

Stochastik

selbstorganisiert erlernen



Ursula Pirkl

Umschlag Vorderseite (Innen)

(unbedruckt)

Hier können Sie noch Vorlagen einfügen

Oberstudienrätin Ursula Pirkl
Stochastik
selbstorganisiert lernen



Kapitel

Grundbegriffe

Einführung in die Stochastik

Kapitel 1	
Grundbegriffe	7
Kapitel 2	
Baumdiagramme und mehrstufige Zufallsversuche	13
Kapitel 3	
Vierfeldertafel	23
Kapitel 4	
Kombinatorische Abzählverfahren	35
Bedingte Wahrscheinlichkeiten	
Kapitel 5	
Einführung in den Begriff der bedingten Wahrscheinlichkeit	47
Kapitel 6	
Definition der Wahrscheinlichkeitsverteilung	61
Wahrscheinlichkeitsverteilung	
Kapitel 7	
Erwartungswert	63
Kapitel 8	
Variation und Standardabweichung	69
Kapitel 9	
Normalverteilung	75
Hypothesentests	
Kapitel 10	
Zweiseitiger Signifikanztest	97
Kapitel 11	
Einseitiger Signifikanztest – Fehler beim Testen	109
Kapitel 12	
Vertrauensintervall	135
Die Betrachtungen zur Normalverteilung	
Kapitel 14	
Anwendung der Normalverteilung	149

VORSCHAU
 Lehrerselbstverlag
 schuldruckportal.de

Gesamtwerk selbstorganisiert erlernen (Bestellnummer 02-032-278)

Alle Rechte vorbehalten. All rights reserved.
 Nachdruck, auch auszugsweise, vorbehaltlich der Rechte,
 aus §§ 53, 54 UrhG ergeben, nicht gestattet.
 LehrerselbstVerlag
 Lehrende & Freunde GmbH, Koblenz (Germany)
 www.lehrerselbstverlag.de
 www.f-druck.de

Oberstudienrätin Ursula P...

Stochastik
 selbstorganisiert erlernen

VORSCHAU
 Lehrerselbstverlag
 schuldruckportal.de

Mathematik

Bestellnummer 02-032-278



VORSCHAU
 LehrerselbstVerlag
 schuldruckportal.de

VORSCHAU
 LehrerselbstVerlag
 schuldruckportal.de

Vorwort	5
Einführung in die Stochastik	
Kapitel 1 Grundbegriffe	7
Kapitel 2 Baumdiagramme und mehrstufige Zufallsexperimente	13
Kapitel 3 Vierfeldertafel	23
Kapitel 4 Kombinatorik und Abszissenverfahren	35
Bedingte Wahrscheinlichkeiten	
Kapitel 5 Definition und Begriff der bedingten Wahrscheinlichkeit	47
Kapitel 6 Definition der Wahrscheinlichkeitsverteilung	61
Wahrscheinlichkeitsverteilung	
Kapitel 7 Erwartungswert	63
Kapitel 8 Varianz und Standardabweichung	69
Kapitel 9 Binomialverteilung	75
Kapitel 10 Zweiseitiger Signifikanztest	97
Kapitel 11 Einseitiger Signifikanztest – Fehler beim Testen	109
Kapitel 12 Vertrauensintervall	135
Kapitel 13 Zentraler Grenzwertsatz und Anwendungen zur Normalverteilung	139
Kapitel 14 Anwendung der Normalverteilung	149

Alle Rechte vorbehalten. All rights reserved.
 Nachdruck, auch auszugsweise, vorbehaltlich der Rechte,
 die sich aus §§ 53 ff. UrhG nicht gestattet.

LehrerselbstVerlag
 Solingen, Germany 2014

www.f-druck.de

Vorwort

Wie schon in den bereits erschienenen Arbeitsbüchern zum Thema Integralrechnung und Lineare Algebra beruht die Erarbeitung der Zusammenhänge auch im Themenbereich Stochastik auf selbstorganisierten Lernformen. Die Lernenden werden anhand von geeigneten Aufgaben- und Fragestellungen an den Stoff herangeführt, indem Erläuterungen sowie Erklärungen frei formuliert oder entsprechende Lückentexte ausgefüllt und Berechnungen selbst durchgeführt bzw. ergänzt werden. Durch das Lesen und Erfassen der Texte in allen Details, das Formulieren von Erläuterungen und das schrittweise Beschreiben von Lösungswegen werden daher nicht nur fachsystematischen auch Kompetenzen im Umgang mit der Fachterminologie erworben. Für einen mittleren bis hohen sprachlichen Anteil der Aufgabenstellungen ist der Erwerb dieser Kompetenzen erforderlich.

Zielgruppe

Die selbstorganisierte Erarbeitung wird hier anhand der Erarbeitung mathematischer Zusammenhänge, die dem die Lernenden nach und nach auch komplexer zusammenhängen zu erarbeiten, was in einer Gesellschaft, in der kein langes Leben in einem immer höheren Stellenwert erlangt, die wachsende Bildungserwartung erhält. Die erworbenen Kompetenzen stellen ferner eine grundlegende Notwendigkeit für die erfolgreiche Bewältigung eines höheren Studiums dar. Dieses Arbeitsbuch ist für Schülerinnen und Schüler einer gymnasialen Oberstufe geeignet, sondern auch für Studienanfänger mit Fachhochschulreife, die diesen Themenbereich im Rahmen des Mathematikunterrichts der Schule nicht behandelt haben, jedoch einen Studiengang wählen, in dem Kenntnisse im Bereich der Wahrscheinlichkeitsrechnung vorausgesetzt werden. Die Lernenden können Lernfortschritt selbst überprüfen und sich bei Bedarf ab September 2014 über das Internet informieren.

Handhabung des Buchs

Die Inhalte des Arbeitsbuchs stellen die grundlegenden Zusammenhänge zu einzelnen Themengebieten der Stochastik zur Verfügung und sind damit als Basis für die Bewältigung der Aufgabenstellungen der Abiturprüfung zu sehen. Im Folgenden werden komplexere Aufgabenstellungen mit anwendenden Lösungswegen und methodische Konzepte beschrieben, die die Lernenden und Schüler mit Zugang zum Fach Stochastik mit dem Themenbereich Stochastik gut zurecht kommen, zumal hinsichtlich der Handhabung von Rechenkalkülen fehlende Kenntnisse hier nur eine untergeordnete Rolle spielen.

Dieses Buch ist für Lernende mit und ohne Vorkenntnisse in der Wahrscheinlichkeitsrechnung geeignet. Die Lernenden werden Begriffe und Arbeitsmethoden, die bereits

in den Bildungsstandards der Sekundarstufe I verankert sind, im ersten Kapitel wiederholt vertieft werden. Verweise auf die Grundlagen, Erklärungen und erweiternde Betrachtungen zu einzelnen Themenbereichen können an den entsprechenden Stellen am Ende eines Kapitels vertieft werden. Anregungen und Verbesserungswünsche werden gerne entgegengenommen und können an service@lehrerselbstverlag.de oder per Adresse service@lehrerselbstverlag.de zugesandt werden.

Die einzelnen Kapitel des Arbeitsbuchs aufeinander aufbauen und Querverweise auf Inhalte aus vorangegangenen Kapiteln enthalten sind, dies notwendig, dass die einzelnen Themen in der vorgegebenen Reihenfolge bearbeitet werden, wobei ergänzende oder erweiternde Betrachtungen stets optional sind.

Kapitel 1 bis 3

Im Kapitel 1 werden zunächst Begriffe und Definitionen zur Verfügung gestellt und elementare Arbeitsweisen, wie beispielsweise der Umgang mit Baumdiagrammen, wiederholt und vertieft. Zusätzlich, insbesondere jedoch im Hinblick auf den allgemeinen Additionssatz und die gemeinsame Verwendung bei der Behandlung von bedingten Wahrscheinlichkeiten, erfolgt hier anhand von Vierfeldertafeln eine ausführliche Betrachtung abhängiger und unabhängiger Zufallsversuche.

Kapitel 4

Die Betrachtungen zu kombinatorischen Abzählverfahren sind nur auf elementare Problemstellungen beschränkt. Vertiefende und komplexe Aufgabenstellungen können ergänzend eingefügt werden. Eine Herleitung für die formelmäßige Betrachtung des ungeordneten Ziehens ohne Zurücklegen, was in den Unterlagen auch als „Lottoproblem“ bezeichnet wird, erfolgt über ein anschauliches Beispiel unter Einbeziehung des Pascalschen Dreiecks. Der Vertiefung des Pascalschen Dreiecks an dieser Stelle in den Unterlagen kommt eine vorbereitende Aufgabe zur Binomialverteilung in Kapitel 9 zu.

Die Komplexität der Aufgabenstellungen zu bedingten Wahrscheinlichkeiten stellt eine Herausforderung für die Lernenden. In den Unterlagen wird ein Schwerpunkt auf die Erarbeitung der Aufgabentexte gelegt, indem die Textaufgaben auf Standardformulierungen zurückgeführt werden. Die Lösung der Aufgaben erfolgt dann über die Anwendung der beiden möglichen Baumdiagramme und die Vierfeldertafel, die bereits aus Kapitel 1 bekannt ist. Vertiefend werden dann der Satz von Bayes und Vierfeldertafeln ergänzt.

Kapitel 6 bis 8

Die Betrachtungen der Wahrscheinlichkeitsverteilungen werden durch Beispiele zur Erläuterung der Zusammenhänge zwischen den verschiedenen Verteilungen des Erwartungs-

VORSCHAU
Lehrerselbstverlag
schuldruckportal.de

VORSCHAU
Lehrerselbstverlag
schuldruckportal.de

wertes und der Standardabweichung eine tabellarische Form verwendet wird, mit der für diesen Aufgabentyp eine Systematisierung des Lösungswegs erreicht wird.

Kapitel 9

Nachdem der Begriff der Wahrscheinlichkeitsverteilung verankert ist, nimmt die Binomialverteilung nun eine zentrale Rolle ein. Anhand des schon bekannten Pascalschen Dreiecks wird die Formel von Bernoulli beispielorientiert hergeleitet, wobei der Bezug zur Bezeichnung Binomialverteilung für die Lernenden anschaulich ersichtlich wird. Der Schwerpunkt dieses Kapitels liegt dann in der Handhabung entsprechender Tabellen. Auch kann hier die Verwendung von Taschenrechnern einfließen.

Da Erwartungswerte und Sigma-Umgebungen der Binomialverteilung beim Testen von Hypothesen eine hervorgehobene Rolle spielen, wird dieser Bereich ausführlich behandelt. Anhand von exemplarisch ausgewählten Verteilungen wird zum Beispiel das Aussehen eines Säulendiagramms einer Binomialverteilung in Abhängigkeit von n und p verdeutlicht. Auch werden für mehrere Verteilungen die Bedeutungen der Sigma-Umgebungen erarbeitet.

Kapitel 10

Der Einsatz in den meisten von Hypothesen erfordern in meinem Unterricht ausschließlich zweiseitigen Tests direkt im Anschluss an die Bestimmung der Sigma-Umgebung über ein Schülerversuch. Im Rahmen von gegebenen Aufgabenstellungen nimmt jeder Lernende aus dem gleichen Behälter eine Stichprobe vorzunehmen. Ich verwende hier eine Mischung aus Linsen und Sonnenblumenkernen. Langjährige Beobachtungen des Versuchs zeigen, dass bei dieser Mischung bei jeder Lerngruppe Stichproben entstehen, die im Bereich des Erwartungswertes liegen, aber auch jedes Mal Stichproben auftreten, die außerhalb der 3σ-Bereiche liegen. (Fallbeispiel: ...)

... werden, können die ...
lenbeispiel ...
Pers...

... die günstigste Ergebnisse auftreten,
... Voraussetzung für das Verständnis der
... Statistik beim Testen von Hypothesen dar.
... Bewertung des Tests erfolgt zunächst über die An-
... der Behandlung zweiseitiger Tests mit Fehlerbetrach-
... Schwerpunkt bei ...
... stellungen zu Tests ...
... von
... Annahme- und ...
... Hüllen-
... kurven von Binomialverteilung, ...
... einbart,
... dass die G ...
... wählt wer-
... den, da ...
... ste Wert im Ablehnungsbereich
... ange...

Kapitel 11

Aufbauend auf die zweiseitigen Tests werden nun, anhand eines Beispiels aus der Medizin, die einseitigen Tests behandelt. Um die Komplexität der formalen Darstellung zu reduzieren und für die Lernenden leichter zugänglich ein einheitliche Lösungssystematik aufzustellen, wird auf einen tabellarisch dargestellten Lösungsweg hingearbeitet. Hierbei stehen die Formalisierung von Annahme- und Ablehnungsbereichen mithilfe von Hüllkurven der Binomialverteilung im Mittelpunkt. Erfahrungen haben gezeigt, dass die Lernenden, welche die hier vorgeschlagene Vorgehensweise anwenden und die Berechnung für den Alpha- und Betafehler markieren, die Bestimmung der Grenzen der Ablehnungsbereiche sich überbieten.

Ferner stellt die sprachliche Ausformulierung der Nullhypothese und der Alternativhypothese im Kopf der Tabelle eine wesentliche Rolle bei der Formulierung der beiden Fehler und deren Konsequenzen dar. Anhand der vorgegebenen Formulierungshinweise „...“, wird irrtümlich angenommen, dass „...“ kann man Bezug auf die im Kopf der Tabelle angegebenen Hypothesen nehmen und so eine für beliebige Testsituationen systematisch erfassbare Fehleranalyse vornehmen. In der Praxis hat es sich gezeigt, dass die Lernenden, nachdem sie mithilfe der systematischen Vorgehensweise anhand der Tabelle die Zusammenhänge beim Testen erfasst haben, die Aufgabenstellung auch ohne Tabelle richtig bearbeiten.

Betrachtungen der Abhängigkeit der Fehler vom Stichprobenumfang und den Wahrscheinlichkeiten der beiden Hypothesen sowie die Behandlung von Operationscharakteristiken können als erweiternde Aufgabenstellung interpretiert werden.

Kapitel 12

Für Testsituationen, die mit den zur Verfügung gestellten Tabellen zur Binomialverteilung nicht bearbeitet werden können, kommt für näherungsweise zu bestimmende Lösungen die Normalverteilung zum Tragen. Dazu werden die Normalformierungsverfahren selbst und anschließend von weiteren grundlegenden Testsituationen mit der Tabelle zur Normalverteilung erarbeitet. Nebenbei, wie die Bestimmung des Stichprobenumfangs in Abhängigkeit von den gegebenen Fehlern oder die Betrachtung von Intervallen haben hier eine vertiefende bzw. ergänzende Funktion.

Dank

Für Anregungen zur inhaltlichen und formalen Gestaltung danke ich meinem Mentor Gerald Pirkl, meiner langjährigen Freundin und Lehrstuhlleiterin Renate Benz-Heinbücher und meinem Kollegen Dr. Torsten-Karl Stempel von der Hochschule ...

Kapitel 1: Grundbegriffe

Aufgabe 1.1

Lesen Sie den folgenden Text und formulieren bzw. veranschaulichen Sie so gut wie möglich, was man unter den aufgeführten Fachbegriffen versteht.

Wie in jedem anderen Gebiet der Mathematik, hat man sich auch in der Stochastik bzw. Wahrscheinlichkeitsrechnung auf die Verwendung einer Reihe von Fachbegriffen geeinigt. Im Folgenden werden grundlegende Fachbegriffe am Beispiel des Würfels erklärt.



Wirft man einen Würfel, weiß man zwar, dass die Zahlen 1 bis 6 fallen können, da man jedoch nicht vorhersehen kann, welche dieser Zahlen fällt, spricht man von einem **Zufallsversuch**.

Fällt beim Würfeln zufällig die Zahl 4, so nennt man das als **Ergebnis** des Zufallsversuchs, die Menge aller Ergebnisse, hier also die Zahlen 1 bis 6, wird im Allgemeinen als **Ergebnismenge S** bezeichnet.

Ist es beim Würfeln noch ein Spiel, ist es wichtig, dass die geworfene Zahl eine gerade Zahl ist, dann spricht man nicht vom Ergebnis des Zufallsversuchs, sondern man sagt, dass das **Ereignis** „gerade Zahl“ eingetreten ist, wenn eine der Zahlen der Menge A = {2, 4, 6} fällt. Das **Ereignis** ist nicht eingetreten, wenn eine ungerade Zahl aus der Menge B = {1, 3, 5} gefallen ist. Alle Teilmengen der Ergebnismenge S, wie die Mengen A und B sowie beispielsweise auch die Menge der durch 3 teilbaren Zahlen C = {3, 6} werden als **Ereignisse** bezeichnet.

Wenn sich die Ereignisse, wie die geraden Zahlen der Menge A und die ungeraden Zahlen der Menge B, gegenseitig ausschließen, jedoch gemeinsam die gesamte Ergebnismenge S bilden, so bezeichnet man A und B als **Wahrscheinlichkeitsereignisse** A und B und A als **Ereignis** und **Gegenereignis**. Man verwendet in diesem Fall auch die Schreibweise Ereignis A und Gegenereignis \bar{A} bzw. Ereignis B und Gegenereignis \bar{B} .

Zufalls-
experiment: _____

Ergebnismenge S: _____

Ereignis E: _____

Ereignis \bar{E} : _____

Gegenereignis zu A (Augenzahl _____): _____

Aufgabe 1.2

Der Begriff Wahrscheinlichkeit

Bearbeiten Sie die Aufgaben mithilfe der folgenden Definition und Angaben zur Schreibweise und zum Berechnen von Wahrscheinlichkeiten.

Definition Laplace-Versuch

Wenn man mit einem idealen Würfel spielt, weiß man, dass die Wahrscheinlichkeit, eine 1 zu würfeln genauso groß ist wie beispielsweise für eine 4 oder jedes andere Ergebnis der Ergebnismenge, d.h. jedes Ergebnis hat die gleiche Wahrscheinlichkeit. Die Wahrscheinlichkeiten kann man hier aus der Geometrie des Würfels erschließen. Man bezeichnet Zufallsversuche, bei denen diese Merkmale eintreffen als **Laplace-Versuche**.

Schreibweise für Wahrscheinlichkeiten

- Kleines **p** für die Wahrscheinlichkeit für ein Ergebnis
- Die Wahrscheinlichkeit für ein Ereignis kann wie für das Beispiel „Würfeln einer geraden Zahl“ in der Schreibweise P durch ein großes P mit einer ergänzenden Information, um welches Ereignis es sich handelt, dargestellt werden.

Berechnen von Wahrscheinlichkeiten

Berechnen der Wahrscheinlichkeit p eines Ereignisses bei einem Laplace-Versuch: $p = \frac{\text{Anzahl der günstigen Möglichkeiten}}{\text{Anzahl aller Möglichkeiten}}$

Berechnen der Wahrscheinlichkeit **P(E)** eines Ereignisses bei einem Laplace-Versuch: $P(E) = \frac{\text{Anzahl der günstigen Möglichkeiten für E}}{\text{Anzahl aller Möglichkeiten}}$

Berechnen von $P(A \cap B)$ eines Ereignisses bei einem Laplace-Versuch: $P(A \cap B) = \dots$

Dem bisher betrachteten Würfeln gibt es eine Reihe anderer Laplace-Versuche, die einem Laplace-Versuch zugeordnet werden können. Geben Sie jeweils an, ob es sich um einen Laplace-Versuch handelt, und ermitteln Sie, wenn möglich, die gesuchten Wahrscheinlichkeiten für das Ergebnis p des Zufallsversuchs und die Wahrscheinlichkeit $P(E)$ für das jeweils angegebene Ereignis E .

Zufallsversuch	Laplace-Versuch	Wahrscheinlichkeit für ein Ergebnis p	Wahrscheinlichkeit für das Ereignis $P(E)$
Idempotenzgesetz 	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(E) = \dots$

Zufallsversuch	Laplace-Versuch	Wahrscheinlichkeit für ein Ergebnis p	Wahrscheinlichkeit für das Ereignis $P(E)$
Roulette einmal drehen (37 Felder)	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(\text{rot}) = \dots$
Eine Karte aus einem Skatspiel ziehen (32 Karten)	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(\text{Bube}) = \dots$
Glücksrad mit gleichen Sektoren einmal drehen 	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(\text{teilbar durch 3}) = \dots$
Glücksrad mit ungleichen Sektoren einmal drehen 	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(\text{Zahl} > 4) = \dots$
Ziehen einer Kugel aus einer Urne mit 10 verschiedenfarbigen Kugeln 	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(\text{schwarz}) = \dots$
Ziehen einer Kugel aus einer Urne mit 10 verschiedenfarbigen Kugeln 	ja <input type="radio"/> nein <input type="radio"/>	$p = \dots$	$P(\text{schwarz}) = \dots$

Weitere Laplace-Versuche:

Aufgabe 1.3

Wertebereich bei Wahrscheinlichkeiten

Ermitteln Sie den Wertebereich der Wahrscheinlichkeiten, geben Sie die Wertebereiche an und begründen Sie die Merksätze:

- a) $P(\text{eine 1 fällt}) = \dots$
- b) $P(\text{eine der Zahlen 1 bis 3 fällt}) = \dots$
- c) $P(\text{eine der Zahlen 1 oder 2 fällt}) = \dots$
- d) $P(\text{eine der Zahlen 1 bis 4 fällt}) = \dots$
- e) $P(\text{eine der Zahlen 1 bis 5 fällt}) = \dots$
- f) $P(\text{eine der Zahlen 1 bis 6 fällt}) = \dots = \dots$
- g) $P(\text{keine 7 fällt}) = \dots$ (sicheres Ereignis)

Welche Aussagen über die Wahrscheinlichkeiten beim Würfeln kann man schließen:
 Welcher größte Wert, den die Wahrscheinlichkeit für ein Ereignis E annehmen kann, wenn $P(E) = \dots$
 Welcher kleinste Wert, den die Wahrscheinlichkeit für ein Ereignis E annehmen kann, wenn $P(E) = \dots$

Für den Wert der Wahrscheinlichkeit P(E) eines Ereignisses gilt:
 $0 \leq P(E) \leq 1$

Beim Würfeln sind die Ereignisse E: „die geworfene Zahl ist durch 3 teilbar“ und \bar{E} : „die geworfene Zahl ist nicht durch 3 teilbar“ Ereignis und Gegenereignis.

Begründen Sie mit diesem Beispiel, dass gilt: $P(E) + P(\bar{E}) = 1$

Bei einem Element ist die Summe aus der Wahrscheinlichkeit P(E) für ein Ereignis E und der Gegenwahrscheinlichkeit P(\bar{E}) immer 1.
 $P(E) + P(\bar{E}) = 1$
Wichtig für viele Anwendungen ist dieser Zusammenhang in der Form:
 $P(\bar{E}) = 1 - P(E)$

Aufgabe 1.4

Die empirische Wahrscheinlichkeit – Gesetz der großen Zahlen

Die Wahrscheinlichkeit p = 1/6 fällt.



Da hier keine Flächen angegeben sind, kann die Wahrscheinlichkeit, mit der eine Augenzahl angezeigt wird, nicht bestimmt werden. Die Flächen sind für die Möglichkeit der Wahrscheinlichkeit, das Fallen einer bestimmten Augenzahl über die empirischen Daten zu bestimmen.

Ein Experiment mit idealer Würfel wurde in 17 voneinander unabhängigen Zufallsversuchen durchgeführt, oft, geworfen. Das entstandene Zahlenmaterial ist in den Tabellen dargestellt und von Excel als Diagramm ausgewertet.

a) Erläutern Sie anhand der Tabellen, was man bei einem Zufallsversuch mit der absoluten Häufigkeit H(i) und der relativen Häufigkeit h(i) versteht, und geben Sie die Formel an, mit der die relativen Häufigkeiten h(i) berechnet werden. Bezeichnungen: Nummer des Zufallsversuchs Nr., absolute Häufigkeit H(i), Anzahl der Würfe: n

Tabelle absolute Häufigkeiten H(i)

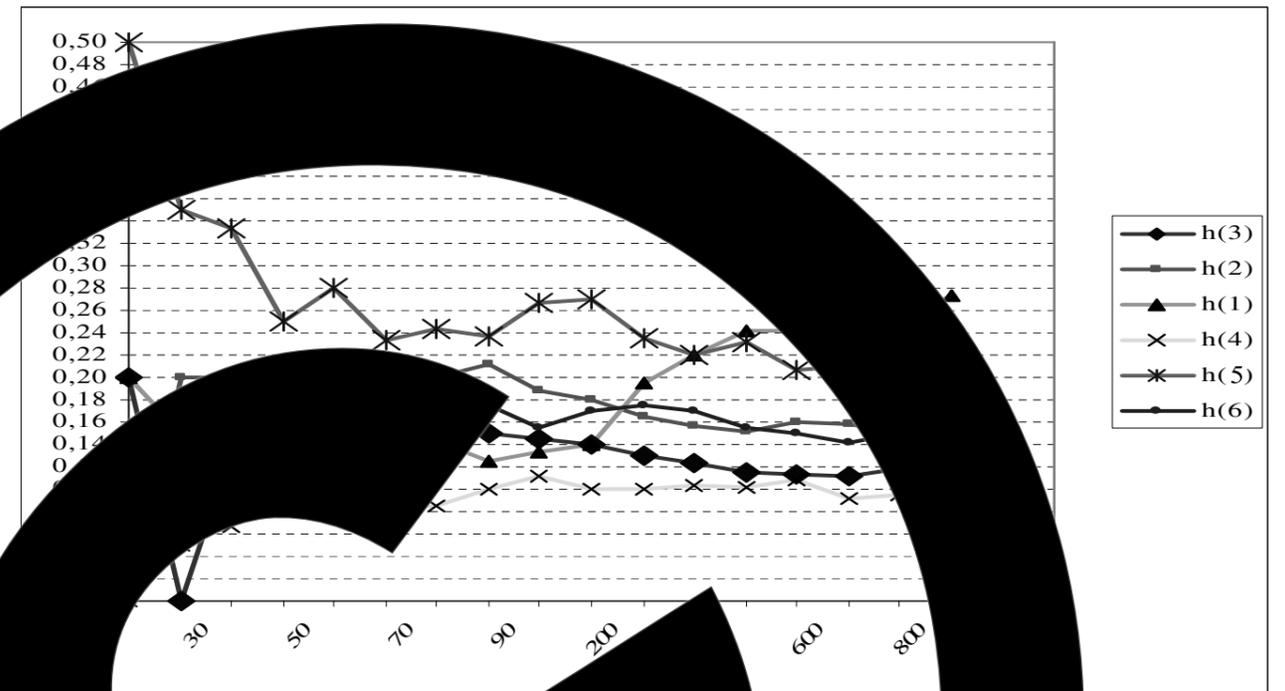
Nr	n	H(1)	H(2)	H(3)	H(4)	H(5)	H(6)
1	10	2	0	2	0	5	1
2	20	0	4	3	1	7	3
3	30	4	6	4	2	10	4
4	40	7	7	6	4	10	6
5	50	8	9	6	4	14	5
6	60	10	12	8	5	14	1
7	70	10	14	10	6	13	7
8	80	12	17	10	8	14	14
9	90	13	17	12	1	24	1
10	100	14	18	14	1	24	1
11	200	26	33	20	7	38	3
12	300	37	47	31	6	68	2
13	400	46	59	41	9	82	2
14	500	57	72	51	10	103	7
15	600	67	95	61	13	108	8
16	700	75	105	67	14	144	11
17	800	93	122	82	16	160	14

Tabelle relative Häufigkeiten h(i)

Nr	n	h(1)	h(2)	h(3)	h(4)	h(5)	h(6)
1	10	0,20	0,00	0,20	0,00	0,50	0,10
2	20	0,00	0,20	0,15	0,05	0,35	0,15
3	30	0,13	0,20	0,13	0,07	0,33	0,13
4	40	0,18	0,18	0,15	0,10	0,25	0,15
5	50	0,16	0,18	0,12	0,08	0,28	0,18
6	60	0,17	0,20	0,13	0,08	0,23	0,18
7	70	0,14	0,20	0,14	0,09	0,24	0,19
8	80	0,15	0,21	0,13	0,10	0,24	0,18
9	90	0,14	0,19	0,13	0,11	0,27	0,16
10	100	0,14	0,18	0,14	0,10	0,27	0,17
11	200	0,13	0,17	0,10	0,10	0,24	0,18
12	300	0,12	0,16	0,10	0,10	0,22	0,17
13	400	0,12	0,15	0,10	0,10	0,23	0,16
14	500	0,11	0,16	0,10	0,11	0,21	0,15
15	600	0,11	0,16	0,10	0,09	0,21	0,14
16	700	0,11	0,15	0,10	0,10	0,21	0,15
17	800	0,12	0,15	0,10	0,10	0,20	0,14

absolute Häufigkeit: $H(i) = n \cdot h(i)$
relative Häufigkeit: $h(i) = \frac{H(i)}{n}$
Formel: $h(i) = \frac{H(i)}{n}$

b) Erläutern Sie die Bedeutung des Diagramms.



c) Beurteilen Sie, ob die 800 Würfe ausreichen, um eine Aussage zur Wahrscheinlichkeit mit machen zu können, mit der die Flächen 1 bis 6 bei dem vorliegenden nicht idealen Würfel angeordnet wurden. Erläutern Sie dabei auch, wie man an einem Diagramm für die relativen Häufigkeiten ablesen könnte, ob man von der Wahrscheinlichkeit, für das Auftreten des Ereignisses berechnen kann.

Das Gesetz der großen Zahlen

Wenn Sie für Aufgabe 1b) die richtige Formel ermittelt haben, ist Ihnen sicherlich aufgefallen, dass die Berechnung der relativen Häufigkeit $h(E)$ für ein Ereignis E der Definition der Wahrscheinlichkeit sehr ähnlich ist. Für eine große Anzahl von Durchführungen des Zufallsversuchs nehmen die Werte der Werte für die relative Häufigkeit ab. Wenn sich der Wert stabilisiert, kann man den für die relative Häufigkeit $h(E)$ ermittelten Wert als Wahrscheinlichkeit $P(E)$ annehmen. Man spricht dann von einer empirischen Wahrscheinlichkeit. Folgendes gilt:

$$h(E) = \frac{\text{absolute Häufigkeit des Ereignisses } E}{\text{Anzahl aller Zufallsversuche } n}$$

$$h(E) \approx P(E) \quad \text{für } n \text{ gilt: } P(E) \approx h(E)$$

Ergänzen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert erlernen



Kapitel

Baumdiagramme und
mögliche Zufallsversuche



Einführung in die Stochastik

Kapitel 1	
Grundbegriffe	7
Kapitel 2	
Baumdiagramme und mehrstufige Zufallsversuche	13
Kapitel 3	
Vierfeldertafel	23
Kapitel 4	
Kombinatorische Abzählverfahren	35
Bedingte Wahrscheinlichkeiten	
Kapitel 5	
Einführung in den Begriff der bedingten Wahrscheinlichkeit	47
Kapitel 6	
Definition der Wahrscheinlichkeitsverteilung	61
Wahrscheinlichkeitsverteilung	
Kapitel 7	
Erwartungswert	63
Kapitel 8	
Variation und Standardabweichung	69
Kapitel 9	
Normalverteilung	75
Hypothesentests	
Kapitel 10	
Zweiseitiger Signifikanztest	97
Kapitel 11	
Einseitiger Signifikanztest – Fisher beim Testen	109
Kapitel 12	
Vertrauensintervall	135
Die Betrachtungen zur Normalverteilung	
Kapitel 14	
Anwendung der Normalverteilung	149

Gesamtwerk selbstorganisiert erlernen (Bestellnummer 278)

Alle Rechte vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte, die aus §§ 53, 54 UrhG ergeben, nicht gestattet.

LehrerSelbstVerlag

LehrerSelbstVerlag GmbH, Koblenz (Germany)

www.lehrerSelbstVerlag.de

www.f-druck.de

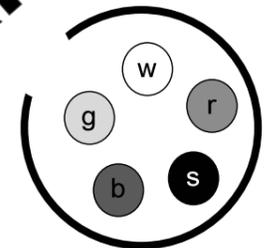
Kapitel 2: Baumdiagramme und mehrstufige Zufallsversuche

Bei den bisherigen Betrachtungen wurden Zufallsversuche untersucht, bei denen beispielsweise nur die Wahrscheinlichkeit ermittelt hat, wenn ein Würfel einmal gewürfelt oder die einzige Kugel aus einer Urne gezogen wurde. Man bezeichnet diese Zufallsversuche als **einstufige Versuche**.

Zufallsversuche können jedoch auch mehrmals durchgeführt werden, die aus mehreren Stufen bestehen. Man bezeichnet diese Zufallsversuche als **mehrstufige Versuche**. Für die Berechnung von Wahrscheinlichkeiten ist dann die Verwendung von Baumdiagrammen hilfreich. Ausgehend von den Kenntnissen zu Baumdiagrammen, die Sie bereits in der Sekundarstufe I erworben haben, werden hier die grundlegenden Regeln wiederholt und

Dazu wird zunächst ein einfacher Zufallsversuch zugrunde gelegt, bei dem aus einer Urne zwei Kugeln gezogen werden. Es wird dabei unterschieden, ob eine gezogene Kugel vor dem nächsten Ziehen zurückgelegt wird oder nicht. Man spricht dann zwischen **mit** und **ohne Zurücklegen**.

In der rechts abgebildeten Urne befinden sich jeweils eine weiße, eine schwarze, eine rote, eine grüne und eine gelbe Kugel. Es sollen zwei Kugeln gezogen werden.



Aus der Urne wird beim ersten Zug die blaue Kugel gezogen. Geben Sie den Inhalt der Urne beim zweiten Ziehen an, wenn

- a) zurückgelegt gezogen wird: _____
- b) ohne Zurücklegen gezogen wird: _____
- c) Vervollständigen Sie:

Beim Ziehen der Urne ändert sich der Inhalt der Urne nicht. Daher ergeben sich bei beiden Ziehungen die gleiche Wahrscheinlichkeit, eine Kugel in einer jeweiligen Farbe zu ziehen.

Beim Ziehen ohne Zurücklegen verändert sich der Inhalt der Urne. Daher ergeben sich bei der ersten und bei der zweiten Ziehung unterschiedliche Wahrscheinlichkeiten, eine Kugel in der jeweiligen Farbe zu ziehen.

Aufgabe

Aus der Urne werden zwei Kugeln gezogen. Zeichnen Sie ein vollständiges Baumdiagramm mit allen Ergebnissen beim Ziehen von zwei Kugeln mit und ohne Zurücklegen aus der oben abgebildeten Urne an. Die Wahrscheinlichkeit für eine Kombination xy aus zwei gezogenen Kugeln wird angegeben, wobei x immer die zuerst gezogene Kugel und y immer die zweite gezogene Kugel bezeichnet. D.h., wenn eine weiße Kugel und dann schwarz gezogen wird gilt: $P(\text{weiß, schwarz})$. Ergänzen Sie sich das Baumdiagramm und die Wahrscheinlichkeiten hinsichtlich der Ziehungen mit und ohne Zurücklegen.

Abb. 2.1 Baumdiagramm für Ziehen mit Zurücklegen

Abb. 2.2 Baumdiagramm für Ziehen ohne Zurücklegen



Wendet man sich für die Berechnung von Wahrscheinlichkeiten an, ergibt sich für die Wahrscheinlichkeit, eine bestimmte Kugelkombination xy zu ziehen, folgende Werte:

Ziehen mit Zurücklegen: $P(xy) = \frac{1}{5} = 0,2$

Ziehen ohne Zurücklegen: $P(xy) = \frac{1}{20} = 0,05$

Aufgabe 2.3

Einfache Additionsregel

Mithilfe der ausführlichen Baumdiagramme und der Kenntnis, dass die Kombination xy beim Ziehen mit Zurücklegen die Wahrscheinlichkeit $p = \frac{1}{25} = 0,04$ hat und beim Ziehen ohne Zurücklegen die Wahrscheinlichkeit $p = \frac{1}{20} = 0,05$ hat, sollen nun Berechnungen für die Aufgabenstellung mit Baumdiagrammen erarbeitet werden. Übertragen Sie die Vorgehensweise der Beispiele 1 und 2 auf die Aufgabenstellung, und ergänzen Sie dann den Text für die Pfadadditionsregel.

Beispiel 1

Bestimmung der Wahrscheinlichkeit des Ereignisses: „Unter den zwei gezogenen Kugeln befindet sich mindestens eine rote Kugel“, das abgelesen wird durch Abzählen im Baumdiagramm.

Ziehen mit Zurücklegen

Aus dem Baumdiagramm kann man durch Abzählen entnehmen, dass von 25 Möglichkeiten es 9 gibt, mindestens eine rote Kugel zu erhalten.

$$P(\text{mit rot}) = P(wr) + P(sr) + P(rw) + P(rs) + P(rr) + P(rb) + P(rb) + P(rg) + P(br) + P(br) + P(gr)$$

$$= \frac{9}{25} = 0,36$$

Ziehen ohne Zurücklegen

Aus dem Baumdiagramm kann man durch Abzählen entnehmen, dass es 8 von 20 Möglichkeiten gibt, eine rote Kugel zu erhalten. Damit ergibt sich:

$$P(\text{mit rot}) = P(wr) + P(sr) + P(rw) + P(rs) + P(rb) + P(rb) + P(br) + P(gr)$$

$$= 8 \cdot \frac{1}{20} = \frac{8}{20} = \frac{2}{5} = 0,4$$

Beispiel 2

Bestimmung der Wahrscheinlichkeit des Ereignisses: „Es werden beim zweimaligen Ziehen zwei gleichfarbige Kugeln gezogen.“

Ziehen mit Zurücklegen

Durch Abzählen ergibt sich:

$$P(\text{gleichfarbig}) = 0 \quad (\text{unmögliches Ereignis})$$

$$= 5 \cdot \frac{1}{25} = \frac{5}{25} = \frac{1}{5} = 0,2$$

Ziehen ohne Zurücklegen

Durch Abzählen ergibt sich:

$P(\text{gleichfarbig}) = 0$ (unmögliches Ereignis)

Aufgabenstellung

Zeigen Sie analog zu den Beispielen, dass sich für die Wahrscheinlichkeit, unter den zwei gezogenen Kugeln genau eine bunte Kugel und eine rote oder eine blaue Kugel zu erhalten, die Werte 0,48 und 0,6 ergeben, wenn man die Kugeln mit/ohne Zurücklegen zieht.

Ziehen mit Zurücklegen

Durch Abzählen ergibt sich:

$$P(\text{genau eine bunte}) =$$

Ziehen ohne Zurücklegen

Durch Abzählen ergibt sich:

$$P(\text{genau eine bunte}) =$$

Ergänzen Sie anhand der Erkenntnisse aus den Beispielen und der Aufgabe den Satz für die folgende Regel:

Einfache Additionsregel

Besteht ein Ereignis E aus mehreren einzelnen disjunkten Ereignissen e_1, e_2, e_3, \dots , mit $E = \{e_1, e_2, e_3, \dots\}$, dann kann die Wahrscheinlichkeit $P(E)$ der einzelnen Wahrscheinlichkeiten der Ereignisse e_i berechnet werden:

$$P(E) = P(e_1) + P(e_2) + P(e_3) + \dots$$

*die einzelnen Ereignisse e_i haben nur unterschiedliche Elemente

Aufgabe 2.4

Reduzierte Baumdiagramme zum Ziel zurück

Das Zeichnen eines vollständigen Baums mit allen Ereignissen und Gegenereignissen, wie in den Abbildungen 2.1 und 2.2, erlaubt oft umfangreiche Baumdiagramme, bei denen das Heraussuchen bzw. Abzählen der für die Aufgabe günstigen Ereignisse aufwendig ist. Es sollen daher nun effektivere Methoden als Anwenden mit Baumdiagrammen entwickelt werden. Als Beispiel dient das schon in Aufgabe 2.3 Beispiel 1 bekannte Ereignis: „Man erhält beim zweimaligen Ziehen mindestens eine rote Kugel.“ Also $P(\text{mit rot})$.

Beginnen Sie dazu zunächst noch einmal mit der oben verwendeten Abzählmethode anhand des umfangreichen Baumes in Abb. 2.1 für das Ziehen mit Zurücklegen die Wahrscheinlichkeiten für die folgenden Ereignisse. Geben Sie das Ergebnis jeweils als Bruch an.

Abkürzung für rotgezogen: r und Abkürzung für das Gegenereignis nicht rot gezogen: \bar{r}

– Beide Kugeln sind rot $P(rr)$: $P(rr) = \frac{4}{25}$

– Die erste Kugel ist rot $P(r)$: $P(r) = \frac{4}{5}$

– Die zweite Kugel ist rot $P(\bar{r}\bar{r})$: $P(\bar{r}\bar{r}) = \frac{16}{25}$

– Die erste Kugel ist rot $P(\bar{r}\bar{r})$: $P(\bar{r}\bar{r}) = \frac{16}{25}$

Die Wahrscheinlichkeiten für die vier Ereignisse sollen nun auf einem **reduzierten Baumdiagramm** (Abb. 2.3) noch die für die Aufgabenstellung benötigten Ergebnisse rot r und nicht rot \bar{r} ersichtlich sein.

Anstatt jedes einzelne Ziehen zu zeichnen, fasst man mehrere zusammen. Somit führt man zu einem Ergebnis, sondern auch zu einem Gegenereignis.

Das Wahrscheinlichkeitsastens berücksichtigt man dann, indem man an einen Ast die Wahrscheinlichkeit für das Eintreten des Ereignis bei der Ziehung, hier die Wahrscheinlichkeiten für \bar{r} eintrifft.

Für die Berechnung der Wahrscheinlichkeiten des hier betrachteten Ereignisses $P(\text{mit rot})$ muss sich das reduzierte Baumdiagramm an. Ergänzen Sie die fehlenden Wahrscheinlichkeiten bei den Wahrscheinlichkeiten.



In beiden Stufen des Zufallsversuchs wird nun nur noch zwischen den Ereignissen rot und nicht rot unterschieden.

Die Wahrscheinlichkeiten für die Ereignisse $P(rr)$, $P(r\bar{r})$, $P(\bar{r}r)$ und $P(\bar{r}\bar{r})$ erhält man, indem man die Wahrscheinlichkeiten der Ereignisse r bzw. \bar{r} entlang der zugehörigen Äste der Stufen, d.h. entlang des Pfades, miteinander multipliziert.

Die Wahrscheinlichkeit eine rote Kugel zu ziehen ist in beiden Stufen $\frac{4}{5}$ und keine rote Kugel zu ziehen in beiden Stufen $\frac{1}{5}$.

