Wiederholung

Verteilung	Parameter	Ω	$\mathbb{P}$
Laplace	–	eine endliche Menge	$\mathbb{P}(\omega) = \frac{1}{\lvert\Omega\rvert} \quad \forall \omega \in \Omega$
Bernoulli	$p \in [0,1]$	$\{0,1\}$	$be_p(0) = 1-p$ $be_p(1) = p$
Hypergeom.	$N, n, R \in \mathbb{Z}_{\geq 0}$ $n, R \leq N$	$\{0,1,2,\dots,n\}$	$h_{N,n,R}(k) = \frac{\binom{R}{k} \cdot \binom{N-R}{n-k}}{\binom{N}{n}}$
Binomial	$n \in \mathbb{Z}_{\geq 0}$, $p \in [0,1]$	$\{0,1,2,\dots,n\}$	$b_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$ — p = W-keit für Erfolg
Poisson	$\lambda \in (0,\infty)$	$\mathbb{Z}_{\geq 0} = \{0,1,2,\dots\}$	$p_\lambda(k) = \frac{\lambda^k e^{-\lambda}}{k!}$
Geometrische	$q \in (0,1)$	$\mathbb{N} = \{1,2,3,\dots\}$	$g_q(k) = q^{k-1}(1-q)$ — $q = 1-p$

Erwartungswert

Situation

Sie haben 1,50 € und Sie wollen mehr Geld besitzen.

Drei Optionen:

(i) Sie gehen ins Kasino und spielen „Rot" im Roulette.

Gewinnen → + 1,50 €

Verlieren → − 1,50 €

(ii) Sie spielen Lotto.

# richtigen Zahlen	Gewinn
≤ 2	0 € − 1,50 €
3	+ 10,40 € − 1,50 €
4	+ 42,40 € − 1,50 €
5	+ 3 340,60 € − 1,50 €
6	+ 574.596,50 € − 1,50 €

(iii) Sie zahlen das Geld in Ihr Bankkonto ein.

Nach einem Jahr werden Sie sicher 1,51 € haben.

Welche Option ist die Beste? (Umfrage im Hörsaal: (i) ≈ 7, (ii) ≈ 6, (iii) ≈ 20)

Gedankenexperiment

Was passiert im Mittel, wenn wir diese Optionen mehrmals wiederholen würden?

Roulette

Wenn wir m-mal gespielt haben, sei m_1 die Anzahl der Spiele in denen wir gewonnen haben. Sei $m_0 = m - m_1$ die Anzahl der Spiele in denen wir verloren haben

Im Mittel: $\frac{3{,}00\,€ \cdot m_1 + 0{,}00\,€ \cdot m_0}{m} = 3{,}00\,€ \cdot \frac{m_1}{m}$.

Das ist nur ein Gedankenexperiment, also wissen wir die Werte von m und m_1 nicht.

Aber wir können erwarten, dass wenn m groß genug ist, dann

$$\frac{m_1}{m} \approx \mathbb{P}(\text{gewinnen}) = \mathrm{be}_{\frac{18}{37}}(1) \quad \text{und}$$

$$\frac{m_0}{m} \approx \mathbb{P}(\text{verlieren}) = \mathrm{be}_{\frac{18}{37}}(0) \quad \text{ist.}$$

$$\Rightarrow \underline{\text{Im Mittel}} = 3{,}00\,€ \cdot \frac{m_1}{m} + 0{,}00\,€ \cdot \frac{m_0}{m}$$
$$\approx 3{,}00\,€ \; \frac{18}{37}$$
$$\approx 1{,}46\,€.$$

Lotto

Wir spielen m-mal Für $0 \leq i \leq 1$, sei m_i die Anzahl der Spiele, wobei wir genau i richtige Zahlen ausgewählt haben.

Im Mittel:

$$\frac{(0{,}00\,€ \cdot (m_0+m_1+m_2) + 10{,}40\,€ \cdot m_3 + 42{,}40\,€ \cdot m_4 + 3.340{,}60\,€ \cdot m_5 + 574.596{,}50\,€ \cdot m_6}{m}$$

Die Hypergeometrische Verteilung passt:

$\frac{m_i}{m} \approx \mathbb{P}(\text{genau } i \text{ richtige Zahlen}) = h_{49,6,6}(i)$

<u>Im Mittel</u> $\approx 0{,}33$ € .

In mathematischer Sprache:

Wir haben mit einem W-raum angefangen:

Roulette : $(\{0,1\}, \mathrm{be}_{\frac{18}{37}})$

Lotto : $(\{0,1,\dots,6\}, h_{49,6,6})$

und eine Zufallsvariable (die beschreibt, wieviel Geld wir besitzen werden):

$X : \{0,1\} \to \mathbb{R}$

$X(0) = 0{,}00$ €

$X(1) = 3{,}00$ €

$Y : \{0,1,\dots,6\} \to \mathbb{R}$

$Y(0) = Y(1) = Y(2) = 0{,}00$ €

$Y(3) = 10{,}40$ € , ...

Dann haben wir uns gefragt: Was ist der Wert dieser Zufallsvariablen im Mittel?

Formel : $\sum_{\omega \in \Omega} X(\omega) \underbrace{\mathbb{P}(\omega)}_{\text{Wahrscheinlichkeit der Elementarereignisse}}$

Wert der Zufallsvariable

<u>Definition</u> (Erwartungswert)

Sei $(\Omega, \mathbb{P})$ ein diskreter W-raum und $X : \Omega \to \mathbb{R}$ eine Zufallsvariable.

Der <u>Erwartungswert von X</u> <u>existiert</u>, wenn die Reihe $\sum_{\omega \in \Omega} X(\Omega) \mathbb{P}(\omega)$ <u>absolut konvergent</u> (*) ist. In diesem Fall ist der Erwartungswert von X,

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\omega).$$

(*) Notiz :

1	$+1$	$+(-1)$	$+1$	$+1$	$+(-1)$	...
1	2	1	2	3	2	...

Sei $(a_n)_{n \geq 1}$ eine Folge von reellen Zahlen

$$\sum_{n=1}^{\infty} a_n = \lim_{n \to \infty} \sum_{n=1}^{N} a_n$$

$$a_n = \begin{cases} 1 & \text{wenn } n \text{ ungerade} \\ -1 & \text{wenn } n \text{ gerade} \end{cases} \Rightarrow \sum_{n=1}^{\infty} a_n = 1 + (-1) + 1 + (-1) + \dots$$

1	0	1	0	...

Eine Reihe $\sum_{n=1}^{\infty} |a_n|$ ist absolut konvergent, wenn die Reihe $\sum_{n=1}^{\infty} |a_n| < \infty$.

Bemerkung

Wenn Ω endlich ist, ist die Reihe $\sum_{\omega \in \Omega} X(\omega)\, \mathbb{P}(\omega)$ immer absolut konvergent.

$\Rightarrow$ Jede Zufallsvariable hat einen Erwartungswert.

Bemerkung

Wenn $\Omega \subseteq \mathbb{R}$, gibt es immer die Zufallsvariable

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

$$Id(\omega) = \omega.$$

Der Erwartungswert solcher W-räume ist der Erwartungswert der Identität, wenn er existiert.

Beispiel (Bernoulliraum)

$\Omega = \{0,1\} \subseteq \mathbb{R}$

$$\begin{aligned} \mathbb{E}[Id_\Omega] &= \sum_{\omega \in \Omega} Id_\Omega(\omega)\, be_p(\omega) \\ &= Id_\Omega(0) \cdot be_p(0) + Id_\Omega(1) \cdot be_p(1) \\ &= 0 \cdot (1-p) + 1 \cdot p \\ &= \underline{p}. \end{aligned}$$

z.B. Im Roulette war

$X(0) = 0{,}00$ €

$X(1) = 3{,}00$ €

Der EW hängt von X ab

In Allgemeinheit : $X : \{0,1\} \to \mathbb{R}$, $\mathbb{E}[X] = \underline{X(0)(1-p) + X(1)\,p}$

Gegenbeispiel

$\Omega = \mathbb{N}$. $\mathbb{P}$ ist ein W-maß, wobei $\mathbb{P}(i) > 0 \quad \forall i \in \mathbb{N}$.

(z.B. die geometrische Verteilung)

$$X(i) = \frac{(-1)^i}{\mathbb{P}(i)} \qquad \left(\text{z.B.} \quad X(i) = \frac{(-1)^i}{q^{i-1}(1-q)} \right)$$

Dann ist die Reihe $\sum_{i \in \mathbb{N}} X(i)\, \mathbb{P}(i) = \sum_{i=1}^{\infty} (-1)^i$

nicht absolut konvergent.

$\Rightarrow$ Der EW existiert nicht.

Wiederholung

Definition (Erwartungswert)

Sei $(\Omega, \mathbb{P})$ ein diskreter W-raum und sei $X : \Omega \to \mathbb{R}$ eine reellwertige Zufallsvariable. Der Erwartungswert von X existiert, wenn die Reihe $\sum_{\omega \in \Omega} X(\Omega)\,\mathbb{P}(\omega)$ absolut konvergent ist (d.h. $\sum_{\omega \in \Omega} |X(\omega)|\,\mathbb{P}(\omega) < \infty$).

In diesem Fall ist der Erwartungswert von X,

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\,\mathbb{P}(\omega).$$

Bemerkung

- Wenn Ω endlich ist, existiert der EW von jeder Zufallsvariable $X : \Omega \to \mathbb{R}$.
- Wenn $\Omega \subseteq \mathbb{R}$, ist die identische Abbildung $\mathrm{Id} : \Omega \to \mathbb{R}$, $\mathrm{Id}(\omega) = \omega$, eine reellwertige Zufallsvariable. Dadurch können wir die Erwartungswerte solcher Verteilungen definieren

Theorie		Anwendung der Theorie		Praxis
Standardverteilung und ihre Anwendung	$\longleftrightarrow$ Zufallsvariable	Wahrscheinlichkeitsräume $(\Omega, \mathbb{P})$	$\longleftrightarrow$ modellieren	Zufallige Situation
zum Beispiel:				Im Roulette:
$(\{0,1\}, be_{\frac{18}{37}})$	$X(0)=0$ $\longleftrightarrow$ $X(1)=2$	$\Omega = \{0,2\}$, $\mathbb{P}(0) = 19/37$, $\mathbb{P}(2) = 18/37$	$\longleftrightarrow$	Spiel „Rot"
$(\{0,1\}, be_{\frac{1}{37}})$	$X(0)=0$ $\longleftrightarrow$ $X(1)=36$	$\Omega = \{0,2\}$, $\mathbb{P}(0) = 19/37$, $\mathbb{P}(2) = 18/37$	$\longleftrightarrow$	Spiel „7"

Erwartungswert von der Binomialverteilung

- $(\{0,1,\dots,n\},\ b_{n,p})$
- ZV : $\mathrm{Id} = X : \{0,\dots,n\} \to \mathbb{R}$
- Ω ist endlich $\Rightarrow$ $\mathbb{E}[X]$ existiert.

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\omega) \\
&= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n\,(n-1)!}{(k-1)!((n-1)-(k-1))!} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= n \sum_{k=1}^{n} \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= n \sum_{k=1}^{n} \binom{n-1}{k-1} p\cdot p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\
[k\to k-1] \quad &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k} (1-p)^{(n-1)-k} \\
[\text{binomial Lehrsatz}] \quad &= np\,(p+(1-p))^{n-1} \\
&= np\, 1^{n-1}
\end{aligned}
$$

$$\boxed{\mathbb{E}[X] = np}$$

Alternativer weg

$$= np \sum_{k=0}^{n-1} b_{n-1,p}(k)$$

$$= np\; 1$$

Eigenschaften des Erwartungswertes

Behauptung

Sei $(\Omega, \mathbb{P})$ ein diskreter W-raum, und sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Der EW von X existiert genau dann, wenn die Reihe $\sum_{k\in X(\Omega)} k\,\mathbb{P}(X=k)$ absolut konvergent ist.

In diesem Fall ist

$$\mathbb{E}[X] = \sum_{k\in X(\Omega)} k\,\mathbb{P}(X=k).$$

$X(\Omega)$ ist die Menge aller möglichen Werte von X

Beweis

Wir prüfen, ob der EW existiert oder nicht.

unterteilen die Summanden nach den Wert von $X(\omega)$

$$
\begin{aligned}
\sum_{\omega\in\Omega} |X(\omega)|\,\mathbb{P}(\omega) &= \sum_{k\in X(\Omega)} \sum_{\substack{\omega\in\Omega \\ X(\omega)=k}} |X(\omega)|\,\mathbb{P}(\omega) \\
&= \sum_{k\in X(\Omega)} \sum_{\substack{\omega\in\Omega: \\ X(\omega)=k}} |k|\,\mathbb{P}(\omega)
\end{aligned}
$$

$$= \sum_{k \in X(\Omega)} |k| \sum_{\substack{\omega \in \Omega: \\ X(\omega)=k}} \mathbb{P}(\omega)$$

Ein Ereignis in $(\Omega, \mathbb{P})$

$$= \sum_{k \in X(\Omega)} |k| \; \mathbb{P}(X=k)$$

Also $\mathbb{E}[X]$ existiert $\iff \sum_{\omega \in \Omega} X(\omega)\,\mathbb{P}(\omega)$ ist absolut konvergent

$\iff \sum_{\omega \in \Omega} |X(\omega)|\,\mathbb{P}(\omega) < \infty$

$\iff \sum_{k \in X(\Omega)} |k| \; \mathbb{P}(\omega) < \infty$

$\iff \sum_{k \in X(\Omega)} k \; \mathbb{P}(\omega)$ ist absolut konvergent ✓

Mit den gleichen Manipulationen:

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\,\mathbb{P}(\omega) = \sum_{k \in X(\Omega)} k \; \mathbb{P}(X=k).$$ □

Satz (Die Linearität des Erwartungswertes)

Sei $(\Omega, \mathbb{P})$ ein diskreter W-raum, seien $X, Y : \Omega - \mathbb{R}$ Zufallsvariablen, deren Erwartungswerte existieren, und sei $c \in \mathbb{R}$ eine Konstante

Dann gilt:

(i) Der Erwartungswert von $c \cdot X$ existiert und

$$\mathbb{E}[cX] = c\,\mathbb{E}[X].$$

Der Erwartungswert von $c+X$ existiert auch und

$$\mathbb{E}[X+c] = \mathbb{E}[X] + c.$$

(ii) Der Erwartungswert von $X+Y$ existiert und

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

(iii) Wenn $X(\omega) \leq Y(\omega)$ für alle $\omega \in \Omega$, ist

$$\mathbb{E}[X] \leq \mathbb{E}[Y].$$

(iv) Für alle Ereignisse $E \subseteq \Omega$, ist

$$\mathbb{E}[\mathbb{1}_E] = \mathbb{P}(E).$$

Beweis

Im Behrends Buch. □

Der Erwartungswert der Binomialverteilung

$(\{0,1,\dots,n\}, b_{n,p})$

→ Wir wiederholen ein Experiment n-mal, wobei die W-keit von Erfolg p ist, und zählen die erfolgreichen Experimente.

Für $1 \leq i \leq n$ sei E_i das Ereignis, dass das i^{te} Experiment erfolgreich ist.

$\mathbb{E}[\mathbb{1}_{E_i}] = \mathbb{P}(E_i) = p$.

Und $\sum_{i=1}^{n} \mathbb{1}_{E_i}$ zählt die erfolgreichen Experimente.

Die Verteilung dieser Summe ist die Binomialverteilung zu den Parametern n und p.

Dann,

$$\mathbb{E}\Big[\sum_{i=1}^{n} \mathbb{1}_{E_i}\Big] \overset{[\text{Linearität des Erwartungswertes}]}{=} \sum_{i=1}^{n} \mathbb{E}[\mathbb{1}_{E_i}]$$

$$= \underline{np}.$$

Varianz und Streuung

Situation 1

Zwei Wetten:

Wir werfen eine faire Münze:

a) Kopf ⟶ Sie gewinnen 1 €

Zahl ⟶ Sie verlieren 1 €

b) Kopf ⟶ Sie gewinnen 1.000 €

Zahl ⟶ Sie verlieren 1.000 €

Sei X Ihr Gewinn in der ersten Wette.

$$\mathbb{E}[X] = \tfrac{1}{2}\cdot 1 + \tfrac{1}{2}\cdot(-1)$$
$$= \underline{0}.$$

Sei Y Ihr Gewinn in der ersten Wette.

$$\mathbb{E}[X] = \tfrac{1}{2}\cdot 1.000 + \tfrac{1}{2}\cdot(-1.000)$$
$$= \underline{0}.$$

Situation 2

a Sie gewinnen 3 € mit W-keit $\frac{1}{2}$.

Sie gewinnen 1 € mit W-keit $\frac{1}{2}$.

$$\mathbb{E}(X) = \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 1$$
$$= \underline{2}.$$

b Sie gewinnen 2.000.000 € mit W-keit $\frac{1}{2.000.000}$.

Sie gewinnen nichts mit W-keit $\frac{999.999}{2.000.000}$.

$$\mathbb{E}[X] = \frac{1}{1.000.000} \cdot 2.000.000 + \frac{999.999}{1.000.000} \cdot 0$$
$$= \underline{2}.$$

Die Varianz

Frage

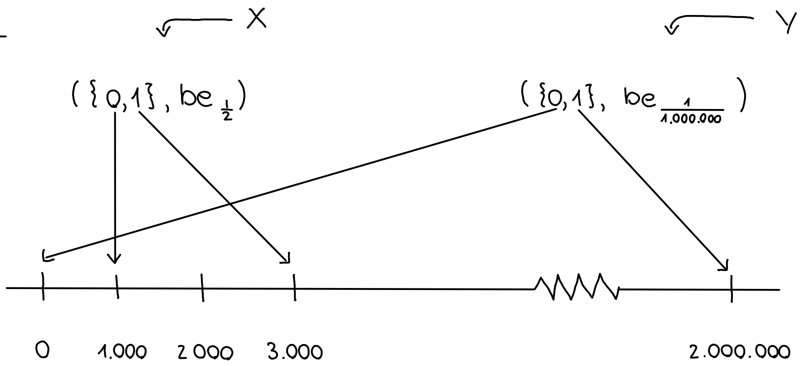
Wie gut approximiert der Erwartungswert die Verteilung?

In anderen Worten

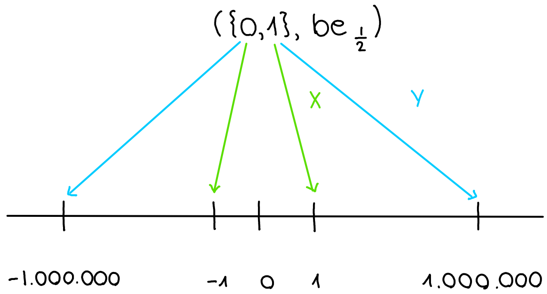
Was ist der Erwartungswert des Abstandes zwischen der Zufallsvariable und ihrem Erwartungswert?

Beispiel

Situation ①



Situation ②



$\mathbb{E}[X] = \mathbb{E}[Y]$

Wie messen wir den Abstand zwischen X und $\mathbb{E}[X]$?

Abstand $= X(\omega) - \mathbb{E}[X]$ ist **keine** gute Idee!

Die Werte dieses „Abstandes" müssen sich aufheben.

$\mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[\mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[X] = 0.$

$\Rightarrow$ Sollte nicht negativ sein.

Zwei natürliche Möglichkeiten:

(i) $\quad |X(\omega) - \mathbb{E}[X]| \geq 0$

(ii) $\quad (X(\omega) - \mathbb{E}[X])^2 \geq 0$

Definition (Varianz)

Sei $(\Omega, \mathbb{P})$ ein W-raum und $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert existiert.

Wir definieren die Varianz von X als

$$V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2],$$

falls dieser Erwartungswert existiert.

Bemerkung

Wenn Ω endlich ist, existiert die Varianz jeder Zufallsvariable auf Ω.

Beispiel

Situation ②

ⓐ
$$\begin{aligned} V(X) &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2] \\ &= (-1)^2 \tfrac{1}{2} + 1^2 \cdot \tfrac{1}{2} \\ &= 1 . \end{aligned}$$

ⓑ
$$\begin{aligned} V(Y) &= \mathbb{E}[(Y - \mathbb{E}[Y])^2] \\ &= \mathbb{E}[Y^2] \\ &= (-1.000.000)^2 \cdot \tfrac{1}{2} + 1.000.000^2 \cdot \tfrac{1}{2} \\ &= 1.000.000.000.000 \end{aligned}$$

Situation ①

ⓐ
$$\begin{aligned} V(X) &= \mathbb{E}[(X - 2.000)^2] \\ &= \sum_{\omega \in \{0,1\}} (X - 2.000)^2 \, \mathbb{P}(\omega) \\ &= (-1000)^2 \cdot \tfrac{1}{2} + 1000^2 \cdot \tfrac{1}{2} \\ &= 1.000.000 \end{aligned}$$

ⓑ
$$\begin{aligned} V(Y) &= \mathbb{E}[(Y - 2.000)^2] \\ &= \sum_{\omega \in \{0,1\}} (Y(\omega) - 2000)^2 \, \mathbb{P}(\omega) \\ &= (-2.000)^2 \left(1 - \tfrac{1}{1.000.000}\right) + (2.000.000.000 - 2000)^2 \cdot \tfrac{1}{1.000.000} \\ &= 3\,999.996.000.000 \end{aligned}$$

Die Streuung

Problem

Die Varianz und der Erwartungswert haben verschiedene Einheiten.

$\mathbb{E}[X] \to €$, $Var(X) \to €^2$

Definition (Streuung)

Sei $(\Omega, \mathbb{P})$ ein W-raum und $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Varianz existiert. Wir definieren die Streuung von X als

$$\sigma(X) = \sqrt{Var(X)} .$$

Bemerkung

$V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] \geq \mathbb{E}[0] = 0$

$\Rightarrow$ $\sigma(X)$ ist wohldefiniert

Eigenschaften von Varianz und Streuung

Satz

Sei $(\Omega, \mathbb{P})$ ein diskreter W-raum, $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert und Varianz existiert, und sei $c \in \mathbb{R}$.

Dann

(i) Die Varianz von cX existiert und

$Var(cX) = c^2 \; Var(X)$ und

$\sigma(cX) = |c| \; \sigma(X)$.

(ii) Die Varianz von $X+c$ existiert und

$V(X+c) = V(X)$ und

$\sigma(X+c) = \sigma(X)$.

(iii) $\mathbb{E}[X^2]$ existiert und

$V(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Beweis

(i) Die Varianz von cX existiert genau dann, wenn

$\mathbb{E}[cX]$ existiert und $\mathbb{E}[(cX-\mathbb{E}[cX])^2]$ existiert.

$\mathbb{E}[cX]$ existiert, und $\mathbb{E}[cX] = c\,\mathbb{E}[X]$.

