

SPEZIELLE MASSZAHLEN I:ETA UND ETA²

Martin-Luther-Universität Halle-Wittenberg

Institut für Soziologie

Übung Einführung in die deskriptive Statistik

Agenda

- Einfaktorielle Varianzanalyse: Eta und Eta²
 - Verwendung
 - PRE-Logik
 - Berechnung aus Rohdaten
 - Berechnung mithilfe von SPSS-Outputs
 - Interpretation

Fallbeispiel: Shoppingdauer

Ein Marktforscher möchte wissen, ob die Shoppingdauer eines Befragten von seinem Geschlecht abhängt. Hierbei steht bei dem Geschlecht der Wert 0 für Frauen und 1 für Männer. Für 10 Befragte erhält er hierbei folgende Werte:

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	0	0	0	0	1	1	1	1	1
Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103

Quiz: Welches Skalenniveau haben die beiden Variablen? Handelt es sich um eine symmetrische oder asymmetrische Fragestellung?

Fallbeispiel: Analyse

asymmetrische
Fragestellung:
Geschlecht(X) →
Einkaufszeit(Y)

Ein Marktforscher möchte wissen, ob die Shoppingdauer eines Befragten von seinem Geschlecht abhängt.

Geschlecht: nominal
Einkaufsdauer: metrisch

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	0	0	0	0	1	1	1	1	1
Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103

geeignetes
Zusammenhangsmaß:
Eta/Eta²

Einfaktorielle Varianzanalyse: η^2 und η

- Verwendung:

- bei einer mindestens nominalen unabhängigen Variablen (X) und einer metrischen abhängigen Variablen (Y)
- unabhängige Variable kann mehr als zwei Ausprägungen haben

- Berechnung:

- $\eta^2 = PRE = \frac{E_0 - E_1}{E_0}$

- $\eta = \sqrt{\eta^2}$



Wie bestimmen sich hier
E0- und E1-Fehler?

Eta und Eta²: E_0 -Fehler (Gesamtvariation)

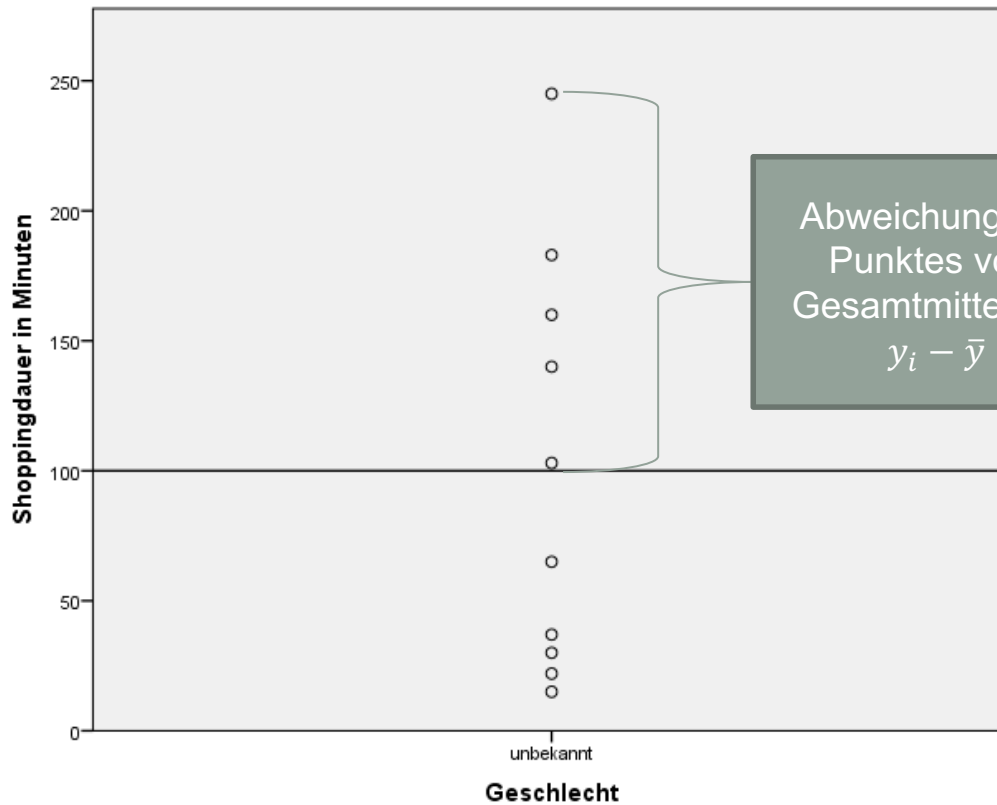
E_0 -Fehler: Fehler ohne Kenntnis der unabhängigen X-Variablen

Person	1	2	3	4	5	6	7	8	9	10
Y: Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103

Quiz: Welcher Wert, wäre hier der beste Tipp, wenn man den Fehler, den man machen will möglichst gering halten will?

Quiz: Mit welchem Maß lässt sich die entsprechende Abweichung beschreiben?

