SS 2013

Diskrete Wahrscheinlichkeitstheorie

Javier Esparza

Fakultät für Informatik
TU München

http://www7.in.tum.de/um/courses/dwt/ss13

Sommersemester 2013

Vorlesungsinhalt

- Diskrete Wahrscheinlichkeitsräume
  - Ereignisse, Zufallsvariablen, Unabhängigkeit
  - Binomial-, geometrische-, Poisson-Verteilung
  - Ungleichungen von Markov und Chebyshev, Chernoff-Schranken
- Kontinuierliche Wahrscheinlichkeitsräume
  - Normalverteilung, Exponentialverteilung
  - Zentraler Grenzwertsatz
- Induktive Statistik
  - Schätzvariablen
  - Konfidenzintervalle
  - Testen von Hypothesen
- Markov-Ketten
  - Ankunftsw'keiten, stationäre Verteilung
  - Warteschlangen

Literatur I

- N. Henze:
  Stochastik für Einsteiger, 9. Auflage
  Vieweg+Teubner, 2012
- T. Schickinger, A. Steger:
  Diskrete Strukturen - Band 2,
  Springer Verlag 2001
- M. Greiner, G. Tinhofer:
  Stochastik für Informatiker,
  Carl Hanser Verlag, 1996
- H. Gordon:
  Discrete Probability,
  Springer-Verlag, 1997

Literatur II

- R. Motwani, P. Raghavan:
  Randomized Algorithms,
  Cambridge University Press, 1995
- L. Fahrmeir, R. Künstler, I. Pigeot, G. Tutz:
  Statistik - Der Weg zur Datenanalyse,
  Springer-Verlag, 1997

Was bedeutet Zufall?

- Der Heilige Augustinus über die Zeit:

  > *Was also ist die Zeit? Wenn niemand mich danach fragt, weiß ich es; wenn ich es jemand auf seine Frage hin erklären will, weiß ich es nicht.*

- Pragmatische Sicht: **Zufall = mangelndes Wissen**.
  Bei gewissen Vorgängen wissen wir für eine sichere Vorhersage nicht genug: Ziehung von Lottozahlen, Regenfall über München, Absturz eines informatischen Systems.
- Es können trotzdem probabilistische Vorhersagen gemacht werden.
- Ziel der Vorlesung: lernen, korrekte Vorhersagen dieser Art zu machen und zu identifizieren.

Es wird gewürfelt . . . I

Am Ende der Vorlesung werden Sie in der Lage sein, folgende Fragen zu beantworten:

(1) Ein fairer Würfel wird 10 Mal geworfen. Mit welcher W'keit kriegt man
   - keine Sechs? (≈ 0.162)
   - genau 1 Mal eine Sechs? (≈ 0.323)
   - genau 3 Mal eine Sechs? (≈ 0.155)

(2) Ein fairer Würfel wird 600000 Mal geworfen. Mit welcher W'keit kriegt man mindestens 100100 Sechser? (≈ 0.36)

(3) Drei faire Würfel werden gleichzeitig geworfen. Mit welcher W'keit
   - beträgt die Summe der Augenzahlen mindestens 14? ($35/216$)
   - ist die größte Augenzahl höchstens 3? ($1/8$)
   - sind alle drei Augenzahlen verschieden? ($5/9$)

Es wird gewürfelt ... II

(4) Sechs faire Würfel werden mit den Zahlen 1 bis 6 nummeriert und dann gleichzeitig geworfen. Wieviele Würfel fallen im Mittel mit der eigenen Nummer als Augenzahl? (1.2)

(5) Hundert faire Würfel werden gleichzeitig geworfen. Mit welcher W'keit beträgt die Summe der Augenzahlen höchstens 310? (≈ 0.01)

(6) Drei faire Würfel werden 1000 Mal gleichzeitig geworfen. Mit welcher W'keit kriegt man kein einziges Mal drei Sechser? (≈ 0.01)

(7) 10 faire Würfel werden mit den Zahlen 1 bis 10 numeriert und dann 10000 Mal gleichzeitig geworfen. Das Ergebnis eines Wurfes wird als Vektor $(a_1, \ldots, a_{10})$ mit $1 \leq a_i \leq 6$ dargestellt. Mit welcher W'keit wiederholt sich kein Ergebnis? (≈ 0.437)

Es wird gewürfelt . . . III

(8) Wie oft muss man einen fairen Würfel im Mittel werfen, um

- die erste Sechs
- drei Sechser hintereinander
- alle sechs Augenzahlen mindestens einmal

zu kriegen? ($6, 258, 14.7$)

(9) Anna versucht, das Ergebnis eine Reihe von Würfen eines fairen Würfels zu raten. Der Würfel wird 600 Mal geworfen. Wie oft muss Anna das richtige Ergebnis raten, so dass die W'keit, das es aus Zufall geschieht, höchstens 1% beträgt? (150)

(10) Ein Würfel wird 15 Mal geworfen und man bekommt 1 Mal die eins, 3 Mal die 2, 4 Mal die 3, 3 Mal die 4, 4 Mal die 5 und 0 Mal die 6. Mit welcher Konfidenz kann behauptet werden, der Würfel sei nicht fair? (0.95)

Es wird gewürfelt . . . IV

Zwei der Ergebnisse in (1)-(10) sind falsch.
Sie können schon heute diese zwei Ergebnisse experimentell bestimmen.

Zufall in der Informatik

- Hardware-Fehler können durch Zufallsvorgänge (z.B. Strahlung) eintreten.

  Zuverlässigkeitsanalyse

- Download-Zeiten von Web-Seiten, Antwort-Zeiten von Diensten können nicht präzise vorhergesagt werden.

  Leistungsanalyse

- Viele Programme verwenden Zufallsbits, um mit großer Wahrscheinlichkeit gewisse Problemsituationen (z.B. Verklemmungen) zu vermeiden:

  Randomisierte Algorithmen

- Viele Sicherheitsprotokolle beruhen auf Annahmen über die W'keit, dass ein Angreifer Information über Schlüsseln gewinnt:

  Sicherheitsanalyse

Thema der Vorlesung: Zufallsexperimente

Das Experiment wird unter genau festgelegten Versuchsbedingungen durchgeführt.

- Ungenau: zwei Zahlen werden zufällig gewählt.
- Genau: ein Paar $(x, y) \in [0, 2^{32} - 1] \times [0, 2^{32} - 1]$ wird zufällig gewählt, jedes Paar hat die gleiche W'keit.

Die Menge der möglichen Ergebnisse (Ausgänge) ist vor der Durchführung des Experimentes bekannt.

Das Experiment kann zumindest prinzipiell beliebig oft unter gleichen Bedingungen wiederholt werden.

Mathematische Modellierung von Zufallsexperimenten

Drei faire Würfel werden geworfen.
Frage: Was ist die W'keit, dass die Summe der Augenzahlen größer als 12 ist?

- Die Menge der möglichen Ausgänge wird aus der Beschreibung gewonnen.
  Ausgänge: $(1,1,1),\ldots,(6,6,6)$
- Die W'keiten der Ausgänge werden ebenfalls aus der Beschreibung gewonnen.
  $\Pr[(1,1,1)] = \ldots = \Pr[(6,6,6)] = \dfrac{1}{216}$
- Die Frage, die uns interessiert, wird auf die Berechnung der W'keit einer Menge von möglichen Ausgängen reduziert.
  $\Pr[\{(5,4,4),(4,5,5),\ldots,(6,6,6)\}]$
- Zwei Fehlerquellen: falsche Modelle und falsche Berechnungen!

Mathematische Modellierung von Zufallsexperimenten II

Die Reduktion der Frage auf W'keiten ist nicht immer so direkt.

Drei faire Würfel werden geworfen.
Frage: Sollte man folgende Wette eingehen? Wenn insgesamt über 12 Augenzahlen angezeigt werden, gewinnt man 1 Euro. Sonst verliert man 2 Euro.

Teil I

Grundbegriffe

1. Diskrete Wahrscheinlichkeitsräume

Diskrete Wahrscheinlichkeitskeitsräume

Definition 1

1. Ein diskreter W'keitsraum ist durch eine (endliche oder abzählbar unendliche) Ergebnismenge $\Omega = \{\omega_1, \omega_2, \ldots\}$ von Elementarereignissen gegeben.
2. Jedem Elementarereignis ω_i ist eine Wahrscheinlichkeit $\Pr[\omega_i]$ zugeordnet, wobei wir fordern, dass $0 \le \Pr[\omega_i] \le 1$ und
$$\sum_{\omega \in \Omega} \Pr[\omega] = 1.$$
3. Eine Menge $E \subseteq \Omega$ heißt Ereignis. Die Wahrscheinlichkeit $\Pr[E]$ eines Ereignisses ist definiert durch
$$\Pr[E] := \sum_{\omega \in E} \Pr[\omega].$$

Modelle I

Beispiel 2

Zwei faire Würfel werden geworfen.

Frage: Mit welcher W'keit beträgt die Gesamtzahl der Augen mindestens 10?

Modell 1: $\Omega = \{\ (1,1), \ldots, (6,6)\ \}$, $\Pr[(i,j)] = \dfrac{1}{36}$.

mind-10 $= \{(4,6), (5,5), \ldots, (6,6)\}$, $\Pr[\text{mind-10}] = \dfrac{1}{6}$.

Modell 2: $\Omega = \{2, 3, 4, \ldots, 10, 11, 12\}$,
$\Pr[2] = \dfrac{1}{36}, \ldots, \Pr[7] = \dfrac{1}{6}, \ldots, \Pr[12] = \dfrac{1}{36}$.

mind-10 $= \{10, 11, 12\}$, $\Pr[\text{mind-10}] = \dfrac{3}{26} + \dfrac{2}{36} + \dfrac{1}{36} = \dfrac{1}{6}$.

Modelle II

Modell 3: $\Omega = \{2, 3, 4, \ldots, 10, 11, 12\}$, $\Pr(i) = \frac{1}{11}$.

$\Pr[\text{mind-10}] = \frac{3}{11}$.

Model 3 ist aus mathematischer Sicht zulässig, jedoch falsch. Es entspricht nicht dem, was man unter “faire Würfel” versteht!

Physikalische Analogie

Elementarereignis ω	$\to$	Objekt
W'keit $\Pr[\omega]$	$\to$	Gewicht des Objekts
$\sum_{\omega \in \Omega} \Pr[\omega] = 1$	$\to$	Gesamtgewicht aller Objekte: 1 Kg.
Ereignis E	$\to$	Objektmenge
$\Pr[E] := \sum_{\omega \in E} \Pr[\omega]$	$\to$	Gesamtgewicht der Objektmenge

Vereinigung, Schnitt, Komplement von Ereignissen

Wenn A und B Ereignisse sind, dann sind auch $A \cup B$, $A \cap B$, $A \setminus B$ etc. Ereignisse.

Das Ereignis $\bar{E} = \Omega \setminus E$ heißt komplementäres Ereignis zu E oder komplement von E.

Zwei Ereignisse A und B heißen disjunkt wenn $A \cap B = \emptyset$ gilt.

Einfache Eigenschaften I

Die folgenden Eigenschaften sind einfache Konsequenzen des Axioms $\sum_{\omega\in\Omega} \Pr[\omega] = 1$ und der Definition $\Pr[E] := \sum_{\omega\in E} \Pr[\omega]$.

Für Ereignisse $A, B, A_1, A_2 \ldots$ gilt:

- $\Pr[\emptyset] = 0$, $\Pr[\Omega] = 1$.
- $0 \leq \Pr[A] \leq 1$.
- $\Pr[\bar{A}] = 1 - \Pr[A]$.
- Wenn $A \subseteq B$, so folgt $\Pr[A] \leq \Pr[B]$.
- Wenn $A \cap B = \emptyset$, so folgt $\Pr[A \cup B] = \Pr[A] + \Pr[B]$.
- Allgemein: $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$.

  (Intuitiv: $\Pr[A] + \Pr[B]$ zählt das "Gewicht" der Objekte in $A \cap B$ zweimal. Korrektur: $\Pr[A \cap B]$ abziehen.)

Einfache Eigenschaften II

- (Additionssatz) Wenn $A_1, \ldots, A_n$ paarweise disjunkt sind (also wenn für alle Paare $i \neq j$ gilt, dass $A_i \cap A_j = \emptyset$), so folgt
$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \Pr[A_i].$$

  Für eine unendliche Menge von paarweise disjunkten Ereignissen $A_1, A_2, \ldots$ gilt analog
$$\Pr\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \Pr[A_i]\ .$$

Einfache Eigenschaften III

- (Boole'sche Ungleichung) Für beliebige $A_1, \ldots, A_n$ gilt

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \Pr[A_i]\,.$$

Analog gilt für eine unendliche Folge $A_1, A_2, \ldots$, dass

$$\Pr\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} \Pr[A_i]\,.$$

Die Siebformel I

Der Additionssatz gilt nur für paarweise disjunkte Ereignisse. Für den allgemeinen Fall erhalten wir folgenden

Satz 3 (Siebformel, Prinzip der Inklusion/Exklusion)

Für Ereignisse $A_1, \ldots, A_n$ *(*$n \geq 2$*) gilt:*

$$\begin{aligned}
\Pr\left[\bigcup_{i=1}^{n} A_i\right] &= \sum_{i=1}^{n} \Pr[A_i] - \sum_{1 \leq i_1 < i_2 \leq n} \Pr[A_{i_1} \cap A_{i_2}] + - \ldots \\
&+ (-1)^{l-1} \sum_{1 \leq i_1 < \ldots < i_l \leq n} \Pr[A_{i_1} \cap \ldots \cap A_{i_l}] + - \ldots \\
&+ (-1)^{n-1} \cdot \Pr[A_1 \cap \ldots \cap A_n]\,.
\end{aligned}$$

Die Siebformel II

Insbesondere gilt für zwei Ereignisse A und B

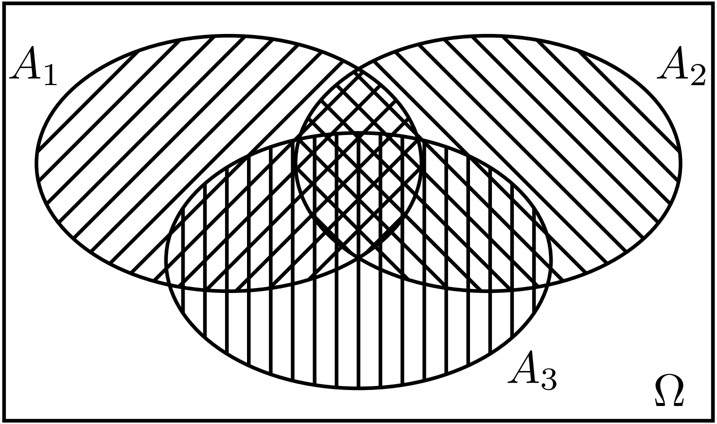
$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] \ .$$

Für drei Ereignisse A_1, A_2 und A_3 erhalten wir

$$\begin{aligned}\Pr[A_1 \cup A_2 \cup A_3] \quad = \quad & \Pr[A_1] + \Pr[A_2] + \Pr[A_3] \\ & -\Pr[A_1 \cap A_2] - \Pr[A_1 \cap A_3] \\ & -\Pr[A_2 \cap A_3] \\ & +\Pr[A_1 \cap A_2 \cap A_3] \ .\end{aligned}$$

Die Siebformel III

Beweis: Nur für $n = 3$:



Durch die im Satz angegebene Summe wird jedes Flächenstück genau einmal gezählt.

Laplacesche W'keitsräume

Definition 4

Ein diskreter W'keitsraum $(\Omega, \Pr)$ heißt Laplacescher W'keitsräum, wenn $\Pr[\omega] = \Pr[\omega']$ für alle $\omega, \omega' \in \Omega$.

Sei $(\Omega, \Pr)$ ein Laplacescher W'keitsräum. Dann:

- Ω ist endlich.
- Für jedes Ereignis E gilt $\Pr[E] = \dfrac{|E|}{|\Omega|}$

Bemerkung: Unwissen über die W'keiten bedeutet nicht, dass ein W'keitsraum ein Laplacescher W'keitsraum sein muss!

Wenn nach unserem Kenntnis der Gesetze (physikalische, psychologische, soziologische ...), die das Ergebnis des Experiments bestimmen, kein Elementarereignis "bevorzugt" wird, dann kann die Laplace-Annahme sinnvoll sein.

Experimentelle Wahl der W'keiten

Definition 5

$$\begin{aligned}\text{relative Häufigkeit von } E \quad &:= \quad \frac{\text{absolute Häufigkeit von } E}{\text{Anzahl aller Beobachtungen}} \\ &= \quad \frac{\text{Anzahl Eintreten von } E}{\text{Anzahl aller Beobachtungen}}.\end{aligned}$$

Um die W'keiten zu wählen: Wiederhole das Experiment und wähle die relativen Häufigkeiten.

Basiert auf das Prinzip („Gesetz der großen Zahlen“):

> *Wenn man das gleiche Zufallsexperiment sehr oft wiederholt, konvergieren die relativen Häufigkeiten der Ereignisse gegen deren Wahrscheinlichkeiten.*

Historische Anfänge der Wahrscheinlichkeitstheorie

Die ersten Hinweise auf mathematische Untersuchungen zu Problemen der Wahrscheinlichkeitstheorie finden sich in einem Briefwechsel zwischen den französischen Mathematikern

Pierre Fermat (1601–1665)

und

Blaise Pascal (1623–1662).

Pascal beschäftigte sich neben der Mathematik auch mit Fragestellungen aus dem Bereich der Physik und auch aus der Informatik! Sein Vater hatte als Steuerinspektor in Rouen umfangreiche Rechnungen durchzuführen und so wurde Pascal zum Bau einer mechanischen Rechenmaschine, der so genannten *Pascaline*, motiviert.

In dem Briefwechsel taucht bereits der Ansatz $\Pr[E] = |E|/|\Omega|$ zur Berechnung der Wahrscheinlichkeit von E auf. Auch den Begriff des Erwartungswerts kann man dort schon finden. Weder Fermat noch Pascal publizierten ihre Überlegungen zur Wahrscheinlichkeitstheorie. Der Niederländer

Christiaan Huygens (1629–1695)

entwickelte ebenfalls Methoden zum Arbeiten mit Wahrscheinlichkeiten aus. Er publizierte im Jahre 1657 auch eine kleine Arbeit mit dem Titel „De ratiociniis in ludo aleae“ (Über die Gesetzmäßigkeiten beim Würfelspiel).

2. Mehrstufige Experimente

Mehrstufige Experimente I

Experimente bestehen oft aus Teilexperimenten, die der Reihe nach durchgeführt werden.

Welches Teilexperiment als nächstes durchgeführt wird, kann vom Ausgang des vorigen Teils abhängen.

Elementarereignisse = mögliche Sequenzen von “Teilereignissen”.

Die W’keit eines Elementarereignisses ist das Produkt der W’keiten der Teilereignisse.

Beispiel 6

Eine faire Münze wird geworfen. Zeigt sie Kopf, werden zwei faire Würfel geworfen. Zeigt sie Zahl, wird nur ein fairer Würfel geworfen.

Frage: Was ist die W'keit, höchstens 4 Augen zu bekommen?

Modell:

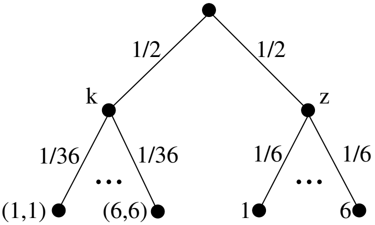
$$\Omega = \{\mathsf{k}\,(1,1),\ldots,\mathsf{k}\,(6,6),\mathsf{z}\,1,\ldots,\mathsf{z}\,6\}$$

$$\Pr[\mathsf{k}\,(i,j)] = \frac{1}{2}\cdot\frac{1}{36} \qquad \Pr[\mathsf{z}\,i] = \frac{1}{2}\cdot\frac{1}{6}$$

W'keit, höchstens 4 Augen zu bekommen: $\dfrac{1}{2}\cdot\dfrac{6}{36}+\dfrac{1}{2}\cdot\dfrac{4}{6}=\dfrac{5}{12}$

Mehrstufige Experimente III

Mehrstufige Experimente können als Bäume visualisiert werden. Die Elementarereignisse sind die Pfade des Baumes.



Jeder Knoten repräsentiert ein Ereignis (die Menge der Pfade, die den Knoten besuchen).

Die W'keit dieses Ereignisses ist das Produkt der W'keiten im Pfad von der Würzel zum Knoten.

Mehrstufige Experimente IV

Beispiel 7

Eine faire Münze wird so lange geworfen, bis die gleiche Seite zweimal hintereinander fällt.

Frage: Mit welcher W'keit ist die Anzahl der Würfe gerade?

Modell:

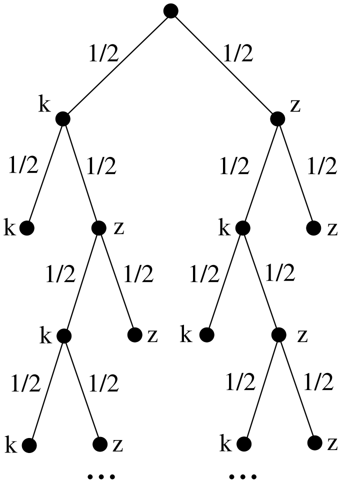
$$\Omega = \{\text{kk, zz, kzz, zkk, kzkk, zkzz} \ldots\} \qquad \Pr[\omega] = 2^{-|\omega|}$$

Es gilt:

$$\Pr[\text{gerade}] = \Pr[\text{kk}] + \Pr[\text{zz}] + \Pr[\text{kzkk}] + \ldots$$

Mehrstufige Experimente V

Visualisierung als Baum:



Mehrstufige Experimente VI

"Abstrakteres" Modell: die Elementarereignisse sind die Anzahlen der Würfe.

$$\Omega = \{2, 3, 4 \ldots\}\,.$$

W'keiten der Elementarereignisse:

$$\begin{aligned}
\Pr[2i] &= \Pr[\,(\mathsf{kz})^{(i-1)}\mathsf{kk}\,] + \Pr[\,(\mathsf{zk})^{(i-1)}\mathsf{zz}\,] \\
&= \left(\frac{1}{2}\right)^{2i} + \left(\frac{1}{2}\right)^{2i} = \left(\frac{1}{2}\right)^{2i-1} \\
\Pr[\mathsf{gerade}] &= \sum_{i=1}^{\infty} \Pr[2i] = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^{2i-1} = \frac{1}{2} \sum_{i=0}^{\infty} \left(\frac{1}{4}\right)^{i} \\
&= \frac{1}{2} \cdot \frac{1}{3/4} = \frac{2}{3}
\end{aligned}$$

Mehrstufige Experimente VII

Beispiel 8

Sie nehmen an einer Spielshow im Fernsehen teil, bei der Sie eine von drei verschlossenen Türen auswählen sollen. Hinter einer Tür wartet der Preis, ein Auto, hinter den beiden anderen stehen Ziegen. Sie zeigen auf eine Tür. Sie bleibt vorerst geschlossen. Der Moderator weiß, hinter welcher Tür sich das Auto befindet; mit den Worten “Ich zeige Ihnen mal was” öffnet er eine andere Tür, und eine meckernde Ziege schaut ins Publikum. Er fragt: “Bleiben Sie bei Nummer eins, oder wählen sie Nummer zwei? ”

Frage: Welche Tür hat bessere Chancen?

