Applicable Statistics for the Clinical Research Professional

San Francisco Bay Area SOCRA Chapter Meeting
July 8th

Shashi Sathyanarayana, Ph.D. CSSBB

Numeric Insight, Inc
www.numericinsight.com

practical and exceptional number crunching, and scientific programming

Let us hit the ground running

- The time taken for cessation of bleeding was recorded for a large number of persons whose fingers had been pricked. The mean time was found to be 1.407 minutes and the standard deviation was 0.588 min.
- In order to stably measure bleeding time, a standard pressure must be applied to the upper arm
- In an effort to determine whether pressure applied to the upper arm increases bleeding time, six persons had pressure equal to 20 mmHg applied to their upper arms and had their fingers pricked.
- For these six persons, the time taken (in minutes) for bleeding to stop were

  1.15

  1.75

  1.32

  1.28

  1.39

  2.50

Clinical question: Does pressure increase bleeding time?

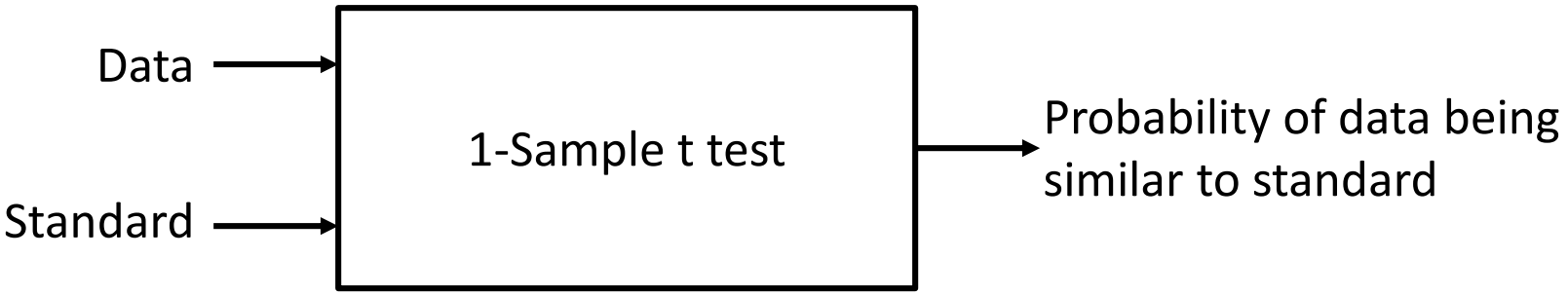
Basic Statistics: A Primer for the Biomedical Sciences (Wiley Series in Probability and Statistics) by Olive Jean Dunn and Virginia A. Clark (Jul 27, 2009). Fourth Edition pp.91

The doctor ordered this test …

- … called the “1-Sample t test for equality of means”
- It is used to answer the question

  “Has something changed relative to a well established standard?”
- Or a little more formally

  “How likely is it that my experimental data is similar to the well established standard data?”
- We need to crank the numerical engine shown below



