# Power and Sample Size

Martin L. Lee, PhD

# Overview of Lecture

- Basic parameters of power/sample size calculations
- Formulae for different settings
- Software

# What we need to know: Study design

- Randomization: yes/no, type
- Matching
- Number of groups
- Measurement type (quantitative/qualitative/time-to-event-censored)
- Repeated measures
- Multivariate adjustment
- One vs. two-tailed

# What we need to know: Study goals

- Estimation versus hypothesis testing
- Superiority versus equivalence/lack of inferiority
- Δ: minimum clinical important difference (superiority); equivalence definition

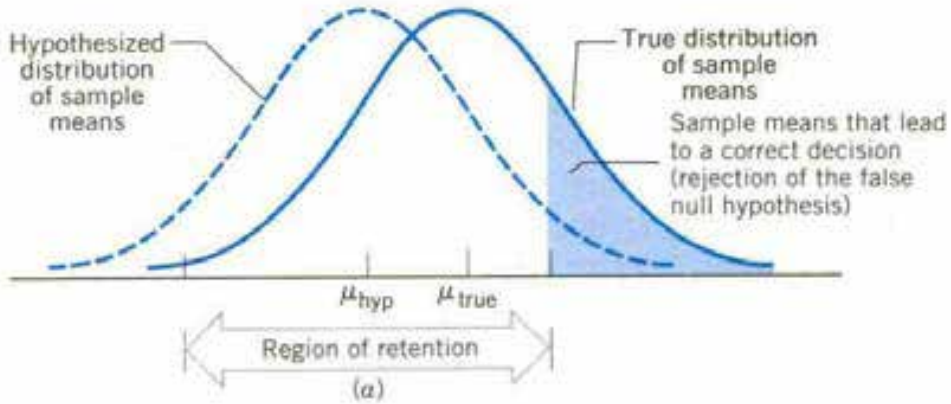
# Statistical errors in hypothesis testing

| | **Significance Difference is** | |
|---|---|---|
| | Present | Absent |
| | (Ho *not* true) | (Ho is true) |
| **Test result** | | |
| Reject Ho | No error | Type I err. |
| | 1-β | α |
| Accept Ho | Type II err. | No error |
| | β | 1-α |

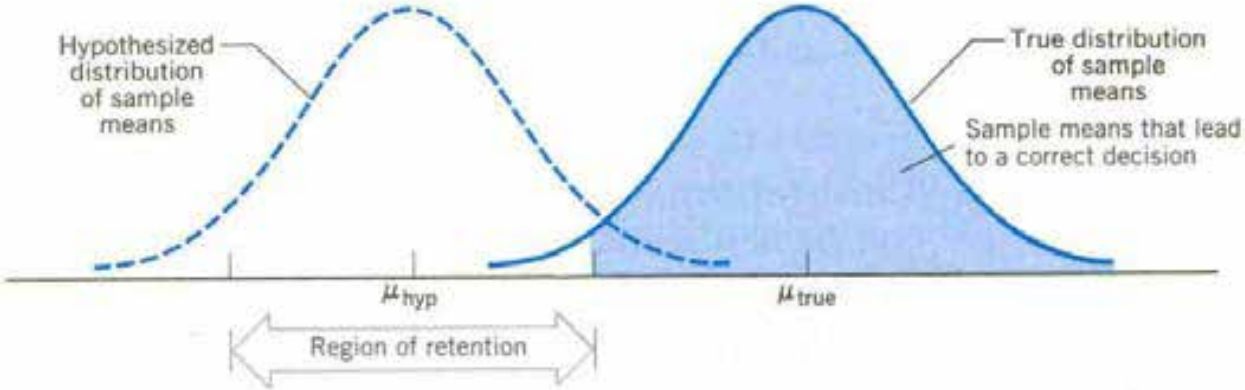
# Choosing error rates

- $\alpha$: “consumer” risk; would like to choose as small as possible, but impractical; 5% is a compromise
- $\beta$: “producer” risk; depends on true $\Delta$ (inverse relationship) and sample size

