outline

1. Equivalence Studies
2. Missing Data
3. The Cross-Over Design
4. Clustered Binary Data
    - Hypothesis test
    - Regression Models: GEE model
5. Longitudinal Data Analysis

Equivalence Studies

Example
Cancer

- Suppose we want to design a clinical trial to compare two surgical treatments for early-stage breast cancer.
  - simple mastectomy (standard treatment, yields a 5-year survival rate of 80%)
  - a more conservative tumor resection (experimental treatment)

- Acceptable: if it can be shown to be no more than 10% inferior to the standard treatment in terms of 5-year survival.

- How can we test whether the experimental treatment is acceptable?

Equivalence study

- Remember
  - studies in which the null hypothesis is that two treatments are equally effective **vs.** the alternative hypothesis that the effects of the two treatments are different from each other
  - effectiveness in each treatment group is expressed as a binomial proportion.

- These types of studies, which constitute the majority of clinical trials, are referred to as *superiority/non-inferiority/equivalence studies*.

Superiority/Non-inferiority test

- p_1: the survival rate for the standard treatment
- p_2: the survival rate for the experimental treatment
- Superiority test
  - Higher means better, $H_1: p_2 > p_1 + \delta, \delta > 0$
  - Higher means worse, $H_1: p_2 < p_1 - \delta, \delta > 0$
- Non-inferiority test
  - Higher means better, $H_1: p_2 > p_1 - \delta, \delta > 0$
  - Higher means worse, $H_1: p_2 < p_1 + \delta, \delta > 0$

Inference Based on CI Estimation

- The cancer example: $p_1 = 0.8$ (standard)
- Non-inferiority test

$$H_1: p_2 > p_1 - 0.1 = 0.7$$

and $H_0: p_1 - p_2 \geq 0.1$

- Then $\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\hat{p}_1\hat{q}_1/n_1 + \hat{p}_2\hat{q}_2/n_2}} \sim N(0,1)$ and the rejection area is

$$\frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\hat{p}_1\hat{q}_1/n_1 + \hat{p}_2\hat{q}_2/n_2}} < -z_{1-\alpha}$$

- A lower one-sided 100% × (1 – α) CI for $p_1 - p_2$

$$p_1 - p_2 < \hat{p}_1 - \hat{p}_2 + z_{1-\alpha}\sqrt{\hat{p}_1\hat{q}_1/n_1 + \hat{p}_2\hat{q}_2/n_2}$$

- If the upper bound of this one-sided CI does not exceed δ, where δ is a pre-specified threshold.

Cancer
Non-inferiority Example

- Suppose we have a clinical trial with 100 patients each on the standard treatment and on the experimental treatment.
  - 80% of patients on the standard treatment
  - 75% of the patients on the experimental treatment

- Can the treatments be considered equivalent if the threshold is 10%?

- $p_1 - p_2 < 0.80 - 0.75 + z_{0.95}\sqrt{0.80 \times \frac{0.20}{100} + 0.75 \times \frac{0.25}{100}} = 0.147$

- The upper bound of the lower 95% CI exceeds 10%, so the treatments *cannot* be considered non-inferior.
  - although the observed survival rates are only 5% apart, the underlying rates may differ by as much as 15%

Sample-Size Estimation (Non-inferiority)

1. Assume the experimental treatment sample size (n_2) is k times as large as the standard treatment sample size (n_1)
   - $n_2 = kn_1$

1. Want to establish equivalence with a probability $1-\beta$. That is,

$$\Pr\left(\hat{p}_1 - \hat{p}_2 + z_{1-\alpha}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n}_1 + \frac{\hat{p}_2\hat{q}_2}{n_2}} \leq \delta\right) = 1-\beta$$

3. So we require
   - $n_1 = \frac{(p_1q_1 + p_2q_2/k)(z_{1-\alpha}+z_{1-\beta})^2}{[\delta-(p_1-p_2)]^2}$ subjects in group 1
   - $n_2 = kn_1$ subjects in group 2

Cancer Revisited

- Estimate the required sample size for the study if

I. we want a probability of 80% for establishing equivalence

– $1 - \beta = 0.80$

II. the sample sizes are the same in the two groups,

– $n_1 = n_2, k = 1$

III. the underlying 5-year survival rate in both groups is 80%

– $p_1 = \mathrm{p}_2 = 0.80, q_1 = q_2 = 0.20$

IV. the threshold for equivalence is 10%

– $\delta = 0.10$

$$n_1 = \frac{0.80 \times 0.20 \times 2(z_{0.95} + z_{0.80})^2}{0.10^2} = 197.6 = n_2$$

Missing data

Multiple imputation

Introduction

- As you can see from your project data, most epidemiologic and clinical studies have missing (or incomplete) data, for many different reasons.

