

# STAT 209: Take-Home Problem Set 1

Rebecca Weiss

February 14, 2013

## 1 Question 1: Observational studies, Associations

### 1.1 Part A: Spurious correlation

```
> rm(list=ls(all=T))
> gc()

      used (Mb) gc trigger (Mb) max used (Mb)
Ncells 198123 10.6      407500 21.8    350000 18.7
Vcells 317488  2.5      905753  7.0    863120  6.6

> #Libraries
> library(mediation)
> #reconstruct the correlation matrix
> data = as.data.frame(cbind(
  c(1, -0.03, 0.39, -0.05, -0.08),
  c(-0.03, 1, 0.07, -0.23, -0.16),
  c(0.39, 0.07, 1, -0.13, -0.29),
  c(-0.05, -0.23, -0.13, 1, 0.34),
  c(-0.08, -0.16, -0.29, 0.34, 1)))
> rownames(data) = c('Exercise','Hardiness','Fitness','Stress','Illness')
> colnames(data) = c('Exercise','Hardiness','Fitness','Stress','Illness')
> #We need to see if the bivariate correlation of -0.27 changes compared to the partial correlation
> #let's look at the correlation between Fitness and Illness controlling for stress
> ryx = -0.29
> ryz = -0.13
> rxz = 0.34
```

```

> ryx.z = (ryx - (ryz*rxz)) / ( sqrt(1-(ryz^2)) * sqrt(1-(rxz^2)) )
> ryx.z
[1] -0.263608

> #this produces a partial correlation that is close to the zero order correlation between
> #Fitness and Illness. This implies that there is a direct effect, controlling for stress.
> #Here's an attempt at a significance test. First, compute the partial correlation matrix:
> cor_mat = as.matrix(data)
> pcor_mat = solve(cor_mat) * -1
> diag(pcor_mat) = -diag(pcor_mat)
> pcor_mat = cov2cor(pcor_mat)
> cor.test(cor_mat, pcor_mat)

        Pearson's product-moment correlation

data:  cor_mat and pcor_mat
t = 45.6318, df = 23, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9874112 0.9976215
sample estimates:
      cor
0.9945225

> #This shows that the correlation and partial correlation are not significantly different
> #from each other at the 95% CI (because the correlation between these two matrices is nonzero)

```

## 1.2 Part B: Stress as mediator

```

> #STANDARDIZED METRICS
> #first, get the coef predictors for stress and fitness
> illness = c(-.08,-.16,-.29,.34)
> illness_pred_mat = data[1:4, 1:4]
> illness_coefs = solve(illness_pred_mat, illness)
> stress = c(-.05, -.23, -.13)
> stress_pred_mat = data[1:3, 1:3]
> stress_coefs = solve(stress_pred_mat, stress)

```

```

> #perform mediation analysis via sobel test
> names(illness_coefs) = names(data)[1:4]
> names(stress_coefs) = names(data)[1:3]
> #get standard errors from correlation matrix
> illness_rsqr = illness_coefs %*% illness
> stress_rsqr = stress_coefs %*% stress
> #for illness
> n = 373
> illness_cov_beta = (1-illness_rsqr)[1] * solve(illness_pred_mat)
> illness_serr = NA
> for (i in 1:length(diag(illness_cov_beta))){
      illness_serr[i] = sqrt(as.numeric(diag(illness_cov_beta)[i])/n) * (n / (n-length(illness)))
    }
> names(illness_serr) = names(illness_coefs)
> #for stress
> stress_cov_beta = (1-stress_rsqr)[1] * solve(stress_pred_mat)
> stress_serr = NA
> for (i in 1:length(diag(stress_cov_beta))){
      stress_serr[i] = sqrt(as.numeric(diag(stress_cov_beta)[i])/n) * (n / (n-length(illness)))
    }
> names(stress_serr) = names(stress_coefs)
> #here's some sloppy coding:
> illness_coefs = as.data.frame(t(illness_coefs))
> stress_coefs = as.data.frame(t(stress_coefs))
> illness_serr = as.data.frame(t(illness_serr))
> stress_serr = as.data.frame(t(stress_serr))
> #first step: indirect effect of stress on illness (mediating fitness and illness)
> indirect_effect = stress_coefs$Fitness * illness_coefs$Stress
> indirect_effect
[1] -0.03166258
> #get the effect variance and standard error
> effect_var = ((stress_coefs$Fitness)^2 * (illness_serr$Stress)^2) +
  ( (illness_coefs$Stress)^2 * (stress_serr$Fitness)^2 )
> effect_serr = sqrt(effect_var)

