LECTURE 12

CONFIDENCE INTERVAL AND HYPOTHESIS TESTING

INTERVAL ESTIMATION

- **Point estimation of θ :** The inference is a guess of a single value as the value of θ. No accuracy associated with it.

- **Interval estimation for θ :** Specify an interval in which the unknown parameter, θ, is likely to lie. It contains measure of accuracy through variance.

INTERVAL ESTIMATION

- An interval with random end points is called a random interval. E.g.,

$$\Pr\left\{\frac{5\bar{X}}{8} \le \theta \le \frac{5\bar{X}}{3}\right\} = 0.95$$

$\left(\frac{5\bar{X}}{8}, \frac{5\bar{X}}{3}\right)$ is a random interval that contains the true value of θ with probability *0.95*.

- An interval $(l(x_1,x_2,\ldots,x_n),\ u(x_1,x_2,\ldots,x_n))$ is called a **100γ % confidence interval** (CI) for θ if

$$\Pr\left\{l\left(x_1,x_2,\cdots,x_n\right)\le\theta\le u\left(x_1,x_2,\cdots,x_n\right)\right\}=\gamma$$

where $0<\gamma<1$.

- The observed values $l(x_1,x_2,\ldots,x_n)$ is a **lower confidence limit** and $u(x_1,x_2,\ldots,x_n)$ is an **upper confidence limit**. The probability γ is called the **confidence coefficient** or the **confidence level**.

- If $\Pr(l(x_1,x_2,\ldots,x_n)\leq\theta)=\gamma$, then $l(x_1,x_2,\ldots,x_n)$ *is* called a one-sided lower ***100γ %*** **confidence limit** for θ.

- *If* $\Pr(\theta\leq u(x_1,x_2,\ldots,x_n))=\gamma$, then $u(x_1,x_2,\ldots,x_n)$ is called a one-sided upper ***100γ %*** **confidence limit** for θ.

APPROXIMATE CI USING CLT

- *Let* $X_1, X_2, \ldots, X_n$ be a r.s.
- By CLT,

Non-normal random sample

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} N(0,1)$$

The approximate 100(1−α)% random interval for μ:

$$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

The approximate 100(1 −α)% CI for μ:

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- Usually, σ is unknown. So, the approximate $100(1-\alpha)\%$ CI for μ:

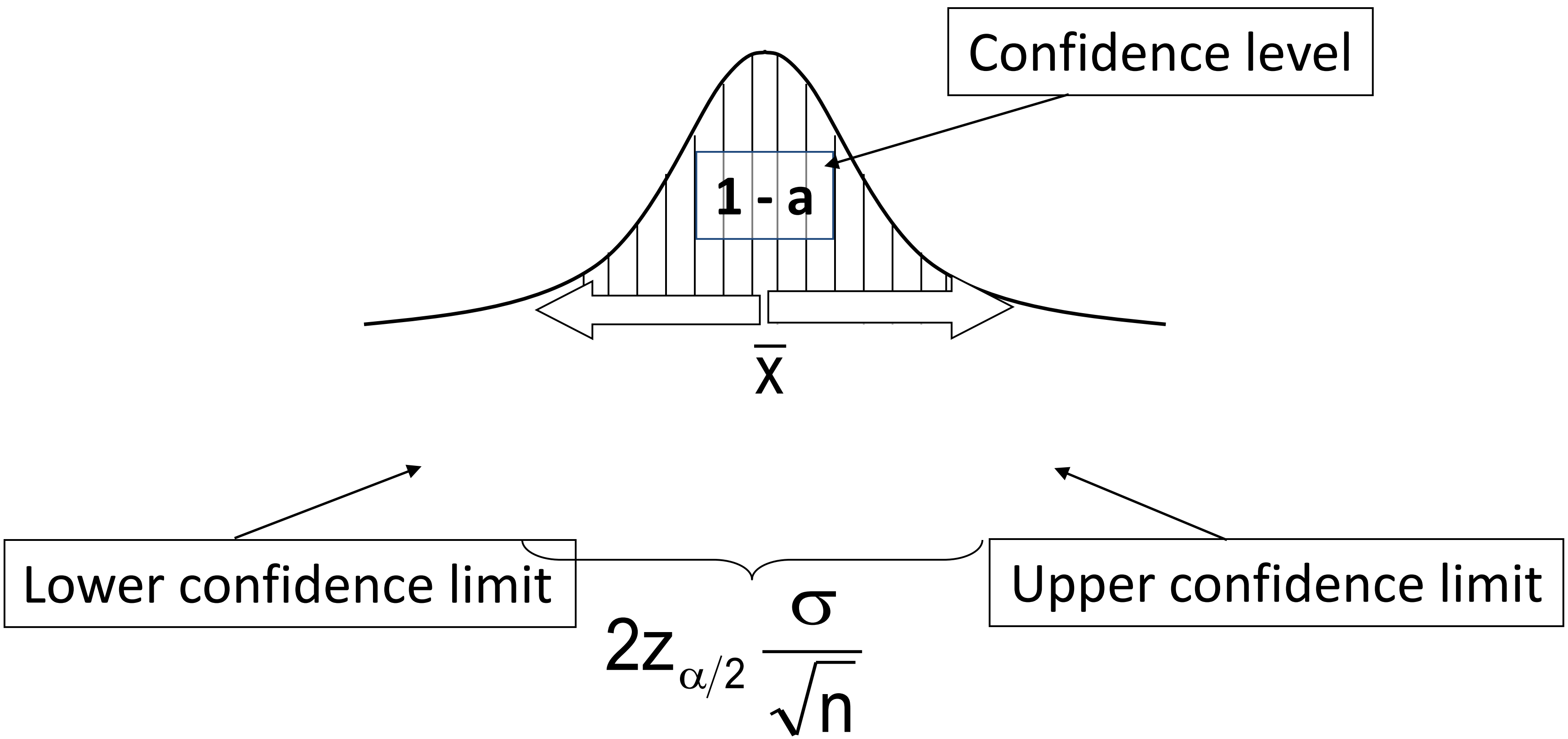
Non-normal random sample

$$\bar{x} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

- When the sample size $n \geq 30$, $t_{\alpha/2,n-1} \sim N(0,1)$.

$$\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}$$

GRAPHICAL DEMONSTRATION OF THE CONFIDENCE INTERVAL FOR μ



Example

Suppose that σ= 1.71 and n=100. Use a 90% confidence level.

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \bar{x} \pm 1.645\frac{1.71}{\sqrt{100}} = \bar{x} \pm .28$$

Recalculate the confidence interval for 95% confidence level.

$$\bar{x} \pm 1.96\frac{1.71}{\sqrt{100}} = \bar{x} \pm .34$$

The width of the 90% confidence interval = 2(.28) = .56
The width of the 95% confidence interval = 2(.34) = .68

Because the 95% confidence interval is wider, it is more likely to include the value of μ.