- Can not do multivariate analyses?
  - such as multiple regression or multiple logistic regression

- Easy solution:
  - Delete it and use the complete data to analysis.

- Work?
  - Small amount: little bias, so OK
  - Nontrivial ($\geq 10\%$) or nonrandom missing: large bias, so NO

Example
Aging

- Determine the longitudinal course of aging and to identify risk factors that affect subsequent mortality

- A multiple logistic regression analysis
  - 2341 elderly participants ages 71–103 in 1988–1989
  - predict mortality through 1991

- predictor variables:
  - $x_1 = \text{age(yrs)}$
  - $x_2 = \text{sex(coded as 1 if male and 2 if female)}$
  - $x_3 = \text{physical–performance score (a 13-item scale from 0 to 12)}$
  - $x_4 = \text{selfassessed health score (a scale from 1 to 4: excellent to poor)}$

missing for 550 elderly participants who were either unwilling or unable to perform the test.

Descriptive statistics

Aging

Descriptive statistics for 2341 older residents of East Boston interviewed in 1988–1989

	Completed the physical-performance evaluation			
	Yes		no	
Dead by 12/31/1991	No	Yes	No	Yes
n	1527	264	416	134
Age, median (IQR)[a]	77 (74-81)	78 (74–84)	78 (74–83)	85 (78–90)
Male, %	32.0	45.1	30.3	47.0
Physical performance median (IQR)	8 (5-10)	6 (2-8)	---	---
Self-assessed health median (IQR)	2 (2–3)	3 (2–3)	2 (2–3)	3 (2–3)

Analysis with complete data

Aging

the logistic-regression analyses for the 1791 elderly participants with complete data

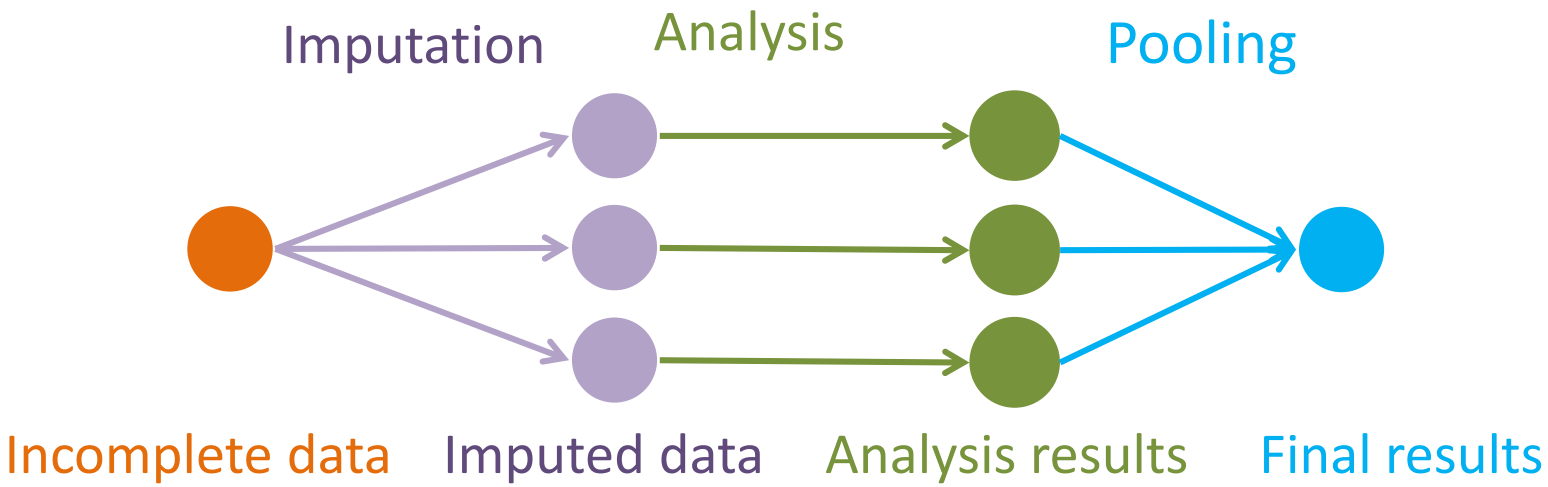
	$\widehat{\boldsymbol{\beta}}$	$\boldsymbol{sd}(\widehat{\boldsymbol{\beta}})$	z value	p-value
Intercept	-4.76	1.17		
Age	0.033	0.013	2.538	0.011
Male	0.92	0.15	6.133	<0.001
Self-assessed health	0.38	0.095	4.000	<0.001
Physical performance	-0.14	0.023	-6.087	<0.001