Eta und Eta²: E_0 -Fehler II (Gesamtvariation)



Abweichung des
Punktes vom
Gesamtmittelwert
 $y_i - \bar{y}$

bester Tipp:
Gesamtmittelwert \bar{y}

Bestimmung E_0 -Fehler:

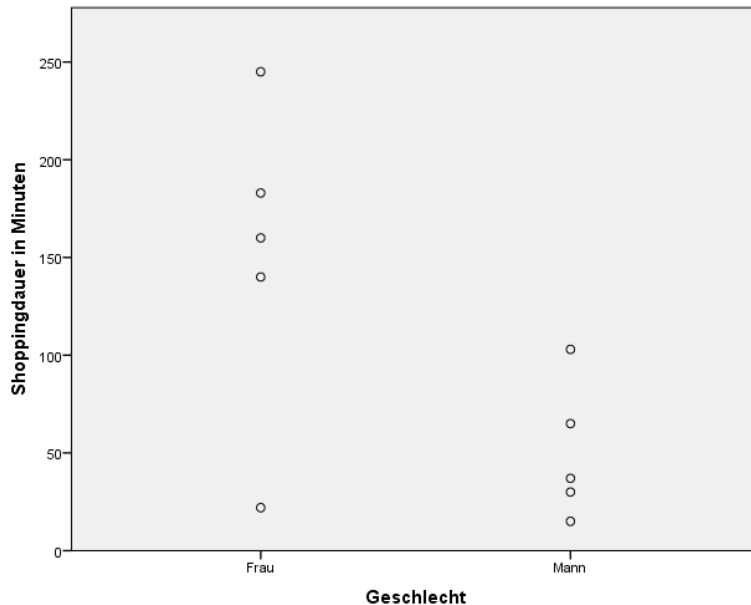
$$E_0 = SAQ_Y = \sum_{i=1}^n (y_i - \bar{y})^2$$

Eta und Eta²: E_1 -Fehler I

(Fehlervariation bzw. nicht-erklärte Variation)

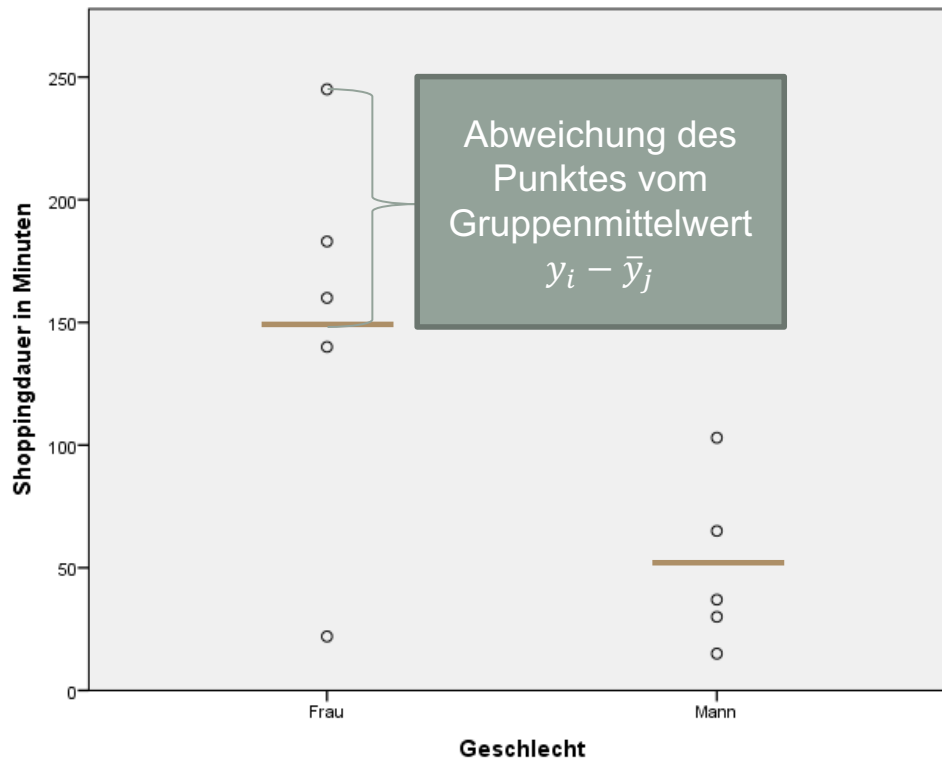
E_1 -Fehler: Fehler mit Kenntnis der unabhängigen X-Variablen

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	0	0	0	0	1	1	1	1	1
Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103



Quiz: Was wäre hier der beste Tipp, wenn wir wissen, dass ein Befragter eine Frau ist? Was wäre der beste Tipp für die Shoppingdauer von Männern?

Eta und Eta²: E₁-Fehler (Fehlervariation bzw. nicht-erklärte Variation)



Was wäre hier der beste Tipp für die Zeit, wenn wir das Geschlecht kennen?