Und $\mathbb{E}[(cX-\mathbb{E}[cX])^2]$ existiert genau dann, wenn

$\sum\limits_{\omega\in\Omega} (cX(\omega)-\mathbb{E}[cX])^2\,\mathbb{P}(\omega)$ absolut konvergent ist

$$\begin{aligned}\sum_{\omega\in\Omega} (cX(\omega)-\mathbb{E}[cX])^2\,\mathbb{P}(\omega) &= \sum_{\omega\in\Omega} c^2\,(X(\omega)-\mathbb{E}[X])^2\,\mathbb{P}(\omega)\\ &= c^2 \sum_{\omega\in\Omega} (X(\omega)-\mathbb{E}[X])^2\,\mathbb{P}(\omega)\\ &= c^2\,\mathrm{Var}(X) \qquad < \infty \qquad \text{, weil } V(X)<\infty\end{aligned}$$

(ii)
$$\begin{aligned}V(X+c) &= \mathbb{E}[(X+c-\mathbb{E}[X+c])^2]\\ &= \mathbb{E}[(X-\mathbb{E}[X])^2]\\ &= V(X).\end{aligned}$$

(iii) Wir definieren $c = \mathbb{E}[X]$.

$\Rightarrow$ $\mathbb{E}[2cX]$ existiert und

$\mathbb{E}[2cX] = 2c\,\mathbb{E}[X]$

existiert

$$\begin{aligned}\Rightarrow \mathbb{E}[(X-\mathbb{E}[X])^2+2cX] &= \mathbb{E}[X^2-2cX+\mathbb{E}[X]^2+2cX]\\ &= \mathbb{E}[X^2+\mathbb{E}[X]^2]\\ &= \mathbb{E}[X^2]+\mathbb{E}[X]^2\end{aligned}$$

$$\begin{aligned}\text{Aber } \mathbb{E}[(X-\mathbb{E}[X])^2+2cX] &= \mathbb{E}[(X-\mathbb{E}[X])^2]+2c\,\mathbb{E}[X]\\ &= V(X)+2\,\mathbb{E}[X]^2\end{aligned}$$

, und $\mathbb{E}[X^2+c^2]$ existiert.

$\Rightarrow$ $V(X) = \mathbb{E}[X^2]-\mathbb{E}[X]^2$

und $\mathbb{E}[X^2] = \mathbb{E}[X^2+c^2+(-c^2)]$ (d h er existiert)

$\square$

Wiederholung

Definition (Varianz und Streuung)

Sei $(\Omega, \mathbb{P})$ ein W-raum und $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert existiert.

Wir definieren die Varianz von X als

$$V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{\omega \in \Omega} (X(\omega) - \mathbb{E}[X])^2 \, \mathbb{P}(\omega)$$

falls dieser Erwartungswert existiert. Dann definieren wir die Streuung von X als

$$\sigma(X) = \sqrt{V(X)} .$$

Satz

Sei $(\Omega, \mathbb{P})$ ein W-raum, $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Varianz existiert, und $c \in \mathbb{R}$ eine Konstante

Dann gilt.

(i) Die Varianz von cX existiert und

$$\mathrm{Var}(cX) = c^2 \, \mathrm{Var}(X) \quad \text{und} \quad \sigma(cX) = |c| \, \sigma(X) .$$

(ii) Die Varianz von $X+c$ existiert und

$$V(X+c) = V(X) \quad \text{und} \quad \sigma(X+c) = \sigma(X) .$$

(iii) $\mathbb{E}[X^2]$ existiert und $V(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$.

Bemerkung

Sei $(a_n)_{n \geq 1}$ eine Folge von nicht negativen reellen Zahlen.

Die Reihe $\sum_{n \geq 1} a_n$ ist absolut konvergent genau dann, wenn

$$\sum_{n \geq 1} a_n \left(= \sum_{n \geq 1} |a_n| \right) < \infty$$

Der Bernoulliraum

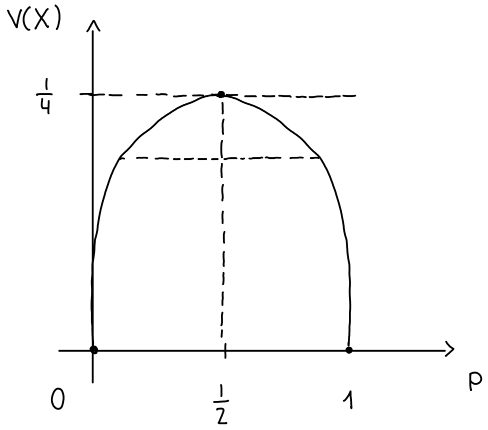
Behauptung

Sei X eine Zufallsvariable, deren Verteilung der Bernoulliraum zum Parameter p ist. Dann gilt:

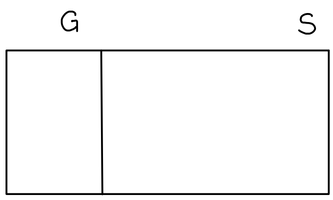
$$V(X) = p(1-p) .$$

Beweis

H.A. $\square$

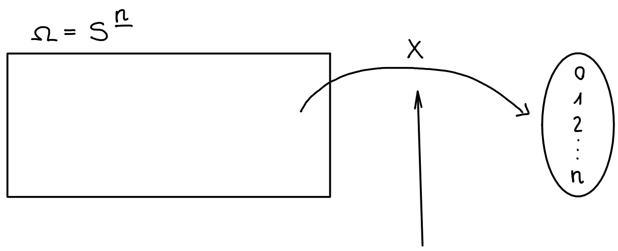


Die Hypergeometrische Verteilung



$|S| = N$

$|G| = R$



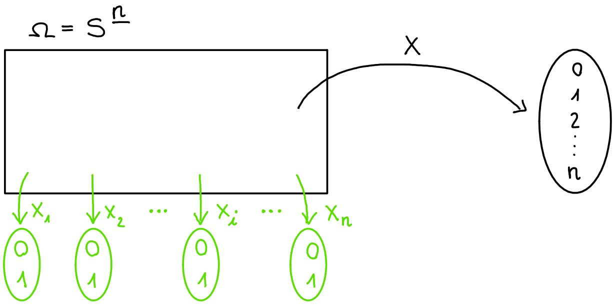
zählt die günstigen Elemente, die wir ausgewählt haben.

- Der Induzierte W-raum auf $\{0, 1, \ldots, n\}$ ist die hypergeometrische Verteilung zu den Parametern N, n und R.
- Ω ist endlich $\Rightarrow$ $V(X)$ existiert.
- Zuerst müssen wir den Erwartungswert von X berechnen

Behauptung

Der Erwartungswert der hypergeometrischen Verteilung zu den Parameter N, n und R ist $\frac{nR}{N}$.

Beweis



wir werden n Zufallsvariablen definieren, die den Bernoulliraum als Verteilung haben.

Wir definieren die Ereignisse E_i, $1 \leq i \leq n$, wobei

$E_i = \{$ das i^{te} Element ist günstig $\}$.

Seien $X_i = \mathbb{1}_{E_i}$ die Indikatorfunktion für diese Ereignisse, dann gilt

$$X = \sum_{i=1}^{n} X_i .$$

[Linearität des Erwartungswertes]

$$\mathbb{E}[X] = \mathbb{E}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \mathbb{P}(E_i) .$$

$$\mathbb{P}(E_i) = \frac{|E_i|}{|S^{\underline{n}}|} = \frac{|E_i|}{|S|^{\underline{n}}}$$

Wieviele n-Permutationen gibt es, wobei das i^{te} Element günstig ist?

$$\boxed{\frac{N-1}{1}} \quad \frac{N-2}{2} \quad \frac{N-3}{3} \quad \frac{N-4}{4} \quad \cdots \quad \frac{N-(i-1)}{i-1} \quad \boxed{\frac{R}{i}} \quad \frac{N-i}{i+1} \quad \cdots \quad \frac{N-(n-3)}{n-2} \quad \frac{N-(n-2)}{n-1} \quad \frac{N-(n-1)}{n}$$

← Möglichkeiten für jede Stelle
← Stellen

- Zuerst reservieren wir ein günstiges Element für die i^{te} Stelle.

  → R Möglichkeiten

- Dann wählen wir das erste Element aus

  → $N-1$ Möglichkeiten

- $|E_i| = (N-1)(N-2)\ldots(N-(i-1))\,R\,(N-i)\ldots(N-(n-1)) = R \cdot \frac{(N-1)!}{(N-n)!}$

$$\Rightarrow \quad \mathbb{P}(E_i) = \frac{|E_i|}{|S|^{\underline{n}}}$$

$$= \frac{\left(R \cdot \frac{(N-1)!}{\cancel{(N-n)!}}\right)}{\left(\frac{N!}{\cancel{(N-n)!}}\right)}$$

$$\mathbb{P}(E_i) = \frac{R}{N}$$

$$\Rightarrow \quad \mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{P}(E_i)$$

$$= \sum_{i=1}^{n} \frac{R}{N}$$

$\left[\frac{R}{N} \text{ hängt nicht von } i \text{ ab}\right]$

$$= \frac{nR}{N}. \qquad \square$$

Behauptung

Die Varianz der hypergeometrische Verteilung zu den Parametern N, n und R ist

$$V(X) = \frac{nR\,(N-R)(N-n)}{N^2\,(N-1)}$$

Beweis

Zuerst berechnen wir $\mathbb{E}[X(X-1)]$.

$$\mathbb{E}[X(X-1)] = \sum_{k=0}^{n} k\,(k-1)\;\mathbb{P}(X=k)$$

$$= \sum_{k=0}^{n} k\,(k-1)\;\frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}$$

$\left[\text{Sei } b \geq 2 \text{ : } \binom{a}{b} = \frac{a(a-1)}{b(b-1)} = \binom{a-1}{b-1}\right]$

$$= \sum_{k=2}^{n} k\,(k-1)\;\frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}$$

$(= 0$ wenn $n \leq 1)$

$$= \sum_{k=2}^{n} \frac{k(k-1)\,\frac{R(R-1)}{k(k-1)}\binom{R-2}{k-2}\binom{N-R}{n-k}}{\frac{N(N-1)}{n(n-1)}\binom{N-2}{n-2}}$$

$$= \frac{n(n-1)\,R(R-1)}{N(N-1)} \sum_{k=2}^{n} \frac{\binom{R-2}{k-2}\binom{N-R}{n-k}}{\binom{N-2}{n-2}}$$

$$= \frac{n(n-1)\,R(R-1)}{N(N-1)} \sum_{k=2}^{n} \frac{\binom{R-2}{k-2}\binom{(N-2)-(R-2)}{(n-2)-(k-2)}}{\binom{N-2}{n-2}}$$

$$= \frac{n(n-1)\,R(R-1)}{N(N-1)} \sum_{k=0}^{n-2} \frac{\binom{R-2}{k}\binom{(N-2)-(R-2)}{(n-2)-k}}{\binom{N-2}{n-2}}$$

$$= \frac{n(n-1)\,R(R-1)}{N(N-1)} \sum_{k=0}^{n-2} \underbrace{h_{N-2,\,n-2,\,R-2}(k)}_{=1}$$

$$\Rightarrow \quad \mathbb{E}[X(X-1)] = \frac{n(n-1)\,R(R-1)}{N(N-1)} . \qquad \checkmark$$

$$\mathbb{E}[X^2] = \mathbb{E}[X(X-1)+X]$$

[L d $\mathbb{E}$]

Linearität des Erwartungswertes

$$= \mathbb{E}[X(X-1)] + \mathbb{E}[X]$$

$$= \frac{n(n-1)\,R(R-1)}{N(N-1)} + \frac{nR}{N} .$$

$$V(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$= \frac{n(n-1)\,R(R-1)}{N(N-1)} + \frac{nR}{N} - \left(\frac{nR}{N}\right)^2$$

$$= \frac{nR}{N^2(N-1)} \big((n-1)(R-1)N + N(N-1) - nR(N-1) \big)$$

$$= \frac{nR(N-R)(N-n)}{N^2(N-1)} \qquad \square$$

Bemerkung

$$X = \sum_{i=1}^{n} X_i$$

$$X \sim (\{0,1,\dots,n\},\ h_{N,n,R})$$

$$V(X) = \frac{nR(N-R)(N-n)}{N^2(N-1)} = n\left(\frac{R}{N}\right)\left(1-\frac{R}{N}\right)\frac{N-n}{N-1}$$

$$X_i \sim (\{0,1\},\ \mathrm{be}_{\frac{R}{N}})$$

$$\sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} \frac{R}{N}\left(1-\frac{R}{N}\right) = n\left(\frac{R}{N}\right)\left(1-\frac{R}{N}\right)$$

wenn $n \geq 2$, $\quad V(X) < \sum_{i=1}^{n} V(X_i)$.

Aber

$$Y \sim (\{0,1,\dots,n\},\ b_{n,p})$$

$$Y = \sum_{i=1}^{n} Y_i$$

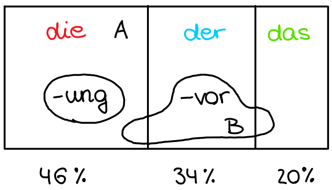
$$Y_i \sim (\{0,1\},\ \mathrm{be}_p)$$

Hausaufgabe

$$V(Y) = n \cdot p\,(1-p) = \sum_{i=1}^{n} p\,(1-p) = \sum_{i=1}^{n} V(Y_i)\,.$$

Bedingte Wahrscheinlichkeit

Situation



$\Omega = \{$ deutsche Wörter $\}$

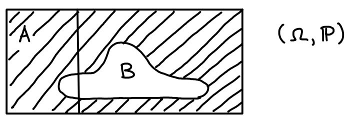
Die Nomina enthalten Informationen über den Artikel.

Diese Informationen verändern die W-keiten.

Wie können wir die neuen W-keiten berechnen?

In größerer Allgemeinheit:

- Ein W-raum $(\Omega, \mathbb{P})$
- Ein Ereignis $A \subseteq \Omega$, für welches wir uns interessieren.
- Ein Ereignis $B \subseteq \Omega$, worüber wir wissen, dass es schon aufgetreten ist.



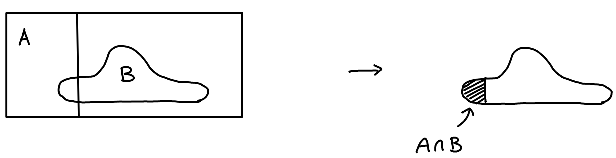
$\Rightarrow$ Alle $\omega \in \Omega$, die nicht in B liegen, sind unerheblich.

Deshalb brauchen wir, wenn A auftritt, dass $A \cap B$ aufgetreten ist.

Ein neuer W-raum $(\Omega', \mathbb{P}')$ für den gilt

$\Omega' = B$

$A \mapsto A \cap B$.



- Wir brauchen folgende zwei Sachen:

  → $\mathbb{P}'(B)$

  $$\frac{\mathbb{P}'(E_1)}{\mathbb{P}(E_1)} = \frac{\mathbb{P}'(E_2)}{\mathbb{P}(E_2)} \qquad \text{für alle } E_1, E_2 \subseteq B \, .$$

<u>Definition</u> (Bedingte Wahrscheinlichkeit)

Sei $(\Omega, \mathbb{P})$ ein W-raum und seien $A, B \subseteq \Omega$ Ereignisse,

sodass $\mathbb{P}(B) > 0$

Dann definieren wir die <u>bedingte W-keit von A unter B</u> als

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \, .$$

Wiederholung

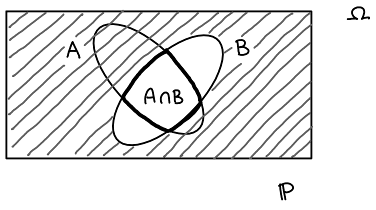
Definition (Bedingte Wahrscheinlichkeit)

Sei $(\Omega, \mathbb{P})$ ein W-raum und seien $A, B \subseteq \Omega$ Ereignisse,

wobei $\mathbb{P}(B) > 0$

Die bedingte W-keit von A unter der Vorraussetzung B ist

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} .$$



$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

$(B, \mathbb{P}(\cdot | B))$ ist ein diskreter W-raum:

- B ist höchstens abzählbar

  ($B \subseteq \Omega$, Ω ist höchstens abzählbar)

- Für jedes $\omega \in B$, $\mathbb{P}(\{\omega\} | B) \geq 0$

  ($\mathbb{P}(\{\omega\} | B)$, $\mathbb{P}(B) \geq 0$)

- $\sum_{\omega \in \Omega} \mathbb{P}(\{\omega\} | B) = \sum_{\omega \in \Omega} \frac{\mathbb{P}(\{\omega\} \cap B)}{\mathbb{P}(B)} = \frac{\sum_{\omega \in \Omega} \mathbb{P}(\{\omega\})}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1 .$ ✓

Beispiel

Situation

Wir würfeln einen fairen Würfel und wir gewinnen,

wenn die Augensumme größer / gleich zehn ist.

W-raum . Laplace auf $\Omega = [6]^2$.

Sei A das Ereignis, dass wir gewinnen.

Das heißt:

$A = \{ (4,6), (5,6), (6,4), (6,5, (6,6) \}$

$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{6}{6^2} = \underline{\frac{1}{6}} .$

Was passiert, wenn wir die erste Augenzahl schon gesehen haben?

Für $1 \leq i \leq 6$ sei B_i das Ereignis, dass die erste Augenzahl i ist. ($B_i = \{(i,1),(i,2),\dots,(i,6)\}$)

Fairer Würfel $\Rightarrow \mathbb{P}(B_i) = \frac{1}{6}$.

$\mathbb{P}(A \mid B_i)$?

i	$A \cap B_i$	$\mathbb{P}(A \mid B_i)$
1	$\emptyset$	0
2	$\emptyset$	0
3	$\emptyset$	0
4	$\{(4,6)\}$	$(\frac{1}{36})/(\frac{1}{6}) = \frac{1}{6}$
5	$\{(5,5),(5,6)\}$	$(\frac{2}{36})/(\frac{1}{6}) = \frac{2}{6}$
6	$\{(6,4),(6,5),(6,6)\}$	$(\frac{3}{36})/(\frac{1}{6}) = \frac{3}{6}$

Unabhängigkeit

Definition (Unabhängigkeit)

Zwei Ereignisse $A, B \subseteq \Omega$ in einem diskreten W-raum sind <u>unabhängig</u>, wenn

$$\boxed{\mathbb{P}(A \mid B) = \mathbb{P}(A)}.$$

Bemerkung

Seien A, B zwei Ereignisse, sodass

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A).$$

können wir auch (im Gegensatz zur vorherigen Definition) nutzen wenn $\mathbb{P}(A) = 0$ oder $\mathbb{P}(B) = 0$.

Dann gilt

$$\boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)}$$

wird oft als Definition für Unabhängigkeit verwendet.

und deshalb

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\cancel{\mathbb{P}(A)} \cdot \mathbb{P}(B)}{\cancel{\mathbb{P}(A)}} = \mathbb{P}(B).$$

(wir nehmen an, dass auch $\mathbb{P}(A) > 0$)

Bemerkung

- $\emptyset$ und Ω sind unabhängig von A, für jedes Ereignis $A \subseteq \Omega$.

$\mathbb{P}(\emptyset \cap A) = \mathbb{P}(\emptyset) = 0 = \mathbb{P}(\emptyset) \cdot \mathbb{P}(A)$. ✓

$\mathbb{P}(\Omega \cap A) = \mathbb{P}(A) = 1 \cdot \mathbb{P}(A) = \mathbb{P}(\Omega)\ \mathbb{P}(A)$. ✓

Achtung

Seien $A, B \subseteq \Omega$ Ereignisse, sodass:

- $\mathbb{P}(A), \mathbb{P}(B) > 0$
- A und B disjunkt sind.

Dann sind A und B <u>nicht</u> unabhängig.

$\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0 \neq \mathbb{P}(A)\,\mathbb{P}(B)$.

Der Satz der totalen Wahrscheinlichkeit

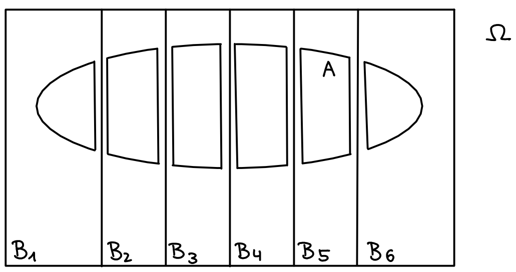
Satz (Satz der totalen W-keit)

Sei $(\Omega, \mathbb{P})$ ein W-raum und sei $\Omega = \dot{\bigcup}_{i \in I} B_i$ eine Zerlegung von Ω in paarweise disjunkte Ereignisse B_i, $i \in I$, sodass

- I höchstens abzählbar ist,
- $\mathbb{P}(B_i) > 0$ ist, für jedes $i \in I$

Dann gilt für jedes Ereignis $A \subseteq \Omega$:

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid B_i)\,\mathbb{P}(B_i).$$



Dieser Satz ist die Summenregel der W-theorie:

Wenn die Berechnung von $\mathbb{P}(A)$ schwierig ist, unterteilen wir die Elemente in Ω nach B_i , $i \in I$,

sodass $\mathbb{P}(A|B_i)$ einfacher zu berechnen ist.

Beweis

Weil $\Omega = \dot{\bigcup}_{i\in I} B_i$ eine Zerlegung ist, haben wir die disjunkte Vereinigung:

$$A = \dot{\bigcup_{i\in I}} (A \cap B_i)$$

$$\begin{aligned} \Rightarrow \quad \mathbb{P}(A) &= \mathbb{P}\Big(\dot{\bigcup_{i\in I}} (A\cap B_i)\Big) \\ &= \sum_{i\in I} \mathbb{P}(A\cap B_i) \\ &= \sum_{i\in I} \mathbb{P}(A|B_i)\, \mathbb{P}(B_i) \qquad \square \end{aligned}$$

Beispiel

Situation

Wir spielen Roulette mit Ronaldo. Wir werfen eine faire Münze:

Kopf $\rightarrow$ wir wetten auf „Rot"

Zahl $\rightarrow$ wir wetten auf „7"

Was ist die W-keit vom Gewinn?

Ereignisse

A = {wir gewinnen}

B_1 = {wir wetten „Rot"}

B_2 = {wir wetten „7"}

$\Omega = B_1 \dot{\cup} B_2$ $\quad$ ($\Rightarrow$ Ω = {(unsere Wette, das Ereignis)} = $\{\underset{\text{Kopf}}{7}, \underset{\text{Zahl}}{\text{Rot}}\} \times \{0,1,2,\dots,36\}$)

Die bedingte W-keit

$\mathbb{P}(A|B_1)$ = $\mathbb{P}$(wir gewinnen, wenn wir „Rot" wetten }

$= \frac{18}{37}$

$\mathbb{P}(A|B_2)$ = $\mathbb{P}$(wir gewinnen, wenn wir „7" wetten }

$= \frac{1}{37}$

Weil die Münze fair ist, ist

$$\mathbb{P}(B_1) = \frac{1}{2} \quad \text{und} \quad \mathbb{P}(B_2) = \frac{1}{2}.$$

Deshalb

$$\mathbb{P}(A) \overset{[\text{sdtW}]}{=} \mathbb{P}(A|B_1)\,\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\,\mathbb{P}(B_2)$$

$$= \frac{18}{37} \cdot \frac{1}{2} + \frac{1}{37} \cdot \frac{1}{2}$$

$$= \frac{19}{74}.$$

Der Satz von Bayes

Manchmal, wenn wir schon wissen, dass A aufgetreten ist, möchten wir auch wissen, in welchen Fall $(B_i, i \in I)$ wir liegen.

Das heißt, dass wir die bedingte W-keit $\mathbb{P}(B_k|A)$ berechnen möchten.

