



Wahrscheinlichkeit und Statistik
für Bachelor Mechatronik
Studiengang FTB-MEC-VZ

Georgia Thurner

24. September 2024

Inhaltsverzeichnis

1	Einleitung	1
2	Explorative Statistik	5
2.0.1	Datenquellen	5
2.0.2	Datenanalyse	6
3	Beschreibende Statistik	7
3.1	Grundbegriffe	7
3.1.1	Merkmalstypen	7
3.1.2	Skalierung von Merkmalen	9
3.1.3	Häufigkeit	10
3.1.4	Klassierung der Daten	11
3.2	Graphische Darstellung von Daten	12
3.2.1	Diagramme zur Häufigkeits-Darstellung	12
3.2.2	Empirische Verteilungsfunktion	16
3.3	Lage- und Streumaße	17
3.3.1	Modalwert	17
3.3.2	Quantile, Quartilsabstand	17
3.3.3	Box-Plot	18
3.3.4	Arithmetisches Mittel	21
3.3.5	Spannweite	22
3.3.6	Empirische Varianz und empirische Standardabweichung	22
3.4	Korrelationsanalyse	24
3.4.1	Zweidimensionale Häufigkeitsverteilungen	24
3.4.2	Durchführung Korrelationsanalyse	27
4	Wahrscheinlichkeitsrechnung	31
4.1	Grundbegriffe	31
4.2	Begriffe und Verknüpfung von Ereignissen	32
4.3	Wahrscheinlichkeit	34
4.3.1	Laplace-Experiment	35
4.3.2	Beispiele zur Berechnung von Wahrscheinlichkeiten	36
4.4	Bedingte Wahrscheinlichkeiten	37
4.4.1	Summen- und Produktregeln	41
4.5	Zufallsvariable	42
4.5.1	Wahrscheinlichkeitsfunktion	43
4.6	Verteilungsfunktion	43

4.7	Kennwerte einer Wahrscheinlichkeitsverteilung	46
4.7.1	Erwartungswert, Varianz und Standardabweichung einer diskreten Zufallsvariablen	47
4.7.2	Erwartungswert, Varianz und Standardabweichung einer stetigen Zufallsvariablen	48
4.7.3	Erwartungswert, Varianz und Standardabweichung einer linearen Funktion der Zufallsvariablen X	49
4.8	Spezielle Wahrscheinlichkeitsverteilungen	51
4.8.1	Binomialverteilung	51
4.8.2	Hypergeometrische Verteilung	53
4.8.3	Poissonverteilung	55
4.8.4	Stetige Gleichverteilung	57
4.8.5	Exponentialverteilung	58
4.8.6	Weibull-Verteilung	60
4.8.7	Gaußsche Normalverteilung	61
4.8.8	6σ	63
4.8.9	Berechnung der Wahrscheinlichkeiten der Normalverteilung	63
4.9	Zentraler Grenzwertsatz	65
4.9.1	Summe zweier binomialverteilter Zufallsvariablen	65
4.9.2	Summe zweier normalverteilter Zufallsvariablen	65
4.9.3	Summe zweier beliebiger Zufallsvariablen	66
4.9.4	Unabhängige Zufallsvariablen mit gleichem Erwartungswert und gleicher Varianz	66
4.9.5	Zentraler Grenzwertsatz	66
5	Schließende Statistik	69
5.1	Parameterschätzungen	69
5.1.1	Vor-Nachteile	70
5.1.2	Qualitätskriterien für Punktschätzer	70
5.1.3	Schätzfunktionen der wichtigsten Parameter	71
5.1.4	Schätzfunktionen bei speziellen Verteilungen	71
5.1.5	Maximum-Likelihood-Methode	72
5.2	Intervallschätzungen	73
5.3	Hypothesentest	79
5.3.1	Unterschied Hypothesentest zu Parameterschätzung	79
5.3.2	Definition Hypothesentests	80
5.3.3	Aufbau von Hypothesentests	81
5.3.4	(Mögliche) Fehlentscheidungen	83
5.3.5	Gauß-Test: Test für μ eines normalverteilten Merkmals bei bekanntem σ	83
5.3.6	t -Test: Test für μ eines normalverteilten Merkmals bei unbekanntem σ	85
5.3.7	χ^2 -Test: Test für σ^2 eines normalverteilten Merkmals	85
5.3.8	Binomialtest	86

5.3.9	Anwendungen	89
-------	-----------------------	----

Abbildungsverzeichnis

1.1	Weltstatistiktag	2
3.1	Balkendiagramm	12
3.2	Säulendiagramm	13
3.3	Stabdiagramm	13
3.4	Liniendiagramm	14
3.5	Kreisdiagramm	14
3.6	Histogramm	16
3.7	Box Plot	19
3.8	Box Plot	20
3.9	Streudiagramm	25
3.10	Anscombe 1	28
3.11	Anscombe 2	28
3.12	Anscombe 3	29
3.13	Anscombe 4	29
4.1	Wahrscheinlichkeitsbaum	38
4.2	Wahrscheinlichkeitsbaum	40
4.3	Wahrscheinlichkeitstableau	40
4.4	Wahrscheinlichkeitsdicht	45
4.5	Binomialverteilung	52
4.6	Hypergeometrische Verteilung	54
4.7	Poissonverteilung	56
4.8	Gleichverteilung	58
4.9	Exponentialverteilung	59
4.10	Weibullverteilung	61
4.11	Gaußsche Normalverteilung	62
5.1	Konfidenzintervall	75
5.2	t-student	77

1

Einleitung

Für wen?

Dieses Manuskript ist für die Studentinnen und Studenten im Bachelor-Studiengang Mechatronik an der FHV in Dornbirn erstellt.


Wozu Statistik?

Anwendung findet die Statistik praktisch in allen Disziplinen. Einige Beispiele:

- Medizin (statistische Auswertungen von Therapiestudien)
- Klimaforschung
- Qualitätskontrolle
- Wettervorhersagen (es regnet mit Wahrscheinlichkeit 70%)
- Versicherungswesen (Prämienkalkulation)
- Verkehrswesen (Studium von Warteschlangen - Ampelsteuerung)
- Epidemiologie (Modelle für die Ausbreitung von Krankheiten)
- Meinungsforschung
- Betriebswirtschaft
- Bankwesen (Portfolio-Analyse, Marketing Strategien)
- Telekommunikation (Modellierung von Verteilungen von Gesprächsdauern)
- Quantenphysik

Weltstatistiktag

Ein weiterer Hinweis auf die anerkannte Bedeutung der Statistik ist z.B. der Weltstatistiktag

Wien, 2020-10-19 – Jeden 20. Oktober macht der europäische Statistiktag auf die Bedeutung von vertrauenswürdigen Statistiken für Gesellschaft, Wirtschaft und Politik aufmerksam. Zusätzlich dazu wird an diesem Tag seit 2010 alle fünf Jahre der  **UN-Weltstatistiktag** begangen – am 20.10.2020 zum nunmehr dritten Mal, diesmal unter dem Motto "Connecting the world with data we can trust".

„Das Bedürfnis nach glaubwürdigen, wissenschaftlich objektiv erhobenen Zahlen und Daten wächst stetig. Gerade die Corona-Krise hat dies erneut verdeutlicht. Der neue Konjunkturmonitor der Statistik Austria bündelt mehr als 80 Konjunkturmerkmale, die vierzehntägig aktualisiert werden. Mit diesem Tool liefert die Statistik Austria umfassende Informationen für Interessierte aus Gesellschaft, Politik und Wirtschaft und damit einen weiteren Beitrag für eine faktenbasierte öffentliche Debatte“, kündigt Statistik-Austria-Generaldirektor Tobias Thomas an.

Abbildung 1.1: Statistik Austria: Pressemitteilung 12.354-194/20

Literatur

Als begleitende oder weiterführende Literatur kann zusätzlich folgendes herangezogen werden:

- Göllmann, Hübl, Pulham, Ritter et al.: Mathematik für Ingenieure: Verstehen, Rechnen, Anwenden. Band 1, Springer Vieweg Verlag
- H. Schiefer, F. Schiefer: Statistik für Ingenieure. Springer Verlag
- Degen, Lorscheid: Statistik Lehrbuch Oldenbourg Verlag
- Degen, Lorscheid: Statistik Aufgabensammlung Oldenbourg Verlag
- G. Teschl, S. Teschl: Mathematik für Informatiker. Band 2 Analysis und Statistik. Springer Vieweg Verlag
- L. Papula: Mathematik für Ingenieure und Naturwissenschaftler. Band 3 Verlag Vieweg
- Bronstein u.a.: Taschenbuch der Mathematik Edition Harri Deutsch

Begriffserläuterungen

In der Wahrscheinlichkeitstheorie untersucht man zufällige Prozesse mit festen als bekannt angegebenen Wahrscheinlichkeiten.

In der Statistik sollen aus beobachteten Daten Schlüsse über unbekannte Wahrscheinlichkeiten gezogen werden.

Die Wahrscheinlichkeitstheorie liefert dabei die mathematischen Grundlagen für die theoretische Untersuchung statistischer Verfahren.

Mit Statistiken grosser Datenmengen versucht man im Allgemeinen Phänomene und Zusammenhänge zu erkennen.

Der Begriff Stochastik fasst die beiden Teilgebiete

Statistik und Wahrscheinlichkeitstheorie

zusammen.

Inhalte

In den Ingenieurberufen wird die Statistik für Themen wie Qualitätskontrolle, statistische Prozesskontrolle eingesetzt. Durch Anwendung statistischer Methoden können Ergebnisse ohne großen Versuchsaufbau und damit kostengünstiger ermittelt werden.

Um hier aber statistisch fundierte Aussagen zu treffen, ist die Kenntnis statistischer Methoden erforderlich.

Die Statistik besteht im Wesentlichen aus drei Teilgebieten:

1. Die **deskriptive Statistik** (auch beschreibende Statistik):
Vorliegende Daten werden in geeigneter Weise beschrieben, aufbereitet und zusammengefasst. Mit ihren Methoden verdichtet man quantitative Daten zu Tabellen, graphischen Darstellungen und Kennzahlen. Bei einigen Institutionen ist, wie bei der amtlichen Statistik, die Erstellung solcher Statistiken die Hauptaufgabe.
2. Die **induktive Statistik** (auch schließende Statistik)
In der induktiven Statistik leitet man aus den Daten einer Stichprobe Eigenschaften einer Grundgesamtheit ab. Die Wahrscheinlichkeitstheorie liefert die Grundlagen für die erforderlichen Schätz- und Testverfahren.
3. Die **explorative Statistik** (auch Data-Mining):
Dies ist methodisch eine Zwischenform der beiden vorgenannten Teilbereiche. Sie bekommt als Anwendungsform zunehmend Bedeutung. Mittels deskriptiver Verfahren und induktiver Testmethoden sucht sie systematisch mögliche Zusammenhänge (oder Unterschiede) zwischen Daten in vorhandenen Datenbeständen und will sie zugleich in ihrer Stärke und Ergebnissicherheit bewerten. Die so gefundenen Ergebnisse lassen sich als Hypothesen verstehen, die erst, nachdem darauf aufbauende, induktive Testverfahren mit entsprechenden Versuchsplanungen sie bestätigen, als statistisch gesichert gelten können.

Dieses Manuskript behandelt die folgenden Themen in der angegebenen Reihenfolge: explorative Statistik (kurz) - deskriptive Statistik - Wahrscheinlichkeitstheorie - induktive Statistik.

2

Explorative Statistik

Im Jahr 1962 schlug John W. Tukey (US-amerikanischer Statistiker 1915 - 2000) in seinem Aufsatz "The Future of Data Analysis" eine neue wissenschaftliche Disziplin namens Datenanalyse vor. Der Fachbereich der explorativen Datenanalyse wurde mit Tukeys im Jahr 1977 erschienenem Buch "Exploratory Data Analysis" begründet. Tukey stellte darin einfache Diagramme (z.B. Box-Plots und Streudiagramme) vor, die in Kombination mit zusammenfassenden Statistiken (Mittelwert, Median, Quantile usw.) dabei helfen, ein Bild eines Datensatzes zu erhalten.

Mit der zunehmenden Verfügbarkeit von Rechenleistung und leistungsfähigen Datenanalyseprogrammen hat sich die explorative Datenanalyse weit über ihren ursprünglichen Rahmen hinaus weiterentwickelt. Die wichtigsten Triebkräfte dieser Disziplin waren die rasche Entwicklung neuer Technologien, der Zugang zu mehr und umfangreicheren Daten und der verstärkte Einsatz der quantitativen Analyse in einer Vielzahl von Disziplinen.

2.0.1 Datenquellen

Es gibt zahlreiche unterschiedlichste Datenquellen: Sensormessungen, Ereignisse, Text, Bilder und Videos. Das Internet der Dinge (engl. Internet of Things (IoT)) produziert ständig neue Informationsfluten. Ein Großteil dieser Daten liegt unstrukturiert vor: Bilder sind nichts anderes als eine Zusammenstellung von Pixeln, wobei jedes Pixel RGB-Farbinformationen (Rot, Grün, Blau) enthält. Texte sind Folgen von Wörtern und Nicht-Wortzeichen, die oft in Abschnitte, Unterabschnitte usw. gegliedert sind. Clickstreams sind Handlungsverläufe eines Nutzers, der mit einer Anwendung oder einer Webseite interagiert.

Tatsächlich besteht eine große Herausforderung der Datenwissenschaft darin, diese Flut von Rohdaten in verwertbare Informationen zu überführen. Um die hierfür statistische Konzepte anwenden zu können, müssen unstrukturierte Rohdaten zunächst aufbereitet und in eine strukturierte Form überführt werden. Eine der am häufigsten vorkommende Form strukturierter Daten ist eine Tabelle mit Zeilen und Spalten.

2.0.2 Datenanalyse

Die explorative Datenanalyse (kurz: EDA) ist eine Untersuchungsmethode, bei der man mithilfe von zusammengefassten statistischen Kenngrößen und grafischen Tools mehr über die Daten in Erfahrung bringen will und ermitteln will, was man in ihnen entdecken kann.

Typische Arbeitsschritte der EDA sind:

- Daten auf Vollständigkeit überprüfen
- Unstimmigkeiten in den Daten wie Ausreißer oder ungewöhnliche Beobachtungen ausmachen
- Übersicht über die Daten erstellen: Typ, Skalierung; Anzahl; Lage- und Streumaße
- Graphische Darstellung der Daten
- Muster in den Daten entdecken
- potenzielle Zusammenhänge zwischen den Merkmalen suchen und interessante Fragen oder Hypothesen aufstellen
- Die aufgestellten Hypothesen mithilfe formeller statistischer Methoden überprüfen

3

Beschreibende Statistik

Die beschreibende (deskriptive) Statistik will große Datenmengen fassbar machen. Dies geschieht einerseits durch übersichtliche Darstellung der Daten mittels Diagrammen und andererseits durch Bestimmung von markanten Kenngrößen dieser Daten.

3.1 Grundbegriffe

Begriff	Erläuterung
Merkmal	zu untersuchender Sachverhalt
Merkmalsausprägung	Konkretisierung des Sachverhalts
Merkmalsträger	Objekt, welches das Merkmal aufweist
Grundgesamtheit	Menge aller Merkmalsträger
Stichprobe	zufällig ausgewählte Teilmenge der Grundgesamtheit
Stichprobenumfang	Anzahl der Elemente der Stichprobe

Beispiel

Ein Betrieb produziere Kondensatoren mit einer vorgeschriebenen Kapazität von $100\mu\text{F}$. Im Produktionsablauf fallen jedoch nicht alle Kondensatoren identisch aus. Die Produktion toleriert Kondensatoren in einem Bereich von $\pm 1\mu\text{F}$, die anderen werden entsorgt. Um möglichst wenig Ausschuss zu haben, wird eine laufende Qualitätskontrolle durchgeführt. Dafür werden stündlich 10 Kondensatoren zufällig der Produktion entnommen und deren Kapazität bestimmt.

Merkmal	
Merkmalsausprägung	
Merkmalsträger	
Grundgesamtheit	
Stichprobe	
Stichprobenumfang	

3.1.1 Merkmalstypen

Merkmale können in unterschiedlichsten Ausprägungen auftreten, zB nur verbal und nicht durch Zahlen beschrieben (zufrieden, unzufrieden,...). Abhängig davon kann man

die Merkmale mathematisch bearbeiten, z.B. ein arithmetisches Mittel berechnen, oder nicht.

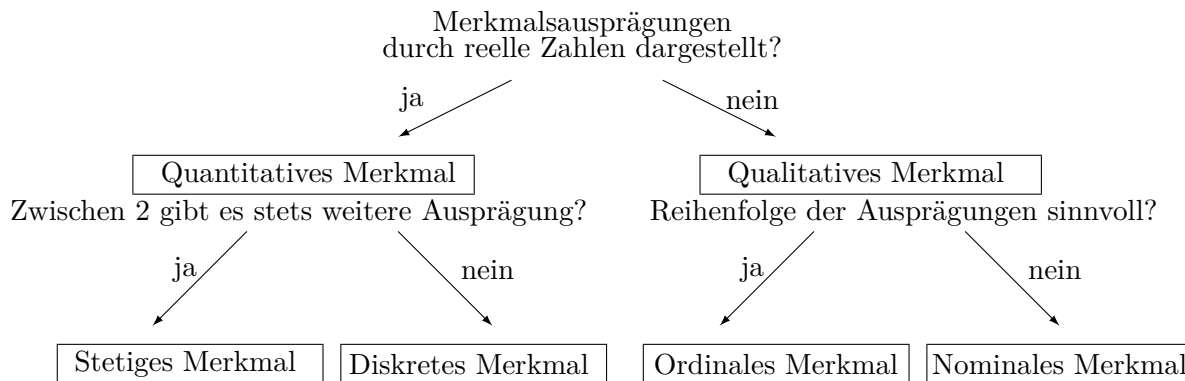
Beispiele

Merkmalsträger	Merkmal	Merkmalsausprägung
Mensch	Augenfarbe	grün, blau, braun
Österreichische Stadt	Einwohnerzahl	0,1,2,...mehrere Millionen
Student	Masse	reelle Zahlen zwischen (0;300)

Merkmale lassen sich in verschiedene Typen einteilen

- Qualitatives Merkmal
 - Ihre Ausprägungen sind verbal und nicht durch Zahlen beschrieben.
 - Werden diesen Merkmalsausprägungen Zahlen zugeordnet, so kann ein qualitatives Merkmal in ein quantitatives überführt werden.
 - Weitere Unterteilung
 - ordinale Merkmale
 - Die Merkmale können in eine sinnvolle Reihenfolge gebracht werden ("ist besser als") (zB Zufriedenheit)
 - nominale Merkmale
 - Es gibt keine sinnvolle Anordnung der Merkmale (zB Farbe)
- Quantitatives Merkmal
 - Merkmale, deren Ausprägungen durch reelle Zahlen dargestellt werden.
 - Diese werden weiter unterteilt in
 - diskrete Merkmale
 - wenn die Merkmalsausprägungen nur abzählbar viele Werte besitzen
 - stetig
 - wenn zwischen 2 Merkmalsausprägungen stets eine weitere Merkmalsausprägung sein kann

Zuordnung zu den Merkmalstypen



Beispiele

- qualitative Merkmale: Farben, Geschlechter, Nationalitäten
- quantitative Merkmale: Körpergröße, Uhrzeit, Datum

Anmerkung

Die Länge von Verschnittresten ist sicherlich ein quantitatives Merkmal, da man die Länge in Längeneinheiten misst. Ob man es als stetiges oder diskretes Merkmal betrachtet, hängt von der Messgenauigkeit ab.

Wenn man nur auf den Millimeter genau misst, also zB 17mm, 18mm handelt es sich streng genommen um ein diskretes Merkmal.

Falls man aber eine (im Prinzip) unendliche Messgenauigkeit annimmt, kann man das Merkmal auch als stetig auffassen.

3.1.2 Skalierung von Merkmalen

Hat man ein Merkmal, so ist für die Wahl der geeigneten statistischen Methode die Skalierung des Merkmals wichtig

- **nominalskaliertes Merkmal**
Für die Merkmale gibt es keine natürliche Reihenfolge
Beispiele: Postleitzahlen, Hausnummern, KFZ-Kennzeichen, Farben
- **ordinalskaliertes Merkmal**
Die Merkmalsausprägungen können in eine Rangordnung gebracht werden.
Beispiele: Intelligenzquotient, Schulnoten
- **intervallskaliertes Merkmal**
Zusätzlich können die Abstände zwischen den Ausprägungen des Merkmals verglichen werden. Die Intervallskala besitzt jedoch keinen natürlichen Nullpunkt. Differenzen der Ausprägungen sind sinnvoll berechenbar, nicht aber Quotienten

Beispiele: Temperaturskala in °C; nicht Schulnoten (da Differenzen wie zB Note 3-Note 2 = ? nicht sinnvoll sind)

- verhältnisskaliertes Merkmal
Die Ausprägungen besitzen einen natürlichen Nullpunkt. Differenzen und Quotienten der Ausprägungen sind sinnvoll berechenbar.
Beispiele: Flächen, Gewichte, Temperatur in Kelvin, Entfernungen, Alter, Einkommen

Abhängigkeit der einsetzbaren statistischen Operationen vom Skalenniveau

Das Skalenniveau bestimmt, welche statistischen Operationen mit einem entsprechend skalierten Merkmal zulässig sind. Dabei können Operationen, die auf einem Skalenniveau zulässig sind, immer auch auf allen höheren Skalenniveaus durchgeführt werden

Skalenniveau in aufsteigender Reihenfolge	statistische Operationen
Nominalskala	Häufigkeit, Modus
Ordinalskala	Median, Quantile
Intervallskala	Quartilsabstand, arithmetisches Mittel, Spannweite Varianz, Standardabweichung
Verhältnisskala	geometrische Mittel

3.1.3 Häufigkeit

Begriff	Erläuterung
absolute Häufigkeit	Anzahl, wie oft eine Merkmalsausprägung vorkommt n_i
Relative Häufigkeit	Anteil, wie oft eine Merkmalsausprägung vorkommt $h_i = \frac{n_i}{n}$
Urliste	Liste der Merkmalsausprägungen in der Reihenfolge der Erhebung
Häufigkeitstabelle	Tabelle mit Merkmalsausprägung, absoluter Häufigkeit, relativer Häufigkeit (oft in %)

Häufigkeitsregeln

- Häufigkeiten sind immer ≥ 0 , d.h. $n_i, h_i \geq 0$
- Die Summe der absoluten Häufigkeiten aller Merkmalsausprägungen ergibt den Stichprobenumfang: $\sum_{i=1}^k n_i = n$
- Die Summe der relativen Häufigkeiten aller Merkmalsausprägungen ergibt 1, d.h. $\sum_{i=1}^k h_i = 1$

Beispiel

Bei der Erhebung der Länge in mm von 20 Schrauben aus der Produktion ergab sich folgende Urliste:

138 164 150 146 144 146 148 148 146 158 140 148 136 148 152 144 150 138 140 152

Hieraus ergibt sich folgende Häufigkeitstabelle:

Merkmalsausprägung	absolute Häufigkeit	relative Häufigkeit	in Prozent
136	1	$\frac{1}{20}$	5%
138	2	$\frac{2}{20}$	10%
140	2	$\frac{2}{20}$	10%
144	2	$\frac{2}{20}$	10%
146	3	$\frac{3}{20}$	15%
148	4	$\frac{4}{20}$	20%
150	2	$\frac{2}{20}$	10%
152	2	$\frac{2}{20}$	10%
158	1	$\frac{1}{20}$	5%
164	1	$\frac{1}{20}$	5%
SUMME	20	1	100%

3.1.4 Klassierung der Daten

Treten sehr viele verschiedene Merkmalsausprägungen auf (zB bei der Ermittlung der Tageshöchsttemperaturen eines Jahres), teilt man diese oft der Übersichtlichkeit halber in Klassen, bzw. Intervalle ein.

Dies bedeutet natürlich einen Informationsverlust, da nach der Klassenbildung keine Aussage mehr über die Verteilung der Werte innerhalb einer Klasse vorliegen.

