Continuous Assessment in Applied Econometrics – 2 hours

Anderson Vil
CY Cergy Paris Université
2026 – 2027
Mock exam 4

Answer the questions in order. All answers must be justified. Calculators are allowed. Throughout, log denotes the natural logarithm. Critical values are read from the tables at the end of the paper. When the number of degrees of freedom is not printed, use the closest lower row, and use the row ∞ above 120. The marking scheme is indicative.

Exercise A – Training volume and marathon times ***(4 points)***

A running club has followed $N = 10$ of its members during their preparation for the Paris marathon. For runner i, x_i is the average weekly training distance over the last three months (in tens of km) and y_i is the finishing time (in minutes). The club secretary only kept the following sums:

$$\sum_i x_i = 50 \quad \sum_i y_i = 2{,}670 \quad \sum_i x_i^2 = 286 \quad \sum_i x_i y_i = 12{,}963 \quad \sum_i y_i^2 = 717{,}804$$

A1. For the model $y_i = \beta_0 + \beta_1 x_i + u_i$, write $X'X$ and $X'y$, compute $(X'X)^{-1}$ and $\hat{\beta} = (X'X)^{-1}X'y$. Interpret $\hat{\beta}_1$.

A2. Using $X'\hat{u} = 0$, show that $\hat{u}'\hat{u} = y'y - \hat{\beta}'X'y$, and compute the sum of squared residuals.

A3. Compute the R^2 and the unbiased estimator $\hat{\sigma}^2$, and deduce the sample correlation between y_i and $\hat{y}_i$.

A4. The coach argues that "without training, nobody finishes a marathon, so the line must go through the origin" and estimates $y_i = \gamma x_i + v_i$ by OLS. Compute $\hat{\gamma}$ and the sum of squared residuals of this model. Then compute $1 - \text{SSR}/\sum_i (y_i - \bar{y})^2$ and $1 - \text{SSR}/\sum_i y_i^2$. State in three lines what each of these two numbers says about the coach's model.

Exercise B – What fills cinema seats? ***(6 points)***

A film distributor studies $N = 1{,}200$ film-weeks: each observation is a film in a given week after its release in France. The dependent variable is the log of weekly admissions. The regressors are the log of the number of screens showing the film, the number of weeks since release w (from 1 to 12) and $w^2/10$, a dummy equal to 1 if the film is a sequel, a dummy equal to 1 during school holidays, and, in column (4), the log of the film's marketing budget.

B1. Interpret the coefficient on the log number of screens in column (3).

B2. Using column (3), compute the exact percentage change in admissions between the 2nd and the 3rd week after release, then between the 8th and the 9th week. What does the comparison reveal about the life of a film in theatres?

B3. The distributor claims that sequels attract more spectators, other things being equal. Test this claim at the 5% level, stating the alternative hypothesis you choose. The distributor's analyst then proposes to add to column (3) an "original film" dummy, equal to 1 if the film is not a sequel, while keeping the constant. What would happen? Justify.

B4. Using the R^2s, test at the 5% level the joint significance of the sequel and school-holiday dummies.

B5. Between columns (3) and (4), the coefficient on log screens falls from 1.085 to 0.816. In the regression of the log marketing budget on a constant and the other regressors of column (3),

Table 1 – Weekly cinema admissions, OLS estimates

Dependent variable	(1) $\log(adm)$	(2) $\log(adm)$	(3) $\log(adm)$	(4) $\log(adm)$
Log number of screens	1.133	1.086	1.085	0.816
	(0.032)	(0.019)	(0.019)	(0.032)
Weeks since release (w)		-0.398	-0.401	-0.401
		(0.021)	(0.020)	(0.020)
$w^2/10$		0.138	0.141	0.140
		(0.016)	(0.015)	(0.015)
Sequel			0.195	0.194
			(0.040)	(0.038)
School holidays			0.180	0.176
			(0.037)	(0.035)
Log marketing budget				0.292
				(0.028)
Constant	-0.124	1.954	1.865	3.382
	(0.183)	(0.124)	(0.123)	(0.187)
N	1,200	1,200	1,200	1,200
R^2	0.508	0.826	0.833	0.847

Notes: standard errors in parentheses. Unit of observation: a film in a given week after its release. *adm*: admissions in France during the week.

the coefficient on log screens is 0.922. Show that these numbers are consistent, and explain in two sentences, in economic terms, why the coefficient falls.

B6. Admissions are much more variable for films released on few screens than for blockbusters, for given characteristics. Does this affect your answers to B3 and B4? Answer in two or three lines.