- Nummer eins.
- Nummer zwei.
- Beide gleich.

Mehrstufige Experimente VIII

Wir vergleichen zwei Experimente: Im ersten Experiment bleiben Sie immer bei der Tür, die Sie gewählt haben. Im zweiten Experiment wechseln Sie immer die Tür.

Annahmen (genaue Beschreibung der Experimente):

- Das Auto wird mit W'keit $1/3$ hinter Tür 1, 2, oder 3 gestellt.
- Sie wählen eine Tür mit W'keit $1/3$.
- Wenn der Moderator zwei Türen aufmachen darf, dann wählt er eine Tür mit W'keit $1/2$.

Elementarereignis: (Tür-des-Autos
Ihre-erste-Wahl
Wahl-des-Moderators
Ihre-zweite-Wahl
Ergebnis)

Mehrstufige Experimente IX

- Experiment "Ich bleibe bei meiner ersten Wahl"
  Modell:

$$\Omega_b = \{\, 1\,1\,2\,1\,A\,,\; 1\,1\,3\,1\,A\,,\; 1\,2\,3\,2\,Z\,, \ldots,\; 3\,3\,2\,3\,A\,\}$$

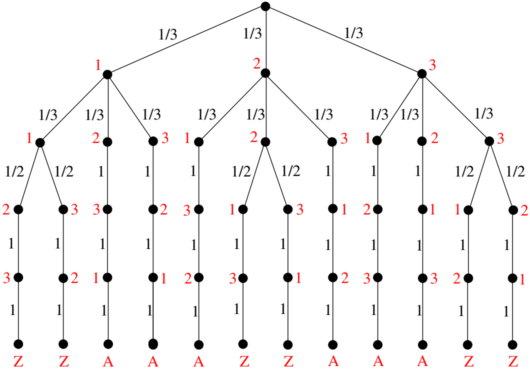
$$\Pr[1\,1\,2\,1\,A] = \frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{2} \qquad \Pr[1\,2\,3\,1\,A] = \frac{1}{3}\cdot\frac{1}{3}$$

$$\Pr[\text{Auto}] = \frac{1}{3}$$

Mehrstufige Experimente X

- Experiment "Ich wechsle"

$$\Omega_w = \{\, 1\,1\,2\,3\,Z\ ,\ 1\,1\,3\,2\,Z\ ,\ 1\,2\,3\,1\,A\ ,\ldots,\ 3\,3\,2\,1\,Z\,\}$$



$$\Pr[\text{Auto}] = \frac{2}{3}$$

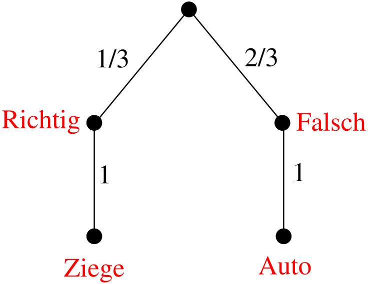
Mehrstufige Experimente XI

- Wir gewinnen ein abstrakteres Modell des Experiments “Ich wechsle” durch “Aggregation”.
- Alle Sequenzen, in denen Sie eine falsche Tür am Anfang wählen, werden aggregiert. In diesen Sequenzen gewinnen Sie immer das Auto.
- Alle Sequenzen, in denen Sie die richtige Tür am Anfang wählen, werden aggregiert. In diesen Sequenzen gewinnen Sie nie das Auto.

Mehrstufige Experimente XI

Ein viel einfacheres Modell:

$$\Omega_w = \{\mathsf{Richtig.Ziege}, \mathsf{Falsch.Auto}\}$$



$$\Pr[\mathsf{Auto}] = \Pr[\mathsf{Falsch.Auto}] = \frac{2}{3}.$$

3. Zufallsvariablen, Erwartungswert, Varianz

Zufallsvariablen I

Anstatt der Ereignisse selbst sind wir oft an „Auswirkungen“ oder „Merkmalen“ der (Elementar-)Ereignisse interessiert.

Definition 9

Sei ein Wahrscheinlichkeitsraum auf der Ergebnismenge Ω gegeben. Eine Abbildung

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

heißt (numerische) Zufallsvariable mit Wertebereich

$$W_X := X(\Omega) = \{x \in \mathbb{R} \mid \exists\, \omega \in \Omega \text{ mit } X(\omega) = x\}$$

Eine Zufallsvariable über einer endlichen oder abzählbar unendlichen Ergebnismenge Ω heißt diskret.

$\Pr[X = a]$ ist eine Abkürzung für $\Pr[\,\{\omega \in \Omega \mid\; X(\omega) = a\}\,]$.

Zufallsvariablen II: Dichte und Verteilung

Definition 10

- Die Funktion

$$\begin{array}{rcl} f_X : \mathbb{R} & \to & [0,1] \\ x & \mapsto & f_X(x) := \Pr[X = x] \end{array}$$

  nennt man (diskrete) Dichte oder Zähldichte der Zufallsvariablen X.

- Die Funktion

$$\begin{array}{rcl} F_X : \mathbb{R} & \to & [0,1] \\ x & \mapsto & F_X(x) := \displaystyle\sum_{x' \in W_X \,:\, x' \le x} \Pr[X = x'] \end{array}$$

  heißt (kumulative) Verteilung der Zufallsvariablen X.

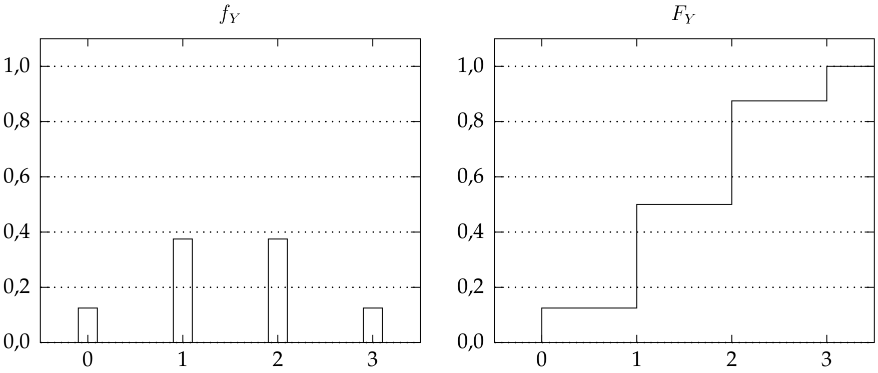
Zufallsvariablen III

Beispiel 11

Wir werfen drei faire Münzen. Sei Y die Zufallsvariable, die die Anzahl der Köpfe angibt. Für die Zufallsvariable Y haben wir

$$
\begin{aligned}
\Pr[Y=0] &= \Pr[\ \{zzz\}\ ] &&= \frac{1}{8}\\
\Pr[Y=1] &= \Pr[\ \{kzz, zkz, zzk\}\ ] &&= \frac{3}{8}\\
\Pr[Y=2] &= \Pr[\ \{kkz, kzk, zkk\}\ ] &&= \frac{3}{8}\\
\Pr[Y=3] &= \Pr[\ \{kkk\}\ ] &&= \frac{1}{8}
\end{aligned}
$$

Zufallsvariablen IV



Dichte und Verteilung von Y

Bernoulli-Verteilung I

Beispiel 12

Ein Würfel wird geworfen. Sei X die Zufallsvariable mit $W_X = \{0, 1\}$ und $X(n) = 1$ falls die Augenzahl $n = 3, 6$ ist und $X(n) = 0$ sonst. Es gilt

$$f_X(1) = 1/3 \qquad f_X(0) = 2/3$$

Definition 13

Eine Zufallssvariable X mit $W_X = \{0, 1\}$ und Dichte

$$f_X(1) = p \qquad f_X(0) = 1 - p$$

für einen Wert $0 \leq p \leq 1$ heißt Bernoulli-verteilt.
Den Parameter p nennen wir Erfolgswahrscheinlichkeit.

Bernoulli-Verteilung II

Der Name der Bernoulli-Verteilung geht zurück auf den Schweizer Mathematiker Jakob Bernoulli (1654–1705). Wie viele andere Mathematiker seiner Zeit hätte auch Bernoulli nach dem Wunsch seines Vaters ursprünglich Theologe werden sollen. Sein Werk *ars conjectandi* stellt eine der ersten Arbeiten dar, die sich mit dem Teil der Mathematik beschäftigen, den wir heute als Wahrscheinlichkeitstheorie bezeichnen.

Binomialverteilung I

Beispiel 14

Eine Münze wird 10-Mal geworfen. Die W'keit von 'Kopf' beträgt $1/3$ bei jedem Wurf.

Modell:

$$\Omega = \{k, z\}^{10} \qquad \Pr[\omega] = \left(\frac{1}{3}\right)^{\mathcal{K}(\omega)} \left(\frac{2}{3}\right)^{10-\mathcal{K}(\omega)}$$

wobei $\mathcal{K}(\omega)$ die Anzahl der Köpfe in ω darstellt.

Es gilt z.B. $\mathcal{K}(kkzzkzkzzz) = 4$ und $\mathcal{K}(z^{10}) = 0$.

Bemerkung: Alle Elementarereignisse mit derselben Anzahl von Köpfen haben die gleiche W'keit.

Binomialverteilung II

$\mathcal{K}\colon \Omega \to \mathbb{R}$ ist eine Zufallsvariable mit $W_{\mathcal{K}} = \{0, \ldots n\}$.

Es gibt $\binom{10}{x}$ Elementarereignisse in Ω mit $\mathcal{K}(\omega) = x$.

Für die Dichte von $\mathcal{K}$ gilt also

$$f_{\mathcal{K}}(x) = \binom{10}{x} \left(\frac{1}{3}\right)^{x} \left(\frac{2}{3}\right)^{10-x}$$

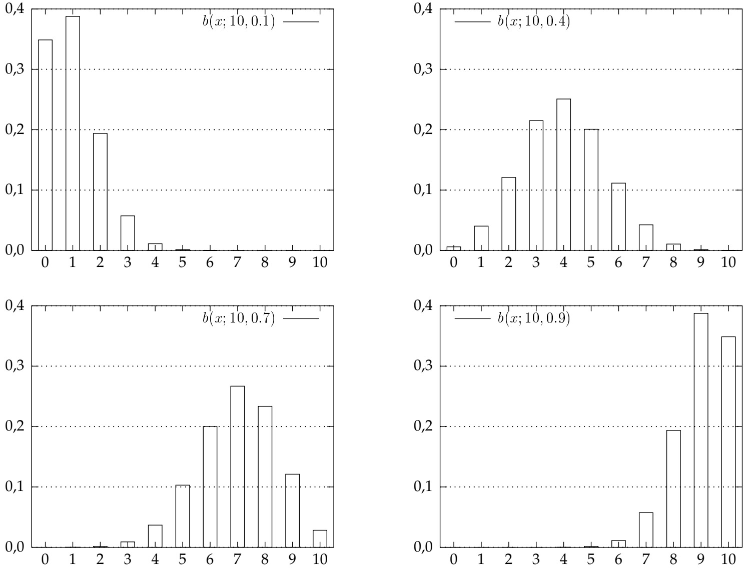
Definition 15

Eine Zufallssvariable X mit $W_X = \{0, \ldots, n\}$ und Dichte

$$f_X(x) = \binom{n}{x} p^x q^{n-x}$$

für einen Wert $0 \leq p \leq 1$ und $q := 1 - p$ heißt binomialverteilt.

Binomialverteilung III



Dichte der Binomialverteilung

Geometrische Verteilung I

Beispiel 16

Eine Münze wird geworfen bis 'Kopf' fällt. Die W'keit von Kopf beträgt $1/3$.

Modell:

$$\Omega = Z^* K \qquad \Pr[Z^n K] = \left(\frac{2}{3}\right)^n \cdot \frac{1}{3}$$

Sei $X : \Omega \to \mathbb{N}$ die Zufallsvariable, die die Anzahl der Würfe angibt, d.h. $X(Z^n K) = n + 1$.

Es gilt:

$$f_X(x) = \left(\frac{1}{3}\right)^{x-1} \cdot \frac{2}{3}$$

Geometrische Verteilung II

Definition 17

Eine Zufallssvariable X mit $W_X = \mathbb{N}$ und Dichte

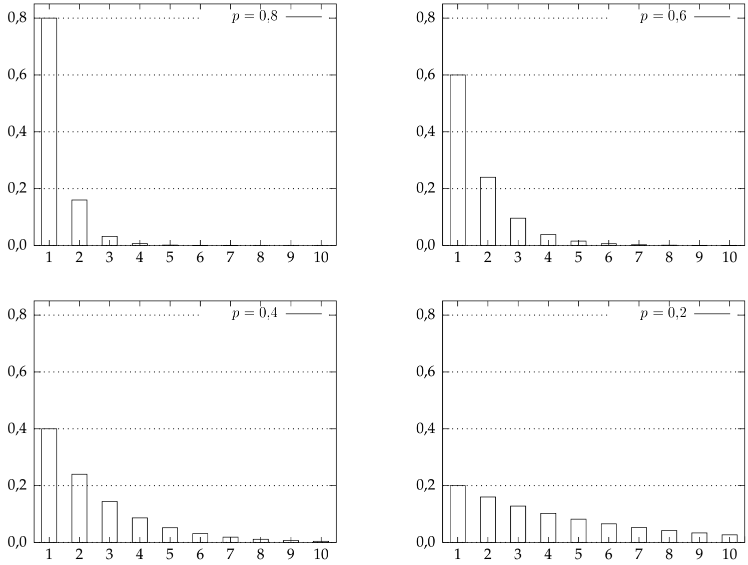
$$f_X(i) = pq^{i-1}$$

für ein $p \in [0,1]$ und $q := 1 - p$ ist geometrisch Verteilt.

Die kumulative Verteilung einer geometrisch verteilten Zufallsvariablen X mit Erfolgswahrscheinlichkeit p ist die Funktion

$$F_X(x) = \Pr[X \leq x] = \sum_{i=1}^{x} pq^{i-1} = p \cdot \sum_{i=0}^{x-1} q^i = p \cdot \left(\frac{1-q^x}{1-q} \right) = 1 - q^x$$

Geometrische Verteilung III



Dichte der geometrischen Verteilung

Mehrere Zufallsvariablen I: Gemeinsame Dichte

Seien $X, Y : \Omega \to \mathbb{R}$ zwei Zufallsvariablen über demselben W'keitsraum.

Wir definieren

$$\Pr[X = x, Y = y] := \Pr[\ \{\omega \mid X(\omega) = x \wedge Y(\omega) = y\}\ ]\,.$$

Die Funktion

$$f_{X,Y}(x, y) := \Pr[X = x, Y = y]$$

heißt gemeinsame (diskrete) Dichte der Zufallsvariablen X und Y.

Mehrere Zufallsvariablen II: Gemeinsame Dichte

Beispiel 18

Die Studierenden von Prof. Evilsparza haben zwei Klausuren geschrieben. Die möglichen Noten sind $1, 2, \ldots, 5$. Ein Studierender besteht, wenn seine Bestnote eine 1 ist oder seine Schnittnote mindestens 3 beträgt. Prof. Evilsparza hat keine Lust zu korrigieren, und verteilt die Noten zufällig.

Modell:

$$\Omega = \{(1,1), \ldots, (5,5)\} \qquad \Pr[(i,j)] = \frac{1}{25}$$

Seien B, S zwei Zufallsvariablen, die die Bestnote und die Schnittnote angeben.

Frage: Welche gemeinsame Dichte haben B und S?

Mehrere Zufallsvariablen III: Gemeinsame Dichte

Die gemeinsame Dichte $f_{B,S}$ von B und S ist:

$B \setminus S$	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
1	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	0	0	0	0
2	0	0	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	0	0	0
3	0	0	0	0	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	0	0
4	0	0	0	0	0	0	$\frac{1}{25}$	$\frac{2}{25}$	0
5	0	0	0	0	0	0	0	0	$\frac{1}{25}$

Mehrere Zufallsvariablen IV: Randdichte

Die Funktionen f_X und f_Y, definiert durch

$$f_X(x) = \sum_{y \in W_Y} f_{X,Y}(x,y) \quad \text{bzw.} \quad f_Y(y) = \sum_{x \in W_X} f_{X,Y}(x,y)$$

nennen wir Randdichten.

Die Ereignisse „$Y = y$“ bilden eine Partitionierung des Wahrscheinlichkeitsraumes, und es gilt daher

$$\Pr[X = x] = \sum_{y \in W_Y} \Pr[X = x, Y = y] = f_X(x)\,,$$

Die Dichten der einzelnen Zufallsvariablen entsprechen also genau den Randdichten.

Mehrere Zufallsvariablen V: Randdichte

Die Randdichten f_B, f_S des Beispiels sind:

$B \setminus S$	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	f_B
1	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	0	0	0	0	$\frac{9}{25}$
2	0	0	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	0	0	0	$\frac{7}{25}$
3	0	0	0	0	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	0	0	$\frac{5}{25}$
4	0	0	0	0	0	0	$\frac{1}{25}$	$\frac{2}{25}$	0	$\frac{3}{25}$
5	0	0	0	0	0	0	0	0	$\frac{1}{25}$	$\frac{1}{25}$
f_S	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{5}{25}$	$\frac{4}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{1}{25}$	1

Mehrere Zufallsvariablen VI: Randdichte

Die W'keit zu Bestehen beträgt

$$
\begin{aligned}
\Pr[B = 1 \cup S \leq 3] &= \Pr[B = 1] + \Pr[S \leq 3] - \Pr[B = 1, S \leq 3] \\
&= f_B(1) + \sum_{i=1}^{3} f_S(i) - \sum_{i=1}^{3} f_{B,S}(1, i) \\
&= \frac{9}{25} + \left(\frac{1}{25} + \frac{2}{25} + \frac{3}{25} + \frac{4}{25} + \frac{5}{25}\right) \\
&\quad - \left(\frac{1}{25} + \frac{2}{25} + \frac{2}{25} + \frac{2}{25} + \frac{2}{25}\right) \\
&= \frac{15}{25} = \frac{3}{5}
\end{aligned}
$$

Bemerkung: Die gemeinsame Dichte ist nicht notwendig durch die beiden Randdichten eindeutig bestimmt!

Mehrere Zufallsvariablen VII: Gemeinsame Verteilung

Für zwei Zufallsvariablen X, Y definiert man die gemeinsame (kumulative) Verteilung

$$
\begin{aligned}
F_{X,Y}(x,y) &= \Pr[X \leq x, Y \leq y] = \Pr[\ \{\omega \mid X(\omega) \leq x, Y(\omega) \leq y\}\ ] \\
&= \sum_{x' \leq x} \sum_{y' \leq y} f_{X,Y}(x',y')\,.
\end{aligned}
$$

Die (kumulative) Randverteilung oder Marginalverteilung ergibt sich gemäß

$$F_X(x) = \sum_{x' \leq x} f_X(x') = \sum_{x' \leq x} \sum_{y \in W_Y} f_{X,Y}(x',y)$$

sowie

$$F_Y(y) = \sum_{y' \leq y} f_Y(y') = \sum_{y' \leq y} \sum_{x \in W_X} f_{X,Y}(x,y')\,.$$

Mehrere Zufallsvariablen VIII: Gemeinsame Verteilung

Die gemeinsame Verteilung $F_{B,S}$ und die Randdichten F_B, F_S von B und S sind:

$F_{B,S}$	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	F_B
1	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{5}{25}$	$\frac{7}{25}$	$\frac{9}{25}$	$\frac{9}{25}$	$\frac{9}{25}$	$\frac{9}{25}$	$\frac{9}{25}$	$\frac{9}{25}$
2	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{6}{25}$	$\frac{10}{25}$	$\frac{14}{25}$	$\frac{16}{25}$	$\frac{16}{25}$	$\frac{16}{25}$	$\frac{16}{25}$	$\frac{16}{25}$
3	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{6}{25}$	$\frac{10}{25}$	$\frac{15}{25}$	$\frac{19}{25}$	$\frac{21}{25}$	$\frac{21}{25}$	$\frac{21}{25}$	$\frac{21}{25}$
4	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{6}{25}$	$\frac{10}{25}$	$\frac{15}{25}$	$\frac{19}{25}$	$\frac{22}{25}$	$\frac{24}{25}$	$\frac{24}{25}$	$\frac{24}{25}$
5	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{6}{25}$	$\frac{10}{25}$	$\frac{15}{25}$	$\frac{19}{25}$	$\frac{22}{25}$	$\frac{24}{25}$	$\frac{25}{25}$	$\frac{25}{25}$
F_S	$\frac{1}{25}$	$\frac{3}{25}$	$\frac{6}{25}$	$\frac{10}{25}$	$\frac{15}{25}$	$\frac{19}{25}$	$\frac{22}{25}$	$\frac{24}{25}$	$\frac{25}{25}$	1

Beispiel 19

Aus einem Skatblatt mit 32 Karten ziehen wir zufällig eine Hand von zehn Karten sowie einen Skat von zwei Karten. Unter den Karten gibt es vier Buben. Die Zufallsvariable X zählt die Anzahl der Buben in der Hand, während Y die Anzahl der Buben im Skat angibt. Wir berechnen die Dichten von X und Y.