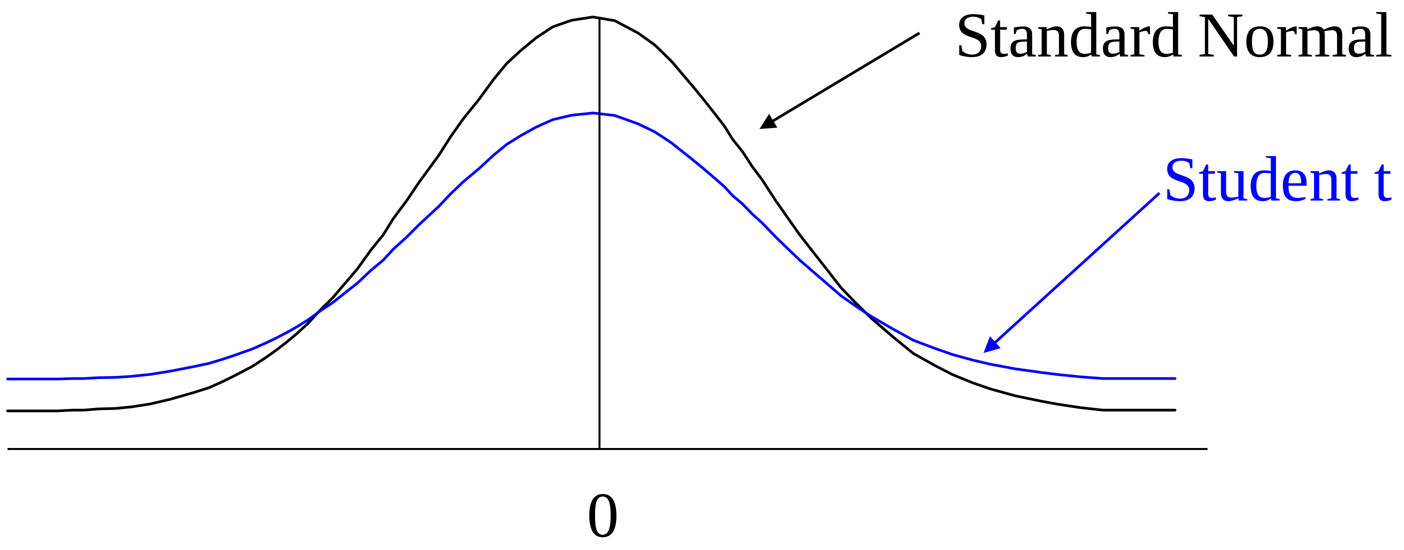
With 95% confidence interval, we allow ourselves to make 5% error; with 90% CI, we allow for 10%.

THE WIDTH OF THE CONFIDENCE INTERVAL

The width of the confidence interval is affected by

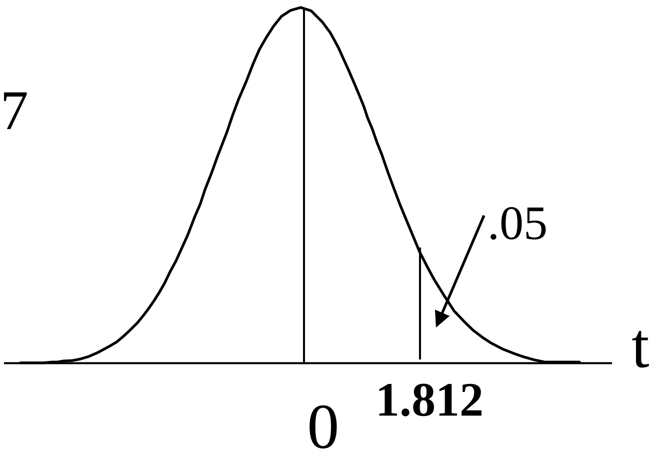
- the population standard deviation (σ)
  - ➢To maintain a certain level of confidence, a larger standard deviation requires a larger confidence interval.
- the confidence level ($1-\alpha$)
  - ➢Larger confidence level produces a wider confidence interval
- the sample size (n).
  - ➢Increasing the sample size decreases the width of the confidence interval while the confidence level can remain unchanged.

INFERENCE ABOUT THE POPULATION MEAN WHEN σ IS UNKNOWN



FINDING T-SCORES UNDER A T-DISTRIBUTION (T-TABLES)

Degrees of Freedom	$t_{.100}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055



$$t_{0.05,\,10} = 1.812$$

EXAMPLE

- A new breakfast cereal is test-marked for 1 month at stores of a large supermarket chain. The result for a sample of 16 stores indicate average sales of $1200 with a sample standard deviation of $180. Set up 99% confidence interval estimate of the true average sales of this new breakfast cereal. Assume normality.

$$n = 16, \bar{x} = \$1200, s = \$180, \alpha = 0.01$$

$$\Rightarrow t_{\alpha/2, n-1} = t_{0.005, 15} = 2.947$$

- 99% CI for μ:

$$\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} = 1200 \pm 2.947\frac{180}{\sqrt{16}} = 1200 \pm 132.6015$$

(1067.3985, 1332.6015)

With 99% confidence, the limits 1067.3985 and 1332.6015 cover the true average sales of the new breakfast cereal.

HYPOTHESIS TESTING

- A hypothesis is a statement about a population parameter.
- The goal of a hypothesis test is to decide which of two complementary hypothesis is true, based on a sample from a population.

- **STATISTICAL TEST:** The statistical procedure to draw an appropriate conclusion from sample data about a population parameter.
- **HYPOTHESIS:** Any statement concerning an unknown population parameter.
- **Aim of a statistical test:** test an hypothesis concerning the values of one or more population parameters.

NULL AND ALTERNATIVE HYPOTHESIS

- **NULL HYPOTHESIS=H_0**
  - E.g., a treatment has no effect or there is no change compared with the previous situation.
- **ALTERNATIVE HYPOTHESIS=H_A**
  - E.g., a treatment has a significant effect or there is development compared with the previous situation.

- **Sample Space, $\mathscr{A}$:** Set of all possible values of sample values $x_1, x_2, \ldots, x_n$. $(x_1, x_2, \ldots, x_n) \in \mathscr{A}$
- **Parameter Space, Ω:** Set of all possible values of the parameters.

$$\Omega = \Omega_0 \cup \Omega_1$$

$$H_0: \theta \in \Omega_0$$

$$H_1: \theta \in \Omega_1$$

- **Critical Region**, *C* is a subset of $\mathscr{A}$ which leads to rejection region of H_0.

$$\text{Reject } H_0 \text{ if } (x_1, x_2, \ldots, x_n) \in C$$

$$\text{Not Reject } H_0 \text{ if } (x_1, x_2, \ldots, x_n) \in C'$$

- A test is a rule which leads to a decision to **fail to reject** or **reject** H_0 on the basis of the sample information.

- **Test Statistic:** The sample statistic on which we base our decision to reject or not reject the null hypothesis.
- **Rejection Region:** Range of values such that, if the test statistic falls in that range, we will decide to reject the null hypothesis, otherwise, we will not reject the null hypothesis.
- If the hypothesis completely specify the distribution, then it is called a **simple hypothesis**. Otherwise, it is **composite hypothesis**. $\theta=(\theta_1, \theta_2)$

$$H_0: \theta_1 = 3 \Rightarrow f(x;3,\ \theta_2)$$
$$H_1: \theta_1 = 5 \Rightarrow f(x;5,\ \theta_2)$$

Composite Hypothesis

If θ_2 is known, simple hypothesis.

	H_0 is True	H_0 is False
Reject H_0	**Type I error** **P(Type I error) = α**	**Correct Decision**
Do not reject H_0	**Correct Decision** **$1-\alpha$**	**Type II error** **$1-\beta$** **P(Type II error) = β**

Tests are based on the following principle:
Fix α, minimize β.

$\Pi(\theta)$=Power function of the test for all $\theta \in \Omega$.

