The Analysis of $2 \times 2 \times 2$ (Multiway) Contingency Tables: Explaining Simpson’s Paradox and Demonstrating Racial Bias in the Imposition of the Death Penalty

It is the mark of a truly intelligent person to be moved by statistics – George Bernard Shaw

September 14, 2016
1) Simpson’s Paradox example from *Car Talk*: Bluto and Popeye
2) Berkeley sex bias in graduate school admissions
3) Death penalty imposition and race
An enjoyable diversion on Saturday mornings is the NPR radio show, *Car Talk*, with Click and Clack, The Tappet Brothers (aka Ray and Tom Magliozzi).

A regular feature of the show, besides giving advice on cars, is The Puzzler; a recent example on September 22, 2012 gives a nice introductory example of one main topic of this chapter, Simpson’s paradox.

It is called, Take Ray Out to the Ball Game, and is given in full in your Module Three reading.

It is about two rookie players, Bluto and Popeye, who started the season on opening day and made a wager as to which one would have the best batting average at the end of the season.
Putting the data about Bluto and Popeye in the form of a $2 \times 2$ table that gives batting averages both before and after the All-Star break as well as for the full year should help see what is happening:

<table>
<thead>
<tr>
<th></th>
<th>Before Break</th>
<th>After Break</th>
<th>Full Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bluto</td>
<td>$\frac{150}{500} = .300$</td>
<td>$\frac{40}{100} = .400$</td>
<td>$\frac{190}{600} = .317$</td>
</tr>
<tr>
<td>Popeye</td>
<td>$\frac{25}{100} = .250$</td>
<td>$\frac{187.5}{500} = .375$</td>
<td>$\frac{212.5}{600} = .354$</td>
</tr>
</tbody>
</table>

Thus, the batting averages of Popeye before and after the break (.250 and .375) can be less than for Bluto (.300 and .400), even though for the full year, Popeye’s average of .354 is better than Bluto’s .317.
This type of counterintuitive situation is referred to as a “reversal paradox” or more usually by the term, “Simpson’s paradox.”

The unusual phenomenon presented by the example above occurs frequently in the analysis of multiway contingency tables.

Basically, various relations that appear to be present when data are conditioned on the levels of one variable, either disappear or change “direction” when aggregation occurs over the levels of the conditioning variable.
Berkeley Sex Bias Example

A well-known real-life example is the Berkeley sex bias case applicable to graduate school (Bickel, Hammel, & O’Connell, 1975).

The table below shows the aggregate admission figures for the fall of 1973:

<table>
<thead>
<tr>
<th></th>
<th>Number of applicants</th>
<th>Percent admitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>8442</td>
<td>44</td>
</tr>
<tr>
<td>Women</td>
<td>4321</td>
<td>35</td>
</tr>
</tbody>
</table>

Given these data, there appears to be a *prima facie* case for bias because a lower percentage of women than men is admitted.
Although a bias seems to be present against women at the aggregate level, the situation becomes less clear when the data are broken down by major. Because no department is significantly biased against women, and in fact, most have a small bias against men, we have another instance of Simpson’s paradox.

Apparently, women tend to apply to competitive departments with lower rates of admission among qualified applicants (for example, English); men tend to apply to departments with generally higher rates of admission (for example, Engineering).
A different example showing a similar point can be given using data on the differential imposition of a death sentence depending on the race of the defendant and the victim. These data are from twenty Florida counties during 1976-1977 (Radelet, 1981):

<table>
<thead>
<tr>
<th>Defendant</th>
<th>Death Penalty</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>19 (12%)</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>17 (10%)</td>
<td>149</td>
<td></td>
</tr>
</tbody>
</table>

Because 12% of white defendants receive the Death penalty and only 10% of blacks, at this aggregate level there appears to be no bias against blacks.
But when the data are disaggregated, the situation appears to change:

<table>
<thead>
<tr>
<th>Victim</th>
<th>Defendant</th>
<th>Death Penalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>White</td>
<td>19 (13%) 132</td>
</tr>
<tr>
<td>White</td>
<td>Black</td>
<td>11 (17%) 52</td>
</tr>
<tr>
<td>Black</td>
<td>White</td>
<td>0 (0%) 9</td>
</tr>
<tr>
<td>Black</td>
<td>Black</td>
<td>6 (6%) 97</td>
</tr>
</tbody>
</table>

When aggregated over victim race, there is a higher percentage of white defendants (12%) receiving the death penalty than black defendants (10%), so apparently, there is a slight race bias against whites.
But when looking within the race of the victim, black defendants have the higher percentages of receiving the death sentence compared to white defendants (17% to 13% for white victims; 6% to 0% for black victims).

