Assignment 4

Daniel Koslovsky

2025-05-16

```r
library(knitr)
library(fixest)
library(modelsummary)
library(haven)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4      v readr     2.1.5
## v forcats   1.0.0      v stringr   1.5.1
## v ggplot2   3.5.1      v tibble    3.2.1
## v lubridate 1.9.3      v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

```r
library(estimatr)
library(stringr)
library(openxlsx)


#Question 1
lo_table <- read.xlsx("prepped_LO_table.xlsx")
lo_table <- data.frame(lo_table)

#Remove sub-indicators and only analyze primary indicators
lo_table <- lo_table |>
  filter(str_starts(Indicator, "I-"))


#Question 2
lo_table_w_pvalues <- lo_table |>
  mutate(chi_sq_pvalues = pchisq(Chi.sq, df = 1, lower.tail = FALSE))

lopez_table_w_sig_vars <- lo_table_w_pvalues |>
  mutate(five_perc_sig = ifelse(chi_sq_pvalues <= .05, 1, 0))

table(lopez_table_w_sig_vars$five_perc_sig)
```
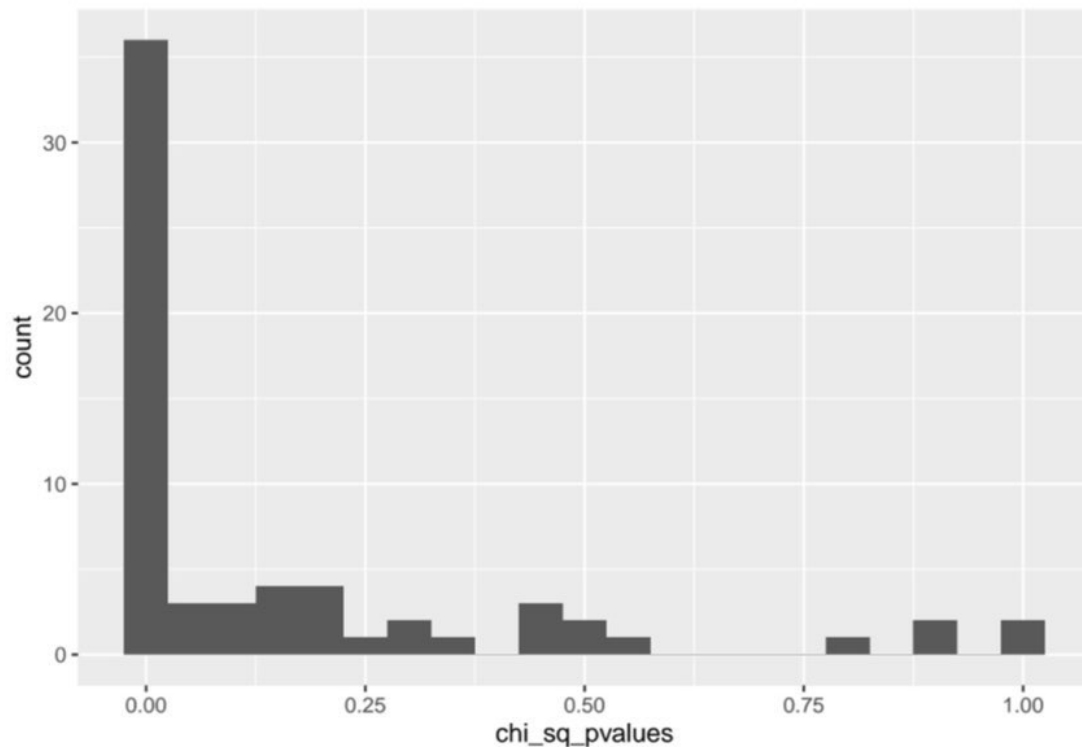
```
##
```

```
##  0  1
## 27 38
```

The author's analysis finds 38 indicators that are significant at the 5% level.

```
#Question 3
lo_table_w_pvalues |>
  ggplot(aes(x = chi_sq_pvalues)) +
  geom_histogram(binwidth = .05)
```


There are a high concentration of p-values under .05 indicating that there is a some predictive signal among the indicators.

