Chapter 6: The Normal Distribution

Learning Objectives

Upon successful completion of Chapter 6, you will be able to:

- Identify shapes of distributions as symmetrical or skewed.
- Identify the properties of the normal distribution.
- Find the area under the standard normal distribution, give z-values.
- Find probabilities for a normally distributed variable by transforming it into a standard score.
- Given area or probability for a normal distribution, find the z-value and the x-value.

I. Introduction

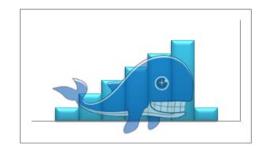
- Many continuous variables have distributions that are bell-shaped and are called approximately normally distributed variables.
- A normal distribution is also known as the bell curve or the Gaussian distribution.

II. Distribution Shapes

A. Negatively Skewed or Left Skewed Distribution

The majority of the data values fall to the right of the mean or cluster at the upper end of the distribution.

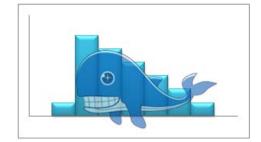
Remember: the tail is to the LEFT



B. Positively Skewed or Right Skewed Distribution

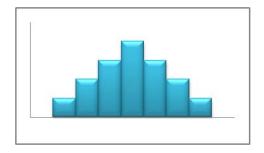
The majority of the data values fall to the left of the mean or cluster at the lower end of the distribution.

Remember: the tail is to the RIGHT



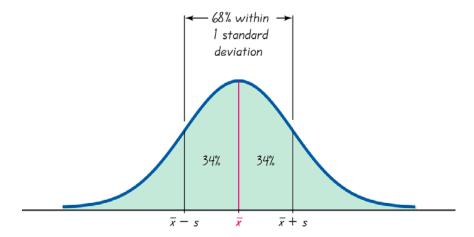
C. Symmetrical Distribution

In a **symmetrical distribution**, the data values are evenly distributed on both sides of the mean.



III. Normal Distributions

The normal distribution is a symmetrical, continuous, bell-shaped distribution of variable.



A. Equation for a Normal Distribution

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

Where:

e ≈ 2.718

 $\pi \approx 3.14$

 μ = population mean

 σ = population standard deviation

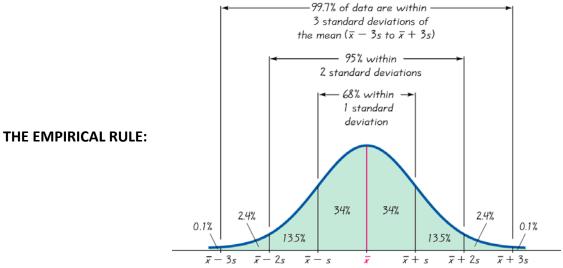
Notice: The shape and position of the normal distribution curve depend on 2 parameters: the mean and the standard deviation.

B. Normal Distribution Properties

- Bell-shaped
- The mean, median, and mode are equal and located at the center of the distribution.
- Uni-modal (only 1 mode).
- Symmetrical about the mean.
- The curve is continuous i.e., there are no gaps or holes. For each value of X, here is a corresponding value of Y.
- The curve never touches the x-axis. Theoretically, no matter how far in either direction the curve extends, it never meets the x-axis, but it gets increasingly closer to the x-axis.
- The total area under the normal distribution curve is equal to 1.00 or 100%.
- .50 or 50% of the area lies to the *left of the mean*.
- .50 or 50% of the area lies to the *right of the mean*
- P(a < z < b) is the area below the normal curve, above the z-axis between z = a and z = b

C. Empirical Rule for All Bell-shaped Distributions

- Approximately 68% of the data values fall within 1 standard deviation of the mean.
- Approximately 95% of the data values fall within 2 standard deviations of the mean.
- Approximately 99.75% of the data values fall within 3 standard deviations of the mean.



D. Types of Normal Distributions

- Standard Normal with mean 0 and standard deviation 1: use given values for z with the tabled probability values.
- Normal Distributions with mean different from 0 and any value for the standard deviation. A different table of probabilities is needed for each paired mean and standard deviation. To avoid this problem, transform all normal distributions to the standard normal using z-scores:
- I. Find the z-score

$$z = \frac{value - mean}{standard\ deviation}$$

$$z = \frac{x - \mu}{\sigma} \qquad \qquad z = \frac{x - \bar{x}}{s}$$

II. Use the z-values to find the appropriate probability from the standard normal table.

Note: the z score is the number of standard deviation that a particular X value is away from the mean; If z is positive, x is above the mean. If z is negative, x is below the mean.