Vorgehen:

- Bei der Klassenbildung zerlegt man das Gesamtintervall, in dem alle möglichen Merkmalsausprägungen liegen, in Teilintervalle
- die Teilintervalle überdecken den Bereich der Merkmalsausprägungen komplett
- Die Anzahl der Teilintervalle sollte nicht größer als $\sqrt{\text{Anzahl der Merkmalsausprägungen}}$ sein
- oft wählt man gleich breite, links abgeschlossene, rechts offene Teilintervalle

Beispiel

Jahresgehälter (in Euro) von 25 Mitarbeiter eines Unternehmens

Urliste: 9.000, 18.000, 18.000, 19.000, 20.000, 20.000, 20.000, 21.000, 23.000, 25.000, 25.000, 27.000, 28.000, 28.000, 29.000, 30.000, 31.000, 32.000, 32.000, 34.000, 36.000, 37.000, 38.000, 40.000, 41.000

Klassierte Daten:

Klasse	Intervall	Anzahl
Klasse 1	[0, 10.000)	1
Klasse 2	[10.000, 20.000)	3
Klasse 3	[20.000, 30.000)	11
Klasse 4	[30.000, 40.000)	8
Klasse 5	[40.000, 50.000)	2

3.2 Graphische Darstellung von Daten

3.2.1 Diagramme zur Häufigkeits-Darstellung

Am Beispiel der gemessenen Schraubenlänge

Balkendiagramm

Für jede Merkmalsausprägung (bzw. jede Klasse) wird ein Balken einheitlicher Dicke gezeichnet. Die Länge des Balkens entspricht der absoluten bzw. relativen Häufigkeit.

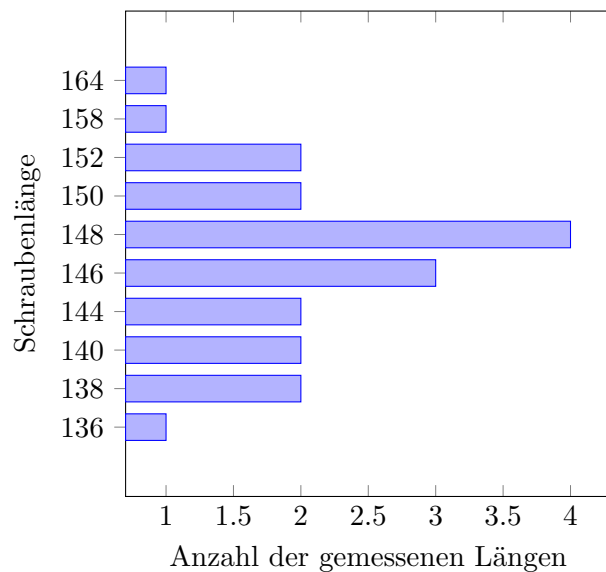


Abbildung 3.1: Balkendiagramm

Säulendiagramm

Für jede Merkmalsausprägung wird eine Säule einheitlicher Breite gezeichnet. Die Länge der Säule entspricht der absoluten bzw. relativen Häufigkeit.

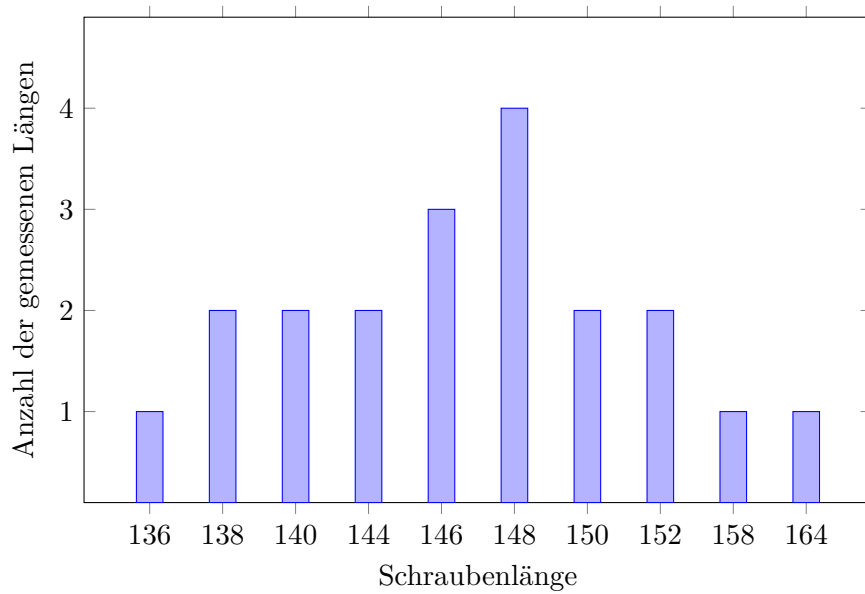


Abbildung 3.2: Säulendiagramm

Stabdiagramm

Für jede Merkmalsausprägung wird eine senkrechte Linie gezeichnet. Die Länge der Linie entspricht der absoluten bzw. relativen Häufigkeit.

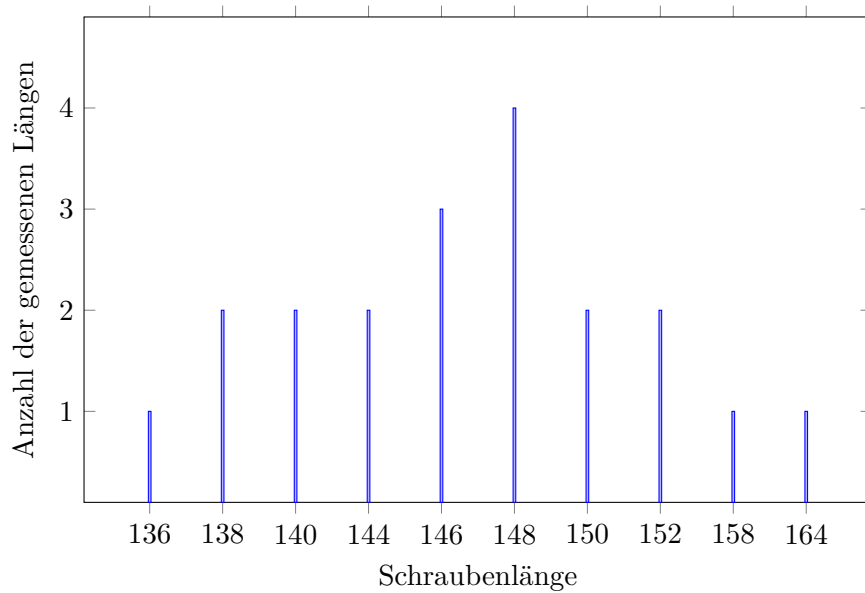


Abbildung 3.3: Stabdiagramm

Liniendiagramm

Ein Liniendiagramm verbindet die einzelnen Häufigkeitswerte der Merkmalsausprägungen.

Ein Liniendiagramm eignet sich nur für Merkmale, für die eine Interpretation der Verbindungslinie sinnvoll ist.

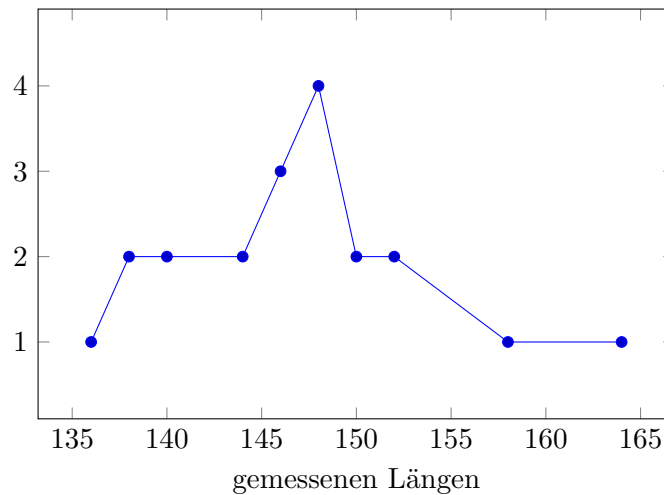


Abbildung 3.4: Liniendiagramm

Kreisdiagramm

Stellt die Teile eines Ganzen dar. Die Dicke jedes Stückes repräsentiert die absolute bzw. relative Häufigkeit der zugehörigen Merkmalsausprägung

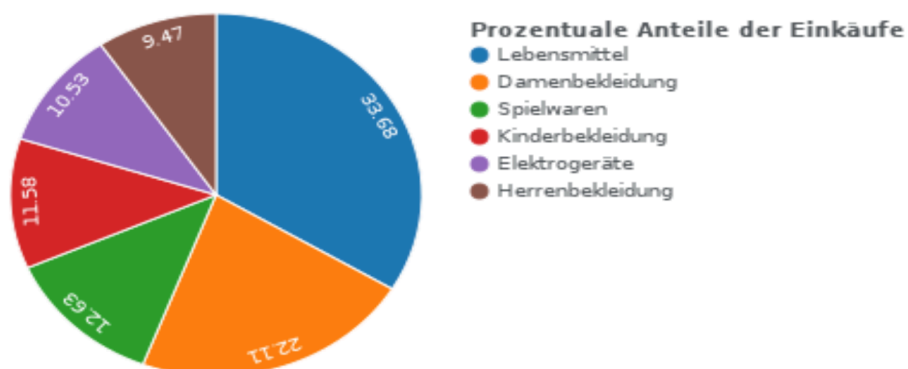


Abbildung 3.5: <https://de.wikipedia.org/wiki/Kreisdiagramm>

Den Winkel des Kreissektors zur Merkmalsausprägung mit der Häufigkeit h_i berechnet

man durch

$$\alpha_i = 360^\circ \cdot h_i$$

Histogramm

Die Häufigkeitsverteilung von quantitativen Merkmalen, kann durch Histogramme grafisch dargestellt werden.

Der Wertebereich der erfassten Daten wird in Klassen eingeteilt. Dabei können die Klassen gleich breit sein oder unterschiedliche Breite aufweisen. Die Intervalle der Klassen grenzen aneinander. Die beiden Klassen am linken und rechten Rand müssen geschlossen sein.

Die relative Häufigkeit jeder Klasse wird mithilfe eines Rechtecks dargestellt.

Der Flächeninhalt des Rechtecks ist proportional zur relativen Häufigkeit der jeweiligen Klasse.

Konstruktion eines Histogramms

- Faustregel für die Anzahl der Klassen ist \sqrt{n} , wobei n die Anzahl Merkmalsausprägungen der erhobenen Daten ist.
- Die Intervallgrenzen der gewählten Klassen werden auf der x-Achse (Abszisse) auf einer Skala eingetragen.
- Bei Klassen unterschiedlicher Breite wird keine y-Achse (Ordinate) angegeben, da die Höhe des Rechtecks nicht der relativen Häufigkeit entspricht. Die relative Häufigkeit wird beim jeweiligen Rechteck angegeben.
- Bei Klassen gleicher Breite kann eine y-Achse zum Abtragen der relativen Häufigkeiten verwendet werden, da die Höhe der Rechtecke der jeweiligen relativen Häufigkeit entspricht.
- Die Rechtecke grenzen entsprechend den Intervallen (Klassen) ohne Abstand aneinander. Im Gegensatz dazu werden Säulendiagramme meist mit Abstand gezeichnet.
- Die Höhe eines Rechtecks ergibt sich aus dem Quotienten von relativer Häufigkeit und Klassenbreite.

Klasse	Intervall	Anzahl	Rechtecksbreite	Rechteckshöhe
Klasse 1	[130, 145)	7	15	$\frac{7}{15} = 0,47$
Klasse 2	[145, 149)	7	4	$\frac{7}{4} = 1,75$
Klasse 3	[149, 165)	6	16	$\frac{6}{16} = 0,375$

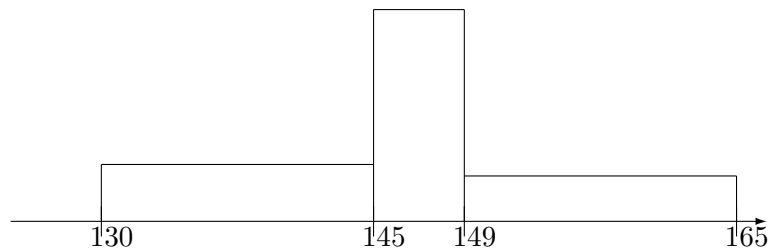


Abbildung 3.6: Histogramm

3.2.2 Empirische Verteilungsfunktion

Verteilungsfunktionen $F(x)$ geben Antworten auf die Frage:

„Wie groß ist der Anteil der Merkmalsträger, die höchstens die Merkmalsausprägung x aufweisen?“

Definition

Sind x_1, x_2, \dots, x_n Ausprägungen eines Merkmals. Dann heißt

$$F_n(x) = \sum_{i=1}^k h_i(x_i) \text{ mit } x_1, \dots, x_k \leq x$$

empirische Verteilungsfunktion.

Eigenschaften

- $F : \mathbb{R} \rightarrow [0, 1]$ monoton wachsend
- $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$

Beispiel

Die Verteilungsfunktion der Schraubenlängen aus dem früheren Beispiel lautet:

$$F_{10}(x) = \begin{cases} 0 & \text{für } x < 136 \\ 0,05 & \text{für } 136 \leq x < 138 \\ 0,15 & \text{für } 138 \leq x < 140 \\ 0,25 & \text{für } 140 \leq x < 144 \\ 0,35 & \text{für } 144 \leq x < 146 \\ 0,5 & \text{für } 146 \leq x < 148 \\ 0,7 & \text{für } 148 \leq x < 150 \\ 0,8 & \text{für } 150 \leq x < 152 \\ 0,9 & \text{für } 152 \leq x < 158 \\ 0,95 & \text{für } 158 \leq x < 164 \\ 1 & \text{für } x \geq 164 \end{cases}$$

3.3 Lage- und Streumaße

3.3.1 Modalwert

Die Merkmalsausprägung, die am häufigsten auftritt heißt **Modalwert** oder **Modus** x_{mod} .
Sie ist nur für diskrete Merkmale definiert.

Beispiel

Bei der Stichprobe (1,1,1,2,5,11,12,26,72) ist der Modus 1

Spezialfälle

- Verschiedene Merkmalsausprägungen treten gleich häufig auf, dann gibt es mehrere Modi:
Bei der Stichprobe (1,1,1,2,5,5,5,11,12,26,26,26,72) ist 1,5 und 26 ein Modus
- Insbesondere gilt, falls alle Merkmalsausprägungen nur einmal auftreten, daß dann alle Merkmalsausprägungen Modalwerte sind
Bei der Stichprobe (1,2,5,11,12,26,72) sind 1,2,5,11,12,26,72 Modalwerte

3.3.2 Quantile, Quartilsabstand

Für n geordnete Merkmalsausprägungen

$$x_1, x_2, x_3, \dots, x_n$$

ist das **p-Quantil**

$$Q_p = \begin{cases} x_{[np]} & \text{für } np \notin \mathbb{N} \\ \frac{x_{np} + x_{np+1}}{2} & \text{für } np \in \mathbb{N} \end{cases}$$

Mit dem p-Quantil will man den Wert ermitteln, für den $100p\%$ der Daten links bzw. $100(1-p)\%$ der Daten rechts von diesem Wert liegen.

Unter der Ceiling-Funktion $[x]$ versteht man die kleinste natürliche Zahl n mit $x \leq n$

Das p-Quantil heißt

Median für $p=0,5$
erstes Quartil für $p=0,25$
drittes Quartil für $p=0,75$

Der **Quartilsabstand** ist $Q_{0,75} - Q_{0,25}$

Beispiele

- Berechnen Sie das 0,1-Quantil für das Alter (in Jahren) einiger Statistik-Studenten
22, 22, 23, 23, 23, 24, 24, 25, 26
Da $9 \cdot 0,1 = 0,9 \notin \mathbb{N}$ ist $Q_{0,1} = x_1 = 22$
- Die p-Quantile sind als statistische Operationen bei ordinalskalierten Merkmalen angegeben.
Bestimmen Sie das 0,5-Quantil und das 0,3-Quantil folgender Merkmalsausprägungen:
unzufrieden - eher unzufrieden - eher zufrieden - zufrieden

3.3.3 Box-Plot

Der Box-Plot ist ein Diagramm, das zur schnellen und übersichtlichen graphischen Darstellung der Verteilung eines Merkmals verwendet wird. Er fasst verschiedene Lagemaße in einer Darstellung zusammen. Der Box-Plot vermittelt einen schnellen Eindruck darüber, in welchem Bereich die Daten liegen und wie sie sich über diesen Bereich verteilen.

Eine weitere wichtige Anwendung des Boxplot ist die Erkennung von Ausreißern im Datensatz, was auf Fehler im Datensatz hinweisen kann.

Ausreißer

Ausreißer in den Daten sind Werte, die sich erheblich von den Übrigen unterscheiden. Ausreißer treten sehr oft in statistischen Daten auf und es ist häufig unklar, wodurch sie zustande gekommen sind. Oft werden Messfehler dafür verantwortlich gemacht und die Ausreißer werden aus den Daten gelöscht. Liegt tatsächlich ein Messfehler vor, ist dieses Vorgehen natürlich sinnvoll, man muss dabei allerdings vorsichtig sein, damit nicht relevante Informationen unterschlagen werden.

Ein Beispiel aus der Wissenschaft für einen leichtfertigen Umgang mit Ausreißern ist die Entdeckung des Ozonlochs über der Antarktis. Dieses wurde im Jahr 1985 von britischen Forschern nachgewiesen. Sie zeigten, dass der Ozongehalt über der Antarktis in diesem Jahr zehn Prozent unter dem normalen Level lag.

Daraufhin kam die Frage auf, wieso der Satellit Nimbus 7, der entsprechende Messinstrumente an Bord hatte und lange schon den Ozongehalt über der Antarktis maß diese Entwicklung nicht bereits früher gemeldet hatte. Es stellte sich heraus, dass der Satellit die geringer werdende Ozonkonzentration schon seit den 70er-Jahren gemessen hatte. Die ungewöhnlich niedrigen Werte hatte das Computersystem des Satelliten in all den Jahren immer fälschlicherweise als Ausreißer klassifiziert und nicht an die Erde weitergeleitet, da das Übermitteln der Daten teuer war und man deshalb keine fehlerhaften Werte umsonst zur Erde senden wollte.

In diesem Fall hat also der fehlerhafte Umgang mit Ausreißern dazu geführt, dass die

Gefahren, die mit dem Ozonloch verbunden sind, erst viele Jahre später, als es möglich gewesen wäre, realisiert wurden.

Varianten des Box-Plots:

Notwendig zur Erstellung des Boxplot sind immer die Merkmalsausprägungen x_1, x_2, \dots, x_n in geordneter Darstellung

1. Box-Plot innerhalb des minimalen und maximalen Wertes

Für die Erstellung dieses Box Plots benötigt man folgende Informationen:

Kennwert	Lage im Box-Plot
Minimum x_1	Strich
Erstes Quartil $Q_{0,25}$	Strich am Beginn der Box
Median $Q_{0,5}$	Strich innerhalb der Box
Drittes Quartil $Q_{0,75}$	Strich am Ende der Box
Maximum x_n	Strich
Quartilsabstand	entspricht der Ausdehnung der Box

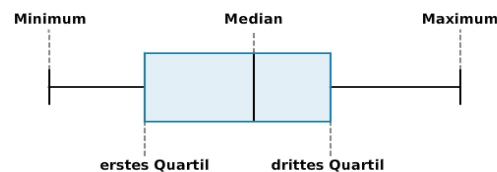


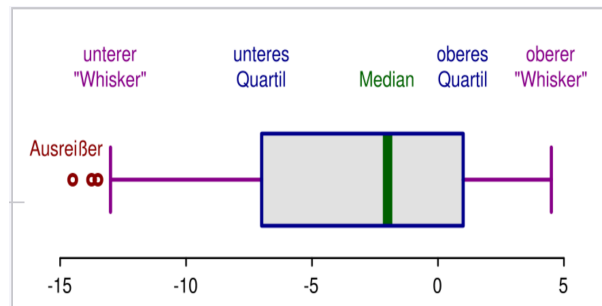
Abbildung 3.7: Box Plot

2. Box-Plot mit Ausreißern

Speziell dann, wenn man bei einem vorliegenden Datensatz sich für Ausreißer interessiert, werden die zwei Linien (Whisker) nicht bis zum Minimumwert, bzw. Maximumwert geführt, sondern man beschränkt deren jeweilige Länge z.B. auf maximal das 1,5 fache des Interquartilsabstands. Dabei endet der Whisker nicht genau nach dieser Länge, sondern bei dem letzten Datenwert, der noch innerhalb dieser Grenze liegt. Die Daten ausserhalb der Linien werden gemäß diesem Vorgehen als Ausreißer gekennzeichnet. Oft werden sie besonders (Kreis, Sternchen) im Plot markiert.

Für die Erstellung dieses Box Plots benötigt man folgende Informationen:

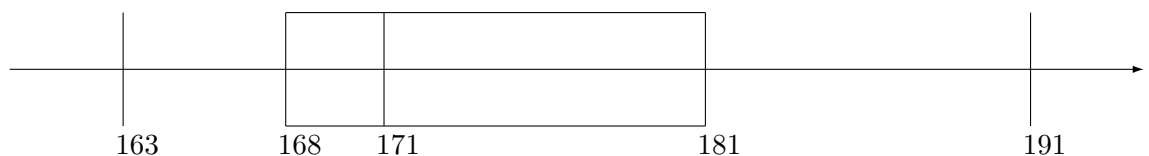
Kennwert	Lage im Box-Plot
Erstes Quartil $Q_{0,25}$	Strich am Beginn der Box
Median $Q_{0,5}$	Strich innerhalb der Box
Drittes Quartil $Q_{0,75}$	Strich am Ende der Box
Quartilsabstand	entspricht der Ausdehnung der Box
Oberer Wert innerhalb des 1,5fachen Quartilsabstandes	oberer Whisker
Unterer Wert innerhalb des 1,5fachen Quartilsabstandes	unterer Whisker
Werte ausserhalb des Bereichs zwischen Whisker	Ausreißer

Abbildung 3.8: <https://de.wikipedia.org/wiki/Box-Plot>**Beispiel**

Zeichnen Sie einen Box-Plot für folgende Größenangaben in cm einiger Studenten

163, 165, 165, 168, 170, 170, 171, 172, 174, 181, 181, 182, 191

$x_{\min} = 163$; $Q_{0,25} = x_4 = 168$; $Q_{0,5} = x_7 = 171$; $Q_{0,75} = x_{10} = 181$; $x_{\max} = 191$

**Informationen aus dem Box Plot**

- In jedem der vier Intervallen $[x_1, Q_{0,25}]$, $[Q_{0,25}, Q_{0,5}]$, $[Q_{0,5}, Q_{0,75}]$, $[Q_{0,75}, x_n]$, liegen jeweils 25% der Daten
- Die Box entspricht dem Bereich, in dem 50% der Daten liegen
- Dort, wo die Intervalle kurz sind häufen sich die Merkmalsausprägungen. Dort, wo die Intervalle lang sind, liegen die Merkmalsausprägungen weiter auseinander
- Die Lage des Medians innerhalb der Box, gibt einen Eindruck der Schiefe der Daten. Ist der Median in der linken Hälfte der Box, so sagt man, daß die Verteilung der

Daten rechtsschief ist, und umgekehrt

- Der Abstand $x_n - x_1$ ist die Spannweite, d.h. die Länge des Bereichs in dem alle Datensätze liegen

3.3.4 Arithmetisches Mittel

- Hat man n Messwerte x_1, x_2, \dots, x_n eines Merkmals, dann heißt

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

das **arithmetische Mittel**

- Hat man k verschiedene Merkmalsausprägungen x_1, \dots, x_k mit ihren relativen Häufigkeit h_1, \dots, h_k , so berechnet sich das arithmetische Mittel

$$\bar{x} = x_1 h_1 + x_2 h_2 + \dots + x_k h_k$$

- Liegt eine Klassierung der Merkmale mit Intervallen $[a_1, a_2), [a_2, a_3), \dots, [a_l, a_{l+1})$ und relativen Häufigkeiten h_1, \dots, h_l vor, so berechnet man das arithmetische Mittel durch

$$\bar{x} = h_1 \cdot \frac{a_1 + a_2}{2} + h_2 \cdot \frac{a_2 + a_3}{2} + \dots + h_l \cdot \frac{a_l + a_{l+1}}{2}$$

Beispiele

1. Aus der Produktion von Karosserieteilen werden 10 Teile gewogen:

Teil	1	2	3	4	5	6	7	8	9	10
Gewicht in N	17,2	17,4	17,2	17,2	17,3	17,3	17,4	17,3	17,6	17,2

Das arithmetische Mittel ist

$$\bar{x} = \frac{1}{10} (17,2 + 17,4 + 17,2 + 17,2 + 17,3 + 17,3 + 17,4 + 17,3 + 17,6 + 17,2) = 17,31$$

2. Die Gewichte der Karosserieteile liegen in folgender Häufigkeitstabelle vor. Berechnen Sie das arithmetische Mittel.

Gewicht N	17,2	17,3	17,4	17,6
relative Häufigkeit	0,4	0,3	0,2	0,1

Das arithmetische Mittel ist

$$\bar{x} = 0,4 \cdot 17,2 + 0,3 \cdot 17,3 + 0,2 \cdot 17,4 + 0,1 \cdot 17,6 = 17,31$$

3. Das Einkommen der Arbeitnehmer eines Betriebs liegt in folgender Häufigkeitstabelle vor. Berechnen Sie das mittlere Einkommen (arithmetisches Mittel)

Intervall in 1000 €	[0,1)	[1, 1,5)	[1,5 , 2)	[2, 2,5)	[2,5 , 3)	[3, 5]
relative Häufigkeit	5%	15%	35 %	25%	10%	10%

Das arithmetische Mittel ist

$$\bar{x} = 0,05 \cdot 0,5 + 0,15 \cdot 1,25 + 0,35 \cdot 1,75 + 0,25 \cdot 2,25 + 0,1 \cdot 2,75 + 0,1 \cdot 4 = 2,0625$$

Vergleich Median - arithmetisches Mittel

In einer Gruppe von zehn Personen haben alle Personen ein Monatseinkommen in unterschiedlicher Höhe. Eine Person erhält 1.000.000 Euro, die übrigen neun bekommen 1.000 Euro, 2.000 Euro, 3.000 Euro usw. bis 9.000 Euro.

