minærva

M

22

SKRIPT

Bedingte Wahrscheinlichkeit und Binomialverteilung

Erklärskript

Die Aufgaben 1049–1090 zu diesem Skript stehen im Aufgabenskript, die Lösungen im Lösungsskript. Verweise auf Aufgaben sind Links.

Mathematik – Klasse 5 bis 10, G8 und G9

Pirmin Gohn

Nutzungshinweis: Dieses Skript darf nur im Rahmen des Unterrichts verwendet werden; insbesondere ist die eigenständige Weitergabe untersagt.

Skriptreihe M · Mathematik (Klasse 5 bis 10, G8 und G9)

M1 Zahlen	M2 Größen	M3 Geometrie I	M4 Brüche	M5 Bruchrechnung	M6 Negative Zahlen	M7 Flächen, Körper	M8 Zuordnungen
M9 Prozent	M10 Terme	M11 Geometrie II	M12 Lineare Funktionen	M13 Wurzeln	M14 Zufall	M15 Strahlensätze	M16 Quadratische Funktionen
M17 Trigonometrie	M18 Exponentialfunktion	M19 Kreis und Körper	M20 Ableitung	M21 Vektoren	**M22 Binomialverteilung**		

Lernen mit minærva

Dieses Skript ist ein Teil von minærva. Alles gehört zusammen: Jedes Thema führt von der Erklärung zum Üben, zum Ausprobieren und ins große Bild.

minaerva.de

Skripte Verstehen
Der Stoff zum Nachlesen: Erklärungen, Herleitungen, Beispiele und typische Fehler.
minaerva.de

noctua Üben
Fragen zu jedem Abschnitt, mit sofortiger Rückmeldung.
noctua.minaerva.de

fabrica Ausprobieren
Versuche und Anschauungen zum Selbermessen und Anfassen, ohne Konto, auch auf dem iPad.
fabrica.minaerva.de

tela Vernetzen
Das Wissensnetz: worauf ein Thema aufbaut und wohin es führt.
tela.minaerva.de

So greifen die Skripte ineinander

Erklärskripte	Mathematik: zweiundzwanzig Themenskripte M1–M22 für die Klassen 5 bis 10 mit dem Stoff, aber ohne Aufgaben. Unter jeder Abschnittsüberschrift steht, in welcher Klasse das Thema im G8 und im G9 drankommt.
Aufgabenskript	Ein gemeinsames Aufgabenskript für alle Themen. Die Erklärskripte verweisen darauf; jede Aufgaben-, Abbildungs- und Tabellennummer gibt es in der Reihe nur einmal.
Lösungsskript	Die Lösungen zu allen Aufgaben, mit denselben Nummern.

Am Ende vieler Abschnitte stehen Links zur passenden Übung, zum Versuch und ins Wissensnetz. In der PDF-Datei lassen sie sich direkt antippen.

Inhaltsverzeichnis

Worum es geht **4**
1 Vierfeldertafel, bedingte Wahrscheinlichkeit **5**
1.1 Bedingte Wahrscheinlichkeit 5
1.2 Baumdiagramm und umgekehrtes Baumdiagramm 6
1.3 Beispiel: ein medizinischer Test 6
2 Stochastische Unabhängigkeit **8**
2.1 Unabhängigkeit sehen 8
2.2 Unabhängigkeit bei Daten 9
3 Zufallsgrößen und Erwartungswert **10**
3.1 Der Erwartungswert 10
3.2 Faire Spiele 11
4 Bernoulli-Ketten, Binomialkoeffizient **12**
4.1 Pfade mit gleich vielen Treffern 12
4.2 Der Binomialkoeffizient 12
4.3 Das Pascalsche Dreieck 13
5 Binomialverteilung **15**
5.1 Kumulierte Wahrscheinlichkeiten 15
5.2 Wie die Parameter das Histogramm formen 16
5.3 Wie oft muss man es versuchen? 17
6 Erwartungswert und Standardabweichung **18**
6.1 Streuung: Varianz und Standardabweichung 18
6.2 Was σ über die Verteilung sagt 19

Worum es geht

Ein Schnelltest zeigt „positiv“. Der Hersteller schreibt, der Test erkenne 95 % aller Erkrankten. Heißt das, dass man mit 95 % Wahrscheinlichkeit krank ist? Die überraschende Antwort lautet: meistens nicht einmal annähernd. Um das zu verstehen, muss man genau sagen, *unter welcher Bedingung* eine Wahrscheinlichkeit gilt. Das ist die Idee der **bedingten Wahrscheinlichkeit**.

