Random Variables

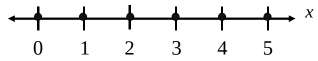
Random Variable

- Represents a numerical value associated with each outcome of a probability distribution.
- Denoted by ***x***, or other variable name
- Examples
  - *x* = number of sales calls a salesperson makes in one day
  - *h* = hours spent on sales calls in one day
  - *d* = outcome of rolling a die

Random Variables

Discrete Random Variable

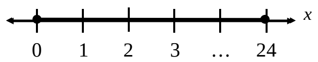
- Has a finite or countable number of possible outcomes that can be listed.
- Example
  - *x* = Number of sales calls a salesperson makes in one day.



Random Variables

Continuous Random Variable

- Has an uncountable number of possible outcomes, represented by an interval on the number line.
- Example
  - *x* = Hours spent on sales calls in one day.



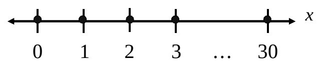
Example: Random Variables

Decide whether the random variable *x* is discrete or continuous.

1. *x* = The number of stocks in the Dow Jones Industrial Average that have share price increases on a given day.

Solution:

Discrete random variable (The number of stocks whose share price increases can be counted.)

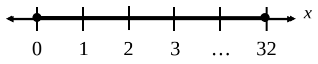


Example: Random Variables

Decide whether the random variable x is discrete or continuous.

2. x = The volume of water in a 32-ounce container.

Solution:
Continuous random variable (The amount of water can be any volume between 0 ounces and 32 ounces)



Discrete Probability Distributions

Discrete probability distribution

- Lists each possible value the random variable can assume, together with its probability.
- Must satisfy the following conditions:

In Words	*In Symbols*
1. The probability of each value of the discrete random variable is between 0 and 1, inclusive.	$0 \le P(x) \le 1$
2. The sum of all the probabilities is 1.	$\Sigma P(x) = 1$

Constructing a Discrete Probability Distribution

Let x be a discrete random variable with possible outcomes $x_1, x_2, \dots, x_n$.

1. Make a frequency distribution for the possible outcomes.
2. Find the sum of the frequencies.
3. Find the probability of each possible outcome by dividing its frequency by the sum of the frequencies.
4. Check that each probability is between 0 and 1 and that the sum is 1.

Exercise 1: Constructing a Discrete Probability Distribution

An industrial psychologist administered a personality inventory test for passive-aggressive traits to 150 employees. Individuals were given a score from 1 to 5, where 1 was extremely passive and 5 extremely aggressive. A score of 3 indicated neither trait. Construct a probability distribution for the random variable x. Then graph the distribution using a histogram.

Score, x	Frequency, f
1	24
2	33
3	42
4	30
5	21

Solution: Constructing a Discrete Probability Distribution

- Divide the frequency of each score by the total number of individuals in the study to find the probability for each value of the random variable.

$$P(1)=\frac{24}{150}=0.16 \quad P(2)=\frac{33}{150}=0.22 \quad P(3)=\frac{42}{150}=0.28$$

$$P(4)=\frac{30}{150}=0.20 \quad P(5)=\frac{21}{150}=0.14$$

- Discrete probability distribution:

x	1	2	3	4	5
P(x)	0.16	0.22	0.28	0.20	0.14

9

Solution: Constructing a Discrete Probability Distribution

x	1	2	3	4	5
P(x)	0.16	0.22	0.28	0.20	0.14

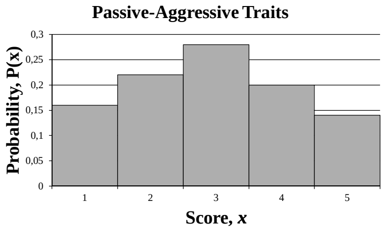
This is a valid discrete probability distribution since

1. Each probability is between 0 and 1, inclusive, $0 \le P(x) \le 1$.
2. The sum of the probabilities equals 1, $\Sigma P(x) = 0.16 + 0.22 + 0.28 + 0.20 + 0.14 = 1$.

10

Solution: Constructing a Discrete Probability Distribution

- Histogram



Because the width of each bar is one, the area of each bar is equal to the probability of a particular outcome.

