Preliminary Statistics

Lecture 5: Hypothesis Testing

SOAS

University of London

Sophie van Huellen (sv8@soas.ac.uk), September 2013

Outline

- Elements/Terminology
- Errors
- Procedure of Testing
- Significance Level
  - Type of errors again
  - Trade-off between Errors
- CI Approach to Hypothesis Testing
- Hypothesis Testing

Elements of Hypothesis Tesing

- In testing we want to make inferences about the unknown population parameters.
- This is done in the form of tests of some particular hypotheses.
- Terminology:
  - **Null Hypothesis**: H_0
  - **Alternative Hypothesis**: H_1
  - **Test Statistic**: A function of the sample upon which the decision will be based.
  - **Decision Rule**: Rule which specifies when to accept or reject the null hypothesis.

Elements of Hypothesis Tesing (cont.)

An Analogy: Criminal Trial

- **Null Hypothesis** → Defendant is innocent
- **Alternative Hypothesis** → Defendant is guilty
- **Test Statistic** → Evidence against the Defendant
- **Decision Rule** → If the jury finds the evidence convincing beyond reasonable doubt, the defendant is found guilty (i.e. if the evidence from the sample seems inconsistent with the null hypothesis, H_0 is rejected).

Errors

- We can never know with certainty if the null or the alternative hypothesis is true.
- 2 Options:
  - Not reject the Null (acquit the defendant)
  - Reject the Null (find the defendant guilty)
- But… Mistakes happen…

Errors (cont.)

Type I and Type II Errors

	H_0 is true	H_0 is false
Do not reject (accept) H_0	Correct	**Type II Error** (Set a guilty person free)
Reject H_0	**Type I Error** (Convict an innocent person)	Correct

Type I Error = Prob (Reject H_0| H_0 true)
Type II Error = Prob (Accept H_0| H_0 false)

Errors (cont.)
Minimise the Probability of Errors

- Avoid completely Type I errors:
  - Always accept the Null – Acquit everyone
  - … but we would make lots of Type II errors.
- Avoid completely Type II errors:
  - Always reject the Null – Convict everyone
  - … but we would make lots of Type I errors.

- Trade-off the two risks
- **In statistical testing, we fix the probability of Type I error!**

Procedure of Testing

1. Specify the Null Hypothesis:

$$\mathrm{H}_0\colon\ \theta = \theta_0.$$

2. Specify the Alternative Hypothesis
   - Two-sides test: $\mathrm{H}_1\colon\ \theta \neq \theta_0$.
   - One-side test: $\mathrm{H}_1\colon\ \theta > \theta_0$ or $\theta < \theta_0$.
3. Design the Test Statistic assuming that the Null is true:

$$Z = \frac{\hat{\theta} - \theta_0}{SE(\hat{\theta})}$$

Procedure of Testing (cont.)

4. Find how the Test Statistic is distributed under the Null:
   - For large samples ($n > 30$), the test statistic is normally distributed:

$$Z \sim SN(0,1)$$

   - For small samples <u>AND</u> if the true variance σ^2 is unknown, the test statistic is distributed as a t:

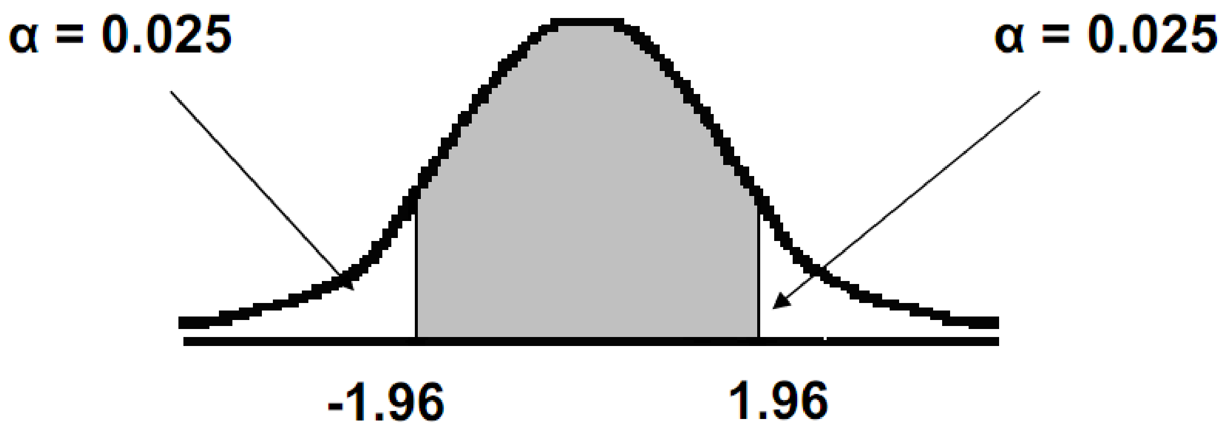
$$Z \sim t_{(n-1)}$$

Procedure of Testing (cont.)

5. Set the *significance level* of the test (conventionally $\alpha = 0.05$ i.e. 5%) and find the critical values in the tails that give the $1-\alpha$ or 95% of the distribution:
   - For a 2-sided test, α would be equally split between the 2 tails,
   - For a 1-sided test, the entire level of significance will be located in one tail.
6. Apply the Decision Rule
   - If $|Z| >$ critical value, reject the Null,
   - Otherwise, Do not Reject the Null.

Procedure of Testing (cont.)
Decision Rule

- Assume:
  - $\alpha = 0.05$ (significance level)
  - $(\alpha - 1) = 0.95$ (confidence coefficient)
- Rejection Rule:
  - $Z > 1.96$ or $Z < -1.96$
  - P-value < 0.05



Procedure of Testing (cont.)

