

Beispiel zur multiplen Regression (Daten aus Wirtz & Nachtigall, 2002, S. 178)

Gegeben seien Körpergewicht (X_1), Lungenvolumen (X_2) und Fitness-Maße (Y) von 10 Personen (*Tabelle 1*). Die Korrelationen zwischen diesen drei Variablen sind in *Tabelle 2* dargestellt. Bei einer einfachen Regression von Y auf X_1 ist der standardisierte Regressionskoeffizient $\beta_1 = -.766$ und die Regressionsgleichung (mit unstandardisiertem Regressionskoeffizienten) lautet: $\hat{Y} = 33.781 - .342 * X_1$. Bei einer einfachen Regression von Y auf X_2 ist der standardisierte Regressionskoeffizient $\beta_2 = .775$ und die Regressionsgleichung (mit unstandardisiertem Regressionskoeffizienten) lautet: $\hat{Y} = -30.674 + .851 * X_2$. Die standardisierten Regressionskoeffizienten bei simultaner Berücksichtigung der Einflüsse von X_1 und X_2 sind $\beta_1 = -.521$ und $\beta_2 = .537$; die Regressionsgleichung (mit unstandardisierten Regressionskoeffizienten) lautet: $\hat{Y} = 1.907 - .232 * X_1 + .589 * X_2$.

Zur Berechnung im Einzelnen¹:

a) Einfache Regression von Y auf X_1 , d.h. $\hat{y} = a + b_1 x_1$:

$$b_1 = r_{yx1} \frac{s_y}{s_{x1}} = -.766 \frac{4.432}{9.930} = -.342$$

$$\beta_1 = b_1 \frac{s_{x1}}{s_y} = -.342 \frac{9.930}{4.432} = -.766$$

(in diesem Fall ist $\beta_1 = r_{yx1}$)

$$a = \bar{y} - b_1 \bar{x}_1 = 9.400 - (-.342) * 71.300 = 33.781$$

Die Regressionsgleichung lautet: $\hat{y} = 33.781 - .342x_1$

Varianzaufklärung: $R^2 = r_{yx1}^2 = -.766^2 = .587$ (d.h. 59 % gemeinsame Varianz)

b) Einfache Regression von Y auf X_2 , d.h. $\hat{y} = a + b_2 x_2$:

$$b_2 = r_{yx2} \frac{s_y}{s_{x2}} = .775 \frac{4.432}{4.036} = .851$$

$$\beta_2 = b_2 \frac{s_{x2}}{s_y} = .851 \frac{4.036}{4.432} = .775$$

(in diesem Fall ist $\beta_2 = r_{yx2}$)

$$a = \bar{y} - b_2 \bar{x}_2 = 9.400 - .851 * 47.100 = -30.674$$

Die Regressionsgleichung lautet: $\hat{y} = -30.674 + .851x_2$

Varianzaufklärung: $R^2 = r_{yx2}^2 = .775^2 = .600$ (d.h. 60 % gemeinsame Varianz)

¹ Die im Folgenden dargestellten Werte sind jeweils mit größerer Genauigkeit berechnet worden als dargestellt, weshalb sich beim Nachvollzug der Rechnungen anhand der dargestellten Werte Abweichungen ergeben können, die auf Rundungsungenauigkeiten zurückzuführen sind.

c) Einfache Regression von X_2 auf X_1 , d.h. $\hat{x}_2 = a + bx_1$:

$$b = r_{x_1x_2} \frac{s_{x_2}}{s_{x_1}} = -.457 \frac{4.036}{9.930} = -.186$$

$$\beta = b \frac{s_{x_1}}{s_{x_2}} = -.186 \frac{9.930}{4.036} = -.457$$

(in diesem Fall ist $\beta = r_{x_1x_2}$)

$$a = \bar{x}_2 - b\bar{x}_1 = 47.100 - (-.186) \cdot 71.300 = 60.354$$

Die Regressionsgleichung lautet: $\hat{x}_2 = 60.354 - .186x_1$

Varianzaufklärung: $R^2 = r_{x_1x_2}^2 = -.457^2 = .209$ (d.h. 21 % gemeinsame Varianz)

d) Multiple Regression von Y auf X_1 und X_2 , d.h. $\hat{y} = a + b_1x_1 + b_2x_2$:

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2} = \frac{-.766 - .775 \cdot (-.457)}{1 - (-.457)^2} = -.521$$

(hier ist $\beta_1 \neq r_{yx_1}$!)

$$\beta_2 = \frac{r_{yx_2} - r_{yx_1}r_{x_1x_2}}{1 - r_{x_1x_2}^2} = \frac{.775 - (-.766) \cdot (-.457)}{1 - (-.457)^2} = .537$$

(hier ist $\beta_2 \neq r_{yx_2}$!)

