Lesson 39 – Normal Distributions

IB Math SL - Santowski

Introduction to the "Normal Distribution"

- Run through the the "dice rolling" simulation (https://www.geogebra.org/m/UsoH4eNl)
- KEY point to make ➔ if we run an "experiment" enough times (i.e collect sufficient data), then our histograms (or our distribution "curves") start taking on a consistent "shape" ➔ this shape will be referred to as the "**normal distribution**"

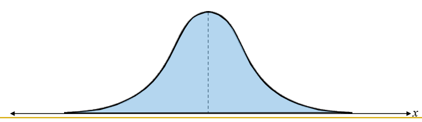
Introduction to the "Normal Distribution"

- Since our distribution came from data we collected, we can analyze the data for key statistical features (parameters) ➔ mean and standard deviation
- We will discuss mean and standard deviation as we analyze our normal distributions

Properties of Normal Distributions

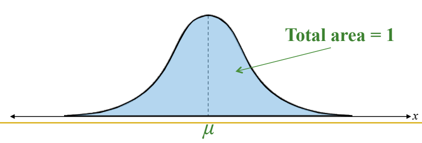
Normal distribution

- 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**.



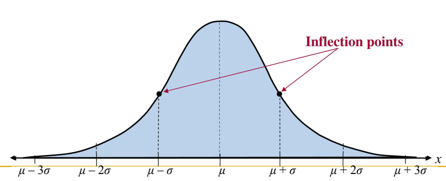
Properties of Normal Distributions

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



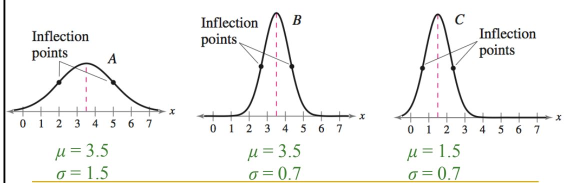
Properties of Normal Distributions

5. Between $\mu - \sigma$ and $\mu + \sigma$ (in the center of the curve), the graph curves downward. The graph curves upward to the left of $\mu - \sigma$ and to the right of $\mu + \sigma$. The points at which the curve changes from curving upward to curving downward are called the **inflection points**.



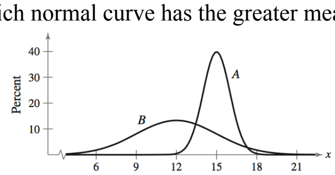
Means and Standard Deviations

- A normal distribution can have any mean and any positive standard deviation.
- The mean gives the location of the line of symmetry.
- The standard deviation describes the spread of the data.



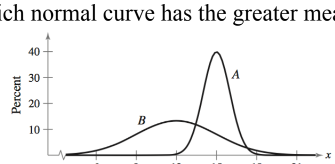
Example: Understanding Mean and Standard Deviation

1. Which normal curve has the greater mean?



Example: Understanding Mean and Standard Deviation

1. Which normal curve has the greater mean?

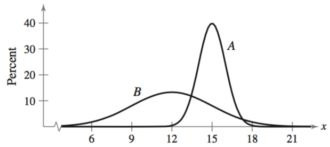


Solution:

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 $x = 12$.)

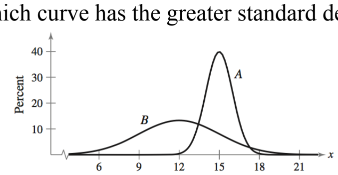
Example: Understanding Mean and Standard Deviation

2. Which curve has the greater standard deviation?



Example: Understanding Mean and Standard Deviation

2. Which curve has the greater standard deviation?

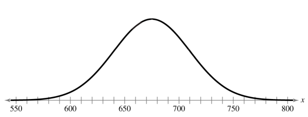


Solution:

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

Example: Interpreting Graphs

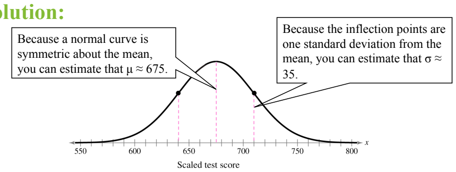
The scaled test scores for the New York State Grade 8 Mathematics Test are normally distributed. The normal curve shown below represents this distribution. What is the mean test score? Estimate the standard deviation.



