Bachelorprüfung WS 2020/2021

Fach: Praxis der empirischen Wirtschaftsforschung

Prüferin: Prof. Regina T. Riphahn, Ph.D.

Vorbemerkungen:

Anzahl der Aufgaben:	Die Klausur besteht aus 3 Aufgaben, die alle bearbeitet werden müssen. **Es wird nur der Lösungsbogen eingesammelt.** Angaben auf dem Aufgabenzettel werden nicht gewertet.
Bewertung:	Es können maximal 60 Punkte erworben werden. Die maximale Punktzahl ist für jede Aufgabe in Klammern angegeben. Sie entspricht der für die Aufgabe empfohlenen Bearbeitungszeit in Minuten.
Erlaubte Hilfsmittel:	• Formelsammlung (ist der Klausur beigefügt) • Tabellen der statistischen Verteilungen (sind der Klausur beigefügt) • Taschenrechner • Fremdwörterbuch
Wichtige Hinweise:	• Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den gesuchten Wert der Freiheitsgrade nicht ausweisen, machen Sie dies kenntlich und verwenden Sie den nächstgelegenen Wert. • Sollte es vorkommen, dass bei einer Berechnung eine erforderliche Information fehlt, machen Sie dies kenntlich und treffen Sie für den fehlenden Wert eine plausible Annahme.

Aufgabe 1: **[18 Punkte]**

Sie möchten den durchschnittlichen Kraftstoffverbrauch pro Kilometer vorhersagen. Der Datensatz enthält folgende Variablen für 345 Autos:

$lp100_i$	Kraftstoffverbrauch pro 100 Kilometer in Liter
$weight_i$	Gewicht des Autos in 100 Kilogramm.
hp	Motorleistung in Pferdestärke (PS)
$foreign_i$	=1, wenn das Auto von einem ausländischen Hersteller stammt, =0 sonst

Sie schätzen das folgende Modell:

$$lp100_i = \beta_0 + \beta_1 ln(weight)_i + \varepsilon_i \tag{1}$$

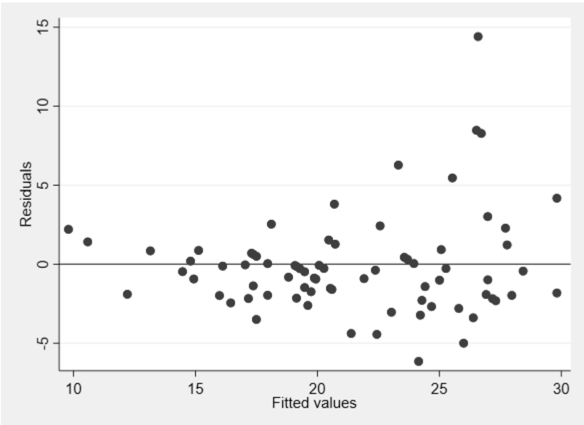
und erhalten die Koeffizientenschätzer: $\hat{\beta}_0 = 4,923$ und $\hat{\beta}_1 = 67,8$ mit $se(\hat{\beta}_0) = 2,267$ und $se(\hat{\beta}_1) = 21,5$

Runden Sie alle Zahlenangaben auf die <u>*dritte*</u> *Nachkommastelle.*

a) Interpretieren Sie den Schätzer $\hat{\beta}_1$ inhaltlich und testen Sie seine statistische Signifikanz am 1% Niveau. (4 Punkte)

b) Ein Kommilitone macht Sie darauf aufmerksam, dass der Koeffizientenschätzer von $ln(weight)$ überschätzt sein könnte. Sie vermuten daraufhin, dass eine relevante Variable ausgelassen wurde und schätzen die Gleichung erneut mit der Variable hp. Welche Bedingungen muss die Variable hp erfüllen, damit das Auslassen zu einer Überschätzung von $\hat{\beta}_1$ führt? Verdeutlichen Sie jeweils am Beispiel von hp die beiden Bedingungen, die dafür erfüllt sein müssen. (3 Punkte)

c) Eine Betrachtung der Residuen kann Aufschluss über die Störterme geben. Betrachten Sie folgende Grafik. Welche der Gauß-Markov Annahme könnte verletzt sein? Begründen Sie Ihre Antwort kurz. (2 Punkte)



