Varianz einer Summe diskreter Zufallsvariabler

Sei X eine diskrete Zufallsvariable, welche die Werte $x_i = x_1 \ldots x_n$ mit der Wahrscheinlichkeit $P(X = x_i)$ annehmen kann. Der Erwartungswert einer solchen Zufallsvariablen ist definiert als

(1) $$E(X) = \mu_X = \sum_{i=1}^{n} x_i \cdot P(X = x_i)$$

Die Varianz ist der Erwartungswert der quadrierten Abweichungen vom Erwartungswert der Zufallsvariablen:

(2) $$\mathrm{Var}(X) = E\left[(X - \mu_X)^2\right] = \sum_{i=1}^{n} (x_i - \mu_X)^2 \cdot P(X = x_i)$$

Für eine andere diskrete Zufallsvariable Y mit den möglichen Werten $y_j = y_1 \ldots y_m$, die mit der Wahrscheinlichkeit $P(Y = y_j)$ eintreten, ist der Erwartungswert entsprechend

(3) $$E(Y) = \mu_Y = \sum_{j=1}^{m} y_j \cdot P(Y = y_j)$$

und die Varianz

(4) $$\mathrm{Var}(Y) = E\left[(Y - \mu_Y)^2\right] = \sum_{j=1}^{m} (y_j - \mu_Y)^2 \cdot P(Y = y_j)$$

Wenn die beiden Zufallsvariablen unabhängig voneinander sind, gilt für die Varianz der Summe $X + Y$:

(5) $$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Beweis:

Entsprechend der Definition der Varianz gilt

$$\mathrm{Var}(X + Y) = E\left[(X + Y - \mu_X - \mu_Y)^2\right] = E\left(\left[(X - \mu_X) + (Y - \mu_Y)\right]^2\right)$$

Die rechte Seite dieser Gleichung wird ausmultipliziert:

(6) $$\mathrm{Var}(X + Y) = E\left[(X - \mu_X)^2 + 2(X - \mu_X) \cdot (Y - \mu_Y) + (Y - \mu_Y)^2\right]$$

Für Erwartungswerte gilt, dass diese entsprechend $E(X + Y) = E(X) + E(Y)$ addiert werden können. Dies auf die rechte Seite von Gleichung (6) angewandt:

$$\mathrm{Var}(X + Y) = E(X - \mu_X)^2 + E(Y - \mu_Y)^2 + E\left[2 \cdot (X - \mu_X) \cdot (Y - \mu_Y)\right]$$

Hierin Gleichung (2) und Gleichung (4) eingesetzt:

(7) $$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + E\left[2 \cdot (X - \mu_X) \cdot (Y - \mu_Y)\right]$$

Der in Gleichung (7) enthaltene Ausdruck $E\left[2(X - \mu_X) \cdot (Y - \mu_Y)\right]$ wird gesondert betrachtet.

Zunächst wird untersucht, ob der konstante Faktor 2, mit dem das Produkt $(X - \mu_X) \cdot (Y - \mu_Y)$ zu multiplizieren ist, auch erst nach der Ermittlung des Erwartungswertes angewandt werden kann. Dies wäre dann der Fall, wenn

$$E\left[2(X - \mu_X) \cdot (Y - \mu_Y)\right] = 2 \cdot E\left[(X - \mu_X) \cdot (Y - \mu_Y)\right]$$

Varianz einer Summe diskreter Zufallsvariabler

Zur Vereinfachung und Verallgemeinerung wird gefragt, ob die folgende Gleichung gültig ist:

(8) $$E(a \cdot X) = a \cdot E(X)$$

Dabei stellt a einen konstanten Faktor dar. Die möglichen Ausprägungen der Zufallsvariablen sind damit $\{a \cdot x_1, a \cdot x_2 \ldots a \cdot x_n\}$. Für $E(a \cdot X)$ gilt, dass diese Ausprägungen mit ihrer Wahrscheinlichkeit multipliziert und addiert werden, sodass

$$\begin{aligned} E(a \cdot X) &= a \cdot x_1 \cdot P(X = x_1) + a \cdot x_2 \cdot P(X = x_2) \ldots + a \cdot x_n \cdot P(X = x_n) \\ &= a \cdot \left[x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) \ldots + x_n \cdot P(X = x_n) \right] \\ &= a \cdot \sum_{i=1}^{n} x_i \cdot P(X = x_i) \\ &= a \cdot E(X) \end{aligned}$$

Gleichung (8) ist also gültig. Damit kann Gleichung (7) folgendermaßen geschrieben werden:

(9) $$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2 \cdot E\left[(X - \mu_X) \cdot (Y - \mu_Y)\right]$$

