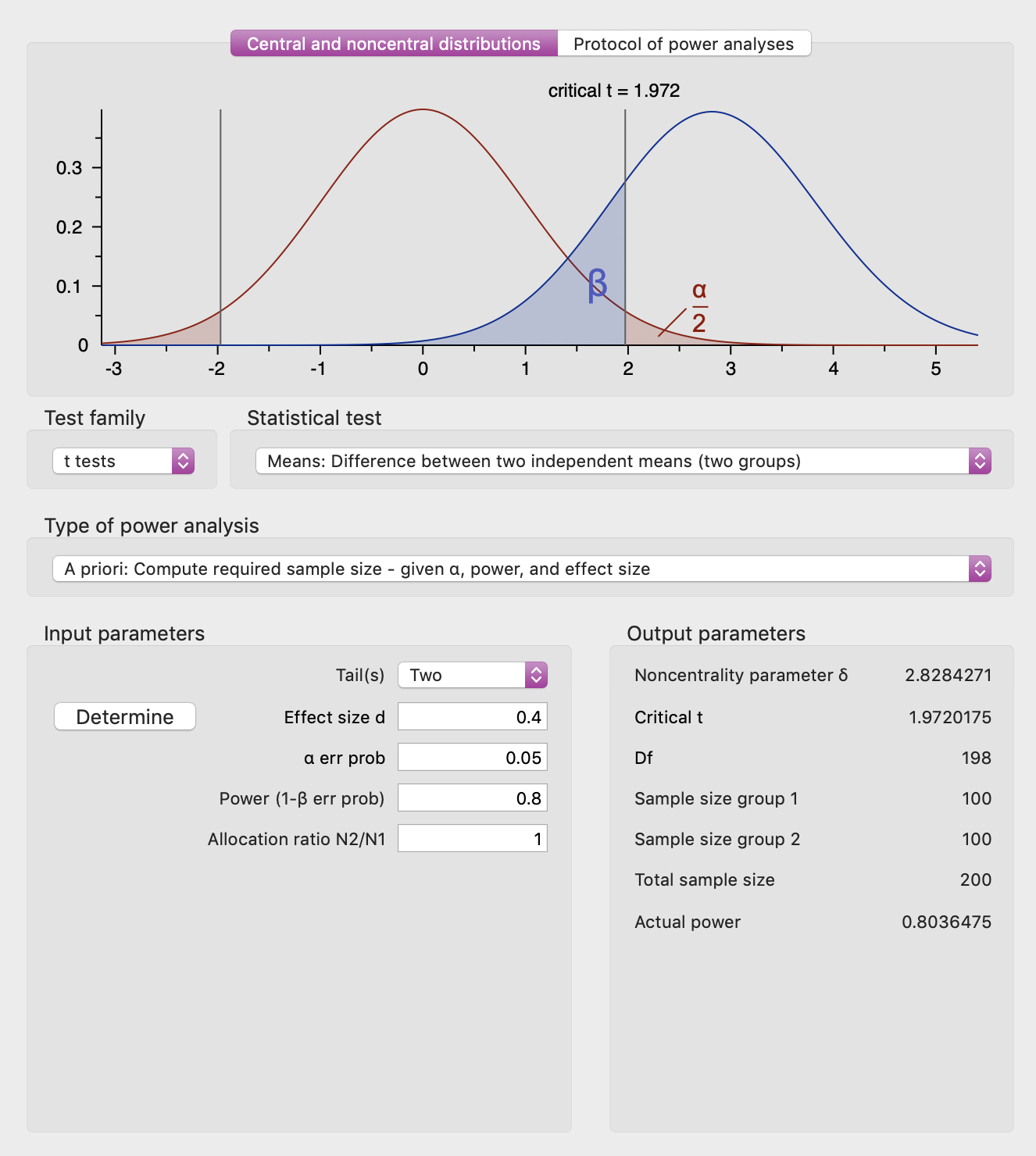
Independent t-test, a priori power analysis

* + - 1. You and your fellow graduate students are planning to conduct a study to compare the average overall GPA of undergraduate students in two different majors at CSU: psychology and physics. Given a desired power level of 0.80 and an alpha threshold of 0.05, what sample size would you need to determine if an effect size of 0.4 exists, in *any* direction, between the mean GPA of students in each major? Assume that the number of students from each major are equal in your sample.



*Answer: You would need an N = 200 to detect this effect.*

1a) In a sentence or two, what does an effect size of *d* = 0.4 *mean*, in terms of the average effect of being in one major versus the other on overall GPA? (hint: consider the formula for cohen’s *d)*

*“d = (Mt - Mc)/SD, where Mt and Mc are the treatment and control group means, respectively, and SD is the pooled standard deviation.”*

* *d = 0.4 indicates that the GPA means of the two majors differ by 2/5 of a standard deviation*
* *the difference in the average GPA of students in one major versus the other is 40% as large as the SD of GPA (the outcome measure) within each of the majors.*
* *The average effect of major is 2/5ths as large as the variability in GPA among students within the same major).*

1b) If the study has sufficient power to detect an effect of *d* = 0.4, what is the full range of effect sizes that can it reliably detect?