Das arithmetische Mittel, der "Durchschnitt" das Monatseinkommen jeder der zehn Personen bei gleichmäßiger Aufteilung der Summe aller Einkommen auf sie, beträgt 104.500 Euro. Allerdings verdient nur eine der zehn Personen mehr als diesen Betrag, die neun anderen deutlich weniger.

Der Median dagegen ist 5.500 Euro. Fünf Personen verdienen mehr als das, fünf Personen weniger.

3.3.5 Spannweite

Die **Spannweite** R ist das einfachste Streuungsmaß, es ist der Bereich in dem die Merkmalsausprägungen liegen: $R = x_{\max} - x_{\min}$

3.3.6 Empirische Varianz und empirische Standardabweichung

Die empirische Varianz ist ein Maß für die Streuung der Merkmalsausprägungen um den arithmetischen Mittelwert. Sie wird definiert als Summe der Abweichungsquadrate geteilt durch $n-1$ (manchmal auch geteilt durch die Anzahl der Messwerte n)

- Hat man n Messwerte x_1, x_2, \dots, x_n eines Merkmals, so heißt

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

die **empirische Varianz**

- Hat man k verschiedene Merkmalsausprägungen x_1, \dots, x_k mit ihren relativen Häufigkeit h_1, \dots, h_k , so berechnet sich die empirische Varianz

$$s^2 = \frac{n}{n-1} \sum_{i=1}^k (x_i - \bar{x})^2 \cdot h_i$$

- Liegt eine Klassierung der Merkmale mit Intervallen $[a_1, a_2), [a_2, a_3), \dots, [a_l, a_{l+1})$ und relativen Häufigkeiten h_1, \dots, h_l vor, so berechnet man die empirische Varianz durch

$$s^2 = \frac{n}{n-1} \sum_{i=1}^l \left(\frac{a_i + a_{i+1}}{2} - \bar{x} \right)^2 \cdot h_i$$

Als **empirische Standardabweichung**, wird die positive Wurzel der empirischen Varianz bezeichnet

$$s = \sqrt{s^2}$$

Berechnen der empirischen Varianz und Standardabweichung der vorhergehenden Beispiele

1. Aus der Produktion von Karosserieteilen werden 10 Teile gewogen:

Teil	1	2	3	4	5	6	7	8	9	10
Gewicht in N	17,2	17,4	17,2	17,2	17,3	17,3	17,4	17,3	17,6	17,2

Das arithmetische Mittel $\bar{x} = 17,31$

Die empirische Varianz $s^2 = 0,0165$

Die empirische Standardabweichung $s = 0,128$

2. Die Gewichte der Karosserieteile liegen in folgender Häufigkeitstabelle vor.

Gewicht in N	17,2	17,3	17,4	17,6
relative Häufigkeit	0,4	0,3	0,2	0,1

Das arithmetische Mittel ist $\bar{x} = 17,31$

Die empirische Varianz

$$s^2 = \frac{10}{9} \left(0,4(17,2 - 17,31)^2 + 0,3 \cdot (17,3 - 17,31)^2 + 0,2 \cdot (17,4 - 17,31)^2 + 0,1 \cdot (17,6 - 17,31)^2 \right) \\ = \frac{10}{9} \cdot 0,0149 = 0,0165$$

Die empirische Standardabweichung $s = 0,128$

3. Das Einkommen der Arbeitnehmer eines Betriebs liegt in folgender Häufigkeitstabelle vor. Berechnen Sie das mittlere Einkommen (arithmetisches Mittel) und die empirische Standardabweichung

Intervall in 1000 €	[0,1)	[1, 1,5)	[1,5 , 2)	[2, 2,5)	[2,5 , 3)	[3, 5]
relative Häufigkeit	5%	15%	35 %	25%	10%	10%

Das arithmetische Mittel $\bar{x} = 2,0625$

Die empirische Varianz

$$s^2 = \frac{6}{5} \left(0,05(0,5 - \bar{x})^2 + 0,15(1,25 - \bar{x})^2 + 0,35(1,75 - \bar{x})^2 + 0,25(2,25 - \bar{x})^2 + 0,1(2,75 - \bar{x})^2 + 0,1(4 - \bar{x})^2 \right) \\ = \frac{6}{5} \cdot 0,68671875 = 0,824$$

Die empirische Standardabweichung $s = 0,908$

Zusammenhang arithmetisches Mittel und Varianz

arithmetisches Mittel	Varianz
Die Summe der Abweichungen aller Messwerte von ihrem arithmetischem Mittel ist Null:	Die Summe der Abweichungsquadrate ist für das arithmetische Mittel ein Minimum:
$\sum_{i=1}^n (x_i - \bar{x}) = 0$	$\sum_{i=1}^n (x_i - \bar{x})^2 = \min!$

3.4 Korrelationsanalyse

In vielen statistischen Untersuchungen werden nicht nur ein Merkmal, sondern gleichzeitig mehrere Merkmale erfasst. Z.B. beim Mikrozensus, werden jedem Befragten ein umfangreicher Fragebogen vorgelegt. Bei der gleichzeitigen Erfassung mehrerer Merkmale interessiert insbesondere:

- Wie stark ist die Abhängigkeit der einzelnen Merkmale?
- Wie lässt sich der Zusammenhang mathematisch beschreiben?

Erhebt man nicht nur ein Merkmal, sondern mehrere Merkmale, dann spricht man von einem mehrdimensionalen Merkmal. Und die Untersuchung des mathematischen Zusammenhangs bezeichnet man als Korrelationsanalyse

ABER: Spurious Correlations.pdf

3.4.1 Zweidimensionale Häufigkeitsverteilungen

Betrachtet man die Merkmale X und Y, die an denselben n Merkmalsträgern erhoben werden, dann ergeben sich als Ergebnis der Erhebung bzw. Messung jeweils Beobachtungswertepaare für jeden Merkmalsträger $i=1..n$

$$(x_i, y_i)$$

Eine graphische Darstellungsform dieser zweidimensionalen Merkmale ist das **Streudiagramm**.

Hierzu werden alle Merkmalsausprägungen (x_i, y_i) als Punkte in einem Koordinatensystem (x-Achse mit x-Werten und y-Achse mit y-Werten) markiert

Beispiel

Student	1	2	3	4	5	6	7	8	9	10
Größe	156	167	168	172	186	192	156	167	172	153
Gewicht	58	63	74	74	95	78	49	62	80	50

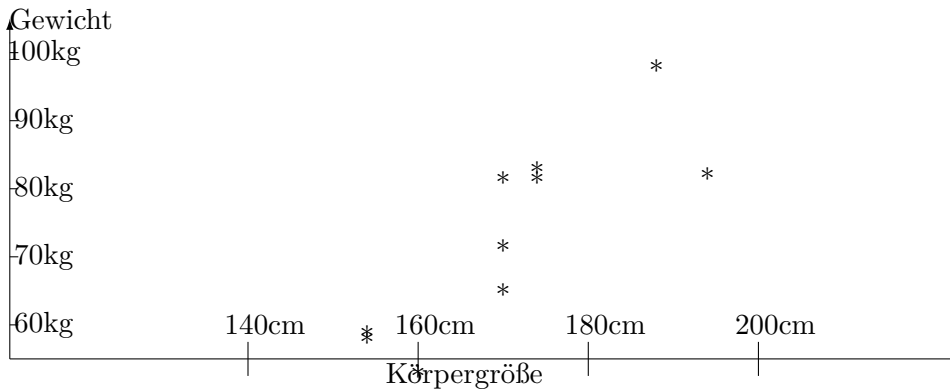


Abbildung 3.9: Streudiagramm

Definitionen

\bar{x}, \bar{y}	arithmetische Mittelwerte der Merkmalsausprägungen von X und Y
s_x^2, s_y^2	Varianz von X, bzw. Y
s_x, s_y	Standardabweichungen von X bzw. Y
s_{xy}	Kovarianz von x,y: $s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$
r_{xy}	Korrelationskoeffizient von x,y (<i>normierte Kovarianz</i>) $r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{s_{xy}}{s_x s_y}$

Bemerkungen

- Bedeutung der Kovarianz:
 - $s_{xy} > 0$ Die Merkmale haben einen positiven Zusammenhang:
hohe (niedrige) x-Werte gehen mit hohen (niedrigen) y-Werten einher.
 - $s_{xy} < 0$ Die Merkmale haben einen negativen Zusammenhang:
hohe (niedrige) x-Werte gehen mit niedrigen (hohen) y-Werten einher.
 - $s_{xy} = 0$ Es besteht kein Zusammenhang zwischen x und y
 Die Kovarianz gibt zwar die Richtung einer Beziehung zwischen zwei Merkmalen an, über die Stärke des Zusammenhangs wird aber keine Aussage getroffen. Um einen Zusammenhang vergleichbar zu machen, muss die Kovarianz normiert werden. Dies führt zum Korrelationskoeffizienten.
- Bedeutung des Korrelationskoeffizienten
 Der Korrelationskoeffizient nimmt Werte zwischen -1 und 1 an und es gilt:
 - $r_{xy} = 0$ Merkmale sind unkorreliert
 - $r_{xy} = 1$ Merkmale sind positiv korreliert
 - $r_{xy} = -1$ Merkmale sind negativ korreliert

Veranschaulichung Kovarianz

Betrachtet man die Merkmale X und Y an 3 Merkmalsträger, dann erhält man die Merkmalspaare

$$(x_1, y_1), (x_2, y_2), (x_3, y_3)$$

und die arithmetischen Mittelwerte \bar{x} und \bar{y}

Schreibt man $\vec{x} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix}$ und analog $\vec{y} = \begin{pmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ y_3 - \bar{y} \end{pmatrix}$

so kann man die Kovarianz als Skalarprodukt interpretieren:

$$s_{xy} = \frac{1}{2} \cdot ((x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + (x_3 - \bar{x})(y_3 - \bar{y}))$$

Veranschaulichung Korrelationskoeffizient

Normiert man die beiden Vektoren (auf die Länge 1), also

$$\vec{x}_e = \frac{1}{|\vec{x}|} \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix} = \frac{1}{\sqrt{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2}} \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix}$$

und analog

$$\vec{y}_e = \frac{1}{|\vec{y}|} \begin{pmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ y_3 - \bar{y} \end{pmatrix} = \frac{1}{\sqrt{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + (y_3 - \bar{y})^2}} \begin{pmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ y_3 - \bar{y} \end{pmatrix}$$

so kann man den Korrelationskoeffizient interpretieren als

Skalarprodukt der zwei auf die Länge 1 normierten Vektoren \vec{x}_e und \vec{y}_e :

$$r_{xy} = \vec{x}_e \cdot \vec{y}_e$$

Mit dieser Interpretation kann man folgern:

$r_{xy} = 0 \iff$ Vektoren stehen senkrecht, sind also unabhängig

$r_{xy} = 1 \iff$ Vektoren sind linear abhängig und

zeigen in die gleiche Richtung

$r_{xy} = -1 \iff$ Vektoren sind linear abhängig und

zeigen in die entgegengesetzte Richtung

Hinweis

Eine gute Erläuterung der Korrelation zweier Merkmale findet man unter youtube unter dem Stichwort Fernuni Hagen; Korrelation:

<https://www.youtube.com/watch?v=X7eeyRX35wM>

3.4.2 Durchführung Korrelationsanalyse

Der Zusammenhang zweier Merkmale wird in drei Schritten analysiert :

1. Zunächst benötigt man eine Vermutung des funktionalen Zusammenhangs der beiden Merkmale
Das Streudiagramm liefert dazu eine Vermutung, also zB eine Gerade, eine Parabel, eine e-Funktion,...
2. Die Funktion hat noch unbestimmte Parameter, die mit Hilfe der Methode der kleinsten Fehlerquadrate (zB Regressionsgerade) so bestimmt werden, dass der Abstand zwischen Messpunkt und Funktion minimal wird.
Genauer: Man minimiert die Summe der vertikalen Abstandsquadrate $\sum_{i=1}^n (y_i - f(x_i))^2$
Für diese Minimierungsaufgabe benötigt man Kenntnisse aus der Analysis zweidimensionaler Funktionen.

Regressionsgerade

Für den Fall, dass man begründet annehmen kann, dass zwischen den Merkmalswerten ein linearer Zusammenhang besteht, kann man zeigen, dass die Koeffizienten der Regressionsgeraden

$$y = kx + d$$

sich folgendermaßen ergeben.

$$k = r_{xy} \frac{s_y}{s_x} \text{ und } d = \bar{y} - k\bar{x}$$

3. Berechnung der Güte der linearen Regression
Der Korrelationskoeffizient gibt Auskunft über die Güte der linearen Regression.
 $r_{xy} = 0$ Merkmale sind unkorreliert
 $r_{xy} = 1$ die Punkte liegen exakt auf einer Geraden mit positiver Steigung
 $r_{xy} = -1$ die Punkte liegen exakt auf einer Geraden mit negativer Steigung

Beispiel: Korrelationsanalyse von Größe und Gewicht

Die Korrelationsanalyse für das vorhergehende Beispiel deutet auf einen linearen Zusammenhang hin.

Die Regressionsgerade ergibt sich zu $y = 98,985x - 98,886$

Der Korrelationskoeffizient ist: $r_{xy} = \dots$

Hinweis

Die 4 Beispiele von Francis Anscombe zeigen, dass alleine die Berechnung einer Regressionsgeraden und des Korrelationskoeffizienten keine vollständige Aussage einer guten linearen Näherung liefern. Eine Vorab-Analyse des Streudiagramms ist stets sinnvoll!

<https://de.wikipedia.org/wiki/Anscombe-Quartett>

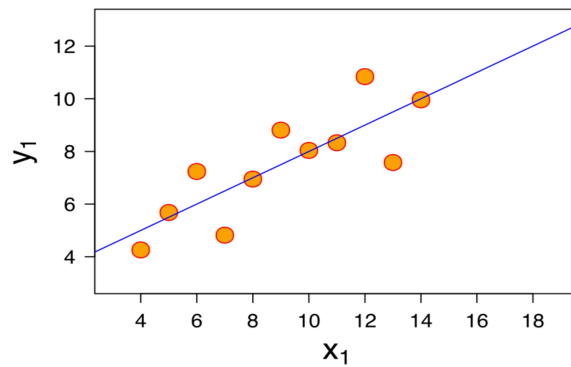


Abbildung 3.10: <https://de.wikipedia.org/wiki/Anscombe-Quartett>

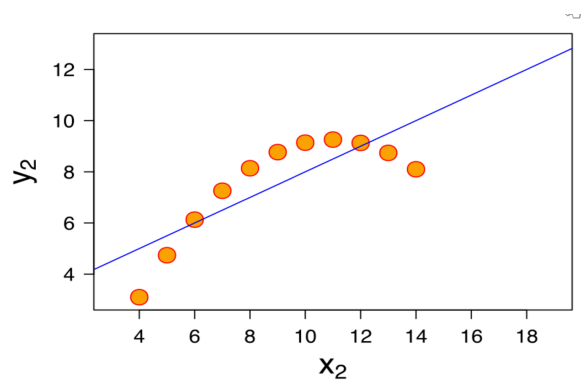
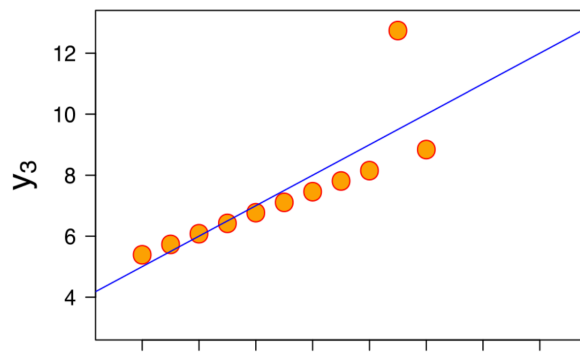
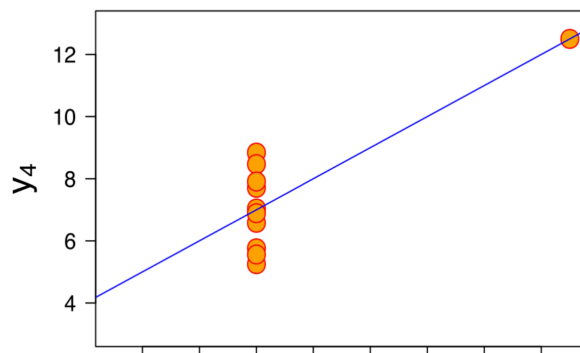


Abbildung 3.11: <https://de.wikipedia.org/wiki/Anscombe-Quartett>

Abbildung 3.12: <https://de.wikipedia.org/wiki/Anscombe-Quartett>Abbildung 3.13: <https://de.wikipedia.org/wiki/Anscombe-Quartett>

Bei allen 4 Streudiagrammen gilt:

Mittelwert von x	\bar{x}	=	9
Varianz von x	s_x^2	=	11
Mittelwert von y	\bar{y}	=	7,5
Varianz von y	s_y^2	=	4,12
Korrelationskoeffizient	$r_{x,y}$	=	0,816
Regressionsgerade	$y = 3 + 0,5x$		

Das Anscombe-Quartett weist auf die Bedeutung der graphischen Datenanalyse hin, die erfolgen soll, bevor man einen Regressionsgerade und den Korrelationskoeffizienten bestimmt.

Nicht lineare Regression

Zahlreiche Beziehungen zeigen deutlich eine nichtlineare Abhängigkeit, zB exponentielle Wachstumsfunktion

$$y = a \cdot e^{bx}$$

In vielen Fällen kann man solche nichtlinearen Funktionen durch geeignete Variablentransformation in lineare verwandeln.

Jedenfalls ist es erforderlich, dass man einen Verdacht über den funktionalen Zusammenhang der beiden Merkmale hat, zB Parabel, e-Funktion. Die dann noch unbekannten Parameter der Funktion kann man auch direkt über die Minimierung der y-Abstandsquadrate bestimmen. Diese Minimierungsaufgabe erfordert Kenntnisse der Analysis mehrerer Variablen. Es gibt diverse SW-Programme, die einem diesen Schritt abnehmen.

4

Wahrscheinlichkeitsrechnung

4.1 Grundbegriffe

Begriff	Beschreibung
Zufallsexperiment	Ein Zufallsexperiment, z.B. das Werfen einer Münze oder eines Würfels, ist ein Experiment, das sich unter gleichen äusseren Bedingungen beliebig oft wiederholen lässt Bei der Durchführung des Experimentes sind mehrere sich gegenseitig ausschliessende Ergebnisse möglich Das Ergebnis des Experimentes lässt sich nicht mit Sicherheit voraussagen, sondern ist zufallsbedingt
Ergebnis ω	Der Ausgang eines Zufallsexperiments heisst Ergebnis. Z. B. ist ein Ergebnis eines Würfelwurfs $\omega = 1$
Ergebnismenge Ω	Die Menge aller möglichen Ergebnisse heisst Ergebnismenge Z. B. Ist die Ergebnismenge des Würfelwurfs: $\Omega = \{1, 2, 3, 4, 5, 6\}$
Ereignis	Jede Teilmenge E der Ergebnismenge Ω heisst Ereignis
Elementarereignis	Ein Elementarereignis ist ein Ereignis mit einem Element

Die Ergebnismenge Ω als auch jedes Ereignis sind somit Mengen!

Arten von Ergebnismengen Ω

- endliche Ergebnismenge
Es gibt eine endliche Anzahl von Ergebnissen: Werfen einer Münze, Würfeln
- abzählbar unendliche Ergebnismenge
Es gibt keine obere Grenze der Anzahl der Ergebnisse, aber man kann die Ergebnisse abzählen
Anzahl der geschriebenen Statistik-Klausuren bis man eine 1 bekommt
Anzahl der Tippfehler in der Bachelorarbeit
- überabzählbar-unendliche Ergebnismenge
Die Elemente in der Ergebnismenge kann man nicht zählen
Messen der Länge eines Gegenstandes; des Gewichts, der Masse,...

Beispiele

- Zufallsexperiment: Wurf einer Münze
 Ergebnis $\omega = \text{Kopf, Zahl}$
 Ergebnismenge $\Omega = \{\text{Kopf, Zahl}\}$
 Ereignis $\{\}, \{\text{Kopf}\}, \{\text{Zahl}\}, \Omega$
 Elementarereignis $\{\text{Kopf}\}, \{\text{Zahl}\}$
- Zufallsexperiment: Wurf eines homogenen Würfels
 Ergebnis $\omega = 1, 2, 3, 4, 5, 6$
 Ergebnismenge $\Omega = \{1, 2, 3, 4, 5, 6\}$
 Ereignis $\{\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1, 2\}, \dots, \Omega$
 Elementarereignis $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$

4.2 Begriffe und Verknüpfung von Ereignissen

Ereignisse sind Mengen, so dass für Ereignisse die Mengenoperationen gelten.

Symbol	Bedeutung
Ω	Das sichere Ereignis
$\emptyset = \{\}$	Das unmögliche Ereignis
$A \cup B$	Vereinigung der Ereignisse A und B Entweder tritt A ein oder B oder A und B gleichzeitig
$A \cap B$	Durchschnitt der Ereignisse A und B A und B treten gleichzeitig ein
\overline{A}	Das Komplement zu A A tritt nicht ein
$A \setminus B$	Differenzmenge von A und B A tritt ein, aber B nicht
$A \cap B = \{\}$	A und B sind disjunkt; A und B können nicht gemeinsam eintreten

Rechenregeln

Für drei Ereignisse A, B und C gelten die folgenden Rechenregeln (analog zu den Regeln der Booleschen Algebra):

- $A \cup B = B \cup A$ Kommutativgesetz
- $A \cap B = B \cap A$
- $A \cup A = A$ Einsgesetz
- $A \cap A = A$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ Assoziativgesetz

- $A \cap (B \cap C) = (A \cap B) \cap C$
- $A \cup \bar{A} = \Omega$
- $A \cap \bar{A} = \{\}$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ Distributivgesetz
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\overline{A \cup B} = \bar{A} \cap \bar{B}$ Regel von de Morgan
- $\overline{A \cap B} = \bar{A} \cup \bar{B}$
- $A \setminus B = A \cap \bar{B}$
- $\bar{A} = \Omega \setminus A$
- $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$
- $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$
- $\{\} \subseteq A \subseteq \Omega$
- $A \subseteq B \iff A = A \cap B$
- $A \subseteq B \iff B = A \cup (B \cap \bar{A})$

Beispiele

- Formulieren Sie eine Frage, die beim Würfelwurf zu folgendem Ereignis führt:

Ereignis	passende Frage
$A = \emptyset$	
$A = \{1, 2\}$	
$A = \{4, 5, 6\}$	
$A = \{2, 4, 6\}$	
$A = \Omega$	

- Beim Wurf einer Münze sei
 A das Ereignis *Zahl liegt oben*,
dann ist die Komplementärmenge
 \bar{A} das Ereignis *Kopf liegt oben*
- Beim Wurf zweier Münzen ergibt sich die folgende Tabelle an möglichen Ereignissen

	Zahl	Wappen
1. Münze	A_{1Z}	A_{1W}
2. Münze	A_{2Z}	A_{2W}

Beispiel für zusammengesetzte Ereignisse sind dann

- Erste Münze zeigt Zahl oder Wappen: $A_{1Z} \cup A_{1W} = \Omega$
- Gleichzeitiges Auftreten von Zahl und Wappen bei der ersten Münze:
 $A_{1Z} \cap A_{1W} = \{\}$
- Erste Münze Zahl, zweite Münze Wappen: $A_{1Z} \cap A_{2W}$
- Ein Würfel wird einmal geworfen. Es werden dabei folgende Ereignisse definiert:
A: Die geworfene Zahl ist kleiner 4
B: Die geworfene Zahl ist ungerade
Bestimmen Sie folgende Ereignisse:
 1. $A = \{1, 2, 3\}$
 2. $B = \{1, 3, 5\}$
 3. $A \cup B = \{1, 2, 3, 5\}$
 4. $A \cap B = \{1, 3\}$
 5. $\overline{A} \cap \overline{B} = \{4, 6\}$
 6. $\overline{A \cap B} = \{2, 4, 5, 6\}$
 7. $A \cap \overline{B} = \{2\}$
 8. $\overline{A \cup B} = \{4, 6\}$
- Eine Urne enthält 3 rote und 5 schwarze Kugeln. Aus der Urne werden nacheinander drei Kugeln ohne zurücklegen entnommen. Folgende Ereignisse werden definiert:
A: Die ersten zwei gezogenen Kugeln haben unterschiedliche Farbe
B: Die zuerst und die zuletzt gezogene Kugel haben dieselbe Farbe
Geben Sie die folgenden Ereignisse an:
$$\Omega = \{RRR, RRS, RSS, RSR, SRR, SRS, SSR, SSS\}$$

$$A = \{RSS, RSR, SRR, SRS\}$$

$$B = \{RRR, RSR, SSS, SRS\}$$

$$A \cap B = \{RSR, SRS\}$$

$$\overline{A}, \overline{B}, A \cup B$$
- In einer Lostrommel befinden sich 15 Lose, davon sind 10 Lose Nieten. Aus der Lostrommel werden nacheinander 2 Lose gezogen. Folgende Ereignisse werden definiert:
A: Es werden nur Nieten gezogen
B: Genau ein Gewinnlos wird gezogen
C: Das zuletzt gezogene Los ist eine Niete
Bestimmen Sie die folgenden Ereignisse:
$$\overline{A \cup B}, B \cap \overline{C}$$

4.3 Wahrscheinlichkeit

Die Wahrscheinlichkeit ist eine Abbildung P von den Ereignissen in das Intervall $[0,1]$.