Exercise C – Do hospital costs have economies of scale? ***(5 points)***

A ministry note reports the following cost function, estimated by OLS on a sample of public hospitals (the number of hospitals is not reported):

$$\widehat{\log C} = \underset{(0.452)}{-1.873} + \underset{(0.038)}{0.912}\log Q + \underset{(0.071)}{0.684}\log w + \underset{(0.052)}{0.241}\log p_E \qquad R^2 = 0.94$$

where C is the hospital's annual operating cost, Q the number of patient stays, w the average nurse wage and p_E the price of energy paid by the hospital. Standard errors are in parentheses. The note also gives $\widehat{\text{Cov}}(\hat{\beta}_2, \hat{\beta}_3) = -0.0021$, where β_2 and β_3 are the coefficients on $\log w$ and $\log p_E$.

C1. Economic theory says that a cost function is homogeneous of degree one in input prices. Write this property as a hypothesis on the parameters and explain it in one sentence.

C2. Test this hypothesis at the 5% level. The number of hospitals is not reported: explain how you proceed and whether your conclusion depends on it.

C3. Write the model that the ministry should estimate if it wants to impose the hypothesis of C1, specifying the dependent variable and the regressors.

C4. The ministry claims that hospitals have economies of scale. Interpret $\hat{\beta}_1$ for a 10% increase in patient stays, translate the claim into a hypothesis on β_1 and test it at the 5% level against the appropriate alternative. What is the smallest number of hospitals for which your conclusion holds?

Exercise D – How much concentrate should a dairy cow get? ***(5 points)***

An agricultural advisory service observes $N = 240$ dairy farms. For farm i, y_i is the average daily milk yield per cow (litres), f_i the daily quantity of concentrate feed per cow (kg), and h_i is equal to 1 if the herd is of the Holstein breed. Its report gives

$$\hat{y}_i = \hat{\beta}_0 + \underset{(0.467)}{2.450}\, f_i \;\; \underset{(0.0329)}{-0.1022}\, f_i^2 + \hat{\beta}_3 h_i \qquad \hat{\sigma} = 2.537 \qquad \widehat{\mathrm{Cov}}(\hat{\beta}_1, \hat{\beta}_2) = -0.01501$$

but the breed coefficient $\hat{\beta}_3$ was left out of the report.

D1. The service's analyst regressed h on a constant, f and f^2, and denoted the residuals by $\hat{r}_i$. She obtained $\sum_i \hat{r}_i y_i = 193.0$ and $\sum_i \hat{r}_i^2 = 44.53$. Recover $\hat{\beta}_3$ and its standard error, and interpret $\hat{\beta}_3$.

D2. In the regression of y on a constant, f and f^2 only, the coefficient on f is 2.988. In the regression of h on a constant, f and f^2, the coefficient on f is 0.1241. Show that these numbers are consistent with the report, and explain the direction of the gap in one sentence.

D3. Milk sells at 0.45 euro per litre and concentrate costs 0.36 euro per kg. The farmer maximises profit by feeding up to the point where one more kg of concentrate yields $0.36/0.45 = 0.80$ litre of milk. Compute this optimal quantity f^*, its standard error by the delta method, and a 95% confidence interval based on the normal approximation. The advisory service currently recommends 10 kg per cow per day. Is this recommendation supported by the data?

D4. The farmers who feed more concentrate may also be the best managers, and good management raises yields for a given feed. Which estimate of D3 would this affect, and in which direction? Answer in two lines.

TABLE G.2 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

Examples: The 1% critical value for a one-tailed test with 25 *df* is 2.485. The 5% critical value for a two-tailed test with large (> 120) *df* is 1.96.

Source: This table was generated using the Stata® function invttail.

Copyright 2016 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

TABLE G.3a 10% 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	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32
	11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25
	12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19
	13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14
	14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10
	15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06
	16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03
	17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00
	18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
	19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96
	20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94
	21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92
	22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90
	23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89
	24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88
	25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87
	26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86
	27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85
	28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84
	29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83
	30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82
	40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76
	60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71
	90	2.76	2.36	2.15	2.01	1.91	1.84	1.78	1.74	1.70	1.67
	120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65
	∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60

Example: The 10% critical value for numerator $df = 2$ and denominator $df = 40$ is 2.44.
Source: This table was generated using the Stata® function invFtail.

Copyright 2016 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

TABLE G.3b 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

Example: The 5% critical value for numerator $df = 4$ and large denominator $df(\infty)$ is 2.37.

Source: This table was generated using the Stata® function invFtail.

Copyright 2016 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

TABLE G.3c 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

Example: The 1% critical value for numerator $df = 3$ and denominator $df = 60$ is 4.13.

Source: This table was generated using the Stata® function invFtail.

Copyright 2016 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.