STATISTICAL INFERENCE PART II

CONFIDENCE INTERVALS AND HYPOTHESIS TESTING

LOCATION PARAMETER

- Let $f(x)$ be any pdf. The family of pdfs $f(x-\mu)$ indexed by parameter μ, is called the location family with standard pdf $f(x)$ and μ is the location parameter for the family.
- μ is a location parameter for $f(x)$ iff the distribution of $X-\mu$ does not depend on μ.

LOCATION PARAMETER

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of a distribution with pdf (or pmf); $f(x; \mu)$; $\mu \in \Omega$. An estimator $t(x_1, \ldots, x_n)$ is defined to be a **location equivariant** iff

$$t(x_1 + c, \ldots, x_n + c) = t(x_1, \ldots, x_n) + c$$

for all values of $x_1, \ldots, x_n$ and a constant c.

- $t(x_1, \ldots, x_n)$ is **location invariant** iff

$$t(x_1 + c, \ldots, x_n + c) = t(x_1, \ldots, x_n)$$

for all values of $x_1, \ldots, x_n$ and a constant c.

SCALE PARAMETER

- Let $f(x)$ be any pdf. The family of pdfs $f(x/\sigma)/\sigma$ *for* $\sigma>0$, indexed by parameter σ, is called the scale family with standard pdf $f(x)$ and σ is the scale parameter for the family.
- σ is a scale parameter for $f(x)$ iff the distribution of X/σ does not depend on σ.

SCALE PARAMETER

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of a distribution with pdf (or pmf); $f(x;\ \sigma);\ \sigma \in \Omega$. An estimator $t(x_1, \ldots, x_n)$ is defined to be a **scale equivariant** iff

$$t(cx_1, \ldots, cx_n) = ct(x_1, \ldots, x_n)$$

for all values of $x_1, \ldots, x_n$ and a constant $c > 0$.

- $t(x_1, \ldots, x_n)$ is **scale invariant** iff

$$t(cx_1, \ldots, cx_n) = t(x_1, \ldots, x_n)$$

for all values of $x_1, \ldots, x_n$ and a constant $c > 0$.

LOATION-SCALE PARAMETER

- Let $f(x)$ be any pdf. The family of pdfs $f(x-\mu)/\sigma$ $for\ \sigma>0$, indexed by parameter (μ,σ), is called the location-scale family with standard pdf $f(x)$ and μ is a location parameter and σ is the scale parameter for the family.
- μ is a location parameter and σ is a scale parameter for $f(x)$ iff the distribution of $(X-\mu)/\sigma$ does not depend on μ and σ.

LOCATION-SCALE PARAMETER

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of a distribution with pdf (or pmf); $f(x;\ \sigma);\ \sigma \in \Omega$. An estimator $t(x_1, \ldots, x_n)$ is defined to be a **location-scale equivariant** iff

$$t(cx_1+d, \ldots, cx_n+d) = ct(x_1, \ldots, x_n)+d$$

for all values of $x_1, \ldots, x_n$ and a constant $c>0$.

- $t(x_1, \ldots, x_n)$ is **location-scale invariant** iff

$$t(cx_1+d, \ldots, cx_n+d) = t(x_1, \ldots, x_n)$$

for all values of $x_1, \ldots, x_n$ and a constant $c>0$.

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.

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

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

INTERVAL ESTIMATION

- 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**.

INTERVAL ESTIMATION

- 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.

METHODS OF FINDING PIVOTAL QUANTITIES

- **PIVOTAL QUANTITY METHOD:**

  If $Q=q(x_1,x_2,\ldots,x_n)$ is a r.v. that is a function of only $X_1,\ldots,X_n$ and θ, then Q is called a **pivotal quantity** if its distribution does not depend on θ or any other unknown parameters (nuisance parameters).

PIVOTAL QUANTITY METHOD

Theorem: Let $X_1, X_2, \ldots, X_n$ be a r.s. from a distribution with pdf $f(x;\theta)$ for $\theta \in \Omega$ and assume that an MLE (or ss) of $\theta, \hat{\theta}$ exists.

- If θ is a location parameter, then $Q = \hat{\theta} - \theta$ is a pivotal quantity.
- If θ is a scale parameter, then $Q = \hat{\theta}/\theta$ is a pivotal quantity.
- If θ_1 and θ_2 are location and scale parameters respectively, then

$$\frac{\hat{\theta}_1 - \theta_1}{\hat{\theta}_2} \text{ and } \frac{\hat{\theta}_2}{\theta_2}$$ are PQs for θ_1 and θ_2.

CONSTRUCTION OF CI USING PIVOTAL QUANTITIES

- If Q is a PQ for a parameter θ and if percentiles of Q say q_1 and q_2 are available such that

$$Pr\{q_1 \leq Q \leq q_2\} = \gamma,$$

Then for an observed sample $x_1, x_2, \ldots, x_n$; a $100\gamma\%$ confidence region for θ is the set of $\theta \in \Omega$ that satisfy $q_1 \leq q(x_1, x_2, \ldots, x_n; \theta) \leq q_2$.

EXAMPLE

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of $Exp(\theta)$, $\theta > 0$. Find a $100\gamma\%$ CI for θ. Interpret the result.

EXAMPLE

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of $N(\mu, \sigma^2)$. Find a $100\gamma\%$ CI for μ and σ^2. Interpret the results.

EXAMPLE

- Let $X_1, X_2, \ldots, X_n$ be a r.s. of $Uniform(\theta, 1)$, $0 < \theta < 1$.

a) Show that $Z = (X_{(1)} - 1)/(\theta - 1)$ is a PQ for θ where $X_{(1)}$ is the first order statistic.

b) Find a 90% CI for θ with equal tail probabilities.

APPROXIMATE CI USING CLT

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

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

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

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

The approximate 100(1 −α)% CI for θ :

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

APPROXIMATE CI USING CLT

- Usually, σ is unknown. So, the approximate 100(1–α)% CI for μ:

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

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

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

The Confidence Interval for μ (σ is known)

- This leads to the following equivalent statement

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

The confidence interval

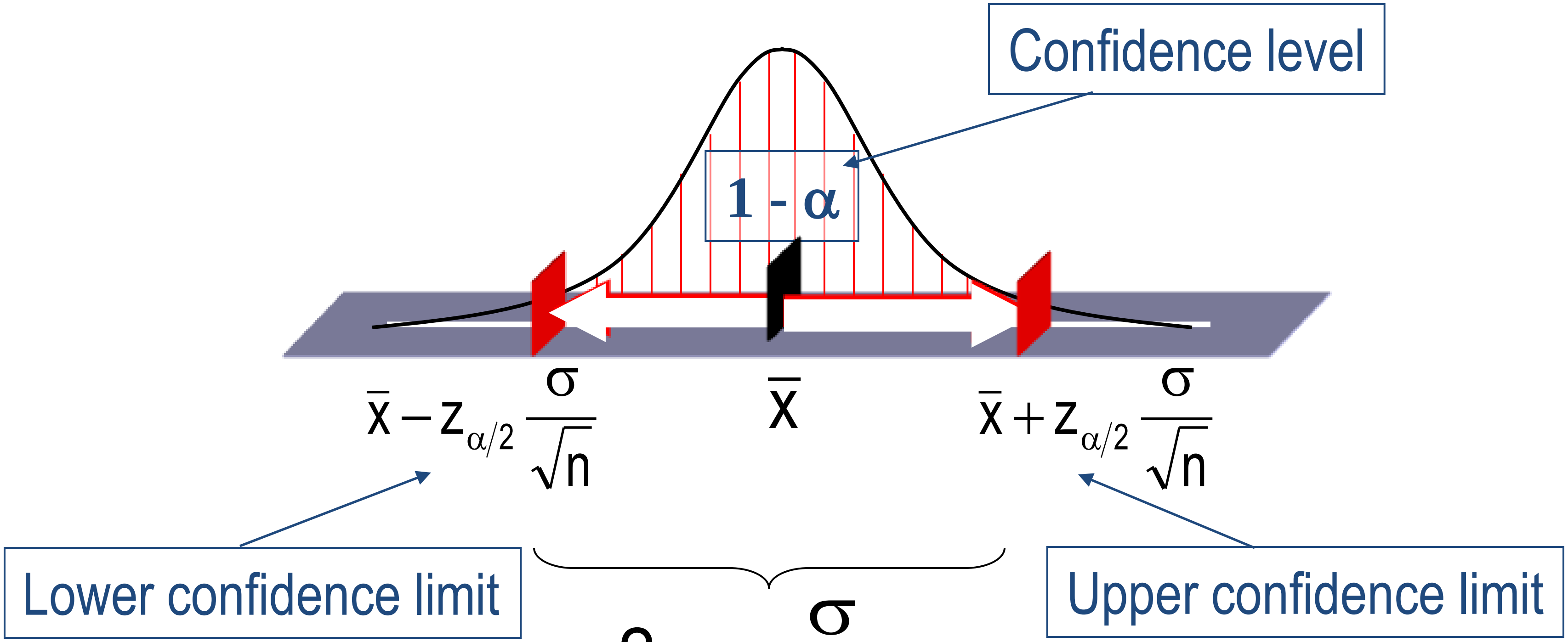
Interpreting the Confidence Interval for μ

$1-\alpha$ of all the values of $\bar{x}$ obtained in repeated sampling from a given distribution, construct an interval

$$\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

that includes (covers) the expected value of the population.

Graphical Demonstration of the Confidence Interval for μ



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

$$2z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

The Confidence Interval for μ (σ is known)

- **Example:** Estimate the mean value of the distribution resulting from the throw of a fair die. It is known that σ = 1.71. Use a 90% confidence level, and 100 repeated throws of the die

•**Solution:** The confidence interval is

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

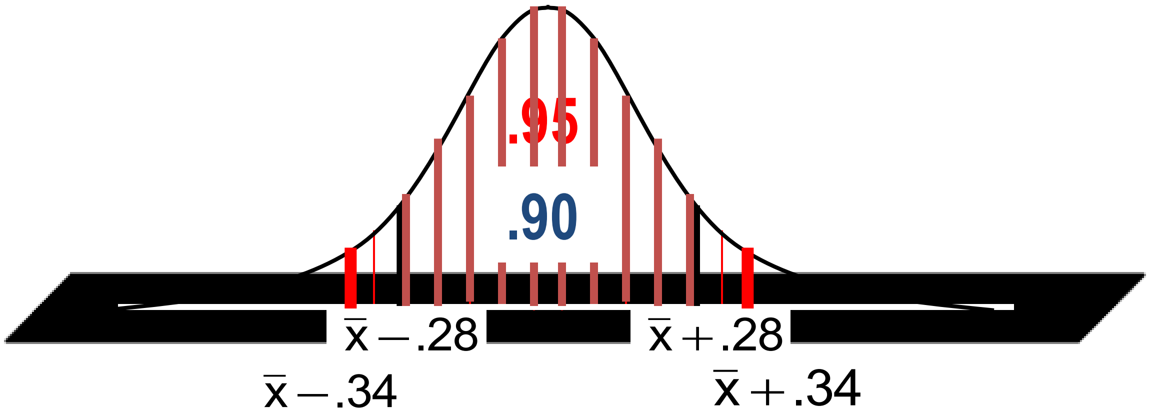
The mean values obtained in repeated draws of samples of size 100 result in interval estimators of the form

[sample mean - .28, Sample mean + .28],

90% of which cover the real mean of the distribution.

