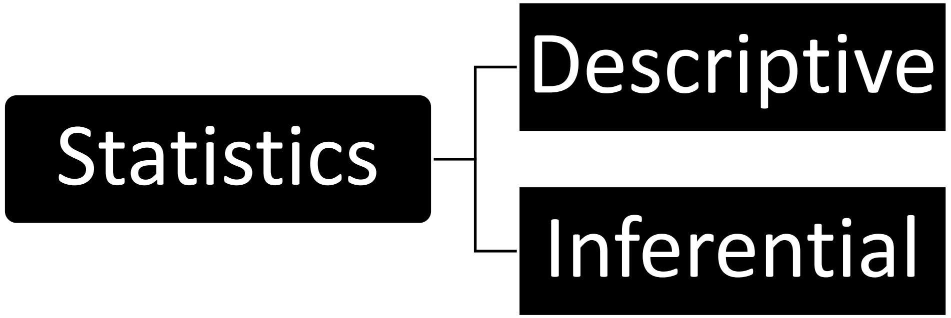
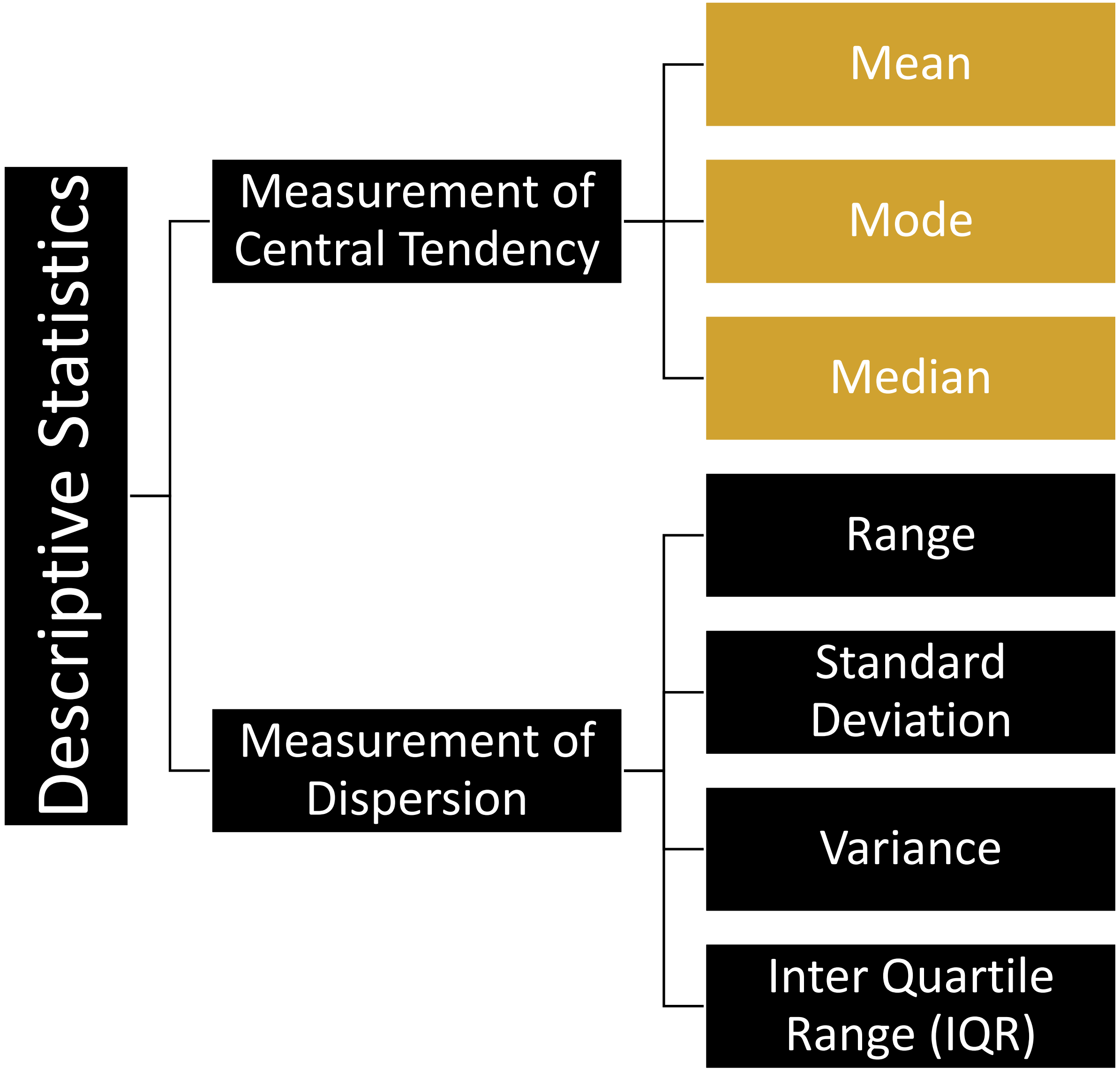
Section 2

Descriptive Statistics

Two Branches:



Hypothesis testing



- Descriptive Statistics
  - Measurement of Central Tendency
    - Mean
    - Mode
    - Median
  - Measurement of Dispersion
    - Range
    - Standard Deviation
    - Variance
    - Inter Quartile Range (IQR)

Basic Statistics

Mean or Average

- Mean = Average
- Mean of 104, 98, 90, 104, 104
- (104+98+90+104+104) / 5 = $\boxed{100}$

$$\bar{x} = \frac{\Sigma x}{n}$$

Mean

Median

- Median is the middle value when arranged in ascending or descending order.
- Median of 104, 98, 90, 104, 104
- Arranged in ascending order:
  - 90, 98, 104, 104, 104

Median

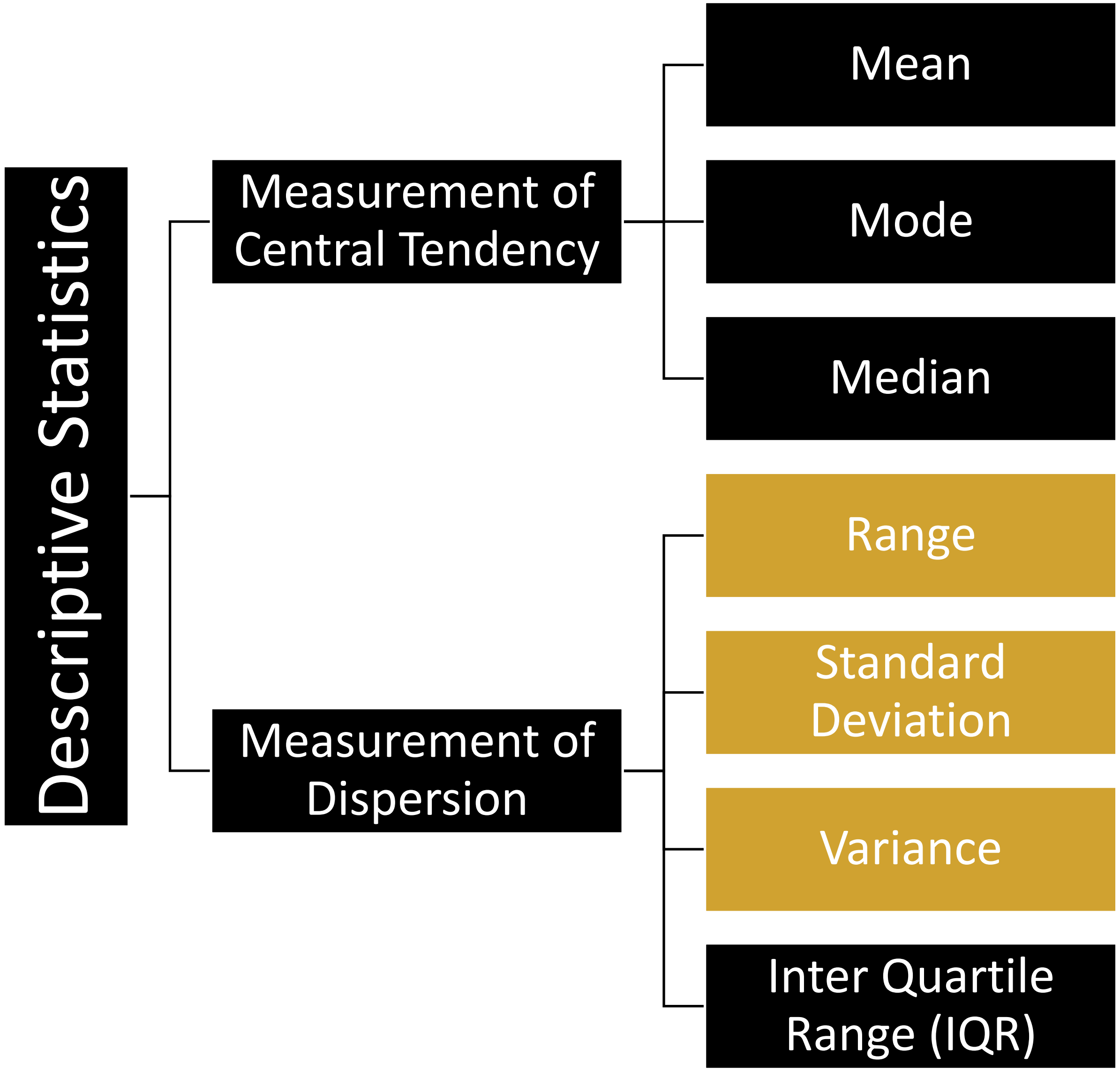
Median

- Median is the middle value when arranged in ascending or descending order.
- Median of 104, 98, 90, 104, 104, **85**
- Arranged in ascending order:
  - 85, 90, 98, 104, 104, 104

$$\frac{98 + 104}{2} = \frac{202}{2} = \boxed{\mathbf{101}}$$

Mode

- Mode is the most occurring number
- Mode of 104, 98, 90, 104, 104
- Is 104 as it occurred three times in the data



- Descriptive Statistics
  - Measurement of Central Tendency
    - Mean
    - Mode
    - Median
  - Measurement of Dispersion
    - Range
    - Standard Deviation
    - Variance
    - Inter Quartile Range (IQR)

Basic Statistics

Range

- ❖ Range = highest – lowest value
- ❖ Range of 104, 98, 90, 85, 104, 104
- ❖ 104 – 85 = $\boxed{19}$

Range

Standard Deviation

- Standard Deviation of 104, 98, 90, 104, 104

1. Find the average (mean) = 100
2. How far each item is from mean (104-100), (98-100), (90-100), (104-100), (104-100)
   - 4, -2, -10, 4, 4
3. Take square of the distance from mean $(4)^2$, $(-2)^2$, $(-10)^2$, $(4)^2$, $(4)^2$
   - 16, 4, 100, 16, 16
4. Take the mean of these squares (16+4+100+16+16)/5 = 30.4 (This is the **Variance**)
5. Square root of variance is the standard deviation

$$\boxed{\sqrt{30.4} = \mathbf{5.51}}$$

Variance and Standard Deviation

Standard Deviation

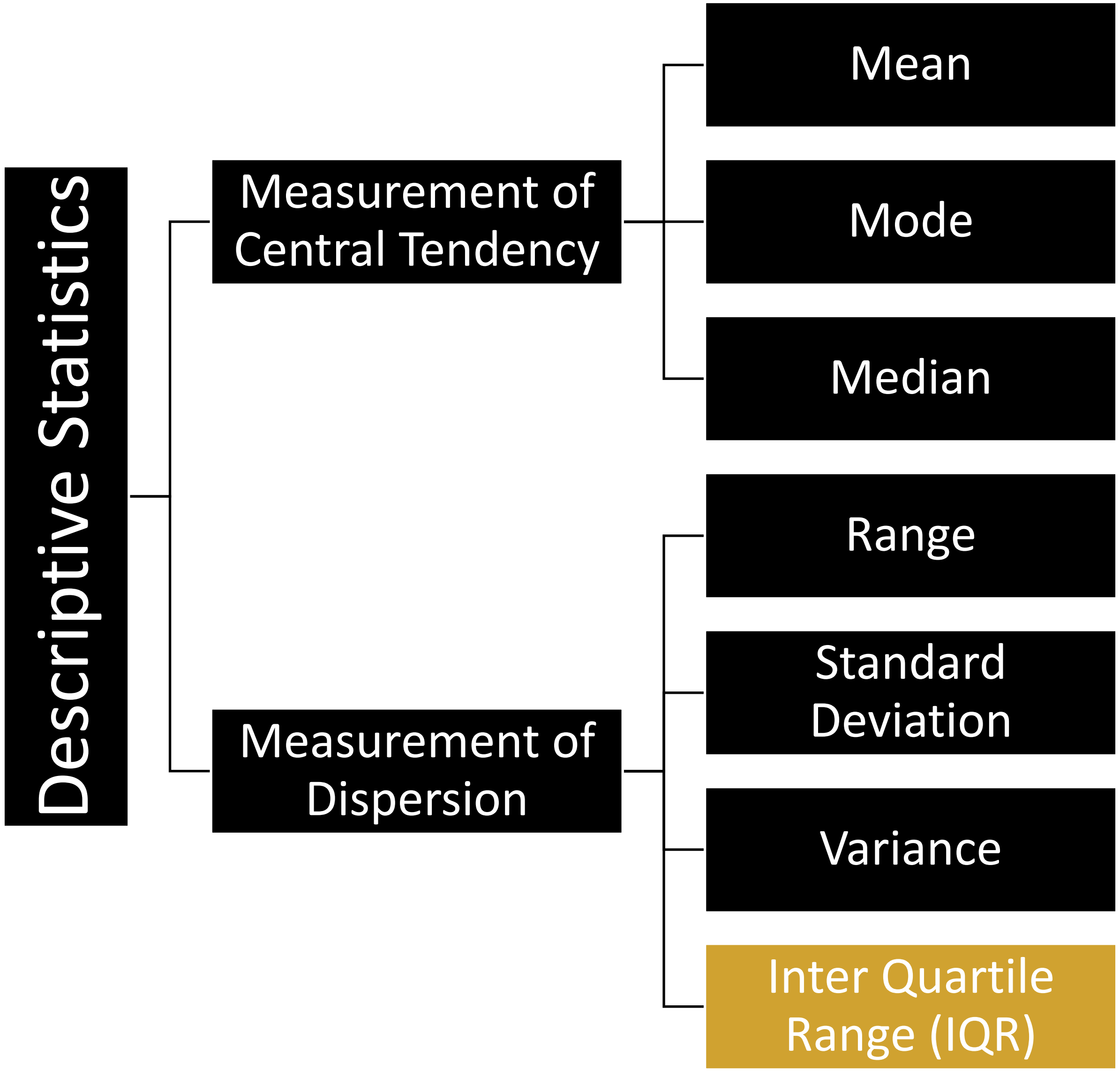
- Sample Standard Deviation

$$\mathrm{s} = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

- Population Standard Deviation

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{N}}$$

Variance and Standard Deviation



- Descriptive Statistics
  - Measurement of Central Tendency
    - Mean
    - Mode
    - Median
  - Measurement of Dispersion
    - Range
    - Standard Deviation
    - Variance
    - Inter Quartile Range (IQR)

Basic Statistics

Median

- Median is the middle value when arranged in ascending or descending order.
- Median of 104, 98, 90, 104, 104, **85**
- Arranged in ascending order:
  - 85, 90, 98, 104, 104, 104

$$\frac{98+104}{2}=\frac{202}{2}=\boxed{\mathbf{101}}$$

Median

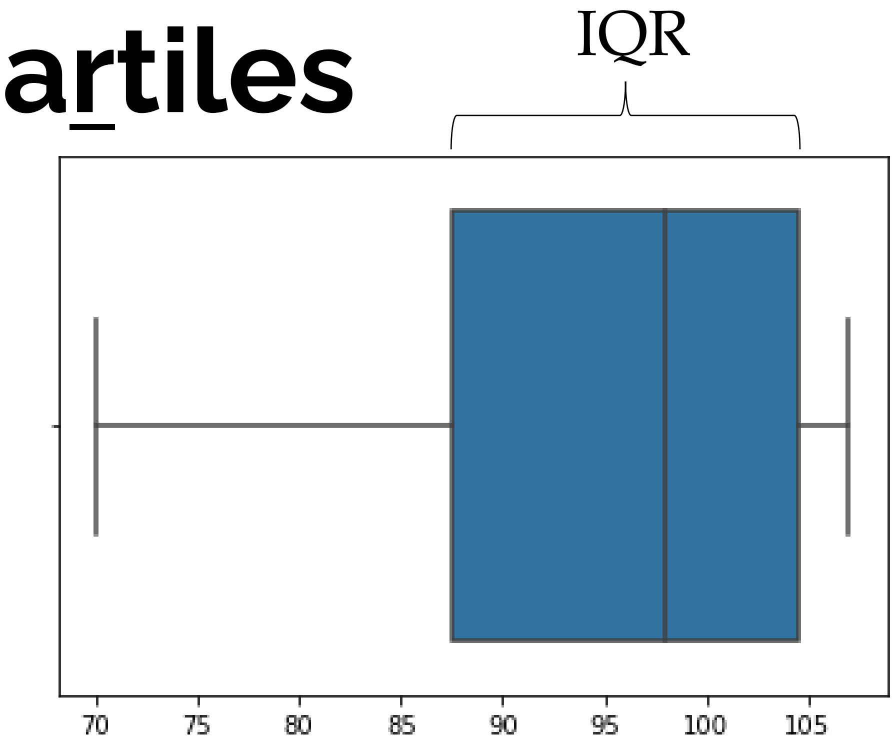
Quartiles

- Median divides the data into two parts. The single point of split is called the Median.
- If we divide the data into four parts, we get three points of splits called Quartiles (Q1, Q2 and Q3).
- Example: 70, 85, 90, 98, 104, 105, 107

Q1 Q2 Q3

Quartile

Quartiles



- Example: 70, 85, 90, 98, 104, 105, 107

Q1 Q2 Q3

Box and Whisker Plot

Quantiles

- Median divides the data into two parts. The single point of split is called the Median.
- If we divide the data into “n” parts, we get (n-1) points of splits called Quantiles.
- Example: Decile(10), Percentile(100)

Quantiles

Section 2

Descriptive Statistics

- Classic Model

$$\frac{\text{Number of outcomes in which the event occurs}}{\text{Total Number of possible outcomes of an experiment}}$$

- Relative Frequency of Occurrence

$$\frac{\text{Number of times an event occurred}}{\text{Total number of opportunities for an event to occur}}$$

Probability

Probability

- Experiment/Trial: Some thing done with an expectation of result.
- Event or Outcome: Result of experiment
- Sample Space: A sample space of an experiment is the set of all possible results of that random experiment.

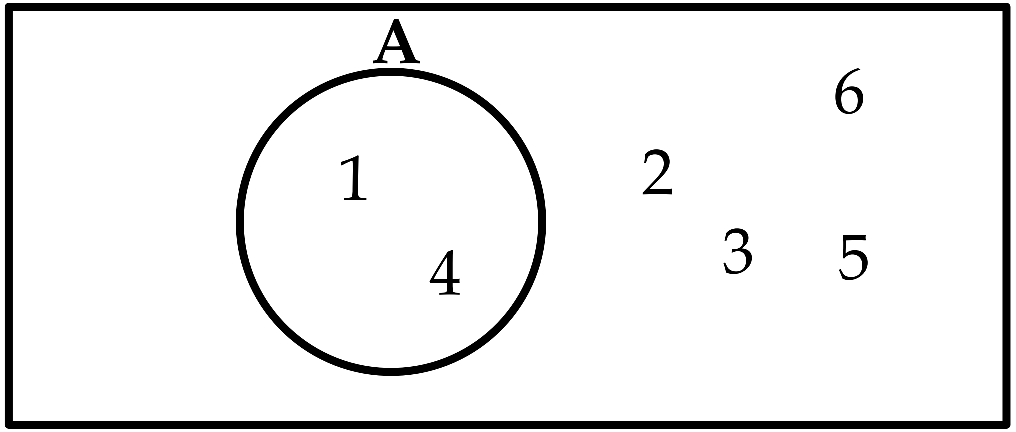
$$\{1, 2, 3, 4, 5, 6\}$$

- Sample space: In roll of two dices

{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6),
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6),
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6),
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6),}

- Venn diagram:

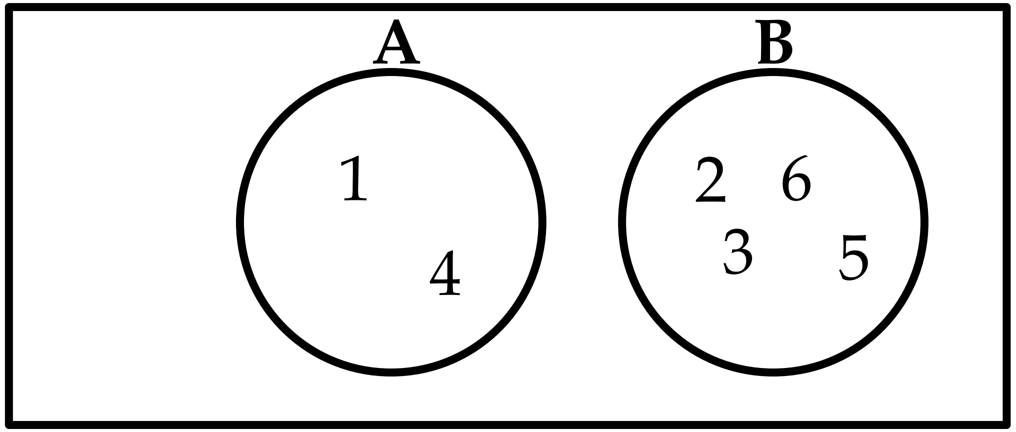
  Event A: Probability of getting 1 or 4 in the roll of a dice.



- Venn diagram:

Event A: Probability of getting 1 or 4 in the roll of a dice.

Event B: Probability of getting 2, 3, 5 or 6



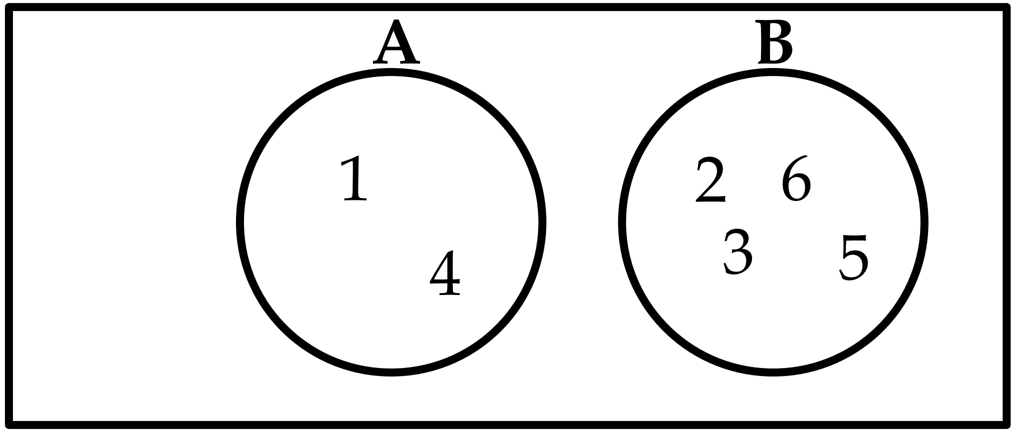
Probability

- Venn diagram:

  Event A: Probability of getting 1 or 4 in the roll of a dice.

  Event B: Probability of getting 2, 3, 5 or 6

Mutually Exclusive Events: When two events cannot occur at the same time

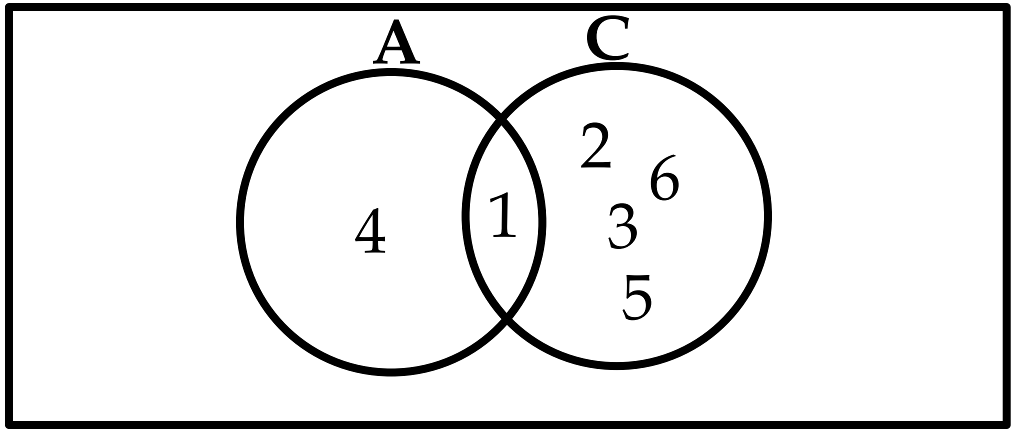


Probability

- Venn diagram:

  Event A: Probability of getting 1 or 4 in the roll of a dice.

  Event C: Probability of getting 1, 2, 3, 5 or 6



Probability

- Union:

  Probability that events A <u>or</u> B occur

  $P(A \cup B)$

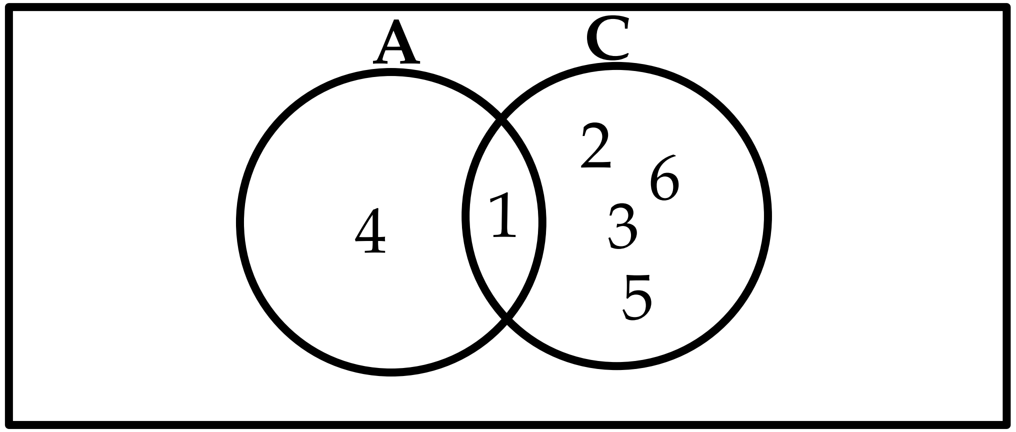
  $\{1, 2, 3, 4, 5, 6\}$

- Intersection:

  Probability that events A <u>and</u> B occur

  $P(A \cap B)$

  $\{1\}$



Probability

- Mutually Exclusive Events: When two events cannot occur at the same time

- Independent Events: The occurrence of Event A does not change the probability of Event B

- Complementary Events: The probability that Event A will <u>NOT</u> occur is denoted by P(A').

Probability

- Rule of Multiplication:

The probability that Events A **and** B both occur =

Probability that Event A occurs

x

Probability that Event B occurs, given that A has occurred

$$\mathbf{P(A \cap B) = P(A)\ P(B|A)}$$

Probability

- Rule of Multiplication:

The probability that Events A **and** B both occur =

Probability that Event A occurs

x

Probability that Event B occurs, given that A has occurred

$$\mathbf{P(A \cap B) = P(A)\ P(B|A)}$$

- In two rolls of dice what is the probability of getting 6 in both? (independent event)

Probability

- Rule of Multiplication:

The probability that Events A **and** B both occur =

Probability that Event A occurs

x

Probability that Event B occurs, given that A has occurred

$$P(A \cap B) = P(A)\, P(B|A)$$

- There are 10 candies in the plate (5 Green, 2 Yellow, 2 Orange and 1 Red). If I pick 2 random ones, what is the probability of getting both Yellow?

Probability

❖ Rule of Addition

The probability that Event A **or** Event B occurs

=

Probability that Event A occurs

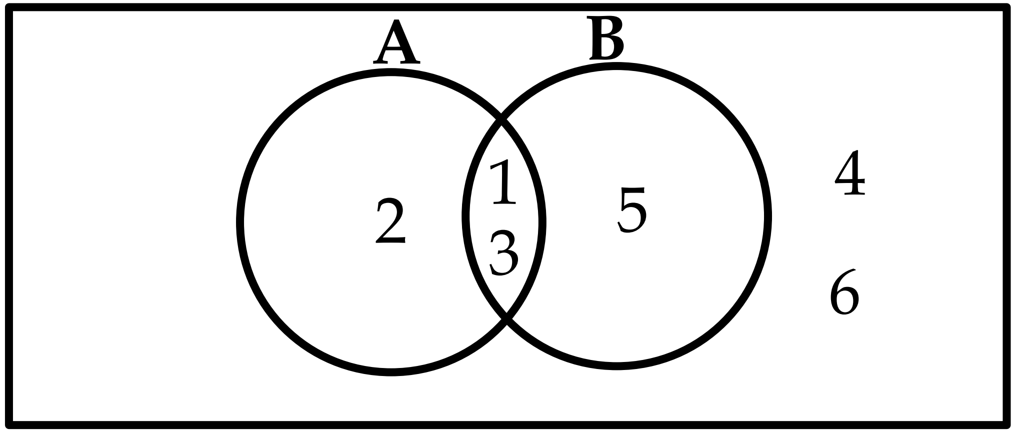
+

Probability that Event B occurs

-

Probability that both Events A and B occur

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Probability

Factorial

- **Factorial** of a non-negative integer n, denoted by n!, is the product of all positive integers less than or equal to n

$$5! = 5 \times 4 \times 3 \times 2 \times 1$$

$$1! = 1$$

$$0! = 1$$

Python code:

```
Import math
Math.factorial(5)
```

- **Permutation:** A set of objects in which position (or order) is important.

- **Combination:** A set of objects in which position (or order) is NOT important.

Permutations
Combinations

	Without repetition	**Python code:**
❖ **Permutation:** Order is important.	$n_{P_r} = \frac{n!}{(n-r)!}$	`import math` `math.perm(5,3)`
❖ **Combination:** Order is NOT important.	$n_{C_r} = \frac{n!}{(n-r)!\,\mathrm{r}!}$	`import math` `math.comb(5,3)`

PERMUTATIONS / COMBINATIONS SUMMARY

Permutation: With repetition

- In case of the lock, total permutations are
- 10 x 10 x 10 x 10 = 10,000

$$n^r$$

Python code:

```
Import math
math.pow(10, 4)
```

Permutations

With repetition

Permutation: Without repetition

- How many ways we can select 3 players out of 5. The first selected becomes the captain, second the vice-captain and third the treasurer.

$$n_{P_r} = \frac{n!}{(n-r)!}$$

$$\mathbf{5}_{P_{\mathbf{3}}} = \frac{\mathbf{5}!}{(\mathbf{5}-\mathbf{3})!} = 5 \text{ x } 4 \text{ x } 3 = 60$$

Python code:

```
import math
math.perm(5,3)
```

Permutations

Without repetition

- **Combination:** Without repetition:

- How many ways we can select 3 players out of 5

$$n_{C_r} = \frac{n!}{(n-r)!\,\mathrm{r}!}$$

$$\mathbf{5}_{C_3} = \frac{5!}{(5-3)!\,3!}$$

$$\mathbf{5}_{C_3} = \frac{5 \text{ x } 4 \text{ x } 3}{3 \text{ x } 2 \text{ x } 1} = 10$$

Python code:

```
import math
math.comb(5,3)
```

Combinations

Without repetition

❖ **Combination:** With repetition:

- ❖ e.g. In the store there are 5 varieties of juice bottles. You want to buy 3 bottles. How many possible combinations you can buy?

~~Without repetition: $^{n}C_r = \frac{n!}{(n-r)!\,\mathrm{r}!}$~~

$$\frac{(r+n\ -1)!}{r!\,(n-1)!}$$

$$\frac{(5+3\ -1)!}{3!(5-1)!} = \frac{7!}{3!4!} = \frac{7\text{x } 6 \text{ x } 5}{3 \text{ x } 2 \text{ x } 1} = 35$$

Python code:

```
n = 5
r = 3
Math.factorial(r+n-1)/(math.factorial(r) * math.factorial(n-1))
```

Combinations

With repetition

	Without repetition	**Python code:**
❖ **Permutation:** Order is important.	$n_{P_r} = \dfrac{n!}{(n-r)!}$	import math math.perm(5,3)
❖ **Combination:** Order is NOT important.	$n_{C_r} = \dfrac{n!}{(n-r)!\,\mathrm{r}!}$	import math math.comb(5,3)

PERMUTATIONS / COMBINATIONS SUMMARY

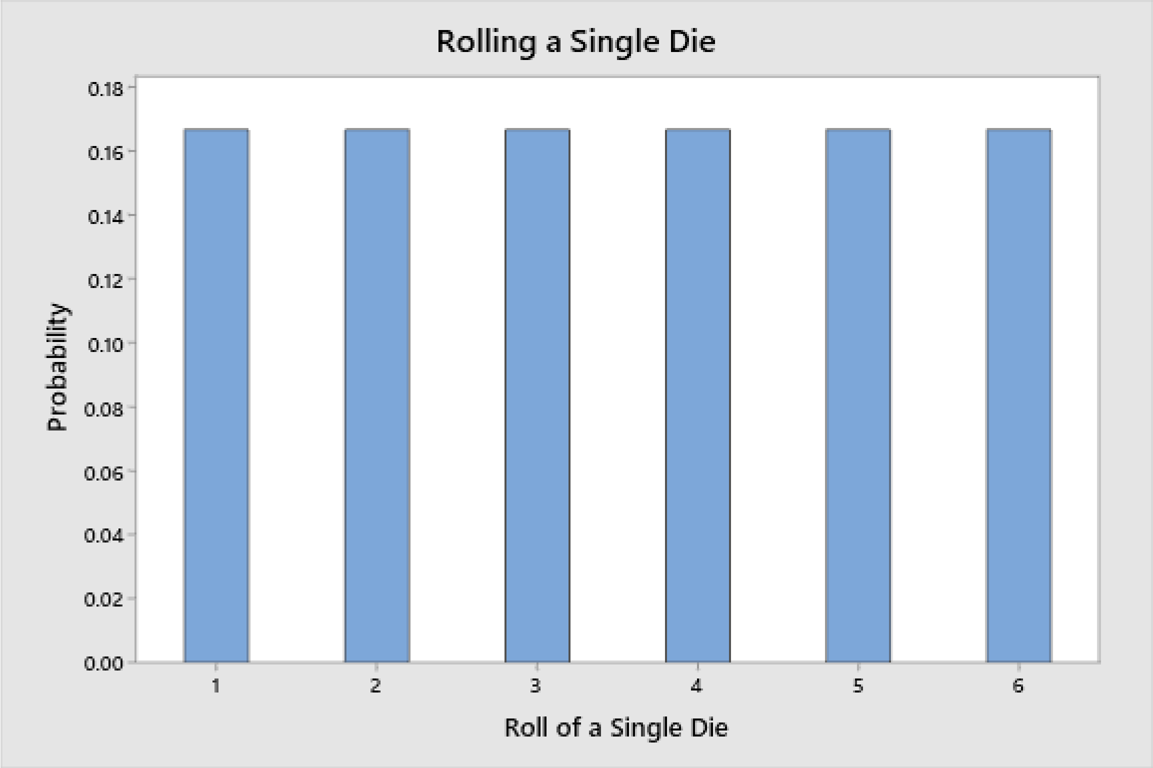
Probability Distributions

01 Binomial

02 Poisson

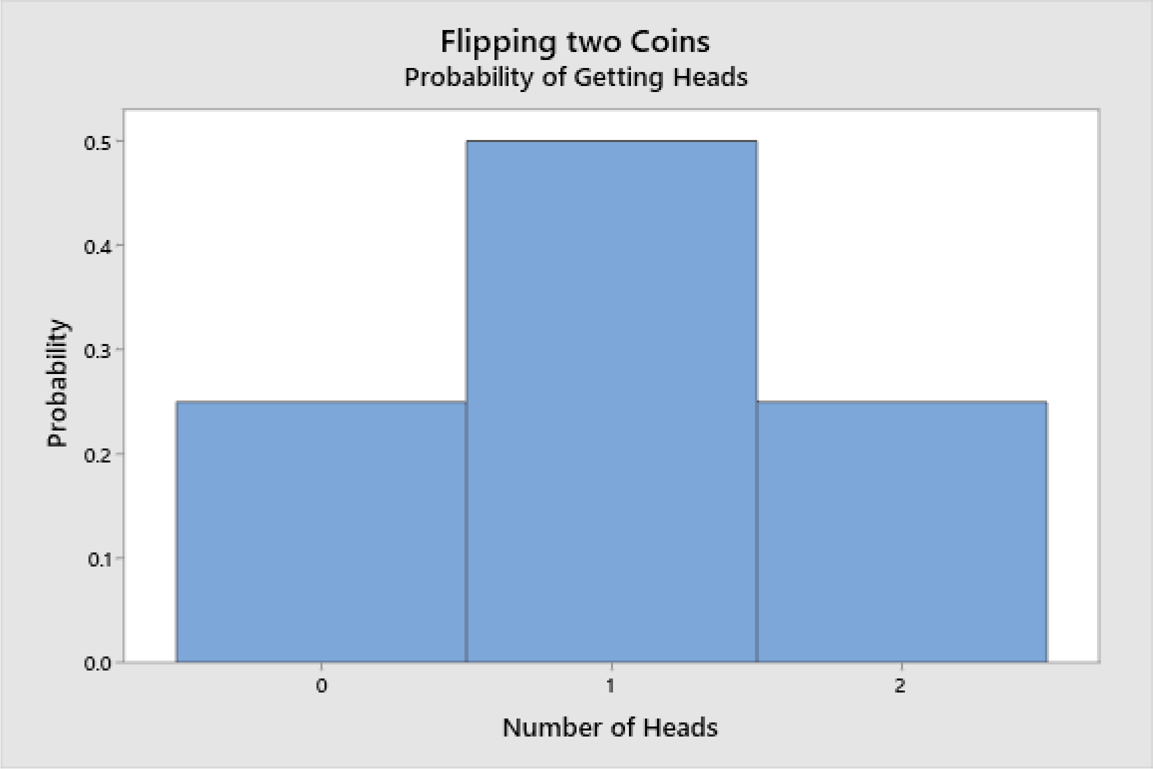
03 Normal

Rolling a Die



Distributions

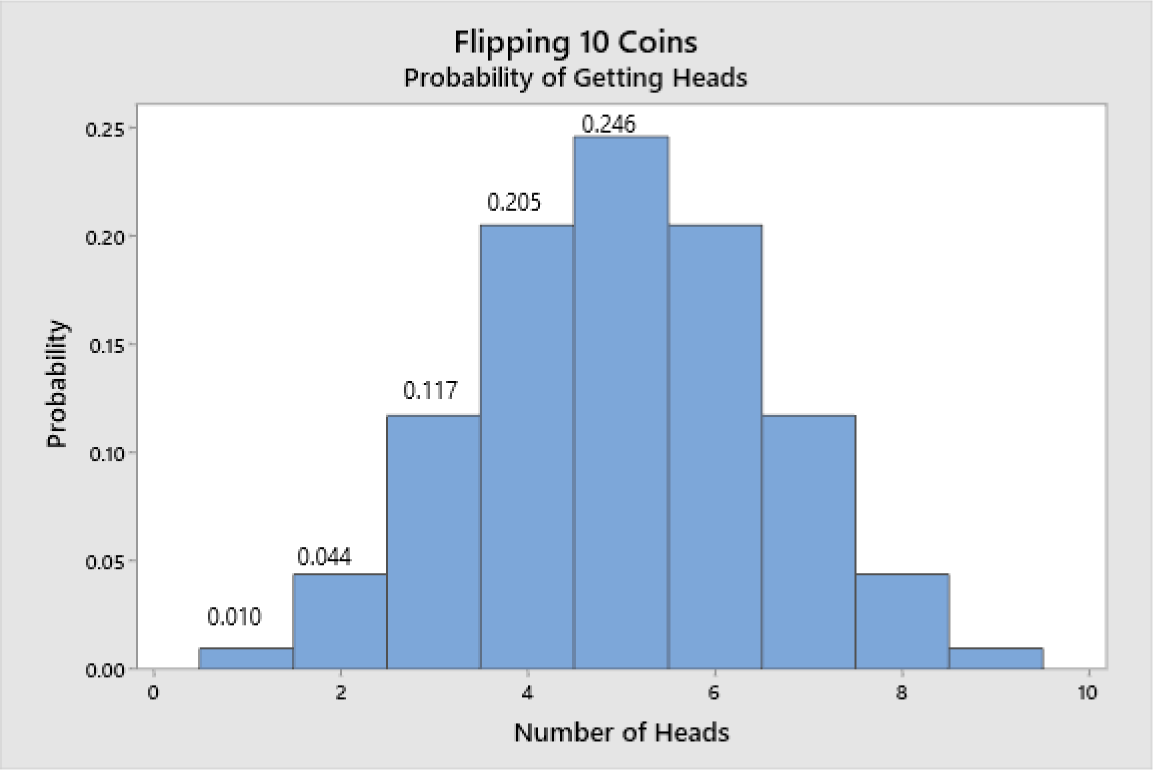
Flipping 2 Coins



1st	2nd
H	H
H	T
T	H
T	T

Distributions

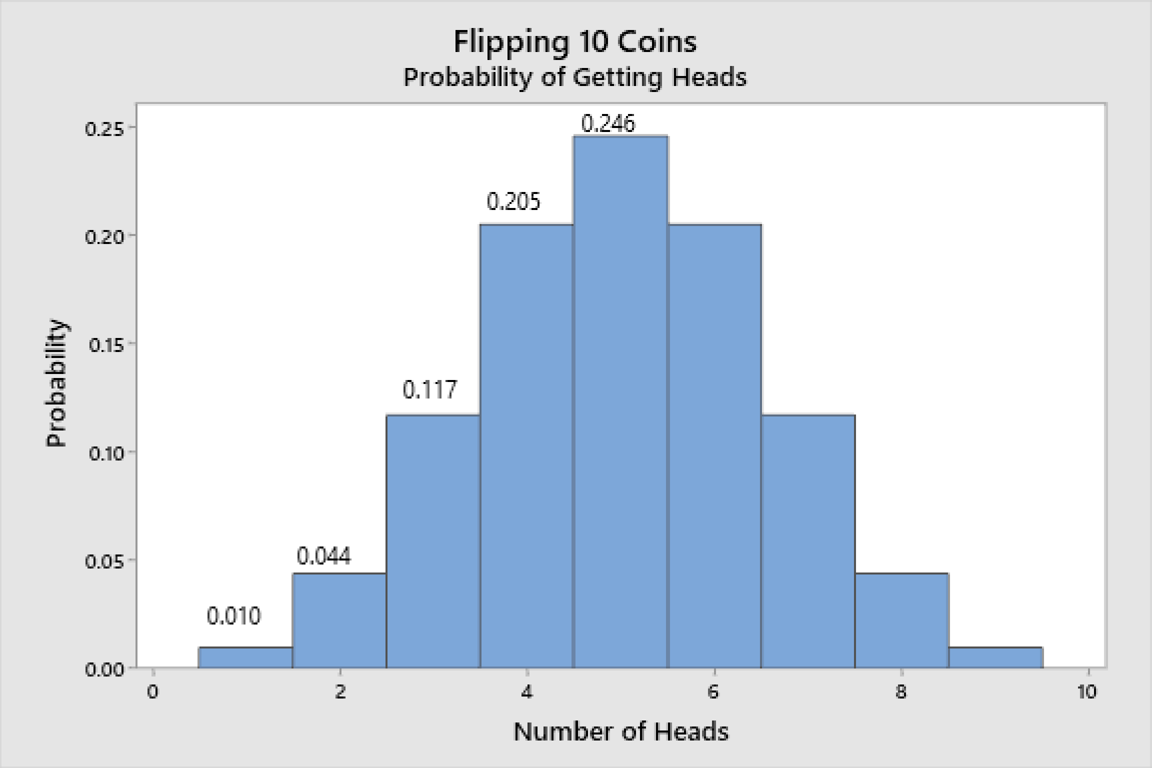
Flipping 10 Coins



Distributions

Probability Distribution

- Discrete (count) vs Continuous (measurement)
- Sum of area = 1.00



Distributions

Binomial Distribution

- A manufacturer has 12% defects rate in production. The buyer decides to test 20 random pieces and will accept the supplier if there are 2 or less defectives. What is the probability of getting accepted?

- **A binomial experiment** has the following properties:
  - The experiment consists of *n* repeated trials.
  - Each trial can result in just two possible outcomes. We call one of these outcomes a success and the other, a failure.
  - The probability of success, denoted by *p*, is the same on every trial.
  - The trials are independent; that is, the outcome on one trial does not affect the outcome on other trials.

Binomial Distribution

$$\mathbf{P(x) = n_{C_x} \cdot p^x \cdot (1 - p)^{n-x}}$$

$$\mathbf{P(x) = \frac{n!}{x!\,(n-x)!} . p^x \cdot (1 - p)^{n-x}}$$

- **x**: The number of successes that result from the binomial experiment.
- **n**: The number of trials in the binomial experiment.
- **p**: The probability of success on an individual trial.
- **P(x)** : Binomial probability - the probability that an *n*-trial binomial experiment results in exactly *x* successes, when the probability of success on an individual trial is *p*.

