Hypothesis Testing
CSE 8340
Empirical Software Engineering

A. Güneş Koru

November 19, 2002

An introduction to the topic that explains basic concepts.

- Hypotheses
- Decision Problem
- The Standard Format of Hypothesis Testing

Hypotheses

- Hypothesis: An assumption or concession made for the sake of argument.

- Hypothesis Testing: Choose between two competing hypotheses about the value of a population parameter using the knowledge obtained from a sample.

  – Simple hypothesis: One value of the population parameter.
    * $\mu = 115$,
    * $\mu_1 - \mu_2 = 0\,(exact\,difference)$, etc.

  – Composite hypothesis: A range of values that the population parameter may assume ($\mu \neq 115$).

- Null Hypothesis (H_0): Status quo. Changes nothing.

- Alternative Hypothesis (H_a): Believed to be true.

- Both can be simple or composite.

Hypotheses (cont...)

- Example: Mean IQ.
  - $H_0 : \mu = 100$, $H_a : \mu > 100$
  - $H_0 : \mu_1 - \mu_2 = 0$, $H_0 : \mu_1 - \mu_2 \neq 0$
- The population parameter should be included in one of these two sets.
- One way to assure is using complementary sets.
- One equality statement as the null hypothesis, a composite alternative hypothesis.
- The values specified by the alternative hypothesis:
  - $One\,Sided\,(tailed)\,test$: Either below or above the value specified in the equality.
  - $Two\,Sided\,(tailed)\,test$: Can be both sides.

Decision Problem

- Accept or reject the null hypothesis based on the evidence.
  - Question: how likely the population parameter can take this value if my null hypothesis is true.
  - Answer is a probability value found by statistical means.
  - Larger sample, more accurate decisions.
- Acceptance or rejection but not proof.

	Reality	
	H_0 is true	H_0 is false
Accept H_0	Correct $(Confidence\,Level = 1 - \alpha)$	$Type\,II\,error$ (β)
Reject H_0	$Type\,I\,error$ (α)	Correct $(Power\,of\,the\,test = 1 - \beta)$

Table 1: Four possible decisions
(Note that column sums are 1)

Decision Problem (cont...)

- α is the $level\,of\,significance$.
- $Confidence\,level = 1 - \alpha$. The complement of $Type\,I$ error.
- $1 - \beta$ is the $power$ of the test.
- One unit change in α does not cause such a change in β.
- If n (sample size) is constant and $\alpha \uparrow$, then $\beta \downarrow$.
- If $n \uparrow$, then $\alpha \downarrow$and $\beta \downarrow$.

The standard format of hypothesis testing

1. State the null and alternative hypotheses.

   - Clear and simple null hypothesis
   - Mutually exclusive null and alternative hypotheses
   - Population parameter should be included in either the null or the alternative hypothesis

2. Determine the appropriate test statistic

   - Test statistic is a random variable used to determine how close a specific sample result falls to one of the hypotheses being tested.
     - Its p.d.f. must be known when it is assumed that the null hypothesis is true.
     - It must contain the parameter being tested.
     - All of its remaining terms must be known and calculable from the sample.
   - If $H_0 : \mu = 130$, the best estimate of μ is $\overline{x}$. Then

the standardization of $\overline{x}$,

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$$

can be used as the test statistic, where μ_0 is the mean specified under the null hypothesis, σ is its known value.

3. Determine the critical regions (Fig. 8.4. p. 315)

   - The set of values that will lead to
     - rejection of H_0: critical region
     - acceptance of H_o: acceptance region
   - Decide on the level of significance, α, how much you can accept wrongly rejecting H_0 when it is true.
   - Social sciences $\alpha = 0.05$ and medical sciences $\alpha = 0.01$ or $\alpha = 0.005$.
   - From a table, look up the z value, that matches the level of significance required.
   - Calculate the critical values using this z value using the above formula.
   - Critical value is the point that separates these two regions (Fig 8.5. p. 316).

4. Compute the value of the test statistic

5. Make the statistical decision and interpretation

Reference:

Donald L. Harnett and James L. Murrey, Introductory Statistical Analysis, 2nd Edition,1980, pp. 305-320, Addison Wesley Publishing Company, ISBN 0-201-02758-5.