Lecture 19 – Sample Size and Summary

Sample size for binomial outcomes

Comparison of two proportions: $H_0: P_1 = P_0$ *vs.* $H_1: P_1 \neq P_0$

Or compare the odds ratio from the model: $\text{logit}(p_i) = \alpha + \beta\ (Group = 1)$

$$H_0: \beta = 0 \text{ vs. } H_1: \beta \neq 0$$

Sample size estimates for these hypotheses are not the same!

- **First tests a difference in two proportions so uses binomial sampling variability**
- **Second tests the variability of the log odds ratios**

Sample size computation for the difference in two binomial proportions

- **Assume n = size of one of the two equal sized groups**
- **1-sided α (use α/2 for 2-sided) and power θ**

Test $H_0: P_1 = P_0$ **and define** $\bar{P} = \dfrac{P_0 + P_1}{2}$

$$\Rightarrow n = \frac{\left(z_{1-\alpha}\sqrt{2\bar{P}(1-\bar{P})} + z_{1-\theta}\sqrt{P_0(1-P_0)+P_1(1-P_1)}\ \right)^2}{(P_1 - P_0)^2}$$

Sample size computation for the odds ratio (Whitemore, 1981)

Assume n = sample size of one of the two equal sized groups

π = proportion of unexposed subjects in the study

1-sided α (use α/2 for 2-sided) and power θ

Test $H_0: \beta = 0$ *vs.* $H_1: \beta = \beta^*$

$$\Rightarrow n = (1+2P_0)\frac{\left(z_{1-\alpha}\sqrt{\frac{1}{1-\pi}+\frac{1}{\pi}} + z_{1-\theta}\sqrt{\frac{1}{1-\pi}+\frac{1}{\pi e^{\beta^*}}}\right)^2}{P_0\ (\beta^*)^2}$$

Most situations require more complex models since several covariates will be included in most models

Rule of thumb due to Peduzzi et al (1996):

Number of parameters (p+1) < 10% of the # of events

Where the number of events is the minimum of the Y=0 and Y=1

Actually more complex than this – depends on the type and distribution of the covariates

