Statistik

Vorlesung 5 - Erwartungswert und Varianz von diskreten Zufallsvariablen

Prof. Dr. Sandra Eisenreich

22. April 2024

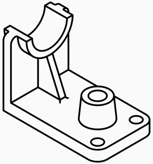
Hochschule Landshut

Agenda

1. Erwartungswert einer diskreten Zufallsvariable
2. Varianz einer diskreten Zufallsvariablen
3. Kovarianzmatrix für Zufallsvektoren

Erwartungswert einer diskreten Zufallsvariable

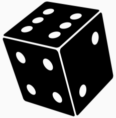
Motivation Erwartungswert



Wahrscheinlichkeit für niO Bauteil $= 0,2\%$. Wie viele defekte Bauteile würden wir bei $N = 1000$ produzierten Bauteilen dann erwarten?

$$0,2\% \cdot 1000 = 2$$

Welchen Wert erwartet man im Schnitt bei sehr vielen Würfen mit einem ungezinkten Würfel?



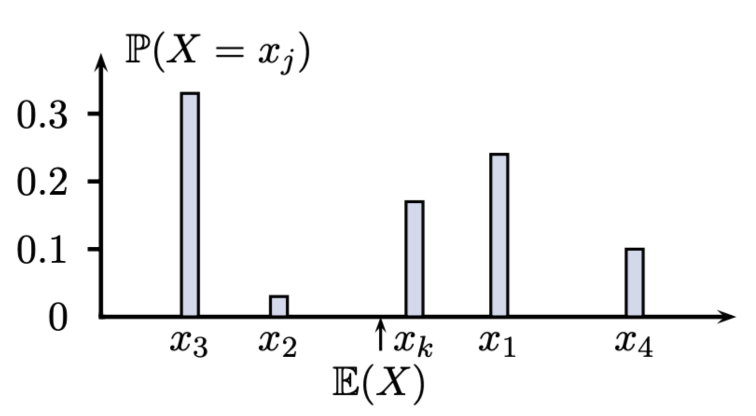
$$\frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \ldots \frac{1}{6} \cdot 6 = \frac{1 + 2 + \ldots + 6}{6} = \frac{21}{6} = 3,5$$

Erklärung. Bei je einem Sechstel aller Würfe erhalten wir 1,. . . , 6, das heißt im Schnitt bei einem Wurf 3,5.

$\rightarrow$ Erwartungswert = “Durchschnittliches Ergebnis”.

Achtung: dieser muss nicht im Ergebnisraum liegen! (z.B. beim Würfeln).

Erwartungswert als (physikalischer) Schwerpunkt



Erwartungswert einer diskreten Zufallsvariablen

Definition (Erwartungswert einer diskreten Zufallsvariablen)

Sei X eine diskrete Zufallsvariable. Falls die Summe

$$E(X) = \sum_{x \in W(X)} x \cdot P(X = x)$$

eindeutig existiert, so heißt sie Erwartungswert von X.

Beispiel: Roulette

Roulettespiel

- die Kugel kann mit gleicher Wahrscheinlichkeit auf eine der Zahlen von 0 bis 36 fallen
  - wir setzen 1€ auf die Zahl 7 (Gewinn: 36 €)
  - wir setzen gleichzeitig 1€ auf die ungeraden Zahlen (Gewinn: 2€)
  - wir setzen gleichzeitig 1€ auf das Tripel $(1, 2, 3)$ (Gewinn: 12€)

Frage: Was ist der Erwartungswert der Zufallsvariable $X =$ Gewinn?

Ergebnis: Roulette

Wir haben 3 Zufallsvariablen:

$$X_A = \begin{cases} 36 & \text{wenn } \omega = 7 \\ 0 & \text{sonst.} \end{cases}$$

$$X_B = \begin{cases} 2 & \text{wenn } \omega \text{ ungerade} \\ 0 & \text{sonst.} \end{cases}$$

$$X_C = \begin{cases} 12 & \text{wenn } \omega = 1, 2, 3 \\ 0 & \text{sonst.} \end{cases}$$

Alle 3 werden gleichzeitig gesetzt (Übung: Sind alle 3 unabhängig?) und wir erhalten:

$$X = X_A + X_B + X_C$$

Ergebnis: Roulette

Die Werte der Zufallsvariable X sind gegeben durch:

$$X(\omega) = \begin{cases} 38 & w = 7 \\ 14 & w \in \{1, 3\} \\ 12 & w = 2 \\ 2 & w \text{ist ungerade und } \notin \{1, 3, 7\} \\ 0 & \text{sonst} \end{cases}$$

und die Verteilung $P(X = k)$ durch:

$$P(X = k) = \begin{cases} \frac{18}{37} & k = 0 \\ \frac{15}{37} & k = 2 \\ \frac{1}{37} & k = 12 \\ \frac{2}{37} & k = 14 \\ \frac{1}{37} & k = 38 \end{cases}$$

Ergebnis: Roulette

Der Erwartungswert ist also

$$E(X) = 2 \cdot \frac{15}{37} + 14 \cdot \frac{2}{37} + 12 \cdot \frac{1}{37} + 38 \cdot \frac{1}{37} = \frac{108}{37}.$$

Abzüglich des Einsatzes von 3€ haben wir $\frac{108}{37} - 3 = -\frac{3}{37}$.

