BST 210
Applied Regression Analysis

"I wish we hadn't learned probability 'cause I don't think our odds are good."

Save the Date!

**** December 19: 11:30am-1:00pm FXB G12 ****

Special Program for our final class!

featuring guest speakers and their seminal contributions to the field of statistics and modeling

And of course a celebration!

Lecture 26
Plan for Today

Sample Size and Power continued…*now using software!*

- Two (or more) sample problems
- Extend to
  - Single covariate logistic models
  - Single covariate survival models
- Extend further to
  - Multivariable logistic
  - Multivariable survival models
- Examples and code (mixed with Lab)

Recall: Sample Size and Power

- No matter the type of data or research question(s) at hand, the motivation behind power and sample size calculations is always the same:

We just want to make sure that—if there <u>really is a clinically meaningful</u> association between our outcome and our predictor of interest—that we are able to design a study that is actually <u>powered to be able to detect</u> this association!

Recall: Sample Size and Power

- Generally speaking:

(1) sample size N = **function(**1-β, α, prevalence of outcome (across groups), clinically meaningful effect size, 1 or 2-sided test, drop-out/missingness rate**)**

and

(2) Power 1-β = **function(**N, α, prevalence of outcome (across groups), clinically meaningful effect size, 1 or 2-sided test, drop-out/missingness rate**)**

Last time we also discussed:

- Relationships between power, SS, and constituents of each
- How to 'correct' for missing or censored outcomes: $\boldsymbol{1 / (1 - \lambda)}$
- 1 sided versus 2 sided calculations
- What influence large or small effect sizes have on SS and power
- That we need an analysis method/model in order to derive power or SS
- Oftentimes such calculations are just the starting point!

Computing Caveats around Sample Size/Power

- SAS is excellent and well documented (regulatory/audit-driven)

- Stata is also great and quite well documented

- R…less so, but still has great flexibility for bringing to life methods out in the literature, and customized means of assessing sample size and power questions

… You will see a bit of each across lecture, lab and HW

Power and Sample Size Calculations: Stata

- [PSS] power -- Power and sample-size analysis for hypothesis tests

Syntax

Compute sample size

power method ... [, power(numlist) power_options ...]

Compute power

power method ..., n(numlist) [power_options ...]

Compute effect size and target parameter

power method ..., n(numlist) power(numlist)
[power_options ...]

Power and Sample Size Calculations: Stata

One sample	
onemean	One-sample mean test (one-sample t test)
oneproportion	One-sample proportion test
onecorrelation	One-sample correlation test
onevariance	One-sample variance test
Two independent samples	
twomeans	Two-sample means test (two-sample t test)
twoproportions	Two-sample proportions test
twocorrelations	Two-sample correlations test
twovariances	Two-sample variances test
Two paired samples	
pairedmeans	Paired-means test (paired t test)
pairedproportions	Paired-proportions test (McNemar's test)
Analysis of variance	
oneway	One-way ANOVA
twoway	Two-way ANOVA
repeated	Repeated-measures ANOVA

Power and Sample Size Calculations in SAS

- **Syntax: POWER Procedure**
- The following statements are available in PROC POWER:

LOGISTIC, MULTREG, ONECORR, ONESAMPLEFREQ, ONESAMPLEMEANS, ONEWAYANOVA, PAIREDFREQ, PAIREDMEANS, PLOT, TWOSAMPLEFREQ, TWOSAMPLEMEANS, TWOSAMPLESURVIVAL, TWOSAMPLEWILCOXON

Power and Sample Size Calculations in R

- In R, there are various libraries that include power and sample size calculators, and you must look around for what you need or program methods yourself.

- R does provide a lot of flexibility for exploring interesting questions within this topic.

