VANDERBILT
School *of* Engineering

CE 3300-01 – RISK, RELIABILITY, AND RESILIENCE ENGINEERING

Hypothesis Testing

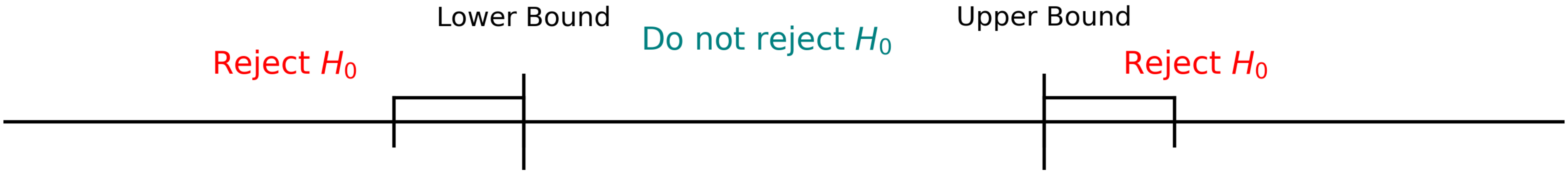
Book Reference: Chapters 6 – 6.3.1, 6.3.2 (examples 6.5 and 6.6)

TA: Mohamad Kazma

March 25th-26th, 2025

Instructor: Dr. Hiba Baroud

Today: Hypothesis Testing



Hypothesis Testing– Learning Objectives

- **Construct** a hypothesis test
- **Apply** the critical and p-value approaches
- **Test** the hypothesis for population parameters

Recall: How to estimate a population parameter?

Given a statistical model of a population, a **point estimate** is a value used to estimate a model parameter.

1. Select a random and representative sample.
2. Collect information from members of the sample.
3. Calculate the value of the sample statistic of interest.
4. Assign values to the population parameter.

How to “assign” values to the population parameter?

Point Estimate	Interval	Test
$\bar{x} \to \mu$ $x \sim \mathrm{Bin}(n, p)$ $\bar{x} = np = \mu$	Confidence level: $(1 - \alpha)\%$ Significance level: $\alpha\%$ Point Estimate $\pm$ margin of error [,] Confidence Interval	$\mu > 70$? $\mu \neq 0$?

Read: Chapter 6 – 6.2.1

What is a Hypothesis Testing?

A **hypothesis** is a statement regarding a parameter (or parameters) of the population.

Hypotheses

H_o**: null hypothesis**

- A statement about a population parameter that is assumed to be true until declared false

H_1 **: alternative hypothesis**

- A statement about a population parameter that will be assumed true if the null hypothesis is found to be false

Claim

e.g., Nashville officials claim a particular unemployment rate, does a sample verify this claim?

Decision

e.g., In a jury trial, we assume a person is innocent until guilt is proven

The Decision: Falling in the rejection region

Notes:
- NEVER "accept the null"
  - The null hypothesis is assumed to be true at first, we either reject it or we don't
- The sample data is the amount of evidence we have to make a decision