Im zweiten Teil geht es um Zufallsexperimente, die aus vielen gleichen Einzelversuchen bestehen: zehn Fragen raten, hundert Freiwürfe werfen, zweihundert Lose ziehen. Man fragt dann nicht mehr nach einem einzelnen Ergebnis, sondern nach einer Anzahl: Wie viele Treffer sind zu erwarten, und wie stark schwankt die Zahl? Darauf antwortet die **Binomialverteilung**, eine der wichtigsten Verteilungen der ganzen Stochastik.

Vorausgesetzt werden Baumdiagramme, die Pfadregeln und das Zählprinzip aus M14.

Bedingte Wahrscheinlichkeit und Binomialverteilung im Lehrplan. G8: Klasse 10. G9: Klasse 10–11. Einzelne Abschnitte können davon abweichen; die Marken unter den Überschriften sagen es genau.

Lehrplan-Marken. Unter jeder Abschnittsüberschrift steht, in welcher Klassenstufe das Thema im achtjährigen (G8) und im neunjährigen Gymnasium (G9) vorkommt, z. B. G8 · Kl. 6 G9 · Kl. 6. Ein Strich (G9 –) heißt: Im Lehrplan dieses Zugs kommt der Abschnitt nicht vor. Grundlage ist der Bildungsplan Baden-Württemberg (G8: Fassung 2016, G9: Fassung vom 21.05.2026). Der Bildungsplan nennt nur Doppeljahrgänge (5/6, 7/8, 9/10); die Klasse ist eine Empfehlung nach der üblichen Reihenfolge der Schulbücher.

Die Aufgaben zu diesem Skript stehen im Aufgabenskript der Reihe, die Lösungen im Lösungsskript. Jeder Abschnitt endet mit Verweisen auf passende Übungen in **noctua** und Versuche in der **fabrica**.

1 Vierfeldertafel, bedingte Wahrscheinlichkeit

G8 · Kl. 10 G9 · Kl. 10

Alle 160 Schüler der zehnten Klassen einer Schule wurden gefragt: „Spielst du ein Instrument?“ und „Bist du in einem Sportverein?“ 64 spielen ein Instrument, 88 sind im Sportverein, und 32 tun beides. Wie passen diese Zahlen zusammen? Am übersichtlichsten trägt man sie in eine Tabelle ein, in der jede Person in genau einem der vier inneren Felder steht (Tabelle 64).

	S: Sportverein	$\overline{S}$: kein Verein	Summe
I: Instrument	32	32	64
$\overline{I}$: kein Instrument	56	40	96
Summe	88	72	160

Tabelle 64: Vierfeldertafel der Umfrage (absolute Häufigkeiten). Gegeben waren 64, 88, 32 und 160, der Rest folgt durch Ergänzen.

Aus den vier gegebenen Zahlen folgt alles andere durch Ergänzen: In der Zeile I fehlen $64 - 32 = 32$, in der Spalte S fehlen $88 - 32 = 56$, und so weiter. Am Ende muss jede Zeile und jede Spalte aufgehen.

Definition 1.1 (Vierfeldertafel)
Betrachtet man zwei Ereignisse A und B und ihre Gegenereignisse $\overline{A}$ und $\overline{B}$, so gibt es vier Kombinationen: $A \cap B$ („A und B“), $A \cap \overline{B}$, $\overline{A} \cap B$ und $\overline{A} \cap \overline{B}$. Eine **Vierfeldertafel** enthält in den vier inneren Feldern die Häufigkeiten oder Wahrscheinlichkeiten dieser Kombinationen, am Rand ihre Zeilen- und Spaltensummen.

Das Zeichen $\cap$ heißt „geschnitten“: $A \cap B$ ist das Ereignis, dass A *und* B eintreten. Wählt man eine der 160 Personen zufällig aus, so ist jede gleich wahrscheinlich (Laplace-Experiment), und man erhält die Wahrscheinlichkeiten, indem man durch 160 teilt (Tabelle 65). Zum Beispiel ist $P(I \cap S) = \frac{32}{160} = 0{,}2$.

	S	$\overline{S}$	Summe
I	0,2	0,2	0,4
$\overline{I}$	0,35	0,25	0,6
Summe	0,55	0,45	1

Tabelle 65: Dieselbe Umfrage als Vierfeldertafel mit Wahrscheinlichkeiten für eine zufällig ausgewählte Person.

