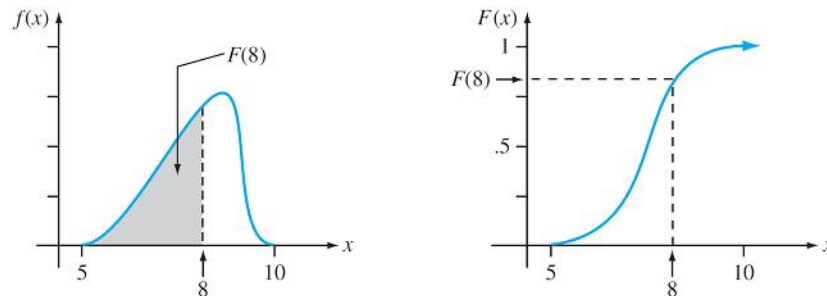
Integrating functions in R

The Cumulative Distribution Function

The **cumulative distribution function** $F(x)$ for a continuous rv X is defined for every number x by

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y) dy$$

For each x , $F(x)$ is the area under the density curve to the left of x . This is illustrated in Figure 4.5, where $F(x)$ increases smoothly as x increases.



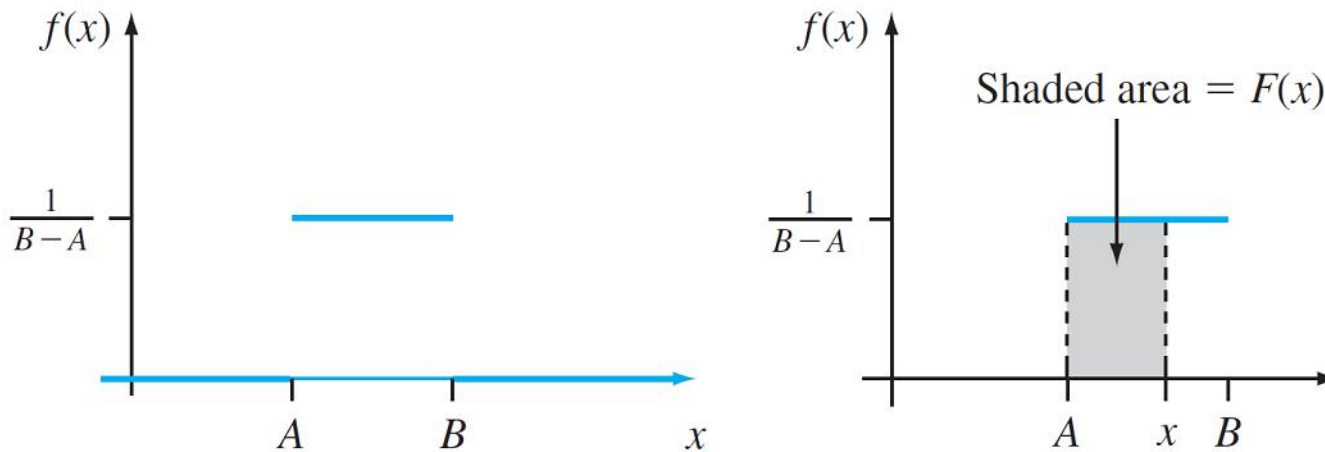
A pdf and associated cdf

Figure 4.5

Example

Let X , the thickness of a certain metal sheet, have a uniform distribution on $[A, B]$.

The density function is shown in Figure 4.6.



The pdf for a uniform distribution

Figure 4.6

Example , cont

cont' d

For $x < A$, $F(x) = 0$, since there is no area under the graph of the density function to the left of such an x .

For $x \geq B$, $F(x) = 1$, since all the area is accumulated to the left of such an x . Finally for $A \leq x \leq B$,

$$F(x) = \int_{-\infty}^x f(y)dy = \int_A^x \frac{1}{B - A} dy = \frac{1}{B - A} \cdot y \Big|_{y=A}^{y=x} = \frac{x - A}{B - A}$$

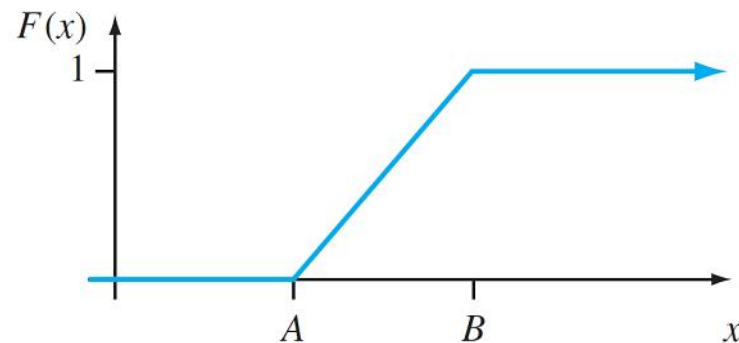
Example , cont

cont' d

The entire cdf is

$$F(x) = \begin{cases} 0 & x < A \\ \frac{x - A}{B - A} & A \leq x < B \\ 1 & x \geq B \end{cases}$$

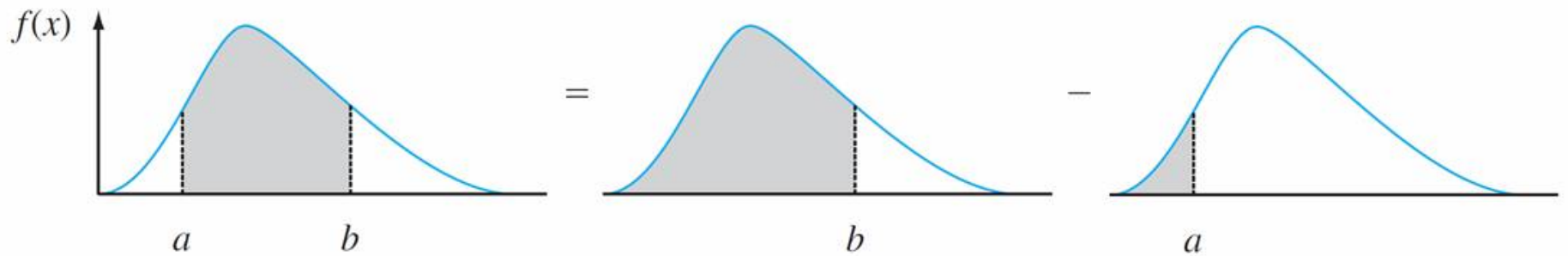
The graph of this cdf appears in Figure 4.7.



