*YOUR TA’S NAME*:

*Lecture Worksheet*

*Thursday 11/5/2020*

**MAIN POINTS OF LECTURE**

1. Assumptions that must hold for hypothesis tests to be valid
	1. The functional form of the relationship between X and Y is appropriately specified; usually this means checking for linearity
	2. There are no extreme outliers
	3. The variability of the prediction errors is constant across the observed values of X (assumption of homoskedasticity)
	4. The values ofY are normally distributed at each value of X (assumption of normality)
	5. The observations are independent
2. Hypothesis Tests About 2YX
	1. R2YX is a sample estimate of population parameter 2YX
	2. If 2YX equals zero, then X does nothing to explain variability in Y
	3. This is a one-sided test because 2YX cannot possibly be less than zero
	4. Critical value is F with dfNUM=1 and dfDENOM=N—2
	5. Test statistic is 
3. Hypothesis Tests About YX
	1. rYX is a sample estimate of population parameter YX
	2. If YX equals zero, then there is no correlation between X and Y
	3. The critical value is a Z score
	4. Test statistic is where 
4. Hypothesis Tests About Slope YX
	1. bYX is a sample estimate of population parameter YX
	2. If YX equals zero, then the regression of Y on X has a zero slope
	3. Critical value is t with n-2 degrees of freedom
	4. Test statistic is 
5. Confidence Intervals for Slope YX
	1. where t\* is a t value with n-1 degrees of freedom

**QUESTIONS**

Mean of X: 6.50 SD of X: 2.95

Mean of Y: 7.15 SD of Y: 1.46

rXY: 0.51 n: 20

1. [From the recorded lecture] Test the hypothesis that 2YX --- the population proportion of variation in Y explained by X --- is zero in the population; use =0.05

With dfNUM=1 and dfDENOM=N-2=18 and =0.05, the critical value F\*=4.41

To compute the F statistic, we need:

$SS\_{TOTAL}=(s\_{Y}^{2})(N-1)$ = (1.462)(19) = 40.5

$SS\_{REGRESSION}=(R\_{YX}^{2})(SS\_{TOTAL})$ = (0.512)(40.5) = 10.53

$SS\_{ERROR}=SS\_{TOTAL}-SS\_{REGRESSION}$= 40.5-10.53 = 29.97

$$F\_{1,N-2}=\frac{SS\_{REGRESSION}/1}{SS\_{ERROR}/N-2}=\frac{10.53/1}{29.97/18}=6.3$$

Since F exceeds the critical value F\*, we reject the null hypothesis.

1. [From the recorded lecture] Test the hypothesis that YX --- the population correlation between X and Y --- is zero in the population; use =0.05

With =0.05 for this two-sided test, the critical value Z\*=1.96

$$Z\_{r}=\left(\frac{1}{2}\right)ln\left(\frac{1+r\_{YX}}{1-r\_{YX}}\right)=\left(\frac{1}{2}\right)ln\left(\frac{1+0.51}{1-0.51}\right)=0.563$$

$$Z=\frac{Z\_{r}-0}{\sqrt{1/N-3}}=\frac{0.563-0}{\sqrt{1/17}}=2.32$$

Since Z exceeds the critical value Z\*, we reject the null hypothesis

1. [From the recorded lecture] Test the hypothesis that YX --- the population slope relating Y to X --- is zero in the population; use =0.05

$$b\_{YX}=r\_{YX}\frac{s\_{Y}}{s\_{X}}=0.51\frac{1.46}{2.95}=0.252$$

With =0.05 for this two-sided test, the critical value t\* with df=N-2=18 is 2.101

$$t\_{N-2}=\frac{b\_{YX}-0}{s\_{b}}=\frac{b\_{YX}-0}{\sqrt{\frac{MS\_{ERROR}}{\left(s\_{X}^{2}\right)\left(N-1\right)}}}=\frac{0.252-0}{\sqrt{\frac{29.97/18}{(2.95^{2})(19)}}}=2.51$$

Since t exceeds the critical value t\*, we reject the null hypothesis

1. [From the recorded lecture] Construct a 95% confidence interval for YX --- the population slope relating Y to X

$C.I.=b\_{1}\pm t\*\sqrt{\frac{MS\_{ERROR}}{\left(s\_{X}^{2}\right)\left(N-1\right)}}$where t\* is a t value with n-2 degrees of freedom. So in this case, t\*=2.101

So, $C.I.=b\_{1}\pm t\*\sqrt{\frac{MS\_{ERROR}}{\left(s\_{X}^{2}\right)\left(N-1\right)}}=0.252\pm 2.101\sqrt{\frac{29.97/18}{(2.95^{2})(19)}}=0.252\pm 0.211$.

This, the 95% confidence interval ranges from 0.041 to 0.463.