

Übung 6: Rauchen und Schwangerschaft

Wir interessieren uns für den Einfluss des Rauchens während der Schwangerschaft, quantifiziert durch die Zahl der pro Tag von der Mutter gerauchten Zigaretten, auf die Gesundheit von Neugeborenen, gemessen durch deren Geburtsgewicht. Da aber auch viele andere Faktoren existieren, die einerseits vermutlich das Geburtsgewicht beeinflussen und andererseits mit dem Rauchverhalten korrelieren, sollten wir weitere erklärende Variablen berücksichtigen. Da Sie die Daten aus einer amerikanischen Studie erhalten haben, ist die Messeinheit für das Geburtsgewicht in Unzen gegeben.

cig: pro Tag Konsum von Zigaretten

bwgth: Geburtsgewicht des Neugeborenen (**b**aby **w**eight) in Unzen.

faminc: Familieneinkommen

male: Dummy Variable = 1 wenn Neugeborenes männlich ist

white: Dummy Variable = 1 wenn Neugeborenes weiss ist.

Verwenden Sie für diese Aufgabe die Datei Rauchen und Schwangerschaft.gdt auf Moodle.

1. Analyse der Daten.

i. Wie viele Frauen sind in der Stichprobe enthalten?

gretl Hauptfenster: Stichprobe/Zeige Status

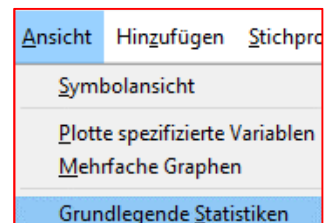


ii. Wie hoch ist der durchschnittliche Zigarettenkonsum pro Tag? Ist dieser Durchschnittswert eine repräsentativ die typische Frau aus der Stichprobe?

gretl: Ansicht/Grundlegende Statistiken → Auswahl: faminc, motheduc und cig

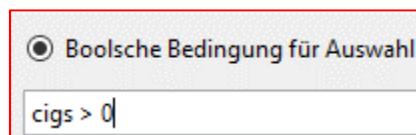
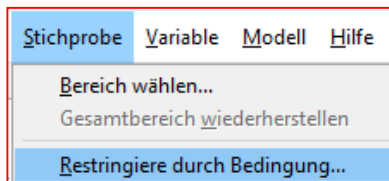
	arith. Mittel	Median	Minimum	Maximum
faminc	29,027	27,500	0,50000	65,000
motheduc	12,936	12,000	2,0000	18,000
cigs	2,0872	0,00000	0,00000	50,000

	Std. Abw.	Var'koeff.	Schiefe	Überwölbung
faminc	18,739	0,64559	0,61762	-0,52660
motheduc	2,3767	0,18373	-0,032120	0,64824
cigs	5,9727	2,8616	3,5604	14,934



iii. Wie viele Frauen Rauchen während der Schwangerschaft? Was ist der Anteil von Raucherinnen in der Stichprobe?

gretl Hauptfenster: Stichprobe/Restringere durch Bedingung/ Boolesche Bedingung: $cigs > 0$ → Dadurch werden die Raucherinnen beibehalten und die Nichtraucherinnen entfernt!



iv. Wie hoch ist der durchschnittliche Zigarettenkonsum pro Tag unter den Raucherinnen?

	arith. Mittel	Median	Minimum	Maximum
faminc	20,917	18,500	0,50000	65,000
motheduc	11,637	12,000	6,0000	18,000
cigs	13,665	10,000	1,0000	50,000

	Std. Abw.	Var'koeff.	Schiefe	Überwölbung
faminc	15,142	0,72392	1,0458	0,95217
motheduc	1,7753	0,15256	0,15604	1,6180
cigs	8,6909	0,63599	1,3020	2,5502

Ansicht	Hinzufügen	Stichprobe
Symbolansicht		
Plote spezifizierte Variablen		
Mehrfache Graphen		
Grundlegende Statistiken		

Hinweis: Die Stichprobe ist wieder auf den Gesamtbereich wiederherzustellen!

gretl Hauptfenster: Stichprobe/Gesamtbereich *wiederherstellen*

Stichprobe	Variable	Modell	Hilfe
Bereich wählen...			
Gesamtbereich wiederherstellen			

- v. Wie hoch ist der durchschnittliche Familieneinkommen? Vergleichen Sie zwischen der Stichprobe und Teilmenge der Raucherinnen.
- vi. Wie viele Neugeborene sind in der Stichprobe weiss?