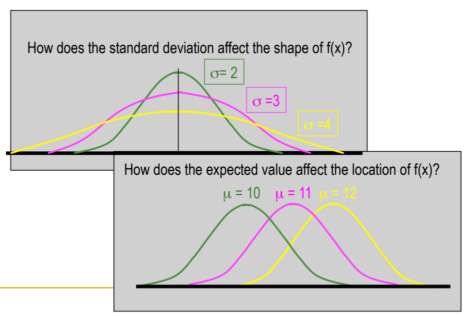
Example: Interpreting Graphs

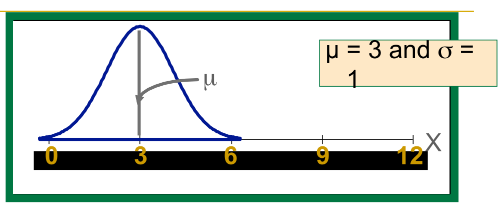
The scaled test scores for the New York State Grade 8 Mathematics Test are normally distributed. The normal curve shown below represents this distribution. What is the mean test score? Estimate the standard deviation.

Solution:



The effects of μ and σ

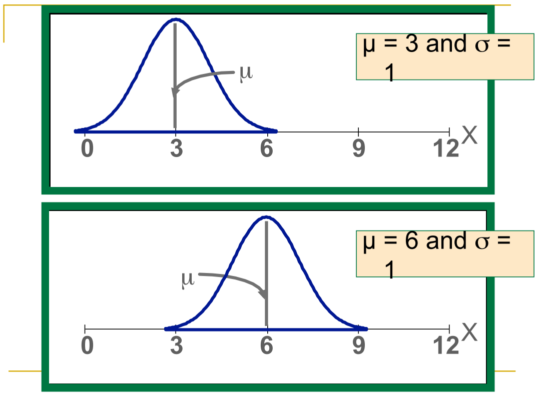


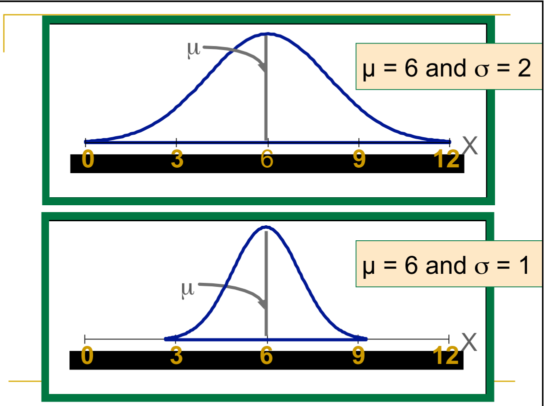


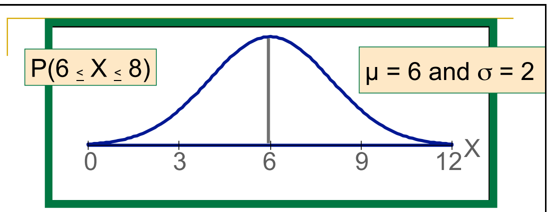
A family of bell-shaped curves that differ only in their means and standard deviations.

μ = the mean of the distribution

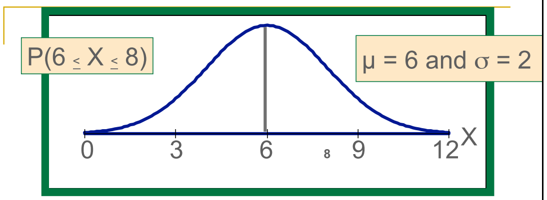
σ = the standard deviation





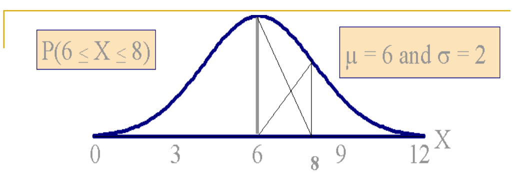


Probability = area under the density curve
P(a ≤ X ≤ b) = area under the density curve between a and b.



Probability = area under the density curve

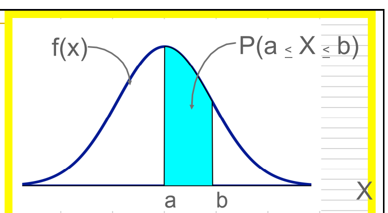
P(**6** ≤ X ≤ **8**) = area under the density curve between **6** and **8**.



Probability = area under the density curve

P(**6** ≤ X ≤ **8**) = area under the density curve between **6** and **8**.

Probabilities: area under graph of f(x)



P(a ≤ X ≤ b) = area under the density curve between a and b.