Die Summe der Wahrscheinlichkeiten an den Ästen einer Verzweigung ist immer 1.

Die Summe der Wahrscheinlichkeiten der Ereignisse aller Pfade hier also $P(rr) + P(r\bar{r}) + P(\bar{r}r) + P(\bar{r}\bar{r})$ ist immer 1.

Berechnung der gesuchten Wahrscheinlichkeit für das Ereignis $P(\text{mit rot})$:

Weg 1: $P(\text{mit rot}) = P(rr) + P(r\bar{r}) = \frac{4}{25} + \frac{4}{25} = \frac{8}{25}$

Weg 2: $P(\text{mit rot}) = 1 - P(\bar{r}\bar{r}) = 1 - \frac{16}{25} = \frac{9}{25}$

Erläutern Sie, welche Regeln für die Berechnung von $P(\text{mit rot})$ auf den beiden Wegen verwendet wurden:

Aufgabe 2.5

Reduzierte Baumdiagramme beim Ziehen ohne Zurücklegen

Beim Ziehen ohne Zurücklegen ergibt sich für das Ereignis E: „Man erhält im n-fachen Ziehen mindestens eine rote Kugel“, das folgende reduzierte Baumdiagramm.

Da nach dem ersten Ziehen keine rote Kugel mehr vorhanden ist, fehlt der Ast rot in der zweiten Stufe des Baums und es gilt: $P(rr) = 0$. Nicht rot zu ziehen ist damit das sichere Ereignis.



a) Begründen Sie, wie sich die Wahrscheinlichkeiten an den Ästen der zweiten Stufe des Baumdiagramms von den Wahrscheinlichkeiten der ersten Stufe unterscheiden.

b) Zeigen Sie, dass die Summe der Wahrscheinlichkeiten in jeder Verzweigung den Wert 1 annimmt.

c) Berechnen Sie die Wahrscheinlichkeit des Ereignisses E durch die Addition der Wahrscheinlichkeiten der einzelnen Pfade die Werte $\frac{1}{5}$ und $\frac{4}{20} = \frac{1}{5}$ ergeben, und tragen Sie die Werte in die entsprechenden Stellen in der Abbildung ein.

d) Erläutern Sie durch eine kurze Begründung, wie man aus den Wahrscheinlichkeiten der einzelnen Ereignisse an den Ästen die Wahrscheinlichkeit des Ereignisses E erhält.

Aufgabe 2.6

Pfadmultiplikationsregel

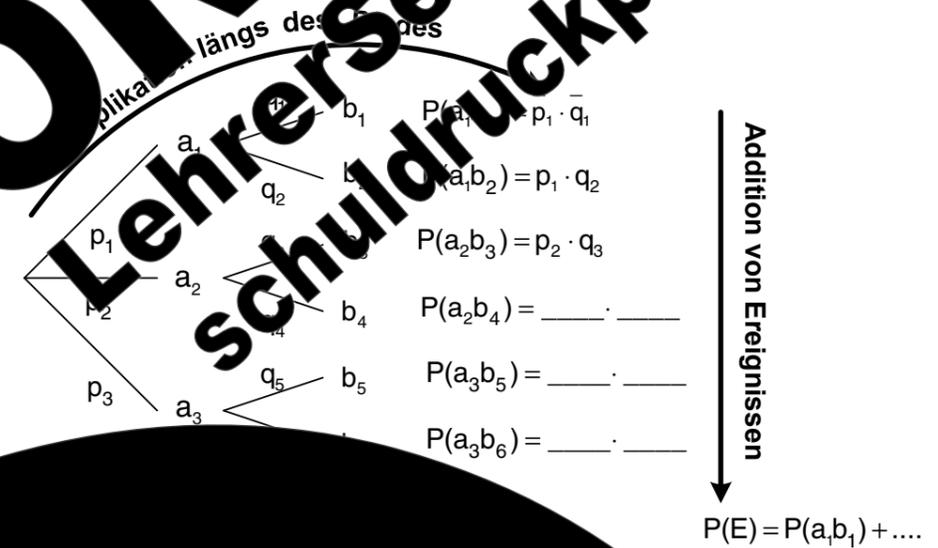
Ergänzen Sie den folgenden Satz anhand der Betrachtungen von Aufgabe 2.5 mit einer gültigen Regel:

Pfadmultiplikationsregel
Die Wahrscheinlichkeit eines Ereignisses bei einem n-fachen Ziehen ohne Zurücklegen berechnet man durch _____ der Einzelwahrscheinlichkeiten längs des Pfades, der zu dem gesuchten Ereignis führt.

Aufgabe 2.7

Zusammenfassung Aufgaben mit Baumdiagrammen

Ergänzen Sie an den entsprechenden Stellen des abgebildeten Baumdiagramms die folgenden Texte:

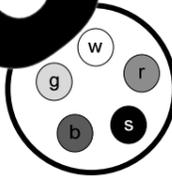


Die Wahrscheinlichkeiten der Ereignisse $P(a_i, b_j)$ werden durch die Multiplikation der Einzelwahrscheinlichkeiten entlang eines Pfades berechnet.

- Die Wahrscheinlichkeit des Ereignisses $P(E)$, das sich aus Ereignissen $P(a_i, b_j)$ bilden lässt, wird durch die Addition der zu E passenden Wahrscheinlichkeiten berechnet.
- Die Summe aller Wahrscheinlichkeiten an den Ästen einer Verzweigung ist immer 1.
- Die Summe aller Ereignisse $P(a_i, b_j)$ ist immer 1.

Aufgabe 2.7

Bestimmen anhand der bereits bekannten Urne (s. rechts) für das Ziehen und ohne Zurücklegen die Wahrscheinlichkeit des Ereignisses E, mindestens eine weiße Kugel zu erhalten, indem Sie Ihre Kenntnisse über reduzierte Baumdiagramme, Pfadmultiplikationsregel und einfache Additionsregel anwenden. Ergänzen Sie die Angaben im Baumdiagramm, und vergleichen Sie Ihre Ergebnisse mit Aufgabe 2.3.



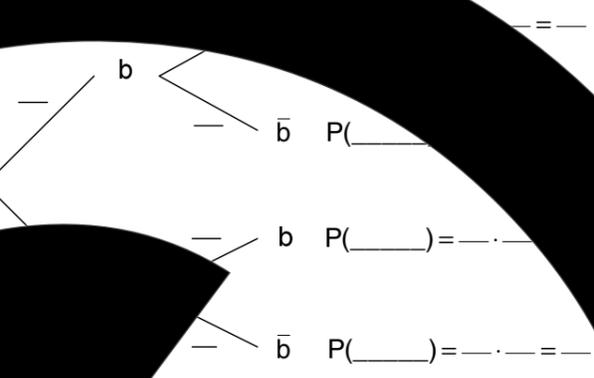
Verwendete Abkürzungen: bunt b und nicht bunt \bar{b}

Ziehen mit Zurücklegen:



Wahrscheinlichkeit für die Berechnung von P(E):
Weg 1: $P(E) = P(bb) + P(b\bar{b}) + P(\bar{b}b) = \dots + \dots + \dots = \dots$
Weg 2: $P(E) = 1 - P(\bar{E}) = 1 - P(\bar{b}\bar{b}) = 1 - \dots = \dots$

Ziehen ohne Zurücklegen:

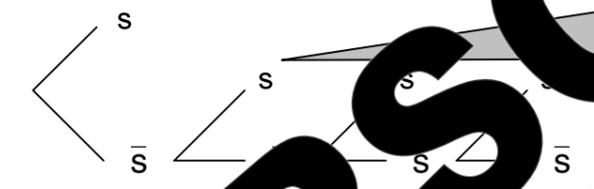


Zwischenstufe für die Berechnung von P(E):
 $P(E) = P(\dots) + P(\dots) + P(\dots) = \dots + \dots + \dots = \dots$
 $P(\bar{E}) = 1 - \dots = 1 - \dots = \dots$

Aufgabe 2.8

Ergänzen Sie die Wahrscheinlichkeiten an den Ästen der Baumdiagramme, und bestätigen Sie durch Rechnung die angegebenen Ergebnisse.

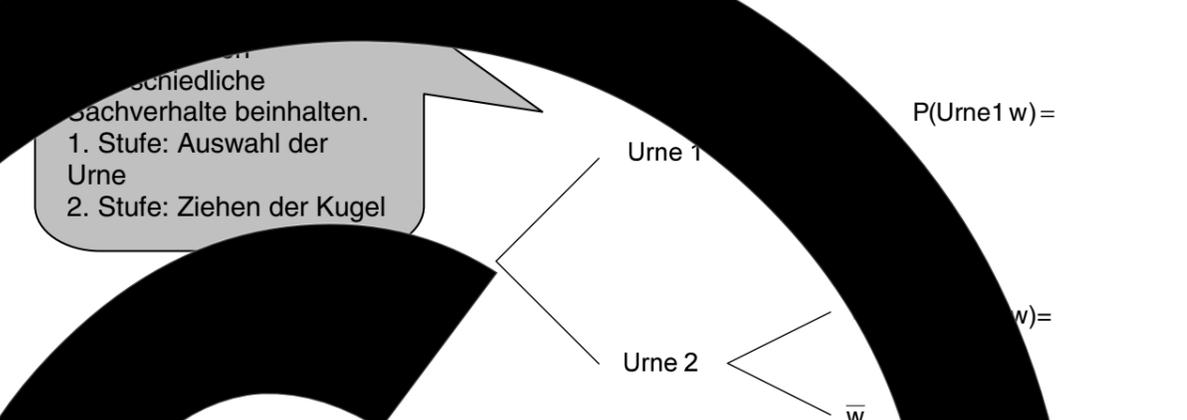
a) Aus der bereits bekannten Urne soll so lange ohne Zurücklegen gezogen werden, bis die schwarze Kugel gezogen wird. Ermitteln Sie die Wahrscheinlichkeiten an den Ästen des zugehörigen Baumdiagramms und bestätigen Sie die Wahrscheinlichkeiten an den angegebenen Ereignisse.



Im ersten Zug schwarz: $P(s) = \dots = 0,2$
Im zweiten Zug schwarz: $P(\bar{s}s) = \dots = 0,2$
Im dritten Zug schwarz: $P(\bar{s}\bar{s}s) = \dots = 0,2$
Im vierten Zug schwarz: $P(\bar{s}\bar{s}\bar{s}s) = \dots = 0,2$
Im fünften Zug schwarz: $P(\bar{s}\bar{s}\bar{s}\bar{s}s) = \dots = 0,2$

Der Zufallsvorgang ist nicht beendet, wenn die schwarze Kugel gezogen wurde. Damit muss auf diesen Pfaden auch kein weiterer Ast gezeichnet werden.

b) Zusätzlich zur ersten Urne gibt es eine zweite Urne, in der eine weiße und eine schwarze Kugel liegt. Es wird zuerst eine der beiden Urnen ausgewählt und dann aus dieser ausgewählten Urne eine Kugel entnommen. Bestätigen Sie, dass die Wahrscheinlichkeit, eine weiße Kugel zu ziehen, 35% beträgt. Ergänzen Sie das zugehörige Baumdiagramm die fehlenden Wahrscheinlichkeiten.



$P(\text{Urne 1 } w) = \dots$
 $P(\text{Urne 2 } w) = \dots$
... = 0,35
... = 0,35

Raum für Notizen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Vierfeldertafel



Einführung in die Stochastik

Kapitel 1
Grundbegriffe 7

Kapitel 2
Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
Vierfeldertafel 23

Kapitel 4
Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
Erwartungswert 63

Kapitel 8
Varianz und Kovarianzschätzung 69

Kapitel 9
Normalverteilung 75

Hypothesentests

Kapitel 10
Zweiseitiger Signifikanztest 97

Kapitel 11
Einseitiger Signifikanztest – Fisher beim Testen 109

Kapitel 12
Vertrauensintervall 135

..... 139

Kapitel 14
Anwendung der Normalverteilung 149

Gesamtwerk selbstorganisiert erlernen (Bestellnummer 278)

Alle Rechte vorbehalten. All rights reserved.
Druck, auch auszugsweise, vorbehaltlich der Rechte,
aus §§ 53, 54 UrhG ergeben, nicht gestattet.
SelbstVerlag
Verlag des Autors & Freunde GmbH, Koblenz (Germany)
www.f-druck.de

www.f-druck.de

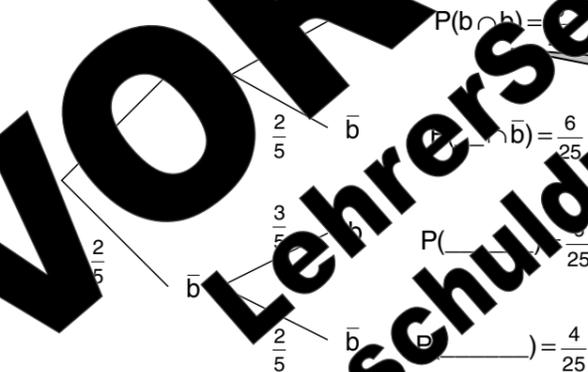
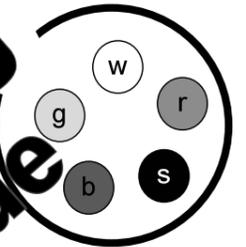
Kapitel 3: Vierfeldertafel

Als Alternative zu Baumdiagrammen kann man bei geeigneten Zufallsversuchen ein Baumdiagramm durch eine sogenannte **Vierfeldertafel** ersetzen. Die Verwendung der Vierfeldertafel ist bei der Behandlung von bedingten Wahrscheinlichkeiten (s. Kapitel 5) sehr nützlich.

Aufgabe 3.1

a) **Struktur der Vierfeldertafel bei stochastisch unabhängigen Zufallsversuchen, wie beim Ziehen mit Zurücklegen**

Aus der durch Kapitel 2 bekannten Urne mit zwei Kugeln mit Zurücklegen, die gezogen werden und das Ereignis „Es sind mindestens zwei bunte Kugeln gezogen“ erneut betrachtet werden. Das zugehörige Baumdiagramm haben Sie bereits in Aufgabe 2.7 erstellt. Die Schreibweise für die Ereignisse soll allerdings auf eine etwas vereinfachte Weise weitergeleitet werden. Wenn das Ereignis „Zwei bunte Kugeln gezogen“ eintritt, ist ja rot und bunt gezogen worden. Das wird durch die Schreibweise $b \cap r$ gekennzeichnet, wobei das Zeichen \cap bedeutet. Ergänzen Sie fehlende Angaben in den angegebenen Ereignissen.



Achtung:
Neue Schreibweise

Die folgende Abbildung zeigt eine Vierfeldertafel für diese Aufgabenstellung. Verdeutlichen Sie sich, dass die Wahrscheinlichkeiten, welche man an die Äste des Baumdiagramms schreibt, in den hellgrauen äußeren Feldern der Vierfeldertafel erscheinen und die Wahrscheinlichkeiten für die Ereignisse in den inneren Feldern eingetragen werden.

	Zweite Ziehung	
	r	b-bar
erste Ziehung	$\frac{9}{25}$	$\frac{6}{25}$
b-bar	$\frac{6}{25}$	$\frac{4}{25}$
	$\frac{15}{25} = \frac{3}{5}$	$\frac{10}{25} = \frac{2}{5}$

Verdeutlichen Sie die folgenden Zusammenhänge bei der Vierfeldertafel:

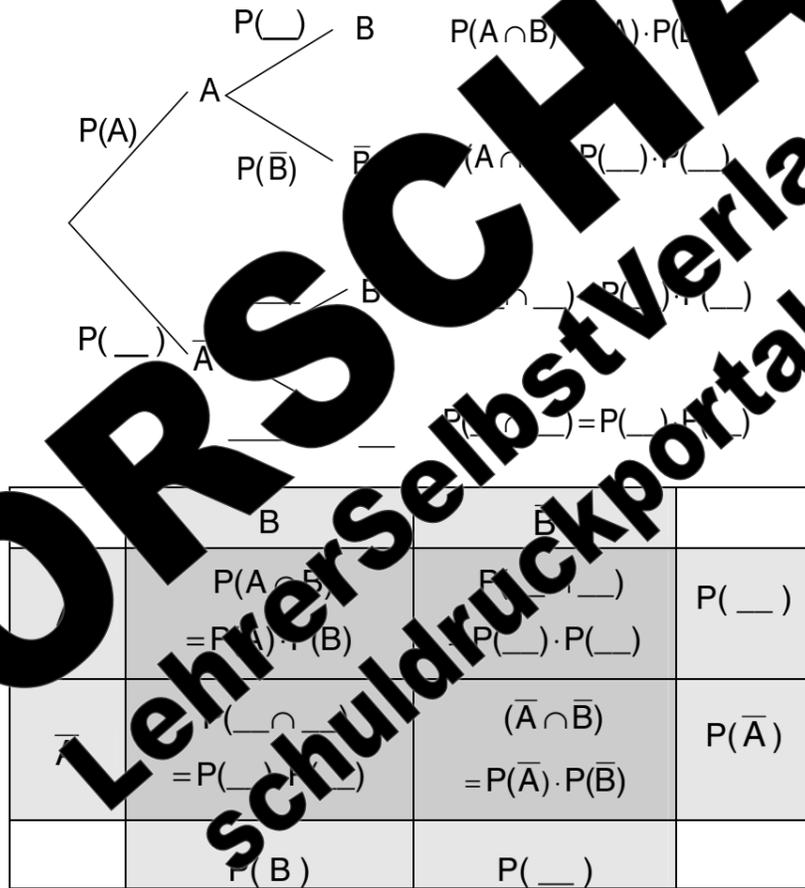
Die Summe zweier Werte in den inneren Feldern ergibt immer den Wert der entsprechenden Zeile bzw. Spalte in den äußeren Feldern.

Die Summe der Werte in der äußeren Zeile bzw. Spalte ergeben immer den Wert der entsprechenden Zeile bzw. Spalte in den äußeren Feldern.

Die inneren Werte lassen sich durch Multiplikation der beiden Werte der äußeren Zeile und Spalte ermitteln. Diese Schreibweise ist allerdings nur dann ein Zufallsversuch vorliegt, bei dem die beiden Züge des Zufallsversuchs **stochastisch unabhängig**, voneinander unabhängig sind. Dies ist beim Ziehen mit Zurücklegen erfüllt.

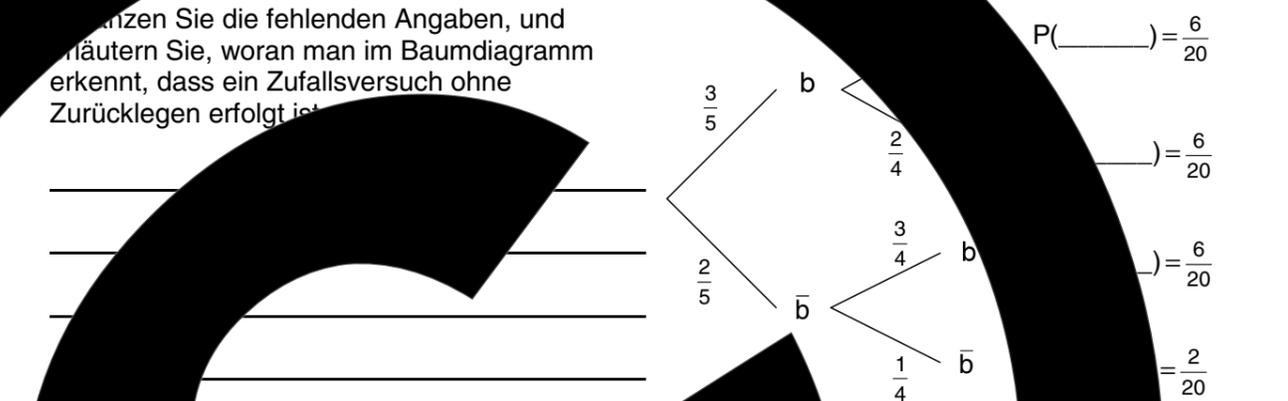
Aufgabe 3.2

a) Ergänzen Sie im folgenden Überblick für den Zusammenhang von Baumdiagramm und Vierfeldertafel bei einem **stochastisch unabhängigen** Zufallsversuch die fehlenden Angaben.



b) Struktur von Zufallsversuchen, wie beim Ziehen ohne Zurücklegen aus einer Urne mit 20 Kugeln, die 6 rot und 14 blau sind.

Die Aufgabenstellung von Aufgabe a) wird nun für einen stochastisch unabhängigen Zufallsversuch (Ziehen aus einer Urne mit 20 Kugeln) betrachtet. Auch die Aufgabenstellung haben Sie bereits in Aufgabe a) gesehen. Ergänzen Sie die fehlenden Angaben, und erläutern Sie, woran man im Baumdiagramm erkennt, dass ein Zufallsversuch ohne Zurücklegen erfolgt ist.



Vierfeldertafel:

Da beim Ziehen ohne Zurücklegen die Wahrscheinlichkeiten der zweiten Stufe eines Zufallsversuches von dem Ausgang der Ziehung in der ersten Stufe abhängen, bezeichnet man diesen Zufallsversuch mit dieser Eigenschaft als **abhängige** bzw. **stochastisch abhängige** Zufallsversuch. Verdeutlichen Sie sich die Auswirkungen der Abhängigkeit auf die Zusammenhänge bei der Vierfeldertafel:

	b zweite Ziehung	\bar{b} zweite Ziehung	
b erste Ziehung	$\frac{6}{20} \cdot \frac{5}{19} = \frac{30}{380}$	$\frac{6}{20} \cdot \frac{14}{19} = \frac{84}{380}$	$\frac{12}{19}$
\bar{b} erste Ziehung	$\frac{14}{20} \cdot \frac{6}{19} = \frac{84}{380}$	$\frac{14}{20} \cdot \frac{14}{19} = \frac{196}{380}$	$\frac{20}{19}$
	$\frac{12}{19}$	$\frac{20}{19}$	1

- Die vertikale und horizontale Summe der Wahrscheinlichkeiten in den inneren Feldern ergibt immer den Wert, der den entsprechenden Zeilen bzw. Spalten in den äußeren Feldern entspricht.
- Die Summe der Werte der äußeren Zeile bzw. Spalte ergibt immer 1.
- Die Wahrscheinlichkeiten für die erste Stufe des Baumes werden an den entsprechenden Stellen in den äußeren Feldern angegeben.
- Die Wahrscheinlichkeiten für die zweite Stufe des Baumes kann man der Vierfeldertafel nicht direkt ablesen, sondern muss sie den äußeren Feldern entnehmen.
- Die Wahrscheinlichkeiten der inneren Felder lassen sich **nicht** durch Multiplikation der beiden zugehörigen Werte der äußeren Felder in der gleichen Zeile und Spalte ermitteln.

Aufgabe 3.3

In einer Vierfeldertafel sind die Wahrscheinlichkeiten für einen stochastisch unabhängigen und einen stochastisch abhängigen Zufallsversuch angegeben. Fassen Sie die Gemeinsamkeiten und Unterschiede zusammen. Ergänzen Sie die folgenden Texte:

Bei einem stochastisch unabhängigen Zufallsversuch sind die in den äußeren Feldern angegebenen Wahrscheinlichkeiten den Werten $P(A)$ und $P(B)$ an den Ästen der ersten und zweiten Stufe eines Baumdiagramms, während bei einem stochastisch abhängigen Zufallsversuch nur die Wahrscheinlichkeiten an den Ästen der ersten Stufe eines Baumdiagramms angegeben sind. Die in den äußeren Feldern der Vierfeldertafel angegebenen Wahrscheinlichkeiten sind die $P(A)$ und $P(B)$.

b) Ergänzen Sie die folgenden Texte: Die Wahrscheinlichkeiten der äußeren Felder durch entsprechende Summen der inneren Felder ermitteln.

c) Bei einem stochastisch unabhängigen Zufallsversuch kann man die Werte für die Wahrscheinlichkeiten der inneren Felder durch entsprechende Multiplikation der Wahrscheinlichkeiten der äußeren Felder ermitteln. Das ist bei einem stochastisch unabhängigen Zufallsversuch _____.

d) Die Summe der Wahrscheinlichkeiten in der Zeile und in der Spalte der inneren Felder ergeben bei einem _____ und einem _____ Zufallsversuch bei jeder Vierfeldertafel den Wert 1.

e) Die Summe der Werte für die Wahrscheinlichkeiten aller inneren Felder ergibt _____.

Übungen

3.1 Prüfung auf stochastische Unabhängigkeit bzw. Unabhängigkeit der Vierfeldertafel

a) Begründen Sie, dass eine Bedingung, dass die rechte Vierfeldertafel zu einem stochastisch unabhängigen Zufallsversuch gehört.

	Merkmal B	Merkmal \bar{B}	
Merkmal A	0,15	0,25	0,4
Merkmal \bar{A}	0,1	0,45	0,6
	0,25	0,7	1

Die rechte Vierfeldertafel zu einem stochastisch unabhängigen Zufallsversuch gehört.

	Merkmal D	Merkmal \bar{D}	
Merkmal C	0,08	0,12	0,2
Merkmal \bar{C}	0,32	0,48	0,8
	0,4	0,6	1

Ü3.2 Aus einer Urne mit 2 roten und 4 schwarzen Kugeln werden zwei Kugeln gezogen. Die Kugeln werden mit Zurücklegen gezogen. Erstellen Sie ein Baumdiagramm. Übertragen Sie die Werte für Wahrscheinlichkeiten und Ereignisse an die entsprechenden Stellen der Vierfeldertafel, und prüfen Sie, ob die Regeln für die Addition und die Multiplikation bei der Vierfeldertafel erfüllt sind.

Baumdiagramm:

Vierfeldertafel:

Die Regeln für die Addition sind _____

Die Regeln für die Multiplikation sind _____

Ü3.3 Bei einem Glücksspiel gibt es eine Urne mit 2 schwarzen und drei weißen Kugeln. Es soll ermittelt werden, wie die Mischung aus schwarzen und weißen Kugeln bei einer zweiten Urne aussehen muss, damit die Wahrscheinlichkeit zwei schwarze Kugeln zu ziehen bei 10% liegt, wenn aus jeder Urne ein Kugel gezogen wird.



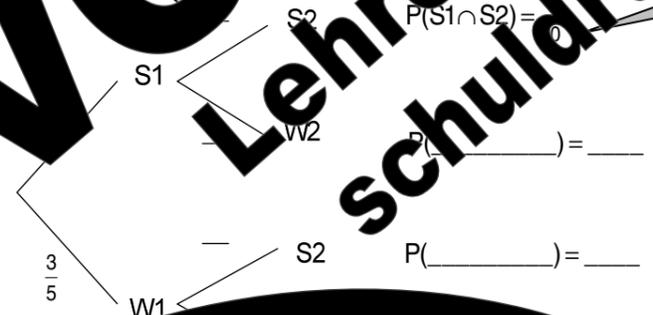
Abkürzungen: S1 schwarze Kugel aus Urne 1, W1 weiße Kugel aus Urne 1
S2 schwarze Kugel aus Urne 2, W2 weiße Kugel aus Urne 2

a) Begründen Sie, warum es sich bei diesem Spiel um einen stochastisch unabhängigen Zufallsversuch handelt, obwohl es sich um einen Versuch mit Zurücklegen handelt.

b) Ermitteln Sie die Wahrscheinlichkeiten aller Ereignisse mithilfe eines Baumdiagramms:

Wahrscheinlichkeit eines Ereignisses ist bekannt.

Wahrscheinlichkeit eines Ereignisses ist bekannt



Hilfe:

Die fehlenden Wahrscheinlichkeiten in der zweiten Stufe des Baumes lassen sich durch Rückrechnen aus der angegebenen Wahrscheinlichkeit für das Ereignis „zwei schwarze Kugeln“ und den Pfadwahrscheinlichkeiten in der ersten Stufe berechnen. Ergänzen Sie die fehlenden Wahrscheinlichkeiten in der zweiten Stufe und zeigen Sie, dass es in Urne 2 eine schwarze und vier weiße Kugeln geben muss.

Alternativ zum Baumdiagramm kann die Lösung auch mit einer Vierfeldertafel ermittelt werden. Berechnen Sie die Felder, ohne auf die Ergebnisse des Baumdiagramms zurückzugreifen, und vergleichen Sie anschließend die Werte.

	S2	
S1		$\frac{3}{5}$
		1

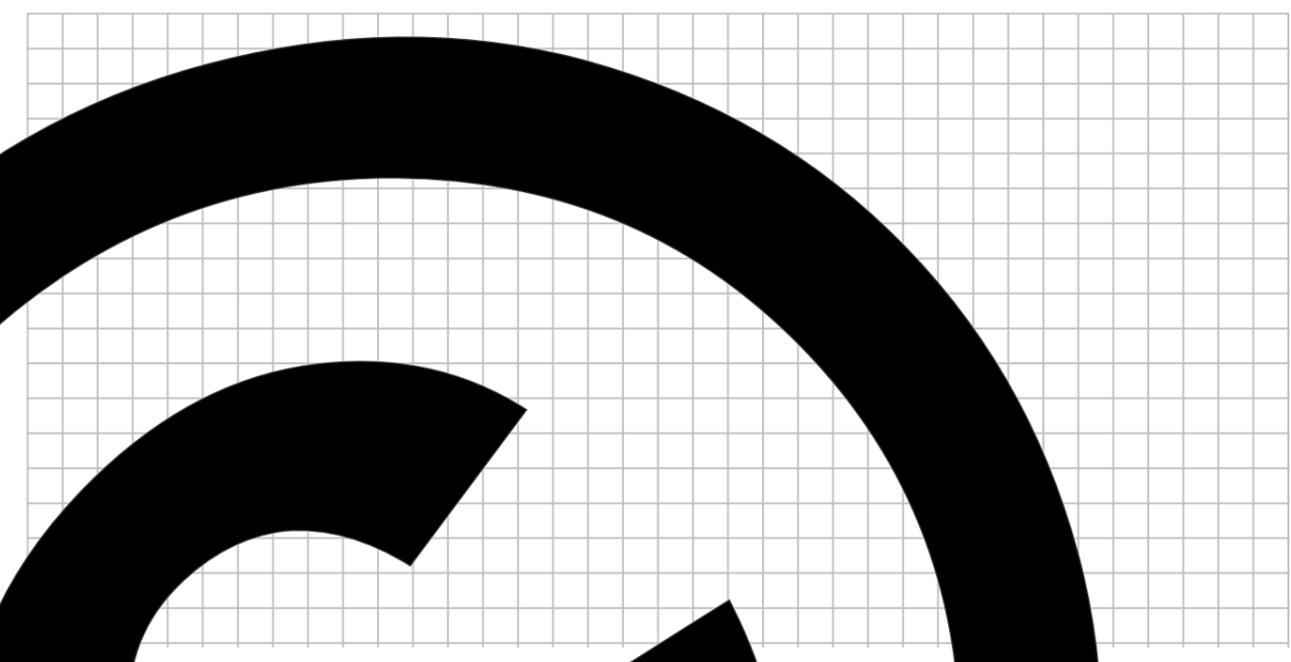
Hilfe:

Da es sich um einen unabhängigen Zufallsversuch handelt, kann man die Additions- und Multiplikationsregel beim Ausfüllen der Vierfeldertafel verwenden. Beispielsweise kann $P(S_2) = \frac{2}{5}$ aus dem Satz $\frac{2}{5} \cdot P(S_2) = \frac{1}{10}$ berechnet werden, was ergibt $P(S_2) = \frac{1}{10} \cdot \frac{5}{2} = \frac{1}{4}$

Ü3.4 Ein Falschspieler wirft gleichzeitig eine ideale und eine gefälschte Münze. Mit welcher Wahrscheinlichkeit zeigt die gefälschte Münze Wappen an, wenn das Ereignis „Beide Münzen zeigen Zahl“ eine Wahrscheinlichkeit von 30% hat. Berechnen Sie die gesuchte Wahrscheinlichkeit zunächst mit einer Vierfeldertafel, und überprüfen Sie Ihr Ergebnis mit einem Baumdiagramm.



Einem Würfelspieler werden zwei Würfel geworfen. Ein Spieler verliert, wenn mindestens eine 1 fällt. Wie muss der Falschspieler einen der beiden Würfel fälschen, wenn er mit einer Wahrscheinlichkeit von 75% gewinnen will? Lösen Sie die Aufgabe mithilfe einer Vierfeldertafel und vergleichsweise mit einem Baumdiagramm.



Aufgabe 3.5

Die Vierfeldertafel mit absoluten Werten

Es gibt häufig Aufgabenstellungen, bei denen keine Angaben zu den Wahrscheinlichkeiten eines Merkmals gegeben sind, sondern absolute Zahlen vorgegeben werden. Auch in diesem Fall kann man mit Vierfeldertafeln arbeiten.

Zahlenbeispiel:

An einer Schule mit 720 Mädchen und 480 Jungen sind während einer Grippeperiode 200 Schüler und Schülerinnen erkrankt. Die Sekretärin hat 144 Krankmeldungen bei Mädchen gezählt. Für eine Statistik soll untersucht werden, ob die Grippeerkrankung vom Geschlecht abhängt und stochastisch unabhängig ist.

Abkürzungen für die Vierfeldertafel: $M \cap K$ = Mädchen, die erkrankt sind, $\bar{M} \cap \bar{K}$ = nicht erkrankt

Info:
Aussehen der Vierfeldertafel für absolute Werte

	K	\bar{K}	
M	$M \cap K$	$M \cap \bar{K}$	Anzahl M
\bar{M}	$\bar{M} \cap K$	$\bar{M} \cap \bar{K}$	Anzahl \bar{M}
	Anzahl K	Anzahl \bar{K}	Gesamtzahl

Erläutern Sie im Kontext der Aufgabenstellung die Bedeutung von:

$M \cap K$ _____

$\bar{M} \cap \bar{K}$ _____

$M \cap \bar{K}$ _____

$\bar{M} \cap K$ _____

Schritt 1:
Aus dem gegebenen Text schließen.

	K	\bar{K}	
M	144		720
\bar{M}			
	240		

Schritt 2:
Durch Anwendung der Additionsregel fehlende Werte ermitteln.

Erläutern Sie, warum der Wert $M \cap K$ eingetragen wird.

Erläutern Sie, wie der Wert 240 ermittelt wird.

Schritt 3:
Ermitteln der Wahrscheinlichkeiten für die Felder der Vierfeldertafel

	K	\bar{K}	
M	0,12	0,48	
\bar{M}	0,08	0,32	0,4
	0,2	0,8	1

a) Erläutern Sie, wie die Werte für die Felder berechnet wurden.

b) Begründen Sie, warum man bei der Untersuchung auf stochastische Abhängigkeit die Werte der Vierfeldertafel als relative Anteile (also über Wahrscheinlichkeiten) angibt.

c) Begründen Sie anhand einer Rechnung, dass die Erkrankung an Grippe stochastisch unabhängig ist.

Erläutern Sie die Berechnungen:

Aufgabe 3.6

Die Vierfeldertafel und der allgemeine Additionssatz

In Kapitel 2 haben Sie die einfache Additionsregel für die Berechnung der Wahrscheinlichkeit von Ereignissen kennengelernt. Lesen Sie ggf. zur Wiederholung in Aufgabe 2.1a nach, was man unter dieser Regel versteht. Anhand des folgenden Beispiels soll gezeigt werden, dass die einfache Additionsregel nicht immer angewendet werden kann und zu einer **allgemeinen Additionsregel** erweitert werden muss.

Zahlenbeispiel:

In einer Glasbläserei entstehen bei der Fertigung zwei Fehler. Die Gläser können Luftschüsse oder Formfehler oder auch beide Fehler haben. Luftschüsse werden bei 10% der Gläser und Formfehler bei 8% der Gläser festgestellt. Die Fehler treten unabhängig voneinander auf. Ein fehlerbehaftetes Glas wird als Ausschuss aussortiert und eingeschmolzen. Es soll ermittelt werden, wie viel Prozent Ausschuss entsteht.

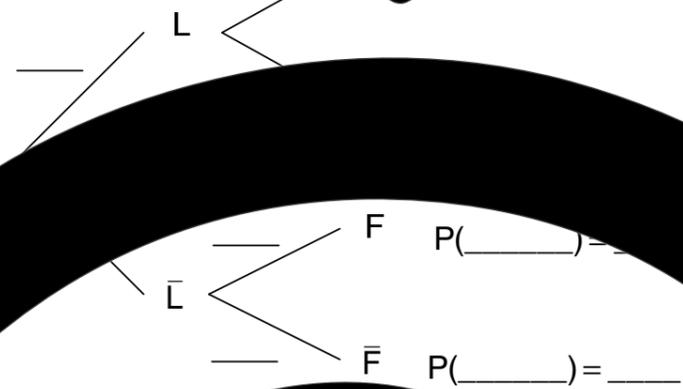
Abkürzungen: L Luftschuss, \bar{L} kein Luftschuss, F Formfehler, \bar{F} kein Formfehler

Die Lösung der Aufgabenstellung soll auf drei verschiedene Weisen ermittelt werden:

1. Lösung über das Baumdiagramm

Das Baumdiagramm für aus, und zeigen Sie durch entsprechende Rechnung, dass 12,6% der Gläser einen Fehler haben und damit Ausschussware sind:

$P(L \cap F) = \dots$



2. Lösung über die Vierfeldertafel.

Füllen Sie die Vierfeldertafel aus, und zeigen Sie, dass sich für P(Ausschuss) falls ein Wert von 12,6% ergibt.

	F	\bar{F}	
L			
\bar{L}			

3. Lösung über die einfache Additionsregel

$P(\text{Ausschuss}) = P(L) + P(F) - P(L \cap F) = 0,10 + 0,08 - 0,08 = 0,10 = 10\%$

Dieses Ergebnis stimmt mit den ausführlich ermittelten Ergebnissen nicht überein und muss **falsch** sein.

Der Fehler liegt auf der Hand!

Die Lösung 3 entstandener Fehler kann mithilfe der Vierfeldertafel sehr anschaulich analysiert und gehoben werden. Dazu werden die am Baumdiagramm beispielhaft erfolgten Überlegungen mithilfe der Vierfeldertafel veranschaulicht dargestellt.

	F	\bar{F}	
L	$P(F \cap L)$	$P(\bar{F} \cap L)$	$P(L)$
\bar{L}	$P(F \cap \bar{L})$	$P(\bar{F} \cap \bar{L})$	$P(\bar{L})$
	$P(F)$	$P(\bar{F})$	

Lösungsweg 2

ganzen Sie: Der Ausschussanteil setzt sich aus ... Fehlern und aus ... Gläsern mit einem der beiden Fehler zusammen und ... berechnet.

Lösungsweg 3 in aller Kürze dargestellt:

$P(\text{Ausschuss}) = P(L) + P(F) - P(L \cap F) = 0,10 + 0,08 - 0,054 = 0,126 = 12,6\%$

... $P(\text{Ausschuss})$ aus der Summe von $P(L)$ und $P(F)$ berechnet ... dass das ... beiden Fehler, nämlich $P(F \cap L)$... Rechnung

... heißt und das ist ...

Begründen Sie allgemein, dass der Ansatz: $P(\text{Ausschuss}) = P(L) + P(F) - P(F \cap L)$ das richtige Ergebnis liefert, und weisen Sie dies auch am Zahlenbeispiel der obigen Aufgabenstellung nach.

Begründung: _____

Nachweis am Zahlenbeispiel: $P(\text{Ausschuss}) = P(L) + P(F) - P(F \cap L) = \dots = \dots\%$

Verallgemeinerte Formel

Damit in unserem Zahlenbeispiel ein Glas zur Ausschussware erklärt es aus, dass mindestens einer der beiden Fehlertritt. Alternativ formuliert kann man sagen: Es tritt ein Formfehler oder ein Lufteinguss auf. Mathematisch wird die Oderbeziehung mit dem Zeichen \cup angegeben. Für zwei beliebige Ereignisse A bzw. B bezeichnet weiter folgt: $P(A \text{ oder } B) = P(A \cup B)$

Für zwei Ereignissen Additionsregel ergibt sich:

Allgemeine Additionsregel

Für zwei Ereignisse A und B gilt: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Ra...

Erg...

VORSCHAU
Lehrerselbstverlag
Schuldruckportal.de

Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



VORSCHAU
Lehrerselbstverlag
Schuldruckportal.de

Kapitel

Kombinatorische Abzählverfahren



VORSCHAU

Einführung in die Stochastik

Kapitel 1
Grundbegriffe 7

Kapitel 2
Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
Vierfeldertafel 23

Kapitel 4
Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
Erwartungswert 63

Kapitel 8
Varianz und Standardabweichung 69

Kapitel 9
Normalverteilung 75

Hypothesentests

Kapitel 10
Zweiseitiger Signifikanztest 97

Kapitel 11
Einseitiger Signifikanztest – Fisher beim Testen 109

Kapitel 12
Vertrauensintervall 135

Kapitel 13
Die Betrachtungen zur Normalverteilung 139

Kapitel 14
Anwendung der Normalverteilung 149

VORSCHAU
Lehrersebstverlag
schuldruckportal.de

Gesamtheit selbstorganisiert erlernen (Bestenfalls 278)

Sie sind vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte,
aus §§ 53, 54 UrhG ergeben, nicht gestattet

LehrersebstVerlag

Lehrersebstverlag & Freunde GmbH, Koblenz (Germany)

schuldruckportal.de

www.f-druck.de

Kapitel 4: Kombinatorische Abzählverfahren

Wir haben bisher in erster Linie Zufallsversuche betrachtet, bei denen man die Wahrscheinlichkeit eines Ereignisses mithilfe von Baumdiagrammen, die aus wenigen Ästen und wenigen Endknoten bestanden, berechnen konnte. Ergeben sich bei einer Aufgabenstellung jedoch Baumdiagramme mit vielen Ästen und Stufen, können diese sehr umfangreich und unübersichtlich werden. Man kann in diesen Fällen auf rechnerische Möglichkeiten, welche als **kombinatorische Abzählverfahren** bezeichnet werden, zurückgreifen.