$= P(\text{Reject } H_0 | \theta) = P((x_1, x_2, \ldots, x_n) \in C | \theta)$

$$\underset{\theta\in\Omega_0}{\Pi(\theta)} = \mathrm{P}\left(\text{Reject } \mathrm{H}_0 \middle| H_0 \text{ is true}\right)$$

$$\rightarrow P\left(Type\ I\ error\right) = \alpha(\theta)$$

Type I error=Rejecting H_0 when H_0 is true

$$\alpha(\theta) \xrightarrow[\theta\in\Omega_0]{\max} \alpha \Rightarrow \text{max. prob. of Type I error}$$

$$\underset{\theta\in\Omega_1}{\Pi(\theta)} = \mathrm{P}\left(\text{Reject } \mathrm{H}_0 \middle| H_1 \text{ is true}\right)$$

$$\rightarrow 1 - P\left(\text{Not Reject } \mathrm{H}_0 \middle| H_1 \text{ is true}\right) = 1 - \beta(\theta)$$

$$\beta(\theta) \xrightarrow[\theta\in\Omega_1]{\max} \beta \Rightarrow max.\ prob.of\ Type\ II\ error$$

PROCEDURE OF STATISTICAL TEST

1. Determining H_0 and H_A.
2. Choosing the best test statistic.
3. Deciding the rejection region (Decision Rule).
4. Conclusion.

HYPOTHESIS TEST FOR POPULATION MEAN, μ

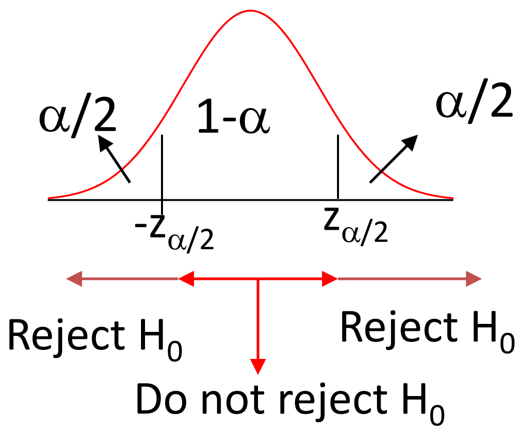
- **σ KNOWN AND X~N(μ, σ²) OR LARGE SAMPLE CASE:**

Two-sided Test | Test Statistic | Rejecting Area

$H_0: \mu = \mu_0$

$H_A: \mu \neq \mu_0$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$



- Reject H_0 if $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$.

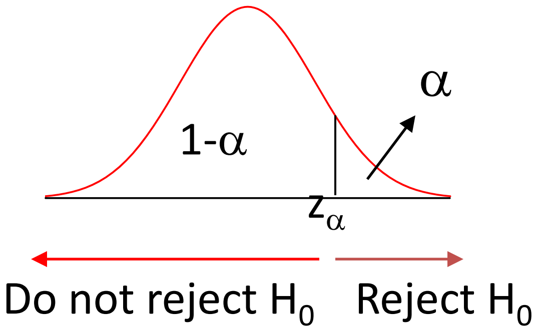
HYPOTHESIS TEST FOR POPULATION MEAN, μ

One-sided Tests | Test Statistic | Rejecting Area

1. $H_0: \mu = \mu_0$

   $H_A: \mu > \mu_0$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

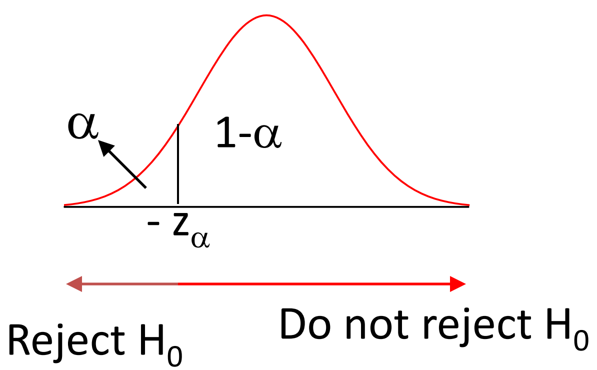


- Reject H_0 if $z > z_\alpha$.

2. $H_0: \mu = \mu_0$

   $H_A: \mu < \mu_0$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$



- Reject H_0 if $z < -z_\alpha$.

POWER OF THE TEST AND P-VALUE

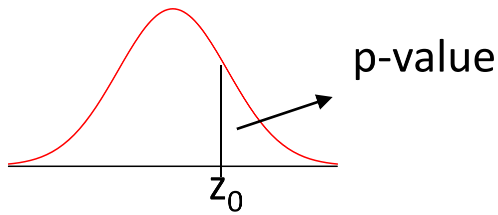
- $1-\beta$ = Power of the test

  = P(Reject $H_0 | H_0$ is not true)
- p-value = Observed significance level = Probability of obtaining a test statistics at least as extreme as the one that you observed by chance, OR, the smallest level of significance at which the null hypothesis can be rejected OR the maximum value of α that you are willing to tolerate.

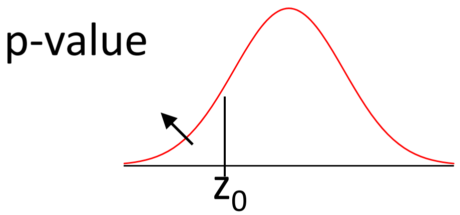
CALCULATION OF P-VALUE

- Determine the value of the test statistics, $z_0 = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$
- For One-Tailed Test:

p-value= $P(z > z_0)$ if H_A: $\mu>\mu_0$



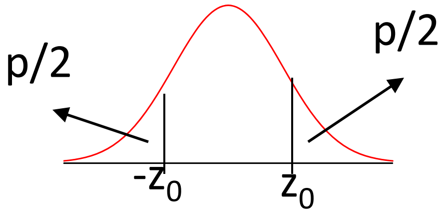
p-value= $P(z < z_0)$ if H_A: $\mu<\mu_0$



- For Two-Tailed Test

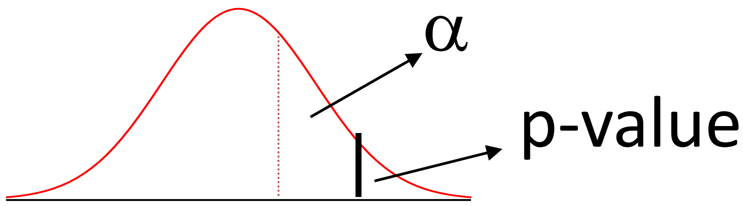
p=p-value = $2.P(z>z_0)$

p=p-value = $2.P(z<-z_0)$



DECISION RULE BY USING P-VALUES

- REJECT H_0 IF p-value $< \alpha$



- DO NOT REJECT H_0 IF p-value $\geq \alpha$

Example

- Do the contents of bottles of catsup have a net weight below an advertised threshold of 16 ounces?
- To test this 25 bottles of catsup were selected. They gave a net sample mean weight of $\bar{X} = 15.9$. It is known that the standard deviation is $\sigma = .4$ We want to test this at significance levels 1% and 5%.

