Die Vierfeldertafel und bedingte Wahrscheinlichkeiten

In der Regel trifft man auf Vierfeldertafeln mit absoluten Häufigkeiten:

	Raucher (R)	Nichtraucher ($\bar{R}$)	Summe
treibt Sport (S)	12	27	39
treibt keinen Sport ($\bar{S}$)	28	33	61
Summe	40	60	100

Teilt man alle Einträge durch die Gesamtsumme unten rechts, dann erhält man die entsprechende Vierfeldertafel mit den relativen Häufigkeiten:

	Raucher (R)	Nichtraucher ($\bar{R}$)	Summe
treibt Sport (S)	0,12	0,27	0,39
treibt keinen Sport ($\bar{S}$)	0,28	0,33	0,61
Summe	0,40	0,60	1

Möchte man nun geeignete Wahrscheinlichkeiten ablesen, dann ergibt sich folgende Tabelle:

	Raucher (R)	Nichtraucher ($\bar{R}$)	Summe
treibt Sport (S)	$P(R \cap S) = 0{,}12$	$P(\bar{R} \cap S) = 0{,}27$	$P(S) = 0{,}39$
treibt keinen Sport ($\bar{S}$)	$P(R \cap \bar{S}) = 0{,}28$	$P(\bar{R} \cap \bar{S}) = 0{,}33$	$P(\bar{S}) = 0{,}61$
Summe	$P(R) = 0{,}4$	$P(\bar{R}) = 0{,}6$	$P(\Omega) = 1$

Wie man sofort sieht, tauchen **bedingte Wahrscheinlichkeiten** nicht auf! Allerdings lassen sich diese, nachdem die erste Pfadregel: $P(B) \cdot P_B(A) = P(A \cap B)$ mit $P(B) > 0$ umgestellt wurde, leicht berechnen:

$P_S(R) = \frac{P(R \cap S)}{P(S)} = \frac{0{,}12}{0{,}39} = \frac{4}{13} \approx 0{,}31$

$P_S(\bar{R}) = \frac{P(\bar{R} \cap S)}{P(S)} = \frac{0{,}27}{0{,}39} = \frac{9}{13} \approx 0{,}69$

$P_{\bar{S}}(R) = \frac{P(R \cap \bar{S})}{P(\bar{S})} = \frac{0{,}28}{0{,}61} = \frac{28}{61} \approx 0{,}46$

$P_{\bar{S}}(\bar{R}) = \frac{P(\bar{R} \cap \bar{S})}{P(\bar{S})} = \frac{0{,}33}{0{,}61} = \frac{33}{61} \approx 0{,}54$

$P_R(S) = \frac{P(R \cap S)}{P(R)} = \frac{0{,}12}{0{,}4} = \frac{3}{10} = 0{,}3$

$P_R(\bar{S}) = \frac{P(R \cap \bar{S})}{P(R)} = \frac{0{,}28}{0{,}4} = \frac{7}{10} = 0{,}7$

$P_{\bar{R}}(S) = \frac{P(\bar{R} \cap S)}{P(\bar{R})} = \frac{0{,}27}{0{,}6} = \frac{9}{20} = 0{,}45$

$P_{\bar{R}}(\bar{S}) = \frac{P(\bar{R} \cap \bar{S})}{P(\bar{R})} = \frac{0{,}33}{0{,}6} = \frac{11}{20} = 0{,}55$

Zwei Ereignisse mit positiven Wahrscheinlichkeiten werden als **stochastisch unabhängig** bezeichnet, genau dann wenn $P_B(A) = P(A)$ bzw. $P_A(B) = P(B)$. Das vorangegangene Ereignis darf also keinen Einfluss auf das nachfolgende Ereignis haben. Wie sieht es bei unserem Beispiel aus? $P_S(R) = \frac{4}{13} \neq \frac{4}{10} = P(R)$ Das bedeutet die beiden Ereignisse R und S sind **stochastisch abhängig**. Es gibt also einen Zusammenhang zwischen dem Rauchen und dem Sport. Man würde hier mehr Sportler unter den Nichtrauchern vermuten als unter den Rauchern. Betrachtet man dazu die bedingten Wahrscheinlichkeiten $P_{\bar{R}}(S) = 0{,}45$ und $P_R(S) = 0{,}3$ dann sieht man das auch.

Alternativ kann man natürlich auch andere Kombinationen untersuchen und erhält das gleiche Ergebnis:
$P_R(S) = 0{,}3 \neq 0{,}39 = P(S)$ oder
$P_{\bar{S}}(R) = 0{,}46 \neq 0{,}4 = P(R)$ oder
$P_{\bar{R}}(S) = 0{,}45 \neq 0{,}39 = P(S) \Longrightarrow R$ und S sind stochastisch abhängig.