ECON 214
Elements of Statistics for Economists

Session 10 – Hypothesis Testing – Part 1

Lecturer: Dr. Bernardin Senadza, Dept. of Economics
Contact Information: bsenadza@ug.edu.gh

UNIVERSITY OF GHANA
College of Education
School of Continuing and Distance Education
2014/2015 – 2016/2017

Session Overview

- A hypothesis is a statement about the value of a population parameter developed for the purpose of testing.
- Hypothesis testing is a procedure, based on sample evidence and probability theory, used to determine whether the hypothesis is a reasonable statement and should not be rejected, or is unreasonable and should be rejected.
- In this session and the next, we explore the concept of hypothesis testing under various scenarios.

Session Overview

- At the end of the session, the student will
  - Be able to define a hypothesis and hypothesis testing
  - Be able to describe the five-step hypothesis testing procedure
  - Be able to distinguish between a one-tailed and a two-tailed test of hypothesis
  - Be able to conduct a test of hypothesis about a population mean

Session Outline

The key topics to be covered in the session are as follows:

- The principle of hypothesis testing
- Steps in hypothesis testing
- One-tailed versus two-tailed tests
- The p-value approach to making decisions about a hypothesis

Reading List

- Michael Barrow, "Statistics for Economics, Accounting and Business Studies", 4th Edition, Pearson
- R.D. Mason , D.A. Lind, and W.G. Marchal, "Statistical Techniques in Business and Economics", 10th Edition, McGraw-Hill

Topic One

THE PRINCIPLE OF HYPOTHESIS TESTING

The principle of hypothesis testing

- A hypothesis is a claim or statement about the value of either a single population parameter or about the values of several population parameters.
- **Example**: Women are paid less, on average, than men.
- Hypothesis testing is about making decisions.
  - It is a procedure, based on sample evidence and probability theory, used to determine whether the hypothesis is a reasonable statement and should not be rejected, or is unreasonable and should be rejected.

The principle of hypothesis testing

- In hypothesis testing there are two conflicting statements about the value of a population parameter(s).
  - **The Null hypothesis (H_0)**
  - **The Alternative hypothesis(H_1 or H_a).**
- Null Hypothesis H_0: A statement about the value of a population parameter(s).
- Alternative Hypothesis H_1: A statement that is accepted if sample data provide evidence that the null hypothesis is false.
- **Example**: *The mean age of Level 200 students is 20 years verses mean age is not 20 years*.

The principle of hypothesis testing

- To test the validity of this hypothesis, we must select a sample from the population, calculate sample statistics and based on certain decision rules, either accept or reject the hypothesis.
- The null hypothesis is initially *presumed* to be true.
  - The analogy of a court of law is a good one here.
  - The accused is presumed innocent (null hypothesis) unless the evidence proves otherwise.
- Evidence is gathered, to see if it is consistent with the hypothesis.

The principle of hypothesis testing

- If it is, the null hypothesis continues to be considered ‘true’ (later evidence might change this).
- If not, the null is rejected in favour of the alternative hypothesis
  - That is, innocence is rejected in favour of a guilty verdict.

The principle of hypothesis testing

- Decision making is never perfect and mistakes can be made.
- There are 2 types of error in hypothesis testing
  - **Type I error**: rejecting the null when it is true (*convicting the innocent*).
  - **Type II error**: accepting the null when it is false (*letting the guilty go free*).

The principle of hypothesis testing

	True situation	
Decision	**H_0 true**	**H_0 false**
Accept H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

Topic Two

STEPS IN HYPOTHESIS TESTING

Steps in hypothesis testing

- In hypothesis testing we must make a decision about the hypothesis;
  - whether it is true or untrue.
- Where do we place the decision line?
- We wish to avoid both Type I and Type II errors.
- We can alter the decision rule to do this.
- Unfortunately, reducing the chance of making a Type I error generally means increasing the chance of a Type II error; hence a trade off.

Steps in hypothesis testing

- **Example**: Accepting a 10-2 majority from the jury to convict (rather than unanimity) reduces the risk of the guilty going free (Type II error), but increases the risk of convicting the innocent (Type I error).
- We wish to make Type I error as small as possible.
- Set the Type I error probability to a particular value: *by convention this is 5%*.
- This is known as the **significance level of the test** and is denoted **α** (probability of rejecting the Null when it is in fact true).
- It is complementary to the confidence level:

  5% significance level ≡ 95% confidence level.