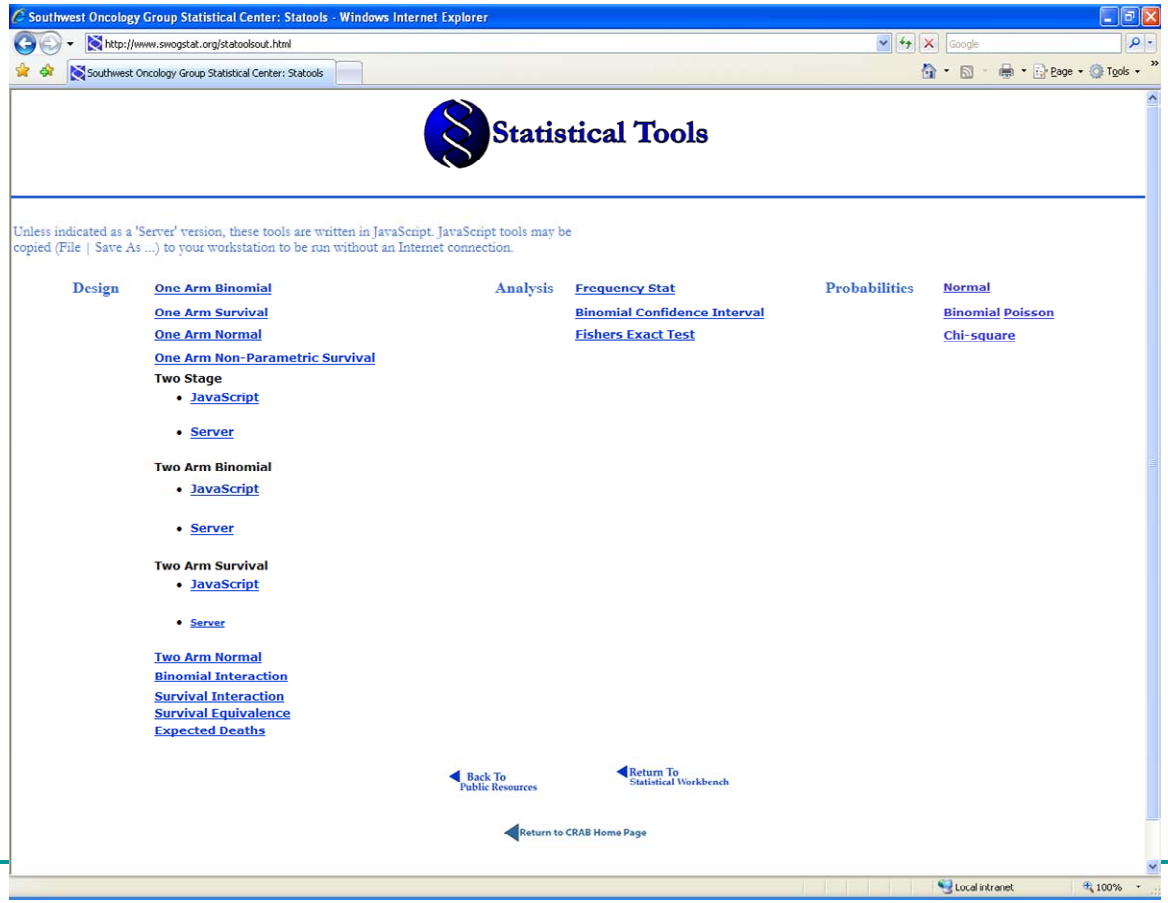
Sample size programs

