Diese Aufgabe wurde zur Klausurvorbereitung am 4.2.09 in der Vorlesung vorgerechnet. Es ist eine ***Ergänzung von Aufgabe 1 des "Klausuraufgaben" downloads*** [1] (Nummer der Aufgabe 41-0806)

Auf dieser Seite die Aufgabenstellung, auf den nächsten Seiten die Lösung

Nr.	Punktzahl	

Die Schätzung einer linearen einfachen Regressionsfunktion möge ergeben haben $\hat{y}_t = 0 + 2x_t$ (mit t = 1,…, T und T = 100) also $\hat{\alpha} = 0$ und $\hat{\beta} = 2$ und die folgenden Ergebnisse $\overline{y} = 0.2$, $\overline{x} = 0.1$, $\sum_{t=1}^{T} x_t y_t = 4$, $\sum_{t=1}^{T} x_t^2 = 2$, $\sum_{t=1}^{T} y_t^2 = 10$, $\hat{\sigma}_u^2 = 2/98$

a) Man berechne und interpretiere* die geschätzten Varianzen von $\hat{\beta}$ und $\hat{\alpha}$ mit

$$\hat{\sigma}^2{}_{\hat{\beta}} = \frac{\hat{\sigma}^2}{S_{xx}} \text{ und } \hat{\sigma}^2_{\hat{\alpha}} = \hat{\sigma}^2\left(\frac{1}{T} + \frac{\overline{x}^2}{S_{xx}}\right) = \overline{X}^2 \hat{\sigma}^2_{\hat{\beta}}$$

wobei $\overline{X}^2 = \sum x_t^2 / T \neq (\overline{x})^2$ das zweite Anfangsmoment ist

(verifizieren Sie auch, dass beide Arten der Berechnung von $\hat{\sigma}^2_{\hat{\alpha}}$ zum gleichen Ergebnis führen)

b) Man bestimme die Korrelation zwischen $\hat{\alpha}$ und $\hat{\beta}$ (die geschätzte Kovarianz ist $\hat{\sigma}_{\hat{\alpha}\hat{\beta}} = -\overline{x}\hat{\sigma}^2_{\hat{\beta}}$)

c) Bestimmen Sie ein *Konfidenzintervall* (zweiseitig symmetrisch 95%) für α und *testen* Sie die Hypothese H_0: $\beta = 0$

(Irrtumswahrscheinlichkeit bzw. Signifikanzniveau 5 % und 1% zweiseitig). Der Tabellenwert ist der Wert der ……. Verteilung mit dem Parameter v = ……. Für die gesuchten Tabellenwerte erhält man

0.9	0.95	0.975	0.99	0.995
1.29	1.66	1.98	2.36	2.63

d) Wenn man ein Konfidenzintervall für die Varianz der Störgröße u berechnen möchte, wie ist dann vorzugehen. Eine kurze Beschreibung des Vorgehens reicht aus. Berechnungen sind nicht durchzuführen.

* **Hinweis:** Mit "Interpretation" ist hier gemeint: wovon hängt die geschätzte Varianz von …. ab, und warum ist das plausibel.

[1] Es heißt dort: In einem Modell wird die Aktienrendite y_t wie folgt mit der Rendite des Marktportfolios x_t (t = 1, …, T= 100 Monatswerte) erklärt …. wobei $\overline{x}$ und $\overline{y}$ die Mittelwerte der Variablen bezeichnen. Die Renditen sind in Dezimalzahlen gemessen (d.h. 0.1 = 10 Prozent).

Nr.	Punktzahl	Lösung in blauer Schrift

Die Schätzung einer linearen einfachen Regressionsfunktion möge ergeben haben $\hat{y}_t = 0 + 2x_t$ (mit t = 1,…, T und T = 100) also $\hat{\alpha} = 0$ und $\hat{\beta} = 2$ und die folgenden Ergebnisse $\bar{y} = 0.2, \bar{x} = 0.1, \sum_{t=1}^{T} x_t y_t = 4, \sum_{t=1}^{T} x_t^2 = 2, \sum_{t=1}^{T} y_t^2 = 10, \quad \hat{\sigma}_u^2 = 2/98$

a) Man berechne und interpretiere die geschätzten Varianzen von $\hat{\beta}$ und $\hat{\alpha}$ mit

$$\hat{\sigma}^2{}_{\hat{\beta}} = \frac{\hat{\sigma}^2}{S_{xx}} \text{ und } \hat{\sigma}_{\hat{\alpha}}^2 = \hat{\sigma}^2\left(\frac{1}{T} + \frac{\bar{x}^2}{S_{xx}}\right) = \overline{X}^2 \hat{\sigma}_{\hat{\beta}}^2 \quad \text{mit} \quad \overline{X}^2 = \sum x_t^2 / T$$

Berechnungen

Es ist eine gewisse Schwierigkeit, zuerst S_{xx} auszurechnen. Es ergibt sich durch Ausmultiplizieren $S_{xx} = \sum (x - \bar{x})^2 = \sum x^2 - T \cdot (\bar{x})^2$ = 2 - 100(0,1)^2 = 2 – 1 = 1: Damit erhält man wegen $\hat{\sigma}_u^2 = \hat{\sigma}^2 = 2/98$, so dass auch $\hat{\sigma}^2{}_{\hat{\beta}}$ = 2/98 ist

(verifizieren Sie auch, dass beide Arten der Berechnung von $\hat{\sigma}_{\hat{\alpha}}^2$ zum gleichen Ergebnis führen)

Hierzu reicht es aus, beide Berechnungsarten (nach den beiden oben angegebenen Formeln) durchzuführen und zu sehen, dass sie zum gleichen Ergebnis führen:

$\hat{\sigma}_{\hat{\alpha}}^2 = \hat{\sigma}^2\left(\frac{1}{T} + \frac{\bar{x}^2}{S_{xx}}\right)$ ergibt $\frac{2}{98}\left(\frac{1}{100} + \frac{0{,}01^2}{1}\right) = \frac{0{,}04}{98}$ = 0,000408 und

$\hat{\sigma}_{\hat{\alpha}}^2 = \overline{X}^2 \hat{\sigma}_{\hat{\beta}}^2$ ergibt $\overline{X}^2 \hat{\sigma}_{\hat{\beta}}^2 = \frac{2}{100} \cdot \frac{2}{98}$ = 0,000408, also den gleichen Wert

(Man beachte, die Varianz der x-Werte ist $s_x^2 = S_{xx}/T = \overline{X}^2 - (\bar{x})^2$ = 1/100 = 2/100 – 0,01)

Interpretation

wovon hängt die geschätzte Varianz von $\hat{\beta}$ ab (hier und bei der Frage nach am besten kleine Skizzen zeichnen, um zu zeigen wie das Streuungsdiagramm aussieht, wenn $\hat{\sigma}_u^2 = \hat{\sigma}^2$, S_{XX} oder klein bzw. groß ist):

Die beiden Faktoren sind

1. die Streuung um die Regressionsgerade (groß wenn $\hat{\sigma}_u^2 = \hat{\sigma}^2$ groß ist, dann ist die Schätzung unsicher) und
2. von dem Streubereich der x-Werte (S_{xx}). Wenn S_{xx} klein ist, dann ist die Schätzung unsicher; Extremfall: wenn S_{xx} = 0 (dann ist auch die Varianz der x Werte 0 und alle Punkte im Streuungsdiagramm liegen übereinander, eine Schätzung der Regressionsgerade ist dann überhaupt nicht möglich

bei der Varianz von $\hat{\alpha}$ also $\hat{\sigma}_{\hat{\alpha}}^2 = \overline{X}^2 \hat{\sigma}_{\hat{\beta}}^2$ kommt noch ein dritter Einflussfaktor hinzu

3. die Größe $\overline{X}^2$: ist dies klein, dann liegen die Punkte im Streuungsdiagramm in der Nähe der y-Achse und die Schätzung des Ordinatenabschnitts α ist dann sicherer als wenn $\overline{X}^2$ groß ist (auch hier ein Skizze nützlich)

b) Man bestimme die Korrelation zwischen $\hat{\alpha}$ und $\hat{\beta}$ (die geschätzte Kovarianz ist $\hat{\sigma}_{\hat{\alpha}\hat{\beta}} = -\bar{x}\hat{\sigma}^2_{\hat{\beta}}$)

= - 0,1. 2/98 = -0,00204 und damit ist $r_{\hat{\alpha}\hat{\beta}} = \hat{\sigma}_{\hat{\alpha}\hat{\beta}} \Big/ \sqrt{\hat{\sigma}^2_{\hat{\alpha}}\hat{\sigma}^2_{\hat{\beta}}}$ und hier ergibt sich eine Kontrollmöglichkeit, weil r absolut nicht größer als 1 sein kann. Man erhält hier $r_{\hat{\alpha}\hat{\beta}} = -0{,}1 \cdot 2/98 \Big/ \sqrt{(2/98)\cdot(0{,}04/98)}$ und somit r = - 0,707 = $-1/\sqrt{2}$.

c) Bestimmen Sie ein *Konfidenzintervall* (zweiseitig symmetrisch 95%) für α und *testen* Sie die Hypothese H_0: β = 0

(Irrtumswahrscheinlichkeit bzw. Signifikanzniveau 5 % und 1% zweiseitig). Der Tabellenwert ist der Wert der …t…. Verteilung

mit dem Parameter v = T – 2 = 100 – 2 = 98…Freiheitsgrade…).

Für die gesuchten Tabellenwerte erhält man

0.9	0.95	0.975	0.99	0.995
1.29	1.66	1.98	2.36	2.63

Konfidenzintervall zweiseitig symmetrisch für α:

Tabellenwert ist dann 1,98: die Grenzen sind gegeben mit

$\hat{\alpha} \pm t \cdot \hat{\sigma}_{\hat{\alpha}} = 0 \pm 1{,}98 \cdot \sqrt{0{,}04/98}$ also – 0,04 und + 0,04

Test von H_0: β = 0

Prüfgröße $\hat{\beta}/\hat{\sigma}_{\hat{\beta}}$ also $2\Big/\sqrt{2/98}$ = 14 was weit größer ist als 1,98 oder 2,63, so dass H_0 abzulehnen ist: β ist signifikant von 0 verschieden auf dem 5% und 1% und einem noch viel geringeren Niveau (der prob value ist praktisch 0,000…).

d) Wenn man ein Konfidenzintervall für die Varianz der Störgröße u berechnen möchte, wie ist dann vorzugehen. Eine kurze Beschreibung des Vorgehens reicht aus. Berechnungen sind nicht durchzuführen.

Der Punktschätzer ist $\hat{\sigma}^2_u = 2/98$ = 0,02041. Man sucht dann die untere Grenze G_u (Wahrscheinlichkeit 0,025) und die oberer Grenze G_o (bei 0,975) gemäß χ^2 Verteilung bei t – 2 = 98 Freiheitsgraden und erhält die

untere Grenze des Konfidenzintervalls mit 0,02041/G_o
obere Grenze … mit 0,02041/G_u

Der Punktschätzer (0,02…) wird nicht genau in der Mitte des Intervalls liegen weil die χ^2 Verteilung linkssteil [= rechtsschief] ist-