Merksatz
In jeder Vierfeldertafel ergeben die beiden inneren Felder einer Zeile die Zeilensumme und die beiden inneren Felder einer Spalte die Spaltensumme. Die vier inneren Felder zusammen ergeben die Gesamtzahl bzw. 1.

1.1 Bedingte Wahrscheinlichkeit

Nun wählt man nur unter den Instrumentalisten eine Person zufällig aus. Wie wahrscheinlich ist sie im Sportverein? Jetzt zählt nur noch die Zeile I der Tafel: Von den 64 Personen dieser Zeile sind 32 im Verein, die Wahrscheinlichkeit ist also $\frac{32}{64} = 0{,}5$. Das ist etwas weniger als die 0,55 für alle Schüler. Man schreibt $P_I(S) = 0{,}5$ und liest „die Wahrscheinlichkeit von S unter der Bedingung I“.
Andersherum gefragt: Wählt man unter den Vereinssportlern zufällig eine Person aus, so spielt sie mit der Wahrscheinlichkeit $P_S(I) = \frac{32}{88} = \frac{4}{11} \approx 0{,}36$ ein Instrument. Jetzt zählt nur die Spalte S.

Definition 1.2 (bedingte Wahrscheinlichkeit)
Ist $P(A) > 0$, so heißt

$$P_A(B) = \frac{P(A \cap B)}{P(A)}$$

die **bedingte Wahrscheinlichkeit** von B unter der Bedingung A. Sie gibt an, wie wahrscheinlich B ist, wenn man schon weiß, dass A eingetreten ist.

Begründung Weiß man, dass A eingetreten ist, kommen nur noch die Fälle aus A infrage, und von ihnen sind die Fälle aus $A \cap B$ günstig. Mit Häufigkeiten ist also $P_A(B) = \frac{|A \cap B|}{|A|}$, wobei $|A|$ die Anzahl der Fälle in A ist. Teilt man Zähler und Nenner durch die Gesamtzahl n, erhält man $\frac{P(A \cap B)}{P(A)}$. In der Umfrage: $\frac{32}{64} = \frac{32:160}{64:160} = \frac{0{,}2}{0{,}4} = 0{,}5$.

Bemerkung
In vielen Büchern schreibt man statt $P_A(B)$ auch $P(B \mid A)$. Die Bedingung steht dann hinter dem Strich.

Typischer Fehler
Falsch: $P_I(S) = P_S(I)$
Richtig: $P_I(S) = \frac{32}{64} = 0{,}5$, aber $P_S(I) = \frac{32}{88} \approx 0{,}36$

Bei $P_I(S)$ wird durch die Zahl der Instrumentalisten geteilt, bei $P_S(I)$ durch die Zahl der Vereinssportler. Der Zähler ist gleich, der Nenner nicht. Die Bedingung legt fest, *wodurch* man teilt.

1.2 Baumdiagramm und umgekehrtes Baumdiagramm

Dieselbe Umfrage lässt sich als zweistufiges Baumdiagramm darstellen, und zwar auf zwei Arten: Man kann zuerst nach dem Instrument und dann nach dem Verein verzweigen oder umgekehrt (Abbildung 372). Auf den Ästen der zweiten Stufe stehen dabei *bedingte* Wahrscheinlichkeiten: Nach dem Ast I ist der Ast S mit $P_I(S) = \frac{1}{2}$ beschriftet, nach dem Ast $\overline{I}$ mit $P_{\overline{I}}(S) = \frac{56}{96} = \frac{7}{12}$.

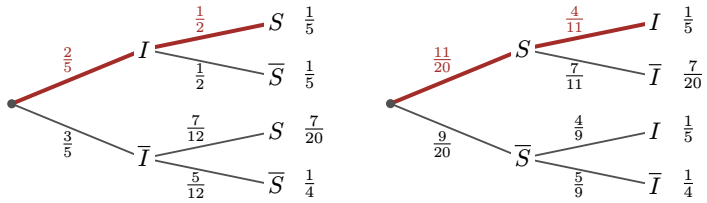


Abbildung 372: Links zuerst nach dem Instrument, rechts zuerst nach dem Verein verzweigt (umgekehrtes Baumdiagramm). Hervorgehoben ist jeweils der Pfad zu $I \cap S$; beide Pfade ergeben $P(I \cap S) = \frac{1}{5}$.