Binomial Distribution

- If you flip a coin 4 times, what is the probability of getting 1 head?

$$\mathbf{P(x) = \frac{n!}{x!\,(n-x)!}\,.\,p^{x} \cdot (1-p)^{n-x}}$$

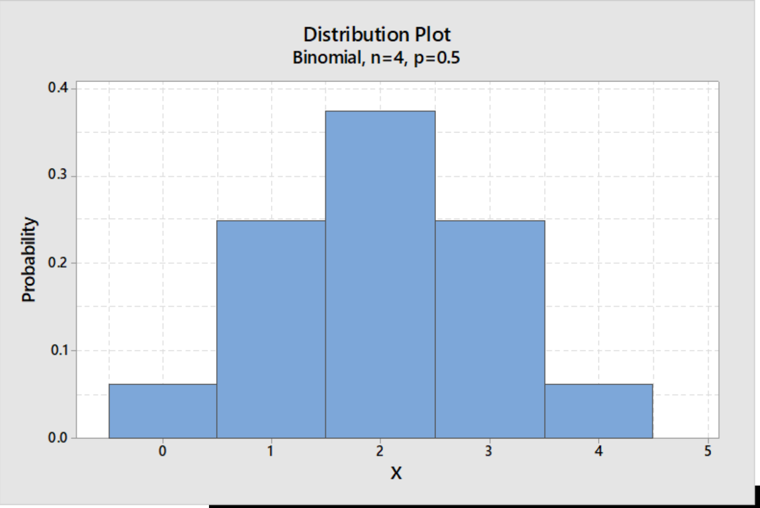
$$\mathbf{P(1) = \frac{4!}{1!\,(4-1)!}\,.\,0.5^{1} \cdot (1-0.5)^{4-1}}$$

$$\mathbf{P(1) = (4).\,(0.5)^{1} \cdot (0.5)^{3}}$$

$$\mathbf{P(1) = (4).\,(0.5)^{4} = 0.25}$$

BINOM.DIST(1,4,0.5,FALSE) = 0.25

BINOM.DIST(2,4,0.5,FALSE) = 0.375



Binomial Distribution

- The mean of the distribution (μ_x) is

  $$n \,.\, p$$

- The variance (σ^2_x) is

  $$n \,.\, p \,.\, (\,1 - p\,)$$

- The standard deviation (σ_x) is

  $$\sqrt{\mathbf{n} \cdot \mathbf{p} \cdot (\mathbf{1} - \mathbf{p})}$$

Binomial Distribution

- The mean of the distribution (μ_x) is

  $$\boldsymbol{n\ .\ p} = 4 \times 0.5 = 2$$

- The variance (σ^2_x) is

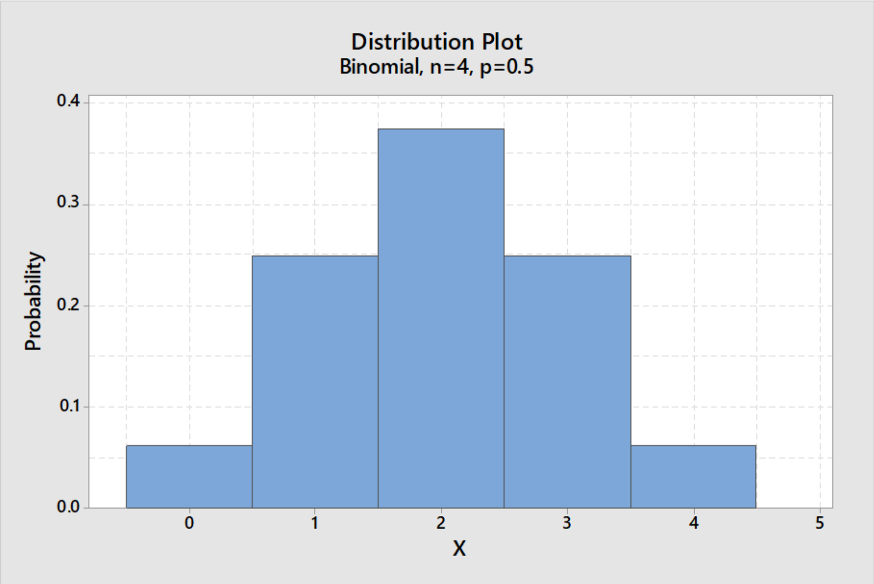
  $$\boldsymbol{n\ .\ p\ .\ (\ 1 - p\ )}$$

  $$4 \times 0.5 \times 0.5 = 1$$

- The standard deviation (σ_x) is

  $$\sqrt{\mathbf{n \cdot p \cdot (1 - p)}}$$

  $$\sqrt{4 \times 0.5 \times 0.5} = 1$$



Binomial Distribution

- A manufacturer has 12% defects rate in production. The buyer decides to test 20 random pieces and will accept the supplier if there are 2 or less defectives. What is the probability of getting accepted?

Binomial Distribution

- p = 0.12, n = 20, x = 0, 1, 2

$$\mathbf{P(x) = \frac{n!}{x!\,(n-x)!}\,.\,p^{x} \cdot (1-p)^{n-x}}$$

$$\mathbf{P(0) = \frac{20!}{0!\,(20-0)!}\,.\,0.12^{0} \cdot (1-0.12)^{20-0}}$$

$$\mathbf{P(1) = \frac{20!}{1!\,(20-1)!}\,.\,0.12^{1} \cdot (1-0.12)^{20-1}}$$

$$\mathbf{P(2) = \frac{20!}{2!\,(20-2)!}\,.\,0.12^{2} \cdot (1-0.12)^{20-2}}$$

P (0,1, 2) = 0.077563 + 0.211535 + 0.274034 = 0.563132

Binomial Distribution

- $p = 0.12$, $n = 20$, $x = 0, 1, 2$
- The mean of the distribution (μ_x) is

  $\boldsymbol{n \cdot p} =$ 20 x 0.12 = 2.4

- The variance (σ^2_x) is

  $\boldsymbol{n \cdot p \cdot (1 - p)}$

  20 x 0.12 x 0.88 = 2.112

- The standard deviation (σ_x) is

  $\sqrt{\mathbf{n \cdot p \cdot (1 - p)}}$ = 1.453

Binomial Distribution

Probability Distributions

01 Binomial

02 Poisson

03 Normal

Binomial vs Poisson Distribution

Similarities:

- Both are for discrete distribution
- Both measure the number of successes

Differences:

- In Poisson distribution the possibilities of success are infinite.

A **Poisson experiment** has the following properties:

- The experiment results in outcomes that can be classified as successes or failures.
- The average number of successes (μ) that occurs in a specified region is known.
- Outcomes are random. Occurrence of one outcome does not influence the chance of another outcome of interest.
- The outcomes of interest are rare relative to the possible outcomes.
  - Example: Road accidents, queue at the counter

Poisson Distribution

$$P(x,\mu) = e^{-\mu} \cdot \frac{\mu^x}{x!}$$

- *e*: A constant equal to approximately 2.71828. (Actually, *e* is the base of the natural logarithm system)
- μ: The mean number of successes that occur in a specified region.
- *x*: The actual number of successes that occur in a specified region.
- P(*x*; μ): The **Poisson probability** that exactly *x* successes occur in a Poisson experiment, when the mean number of successes is μ.

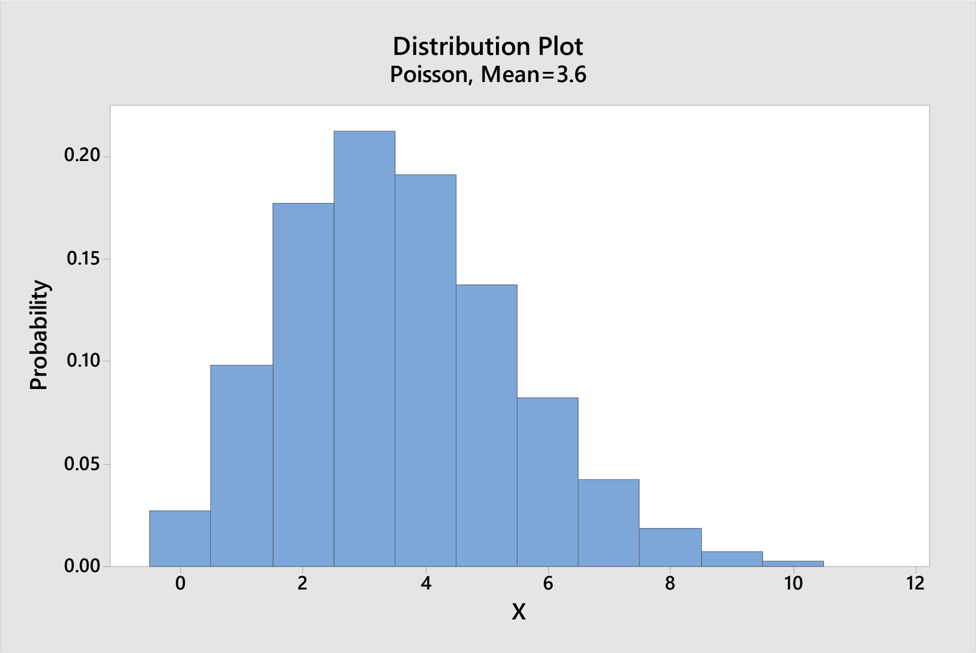
Poisson Distribution

$$P(x, \mu) = e^{-\mu} \cdot \frac{\mu^{x}}{x!}$$

- On a booking counter on the average 3.6 people come every 10 minutes on weekends. What is the probability of getting 7 people in 10 minutes?
- $\mu = 3.6$, $x=7$
- $P(x; \mu) = (e^{-\mu}) (\mu^{x}) / x! = (e^{-3.6}) (3.6^{7}) / 7!$
- $=0.02732 \times 7836.41 / 5040 = 0.0424$

Poisson Distribution

- On a booking counter on the average 3.6 people come every 10 minute on weekends. What is the probability of getting 7 people in 10 minutes?
- P(*7*; 3.6) = 0.0424



Poisson Distribution

- The Poisson distribution has the following properties:
- The mean of the distribution is equal to μ .
- The variance is also equal to μ .

Poisson Distribution

Probability Distributions

01 Binomial

02 Poisson

03 Normal

- Symmetrically distributed

- Long Tails / Bell Shaped

- Mean/ Mode and Median are same

Normal Distribution

- Two factors define the shape of the curve:
  - Mean
  - Standard Deviation

Normal Distribution

- About 68% of the area under the curve falls within **1 standard deviation** of the mean.

- About 95% of the area under the curve falls within **2 standard deviations** of the mean.

- About 99.7% of the area under the curve falls within **3 standard deviations** of the mean.

Normal Distribution

- The total area under the normal curve = 1.
- The probability of any particular value is 0.
- The probability that X is greater than or less than a value = area under the normal curve in that direction

Normal Distribution

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

- where x is a normal random variable,
- μ = mean,
- σ = standard deviation,
- π is approximately 3.14159,
- e is approximately 2.71828.

Normal Distribution

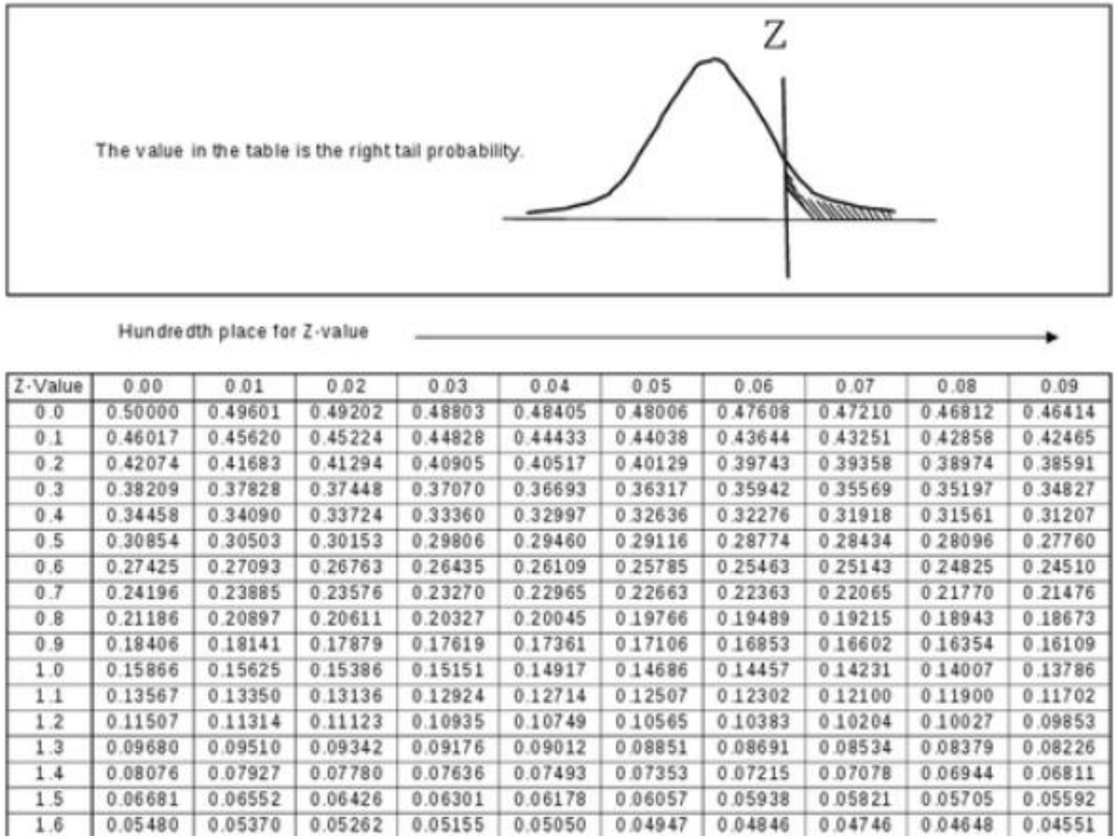
Z Value / Standard Score

- How many standard deviations an element is from the mean.
- $z = (x - \mu) / \sigma$

- z is the z-score,
- x is the value of the element,
- μ is the population mean,
- σ is the standard deviation.

Normal Distribution

Normal Distribution



The value in the table is the right tail probability.

Hundredth place for Z-value →

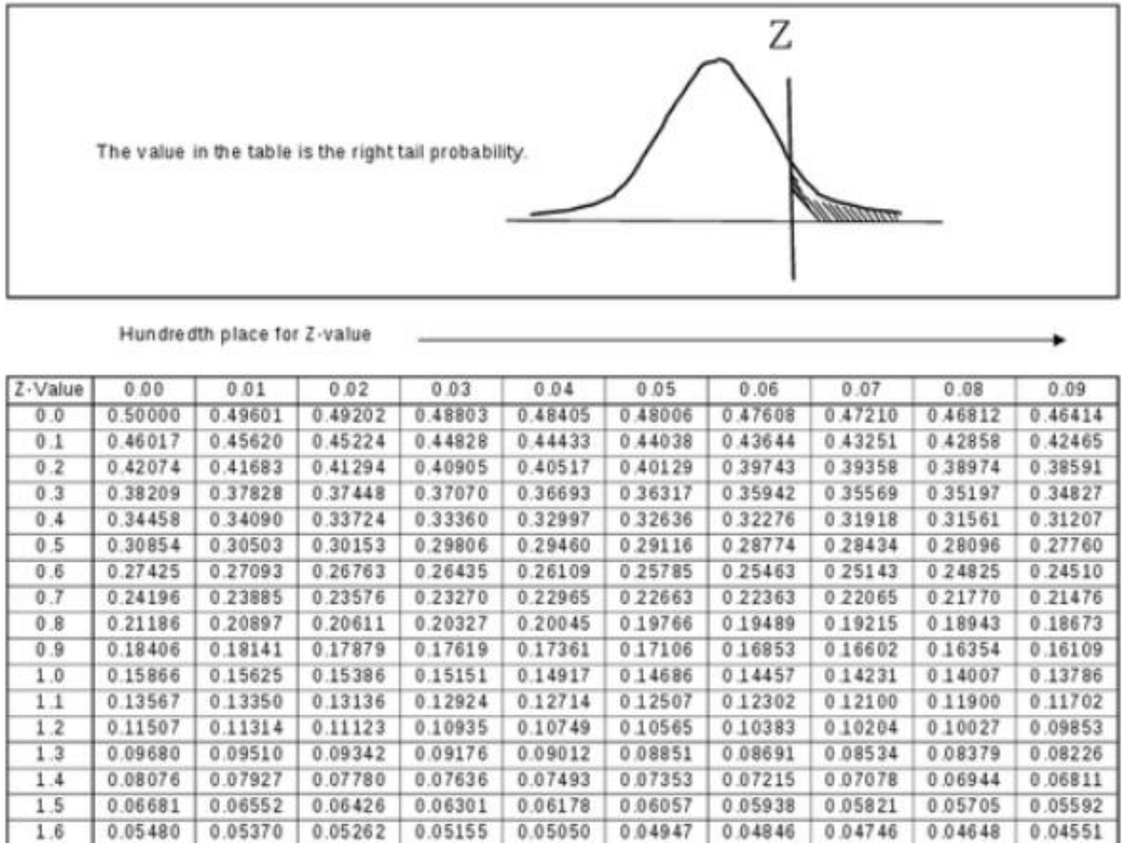
Z-Value	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.49601	0.49202	0.48803	0.48405	0.48006	0.47608	0.47210	0.46812	0.46414
0.1	0.46017	0.45620	0.45224	0.44828	0.44433	0.44038	0.43644	0.43251	0.42858	0.42465
0.2	0.42074	0.41683	0.41294	0.40905	0.40517	0.40129	0.39743	0.39358	0.38974	0.38591
0.3	0.38209	0.37828	0.37448	0.37070	0.36693	0.36317	0.35942	0.35569	0.35197	0.34827
0.4	0.34458	0.34090	0.33724	0.33360	0.32997	0.32636	0.32276	0.31918	0.31561	0.31207
0.5	0.30854	0.30503	0.30153	0.29806	0.29460	0.29116	0.28774	0.28434	0.28096	0.27760
0.6	0.27425	0.27093	0.26763	0.26435	0.26109	0.25785	0.25463	0.25143	0.24825	0.24510
0.7	0.24196	0.23885	0.23576	0.23270	0.22965	0.22663	0.22363	0.22065	0.21770	0.21476
0.8	0.21186	0.20897	0.20611	0.20327	0.20045	0.19766	0.19489	0.19215	0.18943	0.18673
0.9	0.18406	0.18141	0.17879	0.17619	0.17361	0.17106	0.16853	0.16602	0.16354	0.16109
1.0	0.15866	0.15625	0.15386	0.15151	0.14917	0.14686	0.14457	0.14231	0.14007	0.13786
1.1	0.13567	0.13350	0.13136	0.12924	0.12714	0.12507	0.12302	0.12100	0.11900	0.11702
1.2	0.11507	0.11314	0.11123	0.10935	0.10749	0.10565	0.10383	0.10204	0.10027	0.09853
1.3	0.09680	0.09510	0.09342	0.09176	0.09012	0.08851	0.08691	0.08534	0.08379	0.08226
1.4	0.08076	0.07927	0.07780	0.07636	0.07493	0.07353	0.07215	0.07078	0.06944	0.06811
1.5	0.06681	0.06552	0.06426	0.06301	0.06178	0.06057	0.05938	0.05821	0.05705	0.05592
1.6	0.05480	0.05370	0.05262	0.05155	0.05050	0.04947	0.04846	0.04746	0.04648	0.04551

Z Value / Standard Score

- ❖ Perfume bottles are filled with the average volume of 150 cc and the standard deviation of 2 cc.
- ❖ What percent of bottles will have the volume more than 153 cc?

$\mu = 150$ cc
$\sigma = 2$ cc

$z = (x - \mu) / \sigma = (153 – 150)/2 = 1.5$



The value in the table is the right tail probability.

Hundredth place for Z-value →

Z-Value	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.49601	0.49202	0.48803	0.48405	0.48006	0.47608	0.47210	0.46812	0.46414
0.1	0.46017	0.45620	0.45224	0.44828	0.44433	0.44038	0.43644	0.43251	0.42858	0.42465
0.2	0.42074	0.41683	0.41294	0.40905	0.40517	0.40129	0.39743	0.39358	0.38974	0.38591
0.3	0.38209	0.37828	0.37448	0.37070	0.36693	0.36317	0.35942	0.35569	0.35197	0.34827
0.4	0.34458	0.34090	0.33724	0.33360	0.32997	0.32636	0.32276	0.31918	0.31561	0.31207
0.5	0.30854	0.30503	0.30153	0.29806	0.29460	0.29116	0.28774	0.28434	0.28096	0.27760
0.6	0.27425	0.27093	0.26763	0.26435	0.26109	0.25785	0.25463	0.25143	0.24825	0.24510
0.7	0.24196	0.23885	0.23576	0.23270	0.22965	0.22663	0.22363	0.22065	0.21770	0.21476
0.8	0.21186	0.20897	0.20611	0.20327	0.20045	0.19766	0.19489	0.19215	0.18943	0.18673
0.9	0.18406	0.18141	0.17879	0.17619	0.17361	0.17106	0.16853	0.16602	0.16354	0.16109
1.0	0.15866	0.15625	0.15386	0.15151	0.14917	0.14686	0.14457	0.14231	0.14007	0.13786
1.1	0.13567	0.13350	0.13136	0.12924	0.12714	0.12507	0.12302	0.12100	0.11900	0.11702
1.2	0.11507	0.11314	0.11123	0.10935	0.10749	0.10565	0.10383	0.10204	0.10027	0.09853
1.3	0.09680	0.09510	0.09342	0.09176	0.09012	0.08851	0.08691	0.08534	0.08379	0.08226
1.4	0.08076	0.07927	0.07780	0.07636	0.07493	0.07353	0.07215	0.07078	0.06944	0.06811
1.5	0.06681	0.06552	0.06426	0.06301	0.06178	0.06057	0.05938	0.05821	0.05705	0.05592
1.6	0.05480	0.05370	0.05262	0.05155	0.05050	0.04947	0.04846	0.04746	0.04648	0.04551

Normal Distribution

Normal Distribution

Z Value / Standard Score

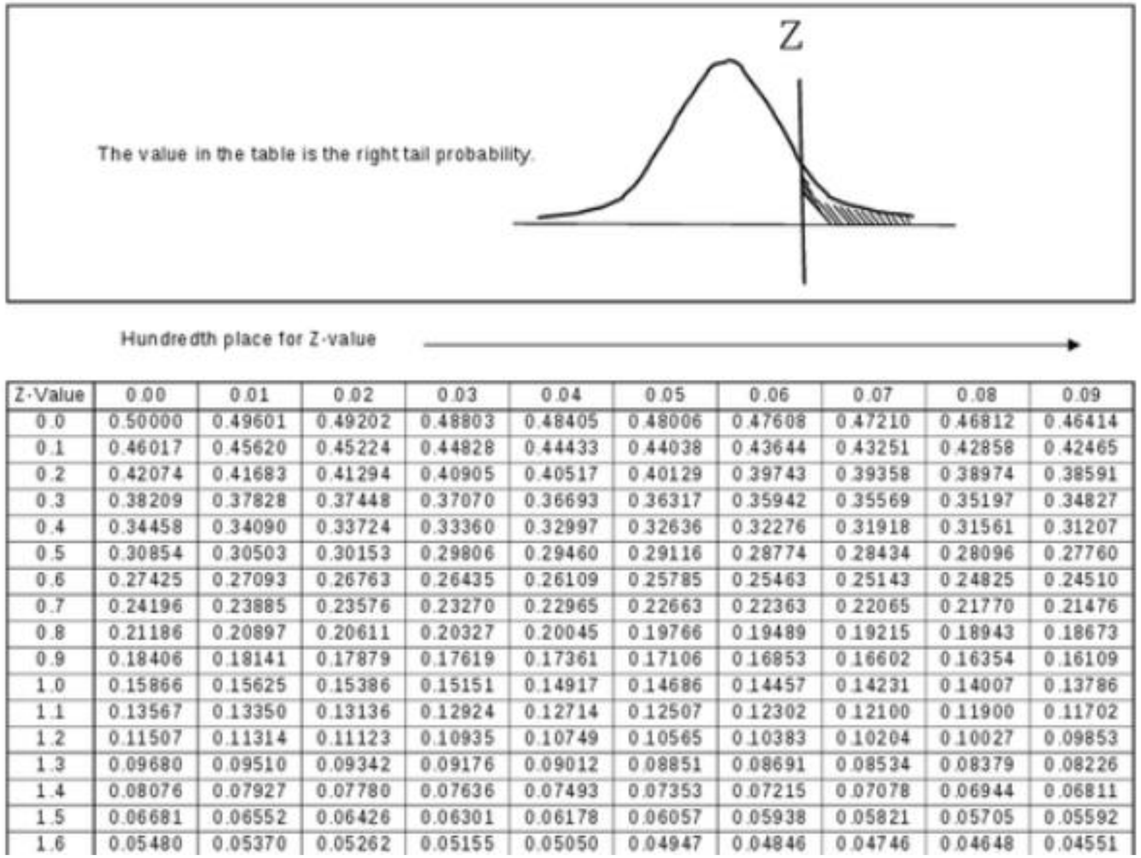
- ❖ Perfume bottles are filled with the average volume of 150 cc and the standard deviation of 2 cc.
- ❖ What percent of bottles will have the volume between 148 and 152 cc?

$\mu = 150$ cc
$\sigma = 2$ cc

z1 = (x - μ) / σ = (148 – 150)/2 = -1

z2 = (x - μ) / σ = (152 – 150)/2 = 1

$$P(x) = 1 - 0.15866 - 0.15866 = 0.68268$$



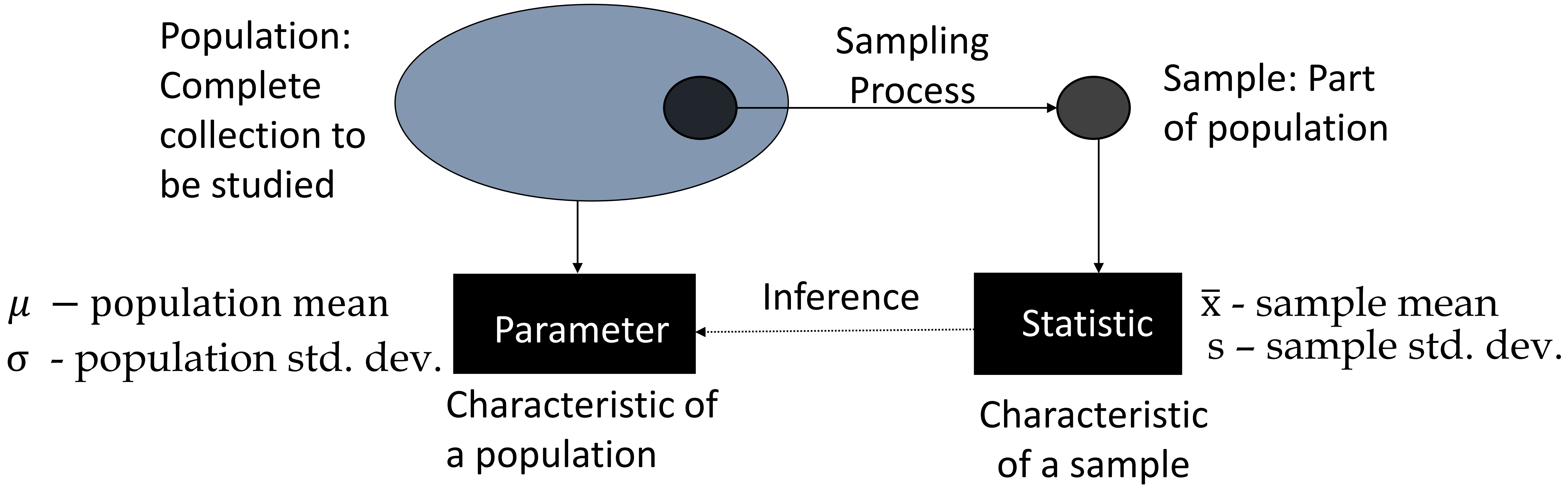
The value in the table is the right tail probability.

Hundredth place for Z-value

Z-Value	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.49601	0.49202	0.48803	0.48405	0.48006	0.47608	0.47210	0.46812	0.46414
0.1	0.46017	0.45620	0.45224	0.44828	0.44433	0.44038	0.43644	0.43251	0.42858	0.42465
0.2	0.42074	0.41683	0.41294	0.40905	0.40517	0.40129	0.39743	0.39358	0.38974	0.38591
0.3	0.38209	0.37828	0.37448	0.37070	0.36693	0.36317	0.35942	0.35569	0.35197	0.34827
0.4	0.34458	0.34090	0.33724	0.33360	0.32997	0.32636	0.32276	0.31918	0.31561	0.31207
0.5	0.30854	0.30503	0.30153	0.29806	0.29460	0.29116	0.28774	0.28434	0.28096	0.27760
0.6	0.27425	0.27093	0.26763	0.26435	0.26109	0.25785	0.25463	0.25143	0.24825	0.24510
0.7	0.24196	0.23885	0.23576	0.23270	0.22965	0.22663	0.22363	0.22065	0.21770	0.21476
0.8	0.21186	0.20897	0.20611	0.20327	0.20045	0.19766	0.19489	0.19215	0.18943	0.18673
0.9	0.18406	0.18141	0.17879	0.17619	0.17361	0.17106	0.16853	0.16602	0.16354	0.16109
1.0	0.15866	0.15625	0.15386	0.15151	0.14917	0.14686	0.14457	0.14231	0.14007	0.13786
1.1	0.13567	0.13350	0.13136	0.12924	0.12714	0.12507	0.12302	0.12100	0.11900	0.11702
1.2	0.11507	0.11314	0.11123	0.10935	0.10749	0.10565	0.10383	0.10204	0.10027	0.09853
1.3	0.09680	0.09510	0.09342	0.09176	0.09012	0.08851	0.08691	0.08534	0.08379	0.08226
1.4	0.08076	0.07927	0.07780	0.07636	0.07493	0.07353	0.07215	0.07078	0.06944	0.06811
1.5	0.06681	0.06552	0.06426	0.06301	0.06178	0.06057	0.05938	0.05821	0.05705	0.05592
1.6	0.05480	0.05370	0.05262	0.05155	0.05050	0.04947	0.04846	0.04746	0.04648	0.04551

Section 6

Inferential Statistics



Population: Complete collection to be studied

Sampling Process

Sample: Part of population

μ – population mean

σ - population std. dev.

Parameter

Characteristic of a population

Inference

Statistic

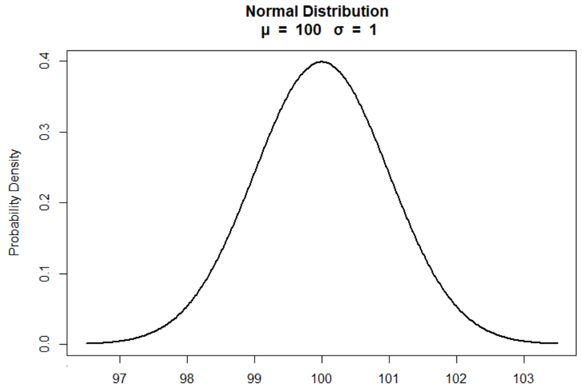
$\bar{x}$ - sample mean

s – sample std. dev.

Characteristic of a sample

SAMPLING

SAMPLING – CONFIDENCE INTERVAL

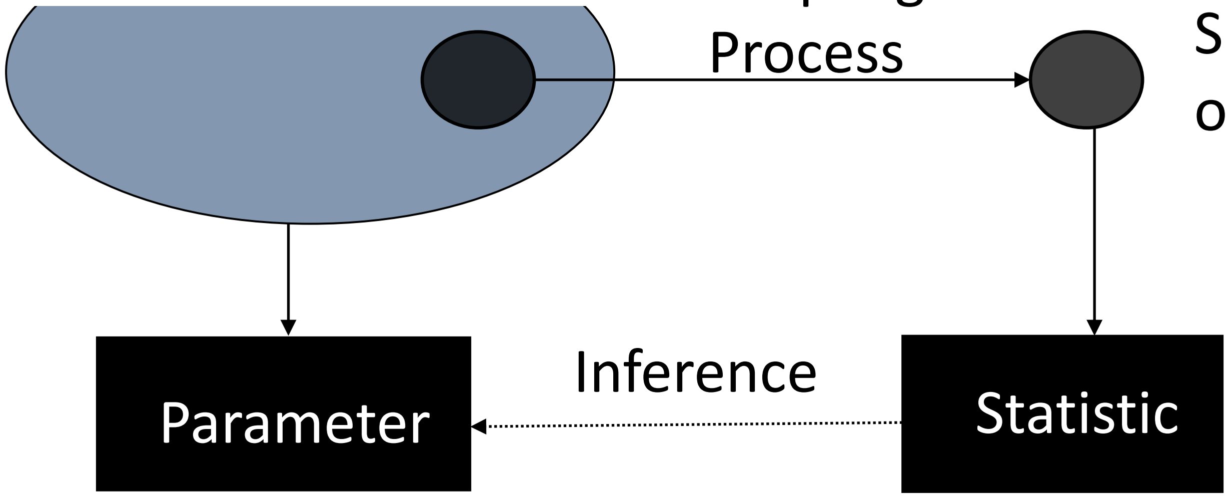


102 101
99 100

Mean = 100.5

$$CI = \bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$CI = 100.5 \pm 1$$



Population: Complete collection to be studied

Sampling Process

Sample: Part of population

Inference

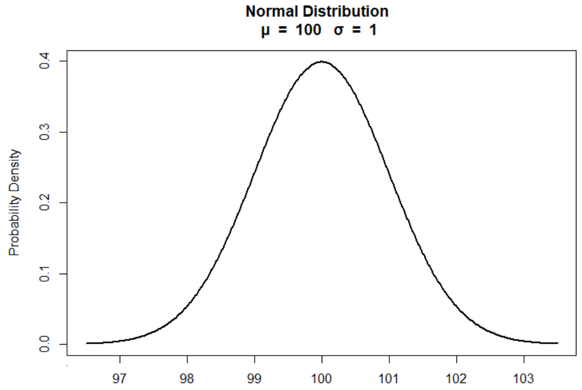
Parameter

Characteristic of a population

Statistic

Characteristic of a sample

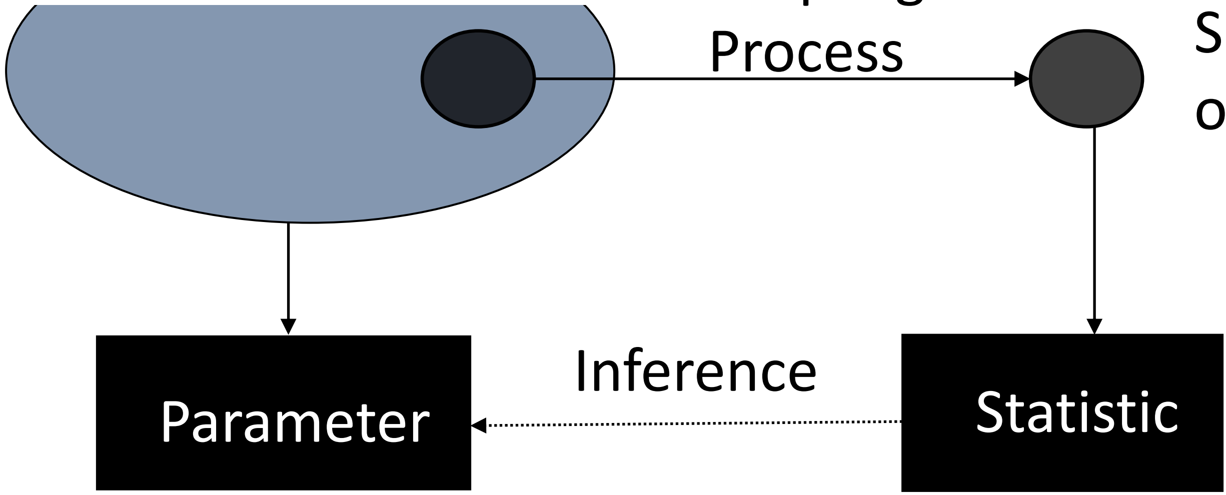
SAMPLING – HYPOTHESIS TEST



102 101
99 100

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Population: Complete collection to be studied



Sampling Process

Sample: Part of population

Parameter

Inference

Statistic

Characteristic of a population

Characteristic of a sample

- For almost all populations, the sampling distribution of the mean can be approximated closely by a normal distribution, provided the sample size is sufficiently large.

Central Limit Theorem

- If a variable has a mean of μ and the variance σ^2, as the sample size n increases, the sample mean approaches a normal distribution with mean $\mu_{\bar{x}}$ and variance $\boldsymbol{\sigma_{\bar{x}}^2}$

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

$$\boldsymbol{\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}}$$

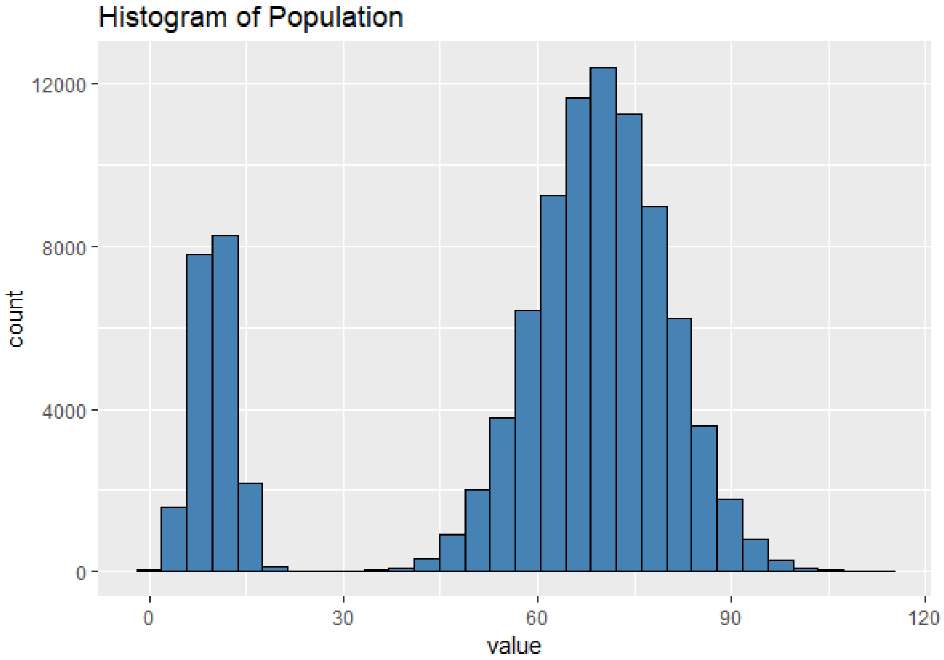
$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

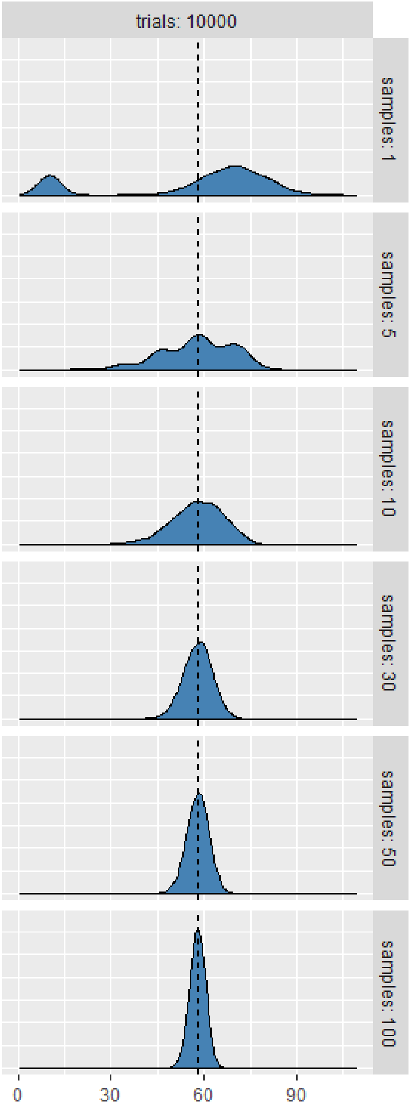
Central Limit Theorem

- Standard deviation of the sampling distribution of the sample mean
  - Called “standard error of the mean”

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

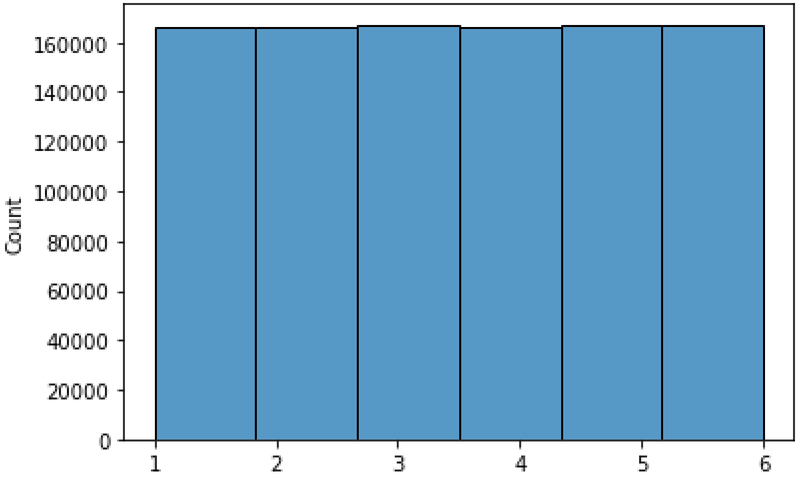
Central Limit Theorem



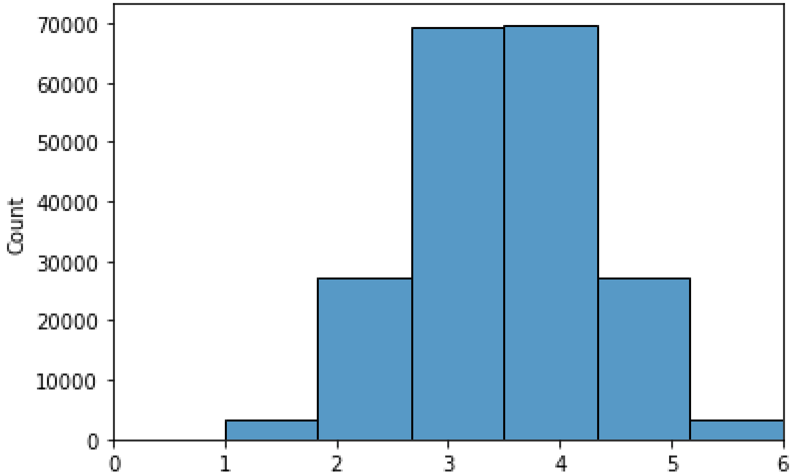


CENTRAL LIMIT THEOREM

CENTRAL LIMIT THEOREM

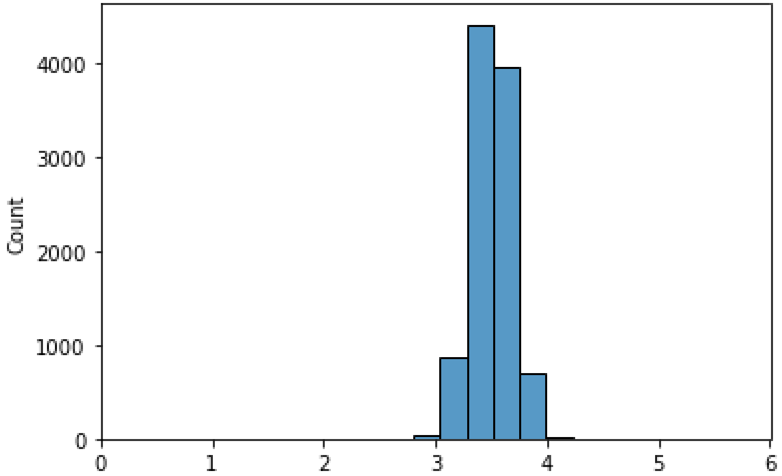


Sample size = 5



$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

Sample size = 100



Hypothesis Testing

01 *Statistical vs practical significance*

02 *Hypothesis testing steps*

03 *Type 1 and 2 errors*

04 *The p Value*

Hypothesis Testing

- A statistical hypothesis test is a method of statistical inference.
- Commonly used tests include:
  - Compare sample statistic with the population parameter
  - Compare two datasets

Hypothesis Testing

Statistical Significance

- Case of a perfume making company:
- Mean Volume 150 cc and sd=2 cc

Practical Significance

- Practical significance of an experiment tells us if there is any actionable information from the result.
- Large samples can find out statistical difference for very small difference. These small differences might not have practical significance.

Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g., t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing

Hypothesis Testing

- Null Hypothesis: The person is innocent
- Alternate Hypothesis: The person is guilty. You need to provide proof of this.
- Court conclusion is: Guilty or Not Guilty (not the innocent)

Hypothesis Testing

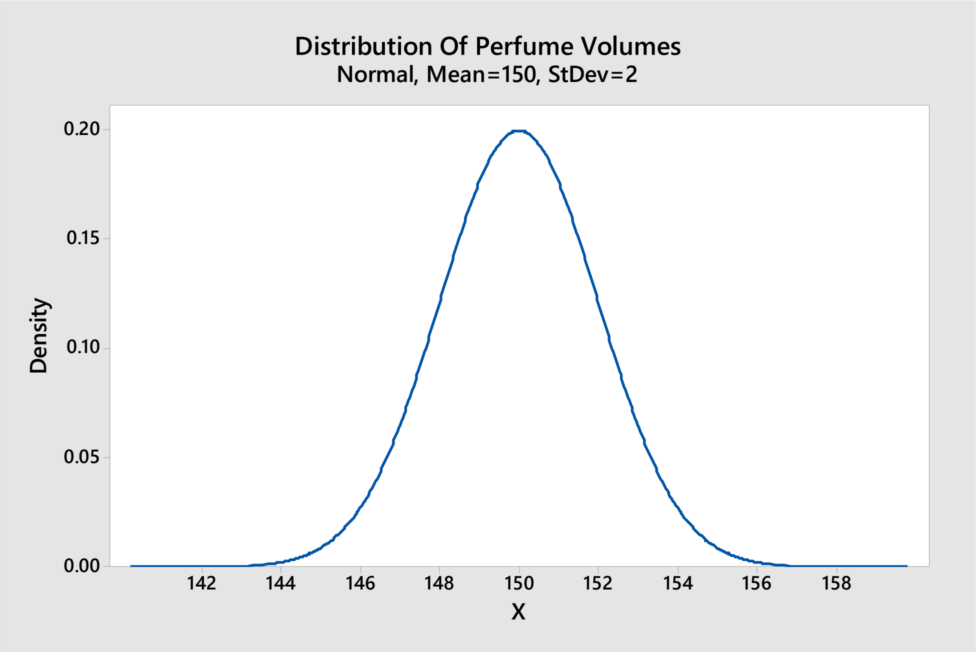
Hypothesis Testing

- Null Hypothesis: The person is innocent
- Alternate Hypothesis: The person is guilty. You need to provide proof of this.
- In statistical terms you:
  - Reject the Null Hypothesis, or
  - Fail to reject the Null Hypothesis (not accept the Null Hypothesis)

Hypothesis Testing

Hypothesis Testing

- Null Hypothesis: The machine is filling the bottles with 150 cc
- Alternate Hypothesis: The machine is "not" filling the bottles with 150 cc.