Der in Gleichung (9) enthaltene Ausdruck $(X - \mu_X) \cdot (Y - \mu_Y)$ wird ausmultipliziert:

$$\begin{aligned} E\left[(X - \mu_X) \cdot (Y - \mu_Y)\right] &= E(X \cdot Y - X \cdot \mu_Y - \mu_X \cdot Y + \mu_X \cdot \mu_Y) \\ &= E(X \cdot Y) - E(\mu_Y \cdot X) - E(\mu_X \cdot Y) + E(\mu_X \cdot \mu_Y) \end{aligned}$$

Da μ_X und μ_Y hier konstante Faktoren sind, können diese Faktoren nach Gleichung (8) vor den Erwartungswert gesetzt werden, sodass

$$E\left[(X - \mu_X) \cdot (Y - \mu_Y)\right] = E(X \cdot Y) - \mu_Y \cdot E(X) - \mu_X \cdot E(Y) + \mu_X \cdot \mu_Y \cdot E(1)$$

Hierin sind μ_X und $E(X)$ sowie μ_Y und $E(Y)$ identisch, und der Erwartungswert von 1 ist 1:

$$E\left[(X - \mu_X) \cdot (Y - \mu_Y)\right] = E(X \cdot Y) - E(X) \cdot E(Y) - E(X) \cdot E(Y) + E(X) \cdot E(Y)$$

(10) $$E\left[(X - \mu_X) \cdot (Y - \mu_Y)\right] = E(X \cdot Y) - E(X) \cdot E(Y)$$

Gleichung (10) in Gleichung (9) eingesetzt:

(11) $$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2 \cdot \left[E(X \cdot Y) - E(X) \cdot E(Y)\right]$$

Gleichung (11) würde mit Gleichung (5) übereinstimmen und die aufgestellte Behauptung wäre bewiesen, wenn

(12) $$E(X \cdot Y) = E(X) \cdot E(Y)$$

Dass diese Gleichung für unabhängige Zufallsvariable gilt, sei im Folgenden bewiesen.

Die möglichen Werte der Zufallsvariablen $X \cdot Y$ bestehen aus allen möglichen Produkten $x_i \cdot y_j$. Jedes x_i muss mit jedem y_j kombiniert werden. So gibt es für x_1 folgende Kombinationen:

$$x_1 \cdot y_1,\ x_1 \cdot y_2,\ x_1 \cdot y_3,\ \ldots,\ x_1 \cdot y_m$$

Ebenso kann x_2 mit allen Ausprägungen von Y kombiniert werden:

$$x_2 \cdot y_1,\ x_2 \cdot y_2,\ x_2 \cdot y_3,\ \ldots,\ x_2 \cdot y_m$$

Entsprechend müssen die Produkte mit x_3 bis x_n gebildet werden, sodass sich als letzte Kombination ergibt:

$$x_n \cdot y_1,\ x_n \cdot y_2,\ x_n \cdot y_3,\ \ldots,\ x_n \cdot y_m$$

Alle diese Kombinationen von Ereignissen sind mit der Wahrscheinlichkeit ihres Eintritts zu multiplizieren und zum Erwartungswert zu addieren. Wenn die Ereignisse x_i und y_j unabhängig voneinander sind, ist die Wahrscheinlichkeit ihres gemeinsamen Eintretens das Produkt ihrer Wahrscheinlichkeiten $P(X = x_i) \cdot P(Y = y_i)$. Damit gilt

$$\begin{aligned}
E(X \cdot Y) = \quad & x_1 \cdot y_1 \cdot P(X = x_1) \cdot P(Y = y_1) \\
+ \quad & x_1 \cdot y_2 \cdot P(X = x_1) \cdot P(Y = y_2) \\
+ \quad & x_1 \cdot y_3 \cdot P(X = x_1) \cdot P(Y = y_3) \\
\vdots \quad & \\
+ \quad & x_1 \cdot y_m \cdot P(X = x_1) \cdot P(Y = y_m) \\
+ \quad & x_2 \cdot y_1 \cdot P(X = x_2) \cdot P(Y = y_1) \\
+ \quad & x_2 \cdot y_2 \cdot P(X = x_2) \cdot P(Y = y_2) \\
\vdots \quad & \\
+ \quad & x_2 \cdot y_m \cdot P(X = x_2) \cdot P(Y = y_m) \\
\vdots \quad & \\
+ \quad & x_n \cdot y_1 \cdot P(X = x_n) \cdot P(Y = y_1) \\
+ \quad & x_n \cdot y_2 \cdot P(X = x_n) \cdot P(Y = y_2) \\
\vdots \quad & \\
+ \quad & x_n \cdot y_m \cdot P(X = x_n) \cdot P(Y = y_m)
\end{aligned}$$

