Business Analytics

Lecture 6

Hypothesis Testing

Dr Yufei Huang

UCL SCHOOL OF MANAGEMENT

Review of Lectures

- Lecture 1: Central tendency and spread measures
- Lecture 2: Probability
- Lecture 3: Random variables
- Lecture 4: The normal distribution
- Lecture 5: Sampling, the central limit theorem, and confidence intervals

Preview of Lectures

- **Lecture 6 (Today): Hypothesis Testing**
- Lecture 7: Simple regression
- Lecture 8: Multiple and nonlinear regressions
- Lecture 9: Revision
  Launch Mock-exam questions
- Lecture 10: Applications of Business Analytics
  Course summary

Hypothesis Testing: Example

Alice is a coffee lover, from a year's data, on average she drinks 10 coffee per week, with standard deviation 2. If there is a week during which she drinks 15 coffee, is it a normal week for her? How about the week during which she drinks 9 coffee?

Big Picture and Motivation

- Hypothesis testing is a **framework** that you can use to test a statement, a concern, or a belief; It is a probabilistic/statistical **language** of “evidence”; It is also the **basis** of scientific rigor.

- Applications
  - Pharmaceutical drug testing, biomedical tests
  - Engineering
  - Strategy
  - Human resources management
  - Operations (production, manufacturing, etc.)
  - Marketing
  - Finance
  - …

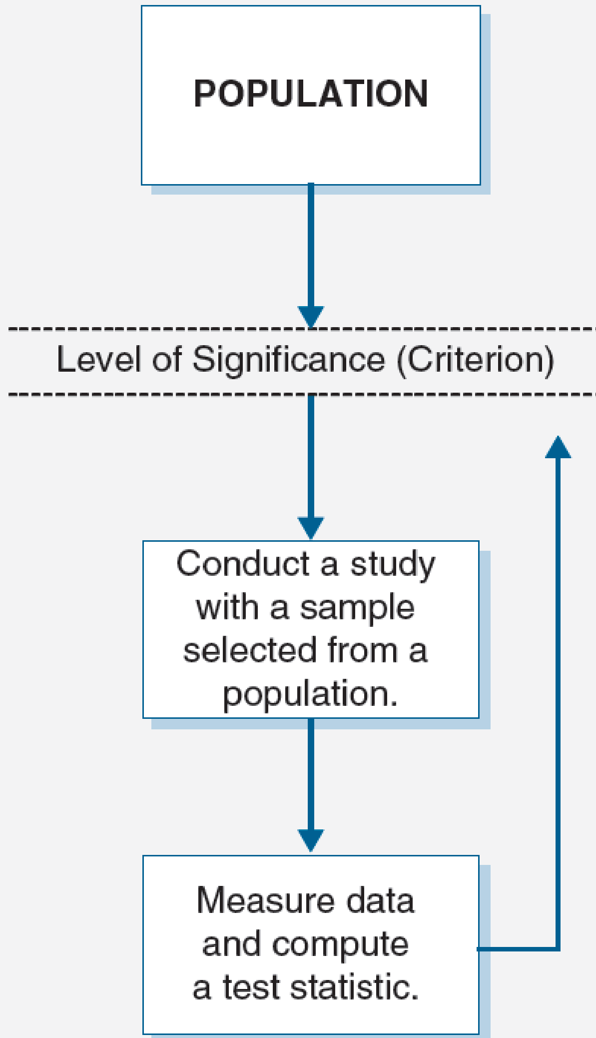
Framework for Hypothesis Testing

STEP 1: State the hypotheses. A researcher states a null hypothesis about a value in the population (H_0) and an alternative hypothesis that contradicts the null hypothesis.

STEP 2: Set the criteria for a decision. A criterion is set upon which a researcher will decide whether to retain or reject the value stated in the null hypothesis.

A sample is selected from the population, and a sample mean is measured.

