

Klausurvorbereitung - Statistik

Aufgabe 1

Studenten der Politikwissenschaft der Johannes Gutenberg-Universität wurden befragt, seit wie vielen Semestern sie eingeschrieben sind.

Berechnen Sie für die folgenden Antworten:

Index i	1	2	3	4	5	6	7	8	9	10	11	12	13
Semesterzahl x_i	4	11	13	2	4	7	19	8	4	6	7	6	3

- a) den Modus
- b) den Median
- c) das arithmetische Mittel
- d) die Spannweite, die Standardabweichung und die Varianz!
- e) Ist die Verteilung symmetrisch, links- oder rechtsgipflig (bzw. links- oder rechtssteil)?

Aufgabe 2

In einer Stichprobe finden Sie folgenden Zusammenhang zwischen dem Geschlecht und der Wahl der Republikaner:

Wahl der REP	Geschlecht		Summe
	männlich	weiblich	
Ja	70	30	100
Nein	930	1170	2100
Summe	1000	1200	2200

- a) Wie hoch ist der Anteil der Republikaner in der Stichprobe? Wieviel Prozent der Männer bzw. der Frauen wählen die Republikaner?
- b) Um wieviel *Prozentpunkte* unterscheidet sich der Stimmenanteil der REP unter Männern und Frauen? Um wieviel *Prozent* ist der Anteil der REP unter den Männern höher als unter den Frauen?
- c) Berechnen Sie das Zusammenhangsmaß V .
- d) Testen Sie mit einer Irrtumswahrscheinlichkeit von 5 % die Hypothese, daß ein signifikanter Zusammenhang besteht.
- e) Welchen Fehler schließen Sie mit diesem Test mit großer Sicherheit aus? Antworten Sie bitte in einem ganzen Satz, der Name des Fehlers genügt nicht.

Aufgabe 3

Sie messen bei 6 Personen sowohl Körpergröße als auch Körpergewicht. Dabei ergeben sich die folgenden Werte.

	Körpergröße	Gewicht
Michael	193	85
Nicole	162	55
Huy	183	72
Tajana	168	58
Katja	170	60
Cornelius	179	76

- a) Stellen Sie die Beziehung zwischen Körpergröße und Körpergewicht graphisch (in einem Streudiagramm) dar !
- b) Beschreiben Sie den Zusammenhang zwischen Körpergröße und -gewicht als eine mathematische Funktion (Modell der linearen Regression). Betrachten Sie dabei das Körpergewicht als abhängige Variable. Verwenden Sie zur Berechnung der Geradengleichung und des Bestimmtheitsmaßes R^2 eine Hilfstabelle.

- c) Beschreiben Sie verbal die Bedeutung des Regressionskoeffizienten und der Regressionskonstanten. Welcher Regressionskoeffizient resultiert, wenn die Regressionsgerade waagrecht verlief?
- d) Wieviel Prozent der Varianz des Körpergewichts wird durch die Körpergröße erklärt?
- e) Welches Gewicht erwarten Sie auf Grund der Regressionsgleichung für eine Person mit einer Körpergröße von 210 cm
- f) Welchen Wert nimmt die Korrelation zwischen Körpergröße und -gewicht an?

Aufgabe 4

Die Wahl zum Bundestag am 2.12.1990 erbrachte in vier Bundesländern folgende Ergebnisse:

Ergebnis der Wahl zum Bundestag am 2.12.1990				
Land	Einheit	gültige Zweitstimmen	SPD	FDP
Bayern	1000	6367,2	1698,0	551,9
Berlin	1000	2012,7	616,3	183,8
Hessen	1000	3443,0	1308,2	374,2
Rheinland-Pfalz	1000	2363,7	853,1	245,3

Die Wahl zum Bundestag am 16.10.1994 erbrachte in vier Bundesländern folgende Ergebnisse:

Ergebnis der Wahl zum Bundestag am 16.10.1994				
Land	Einheit	gültige Zweitstimmen	SPD	FDP
Bayern	1000	6693,0	1984,0	430,1
Berlin	1000	1950,1	663,1	100,6
Hessen	1000	3485,9	1296,8	283,2
Rheinland-Pfalz	1000	2426,2	955,4	168,5

- a) Ermitteln Sie für die vier Bundesländer bei beiden Wahlen jeweils den Stimmenanteil der SPD an den gültigen Zweitstimmen!
- b) Um wieviel *Prozentpunkte* hat sich der Stimmenanteil der SPD 1994 im Vergleich zu 1990 in den vier Bundesländer verändert?
- c) Um wieviel *Prozent* hat sich der Stimmenanteil der SPD 1994 im Vergleich zu 1990 in den vier Bundesländer verändert?
- d) Führen Sie a) bis c) auch für den Stimmenanteil der FDP durch!