```
#Question 4
lopez_table_w_bonf_pvals <- lopez_table_w_sig_vars |>
  mutate(
    bonferroni_pvalue = .05 / 65,
      bonferroni_significant = ifelse(chi_sq_pvalues <= bonferroni_pvalue, 1, 0)
  )

table(lopez_table_w_bonf_pvals$bonferroni_significant)
```

```
##
##  0  1
## 40 25
```

```
#Indicators that are still significant
lo_table_bonf_sig_indicators <- lopez_table_w_bonf_pvals |>
  filter(bonferroni_significant == 1) |>
  select(Indicator)

kable(lo_table_bonf_sig_indicators, caption = "Bonferroni Significant Indicators")
```

Table 1: Bonferroni Significant Indicators

Indicator
I-11. They have shown exaggerated jealousy or have suspected their partner of infidelity in the last 6 months
I-12. They have shown controlling behaviors in the last 6 months
I-13. In the last 6 months they have shown harassing behaviors
I-15. In the last year they have disrespected an authority figure
I-18. Presence of problems in their life in the last 6 months.
I-19. Presence of a criminal record
I-20. Presence of past breakings of sentence conditions
I-21. Presence of physical/sexual aggression records
I-24. Presence of suicidal ideas or attempts
I-25. Presence of any type of addiction or substance abuse (alcohol and drugs)
I-26. Presence of a gender or domestic violence history within their family
I-40. The victim has expressed their desire to end the relationship to the aggressor at least 6 months ago
I-1. The aggressor has contacted the victim through technological means
I-3. The aggressor has physically gotten closer to the victim
I-5. The aggressor has distanced themselves from the victim
I-6. The aggressor shows a peaceful attitude
I-7. The aggressor exteriorizes a respectful attitude towards the law
I-8. The aggressor shows regret
I-9. The aggressor joins Help programs
I-10. The aggressor follows the requirements of the charges and of the family separation
I-11. The victim has resumed living with the aggressor
I-12. The victim doesn't report any episodes to file a complaint about, and/or turns
I-13. The victim engages in activities that threaten their own security
I-21. Presence of a conflict about their responsibilities in the children's care
I-22. The victim feels that the risk level is nonexistent

```
#Indicators that are no longer signficant
lo_table_bonf_sig_indicators <- lopez_table_w_bonf_pvals |>
  filter(bonferroni_significant == 0, five_perc_sig == 1) |>
  select(Indicator)

kable(lo_table_bonf_sig_indicators, caption = "Indicators Significant Unadjusted but Not Bonferroni")
```

Table 2: Indicators Significant Unadjusted but Not Bonferroni

Indicator
I-6. Use of weapons/objects against partner
I-10. An escalation in aggression or threats in the last 6 months
I-14. In the last year they have done material damage

Indicator
I-16. In the last year they have committed an aggression against third parties and/or animals.
I-17. In the last year they have made threats and slights towards third parties.
I-22. Presence of gender violence records against other partners
I-30. Victim is a foreigner
I-39. The victim has suffered gender violence by the hands of other aggressors in the past
I-43. The victim thinks that the aggressor is capable of attacking them in a violent way or even of killing them
I-4. The aggressor is an escaped convict or is in an unknown location
I-16. Judicial proceedings to carry out a separation/divorce unwanted by the aggressor
I-19. The aggressor has a stable economic and labor situation
I-20. The aggressor has favorable family and social support

The number of significant indicators drops from 38 to 25 when implementing the Bonferroni control procedure.

```
#Question 5
lopez_table_w_bh_test <- lopez_table_w_bonf_pvals |>
  arrange(chi_sq_pvalues) |>
  mutate(
    pvalue_rank = row_number(),
    q = .05,
    bh_pvalue = q * (pvalue_rank / 65),
    bh_significance = ifelse(chi_sq_pvalues <= bh_pvalue, 1, 0)
    )

lo_table_bh_sig_indicators <- lopez_table_w_bh_test |>
  filter(bh_significance == 1) |>
  select(Indicator)

