Vorwort: Die folgenden Ausführungen befassen sich mit einem mathematischen Thema, das aber nicht wie im Mathe-Buch, sondern verständlich erläutert vorgestellt wird. Da der Verfasser aus der Elektrotechnik kommt, hat er zum Teil andere Sichtweisen auf die Dinge. Vielleicht hilft gerade diese Besonderheit beim Verständnis der beschriebenen Mathematik.

Achtung: Trotz aller Sorgfalt bei der Erarbeitung kann keine Garantie für die Richtigkeit der Inhalte übernommen werden.

Inhaltsübersicht Stochastik

Grundbegriffe Stochastik

Ergebnis e oder ω

Einzelergebnis einer Beobachtungsreihe – z. B. mit einem Würfel wird die Zahl 2 gewürfelt.

Ergebnismenge Ω

Menge aller möglichen Ergebnisse – z. B. beim Würfeln können die Zahlen 1, 2, 3, 4, 5 und 6 gewürfelt werden – diese bilden die zugehörige Ergebnismenge Ω.

Ereignis E

Zusammenfassungen von Ergebnissen zu Teilmengen der Ergebnismenge heißen Ereignisse – d. h. der Begriff Ereignis steht für eine Menge. Zum Beispiel könnte man sagen, dass das Ereignis A eintritt, wenn beim Würfeln eine gerade Zahl entsteht, das Ereignis B dagegen bei einer ungeraden Zahl. Beide Ereignisse stehen dann für mehrere Möglichkeiten – sie umfassen also beide eine Teilmenge der Ergebnismenge. Das Ereignis A umfasst die Zahlen 2, 4 und 6, das Ereignis B die Zahlen 1, 3 und 5.

Oft kommt es aber auch vor, dass man den Begriff Ereignis nur für ein einzelnes Ergebnis verwendet – zum Beispiel das Würfeln der Zahl 2. Das hängt immer von der jeweiligen Problemlage ab. Dann heißt das Ereignis Elementarereignis.

Das Ereignis, welches keine Ergebnisse enthält, heißt unmögliches Ereignis – Bezeichnung mit dem Symbol ∅.

Das Ereignis Ω, das alle Ergebnisse enthält, heißt das sichere Ereignis.

Ereignisfeld F

Menge aller Ereignisse eines Zufallsexperiments. Denken wir noch einmal an den Würfel. Der hat ja bekanntlich sechs Zahlen (Ergebnismenge mit sechs Elementen), die wir aber in zwei Ereignisse eingeteilt haben – Ereignis A für gerade und Ereignis B für ungerade Zahlen. Das Ereignisfeld umfasst dann nicht die sechs möglichen Zahlen des Würfels, sondern nur die beiden Ereignisse A und B.

Durchschnitt = empirischer Mittelwert m

Die absoluten Häufigkeiten der einzelnen Ereignisse werden mit dem Ereignis selbst multipliziert. Zum Beispiel wie oft wurde die Zahl 3 gewürfelt, multipliziert mit der Zahl 3 selber. Diese Produkte müssen nun addiert und durch die Anzahl Beobachtungen n dividiert werden.

Liefern 20 Würfelversuche

3mal die Zahl 1,
4mal die Zahl 2,
3mal die Zahl 3,
5mal die Zahl 4,
3mal die Zahl 5 und
2mal die Zahl 6,

dann beträgt der Durchschnitt:

$$m = \frac{3 \cdot 1 + 4 \cdot 2 + 3 \cdot 3 + 5 \cdot 4 + 3 \cdot 5 + 2 \cdot 6}{20} = 3{,}35$$

(Empirischer) Median

Sortierung aller Ereignisse = Messwerte nach ihrer Größe, wobei Messungen mit gleichem Ergebnis auch entsprechend mehrfach einzuordnen sind. Bei ungerader Anzahl Messungen n ist der Median der mittlere Wert, bei geraden Anzahlen das arithmetische Mittel der beiden mittleren Werte. Für unseren oben dargestellten Würfelversuch ergibt sich:

$1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 6, 6$

Die beiden mittleren Werte – da es sich um eine gerade Anzahl an Ereignissen handelt – sind die 3 und die 4, womit als Median die 3,5 entsteht.

(Empirischer) Modalwert

Ereignis mit der größten absoluten Häufigkeit unter allen beobachteten Ergebnissen. Für unser Würfel-beispiel trifft das auf die Zahl 4 zu.

Häufigkeit

Absolute Häufigkeit H(e) oder H(E)

Anzahl des Auftretens des Ergebnisses e bzw. des Ereignisses E im Verlaufe einer Beobachtungsreihe – z. B. die Länge 2 m wird 10 mal gemessen – absolute Häufigkeit beträgt 10.

Absolute Häufigkeit H_n(e) oder H_n(E)

Anzahl des Auftretens des Ergebnisses e bzw. des Ereignisses E im Verlaufe einer Beobachtungsreihe von n Versuchen. Der Wert der absoluten Häufigkeit liegt immer zwischen Null und n. Zum Beispiel die Länge 2 m wird bei 15 Messungen 10 mal festgestellt – absolute Häufigkeit beträgt 10 und n beträgt 15.

Relative Häufigkeit h(e) oder h(E)

Absolute Häufigkeit $H(e)$ des Ergebnisses e bzw. des Ereignisses E geteilt durch die Anzahl der Beobachtungen n insgesamt. Der Wert der relativen Häufigkeit liegt immer zwischen Null (Ergebnis trat nie auf) und 1 (Ergebnis trat bei allen Beobachtungen auf). Zum Beispiel die Länge 2 m wird 10 mal bei insgesamt 15 Messungen festgestellt – relative Häufigkeit beträgt $h(e) = 0{,}67$.

Relative Häufigkeit h_n(e) oder h_n(E)

Absolute Häufigkeit $H(e)$ des Ergebnisses e bzw. des Ereignisses E geteilt durch die Anzahl der Beobachtungen n insgesamt – mit Angabe der Anzahl an Beobachtungen. Der Wert der relativen Häufigkeit liegt immer zwischen Null (Ergebnis trat nie auf) und 1 (Ergebnis trat bei allen Beobachtungen auf). Zum Beispiel die Länge 2 m wird 10 mal bei insgesamt 15 Messungen festgestellt – relative Häufigkeit beträgt $h_{15}(e) = 0{,}67$.

Häufigkeiten zweier Ereignisse E und F

Sind zwei Ereignisse E und F unvereinbar, d. h. die Durchschnittsmenge aus E und F ist leer ($E \cap F = \emptyset$), so gilt:

$$H(E \cup F) = H(E) + H(F)$$

Unvereinbar sind zum Beispiel das Messen der Länge 2 m und das Messen der Länge 4 m. Will man nun die Häufigkeit des Auftretens beider Längen wissen (Vereinigungsmenge von E und F, dargestellt durch $E \cup F$), so addieren sich die Einzelhäufigkeiten. Gleiches gilt für die relativen Häufigkeiten.

Für zwei beliebige Ereignisse E und F gilt:

$$H(E \cup F) = H(E) + H(F) - H(E \cap F)$$

Gibt es Teile der Mengen E und F, die einander überdecken oder überlappen, so dürfen diese Teile für die Gesamthäufigkeit nicht doppelt gezählt werden – das würde passieren, wenn man nur die Formel wie bei den unvereinbaren Ereignissen verwendet. Stattdessen ergibt sich die Gesamtmenge aus der Summe der beiden Teilmengen einmal reduziert um den überlappenden Bereich.

Versuchen wir es mit der Verdeutlichung an einem Beispiel. Gefragt ist die Häufigkeit, wie oft Längen unter 3 m und Längen über 2 m gemessen wurden. Also wäre das eine Ereignis die Länge unter 3 m und das andere Ereignis die Länge über 2 m – soweit sicher klar. Addieren wir jetzt die Häufigkeiten der beiden Ereignisse, wird der Bereich über 2 m bis unter 3 m doppelt gezählt, da er in beiden Ereignissen enthalten ist. Also muss die Häufigkeit des Auftretens von Längen in diesem Bereich einmal vom Resultat subtrahiert werden. Damit finden sich nun alle in Frage kommenden Längen – in unserem Beispiel die von Null bis Unendlich – einfach im Resultat wieder. Gleiches gilt für die relativen Häufigkeiten.

Ist $\bar{E}$ das Gegenereignis zum Ereignis E (z. B. Messungen von 2 m als Ereignis E und Messungen ungleich 2 m als Gegenereignis $\bar{E}$), so gilt:

$$H_n(E) + H_n(\bar{E}) = n$$

bzw.

$$h_n(E) + h_n(\bar{E}) = 1$$

Empirisches Gesetz der großen Zahlen

Je mehr Versuchswiederholungen man vornimmt, desto stabiler sind die Werte für relative Häufigkeiten. Wenn wir bei unserer Länge von 2 m bleiben, die in 15 Versuchen 10 mal gemessen wurde, so kommt es sicher vor, dass bei neuen Versuchen mit 15 Messungen die 2 m auch 9 mal oder 11 mal auftauchen. Das sind dann immer gleich ganze 10 % Abweichung in der relativen Häufigkeit. Messen wir 150 oder gar 1500 mal, wirken sich einzeln vorkommende Abweichungen viel weniger aus – die relative Häufigkeit liegt sehr genau bei 0,67.