11

Expected Value

Expected value of a discrete random variable

- Equal to the mean of the random variable.
- $\boldsymbol{E(x) = \mu = \Sigma\, xP(x)}$
- *E(x)* represents the “average” value of all the outcomes
- *E(x)* represents the value that we would expect to get if the trials continued for a very long time

12

Mean

Mean of a discrete probability distribution

- $\boldsymbol{\mu = \Sigma x P(x)}$
- Each value of x is multiplied by its corresponding probability and the products are added
- This formula computes a weighted mean using probability as a weighting factor

13

Exercise 2: Finding the Mean

The probability distribution for the personality inventory test for passive-aggressive traits is given. Find the mean.

Solution:

x	$P(x)$
1	0.16
2	0.22
3	0.28
4	0.20
5	0.14

$\mu = \Sigma x P(x) = 2.94$

14

Variance and Standard Deviation Defining Formulas

Variance of a discrete probability distribution

$\boldsymbol{\sigma^2 = \Sigma(x - \mu)^2 P(x)}$

Standard deviation of a discrete probability distribution

$$\sigma = \sqrt{\sigma^2} = \sqrt{\Sigma(x-\mu)^2 P(x)}$$

15

Exercise 3: Finding the Variance and Standard Deviation – Defining Formula

The probability distribution for the personality inventory test for passive-aggressive traits is given. Find the variance and standard deviation. ($\mu = 2.94$)

x	$P(x)$
1	0.16
2	0.22
3	0.28
4	0.20
5	0.14

= (. −) (.)

16

Solution: Finding the Variance and Standard Deviation – Defining Formula

- Recall μ = 2.94

Score, x	f	$P(x)$	$(x - \mu)^2$	$(x - \mu)^2 * P(x)$
1	24	0.16	3.76	0.602
2	33	0.22	0.88	0.194
3	42	0.28	0.00	0.001
4	30	0.20	1.12	0.225
5	21	0.14	4.24	0.594
TOTAL	150	1.00	10.02	1.616

Standard Deviation: $\sigma = \sqrt{\sigma^2} = \sqrt{1.616} \approx 1.3$

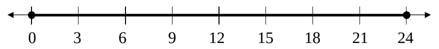
17

Introduction to Normal Distributions

Properties of a Normal Distribution

Continuous random variable

- Has an infinite number of possible values that can be represented by an interval on the number line.



The time spent studying can be any number between 0 and 24.

Continuous probability distribution

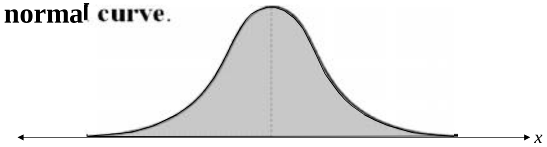
- The probability distribution of a continuous random variable.

18

Properties of Normal Distributions

Normal distribution

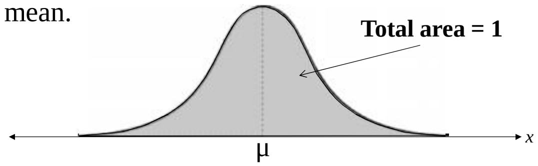
- A continuous probability distribution for a random variable, *x*.
- The most important continuous probability distribution in statistics.
- The graph of a normal distribution is called the **normal curve**.



19

Properties of Normal Distributions

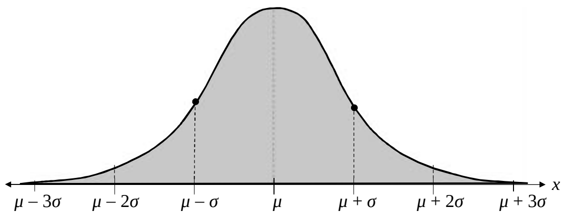
1. The mean, median, and mode are equal.
2. The normal curve is bell-shaped and symmetric about the mean.
3. The total area under the curve is equal to one.
4. The normal curve approaches, but never touches the *x*-axis as it extends farther and farther away from the mean.



20

Properties of Normal Distributions

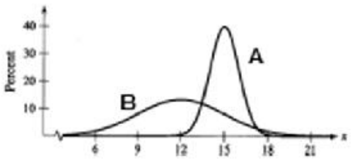
5. Between μ – σ and μ + σ (in the center of the curve), the graph curves downward. The graph curves upward to the left of μ – σ and to the right of μ + σ.



