Some Basic Item Bias Analyses for Ability and Knowledge Tests

Item Difficulty

Once your variables are scored 0 for incorrect and 1 for correct, we can investigate potential bias by comparing the item difficulty (means) in two groups. Below, I compare males and female students on an exam from another course.

Menus

Download the data file adaexam.sav from http://web.pdx.edu/~newsomj/data.htm and open the file in SPSS. We can get all of the means separately for the two groups using the MEANS procedure.

Analyze -> Compare Means-> *Means*…
Move *gender* variable to the *Independent Variable List* and *q1* through *q10* to the Dependent Variable *List*. Click OK.

Syntax[1]

(Before running a syntax file, find your downloaded data file and drag it over to the desktop)

```
GET FILE=' C:\Users\newsomj\Desktop\adaexam.sav'.

Means q1 TO q10 by gender.
```

Item Discrimination Index

Computation of the item discrimination index requires a total score on the test, which I have already computed for the exam data. For your project you will first need to create a new variable which is a composite summed score, using the Transform menu or COMPUTE command in syntax. I called my variable *mctotal*.

Find the Percentiles for the Total Score

Menus

Analyze -> Descriptive Statistics -> Frequencies…

Highlight the scale total score (*mctotal*) on the left and move it over to the box on the righthand side using the arrow button in the middle.

Click on *Statistics* button. Under *Percentile Values*, check the box next to *Percentile(s)*. Then enter values, such as 33 (click add) and 67 (click add), which is the Kelly (1939) method. [note: percentiles need to be in whole numbers not decimals]. Other values can be used.

Syntax

(Before running a syntax file, find your downloaded data file and drag it over to the desktop)

```
GET FILE='C:\Users\newsomj\Desktop\examdata.sav'.

FREQUENCIES VARIABLES=mctotal
  /percentiles=33 67.
```

Obtain the Proportion Correct for Top and Bottom Scorers for the Two Groups

Menus

Create high and low score groups using the cutoff values obtained above:
Transform-> Compute Variable…
Enter a new variable, such as *scoregrp* under Target Variable and a value, such as 1, under Numeric Expression

---

[1] If the lab SPSS on the computer will not access the data file, the following syntax may work instead:

```
CD ' C:\Users\newsomj\Desktop'.
GET FILE='adadata.sav'.
```

Then, click on *If…* button at the bottom. Check *Include if* case satisfies condition, and enter into the box: *mctotal* <= 15
Click ok.

Repeat the above but using *scoregrp* = 0 and *mctotal* > 19
Click ok.

Data -> Sort Cases…
Move over *gender* and *scoregrp*

Data -> Split File… (caution: do not use the first item "Split Into Files")
Check Compare Groups
Move over *gender* and *scoregrp*

Analyze -> Descriptive Statistics-> Descriptives…
Move over the desired variables (e.g. Q1 through Q10)
Click OK

Note: if you do not have enough cases, for any of the subgroups, then try using less extreme percentiles (e.g., you can use above and below the 50th percentile instead).

Syntax
This syntax computes the high and low group variable, `scoregrp`, sorts the file, splits the file, and requests means for all items separately for high and low and male and female.

```
DO IF mctotal LE 15.
COMPUTE scoregrp=1.
ELSE IF mctotal GT 19.
COMPUTE scoregrp=2.
END IF.

SORT CASES BY gender scoregrp.
SPLIT FILE BY gender scoregrp.
DESCRIPTIVES VARIABLES=q1 TO q10.
```

Compute Discrimination Index for All Items for Each Group
Download the MS Excel sheet item discrimination index.xlsx from the data page: http://web.pdx.edu/~newsomj/data.htm and enter values from the output.

R

```
library(lessR)
#you will need to change your location
d = Read("C:/Jason/SPSSWIN/pmclass/adaexam.sav", quiet=TRUE)

#get upper and lower terciles
pivot(d, quantile, c(mctotal),q_num=3)

#subset data frame for females and males
male <- d[.(gender==0), .(q1:q10,mctotal)]
female <- d[.(gender==1), .(q1:q10,mctotal)]

#get upper and lower third quantiles for males
mlower <- d[.(mctotal<=15), .(q1:q10)]
pivot(mlower, c(mean,sd), c(q1,q2,q3,q4,q5,q6,q7,q8,q9,q10))
mupper <- d[.(mctotal>=19), .(q1:q10)]
pivot(mupper, c(mean,sd), c(q1,q2,q3,q4,q5,q6,q7,q8,q9,q10))

#get upper and lower third quantiles for females
flower <- d[.(mctotal<=15), .(q1:q10)]
pivot(flower, c(mean,sd), c(q1,q2,q3,q4,q5,q6,q7,q8,q9,q10))
fupper <- d[.(mctotal>=19), .(q1:q10)]
pivot(fupper, c(mean,sd), c(q1,q2,q3,q4,q5,q6,q7,q8,q9,q10))
```