Stichprobe	Variable	Modell	Hilfe
Bereich wählen...			
Gesamtbereich wiederherstellen			
Restringiere durch Bedingung...			

☐ Boolesche Bedingung für Auswahl eingeben:

☒ Benutze Dummy-Variabel: white

Stichprobe/Restringiere durch Bedingung/ Benutze Dummy Variable/white

→ Dadurch werden die nichtweisse Neugeborenen entfernt!

2. Welchen Einfluss erwarten Sie für die Variablen *cigs* und *faminc* auf das Geburtsgewicht des Neugeborenen (welche Vorzeichen für β_2 und β_3)? Begründen Sie Ihre Antwort.
3. Schätzen Sie das **Modell 1**: $bwght = \beta_1 + \beta_2 \text{cigs} + u$

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	119,772	0,572341	209,3	0,0000	***
cigs	-0,513772	0,0904909	-5,678	1,66e-08	***

Mittel d. abh. Var.	118,6996	Stdabw. d. abh. Var.	20,35396
Summe d. quad. Res.	561551,3	Stdfehler d. Regress.	20,12858
R-Quadrat	0,022729	Korrigiertes R-Quadrat	0,022024

4. Welche Korrelation erwarten Sie zwischen den Variablen *cigs* (Zig-Konsum) und *faminc* (Familieneinkommen)? Erklären Sie, warum die Korrelation positiv oder negativ sein könnte.
5. Analysieren Sie die **Korrelationsstruktur** zwischen den Variablen *bwght*, *cigs* und *faminc*

gretl Hauptfenster: Ansicht/Korrelationsmatrix → Variablen bwght, cigs und faminc auswählen

bwght	faminc	cigs
1,0000	0,1089	-0,1508
	1,0000	-0,1730
		1,0000

Ansicht	Hinzufügen	Stichprobe
Symbolansicht		
Plote spezifizierte Variablen		
Mehrfache Graphen		
Grundlegende Statistiken		
Korrelationsmatrix		

6. Ermitteln Sie die Korrelation zwischen *cigs* und *faminc* mittels einer **Regression**. Einmal für die gesamte Stichprobe, einmal für die Gruppe der Raucherinnen. Wie ändert sich diese Korrelation für diese **Teilmenge** aus der Stichprobe?

Regression für die **gesamte** Stichprobe:

Abhängige Variable: faminc					
	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	30,1598	0,524988	57,45	0,0000	***
cigs	-0,542928	0,0830042	-6,541	8,58e-011	***
Mittel d. abh. Var.	29,02666	Stdabw. d. abh. Var.	18,73928		
Summe d. quad. Res.	472475,2	Stdfehler d. Regress.	18,46324		
R-Quadrat	0,029945	Korrigiertes R-Quadrat	0,029245		

Regression für die **Gruppe der Raucherinnen**

gretl: Stichprobe/Restringere durch Bedingung/cigs > 0

Abhängige Variable: faminc					
	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	23,8077	1,93141	12,33	1,26e-026	***
cigs	-0,211509	0,119344	-1,772	0,0778	*
Mittel d. abh. Var.	20,91745	Stdabw. d. abh. Var.	15,14250		
Summe d. quad. Res.	47668,34	Stdfehler d. Regress.	15,06626		
R-Quadrat	0,014736	Korrigiertes R-Quadrat	0,010045		

Hinweis: Die Stichprobe ist wieder auf den Gesamtbereich wiederherzustellen!

Stichprobe	Variable	Modell	Hilfe
Bereich wählen...			
Gesamtbereich wiederherstellen			

7. Welchen Effekt hat vermutlich die Hinzunahme von *faminc* auf den geschätzten Regressionskoeffizienten b_2 (= b_{cigs})?

