The University of Kansas

Department of Economics

Final **July/25/2019**
Econ 526 - Introduction to Econometrics **Instructor: Caio Vigo Pereira**

Name:

SECTION A - MULTIPLE CHOICE

4% 1. Let X, Y and Z be three random variables. Knowing that $\text{Corr}(X, Y) = 1$ and $\text{Corr}(X, Z) = -1$, if X falls, what can you tell about the direction of the change of Y and Z? based on Quiz 1, A-3

A. Y will raise and Z will raise
B. Y will raise and Z will drop
C. Y will drop and Z will raise
D. Y will drop and Z will drop

4% 2. Consider a multiple linear regression model such as: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$. Under the Gauss-Markov assumptions, what is the distribution of the error term u?

A. $u \sim N(0, 1)$
B. $u \sim N(0, \sigma^2)$
C. $u \sim t_{df}$, where $df = n - k - 1$
D. Gauss-Markov assumptions don't restrict the distribution of the error term u

4% 3. Consider any multiple linear regression. It is known that under the Gauss-Markov assumptions, the OLS estimators are BLUE. What "B" refers to?

A. That the OLS estimators have the smallest variance among the unbiased estimators
B. That $E(\hat{\beta}_j^{OLS}) = \beta_j$ for any $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$
C. That the OLS estimators have the smallest variance among all possible estimators
D. That the OLS estimators are consistent

4% 4. Assume that the **Classical Linear Model (CLM)** assumptions hold. What is the distribution of $\dfrac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)}$?

based on Quiz 7, A-6

A. t_{df}, where $df = n - k - 1$
B. $F_{(q,n-k-1)}$
C. $N(0, k^2)$
D. None of the above

4% 5. The ________ is used to compare across models that have different numbers of explanatory variables but where one is **not** a special case of the other (i.e., **nonnested models**).

A. R^2
B. *t test*
C. *Adjusted* R^2
D. *F test*

2.5% 6. EXTRA POINTS Among the statements below, which one is **NOT** under the *Classical Linear Model* assumptions?
A. the error term u is normally distributed
B. the error term u is independent of the explanatory variables
C. the error term u has mean 0
D. the variance of the error term u is a function of the explanatory variables

2.5% 7. EXTRA POINTS Which of the following can cause the usual OLS t statistics to be invalid (that is, not to have t distributions under the null hypothesis)?
A. Heteroskedasticity
B. Multicollinearity
C. Homoskedasticity
D. Exogenous variables

SECTION B - TRUE OR FALSE

3% 1. We say that an estimator is unbiased if it has the smallest variance among all other estimators. based on Quiz 2, B-4
◯ True ◯ False

3% 2. Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with mean μ, and variance σ^2. Consider the following estimator: $W = (Y_1 + \frac{Y_2}{2} + \frac{Y_n}{2})/2$. Then, W is a **biased** estimator of μ. based on Quiz 2, B-6
◯ True ◯ False

3% 3. Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with mean μ. The *Law of Large Numbers (LLN)* states that $\bar{Y}$ is a consistent estimator of μ. based on Quiz 2, B-3
◯ True ◯ False

3% 4. The following regression model: $log(y) = \beta_0 + \beta_1 log(x_1) + u$ is also known as constant elasticity model. based on Quiz 4, B-4
◯ True ◯ False

3% 5. Consider the following models:

$$\text{Model 1:} \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

$$\text{Model 2:} \quad y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$

Then, $SSR_{model1} \geq SSR_{model2}$ and $SST_{model1} = SST_{model2}$. based on Quiz 5, B-3
◯ True ◯ False

3% 6. Exogenous explanatory variables is not a necessary assumption in order to the OLS estimator to be unbiased, however the assumption $E(u|x_1, \ldots, x_k) = 0$ is necessary.
◯ True ◯ False

3% 7. Multicollinearity violates the Gauss-Markov assumptions, and therefore the OLS estimators are not BLUE.
◯ True ◯ False

3% 8. Given the t statistic, the p-value provides the largest significance level in order to reject the null hypothesis.
◯ True ◯ False

