S1211Q Introduction to Statistics
Lecture 19

Wei Wang

August 1, 2012

P-Value

- To report the result of a hypothesis-testing analysis is to simply say whether the null hypothesis was rejected at a specified level of significance. This type of statement is somewhat inadequate because it says nothing about whether the conclusion was a very close call or quite clear cut.

- P-value is a quantity that conveys much information about the strength of evidence against H_0 and allows an individual decision maker to draw a conclusion at any specified level α.

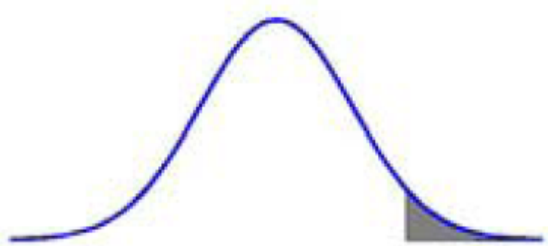
- The P-value (*observed significance level*) is the probability, under the null hypothesis, that the test statistic is more ***extreme*** than the observed statistic.

Example cont.

Ex. (Defective rate cont.) A factory claims that less than 10% of the components they produce are defective. A consumer group is skeptical of the claim and checks a random sample of 300 components and finds that 39 are defective. Is there evidence that 10% of all components made at the factory are defective?

If H_0 is true, $\mathrm{Z} = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim \mathrm{N}(0,1)$

$$\hat{p} = \frac{39}{300} = 0.13 \quad \mathrm{Z} = \frac{0.13 - 0.1}{\sqrt{0.1(1-0.1)/300}} = 1.72$$



P(Z>1.72) = $\boxed{0.0416}$ ← *P-value*

Remarks

- P-value is corresponding to the smallest level of significance at which H_0 would be rejected when a specified test procedure is used on a given data set. The smaller the P-value, the more contradictory is the data to H_0.

- Once the P-value has been determined, the conclusion at any particular level α results from comparing the P-value to α:
  1. P-value $\leq \alpha \rightarrow$ reject H_0 at level α.
  2. P-value $> \alpha \rightarrow$ do not reject H_0 at level α.

- To calculate P-value:
  1. Calculate the test statistic as before.
  2. Compute probability that we will reject the null if the threshold is the test statistic obtained from 1.

- Question: what is the relationship of P-value of the one-sided test and the P-value of the two-sided test?

Two sample tests

- A new drug is claimed to significantly reduce the blood pressure for high blood pressure patients. What kind of tests can we use to verify the claim?

- A new drug is claimed to perform much better in terms of reducing blood pressure than an old drug. What kind of tests can we use to verify the claim?

Things to cover

- As in the one sample testing problem, we will cover the following cases:
  1. Two normal populations with known variance.
  2. Two populations with unknown distribution and large sample size.
  3. Two normal populations with unknown variance.
  4. Two population proportions with large sample size.
  5. Tests about variances. (NOT required.)

- Basic assumptions for comparing population means:
  1. $X_1, X_2, \ldots, X_m$ is a random sample (i.i.d.) from a population with mean μ_1 and variance σ_1^2.
  2. $Y_1, Y_2, \ldots, Y_n$ is a random sample (i.i.d.) from a population with mean μ_2 and variance σ_2^2.
  3. The X and Y samples are independent of one another.

Test statistics

- Since we are comparing the population means, a natural test statistic to use would be the difference of two sample means. Because of independence we have,

$$
\begin{aligned}
E(\bar{X} - \bar{Y}) &= \mu_1 - \mu_2 \\
Var(\bar{X} - \bar{Y}) &= \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}
\end{aligned}
$$

Case I: normal, known variance

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$

Test statistic: $\dfrac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} \sim \mathrm{N}(0,1)$

vs Alternative Hypothesis:

$H_a : \mu_1 - \mu_2 > \Delta_0$, reject if $\dfrac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} > Z_\alpha$

$H_a : \mu_1 - \mu_2 < \Delta_0$, reject if $\dfrac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} < -Z_\alpha$

$H_a : \mu_1 - \mu_2 \neq \Delta_0$, reject if $\dfrac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} < -Z_{\alpha/2}$ or $\dfrac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} > Z_{\alpha/2}$

Questions

- How to compute P-value for case I?
- How to compute type II errors for case I?
- In a balanced design, derive the sample size calculation formula (for alternative “>”):

$$m = n = \frac{(\sigma_1^2 + \sigma_2^2)(Z_\alpha + Z_\beta)^2}{(\Delta' - \Delta_0)^2}$$

Case II: large sample

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$

Test statistic: $\dfrac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}}} \sim A\mathrm{N}(0,1)$

vs Alternative Hypothesis:

$H_a : \mu_1 - \mu_2 > \Delta_0$, reject if $\dfrac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}}} > Z_\alpha$

$H_a : \mu_1 - \mu_2 < \Delta_0$, reject if $\dfrac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}}} < -Z_\alpha$

$H_a : \mu_1 - \mu_2 \neq \Delta_0$, reject if $\dfrac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}}} < -Z_{\alpha/2}$ or $\dfrac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}}} > Z_{\alpha/2}$

Questions

- How to construct confidence interval for $\mu_1 - \mu_2$ in case II?