The cdf for a uniform distribution

Figure 4.7

Using $F(x)$ to Compute Probabilities



Percentiles of a Continuous Distribution

When we say that an individual's test score was at the 85th **percentile of the population**, we mean that 85% of all population scores were below that score and 15% were above.

Similarly, the 40th percentile is the score that exceeds 40% of all scores and is exceeded by 60% of all scores.

Percentiles of a Continuous Distribution

Proposition

Let p be a number between 0 and 1. The **(100 p)th percentile** of the distribution of a continuous rv X , denoted by $\eta(p)$, is defined by

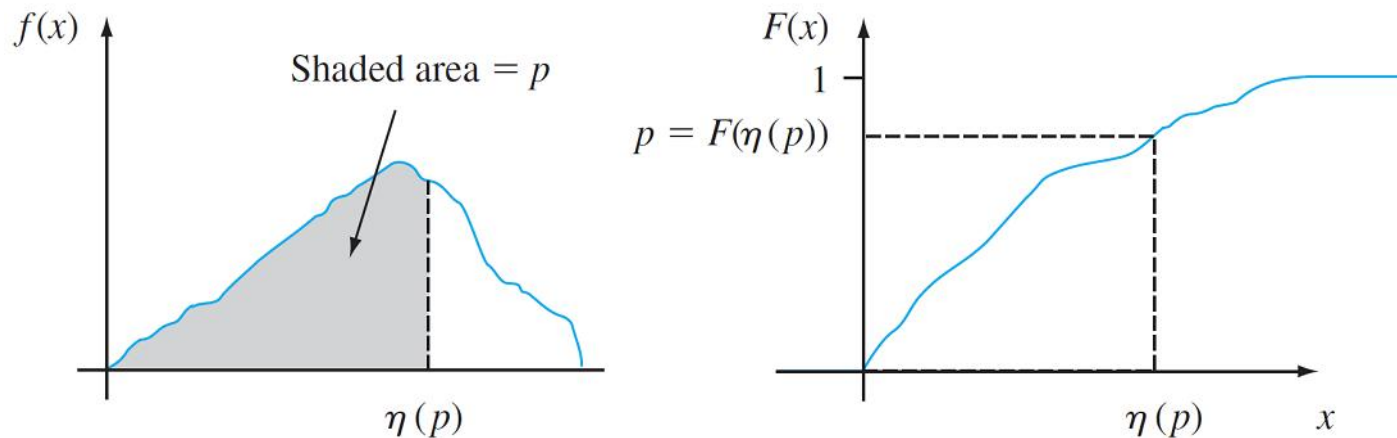
$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y) dy$$

$\eta(p)$ is the specific value such that 100 p % of the area under the graph of $f(x)$ lies to the left of $\eta(p)$ and 100(1 – p)% lies to the right.

Percentiles of a Continuous Distribution

Thus $\eta(.75)$, the 75th percentile, is such that the area under the graph of $f(x)$ to the left of $\eta(.75)$ is .75.

Figure 4.10 illustrates the definition.



The $(100p)$ th percentile of a continuous distribution

Figure 4.10

Example 9

The distribution of the amount of gravel (in tons) sold by a particular construction supply company in a given week is a continuous rv X with pdf

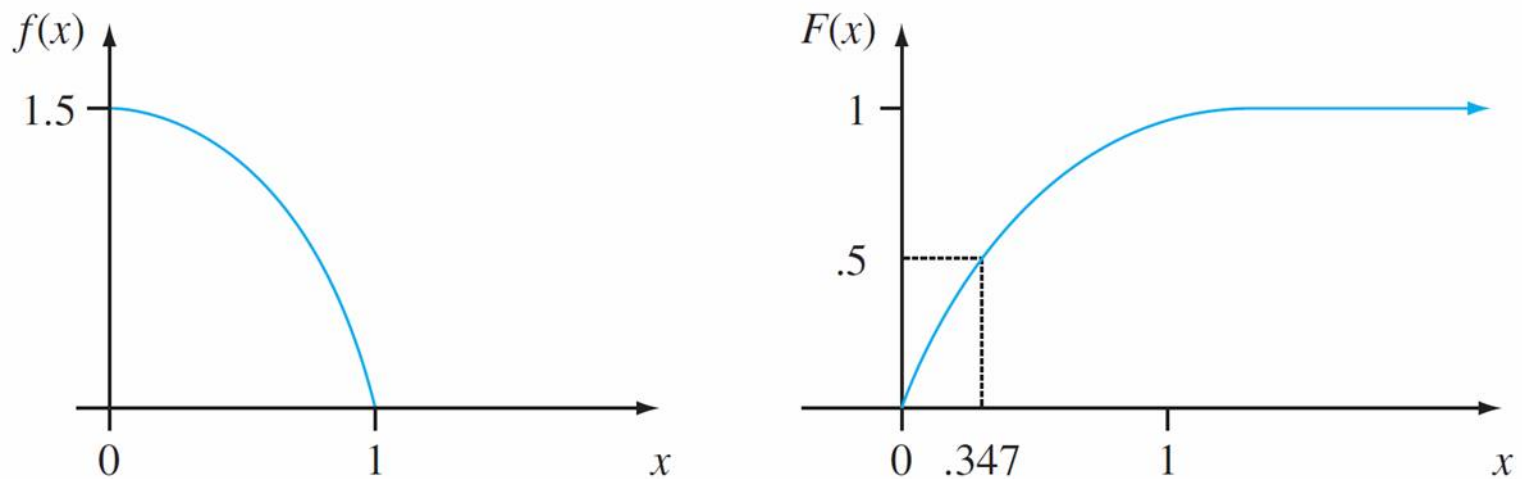
$$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is the cdf of sales for any x ? How do you use this to find the probability that X is less than .25? What about the probability that X is greater than .75? What about $P(.25 < X < .75)$?