Histogramm und Stabdiagramm

Graphische Darstellung der Häufigkeit von Ereignissen in einem Säulendiagramm. Die Höhe der einzelnen Säulen steht für die Größe der absoluten oder relativen Häufigkeit des Sachverhalts, der durch die betreffende Säule veranschaulicht wird. In dem Beispiel mit der zehnfachen Messung von 2 m und der fünffachen Messung von 4 m ergibt das ein Säulendiagramm mit zwei Säulen. Handelt es sich um die Darstellung der absoluten Häufigkeiten, ist die Säule für 2 m 10 Einheiten und die für 4 m 5 Einheiten hoch. Wird im Gegensatz dazu die relative Häufigkeit veranschaulicht, sind die Säulen 0,67 und 0,33 Einheiten hoch.

Beim Stabdiagramm werden die Säulen eines Säulendiagramms einfach durch schmale Stäbe ersetzt – alles andere funktioniert unverändert.

Polygonzugdiagramm

Graphische Darstellung der Häufigkeit von Ereignissen in einem Liniendiagramm in der Regel als Funktion der Zeit, wobei die Punkte der eingetragenen Häufigkeiten einfach von Punkt zu Punkt geradlinig miteinander verbunden werden (es entsteht ein eckiger Linienzug wie ein aufgebogener Polygon). Ein Beispiel für ein Polygonzugdiagramm wäre die Darstellung, wie viele Kunden im Laufe eines Tages in einer Verkaufseinrichtung einkaufen. Sinnvolle Zahlen würden entstehen, wenn man die Kundenzahl für jede volle Stunde der Öffnungszeit im Diagramm einträgt. Man könnte sofort ablesen, in welcher Stunde des Tages der größte Kundenstrom zu bewältigen ist. Polygonzugdiagramm nicht mit Baumdiagramm verwechseln!

(Empirischer) Modalwert

Ereignis mit der größten absoluten Häufigkeit unter allen beobachteten Ergebnissen

Streudiagramm, Streuung und Standardabweichung

Streudiagramm

Graphische Darstellung der bei der Beobachtung ermittelten Zahlenpaare in einem rechtwinkligen Koordinatensystem

Lineare Korrelation

Die Punkte im Streudiagramm lassen sich näherungsweise durch eine lineare Funktion = Gerade annähern.

positive lineare Korrelation	Näherungsgerade monoton wachsend
negative lineare Korrelation	Näherungsgerade monoton fallend

Regressionsgerade

Näherungsgerade, bei der die Quadrate der Abweichungen der Punkte von der gefundenen Geraden am kleinsten sind – Gerade kleinster Quadrate

Formeln:

$$y = a_1 \cdot x + a_0$$

bzw.

$$x = b_1 \cdot y + b_0$$

Die Werte für a_1 und a_0 bzw. b_1 und b_0 kann man berechnen (vgl. Formelsammlung).

(Empirischer) Korrelationskoeffizient r

Zahl zwischen 1 (streng linear ansteigende Punktreihe im Streudiagramm) und -1 (streng linear abfallende Punktreihe im Streudiagramm). Liegt der Wert von r nahe Null, ist im Streudiagramm keine Gerade erkennbar.

(Empirische) Streuung bzw. Standardabweichung

Man berechnet für die einzelnen Ergebnisse e_i das Quadrat ihrer Differenz zum Mittelwert m (das sind die Quadrate der Abweichung für jedes einzelne Ereignis e_i), multipliziert diese Abweichungsgrößen mit der jeweiligen zugehörigen absoluten Häufigkeit (wie oft trat jede Abweichung auf), summiert alle berechneten Produkte und teilt sie abschließend durch die Anzahl n der beobachteten Ereignisse.

Wahrscheinlichkeit

Wahrscheinlichkeit P(E)

Zahl zwischen 0 und 1, deren Größe ein Maß für das Eintreten des Ereignisses E darstellt. Beträgt die Wahrscheinlichkeit $P(E) = 0$, wird das Ereignis E überhaupt nicht eintreten. Bei $P(E) = 1$ tritt das Ereignis E bei jedem Versuch ein.

Laplace-Experiment und Wahrscheinlichkeitsberechnung

Zufallsexperiment mit endlicher Ergebnismenge, bei dem kein Elementarereignis gegenüber den anderen Elementarereignissen in Bezug auf das Eintreten oder Nichteintreten bei einer Versuchsdurchführung bevorzugt wird. Das heißt Chancengleichheit für alle Elementarereignisse wie zum Beispiel beim Lotto – alle Zahlen existieren nur einmal (damit handelt es sich um Elementarereignisse) und sie haben alle die gleiche Chance, gezogen zu werden (keine Bevorzugungen). Bei n Elementarereignissen gilt für die Wahrscheinlichkeit eines einzelnen Elementarereignisses E_i:

$$P(E_i) = \frac{1}{n} = \frac{1}{Anzahl\ der\ möglichen\ Ereignisse}$$

Für alle Ereignisse gilt $0 < P(E_i) < 1$. Für das sichere Ereignis Ω ist $P(\Omega)$=1.

Für beliebige Ereignisse E berechnet sich die Wahrscheinlichkeit über:

$$P(E) = \frac{Anzahl\ der\ Elemente\ in\ E}{Anzahl\ der\ möglichen\ Ereignisse}$$
$$P(E) = \frac{Anzahl\ der\ für\ das\ Ereignis\ E\ günstigen\ Ergebnisse}{Anzahl\ der\ möglichen\ Ereignisse}$$

Bei einem Würfel beträgt die Wahrscheinlichkeit für das Würfeln einer bestimmten Zahl zwischen 1 und 6 somit immer 1/6.

Wahrscheinlichkeit zweier Ereignisse (ODER-Verknüpfung)

Sind zwei Ereignisse E und F unvereinbar, d. h. die Durchschnittsmenge aus E und F ist leer ($E \cap F = \emptyset$), so gilt:

$$P(E \cup F) = P(E) + P(F)$$

Unvereinbar sind zum Beispiel das Messen der Länge 2 m und das Messen der Länge 4 m. Will man nun die Wahrscheinlichkeit des Auftretens beider Längen wissen (Vereinigungsmenge von E und F, dargestellt durch $E \cup F$), so addieren sich die Einzelwahrscheinlichkeiten.

Für zwei beliebige Ereignisse E und F gilt:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

Gibt es Teile der Mengen E und F, die einander überdecken oder überlappen, so dürfen diese Teile für die Gesamtwahrscheinlichkeit nicht doppelt gezählt werden – das würde passieren, wenn man nur die Formel wie bei den unvereinbaren Ereignissen verwendet. Statt dessen ergibt sich die Gesamtmenge aus der Summe der beiden Teilmengen einmal reduziert um den überlappenden Bereich.

Versuchen wir es mit der Verdeutlichung an einem Beispiel. Gefragt ist die Wahrscheinlichkeit von Längenmessungen unter 3 m und Längenmessungen über 2 m. Also wäre das eine Ereignis die Länge unter 3 m und das andere Ereignis die Länge über 2 m – soweit sicher klar. Addieren wir jetzt die Wahrscheinlichkeiten der beiden Ereignisse, wird der Bereich über 2 m bis unter 3 m doppelt gezählt, da er in beiden Ereignissen enthalten ist. Also muss die Wahrscheinlichkeit des Auftretens von Längen in diesem Bereich einmal vom Ergebnis subtrahiert werden. Damit finden sich nun alle in Frage kommenden Längen – in unserem Beispiel die von Null bis Unendlich – einfach im Ergebnis wieder.

Wahrscheinlichkeit von Ereignis und Gegenereignis

Ist $\bar{E}$ das Gegenereignis zum Ereignis E (z. B. Messungen von 2 m als Ereignis E und Messungen ungleich 2 m als Gegenereignis $\bar{E}$), so gilt:

$$P(E) = 1 - P(\bar{E})$$

bzw.

$$P(\bar{E}) = 1 - P(E)$$

Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit des Ereignisses E unter der Bedingung F berechnet sich nach:

$$P_F(E) = \frac{P(F \cap E)}{P(F)} = \frac{Wahrscheinlichkeit\ des\ Eintretens\ von\ E\ und\ F}{Wahrscheinlichkeit\ des\ Eintretgens\ von\ F}$$

Suchen wir dafür ein Beispiel aus der Kraftfahrzeugtechnik: Die Wahrscheinlichkeit für den Umstand, dass ein Auto beim Anlassen nicht anspringt (Ereignis E), ist allgemein sehr gering – sagen wir 0,001 (jeder tausendste Anlassversuch bleibt ohne Erfolg). Liegt allerdings ein Defekt in der elektrischen Anlage des Autos vor (Bedingung F), tritt das Versagen beim Anlassen viel häufiger – nämlich mit der Wahrscheinlichkeit 0,5 – auf (jetzt springt also jedes zweite Auto nicht mehr an). Nehmen wir mal an, ein Defekt in der Fahrzeugelektrik hätte die Wahrscheinlichkeit 0,0125 (jedes achtzigste Auto). So ergibt sich für die Wahrscheinlichkeit für das Nichtanspringen des Motors unter der Bedingung einer gestörten Fahrzeugelektrik auch über die Formel:

$$P_F(E) = \frac{P(F \cap E)}{P(F)} = \frac{0{,}0125 \cdot 0{,}5}{0{,}0125} = 0{,}5$$

Unabhängigkeit von Ereignissen

Sind A und B zwei Ereignisse mit $P(A) > 0$ und $P(B) > 0$, so heißt B von A unabhängig, wenn gilt:

$$P_A(B) = P(B)$$

Andernfalls heißt B von A abhängig. Werden zum Beispiel Kugeln aus einer Urne entnommen und immer wieder zurückgelegt, so spielt es für die nachfolgenden Ziehungen überhaupt keine Rolle, welche Kugeln zuvor gezogen wurden. Hängt B nicht von A ab, so erkennt man an diesem Beispiel, dass dann auch A von B unabhängig ist.