Die Faktoren $x_1 \cdot P(X = x_1), x_2 \cdot P(X = x_2) \ldots x_n \cdot P(X = x_n)$ können jeweils ausgeklammert werden:

$$\begin{aligned}
E(X \cdot Y) = \quad & x_1 \cdot P(X = x_1) \cdot \left[y_1 \cdot P(Y = y_1) + y_2 \cdot P(Y = y_2) \ldots + y_m \cdot P(Y = y_m) \right] \\
+ \quad & x_2 \cdot P(X = x_2) \cdot \left[y_1 \cdot P(Y = y_1) + y_2 \cdot P(Y = y_2) \ldots + y_m \cdot P(Y = y_m) \right] \\
+ \quad & x_3 \cdot P(X = x_3) \cdot \left[y_1 \cdot P(Y = y_1) + y_2 \cdot P(Y = y_2) \ldots + y_m \cdot P(Y = y_m) \right] \\
\vdots \quad & \\
+ \quad & x_n \cdot P(X = x_n) \cdot \left[y_1 \cdot P(Y = y_1) + y_2 \cdot P(Y = y_2) \ldots + y_m \cdot P(Y = y_m) \right]
\end{aligned}$$

Hieraus wiederum $\left[y_1 \cdot P(Y = y_1) + y_2 \cdot P(Y = y_2) \ldots + y_m \cdot P(Y = y_m) \right]$ ausgeklammert:

$$E(X \cdot Y) = \left[x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) \ldots x_n \cdot P(X = x_n) \right] \cdot \left[y_1 \cdot P(Y = y_1) + y_2 \cdot P(Y = y_2) \ldots + y_m \cdot P(Y = y_m) \right]$$

Die beiden Ausdrücke in den eckigen Klammern können mithilfe des Summationszeichens geschrieben werden:

$$E(X \cdot Y) = \sum_{i=1}^{n} x_i \cdot P(X = x_i) \cdot \sum_{j=1}^{m} y_j \cdot P(Y = y_j)$$

Hierin Gleichung (1) und (3) eingesetzt:

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

Das ist Gleichung (12), womit diese bewiesen ist. Aus (12) folgt ohne Weiteres

$$E(X \cdot Y) - E(X) \cdot E(Y) = 0$$

Dies in Gleichung (11) eingesetzt ergibt den Additionssatz für die Varianzen einer Summe unabhängiger Zufallsvariabler:

(13) $$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Die Varianzen der Zufallsvariablen X und Y können zur Varianz der gemeinsamen Zufallsvariablen ihrer Summe addiert werden, wenn die Zufallsvariablen unabhängig voneinander sind.

Eine weitere Bedingung für die Additivität der Varianzen lässt sich aus Gleichung (9) ableiten. Wenn in dieser Gleichung der Ausdruck $E\left[(X-\mu_X)\cdot(Y-\mu_Y)\right]$ den Wert null annimmt, ergibt sich ebenfalls $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Der Ausdruck $E\left[(X-\mu_X)\cdot(Y-\mu_Y)\right]$ nun ist nichts anderes als die Kovarianz von X und Y, denn hierfür gilt

(14) $$\mathrm{Cov}(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i-\mu_X)\cdot(y_j-\mu_Y)\cdot P(X=x_i, Y=y_j) = E\left[(X-\mu_X)\cdot(Y-\mu_Y)\right]$$

Der Zusammenhang von Gleichung (9) lässt sich also auch folgendermaßen formulieren:

(15) $$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y)$$

Gleichung (9) und damit auch Gleichung (15) sind allgemeingültig, d.h. sie gelten für unabhängige wie für abhängige Ereignisse. Die Voraussetzung unabhängiger Ereignisse war für die Ableitung dieser Gleichungen nicht notwendig.

Wenn nun die Kovarianz gleich null sein kann, obwohl die beiden Zufallsvariablen nicht unabhängig sind, ist auch in diesem Fall die Additivität der Varianzen gemäß Gleichung (13) gegeben. Tatsächlich lassen sich Beispiele dafür finden, dass Zufallsvariable zwar abhängig voneinander sind, aber trotzdem nicht korrelieren[1], was durch eine Kovarianz von null angezeigt wird. Auch für diesen speziellen Fall abhängiger Ereignisse gilt $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Subtrahiert man die Gleichungen (11) und (15) voneinander, erhält man eine weitere allgemeingültige Definition der Kovarianz, nämlich

(16) $$\mathrm{Cov}(X,Y) = E(X\cdot Y) - E(X)\cdot E(Y)$$

Dass die rechte Seite dieser Gleichung und damit die Kovarianz gleich null ist, wenn die Zufallsvariablen X und Y unabhängig voneinander sind, wurde bereits gezeigt. Die Kovarianz kann aber auch gleich null sein, wenn die Ereignisse zwar abhängig voneinander sind, aber nicht korrelieren. Anders ausgedrückt: die Bedingung für die Additivität der Varianzen

$$E(X\cdot Y) = E(X)\cdot E(Y)$$

ist stets erfüllt für unabhängige Ereignisse, sie kann aber auch erfüllt sein für abhängige Ereignisse, wenn diese nicht korrelieren.

[1] Vgl. K. Bosch, Großes Lehrbuch der Statistik, München / Wien 1996, S. 215