STAT 215B Assignment 3

Xiaowei Zeng

February 7, 2024

Abstract

The data analysis in Pate and Hamilton (1992) is a let-down since randomization does not justify logistic regression. The main job in this assignment is to carry out an analysis of the Dade County experimental data that is justified by the randomization. The first section conducts some detective tabular analysis of the dataset. The second section employs Fisher exact test and binomial test on the data and then discusses the compatibility of hypotheses with the model in the Pate and Hamilton (PH) setting.

1 Detective Work

1.1 Dataset

I load the "part6_907.txt" in `R`, encode the variables according to the codebook with correspondings, and then add 3 variables to the data based on the `act` column: 1. `assign` (the assigned treatment); 2. `deliver` (the delivered treatment); 3. `comply` (compliance of assigned treatment and delivered treatment). The name and the description of the variables are depicted in Table 1.

Table 1: Variables description

Name	Description
caseno	Case Number
relat	Relationship (married/sep/div/boygirl)
sempl	Whether the suspect is (1) or is not (0) employed
act	Treatment assigned - Treatment delivered
assign	Treatment assigned (arrest/nonarrest)
deliver	Treatment delivered (arrest/nonarrest)
comply	Compliance or Crossover (comply/crossover)

We 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, we need the numbers in Table 2

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, $n_{\cdot 0} = n_{00} + n_{10}$.

Table 2: Contingency table

	nonarrest	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}	$n_{1\cdot}/N_{1\cdot}$
	$n_{\cdot 0}/N_{\cdot 0}$	$n_{\cdot 1}/N_{\cdot 1}$	$n_{\cdot\cdot}/N_{\cdot\cdot}$

1.2 Computing N's

Compute all the N's, including the margins (Table 3), where 139 refers to the unemployed and nonarrested suspects, 124 refers to the unemployed and arrested suspects, 306 refers to the employed and nonarrested suspects, 338 refers to the employed and arrested suspects; 263 refers to the unemployed suspects, 644 refers to the employed suspects; 445 refers to the nonarrested suspects, 462 refers to the arrested suspects; 907 refers to all suspects.

Table 3: Contingency table for N's

	nonarrest	arrest	
unemployed	139	124	263
employed	306	338	644
	445	462	907

1.3 Rate of Unemployment among Suspects

On the right side of Page 693, PH say "Approximately 29 percent of the suspects were unemployed at the time of the presenting incident".

My rate is computed as

$$\frac{N_{0\cdot}}{N_{\cdot\cdot}} = \frac{263}{907} \approx 29.00\%,$$

which aligns well with their report.

1.4 Computing n's

PH provide enough information to recover the n's (Figure 1 in PH). Compute all the n's (Table 4), where 10 refers to the unemployed and nonarrested recidivists, 21 refers to the unemployed and arrested recidivists, 38 refers to the employed and nonarrested recidivists, 21 refers to the employed and arrested recidivists; 31 refers to the unemployed recidivists, 59 refers to the employed recidivists; 48 refers to the nonarrested recidivists, 42 refers to the arrested recidivists; 90 refers to all recidivists.

Table 4: Contingency table for n's

	nonarrest	arrest	
unemployed	10	21	31
employed	38	21	59
	48	42	90

1.5 Rate of Recidivism among Arrestees and Non-arrestees

Divide Table 4 by Table 3, we get Table 5. The percentages in the brackets are the rates of recidivism reported by PH.

Table 5: Table for rate of recidivism

	nonarrest	arrest	
unemployed	7.2% (7.1%)	16.9% (16.7%)	11.8%
employed	12.4% (12.3%)	6.2% (6.2%)	9.2%
	10.8% (10.6%)	9.1% (9.0%)	9.9%

It turns out that the rates are similar, but not the same (within 0.2%). Similarly, compare the ratio in Figure 1 of PH and the ratio computed on my own, there's also a slight difference. This may be attributed to the difference in dataset. For example, perhaps due to censoring or unexpected withdrawn from the research, the number of unemployed non-arrested suspects N_{01} is not 139 when PH compute the rate, leading to a slight bias.

2 Statistical work

PH draw several conclusions from their logistic-regression analyses:

1. "Among employed suspects, arrest had a statistically significant deterrent effect on the occurrence of a subsequent assault."
2. "Among unemployed suspects, significant increases in subsequent assault were associated with arrest."
3. "Among all suspects, there is no statistically significant effect of arrest on the occurrence of a subsequent spouse assault."

We will evaluate each of these conclusions in turn, by comparing the relevant observed rates in my counts table. In Secion 2.1 to Section 2.3, we neither discuss the validity of the hypothesis tests on the dataset nor on the experiment. We only conduct Fisher exact test and binomial proportion test and use the p-values to conclude that either the results support the claim or not support the claim. Then in Section 2.4 and Section 2.5, we discuss the validity of model and test in detail, and then give final conclusions on PH's claims.

