Aufgabe 1: **[29 Punkte]**

Sie interessieren sich für Faktoren, die Treue oder Untreue in Partnerschaften bedingen. Ihnen stehen Befragungsdaten von 601 verheirateten Männern und Frauen zur Verfügung, die Angaben über mögliche Affären gemacht haben. Im Einzelnen enthält Ihr Datensatz folgende Informationen:

naffairs	Anzahl Affären im letzten Jahr
male	Geschlechtsdummy (1= männlich, 0= weiblich)
age	Alter in Jahren
educ	Schulbildung in Jahren
kids	Anzahl Kinder in der Ehe
yrsmarr	bisherige Dauer der Ehe in Jahren
ratemarr	Zufriedenheit mit der Ehe (5= sehr glücklich, …, 1= sehr unglücklich)

Sie schätzen mittels der KQ-Methode folgendes Modell in R:

$$naffairs = \beta_0 + \beta_1 \cdot male + \beta_2 \cdot age + \beta_3 \cdot educ + \beta_4 \cdot kids + \beta_5 \cdot yrsmarr + \beta_6 \cdot yrsmarr^2 + \varepsilon$$

```
Call:
lm(formula = naffairs ~ male + age + educ + kids + yrsmarr + yrsmarr2)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7674 -1.8086 -1.0259 -0.3467 11.5886

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.507097   1.083909   1.390   0.1649
male         0.214307   0.297387   0.721   0.4714
age         -0.044335   0.024072  -1.842   0.0660
educ        -0.019001       ?     -0.314   0.7540
kids        -0.232089   0.383553     ?     0.5453
yrsmarr      0.306197   0.127313   2.405   0.0165
yrsmarr2    -0.007598   0.007134  -1.065   0.2873
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.242 on 594 degrees of freedom
Multiple R-squared: ?,    Adjusted R-squared: 0.03437
F-statistic:  4.56 on ? and ? DF,  p-value: 0.000155
```

a) Bestimmen Sie unter Angabe des Rechenwegs *(9 Punkte)*

(1) den Standardfehler von b_3,

- $b_3/se(b_3) = t_{b3} <=> se(b_3) = b_3/t_{b3}$
- $se(b_3) = -\frac{0{,}019}{-0{,}314} = 0{,}061$

(2) den t-Wert von b_4,

- $b_4/se(b_4) = t_{b4}$
- t_{b4}=-0,232/0,384=-0,605

(3) die Freiheitsgrade des Tests auf Gesamtsignifikanz des Modells,

- dF_1=K-1=6
- dF_2=T-K=601-7=594

(4) das Bestimmtheitsmaß R^2.

- $R^2 = 1 - \frac{SSE}{SST}$
- $\bar{R}^2 = 1 - \frac{SSE/(T-K)}{SST/(T-1)}$
- $<=> 1 - \left((1-\bar{R}^2)\frac{T-K}{T-1}\right) = 1 - \frac{SSE}{SST} = R^2$
- $1 - \left((1-0{,}034)\frac{594}{600}\right) = 0{,}044 = R^2$

b) Betrachten Sie die Schätzergebnisse für die Koeffizienten des Modells. *(10 Punkte)*

(1) Interpretieren Sie b_1 und b_2 inhaltlich und geben Sie an, ob die Koeffizienten statistisch signifikant auf dem 10%-Niveau sind. Begründen Sie Ihre Antwort.

- b_1=0,214: Die erwartete Anzahl von Affären liegt für Männer ceteris paribus um 0,214 höher als für Frauen.
- Der Koeffizient ist nicht signifikant verschieden von Null auf dem 10%-Niveau (p-Wert=0,4714>0,1=α).
- b_2= -0,044: Mit jedem Jahr, um das der Befragte älter wird, sinkt die erwartete Anzahl an Affären ceteris paribus um 0,044**.**
- der Koeffizient ist statistisch signifikant auf dem 10%-Niveau (p-Wert=0,066<0,1=α).

(2) Welchen Betrag hätte b_2 angenommen, wenn das Alter in Monaten statt in Jahren gemessen worden wäre?