- library(Hmisc)
- library(powerMediation)
- library(survival)
- library(powerSurvEpi)

Computing, Power and Sample Size

- And, for all of these packages, there are other procedures or home-grown software possibly available.
- There are also specialized power packages (including nQuery, PASS, and others) available, and things you can find for free on the web.
- Search around, but investigate to be sure that you are doing exactly what you need to do.
- Let’s look at a bunch of examples →

Stata (power): Two-Sample Binomial Test

- The command is:

power twoproportions p_1 p_2, alpha(0.05) n_1(#) n_2(#)

```
. power twoproportions 0.01 0.02, alpha(0.05) n1(3000) n2(1000)

Estimated power for a two-sample proportions test
Pearson's chi-squared test
Ho: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
            N =      4000
           N1 =      3000
           N2 =      1000
        N2/N1 =    0.3333
        delta =    0.0100  (difference)
           p1 =    0.0100
           p2 =    0.0200

Estimated power:

        power =    0.6658
```

Stata (sample size): Two-Sample Binomial Test

We can specify the ratio of sample sizes (if unequal) for a sample size calculation

```
power twoproportions 0.01 0.02, alpha(0.05)
nratio(0.333333) power(0.80)
```

```
Performing iteration ...

Estimated sample sizes for a two-sample proportions test
Pearson's chi-squared test
Ho: p2 = p1  versus  Ha: p2 != p1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.0100  (difference)
           p1 =    0.0100
           p2 =    0.0200
        N2/N1 =    0.3333

Estimated sample sizes:

            N =      5740
           N1 =      4305
           N2 =      1435
        N2/N1 =    0.3333
```