The Confidence Interval for μ (σ is known)

- Recalculate the confidence interval for 95% confidence level.
- Solution: $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = \bar{x} \pm 1.96 \frac{1.71}{\sqrt{100}} = \bar{x} \pm .34$



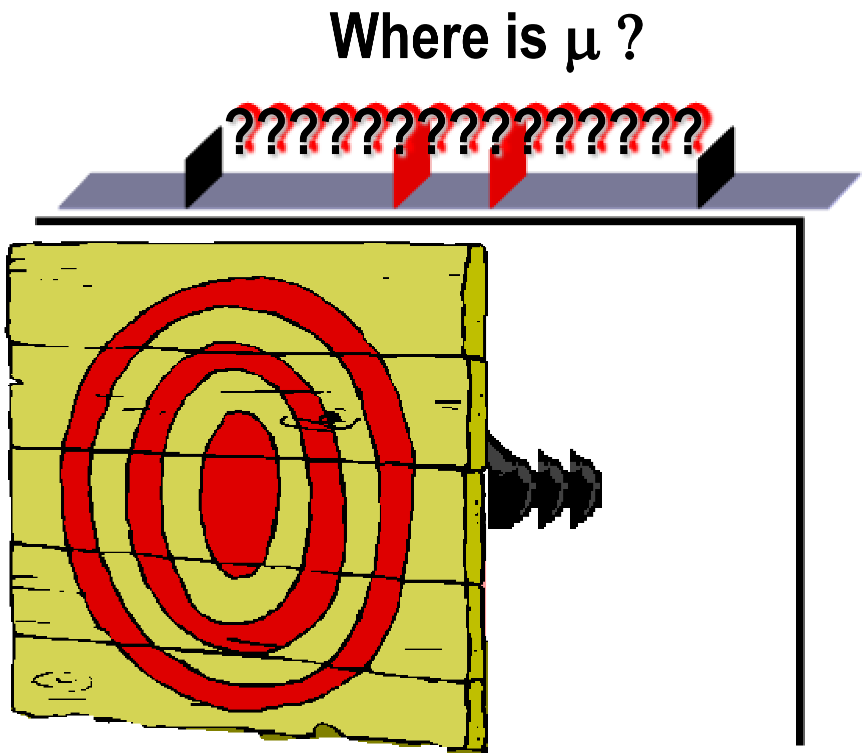
The Confidence Interval for μ (σ is known)

- 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.*

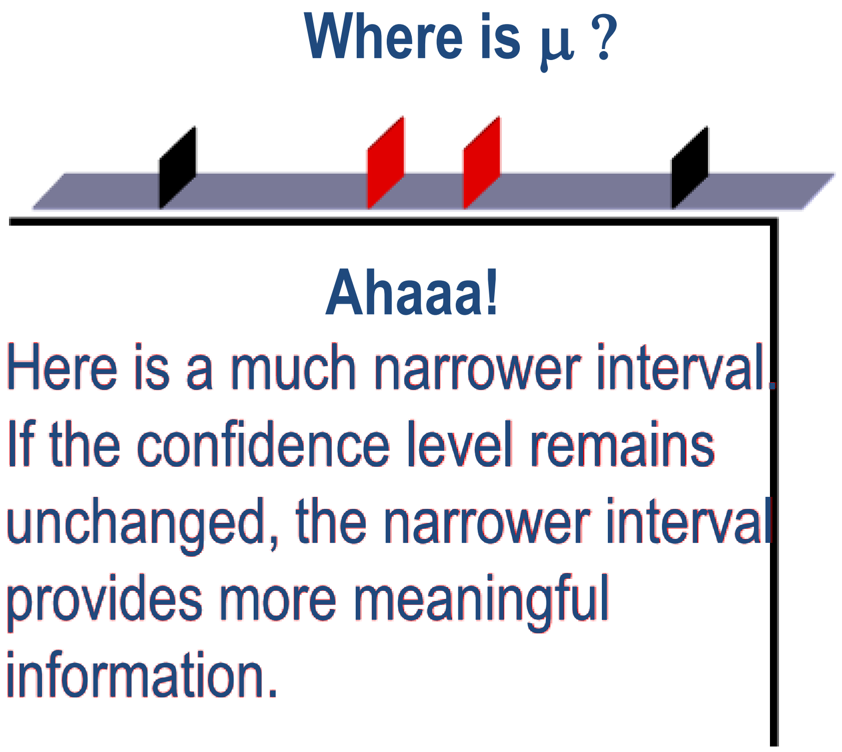
Information and the Width of the Interval

- Wide interval estimator provides little information.



Information and the Width of the Interval

– Wide interval estimator provides little information.



The Width of the Confidence Interval

The width of the confidence interval is affected by

- the population standard deviation (σ)
- the confidence level ($1-\alpha$)
- the sample size (n).

The Affects of σ on the interval width

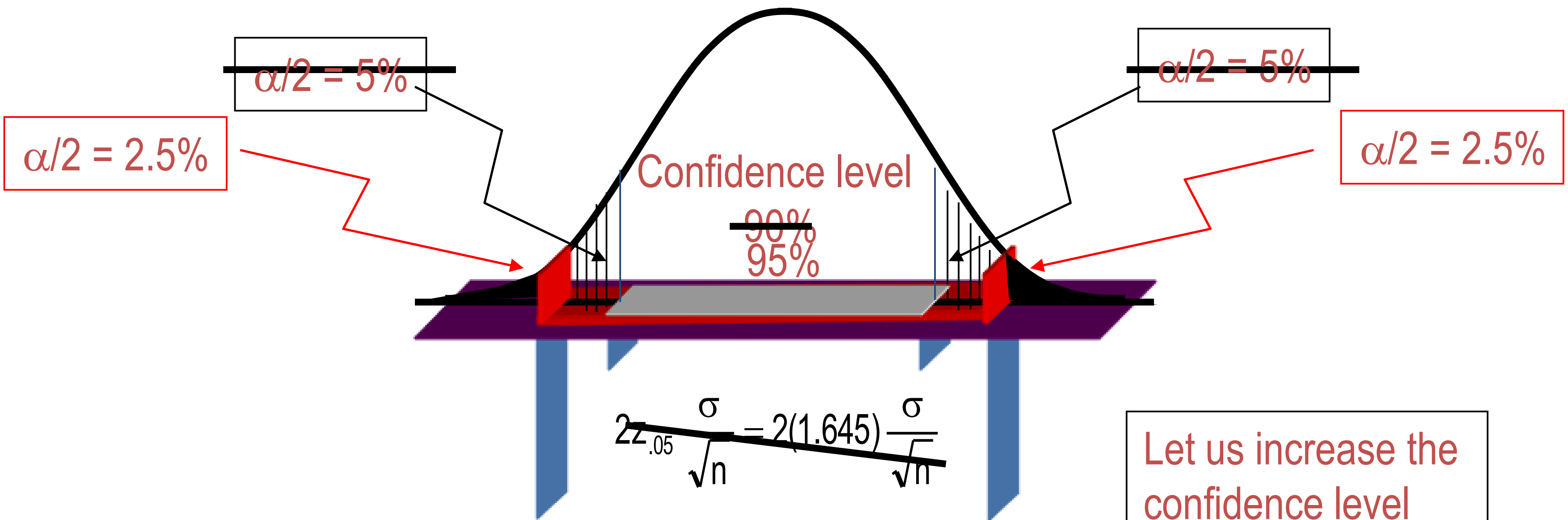
$$\cancel{2z_{.05}\frac{\sigma}{\sqrt{n}} = 2(1.645)\frac{\sigma}{\sqrt{n}}}$$

$$2z_{.05}\frac{1.5\sigma}{\sqrt{n}} = 2(1.645)\frac{1.5\sigma}{\sqrt{n}}$$

Suppose the standard deviation has increased by 50%.

To maintain a certain level of confidence, a larger standard deviation requires a larger confidence interval.

The Affects of Changing the Confidence Level



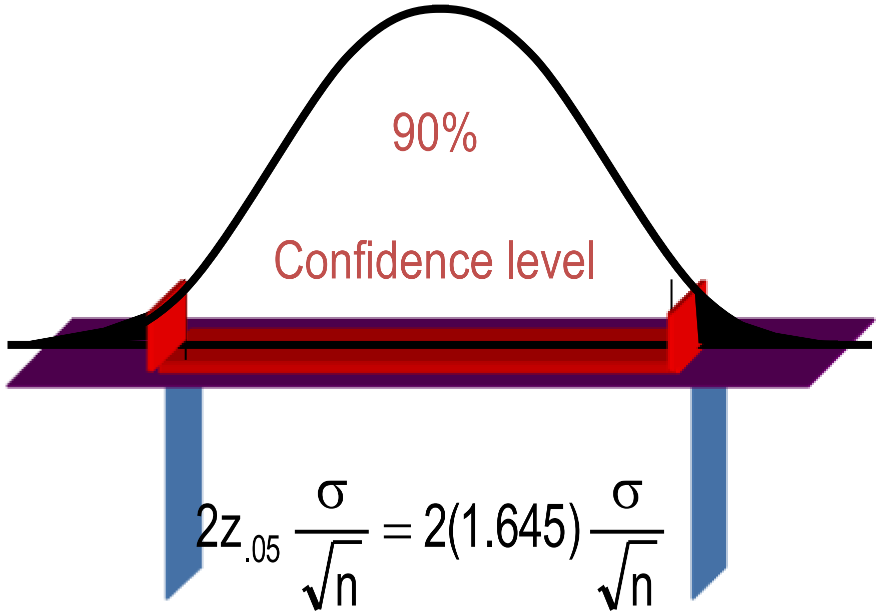
$$\cancel{2z_{.05}\frac{\sigma}{\sqrt{n}} = 2(1.645)\frac{\sigma}{\sqrt{n}}}$$

$$2z_{.025}\frac{\sigma}{\sqrt{n}} = 2(1.96)\frac{\sigma}{\sqrt{n}}$$

Let us increase the confidence level from 90% to 95%.

Larger confidence level produces a wider confidence interval

The Affects of Changing the Sample Size



$$2z_{.05}\frac{\sigma}{\sqrt{n}} = 2(1.645)\frac{\sigma}{\sqrt{n}}$$

Increasing the sample size decreases the width of the confidence interval while the confidence level can remain unchanged.

Selecting the Sample Size

- The required sample size to estimate the mean is

$$n = \left[\frac{z_{\alpha/2}\sigma}{w} \right]^2$$

Selecting the Sample Size

- Example
  - To estimate the amount of lumber that can be harvested in a tract of land, the mean diameter of trees in the tract must be estimated to within one inch with 99% confidence.
  - What sample size should be taken? Assume that diameters are normally distributed with $\sigma = 6$ inches.

Selecting the Sample Size

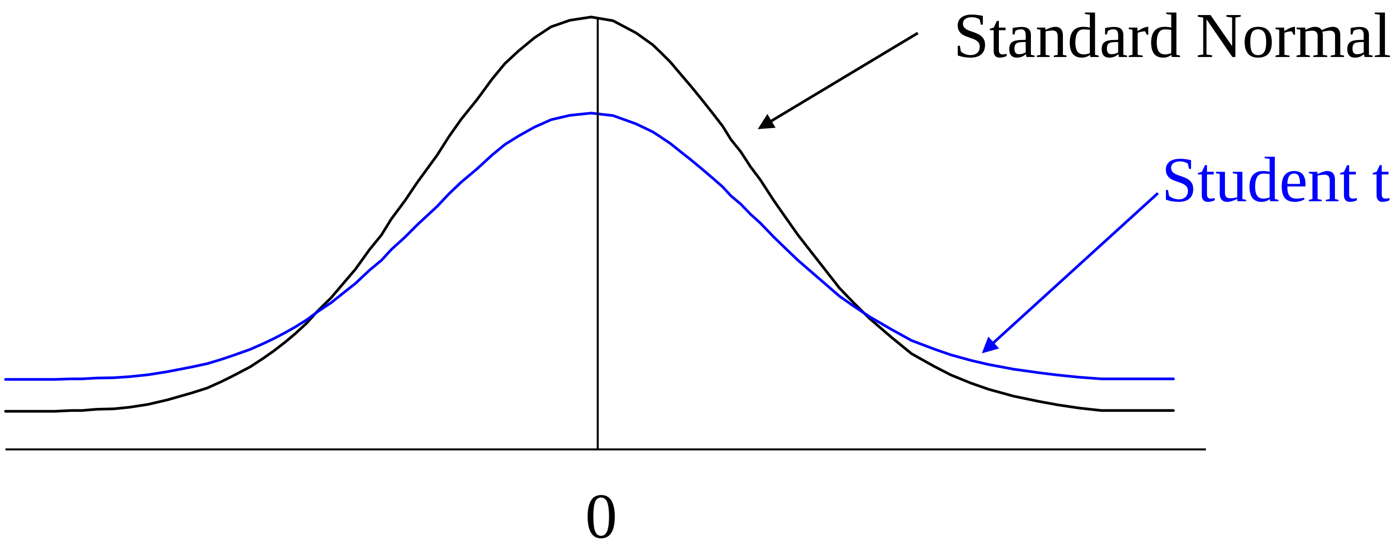
- Solution
  - The estimate accuracy is +/-1 inch. That is w = 1.
  - The confidence level 99% leads to α = .01, thus $z_{\alpha/2} = z_{.005} = 2.575$.
  - We compute

$$n = \left[\frac{z_{\alpha/2}\sigma}{w}\right]^2 = \left[\frac{2.575(6)}{1}\right]^2 = 239$$