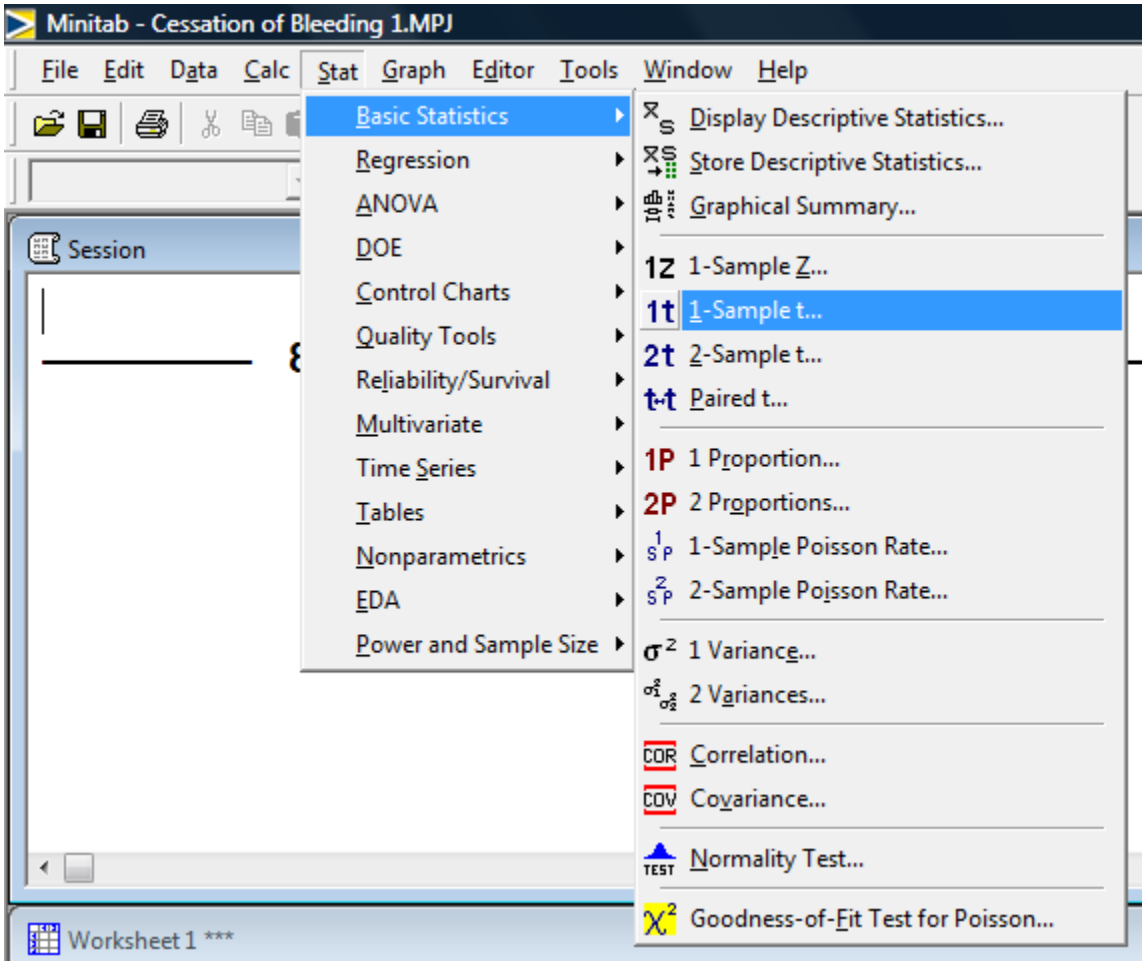
Relax, this test doesn't hurt

Step 1: Simply enter data into Minitab

Step 2: Choose "1-Sample t" from the menu

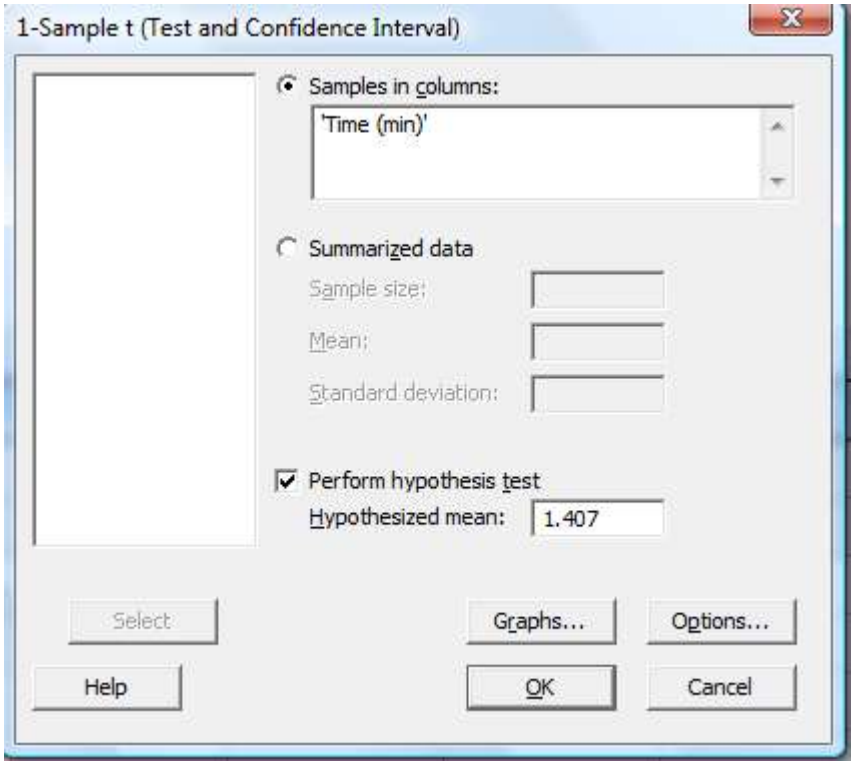
Worksheet 1 ***

↓	C1	C2	C3
	Time (min)		
1	1.15		
2	1.75		
3	1.32		
4	1.28		
5	1.39		
6	2.50		
7			
8			
9			