Stata: Another Example

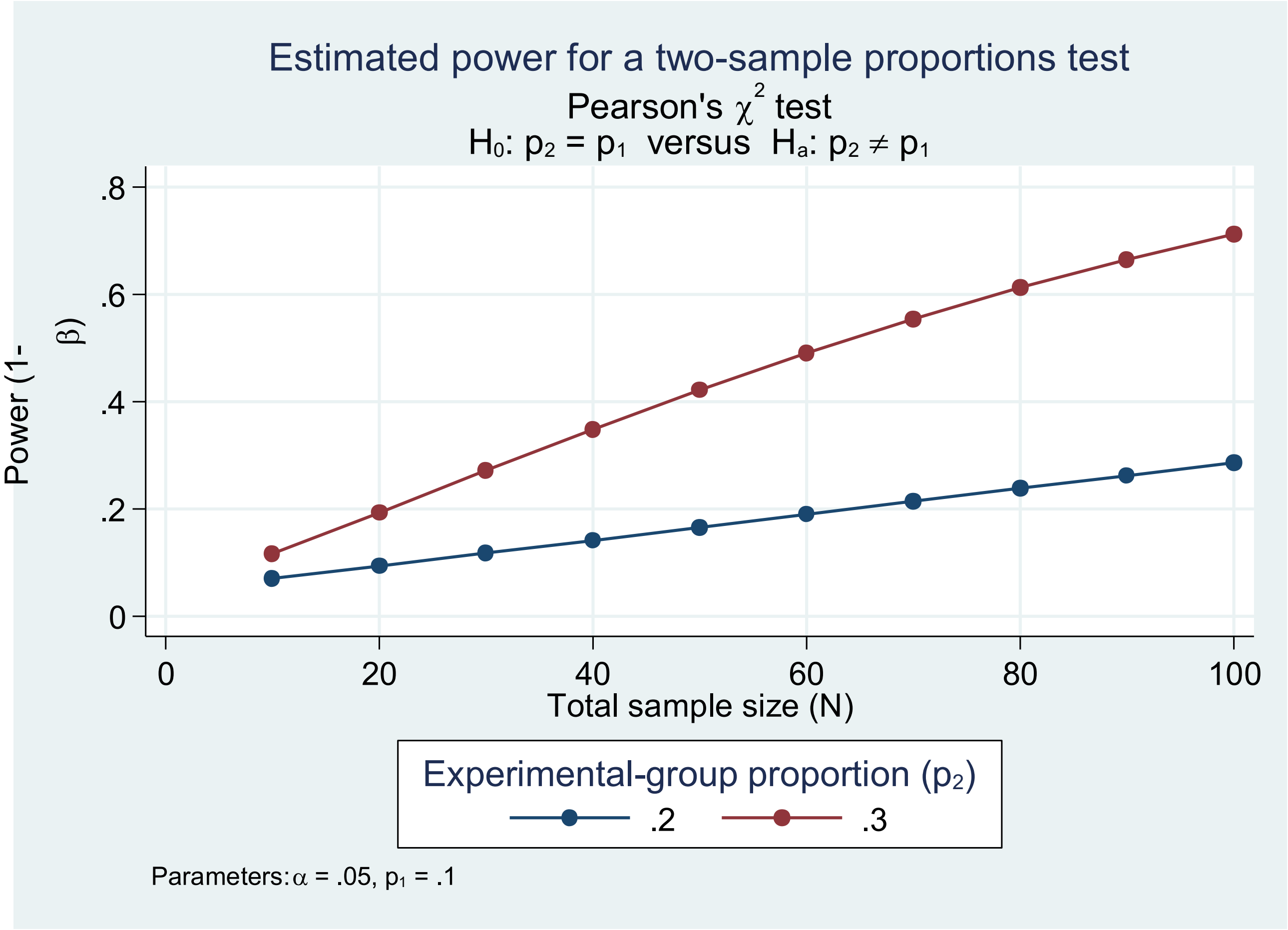
```
. power twoproportions 0.1 (0.2 0.3), n(10 (10) 100)

Estimated power for a two-sample proportions test
Pearson's chi-squared test
Ho: p2 = p1  versus  Ha: p2 != p1

  +-------------------------------------------------------------------+
  |   alpha   power       N      N1      N2   delta      p1      p2 |
  |-------------------------------------------------------------------|
  |     .05  .07035      10       5       5      .1      .1      .2 |
  |     .05  .09349      20      10      10      .1      .1      .2 |
  |     .05   .1171      30      15      15      .1      .1      .2 |
  |     .05    .141      40      20      20      .1      .1      .2 |
  |     .05   .1651      50      25      25      .1      .1      .2 |
  |     .05   .1894      60      30      30      .1      .1      .2 |
  |     .05   .2137      70      35      35      .1      .1      .2 |
  |     .05    .238      80      40      40      .1      .1      .2 |
  |     .05   .2622      90      45      45      .1      .1      .2 |
  |     .05   .2863     100      50      50      .1      .1      .2 |
  |     .05   .1158      10       5       5      .2      .1      .3 |
  |     .05    .193      20      10      10      .2      .1      .3 |
  |     .05   .2712      30      15      15      .2      .1      .3 |
  |     .05   .3479      40      20      20      .2      .1      .3 |
  |     .05   .4214      50      25      25      .2      .1      .3 |
  |     .05   .4904      60      30      30      .2      .1      .3 |
  |     .05   .5541      70      35      35      .2      .1      .3 |
  |     .05   .6122      80      40      40      .2      .1      .3 |
  |     .05   .6647      90      45      45      .2      .1      .3 |
  |     .05   .7115     100      50      50      .2      .1      .3 |
  +-------------------------------------------------------------------+
```

Same Example More Visual

```
. power twoproportions 0.1 (0.2 0.3), n(10 (10) 100) graph
```



Same Example: More Intuition

```
. power twoproportions 0.1 (0.2 0.3), power(.80 .85 .90)

Performing iteration ...

Estimated sample sizes for a two-sample proportions test
Pearson's chi-squared test
Ho: p2 = p1  versus  Ha: p2 != p1
```

alpha	power	N	N1	N2	delta	p1	p2
.05	.8	398	199	199	.1	.1	.2
.05	.85	456	228	228	.1	.1	.2
.05	.9	532	266	266	.1	.1	.2
.05	.8	124	62	62	.2	.1	.3
.05	.85	142	71	71	.2	.1	.3
.05	.9	164	82	82	.2	.1	.3

Stata (power): Two-Sample Log-rank Test

```
. stpower logrank, hratio(.75) alpha(.05) n(400)

Estimated power for two-sample comparison of survivor functions
Log-rank test, Freedman method
Ho: S1(t) = S2(t)

Input parameters:

      alpha =    0.0500  (two sided)
     hratio =    0.7500
          N =       400
         p1 =    0.5000

Estimated number of events and power:

          E =       400
      power =    0.8152
```

So if we only had 400 events total (200 per group) we would have 81.5% power (E for events).

