QUIZ 2 Rubrics

a) Write code to read the categorical data in qz2\_1\_dat.txt (on the course

website) into R, call the first column x, and use the plot() function to make a histogram for x. Hint: you need to literally look at the txt file.

dat = read.table("http://sites.stat.washington.edu/marzban/390/spring21/qz2\_1\_dat.txt", header=T) # 1 point. 0.5 for header option

x = dat[,1]

plot(as.factor(x)) # 1 point for using as.factor(x)

b) Run table(x), look at what it has returned, and reverse engineer what it has done in this case. Explain it in words.

It has returned the frequency of each level of x. # 1 point

c) Write code to make a histogram of x, this time using the function table(). Feel free to experiment, and again, don't worry about the labels of the x, y axes, or the width of the bars, because for us anything that graphically displays frequencies is a good-enough histogram.

plot(table(x)) # 1 point for using table(x)

#only give 0.5 if does hist (table(x))

d) Based on everything we learned above, write code to make a \*relative\* frequency histogram of y. Although, usually, a histogram does not have to be displayed with vertical bars, for this question it should be. So, consult the help pages to see how to display vertical bars. Hint: Look at "type." The width of the bars, however, is NOT important. The axis labels and titles are also NOT important.

plot(H$mids, H$counts/length(y), type="h") # 1 point for H$mids and H$counts; 1 point for division of length(y) or the likes.

1pt for getting correct relative frequencies H$counts/length(y) or H$counts/sum(H$counts)

0.5pt for H$mids as the x-axis in plot. A max of 1 point for using hist in case the relative frequencies are correct.

0.5pt for type= “h”

0pt for hist(y, freq = FALSE), which gives density not relative frequency

e) Write code to make a histogram of y with about 100 breaks; BUT, instead of showing the frequencies on the y-axis, show the log() of the frequencies.

H = hist(y, breaks=100)

plot(H$mids, log(H$counts), type = "h") # 1 point for taking log of freqs

-0.25 if not specifying breaks

-0.25 if not using H$mids as x-axis or not specifying type = “h”

Give 0 if students does log(y) or not taking log on counts

# Total: 7 points