

Probability and Statistics

Prepared by Dr. Alfaisal Hasan

STA 201
Probability and Statistics

Lecture NO 7



Probability and Statistics

Lecture NO 7

STA 201

Probability and Statistics

Prepared by
Dr. Alfaisal Hasan

www.midocean.ae

Chapter

Topics

- Chapter 1: Introduction To Statistics
 - Chapter 2: Descriptive Statistics
 - Chapter 3: Probability
 - Chapter 4: Discrete Probability Distributions
 - Chapter 5: Normal Probability Distributions
 - Chapter 6: Confidence Intervals
 - **Chapter 7: Hypothesis Testing with One Sample**
 - Chapter 8: Hypothesis Testing with Two Samples
 - Chapter 9: Correlation And Regression
 - Chapter 10: Chi-Square Tests and The F -Distribution
- الفصل الأول: مقدمة في الإحصاء
الفصل الثاني: الإحصاء الوصفي
الفصل الثالث: الاحتمالات
الفصل الرابع: التوزيعات الاحتمالية المنفصلة
الفصل الخامس: التوزيع الاحتمالي الطبيعي
الفصل السادس: فترات الثقة
الفصل السابع: اختبار الفرضيات لعينة واحدة
الفصل الثامن: اختبار الفرضيات لعينتين
الفصل التاسع: الارتباط والانحدار
الفصل العاشر: اختبارات كاي تربيع وتوزيع F

Hypothesis Testing with One Sample

7.1 Introduction to Hypothesis Testing

7.1 مقدمة لاختبار الفرضيات

7.2 Hypothesis Testing for the Mean (Large Samples)

7.2 اختبار الفرضيات للمتوسط (عينات كبيرة)

7.3 Hypothesis Testing for the Mean (Small Samples)

7.3 اختبار الفرضيات للمتوسط (عينات صغيرة)

7.4 Hypothesis Testing for Proportions

7.4 اختبار الفرضيات للنسب

7.1 Introduction to Hypothesis testing

Hypothesis test

A process that uses sample statistics to test a claim about the value of a population parameter.

For example: An automobile manufacturer advertises that its new hybrid car has a mean mileage of 50 miles per gallon. To test this claim, a sample would be taken. If the sample mean differs enough from the advertised mean, you can decide the advertisement is wrong.

اختبار الفرضية
عملية تستخدم إحصائيات العينة لاختبار مطالبة حول
قيمة معلمة المحتوى.

على سبيل المثال: تعلن شركة تصنيع سيارات أن سيارتها
الهجينة الجديدة يبلغ متوسط عدد الكيلومترات فيها 50
ميلاً للغالون الواحد. لاختبار هذا الادعاء، سيتم أخذ عينة.
إذا كان متوسط العينة يختلف بدرجة كافية عن
المتوسط المعلن عنه، فيمكنك أن تقرر أن الإعلان خاطئ.

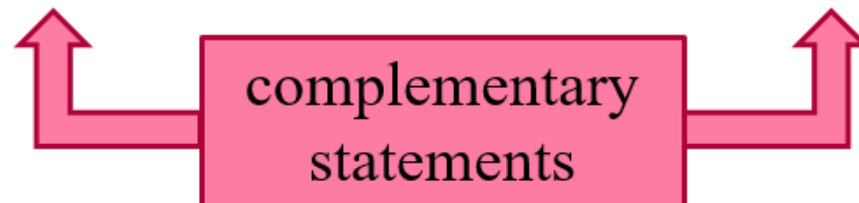
Stating a Hypothesis

Null hypothesis

- A statistical hypothesis that contains a statement of equality such as $\leq$, $=$, or $\geq$.
- Denoted H_0 read “H subzero” or “H naught.”

Alternative hypothesis

- A statement of inequality such as $>$, $\neq$, or $<$.
- Must be true if H_0 is false.
- Denoted H_a read “H sub-a.”



