STAT 401A - Statistical Methods for Research Workers

Pvalues

Jarad Niemi (Dr. J)

Iowa State University

3 Sep 2013

Hypotheses

- Alternative hypothesis (H_A)
  - Your scientific hypothesis about the world.
- Null hypothesis (H_0)
  - The opposite of the alternative hypothesis.
  - Usually a simpler state of affairs.

If δ is a parameter, i.e. expected difference in the response between group A and group B,

- Two-sided hypothesis: $H_0 : \delta = 0$ vs $H_A : \delta \neq 0$
- One-sided hypothesis: $H_0 : \delta \leq 0$ vs $H_A : \delta > 0$
  $H_0 : \delta \geq 0$ vs $H_A : \delta < 0$

P-value

Definition

A statistic is a numerical quantity calculated from data. A test statistic is a statistic used to measure the plausibility of an alternative hypothesis relative to a null hypothesis.

Definition

A pvalue is the probability of observing a test statistic as or more extreme than that observed if the null hypothesis is true.

Definition

The sampling distribution of a test statistic is the distribution of the test statistic under the null hypothesis.

What if the null hypothesis is true?

No difference between groups (treatment has no effect).

Individual	A	B	C	D	Difference of
Response	1	2	3	4	the averages
	Treatment				
Observed	−	−	+	+	2
Scenario 1	−	+	−	+	1
Scenario 2	−	+	+	−	0
Scenario 3	+	−	−	+	0
Scenario 4	+	−	+	−	-1
Scenario 5	+	+	−	−	-2

P-value

The randomization (sampling) distribution of the test statistic is uniform over the numbers:

$$-2 \quad -1 \quad 0 \quad 0 \quad 1 \quad 2^*$$

* indicates the observed test statistic

Let δ be the true difference between the treatments.

- If $H_0 : \delta = 0$ vs $H_A : \delta \neq 0$, then being farther away from 0 is extreme. Since we observed 2 and both -2 and 2 are the same distance from 0, our p-value is $2/6 = 1/3$.
- If $H_0 : \delta \leq 0$ vs $H_A : \delta > 0$, then being positive and farther away from 0 is extreme. Since we observed 2 and nothing is more extreme, our p-value is $1/6$.
- If $H_0 : \delta \geq 0$ vs $H_A : \delta < 0$, then being negative and farther away from 0 is extreme. Since we observed 2 and everything else is more extreme, our p-value is $6/6$.

Summary

- randomization distribution of the test statistic provides a gold standard
- small p-values provide evidence against the null hypothesis