

Übungsaufgabe 2 Milena Rapp

1) Erläutern Sie exemplarisch wozu b^* benutzt wird und wie man diesen inhaltlich interpretiert.

B^* ist der standardisierte Regressionskoeffizient b . Er umfasst einen Wertebereich von -1 bis 1. Die Standardisierung wird wie folgt durchgeführt: der Regressionskoeffizient wird mit dem Quotienten der Standardabweichung von X und der Standardabweichung von Y multipliziert. Bei bivariaten Zusammenhängen entspricht b^* Pearsons r . Mit b^* kann die einflussreichste Variable innerhalb eines Modells identifiziert werden. Die Regressionskoeffizienten werden innerhalb eines Modells hinsichtlich ihrer Stärke vergleichbar, da die Standardisierung eine Betrachtung unabhängig von der individuellen Skalenbreite ermöglicht. Eine Interpretation analog zum Regressionskoeffizienten b ist möglich (jedoch in Standardabweichungen), inhaltlich jedoch nicht wirklich sinnvoll, da die Angabe mit Standardabweichungen nicht intuitiv verständlich ist.

Modell 3			
	b	se	b^*
Alter	0,040***	0,005	0,140
Bildung	1,244***	0,069	0,302
Geschlecht	3,558***	0,160	0,359
N	3039		
R^2	0,211		
Korr. R^2 Sig. Gesamtmodell	0,210***		

Exemplarisch kann hier Modell 3 herangezogen werden. Es zeigt sich, dass das Geschlecht den stärksten Einfluss auf das Einkommen aufweist, das Alter den geringsten Einfluss. Der Regressionskoeffizient b kann nicht verglichen werden, da die drei im Modell enthaltenen Variablen unterschiedliche Skalenbreiten aufweisen.

2)

Führen Sie eine z-Standardisierung für die Originalaltersvariable ($alter_z$) und die auf Null gesetzte Altersvariable $alter0_z$) sowie für „unsere“ Bildungsvariable (0-4).

2.a) Vergleichen Sie die Zahlenwerte, Mean und die Standardabweichung von $alter_z$ und $alter0_z$ und erklären Ihre „Beobachtung“.

Statistiken

		z-Faktorwert:
		Alter mit 18 als Nullpunkt
		Faktorwert(alter)
N	Gültig	3468
	Fehlend	3
Mittelwert		,0000000
Median		,0319708
Standardabweichung		1,0000000
Schiefe		,062
Kurtosis		-,892
Minimum		-1,79595
Maximum		2,37399

Die Kennwerte unterscheiden sich nicht. Durch die z-Transformation ist das arithmetische Mittel der transformierten Variablen immer Null und die Varianz und Standardabweichung immer eins. Da es sich hier um die gleiche Ursprungsvariable handelt, die lediglich umkodiert wurde, weisen die beiden z-transformierten Variablen auf allen Kennwerten die gleichen Zahlen auf.

2.b)

Führen Sie eine Regression von Einkommen auf Alter_0 und Bildung (Modell 1) und eine Regression von Einkommen auf alter0_z und bildung_z (Modell 2) durch und vergleichen Sie die b-Koeffizienten.

MODELL 1

Koeffizienten^a

		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten		
		Regressionskoeffizient B	Standardfehler	Beta	T	Sig.
1	(Konstante)	7,165	,282		25,390	,000
	Alter mit 18 als Nullpunkt	,039	,005	,135	7,496	,000
	Abschluss umkodiert	1,199	,074	,291	16,124	,000

a. Abhängige Variable: NETTOEINKOMMEN<OFFENE+LISTENANGABE>,KAT.

