unabhängig und identisch verteilt (u.i.v.)

Autor und Herausgeber

Alexander Jung*, Department of Computer Science, Aalto University, Espoo, Finland (ORCID 0000-0001-7538-0990)
Editors: Konstantina Olioumtsevits (ORCID 0000-0003-2697-8659), Ekkehard Schnoor (ORCID 0000-0001-8252-1499), Aalto University, Espoo, Finland
*Corresponding author: alex.jung@aalto.fi

Zusammenfassung

Eine Sammlung von Zufallsvariablen (ZV) ist unabhängig und identisch verteilt (i.i.d.), wenn alle die gleiche Wahrscheinlichkeitsverteilung befolgen und sie wechselseitig unabhängig sind. Zwei Würfe desselben Würfels sind i.i.d.: beide Werte haben die gleiche Wahrscheinlichkeitsverteilung, und der erste sagt nichts über den zweiten aus. Die beiden Anforderungen sind getrennt, und jede kann ohne die andere fehlschlagen. Wechselseitige Unabhängigkeit bedeutet, dass die Wahrscheinlichkeit jeder Wertekombination das Produkt einer Wahrscheinlichkeit pro ZV ist. Dies ist stärker als die Unabhängigkeit jedes Paares. Für gemeinsam Gaußsche Zufallsvariablen gelten die beiden Anforderungen genau dann, wenn alle Mittelwerte gleich sind und die Kovarianzmatrix eine skalierte EinheitsMatrix ist.

Definition

Zwei Würfe desselben Würfels sind unabhängig und identisch verteilt: beide Werte folgen der gleichen Wahrscheinlichkeitsverteilung, und der erste sagt nichts über den zweiten aus. Eine Sammlung von Zufallsvariablen (ZV) mit diesen beiden Eigenschaften wird i.i.d. genannt. Formal betrachte Zufallsvariablen $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(m)}$, die alle auf einem gemeinsamen probability space $(\Omega, \Sigma, \mathbb{P})$ definiert sind und Werte in einem measurable Raum $\mathcal{Z}$ annehmen. Sie sind i.i.d., wenn jedes $\mathbf{z}^{(r)}$ der gleichen Wahrscheinlichkeitsverteilung folgt und die Zufallsvariablen wechselseitig unabhängig sind. Der gemeinsame probability space ist es, der die wechselseitige Unabhängigkeit wohldefiniert macht: ein Ergebnis muss alle m Werte gleichzeitig festlegen (siehe ZV). Die beiden Anforderungen sind getrennt, und jede kann ohne die andere fehlschlagen. Zwei Sensoren unterschiedlicher Genauigkeit, jeweils einmal abgelesen, liefern unabhängige Werte, die nicht identisch verteilt sind. Die von einem Sensor zu zwei aufeinanderfolgenden Zeitpunkten gemessene Temperatur ist identisch verteilt und abhängig, da der zweite Wert den ersten fast wiederholt.

Wechselseitige Unabhängigkeit spiegelt sich in einer Faktorisierung der Wahrscheinlichkeitsverteilung wider: Die Wahrscheinlichkeit jeder Kombination ihrer Werte ist das Produkt einer Wahrscheinlichkeit pro ZV. Für beliebige

measurable Mengen $\mathcal{A}_1, \ldots, \mathcal{A}_m \subseteq \mathcal{Z}$,

$$\mathbb{P}\left(\mathbf{z}^{(1)} \in \mathcal{A}_1, \ldots, \mathbf{z}^{(m)} \in \mathcal{A}_m\right) = \prod_{r=1}^{m} \mathbb{P}\left(\mathbf{z}^{(r)} \in \mathcal{A}_r\right). \tag{1}$$

Die Produktregel (1) muss für die gesamte Sammlung gelten, nicht nur für Paare. Seien b_1 und b_2 unabhängige Zufallsvariablen, die jeweils die Werte 0 und 1 mit Wahrscheinlichkeit $1/2$ annehmen, und sei b_3 ihre Summe modulo zwei. Alle drei nehmen die Werte 0 und 1 mit Wahrscheinlichkeit $1/2$ an, sind also identisch verteilt, und je zwei von ihnen sind unabhängig. Die drei zusammen sind es nicht: $\mathbb{P}(b_1 = 1, b_2 = 1, b_3 = 1) = 0$, weil $1 + 1 = 0$ modulo zwei ist, während das Produkt der drei einzelnen Wahrscheinlichkeiten $1/8$ beträgt. Paarweise Unabhängigkeit ist daher schwächer als wechselseitige Unabhängigkeit, und diese drei Zufallsvariablen sind identisch verteilt und paarweise unabhängig, ohne i.i.d. zu sein (Gray and Davisson, 2004, Problem 4.31).