kable(lo_table_bh_sig_indicators, caption = "Benjamini-Hochberg Significant Indicators")
```
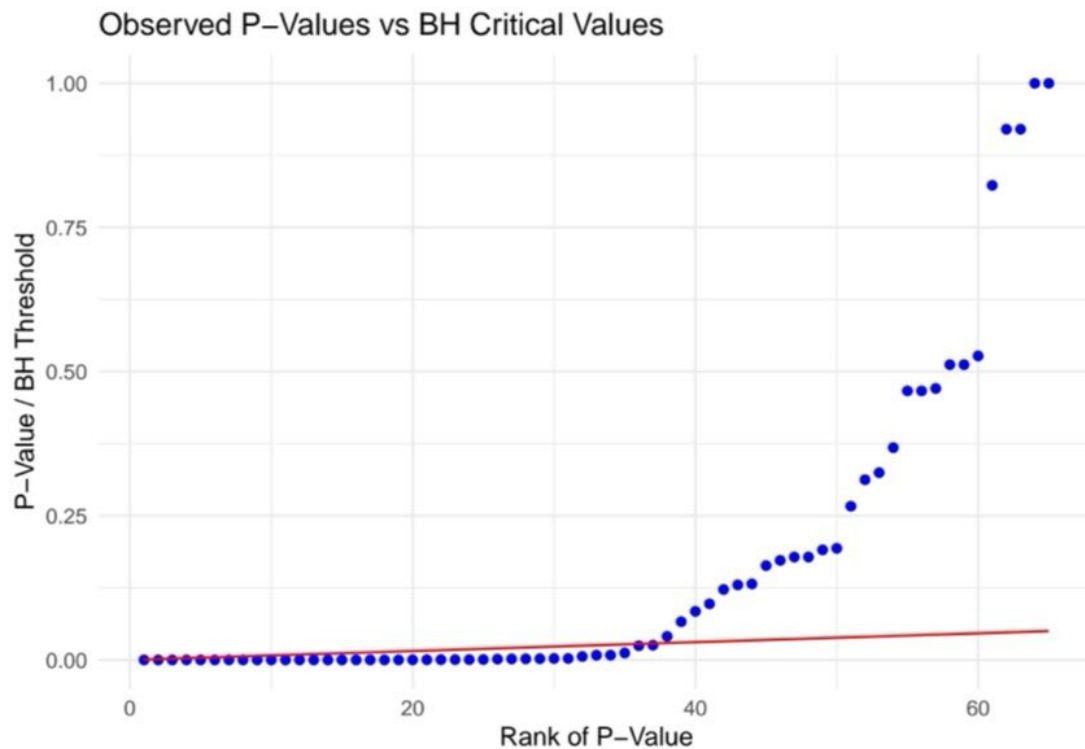
Table 3: Benjamini-Hochberg Significant Indicators

Indicator
I-5. The aggressor has distanced themselves from the victim
I-6. The aggressor shows a peaceful attitude
I-3. The aggressor has physically gotten closer to the victim
I-7. The aggressor exteriorizes a respectful attitude towards the law
I-13. The victim engages in activities that threaten their own security
I-1. The aggressor has contacted the victim through technological means
I-8. The aggressor shows regret
I-12. The victim doesnâ€™t report any episodes to file a complaint about, and/or turns
I-10. The aggressor follows the requirements of the charges and of the family separation
I-21. Presence of a conflict about their responsibilities in the childrenâ€™s care
I-11. The victim has resumed living with the aggressor
I-12. They have shown controlling behaviors in the last 6 months
I-22. The victim feels that the risk level is nonexistent
I-19. Presence of a criminal record
I-11. They have shown exaggerated jealousy or have suspected their partner of infidelity in the last 6 months

Indicator
I-21. Presence of physical/sexual aggression records
I-24. Presence of suicidal ideas or attempts
I-9. The aggressor joins Help programs
I-40. The victim has expressed their desire to end the relationship to the aggressor at least 6 months ago
I-20. Presence of past breakings of sentence conditions
I-18. Presence of problems in their life in the last 6 months.
I-25. Presence of any type of addiction or substance abuse (alcohol and drugs)
I-15. In the last year they have disrespected an authority figure
I-13. In the last 6 months they have shown harassing behaviors
I-26. Presence of a gender or domestic violence history within their family
I-22. Presence of gender violence records against other partners
I-20. The aggressor has favorable family and social support
I-10. An escalation in aggression or threats in the last 6 months
I-6. Use of weapons/objects against partner
I-43. The victim thinks that the aggressor is capable of attacking them in a violent way or even of killing them
I-19. The aggressor has a stable economic and labor situation
I-14. In the last year they have done material damage
I-30. Victim is a foreigner
I-16. Judicial proceedings to carry out a separation/divorce unwanted by the aggressor
I-39. The victim has suffered gender violence by the hands of other aggressors in the past
I-17. In the last year they have made threats and slights towards third parties.
I-4. The aggressor is an escaped convict or is in an unknown location