```

```

> #finding the z value
> z = indirect_effect/effect_serr
> #do we just compare the z against the t distribution for significance? are degrees of freedom n-1?
> print(pt(z, n-1))
[1] 0.03090761

> #UNSTANDARDIZED COEFFICIENTS
> #Unstandarized coefficient for y on x is the covariance of x and y over the variance of
> #x. Since we have all the standard deviations, we simply divide the coefficients by
> #the variance of their corresponding X to get unstandardized coefficients.
> stdevs = c(66.50, 3.80, 18.40, 6.70, 624.80)
> vars = as.data.frame(t(stdevs^2))
> names(vars) = c('Exercise','Hardiness','Fitness','Stress','Illness')
> #now redo all of the previous steps, but with unstandardized coefficients
>
> #first step: indirect effect of stress on illness (mediating fitness and illness)
> indirect_effect = (stress_coefs$Fitness / vars$Fitness) * (illness_coefs$Stress / vars$Stress)
> indirect_effect
[1] -2.083344e-06

> #get the effect variance and standard error
> effect_var = (((stress_coefs$Fitness)^2 / vars$Fitness) * ((illness_serr$Stress)^2 / vars$Stress) +
                ((illness_coefs$Stress)^2 / vars$Stress) * ((stress_serr$Fitness)^2 / vars$Fitness))
> effect_serr = sqrt(effect_var)
> effect_serr
[1] 0.0001371051

> #finding the z value
> z = indirect_effect/effect_serr
> z
[1] -0.01519524

> pt(z, n-1)
[1] 0.4939423

```

This is why we use standardized coefficients. The z-score is nonsense. This is because the original units that the variables were measured in are not compatible to each other. This results in an insignificant finding.

### 1.3 Part C: Effects of errors in measurement

Recall that reliability of a measure is the ratio of slopes. Therefore, we can find the calculations by multiplying the slope coefficient by the reliability. If fitness has reliability 0.8 and stress has reliability 0.7, we can compute the following:

```
> indirect_effect = (stress_coefs$Fitness*0.8) * (illness_coefs$Stress*0.7)
> indirect_effect
[1] -0.01773104
> #get the effect variance and standard error
> effect_var = ((stress_coefs$Fitness * 0.8)^2 * (illness_serr$Stress)^2) +
  ( (illness_coefs$Stress*0.7)^2 * (stress_serr$Fitness)^2 )
> effect_serr = sqrt(effect_var)
> effect_serr
[1] 0.01201017
> #finding the z value
> z = indirect_effect/effect_serr
> z
[1] -1.476335
> pt(z, n-1)
[1] 0.07034993
```

And thus, the relationship is no longer significant.

## 2 Question 2: Potential outcomes, Encouragement designs

### 2.1 Sodium intake on Systolic Blood Pressure

For this question, we are evaluating an experiment that looked at whether sodium intake has some effect on cardiovascular health. Since I don't know much at all about cardiology, I will assume that there was a theoretical reason to presume that a) systolic blood pressure is a useful indicator of cardiovascular health, particularly with respect to hypertension and b) there was a sound theoretical reason to hypothesize that sodium related to cardiovascular health (perhaps naive of me to assume that there are sound reasons behind hypotheses in medical research).