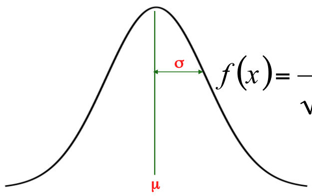
$$P(a \le X \le b) = \int_a^b f(x)dx$$

P(X=a) = 0

P(a ≤ x ≤ b) = P(a < x < b)

The Normal Distribution

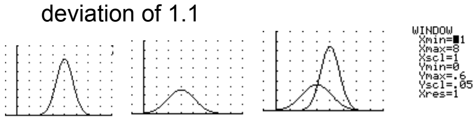
- The Normal distribution has the shape of a "bell curve" with parameters μ and σ^2 that determine the center and spread:
- Additionally, the mean, median and mode are all equal
- The normal curve is symmetrical



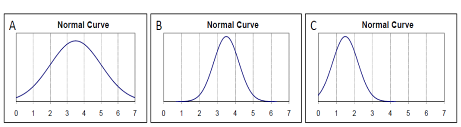
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The Normal Distribution - Example

- Here are scores that IB SL1 students scored on Year 1 June exams where the mean was 4.0 and the standard deviation was 0.7
- Last year, the mean was 3.0 with a standard deviation of 1.1



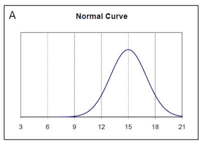
The Normal Distribution

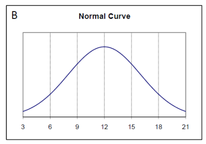


μ=3.5, σ=1.5 μ=3.5, σ=0.7 μ=1.5, σ=0.7

- Curve A and Curve B have the same mean,
- Curve B and Curve C have the same standard deviation,
- Each Curve has a total area of 1.

The Normal Distribution

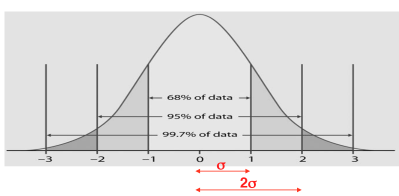




- Which normal curve has a greater mean?
- Which normal curve has a greater standard deviation?

Property of Normal Distributions

- Normal distribution follows the ***68-95-99.7 rule***:
  - 68% of observations are between $\mu - \sigma$ and $\mu + \sigma$
  - 95% of observations are between $\mu - 2\sigma$ and $\mu + 2\sigma$
  - 99.7% of observations are between $\mu - 3\sigma$ and $\mu + 3\sigma$

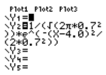
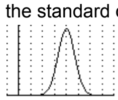
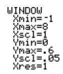


The Normal Distribution - Example

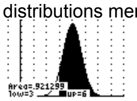
- the IB SL1 students scored on Year 1 June exams where the mean was 4.0 and the standard deviation was 0.7
- (a) what percentage of students scored between 3.3 and 4.7?
- (b) What percentage of students scored between 2.6 and 4.7?
- (c) What percentage of students scored between 1.9 and 5.4?
- (d) If a passing grade was set at 2.6, what percentage of students passed the exam?
- (e) If Honors designations were given to students who scored over 5.4, what percentage of students were given an honors designation?

Working with the Graphing Calculator

- Using a graph
- IB SL1 students scores on Year 1 June exams had a mean of 4.0 and the standard deviation was 0.7

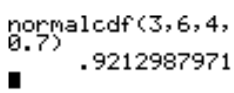




- What percentage of students scored (3 $\leq$ x $\leq$ 6) = 92.1%
- Use shadenorm command on the distributions menu:



Working with the Graphing Calculator

- Using the distribution menus:
- IB SL1 students scores on Year 1 June exams had a mean of 4.0 and the standard deviation was 0.7
- What percentage of students scored (3 < x < 6) ??



Examples

- **Example 1**
- An average light bulb manufactured by the Acme Corporation lasts 300 days with a standard deviation of 50 days. Assuming that bulb life is normally distributed, what is the probability that an Acme light bulb will last at most 365 days?
- **Example 2**
- Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110?

