PLSC 501

Exercise # 5

Andrew Bennett

Please complete the following tasks. I encourage you to work together, but your answers must be your own, and you must show in your work how you arrived at the answers. Please turn in your R-command file via email (you must use R for this exercise). The assignment itself should be submitted via email as a PDF file compiled in LATEX by the beginning of our next Lab, on Tuesday April 13th, 2021.

--Please make tables and figures as professional-looking as possible; do not paste Stata or R output into your documents.

The goal of this exercise is to get you familiar with R, and practice some of the basic skills you have already learned in Stata, specifically: importing data, running regressions, and making tables/plots.

1.—You will need to install the following packages to complete this assignment.

```
install.packages("stargazer")
install.packages("foreign")
install.packages("ggplot2")
install.packages("haven")
install.packages("swirl")
install.packages("readstata13")--> you might need this if you use Stata 13
```

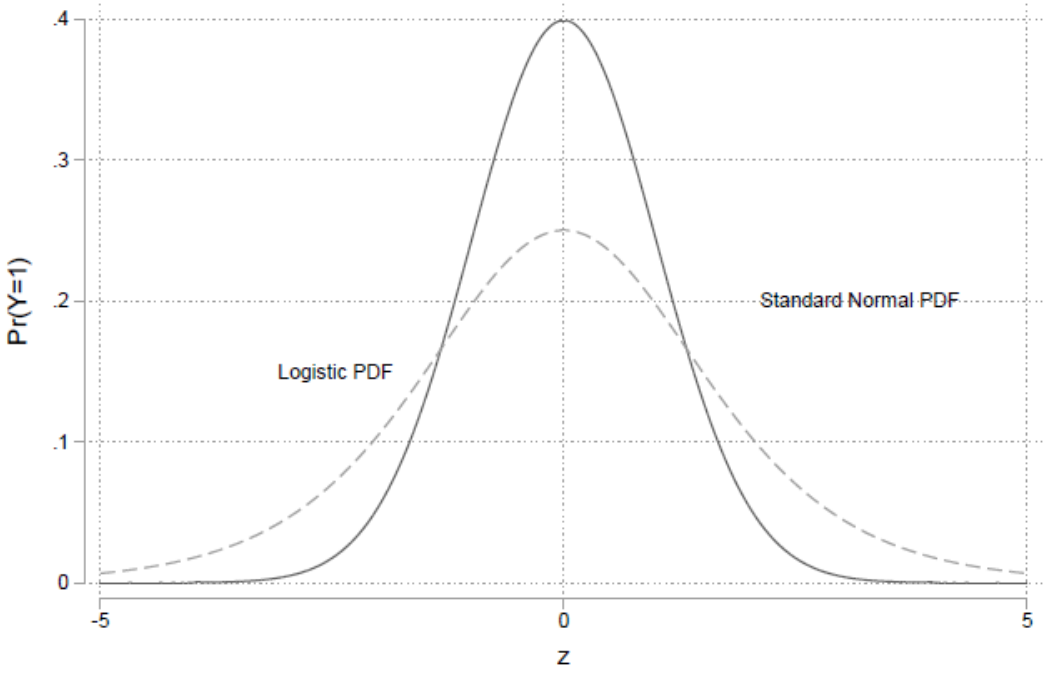
Once you have done so, play around with the "swirl" package (which is an R training guide) for a few hours, or until you feel comfortable with the basics of R. To use swirl, type the following into your R-command window. You can come back to swirl any time when you want more R training.