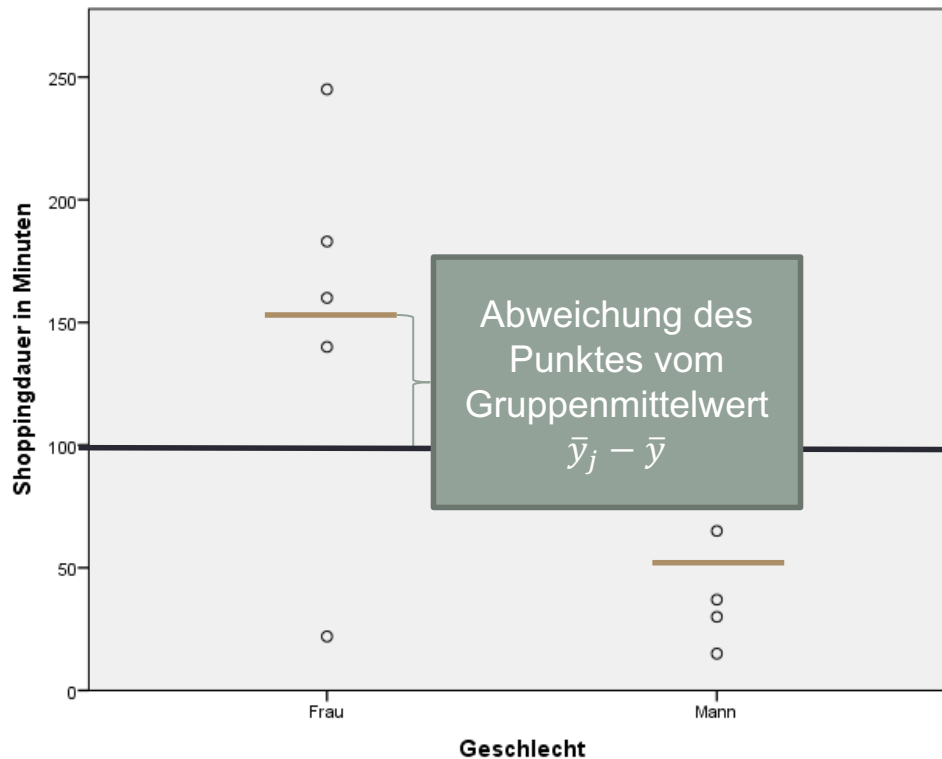
Lösung:
Gruppenmittelwerte \bar{y}_j

Wie lässt sich damit der E₁-Fehler bestimmen?

$$E_1 = \sum_{j=1}^J \sum_{i=1}^{n_j} (y_{j,i} - \bar{y}_j)^2$$

Summe der Variationen aller Gruppen
(hier: Mann und Frau)

Eta und Eta²: $E_0 - E_1$ (erklärte Variation)



$$E_0 - E_1 = SAQ_{zwischen}$$

$$E_0 - E_1 = \sum_{j=1}^J n_j * (\bar{y}_j - \bar{y})^2$$

Summe der Variationen
zwischen den Gruppen

Eta und Eta²: Berechnung und Interpretation I

- Berechnung Eta²:

- $\eta^2 = PRE = \frac{E_0 - E_1}{E_0}$

- $E_0 = SAQ_{Gesamt} = SAQ_Y = \sum_{i=1}^n (y_i - \bar{y})^2$

- $E_1 = SAQ_{innerhalb} = \sum_{j=1}^J \sum_{i=1}^{n_j} (y_{j,i} - \bar{y}_j)^2$

- $E_0 - E_1 = \sum_{j=1}^J n_j * (\bar{y}_j - \bar{y})^2$

E0-Fehler:
Gesamtvariation

E1-Fehler :
Variation innerhalb der
Gruppen

E0-E1 :
Variation zwischen den
Gruppen

- Interpretation Eta²:

- PRE-Maß!

- Durch Kenntnis der unabhängigen Variablen verbessert sich die Prognose der abhängigen Variablen um ... %

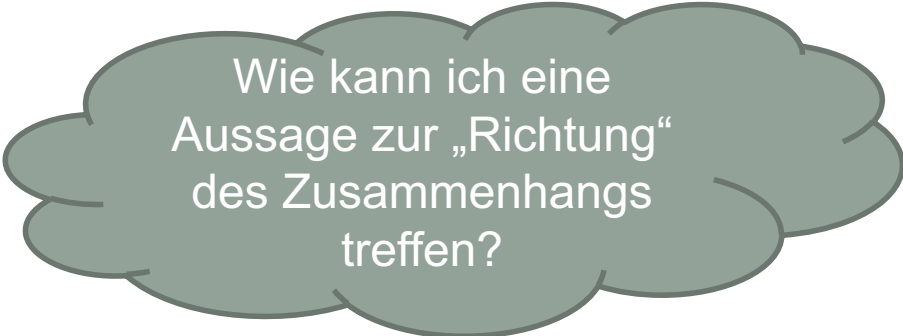
Eta und Eta²: Berechnung und Interpretation II

- Berechnung Eta:

- $\eta = \sqrt{\eta^2}$

- Interpretation Eta:

- normales Zusammenhangsmaß
 - Es besteht kein / ein niedriger / mittlerer / starker Einfluss der X-Variablen auf die y-Variablen.



Wie kann ich eine Aussage zur „Richtung“ des Zusammenhangs treffen?

Eta und Eta²: Berechnung und Interpretation III

- „Richtung“:
 - Vergleich der Mittelwerte der Variablen
 - Der Mittelwert der Gruppe 1 ist niedriger / gleich / höher als der Mittelwert der Gruppe 2.

Fallbeispiel: Shoppingdauer

Ein Marktforscher möchte wissen, ob die Shoppingdauer eines Befragten von seinem Geschlecht abhängt. Für 10 Befragte erhält er hierbei folgende Werte:

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	0	0	0	0	1	1	1	1	1
Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103

- Berechnen Sie η^2 / η als geeignetes Zusammenhangsmaß.
- Interpretieren Sie Ihr Ergebnis vollständig. Gehen Sie hierbei auf η^2 , η und die Mittelwerte ein und geben Sie eine inhaltliche Interpretation.