Answer a few easy questions

Step 3: Specify data column & standard



Step 4: See result

One-Sample T: Time (min)

Test of mu = 1.407 vs not = 1.407

Variable	N	Mean	StDev	SE Mean
Time (min)	6	1.565	0.501	0.204

95% CI	T	P
(1.040, 2.090)	0.77	0.474

Recall the question *"How likely is it that my experimental data is similar to the well established standard data?"*

Answer: A 47% chance. Which is considered "quite likely." There is no strong evidence that pressure increases bleeding time.

80% of all stat terms you need to know

Test of mu = 1.407 vs not = 1.407

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
Time (min)	6	1.565	0.501	0.204	(1.040, 2.090)	0.77	0.474

- Hypothesis test: Where you start with
  - the pessimistic statement that your new method is no better than the old
  - gather data
  - crunch numbers and
  - find the probability of getting the observed result, assuming original statement is correct
- Mean: We all know what this means
- Standard deviation: Think about a 1 lb. box containing 500 jellybeans. Think about another 1 lb. box containing raisins.
- Standard error of mean: If 100 researchers conducted their own bleeding time test, how spread out would their estimates of the mean be?
- Confidence interval: The true mean would be contained within 95 of those researcher's individual confidence intervals.
- p value: The probability of the confidence interval containing the true mean

You're fired!

```
Test of mu = 1.407 vs not = 1.407


Variable     N    Mean   StDev   SE Mean        95% CI           T       P
Time (min)   6   1.565   0.501     0.204   (1.040, 2.090)    0.77   0.474
```

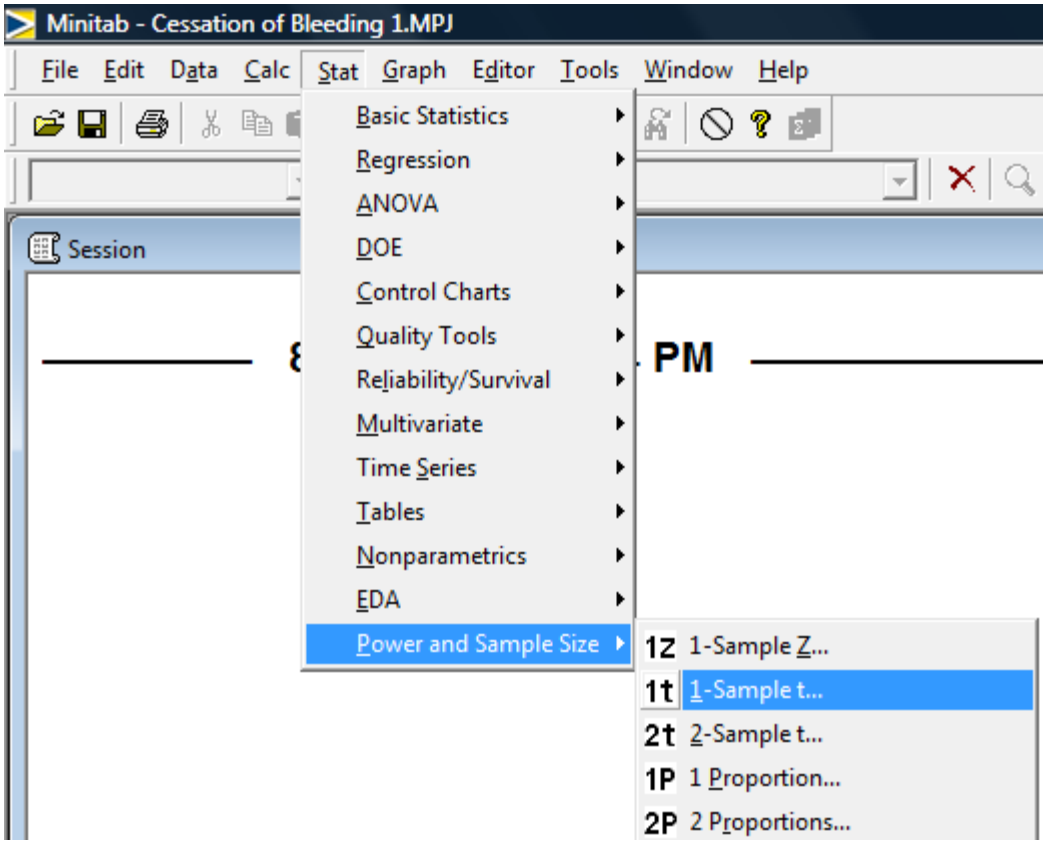
- The researcher is trying to check for a general increase in bleeding time on applying pressure
- However, if multiple researchers attempt to duplicate this study, the experimentally estimated mean bleeding time can vary by at least ±0.204, possibly more (Standard Error)
- This specific 6 sample test has NO POWER
- Even if an effect exists, it can not detect it.