Population Mean

First Case:

$$H_0: \mu =$$

$$H_a: \mu \neq$$

Second Case:

$$H_0: \mu \leq$$

$$H_a: \mu >$$

Third Case:

$$H_0: \mu \geq$$

$$H_a: \mu <$$

Population Proportion

First Case:

$$H_0: p =$$

$$H_a: p \neq$$

Second Case:

$$H_0: p \leq$$

$$H_a: p >$$

Third Case:

$$H_0: p \geq$$

$$H_a: p <$$

Example: Stating the Null and Alternative Hypotheses

Write the claim as a mathematical sentence.
State the null and alternative hypotheses and identify which represents the claim.

اكتب المطالبة في صورة جملة رياضية. اذكر الفرضيات
الصفرية والبديلة وحدد ما يمثل المطالبة.

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

تعلن إحدى الجامعات أن نسبة طلابها الذين يتخرجون خلال
4 سنوات تبلغ 82%.

Solution:

$$H_0: p = 0.82 \longleftarrow \text{Equality condition (Claim)}$$

$$H_a: p \neq 0.82 \longleftarrow \text{Complement of } H_0$$



Example: Stating the Null and Alternative Hypotheses

Write the claim as a mathematical sentence.
State the null and alternative hypotheses and identify which represents the claim.

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

اكتب المطالبة في صورة جملة رياضية. اذكر الفرضيات الصفرية والبديلة وحدد ما يمثل المطالبة.

تعلن إحدى الشركات المصنعة لصنوبر المياه أن متوسط معدل التدفق لنوع معين من الصنوبر أقل من 2.5 جالون في الدقيقة.



Solution:

$H_0: \mu \geq 2.5$ gallons per minute ← Complement of H_a

$H_a: \mu < 2.5$ gallons per minute ← Inequality condition (Claim)

Types of Errors

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

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

تجربة مشروب الكركديه على مجموعة من مرضى ضغط الدم
المرتفع وقياس ضغط الدم لهم قبل وبعد استخدامه لفترة محددة
بانتظام

الفرض الصفري (العدم)

قياس ضغط الدم قبل استخدام الكركديه يساوي قياس ضغط
الدم بعد الاستخدام (يعني لا يوجد فرق)

الفرض البديل

قياس ضغط الدم قبل استخدام الكركديه لا يساوي قياس
ضغط الدم بعد الاستخدام (يعني يوجد فرق)



Example: Identifying Type I and Type II Errors

The USDA limit for salmonella contamination for chicken is 20%. A meat inspector reports that the chicken produced by a company exceeds the USDA limit. You perform a hypothesis test to determine whether the meat inspector's claim is true. When will a type I or type II error occur? Which is more serious?

(Source: United States Department of Agriculture)

حد وزارة الزراعة الأمريكية لتلوث الدجاج بالسالمونيلا هو 20%. أفاد أحد مفتشي اللحوم أن الدجاج الذي تنتجه إحدى الشركات يتجاوز الحد المسموح به لوزارة الزراعة الأمريكية. تقوم بإجراء اختبار الفرضيات لتحديد ما إذا كان ادعاء مفتش اللحوم صحيحًا. متى سيحدث خطأ من النوع الأول أو النوع الثاني؟ أيهما أكثر خطورة؟ (المصدر: وزارة الزراعة الأمريكية)

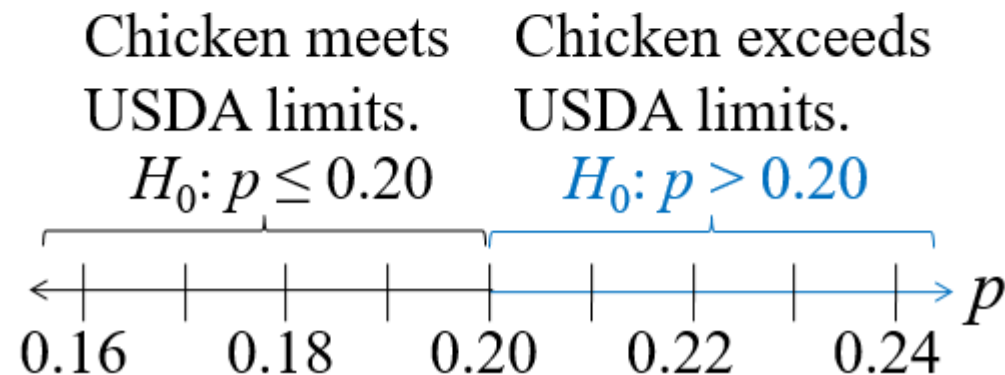


Solution: Identifying Type I and Type II Errors

Let p represent the proportion of chicken that is contaminated.