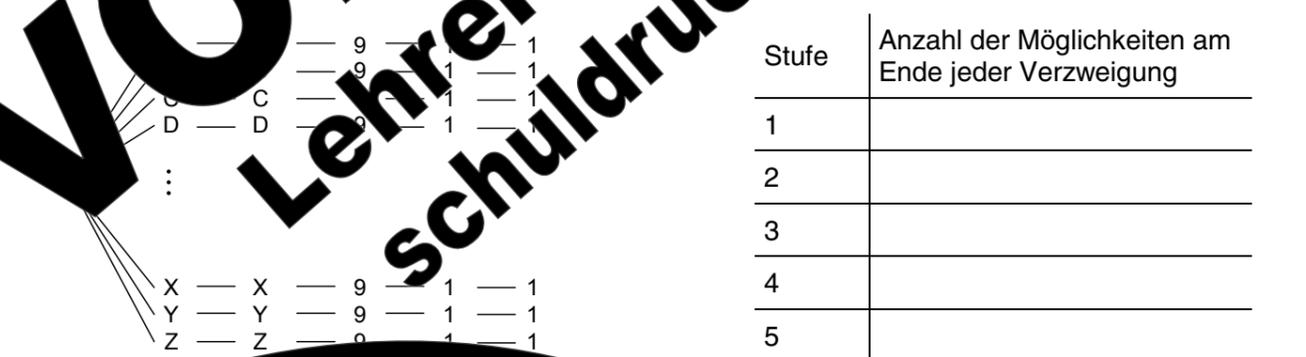
Aufgabe 4.1

Produktregel

Jemand möchte im Kreis Groß-Gerau eine Wunschnummer zulassen und wünscht sich ein Kennzeichen, das nachfolgendes Schema hat und durch die Ziffernsumme zwei weitere gleiche Buchstaben hat und durch die Ziffernsumme die Zahl 911.



a) Für die Berechnung der Anzahl der Möglichkeiten, ein wunschgemäßes Kennzeichen zu erhalten, kann man sich bei dieser Aufgabe auch auf eine andere Weise noch mit einem Baumdiagramm behelfen. Begründen Sie dies mit einem geeigneten Baumdiagramm und der Anzahl der Möglichkeiten an den Verzweigungen in jeder Stufe. Warum es insgesamt genau 26 Kennzeichen mit einer Wunschnummer gibt, wird durch Umlaute ausgeschlossen, es wird vereinfacht und keine weiteren gesetzlichen Einschränkungen bei der Vergabe von Buchstabenkombinationen berücksichtigt.



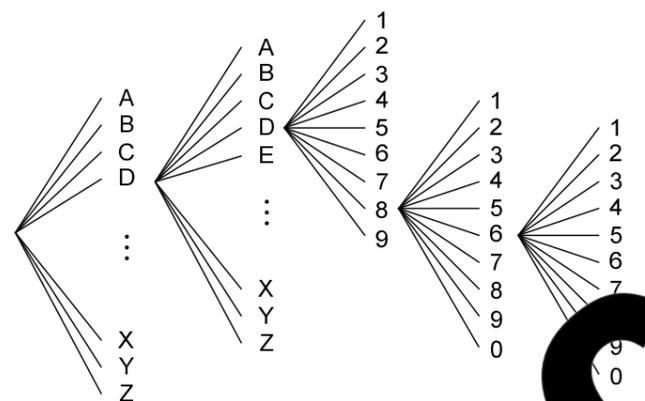
Stufe	Anzahl der Möglichkeiten am Ende jeder Verzweigung
1	
2	
3	
4	
5	

Begründen Sie die Anzahl der Möglichkeiten, ein wunschgemäßes Kennzeichen zu erhalten, indem Sie in der Tabelle für jede Stufe des Baums die Anzahl der Möglichkeiten an den Verzweigungen eintragen, und erklären Sie, warum nur ein Baumdiagramm dargestellt wird.

b) Für die Berechnung der Wahrscheinlichkeit, rein zufällig ein wunschgemäßes Kennzeichen wie oben zu erhalten, sind die Anzahl der möglichen Kennzeichen bestimmen. Vereinfachend soll hier angenommen werden, dass Groß-Gerau nur Nummernschilder mit zwei Buchstaben und drei Ziffern zulassen. Ausschluss von Umlauten und einer Null in der dreistelligen Ziffernfolge wird angenommen. Berechnen Sie die Wahrscheinlichkeit, ein wunschgemäßes Kennzeichen zu erhalten. Begründen Sie dies mit einem Baumdiagramm, warum man die Anzahl der Kennzeichen mit der

$$26 \cdot 9 \cdot 10 \cdot 10 = 585000$$

man kann, indem Sie in der Tabelle für jede Stufe des Baums die Anzahl der Möglichkeiten an den Verzweigungen eintragen, und erklären Sie, warum nur ein Baumdiagramm dargestellt wird.



Stufe	Anzahl der Möglichkeiten am Ende der Verzweigung
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	

Begründung und Erläuterung:

Berechnung der Wahrscheinlichkeit für das Wunschkennzeichen anhand der Definition der Wahrscheinlichkeit:

$$P(\text{Wunschzeichen}) = \frac{\text{Anzahl der Wunschzeichen}}{\text{Anzahl aller Kennzeichen}} = \dots = \dots \%$$

Wenn Sie die Erläuterung... haben, wird ersichtlich, dass man die Anzahl aller Möglichkeiten... Anzahl der Möglichkeiten an den Verzweigungen in jeder Stufe... Diese Berechnungsmethode wird in der Kombinatorik...

Ein Fallversuch bestehe aus k unabhängigen Stufen, wobei in der 1. Stufe n_1 , in der 2. Stufe n_2 ... und in der k-ten Stufe n_k mögliche Ergebnisse gibt. Dann wird die Anzahl aller möglichen Ereignisse N des Produkts $N = n_1 \cdot n_2 \cdot \dots \cdot n_k$ berechnet.

Die Produktregel... von Wahrscheinlichkeiten bei...
herangezogen... unterschieden zwischen:

- Ziehungen mit Zurücklegen.
- Ziehungen ohne Zurücklegen, bei denen die Reihenfolge der Ergebnisse eine Rolle spielt (z.B. bei Lottozahlen). Man bezeichnet diese Stichproben als geordnete Stichproben.
- Ziehungen ohne Zurücklegen, bei denen die Reihenfolge der Ergebnisse keine Rolle spielt (z.B. bei der Auswahl einer Kommission). Man bezeichnet diese Stichproben als ungeordnete Stichproben.

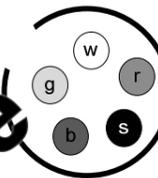
Aufgabe 4.2

Geordnete Stichproben mit Zurücklegen oder Variation

N Anzahl aller Möglichkeiten
 n Anzahl der Kugeln in einer Urne
 k Anzahl der Ziehungen

$$N = n^k$$

- a) Erläutern Sie beispielhaft anhand der bereits aus Kapitel 2 bekannten Urne die Rechenregel zur Ermittlung der Anzahl möglicher Ereignisse, wenn:
- 3 Kugeln mit Zurücklegen gezogen werden.
 - die Reihenfolge der Kugeln eine Rolle spielt, also beispielsweise die Reihenfolge wrs und rsw als unterschiedliche Ereignisse gewertet werden.



b) Ein Fallschloss hat vier Ziffernräder mit den Ziffern 0 bis 9. Begründen Sie, warum die oben angegebene Regel bei dem Zahlenschloss zur Berechnung von allen Kombinationsmöglichkeiten angewendet werden kann, und zeigen Sie, dass es im ungünstigsten Fall mehr als zweieinhalb Stunden dauert, eine vermessene Zahlungskombination durch Probieren herauszufinden, wenn für jede falsche Einstellung eine Sekunde benötigt wird.

Aufgabe 4.3

Geordnete Stichproben ohne Zurücklegen oder Permutationen

N Anzahl aller Möglichkeiten
 n Anzahl der Kugeln in einer Urne
 k Anzahl der Ziehungen

$$N = n(n-1)(n-2) \dots (n-k+1)$$

Wenn $k = n$ gilt $k = n$ $N = n(n-1)(n-2) \dots 1$ (n-Fakultät)

Wenn Sie einen wissenschaftlichen Taschenrechner verwenden, können Sie die Anzahl aller Möglichkeiten n! mit einer entsprechenden Taste oder Taste in Kombination zu berechnen.

a) Bestätigen Sie diese Regel zur Ermittlung der Anzahl aller möglichen Ereignisse im Spielhaus, wenn aus der bekannten Urne (s. Aufg. 4.2) drei Kugeln ohne Zurücklegen gezogen werden und die Reihenfolge der Kugeln eine Rolle spielt. Erläutern Sie auch, warum man die k-te Ziehung mithilfe der Summe $(n - k + 1)$ berechnet.

1. Ziehung: $n = 5$ Kugeln und damit 5 Möglichkeiten

2. Ziehung: $n - 1 = n - 2 + 1 = 5 - 1 = 4$ Kugeln mit 4 Möglichkeiten

3. Ziehung: $n - 2 = n - 3 + 1 = 5 - 2 = 3$ Kugeln und damit 3 Möglichkeiten

Insgesamt gibt es also $N = 5 \cdot 4 \cdot 3 = 60$ möglichen Kombinationen.

Erläuterung:

b) Wenn man aus einer Urne zieht gilt: $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

c) Bestimmen Sie, warum die Formeln $N = n \cdot (n-1) \cdot (n-2) \dots (n-k+1)$ bzw. $N = n!$ bei der folgenden Aufgabe angewendet werden können und berechnen Sie jeweils die Anzahl der unterschiedlichen Kombinationsmöglichkeiten.

Ein Möbelhersteller bietet für ein Regalsystem 10 verschiedene Module an. Zeigen Sie durch Rechnung, dass sich mehr als 150 000 Varianten für die Zusammenstellung der Regalwand ergeben, wenn man im Laden 6 Module kauft bzw. mehr als 3 500 000 unterschiedliche Regalwände, wenn man jedes Module einmal gekauft wird. Zeigen Sie auch, dass man mit 10 Modulen 10! Regalwände zusammenstellen kann.

(2) Auswahl 10 Module im Möbelmarkt $N = 10!$

(3) Variationen $N = 10!$

Ergänzen Sie:

Aufgabe 4.4

Ungeordnete Stichprobe ohne Zurücklegen oder unvollständige Permutation

Beim Lotto werden aus einer Urne mit 49 Kugeln mithilfe des Ziehungsgeräts 6 Kugeln ohne Zurücklegen entnommen, wobei die Reihenfolge, in der die Kugeln gezogen werden, keine Rolle spielt, ob man gewinnt oder verliert. Stichproben, bei denen es, wie im Lotto, nicht auf eine bestimmte Reihenfolge bei der Ziehung ankommt, werden als **ungeordnete Stichproben** bezeichnet. Die Formel, mit der man die Anzahl der Kombinationsmöglichkeiten berechnen kann, soll anhand der folgenden Aufgabenstellungen veranschaulicht werden.

Aufgabenstellung:

Aus einer Urne mit n bunten Kugeln werden k Kugeln ohne Zurücklegen gezogen. Die Anzahl der unterschiedlichen Möglichkeiten, k ungeordnete Kugeln zu entnehmen, soll mit N bezeichnet werden. Ergänzen Sie die fehlenden Angaben in der folgenden Tabelle.

Anmerkung zum Verständnis der Überlegungen:

Um das Zustandekommen der Formel zu verstehen zu können, ist es notwendig zu akzeptieren, dass es neben den Möglichkeiten, für mehrere Kugeln aus einer Urne zu entnehmen, auch immer **eine** Möglichkeit gibt, keine Kugel, also $k = 0$ Kugeln, zu ziehen. Das entspricht der **einen** Möglichkeit, nichts zu tun. Diese Möglichkeit besteht selbst für eine leere Urne, also für eine Urne mit $n = 0$ Kugeln.

1. In der Urne befinden sich $n = 0$ Kugeln

keine Kugel ziehen: $k = 0$	Es gibt eine Möglichkeit N nichts zu tun: $\Rightarrow N = 1$
-----------------------------	--

2. In der Urne befindet sich eine rote Kugel also $n = 1$ Kugeln

keine Kugel ziehen $k = 0$	Es gibt eine Möglichkeit N nichts zu tun: $\Rightarrow N = 1$
eine Kugel ziehen: $k = 1$	eine Möglichkeit eine Kugel zu ziehen: $\Rightarrow N = 1$

3. In der Urne befinden sich eine rote (r) und eine grüne (g) Kugel also $n = 2$ Kugeln.

keine Kugel ziehen $k = 0$	Es gibt eine Möglichkeit N nichts zu tun: $\Rightarrow N = 1$
eine Kugel ziehen: $k = 1$	Es gibt zwei verschiedene Möglichkeiten eine rote oder grüne Kugel zu ziehen: $\Rightarrow N = 2$

4. In der Urne befinden sich eine rote (r) und eine grüne (g) Kugel also $n = 2$ Kugeln.
 $k = 2$ $\begin{matrix} < r-g & \rightarrow rg \\ & g-r & \rightarrow gr \end{matrix}$
 Da zwischen rg und gr ein Unterschied vorliegt, sind dies zwei unterschiedliche Möglichkeiten, die eine rote und eine grüne Kugel zu ziehen: $\Rightarrow N = 2$

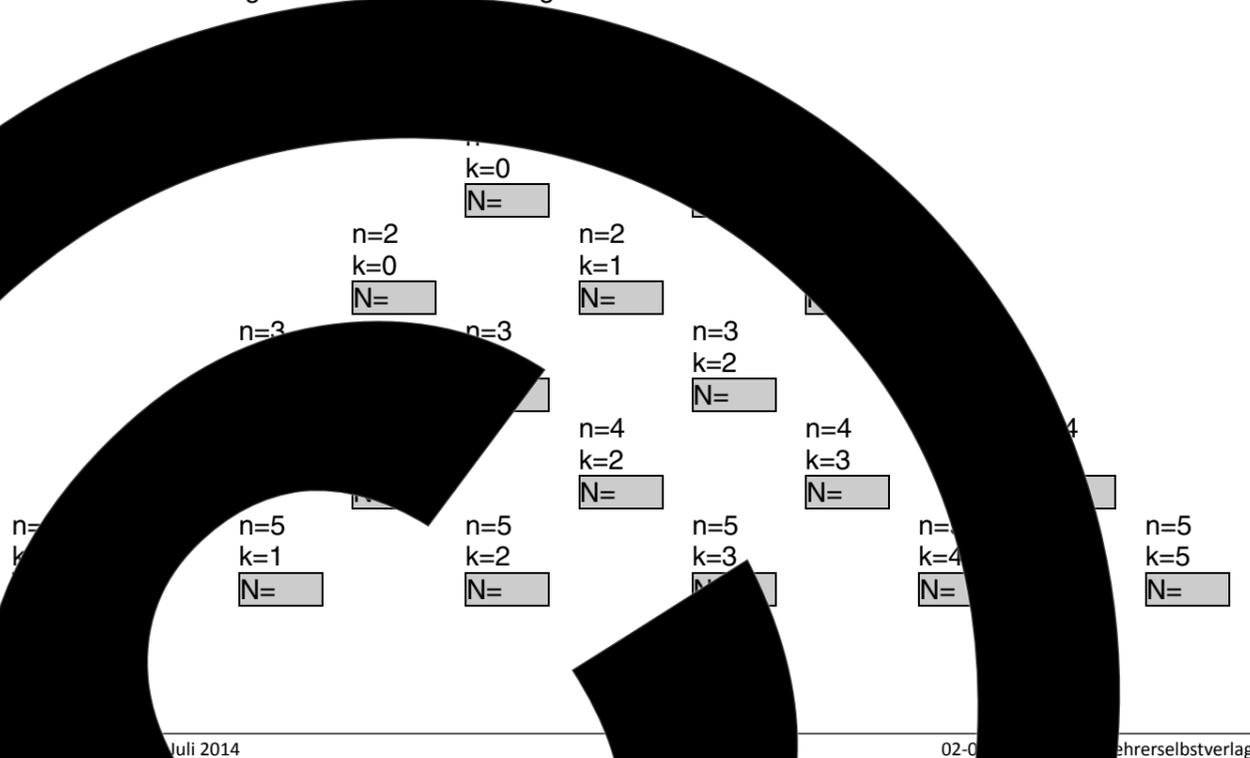
4. In der Urne befinden sich eine rote (r), eine grüne (g) und eine blaue (b) Kugel, $n=3$ Kugeln.

Beachten Sie bei der Bestimmung der Anzahl unterschiedlicher Kombinationen jeweils, dass Ereignisse wie rg und gr nicht unterschieden werden. Jedes Ereignis somit nur als eine Kombinationsmöglichkeit gewertet werden.

keine Kugel ziehen	$k = \underline{\quad}$	$\Rightarrow N = \underline{\quad}$
eine Kugel ziehen	$k = \underline{\quad}$	$\Rightarrow N = \underline{\quad}$
zwei Kugeln ziehen	$k = \underline{\quad}$	$\Rightarrow N = \underline{\quad}$
drei Kugeln ziehen	$k = \underline{\quad}$	$\Rightarrow N = \underline{\quad}$

5. Ergänzen Sie zunächst im folgenden Schema an den entsprechenden Stellen die in Aufgabenteil 1. bis 4. erhaltenen Werte für N ein.

Wie Ihnen vielleicht schon bekannt ist, bilden die Werte von N das sogenannte Pascalsche Dreieck. Informieren Sie sich in geeigneter Weise, wie das Schema für $n=4$ und $n=5$ fortgesetzt wird, und bestimmen Sie diese Werte ohne Verwendung weiterer Betrachtungen von Ziehungen aus Urnen bzw. ohne Erstellung von weiteren Baumdiagrammen.



6. Verallgemeinerungen

Die Werte von N werden auch als **Binomialkoeffizienten** bezeichnet. Hierfür verwendet man die Schreibweise $\binom{n}{k}$ (sprich: „n über k“). Damit ergeben sich bei $n=5$ beispielsweise die oben berechneten Werte:

$$n=5, k=0 \Rightarrow \binom{5}{0} = 1 \quad n=5, k=1 \Rightarrow \binom{5}{1} = 5 \quad n=5, k=2 \Rightarrow \binom{5}{2} = 10 \quad \dots \quad n=5, k=5 \Rightarrow \binom{5}{5} = 1$$

Rechnerische Ermittlung der Binomialkoeffizienten

Die Werte für die Binomialkoeffizienten $\binom{n}{k}$ können auf drei verschiedene Arten ermittelt werden.

a) Mithilfe des Pascalschen Dreiecks

b) Mithilfe des nCr -Taste* des Taschenrechners, Bsp.: $\binom{10}{2}$ Eingabe: 5 nCr 2 ergibt 10

$\binom{10}{4}$ Eingabe: 10 nCr 4 ergibt 210

c) Mithilfe der Formel $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ Bsp.: $\binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = \frac{5 \cdot 2}{1} = 10$

*Englisch: Combinations

7. Berechnen Sie mit dem Taschenrechner, wie viele Möglichkeiten es gibt, eine Mischung aus 8 verschiedenen Saftsorten zu wählen, wenn die Reihenfolge der Auswahl keine Rolle spielt.

$\binom{\quad}{\quad} = \underline{\quad}$

8. In einem Getränkemarkt gibt es 8 verschiedene Sorten Fruchtsaft. Ein Kunde kauft vier Flaschen. Berechnen Sie die Anzahl der Möglichkeiten, eine Mischung aus 4 verschiedenen Saftsorten zusammenzustellen, wenn jede Sorte nur eine Flasche gekauft werden darf.

Die Übungen: _____

In einem Behälter befindet sich eine begrenzte Anzahl n von beliebigen Gegenständen, die man in n_1 Gegenstände mit einem Merkmal 1 und einem Merkmal 2 einteilen kann, wobei es keine Rolle spielt, ob die Gegenstände mit dem jeweiligen Merkmal voneinander unterschieden werden können. Man entnimmt k dieser Gegenstände und berechnet die Wahrscheinlichkeit, wie viele Teile mit dem Merkmal 1 bzw. 2 in der Stichprobe enthalten sind.

Verallgemeinerter Ansatz:

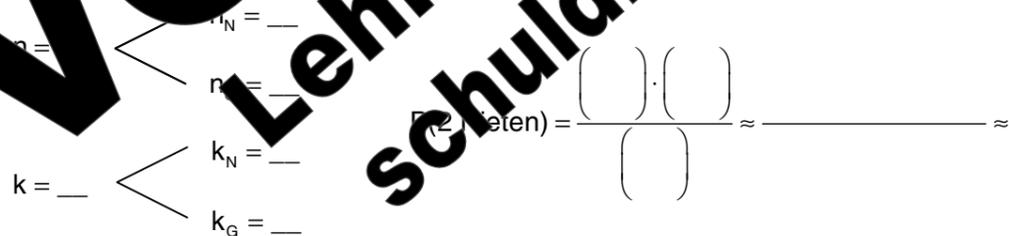
Aufteilen der Gesamtanzahl n bzw. aufteilen der Ziehungen k in die jeweilige Anzahl mit Merkmal 1 und 2.



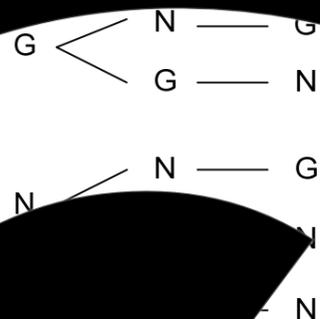
a) Wenden Sie diesen Ansatz auf folgende Aufgabe an, und überprüfen Sie Ihr Ergebnis mithilfe eines Baumdiagramms (Beachten Sie dabei, dass sich um Ziehen ohne Zurücklegen handelt).

In einer Lostrommel befinden sich 20 Lose: 4 Gewinne und 16 Nieten. Zeigen Sie, dass die Wahrscheinlichkeit, unter 5 gezogenen Losem genau 2 Gewinne zu erhalten, bei 15% liegt.

Abkürzungen: Gewinn G, Nieten N



Baumdiagramm



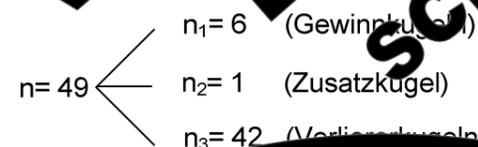
b) Vergleichen Sie den Rechenaufwand der beiden Wege, und beurteilen Sie die Möglichkeit, die Wahrscheinlichkeit, bei 8 gekauften Lose 5 Nieten zu erhalten, über ein Baumdiagramm zu ermitteln.

Ergänzende Übungen: _____

Aufgabe 6

Freizeit: Betrachtungen für Urnen mit mehr als zwei Merkmalen

Beim Lotto wird bei der Ziehung der Zusatzzahl mit ein, so hat man in der Trommel Kugeln mit drei Merkmalen: Gewinnkugel, Zusatzkugel und Verliererkugel. In diesem Fall kann man das Wahrscheinlichkeitschema wie folgt erweitern:



$$P(A) = \frac{\binom{n_1}{k_1} \cdot \binom{n_2}{k_2} \cdot \binom{n_3}{k_3}}{\binom{n}{k}} \quad \text{mit} \quad \begin{matrix} n_1 + n_2 + n_3 = n \\ k_1 + k_2 + k_3 = k \end{matrix}$$

Durch Einsetzen in die obige modifizierte Formel lässt sich die Wahrscheinlichkeit ermitteln.

Zeigen Sie, dass die Wahrscheinlichkeit, bei 5 gezogenen Lose „5 Richtige mit Zusatzzahl“ zu erhalten, eine Chance von etwa 1 zu 2,3 Millionen beträgt.

(mit Zusatzzahl) =

Raum für Notizen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Einführung in den Bereich
der Wahrscheinlichkeiten



Einführung in die Stochastik

Kapitel 1
 Grundbegriffe 7

Kapitel 2
 Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
 Vierfeldertafel 23

Kapitel 4
 Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
 Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
 Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
 Erwartungswert 63

Kapitel 8
 Varianz und Standardabweichung 69

Kapitel 9
 Normalverteilung 75

Hypothesentests

Kapitel 10
 Zweiseitiger Signifikanztest 97

Kapitel 11
 Einseitiger Signifikanztest – Fisher beim Testen 109

Kapitel 12
 Vertrauensintervall 135

..... 139

Kapitel 14
 Anwendung der Normalverteilung 149

Gesamtwerk selbstorganisiert erlernen (Bestenfalls 278)

Alle Rechte vorbehalten. All rights reserved. Nachdruck, auch auszugsweise, vorbehaltlich der Rechte, aus §§ 53, 54 UrhG ergeben, nicht gestattet

SelbstVerlag
 ...tes & Freunde GmbH, Koblenz (Germany)
 Lehrerselbstverlag.de
 www.f-druck.de

Kapitel 5: Einführung in den Begriff der bedingten Wahrscheinlichkeit

Im Kapitel 3 haben Sie sich im Zusammenhang mit den Vierfeldertafeln bereits mit stochastisch abhängigen Zufallsversuchen befasst.

Aufgabe 5.1

a) Nennen Sie zur Wiederholung, mit welchem Kriterium man in einem Baumdiagramm bzw. in einer Vierfeldertafel erkennen kann, dass ein Zufallsversuch stochastisch unabhängig ist.

Kriterium bei einem Baumdiagramm:

Kriterium bei einer Vierfeldertafel:

Bei der Zufallsabhängigkeit mit stochastisch abhängigen Zufallsversuchen entstehen Fragestellungen, die im Begriff der **bedingten Wahrscheinlichkeit** führen. Verdeutlichen Sie sich die Vierfeldertafel für das folgende Beispiel und bestätigen Sie durch Rechnung, dass die Wahrscheinlichkeit an Grippe zu erkranken vom Geschlecht stochastisch unabhängig ist:

Beispiel:

In einer Region, in der 45% der Bevölkerung männlich sind, haben sich 35% der gesamten Bevölkerung mit einem neuen gefährlichen Grippevirus infiziert. Mithilfe der Angabe, dass das Merkmal eine Erkrankung an Grippe mit einer Wahrscheinlichkeit von 15% auftritt und den Allereignissen M und \bar{M} zugeordnet werden können, stellen Sie sich die Vierfeldertafel für die Ereignisse M und G vor.

Verwenden Sie die Formeln der Additionsregel die folgenden Wahrscheinlichkeiten zu bestimmen:

Vierfeldertafel

		G	
M	$P(M \cap G) = 0,2$	$P(M \cap \bar{G}) = 0,25$	$P(M) = 0,45$
\bar{M}	$P(\bar{M} \cap G) = 0,15$	$P(\bar{M} \cap \bar{G}) = 0,40$	$P(\bar{M}) = 0,55$
	$P(G) = 0,35$	$P(\bar{G}) = 0,65$	1

Bestimmen Sie die bedingten Wahrscheinlichkeiten $P(G|M)$ und $P(G|\bar{M})$ für die stochastischen Abhängigkeit:

Aufgabe 5.2

Erstellen von Baumdiagrammen zu der Vierfeldertafel aus Aufgabe 5.1

Zu dieser Vierfeldertafel kann man unter zwei Gesichtspunkten ein Baumdiagramm erstellen, da die erste Stufe des Baumes das Merkmal männlich M oder nicht männlich M-bar, das zweite Merkmal an Grippe erkrankt G oder nicht an Grippe erkrankt G-bar enthalten kann. Ergänzen Sie in den beiden Baumdiagrammen die fehlenden Ereignisse.



Es spielt keine Rolle, ob man hier P(M ∩ G) oder P(G ∩ M) etc. schreibt. Das Ergebnis wird durch das Produkt der Pfadwahrscheinlichkeiten gewonnen und es gilt das Kommutativgesetz. Man kann die Schreibweise so verwenden, wie sie in der Vierfeldertafel verwendet wird.

Beschreibung der Pfade mit Wahrscheinlichkeiten

Da es sich um einen stochastisch abhängigen Zufallsversuch handelt, hängen die Wahrscheinlichkeiten in der zweiten Stufe des Baumes vom Ausgang des Zufallversuchs in der ersten Stufe ab. D.h., dass beispielsweise die Wahrscheinlichkeit P(G) an Grippe zu erkranken in der zweiten Stufe des linken Baumdiagramms davon abhängt, ob in der ersten Stufe das Merkmal M oder das Merkmal M-bar eingetreten ist. Damit ist die Wahrscheinlichkeit in der zweiten Stufe nur den Ausdruck P(G|M) bzw. P(G|M-bar). Die Schreibweise* für die formale Darstellung der Baumdiagramme im Folgenden auf...

1. Stufe: Merkmal Mann oder Frau (nicht Mann)

- Angabe der Wahrscheinlichkeit an Grippe zu erkranken unter der Bedingung, dass man - ein Mann ist P_M(G) - eine Frau ist P_M-bar(G)

- Angabe der Wahrscheinlichkeit an Grippe zu erkranken unter der Bedingung, dass man - ein Mann ist P_G(M) - eine Frau ist P_G(M-bar)

*Anstatt der indizierten Schreibweise P_M(G) bzw. P_M-bar(G) verwendet man auch P(G|M) bzw. P(G|M-bar)

2. Stufe: Merkmal an Grippe erkrankt oder nicht erkrankt

- Angabe der Wahrscheinlichkeit ein Mann zu sein unter der Bedingung, dass man - an Grippe erkrankt ist P_G(M) - nicht an Grippe erkrankt ist P_G(M-bar)

- Angabe der Wahrscheinlichkeit ein Mann zu sein unter der Bedingung, dass man - an Grippe erkrankt ist P_M(G) - nicht an Grippe erkrankt ist P_M-bar(G)

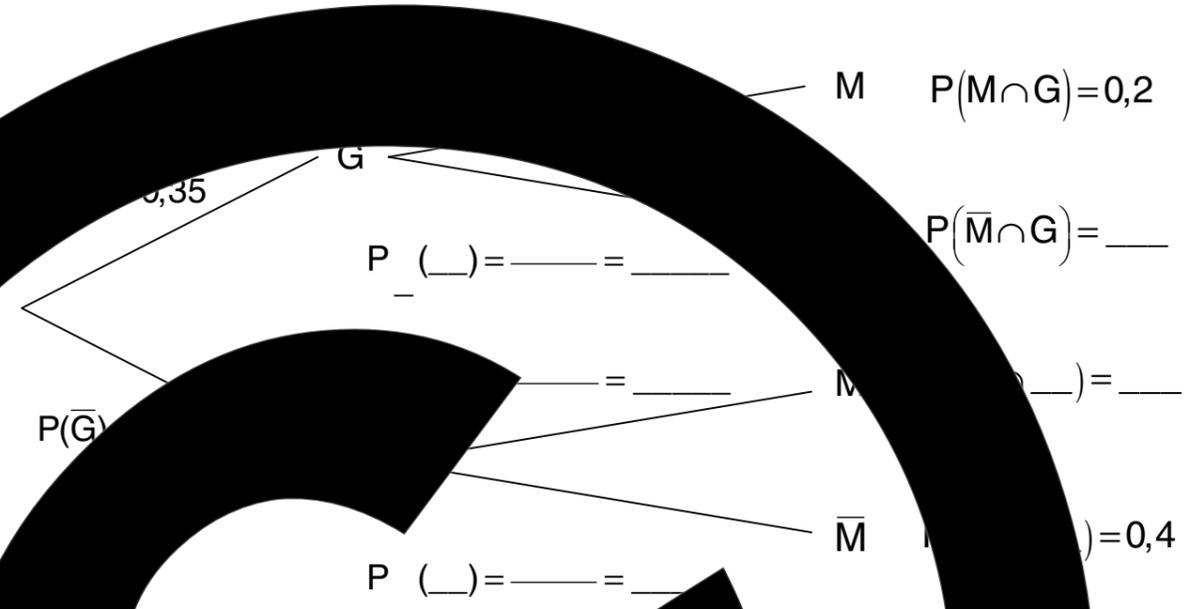
Da die Wahrscheinlichkeiten in der zweiten Stufe beider Baumdiagramme von der Bedingung abhängen, welches Ereignis in der ersten Stufe eingetreten ist, spricht man von bedingten Wahrscheinlichkeiten.

Ergänzen Sie mithilfe der Vierfeldertafel aus Aufgabe 5.1 in beiden Baumdiagrammen die Beschriftung der Pfade, und ermitteln Sie, wie im Beispiel P_M(G) mithilfe der Pfadwahrscheinlichkeitsregel die Wahrscheinlichkeiten für alle Verzweigungen an den Ästen.

Baumdiagramm für die Bedingung M und M-bar in der ersten Stufe:



Baumdiagramm für die Bedingung G und G-bar in der ersten Stufe:



Aufgabe 5.4

Kreuzen Sie jeweils an, ob die Aussagen richtig oder falsch sind.

- a) Die Wahrscheinlichkeiten an den Ästen der 1. Stufe in beiden Baumdiagrammen stimmen mit den entsprechenden Werten der äußeren Felder der Vierfeldertafel überein. ja nein
- b) Die Wahrscheinlichkeiten der Ereignisse am Ende der 1. Stufe in den inneren Werten der Vierfeldertafel und sind daher in beiden Baumdiagrammen identisch. ja nein
- c) Die vier Wahrscheinlichkeiten an den Verzweigungen der Äste in der 2. Stufe in einem Baumdiagramm entsprechen den Werten der äußeren Felder der Vierfeldertafel. ja nein
- d) Die vier Wahrscheinlichkeiten der Ereignisse in einem Baumdiagramm kann man der Vierfeldertafel nicht direkt entnehmen. Für die Werte errechnet man durch entsprechende Quotientenbildung an den inneren und äußeren Feldern der Vierfeldertafel. ja nein
- e) Der Ausdruck $P_G(M)$ gibt die bedingte Wahrscheinlichkeit der zweiten Stufe eines Baumdiagramms des ersten Stufe sich auf das Merkmal Mann oder Frau bezieht. ja nein
- f) Der Ausdruck $P_G(M)$ die bedingte Wahrscheinlichkeit der zweiten Stufe eines Baumdiagramms lassen erste Stufe sich auf das Merkmal an Grippe erkrankt zu sein bezieht. ja nein
- g) Der Ausdruck $P_M(G)$ bedeutet, dass man die Wahrscheinlichkeit Grippe zu haben untersucht, unter der Bedingung eine Frau zu sein. ja nein
- h) Der Ausdruck $P_M(G)$ bedeutet, dass man die Wahrscheinlichkeit eine Frau zu sein untersucht, unter der Bedingung Grippe zu haben. ja nein
- i) Der Ausdruck $P_G(M)$ bedeutet, dass man die Wahrscheinlichkeit Grippe zu haben untersucht, unter der Bedingung ein Mann zu sein. ja nein
- j) Die Wahrscheinlichkeit der zweiten Stufe im Baum enthält die Bedingung, von der die Wahrscheinlichkeit der zweiten Stufe im Baum abhängt. ja nein
- k) Die bedingte Wahrscheinlichkeit in der zweiten Stufe wird berechnet, indem die Wahrscheinlichkeit für das Merkmal in der ersten Stufe durch die Wahrscheinlichkeit für das Merkmal in der ersten Stufe geteilt wird. ja nein
- m) Die bedingte Wahrscheinlichkeit in der zweiten Stufe wird berechnet, indem die Wahrscheinlichkeit für das Merkmal in der zweiten Stufe durch die Wahrscheinlichkeit für das Merkmal in der ersten Stufe geteilt wird. ja nein

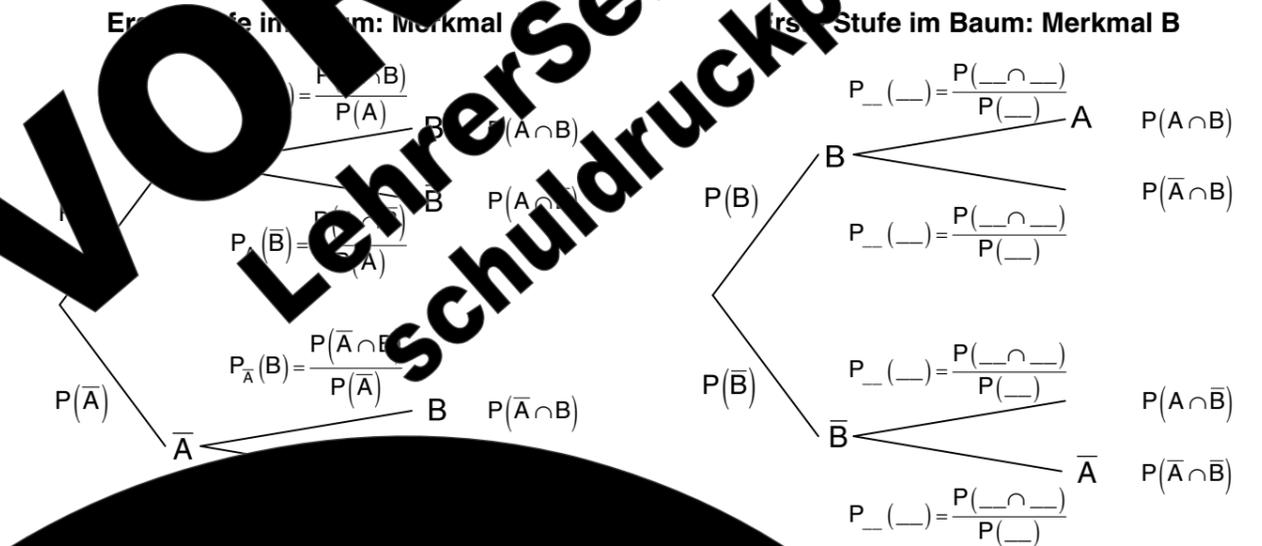
Aufgabe 5.5

Verallgemeinerung des Zusammenhangs zwischen der Vierfeldertafel und den zugehörigen Baumdiagrammen

	B		
A	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
\bar{A}	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
	$P(B)$	$P(\bar{B})$	1

In der Vierfeldertafel kann man die bedingte Wahrscheinlichkeiten, welche in der zweiten Stufe in dem Baumdiagramm erscheinen, nicht direkt ablesen.

Für die Berechnung dieser bedingten Wahrscheinlichkeiten der zweiten Stufe in beiden Baumdiagrammen bildet man jeweils den Quotienten aus der Wahrscheinlichkeit des Ereignisses und der Wahrscheinlichkeit des zugehörigen Merkmals in der ersten Stufe des Zufallsexperimentes. Ergänzen Sie fehlende Angaben.



Skizzieren Sie, wie Sie bedingte Wahrscheinlichkeiten durch die Vierfeldertafel, und damit ohne ein Baumdiagramm, ermitteln können.

Aufgabe 5.6

Standardformulierungen bei Aufgabenstellungen zu bedingten Wahrscheinlichkeiten

Die Interpretation der Aufgabenstellungen bei bedingten Wahrscheinlichkeiten ist oft schwierig. Daher sollen hier anhand des Beispiels Grippeerkrankung ...

1. Fragestellungen, die sich auf die Gesamtheit der Merkmale beziehen

Bei Fragestellungen zu Wahrscheinlichkeiten, welche sich auf die Gesamtheit von Männern und Frauen beziehen, besteht zwischen den beiden Baumdiagrammen kein Unterschied.

- Mit welcher Wahrscheinlichkeit erkrankt man an Grippe (Mann und Grippe): P(M ∩ G) = 0,2
• Mit welcher Wahrscheinlichkeit hat man eine Grippe (Mann und Grippe): P(G) = ...
• Mit welcher Wahrscheinlichkeit ist ein Mann gesund (Mann und Gesund): P(M ∩ G-bar) = ...
• Mit welcher Wahrscheinlichkeit ist eine Frau gesund (Frau und Gesund): P(F ∩ G-bar) = ...

2. Fragestellungen, die sich auf die bedingte Wahrscheinlichkeit der 2. Stufe im Baum beziehen

Bei den folgenden Fragestellungen (Standardformulierungen in Aufgaben) wird bei gegebener Wahrscheinlichkeit des Ereignisses ...

Textalternativen zur Aufgabenstellung:

- Mit welcher Wahrscheinlichkeit trifft man unter den an Grippe Erkrankten auf einen Mann?
• Wie groß ist die Wahrscheinlichkeit, dass man als Mann an Grippe erkrankt?
• Wie groß ist die Wahrscheinlichkeit an Grippe zu erkranken, wenn man ein Mann ist?
• Wie groß ist, unter der Bedingung ein Mann zu sein, die Wahrscheinlichkeit an Grippe zu erkranken?

Lösungsansatz: P_G(M) = P(M ∩ G) / P(G) = 0,2 / 0,45 = 0,44

Berechnen Sie die gesuchten Wahrscheinlichkeiten, und formulieren Sie zusätzlich zur gegebenen Fragestellung eine zweite Fragestellung, welche das Wort „wenn“ enthält.

P_M(G-bar) = P(M ∩ G-bar) / P(M) = ... = ...

P_G(M) = P(M ∩ G) / P(G) = ... = ...

- Mit welcher Wahrscheinlichkeit trifft man unter den Männern auf einen nicht erkrankten.
• Wie groß ist die Wahrscheinlichkeit ...

P_M(G) = P(M ∩ G) / P(M) = ... = ...

P_G(M-bar) = P(M-bar ∩ G) / P(G) = ... = ...

- Mit welcher Wahrscheinlichkeit trifft man unter den Frauen auf eine erkrankte.
• Wie groß ist die Wahrscheinlichkeit ...

Formulieren Sie eine zweite Fragestellung, und ermitteln Sie mit dem passenden mathematischen Ansatz die gesuchte Wahrscheinlichkeit.

Ansatz und Rechenweg:

Ansatz und Rechenweg:

Aufgabe 5.7

Anhand des Beispiels in dieser Aufgabe soll eine weitere Hilfe für die Entschlüsselung Aufgabentexten bei bedingten Wahrscheinlichkeiten gegeben werden.

Zahlenbeispiel:

In Deutschland sind etwa 0,5% der Bevölkerung an einem seltenen Virusmerkel Krebs erkrankt. Man weiß, dass ein spezieller Bluttest 90% der Gesunden und 99% der Erkrankten richtig diagnostiziert.

- a) Wie viel Prozent der Tests fallen positiv aus?
- b) Bei einer Routinekontrolle zeigt der Test ein negatives Ergebnis an. Berechnen Sie die Wahrscheinlichkeiten, dass eine Person bei diesem positiven Ergebnis nicht erkrankt ist.
- c) Bei einer Routinekontrolle zeigt der Test ein negatives Ergebnis an. Berechnen Sie die Wahrscheinlichkeit, dass man trotzdem an dem Krebs erkrankt ist.

Lösungsschritte

Schritt 1:

Festlegen der Merkmale und Abkürzungen:

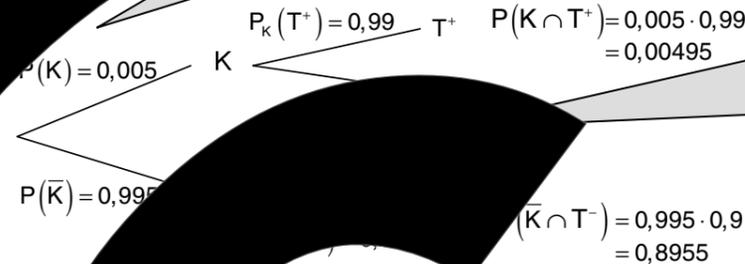
- Person hat Krebs: K
- Person hat keinen Krebs: \bar{K}
- Test ist positiv: T^+
- Test ist negativ: T^-

Schritt 2:

Bauen Sie das Baumdiagramm durch Zuordnung der Zahlenangaben zu der ersten bzw. zweiten Stufe des Baumdiagramms erstelle.