21

Example: Understanding Mean and Standard Deviation

1. Which curve has the greater mean?



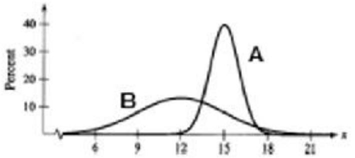
Solution:

Curve *A* has the greater mean (The line of symmetry of curve *A* occurs at *x* = 15. The line of symmetry of curve *B* occurs at *x* = 12.)

22

Example: Understanding Mean and Standard Deviation

2. Which curve has the greater standard deviation?



Solution:

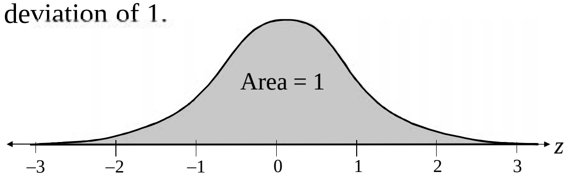
Curve *B* has the greater standard deviation (Curve *B* is more spread out than curve *A*.)

23

The Standard Normal Distribution

Standard normal distribution

- A normal distribution with a mean of 0 and a standard deviation of 1.



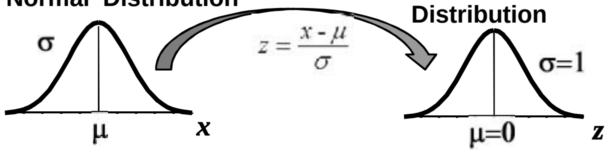
- Any *x*-value can be transformed into a *z*-score by using the formula

$$z = \frac{\text{Value - Mean}}{\text{Standard deviation}} = \frac{x - \mu}{\sigma}$$

24

The Standard Normal Distribution

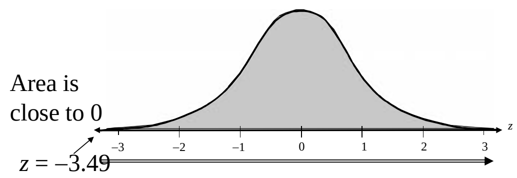
- If each data value of a normally distributed random variable *x* is transformed into a *z*-score, the result will be the standard normal distribution.



- Use the Standard Normal Table to find the cumulative area under the standard normal curve.

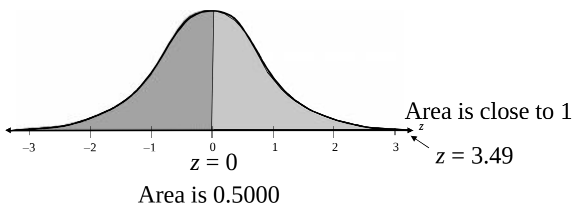
Properties of the Standard Normal Distribution

1. The cumulative area is close to 0 for *z*-scores close to *z* = –3.49.
2. The cumulative area increases as the *z*-scores increase.



Properties of the Standard Normal Distribution

3. The cumulative area for *z* = 0 is 0.5000.
4. The cumulative area is close to 1 for *z*-scores close to *z* = 3.49.



Example: Using The Standard Normal Table

Find the cumulative area that corresponds to a *z*-score of 1.15.

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279

Solution:

Find 1.1 in the left hand column.

Move across the row to the column under 0.05

The area to the left of *z* = 1.15 is 0.8749.

Example: Using The Standard Normal Table

Find the cumulative area that corresponds to a *z*-score of -0.24.

z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

Solution:

Find -0.2 in the left hand column.

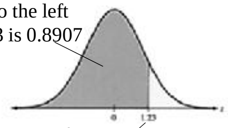
Move across the row to the column under 0.04

The area to the left of *z* = -0.24 is 0.4052.

29

Finding Areas Under the Standard Normal Curve

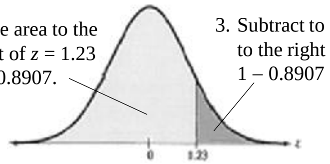
1. Sketch the standard normal curve and shade the appropriate area under the curve.
2. Find the area by following the directions for each case shown.
   a. To find the area to the *left* of *z*, find the area that corresponds to *z* in the Standard Normal Table.



30

Finding Areas Under the Standard Normal Curve

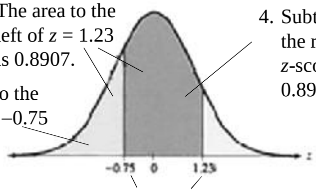
b. To find the area to the *right* of *z*, use the Standard Normal Table to find the area that corresponds to *z*. Then subtract the area from 1.