Eigenschaften eines Erwartungswertes

Wichtige Rechenformeln:

Sind X, Y Zufallsvariablen auf Ω, $a \in \mathbb{R}$ und $A \subset \Omega$ ein Ereignis, so gelten:

(a) $E(X + Y) = E(X) + E(Y)$,

(b) $E(aX) = aE(X)$,

(c) $E(1_A) = P(A)$. Hierbei ist 1_A die Zufallsvariable mit Ausgang 1, falls A eintritt, sonst 0.

Beispiele

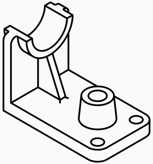
- Was ist der Erwartungswert von $X_A =$ Gewinn im Roulette, wenn wir 1€ auf die Zahl 7 setzen (Gewinn: 36 €)?
- Was ist der Erwartungswert von $X_B =$ Gewinn im Roulette, wenn wir 1€ auf die ungeraden Zahlen setzen (Gewinn: 2€)?
- Was ist der Erwartungswert von $X = X_A + X_B$ (also beiden Spielen)?
- Was ist der Erwartungswert von $X =$ Augensumme bei 10x Würfeln?

Ergebnisse

- $E(X_A) = \frac{1}{37} \cdot 36 = \frac{36}{37}$
- $E(X_B) = \frac{15}{37} \cdot 2 = \frac{30}{37}$
- $E(X) = E(X_A) + E(X_B) = \frac{66}{37}$
- Ist X_1=Augenzahl bei 1x Würfel, so ist

$$E(X) = E(10 \cdot X_1) = 10 \cdot E(X_1) = 35$$

Varianz einer diskreten Zufallsvariablen



Im Schnitt gibt es 2 defekte Bauteile pro 1000, aber das bedeutet nicht dass es immer 2 pro 1000 sind. Je breiter die "Streuung", umso unvorherbarer ist der Produktionsprozess. Diese Streuung nennt man Varianz und sie kann berechnet werden.

Wie berechnen wir die Streuung? Versuch 1

Idee: Berechne die Varianz als Abweichung der Zufallsvariable vom Mittelwert dar: $E(X) = \mu$ und $Y = X - \mu$ stellt die Abweichung des tatsächlichen "Gewinns" vom Erwartungswert

Problem: $E(Y) = E(X - \mu) = E(X) - \mu = 0$,
das heißt die mittlere Varianz wäre 0, weil es Abweichungen in beide richtungen gibt...? $\rightarrow$ kein gutes Maß.

Wie berechnen wir die Streuung? Versuch 2

Idee: Zähle alle Abweichungen positiv mit dem Betrag: $Y = |X - \mu|$ und $E(Y)$ ist mittlere Abweichung

Problem: ein Mathematiker rechnet mit Betragstrichen nur dann, wenn es sich nicht vermeiden lässt: Es gibt immer so hässliche Fallunterscheidungen.

Wie berechnen wir die Streuung? Versuch 3

Idee: Zähle die quadratischen Abweichungen $Y = (X - \mu)^2$ und $E(Y)$ ist mittlere quadratische Abweichung

Das ist das klassische Streuungsmaß und heißt Varianz.

Definition

Ist $E(X) = \mu$ der Erwartungswert der Zufallsvariablen X, so heißt der Wert

$$\sigma^2 := Var(X) := E((X - \mu)^2),$$

falls er existiert, die Varianz von X, und $\sigma = +\sqrt{Var(X)}$ heißt Standardabweichung oder Streuung von X.

Theorem (Rechenregel)

Für eine diskrete Zufallsvariable X mit Erwartungswert μ gilt

$$Var(X) = \sum_{x \in W(X)} (x-\mu)^2 \cdot P(X=x) = \left(\sum_{x \in W(X)} x^2 P(X=x) \right) - \mu^2 = E(X^2) - \mu^2$$

Das bedeutet: $Var(X) = E(X^2) - E(X)^2$

Varianz am Beispiel Roulette

$$
\begin{aligned}
Var(X_A) &= 36^2 P(X_A = 36) + 0^2 P(X_A = 0) - \left(\frac{36}{37}\right)^2 \\
&\approx 34.08 \\
Var(X_B) &= 2^2 \cdot \frac{18}{37} + 0^2 \cdot P(X_B = 0) - \left(\frac{36}{37}\right)^2 \\
&\approx 0.999 \\
Var(X_C) &= 12^2 \cdot \frac{3}{37} - \left(\frac{36}{37}\right)^2 \\
&\approx 10.73
\end{aligned}
$$

Eigenschaften der Varianz: Rechenregeln

Seien X, Y **unabhängige** Zufallsvariablen.