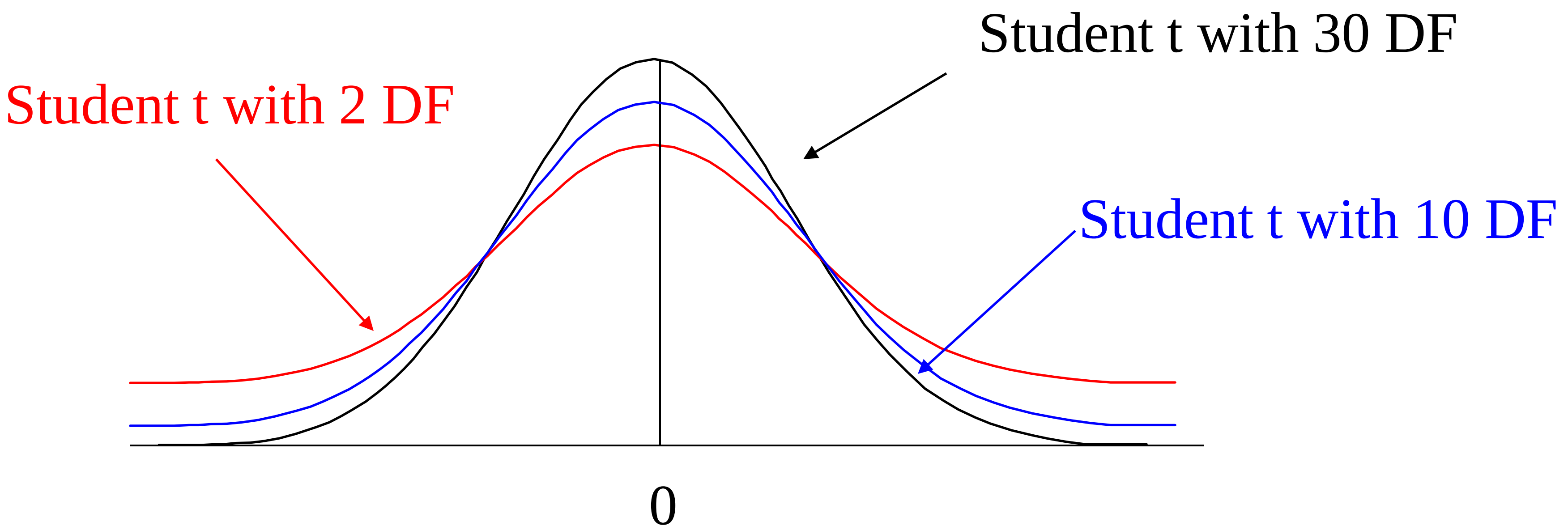
If the standard deviation is really 6 inches, the interval resulting from the random sampling will be of the form $\bar{x} \pm 1$ *If the standard deviation is greater than 6 inches the actual interval will be wider than +/-1.*

Inference About the Population Mean when σ is Unknown

- **The Student t Distribution**



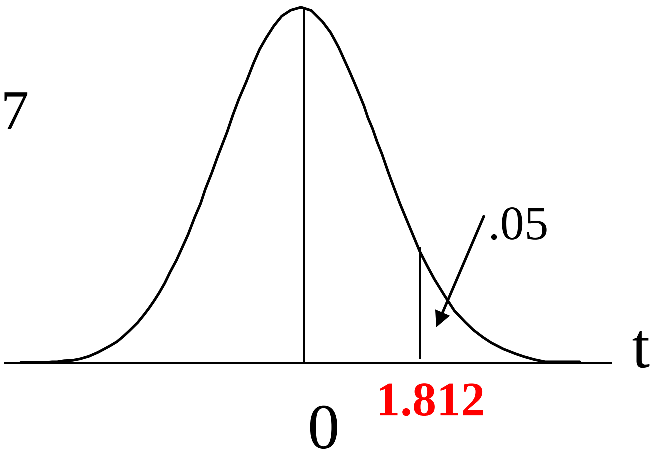
Effect of the Degrees of Freedom on the t Density Function



The "degrees of freedom", (a function of the sample size) determine how spread the distribution is compared to the normal distribution.

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



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

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.

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

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

ANSWER

- **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.

Example

- An investor is trying to estimate the return on investment in companies that won quality awards last year.
- A random sample of 83 such companies is selected, and the return on investment is calculated had he invested in them.
- Construct a 95% confidence interval for the mean return.

Solution (solving by hand)

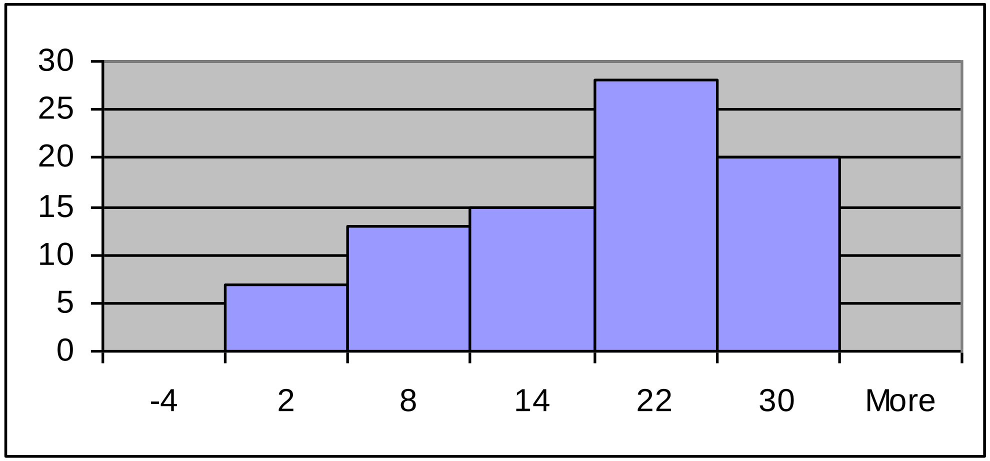
- The problem objective is to describe the population of annual returns from buying shares of quality award-winners.
- The data are interval.
- Solving by hand
  - From the data file we determine $\bar{x} = 15.02$ and $s = 8.31$

$$\bar{x} \pm t_{\alpha/2,n-1} \frac{s}{\sqrt{n}} \cong 15.02 \pm 1.990 \frac{8.31}{\sqrt{83}} = [13.19, 16.85]$$

$t_{.025,82} \cong t_{.025,80}$

Checking the required conditions

- We need to check that the population is normally distributed, or at least not extremely nonnormal.
- There are statistical methods to test for normality
- From the sample histograms we see…



TESTS OF HYPOTHESIS

- 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.

TESTS OF HYPOTHESIS

- **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** states that a treatment has no effect or there is no change compared with the previous situation. The parameter is equal to a single value.

ALTERNATIVE HYPOTHESIS=H_A states that a treatment has a significant effect or there is development compared with the previous situation. The parameter can be greater than or less than or different than the value shown in H_0.

TESTS OF HYPOTHESIS

- Sample Space, $\mathscr{A}$: Set of all possible values of sample values $x_1, x_2, \ldots, x_n$.

$$(x_1, x_2, \ldots, x_n) \in \mathscr{A}$$

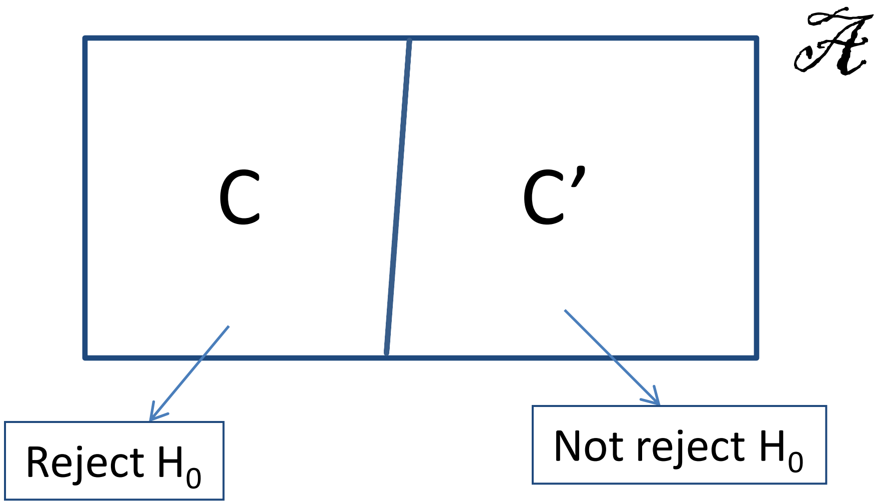
- Parameter Space, Ω: Set of all possible values of the parameters.

Ω=Parameter Space of Null Hypothesis$\cup$ Parameter Space of Alternative Hypothesis

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

TESTS OF HYPOTHESIS

- $\mathscr{A} = C \cup C'$



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

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

TESTS OF HYPOTHESIS

- 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 defines a critical region
- A test is a rule which leads to a decision to accept or reject H_0 on the basis of the sample information.

TEST STATISTIC AND REJECTION REGION

- **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. The probability that the (standardized) test statistic falls in the rejection region is the PROBABILITY OF TYPE I ERROR or SIGNIFICANCE LEVEL FOR THE TEST, which is known as α.

TESTS OF HYPOTHESIS

- If the hypothesis completely specify the distribution, then it is called a **simple hypothesis**. Otherwise, it is **composite hypothesis**.
- $\theta=(\theta_1, \theta_2)$

$$\left.\begin{aligned} H_0&:\theta_1=3 \Rightarrow f(x;3, \theta_2) \\ H_1&:\theta_1=5 \Rightarrow f(x;5, \theta_2) \end{aligned}\right\} \text{Composite Hypothesis}$$

If θ_2 is known, simple hypothesis.

TESTS OF HYPOTHESIS

	H_0 is True	H_0 is False
Reject H_0	**Type I error** **P(Type I error) = α**	**Correct Decision** **$1-\beta$**
Do not reject H_0	**Correct Decision** **$1-\alpha$**	**Type II error** **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(Reject $H_0|\theta$)=P($(x_1, x_2, \ldots, x_n) \in C|\theta$)

TESTS OF HYPOTHESIS

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

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

Type I error=Rejecting H_0 when H_0 is true

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

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

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

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

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.

POWER OF THE TEST AND P-VALUE

- α = Type I error = Significance level of the test. It measures the weight of the evidence favoring rejection of H_0.
- $1-\beta$ = Power of the test

  $= P(\text{Reject } H_0 \mid H_0 \text{ is not true})$
- p-value = Observed significance level = 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.

HYPOTHESIS TEST FOR POPULATION MEAN, μ

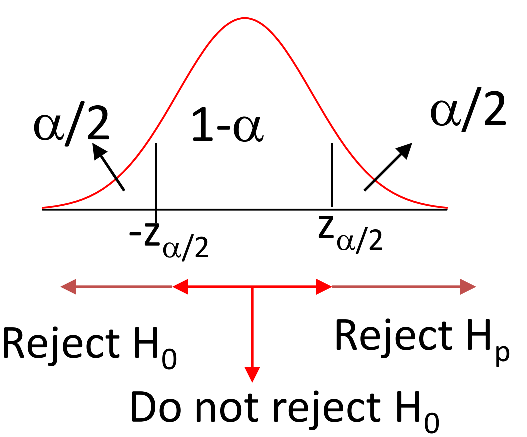
- **σ KNOWN AND X~N(μ, σ²) OR LARGE SAMPLE CASE:**

Two-sided Test | Test Statistic | Rejecting Area

$H_0: \mu = \mu_0$

$H_A: \mu \neq \mu_0$

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



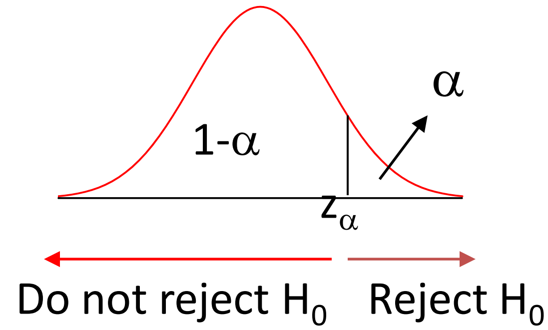
- Reject H_o if $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$.