Es gilt für $0 \leq x \leq 4$ und $0 \leq y \leq 2$

$$\Pr[X = x] = \frac{\binom{4}{x}\binom{28}{10-x}}{\binom{32}{10}} \qquad \Pr[Y = y] = \frac{\binom{4}{y}\binom{28}{2-y}}{\binom{32}{2}}$$

Mehrere Zufallsvariablen X: Gemeinsame Verteilung

Für die gemeinsame Dichte $f_X, Y\ X$ gilt z.B.

$$f_{X,Y}(4,1) = \Pr[X = 4, Y = 1] = 0,$$

und allgemein

$$f_{X,Y}(x,y) = \frac{\binom{4}{x}\binom{28}{10-x}\binom{4-x}{y}\binom{28-(10-x)}{2-y}}{\binom{32}{10}\binom{22}{2}}\,.$$

Zusammengesetzte Zufallsvariablen I

Definition 20

Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen. Die Zufallsvariable $X + Y$ ist die Abbildung

$$\begin{array}{rcl} X + Y : \Omega & \to & \mathbb{R} \\ \omega & \mapsto & X(\omega) + Y(\omega) \end{array}$$

Weitere Zufallsvariablen wie z. B.

$$3X - 2Y + 2 \qquad X \cdot Y \qquad \min\{X, Y\} \qquad X^Y$$

usw. werden analog definiert.

Allgemein: für Zufallsvariablen $X_1, \ldots, X_n$ und $f : \mathbb{R}^n \to \mathbb{R}$ bezeichnet $f(X_1, \ldots, X_n)$ die Zufallsvariable mit

$$f(X_1, \ldots, X_n)(\omega) := f(X_1(\omega), \ldots, X_n(\omega))$$

Zusammengesetzte Zufallsvariablen II

Beispiel 21 (Klausuren nochmal)

Die Studierenden von Prof. Evilsparza haben zwei Klausuren geschrieben. Ein Studierender besteht, wenn seine Bestnote eine 1 ist oder seine Schnittnote mindestens 3 beträgt. Prof. Evilsparza hat keine Lust zu korrigieren und verteilt die Noten zufällig.

Seien X_1, X_2 die Zufallsvariablen, die die Noten der ersten und der zweiten Klausur angeben, z.B. $X_1(1,3) = 1$ und $X_2(1,3) = 3$.

Seien B, S die Zufallsvariablen, die die beste und die Schnittnote angeben.

Es gilt

$$B = \max\{X_1, X_2\} \qquad S = \frac{X_1 + X_2}{2}$$

Erwartungswert I

Drei faire Münzen werden geworfen.
Frage: Wieviele Köpfe fallen im Schnitt ? (Vive la France!)

Eine Münze wird so lange geworfen, bis sie zum ersten Mal „Kopf“ zeigt.
Frage: Wie oft wirft man im Schnitt?

Beim Roulettespiel (Zahlen 0 bis 36) wetten wir immer wieder 1 Euro auf Nummer 5.

Frage: Wieviel gewinnen wir im Schnitt pro Spiel?

Erwartungswert II

Definition 22

Zu einer Zufallsvariablen X definieren wir den Erwartungswert $\mathbb{E}[X]$ durch

$$\mathbb{E}[X] := \sum_{x \in W_X} x \cdot \Pr[X = x] = \sum_{x \in W_X} x \cdot f_X(x)\ ,$$

sofern $\sum_{x \in W_X} |x| \cdot \Pr[X = x]$ konvergiert.

Erwartungswert II

Beispiel 23

Drei faire Münzen werden geworfen. Wieviele Köpfe fallen in Schnitt?

Modell:

$$\Omega = \{K, Z\}^3 \qquad \Pr[\omega] = \frac{1}{8}$$

Sei X die Zufallsvariable, die die Anzahl der Köpfe angibt.

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=0}^{3} i \cdot \Pr[X = i] \\
&= 1 \cdot \Pr[X = 1] \ + \ 2 \cdot \Pr[X = 2] \ + \ 3 \cdot \Pr[X = 3] \\
&= 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} \ = \ \frac{3}{2} \, .
\end{aligned}$$

Erwartungswert III

Beispiel 24

Eine Münze wird so lange geworfen, bis sie zum ersten Mal „Head“ zeigt. Sei k die Anzahl der durchgeführten Würfe. Wenn k ungerade ist, zahlt der Spieler an die Bank k Euro. Andernfalls (k gerade) zahlt die Bank k Euro an den Spieler. Sei G die Zufallsvariable mit

$$G(\omega) := \begin{cases} |\omega| & \text{falls } |\omega| \text{ ungerade,} \\ -|\omega| & \text{falls } \omega \text{ gerade.} \end{cases}$$

Es gilt

$$\Pr[\text{„Anzahl Würfe} = k\text{“}] = (1/2)^k \,.$$

Damit erhalten wir

$$\mathbb{E}[G] = \sum_{k=1}^{\infty} (-1)^{k-1} \cdot k \cdot \left(\frac{1}{2}\right)^k \,.$$

Erwartungswert IV

Da

$$\sum_{k=1}^{\infty} |(-1)^{k-1} \cdot k| \cdot \left(\frac{1}{2}\right)^k \leq \sum_{k=1}^{\infty} k \cdot \left(\frac{1}{2}\right)^k ,$$

existiert der Erwartungswert $\mathbb{E}[G]$.

Es gilt

$$\begin{aligned}
\mathbb{E}[G] &= \sum_{j=1}^{\infty}(2j-1) \cdot \left(\frac{1}{2}\right)^{2j-1} - \sum_{j=1}^{\infty} 2j \cdot \left(\frac{1}{2}\right)^{2j} \\
&= \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^{2j-1} \cdot [(2j-1) - j] \\
&= \frac{1}{2} \cdot \sum_{j=1}^{\infty}(j-1) \cdot \left(\frac{1}{4}\right)^{j-1} = \ldots = \frac{2}{9} .
\end{aligned}$$

Erwartungswert V

Wird jedoch, um das Risiko zu steigern, der zu zahlende Betrag von k Euro jeweils auf 2^k Euro erhöht, also

$$G'(\omega) := \begin{cases} 2^{|\omega|} & \text{falls } |\omega| \text{ ungerade,} \\ 2^{-|\omega|} & \text{falls } \omega \text{ gerade.} \end{cases}$$

dann existiert $\mathbb{E}[G']$ nicht, da

$$\begin{aligned} \mathbb{E}[G'] &= \sum_{i=1}^{\infty} (-1)^{k-1} \cdot 2^k \cdot \left(\frac{1}{2}\right)^k \\ &= \sum_{i=1}^{\infty} (-1)^{k-1} = +1 - 1 + 1 - 1 + - \ldots . \end{aligned}$$

Erwartungswert VI: Verteilungen

- $X \sim \text{Ber}(p)$ bezeichnet, dass X Bernoulli-verteilt ist mit Erfolgsw'keit p.
- $X \sim \text{Bin}(n, p)$ bezeichnet, dass X binomialverteilt ist mit n Versuchen und Erfolgsw'keit p.
- $X \sim \text{Geo}(p)$ bezeichnet, dass X geometrisch verteilt ist mit Erfolgsw'keit p.

Satz 25

- *Wenn* $X \sim$ *Ber*(p), *dann* $\mathbb{E}[X] = p$.
- *Wenn* $X \sim$ *Bin*(n, p), *dann* $\mathbb{E}[X] = np$.
- *Wenn* $X \sim$ *Geo*(p), *dann* $\mathbb{E}[X] = \frac{1}{p}$.

Erwartungswert VII: Bernoulli- und geometrische Verteilung

Bernoulli-Verteilung:

$$\mathbb{E}[X] = 1 \cdot p + 0 \cdot (1 - p) = p$$

Geometrische Verteilung:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^{\infty} i \cdot pq^{i-1} = p \cdot \sum_{i=1}^{\infty} i \cdot q^{i-1} \\
&= p \cdot \frac{\mathsf{d}}{\mathsf{d}\,q} \left(\sum_{i=0}^{\infty} q^i \right) \\
&= p \cdot \frac{\mathsf{d}}{\mathsf{d}\,q} \left(\frac{1}{1-q} \right) \\
&= p \cdot \frac{1}{(1-q)^2} = \frac{1}{p}.
\end{aligned}$$

Erwartungswert VIII: Binomialverteilung

Binomialverteilung:

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{i=0}^{n} i \cdot \binom{n}{i} p^i q^{n-i} \\
&= pq^{n-1} \sum_{i=1}^{n} \binom{n}{i} \cdot i \cdot \left(\frac{p}{q}\right)^{i-1} \\
&= pq^{n-1} \cdot S(p/q) \qquad \text{mit} \qquad S(z) := \sum_{i=1}^{n} \binom{n}{i} \cdot i \cdot z^{i-1}
\end{aligned}
$$

Wir betrachten die Funktion $S(z)$.

Erwartungswert IX: Binomialverteilung

$$
\begin{aligned}
S(z) &= \sum_{i=1}^{n} \binom{n}{i} \cdot i \cdot z^{i-1} &&= \sum_{i=0}^{n} \binom{n}{i} \cdot \frac{\mathsf{d}}{\mathsf{d}\, z} z^{i} \\
&= \frac{\mathsf{d}}{\mathsf{d}\, z} \left(\sum_{i=0}^{n} \binom{n}{i} z^{i} \right) &&= \frac{\mathsf{d}}{\mathsf{d}\, z} (1+z)^{n} = n\,(1+z)^{n-1}
\end{aligned}
$$

Und damit

$$
\begin{aligned}
\mathbb{E}[X] &= pq^{n-1} \cdot S(p/q) \\
&= pq^{n-1} \cdot n \left(1 + \frac{p}{q}\right)^{n-1} \\
&= pq^{n-1} n \left(\frac{1}{q}\right)^{n-1} = np
\end{aligned}
$$

Varianz I

Wir betrachten die beiden folgenden Zufallsexperimente:

1. Wir würfeln (mit einem fairen Würfel): bei gerader Augenzahl erhalten wir 1 Euro, bei ungerader Augenzahl müssen wir 1 Euro bezahlen.
2. Wir würfeln (mit einem fairen Würfel): bei 6 Augen erhalten wir 5 Euro, ansonsten müssen wir 1 Euro bezahlen.

In beiden Fällen ist der erwartete Gewinn $= 0$.

Dennoch sind die „Schwankungen“ im ersten Fall geringer als im zweiten.

Wie können die „Schwankungen“ quantifiziert werden?

Varianz II

Erste Idee: „Schwankungsgrad“ durch $\mathbb{E}[|X-\mu|]$ berechnen, wobei $\mu = \mathbb{E}[X]$ sei. Die Betragsfunktion ist jedoch mathematisch „unhandlich“.

Zweite Idee: „Schwankungsgrad“ berechnen durch

$$\mathbb{E}[(X-\mu)^2] \quad \text{oder} \quad \sqrt{\mathbb{E}[(X-\mu)^2]}\,.$$

Definition 26

Für eine Zufallsvariable X mit $\mu = \mathbb{E}[X]$ definieren wir die Varianz $\mathrm{Var}[X]$ durch

$$\mathrm{Var}[X] := \mathbb{E}[(X-\mu)^2] = \sum_{x \in W_X} (x-\mu)^2 \cdot \Pr[X = x]\,.$$

Die Größe $\sigma := \sqrt{\mathrm{Var}[X]}$ heißt Standardabweichung von X.

Varianz III

Satz 27

Für eine beliebige Zufallsvariable X gilt

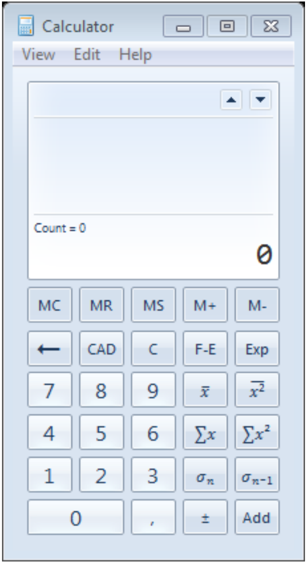
$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2\,.$$

Beweis:

Sei $\mu := \mathbb{E}[X]$. Nach Definition gilt

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2 - 2\mu\cdot X + \mu^2] \\
&= \sum_{\omega\in\Omega} \left(X(\omega)^2 - 2\mu X(\omega) + \mu^2\right) \Pr[\omega] \\
&= \sum_{\omega\in\Omega} X(\omega)^2 \Pr[\omega] - 2\mu \sum_{\omega\in\Omega} X(\omega)\Pr[\omega] + \sum_{\omega\in\Omega} \mu^2 \Pr[\omega] \\
&= \mathbb{E}[X^2] - 2\mu\cdot\mathbb{E}[X] + \mu^2 \\
&= \mathbb{E}[X^2] - 2\mu\cdot\mu + \mu^2 \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2\,.
\end{aligned}$$

Varianz IV



Calculator
View
Edit
Help
Count = 0
0
MC
MR
MS
M+
M-
CAD
C
F-E
Exp
7
8
9
$\bar{x}$
$\overline{x^2}$
4
5
6
$\sum x$
$\sum x^2$
1
2
3
σ_n
σ_{n-1}
0
,
±
Add

Beispiel 28

1. Wir würfeln (mit einem fairen Würfel): bei gerader Augenzahl erhalten wir 1 Euro, bei ungerader Augenzahl müssen wir 1 Euro bezahlen. Es ist

$$\mu = 0 \quad \text{und} \quad \mathrm{Var}[X] = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot (-1)^2 = 1\,.$$

2. Wir würfeln (mit einem fairen Würfel): bei 6 Augen erhalten wir 5 Euro, ansonsten müssen wir 1 Euro bezahlen.
   Es ist

$$\mu = 0 \quad \text{und} \quad \mathrm{Var}[X] = \frac{1}{6} \cdot 5^2 + \frac{5}{6} \cdot (-1)^2 = 5\,.$$

Varianz VI: Bernouilli-Verteilung

Sei X eine Bernoulli-verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p.

$$\begin{aligned}
\mathrm{Var}[X] \quad &= \sum_{x \in W_X} (x - \mathbb{E}[X])^2 \cdot \Pr[X = x] \\
&= (0-p)^2(1-p) + (1-p)^2 p \\
&= p - p^2
\end{aligned}$$

Varianz VII: Geometrische Verteilung

Sei X eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p.

$$
\begin{aligned}
\mathbb{E}[X^2] &= \sum_{i=1}^{\infty} i^2 \cdot pq^{i-1} = \sum_{i=0}^{\infty} (i+1)^2 \cdot pq^i \\
&= p \cdot \left(\sum_{i=0}^{\infty} (i+2)(i+1) \cdot q^i - \sum_{i=0}^{\infty} (i+1) \cdot q^i \right) \\
&= p \cdot \left(\frac{\mathsf{d}^2}{\mathsf{d}\, q^2} \left(\sum_{i=0}^{\infty} q^{i+2} \right) - \frac{\mathsf{d}}{\mathsf{d}\, q} \left(\sum_{i=0}^{\infty} q^{i+1} \right) \right) \\
&= p \cdot \left(\frac{\mathsf{d}^2}{\mathsf{d}\, q^2} \left(\frac{q^2}{1-q} \right) - \frac{\mathsf{d}}{\mathsf{d}\, q} \left(\frac{q}{1-q} \right) \right) \\
&= p \cdot \left(\frac{2}{(1-q)^3} - \frac{1}{(1-q)^2} \right) \\
&= \frac{2-p}{p^2}
\end{aligned}
$$

Momente einer Zufallsvariable

Der Erwartungswert und die Varianz gehören zu den so genannten Momenten einer Zufallsvariablen:

Definition 29

Für eine Zufallsvariable X nennen wir $\mathbb{E}[X^k]$ das k-te Moment und $\mathbb{E}[(X - \mathbb{E}[X])^k]$ das k-te zentrale Moment.

Der Erwartungswert ist also identisch zum ersten Moment, während die Varianz dem zweiten zentralen Moment entspricht.

Teil II

Weitere Grundbegriffe

4. Markov-Diagramme

Markov-Diagramme I

Viele mehrstufige Zufallsexperimente können graphisch mit Markov-Diagrammen dargestellt werden.

Definition 30

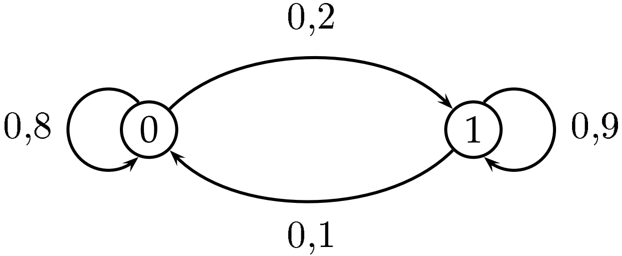
Ein Markov-Diagramm $D = (Q, T, \delta)$ besteht aus

- einer endlichen Menge Q von Zuständen,
- einer Menge $T \subseteq Q \times Q$ von Transitionen, und
- einer W'keitsfunktion $\delta \colon T \to (0, 1]$, die Folgendes erfüllt für jeden Zustand q:

$$\sum_{(q,q') \in T} \delta(q, q') = 1 \; .$$

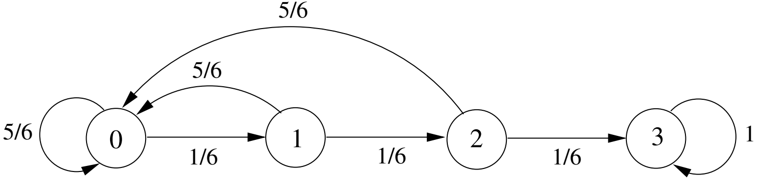
(d.h., die W'keiten der aus q ausgehenden Transitionen addieren sich zu 1, wie es sich gehört!)

Markov-Diagramme II: Beispiel



Beispiel 31

Ein Würfel wird geworfen, bis dreimal hintereinander eine Sechs geworfen wird. Die Sequenzen von Würfen, die mit drei Sechsen Enden, können durch ein Markov-Diagramm dargestellt werden.



Markov-Diagramme IV

Ein (endlicher) Pfad eines Markov-Diagramms ist eine endliche Sequenz $q_0 q_1 \ldots q_k$ von Zuständen mit $k \geq 0$ und $(q_i, q_{i+1}) \in T$ für alle $0 \leq i \leq k-1$. (Ein Zustand darf mehrmals in einem Pfad vorkommen.)

Die Länge von $q_0 q_1 \ldots q_k$ ist k.

Einen Pfad $\pi = q_0 q_1 \ldots q_k$ von D ordnen wir eine W'keit zu:

$$\Pr[\pi] = \begin{cases} 1 & \text{falls } k = 0 \\ \displaystyle\prod_{i=0}^{k-1} \delta(q_i, q_{i+1}) & \text{falls } k > 0 \end{cases}$$

Für eine präfix-freie Menge Π von Pfaden definieren wir

$$\Pr[\Pi] = \sum_{\pi \in \Pi} \Pr[\pi] \,.$$

Markov-Diagramme V

Satz 32

Seien q_1, q_2 *verschiedene Zustände eines Markov-Diagramms* D *mit der Eigenschaft: jeder Zustand von* D *liegt auf mindestens einem Pfad von* q_1 *nach* q_2*.*

Sei $[q_1 \rightsquigarrow q_2]^{q_2=1}$ *die Menge aller Pfade des Diagramms von* q_1 *nach* q_2*, die* q_2 *genau einmal besuchen. Es gilt*

$$\Pr[\ [q_1 \rightsquigarrow q_2]^{q_2=1}\ ] = 1\ .$$

Markov-Diagramme VI

Beweisskizze:

Wir benötigen einige Definitionen. Für jeden Zustand q und jedes $k \geq 1$, sei

- $[q \rightsquigarrow]_k$ die Menge der Pfade aus q der Länge k.
- $[q \rightsquigarrow]_k^{q' \geq 1}$ die Menge der Pfade aus q der Länge k, die q' mindestens einmal besuchen.
- $[q \rightsquigarrow]_k^{q'=0}$ die Menge der Pfade aus q der Länge k, die q' nicht besuchen.
- $[q \rightsquigarrow q']_{\leq k}^{q'=1}$ die Menge der Pfaden von q nach q' der Länge höchstens k

usw.

Markov-Diagramme VII

Für jedes q, q' und jedes $k \geq 1$:

(1) $\Pr[\,[q \rightsquigarrow]_k\,] = 1$. („Massenerhaltung“)
Beweis: Einfache Induktion über k.

(2) Sei n die Anzahl der Knoten von D.
Es gibt ein $\epsilon > 0$, so dass $\Pr[\,[q \rightarrow]_n^{q_2 \geq 1}\,] \geq \epsilon$.
Beweis: weil $[q \rightarrow]_n$ immer mindestens einen Pfad enthält, der q_2 besucht (denn q_2 ist aus jedem Zustand erreichbar).

(3) $\Pr[\,[q \rightsquigarrow]_n^{q_2=0}\,] \leq 1 - \epsilon$ und $\Pr[\,[q \rightsquigarrow]_{2n}^{q_2=0}\,] \leq (1-\epsilon)^2$ und
…und $\Pr[\,[q \rightsquigarrow]_{kn}^{q_2=0}\,] \leq (1-\epsilon)^k$.
Beweis: aus (1) und (2) durch Induktion über $k \geq 0$.

(4) $\Pr[\,[q \rightsquigarrow q']_{\leq k}^{q'=1}\,] = \Pr[\,[q \rightsquigarrow]_k^{q' \geq 1}\,]$.

Beweis: Folgt aus (1) und der Tatsache, dass wenn $\pi \in [q \rightsquigarrow q']_{\leq k}^{q'=1}$ und $\pi\,\pi'$ Länge k hat, dann $\pi\,\pi' \in [q \rightsquigarrow]_k^{q' \geq 1}$.

