Unit 4: Inference for numerical data
2. Comparing means

Sta 101 - Spring 2015

Duke University, Department of Statistical Science

February 25, 2015

Dr. Çetinkaya-Rundel

Slides posted at http://bitly.com/sta101sp15

Announcements

-

1. When comparing means of two groups, ask if paired or independent

- dependent (paired) groups (e.g. pre/post weights of subjects in a weight loss study, twin studies, etc.)

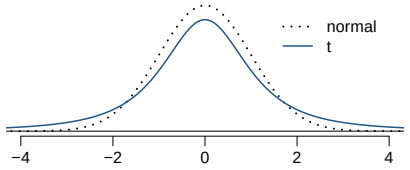
$$SE_{\bar{x}_{diff}} = \frac{s_{diff}}{n_{diff}}$$

- independent groups (e.g. grades of students across two sections)

$$SE_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

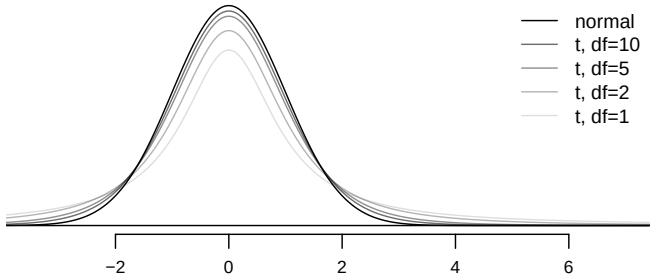
2. T corrects for uncertainty introduced by plugging in s for σ

- Essential when n is small ($n < 30$) since s is more likely to be not be a good estimate for σ when n is small than when n is large
- Could be used when n is large as well
- Also has a bell shape, but its tails are *thicker* than the normal model's
  - Observations are more likely to fall beyond two SDs from the mean than under the normal distribution.
- Extra thick tails are helpful for mitigating the effect of a less reliable estimate for the standard error of the sampling distribution (since n is small)



T distribution

- Always centered at zero, like the standard normal (z) distribution
- Has a single parameter: *degrees of freedom* (*df*)
  - one sample: $df = n - 1$
  - two (independent) samples: $df = min(n_1 - 1, n_2 - 1)$



What happens to shape of the T distribution as *df* increases?

Clicker question

Under the T distribution with 5 degrees freedom, _____ of the data are within one standard deviation of the mean?

(a) 68%
(b) less than 68%
(c) more than 68%

Application exercise: 4.2 Comparing means, Pt 1

See the course webpage for details.

Application exercise: 4.3 Comparing means, Pt 2

See the course webpage for details.

1. When comparing means of two groups, ask if paired or independent
2. T corrects for uncertainty introduced by plugging in s for σ