Satz (Bayes)

Sei $(\Omega, \mathbb{P})$ ein W-raum, und sei $\Omega = \dot{\bigcup}_{i \in I} B_i$ eine Zerlegung von Ω in paarweise disjunkte Ereignisse B_i, sodass

- I höchstens abzählbar ist
- für jedes $i \in I$, $\mathbb{P}(B_i) > 0$.

Dann gilt für jedes Ereignis $A \subseteq \Omega$ und für jedes $k \in I$,

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\,\mathbb{P}(B)}{\sum_{i \in I} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)}$$

Beweis

$$\mathbb{P}(B_k|A) \overset{[\text{Def.}]}{=} \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)}.$$

$$\mathbb{P}(B_k \cap A) = \mathbb{P}(A \cap B_k) = \mathbb{P}(A|B_k)\,\mathbb{P}(B_k).$$

$$\mathbb{P}(A) \overset{[\text{sdtW}]}{=} \sum_{i \in I} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i) \qquad \square$$

Wiederholung

Definition (Unabhängigkeit)

Zwei Ereignisse $A, B \subseteq \Omega$ in einem W-raum sind <u>unabhängig</u>, wenn

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Satz (Satz der totalen W-keit)

Sei $(\Omega, \mathbb{P})$ ein W-raum und sei $\Omega = \dot{\bigcup}_{i \in I} B_i$ eine Zerlegung von Ω in paarweise disjunkte Ereignisse B_i, $i \in I$, wobei I höchstens abzählbar ist, und $\mathbb{P}(B_i) > 0$ $\forall i \in I$. Dann gilt für alle $A \subseteq \Omega$:

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i) .$$

Satz (Bayes)

Sei $(\Omega, \mathbb{P})$ ein W-raum, $\Omega = \dot{\bigcup}_{i \in I} B_i$, so wie im Satz der totalen W-keit. Dann gilt für jedes Ereignis $A \subseteq \Omega$ und für jedes $k \in I$,

$$\mathbb{P}(B_k | A) = \frac{\mathbb{P}(A|B_k)\,\mathbb{P}(B)}{\sum_{i \in I} \mathbb{P}(A|B_i)\,\mathbb{P}(B_i)} .$$

Anwendung von Bayes

Situation

Es gibt eine Krankheit K. Es gibt einen Test, um herauszufinden, ob eine Person K hat oder nicht.

Wir wissen die <u>Prävalenz</u> der Krankheit:

0,3% der Bevölkerung haben K.

Wir werden eine Person der Bevölkerung zufällig auswählen und sie testen.

Es gibt zwei Arten von Fehlern:

- Die Person ist krank, aber unser Test zeigt, dass sie gesund ist.
- Die Person ist gesund, aber unser Test zeigt, dass sie krank ist.

<u>Eigenschaften unseres Tests</u>

- Die <u>Sensivität</u> des Testes ist 99,9 % : das ist die W-keit davon, dass der Test eine kranke Person als krank zeigt
- Die <u>Spezifität</u> des Testes ist 98 % : das ist die W-keit davon, dass der Test eine gesunde Person als gesund erkennen wird.

<u>Frage</u>

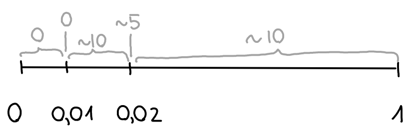
Wir haben eine Person zufällig ausgewählt und getestet, und der Test zeigt, dass die Person krank ist.

Was ist die W-keit, dass diese Person in Wirklichkeit eigentlich gesund ist?

<u>Lösung</u>

Ω ist die Menge von Paaren von Personen und Ereignissen unseres Tests.

Umfrage im Hörsaal:



<u>Ereignisse</u> : $\Omega = B_1 \dot\cup B_2$

B_1 = { die Person ist (wirklich) krank }

B_2 = { die Person ist (wirklich) gesund }

A = { der Test zeigt, dass die Person krank ist }

Wir möchten wissen:

$$\mathbb{P}(B_2 \mid A)$$

Wir wissen:

$\mathbb{P}(B_1) = 0{,}003$ [Prävalenz]

$\mathbb{P}(B_2) = \mathbb{P}(\Omega \setminus B_1) = 1 - \mathbb{P}(B_1) = 0{,}997$

$\mathbb{P}(A \mid B_1) = 0{,}99$ [Sensitivität]

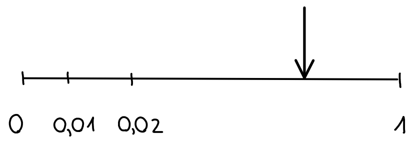
$\mathbb{P}(A \mid B_2) = 1 - \mathbb{P}((\Omega \setminus A) \mid B_2) = 1 - 0{,}98 = 0{,}02$ [Spezifität]

Also wenden wir den Satz von Bayes an:

$$\mathbb{P}(B_2 \mid A) = \frac{\mathbb{P}(A|B_2)\,\mathbb{P}(B_2)}{\mathbb{P}(A|B_1)\,\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\,\mathbb{P}(B_2)}$$

$$= \frac{0{,}002 \cdot 0{,}997}{0{,}99 \cdot 0{,}003 + 0{,}002 \cdot 0{,}997}$$

$$= 0{,}87\ldots \approx 87\,\%$$



Unabhängigkeit von vielen Ereignissen

Situation

Zwei faire Münzen werden geworfen.

- E_1 : die erste Münze zeigt Kopf.
- E_2 : die zweite Münze zeigt Zahl.
- E_3 : die zwei Münzen zeigen verschiedene Seiten.

Der W-raum

$\Omega = \{(K,K), (K,Z), (Z,K), (Z,Z)\}$

ist Laplaceraum ($\mathbb{P} = \frac{1}{4}$).

$E_1 = \{(K,K), (K,Z)\}$

$E_2 = \{(K,Z), (Z,Z)\}$

$E_3 = \{(K,Z), (Z,K)\}$

$\mathbb{P}(E_1) = \frac{|E_1|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}$

$\mathbb{P}(E_2) = \mathbb{P}(E_3) = \frac{1}{2}$

$\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(\{(K,Z)\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(E_1)\,\mathbb{P}(E_2)$

$\Rightarrow$ E_1 und E_2 sind unabhängig.

$\mathbb{P}(E_1 \cap E_3) = \mathbb{P}(\{(K,Z)\}) = \frac{1}{4} = \mathbb{P}(E_1)\,\mathbb{P}(E_3)$

$\Rightarrow$ E_1 und E_3 sind unabhängig.

$\mathbb{P}(E_2 \cap E_3) = \mathbb{P}(E_2)\,\mathbb{P}(E_3)$

$\Rightarrow$ E_2 und E_3 sind unabhängig.

<u>Aber</u>. E_1 und E_2 zusammen bestimmen E_3. $(E_1 \cap E_2 \subseteq E_3)$

<u>Definition</u>

Sei $(\Omega, \mathbb{P})$ ein W-raum und sei I eine beliebige Indexmenge. (I kann auch unendlich oder überabzählbar sein)

Eine Menge $\{E_i : i \in I\}$ von Ereignissen $E_i \subseteq \Omega$ ist <u>unabhängig</u>, wenn

$$\mathbb{P}\Big(\bigcap_{j \in J} E_j\Big) = \prod_{j \in J} \mathbb{P}(E_j)$$

für jede endliche Teilmenge $J \subseteq I$.

<u>Bemerkung</u>

Die Menge $\{E_1, E_2, E_3\}$ von unserem Beispiel ist <u>nicht</u> unabhängig.

$J = \{1,2,3\}$

$$\mathbb{P}\Big(\bigcap_{j \in J} E_j\Big) = \mathbb{P}(\{K,Z\}) = \frac{1}{4} \neq \underbrace{\frac{1}{8}}_{= \prod_{j \in J} \mathbb{P}(E_j)}$$

<u>Bemerkung</u>

unser Beispiel

$$\boxed{\{E_i : i \in I\} \text{ ist unabhängig}} \quad \underset{J = \{i,j\}}{\overset{\not\Leftarrow}{\Longrightarrow}} \quad \boxed{\text{Für jedes Paar } \{i,j\} \subseteq I,\ E_i \text{ und } E_j \text{ sind unabhängig}}$$

Unabhängigkeit von Zufallsvariablen

Beispiel

Wir wählen zufällig eine Person aus.

- X : ihr Monatslohn (in €)
- Y : die Größe ihrer Wohnung (in m^2)
- Z : der Tag ihres Geburtstages

– nicht „unabhängig"

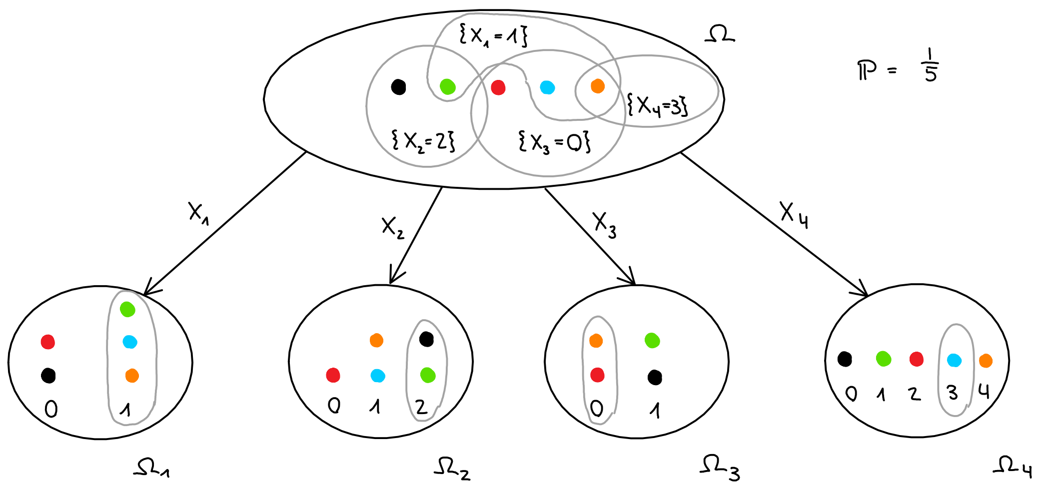
„unabhängig"

Definition

Sei $(\Omega, \mathbb{P})$ ein W-raum und sei I eine beliebige Indexmenge. Für jedes $i \in I$ gibt es eine Zufallsvariable

$X_i : \Omega \to \Omega_i$

Die Menge $\{X_i : i \in I\}$ von Zufallsvariablen ist <u>unabhängig</u>, wenn für beliebige Wahl von Elementarereignissen $\omega_i \in \Omega_i$, die Menge $\{\{X_i = \omega_i\} : i \in I\}$ von Ereignissen unabhängig ist.



Beispiel

Seien E_1, E_2 zwei Ereignisse in $(\Omega, \mathbb{P})$, und seien $\mathbb{1}_{E_1}, \mathbb{1}_{E_2}$ ihre Indikatorfunktionen.

Dann ist $\{\mathbb{1}_{E_1}, \mathbb{1}_{E_2}\}$ unabhängig, genau dann, wenn

$\{E_1, E_2\}$ unabhängig ist.

zz.

$$\boxed{\{E_1, E_2\} \text{ ist unabhängig}} \quad \overset{(i)}{\Leftarrow} \; \overset{(ii)}{\Rightarrow} \quad \boxed{\{\mathbb{1}_{E_1}, \mathbb{1}_{E_2}\} \text{ ist unabhängig}}$$

(i) $\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(\{\mathbb{1}_{E_1} = 1\} \cap \{\mathbb{1}_{E_2} = 1\})$

$[\mathbb{1}_{E_1}, \mathbb{1}_{E_2} \text{ unabh}]$
$= \mathbb{P}(\{\mathbb{1}_{E_1} = 1\}) \cdot \mathbb{P}(\{\mathbb{1}_{E_2} = 1\})$

$= \mathbb{P}(E_1) \cdot \mathbb{P}(E_2)$

$\Rightarrow E_1, E_2$ sind unabhängig

(ii) H.A.

Wiederholung

Die geometrische Verteilung

Situation

Wir wiederholen ein Experiment, das Misserfolgsquote q hat,

bis es zum ersten Erfolg kommt

Wie oft müssen wir das Experiment wiederholen?

$$\mathbb{P}(X=k): \quad \underbrace{\underline{M}\ \underline{M}\ \underline{M} \dots \underline{M}}_{\text{Ersten } k-1 \text{ Experimente}} \ \underbrace{\underline{E}}_{k^{\text{te}} \text{ Experiment}}$$

$$\text{W-keit} \quad q \quad q \quad q \quad q \quad (1-q) \quad \rightarrow \quad q^{k-1}(1-q)$$

Frage

Gibt es einen W-raum Ω und eine Zufallsvariable $X: \Omega \to \mathbb{N}$,

sodass die induzierte Verteilung von X die geometrische

Verteilung ist?

- Nehmen wir $q = \frac{1}{2}$ an.

  $\Rightarrow$ Erfolg und Misserfolg haben die gleiche W-keit.

  $\Rightarrow$ Die Folge von Ergebnissen sind gleichverteilt

- Wir können die Länge der Folgen von oben nicht begrenzen.

  $\Rightarrow$ Wir müssen unendliche Folgen von Experimenten / Ergebnissen bedenken.

Wir werden die Folgen so bezeichnen:

$0 \longleftrightarrow$ Misserfolg

$1 \longleftrightarrow$ Erfolg.

Deshalb ist $\Omega = \{0,1\}^{\mathbb{N}} = \{(\omega_1, \omega_2, \omega_3, \dots) : \omega_i \in \{0,1\} \ \forall i\}$

Zufallsvariable

X zählt wieviele Experimente es gibt bis zum ersten Erfolg.

zum Beispiel

$\omega = (1,0,1,1,1,0,0,0,\dots)$

$X(\omega) = 1$: das erste Experiment war erfolgreich.

$\omega = (0,0,1,1,0,\dots)$

$X(\omega) = 3$

allgemein : $X(\omega) = \min\{i : \omega_i = 1\}$.

Jetzt haben wir eine Menge Ω und eine Zufallsvariable $X: \Omega \to \mathbb{N}$,

die die geometrische Verteilung beschreibt.

Aber $\mathbb{P}$ fehlt !

Zwei Probleme (Übungsblatt 0)

① Ω ist überabzählbar ! $(\Omega = \{0,1\}^{\mathbb{N}})$

② Ω ist unendlich

$\Rightarrow$ es gibt keinen Laplaceraum auf Ω.

Eigenschaften von überabzählbaren W-räumen

„Die Folgen sind gleichverteilt"

$\Leftrightarrow$ $\forall \omega, \omega' \in \Omega$, $\mathbb{P}(\omega) = \mathbb{P}(\omega')$

Behauptung

$\forall \omega \in \Omega$, $\mathbb{P}(\omega) = 0$.

Beweis

Sei $p \in [0,1]$, sodass $\mathbb{P}(\omega) = p$, $\forall \omega \in \Omega$.

$$1 = \mathbb{P}(\Omega) \triangleq \mathbb{P}(\{(1,0,0,\dots), (0,1,0,0,\dots), (0,0,1,0,\dots), \dots\})$$

$$= \mathbb{P}((1,0,0,\dots)) + \mathbb{P}((0,1,0,0,\dots)) + \mathbb{P}((0,0,1,0,\dots)) + \dots$$

$$= p + p + p + \dots$$

$$\Rightarrow p = 0.$$

$\square$

Bedeutet das, dass $\mathbb{P}(E) = 0 \quad \forall E \in \Omega$?

NEIN!

$\mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\omega)$ passt nur, wenn E höchstens abzählbar ist.

Unterscheid zwischen diskreten und überabzählbaren W-räumen

Im diskreten Fall genügt es, die W-keiten von den Elementarereignissen zu definieren.

Im überabzählbaren Fall ist das nicht genug, weil es überabzählbare Ereignisse gibt. In diesem Fall müssen wir die W-keiten von allen Ereignissen $E \subseteq \Omega$ definieren.

$\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$

Bemerkung

Die Potenzmenge $\mathcal{P}(\Omega) = \{E \subseteq \Omega\}$

Aber die Abbildung $\mathbb{P}$ ist keine beliebige Abbildung

Zwei Eigenschaften, die $\mathbb{P}$ haben soll

(E1) $\quad \mathbb{P}(\Omega) = 1$

(E2) $\quad$ Seien $E_1, E_2, E_3, \dots \subseteq \Omega$ eine abzählbare Folge von paarweise disjunkten Teilmengen.

$(E_i \cap E_j = \emptyset \quad , \quad \text{wenn } i \neq j)$

$$\mathbb{P}\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i) \qquad \sim \text{Summenregel}$$

(E3) Sei $\ddot{A}_i : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$ die Abbildung, die die i^{te} Koordinate verändert.

(Das heißt, $\ddot{A}_i(\omega_1, \omega_2, \dots, \omega_{i-1}, \omega_i, \omega_{i+1}, \dots)$

$= (\omega_1, \omega_2, \dots, \omega_{i-1}, 1-\omega_i, \omega_{i+1}, \dots)$.)

Für jede Teilmenge $E \subseteq \{0,1\}^{\mathbb{N}}$ und jedes $i \in \mathbb{N}$,

$$\mathbb{P}(\ddot{A}_i(E)) = \mathbb{P}(E).$$

$(\ddot{A}_i(E) = \{\ddot{A}_i(\omega) : \omega \in E\}.)$

<u>Definition</u>

Sei $S \subset \mathbb{N}$ eine endliche Teilmenge, $S = \{s_1, s_2, \dots, s_n\}$.

Dann

$$\ddot{A}_S(\omega) = \ddot{A}_{s_1}(\ddot{A}_{s_2}(\ddot{A}_{s_3}(\dots(\ddot{A}_{s_n}(\omega))\dots)))$$

verändert alle Koordinaten in S.

$\forall E \subseteq \Omega$, $S \subset \mathbb{N}$, $|S| < \infty$,

$$\mathbb{P}(\ddot{A}_S(E)) = \mathbb{P}(E) .$$

<u>Satz</u> (Vitali, 1905)

Sei $\Omega = \{0,1\}^{\mathbb{N}}$. Es gibt <u>KEINE</u> Abbildung $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ mit den Eigenschaften (E1), (E2) und (E3)

Beweis folgt nach Einführung notwendiger Definitionen.

<u>Definition</u>

Zwei Folgen $\omega, \omega' \in \Omega$ sind äquivalent $(\omega \sim \omega')$, wenn

$$|\{i \in \mathbb{N} : \omega_i \neq \omega_i'\}| < \infty .$$

(Es gibt nur endlich viele Unterschiede zwischen ω und ω'.)

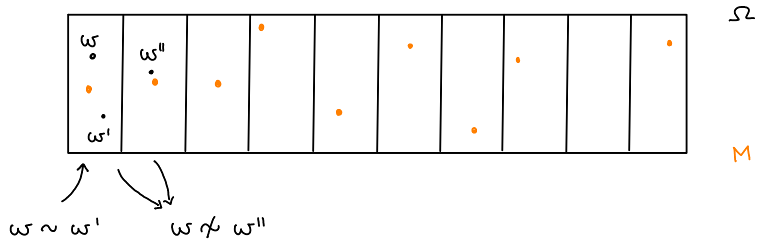
<u>Beispiel</u>

$(0,0,0,\dots) \sim (1,0,0,0,\dots)$

$\sim (\omega_1, \omega_2, \dots)$, wobei nur endlich viele $\omega_i = 1$.

$(0,0,0,\dots) \not\sim (0,1,0,1,0,1,\dots)$

$\not\sim (1,1,1,1,\dots)$



Definition

Sei $\omega \in \Omega$. Dann ist $[\omega]_\sim = \{\omega' \in \Omega : \omega \sim \omega'\}$

die <u>Äquivalenzklasse</u> von ω.

Wir erzeugen eine „verrückte" Menge M:

Wir wählen von jeder Äquivalenzklasse <u>genau eine</u> beliebige Folge

aus und tun sie in M hinein.

$\Rightarrow \ \forall \omega \in \Omega, \quad |M \cap [\omega]_\sim| = 1.$

Nehmen wir an, dass eine Abbildung $\mathbb{P}$ mit den Eigenschaften

(E1), (E2) und (E3) existiert.

Was ist $\mathbb{P}(M)$?

Behauptung

$$\Omega = \dot{\bigcup_{\substack{S \subset \mathbb{N} \\ |S| < \infty}}} \ddot{A}_S(M)$$

Beweis

(i) Sei $\omega \in \Omega$.

Dann $|M \cap [\omega]_\sim| = 1$.

$\Rightarrow \exists \omega' \in (M \cap [\omega]_\sim) \iff \exists \omega' \in M : \omega \sim \omega'$.

$\Rightarrow S = \{ i \in \mathbb{N} : \omega_i \neq \omega'_i \}$ ist endlich.

$$\Rightarrow \quad \omega = \ddot{A}_s(\omega').$$

$$\Rightarrow \quad \omega \in \ddot{A}_s(M)$$

$$\Rightarrow \quad \omega \in \bigcup_{|s|<\infty} \ddot{A}_s(M)$$

Deshalb

$$\bigcup_{|s|<\infty} \ddot{A}_s(M) = \Omega. \qquad \checkmark$$

(ii) $S \neq S'$

$$\Rightarrow \quad \ddot{A}_s(M) \cap \ddot{A}_{s'}(M) = \emptyset. \qquad \checkmark$$

(Beweis siehe Notizen.) $\square$

<u>Wiederspruch</u>

abzählbare Vereinigung

$$1 \overset{(E1)}{=} \mathbb{P}(\Omega) = \mathbb{P}\Big(\dot{\bigcup_{|s|<\infty}} \ddot{A}_s(M)\Big)$$

$$\overset{(E2)}{=} \sum_{|s|<\infty} \mathbb{P}(\ddot{A}_s(M))$$

$$\overset{(E3)}{=} \sum_{|s|<\infty} \mathbb{P}(M)$$

$\lightning$ $\mathbb{P}(M)$ ist nicht wohldefiniert.

$\square$

Wiederholung

Satz (Vitali, 1905)

Sei $\Omega = \{0,1\}^{\mathbb{N}}$. Es gibt KEINE Abbildung $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$

mit den Eigenschaften

(E1) $\mathbb{P}(\Omega) = 1$

(E2) $\forall E_1, E_2, E_3, \ldots$ die paarweise disjunkt sind, gilt

$$\mathbb{P}\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$$

(E3) Sei $\ddot{A}_i : \Omega \to \Omega$ die Abbildung, die die i^{te} Koordinate verändert. Dann gilt für jede Teilmenge $E \subseteq \Omega$ und jede Koordinate $i \in \mathbb{N}$,

$$\mathbb{P}(\ddot{A}_i(E)) = \mathbb{P}(E) .$$

Satz (von Vitali auf Intervallen)

Seien $c, d \in \mathbb{R}$, sodass $c < d$ und sei $\Omega = (c,d]$.

Es gibt keine Abbildung $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ mit den drei Eigenschaften:

(E1) $\mathbb{P}(\Omega) = 1$

(E2) $\forall E_1, E_2, E_3, \ldots \subseteq \Omega$, die paarweise disjunkt sind, gilt

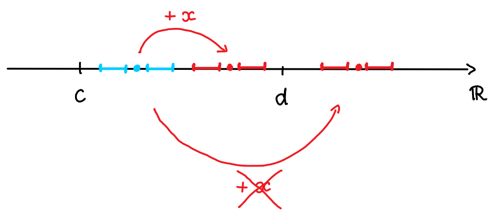
$$\mathbb{P}\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$$

(E3) Für jede Teilmenge $E \subseteq (c,d]$ und jedes $x \in \mathbb{R}$, sodass

$$x + E = \{x + e : e \in E\} \subseteq (c,d]$$

gilt,

$$\mathbb{P}(x + E) = \mathbb{P}(E)$$



σ - Algebren

die Lösung

Wir werden nur für „schöne" Teilmengen von Ω eine W-keit definieren.