Markov-Diagramme VIII

Aus (1)-(4) folgt:

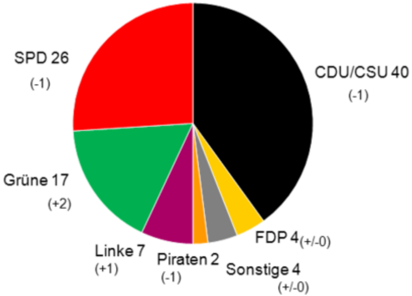
$$
\begin{array}{lll}
 & \Pr[\ [q_1 \leadsto q_2]^{q_2=1}\ ] & \\
= & \lim\limits_{k\to\infty} \Pr[\ [q_1 \leadsto q_2]^{q_2=1}_{\leq k}\ ] & \text{Def. von } [q_1 \leadsto q_2]^{q_2=1} \\
= & \lim\limits_{k\to\infty} \Pr[\ [q_1 \leadsto]^{q_2\geq 1}_{k}\ ] & (4) \\
= & \lim\limits_{k\to\infty} (1 - \Pr[\ [q_1 \leadsto]^{q_2=0}_{k}\ ]) & \Pr[\ [q_1 \leadsto]_k\ ] = 1,\ (1) \\
= & \lim\limits_{k\to\infty} (1 - (1-\epsilon)^k) & \Pr[\ [q_1 \leadsto]^{q_2=0}_{kn}\ ] \leq (1-\epsilon)^k,\ (3) \\
= & 1 &
\end{array}
$$

5. Bedingte Wahrscheinlichkeiten

Bedingte Wahrscheinlichkeiten I

ARD-DeutschlandTREND März 2013 / KW_10

Sonntagsfrage zur Bundestagswahl



Frage: Welche Partei würden Sie wählen, wenn am kommenden Sonntag Bundestagswahl wäre?

Grundgesamtheit: Wahlberechtigte Bevölkerung in Deutschland / Angaben in Prozent
Angaben in Klammern: Vgl. zur Vorwoche

infratest dimap

Bedingte Wahrscheinlichkeiten II

Beispiel 33 (ARD-Deutschlandtrend, März 2013)

Die Zustimmung für die CDU fiel um einen Punkt auf 40%.

66% der Wähler sprechen sich für die Homo-Ehe (sorry, für die rechtliche Gleichstellung gleichgeschlechtlicher Lebenspartnerschaften), bei den CDU-Anhänger jedoch nur 55%.

Eine zufällig gewählte Person wird gefragt, welche Partei sie wählen und ob sie die Gleischstellung befürworten würde.

Modell:

$$\Omega = \{(\text{CDU, Ja}), (\text{CDU, Nein}), (\text{SPD, Ja}), \ldots, (\text{Andere, Nein})\}$$

$$\begin{aligned} \Pr[\text{CDU}] &= \Pr[(\text{CDU, Ja})] + \Pr[(\text{CDU, Nein})] &&= 0.40 \\ \Pr[\text{Ja}] &= \Pr[(\text{CDU, Ja})] + \ldots + \Pr[(\text{Andere, Ja})] &&= 0.66 \end{aligned}$$

Aber was ist die Zahl 0.55?

Bedingte Wahrscheinlichkeiten III

55% der CDU-Anhänger sprechen sich für die Homo-Ehe aus (antworten Ja auf die zweite Frage).

0.55 ist die „W'keit des Ereignisses Ja unter der Bedingung, dass CDU eingetreten ist“. Schreibweise:

$$\Pr[\text{Ja} \mid \text{CDU}] \,.$$

Welche Beziehung gibt es zwischen bedingten und normalen W'keiten?

- Was ist die W'keit von $(\text{CDU, Ja}) = \text{CDU} \cap \text{Ja}$?

$$\left.\begin{array}{l}\text{40\% CDU-Anhänger} \\ \text{55\% davon für die Homo-Ehe}\end{array}\right\} \Longrightarrow 0.40 \cdot 0.55 = 0.22$$

- Es gilt also:

$$\Pr[\text{Ja} \mid \text{CDU}] \cdot \Pr[\text{CDU}] = \Pr[\text{CDU} \cap \text{Ja}]$$

Bedingte Wahrscheinlichkeiten IV

Definition 34

A und B seien Ereignisse mit $\Pr[B] > 0$. Die bedingte Wahrscheinlichkeit $\Pr[A|B]$ von A gegeben B ist definiert als

$$\Pr[A|B] := \frac{\Pr[A \cap B]}{\Pr[B]} \; .$$

Bedingte Wahrscheinlichkeiten V: Eigenschaften

Einige einfache Eigenschaften:

1. $\Pr[B|B] = 1$.
2. $\Pr[A|\Omega] = \Pr[A]$.
3. $\Pr[A \cap B] = \Pr[A \mid B]\Pr[B] = \Pr[B \mid A]\Pr[A]$.
4. (Satz der totalen W'keit) Seien $A_1, \ldots, A_n$ paarweise disjunkt und es gelte $B \subseteq A_1 \cup \ldots \cup A_n$. Dann folgt

$$\Pr[B] = \sum_{i=1}^{n} \Pr[B|A_i] \cdot \Pr[A_i]$$

(aus den bedingungen folgt $\Pr[B] = \sum_{i=1}^{n} \Pr[B \cap A_i]$)

Bedingte Wahrscheinlichkeiten VI

Beispiel 35 (Zweikinderproblem)

Wir nehmen an, dass bei der Geburt eines Kindes beide Geschlechter gleich wahrscheinlich sind.

Prof. Evilsparza braucht zwei kleine Kinder für bösartige Mathe-Experimente. Er besorgt sich eine Liste aller Familien in Garching mit zwei Kindern und wählt zufällig eine davon. Als er an der Tür klingelt macht ein Mädchen auf, schreit und flieht.

Frage: Mit welcher W'keit ist das andere Kind der Familie auch ein Mädchen ?

Modell: $\Omega := \{mm, mj, jm, jj\}$

$$\Pr[mm] = \Pr[mj] = \Pr[jm] = \Pr[jj] = \frac{1}{4}$$

Bedingte Wahrscheinlichkeiten VII

Wir wissen, dass das Ereignis „mindestens ein Mädchen“

$$M := \{mm, mj, jm\}$$

eingetreten ist.

Wir suchen also nach der bedingten W'keit $\Pr[mm|M]$:

$$\Pr[mm|M] = \frac{\Pr[\{mm\} \cap M]}{\Pr[M]} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Bedingte Wahrscheinlichkeiten VIII

Weitere Eigenschaften:

Sei B ein Ereignis mit $\Pr[B] > 0$.

Die bedingten Wahrscheinlichkeiten $\Pr[\bullet|B]$ bilden einen neuen Wahrscheinlichkeitsraum über Ω, denn

$$\sum_{\omega\in\Omega} \Pr[\omega|B] = \sum_{\omega\in\Omega} \frac{\Pr[\omega \cap B]}{\Pr[B]} = \sum_{\omega\in B} \frac{\Pr[\omega]}{\Pr[B]} = \frac{\Pr[B]}{\Pr[B]} = 1.$$

Damit gelten alle Rechenregeln für Wahrscheinlichkeiten auch für bedingte Wahrscheinlichkeiten. Beispielsweise:

$$\begin{aligned} \Pr[A \cup C|B] &\leq \Pr[A|B] + \Pr[C|B] \\ \Pr[\bar{A}|B] &= 1 - \Pr[A|B]\,. \end{aligned}$$

Bedingte Wahrscheinlichkeiten IX: Der Multiplikationssatz

Satz 36 (Multiplikationssatz)

Wenn $\Pr[A_1 \cap \ldots \cap A_n] > 0$, *dann*

$$\begin{aligned}\Pr[A_1 \cap \ldots \cap A_n] &= \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \Pr[A_3|A_1 \cap A_2] \cdot \ldots \\ &\quad \ldots \cdot \Pr[A_n|A_1 \cap \ldots \cap A_{n-1}]\,.\end{aligned}$$

Beweis:

Alle bedingten Wahrscheinlichkeiten sind wohldefiniert, da $\Pr[A_1] \geq \Pr[A_1 \cap A_2] \geq \ldots \geq \Pr[A_1 \cap \ldots \cap A_n] > 0$.

Die rechte Seite der Aussage können wir umschreiben zu

$$\frac{\Pr[A_1]}{1} \cdot \frac{\Pr[A_1 \cap A_2]}{\Pr[A_1]} \cdot \frac{\Pr[A_1 \cap A_2 \cap A_3]}{\Pr[A_1 \cap A_2]} \cdot \ldots \cdot \frac{\Pr[A_1 \cap \ldots \cap A_n]}{\Pr[A_1 \cap \ldots \cap A_{n-1}]}.$$

Offensichtlich kürzen sich alle Terme bis auf $\Pr[A_1 \cap \ldots \cap A_n]$. $\square$

Beispiel 37 (Geburtstagsproblem)

Wie groß ist die Wahrscheinlichkeit, dass in einer m-köpfigen Gruppe mindestens zwei Personen am selben Tag Geburtstag haben?

Umformulierung:
Man werfe b Bälle zufällig und gleich wahrscheinlich in k Körbe.

Frage: Wie groß ist die Wahrscheinlichkeit, dass nach dem Experiment jeder Ball allein in seinem Korb liegt?

Modell:

$$\Omega = \{1, \ldots, b\}^k \qquad \Pr[i_1, \ldots, i_k] = \frac{1}{k^b}$$

Für das Geburtstagsproblem: $k = 365$.

Bedingte Wahrscheinlichkeiten XI

Annahme: $b \leq k$ (sonst ist die W'keit gleich 0).

Wir werfen die Bälle nacheinander. Seien

- $A_i =$ „Ball i landet in einem noch leeren Korb“
- $A =$ „Alle Bälle liegen allein in einem Korb“ .

Nach dem Multiplikationssatz gilt

$$\begin{aligned}
\Pr[A] &= \Pr[A_1 \cap \cdots \cap A_b] \\
&= \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \ldots \cdot \Pr[A_i| \cap_{i=1}^{b-1} A_i].
\end{aligned}$$

Unter der Bedingung, dass die ersten $j-1$ Bälle jeweils in einem leeren Korb landen, bedeutet A_j, dass der j-te Ball in einen der $k-(j-1)$ leeren Körben fällt. Daraus folgt

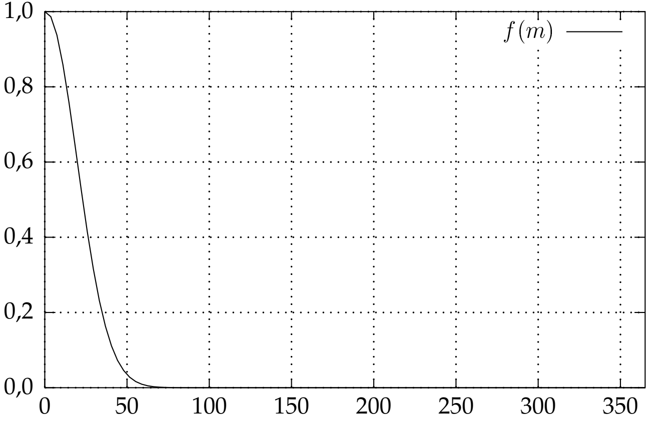
$$\Pr[A_j|A_1 \cap \cdots \cap A_{j-1}] = \frac{k-(j-1)}{k} = 1 - \frac{j-1}{k}.$$

Bedingte Wahrscheinlichkeiten XII

Mit $\Pr[A_j|A_1 \cap \cdots \cap A_{j-1}] = 1 - \frac{j-1}{n}$ erhalten wir

$$
\begin{aligned}
\Pr[A] &= \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \ldots \cdot \Pr[A_b| \cap_{i=1}^{b-1} A_i] \\
&= \prod_{j=1}^{b} \left(1 - \frac{j-1}{k}\right) \\
&\le \prod_{j=1}^{b} e^{-\frac{j-1}{k}} \qquad \left(\text{ wegen } 1 - x \le e^{-x}\right) \\
&= e^{-\frac{1}{n} \cdot \sum_{j=1}^{b-1} j} \\
&= e^{-\frac{b(b-1)}{2k}}
\end{aligned}
$$

Bedingte Wahrscheinlichkeiten XIII



Verlauf von $f(m) := e^{-b(b-1)/(2k)}$ für $n = 365$

Beispiel 38 (Die Lottosensation am 29.6.1995)

Stuttgart (dpa/lsw). Die Staatliche Toto-Lotto GmbH in Stuttgart hat eine Lottosensation gemeldet: Zum ersten Mal in der 40-jährigen Geschichte das deutschen Zahlenlottos wurden zwei identische Gewinnreihen festgestellt. Am 21. Juni dieses Jahres [3016te Ausspielung] kam im Lotto am Mittwoch in der Ziehung A die Gewinnreihe 15-25-27-30-42-48 heraus. Genau die selben Zahlen wurden bei der 1628. Ausspielung im Samstaglotto schon einmal gezogen, nämlich am 20. Dezember 1986. Welch ein Lottozufall: Unter den 49 Zahlen sind fast 14 Millionen verschiedene Sechserreihen möglich.

Bedingte Wahrscheinlichkeiten XV

Frage: Was ist die Wahrscheinlichkeit, dass die erste Wiederholung einer Gewinnreihe ($k = 13,983,816$ verschiedene Sechserreihen) spätestens bei der $b = 3016$-te Ausspielung auftritt?

Antwort: das ist die W'keit, dass mindestens zwei Bälle im selben Korb landen, i.e., $\Pr[\overline{A}]$.

$$\begin{aligned}
\Pr[\overline{A}] &= 1 - \Pr[A] \\
&\approx 1 - e^{-\dfrac{b(b-1)}{2k}} \\
&= 1 - e^{-\dfrac{3016 \cdot 3015}{2 \cdot 13983816}} \\
&\approx 1 - e^{-0.325} \\
&\approx 0,278
\end{aligned}$$

6. Stochastische Unabhängigkeit

Unabhängige Ereignisse I

Beispiel 39 (Zweimaliges Würfeln)

Zwei faire Würfel werden gleichzeitig gewürfelt.

$$\Omega = \{(i,j) \mid 1 \leq i,j \leq 6\} \qquad \Pr[(i,j)] = \frac{1}{36}\,.$$

Wir definieren die Ereignisse

$$A := \text{Augenzahl der ersten Würfel ist gerade}$$
$$B := \text{Augenzahl der zweiten Würfel ist gerade}$$

Nach unserer Intuition gewinnen wir durch das Eintreten von A keine Information in Bezug auf B hinzu.

Im Modell erwarten wir $\Pr[B \mid A] = \Pr[B]$
oder äquivalent $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Unabhängige Ereignisse II

$$
\begin{aligned}
A \cap B &= \{(2,2),(2,4),(2,6),\ldots,(6,2),(6,4),(6,6)\} \\
\Pr[A \cap B] &= \frac{9}{36} = \frac{1}{4} \\
\Pr[A] \cdot \Pr[B] &= \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}
\end{aligned}
$$

Unabhängige Ereignisse III

Intuition: Zwei Ereignisse A, B sind stochastisch unabhängig wenn das Vorwissen, dass B eintritt, keinen Einfluss auf die Wahrscheinlichkeit von A hat.

Definition 40 (Erste Definition)

Zwei Ereignisse $A, B \subseteq \Omega$ mit $\Pr[B] > 0$ sind unabhängig wenn

$$\Pr[A|B] = \Pr[A] \,.$$

Definition 41 (Allgemeinere Definition)

Zwei Ereignisse $A, B \subseteq \Omega$ sind unabhängig wenn

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B].$$

Unabhängige Ereignisse IV

Beide Definitionen sind Äquivalent wenn $\Pr[B] > 0$:

- Wenn $\Pr[A|B] = \Pr[A]$ dann

$$\Pr[A \cap B] = \Pr[A|B] \cdot \Pr[B] = \Pr[A] \cdot \Pr[B]$$

- Wenn $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$ und $\Pr[B] > 0$ dann

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \Pr[A]$$

Unabhängige Ereignisse V

$A :=$ Augenzahl im ersten Wurf ist gerade

$B :=$ Augenzahl im zweiten Wurf ist gerade

$C :=$ Summe der Augenzahlen beider Würfe beträgt 7

Sind A und C unabhängig ?

$$\begin{aligned}
C &= \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\} \\
A \cap C &= \{(2,5),(4,3),(6,1)\} \\
\Pr[A \cap C] &= \frac{3}{36} = \frac{1}{12} \\
\Pr[A] \cdot \Pr[C] &= \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}
\end{aligned}$$

Unabhängige Ereignisse VI

$A :=$ Augenzahl im ersten Wurf ist gerade

$B :=$ Augenzahl im zweiten Wurf ist gerade

$C :=$ Summe der Augenzahlen beider Würfe beträgt 7

A und B sind physikalisch unabhängig und stochastisch unabhängig.

A und C sind physikalisch abhängig aber stochastisch unabhängig.

- Physikalische Unabhängigkeit impliziert stochastische Unabhängigkeit.
- Physikalische Abhängigkeit schliesst stochastische Unabhängigkeit nicht aus.

Unabhängige Ereignismenge I

Wann sind drei Ereignisse A, B, C unabhängig ?

Intuition: Information über den Eintritt von A und B beeinflusst die W'keit von C nicht.

Erster Versuch einer Formalisierung: A, B, C sind unabhängig gdw. $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$, $\Pr[A \cap C] = \Pr[A] \cdot \Pr[C]$, und $\Pr[B \cap C] = \Pr[B] \cdot \Pr[C]$.

Beispiel 42 (Zweimaliges Würfeln)

$A :=$ Augenzahl im ersten Wurf ist gerade,

$B :=$ Augenzahl im zweiten Wurf ist gerade,

$C :=$ Summe der Augenzahlen beider Würfe beträgt 7.

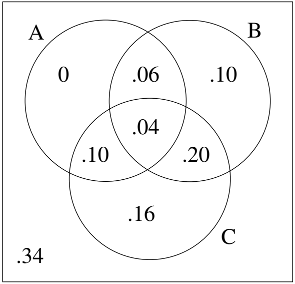
A, B, C sind paarweise unabhängig, aber z.B.

$$\Pr[C | A \cap B] = 0 \neq \Pr[C]$$

Unabhängige Ereignismenge II

Zweiter Versuch einer Formalisierung: A, B, C sind unabhängig gdw. $\Pr[A \cap B \cap C] = \Pr[A] \cdot \Pr[B] \cdot \Pr[C]$.

Beispiel 43 (Glyn George)



$$
\begin{aligned}
\Pr[A] &= 0.20\\
\Pr[B] &= 0.40\\
\Pr[A] &= 0.50\\
\Pr[A \cap B \cap C] &= 0.04\\
\Pr[A|B \cap C] &= 1.00 \neq \Pr[A]
\end{aligned}
$$

Unabhängige Ereignismenge III

Definition 44

Die Ereignisse $A_1, \ldots, A_n$ heißen unabhängig, wenn für alle $1 \leq k \leq n$ und für alle $1 \leq i_1, \ldots, i_k \leq n$ gilt, dass

$$\Pr[A_{i_1} \cap \ldots \cap A_{i_k}] = \Pr[A_{i_1}] \cdot \ldots \cdot \Pr[A_{i_k}]. \tag{1}$$

Eine unendliche Familie von Ereignissen A_i mit $i \in \mathbb{N}$ heißt unabhängig, wenn (1) für jede endliche Teilmenge $I \subseteq \mathbb{N}$ erfüllt ist.

Unabhängige Ereignismenge IV

Lemma 45

Die Ereignisse $A_1, \ldots, A_n$ *sind genau dann unabhängig, wenn für alle* $(s_1, \ldots, s_n) \in \{0,1\}^n$ *gilt, dass*

$$\Pr[A_1^{s_1} \cap \ldots \cap A_n^{s_n}] = \Pr[A_1^{s_1}] \cdot \ldots \cdot \Pr[A_n^{s_n}], \tag{2}$$

wobei $A_i^0 = \bar{A}_i$ *und* $A_i^1 = A_i$.

Unabhängige Ereignismenge V

Beweis (1) $\Rightarrow$ (2):

Induktion über die Anzahl der Nullen in $s_1, \ldots, s_n$.

Wenn $s_1 = \ldots = s_n = 1$ gilt, so ist nichts zu zeigen.

Sonst gelte o.E. $s_1 = 0$ und sei $B = A_2^{s_2} \cap \ldots \cap A_n^{s_n}$.

$$
\begin{aligned}
& \Pr[A_1^{s_1} \cap \cdots \cap A_n^{s_n}] \\
&= \Pr[\bar{A}_1 \cap B] && (s_1 = 0 \text{ und Def. } B) \\
&= \Pr[B] - \Pr[A_1 \cap B] \\
&= \textstyle\prod_{i=2}^{n} \Pr[A_i^{s_i}] - \Pr[A_1] \cdot \prod_{i=2}^{n} \Pr[A_i^{s_i}] && (\text{Ind.Vor}) \\
&= (1 - \Pr[A_1]) \cdot \textstyle\prod_{i=2}^{n} \Pr[A_i^{s_i}] \\
&= \Pr[\bar{A}_1] \cdot \textstyle\prod_{i=2}^{n} \Pr[A_i^{s_i}] \\
&= \textstyle\prod_{i=1}^{n} \Pr[A_i^{s_i}]
\end{aligned}
$$

Unabhängige Ereignismenge VI

Beweis (2) $\Rightarrow$ (1):

Wir zeigen nur (2) $\Rightarrow \Pr[A_1 \cap A_2] = \Pr[A_1] \cdot \Pr[A_2]$.

Es gilt (Satz der totalen Wahrscheinlichkeit):

$$\begin{aligned}
\Pr[A_1 \cap A_2] &= \sum_{s_3,\ldots,s_n \in \{0,1\}} \Pr[A_1 \cap A_2 \cap A_3^{s_3} \cap \ldots \cap A_n^{s_n}] \\
&= \sum_{s_3,\ldots,s_n \in \{0,1\}} \Pr[A_1] \cdot \Pr[A_2] \cdot \Pr[A_3^{s_3}] \cdot \ldots \cdot \Pr[A_n^{s_n}] \\
&= \Pr[A_1] \cdot \Pr[A_2] \cdot \sum_{s_3=0,1} \Pr[A_3^{s_3}] \cdot \ldots \cdot \sum_{s_n=0,1} \Pr[A_n^{s_n}] \\
&= \Pr[A_1] \cdot \Pr[A_2]
\end{aligned}$$

Unabhängige Ereignismenge VII

Lemma 46

Seien A, B und C unabhängige Ereignisse.

1. A *und* $\bar{B}$ *sind unabhängig, sowie* $\bar{A}, B$, *und* $\bar{A}, \bar{B}$.
2. $A \cap B$ *und* C *sind unabhängig.*
3. $A \cup B$ *und* C *sind unabhängig*

Beweis:

(1) Folgt aus Lemma 45 mit $n = 2$.

