

## The Dade County spouse assault experiment

The Police Foundation and the Metro-Dade Police Department ran a big, complicated field experiment. The execution was excellent. The data analysis in Pate and Hamilton (1992) (hereafter PH), on the other hand, is a let-down. Randomization does not justify logistic regression, as we have seen. Your mission: carry out an analysis of the Dade County experimental data that is justified by the randomization.

## Detective work

A simple and effective approach is to compare rates of recidivism (that is, repeat spousal abuse) between the treatment group, who were assigned to arrest, and the control group, who were not. Also of interest is the same rate comparison, in each of two subgroups: unemployed subjects and employed subjects. To pull off these comparisons, you need the numbers in the following table:

	no_arrest	arrest	
unemployed	$n_{00}/N_{00}$	$n_{01}/N_{01}$	$n_{0\cdot}/N_{0\cdot}$
employed	$n_{10}/N_{10}$	$n_{11}/N_{11}$	$n_{1\cdot}/N_{1\cdot}$
	$n_{\cdot 0}/N_{\cdot 0}$	$n_{\cdot 1}/N_{\cdot 1}$	$n_{\cdot\cdot}/N_{\cdot\cdot}$

The  $N$ 's are total subject counts, while the  $n$ 's tally the corresponding number of recidivist subjects. The dots denote summation over an index; for example,  $n_{\cdot 0} = n_{00} + n_{10}$ .

You might expect to find these various counts in an appendix to the paper. But you will not. The only count directly reported in PH is  $N_{\cdot\cdot} = 907$ , the total number of randomized subjects. To figure out the  $N$ 's, use the data file `part6_907.txt`. This file is derived from a dataset available on the National Institute of Justice website. It has one row per subject.<sup>1</sup> With the help of its columns, you can cross-tabulate the total subject counts by employment and assigned-treatment status. In addition to the data file, you have been given an excerpt from the experiment's "Codebook," explaining what the columns are and how the numerical codes are interpreted.

- Compute all the  $N$ 's, including the margins.
- PH report the rate of unemployment among the subjects. Does your rate agree with theirs?

If the data file had a column for recidivism, you'd figure out the  $n$ 's in the same way. Alas, recidivism outcomes appear in a separate file in the dataset—"Part 4". Why don't we just match up the records in these two files? A quote from the Codebook:

---

<sup>1</sup>The original file had 916 rows, for reasons of little interest. Applying an imputation procedure in reverse, this has been reduced to 907.

“Each of the six data files contain at least one variable to identify cases. However, one common case identification variable is not present across all files. At the time of this release, ICPSR had not been successful in linking all files.”

Welcome to applied statistics. As luck would have it, PH provide enough information to recover the  $n$ 's. See their Figure 1.

- Compute all the  $n$ 's, including the margins.
- In a part of the paper separate from Figure 1 and its discussion, PH report the rate of recidivism among arrestees and among non-arrestees. Do your rates agree with theirs?

### Statistical work

PH draw several conclusions from their logistic-regression analyses:

- “Among employed suspects, arrest had a statistically significant deterrent effect on the occurrence of a subsequent assault.”
- “Among unemployed suspects. . . significant *increases* in subsequent assault were associated with arrest.”
- “[Among all suspects, there is] no statistically significant effect of arrest on the occurrence of a subsequent spouse assault.”

Evaluate each of these conclusions in turn, by comparing the relevant observed rates in your hard-won counts table. Report  $p$ -values justified by the randomization that took place. Applicable methods include Fisher's exact test and the two-sample test of equal binomial proportions.

One of the assumptions underlying Fisher's exact test: the total number of observed recidivists (overall, and in each employment-status subgroup) would not change if there had been a different randomization outcome in the Dade County experiment. Discuss whether this assumption is compatible with the Neyman model of the experiment.

On the other hand, the binomial test as rendered in textbooks concerns independent Bernoulli trials. We are not thinking of recidivism outcomes as random coin flips (unlike PH). Instead, the treatment assignment is what's random. How then can the randomization justify the textbook  $p$ -value?