PAIRED TESTS & POWER

10.12.2020

PROBLEM SET 3

* how’s it going?

AN EXPERIMENT

* *Suppose:* we hypothesize that the superior temporal sulcus (STS) responds more to **moving** than to **still** faces

* We perform an ECoG experiment in 20 subjects where we record high gamma responses to:

  * 50 still images of faces

  * 50 moving (video) clips of faces

* How do we test our hypothesis?

A SIMULATION

* Let’s begin by simulating what might happen if our hypothesis is right (simplest case):

* Still face responses are random with mean=0

* Moving face responses are random with mean=1

A 2ND SIMULATION

* What if some subjects' brains just respond more in general?

  * Add a random offset to each subject's responses

* This makes simple tests fail to find an effect!

PAIRED T-TEST

* scipy.stats includes a flavor of t-test which is useful for this!

  * **scipy.stats.ttest_rel** : aka the paired t-test, tests whether two *related* sets of samples have different means

  * This is equivalent to doing a t-test on the differences *within* each subject!

POWER OF STATISTICAL TESTS

* **power** is the probability of rejecting the null hypothesis when the null hypothesis is actually false

* i.e. how often you say “this is significant!” for a real effect

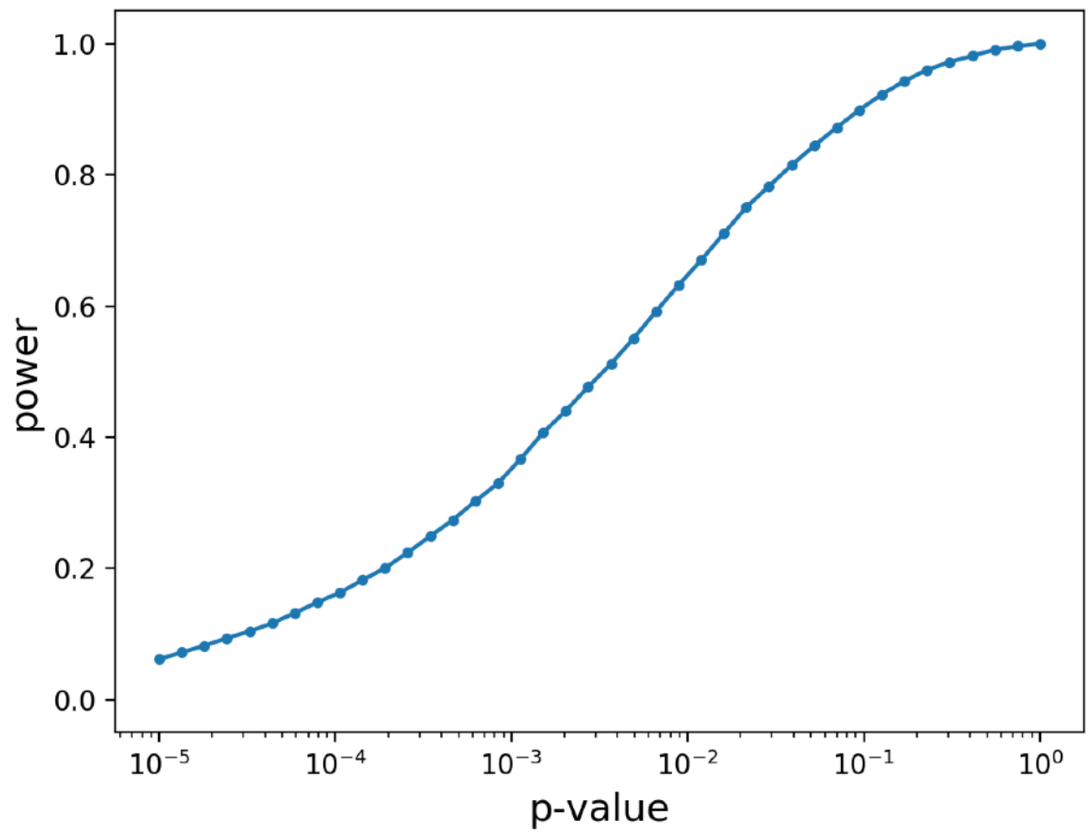
POWER OF STATISTICAL TESTS

* 80% power means that 20% of the time you get a false negative: the test says “not significant” when the effect is real

POWER OF STATISTICAL TESTS

* the power is related to the p-value threshold that you choose for a test

  * smaller threshold = lower power



POWER OF STATISTICAL TESTS

* it’s also related to the **effect size**

POWER OF STATISTICAL TESTS

* finally, power is also related to whether the assumptions of the test are valid and whether the test is mis-specified

POWER OF STATISTICAL TESTS

* e.g. if you have paired samples but use an un-paired t-test, then that could reduce your power to find a real effect

END