Hier muss analysiert werden, welche Zahlenangaben sich auf die erste Stufe eines Baumdiagramms beziehen.

Da sich die Zahlenangaben auf die Bevölkerung beziehen, sind dies die Werte für die erste Stufe des Baumdiagramms.



Die Zahlenangaben zum Bluttest beziehen sich auf die Bedingung, ob eine Person Krebs oder nicht an Krebs erkrankt ist, und sind damit die Werte für die zweite Stufe im Baumdiagramm.

Schritt 3:

Erstellen einer Vierfeldertafel durch Übertragen der Werte aus dem Baumdiagramm. Vervollständigung der Felder mithilfe der Additionsregel.

	T^+	T^-	
K	0,00495	0,00005	0,005
\bar{K}	0,0995	0,8955	0,9045
	0,10445	0,89555	1

Schritt 4:

Analyse der Aufgabenstellung und Berechnung der gesuchten Wahrscheinlichkeiten mit der Vierfeldertafel und gegebenen Formeln für die bedingten Wahrscheinlichkeiten. Dazu kann man die Aufgabenstellung, den Vierfeldersatz und die Additionsregel nutzen.

zu a) Gesucht ist die Wahrscheinlichkeit $P(T^+)$. Das Ergebnis kann sofort aus der Vierfeldertafel abgelesen werden.

$P(T^+) = 0,10445 \approx 10,45\%$

zu b) Mit welcher Wahrscheinlichkeit ist eine Person nicht erkrankt, wenn der Test ein positives Ergebnis anzeigt?

$P_{T^+}(\bar{K}) = \frac{P(\bar{K} \cap T^+)}{P(T^+)} = \frac{0,0995}{0,10445} \approx 95,3\%$

Die Werte für die Berechnung können der Vierfeldertafel entnommen werden.

Bei einem positiven Test ist man mit einer Wahrscheinlichkeit von mehr als 95% nicht erkrankt.

zu c) Mit welcher Wahrscheinlichkeit ist eine Person erkrankt, wenn der Test ein negatives Ergebnis anzeigt?

$P_{T^-}(K) = \frac{P(K \cap T^-)}{P(T^-)} = \frac{0,00005}{0,89555} \approx 0,006\%$

Ein negativer Test kann als sicherer Hinweis, dass man auch nicht erkrankt ist, angesehen werden, da man bei einem negativen Testergebnis nur mit einer Wahrscheinlichkeit von 0,006% trotzdem erkrankt ist.

Die Werte für die Berechnung können der Vierfeldertafel entnommen werden.

Ergänzung:

Anstatt die Vierfeldertafel kann man auch ein Baumdiagramm entwickeln, dessen erste Stufe die Merkmale T^+ und T^- hat, da sich die Aufgabenstellung auf ein Merkmal bezieht, in dessen Hinsicht man nicht mehr die Merkmale K und \bar{K} , sondern die Merkmale T^+ und T^- betrachten muss.

Aufgabe 5.8

Bei 4,5 Millionen Wohneinheiten in Deutschland werden jährlich 200 000 Wohnungsbrände gemeldet. 600 Brände mit tödlichen Folgen für Menschen führen zu Überlegungen, ob Rauchmelder in allen Wohnungen Pflicht werden sollten. Jedoch funktionieren Rauchmelder in Deutschland nicht zu 100% fehlerfrei. Ein handelsüblicher Rauchmelder bei einem Schnäppchenangebot löst im Brandfall mit einer WK von 99,8% einen Alarm aus. Es kann jedoch auch nicht ausgeschlossen werden, dass der Rauchmelder einen Fehlalarm aussendet, wenn es nicht brennt. Solch ein Fehler wird einmal jährlich also mit einer Wahrscheinlichkeit von 1:365 erwartet.

- a) Begründen Sie, warum sich die gegebenen Werte auf ein Baumdiagramm beziehen und dem die Merkmale Brand B und nicht Brand \bar{B} der ersten Stufe zugeordnet werden. Erstellen Sie dieses Baumdiagramm, und ermitteln Sie mit dessen Hilfe eine zugehörige Vierfeldertafel. Verwenden Sie für das Merkmal Alarm A und kein Alarm \bar{A} .

			1

- b) Ergänzen Sie die Vierfeldertafel für die Merkmale Alarm A und kein Alarm \bar{A} der ersten Stufe und Brand B und kein Brand \bar{B} der zweiten Stufe. Erläutern Sie die Zusammenhänge zwischen den beiden Baumdiagrammen und der Vierfeldertafel.

- c) Ermitteln Sie die Wahrscheinlichkeit, dass der Brandmelder im Laufe eines Jahres einen Alarm aussendet.

- d) Berechnen Sie die Wahrscheinlichkeit, dass es bei dem Auslösen des Alarms wirklich brennt. Formulieren Sie diese Aufgabenstellung dazu zunächst mithilfe eines „Wenn-Sollens“.

$$P_{\text{---}}(\text{---}) = \frac{P(\text{---})}{P(\text{---})} = \text{---} \%$$

- e) Berechnen Sie die Wahrscheinlichkeit für einen unbemerkten Wohnungsbrand, weil der Brandmelder keine Alarmmeldung gibt. Formulieren Sie diese Aufgabenstellung dazu auch hier zunächst mithilfe eines „Wenn-Sollens“ und beurteilen Sie anschließend die Sicherheit des Brandmelders.

$$P_{\text{---}}(\text{---}) = \frac{P(\text{---})}{P(\text{---})} = \text{---} = \text{---} = \text{---} \%$$

Beurteilung der Sicherheit:

Aufgabe 5.9

Ergänzende Betrachtung zu Mehrfeldertafeln

In einer Studie zur Berufstätigkeit von Frauen werden in Zufallsland folgende Zahlen erhoben: Von den 34% kinderlosen Frauen sind 95% berufstätig. Während von den 25% der Frauen, die mit einem Kind noch 30% einer Berufstätigkeit nachgehen sind, geben von den 25% der Frauen mit zwei Kindern 80% ihre Berufstätigkeit auf. Unter den Frauen mit drei oder mehr Kindern üben nur noch 10% ihren Beruf aus.

- a) Zeigen Sie, dass in Zufallsland 53,1% der Frauen nicht berufstätig sind.
- b) Zeigen Sie, dass die Wahrscheinlichkeit, unter den berufstätigen Frauen eine Frau mit zwei oder mehr Kindern anzutreffen, nur bei 1% liegt.

Lösungshinweise:

Bei Baumdiagrammen mit mehr als zwei Stufen bzw. bei Mehrfeldertafeln mit mehr als zwei Ästen und zwei Stufen kann man analog zu Baumdiagrammen mit zwei Ästen und zwei Stufen sog. Mehrfeldertafeln erstellen. Tragen Sie im Baumdiagramm an den markierten Stellen die in der Aufgabe gegebenen Zahlenwerte ein und füllen Sie mithilfe dieser Zahlenwerte die Mehrfeldertafel aus. Bestimmen Sie dann die gesuchten Wahrscheinlichkeiten in Aufgabe a) durch Ablesen aus der Mehrfeldertafel und in Aufgabe b) durch Rechnung mit der Vierfeldertafel-entsprechende Formel.



zu b)

Aufgabe 5.10

Ergänzende Betrachtungen zum Begriff der totalen Wahrscheinlichkeit

- a) Verdeutlichen Sie sich anhand der beiden Baumdiagramme aus Aufgabe 5.9 die folgenden Zusammenhänge und den Begriff der **totalen Wahrscheinlichkeit**.

<p>Die erste Stufe im linken Baum in Aufgabe 5.9 bezeichnet Merkmal A.</p> <p>Aus dem Baumdiagramm kann man nur die Wahrscheinlichkeiten der ersten Stufe $P(B)$ und $P(\bar{A})$ direkt ablesen.</p> <p>$P(B)$ und $P(\bar{B})$ können nur durch Addition der entsprechenden Ereignisse berechnet werden.</p> <p>$P(B) = P(A \cap B) + P(\bar{A} \cap B)$ oder $P(B) = P(B A) \cdot P(A) + P(B \bar{A}) \cdot P(\bar{A})$</p> <p>Man bezeichnet die hier berechnete Wahrscheinlichkeit $P(B)$ als totale Wahrscheinlichkeit $P(B)$.</p>	<p>Die erste Stufe im rechten Baum in Aufgabe 5.9 bezeichnet Merkmal B.</p> <p>Aus dem Baumdiagramm kann man nur die Wahrscheinlichkeiten der ersten Stufe $P(B)$ und $P(\bar{B})$ direkt ablesen.</p> <p>$P(A)$ und $P(\bar{A})$ können nur durch Addition der entsprechenden Ereignisse berechnet werden.</p> <p>$P(A) = P(A \cap B) + P(A \cap \bar{B})$ oder $P(A) = P(B A) \cdot P(A) + P(\bar{B} A) \cdot P(A)$</p> <p>Man bezeichnet die hier berechnete Wahrscheinlichkeit $P(A)$ als totale Wahrscheinlichkeit $P(A)$.</p>
---	--

- b) Ergänzen Sie den folgenden Satz, und geben Sie mithilfe der beiden Baumdiagramme aus Aufgabe 5.9 die totalen Wahrscheinlichkeiten $P(\bar{B})$ und $P(\bar{A})$ entsprechend der Definition im Kasten an.

Man berechnet die totale Wahrscheinlichkeit $P(\bar{B})$ bzw. $P(\bar{A})$, indem man im entsprechenden Baum aus Aufgabe 5.9 alle Pfade, die zum Ereignis ___ bzw. ___ führen,

Wenn die 1. Stufe im Baum das Merkmal A bezeichnet, gilt:

$P(\bar{B}) =$ _____

Wenn die 1. Stufe im Baum das Merkmal B bezeichnet, gilt:

$P(\bar{A}) =$ _____

$P(\bar{A}) =$ _____

Zusammenfassung zur Berechnung von bedingten Wahrscheinlichkeiten

Die Aufgabenstellung bezieht sich auf ein Baumdiagramm mit dem Merkmal A in der ersten Stufe.

Häufig sind bei Aufgabenstellungen zu bedingten Wahrscheinlichkeiten Angaben zum Merkmal A also P(A) für die erste Stufe des Baumes und Wahrscheinlichkeiten für Ereignisse oder bedingte Wahrscheinlichkeiten der zweiten Stufe im Baum gegeben, so dass man die Vierfeldertafel ausfüllen kann. Gefragt wird dann in der Regel nach einer bedingten Wahrscheinlichkeit, hier die Wahrscheinlichkeit des Merkmals B, also P(B), der ersten Stufe des Baumdiagramms verknüpft wird, also die Wahrscheinlichkeit für A unter der Bedingung, dass eingetreten ist, berechnet werden soll. Die gesuchte Wahrscheinlichkeit kann dann mit der bekannten Formel berechnet werden.

Die Aufgabenstellung bezieht sich auf ein Baumdiagramm mit dem Merkmal B in der ersten Stufe.

Berechnung der bedingten Wahrscheinlichkeit P_B(A) = $\frac{P(A \cap B)}{P(B)}$

Gleichwertig dazu ist die Formulierung von Bayes, bei der jeweils die totalen Wahrscheinlichkeiten eingesetzt werden.

Bayes $P_B(A) = \frac{P(A) \cdot P_A(B)}{P(A) \cdot P_A(B) + P(\bar{A}) \cdot P_{\bar{A}}(B)}$



Oberstudienrätin Ursula Pirkl

Stochastik selbstorganisiert lernen



Kapitel 5

Definition der Wahrscheinlichkeitsverteilung



Einführung in die Stochastik

Kapitel 1	
Grundbegriffe	7
Kapitel 2	
Baumdiagramme und mehrstufige Zufallsversuche	13
Kapitel 3	
Vierfeldertafel	23
Kapitel 4	
Kombinatorische Abzählverfahren	35
Bedingte Wahrscheinlichkeiten	
Kapitel 5	
Einführung in den Begriff der bedingten Wahrscheinlichkeit	47
Kapitel 6	
Definition der Wahrscheinlichkeitsverteilung	61
Wahrscheinlichkeitsverteilung	
Kapitel 7	
Erwartungswert	63
Kapitel 8	
Variation und Standardabweichung	69
Kapitel 9	
Normalverteilung	75
Hypothesentests	
Kapitel 10	
Zweiseitiger Signifikanztest	97
Kapitel 11	
Einseitiger Signifikanztest – Fisher beim Testen	109
Kapitel 12	
Vertrauensintervall	135
Kapitel 13	
Die Betrachtungen zur Normalverteilung	139
Kapitel 14	
Anwendung der Normalverteilung	149

Gesamt: Selbstorganisiert erlernen (Bestenfalls 278)

Alle Rechte vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte,
aus §§ 53, 54 UrhG ergeben, nicht gestattet

SelbstVerlag

Druck & Freunde GmbH, Koblenz (Germany)

www.f-druck.de

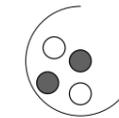
www.f-druck.de

Kapitel 6: Definition der Wahrscheinlichkeitsverteilung

Bei der Berechnung von Wahrscheinlichkeiten mit Baumdiagrammen lernen Sie, dass die Summe der Wahrscheinlichkeiten aller Ereignisse immer den Wert 1 ergibt. Dies ist ein wichtiges Merkmal einer Wahrscheinlichkeitsverteilung bekannt. Es folgt im Hand eines einfachen Beispiels, eine Formalisierung und Erweiterung von Begriffen.

Aufgabe 6.1

Definition der Zufallsvariablen und Darstellung von Wahrscheinlichkeitsverteilungen



Aus einer Urne mit zwei roten und zwei weißen Kugeln werden alle Kugeln ohne Zurücklegen gezogen.

Aufgabenstellung: Berechnen Sie jeweils die Wahrscheinlichkeit, dass man die beiden roten Kugeln nach dem 2. Ziehen, nach dem 3. Ziehen oder erst nach dem letzten Ziehen erhält.

Neue Bezeichnung: Für Vereinfachung und Formalisierung der Schreibweise wird nun eine Zufallsvariable X gegeben. Die Zufallsvariable X gibt hier die Anzahl der Ziehungen an, die man benötigt, damit das Ereignis e , zwei rote Kugeln zu ziehen, eintritt.

$$P(\text{2. rote Kugel im 2. Zug}) = P(X = 2)$$

$$P(\text{2. rote Kugel im 3. Zug}) = P(X = 3)$$

$$P(\text{rote Kugel im 4. Zug}) = P(X = \dots)$$

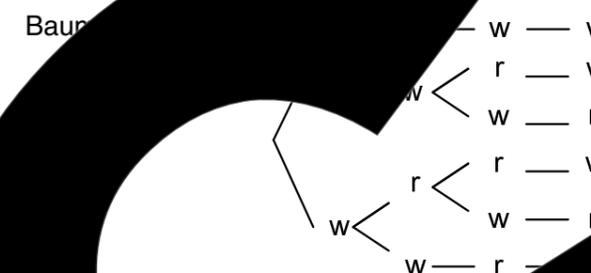
Alternativ:

$$P(e_k) = P(\text{2. rote Kugel im } k. \text{ Zug}) = P(X = k) \text{ für } k = 2, 3, 4$$

die Wahrscheinlichkeit in der Form $P(X = k)$ entsprechend der Aufgabenstellung.

Bearbeiten Sie die folgenden Aufgaben, indem Sie fehlende Werte eintragen.

a) Ermitteln Sie die Wahrscheinlichkeiten an den Ästen des Baumdiagramms und verdeutlichen Sie sich danach die entsprechenden Werte.



Einführung in die Stochastik

Kapitel 1	
Grundbegriffe	7
Kapitel 2	
Baumdiagramme und mehrstufige Zufallsversuche	13
Kapitel 3	
Vierfeldertafel	23
Kapitel 4	
Kombinatorische Abzählverfahren	35
Bedingte Wahrscheinlichkeiten	
Kapitel 5	
Einführung in den Begriff der bedingten Wahrscheinlichkeit	47
Kapitel 6	
Definition der Wahrscheinlichkeitsverteilung	61
Wahrscheinlichkeitsverteilung	
Kapitel 7	
Erwartungswert	63
Kapitel 8	
Variation und Standardabweichung	69
Kapitel 9	
Normalverteilung	75
Hypothesentests	
Kapitel 10	
Zweiseitiger Signifikanztest	97
Kapitel 11	
Einseitiger Signifikanztest – Fisher beim Testen	109
Kapitel 12	
Vertrauensintervall	135
Kapitel 13	
Die Betrachtungen zur Normalverteilung	139
Kapitel 14	
Anwendung der Normalverteilung	149

Gesamtheit selbstorganisiert erlernen (Bestenfalls 278)

Selbstorganisiert erlernen (Bestenfalls 278)

Alle Rechte vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte,
aus §§ 53, 54 UrhG ergeben, nicht gestattet

SelbstVerlag

Verlag des Autors & Freunde GmbH, Koblenz (Germany)

www.f-druck.de

www.f-druck.de

Kapitel 7: Erwartungswert

Aufgabe 7.1

Definition des Erwartungswerts

Zur Information:

Erwartungswerte sind jedem Schüler unter dem Begriff „Mittelwert“ bekannt. Wenn man gefragt wird, welche Zeit man am Morgen durchschnittlich in der Werkstatt für Arbeit benötigt, kann man, wie in der Tabelle unten, aus mehreren Zeitmessungen den Mittelwert bestimmen. Diese Berechnung entspricht prinzipiell der Ermittlung des Erwartungswerts einer Zufallsvariablen bei einem Zufallsversuch.

Berechnung des Mittelwerts aus bekannten Werten

Gemessene Zeit in Minuten	25	30	35	40	45
Absolute Häufigkeit des Messwertes	5	9	10	4	2

Vervollständige die obige Berechnung:

$$\frac{5 \cdot 25 + 9 \cdot 30 + 10 \cdot 35 + 4 \cdot 40 + 2 \cdot 45}{5 + 9 + 10 + 4 + 2} = \frac{500}{38} \approx 13,16 \text{ Minuten}$$

Berechnung des Mittelwerts unter Verwendung der Begrifflichkeit der Wahrscheinlichkeitsrechnung

Verwendet man die Begrifflichkeit der Wahrscheinlichkeitsrechnung, so entspricht jede gemessene Zeit dem Wert x_i , den eine Zufallsvariable X annehmen kann. Die relative Häufigkeit, mit der eine Zeitmessung auftritt, wird als Wahrscheinlichkeit für das Auftreten dieser Zufallsvariable interpretiert. Da bei 30 Messungen die Zeit 40 Minuten auftritt, schreibt man: $P(X = 40) = \frac{4}{30}$. Den Mittelwert aus allen Messungen kann man als Erwartungswert $E(X)$ bezeichnen.

Die Berechnung des Mittelwerts unter Verwendung der Begrifflichkeit der Wahrscheinlichkeitsrechnung dann wie folgt werden:

Werte der Zufallsvariablen x_i	25	30	35	40	45
Wahrscheinlichkeit $P(X = x_i)$ für das Auftreten von x_i	$\frac{5}{30}$	$\frac{9}{30}$	$\frac{10}{30}$	$\frac{4}{30}$	$\frac{2}{30}$

$$E(X) = 25 \cdot \frac{5}{30} + 30 \cdot \frac{9}{30} + 35 \cdot \frac{10}{30} + 40 \cdot \frac{4}{30} + 45 \cdot \frac{2}{30}$$

$$E(X) = \frac{1}{30} (125 + 270 + 350 + 160 + 90) = \frac{995}{30} \approx 33,17 \text{ Minuten}$$

Ergänzen Sie:

Der Erwartungswert $E(X)$ einer Zufallsvariable X wird berechnet, indem man alle Werte x_i der Zufallsvariable X annimmt, jeweils mit den zugehörigen Wahrscheinlichkeiten $P(X = x_i)$ multipliziert und die einzelnen Produkte anschließend addiert.

Allgemein lässt sich der Erwartungswert einer Zufallsvariable X definieren:

Ist X eine Zufallsvariable, welche die Werte x_1, x_2, \dots, x_n annimmt, so heißt die Summe $E(X) = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + \dots + x_n \cdot P(X = x_n)$ Erwartungswert der Zufallsvariable X (der Erwartungswert $E(X)$ wird auch mit dem griechischen Buchstaben μ (Sigma) bezeichnet).

Aufgabe 7.2

Die Wahrscheinlichkeitsverteilung unter dem Gesichtspunkt des Erwartungswerts

Die Aufgabenstellung von Aufgabe 6.1 wird wie folgt erweitert:

Wird nun zusätzlich zu den in Aufgabe 6.1 ermittelten Wahrscheinlichkeiten berechnet, wie viele Züge man bei häufigem Wiederholen des Zufallsversuchs durchschnittlich oder im Mittel benötigt, um zwei rote Kugeln zu erhalten. Ergänzen Sie fehlende Angaben.

Der Erwartungswert beschreibt das Ergebnis, das bei häufiger Durchführung eines Zufallsversuchs zu erwarten ist.

Wahrscheinlichkeitsverteilung aus Aufg. 6.1

x	3	4
$P(X = x_i)$	$\frac{3}{6}$	$\frac{1}{6}$

Berechnung des Erwartungswerts anhand der Definition oben:

$$E(x) = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + x_3 \cdot P(X = x_3)$$

$$E(x) = \dots + \dots + \dots$$

Ergebnis: Man benötigt bei sehr häufiger Durchführung des Spiels durchschnittlich \dots Züge, um zwei rote Kugeln zu erhalten.

Aufgabe 7.3

Wahrscheinlichkeitsverteilung unter dem Gesichtspunkt eines Spielers und des Erwartungswerts.

Eine entscheidende Frage bei Glücksspielen ist nicht nur, ob man gewinnen oder verliert, sondern auch wie viel man gewinnen oder verlieren kann und ob ein Spiel als lohnenswert ist. Betreiber von Spielcasinos müssten Konkurs anmelden, wenn die Betreiber Glücksspiele keinen ausreichenden Gewinn abwerfen, um laufende Kosten abzudecken. Berechnen Sie die folgende Berechnung des Erwartungswerts für den Gewinn bei einem Spiel und übertragen Sie die Ergebnisse auf das anschließende Beispiel.

Die Aufgabenstellung aus Aufgabe 7.2 wird nun so verändert, dass die Berechnung des Gewinns bzw. Verlusts im Mittelpunkt der Betrachtung steht. In einem Spiel mit einem Einsatz von 1,00 € liegt folgender Gewinnplan vor:

Ereignis	Auszahlung
2 rote Kugeln nach der 2. Ziehung	3,00 €
2 rote Kugeln nach der 3. Ziehung	1,00 €
2 rote Kugeln nach der 4. Ziehung	0 €

Für die Gewinn- bzw. Verlustberechnung darf man jetzt nicht nur den ausgezahlten Gewinn laut Gewinnplan betrachten, sondern muss von der Bilanz aus Ein- und Auszahlung ausgehen, d.h. von dem was der Spieler oder der Ausrichter des Spiels nach dem Spiel in der Tasche hat. Man muss sich in der Regel entscheiden, ob die Berechnungen aus Sicht des Spielers oder des Ausrichters erfolgen.

Folgende Tabelle stellt die Gewinn- bzw. Verlustrechnung aus Sicht des Spielers dar:

Ereignis	Auszahlung	Gewinn- /Verlustrechnung
2 rote Kugeln nach der 2. Ziehung	3,00 €	2,00 €
2 rote Kugeln nach der 3. Ziehung	1,00 €	0 €
2 rote Kugeln nach der 4. Ziehung	0 €	-1,00 €

Folgende Tabelle stellt die Gewinn- bzw. Verlustrechnung aus Sicht des Spieleausrichters dar:

Ereignis	Auszahlung	Gewinn- /Verlustrechnung
2 rote Kugeln nach der 2. Ziehung	3,00 €	-2,00 €
2 rote Kugeln nach der 3. Ziehung	1,00 €	0 €
2 rote Kugeln nach der 4. Ziehung	0 €	1,00 €

Wie Sie sehen, ändern Sie nur die Vorzeichen der Zahlen bei der Gewinn-/Verlustrechnung. Das Ergebnis keine Rolle, aus welchem Blickwinkel ein Spiel betrachtet wird. Entscheidend sind die Berechnungen und die entsprechende Wahrscheinlichkeitsverteilung werden dem Spieler durchgeföhrt.

In der Wahrscheinlichkeitsverteilung ändert sich im Vergleich zu Aufgabe 7.2 bei Beachtung monetären Gewinns die Bedeutung der Zufallsvariablen $X = x_i$. Die Zufallsvariable X ordnet nun nicht mehr die Anzahl der Ziehungen an, sondern ordnet den einzelnen x_i den Gewinn bzw. Verlust pro Spiel zu. Für die Beschreibung des eingetretenen Ereignisses beim Ziehen kann man zusätzlich die Variable e_i verwenden. Die Wahrscheinlichkeit für das Eintreten eines Ereignisses e_i bzw. für den Gewinn/Verlust x_i bleibt hinsichtlich der Aufgabe 7.2 unverändert.

Die neue Formulierung der Ereignisse e_i sorgt für mehr Transparenz hinsichtlich der Aufgabenstellung und ist zu empfehlen, aber nicht notwendig.

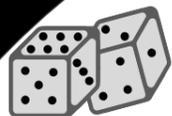
Darstellung des Zufallsversuchs als tabellarische Wahrscheinlichkeitsverteilung			
Auflistung der Ereignisse e_i	e_1 : 2 rote Kugeln nach 2 Ziehungen	e_2 : 2 rote Kugeln nach 3 Ziehungen	e_3 : 2 rote Kugeln nach 4 Ziehungen
Jedem Ereignis e_i wird ein Zahlenwert x_i zugeordnet, der den Gewinn/Verlust darstellt.	2,00 €	0 €	-1,00 €
	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

Der mittlere Gewinn bei vielen durchgeführten Spielen, also der Erwartungswert, wird nun entsprechend zur Definition berechnet.

$$E(X) = x_1 \cdot P(X = x_1) + x_2 \cdot P(X = x_2) + x_3 \cdot P(X = x_3) = 2,00 \text{ €} \cdot \frac{1}{6} + 0 \text{ €} \cdot \frac{2}{6} + (-1,00 \text{ €}) \cdot \frac{3}{6} = -\frac{1}{6} \text{ €} = -0,17 \text{ €}$$

Das bedeutet, dass Sie bei jeder Durchführung des Spiels 17 Cent pro Spiel verlieren.

Aufgabe zur Berechnung des Erwartungswerts bei einem Spiel



Folgendes Glücksspiel wird Ihnen vorgeschlagen:
Es werden drei Würfel geworfen. Sie erhalten die von Ihnen angezeigte Anzahl, wenn diese geradzahlig ist, und Sie erhalten die Hälfte der angezeigten Anzahl, wenn eine ungerade Zahl ergibt.

Entscheiden Sie sich spontan, d.h. ohne die folgenden Rechnungen zu machen, ob Sie sich auf dieses Spiel einlassen würden: ja nein

Übersicht der Augensummen beim Werfen von zwei Würfeln

Ereignis e_i : mögliche Augensumme	Würfelnkombinationen für die möglichen Augensummen: (erster Würfel; zweiter Würfel)	Anzahl der Ereignisse
2	(1;1)	1
3	(1;2)(2;1)	2
4	(1;3)(2;2)(3;1)	3
5	(1;4)(2;3)(3;2)(4;1)	4
6	(1;5)(2;4)(3;3)(4;2)(5;1)	5
7	(2;5)(3;4)(4;3)(5;2)(6;1)	5
8	(3;5)(4;4)(5;3)(6;2)	4
9	(3;6)(4;5)(5;4)(6;3)	4
10	(4;6)(5;5)(6;4)	3
11	(5;6)(6;5)	2
12	(6;6)	1

Tabellarische Wahrscheinlichkeitsverteilung und Berechnung des Erwartungswerts

Ereignis e_i (Augensumme)																			
Zufallsvariable x_i (Einzelauszahlung)																			
$P(X = x_i)$																			

Erwartungswert: _____

Ergebnis: Der Erwartungswert nimmt den Wert Null an. Das heißt, für Sie und den Mitspieler weder ein Gewinn noch ein Verlust. Man erkennt, dass das Spiel fair ist, und weiß nun sicher, auf welches Spiel man sich einlassen kann.

Entscheiden Sie sich spontan, d.h. ohne die folgenden Rechnungen zu machen, ob Sie sich auf dieses Spiel einlassen würden: ja nein

Die Übungsaufgaben: _____

Raum für Notizen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Varianz und Standardabweichung



Einführung in die Stochastik

Kapitel 1
Grundbegriffe 7

Kapitel 2
Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
Vierfeldertafel 23

Kapitel 4
Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
Erwartungswert 63

Kapitel 8
Varianz und Standardabweichung 69

Kapitel 9
Normalverteilung 75

Hypothesentests

Kapitel 10
Zweiseitiger Signifikanztest 97

Kapitel 11
Einseitiger Signifikanztest – Fisher beim Testen 109

Kapitel 12
Vertrauensintervall 135

Kapitel 13
Die Betrachtungen zur Normalverteilung 139

Kapitel 14
Anwendung der Normalverteilung 149

Gesamtheit selbstorganisiert erlernen (Bestenfalls 278)

Copyright vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte,

aus §§ 53, 54 UrhG ergeben, nicht gestattet

SelbstVerlag

Verlag des Lehrers & Freunde GmbH, Koblenz (Germany)

www.lehrerselbstverlag.de

www.f-druck.de

© 2014 Lehrerselbstverlag

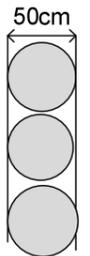
Kapitel 8: Varianz und Standardabweichung

Wie Sie bereits wissen, entspricht die Berechnung des Erwartungswertes $E(X)$ dem Finden eines Mittelwerts für eine Zufallsvariable X . Bei vielen Prozessen in der Arbeit, wie beispielsweise bei Messungen in der Physik oder bei der Herstellung von Bauteilen, reicht es nicht aus, diesen Mittel- bzw. Erwartungswert zu kennen. Vielmehr benötigt man eine Angabe, um und wie viel die gemessenen Werte oder Bauteile von diesem Mittelwert abweichen. Bei Fertigungsprozessen in der Praxis werden diese Abweichungen auch als **Fertigungstoleranzen** bezeichnet. Die Abweichungen vom Erwartungswert werden mithilfe der sog. **Varianz** oder **Standardabweichung**, gewonnen. Verdeutlichen Sie sich anhand von Aufgabe 8.1, wie man diese Größen berechnet. Übertragen Sie Ihre Erkenntnisse dann auf die folgenden Übungsaufgaben.

Aufgabe 8.1

Varianz und Standardabweichung bei Fertigungstoleranzen

Ein Hersteller von kreisförmigen Bauteilen vermutet aufgrund von Beschwerden seiner Kunden, dass seine Schleifmaschine nicht mehr exakt gearbeitet und daher eine neue ersetzt werden muss. Um zu überprüfen, ob diese Vermutung hinsichtlich der Fertigungstoleranz stimmt, lässt er aus der Produktion 100 Scheiben vermessen. Die Scheiben sollen einen Durchmesser von 50 cm haben. Überschreitet die durchschnittliche Abweichung diesen Wert, muss er sich für eine neue Maschine entscheiden und damit hohe Investitionskosten in Kauf nehmen.



Anmerkung: Die Zahlenwerte sind so gewählt, dass die Rechnungen möglichst einfach werden, entsprechen damit allerdings nicht realistischen Gegebenheiten.

Zerlegen Sie durch Rechnung das \sum als Erwartungswert für den Durchmesser der Scheiben der Produktion von 50 cm ergibt. Die Zufallsvariable nimmt hier die Messwerte für den Durchmesser der 100 Scheiben an.

Zufallsvariable $X = x_i$	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
Messwerte für die x_i		49	50	51	52	53		
$P(X = x_i)$	$\frac{1}{100}$	$\sum = 1$						

Erwartungswert:

$$E(x) = P(X = x_1) \cdot x_1 + P(X = x_2) \cdot x_2 + P(X = x_3) \cdot x_3 + P(X = x_4) \cdot x_4 + P(X = x_5) \cdot x_5 + P(X = x_6) \cdot x_6 + P(X = x_7) \cdot x_7$$

$$= \frac{1}{100} \cdot 49 + \frac{1}{100} \cdot 50 + \frac{1}{100} \cdot 51 + \frac{1}{100} \cdot 52 + \frac{1}{100} \cdot 53$$

Bestimmen Sie die Standardabweichung des Erwartungswerts nicht ausreichend für die Entscheidung für die Neuinvestition zu treffen.

Ergänzen Sie für die folgenden Berechnungen die fehlende Angaben.

Für den Erwartungswert wird hier μ verwendet.

Um das Rechnen mit Wurzeln zu vermeiden und um größere Abweichungen stärker gewichtet als die einfachen Abweichungen zu verhalten.

Einfache Abweichung der x_i vom Erwartungswert	$x_1 - \mu$	$x_2 - \mu$	$x_3 - \mu$	$x_4 - \mu$	$x_5 - \mu$	$x_6 - \mu$	$x_7 - \mu$
	47 - 50	48 - _____	49 - _____	50 - _____	51 - _____	52 - _____	53 - _____
$x_i - \mu$	-3	_____	_____	0	_____	_____	_____
Quadrierte Abweichung der x_i vom Erwartungswert $(x_i - \mu)^2$	$(x_1 - \mu)^2$	$(x_2 - \mu)^2$	$(x_3 - \mu)^2$	$(x_4 - \mu)^2$	$(x_5 - \mu)^2$	$(x_6 - \mu)^2$	$(x_7 - \mu)^2$
	_____	_____	_____	_____	_____	_____	_____

Varianz $V(X)$

$$V(x) = P(X=x_1) \cdot (x_1 - \mu)^2 + P(X=x_2) \cdot (x_2 - \mu)^2 + P(X=x_3) \cdot (x_3 - \mu)^2 + P(X=x_4) \cdot (x_4 - \mu)^2 + P(X=x_5) \cdot (x_5 - \mu)^2 + P(X=x_6) \cdot (x_6 - \mu)^2 + P(X=x_7) \cdot (x_7 - \mu)^2$$

$$V(x) = \frac{1}{100} \cdot 9 + \frac{10}{100} \cdot 4 + \dots + \dots + \dots + \dots + \dots + \dots + \dots$$

$$V(x) = 1,62$$

Durch das Ziehen der Wurzel aus dem berechneten Wert für die Varianz wird die **Standardabweichung σ** (Sprich: sigma) bestimmt. Damit wird das Quadrieren der Abweichungen relativiert.

Allgemeine Definition der Standardabweichung

$$\sigma = \sqrt{V(x)} = \sqrt{\sum_{i=1}^n P(X=x_i) \cdot (x_i - E(x))^2}$$

Aufgabe 8.2

Varianz und Standardabweichung bei der Auswertung von Messungen

Wie Sie aus dem naturwissenschaftlichen Unterricht sicherlich wissen müssen, muss bei jeder Messung beachtet werden, dass Messfehler auftreten. Bei der vollständigen Auswertung einer Messung müssen diese Messfehler angegeben werden. Im Allgemeinen wird das als Standardabweichung bezeichnet. Ergänzen Sie die für die folgende Messung fehlenden Angaben.

Zur Bestimmung der Beschleunigung a bei gleichförmig beschleunigter Bewegung wird die folgende Messreihe aufgenommen:

Zeit t in [s]	Geschw. v in $\frac{m}{s}$	Beschleunigung $a = \frac{v}{t}$	Einfache Abweichung von a vom Erwartungswert $a_i - \bar{a}$	Quadrierte Abweichung der a_i vom Erwartungswert $(a_i - \bar{a})^2$
1	0,9	0,9	0,9 - 0,69 = 0,21	0,0441
2	1,0	_____	_____	_____
3	2,4	0,8	_____	_____
4	2,5	0,625	_____	_____
_____	3,1	0,62	_____	_____
_____	4,2	0,7	_____	_____

Mittelwert für a : $E(a) = \bar{a} = 0,69$

Varianz: $V(a) = \frac{1}{6} \cdot (0,0441 + \dots + \dots + \dots + \dots)$

Dieser Messung für die Standardabweichung $\sigma = \sqrt{\dots}$

Übungen

Ü8.1: Varianz und Standardabweichung bei Fertigungstoleranzen

Zwei Maschinen A und B schneiden Stahlstifte auf die Länge von 10 mm. Die Abweichungen bei der Stifflänge auftreten. Bei 30 Stiften werden jeweils die gemessenen Abweichungen in einer Häufigkeitsverteilung dargestellt. Entscheiden Sie intuitiv, welche der Maschinen besser arbeitet bzw. die kleinere Fertigungstoleranz hat, und prüfen Sie durch die Berechnung der Standardabweichung, ob Ihre Entscheidung richtig ist.

- Maschine A arbeitet besser
- Maschine B arbeitet besser

Bei den Stabdiagrammen sind auf der horizontalen Achse die Informationen über die Länge der Stifte aufgetragen und auf der vertikalen Achse die absolute Häufigkeit. Entnehmen Sie die Informationen zu Messwerten und Wahrscheinlichkeiten den beiden Diagrammen, und füllen Sie die Tabelle entsprechend zur Berechnung aus:



Rechnung für Maschine A:

Maschine A							
Messwerte x_i	$x_1 = 7$	$x_2 = 8$	$x_3 = 9$	$x_4 = \underline{\quad}$	$x_5 = \underline{\quad}$	$x_6 = \underline{\quad}$	$x_7 = \underline{\quad}$
relative Häufigkeit: $H(x_i)$ bzw. Wahrscheinlichkeit: $P(X=x_i)$							

Erwartungswert: $E(X) = \underline{\quad}$
 $\underline{\quad} = 10$

Abweichung der Messwerte vom Erwartungswert $x_i - E(X)$							
Quadrierte Abweichung $(x_i - E(X))^2$							

Varianz: $V(X) = \underline{\quad}$
 $\underline{\quad} = 10$
 Standardabweichung: $\sigma = \underline{\quad}$

Rechnung für Maschine B:

Maschine B				
Messwerte x_i	$x_1 = \underline{\quad}$	$x_2 = \underline{\quad}$	$x_3 = \underline{\quad}$	$x_5 = \underline{\quad}$
relative Häufigkeit: $H(x_i)$ bzw. Wahrscheinlichkeit: $P(X=x_i)$				

Erwartungswert: $E(X) = \underline{\quad}$

Abweichung der Messwerte vom Erwartungswert $x_i - E(X)$				
Quadrierte Abweichung $(x_i - E(X))^2$				

Varianz: $V(X) = \underline{\quad}$

$\underline{\quad} = 1,31$

Standardabweichung: $\sigma = \underline{\quad} = 1,31$

Ergebnis:

Anhand des Erwartungswerts kann man nicht entscheiden, welche der Maschinen eine kleinere Toleranz hat, da diese Werte bei beiden Maschinen $\underline{\quad}$ sind.

Maschine B hat die kleinere Fertigungstoleranz als Maschine $\underline{\quad}$, da die Standardabweichung bei Maschine $\underline{\quad}$ ist.

Ihre intuitive Entscheidung war: richtig

Ü8.2 Recherchieren Sie in geeigneten Quellen, welche Bedeutung die Standardabweichung bei Glücksspielen hat.

$\underline{\quad}$
 $\underline{\quad}$
 $\underline{\quad}$

Beide Übungen: $\underline{\quad}$
 $\underline{\quad}$

Raum für Notizen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Binomialverteilung



Einführung in die Stochastik

Kapitel 1
 Grundbegriffe 7

Kapitel 2
 Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
 Vierfeldertafel 23

Kapitel 4
 Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
 Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
 Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
 Erwartungswert 63

Kapitel 8
 Varianz und Standardabweichung 69

Kapitel 9
 Binomialverteilung 75

Hypothesentests

Kapitel 10
 Zweiseitiger Signifikanztest 97

Kapitel 11
 Einseitiger Signifikanztest – Fisher beim Testen 109

Kapitel 12
 Vertrauensintervall 135

Kapitel 13
 Die Betrachtungen zur Normalverteilung 139

Kapitel 14
 Anwendung der Normalverteilung 149

Gesamt: Selbstorganisiert erlernen (Bestenfalls 278)

Alle Rechte vorbehalten. All rights reserved.
 Nachdruck, auch auszugsweise, vorbehaltlich der Rechte,
 aus §§ 53, 54 UrhG ergeben, nicht gestattet.
 Lehrerselbstverlag
 Ursula Pirkel & Freunde GmbH, Koblenz (Germany)
 www.f-druck.de

Kapitel 9: Binomialverteilung

Die Binomialverteilung stellt eine wesentliche Basis bei der statistischen Auswertung von Daten dar und nimmt damit hinsichtlich vieler Anwendungsprobleme in der Wirtschaftswissenschaften und anderen Gebieten einen besonderen Stellenwert in der Wahrscheinlichkeitsrechnung ein. In der folgenden Aufgabe soll verdeutlicht werden, was man unter einer Binomialverteilung versteht.

Aufgabe 9.1

Herleitung der Formel von Bernoulli

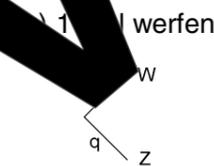
Eine verbeulte Münze zeigt beim Werfen Wappen mit einer unbekanntem Wahrscheinlichkeit p und Zahl z mit einer unbekanntem Wahrscheinlichkeit $q = 1 - p$. Die Münze wird n Mal geworfen.

$$p(\text{Wappen}) = p(W) = p$$

$$p(\text{Zahl}) = p(Z) = q$$

Die Zufallsvariable X ist die oft Wappen gefallen ist.

Ermitteln Sie die durch die gegebenen Baumdiagramme dargestellten Wahrscheinlichkeiten $P(X = k)$ in allgemeiner Form in Abhängigkeit von p und q . Beachten Sie dabei, wie oft der jeweilige Pfad auftritt, und ergänzen Sie die Ausdrücke für die Wahrscheinlichkeiten sowie weitere fehlende Angaben in der zugehörigen und anderen Tabellen.



