EXAMPLE

The Simpson matter

DNA statistics in the Simpson matter

Bruce S. Weir

On 3 October 1995, O.J. Simpson was acquitted of two murders in spite of very strong DNA evidence linking his blood to the crime. Although numerical statements describing the strength of this evidence were made, the DNA profiles included so many loci that the need for presenting numbers in this case, and in others using similarly high numbers of loci, is probably unnecessary. If numbers are to be presented, however, they should be given in the form of likelihood ratios.

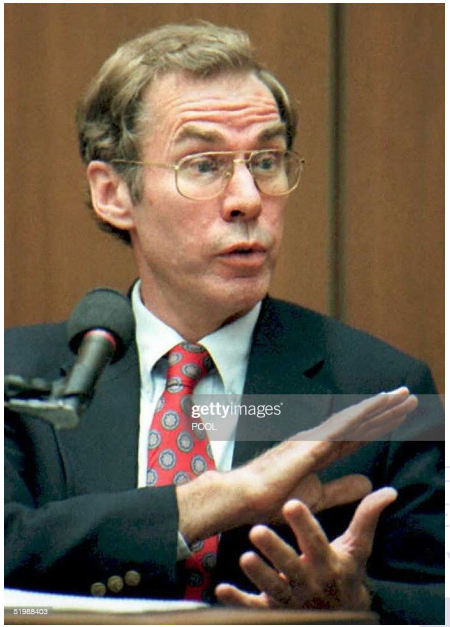
DNA profiling involved Southern analysis of a series of variable number tandem repeat (VNTR) loci that were typed for restriction fragment length polymorphisms (RFLPs) as well as analysis of several other loci amplified by the polymerase chain reaction (PCR). The LAPD analysed only the PCR-amplified locus, *DQα*. CMD used five VNTR loci: *D1S7, D2S44, D7S21, D7S22* and *D12S11*, the PCR locus *DQα* and the PCR Amplitype™ PM system, which is a multiplexed set of five loci: low density lipoprotein receptor (*LDLR*), glycophorin A (*GYPA*), haemoglobin G gammaglobin (*HBGG*), group-specific component (*Gc*) and *D7S8*. (CMD typed the six PCR loci with a reverse dot blot approach using allele-specific oligonucleotide probes immobilized on a nylon membrane strip.) DOJ used 11 VNTR loci with RFLP typing: *D1S7, D1S339, D2S44, D4S139, D5S110, D6S132, D7S467,*

Bruce Weir
NAS member
Director of SISG!

EXAMPLE

The Simpson matter

From my perspective as the statistical expert witness for the prosecution, I am primarily concerned about the way in which the statistical interpretation of matching profiles is presented in court. Of course, I may not be the best person to address these issues, since, whatever the jury may have thought about my testimony, the commentators were not impressed: "Dry as sand and about as digestible," said Peter Aranella (*Los Angeles Times*, June 23, 1995). This commentary is intended to aid the digestion.



EXAMPLE

The Simpson matter

hat

Power of DNA evidence is associated with the “extraordinary” probabilities of matches.

analysed. This occurs because prediction of the frequency that a particular profile occurs in a population is determined by multiplying together the frequencies of all the allelic components of the profile from a particular sample. In the Simpson case, DOJ estimated these allelic fre-

[Freq component 1] x [Freq component 2] x … x [Freq component n] = $\frac{1}{\text{[too big number]}}$

analysed for the trial. Use of these calculations in her closing arguments allowed prosecutor Marcia Clark to state that the chance that a random person would have the profile found on the rear gate at Bundy (Fig. 1*a*) would be 1 in 57 billion.

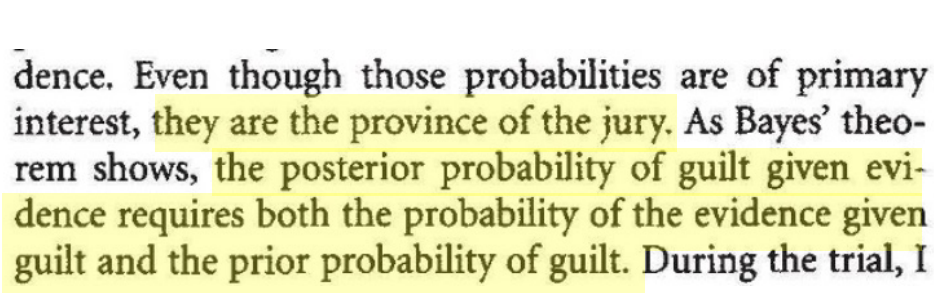
But there were only 5 billion people on the planet in 1995. Using **these** extraordinary numbers is absurd.

