Diskreter Wahrscheinlichkeitsraum, Gleichverteilung

Ein *diskreter Wahrscheinlichkeitsraum* ist eine Menge Ω, (endlich oder abzählbar), zusammen mit einer Abbildung $Pr : \Omega \to [0,1]$ mit

$$\sum_{\omega \in \Omega} Pr[\omega] = 1$$

Für ein $\omega \in \Omega$ nennt man $Pr[\omega]$ die

Wahrscheinlichkeit von ω.

Wichtiger Sonderfall:

Wenn Ω endlich ist und für alle $\omega \in \Omega$ gilt: $Pr[\omega] = \frac{1}{|\Omega|}$, dann spricht man von einer *Gleichverteilung.*

Ereignisse, Zufallsvariablen

Als *Ereignis* bezeichnet man eine Teilmenge $A \subseteq \Omega$.

Die Wahrscheinlichkeit, dass ein Ereignis A *eintritt,* ist

$$Pr[A] = \sum_{\omega \in A} Pr[\omega]$$

Im Fall der Gleichverteilung ist $Pr[\omega] = \frac{1}{|\Omega|}$ für alle ω, also:

$$Pr[A] = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl „gute“ Fälle}}{\text{Anzahl aller Fälle}}$$

Eine *Zufallsvariable* ist eine Abbildung

$$X : \Omega \to \mathbb{R}$$

Erwartungswert

Für eine Zufallsvariable $X : \Omega \to \mathbb{R}$ definieren wir den *Erwartungswert* $E[X]$ wie folgt:

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot Pr[\omega]$$

Bei endlichem Ω ist diese Summe immer definiert. Dagegen muss man in unendlichen Wahrscheinlichkeitsräumen hierbei absolute Konvergenz verlangen!

Für Gleichverteilung ergibt sich: $E[X] = \frac{1}{|\Omega|} \cdot \sum_{\omega \in \Omega} X(\omega)$

Ein Ereignis A kann als Zufallsvariable interpretiert werden, indem wir es mit der durch die charakteristische Funktion χ_A definierten Zufallsvariable identifizieren. (Hier: $\chi_A(x) = 1$ für $x \in A$, $\chi_A(x) = 0$, sonst)

Dann gilt: $$Pr[A] = \sum_{\omega \in A} Pr[\omega] = \sum_{\omega \in \Omega} \chi_A(\omega) \cdot Pr[\omega] = E[\chi_A]$$

Wahrscheinlichkeit von $X = x$

Für eine Zahl $x \in \mathbb{R}$ und eine Zufallsvariable X bietet es sich an, das folgende Ereignis zu betrachten:

$$\{\omega \in \Omega \mid X(\omega) = x\} = X^{-1}(x)$$

Dieses Ereignis nennen wir „$X = x$“ und erhalten:

$$Pr[X = x] = Pr[X^{-1}(x)] = Pr[\{\omega \in \Omega \mid X(\omega) = x\}]$$

Damit kann man den Erwartungswert für X auch als die folgende Summe angeben:

$$E[X] = \sum_{x \in \mathbb{R}} x \cdot Pr[X = x]$$

Natürlich kommen in dieser Summe nur höchstens abzählbar viele von Null verschiedene Summanden vor.

Markov-Ungleichung

Ein wichtiges Hilfsmittel bei Aussagen über Wahrscheinlichkeiten ist die Markov-Ungleichung:

Satz: Sei X eine Zufallsvariable mit $X(\omega) \geq 0$ für alle $\omega \in \Omega$, und sei $E[X] > 0$. Dann gilt:

$$\forall \lambda > 0 : \; Pr[X \geq \lambda \cdot E[X]] \leq \frac{1}{\lambda}$$

Also völlig unabhängig davon, welches Zufallsexperiment wir gerade betrachten: Dass man zufällig über das Doppelte des Erwartungswertes hinauskommt, wird höchstens mit 50 Prozent Wahrscheinlichkeit passieren. Und den Erwartungswert verzehnfachen kann man höchstens mit einer Wahrscheinlichkeit von einem Zehntel!

Beweis der Markov-Ungleichung

Wir haben streng genommen noch nicht definiert, was $Pr[\,X \geq \lambda E[X]\,]$ überhaupt bedeutet.

Es ist $Pr[A]$ für die Menge $A = \{\omega \in \Omega \mid X(\omega) \geq \lambda E[X]\}$.

Damit können wir nun die Markov-Ungleichung beweisen:

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \cdot Pr[\omega] \geq \sum_{\substack{\omega \in \Omega \\ X(\omega) \geq \lambda E[X]}} X(\omega) \cdot Pr[\omega] \geq$$

$$\sum_{\substack{\omega \in \Omega \\ X(\omega) \geq \lambda E[X]}} \lambda E[X] \cdot Pr[\omega] = \lambda E[X] \cdot Pr[\,X \geq \lambda E[X]\,]$$

Division beider Seiten durch $\lambda E[X]$ gibt das gewünschte Ergebnis.

