The University of Kansas

Department of Economics

Midterm **Mar/08/2019**
Econ 526 - Introduction to Econometrics **Instructor: Caio Vigo Pereira**

Name:

SECTION A - MULTIPLE CHOICE

$\boxed{4\%}$ 1. Among the measures of central tendency of a distribution we have:

A. $Var(X)$
B. $Median(X)$
C. $sd(X)$
D. $Cov(X, Y)$

$\boxed{4\%}$ 2. Let X be a discrete random variable. What is the following term?

$$\sum_{j=1}^{m} x_j f_{X|Y}(x_j|y)$$

A. the conditional probability of X given Y
B. the joint distribution of X given Y
C. the joint distribution of Y given X
D. the conditional expectation of X given Y

$\boxed{4\%}$ 3. If X is a random variable such that $E(X) = \mu$ and $Var(X) = \sigma^2$. What is $S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n}(X_i - \bar{X})^2$?

A. an estimator of μ
B. an estimate of σ
C. an estimator of σ^2
D. an estimate of $Cov(\mu, \sigma^2)$

$\boxed{4\%}$ 4. Consider the following simple linear regression model: $y = \beta_0 + \beta_1 x + u$. What is the OLS estimator for β_1?

A. $\bar{y} - \hat{\beta}_1 \bar{x}$

B. $\bar{y} - \beta_1 \bar{x}$

C. $\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})}$

D. $\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

[This statement refers to the following three questions]
Let X_1, X_2, and X_3 be i.i.d. random variables from a population with mean μ and variance σ^2. Consider the following estimators for the mean μ:

$$W = \sum_{i=1}^{3} \frac{1}{i} X_i$$

$$H = \frac{1}{3} \sum_{i=1}^{3} X_i$$

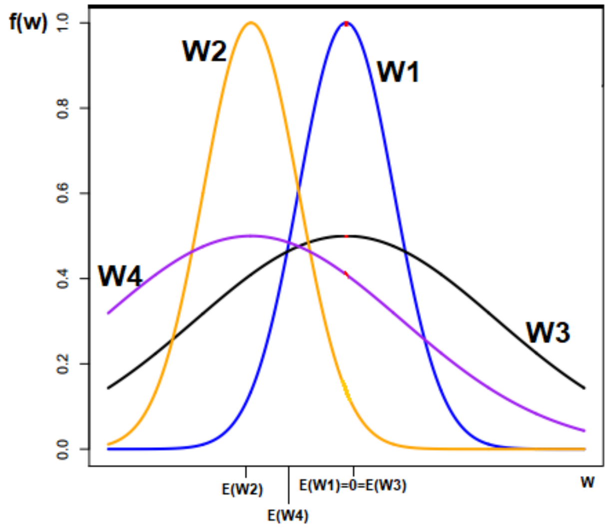
$\boxed{4\%}$ 5. What is the $E(W)$ of the estimator?
A. $\frac{11}{6}\mu$
B. μ
C. 3μ
D. $\frac{1}{3}\mu$

$\boxed{4\%}$ 6. What is the variance of W?
A. $\frac{13}{11}\sigma^2$
B. σ^2
C. $3\sigma^2$
D. $\frac{1}{3}\sigma^2$

$\boxed{4\%}$ 7. What can you tell about the bias of the estimators W and H?
A. W and H are both **unbiased** estimators for the mean μ
B. W is a **biased** and H is an **unbiased** estimator for the mean μ
C. W is an **unbiased** and H is a **biased** estimator for the mean μ
D. W and H are both **biased** estimators for the mean μ

$\boxed{4\%}$ 8. One way to compare estimators that are not necessarily unbiased is to compute the:
A. Variance of the estimator
B. Covariance of the estimator
C. Mean Squared Error of the estimator
D. Standard Deviation of the estimator

Figure 1: The p.d.f. of 4 Estimators of a Population Parameter



4% 9. Figure 1 shows the p.d.f. of 4 estimators of the population parameter θ. Knowing that $\theta = 0$, which estimator(s) for the parameter θ is(are) **biased**?
A. $W1$
B. $W1$ and $W3$
C. $W2$ and $W4$
D. $W1$ and $W2$

4% 10. Refer to Figure 1 again. Knowing that $\theta = 0$, which estimator(s) is(are) relatively efficient?
A. $W1$ is efficient relative to $W3$
B. $W1$ and $W3$ are efficient relative to $W4$
C. $W3$ is efficient relative to $W4$
D. $W3$ is efficient relative to $W1$

