1[γ]. *List the 4 measurement scales discussed in class (not in the book) and say a few words about each.*

Nominal – numbers as names
Ordinal – numbers give order
Interval – distance between numbers are equal, no true zero
Ratio – distance between numbers are equal, has a true zero

2[γ]. *Fill in the statement below with the words marginal, joint and conditional in the correct positions.*

joint = **conditional** x **marginal**

3[γ]. *When X and Y have a conditional (non-independent) relationship, write down an expression for the joint probability, Pr(X=x, Y=y) using conditional and/or marginal probabilities. When the two variables are independent, write an expression for the joint probability using the two marginal probabilities.*

Conditional : Pr(X=x, Y=y) = Pr(Y=y|X=x) * Pr(X=x)

Independent: Pr(X=x, Y=y) = Pr(Y=y) * Pr(X=x)

4[γ]. *(2 points) Calculate the marginal distributions of X and Y below. Calculate E(X) and E(Y). Calculate the 8 conditional probabilities and enter them in the spaces provided. Demonstrate that X and Y below are independent or are not independent (so have a conditional relationship).*

E(X) = **1/3** E(Y) = **1/4**

		Y		
		0	**1**	
X	**0**	1/2	1/6	**2/3**
	1	1/4	1/12	**1/3**
		3/4	**1/4**	

P(X=0|Y=0) = 2/3 **P(X=0|Y=1) = 2/3**

P(X=1|Y=0) = 1/3 **P(X=1|Y=1) = 1/3**

P(Y=0|X=0) = 3/4 **P(Y=1|X=0) = 1/4**

P(Y=0|X=1) = 3/4 **P(Y=1|X=1) = 1/4**

Independent because 1) joints equal the products of the marginals, and 2) the conditionals equal the marginals

5[γ]. *Following the form in the textbook (beginning of appendix 2.1, derivation of equation 2.31), write the expression for var(aX+bY) that makes use of the definition of variance and the expectation operator. Then show the steps to transform the expression into form that we commonly use, involving variances and the covariance. (<u>The purpose of this question is to practice/ensure mastery of the definition of variance and of the use of the expectation operator, as well as of the end formula.)</u>*

To derive Equation (2.31), use the definition of the variance to write $\text{var}(a + bY) = E\{[a + bY - E(a + bY)]^2\} = E\{[b(Y - \mu_Y)]^2\} = b^2E[(Y - \mu_Y)^2] = b^2\sigma_Y^2$.

To derive Equation (2.32), use the definition of the variance to write

$$\begin{aligned}
\text{var}(aX + bY) &= E\{[(aX + bY) - (a\mu_X + b\mu_Y)]^2\} \\
&= E\{[a(X - \mu_X) + b(Y - \mu_Y)]^2\} \\
&= E[a^2(X - \mu_X)^2] + 2E[ab(X - \mu_X)(Y - \mu_Y)] \\
&\quad + E[b^2(Y - \mu_Y)^2] \\
&= a^2\text{var}(X) + 2ab\,\text{cov}(X, Y) + b^2\,\text{var}(Y) \\
&= a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2, \qquad (2.50)
\end{aligned}$$

6[γ]. *Suppose Y_i, i=1, 2, ,…,n, are i.i.d. random variables, each distributed N(4,36). Compute Pr(3 $\leq \bar{Y} \leq$ 5) for a sample size of 64.*

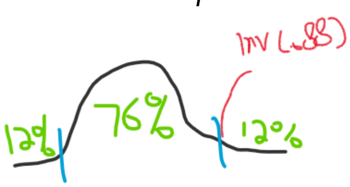
Pr[(Y-4)/√(36/64) ≤ (3-4)/√(36/64)] = Pr[Z ≤ -1.33] = 0.0918

Pr[(Y-4)/√(36/64) ≤ (5-4)/√(36/64)] = Pr[Z ≤ 1.33] = 0.9082

0.9082 – 0.0918 = 0.8164

7[γ]. *Calculate what the cutoff points would be for the standardized Z variable in terms of standard deviations for capturing 76% of the probability, in the middle of the normal distribution. In other words, consider a two-tailed approach which leaves 12% of the probability in each tail and find the values of R that satisfy Pr(-R ≤ Z ≤ R) = 0.76.*

+/- 1.175



0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015

3[Ξ]. If Y is distributed N(1,4), find Pr(Y>0).

8[γ]. *Suppose Y_i, i=1, 2, ,…,n, are i.i.d. random variables, each distributed N(1,16). Compute Pr(1 ≤ Y ≤ 3) for a sample size of 36.*

Pr[(Y-1)/√16 ≤ (1-1)/√16] = Pr[Z ≤ 0] = 0.5000

Pr[(Y-1)/√16 ≤ (3-1)/√16] = Pr[Z ≤ 0.5] = 0.6915

0.6915 – 0.5 = 0.1915

9, What is the difference between an estimator and an estimate?

