Faria, abaixo estão as citações do livro:

AGRESTI, ALAN. *Categorical Data Analysis* 2.ed. Hoboken, New Jersey
Wiley Series in Probability and Statistics, Wiley Interscience, 2002
710 páginas.
ISBN 0-471-36093-7

Estou traduzindo para o português.

página 37:

Na maioria das tabelas de contingência, uma das variáveis, Y, é uma variável resposta e a outra, X, é a variável explanatória.Quando X é prefixada e não aleatória, a noção de distribuição conjunta para X e Y não faz sentido. Entretanto, para cada categoria fixada de de X, Y tem uma distribuição de probabilidade. Desta forma, é apropriado estudar como esta distribuição muda a medida que as categorias de X variam. Dado que o sujeito é classificado na linha i de X , $\pi_{j|i}$ representa a probabilidade de classificação na coluna j de Y, $j = 1, \ldots, J$. Note que $\sum_j \pi_{j|i} = 1$. As probabilidades $\pi_{1|i}, \ldots, \pi_{J|i}$ formam a *probabilidade condicional* de Y para cada categoria i de X. **O principal objetivo de muitos estudos é comparar distribuições condicionais de Y nos vários níveis das variáveis explanatórias.** (o negrito é meu)

Página 39:

Quando X e Y são independentes,

$\pi_{j|i} = \frac{\pi_{ij}}{\pi_{i+}} = \frac{\pi_{i+}\pi_{+j}}{\pi_{i+}} = \pi_{+j}$ para $i = 1, \ldots, I$.

Cada distribuição condicional de Y é idêntica à distribuição marginal de Y. Então, duas variáveis são independentes quando

$\pi_{j|1} = \ldots = \pi_{j|I}$ para $j = 1, \ldots, J$,

quer dizer, a probabilidade de qualquer coluna-resposta é a mesma em cada linha.

Quando Y é uma variável resposta e X uma variável explanatória é mais natural definir independência desta forma do que com a primeira definição. Independência é então referida como *homogeneidade* das distribuições condicionais.