Definition Wahrscheinlichkeit

Eine reelle Funktion P , definiert auf einer Ereignismenge heißt Wahrscheinlichkeit, wenn sie die folgenden drei (Kolmogorow-)Axiome erfüllt:

1. Für jedes Ereignis A ist die Wahrscheinlichkeit eine reelle Zahl zwischen 0 und 1: $0 \leq P(A) \leq 1$
2. Das sichere Ereignis hat die Wahrscheinlichkeit 1: $P(\Omega) = 1$
3. Die Wahrscheinlichkeit einer Vereinigung abzählbar vieler inkompatibler Ereignisse entspricht der Summe der Wahrscheinlichkeiten der einzelnen Ereignisse
Inkompatible Ereignisse sind disjunkte Mengen
 $P(A_1 \cup A_2 \dots) = \sum_{i=1}^{\infty} P(A_i)$

Folgerungen

1. $P(\overline{A}) = 1 - P(A)$
Beweis: A und \overline{A} sind disjunkte Mengen mit $A \cup \overline{A} = \Omega$.
Nach dem 3. Axiom folgt $P(A) + P(\overline{A}) = P(A \cup \overline{A}) = P(\Omega) = 1$ und somit: $P(\overline{A}) = 1 - P(A)$
2. $P(\{\}) = 0$
Beweis: Da $\{\} = \overline{\Omega}$ folgt $P(\{\}) = 1 - P(\Omega) = 1 - 1 = 0$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Beweis: Es ist $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$ und diese Mengen sind disjunkt. Somit ist nach dem 3. Axiom: $P(A \cup B) = P(A \setminus B) + P(B \setminus A) + P(A \cap B)$
 A und B lassen sich jeweils durch die disjunkten Mengen darstellen:
 $A = (A \setminus B) \cup (A \cap B)$ sowie $B = (B \setminus A) \cup (A \cap B)$, so dass gilt:
 $P(A) = P(A \setminus B) + P(A \cap B)$ und $P(B) = P(B \setminus A) + P(A \cap B)$
Daraus ergibt sich insgesamt:
$$\begin{aligned} P(A \cup B) &= P(A \setminus B) + P(B \setminus A) + P(A \cap B) \\ &= P(A \setminus B) + P(A \cap B) + P(B \setminus A) + P(A \cap B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$
4. Sind A und B zwei disjunkte Ereignisse, dann gilt $P(A \cup B) = P(A) + P(B)$
5. Aus $A \subseteq B$ folgt $P(A) \leq P(B)$

4.3.1 Laplace-Experiment

Zufallsexperimente, bei denen nur endlich viele Elementarereignisse möglich und diese gleichwahrscheinlich sind, (wie z.B. beim Werfen einer idealen Münze) nennt man Laplace-Experiment.

Beim Laplace-Experiment mit der Ergebnismenge $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ besitzen alle Elementarereignisse $\{\omega_i\}$ die gleiche Wahrscheinlichkeit

$$P(\{\omega_i\}) = \frac{1}{n} \text{ für } i=1, \dots, n$$

Die Wahrscheinlichkeit eines Ereignisses A ist dann gegeben durch

$$\begin{aligned} P(A) &= \frac{\text{Anzahl der Elementarereignisse in A}}{\text{Anzahl der Elementarereignisse von } \Omega} \\ &= \frac{\text{für A günstige Elementarereignisse}}{\text{alle möglichen Elementarereignisse}} \end{aligned}$$

Beispiele zu Laplace-Experimenten:

1. Für das Ereignis A mit einem idealen Würfel eine 2 zu würfeln gilt die Wahrscheinlichkeit $P(A) = \frac{1}{6}$

2. Wie groß ist die Wahrscheinlichkeit beim einmaligen Würfeln mit einem homogenen Würfel eine gerade Zahl größer als 2 zu würfeln?

$$A = \{4, 6\} \implies P(A) = \frac{1}{3}$$

3. Ein homogener Würfel wird einmal geworfen und folgende Ereignisse festgelegt:

A: Die Augenzahl ist kleiner als 4

B: Die Augenzahl ist eine ungerade Zahl

$$C = \{4, 5\}$$

Bestimmen Sie folgende Wahrscheinlichkeiten:

$$P(A \cap B) = P(\{1, 3\}) = \frac{1}{3}$$

$$P(A \cup B) = P(\{1, 2, 3, 5\}) = \frac{2}{3}$$

$$P(\bar{A} \cap B) = P(\{5\}) = \frac{1}{6}$$

$$P(A \cap C) = P(\{\}) = 0$$

4.3.2 Beispiele zur Berechnung von Wahrscheinlichkeiten

1. Ein Glücksrad hat sechs von 1 bis 6 durchnummerierte Sektoren, teils unterschiedlicher Größe. Folgende zwei Ereignisse seien definiert:

A: Die Nummer ist gerade, d.h. $A = \{2, 4, 6\}$

B: Die Nummer ist ungerade, d.h. $B = \{1, 3, 5\}$

Die Wahrscheinlichkeiten der Elementarereignisse sind:

$$P(1) = P(3) = \frac{1}{8}, P(2) = \frac{1}{4}, P(4) = P(5) = \frac{1}{6} \implies P(6) = \frac{1}{6}$$

Berechnen Sie die Wahrscheinlichkeit für das Eintreten des Ereignisses A, bzw. B

$$P(A) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{4} + \frac{1}{6} + \frac{1}{6} = \frac{7}{12}$$

$$P(B) = \frac{5}{12}$$

2. Ein homogener Würfel wird zweimal geworfen. Wie groß ist dabei die Wahrscheinlichkeit, mindestens einmal die Augenzahl "6" zu erzielen?

Für die beiden Ereignisse:

A: "Augenzahl 6 beim 1. Wurf"

B: "Augenzahl 6 beim 2. Wurf"

gilt: $P(A)=P(B)=\frac{1}{6}$

Das interessierende Ereignis ist $A \cup B$, d.h. eine 6 im 1. oder im 2. oder in beiden Würfeln.

Mit obiger Formel benötigt man noch $P(A \cap B) = \frac{1}{36}$

Damit ist

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$$

3. In einer Urne befinden sich 2 schwarze und 3 rote Kugeln. Es wird einmal gezogen.

- Mit welcher Wahrscheinlichkeit ist die gezogene Kugel schwarz?
- Wie viele schwarze Kugeln müssen mindestens in der Urne liegen, so dass die Wahrscheinlichkeit, eine schwarze Kugel zu ziehen, größer als 0,7 ist?

4.4 Bedingte Wahrscheinlichkeiten

Unter einer bedingten Wahrscheinlichkeit versteht man die Wahrscheinlichkeit für das Eintreten eines Ereignisses A unter der Voraussetzung, dass das Ereignis B bereits eingetreten ist. Natürlich muss B eintreten können, es darf also nicht das unmögliche Ereignis sein.

Definition: Bedingte Wahrscheinlichkeit

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ mit } P(B) \neq 0$$

ist die Wahrscheinlichkeit von A unter der Voraussetzung B

Beispiele

- 27% aller Studierenden in Österreich im WiSe 2022/23 studieren an FHs
S... das Ereignis zu studieren im ..
FH:: Das Ereignis an einer FH zu studieren
 $P(\quad | \quad) = 0,27$
- 75% aller Verkehrsunfälle werden von Männern verursacht
U... das Ereignis, dass ein Unfall passiert
M... das Ereignis, dass ein Mann am Steuer sitzt
 $P(\quad | \quad) = 0,75$

- Von den Studenten, die die Übungsaufgaben nicht bearbeiten haben, bestehen i.a. nur 50% die Statistiklausur
 Ü... das Ereignis die Übungsaufgaben zu bearbeiten
 B... das Ereignis, die Klausur zu bestehen
 $P(\bar{Ü} | \bar{B}) = 0,50$
- Rund ein Drittel aller Krebserkrankungen gehen auf das Konto von Tabakrauch
 K... Ereignis an Krebs zu erkranken R... Ereignis zu rauchen
 $P(K | R) = 0,33$

Eigenschaften

1. Falls $P(A) \neq 0$ und $P(B) \neq 0$, so gilt $\frac{P(A|B)}{P(A)} = \frac{P(B|A)}{P(B)}$
2. Zwei Ereignisse A und B sind unabhängig, wenn für die bedingten Wahrscheinlichkeiten gilt:
 $P(A|B) = P(A)$ und $P(B|A) = P(B)$
3. Für zwei unabhängige Ereignisse A und B gilt: $P(A \cap B) = P(A) \cdot P(B)$

Bedingte Wahrscheinlichkeit am Wahrscheinlichkeitsbaum

Ereignisse A und B in der Grundereignismenge Ω

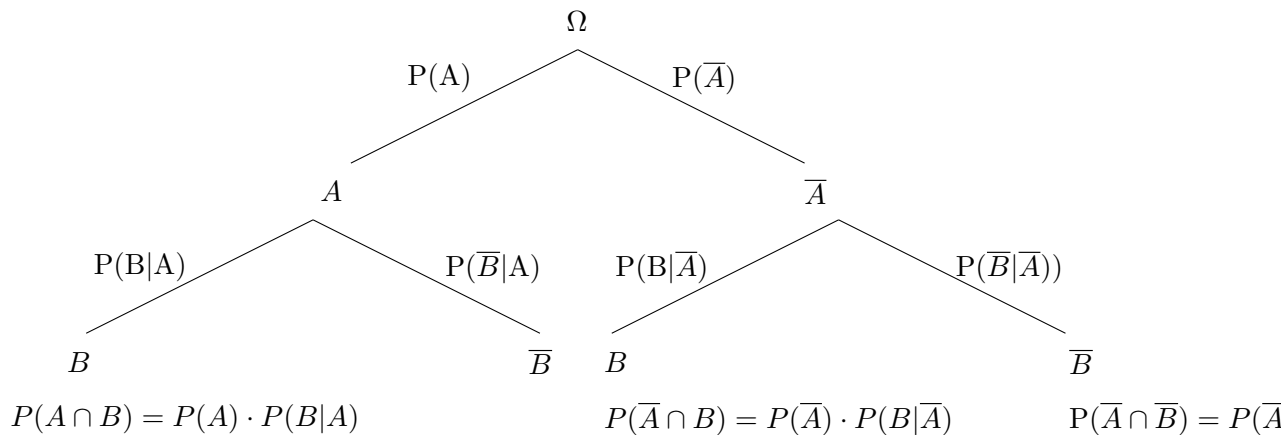


Abbildung 4.1: Wahrscheinlichkeitsbaum allgemein

Beispiel mit drei Vorgehensmethoden

Es handelt sich um eine Statistikvorlesung, zu der man Vorleistungen in Form von Übungsaufgaben erbringen kann

Folgende Annahmen können aus der Erfahrung getroffen werden:

- Von den Studenten, die die Übungsaufgaben bearbeiten, werden 80% die Klausur bestehen
- Von den Studenten, die die Übungsaufgaben nicht bearbeiten, werden 40% die Klausur bestehen
- 90% der Studenten haben die Übungsaufgaben bearbeitet

Wie groß ist die Wahrscheinlichkeit die Klausur zu bestehen?

Lösung:

1. Lösungsmethode: Mengentheoretisch

a) Benennen der Ereignisse

B := das Ereignis die Statistiklausur zu bestehen

U := das Ereignis die Übungsaufgaben zu bearbeiten

b) Umformen der Aussagen in Wahrscheinlichkeitsangaben:

$$0,8 = P(B|U) = \frac{P(B \cap U)}{P(U)}$$

$$0,4 = P(B|\bar{U}) = \frac{P(B \cap \bar{U})}{P(\bar{U})}$$

$$0,9 = P(U)$$

c) Berechnen der gesuchten Grösse. Gesucht $P(B)$:

$$P(B) = P(B \cap U) + P(B \cap \bar{U}) = 0,8P(U) + 0,4P(\bar{U})$$

$$= 0,8 \cdot 0,9 + 0,4 \cdot 0,1 = 0,72 + 0,04 = 0,76$$

\Rightarrow Die Wahrscheinlichkeit die Klausur zu bestehen beträgt 76%

2. Lösungsmethode: Wahrscheinlichkeitsbaum

a) Ausgehend von einer Wurzel, zeichnet man abgehende Äste für jedes Element der Partition. Jeder Ast endet in einem Knoten, der mit U , \bar{U} beschriftet wird. Die Äste beschriftet man mit den Wahrscheinlichkeiten der Partitionen.

b) An jeden Knoten hängt man je zwei weitere Äste, die in den Knoten B und \bar{B} enden, an. Diese Äste beschriftet man mit den Wahrscheinlichkeiten $P(B|U)$ bzw. $P(\bar{B}|U)$ usf.

c) Die Wahrscheinlichkeit $P(B)$ liest man ab, indem man alle Wege betrachtet, die von der Wurzel zum Ereignis B führen. Die Wahrscheinlichkeiten, die auf diesen Wegen liegen werden miteinander multipliziert und die Ergebnisse unterschiedlicher Wege addiert.

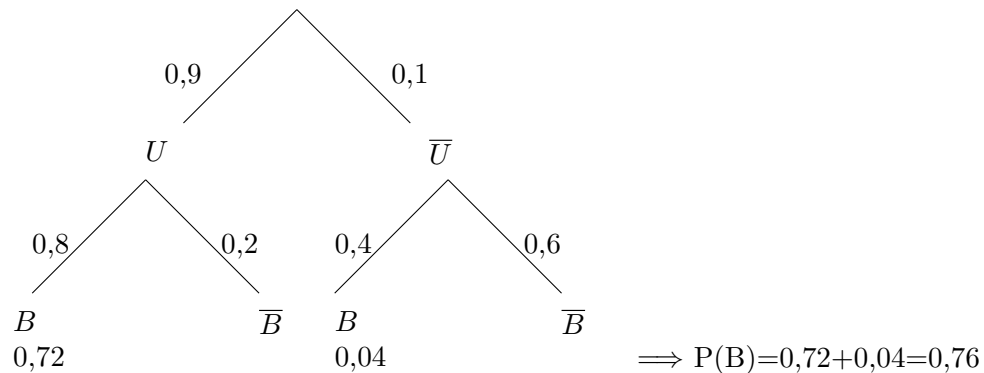


Abbildung 4.2: Wahrscheinlichkeitsbaum

3. Lösungsmethode: Wahrscheinlichkeitstableau

Hier werden alle Wahrscheinlichkeiten aller Kombinationen von Schnittmengen $P(B \cap U), P(B \cap \bar{U}), P(\bar{B} \cap U), P(\bar{B} \cap \bar{U})$ eingetragen

	U	\bar{U}	
B	0,72	0,04	0,76
\bar{B}	0,18	0,06	0,24
	0,9	0,1	1

Abbildung 4.3: Wahrscheinlichkeitstableau

Satz von der totalen Wahrscheinlichkeit:

$$P(A) = P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})$$

Satz von Bayes: $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$

Definition: Stochastische Unabhängigkeit:

Zwei Ereignisse A und B heißen **stochastisch unabhängig** wenn das Eintreten des einen Ereignisses die Wahrscheinlichkeit des Eintretens des anderen Ereignisses nicht beeinflusst.

kurz:

$$P(A|B) = P(A) \text{ bzw. } P(B|A) = P(B)$$

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

Beispiele

- Experiment: Würfeln eines fairen Würfels

Ereignis A .. Augenzahl ist ≥ 5 , d.h. $A = \{5; 6\}$

Ereignis B .. Augenzahl ist gerade, d.h. $B = \{2; 4; 6\}$

Sind die Ereignisse unabhängig?

$$P(A) = \frac{2}{6}; P(B) = \frac{3}{6}; P(A \cap B) = P(\{6\}) = \frac{1}{6}$$

$$P(A) \cdot P(B) = \frac{2}{6} \cdot \frac{3}{6} = \frac{1}{6} = P(A \cap B)$$

\implies

A und B sind unabhängig.

- Experiment: Würfeln eines fairen Würfels

Ereignis C .. Augenzahl ist ≥ 4 , d.h. $C = \{4; 5; 6\}$

Ereignis B .. Augenzahl ist gerade, d.h. $B = \{2; 4; 6\}$

Sind die Ereignisse unabhängig?

$$P(C) = \frac{3}{6}; P(B) = \frac{3}{6}; P(A \cap B) = P(\{4; 6\}) = \frac{1}{3}$$

$$P(C) \cdot P(B) = \frac{3}{6} \cdot \frac{3}{6} = \frac{1}{4} \neq P(A \cap B)$$

\implies

A und B sind abhängig.

4.4.1 Summen- und Produktregeln

Bei zwei Ereignissen A und B gilt

- Summenregel
Sind A und B disjunkte Ereignisse, so ist die Wahrscheinlichkeit dafür, dass entweder das Ereignis A **oder** das Ereignis B eintritt: $P(A) + P(B)$
- Produktregel
Sind A und B unabhängige Ereignisse, so ist die Wahrscheinlichkeit dafür, dass das Ereignis A **und** das Ereignis B eintritt: $P(A) \cdot P(B)$

Hinweis

Eine sehr ausführliche Beschreibung der bedingten Wahrscheinlichkeit findet man unter

<https://www.youtube.com/watch?v=EBZndX0eXcs>

4.5 Zufallsvariable

Um die Methoden der Analysis in der Wahrscheinlichkeitsrechnung einsetzen zu können, braucht man die Begriffe Variable und Funktion.

Als **Zufallsvariable** bezeichnet man eine Funktion, die den Ergebnissen eines Zufallsexperimentes reelle Zahlen zuordnet.

Besteht der Wertebereich der Zufallsvariablen aus endlich oder abzählbar unendlich vielen Werten, dann spricht man von einer **diskreten Zufallsvariablen**

Besteht der Wertebereich der Zufallsvariablen aus \mathbb{R} oder aus Teilintervallen von \mathbb{R} , dann spricht man von einer **stetigen Zufallsvariablen**

Zufallsvariablen werden üblicherweise mit einem Großbuchstaben bezeichnet, z.B. X . Es hat sich dieser etwas irreführende Begriff durchgesetzt, obwohl Zufallsvariablen eigentlich Funktionen sind.

Beispiele:

1. Beim Zufallsexperiment Münzwurf kann man eine Zufallsvariable X definieren, indem man dem Ergebnis ω ="Kopf" den Wert 0 und dem Ergebnis ω ="Zahl" den Wert 1 zuordnet:

$$X(\omega) = \begin{cases} 0 & \text{Kopf} \\ 1 & \text{Zahl} \end{cases}$$

X ist eine diskrete Zufallsvariable.

2. Werden zwei Würfel gleichzeitig geworfen, so ist die Ergebnismenge: (1,1), (1,2), ..., (6,6)

Ordnet man jedem Ergebnis die Augensumme zu, so erhält man eine Zufallsvariable X in der Form:

$$X(1,1)=2, X(1,2)=3, \dots, X(6,6)=12$$

Ordnet man jedem Ergebnis das Produkt zu, so erhält man eine Zufallsvariable X in der Form:

$$X(1,1)=1, X(1,2)=2, \dots, X(6,6)=36$$

Man kann auch folgende Zuordnung durchführen:

$$X(a,b) = \begin{cases} 0 & \text{falls } a \neq b \\ 1 & \text{falls } a = b \end{cases}$$

3. Die Brenndauer T einer aus einem Produktionsvorrat willkürlich herausgegriffenen Glühlampe ist eine stetige Zufallsvariable. Das Elementarereignis $T=t$ tritt ein,

wenn die Brenndauer T gleich der Zeit t ist.

- Die gemessene Länge von Schrauben L ist eine stetige Zufallsvariable, denn es kann jeder reelle Wert vorkommen

4.5.1 Wahrscheinlichkeitsfunktion

diskrete Zufallsvariable

Ist die Zufallsvariable X diskret, so kann man ihre Wahrscheinlichkeitsfunktion in Form einer Tabelle angeben: erste Zeile sind die von X angenommenen Werte x_i und die zweite Zeile sind die dazugehörigen Wahrscheinlichkeiten p_i

Beispiel: X gebe die Augensumme zweier Würfel an:

x_i	2	3	4	...	12
p_i	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$...	$\frac{1}{36}$

stetige Zufallsvariable

Ist die Zufallsvariable X stetig, so wird ihre Wahrscheinlichkeitsfunktion als sogenannte Dichtefunktion angegeben: $p = f(t)$

Beispiel: X gebe die Zerfallsdauer einer radioaktiven Substanz an.
Die Dichtefunktion ist von der Form $f(t) = Ce^{-t}$

4.6 Verteilungsfunktion

Die Verteilung der Zufallsvariablen X wird durch die **Verteilungsfunktion** F_X beschrieben

$$F_X(t) = P(X \leq t) \text{ für } t \in \mathbb{R}$$

Die Verteilungsfunktion gibt an, mit welcher Wahrscheinlichkeit die Zufallsvariable X einen Wert zwischen $-\infty$ und t annimmt.

Häufig ist die Verteilungsfunktion einer Zufallsvariablen nicht bekannt. Man führt dann eine Anzahl von Experimenten durch, um diese Verteilung näherungsweise zu bestimmen.

Eigenschaften:

- F_X ist monoton wachsend, d.h. aus $t_1 < t_2$ folgt $F(t_1) \leq F(t_2)$
- F_X ist rechtsseitig stetig, d.h. $\lim_{h \rightarrow 0} F_X(t + h) = F_X(t)$

$$3. \lim_{t \rightarrow \infty} F_X(t) = 1 \text{ und } \lim_{t \rightarrow 0} F_X(t) = 0$$

Jede Funktion $F : \mathbb{R} \rightarrow [0, 1]$ mit den Eigenschaften

- F ist monoton wachsend
- F ist rechtsseitig stetig
- $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$

ist eine Verteilungsfunktion.

Diskrete Verteilungen

Nimmt die Zufallsvariable X nur diskrete Werte x_i ($i=1,2,\dots$) an, mit den Wahrscheinlichkeiten $P(X = x_i) = p_i$ ($i = 1, 2, \dots$), so ist die Verteilungsfunktion diskret und lautet:

$$F(x) = \sum_{x_i \leq x} p_i$$

Stetige Verteilungen

Ist die Zufallsvariable kontinuierlich, so ist die Wahrscheinlichkeit dafür, dass sie einen bestimmten Wert x_i annimmt gleich 0. Man betrachtet daher die Wahrscheinlichkeit dafür, dass X in einem endlichen Intervall $[a, b]$ liegt. Lässt sich dieses mit Hilfe einer Funktion $f(t)$, der **Dichtefunktion**, in der Form

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

darstellen, dann spricht man von einer stetigen Verteilungsfunktion:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

Flächeninterpretation, Quantil bei einer stetigen Verteilung

Durch die Einführung der Dichtefunktion kann die Wahrscheinlichkeit

$$P(X \leq x) = F(x)$$

als die Fläche zwischen der Dichtefunktion $f(t)$ und der Abszisse im Intervall $(-\infty, x]$ interpretiert werden.