SECTION B - TRUE OR FALSE

3% 1. Let X and Y be two random variables. Knowing that $Cov(X,Y) \neq 0$, then

$$Var(X+Y) = Var(X) + Var(Y)$$

◯ True ◯ False

3% 2. You have a cross-sectional dataset with an independent variable X and a dependent variable Y. You find a positive correlation between X and Y. Then you can conclude that X causes Y.
◯ True ◯ False

3% 3. During one month, you collected the daily returns of Apple Inc. stock (AAPL). Therefore, this is a time series data.
◯ True ◯ False

3% 4. Consider the following simple linear regression model: $y = \beta_0 + \beta_1 x + u$. The essential assumption to derive the estimators of β_0 and β_1 through the Method of Moments is $E(u|X) = 0$.
◯ True ◯ False

3% 5. Consider the following simple linear regression model: $y = \beta_0 + \beta_1 x + u$. When we derive the estimators for β_0 and β_1 we get 3 First Order Conditions.
◯ True ◯ False

SECTION C - SHORT ANSWER

1. Assume that an online store has a customer rating system for the products that they sell on their website. After a customer buys any product online, s/he is invited to send his/her review about the product acquired. The customer can choose any integer number from 0 to 5 to give his/her overall evaluation of the product. This is done choosing the number of stars in the online store website. With these customer evaluations, the online store compute the following formula:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

where X_i is the number of stars given by the customer i. Then, the online store shows this customer rating score (i.e., $\bar{X}$) in their website for all products that they sell. This is done showing the number of stars, where the number of stars is equal to $\bar{X}$.

Now consider that for two products having the same features you see the following customer reviews:

(notice that the number to the right of the customer rating score $\bar{X}$ represents n, i.e., the number of customer reviews). Assuming that the customers review X_i, for $i = 1, 2, \ldots, n$ represents a random sample from the population, answer the questions below:

4% (a) If you are a customer comparing both products (that have the same features, i.e, you can consider both of them roughly the same) and you want to base your decision to buy either product A or product B using the customer reviews rating (i.e., you are interested to know the true mean), based on your statistics knowledge, which one would you choose? Give the name of the theorem that you are using in your answer. (Disregard any possible price difference between both products) **[3 lines answer]**

6% (b) Explain in your own words this theorem. **[4 lines answer]**

2. Suppose you want to study the effects of the number of students per classroom in algebra courses and students' performance in algebra courses for high schools in Kansas. You collected a random sample and now you have data for the above two variables. You called them as *number_students* (which refers to the number os students per classroom in algebra courses), and *students_performance* (which refers to the students' performance in algebra courses - measured as their final grade in a scale from 0 to 4). Therefore, you want to know how *number_students* explains *students_performance*.

2% (a) Using the variables names, write the simple linear regression model. **[1 line answer maximum]**

5% (b) Knowing that the OLS estimate for the intercept is 3.5, and for the slope is -0.01, write the estimated OLS regression line (or SRF) using the variables names. **[1 line answer - don't exceed it]**

3% (c) What is the predicted value for whichever is your dependent variable for a classroom with 20 students? **[1 line answer - don't exceed it]**

3. *(This question refers to Table 1).* In this table you have a random sample with 50 data points from a population, i.e., your observations are $\{(x_i, y_i) : i = 1, 2, \ldots, n\}$, where $n = 50$.

   Considering the following econometric model $y = \beta_0 + \beta_1 x + u$, answer the questions below.

2% (a) Find the value of $\boxed{\mathbf{A}}$ (located at the bottom - last row - of the table)? **[1 line answer maximum]**

3% (b) One of the columns shows $(y_i - \hat{y}_i)$. What is $(y_i - \hat{y}_i)$ and what is the value of $\boxed{\mathbf{B}}$? **[1 line answer maximum]**

5% (c) What is the OLS estimate of β_1 ?

5% (d) What is the OLS estimate of β_0 ?

5% (e) What SST stands for? What SSE stands for? What SSR stands for? What are their formulas? Your answer should be in the following format: "SST = complete name = formula". **[3 lines answer maximum]**

5% (f) What is the R^2 of the regression? Interpret the result. **[2 lines answer maximum]**

Table 1:

Obs. #	y_i	x_i	$(y_i - \bar{y})$	$(x_i - \bar{x})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$\hat{y}_i$	$(y_i - \hat{y}_i)$	$(\hat{y}_i - \bar{y})^2$	$(y_i - \hat{y}_i)^2$
1	140	80	45	42	2025	1764	1890	152.94	-12.94	3356.93	167.42
2	80	20	-15	-18	225	324	270	70.17	9.83	616.58	96.65
3	107	50	12	12	144	144	144	111.55	-4.55	274.03	20.74
4	68	30	-27	-8	729	64	216	83.96	-15.96	121.79	254.85
5	52	10	-43	-28	1849	784	1204	56.37	-4.37	1491.97	19.13
6	90	40	-5	2	25	4	-10	97.76	-7.76	7.61	60.20
7	60	10	-35	-28	1225	784	980	56.37	3.63	1491.97	13.15
8	101	40	6	2	36	4	12	97.76	3.24	7.61	10.50
9	45	10	-50	-28	2500	784	1400	56.37	-11.37	1491.97	129.37
10	110	30	15	-8	225	64	-120	83.96	26.04	121.79	677.87
11	50	10	-45	-28	2025	784	1260	56.37	-6.37	1491.97	40.63
12	80	30	-15	-8	225	64	120	83.96	-3.96	121.79	15.71
13	150	70	55	32	3025	1024	1760	139.14	10.86	1948.69	117.85
14	50	20	-45	-18	2025	324	810	70.17	-20.17	616.58	406.79
15	77	10	-18	-28	324	784	504	56.37	20.63	1491.97	425.43
16	132	70	37	32	1369	1024	1184	139.14	-7.14	1948.69	51.04
17	139	70	44	32	1936	1024	1408	139.14	-0.14	1948.69	0.02
18	114	60	19	22	361	484	418	125.35	-11.35	921.06	128.80
19	34	0	-61	-38	3721	1444	2318	42.58	-8.58	2747.96	73.60
20	107	40	12	2	144	4	24	97.76	9.24	7.61	85.40
21	94	40	-1	2	1	4	-2	97.76	-3.76	7.61	14.13
22	100	40	5	2	25	4	10	97.76	2.24	7.61	5.02
23	40	0	-55	-38	3025	1444	2090	42.58	-2.58	2747.96	6.65
24	70	20	-25	-18	625	324	450	70.17	-0.17	616.58	0.03
25	180	90	85	52	7225	2704	4420	166.73	13.27	5145.77	175.99
26	160	80	65	42	4225	1764	2730	152.94	7.06	3356.93	49.86
27	70	0	-25	-38	625	1444	950	42.58	27.42	2747.96	751.91
28	127	40	32	2	1024	4	64	97.76	29.24	7.61	855.04
29	108	60	13	22	169	484	286	125.35	-17.35	921.06	300.99
30	105	50	10	12	100	144	120	111.55	-6.55	274.03	42.95
31	50	10	-45	-28	2025	784	1260	56.37	-6.37	1491.97	40.63
32	137	70	42	32	1764	1024	1344	139.14	-2.14	1948.69	4.60
33	140	60	45	22	2025	484	990	125.35	14.65	921.06	214.65
34	35	0	-60	-38	3600	1444	2280	42.58	-7.58	2747.96	57.44
35	56	0	-39	-38	1521	1444	1482	42.58	13.42	2747.96	180.12
36	85	30	-10	-8	100	64	80	83.96	1.04	121.79	1.07
37	153	90	58	52	3364	2704	3016	166.73	-13.73	5145.77	188.62
38	46	10	-49	-28	2401	784	1372	56.37	-10.37	1491.97	107.62
39	77	20	-18	-18	324	324	324	70.17	6.83	616.58	46.66
40	160	90	65	52	4225	2704	3380	166.73	-6.73	5145.77	45.35
41	33	20	-62	-18	3844	324	1116	70.17	-37.17	616.58	1381.53
42	179	90	84	52	7056	2704	4368	166.73	12.27	5145.77	150.45
43	79	20	-16	-18	256	324	288	70.17	8.83	616.58	77.99
44	154	70	59	32	3481	1024	1888	139.14	14.86	1948.69	220.70
45	54	10	-41	-28	1681	784	1148	56.37	-2.37	1491.97	5.64
46	133	60	38	22	1444	484	836	125.35	7.65	921.06	58.54
47	96	40	1	2	1	4	2	97.76	-1.76	7.61	3.09
48	127	70	32	32	1024	1024	1024	139.14	-12.14	1948.69	147.48
49	65	10	-30	-28	900	784	840	56.37	8.63	1491.97	74.41
50	51	10	-44	-28	1936	784	1232	56.37	-5.37	1491.97	28.88
Sum	**4,750**	**1,900**	$\boxed{\text{A}}$	$\boxed{\text{not provided}}$	**84,154**	**40,000**	**55,180**	**4,750**	$\boxed{\text{B}}$	$\boxed{\text{not provided}}$	**8,033**