Medical Biometry III

(Biostatistics 513)

Instructor:

David Yanez

These notes are based on material developed by Norman Breslow, Patrick Heagerty, Mary Lou Thompson, Pat Wahl, Jim Hughes and others.

Introduction

Review

Biostat 511: *a bit of everything...*

- Data summaries (means, medians...)
- EDA (Exploratory data analysis)
- CDA (confirmatory data analysis):
  - hypothesis testing
  - p-values, statistical significance
  - confidence intervals
  - power and sample size
- 1-sample inference, 2-sample inference
  - means
  - proportions

Review

Biostat 512: *continuous response variables*

- Simple linear regression
  - transformation(s) (Y and/or X)
  - residuals
- Multiple regression
  - confounding
  - interaction (effect modification)
  - diagnostics
  - factors and dummy variables
- ANOVA, ANCOVA

Review

Biostat 513: *categorical response variables and (censored) time-to-event outcomes*

- Contingency tables
- 2 x 2 Tables
- Stratified methods (Mantel-Haenszel)
- Logistic Regression (binary data)
- Survival Data ("censored" data)

  - Kaplan-Meier curves

  - Cox proportional hazards model

Categorical Data

Part I

Overview

1) **Types of variables**
2) **Association between two categorical variables**
   - Contingency (two-way) tables
   - χ^2 test of homogeneity
   - χ^2 test of independence
   - Testing for trend in proportions
   - Using STATA
3) **2 x 2 Tables**
   - Sampling designs
   - Testing for association
   - Estimation of effects
   - Paired binary data
   - Small sample methods
     - Fisher's exact test
   - Using STATA

Overview

4) Stratified Tables

- Causality
- Confounding
- Effect modification
- Testing for a common OR
  - Mantel-Haenszel Test
- Estimation of a common OR
  - Mantel-Haenszel

5) Measures of Accuracy and Agreement

- Sensitivity & Specificity
- ROC curves
- Kappa statistics

Scales of Measurement

- Nominal
  - Order of categories irrelevant, e.g. gender, color, brand
- Ordinal
  - Order of categories meaningful, e.g. “better”, “same”, “worse”
- Interval (quantitative)
  - Arbitrary origin (0 point) and scale, e.g. Temperature (F vs C)
- Ratio (quantitative)
  - Fixed origin, arbitrary scale
  - Distance (miles vs kilometers), elapsed time

Factors and Contingency Tables

Definition: A **factor** is a categorical (discrete) variable taking a small number of values that represent the *levels* of the factor.

Factors may be *nominal*, *ordinal* or *quantitative*.

Quantitative factors often arise by *grouping* of continuous variables into categories

Examples

- Gender with two levels:

  1 = Male and 2 = Female

- Disease status with three levels:

  1 = Progression, 2 = Stable, 3 = Improved

- Age (categorized) with 4 levels:

  1 = 20-29 yrs, 2 = 30-39, 3 = 40-49, 4 = 50-59

Factors and Contingency Tables

Data description is facilitated by one-way, two-way or multi-way *tables of frequencies* of factor levels and their combinations

- To assess whether two factors are related, we often construct an R x C table that *cross-classifies* the observations according to the 2 factors.
- Examining two-way tables of Factor A vs Factor B at each level of a third Factor C shows how the A/B association may be *explained* or *modified* by C.

Tests: We can test whether the factors are related using a χ^2 test. Depending on the hypotheses / design we may use

1. χ^2 test of homogeneity
2. χ^2 test of independence
3. χ^2 test for trend in proportions

Categorical Data

Example 1: Education level versus willingness to participate , if the study were to start tomorrow, in a study of a vaccine to prevent HIV infection. (Cell counts, row percents and row totals are given.)

	definitely not	probably not	probably	definitely	Total
< high school	52 7.4%	79 11.3%	342 48.9%	226 32.3%	699
high school	62 6.9%	153 17.1%	417 46.6%	262 29.3%	894
some college	53 4.2%	213 16.8%	629 49.5%	375 29.5%	1270
college	54 4.9%	231 21.0%	571 51.9%	244 22.2%	1100
some post college	18 6.5%	46 16.6%	139 50.2%	74 26.7%	277
graduate/ prof	25 4.1%	139 22.8%	330 54.1%	116 19.0%	610
Total	264 5.4%	861 17.8%	2428 50.1%	1297 26.7%	4850

Q: Why might *row* percents be more appropriate than *column* percents here?

Categorical Data

Example 2: From Doll and Hill (1952) – study of British doctors. The table displays the retrospective daily average number of cigarettes smoked daily for lung cancer patients and controls.

	Daily # cigarettes						
	None	< 5	5-14	15-24	25-49	50+	Total
Cancer	7 0.5%	55 4.1%	489 36.0%	475 35.0%	293 21.6%	38 2.8%	1357
Control	61 4.5%	129 9.5%	570 42.0%	431 31.8%	154 11.3%	12 0.9%	1357
Total	68	184	1059	906	447	50	2714

Test of Homogeneity

In **Example 2** we want to test whether the smoking frequency is the same for each of the populations sampled, i.e. we want to test whether the two groups are **homogeneous** with respect to a characteristic, namely smoking. The concept is similar to a t-test, but the response is categorical.

H_0: smoking frequencies are the same in both groups

H_A: smoking frequencies are not the same

Q: What does H_0 predict we would observe if all we knew were the marginal totals?

	Daily # cigarettes						
	None	< 5	5-14	15-24	25-49	50+	Total
Cancer							1357
Control							1357
Total	68	184	1059	906	447	50	2714

Test of Homogeneity

A: H_0 predicts the following **expectations**:

	Daily # cigarettes						
	None	< 5	5-14	15-24	25-49	50+	Total
Cancer	34	92	529.5	453	223.5	25	1357
Control	34	92	529.5	453	223.5	25	1357
Total	68	184	1059	906	447	50	2714

Each group has the same proportion of smokers in each cell as the overall **marginal proportion.** The "equal" expected number for each group is the result of the equal sample size in each group

Q: What would change if there were half as many cases as controls?

Test of Homogeneity

	Daily # cigarettes						
	None	< 5	5-14	15-24	25-49	50+	Total
Cancer	O_{11}=7	O_{12}=55	489	475	293	38	1357
Control	O_{21}=61	129	570	431	154	12	1357
Total	68	184	1059	906	447	50	2714

More generally, if:

- O_{ij} is the observed frequency in row *i* and col *j*
- $m_i = \Sigma_j O_{ij}$ is the row *i* total
- $n_j = \Sigma_i O_{ij}$ is the column *j* total
- $N = \Sigma_i m_i = \Sigma_j n_j$ is the grand total

Then (under H_0) the expected frequency in row *i* and column *j* is

$$E_{ij} = (m_i \times n_j)/N$$

Test of Homogeneity

Summing the differences between the observed and expected counts provides an overall assessment of the adequacy of H_0.

$$X^2 = \sum_{i,j} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \sim \chi^2\left((R-1)\times(C-1)\right)$$

X^2 is known as the **Pearson's Chi-square Statistic.**

Test of Homogeneity

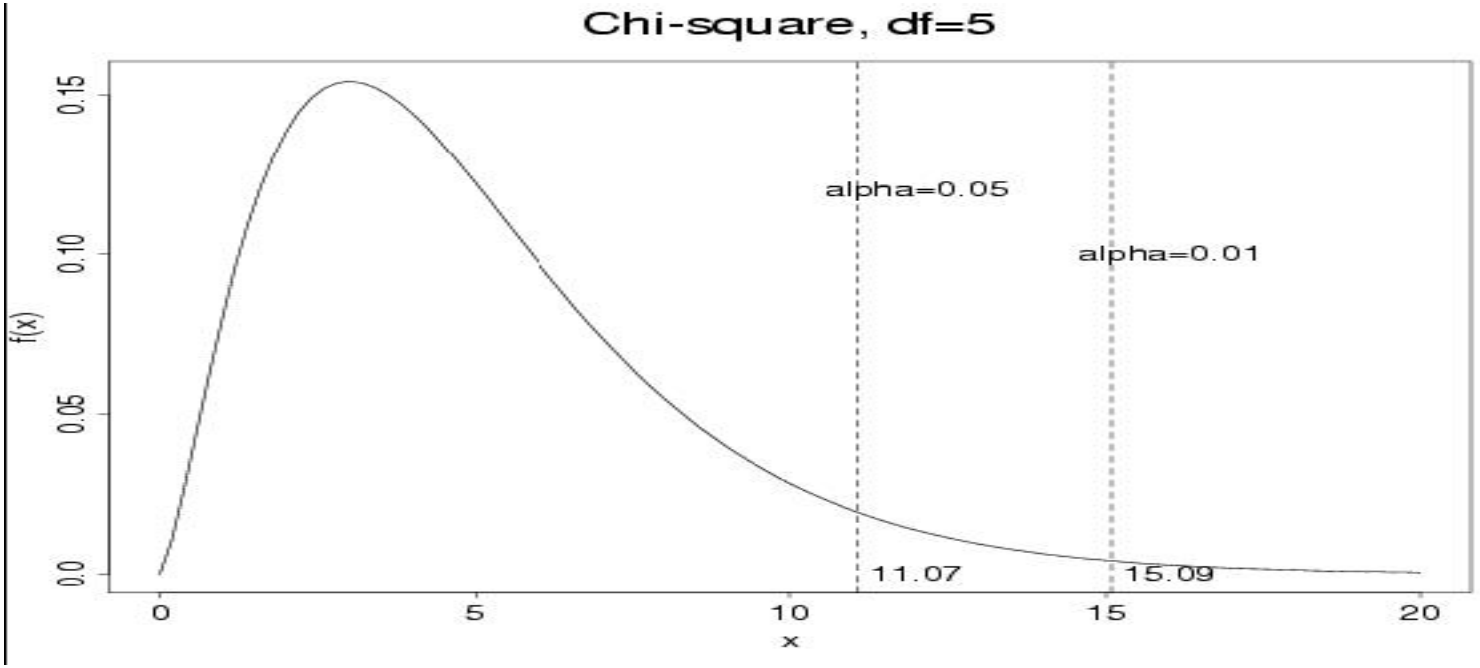
In Example 2 the contributions to the X^2 statistic are:

	Daily # cigarettes						
	None	< 5	5-14	15-24	25-49	50+	Total
Cancer	$\frac{(7-34)^2}{34}$	$\frac{(55-92)^2}{92}$	etc.				
Control	$\frac{(61-34)^2}{34}$	$\frac{(129-92)^2}{92}$					
Total							

	Daily # cigarettes						
	None	< 5	5-14	15-24	25-49	50+	Total
Cancer	21.44	14.88	3.10	1.07	21.61	6.76	
Control	21.44	14.88	3.10	1.07	21.61	6.76	
Total							

$$X^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 137.7$$

Chi-square Distribution



Looking in χ^2 table, we find that $Q^{.95}_{\chi^2(5)} = 11.07$.

