

Zusammenfassung MISTO

(Angewandte Statistik)

Deskriptive Statistik

o Stamm-Blatt-Diagramme

Urliste: 41, 43, 52, 57, 58, 60, 79, 80

Diagramm: Bei zweistelligen Blättern

4	13	77	→	0 77	→	0 8
5	278	104	→	1 04	→	1 0
6	0	132	→	1 32	→	1 3
7	9	227	→	2 27	→	2 3
8	0					

Schritt 1: Teile den Datenbereich in Intervalle gleicher Breite $da, 5$ oder 1 mal einer Potenz von 10 ein. Trage die erste(n) Ziffer(n) der Werte in jeweiligen Intervall links von einer senkrechten Linie der Grösse nach geordnet ein. Dies ergibt den Stamm.

Schritt 2: Runde die beobachteten Werte auf die Stelle, die nach den Ziffern des Stamms kommt. Die resultierenden Ziffern ergeben die Blätter. Diese werden zeilenweise und der Grösse nach geordnet rechts vom Stamm eingetragen.

Histogramme

o Häufigkeitstabelle

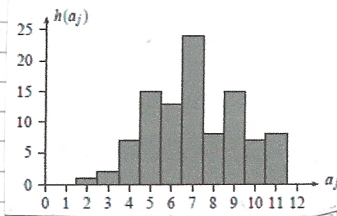
- Die absolute Häufigkeit $h(a_j)$ der Ausprägung a_j ist die Anz. der $x_i \in X$ mit $x_i = a_j$
- Die relative Häufigkeit $f(a_j)$ der Ausprägung a_j ist $f(a_j) = \frac{h(a_j)}{n}$

Bspl: Histogramm anhand einer Häufigkeitstabelle

j	1	2	3	4	5	6	7	8	9	10	11
a_j	2	3	4	5	6	7	8	9	10	11	12
$h(a_j)$	1	2	7	15	13	24	8	15	7	8	0
$f(a_j)$	0.01	0.02	0.07	0.15	0.13	0.24	0.08	0.15	0.07	0.08	0.0

Abbildung 2.3: Stabdiagramm für relative Häufigkeiten $f(x)$ und Verteilungsfunktion $H(x)$ Die Verteilungsfunktion $H(x)$ ist eine monoton steigende Treppenfunktion für die gilt:

$$H(x) = \sum_{u \leq x} f(u), \quad 0 \leq H(x) \leq 1, \quad \sum_{a \leq x \leq b} f(x) = H(b) - H(a).$$



- Ist Ausprägung konkret $\rightarrow$ Bestimmung der absoluten bzw. relativen Häufigkeit kein Problem.
- Ist Auspr. stetig (z.B. Körperlänge) $\rightarrow$ Daten in Klassen einteilen

Anz. Klassen: $\sqrt{n}$ n = Umfang d. StichprobeEmpirische Verteilungsfunktion $0 \leq H(x) \leq 1$

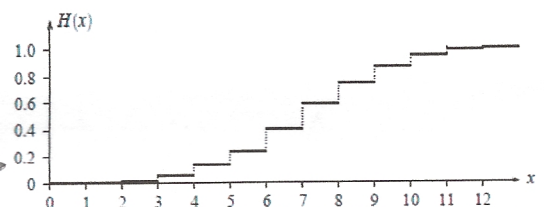
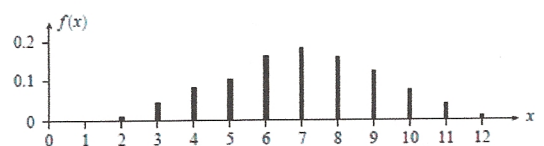
- Die Summenhäufigkeitsfkt $H(x)$ erhält man folgendermassen:
Für die Ausprägung x bestimmt man die Anzahl von Beobachtungswerten in der Stichprobe, die kleiner oder gleich x sind und bildet dann die relative Häufigkeit.

Bspl: Sei X eine geordnete Stichprobe vom Umfang n .

$$\rightarrow \text{Verteil.fkt: } H(x) = \frac{1}{n} \sum_{u \leq x} h(u) = \sum_{u \leq x} f(u)$$

Zwei Würfel werden 1000 mal geworfen. Das Merkmal X ist die Summe der Augenzahlen, dessen Ausprägungen in der folgenden Häufigkeitstabelle vorliegen.

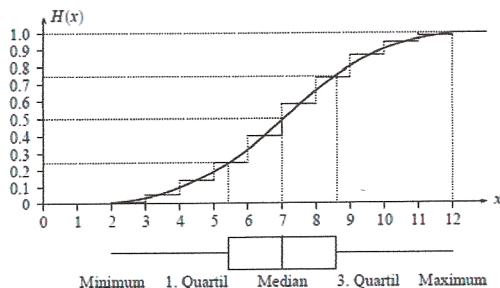
Ausprägung x	Häufigkeit $h(x)$	relative Häufigkeit $f(x)$	absolute Summenhäufigkeit $G(x)$	Verteilungsfunktion $H(x)$
2	12	0.012	12	0.012
3	46	0.046	58	0.058
4	83	0.083	141	0.141
5	103	0.103	244	0.244
6	160	0.160	404	0.404
7	180	0.180	584	0.584
8	159	0.159	743	0.743
9	125	0.125	868	0.868
10	77	0.077	945	0.945
11	43	0.043	988	0.988
12	12	0.012	1000	1.000

In der Abbildung 2.3 ist die relative Häufigkeit $f(x)$ als Stabdiagramm dargestellt. Die Verteilungsfunktion $H(x)$ ist eine Treppenfunktion.