Hinweis: Benutzen Sie Ihr Ergebnis aus Frage 6

8. Schätzen Sie das **Modell 2**: $bwght = \beta_1 + \beta_2 \text{ cigs} + \beta_3 \text{ faminc} + u$

Abhängige Variable: bwght					
	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	116,974	1,04898	111,5	0,0000	***
cigs	-0,463408	0,0915768	-5,060	4,75e-07	***
faminc	0,0927647	0,0291879	3,178	0,0015	***
Mittel d. abh. Var.	118,6996	Stdabw. d. abh. Var.	20,35396		
Summe d. quad. Res.	557485,5	Stdfehler d. Regress.	20,06282		
R-Quadrat	0,029805	Korrigiertes R-Quadrat	0,028404		
F(2, 1385)	21,27392	P-Wert(F)	7,94e-10		
Log-Likelihood	-6130,414	Akaike-Kriterium	12266,83		
Schwarz-Kriterium	12282,54	Hannan-Quinn-Kriterium	12272,70		

9. Vermuten Sie, dass die Berücksichtigung dieser Dummy-Variable einen deutlichen Effekt auf b_{cigs} und b_{faminc} oder deren Standardfehler hat? Warum bzw. warum nicht? Überprüfen Sie Ihre Vermutung anschliessend.

10. Es soll nun die Dummy-Variable *male* als zusätzlicher Regressor hinzugefügt werden (Wert 1, wenn das Neugeborene männlich ist, 0 für weiblich).

Schätzen Sie das **Modell 3**: $bwght = \beta_1 + \beta_2cigs + \beta_3faminc + \beta_4male + u$

Abhängige Variable: bwght					
	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	115,228	1,20788	95,40	0,0000	***
cigs	-0,461046	0,0913378	-5,048	5,07e-07	***
faminc	0,0968798	0,0291453	3,324	0,0009	***
male	3,11397	1,07640	2,893	0,0039	***
Mittel d. abh. Var.	118,6996	Stdabw. d. abh. Var.	20,35396		
Summe d. quad. Res.	554134,6	Stdfehler d. Regress.	20,00965		
R-Quadrat	0,035636	Korrigiertes R-Quadrat	0,033546		
F(3, 1384)	17,04780	P-Wert (F)	7,10e-11		
Log-Likelihood	-6126,230	Akaike-Kriterium	12260,46		
Schwarz-Kriterium	12281,40	Hannan-Quinn-Kriterium	12268,29		

11. Interpretieren Sie b_{faminc} im Modell 3.

Hinweis: Das Geburtsgewicht *bwght* ist hier in *Unzen* angegeben (1 Unze = 28.35 Gramm), das Einkommen der Familie *faminc* ist in *\$1000* -Einheiten angegeben.

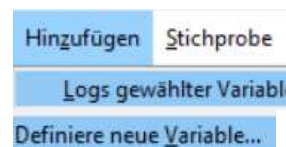
12. Schätzen Sie das Modell 3 mit dem Geburtsgewicht des Neugeborenen in *Gramm* ausgedrückt.

Modell 4: $bwghtgr = \beta_1^* + \beta_2^*cigs + \beta_3^*faminc + \beta_4^*male + u$

Hinweis: 1 Unze = 28.35 Gramm → Variable *bwghtgr* = *bwght* x 28.35

gretl Hauptfenster: Hinzufügen/ Definiere neue Variable/

bwghtgr = *bwght* x 28.35



Abhängige Variable: bwghtgr					
	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	3266,71	34,2434	95,40	0,0000	***
cigs	-13,0706	2,58943	-5,048	5,07e-07	***
faminc	2,74654	0,826268	3,324	0,0009	***
male	88,2810	30,5158	2,893	0,0039	***
Mittel d. abh. Var.	3365,133	Stdabw. d. abh. Var.	577,0349		
Summe d. quad. Res.	4,45e+08	Stdfehler d. Regress.	567,2737		
R-Quadrat	0,035636	Korrigiertes R-Quadrat	0,033546		
F(3, 1384)	17,04780	P-Wert (F)	7,10e-11		