# The effect of Δ on β-error

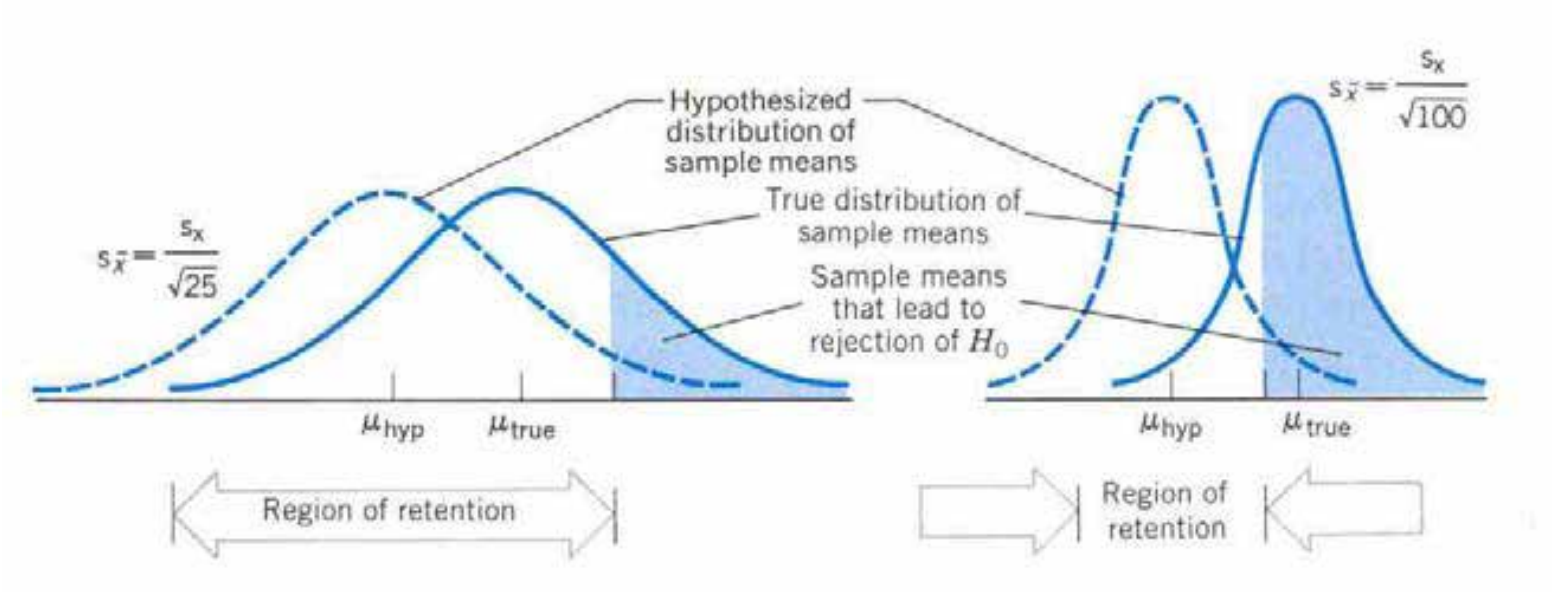



(a)




(b)

# The effect of n on β-error



# Information needed for sample size calculation

- Errors: $\alpha$ and $\beta$
- $\Delta$
- Quantitative data: estimate of SD (literature or “rule of 4”)
- Effect size: $\Delta$/SD or difference in proportions for qualitative data

# Sample size for difference between two means

- Hypotheses:

  $H_o$: $\mu_1 = \mu_2$ vs. $H_a$: $\mu_1 \neq \mu_2$
- Let $n_1$ and $n_2$ be the sample sizes for group 1 and 2, respectively; $N = n_1 + n_2$ ; $r = n_1 / n_2$ ; σ: standard deviation of the variable of interest.
- The sample size is adjusted slightly for the use of the t-test

