

STATISTIK I – Lösung 08**Bravais-Pearson-Korrelationskoeffizient**

Eine Hochschule befragt zehn Absolventen/innen fünf Jahre nach ihrem Abschluss nach ihrem aktuellen Nettogehalt und ordnet die Summen den (metrisch skalierten!) Punkten in ihrer Abschlussklausur zu. Existiert eine lineare Korrelation zwischen dem Abschneiden in der Klausur und dem Nettogehalt nach fünf Jahren Berufstätigkeit?

Student Nr.	Klausurbewertung	Nettogehalt (in EUR)
1	98	2340
2	72	3750
3	65	1890
4	71	2420
5	56	2830
6	93	3110
7	66	1970
8	82	2480
9	89	2860
10	97	3333

Berechnen und interpretieren Sie den Bravais-Pearson-Korrelationskoeffizienten.

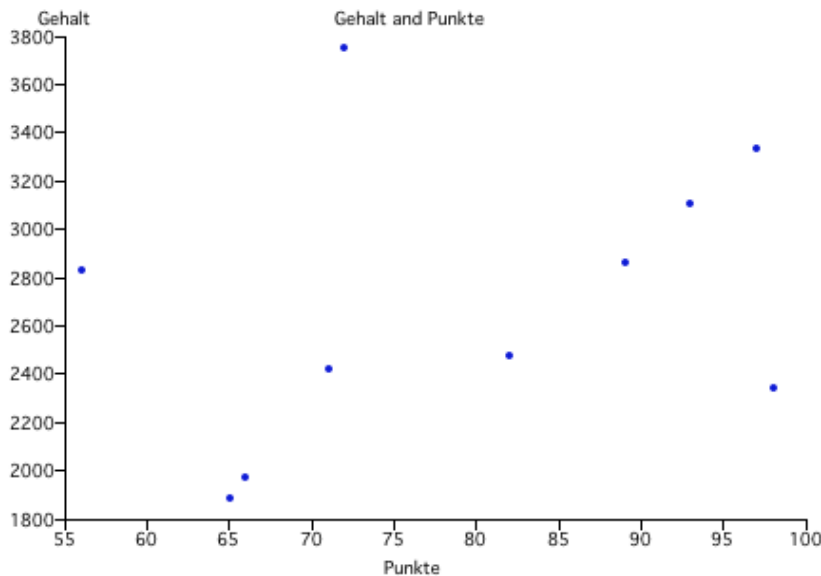
Der erste Schritt besteht auch hier wieder im Anlegen der benötigten Hilfstabelle.

Nr.	x	y	x ²	y ²	(x * y)
1	98	2340	9604	5475600	229320
2	72	3750	5184	14062500	270000
3	65	1890	4225	3572100	122850
4	71	2420	5041	5856400	171820
5	56	2830	3136	8008900	158480
6	93	3110	8649	9672100	289230
7	66	1970	4356	3880900	130020
8	82	2480	6724	6150400	203360
9	89	2860	7921	8179600	254540
10	97	3333	9409	11108889	323301
Σ	789	26983	64249	75967389	2152921
∅	78,90	2698,30	//	//	//

Im zweiten Schritt werden die Zahlen aus der Hilfstabelle in die entsprechende Formel eingesetzt:

$$r = \frac{\sum_{i=1}^n (x_i * y_i) - n * \bar{x} * \bar{y}}{\sqrt{\sum_{i=1}^n (x_i^2) - n * \bar{x}^2} * \sqrt{\sum_{i=1}^n (y_i^2) - n * \bar{y}^2}} = \frac{2152921 - 10 * 78,90 * 2698,30}{\sqrt{64249 - 10 * 78,90^2} * \sqrt{75967389 - 10 * 2698,30^2}} = 0,3017$$

Im dritten Schritt ist das Ergebnis von $r = 0,3$ nun zu interpretieren. Der Wert des Korrelationskoeffizienten deutet auf eine schwach-positive Korrelation hin – es existiert also offenkundig kein linearer Zusammenhang zwischen den Variablen. Dass dennoch ein Zusammenhang – beispielsweise ein monotoner – bestehen könnte, zeigt der Blick auf das Streudiagramm (in diesem Fall erstellt mit SSP).



Tendenziell scheinen hier – von drei aus der Reihe fallenden Werten abgesehen – große Werte von x nämlich durchaus mit großen Werten von y einherzugehen (und umgekehrt). Ein linearer Zusammenhang ist allerdings bereits rein visuell betrachtet maximal schwach vorhanden. Eventuell würde sich an dieser Stelle also die Berechnung des Rangkorrelationskoeffizienten nach Spearman oder des Konkordanzkoeffizienten nach Kendall lohnen, die beide monotone Zusammenhänge abbilden.

Zur inhaltlichen Interpretation sei an dieser Stelle noch einmal festgestellt, dass die Aussage „Das Nettogehalt wird durch die Studienleistung beeinflusst“ auch bei einer starken linearen Korrelation falsch gewesen wäre, da der Bravais-Pearson-Korrelationskoeffizient keinerlei Aufschluss über die Wirkungsrichtung eines möglichen inhaltlichen Zusammenhangs gibt. Auch wenn die alternative Wirkungsrichtung – das Nettogehalt fünf Jahre nach dem Studium wirkt sich rückwirkend auf die Punktzahl der Abschlussklausur aus – rein logisch nicht in Frage kommt, wäre jede Interpretation, die sich inhaltlich so weit vorwagt, nicht mehr allein mit dem Bravais-Pearson-Ergebnis begründbar.

Auch in diesem Beispielfall kann (und sollte) übrigens nicht ausgeschlossen werden, dass es sich auch um eine rein zufällige Scheinkorrelation oder um eine indirekte Korrelation über eine dritte Variable z handeln könnte. Letzteres wäre im vorliegenden Fall sogar durchaus plausibel, da das Nettogehalt sich nach fünf Jahren sicherlich nicht mehr direkt an den Punkten der Abschlussklausur orientieren wird, beide Variablen aber über eine dritte Variable z (beispielsweise die Motivation oder die persönliche Leistungsfähigkeit) indirekt miteinander verbunden sein könnten.