Brian Kang

HW Set #3

# 1.30

d. Prop(x>5) = 1-1 = virtually **0**, prop(x>-5) = 1-0 = virtually **1**

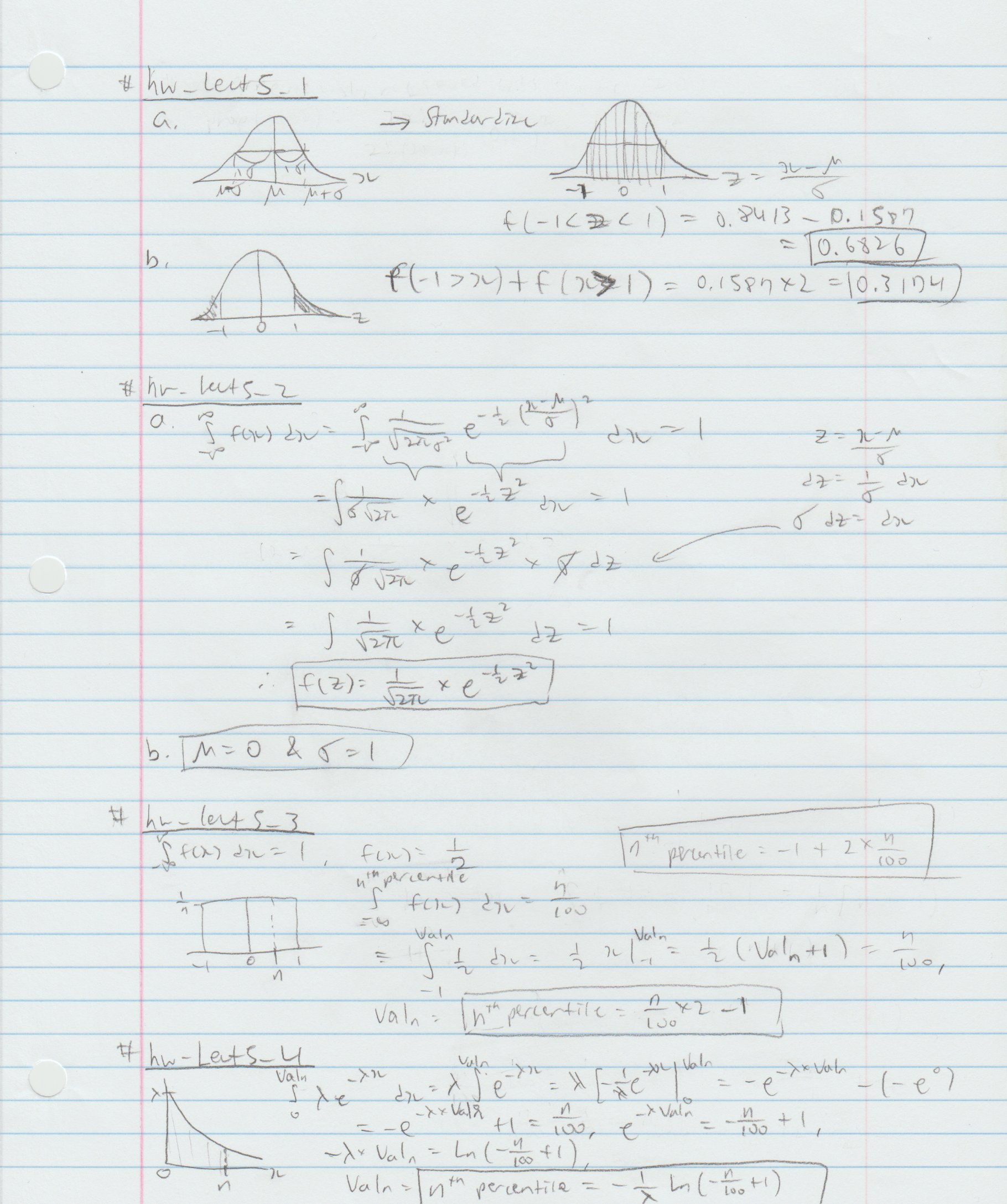
e. prop(abs(z)<2.5) = prop(-2.5<z<2.5) = 0.9938-0.0062 = **0.9876**