MODELL 2

Koeffizienten ^a					
Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	Sig.
		Regressionskoeffizient B	Standardfehler	Beta	
1	(Konstante)	,005	,017		,794
	z-Faktorwert: Alter mit 18 als Nullpunkt	,135	,018	,135	,000
	z-Faktorwert: Abschluss umkodiert	,291	,018	,291	,000

a. Abhängige Variable: z-Faktorwert(einkommen)

Aufgrund der unterschiedlichen Skalenbreite (Ursprungsskala vs. Z-standardisierte Skala) weisen die b-Koeffizienten der beiden Modelle unterschiedliche Werte auf. Mit jedem zusätzlichen Altersjahr steigt das Einkommen um 0,039 Einheiten auf der Einkommensskala an, mit jedem nächst höherem Bildungsabschluss steigt das Einkommen um 1,199 Einheiten auf der Einkommensskala an.

Die b-Koeffizienten von Modell 2 entsprechen den b*-Koeffizienten von Modell 1 und Modell 2 und können wie standardisierte Koeffizienten interpretiert werden. Es wird deutlich, dass der Bildungsabschluss einen stärkeren Einfluss auf das Einkommen ausübt, als das Alter.

2.c) Wie erklären Sie die Werte b und b* in Modell 2? TIPP: Verwenden Sie bei Modell 2 das z-transformierte Einkommen als abhängige Variable.

Die in Modell 2 enthaltenen Variablen wurden alle vor der Durchführung der Regression z-standardisiert. Dadurch entsprechen die b-Koeffizienten den b*-Koeffizienten. Es ist egal, ob die Standardisierung der Variablen vor der Durchführung der Regression selbst vorgenommen wird oder von SPSS bei der Durchführung der Regression intern berechnet wird. Somit entspricht der b-Koeffizient von Modell 1 dem b*-Koeffizienten von Modell 1 und Modell 2.

3. Erstellen Sie ein multivariates Regressionsmodell mit Y=Einkommen. Versuchen Sie dabei den R² – Wert so groß wie nur irgendwie möglich zu bekommen. Jeder

schmutzige Trick der Sozialforschung ist erlaubt (und in diesem Fall erwünscht).
Fügen Sie die entsprechenden Teile des SPSS-Outputs in Ihre Abgabe ein.

Modellzusammenfassung				
Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,966 ^a	,934	,919	1,179

a. Einflußvariablen : (Konstante), WIRTSCHAFTSLAGE, BEFR. HEUTE , ZWEITSTIMME, LETZTE BUNDESTAGSWAHL , PRIVATER MUSIKUNTERRICHT (AUCH GESANG) , Alter mit 18 als Nullpunkt, DAUER DER ARBEITSLOSIGKEIT IN WOCHEN , BEFR.: FACHHOCHSCHULABSCHLUSS , ATTRAKTIVITAET: SELBSTEINSCHAETZUNG , WIRTSCHAFTSLAGE IN DER BRD HEUTE , FERNSEHGESAMTDAUER PRO TAG IN MINUTEN , REDUZIERTE HAUSHALTSGROESSE , HAEUFIGKEIT VON FERNSEHEN PRO WOCHE , Geschlecht mit w=0, OBEN-UNTEN- SKALA: SELBSTEINSTUF., BEFR. , SUBJEKTIVE SCHICHTEINSTUFUNG, BEFR. , FAMILIENBESITZ: IMMOBILIEN? , BEFR.: NETTOEINKOMMEN, OFFENE ABFRAGE , BEFR.: HOCHSCHULABSCHLUSS , Abschluss umkodiert, HAUSHALTSNETTOEINKOMMEN: OFFENE ABFRAGE

Durch die Hinzunahme von Variablen kann R^2 nur anwachsen (nicht kleiner werden). Zudem kann über eine Korrelationsmatrix geprüft werden, welche Variablen stark miteinander korrelieren. Durch Hinzunahme dieser Variablen kann das R^2 effektiv erhöht werden. Durch die Hinzunahme von X-Variablen, die fast das Gleiche messen wie die Y-Variable (z.B. Haushaltseinkommen) erhöht sich das R^2 ebenfalls deutlich.