But, whether excluding these people affected the estimates of the regression parameters?

If yes, how to incorporate participants with missing data into the analysis?

Multiple imputation

- Imputation is defined roughly as the estimation of a missing variable (or variables) as a function of other covariates that are present, including the outcome.
- Multiple imputation is a Monte Carlo technique in which the missing values are replaced by *m*>1 simulated versions, where *m* is typically small (e.g. 3-10, m=3 in the figure).



Procedure(I)

1. Covariates: $x_1, \ldots, x_k$
   - $x_1, \ldots, x_{k-1}$ are complete (n subjects);
   - x_k is available for N_{obs} subjects and is missing for N_{mis} subjects

   A binary outcome variable: y

2. Run a multiple regression analysis of x_k on $x_1, \ldots, x_{k-1}$ and y based on the complete data

$$x_k = \alpha + \gamma_1 x_1 + \cdots + \gamma_{k-1} x_{k-1} + \delta y + \varepsilon$$

3. For $i = 1, \ldots, N_{mis}$ subjects with missing data on x_k, calculate an estimated value for x_k, denoted by $x_{i,k,1}$ for ith subject

$$\widehat{x_k} = \hat{\alpha} + \hat{\gamma_1} x_1 + \cdots + \widehat{\gamma_{k-1}} x_{k-1} + \hat{\delta} y$$

- Two source of bias compiled in $\widehat{x_k}$: estimation bias of $\hat{\alpha}, \hat{\gamma_1}, \ldots, \widehat{\gamma_{k-1}}, \hat{\delta}$, and the residual $e = x_k - \widehat{x_k}$)
- Estimated value $x_{i,k,1} = \xi(\hat{\alpha}) + \xi(\hat{\gamma_1}) x_1 + \cdots + \xi(\widehat{\gamma_{k-1}}) x_{k-1} + \xi(\hat{\delta}) y + \xi \cdot MSE$, where $\xi(X)$ indicates a monte carlo sample representing X and ξ represents N(0,1)

Procedure(II)

4. Run a logistic regression

$$\ln \frac{p}{1-p} = \alpha + \beta_1 x_1 + \cdots + \beta_k x_k$$

5. Repeat steps 3 and 4 for $m - 1$ additional imputations

6. The estimates from the *m* separate imputations are then combined into an overall estimate for β_j

$$\hat{\beta}_j = \sum_{q=1}^{m} \hat{\beta}_{j,q} / m$$

and an overall variance given by

$$Var(\hat{\beta}_j) = \sum_{q=1}^{m} \frac{Var(\hat{\beta}_{j,q})}{m} + \left[\frac{(m+1)}{m}\right] \sum_{q=1}^{m} \frac{(\hat{\beta}_{j,q} - \hat{\beta}_j)^2}{m-1}$$

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

Revisited

Aging

Summary of linear-regression model predicting physical-performance score, and 5 draws of these regression parameters

Variable	Estimate (SE)	5 Drawn values of parameters				
		1	2	3	4	5
Age (per year)	−0.22 (0.013)	−0.23	−0.23	−0.23	−0.21	−0.22
Male gender	1.18 (0.15)	1.21	1.08	1.29	1.25	1.22
Self-assessed health	−1.38 (0.089)	−1.33	−1.43	−1.24	−1.43	−1.35
Dead by 1991	−1.30 (0.20)	−1.12	−1.47	−1.37	−1.60	−1.05
Intercept	27.2 (1.1)	27.6	27.7	27.5	26.3	27.2

