Exploration librairie swirl : Introduction to Statistical Inference

> swirl()

Welcome to swirl! Please sign in. If you've been here before, use the same name as you did then. If you are new, call yourself something unique.

What shall I call you? ***jlbellier***

Please choose a course, or type 0 to exit swirl.

1: Statistical Inference
2: Take me to the swirl course repository!

Selection: ***1***

Please choose a lesson, or type 0 to return to course menu.

1: Introduction	*2: Probability1*	*3: Probability2*
4: ConditionalProbability	*5: Expectations*	*6: Variance*
7: CommonDistros	*8: Asymptotics*	*9: T Confidence Intervals*
10: Hypothesis Testing	*11: P Values*	*12: Power*
13: Multiple Testing	*14: Resampling*	

Selection: ***1***

0%

Introduction to Statistical_Inference. (Slides for this and other Data Science courses may be found at github https://github.com/DataScienceSpecialization/courses. If you care to use them, they must be downloaded as a zip file and viewed locally. This lesson corresponds to Statistical_Inference/Introduction.)

======== ***10%***

In this lesson, we'll briefly introduce basics of statistical inference, the process of drawing conclusions "about a population using noisy statistical data where uncertainty must be accounted for". In other words, statistical inference lets scientists formulate conclusions from data and quantify the uncertainty arising from using incomplete data.

========== ***20%***

Which of the following is NOT an example of statistical inference?

1: Testing the efficacy of a new drug
2: Constructing a medical image from fMRI data
3: Recording the results of a statistics exam
4: Polling before an election to predict its outcome

*Selection: **3***

All that hard work is paying off!

========== ***30%***

So statistical inference involves formulating conclusions using data AND quantifying the uncertainty associated with those conclusions. The uncertainty could arise from incomplete or bad data.

========== ***40%***

Which of the following would NOT be a source of bad data?

1: Small sample size
2: Selection bias
3: A randomly selected sample of population
4: A poorly designed study

*Selection: **3***

Your dedication is inspiring!

========== ***50%***

So with statistical inference we use data to draw general conclusions about a population. Which of the following would a scientist using statistical inference techniques consider a problem?

1: Our data sample is representative of the population
2: Our study has no bias and is well-designed
3: Contaminated data

*Selection: **3***

Your dedication is inspiring!

========== ***60%***

Which of the following is NOT an example of statistical inference in action?

1: Determining a causative mechanism underlying a disease
2: Estimating the proportion of people who will vote for a candidate
3: Testing the effectiveness of a medical treatment
4: Counting sheep

*Selection: **4***

That's the answer I was looking for.

========== ***70%***

We want to emphasize a couple of important points here. First, a statistic (singular) is a number computed from a sample of data. We use statistics to infer information about a population. Second, a random variable is an outcome from an experiment. Deterministic processes, such as computing means or variances, applied to random variables, produce

additional random variables which have their own distributions. It's important to keep straight which distributions you're talking about.

========== 80%

Finally, there are two broad flavors of inference. The first is frequency, which uses "long run proportion of times an event occurs in independent, identically distributed repetitions." The second is Bayesian in which the probability estimate for a hypothesis is updated as additional evidence is acquired. Both flavors require an understanding of probability so that's what the next lessons will cover.

========== 90%

Congrats! You've concluded this brief introduction to statistical inference.

==========100%