Power and Sample Size

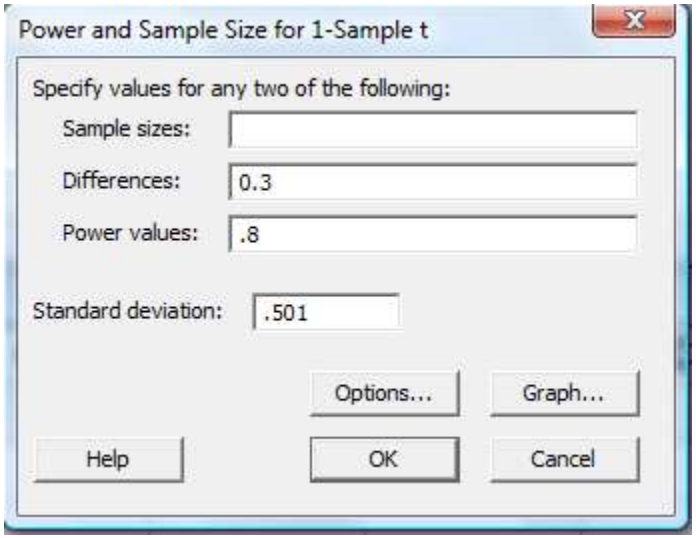
- The researcher wants to be able to detect a 0.3 min difference in bleeding time (effect size, based on clinical knowledge)
- The researcher wants the test to be reliable
- How reliable do they want it to be? The researcher wants to ensure that at least 80% of the time, the t-test reports a positive effect (low p-value) if in fact pressure does change bleeding time
- The researcher wants to make the test reliable (and credible) by testing more subjects
- It is EASY to figure out how many subjects to test.

Power and Sample Size (contd)

Step 1: From the Stat menu, choose “Power and Sample Size” and “1-Sample t”



Step 2: Fill in details



Power and Sample Size (contd)

Power and Sample Size

1-Sample t Test

Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05 Assumed standard deviation = 0.501

Difference	Sample Size	Target Power	Actual Power
0.3	24	0.8	0.802115

- The researcher realized that the experiment had to be repeated with 24 subjects to be 80% certain of catching a 0.3 min difference in bleeding time
- Using clinical input from colleagues, the researcher decided to set the pressure at 40mmHg next time
- Getting really serious, the researcher doubted the validity of the “well established standard” of 1.407 min of bleeding time without pressure, and started thinking
- “Time to do a clinical trial”

“Time to do a clinical trial”