Die Pfadregeln aus M14 gelten unverändert. Die Produktregel entlang des hervorgehobenen Pfades links lautet $\frac{2}{5} \cdot \frac{1}{2} = \frac{1}{5}$, und das ist nichts anderes als die Definition der bedingten Wahrscheinlichkeit, nach $P(A \cap B)$ aufgelöst. Schon beim Ziehen ohne Zurücklegen in M14 standen auf der zweiten Stufe bedingte Wahrscheinlichkeiten, nur hießen sie dort noch nicht so.

Satz 1.1 (Produktregel mit bedingter Wahrscheinlichkeit)
Für zwei Ereignisse A und B mit $P(A) > 0$ gilt

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

Begründung Man multipliziert die Definition $P_A(B) = \frac{P(A \cap B)}{P(A)}$ auf beiden Seiten mit $P(A)$.

Den rechten, umgekehrten Baum braucht man, wenn man von der zweiten Stufe auf die erste zurückschließen will. Dazu füllt man die Vierfeldertafel aus dem ersten Baum und liest den zweiten daraus ab.

Vorgehen
Gegeben ist ein Baum, der zuerst nach A und dann nach B verzweigt. Gesucht ist $P_B(A)$.
a) Pfadwahrscheinlichkeiten mit der Produktregel berechnen und in die Vierfeldertafel eintragen.
b) $P(B)$ als Spaltensumme bestimmen (Summenregel: alle Pfade, die zu B führen).
c) $P_B(A) = \frac{P(A \cap B)}{P(B)}$ berechnen.
d) Bei Bedarf das umgekehrte Baumdiagramm vollständig beschriften.

1.3 Beispiel: ein medizinischer Test

Eine Krankheit hat einer von hundert Menschen; man sagt, ihr **Anteil in der Bevölkerung** beträgt 1 %. Ein Test erkennt 95 % der Kranken (er zeigt bei ihnen „positiv“) und 90 % der Gesunden (er zeigt bei ihnen „negativ“). Eine zufällig ausgewählte Person bekommt ein positives Ergebnis. Wie wahrscheinlich ist sie krank?

Viele antworten spontan „95 %“. Ein **Häufigkeitsbaum** zeigt, warum das falsch ist: Man stellt sich 10000 getestete Personen vor und verteilt sie auf die Äste (Abbildung 373).

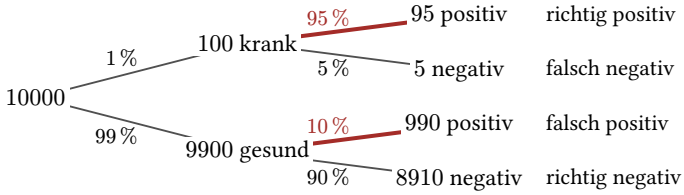


Abbildung 373: Häufigkeitsbaum für 10000 getestete Personen. Hervorgehoben sind die beiden Äste mit positivem Ergebnis.

Positiv getestet werden $95 + 990 = 1085$ Personen, und nur 95 davon sind krank:

$$P_{\text{positiv}}(\text{krank}) = \frac{95}{1085} \approx 0{,}088.$$

Eine positiv getestete Person ist also nur mit etwa 9 % Wahrscheinlichkeit krank. Der Grund: Es gibt so viele Gesunde, dass schon ihre 10 % Fehlalarme (990 Personen) die 95 richtig erkannten Kranken weit übertreffen. Ein negatives Ergebnis ist dagegen sehr verlässlich: $P_{\text{negativ}}(\text{gesund}) = \frac{8910}{8915} \approx 0{,}9994$. In der Vierfeldertafel stehen dieselben Zahlen (Tabelle 66).

	positiv	**negativ**	**Summe**
krank	95	5	100
gesund	990	8910	9900
Summe	1085	8915	10000

Tabelle 66: Vierfeldertafel des Tests für 10000 Personen. Für $P_{\text{positiv}}(\text{krank})$ zählt nur die markierte Spalte.

Bemerkung

In der Medizin heißen die beiden Kennzahlen eines Tests **Sensitivität** ($P_{\text{krank}}(\text{positiv})$, hier 95 %) und **Spezifität** ($P_{\text{gesund}}(\text{negativ})$, hier 90 %). Für die getestete Person zählt aber die umgekehrte Frage $P_{\text{positiv}}(\text{krank})$. Sie hängt stark davon ab, wie häufig die Krankheit überhaupt ist. Deshalb wird ein positives Ergebnis bei einem Reihentest immer durch einen zweiten, genaueren Test überprüft.