STEP 3: Compute the test statistic. This will produce a value that can be compared to the criterion that was set before the sample was selected.



STEP 4: Make a decision.
If the probability of obtaining a sample mean is less than 5% when the null is true, then reject the null hypothesis.
If the probability of obtaining a sample mean is greater than 5% when the null is true, then retain the null hypothesis.

Four Steps

- Step 1. Formulate the hypothesis
  - Null Hypothesis (H_0)
  - Alternate Hypothesis (H_A)

- Step 2. Set the criteria for a decision

- Step 3. Acquire an objective test statistic (e.g. from evidence)

- Step 4. Does the objective test statistic represent an overwhelming evidence against the Null Hypothesis? (i.e., p-value < α? or equivalently, compare the test statistic with the critical z-value or t-value)
  - If yes, reject the null (H_0) and accept the alternative (H_A) as the truth.
  - If no, accept the null (H_0) as the truth.

More about Step 4: P-value & Significance

- P-value: probability of observing the statistic (<u>or more extreme</u>) given that the null hypothesis is true.
  - Must compute

- Significance level α: a threshold probability (e.g. 0.05, or 0.1) that determines whether or not the evidence is overwhelming.
  - Typically given

- If the p-value is less than the significance level α, then you can reject the null hypothesis

- How do we find the p-value? Use the sample distribution (normal)!

More about Step 4: T-Statistics

- Instead of comparing p-value with significance level α, we can also compare the test statistics with z-value (or t-value, if use t-distribution)

- We compare the test-statistic with the z-value, z_{α}, which corresponds to the required significance level α. Criteria depend on the type of test.

- Then, we decide to accept H_0 or not.

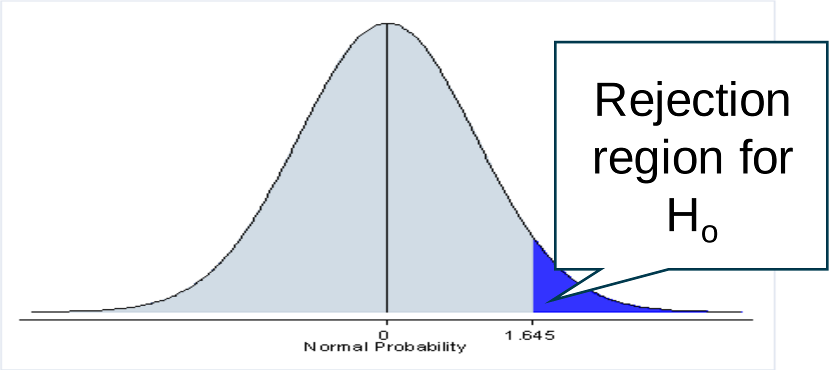
One-tailed and Two-tailed Tests

- One right-tailed test

  Hypotheses:

  H_0: $\mu \leq M$,

  H_1: $\mu > M$

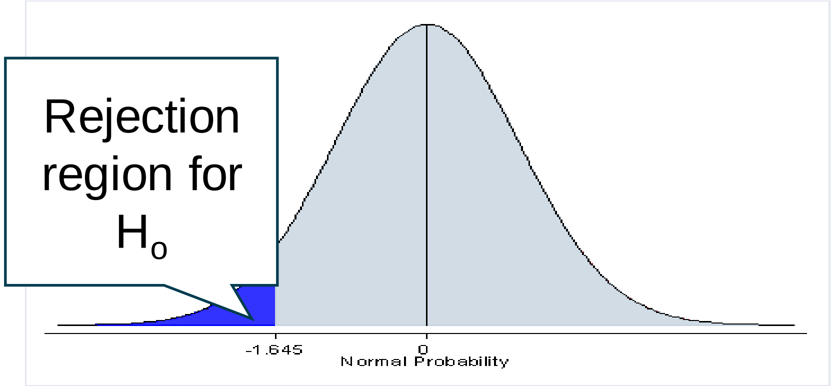


- One left-tailed test

  Hypotheses:

