Statistics

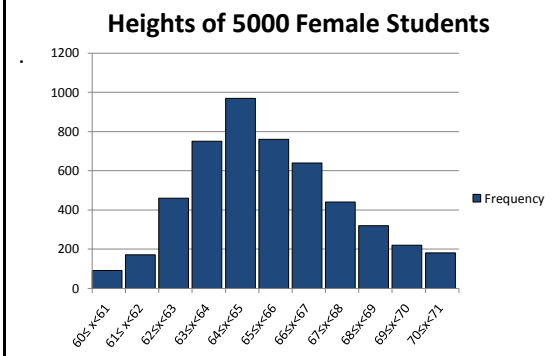
Normal distributions

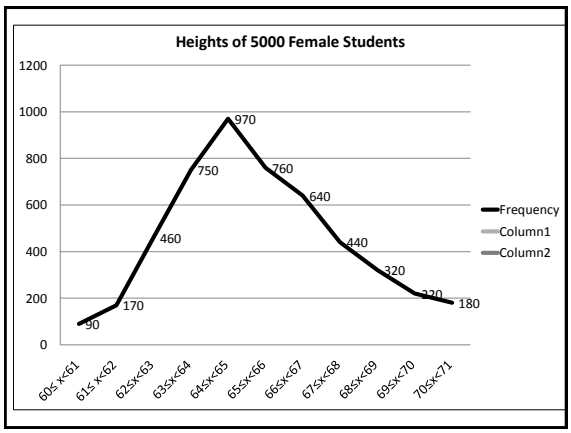
Continuous Random Variable

- **A variable whose values are not countable**
  - **Can take on any value over an interval(s)**
    - **Can assume an infinite number of values**
- **The probability that a continuous random variable assumes a single value is always 0.**

Heights of 5000 Female Students

Inches	Freq (*f*)	Relative Frequency
60 ≤ x < 61	90	.018
6 1 ≤ x < 62	170	.034
62 ≤ x < 63	460	.092
63 ≤ x < 64	750	.15
64 ≤ x < 65	970	.194
65 ≤ x < 66	760	.152
66 ≤ x < 67	640	.128
67 ≤ x < 68	440	.088
68 ≤ x < 69	320	.064
69 ≤ x < 70	220	.044
70 ≤ x < 71	180	.036

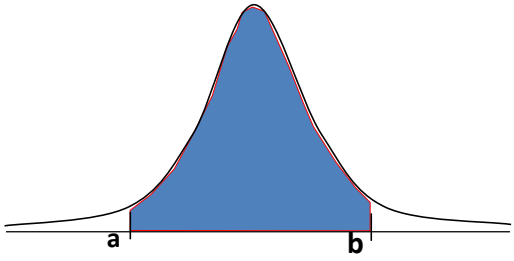




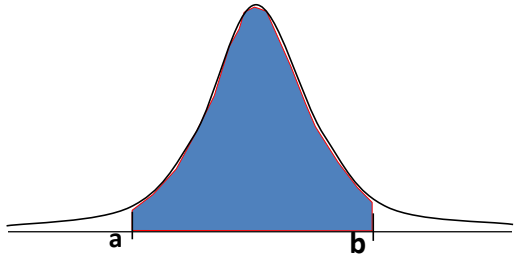
Probability Distribution of cont. ran. variable has two characteristics

- **The probability that x assumes a value in any interval lies in the range 0 to1.**
- **The total probability of all the (mutually exclusive) intervals within which x can assume a value is 1.0.**

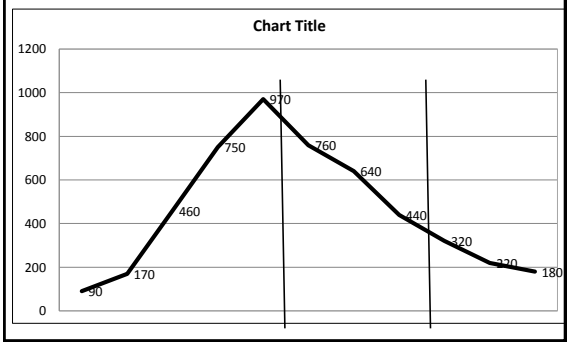
Shaded area is between 0 and 1



Shaded area gives the probability P(a≤ x ≤ b)



Shaded Area gives P(65 ≤ x ≤ 68

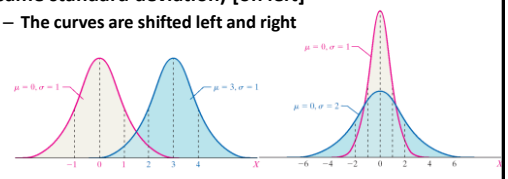


For continuous random variables

- **Probability is always calculated for an interval**
- **The probability that a crv assumes a *single* value is always zero.**
  - **The area of a line is zero**
- **In other words, P(a≤ x ≤ b) is the same as P(a< x < b) in a probability sense**