Aufgabe 5

Sie befragen 1000 zufällig ausgewählte wahlberechtigte Bundesbürger bezüglich ihrer Wahlabsicht. 4 Prozent der Befragten geben an, die Republikaner wählen zu wollen.

- a) Welches Intervall schließt mit 99 %iger Sicherheit den Stimmenanteil für die Republikaner in der gesamten wahlberechtigten Bevölkerung ein?
- b) In welchem Intervall würden Sie den wahren Wert vermuten, wenn Sie den Stichprobenumfang vervierfachen?
- c) Wenn Sie beide Intervalle vergleichen, was stellen Sie fest?
- d) Wiederholen Sie die Schritte a) bis c) für den Stimmenanteil von B90/Grüne (7 Prozent)!

Aufgabe 6

In einer Untersuchung wurde unter anderem der Zusammenhang zwischen der Zufriedenheit mit der Regierungspolitik – gemessen von 1 (sehr unzufrieden mit der Regierung) bis 11 (sehr zufrieden mit der Regierung) – und dem Wahlverhalten (1: SPD, 2: CDU/CSU, 3: B90/Grüne, 4: FDP, 5: PDS, 6: Sonstige) analysiert.

- a) Berechnen Sie die Regressionsgerade!
- b) Wie läßt sich der Regressionskoeffizient interpretieren?
- c) Welches Wahlverhalten würden Sie für einen x-Wert von 6 (auf der Zufriedenheitsskala) voraussagen?

Aufgabe 7

In einer repräsentativen Zufallsstichprobe einer bekannten Mainzer Karnevalsveranstaltung ergibt sich folgende Altersverteilung:

18 30 21 50 62 45 36 91 33 37 48 19 58 55 56 24 31

Wir unterstellen, daß sowohl die Stichprobenziehung als auch die Befragung gegen 23 Uhr noch zu validen Meßergebnissen führte.

Bestimmen Sie

- das arithmetische Mittel der Meßwerte, den Median der Verteilung
- die Varianz, sowie die geschätzte Varianz in der Grundgesamtheit.

Aufgabe 8

Sie wollen die These überprüfen, zwischen dem Berufsstatus und dem Wahlverhalten bestehe ein Zusammenhang. 1998 ergab sich dazu in einer westdeutschen Stichprobe folgende Verteilung.

Wahlverhalten	Berufsstatus			Summe
	Selbständige	Angestellte und Beamte	Arbeiter	
CDU/CSU	80	321	154	555
SPD	32	401	219	652
Andere	49	162	41	252
Summe	161	884	414	1459

- Berechnen Sie einen auf Chi-Quadrat beruhenden Koeffizienten, der den Zusammenhang angemessen beschreibt.
- Erlaubt das Ergebnis mit einer Irrtumswahrscheinlichkeit von 5 Prozent den Schluß auf einen Zusammenhang in der Grundgesamtheit?

Aufgabe 9

Sie ziehen eine Stichprobe an einer deutschen Universität und berechnen anhand einer Reihe von Items den Intelligenzquotienten (IQ) für jeden der 1200 Befragten. Für das arithmetische Mittel in der Stichprobe errechnen sie einen Wert von 95, für die Standardabweichung bei sämtlichen Studenten der Universität schätzen Sie $\hat{S} = 15$. (Es wurden nur Studenten befragt!)

- Welche Voraussetzung muß gegeben sein, um auf Basis der Stichprobe eine Intervallschätzung für die gesamte Universität zu geben?
- Innerhalb welcher Grenzen bewegt sich der IQ aller Studenten an der Universität (bei einer Wahrscheinlichkeit von 99 %) ?
- Von welchen Faktoren ist die Intervallbreite der Schätzung abhängig?

Aufgabe 10

Bei einer Untersuchung zur generationalen Mobilität ergibt sich die untenstehende Tabelle.

Schicht des Sohnes	Schicht des Vaters			Summe
	Unterschicht	Mittelschicht	Oberschicht	
Unterschicht	126	22	8	156
Mittelschicht	70	465	14	549
Oberschicht	4	113	78	195
Summe	200	600	100	900

- Wieviel Prozent der Unterschicht-Söhne sind in eine höhere Schicht aufgestiegen?
- Welches Maß kann den Zusammenhang zwischen der Schichtzugehörigkeit des Vaters und der des Sohnes angemessen beschreiben?
- Welchen Wert nimmt der Koeffizient für diese Tabelle an?

Aufgabe 11

Berechnen Sie für die folgende Notenverteilung einer ESF und Statistik-Klausur:

Noten	1	2	3	4	5	6
Häufigkeiten	1	3	7	3	9	2

- das arithmetische Mittel
- die Varianz und die Standardabweichung!