Zwei Ereignisse A und B heißen auch genau dann unabhängig, wenn gilt:

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

Wer schon mit dem Baumdiagramm gearbeitet hat (vergleiche Abschnitt über mehrstufige Zufallsexperimente), dem ist bekannt, dass zum Beispiel für die Berechnung der Wahrscheinlichkeit des Ziehens einer roten und einer gelben Kugel aus einer Urne mit Zurücklegen nur die am passenden Pfad im Baumdiagramm eingetragenen Wahrscheinlichkeiten miteinander zu multiplizieren sind.

Haben wir es mit noch mehr Ereignissen A_1, A_2, A_3, … A_n zu tun, verwandelt sich die Formel – sofern diese Ereignisse alle unabhängig sind – um in:

$$P(A_1 \cap A_2 \cap A_3 \dots \cap A_n) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot \dots \cdot P(A_n)$$

Wie bei den mehrstufigen Zufallsexperimenten ausführlich erläutert, steckt hier nichts als das Baumdiagramm für Urnenexperimente mit Zurücklegen dahinter. Die Wahrscheinlichkeit für das Ziehen einer bestimmten Kugel oder Kugelfarbe ist zeitlich unveränderlich und damit von anderen gezogenen Kugeln unabhängig, da die Kugelanzahlen in der Urne vor jedem Ziehen wieder der Ausgangssituation entsprechen.

Multiplikationsregel

Sind E und F zwei Ereignisse mit $P(E) > 0$, so gilt:

$$P(E \cap F) = P(E) \cdot P_E(F)$$

Hat zum Beispiel ein Hersteller mit seinen Produkten einen Marktanteil von 30 % und 5 % seiner Produkte werden in der Garantiezeit reklamiert, so beträgt die Wahrscheinlichkeit, dass ein Produkt von diesem Hersteller kommt und reklamiert wird:

$$P(Hersteller \cap Reklamation) = P(Hersteller) \cdot P_{Hersteller}(Reklamation)$$
$$P(Hersteller \cap Reklamation) = 0{,}3 \cdot 0{,}05 = 0{,}015$$

Totale Wahrscheinlichkeit

Es seien E_1, E_2, E_3, … E_n Ereignisse mit den folgenden Eigenschaften:

1. $P(E_i) > 0$ für alle $i = 1 \dots n$
2. Die Ereignisse sind paarweise unvereinbar, d. h. $E_i \cap E_j = \emptyset$ für $i \neq j$.
3. $E_1 \cap E_2 \cap E_3 \cap \dots \cap E_n = \Omega$

Dann gilt für jedes beliebige Ereignis F:

$$P(F) = P(E_1) \cdot P_{E1}(F) + P(E_2) \cdot P_{E2}(F) + P(E_3) \cdot P_{E3}(F) + \cdots + P(E_n) \cdot P_{En}(F)$$

Gibt es zum Beispiel für ein bestimmtes Produkt einen Hersteller H1 mit 30 % Marktanteil, einen Hersteller H2 mit 45 % Marktanteil und einen Hersteller H3 mit 25 % Marktanteil (Marktanteile alle größer Null und unvereinbar sowie insgesamt 100 % – entspricht den Bedingungen 1 bis 3) und die relativen Reklamationshäufigkeiten in der Garantiezeit betragen:

Hersteller H1	5 %
Hersteller H2	7 %
Hersteller H3	3 %,

so ergibt sich für die Wahrscheinlichkeit, dass ein solches Produkt in der Garantiezeit reklamiert wird:

$$P(Rekla) = P(H1) \cdot P_{H1}(Rekla) + P(H2) \cdot P_{H2}(Rekla) + P(H3) \cdot P_{H3}(Rekla)$$
$$P(Rekla) = 0{,}3 \cdot 0{,}05 + 0{,}45 \cdot 0{,}07 + 0{,}25 \cdot 0{,}03 = 0{,}054$$

Pfad-Regel (UND-Verknüpfung)

In mehrstufigen Zufallsexperimenten ist die Wahrscheinlichkeit jedes Elementarereignisses (jedes Pfades im Baumdiagramm) gleich dem Produkt der Wahrscheinlichkeiten an jeder Verbindungslinie = Pfad, der diesem Elementarereignis entspricht. Formelmäßig ergibt sich:

$$P(A_1 \cap A_2 \cap A_3 \cap \ldots \cap A_n) = P(A_1) \cdot P_{A1}(A_2) \cdot P_{A1\cap A2}(A_3) \cdot \ldots \cdot P_{A1\cap A2\cap\ldots\cap An-1}(A_n)$$

Diese Formel sieht eher grausam aus, hat aber als Hintergrund nichts weiter als die schon beschriebene Verfahrensweise im Baumdiagramm. Sucht man die Wahrscheinlichkeit für einen bestimmten Endzustand im Diagramm, so multipliziert man die Wahrscheinlichkeit an der ersten Verbindungslinie des passenden Pfades $P(A_1)$ mit der Wahrscheinlichkeit an der nächsten Verbindungslinie auf diesem Pfad. Und die gibt nichts anderes als die Wahrscheinlichkeit für das Erreichen eines Zustandes A_2 unter der Bedingung, dass man vorher den Zustand A_1 erreicht hat, an. Damit wären zumindest die ersten beiden Faktoren rechts vom Gleichheitszeichen der Formel erklärt. Hat das Baumdiagramm nun noch eine dritte Stufe, wird das bisherige Ergebnis mit der Wahrscheinlichkeit an der nächsten passenden Verbindungslinie multipliziert. Das setzt aber voraus, dass wir zuvor die Zustände A_1 und A_2 durchlaufen hatten – vergleiche dritter Faktor in der Formel.

Satz von Bayes

Bilden die Ereignisse E_1, E_2, E_3, … E_n eine Zerlegung, d. h. $E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_n = \Omega$ und ist F ein beliebiges Ereignis mit $P(F) > 0$, so gilt für jedes $i = 1 \ldots n$:

$$P_F(E_i) = \frac{P(E_i) \cdot P_{Ei}(F)}{P(E_1) \cdot P_{E1}(F) + P(E_2) \cdot P_{E2}(F) + P(E_3) \cdot P_{E3}(F) + \cdots + P(E_n) \cdot P_{En}(F)}$$

Damit besteht eine Möglichkeit, eine bedingte Wahrscheinlichkeit zu berechnen, bei der das Ereignis und die dafür zugrunde liegende Bedingung gegenüber den gegebenen Größen vertauscht sind. Man berechnet in der obigen Formel die Größe $P_F(E_i)$ mit Hilfe der Zahlen für $P_{Ei}(F)$. Die Größen E_i und F wechseln ihre Position.

Demonstrieren wir das an einem Beispiel: In einer Schulklasse sind zwei Drittel der Schüler Jungen und ein Drittel Mädchen. 40 % der Jungen sind größer als 1,80 m. Bei den Mädchen trifft das nur auf 5 % zu. Mit welcher Wahrscheinlichkeit ist ein zufällig ausgewähltes Klassenmitglied größer 1,80 m ein Mädchen?

Ausgewählter Schüler ist ein Junge –	Ereignis J	$P(J) = 0{,}67$
Ausgewählter Schüler ist ein Mädchen –	Ereignis M	$P(M) = 0{,}33$
Ausgewählter Schüler ist größer als 1,80 m –	Ereignis A	$P_J(A) = 0{,}4$
		$P_M(A) = 0{,}05$

Gesucht ist jetzt die Größe $P_A(M)$ – die Wahrscheinlichkeit für das Auswählen eines Mädchens, wenn die Schülergröße über 1,80 m liegt.

$$P_A(M) = \frac{P(M) \cdot P_M(A)}{P(J) \cdot P_J(A) + P(M) \cdot P_M(A)} = \frac{0{,}33 \cdot 0{,}05}{0{,}67 \cdot 0{,}4 + 0{,}33 \cdot 0{,}05} = 0{,}059$$

Geometrische Wahrscheinlichkeit

Ist die Ergebnismenge Ω ein endliches Flächenstück in der Ebene und F ein Ereignisfeld von Teilflächen von Ω, dann nennt man die Funktion P, die jedem Ereignis A die Wahrscheinlichkeit

$$P(A) = \frac{Flächeninhalt\ von\ A}{Flächeninhalt\ von\ \Omega}$$

zuordnet, die geometrische Wahrscheinlichkeitsfunktion.

Ein Beispiel soll das Verständnis erleichtern: Zwei Personen verabreden sich an einem Treffpunkt zwischen 10.00 Uhr und 10.30 Uhr. Wie groß ist die Wahrscheinlichkeit, dass keine von beiden länger als 10 Minuten warten muss? Dieses Problem lässt sich mit normalen Methoden nicht mehr lösen, da sich – wenn man die Eintreffzeiten nicht auf ganze Minuten beschränkt – unendlich viele Wertepaare für das Eintreffen der ersten und das Eintreffen der zweiten Person bilden lassen. Die erste Dimension des Zufallsexperiments ist die Eintreffzeit der ersten Person, die zweite Dimension die Eintreffzeit der zweiten Person – so kommt man zur Darstellung durch Flächen, die ja bekanntlich zweidimensional sind. Existiert gar noch eine dritte Dimension (zum Beispiel Eintreffen einer dritten Person), entwickelt sich die Lösung des Problems zu einer räumlichen Angelegenheit mit drei Dimensionen.