Quantile

- Teilen Grundgesamtheit in 4 gleichgr. Teile
- Erlauben feinere Unterteilung

Boxplot



Für die folgende Interpretation betrachten wir die geglättete Kurve. Das 1. Quartil bestimmt sich aus $a = 0.25$, d.h. $H(x) = 0.25$, somit liest man aus der Graphik ungefähr $x = 5.4$ heraus und für $a = 0.5$ ergibt sich $x = 7$, ferner für $a = 0.75$ $x = 8.6$. Somit

$$x_{0.25} = 5.4, \quad x_{0.5} = 7, \quad x_{0.75} = 8.6.$$

Der Wert $x_{0.25} = 5.4$ bedeutet, dass alle Werte in der Stichprobe mit $x \leq 5.4$ 25% aller Stichprobenwerte ausmacht. Analog für die anderen Werte. Die Interpretation des Boxplots ist:

- Die Hälfte aller Werte liegen innerhalb der Box
- 50% aller Werte sind kleiner als der Median und 50% aller Werte sind grösser als der Median.
- Das 1. Quartil ist der Wert, für den gilt, $\frac{1}{4}$ aller Stichprobenwerte sind kleiner als dieser Wert. Das 3. Quartil ist der Wert, für den gilt, $\frac{3}{4}$ aller Stichprobenwerte sind kleiner als dieser Wert.

Statistische Masszahlen

o Lagemasse

□ Arithm. Mittel

Aufsummieren aller beobachteten Werte und diese Summe durch Anz. Beob. teilen

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Satz: Für Häufigkeitsdaten mit Ausprägungen $a_1, \dots, a_k$ und relativen Häufigkeiten $f(a_1), \dots, f(a_k)$ gilt:

$$\bar{x} = \sum_{i=1}^k a_i \cdot f(a_i)$$

□ Modalwert (Modus)

Gibt an, welche Ausprägung am häufigsten vorkommt.

Der Modalwert x_{mod} bei diskreten Merkmalen: Ausprägung mit grösster Häufigkeit

Der Modalwert x_{mod} bei stetigen Merkmalen: Maximum der Parabel, welche im Histogramm durch die Punkte über den Klassenmitten der am stärksten besetzten Klasse und der beiden Nachbar Klassen geht.

□ Median

Wird so in die Datenmitte platziert, dass 50% unterhalb und 50% oberhalb sind

Für ungerades n ist der Median $\tilde{x}$ die mittlere Beobachtung der geordneten Urliste und für gerades n das arithmetische Mittel der beiden in der Mitte liegenden Beob. das heisst:

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & \text{für } n \text{ ungerade} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{für } n \text{ gerade} \end{cases}$$

□ Quartile

- Indexwerte der Quartile

$$Q_1: \left[\frac{1}{2} \left(1 + \frac{n+1}{2} \right) \right] = \left[\frac{n+3}{4} \right]$$

$$Q_2: \text{Median}$$

$$Q_3: \left[\frac{1}{2} \left(\frac{n+1}{2} + n \right) \right] = \left[\frac{3n+1}{4} \right]$$

- Quartilswerte für gemittelten Indexwert:

→ linear interpolieren

$$Q_1 = (x_{n_1} - x_{n_2}) \cdot \frac{n_1+3}{4} + x_{n_2} \cdot n_2 - x_{n_1} \cdot n_1$$

$$Q_3 = (x_{n_2} - x_{n_1}) \cdot \frac{3n_1+1}{4} + x_{n_1} \cdot n_2 - x_{n_2} \cdot n_1$$

Bsp1: Indexwerte: $Q_1 = 2,25$ $Q_2 = 4,75$

Quartilswerte:

$$Q_1: n_1 = 2 \quad n_2 = 3$$

$$\rightarrow (4-3) \cdot \frac{2+3}{4} + 3 \cdot 3 - 4 \cdot 2 = 3,25$$

$$Q_3: n_1 = 4 \quad n_2 = 5$$

$$\rightarrow (6-5) \cdot \frac{13}{4} + 5 \cdot 5 - 6 \cdot 4 = 5,75$$

Zusammenfassung MSTO

o Streuungsmasse

□ Standardabweichung, Varianz

Die empirische Varianz der Werte $x_1, \dots, x_n$

$$\text{ist } s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Die emp. Standardabw. s ist die Wurzel aus der Varianz $s = \sqrt{s^2}$

$$(\bar{x} = \text{Mittelwert } x_1, x_2, \dots, x_n)$$

Satz: Für die Stichprobe $x_1, \dots, x_n$ mit Ausprägungen $a_1, \dots, a_n$ gilt

$$s^2 = \sum_{i=1}^n (a_i - \bar{x})^2 f(a_i)$$

Satz: Für die emp. Varianz gilt

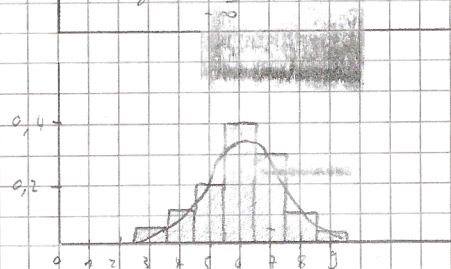
$$s^2 = \frac{1}{n} \left(\sum_{i=1}^n x_i^2 - n \bar{x}^2 \right)$$

Alle Messwerte im Quadrat summieren

Dichtekurven & Normalverteilung

Eine Fkt $f(x)$ ist eine Dichtekurve, wenn $f(x) \geq 0$ ist und die von $f(x)$ überdeckte Gesamtfläche = 1 ist.

$$\text{Also gilt: } \int_{-\infty}^{\infty} f(x) dx = 1$$



o Normalverteilung

- Sind symmetrisch, glockenförmig und haben ein Maximum

Die Dichtekurven einer Normalverteilung

ist für jedes $x \in \mathbb{R}$ durch

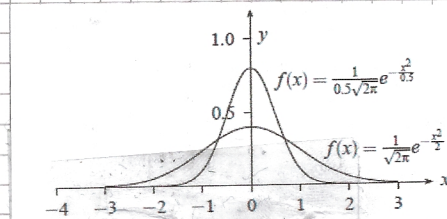
$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ definiert.}$$

Für gegebene Werte der Parameter $\mu \in \mathbb{R}$ und $\sigma > 0$ ist $f(x)$ eindeutig

$$\sigma = s \quad \mu = \bar{x}$$

Spezifiziert Daten heisst Mittelwert und Standardabweichung von $f(x)$

- Maximum an der Stelle $x = \mu$
- Je kleiner σ desto steiler

Bsp: $\mu = 0, \sigma = 1$ und $\sigma = 0,5$ 

Satz:

