Wahrscheinlichkeitstheorie und Statistik

Dr. D. Adams

Fachhochschule Nordwestschweiz
Hochschule für Technik
IMN

FS 2018

Fachhochschule Nordwestschweiz
Hochschule für Technik

Bibliographie

Lothar Papula.
Mathematik für Ingenieure und Naturwissenschaftler – Ein Lehr- und ArbeitsBuch für das Grundstudium, volume 2.
Vieweg + Teubner, Wiesbaden, 2009.

T. Sigg.
Grundlagen der Differentialgleichungen für Dummies.
Für Dummies Series. Wiley VCH Verlag GmbH, 2012.
ISBN 9783527707959.

Griechisches Alphabet

Grossbuchst.	Kleinbuchst.	
A	α	Alpha
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ϵ, ε	Epsilon
Z	ζ	Zeta
H	η	Eta
Θ	θ, ϑ	Theta
I	ι	Iota
K	κ, $\varkappa$	Kappa
Λ	λ	Lambda
M	μ	My
N	ν	Ny
Ξ	ξ	Xi
O	o	Omikron
Π	π, ϖ	Pi
P	ρ, ϱ	Rho
Σ	σ, ς	Sigma
T	τ	Tau
Y	υ	Ypsilon
Φ	ϕ, φ	Phi
X	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega

Beschreibende Statistik

Was ist Stochastik

Definition (Stochastik)

Beschreibung und Untersuchung von Ereignissen, die vom Zufall beeinflusst werden.
Wahrscheinlichkeitstheorie + Statistik

Definition (Statistik)

Analyse von Daten, die durch Zufall beeinflusst sind

Einsatzgebiete Statistik:

- Technik, Physik
- Meteorologie
- Ökonomie

Arbeitsweise Statistik

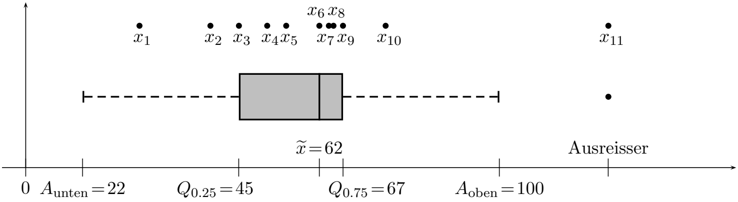
- Formulierung Problem
- Planung Experiment
- Ausführung Experiment
- Beschreibung experimentelle Daten
- Schluss von Stichprobe auf Grundgesamtheit

Beispiel Neonröhren

- Formulierung Problem:
  Wie gross ist Lebensdauer der Neonröhren, die an FHNW verwendet werden?
- Planung Experiment:
  Test einer Röhre genügt nicht. Alle können nicht getestet werden. Wir testen 11 Röhren.
- Ausführung Experiment:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
24	39	45	51	55	62	64	65	67	76	123

  Angaben in Monaten
- Beschreibung experimentelle Daten:
  Lageparameter: Durchschnitt und Standardabweichung
  $\overline{x} = 61 \pm 25.25$
- Schluss von Stichprobe auf Grundgesamtheit:
  Durchschnittliche Lebensdauer 61 Monate


x_6 x_8
x_1
x_2
x_3
x_4
x_5
x_7
x_9
x_{10}
x_{11}
$\widetilde{x}=62$
Ausreisser
0
$A_{\text{unten}}=22$
$Q_{0.25}=45$
$Q_{0.75}=67$
$A_{\text{oben}}=100$

Definition (Mittelwert)

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

Stichprobenumfang: n

Definition (Grundgesamtheit)

Menge der Elemente, die untersucht werden soll.

Definition (Standard-Abweichung)

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$$

Definition (Stichprobe)

Teilmenge der Grundgesamtheit, die untersucht wird.

Metrische / diskrete / quantitative Grössen

Quantitative Merkmale unterteilt in **diskrete** und **metrische**

- **diskrete** Merkmale werden meistens von mehreren Merkmalsträgern angenommen
- bei **diskreten** Merkmalen ist es sinnvoll zu zählen wie of eine Merkmals-Ausprägung angenommen wird
- bei **metrischen** Grössen gibt es zu fast jedem Merkmalsträger eine, von anderen verschiedene Merkmal-Ausprägung (sogar bei grossen Stichproben)

quantitativen Grössen ist die Ordnung meist willkürlich: rot¿grün; Zürich¡ Zug ¡ Aarau

Metrische (stetige) Grössen (m), diskrete Grössen (d), qualitative Merkmale (q)

