Übung zur Vorlesung Ökonometrie
Masterstudiengang
Sommersemester 2025

Aufgabenblatt 2

1. Sei $\mathbf{z}=(v,w)'$, worin v und w Zufallsvariablen sind. Zeigen Sie, dass

$$\operatorname{Var}(\mathbf{z}) = \begin{bmatrix} \operatorname{Var}(v) & \operatorname{Cov}(v,w) \\ \operatorname{Cov}(v,w) & \operatorname{Var}(w) \end{bmatrix}$$

   ist.

2. Betrachten Sie das Modell $\mathbf{y} = X\beta + \mathbf{u}$ mit $E(\mathbf{u}|X) = 0$, worin X eine Regressormatrix der Dimension 30×3 ist und $\beta = (\beta_0, \beta_1, \beta_2)'$ ist. Sie haben den OLS-Schätzvektor $\widehat{\beta}$ ermittelt.

   (a) Definieren Sie die Kovarianzmatrix $\operatorname{Var}(\widehat{\beta}|X)$ von $\widehat{\beta}$. Nutzen Sie zur Vereinfachung der Notation die Tatsache, dass $\widehat{\beta}$ erwartungstreu ist.

   (b) In Ihrem Datensatz liegen folgende konkrete Werte vor:

$$(X'X)^{-1} = \begin{pmatrix} 0.035004 & 0.007120 & -0.011401 \\ 0.007120 & 0.039085 & -0.031120 \\ -0.011401 & -0.031120 & 0.112720 \end{pmatrix}, \quad \hat{u}'\hat{u} = 39.090258$$

   Ermitteln Sie $\operatorname{Var}(\widehat{\beta}|X)$ unter Annahme MLR 5 $\Sigma(\mathbf{u}) = \sigma^2 I$.

   (c) Geben Sie den Standardfehler von $\widehat{\beta}_2$ an.

3. Sie wollen eine lineare Regression $\mathbf{y} = X\beta + \mathbf{u}$ schätzen. Allerdings ist die abhängige Variable $\mathbf{y}$ unbeobachtbar, weil die Daten nur mit Messfehlern erhoben werden. Die gemessenen Daten $\mathbf{z}$ hängen mit der tatsächlichen Variablen $\mathbf{y}$ zusammen gemäß $\mathbf{z} = \mathbf{y} + \mathbf{v}$, wobei $\mathbf{v}$ ein Vektor von zufälligen Messfehlern mit $E(\mathbf{v}) = 0$ ist. Die Zufallsfehler $\mathbf{v}$ und $\mathbf{u}$ sind unabhängig voneinander und von X. Bis auf die Messfehler erfüllt das Modell alle klassischen Annahmen, insbesondere ist die Spezifikation korrekt und es gilt $E(\mathbf{u}|X) = 0$.

   (a) Sie schätzen das Modell mit $\mathbf{z}$ als der abhängigen Variablen mit OLS, und ignorieren dabei zunächst die Messfehler. Geben Sie die Formel für den Schätzer $\widehat{\beta}$ an.

   (b) Ermitteln Sie, ob der unter (a) genannte Schätzer erwartungstreu ist (Hinweis: setzen Sie in der Schätzformel nacheinander die in der Aufgabenstellung gegebenen Ausdrücke für $\mathbf{z}$ und dann für $\mathbf{y}$ ein und bilden Sie den bedingten Erwartungswert von $\widehat{\beta}$ gegeben X).

   (c) Begründen Sie möglichst genau, ob die Standardfehler der geschätzten Regressionsparameter verglichen mit dem Fall ohne Messfehler größer, kleiner oder gleich groß sind.

4. Sie schätzen das Modell $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$ aus einer Stichprobe vom Umfang $n = 42$, und erhalten den Schätzer $\widehat{\beta} = (2.2, 1.1, 0.85)'$ und die zugehörigen Standardfehler $se(\widehat{\beta}) = (1.3, 1.2, 0.2)'$.