Fallbeispiel Shoppingdauer II

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	0	0	0	0	1	1	1	1	1
Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103

- Berechnung:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$

- $E_0 = SAQ_Y = SAQ_{Gesamt} = \sum_{i=1}^n (y_i - \bar{y})^2$

- $E_1 = SAQ_{innerhalb} = \sum_{j=1}^J \sum_{i=1}^{n_j} (y_{j,i} - \bar{y}_j)^2$

Berechnen Sie zunächst den E0-Fehler. Bilden Sie hierzu zunächst den Gesamtmittelwert \bar{y} und berechnen Sie anschließend die Gesamtvariation SAQ_Y .

Fallbeispiel Shoppingdauer III

Person	y_i
1	22
2	140
3	160
4	183
5	245
6	15
7	30
8	37
9	65
10	103
	$\bar{y} = 100$

Fallbeispiel Shoppingdauer III

Person	y_i	$(y_i - \bar{y})^2$
1	22	6084
2	140	1600
3	160	3600
4	183	6889
5	245	21025
6	15	7225
7	30	4900
8	37	3969
9	65	1225
10	103	9
	$\bar{y} = 100$	$SAQ_Y = 56526$

$$E_1 = SAQ_Y = SAQ_{gesamt} = 56526$$

Fallbeispiel Shoppingdauer IV

Person	1	2	3	4	5	6	7	8	9	10
Geschlecht	0	0	0	0	0	1	1	1	1	1
Einkaufsdauer in Minuten	22	140	160	183	245	15	30	37	65	103

- Berechnung:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$

- $E_0 = 56526$

- $E_1 = SAQ_{innerhalb} = \sum_{j=1}^J \sum_{i=1}^{n_j} (y_{j,i} - \bar{y}_j)^2$

Berechnen Sie nun den E1-Fehler. Bilden Sie hierzu zunächst die gruppenspezifischen Mittelwerte für Männer und Frauen und bestimmen Sie für beide Gruppen die Variation. Addieren Sie abschließend beide Werte.

Fallbeispiel Shoppingdauer V

Gruppe 1: Frauen

Person	$y_{\text{Frau},i}$	$(y_i - \bar{y}_{\text{Frau}})^2$
1	22	16384
2	140	100
3	160	100
4	183	1089
5	245	9025
	$\bar{y}_{\text{Frau}} = 150$	$SAQ_{\text{Frau}} = 26698$

Gruppe 2: Männer

Person	$y_{\text{Mann},i}$	$(y_i - \bar{y}_{\text{Mann}})^2$
6	15	1225
7	30	400
8	37	169
9	65	225
10	103	2809
	$\bar{y}_{\text{Mann}} = 50$	$SAQ_{\text{Mann}} = 4828$

- $E_1 = SAQ_{\text{innerhalb}} = \sum_{j=1}^J \sum_{i=1}^{n_j} (y_{j,i} - \bar{y}_j)^2 = SAQ_{\text{Frau}} + SAQ_{\text{Mann}}$
- $E_1 = 26698 + 4828 = 31526$

Fallbeispiel Shoppingdauer VI

- Alternativ: Berechnung der erklärten Variation (E0-E1):

- $E_0 - E_1 = SAQ_{zwischen} = \sum_{j=1}^J (n_j * (\bar{y}_j - \bar{y})^2)$

j	Bezeichnung	n_j	\bar{y}_j	$\bar{y}_j - \bar{y}$	$(\bar{y}_j - \bar{y})^2$	$n_j \cdot (\bar{y}_j - \bar{y})^2$
1	Frau	5	150	+50	2500	12500
2	Männer	5	50	-50	2500	12500
Gesamt	Gesamt	10	100			$E_0 - E_1 = 25000$

Fallbeispiel Shoppingdauer VI

- gegeben:

- $\bar{y}_{\text{Frau}} = 150$

- $\bar{y}_{\text{Mann}} = 50$

- $E_0 = 56526$

- $E_1 = 31526$

- Formeln:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$

- $\eta = \sqrt{\eta^2}$

Fallbeispiel Shoppingdauer VII

- Berechnung:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$

- $\eta^2 = \frac{56526 - 31526}{56526} = \frac{25000}{56526}$

- $\eta^2 = 0,4427$

- $\eta = \sqrt{\eta^2}$

- $\eta = 0,6650$

$$\bar{y}_{\text{Frau}} = 150$$

$$\bar{y}_{\text{Mann}} = 50$$

$$E_0 = 56526$$

$$E_1 = 31526$$

Fallbeispiel Shoppingdauer VIII

- Interpretation η^2 :
 - $\eta^2 = 0,4427$
 - Durch Kenntnis des Geschlechts lässt sich im vorliegenden Beispiel 44,3% der Streuung der Shoppingdauer erklären.
- Interpretation η :
 - $\eta = 0,6650$
 - Es besteht ein hoher Einfluss des Geschlechts auf die Shoppingdauer.
- Interpretation Mittelwerte:
 - Es fällt auf, dass Frauen mit durchschnittlich 150 Minuten deutlich länger einkaufen als Männer mit durchschnittlich 50 Minuten.