EXAMPLE

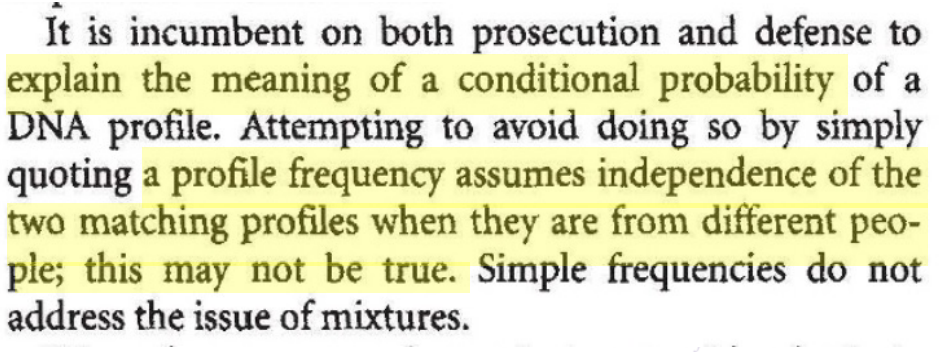
The Simpson matter

The proper way to do it is to use conditional probabilities and a likelihood ratio.

We might like to determine the conditional probabilities of **guilt** given the **evidence**. But we can't do that.



dence. Even though those probabilities are of primary interest, they are the province of the jury. As Bayes' theorem shows, the posterior probability of guilt given evidence requires both the probability of the evidence given guilt and the prior probability of guilt. During the trial, I



It is incumbent on both prosecution and defense to explain the meaning of a conditional probability of a DNA profile. Attempting to avoid doing so by simply quoting a profile frequency assumes independence of the two matching profiles when they are from different people; this may not be true. Simple frequencies do not address the issue of mixtures.

Instead, the right way to do it is to estimate the probability of the **evidence** given **guilt**.

EXAMPLE

The Simpson matter

tion, C^{**}, was that OS did not commit the murders. The strength of the DNA evidence should have been presented as the ratio of the conditional probabilities of that evidence under the two alternatives: $Pr(E|C)/Pr(E|C^{**})$. During the

$$Pr(E \mid C) / Pr(E \mid C^{**})$$

- Evidence
- Prosecution's argument: OJS committed murders
- Defense's argument: OJS did not commit murders

But Judge Ito did not allow presentation of likelihood ratios.

Also, based on ambiguity in genotyping of one of the samples, the defense insisted that the statistical testimony that was presented must include the possibility of a fourth person (beyond OJS, NBS, and RG).

EXAMPLE

The Simpson matter

require four contributors.) The requirement that likelihood ratios not be presented meant that only $Pr(E|C^{**})$ numbers were given, which ironically could have hurt the defense's case. For four unknown contributors, the prob-

Moreover, the inclusion of a fourth possible, unknown contributor in the sample also weakened the defense's case.

defense's case. For four unknown contributors, the probability for the RFLP profile in item 31 varies between 1 in 240 million and 1 in 2.7 billion, depending on the racial databases used. Under explanation C, the probability of

four in explanation C** for the denominator. Thus, if the defense had been willing to concede that one contributor was known, such as Bronco owner OS, then there would have been three unknown people in the denominator and two in the numerator. The likelihood ratio then becomes less than 1,000, a far cry from 1 in 240 million.

EXAMPLE

The Simpson matter

There were a lot of things that went wrong with interpreting DNA evidence testimony in the court, not to mention aspects outside statistical analysis of the DNA evidence.

The jurors for the most part did not understand, or know how to consider, the DNA testimony.

That part of the trail went on for 8 hours a day for almost 9 weeks.