- b_2 hätte einen Betrag von -0,044/12= -0,003 angenommen (Umskalieren der Variable durch Multiplikation mit 12 → Division des Koeffizienten durch 12).

(3) Bestimmen Sie den marginalen Effekt der Ehedauer auf die erwartete Anzahl an Affären i) für Ehepaare, die seit 20 Jahren verheiratet sind, und ii) für Ehepaare, die seit 30 Jahren verheiratet sind. Zeigen Sie Ihren Rechenweg.

- $\frac{\partial \widehat{naffairs}}{\partial yrsmarr} = b_5 + 2b_6 \cdot yrsmarr$
- $\frac{\partial \widehat{naffairs}}{\partial yrsmarr}\Big|_{yrsmarr=20} = 0{,}306 - 2 \cdot 0{,}007 \cdot 20 = 0{,}0023$
- $\frac{\partial \widehat{naffairs}}{\partial yrsmarr}\Big|_{yrsmarr=30} = 0{,}306 - 2 \cdot 0{,}007 \cdot 30 = -0{,}1497$

(4) Um welchen Betrag unterscheidet sich ceteris paribus die erwartete Anzahl an Affären für i) Männer aus Ehen mit zwei Kindern und ii) Frauen aus kinderlosen Ehen? Zeigen Sie Ihren Rechenweg.

- Männer mit zwei Kindern:
  $\widehat{naffairs} = b_0 + b_1 \cdot 1 + b_2 \cdot age + b_3 \cdot educ + b_4 \cdot 2 + b_5 \cdot yrsmarr + b_6 \cdot yrsmarr^2$
- Frauen ohne Kinder:
  $\widehat{naffairs} = b_0 + b_1 \cdot 0 + b_2 \cdot age + b_3 \cdot educ + b_4 \cdot 0 + b_5 \cdot yrsmarr + b_6 \cdot yrsmarr^2$
- age, educ, yrsmarr identisch → Differenz i)-ii) =
  $b_1 + 2 \cdot b_4 = 0{,}214 + 2 \cdot -0{,}232 = -0{,}25$

Für Männer aus Ehen mit zwei Kindern ist die erwartete Anzahl an Affären ceteris paribus um 0,25 kleiner als für Frauen aus kinderlosen Ehen.

c) Testen Sie auf dem 5%-Niveau, ob der Effekt eines weiteren Kindes in der Ehe den Effekt, männlich zu sein genau ausgleicht. (Hinweis: $cov(b_1,b_4)= -0{,}008$)) Geben Sie dazu die Null- und Alternativhypothese, die Teststatistik, den kritischen Wert, Ihre Schlusslogik und das Testergebnis an. *(6 Punkte)*

- H_0: $\beta_1+\beta_4=0$, H_1: $\beta_1+\beta_4\neq0$
- $t=\frac{(b_1+b_4)-0}{se(b_1+b_4)}$
- $Var(b_1+b_4)=Var(b_1)+Var(b_4)+2\cdot Cov(b_1,b_4)=$
- $0{,}297^2+0{,}383^2+2\cdot(-0{,}008)=0{,}219$
- $se(b_1+b_4)=0{,}219^{0{,}5}=0{,}468$
- $t=\frac{(0{,}214-0{,}232)-0}{0{,}468}= -0{,}038$
- $t_{c(\alpha=5\%,\ T-K=594)}=\pm1{,}96$
- Da der empirische t-Wert kleiner als der kritische t-Wert ist, kann die Hypothese, dass der Effekt eines weiteren Kindes den Effekt, männlich zu sein, genau ausgleicht, nicht verworfen werden.

d) Jemand bezweifelt die Glaubwürdigkeit Ihrer Schätzergebnisse, da die Zufriedenheit in der Ehe vernachlässigt wurde. (*4 Punkte)*

(1) Welche Konsequenzen hätte dies für die Eigenschaften des Kleinstquadrate-Schätzers? Begründen Sie Ihre Antwort knapp verbal.

- Der Schätzer wäre verzerrt und inkonsistent, da die Zufriedenheit mit der Ehe nicht im Modell berücksichtigt wurde und diese ausgelassene Variable mit der Ehedauer korreliert sein könnte. („omitted variable"-Problematik).

