Confidence Intervals

18.05 Spring 2014

You should have downloaded studio11.zip and unzipped it into your 18.05 working directory.

Confidence Intervals Applet

Open the applet:
http://mathlets.org/mathlets/confidence-intervals/

1. Play around with the applet. Make sure you understand how it measures if a confidence interval is correct.

2. Read the help page.

3. What is random each time you click the 'Run N trials' button?

4. Fix the parameter settings and run many trials.

(a) Does the confidence interval contain the true mean the correct percentage of the time?

(b) What can you say about the size of the z and t-intevals over repeated trials?

5. How does increasing c change the confidence intervals? Why?

6. How does increasing n, μ or σ change the intervals? Why?

Review: $\chi^2(df)$ confidence intervals for σ^2

- Range: $[0, \infty]$
- Parameter: df = degrees of freedom

Data: $x_1, \ldots, x_n \sim \mathrm{N}(\mu, \sigma^2)$, where μ and σ are unknown.

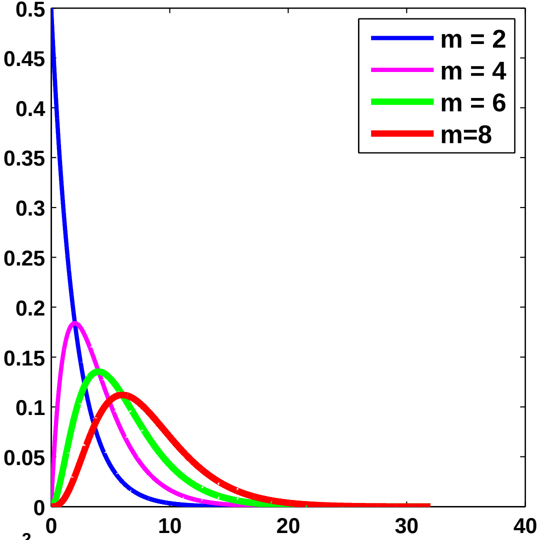
Test statistic: $r = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

$1 - \alpha$ confidence interval for σ^2:

$$\left[\frac{r}{c_{\alpha/2}}, \frac{r}{c_{1-\alpha/2}} \right],$$

$c_{\alpha/2}$ is the right-tail critical value.

χ^2



χ^2_m probability densities for various degrees of freedom

R Problem 1: Confidence intervals for σ^2

Write R code that:

(a) Simulates sampling 17 samples from a $N(2, 3^2)$ distribution.

(b) Computes the 90% confidence interval for σ^2 from the sample.

Stock market volatility

Data file for studio: `studio11SP500data.csv`

• Contains the daily percentage change in the *Standard and Poors 500* stock index over the 14 years.

Volatility:

- Let σ^2 be the variance of the daily percentage change.
- By definition **volatility** $= \sigma$.
- High volatility implies large, fast changes in the value of the index.

Question: Is the volatility of the stock market independent of the day of the week, or are there certain weekdays when volatility tends to be higher?

R Problem 2: Stock market volatility

1. Use the code in `studio11.r` to load the percentage change data for Mondays and Fridays.

(This code also does a little data exploration using plots and a table.)

2. Let σ_M^2 be the true variance of the percent returns on Mondays. Likewise σ_F^2 for Fridays.

3. Use `?var.test` to learn about the function `var.test()`

4. Use `var.test()` to compute a 95% confidence interval for the ratio of the variances. Use the result to decide if one of Mondays or Fridays is more volatile than the other.

Understanding `var.test()`

Notation: $F(\text{df1}, \text{df2}) = F$ distribution with (df1, df2) degrees of freedom.

Theorem. If $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ are independent samples from normal distributions with the **same variance** then the ratio of sample variances follows an F distribution:

$$F = \frac{\text{var}(x_i)}{\text{var}(y_j)} \sim F(n-1, m-1).$$

• Now assume that the normal distributions have **different variances**, σ_x^2, σ_y^2.

Problem: (a) Use the F statistic, critical values of the F distribution and the theorem to determine the $1-\alpha$ confidence interval for the ratio of variances σ_x^2/σ_y^2.

(b) Code your answer in R and show you get the same results as we did using `var.test(x, y)`.