Solche schönen Teilmengen werden „Ereignisse" genannt.

Welche Teilmengen von Ω sind uns wichtig?

Beispiel

Wie lange würde diese Vorlesung dauern?

$\Omega = (0, 120]$

Sinnvolle Fragen

- Dauert die Vorlesung länger als 90 min?

  $E = (90, 120]$

- Wie wahrscheinlich ist es, dass die Vorlesung zwischen 85 und 95 min dauert?

  $E = (85, 95]$

$\Rightarrow$ Die wichtigen Teilmengen sind Intervalle $(a,b] \subseteq \Omega$

Aber (leider) sind Intervalle nicht genug:

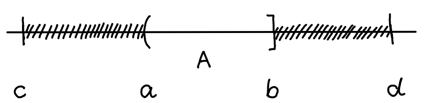
Wir möchten auch ein paar logische Operationen machen können

Seien $A, B \subseteq \Omega$ zwei Ereignisse.

Wie wahrscheinlich ist es, dass ...

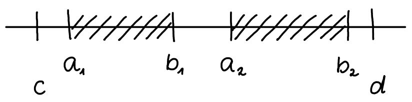
- A <u>nicht</u> auftritt?

  $\Rightarrow$ wir brauchen, dass das Komplement $\Omega \setminus A$ auch ein Ereignis ist.



- dass entweder A oder B auftritt?

  $\Rightarrow$ Die Vereinigung $A \cup B$ soll auch ein Ereignis sein.



- dass A und B auftreten?

  $\Rightarrow$ $A \cap B$ soll auch ein Ereignis sein.

Definition (σ-Algebra)

Sei Ω eine Menge. Eine σ-Algebra auf Ω ist eine Menge

von Teilmengen von Ω.

$\mathcal{E} \subseteq \mathcal{P}(\Omega)$, sodass

(S1) $\Omega \in \mathcal{E}$.

(S2) $\forall E \in \mathcal{E}$, gilt $\Omega \setminus E \in \mathcal{E}$

(S3) $\forall E_1, E_2, E_3, \ldots \in \mathcal{E}$, gilt $\bigcup_{i=1}^{\infty} E_i \in \mathcal{E}$.

Bemerkung

(i) In (S3) erlauben wir abzählbare Vereinigungen.

(ii) Wir können Durchschnitt durch Komplement und

Vereinigung erzeugen:

z.B. $A \cap B = \Omega \setminus ((\Omega \setminus A) \cup (\Omega \setminus B))$

Beispiele

Behauptung

Sei Ω eine beliebige Menge. Die folgenden Teilmengen

von $\mathcal{P}(\Omega)$ sind σ-Algebren auf Ω.

(i) $\mathcal{E}_1 = \{\emptyset, \Omega\}$

(ii) $\mathcal{E}_2 = \mathcal{P}(\Omega)$

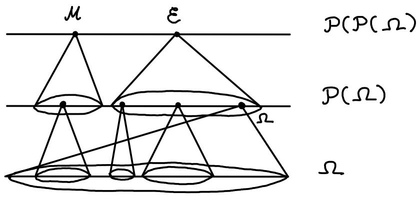
(iii) $\mathcal{E}_3 = \{E \subseteq \Omega \mid$ entweder E oder $\Omega \setminus E$ ist höchstens

abzählbar$\}$

Beweis

H.A. $\square$

Drei Ebenen von Mengen



Erzeugte σ-Algebra

Behauptung

Sei Ω eine Menge und sei $\mathcal{M} \subseteq \mathcal{P}(\Omega)$ eine Menge von Teilmengen von Ω.

Es gibt genau eine Menge von Teilmengen von Ω, $\sigma(\mathcal{M}) \subseteq \mathcal{P}(\Omega)$, sodass

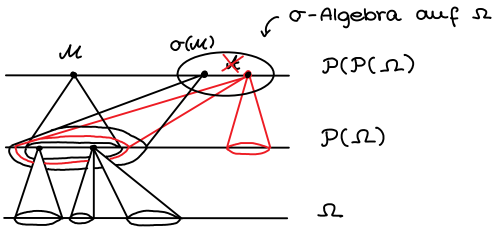
(ES1) $\mathcal{M} \subseteq \sigma(\mathcal{M})$

(ES2) $\sigma(\mathcal{M})$ ist eine σ-Algebra auf Ω

(ES3) Sei $\mathcal{A}$ eine σ-Algebra auf Ω sodass $\mathcal{M} \subseteq \mathcal{A}$.

Dann gilt $\sigma(\mathcal{M}) \subseteq \mathcal{A}$.

$\sigma(\mathcal{M})$ ist die <u>von $\mathcal{M}$ erzeugte σ-Algebra</u>.



Bemerkung

Die Eigenschaft (ES3) bedeutet, dass $\sigma(\mathcal{M})$ die <u>kleinste</u> σ-Algebra auf Ω, die alle Mengen in $\mathcal{M}$ enthält, ist.

Beweis

H.A. $\square$

Wenn $\mathcal{M}$ die Menge von Teilmengen von Ω ist, die alle Mengen, für die wir eine W-keit definieren möchten, enthält, ist $\sigma(\mathcal{M})$ die Menge aller Mengen, für die wir eine W-keit definieren müssen.

Definition

Sei $\Omega = \mathbb{R}$, und sei $\mathcal{M} = \{ (a,b] : a \le b, \; a, b \in \mathbb{R} \}$.

Die Borel'sche σ-Algebra auf $\mathbb{R}$, $\mathcal{B}$, ist die von M erzeugte σ-Algebra:

$$\mathcal{B} = \sigma(\mathcal{M}).$$

Eine Menge $B \in \mathcal{B}$ ist eine Borel-Menge.

Allgemeine W-räume

Definition (W-maß)

Sei Ω eine Menge und sei $\mathcal{E} \subseteq \mathcal{P}(\mathcal{M})$ eine σ-Algebra auf Ω.

Eine Abbildung $\mathbb{P} : \mathcal{E} \to [0,1]$ ist ein W-maß (oder eine Verteilung), wenn

(WM1) $\mathbb{P}(\Omega) = 1$.

(WM2) Für paarweise disjunkte Mengen $E_1, E_2, E_3, \ldots \in \mathcal{E}$, gilt,

$$\mathbb{P}\left(\dot{\bigcup_{i=1}^{\infty}} E_i \right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$$

[(WM2) heißt σ-Additivität]

Definition (W-raum)

Ein W-raum ist ein Tripel $(\Omega, \mathcal{E}, \mathbb{P})$, wobei:

(W1) Ω ist eine beliebige Menge

(W2) $\mathcal{E}$ ist eine σ-Algebra auf Ω.

(W3) $\mathbb{P}$ ist ein W-maß auf $(\Omega, \mathcal{E})$.

Wiederholung

Definition (σ-Algebra)

Sei Ω eine Menge. Eine σ-Algebra ist eine Menge von Teilmengen von Ω, $\mathcal{E} \subseteq \mathcal{P}(\Omega)$, sodass

(S1) $\Omega \in \mathcal{E}$.

(S2) $\forall E \in \mathcal{E}$, $\Omega \setminus E \in \mathcal{E}$.

(S3) $\forall E_1, E_2, E_3, \ldots \in \mathcal{E}$, $\bigcup_{i=1}^{\infty} E_i \in \mathcal{E}$

Definition (Borelsche σ-Algebra)

Sei $\Omega = \mathbb{R}$ und sei $M = \{ (a,b] : a, b \in \mathbb{R}, a \leq b \}$.

Die Borel'sche σ-Algebra auf $\mathbb{R}$, $\mathcal{B} = \sigma(\mathbb{R})$, ist die von $\mathcal{M}$ erzeugte σ-Algebra.

Eine Menge $B \in \mathcal{B}$ wird eine Borel-Menge genannt.

Definition (W-maß, W-raum)

Sei Ω eine Menge und sei $\mathcal{E}$ eine σ-Algebra auf Ω.

Eine Abbildung $\mathbb{P} : \mathcal{E} \to [0,1]$ ist ein W-maß (oder eine Verteilung), wenn:

(WM1) $\mathbb{P}(\Omega) = 1$

(WM2) $\forall E_1, E_2, E_3, \ldots \in \mathcal{E}$, die paarweise disjunkt sind,

gilt $\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i \right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$

Das Tripel $(\Omega, \mathcal{E}, \mathbb{P})$ ist ein W-raum

---

Beispiele

Beispiel 0

Sei $(\Omega, \mathbb{P})$ ein diskreter W-raum.

Dann ist $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ auch ein W-raum.

- (W1): Ω ist eine beliebige Menge. ✓
- (W2): $\mathcal{P}(\Omega)$ ist eine σ-Algebra auf Ω. ✓
- (W3): $\mathbb{P}$ ist ein W-maß auf $(\Omega, \mathcal{P}(\Omega))$. ✓

Beispiel 1 (Die Dirac-Verteilung)

Sei Ω eine beliebige Menge und sei $\mathcal{E}$ eine σ-Algebra auf Ω.

Sei Dirac-Verteilung ist das W-maß $\delta_x : \mathcal{E} \to [0,1]$,

wobei,

$$\delta_x(E) = \begin{cases} 0 & , \text{ wenn } x \notin E \\ 1 & , \text{ wenn } x \in E \end{cases}$$

$(\Omega, \mathcal{E}, \delta_x)$ ist ein W-raum.

- (W1): ✓
- (W2): ✓
- (W3): δ_x ist ein W-maß:

(WM1): Weil $x \in \Omega$, $\delta_x(\Omega) = 1$. ✓

(WM2): Seien $E_1, E_2, E_3, \dots$ paarweise disjunkte Mengen

1. Fall: $x \notin \dot{\bigcup}_{i=1}^{\infty} E_i$

$\Rightarrow x \notin E_i \quad \forall i \in \mathbb{N}$

$$\Rightarrow \delta_x\left(\dot{\bigcup}_{i=1}^{\infty} E_i\right) = 0 = \sum_{i=1}^{\infty} 0 = \sum_{i=1}^{\infty} \delta_x(E_i) \quad ✓$$

2. Fall: $x \in \dot{\bigcup}_{i=1}^{\infty} E_i$

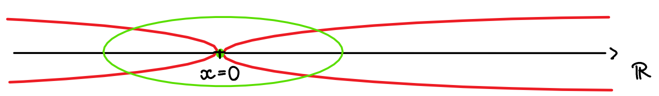
$\Rightarrow$ es gibt genau ein $k = \mathbb{N}$, sodass $x \in E_k$. (die E_i sind paarweise disjunkt.)

$$\delta_x\left(\dot{\bigcup}_{i=1}^{\infty} E_i\right) = 1$$

$$\sum_{i=1}^{\infty} \delta_x(E_i) = \delta_x(E_k) + \sum_{i \neq k} \delta_x(E_i) = 1 + \sum_{i \neq k} 0 = 1. \quad ✓$$

Beispiel von dem Beispiel

$\Omega = \mathbb{R}$, $\mathcal{E} = \mathcal{P}(\mathbb{R})$, $x = 0$.



$\delta_x(E) = 1$. (egal wie klein, solange $x \in E$)

$\delta_x(E) = 0$. (—"— groß, —"— $x \notin E$)

Frage

Wie können wir sinnvolle Verteilung auf $(\mathbb{R}, \mathcal{B})$ erzeugen?

Wir müssen zu jeder Borel-Menge $B \in \mathcal{B}$ eine W-keit zuordnen.

Problem: Es gibt viele Borel-Mengen.

Definition (Maß)

Sei Ω eine Menge und sei $\mathcal{E}$ eine σ-Algebra auf Ω.

Ein Maß ist eine σ-additive Abbildung $\mu : \mathcal{E} \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$.

Das heißt,

$\forall E_1, E_2, E_3, \ldots \in \mathcal{E}$, die paarweise disjunk sind,

$$\mu\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i).$$

Satz (Erweiterung von Carathéodory)

Sei $\mathcal{M} = \{(a,b], \; a,b \in \mathbb{R}, a \leq b\}$, und sei $\mu^* \cdot \mathcal{M} \to \mathbb{R}_{\geq 0}$

eine Abbildung, sodass wenn $E_1, E_2, E_3, \ldots \in \mathcal{B}$ paarweise disjunkt sind und $\dot{\bigcup_{i=1}^{\infty}} E_i \in \mathcal{M}$, gilt

$$\mu^*\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) = \sum_{i=1}^{\infty} \mu^*(E_i).$$

Dann gibt es genau ein Maß $\mu \cdot \mathcal{B} \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$, sodass

für alle $(a,b] \in \mathcal{M}$,

$$\mu((a,b]) = \mu^*((a,b])$$

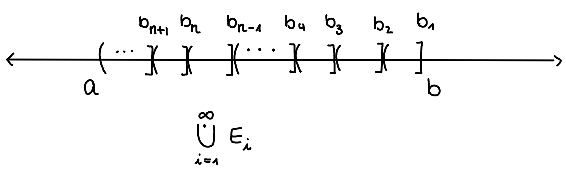
(z.B. es ist möglich, dass $\mu(\mathbb{R}) = \infty$.)

Beispiel (Lebesque-Maß)

Wir definieren $\mu^* : \mathcal{M} \to \mathbb{R}_{\geq 0}$:

$$\mu^*((a,b]) = b - a \geq 0.$$

Seien $E_1, E_2, E_3, \ldots \in \mathcal{M}$ eine paarweise disjunkte Folge, sodass $\dot{\bigcup_{i=1}^{\infty}} E_i = (a,b]$.



Wir brauchen, dass

$$\dot{\bigcup_{i=1}^{\infty}} E_i = (a,b]$$

und o.B.d.A.

$$E_i = (b_{i+1}, b_i]$$

wobei $b_1 \geq b_2 \geq b_3 \geq b_4 \geq \dots$, $b_1 = b$ und $\lim_{n\to\infty} b_n = a$.

zum Beispiel

wenn $\dot{\bigcup_{i=1}^{\infty}} E_i = (0,1]$ ist es möglich, dass

(i) $E_1 = (\frac{1}{2}, 1]$, $E_2 = (0, \frac{1}{2}]$, $E_3 = E_4 = E_5 = \dots = \emptyset = (0,0]$.

(ii) $b_n = \frac{1}{n}$

$E_1 = (\frac{1}{2}, 1)$, $E_2 = (\frac{1}{3}, \frac{1}{2})$, $E_3 = (\frac{1}{4}, \frac{1}{3})$, ...

$$\mu^*\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) = \mu^*((a,b])$$

$$= b-a$$

$$\sum_{i=1}^{\infty} \mu^*(E_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu^*(E_i)$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \mu^*((b_{i+1}, b_i])$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} (b_{i+1}, b_i]$$

$$= \lim_{n\to\infty} \left((b_1 - b_2) + (b_2 - b_3) + (b_3 - b_4) + \dots + (b_{n-1} - b_n) + (b_n - b_{n+1}) \right)$$

$$= \lim_{n\to\infty} (b_1 - b_{n+1})$$

$$= b_1 - \lim_{n\to\infty} b_{n+1}$$

$$= b - a = \mu^*\left(\dot{\bigcup_{i=1}^{\infty}} E_i\right) \quad \checkmark$$

Also können wir den Erweiterungssatz von Carathéodory anwenden, um ein Maß $\lambda : \mathcal{B} \to \mathbb{R}_{\geq 0} \cup \{+\infty\}$ zu erhalten.

λ wird das Lebesgue-Maß auf $\mathbb{R}$ genannt.

Eigenschaften

- $\lambda((a,b]) = \mu^*((a,b])$
  $= b-a$.
- $\lambda(\mathbb{R}) = \infty$ ⟵ Deshalb ist λ kein W-maß. Aber wir können λ benutzen, um ein W-maß zu erzeugen.

  (weil $\lambda(\mathbb{R}) \geq \lambda((0,n]) = n \quad \forall n \in \mathbb{N}$)

Beispiel

Sei $\Omega = (c,d]$, $c, d \in \mathbb{R}$, $c < d$.

Sei $\mathcal{E}$ die Borel'sche σ-Algebra auf Ω.

Das heißt $\mathcal{E} = \{ B \cap \Omega : B \in \mathcal{B} \}$

$= \{ B \in \mathcal{B} : B \subseteq \Omega \}$.

Wir definieren ein W-maß $\mu : \mathcal{E} \to [0,1]$, wobei

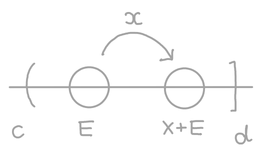
$$\mu(E) = \frac{1}{d-c} \lambda(E)$$

Dann ist $((c,d], \mathcal{E}, \mu)$ ein W-raum, sodass :

① $\mu((c,d]) = 1$

② μ ist σ-additiv

③ $\forall E \in \mathcal{E}$ und $x \in \mathbb{R}$, sodass $\underbrace{x + E}_{= \{x+e \,:\, e \in \Omega\}} \subseteq \Omega$, $\mu(x+E) = \mu(E)$



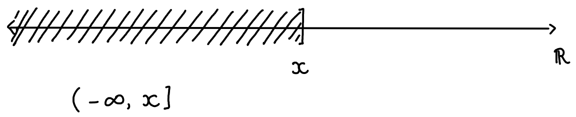
Verteilungsfunktion

Definition

Sei $\mathbb{P}$ eine Verteilung auf $(\mathbb{R}, \mathcal{B})$. Die <u>Verteilungsfunktion</u> $F_{\mathbb{P}}$

ist die Abbildung $F_{\mathbb{P}} : \mathbb{R} \to [0,1]$:

$$F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x])$$

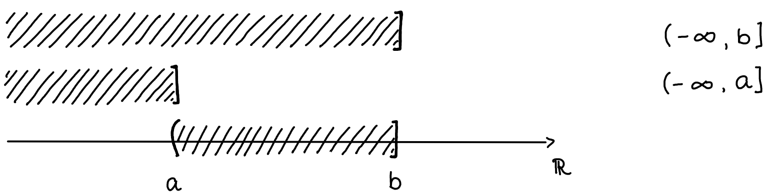


Bemerkung

Die Verteilung $\mathbb{P}$ wird durch ihre Verteilungsfunktion bestimmt.

Sei $(a,b]$ ein Intervall . Dann,

$$\mathbb{P}((a,b]) = \mathbb{P}((-\infty, b] \setminus (-\infty, a])$$



$$\mathbb{P}((a,b]) = \mathbb{P}((-\infty,b] \setminus (-\infty,a])$$

$$= \mathbb{P}((-\infty,b]) - \mathbb{P}((-\infty,a])$$

$$= F_{\mathbb{P}}(b) - F_{\mathbb{P}}(a)$$

Deshalb :

Verteilungsfunktion $F_{\mathbb{P}}$

$\Downarrow$

W-keit von allen Intervallen $(a,b]$

$\Downarrow$ (Erweiterung von Carathéodory)

die Verteilung $\mathbb{P}$.

Wiederholung

Definition (Verteilungsfunktion)

Sei $\mathbb{P}$ eine Verteilung auf $(\mathbb{R}, \mathcal{B})$. Ihre Verteilungsfunktion

ist die Abbildung

$$F_{\mathbb{P}} : \mathbb{R} \to [0,1], \quad \text{wobei} \quad F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x]).$$

Bemerkung

Die Verteilung $\mathbb{P}$ wird durch ihre Verteilungsfunktion

$F_{\mathbb{P}}$ bestimmt.

Beispiele von Verteilungsfunktionen

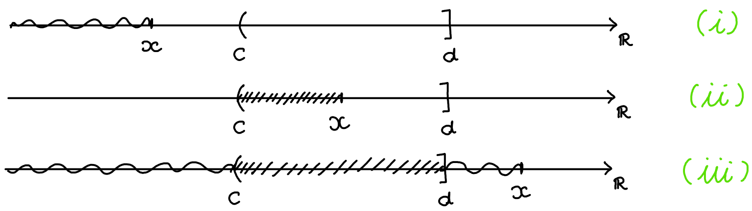
Beispiel ①: Die Gleichverteilung auf $(c,d]$.

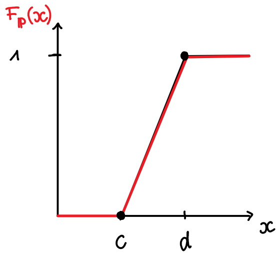
$$\forall E \in \mathcal{B}_{(c,d]}, \quad \mathbb{P}(E) = \frac{\lambda(E)}{d-c}.$$

Insbesondere, wenn $E = (a,b] \subseteq (c,d]$, $\mathbb{P}((a,b]) = \frac{b-a}{d-c}$

$$F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x])$$

$$F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x]) = \mathbb{P}((-\infty, x] \cap (c,d]) = \begin{cases} \mathbb{P}(\emptyset) = 0 & \text{, wenn } x \leq c \quad (i) \\ \mathbb{P}((c,x]) = \frac{x-c}{d-c} & \text{, wenn } c \leq x \leq d \quad (ii) \\ \mathbb{P}((c,d]) = 1 & \text{, wenn } x \geq d \quad (iii) \end{cases}$$





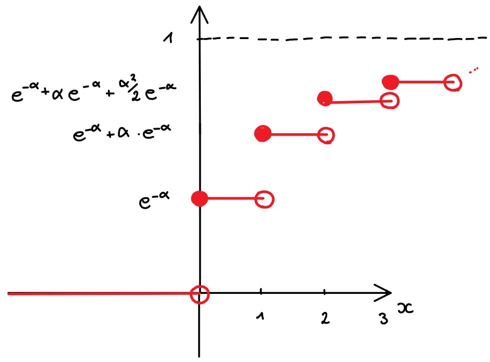
Beispiel ②: Eine diskrete Verteilung

z.B. die Poisson-Verteilung zum Parameter α.

$$\mathbb{P}(\{x\}) = \begin{cases} \frac{\alpha^x e^{-\alpha}}{x!} & \text{, wenn } x \in \mathbb{Z}_{\geq 0} \\ 0 & \text{, sonst} \end{cases}$$

$$F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x]) = \mathbb{P}((-\infty, \lfloor x \rfloor])$$

wobei $\lfloor x \rfloor$ die größte ganze Zahl, die $\leq x$ ist, ist.



Beispiel ③: Wann tritt der Schaden auf?

Wie lange müssen wir warten auf den ersten Schaden?

↳ Es gibt eine Verteilung $\mathbb{P}$ auf $(\mathbb{R}, \mathcal{B})$.

Es gibt eine Verteilungsfunktion:

$$F_{\mathbb{P}}(t) = \mathbb{P}((-\infty, t])$$

Wenn die Wartezeit kleiner gleich t ist, bedeutet das, dass mindestens ein Schadensfall in dem Intervall $[0,t]$ auftritt.

Die Anzahl der Schadensfälle im Intervall $[0,t]$ wird mit der Poisson-Verteilung zum Parameter αt modelliert.