13. Wie ist die Beziehung zwischen den Koeffizienten aus Modell 2 und 3.

14. Interpretieren Sie den Koeffizienten b_{faminc}

15. Durch Diskussionen mit anderen CAS-Teilnehmern haben Sie folgende Modelle zusammengestellt:

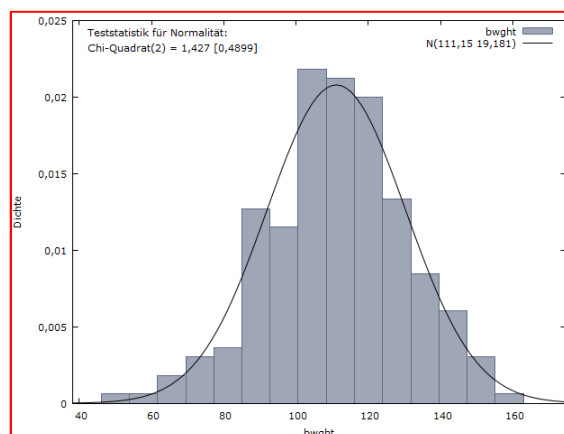
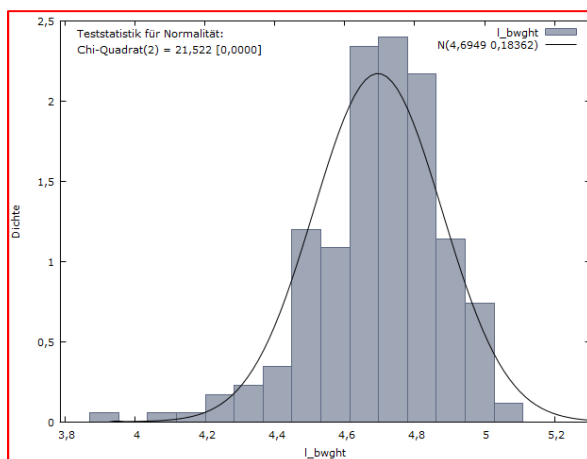
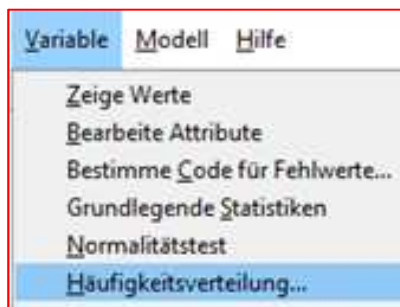
- $bwght = \beta_1 + \beta_2cigs + \beta_3 \ln(faminc) + \beta_4 male + u$
- $\ln(bwght) = \beta_1 + \beta_2cigs + \beta_3 \ln(faminc) + \beta_4 male + u$
- $\ln(bwght) = \beta_1 + \beta_2cigs + \beta_3 faminc + \beta_4 male + u$

Schätzen Sie diese Modelle und interpretieren Sie jeweils den Koeffizienten b_3

Schätzergebnisse:

- i. $bwght = 112.138 - 0.465cigs + 1.927 \ln(faminc) + 3.096 \text{ male}$
- ii. $\ln(bwght) = 4.703 - 0.00406cigs + 0.0169 \ln(faminc) + 0.0258 \text{ male}$
- iii. $\ln(bwght) = 4.729 - 0.0401cigs + 0.000878 \text{ faminc} + 0.0259 \text{ male}$

16. Erstellen Sie ein Histogramm von $\ln(bwght)$ und $bwght$. Welcher Unterschied ist zu vermerken?



17. Ersetzen Sie *faminc* durch *fatheduc* (Ausbildungsdauer des Vaters gemessen in Jahren). Schätzen Sie die folgenden Modelle und interpretieren Sie jeweils den Koeffizienten b_3 :

- i. $bwght = \beta_1 + \beta_2 cigs + \beta_3 \text{ fatheduc} + \beta_4 \text{ male} + u$
- ii. $bwght = \beta_1 + \beta_2 cigs + \beta_3 \ln(\text{fatheduc}) + \beta_4 \text{ male} + u$
- iii. $\ln(bwght) = \beta_1 + \beta_2 cigs + \beta_3 \ln(\text{fatheduc}) + \beta_4 \text{ male} + u$
- iv. $\ln(bwght) = \beta_1 + \beta_2 cigs + \beta_3 \text{ fatheduc} + \beta_4 \text{ male} + u$

Schätzergebnisse:

- i. $bwght = 113.260 - 0.571cigs + 0.411 \text{ fatheduc} + 3.568 \text{ male}$ lin-lin
- ii. $bwght = 106.528 - 0.574cigs + 4.772 \ln(\text{fatheduc}) + 3.524 \text{ male}$ lin-log
- iii. $\ln(bwght) = 4.664 - 0.005cigs + 0.0372 \ln(\text{fatheduc}) + 0.0313 \text{ male}$ log-log
- iv. $\ln(bwght) = 4.716 - 0.0049cigs + 0.0033 \text{ fatheduc} + 0.0317 \text{ male}$ log-lin

18. Schätzen Sie das **Modell 5**:

$$\text{bwght} = \beta_1 + \beta_2 \text{cigs} + \beta_3 \text{parity} + \beta_4 \text{faminc} + \beta_5 \text{motheduc} + \beta_6 \text{fatheduc} + u$$

Die Variable *parity* stellt die Reihenfolge des Neugeborenen unter den Familienkindern. Wenn *parity* = 3 bedeutet es, dass das erfasste Neugeborene das dritte Kind der Frau ist.

