

Universität Ulm,
Institut für Epidemiologie und Medizinische Biometrie,
D-89070 Ulm

**Institut für Epidemiologie und
Medizinische Biometrie**

Leiter: Prof. Dr. D. Rothenbacher
Schwabstr. 13, 89075 Ulm
Tel. +49 731 / 5026901

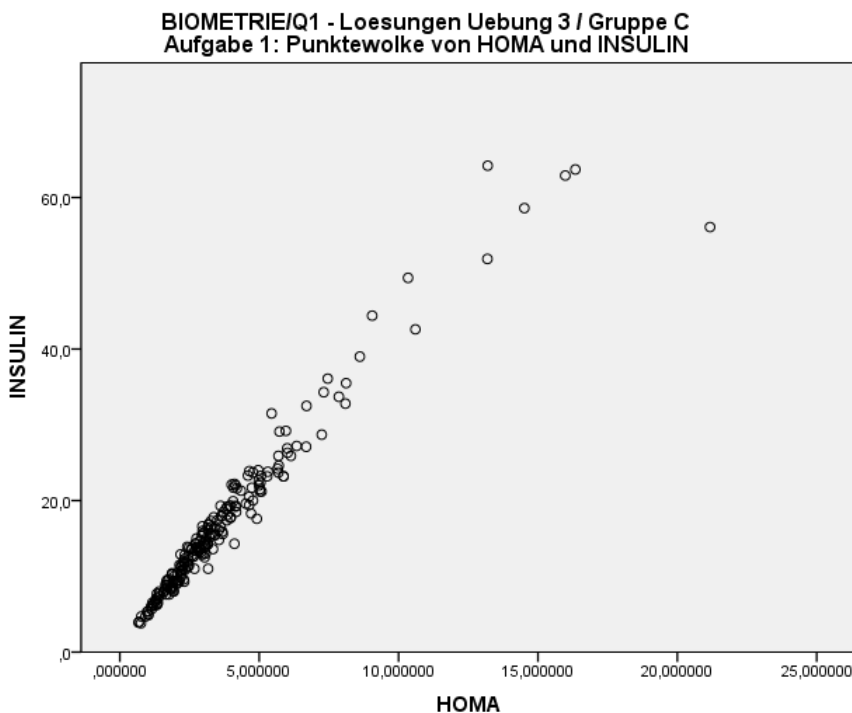
Übung 3 im Fach "Biometrie / Q1"

Aufgabe 1:

Sie möchten untersuchen, ob es einen linearen Zusammenhang zwischen der Insulinresistenz der Probanden (HOMA) und deren Insulinspiegel gibt.

Erstellen Sie eine bivariate Punktwolke (Scatterplot) mit SPSS so, dass auf der x-Achse die Werte von "HOMA" und auf der y-Achse die Werte von "INSULIN" abgetragen werden.

LÖSUNG: Diagramme → Diagrammerstellung



Interpretation:

Nach der Verteilung der Punktwolke lässt sich ein starker positiver Zusammenhang der beiden Merkmale vermuten.

Aufgabe 2:

In der Praxis wird eine Korrelationsberechnung erst angestrebt, wenn aus der Punktwolke hervorgeht, dass ein linearer Zusammenhang der beiden Merkmale bestehen könnte. Berechnen Sie den Korrelationskoeffizienten von "HOMA" und "INSULIN", unabhängig davon, was die Punktwolke zeigt. Wenden Sie die Methode nach Spearman an, weil Sie nicht wissen, ob HOMA und Insulin in unserer Stichprobe normalverteilt sind.

LÖSUNG: **Analysieren → Korrelation → Bivariat**

Korrelationen			INSULIN	HOMA
Spearman-Rho	INSULIN	Korrelationskoeffizient	1,000	,981**
		Sig. (2-seitig)	.	,000
		N	214	213
	HOMA	Korrelationskoeffizient	,981**	1,000
		Sig. (2-seitig)	,000	.
		N	213	213

** . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Korrelationskoeffizienten: **0.981**

Interpretation:

Der Korrelationskoeffizient ist sehr hoch. Es besteht also ein sehr großer Zusammenhang zwischen der Insulinresistenz und dem Insulinspiegel der Probanden.

Aufgabe 3:

Wenden Sie nun auf die Variablen "HOMA" und "INSULIN" das lineare Regressionsmodell an. Wir legen fest, dass "HOMA" das freie und "INSULIN" das abhängige Merkmal ist. Wir untersuchen also, wie gut sich der Insulinspiegel eines Probanden von seiner Insulinresistenz ableiten lässt. Lassen Sie neben den Ergebnissen der Berechnung auch eine Punktwolke mit der Regressionsgerade ausgeben.

