

Beispielklausur (Stand: Januar 2019)

Bitte beachten Sie!

- Klausurdauer: 90 Minuten
- Die Klausur ist „open book“. Es sind ein Taschenrechner, die Unterlagen der Vorlesung sowie maximal zwei Statistiklehrbücher erlaubt.
- Bei Berechnungen können Sie jeweils auf zwei Nachkommastellen runden.
- Teilen Sie nicht nur das Ergebnis einer Rechnung mit, sondern auch den *Rechenweg* und die *Zwischenergebnisse*!
- Wenn Sie die Rückseite des Blatts nutzen, kennzeichnen Sie dies auf der Vorderseite.
- Es sind maximal 90 Punkte zu erreichen (**in dieser Beispielklausur ausnahmsweise 100 Punkte!!**). Mindestens 45 sind zum Bestehen der Klausur erforderlich.

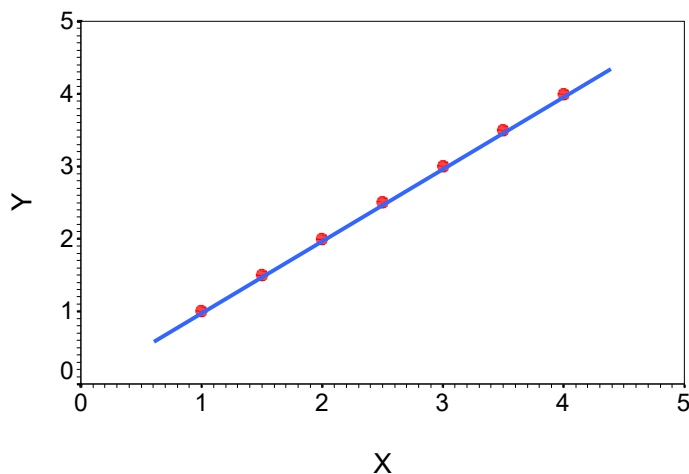
BITTE IN DRUCKBUCHSTABEN AUSFÜLLEN!

| | |
|----------------------|-------------------|
| Nachname:..... | Vorname:..... |
| Matrikelnummer:..... | Studiengang:..... |

NICHT AUSFÜLLEN!

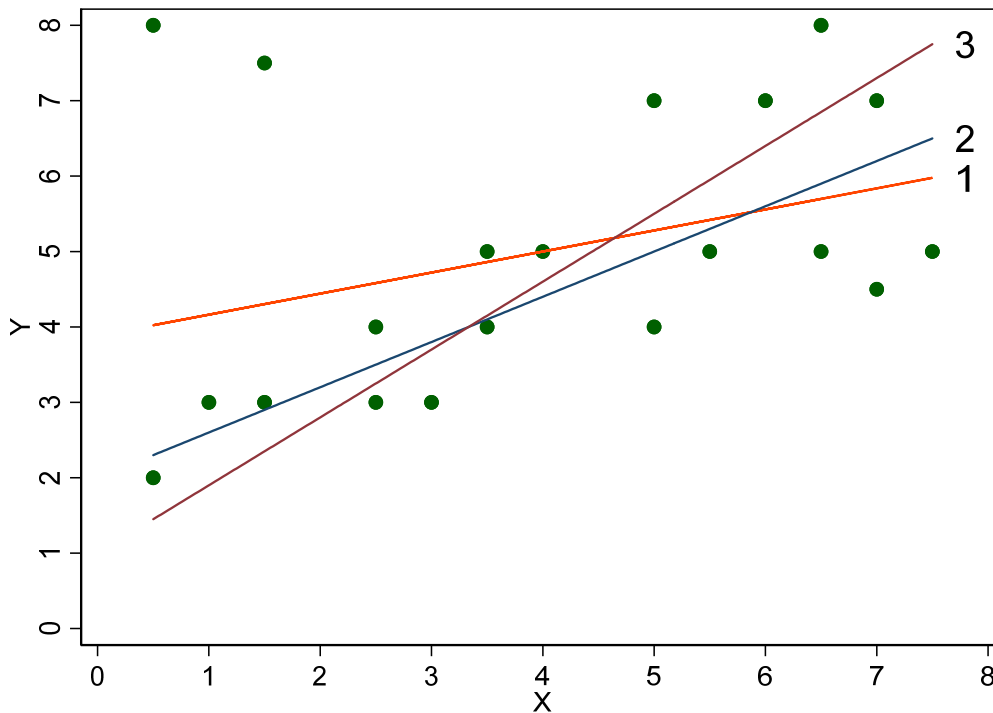
| Aufgabe | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Summe | Note |
|---------|-----|-----|-----|------|------|------|------|-------|------|
| Punkte | (5) | (5) | (5) | (25) | (20) | (30) | (10) | | |