Case III: normal, unknown variance

$H_0 : \mu_1 - \mu_2 = \Delta_0$

Test statistic: $\dfrac{\bar{X} - \bar{Y} - \Delta_0}{\sqrt{\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}}} \sim t_\nu$, ν is the df of the t-distribution and it's approximately estimated

by the sampled data: $\nu = \dfrac{\left(\dfrac{s_1^2}{m} + \dfrac{s_2^2}{n}\right)^2}{\dfrac{(s_1^2/m)^2}{m-1} + \dfrac{(s_2^2/n)^2}{n-1}}$, and round ν town to the nearest integer.

Case III cont.

vs Alternative Hypothesis:

$$H_a : \mu_1 - \mu_2 > \Delta_0 \text{, reject if } \frac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} > t_{\alpha,\nu}$$

$$H_a : \mu_1 - \mu_2 < \Delta_0 \text{, reject if } \frac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} < -t_{\alpha,\nu}$$

$$H_a : \mu_1 - \mu_2 \neq \Delta_0 \text{, reject if } \frac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} < -t_{\alpha/2,\nu} \text{ or } \frac{\overline{X} - \overline{Y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} > t_{\alpha/2,\nu}$$

Questions

- How to compute P-values of the test?
- How to construct confidence interval for $\mu_1 - \mu_2$ in case III?
- What if we know that $\sigma_1^2 = \sigma_2^2$?

  The *pooled estimator* of $\sigma^2 = \sigma_1^2 = \sigma_2^2$ is given by

$$S_p^2 = \frac{m-1}{m+n-2} \cdot S_1^2 + \frac{n-1}{m+n-2} \cdot S_2^2$$

Case IV

$H_0 : p_1 - p_2 = 0$

Test statistic: $\dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{m}+\dfrac{1}{n}\right)}}$, $\hat{p} = \dfrac{m}{m+n}\hat{p}_1 + \dfrac{n}{m+n}\hat{p}_2$ (the *weighted* average of $\hat{p}_1$

and $\hat{p}_2$)

Case IV cont.

vs Alternative Hypothesis:

$$H_a : p_1 - p_2 > 0 \text{, reject if } \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m}+\frac{1}{n}\right)}} > Z_\alpha$$

$$H_a : p_1 - p_2 < 0 \text{, reject if } \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m}+\frac{1}{n}\right)}} < -Z_\alpha$$

$$H_a : p_1 - p_2 \neq 0 \text{, reject if } \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m}+\frac{1}{n}\right)}} > Z_{\alpha/2} \text{ or } \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m}+\frac{1}{n}\right)}} < -Z_{\alpha/2}$$

Remarks and Questions

- Type II error and sample size calculation on book p.356.

- Sample size calculation formula (balanced design):

$$m = n = \frac{\left[Z_\alpha \sqrt{(p_1 + p_2)(q_1 + q_2)/2} + Z_\beta \sqrt{p_1 q_1 + p_2 q_2} \right]^2}{(p_1 - p_2)^2}$$

- Compare the sample size calculation formula with case I. What is the difference between the two?

Tests on variances

- How to do hypothesis testing when the null is: H_0: $\sigma^2 = \sigma_0^2$.
- ➢ *Chi-square test*.

- In the pooled t-test, we assume the two populations have the same variance. How to test if the two variances are indeed the same?
- ➢ *F test*.

Example

<u>Ex.</u> Suppose you wish to test the effect of Prozac (antidepressants) on the well-being of depressed individuals, using a standardized "well-being scale" that sums Likert-type items to obtain a score that could range from 0 to 20. Higher scores indicate greater well-being (that is, Prozac is having a positive effect). 9 individuals participate in the experiment, the scores before and after using Prozac are recorded. Is the effect of Prozac significant?

	moodpre	moodpost
1	3.00	5.00
2	.00	1.00
3	6.00	5.00
4	7.00	7.00
5	4.00	10.00
6	3.00	9.00
7	2.00	7.00
8	1.00	11.00
9	4.00	8.00

Paired t-test

- As in the previous example, the data is paired, the two scores (before and after) recorded for each individual are dependent, but the between individuals the pairs are independent.

- Thus in order to test H_0: $\mu_1 - \mu_2 = 0$, one has to look at the difference of each pair. The problem eventually becomes a one sample t-test problem.