(2) Können Sie das Problem lösen, indem Sie die Variable *ratemarr* in das Modell mit aufnehmen? Begründen Sie kurz verbal.

- Wird *ratemarr* in das Modell aufgenommen, ist *ratemarr* nicht mehr im Störterm enthalten.
- Die Ursache für die Verzerrung wegen ausgelassener Variablen wäre somit nicht mehr vorhanden und das Problem ausgelassener Variablen gelöst.

Aufgabe 2 **[30 Punkte]**

Im Jahr 2025 leiten Sie die Abteilung für empirische Fragen im Verkehrsministerium. 2017 wurde eine Geschwindigkeitsbeschränkung auf deutschen Autobahnen eingeführt. Im Rahmen einer Untersuchung zur Anzahl tödlicher Verkehrsunfälle auf deutschen Straßen liegt Ihnen ein Datensatz mit monatlichen Beobachtungen für den Zeitraum 2014 bis 2022 vor, der folgende Variablen enthält:

ltvunf	logarithmierte Anzahl tödlicher Verkehrsunfälle pro Monat
alq	Arbeitslosenquote pro Monat in Prozent (1-100)
geschw	Dummy: Gesetz zur Geschwindigkeitsbeschränkung ist in Kraft (1=ja, 0=nein)
z	Zeittrend für die beobachteten Monate (1-108)
z·alq	Interaktionsterm zwischen z und alq

Eine Kleinstquadrateschätzung der Spezifikation

$$ltvunf=\beta_1+\beta_2\cdot alq+\beta_3\cdot geschw+\beta_4\cdot z+\beta_5\cdot(z\cdot alq)+\varepsilon$$

liefert folgenden R-Output:

```
Call:
lm(formula = ltvunf ~ alq + geschw + z + zalq)
```

```
Residuals:
     Min       1Q   Median       3Q      Max
-0.23377 -0.08376  0.00126  0.06896  0.20824

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.1959332  0.2789189  22.214  < 2e-16 ***
alq         -0.0489752  0.0094751  -5.169 1.18e-06 ***
geschw       0.0108267  0.0411524   0.263    0.793
z           -0.0032602  0.0041746  -0.781    0.437
z·alq        0.0019713  0.0031341   0.629    0.531
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1049 on 103 degrees of freedom
Multiple R-squared: 0.3698,     Adjusted R-squared: 0.3453
F-statistic: 15.11 on 4 and 103 DF,  p-value: 9.448e-10
```

a) Beantworten Sie folgende Fragen: *(6 Punkte)*

(1) Warum werden Dummy-Variablen typischerweise verwendet? Beschreiben Sie die allgemeine Wirkungsweise von Dummy-Variablen im linearen Regressionsmodell kurz und präzise.

- Dummies werden verwendet, um einen Niveauunterschied in der abhängigen Variable zwischen zwei verschiedenen Gruppen darzustellen. Dabei bewirkt ein Dummy eine y-Achsenabschnittsverschiebung (beeinflusst die Konstante im Modell) und damit eine Parallelverschiebung der Regressionsgerade im einfachen Modell.

(2) Interpretieren Sie für das geschätzte multiple lineare Modell den Einfluss der Variable *geschw* auf die abhängige Variable inhaltlich. Lässt der p-Wert auf einen signifikanten Parameter schließen?

- Die mittlere Anzahl tödlicher Verkehrsunfälle lag in den Jahren, in denen das Gesetz zur Geschwindigkeitsbeschränkung in Kraft war, ceteris paribus im Durchschnitt um ca. 1,1% höher als in den Jahren, in denen das Gesetz nicht in Kraft war. Der p-Wert ist mit p = 0,793 größer als jedes gängige Signifikanzniveau, daher kann die H_0: $\hat{\beta}_3 = 0$ nicht verworfen werden. Der Parameter ist insignifikant.