Hypothesis Testing

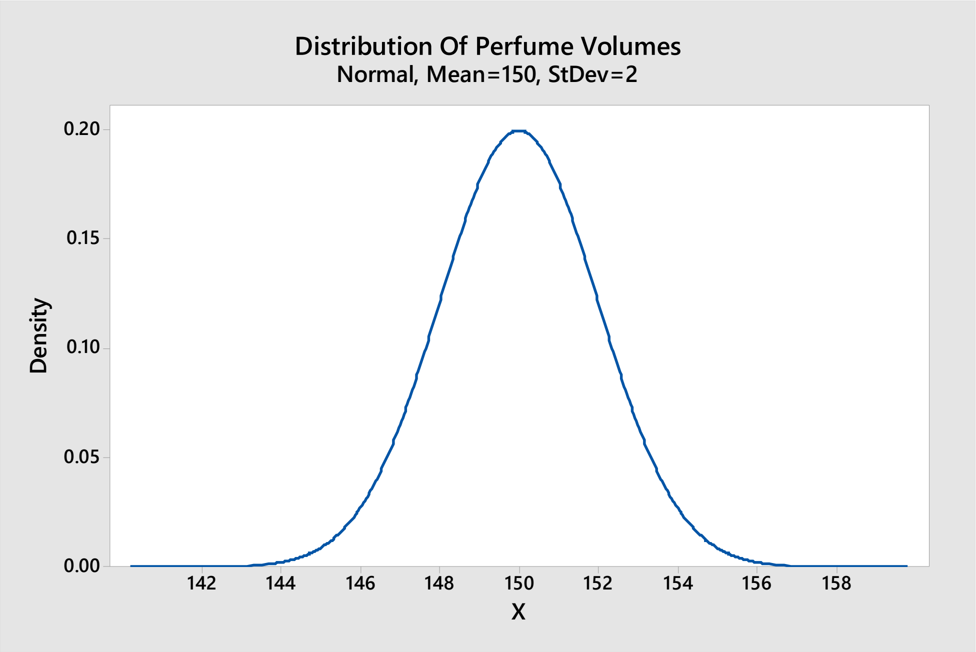
Hypothesis Testing

- Lower Tail Tests
  - $H_0: \mu \geq 150cc$
  - $H_a: \mu < 150cc$
- Upper Tail Tests
  - $H_0: \mu \leq 150cc$
  - $H_a: \mu > 150cc$
- Two Tail Tests
  - $H_0: \mu = 150cc$
  - $H_a: \mu \neq 150cc$

Hypothesis Testing

Hypothesis Testing

- What would you conclude if you pick one sample and find the volume as:
  - 147 cc ….. or
  - 156 cc



Hypothesis Testing

Hypothesis Testing

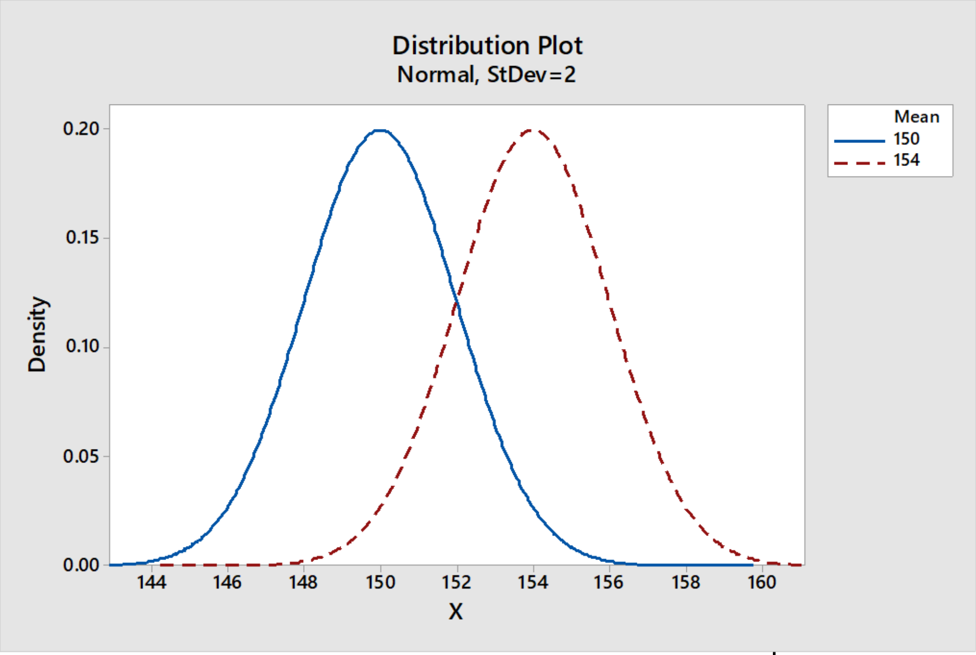
1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing - Errors

Hypothesis Testing - Errors

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

Hypothesis Testing - Errors



$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

Hypothesis Testing - Errors

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

	Type I error (alpha)	Type II error (beta)
Name	Producer’s risk/ **Significance level**	Consumer’s risk
1 minus error is called	**Confidence level**	Power of the test
Example of Fire Alarm	False fire alarm leading to inconvenience	Missed fire leading to disaster
Effects on process	Unnecessary cost increase due to frequent changes	Defects may be produced
Control method	Usually fixed at a pre-determined level, 1%, 5% or 10%	Usually controlled to < 10% by appropriate sample size
Simple definition	Innocent declared as guilty	Guilty declared as innocent

Confidence Level:

C = 0.90, 0.95, 0.99 (90%, 95%, 99%)

Level of Significance or Type I Error:

α = 1 – C (0.10, 0.05, 0.01)

	Type I error (alpha)	Type II error (beta)
Name	Producer's risk/ **Significance level**	Consumer's risk
1 minus error is called	**Confidence level**	Power of the test
Example of Fire Alarm	False fire alarm leading to inconvenience	Missed fire leading to disaster
Effects on process	Unnecessary cost increase due to frequent changes	Defects may be produced
Control method	Usually fixed at a pre-determined level, 1%, 5% or 10%	Usually controlled to < 10% by appropriate sample size
Simple definition	Innocent declared as guilty	Guilty declared as innocent

Hypothesis Testing - Errors

Power

- Power = 1 – β (or 1 - type II error)
- Type II Error: Failing to reject null hypothesis when null hypothesis is false.
- Power: Likelihood of rejecting null hypothesis when null hypothesis is false.
- Or: Power is the ability of a test to correctly reject the null hypothesis.

		True State of Nature	
		H_0 Is true	H_a Is true
Conclusion	Support H_0 / Reject H_a	Correct Conclusion	Type II Error
	Support H_a / Reject H_0	Type I Error	Correct Conclusion (Power)

Hypothesis Testing

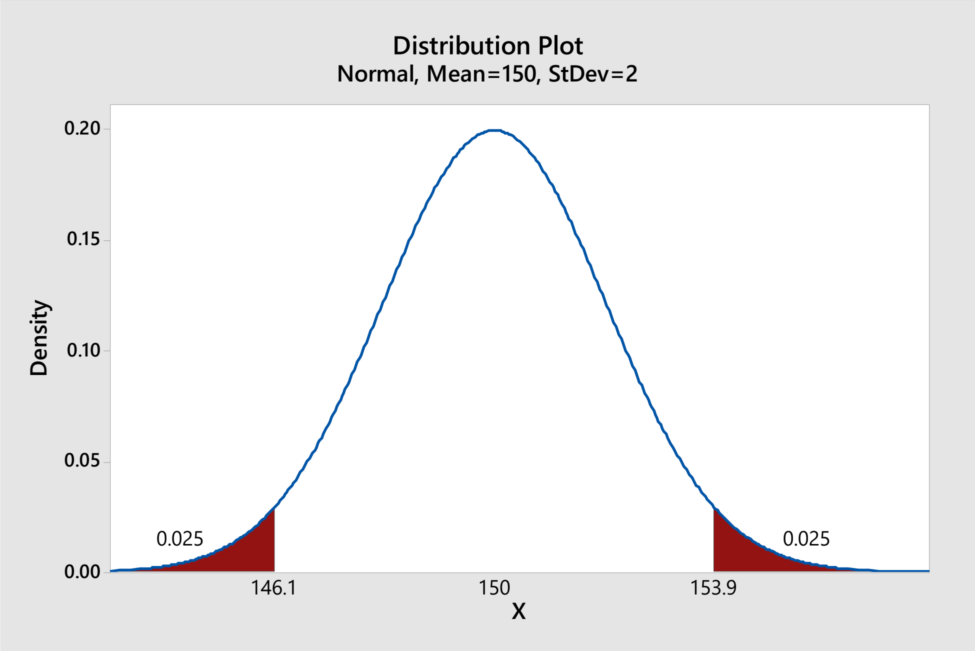
1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing

Hypothesis Testing

- Let's pick 1 bottle from the production line and the volume is 153.8 cc
- With 95% confidence level we will fail to reject the null hypothesis.

H_0: $\mu = 150cc$
H_a: $\mu \neq 150cc$

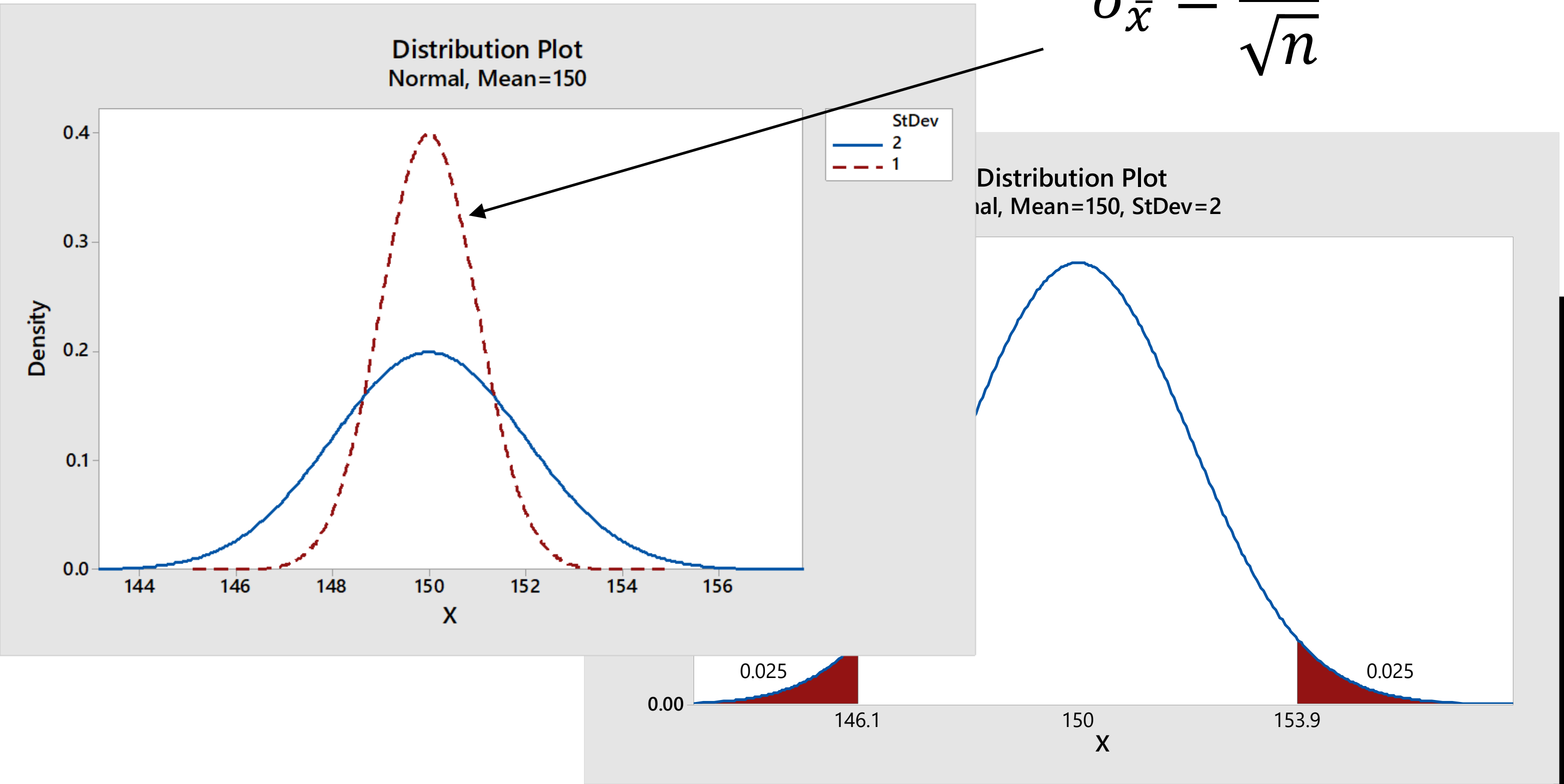


Hypothesis Testing

Hypothesis Testing

- Let’s pick 4 bottles from the production line and find the average volume as 153.8 cc.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

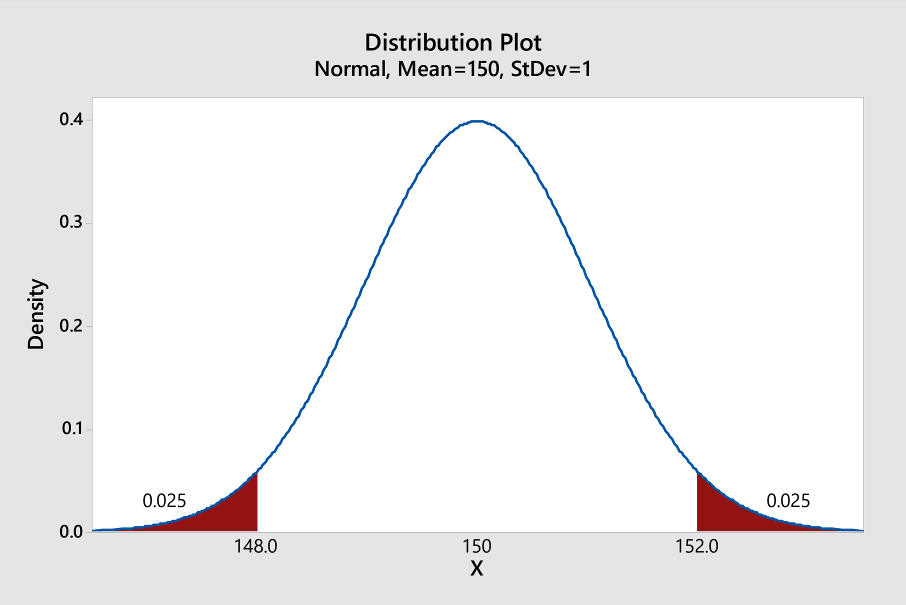


Hypothesis Testing

Hypothesis Testing

- Let’s pick 4 bottles from the production line and find the average volume as 153.8 cc.
- With 95% confidence level we will reject the null hypothesis.

H_0: $\mu = 150cc$
H_a: $\mu \neq 150cc$

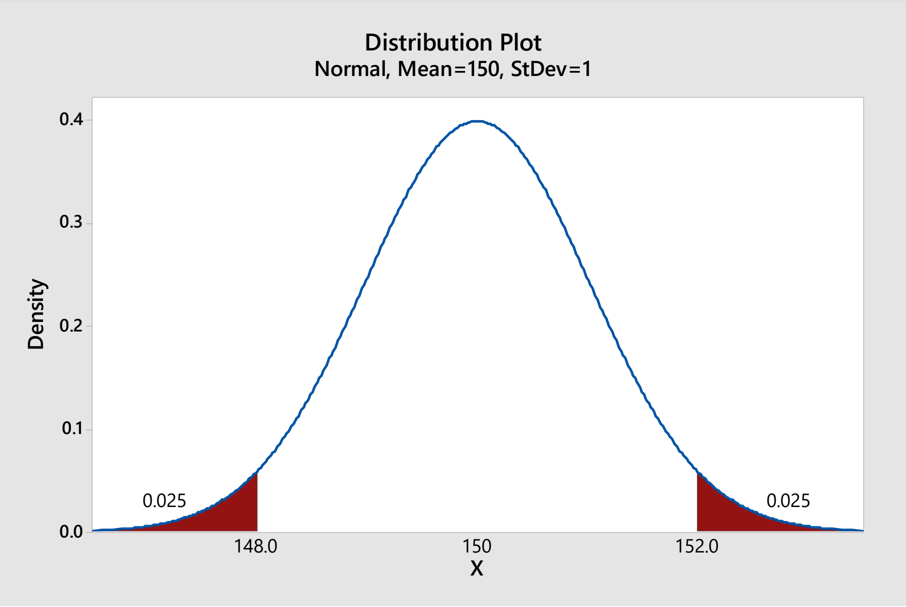


Hypothesis Testing

Hypothesis Testing

$$z_{cal} = \frac{(\bar{X} - \mu)}{\sigma_x/\sqrt{n}} \qquad z_{cal} = \frac{(153.8 - 150)}{2/\sqrt{4}} = 3 \cdot 8$$

For α = 0.05 Two Tails means 0.025 on both tails. Z Critical = 1.96



H_0: μ = 150cc
H_a: μ ≠ 150cc

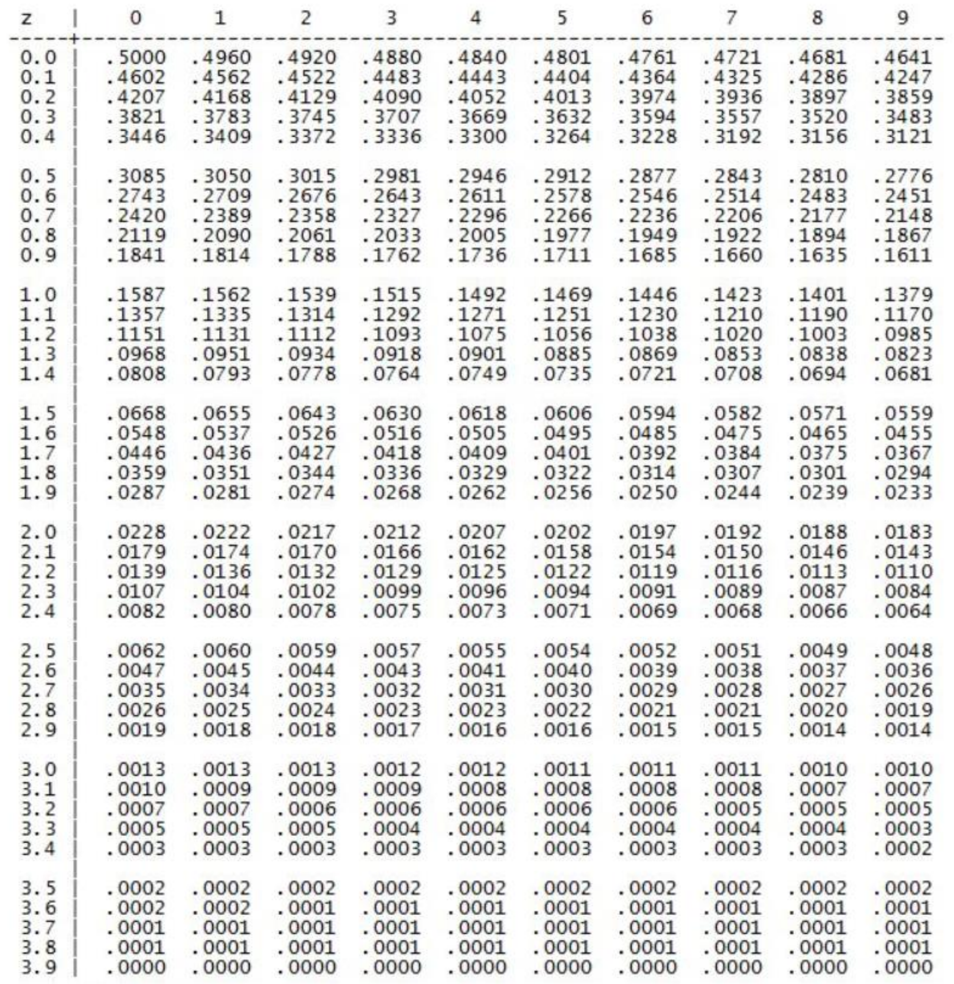
Hypothesis Testing

Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic (e.g t or z score)
5. Critical test statistic
6. Interpret the results.

Hypothesis Testing

Hypothesis Testing



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

- α = 0.01 Two Tails means 0.005 on both tails. Z Critical = 2.575
- α = 0.05 Two Tails means 0.025 on both tails. Z Critical = 1.96
- α = 0.10 Two Tails means 0.05 on both tails. Z Critical = 1.645
- α = 0.05 Single Tails
  - Z Critical = 1.645

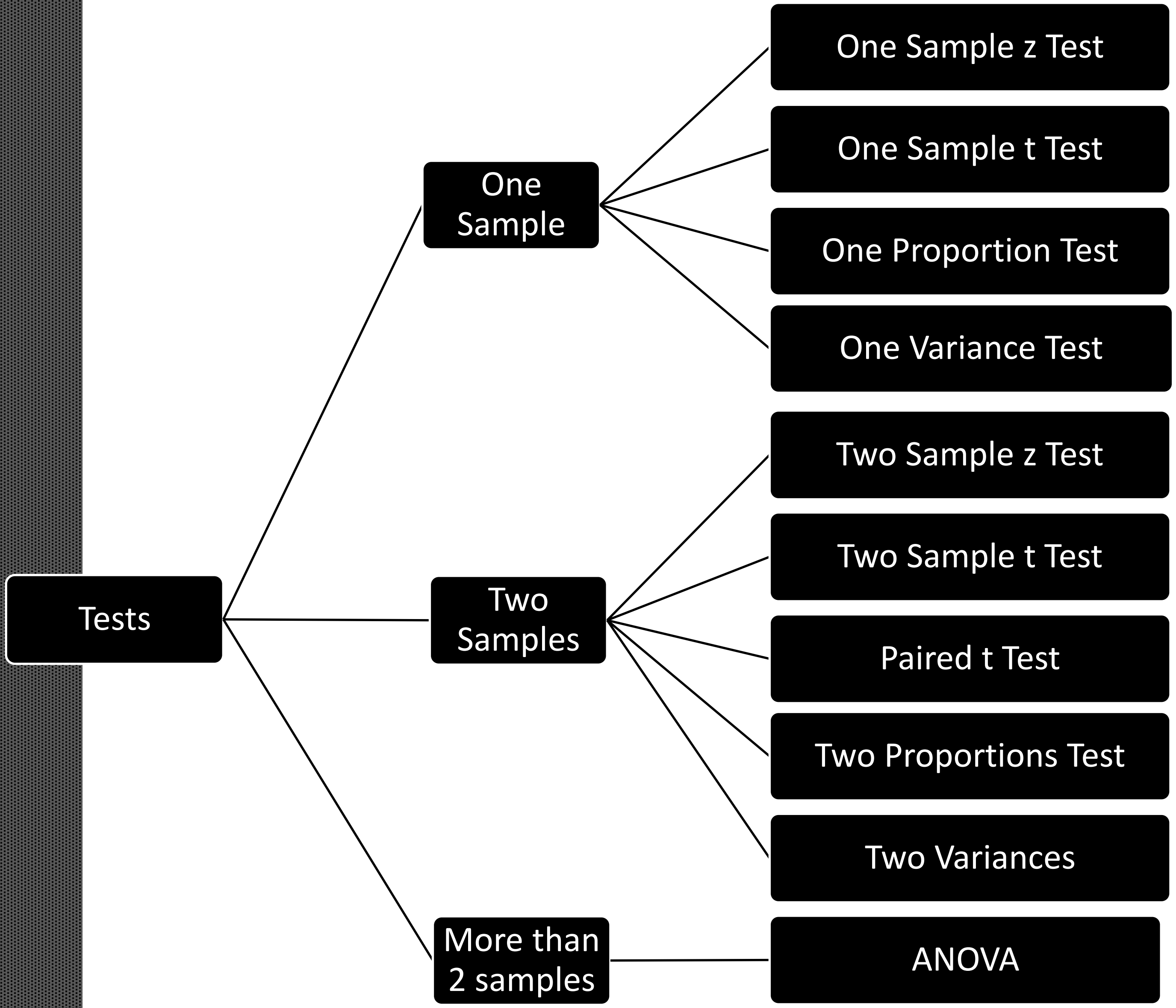
p Value

- p value is the lowest value of alpha for which the null hypothesis can be rejected. (Probability that the null hypothesis is correct)
- For example, if $p = 0.045$ you can reject the null hypothesis at $\alpha = 0.05$
- p is low the null must go / p is high the null fly.

Section 7

Hypothesis Testing Part 1

Tests for means, variances and proportions



Tests

One Sample

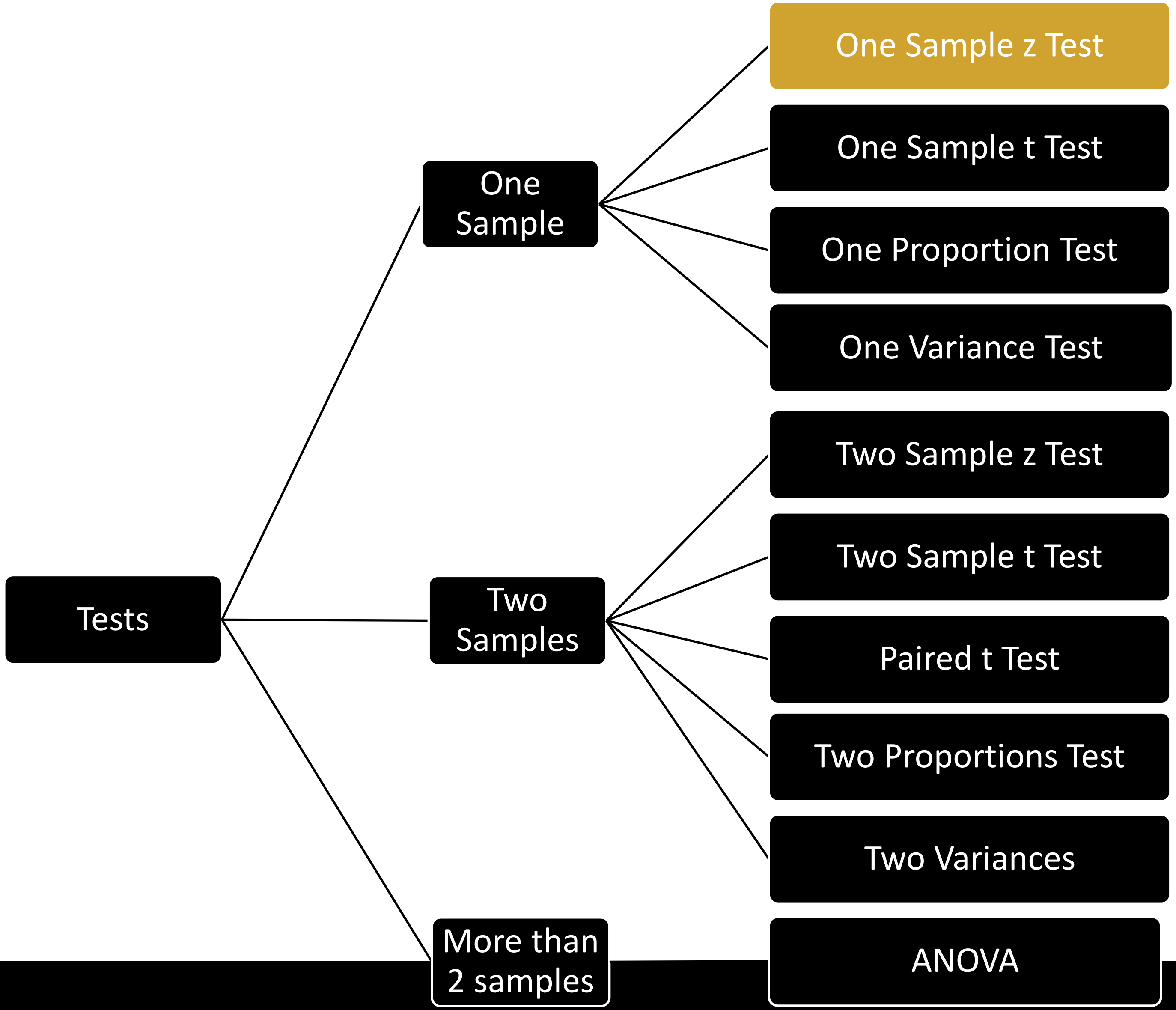
- One Sample z Test
- One Sample t Test
- One Proportion Test
- One Variance Test

Two Samples

- Two Sample z Test
- Two Sample t Test
- Paired t Test
- Two Proportions Test
- Two Variances

More than 2 samples

- ANOVA



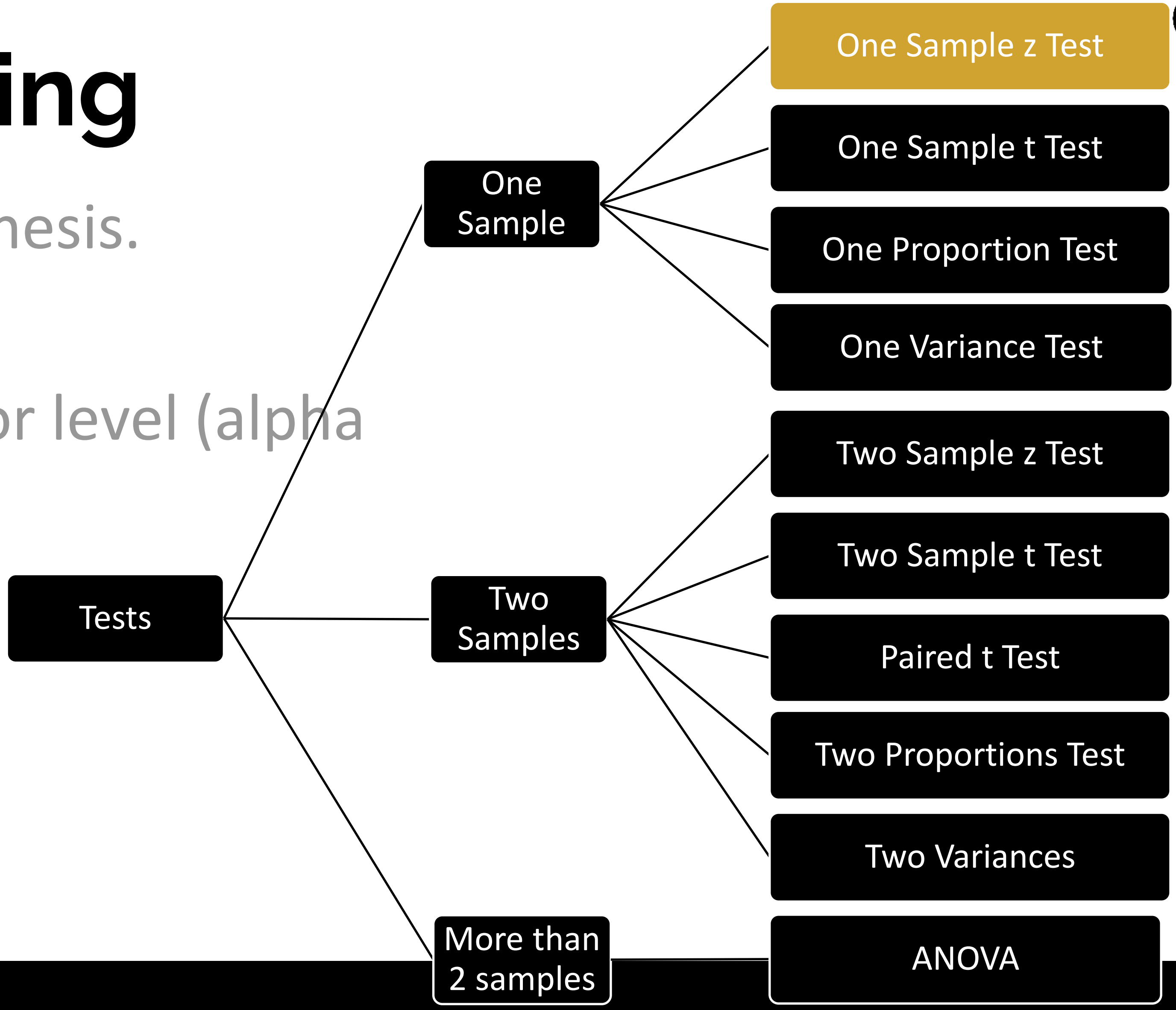
- Tests
  - One Sample
    - One Sample z Test
    - One Sample t Test
    - One Proportion Test
    - One Variance Test
  - Two Samples
    - Two Sample z Test
    - Two Sample t Test
    - Paired t Test
    - Two Proportions Test
    - Two Variances
  - More than 2 samples
    - ANOVA

One Sample z Test

Hypothesis Testing

1. State the Alternate Hypothesis.
2. State the Null Hypothesis.
3. Select a probability of error level (alpha level). Generally 0.05
4. Calculate the test statistic

   (e.g t or z score)
5. Critical test statistic
6. Interpret the results.



- Tests
  - One Sample
    - One Sample z Test
    - One Sample t Test
    - One Proportion Test
    - One Variance Test
  - Two Samples
    - Two Sample z Test
    - Two Sample t Test
    - Paired t Test
    - Two Proportions Test
    - Two Variances
  - More than 2 samples
    - ANOVA

One Sample z Test

Conditions for z Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the population standard deviation is known *** OR ***
  - Sample size ≥ 30

One Sample z Test

Calculated Test Statistic

$H_0: \mu = 150cc$
$H_a: \mu \neq 150cc$

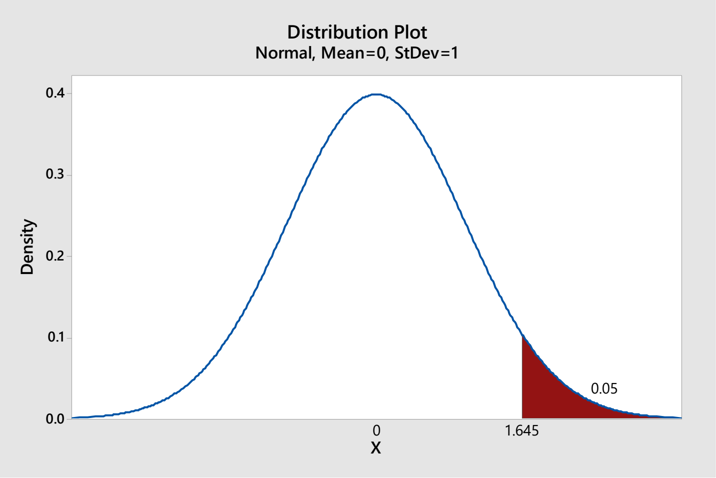
$$z_{cal} = \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}}$$

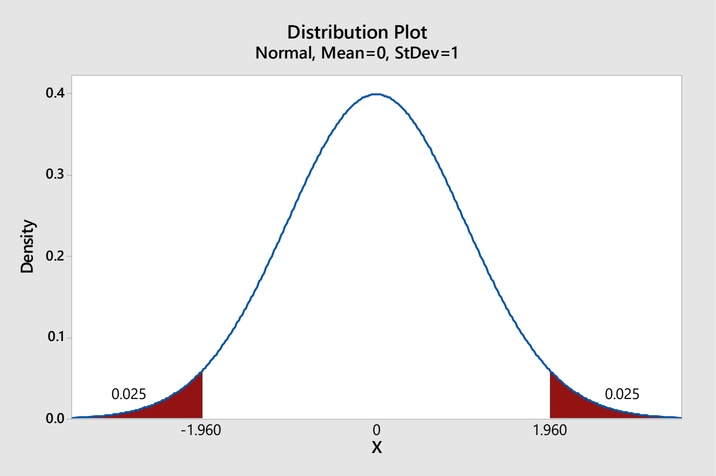
- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume changed? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = ?

One Sample z Test

Critical Test Statistic

z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- ❖ α = 0.05 One Tail
- ❖ Z Critical = 1.645

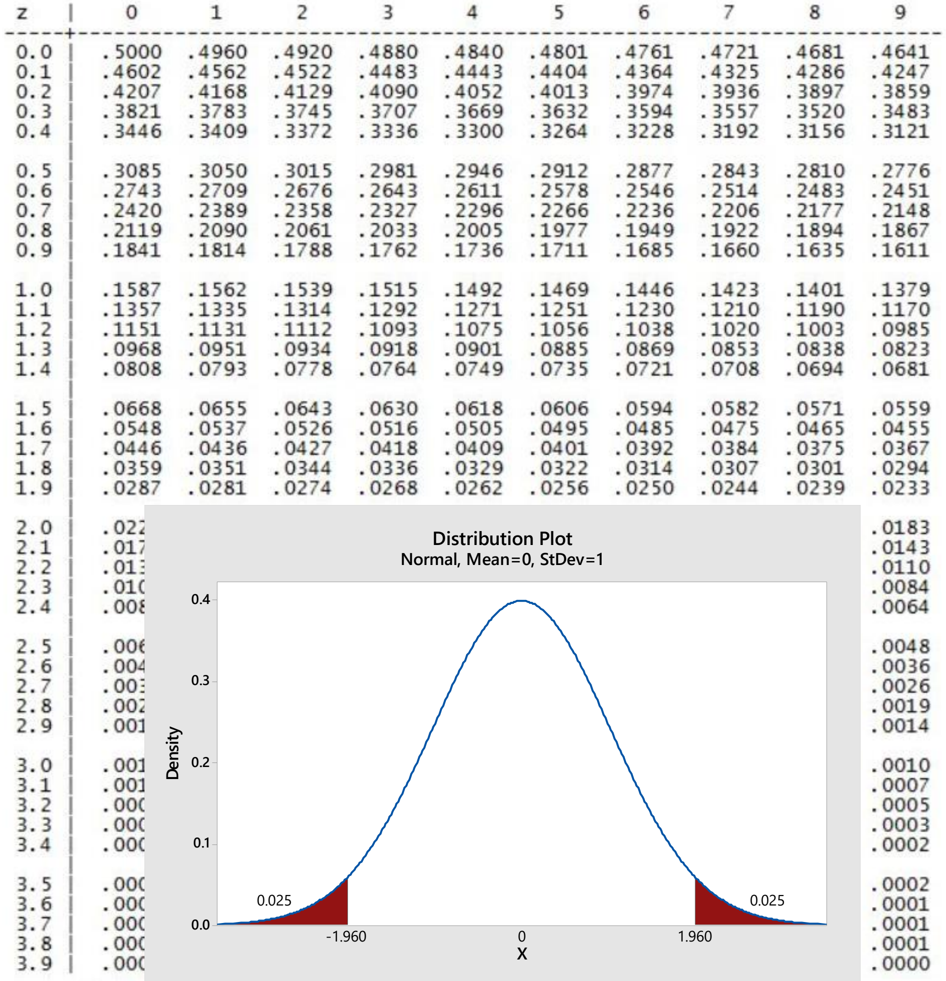
- ❖ α = 0.10 One Tail
- ❖ Z Critical = 1.282

- ❖ α = 0.05 Two Tails
- ❖ Z Critical = 1.96

- ❖ α = 0.10 Two Tail
- ❖ Z Critical = 1.645

One Sample z Test

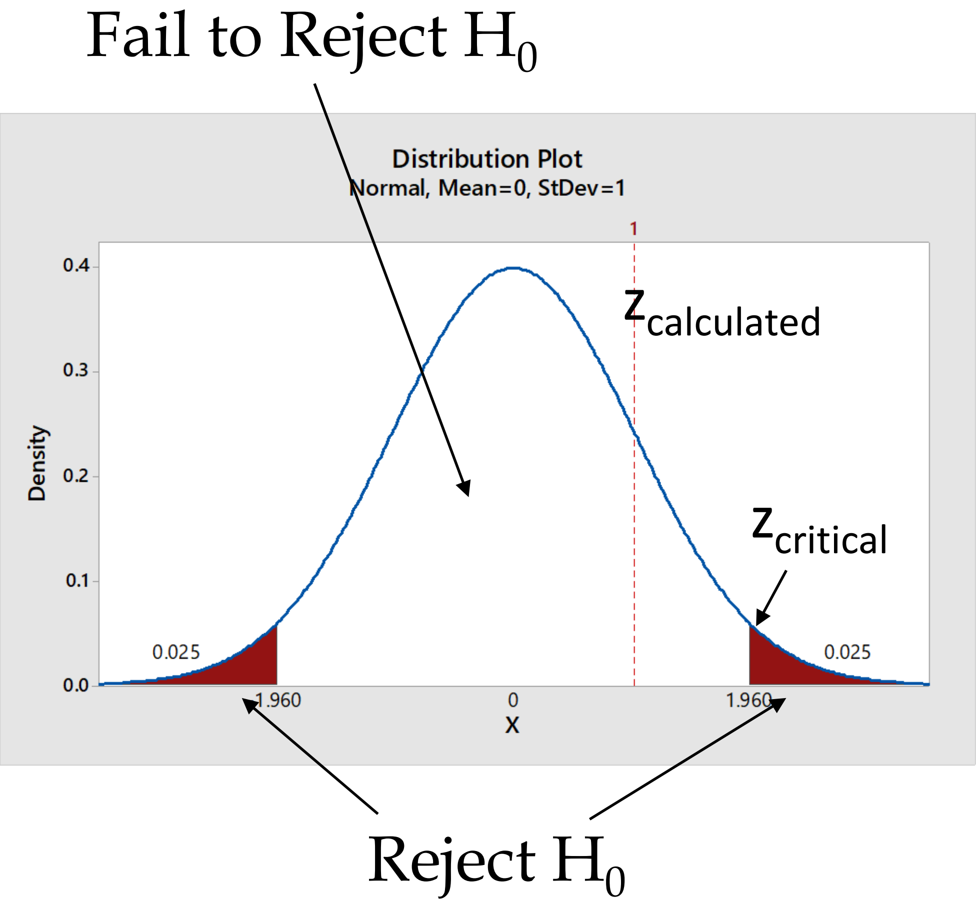
Critical Test Statistic



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.022									.0183
2.1	.017									.0143
2.2	.01									.0110
2.3	.010									.0084
2.4	.008									.0064
2.5	.006									.0048
2.6	.004									.0036
2.7	.003									.0026
2.8	.002									.0019
2.9	.001									.0014
3.0	.001									.0010
3.1	.001									.0007
3.2	.000									.0005
3.3	.000									.0003
3.4	.000									.0002
3.5	.000									.0002
3.6	.000									.0001
3.7	.000									.0001
3.8	.000									.0001
3.9	.000									.0000

- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume changed? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = 1.96

Interpret the Results



Fail to Reject H_0

Reject H_0

- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume changed? (95% confidence)
- $z_{calculated} = (150.2-150)/[2 / \text{sqrt}(100)] = 0.2/0.2 = 1$
- $z_{critical} = 1.96$