$$b_1 = \beta_1 \frac{s_y}{s_{x_1}} = -.521 \frac{4.432}{9.930} = -.232$$

$$b_2 = \beta_2 \frac{s_y}{s_{x_2}} = .537 \frac{4.432}{4.036} = .589$$

$$a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 = 9.400 - (-.232) \cdot 71.300 - .589 \cdot 47.100 = 1.786$$

Die Regressionsgleichung lautet: $\hat{y} = 1.786 - .232x_1 + .589x_2$

Varianzaufklärung: $R^2 = \beta_1 \cdot r_{yx_1} + \beta_2 \cdot r_{yx_2} = (-.521 \cdot -.766) + (.537 \cdot .775) = .815$ (d.h. durch die beiden unabhängigen Variablen werden 82 % der Varianz der abhängigen Variablen „aufgeklärt“)

Die konzeptuellen Modelle der vier Regressionsanalysen (mit standardisierten Regressionskoeffizienten) sehen wie folgt aus:

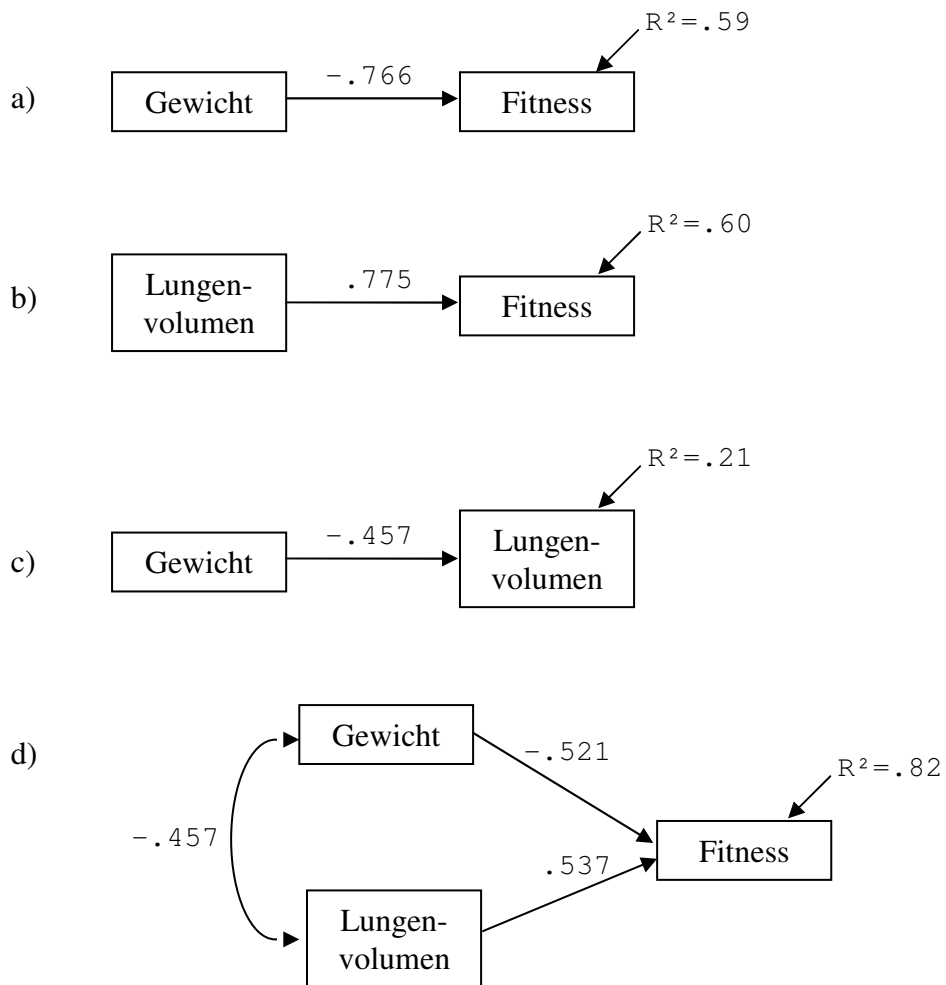


Tabelle 1: Gewicht, Lungenvolumen und Fitness von 10 Personen

	<i>Gewicht</i> (X_1)	<i>Lungenvolumen</i> (X_2)	<i>Fitness</i> (Y)
	87	42	1
	73	43	6
	66	44	7
	62	54	15
	68	45	12
	92	46	4
	60	50	12
	70	46	13
	71	54	14
	64	47	10
<i>aM</i>	71.30	47.10	9.40
<i>s</i>	9.93	4.04	4.43

Tabelle 2: Korrelationsmatrix für Gewicht, Lungenvolumen und Fitness

	X_1	X_2	Y
Gewicht (X_1)	1.000		
Lungenvolumen (X_2)	-.457	1.000	
Fitness (Y)	-.766	.775	1.000