1.1 Korrekturen zu Kapitel 1

1.3.2

Tabelle 1.3. Bedingte Erwartungswerte von Y, $E(Y|X)$

	X=10	X=20	X=30
	$P(X = 10) = 1/4$	$P(X = 20) = 5/12$	$P(X = 30) = 1/3$
	$P(y\|X = 10)$	$P(y\|X = 20)$	$P(y\|X = 30)$
$Y = 10$	1/3	1/5	-
$Y = 20$	1/3	1/5	1/4
$Y = 30$	1/3	1/5	1/4
$Y = 40$	-	1/5	1/4
$Y = 50$	-	1/5	1/4
$E(Y\|X)$	20,00	30,00	35,00
$Var(Y\|X)$	66,67	200,00	125,00

Ist Y eine stetige Zufallsvariable, berechnet sich der bedingte Erwartungswert als

$$E(Y|X) = \int_{-\infty}^{\infty} f(y|x)y\ dy. \tag{1.37}$$

Basierend auf den in **Tabelle 1.3** dargestellten Werten ist der bedingte Erwartungswert von Y gegeben $X = 20$ $E(Y|X = 20) = 30$ und die entsprechende bedingte Varianz $Var(Y|X = 20) = 200,00$.

Übungsaufgaben

1.7

a) Berechnen Sie $P(Y \leq 2{,}30)$, wenn $Y \sim F^5_{100}$.
b) Berechnen Sie $P(Y \leq 2{,}85)$, wenn $Y \sim F^{15}_{10}$.
c) Berechnen Sie $P(Y > 2{,}69)$, wenn $Y \sim F^{40}_{20}$.

1.8 X und Z seien zwei unabhängige Zufallsvariablen, die einer Standardnormalverteilung folgen. Nehmen Sie an, die Zufallsvariable Y berechne sich als $Y = X^2 + Z$.

a) Zeigen Sie, dass $E(Y|X) = X^2$.
c) Zeigen Sie, dass $E(XZ) = 0$.
d) Zeigen Sie, dass $Cov(X, Y) = 0$.

Übungsaufgaben

5.10 Sir Francis Galton, ein Cousin von Charles Darwin, untersuchte gegen Ende des 19. Jahrhunderts die Beziehung zwischen der Größe von Kindern und der Größe ihrer Eltern. Sie aktualisieren die Untersuchung dieses Zusammenhangs durch Daten von 500 Kommilitonen und schätzen folgendes Modell:

$$\begin{aligned} \text{Student} &= \underset{(7,2)}{49,9} + \underset{(0,10)}{0,73}\ \text{Eltern} \\ R^2 &= 0,45 \end{aligned}$$

wobei *Student* die Größe (in cm) der Kommilitonen und *Eltern* die durchschnittliche Größe (in cm) der jeweiligen Eltern der Kommilitonen repräsentiert. Die Werte in Klammern sind geschätzte Standardfehler.

a) Interpretieren Sie die geschätzten Koeffizienten quantitativ.
b) Wenn erwartet würde, dass Kinder im Durchschnitt die gleiche Größe hätten wie ihre Eltern, dann würde dies zwei Hypothesen implizieren, von denen eine den Steigungskoeffizienten und die andere den Achsenabschnitt betrifft.
   (i) Was sollte die Nullhypothese für den Achsenabschnitt sein? Berechnen Sie die relevante t-Statistik und führen Sie einen Hypothesentest für eine Irrtumswahrscheinlichkeit von 1% durch.
   (ii) Was sollte die Nullhypothese für den Steigungskoeffizienten sein? Berechnen Sie die relevante t-Statistik und führen Sie einen Hypothesentest für eine Irrtumswahrscheinlichkeit von 5% durch.
c) Was ist die vorhergesagte Größe eines Kindes, dessen Eltern eine durchschnittliche Größe von 170 cm haben?
d) Gegeben, dass der Achsenabschnitt positiv und der Steigungskoeffizient zwischen 0 und 1 liegt, was lässt sich über die Größe von Studenten sagen, die relativ große Eltern haben? Was lässt sich über die Größe von Studenten sagen, die relativ kleine Eltern haben?

8.1 Korrekturen zu Kapitel 8

Übungsaufgaben

8.4 Unter Verwendung von Daten des Sozioökonomischen Panels (SOEP) für das jahr 2002 wurde folgende Longleichung geschätzt:

$$\ln(w_i) = \beta_0 + \beta_1 S_i + \beta_2 B_i + \beta_3 B_i^2 + \beta_4 F_i + \beta_5 F_i \cdot S_i + \beta_6 R_i + \varepsilon_i,$$

wobei:

- S_i: Jahre der Schulausbildung von Individuum i
- B_i: Jahre der Arbeitsmarkterfahrung von Individuum i
- F_i: Dummy Variable, die den Wert 1 annimmt, wenn Individuum i weiblich ist, und den Wert 0 sonst.
- R_i: Anzahl der Zigaretten,die Individuum i durchschnittlich pro Tag konsumiert.
- ε_i: normalverteilte Fehlerterm mit Mittelwert 0 und Varianz σ^2

Die Schätzergebnisse werden in folgender Tabelle zusammengefasst.

Variable	Koeffizient	Standardfehler
Konstante	6,516	0,049
Jahre der Schulausbildung	0,084	0,002
Arbeitsmarkterfahrung	0,031	0,003
Arbeitsmarkterfahrung^2	-0,004	0,001
Frau	-0,201	0,015
Frau · Jahre der Schulausbildung	-0,015	0,005
Zigaretten pro Tag	-0,024	0,010
R^2	0,301	
Beobachtungen	3.159	

a) Interpretieren Sie die in der Tabelle dargestellten Ergebnisse ökonomisch (d.h. auch ihre quantitative Dimension). Sind die Koeffizienten statistisch signifikant von Null verschieden?

b) Wie würden Sie vorgehen, wenn Sie testen wollten, ob die Variablen B_i und ${B_i}^2$ gemeinsam einen statistisch signifikanten Einfluss auf $ln(w_i)$ haben?

c) Aufgrund einer potenziellen Endogenität der Variable R_i wurde in einem zweiten Schritt eine Instrumentenvariablen-Regression geschätzt. Als Instrument diente dabei eine Dummy-Variable RJ_i, die den Wert 1 annimmt, wenn eine Person schon vor Erreichen des 16. Lebensjahrs regelmäßig geraucht hat, und den Wert 0 sonst. Die folgende Regression

$$R_i = \beta_0 + \delta_1 S_i + \delta_2 B_i + \delta_3 {B_i}^2 + \delta_4 F_i + \delta_5 F_i \cdot S_i + \delta_6 RJ_i + \nu_i$$

ergab für den Parameter der Variablen RJ_i einen Wert von $\hat{\delta}_6 = 0.424$ mit einem Standardfehler von 0.158. Warum denken Sie, dass der geschätzte

Koeffizient der Variablen R_i in Spalte (1) der obigen Tabelle nicht den kausalen Effekt des Zigarettenkonsums auf den Lohn angibt? Diskutieren Sie die Validität des verwendeten Instruments.

d) Wie würden Sie im gegebenen Kontext vorgehen um zu testen, ob R_i endogen ist?