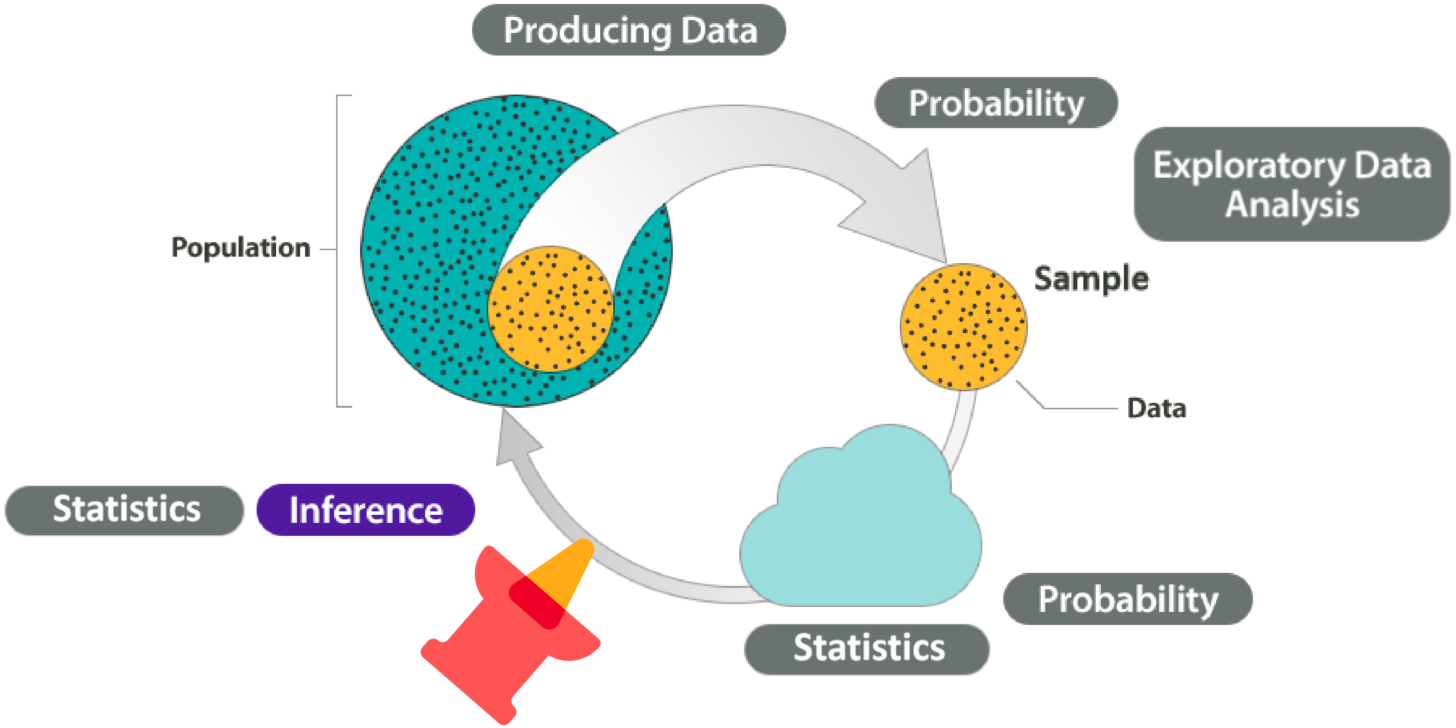
CDS 533
Statistics for Data Science

Instructor: Lisha Yu
Division of Artificial Intelligence
School of Data Science
Lingnan University

Fall 2024

Big Picture of Statistics



Statistical Inference (Hypothesis Testing)

Inferential Statistics

Inferential statistics

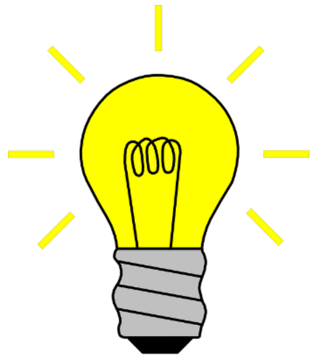
- Draw conclusion from data
- Sample
  - Describe data
- Use **sample statistic** to infer **population parameter**
  - Estimation
  - Hypothesis testing

Real World Problem

Suppose you are a quality analyst at a bulb manufacturing company and analyze the reliability of bulbs. Historically, 70% of the bulbs pass the reliability test.

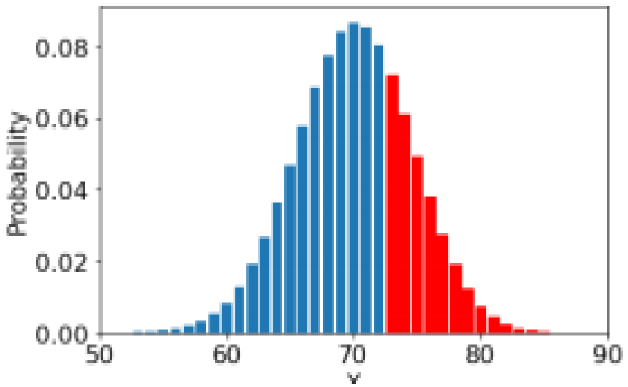
Now, a slightly altered manufacturing process(B) has been introduced to produce the bulbs.

Can you conclude whether the new process improves the reliability of the bulbs?

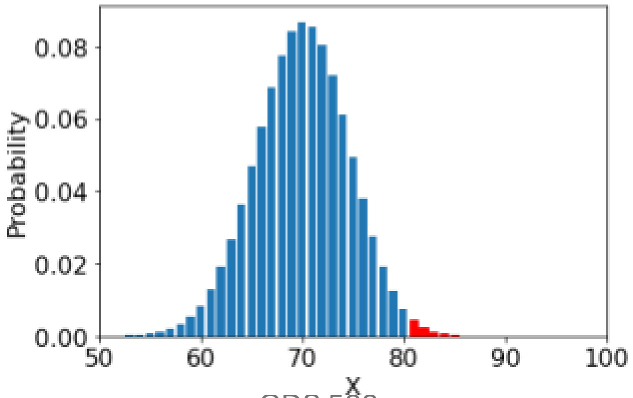


Gathering Evidence for Statistical Inference

We selected a random sample of 100 bulbs out of which 73 are reliable. Does this provide strong evidence that the new manufacturing process is more reliable?



A similar experiment was run with yet another manufacturing process (C). A sample of 100 bulbs produced using this process had 81 reliable bulbs.



Why Hypothesis

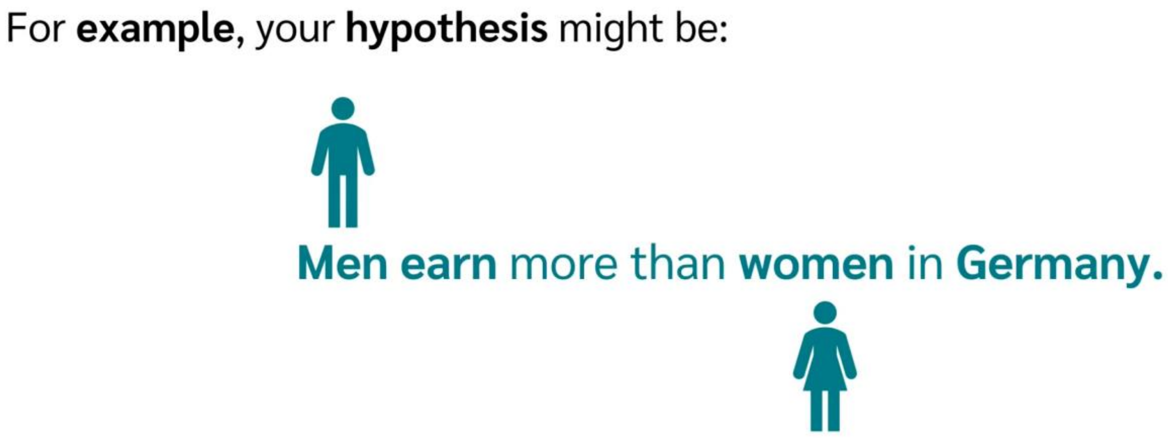
Estimation	The problem of estimation is considered, when there is no previous knowledge of the population parameter. The problem is simpler in that case. A random sample is taken, a sample statistic is computed and an appropriate point and interval estimate is suggested.
Hypothesis Testing	Often the interest is not in the numerical value of the point estimate of the parameter, but in knowing the plausibility of a hypothesis about the population parameter by using sample data. Estimation is not enough to arrive at a conclusion in such cases.

Definitions

In statistics,

a **hypothesis** is a statement or claim about a property of a population.

- The "property of a population" is often the value of a population parameter (such as mean or variance)



Definitions

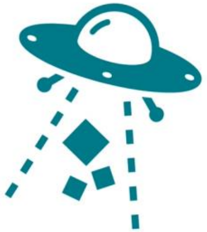
Our goal now is to **test** this **hypothesis.**

So we want to know whether to **reject or retain** the hypothesis.

Example:

Hypothesis

Men earn more than women in Germany.



The collected data, we can then analyze with the help of a **hypothesis test.**

In statistics,

a **hypothesis test** (or test of significance) is a standard procedure for testing a claim about the value of a population parameter in the light of the evidence gathered from the sample.