E. Standard Normal Distribution

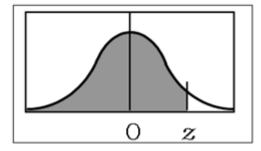
- Mean is 0
- Standard deviation is 1
- Variance is 1

Notation: $N(\mu, \sigma)$ OR N(0, 1) for the standard normal.

• The probability to the left of z is found in Table E, located on the inside cover of the textbook. Be sure the table in your textbook functions this way.

The area under the curve is more important than the frequencies because the area corresponds to the probability!

The area under the normal curve = probability



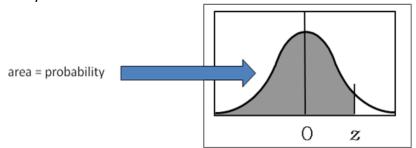
IV.Standard Normal Probability

A. Type 1: Area of Probability to the Left of z

I. Method:

Use Table E giving the area of probability to the left of z:

- 1. Use the z column to find the row for the units and tenths.
- 2. Move to the appropriate hundredths column across the top of the table.
- 3. The Intersection of the row for the units and tenths with the column for hundredths is the probability to the left of z.

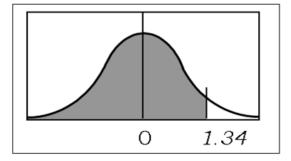


II. <u>Example</u>:

Find the area under the normal distribution curve.

$$P(z < 1.34) =$$

- Sketch the curve.
- Mark the mean.
- Shade the area.
- Use the table to find the area to the left of 1.34



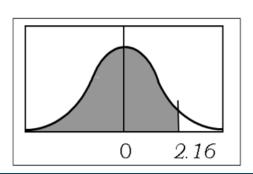
Note: < points left and the area to the left of 1.34 is shaded.

P (z < 1.34) = .9099 is at the intersection of the z-column row 1.3 and the top column labeled .04.

Question 1

Find P (z < 2.16) =

- Sketch the curve.
- Mark the mean.
- Shade the area.
- Use the table to find the area to the left of 2.16



B. Type II: Area to the Right of Any z-Value

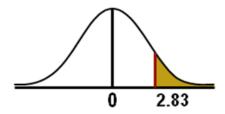
- 1. Look up the z-value in the table to find the area to the left of z.
- 2. Subtract this area from 1.00 to find the area to the right of z.

I. Method:

- 1. Draw the normal curve labeling the mean.
- 2. Shade the area requested.
- 3. Find the probability to the left of the z-value.
- 4. Perform the appropriate arithmetic operation using 1.0 when needed.
 - 1.0 is the area under the entire normal probability curve because the sum of all probability is 1.00.

II. <u>Example</u>:

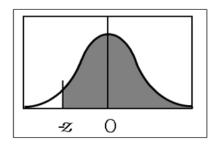
$$P(z > 2.83) =$$



- 1. Find the area to the left of 2.83 because it is given in the table. P(z < 2.83) = .9977
- 2. Subtract the area on the left (.9977) from the total area (1.00) to find the area on the right: P(z > 2.83) = 1.00 .9977 = .0023

Question 2

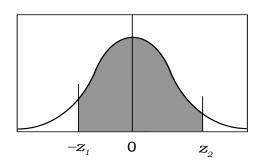
$$P(z > -2.45) =$$



Note: > points to the right. We need to find the area to the right to -2.45.

C. Type III: Area Between Two z-values

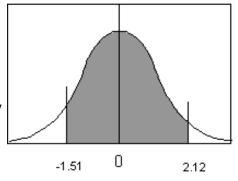
- I. Method
 - 1. Look up both z-values to find the areas to their left.
 - 2. Subtract the smaller area from the larger area to find the area between the two z-values.