d) Als nächstes schätzen Sie das Modell

$$lp100_i = \beta_0 + \beta_1 weight_i + \beta_2 hp_i + \varepsilon_i \tag{2}$$

Sie vermuten jedoch, dass sich die Parameter des Modells (2) für deutsche und ausländische Hersteller unterscheiden. Erläutern Sie kurz das Vorgehen und die Entscheidungslogik des Chow-Tests auf Strukturbruch, den Sie mittels der Variable $foreign$ durchführen können. (3 Punkte)

e) Um die Vermutung aus Aufgabe d) zu überprüfen, führen Sie nun einen Chow-Test auf Strukturbruch für das Modell (2) am 1%-Niveau durch. Geben Sie Hypothesen, Teststatistik, kritischen Wert und Ihre Testentscheidung an. (6 Punkte)
Hinweise: $SSR_{pooled} = 299,567$, $SSR_1 = 135,674$ (für $foreign = 0$), $SSR_2 = 155,210$ (für $foreign = 1$)

Aufgabe 2: **[20 Punkte]**

Sie wollen die Determinanten des monatlichen Bruttoeinkommens untersuchen. Ihnen liegen folgende Variablen für 536 Personen vor:

$income_i$	monatliches Bruttoeinkommen in Euro
$female_i$	=1, wenn Person weiblich; =0,wenn männlich
$educ_i$	Bildung in Jahren
age_i	Alter in Jahren
$experience_i$	Berufserfahrung in Jahren

Sie schätzen das folgende Modell:

$$income_i = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 educ_i + \beta_4 experience_i + \beta_5 experience \cdot educ_i + u_i$$

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	0,287(a)	0,136	???	0,415

Koeffizienten[a]

Modell	Nicht standardisierte Koeffizienten: RegressionskoeffizientB	Standardfehler	T	Signifikanz
$(Konstante)$	940,473	64,427	14,5970	0.000
$female$	-167,352	57,195	-2,926	0,012
age	20,512	4,483	4,576	0,000
$educ$	24,398	7,190	3,161	0,007
$experience$	31,454	12.364	2,544	0,093
$experience \cdot educ$	12,955	1,872	6,920	0,002

a. Abhängige Variable: *income*

Runden Sie alle Zahlenangaben auf die dritte Nachkommastelle.

a) Interpretieren Sie den geschätzten Koeffizienten von *female* inhaltlich. Ist der Effekt statistisch signifikant? (2 Punkte)

b) Interpretieren Sie das R^2 der Schätzung und berechnen Sie das korrigierte Bestimmtheitsmaß $\bar{R}^2$. (3 Punkte)

c) Welchen Wert würde der Koeffizient von *experience* (β_4) annehmen, wenn man Berufserfahrung statt in Jahren in 10 Jahren misst? (2 Punkte)

d) Berechnen Sie den marginalen Effekt der Berufserfahrung für eine Person mit 15 Jahren Bildung. Erläutern Sie das Ergebnis. (2 Punkte)

e) Sie vermuten, dass der Alterseffekt vom Geschlecht abhängt. Sie schätzen folgendes Modell:

$$income_i = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 female_i \cdot age_i + u_i$$

Skizzieren Sie graphisch den Zusammenhang zwischen *age* und *income* für Frauen und Männer. Beschriften Sie die Achsen, die Achsenabschnitte und die Steigungen der Geraden. *Hinweis:* $\hat{\beta}_0 > 0$, $\hat{\beta}_1 < 0$, $\hat{\beta}_2 > 0$, $\hat{\beta}_3 < 0$. (4 Punkte)

f) Sie vermuten, dass der Einfluss des Alters auf das monatliche Einkommen nicht linear verläuft. Sie nehmen daher das quadrierte Alter mit in Ihr Modell auf und schätzen folgendes Modell:

$$income_i = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 age^2{}_i + \beta_4 educ_i + \beta_5 experience_i + \beta_6 experience \cdot educ_i + u_i$$