Example 9

cont' d

The graphs of both $f(x)$ and $F(x)$ appear in Figure 4.11.



The pdf and cdf for Example 4.9

Figure 4.11

Example 9

cont' d

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

How do we find the 40th percentile of this distribution?

How would you find the median of this distribution?

Percentiles of a Continuous Distribution

Definition

The **median** of a continuous distribution, denoted by $\tilde{\mu}$, is the 50th percentile, so $\tilde{\mu}$ satisfies $.5 = F(\tilde{\mu})$. That is, half the area under the density curve is to the left of $\tilde{\mu}$ and half is to the right of $\tilde{\mu}$. The 25th percentile is called the **lower quartile** and the 75th percentile is called the **upper quartile**.

Expected Values

Definition

The **expected** (or **mean**) **value** of a continuous r.v. X with the pdf $f(x)$ is:

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

Example, cont

Back to the gravel example, the pdf of the amount of weekly gravel sales X is:

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

What is EX ?

Expected Values of functions of r.v.

If $h(X)$ is a function of X , then

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

For $h(X) = aX + b$, a linear function,

$$E[h(X)] = E[aX + b] = aE[X] + b$$

Variance

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

$$\begin{aligned}\sigma_X^2 = V(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx \\ &= E[(X - E(X))^2] \\ &= E(X^2) - E(X)^2\end{aligned}$$

The **standard deviation** (SD) of X is

$$\sigma_X = \sqrt{V(X)}$$

When $h(X) = aX + b$,

$$V[h(X)] = V[aX + b] = a^2 \cdot \sigma_X^2 \text{ and } \sigma_{aX+b} = |a| \cdot \sigma_X$$

Example, cont.

How do we compute the variance for the weekly gravel sales example?

The Normal Distribution

The normal distribution is probably the most important distribution in all of probability and statistics.

Many populations have distributions that can be fit very closely by an appropriate normal (or Gaussian, bell) curve.

Examples: height, weight, and other physical characteristics, scores on various tests, etc.

The Normal Distribution

Definition

A continuous r.v. X is said to have a **normal distribution** with parameters μ and $\sigma > 0$ (or μ and σ^2), if the pdf of X is

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \text{ where } -\infty < x < \infty$$

- e has approximate value 2.71828
- π has approximate value 3.14159.

The Normal Distribution

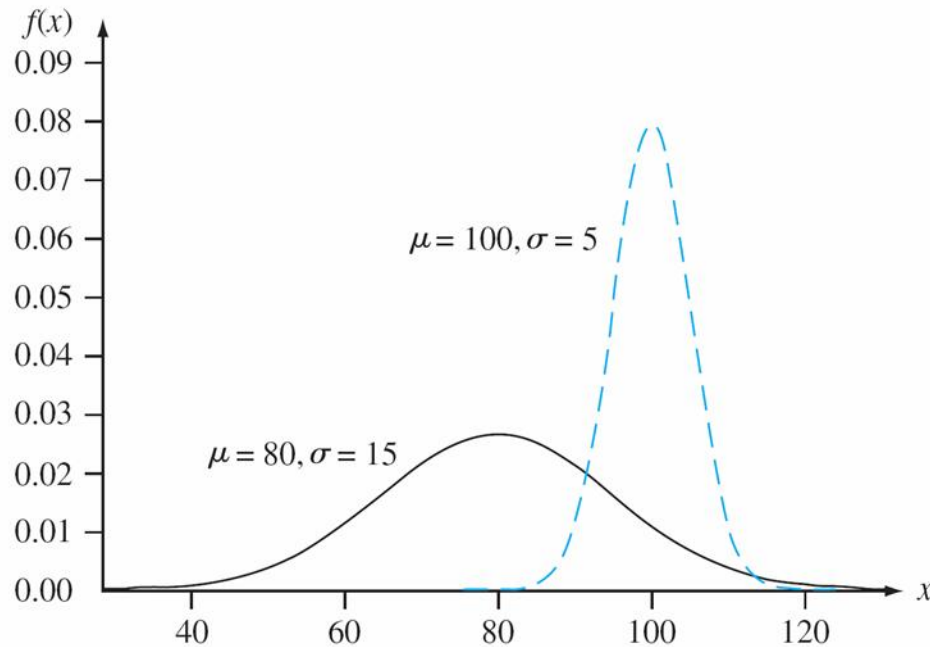
The statement that X is normally distributed with parameters μ and σ^2 is often abbreviated $X \sim N(\mu, \sigma^2)$.

Clearly $f(x; \mu, \sigma) \geq 0$, but a somewhat complicated calculus argument must be used to verify that $\int_{-\infty}^{\infty} f(x; \mu, \sigma) dx = 1$.

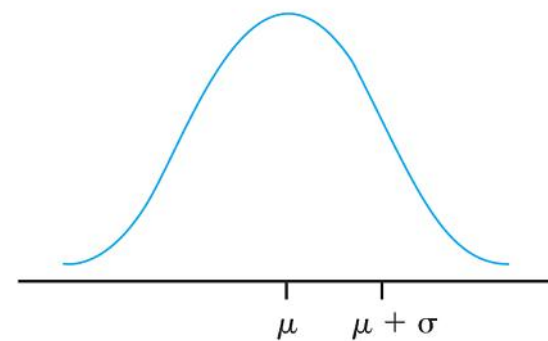
Similarly, it can be shown that $E(X) = \mu$ and $V(X) = \sigma^2$, so the parameters are the mean and the standard deviation of X .