A. Both an estimator and an estimate are functions of a sample of data to be drawn randomly from a population.
B. An estimate is a function of a sample of data to be drawn randomly from a population whereas an estimator is the numerical value of the estimator when it is actually computed using data from a specific sample.
C. Both an estimator and an estimate are numerical values computed using data from a specific sample.
D. An estimator is a function of a sample of data to be drawn randomly from a population whereas an estimate is the numerical value of the estimator when it is actually computed using data from a specific sample.

10.(1.5 points) In a survey of 10,000 likely voters, 8,000 responded that they would vote for the incumbent and 2,000 responded that they would vote for the challenger. Let p denote the fraction of all likely voters who preferred the incumbent at the time of the survey, and let p-hat be the fraction of survey respondents who prefer the incumbent.

a. Use the survey results to estimate p. **p-hat = 0.8**

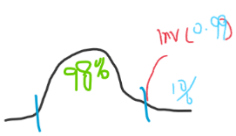
b. Use the estimator of the variance to calculate the standard error of your estimate. If you cannot calculate a square root, simply leave it under the square root symbol. Neatness counts.

Var(p-hat) = (0.8 * 0.2) / 10,000

SE(p-hat) = 0.4 / 100 = 0.004

c. *Construct a 98% (not 95%, not 99%) confidence interval for p.*

0.8 +/- 2.33 * 0.004



1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936

11. (3 points) Below is a comparison of test scores in a recent Eco 4000 class. The goal is to assess if the difference between the average test score for women is statistically significantly different from the average test score for men. Create a 95% confidence interval for the true difference between the performance of women and men. Use large-sample techniques even though there are small numbers in the problem for the sake of convenience. Do not use 1.96 as the critical value, but instead use 2.00, again for convenience. After creating your confidence interval, give your conclusion as to whether or not the difference is statistically significant at the 95% confidence level.

	Average Scores	Standard Deviation	Number of Participants n
Women	23	9	64
Men	22.25	4	16

The expression for the standard error of the difference is $\sqrt{(9^2/64 + 4^2/16)} = \sqrt{(145/64)}$. If we "round" this to $\sqrt{(144/64)}$, then the square root is 12/8 or 3/2.

The confidence interval is 0.75 +/- 2 * $\sqrt{(145/64)}$. That is enough for full credit.

This rounds to 0.75 +/- 2 * 3/2 = -2.25 to 3.75.

12. *Changing the units of measure—that is, measuring test scores in 100s, will do all of the following except for changing the:*

A. numerical value of the intercept.

B. interpretation of the effect that a change in X has on the change in Y.

C. residuals.

D. numerical value of the slope estimate.

13. *(1.5 points) Suppose that a researcher, using data on class size(CS) and average test scores from 625 third-grade classes, estimates the OLS regression*

TestScores-hat = 300 – 6 X CS, R-squared = 0.08, SER = 11.5

a. A classroom has 21 students. What is the regression's prediction for the classroom's average test score?

TestScores-hat = 300 – 6 * 21 = 174

b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in average test score?

Predicted change in test scores = (– 6) * 4 = -24

c. The sample average class size across the 625 classrooms is 20. What is the sample average test score across the 625 classrooms?

TestScores-hat = 300 – 6 * 20 = 180

14. (6 points) Use the values below. Complete the regression output on the following page.

X	Y
-3	1
-1	-3
-1	1
0	-3
0	-1

	X	Y	X-mean	Y-mean	X-mean squared	X-mean * Y-mean	Y-hat	U-hat	Y-mean squared	y-hat - mean squared	U-hat squared
	-3	1	-2	2	4	-4	1	0	4	4	0
	-1	-3	0	-2	0	0	-1	-2	4	0	4
	-1	1	0	2	0	0	-1	2	4	0	4
	0	-3	1	-2	1	-2	-2	-1	4	1	1
	0	-1	1	0	1	0	-2	1	0	1	1
sum	-5	-5	0	0	6	-6	-5	0	16	6	10
* 1 /n	-1	-1					-1		TSS	ESS	SSR
* 1/(n-1)											
* 1/(n-2)											3.333333

B1-hat =	-1	R^2=	
B0-hat =	-2	SER=	=3.33^0.5

Multiple R
R Square
Adjusted R Square
Standard Error
Observations

ANOVA

	SS	*MS*
Regression	6	6

Residual	10	3.333333333
Total	16	

	Standard Error	*t Stat*
Intercept	1.105541597	-1.809068067
X	0.745355992	-1.341640786