Ist A_1, A_2, A_3, … A_n, … eine unendliche Folge von paarweise unvereinbaren Ereignissen und A die Vereinigungsmenge dieser Ereignisse, so gilt für die geometrische Wahrscheinlichkeitsfunktion P:

$$P(A) = \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{\infty} P(A_i)$$

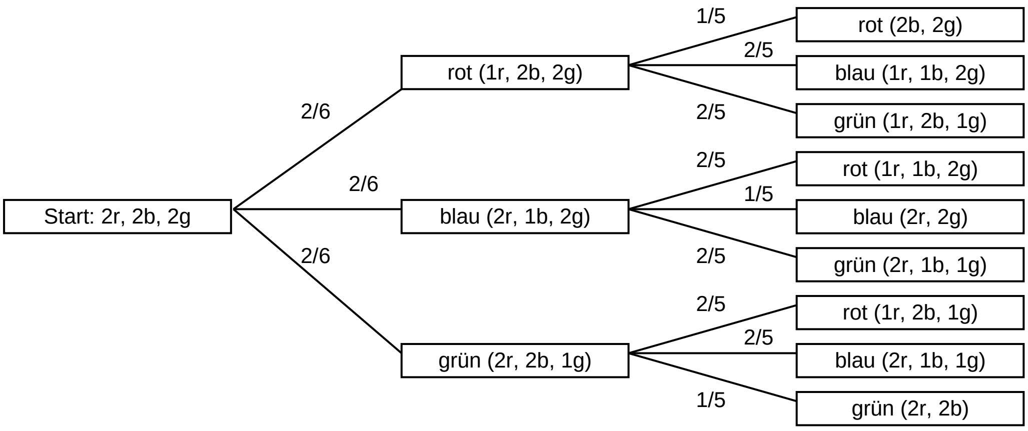
Monte-Carlo-Methode

Verfahren zur Berechnung zum Beispiel von Flächeninhalten beliebiger geometrischer Körper. Die zu untersuchende Fläche wird in eine sie umgebende Fläche bekannten Inhalts eingezeichnet bzw. eingeschrieben. Nun werden große Anzahlen von Zufallszahlen erzeugt, die Punkte in der umgebenden Fläche bezeichnen. Nach hinreichend vielen Punkten kann gesagt werden, mit welcher Wahrscheinlichkeit ein Punkt in der zu untersuchenden Fläche liegt. Die Multiplikation dieser Wahrscheinlichkeit mit dem bekannten Inhalt der umgebenden Fläche erzeugt einen Schätzwert für den Flächeninhalt der zu untersuchende Fläche.

Mehrstufige Zufallsexperimente

Mehrfaches Ziehen ohne Zurücklegen – Baumdiagramm

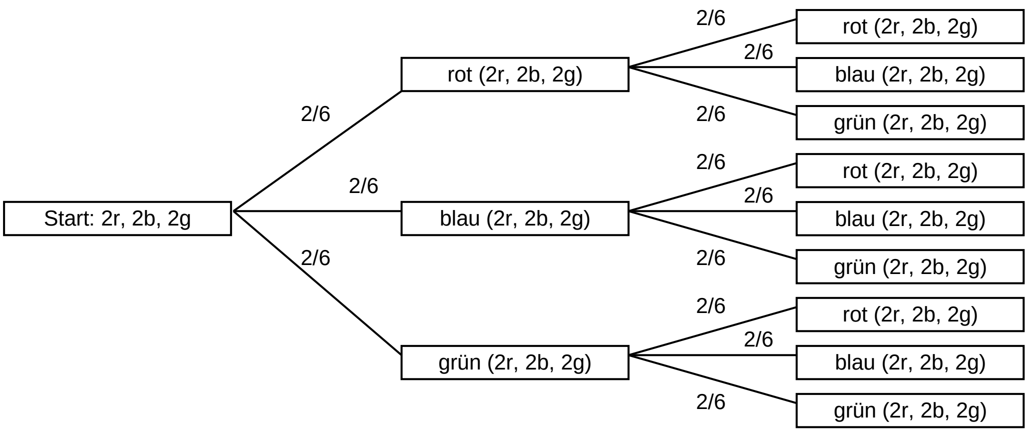
Einfachstes Hilfsmittel zum Analysieren dieses Vorgangs ist ein Baumdiagramm. Man zeichnet dieses von links nach rechts oder von oben nach unten und hält für jeden Zustand den jeweiligen Urneninhalt fest. Zum Beispiel aus einer Urne mit 2 roten, 2 blauen und 2 grünen Kugeln (alle Kugeln gleich groß) werden nacheinander 2 Kugeln ohne Zurücklegen gezogen. In den Feldern des Baumdiagramms steht die jeweils gezogene Kugelfarbe und die noch in der Urne befindlichen Kugelanzahlen:



An die Verbindungslinien = Pfade zwischen den einzelnen Zuständen schreibt man die Wahrscheinlichkeit für das Eintreten des jeweils nachfolgenden Urnenzustands. Die Gesamtwahrscheinlichkeit für das Eintreten eines bestimmten Endzustands ergibt sich aus dem Produkt der vom Start bis zu diesem Endzustand vermerkten Einzelwahrscheinlichkeiten.

Mehrfaches Ziehen mit Zurücklegen – Baumdiagramm

Das Baumdiagramm zeigt in allen Feldern im Gegensatz zum Beispiel ohne Zurücklegen identische Urneninhalte an. Alle Wahrscheinlichkeiten für das Eintreten eines bestimmten nachfolgenden Urneninhalts sind gleich groß, da immer wieder die gleiche Kugelanzahl in der Urne liegt. Das heißt im Beispiel sind die Wahrscheinlichkeiten für das Ziehen einer roten, einer blauen oder einer grünen Kugel immer 2/6.



Summenregel

Gemäß der Wahrscheinlichkeit zweier unvereinbarer Ereignisse ist die Wahrscheinlichkeit eines Ereignisses

$F = \{\omega_1;\ \omega_2;\ \omega_3; \dots;\ \omega_n\}$

gleich der Summe der Wahrscheinlichkeiten der Elementarereignisse

$E_i = \{\omega_i\}$,

die F nach sich ziehen:

$F = E_1 \cup E_2 \cup E_3 \cup \dots \cup E_n$

$P(F) = P(E_1) + P(E_2) + P(E_3) + \cdots + P(E_n)$

Zum Beispiel werden aus einer Urne Kugeln mit aufgedruckten Zahlen mit Zurücklegen gezogen. Das Ereignis F sei das Ziehen einer Kugel mit einer aufgedruckten geraden Zahl. Bei 10 enthaltenen Kugeln mit den Zahlen 1 bis 10 berechnet sich die Wahrscheinlichkeit für das Ereignis F aus der Summe der Einzelwahrscheinlichkeiten für das Ziehen der enthaltenen geraden Zahlen 2, 4, 6, 8 und 10. Da die Wahrscheinlichkeit für das Ziehen einer bestimmten Zahl bei insgesamt 10 Möglichkeiten gleich 0,1 ist, ergibt sich für die Wahrscheinlichkeit für das Ziehen einer geraden Zahl

$Ziehen\ gerade\ Zahl = \{Ziehen\ 2;\ Ziehen\ 4;\ Ziehen\ 6;\ Ziehen\ 8;\ Ziehen\ 10\}$

$Ziehen\ gerade\ Zahl = Ziehen\ 2 \cup Ziehen\ 4 \cup Ziehen\ 6 \cup Ziehen\ 8 \cup Ziehen\ 10$

$P(Ziehen\ gerade\ Zahl) = P(Ziehen\ 2) + P(Ziehen\ 4) + P(Ziehen\ 6) + P(Ziehen\ 8) + P(Ziehen\ 10)$

$P(Ziehen\ gerade\ Zahl) = 0{,}1 + 0{,}1 + 0{,}1 + 0{,}1 + 0{,}1 = 0{,}5$

n-Tupel

Zieht man aus einer beliebigen Menge n Mal mit oder ohne Zurücklegen, ist das Ergebnis dieser Ziehung bzw. dieses Zufallsexperiments eine Folge von n Teilergebnissen, die man als n-Tupel bezeichnet. Jedes n-Tupel steht für genau einen Pfad durch das Baumdiagramm vom Start bis zu einem Zustand der letzten Stufe. Im obigen Beispiel sind die Ergebnisse 2-Tupel, bestehend aus zwei Kugelfarben.

Pfad-Regel

In mehrstufigen Zufallsexperimenten ist die Wahrscheinlichkeit jedes Elementarereignisses (jedes Pfades im Baumdiagramm) gleich dem Produkt der Wahrscheinlichkeiten an jeder Verbindungslinie = Pfad, der diesem Elementarereignis entspricht. Formelmäßig ergibt sich:

$P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n) = P(A_1) \cdot P_{A1}(A_2) \cdot P_{A1 \cap A2}(A_3) \cdot \dots \cdot P_{A1 \cap A2 \cap \dots \cap An-1}(A_n)$

Diese Formel sieht eher grausam aus, hat aber als Hintergrund nichts weiter als die schon beschriebene Verfahrensweise im Baumdiagramm. Sucht man die Wahrscheinlichkeit für einen bestimmten Endzustand im Diagramm, so multipliziert man die Wahrscheinlichkeit an der ersten Verbindungslinie des passenden Pfades $P(A_1)$ mit der Wahrscheinlichkeit an der nächsten Verbindungslinie auf diesem Pfad. Und die gibt nichts anderes als die Wahrscheinlichkeit für das Erreichen eines Zustandes A_2 unter der Bedingung, dass man vorher den Zustand A_1 erreicht hat, an. Damit wären zumindest die ersten beiden Faktoren rechts vom Gleichheitszeichen der Formel erklärt. Hat das Baumdiagramm nun noch eine dritte Stufe, wird das bisherige Ergebnis mit der Wahrscheinlichkeit an der nächsten passenden Verbindungslinie multipliziert. Das setzt aber voraus, dass wir zuvor die Zustände A_1 und A_2 durchlaufen hatten – vergleiche dritter Faktor in der Formel.

