# Transcript: EP 137 - C964 and D195 - Introduction to Inferential Statistics - Statistical Significance with Jim Ashe

*The following transcript is a verbatim account of the video or audio file accompanying this transcript.*

Speaker: Jim Ashe

WGU's IT Audio Series; flexible, portable, profound.

Hi. This is Dr. Jim Ashe, with the computer science team. When people hear statistics, they most often think of charts, graphs, and measurements, histograms, scatter plots, means, medians, etc. Things that describe data. We call these descriptive statistics. In this podcast, I'm going to discuss using data to draw conclusions inferring beyond the known data. This called inferential statistics, and for your capstone, it's a type of non descriptive method. It's anything that uses a sample to draw inferences. It provides a method for attaching meaning to observations. Every college graduate should be familiar with this process, especially students of a hard science. Like you. Say you collect some data, you randomly grab a sample of, say, 30 US adults and find they have an average height of 66.52 inches and you know that US 14-year-olds have an average height of 64.98 inches. That's descriptive, but what can we conclude? Well, maybe US adults are taller on average than 14-year-olds. Sounds like a reasonable conclusion, but maybe you just happen to grab a sample of adults that were relatively tall. What we mean is that the adults are probably taller than 14-year-olds. When the probability is something we can mathematically measure, we can take these collisions beyond intuitive arguments and give them scientific rigor. Rigorous inductive arguments using probability, leverage the so-called rare event rule. An example, a simple fair coin flipping game. Heads I win, tails you win. Let's start the game. Round 1, heads, I win. Round 2, heads, I win again. Round 3, heads, round 4, heads, round 5, heads, round 6, heads, round 7, heads, round 8, heads. I'm really good at this game. At what point did you say. "This seems a little fishy." Maybe around round 5, but certainly by round 8. Why? Well the odds, the probability of getting heads eight times in a row is very small, around 0.039 percent, rounded off, if the game is fair. I calculated those odds assuming that the odds of getting heads was 50 percent. At some point, you decided to reject that assumption because of what we call the rare event rule. If under an assumption, an observed event occurs with a very small probability, then that assumption is probably false. What's a very small probability? That's the statistical significance, usually denoted with the Greek letter six Alpha, and this value is up to you. If you got suspicious at round 5, the probability of getting five heads in a row is 3.125 percent. The very small was less than five percent. You rejected that the coin was fair using the statistical significance of five percent. If you got suspicious at round 7, the probability of seven heads in a row is 0.78125 percent. You rejected that the coin was fair using an Alpha, a statistical significance of one percent. I could say the coin is unfair after round 1 with an Alpha of 55 percent, but that's a pretty weak statement. The smaller the Alpha, the more convincing the argument, but the hard it is to get a result. This coin example is relatively simple because the probability was easy to compute. However, no matter the complexity of the probability space or inferential method, how you reach your conclusion is just as simple. You compare the probability of event to an Alpha and reject or fail to reject the assumption. Finding the probability can often be technically difficult. You'll find lots of online resources discussing z-scores, t-scores, probability distributions, etc., and of course, formulas, lots and lots of formulas, this stuff. Most importantly, the probability is found using software. Python, JASP, R, even Excel. You don't have to understand everything about a combustion engine to drive a car. You need to know which paddle makes it go, which pedal makes it stop, and how to turn. You also need to know where we're going. For us, that's comparing a probability to an Alpha and knowing which conclusion to make. Let's go back to where we started. After selecting a random sample of 30 adults, we find that the average US height of adults is 66.52 inches, and we know that 14-year-olds have an average height of 64.98 inches. It sounds like adults are taller, but maybe we just happen to grab a sample of tall adults. Maybe our results are due to chance. Though unlikely it is possible, we just happen to select 30 tall people. How unlikely that is, is what we need to know. Let's use what we've learned. Applying the rare event rule is a proof by contradiction, so we start by assuming the opposite of our claim. Step 1, assume that the average adult height is not greater than or equal to 14-year-olds. This is the null hypothesis, says there is no difference, the opposite of what we're trying to do here. Step 2, pick a significance. Let's say five percent. Using one in five percent is pretty common. That's our Alpha. Step 3, find the probability. Imagine a bucket of all the samples of 30. We're randomly reaching and grab one. As you might imagine, most samples will have average which is close to the true average. Fewer and fewer have averages larger or smaller, creating a shape that looks more and more like a bell. By the central limit theorem, as samples get large, this shape eventually matches a normal distribution that that perfect bell shaped. Using a distribution, we can calculate our profitability. Let's calculate. That's my computer calculating stuff. It gives me a probability of 0.038 or 3.8 percent. This is the probability that that observed difference, that sample in average height is due simply a chance, assuming that the groups are no different. Step 3, compare, 3.8 percent is less than five percent or Alpha, so our results are significant at five percent, so we reject the null hypothesis and we accept the only alternative called appropriately enough, the alternative hypothesis, there is a difference. Formerly we say, there is sufficient evidence to support the claim that US average adult height is greater than the average height of 14-year-olds. Note, if we had used the significance of one percent and Alpha one percent because 3.4 percent is greater than one percent, our results would instead been inconclusive. There is not sufficient evidence to support the claim that the average US adult height is greater than the average height of 14-year-olds. But then grabbing just a slightly larger sample of 40, I get a similar sample average, but the probability of grabbing that average is 0.3 percent, which is significant in an Alpha one percent, and that's it, that we use a simple one-sample t-test for this example. More complex methods are generally only more complex in how the probability is found, but how conclusions are reached by comparing that probability to a statistical significance is the same. Nowhere does WGU directly require our computer science students to use inferential statistics. However, it might be used in your capstone, C9 64, where machine learning can be used to create regression models, the accuracy of which needs to be measured. You can do that by using the ratio correct predictions, maybe within a margin of error to observations, or you could use a correlation coefficient, a statistical measurement of the relationship and show that that relation is statistically significant, or you could do both. That's it for this podcast. Thank you for listening. If you have any questions, suggestions, or recommendations for anything stat, math, computer science, or data analytics related, please contact me at Jim.Ashe@WGU.edu. That's JIM.ASHE@WGU.EDU.

Schedule time with your instructor to explore more deeply. WGU, a new kind of you.