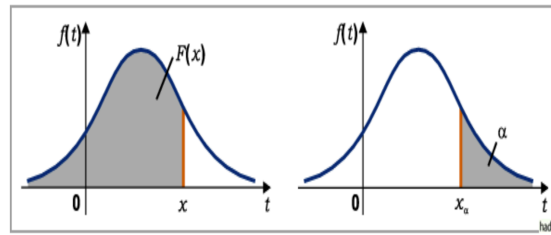


Abbildung 4.4: Wahrscheinlichkeitsdichte und Quantil

Ist $P(X \leq x) = \alpha$
 so nennt man die zugehörige Abszisse $x = x_\alpha$ **Quantil** der Verteilung.

Beispiele

1. Zufallsexperiment: Werfen mit zwei Würfeln

X sei die Zufallsvariable, die jedem Ergebnis die Augensumme zugeordnet.

Unter $X = x_i$ versteht man das Ereignis mit der Augensumme x_i . Die Wahrscheinlichkeit dieser Ereignisse kann man in einer Tabelle zusammenfassen:

x_i	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Die Verteilungsfunktion $F(x)$ ist somit

x	2	3	4	5	6	7	8	9	10	11	12
$F(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

2. Zufallsexperiment: Ziehen von Kugeln aus einer Urne mit Zurücklegen

Eine Urne enthalte 3 weiße und 2 schwarze Kugeln. Es werden nacheinander unter Berücksichtigung der Reihenfolge 3 Kugeln gezogen, wobei die Kugel jeweils wieder zurückgelegt wird.

Die Wahrscheinlichkeit eine weiße Kugel zu ziehen ist $\frac{3}{5}$

Die Wahrscheinlichkeit eine schwarze Kugel zu ziehen ist $\frac{2}{5}$

Da die Kugeln wieder zurückgelegt werden sind die Ziehungen voneinander unabhängig

Die Zufallsvariable X sei

$X =$ "Anzahl der erhaltenen schwarzen Kugeln bei den 3 Ziehungen mit Berücksichtigung der Reihenfolge und bei Zurücklegen."

Die Wahrscheinlichkeit der verschiedenen Ereignisse lassen sich in einer Tabelle aufstellen:

x_i	0	1	2	3
$P(X = x_i)$	$\frac{3 \cdot 3 \cdot 3}{5 \cdot 5 \cdot 5}$	$3 \cdot \frac{2 \cdot 3 \cdot 3}{5 \cdot 5 \cdot 5}$	$3 \cdot \frac{3 \cdot 2 \cdot 2}{5 \cdot 5 \cdot 5}$	$\frac{2 \cdot 2 \cdot 2}{5 \cdot 5 \cdot 5}$
	$= \frac{27}{125}$	$= \frac{54}{125}$	$= \frac{36}{125}$	$= \frac{8}{125}$

Die dazugehörige Verteilungsfunktion lautet:

x	0	1	2	3
$F(x)$	$\frac{27}{125}$	$\frac{27}{125} + \frac{54}{125} = \frac{81}{125}$	$\frac{81}{125} + \frac{36}{125} = \frac{117}{125}$	$\frac{117}{125} + \frac{8}{125} = \frac{125}{125} = 1$

3. Sei X eine stetige Zufallsvariable mit der Dichtefunktion

$$f(t) = \begin{cases} 0,02t & \text{mit } 0 \leq t \leq 10 \\ 0 & \text{sonst} \end{cases}$$

Die Verteilungsfunktion $F(x)$ ergibt sich als:

$$F(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ \int_0^x 0,02t \cdot dt = 0,01x^2 & \text{für } 0 < x \leq 10 \\ 1 & \text{für } x > 10 \end{cases}$$

Anwendung der Verteilungsfunktion

Ist X eine Zufallsvariable, so lässt sich mit Hilfe der Verteilungsfunktion F_X von X leicht die Wahrscheinlichkeit berechnen, dass X Werte in einem bestimmten Intervall annimmt:

Intervall	Wahrscheinlichkeit	diskrete Verteilung	stetige Verteilung
$(a, b]$	$P(a < X \leq b)$	$= F(b) - F(a)$	$= F(b) - F(a)$
(a, b)	$P(a < X < b)$	$= F(b) - P(b) - F(a)$	$= F(b) - F(a)$
(b, ∞)	$P(X > b)$	$= 1 - F(b)$	$= 1 - F(b)$
$[b, \infty)$	$P(X \geq b)$	$= 1 - F(b) + P(b)$	$= 1 - F(b)$

4.7 Kennwerte einer Wahrscheinlichkeitsverteilung

Zur groben Charakterisierung einer Verteilung werden vor allem die Parameter

- Mittelwert oder Erwartungswert μ
- Varianz σ^2
- Standardabweichung σ

einer Zufallsvariablen verwendet.

Diese Kennwerte haben folgende Bedeutung:

Erwartungswert	Dies entspricht dem Wert den die Zufallsvariable X, im Mittel annimmt. Der Erwartungswert kann als Schwerpunkt der Verteilung interpretiert werden
Varianz	Ist ein Maß für die Streuung der Wahrscheinlichkeit um den Erwartungswert
Standardabweichung	Quadratwurzel aus der Varianz

4.7.1 Erwartungswert, Varianz und Standardabweichung einer diskreten Zufallsvariablen

Die diskrete Zufallsvariable X mit den Wahrscheinlichkeitswerte $p_i = P(X = x_i)$ hat die folgenden Kennwerte

1. Mittelwert μ

$$\mu = E(X) = \sum_i x_i \cdot p_i$$

2. Varianz σ^2

$$\sigma^2 = Var(X) = \sum_i (x_i - \mu)^2 \cdot p_i$$

3. Standardabweichung σ

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

Bemerkungen

1. Ist die Zufallsvariable X gleichverteilt, d.h. ist $p_1 = p_2 = \dots = p_n = \frac{1}{n}$, so ist der Erwartungswert gerade das arithmetische Mittel:

$$\mu = x_1 \cdot \frac{1}{n} + x_2 \cdot \frac{1}{n} + \dots + x_n \cdot \frac{1}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

2. Bei einer symmetrischen Verteilung mit dem Symmetriezentrum x_0 gilt: $E(X) = x_0$
3. Für die Varianz gilt : $\sigma^2 \geq 0$
4. Die Varianz σ^2 ist der Erwartungswert der Zufallsvariable $Z = (X - \mu)^2$
Diese Zufallsvariable beschreibt die mittlere quadratische Abweichung vom Mittelwert μ

5. Für die Varianz σ^2 gilt auch die folgende Berechnung:

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

$$\text{denn: } \sigma^2 = E((X - \mu)^2) = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) + \mu^2 E(1) = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2$$

Beispiele

Berechnen Sie Erwartungswert, Varianz und Standardabweichung der folgenden diskreten Zufallsvariablen X

x_i	1	2	3	4
$f(x_i)$	1/8	3/8	3/8	1/8

2. Zufallsvariable "Wurf eines homogenen Würfels" ist gleichverteilt:

x_i	1	2	3	4	5	6
$f(x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

3. Beim Wurf mit zwei homogenen Würfeln ist die diskrete Zufallsvariable "X=erzielte Augensumme" folgendermaßen verteilt

x_i	2	3	4	5	6	7	8	9	10	11	12
$f(x_i)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

4. Die Wahrscheinlichkeitsverteilung für das Zufallsexperiment: "nacheinander Ziehen von Kugeln aus einer Urne mit Zurücklegen", wobei die Urne 3 weiße und 2 schwarze Kugeln enthält, ist:

x_i	0	1	2	3
$P(X = x_i)$	$\frac{27}{125}$	$\frac{54}{125}$	$\frac{36}{125}$	$\frac{8}{125}$

4.7.2 Erwartungswert, Varianz und Standardabweichung einer stetigen Zufallsvariablen

Die stetige Zufallsvariable X mit der Dichtefunktion $f(x)$ hat die folgenden Kennwerte

1. Mittelwert μ

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$$

2. Varianz σ^2

$$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) \cdot dx$$

3. Standardabweichung σ

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

Bemerkungen

1. Ist die Dichtefunktion symmetrisch mit dem Symmetriezentrum x_0 , so ist $\mu = x_0$

2. Auch bei einer stetigen Verteilung erfolgt die Berechnung der Varianz meist durch

$$\sigma^2 = E(X^2) - \mu^2$$
3. Physikalische Analogie für den Erwartungswert und für die Varianz:
 Physikalisch entspricht der Erwartungswert dem Schwerpunkt. Man muss sich dabei die Massen $R(X=x_i)$ an den Positionen x_i entlang vom Zahlenstrahl x plaziert vorstellen.
 Physikalisch entspricht die Varianz dem Trägheitsmoment, wenn man den oben beschriebenen Zahlenstrahl um eine Achse dreht, die senkrecht auf den Zahlenstrahl steht und die durch den Schwerpunkt verläuft

Beispiele

1. Die Lebensdauer T eines elektronischen Bauelements kann als exponentialverteilte Zufallsvariable mit der Dichtefunktion

$$f(t) = \begin{cases} 0 & \text{für } t < 0 \\ \lambda \cdot e^{-\lambda t} & \text{für } t \geq 0 \end{cases}$$
 betrachtet werden. Die mittlere Lebensdauer ist durch den Erwartungswert gegeben. Berechnen Sie den Erwartungswert, die Varianz und die Standardabweichung.
2. X sei eine stetige Zufallsvariable mit der linearen Dichtefunktion

$$f(x) = \begin{cases} 0,02x & \text{für } 0 \leq x \leq 10 \\ 0 & \text{sonst} \end{cases}$$
 Berechnen Sie Erwartungswert, Varianz und Standardabweichung.

4.7.3 Erwartungswert, Varianz und Standardabweichung einer linearen Funktion der Zufallsvariablen X

X sei eine diskrete oder stetige Zufallsvariable mit $E(X) = \mu_x$ und $Var(X) = \sigma_x^2$. Die Kennwerte einer linearen Funktion von X lassen sich hieraus bestimmen:

Sei $Z = aX + b$ mit $a, b \in \mathbb{R}$, dann ist

1. Erwartungswert μ_z

$$\mu_z = E(Z) = E(aX + b) = aE(X) + b = a\mu_x + b$$

2. Varianz σ_z^2

$$\sigma_z^2 = Var(Z) = Var(aX + b) = a^2 Var(X) = a^2 \sigma_x^2$$

3. Standardabweichung σ_z

$$\sigma_z = \sqrt{Var(Z)} = |a| \sigma_x$$

Nachweis für die Varianz von Z

$$\begin{aligned}
\text{Var}(Z) &= E(Z^2) - \mu_z^2 = E((aX + b)^2) - (a^2\mu_x^2 + 2ab\mu_x + b^2) \\
&= E(a^2X^2 + 2abX + b^2) - a^2\mu_x^2 - 2ab\mu_x - b^2 \\
&= a^2E(X^2) + 2abE(X) + b^2 - a^2\mu_x^2 - 2ab\mu_x - b^2 \\
&= a^2E(X^2) - a^2\mu_x^2 \\
&= a^2(E(X^2) - \mu_x^2) = a^2\text{Var}(X)
\end{aligned}$$

Bemerkungen

1. Für folgende Sonderfälle gilt:

$$Z=b \quad E(Z)=b, \text{Var}(Z)=0, \sigma_z = 0$$

$$Z=X+b \quad E(Z)=E(X)+b, \text{Var}(Z)=\text{Var}(X), \sigma_z = \sigma_x$$

$$Z=aX \quad E(Z)=aE(X), \text{Var}(Z)=a^2 \text{Var}(X), \sigma_z = |a|\sigma_x$$

Beispiele

1. Die Zufallsvariable X besitze die Kennwerte $\mu_x = 10$ und $\sigma_x^2 = 1$.
Berechnen Sie Erwartungswert, Varianz und Standardabweichung der Zufallsvariablen $Z=2X+1$
2. Die Zufallsvariable X besitze die Kennwerte $\mu_x = 10$ und $\sigma_x^2 = 1$.
Berechnen Sie Erwartungswert, Varianz und Standardabweichung der Zufallsvariablen $Z = \frac{X - \mu_x}{\sigma_x}$.

Diese Transformation heißt **Standardtransformation**, d.h. Z ist die zu X gehörende standardisierte Zufallsvariable

Hinweise

- Ein Video zu Erwartungswert und Varianz einer diskreten Zufallsvariablen findet man hier

<https://www.youtube.com/watch?v=seQkeVD01fU>

- Ein Video zur Dichtefunktion:

<https://www.youtube.com/watch?v=TZxJ5zZvskc>

- Ein Video zu Erwartungswert einer stetigen Zufallsvariablen:

<https://www.youtube.com/watch?v=FdcCJw2uOPg>

4.8 Spezielle Wahrscheinlichkeitsverteilungen

Zur Bearbeitung einer statistischen Fragestellung, wie z.B.

Die Wartezeit in einem Restaurant wurde in 10 voneinander unabhängigen Messungen bestimmt. Daraus soll nun abgeschätzt werden, nach welcher Zeit 80% der Gäste bedient werden

lässt sich ein generelles Vorgehen angeben

1. Festlegen der Zufallsvariablen
2. Passende Verteilung bestimmen
3. Notwendige Parameter für die Verteilungsfunktion ermitteln
4. Mathematische Formulierung der Fragestellung
5. Berechnung der mathematischen Formulierung und Beantwortung der Frage

Um die passende Verteilung zu bestimmen, bietet die Statistik eine Schar von verschiedenen Verteilungsfunktionen an, die für bestimmte Fälle geeignet sind:

4.8.1 Binomialverteilung

Die **Binomialverteilung** ist eine der wichtigsten diskreten Wahrscheinlichkeitsverteilungen. Sie findet insbesondere Anwendung bei der Qualitätsüberprüfung am Warenausgang. Sie liegt vor, wenn:

- eine Folge n gleichartiger Versuche durchgeführt wird
- bei denen das gewünschte Ereignis jeweils mit der konstanten Wahrscheinlichkeit p eintritt

Sie gibt an, mit welcher Wahrscheinlichkeit sich insgesamt k Erfolge einstellen

Die Binomialverteilung findet überall dort Anwendung, wo z.B.

- alternative Entscheidungen zu treffen sind, wie z.B. bei Münzwurf oder der Qualitätskontrolle ("einwandfrei", Ausschuß)
- eine bestimmten Eigenschaft in einer Stichprobe bestimmt werden soll, bei der die Reihenfolge der Entnahme keine Rolle spielt
- die Gesamtanzahl von defekten Bauteilen, die unter identischen Bedingungen hergestellt worden sind, bestimmt werden sollen

Kennwerte der Binomialverteilung

p Trefferwahrscheinlichkeit;
 n Anzahl Versuche
 k Anzahl Treffer

- Wahrscheinlichkeitsfunktion

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad \text{mit} \quad \binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

- Verteilungsfunktion:

$$F(x) = P(X \leq x) = \sum_{k \leq x} \binom{n}{k} p^k q^{n-k}$$

- Erwartungswert $E(X) = np$
- Varianz $Var(X) = np(1 - p)$
- Standardabweichung $\sqrt{Var(X)} = \sqrt{np(1 - p)}$

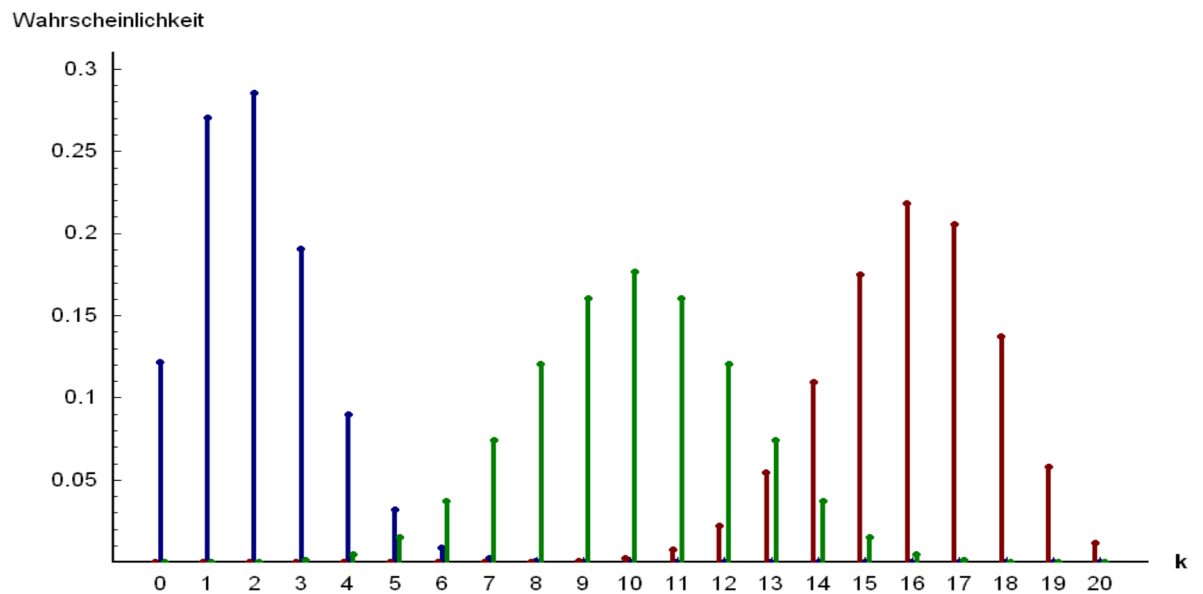


Abbildung 4.5: Wikipedia: Binomialverteilung

Beispiel

Eine Abfüllmaschine füllt erfahrungsgemäß 15% der Bierflaschen ungenau ab. Man überprüft 20 Bierflaschen. Berechnen Sie, wie viele ungenau abgefüllte Bierflaschen zu erwarten sind und wie hoch die Varianz und Standardabweichung sind.

Lösung:

Die Zufallsvariable X zähle die ungenau abgefüllten Bierflaschen. X aus dem angegebenen Beispiel ist Binomialverteilt mit $n=20$ und $p=0,15$.

Es ist $E(X)=np=3$ und $\text{Var}(X)=np(1-p)=3 \cdot 0,85=2,55$ und Standardabweichung $\sqrt{2,55} \approx 1,6$

4.8.2 Hypergeometrische Verteilung

Eine hypergeometrische Verteilung wird angenommen, wenn eine Gesamtmenge Teile mit einer interessierenden Eigenschaft besitzt, und man daran interessiert ist zu wissen, wie groß die Wahrscheinlichkeit bei Entnahme einiger Elemente, ist, dabei die Eigenschaft vorzufinden. Nach Entnahme der Teile werden diese nicht wieder zurückgelegt, somit ändert sich die Wahrscheinlichkeit ein entsprechendes Teil zu ziehen, mit jedem Zug. Sie findet insbesondere Anwendung bei der Qualitätsüberprüfung am Wareneingang

Für die hypergeometrische Verteilung ist somit die Kenntnis folgender Informationen erforderlich:

N : Anzahl der Elemente in der Grundgesamtheit

M : Anzahl der Elemente in der Grundgesamtheit mit der Eigenschaft A

n : Anzahl der entnommenen Elemente

Die hypergeometrische Verteilung spielt z.B. bei den Qualitäts- und Endkontrollen eines Herstellers oder den Abnahmekontrollen eines Kunden eine große Rolle.

Kennwerte der hypergeometrischen Verteilung

- Wahrscheinlichkeitsfunktion der hypergeometrischen Verteilung:

$$P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

- Verteilungsfunktion:

$$F(x) = \sum_{k \leq x} \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}$$

- Erwartungswert: $E(X) = n \cdot \frac{M}{N}$

- Varianz: $Var(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)}$
- Standardabweichung: $\sqrt{Var(X)} = \sqrt{\frac{nM(N-M)(N-n)}{N^2(N-1)}}$

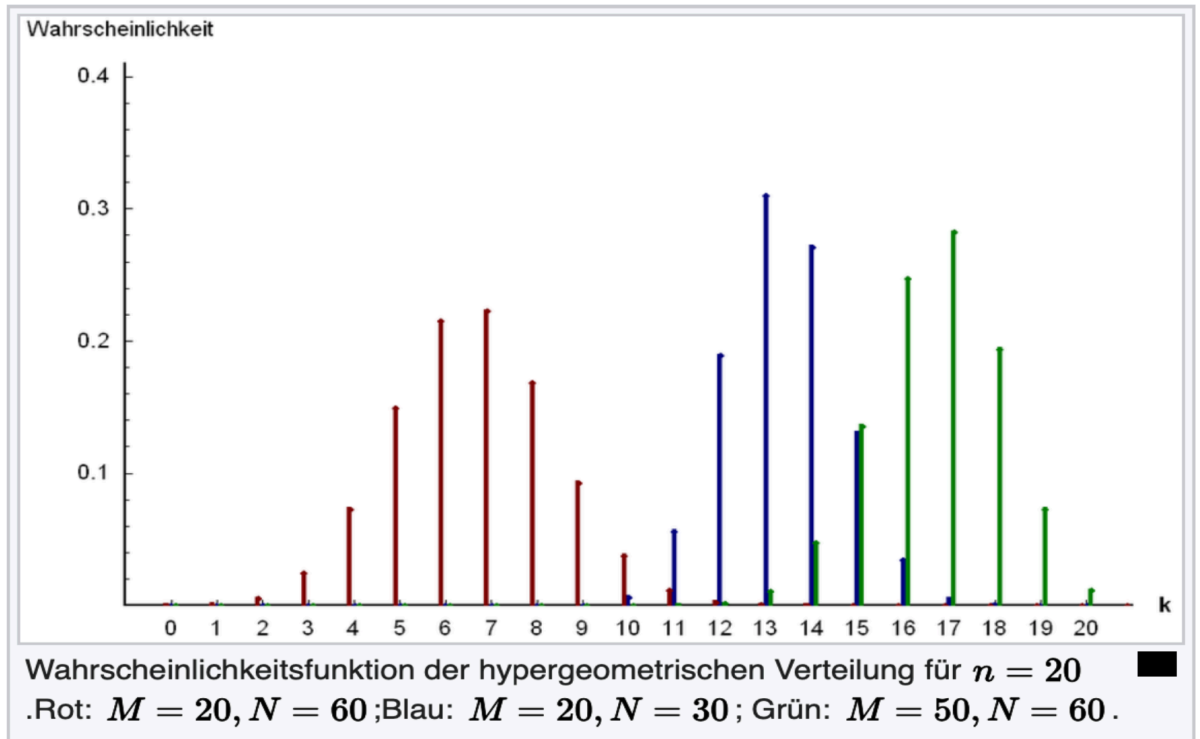


Abbildung 4.6: Wikipedia: Hypergeometrische Verteilung

Bemerkung

Der Unterschied der hypergeometrischen Verteilung zur Binomialverteilung ist nur, dass die Binomialverteilung davon ausgeht, dass sich die Wahrscheinlichkeit eines Treffers pro Zug nicht ändert. Sie funktioniert also nach dem Prinzip Ziehen mit Zurücklegen. Die hypergeometrische Verteilung hingegen geht davon aus, dass die Menge aus der gezogen wird begrenzt ist und dass alles was gezogen wurde nicht mehr zurückgelegt wird. Nach jedem Zug (bzw. nach jeder Wiederholung des Experiments) ändert sich also die Wahrscheinlichkeit.

Ist die Grundmenge N sehr groß und die entnommenen Elemente n dazu klein, so kann oft die Wahrscheinlichkeit $\frac{M}{N}$ als konstant angesehen werden und statt der hypergeome-

trischen Verteilung die Binomialverteilung verwendet werden

Beispiel

Eine Lieferung enthält $N=100$ Transistoren, die aus einer Massenproduktion mit 5% Ausschuß stammen. Bei der Anlieferung der Ware wird vom Kunden eine Abnahmekontrolle in Form einer Stichprobe vom Umfang $n=4$ ohne Zurücklegen durchgeführt. Die entnommenen Transistoren werden dabei auf ihre Funktionstüchtigkeit überprüft. Mit welcher Wahrscheinlichkeit enthält die Stichprobe nur einwandfreie Ware?