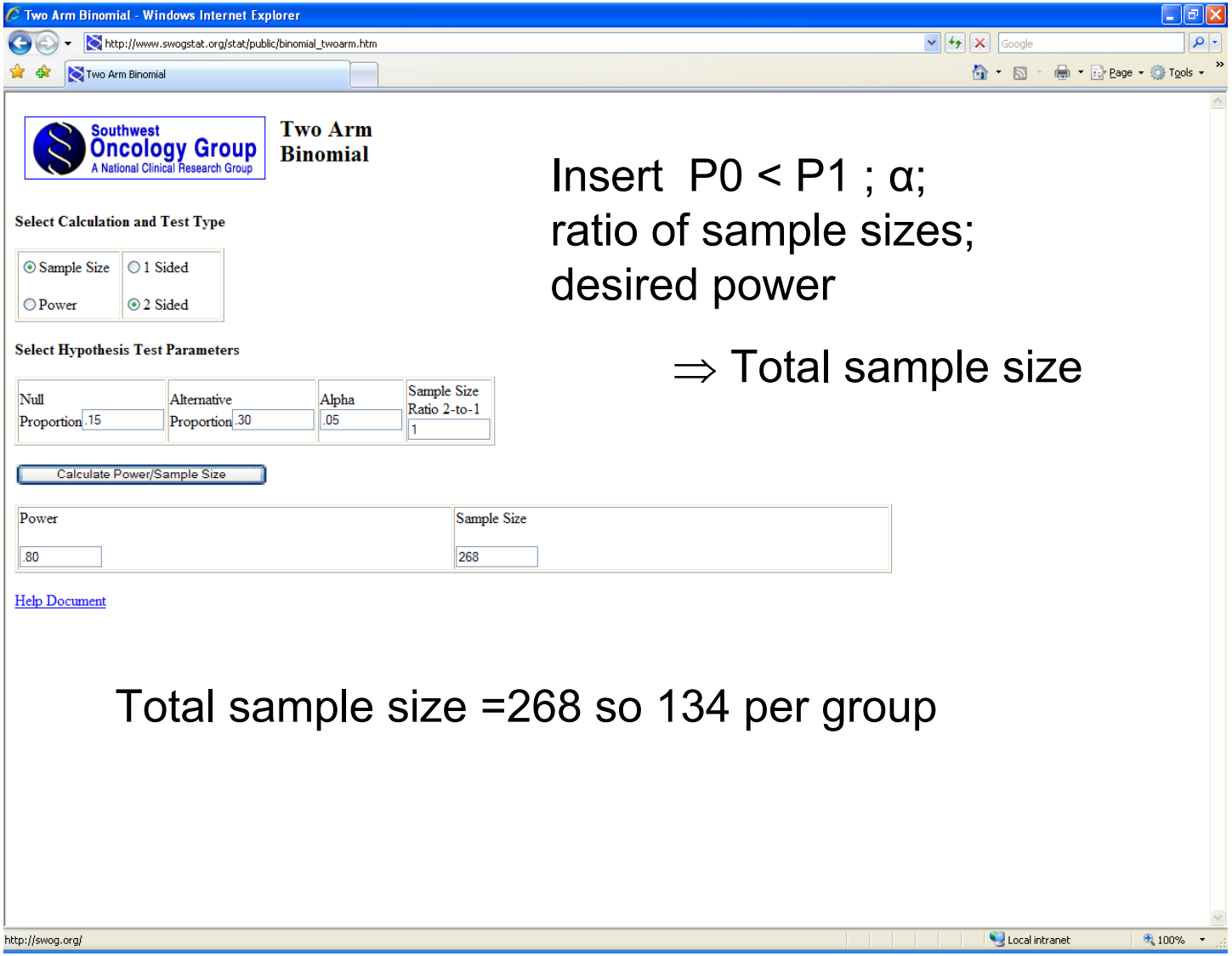
NQuery

NPASS

SIZ (available in many of the computer labs)