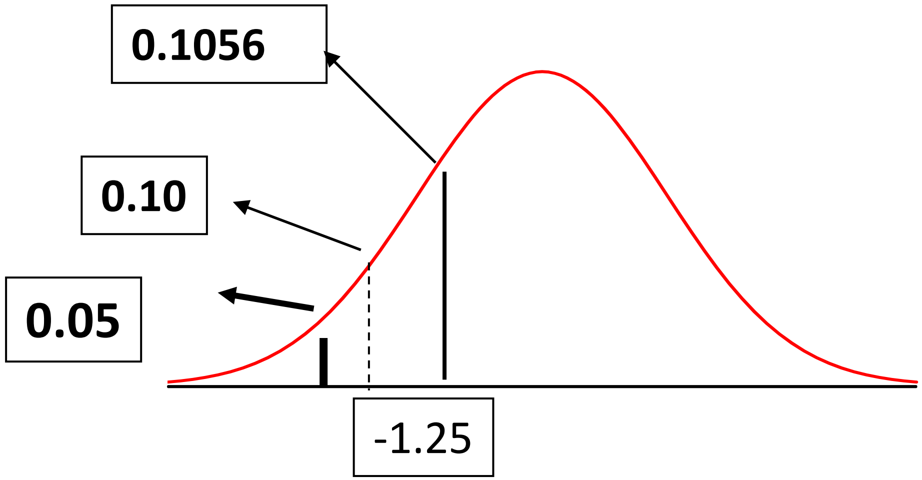
$$Z = \frac{15.9 - 16}{\left\{\frac{.4}{\sqrt{25}}\right\}} = -1.25$$

The p-value is the probability of getting a score worse than this (relative to the alternative hypothesis) i.e.,

$$P(Z < -1.25) = .1056$$

Compare the p-value to the significance level. Since it is bigger than both 1% and 5%, we do not reject the null hypothesis.

- The p-value for this test is 0.1056



- Thus, do not reject H_0 at 1% and 5% significance level. We do not have enough evidence to say that the contents of bottles of catsup have a net weight of less than 16 ounces.

TEST OF HYPOTHESIS FOR THE POPULATION MEAN (σ UNKNOWN)

- For samples of size n drawn from a Normal Population, the test statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Has a Student t-distribution with $n-1$ degrees of freedom

Example

5 measurements of the tar content of a certain kind of cigarette yielded 14.5, 14.2, 14.4, 14.3 and 14.6 mg per cigarette. Show the difference between the mean of this sample $\overline{x} = 14.4$ and the average tar content claimed by the manufacturer, μ=14.0, is significant at α=0.05.

$$s^2 = \frac{\sum_{i=1}^{5}(x_i - \overline{x})^2}{n-1} = \frac{(14.5-14.4)^2 + \ldots + (14.6-14.4)^2}{5-1} = 0.025$$

$$s = 0.158$$

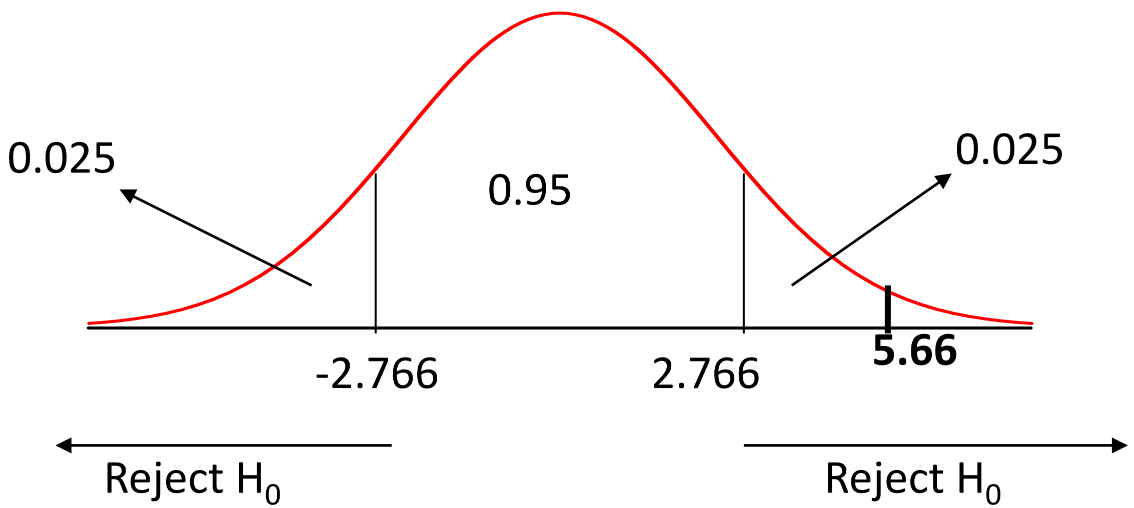
- $H_0: \mu = 14.0$
  $H_A: \mu \neq 14.0$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{14.4 - 14.0}{0.158/\sqrt{5}} = 5.66$$

$$t_{\alpha/2, n-1} = t_{0.025, 4} = 2.766$$

Decision Rule: Reject H_o if $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$.

Reject H_0 at $\alpha = 0.05$. Difference is significant



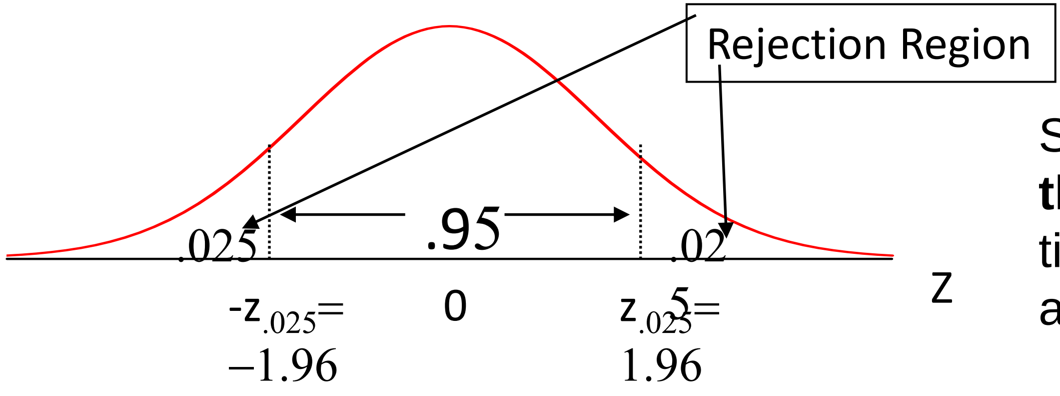
p-value = 2.P(t > 5.66) = 2(0.0024)=0.0048

Since p-value = 0.0048 < α = 0.05, reject H_0.

Example

At a certain production facility that assembles computer keyboards, the assembly time is known (from experience) to follow a normal distribution with mean (m) of 130 seconds and standard deviation (s) of 15 seconds. The production supervisor suspects that the average time to assemble the keyboards does not quite follow the specified value. To examine this problem, he measures the times for 100 assemblies and found that the sample mean assembly time ($\overline{X}$) is 126.8 seconds. Can the supervisor conclude at the 5% level of significance that the mean assembly time of 130 seconds is incorrect?

- We want to prove that the time required to do the assembly is different from what experience dictates: $\mathbf{H_A : \mu \neq 130}$
- The sample mean is $\mathbf{\bar{X} = 126.8}$
- The standard deviation is $\mathbf{\sigma = 15}$
- The standardized test statistic : $\mathbf{Z = \dfrac{126.8 - 130}{\left\{\dfrac{15}{\sqrt{100}}\right\}} = -2.13}$



Since –2.13<-1.96, **we reject the null hypothesis** that the time required to do the assembly is 130 seconds.

DECISION RULE

- Reject H_o if $z < -1.96$ or $z > 1.96$.

  In terms of $\bar{X}$, reject H_0 if

$$\bar{X} < 130 - 1.96\frac{15}{\sqrt{100}} = 127.6$$

$$or\ \bar{X} > 130 + 1.96\frac{15}{\sqrt{100}} = 132.94$$

$$\text{p} - \text{value} = 2.\text{P}(\text{Z} < -2.13) = 2(0.0166) = 0.0332$$

So, since 0.0332 < 0.05, we reject the null.