68.3% aller Beobachtungen liegen im Intervall $[\mu - \sigma, \mu + \sigma]$,95.5% aller Beobachtungen liegen im Intervall $[\mu - 2\sigma, \mu + 2\sigma]$,99.7% aller Beobachtungen liegen im Intervall $[\mu - 3\sigma, \mu + 3\sigma]$.Mehrdimensionale Verteilungen $y = mx + c \rightarrow$ Ausgleichsgerade Punktwolke

$$m = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}$$

$$c = \bar{y} - m \bar{x}$$

Die empirische Kovarianz der Punktwolke $\{(x_i, y_i) | i=1, \dots, n\}$ ist definiert als $\text{cov}(\{(x_i, y_i)\}) =$

$$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Das Vorzeichen von $\text{cov}(x, y)$ sagt aus, ob die Punktwolke ansteigt oder fällt

$$= \text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$$

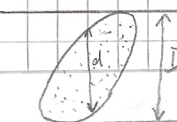
Der empirische Korrelationskoeff.

$r(x, y)$ der Punktwolke $\{(x_i, y_i) | i=1, \dots, n\}$ ist definiert als:

$$r(x, y) = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})}{s(x)} \cdot \frac{(y_i - \bar{y})}{s(y)}$$

(s(x) & s(y) = numerische Standardabw.)

$$\text{Ellipsenregel: } r^2(x, y) \approx 1 - \left(\frac{\sigma}{s}\right)^2$$



Ist $r(x, y) = 0$, so heissen x und y unkorreliert

Wahrscheinlichkeitsrechnung

Ergebnisraum (ER)

Der Ergebnisraum Ω ist die Menge aller Ergebnisse ω eines Zufallversuchs (ZV)

Ereignis

Sei Ω d. ER eines ZV. Ein Ereignis A ist eine Teilmenge von Ω . Das Eig. A tritt ein, falls das Ergebnis ω eines ZV Element von A ist: $\omega \in A \subseteq \Omega$

Das Eig. A tritt nicht ein, falls das Ergebnis ω eines ZV nicht Element ω ist: $\omega \notin A \subseteq \Omega$. Die elem. Teilmengen $\{\omega\} \subseteq \Omega$ werden als Elementarereignisse bezeichnet.

Verknüpfung v. Ereignissen

- Kommutativgesetz: $A \cap B = B \cap A$, $A \cup B = B \cup A$
 Assoziativgesetz: $(A \cap B) \cap C = A \cap (B \cap C)$, $(A \cup B) \cup C = A \cup (B \cup C)$
 Distributivgesetz: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
 De Morgansche Regel: $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$
 Aus $A \subset B$ folgt: $\overline{B} \subset \overline{A}$
 Differenzmenge: $A \setminus B = A \cap \overline{B}$

Definition 1.3

- Das unmögliche Ereignis wird durch die leere Menge $\emptyset$ repräsentiert. Es ist das Ereignis, das nie eintritt.
- Das sichere Ereignis wird durch die Menge Ω , d.h. den ganzen Ergebnisraum repräsentiert. Es ist das Ereignis das immer eintritt.
- Das Ereignis A und B wird durch $A \cap B$ repräsentiert. Es tritt genau dann ein, wenn A eintritt und B eintritt.
- Das Ereignis A oder B wird durch $A \cup B$ repräsentiert. Es tritt genau dann ein, wenn A eintritt oder B eintritt oder wenn beide eintreten.
- Das komplementäre Ereignis $\overline{A}$ von A wird durch $\Omega \setminus A = \{\omega \in \Omega | \omega \notin A\}$ repräsentiert. Es tritt genau dann ein, wenn A nicht eintritt.

Klassische Wahrscheinlichkeit

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anz. Elemente von } A}{\text{Anz. Elemente von } \Omega}$$

Geometrische Wahrscheinlichkeit

$$P(A) = \frac{\text{Länge von } A}{\text{Länge von } \Omega} \quad \text{bei } n=1$$

$$P(A) = \frac{\text{Flächeninhalt v. } A}{\text{Flächeninhalt v. } \Omega} \quad \text{bei } n=2$$

$$P(A) = \frac{\text{Volumen v. } A}{\text{Volumen v. } \Omega} \quad \text{bei } n=3$$

$$A \subseteq \mathbb{R}^n$$

Statistische Wahrscheinlichkeit

→ Relative Häufigkeit

"Ist nach n Versuchswiederholungen n_A -mal das Ereignis A eingetreten, so ist

$$h_n(A) = \frac{n_A}{n} \quad \text{die relative Häufigkeit v. } A.$$

Die statistische Wahrsch. von A ist dann ein Grenzwert $\lim_{n \rightarrow \infty} h_n(A)$

Axiomatische Wahrscheinlichkeit

Definition 3.1

Sei Ω ein Ergebnisraum. Eine Ereignisalgebra $\mathcal{A}$ ist eine Teilmenge der Potenzmenge $\mathcal{P}(\Omega)$ mit den Eigenschaften:

- $\Omega \in \mathcal{A}$.
- $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$.
- $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$ und $A \cup B \in \mathcal{A}$.

Ein Wahrscheinlichkeitsmass ist eine Abbildung $P: \mathcal{A} \rightarrow [0, 1]$ mit den Bedingungen:

- $P(\emptyset) = 0$ und $P(\Omega) = 1$.
- $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$.



Das Tripel $(\Omega, \mathcal{A}, P)$ bezeichnet man als Wahrscheinlichkeitsraum.