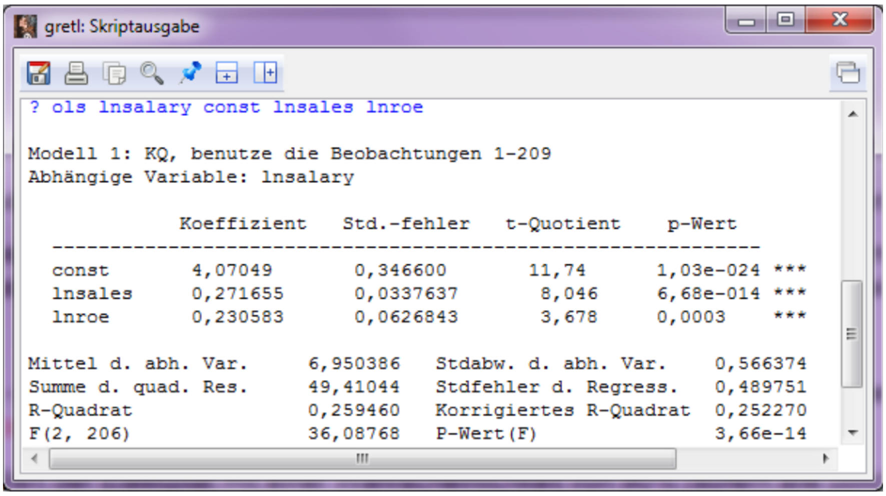
   (a) Sie hatten aus theoretischen Überlegungen erwartet, für β_2 einen Schätzwert von 1 zu erhalten. Mit welcher Art von Test können Sie überprüfen, ob das Ergebnis damit vereinbar ist?

   (b) Formulieren Sie einen zweiseitigen Test der entsprechenden Nullhypothese. Wie lautet die Teststatistik, und welcher Verteilung folgt sie (unter welcher Annahme)?

   (c) Zu welcher Testentscheidung gelangen Sie im vorliegenden Fall?

   (d) Formulieren Sie das Testergebnis in Worten, und ziehen Sie eine Schlussfolgerung.

5. Zahlen umsatzstärkere Unternehmen höhere Managergehälter? Der Wooldridge - Datensatz `ceosal1` enthält Daten über 209 Unternehmen. Sie ermitteln folgenden Zusammenhang zwischen den (logarithmischen) Vorstandsgehältern (ln `salary`), der Eigenkapitalrendite (ln `roe`) und dem Umsatz (ln `sales`) der Unternehmen in der Stichprobe:




```
? ols lnsalary const lnsales lnroe

Modell 1: KQ, benutze die Beobachtungen 1-209
Abhängige Variable: lnsalary

             Koeffizient   Std.-fehler   t-Quotient    p-Wert
  ---------------------------------------------------------------
  const       4,07049       0,346600      11,74       1,03e-024 ***
  lnsales     0,271655      0,0337637      8,046      6,68e-014 ***
  lnroe       0,230583      0,0626843      3,678      0,0003    ***

Mittel d. abh. Var.      6,950386   Stdabw. d. abh. Var.     0,566374
Summe d. quad. Res.      49,41044   Stdfehler d. Regress.    0,489751
R-Quadrat                0,259460   Korrigiertes R-Quadrat   0,252270
F(2, 206)                36,08768   P-Wert(F)                3,66e-14
```

   (a) Nehmen Sie an, die Störgröße sei normalverteilt, und konstruieren Sie ein 90 % - Konfidenzintervall für die Elastizität von `salary` bezüglich `sales`.

   (b) Erläutern Sie verbal die Interpretation des Ergebnisses.

6. Sie haben eine lineare Regression mit 5 Regressoren sowie einer Konstanten basierend auf einer Stichprobe mit 67 Beobachtungen geschätzt. Der Fehlerterm ist normalverteilt. Für einen Koeffizienten, der Sie besonders interessiert, erhalten Sie einen Schätzer von $\widehat{\beta}_j = 0.82$. Ihr Regressionsprogramm gibt außerdem die Information: 't - Quotient $= 8$´ an. Die folgende Tabelle weist kritische Werte c der t - Verteilung aus:

	$P(t > c) =$				
Freiheitsgrade ↓	0.1	0.05	0.025	0.01	0.005
61	1.2956	1.6702	1.9996	2.3890	2.6589
62	1.2954	1.6698	1.9990	2.3880	2.6575
63	1.2951	1.6694	1.9983	2.3870	2.6561

(a) Erläutern Sie kurz, wie der üblicherweise von Regressionsprogrammen ausgegebene 't - Quotient' (gleichbedeutend t - Wert, t - Statistik) generell interpretiert werden muss.