SPSS: Eta² über ANOVA-Tabelle

ANOVA-Tabelle

		Quadratsumme	df	Quadratische		
Shoppingdauer in Minuten * Geschlecht	Zwischen Gruppen (Kombiniert)	25000,000	1	25000,000		
	Innerhalb der Gruppen	31526,000	8	3940,750		
	Gesamtsumme	56526,000	9	6240,750		

$$E_0 - E_1 = SAQ_{\text{zwischen}}$$

$$E_1 = SAQ_{\text{innerhalb}}$$

$$E_0 = SAQ_{\text{gesamt}}$$

Zusammenhangsmaße

	Eta	Eta-Quadrat
Shoppingdauer in Minuten * Geschlecht	,665	,442

$$\eta^2 = \frac{E_0 - E_1}{E_0}$$

$$\eta = \sqrt{\eta^2}$$

Exkurs: Zusammenhang Varianzanalyse mit zwei Gruppen und bivariate Dummyregression

ANOVA-Tabelle

		Quadratsumme	df	Quadratischer Mittelwert	F	Sig.
Shoppingdauer in Minuten * Geschlecht	Zwischen Gruppen (Kombiniert)	25000,000	1	25000,000	6,344	,036
	Innerhalb der Gruppen	31526,000	8	3940,750		
	Gesamtsumme	56526,000	9			

Zusammenhangsmaße

	Eta	Eta-Quadrat
Shoppingdauer in Minuten * Geschlecht	,665	,442

Modellübersicht

Modell	R	R-Quadrat	Angepasstes R-Quadrat	Standardfehler der Schätzung
1	,665 ^a	,442	,373	62,775

a. Prädiktoren: (Konstante), Geschlecht

ANOVA^a

Modell		Quadratsumme	df	Mittel der Quadrate	F	Sig.
1	Regression	25000,000	1	25000,000	6,344	,036 ^b
	Residuum	31526,000	8	3940,750		
	Gesamtsumme	56526,000	9			

a. Abhängige Variable: Shoppingdauer in Minuten

b. Prädiktoren: (Konstante), Geschlecht

Bei beiden stimmen die Werte in den ANOVA-Tabellen überein, ebenso entspricht Eta-Quadrat immer R-Quadrat.

Pearsons r und Eta sind vom Betrag her identisch, allerdings ist zu beachten, dass Eta kein Vorzeichen aufweist, also immer positiv ist, während Pearson r eine Richtungsangabe beinhaltet.

Exkurs: Zusammenhang Varianzanalyse mit zwei Gruppen und Dummyregression II

Koeffizienten^a

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	t	Sig.
		B	Standardfehler	Beta		
1	(Konstante)	150,000	28,074		5,343	,001
	Geschlecht	-100,000	39,703	-,665	-2,519	,036

Geschlecht:
0 – Frau
1 – Mann

a. Abhängige Variable: Shoppingdauer in Minuten

Quiz: Was sagte uns hier nochmal die Regressionskonstante und das Regressionsgewicht? Wie würde das Modell aussehen, wenn wir die Codierung von Mann und Frau tauschen, also 0=Männer, 1=Frauen?

Wiederholung: Dummyregression

Koeffizienten^a

Modell	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	t	Sig.
	B	Standardfehler	Beta		
1 (Konstante)	50,000	28,074		1,781	,113
rekodiertes Geschlecht	100,000	39,703	,665	2,519	,036

a. Abhängige Variable: Shoppingdauer in Minuten

Geschlecht:
0 – Mann
1 – Frau

SPSS: Mittelwerte, Varianzen und Fallzahlen für Gruppen bestimmen

Shoppingdauer in Minuten

Geschlecht	Mittelwert	Fallzahl	Varianz
Frau	150,00	5	6674,500
Mann	50,00	5	1207,000
Gesamtsumme	100,00	10	6280,667