We are given tables 3 and 5. The design of this study was as follows: participants were randomly assigned to either a treatment or control condition. In treatment, participants were encouraged to reduce their sodium intake through an intervention program consisting of counseling, dietary changes, group meetings, and so forth. In control, participants were *not encouraged* to make changes to their diet or lifestyle. Systolic blood pressure and urinary sodium content were the measures of interest.

This is an encouragement design because of the very nature of treatment. Much like Holland's example with students being encouraged to study more, and then evaluating the effect of amount of time studied on some outcome score, this experiment has two paths: the effect of assignment on the independent variable and the effect of the independent variable on the outcome variable. Participants in treatment are encouraged to reduce their sodium intake, a treatment to which we assume must have varying degrees of success amongst assigned participants. We note the average changes between the treatment and control conditions on sodium intake as well as systolic blood pressure, note significant differences via our favorite statistical inference test, and declare sodium intake to have a strong negative effect on cardiovascular health.

### 2.1.1 Dose-response between systolic BP and sodium intake

To estimate the unit-level effects, it is useful to revisit Holland's formulations.

First, with respect to treatment, we have assignment  $S$  (or in Rogosa's slides,  $G$ ), which results in the following variable  $R(u, s)$ :

$R(u, t) =$  amount  $u$  decreases sodium intake if encouraged

$R(u, c) =$  amount  $u$  decreases sodium intake if not encouraged

Thus, outcome is not simply a function of  $S$ , but also of  $R$ , or more formally  $Y(u, s, r)$ :

$Y(u, t, r) =$  systolic blood pressure for  $u$  if  $u$  is encouraged to decrease sodium intake and  $u$  decreases sodium intake by  $r$  amount

$Y(u, c, r) =$  systolic blood pressure for  $u$  if  $u$  is not encouraged to decrease sodium intake and  $u$  decreases sodium intake by  $r$  amount

One important point to realize about encouragement designs is that not all combinations of  $R(u, s)$  and  $Y(u, s, r)$  are observable. This results in a problem: we have no counterfactual data. This is why a simple inferential test of sample means is problematic. We must instead evaluate average causal effects over these variables, using notation from equations 24-27 from Holland:  $\rho(u)$ ,  $(r - r')\beta(u)$ ,  $\tau(u)$ , and the direct + indirect effect  $\tau(u) + \rho(u)\beta(u)$ .

**Difference in amount of sodium intake for  $u$  when in treatment vs. when in control (from table 3, from baseline between conditions)**

$$R_t(u) - R_c(u) = \rho(u)$$

$$154.6 - 156.4 = -1.80000000000001$$

**TABLE 3. Changes in Urinary Sodium Excretion at 6, 12, and 18 Months of Follow-up, Total and Within-Sex Subgroup**

Follow-up Variable	Active			Control			Difference	P
	n	Mean	SD	n	Mean	SD		
All subjects								
Baseline	314	154.6	59.9	401	156.4	60.5	-1.8	.70
6 Months	238	103.0	61.3	336	159.3	81.0	-56.4	<.0001
12 Months	244	100.2	60.9	342	152.1	73.6	-51.9	<.0001
18 Months	242	99.4	60.0	341	146.5	79.2	-47.2	<.0001
18-Month change from baseline	232	-55.2	76.9	330	-11.3	77.7	-43.9	<.0001
18-Month % change from baseline*	232	-42.4	40.4	330	-12.4	45.4	-30.1	.0001
Men								
Baseline	219	166.4	61.5	288	167.3	60.9	-0.9	.88
6 Months	173	107.7	67.1	245	169.1	80.6	-61.4	<.0001
12 Months	173	105.0	62.2	252	160.5	74.1	-55.5	<.0001
18 Months	176	105.0	64.4	246	156.1	82.9	-51.1	<.0001
18-Month change from baseline	166	-60.5	82.9	238	-14.2	80.6	-46.3	<.0001
18-Month % change from baseline*	166	-43.7	41.1	238	-14.0	42.7	-29.7	.0001
Women								
Baseline	95	127.5	46.1	113	128.8	50.2	-1.3	.85
6 Months	65	90.3	40.3	91	133.0	76.4	-42.7	<.0001
12 Months	71	88.5	56.3	90	128.7	67.3	-40.2	.0001
18 Months	66	84.4	42.9	95	121.8	62.9	-37.4	<.0001
18-Month change from baseline	66	-42.0	57.9	92	-4.0	69.5	-38.0	.0004
18-Month % change from baseline*	66	-39.1	37.8	92	-7.9	52.4	-31.1	<.0001

Urinary sodium excretion values are in millimoles per 24 hours.