Anzahl der Möglichkeiten für die geordnete Auswahl von k Elementen aus einer Menge mit n Elementen ohne Zurücklegen

Die Anzahl geordneter Teilmengen mit je k Elementen aus n möglichen beträgt

$$\frac{n!}{(n-k)!}$$

wobei unterschieden wird, ob zum Beispiel eine bestimmte Kugelfarbe beim Ziehen aus einer Urne als erste, zweite, ... oder n-te gezogen wird. Zieht man vier Kugeln aus zehn, ergeben sich dadurch

$$10 \cdot 9 \cdot 8 \cdot 7 = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{10!}{6!} = \frac{n!}{(n-k)!}$$

Möglichkeiten. Die erste Kugel wird noch aus 10 verschiedenen Kugeln gewählt, die zweite nur noch aus 9, die dritte nur noch aus 8 und die vierte letztlich nur noch aus 7 Kugeln.

Anzahl der Möglichkeiten für die ungeordnete Auswahl von k Elementen aus einer Menge mit n Elementen ohne Zurücklegen

Das bisherige Ergebnis für die geordnete Auswahl von Elementen enthält Kugelfolgen, die durch Umsortieren ineinander überführbar sind. Interessiert aber nur die Anzahl bestimmter Kugelarten in einer Ziehung, nicht aber die dabei aufgetretene Reihenfolge, sind die betreffenden Ziehungsergebnisse identisch – die bisher gefundene Anzahl an geordneten Möglichkeiten ist also zu groß und zwar genau um den Faktor $k!$. Damit ergibt sich für die Anzahl ungeordneter Teilmengen mit je k Elementen aus n möglichen

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

Andere Formulierung: Ist M eine Menge mit n Elementen, dann gibt es

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

Teilmengen von M mit k Elementen. Der Ausdruck

$$\binom{n}{k}$$

heißt auch Binomialkoeffizient von n und k – gesprochen n über k.

Zufallsgrößen

Zufallsgröße

Einfach formuliert: Jedem Ergebnis e oder ω der Ergebnismenge Ω einer Beobachtungsreihe wird eine reelle Zahl $X(e)$ oder $X(\Omega)$ zugeordnet. Dabei hat diese Zuordnung nichts zufälliges an sich – sie ist eindeutig, aber häufig nicht umkehrbar eindeutig. Zufällig ist nur – so wie schon bisher – das Auftreten bestimmter Ergebnisse e oder ω im Verlaufe einer Beobachtungsreihe.

Zufallsgrößen werden definiert, um die Ergebnisse von Beobachtungsreihen mit mathematischen Methoden weiter verarbeiten zu können. Da es sich bei den Zufallsgrößen um reelle Zahlen handelt, kommen auch alle mathematischen Operationen für reellwertige Funktionen in Frage.

Nähern wir uns der Thematik ganz sacht: Eine Münze wird zweimal geworfen. Die Anzahl der Wappen wird als Zufallsgröße festgehalten. Welche Ergebnisse sind möglich und wie äußern die sich als Zufallsgröße?

1. Wurf: Wappen,	2. Wurf: Wappen,	Zufallsgröße $X = 2$
1. Wurf: Wappen,	2. Wurf: Zahl,	Zufallsgröße $X = 1$
1. Wurf: Zahl,	2. Wurf: Wappen,	Zufallsgröße $X = 1$
1. Wurf: Zahl,	2. Wurf: Zahl,	Zufallsgröße $X = 0$

So einfach – aber trotzdem schon hier nicht umkehrbar eindeutig. Zwar ist die Zuordnungsvorschrift für die Zufallsgröße festgelegt und eindeutig, aber ich kann aus der Zufallsgröße nur bedingt auf die Ergebnisse in der Beobachtungsreihe schließen. Für $X = 1$ kann man nicht sagen, ob erst das Wappen oder erst die Zahl oben lag – beides ist möglich.

Diskrete Zufallsgröße

Eine diskrete Zufallsgröße nimmt nur endlich viele oder abzählbar unendlich viele Werte an. Die Wertemenge ist eine endliche oder eine abzählbar unendliche Menge.

Das Beispiel unseres Münzwurfs liefert also schon eine erste diskrete Zufallsgröße, weil nur die Zahlen von 0 bis 2 möglich sind. Abzählbar unendlich wäre bei der Serienherstellung eines Produktes das Zählen der Produkte bis zum ersten fehlerhaften. Hier entstehen unter Umständen sehr große Zahlen – deshalb abzählbar *unendlich*.

Verteilungstabelle

Sind für die Zufallsgröße X nur endlich viele Werte möglich, kann man diese in eine Verteilungstabelle (erste Zeile) eintragen und ihnen die entsprechenden Wahrscheinlichkeiten (in der zweiten Zeile) zuordnen.

Für den Münzwurf ergibt sich:

Mögliche Werte x	0	1	2
$P(X = x)$	$\frac{1}{4} = 0{,}25$	$\frac{2}{4} = 0{,}5$	$\frac{1}{4} = 0{,}25$

In jeder Verteilungstabelle muss die Summe aller Wahrscheinlichkeiten in der zweiten Zeile Eins ergeben, sonst hat sich ein Fehler eingeschlichen. Grafisch lässt sich eine Verteilungstabelle mit Hilfe eines Histogramms = Säulendiagramm oder eines Stabdiagramms veranschaulichen.

Verteilungsfunktion

Als Verteilungsfunktion einer Zufallsgröße X bezeichnet man die für alle reellen Zahlen x definierte Funktion:

$F_X(x) = P(X < x)$

$F_X(x)$ gibt also die Wahrscheinlichkeit an, mit der die Zufallsgröße X kleiner dem (reellen) Wert x ist. Dabei ist x die Variable, in deren Abhängigkeit $F_X(x)$ berechnet und somit auch in einem Diagramm gezeichnet wird. Durch die x-Achse wird x dargestellt und auf der y-Achse $F_X(x)$ abgetragen.

Bemühen wir wieder den sehr überschaubaren Münzwurf:

Mögliche Werte x	≤ 0	$0 < x \leq 1$	$1 < x \leq 2$	$2 < x$
$P(X < x)$	0	$\frac{1}{4} = 0{,}25$	$\frac{1}{4} + \frac{2}{4} = 0{,}75$	$\frac{1}{4} + \frac{2}{4} + \frac{1}{4} = 1$

Jetzt muss man genau aufpassen, was durch welches Feld dargestellt wird:

In der ersten Zahlenspalte geht es um den Wertebereich $x \leq 0$. Die Wahrscheinlichkeit dafür, dass die Anzahl der geworfenen Wappen kleiner Null ist – nichts anderes heißt ja $P(X < x)$ in diesem Fall – ist natürlich gleich Null. Weniger als keine Wappen geht nicht.

In der zweiten Zahlenspalte wird der Wertebereich $0 < x \leq 1$ untersucht. Die Wahrscheinlichkeit dafür, dass die Anzahl der geworfenen Wappen kleiner Eins ist, beträgt 0,25, da hier nur der Fall keine Wappen in Frage kommt.

In der dritten Zahlenspalte wird der Wertebereich $1 < x \leq 2$ betrachtet. Die Wahrscheinlichkeit dafür, dass die Anzahl der geworfenen Wappen kleiner Zwei ist, beträgt immerhin schon 0,75, da hier jetzt die Fälle keine Wappen und ein Wappen zutreffen.

Und in der vierten Zahlenspalte wird der Wertebereich $2 < x$ analysiert. Die Wahrscheinlichkeit dafür, dass die Anzahl der geworfenen Wappen kleiner Unendlich ist – die Größe x bleibt ja durch $2 < x$ nach oben unbegrenzt, beträgt 1, da hier alle möglichen Zufallsgrößen beim Münzwurf eingeschlossen werden.

Verallgemeinert man dieses Beispiel, hat jede Verteilungsfunktion $F_X(x) = P(X < x)$ für hinreichend kleine x den Wert Null – wir befinden uns auf der x-Achse unterhalb des Bereiches, wo die Wahrscheinlichkeit(en) für das Eintreten bestimmter Ereignisse größer Null ist bzw. sind. Genau in diesem Bereich der x-Achse, wo also tatsächlich Wahrscheinlichkeiten größer Null vorliegen, steigt die Verteilungsfunktion von Null beginnend an. Sobald der Wert für x diesen Bereich vollständig passiert hat – alle Wahrscheinlichkeitswerte größer Null liegen jetzt unterhalb vom aktuellen x-Wert – erreicht die Verteilungsfunktion den Wert Eins. Die Summe aller möglichen Wahrscheinlichkeiten ist ja bekanntlich gleich Eins und an diesem Sachverhalt ändert sich auch nichts mehr, wenn x gegen Unendlich strebt.

Somit sehen alle Verteilungsfunktionen prinzipiell gleich aus: „Links“ verlaufen sie auf dem Wert Null, irgendwann steigen sie mehr oder weniger steil auf den Wert Eins an und den verlassen sie nach „rechts“ auch nicht mehr. Diese grundsätzliche Ähnlichkeit aller Verteilungsfunktionen eröffnet die Möglichkeit, allgemeingültige Zahlentabellen für bestimmte Typen von Verteilungsfunktionen abzudrucken. Würde jede Verteilungsfunktion vollkommen anders aussehen, ginge das nicht.