CALCULATING THE PROBABILITY OF TYPE II ERROR

H_o: $\mu = 130$

H_A: $\mu \neq 130$

- Suppose we would like to compute the probability of not rejecting H_0 given that the null hypothesis is false (for instance μ=135 instead of 130), i.e.

$$\beta = P(\text{not rejecting } H_o | H_o \text{ is false}).$$

Assuming μ=135 this statement becomes:

$$P(127.06 < \overline{x} < 132.94 \,|\, \mu = 135)$$
$$= P\left(\frac{127.06 - 135}{15/\sqrt{100}} < Z < \frac{132.94 - 135}{15/\sqrt{100}}\right)$$
$$= P(-5.29 < Z < -1.37) = .0853$$

TESTING HYPOTHESIS ABOUT POPULATION PROPORTION, p

ASSUMPTIONS:

1. The experiment is binomial.
2. The sample size is large enough.

x: The number of success

The sample proportion is

$$\hat{p} = \frac{x}{n} \sim N(p, \frac{pq}{n})$$

approximately for large n (np $\geq$ 5 and nq $\geq$ 5).

Two-sided Test

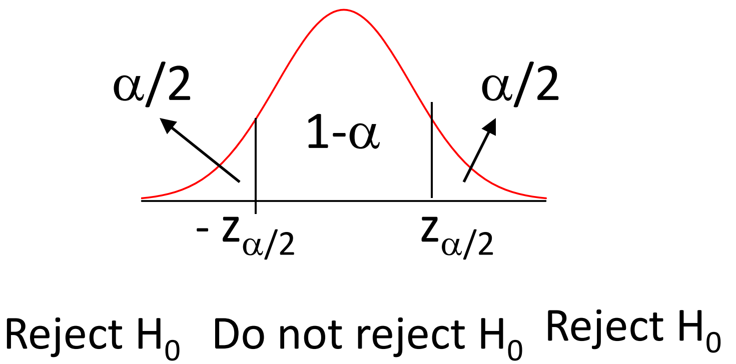
$H_0: p = p_0$

$H_A: p \neq p_0$

Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

Rejecting Area



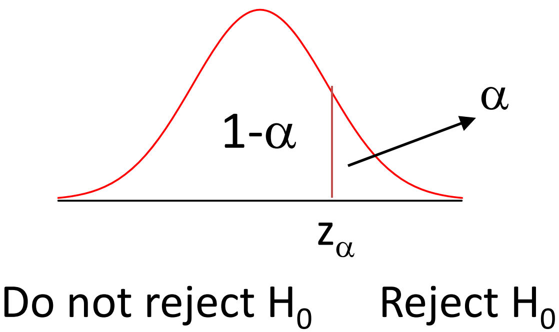
- Reject H_o if $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$.

One-sided Tests	Test Statistic	Rejecting Area

1. H_0: p= p_0

   H_A: p > p_0

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$

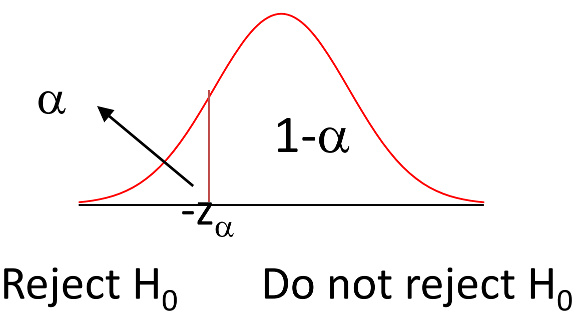


- Reject H_o if z > z_α.

2. H_0: p = p_0

   H_A: p < p_0

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}}$$



- Reject H_o if z < - z_α.

Example

- Mom's Home Cokin' claims that 70% of the customers are able to dine for less than \$5. Mom wishes to test this claim at the 92% level of confidence. A random sample of 110 patrons revealed that 66 paid less than \$5 for lunch.

$$H_o: p = 0.70$$

$$H_A: p \neq 0.70$$

- $x = 66$, $n = 110$ and $p = 0.70$

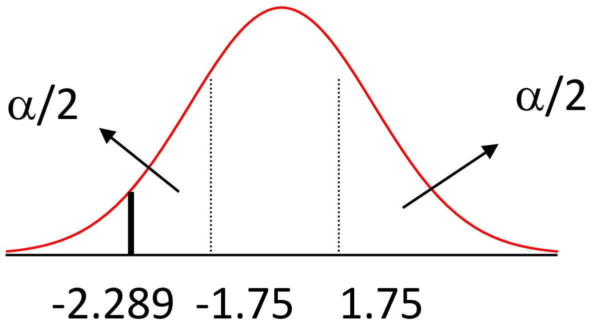
$$\Rightarrow \hat{p} = \frac{x}{n} = \frac{66}{110} = 0.6$$

- $\alpha = 0.08$, $z_{\alpha/2} = z_{0.04} = 1.75$
- Test Statistic:

$$z = \frac{0.6 - 0.7}{\sqrt{(0.7)(0.3)/110}} = -2.289$$

- DECISION RULE:

  Reject H_0 if $z < -1.75$ or $z > 1.75$.
- CONCLUSION: Reject H_0 at $\alpha = 0.08$. Mom's claim is not true.



- p-value = 2. $P(z < -2.289) = 2(0.011) = 0.022$

  The smallest value of α to reject H_0 is 0.022.

  Since p-value = $0.022 < \alpha = 0.08$, reject H_0.

CONFIDENCE INTERVAL APPROACH

- Find the 92% CI for p.

$$\hat{p} \mp z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.6 \mp 1.75\sqrt{\frac{(0.6)(0.4)}{110}}$$

92% CI for p: $0.52 \le p \le 0.68$

- Since $p = 0.7$ is not in the above interval, reject H_0. Mom has overestimated the percentage of customers that pay less than 5$ for a meal.

INFERENCE ABOUT THE POPULATION VARIANCE (s^2)