Testing for the Mean

1. Null Hypothesis:

$$\mathrm{H}_0\colon \mu = \mu_0$$

2. Alternative Hypothesis:

$$\mathrm{H}_1\colon \mu \neq \mu_0$$

3. Test Statistic under the Null:

$$Z = \frac{\hat{\mu} - \mu_0}{SE(\hat{\mu})}$$

4. Distribution of the Test Statistic, under the Null:

$$\mu \sim N\left(\mu_0, \frac{\sigma^2}{N}\right) \Rightarrow Z \sim SN(0,1)$$

5. Critical Values:

$$-z_{0.025} = -(z_{0.025}) = -(1.96)$$

6. Decision Rule $|Z| > 1.96$ reject the Null.

Significance Level

- The significance level corresponds to the probability of Type I error by construction.
- The Testing procedure says:
  - If the Null hypothesis were true, then:

$$\mu \sim N\left(\mu_0, \frac{\sigma^2}{N}\right) \Rightarrow Z \sim SN(0,1)$$

- The Decision Rule says:
  - Whenever we find a test statistic in the extreme 5% of the true distribution, reject the null (though we assumed it is true),
  - Hence: P(Reject Null | Null true) = α%.

Significance Level (cont.)
Types of Error again

- **Type I error** is the error of rejecting H_0 when it is true.
  - The probability of a type I error is α, the level of significance of the statistical test.
- **Type II error** is the error of not rejecting H_0 (the null hypothesis) when it is false.
  - The probability of committing a type II error is designated as β.
- The **power of a test** is the probability of rejecting a false hypothesis.
  - $(1 - \beta)$, not committing a type II error, is the power of the test.

Significance Level (cont.)
Types of Error again

		Decision of the test	
		Do not reject H_0	Reject H_0
True parameter	H_0 is true	Correct decision (prob. $(1-\alpha)$ = confidence interval)	**Type I error** (prob. α = significance level)
	H_0 is false	**Type II error** (prob. β)	Correct decision (prob. $(1-\beta)$ = power of the test)

Significance Level (cont.)
Trade Off between Type I and Type II

- Unfortunately decreasing the probability of a Type I error has the effect of increasing the probability of a Type II error and vice versa.

- Only by increasing the sample size can the probability of one type of error be reduced without increasing the probability of the other.

Confidence Interval Approach to Hypothesis Testing

- Construct a confidence interval around the point estimate.
- If the hypothesised value falls outside of the interval, reject the null hypothesis.
- The values within the confidence interval are the set of all acceptable hypothesis.

- *Example:*
  - Suppose the random variable Y (the height of basketball players) is normally distributed with mean μ and variance σ^2
  - We have a random sample of n basketball players' height, Y_i, 1,2,…n

Confidence Interval Approach to Hypothesis Testing (cont.)

Example (cont.):

- Estimating the ture mean (average heigth of players)
- Estimator for μ:

$$\hat{\mu} = \frac{\sum Y_i}{n}$$

- Distribution of the sample mean:

$$\hat{\mu} \sim N(\mu, \frac{\sigma^2}{n})$$

- Estimate for the standard error of the sample mean:

$$SE(\hat{\mu}) = \frac{s}{\sqrt{n}}$$

Confidence Interval Approach to Hypothesis Testing (cont.)

Example (cont.):

- **Point Estimate:**

  Suppose we get a sample $n = 100$ and find an estimate for sample mean $\hat{\mu} = 2.05m$ and an estimate for $\mathrm{SE}(\hat{\mu}) = 0.05m$.

  - What can we infer about the true (population) mean?
  - *Interpretation:* “The true mean height of all basketball players is 2.05m”

Confidence Interval Approach to Hypothesis Testing (cont.)

Example (cont.):

- **Interval Estimate**
  - Pinpoint how confident we are that the true value lies within a range around the sample mean estimate, i.e. 2.05m.
  - 95% Confidence Interval:
    $$P(2.05 - 1.96 * 0.05 < \mu < 2.05 + 1.96 * 0.05) = 95\%$$
    $$P(1.952 < \mu < 2.148) = 0.95$$
  - *Interpretation:* "With 95% confidence, the true average height of all basketball players will lie between 1.952m and 2.148m"

Hypothesis Testing

Example (cont.):

- Assume we have reasons to believe that the true average height of all players is $\mu_0 = 1.99$m.
- Check this Hypothesis:
  - State the Null: H_0: $\mu = \mu_0 = 1.99$
  - State the Alternative: H_1: $\mu \neq \mu_0 \neq 1.99$
  - Construct the test statistic assuming H_0 is true:

$$\tau = \frac{\hat{\mu} - \mu_0}{SE(\hat{\mu})} = \frac{2.05 - 1.99}{0.05} = 1.2$$

Hypothesis Testing (cont.)

- Distribution of the test statistic, under the null hypothesis:
  - We have a large enough sample so that:
    $$\tau \sim SN(0,1)$$
  - The critical values for the 5% significance level of the distribution of the test statistic are: $\pm$ 1.96.
- Decision Rule: If $\tau = 1.2 < 1.96$ accept the null.
- *Interpretation:* "At the 95% confidence level, the true mean height of all players is not statistically different from 1.99."

Hypothesis Testing (cont.)
Confidence Interval Approach

- The result obtained by the hypothesis test can also be viewed from the Confidence Interval-side.

- Since the hypothesised value (1.99m) is included in the confidence interval for the true mean [1.952m; 2.148m], we cannot reject the null hypothesis, that is 1.99 is one of the acceptable values for the true mean.