Ordnen sie zu:

- Windgeschwindigkeit auf dem Arbeitsweg
- Sonnenschein-Dauer am letzten Tag im Monat
- Anzahl Regentage im April
- Luftdruck während 24 Stunden
- Stau-Stunden am Gotthard
- Anzahl Lastwagen durch Belchentunnnel pro Stunde
- Zivilstand der Studierenden in Klasse
- steuerbares Einkommen der Studierenden in Klasse
- abgeschlossene Diplome der Studierenden in Klasse

Metrische (stetige) Grössen (m), diskrete Grössen (d), qualitative Merkmale (q)

Ordnen sie zu:

- Windgeschwindigkeit (m)
- Sonnenschein-Dauer am letzten Tag im Monat (m)
- Anzahl Regentage im April (d)
- Luftdruck (m)
- Stau-Stunden am Gotthard (m)
- Anzahl Lastwagen durch Belchentunnnel (d/m)
- Zivilstand (q)
- steuerbares Einkommen (m)
- abgeschlossene Diplome (q)

Darstellung von Daten

Graphische Darstellung

- Häufigkeitstabellen (Tabelle, m/d)
- Histogramme (Plot, m/d)
- Kreisdiagramme (q)

Häufigkeitstabellen:

- Anzahl Klassen (Richtwert) $k \approx \sqrt{n}$
- Klassenbreite $d \approx \frac{x_{\max} - x_{\min}}{k}$
- Intervalle $[a_i, a_{i+1}[$ (Daten auf Grenzen konsistent zu rechten Klasse gezählt)

Häufigkeitsverteilungen

BSP 1

Visualisieren Sie die Anzahl Betriebsstörungen an Baumaschinen.

i	0	1	2	3
h_i	48	38	10	4

Histogramm

BSP 2

Erstellen Sie ein Diagramm zum Verhältnis Studentinnen zu Studenten in der Klasse

BSP 3

Erstellen Sie eine Häufigkeitstabelle und ein Histogramm zu den Daten der Zugfestigkeit (Walzdraht, S. 6)

Lageparameter

Definition (Median)

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ ungerade} \\ 1/2 \cdot \left[x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right] & n \text{ gerade} \end{cases}$$

Definition (Quartile $Q_{0.25}$, $Q_{0.75}$)

Oberhalb von $Q_{0.75}$, liegt $1/4$ der Messungen, unterhalb von $Q_{0.25}$ liegt $1/4$ der Messungen

Definition (Ausreissergrenzen)

$A_{\mathrm{unten}} = Q_{0.25} - 1.5 \cdot d_Q$ und $A_{\mathrm{oben}} = Q_{0.75} + 1.5 \cdot d_Q$
Quartilsweite: $d_Q = Q_{0.75} - Q_{0.25}$

Lageparameter (praktisch)

Liste der Länge n

- Ordne die Liste und berechne den Median
- Teile die Werte in zwei Listen auf
  - n ungerade: Der Median wird aus beiden Listen ausgeschlossen.
  - n gerade: Teile die Liste in exakt gleich lange Listen auf
- $Q_{0.25}$ ist der Median der ersten Liste, $Q_{0.75}$ ist der Median der zweiten Liste

Median/Quartile/Ausreisser

5.3	3.8	4.0	19.5	5.0	4.9	2.2	4.1	3.1	5.5

Ordnen:

2.2	3.1	3.8	4.0	4.1	4.9	5.0	5.3	5.5	19.5

Mittelwert: $\overline{x} = 5.74$
Median: $\tilde{x} = \frac{1}{2}(4.1 + 4.9) = 4.5$
Quartile: $Q_{0.25} = 3.8$ und $Q_{0.75} = 5.3$
Aussreissergrenzen: $A_{\mathrm{unten}} = Q_{0.25} - 1.5 \cdot d_Q = 1.55$ und
$A_{\mathrm{oben}} = Q_{0.75} + 1.5 \cdot d_Q = 7.55 \Rightarrow 19.5$ ist Ausreisser
Erstelle Box- und Whiskperplot
Ohne Ausreisser (Mittelwert/Median):
Mittelwert: $\overline{x} = 4.2$
Median: $\tilde{x} = \frac{1}{2}(4.1 + 4.9) = 4.1$

Mittelwert vs. Median

Median ist stabiler gegenüber Ausreissern als Mittelwert

Mittlere Geschwindigkeit?