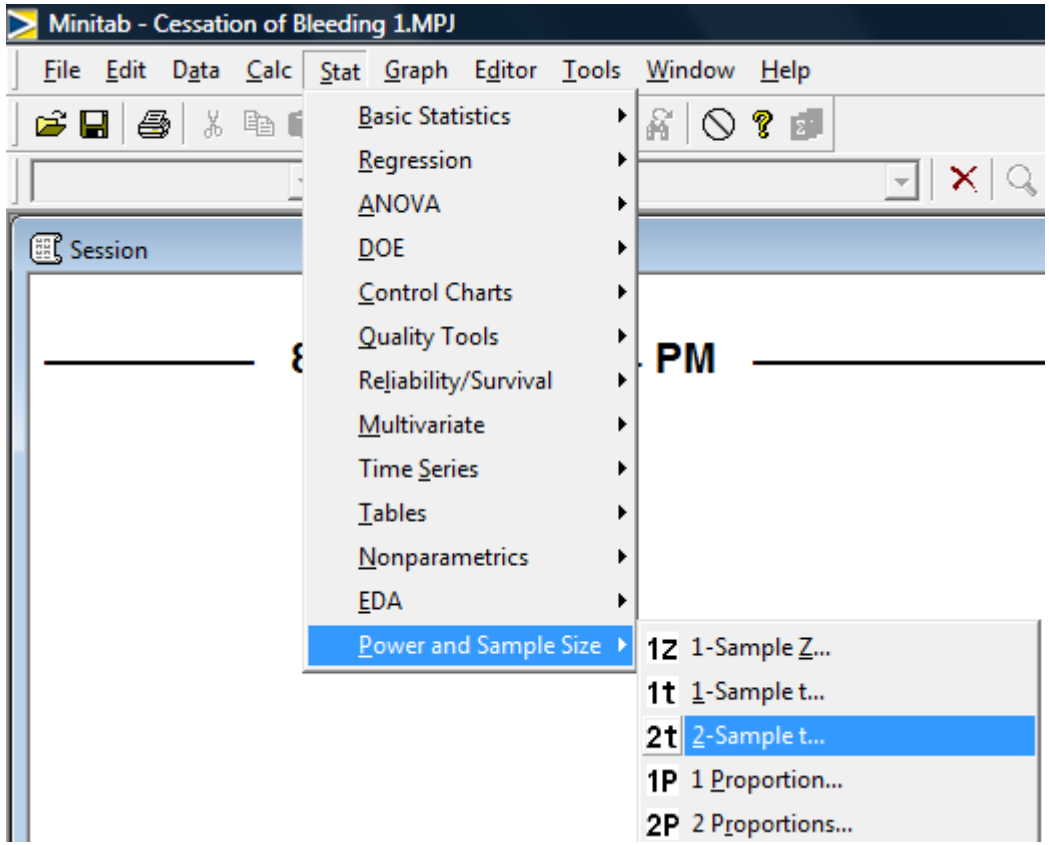
What is my position so far [thinks the researcher]:

- A difference in bleeding time of 0.3 min is worth knowing about (if I can catch this effect size 80% of the time)
- With all this experience, a standard deviation of 0.5 min is to be expected
- Better use a pressure of 40mmHg
- I must go prick a bunch of cuffed people and prick an equal number of noncuffed people
- I need to figure out how many …

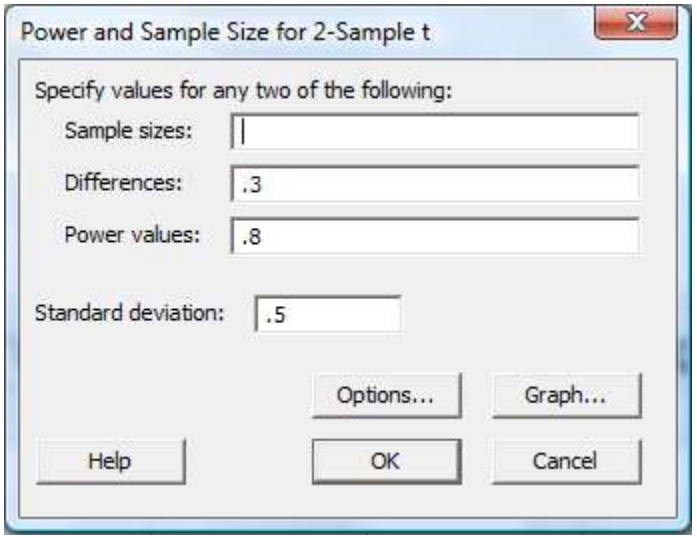
- Again, it is EASY to figure out how many subjects to test.

Power and Sample Size again

Step 1: From the Stat menu, choose “Power and Sample Size” and “2-Sample t”



Step 2: Fill in details



Power and Sample Size again (contd)

Power and Sample Size
2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05 Assumed standard deviation = 0.5

Difference	Sample Size	Target Power	Actual Power
0.3	45	0.8	0.803697

The sample size is for each group.