2.1 Evaluation of Conclusion 1

I first construct a contingency table with treatment and outcome variables among all the employed suspects (Table 6).

The p-value of **one-sided** Fisher exact test is $0.0047 < 0.05$; the p-value of **one-sided** binomial proportion test is $0.0048 < 0.05$. Both of the p-values indicate that among employed suspects, arrest has a statistically significant deterrent effect on the occurrence of a subsequent assault, which supports PH's conclusion 1.

Table 6: Contingency table among employed suspects

	nonrecid	recid
nonarrest	268	38
arrest	317	21

2.2 Evaluation of Conclusion 2

I construct a contingency table with treatment and outcome variables among all the unemployed suspects (Table 7).

Table 7: Contingency table among unemployed suspects

	nonrecid	recid
nonarrest	129	10
arrest	103	21

The p-value of **one-sided** Fisher exact test is $0.0118 < 0.05$; the p-value of **one-sided** binomial proportion test is $0.0121 < 0.05$. Both of the p-values indicate that among unemployed suspects, arrest has a statistically significant positive effect on the occurrence of a subsequent assault, which supports PH's conclusion 2.

2.3 Evaluation of Conclusion 3

I construct a contingency table with treatment and outcome variables among all the suspects without regard to the employment status (Table 8).

Table 8: Contingency table among all suspects

	nonrecid	recid
nonarrest	397	48
arrest	420	42

The p-value of **two-sided** Fisher exact test is $0.44 > 0.05$; the p-value of **two-sided** binomial proportion test is $0.46 > 0.05$. Both of the p-values indicate that among all suspects, there is no statistically significant effect of arrest on the occurrence of a subsequent spouse assault, which supports PH's conclusion 3.

2.4 Assumption Compatibility

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.

$$H_0(\text{Fisher}) : Y_i(1) = Y_i(0), \forall i = 1, .., N$$

This hypothesis is usually regarded as sharp because each suspect is assumed to have zero treatment effect no matter which group it is assigned to. The only randomness of the Fisher exact test comes from the treatment-assignment (randomization), and the p-value can be justified by the randomization test itself. The randomization process will not change the total number of outcome because the total number of control and assignment are fixed.

However, in the PH example (Neyman model), this number is random (on Page 693, "Eligible cases were randomly assigned to an arrest or a no-arrest response"). Though the Neyman model assumes each observation has two potential outcomes which are fixed values, the total number observed recidivists should be random instead of a fixed number.

Thus, under this context, the assumption of Fisher exact test is **not compatible** with Neyman model in PH setting.

2.5 Validity of Binomial Test

The binomial test as rendered in textbooks concerns independent Bernoulli trials. The reasons why we don't think of recidivism outcomes as random coin flips (unlike PH) are as follows:

1. **Independence assumption:** The 907 suspects are supervised by only 396 officers, namely at least 115 officers contribute to three or more cases, suggesting that the observations may not be random or independent and may be correlated with other observations to some extent.
2. **Resampling with replacement:** The binomial test assume the resampling with replacement, but in PH setting, the resampling is conducted without replacement, where the standard error of the test statistic will shrink.

When it comes to the bias direction in the test, we can conclude that the binomial test is conservative so though the reasoning is not true, we can still trust some of the inference before for the following reasons:

1. On one hand, the correlation between suspects implies a lower variance, which will lead to a larger test statistic and a lower p-value, indicating a smaller probability of rejecting the null hypothesis.
2. On the other hand, the resampling without replacement also overestimates the standard error in the test statistic, namely underestimates the test statistic (since the standard error is in the denominator) and overestimates the p-value.

These two aspects all lead to the same result that the binomial test conducted myself is more conservative than the true binomial test.

For the binomial tests among employed and unemployed subgroup respectively, the p-values show significant results, and thus the true binomial test must also reject the null hypothesis and give significant results. However, for the binomial test among all suspects, the p-value is large and we cannot make any informative conclusion about whether the effect is significant or not.

There are some solutions to the aforementioned violation of assumptions of binomial test:

1. **Independence assumption:** To deal with the correlation between observations, we can regard the suspects under supervision of the same officer as a cluster, and then replace the permutation test within individuals in the previous analysis with permuting within clusters.

2. **Resampling with replacement:** We can reproduce the resampling process with replacement in the experiment so that the binomial assumption will not be violated.

Therefore, though the assumptions of binomial test are violated in the PH setting, we can make use of the conservativeness of binomial test to obtain informative conclusion, or we can adopt the aforementioned strategies to deal with the randomness issue and obtain a justified p-value from the analysis on the current Dade County experimental data.