Bspl: $\Omega = \{1, 2, 3\}$

$\mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$
 ↓
 Potenzmenge = 2^n Teilmengen (2^n) $n = \text{Anz. Elem.}$

1) $\mathcal{A}_1 = \{\{1, 2, 3\}, \emptyset\} \rightarrow$ kleinste Ereignisalgebra

2) $\mathcal{A}_2 = \{\{1, 2, 3\}, \emptyset, \{2\}, \{1, 3\}\}$

3) $\mathcal{A}_3 = \{\{1, 2, 3\}, \emptyset, \{1, 3\}, \{2, 3\}, \{3\}, \{1, 2\}, \{1\}, \{2\}\}$
 (komplement v. 2)

Folgerung 3.1

Sei Ω ein Ergebnisraum und $\mathcal{A}$ eine Ereignisalgebra, dann gilt:

- $\Omega \in \mathcal{A} \Rightarrow \emptyset \in \mathcal{A}$.
- Für $A, \overline{A} \in \mathcal{A}$ gilt $P(A) + P(\overline{A}) = 1$.
- Für beliebige Ereignisse $A, B \in \mathcal{A}$ gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- $P(A \setminus B) = P(A) - P(A \cap B)$.



Zusammenfassung MSTODiskreter endlicher Stichprobenraum

Satz: Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum mit dem endlichen, diskreten Ergebnisraum $\Omega = \{\omega_1, \dots, \omega_n\}$.

Seien $0 \leq P(\{\omega_k\}) \leq 1$ die Wahrscheinlichkeiten der Elementarereignisse $\{\omega_k\} \in \Omega, k=1, \dots, n$.

Sei $A \in \mathcal{A}$ ein Ereignis. Die Wahrsch. $P(A)$ lässt sich berechnen durch:

$$P(A) = \sum_{\omega_k \in A} P(\{\omega_k\})$$

$\Rightarrow$ Die Summe aller Wahrscheinlichkeiten der Elementarereignisse aus Ω ist Eins:

$$\sum_{\omega_k \in \Omega} P(\{\omega_k\}) = 1$$

Max und Moritz schiessen je auf eine Flasche und ermitteln so ihre Trefferwahrscheinlichkeiten. Die Wahrscheinlichkeit, dass Max trifft, ist $\frac{1}{3}$. Die Wahrscheinlichkeit, dass Moritz trifft, ist $\frac{1}{3}$. Jeder hat zwei Steine und sie schiessen abwechselungsweise auf die Flasche um sie zu treffen. Max beginnt. Wie gross ist die Wahrscheinlichkeit, dass Max die Flasche mit seinen zwei Steinen trifft?

Lösung mit Hilfe des Baumes in Abbildung 3.7. Man nimmt alle Pfade, wo Max trifft.

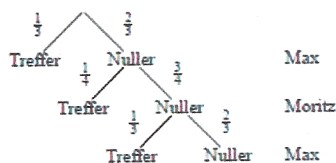


Abbildung 3.7: Baum für zweimaliges Schiessen

Somit ist die Wahrscheinlichkeit, dass Max trifft:

$$P(\text{Max trifft}) = \frac{1}{3} + \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{2}$$

Max und Moritz schiessen wieder abwechselungsweise auf eine Flasche. Max beginnt. Wie oft muss Max schiessen, damit er die Flasche mit einer Wahrscheinlichkeit von mindestens 0.66 trifft?

Lösung wieder mit Hilfe des Baumes in Abbildung 3.7, wobei er fortgesetzt werden muss. Die Wahrscheinlichkeiten sind:

$$\text{mit dem 1. Schuss von Max: } P(\text{Max trifft}) = \frac{1}{3}$$

$$\text{mit dem 2. Schuss von Max: } P(\text{Max trifft}) = \frac{1}{3} + \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{2}$$

$$\text{mit dem 3. Schuss von Max: } P(\text{Max trifft}) = \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + \left(\frac{1}{2}\right)^2 \cdot \frac{1}{3} = \frac{1}{3} \left[1 + \frac{1}{2} + \frac{1}{4}\right] = \frac{7}{12}$$

$$\text{mit dem } n. \text{ Schuss von Max: } P(\text{Max trifft}) = \frac{1}{3} \left[1 + \frac{1}{2} + \dots + \left(\frac{1}{2}\right)^{n-1}\right] = \frac{1}{3} \cdot \frac{2^n - 1}{2^{n-1}}$$

Somit

$$\begin{aligned} \frac{1}{3} \cdot \frac{2^n - 1}{2^{n-1}} &\geq 0.66 \\ 2^n - 1 &\geq 1.98 \cdot 2^{n-1} \\ 2^n - 1 &\geq 0.99 \cdot 2^n \\ 0.01 \cdot 2^n &\geq 1 \\ 2^n &\geq 100 \\ n &\geq 7. \end{aligned}$$

Unendlicher diskreter Ergebnisraum

Beispiel: Sei X die Anz. Würfe bis zum Erscheinen von Kopf.

Ereignis A : X ist eine ungerade Zahl

Ereignis B : X ist eine gerade Zahl

Ereignis C : $X > 5$

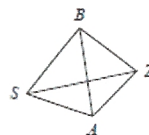
$$\begin{aligned} \rightarrow P(A) &= P(\{1\}) + P(\{3\}) + P(\{5\}) + \dots \\ &= \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \dots = \frac{\frac{1}{2}}{1 - \frac{1}{4}} = \frac{2}{3} \end{aligned}$$

$$\begin{aligned} P(B) &= P(\{2\}) + P(\{4\}) + P(\{6\}) + \dots \\ &= \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \dots = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3} \end{aligned}$$

$$P(C) = 1 - P(\bar{C}) = 1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{8} - \frac{1}{16} - \frac{1}{32} = \frac{1}{32}$$

Beispiel 3.14

Ein Käfer läuft auf einem regulären Tetraeder aus Draht. Er startet in S und braucht um eine Kante zu durchlaufen 1 Minute. An jeder Ecke wählt er zufällig eine der Kanten. Wie gross ist die Wahrscheinlichkeit, dass er den Punkt Z zum ersten Mal nach genau n Minuten erreicht?



