Zu 29.2: Umformung der Formel für die Testfunktion des Unabhängigkeits- und Homogenitätstests

$$\begin{aligned}
Q &= \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{(h_{ij}-h_{ij}^{\mathrm{e}})^2}{h_{ij}^{\mathrm{e}}} \\
&= \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{h_{ij}^2}{h_{ij}^{\mathrm{e}}} - 2\sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{h_{ij}h_{ij}^{\mathrm{e}}}{h_{ij}^{\mathrm{e}}} + \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{(h_{ij}^{\mathrm{e}})^2}{h_{ij}^{\mathrm{e}}} \\
&= \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{h_{ij}^2}{h_{ij}^{\mathrm{e}}} - 2\sum_{i=1}^{K_x}\sum_{j=1}^{K_y}h_{ij} + \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}h_{ij}^{\mathrm{e}}
\end{aligned}$$

Nun gilt natürlich, dass der Stichprobenumfang gleich der Summe der absoluten Häufigkeiten ist:

$$\sum_{i=1}^{K_x}\sum_{j=1}^{K_y}h_{ij} = n$$

Ebenfalls gilt selbiges natürlich auch für die Randhäufigkeiten:

$$\begin{aligned}
\sum_{i=1}^{K_x}h(x_i) &= n, \\
\sum_{j=1}^{K_y}h(y_j) &= n
\end{aligned}$$

Und damit, nach Einsetzen der Nullhypothese, für die Summe der erwarteten Häufigkeiten:

$$\begin{aligned}
\sum_{i=1}^{K_x}\sum_{j=1}^{K_y}h_{ij}^{\mathrm{e}} &= \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{h(x_i)h(y_j)}{n} \\
&= \frac{1}{n}\left(\sum_{i=1}^{K_x}h(x_i)\right)\left(\sum_{j=1}^{K_y}h(y_j)\right) \\
&= \frac{n^2}{n} = n.
\end{aligned}$$

Damit

$$\sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{(h_{ij}-h_{ij}^{\mathrm{e}})^2}{h_{ij}^{\mathrm{e}}} = \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{h_{ij}^2}{h_{ij}^{\mathrm{e}}} - 2n + n = \sum_{i=1}^{K_x}\sum_{j=1}^{K_y}\frac{h_{ij}^2}{h_{ij}^{\mathrm{e}}} - n, \quad \text{q.e.d.}$$

(q.e.d. = *quod erat demonstrandum*, was zu beweisen war)