II. Example:

Find the area between z = 1.51 and z = 2.12:

- 1. Sketch the curve
- 2. Mark the mean
- 3. Shade the area
- 4. Find the area to the left of each z separately
- 5. Subtract the smaller area from the larger



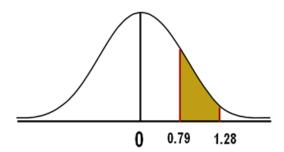
$$P(-1.51 < z < 2.12) = .9830 - .0655$$

= .9175

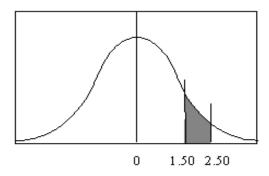
Question 3

Find the area between

z = 0.79 and z = 1.28:

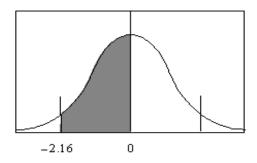


Find the area between z = 1.50 and z = 2.50:



Question 5

Find the area between z = -2.16 and z = 0:



Question 6

P(z > -2.56) =

- a) .4948
- b) .0052
- c) .9947
- d) -.4948

Question 7

P (-1.76 < z < 2.45) =

- a) .9537
- b) .0321
- c) -.0321
- d) None of the above

$$P(1.87 < z < 2.76) =$$

- a) .9664
- b) .0278
- c) -.0278
- d) None of the above

Question 9

$$P(z > 2.61) =$$

- a) .4955
- b) .9955
- c) .0045
- d) -.0045

D. Review: Finding Probability or Area for z-scores

- 1. Sketch the normal curve locating z approximately on the left or right of zero.
- 2. Shade the area requested in the problem.

Type I: Find the probability less than a z-value or use Table E.

Type II: To find the probability greater than z, subtract the probability less than z from 1.00.

Type III: To find the probability between 2 z-values, subtract the smaller area from the larger area.

Question 10

Calculate each probability using the method listed above. The first one has been completed for you.

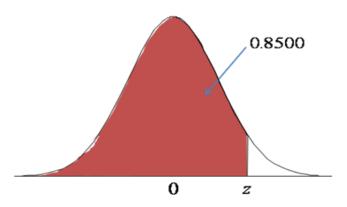
- 1. P(z > 2.56) = .0052
- 2. P(-1.87 < z < 0) =
- 3. P(0 < z < 2.34) =
- 4. P(-1.76 < z < 2.45) =
- 5. P(z > 2.61) =

V. For Standard Normal Probability Given Area/Probability, Find z (Reverse/Backwards of IV)

The normal distribution table may also be used to determine a z-score if we are given the area (working backwards).

A. Example:

What is the z-score associated with the 85th percentile?



Solution, Example: Backwards or Given Area/Probability, Find z

The normal distribution table may also be used to determine a z-score if we are given the area (working backwards).

In Table E, the Normal Probability Table, find the "area" entry that is closest to 0.8500:

Z	0.00	0.01	0.02	0.03		0.04	0.05
1.0				0.8485	0.8500	0.8508	

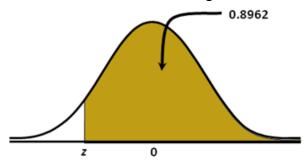
- The area entry closest to 0.8500 is 0.8508.
- The z-score that corresponds to this area is 1.04.
- The 85th percentile in a standard normal distribution is **1.04**.

B. Given Probability, Find Z

Method:

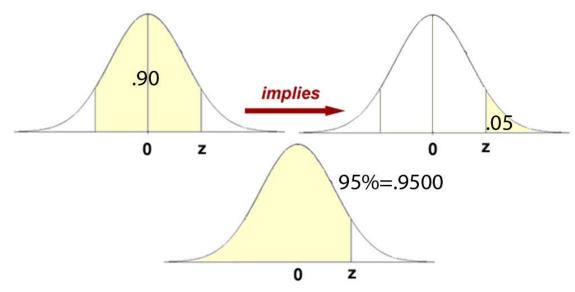
- 1. Sketch the normal curve approximating the location for z on the left or right of 0 depending on the problem.
- 2. Find the probability or area to the left of z.
- 3. Use Table E to locate the area entry *inside the table*.
- 4. To determine the z value, add the z row label for the integer and tenths digits to the column label for the hundredths digits.

Find the z value that corresponds to the given area. If you'd like, review the method discussed above before starting.



Question 12

Find the z value that corresponds to the given area or find the z-values that bound the center 90% of the normal curve.