31

Finding Areas Under the Standard Normal Curve

c. To find the area *between* two *z*-scores, find the area corresponding to each *z*-score in the Standard Normal Table. Then subtract the smaller area from the larger area.

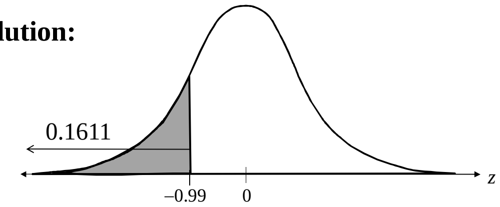


32

Example: Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve to the left of z = -0.99.

Solution:

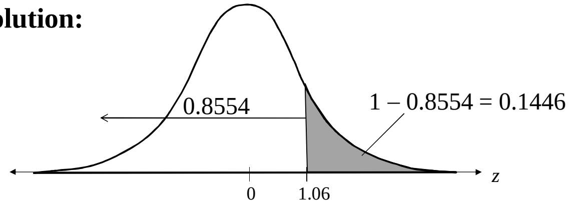


From the Standard Normal Table, the area is equal to 0.1611.

Example: Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve to the right of z = 1.06.

Solution:

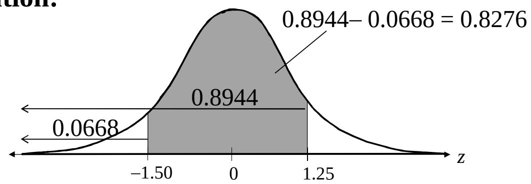


From the Standard Normal Table, the area is equal to 0.1446.

Example: Finding Area Under the Standard Normal Curve

Find the area under the standard normal curve between z = –1.5 and z = 1.25.

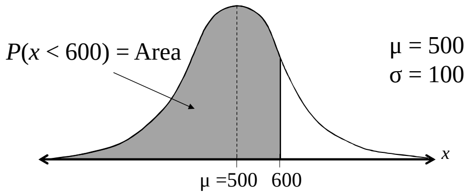
Solution:



From the Standard Normal Table, the area is equal to 0.8276.

Probability and Normal Distributions

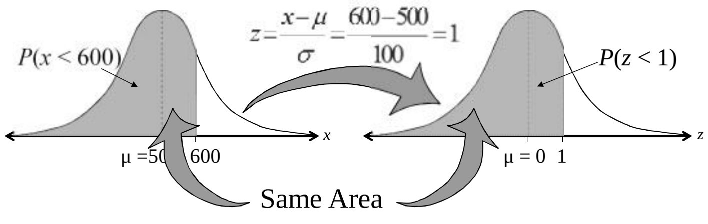
- If a random variable x is normally distributed, you can find the probability that x will fall in a given interval by calculating the area under the normal curve for that interval.



Probability and Normal Distributions

Normal Distribution — $\mu = 500 \quad \sigma = 100$

Standard Normal Distribution — $\mu = 0 \quad \sigma = 1$



$P(x < 500) = P(z < 1)$

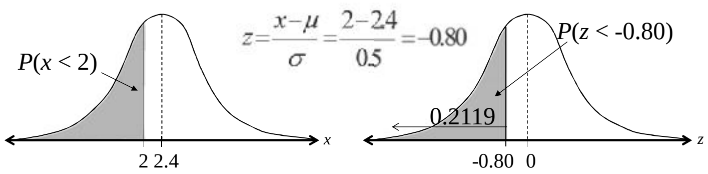
Example: Finding Probabilities for Normal Distributions

A survey indicates that people use their computers an average of 2.4 years before upgrading to a new machine. The standard deviation is 0.5 year. A computer owner is selected at random. Find the probability that he or she will use it for fewer than 2 years before upgrading. Assume that the variable *x* is normally distributed.

Solution: Finding Probabilities for Normal Distributions

Normal Distribution — $\mu = 2.4 \quad \sigma = 0.5$

Standard Normal Distribution — $\mu = 0 \quad \sigma = 1$



$P(x < 2) = P(z < -0.80) = \mathbf{0.2119}$

Example: Finding Probabilities for Normal Distributions

A survey indicates that for each trip to the supermarket, a shopper spends an average of 45 minutes with a standard deviation of 12 minutes in the store. The length of time spent in the store is normally distributed and is represented by the variable *x*. A shopper enters the store. Find the probability that the shopper will be in the store for between 24 and 54 minutes.

