

Einführung in die Statistik für Politikwissenschaftler

Sommersemester 2011

Es können von den Antworten alle, mehrere oder keine Antwort(en) richtig sein. Nur bei einer korrekten Antwort (ohne Auslassungen und überflüssige Ankreuzungen) erhalten Sie die angegebene Punktzahl! Viel Erfolg!

1. Für eine statistische Analyse wird eine Reihe von Merkmalen erfasst. Notieren Sie jeweils mit N, O, I bzw. R, ob es sich um eine Nominal-, Ordinal-, Intervall- oder Ratio-Skala handelt (4 Punkte).

- Das Geburtsjahr des Befragten
- Die Staatsangehörigkeit des Befragten
- Die Selbsteinschätzung auf einer Links-Rechts-Skala (von 0 bis 10)
- Die Kindersterblichkeitsrate pro 1.000 Einwohner eines Landes
- Die Schulden eines Landes in US-Dollar
- Die Beurteilung der Arbeit der Bundesregierung (sehr gut, gut, schlecht, sehr schlecht)
- Die Arbeitslosenrate eines Landes
- Das politische System eines Landes (Anarchie, Diktatur, Monarchie, Theokratie, Demokratie)

2. Wie kann man den Zusammenhang zwischen einer nominal- und einer ordinalskalierten Variablen darstellen? (3 Punkte)

- Es ist nicht möglich, den Zusammenhang zwischen diesen Variablen darzustellen
- Pearson's N
- Streudiagramm
- Chi²

3. Welche der folgenden Aussagen zu Grafiken sind richtig? (3 Punkte)

- Ordinalskalierte Variablen lassen sich nur in Histogrammen darstellen.
- Nominalskalierte Variablen eignen sich aufgrund des fehlenden Nullpunktes nicht für die Darstellung innerhalb eines Balkendiagrammes.
- Man benötigt wenigstens zwei metrische Variablen, um ein Scatterplot zu erstellen.
- Metrische und nominalskalierte Variablen lassen sich nicht innerhalb des gleichen Diagramms darstellen.

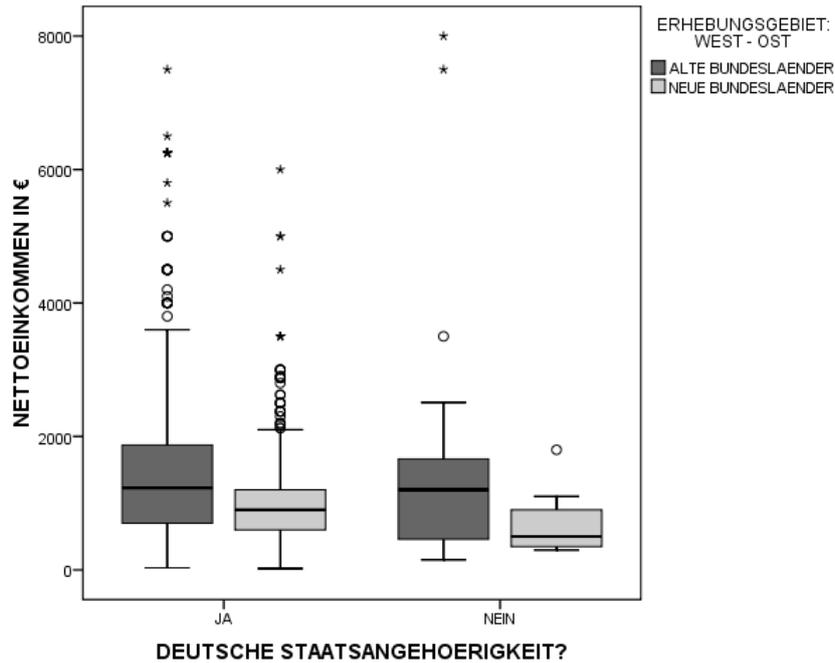
4. Ordnen Sie den Variablen eine passende Darstellungsform zu. Notieren Sie jeweils mit B, G, H oder S, ob es sich um ein Balkendiagramm, Histogramm, Gruppiertes Boxplot oder Streudiagramm handelt (3 Punkte):