HYPOTHESIS TEST FOR POPULATION MEAN, μ

One-sided Tests	Test Statistic	Rejecting Area

1. $H_0: \mu = \mu_0$
   $H_A: \mu > \mu_0$

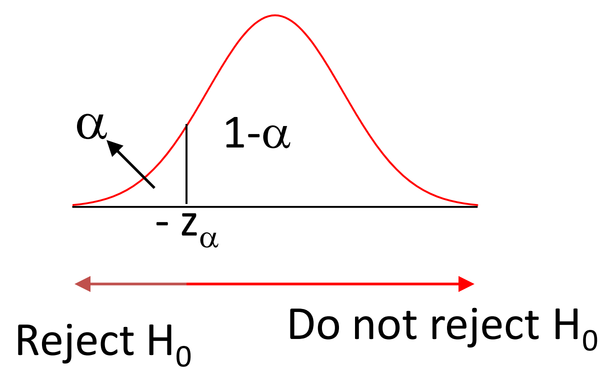
$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$



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

2. $H_0: \mu = \mu_0$
   $H_A: \mu < \mu_0$

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

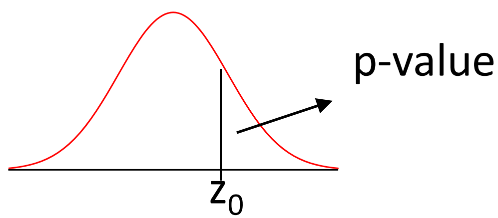


- Reject H_o if $z < - z_\alpha$.

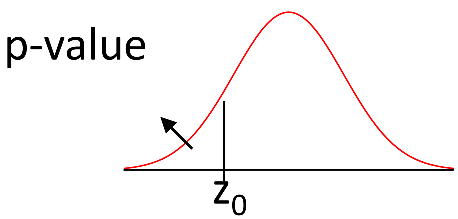
CALCULATION OF P-VALUE

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

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



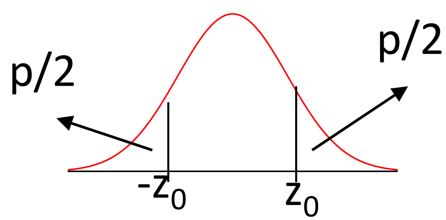
p-value= $P(z < z_0)$ if H_A: $\mu < \mu_0$



- For Two-Tailed Test

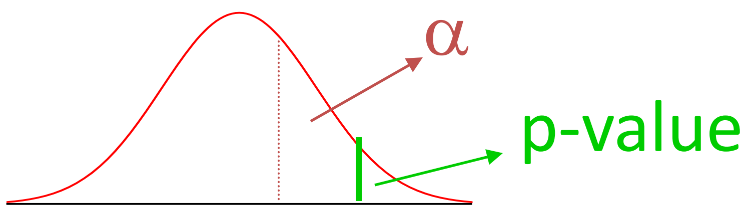
p=p-value = $2.P(z>z_0)$ for $z_0>0$

p=p-value = $2.P(z<z_0)$ for $z_0<0$



DECISION RULE BY USING P-VALUES

- REJECT H_0 IF p-value < α



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

EXAMPLES

- The weights of pots of jam made by a standard process is normally distributed with mean μ=345gr and σ = 2.8gr. A pot produced just before the process closed for the day weight 338.5gr. Is the process working correctly? α = 0.01

$$H_0:\ \mu = 345$$

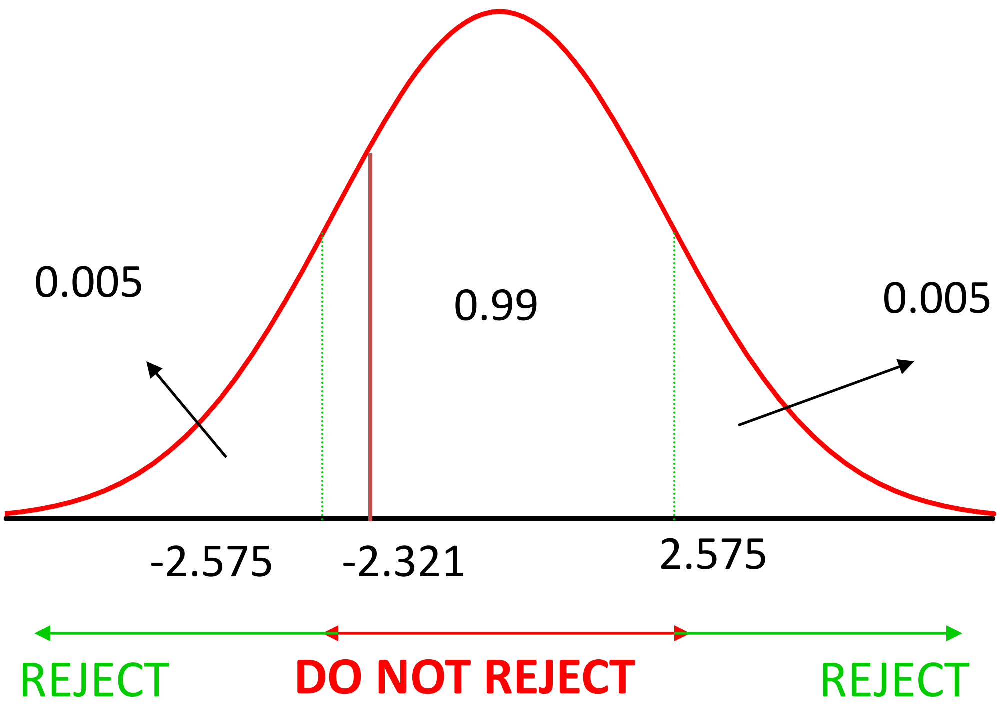
$$H_A:\ \mu \neq 345$$

$$\bar{x} = 338.5 \quad \sigma = 2.8 \quad n = 1 \quad z_{\alpha/2} = z_{0.005} = 2.575$$

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

- The test statistic=

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{338.5 - 345}{2.8} = -2.321$$



- **CONCLUSION:** DO NOT REJECT H_0 AT 1% SIGNIFICANCE LEVEL. THE PROCESS IS WORKING CORRECTLY.

- p-value = 2.P(z<-2.321) = 2.(0.010143)

  =0.02086

- Since p-value = 0.02086 > 0.01, we cannot reject H_0 at 1% significance level

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%.

Computer Output

Excel Output

Test of Hypothesis About MU (SIGMA Known)		
Test of MU = 16 Vs MU less than 16		
SIGMA = 0.4		
Sample mean = 15.9		
Test Statistic: z = -1.25		
P-Value = 0.1056		

- **Minitab Output:**

Z-Test

Test of mu = 16.0000 vs mu < 16.0000

The assumed sigma = 0.400

Variable	N	Mean	StDev	SE Mean	Z	P
Catsup	25	15.9000	0.5017	0.0800	-1.25	0.11

SO DON'T REJECT THE NULL HYPOTHESIS IN THIS CASE

CALCULATIONS

The z-score is: $$Z = \frac{15.9 - 16}{\left\{\frac{.4}{\sqrt{25}}\right\}} = -1.25$$

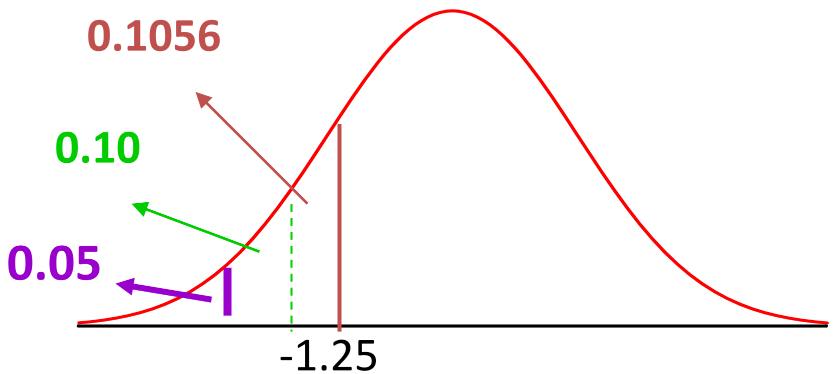
The p-value is the probability of getting a score worse than this (relative to the alternative hypothesis) i.e.,

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

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

P-value for this one-tailed Test

- The p-value for this test is 0.1056



- Thus, do not reject H_0 at 1% and 5% significance level. The contents of bottles of catsup have a net weight of 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 significance at α=0.05.

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

$$s = 0.158$$

SOLUTION

- H_0: $\mu = 14.0$

  H_A: $\mu \neq 14.0$

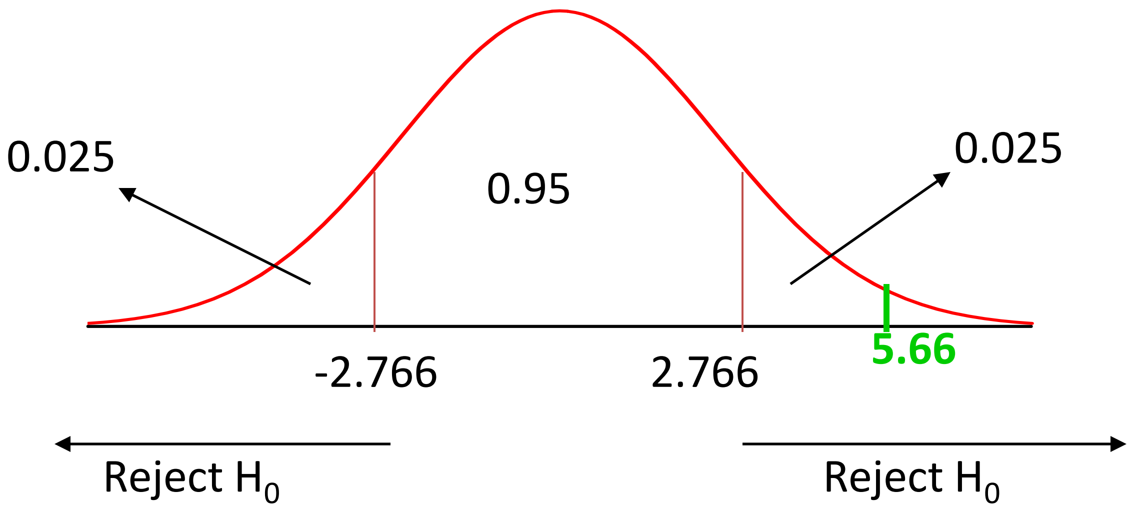
$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{14.4 - 14.0}{0.158 / \sqrt{5}} = 5.66$$

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

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

CONCLUSION

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



P-value of This Test

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

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

Minitab Output

T-Test of the Mean

Test of mu = 14.0000 vs mu not = 14.0000

Variable	N	Mean	StDev	SE Mean	T	P-Value
C1	5	14.4000	0.1581	0.0707	5.66	**0.0048**

CONCLUSION USING THE CONFIDENCE INTERVALS

MINITAB OUTPUT:

Confidence Intervals

Variable	N	Mean	StDev	SE Mean	95.0 % C.I.
C1	5	14.4000	0.1581	0.0707	(14.2036, 14.5964)

- **Since 14 is not in the interval, reject H_0.**

EXAMPLE

- Current output of a (chemical) corporation is 8200 liters/hour of sulfuric acid. An experiment yields a sample of 16 (hourly outputs of the acid) under alternate conditions. $\bar{X} = 8{,}110$ *and* $s=270.5$

$$H_0: \mu = 8200$$

$$H_A: \mu < 8200$$

ANSWER

- The value of the test statistic is:

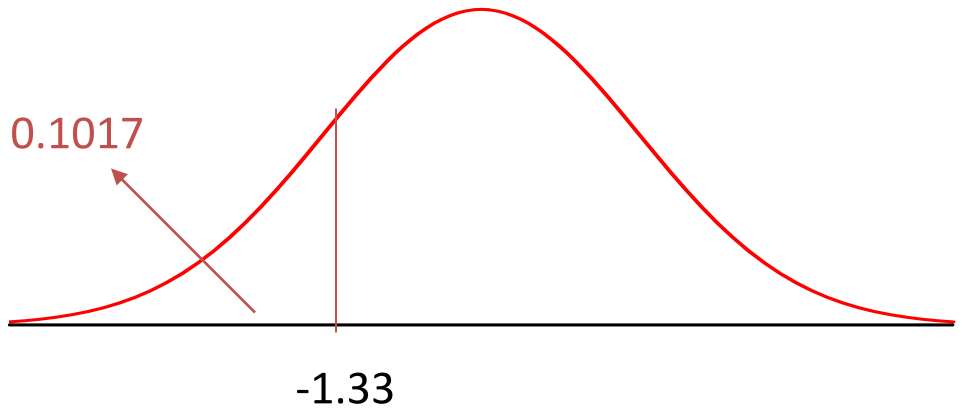
$$t = \frac{8110 - 8200}{\left(\frac{270.5}{\sqrt{16}}\right)} = -1.33$$

Rejection region (α=.05): $t<-t_{\alpha,n-1}=-t_{.05,15}=-1.753$

Conclusion: Do NOT reject H_0 since -1.33 is NOT in the rejection region

P-VALUE

- p-value = $P(t < -1.33) = 0.1017$



- Since p-value = $0.1017 > 0.05$, do not reject H_0.

EXAMPLE

Problem: At a certain production facility that assembles computer keyboards, the assembly time is known (from experience) to follow a normal distribution with mean (μ) of 130 seconds and standard deviation (σ) 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: $H_A : \mu \neq 130$

$$\bar{X} = 126.8$$

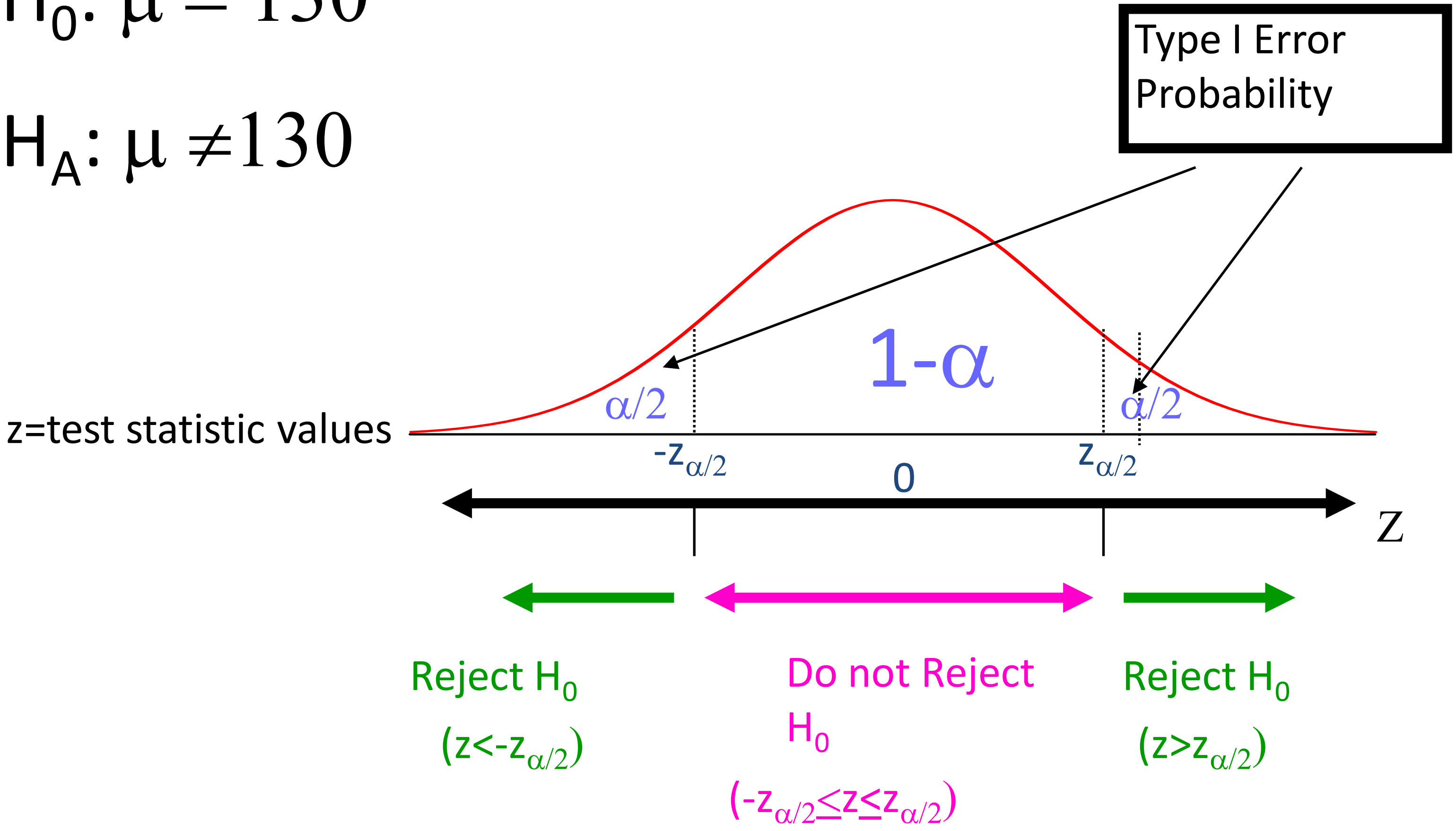
- Since the standard deviation is $\sigma = 15$,
- The standardized test statistic value is:

$$Z = \frac{126.8 - 130}{\left\{\frac{15}{\sqrt{100}}\right\}} = -2.13$$

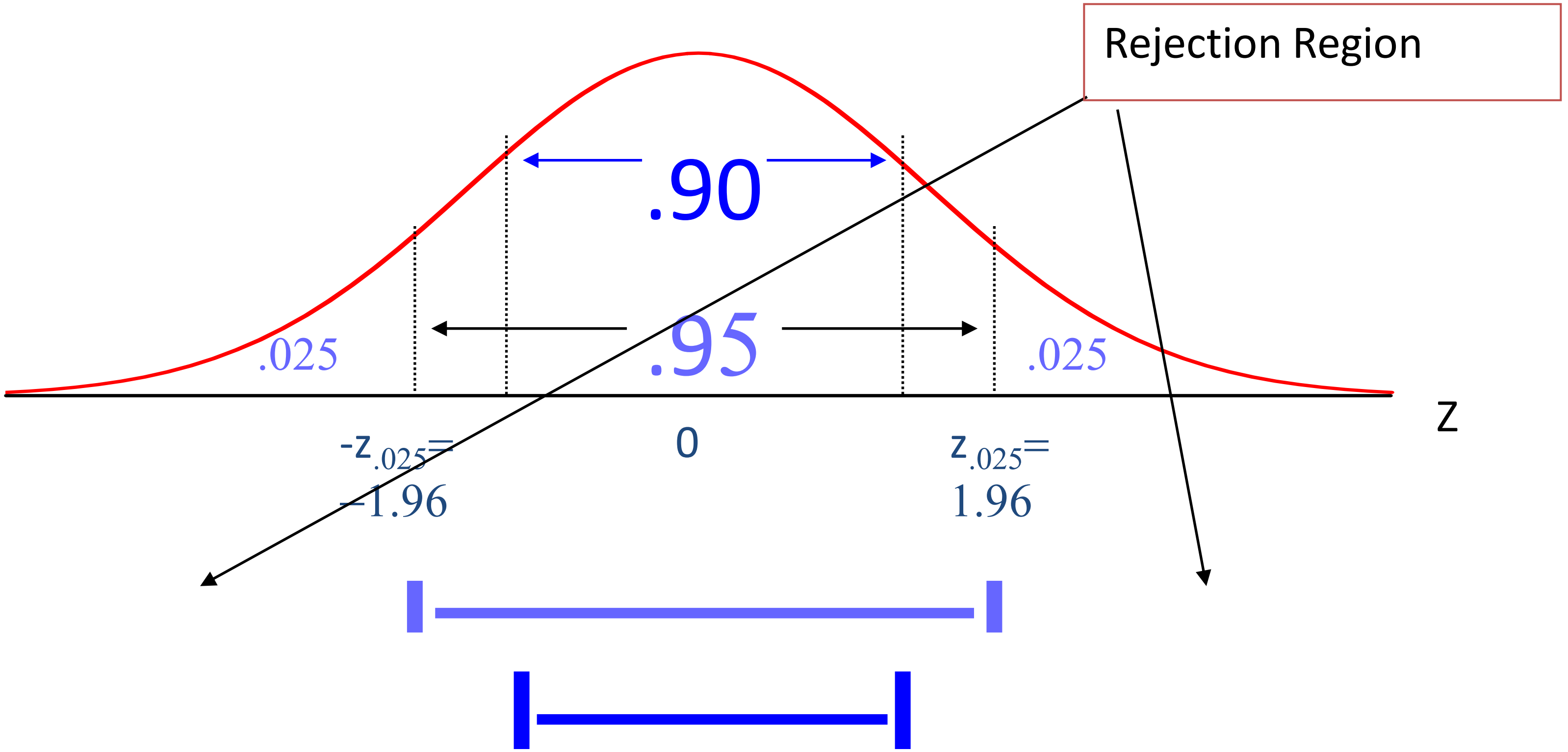
Two-Tail Hypothesis:

$H_0: \mu = 130$

$H_A: \mu \neq 130$



Test Statistic: $z = \dfrac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \dfrac{126.8 - 130}{15 / \sqrt{100}} = -2.13$



CONCLUSION

- Since –2.13<-1.96, it falls in the rejection region.
- Hence, **we reject the null hypothesis** that the time required to do the assembly is still 130 seconds. The evidence suggests that the task now takes either more or less than 130 seconds.

DECISION RULE

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

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

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

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

P-VALUE

- In our example, the p-value is

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

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.

β=P(not rejecting H_o|H_o is false).

