2019ADS2 Week11 Power and Sample Size

by Wanlu Liu, wanluliu@intl.zju.edu.cn

2019-11-25

1. Introduction

This R Markdown file contains a tutorial of assessing the effects of different statistical power and sample size. For those problems that need to workout manually, please include the statistical formula and process of how you get the final number.

2. one sample sample size estimation

An investigator wants to estimate the mean birth weight of infants born full term (approximately 40 weeks gestation) to mothers who are 35 years of age and older The mean birth weight of infants born full-term to mothers under 35 3,510 grams with a standard deviation of 385 grams. How many women older than 35 years of age must be enrolled in the study to ensure that a 95% confidence interval estimate of the mean birth weight of their infants has a margin of error not exceeding 100 grams?

1. step 1: identify experiemnt type, significance leve, criticle value, effect size, variance. This is a one sample, two sided test with normal distribution.

$$\alpha = 0.05$$

$$Z_{1-\alpha/2} = 1.96$$
$$EffectSize = 100$$
$$variance = 385$$

   2.Step 2: calculate the sample size needed with the following formula.

$$n = \frac{Z_{1-\alpha/2} * variance}{EffectSize}$$
$$n = \frac{1.96 * 385}{100}$$
$$n = 56.9$$

2. Step 3. In planning the study, the investigator must consider the fact that some women may deliver prematurely. If women are enrolled into the study during pregnancy, then more than (the sampe size you calcuated before) women will need to be enrolled so that after excluding those who deliver prematurely, we will have our designed number of effective experiment datapoints. For example, if 5% of the women are expected to delivery prematurely (i.e., 95% will deliver full term), then how many women need to be enrolled (n_{total})?

$$n_{total} = 56.9/0.95 = 60$$

4. Step 4. Conclusion. In total, we need to enroll 60 women considering 95% of them deliver full time, and in this case, we can have 57 effective datapoint to investigate the effect of mother's age (older than 35) on the baby weight with avarege of 3510 grams and 385 grams of standar deviation and less than 100 grams error.

2. two sample sample size estimation

An investigator is planning a study to assess the association between coffee consumption and grade point average among college seniors. The plan is to categorize students as heavy drinkers or not using 5 or more cups of coffee on a typical day as the criterion for heavy consumption. Mean grade point averages will be compared between students classified as heavy consumption versus not using a two independent samples test of means. The standard deviation in grade point averages is assumed to be 0.42 and a meaningful difference in grade point averages (relative to coffee consumption status) is 0.25 units. How many college seniors should be enrolled in the study to ensure that the power of the test is 80% to detect a 0.25 unit difference in mean grade point averages? Use a two-sided test with a 5% level of significance.

1. step 1 : identify key values. $\alpha = 0.05$, $\beta = 0.2, 1 - \beta = 0.8$, $Z_{1-\alpha/2} = 1.96$, $Z_{1-\beta} = 0.84$, $\sigma = 0.42$, $\delta(\mu_1 - \mu_2) = 0.25$
2. Step 2: calculate the sample size needed with the following formula.
$$n = 2 * (\frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\frac{\delta}{\sigma}})^2$$
$$n = 2 * (\frac{1.96 + 0.84}{\frac{0.25}{0.42}})^2 = 44.25523$$
3. Step 3. Conclusion. Sample sizes of $n_i = 45$ heavy coffee consumption and 45 who drink few fewer than five cups of coffee per typical day will ensure that the test of hypothesis has 80% power to detect a 0.25 unit difference in mean grade point averages.
4. Step 4. Test this with R power.t.test function.

```
delta=0.25
sigma=0.42
d=delta/sigma
power.t.test( d = d, sig.level = 0.05,
              power = 0.8,
              type ='two.sample',alternative = "two.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 45.28604
##           delta = 0.5952381
##              sd = 1
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

3. relationship between statistical power, sample size, significance leve and effect size.

In this problem, let's investigate the relationship between statistical power, sample size, significance leve and effect size. We will use the a new example.

3.1 Let's assume for a two sample, two-sided test, our significance level is 0.05. The standard variation of our sample is 0.5. If our sample size is 20, what's the statistical power under a effect size (mean difference of two population) of 0.4?

