

Q&A Chat Webinar 5 (Power Analysis)

April 25th 2024 (Theory)

	Questions	Answers
1	third session, may i kindly have the recordings of the previous sessions please. Thank you.	Hello Medy, we have uploaded recordings, PowerPoint slides, and the Q&A chat on our website: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series within a week's time
2	Is it late to get more on DiD? on a quasi experimental data, the event studies, synthetic DiD.	Recordings, slides, and the Q&A chat are on our website: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series
3	How does power analysis help researchers determine the adequacy of their study's sample size?	The purpose of power analysis is exactly to help you understand the adequacy of your sample size! IT has to do with the minimum detectable effect size (MDES) - how big of an effect can you detect (as statistically significant) with the sample you have. Say you find a non-significant coefficient--is it because there really is no effect, or because your sample size is too small?
4	In what ways does knowledge of power analysis impact the interpretation of study results?	If a study is underpowered, it is "harder" or less likely to detect statistically significant effects. For example, while in reality a program has statistically significant impact, an underpowered study may find no impacts.
5	What is correct in Statistics terms, accept H_0 or fail to reject H_0 ?	fail to reject is correct H_0 is correct
6	So, is it then safe to say "power" is a situation of not committing both type 1 and type 2 errors?	power (1-B) is the probability of not making a type ii error
7	type 1 error if we reject null hypothesis (H_0) while it is true right? vice versa	incorrectly fail to reject H_0 when the H_a is true
8	Effect size also do matter in determining the statistical power.	yes
9	Thank you for these explanations. So, the Ideal case should be to have a large effect size and a big sample?	When designing a study and estimating sample size, you want to build the sample on the smallest effect size possible (given the budget)--0.20SD is a small effect size.
10	I've seen many times from many papers that power (1-beta) =0.8. Is it standard practice to play around with this number and degrade it to say 0.5? How does this affect the publishing studies where power is low, say (1-beta=0.5)	No, it is not standard to set power less than 80%. When your power is low, your study is less credible because you do not have sufficient "observations" to detect statistically significant differences between the groups you are studying
11	How large is large sample size? infinite, census?	"If you have a census, then you have everybody. That's large.

		For applied work that we do, we have to take into account the cost of collecting data and the benefits of adding additional units into the study to achieve sufficient power. So, you can define a large sample as the one that gives you sufficient power to detect impacts on your study
12	If comparing two groups, is it okay to increase the sample size for the comparison groups so as to proper matching?	No. The ratio of T and C in your study also affects power. Simply increasing C will not increase your power. The best powered study will be 50-50 split of T/C.
13	how do you find power of the test in case of a two tailed test	In the practical examples that we will do in the second session, those will be done with a two-tailed test. For the theory, it is easier to focus on a one-tailed test as it makes the graphs and explanation easier. Ultimately whether it's a one- or two-tailed test, power is the same (1-B), the likelihood of rejecting the null when it is false.
14	Can we also say power is the probability of rejecting H_0 when it is false?	Yes
15	Hi Ashu, do you mean increasing the sample size within the cluster?	Not sure what this is referring to. In general, you can increase power more by increasing the number of clusters and reducing the units per cluster. Adding more units to existing clusters doesn't help much, due to the homogeneity within clusters.
16	How can we get or determine rho?	I will show you next week that you can get it from existing national survey data.
17	Was k not 8 instead of 15?	on slide #19 we set $k=15$
18	How do we set rho?	Ideally you would obtain rho from national survey data, or from other similar studies.
19	What should be the formula for power calculation for a simple random sample?	You use the same formula but you do not have to correct for clustering. As a result, simple random sampling gives you more power than cluster random sampling, for the same number of observations (N).
20	SD show the level of peakness or whether it is skewed ??	It shows the spread of the distribution.
21	How do you choose the effect size (in terms of the SD used)	The rule of thumb is that 0.20SD is a small effect, .3-0.5 is a medium effect and above 0.5SD is a large effect. All else equal, you want to build your sample using the smallest MDES possible. I usually start with 0.20SD. If that is too costly, I go to 0.25SD and then 0.30SD.
22	Please I would like to know why you started with 0.20SD instead 0.10SD?	0.10SD is a tiny effect size, it is not likely to be of research/policy significance. An intervention that only generates a 0.10SD change isn't doing a good job!
23	Hello, you'll probably explain this later, but I wanted to know where do you adjust if you have more than 1 treatment arm? From the example, Stata gives you total sample size and number per sample which is exactly half?	AS you will always only be comparing one arm with another arm, it doesn't matter if you have 2 treatment arms. You can proceed as if you have just one arm. UNLESS you are also going to compare T1 cs T2, then you have to specify the MDES you expect, which will be probably different from the MDES when comparing a treatment arm with no treatment (C).
24	What if you can't find the baseline values for the indicator/variable of interest?	Go to the literature, or the national household survey, you MUST be able to come up with an idea of the baseline value!!