*d ≥ 0.4*

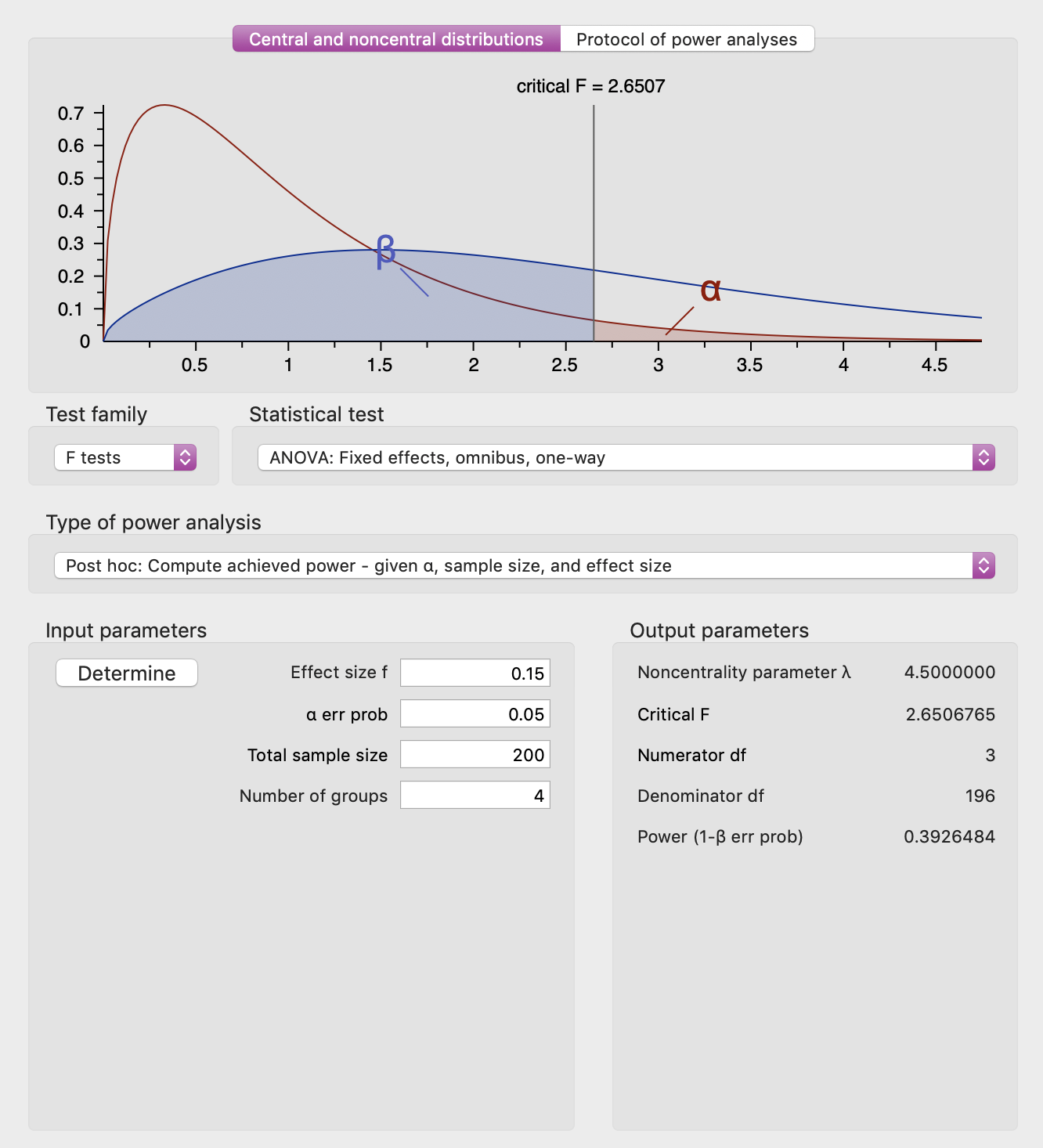
1c) Write an example of the nil hypothesis and one null hypothesis that you might have for the results of this study.

*Nil hypothesis: there is no significant difference between the mean overall GPA for students in each major.*

*Null hypothesis example: Psychology majors will have a mean overall GPA that is 15% higher physics majors (many possibilities here).*

Anova, post-hoc power analysis

* + - 1. You conducted a study to compare impact of four different types of food on self-reported mood: milk chocolate, spaghetti, steamed broccoli, and a turkey sandwich. You recruited a sample of 200 undergraduate students. 50 participants were each assigned to eat one type of food (each participant only ate one of the foods) and they were asked to self-rate their mood after eating it. After you collected all of your data, you conducted a 4-way ANOVA to compare the mean mood rating across the four types of food. At a significance threshold of alpha = 0.05, your results are non-significant. You take your results to your advisor and they ask you if your study was sufficiently powered… oops! You forgot to perform a power analysis beforehand, so you need to do it retroactively. Previous literature indicates that the effect size for the mean difference across the food groups should be around *f* = 0.15. Remembering what you learned about power in PSY 652, you would like to have a power level of at least 0.8. Was your study sufficiently power to detect an effect of *f* = 0.15, if it exists, across the four groups? If not, what power level does your study currently have?