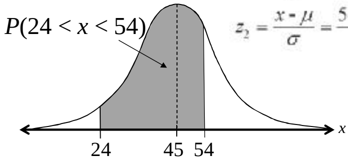
Solution: Finding Probabilities for Normal Distributions

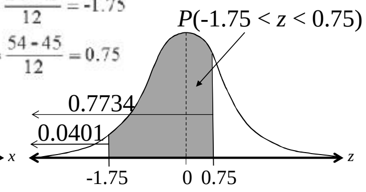
Normal Distribution
$\mu = 45 \quad \sigma = 12$

Standard Normal Distribution
$\mu = 0 \quad \sigma = 1$

$$z_1 = \frac{x - \mu}{\sigma} = \frac{24 - 45}{12} = -1.75$$

$$z_2 = \frac{x - \mu}{\sigma} = \frac{54 - 45}{12} = 0.75$$





$P(24 < x < 54) = P(-1.75 < z < 0.75)$
$= 0.7734 - 0.0401 = \mathbf{0.7333}$

Example: Finding Probabilities for Normal Distributions

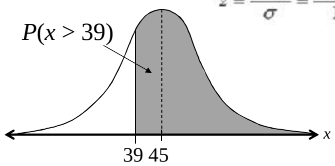
Find the probability that the shopper will be in the store more than 39 minutes. (Recall $\mu = 45$ minutes and $\sigma = 12$ minutes)

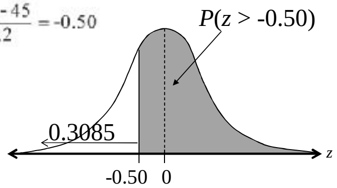
Solution: Finding Probabilities for Normal Distributions

Normal Distribution
$\mu = 45 \quad \sigma = 12$

Standard Normal Distribution
$\mu = 0 \quad \sigma = 1$

$$z = \frac{x - \mu}{\sigma} = \frac{39 - 45}{12} = -0.50$$





$P(x > 39) = P(z > -0.50) = 1 - 0.3085 = \mathbf{0.6915}$

Example: Finding Probabilities for Normal Distributions

If 200 shoppers enter the store, how many shoppers would you expect to be in the store more than 39 minutes?

Solution:
Recall $P(x > 39) = 0.6915$

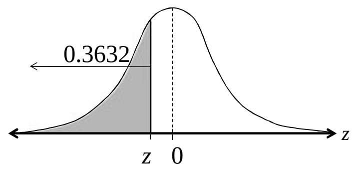
200(0.6915) =138.3 (or about 138) shoppers

Finding values Given a Probability

Example: Finding a *z*-Score Given an Area

Find the *z*-score that corresponds to a cumulative area of 0.3632.

Solution:



46

Solution: Finding a *z*-Score Given an Area

- Locate 0.3632 in the body of the Standard Normal Table.

z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

The *z*-score is -0.35.

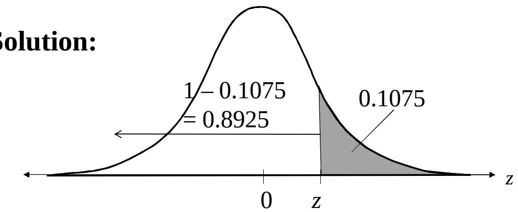
- The values at the beginning of the corresponding row and at the top of the column give the *z*-score.

47

Example: Finding a *z*-Score Given an Area

Find the *z*-score that has 10.75% of the distribution's area to its right.

Solution:



Because the area to the right is 0.1075, the cumulative area is 0.8925.

48

Solution: Finding a *z*-Score Given an Area

- Locate 0.8925 in the body of the Standard Normal Table.

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279

The *z*-score is 1.24.

- The values at the beginning of the corresponding row and at the top of the column give the *z*-score.

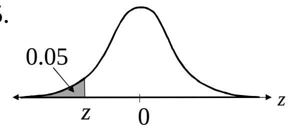
49

Example: Finding a z-Score Given a Percentile

Find the *z*-score that corresponds to P_5.