VI. Probability for Any Normal Curve

Normal Distributions $N(\mu, \sigma)$

- Since each normally distributed variable has a different mean and standard deviation, the shape and location of these curves will vary.
- A different table of probabilities would be needed for each mean, and standard deviation pair.
- To avoid this problem, transform any normal distribution to the standard normal distributions.
- Find the probability for the standard normal distribution and you have the probability for the original normal distribution.

A. Finding the Probability for Any Normally Distributed x value

- Method:
 - 1. Find the z-score for any value. The z-value is the number of standard deviations that a particular X value is away from the mean.

$$z = \frac{value - mean}{standard\ deviation} \quad or \quad z = \frac{x - \mu}{\sigma}$$

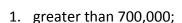
$$z = \frac{x - \bar{x}}{s}$$

- 2. Find the standard normal probability using the appropriate methods.
- II. Example:
 - 3. Heights of 2 year old boys are normally distributed with mean 28 inches and standard deviation 2.4. What is the chance a 2 year old boy is less than 26 inches tall?

$$P(x < 26) = P\left(z < \frac{26 - 28}{2.4}\right) = P(z < -.83) = .20$$

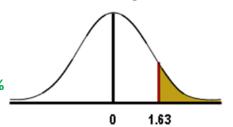
III. Example:

The average daily jail population in the United States is **618,319**. If the distribution is normal and the standard deviation is 50,200, find the probability that on a randomly selected day, the jail population is: $z = \frac{x - \mu}{\sigma}$



$$z = \frac{700,000 - 618,319}{50,200} = 1.63$$

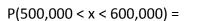
$$P(x > 700,000) = P(z > 1.63) = 1 - .9484 = .0516 \text{ or } 5.16 \%$$

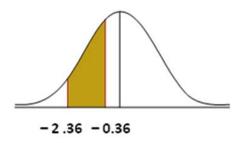


2. between 500,000 and 600,000.

$$z = \frac{600,000 - 618,319}{50,200} = -0.36$$

$$z = \frac{500,000 - 618,319}{50,200} = -2.36$$





The average credit card debt for college seniors is \$3262. If the debt is normally distributed with a standard deviation of \$1100, find these probabilities:

- 1. That an individual senior owes at least \$1000;
- 2. That the senior owes more than \$4000;

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

3. That the senior owes between \$3000 and \$4000.

IV. Notation Reminder

If x is a normal random variable with the mean μ and standard deviation σ , this is often denoted: $X \sim N(\mu, \sigma)$.

Example: Suppose x is a normal random variable with μ = 35 and σ = 6. A convenient notation to identify this random variable is: $x \sim N(35, 6)$.

B. Given Normal Probability, find the x value

We also need to find the x values for $N(\mu, \sigma)$ when probabilities are given.

I. Method:

- a) Given the area or probability, find the probability or area under the normal curve to the left of x.
- b) Use Table E to locate the area to the left of x inside table and determine z. (See page 10)
- c) Find x where:

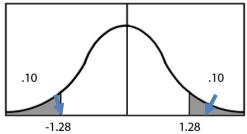
$$z = \frac{x - \mu}{\sigma}$$

$$z\sigma = x - \mu$$

$$x = z \cdot \sigma + \mu$$

II. Example:

If the tallest 10% and the shortest 10% of the population is considered abnormal, what is the range for normal heights of women if heights for women are N (63.6 in, 2.5 in)?



Worlden are W (03.0 III, 2.5 III):

Left answer:

P (z < -1.28) = .10

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

-1.28 = $\frac{x-63.6}{2.5}$

-3.2 = x - 63.6

60.4 = x

Right answer:

P (z > 1.28) = .10

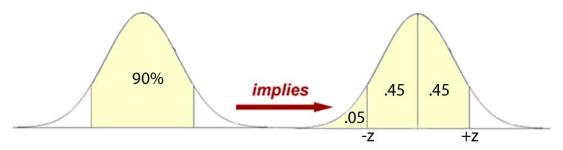
1.28 = $\frac{x-63.6}{2.5}$

3.2 = x - 63.6

Normal height is between 60.4 and 66.8 inches.

III. Example:

The price of a new home in Blair County is normally distributed with a mean price of **\$145,000** and a standard deviation of **\$1,500**. Find the <u>minimum</u> and <u>maximum</u> price to attract the middle 90% of this market.