Overall R^2 0.30; Root MSE 2.88

Table 13.42 Comparison of effects of variables on the risk of death from alternative models; shown are logistic-regression parameters (standard errors)

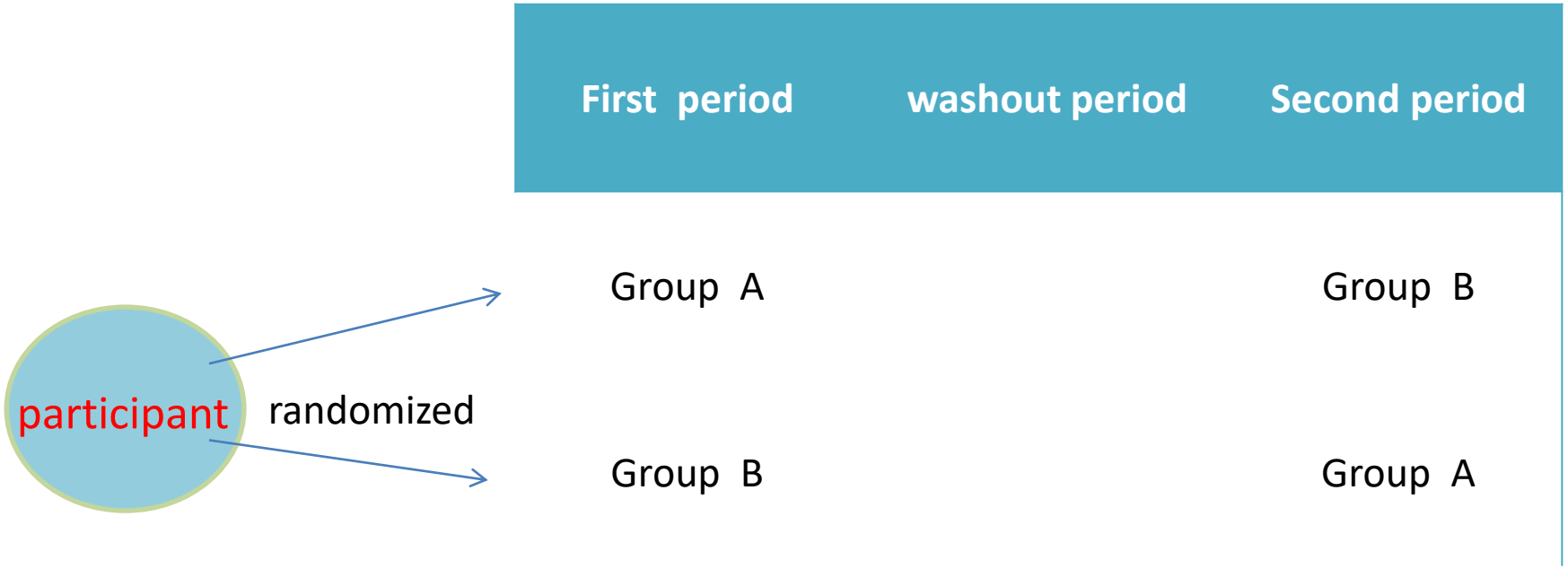
	Analytic method			
	Complete case	No physical performance	Multiple imputation	Indicator method[a]
n (deaths)	1791 (264)	2341 (398)	2341 (398)	2341 (398)
Age	0.033 (0.013)	0.088 (0.009)	0.057 (0.011)	0.068 (0.01)
Male	0.92 (0.15)	0.82 (0.12)	1.00 (0.14)	0.92 (0.12)
Self-assessed health	0.38 (0.095)	0.60 (0.073)	0.39 (0.087)	0.46 (0.076)
Physical performance	−0.14 (0.023)	—	−0.15 (0.024)	−0.12(0.022)
Intercept	−4.76 (1.17)	−10.41 (0.79)	−6.60 (1.04)	−7.42 (0.92)
Indicator of missing performance				−0.47 (0.13)

[a] The indicator method assigns the average performance score (6.8) to those with missing values and includes an indicator variable for this group.

The Cross-Over Design

What is a cross-over design?

- A cross-over design is a type of randomized clinical trial



Some definition

- A washout period in a cross-over design is a period between active drug periods, during which subjects receive no study medication.