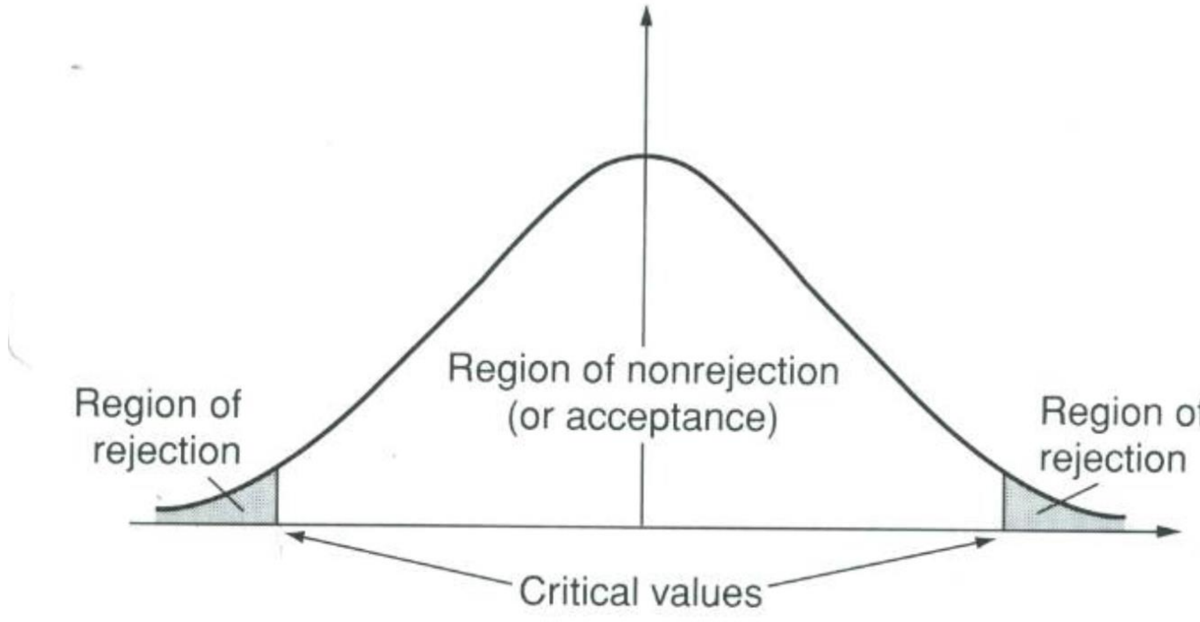


Figure 6.6 Regions of rejection and acceptance page: 260

More on the decision: The approach

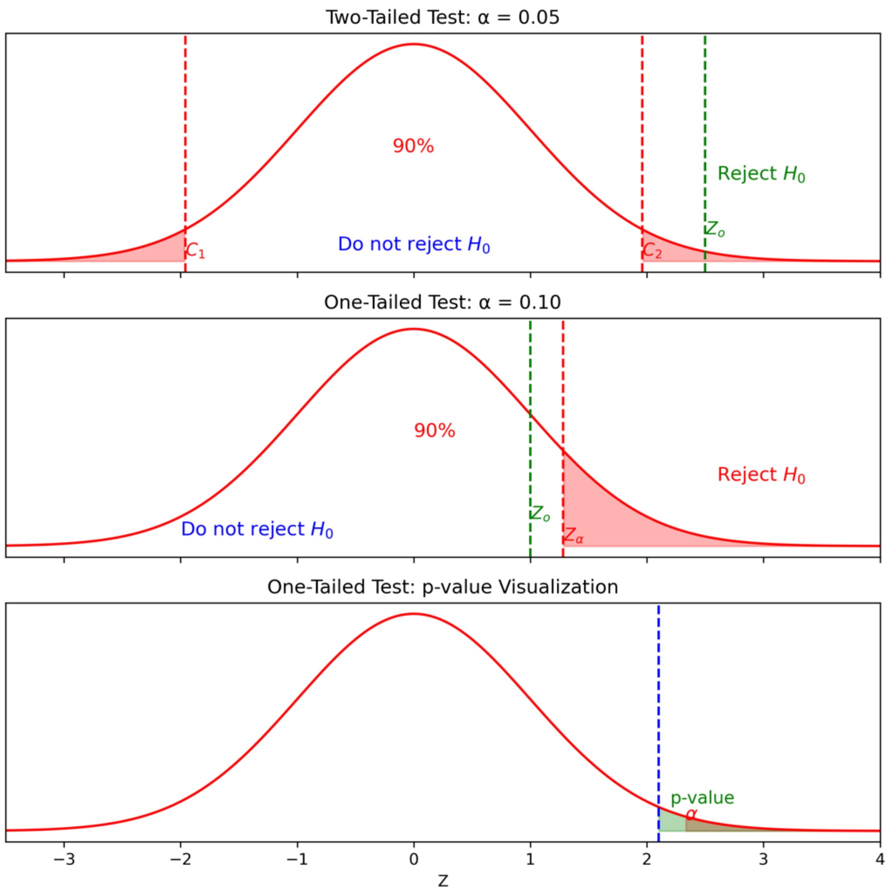
A **hypothesis** is a statement regarding a parameter (or parameters) of the population.