The middle 90% is split into 2 equal parts by the mean, or the area to the left of -z is .05 and the area to the left of +z is .45+.50 = .9500.

1. Find the area in the Table E closest to 0.9500:

z	0.00	0.01	0.02	0.03	0.04		0.05	0.06
1.6					0.9495	0.9500	0.9505	

- 0.9500 is exactly half-way between 0.9495 and 0.9505. Therefore, z = 1.645 and -z = -1.645 by symmetry.
- z = -1.645 and z = 1.645 bound the middle 90% of a normal distribution.

2. Find x where:
$$x = z \cdot \sigma + \mu$$

$$z = \pm 1.645$$
 $\mu = 145,500$ $\sigma = 1500$

Left Answer: Right Answer:

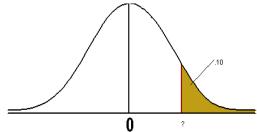
$$x = -1.645 \cdot 1500 + 145,500$$
 $x = +1.645 \cdot 1500 + 145,500$

$$x = 143032.5$$
 $x = 147967.5$

$$x = 143032.5 \text{ or } x = 147967.5$$

143,032.50 and 147,967.50 bound the middle 90% of the market prices for these new homes.

Test scores for the police academy are normally distributed with mean 95 and standard deviation 15. To qualify, a candidate must score in the top 10%. Find the lowest possible score to qualify.



VII. Distribution of Sample Means

- A sampling distribution of sample means is a distribution obtained by using the means for all random samples of the same size taken from the same population.
- Sampling error is the difference between the sample measure and the corresponding population measure due to the fact that the sample is not a perfect representation of the population.

A. Properties of Distribution of Sample Means

- I. The mean of the sample means will be equal to the population mean.
- II. The standard deviation of the sample means will be smaller than the standard deviation of the population, and will be equal to the population standard deviation divided by the square root of the sample size.

B. The Central Limit Theorem (formalizing the properties)

- As the sample size n increases, distribution of the sample means taken with replacement from a population with mean μ and standard deviation σ will approach a normal distribution.
- ullet The mean of the sample means equals the population mean or: $\mu_{ar{X}}=~\mu$
- The standard deviation of the sample means equals: $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ and is called the standard error of the mean.

C. Individual vs. Sample Mean

The central limit theorem is used to answer questions about sample means just like the normal distribution can be used to answer questions about individual values.

• For individuals:
$$Z=rac{x-\mu}{\sigma}$$
For sample means: $Z=rac{ar{x}-\mu}{\sigma/\sqrt{n}}$

Note: Finite Population Correction Factor is not part of this course although it is a topic in the textbook.

Question 15

A.C. Nelson reported that children between 2 and 5 years old watch an average of 14 hours of TV per week. Assuming the variable (hrs of TV watched) is normally distributed with a standard deviation of 3, find the chance that:

- 1. My neighbor John, who is 4, will watch 20 or more hours of TV this week.
- The average number of hours watched by the 25 members of John's play group is 20 hours or more.

Question 16

The average cholesterol content of a certain brand of eggs is 215 milligrams, and the standard deviation is 15 milligrams. Assume the variable is normally distributed.

- 1. If a <u>single egg</u> is selected, find the probability that the cholesterol content will be less than 210 milligrams.
- 2. If a <u>dozen eggs</u> are selected, what is the chance their average cholesterols is 217 or less? Cholesterol for eggs is *N*(215, 15).

VIII. TI-83 Calculator Functions

A. Standard Normal Probability on the TI-83

- Method normalcdf(L,R)
 - 1. Use the Distr Menu accessible with (2nd) VARS
 - 2. Use the normalcdf (left endpoint, right endpoint) = normal cdf(lower value, higher value)
 - Press enter.

(**Note:** the normal pdf is NOT used in this course!)

$$P(a < z < b) = normalcdf(a, b)$$

II. Example:

$$P(2.34 < z < 2.78) = normalcdf(2.34, 2.78) = 0.0069$$

$$P(-1.34 < z < 3.45) =$$

Question 18

$$P(z > 2.34) =$$

Question 19

$$P(z < -1.79) =$$

B. Normal Probability on the TI-83

- I. Method: normal cdf(L,R, μ , σ)
 - 1. Use the Distr Menu accessible with (2nd) VARS
 - 2. Use the normalcdf (left endpoint, right endpoint, mean, standard deviation)
 - 3. Press enter.

normalcdf (L, R, μ , σ)

(Note: the normalpdf is NOT used in this course!)