Q: What is our conclusion?

STATA: Chi-square Test of Homogeneity

```
. input cancer cigs count

1 0 7
1 1 55
1 2 489
1 3 475
1 4 293
1 5 38
0 0 61
0 1 129
0 2 570
0 3 431
0 4 154
0 5 12
. end

. tabulate cancer cigs [freq=count], row chi2 expected
:
:
```

cancer	cigs 0	1	2	3	4	5	Total
0	61	129	570	431	154	12	1,357
	34.0	92.0	529.5	453.0	223.5	25.0	1,357.0
	4.50	9.51	42.00	31.76	11.35	0.88	100.00
1	7	55	489	475	293	38	1,357
	34.0	92.0	529.5	453.0	223.5	25.0	1,357.0
	0.52	4.05	36.04	35.00	21.59	2.80	100.00
Total	68	184	1,059	906	447	50	2,714
	68.0	184.0	1,059.0	906.0	447.0	50.0	2,714.0
	2.51	6.78	39.02	33.38	16.47	1.84	100.00

Pearson chi2(5) = 137.7193 Pr = 0.000

General Chi-square Test Procedure

	Factor Levels				
	1	2	…	C	Total
1	O_{11}	O_{12}	…	O_{1C}	m_1
Group 2	O_{21}				m_2
3	O_{31}				m_3
⋮	⋮				
R	O_{R1}			O_{RC}	m_R
Total	n_1	n_2		n_C	N

1. Compute the expected cell counts under homogeneity assumption: $E_{ij} = m_i n_j / N$

2. Compute the chi-square statistic:

$$X^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

3. Compare X^2 to $\chi^2(df)$ where

$$df = (R\text{-}1) \times (C\text{-}1)$$

Chi-square Test of Independence

The **Chi-squared Test of Independence** is identical in its implementation to the test for homogeneity. The only difference is that the R x C table is formed based on a random sample of N subjects according to the levels of 2 factors (see Example 1). Therefore, the null and alternative hypotheses are different:

H_0: The two factors are statistically independent ($p_{ij} = p_{i.} * p_{.j}$)

H_A: The two factors are not statistically independent

Statistical independence implies that each row has the same relative frequencies (or each column has the same relative frequencies). Thus, the *expected* frequencies are calculated just as for the test of homogeneity where the row (or column) totals are considered *fixed* by *design*.

Chi-square Test of Independence

Example 1 is a situation where individuals are classified according to two factors. In this example, the assumption of independence implies that willingness to participate doesn't depend on the level of education.

	definitely not	**probably not**	**probably**	**definitely**
< high school	52	79	342	226
high school	62	153	417	262
some college	53	213	629	375
college	54	231	571	244
some post college	18	46	139	74
graduate/prof	25	139	330	116

Q: What are the *df* for Example 1?

A: ***df*** = (4-1)x(6-1)=15

Q: What is critical value if $\alpha = 0.05$? A: $Q^{.95}_{\chi^2(15)} = 25.0$

STATA: Chi-square Test of Independence

```
. input educ willing count
0 0 52
0 1 79
0 2 342
0 3 226
1 0 62
1 1 153
1 2 417
1 3 262
2 0 53
2 1 213
2 2 629
2 3 375
3 0 54
3 1 231
3 2 571
3 3 244
4 0 18
4 1 46
4 2 139
4 3 74
5 0 25
5 1 139
5 2 330
5 3 116
. end
```

STATA: Chi-square Test of Independence

```
. tabulate educ willing [freq=count], row chi2
```

Key
frequency
row percentage

educ	willing 0	1	2	3	Total
0	52	79	342	226	699
	7.44	11.30	48.93	32.33	100.00
1	62	153	417	262	894
	6.94	17.11	46.64	29.31	100.00
2	53	213	629	375	1,270
	4.17	16.77	49.53	29.53	100.00
3	54	231	571	244	1,100
	4.91	21.00	51.91	22.18	100.00
4	18	46	139	74	277
	6.50	16.61	50.18	26.71	100.00
5	25	139	330	116	610
	4.10	22.79	54.10	19.02	100.00
Total	264	861	2,428	1,297	4,850
	5.44	17.75	50.06	26.74	100.00

```
Pearson chi2(15) =  89.7235   Pr = 0.000
```

SUMMARY
X^2 Test for RxC Tables

1. Tests of **homogeneity** of a factor across groups or **independence** of two factors rely on **Pearson's X^2 statistic**.

2. **X^2** is compared to a χ^2((R-1)x(C-1)) distribution

   (`display chiprob(df,X²)`).

3. Expected cell counts should be larger than 5.

4. This is a global test without using possible factor ordering. Ordered factors permit a test for trend (next).

X^2 Test for 2xC Tables with Ordered Categories

Example 3: *Smoking and quality of life*

	Self Reported Quality of Health					
Smoke	Poor	Fair	Good	V.Good	Exc.	Total
No	11	27	42	53	11	144
Yes	7	15	16	13	1	52
	18	42	58	66	12	196

- Is there an association between the self report of health and smoking?
  - Pearson chi-squared statistic gives:

$$\sum (O-E)^2 / E = 6.88 \sim \chi^2_4$$

    with p = 0.14.
  - Does that imply that there is no detectable association?
  - Does the Pearson statistic change value if the columns are permuted? Should it?
- Can we take advantage of the natural ordering of the columns?

X^2 Test for 2xC Tables with Ordered Categories

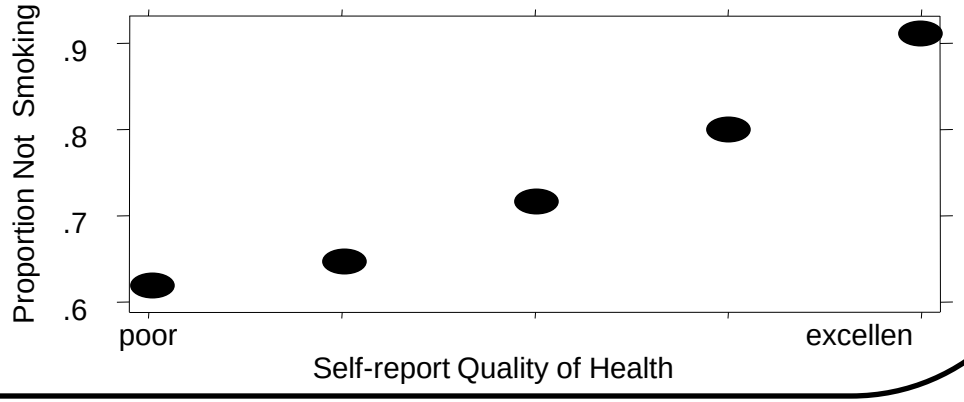
	Self Report Quality of Health					
Smoke	Poor	Fair	Good	V.Good	Exc.	Total
No	11	27	42	53	11	144
Yes	7	15	16	13	1	52
n_j	18	42	58	66	12	196
$\hat{P}_j$	0.61	0.64	0.72	0.80	0.92	0.73

The usual Pearson chi-squared statistic test

$H_o : p_1 = p_2 = \ \ldots \ = p_C$

$H_a : p_k \neq p_j$ for some k, j

ignores the pattern:



X^2 Test for 2xC Tables with Ordered Categories

Consider the alternative hypothesis

H_A: $p_1 \leq p_2 \leq p_3 \ldots \leq p_C$ (or $\geq$) (inequality for at least one pair)

General principle: The more specific the alternative, the more powerful the test (against that alternative)

Method:

1. Define the "doses" (levels of the categories)
   i. Equally spaced : $x_j = 1, 2, 3 \ldots$
   ii. Multiplicative : $x_j = 1, 2, 4, 8, 16 \ldots$
   iii. Log, other, …
2. Compute the test statistic (see Stata output). Result depends on choice of x_j (doses)!
3. Test statistic is $\chi^2(1)$ distributed under H_0.
4. This is known as the Cochran-Armitage test for trend (see Breslow and Day I, 4.5)

STATA

X^2 Test for 2xC Tables with Ordered Categories

```
. input smoke count1 count2 count3 count4 count5
0 11 27 42 53 11
1 7 15 16 13 1
. end

. reshape long count, i(smoke) j(health)
```

smoke	health	count
0	1	11
0	2	27
0	3	42
0	4	53
0	5	11
1	1	7
1	2	15
1	3	16
1	4	13
1	5	1

STATA
X^2 Test for 2xC Tables with Ordered Categories

```
. tabodds smoke health [freq=count]

------------------------------------------------------------------------------
    health  |      cases     controls       odds      [95% Conf.Int]
------------+-----------------------------------------------------------------
          1 |          7           11    0.63636        0.24669   1.64156
          2 |         15           27    0.55556        0.29554   1.04434
          3 |         16           42    0.38095        0.21419   0.67755
          4 |         13           53    0.24528        0.13373   0.44990
          5 |          1           11    0.09091        0.01174   0.70414
------------------------------------------------------------------------------
Test of homogeneity (equal odds): chi2(4)  =     6.85
                                  Pr>chi2  =   0.1443

Score test for trend of odds:     chi2(1)  =     6.63
                                  Pr>chi2  =   0.0100
```

Categorical vs. Continuous

Assume that, instead of a 2-sample t-test, you grouped a continuous outcome, Y, into C categories, as in the following table where C=5:

	G_1	G_2	G_3	G_4	G_5	Total
Sample 1	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	N
Sample 2	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}	N

Q: How much information is lost by grouping?

Efficiency lost by grouping

The following table compares the efficiency with which the difference between two groups can be estimated using a factor variable, relative to $\bar{Y}_1 - \bar{Y}_2$

# Categories	Relative efficiency
2	56.5
3	72.7
4	80.0
5	84.1
6	86.7
7	88.4
8	89.7
9	90.7

SUMMARY
X^2 Tests for RxC Tables

1. Tests of **homogeneity** of a factor across groups or **independence** of two factors rely on **Pearson's X^2 statistic**.
2. X^2 is compared to a $\chi^2(c-1)$ distribution.
3. Ordered factors permit a test for trend. The Cochran-Armitage test statistic is compared to a $\chi^2(1)$ distribution.
4. Creating a factor by categorizing a continuous variable results in a loss of efficiency which decreases as the number of categories increases.

2x2 Tables

Example 1: Pauling (1971)

Patients are randomized to either receive Vitamin C or placebo. Patients are followed-up to ascertain the development of a cold.

	Vit C	Placebo	Total
Cold-Yes	17	31	48
Cold-No	122	109	231
Total	139	140	279

Q: Is treatment with Vitamin C associated with a reduced probability of getting a cold?