Hypothesis Test

But can we just **test every statement?**

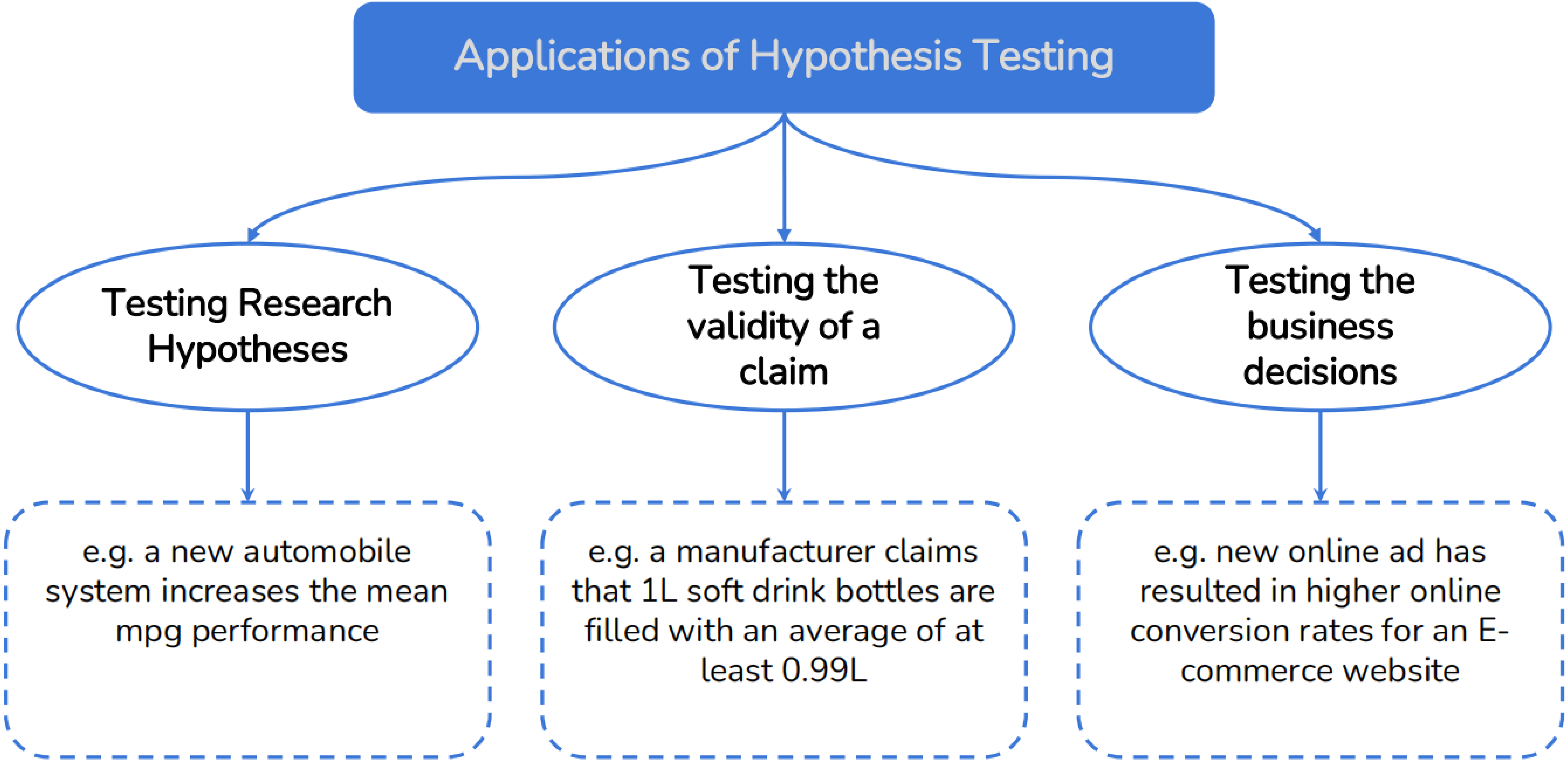
No! Hypotheses are **not simple statements!**

Hypotheses are **formulated** in such a way

that in the further **research process**

they can be tested with collected **data.**

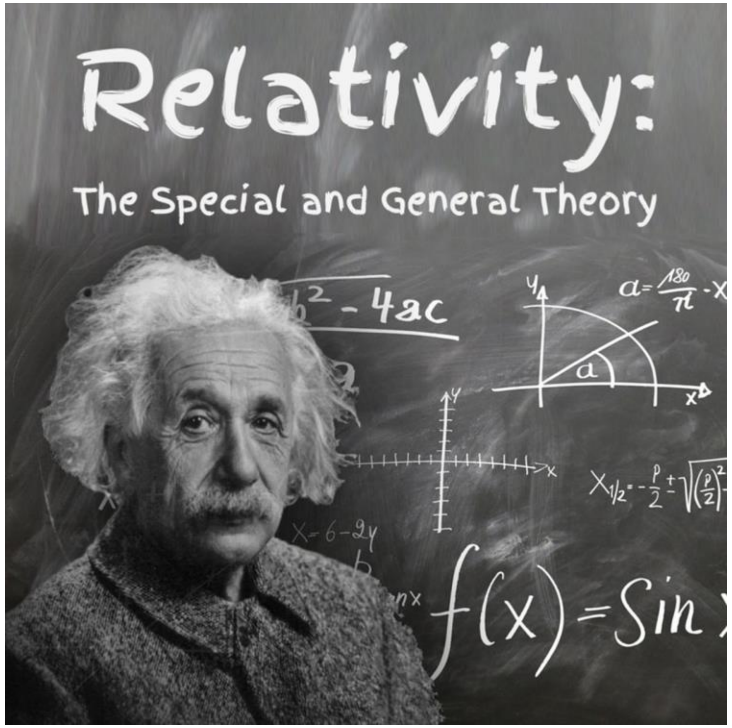
Overview of Applications



Motivation

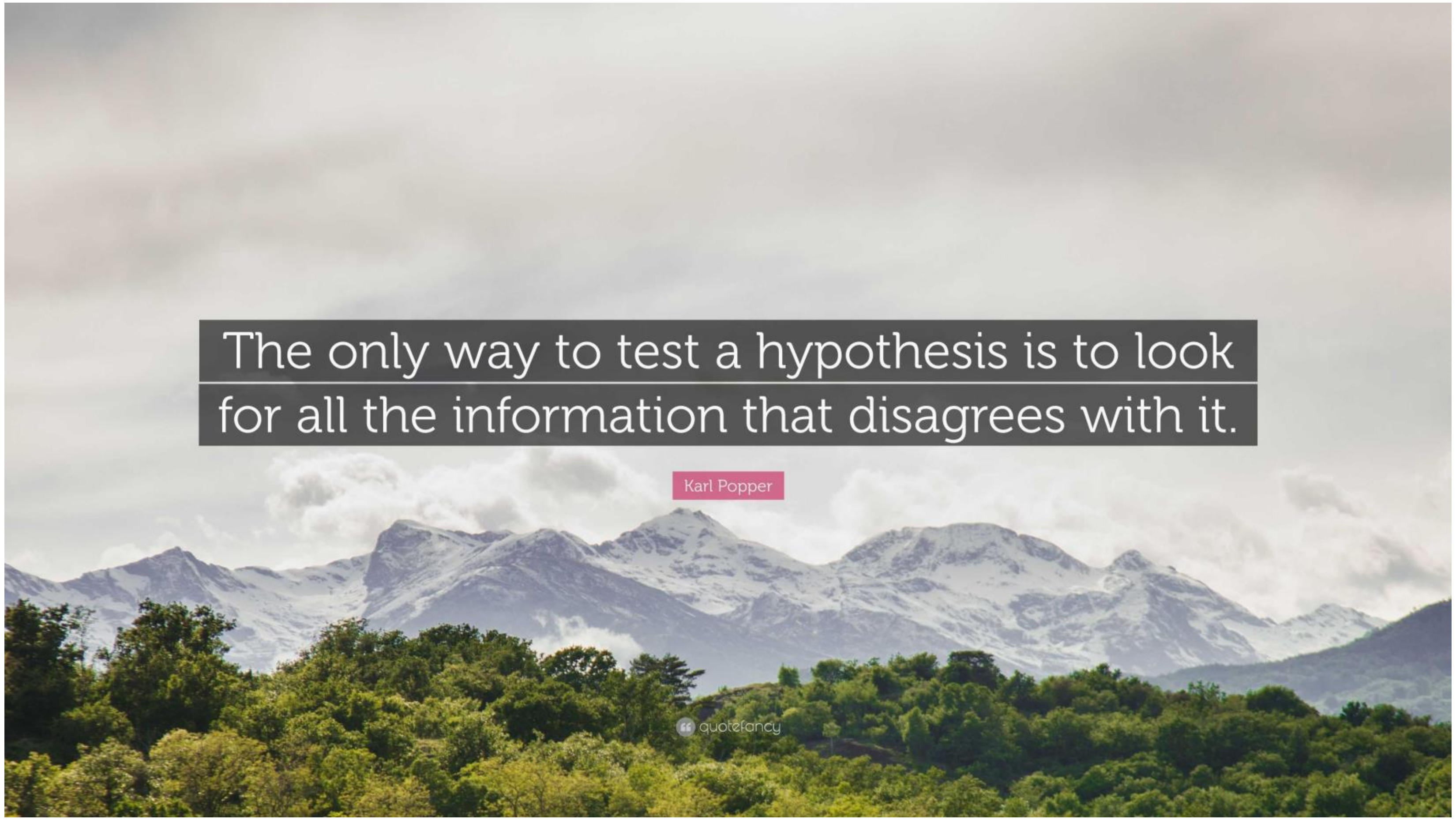
Law of Universal Gravitation

200 year later



General Theory of Relativity

Motivation


The only way to test a hypothesis is to look for all the information that disagrees with it.
Karl Popper
quotefancy

Components of a Formal Hypothesis Test

1. Form Hypothesis
2. State Significance Level
3. Calculate Test Statistic
4. Find Critical Value(s)
5. Conclusion

Forming Hypothesis

Statistical hypothesis

- A statement, or claim, about a population parameter.
- Need a pair of hypotheses
    - one that represents the claim
    - the other, its complement
- When one of these hypotheses is false, the other must be true.

Forming Hypothesis

Null hypothesis (denoted by H_0) is a statement that the value of a population parameter is equal to some claimed value.

- MUST contain condition of **equality**
- No effect or no difference
- Symbolic form: $=, \geq, \leq$
- Read "H subzero" or "H naught"

Alternative hypothesis (denoted by H_a) is the statement that the parameter has a value that differs from the null hypothesis.

- Hypothetical statement that the researcher wants to test.
- Symbolic form : $\neq, >, <$
- Read "H sub-a"

Forming Hypothesis-Examples

State the null and alternative hypotheses that would be used to test each of the following statements:

- A manufacturer claims that the average life of a transistor is less than1000 hours (h)

  $H0 : \mu \geq 1000$

  $H1 : \mu < 1000$

- A pharmaceutical firm maintains that the average time for a certain drug to take effect is 15 mins

  $H0 : \mu = 15$

  $H1 : \mu \neq 15$

- The mean starting salary of graduates is higher than R50000 per annum

  $H0 : \mu = 50000$

  $H1 : \mu > 50000$

Types of Errors

- No matter which hypothesis represents the claim, **always begin the hypothesis test assuming that the equality condition in the null hypothesis is true.**

- At the end of the test, one of two decisions will be made:
  - reject the null hypothesis
  - fail to reject the null hypothesis

- Because your decision is based on a sample, there is the possibility of making the wrong decision.

Types of Errors

	Actual Truth of H_0	
Decision	H_0 is true	H_0 is false
Do not reject H_0	**Correct Decision**	**Type II Error = β**
Reject H_0	**Type I Error = α**	**Correct Decision**

- A **type I error** occurs if the null hypothesis is rejected when it is true.
- A **type II error** occurs if the null hypothesis is not rejected when it is false.

Level of Significance

Level of significance

- Your maximum allowable probability of making a type I error.
  - Denoted by α

- **[Pre-specified]**By setting α at a small value, you are saying that you want the probability of rejecting a true null hypothesis to be small.

- Commonly used levels of significance:
  - $\alpha = 0.10$ $\alpha = 0.05$ $\alpha = 0.01$

	Actual Truth of H_0	
Decision	H_0 is true	H_0 is false
Do not reject H_0	**Correct Decision**	**Type II Error = β**
Reject H_0	**Type I Error = α**	**Correct Decision**

- $\alpha = P(\text{type I error}) = P(\text{rejecting } H_0 \text{ when } H_0 \text{ is true})$
- $\beta = P(\text{type II error}) = P(\text{failing to reject } H_0 \text{ when } H_0 \text{ is false})$

(Power of a hypothesis test: 1-β: rejecting a false null hypothesis.)