  H_0: $\mu \geq M$,

  H_1: $\mu < M$

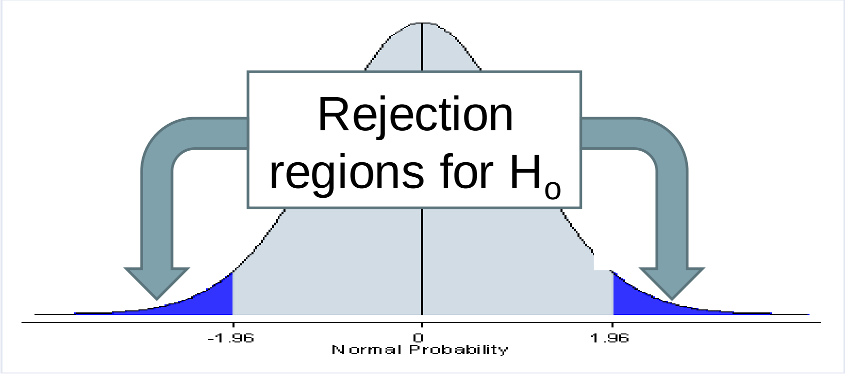


- Two-tailed test

  Hypotheses:

  H_0: $\mu = M$,

  H_1: $\mu \neq M$



:ENT

One-tailed and two-tailed tests

Right tailed	Left tailed	Two-tailed
H_0: μ≤M, H_1: μ>M statistic=($\bar{X}$ -M)/(s/√n) P-value=P(Z>statistic) 1. If P<α, accept H_1 2. If statistic>Z_α, accept H_1	H_0: μ≥M, H_1: μ<M statistic=($\bar{X}$-M)/(s/√n) P-value= P(Z<statistic) 1.If P<α, accept H_1 2.If statistic<Z_α, accept H_1	H_0: μ=M, H_1: μ≠M statistic=($\bar{X}$-M)/(s/√n) If the statistic is positive: P-value=2*P(Z>statistic) If the statistic is negative: P-value=2*P(Z<statistic) 1.If P<α, accept H_1 2.If |statistic|>$Z_{\alpha/2}$, accept H_1 3.If M is outside the (1-α) level confidence interval for μ, accept H_1

Example 1: Connection with Confidence Intervals

A company that delivers packages within London claims that it takes an average of 28 minutes for a package to be delivered from your door to the destination. Suppose that you want to carry out a hypothesis test of this claim.

Set the null and alternative hypotheses:

H_0: $\mu = 28$

H_1: $\mu \neq 28$

Collect sample data:

$n = 100$

$\bar{x} = 31.5$

$s = 5$

Construct a 95% confidence interval for the average delivery times of *all* packages:

$$\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} = 31.5 \pm 1.96 \frac{5}{\sqrt{100}}$$

$$= 31.5 \pm .98 = [30.52,\ 32.48]$$

We can be 95% sure that the average time for all packages is between 30.52 and 32.48 minutes.