- A carry-over effect in a cross-over design is when the effects of one or both study medications taken during the first active drug period have a residual biological effect during the second active drug period.

Example
Sports Medicine

- a cross-over design: comparing Motrin vs. placebo for the treatment of tennis elbow.

- Each participant

1. was randomized to receive either Motrin (group A) or placebo (group B) for a 3-week period.
2. had a 2-week washout period during which they received no study medication.
3. was then "crossed-over" for a second 3-week period to receive the opposite study medication from that initially received.

- How should we compare the efficacy of Motrin vs. placebo?

Assessment of Overall Treatment Effects

- x_{i1}: score of the ith patient in group A in the first period
  x_{i2}: its score in group B in the second period
- y_{i1}: score of the ith patient going to group B in the first period
  y_{i2}: its score going to group B in the second period

- Paired –data method

$$d_{1\mathrm{i}} = x_{i1} - x_{i2},\ d_{2\mathrm{i}} = y_{i2} - y_{i1}$$

Then $(d_{1\mathrm{i}}, i = 1,2, \dots n_1;\ d_{2\mathrm{i}}, i = 1,2, \dots n_2) \sim F$ with mean Δ

$$\overline{d_1} = \sum_{i=1}^{n_1} d_{1i}\,/n_1,\ \overline{d_2} = \sum_{i=1}^{n_2} d_{2i}\,/n_2$$

$$\bar{d} = (\overline{d_1} + \overline{d_2})/2$$

$$H_0: \Delta = 0\ vs.\ H_1: \Delta \neq 0$$

(for the details, see page 668-669)

Assessment of Carry-Over Effects

- Assume that $\bar{x}_i = (x_{i1} + x_{i2})/2 \sim N(\mu_1, \sigma^2)$ and $\bar{y}_i = (y_{i1} + y_{i2})/2 \sim N(\mu_2, \sigma^2)$
- To test $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$
- The test statistic

$$t = \frac{\bar{\bar{x}} - \bar{\bar{y}}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- Where

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{n_1}, \bar{\bar{y}} = \frac{\sum \bar{y}_i}{n_2},$$

and

$$s^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2},$$

$$s_1^2 = \sum_{i=1}^{n_1}(\bar{x}_i - \bar{\bar{x}})^2/(n_1 - 1), s_2^2 = \sum_{i=1}^{n_2}(\bar{y}_i - \bar{\bar{y}})^2/(n_2 - 1)$$

Discussion

- What can we do if we identify a significant carry-over effect?
- In this case, the second-period data are not useful to us because they provide a biased estimate of treatment effects
- We can use an ordinary two-sample t test for independent samples based on the first-period data.
- This test usually has less power than the cross-over efficacy test, or requires a greater sample size to achieve a given level of power

Sample-Size Estimation

- A major advantage of cross-over studies is that they usually require many fewer subjects than the usual randomized clinical trials (which have only 1 period), if no carry-over effect is present.

- Use the cross-over efficacy test given before, we have

$$n = \frac{\sigma_d^2 \left(z_{1-\alpha/2} + z_{1-\beta} \right)^2}{2\Delta^2}$$

- $\sigma_d^2 =$ variance of difference scores
- only applicable if no carry-over effects are present.

Clustered Binary Data

Correlated binary data

Example (Dependent samples)

- Rowe et al. [26] reported on a clinical trial of topically applied 3% vidarbine vs. placebo in treating recurrent herpes labialis.
  - the characteristics of 53 lesions observed on 31 patients receiving vidarbine
  - the characteristics of 69 lesions observed on 39 patients receiving placebo.

- A question of interest is whether the proportion of lesions showing significant shrinkage in the two groups is the same after 7 days.

- The two-sample test for the comparison of binomial proportions is not appropriate here because
  - dependencies in response among lesions observed on the same patient.