(2) Zu zeigen: $\Pr[(A \cap B) \cap C] = \Pr[A \cap B] \cdot \Pr[C]$.

$$\begin{aligned} & \Pr[(A \cap B) \cap C] \\ = \;& \Pr[A \cap B \cap C] \\ = \;& \Pr[A] \cdot \Pr[B] \cdot \Pr[C] \quad (A, B, C \text{ unabh.}) \\ = \;& \Pr[A \cap B] \cdot \Pr[C] \quad (A, B \text{ unabh.}) \end{aligned}$$

Unabhängige Ereignismenge VIII

(3) Zu zeigen: $\Pr[(A \cup B) \cap C] = \Pr[A \cup B] \cdot \Pr[C]$.

$$\begin{aligned}
& \Pr[(A \cup B) \cap C] \\
= \quad & \Pr[(A \cap C) \cup (B \cap C)] \\
= \quad & \Pr[A \cap C] + \Pr[B \cap C] - \Pr[A \cap B \cap C] \\
= \quad & \Pr[C] \cdot (\Pr[A] + \Pr[B] - \Pr[A \cap B]) \qquad (A, B, C \text{ unabh.}) \\
= \quad & \Pr[A \cup B] \cdot \Pr[C]
\end{aligned}$$

Unabhängige Zufallsvariablen I

Definition 47

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen unabhängig, wenn für alle $(x_1, \ldots, x_n) \in W_{X_1} \times \ldots \times W_{X_n}$ gilt

$$\Pr[X_1 = x_1, \ldots, X_n = x_n] = \Pr[X_1 = x_1] \cdot \ldots \cdot \Pr[X_n = x_n]\,.$$

Alternative Definition:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)\,.$$

Noch eine Äquivalente Definition:

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)\,.$$

Unabhängige Zufallsvariablen II

Satz 48

Seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen und $S_1, \ldots, S_n$ beliebige Mengen mit $S_i \subseteq W_{X_i}$. Dann sind die Ereignisse „$X_1 \in S_1$“, ..., „$X_n \in S_n$“ unabhängig.

Beweis:

$$
\begin{aligned}
& \Pr[X_1 \in S_1, \ldots, X_n \in S_n] \\
= \quad & \sum_{x_1 \in S_1} \ldots \sum_{x_n \in S_n} \Pr[X_1 = x_1, \ldots, X_n = x_n] \\
\stackrel{\text{Unabh.}}{=} \quad & \sum_{x_1 \in S_1} \ldots \sum_{x_n \in S_n} \Pr[X_1 = x_1] \cdot \ldots \cdot \Pr[X_n = x_n] \\
= \quad & \left(\sum_{x_1 \in S_1} \Pr[X_1 = x_1] \right) \cdot \ldots \cdot \left(\sum_{x_n \in S_n} \Pr[X_n = x_n] \right) \\
= \quad & \Pr[X_1 \in S_1] \cdot \ldots \cdot \Pr[X_n \in S_n] \,.
\end{aligned}
$$

Unabhängige Zufallsvariablen III

Satz 49

$f_1, \ldots, f_n$ seien reellwertige Funktionen ($f_i : \mathbb{R} \to \mathbb{R}$ für $i = 1, \ldots, n$). Wenn die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig sind, dann gilt dies auch für $f_1(X_1), \ldots, f_n(X_n)$.

Beweis:

Sei $z_i \in W_{f(X_i)}$ für $i = 1, \ldots, n$ und $S_i = \{x;\ f(x) = z_i\}$.

$$\begin{aligned}
& \Pr[f_1(X_1) = z_1, \ldots, f_n(X_n) = z_n] \\
= \quad & \Pr[X_1 \in S_1, \ldots, X_n \in S_n] \\
\overset{\text{Unabh.}}{=} \quad & \Pr[X_1 \in S_1] \cdot \ldots \cdot \Pr[X_n \in S_n] \\
= \quad & \Pr[f_1(X_1) = z_1] \cdot \ldots \cdot \Pr[f_n(X_n) = z_n] .
\end{aligned}$$

$\square$

Unabhängige Zufallsvariablen IV

Satz 50

Seien X, Y unabhängige Zufallsvariablen und $Z := X + Y$. Es gilt

$$f_Z(z) = \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)\,.$$

Beweis: Mit Hilfe des Satzes von der totalen Wahrscheinlichkeit:

$$\begin{aligned}
f_Z(z) &= \Pr[Z = z] = \sum_{x \in W_X} \Pr[X + Y = z | X = x] \cdot \Pr[X = x] \\
&= \sum_{x \in W_X} \Pr[Y = z - x \mid X = x] \cdot \Pr[X = x] \\
&\overset{\text{Unab.}}{=} \sum_{x \in W_X} \Pr[Y = z - x] \cdot \Pr[X = x] \\
&= \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)\,.
\end{aligned}$$

Beispiel 51

Zwei faire Würfel werden geworfen. Seien X, Y die Zufallsvariablen, die die Augenzahl des ersten bzw. des zweiten Würfels angeben. Wir berechnen die Dichte von $Z := X + Y$:

$$\begin{aligned}\Pr[Z = z] &= \sum_{x \in W_X} \Pr[X = x] \cdot \Pr[Y = z - x] \\ &= \sum_{x=1}^{6} \frac{1}{6} \cdot \Pr[Y = z - x] = \sum_{x=\max\{1,z-6\}}^{\min\{6,z-1\}} \frac{1}{36}.\end{aligned}$$

Für $2 \le z \le 7$ erhalten wir

$$\Pr[Z = z] = \sum_{i=1}^{z-1} \frac{1}{36} = \frac{z-1}{36}\ .$$

Unabhängige Zufallsvariablen VI

Und für $7 < z \leq 12$:

$$\Pr[Z = z] = \frac{13 - z}{36}\,.$$

Den Ausdruck

$$\sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)$$

aus Satz 50 nennt man in Analogie zu den entsprechenden Begriffen bei Potenzreihen auch Faltung oder Konvolution der Dichten f_X und f_Y.

7. Satz von Bayes

Beispiel 52 (HIV-Test)

Krankheitstests erkennen machmal kranke Menschen als gesund ("falsch negativ", sehr gefährlich) oder gesunde Menschen als krank ("falsch positiv", weniger gefährlich).

Für den ELISA-Test zur Erkennung von Antikörpern gegen das HIV beträgt die W'keit eines falsch negativen Resultates etwa 0.1%. Die W'keit eines falsch positiven Resultates beträgt 0.2%.

In Deutschland leben ca. 80,000 HIV-Infizierte, etwa 0.1% der Bevölkerung.

Eine zufällig gewählte Person (in Deutschland lebend) wird getestet. Das Ergebnis ist positiv.

Frage: Wie hoch ist die W'keit einer Infektion?

Satz von Bayes II

Seien I, G die Ereignisse "die Person ist infiziert" und "die Person ist gesund".

$$\Pr[I] \approx 0.001 \qquad \Pr[G] = 1 - \Pr[I] \approx 0.999$$

Seien P, N die Ereignisse "das Ergebnis des Tests ist positiv" und "das Ergebnis des Tests ist negativ".

$$\Pr[P|I] = 0.999 \qquad \Pr[N|G] = 0.998$$

Wir suchen $\Pr[I|P]$.

Satz von Bayes III

$$\Pr[I] \approx 0.001 \quad \Pr[G] \approx 0.999 \quad \Pr[P|I] = 0.999 \quad \Pr[N|G] = 0.998$$

Es gilt $\Pr[I|P] \cdot \Pr[P] = \Pr[I \cap P] = \Pr[P|I] \cdot \Pr[I]$.

$$\begin{aligned}
\Pr[I|P] &= \frac{\Pr[P|I] \cdot \Pr[I]}{\Pr[P]} \\
&= \frac{\Pr[P|I] \cdot \Pr[I]}{\Pr[P|G] \cdot P[G] + \Pr[P|I] \cdot P[I]} \\
&= \frac{\Pr[P|I] \cdot \Pr[I]}{(1 - \Pr[N|G]) \cdot P[G] + \Pr[P|I] \cdot P[I]} \\
&= \frac{0.999 \cdot 0.001}{0.002 \cdot 0.999 + 0.999 \cdot 0.001} \\
&= \frac{0.001}{0.003} = \frac{1}{3}
\end{aligned}$$

Satz von Bayes IV

Satz 53 (Satz von Bayes)

Die Ereignisse $A_1, \ldots, A_n$ *seien paarweise disjunkt, mit* $\Pr[A_j] > 0$ *für alle* j*. Ferner sei* $B \subseteq A_1 \cup \ldots \cup A_n$ *ein Ereignis mit* $\Pr[B] > 0$*. Dann gilt für ein beliebiges* $i = 1, \ldots, n$

$$\Pr[A_i|B] = \frac{\Pr[A_i \cap B]}{\Pr[B]} = \frac{\Pr[B|A_i] \cdot \Pr[A_i]}{\sum_{j=1}^{n} \Pr[B|A_j] \cdot \Pr[A_j]}\,.$$

Analog gilt für paarweise disjunkte Ereignisse $A_1, A_2, \ldots$ *mit* $B \subseteq \bigcup_{i=1}^{\infty} A_i$*, dass*

$$\Pr[A_i|B] = \frac{\Pr[A_i \cap B]}{\Pr[B]} = \frac{\Pr[B|A_i] \cdot \Pr[A_i]}{\sum_{j=1}^{\infty} \Pr[B|A_j] \cdot \Pr[A_j]}\,.$$

Satz von Bayes V

Seien $\mathcal{K}_1, \ldots, \mathcal{K}_n$ Krankheiten und $\mathcal{S}_1, \ldots \mathcal{S}_m$ Symptome.

Wir definieren die Ereignisse:

$$K_i = \textit{Der Patient hat die Krankheit } \mathcal{K}_i$$
$$S_j = \textit{Der Patient zeigt das Symptom } \mathcal{S}_j$$

Eine Belegung der Symptome ist ein Element von $\{0,1\}^m$.

Jedem Patient kann eine Belegung $\mathcal{B}$ zugeordnet werden, die angibt, welche Symptome im Patient vorhanden ($\mathcal{B}(j) = 1$) und abwesend ($\mathcal{B}(j) = 0$) sind.

Satz von Bayes VI

Wir definieren das Ereignis:

$$B = \textit{Der Patient zeigt die Belegung } \mathcal{B}.$$

Es gilt $B = \bigcap_{j=1}^{m} U_j$, wobei

$$U_j = \begin{cases} S_j & \text{falls } \mathcal{B}(j) = 1 \\ \overline{S}_j & \text{sonst} \end{cases}$$

Der Satz von Bayes kann nun mit $A_i := K_i$ angewendet werden, unter der Annahme, dass der Patient eine und genau eine der Krankheiten $\mathcal{K}_1, \ldots, \mathcal{K}_n$ hat:

$$\Pr[K_i|B] = \frac{\Pr[B|K_i] \cdot \Pr[K_i]}{\Pr[B]} \; \frac{\Pr[B|K_i] \cdot \Pr[K_i]}{\sum_{j=1}^{\infty} \Pr[B|K_j] \cdot \Pr[K_j]} \, .$$

Satz von Bayes VII

Wenn $S_1 \cap K_j, \ldots, S_m \cap K_j$ unabhängig sind für alle $1 \leq j \leq n$:

$$\begin{aligned}
\Pr[B|K_j] \cdot \Pr[K_j] &= \Pr[B \cap K_j] \\
&= \Pr[\bigcap_{k=1}^{m} U_k \cap K_j] \\
&= \Pr[\bigcap_{k=1}^{m} (U_k \cap K_j)] \\
&\overset{\text{Unabh.}}{=} \prod_{k=1}^{m} \Pr[U_k \cap K_j] \\
&= \prod_{k=1}^{m} (\Pr[U_k|K_j] \cdot \Pr[K_j]) \\
&= \Pr[K_j]^m \cdot \prod_{k=1}^{m} \Pr[U_k|K_j]
\end{aligned}$$

Satz von Bayes VIII

Durch Einsetzen erhalten wir:

$$\Pr[K_i|B] = \frac{\Pr[K_i]^m \cdot \prod_{k=1}^{m} \Pr[U_k|K_i]}{\sum_{j=1}^{n} \Pr[K_j]^m \cdot \prod_{k=1}^{m} \Pr[U_k|K_j]}$$

mit

$$\Pr[U_k|K_i] = \begin{cases} \Pr[S_k|K_i] & \text{falls } \mathcal{B}(k) = 1 \\ \Pr[\overline{S}_k|K_i] = 1 - \Pr[S_k|K_i] & \text{sonst} \end{cases}$$

Beispiel 54 (Diagnose von Hirntumoren)

Microscopic description: The H&E sections show a tumor with a predominately papillary or filiform architecture (Figures 2 and 3). The cells are arranged in a pseudocolumnar fashion forming perivascular palisades about a fibrovascular core. There is marked nuclear pleomorphism with many bizarre nuclei, and a brisk mitotic rate (Figure 4), the nucleoli are indistinct. In areas, the papillary configuration gives way to patternless sheets of cells interrupted by zonal necrosis. Focally, the tumor cells surround microcavities which are either optically empty or contain “wisps” of eosinophilic fibrillary material. Confluent calcospherites are focally prominent. Vascular endothelial proliferation is present. Foci of glial frame indicate invasiveness. Rests of native plexus are occasionally encountered.

Satz von Bayes X

A priori Wahrscheinlichkeiten:

Diagnose	Wahrscheinlichkeit
Grade I	0.25
Grade II	0.25
Grade III	0.20
Grade IV	0.30

Satz von Bayes XI

Bedingte Wahrscheinlichkeiten:

		Gr. I	Gr. II	Gr. III	Gr. IV
Diffuse infiltration	DI	0.364	0.878	0.896	0.819
Necrosis	NE	0.221	0.246	0.716	0.850
Vascular abnormalities	VA	0.648	0.686	0.849	0.926
Vascular occlusions	VO	0.113	0.121	0.518	0.640
Nuclear polymorphism	NP	0.702	0.825	0.971	0.968
Cellular polymorphism	CP	0.362	0.706	0.801	0.833
Visible perycarion	VP	0.494	0.865	0.896	0.887
Typical mitoses	TM	0.121	0.109	0.726	0.853
Atipical mitoses	AM	0.005	0.029	0.556	0.733
Undifferentiated cells	UC	0.413	0.450	0.651	0.943

Satz von Bayes XII

Thomas Bayes (1702–1761) war ein bekannter Theologe und Mitglied der Royal Society. Als sein bedeutendstes Werk gilt sein Beitrag zur Wahrscheinlichkeitstheorie „Essay Towards Solving a Problem in the Doctrine of Chances“. Diese Arbeit wurde erst 1763 publiziert.

Teil III

Rechnen mit Wahrscheinlichkeiten

8. Rechenregeln für Erwartungswert und Varianz

Rechenregeln für den Erwartungswert I

Satz 55

Für jede Zufallsvariable X gilt

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

Beweis:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] \\
&= \sum_{x \in W_X} x \sum_{\{\omega \in \Omega | X(\omega) = x\}} \Pr[\omega] \\
&= \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega]\,.
\end{aligned}$$

Rechenregeln für den Erwartungswert II

Satz 56

Sei X eine Zufallsvariable mit $W_X \subseteq \mathbb{N}_0$. Dann gilt

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr[X \geq i]\,.$$

Beweis:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr[X = i] = \sum_{i=1}^{\infty} \sum_{j=1}^{i} \Pr[X = i] \\
&= \sum_{j=1}^{\infty} \sum_{i \geq j}^{\infty} \Pr[X = i] = \sum_{j=1}^{\infty} \Pr[X \geq j]\,.
\end{aligned}$$

Rechenregeln für den Erwartungswert III

Beispiel 57

Zwei faire Würfel werden geworfen. Der Spieler gewinnt k Euro, wobei k die kleinere der zwei Augenzahlen ist.

Frage: Wieviel Euro gewinnt der Spieler in Durchschnitt?

$$\Omega = \{(1,1),(1,2),\ldots,(6,6)\} \qquad \Pr[(i,j)] = \frac{1}{36}$$

Wir definieren zwei Zufallsvariablen $X_1, X_2 \colon \Omega \to \mathbb{R}$:

$$X_1(i,j) := i \qquad X_2(i,j) := j$$

Wir müssen $\mathbb{E}[\min\{X_1, X_2\}]$ berechnen.

Rechenregeln für den Erwartungswert IV

Wir verwenden Satz 56.

$$
\begin{aligned}
\mathbb{E}[\min\{X_1, X_2\}] &= \sum_{i=1}^{6} \Pr[\min\{X_1, X_2\} \geq i] \\
&= \sum_{i=1}^{6} \Pr[„X_1 \geq i“ \cap „X_2 \geq i“] \\
&\overset{\text{Unabh.}}{=} \sum_{i=1}^{6} \Pr[X_1 \geq i] \cdot \Pr[X_2 \geq i] \\
&= \sum_{i=1}^{6} (\Pr[X_1 \geq i])^2 \\
&= \sum_{i=1}^{6} \left(\frac{(7-i)}{6}\right)^2 \\
&= \frac{91}{36} \approx 2.528
\end{aligned}
$$

Rechenregeln für den Erwartungswert V

Definition 58

Sei X eine Zufallsvariable und A ein Ereignis mit $\Pr[A] > 0$. Die bedingte Zufallsvariable $X|A$ besitzt die Dichte

$$f_{X|A}(x) := \Pr[X = x \mid A] = \frac{\Pr[\text{„}X = x\text{“} \cap A]}{\Pr[A]}\,.$$

Die Definition von $f_{X|A}$ ist zulässig, da

$$\sum_{x \in W_X} f_{X|A}(x) = \sum_{x \in W_X} \frac{\Pr[\text{„}X = x\text{“} \cap A]}{\Pr[A]} = \frac{\Pr[A]}{\Pr[A]} = 1\,.$$

Der Erwartungswert $\mathbb{E}[X|A]$ der Zufallsvariablen $X|A$ berechnet sich entsprechend:

$$\mathbb{E}[X|A] = \sum_{x \in W_X} x \cdot f_{X|A}(x)\,.$$

Rechenregeln für den Erwartungswert VI

Satz 59

Sei X *eine Zufallsvariable. Für paarweise disjunkte Ereignisse* $A_1, \ldots, A_n$ *mit* $A_1 \cup \ldots \cup A_n = \Omega$ *und* $\Pr[A_1], \ldots, \Pr[A_n] > 0$ *gilt*

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X|A_i] \cdot \Pr[A_i]\,.$$

Für paarweise disjunkte Ereignisse $A_1, A_2, \ldots$ *mit* $\bigcup_{i=1}^{\infty} A_k = \Omega$ *und* $\Pr[A_1]$, $\Pr[A_2], \ldots > 0$ *gilt analog*

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{E}[X|A_i] \cdot \Pr[A_i],$$

sofern die Erwartungswerte auf der rechten Seite alle existieren und die Summe $\sum_{i=1}^{\infty} |\mathbb{E}[X|A_i]| \cdot \Pr[A_i]$ *konvergiert.*

Rechenregeln für den Erwartungswert VII

Beweis:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] \\
&= \sum_{x \in W_X} x \cdot \sum_{i=1}^{n} \Pr[X = x | A_i] \cdot \Pr[A_i] \quad \text{(Tot. W'keit)} \\
&= \sum_{i=1}^{n} \Pr[A_i] \sum_{x \in W_X} x \cdot \Pr[X = x | A_i] \\
&= \sum_{i=1}^{n} \Pr[A_i] \cdot \mathbb{E}[X | A_i].
\end{aligned}$$

Der Beweis für den unendlichen Fall verläuft analog.

Beispiel 60

Wir werfen eine Münze so lange, bis zum ersten Mal „Kopf“ erscheint. Dies geschehe in jedem Wurf unabhängig mit Wahrscheinlichkeit p. Wir definieren dazu die Zufallsvariable $X :=$ „Anzahl der Würfe“. Wir haben bereits gesehen, dass

$$\Pr[X = k] = p(1-p)^{k-1}$$

und damit

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k \cdot p(1-p)^{k-1} = p \cdot \frac{1}{(1-(1-p))^2} = \frac{1}{p}\,.$$

Rechenregeln für den Erwartungswert IX

Alternative Berechnungsmethode: (gestützt auf Satz 59)
Sei $K_1 :=$ „Im ersten Wurf fällt Kopf“.

Offensichtlich gilt $\mathbb{E}[X|K_1] = 1$.

Nehmen wir nun an, dass im ersten Wurf nicht „Kopf“ gefallen ist. Wir starten das Experiment neu. Sei X' die Anzahl der Würfe bis zum ersten Auftreten von „Kopf“ im neu gestarteten Experiment.

Wegen der Gleichheit der Experimente gilt $\mathbb{E}[X'] = \mathbb{E}[X]$.

Damit schließen wir $\mathbb{E}[X|\bar{K}_1] = 1 + \mathbb{E}[X'] = 1 + \mathbb{E}[X]$

und erhalten eine einfache Gleichung für $\mathbb{E}[X]$:

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[X|K_1] \cdot \Pr[K_1] + \mathbb{E}[X|\bar{K}_1] \cdot \Pr[\bar{K}_1] \\
&= 1 \cdot p + (1 + \mathbb{E}[X]) \cdot (1 - p)\,.
\end{aligned}$$

mit Lösung $\mathbb{E}[X] = 1/p$.

Rechenregeln für den Erwartungswert X

Formale Ableitung von $\mathbb{E}[X|\bar{K}_1] = 1 + \mathbb{E}[X]$:

$$
\begin{aligned}
\mathbb{E}[X|\bar{K}_1] &= \sum_{n=1}^{\infty} n \cdot \frac{\Pr[X = n \cap \bar{K}_1]}{\Pr[\bar{K}_1]} \\
&= \sum_{n=2}^{\infty} n \cdot \frac{\Pr[z^{n-1}k]}{\Pr[\bar{K}_1]} = \sum_{n=1}^{\infty} (n+1) \cdot \frac{\Pr[z^n k]}{\Pr[\bar{K}_1]} \\
&= \sum_{n=1}^{\infty} (n+1) \cdot \frac{\Pr[\bar{K}_1] \cdot \Pr[z^{n-1}k]}{\Pr[\bar{K}_1]} \\
&= \sum_{n=1}^{\infty} (n+1) \cdot \Pr[z^{n-1}k] \\
&= \sum_{n=1}^{\infty} n \cdot \Pr[z^{n-1}k] + \sum_{n=1}^{\infty} \Pr[z^{n-1}k] = \mathbb{E}[X] + 1
\end{aligned}
$$

Rechenregeln für den Erwartungswert XI

Beispiel 61

Wenn man eine Sechs wirft dann darf man bei Mensch-ärgere-dich-nicht bekanntermaßen nochmal werfen. In Spanien gilt jedoch noch folgende Regel: wirft man dreimal hintereinander eine Sechs, muss man den Stein, den man zuletzt gezogen hat, zurück zum Startpunkt bringen.