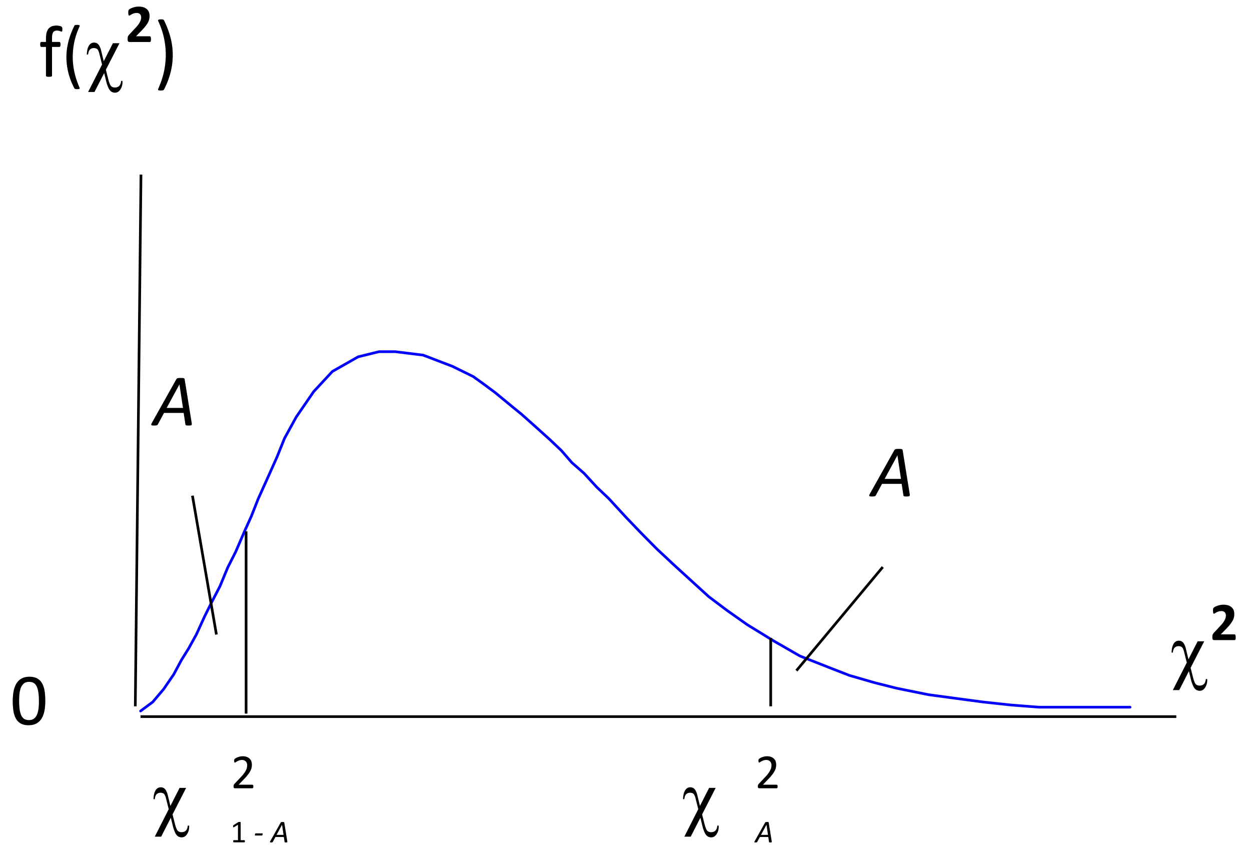
- Test statistic $$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$ Sampling distribution of S^2

which is chi-squared distributed with *n* - 1 degrees of freedom when the population random variable is normally distributed with variance σ^2.

Confidence interval estimator: **LCL** = $\frac{(n-1)\, s^2}{\chi^2_{\alpha/2}}$

UCL = $\frac{(n-1)\, s^2}{\chi^2_{1-\alpha/2}}$

CHI-SQUARE DISTRIBUTION



Example

- Proctor and Gamble told its customers that the variance in the weights of its bottles of Pepto-Bismol is **less than 1.2 ounces squared**. As a marketing representative for P&G, you select 25 bottles and find a variance of 1.7. At the 10% level of significance, is P&G maintaining its pledge of product consistency?

$$H_0: \sigma^2 = 1.2$$

$$H_A: \sigma^2 < 1.2$$

- n=25, s^2=1.7, α=0.10, $\chi^2_{0.90,24} = 15.659$
- **Test Statistics:**

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(24)1.7}{1.2} = 34$$

- **Decision Rule:** Reject H_0 if $\chi^2 < \chi^2_{\alpha,n-1} = 15.6587$
- **Conclusion:** Because χ^2=34 > 15.6587, do not reject H_0.
- We **don't have enough evidence** that suggests the variability in product weights less than 1.2 ounces squared.

Example

- A random sample of 22 observations from a normal population possessed a variance equal to 37.3. Find 90% CI for σ^2.

  90% CI for σ^2:

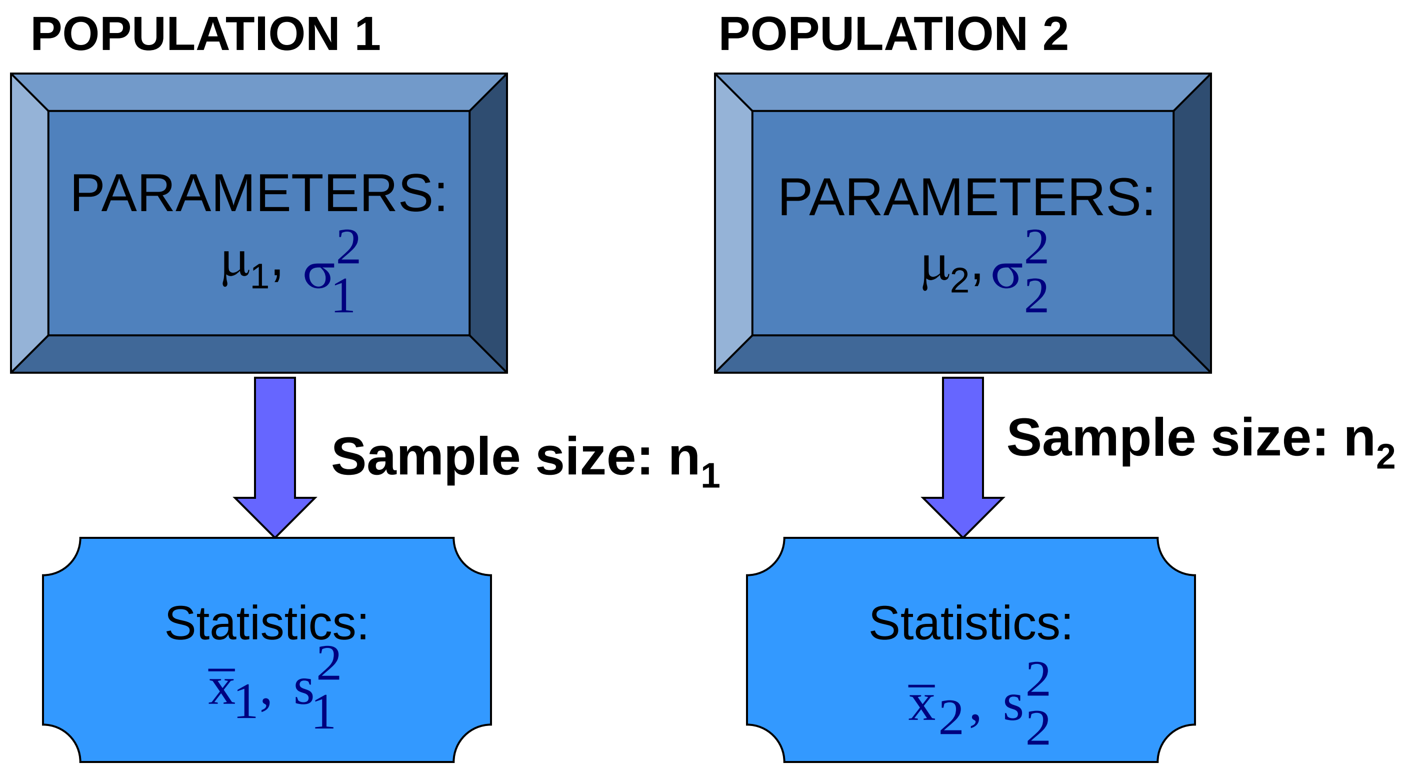
$$\frac{(n-1)s^2}{\chi^2_{0.05,21}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{0.95,21}}$$

$$\frac{(21)37.3}{32.6705} \le \sigma^2 \le \frac{(21)37.3}{11.5913}$$

$$23.9757 \le \sigma^2 \le 67.5765$$

INFERENCE ABOUT THE DIFFERENCE BETWEEN TWO SAMPLES

- **INDEPENDENT SAMPLES**



SAMPLING DISTRIBUTION OF $\bar{X}_1 - \bar{X}_2$

- Consider random samples of n_1 and n_2 from two normal populations. Then,

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

- For non-normal distributions, we can use Central Limit Theorem for $n_1 \geq 30$ and $n_2 \geq 30$.

CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$

- σ_1 and σ_2 are known for normal distribution or large sample $100(1-\alpha)\%$ C.I. for $\mu_1 - \mu_2$ is given by:

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

- If σ_1 and σ_2 are unknown and unequal, we can replace them with s_1 and s_2.

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

We are still using Z table because this is large sample or normal distribution situation.

Example

Is there any significant difference between mean family incomes of two groups?

$$n_1 = 200, \bar{x}_1 = 15530, s_1 = 5160$$

$$n_2 = 250, \bar{x}_2 = 16910, s_2 = 5840$$

Set up a 95% CI for μ_2 - μ_1. $z_{\alpha/2} = z_{0.025} = 1.96$

$$\bar{x}_2 - \bar{x}_1 = 16910 - 15530 = 1380$$

$$s^2_{\bar{x}_2 - \bar{x}_1} = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} = 269550 \Rightarrow s_{\bar{x}_2 - \bar{x}_1} = 519$$

95% CI for μ_2 - μ_1: $\left(\bar{x}_2 - \bar{x}_1\right) \pm 1.96(s_{\bar{x}_2 - \bar{x}_1})$ $363 \le \mu_2 - \mu_1 \le 2397$

With 95% confidence, mean family income in the second group may exceed that in the first group by between $363 and $2397.

TEST STATISTIC FOR $\mu_1 - \mu_2$ WHEN σ_1 AND σ_2 ARE KNOWN

- Test statistic:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- If σ_1 and σ_2 are unknown and unequal, we can replace them with s_1 and s_2.

Example

- Two different procedures are used to produce battery packs for laptop computers. A major electronics firm tested the packs produced by each method to determine the number of hours they would last before final failure.

$$n_1 = 150, \bar{x}_1 = 812\text{hrs}, s_1^2 = 85512$$
$$n_2 = 200, \bar{x}_2 = 789\text{hrs}, s_2^2 = 74402$$

- The electronics firm wants to know if there is a difference in the mean time before failure of the two battery packs. α=0.10

$H_0: \mu_1 = \mu_2 \Rightarrow H_0: \mu_1 - \mu_2 = 0$

$H_A: \mu_1 \neq \mu_2 \Rightarrow H_A : \mu_1 - \mu_2 \neq 0$

Test statistic:

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(812 - 789) - 0}{\sqrt{\frac{85512}{150} + \frac{74402}{200}}} = 0.7493$$

<u>Decision Rule</u> = Reject H_0 if $z<-z_{\alpha/2}=-1.645$ or $z>z_{\alpha/2}=1.645$.

- Not reject H_0. There is not sufficient evidence to conclude that there is a difference in the mean life of the 2 types of battery packs.

σ_1 AND σ_2 ARE UNKNOWN IF $\sigma_1 = \sigma_2$

- A 100(1-α)% C.I. for $\mu_1 - \mu_2$ is given by:

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, n_1+n_2-2} \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

where

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

TEST STATISTIC FOR $\mu_1 - \mu_2$ WHEN $\sigma_1 = \sigma_2$ AND UNKNOWN

- Test Statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

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

Example

- The statistics obtained from random sampling are given as

$$n_1 = 8, \bar{x}_1 = 93, s_1 = 20$$

$$n_2 = 9, \bar{x}_2 = 129, s_2 = 24$$

- It is thought that $\mu_1 < \mu_2$. Test the appropriate hypothesis assuming normality with $\alpha = 0.01$.
- $n_1 < 30$ and $n_2 < 30 \Rightarrow$ t-test
- Because s_1 and s_2 are not much different from each other, use equal-variance t-test (More formally, we can test $Ho: \sigma^2_1 = \sigma^2_2$).

$$H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 < \mu_2 \ (\mu_1 - \mu_2 < 0)$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(7)20^2 + (8)24^2}{8 + 9 - 2} = 15$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{(93 - 129) - 0}{(\sqrt{15})\sqrt{\frac{1}{8} + \frac{1}{9}}} = -19.13$$

- Decision Rule: Reject H_0 if $t < -t_{0.01,8+9-2}$=-2.602
- Conclusion: Since t = -19.13 < $-t_{0.01,8+9-2}$=-2.602, reject H_0 at α = 0.01.

TEST STATISTIC FOR $\mu_1 - \mu_2$ WHEN $\sigma_1 \neq \sigma_2$ AND UNKNOWN

- Test Statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

with the degree of freedom

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

Example

- Does consuming high fiber cereals entail weight loss? 30 people were randomly selected and asked what they eat for breakfast and lunch. They were divided into those consuming and those not consuming high fiber cereals. The statistics are obtained as
- n_1=10; n_2=20

$$\bar{x}_1 = 595.8; \quad \bar{x}_2 = 661.1$$

$$s_1 = 35.7; \quad s_2 = 115.7$$

Because s_1 and s_2 are too different from each other and the population variances are not assumed equal, we can use a t statistic with degrees of freedom

$$df = \frac{\left\{\left(35.7^2/10\right)+\left(115.7^2/20\right)\right\}^2}{\left\{\frac{\left[35.7^2/10\right]^2}{10-1}+\frac{\left[115.7^2/20\right]^2}{20-1}\right\}} = 25.01$$

H_0: $\mu_1 - \mu_2 = 0$

H_A: $\mu_1 - \mu_2 < 0$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(598.8 - 661.1) - 0}{\sqrt{\frac{35.7^2}{30} + \frac{115.7^2}{30}}} = -2.31$$

Reject H_0 if $t < -t_{\alpha,df} = -t_{0.05,\ 25} = -1.708$.

Since $t = -2.31 < -t_{0.05,\ 25} = -1.708$, reject H_0 at $\alpha = 0.05$.

INFERENCE ABOUT THE DIFFERENCE OF TWO MEANS: MATCHED PAIRS EXPERIMENT

- Data are generated from matched pairs; not independent samples.
- Let X_i and Y_i denote the measurements for the i-th subject. Thus, (X_i, Y_i) is a matched pair observations.
- Denote $D_i = Y_i - X_i$ or $X_i - Y_i$.
- If there are n subjects studied, we have

$$D_1, D_2, \ldots, D_n.$$

Then,

$$\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n} \text{ and } s_D^2 = \frac{\sum_{i=1}^{n} D_i^2 - n\bar{D}^2}{n-1} \Rightarrow s_{\bar{D}}^2 = \frac{s_D^2}{n}$$

CONFIDENCE INTERVAL FOR $\mu_D = \mu_1 - \mu_2$

- A $100(1-\alpha)\%$ C.I. for $\mu_D = \mu_1 - \mu_2$ is given by:

$$\overline{x}_D \pm t_{\alpha/2,\, n-1} \frac{s_D}{\sqrt{n}}$$

- For $n \geq 30$, we can use z instead of t.

HYPOTHESIS TESTS FOR $\mu_D = \mu_1 - \mu_2$

- The test statistic for testing hypothesis about μ_D is given by

$$t = \frac{\bar{x}_D - \mu_D}{s_D / \sqrt{n}}$$

with degree of freedom n-1.

Example

- Sample data on attitudes before and after viewing an informational film.