Lösung:

Die Zufallsvariable

X =Anzahl der in der Stichprobe vom Umfang $n=4$ angetroffenen defekten Transistoren

ist hypergeometrisch verteilt. Mit $N=100$, $M=5$ und $n=4$ lautet die Wahrscheinlichkeitsfunktion:

$$f(x) = P(X = x) = \frac{\binom{5}{x} \cdot \binom{100-5}{4-x}}{\binom{100}{4}}$$

Für $x=0$ folgt:

$$P(X = 0) = f(0) = \frac{\binom{5}{0} \cdot \binom{95}{4}}{\binom{100}{4}} = \frac{95 \cdot 94 \cdot 93 \cdot 92}{100 \cdot 99 \cdot 98 \cdot 97} = 0,8119 \approx 81,2\%$$

4.8.3 Poissonverteilung

Die diskrete Poissonverteilung liegt vor, wenn man sich für die Anzahl irgendwelcher Vorkommnisse pro (Zeit-) Einheit interessiert. Sie wird auch verwendet, wenn bei einer Binomialverteilung die Wahrscheinlichkeit für ein bestimmtes Ereignis sehr klein ist

Beispiele sind

- Anzahl X der Unfälle pro Wochendende
- Anzahl X der pro Stunde eintreffenden Telefonanrufe
- Anzahl X der Lackfehler pro cm^2 gefertigtem Blech

Kennwerte der Poissonverteilung

- Wahrscheinlichkeitsfunktion

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \text{ mit } \lambda > 0, k \in \mathbb{N}_0$$

Tipp: Berechnung der Wahrscheinlichkeits- und deren Verteilungsfunktion mit Taschenrechner

$$E(X) = \lambda$$

$$\text{Var}(X) = \lambda$$

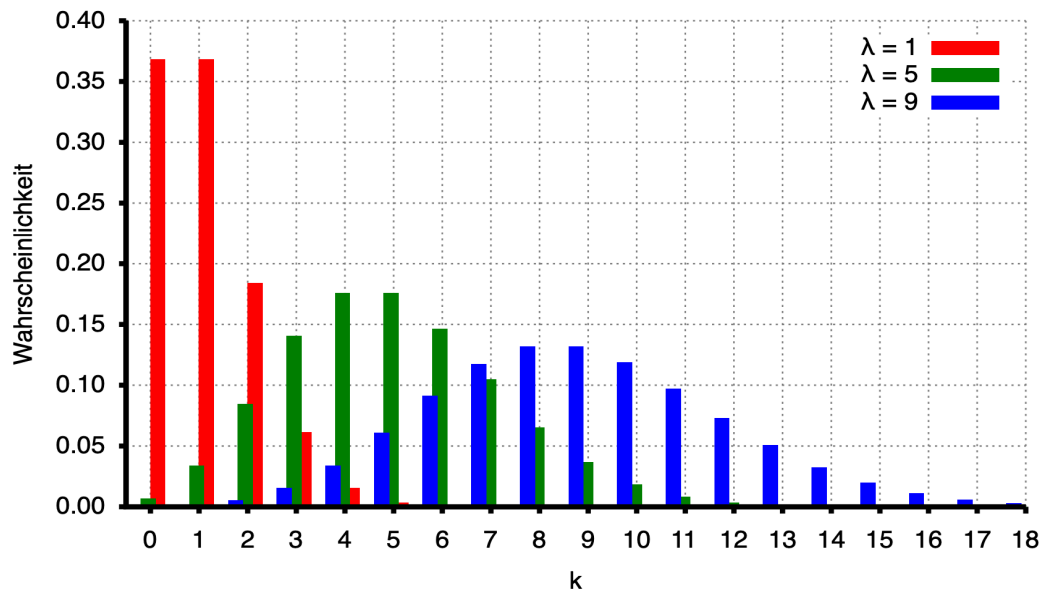


Abbildung 4.7: Wikipedia: Poissonverteilung

Beispiel

Bei einer nächtlichen Autobahnkontrolle auf der A14 im Januar diesen Jahres, registrierte die Polizei pro Stunde 5 Autofahrer mit überhöhter Geschwindigkeit. Wie groß ist die Wahrscheinlichkeit dafür, dass die Polizei bei diesen Kontrollen eine Stunde lang nichts zu tun hat?

Lösung:

Als Zufallsvariable X wird festgelegt:

X ... Anzahl Autofahrer mit überhöhter Geschwindigkeit in einer Stunde

Diese Zufallsvariable wird als poissonverteilt angenommen. Der Parameter λ der Poissonverteilung legt man fest zu $\lambda = E(X) = 5$

Die Frage formuliert man nun mit Hilfe der Wahrscheinlichkeit und der Verteilungsfunktion zu $P(X = 0)$

Diese ergibt sich zu

$$P(X = 0) = \frac{5^0}{0!} e^{-5} = \frac{1}{e^5} \approx 0,7\%$$

4.8.4 Stetige Gleichverteilung

Eine stetige Gleichverteilung liegt vor, wenn

- in einem Intervall $[a,b]$ jeder Wert möglich ist und
- einzelne Ereignisse keine unterschiedlichen Wahrscheinlichkeiten haben können

Kennwerte der stetigen Gleichverteilung

- Dichtefunktion $f(t) = \begin{cases} 0 & \text{für } t < a \\ \frac{1}{b-a} & \text{für } a \leq t \leq b \\ 0 & \text{für } t > b \end{cases}$
- Verteilungsfunktion $F(x) = \begin{cases} 0 & \text{für } t < a \\ \frac{x-a}{b-a} & \text{für } a \leq t \leq b \\ 1 & \text{für } t > b \end{cases}$
- $E(X) = \frac{a+b}{2}$
- $Var(X) = \frac{(b-a)^2}{12}$
- $\sqrt{Var(X)} = \frac{(b-a)}{2\sqrt{3}}$

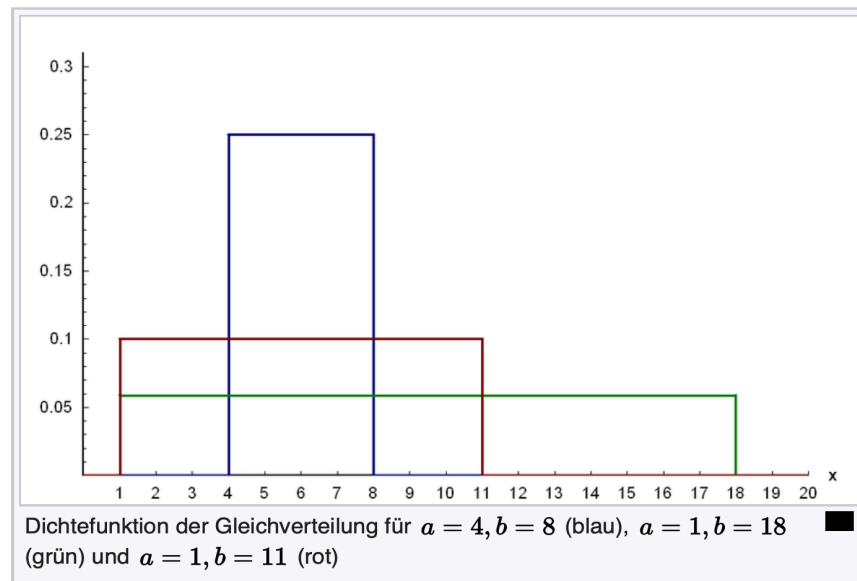


Abbildung 4.8: Wikipedia: stetige Gleichverteilung

Beispiel

Eine S-Bahn fährt alle 10 Minuten. Ein S-Bahngast kommt zu einem beliebigen Zeitpunkt. Seine Wartezeit beträgt 0-10 Minuten.

Sei X die Zufallsvariable **Wartezeit**

Geben Sie die Verteilung und die Dichtefunktion der Verteilung an. Wie lautet der Erwartungswert und die Standardabweichung?

Lösung:

Die Zufallsvariable X kann mit einer stetigen Gleichverteilung beschrieben werden.

$$\text{Die Dichtefunktion lautet: } f(t) = \begin{cases} 0 & \text{für } 0 \leq t \\ \frac{1}{10} & \text{für } 0 < t < 10 \\ 0 & \text{für } 10 \leq t \end{cases}$$

Der Mittelwert ist $\mu = 5$

Die Standardabweichung ist $\sigma = 2,88$

4.8.5 Exponentialverteilung

Die Exponentialverteilung ist eine stetige Verteilung, die bei der Modellierungen von Lebensdauern, Dauern von zufälligen Ereignissen angewendet wird.

Die Exponentialverteilung wird unter anderem zur Abbildung von Zeitintervallen genutzt. Der Parameter λ beschreibt die Zahl der erwarteten Ereignisse pro Zeitintervall mit $\lambda > 0$

Kennwerte der Exponentialverteilung

- Dichtefunktion $f(t) = \begin{cases} 0 & \text{für } x < 0 \\ \lambda \cdot e^{-\lambda \cdot x} & \text{für } 0 \leq x \end{cases}$
- Verteilungsfunktion $F(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-\lambda \cdot x} & \text{für } 0 \leq x \end{cases}$
- $E(X) = \frac{1}{\lambda}$
- $\text{Var}(X) = \frac{1}{\lambda^2}$
- $\sqrt{\text{Var}(X)} = \frac{1}{\lambda}$

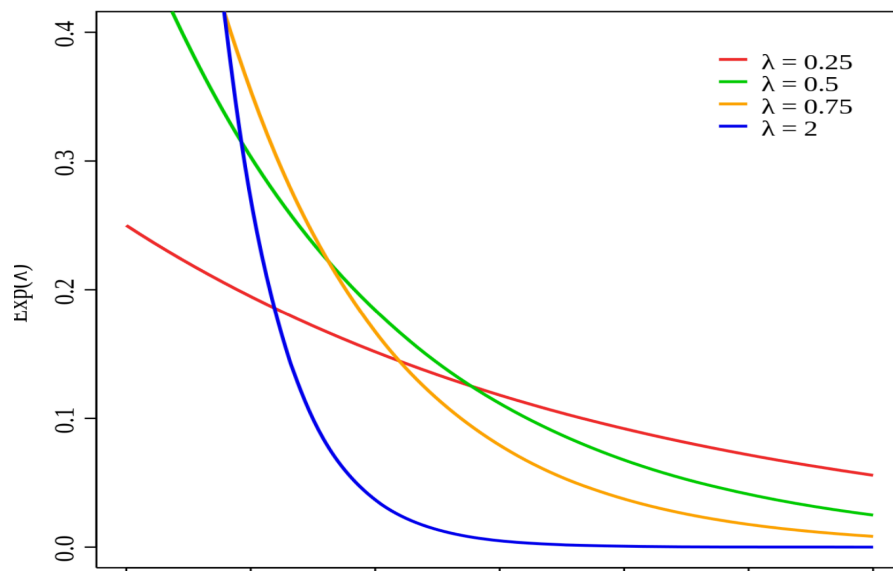


Abbildung 4.9: Wikipedia: Exponentialverteilung

Beispiel

Die Wartezeit in einem Restaurant wurde in 10 voneinander unabhängigen Messungen folgendermassen bestimmt (in min):

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
6,2	1,8	1,5	14,9	4,3	4,8	2,4	5,4	5,5	3,2

Sie wollen daraus nun abschätzen, nach welcher Zeit 80% der Gäste bedient werden.

Lösung:

Die Zufallsvariable X sei die Wartezeit

Für die Verteilungsfunktion von X wird die Exponentialverteilung angesetzt.

Zur Festlegung der Exponentialverteilung benötigt man den Parameter λ .

Aus $\frac{1}{\lambda} = E(X) = \frac{1}{10} (6,2 + 1,8 + \dots) = 5$ folgt $\lambda = \frac{1}{5}$

Die Frage kann man nun folgendermaßen mit Hilfe der Wahrscheinlichkeit und der Verteilungsfunktion formulieren:

Gesuchte ist die Zeit t mit $F(t) = 0,8$

Lösung:

$$1 - e^{-0,2t} = 0,8 \implies t = 8 \text{ min}$$

4.8.6 Weibull-Verteilung

Die Weibull-Verteilung ist eine stetige Verteilung, die zur Modellierung von Windgeschwindigkeiten, wie auch zur Beschreibung der Lebensdauer und Ausfallhäufigkeit von elektronischen Bauelementen herangezogen..

Kennwerte der Weibull-Verteilung

- Dichtefunktion $f(x) = \lambda k (\lambda x)^{k-1} \cdot e^{-(\lambda x)^k}$
- Verteilungsfunktion $F(x) = 1 - e^{-(\lambda x)^k}$
- $E(X) = \lambda^{-1} \Gamma(1 + \frac{1}{k})$
- $Var(X) = \lambda^{-2} \left(\Gamma(1 + \frac{2}{k}) - \Gamma^2(1 + \frac{1}{k}) \right)$

Bei zeitabhängigen Anwendungen wird λ durch den Kehrwert der charakteristischen Lebensdauer T ersetzt. T ist die Zeitspanne, nach der ca. 63% der Einheiten ausgefallen sind.

Für den Formparameter k werden in der Praxis typischerweise Werte im Bereich $0,25 \leq k \leq 5$ verwendet. Für $k=1$ ergibt sich die Exponentialverteilung und für $k \approx 3,602$ ergibt sich eine Verteilung, ähnlich der Normalverteilung

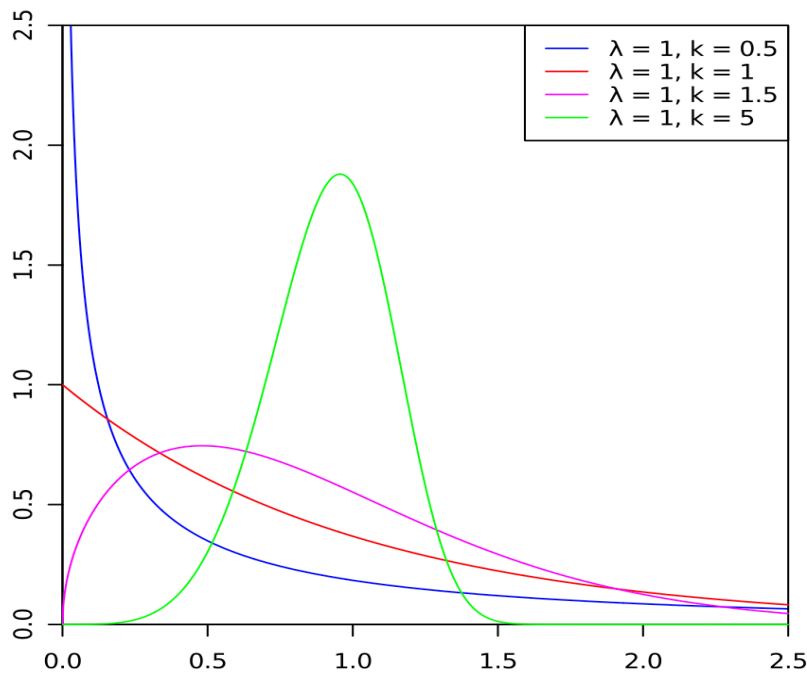


Abbildung 4.10: Wikipedia: Weibullverteilung

4.8.7 Gaußsche Normalverteilung

Die Gaußsche Normalverteilung ist eine stetige Verteilung, die dann unterstellt wird, wenn:

eigentlich ein bestimmter Wert erwartet wird, der aber zufällig nach oben oder unten abweichen kann

Zahlreiche Zufallsvariable, wie z.B. physikalische, technische Meßgrößen genügen dieser Verteilung.

Kennwerte der (Gaußschen) Normalverteilung

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

- Verteilungsfunktion $F(x) = P(-\infty < X \leq x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \cdot dt$
- $E(X) = \mu$
- $Var(X) = \sigma^2$
- Standardabweichung σ

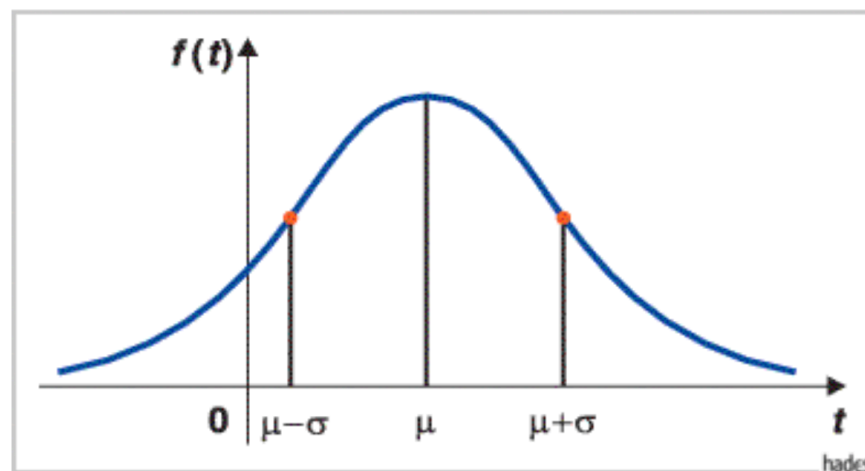


Abbildung 4.11: Wikipedia: Normalverteilung

Bemerkungen

- Die Dichtefunktion der Normalverteilung ist achsensymmetrisch bezüglich $x = \mu$
- Das einzige Maximum liegt bei $x_1 = \mu$
- Die beiden Wendepunkte liegen symmetrisch zum Maximum an den Stellen $x_{2,3} = \mu \pm \sigma$
- Je kleiner die Standardabweichung ist, umso höher liegt das Maximum und umso steiler fällt die Dichtefunktion nach beiden Seiten hin ab

Standardnormalverteilung

Eine Normalverteilung mit den Parametern $\mu = 0$ und $\sigma = 1$ heißt **Standardnormalverteilung**.

Eine normalverteilte Zufallsvariable X mit den Parametern μ und σ lässt sich immer mit Hilfe der Variablentransformation $U = \frac{X - \mu}{\sigma}$ in die standardnormalverteilte Zufallsvariable U überführen.

Die Verteilungsfunktion der Standardnormalverteilung ist

$$\phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{1}{2}t^2} dt$$

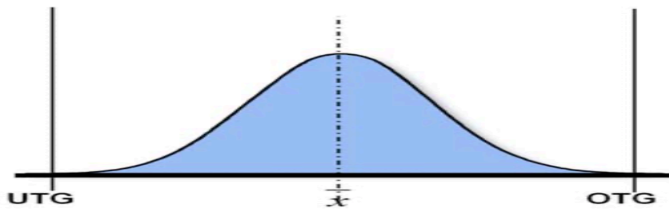
4.8.8 6σ

Six Sigma (6σ) ist ein Managementsystem zur Prozessverbesserung

Zentraler Punkt ist die Beschreibung, Messung, Analyse, Verbesserung und Überwachung von Geschäftsvorgängen mit statistischen Mitteln.

Bedeutung von 6σ :

- Prozessfähigkeit: sagt, wie gut das Ergebnis eines Prozesses mit den Anforderungen übereinstimmt
- OTG: Obere Toleranzgrenze
- UTG: Untere Toleranzgrenze



4.8.9 Berechnung der Wahrscheinlichkeiten der Normalverteilung

1. $P(X \leq x)$

$$\text{Es ist } P(X \leq x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \cdot dt.$$

Mit der Substitution $u = \frac{t-\mu}{\sigma}$ und somit $\frac{du}{dt} = \frac{1}{\sigma}$ folgt:

$$\begin{aligned} P(X \leq x) &= \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \cdot dt = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}u^2} \cdot \sigma du \\ &= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{1}{2}u^2} \cdot du = \phi\left(\frac{x-\mu}{\sigma}\right) \end{aligned}$$

2. $P(X \geq x)$

$$P(X \geq x) = 1 - P(X \leq x) = 1 - \Phi\left(\frac{x - \mu}{\sigma}\right)$$

3. $P(a \leq X \leq b)$

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

4. $P(|X - \mu| \leq k\sigma)$

Oftmals interessieren Intervalle, die symmetrisch um den Erwartungswert μ liegen.

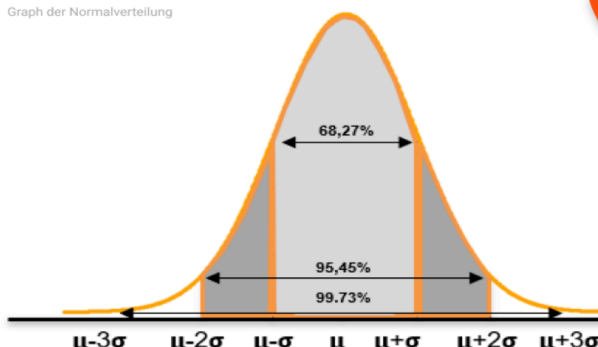
Es ist: $|X - \mu| \leq k\sigma \iff -k\sigma \leq X - \mu \leq k\sigma \iff -k\sigma + \mu \leq X \leq k\sigma + \mu$

und somit ist

$$P(|X - \mu| \leq k\sigma) = P(\mu - k\sigma \leq X \leq \mu + k\sigma) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$$

k=1	$P(X - \mu \leq \sigma) = 2 \cdot \Phi(1) - 1 = 2 \cdot 0,8413 - 1 = 0,6826$	$\approx 68\%$ liegen in diesem Intervall
k=2	$P(X - \mu \leq 2\sigma) = 2 \cdot \Phi(2) - 1 = 2 \cdot 0,9772 - 1 = 0,9544$	$\approx 95\%$ liegen in diesem Intervall
k=3	$P(X - \mu \leq 3\sigma) = 2 \cdot \Phi(3) - 1 = 2 \cdot 0,9987 - 1 = 0,9974$	$\approx 99\%$ liegen in diesem Intervall

Graph der Normalverteilung



Beispiel 1

In einem Werk werden Gewindeschrauben erstellt, deren Durchmesser eine normalverteilte Zufallsvariable X mit Mittelwert $\mu = 10\text{mm}$ und der Standardabweichung $\sigma = 0,2\text{mm}$. Toleriert werden dabei noch zufallsbedingte Abweichungen vom Solldurchmesser bis maximal $0,3\text{mm}$. Welcher Anteil an Ausschußware ist zu erwarten?

Lösung:

Man berechnet zunächst $P(9,7 \leq X \leq 10,3)$, die Wahrscheinlichkeit, dass die Schraubendurchmesser im Toleranzbereich sind. Um obige Gleichung anwenden zu können, ist $0,3\text{mm}$ durch $k\sigma$ zu ersetzen, d.h. $0,3 = k \cdot 0,2 \implies k = 1,5$.

Somit ist

$$P(9,7 \leq X \leq 10,3) = P(|X - 10| \leq k \cdot 0,2) = 2 \cdot \Phi(1,5) - 1 = 2 \cdot 0,9332 - 1 = 0,8664$$

Also sind 86,6% der Schrauben im Toleranzbereich. Der zu erwartende Ausschuß beträgt damit 13,4%.

Beispiel 2

Eine Lieferung enthält $N=100$ Transistoren, die aus einer Massenproduktion mit 5% Ausschuß stammen. Bei der Anlieferung der Ware wird vom Kunden eine Abnahmekontrolle in Form einer Stichprobe vom Umfang $n=4$ ohne Zurücklegen durchgeführt. Die entnommenen Transistoren werden dabei auf ihre Funktionstüchtigkeit überprüft. Mit welcher Wahrscheinlichkeit enthält die Stichprobe nur einwandfreie Ware?

Lösung:

Die Zufallsvariable

X =Anzahl der in der Stichprobe vom Umfang $n=4$ angetroffenen defekten Transistoren

ist hypergeometrisch verteilt. Mit $N=100$, $M=5$ und $n=4$ lautet die Wahrscheinlichkeitsfunktion:

$$f(x) = P(X = x) = \frac{\binom{5}{x} \cdot \binom{100-5}{4-x}}{\binom{100}{4}}$$

Für $x=0$ folgt:

$$P(X = 0) = f(0) = \frac{\binom{5}{0} \cdot \binom{95}{4}}{\binom{100}{4}} = \frac{95 \cdot 94 \cdot 93 \cdot 92}{100 \cdot 99 \cdot 98 \cdot 97} = 0,8119 \approx 81,2\%$$

4.9 Zentraler Grenzwertsatz

4.9.1 Summe zweier binomialverteilter Zufallsvariablen

X binomialverteilte Zufallsvariable: $X \sim B(n_X, p)$

Y binomialverteilte Zufallsvariable: $Y \sim B(n_Y, p)$

X und Y unabhängig

\implies

$Z = X + Y$ binomialverteilte Zufallsvariable: $Z \sim B(n_X + n_Y; p)$.

