QUANTITATIVE TECHNIQUES II

MIDTERM EXAM
A C T U A L

NAME: _____ **PART 1: In-Class**

1. Which of the following represents the Binomial Distribution?

a. $\binom{n}{x} p^{x} q^{1-x}$, where $q = x - p$

b. $\binom{n}{x} p^{x} q^{n}$, where $q = 1 - p$

c. $\binom{n}{x} x^{x} q^{n-x}$, where $q = 1 - p$

d. $\binom{n}{x} p^{x} q^{n-x}$, where $q = 1 - p$

e. All of the above.

f. None of the above.

2. Which of the following represents the Poisson Distribution?

a. $\dfrac{e^{-\lambda} \lambda^{x}}{x!}$

b. $\dfrac{e^{-\lambda} e^{\lambda}}{x!}$

c. $\dfrac{e^{-\lambda} \lambda^{x}}{3.1415\,!}$

d. $\dfrac{e^{-\lambda} x^{\lambda * e}}{\lambda!}$

e. All of the above.

f. None of the above.

3. Which of the following represents the Hypergeometric Distribution?

a. $$\frac{\binom{k}{x}\binom{X-x}{N-n}}{\binom{k-N}{X-n}}$$

b. $$\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

c. $$\frac{\binom{N}{X}\binom{N-n}{X-x}}{\binom{K}{k}}$$

d. All of the above.

e. None of the above.

4. Which of the following represents the Normal Distribution?

a. $$\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(N+n)^2}{2\lambda}}$$

b. $$\frac{1}{\sqrt{2e}}e^{\frac{-(\lambda-e)^2}{2\sigma^2}}$$

c. $$\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

d. All of the above.

e. None of the above.

NAME: _____ **PART 1: In-Class**

5. How many degrees of freedom are there in a 2 x 2 contingency table?
   a. 2
   b. 3
   c. 4
   d. 5
   e. None of the above

6. How many degrees of freedom are there in a 3 x 3 contingency table?
   a. 2
   b. 3
   c. 4
   d. 5
   e. None of the above

7. What is the formula for the POPULATION mean?

8. What is the formula for the POPULATION standard deviation?

9. In simple linear regression, if your error terms are <u>not constant</u>, is it good or bad?

10. Circle either “Constant” or “Non-Constant” for the following:
    a. Heteroscedasticity: Constant or Non-Constant ?
    b. Homoscedasity: Constant or Non-Constant ?

11. If $Y = \beta_0 + \beta_1 X$ is used for the equation for a line,
    a. What represents the intercept of the line?
    b. What represents the slope of the line?

12. In simple linear regression, what is the formula for the 'slope' coefficient?

13. In simple linear regression, what is the formula for the 'intercept' coefficient?

14. How would you express multiple regression with 1 'Y variable' and 4 'X variables'?

15. Which of the following would not represent a continuous variable:.
    a. Gender
    b. Weight
    c. Time
    d. Length

16. Which of the following is considered a dichotomous variable:.
    a. Gender: Male, Not Male
    b. Age: Child, Not Child
    c. Any Binary variable
    d. All of the above

17. In experimental analysis, you can do which of the following?
    a. Infer Causality
    b. Identify Correlations
    c. Both
    d. None of the above

NAME: _____ **PART 1: In-Class**

18. Circle whether each of the following descriptions of variables is considered an "X" or a "Y" variable in regression:

    a. Independent Variable: X or Y
    b. Dependent Variable: X or Y
    c. Explanatory Variable: X or Y
    d. Predictor Variable: X or Y
    e. Response Variable: X or Y
    f. Criterion Variable: X or Y

19. What is a common, (if not the most common), situation where dummy variables are used in regression analysis? (State example, if necessary.)

20. If we have a categorical variable with, say, 4 levels, how many dummy variables should be created for input into our regression model?

21. Pretend we are taking an opinion poll regarding President Bush's decision to invade Iraq. If we are concerned that opinion might change during the survey, what must we do, as each data point is collected, to help prevent our conclusions from being confounded by time?

22. In multiple regression, what plot gives us a quick way to determine non-constant error?

23. In multiple regression, if we want to be more thorough in the determination of non-constant error terms, what series of plots would give us insight?

24. In the study of correlation what is the formula for the r?

25. What is the meaning of r^2, (that is R^2) ?

26. What is the range of values that r^2 can maintain?

27. What is the "ideal" r^2 value?

28. Assume that you have achieved almost the "ideal" r^2 value, which would be good. In terms of your error term plot, how could regression be inappropriate?