Steps in hypothesis testing

- If the cost of making a Type I error is especially high, then set a *lower* significance level, e.g. 1%.
- The significance level is the probability of making a Type I error.
- The decision is made on the basis of a **test statistic**.
- The Test statistic is a value, determined from sample information, used to determine whether or not to reject the null hypothesis.

Steps in hypothesis testing

- A decision to reject the Null hypothesis or not is made by comparing the **critical value** to the test statistic.
- The **Critical value** is the dividing point between the region where the null hypothesis is rejected and the region where it is not rejected.

Steps in hypothesis testing

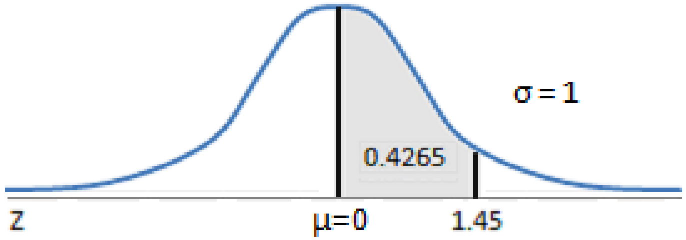
- **Example**: How long do CFLs last?
- A manufacturer of compact fluorescent lamps (CFLs) claims its product lasts at least 5,000 hours, on average.
- A sample of 80 bulbs is tested. The average time before failure is 4,900 hours, with standard deviation 500 hours.
- Should the manufacturer's claim be accepted or rejected?

Example

- $H_0: \mu = 5{,}000$
  $H_1: \mu < 5{,}000$
- **Note**:
  - This is a **one tailed test**, since the rejection region occupies only one side of the distribution (since the H_1 gives a direction. More on this soon).
  - The null hypothesis is always a precise statement (with the equality sign in it).
- Choose significance level of 5% ($\alpha = .05$, meaning the critical value, ***Zc***, is 1.64).
- That is, the Z value that is exceeded with a probability of 0.05 (i.e. $P(Z>Zc)=0.05$, so what is Zc? It is 1.64.

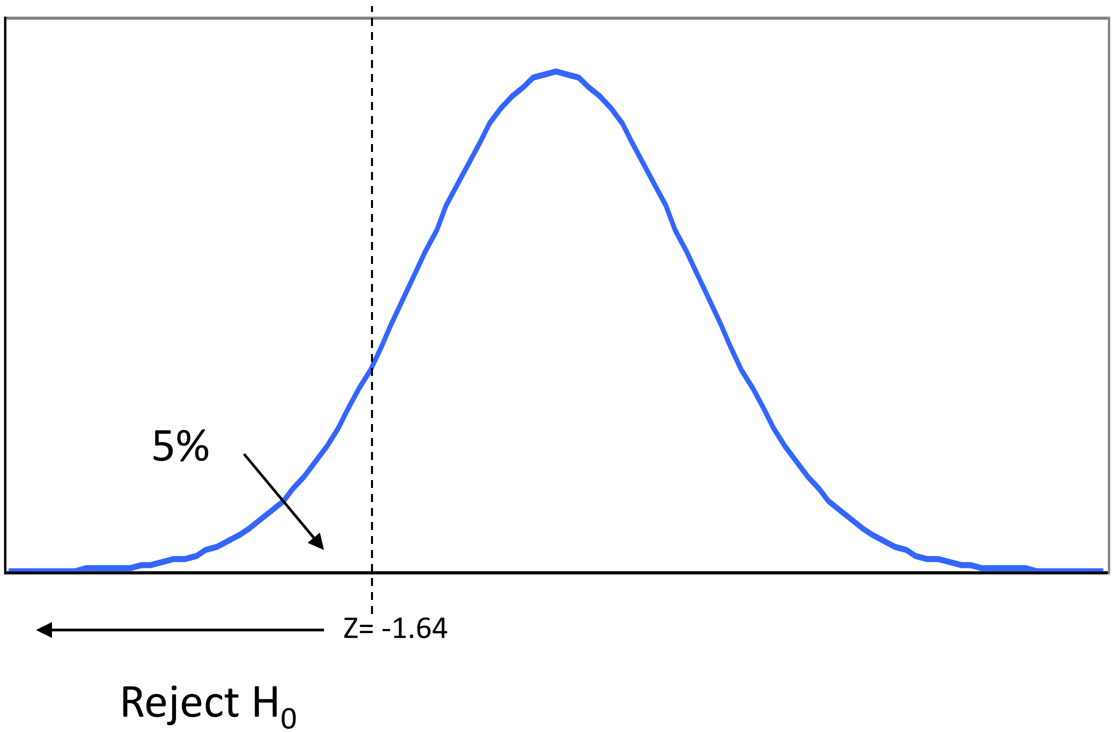
Areas Under the One-Tailed Standard Normal Curve

This table provides the area between the mean and some Z score.
For example, when Z score = 1.45 the area = 0.4265.



