QUIZ 8: (questions are in italics).

a) Write code to compute (and report) the coverage count for a 95% CI for mu2-mu1. Specifically,

- Let the first populations be N(1,2), and the second population be N(1.5,2).

- Let the two samples sizes be 100 and 110, respectively, and

- The number of trials (CIs) be 1000.

- Do NOT use the function t.test(), i.e., use the formulas we have developed in class.

- In those formulas, use the appropriate z\* (not t\*), which should be found in R, not by reading Table 1.

- Start you code with set.seed(1).

- Report the coverage count.

- Hint: revise the code from the previous lab.

n1 = 100

mu1 = 1

sigma1 = 2

n1 = 100

mu2 = 1.5

sigma2 = 2

n2 = 110

n.trial = 1000

set.seed(1)

CI = matrix(nrow=n.trial, ncol=2)

for (i in 1:n.trial) {

x1 = rnorm(n1, mu1, sigma1)

x2 = rnorm(n2, mu2, sigma2) #1pt for sampling x1,x2 with correct parameters, 0.5pt each. -0.25 for each wrong argument

lower = (mean(x2) - mean(x1))-qnorm(.975)\*sqrt(sigma1^2/n1 + sigma2^2/n2)

upper = (mean(x2) - mean(x1))+qnorm(.975)\*sqrt(sigma1^2/n1 + sigma2^2/n2)

#1pt for correctly calculating the CI. No point if used t.test(). -0.5 if used 1.96 instead of qnorm(.975).

#No point deducted if used sample sd instead of sigma1, sigma2

CI[i,] = c(lower,upper)

}

cnt = 0

for (i in 1:n.trial) {

if (CI[i,1] <= mu2-mu1 & CI[i,2] > mu2-mu1)

cnt = cnt+1

}

cnt # 953 . If all is good the count should be around 950.

#1pt for correctly updating and reporting the coverage count. -0.25 if calculated correctly but numerical value not reported.

#3pts total

Part II: Paired and unpaired t-test on paired data.

b) What is it about that data which qualifies it as being \*paired\*? Important: Provide a line of code that supports that claim. If you don't know the answer to this question, proceed to the next parts, first, and then return to this one.

plot(x1,x2) # the scatterplot shows an association. As discussed in the lecture, that's one of the signatures of paired data.

#0.5pt for plot or used cor().

#0.5pt for mentioning “association” “correlation”.

#1pt total

c) Use the function t.test() to report the 95% CI for mu2-mu1 for un-paired data.

t.test(x2,x1)$conf.int # -0.2761204 0.2282966

#1pt for correct CI. -0.25 if numerical value not reported.

#-0.5 if specified alternative = “greater” or “less”

#fine if used wrong data for this question as long as used t.test with correct arguments

d) Use the function t.test() to report the 95% CI for mu2-mu1 for paired data.

t.test(x2,x1,paired=T)$conf.int # -0.14411895 0.09629509

#1pt for paired = T. -0.25 if numerical value not reported, but do not deduct this if deducted 0.25 for previous question.

e) Write code to compute the CI found in part d), but "by hand," i.e., without using the function t.test(), but using the formula we developed in lecture. Make sure to find the value of t\* in R (not from Table 6).

d = x2-x1

lower = mean(d)-qt(.975, df=n-1)\*sd(d)/sqrt(n)

upper = mean(d)+qt(.975, df=n-1)\*sd(d)/sqrt(n)

#1pt for correctly computing the CI.

#-0.5 if used qnorm() instead of qt().

#no point if used the formula for unpaired test

c(lower, upper) # -0.14411895 0.09629509