Public programs http://www.swogstat.org/statoolsout.html

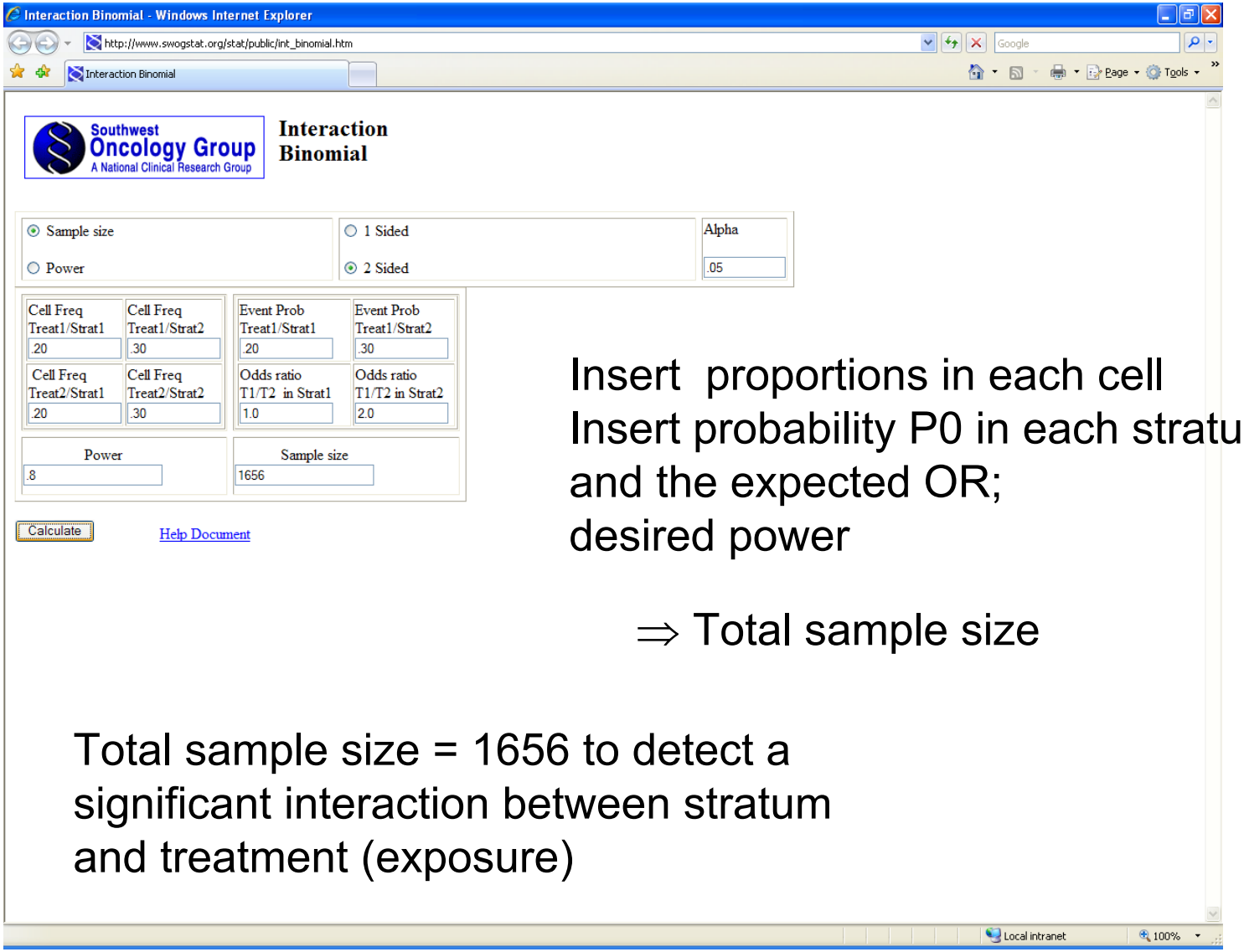




Insert P0 < P1 ; α;
ratio of sample sizes;
desired power

$\Rightarrow$ Total sample size

Total sample size =268 so 134 per group



Insert proportions in each cell
Insert probability P0 in each stratum
and the expected OR;
desired power

$\Rightarrow$ Total sample size

Total sample size = 1656 to detect a significant interaction between stratum and treatment (exposure)

Stata program *sampsi*

- **Used for 2-sample or 1-sample comparisons (continuous or dichotomous data)**
- **Can compute power from fixed sample size or vice-versa**

Calculate the sample size in each group comparing P1 = 30% to P2 = 15%

assuming 2-sided α = 0.05 and power 80%

```
. sampsi .3 .15, power(.8)
Estimated sample size for two-sample comparison of proportions
Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
         power =   0.8000
            p1 =   0.3000
            p2 =   0.1500
         n2/n1 =   1.00
Estimated required sample sizes:
            n1 =      134
            n2 =      134
```

Calculate the power assuming n1=100 and n2=200 assuming 2-sided α = 0.05

```
. sampsi .3 .15, n1(100) n2(200)
Estimated power for two-sample comparison of proportions
Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
            p1 =   0.3000
            p2 =   0.1500
sample size n1 =      100
            n2 =      200
         n2/n1 =     2.00
Estimated power:
         power =   0.8128
```

Calculate the sample size comparing P1 = 30% to P2 = 15% assuming 2-sided α = 0.05 and power 80% if we use a 2:1 allocation ratio for n2/n1

```
. sampsi .3 .15, power(.8) ratio(2)

Estimated sample size for two-sample comparison of proportions
Test Ho: p1 = p2, where p1 is the proportion in population 1
                  and p2 is the proportion in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
         power =   0.8000
            p1 =   0.3000
            p2 =   0.1500
         n2/n1 =   2.00
Estimated required sample sizes:
            n1 =       97
            n2 =      194
```

Unbalanced allocation increases the total sample size

Now assume that we are in fact intending to randomize patients within clinics and are worried about the correlation of patients within the same clinic

Install the add-on Stata ado program *sampclus*

package sxd4 from http://www.stata.com/stb/stb60

TITLE

STB-60 sxd4. Sample size estimation for cluster designed samples

DESCRIPTION/AUTHOR(S)

STB insert by Joanne M. Garrett, University of North Carolina

Support: joanne_garrett@med.unc.edu

After installation, see help sampclus

INSTALLATION FILES (click here to install)

sxd4/sampclus.ado

sxd4/sampclus.hlp

Run the sample size program first

```
. sampsi .3 .15, power(.8)

Estimated sample size for two-sample comparison of proportions
Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:
         alpha =   0.0500  (two-sided)
         power =   0.8000
            p1 =   0.3000
            p2 =   0.1500
         n2/n1 =   1.00
Estimated required sample sizes:
            n1 =      134
            n2 =      134
```

If there is no correlation should not affect the sample sizes

Rho = expected intraclass correlation; numclus = number of clusters;

obsclus = number of observations per cluster

In the example assume we expect to have 10 clinics

```
. sampclus , numclus(10) rho(0)

Sample Size Adjusted for Cluster Design
   n1 (uncorrected) = 134
   n2 (uncorrected) = 134

   Intraclass correlation     = 0
   Average obs. per cluster   = 27
   Minimum number of clusters = 10

   Estimated sample size per group:
      n1 (corrected) = 134
      n2 (corrected) = 134
```