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817

Rejection region



- Note: Because the H1 says "less than", i.e. H_1: $\mu < 5{,}000$, it means the rejection region is in the left tail of the normal curve

Example cont'd

- So we must reject the Null hypothesis if Test Statistic is less than -1.64 (since rejection region is in the left tail of normal curve).
- Our sample evidence gives us 4,900, which is less than the hypothesised value of 5,000.
- Question is: **Is 4,900 far enough below 5,000**?
- Is it more than 1.64 standard errors below 5,000? (1.64 standard errors below the mean cuts off the bottom 5% of the Normal distribution).

Example cont'd

- The question we want to ask is: Is the mean indeed less than 5000, or the sample value of 4900 obtained was due to chance (sampling variability?)
- We compute our Test statistic as:

$$Z = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{4900 - 5000}{500/\sqrt{80}} = -1.79$$

Example cont'd

- 4,900 is 1.79 standard errors below 5,000, so falls into the rejection region (bottom 5% of the distribution).
- Hence, we can reject H_0 at the 5% significance level or, equivalently, with 95% confidence.
- If the ***true* mean** were 5,000, there is less than a 5% (3.67%) chance of obtaining sample evidence such as $\overline{X} = 4{,}900$ from a sample of n = 80.

Formal layout of a problem

1. State the hypotheses

   H_0: $\mu = 5{,}000$
   H_1: $\mu < 5{,}000$

2. Choose significance level (probability of rejecting H_0 when true or committing type I error): 5%

3. Look up critical value and state decision rule: *Zc =* 1.64; reject if Z>Zc or –Z<-Zc [or reject if |Z|>Zc]

4. Calculate the test statistic: *Z* = -1.79

5. Decision: reject H_0 since -1.79 < -1.64 and falls into the rejection region

Topic Three

ONE-TAILED VERSUS TWO-TAILED TESTS

One verses two tailed tests

- Should you use a **one-tailed** (H_1: $\mu < 5{,}000$) or **two-tailed** (H_1: $\mu \neq 5{,}000$) test?
- If you are only concerned about falling one side of the hypothesized value (as in the last example, as we would not worry if the bulbs lasted *longer* than 5,000 hours) use the one tailed test.
- You would not want to reject H_0 if the sample mean were anywhere above 5,000.