- For good measure, the researcher decided to enroll 50 persons to be cuffed and pricked (CP group) and 50 persons to be just pricked (JP group)
- That was a wise move. Because the final data tally turned out to be 39 (CP) and 43 (JP). Here's the data from the clinical trial

Group	N	Mean	StdDev
CP	39	2.192	0.765
JP	43	1.524	0.614

The clinical question, again: Does pressure increase bleeding time?

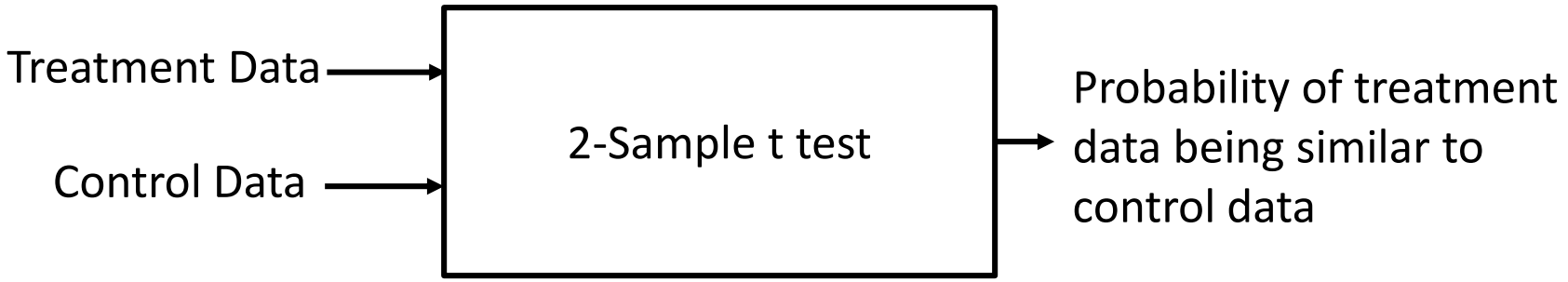
Ibid.

Now, the doctor ordered *this* test …

- … called the "2-Sample t test for equality of means"
- It is used to answer the question

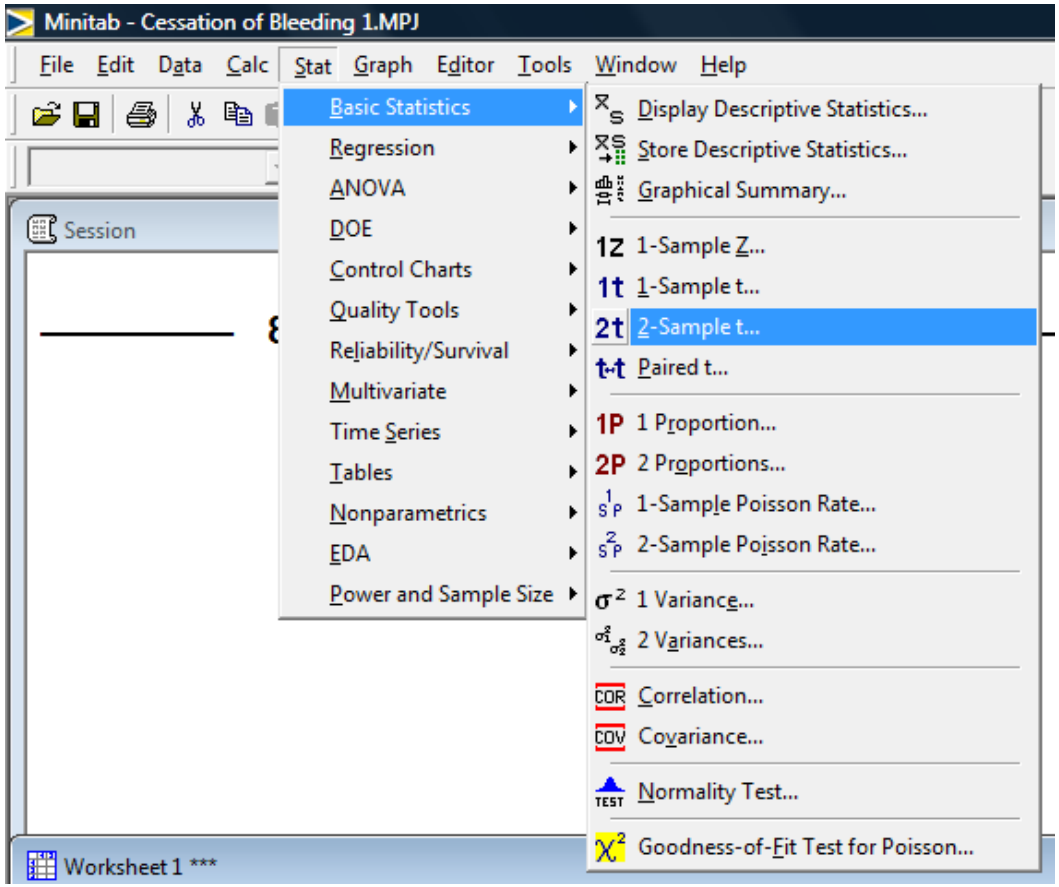
  "Are these two methods of treatments the same?"
- Or a little more formally