29. What might we consider doing if our X and Y variables have a curvilinear relationship?

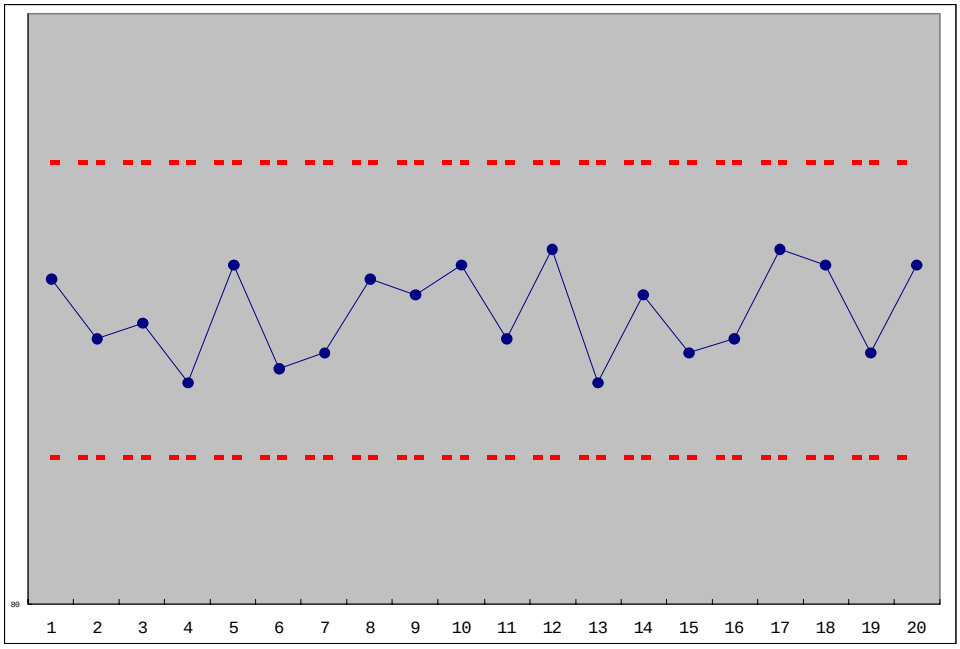
30. What is "The Curse of Dimensionality"?

31. In multiple linear regression, beyond how many continuous "X Variables" do we begin to experience serious challenges justifying the assumptions of regression?

32. What is the formula for the "Odds of p"?

33. What analysis tool is useful if our Y variable has the type yes/no or success/failure?

34. If the following chart represented a quality control chart, would the process be “In Control” or “Out of Control”?



35. Which of the following is the Capability Index:

a. $Cp = \frac{6*(U-L)}{x_i}$

b. $Cp = \frac{U-L}{6\sigma}$

c. $Cp = \frac{U-\lambda^x}{6\mu}$

d. $Cp = \frac{U-\sigma}{U-L}$

e. All of the above.

f. None of the above.

36. Who is responsible for setting the UCL (Upper Control Limits) and LCL (Lower Control Limits) of the chart?

NAME: _____ **PART 1: In-Class**

37. A real estate developer wants Northwest Realty Trust, Inc., to develop a regression model to predict the number of housing starts Y in their region from the interest rate on long-term Treasury bills X_1, the federal funds short-term interest rate X_2, and the short-term interest rate on money market deposits X_3. Use multiple regression to predict the number of housing starts using the following data which was collect for 15 periods in the table below:

Period	Housing Starts (x100)	Treasury Bill Rate (%)	Federal Funds Rate (%)	Money Market Rate (%)
1	108	12.9	10.3	9.5
2	121	14.3	11.4	12.2
3	129	12.6	8.3	8.4
4	95	12.9	12.3	10.8
5	124	14.8	10.3	9.4
6	95	13.7	9.4	8.7
7	112	14.2	11.6	11.3
8	72	12.3	8.7	9
9	117	13.0	10.4	9.1
10	110	13.3	9.0	8
11	68	11.1	10.6	9.4
12	91	12.8	10.1	9.2
13	112	12.8	9.6	10.4
14	98	13.8	11.1	9.9
15	90	13.2	10.0	10.2

Regression

Variables Entered/Removed [b]

Model	Variables Entered	Variables Removed	Method
1	MONEY, TREASURY, FED [a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: STARTS