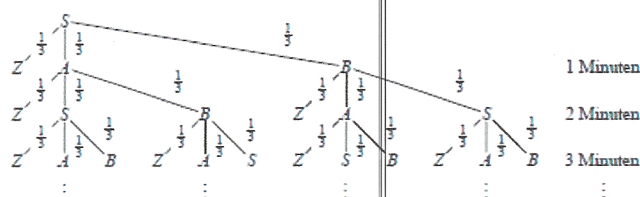
Bezeichnen wir z.B. den Weg von S nach Z über A und B mit (S, A, B, Z) , dann ist der Ergebnisraum unendlich und diskret:

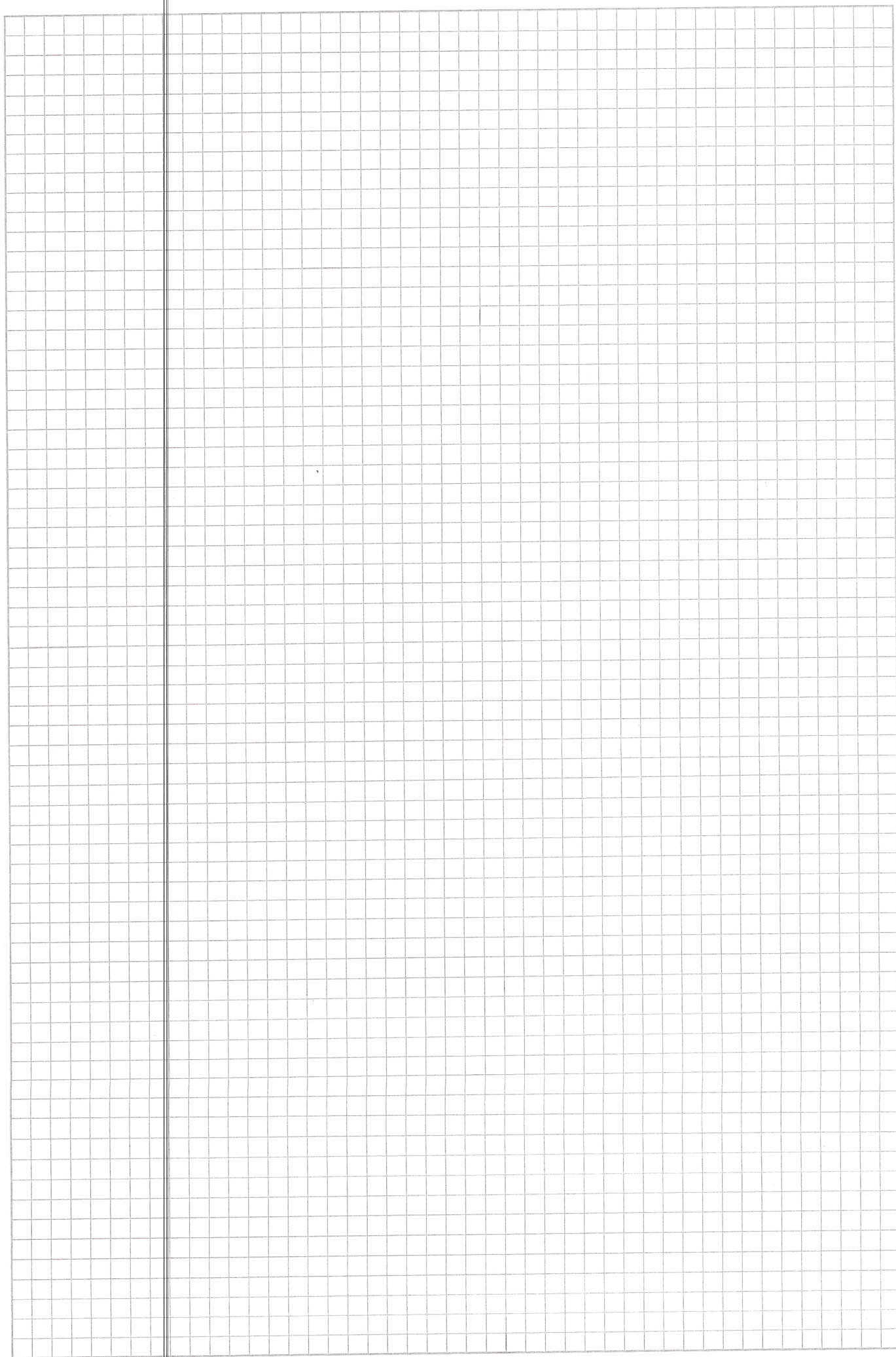
$$\Omega = \{(S, Z), (S, A, Z), (S, B, Z), (S, A, B, Z), (S, A, S, Z), (S, B, A, Z), (S, B, S, Z), (S, A, B, S, Z), \dots\}$$

Die Wahrscheinlichkeiten der einzelnen Pfade erhält man durch einen Baum, der in Abbildung 3.9 skizziert ist.

Also sind die Wahrscheinlichkeiten, dass er das Ziel Z erreicht:

$$\begin{aligned} \text{in 1 Minute: } & \frac{1}{3} \\ \text{in 2 Minuten: } & 2 \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{2}{9} \\ \text{in 3 Minuten: } & 4 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{4}{27} \\ \text{in } n \text{ Minuten: } & \frac{2^{n-1}}{3^n} = \frac{1}{3} \left(\frac{2}{3}\right)^{n-1} \end{aligned}$$





MSO Zusammenfassung 2

Bedingte Wahrscheinlichkeit

Definition: Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ Ereignisse eines Zufallsversuchs mit $P(B) > 0$. Die bedingte Wahrsch. $P(A|B)$ von A unter B ist

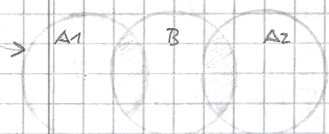
$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

Analog für

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) > 0$$

Folgerung: Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und seien $A_1, A_2, B \in \mathcal{A}$ Ereignisse in Ω . Es gilt

$$P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B), \text{ falls } A_1 \cap A_2 = \emptyset, \text{ und } P(B) > 0$$



Produktsatz: Seien $A, B \in \mathcal{A}$ und $P(B) > 0$, dann gilt:

$$P(A \cap B) = P(A|B) \cdot P(B) \text{ oder } P(A \cap B) = P(B|A) \cdot P(A)$$

