Dr. Quapp: Statistik für Mathematiker mit SPSS

Hinweise zur 4. Übung – Partielle Korrelation

3.] Es wird ein Zusammenhang vermutet zwischen der Zahl der Geburten Y eines Landes (je 100 Einwohner) und der Zahl der Störche X (pro 100 km^2)? Der Zusammenhang läßt sich aber stark "auspartialisieren", wenn man den Industrialisierungsgrad Z mit beachtet (Anteil der Industrieproduktion am Sozialprodukt) Datei stoerche.sav von D:

Land	Y	X	Z
Deutschl	0.60	0.81	55.5
Frankr	0.75	0.85	42.1
England	0.65	0.72	47.3
USA	0.91	1.12	49.2
Indien	1.72	1.88	22.8
Aegypten	1.93	2.21	21.2
China	1.66	2.13	23.4
Peru	1.35	1.21	35.3
Burma	1.57	1.88	27.3
Sudan	2.02	2.33	18.1

Bestimmen Sie die bivariaten und die partiellen Korrelationen und deuten Sie diese.

Neben direkten bivariaten Relationen zwischen Zufallsvariablen sind auch sogenannte versteckte Zusammenhänge von großem Interesse. Die (fiktiven) Daten der Aufgabe sollen dies an einem populären Beispiel verdeutlichen.
Der partielle Korrelationskoeffizient ist definiert durch:

$$\rho_{XY,Z} = \frac{\rho_{XY} - \rho_{YZ}\ \rho_{XZ}}{(1-\rho_{YZ}^2)\ (1-\rho_{XZ}^2)}$$

Im Fenster $--$ > Statistik $--$ > Korrelation $--$ > Partielle K. können im Zusatzfeld Optionen noch die bivariaten Korrelationskoeffizienten mitbestimmt werden. Ansonsten wird Z als Kontrollvariable verwendet, und ergibt eine spürbare Absenkung der Korrelation zwischen X und Y von 0.97 auf 0.71.

4.] Laden Sie die Datei explo.sav von D:

	Stoff	T	G	W
1	*Nitroglyzerin*	4250	7450	515
2	*Gurdynamit*	3420	6650	315
3	*Sprenggelatine*	4300	7800	520
4	*Tetranitropentaerythrit*	3700	6100	415
5	*Hexogen*	3800	8200	480
6	*Trinitrobenzol*	3540	7000	300
7	*Dinitrobenzol*	2500	6100	250
8	*Trinitrotoluol*	2820	6700	285
9	*Trinitroanisol*	2950	6600	290
10	*Pikrins"aure*	3230	7100	305
11	*Trinitrokresol*	2700	6850	275
12	*Hexanitrodiphenylamin*	3450	7100	320
13	*Tetranitromethylanilin*	3370	7200	340

Die Werte geben die Wirkung W (in cm^3) von verschiedenen Sprengstoffen an, wobei vermutet wird, daß diese abhängt von der Explosionstemperatur T (in ^{o}C), und von der Detonationsgeschwindigkeit G (in m/sec).
a) Bestimmen Sie durch lineare Regression die beste Funktion der Art

$$W = a + b\;T + c\;G\;.$$

b) Bestimmen Sie die bivariaten und die partiellen Korrelationen und deuten Sie diese.

Dieses Beispiel ist etwas realistischer: Es zeigt die durchaus qualitativ verschiedene Abhängigkeit der Variable W von den beiden Einflußfaktoren.
Im Fenster $--\ >$ Statistik $--\ >$ Korrelation $--\ >$ Partielle K. können im Zusatzfeld Optionen noch die bivariaten Korrelationskoeffizienten mitbestimmt werden. Ansonsten wird T oder G als Kontrollvariable verwendet; ersteres ergibt eine spürbare Absenkung der Korrelation zwischen G und W von 0.64 auf 0.27, während der Einfluß von G noch messbar ist, aber nicht den wirklichen Zusammenhang ausmacht: Es geht nur die Korrelation zwischen T und W von 0.91 auf 0.85 runter.
Die lineare Regression für W erfolgt 2-dimensional im Fenster $--\ >$ Lineare Regression. Es sollten noch die vorhergesagten Werte gespeichert werden. Als abhängige Variable dient W. Die gesuchte Funktion ist

$$W = -277,337 + 0,14\;T + 0,023\;G\;,$$

und das normierte β gibt die Relation zwische den beiden Einflußfaktoren auch wieder mit 0.819 zu 0.144 an, zugunsten von T, gegen G.
Im Streudiagramm kann unter 3D noch die Punktwolke (T, G, W) betrachtet werden, und eine 3D-Rotation (nach Doppelklick) ausgeführt, kann auch einen Eindruck der Daten vermitteln: Aber aus dem Bild wird überhaupt nicht der verschiedene Einfluß von T oder G ersichtlich!