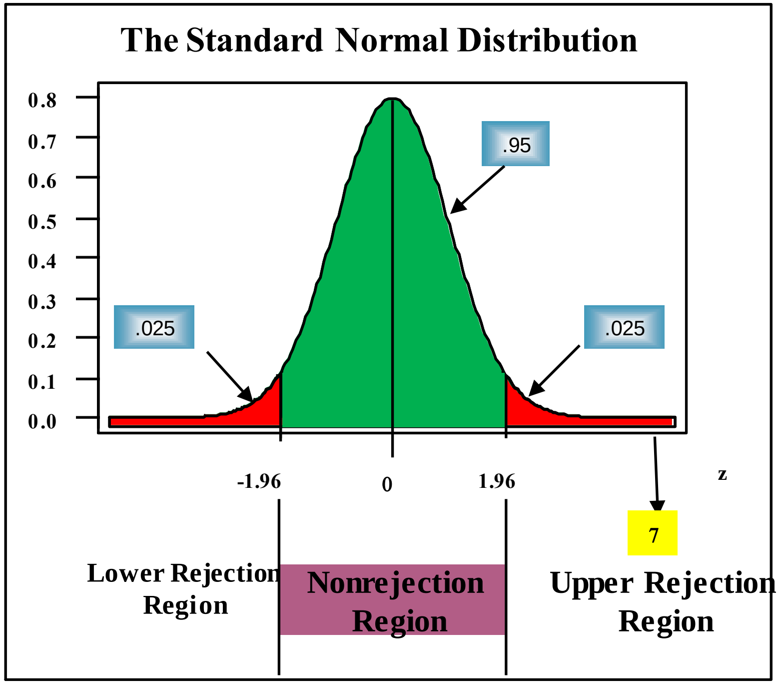
Since the asserted value, 28 minutes, is not in this 95% confidence interval, we may reasonably reject the null hypothesis.

ENT

Example 1: Two-tailed Testing

$$ts = \frac{\text{best guess - null value}}{\text{standard deviation of best guess at } H_0}$$

$$ts = \frac{\overline{X} - \mu}{s / \sqrt{n}} = \frac{31.5 - 28}{5 / \sqrt{100}} = 7$$



$$\text{P-value} = 2P(Z > ts) = 2P(Z > 7) = 2*(0.5 - 0.499999999) \approx 0.000000002$$

Example 2: One-tailed testing

A certain kind of packaged food bears the following statement on the package: "Average net weight 12 oz." Suppose that a consumer group has been receiving complaints from users of the product who believe that they are getting smaller quantities than the manufacturer states on the package. The consumer group wants, therefore, to test the hypothesis that the average net weight of the product in question is 12 oz. versus the alternative that the packages are, on average, underfilled. A random sample of 144 packages of the food product is collected, and it is found that the average net weight in the sample is 11.8 oz. and the sample standard deviation is 6 oz. Given these findings, is there evidence the manufacturer is underfilling the packages?

H_0: $\mu \geq 12$
H_1: $\mu < 12$
n = 144
For $\alpha = 0.05$, the critical value of z is -1.645

The test statistic is:

Do not reject H_0 if: [$z \geq -1.645$]
Reject H_0 if: [$z < -1.645$]

$n = 144$

$\bar{x} = 11.8$

$s = 6$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{11.8 - 12}{\frac{6}{\sqrt{144}}}$$

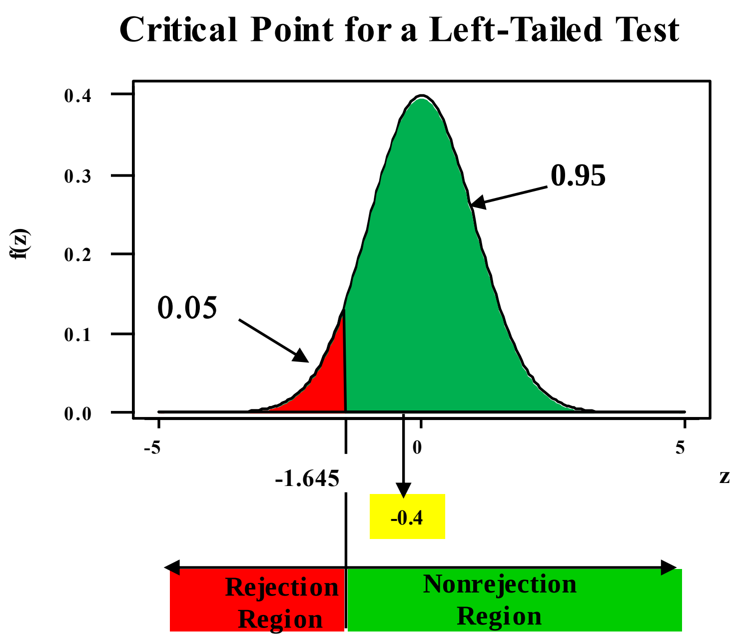
$$= \frac{-.2}{.5} = -0.4 \Rightarrow \text{Do not reject } H_0$$