Unabhängigkeit von zwei Ereignissen

Definition: Seien A und B beliebige Ereignisse eines Zufallsversuchs. Das Ereignis A heißt stochastisch unabhängig von B, wenn das Eintreffen von B die Wahrsch. für A nicht beeinflusst.

$$P(A|B) = P(A)$$

Folgerung: Seien A und B beliebige unabhängige Ereignisse eines Zufallsversuchs. Dann gilt:

$$P(A \cap B) = P(A) \cdot P(B)$$

Die Wahrsch. für das Ereignis

A und B ist das Produkt der Wahrsch. von A, resp. B

Totale Wahrscheinlichkeit

Satz: Sei $B \in \mathcal{A}$ ein Ereignis, und die Ereignisse $A_1, \dots, A_n \in \mathcal{A}$ seien eine Partition von Ω , d.h.

$$\bigcup_{k=1}^n A_k = \Omega \text{ und } A_i \cap A_j = \emptyset \text{ für } i \neq j,$$

$$i, j = 1, \dots, n$$

Dann gilt:

$$P(B) = \sum_{k=1}^n P(B|A_k) \cdot P(A_k)$$

Satz von Bayes

Satz: Seien die Ereignisse $A_1, \dots, A_n$ eine Partition vom Ergebnisraum Ω , und sei $B \in \mathcal{A}$ ein Ereignis mit $P(B) > 0$.

Dann gilt:

$$P(A_k|B) = \frac{P(B|A_k) \cdot P(A_k)}{\sum_{i=1}^n P(B|A_i) \cdot P(A_i)}, \quad k=1, \dots, n$$

Diskrete Zufallsvariablen

Definition: Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Eine Zufallsvariable $X: \Omega \rightarrow \mathbb{R}, x = X(\omega)$ heißt diskret, wenn sie endlich viele oder abzählbar unendlich viele Werte $x_1, x_2, \dots, x_k, \dots$ annehmen kann, d.h. wenn die Menge $\Omega' = \{\omega = X(\omega) | \omega \in \Omega\} = \{x_1, x_2, \dots, x_k, \dots\}$ endlich viele oder abzählbar unendlich viele Elemente hat. Die Wahrsch.-Fkt $g: \Omega' = X(\Omega) \subseteq \mathbb{R} \rightarrow [0, 1]$ der diskret. Zufallsvariablen X ist definiert durch

$$g(x_k) = P(X^{-1}(x_k))$$

Satz: Sei $(\Omega, \mathcal{A}, P)$ ein Wahrsch.-raum, X eine Zufallsvariable und q eine Wahrsch.fkt. Sei $A, B \subseteq \Omega'$, dann gilt $q(A) = \sum_{x_i \in A} q(x_i)$

Damit wird ein Wahrsch.mass die finiert, d.h.

$$1) q(\emptyset) = 0, q(\Omega') = 1$$

$$2) A \cap B = \emptyset \Rightarrow q(A \cup B) = q(A) + q(B)$$

Verteilung einer diskreten Zufallsvar

Definition: Das Wahrsch.mass

$$P_x: \mathcal{A} \rightarrow [0, 1], P_x(A) = \sum_{x_i \in A} q(x_i)$$

heißt Verteilung der Zufallsvariablen X .

Die Abbildung

$$F_x: \mathbb{R} \rightarrow \mathbb{R}, F_x(x) = P(X \leq x) = \sum_{x_i \leq x} q(x_i)$$

Erwartungswert

Definition: Der Erwartungswert $E(X)$ der Zufallsvariablen X ist:

$$E(X) = \sum_{x_k \in \Omega'} x_k \cdot q(x_k) \text{ oder}$$

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\omega)$$

• Rechenregeln für den Erwartungswert

Satz: Seien X und Y Zufallsvariablen und $\lambda, \mu \in \mathbb{R}$, dann gilt

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$$

Der Indikator eines Ereignisses

Geg: $(\Omega, \mathcal{A}, P)$ und Ereignis $A \subseteq \Omega$

Indikator von A : $I_A(\omega) = \begin{cases} 1, & \text{falls } \omega \in A \\ 0, & \text{falls } \omega \in \bar{A} \end{cases}$

Satz: Sei I_A der Indikator eines Ereignisses $A \subseteq \Omega$, dann ist der Erwartungswert: $E(I_A) = P(A)$

Eigenschaften:

$$1) I_A^2 = I_A$$

$$2) I_{\bar{A}} = 1 - I_A$$

$$3) I_{A_1 \cap A_2 \cap \dots \cap A_n} = I_{A_1} \cdot I_{A_2} \cdot \dots \cdot I_{A_n}$$

$$4) I_{A_1 \cup A_2 \cup \dots \cup A_n} = 1 - (1 - I_{A_1})(1 - I_{A_2}) \cdot \dots \cdot (1 - I_{A_n})$$

Varianz und Streuung

Definition: Die Varianz der Zufallsvar. X ist definiert durch

$$V(X) = E((X - \mu)^2) = \sum_{k=1}^n (x_k - \mu)^2 q(x_k),$$

mit $\mu = E(X)$.

Die Streuung/Standardabweichung ist

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

Satz: Für die Varianz gilt:

$$V(X) = E(X^2) - (E(X))^2,$$

resp. für die Streuung gilt:

$$\sigma = \sqrt{E(X^2) - (E(X))^2}$$

• Rechenregeln für die Varianz

Satz: Sei X eine Zufallsvariable und $\lambda, \mu \in \mathbb{R}$, dann gilt

$$V(\lambda X + \mu) = \lambda^2 V(X)$$

MS10 Zusammenfassung 2

Spezielle diskrete Verteilungen

◦ Binomialverteilung

Definition: Eine Zufallsvariable X heisst binomial verteilt mit den Parametern n und p , wenn sie die Wahrschkft

$$q: \Omega' \rightarrow [0,1], q(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

besitzt. Dabei ist n die Anzahl Wdh. eines Bernoulli-Versuchs mit den Ergebnissen 0 und 1. Dabei ist p die Wahrsch. des Elementarereignisses 1