(3) Wie lautet die Arbeitslosenquote, wenn der marginale Effekt von *z* auf die durchschnittliche monatliche Anzahl tödlicher Verkehrsunfälle 0,02 Prozent beträgt?

$$\frac{\partial ltvunf}{\partial z} = \hat{\beta}_4 + \hat{\beta}_5 \cdot alq$$

$$alq = \left(\frac{\partial ltvunf}{\partial z} - \hat{\beta}_4\right) / \hat{\beta}_5 = (0{,}02 + 0{,}003)/0{,}002 = 11{,}5$$

Der marginale Effekt wurde für eine Arbeitslosenquote von 11,5 Prozent bestimmt.

b) Ein Kollege weist Sie auf die Möglichkeit autokorrelierter Residuen hin. Was könnte dazu führen? Geben Sie ein Modell für die Residuen an. *(2,5 Punkte)*

- Im vorliegenden Fall handelt es sich um Zeitreihendaten. Im Fall autokorrelierter Residuen sind zeitlich benachbarte Residuen korreliert. Es ist z.B. plausibel, von positiver Autokorrelation erster Ordnung auszugehen, da in den Sommermonaten aufgrund des Urlaubsverkehrs systematisch mehr Verkehrsunfälle mit Todesfolge zu erwarten sind als in den Wintermonaten. (Alternativ: Es ist plausibel, von positiver Autokorrelation erster Ordnung auszugehen, da in den Wintermonaten wetterbedingt systematisch mehr Verkehrsunfälle mit Todesfolge zu erwarten sind als in den Sommermonaten).
- Z.B.: $e_t = \rho \cdot e_{t-1} + \nu_t$

c) Welche Annahme des Kleinstquadrate-Verfahrens wird bei Autokorrelation verletzt? Was sind die Folgen autokorrelierter Störterme für Kleinstquadrate-Schätzer? *(3 Punkte)*

- $\text{cov}(e_t, e_s) = 0$ gilt nicht.
- Der KQ-Schätzer ist weiterhin unverzerrt, aber nicht mehr effizient.
- Die Standardfehler der KQ-Schätzer sind nicht mehr korrekt, daher können Konfidenzintervalle und Hypothesentests irreführend sein.

d) Führen Sie auf dem 5%-Signifikanzniveau einen Durbin-Watson-Test auf positive Autokorrelation durch. Geben Sie die Null- und Alternativhypothese, die approximative Teststatistik sowie die Testentscheidung an (Hinweis: $\sum_{t=2}^{T} \hat{e}_t \cdot \hat{e}_{t-1} = 0{,}645$, $\sum_{t=2}^{T} \hat{e}_{t-1}^2 = 1{,}123$, $\sum_{t=2}^{T} (\hat{e}_t - \hat{e}_{t-1})^2 = 1{,}12$, $\sum_{t=1}^{T} \hat{e}_t^2 = 1{,}315$).

(5 Punkte)

- $H_0: \rho \leq 0$; $H_1: \rho > 0$
- $\hat{\rho} = \frac{\sum_{t=2}^{T} \hat{e}_t \cdot \hat{e}_{t-1}}{\sum_{t=2}^{T} \hat{e}_{t-1}^2} = \frac{0{,}645}{1{,}123} = 0{,}575$
- $d \approx 2(1 - \hat{\rho}) = 0{,}851$
- Alternativ: $d = \frac{\sum_{t=2}^{T} (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{T} \hat{e}_t^2} = 0{,}851$
- $d_{LC} = 1{,}59$
  $d_{Uc} = 1{,}76$
- Da $d < d_{Lc}$ muss die $H_0: \rho \leq 0$ verworfen werden. Das Testergebnis lässt auf einen Prozess positiver Autokorrelation erster Ordnung schließen.

e) Sie vermuten, dass es systematische Unterschiede in den Regressionsergebnissen für die beiden Regime mit und ohne Gesetz zur Geschwindigkeitsbeschränkung gibt. *(13,5 Punkte)*

(1) Nennen Sie einen Test, mit dem Sie dies überprüfen können und beschreiben Sie seine Vorgehensweise allgemein knapp aber präzise.