```r
delta=0.4
sigma=0.5
d=delta/sigma
power.t.test( n=20, d = d, sig.level = 0.05,
              type ='two.sample',alternative = "two.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 20
##           delta = 0.8
##              sd = 1
##       sig.level = 0.05
##           power = 0.6933994
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

3.2 What happens to our statistical power (increase or decrease) if we increase our significance level to 0.1.

```r
delta=0.4
sigma=0.5
d=delta/sigma
power.t.test( n=20, d = d, sig.level = 0.5,
              type ='two.sample',alternative = "two.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 20
##           delta = 0.8
##              sd = 1
##       sig.level = 0.5
##           power = 0.9676752
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

3.3 What happens to our statistical power (increase or decrease) if we decrease our sample size to 10?

```r
delta=0.4
sigma=0.5
d=delta/sigma
power.t.test( n=10, d = d, sig.level = 0.5,
              type ='two.sample',alternative = "two.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 10
##           delta = 0.8
```

```
##              sd = 1
##       sig.level = 0.5
##           power = 0.8649516
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

3.4 What happens to our statistical power (increase or decrease) if we increase our effect size to 0.8?

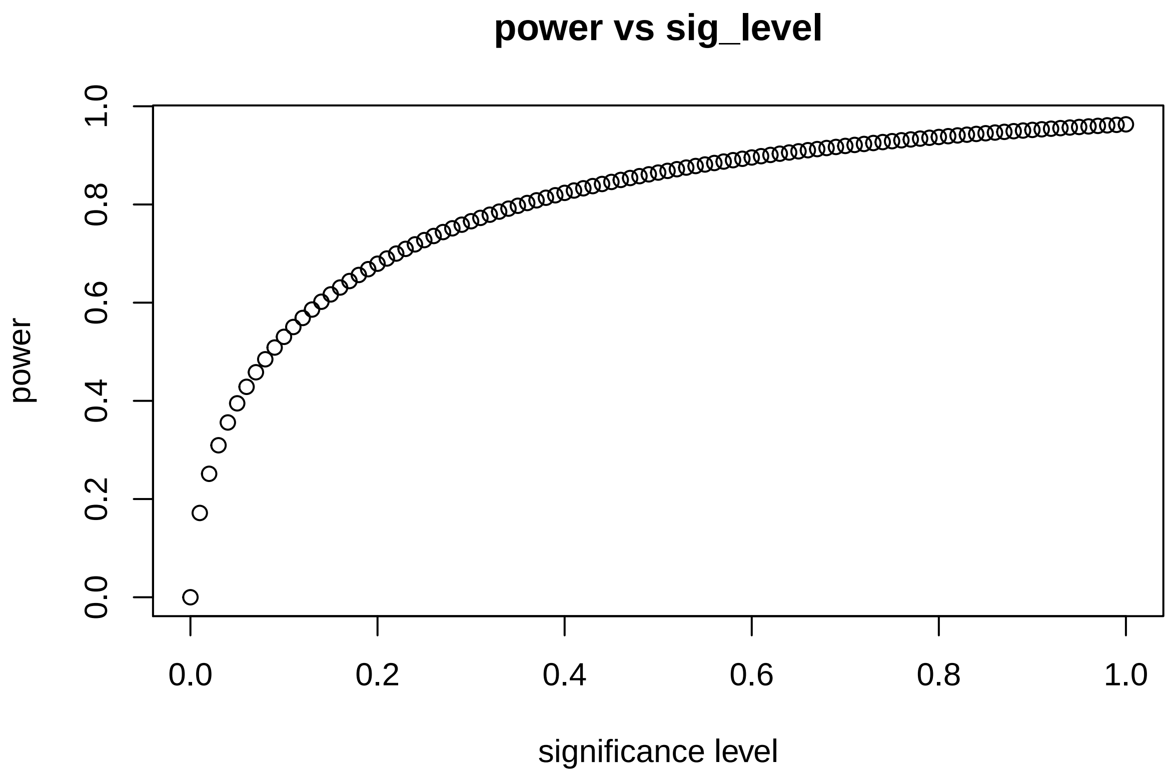
```
delta=0.8
sigma=0.5
d=delta/sigma
power.t.test( n=10, d = d, sig.level = 0.5,
              type ='two.sample',alternative = "two.sided")
```

```
##
##      Two-sample t test power calculation
##
##               n = 10
##           delta = 1.6
##              sd = 1
##       sig.level = 0.5
##           power = 0.998011
##     alternative = two.sided
##
## NOTE: n is number in *each* group
```