1. Wie groß sind α und β bei dieser Regression? Wie groß ist das Bestimmtheitsmaß R^2 ? (5 P)



2. Sie rechnen eine Regression mit Einkommen als aV und Berufserfahrung als uV . Skizzieren Sie einen hypothetischen Residuen-Plot für diese Regression, der auf Nicht-Linearität hindeutet. Vergessen Sie nicht, die Achsen zu beschriften. (5 P)

3. Sie rechnen eine OLS-Regression von Y ($\bar{Y} = 5$) auf X ($\bar{X} = 4$) mit 20 Fällen. Welche der drei Geraden ist die OLS-Regressionsgerade? Begründen Sie Ihr Ergebnis. (5 P)



4. Sie untersuchen, wovon die Punktezahl im Mathetest der Teilnehmer von Statistik I abhängt. Sie rechnen eine Regression auf MATHE (Zahl richtig gelöster Aufgaben). UV sind ABINOTE (Note im Abitur) und MANN (Dummy, 0=Frau, 1=Mann). STATA liefert folgendes Ergebnis:

```
. regress MATHE abinote mann, beta
```

| Source | SS | df | MS | Number of obs = | 299 |
|----------|--------|-----|------------|-----------------|--------|
| Model | 111.6 | 2 | 55.8 | F(2, 296) = | ???? |
| Residual | 2349.3 | 296 | 7.9 | Prob > F = | ???? |
| Total | ?????? | 298 | 8.25799645 | R-squared = | ???? |
| | | | | Adj R-squared = | 0.0389 |
| | | | | Root MSE = | 2.8173 |

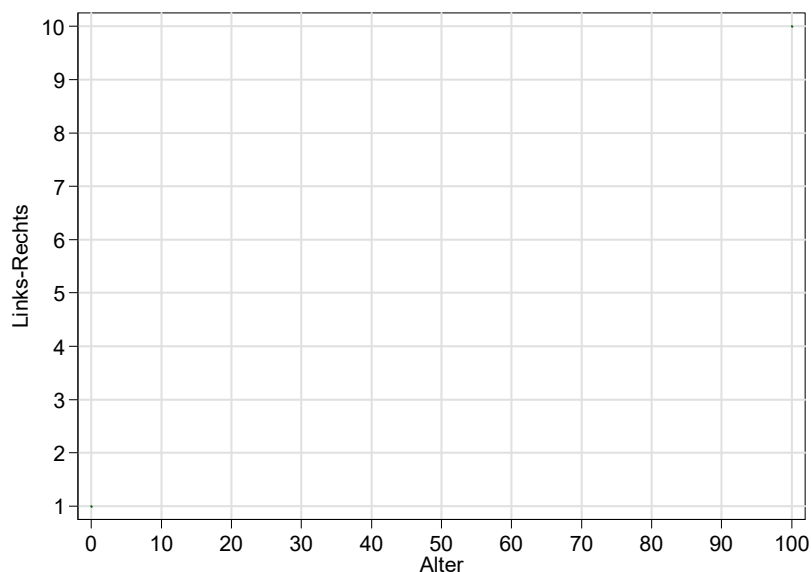
| MATHE | Coef. | Std. Err. | t | P> t |
|---------|-------|-----------|-------|-------|
| abinote | -.84 | .28 | ???? | ????? |
| mann | .86 | ????? | ???? | 0.012 |
| _cons | 10.5 | .7549 | 13.97 | 0.000 |

- Beurteilen Sie die Signifikanz des Gesamtmodells ($\alpha = 0,05$) ($F_{0,95}(2, 296) = 3,0$). (5 P)
- Bestimmen und interpretieren Sie R^2 . (6 P)
- Interpretieren Sie die Regressionskoeffizienten. (4 P)
- Beurteilen Sie die Signifikanz des Effekts der Abiturnote ($\alpha = 0,05$). (6 P)
- Beurteilen Sie die Signifikanz des Geschlechtseffektes ($\alpha = 0,01$). (4 P)

5. Sie berechnen mit den Daten des ALLBUS 1994 eine Regression mit der Links-Rechts Skala (Selbsteinstufung des Befragten: 1=Links, 10=Rechts; Mittelwert 5,0) als abhängige Variable. Unabhängige Variablen sind das Alter (in Jahren; Wertebereich 18-93) und eine West/Ost Dummy (0=Westdeutscher, 1=Ostdeutscher). Zusätzlich berücksichtigen Sie einen Interaktionsterm. Sie erhalten folgenden Output:

| Rechts | Coef. | Std. Err. | t | P> t |
|--------------------|-------|-----------|--------|-------|
| Alter | .012 | .002 | 5.967 | 0.000 |
| Ostdeutscher | .700 | .184 | 3.927 | 0.000 |
| Alter*Ostdeutscher | -.020 | .004 | -5.683 | 0.000 |
| _cons | 4.3 | .179 | 27.061 | 0.000 |