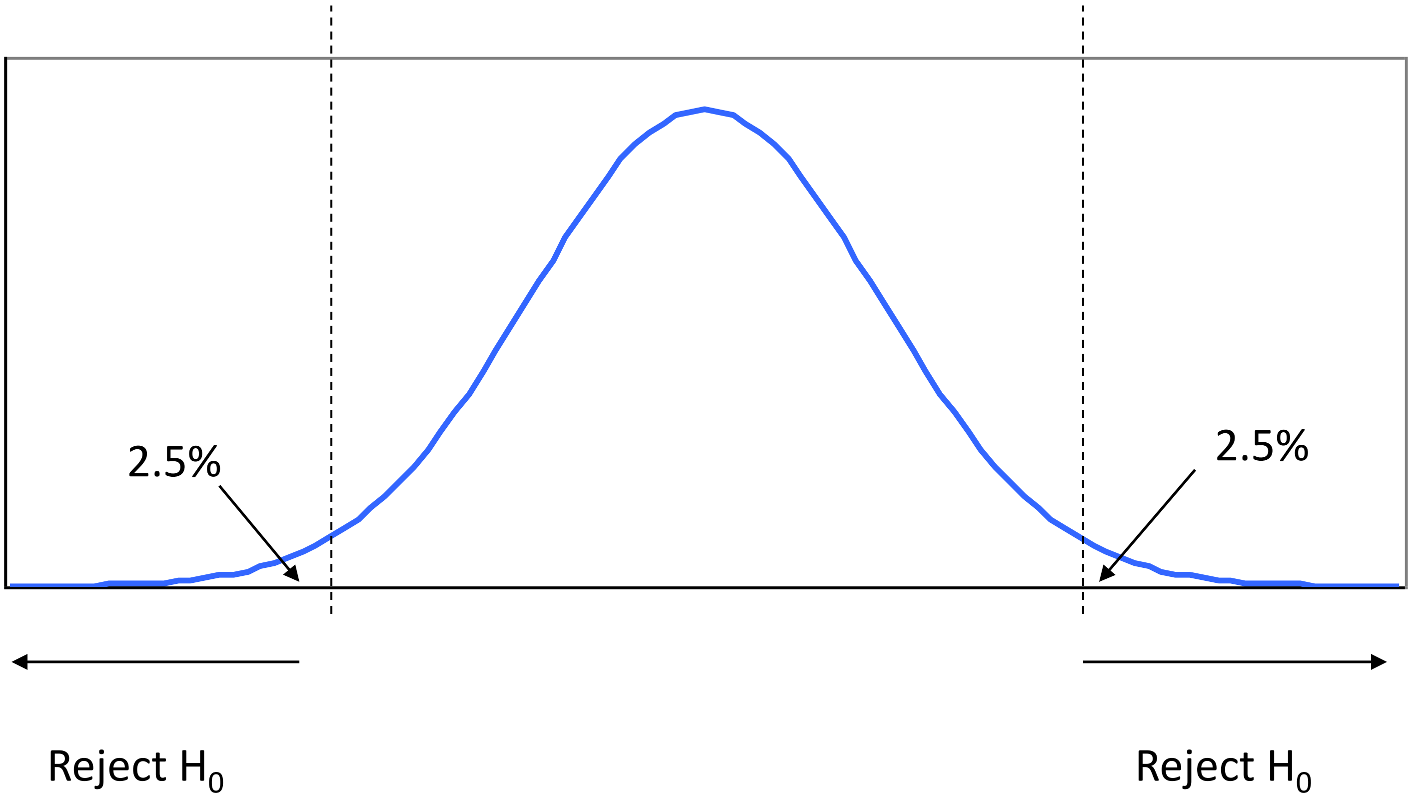
One vs two tailed tests

- If for another reason, you *know* one side is impossible (e.g. demand curves cannot slope upwards), use a one tailed test. Otherwise, use a two tailed test.
- If unsure, choose a two tailed test.
- Never choose between a one or two tailed test on the basis of the sample evidence (i.e. do not choose a one tailed test because you notice that 4,900 < 5,000). The hypothesis should be chosen before looking at the evidence!

Two-tailed test example

- **Example**: It is claimed that an average child spends 15 hours per week watching television. A survey of 100 children finds an average of 14.5 hours per week, with standard deviation 8 hours. Is the claim justified? Test at 5% significance level.
- The claim would be wrong if children spend either *more* or *less* than 15 hours watching TV. The rejection region is split across the two tails of the distribution.
- This is a two tailed test.

A two tailed test – diagram



Solution to the problem

1. $H_0: \mu = 15$
   $H_1: \mu \neq 15$

2. Choose significance level: 5% or α = 0.05

3. Look up critical value: $Z_c = 1.96$; reject H_0 if $Z > Z_c = 1.96$

4. Calculate the test statistic:

$$Z = \frac{\overline{X} - \mu}{s / \sqrt{n}} = \frac{14.5 - 15}{8 / \sqrt{100}} = 0.625$$

5. Decision: we do not reject H_0 since 0.625 < 1.96 and does not fall into the rejection region.

- Topic Four

THE P-VALUE APPROACH

The p-value approach

- There is an alternative way of making the decision.
- Returning to the CFL problem, the test statistic Z = -1.79 cuts off 3.67% in the lower tail of the distribution.
- That is, $P(Z<-1.79)=0.0367$
- .0367 or 3.67% is called the **p-value** for this example.
- Since 0.367 < .05 the test statistic *must* fall into the rejection region for the test.

The p-value approach

- The **p-value** measures the probability of obtaining a sample statistic as extreme as 4900 were the null hypothesis true.
- The level of significance ($\alpha = 0.05$) is the risk level we are willing to tolerate.
- If the **p-value** is less than 0.05, we reject H_0 and we do not reject when the **p-value** is greater than 0.05.

Two ways to reject

Reject H_0 if

- $|Z| > Zc$ i.e. $|-1.79| > 1.64$

or

- the p-value < the significance level (i.e. 3.67% < 5%).

References

- Michael Barrow, “Statistics for Economics, Accounting and Business Studies”, 4th Edition, Pearson
- R.D. Mason , D.A. Lind, and W.G. Marchal, “Statistical Techniques in Business and Economics”, 10th Edition, McGraw-Hill