\*Log transformation was used to compute means and test for differences; first-order Taylor series approximation was used to estimate variances.

Difference in systolic blood pressure for reducing sodium intake by amount  $r$  vs. reducing sodium intake by amount  $r'$  (from table 5, 18 month change from baseline between conditions)

$$Y_{Sr}(u) - Y_{Sr'}(u) = (r - r')\beta(u)$$

$$-5.08 + 3.02 = -2.06$$

Difference in systolic blood pressure when in treatment vs. when in control with same amount of sodium intake  $r$  (from table 5, baseline)

$$Y_{tr}(u) - Y_{cr}(u) = \tau(u)$$

$$124.8 - 125.1 = -0.3$$

**TABLE 5. Blood Pressure Changes and Net Treatment Effects at 6, 12, and 18 Months of Follow-up and at Termination in Active Sodium Intervention and Control Participants**

Change Variable	Baseline		6 Months		12 Months		18 Months		Termination	
	n	Mean±SD	n	Mean±SD	n	Mean±SD	n	Mean±SD	n	Mean±SD
<b>Systolic blood pressure, mm Hg</b>										
Sodium intervention	327	124.8±8.5	305	-5.86±7.95	301	-5.83±7.46	304	-5.08±7.94	327	-4.86±7.81
Control	417	125.1±8.1	397	-3.83±8.46	392	-3.93±7.43	395	-3.02±8.31	417	-3.16±8.11
Δ				-2.03 (-3.26, -0.80)*		-1.90 (-3.02, -0.78)*		-2.06 (-3.28, -0.84)*		-1.69 (-2.85, -0.54)*
<b>Diastolic blood pressure, mm Hg</b>										
Sodium intervention	327	83.7±2.7	305	-3.88±6.42	301	-4.44±5.38	304	-4.35±5.65	327	-4.12±5.71
Control	417	83.9±2.8	397	-2.88±6.32	392	-3.37±5.74	395	-3.18±5.80	417	-3.27±5.73
Δ				-1.00 (-1.95, -0.04)†		-1.06 (-1.90, -0.22)†		-1.17 (2.03, -0.31)*		-0.85 (-1.68, -0.02)†

Values in parentheses show 95% confidence intervals.

\* $P < .01$ , † $P < .05$ .

Overall difference in systolic blood pressure score when in treatment vs. when in control

$$\begin{aligned}
 Y_{tR_t}(u) - Y_{cR_c}(u) &= \tau(u) + \rho(u)\beta(u) \\
 &= 3.40800000000003
 \end{aligned}$$

**Summary:** For every 3.10800000000003 decrease in salt intake, I estimate a unit decrease in systolic blood pressure.

**What's the assumption underlying these estimators?** We are assuming that *individual differences* are not relevant between conditions. In other words, we are assuming that there are no unobserved and influential covariates of participants that differ systematically between treatment versus control conditions, particularly those that interact with the treatment. This presents a potential problem of bias. For example, let's say the encouragement treatment is more effective at introducing novel lifestyle changes amongst young women more so than any other group. Perhaps these unobserved lifestyle changes are also effective at lowering systolic blood pressure. If these lifestyle changes are not observed and accounted for in this model, the effect of sodium intake on systolic blood pressure may be overestimated. For the purposes of this study, it seems rather unlikely that certain subsamples of the sample would react that dramatically different to the intervention treatment, given its ultimate goal of reducing sodium intake.