Hypotheses: $H_0: p \leq 0.2$

$H_a: p > 0.2$ (Claim)



Solution: Identifying Type I and Type II Errors

Hypotheses: $H_0: p \leq 0.2$

$H_a: p > 0.2$ (Claim)

A type I error is rejecting H_0 when it is true.

The actual proportion of contaminated chicken is less than or equal to 0.2, but you decide to reject H_0 .

A type II error is failing to reject H_0 when it is false.

The actual proportion of contaminated chicken is greater than 0.2, but you do not reject H_0 .

Solution: Identifying Type I and Type II Errors

Hypotheses: $H_0: p \leq 0.2$

$H_a: p > 0.2$ (Claim)

- With a type I error, you might create a health scare and hurt the sales of chicken producers who were actually meeting the USDA limits.
- With a type II error, you could be allowing chicken that exceeded the USDA contamination limit to be sold to consumers.
- A type II error could result in sickness or even death.

Level of significance

Level of significance

Your maximum allowable probability of making a type I error.

Denoted by α , the lowercase Greek letter alpha.

By setting the level of significance at a small value, you are saying that you want the probability of rejecting a true null hypothesis to be small.

Commonly used levels of significance:

$$\alpha = 0.10 \quad \alpha = 0.05 \quad \alpha = 0.01$$

$$P(\text{type II error}) = \beta (\text{beta})$$

الحد الأقصى المسموح به لاحتمال ارتكاب خطأ من النوع الأول
يرمز له بالرمز α
من خلال تحديد مستوى الأهمية عند قيمة صغيرة، فإنك تقول أنك تريد
أن يكون احتمال رفض فرضية العدم الحقيقية صغيراً.

Taking decision by Traditional way - First way

1. State the claim mathematically and verbally. Identify the null and alternative hypotheses.

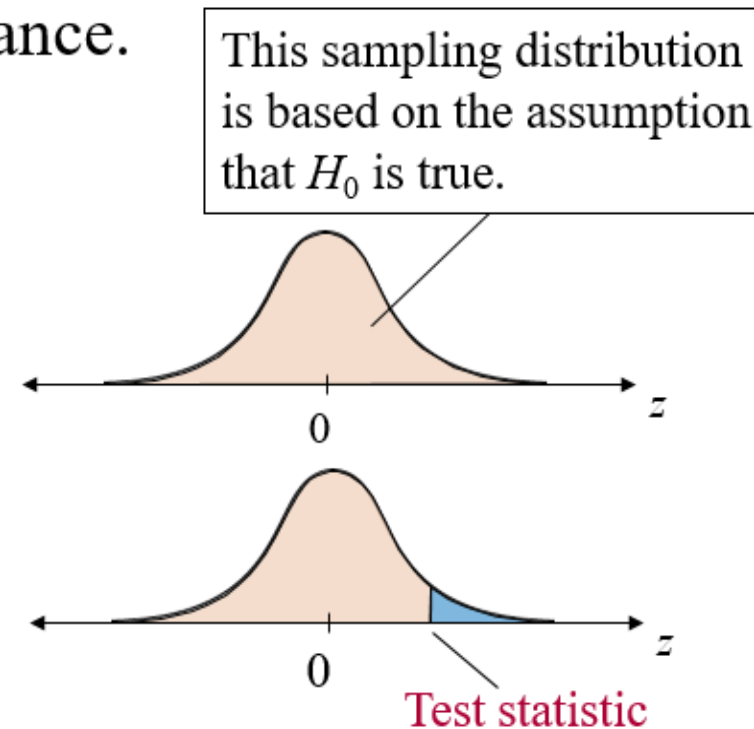
$$H_0: ? \quad H_a: ?$$

2. Specify the level of significance.

$$\alpha = ?$$

3. Determine the standardized sampling distribution and draw its graph.

4. Calculate the test statistic and its standardized value. Add it to your sketch.



Making a Decision by P-value – Second way

P-value (or probability value)

The probability, if the null hypothesis is true, of obtaining a sample statistic with a value as extreme or more extreme than the one determined from the sample data.