Regression Statistics	
Multiple R	**√(ESS/TSS)**
R Square	**ESS/TSS**
Adjusted R Square	XXXXXXX
Standard Error	**SER=√(SSR/(n-2))**
Observations	**n**

$$ESS = \sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2$$

$$TSS = \sum_{i=1}^{n}(Y_i - \overline{Y})^2.$$

$$R^2 = \frac{ESS}{TSS}.$$

$$R^2 = 1 - \frac{SSR}{TSS}.$$

ANOVA

	df	SS	MS
Regression	**1**	**ESS**	**ESS/1**
Residual	**n-2**	**SSR**	**SSR/(n-2)**
Total	**n-1**	**TSS**	

$$SER = s_{\hat{u}} = \sqrt{s_{\hat{u}}^2}, \text{ where } s_{\hat{u}}^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{SSR}{n-2},$$

	Coeffic.
Intercept	**B0-hat**
X Variable 1	**B1-hat**

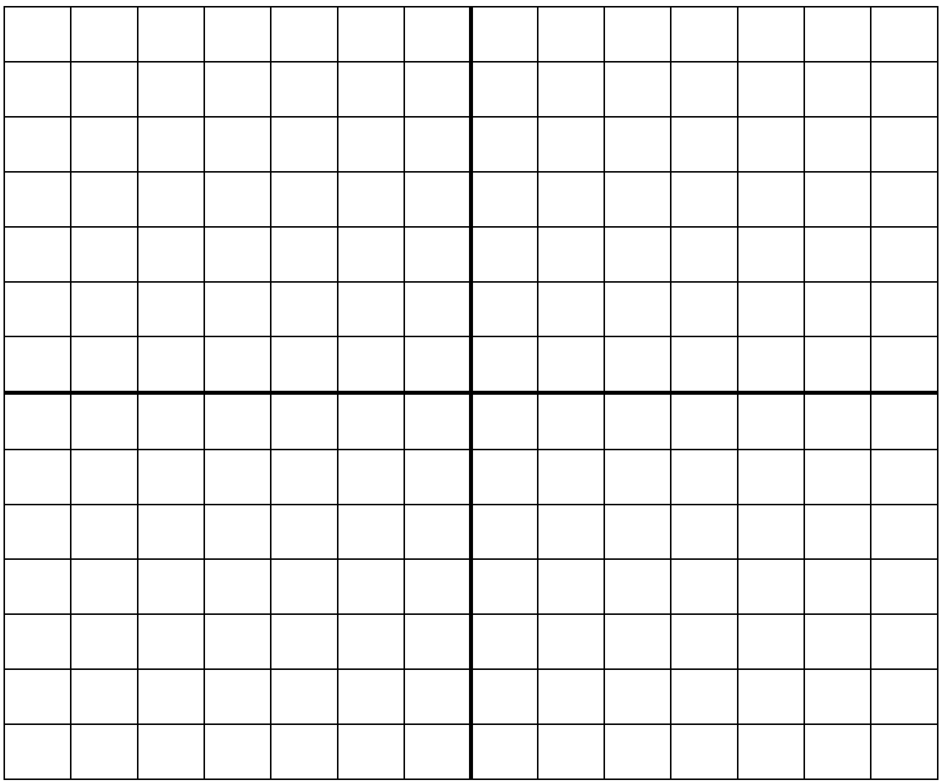
$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}X_iY_i - \overline{X}\,\overline{Y}}{\frac{1}{n}\sum_{i=1}^{n}X_i^2 - (\overline{X})^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X}.$$

RESIDUAL OUTPUT

Observation	*Predicted Y*	*Residuals*
1	**Y-hat**	**U-hat**
2	**Y-hat**	**U-hat**
3	**Y-hat**	**U-hat**
4	**Y-hat**	**U-hat**
5	**Y-hat**	**U-hat**

15. (2 points) On the Cartesian plane below, graph the regression line using the values above. Also, plot the data points as solid points and plot the predicted Y's (the Y-hats) as open circles.



16. [Ξ]. *(2 points) In the regression above do we reject the null hypothesis for* $\hat{\beta}_0$*? For* $\hat{\beta}_1$*?*

[By analyzing the confidence interval of $\widehat{\beta_0}$, if the range contain positive value and negative values together, we have no sufficient information to reject the null hypothesis for $\widehat{\beta_0}$; if not, we should reject it. For, $\hat{\beta}_1$ is the same.]

16. *In landscape orientation recreate the grid of formulas for mean/expected value, variance, covariance, and correlation. Arrange the formulas in four columns: theory (expected value format), when you have the probabilities, when you have all the observations in the population,*

and when you have a sample of the population.