Q: If Vitamin C is associated with reducing colds, then what is the magnitude of the effect?

2x2 Tables

Example 2: Keller (AJPH, 1965)

Patients with (cases) and without (controls) oral cancer were surveyed regarding their smoking frequency (this table collapses over the smoking frequency categories).

	Smoker	**Non-Smoker**	**Total**
Case	484	27	511
Control	385	90	475
Total	869	117	986

Q: Is oral cancer associated with smoking?

Q: If smoking is associated with oral cancer, then what is the magnitude of the risk?

2x2 Tables

Example 3: Norusis (1988)

In 1984, a random sample of US adults were cross-classified based on their income and reported job satisfaction:

	Dissatisfied	**Satisfied**	**Total**
< $15,000	104	391	495
≥ $15,000	66	340	406
Total	170	731	901

Q: Is salary associated with job satisfaction?

Q: If salary is associated with satisfaction, then what is the magnitude of the effect?

2x2 Tables

Each of these tables can be represented as follows:

	E	not E	Total
D	a	b	$(a + b) = m_1$
not D	c	d	$(c + d) = m_2$
Total	$(a + c) = n_1$	$(b + d) = n_2$	N

The question of association can be addressed with **Pearson's** X^2. We compute the **expected** cell counts as follows:

Expected under H_0:

	E	**not E**	**Total**
D	$\mathbf{n_1m_1/N}$	$\mathbf{n_2m_1/N}$	$\mathbf{(a + b) = m_1}$
not D	$\mathbf{n_1m_2/N}$	$\mathbf{n_2m_2/N}$	$\mathbf{(c + d) = m_2}$
Total	$\mathbf{(a + c) = n_1}$	$\mathbf{(b + d) = n_2}$	**N**

2x2 Tables

Example 1: Pauling (1971)

	Vit C	Placebo	Total
Cold-Yes	17	31	48
Cold-No	122	109	231
Total	139	140	279

H_0 : probability of "disease" *does not* depend on treatment

H_A : probability of "disease" *does* depend on treatment

$$X^2 = \frac{N(ad-bc)^2}{n_1 n_2 m_1 m_2}$$

← "quick" computing formula for 2x2 tables

$$= \frac{279(17\times 109 - 31\times 122)^2}{139\times 140\times 48\times 231}$$

$$= 4.81$$

The p-value is $P(\chi^2(1) > 4.81) = 0.028$ (same as 2 sample test of proportions!)

2x2 Tables
Applications in Epidemiology

Example 1: Cold following Vitamin C or Placebo

Cohort sampling

- Sample n_1 "exposed" and n_2 "unexposed" from the population.
- Follow all subjects for a *fixed period of time* (same for everyone).
- Observe ***a*** "cases" or "diseased" among the exposed
- Observe ***b*** "cases" or "diseased" among the unexposed
- This is a **prospective study**.

Sampling model: Two independent binomials

$$a \sim \text{Binomial}(p_1, n_1)$$
$$b \sim \text{Binomial}(p_2, n_2)$$

where

- $p_1 = P(D|E)$=disease probability for exposed
- $p_2 = P(D|\text{notE})$=disease probability for unexposed

2x2 Tables: Cohort Studies
Measures of Association

RD = p_1 - p_2 = risk difference

- Also known as attributable or excess risk
- Measures absolute effect: cases among the exposed that are "attributable" to exposure

RR = p_1 / p_2 = risk ratio (relative risk)

- Measures relative effect of exposure
- Constrained by denominator probability
  - RR ≤ 2 if p_2=0.5
  - RR ≤ 1.25 if p_2=0.8
  - In general RR $\leq 1/p_2$
- The range of RR is $[0, \infty)$. By taking the logarithm, we have $(-\infty, +\infty)$ as the range for ln(RR) and a better approximation to normality for the estimated ln(RR)

Estimated Measures of Association

	Vit C	Placebo	Total
Cold-Yes	17	31	48
Cold-No	122	109	231
Total	139	140	279

$$\hat{RD} = \hat{P}(\text{Cold}|\text{VitC}) - \hat{P}(\text{Cold}|\text{Placebo})$$
$$= 17/139 - 31/140 = 0.122 - 0.221$$
$$= -0.099$$

$$\hat{RR} = \hat{P}(\text{Cold}|\text{VitC}) \,/\, \hat{P}(\text{Cold}|\text{Placebo})$$
$$= (17/139)/(31/140) = 0.122/0.221$$
$$= 0.55$$

Using STATA

```
. csi 17 31 122 109

                |   Exposed   Unexposed  |     Total
----------------+------------------------+----------
          Cases |        17          31  |        48
       Noncases |       122         109  |       231
----------------+------------------------+----------
          Total |       139         140  |       279
                |                        |
           Risk |  .1223022    .2214286  |   .172043
                |                        |
                |      Point estimate    |  [95% Conf. Interval]
                |------------------------+----------------------
 Risk difference |        -.0991264       | -.1868592   -.0113937
      Risk ratio |         .5523323       |  .3209178    .9506203
 Prev. frac. ex. |         .4476677       |  .0493797    .6790822
 Prev. frac. pop |         .2230316       |
                +-------------------------------------------------
                           chi2(1) =     4.81  Pr>chi2 = 0.0283
```

2x2 Tables
Applications in Epidemiology

Example 2: Oral cancer and smoking

Case-control sampling

- Sample m_1 cases (individuals with cancer) and m_2 controls (individuals without cancer) from the population.
- Ask about exposure to smoking in the past
- Observe ***a*** exposed individuals among the cases
- Observe ***c*** exposed individuals among the controls
- This is a **retrospective study**.

Sampling model: Two independent binomials

$$a \sim \text{Binomial}(p_1^E, n_1)$$
$$c \sim \text{Binomial}(p_2^E, n_2)$$

where

- p_1^E=P(E|D)=disease probability for exposed
- p_2^E=P(E|notD)=disease probability for unexposed

2x2 Tables
Applications in Epidemiology

Case-control sampling (example 2)

We can estimate p_1^E and p_2^E but we can't estimate p_1 and $p_2 \Rightarrow$ can't estimate RR. ☹

Instead of the relative risk we can estimate the "**exposure odds ratio**" which Cornfield (1951) showed equivalent to the **disease odds ratio**:

$$\frac{P(E \mid D)/(1-P(E \mid D))}{P(E \mid \overline{D})/(1-P(E \mid \overline{D}))} = \frac{P(D \mid E)/(1-P(D \mid E))}{P(D \mid \overline{E})/(1-P(D \mid \overline{E}))}$$

In other words, **the odds ratio can be estimated regardless of the sampling scheme.**

Absolutely amazing … so what?

The Odds Ratio (OR)

For rare diseases, P(D | E) ≈ 0 so that the **disease odds ratio** then approximates the relative risk:

$$\frac{P(D \mid E)/(1-P(D \mid E))}{P(D \mid \overline{E})/(1-P(D \mid \overline{E}))} \approx \frac{P(D \mid E)}{P(D \mid \overline{E})}$$

Since with case-control data we are able to effectively estimate the exposure odds ratio …

we are then able to equivalently estimate the disease odds ratio …

which for rare diseases approximates the relative risk.

For rare diseases, the odds ratio approximates the relative risk.

Odds Ratio

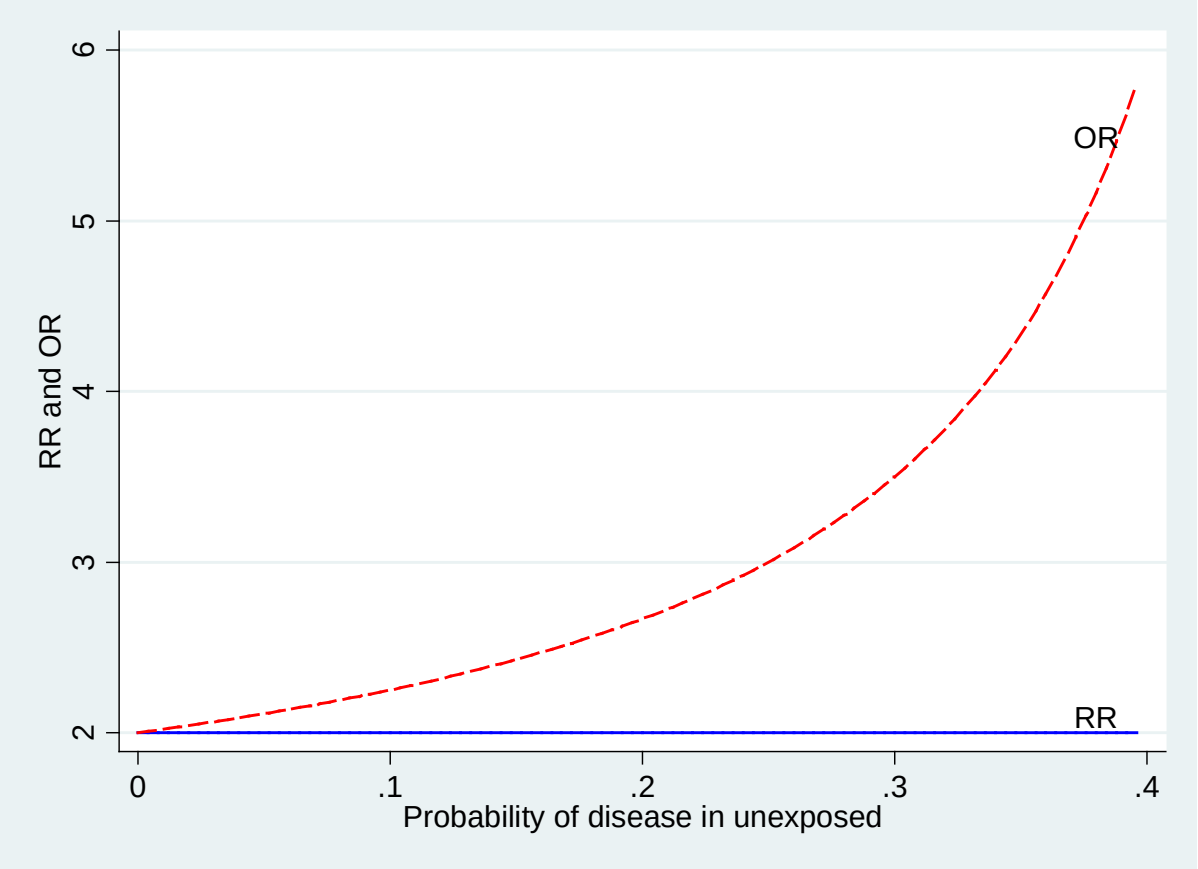
OR = $[p_1/(1- p_1)]/ [p_2/(1- p_2)]$= odds ratio

- Not constrained by denominator
  $0 < OR < \infty$ regardless of p_2
- "Natural" parameter in logistic regression (coming…)
- Asymptotic p-values and CI's are valid even for small to moderate sized samples
- The "disease odds ratio" is the odds of disease for the exposed group divided by the odds of disease for the unexposed group
- Interesting fact: if p_1 and p_2 small then OR ≈ RR

Null hypotheses:

$$p_1=p_2 \Leftrightarrow RD=0 \Leftrightarrow RR=OR=1$$

The Odds Ratio & Relative Risk



For rare diseases, the odds ratio approximates the relative risk.