```
ggplot(lopez_table_w_bh_test, aes(x = pvalue_rank)) +
  geom_point(aes(y = chi_sq_pvalues), color = "blue") +
  geom_line(aes(y = bh_pvalue), color = "red") +
  labs(
    x = "Rank of P-Value",
    y = "P-Value / BH Threshold",
    title = "Observed P-Values vs BH Critical Values"
  ) +
  theme_minimal()
```

Observed P-Values vs BH Critical Values

Using the Benjamini-Hochberg approach, 37 Indicators are significant. Of these, $(.05 * 37 = 1.85)$ about 2 are expected to be false discoveries

```r
#Question 6
lo_table_fdr <- lopez_table_w_bh_test |>
  mutate(false_disc_prob = chi_sq_pvalues * (65 / pvalue_rank))

summary(lo_table_fdr$false_disc_prob)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## 0.000000 0.000107 0.016024 0.191318 0.251681 1.015625
```

```r
#Question 7
summary_table <- lo_table_fdr |>
  mutate(sum_of_tests = bonferroni_significant + bh_significance + five_perc_sig,
         three_test_significance =
           case_when(
             (sum_of_tests == 0) ~ "No significance",
             (sum_of_tests == 3) ~ "Significant for all",
             (bonferroni_significant == 1 & bh_significance == 1 & five_perc_sig == 0) ~ "Bonferroni an
             (bonferroni_significant == 1 & bh_significance == 0 & five_perc_sig == 1) ~ "Bonferroni an
             (bonferroni_significant == 0 & bh_significance == 1 & five_perc_sig == 1) ~ "BH and Unadju
             (bonferroni_significant == 1 & bh_significance == 0 & five_perc_sig == 0) ~ "Bonferroni si
             (bonferroni_significant == 0 & bh_significance == 1 & five_perc_sig == 0) ~ "BH significan
             (bonferroni_significant == 0 & bh_significance == 0 & five_perc_sig == 1) ~ "Unadjusted si
           )
```

```
) |>
  select(Indicator, three_test_significance)