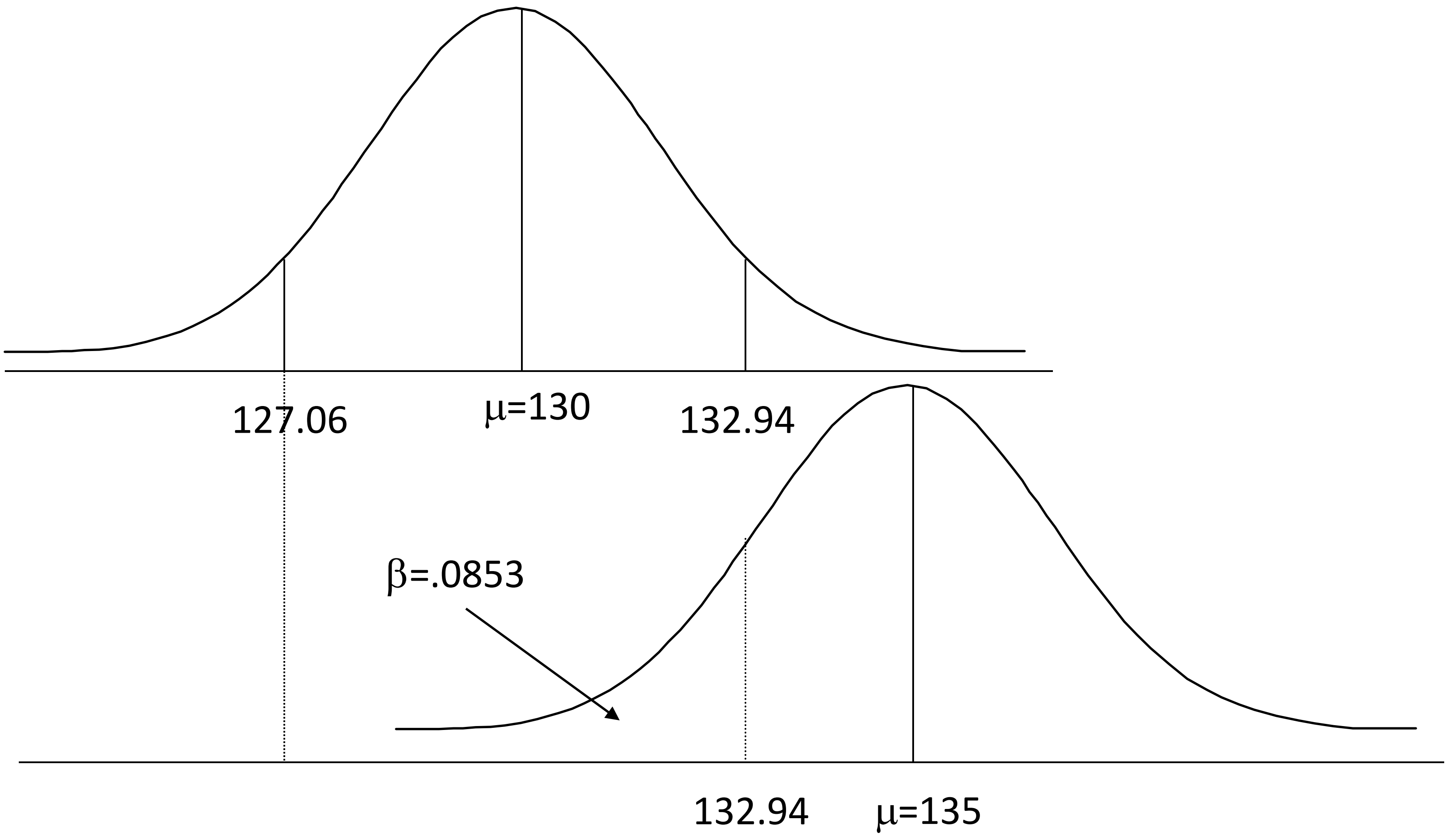
Assuming μ=135 this statement becomes:

$$P(127.06 < \bar{x} < 132.94 | \mu = 135)$$

$$= P(\frac{127.06\text{-}135}{15/\sqrt{100}} < Z < \frac{132.94\text{-}135}{15/\sqrt{100}})$$

$$= P(\text{-}5.29 < Z < \text{-}1.37) = .0853$$

Probability of Type II Error



EXAMPLE

- Consider the test

$$H_o: \mu = 2400$$

$$H_A: \mu > 2400$$

n=50, s=200 and $\alpha = 0.05$

Test Statistic: $z = \dfrac{\bar{x} - 2400}{200/\sqrt{50}} \Leftrightarrow \bar{x} = 2400 + z\dfrac{200}{\sqrt{50}}$

Rejection Region: $z > z_{\alpha/2} = 1.645$ or

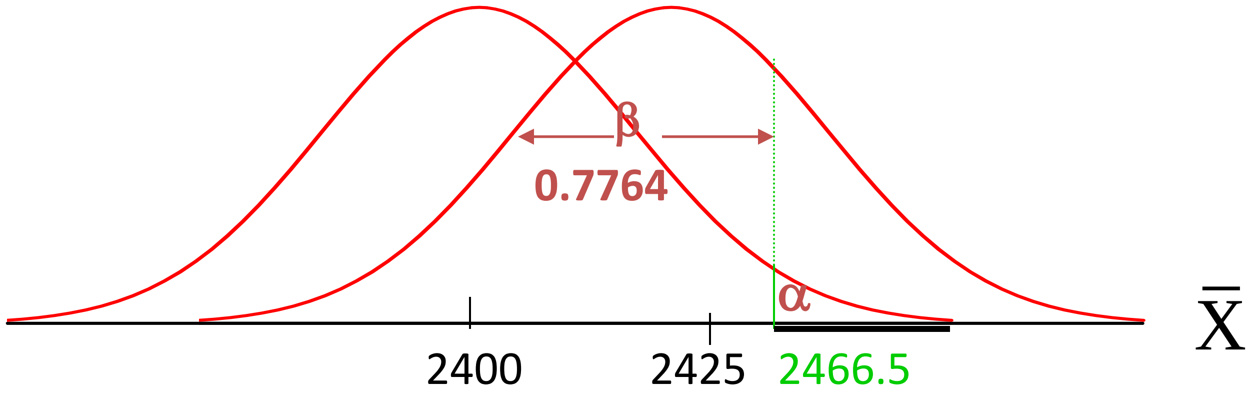
$$\bar{x} > 2400 + 1.645\frac{200}{\sqrt{50}} = 2446.5$$

TYPE II ERROR

- If the actual is μ_A=2425, then

$$\beta = P(\bar{X} \le 2446.5 \mid \mu = 2425) = P\left(\frac{\bar{X} - 2425}{200/\sqrt{50}} \le \frac{2446.5 - 2425}{200/\sqrt{50}}\right)$$

$$= P(Z \le 0.76) = 0.7764$$



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 $nq \geq 5$).

HYPOTHESIS TEST FOR p

Two-sided Test

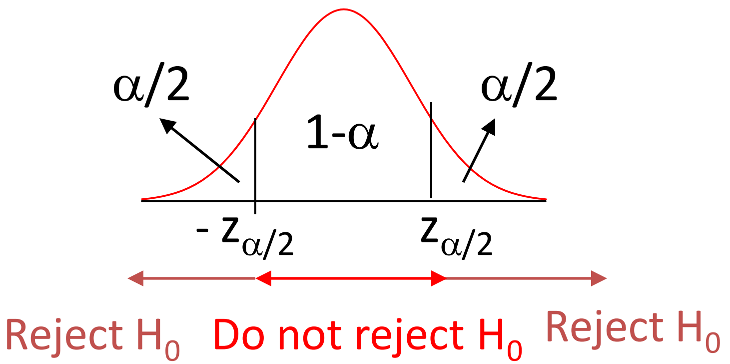
H_0: $p = p_0$

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 $z > z_{\alpha/2}$.

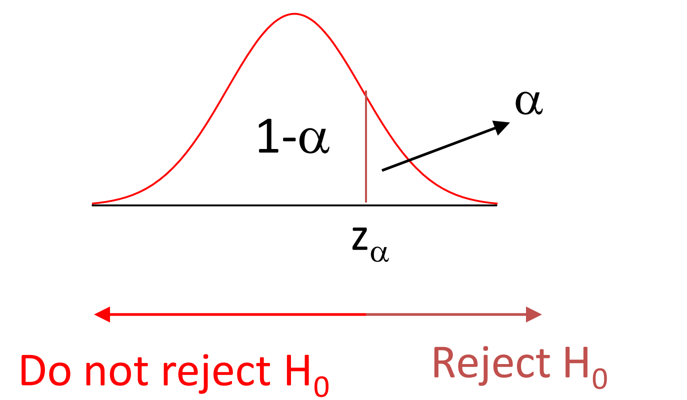
HYPOTHESIS TEST FOR p

One-sided Tests | **Test Statistic** | **Rejecting Area**

1. $H_0: p = p_0$

   $H_A: p > p_0$

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

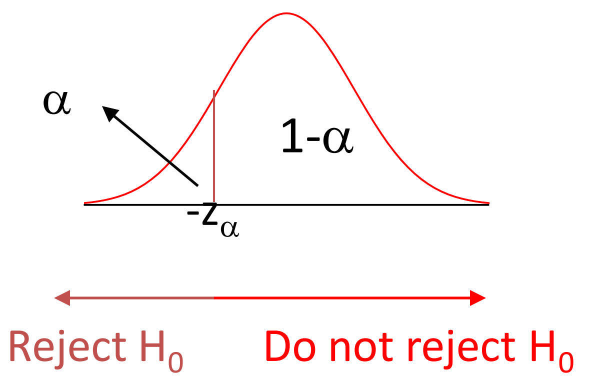


- Reject H_o if $z > z_\alpha$.

2. $H_0: p = p_0$

   $H_A: p < p_0$

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



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

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.

$$H_o: p = 0.70$$

$$H_A: p \neq 0.70$$

ANSWER

- 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:

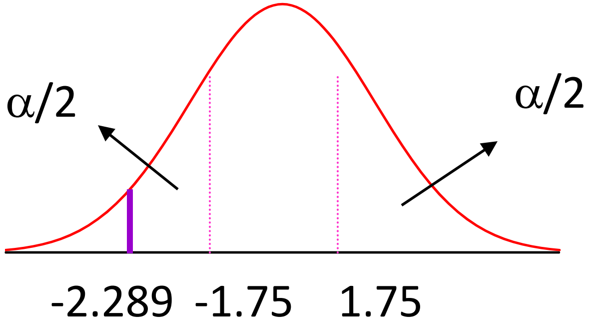
$$z = \frac{0.6 - 0.7}{\sqrt{(0.7)(0.3)/110}} = -2.289$$

CONCLUSION

- **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.

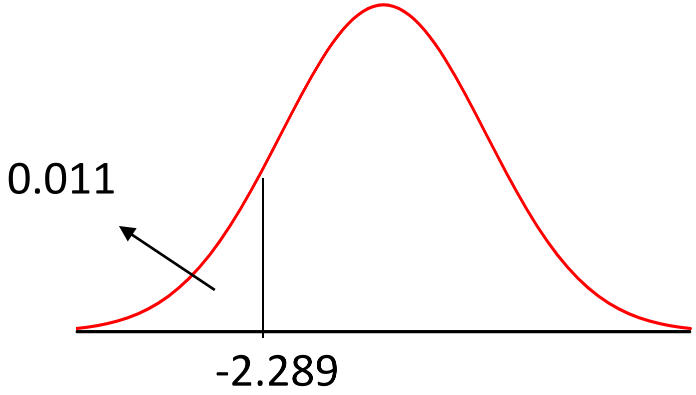


P-VALUE

- 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 < α = 0.08, reject H_0.



CONFIDENCE INTERVAL APPROACH

- Find the 92% CI for p.

$$p \pm z_{\alpha/2}\sqrt{\frac{pq}{n}} = 0.7 \pm 1.75\sqrt{\frac{(0.7)(0.3)}{110}}$$

92% CI for p: $0.623 \le p \le 0.777$

- Since $\hat{p} = 0.6$ is not in the above interval, reject H_0. Mom has underestimated the cost of her meal.