Depends on the nature of the test.

الاحتمال، إذا كانت الفرضية الصفرية صحيحة، للحصول على إحصائية عينة بقيمة متطرفة أو أكثر تطرفاً من تلك المحددة من بيانات العينة.

يعتمد على طبيعة الاختبار.

Making a Decision by P-value – Second way

Decision Rule Based on P -value

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

Decision	Claim	
	Claim is H_0	Claim is H_a
Reject H_0	There is enough evidence to reject the claim	There is enough evidence to support the claim
Fail to reject H_0	There is not enough evidence to reject the claim	There is not enough evidence to support the claim

Example: Interpreting a Decision

You perform a hypothesis test for the following claim. How should you interpret your decision if you reject H_0 ? If you fail to reject H_0 ?

2. H_a (Claim): *Consumer Reports* states that the mean stopping distance (on a dry surface) for a Honda Civic is less than 136 feet.

قمت بإجراء اختبار الفرضية للمطالبة التالية. كيف يجب أن تفسر قرارك إذا رفضت الفرض الصفري؟ إذا فشلت في رفض الفرض الصفري؟
تشير تقارير المستهلك إلى أن متوسط مسافة التوقف (على سطح جاف) لسيارة هوندا سيفيك أقل من 136 قدمًا.



Solution:

- The claim is represented by H_a .
- H_0 is “the mean stopping distance...is greater than or equal to 136 feet.”

Solution: Interpreting a Decision

•If you reject H_0 you should conclude “there is enough evidence to support *Consumer Reports*’ claim that the stopping distance for a Honda Civic is less than 136 feet.”

إذا رفضت الفرض الصفري فيجب أن تستنتج أن "هناك أدلة كافية لدعم ادعاء تقارير المستهلك بأن مسافة التوقف لسيارة هوندا سيفيك أقل من 136 قدمًا".

•If you fail to reject H_0 , you should conclude “there is not enough evidence to support *Consumer Reports*’ claim that the stopping distance for a Honda Civic is less than 136 feet.”

إذا فشلت في رفض الفرض الصفري ، فيجب عليك استنتاج أنه "لا يوجد ما يكفي من الأدلة لدعم ادعاء تقارير المستهلك بأن مسافة التوقف لسيارة هوندا سيفيك أقل من 136 قدمًا".

7.2 Hypothesis Testing for the Mean (Large Samples)

<i>In Words</i>	<i>In Symbols</i>
State the claim mathematically and verbally. Identify the null and alternative hypotheses.	State H_0 and H_a .
Specify the level of significance.	Identify α .
Determine the critical value(s).	
Find the standardized test statistic.	$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \quad \text{if } n \geq 30$

When $n \geq 30$, the sample standard deviation s can be substituted for σ .

Taking Decision – Graphically

Population Mean

First Case:

$$H_0: \mu =$$

$$H_a: \mu \neq$$

Second Case:

$$H_0: \mu \leq$$

$$H_a: \mu >$$

Third Case:

$$H_0: \mu \geq$$

$$H_a: \mu <$$

Population Proportion

First Case:

$$H_0: p =$$

$$H_a: p \neq$$

Second Case:

$$H_0: p \leq$$

$$H_a: p >$$

Third Case:

$$H_0: p \geq$$

$$H_a: p <$$

Two Tail Test

Right Tail Test

Left Tail Test

Example: Testing with Rejection Regions (First way)

Employees in a large accounting firm claim that the mean salary of the firm's accountants is less than that of its competitor's, which is \$45,000. A random sample of 30 of the firm's accountants has a mean salary of \$43,500 with a standard deviation of \$5200. At $\alpha = 0.05$, test the employees' claim.