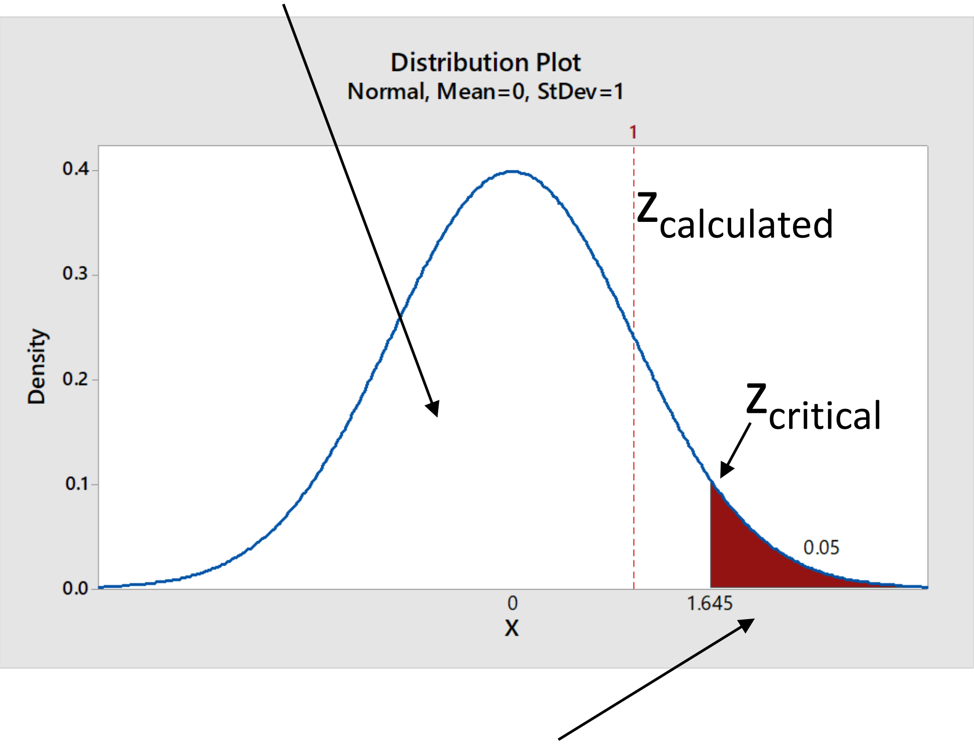
Interpret the Results

One Tail Tests

$H_0: \mu \leq 150cc$

$H_a: \mu > 150cc$

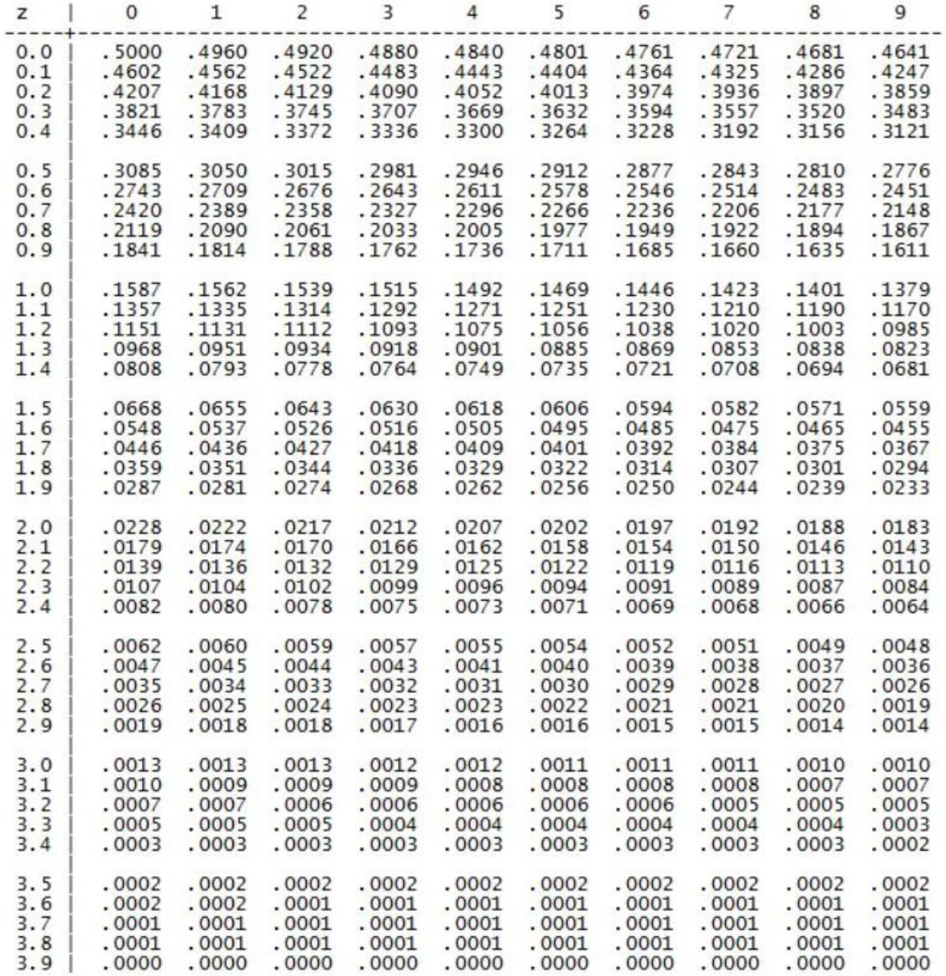
Fail to Reject H_0



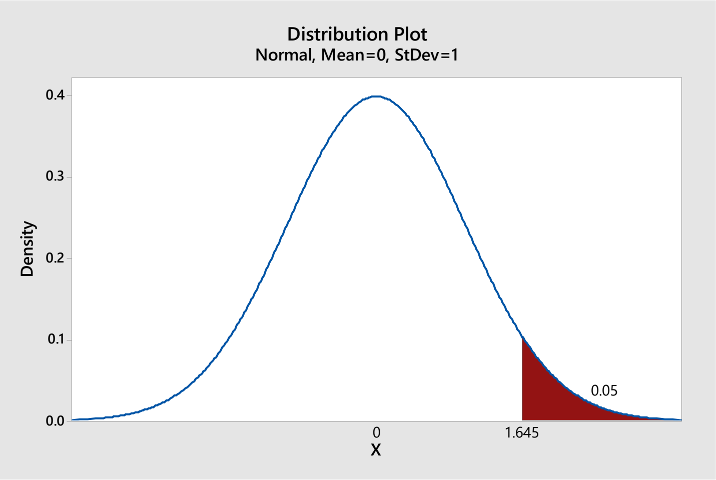
Reject H_0

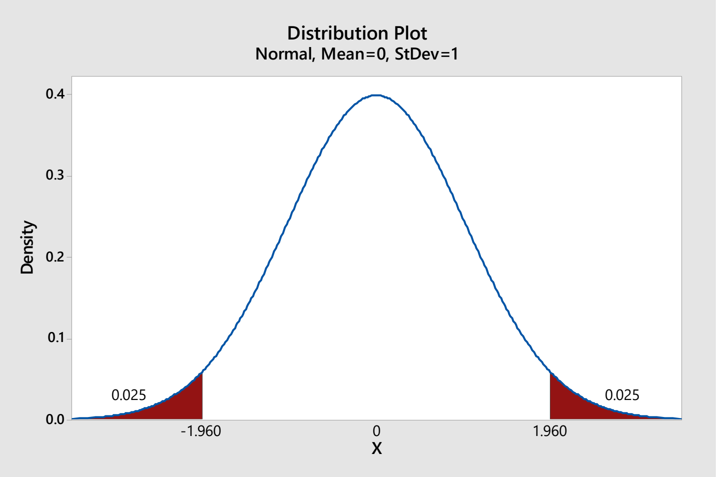
- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume ~~changed~~ increased? (95% confidence)
- $z_{calculated} = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1$
- $z_{critical}$ = ~~1.96~~ 1.645

Critical Test Statistic



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

- α = 0.10 One Tail
- Z Critical = 1.282

- α = 0.05 Two Tails
- Z Critical = 1.96

- α = 0.10 Two Tail
- Z Critical = 1.645

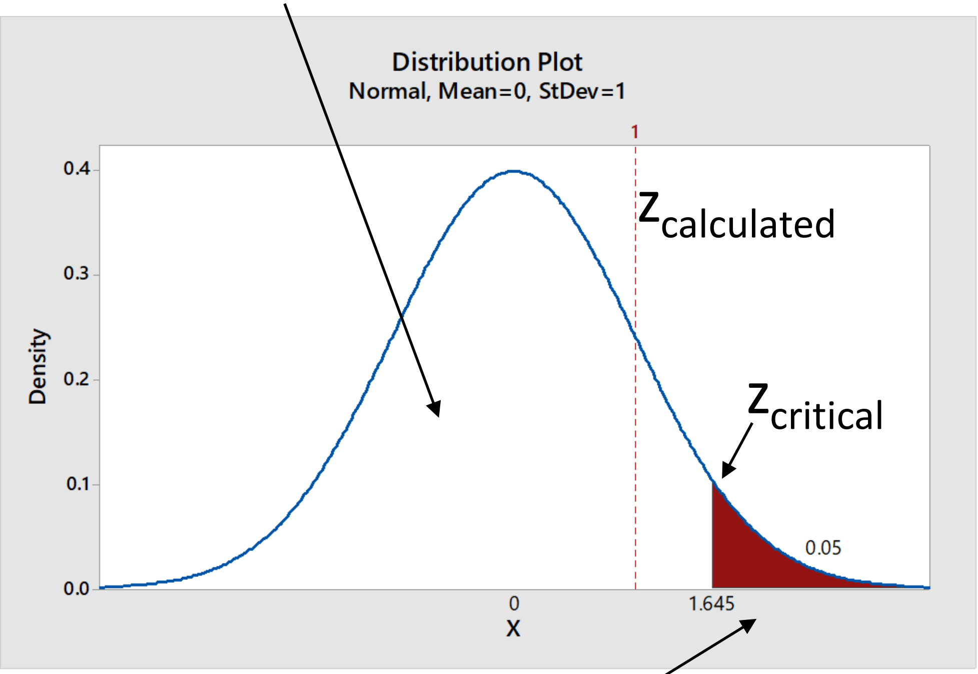
Interpret the Results

One Tail Tests

H_0: $\mu \leq 150$cc

H_a: $\mu > 150$cc

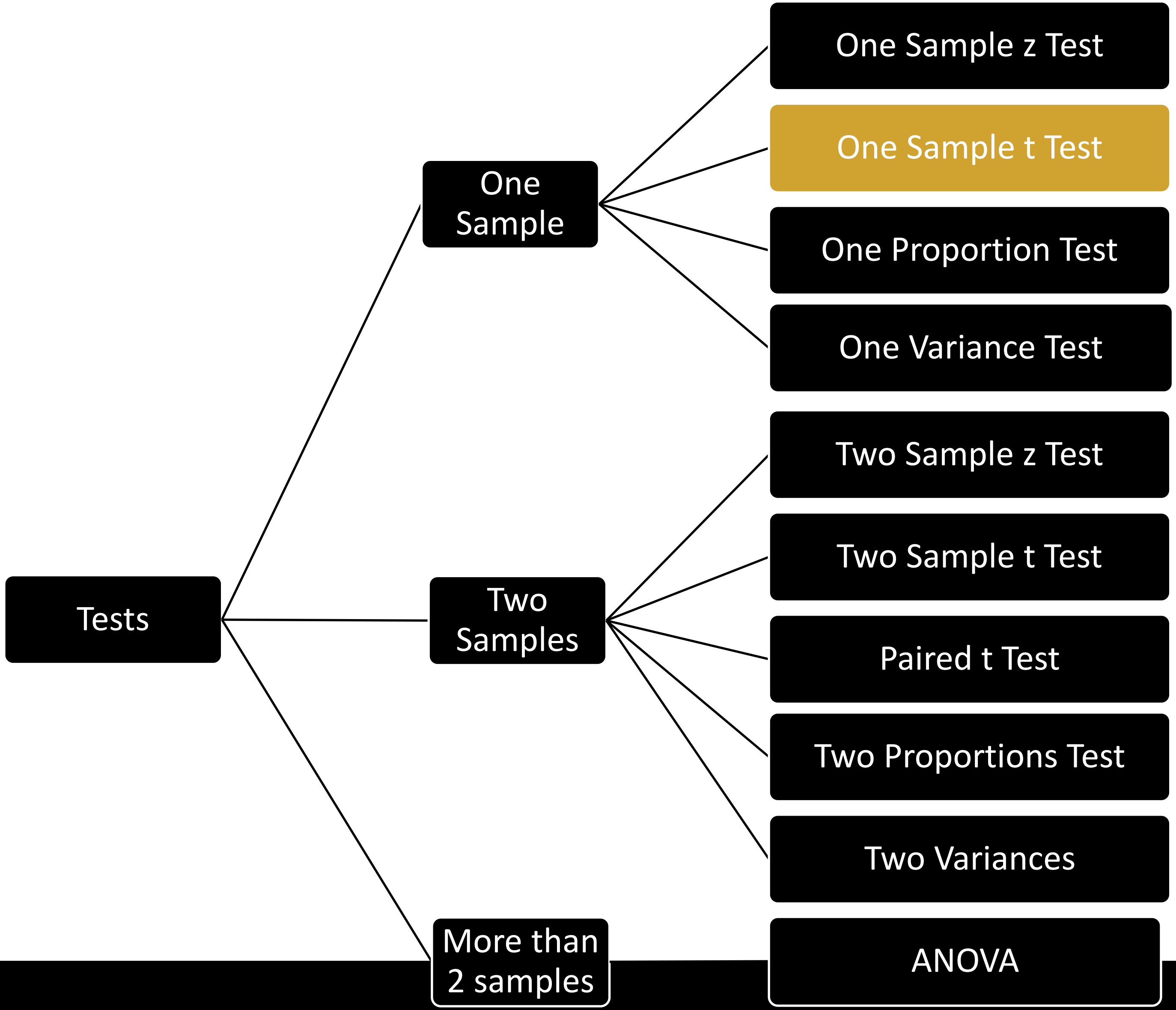
Fail to Reject H_0



Reject H_0

- Example: Perfume bottle producing 150cc with sd of 2 cc, 100 bottles are randomly picked and the average volume was found to be 150.2 cc. Has mean volume ~~changed~~ increased? (95% confidence)
- $z_{calculated}$ = (150.2-150)/[2 / sqrt(100)] = 0.2/0.2 = 1
- $z_{critical}$ = ~~1.96~~ 1.645

One Sample z Test



Tests

One Sample

Two Samples

More than 2 samples

One Sample z Test

One Sample t Test

One Proportion Test

One Variance Test

Two Sample z Test

Two Sample t Test

Paired t Test

Two Proportions Test

Two Variances

ANOVA

One Sample t Test

Conditions for t Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the standard deviation is **unknown** *** AND ***
  - Sample size < 30

One Sample t Test

Conditions for t Test

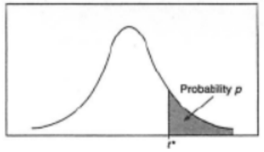
H_0: μ = 150cc
H_a: μ ≠ 150cc

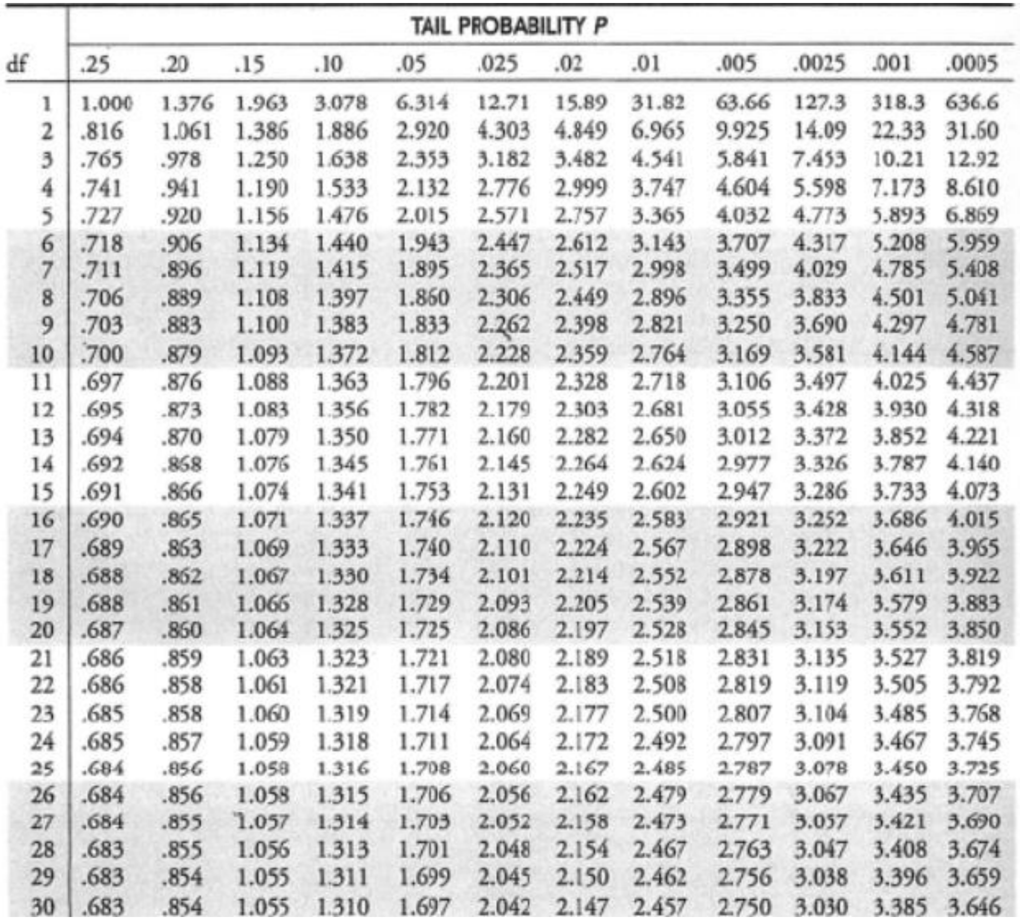
$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

- Example: Perfume bottle producing 150cc, 4 bottles are randomly picked and the average volume was found to be 151cc and sd of the sample bottles was 2 cc. Has mean volume changed? (95% confidence)
- $t_{calculated} = \frac{(\bar{x}-\mu)}{s/\sqrt{n}} = \frac{(151-150)}{2/\sqrt{4}} = \frac{1}{1} = 1$
- $t_{critical}$ = ?

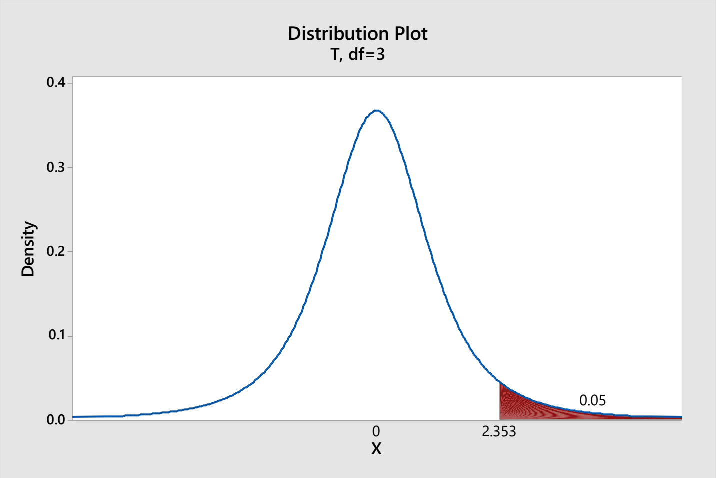
One Sample t Test

Critical Test Statistic

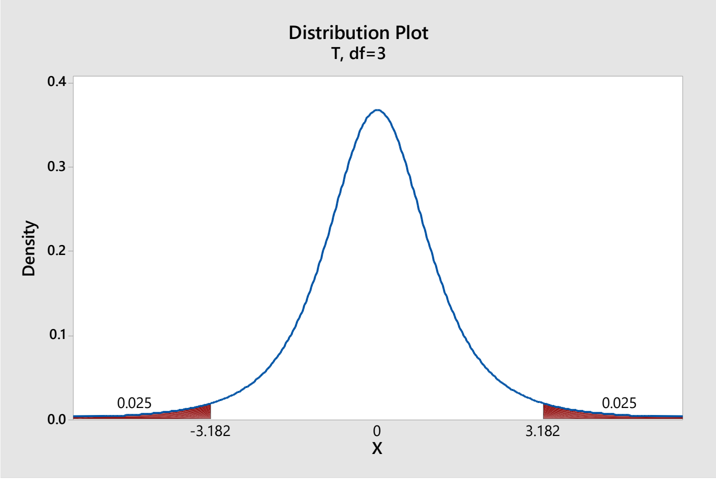




df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



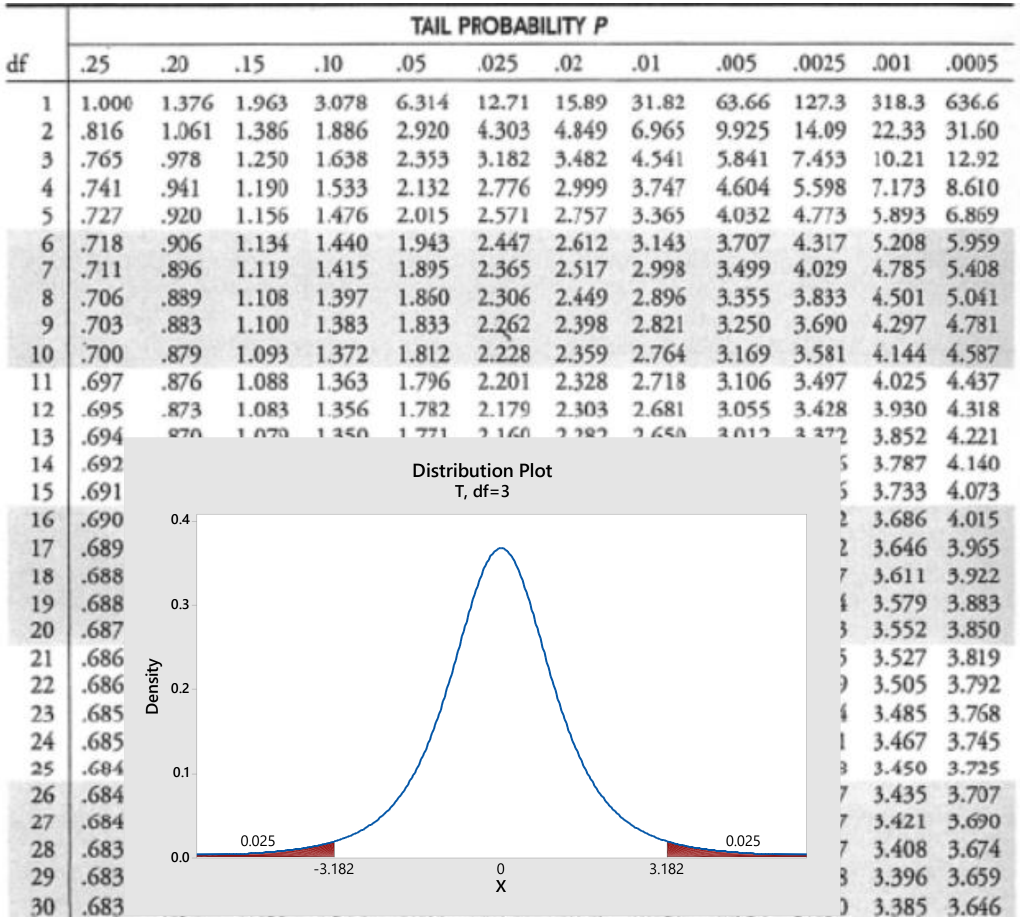
- α = 0.05 One Tails
- Df = 3
- t Critical = 2.353



- α = 0.05 Two Tails
- Df = 3
- t Critical = 3.182

One Sample t Test

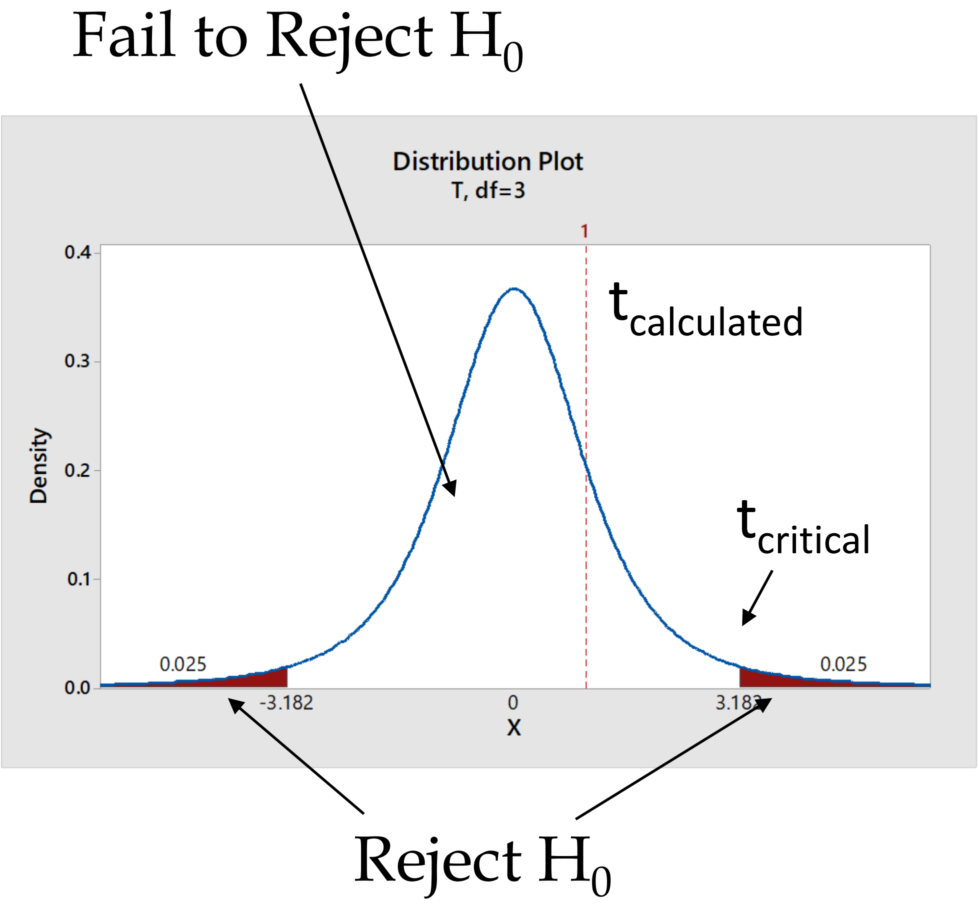
Critical Test Statistic



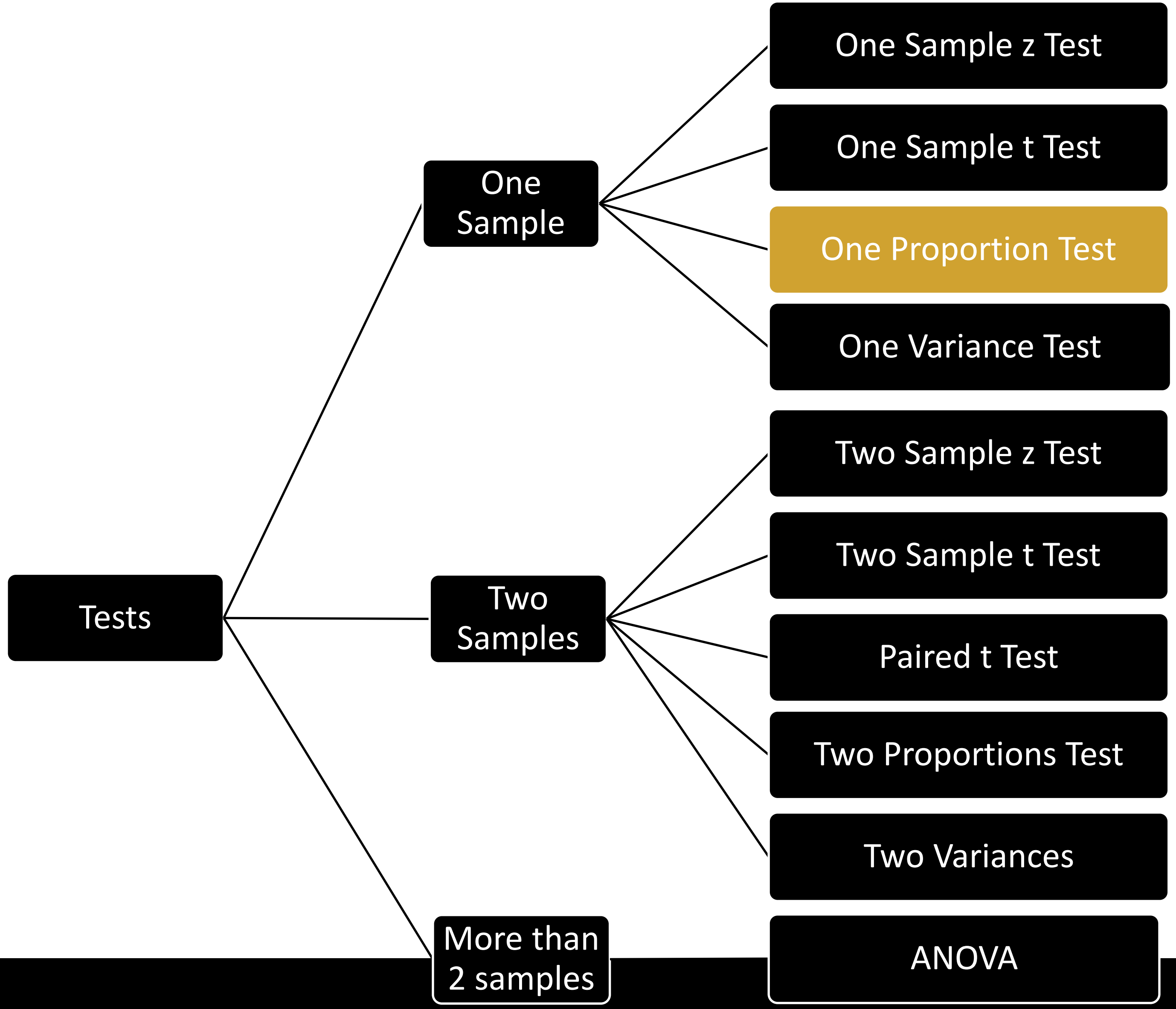
	TAIL PROBABILITY *P*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692										3.787	4.140
15	.691										3.733	4.073
16	.690										3.686	4.015
17	.689										3.646	3.965
18	.688										3.611	3.922
19	.688										3.579	3.883
20	.687										3.552	3.850
21	.686										3.527	3.819
22	.686										3.505	3.792
23	.685										3.485	3.768
24	.685										3.467	3.745
25	.684										3.450	3.725
26	.684										3.435	3.707
27	.684										3.421	3.690
28	.683										3.408	3.674
29	.683										3.396	3.659
30	.683										3.385	3.646

- Example: Perfume bottle producing 150cc, 4 bottles are randomly picked and the average volume was found to be 151cc and sd of the sample bottles was 2 cc. Has mean volume changed? (95% confidence)
- $t_{\text{calculated}} = \dfrac{(\bar{x}-\mu)}{s/\sqrt{n}} = \dfrac{(151-150)}{2/\sqrt{4}} = \dfrac{1}{1} = 1$
- $t_{\text{critical}} = 3.182$

Interpret the Results



- ❖ Example: Perfume bottle producing 150cc, 4 bottles are randomly picked and the average volume was found to be 151cc and sd of the sample bottles was 2 cc. Has mean volume changed? (95% confidence)
- ❖ $t_{calculated} = \frac{(\bar{x}-\mu)}{s/\sqrt{n}} = \frac{(151-150)}{2/\sqrt{4}} = \frac{1}{1} = 1$
- ❖ $t_{critical} = 3.182$



Tests

One Sample

One Sample z Test

One Sample t Test

One Proportion Test

One Variance Test

Two Samples

Two Sample z Test

Two Sample t Test

Paired t Test

Two Proportions Test

Two Variances

More than 2 samples

ANOVA

One Proportion Test

Conditions for One Proportion Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data contains only two categories, such as pass/fail or yes/no
- For Normal approximation:
  - both $np \geq 10$ and $n(1-p) \geq 10$ (data should have at least 10 "successes" and at least 10 "failures") (in some books it is 5)

One Proportion Test

H_0: $p = p_0$

H_a: $p \neq p_0$

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed?
- Can Normality assumption be made?
  - $p_0 = 0.21$, $p=0.14$
  - $np_0 = 0.21x100 = 21$
  - $n(1-p_0)= 0.79x100 = 79$
  - >10 means sample size is sufficient.

One Proportion Test

One Proportion Test

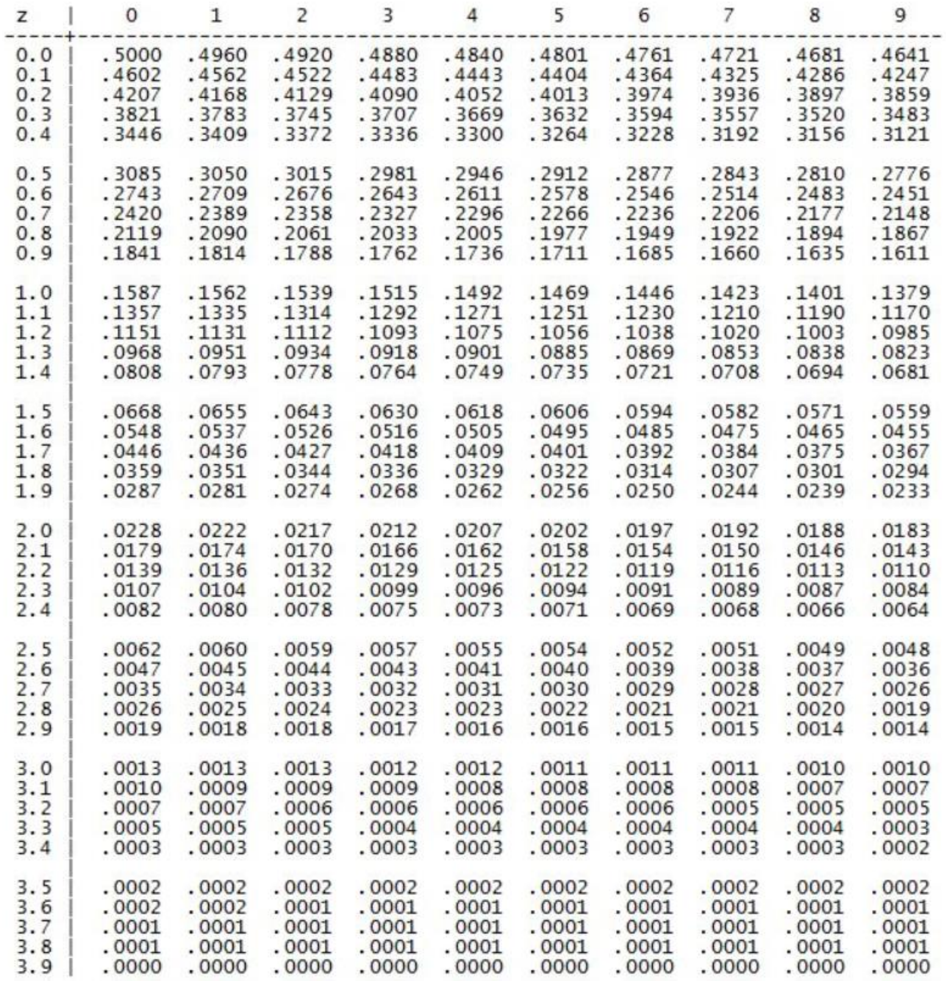
H_0: $p = p_0$
H_a: $p \neq p_0$

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

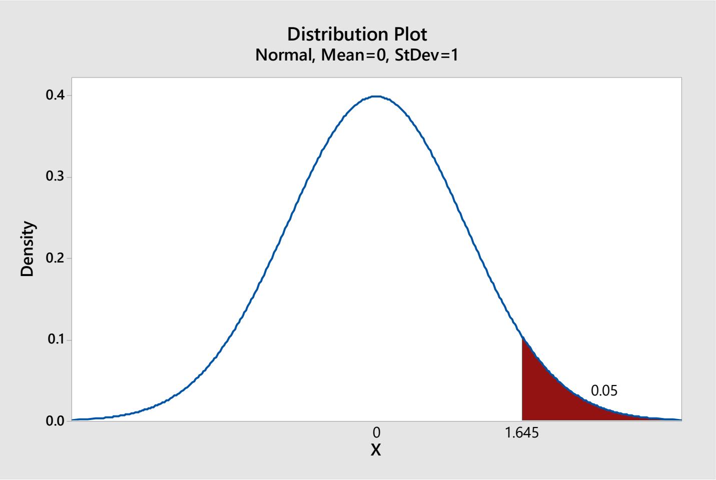
- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed?
- $z_{calculated} = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\mathbf{0.14 - 0.21}}{\sqrt{\frac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{critical} = ?$

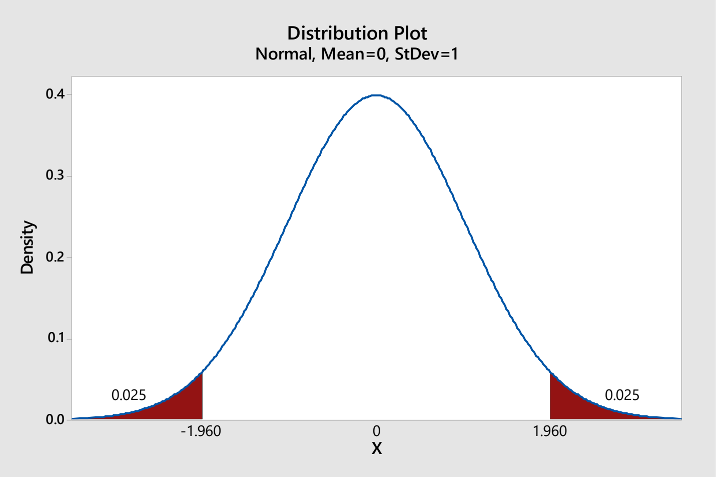
One Proportion Test

Critical Test Statistic



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

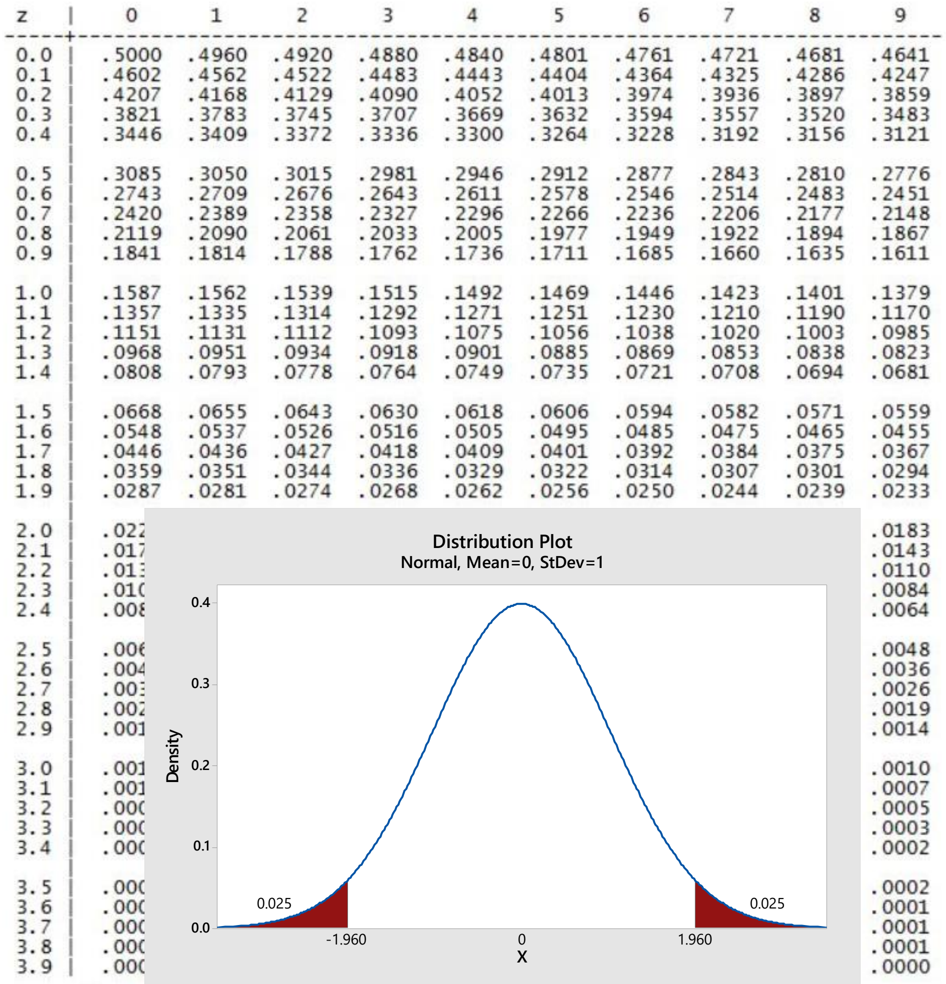
- α = 0.10 One Tail
- Z Critical = 1.282

- α = 0.05 Two Tails
- Z Critical = 1.96

- α = 0.10 Two Tail
- Z Critical = 1.648

One Proportion Test

One Proportion Test



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.02									.0183
2.1	.01									.0143
2.2	.01									.0110
2.3	.01									.0084
2.4	.00									.0064
2.5	.00									.0048
2.6	.00									.0036
2.7	.00									.0026
2.8	.00									.0019
2.9	.00									.0014
3.0	.00									.0010
3.1	.00									.0007
3.2	.00									.0005
3.3	.00									.0003
3.4	.00									.0002
3.5	.00									.0002
3.6	.00									.0001
3.7	.00									.0001
3.8	.00									.0001
3.9	.00									.0000

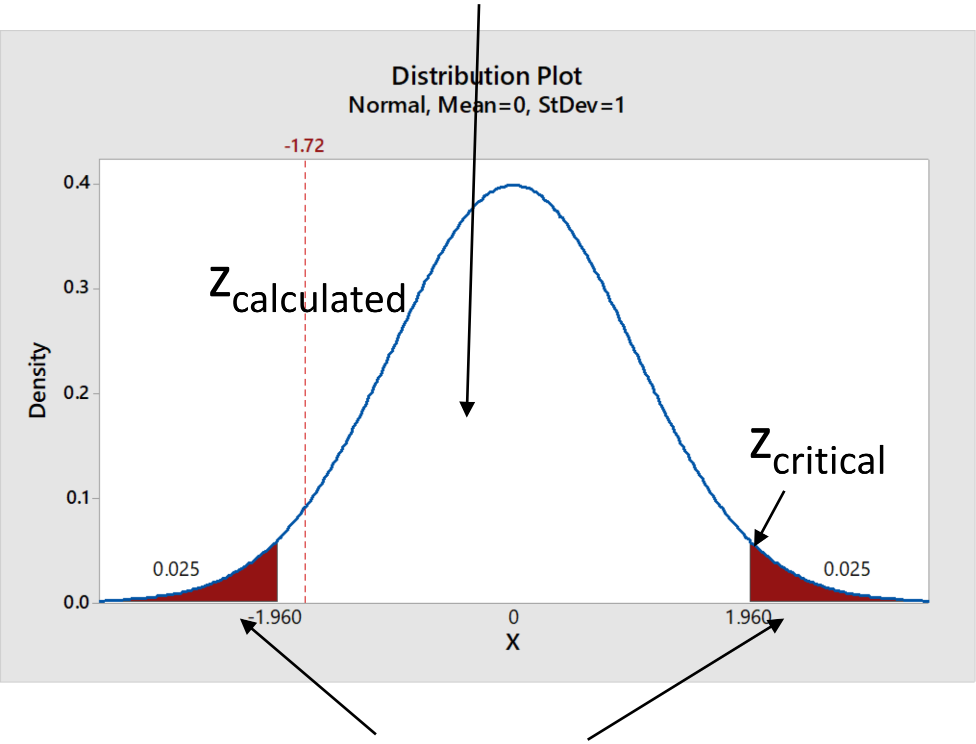
- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed? (95% confidence)
- $z_{\text{calculated}} = \dfrac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{\mathbf{0.14 - 0.21}}{\sqrt{\dfrac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{\text{critical}} = 1.96$

One Proportion Test

Interpret the Results

H_0: $p = p_0$
H_a: $p \neq p_0$

Fail to Reject H_0



Reject H_0

- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit changed? (95% confidence)
- $z_{calculated} = \dfrac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{\mathbf{0.14 - 0.21}}{\sqrt{\dfrac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}} = -1.719$
- $z_{critical} = 1.96$

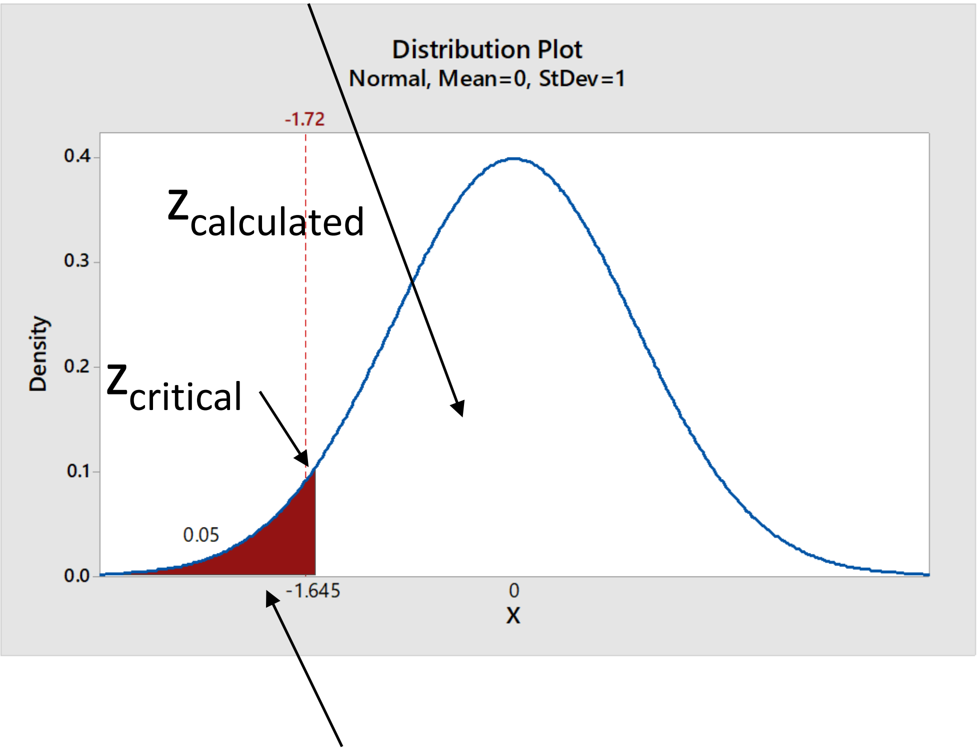
Interpret the Results

One Tail Tests

H_0: $p \geq p_0$

H_a: $p < p_0$

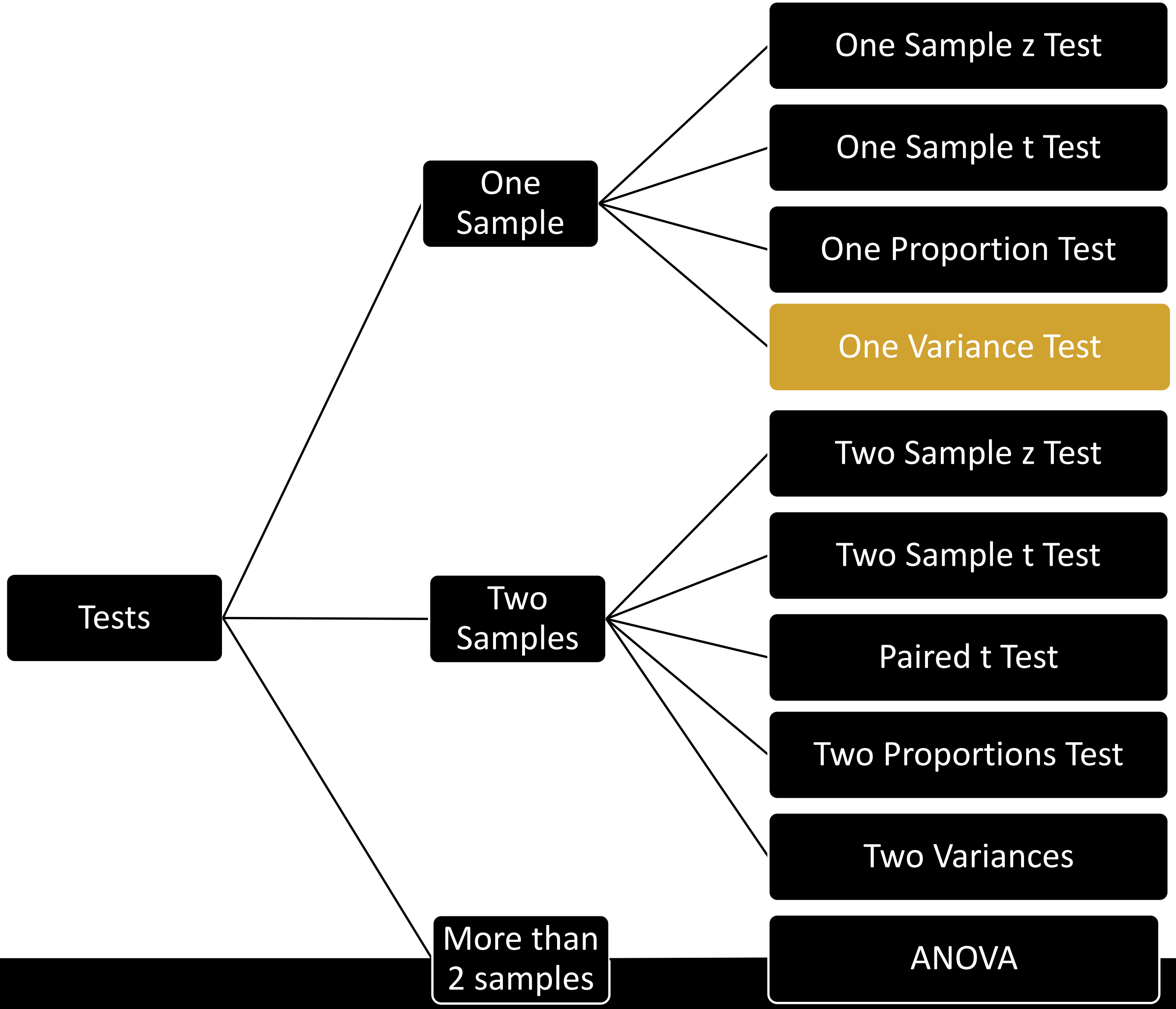
Fail to Reject H_0



Reject H_0

- Example: Smoking rate in a town in past was 21%, 100 samples were picked and found 14 smokers. Has smoking habit reduced at 95% confidence? (one tail)
- $z_{calculated} = \frac{p-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\mathbf{0.14-0.21}}{\sqrt{\frac{\mathbf{0.21(1-0.21)}}{\mathbf{100}}}}$ = -1.719
- $z_{critical}$ = ~~1.96~~ 1.645

One Proportion Test



Tests

One Sample

- One Sample z Test
- One Sample t Test
- One Proportion Test
- **One Variance Test**

Two Samples

- Two Sample z Test
- Two Sample t Test
- Paired t Test
- Two Proportions Test
- Two Variances

More than 2 samples

- ANOVA

One Variance Test

Conditions for One Variance Test

- Random samples
- Each observation should be independent of other
    - Sampling with replacement
    - If sampling without replacement, the sample size should not be more than 10% of the population
- The data follows a Normal Distribution

Variance Tests

- Chi-square test
    - For testing the population variance against a specified value
    - testing goodness of fit of some probability distribution
    - testing for independence of two attributes (Contingency Tables)

- F-test
    - for testing equality of *two* variances from different populations
    - for testing equality of several means with technique of ANOVA.