Deshalb: $(t \geq 0)$

$$\begin{aligned} F_{\mathbb{P}}(t) &= \mathbb{P}((-\infty, t]) \\ &= \mathbb{P}(Z \geq 1) \quad , \text{wobei } Z \sim \mathrm{Poi}(\alpha t) \\ &= 1 - \mathbb{P}(Z = 0) \\ &= 1 - e^{-\alpha t}. \end{aligned}$$

Alternative Schreibweise:

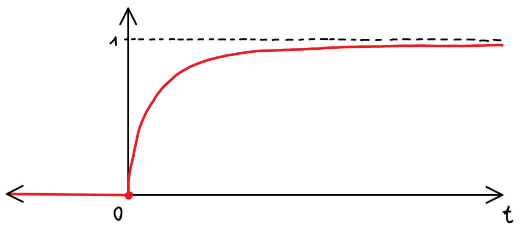
$$= \sum_{k=1}^{\infty} p_{\alpha t}(k)$$

$$= 1 - p_{\alpha t}(0)$$

$$= 1 - e^{-\alpha t}$$

Also,

$$F_{\mathbb{P}}(t) = \begin{cases} 1 - e^{-\alpha t} & , \text{wenn } t \geq 0 \\ 0 & , \text{sonst.} \end{cases}$$



Allgemeine Eigenschaften

Behauptung (Eigenschaften von W-maßen)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum.

Dann gilt:

(a) $\mathbb{P}(\emptyset) = 0$

(b) $\forall E, F \in \mathcal{E}$, wobei $E \subseteq F$, gilt

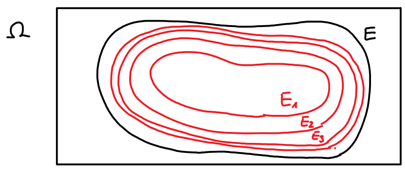
$$\mathbb{P}(E) \leq \mathbb{P}(F).$$ [Monotonität]

(c) $\forall E, F \in \mathcal{E}$ gilt

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$$

(d) $\forall E_1, E_2, E_3, \ldots \in \mathcal{E}$, wobei $E_1 \subseteq E_2 \subseteq E_3 \subseteq E_4 \subseteq \ldots$, gilt

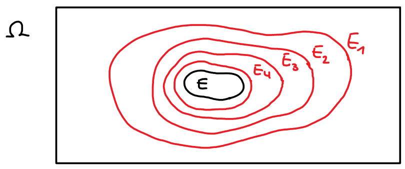
$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \lim_{n \to \infty} \mathbb{P}(E_n)$$ [Stetigkeit von unten]



$$\bigcup_{i=1}^{n} E_i = E_n$$

(e) $\forall E_1, E_2, E_3, \ldots \in \mathcal{E}$, wobei $E_1 \supseteq E_2 \supseteq E_3 \supseteq E_4 \supseteq \ldots$, gilt

$$\mathbb{P}\Big(\bigcap_{i=1}^{\infty} E_i\Big) = \lim_{n \to \infty} \mathbb{P}(E_n)$$



$$\bigcup_{i=1}^{n} E_i = E_n$$

Beweis: Hausaufgabe.

Behauptung (Eigenschaften von Verteilungsfunktionen)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ eine Verteilung, und sei $F_{\mathbb{P}}$ ihre Verteilungsfunktion.

Dann gilt:

(VF1) $F_{\mathbb{P}}$ steigt monoton

(VF2) $\lim_{x \to -\infty} F_{\mathbb{P}}(x) = 0$ und

$\lim_{x \to \infty} F_{\mathbb{P}}(x) = 1$.

(VF3) $F_{\mathbb{P}}$ ist rechtstetig. Das heißt:

$\forall c \in \mathbb{R}: \lim_{\varepsilon \searrow 0} F_{\mathbb{P}}(c+\varepsilon) = F_{\mathbb{P}}(c)$.

Beweis

(VF1) Seien $x, y \in \mathbb{R}$, $x < y$. Dann gilt

$$(-\infty, x] \subseteq (-\infty, y]$$

Die Monotonität von $\mathbb{P}$ $\Rightarrow$ $F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x])$

$$\leq \mathbb{P}((-\infty, y])$$

$$= F_{\mathbb{P}}(y). \quad ✓$$

(VF2) Sei $(x_n)_{n \geq 1}$ eine Folge, wobei $x_n \to -\infty$

Wir definieren:

$E_n = (-\infty, x_n]$.

$E_1 \supseteq E_2 \supseteq E_3 \supseteq \dots$ und $\bigcap_{i=1}^{\infty} E_i = \emptyset$.

Deshalb: $0 = \mathbb{P}(\emptyset) = \lim_{n \to \infty} \mathbb{P}(E_n) = \lim_{n \to \infty} F_{\mathbb{P}}(x_n)$ ✓

(Hier haben wir die Stetigkeit von oben von $\mathbb{P}$ angewendet)

Also, $\lim_{x \to -\infty} F_{\mathbb{P}}(x) = 0$.

$\lim_{x \to \infty} F_{\mathbb{P}}(x) = 1$ $\rightsquigarrow$ sehr ähnlich. ✓

(VF3) Sei $(\varepsilon_n)_{n \geq 1}$ eine Folge, wobei $\varepsilon_n \searrow 0$.

Definieren: $E_n = (-\infty, c + \varepsilon_n]$.

Stetigkeit von oben:

$$
\begin{aligned}
F_{\mathbb{P}}(c) &= \mathbb{P}((-\infty, c]) \\
&= \mathbb{P}\Big(\bigcap_{i=1}^{\infty} (-\infty, c+\varepsilon_i]\Big) \\
&= \lim_{n\to\infty} \mathbb{P}((-\infty, c+\varepsilon_n]) \\
&= \lim_{n\to\infty} F_{\mathbb{P}}(c+\varepsilon_n) \quad \checkmark \qquad \square
\end{aligned}
$$

Bemerkung

Sei $F: \mathbb{R} \to [0,1]$ eine Abbildung mit den Eigenschaften (VF1), (VF2) und (VF3). Dann gibt es eine Verteilung $(\mathbb{R}, \mathcal{B}, \mathbb{P})$, sodass $F_{\mathbb{P}} = F$.

Frage

Wenn wir nur die Verteilungsfunktion F haben, wie können wir die W-keit von einer beliebigen Borel-Menge berechnen?

→ nächsten Donnerstag: Dichtefunktion

Wiederholung

Definition (Verteilungsfunktion)

Sei $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ eine Verteilung. Ihre Verteilungsfkt. ist die Abbildung

$F_{\mathbb{P}} : \mathbb{R} \to [0,1]$, wobei $F_{\mathbb{P}}(x) = \mathbb{P}((-\infty, x])$.

Behauptung

Sei $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ eine Verteilung und sei $F_{\mathbb{P}}$ ihre Verteilungsfkt. Dann gilt:

(VF1) $F_{\mathbb{P}}$ steigt monoton.

(VF2) $\lim\limits_{x \to -\infty} F_{\mathbb{P}}(x) = 0$ und $\lim\limits_{x \to \infty} F_{\mathbb{P}}(x) = 1$

(VF3) $F_{\mathbb{P}}$ ist rechtsstetig. Das heißt, $\forall c \in \mathbb{R}$,

$$\lim_{\varepsilon \searrow 0} F_{\mathbb{P}}(c + \varepsilon) = F_{\mathbb{P}}(c)$$

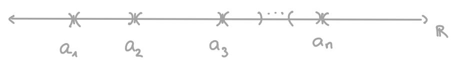
Frage

Wie können wir die W-keiten von beliebigen Borel-Mengen berechnen?

Dichtefunktion

Nehmen Sie an, dass die Verteilungsfkt. $F(x)$ stückweise stetig differenzierbar ist.

stückweise stetig:



$\Rightarrow$ Es gibt eine stückweise stetige Abbildung $F'(x)$.

$\Rightarrow$ $\mathbb{P}((a,b]) = F(b) - F(a)$

$$= \int_{(a,b]} F'(x)\, dx$$

Für eine beliebige Borel-Menge $E \in \mathcal{B}$, ist es sinnvoll zu schreiben

$$\mathbb{P}(E) = \int_E F'(x)\, dx \qquad ? \qquad \text{Ja!}$$

Diese Formel ist sinnvoll, wenn wir das Integral als das Lebesgue-Integral vorstellen.

Am wichtigsten: Wenn eine Funktion f Riemann-integrierbar ist, sind das Lebesque-Integral und das Riemann-Integral gleich.

Definition (Dichtefunktion)

Sei $\rho : \mathbb{R} \to [0,\infty)$ eine Funktion mit den Eigenschaften:

$\rho^{-1}((-\infty,c]) = \rho^{-1}((-\infty,0)) \cup \rho^{-1}([0,c])$ $(c \geq 0)$

$\rightsquigarrow$ $\rho^{-1}((-\infty,0)) = \emptyset$

(DF1) $\forall c \in \mathbb{R}$, $\rho^{-1}((-\infty,c]) \in \mathcal{B}$.

(DF2) $\int_{\mathbb{R}} \rho(x)\,dx = 1$.

(DF3) $\mathbb{P} : \mathcal{B} \to [0,1]$, die durch

$$\mathbb{P}(E) = \int_E \rho(x)\,dx = \int_{\mathbb{R}} \mathbb{1}_E(x)\,\rho(x)\,dx$$

definiert wird, ein W-maß auf $(\mathbb{R}, \mathcal{B})$.

ρ heißt die Dichtefunktion von $\mathbb{P}$.

(und ρ ist die Abbildung von der Verteilungsfunktion $F_{\mathbb{P}}$)

Bemerkung

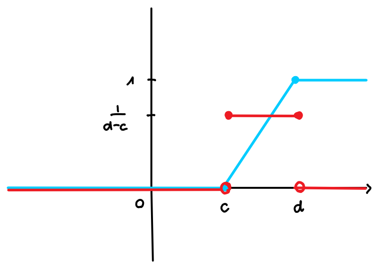
Alle stückweise stetige Funktionen haben die Eigenschaft (DF1).

Beispiel ① (Gleichverteilung auf $(c,d]$)

Verteilungsfunktion:

$$F_{\mathbb{P}}(x) = \begin{cases} 0 & , \text{ wenn } x \leq c \\ \frac{x-c}{d-c} & , \text{ wenn } c \leq x \leq d \\ 1 & , \text{ wenn } x \geq d \end{cases}$$

Dichtefunktion:



$$\rho(x) = F_{\mathbb{P}}'(x) = \begin{cases} \frac{1}{d-c} & , \text{ wenn } c \leq x \leq d \\ 0 & , \text{ sonst} \end{cases}$$

$\mathbb{P}_G(E) = \frac{\lambda(E \cap G)}{\lambda(G)}$: Gleichverteilung auf G wobei $\lambda(G) < \infty$.

Bemerkung

Wir können die Gleichverteilung verallgemeinern:

Sei $D \subseteq \mathbb{R}^n$ eine begrenzte Menge. Die Gleichverteilung auf D hat die Dichtefunktion:

$$\rho_D(x) = \begin{cases} \frac{1}{|D|} & , \text{ wenn } x \in D \\ 0 & , \text{ sonst.} \end{cases}$$

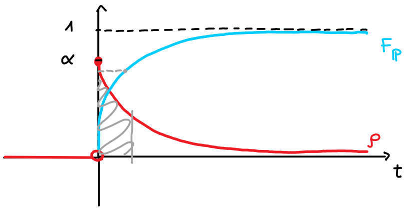
Beispiel ② (Wartezeit)

Verteilungsfunktion:

$$F_{\mathbb{P}}(t) = \begin{cases} 0 & , \text{wenn } t<0 \\ 1-e^{-\alpha t} & , \text{wenn } t \geq 0 . \end{cases}$$

Dichtefunktion:

$$\rho(t) = F_{\mathbb{P}}'(t) = \begin{cases} 0 & , \text{wenn } t<0 \\ \alpha e^{-\alpha t} & , \text{wenn } t \geq 0 \end{cases}$$

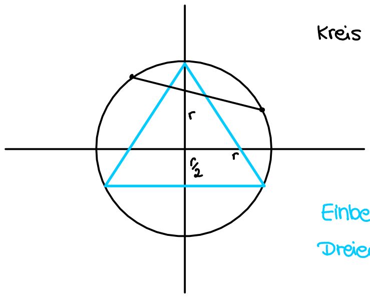


Definition (Exponentialverteilung)

Sei $\alpha \in \mathbb{R}_{>0}$ eine pos. reelle Zahl. Die Exponentialverteilung zum Parameter α ist die Verteilung auf $(\mathbb{R}, \mathcal{B})$ mit der Dichtefunktion

$$\rho(t) = \begin{cases} 0 & , \text{wenn } t<0 \\ \alpha e^{-\alpha t} & , \text{wenn } t \geq 0 \end{cases}$$

Beispiel ③ (Bertrand'sche Paradox)



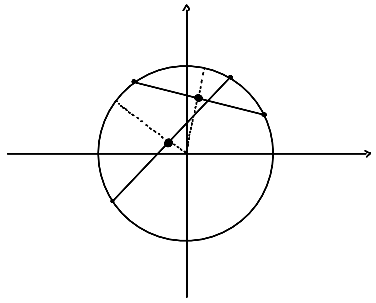
Kreis mit Radius r

Einbeschriebenes gleichseitiges Dreieck

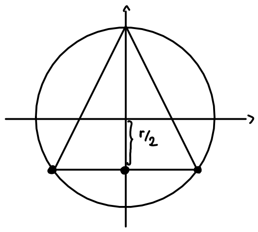
- Eine Sehne wird „rein zufällig" gezogen
- Mit welcher W-keit ist die Sehne länger als die Seiten des eingeschriebenen gleichseitigen Dreiecks?

Bemerkung

- Die Sehne wird durch ihren Mittelpunkt bestimmt.



- Die Länge der Sehne sinkt als der Abstand zwischen dem Mittelpunkt und den Kreismittelpunkt steigt.

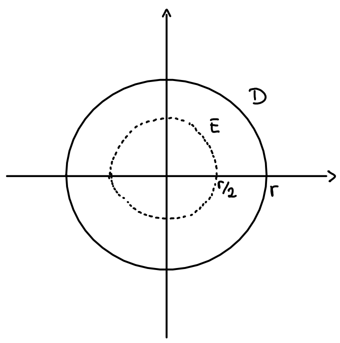


- Die Seite des eingeschriebenen gleichseitigen Dreiecks ist eine Sehne mit Abstand $r/2$.

Deshalb,

Sehne länger als Seite $\Longleftrightarrow$ Abstand vom Kreismittelpunkt $< r/2$.

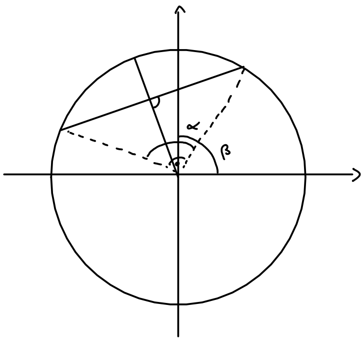
- „rein zufällig" $\Rightarrow$ Der Mittelpunkt der Sehne wird gleichverteilt auf dem Kreis.



E = der Kreis mit Radius $r/2$

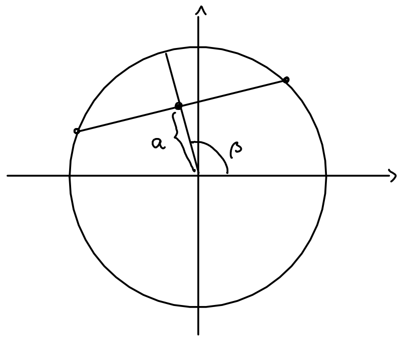
$$\mathbb{P}(E) = \frac{|E|}{|D|} = \frac{\cancel{\pi}\left(\frac{\cancel{r}}{2}\right)^2}{\cancel{\pi}\,\cancel{r^2}} = \frac{1}{4}.$$

Andere Berechnungen



$(\alpha, \beta) \rightsquigarrow$ Sehne

$\to \frac{1}{3}$

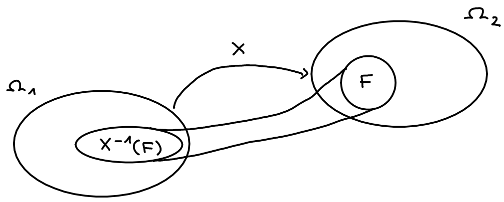


(a, β)

$\to \frac{1}{2}$

Zufallsvariablen

Sei $(\Omega_1, \mathcal{E}_1, \mathbb{P}_1)$ ein W-raum.



Im diskreten Fall

- Jede Abbildung $X: \Omega_1 \to \Omega_2$ ist eine Zufallsvariable.
- $\forall\ F \subseteq \Omega_2$,

  $$\mathbb{P}_X(F) = \mathbb{P}_1(X^{-1}(F))$$

In dem allgemeinen Fall

- Es ist nicht immer möglich, jeder Teilmenge $F \subseteq \Omega_2$ eine W-keit zuzuordnen.

Deshalb brauchen wir auch eine σ-Algebra $\mathcal{E}_2$ auf Ω_2, die alle „wichtigen" Teilmengen von Ω_2 enthält. (*)

(Im diskreten Fall: $\mathcal{E}_2 = \mathcal{P}(\Omega_2)$.)

Problem

$\mathbb{P}_1$ wird nur auf $\mathcal{E}_1$ definiert.

$\Rightarrow$ Wir brauchen, dass $X^{-1}(F) \in \mathcal{E}_1$, wenn wir $\mathbb{P}_X(F)$ so definieren möchten.

Lösung : (*)

Definition (Zufallsvariable)

Sei $(\Omega_1, \mathcal{E}_1, \mathbb{P}_1)$ ein W-raum und sei $(\Omega_2, \mathcal{E}_2)$ ein Ereignisraum. ($\mathcal{E}_2$ ist eine σ-Algebra auf Ω_2)

Eine Abbildung $X: \Omega_1 \to \Omega_2$ ist eine Zufallsvariable, wenn

$$\forall F \in \mathcal{E}_2 , \quad X^{-1}(F) \in \mathcal{E}_1 .$$

Dann wird die Abbildung $\mathbb{P}_X : \mathcal{E}_2 \to [0,1]$, die durch $\mathbb{P}_X(F) = \mathbb{P}_1(X^{-1}(F))$ definiert wird, das induzierte W-maß von X genannt.

Wiederholung

Definition (Zufallsvariable)

Sei $(\Omega_1, \mathcal{E}_1, \mathbb{P}_1)$ ein W-raum, und sei $(\Omega_2, \mathcal{E}_2)$ ein Ereignisraum.

Eine Abbildung $Z : \Omega_1 \to \Omega_2$ ist eine (Ω_2-wertige) Zufallsvariable,

wenn für alle Ereignisse $F \in \mathcal{E}_2$:

$$Z^{-1}(F) \in \mathcal{E}_1 .$$

Definition (Induzierte W-räume)

Sei $(\Omega_1, \mathcal{E}_1, \mathbb{P}_1)$ ein W-raum, sei $(\Omega_2, \mathcal{E}_2)$ ein Ereignisraum, und sei $Z : \Omega_1 \to \Omega_2$ eine Zufallsvariable.

Die Abbildung $\mathbb{P}_Z : \mathcal{E}_2 \to [0,1]$, die durch $\mathbb{P}_Z(F) = \mathbb{P}_1(Z^{-1}(F))$ definiert wird, wird das von Z induziertes W-maß genannt, und der W-raum $(\Omega_2, \mathcal{E}_2, \mathbb{P}_Z)$ wird der von Z induzierte W-raum genannt.

Behauptung

Sei $(\Omega_1, \mathcal{E}_1, \mathbb{P}_1)$ ein W-raum und sei $(\Omega_2, \mathcal{E}_2)$ ein Ereignisraum, wobei $\mathcal{E}_2 = \sigma(\mathcal{M})$. Eine Abbildung $Z : \Omega_1 \to \Omega_2$ ist eine Zufallsvariable genau dann, wenn

$$Z^{-1}(M) \in \mathcal{E}_1 \text{ für alle } M \in \mathcal{M}.$$

Beweis HA. $\square$

Korollar

Sei $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P}_1)$ eine Verteilung. Eine stetige Abbildung $Z : \mathbb{R}^n \to \mathbb{R}$ ist eine Zufallsvariable (in Bezug auf die Borel'sche σ-Algebra)

Beweis

$\mathcal{B} = \sigma(\mathcal{M})$, wobei $\mathcal{M} = \{(-\infty, c],\ c \in \mathbb{R}\}$.

Deshalb ist es genug, wenn wir überprüfen, dass $Z^{-1}((-\infty, c]) \in \mathcal{B}^n$.

$(-\infty, c]$ ist eine abgeschlossene Teilmenge von $\mathbb{R}$, und Z ist stetig.

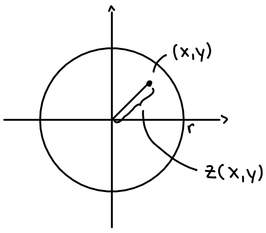
Deshalb ist das Urbild $Z^{-1}((-\infty, c])$ auch eine abgeschlossene Menge (in $\mathbb{R}$), also eine Borel-Menge.

Deshalb ist Z eine Zufallsvariable. $\square$

Beispiel

$Z : \mathbb{R}^2 \to \mathbb{R}$

$Z(x,y) = \sqrt{x^2 + y^2}$



Bedingte Wahrscheinlichkeiten

Definition (bedingte W-keit)

Seien $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum, und seien $A, B \in \mathcal{E}$ Ereignisse, sodass $\mathbb{P}(B) > 0$. Die <u>bedingte W-keit von A unter B</u> ist

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} .$$

Satz (Totalen W-keit)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum, sei $A \in \mathcal{E}$ ein Ereignis, und $\Omega = \dot{\bigcup}_{i \in I} B_i$ eine Zerlegung in paarweise disjunkten Ereignissen $B_i \in \mathcal{E}$, wobei I höchstens abzählbar ist und $\mathbb{P}(B_i) > 0$ für jedes $i \in I$. Dann gilt

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i) .$$

genau wie im diskreten Fall

Satz (Bayes)

Seien $(\Omega, \mathcal{E}, \mathbb{P})$, A, I, $\{B_i : i \in I\}$ wie in der totalen W-keit.

Dann gilt, für jedes $k \in I$,

$$\mathbb{P}(B_k \mid A) = \frac{\mathbb{P}(A \mid B_k)\, \mathbb{P}(B_k)}{\sum_{i \in I} \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i)} . \quad \left(= \frac{\mathbb{P}(A \mid B_k)\, \mathbb{P}(B_k)}{\mathbb{P}(A)} \right)$$

Unabhängigkeit

Definition (Unabhängigkeit von Ereignissen)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum und sei $I \neq \emptyset$ eine Indexmenge.

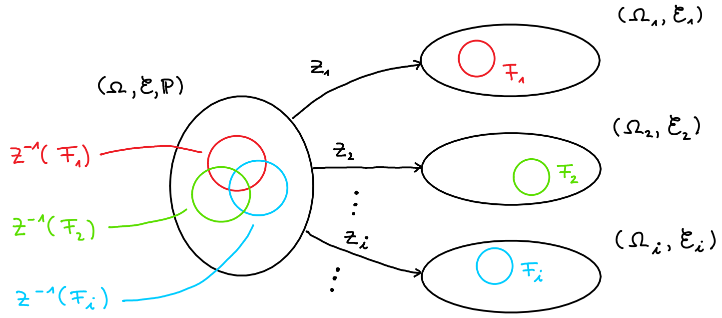
Eine Menge von Ereignissen $\{E_i : i \in I, E_i \in \mathcal{E}\}$ ist

<u>unabhängig</u>, wenn für alle endlichen $J \subseteq I$,

$$\mathbb{P}\left(\bigcap_{j \in J} E_j\right) = \prod_{j \in J} \mathbb{P}(E_j) .$$

Zwei Ereignisse $A, B \in \mathcal{E}$ sind unabhängig, wenn

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) .$$



Definition (Unabhängigkeit von Zufallsvariablen)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum und sei $I \neq \emptyset$ eine Indexmenge.