Für gemeinsam Gaußsche Zufallsvariablen nehmen die beiden Anforderungen eine kompakte Form an. Betrachte einen Zufallsvektor $\mathbf{x} = (x_1, \ldots, x_d)^\top$ von gemeinsam Gaußsche Zufallsvariablen mit Mittelwert Vektor $\boldsymbol{\mu}$ und Kovarianzmatrix $\mathbf{C}$. Seine Komponenten sind i.i.d. genau dann, wenn $\boldsymbol{\mu} = (\mu, \ldots, \mu)^\top$ und $\mathbf{C} = \sigma^2\mathbf{I}$ eine skalierte EinheitsMatrix ist. Tatsächlich ist für $\mathbf{C} = \sigma^2\mathbf{I}$ die gemeinsame Wahrscheinlichkeitsdichte (Murphy, 2012, Gl. (4.1)) ein Produkt univariater Gaußscher Dichten mit Mittelwerte μ_j und der gemeinsamen Varianz σ^2, sodass die Komponenten unabhängig sind, und sie sind genau dann identisch verteilt, wenn die Mittelwerte übereinstimmen. Umgekehrt haben i.i.d. Komponenten gleiche Mittelwerte, gleiche Varianzen und null Kovarianzen. Jede Komponente hat dann die gleiche Varianz σ^2, deren Quadratwurzel σ die Standardabweichung entlang jeder Achse ist. Die Nullen außerhalb der Diagonale machen die Komponenten unkorreliert. Für gemeinsam Gaußsche Zufallsvariablen impliziert dies bereits wechselseitige Unabhängigkeit (Anderson, 2003, Abschn. 2.4), sodass die drei Zufallsvariablen b_1, b_2, b_3 kein gemeinsam Gaußsches Gegenstück haben. Die gleichen Diagonaleinträge und die gleichen Mittelwerte machen die Komponenten identisch verteilt. Abb. 1 zeigt Realisierungen eines solchen Zufallsvektor für $d = 2$.

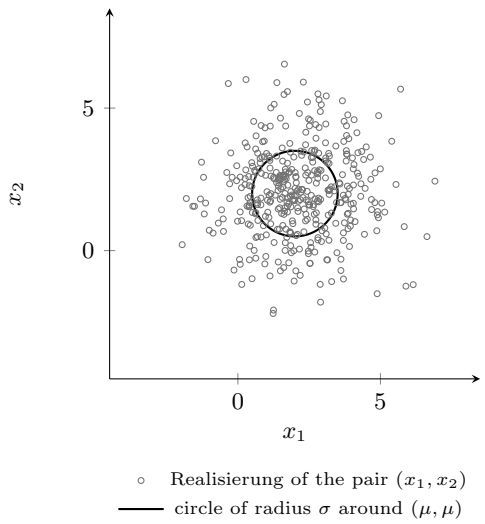

Abbildung 1: Realisierungen eines Paares x_1, x_2 von i.i.d. Gaußsche Zufallsvariablen mit Mittelwert μ und Varianz σ^2, jeweils als offener Kreis gezeichnet. Die durchgezogene Kurve hat den Radius σ und ist bei (μ, μ) zentriert. Identische Wahrscheinlichkeitsverteilungen bewirken, dass die Streuung entlang beider Achsen gleich ist, und wechselseitige Unabhängigkeit lässt die Wolke ungeneigt, sodass die Kovarianzmatrix das σ^2-fache der EinheitsMatrix ist. Daten generiert von `pythondemos/iid.py`

Siehe auch: ZV, Annahme unabhängig und identisch verteilter Daten (u.i.v.-Annahme), Wahrscheinlichkeitsverteilung, probability space, Ereignis, measurable, Realisierung, Gaußsche Zufallsvariable, Kovarianzmatrix, Zufallsvektor, Varianz.

Querverweise

Literatur

1. Anderson (2003). An Introduction to Multivariate Statistical Analysis. 3rd ed., Wiley, Hoboken, NJ.

2. Gray and Davisson (2004). An Introduction to Statistical Signal Processing. Cambridge Univ. Press, Cambridge, U.K.

3. Murphy (2012). Machine Learning: A Probabilistic Perspective. MIT Press, Cambridge, MA, USA.