Inference for the Odds Ratio

As for the relative risk, the range of the odds ratio is [0, ∞). The range of the **ln odds ratio** is (- ∞, +∞). The normal distribution is better as an approximation to the distribution of the estimated *ln* (odds ratio).

```
. cci 484 27 385 90

                                                    Proportion
                |   Exposed   Unexposed  |     Total     Exposed
----------------+------------------------+----------------------
          Cases |       484          27  |       511      0.9472
       Controls |       385          90  |       475      0.8105
----------------+------------------------+----------------------
          Total |       869         117  |       986      0.8813
                |                        |
                |      Point estimate    |  [95% Conf. Interval]
                |------------------------+----------------------
     Odds ratio |         4.190476       |  2.633584    6.836229  (exact)
 Attr. frac. ex. |         .7613636       |  .6202893    .8537205  (exact)
 Attr. frac. pop |          .721135       |
                +-------------------------------------------------
                           chi2(1) =    43.95  Pr>chi2 = 0.0000
```

Interpreting the Odds Ratio

1. What is the *outcome* (i.e. “disease”) of interest?
2. What are the *two groups* (i.e. exposed and unexposed) being contrasted?

$$\text{OR} = \frac{\text{odds of OUTCOME in EXPOSED}}{\text{odds of OUTCOME in UNEXPOSED}}$$

- Close to RR for rare diseases
- Meaningful for both cohort and case-control studies
- OR > 1 $\Rightarrow$ increased risk of OUTCOME with EXPOSURE
- OR < 1 $\Rightarrow$ decreased risk of OUTCOME with EXPOSURE

2x2 Tables
Applications in Epidemiology

Cross-sectional study design:

- Simple random sample from the entire population, not by disease status or exposure status
- Categorical variable with 4 possible outcomes (D,E), (nD,E) (D,nE) (nD,nE)
- Use RD, RR or OR to summarize association
- Cases of disease are **prevalent** cases (compared to incident cases in a prospective or cohort study)

Example 3 (pg 40) is an example of a **cross-sectional** study since only the total sample size for the entire table is fixed in advance. The row totals or column totals are not fixed in advance.

SUMMARY
Measures of Association for 2x2 Tables

RD = $p_1 - p_2$ = risk difference (null: RD = 0)

- measures **absolute effect** of exposure
- sometimes more useful for understanding public health effect of an intervention

RR = p_1 / p_2 = relative risk (null: RR = 1)

- measures **relative effect** of exposure
- bounded above by $1/p_2$

OR = $[p_1(1-p_2)]/[p_2(1-p_1)]$ = odds ratio (null: OR = 1)

- range is 0 to ∞
- approximates RR for rare events
- invariant of switching rows and cols
- key parameter in logistic regression
- good behavior of p-values/CI even for small sample size

SUMMARY
Sampling Designs for 2x2 Tables

1. **Cohort** ("Prospective", "Follow-up")
   - Sample n_1 "exposed" and n_2 "unexposed"
   - Follow everyone for equal period of time
   - Observe incident disease : ***a*** cases among exposed, ***b*** cases among unexposed
   - Model: Two independent binomials

     $a \sim \text{Binomial}(p_1, n_1)$

     $b \sim \text{Binomial}(p_2, n_2)$

     $p_1 = P(D|E)$

     $p_2 = P(D|\text{notE})$
   - Useful measures of association : RR,OR,RD

SUMMARY
Sampling Designs for 2x2 Tables

2. Case-Control (retrospective)

- Sample m_1 "cases" and m_2 "controls"
- Observe exposure history : ***a*** exposed among cases, ***c*** exposed among controls
- Model: Two independent binomials

  a ~ Binomial(p_1^E,m_1)

  c~ Binomial(p_2^E,m_2)

  $p_1^E = P(E|D)$ $\quad$ $p_2^E = P(E|notD)$
- Useful measures of association : OR

3. Cross-sectional

- Sample n individuals from population
- Observe both "exposure" and (prevalent) "disease" status.
- No longitudinal follow-up
- Useful measures of association : RR,OR,RD

SUMMARY
Statistical Inference for RD, RR, OR

Estimation: $\hat{p}_1 = a/n_1 \qquad \hat{p}_2 = b/n_2$

$$\hat{RD} = \hat{p}_1 - \hat{p}_2 \qquad \hat{RR} = \frac{\hat{p}_1}{\hat{p}_2} \qquad \hat{OR} = \frac{\hat{p}_1(1-\hat{p}_2)}{\hat{p}_2(1-\hat{p}_1)}$$

Standard error estimates:

$$SE(\hat{RD}) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{ac}{n_1^3} + \frac{bd}{n_2^3}}$$

$$SE(\log \hat{RR}) = \sqrt{\frac{(1-\hat{p}_1)}{\hat{p}_1 n_1} + \frac{(1-\hat{p}_2)}{\hat{p}_2 n_2}} = \sqrt{\frac{c}{an_1} + \frac{d}{bn_2}}$$

$$SE(\ln \hat{OR}) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

Fisher's Exact Test

Motivation: When a 2×2 table contains cells that have fewer than 5 *expected* observations, the normal approximation to the distribution of the log odds ratio (or other summary statistics) is known to be poor. This can lead to incorrect inference since the p-values based on this approximation are not valid.

Solution: Use Fisher's Exact Test

	E+	E-	Total
D+			m_1
D-			m_2
Total	n_1	n_2	N

Fisher's Exact Test

Example: (Rosner, sec 10.3) A case-control study was conducted among men aged 50-54 who died from CVD over a 1-month period and similarly aged controls. The investigators tried to include equal numbers of men who died from CVD and those that did not. Then, asking a close relative, the dietary habits of the subjects were ascertained.

	High Salt	Low Salt	Total
CVD	5	30	35
Non-CVD	2	23	25
Total	7	53	60

$$OR = \frac{5 \times 23}{2 \times 30} = 1.92$$

Expected counts under H_0:

	High Salt	Low Salt
CVD	4.08	30.92
Non-CVD	2.91	22.09

Fisher's Exact Test

	E+	E-	Total
D+	**a**		m_1
D-			m_2
Total	n_1	n_2	N

- Compare observed table to other possible tables assuming same margins
- Under the null hypothesis,

  H_0 : OR = 1

  the probability of each table can be computed (hypergeometric dist.)
- See how extreme the observed table is compared to all possible tables

Fisher's Exact Test

Example: Cardiovascular disease.

	High Salt	Low Salt	Total
CVD	5	30	35
Non-CVD	2	23	25
Total	7	53	60

Possible Tables:

0		35
		25
7	53	60

1		35
		25
7	53	60

2		35
		25
7	53	60

3		35
		25
7	53	60

4		35
		25
7	53	60

5		35
		25
7	53	60

6		35
		25
7	53	60

7		35
		25
7	53	60

Fisher's Exact Test

For fixed marginal totals, m_1, m_2, n_1, n_2 and overall total, N, the hypergeometric probability of any given constellation of cell counts a, b,c, d is given by:

$$\frac{m_1!m_2!n_1!n_2!}{N!a!b!c!d!}$$

0	35	35
7	18	25
7	53	60

.001

1	34	35
6	19	25
7	53	60

.016

2	33	35
5	20	25
7	53	60

.082

3	32	35
4	21	25
7	53	60

.214

4	31	35
3	22	25
7	53	60

.312

5	30	35
2	23	25
7	53	60

.252

6	29	35
1	24	25
7	53	60

.105

7	28	35
0	25	25
7	53	60

.017

Fisher's Exact Test using STATA

```
. cci 5 30 2 23,exact

                                                Proportion
                |   Exposed   Unexposed  |      Total     Exposed
----------------+------------------------+----------------------
          Cases |         5          30  |         35      0.1429
       Controls |         2          23  |         25      0.0800
----------------+------------------------+----------------------
          Total |         7          53  |         60      0.1167
                |                        |
                |      Point estimate    |  [95% Conf. Interval]
                |------------------------+----------------------
     Odds ratio |         1.916667       |  .2789585    21.62382  (exact)
 Attr. frac. ex. |         .4782609       | -2.584763    .9537547  (exact)
 Attr. frac. pop |          .068323       |
                +-----------------------------------------------
                             1-sided Fisher's exact P = 0.3747
                             2-sided Fisher's exact P = 0.6882
```

The X^2 Test, for comparison

```
. cci 5 30 2 23

                                                    Proportion
                |   Exposed   Unexposed  |     Total     Exposed
----------------+------------------------+----------------
          Cases |         5          30  |        35      0.1429
       Controls |         2          23  |        25      0.0800
----------------+------------------------+----------------
          Total |         7          53  |        60      0.1167
                |                        |
                |      Point estimate    |  [95% Conf. Interval]
                |------------------------+----------------
     Odds ratio |         1.916667       |  .2789585    21.62382  (exact)
 Attr. frac. ex. |         .4782609       | -2.584763    .9537547  (exact)
 Attr. frac. pop |          .068323       |
                +-------------------------------------------------
                           chi2(1) =     0.56  Pr>chi2 = 0.4546
```

Paired Binary Data

Example 4: Sartwell et al (1969)

Is oral contraceptive use associated with thromboembolism? 175 cases with blood clots of unknown origin were matched to controls based on age, race, time and place of hospitalization, parity, marital status and SES.

		Control OC Use	
		Yes	No
Case OC Use	Yes	10	57
	No	13	95

Q: Is OC use associated with thromboembolism?

Q: If OC use is associated with thromboembolism then what is the magnitude of the effect?

Paired Binary Data

Example 4 is an example of **paired binary data**. One way to display these data is the following:

	OC	No OC	Total
Case	67	108	175
Control	23	152	175
Total	90	260	350

Q: Can't we simply use X^2 Test of Homogeneity to assess whether this is evidence for an increase in knowledge?

