

Übung: Schätzung von Probit-Modellen

1. Verkehrsmittelwahl

Reproduzieren Sie die in der Vorlesung besprochenen Resultate des Beispiels zur Verkehrsmittelwahl Auto versus Bus. Die Daten dazu finden Sie im File AutoBus.xls.

Warum ist der Effekt eines zusätzlichen 1-minütigen Zeitnachteils des Busses bei Pendler Nr. 14 am grössten und bei den Pendlern Nr. 6 und 8 sehr klein?

Schätzen Sie ab, wie viele Pendler auf den Bus umsteigen, falls sich die Bus-Fahrzeiten halbieren!

2. Wahlverhalten in amerikanischen Präsidentschaftswahlen

Die Tabelle VoteUSA.xls zeigt die Resultate der amerikanischen Präsidentschaftswahlen von 1976 in den 51 Teilstaaten. Damals stand der Demokrat Jimmy Carter dem Republikaner Gerald Ford gegenüber. Die abhängige Variable y ist auf 1 gesetzt, falls die Mehrheit demokratisch stimmte (dem > rep) und der betreffende Staat somit an Carter ging. Andernfalls ist y auf 0 gesetzt. Die anderen Variablen in der Tabelle sind:

income	mittleres Pro-Kopf Einkommen
school	mittlere Anzahl Schuljahre der Personen, die mehr als 18 Jahre alt sind
urban	prozentualer Bevölkerungsanteil in städtischen Agglomerationen
region	geografische Lage des Staates: 1 "Northeast", 2 "Southeast", 3 "Midwest and Middle South", 4 "West and Mountain regions".

Schätzen Sie ein Probit-Modell für den Wahlausgang ($y = 1$ bzw. 0) mit dem Einkommen ($\log(\text{income})$) und dem "Urbanitätsgrad" ($\text{urban}/100$) als erklärende Variablen!

a) Auch in den USA hat der Anteil der Bevölkerung in städtischen Agglomerationen in den letzten 20 Jahren zugenommen. Welche Partei sollte gemäss ihren Schätzergebnissen davon profitieren? Ist dieser Zusammenhang statistisch signifikant?

b) Weicht das Wahlverhalten in Staaten der Region 4 ("West and Mountain regions") signifikant vom amerikanischen Durchschnitt ab? Beantworten Sie diese Frage, indem sie das Probit-Modell um eine erklärende Variable reg4 erweitern. reg4 wird auf 1 gesetzt, falls der Staat der Region 4 angehört, sonst auf 0. (in Eviews: $\text{reg4} = @\text{recode}(\text{region} = 4, 1, 0)$).

c) Stellen Sie die Wahrscheinlichkeit eines demokratischen Wahlsiegs grafisch in Abhängigkeit des Einkommens dar, wobei Sie urban auf dem Stichprobenmittelwert fixieren und $\text{reg4} = 0$ bzw. $=1$ setzen!

d) Wie gross ist gemäss Ihrem Modell die Wahrscheinlichkeit, dass Oregon demokratisch wählt?

e) "Prognostizieren" Sie die Wahlausgang in den 51 Staaten anhand der naheliegenden Regel: $\text{Prob}(y = 1) > 0.5 \Rightarrow y = 1$
 $\text{Prob}(y = 1) \leq 0.5 \Rightarrow y = 0$

In wie vielen Fällen stimmt das Wahlergebnis mit der Modellprognose überein?

f) Wie wirkt sich eine Einkommenszunahme von 1% auf die Wahrscheinlichkeit eines demokratischen Wahlsiegs in Florida, in Idaho und Minnesota aus? Erklären Sie die grossen Unterschiede!

g) Auf welche Staaten würden Sie als Strategieberater der Demokraten ihre Wahlpropaganda konzentrieren?

Nützliche Eviews-Funktionen:

genr $z = @recode(\text{condition}, x, y)$

Setzt $z = x$ falls die Bedingung zutrifft, andernfalls $z = y$

genr $z = @dnorm(x)$

Berechnet die Dichtefunktion der Standardnormalverteilung an der Stelle x :

$$z = f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right), \text{ mit } \mu = 0 \text{ und } \sigma = 1$$

genr $z = @cnorm(a)$

Berechnet die kumulative Standardnormalverteilung an der Stelle a (Fläche unter der Dichtefunktion bis zur Obergrenze a):

$$z = \text{Prob}(x < a) = \int_{-\infty}^a f(x) dx$$

Die Funktionen können auch verknüpft werden, z.B. mit

genr $z = @recode(@cnorm(w) > 0.5, 1, 0)$

wird z auf 1 gesetzt, falls $@cnorm(w)$ (eine Wahrscheinlichkeit) grösser ist als 0.5. Anderfalls wird z auf 0 gesetzt.