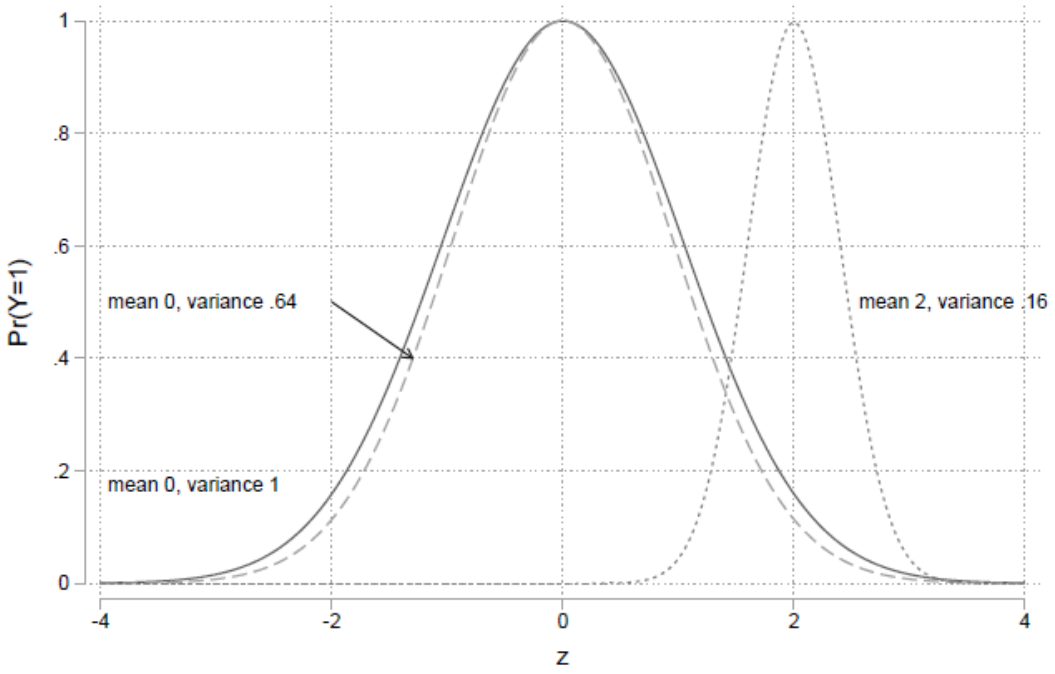
```
library("swirl")
swirl()
```

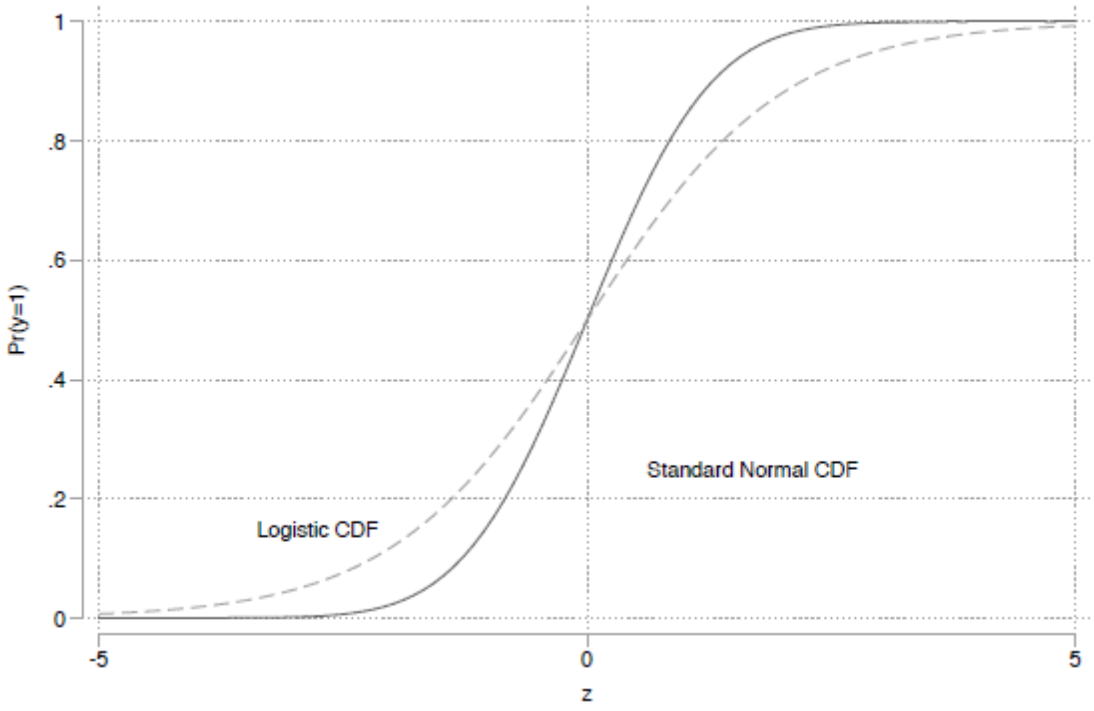
2.—Using the "grade data" from Exercise 1, import the already cleaned data set from Stata (as a dta-file). Once again, create a model with *finalgrade* as your DV and *discfinal* (discussion grade), *schoollevel*, *gender*, and *major* as your IVs. Present the regression results in a professional looking table. The library(haven) package will be needed to import the data, and

the library(stargazer) package will be needed to transform R's regression output into a Latex table.

3.—Replicate **all 3 plots** from below, fully labeled as always. You should have the needed formulas for the normal PDF and logistic PDF from Exercise 1, but will need to look up the formulas for the normal and logistic CDF (hint—the logistic CDF is merely the Logit link function that we used in the previous problem set, and the normal CDF has a built in command in both R and Stata, so I would recommend you locate and use that to make your lives easier. With the exception of the Normal CDF plot, **please do NOT use built in commands**, I want ALL distributions plotted by hand using the Normal/Logistic PDF and Logistic CDF formulas.




1
.8
.6
.4
.2
0
Pr(Y=1)
mean 0, variance .64
mean 2, variance .16
mean 0, variance 1
-4
-2
0
2
4
z


1
.8
.6
.4
.2
0
Pr(y=1)
Standard Normal CDF
Logistic CDF
-5
0
5
z