GSBS 6002

Foundations of Business Analysis

Lecture 7

Dr Bronwyn McCredie

PhD (Finance)

Lecturer and Course Co-ordinator

Faculty of Business and Law

Newcastle Business School

Univariate Statistical Analysis

Learning Objectives

1. Implement the hypothesis-testing procedure
2. Know the difference between Type I and Type II errors
3. Explain and interpret the region of acceptance and p-value methods of assessing statistical significance
4. Understand and apply the one-sample t-test.
5. Know when a univariate $\chi 2$ (chi-square) test is appropriate and how to conduct one.

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

Well, Are They Satisfied or Not?

- Is a satisfaction score of 3.9 good or bad?
- If the satisfaction score was 3.7 last quarter does this mean our service has improved or is the result due to random sampling errors?

LO 1

Types of Statistical Analysis

- Univariate Statistical Analysis
  - *Tests of hypotheses involving only one variable.*
  - *Testing of statistical significance*
- Bivariate Statistical Analysis
  - *Tests of hypotheses involving two variables.*
- Multivariate Statistical Analysis
  - *Statistical analysis involving three or more variables or sets of variables.*

LO 1

Hypothesis Testing

- A *hypothesis* is a formal statement of an unproven proposition that is empirically testable
- Types of Hypotheses
  - *Relational hypotheses*
    - Examine how changes in one variable vary with changes in another.
  - *Hypotheses about differences between groups*
    - Examine how some variable varies from one group to another.
  - *Hypotheses about differences from some standard*
    - Examine how some variable differs from some preconceived standard. These tests typify univariate statistical tests.

LO 1

The Hypothesis-Testing Procedure

- Process
  1. *The specifically stated hypothesis is derived from the research objectives.*
  2. *A sample is obtained and the relevant variable is measured.*
  3. *The measured sample value is compared to the value either stated explicitly or implied in the hypothesis.*
     - If the value is consistent with the hypothesis, the hypothesis is supported.
     - If the value is not consistent with the hypothesis, the hypothesis is not supported.

LO 1

Statistical Analysis: Key Terms

- Hypothesis
  - *a formal statement of an unproven proposition that is empirically testable*
- Null Hypothesis
  - *Statement about the status quo.*
  - *Common view on something, mean view*
  - *No change, no association*
- Alternative Hypothesis
  - *what the researcher really thinks is true e.g. the cause of a phenomenon*
  - *Statement that indicates the opposite of the null hypothesis.*

LO 1

The Hypothesis-Testing Procedure

Examples

- NULL H_0: Monetary policy has no impact on financial markets

  ALTERNATE H_A: Monetary policy has an impact on financial markets

- *H_0: Customer expenditure = \$180*

  H_A: Customer expenditure ≠ \$180

  ← Non Directional Two Sided

- *H_0: Average delivery time ≤ 20mins*

  H_A: Average delivery time > 20mins

  ← Directional One Sided

LO 1

Type I and Type II Errors

- Type I Error
  - *An error caused by rejecting the null hypothesis when it is true.*
  - *Practically, a Type I error occurs when the researcher concludes that a relationship or difference exists in the population when in reality it does not exist.*
  - *Has a probability of alpha (α).*

LO 2

Type I and Type II Errors (cont'd)

- Type II Error
  - *An error caused by failing to reject the null hypothesis when the alternative hypothesis is true.*
  - *Practically, a Type II error occurs when a researcher concludes that no relationship or difference exists when in fact one does exist.*
  - *Has a probability of beta (β).*

LO 2

The Law and Type I and Type II Errors

- Our legal system is based on the concept that a person is innocent until proven guilty.
- Type I error -> will send an innocent person to prison.
- Type II error -> guilty party set free.
- Our society places a high value on avoiding Type I errors, even to the extent that Type II errors are more likely to occur.

LO 2

Significance Levels and p-values

- **Significance Level**
  - *A critical probability associated with a statistical hypothesis test that indicates how likely an inference supporting a difference between an observed value and some statistical expectation is true.*
  - *The acceptable level of Type I error.*
  - *Convention is to use 0.05 (5%).*
- **P-value**
  - *Is the probability of observing a test statistic (or one more extreme) if the null hypothesis is true.*

LO 3

If probability (p-value) < 0.05

- *REJECT* null hypothesis
- Statistically significant

If probability (p-value) > 0.05

- *DO NOT REJECT* null hypothesis
- Differences / Associations are not statistically significant
- Only due to sampling variation

LO 3

Using p-values

- H_0: $\mu = 50g$
- H_A: $\mu \neq 50g$

Assume the significance level is 0.05

- Sample mean = 49.05g
- p-value = 0.454

Should we reject or not reject the null hypothesis?

LO 3

Excel Output

Does the average advertised price differ from $255,000?