$P(0\text{-mal Wappen}) = P(X = 0) = q$

$P(1\text{-mal Wappen}) = P(X = 1) = p$

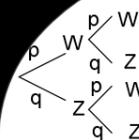
b) 2-mal werfen q^2



$P(1\text{-mal Wappen}) =$

$P(2\text{-mal Wappen}) = P(X = 2) =$

c) 3-mal werfen



$P(0\text{-mal Wappen}) = P(X = 0) =$

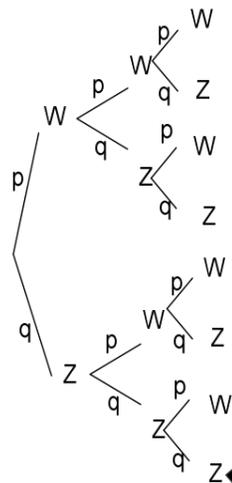
$P(1\text{-mal Wappen}) = P(X = 1) =$

$P(2\text{-mal Wappen}) = P(X = 2) =$

$P(3\text{-mal Wappen}) =$

d) 4-mal werfen

Ergänzen Sie die vierte Stufe im Baumdiagramm, und bestimmen Sie die Wahrscheinlichkeiten.



$P(0\text{-mal Wappen}) = P(X = 0) =$ _____

$P(1\text{-mal Wappen}) = P(X = 1) =$ _____

$P(2\text{-mal Wappen}) = P(X = 2) =$ _____

$P(3\text{-mal Wappen}) = P(X = 3) =$ _____

$P(4\text{-mal Wappen}) = P(X = 4) =$ _____

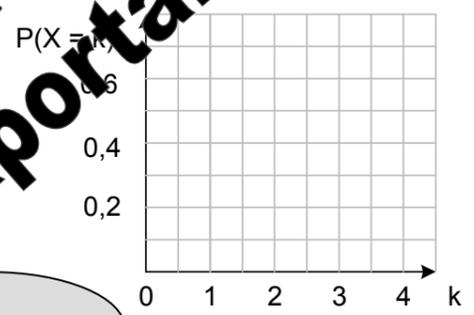
Wahrscheinlichkeitswerte für die Zufallsvariable X. Ergänzen Sie fehlende Angaben.

Ereignisse e_i als Klartext	0 x Wappen	1 x Wappen	2 x Wappen	3 x Wappen	4 x Wappen
Werte k der Zufallsvariable	0	1	2	3	4
Wahrscheinlichkeit für die Zufallsvariable $P(X = k)$ bei n Würfen	$P(X = 0)$	$P(X = 1)$	$P(X = 2)$	$P(X = 3)$	$P(X = 4)$
Anzahl n der Würfe	$n=1$	q	p		
	$n=2$		p^2		
Anzahl n der Würfe	$n=4$	q^4			p^4
	Werte k der Zufallsvariable	$\binom{4}{0} p^0 q^4$	$\binom{4}{1} p^1 q^3$	$\binom{4}{2} p^2 q^2$	$\binom{4}{3} p^3 q^1$
Umformen der Exponenten		$\binom{4}{1} p^1 q^{4-1}$	$\binom{4}{2} p^2 q^{4-2}$	$\binom{4}{3} p^3 q^{4-3}$	$\binom{4}{4} p^4 q^{4-4}$
Verallgemeinerung für n Würfe	$P(X = k) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$ Formel von Bernoulli				

e) Ermitteln Sie für den Münzwurf die Zahlenwerte für die Wahrscheinlichkeitsverteilung mit vier Dezimalstellen genau, wenn die Wahrscheinlichkeit für Wappen bei der verwerflichen Münze $p = 0,3$ beträgt.

Ereignisse e_i als Klartext	0 x Wappen	1 x Wappen	2 x Wappen	3 x Wappen	4 x Wappen
k	0	1	2	3	4
$P(X = k)$	$\binom{4}{0} 0,3^0 \cdot 0,7^4$	$\binom{4}{1} 0,3^1 \cdot 0,7^3$	$\binom{4}{2} 0,3^2 \cdot 0,7^2$	$\binom{4}{3} 0,3^3 \cdot 0,7^1$	$\binom{4}{4} 0,3^4 \cdot 0,7^0$

Zeigen Sie, dass die Summe aller Wahrscheinlichkeiten für die Ereignisse e_i den Wert 1 ergibt, und stellen Sie die berechneten Wahrscheinlichkeiten als Säulendiagramm dar.



Die Wahrscheinlichkeiten $P(X = k)$ werden mit der Formel von Bernoulli berechnet. Da diese Formel aus dem Pascalschen Dreieck und damit aus der Berechnung von Binomialkoeffizienten $\binom{n}{k}$ hervorgeht, bezeichnet man die hier vorliegende Wahrscheinlichkeitsverteilung als **Binomialverteilung**.

Die Formel von Bernoulli wurde anhand des Münzwurfs als ein Zufallsversuch, bei dem die beiden Ausgänge Wappen und Zahl gibt und die Wahrscheinlichkeiten p und $q = 1 - p$ für Wappen und Zahl bei jedem Werfen der Münze gleich bleiben. Diese beiden Eigenschaften sind Grundvoraussetzungen für die Berechnung von Wahrscheinlichkeiten mithilfe der Formel von Bernoulli. Damit gilt:

Eine Zufallsvariable X ist **binomialverteilt**, wenn

- das Experiment ein Ereignis E und dem zugehörigen Gegenereignis \bar{E} nur zwei

Ergebnisse vorstellen kann, dass die Wahrscheinlichkeit p für das Ereignis E und die Wahrscheinlichkeit $q = 1 - p$ für das Gegenereignis \bar{E} in jeder Stufe des Experimentes gleich bleibt.

Aufgabe 9.2

Anwendungsbeispiele für binomialverteilte Zufallsgrößen

Problemstellungen, die man auf einen Zufallsversuch mit gleichbleibenden Wahrscheinlichkeiten in jeder Stufe und zwei Ausgängen zurückführen kann, können mithilfe der Binomialverteilung, also mit der Formel von Bernoulli gelöst werden. Dies soll nun anhand der nachfolgenden drei Beispiele exemplarisch gezeigt werden. Arbeiten Sie die Beispiele durch, erörtern Sie bei fehlende Angaben und formulieren Sie entsprechende Begründungen.

Beispiel 1

Bei der automatisierten Produktion von Biergläsern für die Gastronomie haben 10% der Gläser Fehler, sind daher nicht verkäuflich, werden bei der Kontrolle aufmerkt und weiter eingeschrotet.

- a) Um die Fehlerrate bei der Produktion zu überprüfen, werden der laufenden Produktion 10 Gläser entnommen und auf Fehler geprüft. Berechnen Sie die Wahrscheinlichkeit, dass sich in dieser Stichprobe genau drei Gläser mit einem Fehler befinden.

Die laufende Produktion entnehmen heißt, dass die Wahrscheinlichkeit von 10% für das Auftreten eines Fehlers bei jeder Entnahme gleich ist. Es ständig neue Gläser mit der gleichen Fehlerrate hinzukommen.

Zuordnen der Zufallsgrößen n , p , k und $P(X = k)$ der Formel von Bernoulli.

Es handelt sich um eine binomialverteilte Zufallsvariable, da es nur die beiden Ausgänge „Glas mit Fehler“ und „Glas ohne Fehler“ gibt und die Wahrscheinlichkeit für F bzw. \bar{F} bei jedem Ziehen gleich bleibt.

Anzahl der entnommenen Gläser: $n = \underline{\hspace{2cm}}$

Anzahl der Gläser mit dem Merkmal Fehler: $k = \underline{\hspace{2cm}}$

Wahrscheinlichkeit

Der Anteil in der Stichprobe mit einer Wahrscheinlichkeit von $p = 0,1$ sind drei Gläser mit einem Fehler.

- b) In einer Kiste mit 100 Gläsern befinden sich 100 nicht geprüfte Gläser, von denen 10 Gläser mit Fehler. Wie groß ist die Wahrscheinlichkeit, dass man hier bei der Entnahme von 10 Gläsern genau ein fehlerbehaftetes Glas findet?

Es handelt sich um einen Zufallsversuch, der den Bedingungen der Binomialverteilung entspricht, da man ohne Zurücklegen zieht und die Wahrscheinlichkeit in den verschiedenen Stufen des Zufallsversuchs äquivalent bleibt. Da die Stichprobe jedoch aus einer umfangreichen Menge von 100 Gläsern entnommen wird, kann hier Wahrscheinlichkeiten von Stufe zu Stufe als konstant angesehen werden und die Formel von Bernoulli für die Berechnung einer Näherung angewendet werden.

- i) Berechnen Sie die gesuchte Wahrscheinlichkeit näherungsweise mithilfe der Formel von Bernoulli.

Anzahl der entnommenen Gläser: $n = \underline{\hspace{2cm}}$

Anzahl der Gläser mit Fehler: $k = \underline{\hspace{2cm}}$

Wahrscheinlichkeit für das Merkmal Fehler: $p = \underline{\hspace{2cm}}$

$P(X = \underline{\hspace{2cm}}) = \underline{\hspace{2cm}} = 38,74\%$

- ii) Zeigen Sie, dass sich mithilfe von kombinatorischen Verfahren der exakte Wert $13,8\%$, das heißt nur eine Abweichung von $0,4\%$ zum vorherigen Näherungswert, ergibt.

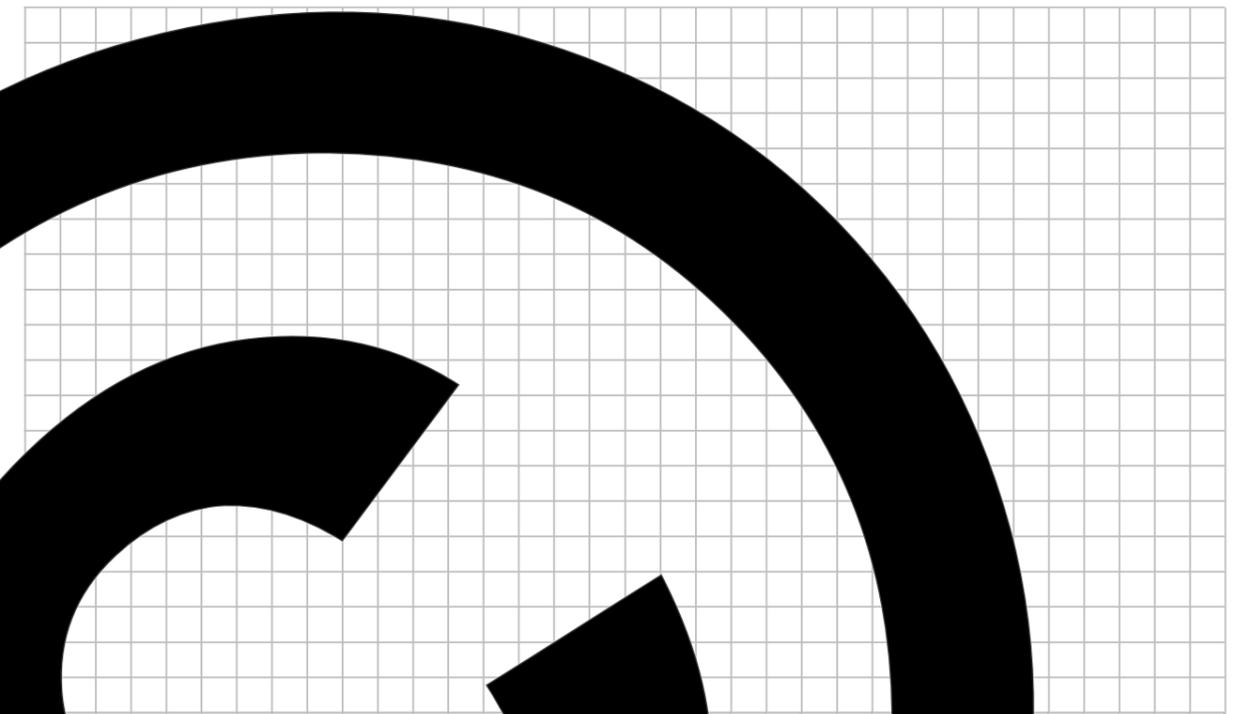
$n = \underline{\hspace{2cm}}$

$k = \underline{\hspace{2cm}}$

$P(X = \underline{\hspace{2cm}})$

Achtung!
Die Bedeutungen von n und k unterscheiden sich in der Kombinatorik von denen bei der Binomialverteilung.

In einem Karton mit 20 ungeprüften Gläsern befinden sich 10% Gläser mit Fehler. Berechnen Sie auch hier die Wahrscheinlichkeit, dass sich unter 10 entnommenen Gläsern eines mit einem Fehler befindet, als Näherungswert mit der Formel von Bernoulli und mit dem kombinatorischen Verfahren. Zeigen Sie durch Rechnung, dass die beiden Werte nun fast 14% voneinander abweichen und die Verwendung der Formel von Bernoulli hier nicht mehr geeignet ist.



Zusammenfassung und Ergänzungen zur Schreibweise

Sie haben in den Beispielen von Aufgabe 9.2 gelernt, unter welchen Bedingungen die Binomialverteilung für Berechnungen von Wahrscheinlichkeiten verwendet werden kann. Es wurde gezeigt, dass von einer binomialverteilten Zufallsvariablen ausgegangen werden kann. Es wurde auch gezeigt, dass jedoch Aufgabenstellungen, die in der Form $P(X < k)$ oder $P(X \leq k)$ vorliegen, einen beträchtlichen Rechenaufwand nach sich ziehen, da die Formel mehrfach für verschiedene Werte berechnet werden muss. Da man diesen Rechenaufwand erheblich verringern kann, wenn man die Binomialverteilung in der Form $F_{n,p}(k)$ darstellt, spielt die Anwendung der Binomialverteilung in der Form $F_{n,p}(k)$ eine große Rolle. Bei der Verwendung der Formel von Bernoulli im Rahmen der Berechnung von binomialverteilten Zufallsvariablen wird häufig die Schreibweise $B_{n,p}(k)$ wie folgt angewandt:

Für eine binomialverteilte Zufallsvariable gilt:

$$P(X = k) = B_{n,p}(k) = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Für $n = 10$, $k = 7$ und $p = 0,8$ schreibt man: $P(X = 7) = B_{10,0,8}(7) = \binom{10}{7} 0,8^7 \cdot 0,2^3 = 0,2013$

Umgang mit der Binomialverteilung

a) Berechnen Sie die binomialverteilten Wahrscheinlichkeiten $P(X = k) = B_{5,0,4}(k)$ mithilfe der Formel von Bernoulli, und tragen Sie Ihre Ergebnisse an den vorgesehenen Stellen ein. Vervollständigen Sie anschließend das Säulendiagramm.

k	$P(X = k)$
0	
1	
2	$B_{5,0,4}(2) =$
3	$B_{5,0,4}(3) =$
4	$B_{5,0,4}(4) =$
5	$B_{5,0,4}(5) =$

Die Wahrscheinlichkeiten können in Tafelwerken und oft auch Schulbüchern in Tafelwerken abgelesen werden und können auch mit einem Taschenrechner oder Computer ermittelt werden. Vergleichen Sie die Binomialverteilung ermittelten Werte mit einer entsprechenden Tabelle aus einem Taschenrechner oder mit einer Liste, die ein Taschenrechner erzeugt. Lesen Sie dazu die Bedienungsanleitung Ihres Taschenrechners nach, wie die Werte einer Binomialverteilung als Liste oder Einzelwerte erzeugt werden können.

b) Im Säulendiagramm ist eine $B_{5,0,7}(k)$ -Verteilung abgebildet. Lesen Sie die Werte für die Wahrscheinlichkeiten $P(X = k)$ aus dem Diagramm näherungsweise ab, und tragen Sie die Werte in der Tabelle ein. Berechnen Sie die Werte anschließend mit der Formel von Bernoulli vier Nachkommastellen genau.

k	P(X = k)	
	Werte abgelesen	Werte berechnet
0		
1		
2		
3		
4		
5		

c) Ermitteln Sie die Wahrscheinlichkeiten in der Form $P(X < k)$, $P(X > k)$, $P(X \leq k)$ und $P(X \geq k)$.

Die Werte im Säulendiagramm sind die Zahlenwerte einer $B_{10,0,6}(k)$ -Verteilung gegeben.

k	$P(X = k)$
0	0,0060
1	0,0360
2	0,0106
3	0,0425
4	0,1115
5	0,2007
6	0,0060

d) Ermitteln Sie die Wahrscheinlichkeit $P(X < 5) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$ im Stabdiagramm, welche Balken hier aufsummiert werden.

$$P(X < 5) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = \dots + \dots + \dots = \dots$$

Die Wahrscheinlichkeiten können ebenfalls mithilfe von Tafelwerken abgelesen werden und können auch mit einem Taschenrechner ermittelt werden. Man verwendet hier die Binomialverteilungsfunktion $F_{n,p}(k)$. Ermitteln Sie den gesuchten Wert in einer derartigen Tabelle, bzw. berechnen Sie den Wert mithilfe Ihres Taschenrechners.

$$P(X < 5) = P(X \leq 4) = F_{10,0,6}(4) = \dots$$

ii) Ermitteln Sie anhand der Tabelle oder mithilfe des Taschenrechners für eine $B_{50;0,5}(k)$ -Verteilung die folgenden Wahrscheinlichkeiten.

$P(X < 20) = P(x \leq 19) = F_{50;0,5}(19) =$ _____

$P(X < 30) =$ _____

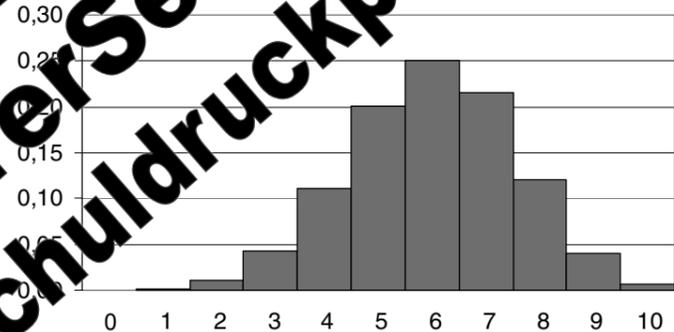
$P(X \leq 25) =$ _____

$P(X \leq 11) =$ _____

$P(X < 40) =$ _____

iii) Markieren Sie in der folgenden Tabelle und im Säulendiagramm die Werte bzw. Balken, die man aufsummieren muss, um diese Verteilung die Wahrscheinlichkeit $P(X \geq 5)$ zu ermitteln, und zeigen Sie die Verwendung der gedruckten Tabellen, nach entsprechende Summenbildung auf den folgenden Beispielen, dass sich der Wert 0,8377 ergibt.

k	$P(X = k)$
0	0,0010
1	0,0050
2	0,0107
3	0,0175
4	0,0250
5	0,0337
6	0,0437
7	0,0547
8	0,0663
9	0,0781
10	0,0900



$P(X \geq 5) =$ _____

und auch nicht direkt mit dem Taschenrechner erhalten. Man muss hier mit den Gegenwahrscheinlichkeiten arbeiten. Da die Gegenwahrscheinlichkeit der Summe aller Werte aller Balken bei einer Verteilung 1 ist, entspricht die Summe des Gegenwahrscheinlichkeits der Summe aller Balken bei einer Verteilung 1. Die Gegenwahrscheinlichkeit der Summe aller Balken bis 4 ist 0,1663. Die Gegenwahrscheinlichkeit der Summe aller Balken bis 4 ist 0,1663. Für die Berechnung bedeutet das:

$P(X \geq 5) = 1 - P(X \leq 4) = 1 - F_{10;0,5}(4) = 1 - 0,1663 = 0,8337$

iv) Bestimmen Sie die folgenden Werte:

$P(X > 5) =$ _____

$P(X < 10) =$ _____

$P(X \leq 10) =$ _____

$P(X < 10) =$ _____

Übungen

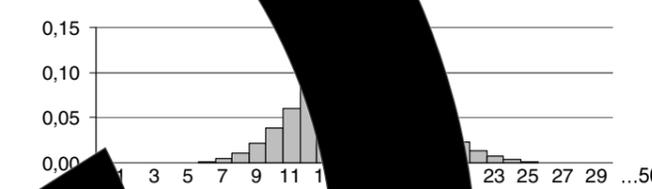
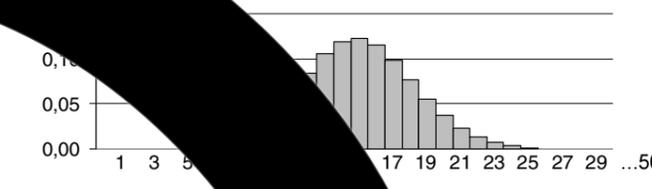
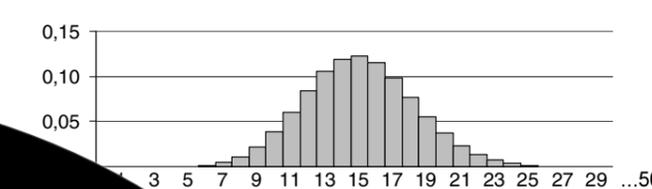
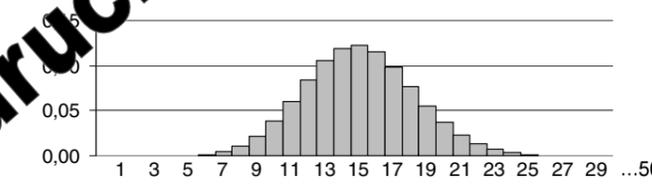
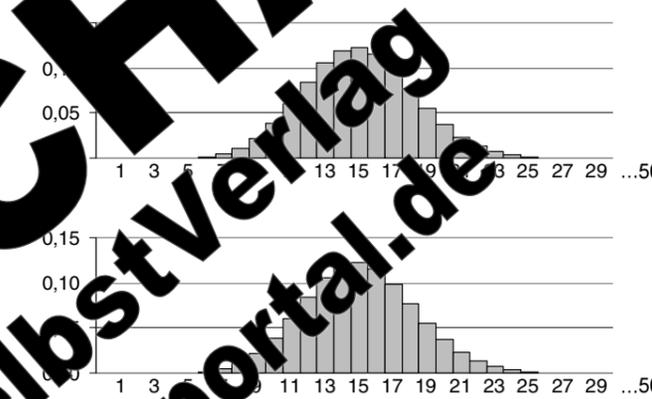
Ü9.1 Ermitteln von Wahrscheinlichkeiten mit der Binomialsummenfunktion $F_{50;0,3}(k)$. Markieren Sie die zu berechnenden Wahrscheinlichkeiten im dargestellten Ausschnitt des Säulendiagramms der $B_{50;0,3}(k)$ -Verteilung, und bestätigen Sie jeweils die angegebenen Wahrscheinlichkeiten mithilfe des Taschenrechners und oder einer entsprechenden Tabelle.

a) $P(X \leq 13) = F_{50;0,3}(\quad)$
 $= 0,3270$

b) $P(X > 17) = 1 - (\quad)$
 $= 1 - 0,6722$
 $= 0,3278$

d) $P(X \leq 9) + P(X \geq 20)$
 $= P(X \leq 9) + 1 - P(X \leq 19)$
 $= F_{50;0,3}(9) + 1 - F_{50;0,3}(19)$
 $= 0,0478 + 1 - 0,6522$
 $= 0,4456$

e) $P(11 < X < 18)$
 $= P(X \leq 17) - P(X \leq 11)$
 $= F_{50;0,3}(17) - F_{50;0,3}(11)$
 $= 0,9522 - 0,0642$
 $= 0,8880$



Aufgabe 9.4

Ermitteln von k bei der Binomialsammenfunktion

Anhand des folgenden Beispiels soll verdeutlicht werden, dass es auch Aufgabenstellungen gibt, bei denen k ermittelt werden muss. Verdeutlichen Sie sich dazu das folgende Beispiel.

Bei der Produktion von T-Shirts einer hochpreisigen Marke sind leider 30% der T-Shirts mit einem Fehler an einer Naht betroffen. Es werden zur Überprüfung der Angaben des Produktionsleiters Stichproben von 50 T-Shirts in jeder Größe untersucht. Der Wert k zählt die T-Shirts mit einem Fehler.

- a) Mit welcher Wahrscheinlichkeit sind höchstens 10, 15, 20 T-Shirts mit den Fehlern an der Naht in einer Stichprobe?

$$P(X \leq 10) = F_{50; 0,3}(10) = 7,9\%$$

$$P(X \leq 15) = F_{50; 0,3}(15) = 57\%$$

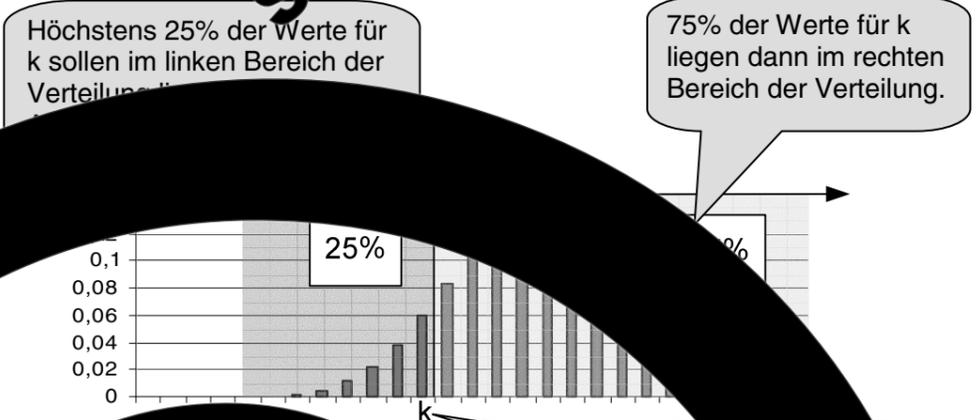
$$P(X \leq 20) = F_{50; 0,3}(20) = 95\%$$

In den Stichproben sind sich mit einer Wahrscheinlichkeit von 7,9% bis zu 10, mit einer Wahrscheinlichkeit von 57% bis zu 15 und mit einer Wahrscheinlichkeit von 95% bis zu 20 T-Shirts mit Fehler an der Naht. Wenn 30% der T-Shirts den Fehler aufweisen.

- b) In der Stichprobe sind sich mit einer Wahrscheinlichkeit von höchstens 25 % bis zu k T-Shirts mit einem Fehler an der Naht. Welche Werte kann k annehmen, wenn die Angaben des Produktionsleiters hinsichtlich der Fehlerquote stimmen?

Bei dieser Fragestellung ist k gesucht. Es werden neue Überlegungen notwendig.

Verdeutlichen Sie sich die Zusammenhänge anhand der folgenden Darstellungen:



Aus einer Tabelle kann man ablesen, dass zu welchem Wert von k die Verteilungsfunktion $F_{50; 0,3}(k) \leq 0,25$ erfüllt ist. Hier ist $k = 10$ erfüllt, denn aus der Tabelle ergibt sich $F_{50; 0,3}(10) = 0,079$.

Der Wert $k = 10$ ist die Verteilungsfunktion $F_{50; 0,3}(k) \leq 0,25$ erfüllt, denn aus der Tabelle ergibt sich $F_{50; 0,3}(10) = 0,079$.

In der Stichprobe befinden sich mit einer Wahrscheinlichkeit von 22,29% bis zu k T-Shirts mit einem Fehler an der Naht, wenn die Angabe des Produktionsleiters hinsichtlich der Fehlerquote stimmen.

Aufgabe 9.5

Allgemeine Formulierung der Aufgabenstellung zur Bestimmung von k

In der Aufgabenstellung sind n, p und $P(X = k)$ gegeben. Gesucht ist der Wert von k. Zur Veranschaulichung der Rechnung wird in den folgenden Abbildungen ein Säulendiagramm ersatzweise eine einhüllende Kurve für das Säulendiagramm, die den für die Rechnung relevanten Balken verwendet.

Die grundlegende Aufgabenstellung wird jeweils durch die folgenden Abbildungen verdeutlicht und anhand der Verteilung $B_{100; 0,5}(k)$ beispielhaft berechnet.

Bis zu welchem Wert von k gilt, dass die Wahrscheinlichkeit $P(X \leq k)$ höchstens 10% ist? Ab welchem Wert von k gilt, dass die Wahrscheinlichkeit $P(X \geq k)$ höchstens 10% ist?



$$P(X \leq k) \leq 0,1$$

$$1 - P(X \leq k-1) \leq 0,1$$

$$P(X \leq k-1) \geq 0,9$$

$$F_{100; 0,5}(k-1) \geq 0,9$$

$$k-1 = 43 \leftarrow \text{aus der Tabelle}$$

$$k = 44$$

$$P(X \geq k) \leq 0,1$$

$$1 - P(X \leq k-1) \leq 0,1$$

$$- P(X \leq k-1) \leq -0,9$$

$$P(X \leq k-1) \geq 0,9$$

$$F_{100; 0,5}(k-1) \geq 0,9$$

$$k-1 = 56 \leftarrow \text{aus der Tabelle}^*$$

$$k = 57$$

*Hinweis zum Ablesen aus einer entsprechenden Tabelle:

1. Tabelle F_{100} auswählen.
2. In der Spalte für die Wahrscheinlichkeit von oben nach unten durchgehen, bis man auf den ersten Wert trifft, der größer oder gleich dem gesuchten Wert ist.
3. Der Wert in der Spalte links daneben ist der Wert für k zuzuordnen ist.

9.2 Bestimmen Sie für die angegebenen Wahrscheinlichkeiten die Werte für k. Die Ergebnisse sind zur Kontrolle anzugeben. Raum für Rechnungen auf der folgenden Seite.

Verteilung	Bedingung	Ergebnis
a) $B_{100; 0,5}(k)$	$P(X \leq k) \leq 0,05$	k = 1
	$P(X \geq k) \leq 0,05$	k = 9
b) $B_{100; \frac{1}{6}}(k)$	$P(X \leq k) \leq 0,025$	k = 9
	$P(X \geq k) \leq 0,025$	k = 25
c) $B_{100; 0,97}(k)$	$P(X \leq k) \leq 0,85$	k = 20
	$P(X \geq k) \geq 0,85$	k = 13
d) $B_{100; 0,8}(k)$	$P(X \leq k) \leq 0,33$	k = 95
	$P(X \geq k) \leq 0,33$	k = 99
e) $B_{100; 0,6}(k)$	$P(X \leq k) \geq 0,2$	k = 68
	$P(X \geq k) \geq 0,95$	k = 1
f) $B_{100; \frac{5}{6}}(k)$	$P(X \leq k) \leq 0,1$	k = 74
	$P(X \geq k) \leq 0,1$	k = 86
g) $B_{100; \frac{5}{6}}(k)$	$P(X \leq k) \leq 0,1$	k = 78
	$P(X \geq k) \leq 0,1$	k = 88

VORSCHAU
Lehrersebstverlag
schuldruckportal.de



Ergebnisse: _____

Aufgabe 9.6

Erwartungswert und Standardabweichung bei der Binomialverteilung

Zeigen Sie mit den in Kapitel 8 verwendeten Verfahren, dass sich für den Erwartungswert und die Standardabweichung bei den folgenden binomialverteilten Zufallsgrößen die angegebenen Werte berechnen lassen.

a) Die Zufallsvariable X ist $B_{5;0,4}(k)$ verteilt. Berechnen Sie den Erwartungswert μ und ergänzen Sie fehlende Angaben in der folgenden Tabelle. Berechnen Sie anschließend die Varianz $V(x)$ sowie die Standardabweichung σ .

$$\mu = _ \cdot 0,0778 + _ \cdot 0,2592 + _ \cdot 0,3456 + _ \cdot 0,2304 + _ \cdot 0,0768 + _ \cdot 0,0102 = _ \approx 2$$

k	Wahrscheinlichkeit $P(X = k)$	Einfache Abweichung vom Erwartungswert	Quadratisierte Abweichung
0	0,0778	0 - 2 =	
1	0,2592	1 - 2 =	
2	0,3456		
3	0,2304		
4	0,0768		
5	0,0102		

$$V(x) = _ \cdot 0,0778 + _ \cdot 0,2592 + _ \cdot 0,3456 + _ \cdot 0,2304 + _ \cdot 0,0768 + _ \cdot 0,0102 = _ \approx 1,197$$

$$\sigma = _ \approx 1,1$$

b) Die Zufallsvariable Y ist $B_{3;0,8}(k)$ verteilt. Berechnen Sie wie in a).

$$\mu = _ \cdot 0,0080 + _ \cdot 0,0240 + _ \cdot 0,0480 + _ \cdot 0,0720 + _ \cdot 0,0960 + _ \cdot 0,1200 = _ \approx 2,4$$

k	Wahrscheinlichkeit $P(X = k)$	Einfache Abweichung vom Erwartungswert	Quadratisierte Abweichung
0	0,0080	0 - 2,4 =	
1	0,0240		
2	0,0480		
3	0,0720		
4	0,0960		
5	0,1200		

$$V(x) = _ \cdot 0,0080 + _ \cdot 0,0240 + _ \cdot 0,0480 + _ \cdot 0,0720 + _ \cdot 0,0960 + _ \cdot 0,1200 = _ \approx 0,69$$

c) Ermitteln Sie für die in Aufgabenteil a) und b) gegebenen Größen die folgenden Produkte, vergleichen Sie die Ergebnisse jeweils mit den berechneten Werten für μ und σ . Tragen Sie anschließend den Merksatz.

zu a) $n \cdot p =$ ____ und $\sqrt{n \cdot p \cdot q} =$ ____ zu b) $n \cdot p =$ ____ und $\sqrt{n \cdot p \cdot q} =$ ____

Für eine binomialverteilte Zufallsgröße X mit den Parametern n und p gilt:
Berechnung des Erwartungswertes $\mu =$ ____
Berechnung der Standardabweichung $\sigma =$ ____

Aufgabe 9.7

Eigenschaften der Binomialverteilung

In den Abbildungen 9.1 bis 9.5 sind jeweils in geeigneter Ausschnittgröße der k-Achse die Säulendiagramme einer Binomialverteilung abgebildet. Betrachten Sie hinsichtlich des Erscheinungsbildes der Verteilungen die nachfolgenden Aufgabenstellungen.

a) Untersuchen Sie, welche Auswirkung die Werte von n und p auf das Erscheinungsbild der Verteilung hinsichtlich der Symmetrieeigenschaft, der Lage und Höhe des höchsten Balkens und der Breite der Verteilung haben.

Je größer n wird, desto symmetrischer ist das Erscheinungsbild der Verteilung, wenn man den höchsten Balken als Symmetrieachse verwendet.

- Je größer n wird, desto ____ wird die Höhe des höchsten Balkens.
- Je größer n wird, desto ____ wird die Breite der Verteilung.
- Wenn p größer wird, verschiebt sich der k-Achse nach ____.

b) Berechnen Sie für die in den Abbildungen 9.1 bis 9.5 gezeigten Verteilungen die Standardabweichung σ sowie jeweils die 2-fache und die 3-fache Standardabweichung 2σ und 3σ . Tragen Sie die Werte an den vorgesehenen Stellen ein.

Überprüfen Sie, welche Beziehung zwischen dem Wert von μ und der Lage des höchsten Balkens der Verteilung besteht.

d) Formulieren Sie einen Zusammenhang zwischen den Symmetrieeigenschaften der Verteilung und dem Wert von σ .

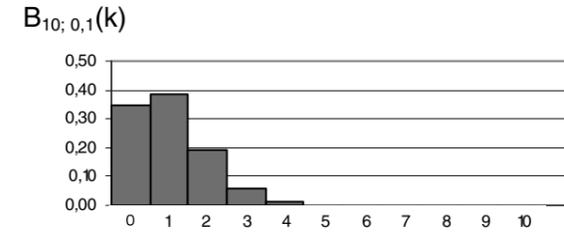


Abb. 9.1

$1\sigma =$ ____

$2\sigma =$ ____

$3\sigma =$ ____

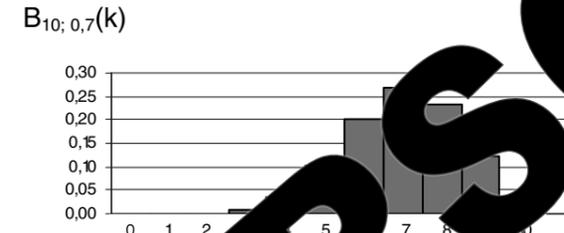


Abb. 9.2

$1\sigma =$ ____

$2\sigma =$ ____

$3\sigma =$ ____

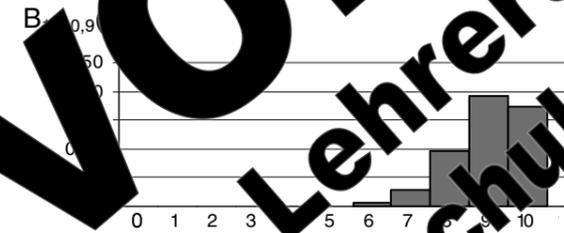


Abb. 9.3

$\mu =$ ____

$1\sigma =$ ____

$2\sigma =$ ____

$3\sigma =$ ____

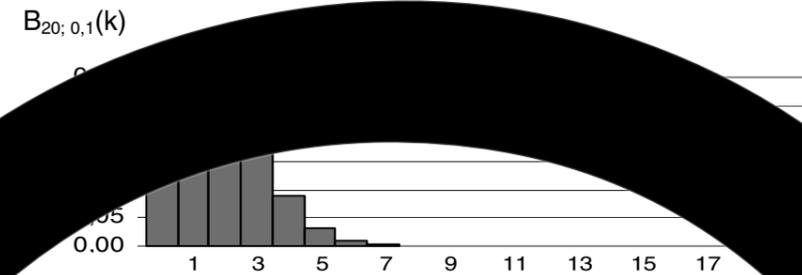


Abb. 9.4

$\mu =$ ____

$1\sigma =$ ____

$2\sigma =$ ____

$3\sigma =$ ____

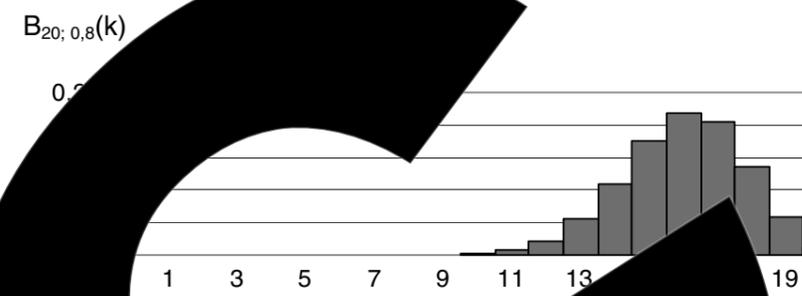


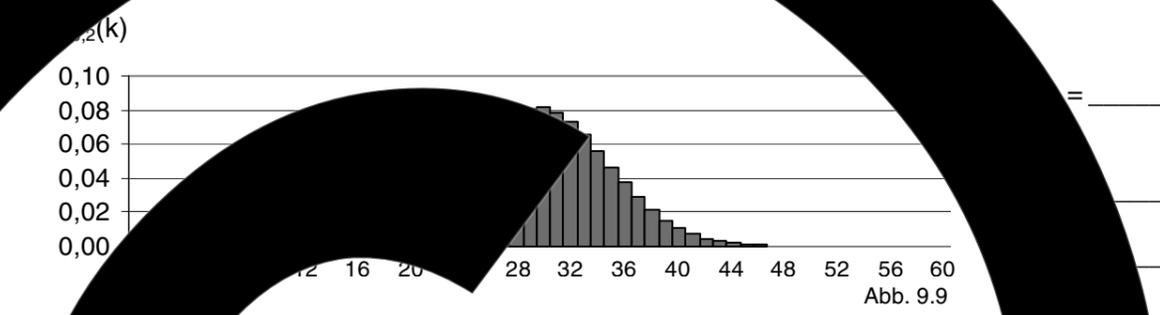
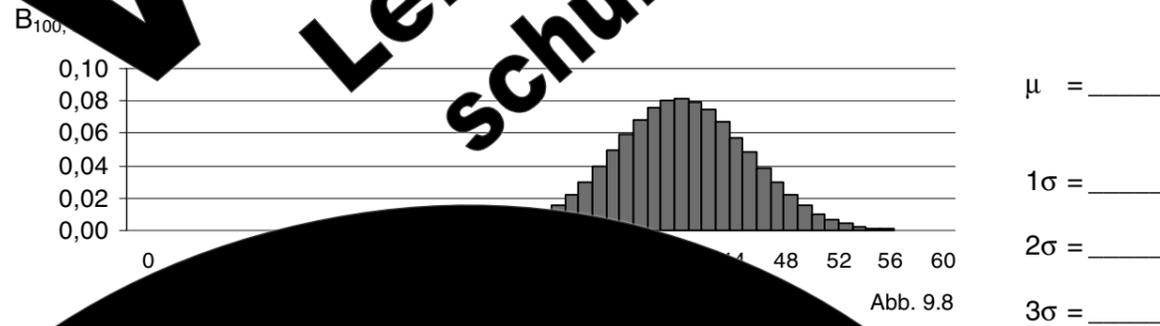
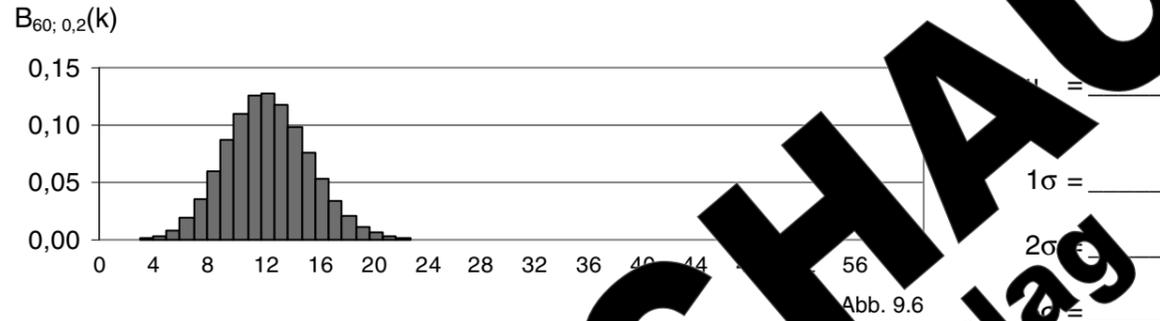
Abb. 9.5

$\mu =$ ____

$1\sigma =$ ____

$2\sigma =$ ____

$3\sigma =$ ____



Aufgabe 9.8

Die σ -Umgebungen bei der Binomialverteilung

Im Rahmen von Aufgabe 9.2 wurden im Beispiel 1 die Wahrscheinlichkeiten p ermittelt, in einer Stichprobe vom Umfang n eine bestimmte Anzahl k an fehlerhaften Gläsern zu erhalten. Würde man in der Praxis viele dieser Stichproben entnehmen, könnte man erwarten, dass die Anzahl k der Gläser mit Fehler in jeder der Stichprobe meistens nicht genau μ ausfällt, sondern μ für das Auftreten eines Fehlers entspricht, jedoch häufig in der Nähe des Erwartungswertes liegt. Fragestellungen zu Wahrscheinlichkeiten hinsichtlich der Abweichung des Wertes von k vom Erwartungswert in einer Stichprobe werden bei einer Binomialverteilung oft mit den sogenannten σ -Umgebungen untersucht. Daher soll in der folgenden Aufgabe der Begriff der σ -Umgebungen geklärt werden.

a) Graphische Darstellung der σ -Umgebungen

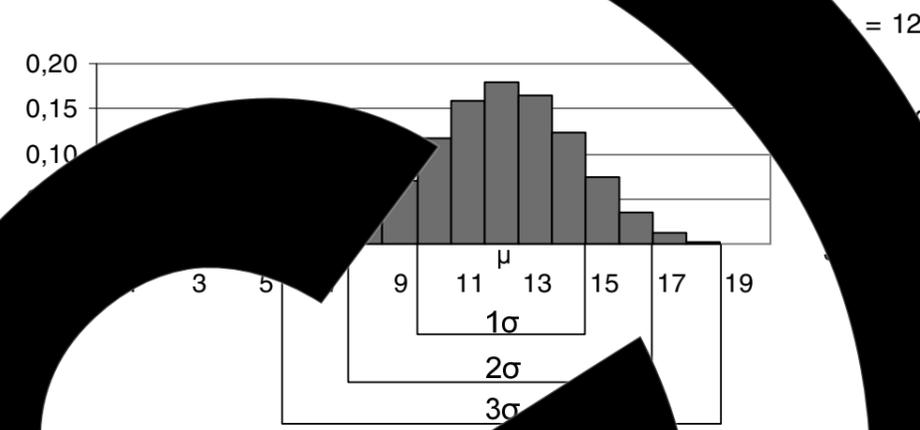
In der Abbildung ist die $B_{100; 0,6}(k)$ -Verteilung abgebildet. Unterhalb der Balken sind drei Bereiche markiert, die als 1σ , 2σ und 3σ -Umgebung in der Binomialverteilung bezeichnet werden. Da bei einer Binomialverteilung die Zufallsvariable nur ganzzahlige Werte annimmt, geht man bei der Ermittlung dieser σ -Umgebungen wie folgt vor:

Beide Ermittlung der σ -Umgebungen wird unabhängig von der Dezimalstelle zum nächsten Balken in Richtung des Erwartungswertes gerundet.