The Normal Distribution

Figure below presents graphs of $f(x; \mu, \sigma)$ for several different (μ, σ) pairs.



Two different normal density curves



Visualizing μ and σ for a normal distribution

The Standard Normal Distribution

The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution**.

A r.v. with this distribution is called a **standard normal random variable** and is denoted by Z . Its pdf is:

$$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \text{ where } -\infty < z < \infty$$

The graph of $f(z; 0, 1)$ is called the *standard normal curve*.

Its inflection points are at 1 and -1 . The cdf of the standard normal is

$$\Phi(z) = \int_{-\infty}^z f(y; 0, 1) dy$$

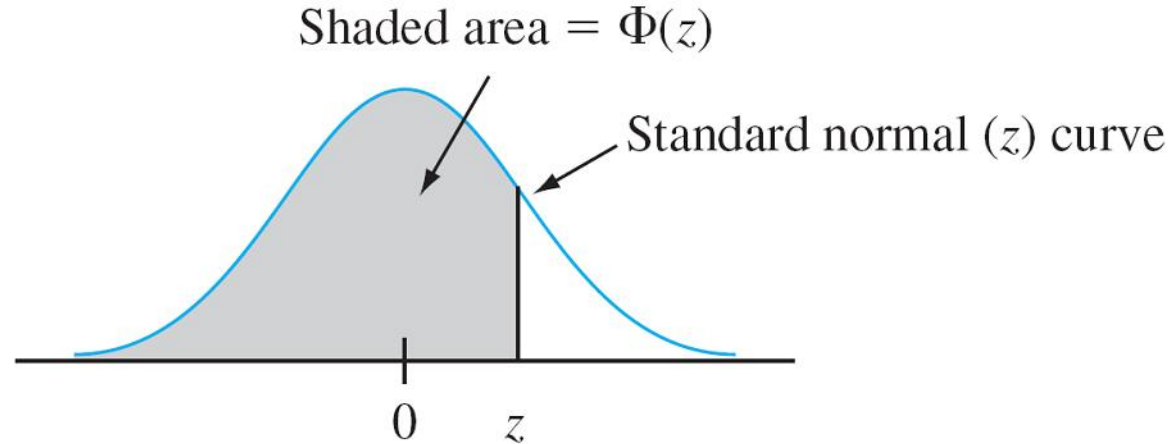
note we use the special notation $\Phi(z)$ for the cdf of Z .

The Standard Normal Distribution

- The standard normal distribution rarely occurs naturally.
- Instead, it is a **reference distribution** from which information about other normal distributions can be obtained via a simple formula.
- These probabilities can then be found “normal tables”.
- This can also be computed with a single command in R.

The Standard Normal Distribution

Figure below illustrates the probabilities tabulated in Table A.3:



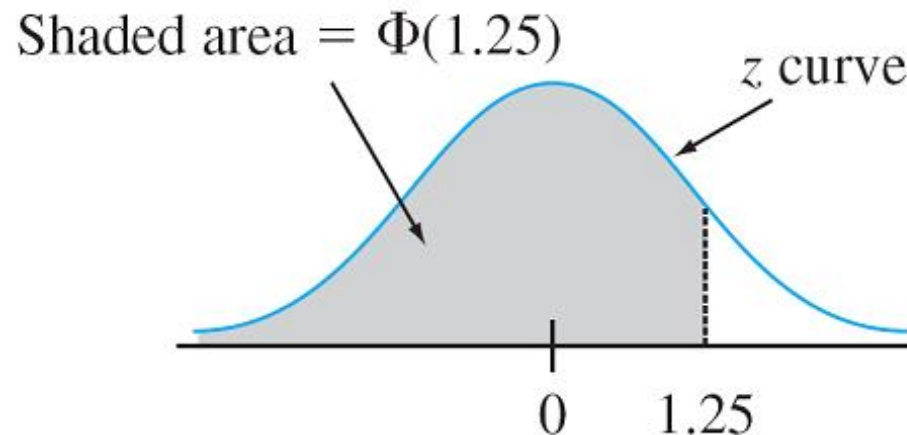
Example

cont' d

$P(Z \leq 1.25) = \Phi(1.25)$, a probability that is tabulated in a normal table.

What is this probability?

Figure below illustrates this probability:



Example , cont.

cont' d

b) $P(Z \geq 1.25) = ?$

c) Why does $P(Z \leq -1.25) = P(Z \geq 1.25)$? What is $\Phi(-1.25)$?

d) How do we calculate $P(-.38 \leq Z \leq 1.25)$?

Example

The 99th percentile of the standard normal distribution is that value of z such that the area under the z curve to the left of the value is .99

Tables give for fixed z the area under the standard normal curve to the left of z , whereas now – we have the area and want the value of z .