- **If we let η = Δ / σ be the "effect size", then:**

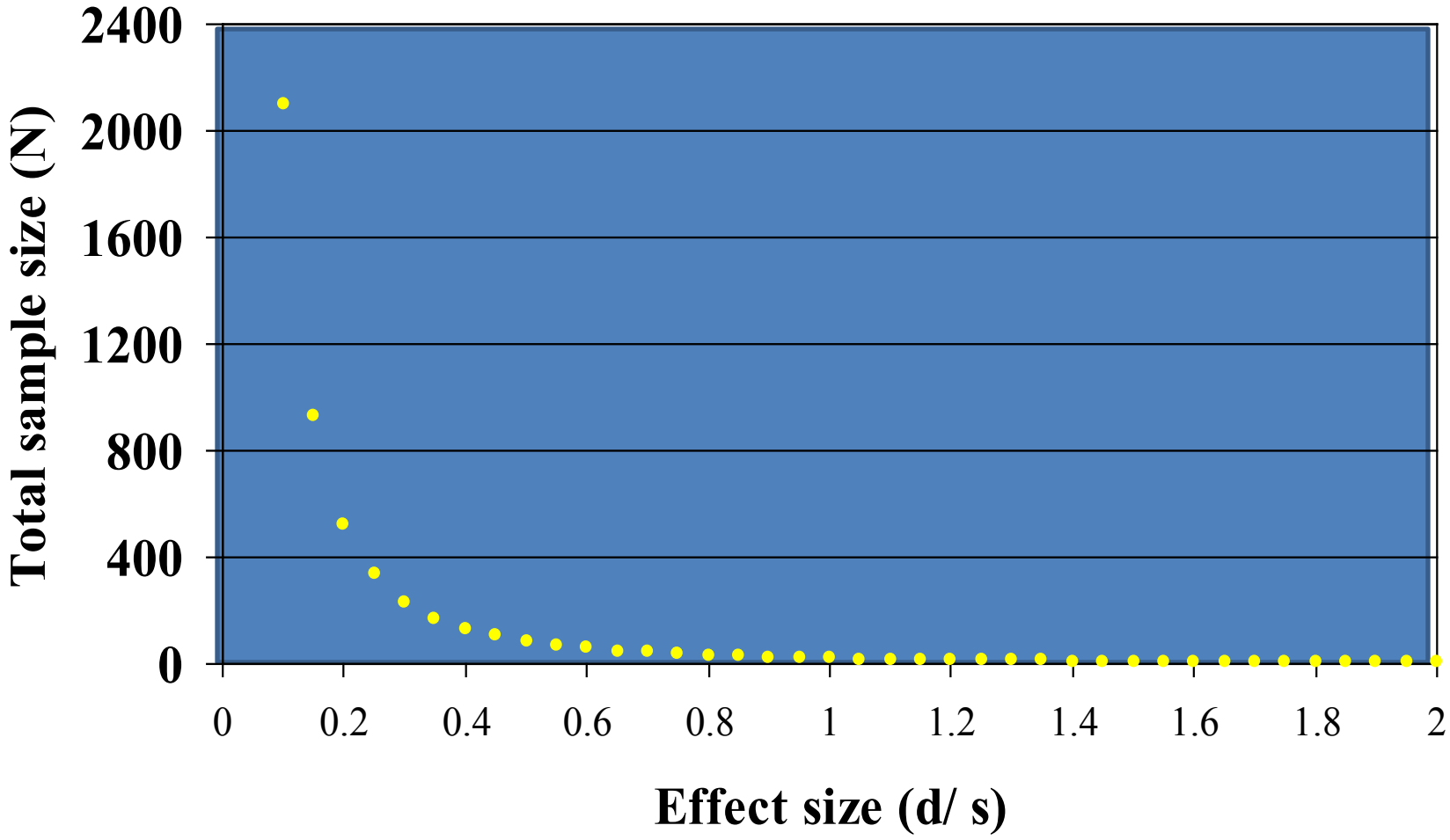
$$N = \frac{(r+1)\left(Z_{\alpha/2} + Z_{1-\beta}\right)^2 \sigma^2}{r\Delta^2}$$

where $Z_{\alpha/2}$ and $Z_{1-\beta}$ are normal deviates

- If $n_1 = n_2$, power = 0.80, alpha = 0.05, then $(Z_{\alpha/2} + Z_{1-\beta})^2 = (1.96 + 1.28)^2 = 10.5$, then the equation is reduced to:

$$N = \frac{21}{\eta^2}$$

# Sample size for 2 means versus effect size



# An example: Oral contraceptive use

- Study of OC use
- p=0.46
- Δ=5.4, which is large
- Power to detect Δ is 0.13

| | n | Mean BP | SD |
|---|---|---|---|
| OC user | 8 | 132.8 | 15.3 |
| Non-OC user | 21 | 127.4 | 18.2 |

# An example: OC use

- With the same $\Delta$ (about 5), average SD of about 16.7 and power of 80%, we need 154 subjects
- Notice $\delta$=0.3 standard deviations here, which is small-medium ES (Cohen, 1988)

ession Survival ROC Variances DOE Tools Window Help

MACRO | DIFFS 2 PROP | 2 S T-TEST | 1 WAY ANOVA | R M ANOVA | CNTRL M C-S | H R LRNK | LINEAR REG | LOGIST REG | HELP | PDF