Statistical Tests

- To decide whether H_0 will be rejected or not , a value, called the **TEST STATISTIC** has to be calculated by using sample results

There are different test statistics for testing:

- **One-Sample** (single population)
  - Mean, Proportion, Variance
- **Two-Samples** (Difference between two populations)
  - Mean, Proportion, Variance

Parameter	Sampling Distribution	Test Statistics
Proportion: p	Normal (Z)	$Z = \frac{\hat{p} - p}{\sqrt{pq/n}}$
Mean: μ	t	$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
Mean: μ	Normal (Z)	$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
Standard Deviation: σ **Or Variance:** σ^2	χ^2	$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

One-Sample

Critical Region

Distribution of test statistic often follows a certain distribution.

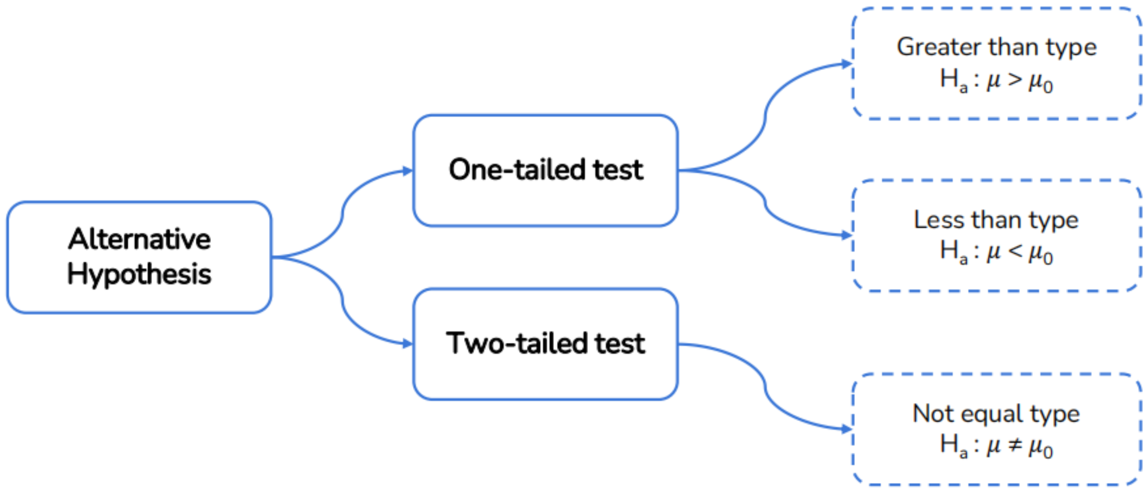
- Distribution can be divided into 2 regions:
  - A region of rejection (Critical region)
  - A region of non- rejection

- **Critical region** - set of all values of the test statistic that would cause a rejection of the null hypothesis.
- **Critical Value** – value or values that separate the critical region from the values of the test statistics that do **not** lead to a rejection of the null hypothesis.

Types of the Test

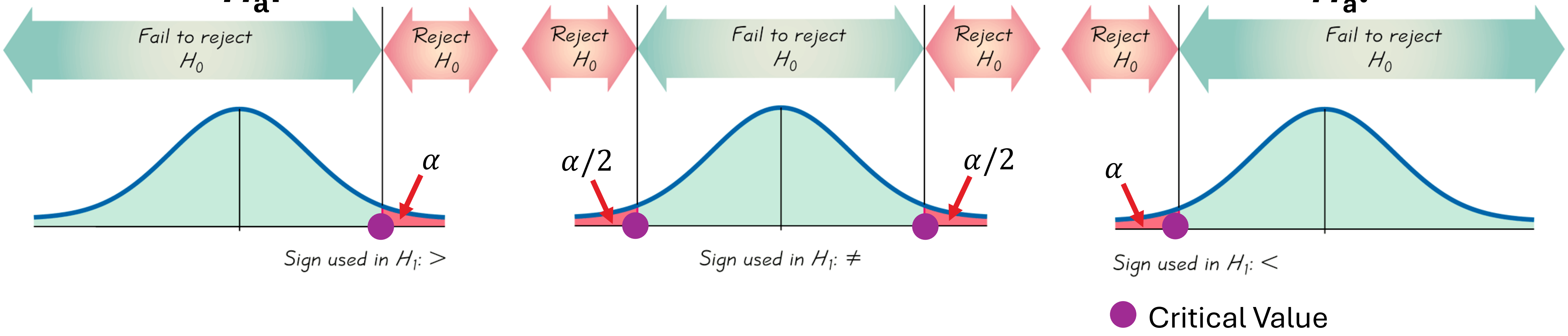
- The type of test depends on the region of the sampling distribution that favors a rejection of H_0.
- This region is indicated by the **alternative hypothesis.**

Three types of hypothesis tests
- Left-tailed test
- Right-tailed test
- Two-tailed test



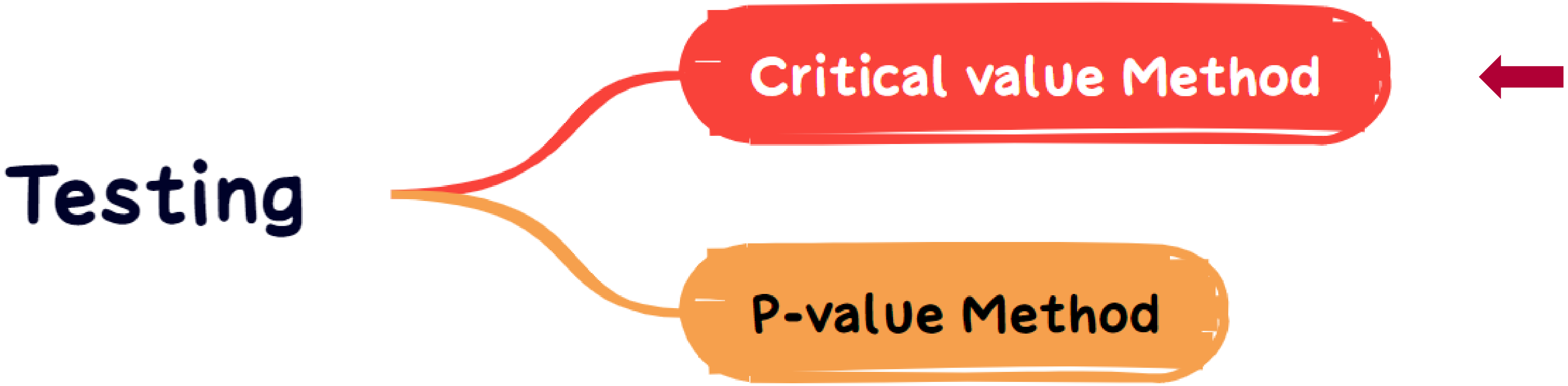
Types of the Test

Right tailed	Two tailed	Left tailed
$H_0: \leq$	$H_0: =$	$H_0: \geq$
$H_a: >$	$H_a: \neq$	$H_a: <$



Test statistic value **does not change** for two-tailed or one-tailed test.
Only the critical value(s) associated with the test statistic changes.

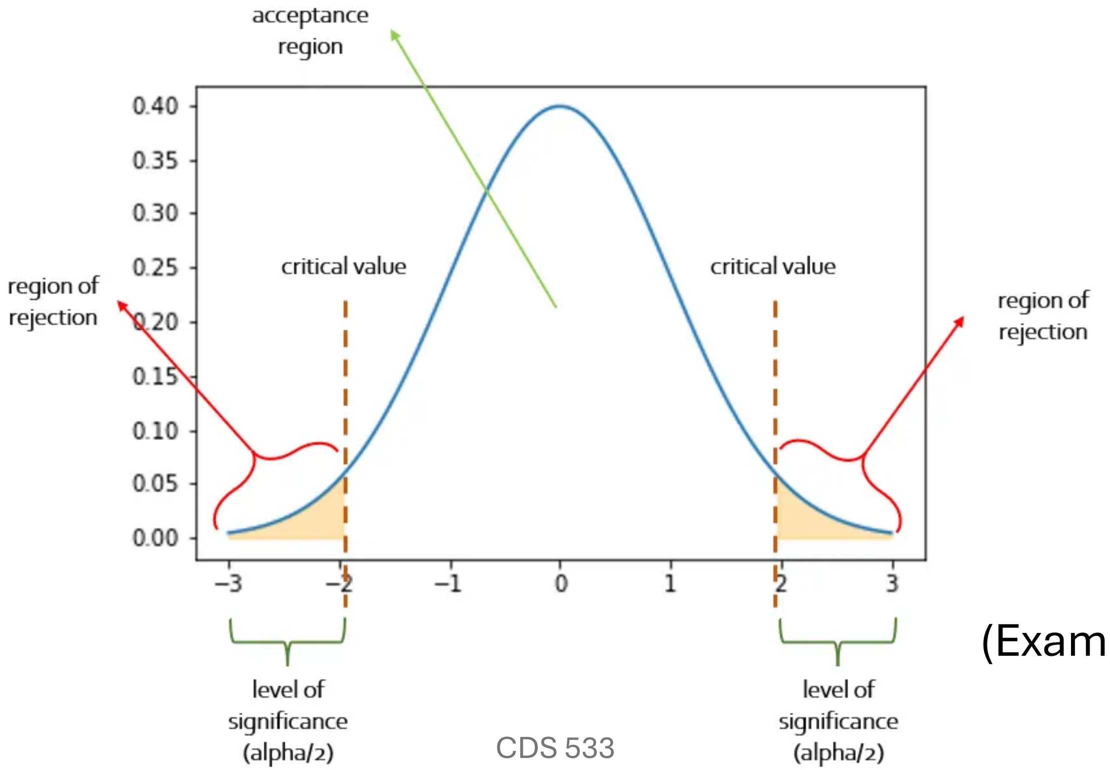
Testing Methods



Critical Value Method

To use a rejection region to conduct a hypothesis test, calculate the **test statistic**. If the standardized test statistic

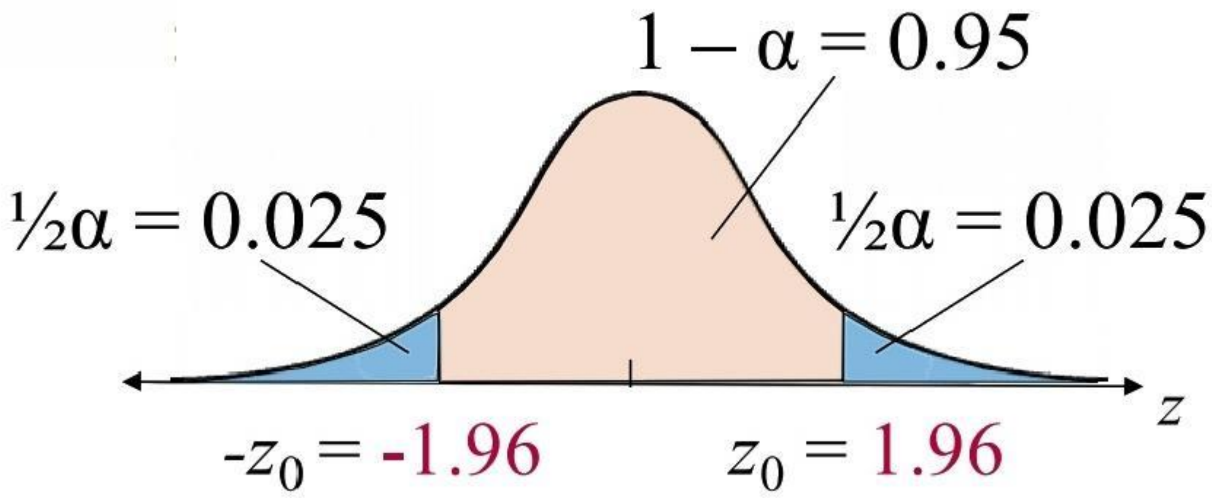
1. is in the rejection region, then reject H_0 .
2. is not in the rejection region, then fail to reject H_0 .