25	Can i get assistance on how to stuck cross sectional data like dhs?	You mean 'stack'? Use the 'append' command in STATA.
26	So, the effect size is implicit from the distribution of the population at the baseline, right?	Yes, the baseline distribution is crucial for the power analysis. But so is the effect size, which comes from you!
27	Should I use design effect for stratified sampling?	Yes
28	Thank you, excellent presentation	You enjoyed my hand drawn graphs. :)
29	Like are there other ways around to determine the sample size if you don't have baseline? Are there assumptions that you need to make? And are there examples of framing these assumptions?	BTW I saw your email and will endeavor to respond soon. You have to look in the literature and find a comparable 'baseline' mean and SD and use that to do the power analysis. But you need to do something, even if it is back-of-the-envelope. The worst thing is to field a survey where the sample size was too small--a total waste of money.
30	If an intervention has already been implemented and we are working to build a counterfactual retrospectively, does this approach still apply in picking the minimum sample to compare with the treated group?	If you're going to actually collect data on the comparison group, then yes you should do the calculations to know how many units to interview. If you are pulling existing data from a national survey, then you will do an 'ex-post' power analysis, like I showed you with the example from Emily.
31	Hi Sir, I want to do an analysis of the effects of creative industries on employment in my country. I have finished with my questionnaire. As I don't have any previous survey about creative industries How can I do to calculate the size of my sample? given the fact that I don't have Mean and SD Thanks. Nguenkwe	You need to go to the literature and find comparable sample characteristics even if from a neighboring country and use that. You do not want to waste your money on a survey with a sample size that is not adequate!
32	Could you kindly also highlight the case where one is dealing with an observational study (e.g. time to an event) and how the observational window also gets into the calculation of how many units to observe	The observation window will affect the MDES. So adjust your MDES to calculate the sample size you need for a given window.
33	Is the drop from 784 to 352 the preferred threshold? So, in practice, I should use 352. I need your response. thank you.	It all depends on the MDES that you are willing to have. If you think your intervention will have a very large impact, then you can go with a larger MDES, and thus smaller sample size.
34	is power size use uniquely for randomize study?	Definitely! How do you know the ideal sample size for each arm (T vs C)?
35	I want to talk about RCT	I think next year we will focus the webinar series on Impact Evaluation. We need to discuss with AERC.

36	Thanks, prof, for the wonderful presentation. How do I apply this to a known population? Taking your example of the cash transfer, what if we want to test the effect of the cash transfer on consumption? How would we compute a representative sample size?	This is the example I just did in the session, no?
37	Ok great thank for this presentation, but we need practical example in the case of cluster power size effect	Next week!

Q&A Chat Webinar 6 (Power Analysis)

May 2nd 2024 (Application)