Für jedes $i \in I$ sei $(\Omega_i, \mathcal{E}_i)$ ein Ereignisraum und

sei $Z_i : \Omega \to \Omega_i$ eine Zufallsvariable.

Die Menge von Zufallsvariablen $\{Z_i : i \in I\}$ ist <u>unabhängig</u>,

wenn für eine beliebige Wahl von Ereignissen $\{F_i : i \in I, F_i \in \mathcal{E}_i\}$,

die Menge von Urbildern $\{Z_i^{-1}(F_i) : i \in I\}$ ist unabhängig.

Unabhängigkeits-kriterien

- $(\Omega, \mathcal{E}, \mathbb{P})$, $I \neq \emptyset$, $(\Omega_i, \mathcal{E}_i)$ für $i \in I$
- $Z_i : \Omega \to \Omega_i$ Zufallsvariable
- Für jedes $i \in I$ gibt es $\mathcal{G}_i \subseteq \mathcal{P}(\Omega_i)$, sodass

$\mathcal{E}_i = \sigma(\mathcal{G}_i)$ und $\forall\, F, G \in \mathcal{G}_i$, $F \cap G \in \mathcal{G}_i$

($\mathcal{G}_i$ ist <u>schnitt-stabil</u>)

Die Menge von Zufallsvariablen $\{Z_i : i \in I\}$ ist unabhängig,

genau dann, wenn $\{ Z_i^{-1}(G_i) : i \in I\}$ unabhängig ist,

für <u>beliebige</u> $G_i \in \mathcal{G}_i$

Korollar

- $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum
- $Z_i : \Omega \to \mathbb{R}$ Zufallsvariable , $i \in I$

Die Menge $\{Z_i : i \in I\}$ ist unabhängig genau dann, wenn

für beliebige endliche $J \subseteq I$ und $c_j \in \mathbb{R}$, $j \in J$,

$$\mathbb{P}\Big(\bigcap_{j\in J} \{Z_j \le c_j\}\Big) = \prod_{j \in J} \mathbb{P}(Z_j \le c_j) = \prod_{j\in J} F_{Z_j}(c_j) .$$

Situation

- Seien $\mathbb{P}_1, \mathbb{P}_2, \dots, \mathbb{P}_n$ Verteilungen auf $(\mathbb{R}, \mathcal{B})$, mit Dichtefunktionen $\rho_1, \rho_2, \dots, \rho_n : \mathbb{R} \to [0,\infty)$.
- Wir erzeugen einen W-raum $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ durch die n-dimensionale Dichtefunktion $\rho : \mathbb{R}^n \to [0,\infty)$

$$\rho(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} \rho_i(x_i) .$$

$$\mathbb{P}(E) = \int \dots \int_E \rho(x_1, x_2, \dots, x_n)\, dx_1 \dots dx_n$$

Behauptung

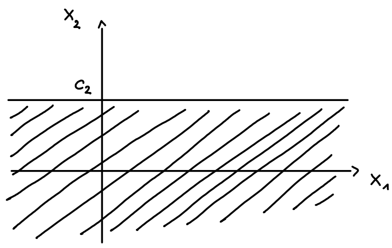
- $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ ist ein W-raum .
- Die Projektion $\mathrm{Proj}_i : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{Proj}_i(x_1, x_2, \dots, x_n) = x_i$$

  ist eine Zufallsvariable , und das induzierte W-maß ist $\mathbb{P}_i$.
- $\{\mathrm{Proj}_i : 1 \le i \le n\}$ ist unabhängig

Beweis

Sei $J \subseteq [n]$ eine beliebige Teilmenge und seien c_j, $j \in J$, beliebige reelle Zahlen.



$J = [2]$

$$\mathbb{P}\Big(\bigcap_{j\in J} \{ Z_j \leq c_j \}\Big) = \int_{\{(x_1,\dots,x_n) \cdot \; x_j \leq c_j \; \forall j \in J\}} \cdots \int \rho(x_1,\dots,x_n)\, dx_1 \dots dx_n$$

$$= \int_{\{(x_1,\dots,x_n) \cdot \; x_j \leq c_j \; \forall j \in J\}} \cdots \int \prod_{i=1}^{n} \rho_i(x_i)\, dx_1 \dots dx_n$$

[Fubini]

$$= \prod_{i=1}^{n} \int_{E_i} \rho(x_i)\, dx_i \,, \text{ wobei } E_i = \begin{cases} \mathbb{R} & \text{, wenn } i \notin J \\ (-\infty, c_i) & \text{, wenn } i \in J \end{cases}$$

$$= \Big(\prod_{j\in J} F_{\mathbb{P}_j}(c_j) \Big) \underbrace{\Big(\prod_{i \notin J} \int_{\mathbb{R}} \rho_i(x_i)\, dx_i \Big)}_{=1, \text{ weil } \rho_i \text{ eine Dichtefunktion ist.}}$$

$$= \prod_{j\in J} F_{\mathbb{P}_j}(c_j)$$

$$= \prod_{j\in J} \mathbb{P}(\mathrm{Proj}_j \leq c_j) \quad \checkmark \qquad \square$$

Wiederholung

Behauptung

Seien $\mathbb{P}_i$, $1 \le i \le n$, Verteilungen auf $(\mathbb{R}, \mathcal{B})$ mit Dichtefunktionen

$\rho_i \; \mathbb{R} \to [0, \infty)$.

Durch die n-dimensionale Dichtefunktion $\rho(x_1, x_2, .., x_n) = \prod_{i=1}^{n} \rho_i(x_i)$

können wir einen W-raum $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ definieren.

Die i^{te} Projektion $Proj_i : \mathbb{R}^n \to \mathbb{R}$ ist eine Zufallsvariable, die

die Verteilung $\mathbb{P}_i$ induziert.

Schließlich ist die Menge von Projektionen $\{Proj_i : 1 \le i \le n\}$ unabhängig.

Summen und Faltungen

Situation

- W-raum $(\Omega, \mathcal{E}, \mathbb{P})$
- Zwei unabhängige Zufallsvariablen $X, Y: \Omega \to \mathbb{R}$ mit Dichtefunktionen

  $f, g : \mathbb{R} \to [0, \infty)$

Frage

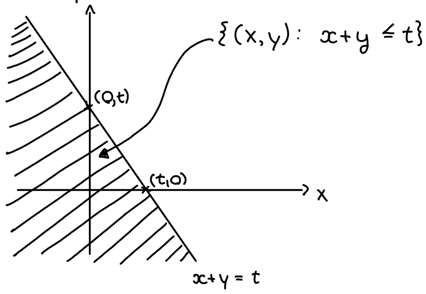
Was ist die Verteilung von $Z = X + Y$?

Strategie

① Wir berechnen die Verteilungsfunktion

② Wir berechnen die Dichtefunktion

Verteilungsfunktion



$F_Z(t) = \mathbb{P}(Z \le t)$

$\quad = \mathbb{P}(\{(x,y): x+y \le t\})$

Weil X und Y unabhängig sind, ist die Dichtefunktion

von dem Paar (X, Y) $\quad \rho(x,y) = f(x)\, g(y)$

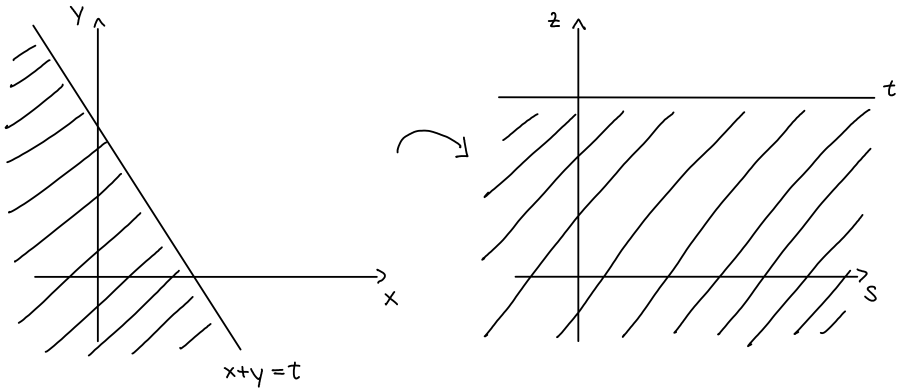
$\mathbb{P}((X,Y) \in E [\subseteq \mathbb{R}^2]) = \iint_E \rho(x,y)\, dx\, dy$

Also,

$$F_Z(t) = \iint_{\{(x,y)\,:\, x+y \leq t\}} f(x)\, g(y) \quad dx\, dy$$

Wir machen eine Variablensubstitution um eine schönere Fläche zu bekommen.

$$\left.\begin{aligned} x &= s \\ y &= z-s \end{aligned}\right\} \Rightarrow \{(x,y) : x+y \leq t\} = \{(s,z) : z \leq t\}$$



Der Jacobi-Operator:

$$\begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial z} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

$$\left|\begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial z} \end{pmatrix}\right| = \left|\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right| = 1$$

$$F_Z(t) = \iint_{\{(x,y)\,:\, x+y \leq t\}} f(x)\, g(x) \quad dx\, dy$$

$$= \iint_{\{(s,z)\,:\, z \leq t\}} f(s)\, g(z-s) \quad ds\, dz$$

$$F_{X+Y}(t) = \int_{-\infty}^{t} \int_{-\infty}^{\infty} f(s)\, g(z-s)\, ds\, dz$$

↳ Verteilungfunktion von $Z = X+Y$.

- Die Dichtefunktion ρ_{X+Y} ist die Ableitung von F_{X+Y} nach t.

Per dem Fundamentalsatz der Analysis ist

$$\rho_{X+Y}(t) = F_{X+Y}'(t)$$

$$\boxed{\rho_{X+Y}(t) = \int_{-\infty}^{\infty} f(s)\, g(t-s)\, ds}$$

Definition (Faltung)

Seien $f, g : \mathbb{R} - \mathbb{R}$ zwei Funktionen. Die Faltung von f und g ist die Abbildung $f*g : \mathbb{R} \to \mathbb{R}$, die durch

$$f*g(t) = \int_{-\infty}^{\infty} f(s)\, g(t-s)\, ds$$

definiert wird, wenn dieses Integral existiert

Satz

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum und seien $X, Y : \Omega \to \mathbb{R}$ zwei unabhängige Zufallsvariablen mit Dichtefunktionen f und g.

Dann ist die Dichtefunktion der Summe von $X+Y$ die Faltung von f und g;

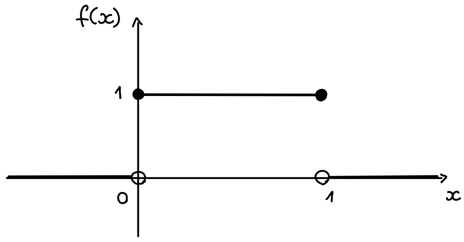
$$\rho_{X+Y}(t) = \int_{-\infty}^{\infty} f(s)\, g(t-s)\, ds$$

Beispiel

Seien X und Y unabhängige Gleichverteilungen auf $[0,1]$.

Was ist die Verteilung von $X+Y$?

$$f(x) = g(x) = \begin{cases} \frac{1}{1-0} = 1 & \text{, wenn } x \in [0,1] \\ 0 & \text{, sonst} \end{cases}$$



Was ist die Faltung von f und g?

$$f * g\,(t) = \int_{-\infty}^{\infty} f(s)\, g(t-s)\, ds$$

$$f(s)\,g(t-s) = \begin{cases} 1 & \text{, wenn } 0 \le s \le 1 \ \underline{\text{und}} \ 0 \le t-s \le 1 \\ 0 & \text{, sonst.} \end{cases}$$

$$f(s)\,g(t-s) = \begin{cases} 1 & \text{, wenn } s \in [0,1] \cap [t-1, t] \\ 0 & \text{, sonst.} \end{cases}$$

$$f*g\,(t) = \int_{[0,1] \cap [t-1,t]} 1 \, ds$$

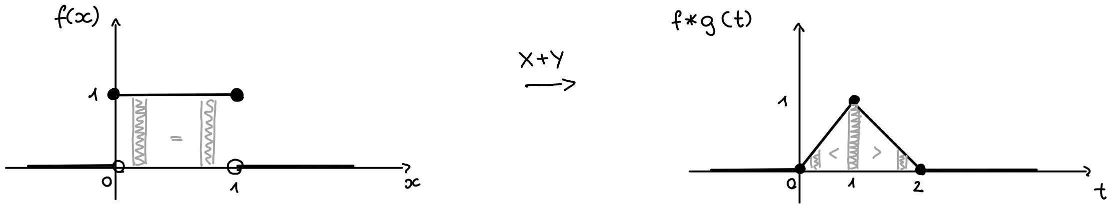
$$= \left| [0,1] \cap [t-1,t] \right|$$

Also,

$$f*g\,(t) = \left| [0,1] \cap [t-1,t] \right|$$

$$= \begin{cases} |\emptyset| = 0 & \text{, wenn } t < 0 \\ |[0,t]| = t & \text{, wenn } 0 \le t \le 1 \\ |[t-1,1]| = 2-t & \text{, wenn } 1 \le t \le 2 \\ |\emptyset| = 0 & \text{, wenn } t > 2 \end{cases}$$

$$= \begin{cases} t & \text{, wenn } 0 \le t \le 1 \\ 2-t & \text{, wenn } 1 \le t \le 2 \\ 0 & \text{, sonst.} \end{cases}$$



Erwartungswert

Im diskreten Fall : $\mathbb{E}[Z] = \sum_{\omega\in\Omega} Z(\omega)\,\mathbb{P}(\omega)$

In den stetigen Fall : $\int$ ρ

Definition (Erwartungswert)

Sei $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ eine Verteilung mit Dichtefunktion $\rho : \mathbb{R}^n \to [0,\infty)$.

Sei $Z : \mathbb{R}^n \to \mathbb{R}$ eine reellwärtige Zufallsvariable. Wenn

$\int_{\mathbb{R}^n} |Z(x)|\,\rho(x)\,dx < \infty$, definieren wir den <u>Erwartungswert von Z</u>

durch

$$\mathbb{E}[Z] = \int_{\mathbb{R}^n} Z(x)\,\rho(x)\,dx$$

Bemerkung

(i) Wenn es eine kompakte Teilmenge $\Omega \subseteq \mathbb{R}^n$ gibt (z.B.: $[c,d] \subseteq \mathbb{R}$) sodass:

- $\rho(x) = 0 \quad \forall x \notin \Omega$
- ρ und Z sind stetig auf Ω.

$$\Rightarrow \int_{\mathbb{R}^n} |Z(x)|\,\rho(x)\,dx = \int_{\Omega} |Z(x)|\,\rho(x)\,dx < \infty$$

(ii) Der EW einer Verteilung ist der EW der Identität auf der Verteilung.

Beispiel ① Die Gleichverteilung auf $[c,d]$

Wir wissen,

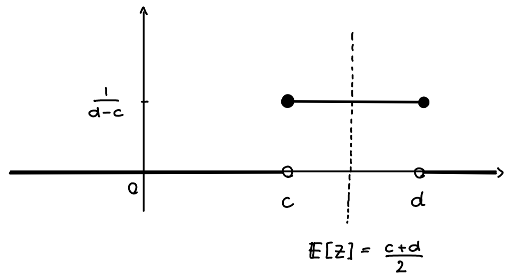
$$\rho(x) = \begin{cases} \frac{1}{d-c}, & \text{wenn } x \in [c,d] \\ 0, & \text{sonst.} \end{cases}$$

$Z = \mathrm{Id}$; $Z(x) = x$.

$$\begin{aligned}\mathbb{E}[Z] &= \int_{-\infty}^{\infty} Z(x)\,\rho(x)\,dx \\ &= \int_c^d x \cdot \frac{1}{d-c}\,dx \\ &= \frac{1}{d-c}\left[\frac{1}{2}x^2\right]_c^d\end{aligned}$$

$$= \frac{d^2 - c^2}{2(d-c)}$$

$$\boxed{\mathbb{E}[Z] = \frac{c+d}{2}}$$



Beispiel ②

$\rho = \mathbb{R} \to [0, \infty)$

$$\rho(t) = \begin{cases} t & \text{, wenn } 0 \leq t \leq 1 \\ 2-t & \text{, wenn } 1 \leq t \leq 2 \\ 0 & \text{, sonst.} \end{cases}$$

$Z : \mathbb{R} \to \mathbb{R}$ ist die Identität.

$$\mathbb{E}[Z] = \int_{\mathbb{R}} x \cdot \rho(x) \, dx = \int_0^1 x \cdot x \, dx + \int_1^2 x(2-x) \, dx$$

$$= \left[\tfrac{1}{3} x^3\right]_0^1 + \left[x^2 - \tfrac{1}{3} x^3\right]_1^2$$

$$= \tfrac{1}{3} + \left(4 - \tfrac{8}{3}\right) - \left(1 - \tfrac{1}{3}\right)$$

$$= 1.$$

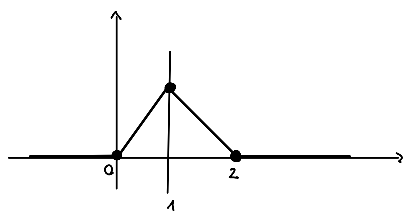
$$= \tfrac{1}{2} + \tfrac{1}{2}$$

$$= \tfrac{1-0}{2} + \tfrac{1-0}{2}$$

$= \mathbb{E}[X] + \mathbb{E}[Y]$, wobei X, Y Gleichverteilungen auf $[0,1]$ sind.

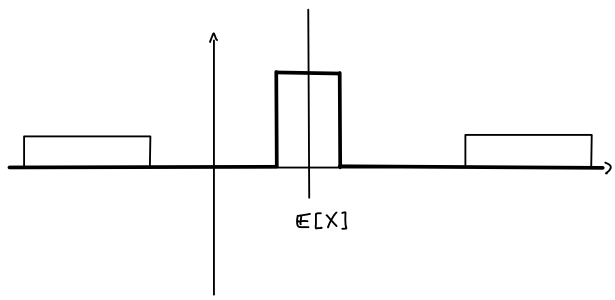
Linearität des EWes:

$\mathbb{E}[Z] = \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$



Definition (Varianz, Streuung)

Sei $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ ein W-raum mit Dichtefunktion $\rho : \mathbb{R}^n \to [0,\infty)$ und sei $Z = \mathbb{R}^n \to \mathbb{R}$ eine Zufallsvariable, deren EW existiert.



Die <u>Varianz von Z</u> ist

$$V(Z) = \mathbb{E}((Z - \mathbb{E}[Z])^2) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2.$$

Die <u>Streuung</u> ist $\sigma(Z) = \sqrt{V(Z)}$.

Wiederholung

Definition (Erwartungswert)

Sei $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ ein W-raum mit Dichtefunktion $\rho : \mathbb{R} \to [0, \infty)$,

und sei $Z : \mathbb{R}^n \to \mathbb{R}$ eine Zufallsvariable.

Wenn $\int_{\mathbb{R}^n} |Z(x)| \, \rho(x) \, dx < \infty$, definieren wir den

Erwartungswert von Z durch

$$\mathbb{E}(Z) = \int_{\mathbb{R}^n} Z(x) \, \rho(x) \, dx .$$

Definition (Varianz)

Sei $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P})$ ein W-raum mit Dichtefunktion ρ, und sei

$Z : \mathbb{R}^n \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert

existiert.

Die Varianz von Z ist

$$V(Z) = \mathbb{E}\left[(Z - \mathbb{E}[Z]^2) \right] = \int_{\mathbb{R}^n} (Z(x) - \mathbb{E}[Z])^2 \, \rho(x) \, dx,$$

wenn dieser Erwartungswert existiert.

In diesem Fall wird die Streuung von Z durch

$$\sigma(Z) = \sqrt{V(Z)}$$

definiert.

Normalverteilung

Definition (Normalverteilung)

Sei $a \in \mathbb{R}$ eine reelle Zahl und sei $\sigma > 0$ eine positive reelle Zahl.

Die Normalverteilung mit Erwartungswert a und Varianz σ^2 ist

der W-raum $(\mathbb{R}, \mathcal{B}, n_{a,\sigma^2})$ mit der Dichtefunktion

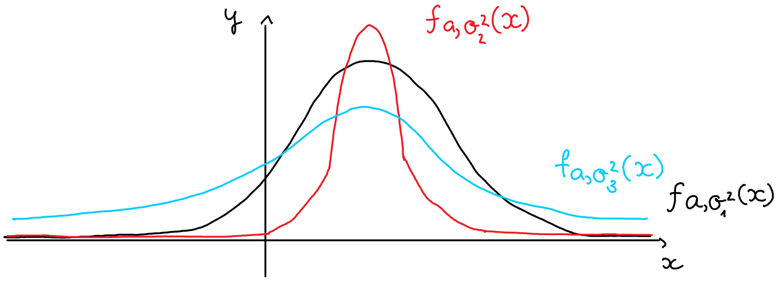
$$f_{a,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} .$$

Bemerkung

(i) Wenn wir die Parameter der Normalverteilung geben, geben wir die Varianz σ^2, nicht die Streuung σ.

(ii) Die Normalverteilung mit Erwartungswert 0 und Varianz 1, $(\mathbb{R}, \mathcal{B}, n_{0,1})$, heißt die Standard - Normalverteilung

Die Dichtefunktion $f_{a,\sigma^2}(x)$ ist die sogenannte „Gaußsche Glockenkurve."



$\sigma_2 < \sigma_1 < \sigma_3$

Wir müssen beweisen, dass $f_{a,\sigma^2}(x)$ wirklich eine Dichtefunktion ist.

(DF 1) Für jedes $c \in \mathbb{R}$, $\rho^{-1}((-\infty, c]) \in \mathcal{B}$. ✓

[f_{a,σ^2} hat die Eigenschaft, weil sie stetig ist]

(DF 2) $\int_{-\infty}^{\infty} f_{a,\sigma^2}(x)\, dx = 1.$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-a)^2}{2\sigma^2}}\, dx$$

Wir machen eine Variablensubstitution:

$u = \frac{x-a}{\sigma} \iff x = \sigma u + a$

$\frac{dx}{du} = \sigma$

$\lim_{x \to -\infty} u = -\infty$

$\lim_{x \to \infty} u = \infty$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-u^2}{2}} \sigma \, du = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}} du.$$

Wenn $g(u) = e^{-\frac{u^2}{2}}$ ist, ist die Ableitung

$$g'(u) = (-u) e^{-\frac{u^2}{2}}$$

Leider gibt es keine einfache Stammfunktion.

Wie können wir das Integral dann berechnen?