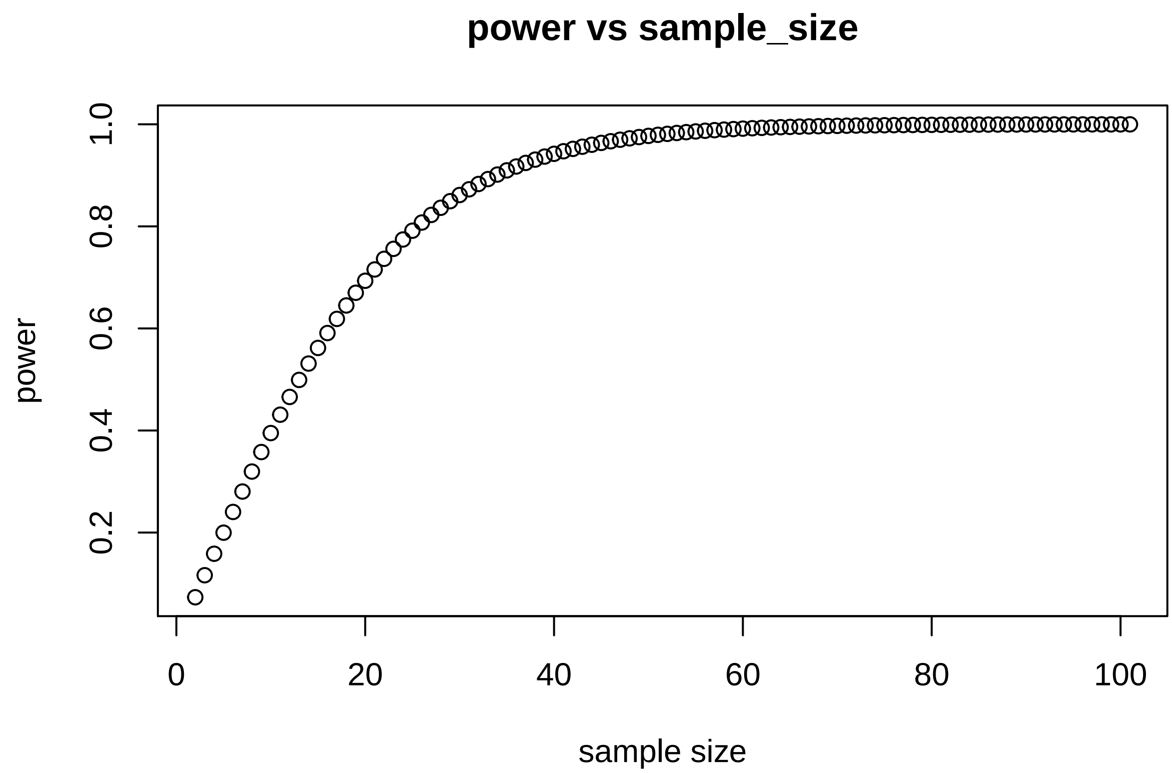
3.5 Advanced challenges (optional). If you can figure it out yourself, you are an absolute 'R master'!

Based on the example from the proble, can you use simulation and R plotting to figure out the relationship between statistical power vs significance level, sample size, and effect size? Three curves with y-axis being the statistical power and x-axis being the different significance level, or sample size, or effect size (each simulate 100 data points for plotting).