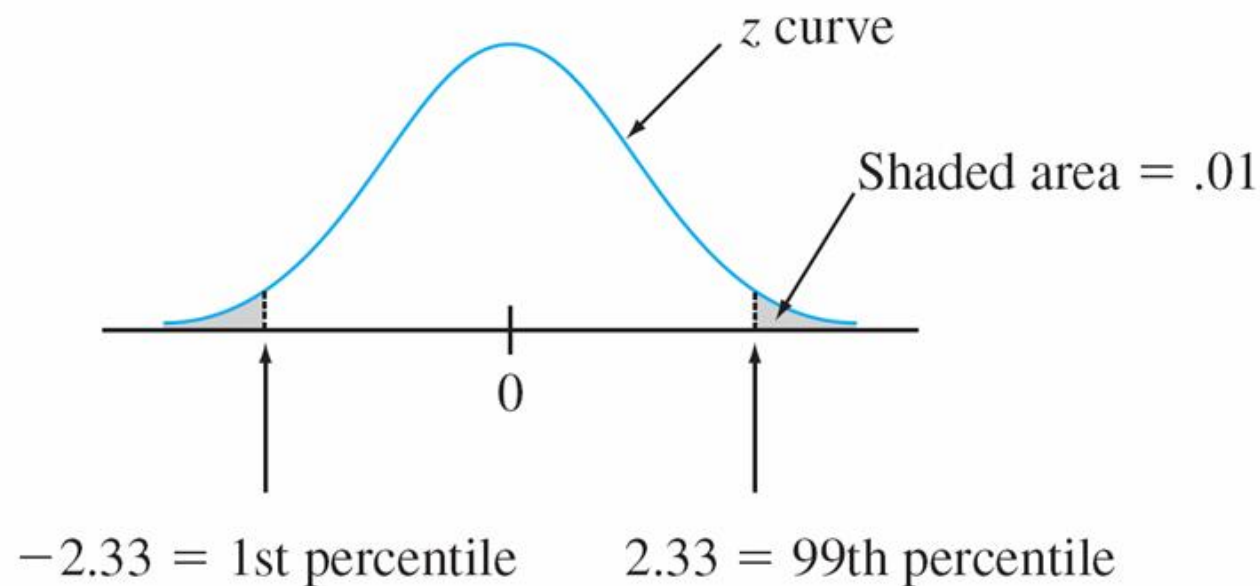
This is the “inverse” problem to $P(Z \leq z) = ?$

How can the table be used for this?

Example

cont' d

By symmetry, the first percentile is as far below 0 as the 99th is above 0, so equals -2.33 (1% lies below the first and also above the 99th).



Percentiles of the Standard Normal Distribution

If p does not appear in a table, what can we do?

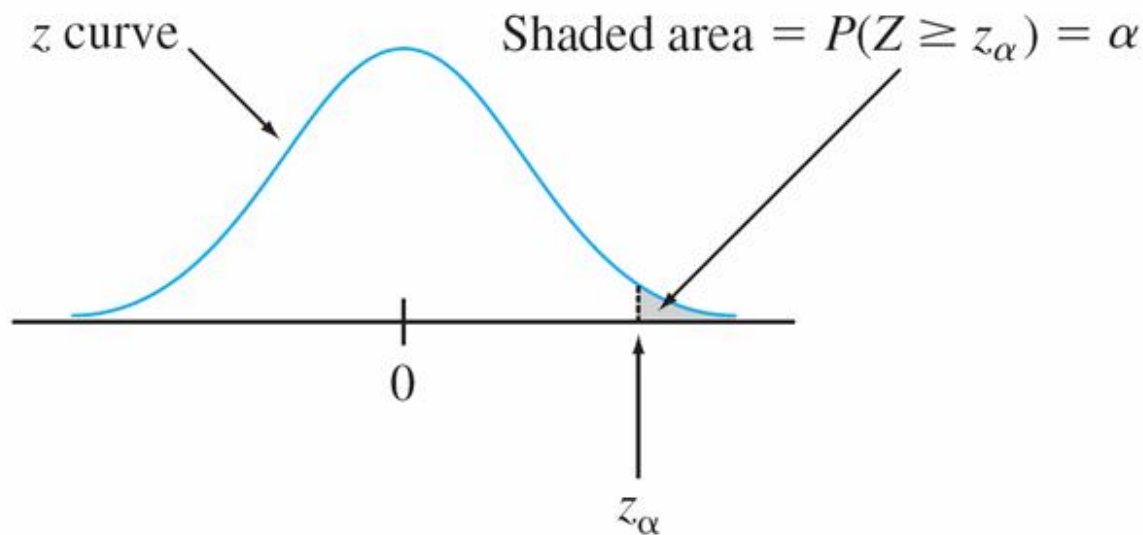
What is the 95th percentile of the normal distribution?

z_α Notation

In statistical inference, we need the z values that give certain tail areas under the standard normal curve.

There, this notation will be standard:

z_α will denote the z value for which α of the area under the z curve lies to the right of z_α .



z_α Notation for z Critical Values

For example, $z_{.10}$ captures upper-tail area .10, and $z_{.01}$ captures upper-tail area .01.

Since α of the area under the z curve lies to the right of z_α , $1 - \alpha$ of the area lies to its left.

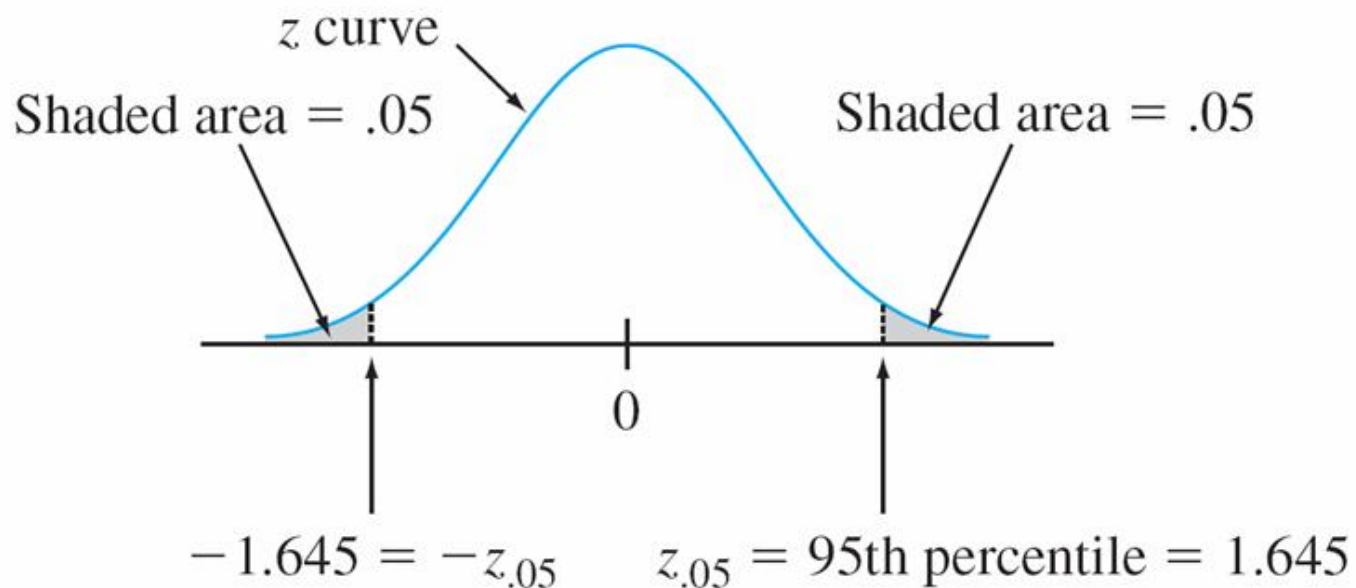
Thus z_α is the $100(1 - \alpha)$ th percentile of the standard normal distribution.

