Detective work

The Dade County spouse assault experiment was executed excellently as a randomized field experiment [Pate and Hamilton (1992), hereafter PH]. Since randomization does not justify logistic regression as presented in PH, we would like to recover the experimental data and conduct another anlysis that is justified by the randomization.

In order to conduct comparion on rates of re-abuse, between the treatment group (arrestees) and the control group, as well as the same rate comparison within the two subgroups (unemployed/employed subjects), we need the numbers in the following table:

	no arrest	arrest	
unemployed	n_{00}/N_{00}	n_{01}/N_{01}	$n_{0.}/N_{0.}$
employed	n_{10}/N_{10}	n_{11}/N_{11}	$n_{1.}/N_{1.}$
	$n_{.0}/N_{.0}$	$n_{.1}/N_{.1}$	$n_{..}/N_{..}$

1. Subject counts

The total subject counts can be directly derived from the codebook and the dataset file *part_607*. After basic data cleaning, we re-group the data based on column *act* into a new two-level column *trass*, i.e. treament assgined, following the *intention-to-treat* principle. So with cross-tabulation, we get all the N's for the table.

```
##      trass
## sempl   0   1
##     0 139 124
##     1 306 338
```

A simple sanity check on the matching of the data to the PH paper is done with the unemployment rate among the subjects, which is 28.9967% from the data and agrees with 29% in the PH paper.

2. Re-abusing subject counts

Without a proper linked dataset on the number of re-abusing subjects, we could only recover the n's from the PH Figure 1 percentages. We round the calculated numbers to the nearest integers to reflect the fact that they are real human counts.

```
##      trass
## sempl  0  1
##     0 10 21
##     1 38 21
```

A similar sanity check would be the re-abusing rates. The rate of re-abuse among arrestees is 9.0909%, which agrees with the 9.0% in PH paper. Also, the re-abusing rate of non-arrestees is 10.7865%, slightly bigger than the PH 10.6% due to rounding.

So, the final result of our data recovery work gives the table below.

	no arrest	arrest	
unemployed	10/139	21/124	31/263
employed	38/306	21/338	59/644
	48/445	42/462	90/907

Statistical work

Based on the recovered data, we would like to re-evaluate the three conclusions drawn in the PH paper based their logistic regression analyses. We would like to compare the relevant observed rates in the above table with serveral methods, including Fisher's exact test and two-sample test of equal binomial proportions.

In more details, the two-sample test of equal binomial proportions use the test statistic

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)(\frac{1}{n_1} + \frac{1}{n_2})}}$$

where $p_1 = n_1/N_1$, $p_2 = n_2/N_2$, $p = (n_1 + n_2)/(N_1 + N_2)$, and follows normal distribution check.

1. re-abuse ~ arrest | employed

We would like to assess the conclusion that "Among employed suspects, arrest had a statistically significant deterrent effect on the occurrence of a subsequent assault", with the hypothesis testing given below.

Among employed suspects,

H0: the occurrence rate of a subsequent assault is the same between the arrestees and non-arrestees
HA: the re-abuse rate is higher among the non-arrestees than that of the arrestees

Employed suspects:

```
##           non-arrest arrest
## reabuse           38     21
## non-recid        268    317
##
##  Fisher's Exact Test for Count Data
##
## data:  cl1
## p-value = 0.004715
## alternative hypothesis: true odds ratio is greater than 1
## 95 percent confidence interval:
##  1.297   Inf
## sample estimates:
## odds ratio
##      2.138
```

Using the test statistic given above, the p-value for the two-sample binomial test is 0.0032, which is close to that of Fisher's exact test. Both of them suggest that we can reject the null hypothesis and the drawn conclusion is acceptable.

2. re-abuse ~ arrest | unemployed

We would like to assess the conclusion that "Among unemployed suspects, ... significant increases in subsequent assault were associated with arrest", with the hypothesis testing given below.

Among unemployed suspects,

H0: the occurrence rate of a subsequent assault is the same between the arrestees and non-arrestees
HA: the re-abuse rate is lower among the non-arrestees than that of the arrestees

Unemployed suspects:

```
##           non-arrest arrest
## reabuse           10     21
## non-recid        129    103
##
##  Fisher's Exact Test for Count Data
##
## data:  cl2
## p-value = 0.01184
## alternative hypothesis: true odds ratio is less than 1
## 95 percent confidence interval:
##  0.0000 0.7901
## sample estimates:
## odds ratio
##     0.3816
```

Similarly, using the test statistic given above, the p-value for the two-sample binomial test is 0.0072, which is also close to that of Fisher's exact test. Moreover, both of them suggest that we can reject the null hypothesis and the drawn conclusion in PH is acceptable.

3. re-abuse $\sim$ arrest

We would like to assess the final conclusion that "[Among all suspects, there is] no statistically significant effect of arrest on the occurrence of a subsequent spouse assault", with the hypothesis testing given below.

Among all suspects,

H0: the occurrence rate of a subsequent assault is the same between the arrestees and non-arrestees
HA: the re-abuse rate is different among the non-arrestees from the arrestees

All suspects:

```
##           non-arrest arrest
## reabuse           48     42
## non-recid        397    420
##
##  Fisher's Exact Test for Count Data
##
## data:  cl3
## p-value = 0.4372
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.7635 1.9199
## sample estimates:
## odds ratio
##      1.209
```

Using the test statistic given above, the p-value for the two-sample binomial test is 0.4017, which is close to that of Fisher's exact test. Both of them suggest that we can NOT reject the null hypothesis and the PH conclusion is acceptable.

Randomization

About the Fisher's exact test, the p-value is computed conditionally on the row and column margins being constant. In our application, this assumption suggest that the total number of observed re-abusers would not change if there had been a different randomization outcome in the Dade County experiment.

In the Neyman model of experiment, the Y_i^T and Y_i^C is unknown but deterministic, with only the assignment X_i being random. The total number of the observed re-abusers, in theory, can change in the overall population under a different random assignment. And if the employment rate is fixed, so will the subgroups. However, in reality, this Fisher's exact test assumption could be really close to the truth if the Y^T's and Y^C's share similar success rate. This is verified by the similar p-values calculated by the two sample Binomial test.

On the other hand, the binomial test concerns independent Bernoulli trials. In our experiment, the re-abuse outcomes are not considered as random coin flips, but only the treatment assginment is pure random.

However, the randomization can justify that the two distributions (treatment and control) are independent regardless of their underlying details of the distributions. Even though not exactly binomial, the distributions can be approximated by binomial distributions, generating p-values close to those in the Fisher's exact test, which casts hypergeometric distribution on the 2-by-2 table.