Diskret gleichverteilte Zufallsgröße

Eine Zufallsgröße X mit der Wertemenge $\{x_1, x_2, x_3, \dots, x_n\}$ heißt diskret gleichverteilt über die möglichen Werte $x_1, x_2, x_3, \dots, x_n$, wenn gilt:

$$P(X = x_1) = P(X = x_2) = P(X = x_3) = \cdots = P(X = x_n) = \frac{1}{n}$$

2-Punkt-verteilte Zufallsgröße

Eine Zufallsgröße X mit der Wertemenge $\{x_1, x_2\}$ heißt 2-Punkt-verteilt über die Werte x_1, x_2 mit dem Parameter p $(0 < p < 1)$, wenn gilt:

$$P(X = x_1) = 1 - p$$

sowie

$$P(X = x_2) = p$$

Sind $x_1 = 0$ und $x_2 = 1$, so handelt es sich auch um eine 0-1-Verteilung.

1-Punkt-verteilte Zufallsgröße

Alle Ergebnisse einer Beobachtungsreihe werden auf einen festen Punkt a der Zahlengerade abgebildet. Egal, was passiert, der Wert der Zufallsgröße beträgt also immer nur a. Wenn keine anderen Werte möglich sind, gilt dadurch automatisch:

$$P(X = a) = 1$$

Geometrisch verteilte Zufallsgröße

Eine Zufallsgröße heißt geometrisch verteilt mit dem Parameter p mit $(0 < p < 1)$, wenn sie folgende Bedingungen erfüllt: Die Wertemenge ist die Menge aller natürlichen Zahlen und es gilt für jede natürliche Zahl l:

$$P(X = l) = (1 - p) \cdot p^l$$

Die Wahrscheinlichkeiten $P(X = l)$ bilden für $l = 1, 2, 3, 4, \ldots$ eine geometrische Folge. Diese zeichnet sich durch folgende Eigenschaften aus: Der Quotient q zweier aufeinander folgenden Glieder ist konstant und a_n kann aus a_{n-1} und a_{n+1} durch Berechnung des <u>geometrischen</u> Mittelwerts berechnet werden:

$$q = \frac{a_{n+1}}{a_n}$$

Das gilt für alle n.

Zum Beispiel: 2; 4; 8; 16; 32; 64; ...

Die rekursive Berechnung einer geometrischen Zahlenfolge erfolgt über die Formel:

$$a_{n+1} = a_n \cdot q$$

mit Angabe von a_1.

Die explizite Berechnung einer geometrischen Zahlenfolge erfolgt über die Formel:

$$a_n = a_1 + q^{(n-1)}$$

oder

$a_n = a_0 + q^n$

Geometrischer Mittelwert:

$a_n = \sqrt{a_{n-1} \cdot a_{n+1}}$

Erwartungswert

Für eine Zufallsgröße X mit der Wertemenge $\{x_1, x_2, x_3, \dots, x_n\}$ berechnet sich der zugehörige Erwartungswert $E(X)$ über die Formel:

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

Die Formel hat eine große Ähnlichkeit mit der Berechnung des empirischen Mittelwertes, bei dem die festgestellten numerischen Ergebnisse einer Beobachtungsreihe mit ihrer jeweiligen relativen Häufigkeit multipliziert werden und anschließend ebenfalls die Summe über alle so entstandenen Ausdrücke berechnet wird. Somit ist die häufige Bezeichnung des Erwartungswertes als Mittelwert erstens sachlich korrekt und zweitens nachvollziehbar.

Würfeln wir wieder: Bei einem Würfel beträgt die Wahrscheinlichkeit, dass eine bestimmte Zahl gewürfelt wird, immer 1/6. Als Erwartungswert ergibt sich entsprechend:

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3{,}5$$

Würde man die x-Achse von 1 bis 6 als frei beweglichen Stab betrachten, müsste man diesen an der Stelle 3,5 zum Balancieren auf den Finger legen, da dort der Schwerpunkt ist.

Für die Berechnung eines Erwartungswertes gilt weiterhin:

$$E(X + Y) = E(X) + E(Y)$$
$$E(a \cdot X) = a \cdot E(X)$$
$$E(X + a) = E(X) + a$$

Varianz

Für eine Zufallsgröße X mit der Wertemenge $\{x_1, x_2, x_3, \dots, x_n\}$ berechnet sich die Varianz $V(X)$ über die Formel:

$$V(X) = (x_1 - E(X))^2 \cdot P(X = x_1) + (x_2 - E(X))^2 \cdot P(X = x_2) + (x_3 - E(X))^2 \cdot P(X = x_3) + \cdots + (x_n - E(X))^2 \cdot P(X = x_n)$$

Zunächst ist die Varianz ein Maß für die Abweichungen der einzelnen Ergebnisse einer Beobachtungsreihe vom Erwartungswert = Mittelwert der Zufallsgröße X. Je größer die Differenzen in den Klammern sind, desto größer werden natürlich auch die Quadrate davon.

Welchen Sinn hat nun aber die Quadratbildung? Könnte man auf sie verzichten? Untersuchen wir den Sachverhalt am Würfel. Die Berechnung der Varianz führt für den Würfel mit Erwartungswert $E(X) = 3{,}5$ auf die Formel:

$$V(X) = (1 - 3{,}5)^2 \cdot P(1) + (2 - 3{,}5)^2 \cdot P(2) + (3 - 3{,}5)^2 \cdot P(3) + (4 - 3{,}5)^2 \cdot P(4) + (5 - 3{,}5)^2 \cdot P(5) + (6 - 3{,}5)^2 \cdot P(6)$$

$$V(X) = (1-3{,}5)^2 \cdot \frac{1}{6} + (2-3{,}5)^2 \cdot \frac{1}{6} + (3-3{,}5)^2 \cdot \frac{1}{6} + (4-3{,}5)^2 \cdot \frac{1}{6} + (5-3{,}5)^2 \cdot \frac{1}{6} + (6-3{,}5)^2 \cdot \frac{1}{6}$$
$$V(X) = 2{,}92$$

Würde man auf die Quadrate verzichten, ergibt sich aber als Ergebnis der Wert Null, weil sich die positiven und negativen Abweichungen der Würfelzahlen vom Erwartungswert 3,5 gegenseitig aufheben. Es soll sich aber letztlich jede Abweichung in der Größe der Varianz auswirken. Das geht nur, wenn jede Abweichung – egal ob positiv oder negativ – einen (positiven) Beitrag zur Summe leistet. Also niemals die Quadratbildung vergessen!

Für die Berechnung der Varianz gilt auch:

$$V(X) = E\left[\left(X - E(X)\right)^2\right] = E(X^2) - \left(E(X)\right)^2$$
$$V(a \cdot X) = a^2 \cdot V(X)$$
$$V(X + a) = V(X)$$

Die erste Formel lassen wir einfach im Raum stehen. Bei der zweiten Formel wirkt sich natürlich wieder das schon beschriebene Quadrat aus. Und bei der dritten Formel ist zu sagen, dass die Addition eines konstanten Wertes a auf die Größe der Abweichungen keinen Einfluss hat – die Abweichungen finden dann eben halt um den Wert $X + a$ statt um den Wert X statt. Setzt man den Ausdruck aus der dritten Formel in die erste Formel ein, heben sich alle Anteile mit a gegenseitig auf.

Standardabweichung

Für die Fälle, wo das Quadrat bei der Berechnung der Varianz zum Beispiel bei weitergehenden Untersuchungen doch stört, wurde die Standardabweichung σ der Zufallsgröße X definiert:

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

Hypergeometrisch verteilte Zufallsgröße

Sind N, $K \leq N$ und $n \leq N$ natürliche Zahlen und ist X eine Zufallsgröße mit der Wertemenge $\{0, 1, 2, 3, \dots, n\}$, dann heißt X hypergeometrisch verteilt, wenn für alle $k \epsilon \{0, 1, 2, 3, \dots, n\}$ gilt:

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

Der Erwartungswert $E(X)$ und die Varianz $V(X)$ einer hypergeometrisch verteilten Zufallsgröße X mit den Parametern N, K und n berechnen sich nach den Formeln:

$$E(X) = \frac{K \cdot n}{N}$$
$$V(X) = \frac{K \cdot n \cdot (N-K) \cdot (N-n)}{N^2 \cdot (N-1)}$$

Nehmen wir die praktische Verdeutlichung am Beispiel einer Urne mit 10 gleich großen Kugeln vor, von denen nur zwei Kugeln rot gefärbt sind. Auf gut Glück werden vier Kugeln ohne Zurücklegen entnommen. Gefragt ist die <u>Anzahl</u> der roten Kugeln in der entnommenen Probe.

Es liegt eine Gesamtheit von $N = 10$ Kugeln vor. Darin sind $K = 2$ rote Kugeln enthalten. Die Zufallsgröße X, die jeder Stichprobe vom Umfang $n = 4$ die Anzahl k der darin enthaltenen roten Kugeln zuordnet, hat die möglichen Werte $k\epsilon\{0, 1, 2\}$. Für das geordnete Auswählen von vier beliebigen Kugeln aus den 10 gibt es

$$10 \cdot 9 \cdot 8 \cdot 7 = 5040 = \frac{10!}{6!} = \frac{N!}{(N-n)!}$$

Möglichkeiten (beim Ziehen der ersten Kugel sind noch 10 Kugeln in der Kiste, die zweite Kugel kann nur noch aus 9 gezogen werden, die dritte nur noch aus 8 und die vierte nur noch aus 7). Diese Zahl berücksichtigt aber nicht, dass es unerheblich ist, in welcher Reihenfolge bestimmte Kugelkombinationen gezogen werden. Uns ist es egal, ob zum Beispiel eine in den vier Kugeln enthaltene rote Kugel als erste, als zweite, als dritte oder gar als vierte Kugel gezogen wurde. Die betreffenden vier Ergebnisse sind durch Umsortieren ineinander überführbar und für uns somit identisch, da wir uns nur für die Kugelanzahlen, nicht aber für die Kugelreihenfolgen interessieren. Der entsprechende Korrekturfaktor, der unser bisheriges Ergebnis auf die tatsächlich unterschiedlichen Fälle reduziert, lautet:

$$4 \cdot 3 \cdot 2 \cdot 1 = 24 = 4! = n!$$

Die bisherige Formel für das geordnete Ziehen von vier Kugeln ohne Zurücklegen verwandelt sich damit in:

$$\frac{10!}{4! \cdot 6!} = 210 = \frac{N!}{n! \cdot (N-n)!} = \binom{N}{n}$$

womit die bereits bekannte Formel für das ungeordnete Auswählen von n Elementen aus einer Menge mit N Elementen entstanden ist, die mit der Berechnung des entsprechenden Binomialkoeffizienten

$$\binom{N}{n}$$

übereinstimmt. Der Ausdruck

$$\binom{K}{k}$$

gibt an, wie viele Möglichkeiten es jeweils für das Ziehen von keiner, von einer oder von allen beiden roten Kugeln aus den vorhandenen $K = 2$ roten Kugeln in der Urne gibt – $k\epsilon\{0, 1, 2\}$.