Critical value approach

- Determine c-value based on α
- Compare c-value and test statistic (Z_o)

p-value approach

- Determine the probability in the tail based on the test statistic
- Compare p-value and α



Critical value approach

1. State the null and alternative hypotheses
2. Select the distribution to use (similar conditions used in the confidence interval) and determine the rejection and non-rejection regions
3. Calculate the value of the test statistic
4. Determine the critical values (c-values)
5. Compare the test statistic and c-values
6. Make a decision

Test statistic – Hypothesis testing for the mean

- σ known

$$Z_o = \frac{\bar{x} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$$

- σ unknown

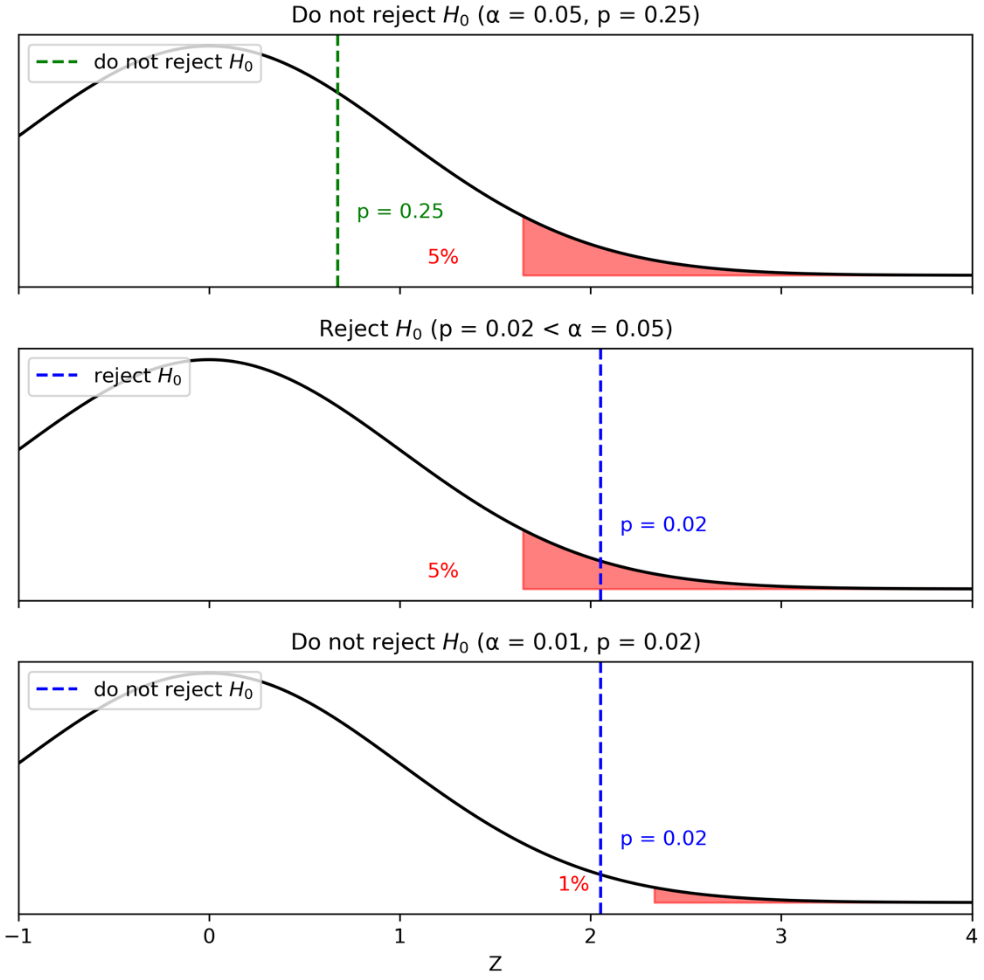
$$t_o = \frac{\bar{x} - \mu_o}{\frac{s}{\sqrt{n}}}$$

μ_o under H_o is True

$$H_o : \mu = 70 \quad H_o : \mu \leq 2$$
$$H_1 : \mu \neq 70 \quad H_1 : \mu > 2$$

p - value approach

1. State the null and alternative hypotheses
2. Select the distribution to use (*similar conditions used in the confidence interval*) and determine the rejection and non-rejection regions
3. Calculate the value of the test statistic
4. Calculate p-value
5. Compare p-value and α
6. Make a decision



Example 1 – Long-distance Calls

$$\sigma = 2.65 \quad \bar{x} = 13.71 \quad \alpha = 2\% \quad n = 150$$

A Telephone company provides long-distance telephone service in an area.