Strecke [km]	20	20	20
Geschwindigkeit [km/h]	40	120	80

$$\tilde{v} = \frac{\sum_i^n s_i}{\sum_i^n t_i} = \frac{\sum_i^n s_i}{\sum_i^n \frac{s_i}{v_i}}$$

$$\overline{v}_{\mathrm{h}} = 65.45 \text{ km/h}$$

Mittlere Verzinsung?

Jahr	1	2	3
Zins %	8.5	12.2	-4.5

$$K = K_0 \cdot (1 + r_1) \cdot (1 + r_2) \cdot (1 + r_3)$$

$$\overline{v}_{\mathrm{g}} = 5.15 \text{ \%}$$

Lageparameter

Definition (Arithmetisches Mittel)

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

Definition (Harmonisches Mittel)

$$\frac{1}{\overline{x}_\mathrm{h}} = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{1}{x_i}$$

Definition (Geometrisches Mittel)

$$\overline{x}_\mathrm{g} = (x_1 \cdot x_2 \ \ldots \ x_n)^{1/n} = \left(\prod_{i=1}^{n} x_i \right)^{1/n}$$

Zufall und Ereignis

5 Zufallsexperimente

1. **Permutation**: 4 Bücher ins Regal stellen. $4 \cdot 3 \cdot 2 \cdot 1 = 24$ Möglichkeiten.
   Bei n Büchern $n!$ Möglichkeiten.
2. **Geordnete Stichprobe ohne Zurücklegen**: Aus 10 Büchern 3 ins Regal stellen. $10 \cdot 9 \cdot 8 = \frac{10 \cdot 9 \ldots 2 \cdot 1}{7 \cdot 8 \ldots 2 \cdot 1} = 720$ Möglichkeiten.
   Aus n Büchern k ins Regal stellen. $\frac{n!}{(n-k)!}$ Möglichkeiten.
3. **Variation** (Geordnete Stichprobe mit Zurücklegen): Zahlenschloss mit 3 Ringen, auf jedem Ring hat es 10 Zahlen (0,1,2, ..., 9): $10 \cdot 10 \cdot 10 = 10^3 = 1000$ Möglichkeiten
   k Ringe mit n Zahlen darauf. n^k Möglichkeiten

5 Zufallsexperimente (forts)

1. **Kombination** (Ungeordnete Stichprobe ohne Zurücklegen): 3-er Gruppen aus 10 Studierenden: geordnete 3er Gruppen $\frac{10!}{7!} = 720$ Möglichkeiten; Möglichkeiten 3 Studierende in 3er Gruppe anzuordnen $3 \cdot 2 \cdot 1 = 6 \Rightarrow \frac{10!}{7! \cdot 3!} = 120$ Möglichkeiten k-er Gruppen aus n Studierenden: $\frac{n!}{(n-k)! \cdot k!}$ Möglichkeiten
2. **Ungeordnete Stichprobe mit Wiederholung**: $k = 6$ identische Kugeln auf $n = 9$ Kisten verteilen (in jeder Kiste dürfen mehrere Kugeln liegen): $\frac{(n+k+1)!}{k!} = 3003$

Definition (Zufallsexperiment)

Vorgang

- beliebig oft wiederholbar und
- Ausgang ungewiss

(innerhalb einer Menge möglicher Ergebnisse).

Definition (Stichprobenraum)

Menge S aller Ausfallsmöglichkeiten.

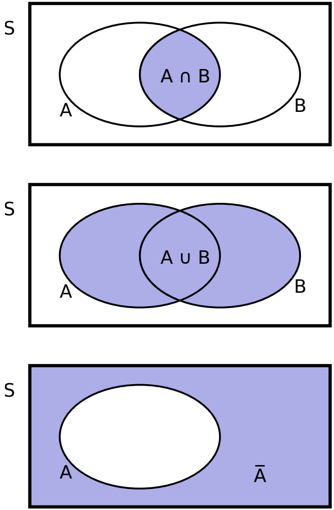
Geben Sie den Stichprobenraum an

- Werfen einer Münze
- Werfen eines Würfels
- Ziehung einer Lottozahl
- Kontrolle eines Haushalts durch die Billag

- Werfen einer Münze $S = \{K, Z\}$
- Werfen eines Würfels $S = \{1, 2, 3, 4, 5, 6\}$
- Ziehung einer Lottozahl $S = \{1, 2, 3, \ldots, 44, 45\}$
- Billag $S = \{\text{Fernseher}, \text{Radio}, 0\}$

Definition (Verknüpfung von Ereignissen)

- und-Verknüpfung $A \cap B$
- oder-Verknüpfung $A \cup B$
- Gegenereignis $\overline{A}$