Stata (Sample Size): Two-Sample Log-rank Test

```
. stpower logrank, hratio(.75) alpha(.05) power(.9)

Estimated sample sizes for two-sample comparison of survivor functions
Log-rank test, Freedman method
Ho: S1(t) = S2(t)

Input parameters:

      alpha =    0.0500  (two sided)
     hratio =    0.7500
      power =    0.9000
         p1 =    0.5000

Estimated number of events and sample sizes:

          E =       516
          N =       516
         N1 =       258
         N2 =       258
```

So if we had 516 events (258 per group) we would have 90% power (E for events, N for sample size, no censoring).

Stata (Sample Size) : Two-Sample Log-rank with Withdrawals

```
. stpower logrank, hratio(.75) power(.9) wdprob(.2)

Estimated sample sizes for two-sample comparison of survivor functions
Log-rank test, Freedman method
Ho: S1(t) = S2(t)

Input parameters:

      alpha =    0.0500  (two sided)
     hratio =    0.7500
      power =    0.9000
         p1 =    0.5000
 withdrawal =     20.00%

Estimated number of events and sample sizes:

          E =       516
          N =       644
         N1 =       322
         N2 =       322
```

The common probability of withdrawing is assumed to be 20% in each group. Here, 516/(1 – 0.2) = 645.

Stata (Sample Size): Two-Sample Log-rank with Withdrawals

```
. stpower logrank, hratio(.75) power(.9) wdprob(.4)

Estimated sample sizes for two-sample comparison of survivor functions
Log-rank test, Freedman method
Ho: S1(t) = S2(t)

Input parameters:

      alpha =    0.0500  (two sided)
     hratio =    0.7500
      power =    0.9000
         p1 =    0.5000
 withdrawal =     20.00%

Estimated number of events and sample sizes:

          E =       516
          N =       860
         N1 =       430
         N2 =       430
```

The common probability of withdrawing is assumed to be 40% in each group. Here, 516/(1 – 0.4) = 860.

Bottom Line on SS and Power…

- Often, the sample sizes you estimate from simplistic assumptions are just a starting point – you probably need to adjust for

    - dropouts or dropins
    - noncompliance
    - censored observations
    - missing values
    - covariates and confounders

  or other related factors

- Not only that, investigators often anticipate effect sizes larger than what may be realistic

Extensions to Logistic and Survival Models

- In practice, we often need power and sample size calculations for more complex settings than one-sample or two-sample problems, as we often expect to adjust for other factors in a primary analysis of any research hypothesis

- Rosner's *Fundamentals of Biostatistics* provides these for stratified 2x2 tables (Mantel-Haenszel approach) and stratified incidence rates (Poisson comparisons) – you need to specify a common OR or IRR under the alternative, and other details

Next: Power for Logistic Regression

- Logistic regression power calculations for a single binary covariate (based on Wald test) reduce down to the two-group binomial proportion comparisons (based on Score test) we've considered previously (because the two tests are asymptotically equivalent)

- What about for a continuous covariate?

- Stata, SAS and R offer ways to do this, but with some specific parameters required. In particular, we must specify the probability of the outcome *at the mean value of the covariate x,* and the probability of the outcome *at one standard deviation higher than the mean value of x*

- These probability estimates then go into the calculations involving the ORs and related information for the SS and power computation

Power: Logistic Regression w/continuous covariate

Let's have an example:

- Suppose you are designing a study for which you anticipate the probability of success at the mean value of your continuous covariate to be 0.25.
- Furthermore, suppose that at one standard deviation higher than the mean value of your continuous covariate, the probability of success is 0.35
- What is the interpretation of this in terms of odds ratios?