Linearität des Erwartungswertes

Für zwei Zufallsvariablen X und Y und zwei reelle Zahlen a, b kann man eine weitere Zufallsvariable $aX + bY$ definieren durch

$$(aX + bY)(\omega) = a \cdot X(\omega) + b \cdot Y(\omega)$$

Hierbei handelt es sich also um eine Linearkombination der beiden gegebenen Zufallsvariablen X und Y. Auch $aX + bY$ ist dann eine Funktion von Ω nach $\mathbb{R}$.

Der Satz sagt, dass der Erwartungswert einer Linearkombination als Linearkombination der Erwartungswerte zu berechnen ist:

Satz: $E[aX + bY] = a \cdot E[X] + b \cdot E[Y]$

Beweis des Satzes

Es gilt

$$
\begin{aligned}
E[aX + bY] &= \sum_{\omega\in\Omega} (aX + bY)(\omega) Pr[\omega] \\
&= \sum_{\omega\in\Omega} (aX(\omega) + bY(\omega)) Pr[\omega] \\
&= a\Big(\sum_{\omega\in\Omega} X(\omega)\cdot Pr[\omega]\Big) + b\Big(\sum_{\omega\in\Omega} Y(\omega)\cdot Pr[\omega]\Big) \\
&= a\cdot E[X] + b\cdot E[Y]
\end{aligned}
$$

Unabhängige Zufallsvariablen

Es seien $X, Y : \Omega \to \mathbb{R}$ zwei Zufallsvariablen. Wir sagen, dass X und Y *unabhängige* Zufallsvariablen sind, falls

$$\forall x, y \in \mathbb{R} : \quad Pr[X = x \wedge Y = y] = Pr[X = x] \cdot Pr[Y = y]$$

Satz: Für unabhängige Zufallsvariablen X und Y gilt:

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

Der Beweis ergibt sich direkt aus der Definition der Unabhängigkeit von X und Y; die Details bleiben den interessierten Teilnehmern überlassen.

Varianz

Als Varianz der Zufallsvariablen X bezeichnet man den Erwartungswert des Quadrats der Differenz von X und $E[X]$:

$$
\begin{aligned}
Var[X] &= E[(X - E[X])^2] && \text{Quadrat ausrechnen} \\
&= E[X^2 - 2 \cdot E[X]\,X + E[X]^2] && \text{Erwartungswert linear} \\
&= E[X^2] - 2 \cdot E[X] \cdot E[X] + E[X]^2 \\
&= E[X^2] - E[X]^2
\end{aligned}
$$

Insbesondere kann man daraus schließen, dass der Erwartungswert des Quadrats mindestens so groß ist wie das Quadrat des Erwartungswerts.

Additivität der Varianz

Satz: Wenn X und Y unabhängige Zufallsvariablen sind, dann gilt:

$$Var[X+Y] = Var[X] + Var[Y]$$

Beweis: Wir müssen benutzen, dass $E[XY] = E[X]\,E[Y]$ gilt:

$$\begin{aligned}
Var[X+Y] &= E[(X+Y)^2] - E[X+Y]^2 \\
&= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\
&= E[X^2] + 2E[X]E[Y] + E[Y^2] \\
&\qquad - E[X]^2 - 2E[X]E[Y] - E[Y]^2 \\
&= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 \\
&= Var[X] + Var[Y]
\end{aligned}$$

Aufgabe 3.1

Ein Jäger hat Treffsicherheit 50%. Wie groß ist die Wahrscheinlichkeit, dass er bei 10 Schüssen mindestens dreimal trifft?

Lösung:

W'keit, dass er 10 mal daneben schießt:	$\frac{1}{2^{10}}$
W'keit für 9 mal daneben und einmal treffen:	$\frac{10}{2^{10}}$
W'keit für 8 mal daneben und zweimal treffen:	$\frac{45}{2^{10}}$

In allen übrigen Fällen trifft der Jäger mindestens dreimal.

Also ergibt sich: $\frac{2^{10}-56}{2^{10}} = \frac{968}{1024}$, das heißt fast 95%.

Kombinatorik

SPIELEN mit Zahlen...

In dieser Einheit spielen wir mit:

Binomialkoeffizienten

$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$

Anzahl der k-elementigen Teilmengen einer n-elementigen Menge

und

Partitionszahlen.

Auf wieviele Arten kann man n als Summe von k positiven ganzen Zahlen schreiben? Diese Anzahl nennen wir $P(n, k)$.

Die nächste Einheit wird sich dann mit den *Catalan-Zahlen* und deren Bezug zu *Dyck-Wörtern* beschäftigen.