- Für $a, b \in \mathbb{R}$ gilt: $Var(aX + b) = a^2 Var(X)$
- $E(X \cdot Y) = E(X) \cdot E(Y)$
- $Var(X + Y) = Var(X) + Var(Y)$.
- Für **unabhängige** Zufallsvariablen X_i gilt:
  $Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$
- Tschebyscheff-Ungleichung (wie wahrscheinlich ist ein Ergebnis weiter als ϵ vom Erwartungswert weg?): Für jedes $\varepsilon > 0$ gilt:
  $P(|X - E(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} Var(X)$

Warum gilt $Var(X+Y) = Var(X) + Var(Y)$?

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

Standardisierung: Motivation

In Data Science und Machine Learning ist es **essentiell**, dass alle Daten im gleichen Wertebereich liegen und oft auch vergleichbar sind, das heißt dieselbe Varianz aufweisen. Wie?

- in den gleichen Wertebereich bringen: durch Abziehen des Erwartungswerts = Mittelwerts von jedem Merkmal.
- vergleichbar machen: die Varianz aller Merkmale auf 1 bringen.

Das nennt man Standardisierung der Daten.

Standardisierung eine Zufallsvariablen

Definition

Ist X eine Zufallsvariable mit dem Erwartungswert μ und Varianz σ^2, so heißt die Zufallsvariable

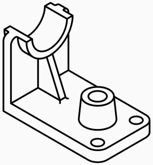
$$X^* := \frac{X - \mu}{\sigma}$$

die Standardisierte von X.

Es gilt für $X^* = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$

$$\begin{aligned}
E(X^*) &= E\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma}E(X) - \frac{\mu}{\sigma} = 0 \\
Var(X^*) &= Var\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}Var(X) = 1
\end{aligned}$$

Kovarianzmatrix für Zufallsvektoren



X= “Die Abweichung eines Bauteils in x-Richtung auf ein Zehntel mm gerundet”; Y, Z ebenso.

Wenn ein Bauteil zum Beispiel etwas verdreht gemessen wird, hat man überall Abweichungen, obwohl eventuell gar kein Fehler vorliegt $\Rightarrow X, Y$ und Z sind nicht unabhängig voneinander.

Die Wechselwirkungen stecken in den Kovarianzen $Cov(X, Y), Cov(Y, Z), Cov(X, Z)$.

Was passiert, wenn die Zufallsvariablen X und Y nicht unabhängig sind? Dann gilt:

$$\begin{aligned} Var(X+Y) &= Var(X) + Var(Y) + 2E(XY) - 2E(X)E(Y) \\ &= Var(X) + Var(Y) + 2 \cdot Cov(X,Y) \end{aligned}$$

Definition (Kovarianz)

$$Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{2}(Var(X+Y) - Var(X) - Var(Y))$$

heißt Kovarianz zwischen X und Y (ein Maß für ihre Abhängigkeit).

Eigenschaften der Kovarianz

Theorem

Sind X, Y Zufallsvariablen und a, b reelle Zahlen, so gelten:

(a) $Cov(X, Y) = Cov(Y, X)$ *und* $Cov(X, X) = Var(X)$

(b) $Cov(X + a, Y + b) = Cov(X, Y)$

(c) *Sind* X*,* Y *stochastisch unabhängig, so folgt* $Cov(X, Y) = 0$.

Kovarianzmatrix

Definition

Seien $X_1, \ldots, X_d$ Zufallsvariablen. Dann definiert man ihre Kovarianzmatrix als die $d \times d$-Matrix

$$Cov(X_1, \ldots, X_d) = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \ldots & Cov(X_1, X_d) \\ Cov(X_2, X_1) & Var(X_2) & \ldots & Cov(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_d, X_1) & Cov(X_d, X_2) & \ldots & Var(X_d) \end{pmatrix}$$

Es gilt:

- Die Kovarianzmatrix ist symmetrisch, da $Cov(X_i, X_j) = Cov(X_j, X_i)$.
- Die Kovarianzmatrix ist diagonal $\Leftrightarrow$ Die Zufallsvariablen sind unabhängig.

Literatur

- Hartmann, Peter; Mathematik für Informatiker, Springer-Vieweg; 7. Auflage; 2019
- Henze, Norbert; Stochastik für Einsteiger; Springer; 10. Auflage; 2013