Satz (Gaußsches Integral)

$$\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = \sqrt{2\pi}.$$

Beweis

Sei $I = \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du$

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du\right)^2$$

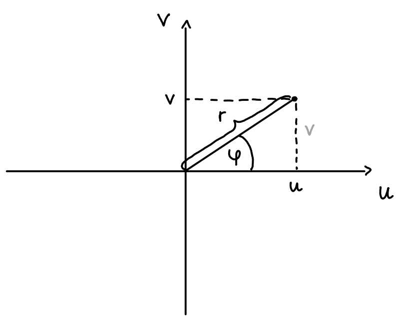
$$= \left(\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du\right)\left(\int_{-\infty}^{\infty} e^{-\frac{v^2}{2}} dv\right)$$

[Satz von Fubini]

$$= \iint_{\mathbb{R}^2} e^{-\frac{u^2}{2}} \cdot e^{-\frac{v^2}{2}} du\, dv.$$

$$= \iint_{\mathbb{R}^2} e^{-\frac{u^2+v^2}{2}} du\, dv.$$

Jetzt machen wir eine Variablensubstitution:



$$u = r \cdot \cos\varphi$$

$$v = r \cdot \sin\varphi$$

$$u^2 + v^2 = r^2 \cos^2\varphi + r^2 \sin^2\varphi = r^2$$

Jacobi-Operator:

$$\begin{pmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \varphi} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

Seine Determinante:

$$\left| \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \right| = r\cos^2\varphi + r\sin^2\varphi = r \; .$$

$$I^2 = \iint_{\mathbb{R}^2} e^{-\frac{u^2+v^2}{2}} \, du \, dv$$

$$= \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} \, dr \, d\varphi$$

Dieses Integral können wir berechnen, denn wir haben eine Stammfunktion (vorher genannt). Berechnen das innere Integral:

$$= \int_0^{2\pi} \left[-e^{-\frac{r^2}{2}} \right]_0^{\infty} d\varphi$$

$$= 2\pi \left[-e^{-\frac{r^2}{2}} \right]_0^{\infty}$$

$$= 2\pi \, (0-(-1))$$

$$= 2\pi$$

$\Rightarrow \quad I = \sqrt{2\pi} \qquad \square$ (Bemerke: I ist positiv.)

$\Rightarrow$ Die Normalverteilung ist eine Verteilung.

Ihre Verteilungsfunktion:

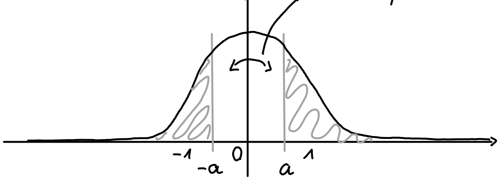
$$F_{a,\sigma^2}(z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} \, e^{-\frac{(x-a)^2}{2\sigma^2}} \, dx \, .$$

Wir brauchen einen Rechner / eine Tabelle, um $F_{a,\sigma^2}(z)$ zu berechnen.

Es ist genug,

$$\Phi(z) := F_{0,1}(z)$$

zu wissen.



$\Phi(-a) = \mathbb{P}(Z \le -a)$

[Achsen-Sym]

$= \mathbb{P}(Z \ge a)$

$= 1 - \mathbb{P}(Z \le a)$

$= 1 - \Phi(a)$

Das Gesetz der großen Zahlen

Situation

- Ein W-raum $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ mit Dichtefunktion $\rho : \mathbb{R} \to [0,\infty)$.
- Eine Zufallsvariable $X : \mathbb{R} \to \mathbb{R}$.
- „vielmals wiederholen" $\longleftrightarrow$ Eine Folge $X_1, X_2, X_3, \ldots$ von Zufallsvariablen, die identisch verteilt sind, mit Verteilung $\mathbb{P}_X$, <u>und</u> die Menge $\{X_1, X_2, ..\}$ ist unabhängig.

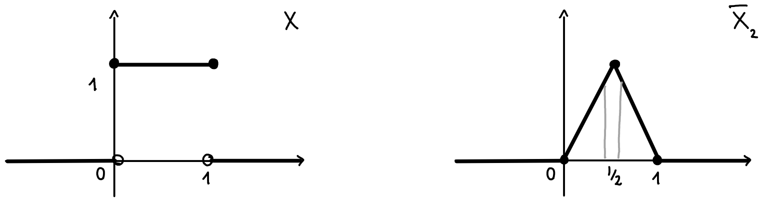
„durchschnittliche Wert" $\longleftrightarrow$ $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Frage

Was können wir über die Verteilung von $\overline{X}_n$ sagen?

Wir werden sehen, dass $\overline{X}_n$ nah $\mathbb{E}[X]$ $(= \mathbb{E}[X_i])$ liegt.

Beispiel Gleichverteilung auf $[0,1]$



$\overline{X}_2 = \frac{1}{2}(X_1 + X_2)$

<u>Wir brauchen:</u>

Satz (Tschebyscheff - Ungleichung)

- $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ ein W-raum mit Dichtefunktion ρ.
- $Z : \mathbb{R} \to \mathbb{R}$ eine Zufallsvariable, deren Varianz existiert.

Dann, für $\lambda > 0$,

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq \lambda) \leq \frac{V(Z)}{\lambda^2}$$

Satz (Markov - Ungleichung)

- $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ ein W-raum mit Dichtefunktion ρ.
- $Z: \mathbb{R} \to \mathbb{R}_{\geq 0}$ eine nicht negative Zufallsvariable, deren Erwartungswert existiert.

Dann gilt, für jedes $a > 0$,

$$\mathbb{P}(Z \geq a) \leq \frac{\mathbb{E}(Z)}{a}$$

Wiederholung

Satz (Markov - Ungleichung)

Sei $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ ein W-raum mit Dichtefunktion ρ, und sei $Z: \mathbb{R} \to [0,\infty)$ eine nicht negative Zufallsvariable, deren Erwartungswert existiert.

Dann gilt, für jedes $a > 0$,

$$\mathbb{P}(Z \geq a) \leq \frac{\mathbb{E}[Z]}{a}.$$

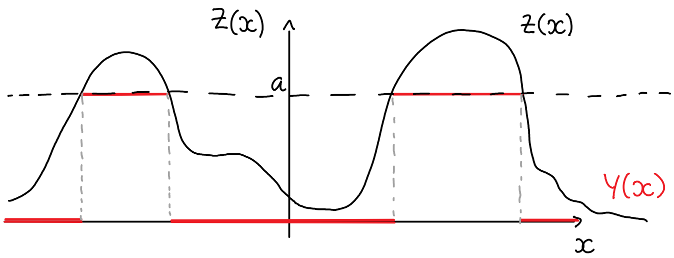
Satz (Tschebyscheff - Ungleichung)

Sei $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ ein W-raum mit Dichtefunktion ρ, und sei $Z: \mathbb{R} \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert und Varianz existieren.

Dann gilt, für jedes $\lambda > 0$,

$$\mathbb{P}(|Z - \mathbb{E}[Z]| \geq \lambda) \leq \frac{V(Z)}{\lambda^2}.$$

Beweis (Markov - Ungleichung)



Wir definieren eine neue Zufallsvariable $Y: \mathbb{R} \to [0,\infty)$

$$Y(x) = \begin{cases} a, & \text{wenn } Z(x) \geq a \\ 0, & \text{wenn } 0 \leq Z(x) < a. \end{cases}$$

$$(Y(x) = a \cdot \mathbb{1}_{\{Z \geq a\}}.)$$

Merken Sie sich, dass

$$Y(x) \leq Z(x), \quad \forall x \in \mathbb{R}.$$

$\Rightarrow \mathbb{E}[Y] \leq \mathbb{E}[Z]$ $\qquad \rightarrow$ (Monotonität des Erwartungswertes)

$$\mathbb{E}[Y] = \int_{\mathbb{R}} Y(x)\, \rho(x)\, dx$$

$$= \int_{\{x : Z(x) \geq a\}} a\, \rho(x)\, dx$$

$$= a \int_{\{x : Z(x) \geq a\}} \rho(x)\, dx$$

$$= a \cdot \mathbb{P}(Z \geq a)$$

Also,

$$a \cdot \mathbb{P}(Z \geq a) = \mathbb{E}[Y] \leq \mathbb{E}[Z].$$

$$\Rightarrow \mathbb{P}(Z \geq a) \leq \frac{\mathbb{E}[Z]}{a}. \qquad \square$$

Beweis (Tschebyscheff - Ungleichung)

Wir definieren die Zufallsvariable $T : \mathbb{R} \to [0, \infty)$,

$$T(x) = (Z(x) - \mathbb{E}[Z])^2.$$

Dann ist

$$\mathbb{E}[T] = \mathbb{E}[(Z - \mathbb{E}[Z])^2] = V(Z).$$

Wir wenden die Markov - Ungleichung mit $a = \lambda^2$ an:

$$\mathbb{P}(T \geq \lambda^2) \leq \frac{\mathbb{E}[T]}{\lambda^2} = \frac{V(Z)}{\lambda^2}$$

$$\{T \geq \lambda^2\} = \{(Z - \mathbb{E}[Z])^2 \geq \lambda^2\}$$

$$= \{|Z - \mathbb{E}[Z]| \geq \lambda^2\}$$

$$\Rightarrow \mathbb{P}(|Z - \mathbb{E}[Z]| \geq \lambda) \leq \frac{V(Z)}{\lambda^2} \qquad \square$$

Behauptung

Diese Ungleichung ist nur hilfreich, wenn

$$\lambda^2 \geq V(Z), \text{ oder } \lambda \geq \sigma(Z).$$

Das Gesetz der großen Zahlen

Satz (Schwaches Gesetz der großen Zahlen)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum und seien $X_1, X_2, \dots, X_n : \Omega \to \mathbb{R}$ Zufallsvariablen, sodass für jedes $i \in [n]$, $\mathbb{E}[i]$ und $V(i)$ existieren, und die Menge $\{X_1, X_2, \dots, X_n\}$ unabhängig ist.

Sei $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$ der Durchschnitt.

Dann gilt für jedes $\varepsilon > 0$,

$$\mathbb{P}\left(\left| \overline{X}_n - \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i] \right| \geq \varepsilon \right) \leq \frac{\sum_{i=1}^{n} V(X_i)}{n^2 \varepsilon^2}$$

Wichtiger Fall

- $X_1, X_2, \dots, X_n$ sind identisch verteilt, wie eine ZV $X : \Omega \to \mathbb{R}$.
- $\mathbb{E}[X_i] = \mathbb{E}[X] \quad \forall i$.

  $\Rightarrow \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i] = \mathbb{E}[X]$.

- $V(X_i) = V(X) \quad \forall i$.

  $\Rightarrow \sum_{i=1}^{n} V(X_i) = n\, V(X)$.

$$\Rightarrow \mathbb{P}\left(\left| \overline{X}_n - \mathbb{E}[X] \right| \geq \varepsilon \right) \leq \frac{n\, V(X)}{n^2 \varepsilon^2} = \frac{V(X)}{n \varepsilon^2} \xrightarrow{n \to \infty} 0$$

Beweis

Wir wenden die Tschebysheff-Ungleichung an, mit $Z = \overline{X}_n$ und $a = \varepsilon$.

- $\mathbb{E}[\overline{X}_n] = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} X_i \right] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i]$.

  [Liniarität des Erwartungswertes]

- $V(\overline{X}_n) = V\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i \right) = \frac{1}{n^2} \sum_{i=1}^{n} V(X_i)$

  [die ZV X_i sind unabhängig]

$V(cX) = c^2 V(X)$

$V(cX) = \mathbb{E}[c^2 (X - \mathbb{E}[X])^2] = c^2\, \mathbb{E}[(X - \mathbb{E}[X])^2]$

Per Tschebyscheff:

$$\mathbb{P}(|\bar{X}_n - \mathbb{E}[\bar{X}_n]| \geq \varepsilon) \leq \frac{V(\bar{X}_n)}{\varepsilon^2}$$

$$\|$$

$$\mathbb{P}(|\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i]| \geq \varepsilon) \leq \frac{\sum_{i=1}^{n} V(X_i)}{n^2 \varepsilon^2}$$

$\square$

Bemerkung

- Wenn $X_1, X_2, X_3, \ldots$ unabhängig und identisch verteilt wie X sind, dann, für jedes $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|\bar{X}_n - \mathbb{E}[X]| \geq \varepsilon) = 0$$

passt auch, wenn nur $\mathbb{E}[X]$ existiert.

Satz (Starkes Gesetz der großen Zahlen)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum und sei $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert und Varianz existieren. Seien $X_1, X_2, X_3, \ldots : \Omega \to \mathbb{R}$ Zufallsvariablen, die identisch verteilt wie X sind und paarweise unabhängig voneinander sind.

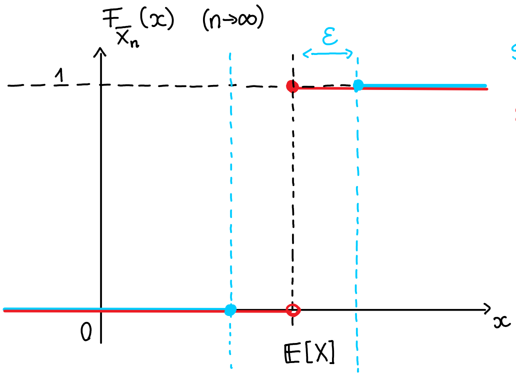
Sei, für jedes $n \in \mathbb{N}$, $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ der Durchschnitt der ersten n Zufallsvariablen.

Dann gilt,

$$\mathbb{P}(\lim_{n\to\infty} \bar{X}_n = \mathbb{E}[X]) = 1 .$$

Dieser Satz wird in dieser Vorlesung nicht beweisen, aber wir sollen die Aussage und den Unterschied zum schwachen Gesetz verstehen.

Starkes gegen Schwaches



Schwaches Gesetz

Starkes Gesetz

Schwaches Gesetz

Die Durchschnitte $\overline{X}_n$ konvergieren in W-keit gegen $\mathbb{E}[X]$.

$$\forall \varepsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}\left(\left|\overline{X}_n - \mathbb{E}[X]\right| \geq \varepsilon \right) = 0.$$

Starkes Gesetz

Die Durchschnitte $\overline{X}_n$ konvergieren fast sicher gegen $\mathbb{E}[X]$.

$$\mathbb{P}\left(\lim_{n\to\infty} \overline{X}_n = \mathbb{E}[X] \right) = 1.$$

Konvergenz fast sicher $\overset{\Rightarrow}{\not\Leftarrow}$ Konvergenz in W-keit

Satz (Zentraler Grenzwertsatz)

Sei $(\Omega, \mathcal{E}, \mathbb{P})$ ein W-raum, und sei $X: \Omega \to \mathbb{R}$ eine Zufallsvariable, deren Erwartungswert und Varianz existieren, sodass

$$V(X) = \sigma(X)^2 > 0 .$$

Seien $X_1, X_2, X_3, \ldots : \Omega \to \mathbb{R}$ unabhängige Zufallsvariablen, die identisch verteilt wie X sind.

Für jedes $n \in \mathbb{N}$ definieren wir $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.

Sei Z ein Zufallsvariable, die wie die Standard-Normalverteilung verteilt ist.

(Satz weiter auf der nächsten Seite!)

Dann gilt, für ein beliebiges Intervall $[a,b] \subseteq \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\sqrt{n}(\overline{X}_n - \mathbb{E}[X])}{\sigma(X)} \in [a,b] \right) = \mathbb{P}(Z \in [a,b])$$

$$Y = \frac{\sqrt{n}(\overline{X}_n - \mathbb{E}[X])}{\sigma(X)}$$

$$\mathbb{E}[Y] = \mathbb{E}\left[\frac{\sqrt{n}}{\sigma(X)} (\overline{X}_n - \mathbb{E}[X]) \right]$$

$$= \frac{\sqrt{n}}{\sigma(X)} \left(\mathbb{E}[\overline{X}_n] - \mathbb{E}[X] \right)$$

$$= 0 \qquad = \mathbb{E}[Z] \qquad \checkmark$$

$$V(Y) = V\left(\frac{\sqrt{n}}{\sigma(X)} (\overline{X}_n - \mathbb{E}[X]) \right)$$

$$= \frac{\sqrt{n}}{\sigma(X)} \quad V(\overline{X}_n - \mathbb{E}[X])$$

$$= \frac{n}{V(X)} \underbrace{V(\overline{X}_n)}_{= \frac{V(X)}{n}} \qquad = 1 \quad = \quad V(Z) \qquad \checkmark$$

Wiederholung

Zentraler Grenzwertsatz

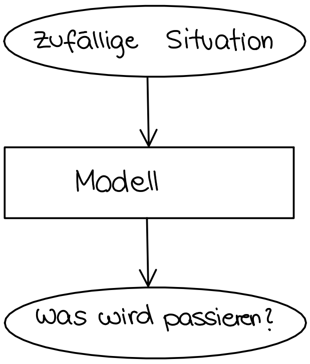
Sei Z eine Zufallsvariable, die der Standard-Normalverteilung entspricht. Für alle $c \leq d \in \mathbb{R}$

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\sqrt{n}}{\sigma} (\overline{X}_n - a) \in [c,d] \right) = \mathbb{P}(Z \in [c,d])$$

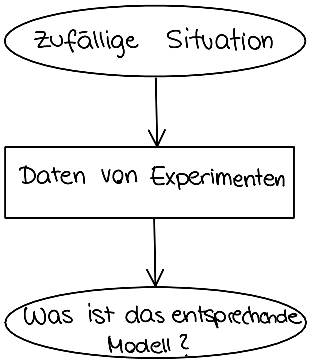
Das heißt, dass $\frac{\sqrt{n}}{\sigma} (\overline{X}_n - a) \to Z$ in Verteilung.

Statistik

W-theorie:



Statistik:



Zentrales Beispiel

Situation

Ein Importeur erhält eine Lieferung von $N = 10.000$ Orangen. Er möchte wissen, wieviel, sagen wir w davon faul sind, weil er den Preis nur zahlen muss, wenn höchstens 5 % faul sind.

Er macht eine Stichprobe von $n = 50$ Orangen und stellt fest, dass x faul sind.

Was kann er jetzt über w sagen?

1. Idee

Er kann versuchen mit x den Wert von ω zu schätzen.

Er kann sagen

$$\omega = T(x) \quad \rightarrow \text{Schätzer von } \omega.$$

Zum Beispiel, die Verteilung von x ist die hypergeometrische Verteilung zu den Parametern $N, R = \omega, n$

Der EW ist $\frac{\omega}{N} \cdot n$.

Wir können einsetzen $x = \frac{\omega}{N} \cdot n \iff \omega = N \cdot \frac{x}{n}$

2. Idee

Er kann sich nicht sicher sein, dass $T(x)$ der richtige Wert von ω ist. Stattdessen kann er ein Intervall $C(x) \subset \mathbb{R}$ suchen, sodass mit hoher W-keit $\omega \in C(\omega)$ gilt.

Solche Intervalle heißen <u>Konfidenzintervalle</u>.

Ziel: $C(x)$ so klein wie möglich.

3. Idee

Schließlich muss er entscheiden, ob er die Lieferung akzeptieren wird oder nicht. Wie soll er diese Entscheidung fällen?

→ Testen von Hypothesen.

Das parametrische Modell

Hier haben wir nicht nur einen Raum, sondern eine Menge von W-räumen. Für jeden möglichen Wert vom Parameter ω gibt es ein W-maß $\mathbb{P}_\omega = h_{N,\omega,n}$.

<u>Allgemein:</u>

- Die Beobachtungsereignisse x bilden den Stichprobenraum $\mathcal{X}$, der mit einer σ-Algebra $\mathcal{E}$

versehen ist.

- Für jeden möglichen Wert des unbekannten Parameters θ gibt es ein entsprechendes W-maß $\mathbb{P}_\theta : \mathcal{E} \to [0,1]$
- Θ ist der Parameterraum : Die Menge der möglichen Werte des Parameters.

Definition (Parametrisches Modell)

Ein parametrisches Modell ist eine Menge von möglichen W-räumen

$$\mathcal{M} := \{ (\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta \},$$

wobei $\Theta \subseteq \mathbb{R}^n$ der Parameterraum ist.

Beispiel ①

Unser Orangenbeispiel hat das parametrische Modell $\mathcal{M} = \{ (\{0,\dots,50\}, \mathcal{P}(\{0,\dots,50\}), h_{10.000,\omega,50}) : \omega \in \{0,\dots,10.000\}\}$

Beispiel ②

Das Intervall $[0,\theta]$ sei gegeben, wobei $\theta > 0$ unbekannt ist

Es werden n unabhängige gleichverteilte Zufallsvariablen aus $[0,\theta]$ gezogen. Was ist das Modell?

$\mathcal{M} = \{ ([0,\infty)^n, \mathcal{B}([0,\infty)^n), \mathbb{P}_\theta^n) : \theta \in \mathbb{R}^+ \}$, wobei

$\mathbb{P}_\theta^n$ die Dichtefunktion

$$f_\theta^n(x_1,\dots,x_n) = \prod_{i=1}^{n} f_\theta(x_i) = \begin{cases} \frac{1}{\theta^n} & \text{, wenn } x_i \in [0,\theta] \ \forall i \in [n] \\ 0 & \text{, sonst} \end{cases}$$

(Gleichverteilung auf $[0,\theta]^n$.)

Parameterschätzung

Definition (Schätzer)

Sei $\{(\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta\}$ ein parametrisches Modell.

Sei W eine Menge und sei $\mathcal{F}$ eine σ-Algebra auf W.

Sei $\tau : \Theta \to W$ eine Abbildung.

Zu einer Beobachtung $x \in \mathcal{X}$ wollen wir $T(x) \in W$

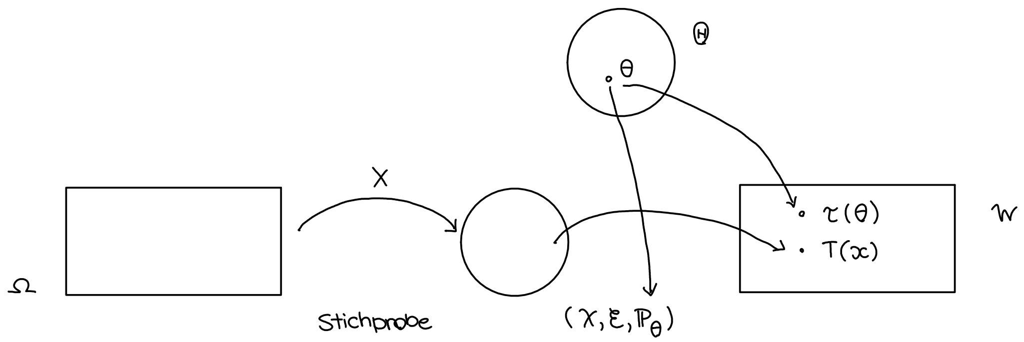
angeben, das den Parameter $\tau(\theta)$ schätzt

$T(x)$ heißt Schätzer von $\tau(\theta)$ und die ZV

$$T(X) : \Omega \to W$$

$$\omega \mapsto T(X(\omega))$$

heißt eine Statistik.



Beispiel (Orangenbeispiel)

Wir wissen, dass die Beobachtung der hypergeometrischen Verteilung zu den Parametern N, θ, n entspricht.

Der EW dieser Verteilung ist

$$\frac{\theta}{N} \cdot n \qquad = \mathbb{E}_\theta[X]$$

Wir können vermuten, dass die Beobachtung x nah diesem EW liegt.

$$x \approx \frac{\theta}{N} \cdot n$$

$\Rightarrow \quad \theta = \frac{x}{n} \cdot N = T(x)$

ein Schätzer von θ.

Erwartungstreuer Schätzer

Es ist nicht möglich, dass der Schätzer immer richtig ist.