UCL SCHOOL OF MANAGEMENT

Example 2: One-tailed testing

$$ts = \frac{\text{best guess - null value}}{\text{standard deviation of best guess at } H_0}$$

$$ts = z = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{11.8 - 12}{6/\sqrt{144}} = -0.4$$



$$\text{P-value} = P(Z < ts) = P(Z < -0.4) = 0.5 - 0.1554 \approx 0.3446$$

Exercise

A study was conducted to determine customer satisfaction from real estate deals. Suppose that before the financial crisis, the average customer satisfaction rating, on a scale of 0 to 100, was 77. A survey questionnaire was sent to a random sample of 50 residents who bought new houses recently. The average satisfaction rating for this sample was found to be 84. Assume that the standard deviation of satisfaction level in the population is 28. Did customer satisfaction increase? Answer to a significance level of $p=0.05$.

Exercise

Solution. Hypotheses: H_0: $\mu \leq 77$, H_1: $\mu > 77$. Note that, this is a one-tailed test.

Calculation of test statistic

The sample size is larger than 30, so, by the Central Limit Theorem, the sample mean will be normally distributed with mean 77 and std 28/50^0.5=3.96.

Therefore, $\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} = \frac{\overline{X} - 77}{3.96}$ will follow a standard normal distribution

Our test statistic is: $Z = \frac{\overline{X} - 77}{3.96} = \frac{84 - 77}{3.96} = 1.7676 > 1.64$

The p value is: $P(Z>1.7676)=1-P(Z<1.7676)=0.0385<0.05$

Conclusion: as the probability of obtaining a sample mean of 84 is smaller than the required significance level, we accept H_1 and reject H_0: customer satisfaction increased.

Types of errors

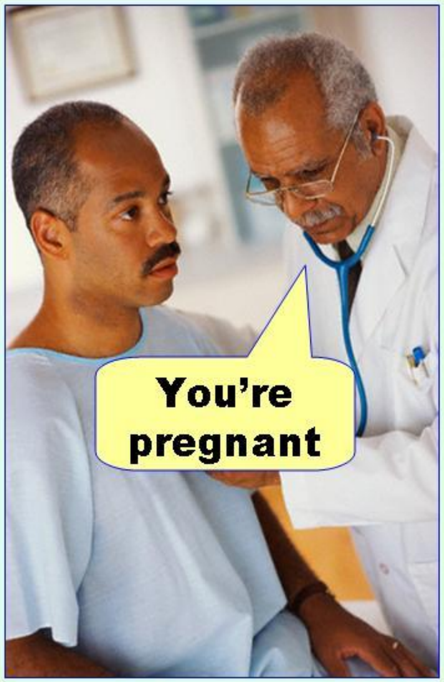
- We have to make an accept/reject decision
- We cannot work on a 100% confidence level, so there can always be errors.
- No error:
  - We accept H_0 and H_0 is true
  - We reject H_0 and H_0 is false
- Error:
  - We reject H_0 and H_0 is true: Type I error
  - We accept H_0 and H_0 is false: Type II error

UCL SCHOOL OF MANAGEMENT

Example

Type I error

(false positive)



Type II error

(false negative)

Not pregnant
=> H_0 is true

Pregnant
=> H_0 is false

H_0 : you are not pregnant

H_1 : you are pregnant

Type I error:

H_0 is true,

We reject H_0.

Type II error:

H_0 is false

We accept H_0

Reference

- Chapters 7 of:

Aczel, A., & J. Sounderpandian. 2008. Complete Business Statistics. McGraw-Hill/Irwin, Seventh Edition.

UCL
SCHOOL OF MANAGEMENT