Typischer Fehler

Falsch: „Der Test erkennt 95 % der Kranken. Ich bin positiv, also bin ich zu 95 % krank.“
Richtig: 95 % ist $P_{\text{krank}}(\text{positiv})$. Gefragt ist $P_{\text{positiv}}(\text{krank}) \approx 0{,}09$, also etwa 9 %.

Bedingung und Ereignis sind vertauscht. Wer mit Häufigkeiten statt mit Prozenten rechnet, macht diesen Fehler kaum.

Üben: **noctua** ↗ Ausprobieren: **Medizinischer Test** ↗ · **Baumdiagramm zum Anfassen** ↗ · **Vierfeldertafel und Baum** ↗

2 Stochastische Unabhängigkeit

G8 · Kl. 10 G9 · Kl. 10

In der Umfrage ist $P(S) = 0{,}55$. Unter den Instrumentalisten ist der Anteil der Vereinssportler $P_I(S) = 0{,}5$, unter den übrigen Schülern $P_{\overline{I}}(S) = \frac{7}{12} \approx 0{,}58$. Wer weiß, dass jemand ein Instrument spielt, schätzt die Wahrscheinlichkeit für den Verein also etwas anders ein. Die beiden Merkmale **hängen** (ein wenig) **voneinander ab**.

Ganz anders bei einer Münze und einem Tetraeder (vier gleiche Flächen mit 1 bis 4), die man gleichzeitig wirft: Zu wissen, dass die Münze Wappen zeigt, ändert nichts an der Wahrscheinlichkeit $\frac{1}{4}$ für eine 4. Solche Ereignisse heißen unabhängig. Dann ist $P_A(B) = P(B)$, und die Produktregel wird besonders einfach: $P(A \cap B) = P(A) \cdot P_A(B) = P(A) \cdot P(B)$.

Definition 2.1 (stochastisch unabhängig)

Zwei Ereignisse A und B heißen **stochastisch unabhängig**, wenn

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

gilt. Andernfalls heißen sie **stochastisch abhängig**.

Satz 2.1 (Unabhängigkeit und bedingte Wahrscheinlichkeit)

Ist $P(A) > 0$, so sind A und B genau dann unabhängig, wenn $P_A(B) = P(B)$ gilt: Das Eintreten von A ändert die Wahrscheinlichkeit von B nicht.

Begründung $P_A(B) = P(B)$ heißt $\frac{P(A \cap B)}{P(A)} = P(B)$. Multipliziert man mit $P(A)$, erhält man $P(A \cap B) = P(A) \cdot P(B)$, und umgekehrt.

Man nimmt die Produktform als Definition, weil sie in A und B symmetrisch ist und auch dann einen Sinn hat, wenn $P(A) = 0$ ist.

Beispiel 1

Ein Oktaeder hat acht gleich große Flächen mit den Zahlen 1 bis 8; es wird einmal geworfen. A: „gerade Zahl“ $= \{2, 4, 6, 8\}$, B: „höchstens 4“ $= \{1, 2, 3, 4\}$, C: „Primzahl“ $= \{2, 3, 5, 7\}$.

- $A \cap B = \{2, 4\}$, also $P(A \cap B) = \frac{2}{8} = \frac{1}{4}$. Es ist $P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Die Ereignisse A und B sind unabhängig.
- $A \cap C = \{2\}$, also $P(A \cap C) = \frac{1}{8}$. Aber $P(A) \cdot P(C) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Die Ereignisse A und C sind abhängig: Unter den geraden Zahlen ist nur ein Viertel prim, unter allen Zahlen die Hälfte.

2.1 Unabhängigkeit sehen

Zeichnet man ein Quadrat mit der Fläche 1 und teilt es senkrecht im Verhältnis $P(A) : P\left(\overline{A}\right)$, dann jede Spalte waagerecht nach dem Anteil von B in dieser Spalte, so sind die vier Teilflächen genau die Wahrscheinlichkeiten der vier Felder der Vierfeldertafel (Abbildung 374). Bei Unabhängigkeit ist der Anteil von B in beiden Spalten gleich, und die waagerechte Trennlinie läuft gerade durch. Bei Abhängigkeit hat sie eine Stufe.