Similarly, what does $-z_\alpha$ mean?

Example – critical values

$z_{.05}$ is the $100(1 - .05)$ th = 95th percentile of the standard normal distribution, so $z_{.05} = 1.645$.

The area under the standard normal curve to the left of $-z_{.05}$ is also .05



Nonstandard Normal Distributions

When $X \sim N(\mu, \sigma^2)$, probabilities involving X are computed by “standardizing.” The **standardized variable** is $(X - \mu)/\sigma$.

Subtracting μ shifts the mean from μ to zero, and then dividing by σ scales the variable so that the standard deviation is 1 rather than σ .

Proposition

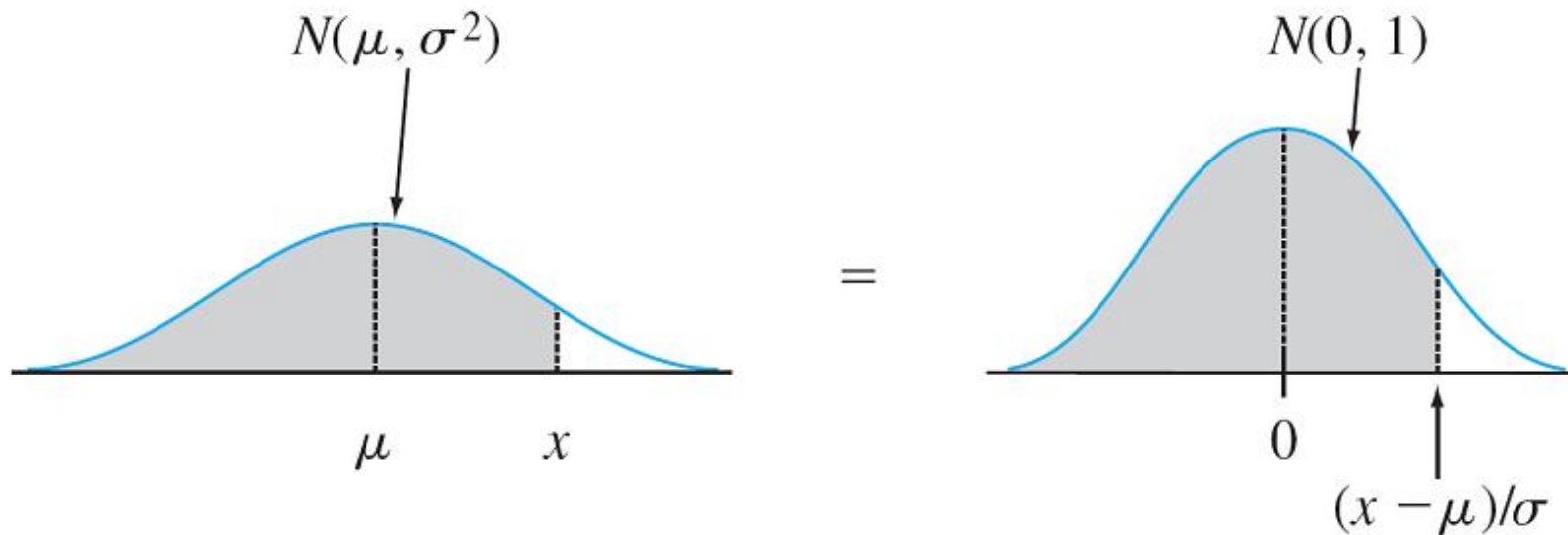
If X has a normal distribution with mean μ and standard deviation σ , then

$$Z = \frac{X - \mu}{\sigma}$$

is distributed standard normal.

Nonstandard Normal Distributions

Why do we standardize normal random variables?



Equality of nonstandard and standard normal curve areas

Nonstandard Normal Distributions

has a standard normal distribution. Thus

$$\begin{aligned}P(a \leq X \leq b) &= P\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)\end{aligned}$$

$$P(X \leq a) = \Phi\left(\frac{a - \mu}{\sigma}\right) \quad P(X \geq b) = 1 - \Phi\left(\frac{b - \mu}{\sigma}\right)$$



THE NORMAL DISTRIBUTION IN R.



Exponential Distribution

The Exponential Distributions

The family of exponential distributions provides probability models that are very widely used in engineering and science disciplines. Examples?

Definition

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

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

The Exponential Distributions

Integration by parts give the following results:

$$\mu = \frac{1}{\lambda} \quad \sigma^2 = \frac{1}{\lambda^2}$$

How would we set up the calculation for EX?

Both the mean and standard deviation of the exponential distribution equal $1/\lambda$.

CDF:

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

The Exponential Distributions

Another important application of the exponential distribution is to model the distribution of lifetimes, or **times to an event**.