RG. Trial commentators almost always transposed the conditional, however. For example, there was a 1 in 1,400 chance that the *DQα + D1S80* profile of the Bronco centre console could have had the type seen if it came from two people other than OS and RG. But Linda Deutsch of the Associated Press (June 26, 1995) misunderstood my testimony and presented it as giving "a chance of 1 in 1,400 that any two people in the population could be responsible for such a stain."

multiple contributors to a stain. For me, one of the most disappointing aspects of the Simpson trial was the refusal by the defense to acknowledge the standard way of presenting numbers for transfer evidence.

Stratified Contingency Tables

Session 8

Module 1 Probability & Statistical Inference

The Summer Institutes
DEPARTMENT OF BIOSTATISTICS
SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

Overview

1. **2 x 2 Tables**
    - Paired Binary Data
2. **Stratified Tables**
    - Confounding
    - Effect Modification

EXAMPLE

2 x 2 Tables

Epidemiological Applications: Matched Case Control Study

213 subjects with a history of acute myocardial infarction (AMI) were **matched** by age and sex with one of their siblings who did not have a history of AMI. The prevalence of a particular polymorphism was compared between the siblings.

Question 1 Is there an association between the polymorphism prevalence and AMI?

Question 2 If there is an association then what is the magnitude of the effect?

EXAMPLE

2 x 2 Tables

Epidemiological Applications: Matched Case Control Study

Q: Can't we simply use Pearson's χ^2 Test to assess whether this is evidence for an association?

A: NO!!! Pearson's χ^2 test assumes that the columns are **independent** samples. In this design the 213 people with AMI are genetically related to the 213 people w/o AMI. This is an example of **paired binary data**.

Exposure Status	Disease Status: AMI	Disease Status: no AMI	TOTAL
Carrier	96	87	183
Noncarrier	117	126	243
TOTAL	213	213	426

EXAMPLE

2 x 2 Tables

Epidemiological Applications: Paired Binary Data

73 pairs
AMI = carrier
No AMI = carrier

14 pairs
AMI = not carrier
No AMI = carrier

23 pairs
AMI = carrier
No AMI = not carrier

103 pairs
AMI = not carrier
No AMI = not carrier

For **paired binary data** we display the results as shown in the table. The cells now represent information about each **pair**.

This analysis explicitly recognizes the heterogeneity of subjects.

The **concordant pairs** (73 and 103) provide no information about the association between AMI and the polymorphism.

🔑 The information regarding the association is in the **discordant pairs,** 14 and 23.

no AMI \ AMI	carrier	non-carrier	TOTAL
carrier	73	14	87
noncarrier	23	103	126
TOTAL	96	117	213

2 x 2 Tables

Epidemiological Applications: Paired Binary Data

For **paired binary data** we display the results as shown in the table.

This analysis explicitly recognizes the heterogeneity of subjects.

$$p_1 = P(\text{carrier} \mid \text{AMI}) = p_{11} + p_{01}$$

$$p_0 = P(\text{carrier} \mid \text{No AMI}) = p_{11} + p_{10}$$

$$H_0 : p_1 = p_0$$

$$H_A : p_1 \neq p_0$$

🔑 The information for testing these hypotheses is contained in the **discordant pairs** (0,1) and (1,0).

no AMI \ AMI	1	0	TOTAL
1	n_{11}	n_{10}	
0	n_{01}	n_{00}	
TOTAL			n

$n_{11}/n =$ **p_{11}**	$n_{10}/n =$ **p_{10}**
$n_{01}/n =$ **p_{01}**	$n_{00}/n =$ **p_{00}**

2 x 2 Tables

Epidemiological Applications: McNemar's Test for Paired Binary Data

Under the null hypothesis we expect equal numbers of (0,1) pairs and (1,0) pairs. We can evaluate this hypothesis using or **McNemar's Test for Paired Binary Data**. The **McNemar's chi-squared statistic** is

$$X^2 = \frac{(n_{10} - n_{01})^2}{n_{10} + n_{01}} \sim \chi^2(1)$$

The **odds ratio** comparing the odds of carrier in those with AMI to odds of carrier in those w/o AMI is estimated by:

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

Confidence intervals can be obtained as described in Breslow and Day (1981), section 5.2, or in Armitage and Berry (1987), chapter 16.