Solution:

The *z*-score that corresponds to P_5 is the same *z*-score that corresponds to an area of 0.05.



The areas closest to 0.05 in the table are 0.0495 (z = -1.65) and 0.0505 (z = -1.64). Because 0.05 is halfway between the two areas in the table, use the *z*-score that is halfway between -1.64 and -1.65. **The *z*-score is -1.645**.

50

Transforming a *z*-Score to an *x*-Score

To transform a standard *z*-score to a data value *x* in a given population, use the formula

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

51

Example: Finding an x-Value

The speeds of vehicles along a stretch of highway are normally distributed, with a mean of 67 miles per hour and a standard deviation of 4 miles per hour. Find the speeds *x* corresponding to *z*-sores of 1.96, -2.33, and 0.

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

- $z = 1.96$: $x = 67 + 1.96(4) = 74.84$ miles per hour
- $z = -2.33$: $x = 67 + (-2.33)(4) = 57.68$ miles per hour
- $z = 0$: $x = 67 + 0(4) = 67$ miles per hour

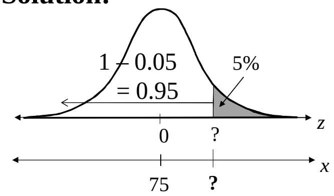
Notice 74.84 mph is above the mean, 57.68 mph is below the mean, and 67 mph is equal to the mean.

52

Example: Finding a Specific Data Value

Scores for a civil service exam are normally distributed, with a mean of 75 and a standard deviation of 6.5. To be eligible for civil service employment, you must score in the top 5%. What is the lowest score you can earn and still be eligible for employment?

Solution:



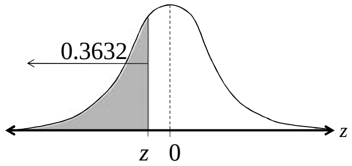
An exam score in the top 5% is any score above the 95th percentile. Find the *z*-score that corresponds to a cumulative area of 0.95.

53

Example: Finding a *z*-Score Given an Area

Find the *z*-score that corresponds to a cumulative area of 0.3632.

Solution:



54

Solution: Finding a *z*-Score Given an Area

- Locate 0.3632 in the body of the Standard Normal Table.

z	.09	.08	.07	.06	.05	.04	.03
−3.4	.0002	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0003	.0004	.0004	.0004	.0004	.0004	.0004
−3.2	.0005	.0005	.0005	.0006	.0006	.0006	.0006
−0.5	.2776	.2810	.2843	.2877	.2912	.2946	.2981
−0.4	.3121	.3156	.3192	.3228	.3264	.3300	.3336
−0.3	.3483	.3520	.3557	.3594	.3632	.3669	.3707
−0.2	.3859	.3897	.3936	.3974	.4013	.4052	.4090
−0.1	.4247	.4286	.4325	.4364	.4404	.4443	.4483
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	.4880

The *z*-score is -0.35.

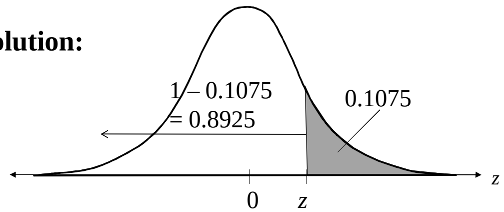
- The values at the beginning of the corresponding row and at the top of the column give the *z*-score.

55

Example: Finding a *z*-Score Given an Area

Find the *z*-score that has 10.75% of the distribution's area to its right.

Solution:



Because the area to the right is 0.1075, the cumulative area is 0.8925.

56

Solution: Finding a *z*-Score Given an Area

- Locate 0.8925 in the body of the Standard Normal Table.

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279

The *z*-score is 1.24.

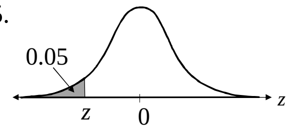
- The values at the beginning of the corresponding row and at the top of the column give the *z*-score.

Example: Finding a z-Score Given a Percentile

Find the *z*-score that corresponds to P_5.

Solution:

The *z*-score that corresponds to P_5 is the same *z*-score that corresponds to an area of 0.05.



The areas closest to 0.05 in the table are 0.0495 (z = -1.65) and 0.0505 (z = -1.64). Because 0.05 is halfway between the two areas in the table, use the *z*-score that is halfway between -1.64 and -1.65. **The *z*-score is -1.645**.