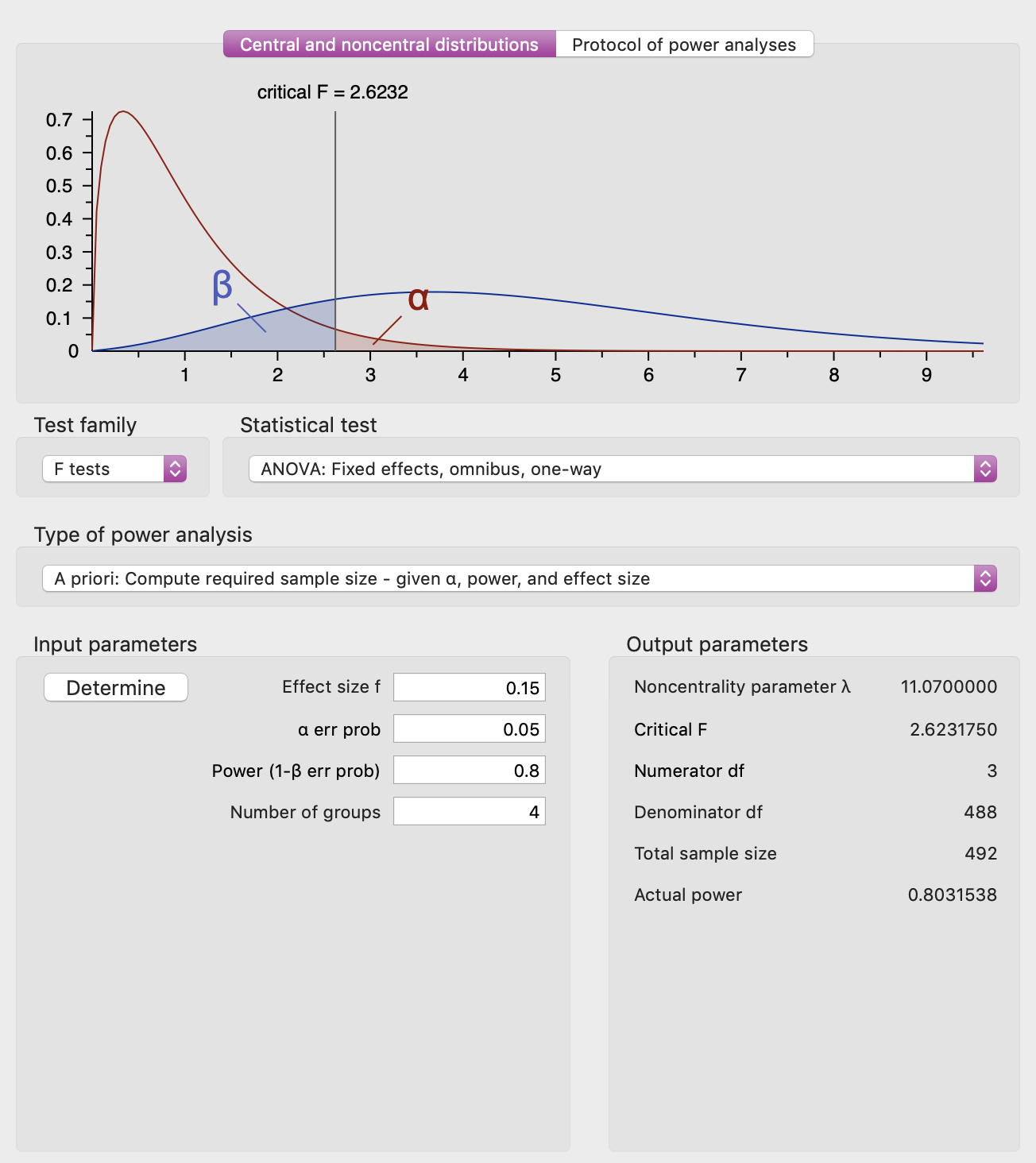


*Underpowered, power = 0.392*

2a) What is the probability of making a Type II error at your current power level?

*Since power = 0.392, probability of a Type II error = 0.608 (therefore ~3 times as likely to succeed as to fail)*

2b) Fortunately, NIH thinks your research is promising and they have given you additional grant funding for this study, so you can recruit more participants. How many more people do you need to recruit to detect an effect if *f* = 0.15 if it exists in your data?



*N=492 total, so need 292 more people (200 + 292)*