According to the company's records, the **average length** of all long-distance calls placed through this company **in 2004 was 12.44 minutes**.

The company **sampled 150 calls** and found a **mean length of 13.71 for this sample**. The **standard deviation** of all such calls is **2.65 minutes.**

Using the **2% significance level**, can you conclude that the **mean length of all current long-distance calls is different from 12.44 minutes**?

$$\mu = \mu_o \rightarrow \ H_o : \mu = 12.44\text{mns}$$

$$\mu \neq \mu_o \rightarrow \ H_1 : \mu \neq 12.44\text{mns}$$

Example 1 – Continued

$\sigma = 2.65 \quad \bar{x} = 13.71 \quad \alpha = 2\%$

$\mu = \mu_o \rightarrow \; H_o : \mu = 12.44\text{mns}$

$\mu \neq \mu_o \rightarrow \; H_1 : \mu \neq 12.44\text{mns}$

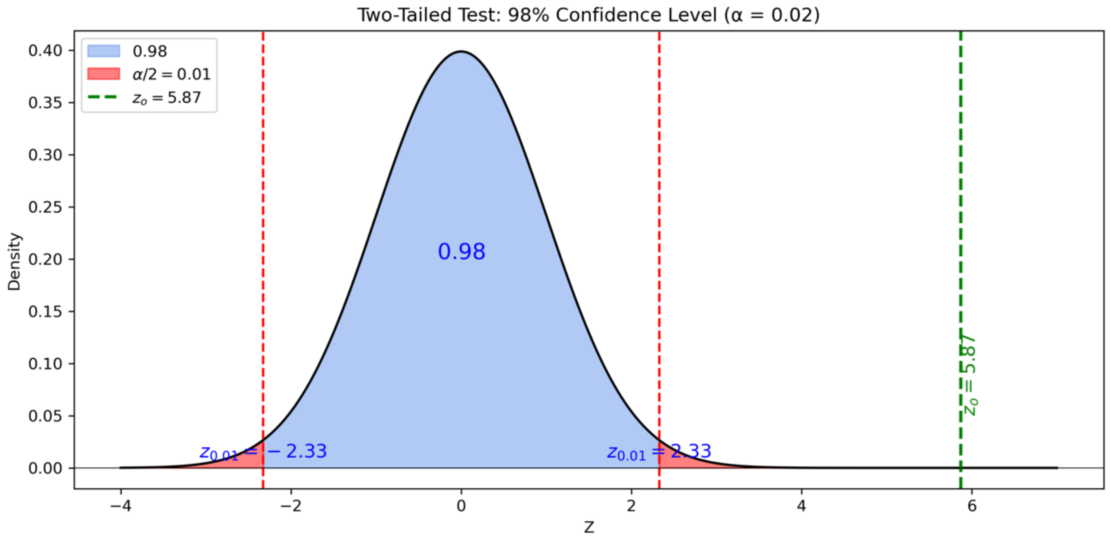
Z-dist because σ is known

$$\rightarrow z_0 = \frac{\overbrace{13.71}^{\bar{x}} - \overbrace{12.44}^{\mu_0}}{\overbrace{\frac{2.65}{\sqrt{\underbrace{150}_{n}}}}^{\sigma}} = 5.87 \rightarrow (z_0)$$

Find critical z-values for α = 0.02 (two-tailed)

Since $z_o = 5.87 > z_{0.01} = 2.33$,

$\rightarrow$ we reject the null hypothesis.



Reject the null hypothesis.
The average long-distance call is not equal to 12.44 minutes even though that was the average through 2004.