Aufgabe 12

In einer Untersuchung wurden unter anderem erhoben: die Selbsteinstufung auf der Links-Rechts-Dimension und die Einstellung zu der Aussage „Der Sozialismus ist eine gute Idee, die bisher nur schlecht ausgeführt wurde.“

Links-Rechts-Wert 1=links, 11=rechts	„Sozialismus ist eine gute Idee“ 1=ablehnend, 5=zustimmend
10	2
9	4
3	5
8	4
6	3
8	4
5	4
5	5
9	5

- Mit welchem Modell kann der Einfluß der unabhängigen Variablen Links-Rechts-Selbsteinstufung auf die Einstellung zum Sozialismus angemessen beschrieben werden?
- Berechnen Sie die Vorhersagegleichung,
- den Korrelationskoeffizienten r und
- das Bestimmtheitsmaß R^2 . Interpretieren Sie das Bestimmtheitsmaß!
- Welchen Wert des Sozialismus-Items würden Sie bei einem Befragten mit einem Links-Rechts-Wert von 1 vorhersagen?

Aufgabe 13

Gegeben sei die folgende Kreuztabelle:

Kernkraftwerke	Wahlabsicht		Summe
	CDU	SPD	
...abschalten	222	478	700
...weiter betreiben	628	322	950
Summe	850	800	1650

- Berechnen Sie Spalten- und Zeilenprozentage. Interpretieren Sie die Prozentwerte!
- Mit welchem auf Chi-Quadrat beruhenden Koeffizienten können Sie den Zusammenhang angemessen beschreiben? Welchen Wert nimmt der Koeffizient an?
- Ist das gefundene Ergebnis statistisch signifikant, d.h. können Sie mit einer Irrtumswahrscheinlichkeit von 5 Prozent davon ausgehen, daß in der Grundgesamtheit nicht die Nullhypothese gilt?

Aufgabe 14

Die Wahl zum Bundestag am 2.12.1990 erbrachte in vier Bundesländern folgende Ergebnisse:

Land	Ergebnis der Wahl zum Bundestag am 2.12.1990			
	Einheit	Wahlberechtigte	gültige Zweitstimmen	ungültige Zweitstimmen
Bayern	1000	8623,6	6367,2	52,6
Berlin	1000	2537,3	2012,7	31,2
Hessen	1000	4294,1	3443,0	41,1
Rheinland-Pfalz	1000	2935,6	2363,7	33,3

- Ermitteln Sie für die vier Bundesländer die Wahlbeteiligung!
- Bestimmen Sie den Anteil der gültigen Stimmen an den abgegebenen Stimmen!

Die Wahl zum Bundestag am 16.10.1994 erbrachte in vier Bundesländern folgende Ergebnisse:

Land	Ergebnis der Wahl zum Bundestag am 16.10.1994			
	Einheit	Wahlberechtigte	gültige Zweitstimmen	ungültige Zweitstimmen
Bayern	1000	8767,5	6693,0	51,2
Berlin	1000	2505,9	1950,1	20,4
Hessen	1000	4290,3	3485,9	47,0
Rheinland-Pfalz	1000	2985,4	2426,2	29,9

- Ermitteln Sie auch hier sowohl die Wahlbeteiligung als auch den Anteil der gültigen Stimmen an den abgegebenen Stimmen!
- Um wieviel *Prozentpunkte* hat sich die Wahlbeteiligung 1994 im Vergleich zu 1990 in den vier Bundesländer verändert.
- Um wieviel *Prozent* hat sich die Wahlbeteiligung 1994 im Vergleich zu 1990 in den vier Bundesländer verändert?

Aufgabe 15

Gegeben sei eine Zufallsstichprobe mit $n=3500$. Sie berechnen auf einer Politiker-Sympathie-Skala (1=sehr gut, 6=ungenügend) für den Bundeskanzler einen Mittelwert von 3,3.

- Berechnen Sie das 95%ige Konfidenzintervall. Verwenden Sie als geschätzte Varianz für die Grundgesamtheit $\hat{S}^2=4$.
- Erhöhen Sie die Sicherheit auf 99 Prozent. Was beobachten Sie bezüglich der Intervallbreite?
- Interpretieren Sie beide Konfidenzintervalle!

Aufgabe 16

- Wieso existieren verschiedene Zusammenhangsmaße? Stellen Sie alle (Ihnen bekannten) Maße vor!
- Gegeben seien zwei Variablen, für die ein Korrelationskoeffizient von $r = 0,56$ berechnet wurde. Wie interpretieren Sie diesen Wert?
- Zu welchem Zweck sollte nun eine Regressionsanalyse gerechnet werden?