Now assume a correlation of 0.05 between two patients in the same clinic

```
. sampclus , numclus(10) rho(0.05)

For this rho, the minimum number of clusters possible is: 14
```

Increase the number of clinics to 14

```
. sampclus , numclus(14) rho(0.05)

Sample Size Adjusted for Cluster Design
   n1 (uncorrected) = 134
   n2 (uncorrected) = 134

   Intraclass correlation     = .05

   Average obs. per cluster   = 425
   Minimum number of clusters = 14

   Estimated sample size per group:
      n1 (corrected) = 2975
      n2 (corrected) = 2975
```

Consider a different approach – if we take 20 patients per clinic

then how many clinics do we need?

```
. sampclus , obsclus(20) rho(0.05)

Sample Size Adjusted for Cluster Design
   n1 (uncorrected) = 134
   n2 (uncorrected) = 134

   Intraclass correlation     = .05

   Average obs. per cluster   = 20
   Minimum number of clusters = 27

   Estimated sample size per group:
      n1 (corrected) = 262
      n2 (corrected) = 262
```

We would need 27 clinics and 524 patients

- **Small correlations can have huge impact**
- **Maximize the number of clusters**
- **Minimize the number of subjects per cluster**
- **Accounting for factors that make patients correlated within a cluster can reduce the intraclass correlation**

Matched case-control designs

Another add-on program from Stata is available

---
help for sampsi_mcc

Calculates Sample Size or Power for Matched Case-Control Studies

sampsi_mcc [, p0(#) alt(#) n1(#) m(#) phi(#) alpha(#) power(#)
solve(string) }]

Description

sampsi_mcc calculates the power and sample size for a matched case control study. The theory behind this command is described in Dupont (1988) Power Calculations for Matched Case-Control Studies, Biometrics. The calculations require the usual alpha and beta values, a possible alternative odds ratio (the null is 1), phi the correlation of exposure between pairs in the case-control set (the default is 0.2) and the probability of exposure in the controls. This command can be combined with samplesize in order to look at multiple calculations and to plot the results.

Updating this command

To obtain the latest version click the following to uninstall the old version

ssc uninstall sampsi_mcc

And click here to install the new version

ssc install sampsi_mcc

Options

p0(#) specifies the exposure probability of controls; default is p0(0.5).
alt(#) specifies the "alternative OR".
n1(#) specifies the number of cases; default is n1(100).
m(#) specifies the number of matched controls per case; default is m(1).
alpha(#) significance level of test; default is a(0.05).
power(#) power of test; default is p(0.9).
solve(string) specifies whether to solve for the sample size or power; default is s(n) solves for n and the only other choice is s(power) solves for power.
phi(#) specifies the correlation of exposure between pairs in the case-control set. The paper recommends the default value 0.2 when it is unknown.

Author: Adrian Mander, MRC Human Nutrition Research, Cambridge, UK.

Reference: Dupont W.D. (1988) Power calculations for matched case-control studies. Biometrics 44: 1157-1168.

Consider a case-control study with 100 cases, exposure probability 50% in the controls and target odds ratio of 1.50 – get the power

```
. sampsi_mcc, p0(0.5) alt(1.5) n(100) s(power)

Estimate power for Matched Case Control Study
Test Ho: Odds ratio=1 Ha: Odds ratio = alt. OR
Assumptions:
                Alpha =    0.0500
(number of controls)M =    1.0000
  Prob. Exp. Controls =    0.5000
              Alt OR =    1.5000
                   N =  100.0000
Estimated Power:
       Power = .24054668
```

Only 24% power – how many cases would we need for 90% power (default value)?

```
. sampsi_mcc, p0(0.5) alt(1.5) s(n)

Estimate Sample Size for Matched Case Control Study
Test Ho: Odds ratio=1 Ha: Odds ratio = alt. OR
Assumptions:
                Alpha =    0.0500
(number of controls)M =    1.0000
  Prob. Exp. Controls =    0.5000
            Alt OR    =    1.5000
                Power =    0.9000
Estimated Number of Cases:
       N = 644
```