Note: Critical value = -1.645

يدعي الموظفون في شركة محاسبة كبيرة أن متوسط الراتب لمحاسبي الشركة أقل من متوسط رواتب منافسيها، وهو 45000 دولار. عينة عشوائية مكونة من 30 من محاسبي الشركة متوسط راتبهم 43,500 دولار مع انحراف معياري قدره 5,200 دولار. اختبر مطالبة الموظفين.
القيمة الحرجة = -1.645



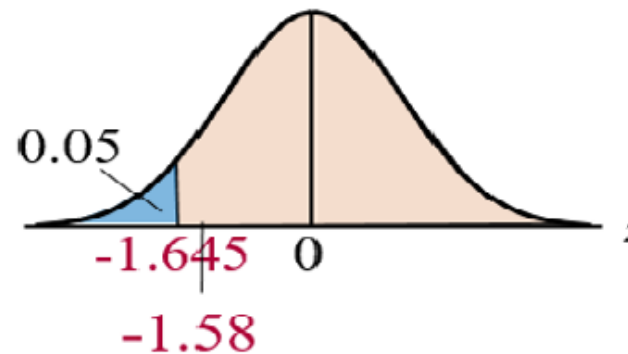
Solution: Testing with Rejection Regions (First way)

- $H_0: \mu \geq \$45,000$
- $H_a: \mu < \$45,000$
- $\alpha = 0.05$
- critical value = -1.645

- **Test Statistic**

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{43,500 - 45,000}{5200/\sqrt{30}} = -1.58$$

- **Rejection Region:**



- **Decision: Fail to reject H_0**

At the 5% level of significance, there is not sufficient evidence to support the employees' claim that the mean salary is less than \$45,000.

Hypothesis Testing by P-value – Second way

Decision Rule Based on P -value

To use a P -value to make a conclusion in a hypothesis test, compare the P -value with α .

1. If $P \leq \alpha$, then reject H_0 .
2. If $P > \alpha$, then fail to reject H_0 .

Example: Hypothesis Testing Using P-values

In an advertisement, a pizza shop claims that its mean delivery time is less than 30 minutes. A random selection of 36 delivery times has a sample mean of 28.5 minutes and a standard deviation of 3.5 minutes. Is there enough evidence to support the claim at $\alpha = 0.01$? Use a P -value. $P\text{-value} = 0.0051$

في أحد الإعلانات، يدعي متجر بيتزا أن متوسط وقت التسليم أقل من 30 دقيقة. يحتوي الاختيار العشوائي لـ 36 وقت تسليم على متوسط عينة يبلغ 28.5 دقيقة وانحراف معياري قدره 3.5 دقيقة. هل هناك أدلة كافية لدعم هذا الادعاء عند

$$P\text{-value} = 0.0051 - \alpha = 0.01$$



Solution: Hypothesis Testing Using P-values

- $H_0: \mu \geq 30 \text{ min}$
- $H_a: \mu < 30 \text{ min}$
- $\alpha = 0.01$
- $P\text{-value} = 0.0051$
- **Decision:** $0.0051 < 0.01$

Reject H_0

At the 1% level of significance, you have sufficient evidence to conclude the mean delivery time is less than 30 minutes.

7.3 Hypothesis Testing for the Mean (Small Samples)

In Words

State the claim mathematically and verbally. Identify the null and alternative hypotheses.

Specify the level of significance and the degrees of freedom

Determine any critical value(s).

Find the standardized test statistic.

In Symbols

State H_0 and H_a .

Identify α .
d.f. = $n - 1$.

Use Table T distribution

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

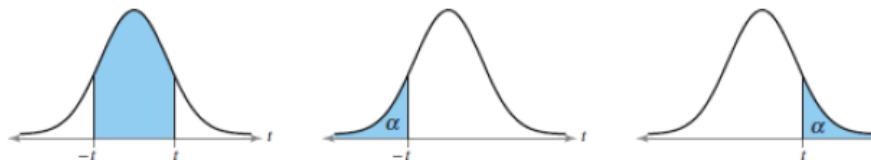
Example: Finding Critical Values for t

Find the critical value t_0 for a left-tailed test given $\alpha = 0.05$ and $n = 21$.