arithmetisches Mittel

Fallzahl

$$\hat{\sigma}_X^2 = \frac{(x_i - \bar{x})^2}{n-1}$$

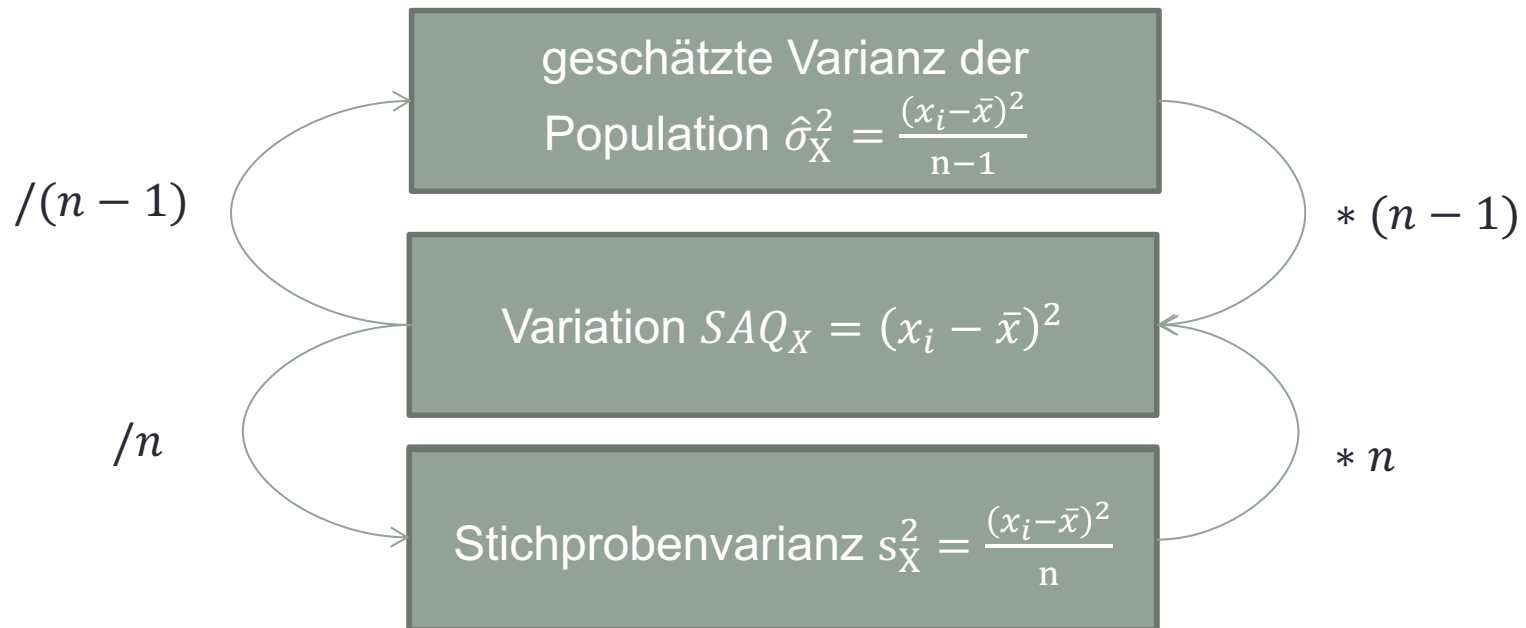
geschätzte
Populationsvarianz

Achtung: Da wir uns in der deskriptiven Statistik befinden haben wir

bisher die Stichprobenstandardabweichung $s_X = \hat{\sigma}_X * \frac{n-1}{n} = \sqrt{\frac{(x_i - \bar{x})^2}{n}}$

verwendet.

FAQ: Warum mit (n-1) multiplizieren?



Aufgabe 1: Klausurpunktzahl

Ein Qualitätssicherungsbeauftragter einer Universität interessiert sich dafür, ob es einen Einfluss des Dozenten, bei dem die Übung belegt wurde, auf die erzielte Punktzahl in der Klausur gibt. Mithilfe von SPSS erhält er folgende Tabelle:

Punktzahl

Dozent	Mittelwert	H	Varianz
Dozent 1	66,00	10	1124,667
Dozent 2	46,00	10	669,556
Gesamtsumme	56,00	20	955,158

- Welches Maß ist hier geeignet?
- Berechnen Sie das Maß!
- Interpretieren Sie das Maß vollständig!

Achtung: Die Variation ergibt sich hier durch $(\text{Fallzahl}-1) \cdot \text{Varianz}$, da wir es mit dem Schätzer für die Grundgesamtheit zu tun haben.

Aufgabe 1a: Analyse

asymmetrische
Fragestellung:
Methode(X) →
Punktzahl(Y)

Ein Qualitätssicherungsbeauftragter einer Universität interessiert sich dafür, ob es einen Einfluss des Dozenten, bei dem die Übung belegt wurde, auf die erzielte Punktzahl in der Klausur gibt.

Methode: nominal
Punktzahl: metrisch

Punktzahl

Dozent	Mittelwert	H	Varianz
Dozent 1	66,00	10	1124,667
Dozent 2	46,00	10	669,556
Gesamtsumme	56,00	20	955,158

geeignetes
Zusammenhangsmaß:
Eta/Eta²

Aufgabe 1b: Lösung

- Berechnung:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$
- $E_0 = SAQ_Y = SAQ_{gesamt}$
- $E_0 = (n_{gesamt} - 1) * \hat{\sigma}_{gesamt}^2$
- $E_0 = (20 - 1) * 955,158$
- **$E_0 = 18148,002$**

Punktzahl

Dozent	Mittelwert	H	Varianz
Dozent 1	66,00	10	1124,667
Dozent 2	46,00	10	669,556
Gesamtsumme	56,00	20	955,158

Aufgabe 1b: Lösung II

- Berechnung:

- $E_1 = SAQ_{Dozent1} + SAQ_{Dozent2}$

- $E_1 = \left((n_{Dozente1} - 1) * \hat{\sigma}_{Dozent1}^2 \right) + \left((n_{Dozent2} - 1) * \hat{\sigma}_{Dozent2}^2 \right)$

- $SAQ_{Dozent1} = (10 - 1) * 1124,667$

- $SAQ_{Dozent1} = 10122,003$

- $SAQ_{Dozent2} = (10 - 1) * 669,556$

- $SAQ_{Dozent2} = 6026,0004$

- $E_1 = 10122,003 + 6026,004$

- $E_1 = \mathbf{16148,007}$

Punktzahl

Dozent	Mittelwert	H	Varianz
Dozent 1	66,00	10	1124,667
Dozent 2	46,00	10	669,556
Gesamtsumme	56,00	20	955,158

Aufgabe 1b: Lösung III

- Berechnung η^2

- $\eta^2 = PRE = \frac{E_0 - E_1}{E_0}$

- $\eta^2 = \frac{18148,002 - 16148,007}{18148,002}$

- $\eta^2 = 0,110205$

- Berechnung η

- $\eta = \sqrt{\eta^2}$

- $\eta = 0,332$

- $E_0 = 18148,002$
- $E_1 = 16148,007$

Aufgabe 1c: Lösung

- Interpretation η^2
 - Durch Kenntnis des Dozenten verbessert sich die Prognose der Klausurpunktzahl um 11,02%
- Interpretation η :
 - Es besteht ein mittlerer Einfluss des Dozenten auf die Klausurpunktzahl.
- Interpretation Mittelwerte:
 - Anhand der Mittelwerte wird deutlich, dass die Klausurpunktzahl bei Dozenten 1 im Durchschnitt 20 Punkte höher liegt als beim 2. Dozenten. (66 bzw. 46 Punkte)