Aber wir können fordern, dass der Schätzer im Mittel richtig ist.

im Mittel = der EW des Schätzers.

Aber der EW hängt von der Verteilung ab, und die Verteilung hängt vom Parameter θ ab.

$$\left.\begin{array}{l} \mathbb{E}_\theta[T(X)] \\ V_\theta(T(X)) \end{array}\right\} \text{ der EW und die Varianz des Schätzers, wenn } \theta \text{ der Parameter ist.}$$

Definition (erwartungstreu)

Sei $\{(\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta\}$ ein parametrisches Modell, und sei

$\tau : \Theta \to W$ ein Parameter, den wir schätzen möchten

(meistens ist $W = \Theta$ und τ die Identität – also schätzen wir θ direkt.)

Ein Schätzer $T : \mathcal{X} \to W$ von $\tau(\theta)$ heißt erwartungstreu, wenn für alle $\theta \in \Theta$,

$$\mathbb{E}_\theta[T(X)] = \tau(\theta).$$

(Informell: der Schätzer ist im Mittel richtig.)

Beispiel 1 (Orangenbeispiel)

Wir haben den Schätzer $T(x) = \frac{x}{n} \cdot N$ von θ.

Behauptung : T ist erwartungstreu.

Beweis : Wir müssen den EW berechnen

$$\mathbb{E}_\theta[T(X)] = \mathbb{E}_\theta\left[\frac{X}{n} \cdot N\right] \underset{\text{Liniarität des EW}}{=} \frac{N}{n} \cdot \mathbb{E}_\theta[X]$$

X hat die hypergeometrische Verteilung zu den Parametern N, θ, n, also

$$\mathbb{E}_\theta[X] = \frac{\theta}{N} \cdot n$$

$$\Rightarrow \mathbb{E}_\theta[T(X)] = \frac{\cancel{N}}{\cancel{n}} \cdot \frac{\theta}{\cancel{N}} \cdot \cancel{n} = \theta \quad \checkmark$$

$\square$

Beispiel 2

Wir haben die Gleichverteilung auf $[0,\theta]$, wobei $\theta > 0$ unbekannt ist.

Wir machen n unabhängige Stichproben $X_1, X_2, \dots, X_n$

$\rightarrow$ Beobachtungen $x_1, x_2, \dots, x_n$

Der EW der Gleichverteilung ist $\frac{\theta}{2}$

Schwaches Gesetz der großen Zahlen : der Stichprobenmittelwert $\frac{1}{n}\sum_{i=1}^{n} x_i$ liegt nah dem EW.

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} x_i \approx \frac{\theta}{2}.$$

$$\Rightarrow T_n(x_1, x_2, \dots, x_n) = \frac{2}{n}\sum_{i=1}^{n} x_i$$

Behauptung : Der Schätzer T_n ist erwartungstreu.

Beweis : $\mathbb{E}_\theta[T_n(X_1, X_2, \dots, X_n)] = \mathbb{E}_\theta\left[\frac{2}{n}\sum_{i=1}^{n} X_i\right]$

$$\underset{\left[\begin{smallmatrix}\text{Liniarität}\\ \text{des EW}\end{smallmatrix}\right]}{=} \frac{2}{n}\sum_{i=1}^{n} \mathbb{E}_\theta[X_i]$$

X_i ist gleichverteilt auf $[0,\theta]$ und deshalb

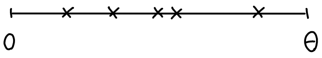
$$\mathbb{E}_\theta[T_n(X_1, \dots, X_n)] = \frac{2}{n}\sum_{i=1}^{n} \mathbb{E}_\theta[X_i]$$

$$= \frac{2}{n}\sum_{i=1}^{n} \frac{\theta}{2}$$

$$= \frac{\cancel{2}}{\cancel{n}} \cdot \frac{\cancel{n}\,\theta}{\cancel{2}} = \theta \quad \checkmark$$

$\square$

2. Schätzer



Wir bekommen n Zufallszahlen in $[0,\theta]$

$\Rightarrow \quad X_1, X_2, \dots, X_n \le \theta$

$\Rightarrow \quad \max\limits_{1 \le i \le n} X_i \le \theta$

Wenn n groß ist, haben wir viele Zahlen im Intervall $[0,\theta]$.

Deshalb können wir erwarten, dass eine davon nah θ liegt.

$\Rightarrow \quad \max\limits_{1 \le i \le n} x_i \approx \theta$

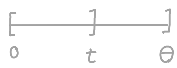
Das motiviert einen anderen Schätzer

$$T_n'(x_1, x_2, \dots, x_n) = \max_{1 \le i \le n} x_i \, .$$

Frage : Ist T_n' erwartungstreu?

Wir berechnen die Verteilungsfunktion von T_n' : $\qquad (0 \le t \le \theta)$

$$\begin{aligned}
\mathbb{P}_\theta(T_n'(X_1, X_2, \dots, X_n) \le t) &= \mathbb{P}_\theta(\max_{1 \le i \le n} X_i \le t) \\
&= \mathbb{P}_\theta(X_i \le t \ \forall i) \\
&\overset{[\text{Unabhängigkeit}]}{=} \prod_{i=1}^{n} \mathbb{P}_\theta(X_i \le t) \\
&\overset{\left[\substack{\text{Gleichverteilung} \\ \text{auf } [0,\theta]}\right]}{=} \prod_{i=1}^{n} \left(\frac{t}{\theta}\right) \\
&= \left(\frac{t}{\theta}\right)^n .
\end{aligned}$$



Die Verteilungsfunktion von T_n' ist $F_\theta(t) = \left(\frac{t}{\theta}\right)^n$. $(0 \le t \le n)$

Die Dichtefunktion ist dann

$$f_\theta(t) = \frac{d}{dt} F_\theta(t) = \begin{cases} \dfrac{n t^{n-1}}{\theta^n} & \text{, wenn } 0 \le t \le n \\ 0 & \text{, sonst.} \end{cases}$$

Deshalb

$$\mathbb{E}_\theta[T_n'(X_1, X_2, \dots, X_n)] = \int_{\mathbb{R}} t \cdot f_\theta(t)\, dt$$

$$= \int_0^\theta \frac{n t^n}{\theta^n}\, dt$$

$$= \frac{n}{\theta} \left[\frac{t^{n+1}}{n+1}\right]_0^\theta = \frac{n}{\theta^n} \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta.$$

Deshalb ist $T_n'(x_1, x_2, \dots, x_n) = \max_{1\le i\le n} x_i$ nicht erwartungstreu, weil

$\mathbb{E}_\theta[T_n'] = \frac{n}{n+1} \cdot \theta \neq \theta.$

Wir können $T_n^*(x_1, x_2, \dots, x_n) = \frac{n+1}{n} T_n'(x_1, x_2, \dots, x_n)$

$= \frac{n+1}{n} \max_{1\le i\le n} x_i$ definieren.

Dann ist

$$\mathbb{E}_\theta[T_n^*] = \mathbb{E}_\theta\left[\frac{n+1}{n} T_n'\right]$$

$$= \frac{n+1}{n} \mathbb{E}_\theta[T_n']$$

$$= \frac{\cancel{n+1}}{\cancel{n}} \cdot \frac{\cancel{n}}{\cancel{n+1}} \cdot \theta = \theta \quad \checkmark$$

$\Rightarrow$ T_n^* ist <u>erwartungstreu</u>. (dieser Schätzer ist besser als der vorherige, da er eine kleinere Varianz hat.)

ML - Schätzer

Frage

Wenn wir x beobachten, welchen Wert von θ ist der wahrscheinlichste?

Aber es gibt keinen W-raum auf $\Theta \to \theta$ hat kein W-keit.

Wir können versuchen ein W-maß auf Θ zu erzeugen.

(Endlicher Fall)

Zuerst haben wir keine Information über θ $\rightsquigarrow$ wir nehmen den

Laplaceraum auf Θ.

$\Rightarrow$ $\mathbb{P}(\theta) = \frac{1}{|\Theta|}$

Aber jetzt erhalten wir Information über θ durch die Beobachtung x.

Wir können den Satz von Bayes anwenden, um die bedingte W-keit von θ zu berechnen.

Ereignisse

$$A = \{X = x\}$$

$$B_\theta = \{\text{der Parameter ist } \theta\}$$

$\{B_\theta : \theta \in \Theta\}$ ist eine Zerlegung des W-raumes.

Wir möchten die bedingte W-keit $\mathbb{P}(B_\theta \mid A)$ berechnen.

Wir wissen $\mathbb{P}(A \mid B_\theta) = \mathbb{P}_\theta(X = x)$.

Bayes

$$\mathbb{P}(B_{\theta^*} \mid A) = \frac{\mathbb{P}(A \mid B_{\theta^*})\, \mathbb{P}(B_{\theta^*})}{\sum_{\theta \in \Theta} \mathbb{P}(A \mid B_\theta)\, \mathbb{P}(B_\theta)}$$

$$= \frac{\mathbb{P}(A \mid B_{\theta^*}) \cdot \frac{1}{|\Theta|}}{\mathbb{P}(A)}$$

$$\mathbb{P}(B_{\theta^*} \mid A) = \frac{\mathbb{P}_{\theta^*}(X = x)}{|\Theta|\, \mathbb{P}(A)}$$

Deshalb wird $\mathbb{P}(B_\theta \mid A)$ maximiert, wenn $\mathbb{P}_\theta(X = x)$ maximiert wird

---

Definition (Likelihood-Funktion)

Sei $\{(\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta\}$ ein parametrisches Modell.

Die Likelihood-Funktion ist die Abbildung $\rho : \mathcal{X} \times \Theta \to [0, \infty)$, wobei

$$\rho(x, \theta) = \begin{cases} \mathbb{P}_\theta(X = x) & \text{, wenn } \mathbb{P}_\theta \text{ diskret ist} \\ f_\theta(x) & \text{, wenn } \mathbb{P}_\theta \text{ stetig ist mit Dichtefunktion } f_\theta \end{cases}$$

Definition (ML-Schätzer)

Der Schätzer $\hat{\Theta} : \mathcal{X} \to \Theta$ von θ ist der Maximum-Likelihood-Schätzer (kurz ML-Schätzer), wenn er die Likelihoodfunktion maximisiest:

$$\rho(x, \hat{\theta}(x)) = \max_{\theta \in \Theta} \rho(x, \theta) .$$

Wiederholung

Definition (Likelihood-Funktion)

Sei $\{(\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta\}$ ein parametrisches Modell.

Die <u>Likelihood-Funktion</u> ist die Abbildung $\rho : \mathcal{X} \times \Theta \to [0, \infty)$,

die durch

$$\rho(x, \Theta) = \begin{cases} \mathbb{P}_\theta(X = x) & \text{, wenn } \mathbb{P}_\theta \text{ diskret ist,} \\ f_\theta(x) & \text{, wenn } \mathbb{P}_\theta \text{ stetig mit Dichtefunktion } f_\theta \text{ ist} \end{cases}$$

definiert wird.

Definition (ML-Schätzer)

Sei $\{(\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta\}$ ein Modell, und sei ρ seine Likelihood-Funktion.

Der Schätzer $\hat{\Theta} : \mathcal{X} \to \Theta$ von θ ist der <u>Maximum-Likelihood-Schätzer</u>

(kurz <u>ML-Schätzer</u>), wenn

$$\rho(x, \hat{\theta}(x)) = \max_{\theta \in \Theta} \rho(x, \Theta) .$$

Beispiel ① (Gleichverteilung)

Hier haben wir : $\Theta = \{\theta : \theta \in \Theta\}$

$\mathbb{P}_\theta$ = Gleichverteilung auf dem Intervall $[0, \theta]$.

Wir machen n Stichproben

→ <u>Beobachtung</u> : $(x_1, x_2, \dots, x_n) \in [0, \theta]^n$

Frage

Was ist der ML-Schätzer $\hat{\theta}(x_1, x_2, \dots, x_n)$?

Zuerst müssen wir die Likelihood-Funktion berechnen.

Die Dichtefunktion auf $[0, \Theta]$ ist

$$f(x) = \begin{cases} \frac{1}{\theta} & \text{, wenn } 0 \le x \le \theta \\ 0 & \text{, sonst.} \end{cases}$$

Wenn wir n unabhängige Stichproben haben, ist die Likelihoodfunktion:

$$p((x_1, x_2, \ldots, x_n), \theta) = \begin{cases} \frac{1}{\theta^n} & \text{, wenn } x_i \leq \theta \ \forall i \\ 0 & \text{, sonst} \end{cases}$$

$\hat{\theta}(x_1, x_2, \ldots, x_n)$ ist der Wert von θ, der die Funktion maximiert.

Wir brauchen $\theta \geq x_i \quad \forall i$.

Das heißt, $\theta \geq \max\limits_{1 \leq i \leq n} x_i$.

Dann ist $p((x_1, x_2, \ldots, x_n), \theta) = \frac{1}{\theta^n}$, eine monoton fallende Funktion.

Um sie zu maximieren, nehmen wir θ so klein wie möglich.

$$\Rightarrow \quad \hat{\theta}(x_1, x_2, \ldots, x_n) = \max_{1 \leq i \leq n} x_i .$$

Bemerkung

Wir haben gesehen, dass dieser Schätzer nicht erwartungstreu ist.

Beispiel 2 (Orangenbeispiel)

Hier ist

$$\mathbb{P}_\theta = h_{N,\theta,n} .$$

Wir haben eine Beobachtung x.

Was ist der ML-Schätzer $\hat{\theta}(x)$?

Die Likelihood-Funktion:

$$p(x, \theta) = \mathbb{P}_\theta(X = x) = h_{N,\theta,n}(x)$$

$$= \frac{\binom{\theta}{x}\binom{N-\theta}{n-x}}{\binom{N}{n}}$$

Wir berechnen den Bruch

$$\frac{p(x, \theta)}{p(x, \theta - 1)} .$$

$$\frac{p(x,\theta)}{p(x,\theta-1)} = \frac{\binom{\theta}{x}\binom{N-\theta}{n-x}}{\binom{\theta-1}{x}\binom{N-(\theta-1)}{n-x}}$$

$$= \frac{\binom{\theta}{x}}{\binom{\theta-1}{x}} \cdot \frac{\binom{N-\theta}{n-x}}{\binom{N-\theta+1}{n-x}} \qquad \left[\binom{a}{b} = \frac{a!}{b!(a-b)!} \right]$$

$$= \frac{\theta}{\theta-x} \cdot \frac{N-\theta-(n-x)+1}{N-\theta+1} \geq 1$$

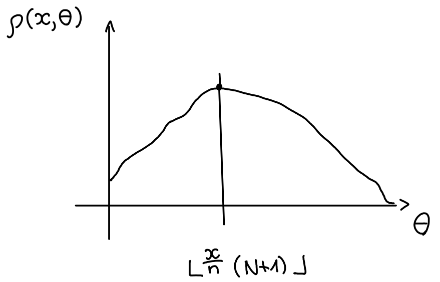
$$\Longleftrightarrow \quad \theta\left(N-\theta-(n-x)+1\right) \geq (\theta-x)(N-\theta+1)$$

$$\Longleftrightarrow \quad \theta N - \theta^2 - n\theta + x\theta + \theta$$

$$\geq \theta N - \theta^2 + \theta - xN + x\theta - x$$

$$\Longleftrightarrow \quad x(N+1) \geq \theta n$$

$$\Longleftrightarrow \quad \theta \leq \left\lfloor \frac{x}{n}(N+1) \right\rfloor$$



Deshalb wird die Likelihood-Funktion maximiert, wenn

$$\theta = \left\lfloor \frac{x}{n}(N+1) \right\rfloor .$$

$$\Rightarrow \quad \hat{\theta}(x) = \left\lfloor \frac{x}{n}(N+1) \right\rfloor$$

<u>Bemerkung</u>

$$\hat{\theta}(x) = \left\lfloor \frac{x}{n}(N+1) \right\rfloor \approx \frac{x}{n} N = T(x)$$

<u>Beispiel 3</u> (Exponentialverteilung)

<u>Modell</u>

$\{ (\mathbb{R}^n_{\geq 0}, \mathcal{B}^n, \mathbb{P}_\theta) : \theta \in \mathbb{R}^n_{\geq 0} \}$, wobei

$\mathbb{P}_\theta$ die Exponentialverteilung zum Parameter θ ist.

Wir beobachten n Wartezeiten.

Was ist der ML-Schätzer $\hat{\theta}(x_1, \dots, x_n)$?

Die Likelihood-Funktion:

$$p((x_1, x_2, \dots, x_n), \theta) = \prod_{i=1}^{n} f_\theta(x_i), \text{ wobei}$$

$$f_\theta(x) = \begin{cases} \theta e^{-\theta x} & \text{, wenn } x \geq 0 \text{ ist} \\ 0 & \text{, sonst.} \end{cases}$$

$$\Rightarrow p((x_1, x_2, \dots, x_n), \theta) = \prod_{i=1}^{n} (\theta \cdot e^{-\theta x_i})$$
$$= \theta^n e^{-\theta \sum_{i=1}^{n} x_i} .$$

Trick

Sei $f : \Theta \to \mathbb{R}_{\geq 0}$ eine positive Funktion.

θ maximiert f genau dann wenn θ die Funktion $\log(f_\theta)$ maximiert.

Wir können den $\log(p((x_1, \dots, x_n), \theta))$ maximieren:

$$\log(p) = n \cdot \log\theta - \theta \sum_{i=1}^{n} x_i \qquad [\log = \log_e = \ln]$$
$$= g(\theta).$$

Um diese Funktion zu maximieren, nehmen wir die Ableitung.

$$\frac{dg}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

$$\frac{dg}{d\theta} = 0 \Leftrightarrow \frac{n}{\theta} - \sum_{i=1}^{n} x_i = 0$$

$$\Leftrightarrow \theta = \frac{n}{\sum_{i=1}^{n} x_i}$$

$$\Rightarrow \hat{\theta}(x_1, x_2, \dots, x_n) = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}} .$$

Konfidenzintervalle

Definition (Konfidenzintervall)

Sei $\{(\mathcal{X}, \mathcal{E}, \mathbb{P}_\theta) : \theta \in \Theta\}$ ein parametrisches Modell, wobei $\Theta \subseteq \mathbb{R}$, und sei $\alpha \in (0,1)$ eine reelle Zahl.

Die Abbildung

$$C : \mathcal{X} \to \mathcal{P}(\mathbb{R}),$$

$$x \mapsto C(x) \qquad \text{Intervall} \subset \mathbb{R},$$

heißt <u>Konfidenzabbildung zum Irrtumsniveau α</u>, falls

$$\mathbb{P}_\theta(\theta \in C(x)) = \mathbb{P}_\theta(\{x : \theta \in C(x)\}) \geq 1 - \alpha \quad \text{gilt} \quad \forall \theta.$$

Bemerkung

Nur X ist zufällig!

Aber die W-keit hängt von θ ab, und wir wissen den Wert von θ nicht.

Wie können wir C erzeugen?

Vorhergehensweise:

(1) Zu $\theta \in \Theta$ bestimmen wir ein (möglichst kleines) $C_\theta \subseteq \mathcal{X}$, sodass

$$\mathbb{P}_\theta(X \in C_\theta) \geq 1 - \alpha.$$

[Hier ist θ festgelegt, deshalb wissen wir die Verteilung $\mathbb{P}_\theta$.]

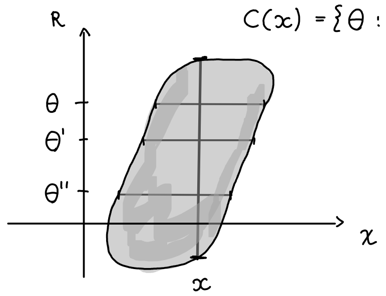
(2) Wir definieren

$$C(x) = \{\theta : x \in C_\theta\} \subseteq \Theta$$

$\forall \theta$,

$$\mathbb{P}_\theta(\theta \in C(X)) = \mathbb{P}_\theta(\{x : \theta \in C(x)\}) \qquad [\text{per Definition von } C(x)]$$

$$= \mathbb{P}_\theta(\{x : x \in C_\theta\}) \geq 1 - \alpha.$$



2

$C(x) = \{\theta : x \in C_\theta\}$ · vertikaler Schnitt.

1

$C_\theta : \mathbb{P}_\theta(X \in C_\theta) \geq 1-\alpha$

$C_{\theta'}$ ...

Beispiel (Binomialverteilung)

$$((\{0,1,\dots,n\}, \mathcal{P}(\{0,1,\dots,n\}), b_{n,\theta}) : \theta \in [0,1] \}$$

Für jedes θ brauchen wir eine Menge C_θ, sodass

$$\mathbb{P}_\theta(X \notin C_\theta) \leq \alpha,$$

wobei X binomialverteilt ist.

Wir wenden die Tschebyscheff - Ungleichung an:

$$C_\theta = (\mathbb{E}_\theta[X] - \varepsilon, \mathbb{E}_\theta[X] + \varepsilon)$$

$$\Rightarrow \quad \mathbb{P}_\theta(X \notin C_\theta) = \mathbb{P}_\theta(|X - \mathbb{E}_\theta[X]| \geq \varepsilon)$$

$$(T-U) \quad \leq \frac{V_\theta(X)}{\varepsilon^2} .$$

Binomialverteilung

$$\mathbb{E}_\theta[X] = n \cdot \theta$$

$$V_\theta(X) = n \cdot \theta \cdot (1-\theta)$$

1 Schritt

$\left[\begin{array}{l}\theta(1-\theta) \leq \frac{1}{4} \\ \text{wenn } 0 \leq \theta \leq 1\end{array}\right]$

$$\leq n \cdot \frac{1}{4}$$

Also,

$$\mathbb{P}_\theta(|X - n\theta| \geq \varepsilon) \leq \frac{n}{4\varepsilon^2}$$

Es ist genug, wenn

$$\frac{n}{4\varepsilon^2} \leq \alpha$$

$$\Leftrightarrow \quad \varepsilon^2 \geq \frac{n}{4\alpha} \quad \text{oder} \quad \varepsilon \geq \frac{\sqrt{n}}{2\sqrt{\alpha}}$$

$$\rightarrow C_\theta = (n\theta - \varepsilon, n\theta + \varepsilon)$$

Wir wollen, dass ε so klein wie möglich ist.

$$\underline{\varepsilon = \frac{\sqrt{n}}{2\sqrt{\alpha}}} \quad .$$

(2) Das Konfidenzintervall:

$$\begin{aligned}
C(x) &= \{\theta : x \in C_\theta\} \\
&= \{\theta : x \in (n\theta - \varepsilon, n\theta + \varepsilon)\} \\
&= \{\theta : |x - n\theta| < \varepsilon\} \\
&= \{\theta : |\tfrac{x}{n} - \theta| < \tfrac{\varepsilon}{n}\} \\
&= \left(\frac{x}{n} - \frac{\varepsilon}{n}, \frac{x}{n} + \frac{\varepsilon}{n}\right) \\
&= \left(\frac{x}{n} - \frac{1}{2\sqrt{\alpha n}}, \frac{x}{n} + \frac{1}{2\sqrt{\alpha n}}\right)
\end{aligned}$$

<u>Beispiel zum Beispiel</u>:

$\alpha = 0{,}025$

Radius $= 0{,}1$

$\rightarrow$ $n = \underline{1.000}$

Radius $= 0{,}01$

$\rightarrow$ $n = \underline{100.000}$