Additionally, we assume that the encouragement was the only factor leading to a difference in behavior between treatment and control. This is because the implicit assumption is that control is acting as treatment would, had it not been for the treatment. Because this is an *encouragement*, it is difficult to be sure that 1) the treatment uniformly affect all those in the response-treatment condition and 2) those in the control didn't also spontaneously decide to make lifestyle



changes. Again, this is unusual and unlikely, given the nature of the behavior the authors were trying to change.

Holland alerts us to the importance of ALICE (additive, linear, causal effects). This is a more formal explanation of this assumption: that the effects of the treatment are constant and don't depend on some intrinsic property of the unit (in this case, participants)

## 2.2 Framing effects on Likelihood of Congressional Outreach

In this study, Brader et al. (2008) presumed that exposure to certain kinds of media regarding immigration policy would induce a emotional response amongst exposed individuals, which would increase the probability that they would want to write their Congressman. In this case, we have a treatment  $S$  of {Latino cue : European cue}. Our response variable is “emo”, which is a 10pt measure of anxiety. Our outcome variable is a binary treatment {0 : 1}, indicating whether a participant agreed to send a message to their Congressman or not.

This resembles an encouragement design, albeit more subtly than the previous example. Here, the encouragement is the frame; a Latino frame is ostensibly supposed to incur a different level of anxiety amongst participants as opposed to a European frame. This effect, in line with much social science, is presumed to be an implicit effect, as it is a psychological construct. Therefore, the measured variable “emo” is supposed to serve as an indicator as to whether the treatment elicited the appropriate response. Additionally, the measured variable “emo” is presumed to be a mediator on “cong\_mesg”, such that the probability that a person will submit a message to their Congressman is increased if anxiety is increased.

First, let's load the data and fit the models, *lm* models “model0”, “model1”, and “model2”.

```
> options(width=60)
> options(continue=" ")
> library(mediation)
> data(framing)
> attach(framing)
> model0 = lm(emo ~ treat)
> model1 = lm(cong_mesg ~ emo)
> model2 = lm(treat ~ cong_mesg)
> detach(framing)
```

### What are the significant effects?

**Treatment on “emo”:** According to the OLS model, the treatment does have a positive and significant effect on anxiety. Being in the treatment condition increases anxiety by 1.48 units. For females, this effect is smaller and less significant, 1.16 units.

	Model 0	Model 1	Model 2
(Intercept)	6.59*** (0.19)	-0.11 (0.07)	0.23*** (0.03)
cong_mesg			0.09 (0.06)
treat	1.48*** (0.38)		
emo		0.06*** (0.01)	
<i>N</i>	265	265	265
<i>R</i> <sup>2</sup>	0.05	0.14	0.01
adj. <i>R</i> <sup>2</sup>	0.05	0.13	0.01
Resid. sd	2.70	0.44	0.44

Standard errors in parentheses

† significant at  $p < .10$ ; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

Table 1: Males and Females

	Model 0	Model 1	Model 2
(Intercept)	6.76*** (0.27)	-0.12 (0.10)	0.27*** (0.04)
cong_mesg			0.01 (0.08)
treat	1.16* (0.52)		
emo		0.06*** (0.01)	
<i>N</i>	139	139	139
<i>R</i> <sup>2</sup>	0.04	0.12	0.00
adj. <i>R</i> <sup>2</sup>	0.03	0.11	-0.01
Resid. sd	2.72	0.43	0.45

Standard errors in parentheses

† significant at  $p < .10$ ; \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

Table 2: Females only

**Treatment on “cong\_mesg”:** According to the OLS model, anxiety does have a positive and significant effect on the probability that someone will send a message to their Congressman. For every unit increase in anxiety, the probability of sending a message increases by 6%. For females, this effect is the same.

**Effect of unit changes in “emo” on probability of sending message to Congressmen:** Upon reviewing model 2, we see that the slope is 0.09 and the intercept is 0.23. The slope is the change in outcome as treatment moves from 0 to 1. The intercept is the baseline probability. Interpreting these results according to those terms means that the baseline probability that someone will send a message to their Congressman, regardless of treatment or anxiety, is 23%. For females, the baseline is higher, 29%.