- $\eta^2 = 0,110205$
- $\eta = 0,332$
- $\bar{y}_{Dozent1} = 66 \text{ Punkte}$
- $\bar{y}_{Dozent2} = 46 \text{ Punkte}$

Aufgabe 1c: SPSS

ANOVA-Tabelle

		Quadratsumme	df	Quadratischer Mittelwert	F	Sig.
Punktzahl * Dozent	Zwischen Gruppen (Kombiniert)	2000,000	1	2000,000	2,229	,153
	Innerhalb der Gruppen	16148,000	18	897,111		
	Gesamtsumme	18148,000	19			

Zusammenhangsmaße

	Eta	Eta-Quadrat
Punktzahl * Dozent	,332	,110

Aufgabe 2: Geschlecht und Arbeitszeit

Eine Arbeitsmarktforscher möchte wissen, ob das Geschlecht der Befragten einen Einfluss auf die wöchentliche Arbeitszeit in Stunden ausübt. Mithilfe von Daten aus dem ALLBUS 2014 erhält er folgende Tabelle:

GESCHLECHT, BEFRAGTE<R>	Mittelwert	H	Varianz
MAENNLICH	43,7907	1063	87,075
WEIBLICH	34,7169	847	124,839
Gesamtsumme	39,7669	1910	124,095

- Welches Maß ist hier geeignet?
- Berechnen Sie das Maß!
- Interpretieren Sie das Maß vollständig!

Aufgabe 2a: Analyse

asymmetrische
Fragestellung:
Geschlecht (X) →
Arbeitszeit (Y)



Eine Arbeitsmarktforscher möchte wissen, ob das Geschlecht der Befragten einen Einfluss auf die wöchentliche Arbeitszeit in Stunden ausübt. Mithilfe von Daten aus dem ALLBUS 2014 erhält er folgende Tabelle

Geschlecht: nominal
Arbeitszeit: metrisch



GESCHLECHT, BEFRAGTE <R>	Mittelwert	H	Varianz
MAENNLICH	43,7907	1063	87,075
WEIBLICH	34,7169	847	124,839
Gesamtsumme	39,7669	1910	124,095

geeignetes
Zusammenhangsmaß:
 Eta/Eta^2



Aufgabe 2b: Lösung

GESCHLECHT, BEFRAGTE<R>	Mittelwert	H	Varianz
MAENNLICH	43,7907	1063	87,075
WEIBLICH	34,7169	847	124,839
Gesamtsumme	39,7669	1910	124,095

- Berechnung E_0 :
 - $E_0 = SAQ_{Gesamt} = 124,095 * (1910 - 1) = 236.897,355$
- Berechnung E_1 :
 - $E_1 = SAQ_{Mann} + SAQ_{Frau}$
 - $SAQ_{Mann} = 87,075 * (1063 - 1) = 92.473,65$
 - $SAQ_{Frau} = 124,839 * (846) = 105.613,794$
 - $E_1 = 198.087,444$

Aufgabe 2b: Lösung II

- Berechnung:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$

- $\eta^2 = \frac{236.897,355 - 198.087,444}{236.897,355}$

- $\eta^2 = 0,1638$

- $\eta = 0,405$

$$E_0 = 236.897,355$$

$$E_1 = 198.087,444$$

Aufgabe 2c: Lösung

- Interpretation η^2/η :
 - η beträgt 0,405. Dies bedeutet, dass ein mittlerer Einfluss des Geschlechts auf die wöchentliche Arbeitszeit in Stunden besteht.
 - η^2 liegt bei 0,163. Durch Kenntnis des Geschlechts lässt sich die Prognose der wöchentlichen Arbeitszeit einer Person um 16,3 % verbessern.

Aufgabe 2c: Lösung II

- Interpretation Mittelwerte:
 - Die durchschnittliche Arbeitszeit der Männer ist mit 43,8 Wochenstunden deutlich höher als die durchschnittliche Arbeitszeit der Frauen mit 34,7 Wochenstunden.
 - Wenn man sich die Varianzen beider Ausprägungen anschaut, fällt auf, dass die Abweichung in den wöchentlichen Arbeitszeiten bei den Frauen höher ist als bei den Männern.

GESCHLECHT, BEFRAGTE<R>	Mittelwert	H	Varianz
MAENNLICH	43,7907	1063	87,075
WEIBLICH	34,7169	847	124,839
Gesamtsumme	39,7669	1910	124,095