EXAMPLE

H_0: p = .10

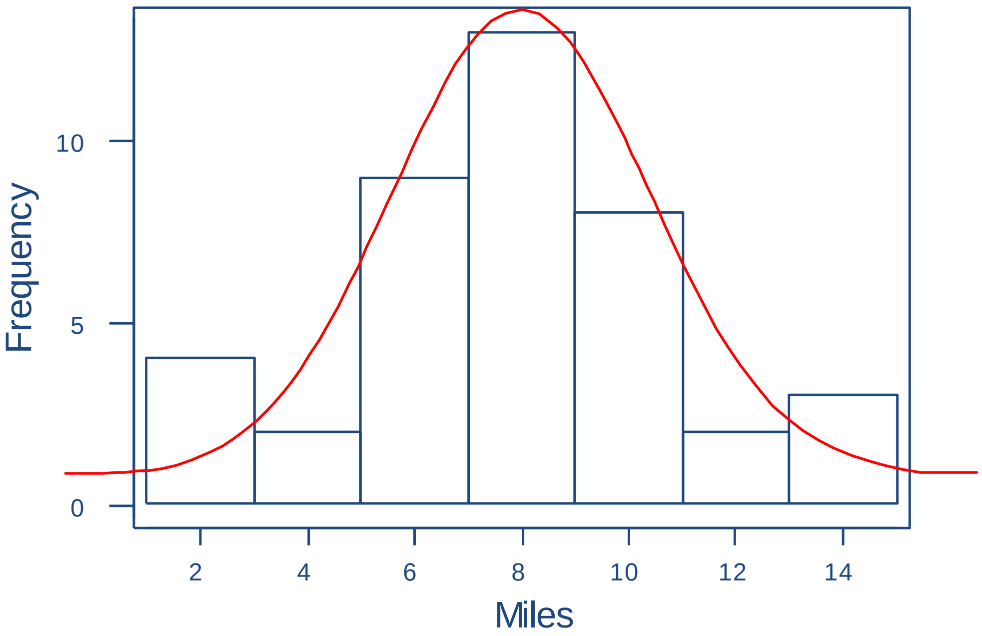
H_A: p > .10

- **Data:** x=52 (number of visitors in sample that would rent the device) in a sample of 400 visitors surveyed.
- **Test Statistic**: $$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{52}{400} - .10}{\sqrt{\frac{(.10)(.90)}{400}}} = 2.0$$
- P-value = P(z > 2) = 0.0228 > α = 0.05
- Not Reject H_0 at α = 0.05

EXCEL OUTPUT

<u>**Test of Hypothesis About p**</u>			
Test of p = 0.1 Vs p greater than 0.1			
Sample Proportion = 0.13			
Test Statistic = 2			
P-Value = 0.0228			

Testing the Normality Assumption



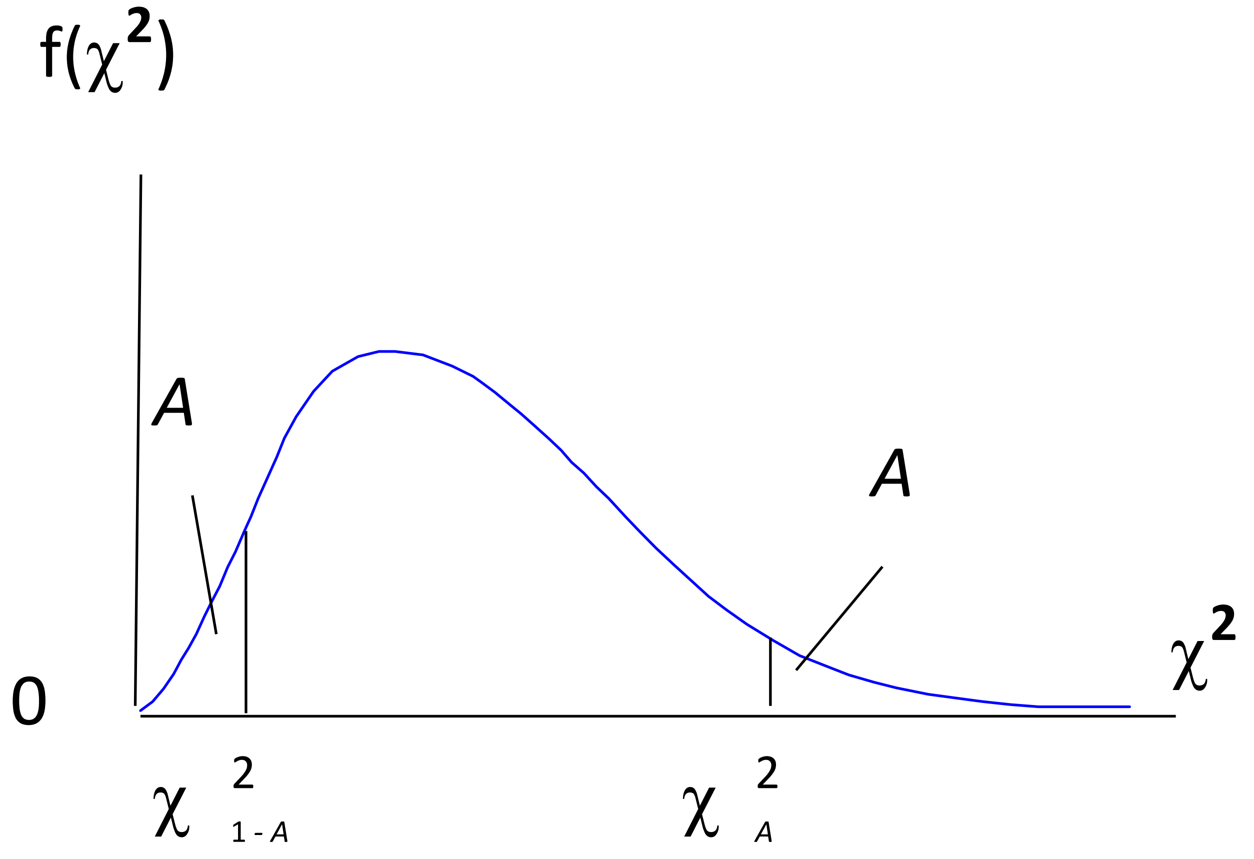
SAMPLING DISTRIBUTION OF s^2

- The statistic

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

is chi-squared distributed with n-1 d.f. when the population random variable is normally distributed with variance σ^2.

CHI-SQUARE DISTRIBUTION



Inference about the Population Variance (σ^2)

- Test statistic

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

which is chi-squared distributed with n - 1 degrees of freedom

Confidence interval estimator:

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

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

Testing the Population Variance (σ^2) 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?

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

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

ANSWER

- 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.
- The evidence suggests that the variability in product weights exceed the maximum allowance.

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:

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

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

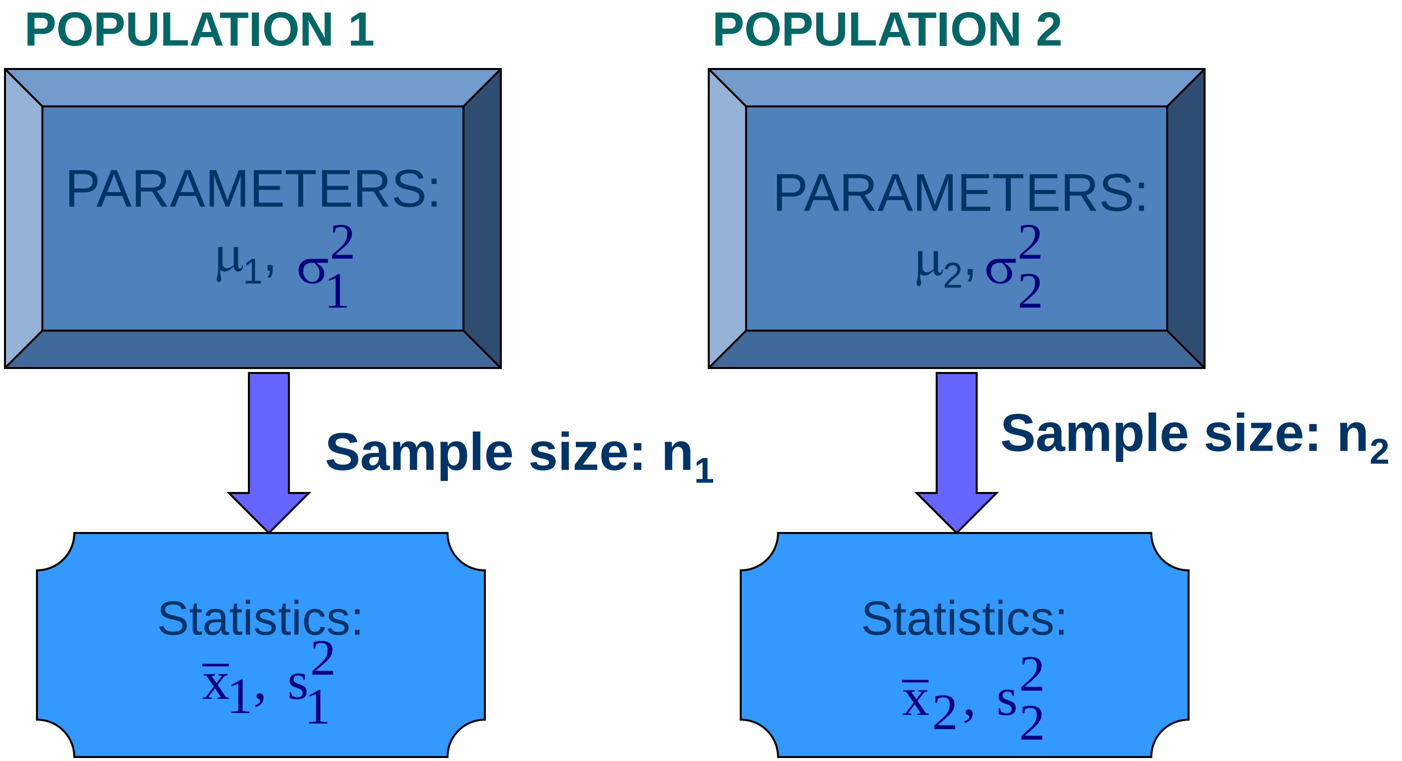
$$23.9757 \le \sigma^2 \le 67.5765$$

INTERPRETATION OF THE CONFIDENCE INTERVAL

- We are 90% confident that the population variance is between 23.9757 and 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 $n_2 \geq 30$.

INFERENCE ABOUT μ_1-μ_2

CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$

σ_1 AND σ_2 ARE KNOWN FOR NORMAL DISTRIBUTION OR LARGE SAMPLE

- A $100(1-\alpha)\%$ 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.

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

EXAMPLE

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

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

- 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: $\left(\bar{x}_2 - \bar{x}_1\right) \pm 1.96(s_{\bar{x}_2 - \bar{x}_1})$

$$363 \leq \mu_2 - \mu_1 \leq 2397$$

INTERPRETATION

- With 95% confidence, mean family income in the second group exceeds 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

SOLUTION

- **STEP 1:** $H_0: \mu_1 = \mu_2 \Rightarrow H_0: \mu_1 - \mu_2 = 0$

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

- **STEP 2:** Test statistic:

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

- **STEP 3**: Decision Rule = Reject H_0 if $z<-z_{\alpha/2}$=-1.645 or $z>z_{\alpha/2}$=1.645.
- **STEP 4:** 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
$\sigma_1 = \sigma_2$

- 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 $\alpha = 0.01$.

SOLUTION

- $n_1 < 30$ 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.

$$H_0: \mu_1 = \mu_2$$

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

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

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

- 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 $\alpha = 0.01$.

Test Statistic for $\mu_1 - \mu_2$ when $\sigma_1 \neq \sigma_2$ and unknown

- Test Statistic:

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

with the degree of freedom

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

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

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

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

SOLUTION

- 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

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

$H_0: \mu_1 - \mu_2 = 0$

$H_A: \mu_1 - \mu_2 < 0$

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

- DECISION RULE:

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

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

MINITAB OUTPUT

- **Two Sample T-Test and Confidence Interval**

Twosample T for Consmers vs Non-cmrs

	N	Mean	StDev	SE Mean
Consmers	10	595.8	35.7	11
Non-cmrs	20	661	116	26

- 95% C.I. for mu Consmers - mu Non-cmrs: **(-123, -7)**

T-Test mu Consmers = mu Non-cmrs (vs <):

T= -2.31 **P=0.015** DF= 25

EXAMPLE

μ_1=mean assembly time (in minutes) using Design A

μ_2=mean assembly time (in minutes) using Design B

$$H_0: (\mu_1 - \mu_2) = 0$$

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

- To decide the correct test, calculate the sample standard deviations.

$$s_1=0.921 \text{ and } s_2 = 1.14$$

Because s_1 approximately equal to s_2, use equal-variance t-test.