**Can we make the same assumptions with this encouragement design as with the sodium intake example in 2a?** No. There are several indicators that individual differences interact with the response to the treatment. First, there are baseline differences between males and females regarding their probability of sending a message to a Congressman regardless of their treatment. Second, the effect of the treatment on anxiety is different. Women do not react as strongly to the treatment with respect to elevated levels of anxiety, and yet the effect of anxiety on probability of sending a message to a Congressman is the same between males and females.

Consider, also, that there it’s not clear that those in the control condition should be considered identical to those in the treatment condition, aside from the framing encouragement. Perhaps there were additional unobserved events that occurred

ALICE

### 3 Question 3: Multilevel Data, Random effects models

#### 3.1 Compute the traditional quantities for “normexam” predicted by prior test score “standLRT”

There are two models that we can use to try to explore the relationship that this metric “standLRT” has with “normexam.” One is a simple OLS, regressing the school average score “schavg” on the outcome variable “normexam.” This model essentially assumes that there are no significant differences between students at various schools. A better model would be to try and account for the fixed effect of student variance within schools.

**Part A: Aggregation bias** First, let’s evaluate this relationship at the school level, using school-level means.

```
> #is the distribution of the schavg is the same as the actual within school mean standLRT
> options(width=60)
> summary(unique(schavg))

    Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
-0.75600 -0.24170 -0.04053 -0.03075  0.21030  0.63770

> summary(tapply(standLRT, school, mean))

    Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
-0.75600 -0.24170 -0.04053 -0.03075  0.21030  0.63770

> #Also, school 48 seems to be really throwing off the boxplot, so I'm going to get rid of it for now
> meannormexam = tapply(Exam$normexam, Exam$school, mean)
> Exam$meannormexam = meannormexam[Exam$school]
> Exam = Exam[!Exam$school %in% 48,]
> Exam_school = Exam[!duplicated(Exam[, 'school']),]
> model_agg = lm(meannormexam ~ schavg, data=Exam_school)

> options(width=60)
> summary(model_agg)

Call:
lm(formula = meannormexam ~ schavg, data = Exam_school)

Residuals:

    Min       1Q   Median       3Q      Max
-1.15701 -0.14005 -0.00513  0.19993  0.66298
```

Coefficients:

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.005325    0.040302   0.132    0.895
schavg      0.880943    0.118093   7.460 3.43e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.3216 on 62 degrees of freedom
Multiple R-squared: 0.473,      Adjusted R-squared: 0.4645
F-statistic: 55.65 on 1 and 62 DF,  p-value: 3.43e-10
```

According to the aggregated model, for every unit increase in “schavg,” normexam will increase by 0.881. The intercept is 0.00532. Thus, we might be tempted to say that the way to improve the normalized exam score is simply to accept the students with the highest entering scores. This is hardly compelling policy and begs the question: are all students within schools essentially identical?

**Part B: Contextual effects** What is the (supposed) estimate of the effect on “normexam” for a student being in a school with one unit higher mean score on “standLRT” and student “held constant?”. This is akin to asking, what is the average slope and intercept over all individual regressions “normexam” on “cstandLRT” fit for each school. As opposed to the aggregated model, the estimates for the slope and intercept are slightly different.