#1.32

e. z#, where prop(z\*) = 1-0.002 = 0.998, z\* = 2.88, so **2.88 right of 0**

where prop(z\*) = 0.002, z\* = -2.88, so equally, **2.88 left of 0**

# hw\_lect5 problems below…



#hw\_lect6 problems below…

#1.55

“number of erroneously received bits out of 20 bits” -> binomial distribution

a. prop(x = at most 2 = 0,1,2) = 0.2852+0.2702+0.1216 = **0.6770**

prop(x=2) = 20!/(2! \* 18!) \* (0.1)^2 \* (0.9)^18 = 0.2852

prop(x=1) = 20!/(1! \* 19!) \* (0.1)^1 \* (0.9)^19 = 0.2702

prop(x=0) = 20!/(0! \* 20!) \* (0.1)^0 \* (0.9)^20 = 0.1216

b. prop(x = at least 5) = 1 - prop(x = 0,1,2,3,4) = 1-0.6770-0.1901-0.0898 = **0.0431**

prop(x=3) = 20!/(3! \* 17!) \* (0.1)^3 \* (0.9)^17 = 0.1901

prop(x=4) = 20!/(4! \* 16!) \* (0.1)^4 \* (0.9)^16 = 0.0898

c. prop(x = more than half = greater than 10 = 11,12,13,…20) = 1 - prop(x=0,1,2,…,10) = 1 – virtually 1 = **virtually 0**

**# hw\_lect6\_1**

# install.packages(c("survival", "Zelig"))

# library(survival)

library(Zelig)

# 2 categorical/discrete: barb2 and gdpw2

# 2 continuous: prsexp2 and prscorr2

data("PErisk")

dat <- PErisk

# subset the function to extract the columns (or variables) I want

modifiedDat <- dat[,c("barb2", "prsexp2", "prscorr2", "gdpw2")]

# save data of barb2 = Black Market Premium to variable,

# for each level of prsexp2 = Lack of Expropriation Risk

cont0 <- modifiedDat[modifiedDat$prsexp2==0,]$barb2

cont1 <- modifiedDat[modifiedDat$prsexp2==1,]$barb2

cont2 <- modifiedDat[modifiedDat$prsexp2==2,]$barb2

cont3 <- modifiedDat[modifiedDat$prsexp2==3,]$barb2

cont4 <- modifiedDat[modifiedDat$prsexp2==4,]$barb2

cont5 <- modifiedDat[modifiedDat$prsexp2==5,]$barb2

# sort each unique level

lv <- sort(unique(modifiedDat$prsexp2))

# horizontal boxplot

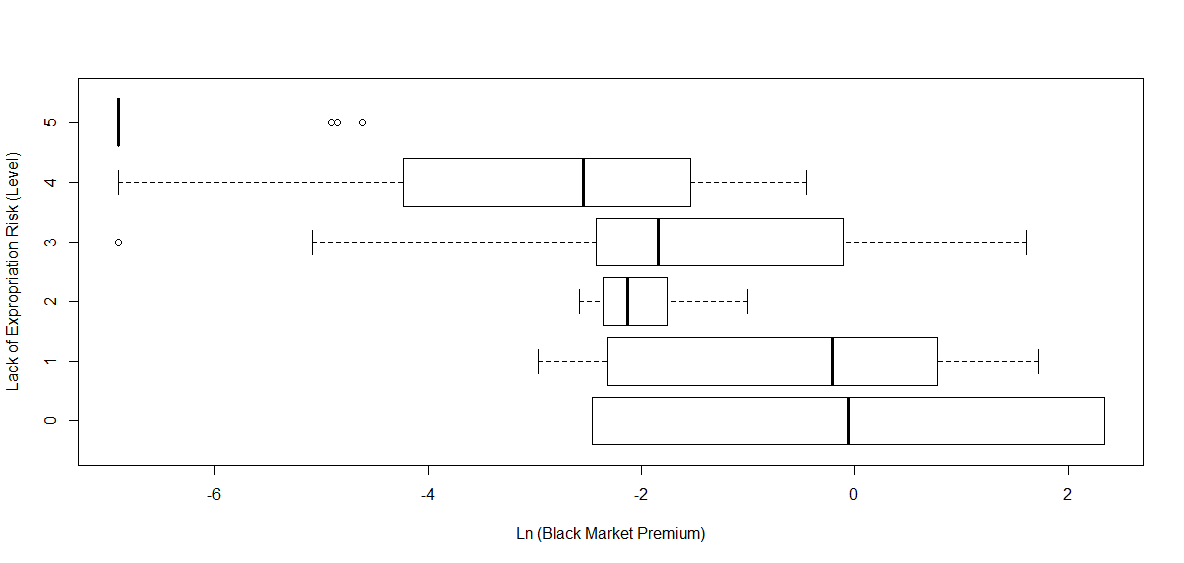
boxplot(cont0, cont1, cont2, cont3, cont4, cont5,