Pooled variance:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(25-1)(.921)^2 + (25-1)(1.142)^2}{25+25-2} = 1.075$$

Test Statistic:
(assuming equal variances)

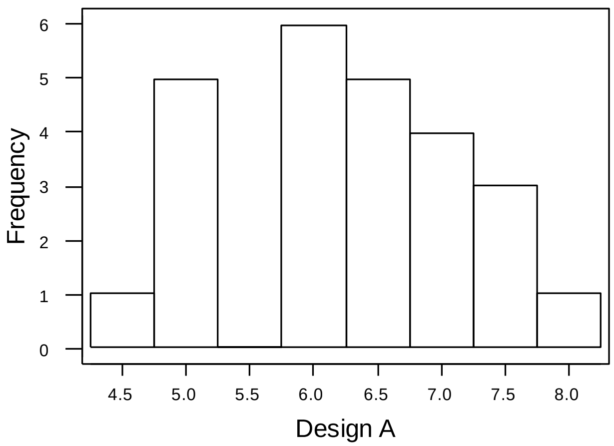
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

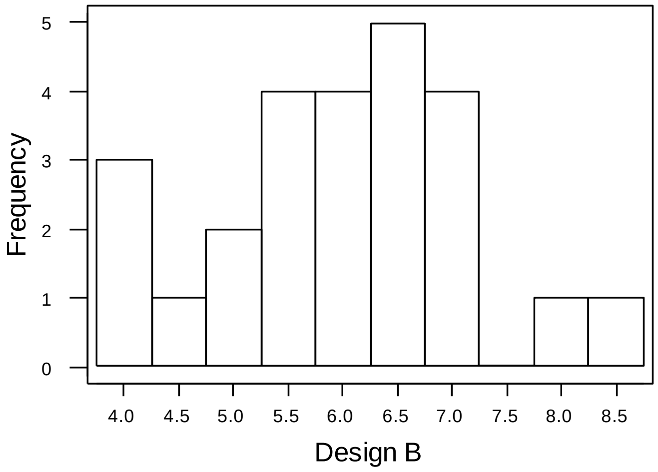
$$= \frac{(6.288 - 6.016) - 0}{\sqrt{1.075\left(\frac{1}{25} + \frac{1}{25}\right)}} = .93$$

Rejection region: $|t| > t_{.025,48} = 2.009$

Conclusion: Do not reject H_0

Normality Assumption





- Although the histograms are not bell-shaped, given the robustness of the t-test the conclusion (not rejecting H_0) may be accurate

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

$$D_1, D_2, \dots, D_n.$$

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:

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

- 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

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

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

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

$t_{0.05,9}$

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

- 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

SOLUTION

$H_0: \mu_D = 0$

$H_A: \mu_D > 0$

- Test Statistics:

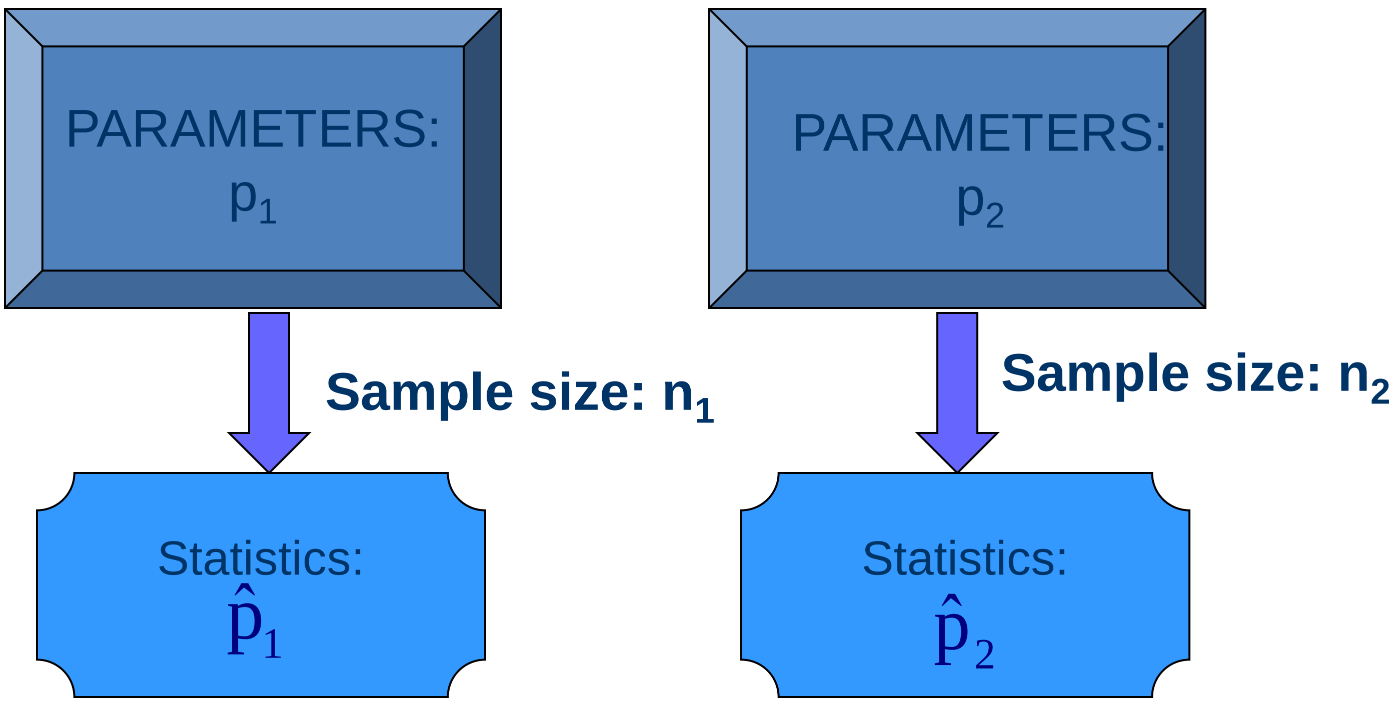
$$t = \frac{\bar{x}_D - \mu_D}{s_D / \sqrt{n}} = \frac{4.55 - 0}{7.22/\sqrt{20}} = 2.82$$

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

Conclusion: Reject H_0 at $\alpha = 0.05$

Inference About the Difference of Two Population Proportions

Population 1 Population 2



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 $\alpha = 0.05$.

SOLUTION

$H_0: p_1 = p_2$

$H_A: p_1 < p_2$

$\hat{p}_1 = \frac{x_1}{n_1} = \frac{117}{318}$ 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:

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

- Decision Rule: Reject H_0 if $z < -z_{0.05}=-1.96$.
- Conclusion: Because $z = -1.4518 > -z_{0.05}=-1.96$, not reject H_0 at $\alpha=0.05$. Manufacturer is wrong.

Example

- In a study, doctors discovered that **aspirin** seems to help **prevent heart attacks**. Half of 22,000 male physicians 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

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

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

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

Test of Hypothesis for p_1-p_2

$H_0: p_1 - p_2 = 0$

$H_A: p_1 - p_2 < 0$

- **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 $100(1-\alpha)\%$ C.I. for $p_1 - p_2$ is given by:

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

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

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

TEST OF HYPOTHESIS

HOW TO DERIVE AN APPROPRIATE TEST

MOST POWERFUL TEST (MPT)

$H_0: \theta = \theta_0 \Rightarrow$ Simple Hypothesis

$H_1: \theta = \theta_1 \Rightarrow$ Simple Hypothesis

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

The Neyman-Pearson Lemma:

$$C = \left\{ (x_1, x_2, \cdots, x_n) : \frac{L(\theta_0)}{L(\theta_1)} \leq k \right\}$$

Reject H_0 if $L = \dfrac{L(\theta_0)}{L(\theta_1)} \leq k$

$$\alpha = P(L \leq k | \theta = \theta_0) \rightarrow k$$

$$\beta = 1 - P(L \leq k | \theta = \theta_1)$$

EXAMPLES

- $X \sim N(\mu, \sigma^2)$ where σ^2 is known.

$$H_0: \mu = \mu_0$$

$$H_1: \mu = \mu_1$$

Find the most powerful test of size α.

EXAMPLES

- On the basis of a r.s. of size 1 from the pdf

$$f(x;\theta) = \theta x^{\theta-1}, 0 < x < 1, \theta > 1$$

$H_0: \theta = \theta_0$
$H_1: \theta = \theta_1$

Use the Neyman-Pearson lemma to derive MPT of size α.

UNIFORMLY MOST POWERFUL (UMP) TEST

- If a test is most powerful against every possible value in a composite alternative, then it will be a UMP test.
- To be able to find UMPT, we use Monotone Likelihood Ratio (MLR).
- If $L=L(\theta_0)/L(\theta_1)$ depends on $(x_1,x_2,\ldots,x_n)$ only through the statistic $y=u(x_1,x_2,\ldots,x_n)$ and L is an **increasing function** of y for every given $\boldsymbol{\theta_0>\theta_1}$, then we have a **monotone likelihood ratio** (MLR) in statistic y.
- *If* L is an **decreasing function** of y for every given $\boldsymbol{\theta_0>\theta_1}$, then we have a monotone likelihood ratio (MLR) in statistic $-y$.

UNIFORMLY MOST POWERFUL (UMP) TEST

- **Theorem:** If a joint pdf $f(x_1, x_2, \ldots, x_n; \theta)$ has MLR in the statistic Y, then a UMP test of size α for $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0$ is to reject H_0 if $Y \geq c$ where $P(Y \geq c | \theta_0) = \alpha$.
- for $H_0: \theta \geq \theta_0$ vs $H_1: \theta < \theta_0$ is to reject H_0 if $Y \leq c$ where $P(Y \leq c | \theta_0) = \alpha$.

EXAMPLE

- $X \sim Uniform(0,\theta)$

a) Single observation

b) Random sample of size n

H_0:θ=3

H_1:θ>3

Find UMPT of size α.

EXAMPLE

- $X \sim Exp(\theta)$

$H_0: \theta \leq \theta_0$

$H_1: \theta > \theta_0$

Find UMPT of size α.

GENERALIZED LIKELIHOOD RATIO TEST (GLRT)

- Derives a test when we have two-sided composite alternative or when we have unknown nuisance parameters
- GLRT is the generalization of MPT and provides a desirable test in many applications but it is not necessarily a UMP test.

GENERALIZED LIKELIHOOD RATIO TEST (GLRT)

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

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

$$L(\theta) = f(x_1, x_2, \cdots, x_n; \theta) \overset{r.s.}{=} f(x_1;\theta), f(x_2;\theta), \cdots, f(x_n;\theta)$$

Let $L(\hat{\Omega}) = \max\limits_{\theta \in \Omega} L(\theta) = L(\hat{\theta})$ and

MLE of θ

$$L(\hat{\Omega}_0) = \max_{\theta \in \Omega_0} L(\theta) = L(\hat{\theta}_0)$$

MLE of θ under H_0

GENERALIZED LIKELIHOOD RATIO TEST (GLRT)

$$\lambda = \frac{L\left(\hat{\Omega}_0\right)}{L\left(\hat{\Omega}\right)}, 0 \le \lambda \le 1 \Rightarrow \text{The Generalized Likelihood Ratio}$$

GLRT: Reject H_0 if $\lambda < \lambda_0$

EXAMPLE

- $X \sim N(\mu, \sigma^2)$

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Derive GLRT of size α.

- $X \sim N(\mu, \sigma^2)$

$$H_0: \mu = \mu_0$$
$$H_1: \mu > \mu_0$$

Derive GLRT of size α.

EXAMPLE

- $X \sim Exp(\theta)$

$$H_0: \theta = \theta_0$$
$$H_1: \theta \neq \theta_0$$

a) Find GLR for this test.

b) Determine the rejection region

c) Find the power function

d) From the acceptance region of this test, find a 100(1-α)% CI for θ

e) Write a general expression for the p-value of this test for a given observed value $\bar{x}$ of $\overline{X}$

ASYMPTOTIC DISTRIBUTION OF $-2\ln\lambda$

- GLRT: Reject H_0 if $\lambda<\lambda_0$
- GLRT: Reject H_0 if $-2\ln\lambda>-2\ln\lambda_0=c$

$$-2\ln\lambda \underset{asympt.}{\overset{under\ \mathrm{H}_0}{\sim}} \chi_k^2$$

where k is the number of parameters to be tested.

$\Rightarrow$Reject H_0 if $-2\ln\lambda>\chi^2_{\alpha,k}$

EXAMPLE

- $X \sim N(\mu, \sigma^2)$

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Derive approximate GLRT of size α.

TWO SAMPLE TESTS

$$X \sim N\left(\mu_1, \sigma_1^2\right), r.s.\, n_1, \bar{x}, s_x^2$$

$$Y \sim N\left(\mu_2, \sigma_2^2\right), r.s.\, n_2, \bar{y}, s_y^2$$

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

Derive GLRT of size α.