kable(summary_table, caption = "Summary Table")
```

Table 4: Summary Table

Indicator	three_test_significance
I-5. The aggressor has distanced themselves from the victim	Significant for all
I-6. The aggressor shows a peaceful attitude	Significant for all
I-3. The aggressor has physically gotten closer to the victim	Significant for all
I-7. The aggressor exteriorizes a respectful attitude towards the law	Significant for all
I-13. The victim engages in activities that threaten their own security	Significant for all
I-1. The aggressor has contacted the victim through technological means	Significant for all
I-8. The aggressor shows regret	Significant for all
I-12. The victim doesnâ€™t report any episodes to file a complaint about, and/or turns	Significant for all
I-10. The aggressor follows the requirements of the charges and of the family separation	Significant for all
I-21. Presence of a conflict about their responsibilities in the childrenâ€™s care	Significant for all
I-11. The victim has resumed living with the aggressor	Significant for all
I-12. They have shown controlling behaviors in the last 6 months	Significant for all
I-22. The victim feels that the risk level is nonexistent	Significant for all
I-19. Presence of a criminal record	Significant for all
I-11. They have shown exaggerated jealousy or have suspected their partner of infidelity in the last 6 months	Significant for all
I-21. Presence of physical/sexual aggression records	Significant for all
I-24. Presence of suicidal ideas or attempts	Significant for all
I-9. The aggressor joins Help programs	Significant for all
I-40. The victim has expressed their desire to end the relationship to the aggressor at least 6 months ago	Significant for all
I-20. Presence of past breakings of sentence conditions	Significant for all
I-18. Presence of problems in their life in the last 6 months.	Significant for all
I-25. Presence of any type of addiction or substance abuse (alcohol and drugs)	Significant for all
I-15. In the last year they have disrespected an authority figure	Significant for all
I-13. In the last 6 months they have shown harassing behaviors	Significant for all
I-26. Presence of a gender or domestic violence history within their family	Significant for all
I-22. Presence of gender violence records against other partners	BH and Unadjusted significant
I-20. The aggressor has favorable family and social support	BH and Unadjusted significant
I-10. An escalation in aggression or threats in the last 6 months	BH and Unadjusted significant
I-6. Use of weapons/objects against partner	BH and Unadjusted significant
I-43. The victim thinks that the aggressor is capable of attacking them in a violent way or even of killing them	BH and Unadjusted significant
I-19. The aggressor has a stable economic and labor situation	BH and Unadjusted significant
I-14. In the last year they have done material damage	BH and Unadjusted significant

Indicator	three_test_significance
I-30. Victim is a foreigner	BH and Unadjusted significant
I-16. Judicial proceedings to carry out a separation/divorce unwanted by the aggressor	BH and Unadjusted significant
I-39. The victim has suffered gender violence by the hands of other aggressors in the past	BH and Unadjusted significant
I-17. In the last year they have made threats and slights towards third parties.	BH and Unadjusted significant
I-4. The aggressor is an escaped convict or is in an unknown location	BH and Unadjusted significant
I-16. In the last year they have committed an aggression against third parties and/or animals.	Unadjusted significant only
I-2. Psychological Violence	No significance
I-42. The victim fears for the physical integrity of the minors or family members under their care	No significance
I-2. The aggressor has contacted the victim through third parties	No significance
I-35. There is a history of gender or domestic violence within their family	No significance
I-18. The aggressor establishes a new sentimental relationship	No significance
I-15. The victim has minors or family member under their care	No significance
I-14. The victim is in a situation of economical dependency	No significance
I-17. The victim establishes a new sentimental relationship, which is not approved by the aggressor	No significance
I-23. Presence of a mental/psychiatric disorder	No significance
I-33. Victim has suicidal ideas/attempts	No significance
I-8. The aggressor is an expert in combat techniques	No significance
I-36. Victim is economically dependent on the aggressor	No significance
I-5. Victimâ€™s defensive reaction to physical aggression	No significance
I-34. Victim has any kind of addiction or engages in substance abuse (alcohol and drugs)	No significance
I-27. Presence of any type of disability	No significance
I-1. Presence of any type of violence carried out by the aggressor	No significance
I-9. Presence of threats/plans aimed to cause the victim harm	No significance
I-32. Victim has a mental/psychiatric disorder	No significance
I-4. Sexual violence	No significance
I-7. The aggressor has access to firearms	No significance
I-38. The victim has withdrawn complaints in the past	No significance
I-37. The victim has minors or family members under their care	No significance
I-28. Victim is in gestation period	No significance
I-31. Victim has no social/family support	No significance
I-41. There have been reports of mutual violence	No significance
I-3. Physical violence	No significance
I-29. Victim suffers from a serious illness	No significance

As in the original paper, indicators that are related to the incident in question and the victim's characteristics are likely to be not good indicators after the adjustments are applied as well. The characteristics having to do with the aggressor's characteristics, like showing a peaceful attitude and showing regret. One pattern among the indicators that were not significant for the Bonferroni adjustment but significant otherwise, seems to be that a lot of them are related to the aggressor's life situation, such as their family and social support and whether they have a stable economic situation.

#Question 8#

FWER control is more important when false positives are more costly than false negatives, the opposite

is true for FDR control. In this research study investigating attributes that correlate well with domestic violence recidivism, the possible costs of a false positive are resources spent by the government to prevent domestic violence recidivism that is in reality not likely to occur. The possible costs of a false negative are that preventable instances of domestic violence recidivism occur. Given that domestic violence is a very traumatic and disruptive event in a person's life, it is likely that the cost of a false negative is more costly to society than a false positive. Moreover, the study is more of an exploratory analysis for what indicators deserve further analysis, than a final determination of what indicators should or should not be incorporated into policy. Therefore, FDR control is more appropriate in this instance.

#Question 9#

This occurs because the slope of the 30 lowest p values are basically zero and the next seven are not much greater than zero, so the BH adjustment has no effect until we reach the 38th p value.

#Question 10#

Accounting for multiple testing does not have a big impact on the conclusions about which indicators are significant. Most of the indicators originally found significant remain so. We are more skeptical of the indicators that were only significant after the FDR adjustment, but can't rule them out. We can likely disregard the indicator that was no longer significant after both adjustments.