One Variance Test

One Variance Test

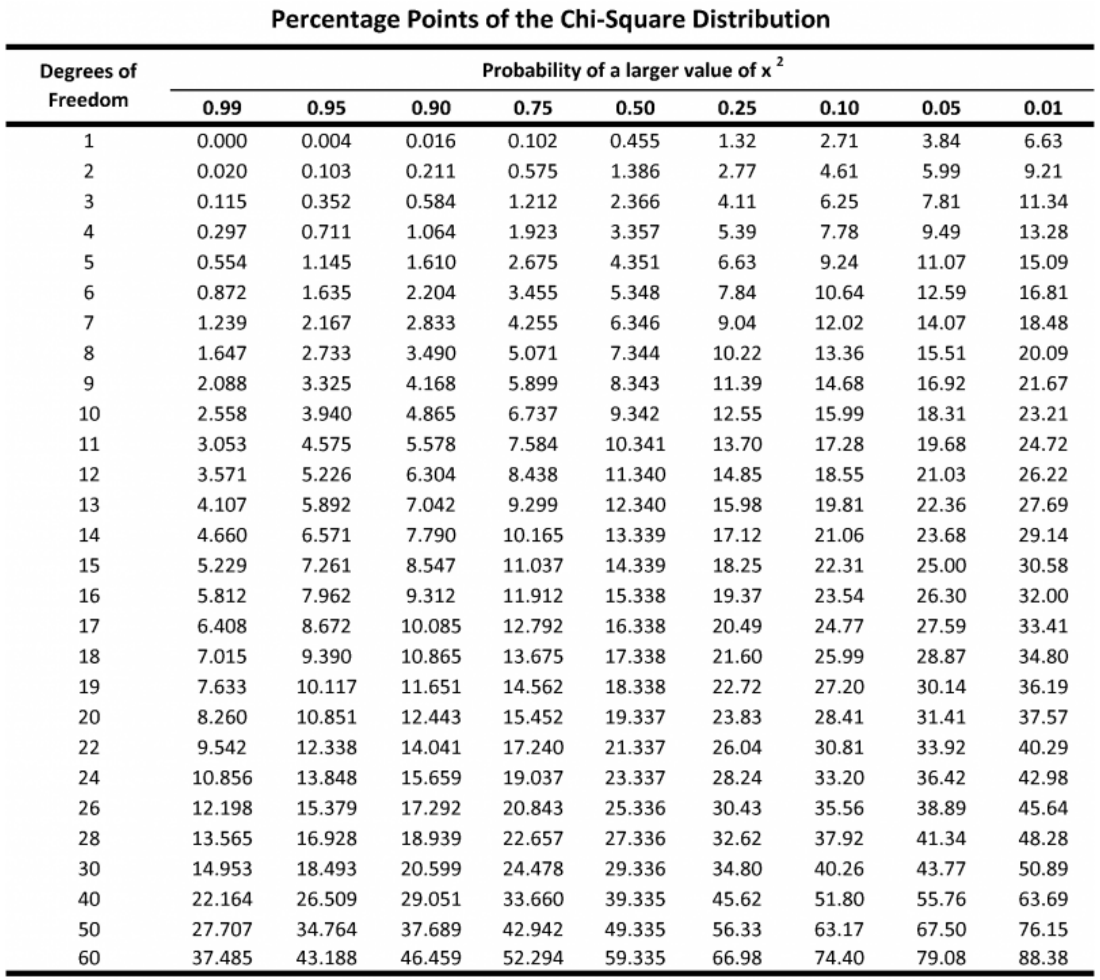
Ho: $s^2 <= \sigma^2$
Ha: $s^2 > \sigma^2$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

- ❖ Example 1: A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **increased** from established 2 cc? 90% confidence level.
- ❖ $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{\mathbf{(50)}\ \boldsymbol{x}\ \mathbf{2.35^2}}{\mathbf{2^2}} = 69.03$
- ❖ What is critical value of Chi Square for 50 degrees of freedom?

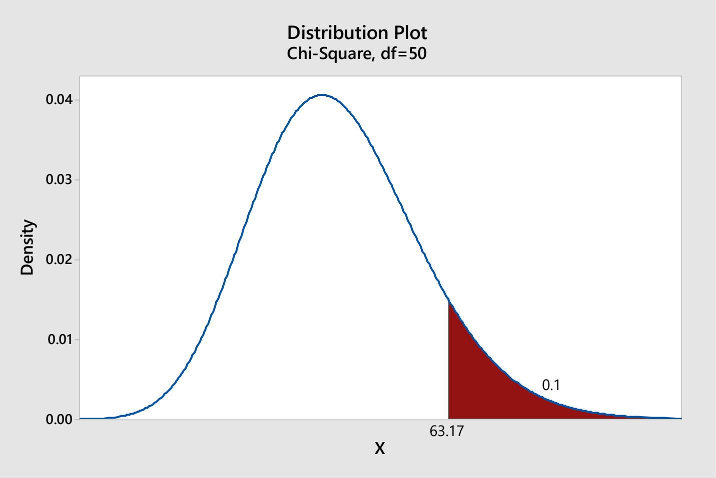
One Variance Test

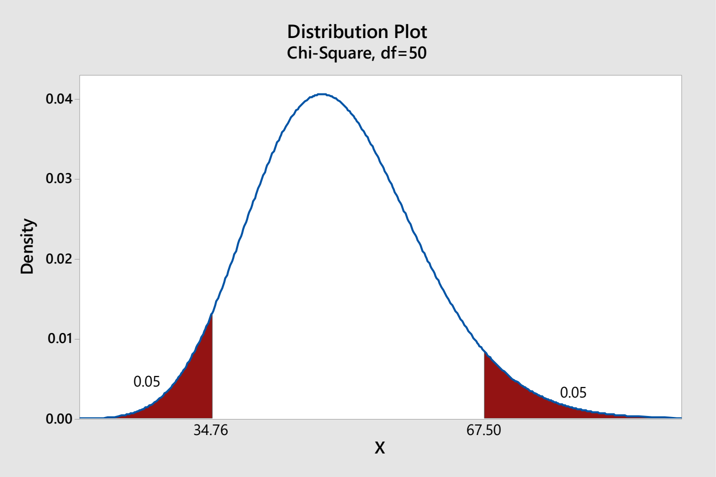
Critical Test Statistic



Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x ²								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



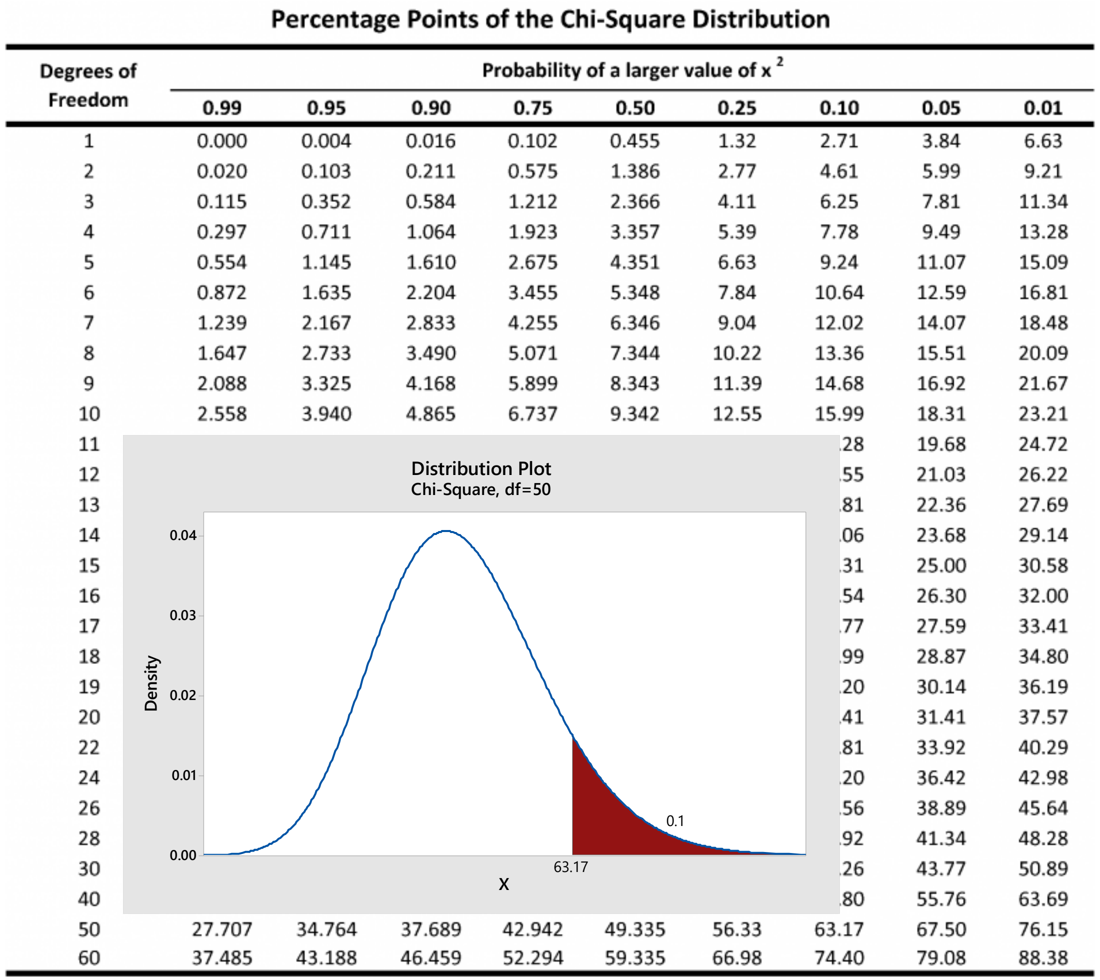


- α = 0.10 One Tail
- Df = 50
- χ^2 Critical = 63.17

- α = 0.10 Two Tail
- Df = 50
- χ^2Critical = 34.76 and 67.50

One Variance Test

Critical Test Statistic



Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x ^{2}								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11							.28	19.68	24.72
12							.55	21.03	26.22
13							.81	22.36	27.69
14							.06	23.68	29.14
15							.31	25.00	30.58
16							.54	26.30	32.00
17							.77	27.59	33.41
18							.99	28.87	34.80
19							.20	30.14	36.19
20							.41	31.41	37.57
22							.81	33.92	40.29
24							.20	36.42	42.98
26							.56	38.89	45.64
28							.92	41.34	48.28
30							.26	43.77	50.89
40							.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

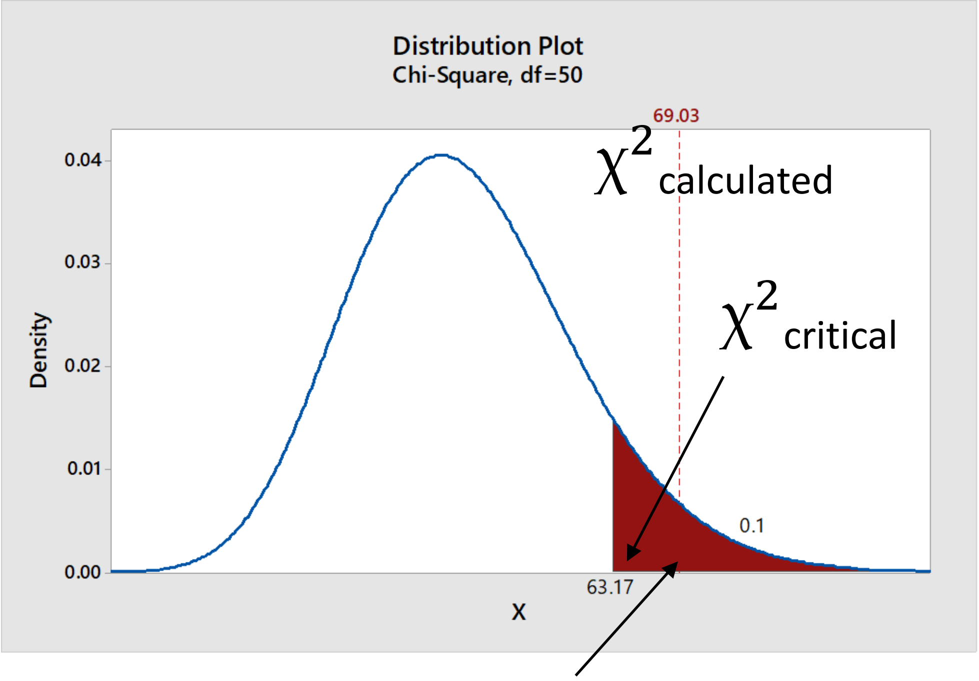
- Example 1: A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **increased** from established 2 cc? 90% confidence level.

- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{(\mathbf{50})\ x\ \mathbf{2.35}^2}{\mathbf{2}^2} = 69.03$

- $\chi^2(\text{critical}) = 63.17$

One Variance Test

$H_o: s^2 <= \sigma^2$
$H_a: s^2 > \sigma^2$



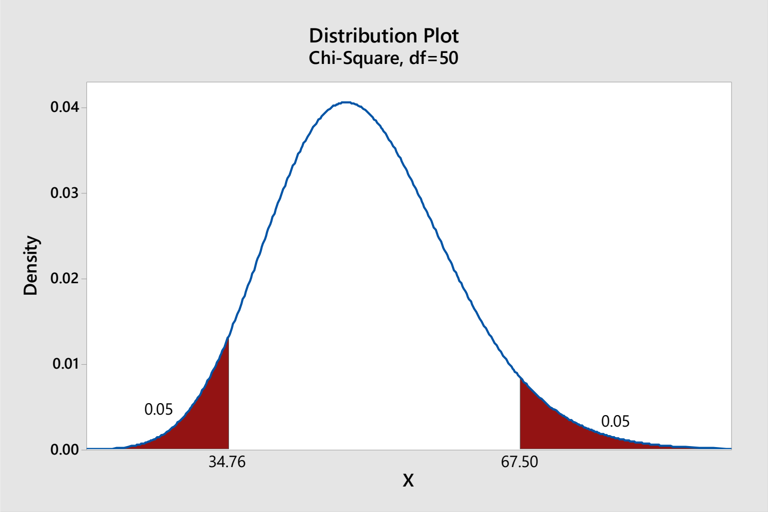
Reject H_0

- Example 1: A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **increased** from established 2 cc? 90% confidence level.
- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{(50)\ x\ 2.35^2}{2^2} = 69.03$
- $\chi^2(\text{critical}) = 63.17$

One Variance Test

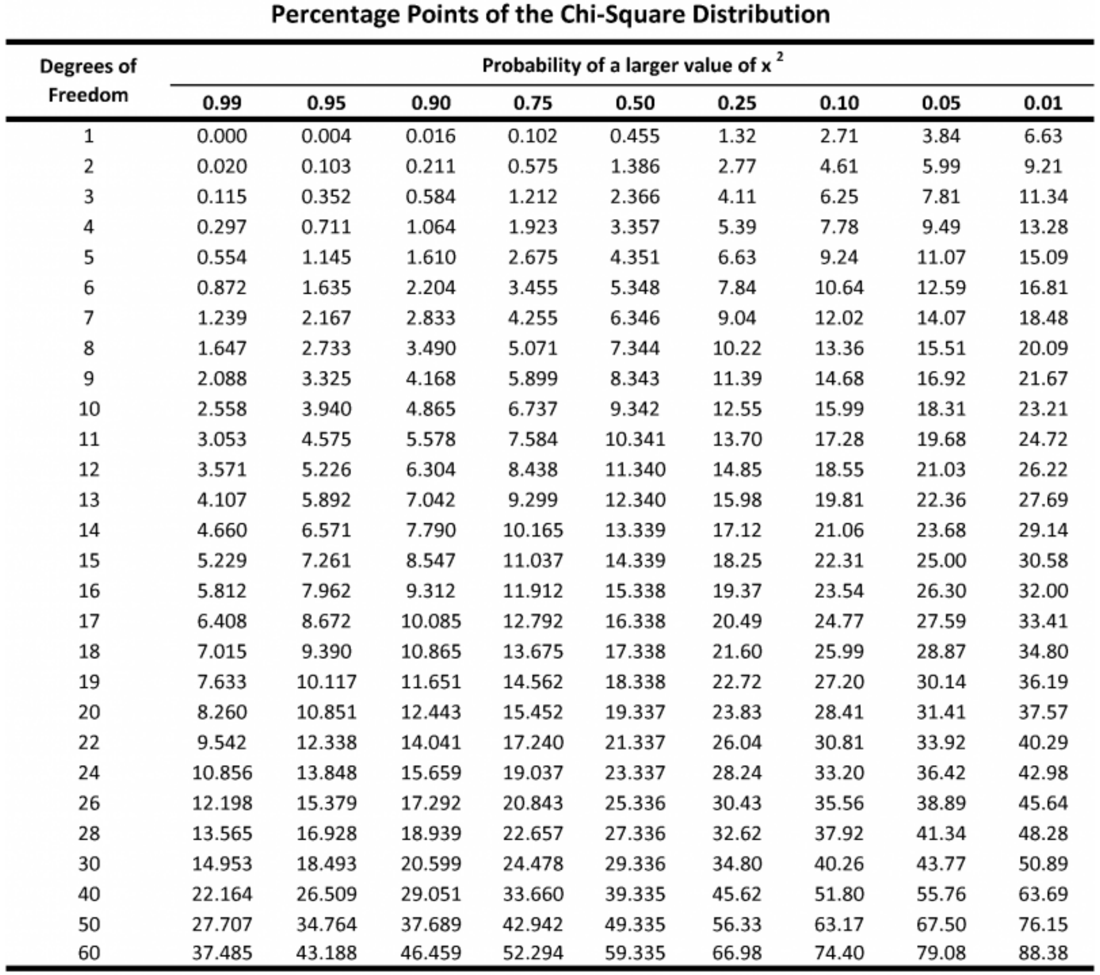
Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x ²								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10								18.31	23.21
11								19.68	24.72
12								21.03	26.22
13								22.36	27.69
14								23.68	29.14
15								25.00	30.58
16								26.30	32.00
17								27.59	33.41
18								28.87	34.80
19								30.14	36.19
20								31.41	37.57
22								33.92	40.29
24								36.42	42.98
26								38.89	45.64
28								41.34	48.28
30								43.77	50.89
40								55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



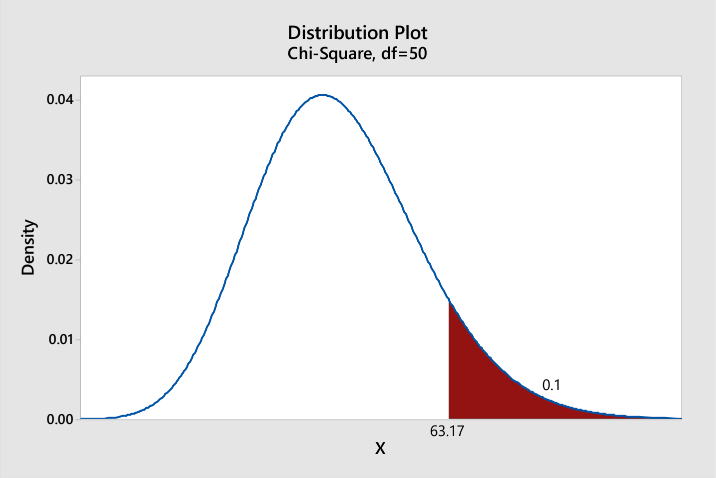
- **Example 2:** A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **changed** from established 2 cc? 90% confidence level.
- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{(50)\ x\ 2.35^2}{2^2} = 69.03$
- What is critical value of Chi Square for 50 degrees of freedom? (two tails test)

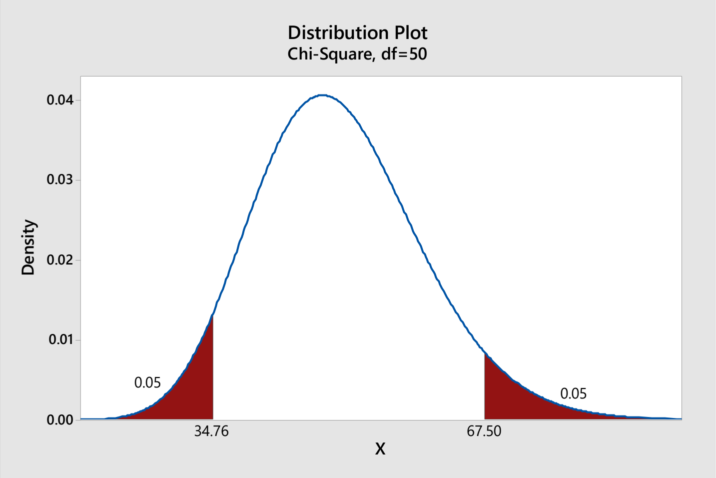
Critical Test Statistic



Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	**0.95**	**0.90**	**0.75**	**0.50**	**0.25**	**0.10**	**0.05**	**0.01**
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38





- α = 0.10 One Tail
- Df = 50
- χ^2 Critical = 63.17

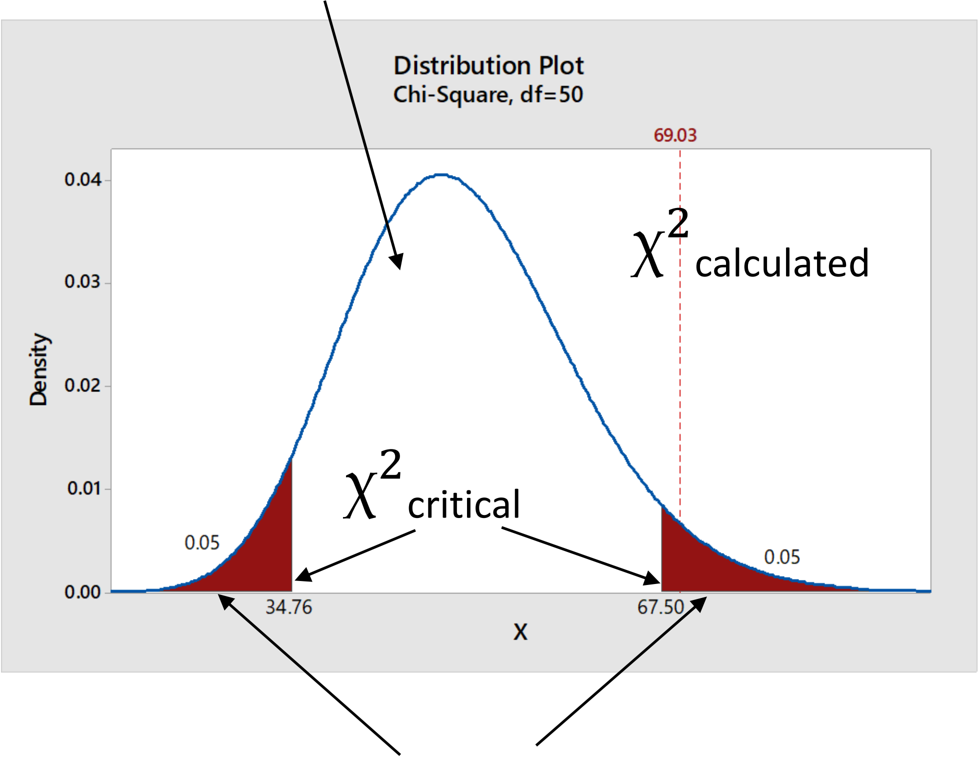
- α = 0.10 Two Tail
- Df = 50
- χ^2Critical = 34.76 and 67.50

One Variance Test

One Variance Test

Ho: $s^2 = \sigma^2$
Ha: $s^2 \neq \sigma^2$

Fail to Reject H_0

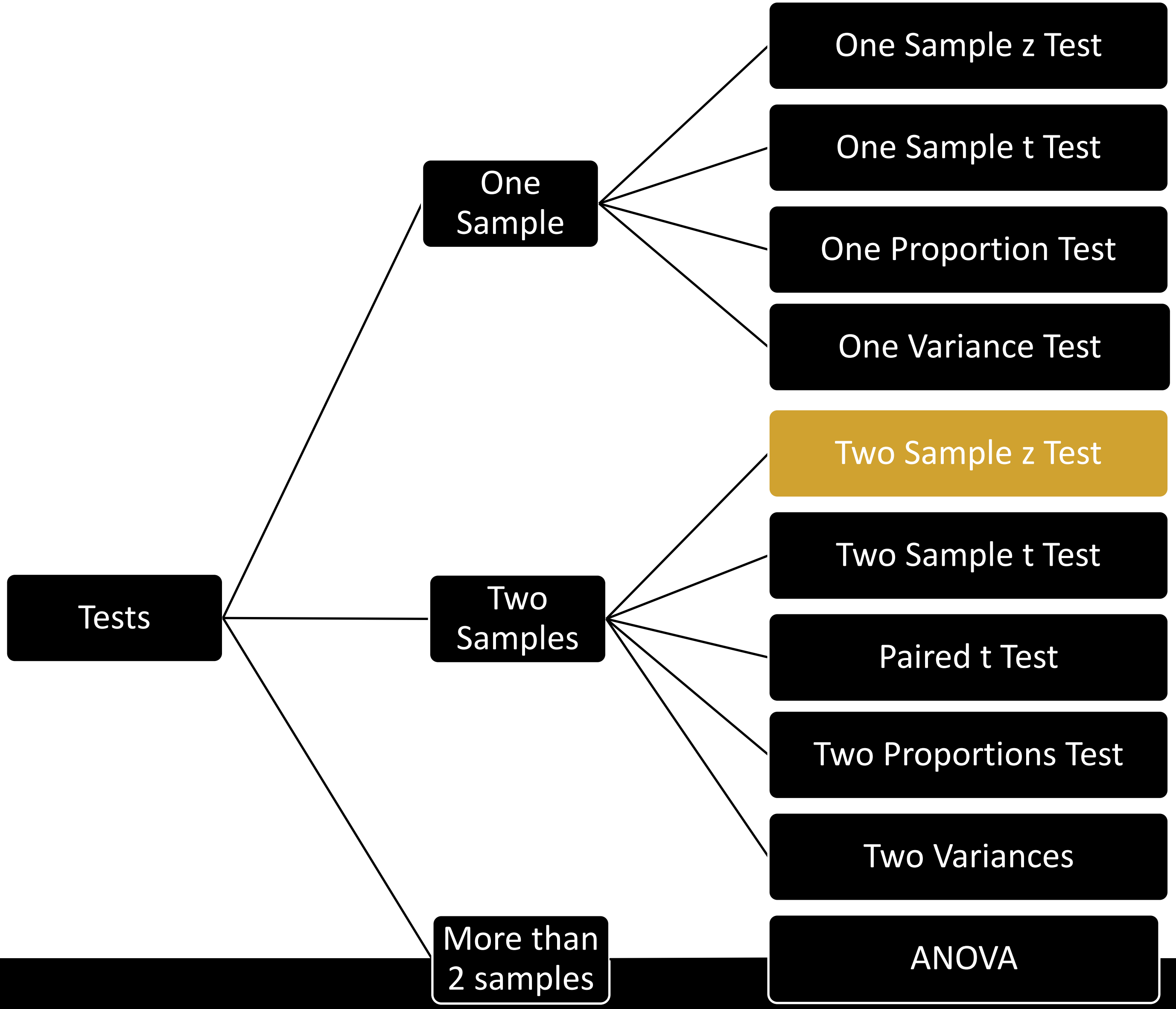


Reject H_0

- **Example 2:** A sample of 51 bottles was selected. The standard deviation of these 51 bottles was 2.35 cc. Has it **changed** from established 2 cc? 90% confidence level.
- $\chi^2(cal.) = \frac{(n-1)s^2}{\sigma^2} = \frac{\mathbf{(50)}\ \boldsymbol{x}\ \mathbf{2.35^2}}{\mathbf{2^2}} = 69.03$
- $\chi^2(\text{critical}) = 34.76$ and 67.50

Section 8

Hypothesis Testing Part 2



- Tests
  - One Sample
    - One Sample z Test
    - One Sample t Test
    - One Proportion Test
    - One Variance Test
  - Two Samples
    - Two Sample z Test
    - Two Sample t Test
    - Paired t Test
    - Two Proportions Test
    - Two Variances
  - More than 2 samples
    - ANOVA

Two Sample Z Test

Conditions for z Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the population standard deviation is known *** OR ***
  - Sample size ≥ 30

Two Sample Z Test

Z Test

One Sample

H_0: μ = 150cc
H_a: μ ≠ 150cc

$$z_{cal} = \frac{(\bar{x} - \mu)}{\sigma/\sqrt{n}}$$

Two Sample

- **Null hypothesis**: $H_0: \mu_1 = \mu_2$
  - or $H_0: \mu_1 - \mu_2 = 0$
- **Alternative hypothesis**: $H_a: \mu_1 \neq \mu_2$

$$z_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Two Sample Z Test

Calculated Test Statistic

H_0: $\mu_1 = \mu_2$
H_a: $\mu_1 \neq \mu_2$

$$z_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$z_{cal} = \frac{(151.2 \ - 151.9)}{\sqrt{\frac{2.1^2}{100} + \frac{2.2^2}{100}}}$$

- Example: From two machines 100 samples each were drawn.

  Machine 1: Mean = 151.2 / sd = 2.1

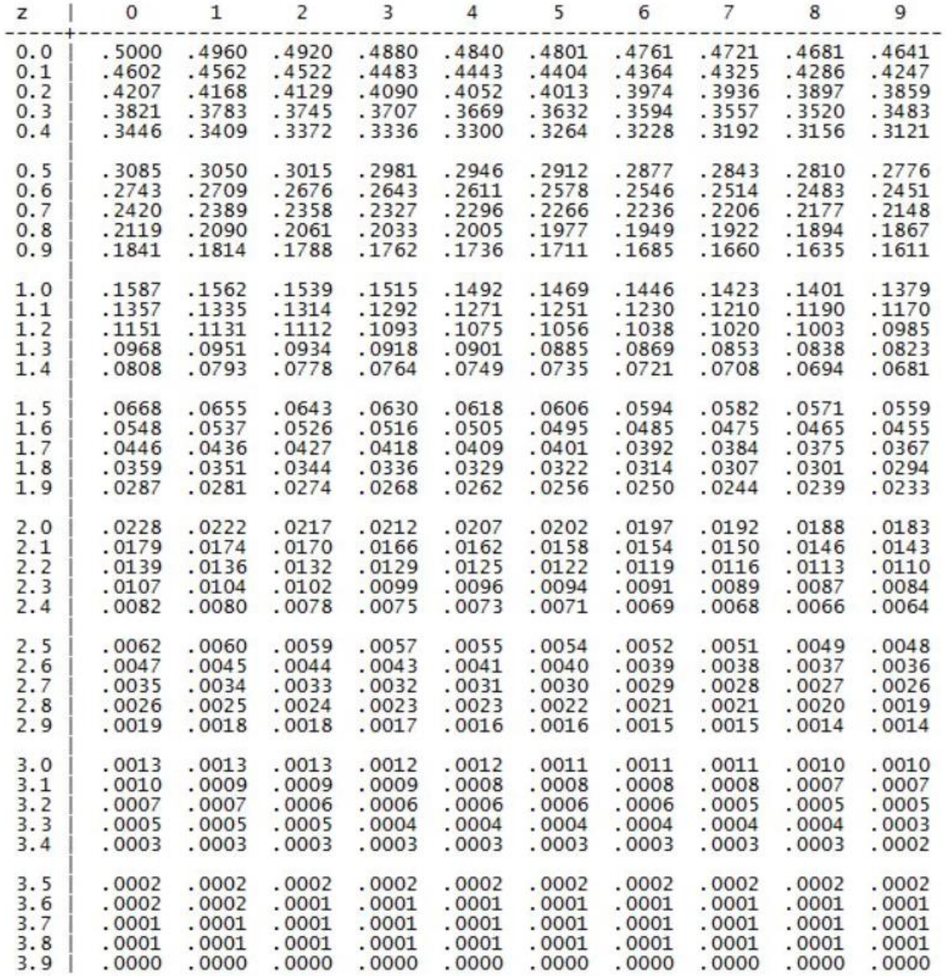
  Machine 2: Mean = 151.9 / sd = 2.2

  Is there difference in these two machines. Check at 95% confidence level.

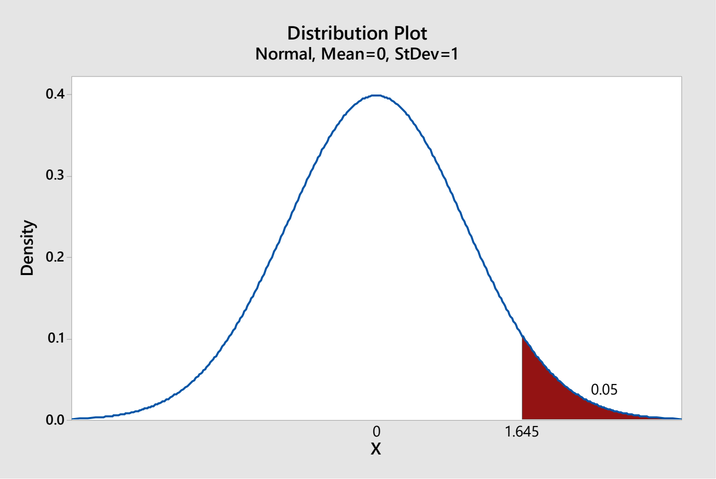
- Zcal = - 0.7 / 0.304 = -2.30
- $z_{critical}$ = ? (for alpha = 0.05, two tail test)

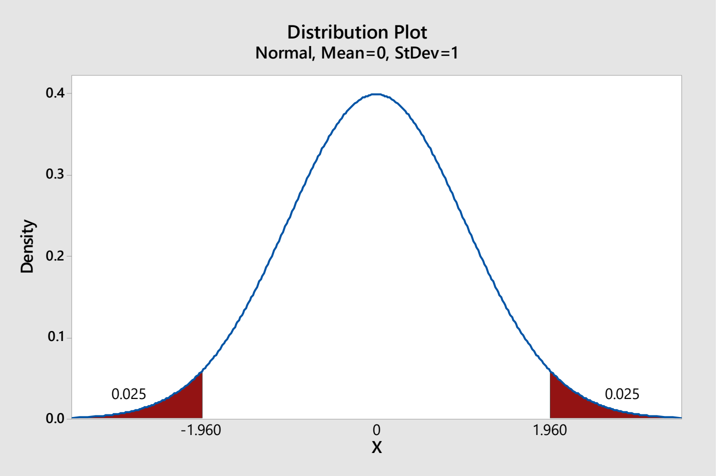
Two Sample Z Test

Critical Test Statistic



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

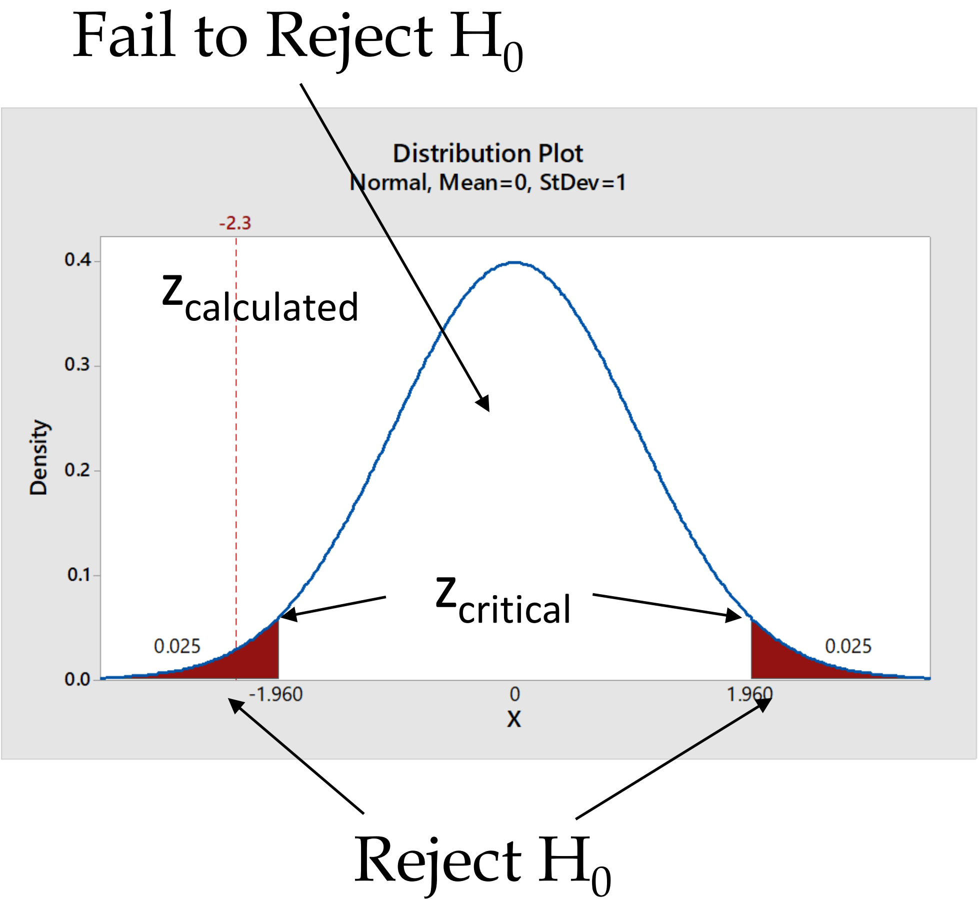
- α = 0.10 One Tail
- Z Critical = 1.282

- **α = 0.05 Two Tails**
- **Z Critical = 1.96**

- α = 0.10 Two Tail
- Z Critical = 1.645

Two Sample Z Test

Interpret the Results



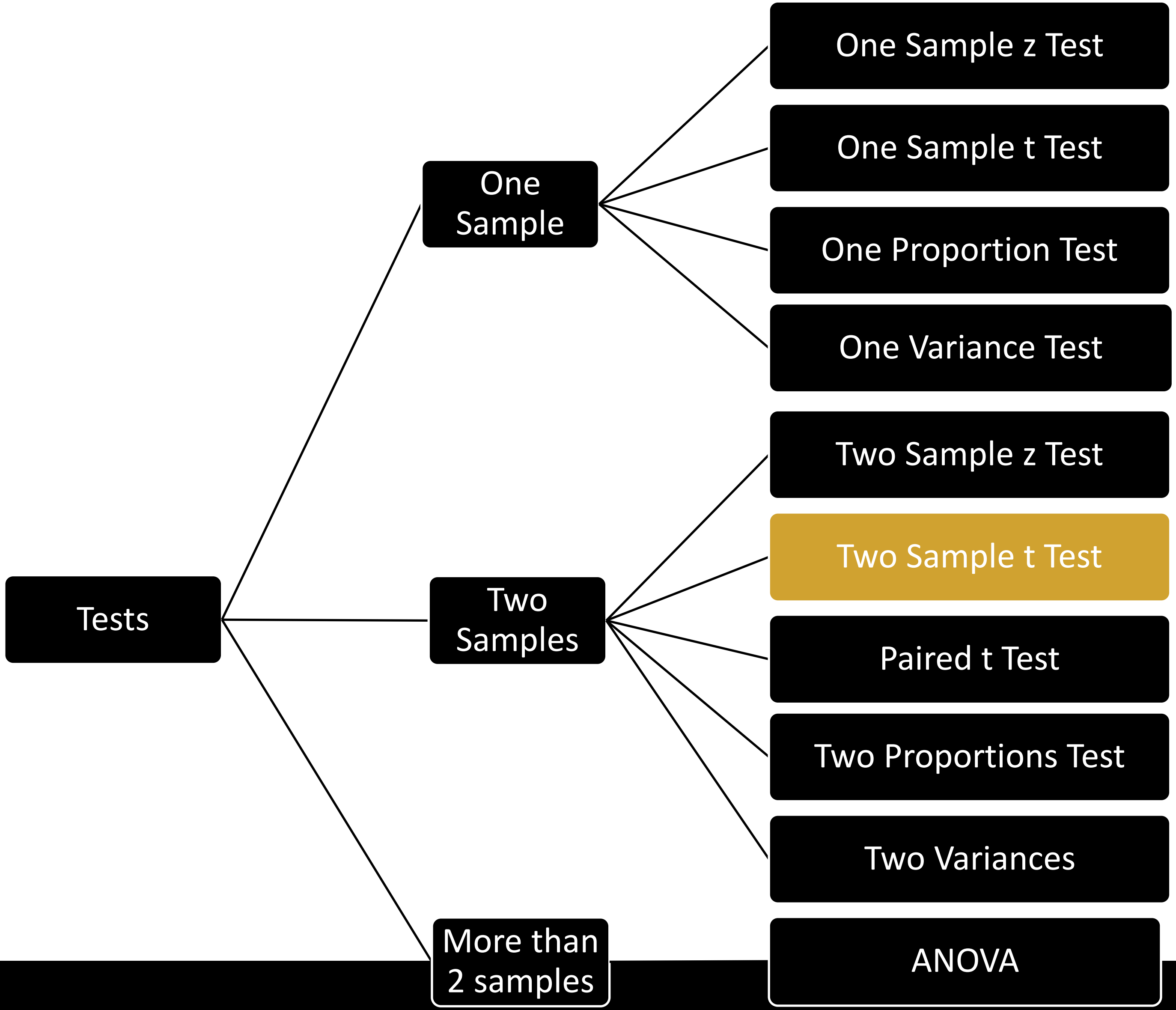
- Example: From two machines 100 samples each were drawn.