- Chow-Test
- Hierbei wird auf Unterschiede in den Regressionen für zwei Regime getestet. Dazu wird das Modell vollständig mit dem Dummy, der die beiden Regime kennzeichnet, interagiert. (Alternativ: Zwei separate Schätzungen für die beiden Regime.)
- Die Schätzung des vollständig interagierten Modells ergibt die Fehlerquadratsumme des unrestringierten Modells, im restringierten Modell wird unterstellt, dass keine Unterschiede zwischen den Regimen bestehen und die Parameter der Interaktionsterme sowie des Regimedummies daher insignifikant sind. (Alternativ: Die Schätzung der separaten Regressionen ergibt $SSE_1 + SSE_2 = SSE_U$. SSE_R ergibt sich aus der Schätzung des Ursprungsmodells ohne den Regimedummy)
- Mit einem F-Test wird getestet, ob die Parameter der Interaktionsterme gleich Null sind.

(2) Führen Sie den Test **für das Beispiel der Aufgabe** durch. Geben Sie mögliche Schätzgleichungen, die Null- und Alternativhypothese, die Teststatistik, den kritischen Wert auf dem 5%-Signifikanzniveau sowie die Testentscheidung an (Hinweis: Unterstellen Sie $SSE_R = 18731{,}98$; $SSE_U = 16297{,}42$).

- Unrestringiertes Modell:
  $ltvunf = \beta_1 + \beta_2 \cdot geschw + \beta_3 \cdot alq + \beta_4 \cdot alq \cdot geschw + \beta_5 \cdot z + \beta_6 \cdot z \cdot geschw +$
  $\beta_7 \cdot zalq + \beta_8 \cdot zalq \cdot geschw + \varepsilon$
- Restringiertes Modell:
  $ltvunf = \beta'_1 + \beta'_3 \cdot alq + \beta'_5 \cdot z + \beta'_7 \cdot zalq + \varepsilon'$
- $H_0: \beta_2 = \beta_4 = \beta_6 = \beta_8 = 0$; $H_1: \beta_2 \vee \beta_4 \vee \beta_6 \vee \beta_8 \neq 0$

- $F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(T-K)} = \frac{(18731{,}98 - 16297{,}42)/4}{16297{,}42/(108-7)} = 3{,}77$
- $F_{c;0,05;4;101} = 2{,}47$
- $F > F_c$ ⇨ Die H_0 muss auf dem 5%-Signifikanzniveau verworfen werden.

(3) Interpretieren Sie das Testergebnis und seine Bedeutung für Ihre Schätzstrategie.

- Der Test deutet an, dass es systematische Unterschiede zwischen den Regimen gibt. Daher sollten separate Regressionen für die beiden Regime oder ein vollständig interagiertes Modell geschätzt werden.

Aufgabe 3: **[10 Punkte]**

In R wurde folgende Funktion programmiert:

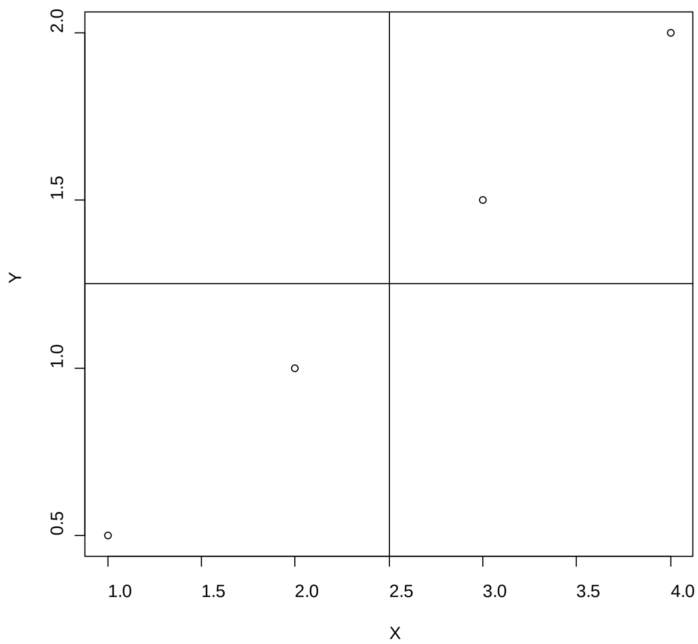
```
I.func <- function(x)
 {
  X <- seq(x:4*x)
  Y <- X/2
  plot(X,Y)
  abline(v=mean(X),h=mean(Y))
  return(sum(X),sum(Y))
}
```

a) Welchen R-Befehl müssen Sie eingeben, um die Funktion auszuführen und welchen, um Änderungen an der Funktion vorzunehmen? *(2 Punkte)*

```
> I.func()
> fix(I.func)
```

b) Stellen Sie alle Ausgaben so dar, wie sie mit dieser Funktion für x = 1 erzeugt werden. *(8 Punkte)*