4.9.2 Summe zweier normalverteilter Zufallsvariablen

X normalverteilte Zufallsvariable: $X \sim N(\mu_X, \sigma_X^2)$

Y normalverteilte Zufallsvariable: $Y \sim N(\mu_Y, \sigma_Y^2)$

X und Y unabhängig

\implies

$Z = X + Y$ normalverteilte Zufallsvariable: $Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

4.9.3 Summe zweier beliebiger Zufallsvariablen

X Zufallsvariable mit $E(X) = \mu_X$ und $Var(X) = \sigma_X^2$

Y Zufallsvariable: $E(Y) = \mu_Y$ und $Var(Y) = \sigma_Y^2$

X und Y unabhängig

Jeweilige Verteilung unbekannt

\Rightarrow

Z = X + Y Zufallsvariable mit

$$E(Z) = \mu_X + \mu_Y$$

$$Var(Z) = \sigma_X^2 + \sigma_Y^2.$$

Verteilung unbekannt

Beispiel

Seien X und Y zwei unabhängige und identisch verteilte Zufallsvariablen mit Varianz $\sigma^2 = 2$.

Was ist die Varianz der Zufallsvariable $Z = 1 + X - Y - Y$?

LSG.:

$$Z = 1 + X - 2Y \Rightarrow$$

$$\sigma_Z^2 = \sigma_1^2 + \sigma_X^2 + (-2)^2 \sigma_Y^2 = 0 + 2 + 4 \cdot 2 = 10$$

4.9.4 Unabhängige Zufallsvariablen mit gleichem Erwartungswert und gleicher Varianz

Für $i=1..n$

X_i Zufallsvariable mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$

X_i unabhängig; Verteilung unbekannt

\Rightarrow

$S = \sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n$ Zufallsvariable mit

$$E(S) = E(X_1) + E(X_2) + \dots + E(X_n) = n\mu$$

$$Var(S) = Var(X_1) + Var(X_2) + \dots + Var(X_n) = n\sigma^2$$

\Rightarrow

$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ Zufallsvariable mit

$$E(\bar{X}) = E\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu$$

$$Var(\bar{X}) = Var\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n\right) = \frac{1}{n^2} \sum_{i=1}^n Var(X_i) = \frac{\sigma^2}{n}$$

4.9.5 Zentraler Grenzwertsatz

Ist (X_n) eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit identischen Erwartungswerten $E(X_i) = \mu$ und

identischen Varianzen $Var(X_i) = \sigma^2$.

Dann ist für $n \geq 30$:

$S_n = \sum_{i=1}^n X_i$ und $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ normalverteilt,

unabhängig von der tatsächlichen Verteilung der einzelnen X_i :

$$P(S_n \leq x) \approx \phi\left(\frac{x - n \cdot \mu}{\sigma \cdot \sqrt{n}}\right)$$

bzw.

$$P(\bar{X}_n \leq x) \approx \phi\left(\frac{x - \mu}{\sigma} \cdot \sqrt{n}\right)$$

siehe hierzu:

www.onlinestatbook.com/stat_sim/sampling_dist/index.html

Beispiel

Beispielweise wollen Sie bestimmen, wieviel Euro Personen pro Monat in Frankreich für Fastfood ausgeben. Aus diesem Anlass ziehen Sie nacheinander aus der Population der französischen Staatsbürger Zufallsstichproben mit immer derselben Größe n . Hierbei ziehen sie jedoch nicht nur eine oder zwei Stichproben, sondern sehr, sehr viele gleichgroße Zufallsstichproben. Für jede Stichprobe berechnen Sie daraufhin den Mittelwert der Ausgaben für Fastfood pro Monat. So erhalten Sie sehr viele Mittelwerte, die sich von Stichprobe zu Stichprobe leicht unterscheiden (je nachdem wie viele Fastfood-Fans Sie zufällig in der Stichprobe haben). Wenn Sie nun alle Mittelwerte in eine Liste schreiben und der Größe nach ordnen entsteht eine neue Verteilung? die Verteilung der Stichproben-Mittelwerte., die nach dem zentralen Grenzwertsatz einer Normalverteilung entspricht.

5

Schließende Statistik

Eine grundlegende Aufgabe der Statistik besteht darin, Kenntnisse und Informationen über die Eigenschaften einer bestimmten Menge von Objekten zu gewinnen, ohne daß man dabei alle Objekte in die Untersuchung mit einbeziehen muß. Letzteres ist meist auch gar nicht möglich, da z.B.

- zu hoher Zeit- und Kostenaufwand
- Die Anzahl der Objekte ist zu groß
- Die Objekte könnten bei der Untersuchung zerstört werden (vgl. Messen der Lebensdauer)

Dieses Ziel hat die schließende, bzw. induktive Statistik. Bei der schließenden Statistik geht es darum aus einer Stichprobe mit Hilfe der Wahrscheinlichkeitsrechnung Rückschlüsse auf die zugrundeliegende Grundgesamtheit zu ziehen, bzw. Vermutungen über die Grundgesamtheit zu stützen oder zu verwerfen.

Im Folgenden werden im Wesentlichen zwei Fragestellungen behandelt.

1. Zum einen die Frage, wie man die unbekannten Parameter einer Verteilung bestimmt, wenn man zwar die Art der Verteilung kennt, nicht aber die notwendigen Parameter. Dies haben wir bei den vorangegangenen Beispielen oftmals so gelöst, dass wir den Erwartungswert durch den Mittelwert einer Stichprobe beschrieben haben. Dieses Vorgehen wird hier nun (im Nachhinein) untermauert.
2. Zum anderen werden Vermutungen (Hypothesen) über den Wert eines Parameters der Verteilung getestet. Dies ist immer dann wichtig, wenn man einen Standpunkt vertreten will.

5.1 Parameterschätzungen

Für die Schätzung der Parameter, wie z.B. dem Mittelwert oder der Varianz einer Verteilung, gibt es zwei Wege

1. Punktschätzung

Als **Punktschätzer** wird eine Schätzfunktion, die jeder Stichprobe einen Wert zuordnet, der eine gewisse Eigenschaft der zugrundeliegenden Wahrscheinlichkeit schätzen soll. In den meisten Anwendungen ist die interessierende Größe ein Parameter der Wahrscheinlichkeitsverteilung (wie z. B. der Erwartungswert μ einer Normalverteilung)

2. Intervallschätzung

Ein **Intervallschätzer** ist eine Schätzfunktion, die jedem Ausgang eines statistischen Experimentes ein Intervall zuordnet und nicht nur einen einzelnen Wert.

Intervallschätzer bilden die Grundlage für die Bestimmung von Konfidenzintervallen. Dies sind die Intervalle, bei denen eine vorgegebene Überdeckungswahrscheinlichkeit garantiert ist.

5.1.1 Vor-Nachteile

	Vorteil	Nachteil
Punktschätzung		
Intervallschätzung		

5.1.2 Qualitätskriterien für Punktschätzer

Für die Schätzung der unbekannten Parameter einer Wahrscheinlichkeitsverteilung werden spezielle Funktionen benötigt. Sie ermöglichen die näherungsweise Berechnung dieser Parameter unter Verwendung einer konkreten Stichprobe, die man der Grundgesamtheit entnimmt.

An diese Schätzfunktionen T hat man gewisse Anforderungen:

- sie sollen **erwartungstreu** sein, d.h. im Mittel tatsächlich den unbekannten Parameter λ als Ergebnis haben.
 $E(T) = \lambda$
- sie sollen **konsistent** sein, d.h. die Ergebnisse der Schätzfunktion sollen mit zunehmender Zahl von Werten den unbekannten Parameter annähern.
 $\lim_{n \rightarrow \infty} Var(T) = 0$

5.1.3 Schätzfunktionen der wichtigsten Parameter

Gesuchter Parameter der Verteilung der Zufallsvariablen X	erwartungstreue und konsistente Schätzfunktion für den Parameter
Erwartungswert $E(X) = \mu$	$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
Varianz $\text{Var}(X) = \sigma^2$	$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
Anteilswert p für Ereignis A	$P = \frac{X}{n}$ X.. Anzahl der Erfolge bei n-facher Ausführung

5.1.4 Schätzfunktionen bei speziellen Verteilungen

Verteilung	Parameter	Schätzfunktion
Binomialverteilung	p	
Poissonverteilung	λ	
Exponentialverteilung	λ	
Normalverteilung	μ, σ	

Beispiele

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

- Der Erwartungswert der Zufallsvariable X wird aus n unabhängigen Wiederholungen von X geschätzt. Das arithmetische Mittel der n Beobachtungswerte ist ein erwartungstreuer Punktschätzer für den Erwartungswert.

$$E(T) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \mu$$

ist ein konsistenter Punktschätzer für den Erwartungswert.

$$\text{Var}(T) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n} \sigma^2 \rightarrow 0 \text{ für } n \rightarrow \infty$$

- Die Varianz der Zufallsvariable X wird aus n unabhängigen Wiederholungen von X geschätzt. Dann ist $\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2$ ein erwartungstreuer Punktschätzer für die Varianz.

$$\begin{aligned}
E(T) &= E\left(\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^n ((X_i - \mu) - (\bar{X} - \mu))^2\right) \\
&= \frac{1}{n-1} E\left(\sum_{i=1}^n ((X_i - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu) + (\bar{X} - \mu)^2)\right) \\
&= \frac{1}{n-1} \left(E\left(\sum_{i=1}^n ((X_i - \mu)^2 - 2(\bar{X} - \mu)n \frac{1}{n} \sum_{i=1}^n (X_i - \mu)) + n(\bar{X} - \mu)^2\right)\right) \\
&= \frac{1}{n-1} \left(E\left(\sum_{i=1}^n ((X_i - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2)\right)\right) \\
&= \frac{1}{n-1} \left(E\left(\sum_{i=1}^n ((X_i - \mu)^2 - n(\bar{X} - \mu)^2)\right)\right) \\
&= \frac{1}{n-1} \left(\sum_{i=1}^n E((X_i - \mu)^2) - nE(\bar{X} - \mu)^2\right) \\
&= \frac{1}{n-1} \left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \sigma^2
\end{aligned}$$

3. $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2$ ist kein erwartungstreuer Punktschätzer für die Varianz.

$$\begin{aligned}
E(T) &= E\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2\right) = \frac{1}{n} E\left(\sum_{i=1}^n ((X_i - \mu) - (\bar{X} - \mu))^2\right) \\
&= \dots \\
&= \frac{1}{n} \left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \frac{n-1}{n} \sigma^2
\end{aligned}$$

Beispiel

Die Lebensdauer T eines bestimmten elektronischen Bauelements genüge einer Exponentialverteilung mit dem unbekannten Parameter λ . Man ermittelt einen Schätzwert λ für diesen Parameter anhand der folgenden Stichprobe:

i	1	2	3	4	5	6	7	8
t_i	950	980	1150	770	1230	1210	990	1120

Aus der Stichprobe erhält man den Mittelwert

$$\bar{t} = \frac{1}{8} \sum_{i=1}^8 t_i = \frac{1}{8} (950 + 980 + \dots + 1120) = 1050$$

Somit ist

$$\lambda = \frac{1}{\bar{t}} = \frac{1}{1050} \approx 0,00095$$

ein Schätzwert für den unbekannten Parameter λ

5.1.5 Maximum-Likelihood-Methode

Die Maximum-Likelihood (ML)-Methode zählt zu den gängigsten Verfahren zur Schätzung von Parametern einer Grundgesamtheit auf Basis einer Stichprobe. Ihr Grundgedanke ist, den Wert eines Parameters so zu bestimmen, dass das Auftreten der konkreten Beobachtungen in der Stichprobe am wahrscheinlichsten ist.

Verfahren

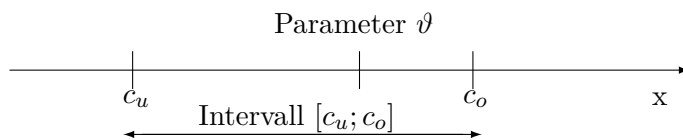
1. Aufstellen der ML-Funktion

$$L(\Theta) = \prod_{i=1}^n f_{X_i}(p)$$
2. Logarithmieren: $\log(L(\Theta)) = \sum_{i=1}^n \log f_{X_i}(\Theta)$
3. Bestimmen der Ableitung $\frac{d}{d\theta} \log(L(\Theta))$
4. nach gesuchtem Parameter auflösen

5.2 Intervallschätzungen

Zuvor haben wir uns damit beschäftigt, näherungsweise einen unbekannten Parameter der Wahrscheinlichkeitsverteilung zu schätzen. Diesen erhielten wir durch eine Schätzfunktion aus einer konkreten Stichprobe. Diese sogenannte Punktschätzung ermöglicht jedoch keine Aussage über die Genauigkeit der Schätzung. Der aus einer Zufallsstichprobe gewonnene Schätzwert kann noch erheblich vom tatsächlichen Wert abweichen.

Anders schaut es aus, wenn man darauf verzichtet einen Wert als Parameter zu haben, sondern nur ein Intervall für den gesuchten Parameter angibt. Dann kann man die Wahrscheinlichkeit berechnen, dass das Intervall tatsächlich den unbekannten Parameter überdeckt. Bzw. kann man umgekehrt, die Wahrscheinlichkeit vorgeben, sozusagen als Anforderung, und dann ein Intervall berechnen, das diesen Anforderungen genügt.

**Definitionen**

Ein **Konfidenzintervall** ist ein Intervall, das bei unendlicher Wiederholung eines Zufallsexperiments mit einer gewissen Wahrscheinlichkeit (dem **Konfidenzniveau** $1-\alpha$) die wahre Lage des Parameters einschließt.

Man nennt α die **Irrtumswahrscheinlichkeit**.

Ein häufig verwendetes Konfidenzniveau ist 95 %, so dass in diesem Fall (wenn man das Zufallsexperiment auf identische Art und Weise wiederholt) das Konfidenzintervall in näherungsweise 95 % aller Fälle den unbekannten Parameter überdecken wird.

I: Konfidenzintervall für den Erwartungswert einer Normalverteilung bei bekannter Varianz**Herleitung**

- Schätzfunktion für den Erwartungswert ist $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
- \bar{X} ist normalverteilt mit $E(\bar{X}) = \mu$ und $Var(\bar{X}) = \frac{\sigma^2}{n}$
- $U := \frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n}$ ist standardnormalverteilt
- Gesucht: symmetrisches Intervall um 0 mit $P(|U| \leq c) = 1 - \alpha$
- Bestimme $c = z_{1-\frac{\alpha}{2}}$ das $(1 - \frac{\alpha}{2})$ -Quantil der Standardnormalverteilung
- $-c \leq \frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n} \leq c \implies -\frac{c\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq \frac{c\sigma}{\sqrt{n}} \implies -\bar{X} - \frac{c\sigma}{\sqrt{n}} \leq -\mu \leq -\bar{X} + \frac{c\sigma}{\sqrt{n}}$
 $\implies \bar{x} - \frac{c\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{c\sigma}{\sqrt{n}}$

ZusammenfassungBestimme \bar{x} aus der Stichprobe

zweiseitiges Konfidenzintervall	$\left[\bar{x} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right]$
einseitiges unteres Konfidenzintervall	$\left(-\infty, \bar{x} + z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right]$
einseitiges oberes Konfidenzintervall	$\left[\bar{x} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, \infty \right)$

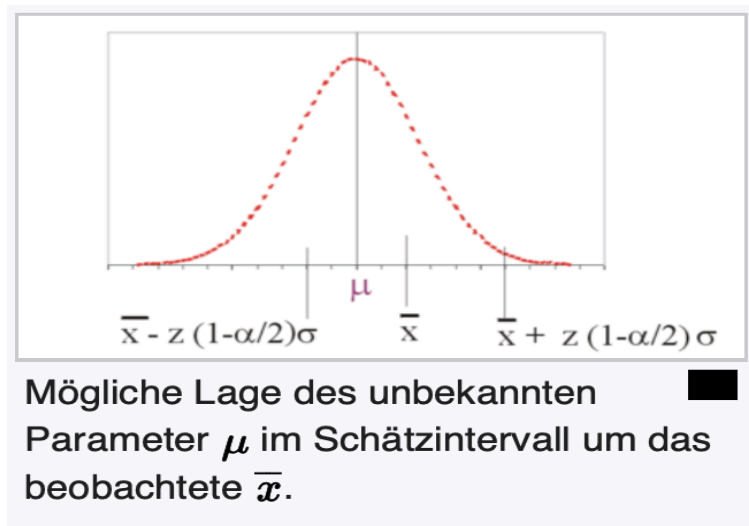


Abbildung 5.1: Wikipedia: Konfidenzintervall

Beispiel:

Auf 12 Versuchsflächen wurden Rapssorten für die Biokraftstoffproduktion angebaut. Folgende Erträge (in $10^4 m^2$) erbrachten die Flächen:

35,6 33,7 37,8 31,2 37,2 34,1 35,6 36,6 37,1 34,9 35,6 34,0

Aus Erfahrung weiß man, dass die Erträge als normalverteilt angesehen werden können.

Desweiteren ist die Varianz der Verteilung bekannt $\sigma^2 = 3,24$

Gesucht ist das zweiseitige 0,95-Konfidenzintervall für den durchschnittlichen Ertrag.

LSG.:

$n = 12$ $\bar{x} = 35,283$ $\alpha = 0,05$ $z_{1-\frac{\alpha}{2}} = z_{0,975} = 1,96$ $\sigma = \sqrt{3,24} = 1,8$ Das Konfidenzintervall lautet somit:

$$\left[\bar{x} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right] = \left[35,283 - 1,96 \cdot \frac{1,8}{\sqrt{12}}; 35,283 + 1,96 \cdot \frac{1,8}{\sqrt{12}} \right] = [34,265; 36,302]$$

Mindeststichprobenumfang

Die Länge eines zweiseitigen Konfidenzintervalls für μ berechnet sich zu

$$l = 2 \cdot z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Daraus kann man umgekehrt den Mindeststichprobenumfang für ein Konfidenzintervall der Länge l berechnen:

$$n \geq 4 \cdot \sigma^2 \cdot \left(\frac{z_{1-\alpha/2}}{l} \right)^2$$

Beispiel

Für das vorhergehende Beispiel der Rapsfelder möchte man ein Konfidenzintervall der Länge 0,5 bestimmen.

Lösung:

Hierfür ist eine größere Anzahl an Versuchsflächen erforderlich:

$$n \geq 4 \cdot 3,24 \cdot \left(\frac{1,96}{0,5}\right)^2 = 199,148$$

Man benötigt für die Länge 0,5 des Konfidenzintervalls mindestens 200 Versuchsflächen

II: Konfidenzintervall für den Erwartungswert einer Normalverteilung bei unbekannter Varianz

Herleitung

- Schätzfunktion für den Erwartungswert ist $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
- \bar{X} ist normalverteilt mit $E(\bar{X}) = \mu$ und $Var(\bar{X}) = \frac{\sigma^2}{n}$
- $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$ ist die Schätzfunktion für die unbekannt Standardabweichung
- $U := \frac{\bar{X} - \mu}{S} \cdot \sqrt{n}$ ist t-student-verteilt mit f=n-1 Freiheitsgraden
- Gesucht: symmetrisches Intervall um 0 mit $P(|U| \leq c) = 1 - \alpha$
- Bestimme $c = t_{n-1; 1-\frac{\alpha}{2}}$ das $(1 - \frac{\alpha}{2})$ -Quantil der t-student-Verteilung
- $-c \leq \frac{\bar{X} - \mu}{S} \cdot \sqrt{n} \leq c \implies -\frac{cS}{\sqrt{n}} \leq \bar{X} - \mu \leq \frac{cS}{\sqrt{n}} \implies -\bar{X} - \frac{cS}{\sqrt{n}} \leq -\mu \leq -\bar{X} + \frac{cS}{\sqrt{n}} \implies \bar{x} - \frac{cs}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{cs}{\sqrt{n}}$

Zusammenfassung

Bestimme \bar{x} und s aus der Stichprobe

zweiseitiges Konfidenzintervall	$\left[\bar{x} - t_{n-1; 1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1; 1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right]$
einseitiges unteres Konfidenzint.	$\left(-\infty, \bar{x} + t_{n-1; 1-\alpha} \cdot \frac{s}{\sqrt{n}} \right]$
einseitiges oberes Konfidenzint.	$\left[\bar{x} - t_{n-1; 1-\alpha} \cdot \frac{s}{\sqrt{n}}, \infty \right)$

Bemerkung

- Die Student-t-Verteilung geht auf William Gosset zurück. Er arbeitete als Chemiker bei der Guinness-Brauerei und beschäftigte sich mit der Statistik kleiner Stichproben.

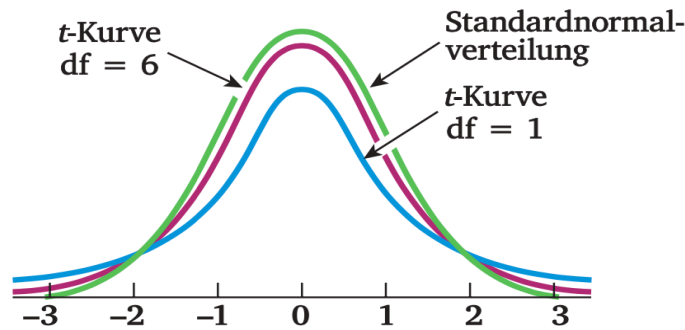


Abbildung 5.2: <https://matheguru.com/stochastik/t-verteilung-students-t-verteilung.html>

- Bei Stichproben $n > 30$ kann die unbekannte Standardabweichung σ durch die Stichprobenstandardabweichung s ersetzt werden und dann ein Intervall für den Erwartungswert bei bekannter Varianz bestimmt werden

Beispiel

Wir gehen nun bei den 12 Rapsfeldern von einer unbekannten Varianz aus (dies ist meist realistischer) und bestimmen das zweiseitige 0,95-Konfidenzintervall für den Erwartungswert:

LSG.:

$$n = 12 \quad \bar{x} = 35,283 \quad \alpha = 0,05 \quad t_{n-1;1-\frac{\alpha}{2}} = t_{11;0,975} = 2,201 \quad s^2 = 3,4288$$

Damit ergibt sich das Konfidenzintervall zu:

$$\begin{aligned} & \left[\bar{x} - t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right] \\ &= \left[35,283 - 2,201 \cdot \frac{1,852}{\sqrt{12}}; 35,283 + 2,201 \cdot \frac{1,852}{\sqrt{12}} \right] \\ &= [34,11; 36,46] \end{aligned}$$

III: Konfidenzintervall für die Varianz einer Normalverteilung**Herleitung**

- Schätzfunktion für die Varianz ist $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

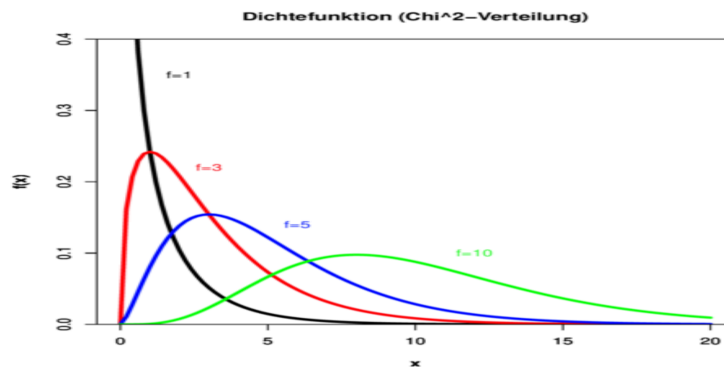
- $Z := (n-1) \cdot \frac{S^2}{\sigma^2}$ ist χ^2 -verteilt
- Gesucht: Intervall mit $P(c_1 \leq Z \leq c_2) = 1 - \alpha$
- Bestimme $c_1 = \chi_{n-1; \frac{\alpha}{2}}^2$ und $c_2 = \chi_{n-1; 1-\frac{\alpha}{2}}^2$ die Quantile der χ^2 -Verteilung
- $c_1 \leq (n-1) \frac{S^2}{\sigma^2} \leq c_2 \implies (n-1) \frac{S^2}{c_2} \leq \sigma^2 \leq (n-1) \frac{S^2}{c_1}$
 $\implies (n-1) \frac{s^2}{c_2} \leq \sigma^2 \leq (n-1) \frac{s^2}{c_1}$

Zusammenfassung

Bestimme \bar{x} und s aus der Stichprobe

zweiseitiges Konfidenzintervall	$\left[\frac{(n-1) \cdot s^2}{\chi_{n-1; 1-\frac{\alpha}{2}}^2}, \frac{(n-1) \cdot s^2}{\chi_{n-1; \frac{\alpha}{2}}^2} \right]$
einseitiges unteres Konfidenzint.	$\left[0, \frac{(n-1) \cdot s^2}{\chi_{n-1; \alpha}^2} \right]$
einseitiges oberes Konfidenzint.	$\left[\frac{(n-1) \cdot s^2}{\chi_{n-1; 1-\alpha}^2}, \infty \right)$

Bemerkung



Beispiel

Wir suchen nun für das Beispiel der 12 Rapsfelder ein zweiseitiges 0,95-Konfidenzintervall für die unbekannte Varianz:

Lösung:

$$n = 12 \quad \bar{x} = 35,283 \quad \alpha = 0,05 \quad \chi_{n-1}^2 \left(1 - \frac{\alpha}{2}\right) = \chi_{11}^2(0,975) = 21,92 \quad \chi_{11}^2(0,025) = 3,816 \quad \hat{\sigma}^2 = 3,4288$$

Damit ergibt sich das Konfidenzintervall zu:

$$\left[\frac{(n-1) \cdot \hat{\sigma}^2}{\chi_{n-1}^2 \left(1 - \frac{\alpha}{2}\right)}; \frac{(n-1) \cdot \hat{\sigma}^2}{\chi_{n-1}^2 \left(\frac{\alpha}{2}\right)} \right] = \left[\frac{11 \cdot 3,4288}{21,92}; \frac{11 \cdot 3,4288}{3,816} \right] = [1,721; 9,884]$$

IV: Konfidenzintervall für die Standardabweichung einer Normalverteilung

Hierzu darf man aus den vorigen Intervallgrenzen für die Varianz die Wurzel ziehen.

zweiseitiges Konfidenzintervall	$\left[\sqrt{\frac{(n-1) \cdot s^2}{\chi_{n-1;1-\frac{\alpha}{2}}^2}}, \sqrt{\frac{(n-1) \cdot s^2}{\chi_{n-1;\frac{\alpha}{2}}^2}} \right]$
einseitiges unteres Konfidenzint.	$\left[0, \sqrt{\frac{(n-1) \cdot s^2}{\chi_{n-1;\alpha}^2}} \right]$
einseitiges oberes Konfidenzint.	$\left[\sqrt{\frac{(n-1) \cdot s^2}{\chi_{n-1;1-\alpha}^2}}, \infty \right)$

Beispiel

Aufbauend auf dem letzten Beispiel erhält man das 0,95-Konfidenzintervall für die Standardabweichung aus:

$$[\sqrt{1,721}; \sqrt{9,884}] = [1,312; 3,144]$$

5.3 Hypothesentest

5.3.1 Unterschied Hypothesentest zu Parameterschätzung

Bei Hypothesentests:

- Fragestellung ist etwas anders als bei Schätzungen.
- Statt einen unbekannten Parameter zu schätzen, hat man hier eine Vermutung (Hypothese) über einen unbekannten Parameter.
- der Test dient dazu, anhand vorliegender Beobachtungen eine begründete Entscheidung über die Gültigkeit oder Ungültigkeit der Hypothese zu treffen.