Satz: Sei X eine binomial verteilte Zufallsvariable. Der Erwartungswert und die Streuung sind dann:

$$E(X) = \sum_{x \in \Omega'} q(x) \cdot x = n \cdot p$$

$$\sigma = \sqrt{n \cdot p \cdot (1-p)}$$

◦ Geometrische Verteilung

Anz. Bernoulli-Versuche bis zum ersten Erfolg

Definition: Eine Zufallsvar. X heisst geometrisch verteilt, wenn sie die Wahrschkft

$$q: \Omega' = \mathbb{N} \rightarrow [0,1], q(x) = (1-p)^{x-1} p$$

besitzt. Dabei ist X die Anz. Wdh. eines Bernoulli-Versuchs bis der Treffer zum ersten Mal auftritt und p ist die Wahrsch. des Treffers

Satz: Sei X eine geometrisch verteilte Zufallsvariable. Der Erwartungswert und die Streuung sind:

$$E(X) = \frac{1}{p}, \quad \sigma = \sqrt{\frac{1-p}{p^2}}$$

◦ Hypergeometrische Verteilung

Satz: Eine Zufallsvariable X heisst hypergeometrisch verteilt mit den Parametern n, M, N , wenn sie die Wahrschkft

$$q(x) = \begin{cases} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}, & \text{für } x \in \{0, 1, 2, \dots, n\} \\ 0, & \text{sonst} \end{cases}$$

besitzt. Erwartungswert und Streuung sind

$$E(X) = n \frac{M}{N}, \quad \sigma = \sqrt{n \frac{M}{N} \left(1 - \frac{M}{N}\right) \left(1 - \frac{n-1}{N-1}\right)}$$

◦ Poisson Verteilung

Satz: Eine Zufallsvariable X hat eine Poisson-Verteilung mit dem Parameter μ , wenn sie die Wahrschkft

$$q(x) = \frac{\mu^x}{x!} e^{-\mu}$$

besitzt. Erwartungswert und Streuung sind $E(X) = \mu, \sigma^2 = \mu \rightarrow \sigma = \sqrt{\mu}$

→ M = Anz. Elemente in der Grundmenge mit dem Merkmal

N = Anz. Elem. in d. Grundmenge

n = Anz. Elem. in d. Stichprobe

x = Anz. Elem. in d. Stichp. mit d. Merkmal

$q(x)$ = Wahrsch., dass es x Elem. mit d. Merkmal in der Stichprobe hat

Bei max: z.B. max 2 Äpfel → $\sum_{x=0,1,2} q(x) = q(0) + q(1) + q(2)$

Bei min: z.B. mind 2 Äpfel → $1 - q(0) - q(1)$

Beispiele TR

Hypergeom A1) $\sum_{x=0}^2 (\text{hypergeom}^{N, M, n, x}(10, 6, 5, x))$

Binom A5) $\text{binom}^{n, p, x}(10, \frac{3}{7}, \frac{x}{7}) = 9,059$

Poisson A8) $\mu = 1 \quad P(x \geq 5) = 1 - q(0) - \dots - q(4)$

$\Rightarrow 1 - \sum_{x=0}^4 (\text{poissonapprox}^{\mu, x}(1, x)) \rightarrow \text{ans} - 365 = \underline{\underline{1,34}}$

MS10 Zusammenfassung 3

Stetige Zufallsvariablen

Definition: Eine Zufallsvariable X ist stetig, wenn gilt:

$P(\{\omega \in \Omega \mid X(\omega) = x\}) = 0$, für alle $x \in \mathbb{R}$, d.h. die Wertemenge der Zufallsvariablen X ist ein Kontinuum, und wenn ihre Verteilungsfunktion

$F_X(x) = P(X(\omega) \leq x)$ mit Hilfe einer nicht-negativen, integrierbaren Funktion $f(x)$ durch das Integral $F(x) = \int_{-\infty}^x f(t) dt$ ausgedrückt werden kann.

Die Funktion $f(x)$ heißt dann die Dichtefunktion/Dichte der Verteilung von X .

Folgerung: Für die Verteilungsfkt. $F_X(x)$, resp. Dichte $f(x)$ einer Zufallsvariablen X folgt:

a) Die Wahrscheinlichkeit dafür, dass eine Beobachtung von X im Intervall $(a, b]$ liegt, berechnen wir aus der Gleichung

$$P(X(\omega) \leq b) = P(X(\omega) \leq a) + P(a < X(\omega) \leq b)$$

Also:

$$P(a < X(\omega) \leq b) = \int_a^b f(t) dt$$

b) Es gilt:

$$P(a \leq X(\omega) \leq b) = P(a < X(\omega) \leq b)$$

c) Die Dichte $f(x)$ erfüllt die Gleichung

$$\int_{-\infty}^{\infty} f(t) dt = 1$$

$$E(X) = \int_{-\infty}^{\infty} t \cdot f(t) dt$$

$$\text{Var}(X) = \int_{-\infty}^{\infty} [t - E(X)]^2 \cdot f(t) dt$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

Gleichverteilung

- Verteilung einer Variablen X , deren Dichte gegeben ist durch

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a < x < b \\ 0 & \text{sonst} \end{cases}$$

$$E(X) = \frac{1}{b-a} \int_a^b t \cdot dt = \frac{a+b}{2}$$

$$\begin{aligned} \text{Var}(X) &= \frac{1}{b-a} \int_a^b t^2 dt - \left[\frac{a+b}{2} \right]^2 \\ &= \frac{(b-a)^2}{12} \end{aligned}$$

Normalverteilung

Definition: Eine Zufallsvariable X mit Wertebereich $\mathbb{R}$ heißt normalverteilt mit den Parametern $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}^+$, wenn ihre Dichte gegeben ist durch

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Dichte: $f(x)$

- Verteilungsfkt: Φ

$$\Phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Beispiel 6.1

Die Abbildung 6.1 zeigt die Dichtefunktion $\phi(x)$ und die Normalverteilung $\Phi(x)$ für $\sigma \approx 0.5$ und $\sigma = 1$ mit $\mu = 2$.