Ausgegeben werden eine Grafik und zwei Werte:




```
> [1] 10
> [2] 5
```

Aufgabe 4: **[10 Punkte]**

Welche Antwort ist richtig? Bitte kreuzen Sie die zutreffende Antwort an. Zu jeder Frage gibt es nur eine richtige Antwort. Für jede korrekt angekreuzte Antwort gibt es 1 Punkt, für jede falsch angekreuzte Antwort wird 1 Punkt abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

1.	Welchen R-Befehl kann man **nicht** verwenden, um SST zu berechnen? (y sei die abhängige Variable einer linearen Regression)?	
	☒	`> sum((y – mean(y)^2))`
	☐	`> sum((y – mean(y))^2)`
	☐	`> sum({y – mean(y)}^2)`
2.	Welche Option ist in R beim Einlesen von Daten anzugeben, wenn die Daten in der ersten Zeile bereits Variablennamen enthalten?	
	☒	`> (..., header=T)`
	☐	`> (..., first.row=T)`
	☐	`> (..., varnames=T)`
3.	Welches Objekt wird mit dem R-Befehl `t(A)` generiert?	
	☐	Vektor mit *t*-Werten des Modells A
	☐	Zeitreihe t des Vektors A
	☒	Transponierte der Matrix A
4.	Mit welchem der folgenden Parameter des Befehls `> plot()` kann man die Beschriftung der *x*-Achse ändern?	
	☐	`xlim`
	☒	`xlab`
	☐	`xaxis`
5.	Welche Kenngröße wird mit dem R-Befehl `sum((x-mean(x))^2)/(length(x)-1)` ausgegeben?	
	☐	Schiefe der Verteilung der Zufallsvariable *x*
	☒	Varianz der Zufallsvariable *x*
	☐	Dichteverteilung der Zufallsvariable *x*
6.	Mit welchem R-Befehl erzeugen Sie einen Vektor *x*, der ungerade Zahlen zwischen 1 und 99 enthält?	
	☒	`> x <- seq(1,99, by=2)`
	☐	`> x <- seq(1,99, x[2]-1)`
	☐	`> x <- seq(1,99, odd=T)`
7.	Welchen der folgenden R-Befehle können Sie **nicht** verwenden, um die vorhergesagten Werte eines linearen Modells zu generieren?	
	☐	`> predict(lm(y ~ x))`
	☒	`> summary(lm(y ~ x))`
	☐	`> fitted.values(lm(y ~ x))`
8.	Bei welchem der folgenden R-Befehle erhalten Sie keine Fehlermeldung?	
	☐	`> pf(.99;10,3)`
	☐	`> pf(.99,10;3)`
	☒	`> pf(.99,10,3)`
9.	Welchen R-Befehl müssen Sie verwenden, um den Datensatz `Daten.txt` einzulesen?	

	❑	`> get.data("C:/StatistikII/Daten.txt")`
	☒	`> read.table("C:/StatistikII/Daten.txt")`
	❑	`> read.data("C:/StatistikII/Daten.txt")`
10.	Welchen R-Befehl kann man verwenden, um einen Chow-Test durchzuführen?	
	❑	`> chow.test(mod1.kq, mod2.kq)`
	☒	`> anova(mod1.kq, mod2.kq)`
	❑	`> F.test(mod1.kq, mod2.kq)`

Aufgabe 5 **[26 Punkte]**

Wahr oder falsch? Tragen Sie für jede der folgenden Aussagen ein „w“ für „wahr“ oder ein „f“ für „falsch“ ein. Für jede richtige Antwort gibt es 1 Punkt, für jede falsche Antwort wird 1 Punkt abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

