Problem Set 4

ylelkes

September 25, 2015

We'll be using dplyr and reshape2 for this exercise.

Here's a recent study that was run examining the effects of media portrayals of victim suffering on attitudes towards gays.

Download the study material. As before, use markdown for this problem set, but this time print it to pdf. We are interested in Goldman's key dependent measures.

Among heterosexual respondents, there are four dependent variables: sympathy toward the victim (Q2); overall favorability toward gays and lesbians (Q17); support for gay rights policies (Q16); and pity toward gays and lesbians (Q18_2).

1. First, let's select only the columns we're interested in with dplyr (the experimental condition, the four dependent measures) reshape or melt the dataframe so that the four dependent measures that were asked among heterosexuals are "stacked.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(reshape2)
goldman <- foreign::read.spss("~/Downloads/Goldman507/TESS163_Goldman_Client.sav",to.data.frame=T,use.va
```

```
## Warning in
## foreign::read.spss("~/Downloads/Goldman507/TESS163_Goldman_Client.sav", :
## ~/Downloads/Goldman507/TESS163_Goldman_Client.sav: Unrecognized record
## type 7, subtype 14 encountered in system file

## Warning in
## foreign::read.spss("~/Downloads/Goldman507/TESS163_Goldman_Client.sav", :
## ~/Downloads/Goldman507/TESS163_Goldman_Client.sav: Unrecognized record
## type 7, subtype 18 encountered in system file
```

```
goldman2 <- select(goldman,XTESS163,Q2,Q17,Q16,Q18_2)
meltedgoldman <- melt(goldman2,id.vars = "XTESS163")
```

```
## Warning: attributes are not identical across measure variables; they will
## be dropped
```

2. Next, use the various functions in dplyr to first filter out rows so that we only have data among heterosexuals. You'll need to take a look at the codebook here.

```
meltedgoldman$XTESS163 <- as.numeric(meltedgoldman$XTESS163)
goldman3 <- filter(meltedgoldman,XTESS163<5)
```

3. Next, using dplyr, generate a table that contains the mean, standard deviation, and N for each condition and dependent measures. Print this table using xtable.

```
goldman3$value[goldman3$value<0]=NA
group_by(goldman3,variable,XTESS163) -> groupgoldman
table1 <- summarise(groupgoldman,mean=mean(value,na.rm=T),sd=sqrt(var(value,na.rm=T)),N=(n()))
```

% latex table generated in R 3.1.2 by xtable 1.7-4 package % Sat Oct 17 11:39:51 2015

	variable	XTESS163	mean	sd	N
1	Q2	1.00	3.35	0.68	289
2	Q2	2.00	4.02	0.82	285
3	Q2	3.00	4.36	0.80	281
4	Q2	4.00	2.12	0.83	289
5	Q17	1.00	5.73	3.18	289
6	Q17	2.00	5.59	3.06	285
7	Q17	3.00	5.82	3.09	281
8	Q17	4.00	5.69	3.05	289
9	Q16	1.00	2.74	1.66	289
10	Q16	2.00	2.79	1.54	285
11	Q16	3.00	2.89	1.61	281
12	Q16	4.00	2.82	1.58	289
13	Q18_2	1.00	3.52	1.22	289
14	Q18_2	2.00	3.57	1.11	285
15	Q18_2	3.00	3.50	1.18	281
16	Q18_2	4.00	3.53	1.17	289

4. Now mutate the table to include the standard errors and upper and lower 95 percent confidence intervals.

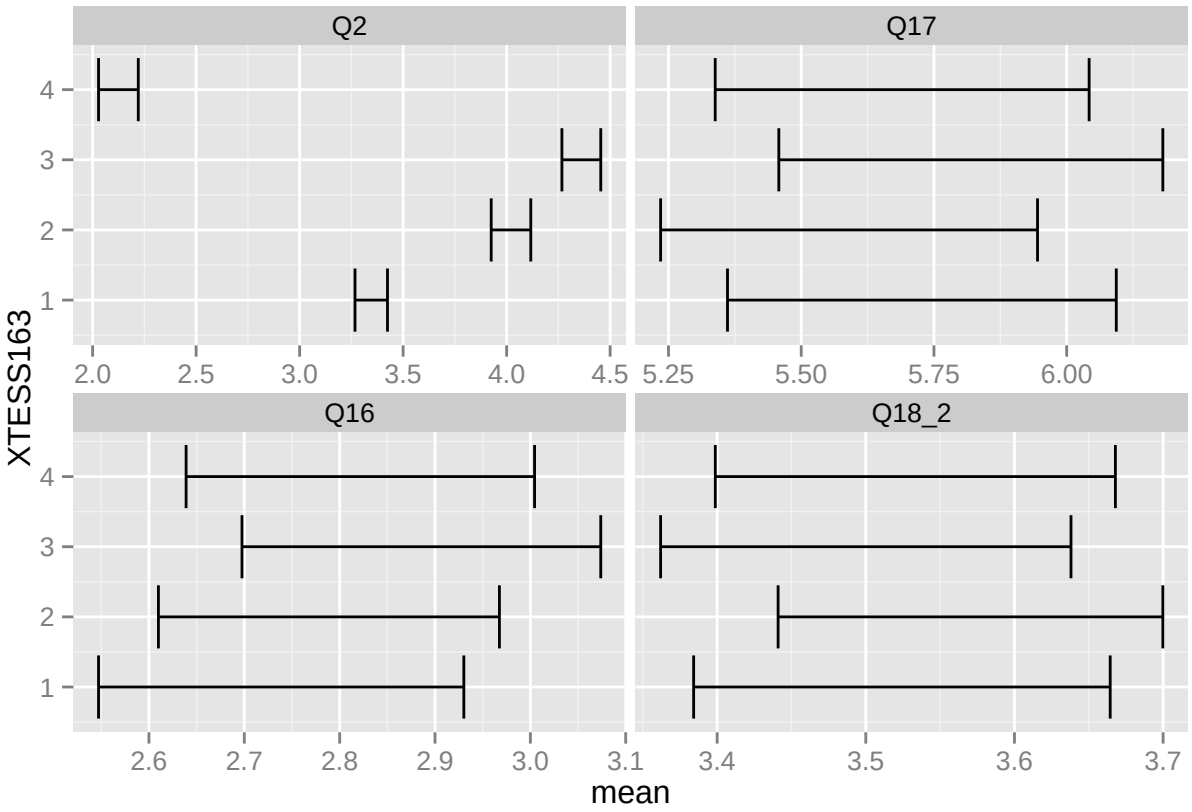
```
mutatedtab <- mutate(table1,se=sd/sqrt(N),upper=mean+1.96*se,lower=mean-1.96*se)
```

5. Finally, create a plot where you illustrate these results. Have a facet for each dependent measure, the y-axis should index condition, the x-axis the score on the dependent measure. Include the 95 percent confidence intervals. Allow the scales on the x-axis to be "free", i.e., vary by dependent measures.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