Normal Curves

- Two normal curves with different means (but the same standard deviation) [on left]
  - The curves are shifted left and right

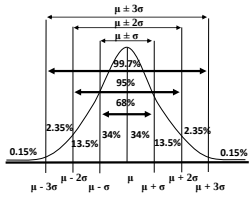


- Two normal curves with different standard deviations (but the same mean) [on right]
  - The curves are shifted up and down

Properties of the Normal Density Curve

- **It is symmetric about its mean, μ**
- **Because mean = median = mode, the highest point occurs at x = μ**
- **It has inflection points at μ – σ and μ + σ**
- **Area under the curve = 1**
- **Area under the curve to the right of μ equals the area under the curve to the left of μ, which equals ½**
- **As x increases or decreases without bound (gets farther away from μ), the graph approaches, but never reaches the horizontal axis (like approaching an asymptote)**
- **The Empirical Rule applies**

Empirical Rule



Normal Probability Density Function

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

where μ is the mean and σ is the standard deviation of the random variable x

Area under a Normal Curve

The area under the normal curve for any interval of values of the random variable X represents either

- **The proportion of the population with the characteristic described by the interval of values**

or

- **The probability that a randomly selected individual from the population will have the characteristic described by the interval of values**

[the area under the curve is either a proportion or the probability]

Key Concept

This section presents the *standard normal distribution* which has three properties:

1. It's graph is bell-shaped.

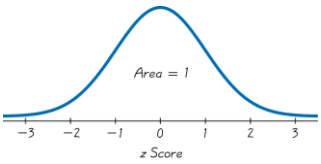
2. It's mean is equal to 0 ($\mu = 0$).

3. It's standard deviation is equal to 1 ($\sigma = 1$).

Develop the skill to find areas (or probabilities or relative frequencies) corresponding to various regions under the graph of the standard normal distribution. Find *z*-scores that correspond to area under the graph.

Standard Normal Distribution

The standard normal distribution is a normal probability distribution with $\mu = 0$ and $\sigma = 1$. The total area under its density curve is equal to 1.

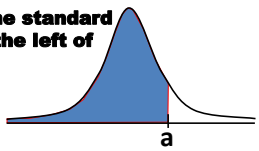


Z values or Z scores

- **The units marked on the horizontal axis of the standard normal curve**
- **A specific value of z gives the distance between the mean and the point represented by z in terms of standard deviation**
  - **z can be thought of as the number of standard deviations away from the mean.**

Example 1

Determine the area under the standard normal curve that lies to the left of



A. Z = -3.49
ANS: 0.0002
B. Z = -1.99
ANS: 0.0233
C. Z = 0.92
ANS: 0.8212
D. Z = 2.90
ANS: 0.9981

Notation

P($a < z < b$)

denotes the probability that the *z* score is between *a* and *b*.

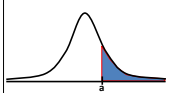
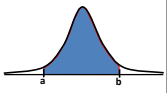
P($z > a$)

denotes the probability that the *z* score is greater than *a*.

P($z < a$)

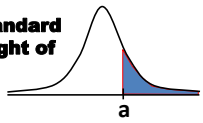
denotes the probability that the *z* score is less than *a*.

Calculating the Area under Standard Normal Curve

Approach	Graphically	Solution
Find the area to the left of z_a *P(Z < a)*	Shade the area to the left of z_a 	Use Table IV to find the row and column that correspond to z_a. The area is the value where the row and column intersect.
Find the area to the right of z_a *P(Z > a) or 1 – P(Z < a)*	Shade the area to the right of z_a 	Use Table IV to find the area to the left of z_a. The area to the right of z_a is 1 – area to the left of z_a.
Find the area between z_a and z_b *P(a < Z < b)*	Shade the area between z_a *and* z_b 	Use Table IV to find the area to the left of z_b and to the left of z_a. The area between is $area_{zb}$ – $area_{za}$.

Example 2

Determine the area under the standard normal curve that lies to the right of



a) Z = -3.49
ANS: 0.9998

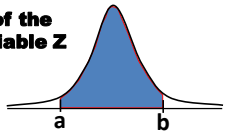
b) Z = -0.55
ANS: 0.7088

c) Z = 2.23
ANS: 0.0129

d) Z = 3.45
ANS: 0.0003

Example 3

Find the indicated probability of the standard normal random variable Z



a) P(-2.55 < Z < 2.55)
ANS: 0.9892

b) P(-0.55 < Z < 0)
ANS: 0.2088

c) P(-1.04 < Z < 2.76)
ANS: 0.8479

Key Concept

This section presents methods for working with normal distributions that are not standard.