Für die 1σ -Umgebung gilt:
 Linke Grenze: $60 - 2,2 = 57,8$ Damit liegt der Balken Nr. 10 in der 1σ -Umgebung
 Rechte Grenze: $60 + 2,2 = 62,2$ Damit liegt der Balken Nr. 14 in der 1σ -Umgebung

Für die 2σ -Umgebung gilt:
 Linke Grenze: $60 - 4,4 = 55,6$ Damit liegt der Balken Nr. 8 in der 2σ -Umgebung
 Rechte Grenze: $60 + 4,4 = 64,4$ Damit liegt der Balken Nr. 16 in der 2σ -Umgebung

Für die 3σ -Umgebung gilt:
 Linke Grenze: $60 - 6,6 = 53,4$ Damit liegt der Balken Nr. 6 in der 3σ -Umgebung
 Rechte Grenze: $60 + 6,6 = 66,6$ Damit liegt der Balken Nr. 18 in der 3σ -Umgebung



i) Markieren Sie in den Abbildungen 9.2, 9.5, 9.6, 9.8 und 9.9 entsprechend zu Beispiel 1 oben jeweils die 1σ-, 2σ- und 3σ-Umgebung.

ii) Vergleichen Sie, welchen Bereich die 3σ-Umgebung bei allen in 1) markierten Verteilungen abdeckt und welche Markierungen in den jeweiligen Verteilungen abdeckt und welche Grenzen.

Bei allen Verteilungen liegen unabhängig von n und k fast alle Werte innerhalb der 3σ-Umgebung.

iii) Wie Sie sicher erkannt haben, liegen bei allen markierten Verteilungen fast alle Werte von k innerhalb der 3σ-Umgebung. Schließen Sie mit einer persönlichen Angabe ab, welchen Bereich bei diesen Verteilungen jeweils die 1σ-, 2σ- und 3σ-Umgebung abdeckt.

Der Wert für k liegt bei etwa 70% innerhalb des 1σ-Bereichs.

Der Wert für k liegt bei etwa 95% innerhalb des 2σ-Bereichs.

Der Wert für k liegt bei etwa 99,7% innerhalb des 3σ-Bereichs.

b) Berechnen Sie die 1σ-Umgebung für eine $B_{100, 0,3}(k)$ -Verteilung mit $\mu = 30$

Vergleichen Sie sich mit der Berechnung 21 anhand des Beispiels für die 1σ-Umgebung und gehen Sie für die 2σ- und 3σ-Umgebungen analog dazu vor.

Gegeben: $B_{100, 0,3}(k)$ -Verteilung mit $\mu = 30$

1σ-Umgebung: $1\sigma = \sqrt{100 \cdot 0,3 \cdot 0,7} = 4,58$

Grenzen des 1σ-Bereichs berechnen:

$$\mu - 1\sigma \leq x \leq \mu + 1\sigma$$

Da es für k bei der Binomialverteilung nur ganzzahlige Werte gibt, müssen die berechneten Werte gerundet werden. Um innerhalb der 1σ-Umgebung zu bleiben, wird der linke Wert aufgerundet und der rechte Wert immer abgerundet.

$$26 \leq x \leq 34$$

Berechnen der Wahrscheinlichkeit, dass ein Wert für k bei der Binomialverteilung innerhalb der 1σ-Umgebung liegt:

$$P(26 \leq x \leq 34)$$

$$0,674 = 67,4\%$$

Da es für k bei der Binomialverteilung nur ganzzahlige Werte gibt, müssen die berechneten Werte gerundet werden. Um innerhalb der 1σ-Umgebung zu bleiben, wird der linke Wert aufgerundet und der rechte Wert immer abgerundet.

Zum Erwartungswert hin runden!

Zum Erwartungswert hin runden!

2σ-Umgebung: $2\sigma = 9,2$

Grenzen des 2σ-Bereichs berechnen:

$$\mu - 2\sigma \leq x \leq \mu + 2\sigma$$

$$30 - 9,2 \leq x \leq 30 + 9,2$$

$$20,8 \leq x \leq 39,2$$

$$21 \leq x \leq 39$$

Berechnen der Wahrscheinlichkeit, dass ein Wert für k bei der Binomialverteilung innerhalb der 2σ-Umgebung liegt:

$$P(21 \leq x \leq 39) = P(x \leq 39) - P(x \leq 20)$$

$$= F_{100, 0,3}(39) - F_{100, 0,3}(20)$$

$$= 0,996 - 0,004 = 0,992 = 99,2\%$$

3σ-Umgebung: $3\sigma = 13,74$

Grenzen des 3σ-Bereichs berechnen:

$$\mu - 3\sigma \leq x \leq \mu + 3\sigma$$

$$16,26 \leq x \leq 43,74$$

$$16 \leq x \leq 44$$

Berechnen der Wahrscheinlichkeit, dass ein Wert für k bei der Binomialverteilung innerhalb der 3σ-Umgebung liegt:

$$P(16 \leq x \leq 44) = P(x \leq 44) - P(x \leq 15)$$

$$= F_{100, 0,3}(44) - F_{100, 0,3}(15)$$

$$= 0,996 - 0,004 = 0,992 = 99,6\%$$

VORSCHAU
Lehrersebstverlag
schuldruckportal.de

VORSCHAU
Lehrersebstverlag
schuldruckportal.de

Literaturwerte zu den σ -Umgebungen

Sie haben in Aufgabe 9.7 festgestellt, dass die Symmetrieeigenschaft der Binomialverteilung von n und damit auch von σ abhängig ist. Je größer der Wert von n bzw. σ wird, desto symmetrischer ist die Verteilung bezüglich des höchsten Balkens. Die berechneten Wahrscheinlichkeiten für die σ -Umgebungen werden ebenfalls umso genauer, je größer der Wert von n ist.

Wenn der Wert von σ größer als drei ist, also $\sigma > 3$ gilt, ist die Laplace-Bedingung erfüllt, und es können mit guter Näherung die folgenden Werte verwendet werden:

Für eine binomialverteilte Zufallsvariable X ist die Standardabweichung $\sigma = \sqrt{n \cdot p \cdot q} > 3$, die Laplace-Bedingung erfüllt, und es gelten die folgenden Regeln:

- 1 σ -Regel: $P(\mu - \sigma \leq X \leq \mu + \sigma) = 68,3\%$
- 2 σ -Regel: $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 95,4\%$
- 3 σ -Regel: $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 99,7\%$

Raum für Notizen



VORSCHAU
Lehrerselbstverlag
schuldruckportal.de

Oberstudienrätin Ursula Pirkl

Stochastik
selbstorganisiert lernen



Kapitel

Zweiseitiger Signifikanztest



VORSCHAU
Lehrerselbstverlag
schuldruckportal.de

Einführung in die Stochastik

Kapitel 1
Grundbegriffe 7

Kapitel 2
Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
Vierfeldertafel 23

Kapitel 4
Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
Erwartungswert 63

Kapitel 8
Varianz und Standardabweichung 69

Kapitel 9
Normalverteilung 75

Hypothesentests

Kapitel 10
Zweiseitiger Signifikanztest 97

Kapitel 11
Einseitiger Signifikanztest – Fehler beim Testen 109

Kapitel 12
Vertrauensintervall 135

Kapitel 13
Die Betrachtungen zur Normalverteilung 139

Kapitel 14
Anwendung der Normalverteilung 149

Kapitel 10: Zweiseitiger Signifikanztest

Aufgabe 10.1

Grundidee des Testens

Grundsätzlich entspricht das Testen von Hypothesen der Erstellung einer Wahlprognose. Da man nicht alle Wahlberechtigten nach ihrer Stimmabgabe befragen kann, beschränkt man sich bei den Umfragen auf eine mehr oder weniger zufällig ausgewählte Gruppe von Personen und schließt von deren Wahlverhalten auf den Ausgang einer Wahl. Grundsätzliche Vorgehensweisen bei der Auswertung von solchen Erhebungen und auch Fehlern, die dabei eintreten können, sind Gegenstand der hier zu bearbeitenden Thematik.

Als einführendes Beispiel soll nun ein Vorgehen zur Feststellung der Anteile an S- und L-Körnern betrachtet werden.

Aufgabenstellung:

Ein Geflügelzüchter stellt bei einer Züchtergemeinschaft eine Mischung aus S-Körnern und L-Körnern. Für eine bestimmte Sorte müssen 33% S-Körner in der Mischung vorhanden sein. Die Einhaltung dieser Vorgabe wird aufgrund einer zufällig entnommenen Stichprobe überprüft.

1. Entnahme der Stichprobe

Zur Überprüfung soll nun praktisch durch die Entnahme einer Stichprobe von jedem Schüler/Lehrer S aus einem Behälter mit S-Körnern und L-Körnern simuliert werden. Zählen Sie alle entnommenen Körner, tragen Sie die Werte für **Ihre gezogene Stichprobe** in den folgenden Feldern ein und berechnen Sie Ihren Erwartungswert. (Sollten Sie den Zufallsversuch nicht durchgeführt haben, verwenden Sie folgende Werte: Umfang der Stichprobe: $n = 210$, Anzahl S-Körner: 85)

Prozentualer Anteil der S-Körner laut Hersteller: $p = 33\% = 0,33$

Stichprobenumfang n Ihrer Stichprobe: $n =$ _____

Anzahl der S-Körner k in Ihrer Stichprobe: $k =$ _____

Berechnen Sie den Erwartungswert μ für die Anzahl der S-Körner in Ihrer Stichprobe (auf eine ganze Zahl gerundet) für den Sollwert von 33% S-Körnern in der Mischung.

Erwartungswert für die Anzahl der S-Körner $\mu =$ _____

c) Vergleichen Sie die Anzahl der S-Körner mit der Anzahl der L-Körner im Versuch und ziehen Sie eine Schlussfolgerung. Was stellen Sie fest?

Die Anzahl der gezogenen S-Körner in einer Stichprobe stimmt in der Regel mit dem Erwartungswert für die Anzahl von S-Körnern überein, man zieht mehr oder weniger Körner als der Erwartungswert angibt. Wenn man genau den Erwartungswert erhält, müsste man ... sein, ab ...

welcher Abweichung der Stichprobe vom Erwartungswert μ die Lieferung reklamiert wird, von dem man davon ausgehen kann, dass die Vorgabe von $p = 33\%$ nicht eingehalten wurde und das Zuchtergebnis gefährdet wird. Entscheiden Sie nun intuitiv, ob Sie auf Ihrer Seite die Lieferung beanstanden würden.

Als Geflügelzüchter würde ich die Stichprobe beanstanden bzw. reklamieren oder Nein

2. Anwenden der Binomialverteilung für die Beurteilung Ihrer Stichprobe

Eine Hilfe für die Beurteilung Ihrer Stichprobe und die Entscheidung, ob man hier im Falle des Geflügelzüchters die Lieferung reklamieren würde oder nicht liefert, ist die Anwendung von Eigenschaften der Binomialverteilung auf Ihre Stichprobe. Die typische Vorgehensweise wird nun zurück anhand der Abbildung unten erläutert.

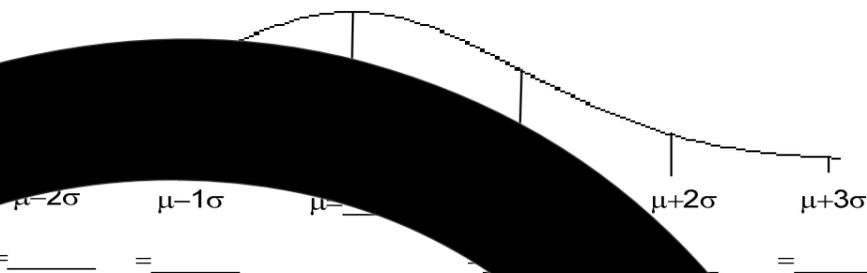
- a) Stellen Sie die Werte Ihrer Stichprobe in Form $B_{n,p}(k)$ dar: $B_{\dots, \dots}(\dots)$
- b) Berechnen Sie für Ihre Stichprobe die einfache, zweifache und dreifache Standardabweichung

Einfache Standardabweichung $1\sigma = \dots$

Zweifache Standardabweichung $2\sigma = \dots$

Dreifache Standardabweichung $3\sigma = \dots$

- c) In der folgenden Abbildung sind die Hüllkurven für eine idealisierte Binomialverteilung und nur die Balken dargestellt, welche die 1σ -, 2σ - bzw. 3σ -Umgebung begrenzen. Berechnen Sie für Ihre Stichprobe aus den oben ermittelten Werten die **ganzzahligen** Werte für diese Grenzen, und tragen Sie die Zahlenwerte in der Abbildung unten ein.



- d) Markieren Sie, farblich abgelesen, mit einem weiteren senkrechten Balken diejenige Stelle der Verteilung der Wahrscheinlichkeiten, die dem Erwartungswert μ entspricht, und ergänzen Sie den folgenden Text.

Die Anzahl der S-Körner in einer Stichprobe liegt

- innerhalb des 1σ -Bereichs.
- innerhalb des 2σ -Bereichs.
- innerhalb des 3σ -Bereichs.

Die Wahrscheinlichkeit, dass bei einer Mischung mit 33% S-Körnern eine Stichprobe zufällig in diesen Bereich fällt, ist gleich / kleiner / größer als _____%.

- e) Entscheiden Sie nun erneut, ob Sie die Lieferung reklamieren würden. Begründen dies.

Beanstanden der Lieferung: Ja oder Nein

Begründung: _____

Zusammenfassung und Veranschaulichung

Ob Sie die Lieferung in diesem Beispiel aufgrund Ihrer Stichprobe annehmen oder beanstanden würden, hängt ab von dem, in welchem Bereich der Binomialverteilung der Wert für die Anzahl der S-Körner fällt. Je näher die Anzahl der S-Körner in der Stichprobe am Erwartungswert liegt, desto eher wird man davon ausgehen, dass die Forderung für eine 33%-Mischung erfüllt ist. Je weiter die Anzahl der S-Körner von der Stichprobe vom Erwartungswert entfernt liegt, desto eher wird man von einem falschen Meinung ausgehen und die Lieferung beanstanden. Betrachtet man bei diesem Versuch eine Vielzahl von Stichproben, so wird deutlich, dass fast alle Schüler und Schülerinnen Stichproben entnommen haben, die innerhalb der zu der Stichprobe gehörenden 3σ -Umgebung liegen. Durch Zufall sind jedoch auch Stichproben vorhanden, die eine große Abweichung vom Erwartungswert aufweisen und außerhalb dieser 3σ -Umgebung liegen. Aufgrund der Stichprobe würde man dann davon ausgehen, dass die geforderte 33%-Mischung nicht eingehalten wurde, obwohl in der Mischung wirklich 33% S-Körner vorhanden sind. Würde man in diesem Fall eine Lieferung beanstanden, so begeht man einen Irrtum und handelt sich ggf. Ärger mit dem Lieferanten ein, wenn dieser nachweisen kann, dass seine Lieferung in Ordnung ist. Die Wahrscheinlichkeit, genau diesen Fehler zu machen, bezeichnet man als **Irrtumsrisiko**.

Bestimmung des Irrtumsrisikos in Abhängigkeit vom Stichprobenumfang

Als Grundlage für die Betrachtungen des Irrtumsrisikos soll weiterhin der Stichprobenversuch aus Aufgabe 10.1 dienen. Allerdings werden nun einheitlich die folgenden Parameter für die entnommene Stichprobe zugrunde gelegt:

Wahrscheinlichkeit $p = 0,33$

Anzahl der S-Körner $n = 231$

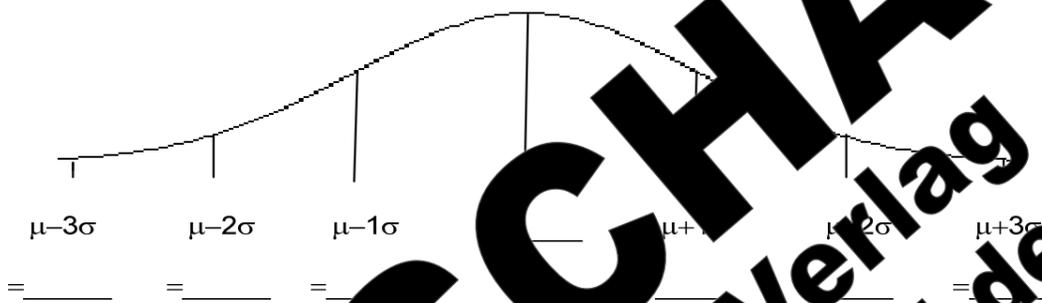
Erwartungswert für S-Körner: $\mu = np = 231 \cdot 0,33 \approx 76$

Einfache Standardabweichung: $1\sigma \approx 7,14$

Zweifache Standardabweichung $2\sigma \approx 14,28$

Dreifache Standardabweichung $3\sigma \approx 21,42$

- a) Tragen Sie die Grenzen für die jeweiligen σ -Umgebungen als **ganzzahlige** Werte in der folgenden Graphik ein. Achten Sie darauf, dass man bei der Bestimmung der σ -Umgebungen immer zum Erwartungswert hin rundet. (Vgl. Kapitel 9 Aufg. 9.8)



- b) Es werden von 8 verschiedenen Personen zufällig Stichproben entnommen. Für die Anzahl der S-Körner ergeben sich folgende Werte.

Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7	Person 8
59	71	65	93	98	54	89	

Markieren Sie die Grenzen für die ermittelte Anzahl an S-Körnern bei den 6 Stichproben in der Abbildung oben mithilfe von vertikalen senkrechten Balken und ergänzen Sie:

Die Stichproben der Personen _____ liegen mit _____ S-Körnern innerhalb der 1 σ -Umgebung.

Die Stichproben der Personen _____ liegen mit _____ S-Körnern innerhalb der 2 σ -Umgebung.

Die Stichproben _____ mit _____ S-Körnern innerhalb der 3 σ -Umgebung.

Die Stichproben _____ mit _____ S-Körnern außerhalb der 3 σ -Umgebung.

Basierend auf dem Testergebnis muss nun jede Person mithilfe der Ergebnisse der Stichprobe entscheiden, ob man davon ausgehen kann, dass die Lieferung in Ordnung ist und angenommen werden sollte, oder ob man davon ausgehen muss, dass das Mischgut nicht in Ordnung ist, weil die Stichprobenerwartungswert abweicht. Eine Entscheidung wird hier mithilfe der σ -Umgebungen und den sogenannten **Annahme-** und **Ablehnungsbereichen** für eine Stichprobe vorgenommen. Die σ -Umgebungen als Grundlage für die Entscheidung, so ergeben sich die Bereiche für die Ermittlung dieser Annahme- und Ablehnungsbereiche.

- c) Betrachtung der 3 σ -Umgebung

Bei einer Binomialverteilung liegen mehr als 99% aller zufällig entnommenen Stichproben im **3 σ -Bereich**, d.h. in weniger als 1% der Fälle liegen Stichprobenergebnisse außerhalb dieses Bereiches. Es ist üblich, diese große Abweichung vom Erwartungswert als **signifikante Abweichung** zu bezeichnen.

Werte innerhalb des 3 σ -Bereichs geben den **Annahmehereich** für die Stichprobe an. Werte außerhalb des 3 σ -Bereichs geben den **Ablehnungsbereich** für die Stichprobe an.

Die **Grenzwerte** für den **Ablehnungsbereich** sind: $\mu - 3\sigma$ links und $\mu + 3\sigma$ rechts = _____

Die Stichproben mit den Werten _____ liegen hier im **Annahmehereich** für die Lieferung und hier angenommen.

Die Stichproben mit den Werten _____ liegen im **Ablehnungsbereich** der 3 σ -Umgebung und man wird sich hier mit der Entscheidung die Lieferung zu reklamieren.

Bei einem **Wartumsrisiko** für eine unberechtigte Ablehnung bzw. Reklamation gibt die Wahrscheinlichkeit an, dass eine Stichprobe zufällig außerhalb der 3 σ -Umgebung liegt, und hat hier den sehr kleinen Wert von _____. man hat ein sehr kleines Risiko, sich wegen einer unberechtigten Reklamation mit dem Lieferanten Ärger einzuhandeln, ganz anders bei der Annahme der Lieferung ein relativ großes Risiko ein, dass die Lieferung nicht stimmt und sich die Tiere nicht gut entwickeln.

- d) Betrachtung der 2 σ -Umgebung

Bei einer Binomialverteilung liegt die Wahrscheinlichkeit, dass eine Stichprobe in den **2 σ -Bereich** fällt, bei 95%, d.h. 5% aller Stichproben liegen durch Zufall nicht in diesem Bereich. Es ist üblich, diese recht große Abweichung vom Erwartungswert als **signifikante Abweichung** zu bezeichnen.

Werte innerhalb des 2 σ -Bereichs geben den **Annahmehereich** für die Stichprobe an. Werte außerhalb des 2 σ -Bereichs geben den **Ablehnungsbereich** für die Stichprobe an.

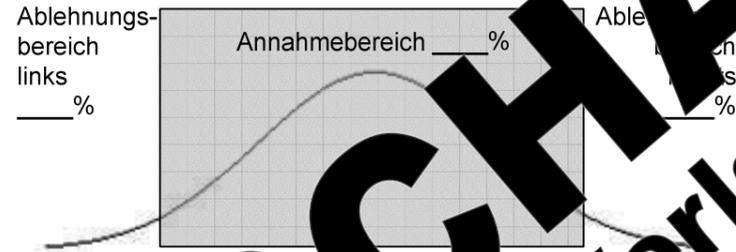
Die **Grenzwerte** für den **Ablehnungsbereich** sind: $\mu - 2\sigma$ links und $\mu + 2\sigma$ rechts = _____

Die Stichproben mit den Werten _____ liegen im **Annahmehereich** für die Lieferung und hier angenommen.

Die Stichproben mit den Werten _____ liegen im **Ablehnungsbereich**, und die Lieferung wird hier nicht angenommen.

Bei einem **Wartumsrisiko** für eine unberechtigte Ablehnung bzw. Reklamation gibt die Wahrscheinlichkeit an, dass eine Stichprobe zufällig außerhalb der 2 σ -Umgebung liegt, und hat hier den Wert von _____. D.h., man hat ein relativ kleines Risiko, sich wegen einer unberechtigten Reklamation mit dem Lieferanten Ärger einzuhandeln.

Lösungsweg für ein Irrtumsrisiko von 16%



Rechnung:



(Kontrollergebnis: $A = \{37, \dots, 40\}$)

Aufgabe 10.4

Irrtumsrisiko bei vorgegebenem Wert für k berechnen

Bei dem Körnerversuch aus Aufgabe 10.1 ($p = \frac{1}{3}$) wird festgelegt, dass eine Körnerlieferung reklamiert, wenn sich in einer Stichprobe von 100 Körnern weniger als 20 oder mehr als 45 S-Körner befinden. Es soll das Irrtumsrisiko, mit dem dieser Test arbeitet, ermittelt werden. Zur Veranschaulichung des Annahme- und Ablehnungsbereichs ist eine Normalverteilungskurve, anstelle einer Binomialverteilung, mit einzelnen Säulen vereinfacht dargestellt. Die Hüllkurve, in der die Balken eingetragen werden, verläuft durch den Mittelwert $\mu = 33,3$.

a) Verdeutlichen Sie sich den Lösungsweg anhand des folgenden Beispiels.

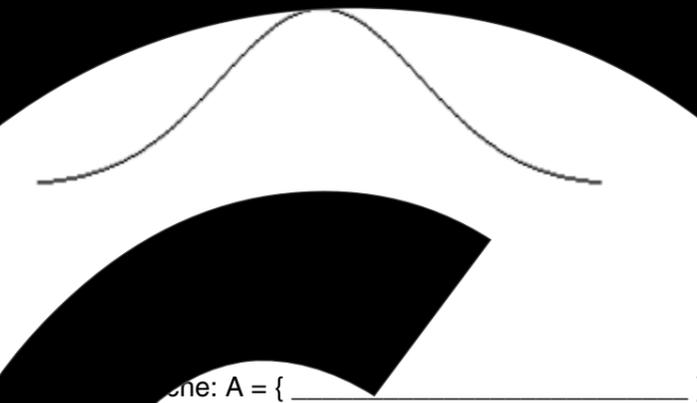
Visualisieren der Annahme- und Ablehnungsbereiche: Rechnung:



Der Test arbeitet mit einem Irrtumsrisiko von $\dots\%$

b) Bestimmen Sie das Irrtumsrisiko, wenn die Lieferung für weniger als 20 und mehr als 46 S-Körner in der Stichprobe abgelehnt wird. (Kontrollergebnis: $\alpha = 1,48\%$)

Visualisieren der Annahme- und Ablehnungsbereiche: Rechnung:



Übungen 10.1

Ü10.1 Die Freie Wählergemeinschaft FWG in Zufallsstadt hatte bei der letzten Kommunalwahl 20% der Stimmen erhalten. Um herauszufinden, ob sich der Stimmenanteil bei den kommenden Kommunalwahlen ändern wird, wird rechtzeitig vor der kommenden Kommunalwahl eine Umfrage unter den Wählern durchgeführt. Da die Umfrage durch Zufall ein Ergebnis liefern kann, das man falsche Rückschlüsse auf den momentanen Stimmenanteil ziehen könnte, wird bei der Auswertung das Irrtumsrisiko berücksichtigt.

- a) Bei der Umfrage werden 400 Personen befragt, und das Irrtumsrisiko soll weniger als 5% betragen. Bestimmen Sie den Annahme- und Ablehnungsbereich bei diesem Test, veranschaulichen Sie die Bereiche anhand einer Graphik und beurteilen Sie, ob man bei dieser Umfrage von einem unveränderten Stimmenanteil ausgehen kann, wenn 200 Personen angeben, die FWG wählen wollen.
- b) Der Parteivorstand legt fest, dass nur 100 Personen befragt werden sollen und dass man nur dann, wenn weniger als 20 Personen angeben die FWG nicht zu wählen, von einem veränderten Stimmenanteil ausgehen wird. Veranschaulichen Sie die Aufgabenstellung anhand einer Graphik und berechnen Sie, mit welchem Irrtumsrisiko der Test arbeitet.
- c) Der Parteivorstand legt das Irrtumsrisiko bei einer Umfrage unter 100 Personen auf einen Wert von 5% fest. Veranschaulichen Sie die Aufgabenstellung anhand einer Graphik und berechnen Sie den Annahmebereich für diese Stichprobe, wenn der Stimmenanteil nach wie vor bei 20% liegt.

VORSCHAU
LehrerSelbstVerlag
schuldruckportal.de

VORSCHAU
LehrerSelbstVerlag
schuldruckportal.de

de Übungen: _____

Raum für Notizen



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Signifikanztest – er beim Testen



Einführung in die Stochastik

Kapitel 1	
Grundbegriffe	7
Kapitel 2	
Baumdiagramme und mehrstufige Zufallsversuche	13
Kapitel 3	
Vierfeldertafel	23
Kapitel 4	
Kombinatorische Abzählverfahren	35
Bedingte Wahrscheinlichkeiten	
Kapitel 5	
Einführung in den Begriff der bedingten Wahrscheinlichkeit	47
Kapitel 6	
Definition der Wahrscheinlichkeitsverteilung	61
Wahrscheinlichkeitsverteilung	
Kapitel 7	
Erwartungswert	63
Kapitel 8	
Variation und Standardabweichung	69
Kapitel 9	
Normalverteilung	75
Hypothesentests	
Kapitel 10	
Zweiseitiger Signifikanztest	97
Kapitel 11	
Einseitiger Signifikanztest – Fehler beim Testen	109
Kapitel 12	
Vertrauensintervall	135
Die Betrachtungen zur Normalverteilung	
Kapitel 14	
Anwendung der Normalverteilung	149

Gesamtheit selbstorganisiert erlernen (Bestenfalls 278)

Sie sind vorbehalten. All rights reserved.
 auch auszugsweise, vorbehaltlich der Rechte,
 aus §§ 53, 54 UrhG ergeben, nicht gestattet

SelbstVerlag
 tes & Freunde GmbH, Koblenz (Germany)
 Lehrerselbstverlag.de
 www.f-druck.de

www.f-druck.de

Kapitel 11: Einseitiger Signifikanztest – Fehler beim Testen

Nachdem Sie sich in Kapitel 10 anhand von zweiseitigen Tests schon mit dem Erkennen von Zusammenhängen beim Testen von Hypothesen beschäftigt haben, erfolgen in diesem Kapitel anhand von sogenannten **einseitigen** Tests weitergehende Betrachtungen vor allem hinsichtlich der Fehler, die bei der Beurteilung einer Stichprobe auftreten können.

Aufgabe 11.1

Aufgabenstellung

Ein Forschungslabor hat zur Behandlung einer seltenen, jedoch meist tödlich verlaufenden Krankheit eine verbesserte Therapie entwickelt. Während die alte Therapie in maximal 20% der Fälle eine Heilung ermöglichte, wird aufgrund erster Versuche vermutet, dass die neue Therapie nun eine Heilung bei mehr als 20% der Erkrankten erzielt. Man geht davon aus, dass nun mindestens 40% der Erkrankten geheilt werden können. Es stehen eine kleine Gruppe von 20 Patienten zur Verfügung, an denen die Wirksamkeit der neuen Therapie getestet werden kann. Der Geschäftsführer der Forschungseinrichtung ist, dass die Forschungsergebnisse kurzfristig als Innovation auf einer Ärztenversammlung veranschaulicht und das Recht an einer Pharmakonzern zu kaufen will, wenn **mindestens 8** Personen in dieser Studie geheilt werden. Beurteilen Sie, ob die Entscheidungsregel des Geschäftsführers sinnvoll ist.

1. Schritt: Formulieren von Hypothesen und Zuordnen der Werte aus der Aufgabenstellung

Es werden im Kontext der Aufgabenstellung zwei sich gegenseitig ausschließende Hypothesen formuliert.

- Die **Nullhypothese H_0** bezieht sich **meist** auf einen **gesicherten Wert**. Hier nimmt man an, dass die neue Therapie genauso hilft wie die alte Therapie und in 20% der Fälle heilt.
- Die **Alternativhypothese H_1** bezieht sich hier auf die ungesicherte Annahme, dass die neue Therapie besser wirkt. Man nimmt hier 40% an.

	Nullhypothese H_0			Alternativhypothese H_1		
Die neue Therapie wirkt wie die alte Therapie.				Die neue Therapie ist besser.		
Erwartungswert und Standardabweichung	$n = 20$	$p_0 = 0,2$	$\mu_0 = 4$	$n = 20$	$p_1 = 0,4$	$\mu_1 \geq 8$

Die Verteilung der Testergebnisse wird durch die Nullhypothese H_0 mit der kleineren Wahrscheinlichkeit p in der linken Spalte der Tabelle zu notieren. Da in dieser Aufgabe p_0 kleiner als p_1 ist, wird demnach p_0 hier in der linken Spalte notiert. (Begründung wird in Schritt 2 deutlich.)

Anmerkung: Die Verteilung der Testergebnisse wird durch die Nullhypothese H_0 mit der größeren Wahrscheinlichkeit p in der rechten Spalte der Tabelle zu notieren. Da in dieser Aufgabe p_1 größer als p_0 ist, wird demnach p_1 hier in der rechten Spalte notiert.

2. Schritt: Festlegen und Visualisieren des Annahme- und Ablehnungsbereichs

Für die folgenden Betrachtungen wird hinsichtlich der Visualisierung für die Annahme der Nullhypothese H_0 der in der Aufgabenstellung angegebene Wert $p = 0,05$ verwendet. Färben Sie in der Abbildung alle Balken, welche zur Verteilung von H_0 gehören und alle Balken, welche zur Verteilung von H_1 gehören mit zwei verschiedenen Farben. Zur Übersichtlichkeit werden die Balken jeweils mit einem Zwischenabstand dargestellt.



Markieren der Ablehnungsbereiche für die beiden Hypothesen

Begründen Sie, warum der Bereich $k \leq 7$ als Ablehnungsbereich für H_1 und der Bereich $k \geq 8$ als Ablehnungsbereich für H_0 bezeichnet werden kann.

... ist, detaillierte Balkendiagramme wie in Abb. 11.1, visualisiert man die ...
... entsprechen zu den Überlegungen beim zweiseitigen Test und der Hüllkurven von ...
... idealisierten Binomialverteilungen. Hier werden, wie in Abb. 11.2, nur die beiden Balken ...
... der Entscheidungsgrenze eingetragen.

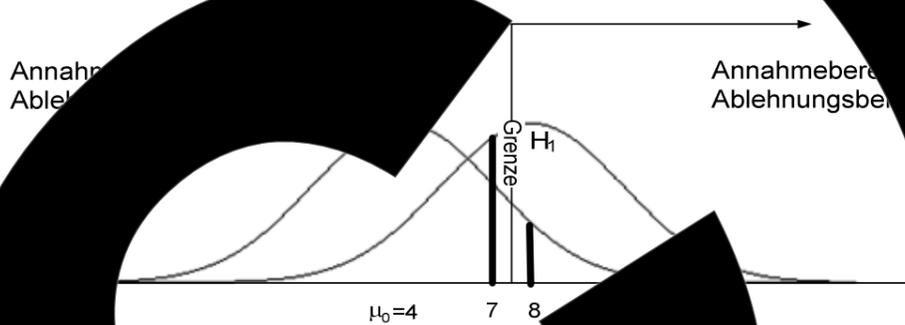


Abb. 11.2

3. Schritt: Betrachtung von Fehlentscheidungen

Für welche der beiden Hypothesen man sich in der Praxis nun entscheidet, hängt vom Ergebnis der Stichprobe ab. Wie Sie im Körnerversuch in den Aufgaben 10.1 und 10.2 gesehen haben, kann eine Stichprobe durchaus ein Ergebnis liefern, das zu einer falschen Entscheidung führt. Wir haben hier gesehen, dass es Stichproben gab, die den Schluss zogen, es sei sich in dem Glas keine 33% S-Körner befinden, obwohl das Mischungsverhältnis tatsächlich bei 33% liegt. Dies führt beim Körnerversuch zu der Fehlentscheidung, die Lieferung unbrauchbar zu erklären und zu reklamieren. Verdeutlichen Sie sich anhand der folgenden Tabelle die Begriffe für die verschiedenen möglichen Fehlentscheidungen hinsichtlich der hier beschriebenen Testsituation. Wirks mit einer Medikamenten...

In dieser Spalte geht man davon aus, dass die Nullhypothese zutrifft.
In dieser Spalte geht man davon aus, dass die Hypothese H1 in Wirklichkeit zutrifft.

Formulieren des Fehlers:	Formulieren des Fehlers:
Obwohl die neue Therapie besser ist, wird aufgrund der Stichprobe irrtümlich angenommen, dass die neue Therapie besser ist.	Obwohl die neue Therapie besser als die alte ist, wird aufgrund der Stichprobe irrtümlich angenommen, dass die neue Therapie nicht besser ist.
Konsequenz der Fehlentscheidung: Blamage durch die Veröffentlichung eines nicht zutreffenden Forschungsergebnisses. (Wird in Wissenschaft und Forschung als schlimmer Fehler angesehen.)	Konsequenz der Fehlentscheidung: Die Erkrankten erhalten kein besseres Medikament, bzw. es wird ggf. unnötigerweise Geld für weitere Forschung ausgegeben.

Wird in der Praxis eine Stichprobe zwei unterschiedliche Hypothesen getestet, so unterscheidet man im Hypothesentest daher zwischen dem **Fehler 1. Art** (oder **Fehler α**) und dem **Fehler 2. Art** (oder **Fehler β**).

Die Reihenfolge der Hypothesen ist festgelegt, die Reihenfolge der Entscheidungen wie folgt vorgenommen:

α -Fehler: – Man ordnet der Nullhypothese H_0 den α -Fehler zu.
– Der α -Fehler ist im Allgemeinen der Fehler, welcher die schwereren Folgen hat.
– Ein schwererer Fehler ist, sollte die Wahrscheinlichkeit für diesen Fehler so klein und nicht größer als die für den α -Fehler sein.

β -Fehler: – Man ordnet der Alternativhypothese H_1 den β -Fehler zu.
– Der β -Fehler ist im Allgemeinen der Fehler, der als weniger schwer angesehen wird.
– Der β -Fehler sollte in der Regel nicht kleiner als der α -Fehler sein, da die Konsequenzen dieses Fehlers ebenfalls nicht unerheblich sind, so gestaltet werden, dass auch dieser Fehler möglichst klein wird.

Beide Fehler können nun, um der Aufgabenstellung eine übersichtliche Struktur zu geben, in Verbindung mit den Werten für die Entscheidungsgrenze ebenfalls tabellarisch dargestellt werden.

In dieser Spalte der Tabelle geht man davon aus, dass die Nullhypothese H_0 zutrifft und das neue Medikament nicht besser ist. Befinden sich in der Stichprobe weniger als 8 geheilte Patienten, spricht das für die Gültigkeit der Nullhypothese H_0 und man entscheidet sich zu Recht für H_0 . Erhält man jedoch durch die Stichprobe mit weniger geheilten Personen, entscheidet man sich zu Unrecht gegen die Nullhypothese und begeht mit dem α -Fehler.

In dieser Spalte der Tabelle geht man davon aus, dass die Nullhypothese H_0 zutrifft und das neue Medikament nicht besser ist. Befinden sich in der Stichprobe mindestens 8 geheilte Patienten, spricht das für die Gültigkeit der Nullhypothese H_0 und man entscheidet sich zu Recht für H_0 . Erhält man jedoch durch die Stichprobe mit weniger geheilten Personen, begeht man einen β -Fehler, da man sich zu Unrecht gegen die Alternativhypothese entscheidet.

Ausgang des Zufallsversuchs

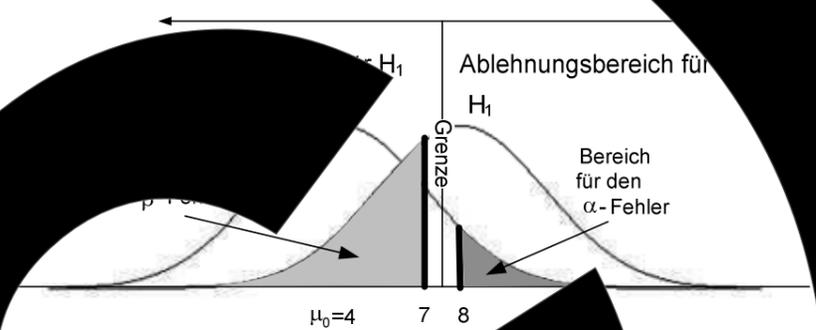
$k \leq 7$	richtige Entscheidung	β -Fehler
$k \geq 8$	α -Fehler	richtige Entscheidung

4. Schritt: Visualisieren und Berechnen des α - und β -Fehlers

Der α -Fehler gibt die Wahrscheinlichkeit an, dass der Wert von k in der Stichprobe in den Ablehnungsbereich von H_0 fällt. Bei der Aufgabe betrifft das alle Balken der Verteilung von H_0 , welche rechts von der Entscheidungsgrenze liegen (vgl. Abb. 11.1).

Der β -Fehler gibt die Wahrscheinlichkeit an, dass der Wert von k in der Stichprobe in den Ablehnungsbereich von H_0 fällt. Bei der Aufgabe betrifft das alle Balken der Verteilung von H_1 , welche links von der Entscheidungsgrenze liegen (vgl. Abb. 11.2).

Um die Fehler zu berechnen, ist es sinnvoll, die Verteilungen zu visualisieren. Man kann dazu die Abbildungen 11.1 bzw. 11.2 zugrundelegen, und die Fehler entsprechend der Markierungen berechnen.