```
ggplot(mutatedtab,aes(y=XTESS163,x=mean))+geom_errorbarh(aes(xmin=lower,xmax=upper))+facet_wrap(~variabl
```



6. Write up a few lines explaining the results of this experiment.

7. Next let's see if sympathy for the victim in the story depends on the sexual orientation of the subject, which you can find out by looking at XTESS163. Run a t-test to see if Q2 varies between gay and straight people. Report these results in APA format, i.e., the mean among straight people is X, the mean among gay people is X. But report the numbers dynamically–i.e., extract them from the t-test output and embed them with inline R code. Someone from the psych department at the UVA made a neat package that allows us to automatically print out some of this information with the st() function.

```
goldman$Q2[goldman$Q2<0]=NA
goldman$gaystraight<- car::recode(goldman$XTESS163,"1:4='Straight';else='Gay'")
ttestout <- t.test(Q2~gaystraight,goldman)
library(staTeX)
```

```
devtools::install_github(repo = "staTeX",username = "SachaEpskamp")
library(staTeX)
```

The mean among straight people was 3.45, and the mean among gay people was 3.95, `st(ttestout)`.

8. Let's see if there is an interaction between condition and being straight on Q2. You'll have to create variables indicating if a respondent was gay or straight, and a variable indicating if they saw the hero story or the victim story (ignore other conditions as missing.) Go back to the original dataset, and create a table which reports the means and standard deviations which groups by sexual orientation and by condition.

```
goldman$gaystraight<- car::recode(goldman$XTESS163,"1:4='Straight';else='Gay'",as.factor=T)
goldman$condition <- car::recode(goldman$XTESS163,"c(3,7)='Victim';c(2,6)='Hero';else=NA",as.factor=T)
goldman$Q2  <- as.numeric(goldman$Q2)
goldman1 <- select(goldman,gaystraight,condition,Q2)
goldman2 <- group_by(goldman1,gaystraight,condition)
goldman3 <- na.omit(summarise(goldman2,mean=mean(Q2,na.rm=T),sd=sqrt(var(Q2,na.rm=T))))
```

9. Finally, transform that table into a two x two table using dcast, which presents the means for each combination. Remove NAs, and again present the table using xtable.

```
library(reshape2)
castedtab <- dcast(goldman3,gaystraight~condition,value.var="mean")
```

10. Finally, using st() again, report the results from a 2x2 anova in APA style. HINT: To report one statistic at a time you'll have to do st(X)[1], st(X)[2], etc.

```
anova12 <- aov(Q2~condition*gaystraight,goldman2)
```

The effect of condition on Q2 was ($F(1, 1125) = 58.328$, $p < 0.001$). The effect of sexual orientation on Q2 was ($F(1, 1125) = 2.854$, $p = 0.091$). The two way interaction was not significant ($F(1, 1125) = 0.132$, $p = 0.717$).