Binomialkoeffizienten

Binomialkoeffizienten tauchen im Besonderen immer dann auf, wenn als *Zufallsexperiment* von n Kugeln zufällig k Kugeln gezogen werden, z.B. im Sport oder bei Lotterien.

1. Variationsmöglichkeit: Werden gezogene Kugeln zurückgelegt?
2. Variationsmöglichkeit: Kommt es auf die Reihenfolge an, in der die Kugeln gezogen werden?

Satz: Werden k aus n Kugeln mit Zurücklegen gezogen, so gibt es n^k mögliche Ergebnisse mit Reihenfolge, bzw. $\binom{n+k-1}{k}$ mögliche Ergebnisse ohne Reihenfolge.
Ohne Zurücklegen sind es $k! \cdot \binom{n}{k}$ mögliche Ergebnisse mit Reihenfolge bzw. $\binom{n}{k}$ ohne Reihenfolge.

Beweis des Satzes

Die Fälle „mit/mit" und „ohne/ohne" sind klar.

Mit Reihenfolge, aber ohne Zurücklegen:

Zunächst wie im Fall „ohne/ohne". Aber dann gibt es $k!$ Möglichkeiten, die k gezogenen Kugeln anzuordnen.

Ohne Reihenfolge, aber mit Zurücklegen:

Wichtig ist nur, wie oft jede Kugel gezogen wird.
Das Ergebnis ist also eine Folge $a_1, \ldots, a_n$ mit $\sum_{i=1}^{n} a_i = k$.

Ein solches Ergebnis stellen wir in der folgenden Form dar:

$$1^{a_1}01^{a_2}01^{a_3} \ldots 1^{a_{n-1}}01^{a_n}$$

Also durch einen 0-1-String mit k Einsen und $n-1$ Nullen.
D.h., wir haben einen String der festen Länge $n+k-1$, in dem die $n-1$ Positionen der Nullen das Ergebnis festlegen.
Hierfür gibt es $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$ Möglichkeiten.

Partitionszahlen

Wird eine n-elementige Menge in k nichtleere Teilmengen zerlegt, erzeugt das immer eine Aufteilung von n als Summe

$$n = n_1 + n_2 + \cdots + n_k$$

wobei n_i die Größe des i-ten Teils ist. Wir ordnen die n_i so an, dass für alle i gilt: $n_i \geq n_{i+1} \geq 1$.

Die *Partitionszahl $P(n, k)$* definieren wir jetzt als die Anzahl verschiedener Möglichkeiten, n auf diese Weise in k Summanden zu zerlegen.

So gilt zum Beispiel $P(7, 3) = 4$, denn 7 lässt sich auf folgende vier Arten in drei Summanden zerlegen:

5+1+1 4+2+1 3+3+1 3+2+2

$P(n)$ sei die Summe $\sum_k P(n, k)$.

Was ist $P(7)$?

Rekursionsformel für Partitionszahlen

Es gilt:

Satz:
$$\begin{aligned} P(n,k) &= P(n-1,k-1) + P(n-k,k) \\ &= \sum_{j \leq k} P(n-k,j) \\ &= \sum_{j \geq 0} P(n-jk-1,k-1) \end{aligned}$$

Für $k \geq \frac{n}{2}$ gilt außerdem: $P(n,k) = P(n-k)$.

Beweis: Wir überlegen uns zuerst die letzte Behauptung. Alle k Summanden sind größer als 0, d.h. wenn wir zunächst k mal eine 1 legen, bleibt noch die Teilsumme $n-k$ zu verteilen. Dafür gibt es $P(n-k)$ Möglichkeiten.

Wieso muss man hier $k \geq \frac{n}{2}$ fordern?

Beweis

Nun zur ersten Aufteilung:

Die Summendarstellungen $n = n_1 + \cdots + n_k$ können wir in zwei Fälle unterteilen, nämlich für $n_k = 1$ und $n_k \geq 2$. Wenn $n_k = 1$ ist, ist noch $n - 1$ auf die ersten $k - 1$ Zahlen zu verteilen. Dafür gibt es $P(n-1, k-1)$ Möglichkeiten. Wenn aber $n_k > 1$ gilt, dann ist noch $n - k$ auf k Summanden zu verteilen. Daraus die Beh.

Die erste Summendarstellung folgt aus dieser Zerlegung iterativ:

$$P(n,k) = P(n-1,k-1) + P(n-k,k) = P(n-2,k-2) + P(n-k,k-1) + P(n-k,k) = \ldots$$

Hierbei wurde immer ein $P(n-j, k-j)$ weiter zerlegt. Wenn wir stattdessen das $P(n-k, k)$ weiter zerlegen, erhalten wir die zweite Summendarstellung:

$$P(n,k) = P(n-1,k-1) + P(n-k,k) = P(n-1,k-1) + P(n-k-1,k-1) + P(n-2k,k) = \ldots$$

Das komplettiert den Beweis.