Aus Abb.11.3 kann man nun die Ansätze für die Berechnung der beiden Fehler entnehmen. Berechnen Sie die beiden Fehler, indem Sie die entsprechenden Felder in der Tabelle und ergänzen. Für die Wahrscheinlichkeit der Alternativhypothese H_1 wird der Wert $p_1 = 0,4$ verwendet. (vgl. Aufgabenstellung)

Berechnung der Fehler	Berechnung des α -Fehlers unter Verwendung der Verteilung von H_0	Berechnung des β -Fehlers unter Verwendung der Verteilung von H_1
	$\alpha = P(X \geq \text{---})$ $\alpha = 1 - P(X \leq \text{---})$ $\alpha = 1 - F_{20;0,2}(\text{---})$ Verteilung von H_0	$\beta = P(X \leq \text{---})$ $\beta = F_{20;0,4}(\text{---})$ Verteilung von H_1
	$\alpha = \text{---}$	$\beta = \text{---}$
	$\alpha = \text{---}$	$\beta = \text{---}$
	$\alpha = \text{---}$	$\beta = \text{---}$

5. Schritt: Formulieren der Antwort

Aufgrund der Entscheidungsgrenze des Geschäftsführers ist ein sehr kleiner α -Fehler von 3,2% auf. Damit ist die Wahrscheinlichkeit, sich durch die Veröffentlichung eines falschen Forschungsergebnisses in der Öffentlichkeit zu blamieren, gering. Allerdings ist der β -Fehler mit fast 42% sehr groß.

Aufgabe 11.2

Berechnen Sie, warum man bei gleichem Stichprobenumfang den β -Fehler von Aufgabe 11.1 nur verkleinern kann, wenn man einen größeren α -Fehler zulässt.

Auf der folgenden Seite ist eine Tabelle abgebildet, die die in Aufgabe 11.1 dargestellten Ergebnistabellen zusammensetzt. Um die Vorgehensweise bei der Berechnung des Test zusammenzufassen und zu vertiefen, soll diese Tabelle auf Grundlage der Aufgabenstellung für einen Stichprobenumfang von **50 Personen** ausgefüllt werden. Der Geschäftsführer legt hier fest, dass man von einer verbesserten Therapie ausgehen will, wenn mindestens 15 Personen geheilt werden. Orientieren Sie sich an der modifizierten Aufgabensituation und führen Sie die Schritte 1 bis 4 durch.

Schritt 1: Formulieren Sie die Hypothesen, ermitteln Sie jeweils die Werte für μ_0 und μ_1 . Schritt 2: Skizzieren Sie die beiden Verteilungen mit H_0 und H_1 und zeichnen Sie die Entscheidungsgrenze sowie die zugehörigen Balken für die Entscheidungsregel an geeigneten Stellen in der Abbildung ein. Schritt 3: Formulieren Sie die Fehler, notieren Sie die zugehörigen Werte für α und β und ordnen Sie die Fehler den α -Fehler und β -Fehler zu. Schritt 4: Markieren Sie die Fehlerbereiche in der Abbildung und berechnen Sie die Fehler. Verwenden Sie für H_1 wieder $p_1 = 0,4$.

Formulieren der Hypothesen und Zuordnung der Zahlenwerte	Nullhypothese H_0 Die neue Therapie ist nicht besser als die alte Therapie. $n = 50$ $p_0 = 0,2$ $\mu_0 = 10$	Alternativhypothese H_1 Die neue Therapie ist besser als die alte Therapie. $n =$ $p_1 =$ $\mu_1 > 10$
Markieren Ablehnungsbereiche und Fehler		
Formulieren der Fehler und Konsequenz	Formulieren des Fehler 1. Obwohl die neue Therapie besser als die alte ist, wird aufgrund der Stichprobe irrtümlich angenommen, dass die neue Therapie besser ist. Konsequenz der Fehlentscheidung: Blamage durch die Veröffentlichung nicht zutreffender Forschungsergebnisse. Fehler ist gravierend / weniger gravierend.	Formulieren des Fehler 2. Obwohl die neue Therapie besser als die alte ist, wird aufgrund der Stichprobe irrtümlich angenommen, dass die neue Therapie nicht besser ist. Konsequenz der Fehlentscheidung: Die Erkrankten erhalten kein besseres Medikament bzw. es wird ggf. unnötigerweise Geld für weitere Forschung ausgegeben. Fehler ist gravierend / weniger gravierend.
	$k \leq g_r - 1$ richtige Entscheidung	β-Fehler
	$k \geq g_r$ richtige Entscheidung	
	Berechnung des β-Fehlers unter Verwendung der Verteilung von H_1, wenn $p_1 = 0,4$ gilt	
	$P(X \geq g_r) \leq \alpha$ $1 - P(X \leq g_r - 1) \leq 0,01$ $P(X \leq g_r - 1) \geq 0,99$ $F_{50; 0,2}(g_r - 1) \geq 0,99$ $g_r = 18$	$\beta = P(X \geq g_r H_1)$ $\beta = 0,24$ $\beta \approx 24\%$

Antwort: In der Stichprobe mindestens 18 geheilte Personen befinden sich mit einem Risiko von weniger als 1% behaupten, dass die neue Therapie besser ist als die alte. Der β -Fehler 2. Art beträgt dann 24%.

Übung 11.2

Verwenden Sie für diese Übung die Aufgabenstellung von Aufgabe 11.1. Verwenden Sie dabei für die Stichprobe nun **100 Patienten** und zeigen Sie, indem Sie die folgenden Berechnungen ausführen, dass der β -Fehler bei diesem veränderten Stichprobenumfang einen Wert von 24% annimmt. Da sich wegen der unveränderten Aufgabenstellung bei der Formulierung der Hypothesen und der Konsequenzen keine Veränderungen ergeben, ist dieser Schritt hier nicht erforderlich und erscheint in dieser Tabelle nicht noch einmal.

Formulieren der Hypothesen und Zuordnung der Zahlenwerte	Nullhypothese H_0 Die neue Therapie ist nicht besser als die alte Therapie. $n =$ $p_0 =$ $\mu_0 =$	Alternativhypothese H_1 Die neue Therapie ist besser als die alte Therapie. $n =$ $p_1 =$ $\mu_1 =$
Markieren Ablehnungsbereiche und Fehler		
	$k \leq g_r - 1$ richtige Entscheidung	β-Fehler
	$k \geq g_r$ richtige Entscheidung	
	Berechnung des β-Fehlers unter Verwendung der Verteilung von H_1, wenn $p_1 = 0,4$ gilt	
	$P(X \geq g_r) \leq \alpha$ $1 - P(X \leq g_r - 1) \leq 0,01$ $P(X \leq g_r - 1) \geq 0,99$ $F_{100; 0,2}(g_r - 1) \geq 0,99$ $g_r = 18$	$\beta = P(X \geq g_r H_1)$ $\beta = 0,24$ $\beta \approx 24\%$

Aufgabe 11.5

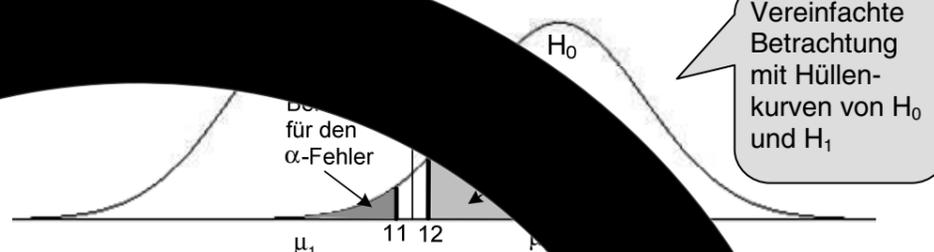
Bei Einnahme des verbesserten Medikaments (vgl. Aufgabenstellung 11.3) wird erwartet mehr Erkrankte geheilt. Allerdings treten nun bei 30% der Patienten Nebenwirkungen auf. Durch eine Beimischung eines weiteren Wirkstoffes soll erreicht werden, dass noch bei 10% der Patienten Nebenwirkungen auftreten. Für einen Test des modifizierten Medikaments stehen 50 Personen zur Verfügung. Man entscheidet, dass man von einer Vergrößerung der Nebenwirkung ausgehen will, wenn in dieser Stichprobe weniger als 12 Personen Nebenwirkungen haben.

- a) Ergänzen Sie die Tabelle unten, erläutern Sie die Fehler bei diesem Test an den Beispielen, und zeigen Sie, dass der α -Fehler einen Wert von etwa 14% hat und der β -Fehler 9% beträgt, wenn man für H_1 die Wahrscheinlichkeit $p_1 = 0,1$ verwendet.
- b) Begründen Sie, warum dieser Test nicht ausreicht.

zu a)

Die Alternativhypothese H_1 ist eine linkssteigende Verteilung, die die linke Spalte zu wählen, da die Wahrscheinlichkeit für H_1 kleiner ist als die Wahrscheinlichkeit von H_0 ist und somit die Verteilung von H_1 links von der Verteilung der Nullhypothese H_0 liegt.

Formulierung des Fehlers und Zahlenwerte	Alternativhypothese H_1	Nullhypothese H_0
	Die neue Therapie hat weniger Nebenwirkungen als die alte. $n = 50$ $p_1 < 0,3$ $\mu_1 < 15$	Die neue Therapie hat die gleichen Nebenwirkungen wie die alte. $n = 50$ $p_0 = \dots$ $\mu_0 = \dots$
Visualisieren und Markieren der Ablehnungsbereiche	Annahmebereich für H_1 Ablehnungsbereich für H_0	Annahmebereich für H_0 Ablehnungsbereich für H_1



Bezug auf H_1	Formulieren des Fehlers:	Formulieren des Fehlers:	Bezug auf H_0
	Obwohl _____	Obwohl _____	

Formulieren der Fehler und Konsequenzen	wird irrtümlich angenommen, dass _____	wird irrtümlich angenommen, dass _____
	Konsequenz der Fehlentscheidung:	Konsequenz der Fehlentscheidung:
	Fehler ist gravierend / weniger gravierend.	Fehler ist gravierend / weniger gravierend.
k _____	richtige Entscheidung	_____ -Fehler
k _____	_____ -Fehler	richtige Entscheidung
Berechnung des β -Fehlers unter Verwendung der Verteilung von H_1 und der Wahrscheinlichkeit $p_1 = 0,1$		Berechnung des α-Fehlers unter Verwendung der Verteilung von H_0 und der Wahrscheinlichkeit $p_0 = 0,3$
$\beta = P(x \geq \dots)$		$\alpha = P(x \leq \dots)$
$\beta = 1 - P(x \leq \dots)$		$\alpha = F_{50;0,3}(\dots)$ Verteilung von H_0
$\beta = 1 - F_{20;0,1}(\dots)$ Verteilung von H_1		$\alpha = \dots$
		$\alpha = \dots \%$

zu b): _____

Aufgabe 11.6
Die Aufgabenstellung 11.5 zeigt, dass der α -Fehler größer als der β -Fehler ist und dass dies mit als nicht akzeptabel angesehen werden kann, soll eine Entscheidungsregel gefunden werden, bei welcher der α -Fehler kleiner als 5% wird. Ermitteln Sie eine geeignete Entscheidungsregel für die Aufgabe 11.3 bzw. Übung 11.2.
Ergebnis: In der Stichprobe dürfen sich höchstens 12 Personen mit Nebenwirkungen befinden; $\beta = 5,8\%$

Formulieren der Hypothesen und Zuordnung der Zahlenwerte	hypothese H_0 $n = \dots$ $p = \dots$ $\mu = \dots$	hypothese H_1 $n = \dots$ $\mu = \dots$
Markieren Ablehnungsbereiche und Fehler	Ablehnungsbereich für H_0 	Ablehnungsbereich für H_1 
Formulieren des Fehlers:	Obwohl H_0 angenommen, wird aufgrund der Stichprobe irrtümlich H_1 angenommen.	Obwohl H_1 angenommen, wird aufgrund der Stichprobe irrtümlich H_0 angenommen.
Konsequenz der Fehlerscheidung:	Fehler ist gravierend / weniger gravierend.	Fehler ist gravierend / weniger gravierend.
α -Fehler	β -Fehler	α -Fehler
Berechnungen zum α -Fehler	Berechnungen zum β -Fehler	Berechnungen zum α -Fehler

Aufgabe 11.7

Links- und rechtsseitiger Test

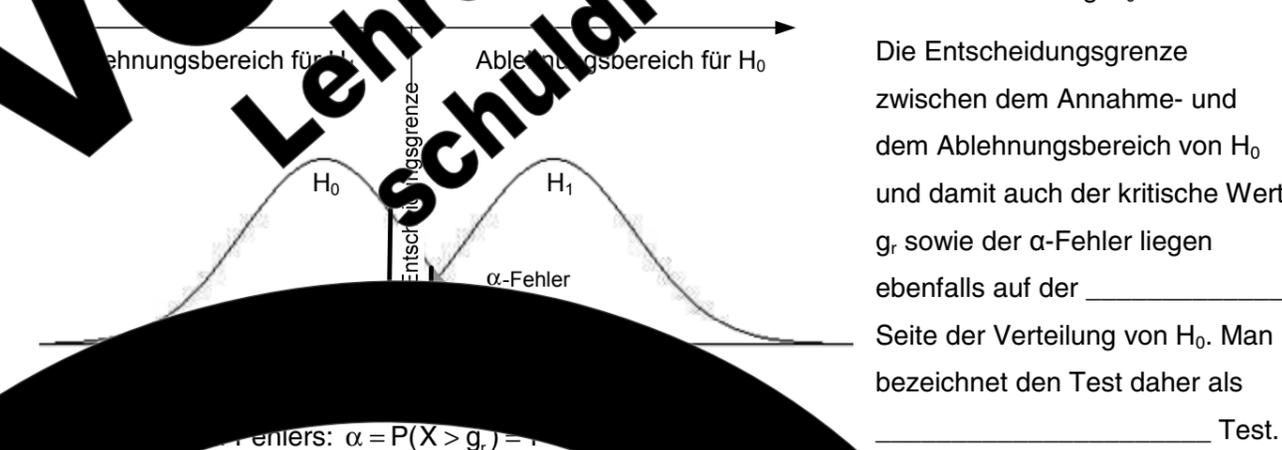
Verdeutlichen Sie sich die beiden folgenden Abbildungen und ergänzen Sie die Lücken im Text.

Linksseitiger Test



Berechnen des α -Fehlers: $\alpha = P(X < g_l - 1)$ mit $\mu = \mu_0$

Rechtsseitiger Test



Fehlers: $\alpha = P(X > g_r) = 1 - P(X \leq g_r)$

Übersicht der Testarten in den bisher behandelten Aufgaben

Testart	Berechnen von	Aufgabe/Übung
linksseitig	α	Aufgabe 11.1, Übung 11.1
linksseitig	α und β	Aufgabe 11.3, Übung 11.3
rechtsseitig	α	Aufgabe 11.4
rechtsseitig	α und β	Aufgabe 11.5

Die Aufgabenstellung für die Tabelle zur Bearbeitung einseitiger Tests in ergänzenden Aufgaben befindet sich auf der Folgeside.

Die Übungen: _____

Formulieren der Hypothesen und Zuordnung der Zahlenwerte	hypothese H_0 $n = \dots$ $p = \dots$ $\mu = \dots$	hypothese H_1 $n = \dots$ $p = \dots$ $\mu = \dots$
Markieren Ablehnungsbereiche und Fehler		
Formulieren des Fehlers:	Formulieren des Fehlers:	Formulieren des Fehlers:
Obwohl H_0 wahr ist, wird aufgrund der Stichprobe irrtümlich H_1 angenommen, α	Obwohl H_1 wahr ist, wird aufgrund der Stichprobe irrtümlich H_0 angenommen, β	Obwohl H_0 wahr ist, wird aufgrund der Stichprobe irrtümlich H_0 angenommen, α
Konsequenz der Fehlentscheidung:	Konsequenz der Fehlentscheidung:	Konsequenz der Fehlentscheidung:
Fehler ist gravierend / weniger gravierend.	Fehler ist gravierend / weniger gravierend.	Fehler ist gravierend / weniger gravierend.
kritische Zahl	kritische Zahl	kritische Zahl
Berechnung zum α Fehler	Berechnung zum β Fehler	Berechnung zum α Fehler

VORSCHAU
LehrerSelbstVerlag
schuldruckportal.de

Aufgabe 11.8

Ergänzende Betrachtungen zur Abhängigkeit der Fehler vom Stichprobenumfang

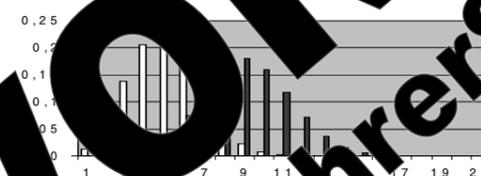
Wenn man die Werte für die beiden Fehler aus Aufgabe 11.7 sowie n und α vergleicht, erkennt man, dass bei einem gleich bleibenden α -Fehler der β -Fehler bei einem kleineren Stichprobenumfang größer wird. Dieser Zusammenhang soll anhand der folgenden Beispiele näher untersucht werden. Dazu werden die im Folgenden abgebildeten Verteilungen herangezogen. (Um die Übersichtlichkeit zu erhalten sind die Balken Zwischenabstände.) Der kritische Wert für die Entscheidungsregel wird jeweils so gewählt, dass der α -Fehler etwa bei 3% liegt.

a) Markieren Sie jeweils näherungsweise die Grenzen zwischen den Ablehnungs- und Annahmebereichen. Ergänzen Sie die fehlenden Berechnungen:

Verteilungen $H_0: p = 2\%$ $H_1: p = 4\%$ $n = 20$ $n = 100$
 Größe des α -Fehlers $\alpha = 3,1\%$
 Größe des β -Fehlers $\beta = 41,6\%$ $\beta = 9,6\%$ $\beta = 9,6\%$ $\beta = 9,6\%$ $\beta = 9,6\%$

$B_{20;0,2}(k)$ und $B_{20;0,4}(k)$

kritische Zahl $k = 1$

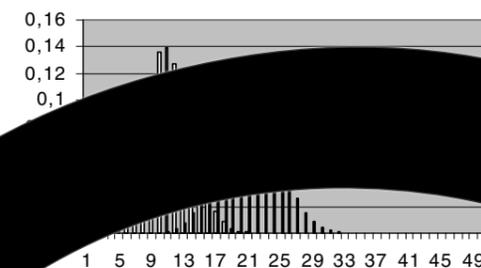


Wert für α aus Aufgabe 11.1
 $\alpha = 3,1\%$

Wert für β aus Aufgabe 11.1
 $\beta = 41,6\%$

$B_{20;0,2}(k)$ und $B_{20;0,4}(k)$

kritische Zahl $k = 16$



$\alpha = P(x \geq \dots)$
 $\alpha = 1 - P(x \leq \dots)$
 $\alpha = 1 - F_{\dots; \dots}(\dots)$

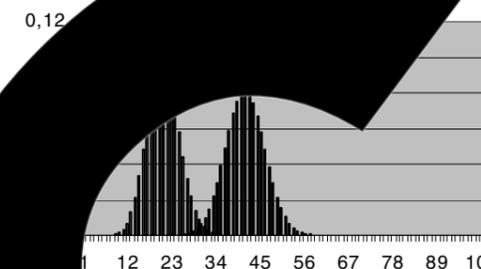
$\beta = P(x \leq \dots)$
 $\beta = F_{\dots; \dots}(\dots)$

$\beta = \dots$

$\beta = 9,6\%$

$B_{100;0,2}(k)$ und $B_{100;0,4}(k)$

kritische Zahl $k = 12$



$\alpha = P(x \geq \dots)$
 $\alpha = 1 - P(x \leq \dots)$
 $\alpha = 1 - F_{\dots; \dots}(\dots)$

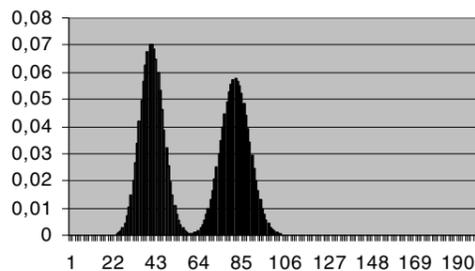
$\beta = P(x \leq \dots)$
 $\beta = F_{\dots; \dots}(\dots)$

$\alpha = \dots$

$\alpha = \dots$

$B_{200; 0,2}(k)$ und $B_{200; 0,4}(k)$

kritische Zahl $k = 51$



$\alpha = P(x \geq \dots)$
 $\alpha = 1 - P(x \leq \dots)$
 $\alpha = 1 - F_{\dots; \dots}(\dots)$
 $\beta = \dots$
 $\beta = F_{\dots; \dots}(\dots)$
 $\beta = 0\%$
 $\beta = 35\%$

- b) Formulieren Sie ein Ergebnis, indem Sie den folgenden Satz ergänzen:
Wenn der gravierendere Fehler kleiner bleiben soll, kann man β -Fehler ebenfalls verkleinern, indem man den Umfang n ...
- c) Erläutern Sie die Entstehung des Scheinungsbildes bei beiden Verteilungen, warum der β -Fehler bei wachsendem Stichprobenumfang abnimmt.

Informationen zum Umfang n in der Praxis des Testens

1. Priorität ... gehalten werden. Hier werden in ... 1% angestrebt.
... Kosten eines Tests, die beispielsweise ... Testverfahren oder hohen Personalaufwand bestimmt werden können. In der Praxis im Allgemeinen niedrig gehalten werden. Daher wird man den Umfang der Probe nur so groß wählen, dass bei einem möglichst kleinen α -Fehler ein ... vertretbares Maß annimmt. Die Größe der Stichprobe wird daher vom erlaubten Fehler bestimmt.
Die Größe ... beispielsweise bei Tests in der ... kann auch durch ... nach oben begrenzt sein ... technische ...

Folgerung: ... Fehler 2. Art als unwichtig zu bewerten, wird man n klein ... Fehler 2. Art ebenfalls gravierend, wird man versuchen, n ...

Aufgabe 11.9

Ergänzende Betrachtungen zur Abhängigkeit des β -Fehlers von der Wahrscheinlichkeit p_1 der Alternativhypothese H_1

Für die folgenden Betrachtungen wird für den Umfang der Stichprobe der Wert $n = 20$ und als Annahmebereich für H_0 der Bereich $A = \{0, \dots, 7\}$ gewählt. Die folgenden Beispiele unten für H_0 eine Wahrscheinlichkeit von $p_0 = 20\%$ verwendet wird, für die der α -Fehler $\alpha = 3,21\%$ jeweils gleich (vgl. Aufgabe 11.1).

Für die Hypothese H_1 wird die Wahrscheinlichkeit p beginnend bei 20% jeweils um 10% vergrößert. Markieren Sie in allen Abbildungen die Entscheidungsgrenze und bestimmen Sie die für den jeweiligen β -Fehler angegebenen Werte durch Berechnung (Zur besseren Übersichtlichkeit erhalten die Balken Zwischenabstände.)

Verteilungen für $n = 20$ Abhängigkeit der Verteilungen von der Größe des β -Fehlers

$H_0: p_0 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $H_1: p_1 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $\beta = P(x \leq \dots)$
 $\beta = F_{20; \dots}(\dots)$
 $\beta = \dots = 96,7\%$

$H_0: p_0 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $H_1: p_1 = 30\% \Rightarrow B_{20; 0,3}(k)$
 $\beta = P(x \leq \dots)$
 $\beta = F_{20; \dots}(\dots)$
 $\beta = \dots = 77,23\%$

$H_0: p_0 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $H_1: p_1 = 40\% \Rightarrow B_{20; 0,4}(k)$
 $\beta = P(x \leq \dots)$
 $\beta = F_{20; \dots}(\dots)$
 $\beta = \dots = 41,59\%$

$H_0: p_0 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $H_1: p_1 = 50\% \Rightarrow B_{20; 0,5}(k)$
 $\beta = P(x \leq \dots)$
 $\beta = F_{20; \dots}(\dots)$
 $\beta = \dots = 13,16\%$

$H_0: p_0 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $H_1: p_1 = 60\% \Rightarrow B_{20; 0,6}(k)$
 $\beta = P(x \leq \dots)$
 $\beta = F_{20; \dots}(\dots)$
 $\beta = \dots = 2,1\%$

$H_0: p_0 = 20\% \Rightarrow B_{20; 0,2}(k)$
 $H_1: p_1 = 70\% \Rightarrow B_{20; 0,7}(k)$
 $\beta = P(x \leq \dots)$
 $\beta = F_{20; \dots}(\dots)$
 $\beta = \dots = 0,13\%$

Ergänzen Sie anhand der Untersuchung die folgenden Sätze.

Je _____ die Werte für die Wahrscheinlichkeiten p_0 und p_1 der Nullhypothese H_0 und der Alternativhypothese H_1 voneinander abweichen, desto weniger überlappen sich die beiden Verteilungen. Das bedeutet, dass bei vorgegebenem konstantem α -Fehler der β -Fehler _____.

Als Folge für die Stichprobengröße ergibt sich demnach:

- Bei einer _____ Differenz der Wahrscheinlichkeiten p_0 und p_1 erhält man auch bei kleinem Stichprobenumfang einen kleinen β -Fehler.
- Bei einer kleinen Differenz der Wahrscheinlichkeiten p_0 und p_1 kann man bei vorgegebenem α -Fehler den β -Fehler nur klein halten, wenn man einen _____ Stichprobenumfang wählt.

Aufgabe 11.10

Erweiterndes Verständnis der Betrachtung zum Begriff Operationscharakteristik

Wie Sie in Aufgabe 11.9 und 11.9 sich mit Sicherheit erkannt haben, hängt die Größe des β -Fehlers von der Differenz der Wahrscheinlichkeiten p_0 der Nullhypothese und p_1 der Alternativhypothese sowie vom Stichprobenumfang n ab. Diesen Zusammenhang kann man in einem Diagramm, in dem man auf der horizontalen Achse die Werte für die Wahrscheinlichkeiten p_1 und auf den vertikalen Achse den Wert für den jeweiligen β -Fehler abträgt, darstellen. Ein derartiges Diagramm wird als **Operationscharakteristik** oder kurz **OC-Funktion** bezeichnet, und es gilt $\beta = OC(p_1)$.

a) OC-Funktion für einen rechtsseitigen Test

In der folgenden Tabelle sind die Ergebnisse von Aufgabe 11.9, ergänzt durch einige weitere Werte für p_1 , angegeben. Tragen Sie diese Werte als Punkte im dazugehörigen Diagramm ein und verbinden Sie diese Punkte zu einer Kurve.

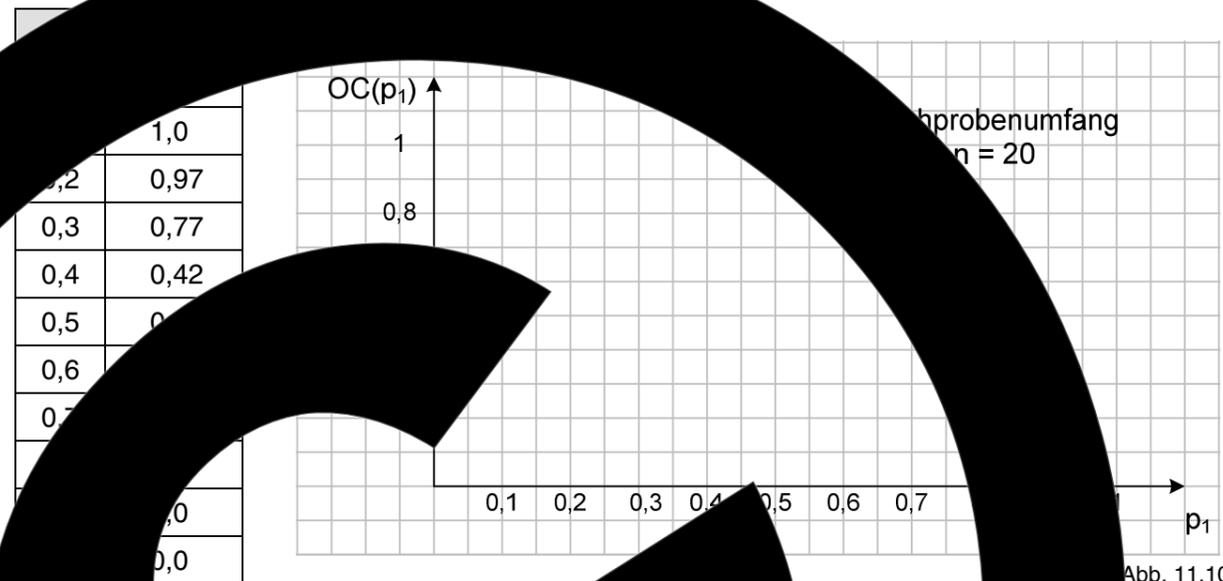


Abb. 11.10.1

Die folgende Tabelle enthält für einen Stichprobenumfang von $n = 100$ die Ergebnisse eines rechtsseitigen Tests. Fertigen Sie ebenfalls den Graphen der OC-Funktion.

p_1	β -Fehler
0,0	1,0
0,1	1,0
0,2	0,98
0,3	0,38
0,4	0,01
0,5	0,00
0,6	0,00
0,7	0,00
0,8	0,00
0,9	0,00
1,0	0,00

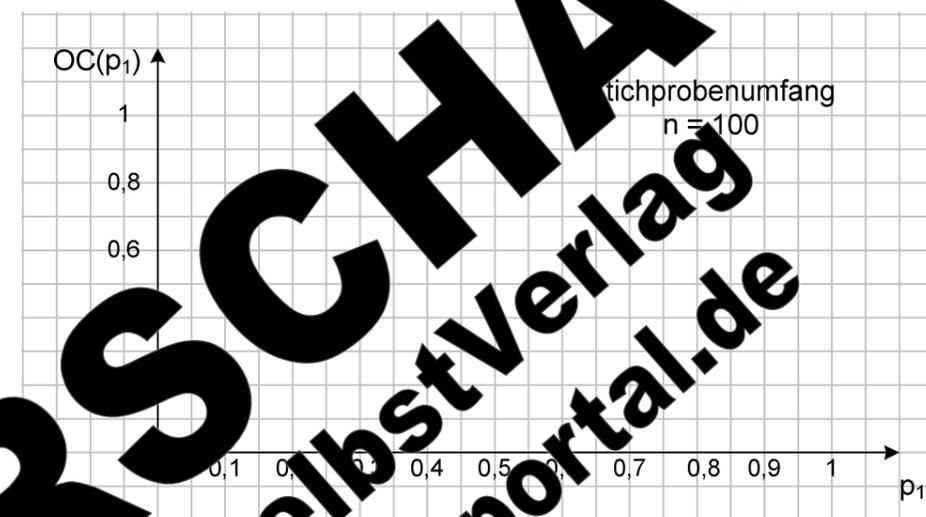


Abb. 11.10.2

Erläutern Sie, welche Gemeinsamkeiten und Unterschiede die Graphen in Abb. 11.9.1 und 11.9.2 aufweisen. Ergänzen Sie die Lücken im Text vollständig.

Wenn bei einem rechtsseitigen Test die Wahrscheinlichkeit p_1 wesentlich kleiner als p_0 ist, nimmt der β -Fehler _____ ab. In diesem Fall liegt die Verteilung von H_1 vollständig im Annahmebereich von H_0 . Wenn bei einem rechtsseitigen Test die Wahrscheinlichkeit p_1 und p_0 etwa _____ sind, liegt der Wert für den β -Fehler nahe bei 1. In diesem Fall liegt der größte Teil der Verteilung von H_1 im Ablehnungsbereich von H_0 . Wenn bei einem rechtsseitigen Test die Wahrscheinlichkeit p_1 und p_0 etwa _____ sind, nimmt der β -Fehler ab. Je größer der Unterschied zwischen p_1 und p_0 wird der β -Fehler. Bei _____ Unterschieden nimmt der β -Fehler schließlich _____ ab, da sich die beiden Verteilungen nicht mehr überlappen und die Verteilung von H_1 vollständig in den Ablehnungsbereich von H_0 fällt.

Je größer n ist, desto kleiner ist der Bereich, in dem sich die beiden Verteilungen auch bei nahe beieinander liegenden Werten von p_1 und p_0 überlappen (vgl. Aufgabe 11.9). Je größer n ist, desto kleiner nimmt der β -Fehler für einen gegebenen Wert von p_1 sehr schnell _____ ab. In der Operationscharakteristik erkennt man daran, dass die Kurve in der Umgebung von p_0 bei _____ abfällt. Je größer n ist, desto steiler verläuft die Kurve als bei kleinen Werten von n . Man sieht auch, dass bei großem n die Schärfe des Tests _____ abnimmt.

b) Operationscharakteristik bei einem linksseitigen Test

Begründen Sie, warum die OC-Funktion in Abb. 11.10.3 einen linksseitigen Test (vgl. Aufg. 10.7) beschreibt, indem Sie den Text vervollständigen:

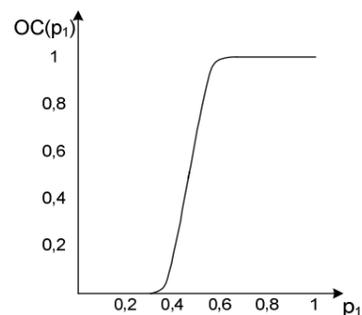


Abb. 11.10.3

Die Verteilung von H_1 liegt bei einem linksseitigen Test _____
von der Verteilung von H_0 . Wenn die Wahrscheinlichkeit p_1 viel
_____ als die Wahrscheinlichkeit von H_0 ist, so gibt es keine
Überlappung der Verteilungen H_0 und H_1 . Die Verteilung von H_1 liegt
dann vollständig im _____ Bereich von H_0 . Damit
_____ Fehler bei einem linksseitigen Test für _____ Werte
von p_1 den Wert Null an. Der β -Fehler _____,
_____ an sich der Wert von p_1 dem Wert von p_0 annähert, und nimmt
den Wert _____, wenn gilt: $p_1 < p_0$.

c) Operationscharakteristik bei einem zweiseitigen Test

Begründen Sie, warum die OC-Funktion in Abb. 11.10.4 einen zweiseitigen Test beschreibt, indem Sie den Text vervollständigen:

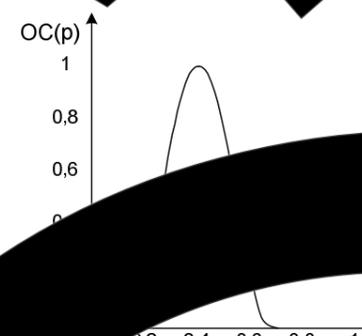


Abb. 11.10.4

Da eine Hypothese bei einem zweiseitigen Test nach links und
nach rechts getestet wird (vgl. Aufg. 10.3 und 10.4), tritt der
Fehler auf _____ Seiten der Hypothese auf. Liegt der
_____ der Wahrscheinlichkeit, welcher der
_____ kommt der Funktionswert der
Operationscharakteristik _____ Wert
an. Der Funktionswert von _____, wenn p
_____ und die Wahrscheinlichkeit der _____ stark von einander
_____.

Übung 11.3

Zusammenfassende Übungsaufgabe zu einseitigen Tests

Testsituation:

In der Automatisierungstechnik werden zur Visualisierung von Abläufen Geräte mit LED-beleuchteten Displays eingesetzt. Diese Geräte sind oft extremen Bedingungen wie beispielsweise großen Temperaturschwankungen, ausgesetzt und haben dadurch eine begrenzte Lebensdauer. Halten die Displays dieser Beanspruchung nicht stand, werden sie nicht richtig aus, so wird meist das gesamte Produkt unbrauchbar und muss ersetzt werden, was zu hohen Kosten für die Erfüllung von Garantieleistungen entstehen können. Daher wird die vom Hersteller angegebene Lebensdauer von einer Firma, die diese Displays einsetzen möchte, geprüft. Man verwendet dazu Thermoschmelzröhren, in denen die Displays untergebracht werden können und in dem sie nur in Zweitelndzyklen von 10°C auf 70°C und zurück, wobei ein Zyklus als Simulation für einen Arbeitstag angesehen wird. Sollte sich bei dem Test herausstellen, dass die Lebensdauer der Displays den Anforderungen im Endgerät nicht entsprechen, bedeutet dies eine Zeitverzögerung für die Produktion, die Verbesserung in dem Hersteller eingefordert werden müssen und die die Testhersteller erforderlich sind. Dies kann zur Folge haben, dass die Konkurrenz Wettbewerbsvorteile hat.

a) Begründen Sie anhand einer Rechnung, warum ein Test, in dem geprüft werden soll, ob die Herstellerangaben zur Lebensdauer der Displays zutrifft über einen Zeitraum von mindestens 5 Monaten laufen muss, wenn für die Endgeräte eine Lebensdauer von 5 Jahren angesetzt werden soll.

Der Hersteller der Displays gibt an, dass mindestens 95% seiner Ware unter extremen Temperaturschwankungen eine Lebensdauer von fünf Jahren hat. Da die Geschäftsleitung aus Erfahrungen mit anderen Herstellern von vergleichbaren Displays weiß, dass nur 80% der Displays die geforderte Lebensdauer von 5 Jahren haben, wird folgender Test vorgeschlagen: Die Geschäftsleitung legt fest, dass der Test mit der Nullhypothese: „80% der Displays haben eine Lebensdauer von 5 Jahren“ bzw. der Alternativhypothese: „95% der Displays haben eine Lebensdauer von 5 Jahren“ bei einem Signifikanzniveau von 1% durchgeführt werden soll. Formulieren Sie die Fehler, die bei diesem Test auftreten können und welche Konsequenzen diese haben. Geben Sie die zugehörige Entscheidungsregel und den Bereich der Ablehnung $\{47, \dots, 50\}$, $\beta = 24\%$). Verwenden Sie die Ergebnisse der Aufgabe 11.2.

Begründen Sie anhand der Konsequenzen für die Produktion, dass das Signifikanzniveau von 1% nur bedingt angemessen ist, und erläutern Sie, wie das Signifikanzniveau des β -Fehler nur verkleinert werden kann, wenn ein größerer α -Fehler vorgezogen wird.

Berechnen Sie beide Fehler, wenn folgende Entscheidungsregel festgelegt wird: Die Nullhypothese geht davon aus, dass die Angaben des Herstellers zur Lebensdauer der Displays zutrifft, wenn mindestens 45 Displays bei dem Test ausfallen. Berechnen Sie α und β für $n = 50$, $\alpha = 1,85\%$, $\beta = 10,4\%$.

Formulieren der Hypothesen und Zuordnung der Zahlenwerte	hypothese H_0 n = ____ p ____ μ ____	hypothese H_1 n = ____ p ____ μ ____
Markieren Ablehnungsbereiche und Fehler		
Formulieren des Fehlers:	Obwohl _____, _____ wird aufgrund der Stichprobe irrtümlich _____ angenommen, _____	Obwohl _____, _____ wird aufgrund der Stichprobe irrtümlich _____ angenommen, _____
Formulieren des Fehlers:	Obwohl _____, _____ wird aufgrund der Stichprobe irrtümlich _____ angenommen, _____	Obwohl _____, _____ wird aufgrund der Stichprobe irrtümlich _____ angenommen, _____
Konsequenz der Fehlentscheidung:	_____	_____
Fehler ist _____	_____	Fehler ist gravierend / weniger gravierend.
Berechnung zum _____ Fehler	_____	Berechnung zum _____ Fehler

Raum zu Berechnungen zur Übung 11.3 d)

Übung 11.4

Nicht immer ist bei Signifikanztests die Wahrscheinlichkeit der Alternativhypothese H_1 in diesem Fall wird auf Grundlage des gegebenen Signifikanzniveaus oft nur der Annahmebereich für die Nullhypothese ermittelt und auch kein Fehler 2. Art berechnet. Die folgende Problemstellung soll nun beispielhaft für diese Testvariante bearbeitet werden.

Aus mehrjährigen Beobachtungen weiß man, dass 70% der Bevölkerung eines Einkaufszentrums die Stadt A für einen Einkaufsbummel wählen und dazu mit dem Auto aus der Innenstadt anfahren. Um die Feinstaubbelastung zu senken, soll eine kostenfreie Parkside-Angebote an Stadtrand eingerichtet und die Parkgebühren in der Innenstadt gestrichelt werden. Der Einzelhandel aufgrund der Maßnahme eine Abwanderung von Kunden in andere Städte und damit einen Umsatzrückgang befürchtet, soll zunächst auf einem Signifikanzniveau von 3% eine entsprechende Umfrage unter 100 Passanten durchgeführt werden.

a) Begründen Sie, warum es sinnvoll ist, ein Test für die Hypothese H_0 der Wert $p_0 = 0,7$ und für die Hypothese H_1 eine Wahrscheinlichkeit $p_1 < 0,7$ anzunehmen.

b) Formulieren Sie beide Hypothesen und begründen Sie, warum es sinnvoll ist, die Hüllkurve für die Verteilung der Hypothese H_1 zu skizzieren, obwohl der Wert von p_1 nicht gegeben wird. Ergänzen Sie die Entscheidungsgrenze und fehlende Werte in der Tabelle sowie in der Abbildung und markieren Sie den Bereich für den α -Fehler.

	Alternativhypothese H_1		Nullhypothese H_0	
Ergebnis und Entscheidung über Zahlenwerte	$n = 100$	$p_1 < 0,7$	$\mu_1 < ______$	$n = ______$
Markieren Ablehnungsbereich	für H_1		Ablehnungsbereich für H_0	
		μ_1	μ_0	

c) Ermitteln Sie den Annahmebereich für die Nullhypothese und formulieren Sie eine geeignete Antwort.



d) Diskutieren Sie die Konsequenzen und deren Konsequenzen.

α -Fehler: _____

β -Fehler: _____

Ergebnis: _____

Raum für Notizen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Vertrauensintervall



Einführung in die Stochastik

Kapitel 1
Grundbegriffe 7

Kapitel 2
Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
Vierfeldertafel 23

Kapitel 4
Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
Erwartungswert 63

Kapitel 8
Varianz und Standardabweichung 69

Kapitel 9
Normalverteilung 75

Hypothesentests

Kapitel 10
Zweiseitiger Signifikanztest 97

Kapitel 11
Einseitiger Signifikanztest – Fehler beim Testen 109

Kapitel 12
Vertrauensintervall 135

Erweiternde Betrachtungen zur Normalverteilung 139

Kapitel 14
Anwendung der Normalverteilung 149

VORSCHAU
Lehrersebstverlag
schuldruckportal.de

Gesamtes Werk selbstorganisiert erlernen (Bestenfalls 278)

Sie behalten alle Rechte vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte,
aus §§ 53, 54 UrhG ergeben, nicht gestattet

SelbstVerlag
F. Druck & Freunde GmbH, Koblenz (Germany)
www.f-druck.de

www.f-druck.de

© 2014 Lehrersebstverlag

© 2014 Ursula Pirkel, Juli 2014

Kapitel 12: Erweiternde Betrachtungen zur Ermittlung eines Vertrauensintervalls

Aufgabe 12.1

Bei Wahlprognosen stützt sich die Vorhersage für das Abschneiden einer Partei an Wahltag auf repräsentative Umfragen. Hierbei wird auf Grundlage dieser Umfrage ein Intervall ermittelt, in dem mit einer vorgegebenen Sicherheit der zu erwartende Prozentsatz der Wähler, die für eine Partei liegen. Dieses Intervall wird als **Vertrauensintervall** oder **Konfidenzintervall** bezeichnet.