Subject	Before	After	Difference
i	X_i	Y_i	$D_i=Y_i-X_i$
1	41	46.9	5.9
2	60.3	64.5	4.2
3	23.9	33.3	9.4
4	36.2	36	-0.2
5	52.7	43.5	-9.2
6	22.5	56.8	34.3
7	67.5	60.7	-6.8
8	50.3	57.3	7
9	50.9	65.4	14.5
10	24.6	41.9	17.3

$$\overline{x}_D = 7.64, s_D = 12,57$$

- 90% CI for $\mu_D = \mu_1 - \mu_2$:

$$\overline{x}_D \pm t_{\alpha/2,n-1} \frac{s_D}{\sqrt{n}} = 7.64 \pm 1.833 \frac{12.57}{\sqrt{10}}$$

$t_{0.05,\ 9}$

$$0.36 \le \mu_D = \mu_1 - \mu_2 \le 14.92$$

- With 90% confidence, the mean attitude measurement after viewing the film exceeds the mean attitude measurement before viewing by between 0.36 and 14.92 units.

Example

- How can we design an experiment to show which of two types of tires is better? Install one type of tire on one wheel and the other on the other (front) wheels. The average tire (lifetime) distance (in 1000's of miles) is: $\bar{X}_D = 4.55$ with a sample difference s.d. of $s_D = 7.22$
- There are a total of n=20 observations

H_0: μ_D=0

H_A:μ_D>0

- Test Statistics:

$$t = \frac{\bar{x}_D - \mu_D}{s_D / \sqrt{n}} = \frac{4.55 - 0}{7.22 / \sqrt{20}} = 2.82$$

Reject H_0 if $t > t_{.05,19}$=1.729,

Conclusion: Reject H_0 at α=0.05

INFERENCE ABOUT THE DIFFERENCE OF TWO POPULATION PROPORTIONS

- SAMPLING DISTRIBUTION OF $\hat{p}_1 - \hat{p}_2$
- A point estimator of p_1-p_2 is

$$\hat{p}_1 - \hat{p}_2 = \frac{x_1}{n_1} - \frac{x_2}{n_2}$$

- The sampling distribution of $\hat{p}_1 - \hat{p}_2$ is

$$\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2})$$

if $n_i p_i \geq 5$ and $n_i q_i \geq 5$, i=1,2.

STATISTICAL TESTS

- Two-tailed test

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\hat{q}\left\{\frac{1}{n_1} + \frac{1}{n_2}\right\}}} \quad \text{where } \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$H_o: p_1 = p_2$

$H_A: p_1 \neq p_2$

Reject H_0 if $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$.

- One-tailed tests

$H_o: p_1 = p_2$

$H_A: p_1 > p_2$

Reject H_0 if $z > z_{\alpha}$

$H_o: p_1 = p_2$

$H_A: p_1 < p_2$

Reject H_0 if $z < -z_{\alpha}$

Example

- A manufacturer claims that compared with his closest competitor, fewer of his employees are union members. Of 318 of his employees, 117 are unionists. From a sample of 255 of the competitor's labor force, 109 are union members. Perform a test at $\alpha = 0.05$.

  H_0: $p_1 = p_2$

  H_A: $p_1 < p_2$

$\hat{p}_1 = \frac{x_1}{n_1} = \frac{117}{318}$ and $\hat{p}_2 = \frac{x_2}{n_2} = \frac{109}{255}$, so pooled sample proportion is $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{117 + 109}{318 + 255} = 0.39$

Test Statistic:

$$z = \frac{(117/318 - 109/255) - 0}{\sqrt{(0.39)(1-0.39)\left(\frac{1}{318} + \frac{1}{255}\right)}} = -1.4518$$

- Reject H_0 if $z < -z_{0.05} = -1.96$.
- Because $z = -1.4518 > -z_{0.05} = -1.96$, not reject H_0 at $\alpha = 0.05$. Manufacturer is wrong. There is no significant difference between the proportions of union members in these two companies.

Example

- In a study, doctors discovered that aspirin seems to help prevent heart attacks. Half of 22,000 male participants took aspirin and the other half took a placebo. After 3 years, 104 of the aspirin and 189 of the placebo group had heart attacks. Test whether this result is significant.
- p_1: proportion of all men who regularly take aspirin and suffer from heart attack.
- p_2: proportion of all men who do not take aspirin and suffer from heart attack

$$\hat{p}_1 = .009455 = \frac{104}{11000}$$

$$\hat{p}_2 = .01718 = \frac{189}{11000};$$

$$\textit{Pooled} \text{ sample proportion: } \hat{p} = \frac{104+189}{11000+11000} = .01332$$

$$H_0: p_1 - p_2 = 0$$

$$H_A: p_1 - p_2 < 0$$

- Test Statistic:

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

$$= \frac{.009455 - .01718}{\sqrt{(.01332)(.98668)\left(\frac{1}{11,000} + \frac{1}{11,000}\right)}} = -5.02$$

Conclusion: Reject H_0 since p-value=$P(z<-5.02) \approx 0$

CONFIDENCE INTERVAL FOR p_1-p_2

A 100(1-α)% C.I. for $p_1 - p_2$ is given by:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = -.077 \pm 1.96 * .00156$$

$$-.08 < p_1 - p_2 < -.074$$

INFERENCE ABOUT COMPARING TWO POPULATION VARIANCES

- Sampling distribution of σ_1^2 / σ_2^2
- For independent r.s. from normal populations

$$\frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

- (1-α)100% CI for σ_1^2 / σ_2^2

$$\frac{s_1^2}{s_2^2}\left[\frac{1}{F(1-\alpha/2, n_1-1, n_2-1)}\right] < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\left[\frac{1}{F(\alpha/2, n_1-1, n_2-1)}\right]$$

STATISTICAL TESTS

- Two-tailed test

H_o: $\sigma_1^2 = \sigma_2^2$ (or $\frac{\sigma_1^2}{\sigma_2^2} = 1$) $\quad F = \frac{s_1^2}{s_2^2}$

H_A: $\sigma_1^2 \neq \sigma_2^2$

Reject H_0 if $F < F_{\alpha/2,n1-1,n2-1}$ or $F > F_{1-\alpha/2,n1-1,n2-1}$.

- One-tailed tests

H_o: $\sigma_1^2 = \sigma_2^2$

H_A: $\sigma_1^2 > \sigma_2^2$

Reject H_0 if $F > F_{1-\alpha, n1-1,n2-1}$

H_o: $\sigma_1^2 = \sigma_2^2$

H_A: $\sigma_1^2 < \sigma_2^2$

Reject H_0 if $F < F_{\alpha,n1-1,n2-1}$

Example

- A scientist would like to know whether group discussion tends to affect the *closeness* of judgments by property appraisers. Out of 16 appraisers, 6 randomly selected appraisers made decisions after a group discussion, and rest of them made their decisions without a group discussion. As a measure of *closeness* of judgments, scientist used variance.
- Hypothesis: Groups discussion will reduce the variability of appraisals.

	Appraisal values (in thousand $)	Statistics
With discussion	97, 101,102,95,98,103	n_1=6 s_1^2=9.867
Without discussion	118, 109, 84, 85, 100, 121, 115, 93, 91, 112	n_2=10 s_2^2=194.18

Ho: $\frac{\sigma_1^2}{\sigma_2^2} \geq 1$

H1: $\frac{\sigma_1^2}{\sigma_2^2} < 1$

$$F = \frac{s_1^2}{s_2^2} = \frac{9.867}{194.18} = 0.05081 < F(0.05,5,9) = 0.21$$

Reject Ho. Group discussion reduces the variability in appraisals.