(b) Wie ist die Information 't - Quotient = 8' im vorliegenden Fall konkret zu deuten?

(c) Testen Sie die Nullhypothese: $H_0 : \beta = 1$ gegen die Alternative $H_1 : \beta \neq 1$ auf dem 5% - Signifikanzniveau.

(d) Testen Sie die Nullhypothese: $H_0 : \beta = 0.8$ gegen die Alternative $H_1 : \beta \neq 0.8$ auf dem 1% - Signifikanzniveau.

(e) Testen Sie die Nullhypothese: $H_0 : \beta < 0.62$ gegen die Alternative $H_1 : \beta > 0.62$ auf dem 5% - Signifikanzniveau.

(f) Geben Sie ein 90% - Konfidenzintervall für den Koeffizienten an.

(g) Interpretieren Sie in einem Satz das Ergebnis aus der vorigen Teilaufgabe.

7. Geben Sie zu jeder der folgenden Aussagen an, ob Sie sie für richtig, falsch oder bei gegebener Information unentscheidbar halten. Wenn Sie eine Aussage für falsch oder unentscheidbar halten, besteht die korrekte Aufgabenlösung in einer kurzen Begründung, worin der von Ihnen entdeckte Fehler besteht oder welche Information zur Entscheidung fehlt. Aussagen zu Tests beziehen sich jeweils auf zweiseitige Hypothesen.

(a) Wenn der p - Wert einer Teststatistik größer als 0.05 ist, kann die Nullhypothese bei jedem geringeren Signifikanzniveau als 5% abgelehnt werden.

(b) Wenn der p - Wert des t - Tests mit der Nullhypothese, ein Regressionsparameter sei gleich null, 0.03 beträgt, kann die Nullhypothese auf einem Signifikanzniveau von 5% abgelehnt werden.

(c) Wenn der p - Wert des t - Tests mit der Nullhypothese, ein Regressionsparameter sei gleich null, 0.05 beträgt, dann ist der Betrag des Wertes der t - Statistik gleich dem 0.95 - Quantil der t - Verteilung mit der entsprechenden Anzahl von Freiheitsgraden.

(d) Wenn der p - Wert des t - Tests mit der Nullhypothese, ein Regressionsparameter sei gleich null, 0.04 beträgt, dann muss der Wert der t - Statistik größer als das 0.975 - Quantil der t - Verteilung mit der entsprechenden Anzahl von Freiheitsgraden sein.

(e) Wenn der p - Wert des t - Tests mit der Nullhypothese, ein Regressionsparameter sei gleich null, 0.05 unterschreitet, ist der Parameter mit einer Wahrscheinlichkeit von weniger als 5% null.

(f) Wenn der p - Wert des t - Tests mit der Nullhypothese, ein Regressionsparameter sei gleich null, 0.05 unterschreitet, dann überdeckt das 95%-Konfidenzintervall die null nicht.

(g) Wenn der p-Wert des t-Tests mit der Nullhypothese, ein Regressionsparameter sei gleich null, 0.12 beträgt, dann überdeckt das 90%-Konfidenzintervall die null.

(h) Wenn das 95% - Konfidenzintervall um einen Regressionsparameter die null umschließt, ist der geschätzte Parameter statistisch nicht signifikant von null verschieden.

(i) Wenn die Schätzung eines Regressionsparameters 2 ergibt, und das 90% - Konfidenzintervall mit $[1.6, 2.4]$ geschätzt wird, beträgt die Wahrscheinlichkeit, dass das Intervall den wahren Wert beinhaltet 90%.

(j) Wenn die Schätzung eines Regressionsparameters 2 ergibt, und das 90% - Konfidenzintervall mit $[1.6, 2.4]$ geschätzt wird, beträgt die Wahrscheinlichkeit, dass der geschätzte Parameter außerhalb des Intervalls liegt 10%.