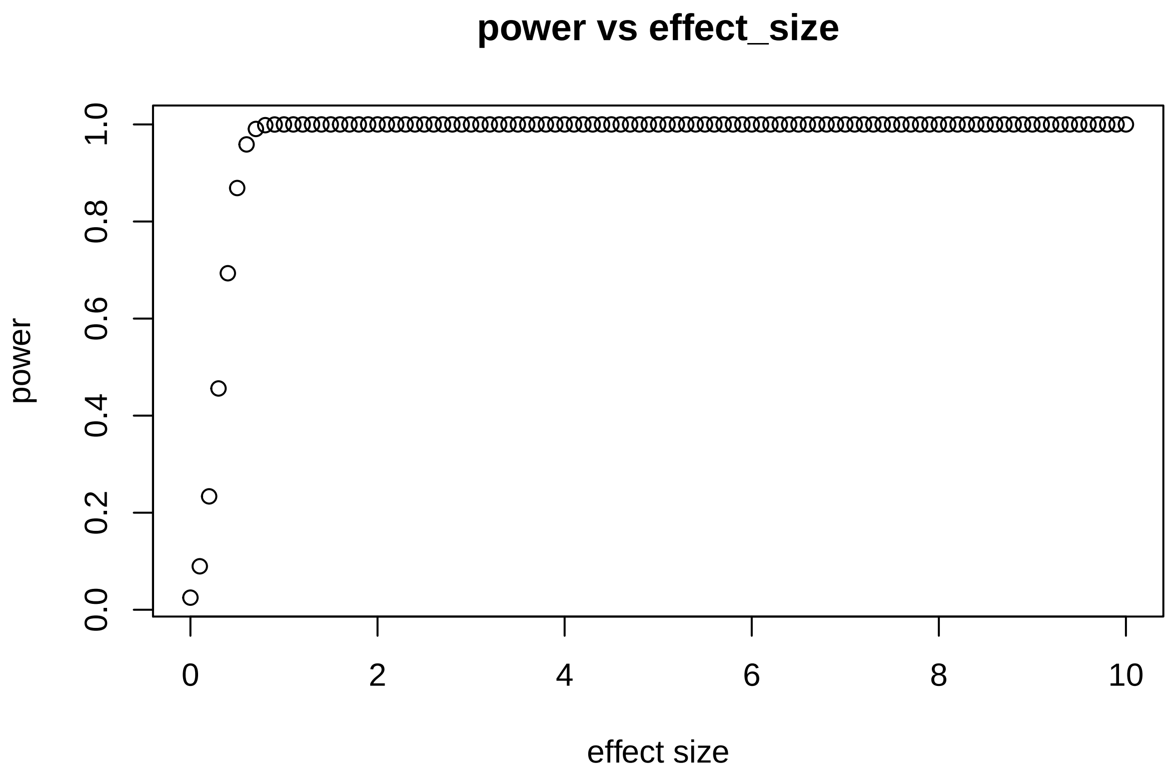
```
delta=0.4
sigma=0.5
d=delta/sigma
sig=seq(0,1,by=0.01)
power=rep(0,length(sig)) #initialize power vector
for (i in 1:length(sig)){
  power[i]=power.t.test( n=10, d = d, sig.level = sig[i],
              type ='two.sample',alternative = "two.sided")$power
}
plot(x=sig,y=power,xlab="significance level",ylab="power",main="power vs sig_level")
```




```
delta=0.4
sigma=0.5
d=delta/sigma
sample=c(2:101)
power=rep(0,length(sample)) #initialize power vector
for (i in 1:length(sample)){
  power[i]=power.t.test( n=sample[i], d = d, sig.level = 0.05,
              type ='two.sample',alternative = "two.sided")$power
}
plot(x=sample,y=power,xlab="sample size",ylab="power",main="power vs sample_size")
```




```r
sigma=0.5
delta=seq(0,10,by=0.1)
d=delta/sigma
power=rep(0,length(delta)) #initialize power vector
for (i in 1:length(delta)){
  power[i]=power.t.test( n=20, d = d[i], sig.level = 0.05,
              type ='two.sample',alternative = "two.sided")$power
}
plot(x=delta,y=power,xlab="effect size",ylab="power",main="power vs effect_size")
```



4. relationship between sample size vs p-value

Let's see the relationship between sample size vs p-value.