Iterations | Template

Axes/Legend/Grid | Plot Text

**Effect Size**

Means

Mean1 (Mean of Group 1):

10

Mean2 (Mean of Group 2):

11, 12, 13

Standard Deviations

S1 (Standard Deviation Group 1):

2.5

S2 (Standard Deviation Group 2):

S1

☐ Known Standard Deviation

Standard Deviation Estimator

**Test**

Alternative Hypothesis:

Ha: Mean1 <> Mean2

Nonparametric Adjust. (Mann-Whitney Test):

Ignore

FIND (SOLVE FOR):
Select the parameter to be solved for in terms of the other parameters.

Note that this is the parameter displayed on the vertical axis of the plot.

Reset | Guide Me

# Output from PASS 2008: QOL scale example

| Power | N1 | N2 | Ratio | Alpha | Beta | Mean1 | Mean2 | S1 | S2 |
|---|---|---|---|---|---|---|---|---|---|
| 0.90148 | 133 | 133 | 1 | 0.050 | .0985 | 210.00 | 11.00 | 2.50 | 2.50 |
| 0.80365 | 100 | 100 | 1 | 0.050 | .1963 | 510.00 | 11.00 | 2.50 | 2.50 |
| 0.90150 | 34 | 34 | 1 | 0.050 | .0985 | 010.00 | 12.00 | 2.50 | 2.50 |
| 0.80749 | 26 | 26 | 1 | 0.050 | .1925 | 110.00 | 12.00 | 2.50 | 2.50 |
| 0.90719 | 16 | 16 | 1 | 0.050 | .0928 | 110.00 | 13.00 | 2.50 | 2.50 |
| 0.80208 | 12 | 12 | 1 | 0.050 | .1979 | 210.00 | 13.00 | 2.50 | 2.50 |

# Sample size for 2 means: hypothesis of equivalence

- $H_0$: $|\mu_1 - \mu_2| \geq \Delta$, where $\Delta$ is the definition of clinical bioequivalence

$$N = \frac{(r+1)\left(Z_{\alpha/2} + Z_{1-\beta}\right)^2 \sigma^2}{r(\delta - \Delta)^2}$$

where $\delta$=true difference

# Sample size for difference between 2 proportions

- Hypotheses:

  $H_o$: $p_1 = p_2$ vs. $H_a$: $p_1 = p_2 + d$ .

- Let $p_1$ and $p_2$ be the sample proportions (e.g. estimates of $p_1$ and $p_2$) for group 1 and group 2. Then, the sample size to test the hypothesis is:

$$n = \frac{\left(Z_{\alpha/2}\sqrt{2p(1-p)} + Z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right)^2}{(p_1-p_2)^2}$$

where: $n$ = sample size for <u>each</u> group ; $p = (p_1 + p_2) / 2$ ; $Z_{a/2}$ and $Z_{1-b}$ are normal deviates

# Sample size for 2 proportions: an example

- **Example**: The rate of adverse events for drug A is expected to be 70%, and for drug B 60%. A study is planned to show the difference at the significance level of 1% and power of 90%.
- The sample size can be calculated as follows:
  - $p_1 = 0.6$; $p_2 = 0.7$; $p = (0.6 + 0.7)/2 = 0.65$; $Z_{0.01} = 2.58$; $Z_{1-0.9} = 1.28$.
  - The sample size required for each group should be:

$$n = \frac{\left(2.58\sqrt{2 \times 0.65 \times 0.35} + 1.28\sqrt{0.6 \times 0.4 + 0.7 \times 0.3}\right)^2}{(0.6 - 0.7)^2} \approx 675$$

# Sample size for two proportions vs. effect size

| | Difference from $p_1$ by: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** |
| $P_1$ | | | | | | | | |
| 0.1 | 424 | 131 | 67 | 41 | 28 | 19 | 14 | 10 |
| 0.2 | 625 | 173 | 82 | 47 | 30 | 20 | 14 | 9 |
| 0.3 | 759 | 198 | 89 | 50 | 30 | 19 | 13 | 8 |
| 0.4 | 825 | 206 | 89 | 47 | 28 | 17 | . | . |
| 0.5 | 825 | 198 | 82 | 41 | 22 | . | . | . |
| 0.6 | 759 | 173 | 67 | 31 | . | . | . | . |
| 0.7 | 625 | 131 | 45 | . | . | . | . | . |
| 0.8 | 424 | 73 | . | . | . | . | . | . |

