

# 431 Class 18

Thomas E. Love

2017-10-31

## Gelman on Statistical Significance

"... we use the term statistically significant in the conventional way, to mean that an estimate is **at least two standard errors away** from some "null hypothesis" or prespecified value that would indicate no effect present. An estimate is statistically insignificant if the observed value could reasonably be explained by simple chance variation, much in the way that a sequence of 20 coin tosses might happen to come up 8 heads and 12 tails; we would say that this result is not statistically significantly different from chance. More precisely, the observed proportion of heads is 40 percent but with a standard error of 11 percent - thus, the data are less than two standard errors away from the null hypothesis of 50 percent, and the outcome could clearly have occurred by chance. Standard error is a measure of the variation in an estimate and gets smaller as a sample size gets larger, converging on zero as the sample increases in size."

Gelman's blog (2017-10-28)

## Today's Agenda

- Discussion of the Class 17 In-Class Survey
- Comparing More than Two Populations: The Analysis of Variance

# Project Task C

See README for Class 18, and README for Project Task C, please.

The Google Form for the survey is at  
<https://goo.gl/forms/bB1xJ16NnLihP9Gu1>

Everyone must fill out the survey, regardless of whether you are working in a group.

It's due at noon on Wednesday 2017-11-08, as is the Task C Word template.

# In-Class Survey from Class 17

## In-Class Survey (class17a data)

We chose (using a computer) a random number between 0 and 100.  
Your number is  $X = 10$  (or 65).

- 1 Do you think the percentage of countries which are in Africa, among all those in the United Nations, is higher or lower than  $X$ ?
- 2 Give your best estimate of the percentage of countries which are in Africa, among all those in the United Nations.

## The facts

- There are 193 sovereign states that are members of the UN.
- The African regional group has 54 member states, so that's 28%.
- UN regions for countries are this [Wikipedia link](#)
- The class17a data set contains the answers to these questions from 185 students asked the same questions in the same way over the past four years (since 2014).

# In-Class Survey from Class 17

## In-Class Survey (class17a data)

We chose (using a computer) a random number between 0 and 100.  
Your number is  $X = 10$  (or 65).

- 1 Do you think the percentage of countries which are in Africa, among all those in the United Nations, is higher or lower than  $X$ ?
- 2 Give your best estimate of the percentage of countries which are in Africa, among all those in the United Nations.

## The facts

- There are 193 sovereign states that are members of the UN.
- The African regional group has 54 member states, so that's 28%.
- UN regions for countries are this [Wikipedia link](#)
- The class17a data set contains the answers to these questions from 185 students asked the same questions in the same way over the past four years (since 2014).

# In-Class Survey from Class 17

## In-Class Survey (class17a data)

We chose (using a computer) a random number between 0 and 100.  
Your number is  $X = 10$  (or 65).

- 1 Do you think the percentage of countries which are in Africa, among all those in the United Nations, is higher or lower than  $X$ ?
- 2 Give your best estimate of the percentage of countries which are in Africa, among all those in the United Nations.

## The facts

- There are 193 sovereign states that are members of the UN.
- The African regional group has 54 member states, so that's 28%.
- UN regions for countries are this [Wikipedia link](#)
- The class17a data set contains the answers to these questions from 185 students asked the same questions in the same way over the past four years (since 2014).

# In-Class Survey from Class 17

## In-Class Survey (class17a data)

We chose (using a computer) a random number between 0 and 100.  
Your number is  $X = 10$  (or 65).

- 1 Do you think the percentage of countries which are in Africa, among all those in the United Nations, is higher or lower than  $X$ ?
- 2 Give your best estimate of the percentage of countries which are in Africa, among all those in the United Nations.

## The facts

- There are 193 sovereign states that are members of the UN.
- The African regional group has 54 member states, so that's 28%.
- UN regions for countries are this [Wikipedia link](#)
- The class17a data set contains the answers to these questions from 185 students asked the same questions in the same way over the past four years (since 2014).

# In-Class Survey from Class 17

## In-Class Survey (class17a data)

We chose (using a computer) a random number between 0 and 100.  
Your number is  $X = 10$  (or 65).

- 1 Do you think the percentage of countries which are in Africa, among all those in the United Nations, is higher or lower than  $X$ ?
- 2 Give your best estimate of the percentage of countries which are in Africa, among all those in the United Nations.

## The facts

- There are 193 sovereign states that are members of the UN.
- The African regional group has 54 member states, so that's 28%.
- UN regions for countries are this [Wikipedia link](#)
- The class17a data set contains the answers to these questions from 185 students asked the same questions in the same way over the past four years (since 2014).



# Analysis of Variance (Section 28, Course Notes)

## Analysis of Variance to Compare More Than Two Population Means using Independent Samples

Suppose we want to compare more than two population means, and we have collected three or more independent samples.

This is analysis of a continuous outcome variable on the basis of a single categorical factor — in fact, it's often called **one-factor** ANOVA or **one-way** ANOVA to indicate that the outcome is being split up into the groups defined by a single factor.

- $H_0$ : population means in each group are the same
- $H_A$ :  $H_0$  isn't true; at least one  $\mu$  differs from the others

When there are just two groups, then this boils down to an F test that is equivalent to the Pooled t test.

## One-Way ANOVA

If we have a grouping factor with  $k$  levels, then we are testing:

•  $H_0: \mu_1 = \mu_2 = \dots = \mu_k$  VS

# On $p$ values and statistical significance

## The Value of a $p$ -Valueless Paper

Jason T. Connor (2004) *American J of Gastroenterology* 99(9): 1638-40.  
Abstract: As is common in current biomedical research, about 85% of original contributions in *The American Journal of Gastroenterology* in 2004 have reported  $p$ -values. However, none are reported in this issue's article by Abraham et al. who, instead, rely exclusively on effect size estimates and associated confidence intervals to summarize their findings. **Authors using confidence intervals communicate much more information in a clear and efficient manner than those using  $p$ -values. This strategy also prevents readers from drawing erroneous conclusions caused by common misunderstandings about  $p$ -values.** I outline how standard, two-sided confidence intervals can be used to measure whether two treatments differ or test whether they are clinically equivalent.  
DOI: 10.1111/j.1572-0241.2004.40592.x

## Editorial from JAMA Cardiology 2016-10-12