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.652[a]	.425	.268	1.53762

a. Predictors: (Constant), MONEY, TREASURY, F

QUANTITATIVE TECHNIQUES II

MIDTERM EXAM

A C T U A L

NAME: _____ **PART 1: In-Class**

ANOVA [b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regressio	19.237	3	6.412	2.712	.096[a]
	Residual	26.007	11	2.364		
	Total	45.244	14			

a. Predictors: (Constant), MONEY, TREASURY, FED

b. Dependent Variable: STARTS

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-4.022	6.354		-.633	.540
	TREASURY	1.322	.486	.667	2.719	.020
	FED	-.440	.575	-.271	-.765	.460
	MONEY	.142	.581	.089	.244	.812

a. Dependent Variable: STARTS

a. Use the tables above to determine the regression equation $Y'=a+b_1X_1+b_2X_2+b_3X_3$

b. Predict the number of housing starts if $X_1=13$, $X_2=10$, and $X_3=10$.

c. State a reasonable hypothesis in statistical terms (notation)

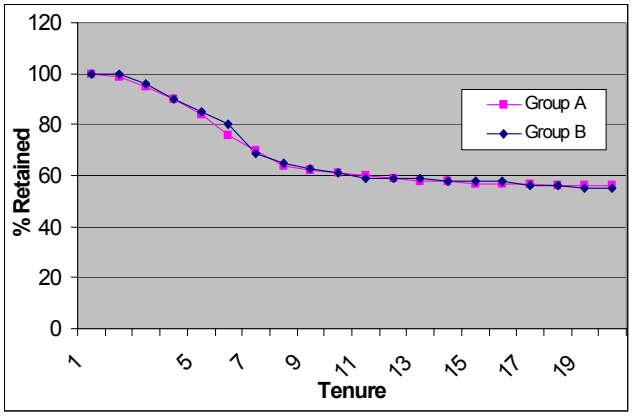
d. Which variable has the “best” p-value (Sig.)?

e. Which variable(s) should be removed, if any, from the regression model

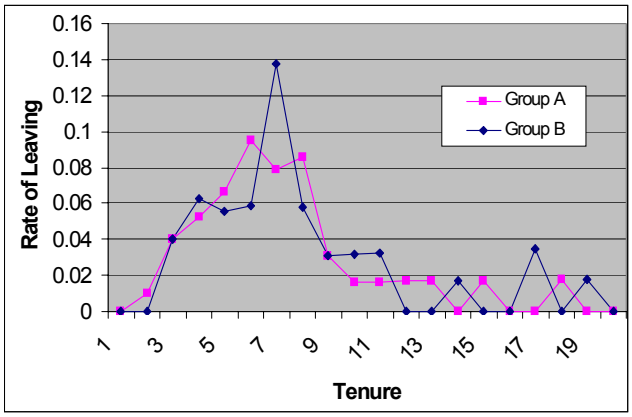
f. What is the value for r^2, (that is R^2)?

38. We are interested in better understanding our employee retention. What is the name of the quantitative technique which is appropriate for this type of analysis?

39. Does the following chart represent the Survival Function or the Hazard Rate?



40. Does the following chart represent the Survival Function or the Hazard Rate?



41. Which of the following are “flavors” of this method?

a. Cox Proportional Hazard Rate, Actuarial/Life Tables & Kaplan-Meier

b. Chi-Square, F-Test & Z-Test

c. Loglikelihood, Logistic & Neural Networks

d. Mann-Whitney, Fisher Exact & Factor Analysis

e. Time Series, Decision Analysis & Queuing Theory

QUANTITATIVE TECHNIQUES II

MIDTERM EXAM
A C T U A L

NAME: _____ **PART 1: In-Class**

42. We do not have the time or money to interview all 306 employees on the city payroll. Therefore, we decide to interview a simple random sample of 16, selected from the personnel list which the city clerk kindly provided. (For privacy purposes, the data has been transformed to hide the actual salaries.) Using a regression method from class and the table below, answer the following questions.

Respondent	Education	Income						
1	4	10,516						
2	6	8,212						
3	8	10,011						
4	8	12,405						
5	10	11,598						
6	10	15,336						
7	11	10,186						
8	12	12,444						
9	12	16,908						
10	12	18,347						
11	14	16,326						
12	15	12,772						
13	17	20,018						
14	18	16,526						
15	18	19,414						
16	20	18,822						

