Numerical Analysis — Problem Set 9 — Statistical Significance — Paired t-Squared Test

Due Tuesday, Nov. 15 (beginning of class)

In our discussion on Friday, Nov. 11 we all understand, that just because an event is unlikely does not mean that it is impossible. For example, Clara could visit a casino, play a perfectly fair game where you win half the time, but nonetheless lose the first three times she plays it. It does not mean the game is rigged.

Or a heart medication might be studied, and 5 people in the medicated group might get heart attacks, whereas 10 people in the unmedicated group might get heart attacks, but this does not mean that the medication is effective. It might just be good luck for the medicated patients, and bad luck for the unmedicated ones, and have nothing to do with the effectiveness of the medication.

All statistical studies suffer from this kind of uncertainty. The result could be due to chance, not a real effect. Often studies that showed promising results are repeated, and the good luck disappears. The new study fails to find an effect. There is even a parody journal, called The Journal of Irreproducible Results, *whose title is about this problem. A great tool is being able to critically look at scientific claims, and decide for yourself whether they are significant.*

1. Quantifying Unlikeliness

We looked closely at 20 coin tosses. The formula for the fraction of the 20 coin tosses that will have k heads was

$$\frac{1}{2^{20}} \frac{20!}{k!\,(20-k)!}$$

We graphed this for $k = 0,\ 1,\ 2,\ \ldots,\ 20$, and saw that it peaked around 18% at $k = 10$.

If the coin is weighted, so that the probability of getting heads on each toss is w, then the probability is instead:

$$w^k (1 - w)^{20-k} \frac{20!}{k!\,(20-k)!}$$

Make sure you see that if $w = \frac{1}{2}$ then you get the original formula.

If instead of 20, we use the variable n for the number of tosses, then the formula is

$$w^k(1 - w)^{n-k} \frac{n!}{k!\,(n-k)!}$$

(a) Write a little program that expects w in a register, expects n in another register, and the user just types k and hits R/S, and boom, the answer comes out.

(b) Use your program to make a table and a graph of this function for

$n = 9, \; w = 1/3$

To make the graph, you are going to have to run your program for the 10 different values of $k = 0, \; 1, \; 2, \; \ldots, \; 9$.

(c) Do the same thing again for

$n = 21, \; w = 1/3$

You are going to have to run your program for 22 different values of $k = 0, \; 1, \; 2, \; \ldots, \; 21$.

(d) You expect to get an average of 3 heads out of 9 tosses with this weighted coin. By adding up the percentages in your table for part (c) for 6, 7, 8, and 9, what are the odds of getting 6 or more heads?

(e) You expect to get an average of 7 heads of 21 tosses. By adding up the percentages in your table for part (d) for 14, 15, 16, 17, 18, 19, 20, and 21, what are the odds of getting 14 or more heads?

Discussion: Getting twice as many heads as expected is a lot less likely if there are 21 rolls than 7 rolls, I hope you agree after doing (d) and (e). If you got twice as many heads as expected after 21 rolls, you'd be really suspicious that $w = \frac{2}{3}$ *not* $w = \frac{1}{3}$.

2. The Paired t – Test

In[2]:=
```
TableForm[
 {{"", "Medicated (Left) Arm", "Unmedicated (Right) Arm"}, {"Bri", 18, 25},
  {"Bra", 22, 21}, {"Che", 5, 7}, {"Cla", 5, 17}, {"Max", 0, 20}}]
```

Out[2]//TableForm=

	Medicated (Left) Arm	Unmedicated (Right) Arm
Bri	18	25
Bra	22	21
Che	5	7
Cla	5	17
Max	0	20

The average number of square centimeters of poison oak rash on the medicated arms, was 10. For the unmedicated arms, it is 18. The paired *t* – test tells you whether this was might be due to chance. Of course, just like with the casino, there is always some possibility the study result was just good luck (for the medication manufacturer's sales department), and that actually the medication has no effect. The paired *t*-test quantifies how unlikely it is that the effect is due to chance.

(a) What is the mean of the left column? What is the mean of the right column? What is the difference (right - left) of the means? Make another column above that is just the difference (right - left).

(b) Look at the JMP documentation I passed out and the *HP-25 Applications Programs* book. Make a table of the differences (between the right arm and the left arm).

(c) Using the definitions for std deviation and std error of the mean, compute both of those.

(d) Apply the paired *t* – test program the data. What does the program have to say about the significance?

3. The Paired *t* – Test Again

I'm going to re-do the fake data, and make it jump around less. The difference in the means will still be 8. There are still 5 subjects.

In[6]:=
```
TableForm[
 {{"", "Medicated (Left) Arm", "Unmedicated (Right) Arm"}, {"Bri", 18, 25},
  {"Bra", 21, 27}, {"Che", 3, 12}, {"Cla", 7, 17}, {"Max", 12, 20}}]
```

Out[6]//TableForm=

	Medicated (Left) Arm	Unmedicated (Right) Arm
Bri	18	25
Bra	21	27
Che	3	12
Cla	7	17
Max	12	20

Do all of (a)-(d) for this less jumpy data. If things go well, even though the difference in the means is still 8, it is much less likely that that difference occurred by chance.

Discussion: It is a bit of a stretch, because we haven't done the theory of the t-test. We have just claimed it is a good test. I haven't been through the theory myself. But I can see that what the t-test is doing is very much like what we did in 1(d) and 1(e).

More Thought: Try to compare 2 and 3 and see why making the data less jumpy makes the difference of 8 more significant.