- Die Schulden der Befragten (in €), geordnet nach dem Bundesland der Befragten
- Die Militärausgaben verschiedener Länder (in US-\$) in Abhängigkeit von der Alphabetisierungsrate
- Anzahl der Straftaten pro 100 Einwohner in verschiedenen Ländern in Abhängigkeit von der Arbeitslosenquote der Länder
- Die Wahlabsicht der Befragten
- Anzahl der Polizisten pro 1.000 Einwohner in verschiedenen Ländern
- Bildungsniveau der Befragten

5. Welche der folgenden Aussagen sind richtig? (3 Punkte)

- Bei einer Standardnormalverteilung befinden sich Modus, Median und arithmetisches Mittel auf dem Erwartungswert.
- Der Median ist weniger anfällig gegenüber Ausreißern und Extremwerten als das arithmetische Mittel.
- Bei einer linkssteilen Verteilung befindet sich der Median rechts vom Modus.
- Bei einer rechtsschiefen Verteilung befindet sich der Median rechts vom Modus.

6. Erläutern Sie folgende Grafik (3 Punkte):



- Es wurde ein gruppiertes Boxplot für die Verteilung des Nettoeinkommens in den alten und neuen Bundesländern, jeweils für deutsche und andere Staatsangehörige erstellt.
- Mehr als 75% aller Fälle liegen unter 2.000 €.
- Aufgrund der Quartilsgleichheit innerhalb der Boxen kann von einer Normalverteilung der Nettoeinkommen ausgegangen werden.
- Die geringste Spannweite besitzt die Verteilung des Nettoeinkommens der Befragten anderer Staatsangehörigkeit aus den neuen Bundesländern.

7. In einer Studie wurden das Geschlecht und der Besitz eines Führerscheins erfasst. Welche Aussagen sind richtig? (3 Punkte)

	männlich	weiblich
Führerschein	60	50
kein Führerschein	30	25

- Insgesamt wurden 165 Personen befragt.
- Die Kreuztabelle spricht für vollkommene statistische Unabhängigkeit der beiden Merkmale.
- Männliche Teilnehmer haben mit einer höheren Wahrscheinlichkeit einen Führerschein.
- Nach einer Berechnung von χ^2 würde dieses den Wert von 0 annehmen.

8. Aus einem Länder-Datensatz ergibt sich folgende Korrelationsmatrix. Welche Aussagen sind richtig? (3 Punkte)

Korrelationen

		Alphabetisierungsrate (%)	Bevölkerungswachstum (% pro Jahr)	Mitgliedschaft OECD	Bruttoinlandsprodukt / Kopf
Alphabetisierungsrate (%)	Korrelation nach Pearson	1,000	-,699	-,419	,552
	Signifikanz (2-seitig)		,000	,000	,000
	N	107	107	107	107
Bevölkerungswachstum (% pro Jahr)	Korrelation nach Pearson	-,699	1,000	,480	-,521
	Signifikanz (2-seitig)	,000		,000	,000
	N	107	109	109	109
Mitgliedschaft OECD	Korrelation nach Pearson	-,419	,480	1,000	-,814
	Signifikanz (2-seitig)	,000	,000		,000
	N	107	109	109	109
Bruttoinlandsprodukt / Kopf	Korrelation nach Pearson	,552	-,521	-,814	1,000
	Signifikanz (2-seitig)	,000	,000	,000	
	N	107	109	109	109

- Zwischen Alphabetisierungsrate und Bevölkerungswachstum besteht ein signifikant positiver Zusammenhang.
- Ein geringes BIP pro Kopf geht mit einem hohen Bevölkerungswachstum einher.
- Es wurden insgesamt 109 Länder untersucht.
- Zwischen BIP pro Kopf und Alphabetisierungsrate besteht ein positiver Zusammenhang, dieser ist jedoch nicht signifikant.