(Example of Two-tails Test)

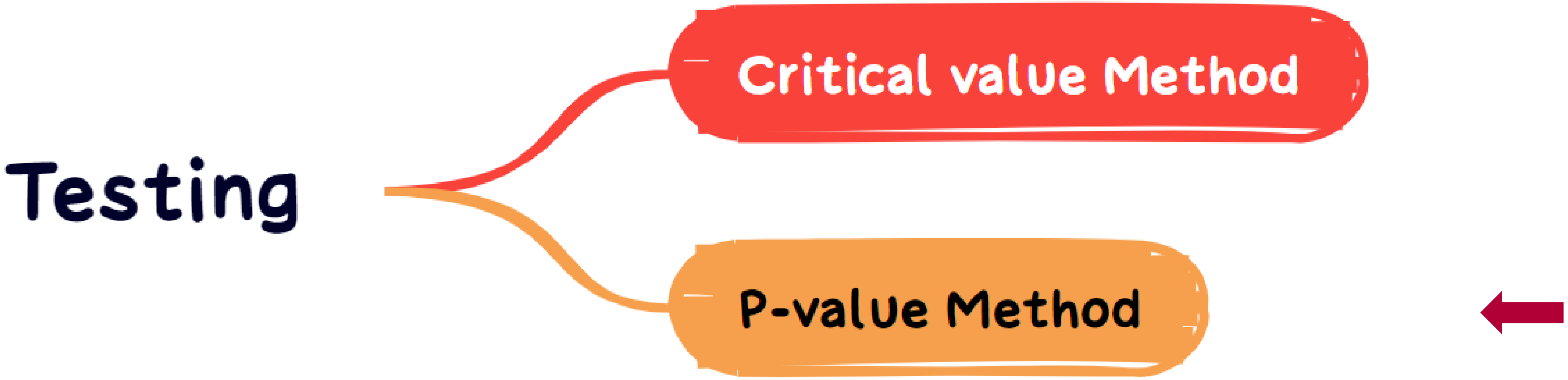
Critical Value Method: Example

Find the critical value and rejection region for a two-tailed test with $a = 0.05$. The Sampling distribution is normal.



The rejection regions are
to the left of $-z_0$=-1.96 and to the right of z_0=1.96

Testing Methods



P-Value Method

P-value (or probability value)

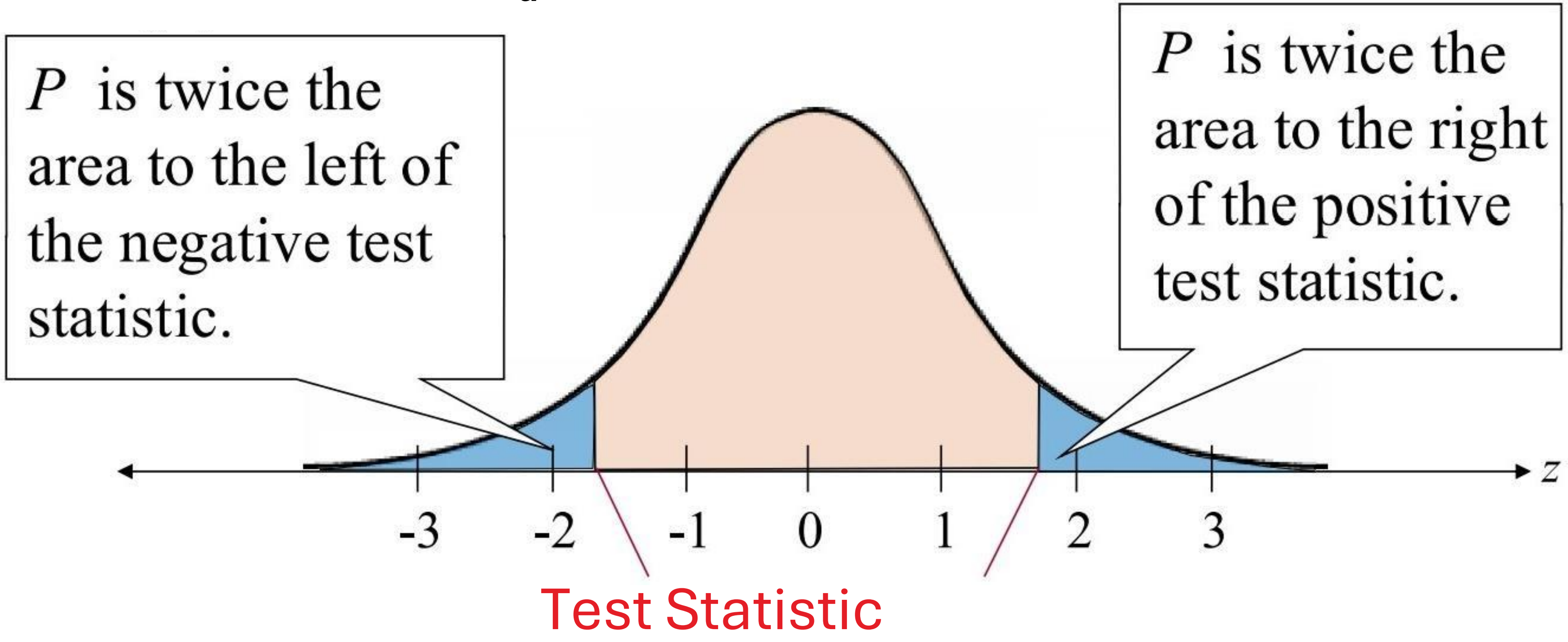
- The probability, if the null hypothesis is true, of obtaining a sample statistic with a value **as extreme or more extreme than** the one determined from the sample data, assuming that the null hypothesis is true.
- Use test statistic to find the **probability**.
- Depends on the Types of the test.