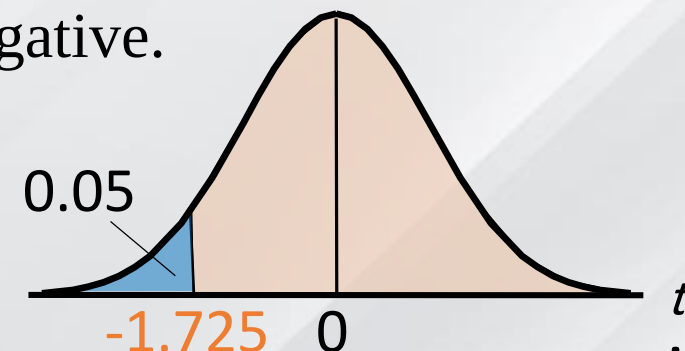
Solution:

- The degrees of freedom are $d.f. = n - 1 = 21 - 1 = 20$.
- Look at $\alpha = 0.05$ in the “One Tail, α ” column.
- Because the test is left-tailed, the critical value is negative.

Table 5 — t-Distribution



	Level of confidence, c					
	0.50	0.80	0.90	0.95	0.98	0.99
	One tail, α					
	0.25	0.10	0.05	0.025	0.01	0.005
	Two tails, α					
	0.50	0.20	0.10	0.05	0.02	0.01
1	1.000	3.078	6.314	12.706	31.821	63.657
2	.816	1.886	2.920	4.303	6.965	9.925
3	.765	1.638	2.353	3.182	4.541	5.841
4	.741	1.533	2.132	2.776	3.747	4.604
5	.727	1.476	2.015	2.571	3.365	4.032
18	.688	1.330	1.734	2.101	2.552	2.878
19	.688	1.328	1.729	2.093	2.539	2.861
20	.687	1.325	1.725	2.086	2.528	2.845
21	.686	1.323	1.721	2.080	2.518	2.831
22	.686	1.321	1.717	2.074	2.508	2.819
23	.685	1.319	1.714	2.069	2.500	2.807



Example: Testing μ with a Small Sample – First Way

A used car dealer says that the mean price of a 2005 Honda Pilot LX is at least \$23,900. You suspect this claim is incorrect and find that a random sample of 14 similar vehicles has a mean price of \$23,000 and a standard deviation of \$1113. Is there enough evidence to reject the dealer's claim at $\alpha = 0.05$? Assume the population is normally distributed.

(Adapted from Kelley Blue Book)

يقول أحد تجار السيارات المستعملة أن متوسط سعر سيارة هوندا موديل 2005 يبلغ 23,900 دولار على الأقل. أنت LX بايلوت تشك في أن هذا الادعاء غير صحيح وتجد أن عينة عشوائية مكونة من 14 مركبة مماثلة يبلغ متوسط سعرها 23000 دولار وانحراف معياري قدره 1113 دولارًا. هل هناك أدلة كافية لرفض دعوى التاجر؟

$\alpha = 0.05$



Solution: Testing μ with a Small Sample

- $H_0: \mu \geq \$23,900$

- $H_a: \mu < \$23,900$

- $\alpha = 0.05$

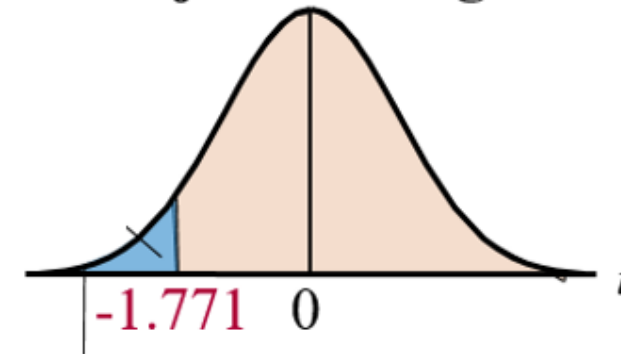
- $\text{df} = 14 - 1 = 13$

- critical value = -1.771

- **Test Statistic:**

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{23,000 - 23,900}{1113/\sqrt{14}} \approx -3.026$$

- **Rejection Region:**



-3.026

- **Decision: Reject H_0**

At the 0.05 level of significance, there is enough evidence to reject the claim that the mean price of a 2005 Honda Pilot LX is at least \$23,900

7.4 Hypothesis Testing for the Proportion p

Verify that $np \geq 5$ and $nq \geq 5$

In Words

State the claim mathematically and verbally. Identify the null and alternative hypotheses.