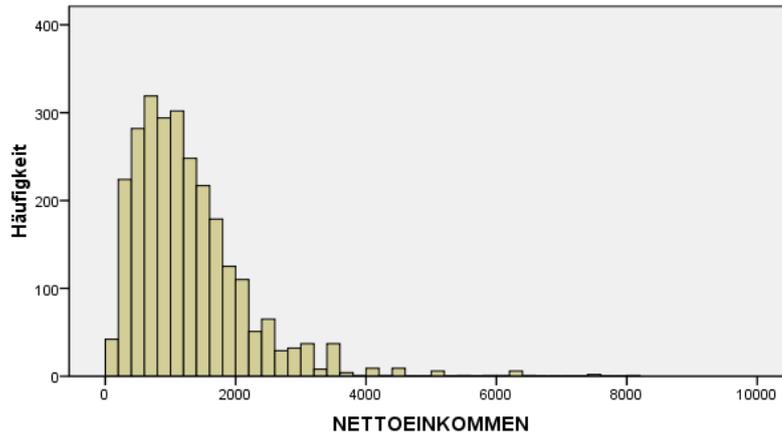
9. Welche der folgenden Aussagen zu Korrelationen sind korrekt? (3 Punkte)

- Die Spearman-Korrelation basiert auf einem Rangvergleich.
- Überschreitet Pearson's r den Wert von 0,5 muss von Kausalität ausgegangen werden.
- Wenn Korrelation vorliegt, liegt auch immer Kausalität vor.
- Spearman's r und Pearson's r können Werte von -1 bis +1 annehmen.

10. Welche der folgenden Aussagen zu Ausreißern und Extremwerten sind richtig? (3 Punkte)

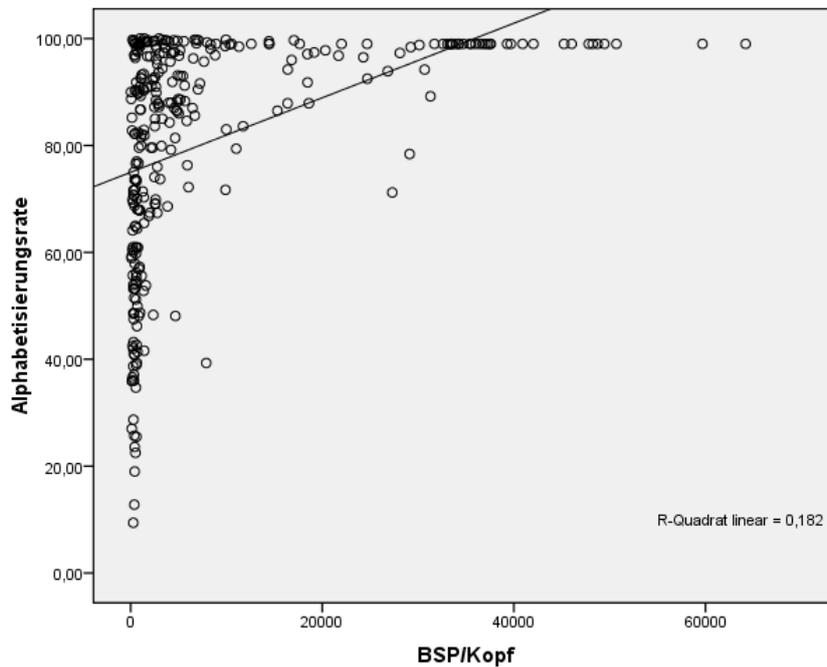
- Ein Mindestmaß an Ausreißern und Extremwerten ist notwendig, um ein Boxplot zu erstellen.
- Die Spannweite ist aufgrund ihrer Berechnung nicht anfällig gegenüber Ausreißern.
- Das Vorhandensein von Extremwerten macht eine Stichprobe ungültig, um sie in weitere Berechnungen einfließen zu lassen.
- Extremwerte befinden sich 3 Kastenlängen oder mehr von der Box entfernt.

11. Erläutern Sie die folgende Grafik (3 Punkte):



- Es handelt sich um ein Histogramm
- Aufgrund der Darstellung kann sicher von einer Normalverteilung des Nettoeinkommens ausgegangen werden.
- Die Verteilung ist rechtsschief
- Der Median liegt bei ungefähr 5.000 Euro

12. Erläutern Sie folgende Grafik (3 Punkte):



- In dem vorliegenden Histogramm wird die Verteilung des Zusammenhangs zwischen Alphabetisierungsrate und BSP pro Kopf durch Ausreißer verzerrt.
- Das R^2 deutet auf einen starken linearen Zusammenhang hin.
- R^2 deutet auf einen leichten negativen Zusammenhang von 0,182 hin.
- Die Regressionsgerade im Streudiagramm stellt den Zusammenhang beider Variablen sehr gut dar.