Hier muss man jeweils das k einsetzen, für das man die Wahrscheinlichkeit berechnen will. Und hinter der Größe

$$\binom{N-K}{n-k}$$

verbirgt sich nichts anderes als die Anzahl an Möglichkeiten, $n - k$ nicht rote Kugeln aus der Urne zu ziehen. $N - K$ ist ja die Gesamtkugelanzahl verringert um die Anzahl roter Kugeln und $n - k$ stellt die Anzahl gezogener nicht roter Kugeln dar. Auch hier muss man je nach auszurechnender Wahrscheinlichkeit das k entsprechend einsetzen. Will man die Ausgangsformel nach diesen Erläuterungen in Bezug auf unser Kugelbeispiel verbal aufstellen, führt das zu dem Ausdruck:

$$P(X = k) = \frac{(Anzahl\ der\ Mögl.\ für\ das\ Ziehen\ k\ roter\ Kugeln) \cdot (Anzahl\ der\ Mögl.\ für\ das\ Ziehen\ n-k\ nicht\ roter\ Kugeln)}{(Anzahl\ der\ Mögl.\ für\ das\ Ziehen\ n\ beliebiger\ Kugeln\ aus\ N)}$$

Über dem Bruchstrich versteckt sich das Baumdiagramm. In der ersten Klammer steht die Anzahl an Möglichkeiten für das Ziehen von k roten Kugeln. Für jede dieser Varianten gibt es nun aber noch eine bestimmte Anzahl an Möglichkeiten zum Ziehen der dazugehörigen $n - k$ nicht roten Kugeln, die in der zweiten Klammer steht. Um die Gesamtanzahl passender Ziehungsvarianten zu ermitteln, muss man entsprechend die erste Klammer mit der zweiten Klammer multiplizieren. Jede einzelne Variante des Ziehens der k roten Kugeln muss ja mit allen in Frage kommenden Reihenfolgen der nicht roten Kugeln multipliziert werden. Und unter dem Bruchstrich steht die Gesamtzahl an möglichen

Ziehungsergebnissen für n Kugeln aus insgesamt N. Das Laplace-Experiment mit der zugehörigen Berechnung der Wahrscheinlichkeit lässt grüßen:

$$P(X) = \frac{Anzahl\ der\ f\text{ür}\ das\ Ereignis\ X\ g\ddot{u}nstigen\ Ergebnisse}{Anzahl\ der\ m\ddot{o}glichen\ Ereignisse}$$

Zahlenmäßig lässt sich die folgende Verteilungstabelle berechnen:

Mögliche Werte k	0	1	2
$P(X = k)$	$\frac{\binom{2}{0}\cdot\binom{8}{4}}{\binom{20}{4}} = 0{,}33$	$\frac{\binom{2}{1}\cdot\binom{8}{3}}{\binom{20}{4}} = 0{,}53$	$\frac{\binom{2}{2}\cdot\binom{8}{2}}{\binom{20}{4}} = 0{,}13$

Denken Sie an den Rechentest bei Verteilungstabellen: Die Summe aller Wahrscheinlichkeiten muss Eins ergeben. Stimmt das so wie hier, ist die Gefahr eines Rechenfehlers schon mal sehr gering ...

Bernoulli-Experiment

Ein Zufallsexperiment, bei dem nur zwei Ergebnisse möglich sind, bezeichnet man als Bernoulli-Experiment. Die Ergebnismenge $\Omega = \{e, m\}$ umfasst also zwei Elemente. Das Eintreten eines Ereignisses $E = \{e\}$ wird häufig als Erfolg oder Treffer und das Eintreten des Gegenereignisses $\bar{E} = \{m\}$ als Misserfolg oder Niete interpretiert.

Als Beispiele für Bernoulli-Experimente kommen zum Beispiel in Frage: der Münzwurf, das Würfeln einer bestimmten Zahl (Zahl 3 oder nicht Zahl 3), das Ziehen eines Gewinns bei einer Lotterie, das Geschlecht eines Menschen.

Die Wahrscheinlichkeitsfunktion P ist bestimmt durch eine Zahl p mit $(0 \leq p \leq 1)$:

$P(E) = p$

und

$P(\bar{E}) = 1 - p$

Es gilt weiterhin:

$P(\emptyset) = 0$

und

$P(\Omega) = 1$

Die Zufallsgröße X mit

$X(e) = 1$

und

$X(m) = 0$

ist 0-1-verteilt mit dem Parameter p. Daraus ergibt sich die folgende Verteilungstabelle:

Mögliche Werte x x	1	0
$P(X = x)$	p	$1 - p$

Bernoulli-Kette

Zufallsversuche, bei denen mehrmals hintereinander und unabhängig voneinander die gleichen Bernoulli-Experimente durchgeführt werden, heißen Bernoulli-Ketten. Typisch ist ja zum Beispiel für Qualitätskontrollen bei einer laufenden Produktion, dass Proben mehrmals hintereinander und unabhängig voneinander entnommen werden.

Simulieren lässt sich eine solche Bernoulli-Kette mit einem Urnenmodell: In einer Urne befinden sich n gleich große Kugeln, von denen m weiß und $n - m$ schwarz sind. Jedes Bernoulli-Experiment besteht im Ziehen einer Kugel mit Zurücklegen. Der Parameter p kann über

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

berechnet werden. p ist hier die Wahrscheinlichkeit für das Ziehen einer weißen Kugel. Ohne Zurücklegen handelt es sich um keine Bernoulli-Kette, da sich die Wahrscheinlichkeiten von Kugel zu Kugel verändern. Die noch in der Urne befindlichen Kugeln nehmen ja von Zug zu Zug ab.

Die zu einer Bernoulli-Kette der Länge n gehörende Ergebnismenge hat stets folgendes Aussehen:

$$\Omega_n = Menge\ aller\ geordneten\ n - Tupel\ aus\ \{e, m\}$$

Wird zum Beispiel eine Münze zweimal geworfen, wobei das Obenliegen der Zahl als Erfolg e gelte, umfasst die Ergebnismenge Ω die Ergebnisse:

$$\Omega_2 = \{(e;\ e); (e;\ m); (m;\ e); (m;\ m)\}$$

Ω_2 hat damit $2^2 = 4$ Elemente. Bei vierfachem Münzwurf würden sich $2^4 = 16$ Elemente ergeben.

Die Gesamtwahrscheinlichkeit für das Eintreten eines bestimmten Elementarereignisses ergibt sich auf Grund der Unabhängigkeit der Teilversuche immer aus dem Produkt der Einzelwahrscheinlichkeiten für die Teilversuche – die Ergebnisse der Teilversuche müssen UND-verknüpft werden:

$$E = E_1 \cap E_2 \cap E_3 \cap \dots \cap E_n$$
$$P(E) = P(E_1 \cap E_2 \cap E_3 \cap \dots \cap E_n)$$
$$P(E) = P(E_1) \cdot P(E_2) \cdot P(E_3) \cdot \dots \cdot P(E_n)$$
$$P(E) = p^n$$

Damit erhält man für die Wahrscheinlichkeit eines beliebigen Elementarereignisses F mit n Versuchen, von denen k erfolgreich und entsprechend $n - k$ nicht erfolgreich verliefen, den Ausdruck:

$$P(F) = p^k \cdot (1 - p)^{n-k}$$

Die Reihenfolge der Erfolge und Misserfolge spielt dabei keine Rolle. Da die beiden Ausdrücke p^k (steht für die Erfolge) und $(1 - p)^{n-k}$ (steht für die Misserfolge) multiplikativ miteinander verbunden sind und die Faktoren bei einer Multiplikation bekanntlich ohne Auswirkung auf das Ergebnis vertauscht werden können, kommt es auch hier nur auf die Anzahlen von Erfolgen und Misserfolgen an. Stimmen diese Anzahlen bei unterschiedlichen Reihenfolgen überein, führen sie zum gleichen Ergebnis.