Example 1 – Continued

$$\sigma = 2.65 \quad \bar{x} = 13.71 \quad \alpha = 2\%$$

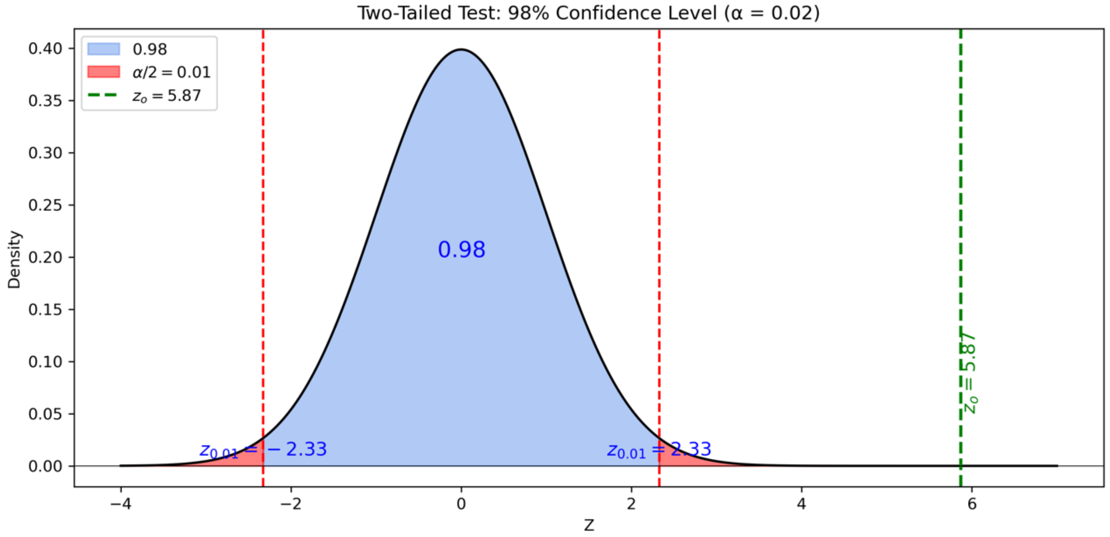
$$\mu = \mu_o \rightarrow \quad H_o : \mu = 12.44\text{mns}$$
$$\mu \neq \mu_o \rightarrow \quad H_1 : \mu \neq 12.44\text{mns}$$

Z-dist because σ is known

$$\rightarrow z_0 = \frac{\overbrace{13.71}^{\bar{x}} - \overbrace{12.44}^{\mu_0}}{\frac{\overbrace{2.65}^{\sigma}}{\sqrt{\underbrace{150}_{n}}}} = 5.87 \rightarrow (z_0)$$

Calculate p-value and compare to alpha

$$\rightarrow \ p\text{-value}$$
$$= 2 \times P\left(Z_o \geqslant 5.87\right)$$
$$= 2 \times 10^{-9}$$
$$\rightarrow p\text{-value} \ < \alpha$$



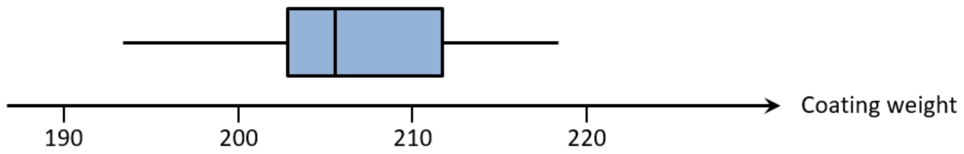
Example 2 – Pipe Coatings

A galvanized coating process is used to coat large pipes. The coating process could lead to too much weight on a pipe.

Production standards call for a **true average weight of 200 lb. per pipe.**
Based on the descriptive statistics of a **sample of 30 pipes**, is there a concern about the extra **weight at the 90% confidence level**?

$$\alpha = 10\% \qquad \bar{X} \sim T\left(\mu, \frac{s}{\sqrt{n}}\right)$$

Variable	N	Mean	Median	StDev
ctg wt	30	206.73	206.00	6.35



$$\rightarrow \quad \begin{array}{l} \mu \leq \mu_o \rightarrow H_o : \mu \leq 200 \\ \mu > \mu_o \rightarrow H_1 : \mu > 200 \end{array}$$

Example 2 – Continued

$$\alpha = 10\% \quad \bar{X} \sim T\left(\mu, \frac{s}{\sqrt{n}}\right)$$

$$\mu = \mu_o \rightarrow H_o : \mu \leq 200$$
$$\mu \neq \mu_o \rightarrow H_1 : \mu > 200$$