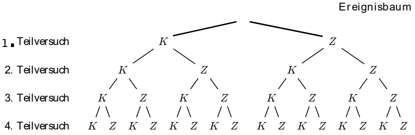
Theorem (Morgansche Regeln)

- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Beschreiben Sie in Worten und Mengen

$$S = \{1,2,3,4,5,6\},\ A = \{1,3,5\},\ B = \{2,4,6\},\ C = \{3,6\}$$

- $\overline{A} = \{2,4,6\} = B$, Würfeln einer geraden Zahl
- $\overline{B} = \{1,3,5\} = A$, Würfeln einer ungeraden Zahl
- $A \cup B = \{1,2,3,4,5,6\} = S$, Sicheres Ereignis
- $A \cap B = \emptyset$, unmögliches Ereignis
- $A \cap C = \{3\}$, Würfeln der 3



Theorem (Produktregel)

Besteht ein zusammengesetzter Versuch aus m unabhängigen Teilversuchen mit jeweils $[n_1, n_2, n_3, n_4, \ldots, n_m$ *Ausfallsmöglichkeiten, dann besitzt der Versuch*

$$n_1 \cdot n_2 \ \ldots \ \cdot n_m \ \textit{Ausfälle.}$$

Permutationen der Ziffern 1 bis 5?

$$P(5) = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$$

Variation 3-ter Ordnung mit Zurücklegen.

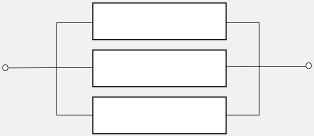
Anzahl mögliche dreistellige Zahlen?

Ziffern 1 bis 9.

$$V(9;3) = 9^3 = 729$$

Variation 3-ter Ordnung mit Zurücklegen.

Anzahl mögliche Schaltungen



5 verschiedene Widerstände $R_1, \ldots, R_5$
Jeder Widerstand nur ein Mal verwenden

$$C(5;3) = \binom{5}{3} = \frac{5!}{3! \cdot 2!} = 10$$

Kombination 3-ter Ordnung ohne Zurücklegen.

Anzahl mögliche ungeordneter Stichproben einer Lieferung von Batterien?

Gelieferte Batterien: 100, Stichprobe: 10

$$C(12;3) = \binom{100}{10} = \frac{100!}{10! \cdot 90!} = 17\,310\,309\,456\,440$$

Kombination 10-ter Ordnung ohne Zurücklegen.

Anzahl mögliche dreistellige Zahlen?

Ziffern 1 bis 9. Jede Ziffer nur ein Mal verwenden. (Ziffern 1 bis 9: kein Probleme mit führenden 0)

$$V(9;3) = \frac{9!}{(9-3)!} = 7 \cdot 8 \cdot 9 = 504$$

Variation 3-ter Ordnung ohne Zurücklegen.

Pferdetoto

Dreierwette: Zieleinflauf der ersten drei Pferde. Anzahl Möglichkeiten bei 10 Pferden?

$$V(10;3) = \frac{10!}{(10-3)!} = 720$$

Variation 3-ter Ordnung ohne Wiederholungen.

Anzahl mögliche Schaltungen

5 verschiedene Widerstände $R_1, \ldots, R_5$
Jeder Widerstand bis zu 3 Mal verwenden

$$C_w(5;3) = \binom{5+3-1}{3} = \frac{7!}{4! \cdot 3!} = 35$$

Kombination 3-ter Ordnung mit Zurücklegen.

	ohne Wiederholung	mit Wiederholung	
Kombination k-ter Ordnung	$C(n;k) = \binom{n}{k}$	$C_w(n;k) = \binom{n+k-1}{k}$	ungeordnete Stichproben
Variation k-ter Ordnung	$V(n;k) = \frac{n!}{(n-k)!}$	$V_w(n;k) = n^k$	geordnet Stichproben
	Ziehung ohne Zurücklegen	Ziehung mit Zurücklegen	

Definition (Binomialkoeffizient)

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Definition (Wahrscheinlichkeit (theoretisch))

$$p = \frac{g}{m}$$

g: Anzahl günstige Fälle
m: Anzahl mögliche Fälle

Wahrscheinlichkeit 6 beim Würfeln?

$$m = 6, \quad g = 1 \;\Rightarrow\; p = \frac{1}{6}$$

Wahrscheinlichkeit Kopf beim Werfen Münze?

$$m = 2, \quad g = 1 \;\Rightarrow\; p = \frac{1}{2}$$

6 Richtige Zahlen beim Lotto (6 mal 45 Zahlen)?

$$m = \binom{45}{6} = 8\,145\,060$$

$$g = 1 \;\Rightarrow\; p = \frac{1}{m} = 0.000\,000\,123$$

Vergleiche Wahrscheinlichkeit mit Strecke Brugg-Paris

Strecke: 611 000 m.

$$p \approx \frac{0.07}{611\,000}$$

ca. 7 cm!