Power: Logistic Regression w/continuous covariate

- Well, the same as always! The odds ratio comparing these two probabilities equals

  $$(0.35/0.65) \,/\, (0.25/0.75) = 1.615$$

- This is for a one SD increase in *x* (*not* a one unit change in *x*)

- If we know (or anticipate) the OR for a *one unit* increase in *x*, then that OR raised to the SD(*x*) power gives this OR for a one SD(*x*) unit increase in *x*

- Thus, you can figure out what is needed if you have available the probability of success at the mean value of *x* and the odds ratio associated with a one unit or a one SD(*x*) unit increase in *x*

Power: Logistic Regression w/continuous covariate

```
. powerlog, p1(.25) p2(.35) help

 Logistic regression power analysis
 One-tailed test: alpha=.05  p1=.25  p2=.35  rsq=0  odds
ratio=1.615384615384615

 power          n
 0.60         104
 0.65         118
 0.70         134
 0.75         152
 0.80         174
 0.85         201
 0.90         238

Explanation of terms
p1  -- the probability that the response variable equals 1
         when the predictor is at the mean
p2  -- the probability that the response variable equals 1
         when the predictor is one standard deviation above the mean
rsq -- the squared multiple correlation between the predictor
         variable and all other variables in the model
```

Power: Logistic Regression w/continuous covariate

```
. powerlog, p1(.10) p2(.35)

Logistic regression power analysis
One-tailed test: alpha=.05  p1=.1  p2=.35  rsq=0  odds ratio=4.846153846153846

power          n
0.60         145
0.65         156
0.70         169
0.75         184
0.80         200
0.85         221
0.90         248
```

Power for <u>Multivariable</u> Logistic Regression

- Finally, what if there are additional covariates in the model that you want to adjust for in your ultimate analysis? What might be a reasonable approach to somehow accounting for such variables which might end up associating with one another, when we are way back at the start determining sample size?

- All the logistic regression power calculations on previous slides can adjust for additional covariates if we have….

  the R^2 value!

  (that is, the square of the multiple correlation coefficient between the other covariates and the *x* covariate of primary interest, acting as the exposure variable)

- Basically, what is R^2 from the linear regression model predicting *x* from your other covariates?

- No additional covariates? $R^2 = 0$

Power for Logistic Regression

```
. powerlog, p1(.25) p2(.35) rsq(.2) help

 Logistic regression power analysis
 One-tailed test: alpha=.05  p1=.25  p2=.35  rsq=.2  odds
ratio=1.615384615384615

 power          n
 0.60         130
 0.65         147
 0.70         167
 0.75         190
 0.80         217
 0.85         251
 0.90         297

Explanation of terms
p1  -- the probability that the response variable equals 1
         when the predictor is at the mean
p2  -- the probability that the response variable equals 1
         when the predictor is one standard deviation above the mean
rsq -- the squared mulitple correlation between the predictor
         variable and all other variables in the model
```

Power for Logistic Regression

```
. powerlog, p1(.25) p2(.35) rsq(.4)

 Logistic regression power analysis
 One-tailed test: alpha=.05  p1=.25  p2=.35  rsq=.4  odds ratio=1.615384615384615

 power          n
 0.60         173
 0.65         196
 0.70         223
 0.75         253
 0.80         290
 0.85         335
 0.90         397
```

Next: Power for Cox Regression

- Cox regression power calculations for a single binary covariate reduce down to the two-group log-rank comparisons we've considered previously (again due to asymptotic equivalence)

- What about for a continuous covariate?