W	Je stärker die Konzentration eines Merkmals, umso größer der Abstand zwischen Lorenzkurve und Diagonale.
F	Heteroskedastie führt zu verzerrten Schätzern für den Achsenabschnittsparameter.
F	Um eine saisonbereinigte Zeitreihe zu erstellen, können lineare, exponentielle oder logistische Saisonmodelle genutzt werden.
F	Eine Variable z kann als Instrument für eine erklärende Variable x genutzt werden, wenn ihre Standardabweichung nicht mit x korreliert.
W	Eine Vorhersage auf Basis eines linearen Modells ist genau dann unverzerrt, wenn der Erwartungswert des Vorhersagefehlers 0 beträgt.
F	Der KQ Schätzer ist inkonsistent, wenn cov(x,e) = 0.
W	Der Herfindahl-Index ist ein absolutes Konzentrationsmaß.
F	Im linearen Modell gibt die Regressionskonstante den Mittelwert der abhängigen Variablen an.
W	Nach dem Method of Moments Schätzverfahren, lassen sich Bevölkerungsparameter durch analoge Parameter der Stichprobe schätzen.
W	Unverzerrtheit und Effizienz stellen unterschiedliche Anforderungen an die Auswahl der erklärenden Variablen.
W	Der Paascheindex ist kommensurabel.
W	Multikollinearitätsprobleme lassen sich über eine Erhöhung der Beobachtungszahl reduzieren.
F	Der Kleinstquadrateschätzer ist umso unverzerrter, je größer die Stichprobe.
F	Um nichtlineare Zusammenhänge zwischen erklärenden und abhängigen Variablen abzubilden, muss mehr als eine erklärende Variable im Modell sein.
W	Ein Typ I Fehler liegt vor, wenn eine Nullhypothese verworfen wurde, obwohl sie zutrifft.
W	Bei einem t-Test wird die Nullhypothese verworfen, wenn der p-Wert kleiner als α ist.
W	Bei heteroskedastischen Störtermen bleiben die Kleinstquadrateschätzer der Steigungsparameter unverzerrt.

W	Die Hypothese positiver Autokorrelation erster Ordnung kann mittels eines t-Tests getestet werden.
F	Eine Vorhersage auf Basis eines log-linearen Modells ist genau dann unverzerrt, wenn der Vorhersagefehler für mindestens eine Beobachtung 0 beträgt.
F	Beim F-Test entspricht die Zahl der Freiheitsgrade im Nenner der Anzahl der getesteten Restriktionen.
W	Man kann einen Lagrange Multiplikator-Test verwenden, um auf Autokorrelation erster Ordnung zu testen.
W	Das Modell mit autokorrelierten Störtermen unterstellt, dass Zufallsschocks über eine Periode hinaus wirken.
F	Werden logarithmierte erklärende Variablen genutzt, so muss der geschätzte Steigungsparameter positiv sein.
F	Das Verfahren gleitender Durchschnitte kann zur Preisbereinigung verwendet werden.
F	In einem Modell kann jede Variable W als Instrumentvariable genutzt werden, die nicht in der Modellgleichung vorkommt.
F	$\frac{\sum p_0 \cdot q_0}{\sum p_0 \cdot q_t}$ ist ein Paasche-Mengenindex.

Aufgabe 6 **[15 Punkte]**

Welche Antwort ist richtig? Bitte kreuzen Sie die zutreffende Antwort an. Zu jeder Frage gibt es nur eine richtige Antwort. Für jede korrekt angekreuzte Antwort gibt es 1 Punkt, für jede falsch angekreuzte Antwort wird 1 Punkt abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