10 Nüsse davon 3 verdorben.

Wahrscheinlichkeit mit einem Griff 2 gute Nüsse zu finden.

$$m = \binom{10}{2} = 45$$

$$g = \binom{7}{2} = 21$$

Hier ist es einfacher mit geordneten Stichproben zu rechnen, obwohl eine ungeordnete Stichprobe !

Geordnete Stichprobe vs. ungeordnete Stichprobe

Bei vielen Experimenten können geordnete Stichproben betrachtet werden. Das vereinfach oft die Rechnungen. Dabei muss m und g konsequent für geordnete Stichproben berechnet werden

Definition (Wahrscheinlichkeit (experimentell))

$$h(s_i) = \frac{\text{Anz.desAuftretensvon} s_i}{\text{Anz.Versuche}} = \frac{n_i}{N}$$

Für Wahrscheinlichkeiten $p(s_i)$ (oder $h(s_i)$ gilt:)

Theorem (Axiome der Wahrscheinlichkeit)

Sei $S = \{s_1, s_2, \ldots s_n\}$ der Stichprobenraum des Versuchs (d.h. $s_i \cup s_j = 0$ für $i \neq j$):

- $0 \leq p(s_i) \leq 1$
- $p(s_1) + p(s_2) + \ldots + p(s_n) = 1$
- *Spezielles Ereignis, z.B $A = \{s_1, s_2, s_3\}$:*

$$p(A) = p(s_1) + p(s_2) + p(s_3)$$

Experimentelle Wahrscheinlichkeit (Simulation)

Führen Sie mit matlab folgenden Versuch durch: Mit einem Würfel wird 1000 mal gewürfelt. Bei jedem Wurf wird notiert, wie viele mal insgesamt eine 3 gewürfelt wurde.
Befehle: `for i=1:mmax ... end, r>0, randi([1 3])`

```
mmax=10^3;
sta=(1:mmax+1)*0; % leere Liste initialisieren
for i=1:mmax
  r = randi([1 6]);    % Würfeln
  sta(i+1)=(sta(i)+(r==3));
end
```

Berechnen und plotten Sie die Wahrscheinlichkeit, dass eine 3 gewürfelt wird als Funktion der Versuchsnummer.
Gegen welchen Wert konvergiert die relative Häufigkeit? Befehle: `plot, ./ , end`

Experimentelle Wahrscheinlichkeit forts.

```
relh=sta./(1:mmax+1);
plot((1:mmax+1),relh)
relh(end-2:end)
```

Theorem (Additionssatz)

allgemein:

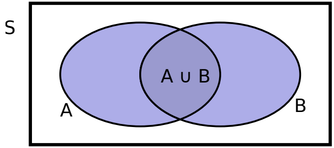
$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

also für A und B elementfremd:

$$p(A \cup B) = p(A) + p(B)$$

Wahrscheinlich Gegenereignis

$q = p(\overline{A}) = 1 - p(A)$



Theorem (Multiplikationssatz)

A und B beziehen sich auf Teilversuche:

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

Würfel

$A = \{1,3,5\}$ und $B = \{3,4,5,6\}$. Berechnen Sie $p(A \cup B)$.

$$\begin{aligned}
p(A \cup B) &= p(A) + p(B) - p(A \cap B) \\
&= \left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) + \left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) - \left(\frac{1}{6} + \frac{1}{6}\right) \\
&= \frac{5}{6}
\end{aligned}$$

Wahrscheinlichkeit drei mal hintereinander Sechs zu würfeln.

$$p(A_1 \cap A_2 \cap A_3) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{6^3} = 0.00463$$

Beispiel Versuch mit p und q

p Wahrscheinlichkeit für Erfolg, $q = 1 - p$ Wahrscheinlichkeit für MissErfolg. Für n Versuche berechnen Sie

- Wahrsch. für n Erfolge p^n
- Wahrsch. für n Misserfolge $q^n = (1 - p)^n$
- Wahrsch. mind. ein Erfolg $1 - q^n$
- Wahrsch. erster Erfolg beim letzten Versuch: $q^{n-1} \cdot p$