Modell 14: KQ, benutze die Beobachtungen 1-1388 (n = 1191)
Fehlende oder unvollständige Beobachtungen entfernt: 197
Abhängige Variable: bwght

	Koeffizient	Std.-fehler	t-Quotient	p-Wert
const	114,524	3,72845	30,72	6,87e-153 ***
cigs	-0,595936	0,110348	-5,401	8,02e-08 ***
parity	1,78760	0,659406	2,711	0,0068 ***
faminc	0,0560414	0,0365616	1,533	0,1256
motheduc	-0,370450	0,319855	-1,158	0,2470
fatheduc	0,472394	0,282643	1,671	0,0949 *
Mittel d. abh. Var.	119,5298	Stdabw. d. abh. Var.	20,14124	
Summe d. quad. Res.	464041,1	Stdfehler d. Regress.	19,78878	
R-Quadrat	0,038748	Korrigiertes R-Quadrat	0,034692	
F(5, 1185)	9,553500	P-Wert (F)	5,99e-09	
Log-Likelihood	-5242,220	Akaike-Kriterium	10496,44	
Schwarz-Kriterium	10526,94	Hannan-Quinn-Kriterium	10507,93	

Abgesehen von Konstante war p-Wert am höchsten für Variable 6 (motheduc)

- Warum reduziert gretl hier jeweils die Zahl der einbezogenen Familien bei diesen Schätzungen (Frage 13)? Könnte das Konsequenzen bzgl. der Repräsentativität der "selektierten" Familien haben?
- Spielt die Reihenfolge des Neugeborenen eine Rolle für das Geburtsgewicht? Interpretieren Sie den Koeffizienten b_3 .
- Sind alle Steigungskoeffizienten **gemeinsam** signifikant?

19. Testen Sie die **Nullhypothese** im Modell 5, dass die **Elternausbildung** keinen Effekt auf das Gewicht des Neugeborenen hat.

i. Mittels gretl Test

gretl: Tests / Variable weglassen → Schätze reduziertes Modell → interpretieren Sie den p-Wert.

Nullhypothese: Die Regressionskoeffizienten sind Null für die Variablen **motheduc, fatheduc**
Teststatistik: F(2, 1185) = 1,43727, p-Wert 0,23799

ii. Bestimmen Sie den kritischen Wert F_c mittels gretl. Was ist Ihre **Schlussfolgerung**?

gretl Hauptfenster: Werkzeuge / Statistische Tabellen / F / rechtsseitige Wahrscheinlichkeit = 0.05

iii. Berechnen Sie den F-Wert mittels **Bestimmtheitsmass R^2** durch eigene Schätzung des restringierten Modells.

Hinweis: Die Schätzung des restringierten Modells sollte mit den gleichen Daten wie im Modell 6 erfolgen. Deshalb muss eine Proxy-Variable $\text{fath_moth} = \text{fatheduc} \times \text{motheduc}$ gebildet. Die Stichprobe kann mittels diese Proxy-Variable reduziert werden.

gretl: Hinzufügen / Definiere neue Variable: $\text{fath_moth} = \text{fatheduc} \times \text{motheduc}$

gretl: Stichprobe / Restringiere durch Bedingung / fath_moth > 0

☒ Boolesche Bedingung für Auswahl eingeben:

☐ Benutze Dummy-Variable: male

197 Beobachtungen entfernt

Modell 17: KQ, benutze die Beobachtungen 1-1191
Abhängige Variable: bwght

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	114,524	3,72845	30,72	6,87e-153	***
cigs	-0,595936	0,110348	-5,401	8,02e-08	***
parity	1,78760	0,659406	2,711	0,0068	***
faminc	0,0560414	0,0365616	1,533	0,1256	
motheduc	-0,370450	0,319855	-1,158	0,2470	
fatheduc	0,472394	0,282643	1,671	0,0949	*
Mittel d. abh. Var.	119,5298	Stdabw. d. abh. Var.	20,14124		
Summe d. quad. Res.	464041,1	Stdfehler d. Regress.	19,78878		
R-Quadrat	0,038748	Korrigiertes R-Quadrat	0,034692		

Regression mit nur 1191
Beobachtungen → 197
Beobachtungen wurden entfernt!