A: NO!!! The X^2 tests assume that the rows are **independent** samples. In this design, the controls are constrained to be similar to the controls in many respects (recall paired t-test vs two sample t-test)

Paired Binary Data

For paired binary data we display the results as follows:

		Control OC	
		Yes	No
Case OC	Yes	n_{11}	n_{10}
	No	n_{01}	n_{00}

- (Yes,Yes) and (No,No) pairs provide no information about effect of OC use. These are known as the **concordant pairs**.
- The information regarding OC use is in the **discordant pairs**, (No, Yes) and (Yes, No).

p_1 = Pr (OC use in cases)
p_2 = Pr(OC use in controls)

$H_0 : p_1 = p_2$
$H_A : p_1 \neq p_2$

Paired Binary Data
McNemar's Test

Under the null hypothesis, $H_0: p_1 = p_2$, we expect equal numbers of (01) and (10) discordant pairs ($E[n_{01}] = E[n_{10}]$). Specifically, under the null:

$$M = n_{01} + n_{10}$$

$$n_{10} \mid M \sim Bin\left(M, \frac{1}{2}\right)$$

$$Z = \frac{n_{10} - M\frac{1}{2}}{\sqrt{M\frac{1}{2}\left(1-\frac{1}{2}\right)}}$$

Under H_0, $Z^2 \sim \chi^2(1)$, and forms the basis for **McNemar's Test for Paired Binary Responses**.

The odds ratio comparing OC use to no OC use is estimated by:

$$\hat{OR} = \frac{n_{01}}{n_{10}}$$

CI: Breslow and Day (1981), sec. 5.2.

Paired Binary Data
McNemar's Test

Example 4: OC use and thromboembolism

```
. mcci 10 57 13 95

                 | Controls                |
Cases            |   Exposed   Unexposed   |      Total
-----------------+-------------------------+------------
         Exposed |        10          57   |         67
       Unexposed |        13          95   |        108
-----------------+-------------------------+------------
           Total |        23         152   |        175

McNemar's chi2(1) =     27.66    Prob > chi2 = 0.0000
Exact McNemar significance probability       = 0.0000

Proportion with factor
        Cases       .3828571
        Controls    .1314286     [95% Conf. Interval]
                    ---------    --------------------
        difference  .2514286      .1597329   .3431243
        ratio       2.913043      1.918355   4.423488
        rel. diff.  .2894737      .1985361   .3804113

        odds ratio  4.384615      2.371377   8.731311   (exact)
```

Paired Binary Data

Paired data analyses arise in a number of situations …

- Matched case-control studies (as above)
- Repeated tests on an individual over time (e.g. before-after)
- Paired observations on an individual (e.g. two eyes)
- Twin studies
- Other …

Rater Agreement

- Sometimes pair-matched categorical data come from two raters who categorize the same subject/object
  - Two pathologists deciding if a biopsy is cancer
  - Two tests for HPV
- Key: Interest is in characterizing the amount of agreement, not comparing the probability of a “success”
- Key: Neither rater is considered perfect (no “gold standard”)

Rater Agreement

Example: Dietary questionnaire administered several months apart. A question was asked regarding the number of servings of beef consumed per week

		Survey 2	
		≤1/week	>1/week
Survey 1	≤1/week	136	92
	>1/week	69	240

Simple agreement:

- (136+240)/537 = 0.70
- CI, tests follow from binomial
- Is this enough?

Rater Agreement

- Consider the following hypothetical examples:

		rater2		
		Yes	No	Total
rater 1	Yes	95	0	95
	No	5	0	5
	Total	100	0	100

agreement = 95%

		rater2		
		Yes	No	Total
rater 1	Yes	45	5	50
	No	5	45	50
	Total	50	50	100

agreement = 90%

- Which pair of raters are agreeing better?

Rater Agreement

Kappa

- Chance-corrected agreement

		rater2		
		Yes	No	Total
rater 1	Yes	a	b	a+b
	No	c	d	c+d
	Total	a+c	b+d	N

- $\kappa = \frac{p_o - p_e}{1 - p_e}$ where

  $p_o = (a + d)/N$

  $p_e = (a+b)*(a+c)/N^2 + (b+d)(c+d)/N^2$

- Range $-p_e/(1 - p_e)$ to 1
- Test of Ho: $\kappa = 0$ possible; CI more interesting
- Extensions to multiple raters and multiple categories possible

Rater Agreement

- Consider the following hypothetical examples:

		rater2		
		Yes	No	Total
rater 1	Yes	95	0	95
	No	5	0	5
	Total	100	0	100

agreement = 95%

kappa = 0.0

		rater2		
		Yes	No	Total
rater 1	Yes	45	5	50
	No	5	45	50
	Total	50	50	100

agreement = 90%

kappa = .8

Landis and Koch (1977)

κ	**Interpretation**
< 0	Poor agreement
0.01 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.0	Almost perfect

Rater Agreement

		Survey 2	
		≤1/week	>1/week
Survey 1	≤1/week	136	92
	>1/week	69	240

```
. kap survey1 survey2, tab

           |        survey2
   survey1 |         0          1 |     Total
-----------+----------------------+----------
         0 |       136         92 |       228
         1 |        69        240 |       309
-----------+----------------------+----------
     Total |       205        332 |       537

             Expected
Agreement    Agreement      Kappa    Std. Err.         Z      Prob>Z
------------------------------------------------------------------------
  70.02%       51.78%     0.3782      0.0430       8.80       0.0000
```

- 95% CI: $0.378 \pm 1.96 * .043$

REVIEW
Two-way Tables

Q: Is there an association?

- R x C Tables
  - Chi-square tests of Homogeneity & Independence
- 2 x 2 Tables
  - Chi-square test
  - Fisher's exact test
  - Paired data and McNemar's test

REVIEW
Two-way Tables

Q: What is the magnitude of the association?

- Exploratory Data Analysis (EDA)
  - Row and/or column percentages
- Risk difference: $p_1 - p_2$
  - Risk = p = probability of an outcome
  - Confidence intervals and tests
- Relative Risk (risk ratio): p_1 / p_2
  - Confidence intervals and tests
- Odds Ratio: $[p_1(1-p_2)]/[p_2(1-p_1)]$
  - Confidence intervals and tests
- Kappa (chance-corrected agreement)
  - Confidence intervals

CONSIDERATIONS
in Data Analysis

"In an analysis, the basic questions to consider are the degree of association between risk for disease and the factors under study, the extent to which the observed associations may result from ***bias****,* ***confounding*** *and/or* ***chance****, and the extent to which they may be described as* ***causal****."*

Breslow & Day (1980) Volume I.

Stratified Tables

- Often, a third measure influences the relationship between the two primary measures (i.e. disease and exposure).
- In such cases, how can we tell if the "exposure" is the cause of the disease? (what do we mean by cause?)
- How do we "remove or control for the effect" of the third measure?

Example: Effect of seat belt use on accident fatality

	Seat Belt	
Driver	Worn	Not worn
dead	10	20
alive	40	30
Total	50	50
Fatality Rate	10/50 (20%)	20/50 (40%)

Stratified Tables

But, suppose...

	Impact Speed			
	$\leq$ 40 mph		> 40 mph	
Driver	seat belt		seat belt	
	worn	not	worn	not
dead	3	2	7	18
alive	27	18	13	12
Total	30	20	20	30
Fatality Rate	10%	10%	35%	60%

How does this affect your inference?

Stratified Tables - EM

Effect modification (aka Interaction):

- Magnitude or direction of effect varies between subgroups
- *Depends on what effect measure is chosen* (i.e. may observe EM for RR, not RD)
- In practice, is variation in effect between subgroups greater than expected by chance?
- If EM present, generally better to provide subgroup-specific results
- If pooling, need to consider if relative frequencies of subgroups in sample is similar to population

Stratified Tables - Confounding

"Condom Use increases the risk of STD"

		STD rate	
Condom Use	Yes	6/200	(3.0%)
	No	14/540	(2.6%)

BUT ...

		STD rate	
# Partners < 3			
Condom Use	Yes	1/100	(1%)
	No	10/500	(2%)
# Partners ≥ 3			
Condom Use	Yes	5/100	(5%)
	No	4/40	(10%)

Explanation: Individuals with more partners are more likely to use condoms. But individuals with more partners are also more likely to get STDs.

Assessing Causality in Observational Studies

From association to causation according to A. Bradford Hill:

Strength of association

- Helps rule out bias from unmeasured confounders

Consistency of association

- In different population groups
- Using different study designs

Specificity of association

- In disease subcategories
  - specific cell types of cancer
- In exposure groups
  - chest irradiation in prepubertal girls

Relationship in time (exposure precedes effect)

Biological gradient or dose-response

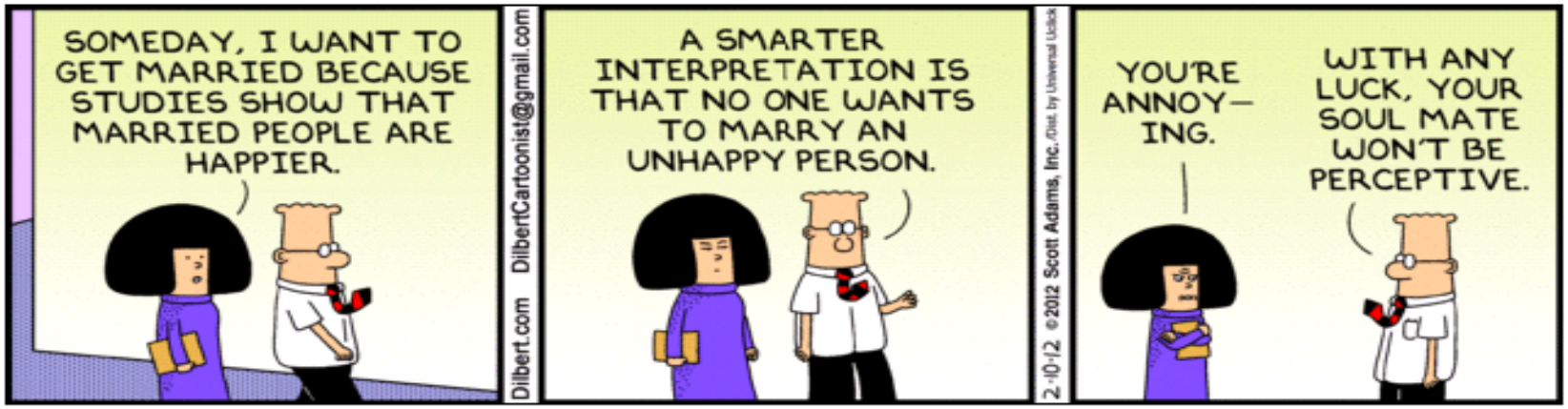
Assessing Causality in Observational Studies

Lack of alternative explanations:

- Adjustment for measured confounders, selection factors
- Sensitivity analyses to gauge impact of unmeasured confounders

Coherence of the Evidence: considerations external to the study

- Time trends in population incidence
- Laboratory studies



Causal Inference Concepts

- Potential Outcomes or "Counterfactuals"
  - Response of subject i if treated – $Y_i(1)$
  - Response of subject i if not treated – $Y_i(0)$
- In practice, we only observe one of these …
- But the **causal effect** involves both …

$$\Delta_i = Y_i(1) - Y_i(0)$$

- *We can never observe* Δ_i !