II. Example:

For a normal random variable with mean 50 and standard deviation 2, find the proportion of scores greater than 45.

Question 20

For z, find
$$P(z < 2.87)$$

Question 21

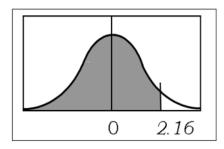
For a normal random variable x with mean 12 and variance 6, find the probability that x is more than 16.

IX. Summary

- Many variables such as heights, weights, and temperatures have normally distributed data.
- The standard normal distribution has a mean of 0 and a standard deviation of 1.
- All normal distributions are bell-shaped with equal mean, median, and mode.
- For individual data values from a normal distribution use: $Z = \frac{\bar{x} \mu}{\sigma}$
- For group averages with n \geq 30, use: $Z = \frac{x-\mu}{(\frac{\sigma}{\sqrt{n}})}$

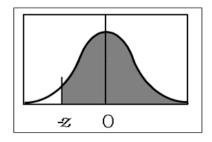
P(z < 2.16) = .9846 is at the intersection of the Z-column row 2.1 and the top column labeled .06

- 1. Sketch the curve.
- 2. Mark the mean.
- 3. Shade the area.
- 4. Use the table to find the area to the left of 2.16



Answer: Question 2

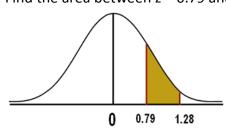
Find the area to the right of -2.45



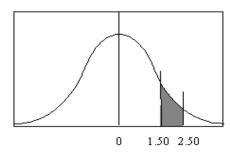
$$P(z > -2.45) = 1 - .0071 = .9929$$

Answer: Question 3

Find the area between z = 0.79 and z = 1.28:



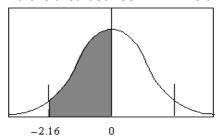
Find the area between z = 1.50 and z = 2.50:



P(1.50 < z < 2.50) = .9938 - .9332 = .0606

Answer: Question 5

Find the area between z = -2.16 and z = 0:



P(-2.16 < z < 0) = .5000 - .0154 = .4846

Answer: Question 6

P(z > -2.56) =

- a.) .4948
- b.) .0052
- c.) .9948
- d.) -.4948

Answer: Question 7

P(-1.76 < z < 2.45) =

- a.) .9537
- b.) .0321
- c.) -.0321
- d.) None of the above.

Answer: Question 8

P(1.87 < z < 2.76) =

- a.) .9664
- b.) .0278
- c.) -.0278
- d.) None of the above.

P(z > 2.61) =

- a.) .4955
- b.) .9955
- c.) .0045
- d.) -.0045

Answer: Question 10

Calculate each probability. The first one has been completed for you.

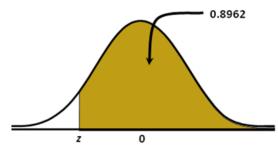
- 1. P(z > 2.56) = .0052
- 2. P(-1.76 < z < 2.45) = .9929 .0392 = .9537
- 3. P(-1.87 < z < 0) = .5 .0307 = .4693
- 4. P(z > 2.61) = 1 .9955 = .0045
- 5. P(0 < z < 2.34) = .9904 .5000 = .4904

Answer: Question 11

Find the z value that corresponds to the given area.

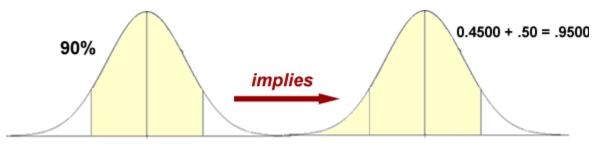
Method:

- 1. Sketch
- 2. The area or probability to the left of z is: 1 .8962 = .1038
- 3. .1038 is located inside Table E at the intersection of row -1.2 and column .06.
- 4. z = -1.2 + -.06 = -1.26



Thus, P(z > -1.26) = 0.8962

Find the z-values that correspond to the given area or find the z-values that will bound the center 90% of the normal curve.