  "How likely is it that the two methods of treatment are similar?"
- We need to crank the numerical engine shown below

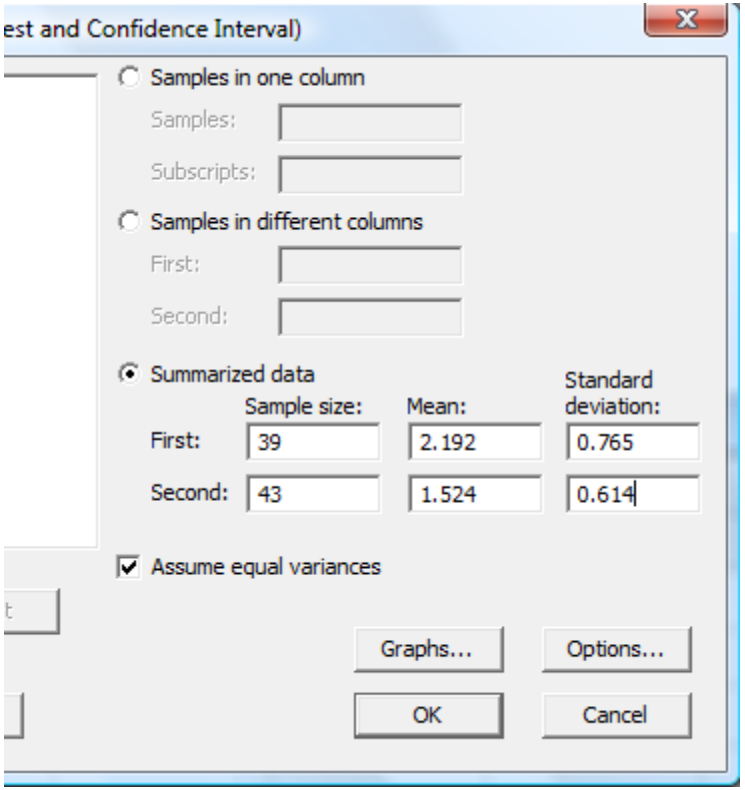


Relax, this test doesn't hurt either

Step 1: Choose "2-Sample t" from the menu



Step 2: Enter data (summarized, in this particular case)



Presto: 2-Sample t test results

Two-Sample T-Test and CI

Sample	N	Mean	StDev	SE Mean
1	39	2.192	0.765	0.12
2	43	1.524	0.614	0.094

Difference = mu (1) - mu (2)
Estimate for difference: 0.668
95% CI for difference: (0.364, 0.972)
T-Test of difference = 0 (vs not =): T-Value = 4.38 P-Value = 0.000 DF = 80
Both use Pooled StDev = 0.6899

Recall the question *"How likely is it that the two methods of treatment are similar?"*

Answer: VERY unlikely, with immeasurably small probability. There is strong evidence that pressure increases bleeding time.

A Power and Sample Size calculation shows that this is a very powerful test.

Power and Sample Size

2-Sample t Test
Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05 Assumed standard deviation = 0.5

Difference	Sample Size	Power
0.5	39	0.991791
0.5	43	0.995645

The sample size is for each group.

Wake Up!

- A clinical trial is a formal experiment to compare the efficacy and safety of two or more treatments.
- Trials require review and approval by a local body called the Institutional Review Board (IRB).
- Most clinical trials are registered. The website of the public registry maintained by the US government is www.clinicaltrials.gov
- Examples of mainstream medical journals that report clinical trials are:
  - Journal of the American Medical Association
  - New England Journal of Medicine
  - Lancet
  - British Medical Journal
  - Eurointervention

Legal Notice