4.1 First, for sample A, let's generate 5 random number from normal distribution with mean of 10 and sd of 5. Then, for sample B, let's generate 5 random number from normal distribution with mean of 11 and sd of 5. Now we want to compare whether there is any significance difference between the mean of sample A and B, what should we do? write out the R code. Is there any significant difference for the mean of sample A and B?

```
set.seed(13)
a=rnorm(5,mean=10,sd=5)
b=rnorm(5,mean=11,sd=5)
t.test(a,b)
```

```
##
##  Welch Two Sample t-test
##
## data:  a and b
## t = -0.10438, df = 7.6853, p-value = 0.9195
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.636179  5.151360
## sample estimates:
## mean of x mean of y
##  13.37906  13.62147
```

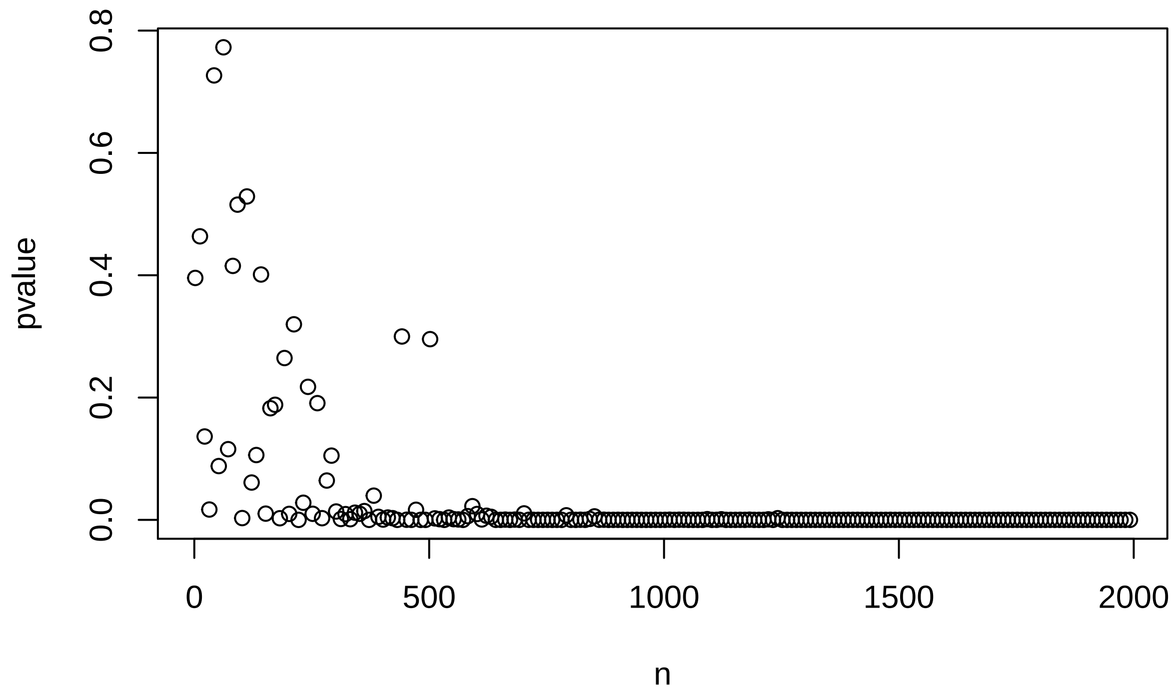
4.2 What if we now we increase the sample size to 500 (instead of 5) for sample A and B? Is there any significant difference for the mean of sample A and B? Write out R code.

```
a=rnorm(500,mean=10,sd=5)
b=rnorm(500,mean=11,sd=5)
t.test(a,b)
```

```
##
##  Welch Two Sample t-test
##
## data:  a and b
## t = -3.3992, df = 988.45, p-value = 0.0007028
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.696671 -0.454689
## sample estimates:
## mean of x mean of y
##  9.926613 11.002293
```

4.3 Advanced challenges (optional). If you can figure it out yourself, you are an absolute 'R master'! Just like what we did for 3.5, can you use simulation and plotting to visualize the relationship between sample size and p-value? Plot out the curve with p-value on the y-axis and different sample number on the x-axis (with at least 100 datapoints).

```
set.seed(13)
n=seq(2,2000,by=10)
a=list() #initialize a, why I use list to initialize a here?
b=list() #initialize a
pvalue=rep(0,length(n))
for (i in 1:length(n)){
  a[[i]]=rnorm(n[i],mean=10,sd=5)
  b[[i]]=rnorm(n[i],mean=11,sd=5)
  pvalue[i]=t.test(a[[i]],b[[i]])$p.value
}
plot(x=n,y=pvalue)
```



R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.