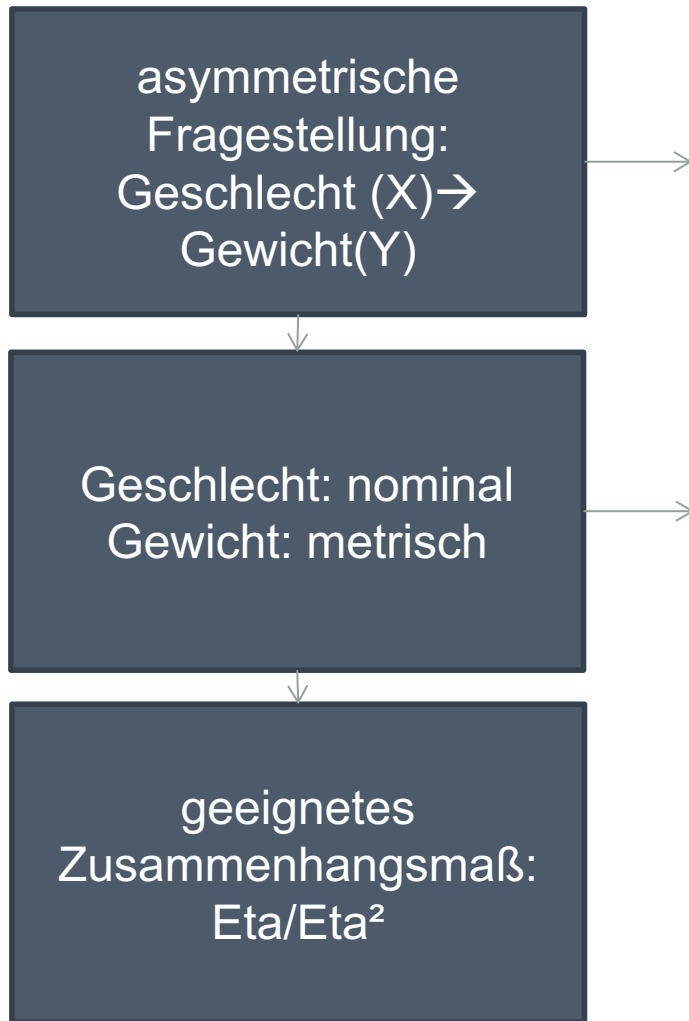
Aufgabe 3: Geschlecht und Gewicht

Eine Gesundheitsforscherin möchte wissen, ob das Geschlecht der Befragten einen Einfluss auf das Gewicht in kg ausübt. Mithilfe von Stata erhält sie für Daten aus dem ALLBUS 2014 folgende Tabelle:

GESCHLECHT, BEFRAGTE<R>	Mittelwert	H	Standardabw eichung
MAENNLICH	85,30	1756	15,113
WEIBLICH	70,17	1665	14,378
Gesamtsumme	77,94	3421	16,584

- Welches Maß ist hier geeignet?
- Berechnen Sie das Maß!
- Interpretieren Sie das Maß vollständig!

Aufgabe 3a: Analyse



Eine Gesundheitsforscherin möchte wissen, ob das Geschlecht der Befragten einen Einfluss auf das Gewicht in kg ausübt.

GESCHLECHT, BEFRAGTE <R>	Mittelwert	H	Standardabweichung
MAENNLICH	85,30	1756	15,113
WEIBLICH	70,17	1665	14,378
Gesamtsumme	77,94	3421	16,584

Aufgabe 3b: Lösung

GESCHLECHT, BEFRAGTE<R>	Mittelwert	H	Standardabw eichung
MAENNLICH	85,30	1756	15,113
WEIBLICH	70,17	1665	14,378
Gesamtsumme	77,94	3421	16,584

- Berechnung E_0 :

- $E_0 = SAQ_{Gesamt} = 16,584^2 * (3421 - 1) = 940.599,3715$

- Berechnung E_1 :

- $E_1 = SAQ_{Mann} + SAQ_{Frau}$

- $SAQ_{Mann} = 15,113^2 * (1756 - 1) = 400.846,8596$

- $SAQ_{Frau} = 14,378^2 * (1665 - 1) = 343.993,535$

- $E_1 = 744.840,3946$

Aufgabe 3b: Lösung II

- Berechnung:

- $\eta^2 = \frac{E_0 - E_1}{E_0}$

- $\eta^2 = \frac{940.599,3715 - 744.840,3946}{940.599,3715}$

- $\eta^2 = 0,2081$

- $\eta = 0,456$

$$E_0 = 940.599,3715$$

$$E_1 = 744.840,3946$$

Zusammenhangsmaße

	Eta	Eta-Quadrat
GEWICHT IN KG, BEFRAGTE<R> * GESCHLECHT, BEFRAGTE<R>	,456	,208

Aufgabe 3c: Lösung

- Interpretation η^2/η :
 - η beträgt 0,456. Dies bedeutet, dass ein mittlerer Einfluss des Geschlechts auf das Gewicht besteht.
 - η^2 liegt bei 0,208. Durch Kenntnis des Geschlechts lässt sich die Prognose der Gewichts um 20,8 % verbessern.

Aufgabe 3c: Lösung II

- Interpretation Mittelwerte:
 - Betrachtet man die Mittelwerte, so fällt auf das Männer im Durchschnitt mit 85,3 kg etwa 15,1 kg schwerer sind als Frauen, die im Durchschnitt 70,2 kg wiegen.
 - Wenn man sich die Standardabweichungen beider Variablen anschaut, fällt auf, dass das Gewicht bei dem Männern etwas mehr streut als bei den Frauen.

GESCHLECHT, BEFRAGTE<R>	Mittelwert	H	Standardabw eichung
MAENNLICH	85,30	1756	15,113
WEIBLICH	70,17	1665	14,378
Gesamtsumme	77,94	3421	16,584

Literaturhinweise

- Kerstin Völkl / Christoph Korb (2018): Deskriptive Statistik. Eine Einführung für Politikwissenschaftlerinnen und Politikwissenschaftler. S. 234-243.
- Hans Benninghaus (2007): Deskriptive Statistik. Eine Einführung für Sozialwissenschaftler. S. 228-250.