Schätzungen	Hypothesentest
Schätzen eines unbekannten Parameters	es liegt eine Vermutung über einen Parameter der Verteilung vor
Test gibt Schätzwert bzw. Konfidenzintervall	Test gibt Entscheidung über Gültigkeit der Vermutung

5.3.2 Definition Hypotestentests

Beschreibung

Unter einer **Hypothese** versteht man irgendwelche Annahmen, Vermutungen oder Behauptungen über die Wahrscheinlichkeitsverteilung einer Zufallsvariablen oder Grundgesamtheit und deren Parameter.

Ein **Parametertest** ist ein statistisches Prüfverfahren für einen unbekannten Parameter in der Wahrscheinlichkeitsverteilung einer Zufallsvariablen oder Grundgesamtheit, wobei die Art der Verteilung (d.h. der Verteilungstyp, wie z.B. Binomialverteilung etc) als bekannt vorausgesetzt wird. Ein solcher Test dient der Überprüfung einer Hypothese. Die zu überprüfende Hypothese wird meist als **Nullhypothese** H_0 bezeichnet. Ihr wird oft eine Alternativhypothese H_1 gegenübergestellt. Es ist dann das Ziel eines Parametertests eine Entscheidung darüber zu ermöglichen, ob man die Nullhypothese H_0 beibehalten kann, da die Auswertung des verwendeten Stichprobenmaterials in keinem Widerspruch zur Nullhypothese steht oder ob man sie zugunsten der Alternativhypothese H_1 verwirft. Mit einem Parametertest kann also über Ablehnung oder Beibehaltung einer aufgestellten Hypothese entschieden werden.

Beispiel

Ein Großhändler bestellt beim Hersteller einen größeren Posten eines elektronischen Bauelements und vereinbart dabei, daß die Ware einen maximalen Ausschußanteil von $p_0 = 1\%$ enthalten darf. Bei der Anlieferung wird dies mit einem Test geprüft. Der Großhändler wird daher die Nullhypothese

$$H_0 : p \leq p_0 = 1\%$$

gegen die Alternativhypothese

$$H_1 : p > p_0 = 1\%$$

testen. Dies ist ein einseitiger Paramtertest, da hier die Alternativhypothese nur Werte $p > p_0$ zuläßt.

Meist lässt sich nicht mit Sicherheit sagen, ob eine Hypothese stimmt oder nicht. Man versucht daher, die Wahrscheinlichkeiten für Fehlentscheidungen zu kontrollieren.

Vergleich mit Gerichtsverfahren (Wikipedia: Statistischer Test)

Ein Hypothesentest lässt sich im Prinzip mit einem Gerichtsverfahren vergleichen.

Das Verfahren hat (meistens) den Zweck, festzustellen, ob es ausreichend Beweise gibt, den Angeklagten zu verurteilen.

Es wird dabei immer von der Unschuld eines Verdächtigen ausgegangen, und solange große Zweifel an den Belegen für ein tatsächliches Vergehen bestehen, wird ein Angeklagter freigesprochen.

Nur wenn die Indizien für die Schuld eines Angeklagten deutlich überwiegen, kommt es zu einer Verurteilung.

Es gibt demnach zu Beginn des Verfahrens die beiden Hypothesen

H_0 : der Verdächtige ist unschuldig

und

H_1 : der Verdächtige ist schuldig

Erstere nennt man Nullhypothese, von ihr wird vorläufig ausgegangen. Die zweite nennt man Alternativhypothese. Sie ist diejenige, die zu beweisen versucht wird.

Um einen Unschuldigen nicht zu leicht zu verurteilen, wird die Hypothese der Unschuld erst dann verworfen, wenn ein Irrtum sehr unwahrscheinlich ist.

Man spricht auch davon, die Wahrscheinlichkeit für einen Fehler 1. Art (also das Verurteilen eines Unschuldigen) zu kontrollieren.

Naturgemäß wird durch dieses Vorgehen die Wahrscheinlichkeit für einen Fehler 2. Art (also das Freisprechen eines Schuldigen) groß.

Aufgrund der stochastischen Struktur des Testproblems lassen sich wie in einem Gerichtsverfahren Fehlentscheidungen grundsätzlich nicht vermeiden. Man versucht in der Statistik allerdings optimale Tests zu konstruieren, die die Fehlerwahrscheinlichkeiten minimieren.

5.3.3 Aufbau von Hypothesentests

- Nullhypothese H_0 aufstellen
- Gegenteil dieser Aussage Alternativhypothese H_1 notieren
- Tipp zum Aufstellen der Hypothesen:
 - Was gezeigt bzw. bewiesen werden soll, gehört in die Alternativhypothese
 - Das Gleichheitszeichen gehört immer in die Nullhypothese
 - Beim Aufstellen der Nullhypothese geht man davon aus : alles bleibt beim Alten, nichts hat sich geändert.
- Mögliche Testprobleme (Θ ist der unbekannte Parameter; Θ_0 steht für den vermuteten Wert)

- zweiseitig:
 $H_0 : \Theta = \Theta_0, H_1 : \Theta \neq \Theta_0$
- rechtseinseitig:
 $H_0 : \Theta \leq \Theta_0, H_1 : \Theta > \Theta_0$
- links einseitig:
 $H_0 : \Theta \geq \Theta_0, H_1 : \Theta < \Theta_0$
- Wahl des geeigneten Tests (Testgröße)
- Bestimmung des kritischen Bereiches zum Signifikanzniveau α , das vor Realisierung der Stichprobe feststehen muss. Typische Werte für α sind 0,1; 0,05; 0,01
 Mit dem Signifikanzniveau bestimmt man den Ablehnungsbereich
- Berechnen von Testgrößen
 Dies ist eine Funktion der Stichprobenvariablen, die sensibel für das Testproblem ist
- Ergebnis
 Der Ablehnungsbereich umfasst jene Werte der Testgröße, die für die Alternative H_1 sprechen und mit einer Wahrscheinlichkeit $\leq \alpha$ auftreten. α das Signifikanzniveau.
 Bemerkung:
 Liegt ein Versuchsergebnis nun im Annahmehereich, wird dadurch nicht die Hypothese bestätigt, sondern man entscheidet sich durch die vorher festgelegte Entscheidungsregel, sie weiter als richtig anzusehen.
 Die Annahme der Nullhypothese führt immer zur Ablehnung der Alternativhypothese, ist aber kein Beweis dafür, dass die Nullhypothese stimmt. Die Ablehnung der Nullhypothese führt zur Annahme der Alternativhypothese.

Einfaches Beispiel

Ein einfaches Beispiel ist der Münzwurf.

Hier geht man davon aus, dass beide Ereignisse Wappen und Zahl gleichwahrscheinlich sind mit der Erfolgswahrscheinlichkeit $p = 0,5$.

Die Nullhypothese, die bestätigt werden soll: $H_0 : p = 0,5$

Alternativhypothese $H_1 : p \neq 0,5$

Test: Es soll nun 36-mal geworfen werden

X: Anzahl der geworfenen Wappen

Signifikanzniveau auf $\alpha = 5\% = 0,05$ festgelegt.

Der Erwartungswert ist $\mu = n \cdot p = 36 \cdot 0,5 = 18$.

Die Standardabweichung $\sigma = \sqrt{np(1-p)} = \sqrt{36 \cdot 0,25} = \sqrt{9} = 3$

Um eine 95-prozentige Wahrscheinlichkeit abzudecken, findet man in Tabellen für die σ -Umgebung einen Wert für $z = 1,96$.

Das heißt, man kann, nachdem man die Umgebung mit $\mu - 1,96 \cdot 3$ und $\mu + 1,96 \cdot 3$, also $X = 12,12$ und $X = 23,88$, festgelegt hat, die Entscheidungsregel aufstellen:

Verwirf die Annahme, dass die Erfolgswahrscheinlichkeit $p = 0,5$ ist, wenn die Anzahl der Wappen $X < 13$ oder $X > 23$ ist.

5.3.4 (Mögliche) Fehlentscheidungen

Auch wenn es wünschenswert ist, dass der Test aufgrund der vorliegenden Daten richtig entscheidet, besteht die Möglichkeit von Fehlentscheidungen.

Am Ende eines Parametertests ist stets eine Entscheidung zu treffen. Sie kann zugunsten der Nullhypothese oder der Alternativhypothese ausfallen. In beiden Fällen werden Rückschlüsse von einer Stichprobe auf die entsprechende Grundgesamtheit gezogen. Dabei ist zu Bedenken, dass es keine absolut sicheren Rückschlüsse geben kann. Bei einer Testentscheidung besteht immer die Möglichkeit eines Irrtums. D.h. Bei jeder Testentscheidung besteht die Möglichkeit dass die getroffene Entscheidung falsch ist. Dabei gibt es zwei Arten von Fehlern zu unterscheiden:

Definitionen:

1. Ein Fehler 1. Art
liegt vor, wenn eine an sich richtige Nullhypothese H_0 abgelehnt wird. Die Wahrscheinlichkeit für einen Fehler 1. Art wird mit α bezeichnet
2. Ein Fehler 2. Art
wird begangen, wenn eine an sich falsche Nullhypothese H_0 beibehalten wird. Die Wahrscheinlichkeit für einen Fehler 2. Art wird mit β bezeichnet

5.3.5 Gauß-Test: Test für μ eines normalverteilten Merkmals bei bekanntem σ

- Formuliere die Hypothesen
 1. $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$
 2. $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$ linksseitig
 3. $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$ rechtsseitig
- Wähle ein Signifikanzniveau α (z.B.: 0.1, 0.05, 0.01)
- Ziehe eine Stichprobe vom Umfang n , berechne \bar{x} und den dazugehörigen standardisierten Prüfwert

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$
- Bestimme das entsprechende Quantil der Standardnormalverteilung:
 1. $z_{1-\alpha/2}$
 2. $z_{1-\alpha}$

3. $z_{1-\alpha}$

- Verwerfe H_0 , falls

1. $|z| > z_{1-\alpha/2}$

2. $z < -z_{1-\alpha}$

3. $z > z_{1-\alpha}$

Beispiele

1. (YouTube: Daniel Jung) Mindestens 60% der Studenten lieben Mathe.
 $H_0 : p \geq 0,6$ gegen $H_1 : p < 0,6$

2. 30% aller Studenten lieben Mathe!

- Prüfe ob sich die Anzahl verringert hat: $H_1 : p < 0,3$ gegen $H_0 : p \geq 0,3$
- Prüfe ob sich die Anzahl vergrößert hat: $H_1 : p > 0,3$ gegen $H_0 : p \leq 0,3$
- Prüfe ob sich die Anzahl geändert hat $H_1 \neq 0,3$ gegen $H_0 = 0,3$

3. (Göllmann: Mathematik für Ingenieure. Springer Verlag)

Zwischen dem Zulieferer für Schaumstoff und einem LKW-Hersteller gibt es einen Liefervertrag, in dem geregelt ist, dass der LKW-Hersteller einen Preisnachlass fordern darf, wenn nachgewiesen werden kann, dass bei einer Lieferung die Schaumstoffpakete weniger als die vereinbarten (und auf die Pakete aufgedruckten) 100 kg wiegen. Da eine Vollerhebung zu teuer und zu aufwendig ist, haben die Parteien vereinbart, aus jeder Lieferung eine Stichprobe vom Umfang 10 zu ziehen, und sie konnten sich auf ein Signifikanzniveau von 5% einigen.

Beide Parteien gehen davon aus, dass das Gewicht der Schaumstoffpakete normalverteilt ist, wobei generell eine Varianz von $\sigma^2 = 4$ vorliegt.

Bei der letzten Lieferung ergab die Stichprobe folgende Werte in kg:

101, 98, 96, 102, 99, 101, 102, 103, 96, 97

- a) Der LKW-Hersteller will nun wissen, ob er einen Preisnachlass fordern darf.
 $H_0 : \mu \geq 100$ gegen $H_1 : \mu < 100$

D.h. Lehne H_0 ab, falls $z < -z_{1-\alpha}$

$$\mu_0 = 100$$

$$\sigma = 2$$

$$n = 10 \quad \bar{x} = 99,5$$

$$z = \frac{99,5 - 100}{2/\sqrt{10}} = -\frac{0,5 \cdot \sqrt{10}}{2} \approx -0,79$$

$$-z_{1-\alpha} = -0,8289$$

aber $z = -0,79$ ist nicht kleiner als $-z_{1-\alpha} = -0,8289$

Da das nicht stimmt, kann der LKW-Hersteller die Nullhypothese nicht ablehnen.

- b) Nun möchte der Zulieferer nachweisen, dass er nicht nur 100 kg Schaumstoff verpackt hat, sondern sogar mehr als das. Also

$$H_0 : \mu \leq 100 \text{ gegen } H_1 : \mu > 100$$

Dies stimmt ebenfalls nicht. Folglich kann der Zulieferer seine Behauptung ebenfalls nicht belegen.

5.3.6 t -Test: Test für μ eines normalverteilten Merkmals bei unbekanntem σ

- Formuliere die Hypothesen
 1. $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$
 2. $H_0 : \mu \geq \mu_0$ gegen $H_1 : \mu < \mu_0$ linksseitig
 3. $H_0 : \mu \leq \mu_0$ gegen $H_1 : \mu > \mu_0$ rechtsseitig
- Wähle ein Signifikanzniveau α (z.B.: 0.1, 0.05, 0.01)
- Ziehe eine Stichprobe vom Umfang n , berechne \bar{x} und s und den dazugehörigen standardisierten Prüfwert

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$
- Bestimme das entsprechende Quantil der t-Verteilung:
 1. $t_{n-1;1-\alpha/2}$
 2. $t_{n-1;1-\alpha}$
 3. $t_{n-1;1-\alpha}$
- Verwerfe H_0 , falls
 1. $|t| > t_{n-1;1-\alpha/2}$
 2. $t < -t_{n-1;1-\alpha}$
 3. $t > t_{n-1;1-\alpha}$

5.3.7 χ^2 -Test: Test für σ^2 eines normalverteilten Merkmals

- Formuliere die Hypothesen
 1. $H_0 : \sigma^2 = \sigma_0^2$ gegen $H_1 : \sigma^2 \neq \sigma_0^2$
 2. $H_0 : \sigma^2 \geq \sigma_0^2$ gegen $H_1 : \sigma^2 < \sigma_0^2$ linksseitig
 3. $H_0 : \sigma^2 \leq \sigma_0^2$ gegen $H_1 : \sigma^2 > \sigma_0^2$ rechtsseitig
- Wähle ein Signifikanzniveau α (z.B.: 0.1, 0.05, 0.01)

- Ziehe eine Stichprobe vom Umfang n , berechne s^2 und den dazugehörigen standardisierten Prüfwert

$$y = (n-1) \frac{s^2}{\sigma_0^2}$$
- Bestimme das entsprechende Quantil der χ^2 -Verteilung:
 1. $\chi_{n-1;\alpha/2}^2$ und $\chi_{n-1;1-\alpha/2}^2$
 2. $\chi_{n-1;\alpha}^2$
 3. $\chi_{n-1;1-\alpha}^2$
- Verwerfe H_0 , falls
 1. $y < \chi_{n-1;\alpha/2}^2$ oder $y > \chi_{n-1;1-\alpha/2}^2$
 2. $y < \chi_{n-1;\alpha}^2$
 3. $y > \chi_{n-1;1-\alpha}^2$

5.3.8 Binomialtest

Test für p eines binomialverteilten Merkmals

Der Binomialtest wird verwendet, um Hypothesen über die Wahrscheinlichkeit p der Binomialverteilung zu prüfen.

- Formuliere die Hypothesen
 1. $H_0 : p = p_0$ gegen $H_1 : p \neq p_0$ zweiseitig
 2. $H_0 : p \geq p_0$ gegen $H_1 : p < p_0$ linksseitig
 3. $H_0 : p \leq p_0$ gegen $H_1 : p > p_0$ rechtsseitig
- Wähle ein Signifikanzniveau α (z.B.: 0.1, 0.05, 0.01)
- Ziehe eine Stichprobe vom Umfang n , und bestimme N =Anzahl eingetretener Ereignisse
- Bestimme die entsprechende Entscheidungsregel
 1. Grösstes k_u mit $P(X \leq k_u) \leq \frac{\alpha}{2}$ und kleinstes k_o mit $P(X \geq k_o) \leq \frac{\alpha}{2}$
 2. Grösstes k_u mit $P(X \leq k_u) \leq \alpha$
 3. Kleinstes k_o mit $P(X \geq k_o) \leq \alpha$
- Verwerfe H_0 , falls
 1. $N \in \{0, \dots, k_u\} \cup \{k_o, \dots, n\}$
 2. $N \in \{0, \dots, k_u\}$
 3. $N \in \{k_o, \dots, n\}$

Anwendungsbeispiel

Sie möchten anhand eines Signifikanzniveaus von $\alpha = 0,05$ nachweisen, dass der Anteil an unzufriedenen Kunden beim Unternehmen SPEEDTEL geringer als 10% ist. Dazu fragen Sie $n=8$ Personen, von denen zwei unzufrieden sind. Die Zufallsvariable X zähle die unzufriedenen Kunden

1. $H_0 : p \geq 0,1$ gegen $H_1 : p < 0,1$ linksseitig
2. Grösstes k_u mit $P(X \leq k_u) \leq 0,05 : k_u = 0$
3. Verwerfe H_0 falls $N=0$. Da $N=2$ wird H_0 nicht verworfen

Wichtig:

Bei linksseitigen Tests rundet man den Annahmereich auf die nächste ganze Zahl auf und bei rechtsseitigen ab. Bei zweiseitigen Hypothesentests rundet man immer nach innen, also macht man den Bereich eher kleiner.

5.3.9 Anwendungen

Süddeutsche Zeitung 20. März 2019



Wissen > Mathematik > Statistik: P-Wert und Signifikanz in der Kritik

20. März 2019, 19:06 Uhr Statistik

Signifikanter Unfug

Die statistische Signifikanz, gemessen mit dem sogenannten p-Wert, hat in der Wissenschaft eine geradezu götzenhafte Bedeutung erlangt. 800 Forscher beklagen Fehler und fordern ein Umdenken.

Von Patrick Hübner

Ist eine fragwürdige Münze gezinkt oder nicht? Das ist ein klarer Fall für Empirie. Man plane ein Experiment, welches die Hypothese "gezinkt" belegt oder widerlegt. In diesem Fall würde man die Münze wohl mehrmals werfen und notieren, wie oft sie Kopf oder Zahl zeigt.

Nehmen wir an, zwei Forschergruppen gehen ans Werk. Die erste Gruppe wirft die Münze zwanzigmal. Sie fällt zwölfmal auf Kopf. Ist sie also gezinkt? Die Intuition sagt ebenso wie die Formelwerke der Statistik: Nein, die Daten dieses Experiments sind nicht signifikant. Sie sind mit der Nullhypothese konsistent, also der These, dass das Geldstück *nicht* gezinkt ist. Die zweite Forschergruppe wirft die Münze 200-mal und sie fällt 120-mal auf Kopf. Nun ist laut Statistik evident: Mit der Münze ist etwas faul. Und schon gibt es einen herzhaften wissenschaftlichen Diskurs. Zwei Studien kommen zu widersprüchlichen Ergebnissen! Und das, obwohl die Münze in beiden Fällen mit gleicher relativer Häufigkeit Kopf zeigte.

Ist ein derart vereinfachtes Beispiel geeignet, um einen Missstand in der echten Wissenschaft zu illustrieren? Leider ja, wie ein Kommentar in der aktuellen Ausgabe der Zeitschrift *Nature* zeigt. Mehr als 800 Forscher haben den Appell unterschrieben. Die Verfasser, der Zoologe Valentin Amrhein von der Uni Basel sowie die amerikanischen Medizin-Statistiker Sander Greenland und Blake McShane, prangern den Missbrauch statistischer Kenngrößen an, insbesondere das Prinzip der "Signifikanz". Ihre Empfehlung lautet, die Signifikanz komplett abzuschaffen und die Ergebnisse von Experimenten nicht mehr in einfache Ja-Nein-Schemen zu pressen.

Fehlende statistische Signifikanz ist kein Gegenbeweis

Die Autoren zitieren ein frappierendes Beispiel aus der Medizin, das große Ähnlichkeit mit dem eben genannten Münzwurf-Beispiel hat. In zwei Studien wurden Hinweise auf schädliche Nebenwirkungen eines Medikaments überprüft. Es zeigte sich, dass die Patienten beide Male 20 Prozent häufiger an Nebenwirkungen erkrankten. Aber in einem Fall (den die Pharmaindustrie bevorzugte) blieb die Kenngröße für Signifikanz unter einer kritischen Schwelle. Die Studienautoren zogen daraus den Schluss: Seht her, es gibt keine Nebenwirkungen.



Statistische Fundierung zur Aussage:

X .. zähle Anzahl Kopf beim Münzwurf

$$H_0 : p = \frac{1}{2}$$

$$H_1 : p \neq \frac{1}{2}$$

Signifikanzniveau $\alpha = 1\%$

Es handelt sich um einen zweiseitigen Binomialtest

1. $n=20$; $N=12$

Größtes k_u mit $P(X \leq k_u) \leq 0,005 \implies k_u = 3$

Kleinstes k_o mit $P(X \geq k_o) \leq 0,005 \implies k_o = 16$

Verwerfe H_0 , falls $N \in \{0, \dots, 3\} \cup \{16, \dots, 20\}$

H_0 bleibt

2. $n=200$; $N=120$

Größtes k_u mit $P(X \leq k_u) \leq 0,005 \implies k_u = 81$

Kleinstes k_o mit $P(X \geq k_o) \leq 0,005 \implies k_o = 118$

Verwerfe H_0 , falls $N \in \{0, \dots, 81\} \cup \{118, \dots, 200\}$

H_0 verwerfen

Briefzentrum Wien