Unter Berücksichtigung **aller** Beobachtungen ist das R^2 anders! Deshalb ist die Benutzung der Proxy-Variable wichtig!

20. Schätzen Sie das **Modell 6**:

$$\ln(\text{bwght}) = \beta_1 + \beta_2 \text{cigs} + \beta_3 \ln(\text{faminc}) + \beta_4 \text{parity} + \beta_5 \text{male} + \beta_6 \text{white} + u$$

Modell 14: KQ, benutze die Beobachtungen 1-1388
Abhängige Variable: l_bwght

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	4,65771	0,0221653	210,1	0,0000	***
cigs	-0,00435015	0,000851842	-5,107	3,73e-07	***
l_faminc	0,00927740	0,00593081	1,564	0,1180	
parity	0,0159828	0,00563877	2,834	0,0047	***
male	0,0265458	0,0100295	2,647	0,0082	***
white	0,0547875	0,0130518	4,198	2,87e-05	***
Mittel d. abh. Var.	4,760031	Stdabw. d. abh. Var.	0,190662		
Summe d. quad. Res.	48,04116	Stdfehler d. Regress.	0,186446		
R-Quadrat	0,047187	Korrigiertes R-Quadrat	0,043740		
F(5, 1382)	13,68835	P-Wert(F)	4,58e-13		
Log-Likelihood	364,8246	Akaike-Kriterium	-717,6492		
Schwarz-Kriterium	-686,2355	Hannan-Quinn-Kriterium	-705,9010		

- i. Was ist der Effekt auf das Geburtsgewicht, wenn die Mutter 10 Zigaretten pro Tag mehr raucht?
- ii. Wie viel mehr Geburtsgewicht weist ein **männliches** Neugeborenes gegenüber einem **Weiblichen** auf, *ceteris paribus*? Ist der Koeffizient β_5 signifikant auf 5%-Niveau?
- iii. Wie viel mehr Geburtsgewicht weist ein **weisses** Neugeborenes gegenüber der Referenzgruppe auf, *ceteris paribus*? Ist der Koeffizient β_6 signifikant auf 5%-Niveau?

21. Schätzen Sie das **Modell 7**:

$$\ln(\text{bwght}) = \beta_1 + \beta_2 \text{cigs} + \beta_3 \ln(\text{faminc}) + \beta_4 \text{parity} + \beta_5 \text{male} + \beta_6 \text{white} + \beta_7 \text{motheduc} + \beta_8 \text{fatheduc} + u$$

- i. Was ist die Auswirkung eines zusätzlichen Ausbildungsjahres der Mutter auf das Geburtsgewicht?

Modell 13: KQ, benutze die Beobachtungen 1-1388 (n = 1191)
 Fehlende oder unvollständige Beobachtungen entfernt: 197
 Abhängige Variable: l_bwght

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	4,65267	0,0381545	121,9	0,0000	***
cigs	-0,00521438	0,00102675	-5,079	4,42e-07	***
l_faminc	0,0110315	0,00854044	1,292	0,1967	
parity	0,0172014	0,00613350	2,804	0,0051	***
male	0,0341430	0,0107022	3,190	0,0015	***
white	0,0453991	0,0150870	3,009	0,0027	***
motheduc	-0,00297633	0,00297307	-1,001	0,3170	
fatheduc	0,00327634	0,00260843	1,256	0,2093	
Mittel d. abh. Var.	4,767536	Stdabw. d. abh. Var.	0,188013		
Summe d. quad. Res.	39,99114	Stdfehler d. Regress.	0,183861		
R-Quadrat	0,049303	Korrigiertes R-Quadrat	0,043678		
F(7, 1183)	8,764331	P-Wert (F)	1,55e-10		
Log-Likelihood	331,1061	Akaike-Kriterium	-646,2122		
Schwarz-Kriterium	-605,5518	Hannan-Quinn-Kriterium	-630,8901		

22. Schätzen Sie das [Modell 8](#):

$$\text{bwght} = \beta_1 + \beta_2 \text{cigs} + \beta_3 \ln(\text{faminc}) + \beta_4 \text{parity} + \beta_5 \text{male} + \beta_6 \text{white} + \beta_7 \text{motheduc} + \beta_8 \text{fatheduc} + u$$

- i. Wie viel mehr Geburtsgewicht weist ein [männliches](#) Neugeborenes gegenüber der Referenzgruppe auf, ceteris paribus? Ist der Koeffizient β_5 signifikant auf dem 5%-Signifikanzniveau?