- H_0: $\mu = \$255000$
- H_A: $\mu \neq \$255000$

	A	B	C
1	t-Test: Paired Two Sample for Means		
2			
3		*Variable 1*	*Variable 2*
4	Mean	244435.055	255000
5	Variance	1.6182E+10	0
6	Observations	1200	1200
7	Pearson Correlation	#DIV/0!	
8	Hypothesized Mean Difference	0	
9	df	1199	
10	t Stat	-2.8770119	
11	P(T<=t) one-tail	0.00204286	
12	t Critical one-tail	1.64612548	
13	P(T<=t) two-tail	0.00408572	
14	t Critical two-tail	1.96194444	

- We reject the null hypothesis (p-value 0.004). The mean advertised price ($244435) is significantly different from $255000.

LO 3

Region of acceptance

Region of Acceptance *- If the test statistic falls within the region of acceptance, the null hypothesis is not rejected.*

The null hypothesis: the mean is equal to 3.0:

$$H_o : \mu = 3.0$$

The alternative hypothesis: the mean does *not* equal 3.0:

$$H_1 : \mu \neq 3.0$$

$$\text{Critical value} - \text{lower limit} = \mu - ZS_{\bar{X}} \quad \text{or} \quad \mu - Z\frac{S}{\sqrt{n}}$$

$$= 3.0 - 1.96\left(\frac{1.5}{\sqrt{225}}\right)$$

$$= 3.0 - 1.96(.1)$$

$$= 3.0 - 0.196$$

$$= 2.804$$

LO 3

An Example of Hypothesis Testing

$$\text{Critical value} - \text{upper limit} = \mu + ZS_{\bar{X}} \qquad \text{or} \qquad \mu + Z\frac{S}{\sqrt{n}}$$

