Hi Class!

I had a great question from a student and thought I would share the response.  Robert our TA rang in to .. .thank you Robert!  Please read below and think about it as much as you can before live session.  You can get the idea by just thinking about it ... you don't need to write anything down.

  Robert made a Exam/Coffee example ... let me introduce it.  Let's say we have n = 4 people in the study.  Students were randomly assigned to either drink coffee and then take and exam or not drink coffee and then take the same exam.  Let's say the two coffee drinkers got As and the two non coffee drinkers got Bs.  A researcher then comes in and says ..."YES! see ... we can conclude with high confidence that drinking coffee causes higher test grades!"

Think about that for a second.

Contrast this idea now with a much larger study with n = 1000 subjects.  Let's say that half drank coffee and half did not and the 500 coffee drinkers got As and the 500 non coffee drinkers got Bs.  Is this stronger evidence of a coffee effect?

I say YES!

The first study's results could have easily happened by chance.

If coffee had no effect, we could easily have just accidentally / randomly assigned the smartest two students to the coffee group.  To this end... maybe of the 4 people the two with As also both wore Nike tennis shoes and the others did not.  Are we to believe that Nike tennis shoes are causing higher scores and their absence causes lower scores?  Not likely.

**We observed this:**

Student 1 Coffee A

Student 2 Coffee A

Student 3 No Coffee B

Student 4 No Coffee B

But if coffee had no effect ... the "smartest" students had to be assigned somewhere... the other possibilities are:

Student 1 No Coffee A

Student 2 No Coffee A

Student 3 Coffee B

Student 4 Coffee B

Student 1 No Coffee A

Student 2 Coffee A

Student 3 Coffee B

Student 4 No Coffee B

Student 1 No Coffee A

Student 2 Coffee A

Student 3 No Coffee B

Student 4 Coffee B

Student 1 Coffee A

Student 2 NoCoffee A

Student 3 No Coffee B

Student 4 Coffee B

and

Student 1 Coffee A

Student 2 NoCoffee A

Student 3 Coffee B

Student 4 NoCoffee B

That's it ... if coffee had no effect then each of these assignments and outcomes have the same chance of being the experiment that was conducted.  In combinations this is known as 4 C 2 = 6.  The outcome we actually observed had a 1 in 6 chance of happening IF coffee has no effect because if coffee had no effect Student 1 and 2 would get As and Student 3 and 4 would get Bs regardless of if they drank coffee or not.

With the larger study, it is very unlikely that the most extreme result would happen by chance.  There are 1000 C 2 = 499500 different experiments that could have resulted IF coffee had no effect.  So if we conclude that coffee had no effect then we are saying that we know that the result as extreme as the one we saw would happen only 1 in 499,500 times and that we are ok with the fact that this very very unlikely event actually occurred (much as before as when someone would have to believe that if I won the lottery 10 times in a row that I am playing fair.)

The conclusion that makes more sense in the larger study is that lightening did not strike twice (a very rare event did not happen) and what we are witnessing is the fact that coffee actually helps test scores.

The rank sum test quantifies this.  Which is where Robert's example comes in:

Instead of As and Bs let's say that we have a better measure of their performance on the test: a grade between 0 and 100.

Student        Score         Treatment        Rank

Jerry                 99              Coffee              1

Silvia                 96             Coffee             2

Mark                  88            Coffee              3

Emma                83            Coffee              4

John                76              No Coffee         5

Amanda           71             No Coffee         6

Linda                67             No Coffee         7

Tim                   63              No Coffee        8

We happened to witness that the Coffee Students outperformed the non coffee students in every case.  We can either say that coffee had no effect and that we just happened to randomly assign the "smartest" students to drink coffee by chance ... but what would the probably of that be?  It is simply a counting exercise.  First we need to figure out how many different combination of Students and coffee groups there are.  8 C 4.  What we have observed is the most extreme combination in favor of coffee so the bigger 8 C 4 is the more evidence that coffee actually had an effect.  I will leave it here and we will discuss this very problem tomorrow.

Please give this some deep thought before live session.