Causal Inference Concepts

In a randomized study we can get an unbiased estimate of the **average causal effect**

Treatment Assignment	Potential Outcome Tx = 1	Tx = 0
0		$Y_i(0)$
1	$Y_i(1)$	
1	$Y_i(1)$	
0		$Y_i(0)$
1	$Y_i(1)$	
0		$Y_i(0)$
0		$Y_i(0)$
1	$Y_i(1)$	
	$\bar{Y}(1)$	$\bar{Y}(0)$

Causal Inference Concepts

Assume:

$\bar{Y}(1)$ is an unbiased estimate of average response for the population **if everyone were treated**

$\bar{Y}(0)$ is an unbiased estimate of average response for the population **if everyone were not treated**

- Why are these assumptions reasonable for a randomized study ????
- **average causal effect** = $\bar{\Delta} = \bar{Y}(1) - \bar{Y}(0)$

Causal Inference Concepts

- We can estimate the average causal effect when there is nothing (other than exposure) that *systematically* differs between exposed and unexposed groups
- Randomization guarantees this – "no unmeasured confounding"
- Note that there is no single exposure effect – each individual may have a different effect, Δ_i
- Different populations (i.e. young, old) may have different average causal effects (this is an example of …. what?)

Some Views of Confounding

"*A confounding variable is a variable that is associated with both the disease and the exposure variable.*" Rosner (1995)

"*Confounding is the distortion of a disease/exposure association brought about by the association of other factors with both disease and exposure, the latter associations with disease being causal.*" Breslow & Day (1980)

"*If any factor either increasing or decreasing the risk of disease besides the characteristic or exposure under study is unequally distributed in the groups that are being compared with regard to the disease, this itself will give rise to differences in disease frequency in the compared groups. Such distortion, termed confounding, leads to an invalid comparison.*" Lillienfeld & Stolley (1994)

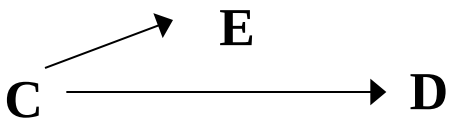
Some Views of Confounding

1. *A confounding factor must be a risk factor for the disease.*
2. *A confounding factor must be associated with the exposure under study in the source population (the population at risk from which the cases are derived).*
3. *A confounding factor must not be affected by the exposure or the disease. In particular, it cannot be an intermediate step in the causal path between the exposure and the disease."*

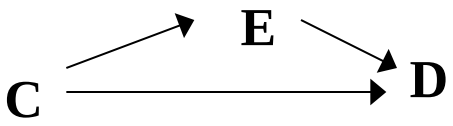
Rothman & Greenland (1998)

Confounding – Causal Diagrams

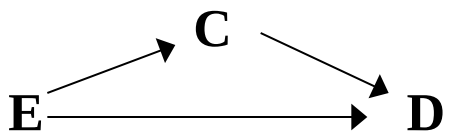
E = Exposure D = Disease C = Potential Confounder



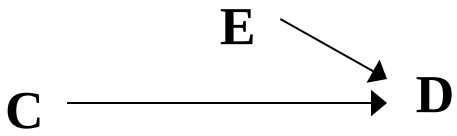
An apparent association between E (say, alcohol consumption) and D (say, lung cancer) is completely explained by C (say, smoking). C is a confounder.



An association between E and D is partly due to variations in C. C is a confounder.



C (say osteoporosis) is in the causal path between E (say calcium deficiency) and D (say hip facture). C is not a confounder.



C has an independent effect on D. C is not a confounder.

Stratified Tables - Confounding

"Condom Use increases the risk of STD"

		STD rate	
Condom Use	Yes	6/200	(3.0%)
	No	14/540	(2.6%)

		STD rate	
# Partners < 3			
Condom Use	Yes	1/100	(1%)
	No	10/500	(2%)
# Partners ≥ 3			
Condom Use	Yes	5/100	(5%)
	No	4/40	(10%)

- ✓ Partners related to disease (STD)
- ✓ Partners related to exposure (condom use)
- ✓ Assume partners not intermediate

Causal Inference Concepts

- Can we estimate (average) causal effects from observational (non-randomized) data?
- The difficulty with observational data is that "exposure" is not randomly assigned. This implies that the average outcome among those actually exposed **may not be equal** to the average outcome that would be observed if everyone was exposed (**selection bias**).

  Examples:

  - E=treatment with AZT, Y=CD4 cell count
  - E=condom use, Y=presence of STD
  - E=occupational manganese, Y=neurological impairment
  - and many more …

Q: What can we do in these situations?

A: Control for imbalances via stratification or regression

Choosing Confounders for Adjustment

- Selection on the basis of statistical significance of association with disease can leave residual confounding effect; not recommended
- Some advocate choice based on ***a priori*** considerations
  - o Study design/protocol specifies particular exposure-disease association under investigation
  - o Confounders selected/measured based on their role as known risk factors for the disease
- Others advocate choice of confounders based on how much they affect RR (OR, RD) when included/ excluded from the model.
- See work by Pearl, Robins and Greenland for more formal criteria
- Report results of up to three planned analyses
  - unadjusted
  - adjusted for primary covariates (known risk factors)
  - adjusted for primary and secondary risk factors (known and suspected risk factors)

Choosing Confounders for Adjustment

"*The object of the stratification … is … to alleviate the distortion in the estimated effect of exposure caused by confounding or selection bias. If another factor is causally related to disease, and if there is a chance it could be correlated with the exposure of interest, statistical adjustment is needed to produce a valid estimate. Since one rarely has good prior information about the degree of association between various risk factors in the population, a reasonable and prudent policy is to take account in the analysis of all known causal factors regardless of whether they may appear to be related to the exposure of interest in the data at hand. Significance testing of their relationship to either disease or exposure is irrelevant to the issue of whether adjustment for such effects modifies the association of interest*."

(Breslow, Ann. Rev. Publ. Health, 1982, p. 38-39)

Adjusting for Confounders via Stratification

Basic idea of adjusted estimates

- Compute separate effect estimates for each stratum of the confounder
- Assess homogeneity of effects across strata
- (Weighted) average effect estimate over strata (adjusted effect)
- Global null hypothesis: no effect in any stratum
- Different methods of pooling, testing have been proposed. We will focus on Mantel-Haenszel methods

Stratified Contingency Tables - Example

Example 1: (Rosner sec 13.4)

Suppose we are interested in the relationship between lung cancer incidence and heavy drinking (defined as ≥ 2 drinks per day). A prospective study is conducted where drinking status is determined at baseline and the cohort is followed for 10 years to determine cancer endpoints. Smoking status is also measured at baseline.

Stratified Contingency Tables - Example

1) Pooled data, not controlling for smoking

	Heavy Drinker		
	Yes	No	
Cancer	33	27	60
No Cancer	1667	2273	3940
	1700	2300	4000

Stratified Contingency Tables - Example

1) Pooled data, not controlling for smoking

```
. cci 33 27 1667 2273
                                                        Proportion
                |   Exposed   Unexposed  |      Total     Exposed
----------------+------------------------+------------------------
          Cases |        33          27  |         60      0.5500
       Controls |      1667        2273  |       3940      0.4231
----------------+------------------------+------------------------
          Total |      1700        2300  |       4000      0.4250
                |                        |
                |      Point estimate    |    [95% Conf. Interval]
                |------------------------+------------------------
     Odds ratio |         1.666533       |    .9677794    2.892948 (exact)
 Attr. frac. ex. |          .399952       |   -.0332934    .6543319 (exact)
 Attr. frac. pop |         .2199736       |
                +-------------------------------------------------
                             chi2(1) =     3.89  Pr>chi2 = 0.0484
```

Stratified Contingency Tables - Example

2) Stratified by smoking at baseline

Smokers

	Heavy Drinking		
	Yes	No	
Cancer	24	6	30
No Cancer	776	194	970
	800	200	1000

```
. cci 24 6 776 194

                                                    Proportion
                |   Exposed   Unexposed  |     Total     Exposed
----------------+------------------------+----------------------
          Cases |        24           6  |        30      0.8000
       Controls |       776         194  |       970      0.8000
----------------+------------------------+----------------------
          Total |       800         200  |      1000      0.8000
                |                        |
                |      Point estimate    |  [95% Conf. Interval]
                |------------------------+----------------------
     Odds ratio |                 1      |  .3911965    3.033018  (exact)
 Attr. frac. ex. |                 0      |  -1.55626    .6702954  (exact)
 Attr. frac. pop |                 0      |
                +-------------------------------------------------
                           chi2(1) =     0.00  Pr>chi2 = 1.0000
```

Stratified Contingency Tables - Example

Non-smokers

	Heavy Drinking		
	Yes	No	
Cancer	9	21	30
No cancer	891	2079	2970
	900	2100	3000

```
. cci 9 21 891 2079

                                                Proportion
                 |   Exposed   Unexposed  |      Total      Exposed
-----------------+------------------------+---------------------
           Cases |         9          21  |         30      0.3000
        Controls |       891        2079  |       2970      0.3000
-----------------+------------------------+---------------------
           Total |       900        2100  |       3000      0.3000
                 |                        |
                 |      Point estimate    |  [95% Conf. Interval]
                 |------------------------+---------------------
      Odds ratio |                1       |   .4015748    2.288393  (exact)
 Attr. frac. ex. |                0       | -1.490196     .5630121  (exact)
 Attr. frac. pop |                0       |
                 +-----------------------------------------------
                            chi2(1) =     0.00  Pr>chi2 = 1.0000
```

Stratified Contingency Tables

Q: How can we combine the information from both tables to obtain

- a single adjusted estimated OR and CI
- an overall test of significance that takes account of the stratification?

A: **Mantel-Haenszel Methods** – assess association between disease and exposure after controlling for one or more confounding variables.

Notation:

	E	$\overline{E}$	
D	a_i	b_i	$(a_i + b_i)$
$\overline{D}$	c_i	d_i	$(c_i + d_i)$
	$(a_i + c_i)$	$(b_i + d_i)$	N_i

where i = 1,2,…,K are the strata.

Combining Epidemiologic Measures

Notes:

1) There are different ways of combining stratum-specific estimates into a common estimate, including Woolf's method and the Mantel-Haenszel method. We will focus on the latter.

2) There are also different ways of testing for heterogeneity, including Mantel-Haenszel and Breslow-Day. The latter tends to be more robust and is most commonly used.