Frage: Wie oft muss man einen fairen Würfel im Mittel werfen, um dreimal hintereinander eine Sechs zu würfeln?

Rechenregeln für den Erwartungswert XII

Als Elementarereignisse für jede Stufe des Experiments (einmal würfeln) nehmen wir 6 (eine Sechs wird gewürfelt) mit $\Pr[6] = 1/6$ und $\overline{6}$ (keine Sechs wird gewürfelt) mit $\Pr[\overline{6}] = 5/6$.

Modell:

$$\Omega = \left((\epsilon + 6 + 66)\overline{6}\right)^* 666 \qquad \Pr[\omega] = \left(\frac{1}{6}\right)^{s(\omega)} \cdot \left(\frac{5}{6}\right)^{|\omega| - s(\omega)}$$

mit $s(\omega) =$ Anzahl von Sechsen in ω.

Wir definieren drei Ereignisse:

$$\begin{aligned} \mathbf{\overline{6}} \quad &= \quad \text{Elemente aus } \Omega\text{, die mit } \overline{6} \text{ anfangen} \\ \mathbf{6\overline{6}} \quad &= \quad \text{Elemente aus } \Omega\text{, die mit } 6\overline{6} \text{ anfangen} \\ \mathbf{66\overline{6}} \quad &= \quad \text{Elemente aus } \Omega\text{, die mit } 66\overline{6} \text{ anfangen} \end{aligned}$$

Es gilt

$$\Omega = \mathbf{\overline{6}} \cup \mathbf{6\overline{6}} \cup \mathbf{66\overline{6}} + \{666\}$$

Rechenregeln für den Erwartungswert XIII

Sei X die Zufallsvariable, die die Gesamtzahl der Würfel angibt.
Mit Satz 59 gilt

$$\begin{aligned} E[X] &= E[X \mid \overline{\mathbf{6}}] \cdot \Pr[\overline{6}] + E[X \mid \mathbf{6}\overline{\mathbf{6}}] \cdot \Pr[\mathbf{6}\overline{\mathbf{6}}] \\ &\quad + E[X \mid \mathbf{66}\overline{\mathbf{6}}] \cdot \Pr[\mathbf{66}\overline{\mathbf{6}}] + E[X \mid 666] \cdot \Pr[666] \end{aligned}$$

Wir haben

$$\begin{aligned} E[X \mid \overline{\mathbf{6}}] &= 1 + E[X] & \qquad E[X \mid \mathbf{66}\overline{\mathbf{6}}] &= 3 + E[X] \\ E[X \mid \mathbf{6}\overline{\mathbf{6}}] &= 2 + E[X] & \qquad E[X \mid 666] &= 3 \end{aligned}$$

und damit

$$E[X] = \frac{5}{6}(1 + E[X]) + \frac{5}{36}(2 + E[X]) + \frac{5}{216}(3 + E[X]) + \frac{3}{216}$$

mit Lösung $E[X] = 258$.

Rechenregeln für den Erwartungswert XIV

Satz 62 (Linearität des Erwartungswerts)

Für beliebige Zufallsvariablen $X_1, \ldots, X_n$ *und* $a_1, \ldots, a_n \in \mathbb{R}$ *gilt*

$$\mathbb{E}[a_1 X_1 + \cdots + a_n X_n] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n]\,.$$

Beweis:

$$\begin{aligned}
&\mathbb{E}[a_1 X_1 + \ldots + a_n X_n] \\
&= \sum_{\omega \in \Omega} (a_1 \cdot X_1(\omega) + \ldots + a_n \cdot X_n(\omega)) \cdot \Pr[\omega] \\
&= a_1 \cdot \left(\sum_{\omega \in \Omega} X_1(\omega) \cdot \Pr[\omega] \right) + \cdots + a_n \cdot \left(\sum_{\omega \in \Omega} X_n(\omega) \cdot \Pr[\omega] \right) \\
&= a_1 \cdot \mathbb{E}[X_1] + \ldots + a_n \cdot \mathbb{E}[X_n]\,.
\end{aligned}$$

$\square$

Beispiel 63

n betrunkene Seeleute torkeln nach dem Landgang in ihre Kojen. Sie haben völlig die Orientierung verloren, weshalb wir annehmen, dass jede Zuordnung der Seeleute zu den n Betten gleich wahrscheinlich ist (genau ein Seemann pro Bett).

Frage: Wie viele Seeleute liegen im Mittel im richtigen Bett?

Die Anzahl der Seeleute im richtigen Bett zählen wir mit der Zufallsvariablen X, die als Summe der Zufallsvariablen $X_1, \ldots, X_n$ dargestellt wird, wobei

$$X_i := \begin{cases} 1 & \text{falls Seemann } i \text{ in seinem Bett liegt,} \\ 0 & \text{sonst.} \end{cases}$$

Offenbar gilt $X := X_1 + \cdots + X_n$.

Rechenregeln für den Erwartungswert XVI

Für die Variablen X_i erhalten wir

$$\Pr[X_i = 1] = \frac{1}{n}$$

da jedes Bett von Seemann i mit gleicher Wahrscheinlichkeit aufgesucht wird.

Daraus folgt

$$\mathbb{E}[X_i] = 0 \cdot \Pr[X_i = 0] \; + \; 1 \cdot \Pr[X_i = 1] = \frac{1}{n},$$

und somit

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{1}{n} = 1\,.$$

Im Mittel sucht also nur ein Seemann sein eigenes Bett auf.

Rechenregeln für den Erwartungswert XVII

Satz 64 (Multiplikativität des Erwartungswerts)

Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[X_1 \cdot \ldots \cdot X_n] = \mathbb{E}[X_1] \cdot \ldots \cdot \mathbb{E}[X_n]\,.$$

Beweis: Wir beweisen den Fall $n = 2$. Der allgemeine Fall ist analog.

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= \sum_{x \in W_X} \sum_{y \in W_Y} xy \cdot \Pr[X = x, Y = y] \\
&\overset{\text{Unabh.}}{=} \sum_{x \in W_X} \sum_{y \in W_Y} xy \cdot \Pr[X = x] \cdot \Pr[Y = y] \\
&= \sum_{x \in W_X} x \cdot \Pr[X = x] \sum_{y \in W_Y} y \cdot \Pr[Y = y] \\
&= \mathbb{E}[X] \cdot \mathbb{E}[Y]\,.
\end{aligned}$$

Rechenregeln für den Erwartungswert XVIII

Für die Gültigkeit von Satz 64 ist die Unabhängigkeit der Zufallsvariablen wirklich notwendig.

Sei $Y = -X$ für eine Zufallsvariable mit einer von Null verschiedenen Varianz. Dann gilt

$$\mathbb{E}[X \cdot Y] = -\mathbb{E}[X^2] \neq -(\mathbb{E}[X])^2 = \mathbb{E}[X] \cdot \mathbb{E}[Y]\,.$$

Rechenregeln für die Varianz I

Satz 65

Für eine beliebige Zufallsvariable X und $a, b \in \mathbb{R}$ gilt

$$\mathrm{Var}[a \cdot X + b] = a^2 \cdot \mathrm{Var}[X]\,.$$

Beweis:

Aus der in Satz 62 gezeigten Linearität des Erwartungswerts folgt $\mathbb{E}[X + b] = \mathbb{E}[X] + b$.

Zusammen mit der Definition der Varianz ergibt sich damit sofort

$$\mathrm{Var}[X + b] = \mathbb{E}[(X + b - \mathbb{E}[X + b])^2] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathrm{Var}[X]\,.$$

Weiter folgt mit Satz 27:

$$\mathrm{Var}[a \cdot X] = \mathbb{E}[(aX)^2] - \mathbb{E}[aX]^2 = a^2\mathbb{E}[X^2] - (a\mathbb{E}[X])^2 = a^2 \cdot \mathrm{Var}[X]\,,$$

und daraus zusammen die Behauptung.

Rechenregeln für die Varianz II

Satz 66 (Additivität der Varianz)

Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ und $X := X_1 + \ldots + X_n$ gilt

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]\,.$$

Beweis:

Fall $n = 2$ mit Zufallsvariablen X und Y.

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2 + 2XY + Y^2] = \mathbb{E}[X^2] + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2]$$
$$\mathbb{E}[X+Y]^2 = (\mathbb{E}[X] + \mathbb{E}[Y])^2 = \mathbb{E}[X]^2 + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y]^2$$

Wir ziehen die zweite Gleichung von der ersten ab und erhalten

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

Mit Hilfe von Satz 27 folgt die Behauptung. $\square$

Rechenregeln für die Varianz IV

Für abhängige Zufallsvariablen $X_1, \ldots, X_n$ gilt Satz 66 im Allgemeinen nicht. Als Beispiel funktioniert wiederum der Fall $X = -Y$:

$$\mathrm{Var}[X+Y] = 0 \neq 2 \cdot \mathrm{Var}[X] = \mathrm{Var}[X] + \mathrm{Var}[Y]\,.$$

9. Indikatorvariablen

Indikatorvariablen I

Indikatorvariablen erlauben die Darstellung von W'keiten als Erwartungswerte.

Dadurch können Rechenregeln für den Erwartungswert auch für die Berechnung von W'keiten angewendet werden.

Definition 67

Zu einem Ereignis A heißt die Zufallsvariable

$$I_A := \begin{cases} 1 & \text{falls } A \text{ eintritt,} \\ 0 & \text{sonst} \end{cases}$$

Indikatorvariable des Ereignisses A.

(Die Variablen $X_1, \ldots, X_n$ im Matrosen-Beispiel sind Indikatorvariablen.)

Indikatorvariablen II

Lemma 68

Seien $A, A_1, \ldots, A_n$ Ereignisse (nicht notwendig unabhängig!).

(1) $\Pr[A] = \mathbb{E}[I_A]$.

(2) $\Pr[A_1 \cap \ldots \cap A_n] = \mathbb{E}[I_{A_1} \cdot \ldots \cdot I_{A_n}]$.

Beweis:

(1) $\mathbb{E}[I_A] = 1 \cdot \Pr[A] + 0 \cdot \Pr[\bar{A}] = \Pr[A]$.

(2) Das Produkt von Indikatorvariablen ist gleich 1 genau dann, wenn alle entsprechenden Ereignisse eintreten. $\square$

Indikatorvariablen III

Beispiel 69

Wir betrachten wieder das Beispiel der betrunkenen Matrosen.

Frage: Welche Varianz hat die Variable $X := X_1 + \ldots + X_n$?

Wir haben $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - 1$.

Mit $X = X_1 + \cdots + X_n$ gilt:

$$\begin{aligned}
\mathbb{E}[X^2] &= \mathbb{E}\left[\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n}\sum_{j\neq i} X_i X_j\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[X_i^2\right] + \sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}\left[X_i X_j\right] && \text{(Linearität)} \\
&= n \cdot \mathbb{E}\left[X_1^2\right] + n(n-1) \cdot \mathbb{E}\left[X_1 X_2\right] && \text{(Symmetrie)}
\end{aligned}$$

Indikatorvariablen IV

Sei A_i das Ereignis, dass der i-te Seemann im richtigen Bett liegt. Mit der Notation der Indikatorvariablen sei $X_i = I_{A_i}$. Dann gilt:

$$\mathbb{E}[X_1 X_2] = \mathbb{E}[I_{A_1} I_{A_2}] = \Pr[A_1 \cap A_2] = \frac{1}{n(n-1)}$$

sowie

$$\mathbb{E}[X_1^2] = 0^2 \cdot \Pr[\bar{A}_1] + 1^2 \cdot \Pr[A_1] = \Pr[A_1] = \frac{1}{n}\,.$$

Wir erhalten

$$\begin{aligned} \mathbb{E}[X^2] &= n \cdot \mathbb{E}\left[X_1^2\right] + n(n-1) \cdot \mathbb{E}\left[X_1 X_2\right] \\ &= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} = 2\,. \end{aligned}$$

und

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2 - 1 = 1.$$

Indikatorvariablen V

Einfacher Beweis für Satz 3 mit Hilfe von Indikatorvariablen:

Zur Erinnerung:

Satz 3 (Siebformel, Prinzip der Inklusion/Exklusion)

Für Ereignisse $A_1, \ldots, A_n$ $(n \geq 2)$ gilt:

$$\begin{aligned}\Pr\left[\bigcup_{i=1}^{n} A_i\right] &= \sum_{i=1}^{n} \Pr[A_i] - \sum_{1 \leq i_1 < i_2 \leq n} \Pr[A_{i_1} \cap A_{i_2}] + - \ldots \\ &+ (-1)^{l-1} \sum_{1 \leq i_1 < \ldots < i_l \leq n} \Pr[A_{i_1} \cap \ldots \cap A_{i_l}] + - \ldots \\ &+ (-1)^{n-1} \cdot \Pr[A_1 \cap \ldots \cap A_n]\,.\end{aligned}$$

Indikatorvariablen VI

Beweis: Sei $B := A_1 \cup \ldots \cup A_n$. Wir betrachten die Indikatorvariablen $I_i := I_{A_i}$ der Ereignisse $A_1, \ldots, A_n$ und die Indikatorvariable $I_{\bar{B}}$ des Ereignisses $\bar{B}$.

Es gilt

$$\prod_{i=1}^{n}(1 - I_i) = \begin{cases} 1 & \text{falls } I_1 = \ldots = I_n = 0, \\ & \text{d.h. wenn } B \text{ nicht eintritt} \\ 0 & \text{sonst} \end{cases} = I_{\bar{B}}$$

Wir erhalten:

$$I_{\bar{B}} = 1 - \sum_{1 \le i \le n} I_i + \sum_{1 \le i_1 < i_2 \le n} I_{i_1} I_{i_2} - \ldots + (-1)^n I_1 \cdot \ldots \cdot I_n.$$

Indikatorvariablen VII

Es folgt

$$
\begin{aligned}
\Pr[B] &= 1 - \Pr[\overline{B}] = 1 - \mathbb{E}[I_{\bar{B}}] && \text{(Lem.68(1))} \\
&= 1 - \mathbb{E}\left[1 - \sum_{1 \le i \le n} I_i + \ldots + (-1)^n I_1 \cdots I_n\right] \\
&= \sum_{1 \le i \le n} \mathbb{E}[I_i] - \sum_{1 \le i_1 < i_2 \le n} \mathbb{E}[I_{i_1} I_{i_2}] \\
&\quad + \ldots + (-1)^n \mathbb{E}[I_1 \cdot \ldots \cdot I_n] && \text{(Linear.)} \\
&= \sum_{1 \le i \le n} \Pr[A_i] - \sum_{1 \le i_1 < i_2 \le n} \Pr[A_{i_1} \cap A_{i_2}] \\
&\quad + \ldots + (-1)^n \Pr[A_1 \cap \ldots \cap A_n] && \text{(Lem.68(2))}
\end{aligned}
$$

10. Formelsammlung

Gesetze zum Rechnen mit Ereignissen I

Im Folgenden seien A und B, sowie $A_1, \ldots, A_n$ Ereignisse. Die Notation $A \uplus B$ steht für $A \cup B$ und zugleich $A \cap B = \emptyset$ (disjunkte Vereinigung). $A_1 \uplus \ldots \uplus A_n = \Omega$ bedeutet also, dass die Ereignisse $A_1, \ldots, A_n$ eine Partition der Ergebnismenge Ω bilden.

$\Pr[\emptyset] = 0$
$0 \leq \Pr[A] \leq 1$
$\Pr[\bar{A}] = 1 - \Pr[A]$
$A \subseteq B \implies \Pr[A] \leq \Pr[B]$

Gesetze zum Rechnen mit Ereignissen II

$\forall i \neq j : A_i \cap A_j = \emptyset \implies$ $\Pr\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \Pr[A_i]$	Additionssatz
$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$	Inklusion/Exklusion, Siebformel
$\Pr\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \Pr[A_i]$	Boolesche Ungleichung
$\Pr[A\|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$ für $\Pr[B] > 0$	Def. bedingte Ws.

Gesetze zum Rechnen mit Ereignissen III

$B \subseteq A_1 \uplus \ldots \uplus A_n \implies$ $\Pr[B] = \sum_{i=1}^{n} \Pr[B \mid A_i] \cdot \Pr[A_i]$	Satz von der totalen Wahrscheinlichkeit
$\Pr[B] > 0,\ B \subseteq A_1 \uplus \ldots \uplus A_n \implies$ $\Pr[A_i \mid B] = \dfrac{\Pr[B \mid A_i] \cdot \Pr[A_i]}{\sum_{i=1}^{n} \Pr[B \mid A_i] \cdot \Pr[A_i]}$	Satz von Bayes
$\Pr[A_1 \cap \ldots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2 \mid A_1] \cdot$ $\ldots \cdot \Pr[A_n \mid A_1 \cap \ldots \cap A_{n-1}]$	Multiplikationssatz
A und B unabhängig $\iff$ $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$	Definition Unabhängigkeit

Gesetze zum Rechnen mit Zufallsvariablen I

Sei X eine diskrete Zufallsvariable. Wenn $\mathbb{E}[X]$ und $\mathrm{Var}[X]$ existieren, dann gilt:

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] \\ &= \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega] \\ \Big(&= \sum_{i=1}^{\infty} \Pr[X \geq i], \quad \text{falls } W_X \subseteq \mathbb{N}_0 \Big) \end{aligned}$$

Erwartungswert

$$\begin{aligned} \mathrm{Var}[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ &= \sum_{x \in W_X} \Pr[X = x] \cdot (x - \mathbb{E}[X])^2 \end{aligned}$$

Varianz

Gesetze zum Rechnen mit Zufallsvariablen II

Seien a, b, a_1, ..., $a_n \in \mathbb{R}$, $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$.

$X_1, \ldots, X_n$ unabhängig	$\iff$	für alle $(a_1, \ldots, a_n)$: $\Pr[X_1 = a_1, \ldots, X_n = a_n]$ $= \Pr[X_1 = a_1] \cdot \ldots \cdot \Pr[X_n = a_n]$
$X_1, \ldots, X_n$ unabhängig	$\implies$	$f_1(X_1), \ldots, f_n(X_n)$ unabhängig

Gesetze zum Rechnen mit Zufallsvariablen III

$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X \mid A_i] \cdot \Pr[A_i]$
$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$
$\mathrm{Var}[a \cdot X + b] = a^2 \cdot \mathrm{Var}[X]$

Gesetze zum Rechnen mit Zufallsvariablen IV

$\mathbb{E}[a_1 X_1 + \ldots + a_n X_n]$ $= a_1 \mathbb{E}[X_1] + \ldots + a_n \mathbb{E}[X_n]$	Linearität des Erwartungswerts
$X_1, \ldots, X_n$ unabhängig $\Longrightarrow$ $\mathbb{E}[X_1 \cdot \ldots \cdot X_n] = \mathbb{E}[X_1] \cdot \ldots \cdot \mathbb{E}[X_n]$	Multiplikativität des Erwartungswerts
$X_1, \ldots, X_n$ unabhängig $\Longrightarrow$ $\mathrm{Var}[X_1 + \ldots + X_n] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]$	Varianz einer Summe

Wichtige Reihen

$\sum_{k=0}^{\infty} 1 \cdot x^k$	$=$	$\frac{1}{1-x}$	$\lvert x \rvert < 1$	Geometrische Reihe
$\sum_{k=1}^{\infty} k \cdot x^{k-1}$	$=$	$\frac{1}{(1-x)^2}$	$\lvert x \rvert < 1$	Ableitung der geometrische Reihe
$\sum_{k=0}^{\infty} \binom{n}{k} \cdot x^k$	$=$	$(1+x)^n$	$\lvert x \rvert < 1$	Binomialreihe
$\sum_{k=0}^{\infty} \frac{1}{k!} \cdot x^k$	$=$	e^x		Exponentialreihe

11. Wichtige Verteilungen

Erwartungswert und Varianz der Binomialverteilung

Sei $X \sim \mathrm{Bin}(n, p)$. Wir haben $\mathbb{E}[X] = np$ berechnet. Für die Varianz haben wir noch keinen Ausdruck berechnet.

Sind $X_1, \ldots, X_n$ unabhängig Bernoulli-verteilte Zufallsvariablen mit der gleichen Erfolgsw'keit p, dann gilt $X_1 + \ldots + X_n \sim \mathrm{Bin}(n, p)$. Mit der Linearität des Erwartungswertes folgt:

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n] = np$$

Mit der Rechenregel für die varianz einer Summe von unabhängigen Zufalsvariablen folgt:

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n] = npq$$

Summe von Binomialverteilugen

Satz 70

Wenn $X \sim \mathrm{Bin}(n_x, p)$ *und* $Y \sim \mathrm{Bin}(n_y, p)$ *unabhängig sind, dann gilt* $(X + Y) \sim \mathrm{Bin}(n_x + n_y, p)$.

Beweis:

Seien $X_1, \ldots, X_{n_x}, Y_1, \ldots, Y_{n_y}$ unabhängige Bernoulli-verteilte Zufallsvariablen mit der gleichen Erfolgsw'keit p. Es gilt

$$X_1 + \ldots + X_{n_x} \sim \mathrm{Bin}(n_x, p) \qquad Y_1 + \ldots + Y_{n_y} \sim \mathrm{Bin}(n_y, p)$$

und

$$X_1 + \ldots + X_{n_x} + Y_1 + \ldots + Y_{n_y} \sim \mathrm{Bin}(n_x + n_y, p)\ .$$

$\square$

Gedächtnislosigkeit der geometrischen Verteilung I

Beispiel 71

Bei mündlichen Prüfungen stellt Prof. Evilsparza irgendeine beliebige Frage und, während der Kandidat antwortet, denkt er an seinen nächsten Urlaub auf den Malediven. Anschließend wirft er eine faire Würfel, und der Kandidat besteht nur dann, wenn der Würfel eine 1 zeigt. An einem Tag prüft Prof. Evilsparza Kandidaten nur, bis einer besteht (danach hat er keine Lust mehr).