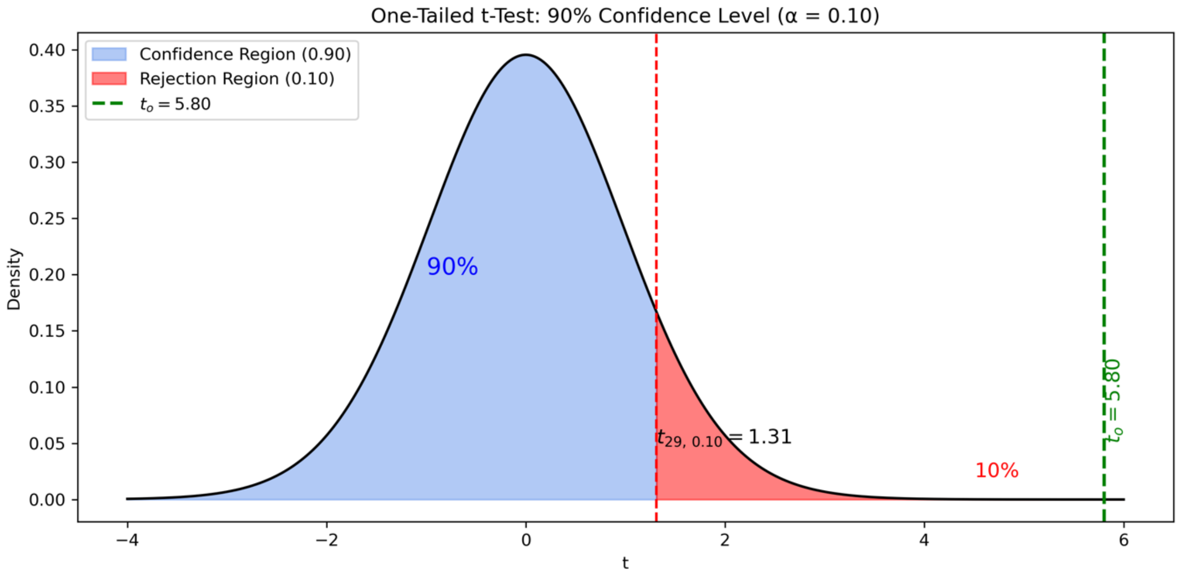
$\rightarrow$ t -dist because σ is unknown

$$t_o = \frac{206.73 - 200}{\frac{6.35}{\sqrt{30}}} = 5.8$$

Find critical t-values for α = 0.1 (one-tailed)

Since $t_o = 5.80 > 1.31$, we reject the null hypothesis.
There is evidence that the average pipe weight exceeds 200 lb .

Reject the null hypothesis.
There is concern about the extra weight.