  Machine 1: Mean = 151.2 / sd = 2.1

  Machine 2: Mean = 151.9 / sd = 2.2

  Is there difference in these two machines. Check at 95% confidence level.
- Zcal = - 0.7 / 0.304 = -2.30
- $z_{critical} = 1.96$
- Conclusion: Reject Null Hypothesis
  - H_0: $\mu_1 = \mu_2$

Two Sample Z Test



Tests

One Sample

- One Sample z Test
- One Sample t Test
- One Proportion Test
- One Variance Test

Two Samples

- Two Sample z Test
- Two Sample t Test
- Paired t Test
- Two Proportions Test
- Two Variances

More than 2 samples

- ANOVA

Two Sample t Test

Conditions for t Test

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- Sampling distribution approximates Normal Distribution
  - Population is Normally distributed and the standard deviation is **unknown** *** AND ***
  - Sample size < 30

Two Sample t Test

Types of t Tests

- If two set of data are independent or dependent.
  - If the values in one sample reveal no information about those of the other sample, then the samples are independent.
    - Example: Volume produced by two machines
  - If the values in one sample affect the values in the other sample, then the samples are dependent.
    - Example: Blood pressure before and after a specific medicine

Two sample t test

Paired t test

Two Sample t Test

One Sample t Test

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

Two Sample t Tests

- ❖ Is variance for two samples equal?

Yes

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

No

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Two Sample t Test

Conditions for t Test

$$H_0: \mu_A = \mu_B$$
$$H_a: \mu_A \neq \mu_B$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$df = n_1 + n_2 - 2$$

- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

  Is the mean different? Calculate with 95% confidence.

  - $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 1.52$
  - $\bar{x}_1 = 151.8$, $\bar{x}_2 = 155.6$

Two Sample t Test

Conditions for t Test

$H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s_p = 1.50$$

- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

Is the mean **different**? Calculate with 95% confidence.

- $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 1.52$
- $\bar{x}_1 = 151.8$, $\bar{x}_2 = 155.6$

Two Sample t Test

Conditions for t Test

$$H_0: \mu_A = \mu_B$$
$$H_a: \mu_A \neq \mu_B$$

$$s_p = 1.50$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$t_{cal} = -\ 4.01$$

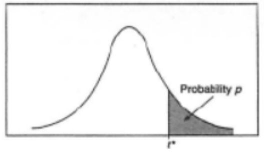
- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

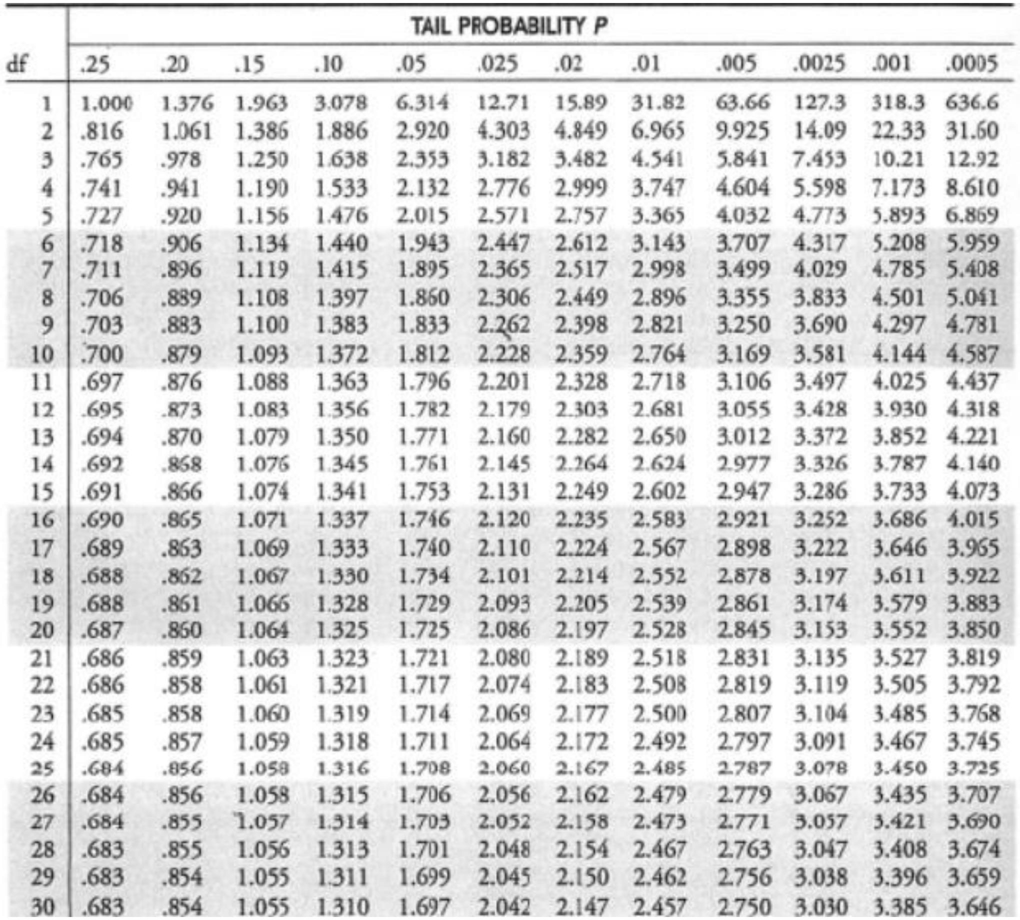
Is the mean different? Calculate with 95% confidence.

- $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 1.52$
- $\bar{x}_1 = 151.8$, $\bar{x}_2 = 155.6$

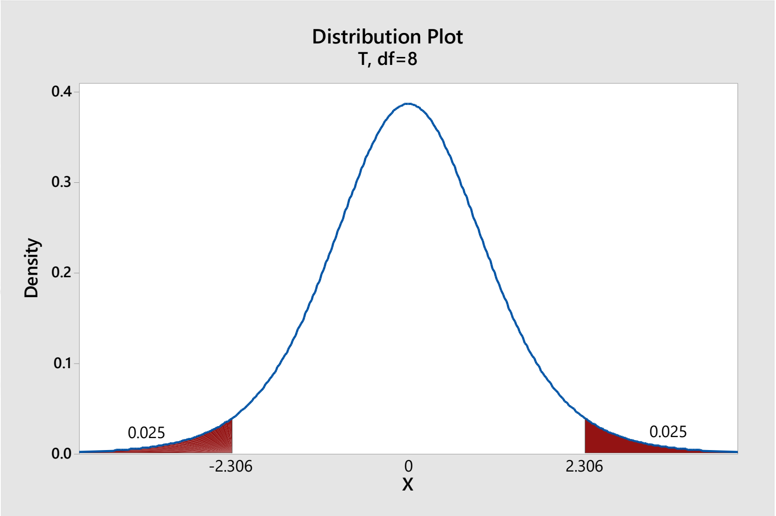
Two Sample t Test

Critical Test Statistic





	TAIL PROBABILITY *P*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



$$df = n_1 + n_2 - 2$$

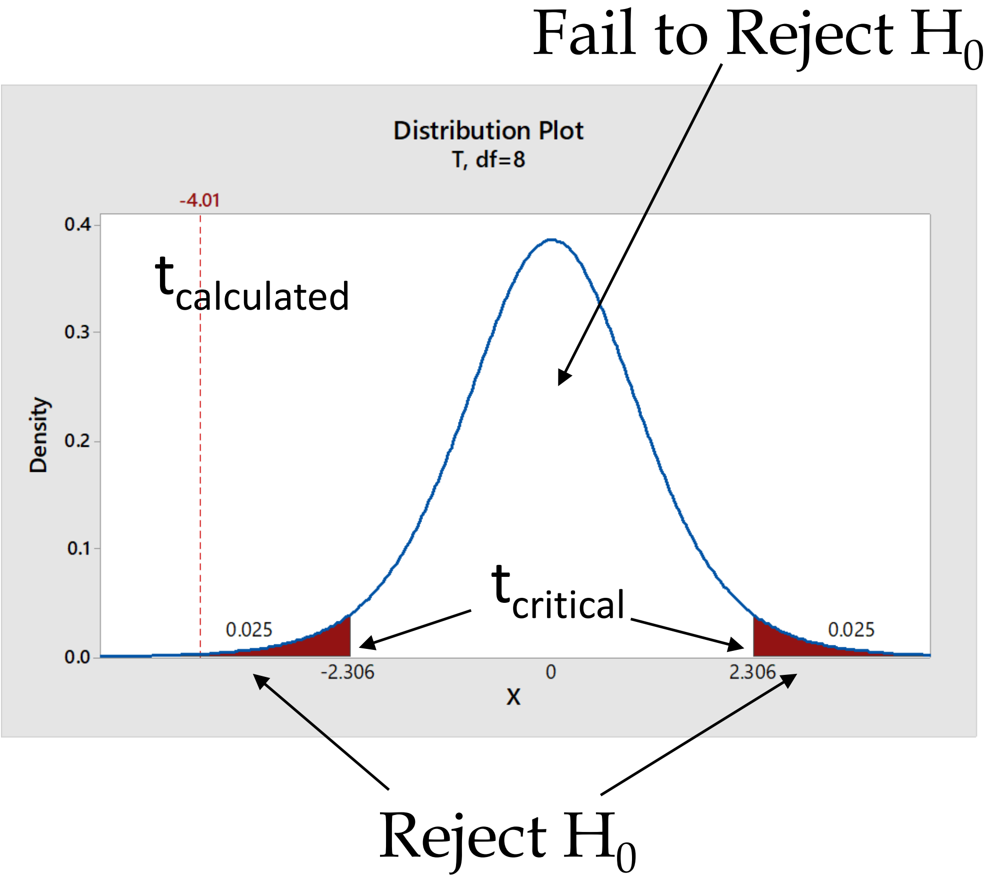
$$df = 8$$

- α = 0.05 Two Tails
- Df = 8
- t Critical = 2.306

Two Sample t Test

$H_0: \mu_A = \mu_B$
$H_a: \mu_A \neq \mu_B$

Interpret the Results



- Example: Samples from two machines A and B have the following volumes in bottles. **(assume equal variance)**
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine B:** 156, 155, 158, 155, 154

  Is the mean **different**? Calculate with 95% confidence.
- $t_{calculated} = -4.01$
- $t_{critical} = 2.306$

Two Sample t Test

One Sample t Test

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

Two Sample t Tests

- ❖ Is variance for two samples equal?

Yes

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

No

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Two Sample t Test

Conditions for t Test

H_0: $\mu_A = \mu_C$
H_a: $\mu_A \neq \mu_C$

$$\xcancel{t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_p\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

- ❖ Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - ❖ **Machine A:** 150, 152, 154, 152, 151
  - ❖ **Machine C:** 144, 162, 177, 150, 140

  Is the mean different? Calculate with 95% confidence.

  - ❖ $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 15.0$
  - ❖ $\bar{x}_1 = 151.8$, $\bar{x}_2 = 154.6$

Two Sample t Test

Conditions for t Test

H_0: $\mu_A = \mu_C$
H_a: $\mu_A \neq \mu_C$

$$t_{cal} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$t_{cal} = -0.41$

- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

Is the mean different? Calculate with 95% confidence.

- $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 15.0$
- $\bar{x}_1 = 151.8$, $\bar{x}_2 = 154.6$

Two Sample t Test

Conditions for t Test

H_0: $\mu_A = \mu_C$
H_a: $\mu_A \neq \mu_C$

$$\cancel{df = n_1 + n_2 - 2}$$

$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

$df = 4.078$ or 4 (round down to nearest integer)

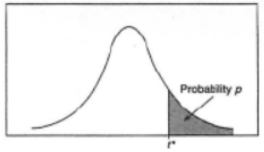
- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

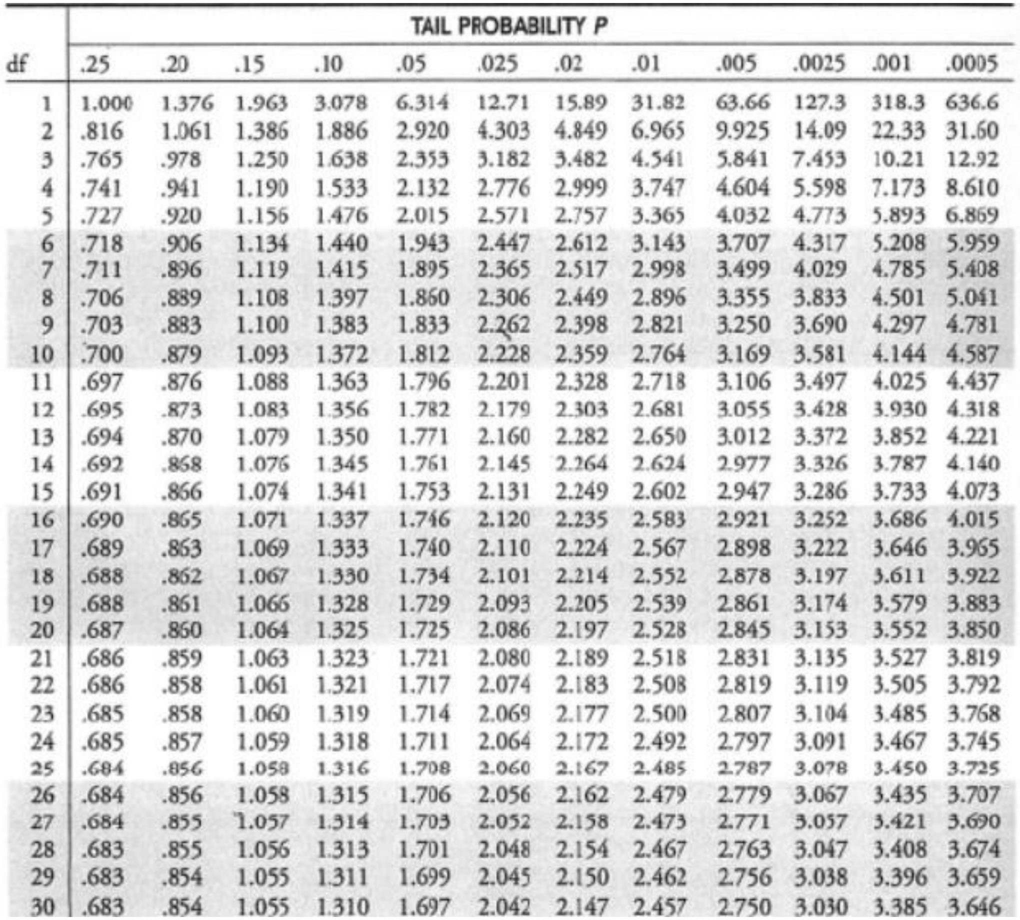
Is the mean different? Calculate with 95% confidence.

- $n_1 = 5$, $n_2 = 5$, $s_1 = 1.48$, $s_2 = 15.0$
- $\bar{x}_1 = 151.8$, $\bar{x}_2 = 154.6$

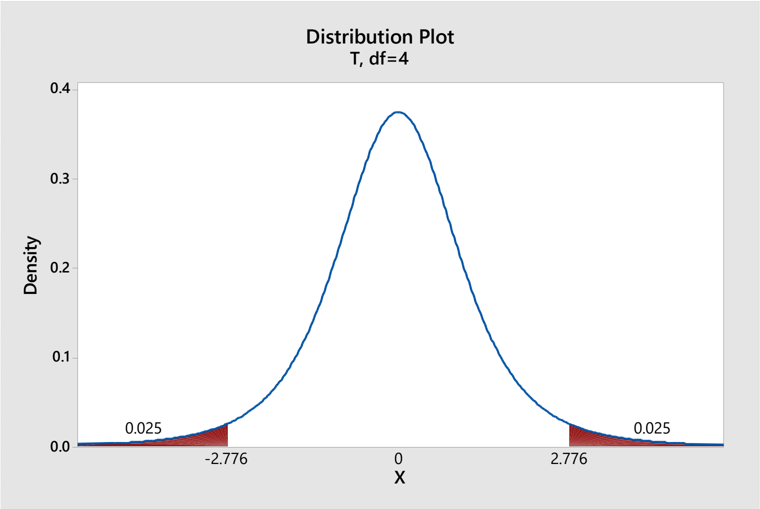
Two Sample t Test

Critical Test Statistic





	TAIL PROBABILITY *P*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



$$df = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

$$df = 4$$

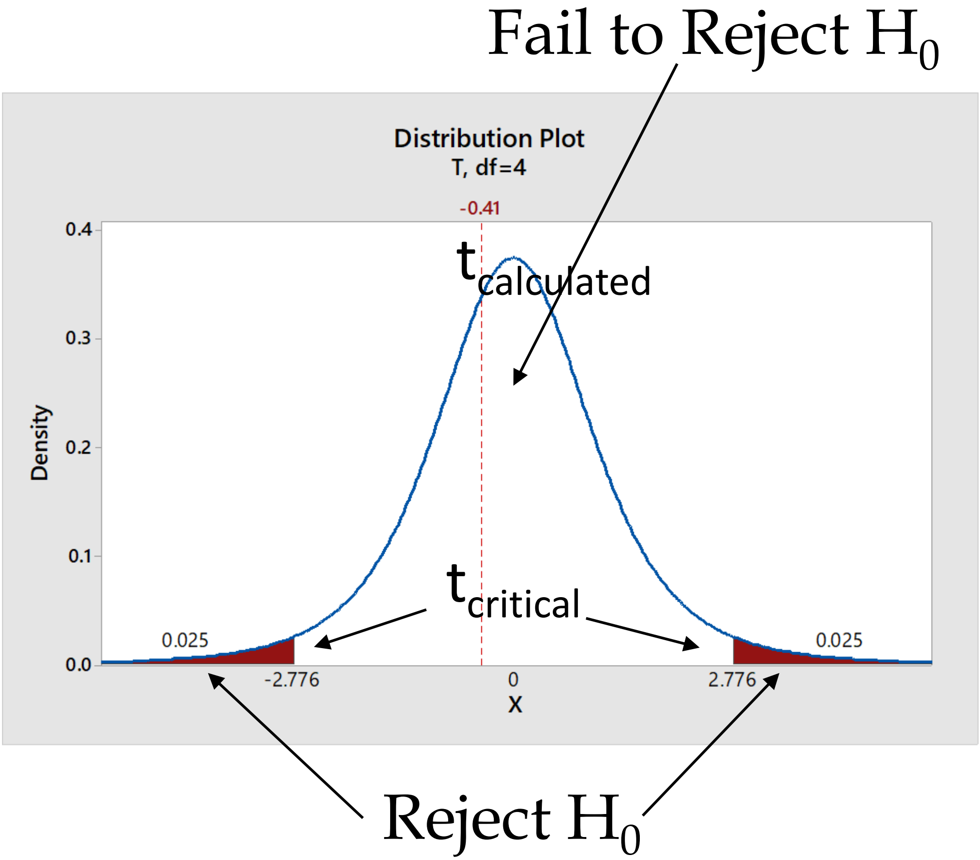
- ❖ α = 0.05 Two Tails
- ❖ Df = 4
- ❖ t Critical = 2.776

Two Sample t Test

Interpret the Results

$H_0: \mu_A = \mu_C$

$H_a: \mu_A \neq \mu_C$



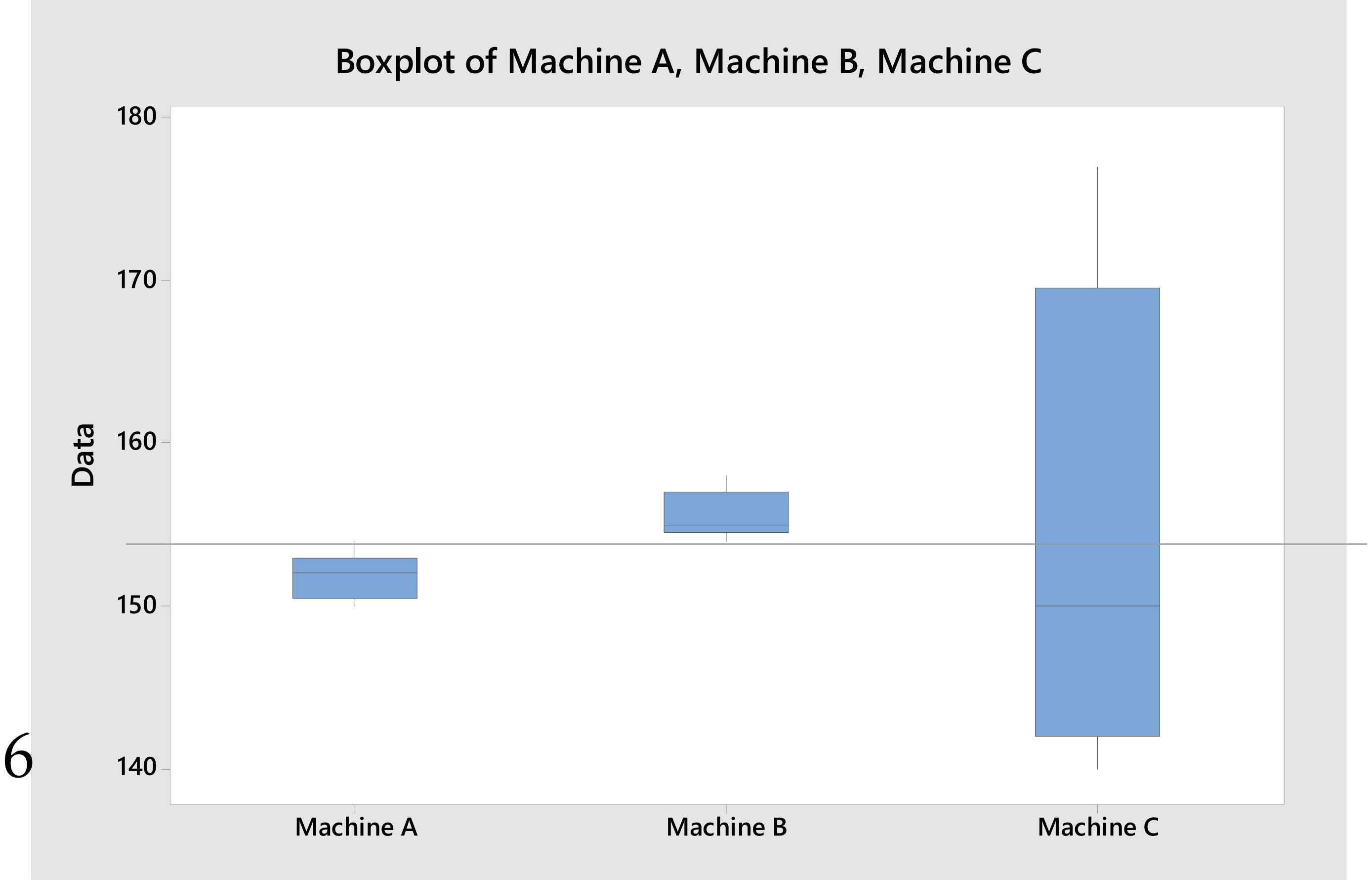
- Example: Samples from two machines A and B have the following volumes in bottles. (**unequal variance**)
  - **Machine A:** 150, 152, 154, 152, 151
  - **Machine C:** 144, 162, 177, 150, 140

  Is the mean different? Calculate with 95% confidence.
- $t_{calculated} = - 0.41$
- $t_{critical} = 2.776$

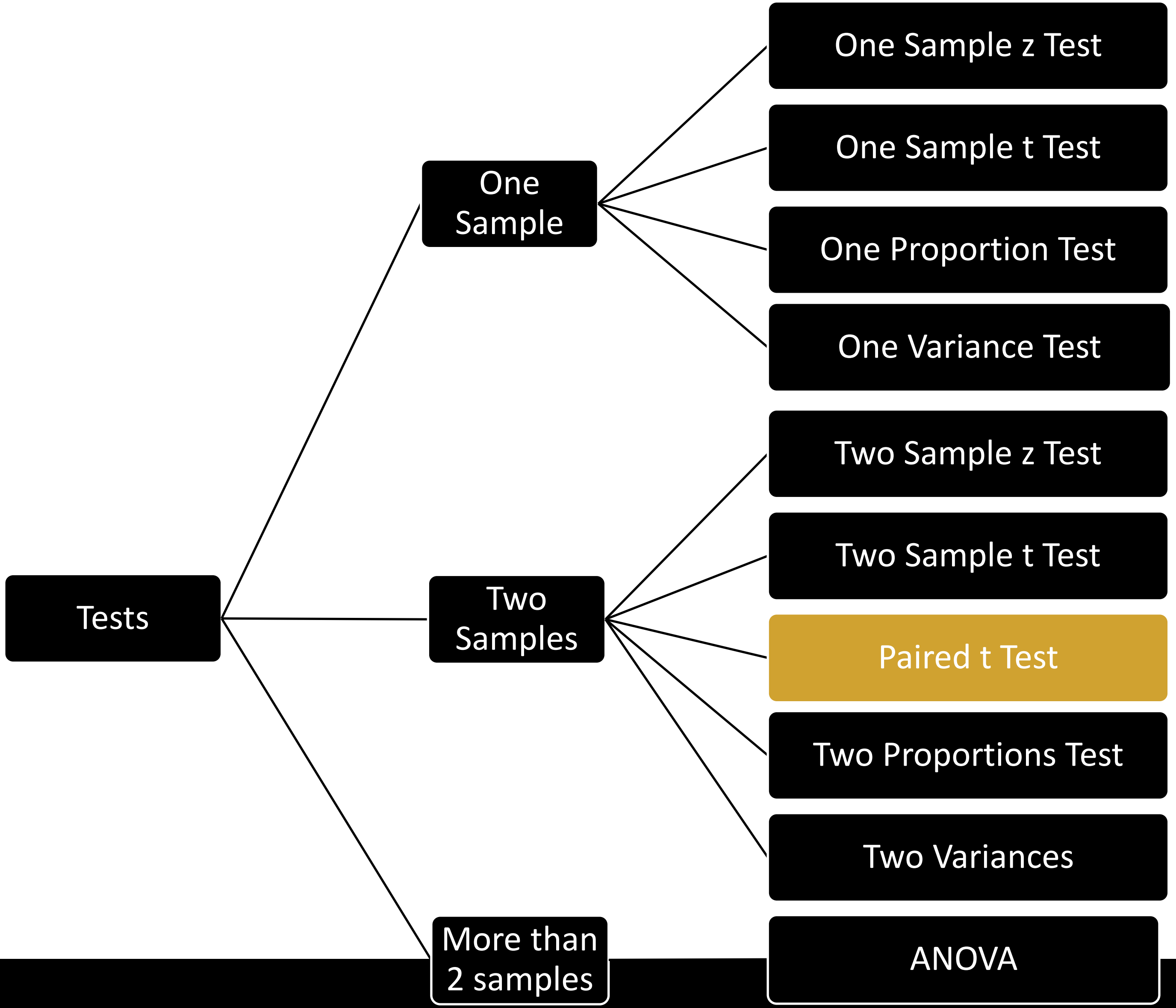
Interpret the Results

Machine A	Machine B	Machine C
150	156	144
152	155	162
154	158	177
152	155	150
151	154	140

$\bar{x}_{\mathbf{A}} = 151.8$, $\bar{x}_{\mathrm{B}} = 155.6$ $\bar{x}_{\mathrm{C}} = 154.6$



Two Sample t Test



Tests

One Sample

- One Sample z Test
- One Sample t Test
- One Proportion Test
- One Variance Test

Two Samples

- Two Sample z Test
- Two Sample t Test
- Paired t Test
- Two Proportions Test
- Two Variances

More than 2 samples

- ANOVA

Paired t Test

Types of t Tests

- If two set of data are independent or dependent.
  - If the values in one sample reveal no information about those of the other sample, then the samples are independent.
    - Example: Blood pressure of male/female
  - If the values in one sample affect the values in the other sample, then the samples are dependent.
    - Example: Blood pressure before and after a specific medicine

Two sample t test

Paired t test

Paired t Test

One Sample t Test

$$t_{cal} = \frac{(\bar{x} - \mu)}{s/\sqrt{n}}$$

Paired t Tests

- ❖ Find the difference between two set of readings as d1, d2 …. dn.
- ❖ Find the mean and standard deviation of these differences.

$$t = \frac{\bar{d}}{s/\sqrt{n}}$$

Paired t Test

Paired t Tests

$H_0: \mu_{before} = \mu_{after}$
$H_a: \mu_{before} \neq \mu_{after}$

Patient	Before	After
1	120	122
2	122	120
3	143	141
4	100	109
5	109	109

- Example: Before and after medicine BP was measured. Is there a difference at 95% confidence level?

Paired t Test

Paired t Tests

$H_0: \mu_{before} = \mu_{after}$
$H_a: \mu_{before} \neq \mu_{after}$

$$t = \frac{\bar{d}}{s/\sqrt{n}}$$

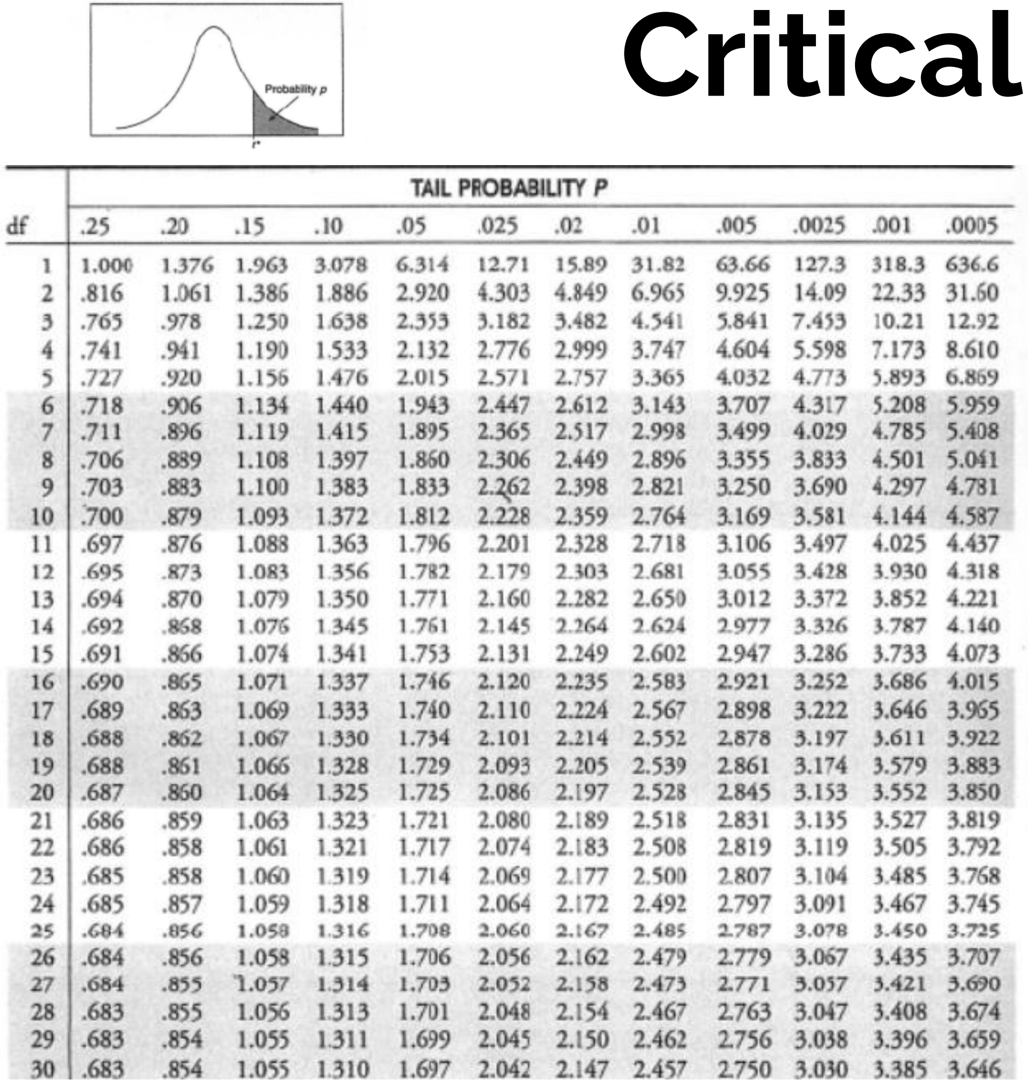
Patient	Before	After	difference
1	120	122	-2
2	122	120	2
3	143	141	2
4	100	109	-9
5	109	109	0

- Example: Before and after medicine BP was measured. Is there a **difference** at 95% confidence level?

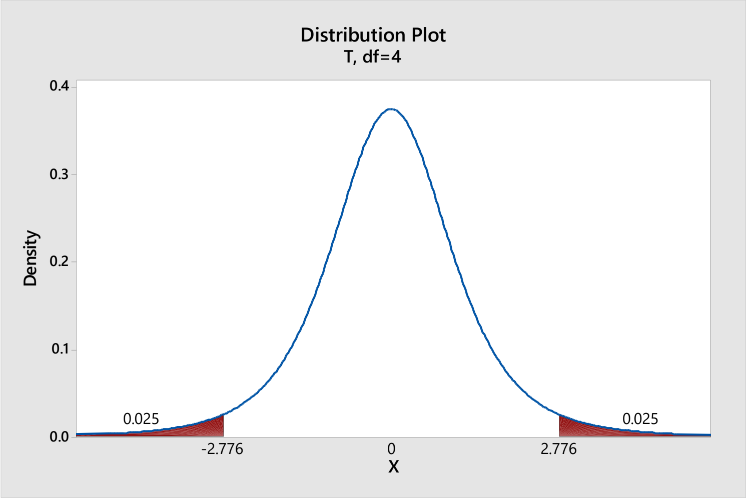
- $\bar{d}$ = -1.4 , s = 4.56 , n = 5
- $t_{cal.}$ = 1.4/2.04 = - 0.69

Paired t Test

Critical Test Statistic



df	TAIL PROBABILITY P											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646



$$df = n - 1$$

$$df = 4$$

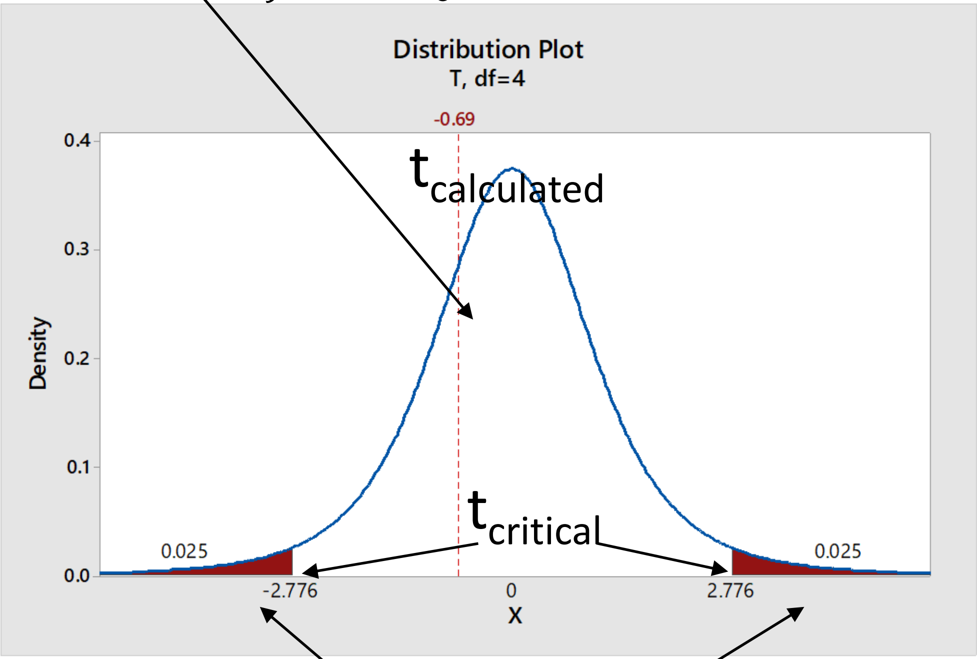
- α = 0.05 Two Tails
- df = 4
- t Critical = 2.776

Paired t Test

Interpret the Results

$H_0: \mu_{before} = \mu_{after}$
$H_a: \mu_{before} \neq \mu_{after}$

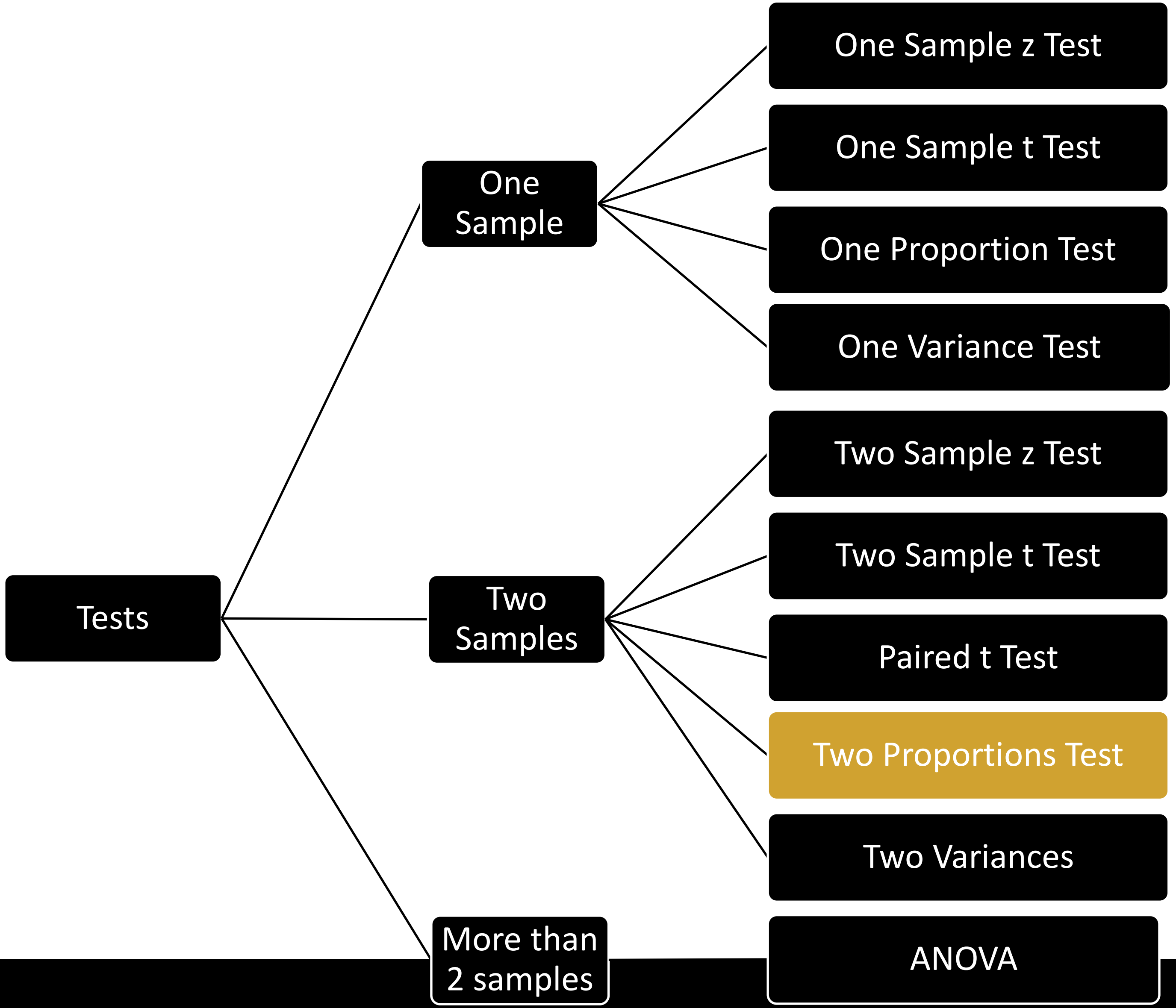
Fail to Reject H_0



Reject H_0

- Example: Before and after medicine BP was measured. Is there a **difference** at 95% confidence level?

- $t_{calculated} = - 0.69$
- $t_{critical} = 2.776$



Tests

One Sample

Two Samples

More than 2 samples

One Sample z Test

One Sample t Test

One Proportion Test

One Variance Test

Two Sample z Test

Two Sample t Test

Paired t Test

Two Proportions Test

Two Variances

ANOVA

Two Proportions Test

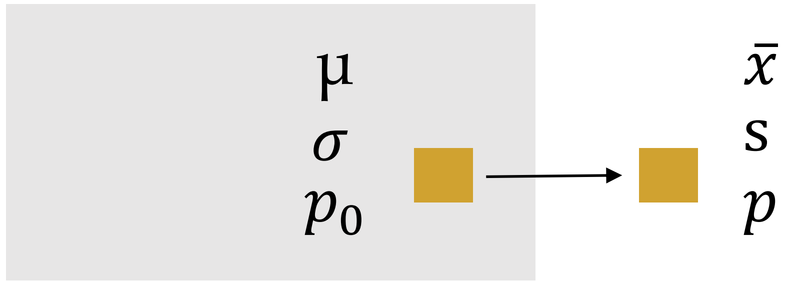
Conditions for Proportions Test

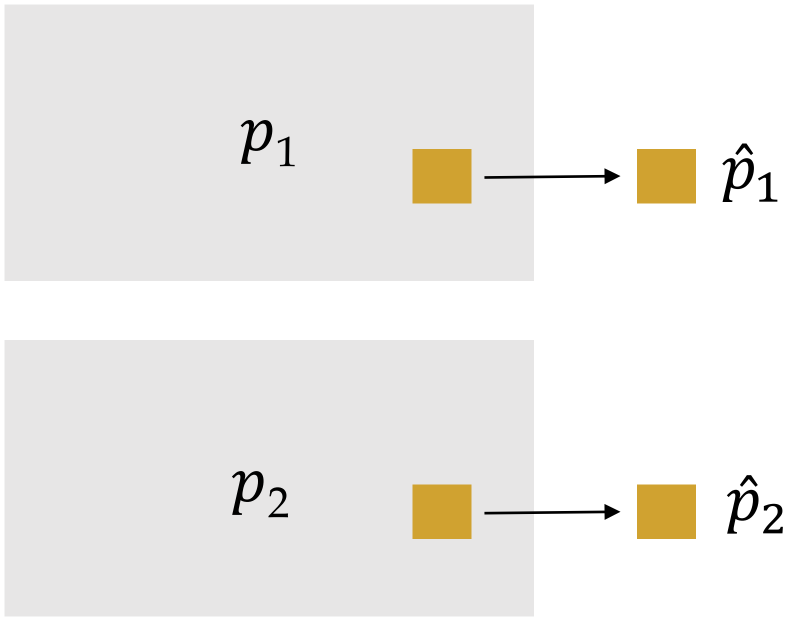
- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data contains only two categories, such as pass/fail or yes/no
- For Normal approximation:
  - both $np \geq 10$ and $n(1-p) \geq 10$ (data should have at least 10 "successes" and at least 10 "failures") **for each sample** (in some books it is 5)

Two Proportions Test

Proportions – Sample vs Population

How do we represent sample and population proportions?





One Proportion Test

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Two Proportions Test

One Proportion Test

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

Two Proportions Tests

Test for **no** difference between proportions

Yes

Pooled

No

Un-pooled

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

H_0: $p_1 - p_2 = d$
H_a: $p_1 - p_2 \neq d$

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2) - (\mathrm{p}_1 - \mathrm{p}_2)}{\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}}$$

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

Test if Normality can be assumed?

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- ❖ Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- ❖ $z_{calculated}$ = ?
- ❖ $z_{critical}$ = ?

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

Test if Normality can be assumed?