Specify the level of significance.

Determine any critical value(s).

Find the standardized test statistic.

In Symbols

State H_0 and H_a .

Identify α .

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

Example: Hypothesis Test for Proportions

Zogby International claims that 45% of people in the United States support making cigarettes illegal within the next 5 to 10 years. You decide to test this claim and ask a random sample of 200 people in the United States whether they support making cigarettes illegal within the next 5 to 10 years. Of the 200 people, 49% support this law. At $\alpha = 0.05$ is there enough evidence to reject the claim?

Note: Critical values = -1.96 and 1.96

تزرع منظمة زغبي الدولية أن 45% من الأشخاص في الولايات المتحدة يؤيدون جعل السجائر غير قانونية خلال الخمس إلى العشر سنوات القادمة. قررت اختبار هذا الادعاء واسأل عينة عشوائية مكونة من 200 شخص في الولايات المتحدة عما إذا كانوا يؤيدون جعل السجائر غير قانونية خلال الخمس إلى العشر سنوات القادمة. من بين 200 شخص، 49% يؤيدون هذا القانون.

Critical values = -1.96 and 1.96

$\alpha = 0.05$

$$np = 200(0.45) = 90 \text{ and } nq = 200(0.55) = 110$$



Solution: Hypothesis Test for Proportions

- $H_0: p = 0.45$

- $H_a: p \neq 0.45$

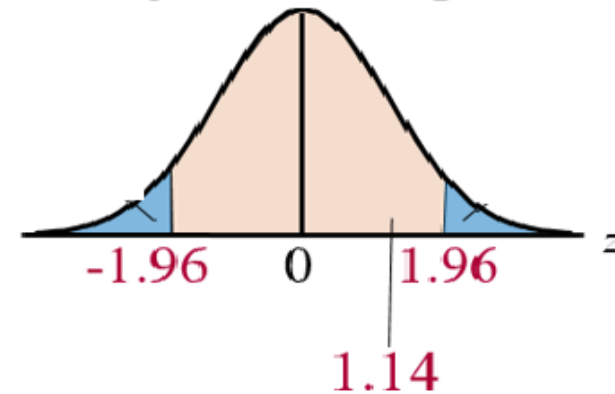
- $\alpha = 0.05$

- critical value = -1.96 and 1.96

- Test Statistic

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.49 - 0.45}{\sqrt{(0.45)(0.55)/200}} \approx 1.14$$

- Rejection Region:



- Decision: Fail to reject H_0

At the 5% level of significance, there is not enough evidence to reject the claim that 45% of people in the U.S. support making cigarettes illegal within the next 5 to 10 years.

Chapter Summary

7.1 Introduction to Hypothesis Testing

7.2 Hypothesis Testing for the Mean (Large Samples)

7.3 Hypothesis Testing for the Mean (Small Samples)

7.4 Hypothesis Testing for Proportions

Un Solved Examples

Chapter

If Null hypotheses $H_0 : P = 0.63$ then Alternative hypotheses $H_a : P \geq 0.63$

إذا كان الفرض الصفري $H_0 : P = 0.63$ فإن الفرض البديل يكون $H_a : P \geq 0.63$

Un Solved Examples

Chapter

If $H_0: \mu \geq 20$, $H_a: \mu < 20$ and $\alpha = 0.01$ and p-value = 0.0041 then we failed to reject Null hypotheses

إذا كان $H_0: \mu \geq 20$, $H_a: \mu < 20$ و $\alpha = 0.01$ and p-value = 0.0041 فإننا نرفض في رفض الفرض الصفري



Thank You