- Similar to the Logistic regression scenario with a continuous covariate. Stata offers a way to do this, based on a method developed by Hsieh and Lavori. Effectively, you need to propose a β coefficient (log hazard ratio) for the covariate you are testing (vs. the null hypothesis $\beta = 0$) plus the standard deviation for the continuous covariate x of interest

Power for Cox Regression

```
. stpower cox +1, sd(1)

Estimated sample size for Cox PH regression
Wald test, log-hazard metric
Ho: [b1, b2, ..., bp] = [0, b2, ..., bp]

Input parameters:

         alpha =    0.0500  (two sided)
            b1 =    1.0000
            sd =    1.0000
         power =    0.8000

Estimated number of events and sample size:

             E =         8
             N =         8
```

Power for Cox Regression

```
. stpower cox +1, sd(0.5)

Estimated sample size for Cox PH regression
Wald test, log-hazard metric
Ho: [b1, b2, ..., bp] = [0, b2, ..., bp]

Input parameters:

         alpha =    0.0500  (two sided)
            b1 =    1.0000
            sd =    0.5000
         power =    0.8000

Estimated number of events and sample size:

             E =        32
             N =        32
```

Power for Cox Regression

```
. stpower cox +1, sd(0.25)

Estimated sample size for Cox PH regression
Wald test, log-hazard metric
Ho: [b1, b2, ..., bp] = [0, b2, ..., bp]

Input parameters:

         alpha =    0.0500  (two sided)
            b1 =    1.0000
            sd =    0.2500
         power =    0.8000

Estimated number of events and sample size:

             E =       126
             N =       126
```

Power for Cox Regression

```
. stpower cox +1, sd(0.25) wdprob(.2)

Estimated sample size for Cox PH regression
Wald test, log-hazard metric
Ho: [b1, b2, ..., bp] = [0, b2, ..., bp]

Input parameters:

         alpha =    0.0500  (two sided)
            b1 =    1.0000
            sd =    0.2500
         power =    0.8000
 withdrawal(%) =     20.00

Estimated number of events and sample size:

             E =       126
             N =       157
```

Simulation Methods

- In some situations, you may wish to develop power calculations based on a pilot data set that you have available
- You might sample (with replacement) from this data set with a larger sample size and determine the proportion of times the null hypothesis was rejected – this would be your estimate of power
- E.g., if you had a pilot data set with 25 observations, you could sample 100 or 150 or 200 observations (with replacement) from the data set to estimate the power

Simulation Methods

- Alternately, you might use descriptive statistics from your pilot data set (e.g., sample means and SDs of covariates, β coefficients that you estimated) and base simulations of power based on that

- Similar methods might be used if you have sufficient data available from a publication

- You likely need to perform a sensitivity analysis, e.g., using a range of β coefficients, to assess power in different situations – your pilot study probably did not accurately estimate your true β coefficient (ie why are you doing another study?)

Power Points!

- Lean on facts regarding relationships between test statistics (asymptotic or otherwise), sample size and power, sided-ness of hypothesis tests, size of effect, and also measures of association…as seen below (homework hint!):

- Keep in mind that RR ~ OR when outcome events are rare
- Also that HR ~ IRR ~ RR when the follow-up time is short (HR is instantaneous measure yet constant over time under PH), so under rare event/short follow-up settings the HR ~ OR
- Whereas if follow-up time is longer (even under rare event setting) HR will not be similar to OR or RR.

Power Points!

- There are many power calculations performed in Stata, SAS, or R (or other packages), but be sure to understand exactly what the package is assuming and doing
- You likely need to adjust for complications in your statistical analyses, and inflate the sample size as needed to ensure sufficient power
- Be sure to use realistic estimates of effect sizes
- In some cases, you might seek statistical consultation with a colleague or a field expert

Coming Up

- Introduction to the analysis of missing data!
- Review for Final Exam
- Save the date! ** December 19: 11:30am-1:00pm **
  Special Program for our final class!
    - featuring guest speakers and their seminal contributions to the field of statistics and modeling

Until then, let's get to some examples! →