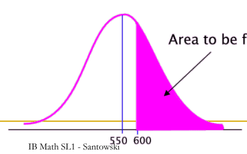
Examples

- **Example 1**
- An average light bulb manufactured by the Acme Corporation lasts 300 days with a standard deviation of 50 days. Assuming that bulb life is normally distributed, what is the probability that an Acme light bulb will last at most 365 days? (90%)

- **Example 2**
- Suppose scores on an IQ test are normally distributed. If the test has a mean of 100 and a standard deviation of 10, what is the probability that a person who takes the test will score between 90 and 110? (68%)

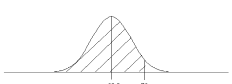
Examples

- Suppose you must establish regulations concerning the maximum number of people who can occupy a lift.
- You know that the total weight of 8 people chosen at random follows a normal distribution with a mean of 550kg and a standard deviation of 150kg.
- What's the probability that the total weight of 8 people exceeds 600kg?

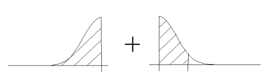


Example

- Suppose that height is normally distributed. The mean height of American women 18-24 is 65.5" with a standard deviation of 2.5". What is the probability that a randomly selected woman is less than 70" tall?



$P(x \leq 70) = ?$

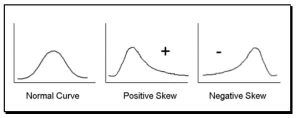


$P(x \leq 65.5) + P(65.5 \leq x \leq 70)$

$= 0.5 + P(65.5 \leq x \leq 70)$

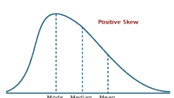
Other variations on distributions

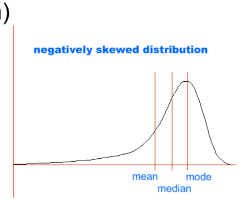
- Not all distributions have the same "bell shaped" curve of the normal distribution
- Sometimes, the mean, median, mode are NOT all identical
- Sometimes the distribution is asymmetrical



Other variations on distributions

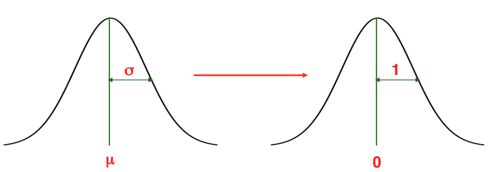
- Distributions can be positively skewed (mean>median>mode) or negatively skewed (mode>median>mean)





Standardization

- If we only have a standard normal table, then we need to ***transform*** our non-standard normal distribution into a standard one
  - This process is called **standardization**



Standardization Formula

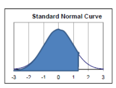
- We convert a non-standard normal distribution into a standard normal distribution using a *linear transformation*
- If X has a N(μ,σ^2) distribution, then we can convert to Z which follows a N(0,1) distribution

$$Z = (X-\mu)/\sigma$$

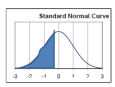
- First, subtract the mean μ from X
- Then, divide by the standard deviation σ of X

Working with Z-scores

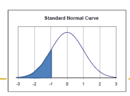
- Find the cumulative area that corresponds
- to a z-score of 1.15



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

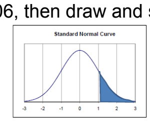


- Find the area under the standard normal curve to the left of z = -0.99, then draw and shade the area under the curve.

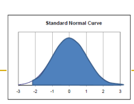


Working with Z-scores

- Find the area under the standard normal curve to the right of z = 1.06, then draw and shade the area under the curve.

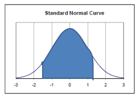


- Find the area under the standard normal curve to the right of z = -2.16, then draw and shade the area under the curve.

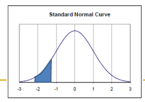


Working with Z-scores

- Find the area under the standard normal curve between *z* = -1.5 and *z* = 1.25, then draw and shade the area under the curve.



- Find the area under the standard normal curve between *z* = -2.16 and *z* = -1.35, then draw and shade the area under the curve.



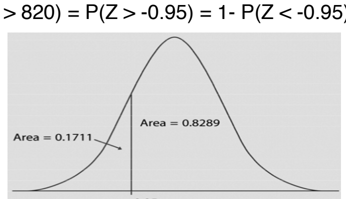
Another Example

- NCAA Division 1 SAT Requirements: athletes are required to score at least 820 on combined math and verbal SAT
- In 2000, SAT scores were normally distributed with mean μ of 1019 and SD σ of 209
- What percentage of students have scores greater than 820 ?

$$Z = (X-\mu)/\sigma = (820-1019)/209 = -199/209 = -.95$$

Another Example

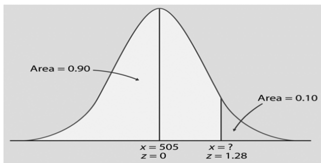
- P(X > 820) = P(Z > -0.95) = 1- P(Z < -0.95)



- P(Z < -0.95) = 0.17 so P(X > 820) = 0.83
- 83% of students meet NCAA requirements

SAT Verbal Scores

- Now, just look at X = Verbal SAT score, which is normally distributed with mean μ of 505 and SD σ of 110
- What Verbal SAT score will place a student in the top 10% of the population?