3) It is also possible to estimate stratum-specific and common attributable risks and relative risks (and their confidence intervals). See Kleinbaum, Kupper & Morgenstein Table 17.16 or Hennekens & Buring, Table 12.7.

Mantel-Haenszel Methods

(1) Estimate the common odds ratio

The Mantel-Haenszel estimate of the odds ratio assumes there is a **common** odds ratio:

$$OR_{pool} = OR_1 = OR_2 = \ldots = OR_K$$

To estimate the common odds ratio we take a weighted average of the stratum-specific odds ratios:

Recall: $\hat{OR}_i = a_i d_i / b_i c_i$

MH estimate: $\hat{OR}_{pool} = \frac{1}{W} \sum_{i=1}^{K} w_i \cdot \hat{OR}_i$

where $w_i = b_i c_i / N_i$

$$W = \sum_{i=1}^{K} w_i$$

Mantel-Haenszel Methods

(2) Confidence interval for common OR

- Robins-Greenland-Breslow expression for var(ln(OR_{MH})) based on estimating equation theory

- Robust: valid for "large" frequencies or "large" number of strata

(3) Test of pooled odds ratio

H_0: common odds ratio is 1.0

H_A: common odds ratio $\neq$ 1.0

Under H_0, the MH test statistic has approximately a Chi-squared distribution with 1 *df*.

Mantel-Haenszel Methods

(4) **Test of effect modification** (homogeneity, interaction)

H_0: $OR_1 = OR_2 = \ldots = OR_K$
H_A: not all stratum-specific OR's are equal

- MH-test statistic compares stratum-specific ln(OR)s and pooled ln(OR), distribution under H_0 is Chi-squared with K-1 *df*
- Breslow & Day derived an alternative Chi-squared test statistic
- The B-D test should only be used with a "small" number of "large" tables

.

Mantel-Haenszel Methods – STATA Example

Lung Cancer and heavy drinking study

```
. list
     +---------------------------------+
     | cancer   drink   number   smoke |
     |---------------------------------|
  1. |      1       1       24       1 |
  2. |      1       0        6       1 |
  3. |      0       1      776       1 |
  4. |      0       0      194       1 |
  5. |      1       1        9       0 |
  6. |      1       0       21       0 |
  7. |      0       1      891       0 |
  8. |      0       0     2079       0 |
     +---------------------------------+
. cc cancer drink [freq=number], by(smoke) bd

         Smoker |       OR      [95% Conf. Interval]    M-H Weight
----------------+-------------------------------------------------
              0 |          1     .4015748   2.288393         6.237 (exact)
              1 |          1     .3911965   3.033018         4.656 (exact)
----------------+-------------------------------------------------
          Crude |   1.666533     .9677794   2.892949               (exact)
   M-H combined |          1     .5521991   1.810941
---------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =     0.00  Pr>chi2 = 1.0000
Test of homogeneity (B-D)      chi2(1) =     0.00  Pr>chi2 = 1.0000

                 Test that combined OR = 1:
                             Mantel-Haenszel chi2(1) =      0.00
                                                Pr>chi2 =    1.0000
```

Stratified Contingency Tables

Example 2:

Bishop (1969) described a study to investigate the effect of the length of antenatal care, and the place where care is received, on infant survival past the first month of life.

	Clinic			
	A		B	
Infant survival	Duration of antenatal care		Duration of antenatal care	
	<1 month	≥ 1month	<1 month	≥ 1month
dead	3	4	17	2
alive	176	293	197	23

Stratified Contingency Tables – Example 2

1) Pooled data, not controlling for clinic

	Duration of antenatal care	
Infant survival	<1 month	≥ 1 month
dead	20	6
alive	373	316

```
. cci 20 6 373 316

                                                        Proportion
                |   Exposed   Unexposed  |      Total     Exposed
----------------+------------------------+------------------------
          Cases |        20           6  |         26      0.7692
       Controls |       373         316  |        689      0.5414
----------------+------------------------+------------------------
          Total |       393         322  |        715      0.5497
                |                        |
                |      Point estimate    |    [95% Conf. Interval]
                |------------------------+------------------------
     Odds ratio |          2.82395       |    1.075539    8.689846 (exact)
 Attr. frac. ex. |         .6458861       |     .070234    .8849232 (exact)
 Attr. frac. pop |         .4968354       |
                +-------------------------------------------------
                             chi2(1) =     5.26  Pr>chi2 = 0.0219
```

Stratified Contingency Tables – Example 2

2) Stratified by clinic

Clinic A

	Duration of antenatal care	
Infant survival	<1 month	≥ 1 month
dead	3	4
alive	176	293

```
. cci 3 4 176 293
                                                        Proportion
                |   Exposed   Unexposed  |      Total     Exposed
----------------+------------------------+------------------------
          Cases |         3           4  |          7      0.4286
       Controls |       176         293  |        469      0.3753
----------------+------------------------+------------------------
          Total |       179         297  |        476      0.3761
                |                        |
                |      Point estimate    |    [95% Conf. Interval]
                |------------------------+------------------------
     Odds ratio |          1.24858       |    .1807193    7.469567 (exact)
 Attr. frac. ex. |         .1990899       |   -4.533444    .8661234 (exact)
 Attr. frac. pop |         .0853242       |
                +-------------------------------------------------
                             chi2(1) =     0.08  Pr>chi2 = 0.7726
```

Stratified Contingency Tables – Example 2

Clinic B

	Duration of antenatal care	
Infant survival	<1 month	≥ 1 month
dead	17	2
alive	197	23

```
. cci 17 2 197 23

                                                      Proportion
                |   Exposed   Unexposed  |      Total     Exposed
----------------+------------------------+------------------------
          Cases |        17           2  |         19      0.8947
       Controls |       197          23  |        220      0.8955
----------------+------------------------+------------------------
          Total |       214          25  |        239      0.8954
                |                        |
                |      Point estimate    |    [95% Conf. Interval]
                |------------------------+------------------------
     Odds ratio |         .9923858       |    .2126309    9.403103 (exact)
 Prev. frac. ex. |         .0076142       |   -8.403103    .7873691 (exact)
 Prev. frac. pop |         .0068182       |
                +-------------------------------------------------
                             chi2(1) =     0.00  Pr>chi2 = 0.9922
```

Mantel-Haenszel Methods – Example 2

Infant mortality study

```
. list
```

	clinic	length	death	count
1.	0	1	1	3
2.	0	1	0	176
3.	0	0	1	4
4.	0	0	0	293
5.	1	1	1	17
6.	1	1	0	197
7.	1	0	1	2
8.	1	0	0	23

```
. cc death length [freq=count], by(clinic) bd tarone

          clinic |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |    1.24858      .1807193   7.469567     1.478992 (exact)
               1 |   .9923858      .2126309   9.403103     1.648536 (exact)
-----------------+-------------------------------------------------
           Crude |    2.82395      1.075539   8.689846              (exact)
    M-H combined |   1.113539      .3759998   3.297789
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =     0.04  Pr>chi2 = 0.8339
Test of homogeneity (B-D)      chi2(1) =     0.04  Pr>chi2 = 0.8338
Test of homogeneity (Tarone)   chi2(1) =     0.04  Pr>chi2 = 0.8339

                  Test that combined OR = 1:
                              Mantel-Haenszel chi2(1) =      0.04
                                              Pr>chi2 =    0.8442
```

Mantel-Haenszel Methods – Example 2

Using Stata **mhodds** command:

```
. mhodds death length clinic [freq=count]

Mantel-Haenszel estimate of the odds ratio
Comparing length==1 vs. length==0, controlling for clinic


     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       1.113539       0.04        0.8442         0.380670   3.257330
```

```
. mhodds death length [freq=count], by(clinic)

Maximum likelihood estimate of the odds ratio
Comparing length==1 vs. length==0
by clinic

   clinic | Odds Ratio        chi2(1)         P>chi2       [95% Conf. Interval]
----------+-------------------------------------------------------------------
        0 |   1.248580           0.08         0.7728         0.27574    5.65363
        1 |   0.992386           0.00         0.9922         0.21474    4.58619

    Mantel-Haenszel estimate controlling for clinic
    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       1.113539       0.04        0.8442         0.380670   3.257330
    ------------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(1)   =    0.04
                                     Pr>chi2   =  0.8341
```

Stratified Data – Summary

1. Compute stratum-specific measures
2. Evaluate stratum-specific estimates by a test of homogeneity. Consider test results in light of sample size.
3. If the homogeneity test result is non-significant then consider a common estimate, pooling across all strata
   (a) calculate an overall (common) summary (OR)
   (b) test for significant association
   (c) calculate confidence interval

Stratified Data – Summary

4. If the homogeneity test result is significant then we are concerned that the ORs vary across strata. We may

   (a) If the direction of association (+) is same and the difference is small in magnitude, then
   - proceed as in 3 above (calculating average summary)
   - report on the test of homogeneity.

   (b) If the direction of the association is different, then
   - report results from test of homogeneity
   - report stratum-specific measures and confidence intervals.
   - does the average make sense at all?

Mantel-Haenszel – Matched Data

Suppose we compute the Mantel-Haenszel estimator for pair-matched data …

For pair i, the only possible tables are …

	E+	E-	Total
D+	0	1	1
D-	0	1	1
Total	0	2	2

$a_i \times d_i = 0$

$b_i \times c_i = 0$

	E+	E-	Total
D+	1	0	1
D-	0	1	1
Total	1	1	2

$a_i \times d_i = 1$

$b_i \times c_i = 0$

	E+	E-	Total
D+	1	0	1
D-	1	0	1
Total	2	0	2

$a_i \times d_i = 0$

$b_i \times c_i = 0$

	E+	E-	Total
D+	0	1	1
D-	1	0	1
Total	1	1	2

$a_i \times d_i = 0$

$b_i \times c_i = 1$

Mantel-Haenszel – Matched Data

In the case of paired data, the Mantel-Haenszel odds ratio estimate is given by :

$$\hat{\psi}_{MH} = \frac{\sum_i a_i d_i / 2}{\sum_i b_i c_i / 2} = \frac{n_{10}}{n_{01}}$$

n_{10} = # pairs with E = 1 for “case” , E = 0 for “control”

n_{01} = # pairs with E = 0 for “case”, E = 1 for “control”

(refer back to slide 65)

Ille-et-Vilaine Case-control Study

Cases: 200 males diagnosed with esophageal cancer in one of the regional hospitals in French department of Ille-et-Vilaine (Brittany) between Jan 1972 and Apr 1974

Controls: Random sample of 778 adult males from electoral lists in each commune (775 with usable data)

Exposures: Detailed dietary interview on consumption of various foods, tobacco and alcoholic beverages

Background: Brittany was a known "hot spot" of esophageal cancer in France and also had high levels of alcohol consumption, particularly of the local (often homemade) apple brandy known as Calvados

Reference: Tuyns AJ, Pequinot G, Jensen OM. (1977) Le cancer de l'oesophage en Ille-et-Vilaine en fonction des niveaux de consommation d'alcohol et de tabac. *Bull Canc* **64**: 45-60.