Berechnen Sie das Alter, welches das geschätzte monatliche Einkommen maximiert.
Hinweis: $\hat{\beta}_2 = 20,512$ *und* $\hat{\beta}_3 = -0,249$. (2 Punkte)

g) Statt des Modells, welches das Alter in Jahren enthält, entscheiden Sie sich für ein *neues* Modell. Sie bilden insgesamt 5 Dummy Variablen für einzelne Alterskategorien und nehmen alle in Ihr neues Modell mit auf. Welches Problem entsteht hierbei? Wie lässt es sich lösen? (3 Punkte)

h) Anstatt des monatlichen Bruttoeinkommens als abhängige Variable wählen Sie nun eine Dummy-Variable als abhängige Variable. Nennen Sie zwei Nachteile des linearen Wahrscheinlichkeitsmodells. (2 Punkte)

Aufgabe 3: **[22 Punkte]**

Sie interessieren sich dafür, ob die Schulklassengröße einen Einfluss auf die Löhne im Erwachsenenalter hat. Sie verfügen über einen US-amerikanischen Datensatz mit den folgenden Variablen für 764 Schüler:

$lohn_i$	Stundenlohn von Person i im Alter 27 (in US$)
$groß_i$	=1, wenn Person i in einer großen Klasse war (> 17 SchülerInnen); =0 sonst
$frau_i$	=1, wenn Frau; =0 sonst
$eltern_i$	Gesamteinkommen der Eltern zum Zeitpunkt der Einschulung von Person i (in 1000 US$)

Sie schätzen folgendes lineares Regressionsmodell mit SPSS:

$$log(lohn_i) = \beta_0 + \beta_1 groß_i + \beta_2 frau_i + \beta_3 eltern_i + u_i$$

Koeffizienten[a]

	Nicht standardisierte Koeffizienten			
Modell	RegressionskoeffizientB	Standardfehler	T	Signifikanz
$(Konstante)$	6,013	0,088	68,32	0,000
groß	-0,065	0,016	???	???
$frau$	-0,362	0,063	???	???
$eltern$	0,118	0,029	4,07	0,000

a. Abhängige Variable: $log(lohn)$

Das R^2 der Schätzung beträgt $0,234$.
Runden Sie alle Zahlenangaben auf die dritte Nachkommastelle.

a) Interpretieren Sie den geschätzten Koeffizienten von $eltern$ sowohl inhaltlich als auch statistisch. (2 Punkte)

b) Berechnen und interpretieren Sie inhaltlich den genauen (!) Effekt der Variable *frau* auf den Stundenlohn. (2 Punkte)

c) Testen Sie auf dem 5%-Signifikanzniveau, ob die Koeffizienten der Variablen *groß* und $frau$ gemeinsam signifikant sind. Bei einer erneuten Schätzung des Modells ohne diese beiden Variablen erhalten Sie ein R^2 von 0,186. Geben Sie Null- und Alternativhypothese, Freiheitsgrade, Teststatistik, kritischen Wert und Testentscheidung an. (5 Punkte)

d) Berechnen und interpretieren Sie das 99%-Konfidenzintervall für den geschätzten Koeffizienten der Variable $frau$. Gehen Sie darauf ein, ob der Koeffizient statistisch signifikant von Null verschieden ist. (4 Punkte)

e) Testen Sie, ob die Variable *groß* einen signifikant negativen Zusammenhang mit dem späteren Einkommen hat. Geben Sie ein konkretes Testverfahren, Null- und Alternativhypothese, Teststatistik, Freiheitsgrade, kritischen Wert und Ihre Testentscheidung für das 5%-Signifikanzniveau an. (5 Punkte)

f) Wie muss ein einfaches lineares Modell spezifiziert sein, damit der geschätzte Steigungsparameter i) eine Semielastizität und ii) eine Elastizität angibt? (2 Punkte)

g) Welchen Effekt hat Heteroskedastie auf Unverzerrtheit und Effizienz des KQ-Schätzers? (2 Punkte)

Formelsammlung – Praxis der empirischen Wirtschaftsforschung

Kapitel 1:

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}x_i(y_i-\bar{y})$$

$$=\sum_{i=1}^{n}(x_i-\bar{x})y_i$$

$$=\sum_{i=1}^{n}x_iy_i-n\bar{x}\cdot\bar{y}$$

$$E\left(\sum_{i=1}^{n}a_iX_i\right)=\sum_{i=1}^{n}a_iE(X_i)$$

$$E\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}E(X_i)$$

$Var(aX + bY) = a^2\,Var(X) + b^2\,Var(Y) + 2ab\,Cov(X,Y)$

Für identisch und unabhängig verteilte Zufallsvariablen Y_i:

$$\bar{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$$

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i-\bar{Y}_i\right)^2$$

Kapitel 2:

$y_i = \beta_0 + \beta_1 x_i + u_i$

$$\hat{\beta}_0=\bar{y}-\hat{\beta}_1\bar{x}$$

$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$\sum_{i=1}^{n}\hat{u}_i=0 \qquad \sum_{i=1}^{n}x_i\hat{u}_i=0$$

$$SST\equiv\sum_{i=1}^{n}(y_i-\bar{y})^2$$

$$SSE\equiv\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$$

$$SSR\equiv\sum_{i=1}^{n}\hat{u}_i^2=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$$

$$R^2=\frac{SSE}{SST}=1-\frac{SSR}{SST}\ ,\ 0\le R^2\le 1$$

$$E\left(\hat{\beta}_0\right)=\beta_0 \qquad E\left(\hat{\beta}_1\right)=\beta_1$$

$$Var(\hat{\beta}_1)=\frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{\sigma^2}{SST_x}$$

$$Var(\hat{\beta}_0)=\frac{\sigma^2\,\frac{1}{n}\sum_{i=1}^{n}x_i^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$\hat{\sigma}^2=\frac{1}{(n-2)}\sum_{i=1}^{n}\hat{u}_i^2=\frac{SSR}{(n-2)}$$

Regression durch den Ursprung:

$$\tilde{\beta}_1=\frac{\sum_{i=1}^{n}x_iy_i}{\sum_{i=1}^{n}x_i^2}$$

Kapitel 3:

$$R^2=\frac{SSE}{SST}=\frac{\left(\sum_{i=1}^{n}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})\right)^2}{\left(\sum_{i=1}^{n}(y_i-\bar{y})^2\right)\left(\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2\right)}$$

Wenn $y=\beta_0+\beta_1\,x_1+\beta_2\,x_2+u$

und $\tilde{y}=\tilde{\beta}_0+\tilde{\beta}_1\,x_1$

dann $\tilde{\beta}_1=\hat{\beta}_1+\hat{\beta}_2\,\tilde{\delta}_1$ mit $x_2=\tilde{\delta}_0+\tilde{\delta}_1x_1$

Allgemein für j = 1, 2, …, k:

$$Var\left(\hat{\beta}_j\right)=\frac{\sigma^2}{SST_j\left(1-R_j^2\right)}$$

$$SST_j=\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2$$

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{n}\hat{u}_i^2}{n-k-1}=\frac{SSR}{n-k-1}$$

$$se(\hat{\beta}_j)=\frac{\hat{\sigma}}{\left[SST_j(1-R_j^2)\right]^{\frac{1}{2}}}$$

MLR.1: Modell der Grundgesamtheit
$y = \beta_0 + \beta_1 x_i + \beta_2 x_2 + \ldots + \beta_k x_k + u$

MLR.2: Zufallsstichprobe der Größe n folgt dem Bevölkerungsmodell.

MLR.3: Keine unabhängige Variable ist konstant. Keine perfekte Kollinearität.

MLR.4: $E\left(u\,|\,x_1,x_2,...,x_k\right)=0$

MLR.5: $Var\left(u\,|\,x_1,x_2,...,x_k\right)=\sigma^2$

MLR.6: u ist von x_1, x_2, …, x_k unabhängig und $u \sim \text{Normal}(0, \sigma^2)$.

Kapitel 4:

$$(\hat{\beta}_j - \beta_j)\big/\mathrm{se}(\hat{\beta}_j) \sim t_{n-k-1}$$

$$-t_{\frac{\alpha}{2},n-k-1} \le \frac{\hat{\beta}_j - \beta_j}{\mathrm{se}(\hat{\beta}_j)} \le t_{\frac{\alpha}{2},n-k-1}$$

$$\hat{\beta}_j - c \cdot \mathrm{se}(\hat{\beta}_j) \le \beta_j \le \hat{\beta}_j + c \cdot \mathrm{se}(\hat{\beta}_j)$$

$$F \equiv \frac{(SSR_r - SSR_u)/q}{SSR_u/(n-k-1)}$$

$$F = \frac{(R_u^2 - R_r^2)/q}{(1-R_u^2)/(n-k-1)}$$

Kapitel 5:

$$\lim_{n\to\infty} P(\left|\hat{\beta}_1 - \beta_1\right| > \varepsilon) \to 0$$

$$\mathrm{plim}\left(\hat{\beta}_1\right) = \beta_1$$

Kapitel 6:

Standardisierung:

$$\frac{y_i - \bar{y}}{\hat{\sigma}_y} = \hat{\beta}_1 \left(\frac{\hat{\sigma}_1}{\hat{\sigma}_y}\right)\left(\frac{x_{i1} - \bar{x}_1}{\hat{\sigma}_1}\right) + \hat{\beta}_2 \left(\frac{\hat{\sigma}_2}{\hat{\sigma}_y}\right)\left(\frac{x_{i2} - \bar{x}_2}{\hat{\sigma}_2}\right)$$
$$+ \ldots + \hat{\beta}_k \left(\frac{\hat{\sigma}_k}{\hat{\sigma}_y}\right)\left(\frac{x_{ik} - \bar{x}_k}{\hat{\sigma}_k}\right) + \frac{\hat{u}_i}{\hat{\sigma}_y}$$

Semielastizität:
$$\%\Delta\hat{y} = 100 \cdot [\exp(\hat{\beta}_j \Delta x_j) - 1]$$

$$\bar{R}^2 = 1 - \frac{SSR/(n-k-1)}{SST/(n-1)} = 1 - \frac{\hat{\sigma}^2}{SST/(n-1)}$$

$$\bar{R}^2 = 1 - (1-R^2)\frac{n-1}{n-k-1}$$

$$P[\hat{y}^0 - t_{\frac{\alpha}{2},n-k-1} \cdot \mathrm{se}(\hat{e}^0) \le y^0 \le \hat{y}^0 + t_{\frac{\alpha}{2},n-k-1} \cdot \mathrm{se}(\hat{e}^0)] = 1\text{-}\alpha$$

$$\widehat{\log y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 \ldots + \hat{\beta}_k x_k$$
$$E(y\,|\,\mathbf{x}) = \exp(\sigma^2/2) \cdot \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k)$$

Kapitel 7:

Regression nach Gruppen
- Modell gepoolt: $y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$

Chow-Test (mit $SSR_P = SSR_{\text{gepooltes Modell}}$):

$$F = \frac{(SSR_P - (SSR_1 + SSR_2))/(k+1)}{(SSR_1 + SSR_2)/\left(n - 2(k+1)\right)}$$

TABLE G.2

Critical Values of the t Distribution

		Significance Level				
1-Tailed:		**.10**	**.05**	**.025**	**.01**	**.005**
2-Tailed:		**.20**	**.10**	**.05**	**.02**	**.01**
Degrees of Freedom	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.499
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	30	1.310	1.697	2.042	2.457	2.750
	40	1.303	1.684	2.021	2.423	2.704
	60	1.296	1.671	2.000	2.390	2.660
	90	1.291	1.662	1.987	2.368	2.632
	120	1.289	1.658	1.980	2.358	2.617
	∞	1.282	1.645	1.960	2.326	2.576

Examples: The 1% critical value for a one-tailed test with 25 *df* is 2.485. The 5% critical value for a two-tailed test with large (> 120) *df* is 1.96.

Source: This table was generated using the Stata® function invttail.

TABLE G.3b

5% Critical Values of the F Distribution

	Numerator Degrees of Freedom									
Denominator Degrees of Freedom	1	2	3	4	5	6	7	8	9	10
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Example: The 5% critical value for numerator $df = 4$ and large denominator df (∞) is 2.37.
Source: This table was generated using the Stata® function invFtail.

TABLE G.3c

1% Critical Values of the *F* Distribution

Denominator Degrees of Freedom	Numerator Degrees of Freedom: 1	2	3	4	5	6	7	8	9	10
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63
90	6.93	4.85	4.01	3.54	3.23	3.01	2.84	2.72	2.61	2.52
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32

Example: The 1% critical value for numerator $df = 3$ and denominator $df = 60$ is 4.13.

Source: This table was generated using the Stata® function invFtail.