Example 2 – Continued

Calculate p-value and compare to alpha

$\rightarrow$ p-value

$$= P(\bar{x} \geqslant 206.73$$

$$= P\left(T \geqslant t_0\right)$$

$$= P(T \geqslant 5.8)$$

T-distribution

$$P(T < 5.8) > 0.999$$

$$\Rightarrow \quad P\text{-value} \ < 0.0001$$

Using R

where $P(T < 5.8)$

$$1 - pt(5.2, 29) = 1.8e-06$$

TABLE A.3 Critical Values of t-Distribution at Confidence Level $(1-\alpha) = p$

d.o.f.	$p = 0.900$	$p = 0.950$	$p = 0.975$	$p = 0.990$	$p = 0.995$	$p = 0.999$
1	3.0777	6.3138	12.7062	31.8205	63.6567	318.3088
2	1.8856	2.9200	4.3027	6.9646	9.9248	22.3271
3	1.6377	2.3534	3.1824	4.5407	5.8409	10.2145
4	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732
5	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934
6	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076
7	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853
8	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008
9	1.3803	1.8331	2.2622	2.8214	3.2498	4.2968
10	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437
11	1.3634	1.7959	2.2001	2.7181	3.1058	4.0247
12	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296
13	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520
14	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874
15	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328
16	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862
17	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458
18	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105
19	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794
20	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518
21	1.3232	1.7207	2.0796	2.5176	2.8314	3.5272
22	1.3212	1.7171	2.0739	2.5083	2.8188	3.5050
23	1.3195	1.7139	2.0687	2.4999	2.8073	3.4850
24	1.3178	1.7109	2.0639	2.4922	2.7969	3.4668
25	1.3163	1.7081	2.0595	2.4851	2.7874	3.4502
26	1.3150	1.7056	2.0555	2.4786	2.7787	3.4350
27	1.3137	1.7033	2.0518	2.4727	2.7707	3.4210
28	1.3125	1.7011	2.0484	2.4671	2.7633	3.4082
29	1.3114	1.6991	2.0452	2.4620	2.7564	3.3962 $\leftarrow t_o = 5.8$
30	1.3104	1.6973	2.0423	2.4573	2.7500	3.3852
31	1.3095	1.6955	2.0395	2.4528	2.7440	3.3749
32	1.3086	1.6939	2.0369	2.4487	2.7385	3.3653
33	1.3077	1.6924	2.0345	2.4448	2.7333	3.3563
34	1.3070	1.6909	2.0322	2.4411	2.7284	3.3479
35	1.3062	1.6896	2.0301	2.4377	2.7238	3.3400
36	1.3055	1.6883	2.0281	2.4345	2.7195	3.3326
37	1.3049	1.6871	2.0262	2.4314	2.7154	3.3256
38	1.3042	1.6806	2.0244	2.4286	2.7116	3.3190
39	1.3036	1.6849	2.0227	2.4258	2.7079	3.3128
40	1.3031	1.6839	2.0211	2.4233	2.7045	3.3069
45	1.3006	1.6794	2.0141	2.4121	2.6896	3.2815
50	1.2987	1.6759	2.0086	2.4033	2.6778	3.2614
55	1.2971	1.6703	2.0040	2.3961	2.6682	3.2451
60	1.2958	1.6706	2.0003	2.3901	2.6603	3.2317
70	1.2938	1.6794	1.9944	2.3808	2.6479	3.2108
80	1.2922	1.6759	1.9901	2.3739	2.6387	3.1953
90	1.2910	1.6750	1.9867	2.3685	2.6316	3.1833
∞	1.2824	1.6449	1.9600	2.3264	2.5759	3.0903

Summary: Hypothesis Testing

Things to keep in mind:

- Equality is in the null
- Don't "accept" a null (either reject or do not reject)
- Setting up the hypotheses is critical

Hypothesis testing for the proportion

$$Z_o = \frac{\hat{p} - p_o}{\sqrt{\left(\frac{p_o(1-p_o)}{n}\right)}}$$

	Two-Tailed Test	Left-Tailed Test	Right-Tailed Test
Sign in the null hypothesis H_0	$=$	$=$ or $\geqq$	$=$ or $\leq$
Sign in the alternative hypothesis H_1	$\neq$	$<$	$>$
Rejection region	In both tails	In the left tail	In the right tail

Example 3 –Computer Sales (Left-tailed Test)

A company sells computers and computer parts by mail.

The company claims that **at least 90%** of all orders are mailed within **72 hours after they are received.**
The quality control department at the company often takes samples to check if this claim is valid. A recently taken **sample of 150 orders** showed that **129 of them were mailed within 72 hours.**
Do you think the company's claim is true? Or is the percentage of orders mailed within 72 hours less than 90%? Use a 2.5% significance level.

$$\alpha = 0.025, \quad n = 150,$$

$$\hat{p} = \frac{129}{150} = 0.86, \quad p_o = 0.9$$

$$\rightarrow$$

$$\mu \geq \mu_o \rightarrow H_o : \mu \geq 0.9$$

$$\mu < \mu_o \rightarrow H_1 : \mu < 0.9$$

Example 3 – Continued

$$\alpha = 0.025, \quad n = 150,$$
$$\hat{p} = \frac{129}{150} = 0.86, \quad p_o = 0.9$$

$\rightarrow$

$$\mu \geq \mu_o \rightarrow H_o : \mu \geq 0.9$$
$$\mu < \mu_o \rightarrow H_1 : \mu < 0.9$$