The 90% is split into 2 equal 0.45 parts by the mean.

Find the area in Table E closest to .9500 = .5000 + .4500

z	0.00	0.01	0.02	0.03	0.04		0.05	0.06
1.6					0.9495	0.9500	0.9505	

0.9500 is exactly half-way between 0.9495 and 0.9505. Therefore, z = 1.645.

Using symmetry, z = -1.645 and z = 1.645 bound the middle 90% of a normal distribution.

Answer: Question 13

The average credit card debt for college seniors is **\$3262**. If the debit is normally distributed with a standard deviation of **\$1100**, find these probabilities:

1. That an individual seniors owes at least \$1000;

$$z = \frac{1000 - 3262}{1100} = \ -2.06$$

$$P(x \ge 1000) = P(z > -2.06) = 1 - .0197 = .9803$$

2. That the senior owes more than \$4000;

$$z = \frac{4000 - 3262}{1100} = 0.67$$

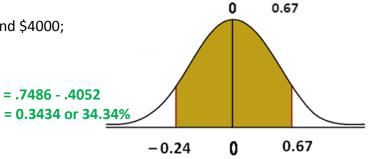
$$P(x > 4000) = P(z > 0.67) = 1 - .7486 = 0.2514 \text{ or } 25.14\%$$

3. That the senior owes between \$3000 and \$4000;

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

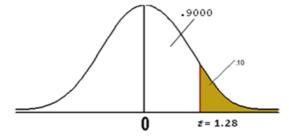
P(3000 < x < 4000) = P(-0.24 < z < 0.67) = .7486 - .4052

$$z = \frac{3000 - 3262}{1100} = -0.24$$



-2.06

Test scores for the police academy are normally distributed with mean 95 and standard deviation 15. To qualify, a candidate must score in the top 10%. Find the lowest possible score to qualify.



$$x = z \cdot \sigma + \mu$$

$$\mu = 95$$

$$\sigma$$
 = 15

$$x = 1.28 \cdot 15 + 95$$

$$x = 114.2$$

114.2 is the lowest score for a candidate to qualify.

Answer: Question 15

A. C. Nelson reported that children between 2 and 5 years old watch an average of 14 hours of TV per week. Assuming the variable (hrs of TV watched) is normally distributed with a standard deviation of 3, find the chance that:

1. My neighbor John, who is 4, will watch 20 or more hours of TV this week.

- for an individual, use N(14, 3)

$$P(x > 20) = P(z > 2) = 1 - .9772 = .0228$$

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

2. The average number of hours watched by the 25 members of John's play group is 20 hours or more.

- for the average of a group, use $N(14, \frac{3}{\sqrt{25}})$

$$P(\overline{X} > 20) = P\left(z > \frac{20 - 14}{\left(\frac{3}{\sqrt{25}}\right)}\right) = P(z > 10) = 0^{+}$$

Answer: Question 16

The average cholesterol content of a certain brand of eggs is 215 milligrams, and the standard deviation is 15 milligrams. Assume the variable is normally distributed.

1. If a <u>single egg</u> is selected, find the probability that the cholesterol content will be less than 210 milligrams.

$$Z = \frac{X - \mu}{\sigma}$$
 $P(x < 210) = P(z > \frac{210 - 215}{15}) = P(z < -.33) = .3707$

2. If a <u>dozen eggs</u> are selected, what is the chance their average cholesterols is 217 or less? Cholesterol for eggs is *N*(215, 15).

$$z = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}} \qquad P(\overline{X} < 217) = P\left(z > \frac{217 - 215}{\left(\frac{15}{\sqrt{12}}\right)}\right) = P(z < .46) = .6772$$

P(-1.34 < z < 3.45) = normalcdf(-1.34, 3.45) = .9096

Answer: Question 18

P(z > 2.34) = normalcdf(2.34, 1000) = .0096

Answer: Question 19

P(z < -1.79) = normalcdf(-1000, -1.79) = .0367

Answer: Question 20 For z, find P (z < 2.87)

normalcdf (-1000, 2.87) = .9979

Answer: Question 21

For a normal random variable x with mean 12 and variance 6, find the probability that x is more than 16.

normalcdf (16, 1000, 12, $\sqrt{6}$) = 0.512