Heute prüft Prof. Evilsparza. Kandidat Nr. 8 kommt in den Warteraum und erfährt, dass bisher 3 Kandidaten geprüft wurden und alle durchgefallen sind.

Frage: Mit welcher W'keit wird er noch geprüft ?

Gedächtnislosigkeit der geometrischen Verteilung II

Modell: $\Omega = F^*B \qquad \Pr[F^kB] = \left(\frac{5}{6}\right)^k \left(\frac{1}{6}\right)$

Sei X die Zufallsvariable, die angibt, wieviele Kandidaten insgesamt geprüft werden (z.B. gilt $X(F^kB) = k+1$). Gesucht wird:

$$\Pr[X > 7 \mid X > 3]$$

Intuitiv ist die Situation im Experiment äquivalent zu: Kandidat Nr. 5 kommt in den Warteraum und erfährt, dass noch kein Kandidat geprüft wurde.

In diesem Fall ist die gesuchte W'keit $\Pr[X > 4]$, und daher erwarten wir

$$\Pr[X > 7 \mid X > 3] = \Pr[X > 4]$$

und im Allgemeinen:

$$\Pr[X > x + y \mid X > x] = \Pr[X > y]\ .$$ Stimmt es?

Gedächtnislosigkeit der geometrischen Verteilung III

Die Variable X is geometrisch verteilt mit Erfolgsw'keit $p = 1/6$.

$$\begin{aligned}
\Pr[X > y] &= \sum_{i=y+1}^{\infty} \Pr[X = i] = \sum_{i=y+1}^{\infty} (1-p)^{i-1} p \\
&= (1-p)^y p \cdot \sum_{i=0}^{\infty} (1-p)^i \\
&= (1-p)^y p \cdot \frac{1}{1-(1-p)} = (1-p)^y \\
\Pr[X > y + x \mid X > x] &= \frac{\Pr[X > y+x, X > x]}{\Pr[X > x]} \\
&= \frac{\Pr[X > y+x]}{\Pr[X > x]} = \frac{(1-p)^{y+x}}{(1-p)^x} \\
&= (1-p)^y
\end{aligned}$$

Gedächtnislosigkeit der geometrischen Verteilung IV

Definition 72

Eine Zufallsvariable X heißt gedächtnislos wenn

$$\Pr[X > x + y \mid X > x] = \Pr[X > y]$$

gilt für alle $x, y \in \mathbb{R}$.

Intuition: wenn eine Variable gedächtnislos ist, dann kann man die Vergangenheit „vergessen", da sie keine nützliche Information für die Zukunft gibt.

Lemma 73

Geometrisch verteilte Zufallsvariablen sind gedächtnislos.

Es gilt sogar die Umkehrung: eine gedächtnislose, Zufallsvariable X mit $W_X = \mathbb{N}_0$ ist geometrisch verteilt!

Beispiel 74 (Das Coupon-Collector-Problem)

Manche Firmen legen den Verpackungen ihrer Produkte oft kleine Bilder bei, um den Käufer zum Sammeln anzuregen.

Ein Kind kauft Packungen so lange ein, bis es die vollständige Sammlung besitzt. Wir nehmen an, dass jede Packung nur ein Bild enthält und bei jedem Kauf jedes Bild mit gleicher Wahrscheinlichkeit auftritt.

Frage: Wie viele Packungen muss das Kind im Mittel erwerben?

Gedächtnislosigkeit der geometrischen Verteilung VI

Modell bei n Beilagen:

$$\begin{aligned}\Omega &= \{w\,k \mid \; w \in \{1,\ldots,n\}^*, k \in \{1,\ldots,n\}, \text{ und} \\ &\qquad\qquad \text{alle Zahlen bis auf } k \text{ kommen in } w \text{ vor}\} \\ \Pr[w\,k] &= \left(\frac{1}{n}\right)^{|w|+1}\end{aligned}$$

Sei X die Zufallsvariable, die angibt, wie viele Bilder gekaut werden (z.B. $X(1122123) = 7$). Gesucht wird $\mathbb{E}[X]$.

Gedächtnislosigkeit der geometrischen Verteilung VII

Bezeichne Phase i die Schritte vom Erwerb des $(i-1)$-ten Bilds (ausschließlich) bis zum Erwerb der i-ten Bilds (einschließlich).

Sei etwa $n = 4$, und seien die Beilagen mit den Zahlen $1, 2, 3, 4$ identifiziert. Ein Experiment ist z.B.:

$$\underbrace{2}_{1}, \underbrace{2, 1}_{2}, \underbrace{2, 2, 3}_{3}, \underbrace{1, 3, 2, 3, 1, 4}_{4} .$$

Beobachtung: Phase i endet genau dann, wenn wir eine der $n - i + 1$ Beilagen erhalten, die wir noch nicht besitzen.

Sei X_i die variable, die angibt, wieviele Bilder in der i-ten Phase gekauft werden. Mit der Gedächtnislosigkeit der geometrischen Verteilung folgt, dass X_i geometrisch verteilt mit

$$p = \frac{n - i + 1}{n} \qquad \mathbb{E}[X_i] = \frac{n}{n - i + 1} .$$

Gedächtnislosigkeit der geometrischen Verteilung VIII

Wir haben

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=1}^{n} \mathbb{E}[X_i] \\ &= \sum_{i=1}^{n} \frac{n}{n-i+1} \\ &= n \cdot \sum_{i=1}^{n} \frac{1}{i} = n \cdot H_n,\end{aligned}$$

wobei $H_n := \sum_{i=1}^{n} \frac{1}{i}$ die n-te harmonische Zahl bezeichnet.

Da $H_n = \ln n + O(1) + O\left(\frac{1}{n}\right)$, folgt

$$\mathbb{E}[X] = n \ln n + O(n) \; .$$

Negative Binomialverteilung I

Beispiel 75

Eine Münze wird solange geworfen, bis zum n-te Mal Kopf vorkommt. Die W'keit von Kopf beträgt p.

Frage: Mit welcher W'keit wird die Münze k-Mal geworfen?

Modell:

$$\begin{aligned} \Omega &= \{w \in \{K, Z\}^* K \mid K \text{ kommt in } w\ (n-1)\text{-Mal vor}\} \\ \Pr[w] &= p^n q^{|w|-n} \end{aligned}$$

Sei Z die Zufallsvariable, die angibt, wie oft insgesamt geworfen wird. Gesucht wird die Dichte $f_Z(k)$.

Negative Binomialverteilung II

Falls $Z(w) = k$ ist, so kommt in w genau n-Mal Kopf und $(k-n)$-Mal Zahl.

Dafür gibt es genau

$$\binom{k-1}{n-1}$$

Möglichkeiten, von denen jede mit W'keit

$$p^n (1-p)^{k-n}$$

eintritt. Es gilt also

$$f_Z(k) = \binom{k-1}{n-1} \cdot p^n (1-p)^{k-n} \, .$$

Die Zufallsvariable Z nennt man negativ binomialverteilt mit Ordnung n.

Hypergeometrische Verteilung I

Beispiel 76

In der Legislaturperiode 2008-2013 hat der bayerischer Landtag 187 Abgeordnete, 92 davon bilden die CSU-Fraktion. 23 Abgeordnete haben Familienangehörige beschäftigt. 17 davon gehören zur CSU-Fraktion.

Frage: Kann die Anhäufung bei der CSU ein Zufall sein?

Mathematische Formulierung: 187 Bälle in einem Sack. 92 Bälle sind schwarz, der Rest weiß. 23 Bälle werden zufällig gezogen (ohne Zurücklegen). Sei X die Zufallsvariable, die die Anzahl der schwarzen Bällen, die gezogen wurden, angibt. Gesucht ist $\Pr[X \geq 17]$.

$$\Pr[X = k] = \frac{\binom{92}{k}\binom{95}{23-k}}{\binom{187}{23}} \qquad \Pr[X \geq 17] = \sum_{k=17}^{23} \frac{\binom{92}{k}\binom{95}{23-k}}{\binom{187}{23}} \approx 0.0097$$

Hypergeometrische Verteilung II

Beispiel 77

In der Legislaturperiode 2008-2013 hat der bayerischer Landtag 187 Abgeordnete. 15 Abgeordnete bilden die FDP-Fraktion. 23 Abgeordnete haben Familienangehörige beschäftigt. Keiner gehört zur FDP-Fraktion.

Frage: Kann die „weisse Weste“ der FDP ein Zufall sein?

Mathematische Formulierung: 187 Bälle in einem Sack. 15 Bälle sind gelb, der Rest weiß. 23 Bälle werden zufällig gezogen (ohne zurücklegen). Sei Y die Zufallsvariable, die die Anzahl der gelben Bällen, die gezogen wurden, angibt. Gesucht ist $\Pr[Y \leq 0] = \Pr[Y = 0]$.

$$\Pr[Y = k] = \frac{\binom{15}{k}\binom{172}{23-k}}{\binom{187}{23}} \qquad \Pr[Y = 0] = \frac{\binom{172}{23}}{\binom{187}{23}} = 0.128$$

Hypergeometrische Verteilung III

Eine Zufallsvariable X mit $W_X = \mathbb{N}_0$ ist hypegeometrisch verteilt mit Parametern N, M, n wenn ihre Dichte folgende Gestalt hat:

$$f_X(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Es gilt $\mathbb{E}[X] = n\frac{M}{N}$

Beispiel 78

Beim Kundenservice einer Firma rufen im Durschnitt k Kunden pro Tag an.

Wir betrachten ein diskretes Modell, in dem der Tag in $n \geq k$ gleich lange Zeitintervallen unterteilt wird (jeweils $24/n$ Stunden lang). Wir nehmen an, dass Kunden in jedem Intervall mit der selben W'keit anrufen und dass höchstens ein Kunde pro Zeitintervall anruft.

Damit ruft ein Kunde in ein Intervall mit W'keit $\dfrac{k}{n}$ an. Die Anzahl X der Anrufe an einem Tag ist binomialverteilt:

$$\Pr[X \leq a] = \sum_{i=0}^{a} \binom{n}{i} \left(\frac{k}{n}\right)^{i} \left(1 - \frac{k}{n}\right)^{(n-i)}$$

Poisson-Verteilung II

Die folgende Tabelle zeigt $\Pr[X \leq 5]$ für $k = 3$ und verschiedene Werte von n:

n	$\Pr[X \leq 5]$
5	1
6	0.9844
8	0.9640
24	0.9297
$24 * 60$	0.9163

Frage: Zu welchem Wert konvergiert diese Folge?

Poisson-Verteilung III

Sei X_n die Anzahl der Anrufe bei einer Unterteilung in n Zeitabschnitten. Es gilt

$$X_n \sim \mathrm{Bin}(n, p_n) \qquad \text{mit } p_n = \lambda/n \,.$$

Für ein beliebiges a mit $0 \le a \le n$ ist die W'keit $b(a; n, p_n)$, dass X_n den Wert a annimmt, gleich

$$\begin{aligned} b(a; n, p_n) &= \binom{n}{a} \cdot p_n^a \cdot (1 - p_n)^{n-a} \\ &= \frac{n^{\underline{a}}}{a!} \cdot \frac{\lambda^a}{n^a} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-a} \\ &= \frac{\lambda^a}{a!} \cdot \frac{n^{\underline{a}}}{n^a} \cdot \left(1 - \frac{\lambda}{n}\right)^{n} \cdot \left(1 - \frac{\lambda}{n}\right)^{-a} \,. \end{aligned}$$

Poisson-Verteilung IV

Wir betrachten nun $n \to \infty$ und erinnern uns, dass

$$\lim_{n\to\infty} \frac{n^{\underline{k}}}{n^k} = 1 \quad \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda} \quad \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-k} = 1$$

Damit folgt

$$\begin{aligned}
\lim_{n\to\infty} b(k; n, p_n) &= \lim_{n\to\infty} \left(\frac{\lambda^a}{a!} \cdot \frac{n^{\underline{a}}}{n^a} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-a} \right) \\
&= \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda} \cdot 1 \\
&= \frac{e^{-\lambda}\lambda^k}{k!}.
\end{aligned}$$

Poisson-Verteilung V

Definition 79

Eine Zufallsvariable X mit $W_X = \mathbb{N}_0$ und Dichte

$$f_X(i) = \frac{e^{-\lambda}\lambda^i}{i!} \quad \text{für alle } i \in \mathbb{N}_0\text{, wobei } \lambda > 0$$

heißt Poisson-verteilt mit dem Parameter λ. Wir schreiben $X \sim \mathrm{Po}(\lambda)$.

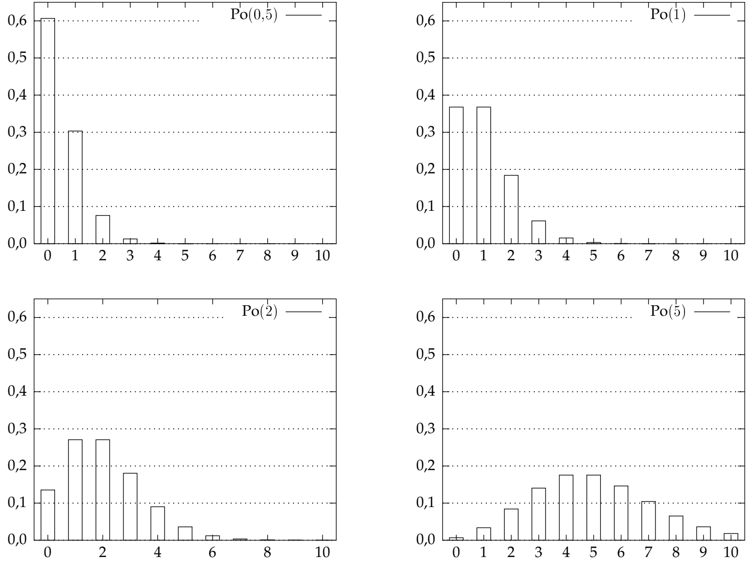
f_X ist eine zulässige Dichte, da

$$\sum_{i=0}^{\infty} f_X(i) = \sum_{i=0}^{\infty} \frac{e^{-\lambda}\lambda^i}{i!} = e^{-\lambda} \cdot e^{\lambda} = 1\,.$$

Es gilt also

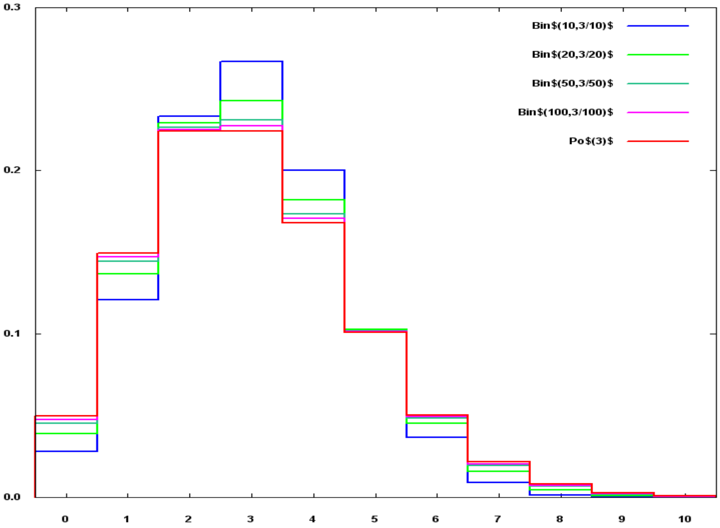
$$\lim_{n\to\infty} b(k; n, \frac{\lambda}{n}) = f_X(k) \qquad \text{für } X \sim \mathrm{Po}(\lambda).$$

Poisson-Verteilung VI



Dichte der Poisson-Verteilung

Poisson-Verteilung VII



Vergleich von Binomial- und Poisson-Verteilung

Poisson-Verteilung VIII

Aus

$$\lim_{n\to\infty} b(k; n, \frac{\lambda}{n}) = f_X(k) \qquad \text{für } X \sim \mathrm{Po}(\lambda).$$

folgt:

Ist p in Vergleich zu n hinreichend klein, so kann man $\mathrm{Po}(np)$ als Approximation der Binomialverteilung $\mathrm{Bin}(n,p)$ verwenden.

Diese Tatsache wird als Gesetz seltener Ereignisse bezeichnet, da die Wahrscheinlichkeit eines einzelnen Treffers relativ klein sein muss, wenn die Approximation gute Ergebnisse liefern soll.

Poisson-Verteilung IX

Die folgenden Voraussetzungen müssen erfüllt sein, damit die Annahme der Poisson-Verteilung gerechtfertigt ist:

- Die Ereignisse treten nie zur gleichen Zeit auf.
- Die Wahrscheinlichkeit, dass ein Ereignis in einem (kleinen) Zeitintervall δt auftritt, ist proportional zur Länge von δt.
- Die Anzahl der Ereignisse in einem festen Zeitintervall hängt nur von dessen Länge ab, nicht aber von der Lage auf der Zeitachse.
- Wenn man zwei disjunkte Zeitintervalle betrachtet, so sind die Anzahlen der Ereignisse in diesen Zeiträumen voneinander unabhängig.

Beispiel 80 (FRM II in Bezug auf Erdbeben in Japan)

FRM II ist analog zu allen deutschen KKWs nach einem Ermessenserdbeben nach der MSK Skala ausgelegt. Die Bemessungsintensität für FRM II lautet I(MSK) = VI–VII mit Eintrittswahrscheinlichkeit $10^{-5}/a$.
(VI: leichte Verputzschäden an Gebäuden. VII: Risse im Verputz, in Wänden und an Schornsteinen.5-0-6.3 auf der Richterskala).

In der seismographischen Region Bayerische Molasse sind aus den vergangenen Jahrhunderten insgesamt 6 tektonische Erdbeben berichtet. Das stärkste mit I(MSK)=VI ereignete sich am 9.10.1935 bei St. Martin in Österreich 135 km östlich von Garching. Dieses und alle anderen Erdbeben lösten nur sehr schwache Bodenbewegungen in Garching aus mit I(MSK) = III–IV.
(III: nur von wenigen Personen gespürt. IV: von vielen Personen gespürt; Geschirr und Fenster klirren.)

Poisson-Verteilung XI

Frage: Mit welcher Wahrscheinlichkeit wird Garching in 2014 einmal oder mehr von einem Erdbeben mit I(MSK)≥VI–VII heimgesucht?

Die Anzahl X der Katastrophen in Garching wird durch eine Poisson-Verteilung mit Parameter $\lambda = 10^{-5}$ pro Jahr modelliert. (Vergleich: nach Wikipedia $\lambda \approx 1000$ für die ganze Welt.)

Damit gilt

$$
\begin{aligned}
\Pr[X \geq 1] &= 1 - \Pr[X = 0] = 1 - e^{-\lambda} \\
&\approx 1 - 0{,}99999000005 \approx 10^{-5}. \\
\Pr[X \geq 2] &= 1 - \Pr[X = 0] - \Pr[X = 1] = 1 - e^{-\lambda} - \lambda e^{-\lambda} \\
&\approx 1 - 0{,}99999000005 - 0{,}00000999990 = 5 \cdot 10^{-11}.
\end{aligned}
$$

Poisson-Verteilung XII: Erwartungswert

Erwartungswert und Varianz

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{i=0}^{\infty} i \cdot \frac{e^{-\lambda}\lambda^i}{i!} \\
&= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} \\
&= \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \\
&= \lambda e^{-\lambda} e^{\lambda} = \lambda\,.
\end{aligned}
$$

Poisson-Verteilung XIII: Varianz

Es gilt

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\
&= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2
\end{aligned}
$$

und

$$
\begin{aligned}
\mathbb{E}[X(X-1)] &= \sum_{i=0}^{\infty} i(i-1) \cdot \frac{e^{-\lambda}\lambda^i}{i!} \\
&= \lambda^2 e^{-\lambda} \sum_{i=2}^{\infty} \frac{\lambda^{i-2}}{(i-2)!} = \lambda^2 e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \\
&= \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2
\end{aligned}
$$

Damit erhalten wir

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

Poisson-Verteilung XIV: Summe

Satz 81

Sind X und Y unabhängige Zufallsvariablen mit $X \sim \mathrm{Po}(\lambda)$ und $Y \sim \mathrm{Po}(\mu)$, dann gilt

$$Z := X + Y \sim \mathrm{Po}(\lambda + \mu)\,.$$

Poisson-Verteilung XV: Summe

Beweis:

$$
\begin{aligned}
f_Z(z) &= \sum_{x=0}^{\infty} f_X(x) \cdot f_Y(z-x) = \sum_{x=0}^{z} \frac{e^{-\lambda}\lambda^x}{x!} \cdot \frac{e^{-\mu}\mu^{z-x}}{(z-x)!} \\
&= e^{-(\lambda+\mu)} \cdot \frac{(\lambda+\mu)^z}{z!} \cdot \sum_{x=0}^{z} \frac{z!}{x!(z-x)!} \left(\frac{\lambda}{\lambda+\mu}\right)^x \left(\frac{\mu}{\lambda+\mu}\right)^{z-x} \\
&= e^{-(\lambda+\mu)} \cdot (\lambda+\mu)^z \frac{1}{z!} \cdot \sum_{x=0}^{z} \binom{z}{x} p^x (1-p)^{z-x},
\end{aligned}
$$

wobei $p := \dfrac{\lambda}{\lambda+\mu}$.