Find critical z-values for α = 0.025 (one-tailed)

$$\rightarrow z_0 = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o(1-p_o)}{n}}} = \frac{0.86 - 0.9}{\sqrt{\frac{0.9\times0.1}{150}}}$$
$$= -1.63$$

$\rightarrow$ For $\alpha = 0.025$, the left-tailed critical z-value is $z_\alpha = -1.96$
$z = -1.63 > -1.96 \Rightarrow$ Do NOT reject H_0

Calculate p-value and compare to alpha

$$\rightarrow p\text{-value} = P(z < -1.63)$$
$$= 1 - 0.948 = 0.0516$$

$\rightarrow$ p-value $> \alpha$

Do not reject H_0

Example 4 – Engine Water

The amount of wear (in 0.0001 inches) after a fixed mileage was determined for each of $n = 8$ internal combustion engines having copper lead as a bearing material, resulting in $\bar{x} = 3.72$ and $s = 1.25$.
Test at the 95% confidence level whether the average amount of wear is more than 3.50.

- Sample size: $n = 8$
- Sample mean: $\bar{x} = 3.72$
- Sample standard deviation: $s = 1.25$
- Test value: $\mu_0 = 3.50$ - Confidence level: $95\% \Rightarrow$ Significance level $\alpha = 0.05$
- Hypothesis type: One-tailed (right)

$$\rightarrow \begin{array}{l} \mu \leq \mu_o \rightarrow H_o : \mu \leq 3.5 \\ \mu > \mu_o \rightarrow H_1 : \mu > 3.5 \end{array}$$

Example 4 – Continued

$$\mu \leq \mu_o \rightarrow H_o : \mu \leq 3.5$$
$$\mu > \mu_o \rightarrow H_1 : \mu > 3.5$$

Calculate p-value and compare to alpha

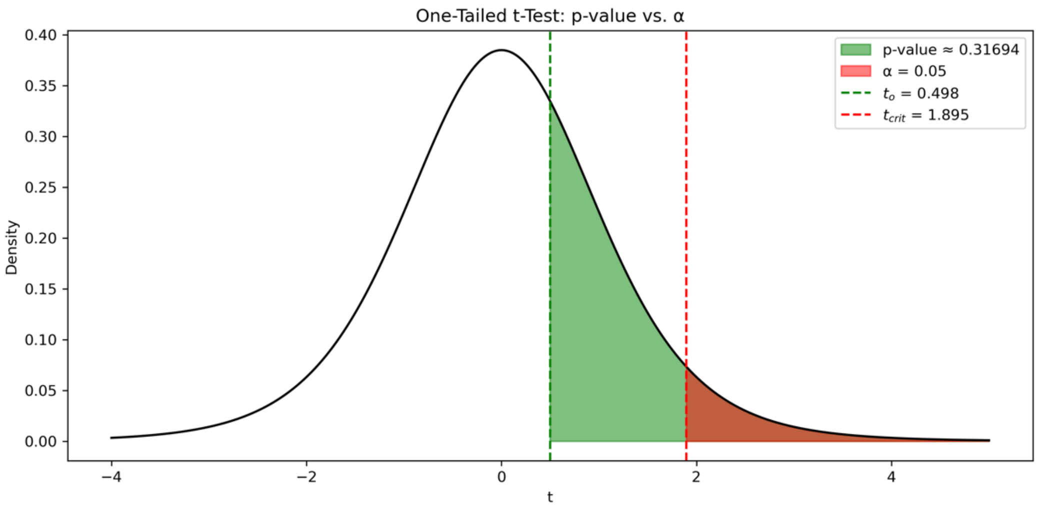
Small sample size, assume population follows a normal distribution
Use the t-distribution since σ is unknown.

$$t_0 = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{3.72 - 3.5}{1.25/\sqrt{8}} = 0.49$$

$$p - \text{ value } = P(T > 0.49)$$
$$= 1 - pt(0.49, 7)$$

R

$$df = n - 1 = 8 - 1 = 7$$



Accept the null hypothesis.

Do not reject the null hypothesis. Although the sample average is greater than 3.5, the true average amount of wear is less than or equal to 3.5.