ylab = "Lack of Expropriation Risk (Level)",

xlab = "Ln (Black Market Premium)",

horizontal = T, at = lv)

axis(2,lv)



# Interpretation:

# I think we can see a rough trend that overall the level of LACK

# of expropriation risk from a government is inversly related to

# the black market premium (high level <-> lower premium). For the

# box plot of level 5, it looks like there were many numbers with the

# same value that the min,1st 2nd 3rd quartiles,max are all equal, so

# producing outliers outside of that compact range. Level 4,3 both have

# a wide spread and similar sizes of range and the box. Level 2 is

# relatively compact, having a small box and range. Level 1 has a big

# box compared to its "whiskers," indicating that there are more

# values in the sorted center than toward the high and low end.

# Level 1 has a very wide box with no "whiskers" (our data indicates

# that this level only has 2 values). It is difficult to clearly say

# that each boxplot is located "higher or lower" due to their spread

# and uncertain true center of the plot, especially when comparing the

# levels 3 & 2 and 1 & 0.

# --------------------------------------------------------------------

**# hw\_lect6\_2**

**# Part II**

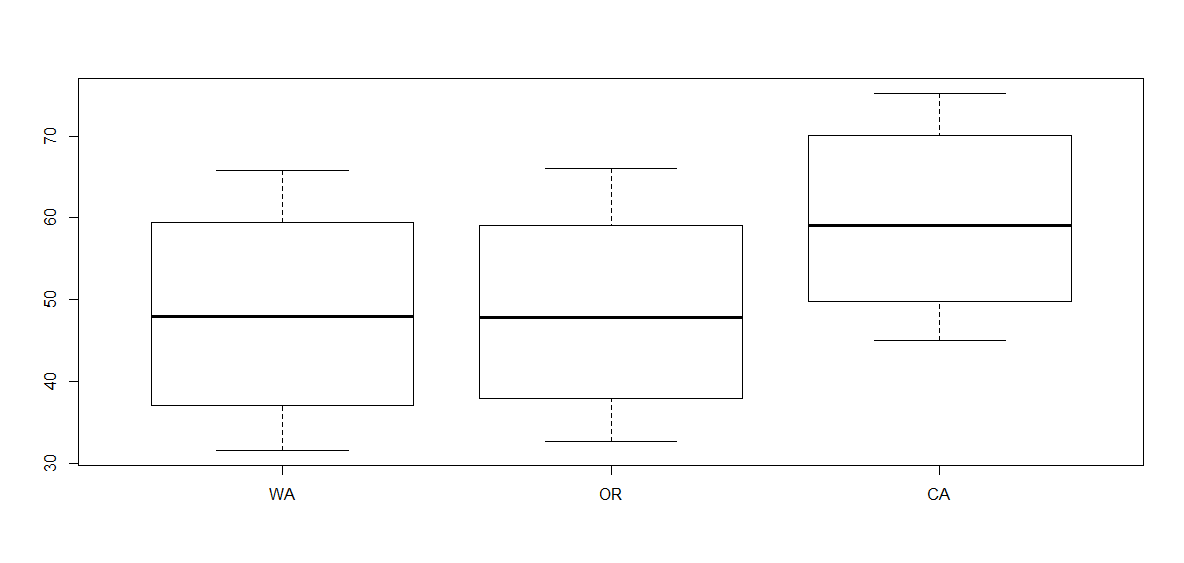
# d)

datt <- read.table(

"https://www.stat.washington.edu/marzban/390/summer18/temp\_dat.txt",

header = T)

boxplot(datt)