#	Question	Answer
1	With a huge sample like KIHBS 2015/2016 (20,000 obs+), please advise whether researchers should still have to worry about statistical power. Thanks!	Live answered. Yes definitely, because once you start cutting the data into say rural, then small holder farmers, then women vs men, you will end up with a small sample!
2	After reduction of large data because of whatever you want to look at. How are you going to add up to meet the needed statistical power. Thank you	To understand the number of observations to achieve, say, 80% power, you would have to conduct a power analysis. We will show you how to conduct power analysis today.
3	If research is conducted in rural area, what SD is required? the overall (national) SD or only the SD for rural people?	SD for the population you are studying, in your case rural area
4	please, if I understand well, we should always do the power analysis even if the study is not an impact evaluation	Yes. Because if you cannot detect impacts/effects of your indicator/outcome variable, it could be because your study is underpowered. In that case, running a power analysis would strengthen your argument. Note that for any research, you are testing one group versus another, males versus females, or the effect of another year of schooling, etc.
5	from where we get power =0.8	we have set power at 80% = 0.8. We are calculating the minimum effect sizes. Power of 80% is the default, just like 5% significance level is the default.
6	Please is the initial mean of 7.04 also a pooled mean?	no. that's mean for women
7	On peut avoir la présentatin à la fin de la séance?	The presentation will be archived here https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series
8	I did not well understand the command please Pr.	we will review the command again in Stata examples.
9	1. Is 1000 sample size or population?	N = 1000 is the sample size in our data
10	what is MEDS Mean?	minimum detectable effect size
11	What is the range to say enough power?	The default, the minimum, is 80%. I usually calculate with 80% and 90% just to see the difference.

12	I am a little confuse in the case of Emily exPost Analysis, 0.18SD is very small effect compared to the three categories, I was expecting that the interpretation would be that Emily's data set is too small and might need to collect more samples.	By three categories, I assume you refer to (small, medium, and large on slide #6), correct? Remember those are just categories to help us understand how big the effect sizes are. The categories in themselves do not have any relationship to the power analysis we are conducting. This is to say, in the Emily example: we conducted a power analysis to determine the minimum detectable effect size (MDES), given our sample size (N=1000), power (0.8), mean (7.04), and sd (1.56). From that analysis, we determined that the MDES is 0.28, meaning we can only detect statistically significant differences between men and women that are larger than 0.28. $0.28 = 0.18 \text{ standard deviation } (0.28/1.56)$
13	I mean clarify how does the small SD (0.18SD) nullify the statistically insignificancy (MDES 0.28) that Emily found?	It doesn't nullify. MDES strengthen Emily's argument that there is no significant differences between the two groups she studied. Had she not run the power analysis, someone could have said maybe the reason Emily did not detect significant impact is because her study was underpowered. We have not actually done the statistical analysis in this study. Right now, we are just assessing the MDES.
14	In the previous example: we get 0.18sd, and said well powered study. when do we say the dataset has not enough power to answer the study questions?	We estimated the minimum detectable effect size in our dataset: we set power =0.8, alpha =0.05; from our data we have N=1000 observations, mean (women) = 7.04, and sd (men+women)=1.56. We found that our MDES =0.28, (which is $= 0.28/1.56 = 0.18 \text{ SD}$). This implies that we are less likely to detect statistically significant effect that are lower than 0.18 SD. We then go back to our regression analysis to see what we have actually estimate. In Emily's case, we showed that she will be able to detect (as statically significant) an effect size of 0.18SD. So her study was well powered.
15	Sure. Why are we not taking the average for the pooled sample, ie combining male and female?	We used mean of one group (which was women), and we wanted to estimate the minimum detectable differences between men and women. We could have also used the mean of men and see what that minimum detectable differences would be. However, we cannot use the mean of both groups (men and women) to detect the differences between the two groups. We only need the mean of one group

16	What if the subjects are homogeneous, does the influence of K reduces on the DEFF?	The more homogenous, the higher is rho, and the greater the design effect (less power, everything else equal).
17	I saw somewhere in the study using a matching design where they have more subject in the control than in the treatment, is this different from the K are referring here (that we rarely distinguish the K - clusters)?	Yes that is different. In general, for matching designs, you should have 3x as many units to draw from in the comparison group, relative to the T group.
18	Hello, Is it possible to get a 10-15 minutes chat with Prof? I have a tax administration dataset and a program that was implemented to improve compliance that I am evaluating as part of my thesis. I have a couple of questions and clarifications relating to what we have covered so far that I could really use some input on.	Please send Jacque an email with your question, plus some small background of the study parameters.
19	What number should I insert if the cluster in the treatment and control group are not the same.	Use the mean, or you can select 'unequal cluster sizes' and then it will give you options for the sizes for T and C.
20	Share the do file	You can access today's STATA files on the website through this link: https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series . Scroll down to May 2nd and then click on STATA files to access these
21	where is the data	https://www.airp3-africa.org/applied-quantitative-analysis-webinar-series