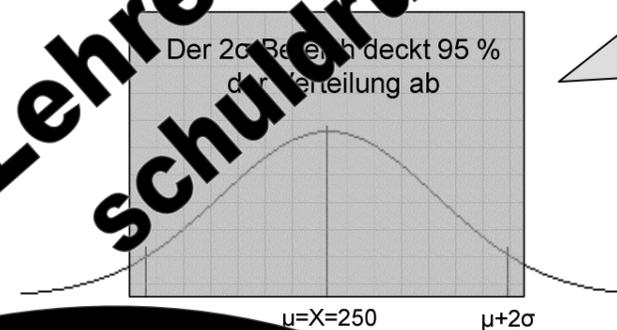
Zahlenbeispiel:

Bei einer repräsentativen Umfrage unter 1000 Personen gaben 250 an, dass sie die Partei A wählen wollen. Es soll auf einem Signifikanzniveau von 5%, über die Verteilung der Wählerumgebung, ein Vertrauensintervall für den zu erwartenden Anteil p ermittelt werden.

Für die Berechnung des Vertrauensintervalls werden zwei Wege vorgestellt:

1. Abschätzen des Vertrauensintervalls über die Wählerumgebung

Für die Ermittlung des Vertrauensintervalls bzw. der Wahrscheinlichkeit p , mit der die Partei gewählt wird, stehen hier nur die Umfrageergebnisse zur Verfügung. Man geht für die zugehörige Verteilung davon aus, dass der Umfragewert $X = 250$ in etwa die Realität widerspiegelt, und verwendet die Mittelwert- bzw. Erwartungswert der Verteilung. Daraus ergibt sich die folgende Abbildung:



Zur Erinnerung:
Wenn die 2σ -Umgebung verwendet wird, teilt sich der Fehler zu je 2,5% auf den linken und rechten Rand auf.

Abb.12.1

Wählen zu wollen, erhält man aus diesen Umfrageergebnissen den Umfragewert $X = 250$ bzw. die Wahrscheinlichkeit p , mit der die Partei gewählt wird, $p = \frac{250}{1000} = 25\%$. Der Vertrauensintervall um diesen Wert herum schwanken. Um die Abweichungen abzuschätzen, d.h. angeben zu können, in welchem Bereich der Wert von p liegen kann, werden die folgenden Betrachtungen durchgeführt. Sie erörtern Sie die einzelnen Rechenschritte und Umformungen.

(1) $\mu = X = 250$
 $p = \frac{250}{1000} = 0,25$

$\mu - 2\sigma \leq X \leq \mu + 2\sigma$

(3) $222,6 \leq X \leq 277,4$

(4) $\frac{222,6}{n} \leq \frac{X}{n} \leq \frac{277,4}{n}$

(5) $\frac{222,6}{1000} \leq \frac{X}{n} \leq \frac{277,4}{1000}$

(6) $22,26\% \leq p \leq 27,74\%$

Ergebnis:
Man kann also abschätzen, dass der Stimmenanteil der Partei zum Zeitpunkt der Umfrage zwischen 22,26% und 27,74% liegt.

2. Berechnung des Vertrauensintervalls für p

Die Herleitung einer Formel zur Berechnung des Konfidenzintervalls und die Lösung der entstehenden Gleichung erscheint auf den ersten Blick im Vergleich zum oben beschriebenen Weg aufwendig. Bei Verwendung einer entsprechenden Rechnertechnologie (z.B. CAS-System) kann man mithilfe der hergeleiteten Formel das Konfidenzintervall für p lediglich durch Eingabe der beiden Werte von n und X computerunterstützt schnell und einfach berechnen lassen.

a) Herleitung

Der Wert für die Anzahl der Stimmen X stimmt mit den Überlegungen überein. Die Schätzung $\hat{p} = \frac{X}{n}$ stimmt mit den Überlegungen überein. In der folgenden Gleichung wird hier eine Gleichung für die Berechnung des Vertrauensintervalls für p aufgestellt. Erläutern Sie die einzelnen Schritte und Umformungen.

(1) $\mu - 2\sigma \leq X \leq \mu + 2\sigma$

(2) $X \geq \mu - 2\sigma$

$\frac{X}{n} \geq \frac{\mu - 2\sigma}{n}$ und $X - \mu \leq 2\sigma$

(4) $-(X - \mu) \leq 2\sigma$ und $X - \mu \leq 2\sigma$

(5) $|X - \mu| \leq 2\sigma$

(6) $|X - np| \leq 2\sqrt{np(1-p)}$

(7) $\left| \frac{X}{n} - p \right| \leq 2\sqrt{\frac{p(1-p)}{n}}$

(8) $\left| \frac{X}{n} - p \right| \leq 2\sqrt{\frac{p(1-p)}{n}}$

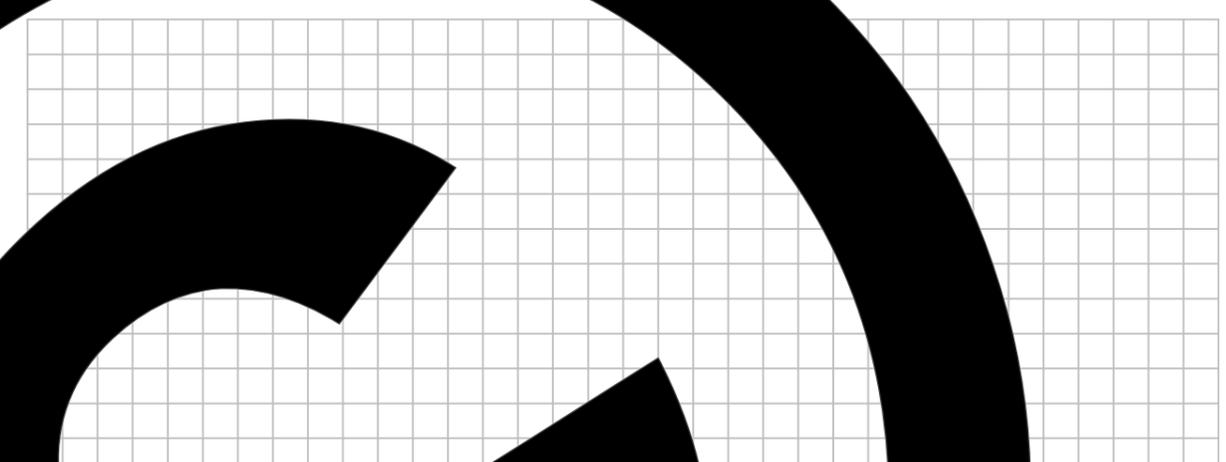
(9) Da nur die beiden Randfälle von p betrachtet werden, kann zur Vereinfachung der Rechnung ein Gleichheitszeichen verwendet werden.

$\left(\frac{X}{n} - p \right)^2 = \frac{4}{n} p(1-p)$

b) Berechnung von p für die gegebenen Umfragewerte

Zeigen Sie, dass die folgende Gleichung nach p bzw. mithilfe Ihres Taschenrechners, dass die Gleichung für p ergibt.

$\left(\frac{X}{n} - p \right)^2 = \frac{4}{n} p(1-p)$



Raum für Notizen:



Oberstudienrätin Ursula Pirkl

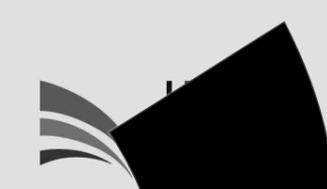
Stochastik

selbstorganisiert lernen



Kapitel

Grundlegende Betrachtung
Normalverteilung



Wie Sie sicherlich erkannt haben, verschiebt der Wert von p die Funktion $f(x)$ nach rechts und links und n beeinflusst die Höhe bzw. Breite der Verteilung und damit den Abstand von a nach b . Die Problem haben auch die Mathematiker De Moivre (1667–1754), Laplace (1749–1827) und Gauss (1777–1855) erkannt. Überlegungen dieser drei Mathematiker waren die Grundlage dafür, eine Funktion $f(x)$ zu entwickeln, die unabhängig von den Werten p und n ist und somit die standardisierte bzw. normierte Verteilung darstellt, mit der Wahrscheinlichkeiten für beliebige Werte von n und p berechnet werden können. Diese Verteilung wird heute als **Normalverteilung** bezeichnet.

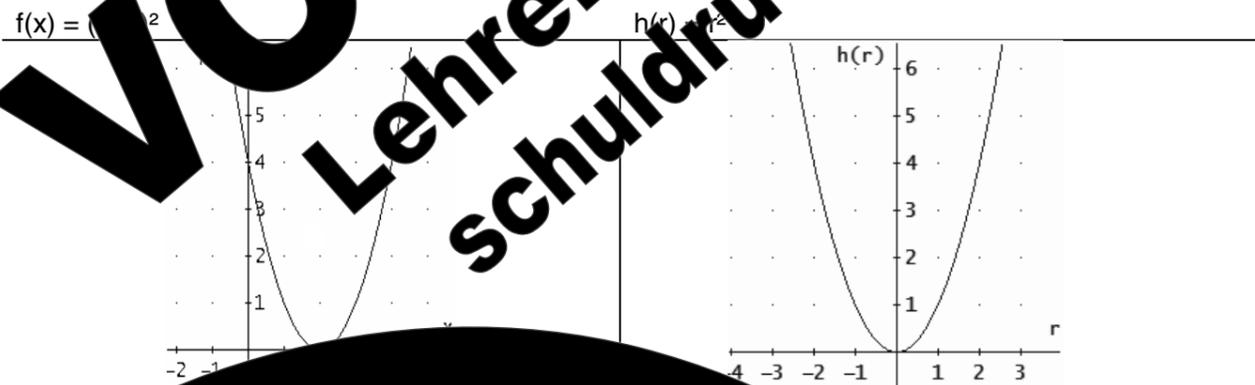
In den nun folgenden Aufgaben wird das Verfahren der Normierung betrachtet und somit erläutert, warum man aus dem Wert der Zufallsvariablen $X = x$ den Erwartungswert μ und die Standardabweichung σ eine neue Variable z errichtet, mit der man dann ähnlich wie bei der Binomialverteilung die Tabelle für Normalverteilung nutzen kann. Hier jedoch einige Kenntnisse bezüglich der Transformation von Koordinatensystemen benötigt werden. Folgt mit Aufgabe 13.2 zunächst ein Einschub, der die notwendigen Schritte hierfür erklärt bzw. wiederholt.

Aufgabe 13.2 Transformation von Koordinatensystemen für das Verschieben von Funktionen

Am Beispiel einer Parabel wird erläutert, welche Transformation eine Funktion verschoben wird.

Achtung: Im transformierten Koordinatensystem wird die horizontale Achse mit r bezeichnet.

Koordinatensystem waagrecht: r -Achse



Gegeben seien die folgenden

$f(0) = 4$	$f(0) = h(-2) = (-2 - 2)^2 = 4$
$f(1) =$	$f(1) = h(-1) =$
$f(2) =$	$f(2) = h(0) = h(2 - 2) =$
$f(3) =$	$f(3) = h(\quad) = h(\quad)$
$f(\quad) =$	$f(4) = h(\quad) = h(\quad)$
	$f(x_0) = h(r_0) = h(\quad) = (\quad)^2$

Sie erkennen, dass die Funktionswerte von $h(r)$ und $f(x)$ identisch sind, wenn man die Verschiebung von $f(x)$ um 2 in Richtung positiver x -Achse berücksichtigt. Dies ist die Verschiebung $r_0 = x_0 - 2$ den gleichen Funktionswert $f(x)$ an $h(r)$.

$f(x) = f(x - 2)$

Aufgabe 13.3 Normierung der Binomialverteilung

a) Verschieben der Binomialverteilung in den Ursprung eines Koordinatensystems

Das Balkendiagramm wird durch eine Verschiebung, entsprechend dem Beispiel, so verschoben, dass der höchste Balken, also der Erwartungswert, an der x -Achse liegt. In der Abbildung werden nur der höchste Balken und die Hüllkurve dargestellt.



Zeichnen Sie an einer beliebigen Stelle $X = k$ im linken Diagramm und an der entsprechenden Stelle im rechten Diagramm einen senkrechten Balken ein und begründen Sie, warum sich auf den horizontalen Achsen aus der x -Koordinate $X = k$ im linken Diagramm, die y -Koordinate $Y = k - \mu$ im rechten Diagramm ergibt.

Begründen Sie, warum für die Wahrscheinlichkeiten $P(X)$ und $P(Y)$ gilt: $P(X = k) = P(Y = k - \mu)$

b) Normierung der Breite

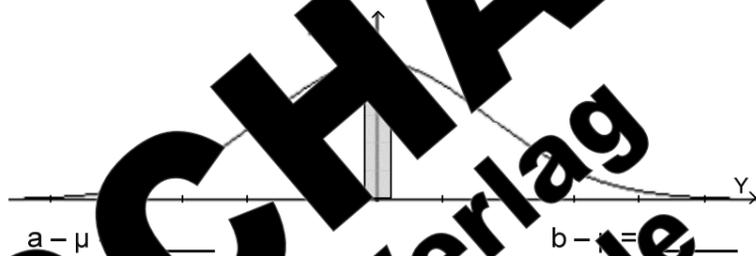
Wie Sie bei der Binomialverteilung für jedes n eine andere Höhe und andere Breite.

Bestimmen Sie für die Binomialverteilungen den Erwartungswert μ und die Standardabweichung σ sowie die Wahrscheinlichkeit für das Eintreten des Erwartungswertes, also $P(X = \mu)$.

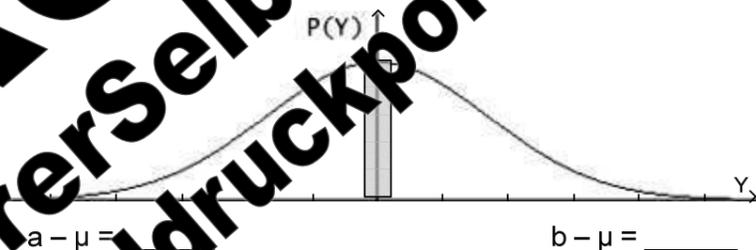
Bestimmen Sie aus der F-Tabelle jeweils die ersten Werte $k = a$ und $k = b$ ab, die $F_{n,p}(a) > 0$ und $F_{n,p}(b) \leq 1$. Ermitteln Sie die um μ verschobenen Werte $a - \mu$ und $b - \mu$ und tragen Sie alle Werte an die vorgesehenen Stellen ein.

Berechnen Sie auch die Breite d der Verteilung über die Bildung der Differenz $a - b = a - \mu - (b - \mu)$ und ergänzen Sie die Werte an den vorgesehenen Stellen ein. Ergänzen Sie den nachfolgenden

- Bsp. 1 $n = 100$
- $p = 0,3$
- $\mu =$ _____
- $P(X = \mu) =$ _____
- $\sigma =$ _____
- $a =$ _____
- $b =$ _____
- Breite $d =$ _____



- Bsp.2 $n = 100$
- $p = 0,5$
- $\mu =$ _____
- $P(X = \mu) =$ _____
- $\sigma =$ _____
- $a =$ _____
- $b =$ _____
- Breite $d =$ _____



Anhand der beiden Werte a und b sowie μ für die Breite erkennt man, dass hier beispielsweise die Randwerte $a - \mu$ unterschiedlich sind und die $B_{100,0,5}(k)$ -Verteilung _____ als die $B_{100,0,3}(k)$ -Verteilung ist.

Damit ergibt sich die Aufgabe 13.1 schon erwähnt, bei der Flächen unter der Normalverteilung zwischen den Grenzen. Um dieses Problem zu lösen, wird die Breite der Hüllkurve für alle k gleich gemacht und somit $a - \mu$ unterschiedlich gemacht.

Man wendet man die Ihnen bereits bekannten Zusammenhänge an. Bei einer Normalverteilung mit $\sigma > 3$ (Laplacebedingung vgl. Aufgabe 9) ist die Wahrscheinlichkeit, dass ein Wert für k in einen der Sigma-Bereiche fällt, für den 1σ -Bereich 68,3%, für den 2σ -Bereich 95,4% und für den 3σ -Bereich 99,7%. Interpretiert man diese Wahrscheinlichkeiten dann decken bei diesen Binomialverteilungen jeweils die gleiche Breite der Normalverteilung ab. Diese Eigenschaft kann man zur Normierung der Breite verwenden.

Berechnen Sie den Quotient h aus d und σ und vergleichen Sie die beiden

für $B_{100,0,3}(k)$ gilt $h_{0,3} = \frac{d}{\sigma} = \frac{100 \cdot 0,3 \cdot 0,7}{\sqrt{100 \cdot 0,3 \cdot 0,7}} \approx 1,05$ und für $B_{100,0,5}(k)$ gilt $h_{0,5} = \frac{d}{\sigma} = \frac{100 \cdot 0,5 \cdot 0,5}{\sqrt{100 \cdot 0,5 \cdot 0,5}} \approx 1,00$

Das Ergebnis zeigt, dass sich bei einer Division durch σ für beide Hüllkurven etwa der gleiche Wert ergibt, also nach Anwendung des Stauchungsfaktors $\frac{1}{\sigma}$ beide Verteilungen Hüllkurven die gleiche Breite annehmen.

Man bezeichnet die waagerechte Achse der mit dem Faktor $\frac{1}{\sigma}$ gestauchten Hüllkurve nun mit z und die senkrechte Achse mit $P(Z)$. Ergänzen Sie die Skalierung der z-Achse in der Abbildung mit den entsprechenden Zahlenwerten.



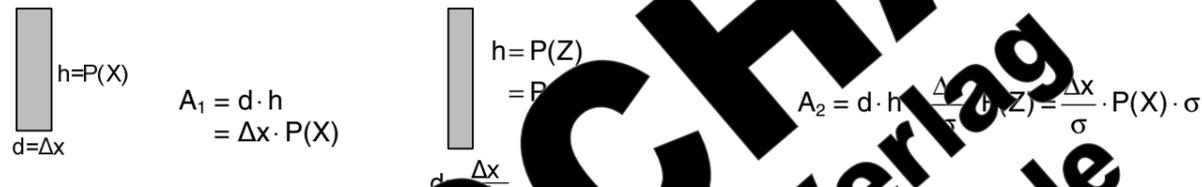
Vergleich des $P(Z)$ -Diagramm mit dem $P(Y)$ -Diagramm und der Ausgangsverteilung im $P(X)$ -Diagramm

- Im $P(X)$ -Diagramm wird für die Berechnung von $P(X = x_i)$ für x_i ein beliebiger Wert k festgelegt, so dass gilt: $X = k$
- Im $P(Y)$ -Diagramm wird der Wert für k nach links um μ verschoben und nun mit y bezeichnet. Für die Berechnung von $P(Y = y)$ gilt: $y = k - \mu$
- Im $P(Z)$ -Diagramm werden die ursprünglich im $P(X)$ -Diagramm festgelegten Werte von k mit z bezeichnet. Die Werte von z sind bezüglich der ursprünglichen Lage von k um μ nach links verschoben, und die Abstände werden aufgrund der Division durch σ gestaucht. Die Position der Werte von z können daher aus dem ursprünglich vorgegebenen Wert k wie folgt berechnet werden: $z = \frac{k - \mu}{\sigma}$

Hüllkurve
 Die Hüllkurve der Verteilung mit dem Faktor $\frac{1}{\sigma}$ ist die Hüllkurve der Verteilung mit dem Faktor $\frac{1}{\sigma}$ und die Höhe der Verteilung bzw. der Balken gleich geblieben ist. Erläutern Sie, welche Eigenschaft für die Summe der Fläche aller Balken bzw. die Fläche unter der Hüllkurve hinsichtlich der Normierung für die Fläche, nämlich: $A_{ges} = \sum_{i=1}^n P(X = x_i) = 1$ bzw. $A_a(b) = \int_a^b f(x) dx = 1$ (Aufgabe 13.1) hat.

c) Normierung der Höhe der Verteilung

Durch das Stauchen der Verteilung ist die ursprüngliche Balkenbreite geworden. Die gleiche Höhe der Balken wird die Gesamtfläche aller Balken daher kleiner. Um diese Veränderung auszugleichen, muss die Höhe der Balken um den gleichen Faktor vergrößert werden. Begründen Sie anhand der Abbildung unten, dass für die neue Höhe $P(Z)$ der Balken gilt: $P(Z) = P(X) \cdot \sigma$.



Die Höhe der Hüllkurve wird durch die Höhe des höchsten Balkens der Binomialverteilung bestimmt. Begründen Sie, warum die Höhe des höchsten Balkens in den beiden Beispielverteilungen aus $B_{100;0,3}$ und $B_{100;0,5}$ (k) jeweils durch den Wert für die Wahrscheinlichkeit $P(X = \mu)$ berechnet wird.

Berechnen Sie nun für beide Verteilungen das Produkt aus $P(X = \mu)$ und σ . Vergleichen Sie die beiden Werte und interpretieren Sie das Ergebnis hinsichtlich der Normierung der Binomialverteilung.

Für $B_{100;0,3}$ (k) gilt: $P(X = \mu) \cdot \sigma =$ _____

Für $B_{100;0,5}$ (k) gilt: $P(X = \mu) \cdot \sigma =$ _____

Für größere Werte von n zeigt sich, dass die Höhe des größten Balkens der Binomialverteilung also die Höhe der Hüllkurve der Normalverteilung annimmt. Um die Balkenbreite ergibt sich die Hüllkurve nebenan abgebildet. **Normalverteilung**



d) Die Glockenkurve von Gauß und ihre Näherungsformel als Funktion Hüllkurve

Gauß hat für die Hüllkurve, die der Form einer Glocke ähnelt und den oben genannten Bedingungen enthält, die folgende Funktionsgleichung entwickelt und mit $\phi(z)$ bezeichnet.

Glockenkurve von Gauß: $\phi(z) = \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{z^2}{2\sigma^2}}$

Da $\phi(z)$ aus einer Verdichtung (Stauchung) der Binomialverteilung entstanden ist, wird $\phi(z)$ auch als **Dichtefunktion** bezeichnet.

Begründen Sie, warum für die Berechnung der gesamten Fläche unter dieser Exponentialfunktion bzw. Glockenkurve der folgende Ansatz verwendet werden kann:

$$A_{ges} = \int_{-\infty}^{\infty} \phi(z) dz = 1$$

Aufgabe 13.4 Die Glockenkurve von Gauß und die Berechnung von Wahrscheinlichkeiten

In der folgenden Abbildung wird hinsichtlich der Berechnung der Wahrscheinlichkeit $P(X \leq k)$ entsprechend zu Aufgabe 13.1 ein Balkendiagramm mit der normierten Hüllkurve der Normalverteilung verglichen.

Balkendiagramm	Normalverteilung
Berechnet man die Wahrscheinlichkeit $P(X \leq k)$ mit Hilfe der F-Tabelle, gilt:	Die Wahrscheinlichkeit $P(X \leq k)$ entspricht nun der Fläche zwischen der x-Achse und unter der Funktion $\phi(z)$, also dem Integral:
$P(X \leq k) = F(k)$	$P(X \leq k) \approx \Phi(z) = \int_{-\infty}^z \phi(z) dz = \int_{-\infty}^z \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{z^2}{2\sigma^2}} dz$
	$\Phi(z)$ wird als Gaußsche Fehlerfunktion bezeichnet und gibt den Wert für die gesamte Wahrscheinlichkeit ($\Phi(\infty) = 1$) an. Damit die Näherung hinreichend genau ist, muss die Bedingung $\sigma > 0$ erfüllt sein.

Da die Funktion $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-0,5z^2}$ keine Stammfunktion besitzt, kann man $\phi(z)$ nicht integrieren. Die Werte für $\Phi(z)$ liegen daher in der **Tabelle zur Gaußschen Summenfunktion** tabellarisch vor.

Damit gilt für die Berechnung von Wahrscheinlichkeiten:
 $P(X \leq k) = F_{n;p}(k) \approx \Phi(z)$

Für sehr große Werte von n (Richtwert $n > 1000$) ist die Dichte der Binomialverteilung extrem schmal und die Glockenkurve deckt die Fläche des Balkens weitgehend exakt ab. In diesem Fall ist die Variable z durch die Standardisierung der Binomialverteilung bestimmt, und die Berechnung erfolgt über den bereits hergeleiteten Zusammenhang:

Die Laplace-Bedingung $\frac{k - np}{\sigma} > 3$ muss erfüllt sein.

Für Binomialverteilungen mit $n > 100$ ist die Annäherung der Glockenkurve an die Balken der Binomialverteilung extrem genau. Um diese Ungenauigkeit auszugleichen, verwendet man in der Regel die Ausgleichszahl 0,5. Das heißt:

$$z = \frac{k - \mu + 0,5}{\sigma}$$

Beispiel
 Eine Binomialverteilung $B_{500;0,3}(k)$ sei verteilt.
 Für die Berechnung der Wahrscheinlichkeit $P(X \leq 140)$ bzw. $P(X > 140)$ ergibt sich:

$$P(X \leq 140) = F_{500;0,3}(140)$$

Genauer Wert mit dem Taschenrechner oder der F-Tabelle ermittelt.

$$P(X > 140) = 1 - P(X \leq 140)$$

Steht keine passende F-Tabelle oder kein Taschenrechner zur Verfügung, wird die WK näherungsweise über die Normalverteilung bestimmt.

Wählen der Variablen z für $k = 140$ mit der Formel für z:

$$z = \frac{k - \mu + 0,5}{\sigma} = \frac{k - np + 0,5}{\sqrt{np(1-p)}} = \frac{140 - 500 \cdot 0,3 + 0,5}{\sqrt{500 \cdot 0,3 \cdot 0,7}} = -0,93$$

(2) Bestimmung des Wertes $\Phi(-0,93)$ durch Ablesen aus der Tabelle der Standardnormalverteilung:

$$P(X \leq 140) = \Phi(-0,93) = 0,1761 = 17,61\%$$

An dem Wert $\Phi(-0,93) = 0,1761$ kann man ablesen, dass die Wahrscheinlichkeit, dass die Zufallsvariable X einen Wert kleiner als -0,93 annimmt, 17,61% beträgt. Für die gesuchte Wahrscheinlichkeit liefert $1 - \Phi(-0,93) = 0,8239 = 82,39\%$.

$$P(X > 140) = 1 - \Phi(-0,93) = 1 - 0,1761 = 0,8239 = 82,39\%$$

Beispiel 2

Die Fernsehshow Mutandenstadel hat in der Zielgruppe Ü70 eine Einschaltquote von 70%. Nach dem Austausch des Moderators befürchtet man, dass die Beliebtheit der Show gesunken ist und als Folge die Einschaltquote zukünftig sinken wird, was einer Verringerung der Werbeeinnahmen für den Sender bedeutet. Daher führt man eine Befragung unter 500 Personen der Zielgruppe durch. Man will von einer gesunkenen Beliebtheit ausgehen, wenn weniger als 335 Personen bei dieser Umfrage angeben, dass ihnen die Sendung noch gefällt. Mit welcher Irrtumswahrscheinlichkeit arbeitet der Test?

Lösung:

Begründen Sie, warum es sich um einen einseitigen Test handelt.

Erläutern Sie die einzelnen Rechenschritte (1) bis (6).

(1) $\alpha = P(X < 335)$

(2) $F_{500;0,7}(335)$

(3) $\alpha \approx \Phi(z)$

(4) $\frac{334 - 350 + 0,5}{10,25} = -1,51$

(5) $\alpha \approx \Phi(-1,51)$

(6) $\alpha \approx 0,0655 = 6,55\%$

Antwort: Die Irrtumswahrscheinlichkeit beträgt 6,6%.

Übungen:

Raum für Notizen:



Oberstudienrätin Ursula Pirkl

Stochastik

selbstorganisiert lernen



Kapitel

Anwendung der Normalverteilung



Einführung in die Stochastik

Kapitel 1
Grundbegriffe 7

Kapitel 2
Baumdiagramme und mehrstufige Zufallsversuche 13

Kapitel 3
Vierfeldertafel 23

Kapitel 4
Kombinatorische Abzählverfahren 35

Bedingte Wahrscheinlichkeiten

Kapitel 5
Einführung in den Begriff der bedingten Wahrscheinlichkeit 47

Kapitel 6
Definition der Wahrscheinlichkeitsverteilung 61

Wahrscheinlichkeitsverteilung

Kapitel 7
Erwartungswert 63

Kapitel 8
Varianz und Standardabweichung 69

Kapitel 9
Normalverteilung 75

Hypothesentests

Kapitel 10
Zweiseitiger Signifikanztest 97

Kapitel 11
Einseitiger Signifikanztest – Fehler beim Testen 109

Kapitel 12
Vertrauensintervall 135

Kapitel 13
Die Betrachtungen zur Normalverteilung 139

Kapitel 14
Anwendung der Normalverteilung 149

Kapitel 14: Anwendung der Normalverteilung

Aufgabe 14.1

Bestimmen der Hilfsgröße z bei gesuchtem kritischen Wert

a) Bestimmen eines linksseitigen kritischen Wertes $k = g_1$ vorliegendem α -Fehler



Rechnung:

Bestimmen von z

$$P(X \leq g_1) \leq 0,025$$

$$P\left(\frac{X - \mu}{\sigma} \leq \frac{g_1 - \mu}{\sigma}\right) \leq 0,025$$

$$\Phi(z) \leq 0,025$$

$$z \leq -1,96$$

2. Bestimmen der Formel für z

$$z = \frac{g_1 - \mu + 0,5}{\sigma}$$

$$-1,96 \leq \frac{g_1 - 15 + 0,5}{2,45}$$

$$-1,96 \cdot 2,45 \leq g_1 - 14,5$$

$$g_1 \leq 36,84$$

3. Rückwärtigen Wert, der $P(X \leq g_1)$ also erfüllt,

Gesamtes Material selbstorganisiert erlernen (Bestenfalls 278)

Alle Rechte vorbehalten. All rights reserved.

Druck, auch auszugsweise, vorbehaltlich der Rechte,
aus §§ 53, 54 UrhG ergeben, nicht gestattet

LehrersebstVerlag

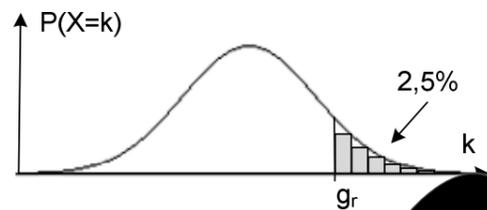
Lehrersebstverlag GmbH, Koblenz (Germany)

www.lehrersebstverlag.de

www.f-druck.de

b) Bestimmen eines rechtsseitigen kritischen Wertes $k = g_r$ bei vorgegebenem Fehler

Darstellung anhand der Binomialverteilung



Gegebene Werte:

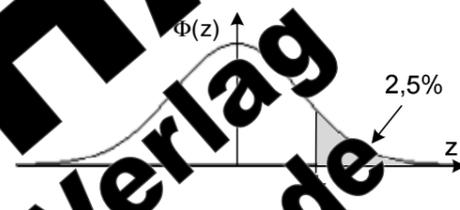
$n = 300$

$p = \frac{1}{6}$

$\alpha = 0,025$

$\Phi(z) \geq 0,975$

Darstellung anhand der Normalverteilung



Rechnung

1. Bestimmen von z

$P(X \geq g_r) = 0,025$

$1 - P(X \leq g_r - 1) = 0,025$

$P(X \leq g_r - 1) = 0,975$

$\Phi(z) \geq 0,975$

$z \geq 1,96$

2. Bestimmen von g_r durch Einsetzen in die Formel für z

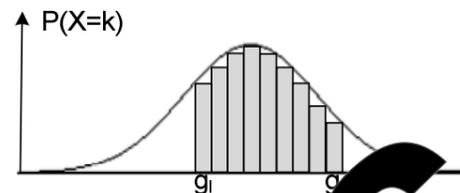
$z = \frac{g_r - 1 - \mu + 0,5}{\sigma} \Rightarrow \frac{g_r - 1 - \mu + 0,5}{\sigma} > -1,96$

$g_r \geq 63,15$

3. Runden auf den nächsten ganzzahligen Wert, der $P(X \geq g_r)$ als α erfüllt, liefert $g_r = 64$.

c) Allgemeine Betrachtungen zum Umgang mit der Normalverteilung und Berechnung der kritischen Werte g_l und g_r bei einer $B_{n,p}(k)$ -Verteilung

Darstellung anhand der Binomialverteilung



Gegebene Werte:

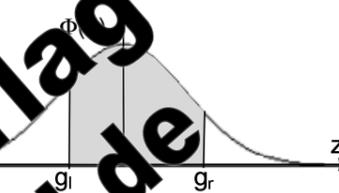
$n = \text{bekannt}$

$p = \text{bekannt}$

$\sigma = \text{bekannt}$

$g_l \leq k$

Darstellung anhand der Normalverteilung



1. Berechnung der Wahrscheinlichkeit $P(g_l \leq X \leq g_r)$

$P(g_l \leq X \leq g_r) = P(X \leq g_r) - P(X \leq g_l - 1)$

$= F_{n,p}(g_r) - F_{n,p}(g_l - 1)$

$= \Phi(z_r) - \Phi(z_l)$

Berechnen von z_l und z_r

$z_l = \frac{g_l - 1 - np + 0,5}{\sqrt{np(1-p)}}$

$z_l = \frac{g_l - np - 0,5}{\sqrt{npq}}$

Bei Ausdrücken der Form $P(X \geq k)$ ergibt sich ein negativer Ausgleichssummand.

$z_r = \frac{g_r - \mu + 0,5}{\sigma}$

$z_r = \frac{g_r - np + 0,5}{\sqrt{np(1-p)}}$

$z_r = \frac{g_r - np + 0,5}{\sqrt{npq}}$

Bei z_l und z_r die Werte von $\Phi(z_r)$

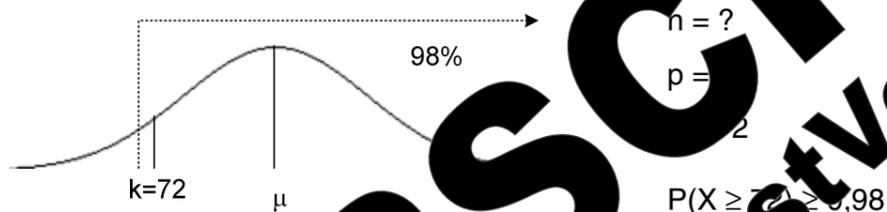
Ergebnis: _____

Aufgabe 14.2

Bestimmen des Stichprobenumfangs bei bekanntem p

Von einem hochwirksamen Schlafmittel ist durch Testreihen bekannt, dass es von 6 Personen keine Nebenwirkungen erzeugt. Für die Zulassung des Arzneimittels soll durch eine Stichprobe überprüft werden, ob diese Angabe zutrifft. Welchen Umfang muss die Stichprobe haben, wenn man mit einer Sicherheit von mindestens 98% mehr als 71 nebenwirkungsfreie Personen erhalten will.

Aufgabenstellung anhand einer Skizze:



a) Die Normalverteilung darf verwendet werden, wenn die Laplace-Bedingung erfüllt ist. Da die Wahrscheinlichkeit p von 1/6 kleiner als 0,5 ist, kann man sich über die Größe der Stichprobe den Wert von σ beeinflussen. Sie führen Rechnung und der folgenden Überlegung, dass der minimale Stichprobenumfang, die Anwendbarkeit der Laplace-Bedingung 65 Personen beträgt.

Aus Laplace-Bedingung

Verwenden der Normalverteilung
 n gilt: $\sigma = \sqrt{\frac{5}{6} \cdot \frac{1}{6}} \Rightarrow \sigma^2 > 9 \Rightarrow n \frac{5}{6} \geq 9 \Rightarrow n = \underline{\hspace{2cm}}$ (minimale Stichprobe)

b) Verdeutlichen Sie sich die Herleitung der Ungleichung (1). Formen Sie die Ungleichung (1) so um, dass sich die Ungleichung (2) ergibt, mit der sich dann n berechnen lässt. Lösen Sie die Ungleichung (2) nach n auf und bestätigen Sie damit die auf der folgenden Seite angegebenen Lösungen.

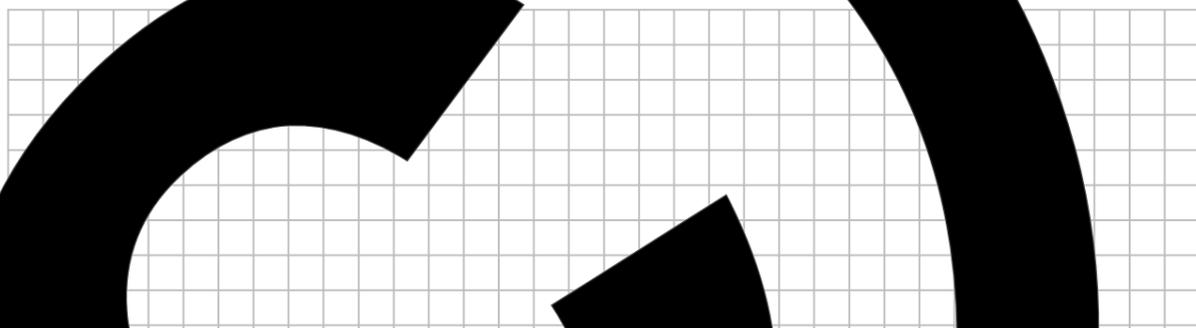
$P(X \geq k) \geq 0,98$

$\Phi\left(\frac{k - n \cdot \frac{1}{6}}{\sqrt{n \cdot \frac{5}{6} \cdot \frac{1}{6}}}\right) \leq 0,02$

$\Phi(z) \leq 0,02$

$z \leq -2,06$ Einsetzen der Werte in Hilfsfunktion

(1) $\frac{71 - n \frac{5}{6} + 0,5}{\sqrt{n \frac{5}{6} \cdot \frac{1}{6}}} \leq -2,06$



(2) $429 - 5n + 4,606\sqrt{n} \leq 0$ Hinweis: Substitution $\sqrt{n} = m$

Lösen Sie die Ungleichung (2) unter Verwendung des Hinweises nach dem auf und bestätigen Sie damit die angegebene Lösung.



$\Rightarrow m^2 - 5m + 4,606 \leq -8,808$

Damit ergibt sich für die gesuchte Anzahl n

$n \geq 81$ also $n = 85$ Personen.

Anmerkung:

Um n zu erhalten, müssen die Werte von m quadriert werden. Da Quadrieren keine Äquivalenzumformung ist, sollte für Gleichungen, bei denen die Lösungen durch Quadrieren der Lösungsvariablen gefunden wurden, eine Probe durch Einsetzen der Ergebnisse in die Ausgangsgleichung vorgenommen werden.

Bei der Probe erkennt man, dass nur $n_1 = 95$ die Lösung ist. Damit ist ab einem Mindeststichprobenumfang von $n = 95$ Personen, gewarant, dass man mit 98% Sicherheit mindestens 72 Personen ohne Nebenwirkungen erhält. Bei einem Stichprobenumfang ist auch die Laplace-Bedingung erfüllt.

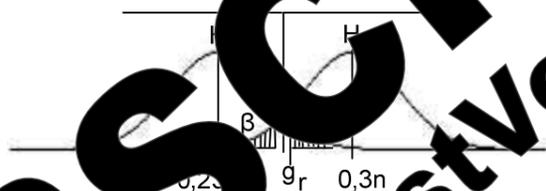
Übungen: _____

Aufgabe 14.3

Stichprobenumfang bei einem rechtsseitigen Test ermitteln, wenn α - und β -Fehler gleichzeitig fest vorgegeben sind

Eine Partei vermutet vor der Wahl, dass ihre Quote von 23% auf 30% gestiegen ist. Um unnötige Wahlkampfkosten zu vermeiden, soll diese Vermutung mit einer Umfrage geprüft werden. Es soll berechnet werden, wie viele Personen befragt werden müssen, wenn der Fehler 1. Art und der Fehler 2. Art gleichzeitig unter 5% liegen sollen.

Verdeutlichen Sie sich den Ansatz und lösen Sie vorstehende Aufgaben anhand dieses Verfahrens.



Bei β Fehler sollen unter 5% liegen, also gilt

$$\begin{aligned}
 P(X \geq g_r) &= \alpha \\
 P(X < g_r) &\leq \beta \\
 P(X \leq g_r - 1) &\geq 0,95 & P(X \leq g_r - 1) &\leq 0,05 \\
 P(X \leq g_r - 1) &\geq 0,95 & F_{n,0,3}(g_r - 1) &\leq 0,05 \\
 \Phi(z) &\geq 0,95 & \Phi(z) &\leq 0,05 \\
 z &\geq 1,65 & z &\leq -1,65 \\
 \frac{g_r - 1 - 0,23n + 0,5}{\sqrt{n \cdot 0,23 \cdot 0,77}} &\geq 1,65 & \frac{g_r - 1 - 0,3n + 0,5}{\sqrt{n \cdot 0,3 \cdot 0,7}} &\leq -1,65 \\
 g_r - 0,23n - 0,5 &\geq 0,694\sqrt{n} & g_r - 0,3n - 0,5 &\leq -0,7561\sqrt{n}
 \end{aligned}$$

Der Ansatz liefert zwei Gleichungen mit den Variablen g_r und n . Zeigen Sie, dass man durch Elimination von g_r und anschließender Einsetzung in die angegebene Ungleichung erhält. Begründen Sie, warum man $n \geq 429,4$ erhält.

$$\begin{aligned}
 \text{II} \quad & g_r - 0,3n - 0,5 \leq -0,7561\sqrt{n} \\
 \text{I} \quad & g_r - 0,23n - 0,5 \geq 0,694\sqrt{n} \\
 \text{II} \quad & -g_r + 0,3n + 0,5 \geq 0,7561\sqrt{n}
 \end{aligned}$$

$$\begin{aligned}
 \sqrt{n} &\geq 20,72 \\
 n &\geq 429,4 \Rightarrow n = 430
 \end{aligned}$$

Antwort: Die Partei muss mindestens 430 Leute befragen, um mit der geforderten Sicherheit ausgehen zu können, dass sich die Quote verbessert hat.

Erläuterungen: _____

**Umschlag
Rückseite
(Innen)**

(unbedruckt)

Hier können Sie noch Vorlagen einfügen

2014-15

VORSCHAU

LehrerSelbstVerlag
schuldruckportal.de