# e)

# There is NO evidence that OR is warmer than WA. All the values

# displayed on the boxplot seem to nearly identical. They have very

# similar range, min, max, IQR, median, 1st and 3rd quartiles. Due

# to the huge overlap of data according to the boxplot, we cannot

# concluded that OR is warmer than WA.

# f)

# I think there is reasonable data to say that CA is warmer than WA.

# Although CA seems to have similar spread (range and IQR) and similar

# location of center within the IQR, around the middle, the values

# themselves indicate that they are approximately 15 units higher.

# Therefore, I think it is understandable to say that the mean of

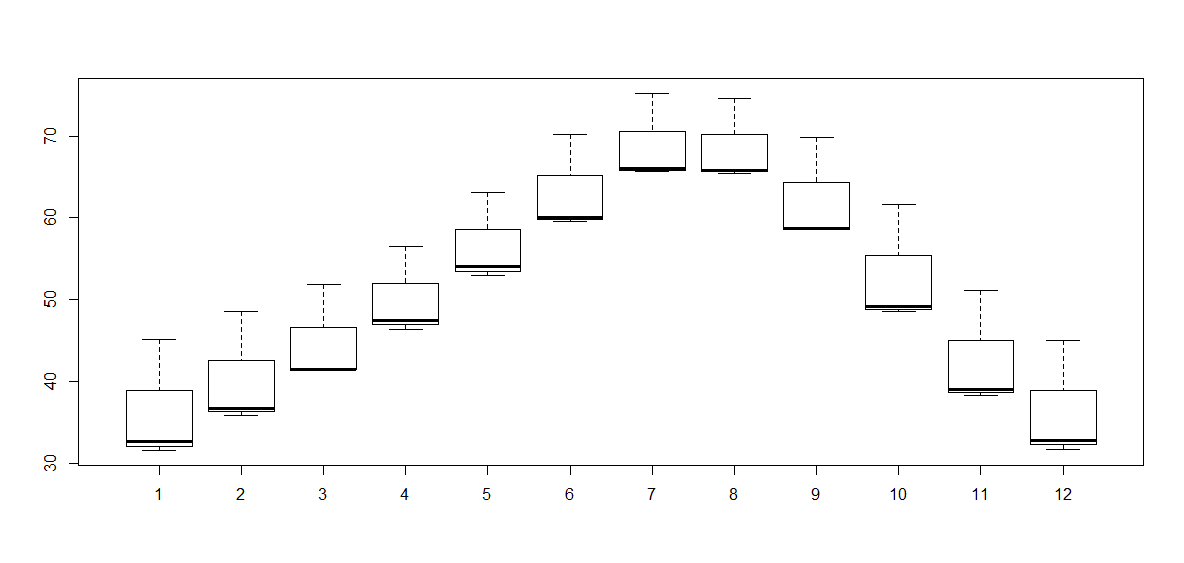
# temperature at CA is higher than that of WA, so I say overall it is

# reasonable to conclude CA is warmer than WA.

# g)

Tdatt <- t(datt)

boxplot(Tdatt)



# h)

# All WA, OR, and CA are in the westcoast. There is NO evidence that

# temperatures in July and August are different in the westcoast for

# similar reasons listed in (e). The boxplots seem to be almost

# identical; the heavily skewed shape (due to only having 3 values),

# the spread of range and box, the center median all are very similar.

# Due to this huge overlap we cannot reach the conclusion that July

# is warmer than Auguest in the westcoast.

**# Part III**

# i)

dattt <- as.matrix(read.table("https://www.stat.washington.edu/marzban/390/summer18/inter\_arrival\_dat.txt", header = T))

# "inter arrival time" is an example of exp. dist.

# hist(dattt) # to check that this is expoential

quantile(dattt, probs = 0.10) # shortest 10%

# gives the value, answer = 0.1608677