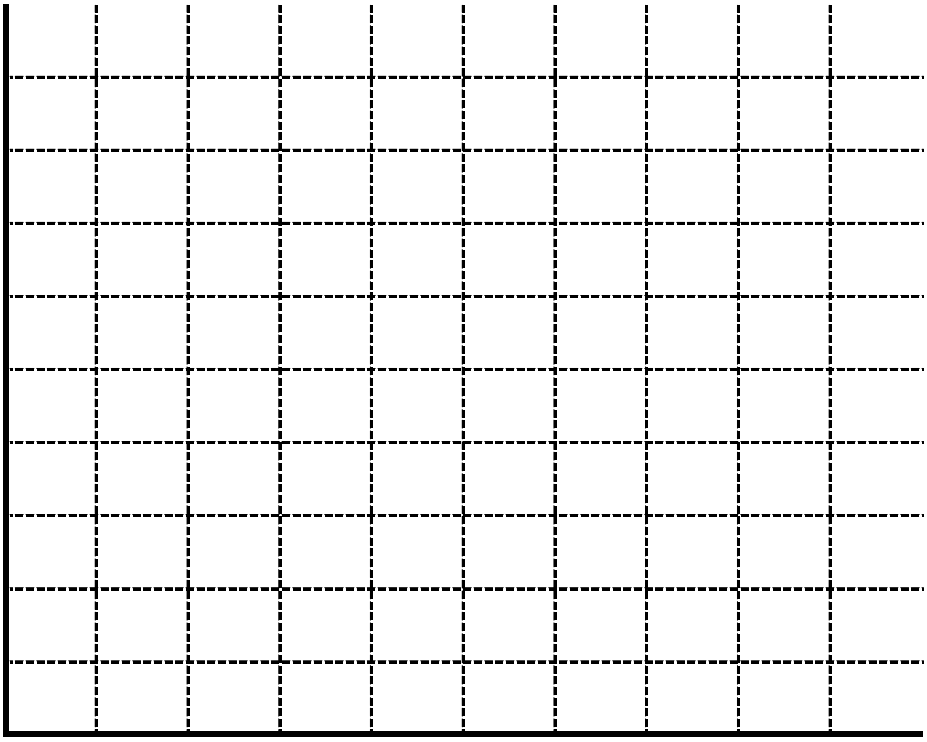
NAME: _____ **PART 1: In-Class**

a. Use Years to predict Income: So which variable should be independent and which should be dependent?

Independent:

Dependent:

b. Sketch the plot of the dependent against the independent.



c. In terms of the B_1 Coefficient, what could be your null hypothesis?

d. What is the B_0 Coefficient?

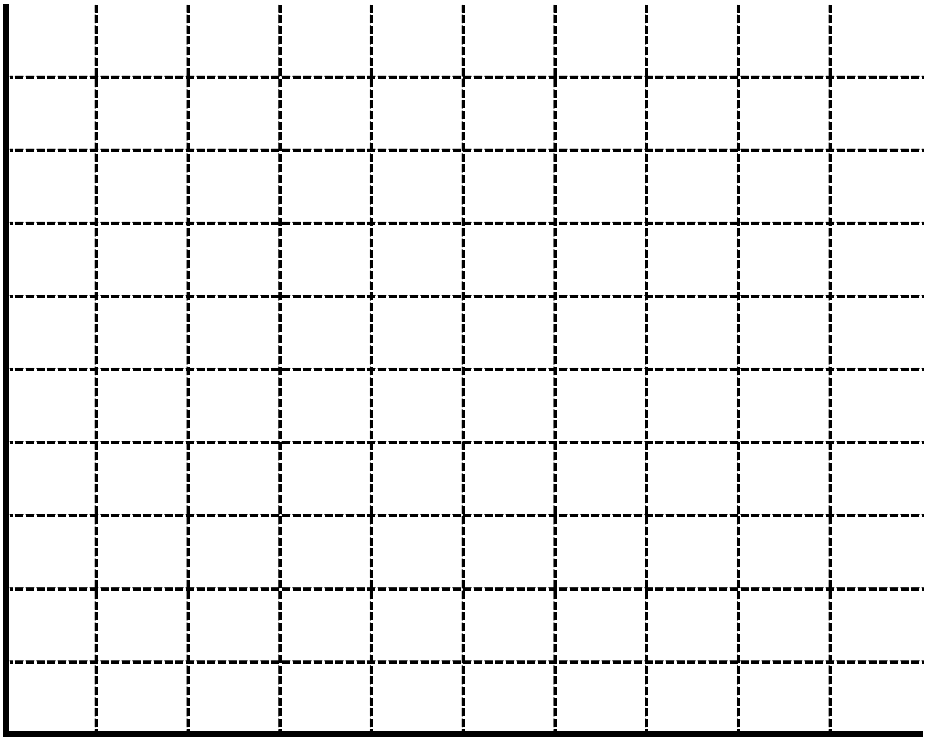
e. What is the B_1 Coefficient?

f. What is the Regression Equation?

g. Fill in the predicted values for each X observation.

h. Fill in the residuals for each X observation.

i. Sketch a plot of the residuals against the predicted values.



j. Do the error terms appear constant?