The conclusion is disconcerting: the value of a victim is worth more if white than if black, and because more whites kill whites, there appears to be a slight bias against whites at the aggregate level.

But for both types of victims, blacks are more likely to receive the death penalty.
The United States has had a troubled history with the imposition of the death penalty.

Two amendments to the Constitution, the Eighth and the Fourteenth, operate as controlling guidelines for how death penalties are to be decided on and administered (if at all). The Eighth Amendment prevents “cruel and unusual punishment”; the Fourteenth Amendment contains the famous “equal protection” clause:

No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.
Various Supreme Court rulings over the years have relied on the Eighth Amendment to forbid some punishments entirely and to exclude others that are excessive in relation to the crime or the competence of the defendant.

One of the more famous such rulings was in *Furman v. Georgia* (1972), which held that an arbitrary and inconsistent imposition of the death penalty violates both the Eighth and Fourteenth Amendments, and constitutes cruel and unusual punishment. This ruling lead to a moratorium on capital punishment throughout the United States that extended to 1976 when another Georgia case was decided in *Gregg v. Georgia* (1976).
The Analysis of 2 × 2 × 2 (Multiway) Contingency Tables: Explaining Simpson's Paradox and Demonstrating Racial Bias in the Imposition of the Death Penalty

Gregg v. Georgia (1976)

The Supreme Court case of *Gregg v. Georgia* reaffirmed the use of the death penalty in the United States.

It held that the imposition of the death penalty does not automatically violate the Eighth and Fourteenth Amendments. If the jury is furnished with standards to direct and limit the sentencing discretion, and the jury’s decision is subjected to meaningful appellate review, the death sentence may be constitutional.

If, however, the death penalty is mandatory, so there is no provision for mercy based on the characteristics of the offender, then it is unconstitutional.

There are many analyses done by Baldus et al. and others on the interrelation between the race of the victim and of the defendant and the imposition of the death penalty.

Most do not show an explicit Simpson’s paradox such as for the Radelet data of the last section, where a black defendant has a higher probability of receiving the death penalty compared to a white defendant.

But universally, the race of the victim plays a crucial part in death penalty imposition – when the victim is white, the probability of receiving the death penalty is substantially higher than for black victims.

The relative risks, for example, are all much greater than the value of 2.0 needed to legally assert specific causation.
In *McCleskey v. Kemp*, the Court held that despite statistical evidence of a profound racial disparity in application of the death penalty, such evidence is insufficient to invalidate a defendant’s death sentence.

The syllabus of this ruling is given in your Module Three readings.

To see additional contemporary commentary, an article by Anthony Lewis lamenting this ruling appeared in the *New York Times* (April 28, 1987), entitled “Bowing To Racism.”

In the course of his essay, Stevens comments on McCleskey v. Kemp and notes that Justice Powell (who wrote the majority opinion) in remarks he made to his biographer, said that he should have voted the other way in the McCleskey 5 to 4 decision.

The doctrine of stare decisis suggests that no amount of statistical evidence will ever be sufficient to declare the death penalty in violation of the “equal protection” clause of the Fourteenth Amendment.
In 1987, the Court held in McCleskey v. Kemp that it did not violate the Constitution for a state to administer a criminal justice system under which murderers of victims of one race received death sentences much more frequently than murderers of victims of another race.

The case involved a study by Iowa law professor David Baldus and his colleagues demonstrating that in Georgia murderers of white victims were eleven times more likely to be sentenced to death than were murderers of black victims.

We give extensive redactions of the majority opinion and dissent in your Module Three readings (by Justice Brennan) for McCleskey v. Kemp.