- a. Welchen Effekt hat „Alter“ für Westdeutsche, welchen für Ostdeutsche? Ist der Unterschied signifikant ($\alpha = 0,05$)? (5 P)
- b. Wie beurteilen Sie den West/Ost Unterschied in der politischen Einstellung? Erstellen Sie hierzu die Regressionsgleichungen für West- und Ostdeutsche. Skizzieren Sie einen Profile-Plot (Alter variiert von 0 bis 100) und interpretieren Sie anhand des Profile-Effect Plots den West/Ost Unterschied. (15 P)



6. Sie berechnen mit den Daten des ALLBUS 1994 eine logistische Regression mit der Wahlabsicht (Sonntagsfrage: 1=CDU, 0=SPD) als abhängige Variable. Unabhängige Variablen sind die Bildung (in Jahren; Wertebereich 7-24), die Konfession (ohne Konfession, katholisch, evangelisch) und das Alter (in Jahren; Wertebereich 18-93). Sie erhalten folgenden Output:

```
Iteration 0: log likelihood = -1062.0
Iteration 1: log likelihood = -1008.3
Iteration 2: log likelihood = -1008.1
Iteration 3: log likelihood = -1008.0
```

Log likelihood = -1008.0

| CDU | Coef. | Std. Err. | z | P> z |
|-------------------------|--------|-----------|--------|-------|
| Bildung | .030 | .021 | 1.324 | 0.185 |
| Konfession: katholisch | 1.250 | .147 | 8.617 | 0.000 |
| Konfession: evangelisch | .503 | .140 | 3.588 | 0.000 |
| Alter | .015 | .003 | 4.837 | 0.000 |
| _cons | -1.900 | .343 | -5.663 | 0.000 |

- Berechnen Sie Pseudo-R² und beurteilen Sie den Modellfit. (5 P)
- Berechnen Sie die Likelihood-Ratio Teststatistik und beurteilen Sie, ob das Modell signifikant ist (bei einem Signifikanzniveau von 5% sind die kritischen χ^2 -Werte: 3,84 (df=1), 5,99 (df=2), 7,81 (df=3), 9,49 (df=4)). (5 P)
- Wie hoch ist die Wahrscheinlichkeit CDU zu wählen für einen 30-jährigen konfessionslosen Realschüler (Bildung=10)? Wie hoch ist sie für einen 30-jährigen katholischen Realschüler? (14 P)
- Wahrscheinlichkeits-Interpretation: Interpretieren Sie den Effekt von "katholisch". Verwenden Sie hierzu die Ergebnisse aus der Teilaufgabe c. (2 P)
- Odds-Interpretation: Wie groß ist der Effekt des "Alters" auf das Odds CDU vs. SPD? (2 P)
- Logit-Interpretation: Beurteilen Sie den Effekt der "Bildung" (Richtung und Signifikanz). (2 P)

7. Sie berechnen mit den Daten des ALLBUS 2002 eine multinomiale logistische Regression mit der Wahlabsicht (Sonntagsfrage: CDU=0, SPD=1, Grüne=2) als abhängige Variable. Unabhängige Variable ist das Geschlecht (Mann=0, Frau=1). Sie erhalten folgenden Output:

| Multinomial logistic regression | | | | Number of obs | = | 1,583 |
|---------------------------------|----------------|-----------|--------|---------------|----------------------|-----------|
| partei | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] | |
| CDU | (base outcome) | | | | | |
| SPD | | | | | | |
| frau | .16 | .108031 | 1.50 | 0.133 | -.0493829 | .374091 |
| _cons | -.16 | .0738231 | -2.24 | 0.025 | -.3101551 | -.0207738 |
| Grüne | | | | | | |
| frau | .58 | .1618767 | 3.64 | 0.000 | .2720447 | .9065898 |
| _cons | -1.58 | .1212234 | -13.07 | 0.000 | -1.822339 | -1.347152 |

- Wie hoch ist die Wahrscheinlichkeit „Grüne“ zu wählen für einen Mann? (4 P)
- Wie hoch ist die Wahrscheinlichkeit „Grüne“ zu wählen für eine Frau? (4 P)
- Berechnen Sie den Geschlechtsunterschied für P(Grüne). (2 P)