Notation

		# of lesions	# of success	Prob. estimate
Group 1	Patient 1	m_{11}	a_{11}	$\hat{p}_{11} = a_{11}/m_{11}$
	2	m_{12}	a_{12}	$\hat{p}_{12} = a_{12}/m_{12}$
	…	…		…
	n_1	m_{1n_1}	a_{1n_1}	$\hat{p}_{1n_1} = a_{1n_1}/m_{1n_1}$
		M_1	A_1	$\hat{p}_1 = A_1/M_1 = \sum_{j=1}^{n_i} m_{1j}\hat{p}_{1j}/M_1$
Group 2	Patient 1	m_{11}	a_{11}	$\hat{p}_{11} = a_{11}/m_{11}$
	2	m_{12}	a_{12}	$\hat{p}_{12} = a_{12}/m_{12}$
	…	…		…
	n_1	m_{1n_1}	a_{1n_1}	$\hat{p}_{1n_1} = a_{1n_1}/m_{1n_1}$
		M_2	A_2	$\hat{p}_2 = A_2/M_2$

$N = n_1 + n_2, M = M_1 + M_2$

Intraclass correlation coefficient

- An estimate of the degree of clustering within individuals/patients is given by the intra-class correlation ρ_I for clustered binary data.

- Definition: the correlation ρ_I between two replicates from the same subject, i.e., between y_{ij} and y_{il}, where y_{ij} is the j-th record of the i-th subject.
- If y_{ij} follows a one-way random-effect ANOVA model,

$$\rho_I = \frac{{\sigma_A}^2}{{\sigma_A}^2 + \sigma^2}$$

- It can be estimated by

$$\hat{\rho}_I = \max\left(\frac{\widehat{\sigma_A}^2}{\widehat{\sigma_A}^2 + \hat{\sigma}^2}, 0\right) = \max\left(\frac{MSB - MSW}{MSB + (n_0 - 1)MSW}, 0\right)$$

(refer to page 6 in lecture 8)

the degree of clustering

- Take each individual/patient as a group in the random-effect ANOVA model, where $y_{ijk} \sim B(1, p_{ij})$ indicates whether or not the k-th lesion shrinks of the j-th patient in the i-th group
- Then
  - Within SS=$\sum_{i=1}^{2}\sum_{j=1}^{n_1}\sum_{k=1}^{m_{ij}}(y_{ijk}-\hat{p}_{ij})^2 = \cdots = \sum_{i=1}^{2}\sum_{j=1}^{n_1} a_{ij}\left(1-\hat{p}_{ij}\right)$
  - Between SS= $\sum_{i=1}^{2}\{\sum_{j=1}^{n_1} m_{ij}\left(\hat{p}_{ij}-\hat{p}_i\right)^2\}$
- Let $MSB = Between\ SS/(N-2)$, $MSW = Within\ SS/(M-N)$
- The estimate of intra-class correlation is

$$\hat{\rho} = (MSB - MSW)/[MSB\ +\ (m_A\ -1)MSW]$$

where $m_A = \left[M - \sum_{i=1}^{2}\left(\sum_{j=1}^{n_i} m_{ij}^2/M_i\right)\right]/(N-2)$

design effect

- The clustering correction factor in group i may be defined as

$$C_i = \sum_{j=1}^{n_i} m_{ij} C_{ij} / M_i$$

where $C_{ij} = 1 + (m_{ij} - 1)\hat{\rho}$

design effect

- Note:
  - if the $\hat{\rho}$ =0, then no clustering is present and $C_i = 1,\ i = 1,2.$
  - If the $\hat{\rho}$ >0, then the $C_i > 1,\ i = 1,2$.

- The design effects in the two samples (C_1, C_2) are used as correction factors to modify the standard test statistic comparing two binomial proportions (Equation 10.3) for clustering effects.

Hypothesis Testing

- To test $H_0: p_1 = p_2$ vs. $H_1: p_1 \neq p_2$
- p_i: the underlying success rate among observations in group i

- The test statistic is

correction factors

$$z = \left[|\hat{p}_1 - \hat{p}_2| - \left(\frac{C_1}{2M_1} + \frac{C_2}{2M_2} \right) \right] / \sqrt{\hat{p}\hat{q}(C_1/M_1 + C_2/M_2)}$$

- Remember: Two-Sample Test for Binomial Proportions

$$z = \left[|\hat{p}_1 - \hat{p}_2| - \left(\frac{1}{2M_1} + \frac{1}{2M_2} \right) \right] / \sqrt{\hat{p}\hat{q}(1/M_1 + 1/M_2)}$$

- Only be used if $M_i \hat{p}\hat{q}/C_i \geq 0.5$ for $i = 1,2$.