P-Value Method

Two-tailed Test

H_0: =

H_a: ≠



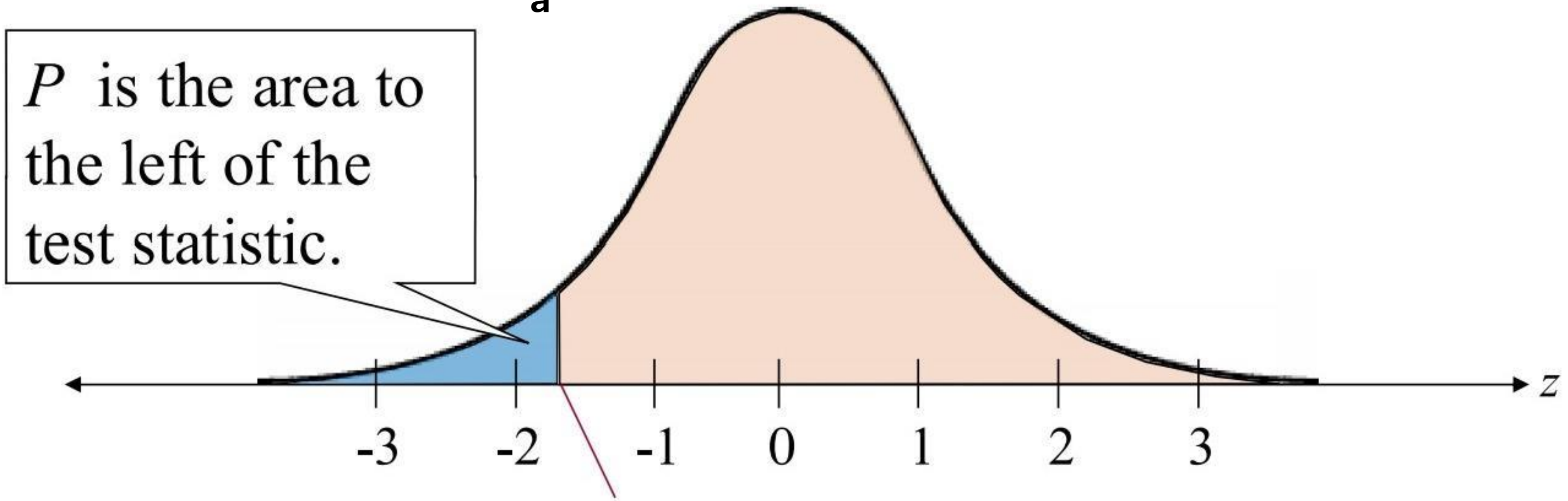
P-Value Method

1. Form Hypothesis
2. Select Significance Level
3. Calculate Test Statistic
4. Find Critical Value(s)
5. Conclusion

Left-tailed Test

H_0: ≥

H_a: <



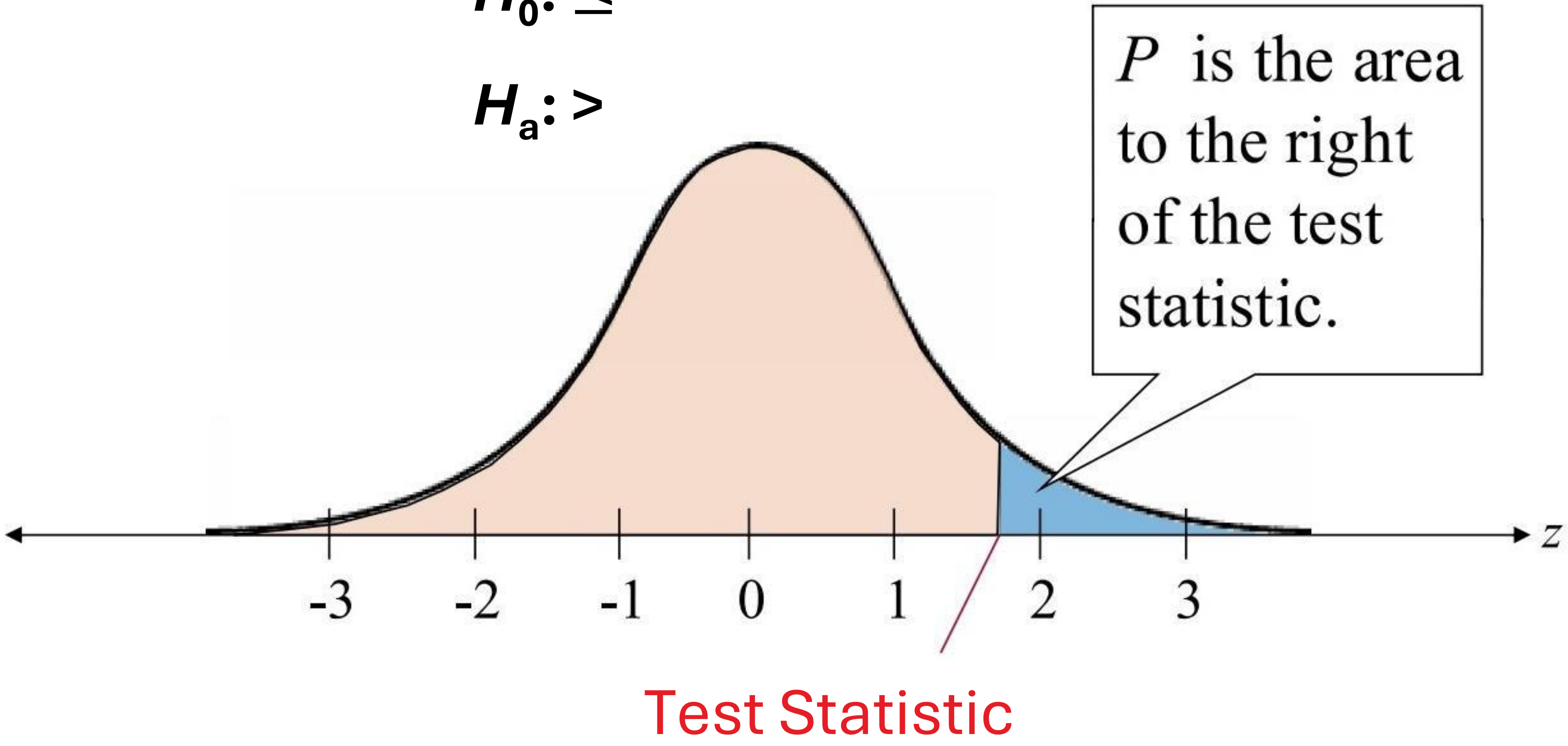
Test Statistic

P-Value Method

Right-tailed Test

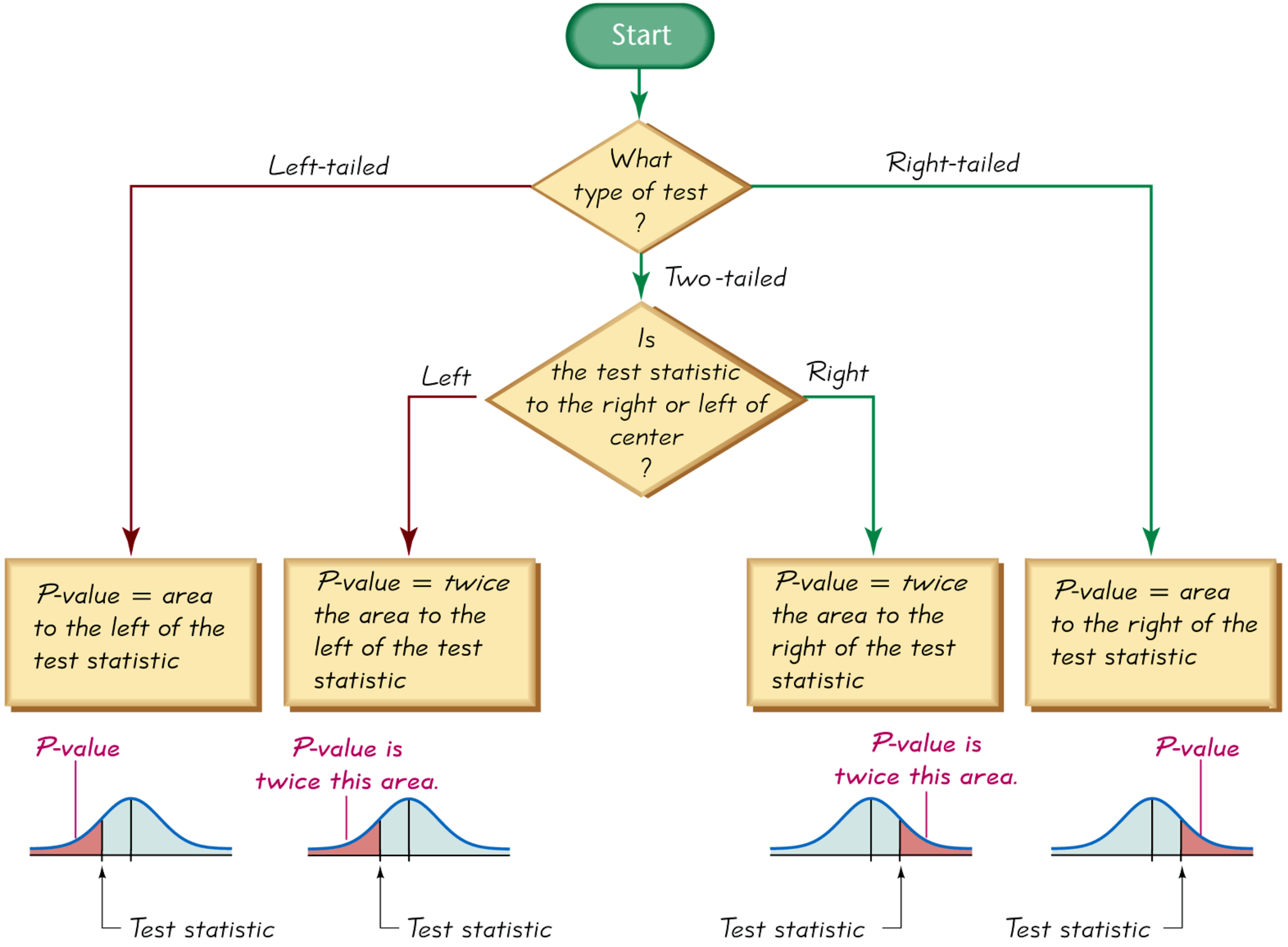
$H_0: \leq$

$H_a: >$



P-Value Method

1. Form Hypothesis
2. Select Significance Level
3. Calculate Test Statistic
4. Find Critical Value(s)
5. Conclusion

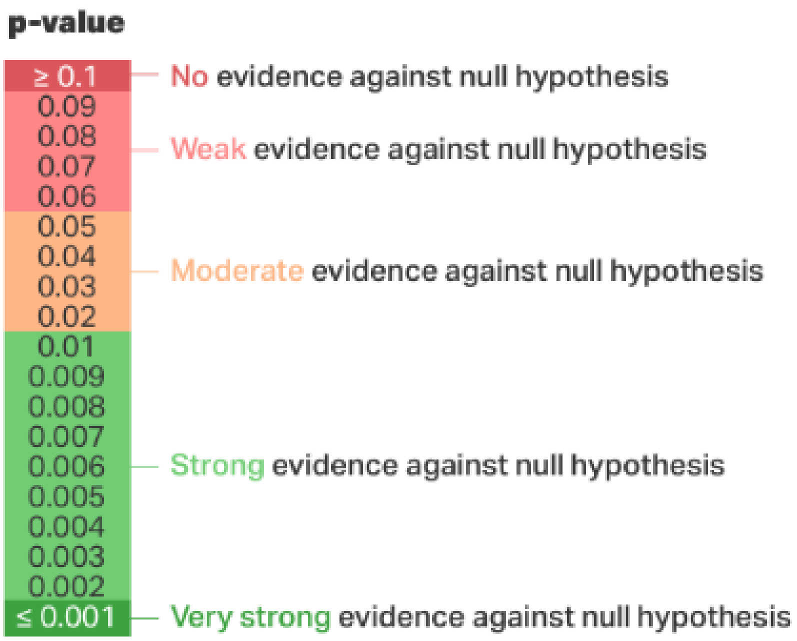


1. Form Hypothesis
2. Select Significance Level
3. Calculate Test Statistic
4. Find Critical Value(s)
5. Conclusion

P-Value Method

Decision Rule Based on P-value

- Compare the P-value with a.
  - If $P \leq a$, reject H_0 .
  - If $P > a$, fail to reject H_0.



Here is a memory tool useful for interpreting the P-value:

If the P is low, the null must go.
If the P is high, the null will fly.

P-Value Method: Example

For each claim, state H_0 and H_a . Then determine whether the hypothesis test is a left-tailed, right-tailed, or two-tailed test. Sketch a normal sampling distribution and shade the area for the P-value.

1. A university publicizes that the proportion of its students who graduate in 4 years is 82%.

Solution:

$$H_0: p = 0.82$$
$$H_a: p \neq 0.82$$

P-Value Method: Example

For each claim, state H_0 and H_a . Then determine whether the hypothesis test is a left-tailed, right-tailed, or two-tailed test. Sketch a normal sampling distribution and shade the area for the P-value.

2. A water faucet manufacturer announces that the mean flow rate of a certain type of faucet is less than 2.5 gallons per minute.

Solution:

H_0: $\mu \geq 2.5$ gpm

H_a: $\mu < 2.5$ gpm

P-Value Method: Example

For each claim, state H_0 and H_a . Then determine whether the hypothesis test is a left-tailed, right-tailed, or two-tailed test. Sketch a normal sampling distribution and shade the area for the P-value.

3. A cereal company advertises that the mean weight of the contents of its 20-ounce size cereal boxes is more than 20 ounces.

Solution:

H_0: $\mu \leq 20$ oz

H_a: $\mu > 20$ oz

Conclusions

1. **Critical Value Method**

- Reject H_0 if the test statistic falls in the critical region
- Fail to reject H_0 if the test statistic does not fall in the critical region

2. **P-Value Method**

- Reject H_0 if the P-value is less than or equal a, i.e. $P \leq \alpha$
- Fail to reject H_0 if the P-value is greater than the a, i.e., $P > a$

We are not proving the null hypothesis.