3% 9. Consider any multiple linear regression. Knowing that you can reject H_0 for a specific parameter at 1% significance level, then you should be able to reject the H_0 at 2% significance level, but not necessarily at 0.1% significance level.
based on Quiz 7, B-3
◯ True ◯ False

3% 10. Consider the following multiple linear regression model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$$

Assume that the 95% confidence interval for β_1 is $[-0.254, 1.723]$.
Therefore, $\hat{\beta}_1$ is statistically different from 0 at 5% significance level.
◯ True ◯ False

SECTION C - SHORT ANSWER

Consider a data set containing a random sample with salary information and career statistics for 269 players in the National Basketball Association (NBA). The dataset consists of the following variables (variable's name and description):

wage	annual salary, thousands $
exper	years as professional player
age	age in years
coll	years played in college
games	average games per year
minutes	minutes per season
guard	=1 if guard
forward	=1 if forward
center	=1 if center
points	points per game
rebounds	rebounds per game
assists	assists per game
draft	draft number
allstar	=1 if ever all star
avgmin	minutes per game

1. (This question refers to **Regression (A)**) Consider the following regression (R output) [**Notice that the significance level "stars" - *, **, *** - were suppressed in this output**]:

REGRESSION (A)

	Dependent variable: log(wage)
exper	0.1289
	(0.0354)
age	-0.0585
	(0.0350)
coll	-0.0556
	(0.0519)
allstar	-0.0038
	(0.1376)
avgmin	0.0507
	(0.0046)
Constant	6.8871
	(0.8442)
Observations	269
R2	0.4913
Adjusted R2	0.4816
Residual Std. Error	0.6346 (df = 263)
F Statistic	50.7930 (df = 5; 263)

4% (a) State the null hypothesis that the number of years played in college has no *ceteris paribus* effect on a NBA's player salary (two-sided). State the alternative hypothesis that there is an effect? [Two lines answer]

4% (b) Test the hypothesis stated above at the 1% significance level. Find the critical value. [Two lines answer]

4% (c) Do you reject the null hypothesis? Explain the statistical significance of your test at 1% significance level. [Two lines answer]

2% (d) Would you include *coll* in a final model explaining NBA players salary in terms of years played in college? Why? Explain. [One line answer]

4% (e) Find the 99% confidence interval for β_{exper}. [One line answer]

4% (f) Is the variable *exper* statistically significant at 1% significance level? [One line answer]

2. (This question refers to **Regression (B)**) Consider the following (additional) regression:

REGRESSION (B)

```
==============================================
                        Dependent variable:
                    ---------------------------
                              log(wage)
----------------------------------------------
exper                         0.1257***
                              (0.0352)

age                           -0.0552
                              (0.0348)

coll                          -0.0370
                              (0.0518)

allstar                       -0.2132
                              (0.1565)

avgmin                        0.0299***
                              (0.0094)

points                        0.0450***
                              (0.0164)

games                         -0.0001
                              (0.0025)

Constant                      6.8158***
                              (0.8361)

----------------------------------------------
Observations                    269
R2                             0.5056
Adjusted R2                    0.4923
Residual Std. Error      0.6280 (df = 261)
F Statistic          38.1252*** (df = 7; 261)
==============================================
Note:              *p<0.1; **p<0.05; ***p<0.01
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.8158193  0.8360829   8.152  1.5e-14 ***
exper        0.1256580  0.0351871   3.571 0.000423 ***
age         -0.0552317  0.0347665  -1.589 0.113350
coll        -0.0369725  0.0518374  -0.713 0.476335
allstar     -0.2131675  0.1565311  -1.362 0.174428
avgmin       0.0299440  0.0093665   3.197 0.001560 **
points       0.0449557  0.0163604   2.748 0.006417 **
games       -0.0001336  0.0025255  -0.053 0.957843
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2% (a) Which variables are statistically significant at 1% significance level. List their names. [*Hint: No computation required.*] [One line answer]

4% (b) Using the data from both regressions, state the null and alternative hypothesis that *points* and *games* are **jointly** statistically significant. Write down the unrestricted and the restricted model. [Four lines answer]

4% (c) Test the hypothesis stated above at the 1% significance level. Find the critical value. Test the same hypothesis again at the 5% significance level. Find the critical value. [Four lines answer]

4% (d) Do you reject the null hypothesis? Explain the statistical significance of your test at 1% and 5% significance levels. [*Hint: Don't forget to use a specific word when explaining the statistical significance.*] [Four lines answer]

3. (This question refers to **Regression (C)**). Answer the questions below knowing that in this data set we classify a basketball player in one of the following three categories: guard, forward and center.

REGRESSION (C)

