Korrektur der Ausführungen im Regressionsbuch (3. Auflage) zum einseitigen Signifikanztest

Urban/Mayerl, 7.7.2009

–> Der folgende Text ersetzt den alten Text:
ab: S. 149, Ende des vorletzten Absatzes "... Wert "β=0" zu betrachten."
bis: S. 150, Ende des dritten Absatzes von oben "... an dieser Stelle keinen Sinn."

–> zu löschen ist der Kasten "Ergänzung 3.7"

Wie zuvor gezeigt, kann also ein zweiseitiger Test der H_0-Hypothese "β=0" dadurch ausgeführt werden, dass ein Konfidenzintervall berechnet wird und dann geschaut wird, ob der β=0-Wert in diesem Intervall liegt. Liegt der Wert nicht darin, so ist die Nullhypothese mit der gewählten Irrtumswahrscheinlichkeit zu verwerfen und b als "signifikant" zu bezeichnen.

Dieser Signifikanztest wird zweiseitig durchgeführt. Denn die zur Nullhypothese "β=0" korrespondierende Alternativhypothese lautet: "$\beta \neq 0$". Ob also der "wahre" Koeffizient β links oder rechts von β=0 liegt, bzw. ob er negativ oder positiv ist, wird damit offen gelassen (vgl. dazu Abb. 3.7).

Ist man jedoch bereit anzunehmen, dass der zu schätzende β-Wert auf jeden Fall positiv sein muss, so kann man den Signifikanztest auch allein "rechts-einseitig" durchführen. Der 5%ige Signifikanzbereich liegt dann ausschließlich auf der rechten Seite der Wahrscheinlichkeitsverteilung und beginnt ab der oberen (positiven) Grenze des 95% Konfidenzintervalls (vgl. die folgende Abb. 3.11.1). Natürlich lässt sich der Test auch "links-einseitig" durchführen. Der Signifikanzbereich liegt dann ausschließlich am linken Ende der Wahrscheinlichkeitsverteilung.

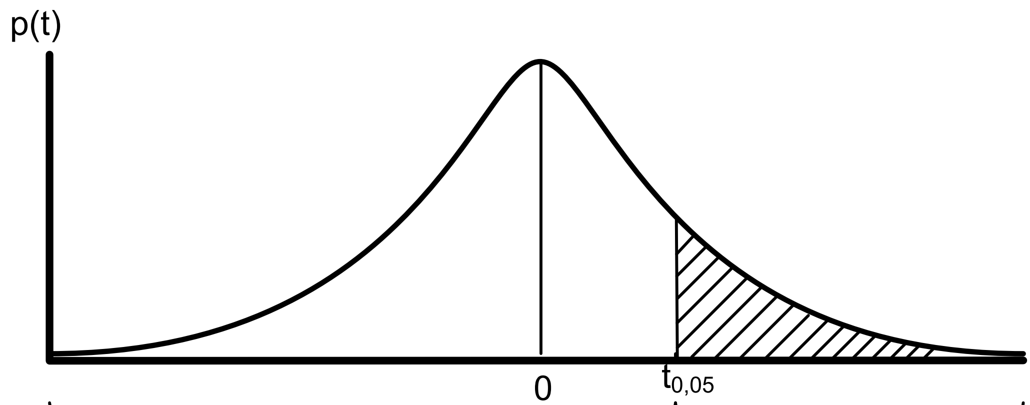
Leider kann man einen einseitigen Signifikanztest nicht nach der oben vorgestellten Logik des zweiseitigen Tests durchführen. Stattdessen braucht man für einen einseitigen Test die theoretische Wahrscheinlichkeitsverteilung der sogenannten t-Statistik. Für große Fallzahlen entspricht ihre Verteilungsform derjenigen einer Normalverteilung (vgl. die folgende Abbildung 3.11.1).

Ein empirischer t-Wert wird berechnet durch Division des geschätzten b-Koeffizienten durch dessen Standardfehler:

$$t = b / SE_b$$

Nach Gl. 3.10 beträgt b=9,37 und SE_b=10,82. Dividieren wir b durch SE_b ergibt sich somit ein t-Wert von 0,87. Mit diesem Wert können wir einen rechts-einseitigen Signifikanztest durchführen. Wir benutzen dazu eine t-Wahrscheinlichkeitsverteilung, so wie sie in Abbildung 3.11.1 dargestellt wird. Entsprechend dieser Abbildung wird der Signifikanzbereich vom Konfidenzbereich durch den Wert "$t_{0.05}$" abgetrennt, so dass 5% aller t-Werte rechts von diesem Wert liegen. Aus Tabelle A1 (im Anhang) lässt sich entnehmen, dass in unserem Beispiel der kritische $t_{0.05}$-Wert für einen einseitigen Test mit 10 Freiheitsgraden (die sich ergeben aus 12 Beobachtungsfällen minus 2 zu schätzender Parameter) einen numerischen Wert von 1,81 aufweist. Somit liegt der empirische t-Wert von 0,87 (s.o.) in der Abbildung links von dem kritischen t-Wert und befindet sich damit im Konfidenzbereich der t-Verteilung. Mithin kann er nicht als signifikant gelten, und es muss auch weiterhin angenommen werden, dass die H_0 (β=0) wahr ist. Ein Schätzwert von b=9,37 wäre somit als eine rein zufällige Abweichung vom Nullwert zu verstehen.

Abbildung 3.11.1: Rechts-einseitiger Signifikanztest



Nach der hier vorgestellten Testlogik könnte auch ein links-einseitiger Signifikanztest durchgeführt werden. Da ein solcher Test allerdings voraussetzt, dass der wahre Wert des Regressionskoeffizienten negativ ist und in unserem Beispiel ein durchgängig negativer Effekt des Alters auf das zu erwartende Einkommen wohl auszuschließen ist, macht ein solcher Test an dieser Stelle keinen Sinn.