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $\hat{p}_1$= 30/200 = 0.15
- $\hat{p}_2$= 10/100 = 0.10
- n_1 = 200, n_2 = 100

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

Test if Normality can be assumed?

$$n_1\hat{p}_1 \geq 10$$
$$n_1(1 - \hat{p}_1) \geq 10$$

$$n_2\hat{p}_2 \geq 10$$
$$n_1(1 - \hat{p}_2) \geq 10$$

- ❖ Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- ❖ $\hat{p}_1$= 30/200 = 0.15
- ❖ $\hat{p}_2$= 10/100 = 0.10
- ❖ n_1 = 200, n_2 = 100

Two Proportions Test

Two Proportions Tests

H_0: $p_1 - p_2 = 0$
H_a: $p_1 - p_2 \neq 0$

$$\bar{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{(n_1 + n_2)}$$

$$\bar{p} = 0.1333$$

- ❖ Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- ❖ $\hat{p}_1$= 30/200 = 0.15
- ❖ $\hat{p}_2$= 10/100 = 0.10
- ❖ n_1 = 200, n_2 = 100

Two Proportions Test

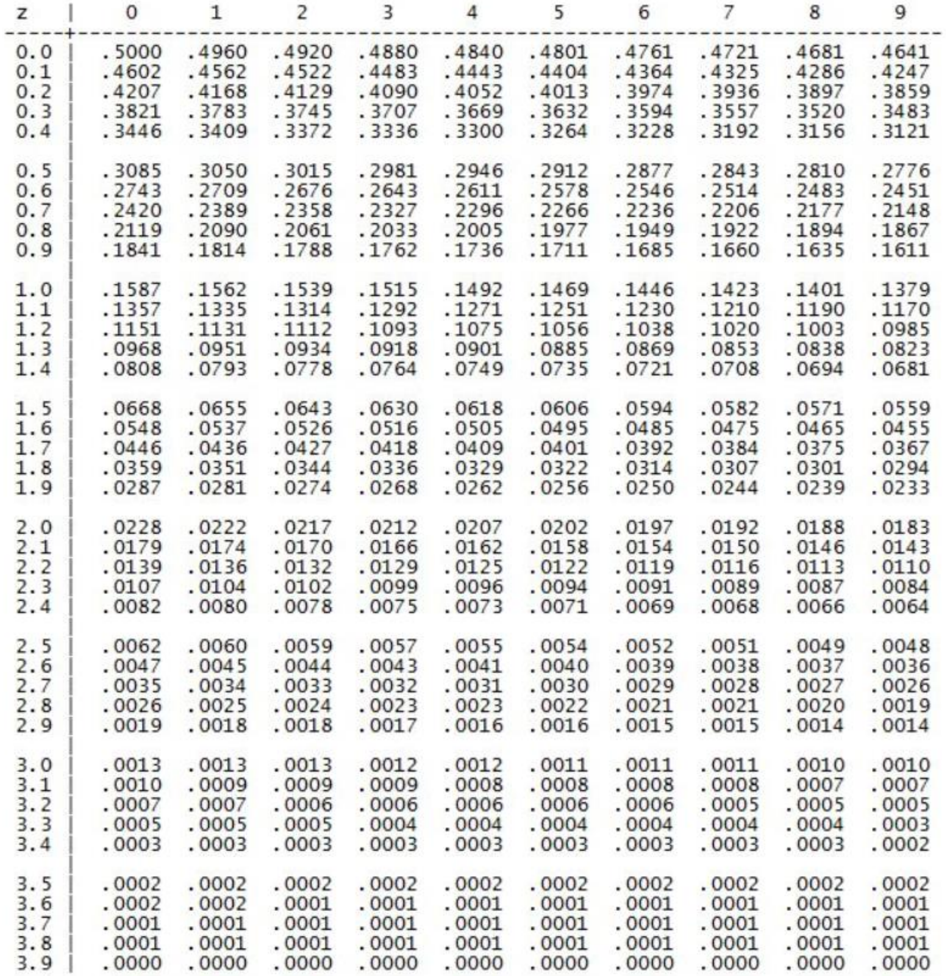
Two Proportions Tests

$$H_0: \ p_1 - p_2 = 0$$
$$H_a: \ p_1 - p_2 \neq 0$$
$$\bar{p} = 0.1333$$
$$z_{cal} = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$
$$\mathrm{z_{cal}} = \frac{(0.15 - 0.10)}{\sqrt{0.133(1-0.133)\left(\frac{1}{200}+\frac{1}{100}\right)}}$$
$$\mathrm{z_{cal}} = 1.20$$

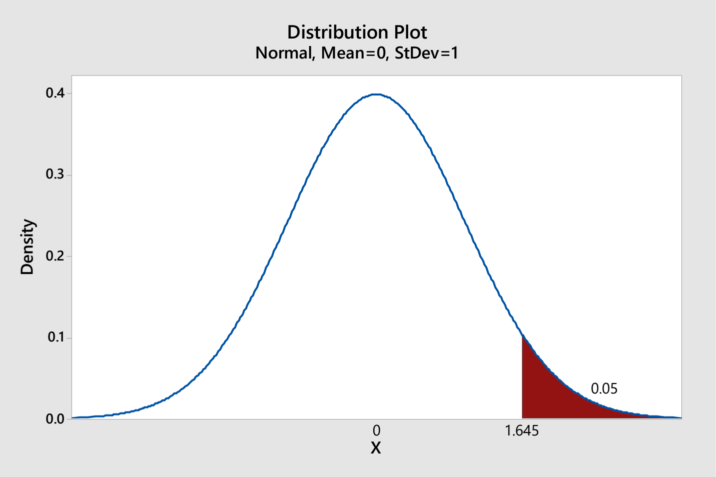
- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $\hat{p}_1$= 30/200 = 0.15
- $\hat{p}_2$= 10/100 = 0.10
- n_1 = 200, n_2 = 100

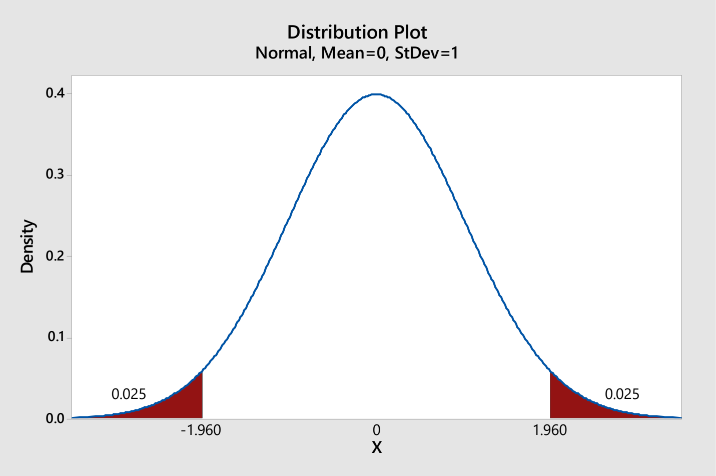
Two Proportions Test

Critical Test Statistic



z	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000





- α = 0.05 One Tail
- Z Critical = 1.645

- α = 0.10 One Tail
- Z Critical = 1.282

- α = 0.05 Two Tails
- Z Critical = 1.96

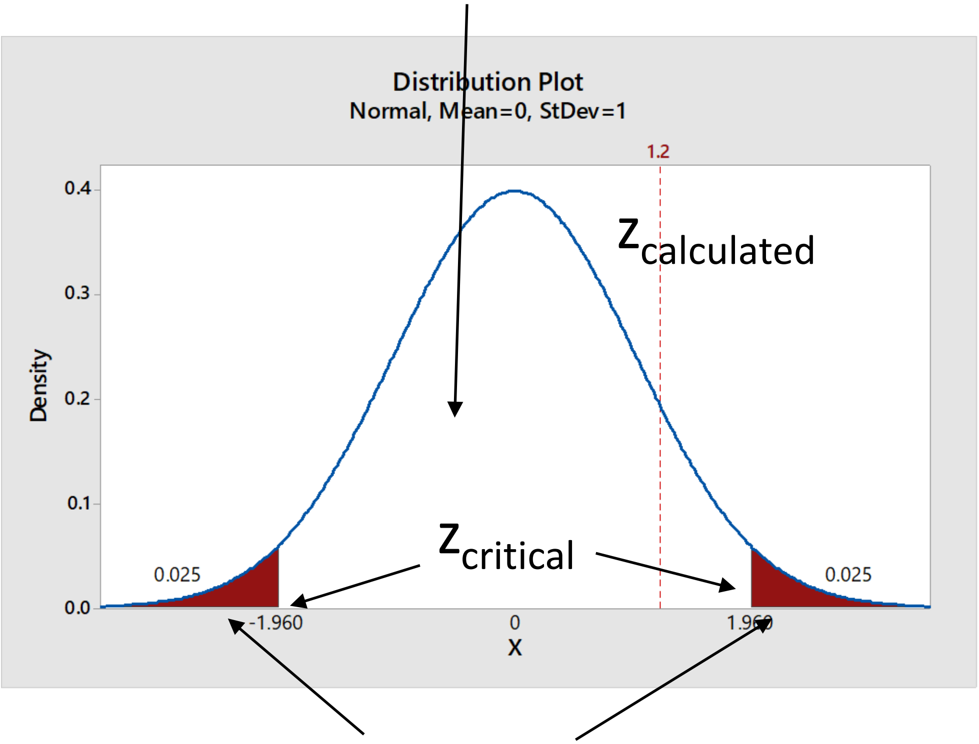
- α = 0.10 Two Tail
- Z Critical = 1.645

Two Proportions Test

Interpret the Results

H_0: $p = p_0$
H_a: $p \neq p_0$

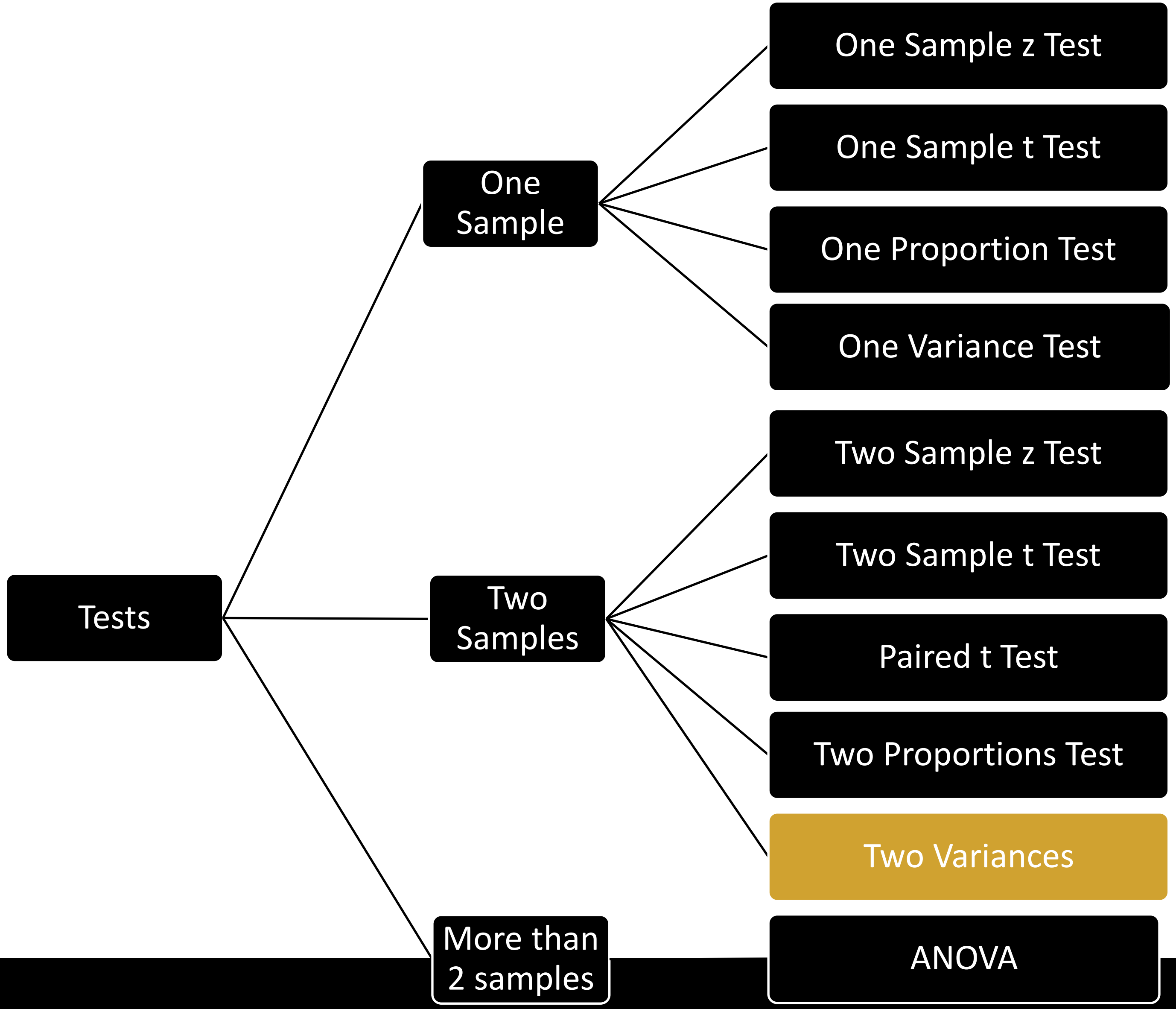
Fail to Reject H_0



Reject H_0

- Example: From vendor A we test 200 pieces and find 30 defectives. From vendor B we test 100 pieces and we find 10 defectives. Is there a significant difference in the quality of these two vendors? Use 95% confidence level.
- $z_{calculated} = 1.20$
- $z_{critical} = 1.96$

Two Proportions Test



- Tests
  - One Sample
    - One Sample z Test
    - One Sample t Test
    - One Proportion Test
    - One Variance Test
  - Two Samples
    - Two Sample z Test
    - Two Sample t Test
    - Paired t Test
    - Two Proportions Test
    - Two Variances
  - More than 2 samples
    - ANOVA

Two Variances Test

Conditions for Variance Tests

- Random samples
- Each observation should be independent of other
  - Sampling with replacement
  - If sampling without replacement, the sample size should not be more than 10% of the population
- The data follows a Normal Distribution

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Two Variances Test

Two Variances Test

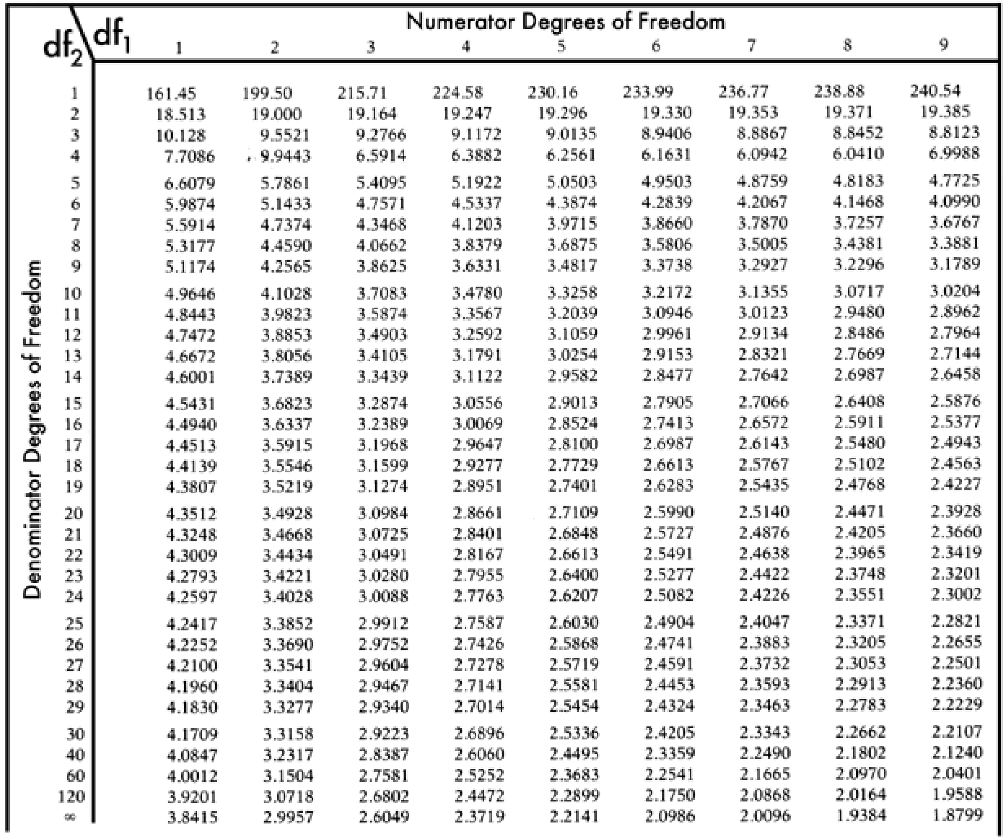
$H_0: \sigma^2{}_1 = \sigma^2{}_2$
$Ha: \sigma^2{}_1 \neq \sigma^2{}_2$

$$F_{cal} = \frac{s_1^2}{s_2^2}$$

- Example: We took 8 samples from machine A and the **standard deviation** was 1.1. For machine B we took 5 samples and the **variance** was 11. Is there a difference in variance at 90% confidence level?
- n1 = 5, $s^2{}_1 = 11$, $df_1 = 4$ (numerator)
- n2 = 8, $s_2 = 1.1$, $s^2{}_2 = 1.21$, $df_2 = 7$ (denominator)
- $F_{\text{calculated}} = 11/1.21 = 9.09$ **(higher value at top)**

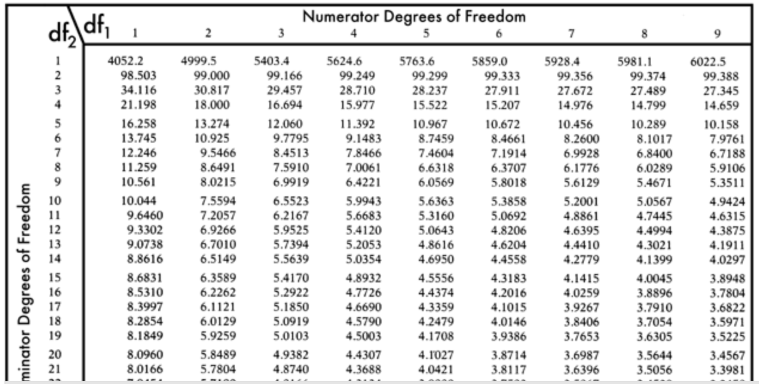
Two Variances Test

Critical Test Statistic



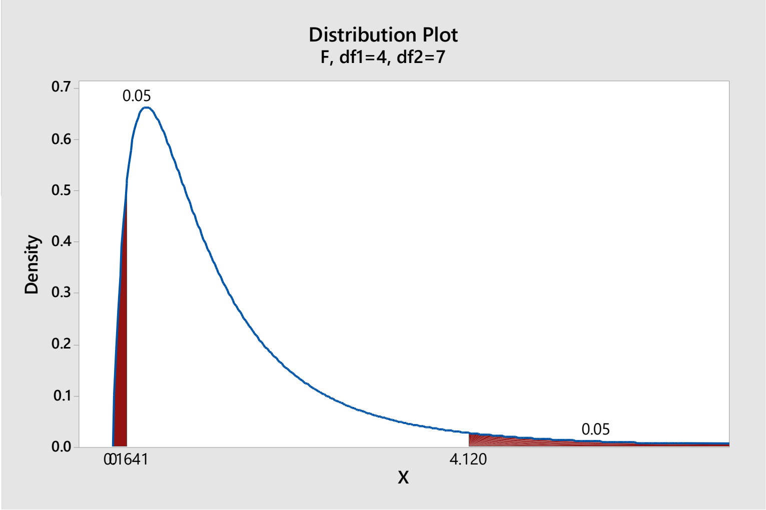
F - Distribution (α = 0.05 in the Right Tail)

df₂ \ df₁	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799



F - Distribution (α = 0.01 in the Right Tail)

df₂ \ df₁	1	2	3	4	5	6	7	8	9
1	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1017	7.9761
7	12.246	9.5466	8.4513	7.8466	7.4604	7.1914	6.9928	6.8400	6.7188
8	11.259	8.6491	7.5910	7.0061	6.6318	6.3707	6.1776	6.0289	5.9106
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9424
11	9.6460	7.2057	6.2167	5.6683	5.3160	5.0692	4.8861	4.7445	4.6315
12	9.3302	6.9266	5.9525	5.4120	5.0643	4.8206	4.6395	4.4994	4.3875
13	9.0738	6.7010	5.7394	5.2053	4.8616	4.6204	4.4410	4.3021	4.1911
14	8.8616	6.5149	5.5639	5.0354	4.6950	4.4558	4.2779	4.1399	4.0297
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948
16	8.5310	6.2262	5.2922	4.7726	4.4374	4.2016	4.0259	3.8896	3.7804
17	8.3997	6.1121	5.1850	4.6690	4.3359	4.1015	3.9267	3.7910	3.6822
18	8.2854	6.0129	5.0919	4.5790	4.2479	4.0146	3.8406	3.7054	3.5971
19	8.1849	5.9259	5.0103	4.5003	4.1708	3.9386	3.7653	3.6305	3.5225
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567
21	8.0166	5.7804	4.8740	4.3688	4.0421	3.8117	3.6396	3.5056	3.3981

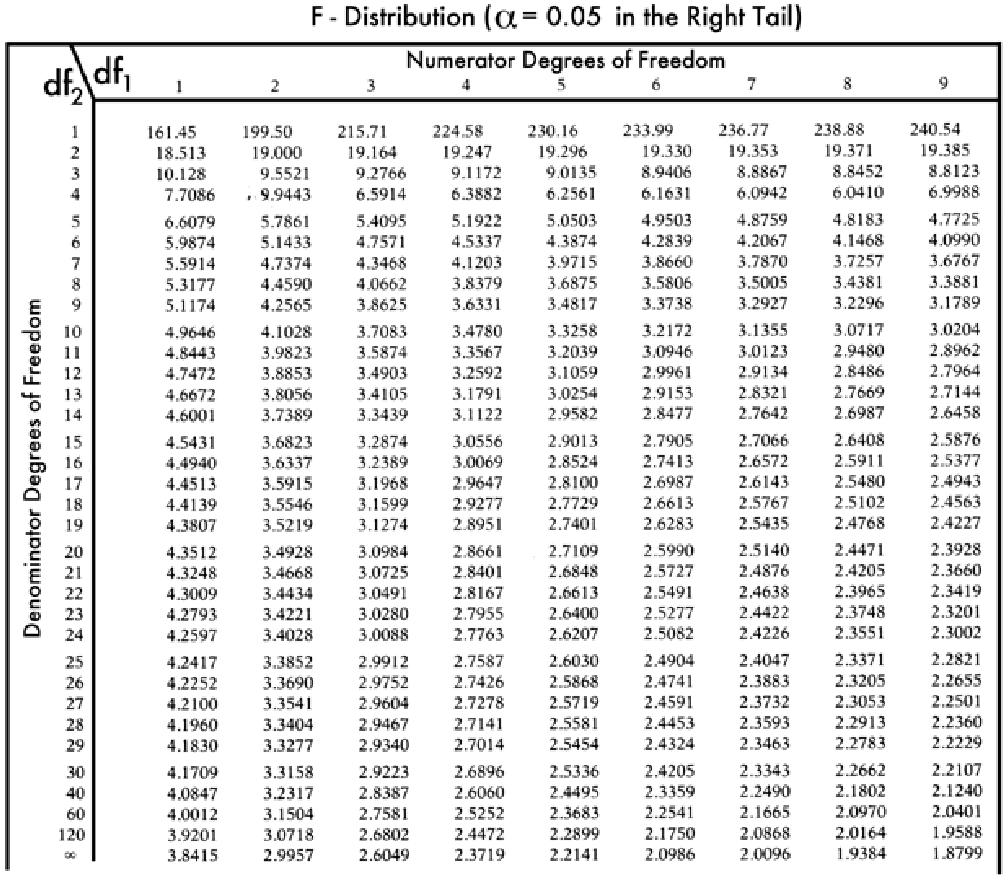


- ❖ Numerator df = 4
- ❖ Denominator df = 7

- ❖ α = 0.10 Two Tail
- ❖ F 0.05, 4, 7 = 4.1203

Two Variances Test

Critical Test Statistic

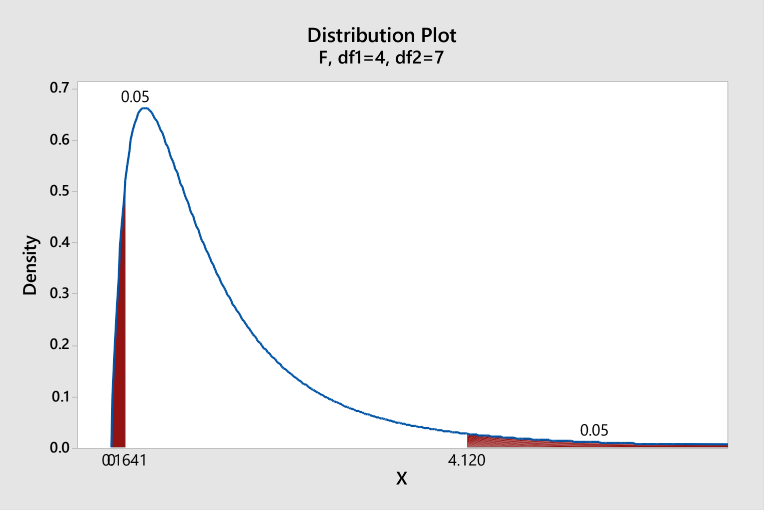


F - Distribution (α = 0.05 in the Right Tail)

Numerator Degrees of Freedom

Denominator Degrees of Freedom

df₂ \ df₁	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799



Distribution Plot

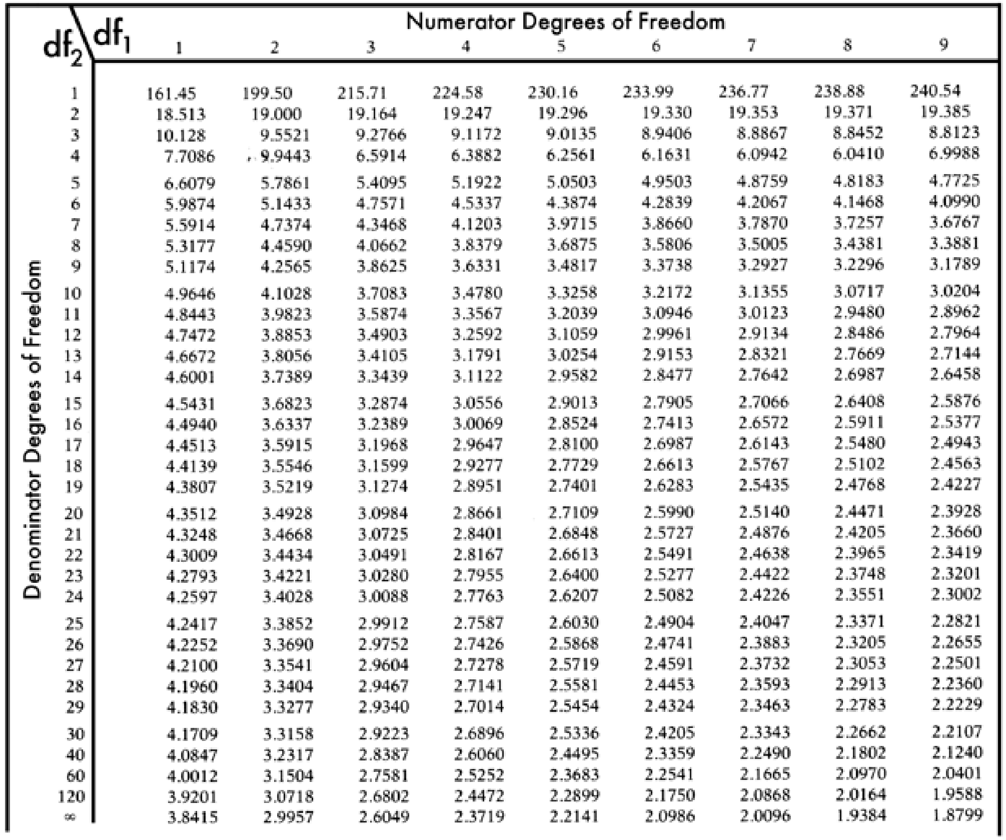
F, df1=4, df2=7

- ❖ Numerator df = 4
- ❖ Denominator df = 7

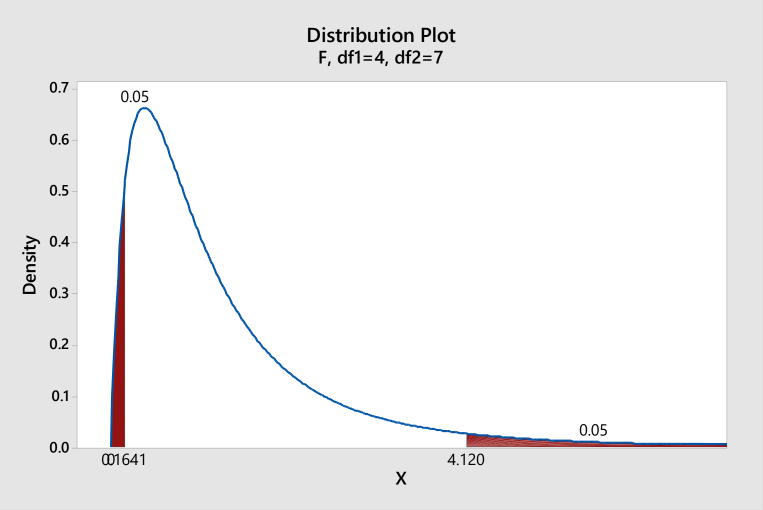
- ❖ α = 0.10 Two Tail
- ❖ F 0.05, 4, 7 = 4.1203
- ❖ F 0.95, 4, 7 = ?

Critical Test Statistic

F - Distribution (α = 0.05 in the Right Tail)



Denominator Degrees of Freedom	Numerator Degrees of Freedom								
df_2 \ df_1	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799



- Numerator df = 4
- Denominator df = 7

- α = 0.10 Two Tail
- $F_{0.05, 4, 7}$ = 4.1203

$$F_{0.95, 4, 7} = 1/\ F_{0.05, 7, 4}$$

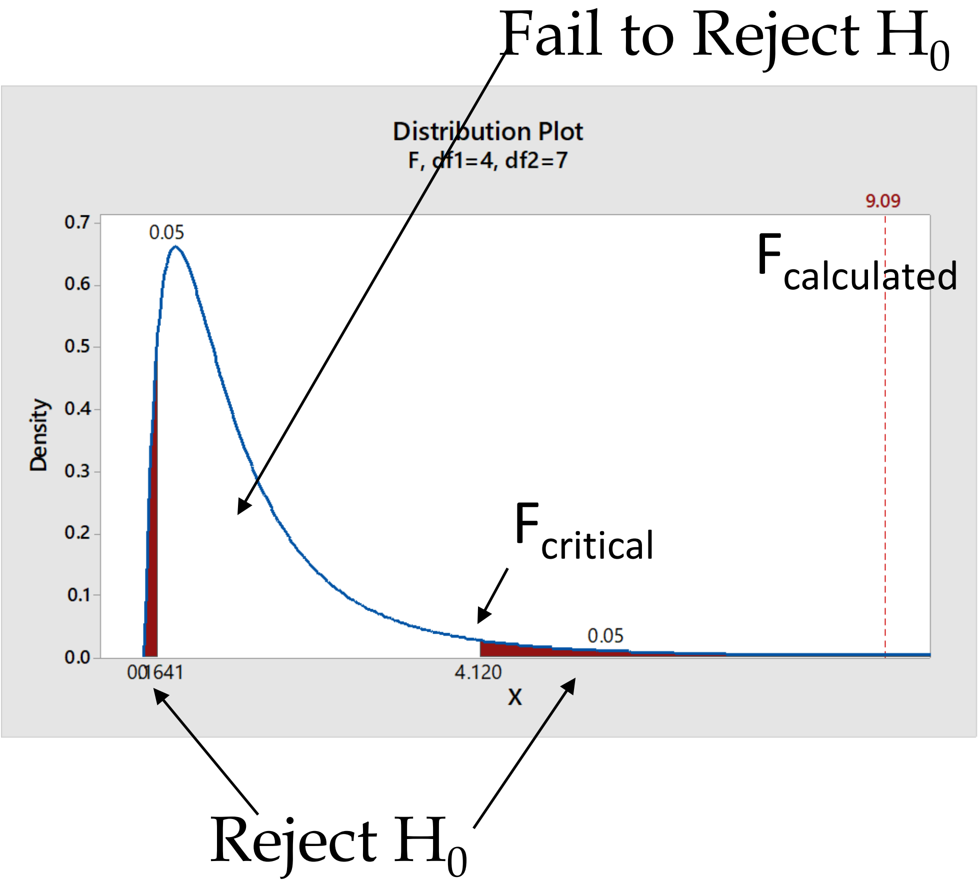
$$F_{0.95, 4, 7} = 1/\ 6.0942 = 0.164$$

Two Variances Test

Two Variances Test

$H_0: \sigma^2_1 = \sigma^2_2$

$Ha: \sigma^2_1 \neq \sigma^2_2$



Fail to Reject H_0

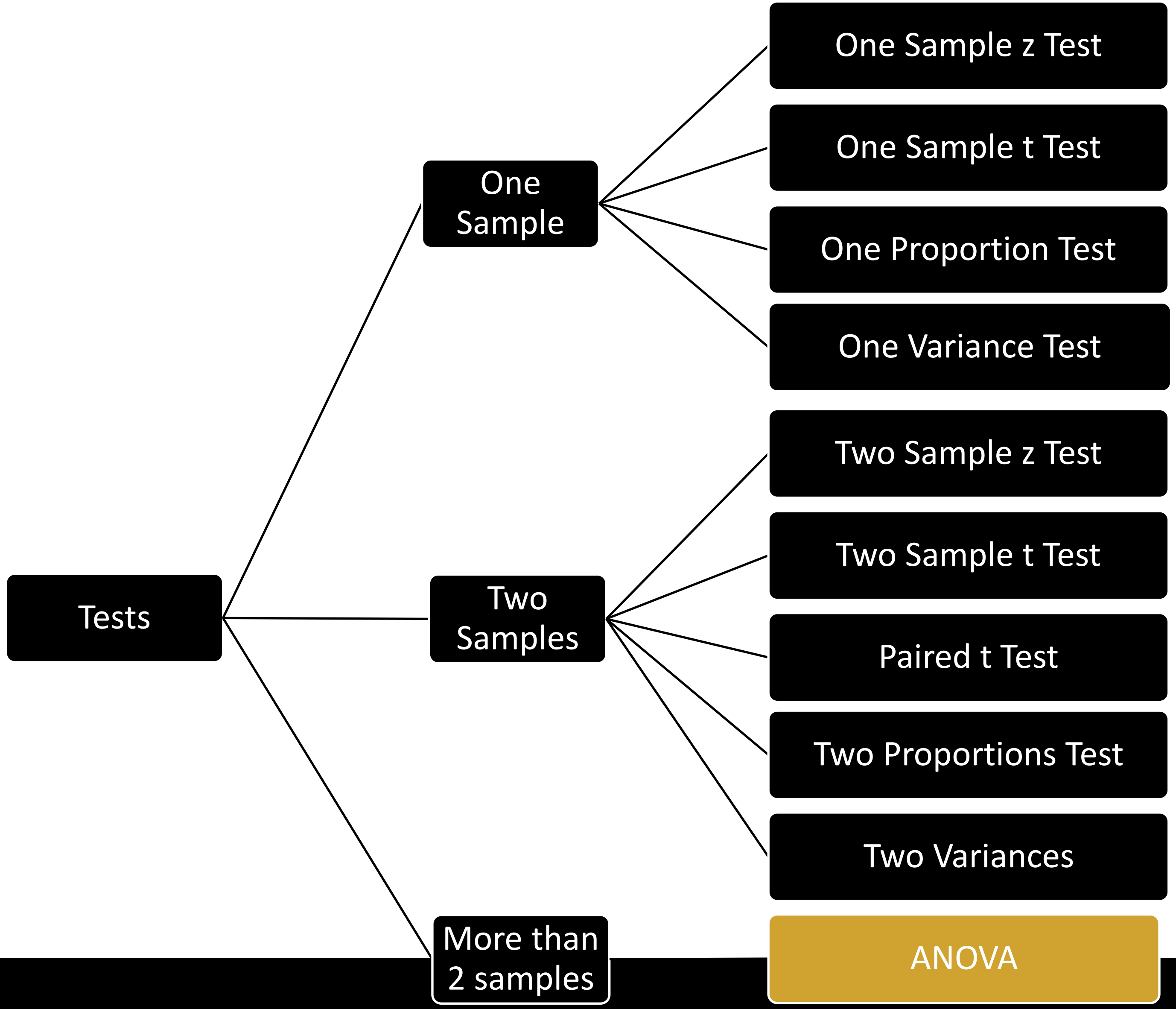
$F_{calculated}$

$F_{critical}$

Reject H_0

- Example: We took 8 samples from machine A and the **standard deviation** was 1.1. For machine B we took 5 samples and the **variance** was 11. Is there a difference in variance at 90% confidence level?
- $n1 = 5, s^2_1 = 11, df_1 = 4$ (numerator)
- $n2 = 8, s_2 = 1.1, s^2_2 = 1.21, df_2 = 7$ (denominator)
- $F_{\ calculated} = 11/1.21 = 9.09$ **(higher value at top)**
- $F_{\ critical} = 0.0164$ and 4.120

Two Variances Test



- Tests
  - One Sample
    - One Sample z Test
    - One Sample t Test
    - One Proportion Test
    - One Variance Test
  - Two Samples
    - Two Sample z Test
    - Two Sample t Test
    - Paired t Test
    - Two Proportions Test
    - Two Variances
  - More than 2 samples
    - ANOVA

ANOVA

Variance Tests

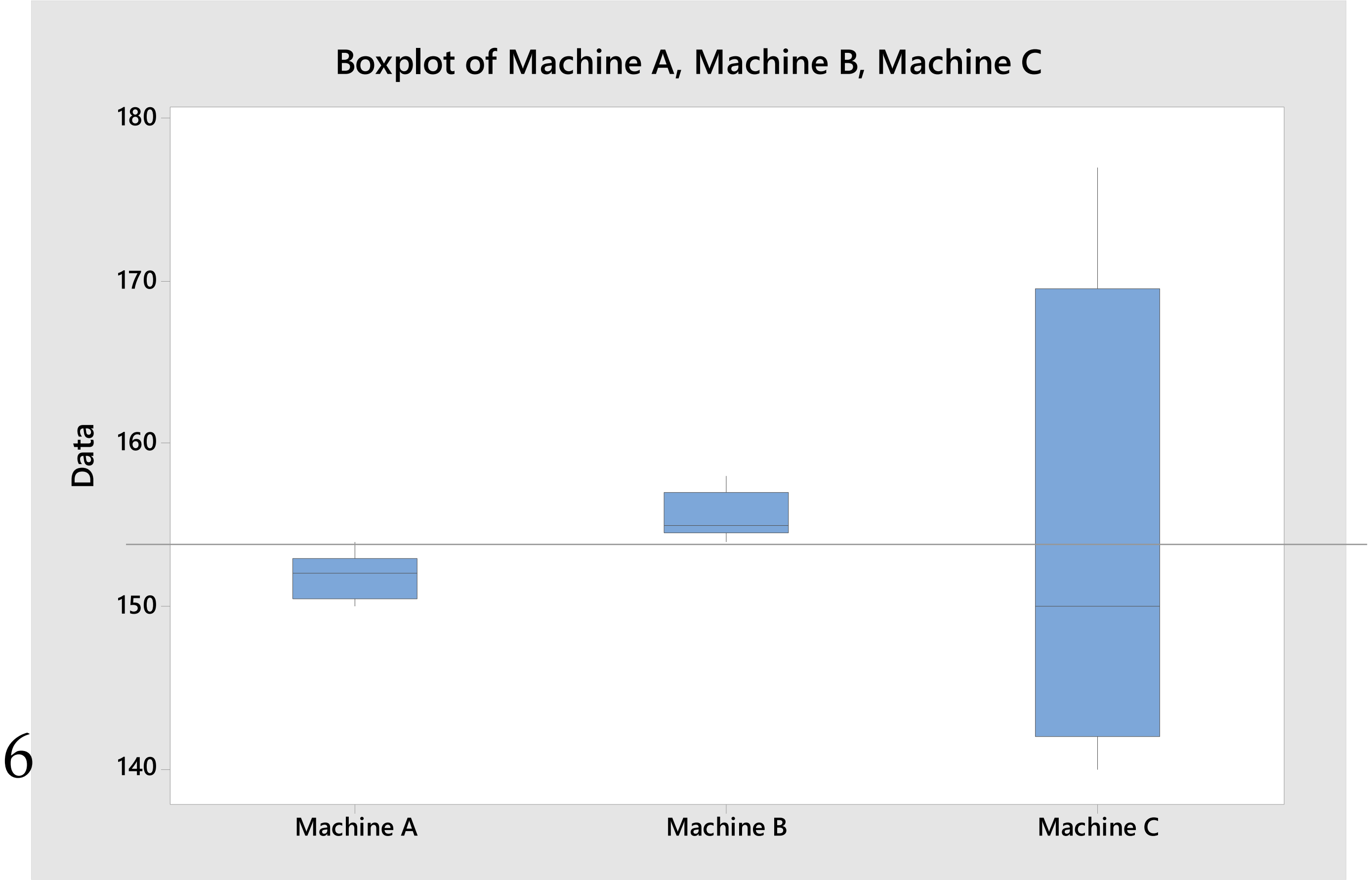
- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Two Sample t Tests

Machine A	Machine B	Machine C
150	156	144
152	155	162
154	158	177
152	155	150
151	154	140

$\bar{x}_{\mathbf{A}} = 151.8$, $\bar{x}_{\mathrm{B}} = 155.6$ $\bar{x}_{\mathrm{C}} = 154.6$



ANOVA

T Test vs ANOVA

2 Sample T Test

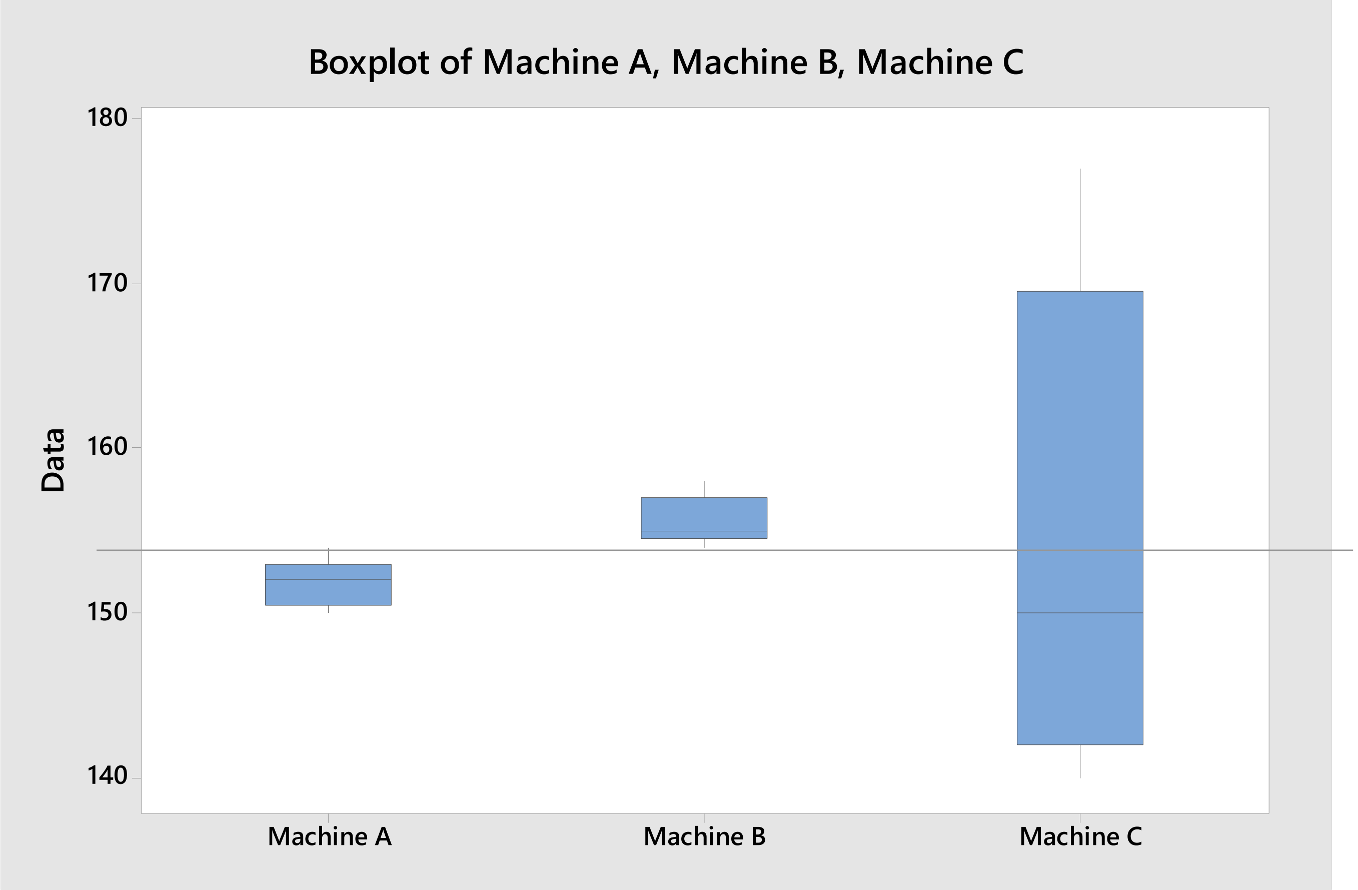
$H_0: \mu_A = \mu_B$

$H_a: \mu_A \neq \mu_B$

ANOVA

$H_0: \mu_A = \mu_B = \mu_C = \mu_D \ldots\ldots = \mu_k$

H_a: At least one of the means is different from others



ANOVA

T Test vs ANOVA

T Test

$H_0: \mu_A = \mu_B$

$H_a: \mu_A \neq \mu_B$

ANOVA

$H_0: \mu_A = \mu_B = \mu_C = \mu_{D\,.....} = \mu_k$

H_a: At least one of the means is different from others

- **Why ANOVA?**
  - We used t test to compare the means of two populations.
  - What if we need to compare more than two populations? With ANOVA we can find out if one or more populations have different mean or comes from a different population.
  - We could have conducted multiple t Test.
  - How many t Test we need to conduct if have to compare 4 sample means? … 6

ANOVA

T Test vs ANOVA

4				x
3			x	3 vs 4
2		x	2 vs 3	2 vs 4
1	x	1 vs 2	1 vs 3	1 vs 4
	1	2	3	4

- **Why ANOVA?**
  - We used t test to compare the means of two populations.
  - What if we need to compare more than two populations? With ANOVA we can find out if one or more populations have different mean or comes from a different population.
  - We could have conducted multiple t Test.
  - How many t Test we need to conduct if have to compare 4 sample means? … 6

T Test vs ANOVA

4				x
3			x	3 vs 4
2		x	2 vs 3	2 vs 4
1	x	1 vs 2	1 vs 3	1 vs 4
	1	2	3	4

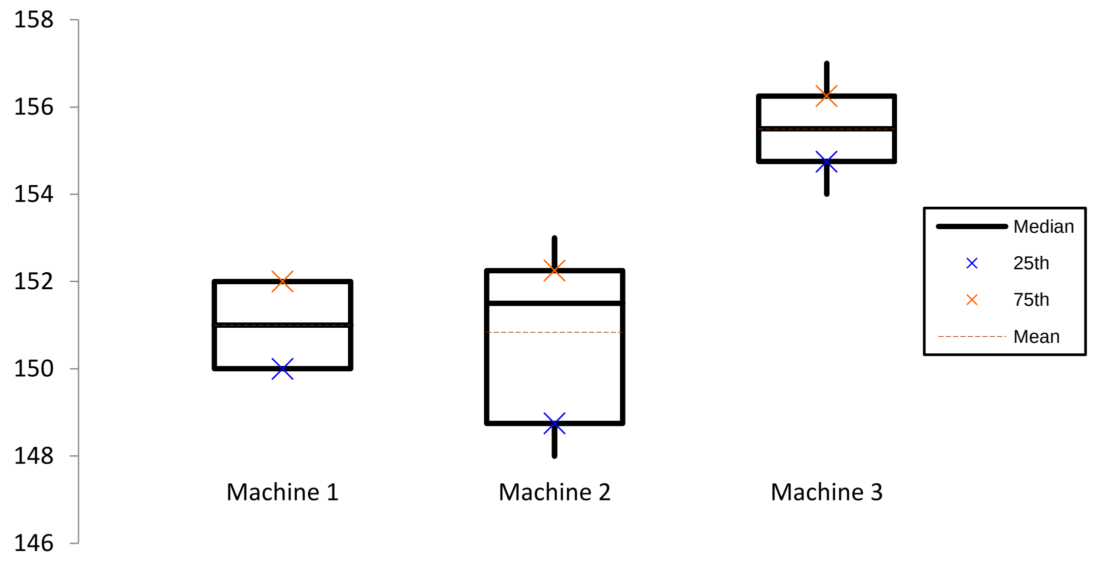
- Why ANOVA?
  - How many t Test we need to conduct if have to compare 4 samples? … 6
  - Each test is done with alpha = 0.05 or 95% confidence.
  - 6 tests will result in confidence level of 0.95x0.95x0.95x0.95x0.95x0.95 = 0.735

ANOVA

Machine 1	**Machine 2**	**Machine 3**
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1$ = 151	**$\bar{x}_2$ = 150.83**	**$\bar{x}_3$ = 155.50**

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

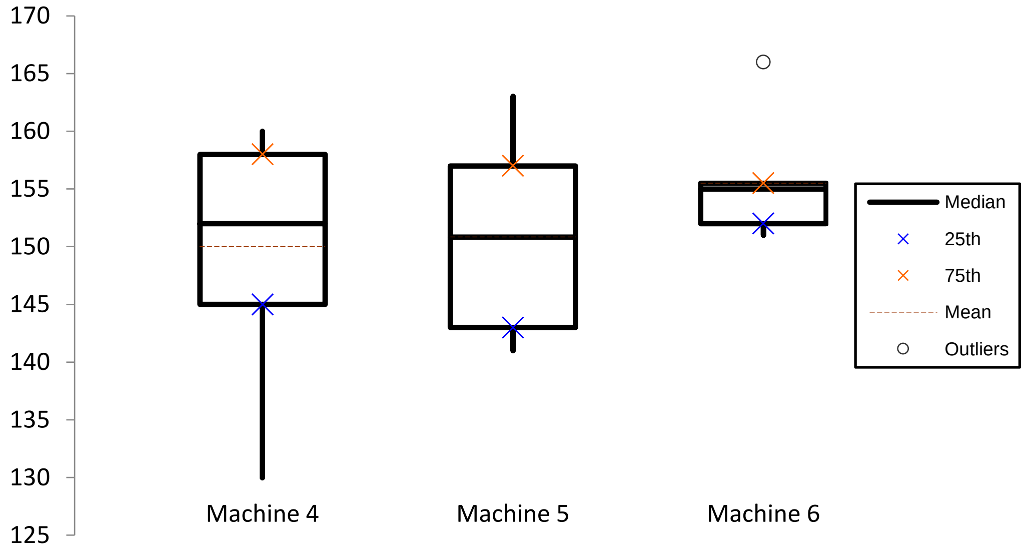


ANOVA

Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4$ = 151.00	**$\bar{x}_5$ = 150.83**	**$\bar{x}_6$ = 155.50**

ANOVA

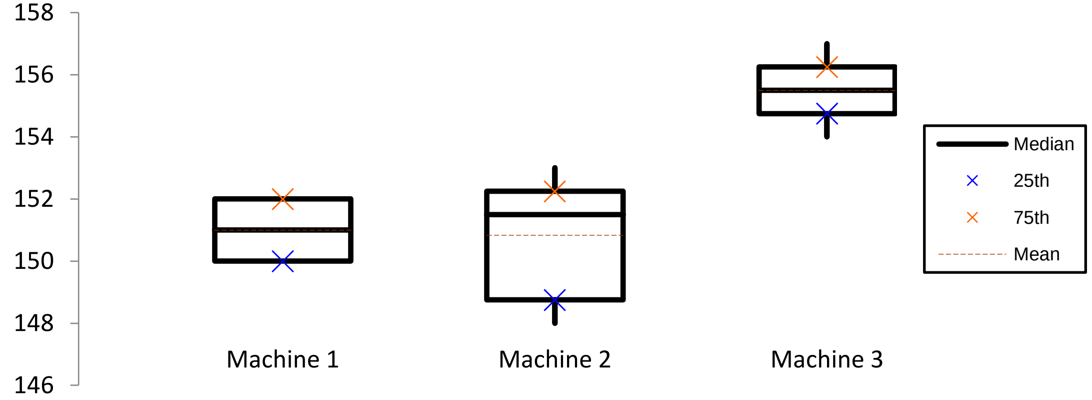
Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4$ = 151.00	$\bar{x}_5$ = 150.83	$\bar{x}_6$ = 155.50

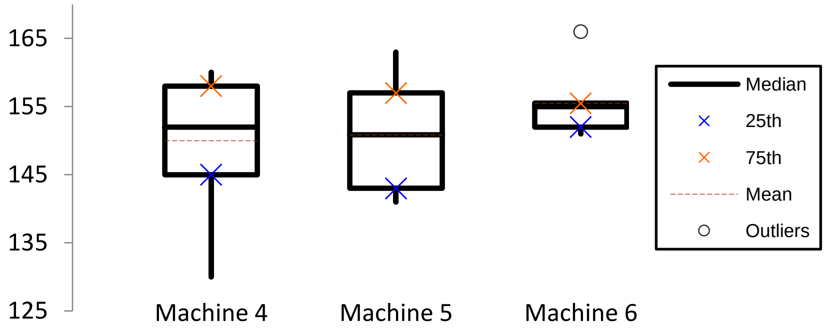


ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1$ = 151.00	**$\bar{x}_2$ = 150.83**	**$\bar{x}_3$ = 155.50**

Machine 4	Machine 5	Machine 6
130	163	166
155	152	154
160	143	155
158	141	151
152	149	152
145	157	155
$\bar{x}_4$ = 151.00	**$\bar{x}_5$ = 150.83**	**$\bar{x}_6$ = 155.50**





ANOVA

- ANOVA is Analysis of Variance
- Variance

$$s^2 = \frac{\sum(x_i - \overline{X})^2}{n-1}$$

- Numerator of this formula is Sum of Squares, the denominator is the degrees of freedom for the sample.