Transforming a *z*-Score to an *x*-Score

To transform a standard *z*-score to a data value *x* in a given population, use the formula

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

Example: Finding an x-Value

The speeds of vehicles along a stretch of highway are normally distributed, with a mean of 67 miles per hour and a standard deviation of 4 miles per hour. Find the speeds *x* corresponding to *z*-sores of 1.96, -2.33, and 0.

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

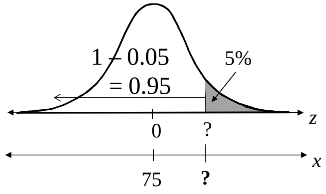
- $z = 1.96$: $x = 67 + 1.96(4) = 74.84$ miles per hour
- $z = -2.33$: $x = 67 + (-2.33)(4) = 57.68$ miles per hour
- $z = 0$: $x = 67 + 0(4) = 67$ miles per hour

Notice 74.84 mph is above the mean, 57.68 mph is below the mean, and 67 mph is equal to the mean.

Example: Finding a Specific Data Value

Scores for a civil service exam are normally distributed, with a mean of 75 and a standard deviation of 6.5. To be eligible for civil service employment, you must score in the top 5%. What is the lowest score you can earn and still be eligible for employment?

Solution:



An exam score in the top 5% is any score above the 95th percentile. Find the *z*-score that corresponds to a cumulative area of 0.95.

61

Example: Sampling Distribution of Sample Means

c. List all the possible samples of size $n = 2$ and calculate the mean of each sample.

Solution:

Sample	$\bar{x}$	Sample	$\bar{x}$
1, 1	1	5, 1	3
1, 3	2	5, 3	4
1, 5	3	5, 5	5
1, 7	4	5, 7	6
3, 1	2	7, 1	4
3, 3	3	7, 3	5
3, 5	4	7, 5	6
3, 7	5	7, 7	7

These means form the sampling distribution of sample means.

62

Example: Sampling Distribution of Sample Means

d. Construct the probability distribution of the sample means.

Solution:

$\bar{x}$	f	Probability
1	1	0.0625
2	2	0.1250
3	3	0.1875
4	4	0.2500
5	3	0.1875
6	2	0.1250
7	1	0.0625

63

Example: Sampling Distribution of Sample Means

e. Find the mean, variance, and standard deviation of the sampling distribution of the sample means.

Solution:

The mean, variance, and standard deviation of the 16 sample means are:

$$\mu_{\bar{x}} = 4 \qquad \sigma_{\bar{x}}^2 = \frac{5}{2} = 2.5 \qquad \sigma_{\bar{x}} = \sqrt{2.5} \approx 1.581$$

These results satisfy the properties of sampling distributions of sample means.

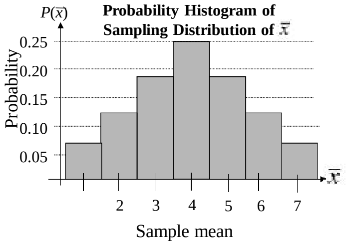
$$\mu_{\bar{x}} = \mu = 4 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{5}}{\sqrt{2}} \approx \frac{2.236}{\sqrt{2}} \approx 1.581$$

64

Example: Sampling Distribution of Sample Means

f. Graph the probability histogram for the sampling distribution of the sample means.

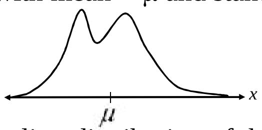
Solution:



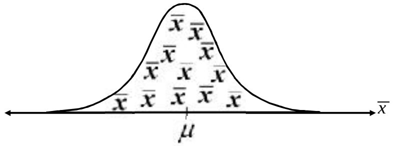
The shape of the graph is symmetric and bell shaped. It approximates a normal distribution.

The Central Limit Theorem

1. If samples of size $n \geq 30$, are drawn from any population with mean = μ and standard deviation = σ,

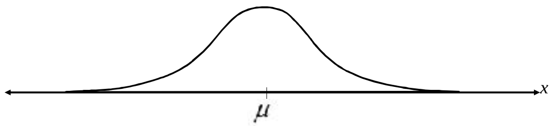


then the sampling distribution of the sample means approximates a normal distribution. The greater the sample size, the better the approximation.