Note how sample size increases near the middle of the range of the baseline proportion

ortions [Proportions]

egression Survival ROC Variances DOE Tools Window Help

MACRO | DIFFS 2 PROP | 2 S T-TEST | 1 WAY ANOVA | R M ANOVA | CNTRL MC-S | H R LRNK | LINEAR REG | LOGIST REG | HELP | PDF

d | Options | Template

Axes/Legend/Grid | Plot Text

**Effect Size**

P1 (Treatment Group Proportion|H1):

.2, .3, .4

P2 (Control Group Proportion):

.1

Odds Ratio and Proportions Estimator

**Test**

Alternative Hypothesis (H1):

Two-Sided

Test Type:

Z Test (Pooled)

P2 (CONTROL GROUP PROPORTION):
Enter a value for the proportion in the control (baseline, standard, or reference) group, P2.

Values must be between 0 and 1.

You can enter a range of values such as
'0.1 0.2 0.3'
or
'0.1 to 0.5 by 0.1'.

Reset | Guide Me

# Output from PASS 2008: Remission rate example

| Power | Sample Size N1 | Sample Size N2 | Prop\|H1 Trtmnt P1 | Prop\|H1 Control P2 | Diff H0 D0 | Diff H1 D1 | Target Alpha | Actual Alpha | Beta |
|---|---|---|---|---|---|---|---|---|---|
| 0.9002 | 266 | 266 | 0.2 | 0.1 | 0 | 0.1 | 0.05 | | 0.0998 |
| 0.8001 | 199 | 199 | 0.2 | 0.1 | 0 | 0.1 | 0.05 | | 0.1999 |
| 0.9023 | 79 | 79 | 0.3 | 0.1 | 0 | 0.2 | 0.05 | 0.0516 | 0.0977 |
| 0.8051 | 59 | 59 | 0.3 | 0.1 | 0 | 0.2 | 0.05 | 0.0507 | 0.1949 |
| 0.9010 | 40 | 40 | 0.4 | 0.1 | 0 | 0.3 | 0.05 | 0.0587 | 0.0990 |
| 0.8126 | 31 | 31 | 0.4 | 0.1 | 0 | 0.3 | 0.05 | 0.0552 | 0.1874 |

# Sample size for difference between 2 proportions: Equivalence

$$n = \frac{\left(Z_{\alpha/2}\sqrt{2p(1-p)} + Z_{1-\beta}\sqrt{p_1(1-p_1) + p_2(1-p_2)}\right)^2}{(\delta - (p_1 - p_2))^2}$$

where, δ is the true difference between the proportions

# Sample size for two proportions when there are multiple controls per case

$$n = \frac{\left(Z_{\alpha/2}\sqrt{((c+1)/c)p(1-p)} + Z_{1-\beta}\sqrt{p_1(1-p_1)/c + p_2(1-p_2)}\right)^2}{(p_1 - p_2)^2}$$

where c = number of cases per control

and $p = (cp_1 + p_2)/(1+c)$

n = number of cases and c x n = number of controls

# Effect of adding more controls on power

| Controls Per Case | POWER |
|---|---|
| 1 | 76.2 |
| 2 | 86.5 |
| 3 | 89.9 |
| 4 | 91.5 |
| 5 | 92.4 |
| 6 | 93.0 |
| 7 | 93.4 |

Note that $\alpha=0.05$, power=0.80 and an odds ratio of 2.0

# Sample size for estimating an odds ratio

- In case-control study the data are usually summarized by an odds ratio (OR), rather then difference between two proportions.
- If $p_1$ and $p_2$ are the proportions of cases and controls, respectively, exposed to a risk factor, then:

$$OR = \frac{p_1(1-p_2)}{p_2(1-p_1)}$$

- If we know the proportion of exposure in the general population *(p)*, the <u>total</u> sample size *N* for estimating an OR is:

$$N = \frac{(1+r)^2 (Z_\alpha + Z_{1-\beta})^2}{r(\ln OR)^2 p(1-p)}$$

- Where $r = n_1 / n_2$ is the ratio of sample sizes for group 1 and group2; *p* is the prevalence of exposure in the controls; and OR is the hypothetical odds ratio. If $n1 = n_2$ (so that $r = 1$) then the fomula is reduced to:

$$N = \frac{4(Z_\alpha + Z_{1-\beta})^2}{(\ln OR)^2 p(1-p)}$$

# Sample size for an odds ratio: example

- **Example**: The prevalence of vertebral fracture in a population is 25%. It is interested to estimate the effect of smoking on the fracture, with an odds ratio of 2, at the significance level of 5% (one-sided test) and power of 80%.
- The total sample size for the study can be estimated by:

$$N = \frac{4(1.64 + 0.85)^2}{(\ln 2)^2 \times 0.25 \times 0.75} = 275$$

# Sample size for rare events

- Goal of study is to estimate the probability of a rare event
- Example: frequency of HIV after infusion of a blood product
- Hopefully, zero events
- What is the 95% upper bound for the probability, $\theta$ ?

# Sample size for rare events

- $\theta \leq 3/n$ ("Rule of 3's")
- Defines n for a desired upper bound

Example: How many patients do we need to study in order to ensure with 95% confidence that the rate of HIV transmission is less than 1%?

Answer: 300 by rule of 3's

# Software

- PASS 2008
- nQuery
- Internet calculators

tion Regression Survival ROC Variances DOE Tools Window Help

OUT MACRO DIFFS 2 PROP 2 S T-TEST 1 WAY ANOVA RM ANOVA CNTRL MC-S HR LRNK LINEAR REG LOGIST REG HELP PDF

kground | Iterations | Template

s | Axes/Legend/Grid | Plot Text

**Effect Size**

Baseline Probability

P0 (Baseline Probability that Y=1):

.50

Alternative Probability

Use P1 or Odds Ratio:

Odds Ratio

P1 (Alternative Probability that Y=1):

0.10 to 0.20 by 0.05

Odds Ratio (Odds1/Odds0):

3

Odds Ratio and Proportions Estimator

Covariates (X1 is the Variable of Interest)

R-Squared of X1 with Other X's:

0

X1 (Independent Variable of Interest):

Binary (X=0 or 1)

Percent of N with X1=1:

Reset all options to their default values.

Reset | Guide Me

n Regression Survival ROC Variances DOE Tools Window Help

OUT MACRO DIFFS 2 PROP 2 S T-TEST 1 WAY ANOVA RM ANOVA CNTRL MC-S HR LRNK LINEAR REG LOGIST REG HELP PDF

ground | Iterations | Template

Axes/Legend/Grid | Plot Text

**Effect Size**

C: Variables Controlled

Number of Independent Variables Controlled:

5

R2(C):

.1

T: Variables Tested

Number of Independent Variables Tested:

1

R2(T|C):

0.05

R2(T|C):
Enter the amount that is added to R2 by those independent variables that are to be tested after fitti
those that are controlled for. This can be a single value or a range of values.

The values must be between zero and one and the total of R2(C)+R2(T|C) must be less than one.

Reset | Guide Me

AT&T Webmail ... | 2 Microsoft O... | PASS 2008 Ho... | PASS: Logistic ... | PASS: Multiple ... | 1:

# Sample size for estimating the mean and proportion

- How close to the true mean
- Confidence around the sample mean

- $n = (Z_{\alpha/2})^2 s^2 / d^2$

  s: standard deviation

  $d$: the accuracy of estimate (how close to the true mean).

  $Z_{\alpha/2}$: normal deviate reflects the confidence level.

- How close to the true proportion
- Confidence around the sample proportion.

- $n = (Z_{\alpha/2})^2 p(1-p) / d^2$

  $p$: proportion to be estimated.

  $d$: the accuracy of estimate (how close to the true proportion).

  $Z_{\alpha/2}$: normal deviate reflects the confidence level.

# Examples of estimating the mean and proportion

- **Example**: we want to estimate the average weight in a population, and we want the error of estimation to be less than 2 kg of the true mean, with 95% confidence and we suppose σ=20
- $n = (1.96)^2\ 20^2 / 2^2 = 384$

- **Example**: The expected prevalence of a disease is around 20%. We want to estimate the true prevalence in a community within 5% with a 95% confidence interval.
- $n = (1.96)^2\ (0.2)(0.8) / 0.05^2 = 246$ individuals.