$$= 3.0 + 1.96\left(\frac{1.5}{\sqrt{225}}\right)$$

$$= 3.0 + 1.96(.1)$$

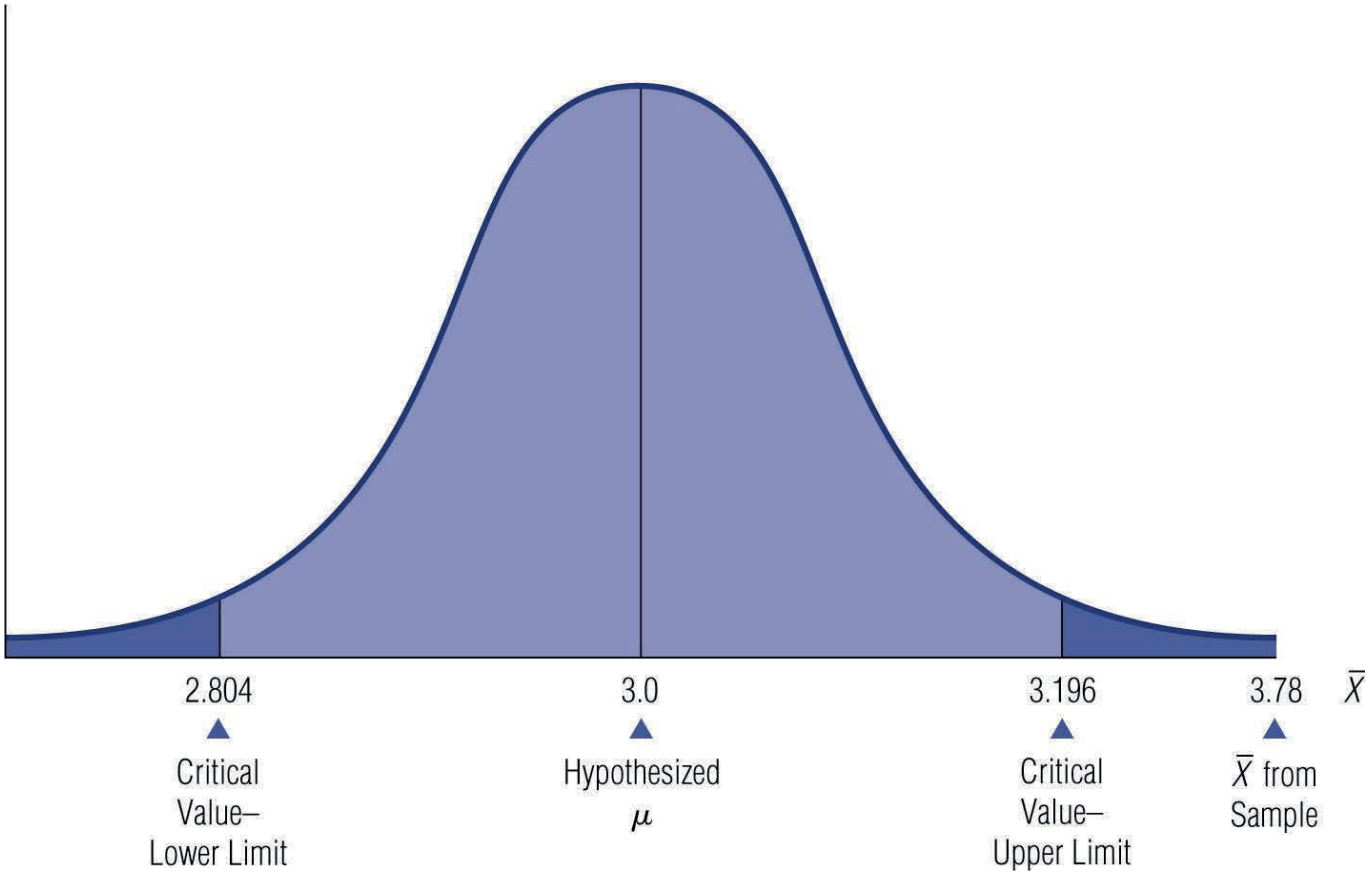
$$= 3.0 + 0.196$$

$$= 3.196$$

LO 3

EXHIBIT 21.3

A Hypothesis Test Using the Sampling Distribution of X-bar under the Hypothesis that μ= 3.0.

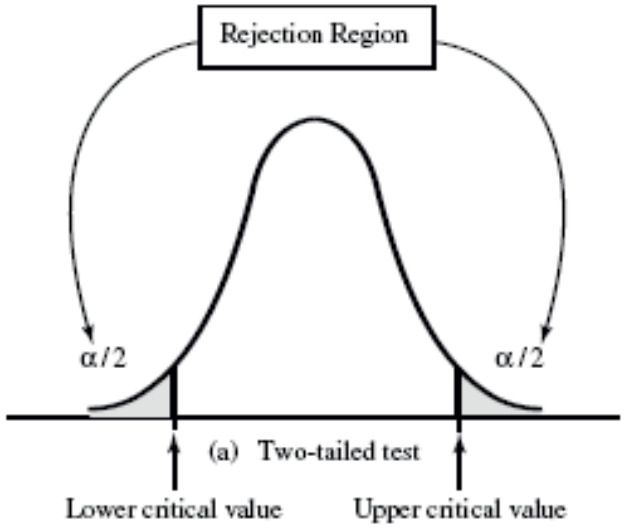


LO 3

Rejection Regions

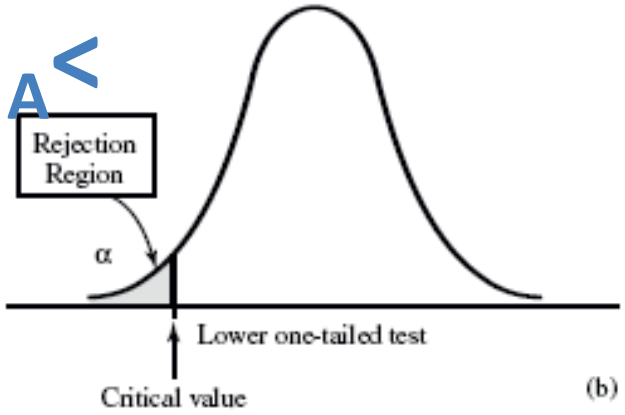
$H_0 =$

$H_A \neq$



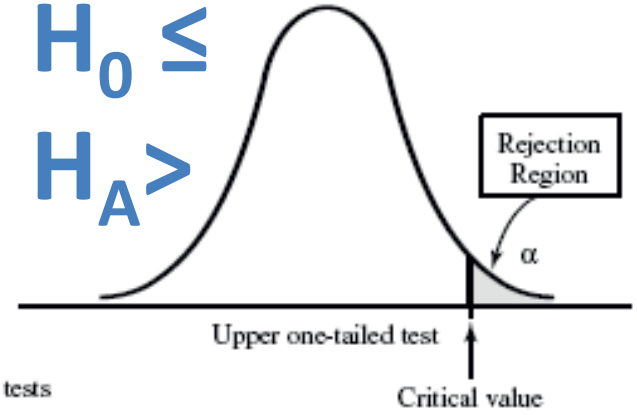
$H_0 \geq$

$H_A <$



$H_0 \leq$

$H_A >$



(b) One-tailed tests

LO 3

Choosing the Appropriate Statistical Technique

- Choosing the correct statistical technique requires considering:
  - *Type of question to be answered*
  - *Number of variables involved*
  - *Level of scale measurement*
- Statistical Techniques:
  - *Tests for difference vs Tests for association.*
  - *Univariate vs Bivariate (or Multivariate) tests.*
  - *Parametric vs Non-Parametric.*

LO 4