Die Wahrscheinlichkeit, beim zweifachen Münzwurf zweimal die Zahl zu erreichen – das entspricht dem Fall {Zahl; Zahl}, beträgt somit

$$P(Zahl, Zahl) = 0{,}5^2 \cdot (1 - 0{,}5)^{2-2} = 0{,}25$$

Eine identische Zahl erhält man für die Fälle {Zahl; Nichtzahl} und {Nichtzahl; Zahl}, wobei in der folgenden ersten Zeile die triviale Berechnung der Wahrscheinlichkeit ohne die obige Formel, nur einfach gemäß Pfad im dazugehörigen Baumdiagramm aufgeschrieben ist – die Faktoren wechseln zwar, ändern aber damit nichts am Ergebnis:

$$P(Zahl, Nichtzahl) = 0{,}5^1 \cdot (1 - 0{,}5)^{2-1} = 0{,}25$$

Für den Fall {Nichtzahl; Nichtzahl} ergibt sich:

$$P(Nichtzahl, Nichtzahl) = 0{,}5^0 \cdot (1 - 0{,}5)^{2-0} = 0{,}25$$

Die zahlenmäßige Gleichheit der Ergebnisse liegt daran, dass bei einer guten Münze die Wahrscheinlichkeit für das Obenliegen der Zahl mit 0,5 genauso groß ist, wie die Wahrscheinlichkeit für das Untenliegen der Zahl. Somit ergibt sich für p, aber auch für $1 - p$ der Wert 0,5. Wer sich nach einem anspruchsvolleren Beispiel sehnt, der sei auf einen Würfel verwiesen.

Berechnen Sie (sich) die Wahrscheinlichkeit für das k-fache Würfeln der Zahlen 3 oder 4 (Erfolg) – alle anderen Zahlen stehen also für Misserfolg – bei n Würfelversuchen. Wir stimmen sicher überein, dass hier $p = 2/6 = 1/3$ gilt. Die restliche Berechnung schaffen Sie ohne Vorgaben.

Binomialverteilung

Die Binomialverteilung geht ebenfalls von einer Folge von Bernoulli-Experimenten – also einer Bernoulli-Kette aus. Das heißt die zugehörigen Experimente liefern nur die beiden Aussagen Erfolg = Treffer oder Misserfolg = Niete. Für die Wahrscheinlichkeit eines beliebigen Elementarereignisses F mit n Versuchen, von denen k erfolgreich und entsprechend $n - k$ nicht erfolgreich verliefen, ergab sich:

$$P(F) = p^k \cdot (1 - p)^{n-k}$$

Im Gegensatz dazu liefert die Binomialverteilung eine Berechnung der Wahrscheinlichkeit des Falles, dass während einer Bernoulli-Kette k Erfolge auftreten, d. h. die entsprechende Zufallsgröße X den Wert k annimmt. Der Unterschied zur obigen Formel besteht darin, dass die Binomialverteilung alle in Frage kommenden Kombinationen von k Erfolgen in n Versuchen berücksichtigen muss. Für eine Auswahl von k aus n Elementen gibt es ja bekanntlich

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

Möglichkeiten (da diese Größe auch für die Koeffizienten in der binomischen Formel zutrifft, heißt die Verteilung Binomialverteilung). Damit ist für die Binomialverteilung schon alles notwendige gesagt bzw. geschrieben. Für die Wahrscheinlichkeit, dass in einer Bernoulli-Kette mit n Experimenten genau k Erfolge auftreten, gilt:

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

Während der Ausdruck $P(F)$ nur *einen* der möglichen Pfade mit k Erfolgen und $n - k$ Misserfolgen durch das Baumdiagramm berücksichtigt, bezieht die Formel für $P(X = k)$ *alle* Pfade mit k Erfolgen und $n - k$ Misserfolgen im Baumdiagramm ein.

Schreibweise für die Binomialverteilung von X mit den Parametern n und p:

$$X \sim B(n; p)$$

Der Ausdruck

$$\binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

wird auch mit der Zeichenfolge

$B(n; p; k)$

abgekürzt.

Bemühen wir hier als Beispiel wieder einmal den Würfel: Bei $n = 4$ Würfen soll das Fallen der Zahl 6 als Erfolg ($p = 1/6$), alle anderen Zahlen als Misserfolg ($1 - p = 5/6$) gelten. Das heißt die Anzahl für das Würfeln der Zahl 6 spiegelt sich in der Variablen k wieder. Wir stimmen sicher überein, dass die Zahl 6 von gar nicht ($k = 0$) bis maximal vier Mal ($k = n = 4$) fallen kann. Untersuchen wir die zugehörigen Wahrscheinlichkeiten $P(X = k)$, ergeben sich die folgenden Zahlenwerte:

Mögliche Werte k	0	1	2	3	4
$P(X = k)$	0,4823	0,3858	0,1157	0,0154	0,0008

Vom normalen Verständnis her kann man den Zahlen sicher folgen, insofern bei 4 Würfen die Wahrscheinlichkeiten dafür, dass die Zahl 6 gar nicht oder nur einmal fällt, am größten sind. Wollen wir noch die Verteilungsfunktion $P(X < k)$ in Zahlen fassen? Hier sind die Ergebnisse (die Tabelle mit der Verteilungsfunktion wird am Beginn dieses Kapitels erklärt):

Mögliche Werte k	$k \leq 0$	$0 < k \leq 1$	$1 < k \leq 2$	$2 < k \leq 3$	$3 < k \leq 4$	$4 < k$
$P(X < k)$	0	0,4823	0,8681	0,9838	0,9992	1

Der Erwartungswert bei einer Binomialverteilung berechnet sich über die Formel:

$$E(X) = n \cdot p$$

Beim gerade erläuterten Würfelbeispiel beträgt der Erwartungswert für die Häufigkeit X der Zahl 6 bei 4 Würfen somit:

$$E(X) = n \cdot p = 4 \cdot \frac{1}{6} = \frac{2}{3}$$

Die Varianz kann bei einer Binomialverteilung über die Beziehung

$$V(X) = n \cdot p \cdot (1 - p)$$

ermittelt werden. Für den Würfel ergibt sich:

$$V(X) = n \cdot p \cdot (1 - p) = 4 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{9}$$

Ungleichung von Tschebyschev

Führt man ein Bernoulli-Experiment mit n Versuchen mehrfach aus, so ist zu erwarten, dass sich die dabei auftretenden absoluten Häufigkeiten X der Erfolge in der Nähe des Erwartungswertes $E(X)$ der zugehörigen binomialverteilten Zufallsgröße X einstellen. Der Begriff Nähe – immer abhängig vom Versuchscharakter des jeweiligen Experiments – wird dabei mit einem reellen Wert $a > 0$ beschrieben. Es wird die Frage gestellt, mit welcher Wahrscheinlichkeit die einzelne Häufigkeit im Intervall

$$[-a + E(X); E(X) + a]$$

liegt. Um den Erwartungswert $E(X)$ wird quasi ein Korridor der Breite $2a$ gelegt – einmal a „nach unten“ und einmal a „nach oben“. Für die betreffende Wahrscheinlichkeit kann man schreiben:

$P(-a + E(X) \leq X \leq E(X) + a) = P(|X - E(X)| \leq a)$

Die Wahrscheinlichkeit dafür, dass sich die Häufigkeit X in der a-Umgebung von $E(X)$ befindet, ist gleich der Wahrscheinlichkeit dafür, dass der Betrag der Differenz von X und $E(X)$ kleiner gleich a ist. Der Herr Tschebyschev hat nun seinerzeit nachgewiesen, dass:

$$P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}$$

gilt (deshalb wurde diese Gleichung auch nach ihm benannt). Geht man zum Gegenereignis über, verwandelt sich die Formel in:

$$P(|X - E(X)| < a) \geq 1 - \frac{V(X)}{a^2}$$

Handelt es sich bei X um eine binomialverteilte Zufallsgröße, erzeugen die zugehörigen Formeln für $E(X)$ und $V(X)$:

$$P(|X - n \cdot p| \geq a) \leq \frac{n \cdot p \cdot (1-p)}{a^2}$$

$$P(|X - n \cdot p| < a) \geq 1 - \frac{n \cdot p \cdot (1-p)}{a^2}$$

Mit der Tschebyschevschen Ungleichung kann man also unter anderem bestimmen, nach wieviel Einzelversuchen eine statistische Größe mit einer bestimmten Wahrscheinlichkeit außerhalb oder innerhalb einer vorher festzulegenden Grenze a liegt.

Für ein Zahlenbeispiel erinnern wir uns des Zusammenhanges

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

mit σ als Standardabweichung. Setzt man diese Größe in die Tschebyschevsche Ungleichung in ihrer zweiten Form ein, verwandelt sich diese in

$$P(|X - E(X)| < a) \geq 1 - \frac{\sigma^2}{a^2}$$

Verwendet man nun für die Zahl a Vielfache der Standardabweichung σ, ergeben sich die Zahlen in der folgenden Tabelle.

Wert für a	Formel	Ergebnis
$1 \cdot \sigma$	$P(\lvert X - E(X)\rvert < \sigma) \geq 1 - \frac{\sigma^2}{\sigma^2}$	0
$2 \cdot \sigma$	$P(\lvert X - E(X)\rvert < 2 \cdot \sigma) \geq 1 - \frac{\sigma^2}{4 \cdot \sigma^2}$	$\frac{3}{4} = 0{,}75 = 75\ \%$
$3 \cdot \sigma$	$P(\lvert X - E(X)\rvert < 3 \cdot \sigma) \geq 1 - \frac{\sigma^2}{9 \cdot \sigma^2}$	$\frac{8}{9} = 0{,}89 = 89\ \%$
$4 \cdot \sigma$	$P(\lvert X - E(X)\rvert < 4 \cdot \sigma) \geq 1 - \frac{\sigma^2}{16 \cdot \sigma^2}$	$\frac{15}{16} = 0{,}94 = 94\ \%$

Die Wahrscheinlichkeit, dass die Zufallsgröße X sich um weniger als die Standardabweichung σ vom Erwartungswert $E(X)$ unterscheidet, ist Null – in diesem engen Bereich um $E(X)$ liegen also überhaupt keine Treffer. Im Bereich der Größe $2 \cdot \sigma$ um $E(X)$, liegen dagegen immerhin schon 75 % aller Treffer usw.