Conclusions

Wording of final conclusion

1. **Reject** the H_0

Conclusion: There **is** sufficient (enough) evidence to conclude.......................................(what ever H_a says)

2. **Fail to reject** the H_0

Conclusion: There **is not** sufficient (enough) evidence to conclude.......................................(what ever H_a says)

Conclusions

Wording of final conclusion

1. **Reject** the H_0

Conclusion: There **is** sufficient (enough) evidence to conclude.....................................(what ever H_a says)

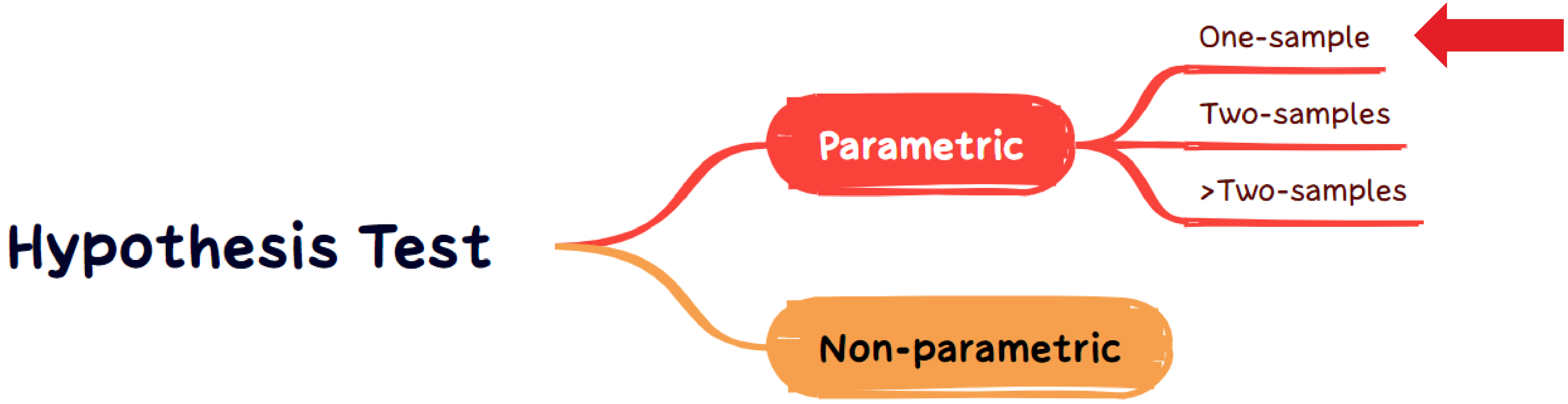
2. **Fail to reject** the H_0

Conclusion: There **is not** sufficient (enough) evidence to conclude.....................................(what ever H_a says)

Steps of Hypothesis Test

1. Determine H_0 and H_a. (and a if necessary)
2. Determine the correct test statistic and calculate.
3. Determine the critical values and sketch a graph.
4. Determine Reject H_0 or Fail to reject H_0
5. State conclusion in simple nontechnical terms.

Hypothesis Testing



One-sample

Hypothesis Test for

Mean

- Large sample
- Small sample

Proportion

Variance

One-sample: Large Sample Mean

Mean with $n \geq 30 \; or \; \sigma \; is \; known$

Z-test

Assumption:

1. The sample is a random sample.
2. The sample is large (n>30): a)CLT applies; b)Can use normal
3. If σ is known

Notation

$\bar{x}$ = sample mean

μ = hypothesized population mean

σ = population standard deviation

n = sample size

One-sample: Large Sample Mean

Mean with $n \geq 30$ $or\ \sigma\ is\ known$

Testing $H_0: \mu = \mu_0$ for $n \geq 30$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: \mu \neq \mu_0$	$\|z\| \geq Z_{1-\alpha/2}$	$z = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{s}{\sqrt{n}}\right)}$
$H_1: \mu > \mu_0$	$z \geq Z_{1-\alpha}$	
$H_1: \mu < \mu_0$	$z \leq -Z_{1-\alpha}$	

Use σ if known

One-sample: Large Sample Mean

Example

It will be cost effective to employee an additional staff member at a well known take away restaurant if the average sales for a day is more than 11,000 per day. A sample of 60 days were selected and the average sales for the 60 days were 11,841 with a standard deviation of 1,630. Test if it will be cost effective to employ an additional staff member. Assume a normal distributed population. Use α = 0.05

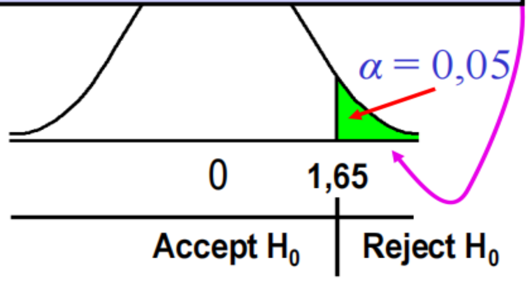
One-sample: Large Sample Mean

- Solution
  - $H_0 : \mu = 11\,000$
  - $H_1 : \mu > 11\,000$
  - $\alpha = 0,05$

$$z = \frac{\bar{x} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{11841 - 11000}{\frac{1630}{\sqrt{60}}} = 3,99$$

  - Reject H_0

At $\alpha = 0,05$ if it will be cost effective to employ an additional staff member – the average monthly income is more than R11 000



Critical value $Z_{1-\alpha} = Z_{0.95} = 1.65$

One-sample: Small Sample Mean

Mean with $n < 30$ and σ is unknown

t-test

Assumption:

1. The sample is a random sample.
2. The sample is small (n≤30): a) Population is approximately normal
3. If σ is unknown

Notation

$\bar{x}$ = sample mean

μ = hypothesized population mean

s = Sample standard deviation

n = sample size

One-sample: Small Sample Mean

Mean with $n < 30$ *and* σ *is unknown*

Testing $H_0: \mu = \mu_0$ for $n < 30$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: \mu \neq \mu_0$	$\lvert t \rvert \geq t_{n-1;1-\alpha/2}$	$t = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{s}{\sqrt{n}}\right)}$
$H_1: \mu > \mu_0$	$t \geq t_{n-1;1-\alpha}$	
$H_1: \mu < \mu_0$	$t \leq -t_{n-1;1-\alpha}$	

One-sample: Small Sample Mean

Example

Government claims that patients will wait less than 30 minutes on average to see a doctor. A random sample of 25 patients revealed that their average waiting time was 28 minutes with a standard deviation of 8 minutes. On a 1% level of significance can we say that the claim from government is correct?

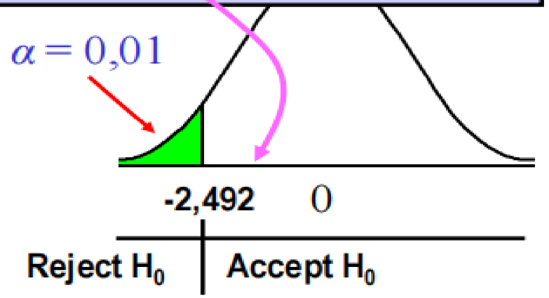
One-sample: Small Sample Mean

- Solution
  - $H_0 : \mu = 30$
  - $H_1 : \mu < 30$
  - $\alpha = 0{,}01$

$$t = \frac{\bar{x} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{28 - 30}{\frac{8}{\sqrt{25}}} = -1{,}25$$

  - Accept H_0

At α = 0,01 we can <u>not</u> say that the average waiting time at clinics is less than 30 minutes



$t_{n-1;1-\alpha} = t_{24;0.99} = -2.492$

One-sample: Proportion

Assumption:

1. The sample is a random sample.
2. The conditions np ≥ 5 and nq ≥ 5 are both satisfied ($\mu = np, \sigma = \sqrt{npq}$)

Notation

n = sample size or number of trials

p = population proportion (used in the null hypothesis)

$\hat{p} = \frac{x}{n}$ = Sample proportion

Notes:

$\hat{p}$ sometimes is given directly

"10%of the observed sports cars are red" is expressed as

$\hat{p} = 0.10$

$\hat{p}$ sometimes must be calculated

"96 surveyed households have cable TV and 54 do not"is calculated using

$\hat{p} = \frac{x}{n} = \frac{96}{(96+54)} = 0.64$

One-sample: Proportion

Testing $H_0: p = p_0$ for $n \geq 30$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: p \neq p_0$	$\lvert z \rvert \geq Z_{1-\alpha/2}$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$
$H_1: p > p_0$	$z \geq Z_{1-\alpha}$	
$H_1: p < p_0$	$z \leq -Z_{1-\alpha}$	

One-sample: Proportion

Example

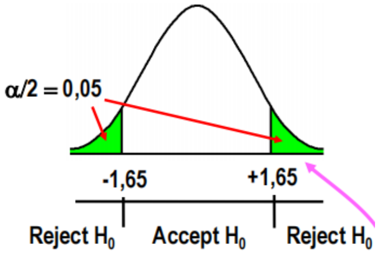
A market research company investigates the claim of a supplier that 35% of potential buyers are preferring their brand of milk. A survey was done in several supermarkets and it was found that 61 of the 145 shoppers indicated that they will buy the specific brand of milk. Assist the research company with the claim of the supplier on a 10% level of significance.

One-sample: Proportion

- Solution
  - The population of interest is the proportion of buyers
  - Want to test the claim that the proportion is 35% = 0,35
  - $H_0 : p = 0,35$
  - The alternative hypothesis must specify that the proportion is not 35%
  - $H_1 : p \neq 0,35$
  - Sample proportion $\hat{p} = \frac{\text{number of successes}}{\text{sample size}} = \frac{x}{n} = \frac{61}{145} = 0,42$
  - Need to test if 0,42($\hat{p}$) is significant different from 0,35(p)

- Solution

At α = 0,10 we can not say that 35% of the clients will prefer the brand of milk



$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0,42 - 0,35}{\sqrt{\frac{0,35(1-0,35)}{145}}} = 1,76$$

- Reject H_0

$Z_{1-\alpha/2} = Z_{0.95} = +/- 1.65$

p-value=2P(z>1.76)=2(1-0.9608)=0.0784

α=0.1

P-value=0.0784<0.1 -> Reject

TABLE A-2 *(continued)* Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.940[illegible]	.9418
1.6	.9452	.9463	.9474	.9484	.9495 *	.9505	.95[illegible]	.9525
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693

One-sample: Variance

Chi-Square Statistic

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

with degrees of freedom df = n – 1, formally denoted as χ^2_{n-1}.

Testing $H_0: \sigma^2 = \sigma^2_0$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: \sigma^2 \neq \sigma^2_0$	$X^2 \leq X^2_{n-1;\alpha/2}$ or $X^2 \geq X^2_{n-1;1-\alpha/2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$
$H_1: \sigma^2 > \sigma^2_0$	$X^2 \geq X^2_{n-1;1-\alpha}$	
$H_1: \sigma^2 < \sigma^2_0$	$X^2 \leq X^2_{n-1;\alpha}$	

One-sample: Variance

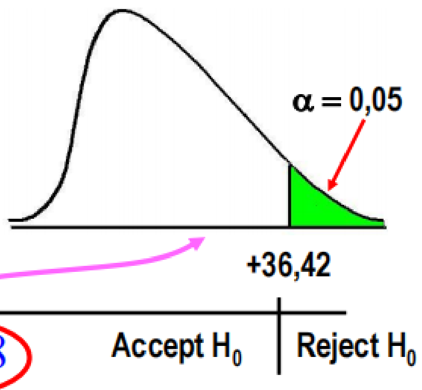
Example

The variation in the content of a 340ml can of beer should be more than $10ml^2$. To test the validity of this, 25 cans of beers revealed a variance of 12 ml^2. On a 5% level of significance can we say the variation in the content of the cans is too large?