A partial reason for the popularity of such applications is the “**memoryless**” **property** of the Exponential distribution.

The Exponential Distributions

Suppose a light bulb's lifetime is exponentially distributed with parameter λ .

Say you turn the light on, and then we leave and come back after t_0 hours to find it still on. What is the probability that the light bulb will last for at least additional t hours?

In symbols, we are looking for $P(X \geq t + t_0 \mid X \geq t_0)$.

How would we calculate this?

The Chi-Squared Distribution

Definition

Let ν be a positive integer. Then a random variable X is said to have a **chi-squared distribution** with parameter ν if the pdf of X is the **gamma density with $\alpha = \nu/2$ and $\beta = 2$** . The pdf of a chi-squared rv is thus

$$f(x; \nu) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4.10)$$

The parameter is called the **number of degrees of freedom** (df) of X . The symbol χ^2 is often used in place of “chi-squared.”

The Weibull Distribution

The Weibull distribution is used similarly to the exponential distribution to model times to an event, but with an extra parameter included for flexibility.

Definition

A random variable X is said to have a **Weibull distribution** with parameters α and β ($\alpha > 0$, $\beta > 0$) if the pdf of X is

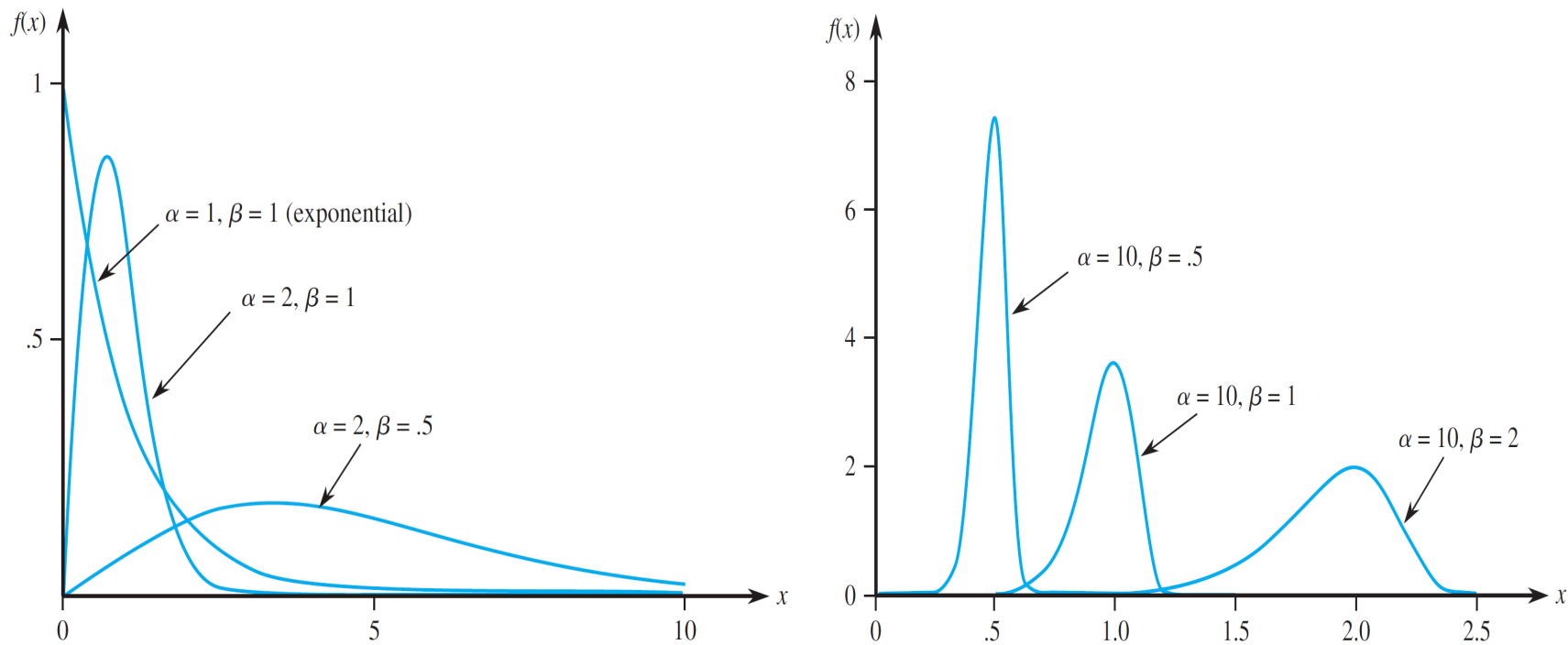
$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is this distribution if alpha = 1?

(4.11)

The Weibull Distribution

Both α and β can be varied to obtain a number of different-looking density curves, as illustrated below.



Weibull density curves

The Weibull Distribution

β is called a scale parameter, since different values stretch or compress the graph in the x direction, and α is referred to as a shape parameter.

β Integrating to obtain $E(X)$ and $E(X^2)$ yields

$$\mu = \beta \Gamma\left(1 + \frac{1}{\alpha}\right) \quad \sigma^2 = \beta^2 \left\{ \Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \right]^2 \right\}$$

The computation of μ and σ^2 requires use of the gamma function.

The Beta Distribution

So far, all families of continuous distributions (except for the uniform distribution) had positive density over an infinite interval.

The beta distribution provides positive density only for X in an interval of finite length $[A,B]$.

The standard beta distribution is commonly used to model variation in the **proportion or percentage** of a quantity occurring in different samples. Examples?