SAT Verbal Scores

- From the table, P(Z >**1.28**) = 0.10
- Need to reverse standardize to get X:

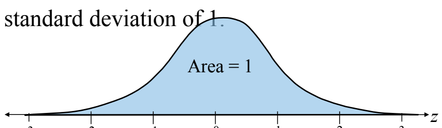
$$X = \sigma \cdot Z + \mu = 110 \cdot 1.28 + 505 = 646$$

- So, a student needs a Verbal SAT score of 646 in order to be in the top 10% of all students

The Standard Normal Distribution

Standard normal distribution

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

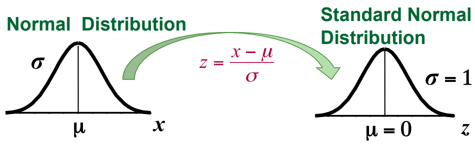


- Any x-value can be transformed into a z-score by using the formula

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

The Standard Normal Distribution

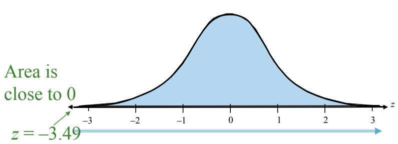
- If each data value of a normally distributed random variable x is transformed into a z-score, the result will be the standard normal distribution.



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

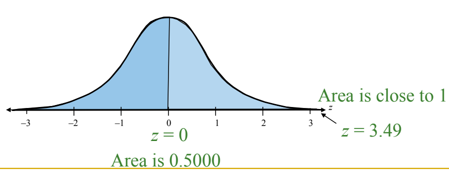
Properties of the Standard Normal Distribution

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

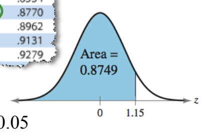
3. The cumulative area for $z = 0$ is 0.5000.
4. The cumulative area is close to 1 for z-scores close to $z = 3.49$.



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.

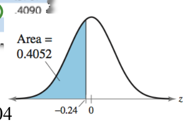
Move across the row to the column under 0.05

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	
−0.0	.4641	.4681	.4721	.4761	.4801	.4840	



Solution:

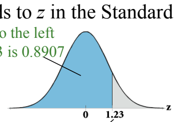
Find –0.2 in the left hand column.

Move across the row to the column under 0.04

The area to the left of $z = –0.24$ is 0.4052.

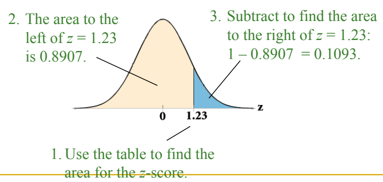
Finding Areas Under the Standard Normal Curve

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.



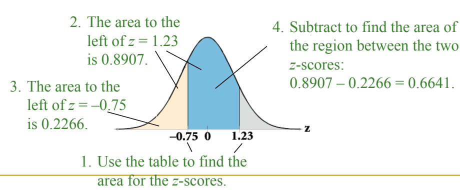
Finding Areas Under the Standard Normal Curve

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.



Finding Areas Under the Standard Normal Curve

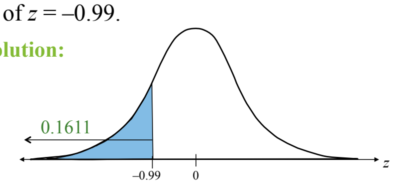
c. To find the area *between* two z-scores, find the area corresponding to each z-score in the Standard Normal Table. Then subtract the smaller area from the larger area.



Example: Finding Area Under the Standard Normal Curve

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

Solution:

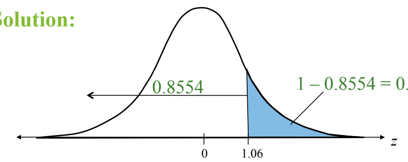


From the Standard Normal Table, the area is equal to 0.1611.

Example: Finding Area Under the Standard Normal Curve

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

Solution:

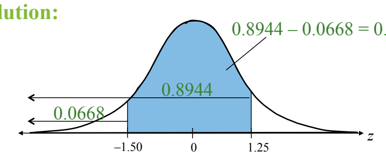


From the Standard Normal Table, the area is equal to 0.1446.

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.