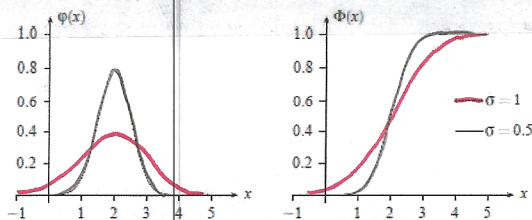


Abbildung 6.1: Dichte $\phi(x)$ und Normalverteilung $\Phi(x)$

Satz: Sei X eine normalverteilte Zufallsvariable mit der Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \text{ Der}$$

Erwartungswert und die Varianz sind:

$$E(X) = \mu, \quad \text{Var}(X) = \sigma^2$$

Standardnormalverteilung

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \text{ mit der Dichte } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

- Die Standardnormalverteilung hat den Erwartungswert $\mu=0$ und die Standardabweichung $\sigma=1$

Satz: Sei $x = X(\omega)$ eine beliebige normalverteilte Zufallsvariable mit Erwartungswert $E(X) = \mu$ und Standardabweichung σ . Also:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{Sei } z = Z(\omega) = \frac{X(\omega) - \mu}{\sigma} = \frac{x - \mu}{\sigma}$$

Dann ist $Z(\omega)$ standardnormalverteilt,

also:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt,$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

und es gilt:

$$P(X \leq x) = P(Z \leq z) \Leftrightarrow \Phi(x) = \Phi(z)$$

Satz: Sei $x = X(\omega)$ eine normalverteilte Zufallsvariable. Dann gilt:

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0,68268$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0,95450$$

Sei $x = X(\omega)$ eine beliebige normalverteilte Zufallsvariable mit Erwartungswert $E(X) = \mu$ und Standardabweichung σ . Wir bilden die neue Zufallsvariable

$$z = Z(\omega) = \frac{X(\omega) - \mu}{\sigma} = \frac{x - \mu}{\sigma}$$

Die Zufallsvariable Z hat einen Erwartungswert $E(Z) = 0$ und eine Standardabweichung $\sigma = 1$, denn

$$E(Z) = E\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma} E(X - \mu) = \frac{1}{\sigma} [E(X) - \mu] = \frac{1}{\sigma} [\mu - \mu] = 0,$$

und

$$\begin{aligned} \text{Var}(Z) &= E(Z^2) - [E(Z)]^2 = E(Z^2) = E\left(\frac{X^2 - 2\mu X + \mu^2}{\sigma^2}\right) \\ &= \frac{1}{\sigma^2} E(X^2 - 2\mu X + \mu^2) \\ &= \frac{1}{\sigma^2} (E(X^2) - 2\mu E(X) + \mu^2) \\ &= \frac{1}{\sigma^2} (\sigma^2 + \mu^2 - 2\mu^2 + \mu^2) = 1. \end{aligned}$$

Daraus folgt für die Standardabweichung von Z : $\sigma = \sqrt{\text{Var}(Z)} = 1$.

MSTO Zusammenfassung 3

6.2.2 Approximation der Binomialverteilung durch die Normalverteilung

Viele Probleme, welche binomial verteilt sind, lassen mit der Normalverteilung approximieren und dadurch einfacher lösen. Die Binomialverteilung einer Zufallsvariablen $x = X(\omega)$ lautet

$$q(x) = \binom{n}{x} p^x (1-p)^{n-x},$$

mit dem Erwartungswert $E(X) = n \cdot p$ und der Standardabweichung $\sigma = \sqrt{np(1-p)}$.

Beispiel 6.4

Ein Bernoulli-Versuch, mit den Ergebnissen 0 und 1 und mit der Wahrscheinlichkeit $p = 0.6$ für 1, wird zehn mal wiederholt, also $n = 10$. Die Zufallsvariable X , welche die Anzahl Einer zählt, ist binomialverteilt mit

$$E(X) = n \cdot p = 6, \quad \sigma = \sqrt{np(1-p)} = 1.55.$$

Wir approximieren die Wahrscheinlichkeitsfunktion $q(x)$ mit der Dichte

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi} \cdot 1.55} e^{-\frac{(x-6)^2}{4.80}}, \quad \mu = E(X).$$

Die Werte der Binomialverteilung sind

x	0	1	2	3	4	5	6	7	8	9	10
q(x)	0.0001	0.0016	0.0106	0.0425	0.1115	0.2007	0.2508	0.2150	0.1209	0.0403	0.0060

Frage:

Wie gross ist die Wahrscheinlichkeit, dass x im Intervall $[4, 7]$ liegt?

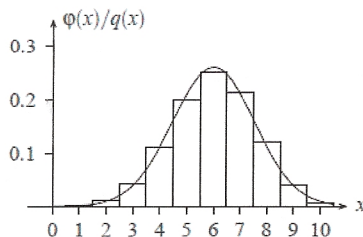
Lösung:

Mit der Binomialverteilung:

$$P(4 \leq X \leq 7) = \sum_{x=4}^7 q(x) = \sum_{x=4}^7 \binom{10}{x} 0.6^x 0.4^{10-x} = 0.1115 + 0.2007 + 0.2508 + 0.2150 = 0.7779.$$

Mit der Normalverteilung:

$$P(4 \leq X \leq 7) = \int_{3.5}^{7.5} \varphi(x) dx = 0.2575 \int_{3.5}^{7.5} e^{-\frac{(x-6)^2}{4.80}} dx = 0.7802.$$



Exponentialverteilung

- interessiert an den Zeit-
differenzen bzw. derer Verteilung

Definition: Eine Zufallsvariable $x = X(\omega)$ mit Wertebereich $\mathbb{R}^+$ heisst exponentialverteilt mit dem Parameter $\lambda > 0$, wenn ihre Dichte gegeben ist durch:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

Satz: Die Verteilungsfkt $F(x)$ einer exponentialverteilten Zufallsvar. X

Ist:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

Satz: Erwartungswert und Varianz einer exponentialverteilten Zufallsvariablen sind:

$$E(x) = \frac{1}{\lambda} \quad , \quad \text{Var}(x) = \frac{1}{\lambda^2}$$