Fehlende oder unvollständige Beobachtungen entfernt: 197
 Abhängige Variable: bwght

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	106,538	4,07630	26,14	3,14e-119	***
cigs	-0,597376	0,109695	-5,446	6,27e-08	***
l_faminc	1,22061	0,912434	1,338	0,1812	
parity	1,91752	0,655284	2,926	0,0035	***
male	3,82465	1,14339	3,345	0,0008	***
white	4,63746	1,61185	2,877	0,0041	***
motheduc	-0,336755	0,317634	-1,060	0,2893	
fatheduc	0,415149	0,278676	1,490	0,1366	
Mittel d. abh. Var.	119,5298	Stdabw. d. abh. Var.	20,14124		
Summe d. quad. Res.	456463,7	Stdfehler d. Regress.	19,64313		
R-Quadrat	0,054445	Korrigiertes R-Quadrat	0,048850		
F(7, 1183)	9,730940	P-Wert (F)	7,99e-12		
Log-Likelihood	-5232,416	Akaike-Kriterium	10480,83		
Schwarz-Kriterium	10521,49	Hannan-Quinn-Kriterium	10496,15		

23. Antworten Sie auf diese Fragen mittels einer [Regression](#).

- ii. Wie viel wiegt ein weibliches Neugeborenes im Durchschnitt in [Kg](#)?

Abhängige Variable: bwght

	Koeffizient	Std.-fehler	t-Quotient	p-Wert	
const	117,167	0,787514	148,8	0,0000	***
male	2,94235	1,09115	2,697	0,0071	***
Mittel d. abh. Var.	118,6996	Stdabw. d. abh. Var.	20,35396		
Summe d. quad. Res.	571612,8	Stdfehler d. Regress.	20,30810		
R-Quadrat	0,005219	Korrigiertes R-Quadrat	0,004501		
F(1, 1386)	7,271438	P-Wert (F)	0,007091		
Log-Likelihood	-6147,782	Akaike-Kriterium	12299,56		
Schwarz-Kriterium	12310,03	Hannan-Quinn-Kriterium	12303,48		

- iii. Wie viel mehr Geburtsgewicht in Gramm weist ein männliches Neugeborenes gegenüber einem Weiblichen auf?
- iv. Warum ist der Steigungskoeffizient kleiner als β_5 im Modell 7

24. Welches Modell würden Sie vorziehen? Begründen Sie Ihre Antwort.

Zusammenstellung der zu vergleichenden Modelle mit den entsprechenden Kriterien.

$$\widehat{\text{Modell 2: bwght}} = 116.97 - 0.463cigs + 0.0927faminc$$

$$\widehat{\text{Modell 3: bwght}} = 115.228 - 0.461cigs + 0.09687faminc + 3.114male$$

$$\widehat{\text{Modell 5: bwght}} = 114.524 - 0.596cigs + 1.787parity + 0.0560faminc - 0.37motheduc + 0.472fatheduc$$

$$\widehat{\text{Modell 6: lnbwght}} = 4.657 - 0.00435cigs + 0.00927lnfaminc + 0.0159parity + 0.0265male + 0.0547white$$

$$\widehat{\text{Modell 7: lnbwght}} = 4.657 - 0.00521cigs + 0.0172parity + 0.0117lnfaminc + 0.0341male + 0.045white - 0.0029motheduc + 0.00327fatheduc$$

$$\widehat{\text{Modell 8: bwght}} = 106.53 - 0.5973cigs + 1.917parity + 1.22lnfaminc + 3.82male + 4.63white - 0.336motheduc + 0.415fatheduc$$

Modell	2	3	5	6	7	8
Abh. Variable	bwght	bwght	bwght	lnbwght	lnbwght	bwght
#Regressoren	3	4	6	6	8	8
\bar{R}^2	0.028	0.0327	0.0346	0.0437	0.0436	0.0488
Akaike	12266	12261.6	10496	-717.64	-646.21	10480
SIC	12282		10526	-686.2	-605.5	10521