Try 80% power

```
. sampsi_mcc, p0(0.5) alt(1.5) s(n) power(0.8)
Estimate Sample Size for Matched Case Control Study
Test Ho: Odds ratio=1 Ha: Odds ratio = alt. OR
Assumptions:
                Alpha =    0.0500
(number of controls)M =    1.0000
  Prob. Exp. Controls =    0.5000
            Alt OR    =    1.5000
                Power =    0.8000
Estimated Number of Cases:
       N = 483
```

If we only had 100 cases could we increase the controls to 5-1 and compute the power again

```
. sampsi_mcc, p0(0.5) alt(1.5) n(100) s(power) m(5)
Estimate power for Matched Case Control Study
Test Ho: Odds ratio=1 Ha: Odds ratio = alt. OR
Assumptions:
                Alpha =    0.0500
(number of controls)M =    5.0000
  Prob. Exp. Controls =    0.5000
               Alt OR =    1.5000
                    N =  100.0000
Estimated Power:
       Power = .41141439
```

Still only 41% - how many cases would we need for 90% power at 5-1

```
. sampsi_mcc, p0(0.5) alt(1.5) s(n) m(5)
Estimate Sample Size for Matched Case Control Study
Test Ho: Odds ratio=1 Ha: Odds ratio = alt. OR
Assumptions:
                Alpha =    0.0500
(number of controls)M =    5.0000
  Prob. Exp. Controls =    0.5000
            Alt OR    =    1.5000
                Power =    0.9000

Estimated Number of Cases:
       N = 343
```

Sensitivity analysis – what it the exposure rate in controls is 30% not 50%

```
. sampsi_mcc, p0(0.3) alt(1.5) s(n) m(5)
Estimate Sample Size for Matched Case Control Study
Test Ho: Odds ratio=1 Ha: Odds ratio = alt. OR
Assumptions:
                Alpha =    0.0500
(number of controls)M =    5.0000
  Prob. Exp. Controls =    0.3000
            Alt OR    =    1.5000
                Power =    0.9000
Estimated Number of Cases:
       N = 376
```

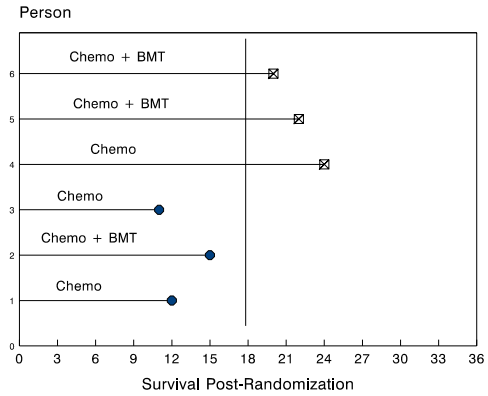
Quick Review

Binary outcomes

- **Use logistic regression where it naturally applies**

⇒ **Avoid dichotimizing a continuous variable**

⇒ **In a cohort study, time to failure may be a better outcome variable (take Biost/Epi 537)**



Study designs

- **Prospective design (exposure groups may be fixed by design)**

  Cohort studies may yield risk P(Y=1 | X=x) and relative risk estimates

- **Retrospective design (case status fixed by design)**

  Case-control studies yield odds ratio estimates

However, both use the same underlying logistic model

$$\log\left(\frac{p_i}{1-p_i}\right) = \text{logit}(p_i) = \alpha + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$$

Under the rare disease assumption interpretation of β is the same, but interpretation of α depends on the sampling design

Odds ratio for a unit change in x_j is $e^{\beta j}$ when other x's are held constant

Adding a constant to x_j does not change $e^{\beta j}$

Multiplying x_j by a constant *c* changes the odds ratio to exp(β_j / c)

Neither manipulation will change the hypothesis test of β_j

Test of $H_0: \beta_p = 0$ given $x_1, \ldots, x_{p-1}$

1. **Wald test uses asymptotic normality of** $\hat{\beta}_p$

$$z = \frac{\hat{\beta}_p}{se(\hat{\beta}_p)} \quad \textbf{or equivalently} \quad \chi^2(1\text{ df}) = z^2 = \left(\frac{\hat{\beta}_p}{se(\hat{\beta}_p)}\right)^2$$