One-sample: Variance

- Solution
  - $H_0 : \sigma^2 = 10$
  - $H_1 : \sigma^2 > 10$
  - $\alpha = 0{,}05$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(25-1)12}{10} = 28{,}8$$



At α = 0,05 we can not say that the variation in the content of the cans is more than $10 ml^2$

  - Accept H_0

$X^2_{n-1;\,1-\alpha} = X^2_{24;\,0.95} = 36.42$

Degrees of Freedom	Area to the *Right* of the Critical Value							
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852
20	7.434	8.260	9.591	10.851	12.443	[illegible]	31.410	34.170
21	8.034	8.897	10.283	11.591	13.240	[illegible]	32.671	35.479
22	8.643	9.542	10.982	12.338	14.042	[illegible]	33.924	36.781
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646

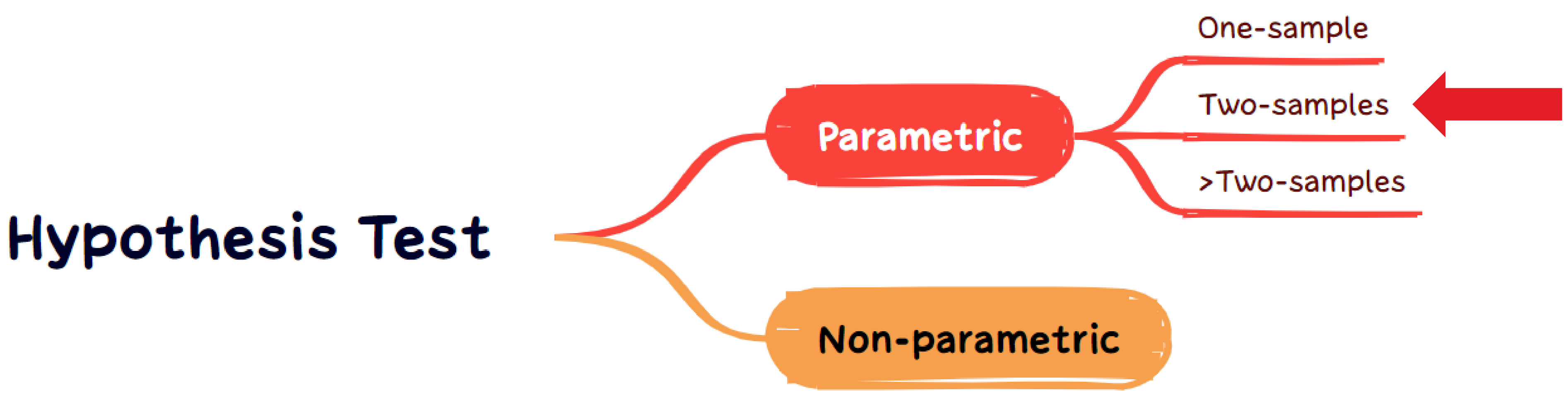
p-value=P(χ2 >28.8) is between 0.9 and 0.1
α=0.05
P-value>0.05 -> Fail to reject.

Comparison with Confidence Interval

The inferences from the hypothesis test and the confidence interval are consistent with each other.

- If we are 99% confident that $98.1 < \mu < 98.4$, then the two-sided p-value for the null hypothesis $\mu = 98.6$ must be less than 0.01, and it is.
- The hypothesis test quantifies evidence against a null hypothesis with a probability (of observing a result at least as extreme as that observed, assuming the null hypothesis is true), on a scale from 0 to 1.
- A confidence interval quantifies uncertainty on the scale of the measurements of the data.
- Thus, confidence intervals provide information in the context of the problem, and thus are more informative to the reader with background knowledge.
- A confidence interval allows the reader to ascertain the practical importance of the inference.

Hypothesis Testing



Two-samples

Hypothesis Test for

Difference between two means

- Large sample
- Small sample

Difference between two proportions

Difference between two Variances

Hypothesis tests for comparing two Populations

- H_0: Population 1 parameter = Population 2 parameter
- H_1: Population 1 parameter $\neq$ Population 2 parameter
- H_1: Population 1 parameter > Population 2 parameter
- H_1: Population 1 parameter < Population 2 parameter

$\boxed{\mu_1 \,/\, \sigma^2{}_1 \,/\, p_1}$ $\boxed{\mu_2 \,/\, \sigma^2{}_2 \,/\, p_2}$

Two-samples: Means

Compares two parameters from two populations.

Sampling methods:

Independent Samples

The sample selected from one population is not related to the sample selected from the second population.

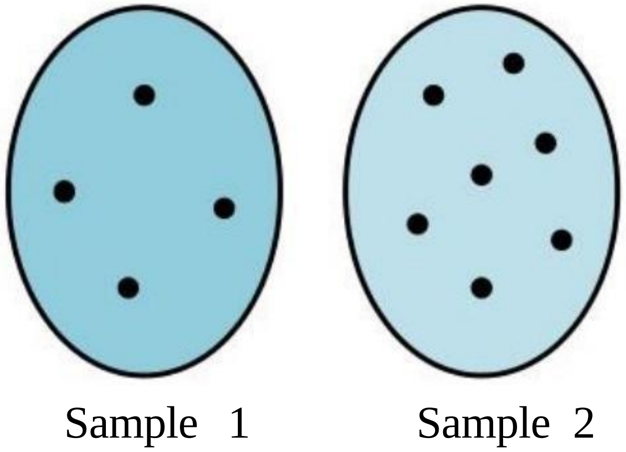
- Large samples
- Small samples

Dependent Samples (paired or matched samples)

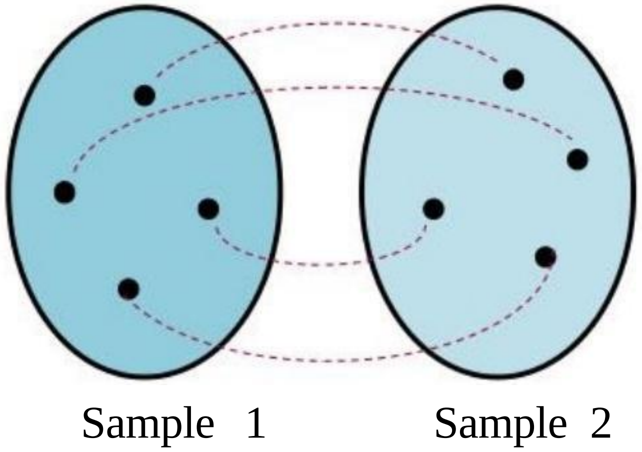
Each member of one sample corresponds to a member of the other sample.

Independent and Dependent Samples

Independent Samples



Dependent Samples



Two-samples: Independent means

$[\boldsymbol{Unequal\ Variance}]\ \sigma_1^2 \neq \sigma_2^2$

Testing $H_0: \mu_1 = \mu_2$ for $n_1 \geq 30$ and $n_2 \geq 30$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: \mu_1 \neq \mu_2$	$\lvert z \rvert \geq Z_{1-\alpha/2}$	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
$H_1: \mu_1 > \mu_2$	$z \geq Z_{1-\alpha}$	
$H_1: \mu_1 < \mu_2$	$z \leq -Z_{1-\alpha}$	

Use σ_1^2 and σ_2^2 if known

Two-samples: Independent means

[***Equal Variance***] $\sigma_1^2 = \sigma_2^2$ (homogeneity of variance)

Testing H_0: $\mu_1 = \mu_2$ for $n_1 < 30$ and $n_2 < 30$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
H_1: $\mu_1 \neq \mu_2$	$\lvert t \rvert \geq t_{n1+n2-2\,;\,1-\alpha/2}$	$t = \dfrac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ with $s_p = \sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$
H_1: $\mu_1 > \mu_2$	$t \geq t_{n1+n2-2\,;\,1-\alpha}$	
H_1: $\mu_1 < \mu_2$	$t \leq -t_{n1+n2-2\,;\,1-\alpha}$	

Two-samples: Independent Means

Example

There is a belief that people staying in Cape Town travel less than people staying in Johannesburg. Random samples of 43 people in Cape Town and 39 in Johannesburg were drawn. For each person the distance travelled during October were recorded. Test the belief on a 5% level of significance.

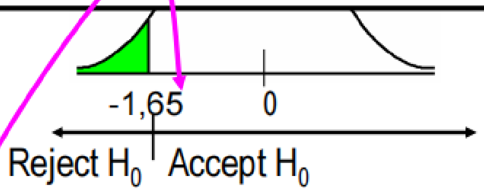
Cape Town	Johanesburg
$\overline{x_1}$ = 604	$\overline{x_2}$ = 633
n_1 = 43	n_2 = 39
s_1 = 64	s_2 = 103

Two-samples: Independent Means

- From the data:

Cape Town	Johan…
$\bar{x}_1 = 604$	$\bar{x}_2 =$
$n_1 = 43$	$n_2 = 39$
$s_1 = 64$	$s_2 = 103$

The belief that people staying in Cape Town travel less than people staying in Johannesburg is **not** true on a 5% level of significance



- $H_0: \mu_1 = \mu_2$
- $H_1: \mu_1 < \mu_2$
- $$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{604 - 633}{\sqrt{\frac{64^2}{43} + \frac{103^2}{39}}} = -1,51$$
- Accept H_0

p-value=P(z <-1.51) is 0.655

α=0.05

P-value>0.05 -> Fail to reject.

TABLE A-2 Standard Normal (z) Distribution: Cumulative Area from the LEFT								
z	.00	.01	.02	.03	.04	.05	.06	.07
−3.50 and lower	.0001							
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307
−1.7	.0446	[illegible]	.0427	.0418	.0409	.0401	.0392	.0384
−1.6	.0548	[illegible]	.0526	.0516	.0505 *	.0495	.0485	.0475
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020

Two-samples: Dependent means (Paired)

$\sigma_1^2 = \sigma_2^2$

	Observation 1	2	3	- - - - - - - - - - - -	n
Sample 1	X_{11}	X_{12}	X_{13}	- - - - - - - - - - - -	X_{1n}
Sample 2	X_{21}	X_{22}	X_{23}	- - - - - - - - - - - -	X_{2n}
Difference (d)	d_1 $(X_{11} - X_{21})$	d_2 $(X_{12} - X_{22})$	d_3 $(X_{13} - X_{23})$		d_n $(X_{1n} - X_{2n})$

$$\bar{d} = \frac{1}{n}\sum d \quad \text{and} \quad s_d = \sqrt{\frac{\sum d^2 - \frac{1}{n}\left(\sum d\right)^2}{n-1}}$$

Two-samples: Dependent means (Paired)

$\sigma_1^2 = \sigma_2^2$

Testing $H_0: \mu_1 = \mu_2$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: \mu_1 \neq \mu_2$	$\lvert t \rvert \geq t_{n-1\,;\,1-\alpha/2}$	$t = \dfrac{\bar{d}}{\left(\dfrac{s_d}{\sqrt{n}}\right)}$
$H_1: \mu_1 > \mu_2$	$t \geq t_{n-1\,;\,1-\alpha}$	
$H_1: \mu_1 < \mu_2$	$t \leq -t_{n-1\,;\,1-\alpha}$	

Two-samples: Paired Means

Example

Blood samples for 10 different types of test were analyzed by the traditional laboratories and by the newly equipped laboratories. The time, in minutes, were captured for each test. Did the time reduce? α = 0.01

Blood sample	Existing lab	New lab
1	47	70
2	65	83
3	59	78
4	61	46
5	75	74
6	65	56
7	73	74
8	85	52
9	97	99
10	84	57

-Calculate the difference for each row

-Calculate the average differences and the standard deviation of the differences

$\bar{x}_d$ =2.2

s_d =19.14

Two-samples: Paired Means

p-value=P(t >0.36) >0.1
α=0.01
P-value>0.01 -> Fail to reject.