ANOVA

ANOVA

$$F = \frac{s_1^2}{s_2^2}$$

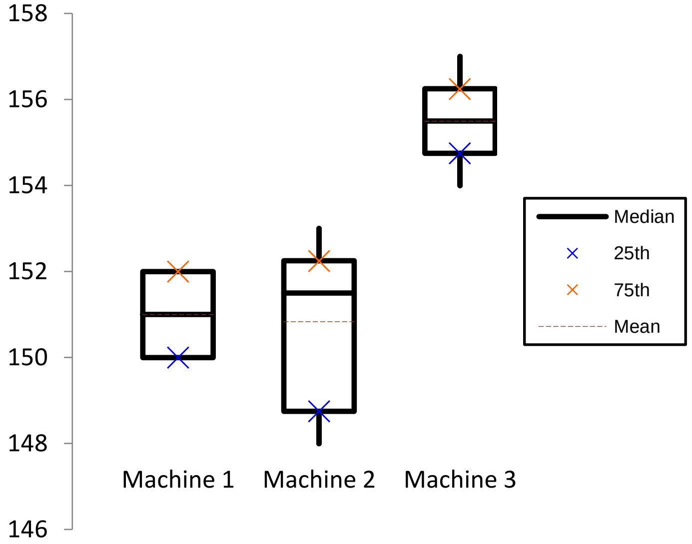
$$F = \frac{\dfrac{\sum (x - \bar{x}_1)^2}{n_1 - 1}}{\dfrac{\sum (x - \bar{x}_2)^2}{n_2 - 1}}$$

$$F = \frac{\dfrac{SS_1}{\mathrm{d}f_1}}{\dfrac{SS_2}{\mathrm{d}f_2}}$$

$$F = \frac{MSS_1}{MSS_2}$$

$$F = \frac{MSS_{between}}{MSS_{within}}$$

$$F = \frac{\dfrac{SS_{\mathrm{between}}}{\mathrm{d}f_{\mathrm{between}}}}{\dfrac{SS_{within}}{\mathrm{d}f_{\mathrm{within}}}}$$



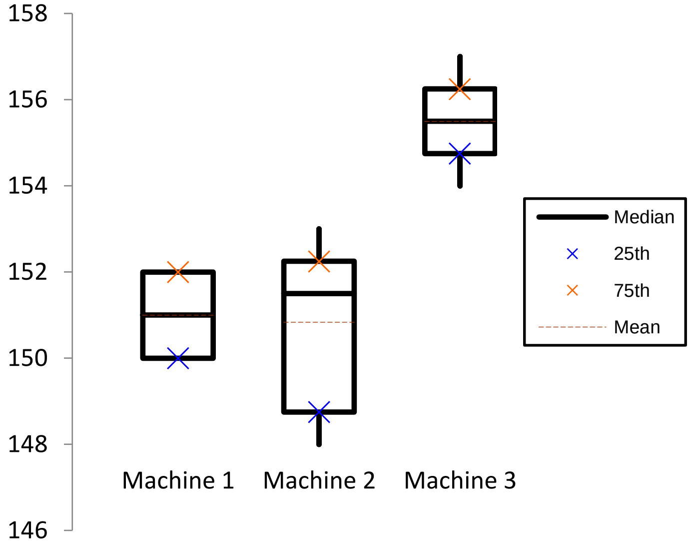
ANOVA

ANOVA

$$F = \frac{\frac{SS_{\text{between}}}{\mathrm{d}f_{\text{between}}}}{\frac{\mathrm{SS}_{within}}{\mathrm{d}f_{\text{within}}}}$$

$$F = \frac{MSS_{between}}{MSS_{within}}$$

❖ SST = SS $_{\text{between(or treatment, or column)}}$ + SS $_{\text{within(or error)}}$



ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1$ = 151	**$\bar{x}_2$ = 150.83**	**$\bar{x}_3$ = 155.50**

- **Null hypothesis**: $H_0: \mu_1 = \mu_2 = \mu_3$
- **Alternative hypothesis**: H_a : Means are not all equal

Check at 95% confidence level.

- $SS_{\text{between(or treatment, or column)}}$
- $SS_{\text{within(or error)}}$

$$F = \frac{\dfrac{SS_{\text{between}}}{\mathrm{d}f_{\text{between}}}}{\dfrac{SS_{within}}{\mathrm{d}f_{\text{within}}}}$$

$$F = \frac{MSS_{between}}{MSS_{within}}$$

ANOVA

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

- $SS_{within} = 4.00+18.83+5.50 = 28.33$

Machine 1	x1 - x̄1	Sqr(x1 - x̄1)	Machine 2	x2 - x̄2	Sqr(x2 - x̄2)	Machine 3	x3 - x̄3	Sqr(x3 - x̄3)	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	

ANOVA

Machine 1	**Machine 2**	**Machine 3**
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1$ = 151	**$\bar{x}_2$ = 150.83**	**$\bar{x}_3$ = 155.50**

❖ SS $_{between}$ = (2.07+2.58+9.36)x6 = 84.06

Machine 1	x1 - x̄1	Sqr(x1 - x̄1)	Machine 2	x2 - x̄2	Sqr(x2 - x̄2)	Machine 3	x3 - x̄3	Sqr(x3 - x̄3)	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	
	-1.44	2.07		-1.61	2.58		3.06	9.36	

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

- $SS_{within} = 4.00+18.83+5.50 = 28.33$
- $SS_{between} = (2.07+2.58+9.36)x6 = 84.06$

Machine 1	x1 - x̄1	Sqr(x1 - x̄1)	Machine 2	x2 - x̄2	Sqr(x2 - x̄2)	Machine 3	x3 - x̄3	Sqr(x3 - x̄3)	
150.00	-1.00	1.00	153.00	2.17	4.69	156.00	0.50	0.25	
151.00	0.00	0.00	152.00	1.17	1.36	154.00	-1.50	2.25	
152.00	1.00	1.00	148.00	-2.83	8.03	155.00	-0.50	0.25	
152.00	1.00	1.00	151.00	0.17	0.03	156.00	0.50	0.25	
151.00	0.00	0.00	149.00	-1.83	3.36	157.00	1.50	2.25	
150.00	-1.00	1.00	152.00	1.17	1.36	155.00	-0.50	0.25	
151.00			150.83			155.50			152.44
		4.00			18.83			5.50	
	-1.44	2.07		-1.61	2.58		3.06	9.36	

ANOVA

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1 = 151$	**$\bar{x}_2 = 150.83$**	**$\bar{x}_3 = 155.50$**

- $SST = SS_{\text{between(or treatment)}} + SS_{\text{within(or error)}}$
- $SST = 84.06 + 28.33 = 112.39$
- **<u>Degrees of freedom</u>**
- Total $df = df_{\text{between(or treatment)}} + df_{\text{within(or error)}}$
- $(N-1) = (C-1) + (N-C)$
- $df_{\text{between}} = 3-1=2,$
- $df_{\text{total}} = 17$, $df_{\text{within}} = 17-2=15$

- $SS_{\text{within}} = 28.33$
- $SS_{\text{between}} = 84.06$

ANOVA

ANOVA

Machine 1	Machine 2	Machine 3
150	153	156
151	152	154
152	148	155
152	151	156
151	149	157
150	152	155
$\bar{x}_1$ = 151	**$\bar{x}_2$ = 150.83**	**$\bar{x}_3$ = 155.50**

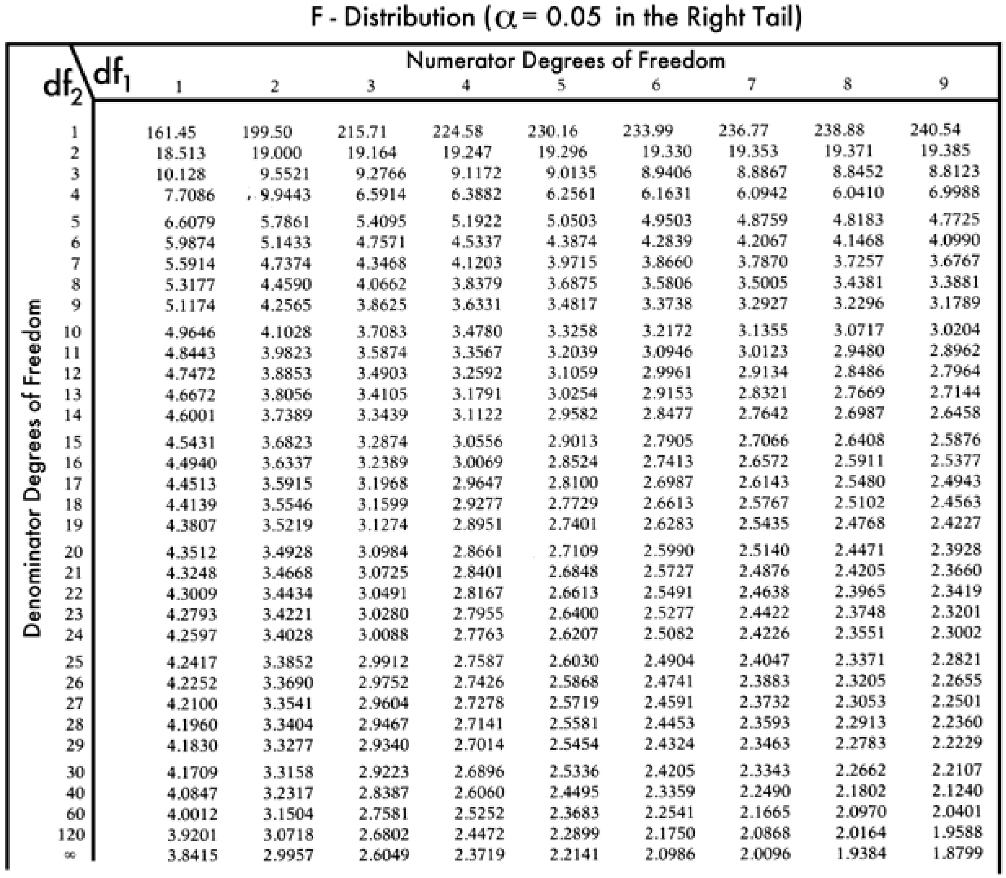
- **Mean Sum of Square = SS / df**
- $MSS_{between} = SS_{between} / df_{between}$
- $MSS_{between} = 84.06 / 2 = 42.03$

- $MSS_{within} = SS_{within} / df_{within}$
- $MSS_{within} = 28.33 / 15 = 1.89$

- $F = MSS_{between} / MSS_{within} = 42.03/1.89 = 22.24$

- $SS_{within} = 28.33$
- $SS_{between} = 84.06$
- $df_{within} = 15$
- $df_{between} = 2$

Critical Test Statistic



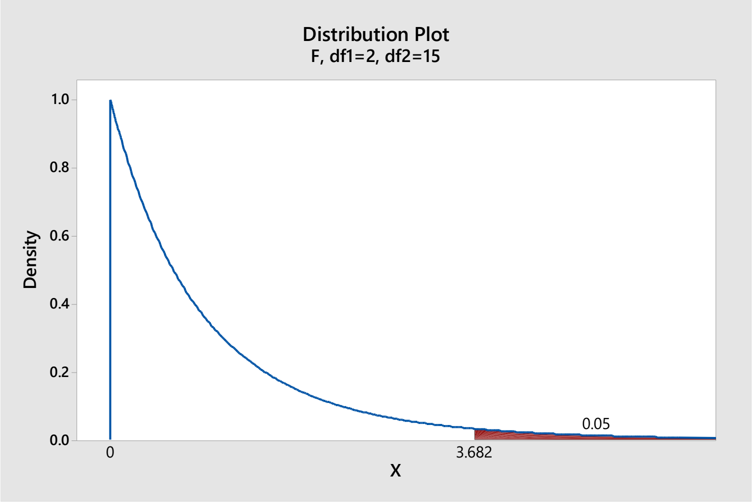
F - Distribution (α = 0.05 in the Right Tail)

df_2 \ df_1	Numerator Degrees of Freedom 1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

Denominator Degrees of Freedom

$$F = \frac{\frac{SS_{\text{between}}}{\mathrm{d}f_{\text{between}}}}{\frac{SS_{within}}{\mathrm{d}f_{\text{within}}}}$$

- ❖ Numerator (between) df = 2
- ❖ Denominator (within) df = 15



Distribution Plot
F, df1=2, df2=15

- ❖ α = 0.05 One Tail
- ❖ $F\ 0.05, 2, 1_5 = 3.68$

ANOVA

- F = $MSS_{between}$ / MSS_{within} = 42.03/1.89 = 22.24

- Compare this with $F_{critical}$
- F (0.05, 2, 15) = 3.68

- Reject Null Hypothesis

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Variance Tests – Chi-Square

Goodness of Fit Test (Chi Square)

- To test if the sample is coming from a population with specific distribution.
- Other goodness-of-fit tests are
  - Anderson-Darling
  - Kolmogorov-Smirnov

Goodness of Fit Test

Goodness of Fit Test (Chi Square)

- H_0: The data follow a specified distribution.
- Ha: The data do not follow the specified distribution.
- Calculated Statistic: $\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$
- Critical Statistic: Chi square for k-1 degrees of freedom for specific alpha.

Goodness of Fit Test

Goodness of Fit Test (Chi Square)

- A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

Head	40
Tail	60

Goodness of Fit Test

- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

Ho: Coin is not biased.
Ha: Coin is biased.
Alpha = 0.05

Head	40
Tail	60

Goodness of Fit – Chi-Square

Ho: Coin is not biased.
Ha: Coin is biased.
Alpha = 0.05

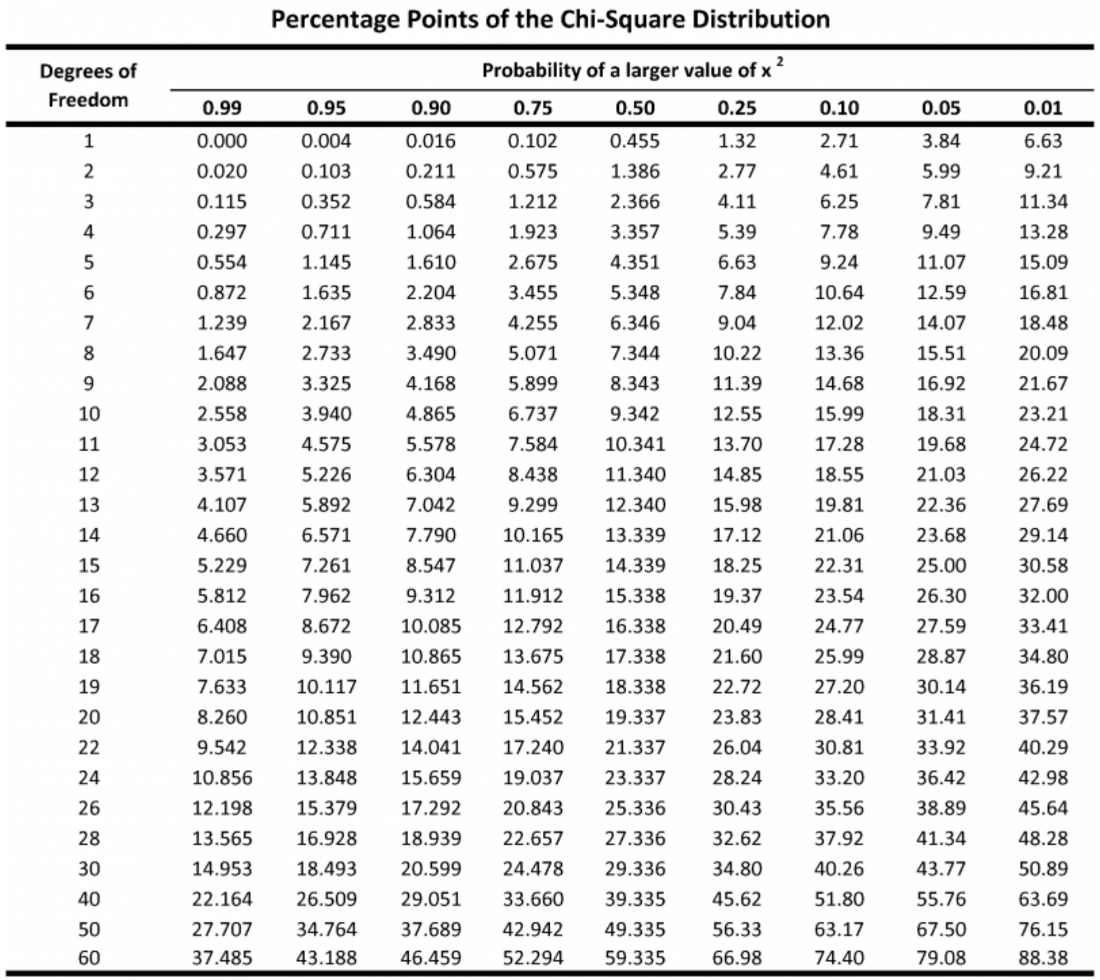
- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Flip	Expected	Observed	O-E	(O-E)^2	(O-E)^2/E
Head	50	40	-10	100	2
Tail	50	60	10	100	2
					X^2 = 4

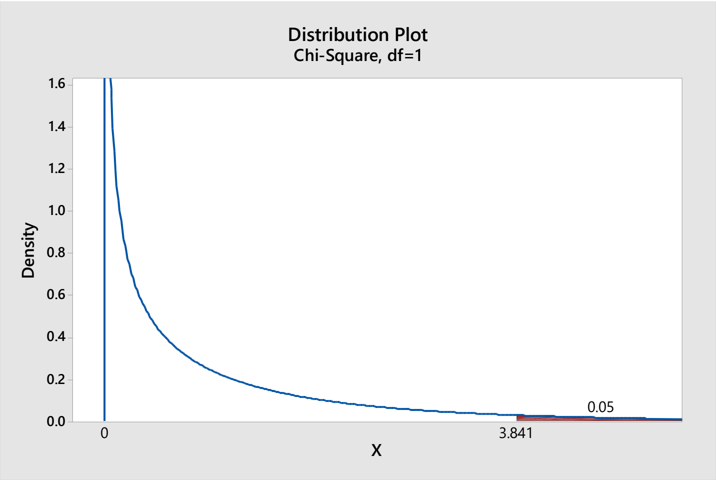
Goodness of Fit – Chi-Square

Critical Test Statistic



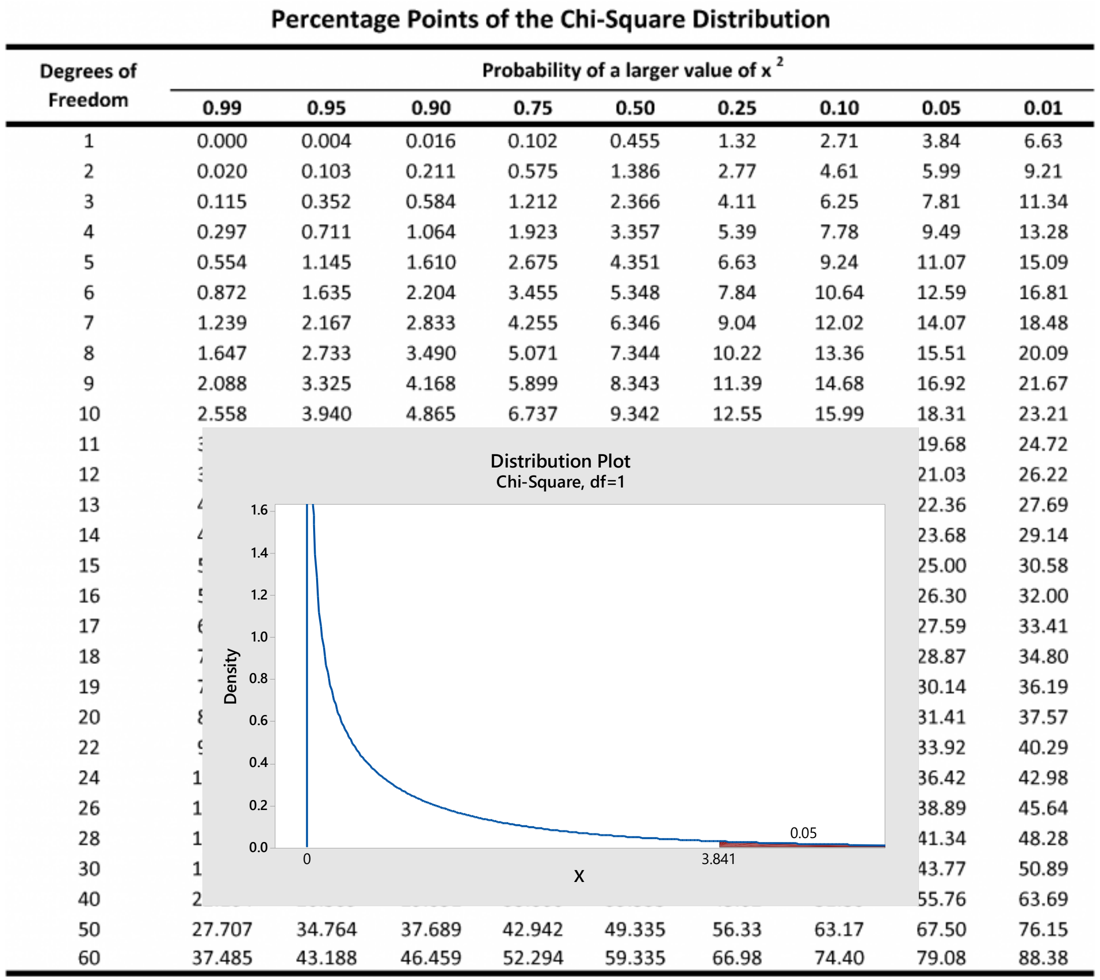
Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x ^{2}								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



- α = 0.05 One Tail
- Df = 1
- χ^2 Critical = 3.84

Critical Test Statistic



Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11								19.68	24.72
12								21.03	26.22
13								22.36	27.69
14								23.68	29.14
15								25.00	30.58
16								26.30	32.00
17								27.59	33.41
18								28.87	34.80
19								30.14	36.19
20								31.41	37.57
22								33.92	40.29
24								36.42	42.98
26								38.89	45.64
28								41.34	48.28
30								43.77	50.89
40								55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

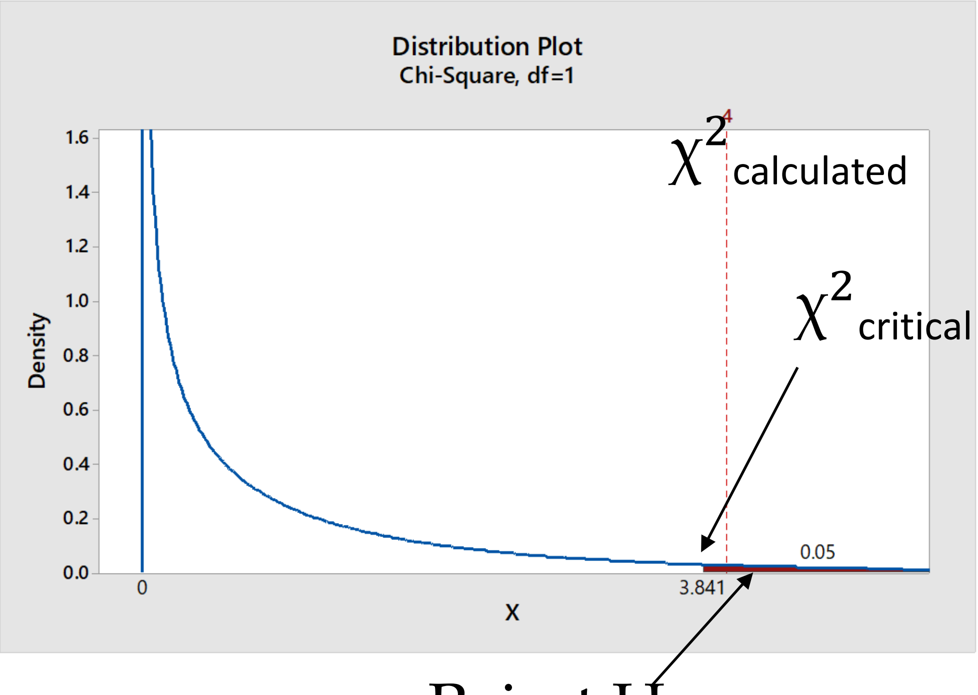
Head	40
Tail	60

- $\chi^2(calculated) = 4.0$

- $\chi^2(\text{critical}) = 3.84$

Goodness of Fit – Chi-Square

Ho: Coin is not biased.
Ha: Coin is biased.
Alpha = 0.05



χ^2 calculated

χ^2 critical

Reject H_0

- Example 1: A coin is flipped 100 times. Number of heads and tails are noted. Is this coin biased? Check with 95% Confidence Level.

Head	40
Tail	60

- $\chi^2(calculated) = 4.0$
- $\chi^2(\text{critical}) = 3.84$

Goodness of Fit – Chi-Square

H0: The data follow a specified distribution.
Ha: The data do not follow the specified distribution.
Alpha = 0.05

- Example 2: A t-shirt manufacturer expects vs actual sale.

Size	**Proportions**	**Counts**
Small	0.1	25
Medium	0.2	41
Large	0.4	91
Extra Large	0.3	68

Goodness of Fit – Chi-Square

- Example 2: A t-shirt manufacturer expects vs actual sale.

H0: The data follow a specified distribution.
Ha: The data do not follow the specified distribution.
Alpha = 0.05

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Size	**Proportions**	**Expected**	**Observed**
Small	0.1	22.5	25
Medium	0.2	45	41
Large	0.4	90	91
Extra Large	0.3	67.5	68

Goodness of Fit – Chi-Square

Variance Tests

- Chi-square test
  - For testing the population variance against a specified value
  - testing goodness of fit of some probability distribution
  - testing for independence of two attributes (Contingency Tables)

- F-test
  - for testing equality of *two* variances from different populations
  - for testing equality of several means with technique of ANOVA.

Variance Tests – Chi-Square

Contingency Tables

- To find relationship between two discrete variables.

	Smoker	Non Smoker	
Male	60	40	100
Female	35	40	75
	95	80	175

	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

Contingency Tables

Contingency Tables

	Smoker	Non Smoker	
Male	60	40	100
Female	35	40	75
	95	80	175

	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

- Null hypothesis is that there is no relationship between the row and column variables.
- Alternate hypothesis is that there is a relationship. Alternate hypothesis does not tell what type of relationship exists.

Calculating Test Statistic (Chi Square)

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

Calculating Test Statistic (Chi Square)

OBSERVED	**Operator 1**	**Operator 2**	**Operator 3**	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

EXPECTED	**Operator 1**	**Operator 2**	**Operator 3**	
Shift 1	122x71/347	110x71/347	115x71/347	71
Shift 2	122x116/347	110x116/347	115x116/347	116
Shift 3	122x160/347	110x160/347	115x160/347	160
	122	110	115	347

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Calculating Test Statistic (Chi Square)

OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	112	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	122x71/347	110x71/347	115x71/347	71
Shift 2	122x116/347	110x116/347	115x116/347	116
Shift 3	122x160/347	110x160/347	115x160/347	160
	122	110	115	347

EXPECTED	Operator 1	Operator 2	Operator 3	
Shift 1	24.96	22.51	23.53	71
Shift 2	40.78	36.77	38.44	116
Shift 3	56.25	50.72	53.02	160
	122	112	115	347

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Calculating Test Statistic (Chi Square)

OBSERVED	**Operator 1**	**Operator 2**	**Operator 3**	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	110	115	347

EXPECTED	**Operator 1**	**Operator 2**	**Operator 3**	
Shift 1	122x71/347	110x71/347	115x71/347	71
Shift 2	122x116/347	110x116/347	115x116/347	116
Shift 3	122x160/347	110x160/347	115x160/347	160
	122	110	115	347

$(O-E)^2/E$	**Operator 1**	**Operator 2**	**Operator 3**	
Shift 1	$(22-24.96)^2/24.96 = 0.35$	0.54	0.01	71
Shift 2	$(28-40.78)^2/40.78 = 4.00$	17.31	4.03	116
Shift 3	$(72-56.25)^2/56.25 = 4.41$	16.26	3.18	160
	122	112	115	347

EXPECTED	**Operator 1**	**Operator 2**	**Operator 3**	
Shift 1	24.96	22.51	23.53	71
Shift 2	40.78	36.77	38.44	116
Shift 3	56.25	50.72	53.02	160
	122	112	115	347

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

Contingency Tables

Calculating Test Statistic (Chi Square)

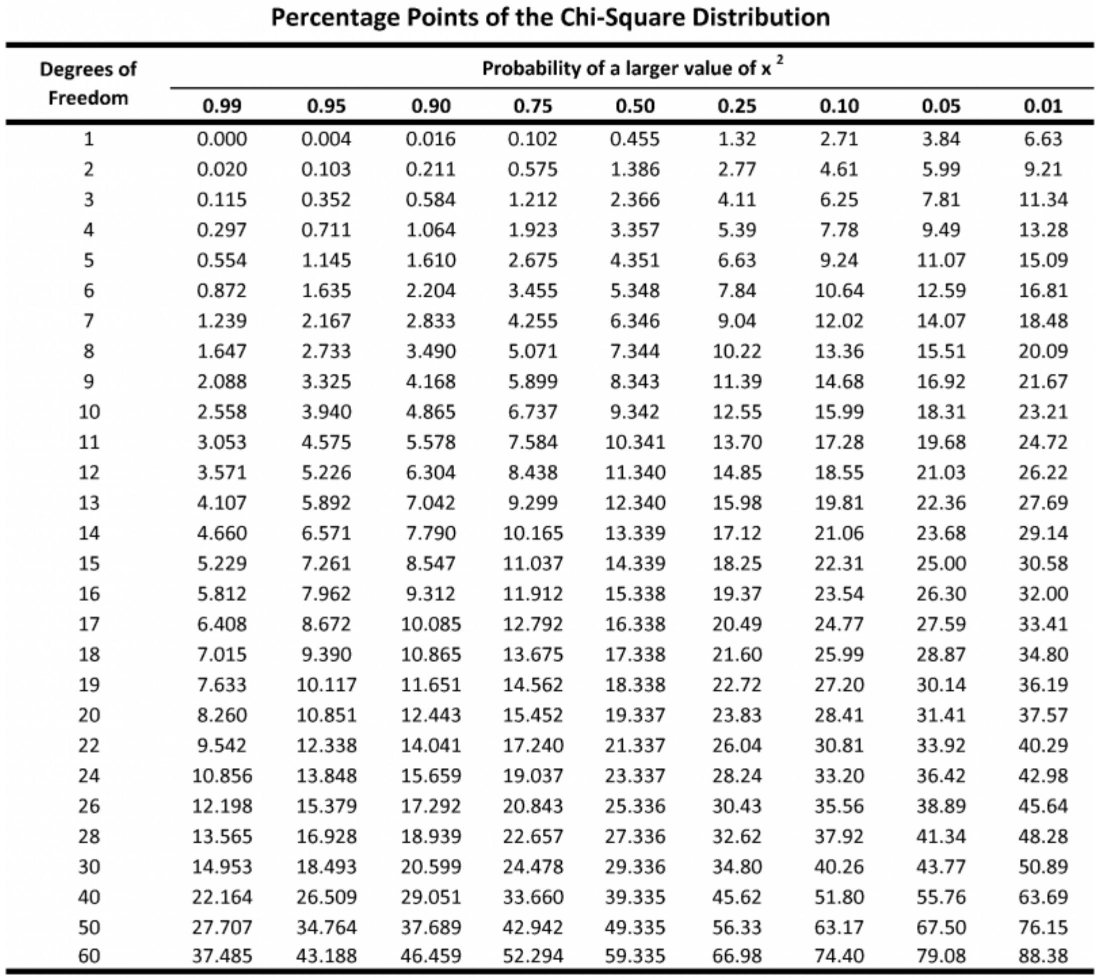
$(O\text{-}E)^2/E$	Operator 1	Operator 2	Operator 3	
Shift 1	$(22-24.96)^2/24.96 = 0.35$	0.54	0.01	71
Shift 2	$(28-40.78)^2/40.78 = 4.00$	17.31	4.03	116
Shift 3	$(72-56.25)^2/56.25 = 4.41$	16.26	3.18	160
	122	112	115	347

$$X^2 = 50.09$$

$$\chi^2 = \sum_{i=1}^{k} \frac{(O-E)^2}{E}$$

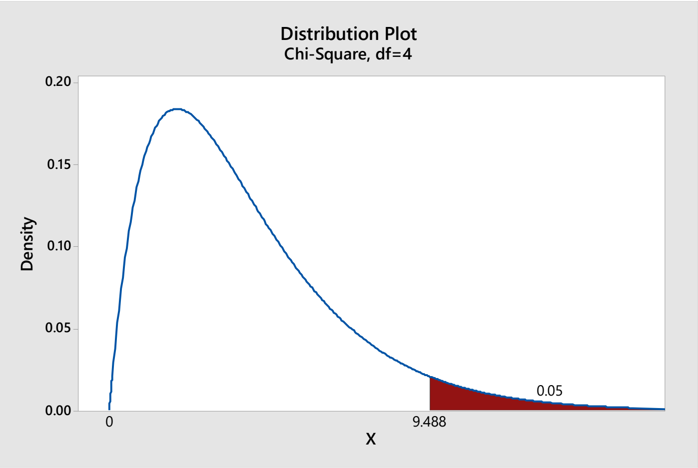
Contingency Tables

Critical Test Statistic



Percentage Points of the Chi-Square Distribution

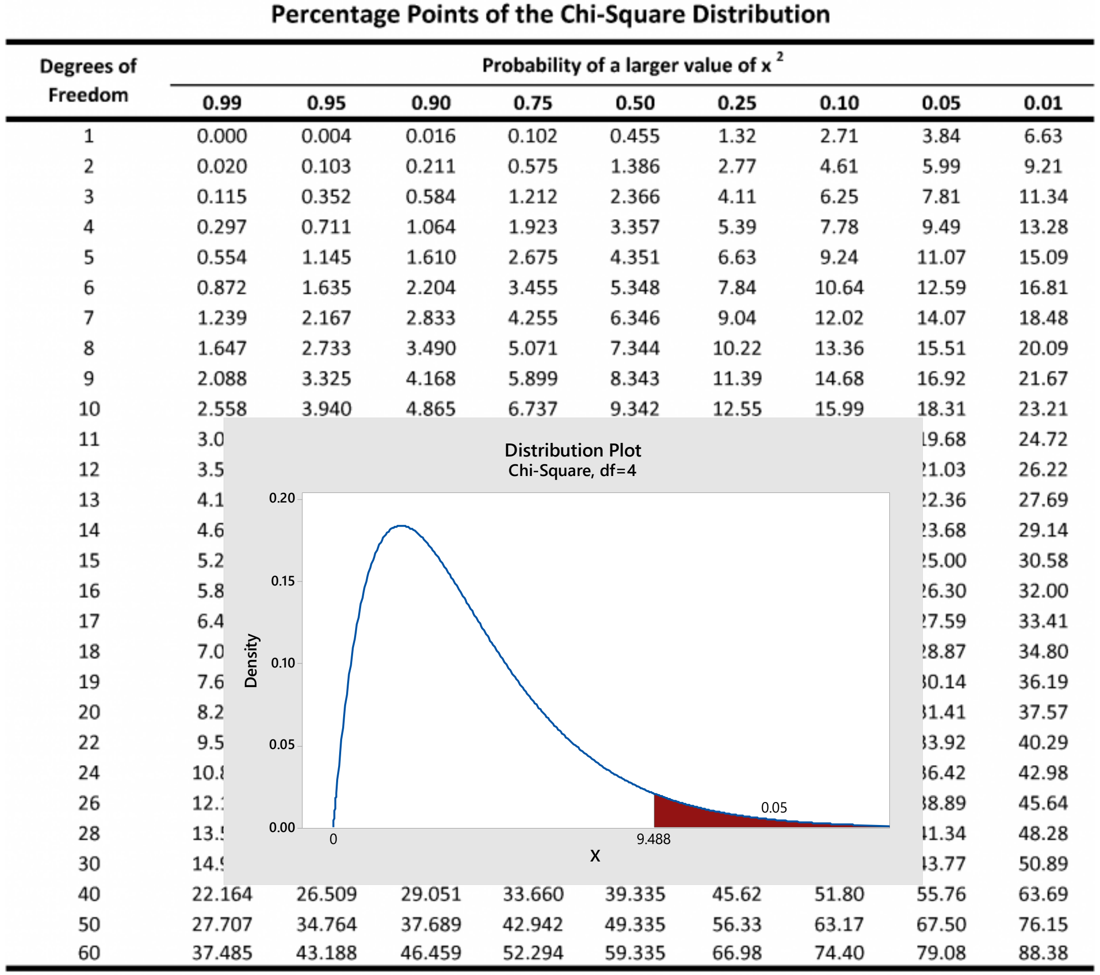
Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.053	4.575	5.578	7.584	10.341	13.70	17.28	19.68	24.72
12	3.571	5.226	6.304	8.438	11.340	14.85	18.55	21.03	26.22
13	4.107	5.892	7.042	9.299	12.340	15.98	19.81	22.36	27.69
14	4.660	6.571	7.790	10.165	13.339	17.12	21.06	23.68	29.14
15	5.229	7.261	8.547	11.037	14.339	18.25	22.31	25.00	30.58
16	5.812	7.962	9.312	11.912	15.338	19.37	23.54	26.30	32.00
17	6.408	8.672	10.085	12.792	16.338	20.49	24.77	27.59	33.41
18	7.015	9.390	10.865	13.675	17.338	21.60	25.99	28.87	34.80
19	7.633	10.117	11.651	14.562	18.338	22.72	27.20	30.14	36.19
20	8.260	10.851	12.443	15.452	19.337	23.83	28.41	31.41	37.57
22	9.542	12.338	14.041	17.240	21.337	26.04	30.81	33.92	40.29
24	10.856	13.848	15.659	19.037	23.337	28.24	33.20	36.42	42.98
26	12.198	15.379	17.292	20.843	25.336	30.43	35.56	38.89	45.64
28	13.565	16.928	18.939	22.657	27.336	32.62	37.92	41.34	48.28
30	14.953	18.493	20.599	24.478	29.336	34.80	40.26	43.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38



- Df = (r-1)(c-1) = 4
- α = 0.05 One Tail
- χ^2 Critical = 9.49

Contingency Tables

Critical Test Statistic



Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x ²								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21
11	3.0							9.68	24.72
12	3.5							1.03	26.22
13	4.1							2.36	27.69
14	4.6							3.68	29.14
15	5.2							5.00	30.58
16	5.8							6.30	32.00
17	6.4							7.59	33.41
18	7.0							8.87	34.80
19	7.6							0.14	36.19
20	8.2							1.41	37.57
22	9.5							3.92	40.29
24	10.							6.42	42.98
26	12.							8.89	45.64
28	13.							1.34	48.28
30	14.							3.77	50.89
40	22.164	26.509	29.051	33.660	39.335	45.62	51.80	55.76	63.69
50	27.707	34.764	37.689	42.942	49.335	56.33	63.17	67.50	76.15
60	37.485	43.188	46.459	52.294	59.335	66.98	74.40	79.08	88.38

OBSERVED	Operator 1	Operator 2	Operator 3	
Shift 1	22	26	23	71
Shift 2	28	62	26	116
Shift 3	72	22	66	160
	122	112	115	347

- $\chi^2(calculated) = 50.09$
- $\chi^2(\text{critical}) = 9.49$