That is, the mean is not 0 or the standard deviation is not 1, or both.

The key concept is that we can use a simple conversion that allows us to standardize any normal distribution so that the same methods of the previous section can be used.

Standardizing a Normal Random Variable

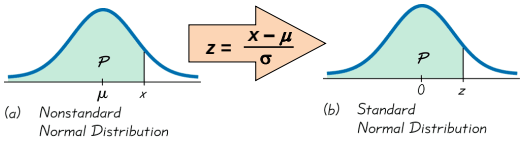
Z statistic: $z = \dfrac{x - \mu}{\sigma}$

Aka z "value" or z "score"

where μ is the mean and σ is the standard deviation of the random variable X

Z is normally distributed with mean of 0 and standard deviation of 1

Converting to a Standard Normal Distribution



(a) Nonstandard Normal Distribution (b) Standard Normal Distribution

Finding a probability i.e. Finding the Area under any Normal Curve

- **Draw a normal curve and shade the desired area**
- **Convert the values of X to Z-scores using**

$$z = \frac{x - \mu}{\sigma}$$

- **Draw a standard normal curve and shade the area desired**
- **Find the area under the standard normal curve. This area is equal to the area under the normal curve drawn in Step 1**

Example 1

A random variable x is normally distributed with μ=10 and σ=3.

What is the area between x_1 = 7 and x_2= 15?

Helpful Hints

1. Don't confuse *z* scores and areas. *z* scores are distances along the horizontal scale, but areas are regions under the normal curve. Table 4 lists *z* scores in the left column and across the top row, but areas are found in the body of the table.
2. Choose the correct (right/left) side of the graph.
3. A z score must be negative whenever it is located in the left half of the normal distribution.
4. Areas (or probabilities) are positive or zero values, but they are never negative.

Example 2

For a general random variable *X* with $\mu = 6$ and $\sigma = 4$.

Calculate $P(4 < X < 11)$

p. 274 #44

- The transmission model of a specific car has a warranty for 40,000 miles. It is known that the life of such transmissions have a normal distribution with a mean of 72,000 miles and a standard deviation of 12,000 miles.
- What percentage of transmissions will fail before the end of the warranty period?

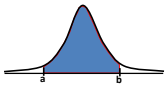
Example 4

- Fast Auto service guarantees that the maximum waiting time for its customers is 20 minutes for oil service on their cars. If also guarantees that any customer who has to wait longer than 20 min will receive a 50% discount on all charges. If is estimated that the mean time taken for oil service is 15 minutes per car with a standard deviation of 2.4 minutes. Suppose the time follows a normal distribution.
- What percentage of customers will receive a 50% discount on their service?

What if you know the area but not x or z?

- Rockport Corp. make electric shavers. The life of model S7J has a normal distribution with a mean of 70 months and a standard deviation of 8 months. The company is trying to determine the warranty period for this shaver. Any shaver needing repair during the warranty period will be replaced free of charge. What should the warranty period be if the company does not want to replace more than 1% of all shavers?

Example 6



For random variable X with

$\mu = 6$

$\sigma = 4$

Find the values that contain 90% of the data around μ

SOLUTION: We are looking for x values. 1) find the z scores: 100% - 90% = 10% to be evenly divided in the tails. So each tail has .05 in it.

What z score corresponds to 0.05?

Use the formula $x = \mu + z\sigma$

Given Probability, Find the Associated Random Variable Value

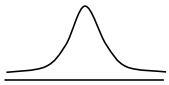
Procedure for Finding the Value of a Normal Random Variable Corresponding to a Specified Proportion, Probability or Percentile

- **Draw a normal curve and shade the area corresponding to the proportion, probability or percentile**
- **Use Table IV or calculator to find the Z-score that corresponds to the shaded area**
- **Obtain the normal value from the fact that $x = \mu + z\sigma$**

Example 1

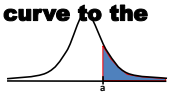
Find the Z-score such that the area under the standard normal curve to the left is 0.1.

ANS: invNorm(0.1) = -1.282 = a



Find the Z-score such that the area under the standard normal curve to the right is 0.35.

ANS: invNorm(1-0.35) = 0.385



p. 279 #54a

- Find the value of z so that the area under the standard normal curve
  a) from 0 to z is (approximately) 0.1965 and z is positive

p. 279 #56a

- **Determine the value of z so that the area under the standard normal curve**
- **a) in the right tail is 0.0250**

p. 279 #62

- **Rockport Corp. make electric shavers. The life of model S7J has a normal distribution with a mean of 70 months and a standard deviation of 8 months. The company is trying to determine the warranty period for this shaver. Any shaver needing repair during the warranty period will be replaced free of charge. What should the warranty period be if the company does not want to replace more than 1% of all shavers?**