EXAMPLE

Effect Modification

Stratified Tables

Often, **a third variable** influences the relationship between the two primary measures (e.g., disease and exposure).

Example (right):
Effect of seat belt use on car accident fatality

		Seat Belt: Worn	Seat Belt: Not Worn
Driver	Dead	10	20
	Alive	40	30
	TOTAL	50	50
	Fatality Rate	10/50 (20%)	20/50 (40%)

Risk difference = RD = 20 – 40 = -20%: “The fatality rate is 20% less if seatbelt is worn.”

EXAMPLE

Effect Modification

Stratified Tables

But suppose we also consider **impact speed**.

How does this affect the inference?

This is an example of effect modification or interaction:

Effects are different in subgroups of a third variable, and the overall effect is intermediate.

		< 40 mph Seat Belt		> 40 mph Seat Belt	
		Worn	Not Worn	Worn	Not Worn
Driver	Dead	3	2	7	18
	Alive	27	18	13	12
	TOTAL	30	20	20	30
	Fatality Rate	3/30 (10%)	2/20 (10%)	7/20 (35%)	18/30 (60%)

EXAMPLE

Effect Modification

Dependence on the effect measure used

Rate of fractures (over 5 years) by age and calcium level in drinking water.

Calcium Level	Age: 20-35 yrs	Age: 55-80 yrs	Overall (pooled)
High	1.1%	11.0%	7.8%
Low	3.3%	13.2%	10.0%
Risk Ratio	0.33	0.83	0.78
Risk Difference	-2.2%	-2.2%	-2.2%

Effect modification depends on the effect measure used!

Effect

There's evidence of effect modification on the risk ratio scale.

No effect

There's no evidence of effect modification on the risk difference scale.

Paws

Work through questions 1-2

EXAMPLE

Confounding

Suppose we are interested in the relationship between lung cancer incidence and heavy drinking.

We conduct a **prospective cohort study** where drinking status is determined at baseline and the cohort is followed for 10 years to determine cancer endpoints.

We also measure smoking status at baseline.

Pooled data
(Not controlled for smoking)

Lung Cancer Status	Heavy Drinker: Yes	Heavy Drinker: No	TOTAL
Yes	33	27	60
No	1667	2273	3940
TOTAL	1700	2300	4000

OR = (33 x 2273) / (1667 x 27) = **1.67**

EXAMPLE

Confounding

A higher proportion of heavy drinkers are smokers — **800/1700** vs **200/2300**

A higher proportion of lung cancer cases are smokers — **30/1000** vs **30/3000**

The comparison of heavy drinkers to not-heavy drinkers is really a comparison of smokers to nonsmokers!

Stratified data (by smoking)

Smokers

Heavy Drinker

Lung Cancer Status	Yes	No	TOTAL
Yes	24	6	30
No	776	194	970
TOTAL	800	200	1000

OR = 1

Nonsmokers

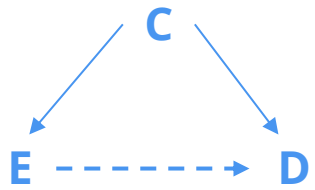
Heavy Drinker

Lung Cancer Status	Yes	No	TOTAL
Yes	9	21	30
No	891	2079	2970
TOTAL	900	2100	3000

OR = 1

Confounding

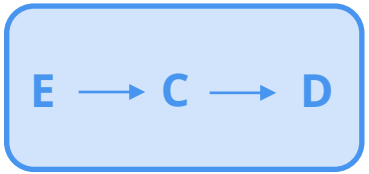
A **confounder** is associated with both the disease and exposure and is not in the causal path between disease and exposure.



An apparent association between E and D is completely explained by C.

C is a confounder.

A confounder is not a **mediator**.



The implicit assumption is that we want to know if E "causes" D

A simple, common example from genetics is the linked gene: we discover a gene which appears to be associated with disease … does it cause the disease or is it merely linked to the true causal gene?