```
> #Centering within-school individual scores using school means
> library(lme4)
> Exam$cstandLRT = Exam$standLRT - Exam$schavg
> model_lmllist = lmList(normexam ~ cstandLRT|school, Exam)
> #fixed effects are the same as taking the mean Intercept and Slopes for the centered standLRT metric
> sapply(coef(model_lmllist), mean)

(Intercept)    cstandLRT
-0.01648814  0.53905282
```

According to this model, the true effect on normexam is 0.539, with the intercept at -0.0165. This is a rather large difference in slope than the aggregated model, which indicates an *aggregation bias* of 0.342.

### 3.2 Gender differences in outcome (the “gender gap”) for the coed (“mixed”) schools

**Part C: Give a point estimate and standard error for gender differences in outcome for these multilevel data** First, let’s formally define the models that we are looking at. We’re interested

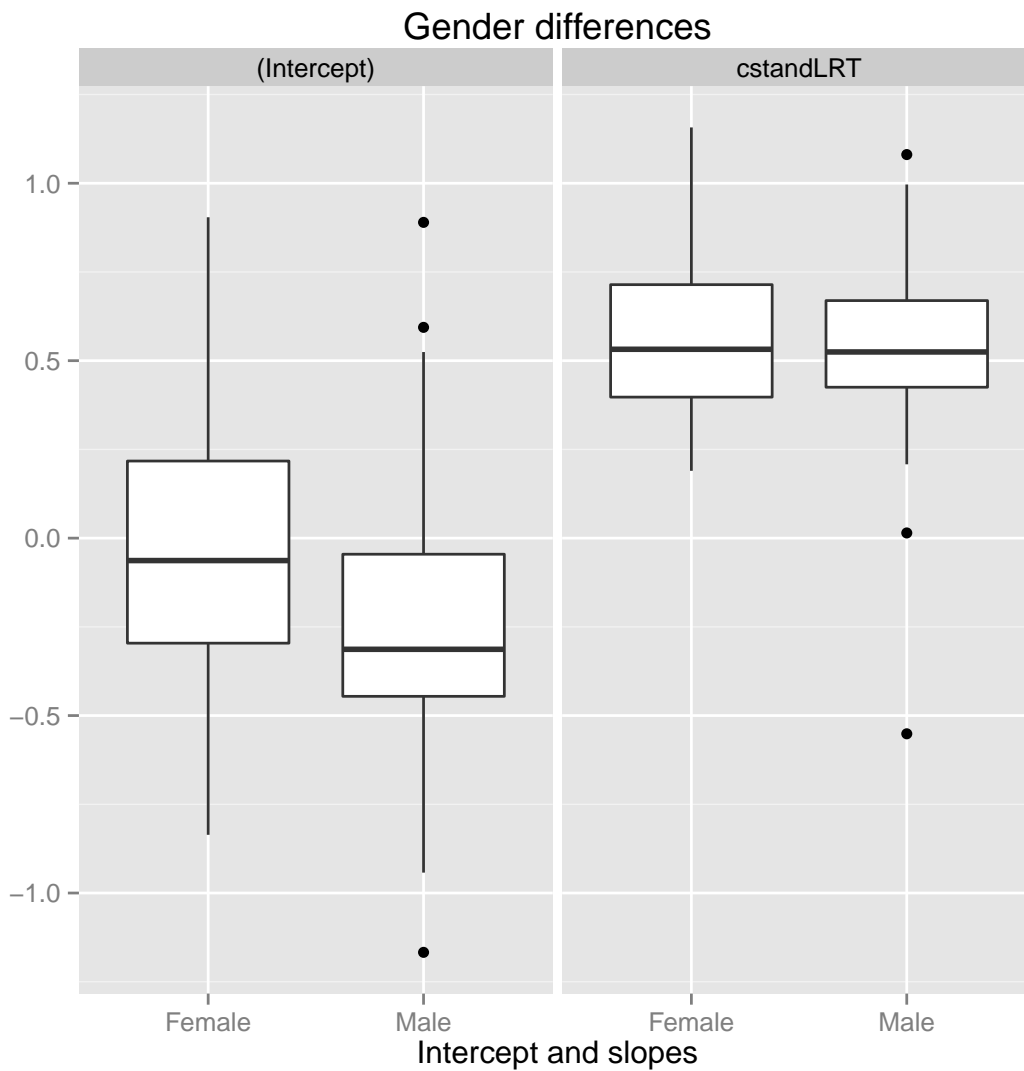
At another level, we recognize that each individual school might operate slightly differently, and that using a mean measure of a school actually results in a loss of information about the statistical relationships between the variables of interest. Level 1 is the within-schools regression, looking at each individual  $j$  in school  $i$ . Level 2 is over all schools, acknowledging the effect of average standLRT performance and gender.

**Level 1:**  $normexam = \alpha_{0i} + \alpha_{1i}cstandLRT_{ij} + \epsilon_{ij}$

**Level 2:**  $\alpha_{0i} = \gamma_{00} + \gamma_{01}schavg + \gamma_{02}sex_i + u_{0i}$

**Level 2:**  $\alpha_{1i} = \gamma_{10} + \gamma_{11}schavg + \gamma_{12}sex_i + u_{1i}$