Ille-et-Vilaine Case-control Study

Scientific Questions:

- What is the estimated effect of 80g+ alcohol consumption on the risk of cancer, controlling for age?
- Do these data provide strong evidence that there is an association between alcohol consumption and the risk of cancer, after adjusting for age?
- Is the effect of alcohol consumption the same for each age category?

Ille-et-Vilaine Data

COL	VAR	RANGE/VALUES
1-2	Age group	1 = 25-34 (years) 2 = 35-44 3 = 45-54 4 = 55-64 5 = 65-74 6 = 75+
3-4	Alcohol	1 = 0-39 (gms/day) 2 = 40-79 3 = 80-119 4 = 120+
5-6	Tobacco	1 = 0-9(gms/day) 2 = 10-19 3 = 20-29 4 = 30+
7-9	Cases	Number of esphogeal cancer cases
10-12	Controls	Number of population controls

Ille-et-Vilaine Data

Read data into Stata

```
. infix age 1-2 alc 3-4 tob 5-6 count1 7-9 count0 10-12 using "tuyns.txt"
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	88	3.386364	1.650021	1	6
alc	88	2.454545	1.123511	1	4
tob	88	2.409091	1.120718	1	4
count1	88	2.272727	2.753169	0	17
count0	88	8.806818	12.13512	0	60

Convert data into better format for tabling

```
. reshape long count, i(age alc tob) j(case)
. expand
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	975	3.271795	1.386713	1	6
alc	975	1.853333	.9063016	1	4
tob	975	1.765128	.9777895	1	4
case	975	.2051282	.4040025	0	1

Ille-et-Vilaine Data

Label variable values:

```
. label define agelab 1 "25-34" 2 "35-44" 3 "45-54" 4 "55-64" 5 "65-
   74" 6 "75+"
. label define alclab 1 "0-39" 2 "40-79" 3 "80-119" 4 "120+"
. label define toblab 1 "0-9" 2 "10-19" 3 "20-29" 4 "30+"
. label values age agelab
. label values alc alclab
. label values tob toblab
```

Ille-et-Vilaine Data

Age distribution of cases and controls:

```
. tab age case, col
           |         case
       age |      0.00       1.00 |     Total
-----------+----------------------+----------
     25-34 |       115          1 |       116
           |     14.84       0.50 |     11.90
-----------+----------------------+----------
     35-44 |       190          9 |       199
           |     24.52       4.50 |     20.41
-----------+----------------------+----------
     45-54 |       167         46 |       213
           |     21.55      23.00 |     21.85
-----------+----------------------+----------
     55-64 |       166         76 |       242
           |     21.42      38.00 |     24.82
-----------+----------------------+----------
     65-74 |       106         55 |       161
           |     13.68      27.50 |     16.51
-----------+----------------------+----------
       75+ |        31         13 |        44
           |      4.00       6.50 |      4.51
-----------+----------------------+----------
     Total |       775        200 |       975
           |    100.00     100.00 |    100.00
```

Ille-et-Vilaine Data

Population age and alcohol frequencies in controls:

```
. tab age alc if case==0, col
```

age	alc 0-39	40-79	80-119	120+	Total
25-34	61	45	5	4	115
	15.80	16.07	5.75	18.18	14.84
35-44	88	76	20	6	190
	22.80	27.14	22.99	27.27	24.52
45-54	77	61	27	2	167
	19.95	21.79	31.03	9.09	21.55
55-64	77	62	19	8	166
	19.95	22.14	21.84	36.36	21.42
65-74	60	28	16	2	106
	15.54	10.00	18.39	9.09	13.68
75+	23	8	0	0	31
	5.96	2.86	0.00	0.00	4.00
Total	386	280	87	22	775
	100.00	100.00	100.00	100.00	100.00

Ille-et-Vilaine Data

- **Scoring may help see patterns**

```
. tab age, summarize(alc) noobs
```

age	Summary of alc: Mean	Std. Dev.	Freq.
25-34	1.60	0.77	116
35-44	1.75	0.83	199
45-54	1.96	0.91	213
55-64	2.02	0.99	242
65-74	1.84	0.89	161
75+	1.57	0.87	44
Total	1.85	0.91	975

Alcohol Exposure & Esophageal Cancer

Consider alcohol "exposure" as a factor with two levels:

```
. generate alcexp = alc
. recode alcexp 1/2=0  3/4=1
. cc case alcexp
```

```
                                                      Proportion
                |   Exposed   Unexposed  |      Total     Exposed
----------------+------------------------+------------------------
          Cases |        96         104  |        200       0.4800
       Controls |       109         666  |        775       0.1406
----------------+------------------------+------------------------
          Total |       205         770  |        975       0.2103
                |                        |
                |      Point estimate    |    [95% Conf. Interval]
                |------------------------+------------------------
     Odds ratio |         5.640085       |    3.937435    8.061794 (exact)
 Attr. frac. ex. |         .8226977       |    .7460276    .8759581 (exact)
 Attr. frac. pop |         .3948949       |
                +-------------------------------------------------
                             chi2(1) =   110.26  Pr>chi2 = 0.0000
```

Ille-et-Vilaine Data

```
. bysort age: tab case alcexp

-> age = 25-34
           |        alcexp
      case |         0          1 |     Total
-----------+----------------------+----------
      0.00 |       106          9 |       115
      1.00 |         0          1 |         1
-----------+----------------------+----------
     Total |       106         10 |       116

-> age = 35-44
           |        alcexp
      case |         0          1 |     Total
-----------+----------------------+----------
      0.00 |       164         26 |       190
      1.00 |         5          4 |         9
-----------+----------------------+----------
     Total |       169         30 |       199

-> age = 45-54
           |        alcexp
      case |         0          1 |     Total
-----------+----------------------+----------
      0.00 |       138         29 |       167
      1.00 |        21         25 |        46
-----------+----------------------+----------
     Total |       159         54 |       213

-> age = 55-64
           |        alcexp
      case |         0          1 |     Total
-----------+----------------------+----------
      0.00 |       139         27 |       166
      1.00 |        34         42 |        76
-----------+----------------------+----------
     Total |       173         69 |       242
```

Ille-et-Vilaine Data

```
. bysort age: tab case alcexp (continued)


-> age = 65-74
           |        alcexp
      case |         0          1 |     Total
-----------+----------------------+----------
      0.00 |        88         18 |       106
      1.00 |        36         19 |        55
-----------+----------------------+----------
     Total |       124         37 |       161

-> age = 75+
           |        alcexp
      case |         0          1 |     Total
-----------+----------------------+----------
      0.00 |        31          0 |        31
      1.00 |         8          5 |        13
-----------+----------------------+----------
     Total |        39          5 |        44
```

Mantel-Haenszel – Ille-et-Vilaine

- Using Stata **cc** command

```
. cc case alcexp, by(age) bd
```

```
             age |       OR      [95% Conf. Interval]    M-H Weight
-----------------+-------------------------------------------------
           25-34 |          .           0          .             0 (exact)
           35-44 |   5.046154     .9268664   24.86538      .6532663 (exact)
           45-54 |   5.665025     2.632894   12.16536      2.859155 (exact)
           55-64 |   6.359477     3.299319   12.28473      3.793388 (exact)
           65-74 |   2.580247     1.131489   5.857261      4.024845 (exact)
             75+ |          .     4.388738          .             0 (exact)
-----------------+-------------------------------------------------
           Crude |   5.640085     3.937435   8.061794               (exact)
    M-H combined |   5.157623     3.562131   7.467743
-----------------------------------------------------------------
Test of homogeneity (B-D)      chi2(5) =     9.32  Pr>chi2 = 0.0968

                                       Test that combined OR = 1:
                                       Mantel-Haenszel chi2(1) =     85.01
                                              Pr>chi2 =    0.0000
```

Mantel-Haenszel – Ille-et-Vilaine

- Using Stata **mhodds** command

```
. mhodds case alcexp age
Mantel-Haenszel estimate of the odds ratio
Comparing alcexp==1 vs. alcexp==0, controlling for age

    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       5.157623      85.01        0.0000         3.494918   7.611359
    ------------------------------------------------------------------
```

Mantel-Haenszel – Ille-et-Vilaine

```
. mhodds case alcexp, by(age)
Mantel-Haenszel estimate of the odds ratio
Comparing alcexp==1 vs. alcexp==0, controlling for age
by age

------------------------------------------------------------------------------
      age | Odds Ratio     chi2(1)      P>chi2    [95% Conf. Interval]
----------+-------------------------------------------------------------------
    25-34 |          .       10.60      0.0011            .          .
    35-44 |   5.046154        6.32      0.0119      1.23889   20.55361
    45-54 |   5.665025       25.94      0.0000      2.66472   12.04350
    55-64 |   6.359477       38.74      0.0000      3.25661   12.41873
    65-74 |   2.580247        6.27      0.0123      1.19482    5.57211
      75+ |          .       13.15      0.0003            .          .
------------------------------------------------------------------------------

    Mantel-Haenszel estimate controlling for age and age
    ------------------------------------------------------------------
     Odds Ratio    chi2(1)        P>chi2        [95% Conf. Interval]
    ------------------------------------------------------------------
       5.157623      85.01        0.0000         3.494918   7.611359
    ------------------------------------------------------------------

Test of homogeneity of ORs (approx): chi2(5)   =    8.28
                                     Pr>chi2   =  0.1416
```

REVIEW

- R x C contingency table
  - Test for homogeneity (Pearson chi-squared)
  - Test for trend (Cochran-Armitage)
- Single 2 x 2 table
  - Different sampling schemes
    - Cohort (row totals fixed)
    - Case-control (column totals fixed)
    - Cross-sectional (grand total fixed)
  - Different measures of association
    - RD (Designs 1 & 3)
    - RR (Designs 1 & 3)
    - OR (Designs 1, 2 & 3)
  - Test of association
    - Pearson chi-squared
    - McNemar's
    - Fisher's exact
  - Agreement
    - Kappa

REVIEW

- Series of 2 x 2 tables
  - Mantel-Haenszel (combined) OR estimate
  - Mantel-Haenszel test for association

    H_o: OR = 1

    H_A: OR constant, $\neq 1$
  - Breslow-Day Tests for Homogeneity (Interaction, Effect Modification)
- Paired binary data as extreme case of stratification of 2 x 2 tables
- These simple methods have served epidemiologists well for many years, and still do!