	POPULATION			SAMPLE
Name	Theory	Have Probabilities	Have All Observations	Have Sample Data
Mean	$\mu_Y = E(Y)$	$\mu_Y = \sum_{i=1}^{k} y_i p_i$	$\mu_Y = \frac{1}{n}\sum_{i=1}^{n} y_i$	$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i$
Variance	$\sigma_Y^2 = E[(Y-\mu_Y)^2]$	$\sigma_Y^2 = \sum_{i=1}^{k} (y_i-\mu_Y)^2 p_i$	$\sigma_Y^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i-\mu_Y)^2$	$s_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i-\bar{Y})^2$
Covariance:				
	$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$	$\sigma_{XY} = \sum_{i=1}^{k}\sum_{j=1}^{l} (x_j-\mu_X)(y_i-\mu_Y)\Pr(X=x_j, Y=y_i)$	$\sigma_{XY} = \frac{1}{n}\sum_{i=1}^{n} (x_i-\mu_X)(y_i-\mu_Y)$	$s_{XY} = \frac{1}{n-1}\sum_{i=1}^{n} (x_i-\bar{X})(y_i-\bar{Y})$
Correlation	$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$	$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$	$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$	$r_{XY} = \frac{s_{XY}}{s_X s_Y}$

17. *(2 points) List the 3 base assumptions for large-sample regression analysis. List at least one assumption that we must add for small-sample regression analysis.*

- $E(u_i|X_i)=0$
- (X_i, Y_i) are iid
- large outliers are unlikely

- errors are distributed normally around the line

18. *$E(u_i|X_i)=0$ says that:*
A. dividing the error by the explanatory variable results in a zero (on average).
B. the sample regression function residuals are unrelated to the explanatory variable.
C. the conditional distribution of the error given the explanatory variable has a zero mean.
D. the sample mean of the Xs is much larger than the sample mean of the errors

[Ξ].

The expected value of a discrete random variable

- A. is computed as a weighted average of the possible outcome of that random variable, where the weights are the probabilities of that outcome.
- B. equals the population median.
- C. is the outcome that is most likely to occur.
- D. can be found by determining the 50% value in the c.d.f.

[Ξ].

The *skewness* of the distribution of a random variable Y is defined as follows:

- A. $\dfrac{E\left[\left(Y^3 - \mu_Y\right)\right]}{\sigma_Y^2}$
- B. $\dfrac{E\left[\left(Y - \mu_Y\right)^3\right]}{\sigma_Y^3}$
- C. $E\left[\left(Y - \mu_Y\right)^3\right]$
- D. $\dfrac{E\left[\left(Y^3 - \mu_Y^3\right)\right]}{\sigma_Y^3}$

[Ξ].

An estimator is

- A. an estimate.
- B. a formula that gives an efficient guess of the true population value.
- C. a nonrandom number.
- D. a random variable.

[Ξ].

An estimator $\hat{\mu}_Y$ of the population value μ_Y is unbiased if

- A. $\overline{Y}$ has the smallest variance of all estimators.
- B. $\hat{\mu}_Y = \mu_Y$.
- C. $\overline{Y} \xrightarrow{p} \mu_Y$.
- D. $E(\hat{\mu}_Y) = \mu_Y$.

[Ξ].

To obtain the slope estimator using the least squares principle, you divide the

- A. sample covariance of X and Y by the sample variance of X.
- B. sample variance of X by the sample covariance of X and Y.
- C. sample covariance of X and Y by the sample variance of Y.
- D. sample variance of X by the sample variance of Y.

[Ξ].

(Requires Appendix) The sample regression line estimated by OLS

- A. will always have a slope smaller than the intercept.
- B. will always run through the point $(\bar{X}, \bar{Y})$.
- C. is exactly the same as the population regression line.
- D. cannot have a slope of zero.

[Ξ].

If the absolute value of your calculated $t-$statistic exceeds the critical value from the standard normal distribution, you can

- A. reject the assumption that the error terms are homoskedastic.
- B. reject the null hypothesis.
- C. conclude that most of the actual values are very close to the regression line.
- D. safely assume that your regression results are significant.

[Ξ].

Consider the following regression line: $\widehat{TestScore} = 698.9 - 2.28 \times STR$. You are told that the $t-$statistic on the slope coefficient is 4.38. What is the standard error of the slope coefficient?

- A. 1.96
- B. -1.96
- C. 0.52
- D. 4.38

[**A**]

1. Analyzing the behavior of unemployment rates across U.S. states in March of 2006 is an example of using

A. time series data.

B. cross–sectional data.

C. panel data.

D. experimental data.

[**A**]

10. For a normal distribution, the skewness and kurtosis measures are as follows:

A. 0 and 3

B. 1 and 2

C. 1.96 and 4

D. 0 and 0

Extra Credit: Write the values below for letters 13-18of the Greek alphabet.

Ν	ν	Nu	n
Ξ	ξ	Xi	x
Ο	o	Omicron	o
Π	π	Pi	p
Ρ	ρ	Rho	r
Σ	σ	Sigma	s