The Beta Distribution

Definition

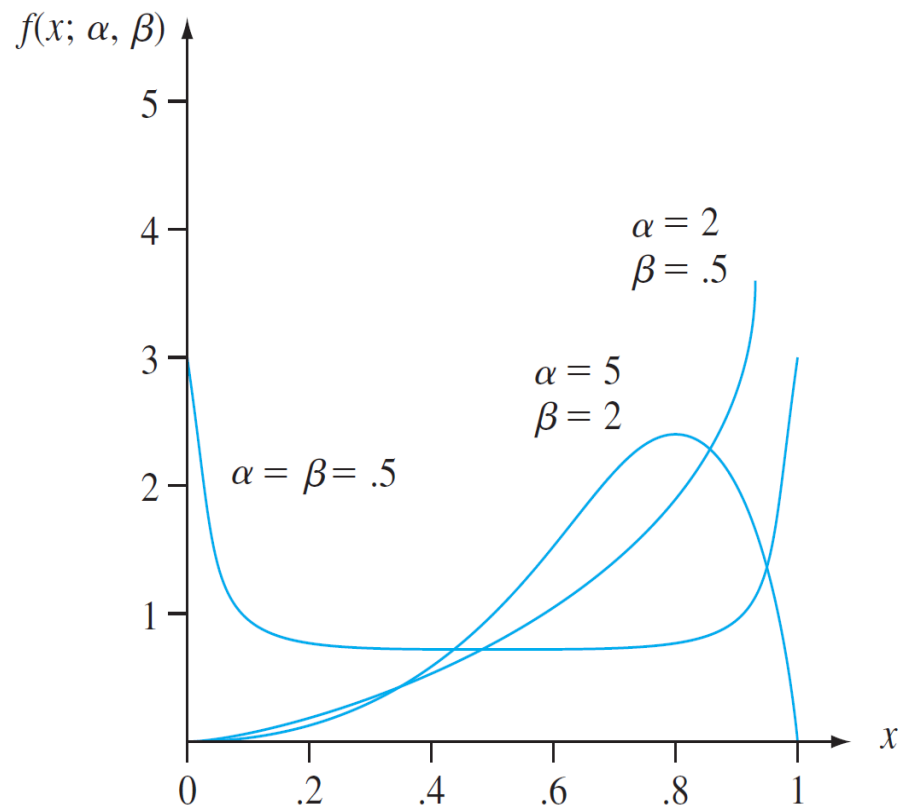
A random variable X is said to have a **beta distribution** with parameters α , β (both positive), A , and B if the pdf of X is

$$f(x; \alpha, \beta, A, B) = \begin{cases} \frac{1}{B - A} \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\frac{x - A}{B - A}\right)^{\alpha-1} \left(\frac{B - x}{B - A}\right)^{\beta-1} & A \leq x \leq B \\ 0 & \text{otherwise} \end{cases}$$

The case $A = 0$, $B = 1$ gives the **standard beta distribution**.

The Beta Distribution

Figure below illustrates several standard beta pdf' s.



Standard beta density curves

The Beta Distribution

Graphs of the general pdf are similar, except they are shifted and then stretched or compressed to fit over $[A, B]$.

Unless α and β are integers, integration of the pdf to calculate probabilities is difficult. Either a table of the incomplete beta function or appropriate software should be used.

The mean and variance of X are

$$\mu = A + (B - A) \cdot \frac{\alpha}{\alpha + \beta} \quad \sigma^2 = \frac{(B - A)^2 \alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

Example 1

Suppose the pdf of the magnitude X of a dynamic load on a bridge (in newtons) is

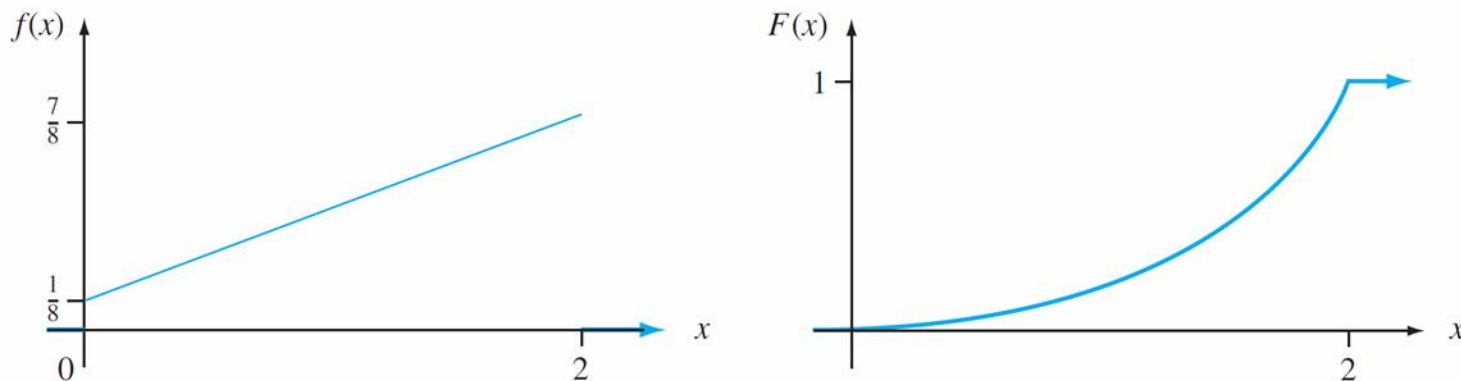
$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

What are $F(x)$, $E(X)$, and $\text{Var}(X)$?

Example 1

cont' d

The graphs of $f(x)$ and $F(x)$ are shown in Figure 4.9.



The pdf and cdf for Example 4.7

Example 1

cont' d

What is the probability that the load is between 1 and 1.5?

The probability that the load exceeds 1?

The average load? The median load?

Example 2

The time that it takes a driver to react to the brake lights on a decelerating vehicle is critical in helping to avoid rear-end collisions.

The article “Fast-Rise Brake Lamp as a Collision-Prevention Device” (*Ergonomics*, 1993: 391–395) suggests that reaction time for an in-traffic response to a brake signal from standard brake lights can be modeled with a normal distribution having mean value 1.25 sec and standard deviation of .46 sec.

What is the probability that reaction time is between 1.00 sec and 1.75 sec?