Lecture 11: Z tests and t tests, continued

Announcements:

- Reading for this week's subject: pp. 347 - 353 in Horvat.

Announcements:

- Reading for this week's subject: pp. 347 - 353 in Horvat.
- Your final grade in this course will be based 75% on exams (midterm and/or final), and 25% on problem sets (and class attendence for regular students, participation, and behavior).

Lecture 11: Z tests and t tests, continued

Announcements:

- Reading for this week's subject: pp. 347 - 353 in Horvat.
- Your final grade in this course will be based 75% on exams (midterm and/or final), and 25% on problem sets (and class attendence for regular students, participation, and behavior).
- EACH WEEKLY PROBLEM SET MUST BE RETURNED THE FOLLOWING TUESDAY IN CLASS! If a problem set is late, you receive no credit (because answers will be posted that day).

Lecture 11: Z tests and t tests, continued

Announcements:

- Reading for this week's subject: pp. 347 - 353 in Horvat.
- Your final grade in this course will be based 75% on exams (midterm and/or final), and 25% on problem sets (and class attendence for regular students, participation, and behavior).
- EACH WEEKLY PROBLEM SET MUST BE RETURNED THE FOLLOWING TUESDAY IN CLASS! If a problem set is late, you receive no credit (because answers will be posted that day).
- Also, Problem sets 4 through 6 are all due IN CLASS ON TUESDAY, NOVEMBER 23.

More announcements

- If you have any questions about the exams, please see me or Tomislav.

More announcements

- If you have any questions about the exams, please see me or Tomislav.
- Next exam will include all questions in this exam that less than 50% of people answered correctly. These are questions 9, 24, 25, 21, 16, and 5. So make sure you understand those questions!

Z test in general

$$Z =$$

Z test in general

$$Z = \frac{\text{observed statistic} - }{}$$

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{}$$

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true}}$$

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of}}$$

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n,

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.
If observed statistic is a binomial variate,

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.

If observed statistic is a binomial variate, its standard deviation is

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.
If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.
If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.
If observed statistic is a Poisson variate,

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.
If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.
If observed statistic is a Poisson variate, its standard deviation is

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.

If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.

If observed statistic is a Poisson variate, its standard deviation is $\sqrt{\lambda}$.

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.
If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.
If observed statistic is a Poisson variate, its standard deviation is $\sqrt{\lambda}$.
If observed statistic is the difference between two sample means,

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.

If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.

If observed statistic is a Poisson variate, its standard deviation is $\sqrt{\lambda}$.

If observed statistic is the difference between two sample means, its standard deviation is

Z test in general

$$Z = \frac{\text{observed statistic} - \text{expected}}{\text{true standard deviation of observed statistic}}$$

If observed statistic is the mean of a sample of n, then its standard deviation is $\sigma/\sqrt{n}$.
If observed statistic is a binomial variate, its standard deviation is $\sqrt{np(1-p)}$.
If observed statistic is a Poisson variate, its standard deviation is $\sqrt{\lambda}$.
If observed statistic is the difference between two sample means, its standard deviation is $\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$.

$= \sqrt{Var(\bar{x}_1) + Var(\bar{x}_2)}$

Example: poll results Z test

$$Z = \frac{\text{obs} - \text{exp}}{\text{sd obs}} = \frac{475 - 510}{\sqrt{npq}}$$

$$\sqrt{npq} = \sqrt{(1000)(0.51)(0.49)}$$

Let's say a poll of 1000 random Croatians found that 47.5% Were in favor of entering the EU. Is this unusual if the true proportion was 51.0% in favor?

$$Z = \frac{-35}{15.8} = -2.21$$

∴ unusual, not in favor

Example: poll results Z test

Let's say a poll of 1000 random Croatians found that 47.5% Were in favor of entering the EU. Is this unusual if the true proportion was 51.0% in favor?

Example: poll results Z test

Let's say a poll of 1000 random Croatians found that 47.5% Were in favor of entering the EU. Is this unusual if the true proportion was 51.0% in favor?

Example: poll results chi square test

χ^2 instead of z

$$\chi^2 = \sum_i \frac{(obs_i - exp_i)^2}{exp_i} = \frac{(475-510)^2}{510} + \frac{(525-490)^2}{490} = 35^2/510 + 35^2/490$$

4.901

Let's say a poll of 1000 random Croatians found that 47.5% Were in favor of entering the EU. Is this unusual if the true proportion was 51.0% in favor?

Observed: 475 favor, 525 against

Expected: 510 favor, 490 against

Example: poll results chi square test

Let's say a poll of 1000 random Croatians found that 47.5% Were in favor of entering the EU. Is this unusual if the true proportion was 51.0% in favor?

Chi square test

Chi square test

$$\chi^2 =$$

Chi square test

$$\chi^2 = \frac{\sum_1^n(\text{count of observed} - \text{count of expected})^2}{\text{count of expected}}$$

Degrees of freedom is the number of counts that are free to vary in this calculation.

Example: poll results chi square test

Let's say a poll of 1000 random Croatians found that 47.5% Were in favor of entering the EU. Is this unusual if the true proportion was 51.0% in favor?

Z test versus t test

$$Z = \frac{\text{obs} - \text{exp}}{\sigma_{\text{obs}}}$$

What if we don't know σ?