```
> options(width=60)
> options(continue=" ")
> #First, subset down to only mixed-gender schools
> Exam_mxd = Exam[Exam$type == 'Mxd',]
> Exam_mxd$sex = factor(Exam_mxd$sex, levels=c('M','F'))
> #FYI, schools 43 and 47 barely count as coed, should drop them
> #table(sex, school)
> Exam_mxd = Exam_mxd[!Exam_mxd$school %in% c(43, 47),]
> #using lmList
> model_lv11_XX = lmList(normexam ~ cstandLRT|school, subset = sex == 'F', data=Exam_mxd)
> model_lv11_XY = lmList(normexam ~ cstandLRT|school, subset = sex == 'M', data=Exam_mxd)
> #using lmer
> library(lme4)
> model_lv12_lmer = lme4::lmer(normexam ~ schavg*cstandLRT + sex*cstandLRT +
                               (1 + cstandLRT|school), data = Exam_mxd)
```



As you can see from the graph, there is a gender effect on the intercept. The boxplot indicates that females have somewhat stronger performance within the schools on average, and the slopes on cstandLRT are relatively the same.

This is supported by the output below:

Linear mixed model fit by REML

Formula: normexam ~ schavg \* cstandLRT + sex \* cstandLRT + (1 + cstandLRT | school)

Data: Exam\_mxd

AIC BIC logLik deviance REMLdev

4591 4647 -2286 4547 4571

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
school	(Intercept)	0.0922683	0.303757	
	cstandLRT	0.0074176	0.086125	0.509

Residual                      0.5292113 0.727469

Number of obs: 2026, groups: school, 33

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.159155	0.058561	-2.718
schavg	0.934706	0.177825	5.256
cstandLRT	0.556726	0.028296	19.675
sexF	0.169522	0.033376	5.079
schavg:cstandLRT	0.133952	0.077442	1.730
cstandLRT:sexF	0.008816	0.035436	0.249

Correlation of Fixed Effects:

	(Intr)	schavg	cstLRT	sexF	sc:LRT
schavg	0.105				
cstandLRT	0.276	0.029			
sexF	-0.276	-0.004	-0.048		
schvg:csLRT	0.033	0.298	0.149	0.002	
cstndLRT:sF	-0.021	0.002	-0.534	-0.003	0.007

#### Part D: Do these gender differences have any association with “intake”

- This model is pegged against males as the baseline. The fixed-effect of -0.159 is the average level of performance on normexam of males.
- The coefficient for sexF is 0.17 and it represents the average performance of females on the normexam relative to males. In other words, females are performing better than males, on average, within schools.
- The coefficient for schavg is 0.935, which is a school’s average normexam performance to their average standLRT scores. in other words, for each unit increase in schavg, normexam will increase by 0.935 on average.
- The coefficient for cstandLRT is the average slope for standLRT for all males in these schools. Contrast that with the coefficient for cstandLRT:sexF, which is the average slope for all females. The slope is smaller, on average, for females.
- The coefficient for schavg:cstandLRT is 0.134, which represents the average change in the within-school cstandLRT slope given a one-increment increase in the school average. In other words, if a school’s average (schavg) were to increase by one unit, the slope for that school’s cstandLRT would increase by 0.935.