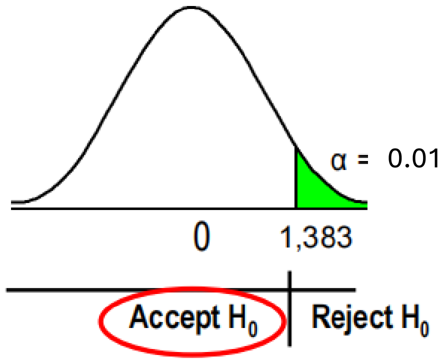
- The hypotheses test for this problem is

  H_0: $\mu_1 = \mu_2$

  H_1: $\mu_1 > \mu_2$

- The statistic is

$$t = \frac{\bar{d}}{\left(\frac{s_d}{\sqrt{n}}\right)} = \frac{2,2}{19,14/\sqrt{10}} = 0.36$$

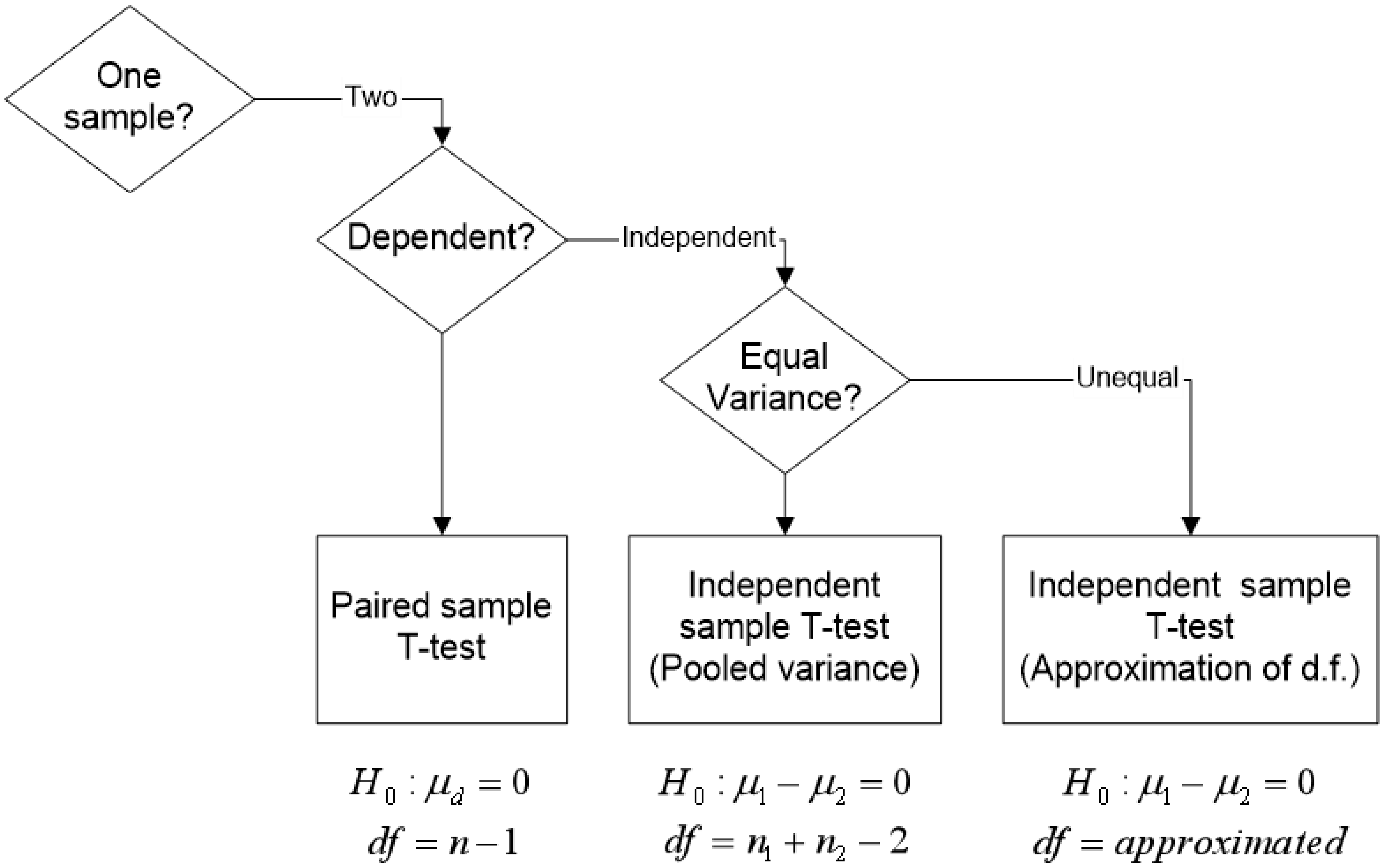


Using α = 0.01, introducing some new equipment the time taken did **not** reduce.

TABLE A-3 *t* Distribution: Critical *t* Values

	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
	Area in Two Tails				
Degrees of Freedom	0.01	0.02	0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341

Two-samples: Means



Two-samples: Proportions

Testing H_0: $p_1 = p_2$ for $n_1 \geq 30$ and $n_2 \geq 30$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
H_1: $p_1 \neq p_2$	$\lvert z \rvert \geq Z_{1-\alpha/2}$	$z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$
H_1: $p_1 > p_2$	$z \geq Z_{1-\alpha}$	
H_1: $p_1 < p_2$	$z \leq -Z_{1-\alpha}$	where $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$

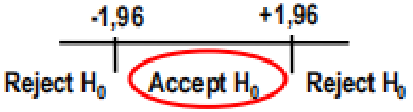
Two-samples: Proportion

Example

A clothing manufacturer introduced two new swim suit ranges on the market. Of the 266 clients asked if they will wear range A, 85 indicated they will. Of the 192 clients asked if they will wear range B, 50 indicated they will. Can we say there is a difference in the preferences of the two ranges. Use $\alpha = 0,05$.

Two-samples: Proportion

- From the data:
  - Range A: $\hat{p}_1 = \frac{85}{266} = 0,32$ Range B: $\hat{p}_2 = \frac{50}{192} = 0,26$
  - $\hat{p} = \dfrac{266(0,32) + 192(0,26)}{266 + 192} = 0,29$
  - H_0: $p_1 = p_2$
  - H_1: $p_1 \neq p_2$
  - $$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0,32 - 0,26}{\sqrt{0,29(1-0,29)\left(\frac{1}{266} + \frac{1}{192}\right)}} = 1,93$$
  - Accept H_0



There is no difference in the preferences of the two ranges if α = 0,05.

p-value=2P(z >1.93) =0.0536
α=0.05
P-value>0.05 -> Fail to reject.

TABLE A-2 *(continued)* Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05
0.0	.5000	.5040	.5080	.5120	.5160	.5199
0.1	.5398	.5438	.5478	.5517	.5557	.5596
0.2	.5793	.5832	.5871	.5910	.5948	.5987
0.3	.6179	.6217	.6255	.6293	.6331	.6368
0.4	.6554	.6591	.6628	.6664	.6700	.6736
0.5	.6915	.6950	.6985	.7019	.7054	.7088
0.6	.7257	.7291	.7324	.7357	.7389	.7422
0.7	.7580	.7611	.7642	.7673	.7704	.7734
0.8	.7881	.7910	.7939	.7967	.7995	.8023
0.9	.8159	.8186	.8212	.8238	.8264	.8289
1.0	.8413	.8438	.8461	.8485	.8508	.8531
1.1	.8643	.8665	.8686	.8708	.8729	.8749
1.2	.8849	.8869	.8888	.8907	.8925	.8944
1.3	.9032	.9049	.9066	.9082	.9099	.9115
1.4	.9192	.9207	.9222	.9236	.9251	.9265
1.5	.9332	.9345	.9357	.9370	.9382	.9394
1.6	.9452	.9463	.9474	.9484	.9495 *	.9505
1.7	.9554	.9564	.9573	[illegible]	.9591	.9599
1.8	.9641	.9649	.9656	[illegible]	.9671	.9678
1.9	.9713	.9719	.9726	.9732	.9738	.9744
2.0	.9772	.9778	.9783	.9788	.9793	.9798

Two-samples: Variance

Testing $H_0: \sigma^2_1 = \sigma^2_2$		
Alternative hypothesis	Decision rule: Reject H_0 if	Test statistic
$H_1: \sigma^2_1 \neq \sigma^2_2$	$F \geq F_{n1-1\,;\,n2-1\,;\,\alpha/2}$	$F = \dfrac{s_1^2}{s_2^2}$
$H_1: \sigma^2_1 > \sigma^2_2$	$F \geq F_{n1-1\,;\,n2-1\,;\,\alpha}$	

Assume population 1 has the larger variance. Thus always: $s^2_1 > s_2^2$

Two-samples: Variance

Example

An important measure to determine service delivery in the banking sector is the variability in the service times. An experiment was conducted to compare the service times of two bank tellers. The results from the experiment:

- Teller A: n_A = 18 and s_A^2= 4,03
- Teller B: n_B = 26 and s_B^2 = 9,49

Can we say that the variance in service time of teller A is less than that variance of teller B on a 5% level of significance.

Remember:
Population 1 has the larger variance.
Thus always: $s^2_1 > s^2_2$

We will then test if the variance in service time of teller B is more that the variation of teller A: $s_B^2 > s_A^2$

Two-samples: Paired Means

- From the data:
  - The results from the experiment:
    - Teller A: $n_A = 18$ and $s^2_A = 4{,}03$
    - Teller B: $n_B = 26$ and $s^2_B = 9{,}49$

$F_{26-1;18-1;0,05} = 2{,}18$

Accept H_0 | Reject H_0

- H_0: $S^2_1 = S^2_2$
- H_1: $S^2_1 > S^2_2$

$$F = \frac{S_1^2}{S_2^2} = \frac{9,49}{4,03} = 2,35$$

- Reject H_0

Variation of teller B is more than the variation of teller A on a 5% level of significance.

TABLE A-5 *(continued)* *F* Distribution (α = 0.05 in the right tail)

or degrees of freedom (df_2)	Numerator degrees of freedom (df_1)					
	10	12	15	20	24	30
1	241.88	243.91	245.95	248.01	249.05	250.10
2	19.396	19.413	19.429	19.446	19.454	19.462
3	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166
4	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459
5	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957
6	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082
7	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758
8	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794
9	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637
10	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996
11	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705
12	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663
13	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803
14	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082
15	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468
16	2.4935	2.4247	2.3522	2.2756	2.23[illegible]	2.1938
17	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477
18	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071
19	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712

Lab Time