Lösung: **Analysieren → Regression → Linear**

Aufgenommene/Entfernte Variablen^b

Modell	Aufgenommene Variablen	Entfernte Variablen	Methode
1	HOMA ^a	.	Einschluß

a. Alle gewünschten Variablen wurden eingegeben.

b. Abhängige Variable: INSULIN

Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,968 ^a	,937	,937	2,6848

a. Einflußvariablen : (Konstante), HOMA

Bestimmtheitsmaß r^2

ANOVA^b

Modell	Quadratsumme	df	Mittel der Quadrate	F	Sig.
1 Regression	22648,045	1	22648,045	3141,931	,000 ^a
Nicht standardisierte Residuen	1520,956	211	7,208		
Gesamt	24169,001	212			

a. Einflußvariablen : (Konstante), HOMA

b. Abhängige Variable: INSULIN

Koeffizienten^a

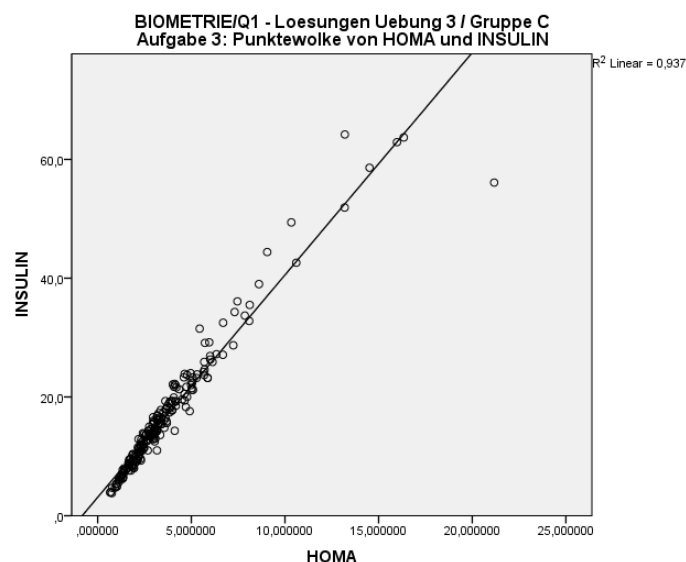
Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
		Regressionskoeffizient B	Standardfehler	Beta		
1	(Konstante)	3,113	,305		10,191	,000
	HOMA	3,744	,067	,968	56,053	,000

a. Abhängige Variable: INSULIN

Y-Achsenabschnitt

Steigung b

Notieren Sie die ermittelten Ergebnisse (3 Nachkommastellen):

a (Y-Achsenabschnitt) **3.113**b (Regressionskoeffizient) **3.744**r² (lineares Bestimmtheitsmaß) **0.937**Geradengleichung der Regressionsgeraden: $y = 3.113 + 3.744x$ **Lösung:**In **Diagramm-Editor** wechseln: **Elemente** → **Anpassungslinie bei Gesamtwert**

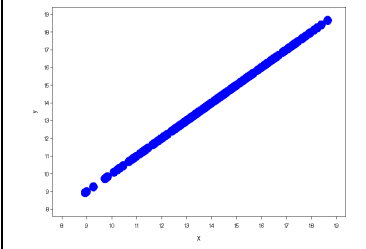
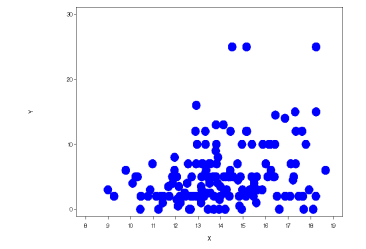
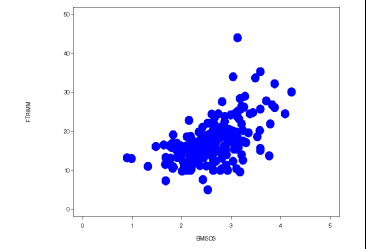
Interpretation:

Wie sich aus der Grafik ersehen lässt, liegen die Wertepaare sehr nahe zur Regressionsgerade. Das sehr hohe Bestimmtheitsmaß bestätigt, dass sich die Wertepaare sehr gut durch die Regressionsgerade beschreiben lassen.

Das lineare Regressionsmodell ist hier also gut geeignet.

Aufgabe 4:

Tippen Sie zu jeder Punktwolke den korrekten Korrelationskoeffizienten, indem Sie das entsprechende Feld markieren:

			
r = 0.52	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
r = 0.94	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
r = 1.00	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Aufgabe 5:

Sie möchten wissen, wie viele Hypertoniker sich unter den Probanden befinden. Hypertonie wird durch einen systolischen Blutdruck über 140 mmHg oder einen diastolischen Blutdruck über 90 mmHg definiert.

Um eine übersichtliche Tabelle erstellen zu können, müssen Sie zuerst die Variablen "RRDIAS" und "RRSYS" in folgende Klassen einteilen:

RRDIAS	$0 < x \leq 70$	RRSYS	$0 < x \leq 120$
	$70 < x \leq 90$		$120 < x \leq 140$
	$90 < x$		$140 < x$

LÖSUNG: Transformieren → Umkodieren in andere Variablen

Erstellen Sie nun eine zweidimensionale Häufigkeitstabelle mit absoluten Häufigkeiten. Verwenden Sie als Zeilenvariable "RRDIAS_recoded" und als Spaltenvariable "RRSYS_recoded".

LÖSUNG: Analysieren → Deskriptive Statistiken → Kreuztabellen**RRDIAS_recoded * RRSYS_recoded Kreuztabelle**

Anzahl

		RRSYS_recoded			Gesamt
		0 - 120	120 - 140	größer 140	
RRDIAS_recoded	0 - 70	49	16	2	67
	70 - 90	58	66	16	140
	größer 90	3	5	4	12
Gesamt		110	87	22	219

Wieviele Personen mit Hypertonie wurden beobachtet?

4 ☐30 ☒34 ☐112 ☐