1.	Der Kleinstquadratschätzer	
	☑	kann bei nicht normalverteilten Residuen das Gauss-Markov-Theorem erfüllen.
	❑	sollte nur bei normalverteilten Residuen verwendet werden.
	❑	sollte nur bei normalverteilten abhängigen Variablen verwendet werden.
2.	Intervallschätzer sind	
	❑	weniger informativ als Punktschätzer;
	❑	nur für die Grundgesamtheit gültig;
	☒	umso enger, je kleiner der geschätzte Standardfehler ist.
3.	Bei einem t-Test der Nullhypothese $\beta \geq k$ am 1 Prozent Signifikanzniveau	
	☒	kann die Nullhypothese nicht verworfen werden, wenn $p > 0.01$;
	❑	ist die t-Verteilung von der Stichprobengröße unabhängig;
	❑	muss die Nullhypothese verworfen werden, wenn $p < 0.10$.
4.	Eine Multiplikation der abhängigen Variable mit 100 führt zu	
	❑	einem um den Faktor 100 erhöhten Achsenabschnittsparameter bei unveränderten Steigungsparametern.
	☒	um den Faktor 100 erhöhten Werten für alle Achsenabschnitts- und Steigungsparameter.
	❑	unveränderten Parametern.
5.	Ausgelassene Variablen	
	☒	können zu verzerrt geschätzten Parametern führen.
	❑	führen zu überhöhten R^2-Werten.
	❑	gibt es nur bei Zeitreihendaten.

6.		Ein Parameter β_2 wird umso präziser geschätzt
	☒	je größer die Streuung in x ist.
	☐	je kleiner sein anschließend bestimmter t-Wert ist.
	☐	je größer die Streuung in y ist.
7.		Werden Interaktionseffekte im Regressionsmodell verwendet,
	☒	können Parameter für verschiedene Teilstichproben verglichen werden.
	☐	kann anhand des RESET-Tests ihre Signifikanz überprüft werden.
	☐	können die geschätzten Koeffizienten nicht mehr interpretiert werden.
8.		Bei gegen unendlich konvergierender Stichprobengröße
	☐	konvergiert der Intervallschätzer der Steigungsparameter gegen das Signifikanzniveau.
	☒	konvergiert die Varianz des KQ Schätzers gegen Null.
	☐	konvergiert das R^2 gegen 1.
9.		Der Typ II Fehler
	☐	tritt auf, wenn die Nullhypothese verworfen wird, obwohl sie zutrifft.
	☐	ist umso wahrscheinlicher, je größer die Stichprobe ist.
	☒	wird unwahrscheinlicher, wenn der Typ I Fehler wahrscheinlicher wird.
10.		Zur Durchführung eines Chow-Tests
	☐	werden Interaktionsterme und ihr quadrierter Wert benötigt.
	☒	kann man die Stichprobe in zwei Teile teilen.
	☐	werden Polynome der vorhergesagten Werte verwendet.
11.		Der R^2-Wert einer Schätzung ist umso höher,
	☒	je geringer die unerklärte Variation der abhängigen Variable.
	☐	je geringer die erklärte Variation der erklärenden Variablen.
	☐	je größer der Achsenabschnitt geschätzt wird.
12.		Welche Interpretation für β_1 ist im einfachen linearen Regressionsmodell $\ln y = \beta_0 + \beta_1 \ln x + e$ zutreffend?
	☐	bei einer Änderung von x um eine Einheit ändert sich y um β_1 Einheiten.
	☐	bei einer Änderung von x um eine Einheit ändert sich y um $100 \cdot \beta_1$ Prozent.
	☒	bei einer Änderung von x um ein Prozent ändert sich y um β_1 Prozent.
13.		Die Hypothese, dass eine erklärende Variable endogen ist,
	☒	lässt sich mit Hilfe des Hausman-Tests prüfen.
	☐	lässt sich mit Hilfe des Goldfeld-Quandt-Tests prüfen.
	☐	lässt sich mit Hilfe des Gauss-Markov-Tests prüfen.
14.		Bei Messfehlern in den erklärenden Variablen
	☐	ist die Präzision der Schätzung reduziert.
	☒	sind die Parameterschätzer inkonsistent.
	☐	sollten generalisierte Kleinstquadrateschätzer verwendet werden.
15.		Ein Lagrange Multiplikator-Test auf Autokorrelation
	☐	ist nur für kleine Stichproben gültig.
	☒	gilt auch, wenn eine der erklärenden Variablen die verzögerte abhängige Variable y_{t-1}

		ist.
	❑	kann nur zum Test auf das Vorliegen eines AR(1)-Prozesses verwendet werden.