```
===============================================
                        Dependent variable:
                    ---------------------------
                               wage
-----------------------------------------------
exper                        224.4807***
                              (48.2237)

age                         -110.9587**
                              (48.3493)

forward                       108.5672
                             (112.4999)

Constant                    3,269.9330***
                            (1,096.8610)

-----------------------------------------------
Observations                    269
R2                             0.1862
Adjusted R2                    0.1770
Residual Std. Error     906.9850 (df = 265)
F Statistic           20.2135*** (df = 3; 265)
===============================================
Note:               *p<0.1; **p<0.05; ***p<0.01
```

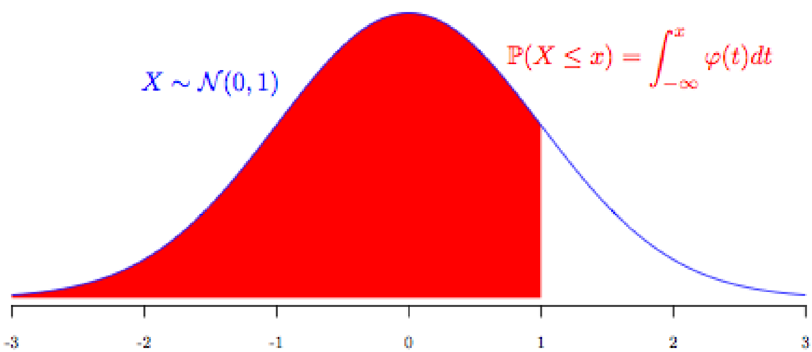
4% (a) State the null and alternative hypothesis of the F statistic for overall significance of a regression. Do you reject the null hypothesis? Explain the statistical significance of your test at 1% significance level. [Three lines answer]

3% (b) What is the estimated average difference in salary between being a forward or not, for players with the same *exper* and *age*? [*Hint: Use the correct measure unit*] [One line answer]

2% (c) All other factors being equal, is there any statistical evidence that being a *forward* player impacts the annual salary of a NBA player? Consider three different significance levels: 1%, 5% and 10% (significance level) in your answer. [Three lines answer]

5% 4. **[Gauss-Markov Theorem]** Under which assumptions does the Gauss-Markov Theorem holds? State and briefly explain each one of them. [One line answer per assumption]

Standard Normal Distribution



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
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
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
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	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Critical Values of the t-distribution

		Significance Level				
1-Tailed:		.10	.05	.025	.01	.005
2-Tailed:		.20	.10	.05	.02	.01
Degrees of Freedom	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	30	1.310	1.697	2.042	2.457	2.750
	40	1.303	1.684	2.021	2.423	2.704
	60	1.296	1.671	2.000	2.390	2.660
	90	1.291	1.662	1.987	2.368	2.632
	120	1.289	1.658	1.980	2.358	2.617
	∞	1.282	1.645	1.960	2.326	2.576

Source: Wooldridge, Jeffrey M. *Introductory Econometrics*, 2015.

1% Critical Values of the F Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom									
	1	2	3	4	5	6	7	8	9	10
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
90	6.93	4.85	4.01	3.54	3.23	3.01	2.84	2.72	2.61	2.52
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

Source: Wooldridge, Jeffrey M. *Introductory Econometrics*, 2015.

5% Critical Values of the F Distribution

		Numerator Degrees of Freedom									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Denominator Degrees of Freedom	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
	26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
	27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
	28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
	29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
	40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
	90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
	120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Source: Wooldridge, Jeffrey M. *Introductory Econometrics*, 2015.