Regression Models for Clustered Binary Data

- We have considered a comparison of two binomial proportions where the units of observation are not independent.
- However, we often would like to consider one or more additional covariates in our modeling.
- For this purpose, we wish to extend logistic regression methods to allow for correlation between subunits within the same cluster.
- A technique called generalized estimating equations (GEE) can perform this type of analysis

What is GEE?

- GEE (Liang and Zeger, 1986) provide a method of inference for a wide variety of models when responses are correlated
  - Linear regression (for continuous outcomes)
  - Logistic regression (for binary outcomes)
  - Poisson regression (for outcomes that are counts)
  - Proportional odds (for ordinal categorical outcomes)

Contd.

- GEE are an analysis method, not models in and of themselves
  - You specify a model that you'd like to fit using GEE
  - Model is specified through
    1. A link function that relates the mean response to the regression equation
       - "link = logit" for logistic regression
       - "link = log" for Poisson regression
    2. An assumed distribution for the response, although distributional assumptions not really strong
       - "binomial" for logistic regression
       - "poisson" for Poisson regression
    3. A working correlation matrix (more on this in the next slide)

Working correlation matrix

Exchabgeable

$$\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$$

m-dependent

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{pmatrix}$$

Unstructured

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_4 & \rho_5 \\ \rho_2 & \rho_4 & 1 & \rho_6 \\ \rho_3 & \rho_5 & \rho_6 & 1 \end{pmatrix}$$

Autoregressive

$$\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

GEE logistic model

- y_{ij}: outcome for the *j*th subunit in the *i*th cluster

$$= \begin{cases} 1, & \text{with probability } p_{ij} \\ 0, & \text{with probability } q_{ij} = 1 - p_{ij} \end{cases}$$

- $x_{ij1}, \ldots, x_{ijk}$ be a set of covariates for the *j*th subunit in the *i*th cluster.

- a model allows for the correlation between outcomes for multiple subunits in the same cluster.

$$\text{logit}(p_{ij}) = \ln \frac{p_{ij}}{1 - p_{ij}} = \alpha + \beta_1 x_{ij1} + \cdots + \beta_k x_{ijk}$$

- where $\text{corr}(p_{ij}, p_{ik}) = \rho$.

compound symmetry or exchangeable correlation structure

Longitudinal Data Analysis

Introduction

- An important application of clustered data methods is in longitudinal data analysis
  - each subject provides repeated measures over time
  - the goal: assess the effect of covariates on the rate of change over time.

- A clinical trial was performed among subjects with retinitis pigmentosa (RP) to compare the rate of decline of ERG (electro-retinogram) amplitude over time among 4 treatment groups.

- How should we compare the rate of decline among the 4 treatment groups?

Example

Ophthalmology

1. The ERG is an objective measure of the electrical activity in the retina.
   - In normals, the average ERG is about 350 µV.
   - In RP patients, it declines over time and is often <10 µV and sometimes <1 µV, after which total blindness often occurs.

2. Subjects were randomized to either
   - group 1 = 15,000 IU of vitamin A per day,
   - group 2 = 400 IU of vitamin E per day,
   - group 3 = 15,000 IU of vitamin A and 400 IU of vitamin E per day,
   - group 4 = placebo
3. were followed annually for 4−6 years.

Longitudinal Data Analysis

$$y_{it} = \alpha + \sum_{j=1}^{3} \beta_j x_{ij} + \gamma t + \sum_{j=1}^{3} \delta_j x_{ij} t + e_{it}$$

- $y_{it} = \ln(\text{ERG amplitude})$ for the *i*thsubject at time *t*
- $x_{ij} = I(\text{the ith subject is in treatment group } j)$

- ϵ_{it} = error term which is assumed to be normally distributed with mean = 0 and variance = σ^2
  - But over time are *not* assumed to be independent
  - exchangeable or compound symmetry correlation structure: $corr(e_{it_1}, e_{it_2}) = \rho \neq 0$
    - reasonable for the above relatively short-term clinical trial

Statistical methods

- Statistical methods for longitudinal data
  1. Univariate and Multivariate Analysis of Variance
  2. generalized estimation equation(GEE)
  3. Mixed-effect regression model(MRM)

- See more in
  1. http://www.uic.edu/classes/bstt/bstt513/
  2. http://tigger.uic.edu/~hedeker/long.html
  3. http://www.ats.ucla.edu/stat/r/examples/alda/ch2.htm