2. **Likelihood ratio test (most reliable)**

Also follows a $\chi^2(1\text{df})$ **for a single covariate**

Controlling for confounders

Unconditional logistic regression

- **continuous confounder**

$$\text{logit}(p_i) = \alpha + \tilde{\beta}\, x_i$$

$$\text{logit}(p_i) = \alpha + \beta\, x_i + \gamma\, z_i$$

Compare estimates of $\tilde{\beta}$ and β

- **effect modifier**

$$\text{logit}(p_i) = \alpha + \beta x_i + \gamma\, z_i + \theta\, x_i z_i$$

Test $H_0: \theta = 0$

- **explicit stratification**

$$\text{logit}(p_i) = \alpha_j + \beta x_i$$

for person *i* in stratum *j* =1 , ... , J

Do not need to test the J α parameters

Can compare estimate of β with and without stratification

Use only if J << # of observations

Conditional logistic regression

- **unmeasured confounder embodied by an observed relationship (sibling, twin , etc)**
- ***1 : m* matching on confounders (typically prior to collection of exposure information)**
- **post-hoc stratification (implicit)**

  Underlying model is $\text{logit}(p_i) = \alpha_j + \beta x_i$

  for person *i* in stratum *j* =1 , ... , J

  Do not need to estimate the J α parameters

  Can compare estimate of β with that of explicit stratification

- **propensity score matching**

  Model the probability of exposure as a function of potential confounders

  Create strata on fine matching of probability of exposure

  Underlying model is $\text{logit}(p_i) = \alpha_j + \beta x_i$

  for person *i* in propensity score stratum *j* =1 , ... , J

  Can now fit a simple conditional logistic regression model linking outcome to exposure adjusted for the propensity to be exposed

Random effects logistic regression

- **Underlying model is** $\text{logit}(p_i) = \alpha_0 + \sigma u_j + \beta x_i$

  where u_j comes from some distribution with mean zero and standard deviation one

  then $\alpha_j = \alpha_0 + \sigma u_j$ **but we do not estimate the u_j's only σ**

  - **Model is a compromise between the unconditional and conditional logistic regression models**

Model diagnostics

Unconditional logistic regression

- **Fitted values**

$$\hat{p}_i = \frac{e^{\hat{\beta}_0 + x_{i1}\hat{\beta}_1 + \ldots + x_{ip}\hat{\beta}_p}}{1 + e^{\hat{\beta}_0 + x_{i1}\hat{\beta}_1 + \ldots + x_{ip}\hat{\beta}_p}}$$

- **Delta-betas**
- **Can use Pearson or Hosmer-Lemeshow tests to check model fit**

Conditional logistic regression

- **Fitted values (1 to *m* matching)**

$$\hat{p}_i = \frac{e^{x_{i1}\hat{\beta}_1 + \ldots + x_{ip}\hat{\beta}_p}}{\sum_{j=0}^{m} e^{x_{j1}\hat{\beta}_1 + \ldots + x_{jp}\hat{\beta}_p}}$$

for i = 0 , 1 , 2, ... , m

- **Delta-betas**
- **More difficult to check goodness-of-fit**
- **Can still use likelihood ratio and Wald tests**

Clustered or longitudinal data

Random effects logistic regression

Fit a term corresponding to subject (longitudinal) or cluster as a random effect

$$\text{logit}(p_i) = \alpha_0 + \sigma u_j + \beta x_i$$

Test $H_0: \sigma = 0$ **versus** $H_1: \sigma > 0$

Generalized estimating equations

Person *i* measured at several time points *j* = 1 , ..., J

Fit a term corresponding to subject *i* at time *j*

$$\text{logit}(p_{ij}) = \alpha + \beta x_{ij}$$

Account for the correlation within a person by

1. **Using a "working correlation" matrix**
2. **Using a robust variance estimate to get the standard errors**

Wald test used to evaluate significance

Multinomial models

For nominal outcomes we have separate models comparing each outcome to some referent outcome

For ordinal outcomes we have three basic approaches:

1. **Adjacent categories: Model the relationship between two adjacent categories but assume the covariates operate in the same manner between any two adjacent categories (special case of multinomial)**
2. **Continuation ratio model: Compare each category to all categories below that category**
3. **Proportional odds model: Dichotomize at each outcome and fit a model assuming a constant covariate relationship for each dichotimization**

ROC models are actually ordinal regression models as well

Can use LR tests and usual machinery in multinomial/ordinal regression models

Good luck on the final and Happy Holidays !