Adjusting the OR via Stratification

Basic idea (also works for RR or RD)

- Compute separate OR for each stratum
- Assess homogeneity of OR's across strata
  Is there EM?
- Pool OR's: used weighted average
  Adjust for confounding
- Global test of pooled OR = 1
  Is there association, after adjustment
- Different methods of pooling, testing have been proposed.
  The Mantel-Haenszel methods are described in the following slides
- Same idea for RR and RD

End

Work through questions 3-4

Next slides: an example of how to adjust OR, RR, or RD given stratification

EXAMPLE

Stratified Contingency Tables

Mantel-Haenszel Methods

A 1985 study identified a group of **509 cancer cases and 489 controls** by mail questionnaire. The main purpose of the study was to look at the **effect of passive smoking on cancer risk**.

In the study **passive smoking** was defined as exposure to the cigarette smoke of a spouse who smoked at least one cigarette/day for at least 6 months.

One potential confounding variable **was smoking by the test subjects themselves** since personal smoking is related to both cancer risk and having a spouse that smokes.

Therefore, it was important to control for **personal smoking** before looking at the relationship between passive smoking and cancer risk.

Example from section 13.5 of Rosner biostatistics textbook; good resource for applying Mantel-Haenszel methods

EXAMPLE

Stratified Contingency Tables

Mantel-Haenszel Methods

Pooled data (not controlled for personal smoking)

Cancer Status	Passive Smoking: Yes	Passive Smoking: No	TOTAL
Case	281	228	509
Control	210	279	489
TOTAL	491	507	998

OR = 1.64

p-value = 0.0001

For information on how to complete these calculations in R:
https://a-little-book-of-r-for-biomedical-statistics.readthedocs.io/en/latest/src/biomedicalstats.html

EXAMPLE

Stratified Contingency Tables

Mantel-Haenszel Methods

Stratified data (by personal smoking)

Personal Smoking: Smokers

Passive Smoking

Cancer Status	Yes	No	TOTAL
Case	161	117	278
Control	130	124	254
TOTAL	291	241	532

OR = 1.31

p-value = 0.1192

Personal Smoking: Nonsmokers

Passive Smoker

Cancer Status	Yes	No	TOTAL
Case	120	111	231
Control	80	155	235
TOTAL	200	266	466

OR = 2.09

p-value = 0.0001

Stratified Contingency Tables

Mantel-Haenszel Methods

Q: How can we combine the information from both stratum-specific tables to obtain an overall test of significance that takes account of the stratification?

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

Disease \ Exposure	yes	no	TOTAL
yes	a_i	b_i	a_i+b_i
no	c_i	d_i	c_i+d_i
TOTAL	a_i+c_i	b_i+d_i	N_i

where i = 1,2,...,K is the number of strata.

Stratified Contingency Tables

Mantel-Haenszel Methods

(1) Test of effect modification
(heterogeneity, interaction)

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

(2) 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:

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

(3) Test of common odds ratio

H_0: common odds ratio is 1.0
H_A: common odds ratio ≠ 1.0

EXAMPLE

Stratified Contingency Tables

Mantel-Haenszel Methods

	case	passive	number	smoke
1.	1	1	120	0
2.	1	0	111	0
3.	0	1	80	0
4.	0	0	155	0
5.	1	1	161	1
6.	1	0	117	1
7.	0	1	130	1
8.	0	0	124	1

Entering the stratum-specific data

```
. cc case passive [freq=number], by(smoke) bd

Personal Smoking |       OR      [95% Conf. Interval]     M-H Weight
-----------------+-------------------------------------------------
               0 |   2.094595      1.41754   3.097165      19.05579  (exact)
               1 |   1.312558     .9184614   1.875813      28.59023  (exact)
-----------------+-------------------------------------------------
           Crude |   1.637406     1.265013   2.119599                (exact)
     M-H combined |   1.625329     1.263955   2.090024
-----------------+-------------------------------------------------
Test of homogeneity (M-H)       chi2(1) =     3.27  Pr>chi2 = 0.0706
Test of homogeneity (B-D)       chi2(1) =     3.27  Pr>chi2 = 0.0704

                Test that combined OR = 1:
                        Mantel-Haenszel chi2(1) =     14.42
                                          Pr>chi2 =    0.0001
```

Calculating the pooled OR and testing whether it is different from 1