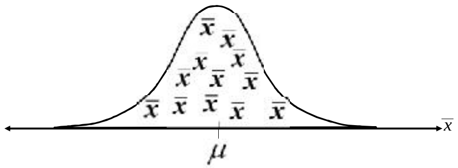


The Central Limit Theorem

2. If the population itself is normally distributed,



the sampling distribution of the sample means is normally distribution for ***any*** sample size *n*.



The Central Limit Theorem

- In either case, the sampling distribution of sample means has a mean equal to the population mean.

$$\mu_{\bar{x}} = \mu$$

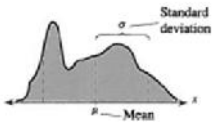
- The sampling distribution of sample means has a variance equal to $1/n$ times the variance of the population and a standard deviation equal to the population standard deviation divided by the square root of *n*.

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$ Variance

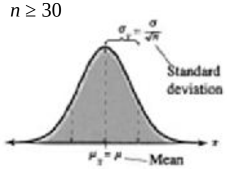
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$ Standard deviation (**standard error of the mean)**

The Central Limit Theorem

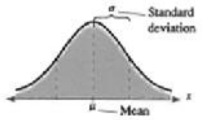
1. Any Population Distribution



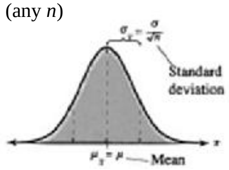
Distribution of Sample Means, $n \geq 30$



2. Normal Population Distribution

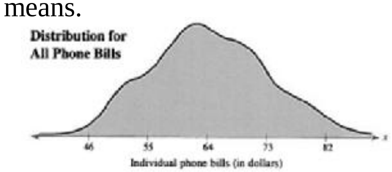


Distribution of Sample Means, (any n)



Example: Interpreting the Central Limit Theorem

Phone bills for residents of a city have a mean of \$64 and a standard deviation of \$9. Random samples of 36 phone bills are drawn from this population and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.



Solution: Interpreting the Central Limit Theorem

- The mean of the sampling distribution is equal to the population mean

$$\mu_{\bar{x}} = \mu = 64$$

- The standard error of the mean is equal to the population standard deviation divided by the square root of *n*.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{36}} = 1.5$$

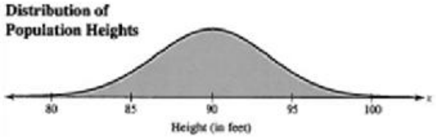
Solution: Interpreting the Central Limit Theorem

- Since the sample size is greater than 30, the sampling distribution can be approximated by a normal distribution with

$$\mu_{\bar{x}} = 64 \qquad \sigma_{\bar{x}} = 1.5$$

Example: Interpreting the Central Limit Theorem

The heights of fully grown white oak trees are normally distributed, with a mean of 90 feet and standard deviation of 3.5 feet. Random samples of size 4 are drawn from this population, and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling distribution of sample means.



73

Example: Probabilities for x and $\bar{x}$

2. You randomly select 25 credit card holders. What is the probability that their mean credit card balance is less than $2500?

Solution:
You are asked to find the probability associated with a sample mean $\bar{x}$.

$$\mu_{\bar{x}} = \mu = 2870 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{900}{\sqrt{25}} = 180$$

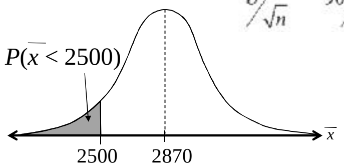
74

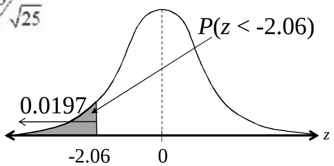
Solution: Probabilities for x and $\bar{x}$

Normal Distribution
$\mu = 2870$ $\sigma = 180$

Standard Normal Distribution
$\mu = 0$ $\sigma = 1$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{2500 - 2870}{900 / \sqrt{25}} \approx -2.06$$





$P(\bar{x} < 2500) = P(z < -2.06) =$ **0.0197**

75

Solution: Probabilities for x and $\bar{x}$

- There is a 34% chance that an individual will have a balance less than $2500.
- There is only a 2% chance that the mean of a sample of 25 will have a balance less than $2500 (unusual event).
- It is possible that the sample is unusual or it is possible that the auditor's claim that the mean is $2870 is incorrect.

76