Da die Summe gleich 1 ist, folgt

$$
f_Z(z) = e^{-(\lambda+\mu)} \cdot (\lambda+\mu)^z \frac{1}{z!} \,.
$$

12. Abschätzen von Wahrscheinlichkeiten

Markov-Ungleichung I

Satz 82 (Markov-Ungleichung)

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}.$$

Äquivalent dazu:

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq 1/t\,.$$

Markov-Ungleichung II

Beweis:

$$\begin{aligned} t \cdot \Pr[X \geq t] &= t \cdot \sum_{x \in W_X,\, x \geq t} \Pr[X = x] \\ &\leq \sum_{x \in W_X,\, x \geq t} x \cdot \Pr[X = x] \\ &\leq \sum_{x \in W_X} x \cdot \Pr[X = x] \\ &= \mathbb{E}[X]\,. \end{aligned}$$

$\square$

Markov-Ungleichung III

Alternativer Beweis:

Es gilt

$$\mathbb{E}[X] = \mathbb{E}[X|X < t] \cdot \Pr[X < t] + \mathbb{E}[X|X \geq t] \cdot \Pr[X \geq t]\,.$$

Wegen $\mathbb{E}[X|X < t] \cdot \Pr[X < t] \geq 0$ und $\mathbb{E}[X|X \geq t] \geq t$ folgt sofort

$$\mathbb{E}[X] \geq t \cdot \Pr[X \geq t]\,.$$

Markov-Ungleichung IV

Die Markov-Ungleichung ist nach Andrey Andreyevich Markov (1856–1922) benannt, der an der Universität von St. Petersburg bei Chebyshev studierte und später dort arbeitete. Neben seiner mathematischen Tätigkeit fiel Markov durch heftige Proteste gegen das Zaren-Regime auf, und nur sein Status als vermeintlich harmloser Akademiker schützte ihn vor Repressalien durch die Behörden. Im Jahr 1913 organisierte er parallel zum dreihundertjährigen Geburtstag der Zarenfamilie Romanov eine Feier zum zweihundertjährigen Geburtstag des Gesetzes der großen Zahlen (s.u.).

Chebyshev-Ungleichung I

Die folgende Abschätzung ist nach Pavnuty Lvovich Chebyshev (1821–1894) benannt, der ebenfalls an der Staatl. Universität in St. Petersburg wirkte.

Satz 83 (Chebyshev-Ungleichung)

Sei X eine Zufallsvariable, und sei $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;\leq\; \frac{\mathrm{Var}[X]}{t^2}\,.$$

Äquivalent dazu:

$$\Pr[|X - \mathbb{E}[X]| \geq t\sqrt{\mathrm{Var}[X]}] \;\leq\; 1/t^2\,.$$

Chebyshev-Ungleichung II

Beweis:

Wir stellen fest, dass

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;=\; \Pr[(X - \mathbb{E}[X])^2 \geq t^2]\,.$$

Setze

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

Dann gilt $\mathbb{E}[Y] = \mathrm{Var}[X]$, und damit mit der Markov-Ungleichung:

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;=\; \Pr[Y \geq t^2] \;\leq\; \frac{\mathbb{E}[Y]}{t^2} \;=\; \frac{\mathrm{Var}[X]}{t^2}\,.$$

$\square$

Chebyshev-Ungleichung III

Beispiel 84

Wir werfen 1000-mal eine ideale Münze und ermitteln die Anzahl X der Würfe, in denen „Kopf" fällt.

Frage: Wie groß ist die W'keit, dass mehr als 550-mal „Kopf" fällt?

Mit $X \sim \mathrm{Bin}(1000, \frac{1}{2})$ gilt

$$\mathbb{E}[X] = \frac{1}{2}n = 500 \qquad \text{und} \qquad \mathrm{Var}[X] = \frac{1}{4}n = 250.$$

Chebyshev-Ungleichung $\Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$:

$$\begin{aligned} \Pr[X \geq 550] &\leq \Pr[|X - 500| \geq 50] \\ &\leq \frac{250}{50^2} = 0{,}1\,. \end{aligned}$$

Chebyshev-Ungleichung IV

Setze nun $n = 10000$ und betrachte wieder eine maximal 10%-ige Abweichung vom Erwartungswert:

$$\mathbb{E}[X] = 5000 \qquad \text{und} \qquad \mathrm{Var}[X] = 2500$$

$$\begin{aligned} \Pr[X \geq 5500] &\leq \Pr[|X - 5000| \geq 500] \\ &\leq \frac{2500}{500^2} = 0{,}01\,. \end{aligned}$$

Satz 85 (Gesetz der großen Zahlen)

Sei X eine Zufallsvariable und seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen mit derselben Verteilung wie X.

Sei

$$Z := \frac{X_1 + \ldots + X_n}{n},$$

Seien $\varepsilon, \delta > 0$ beliebig aber fest. Dann gilt für alle $n \geq \frac{\mathrm{Var}[X]}{\varepsilon\delta^2}$:

$$\Pr[\, |Z - \mathbb{E}[X]| \geq \delta \,] \;\leq\; \varepsilon.$$

Gesetz der großen Zahlen II

Beweis:

Für Z gilt

$$\mathbb{E}[Z] = \frac{1}{n} \cdot (\mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n]) = \frac{1}{n} \cdot n \cdot \mathbb{E}[X] = \mathbb{E}[X],$$

sowie (Unabhängigkeit)

$$\mathrm{Var}[Z] = \frac{1}{n^2} \cdot (\mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]) = \frac{1}{n^2} \cdot n \cdot \mathrm{Var}[X] = \frac{\mathrm{Var}[X]}{n}.$$

Mit der Chebyshev-Ungleichung erhalten wir

$$\Pr[|Z - \mathbb{E}[X]| \geq \delta] = \Pr[|Z - \mathbb{E}[Z]| \geq \delta] \leq \frac{\mathrm{Var}[Z]}{\delta^2} = \frac{\mathrm{Var}[X]}{n\delta^2} \leq \varepsilon,$$

nach Wahl von n. $\square$

Gesetz der großen Zahlen III

Sei X eine Indikatorvariable für ein Ereignis A mit $\Pr[A] = p$.

Somit ist X Bernoulli-verteilt mit $\mathbb{E}[X] = p$.

$Z = \frac{1}{n}(X_1 + \ldots + X_n)$ gibt die relative Häufigkeit an, mit der A bei n Wiederholungen des Versuchs eintritt, denn

$$Z = \frac{\text{Anzahl der Versuche, bei denen } A \text{ eingetreten ist}}{\text{Anzahl aller Versuche}}.$$

Mit Hilfe des obigen Gesetzes der großen Zahlen folgt

$$\Pr[\,|Z - p| \geq \delta\,] \leq \varepsilon,$$

für genügend großes n.

Also nähert sich die relative Häufigkeit von A bei hinreichend vielen Wiederholungen des Experiments mit beliebiger Sicherheit und beliebige Genauigkeit an die „wahre“ Wahrscheinlichkeit p an.

Gesetz der großen Zahlen IV

Die obige Variante eines Gesetzes der großen Zahlen geht auf Jakob Bernoulli zurück, der den Satz in seinem Werk ars conjectandi zeigte.

Es soll betont werden, dass das Gesetz der großen Zahlen die

$$\text{relative Abweichung } \left|\frac{1}{n}\sum_i X_i - p\right|$$

und nicht die

$$\text{absolute Abweichung } \left|\sum_i X_i - np\right|$$

abschätzt!

Chernoff-Schranken I

Die hier betrachtete Art von Schranken ist nach *Herman Chernoff* (*1923) benannt. Sie finden in der komplexitätstheoretischen Analyse von Algorithmen eine sehr häufige Verwendung.

Satz 86

Seien $X_1, \ldots, X_n$ *unabhängige Bernoulli-verteilte Zufallsvariablen mit* $\Pr[X_i = 1] = p_i$ *und* $\Pr[X_i = 0] = 1 - p_i$.

Seien $X := \sum_{i=1}^{n} X_i$ *und* $\mu := \mathbb{E}[X] = \sum_{i=1}^{n} p_i$.

Für jedes $\delta > 0$ *gilt*

$$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \right)^{\mu}.$$

Chernoff-Schranken II

Beweis:

Für $t > 0$ gilt $X \geq (1+\delta)\mu$ iff $e^{tX} \geq e^{t(1+\delta)\mu}$ und so

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}]\,.$$

Mit der Markov-Ungleichung folgt

$$\Pr[X \geq (1+\delta)\mu] = \Pr[e^{tX} \geq e^{t(1+\delta)\mu}] \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}}.$$

Wegen der Unabhängigkeit der Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\exp\left(\sum_{i=1}^{n} tX_i\right)\right] = \mathbb{E}\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}].$$

Chernoff-Schranken III

Weiter ist für $i \in \{1, \ldots, n\}$:

$$\mathbb{E}[e^{tX_i}] = e^{t\cdot 1} p_i + e^{t\cdot 0}(1 - p_i) = e^t p_i + 1 - p_i = 1 + p_i(e^t - 1)\,,$$

und damit

$$\begin{aligned}
\Pr[X \geq (1+\delta)\mu] &\leq \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&\leq \frac{\prod_{i=1}^{n}\exp(p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&= \frac{\exp(\sum_{i=1}^{n} p_i(e^t - 1))}{e^{t(1+\delta)\mu}} = \frac{e^{(e^t - 1)\mu}}{e^{t(1+\delta)\mu}} =: f(t)\,.
\end{aligned}$$

Wir wählen nun t so, dass $f(t)$ minimiert wird, nämlich

$$t := \ln(1+\delta)\,.$$

und erhalten

$$\Pr[X \geq (1+\delta)\mu] \leq \frac{e^{(e^t - 1)\mu}}{e^{t(1+\delta)\mu}} = \frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}}\,.$$

Chernoff-Schranken III

Weiter ist für $i \in \{1, \ldots, n\}$:

$$\mathbb{E}[e^{tX_i}] = e^{t \cdot 1} p_i + e^{t \cdot 0}(1 - p_i) = e^t p_i + 1 - p_i = 1 + p_i(e^t - 1)\,,$$

und damit

$$\begin{aligned}
\Pr[X \geq (1+\delta)\mu] &\leq \frac{\prod_{i=1}^n (1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&\leq \frac{\prod_{i=1}^n \exp(p_i(e^t - 1))}{e^{t(1+\delta)\mu}} \\
&= \frac{\exp(\sum_{i=1}^n p_i(e^t - 1))}{e^{t(1+\delta)\mu}} = \frac{e^{(e^t - 1)\mu}}{e^{t(1+\delta)\mu}} =: f(t)\,.
\end{aligned}$$

Wir wählen nun t so, dass $f(t)$ minimiert wird, nämlich

$$t := \ln(1+\delta)\,.$$

und erhalten

$$\Pr[X \geq (1+\delta)\mu] \leq \frac{e^{(e^t - 1)\mu}}{e^{t(1+\delta)\mu}} = \frac{e^{\delta\mu}}{(1+\delta)^{(1+\delta)\mu}}\,.$$

Chernoff-Schranken IV

Beispiel 87

Wir werfen eine faire Münze n-mal. Wir suchen eine obere Schranke für die W'keit, mit der „Kopf"

$$\frac{n}{2}\,(1 + 10\%) \quad \text{Mal}$$

oder öfter fällt.

n	Chebyshev	Chernoff
1000	$0{,}1$	$0{,}0889$
10000	$0{,}01$	$0{,}308 \cdot 10^{-10}$
n	$\frac{\frac{1}{4}n}{(0{,}1 \cdot \frac{1}{2}n)^2}$	$\left(\frac{e^{0{,}1}}{(1+0{,}1)^{1+0{,}1}}\right)^{n/2}$
	$= \frac{100}{n}$	$\approx 0.9975^n$

Satz 88

Seien $X_1, \ldots, X_n$ *unabhängige Bernoulli-verteilte Zufallsvariablen mit* $\Pr[X_i = 1] = p_i$ *und* $\Pr[X_i = 0] = 1 - p_i$.

Seien $X := \sum_{i=1}^{n} X_i$ *und* $\mu := \mathbb{E}[X] = \sum_{i=1}^{n} p_i$.

Für jedes $0 < \delta < 1$ *gilt*

$$\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^{\mu}.$$

Beweis:

Analog zum Beweis von Satz 86. $\square$

Chernoff-Schranken VI

Abschätzungen, wie sie in Satz 86 und Satz 88 angegeben sind, nennt man auch

tail bounds,

da sie Schranken für die tails, also die vom Erwartungswert weit entfernten Bereiche angeben.

Man spricht hierbei vom

upper tail (vgl. Satz 86) und vom lower tail (vgl. Satz 88).

Die Chernoff-Schranken hängen exponentiell von μ ab!

Chernoff-Schranken VII

Lemma 89

Für $0 \leq \delta < 1$ *gilt*

$$(1-\delta)^{1-\delta} \geq e^{-\delta+\delta^2/2} \quad \textit{und} \quad (1+\delta)^{1+\delta} \geq \delta + \delta^2/3.$$

Beweis:

Nur für die erste Ungleichung. Sei

$$\begin{aligned} f(\delta) &= \ln\left((1-\delta)^{1-\delta}\right) = (1-\delta)\ln(1-\delta) \\ g(\delta) &= \ln\left(e^{-\delta+\delta^2/2}\right) = -\delta + \delta^2/2 \end{aligned}$$

Es reicht zu zeigen $f(\delta) \geq g(\delta)$ für $0 \leq \delta < 1$.

Chernoff-Schranken V

Mit $f(0) = 0 = g(0)$ reicht es sogar $f'(\delta) \geq f'(\delta)$, d.h.

$$((1-\delta)\ln(1-\delta))' = -1 - \ln(1-\delta) \geq -1 + \delta = \left(-\delta + \delta^2/2\right)'$$

Wir haben

$$\begin{array}{rrcl} & -1 - \ln(1-\delta) & \geq & -1 + \delta \\ \iff & \ln(1-\delta) & \leq & -\delta \\ \iff & (1-\delta) & \leq & e^{-\delta} \end{array}$$

Die letzte Ungleichung folgt aus

$$e^{-x} = 1 - x + \frac{1}{2}x^2 - \cdots$$

Chernoff-Schranken VIII

Korollar 90

Seien $X_1, \ldots, X_n$ *unabhängige Bernoulli-verteilte Zufallsvariablen mit* $\Pr[X_i = 1] = p_i$ *und* $\Pr[X_i = 0] = 1 - p_i$*. Dann gelten folgende Ungleichungen für* $X := \sum_{i=1}^n X_i$ *und* $\mu := \mathbb{E}[X] = \sum_{i=1}^n p_i$*:*

1. $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1,81$*,*
2. $\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}$ *für alle* $0 < \delta \leq 1$*,*
3. $\Pr[|X - \mu| \geq \delta\mu] \leq 2e^{-\mu\delta^2/3}$ *für alle* $0 < \delta \leq 1$*,*
4. $\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ *und*
5. $\Pr[X \geq t] \leq 2^{-t}$ *für* $t \geq 2e\mu$*.*

Chernoff-Schranken IX

Beweis:

1 und 2 folgen direkt aus Satz 86 bzw. 88 und Lemma 89.

Aus 1 und 2 zusammen folgt 3.

Die Abschätzung 4 erhalten wir direkt aus Satz 86, da für den Zähler gilt

$$e \leq e^{(1+\delta)}\,.$$

5 folgt aus 4, indem man $t = (1+\delta)\mu$ setzt, $t \geq 2e\mu$:

$$\left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu} \leq \left(\frac{e}{t/\mu}\right)^{t} \leq \left(\frac{1}{2}\right)^{t}\,.$$

$\square$

Beispiel 91

Wir werfen $c \cdot n$ Bälle unabhängig und gleichverteilt in n Körbe.

Frage: Wie groß sollen die Körbe sein, so dass mit großer W'keit kein Korb überläuft ?

Sei $Y_i :=$ Anzahl der Bälle im i-ten Korb für $i = 1, \ldots, n$ und sei

$$Y := \max_{1 \leq i \leq n} Y_i$$

Wir suchen nach einer möglichst langsam wachsenden Funktion $f(n)$ mit

$$\lim_{n \to \infty} \Pr[Y \geq f(n)] = 0$$

Chernoff-Schranken XI

Wir verwenden Aussage 5 von Korollar 90.

Es gilt $p_1 = \ldots = p_{cn} = \dfrac{1}{n}$ und $\mu = \sum_{i=1}^{cn} p_i = c$.

Mit $t = 2\log n$ folgt

$$\Pr[Y_i \geq 2\ln n] \leq 1/n^2 \qquad \text{für } \ln n \geq ec.$$

Daraus ergibt sich

$$\begin{aligned}
\Pr[Y \geq 2\ln n] &= \Pr[\, Y_1 \geq 2\ln n \;\cup \ldots \cup\; Y_{cn} \geq 2\ln n \,] \\
&\leq \sum_{i=1}^{cn} \Pr[\, Y_1 \geq 2\ln n \,] \\
&\leq cn \cdot \frac{1}{n^2} = \frac{c}{n}
\end{aligned}$$

Chernoff-Schranken XII

Literatur:

Torben Hagerup, Christine Rüb:
A guided tour of Chernoff bounds
Inf. Process. Lett. **33**, pp. 305–308 (1990)

13. Erzeugende Funktionen

Erzeugende Funktionen I

Definition 92

Für eine Zufallsvariable X mit $W_X \subseteq \mathbb{N}_0$ ist die (wahrscheinlichkeits-)erzeugende Funktion definiert durch

$$G_X(s) := \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k = \mathbb{E}[s^X] \,.$$

Bei w'keitserzeugenden Funktionen haben wir kein Problem mit der Konvergenz, da für $|s| < 1$ gilt

$$\begin{aligned} |G_X(s)| &= \left| \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k \right| \\ &\leq \sum_{k=0}^{\infty} \Pr[X = k] \cdot |s^k| \leq \sum_{k=0}^{\infty} \Pr[X = k] = 1 \,. \end{aligned}$$

Erzeugende Funktionen II

Gegeben $G_X(s)$ können wir die W'keiten $\Pr[X = i]$ "ablesen".

Mit

$$G'_X(s) = \sum_{k=1}^{\infty} k \cdot \Pr[X = k] \cdot s^{k-1}$$

gilt

$$\Pr[X = 1] = G'_X(0)$$

Analog erhalten wir

$$G_X^{(i)}(0) = \Pr[X = i] \cdot i!$$

also

$$\Pr[X = i] = \frac{G_X^{(i)}(0)}{i!}$$

Erzeugende Funktionen III: Verteilungen

Bernoulli-Verteilung Sei X eine Bernoulli-verteilte Zufallsvariable mit $\Pr[X = 0] = 1 - p$ und $\Pr[X = 1] = p$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = (1-p) \cdot s^0 + p \cdot s^1 = 1 - p + ps\,.$$

Gleichverteilung auf $\{0, \ldots, n\}$

Sei X auf $\{0, \ldots, n\}$ gleichverteilt, d.h. für $0 \leq k \leq n$ ist $\Pr[X = k] = 1/(n+1)$. Dann gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \frac{1}{n+1} \cdot s^k = \frac{s^{n+1} - 1}{(n+1)(s-1)}\,.$$

Erzeugende Funktionen IV: Verteilungen

Binomialverteilung Für $X \sim \mathrm{Bin}(n,p)$ gilt nach der binomischen Formel

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot s^k = (1-p+ps)^n \,.$$

Geometrische Verteilung

Sei X eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p. Dann gilt

$$\begin{aligned} G_X(s) = \mathbb{E}[s^X] &= \sum_{k=1}^{\infty} p(1-p)^{k-1} \cdot s^k \\ &= ps \cdot \sum_{k=1}^{\infty} ((1-p)s)^{k-1} = \frac{ps}{1-(1-p)s} \,. \end{aligned}$$

Erzeugende Funktionen V: Verteilungen

Poisson-Verteilung Für $X \sim \mathrm{Po}(\lambda)$ gilt

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot s^k = e^{-\lambda+\lambda s} = e^{\lambda(s-1)}.$$

Sei X binomialverteilt mit $X \sim \mathrm{Bin}(n, \lambda/n)$. Für $n \to \infty$ folgt

$$G_X(s) = \left(1 - \frac{\lambda}{n} + \frac{\lambda s}{n}\right)^n = \left(1 + \frac{\lambda(s-1)}{n}\right)^n \to e^{\lambda(s-1)}.$$

Man kann beweisen, dass aus der Konvergenz der erzeugenden Funktion die Konvergenz der Verteilung folgt.

Erzeugende Funktionen VI: Erwartungswert und Varianz

Gegeben $G_X(s)$ können wir auch $\mathbb{E}[X]$, $\mathrm{Var}[X]$, und andere Momente "ablesen".

Da

$$G_X(s) := \sum_{k=0}^{\infty} \Pr[X = k] \cdot s^k = \mathbb{E}[s^X]\,,$$

gilt

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k \cdot \Pr[X = k] = G'_X(1)\,.$$

Erzeugende Funktionen VII: Erwartungswert und Varianz

Beispiel 93

Sei X binomialverteilt mit $X \sim \mathrm{Bin}(n,p)$, also

$$G_X(s) = (1-p+ps)^n\,.$$

Dann gilt

$$G'_X(s) = n \cdot (1-p+ps)^{n-1} \cdot p$$

und somit

$$\mathbb{E}[X] = G'_X(1) = np\,.$$

Erzeugende Funktionen VIII: Erwartungswert und Varianz

Ebenso ergibt sich

$$\mathbb{E}[X(X-1)\dots(X-k+1)] = G_X^{(k)}(1)\,,$$

und damit

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\ &= G''_X(1) + G'_X(1) - (G'_X(1))^2\,.\end{aligned}$$

Erzeugende Funktionen IX: Summe

Mit dem folgenden Satz gewinnen wir die erzeugenden Funktionen für Summen von unabhängigen Variablen.

Satz 94

Seien $X_1, \ldots, X_n$ *unabhängige Zufallsvariablem mit* $W_{X_1}, \ldots, W_{X_n} \subseteq \mathbb{N}$ *und sei* $Z := X_1 + \ldots + X_n$. *Es gilt*

$$G_Z(s) = \prod_{i=1}^{n} G_{X_i}(s)$$

Beweis: Nur für $n = 2$

Erzeugende Funktionen IX: Summe

$$
\begin{aligned}
G_Z(s) &= \sum_{k=0}^{\infty} \Pr[Z = k] \cdot s^k \\
&= \sum_{k=0}^{\infty} \left(\sum_{\{(i,j) \in \mathbb{N}^2 \mid i+j=k\}} \Pr[X_1 = i, X_2 = j] \right) \cdot s^k \\
&= \sum_{k=0}^{\infty} \left(\sum_{\{(i,j) \in \mathbb{N}^2 \mid i+j=k\}} \Pr[X_1 = i] \cdot \Pr[X_2 = j] \right) \cdot s^{i+j} \\
&= \sum_{i,j=0}^{\infty} \Pr[X_1 = i] \cdot \Pr[X_2 = j] \cdot s^{i+j} \\
&= \left(\sum_{i=0}^{\infty} \Pr[X_1 = i] \cdot s^i \right) \cdot \left(\sum_{j=0}^{\infty} \Pr[X_2 = j] \cdot s^j \right) \\
&= G_{X_1}(s) \cdot G_{X_2}(s)
\end{aligned}
$$