13. Welche der folgenden Aussagen zur linearen Regression sind richtig? (3 Punkte)

- Besonders bei multiplen linearen Regressionen mit geringer Fallzahl sollte das korrigierte R^2 statt des unkorrigierten R^2 beachtet werden.
- Multikollinearität zwischen den unabhängigen Variablen ist zu vermeiden.
- Ein negatives R^2 deutet auf einen negativen Zusammenhang zwischen den Variablen hin.
- Die Signifikanzresistenz eines Zusammenhangs kann durch die Korrelation der Variablen erklärt werden.

14. Welche Aussagen zur Teststatistik sind zutreffend? (3 Punkte)

- Je größer die Stichprobe ist, desto höher ist die Wahrscheinlichkeit, dass der Mittelwert der Stichprobe dem der Grundgesamtheit entspricht.
- Konfidenzintervalle geben wie Somer's d und Gamma den Zusammenhang ordinalskalierten Variablen an.
- Der t-Test testet den Unterschied zweier Mittelwerte. Aus diesem Grund sollten keine nominalskalierten Variablen verwendet werden, da das arithmetische Mittel ein Lagemaß für metrische Variablen ist.
- Signifikante Werte werden in Tabellen häufig mit Sternen gekennzeichnet. Anschaulicher wäre es jedoch, signifikante Werte in einer Boxplot-Grafik darzustellen.

15. Welche Aussagen über den folgenden SPSS-Ausdruck sind richtig? (3 Punkte)

Gruppenstatistiken

		N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
NETTOEINKOMMEN	MANN	1344	1526,95	982,734	26,806
	FRAU	1300	963,43	622,979	17,278

Test bei unabhängigen Stichproben

	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit				
	F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz
	NETTOEINKOMMEN Varianzen sind gleich	125,997	,000	17,544	2642	,000	563,516
NETTOEINKOMMEN Varianzen sind nicht gleich			17,669	2283,280	,000	563,516	31,892

- Der Levene-Test prüft, ob die Varianzen der Stichproben gleich sind oder nicht. Da die t-Werte in der Tabelle sehr hoch sind, liegt Varianzgleichheit vor. Dies nennt man Homoskedastizität.
- Bei diesem SPSS-Ausdruck wird geprüft, ob die Einkommensunterschiede zwischen Männer und Frauen signifikant sind.
- Der Levene-Test hat bewiesen, dass die Varianzen der beiden Stichproben gleich sind. Diesen Schluss lässt auch die Gruppenstatistik zu.
- Hier wird ein t-Test auf den Unterschied zweier Mittelwerte durchgeführt.

16. Folgende lineare Regression soll den Zusammenhang zwischen BSP pro Kopf und der Lebenserwartung aufzeigen. Welche Aussagen sind richtig? (4 Punkte)

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,598 ^a	,357	,355	8,9281

a. Einflußvariablen : (Konstante), BSP/Kopf

ANOVA^b

Modell		Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
1	Regression	13693,672	1	13693,672	171,792	,000 ^a
	Residuen	24630,582	309	79,711		
	Gesamt	38324,254	310			

a. Einflußvariablen : (Konstante), BSP/Kopf

b. Abhängige Variable: Lebenserwartung

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	61,911	,599		103,436	,000
	BSP/Kopf	,001	,000	,598	13,107	,000

a. Abhängige Variable: Lebenserwartung

- Die Lebenserwartung soll durch das BSP pro Kopf erklärt werden.
- Die Ergebnisse der F-Statistik lassen auf einen linearen Zusammenhang schließen.
- Der Determinationskoeffizient gibt den Anteil der durch die unabhängige Variable erklärten Varianz der abhängigen Variable an.
- Es besteht kein signifikanter Zusammenhang zwischen dem BSP pro Kopf und der Lebenserwartung.