Z test versus t test

What if we don't know σ? We calculate it from the sample as the

Z test versus t test

What if we don't know σ? We calculate it from the sample as the sample standard deviation:

Z test versus t test

What if we don't know σ? We calculate it from the sample as the sample standard deviation:

$$s = \sqrt{\frac{\sum_1^n (x_i - \overline{x})^2}{n-1}}$$

Z test versus t test

What if we don't know σ? We calculate it from the sample as the sample standard deviation:

$$s = \sqrt{\frac{\sum_1^n (x_i - \overline{x})^2}{n-1}}$$

$n - 1$ is called v, or the "degrees of freedom"

Z test versus t test

What if we don't know σ? We calculate it from the sample as the sample standard deviation:

$$s = \sqrt{\frac{\sum_1^n (x_i - \overline{x})^2}{n - 1}}$$

$n - 1$ is called v, or the "degrees of freedom" used in the calculation of s.

Z test versus t test

Z test versus t test

$$Z = \frac{\text{observed} - \text{expected}}{\text{true standard deviation of observed}}$$

Z test versus t test

$$Z = \frac{\text{observed} - \text{expected}}{\text{true standard deviation of observed}}$$

$$Z = \frac{\text{observed} - \text{expected}}{\sigma_{\text{obs}}}.$$

Z test versus t test

$$Z = \frac{\text{observed} - \text{expected}}{\text{true standard deviation of observed}}$$

$$Z = \frac{\text{observed} - \text{expected}}{\sigma_{\text{obs}}}.$$

$$t = \frac{\text{observed} - \text{expected}}{\text{estimated standard deviation of observed}}$$

Z test versus t test

$$Z = \frac{\text{observed} - \text{expected}}{\text{true standard deviation of observed}}$$

$$Z = \frac{\text{observed} - \text{expected}}{\sigma_{\text{obs}}}.$$

$$t = \frac{\text{observed} - \text{expected}}{\text{estimated standard deviation of observed}}$$

$$t = \frac{\text{observed} - \text{expected}}{s_{\text{obs}}}.$$

Z test versus t test

Z test versus t test

Z is normally distributed, with mean of 0 and standard deviation of 1.

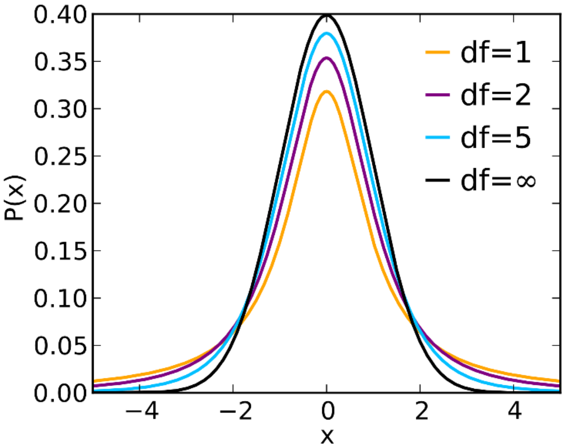
Z test versus t test

Z is normally distributed, with mean of 0 and standard deviation of 1.
t is very close to normally distributed, symmetrical, with mean of 0 and standard deviation slightly higher than 1.

Z test versus t test

Z is normally distributed, with mean of 0 and standard deviation of 1.
t is very close to normally distributed, symmetrical, with mean of 0 and standard deviation slightly higher than 1.
But t is practically indistinguishable from a normal distribution if $n > 30$.

t distribution for various sample sizes (or degrees of freedom)



Example: a sample of heights

$$t = \frac{\text{obs} - \text{exp}}{\text{sd obs}} = \frac{173 - 169}{9.2/\sqrt{15}}$$

Say we have a random sample of 15 women, their sample mean height is $\overline{x} = 173$ cm, and sample standard deviation of height is $s = 9.2$ cm. Is this unusual if the population mean is $\mu = 169$ cm?

$$= \frac{4}{2.38} = 1.683$$

Example: a sample of heights

Say we have a random sample of 15 women, their sample mean height is $\overline{x} = 173$ cm, and sample standard deviation of height is $s = 9.2$ cm. Is this unusual if the population mean is $\mu = 169$ cm?

Example: a sample of heights

Say we have a random sample of 15 women, their sample mean height is $\overline{x} = 173$ cm, and sample standard deviation of height is $s = 9.2$ cm. What is the 95% confidence interval for the population mean? We use qt() to find the critical t.

Example: a sample of heights

Say we have a random sample of 15 women, their sample mean height is $\overline{x} = 173$ cm, and sample standard deviation of height is $s = 9.2$ cm. What is the 95% confidence interval for the population mean? We use qt() to find the critical t.

Example: exam scores

In our Exam 1, the mean score of women was 72.7 and the mean score for men was 70.3. The standard deviation for women was 7.7 and for men was 8.3. There were 14 women and 13 men. Is the sex difference in means unusual, if we expect the difference to be zero?

Example: exam scores

In our Exam 1, the mean score of women was 72.7 and the mean score for men was 70.3. The standard deviation for women was 7.7 and for men was 8.3. There were 14 women and 13 men. Is the sex difference in means unusual, if we expect the difference to be zero?

Example: exam scores

In our Exam 1, the mean score of women was 72.7 and the mean score for men was 70.3. The standard deviation for women was 7.7 and for men was 8.3. There were 14 women and 13 men. Is the sex difference in means unusual, if we expect the difference to be zero?

Degrees of freedom if standard deviations or sample sizes are different in the two groups

Behrens-Fisher problem:
What is the degrees of freedom if $n_1 \neq n_2$ or $s_1 \neq s_2$?

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}.$$