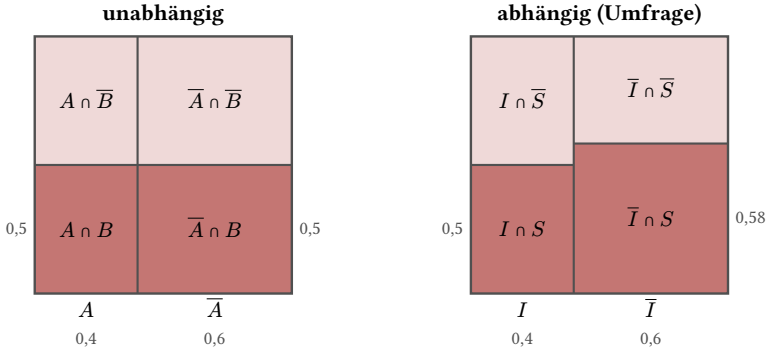


Abbildung 374: Flächendiagramme: links $P(A) = 0{,}4$ und $P_A(B) = P_{\overline{A}}(B) = 0{,}5$, rechts die Umfrage mit $P_I(S) = 0{,}5$ und $P_{\overline{I}}(S) \approx 0{,}58$. Am Rand stehen die Anteile.

Im Baumdiagramm erkennt man Unabhängigkeit daran, dass auf der zweiten Stufe in allen Zweigen dieselben Wahrscheinlichkeiten stehen, wie beim Ziehen *mit* Zurücklegen. Beim Ziehen *ohne* Zurücklegen ändern sie sich, die Ziehungen sind abhängig.

Satz 2.2 (Gegenereignisse)
Sind A und B unabhängig, so sind auch A und $\overline{B}$, $\overline{A}$ und B sowie $\overline{A}$ und $\overline{B}$ unabhängig.

Beweis Für A und $\overline{B}$: In der Zeile A der Vierfeldertafel ist $P\left(A \cap \overline{B}\right) = P(A) - P(A \cap B)$. Wegen der Unabhängigkeit ist das

$$P(A) - P(A) \cdot P(B) = P(A) \cdot (1 - P(B)) = P(A) \cdot P\left(\overline{B}\right).$$

Für die anderen Paare rechnet man genauso. ∎

2.2 Unabhängigkeit bei Daten

In der Umfrage ist $P(I) \cdot P(S) = 0{,}4 \cdot 0{,}55 = 0{,}22$, aber $P(I \cap S) = 0{,}2$. Streng genommen sind die Merkmale abhängig, doch der Unterschied ist klein. Bei echten Daten stimmt die Produktformel fast nie genau, schon weil Anzahlen ganze Zahlen sind. Ob eine kleine Abweichung bloßer Zufall ist oder auf einen echten Zusammenhang hinweist, lernt man in der Oberstufe zu prüfen.

Auch eine deutliche Abhängigkeit sagt nichts über Ursache und Wirkung (siehe M14): An Tagen mit viel Eisverkauf gibt es auch mehr Sonnenbrände. Beides hängt voneinander ab, weil beides vom Sonnenschein abhängt; das Eis verursacht keinen Sonnenbrand.

Typischer Fehler
Falsch: „A und B schließen sich aus, sie haben nichts miteinander zu tun, also sind sie unabhängig."
Richtig: Sind $P(A) > 0$ und $P(B) > 0$ und schließen sich A und B aus, so ist $P(A \cap B) = 0 \neq P(A) \cdot P(B)$. Sie sind abhängig.

Unvereinbare Ereignisse hängen sogar stark voneinander ab: Weiß man, dass A eingetreten ist, dann ist B unmöglich.

Üben: **noctua** ↗ Ausprobieren: **Vierfeldertafel und Baum** ↗

3 Zufallsgrößen und Erwartungswert

G8 · Kl. 10 G9 · Kl. 11

Beim Schulfest gibt es eine Kugelbude: Man zahlt 1 € Einsatz und zieht blind eine Kugel aus einem Beutel mit 10 Kugeln, nämlich 1 goldenen, 2 silbernen und 7 weißen. Für Gold bekommt man 4 € ausgezahlt, für Silber 1 € (also den Einsatz zurück), für Weiß nichts. Lohnt sich das? Hier interessiert nicht die Farbe selbst, sondern eine Zahl, die zu ihr gehört: der Gewinn.

Definition 3.1 (Zufallsgröße und Wahrscheinlichkeitsverteilung)
Eine **Zufallsgröße** X ordnet jedem Ergebnis eines Zufallsexperiments eine Zahl zu. „$X = x$“ ist das Ereignis, dass X den Wert x annimmt. Die Tabelle aller Werte x_1, x_2, …, x_n mit den Wahrscheinlichkeiten $P(X = x_1)$, …, $P(X = x_n)$ heißt **Wahrscheinlichkeitsverteilung** von X. Die Wahrscheinlichkeiten ergeben zusammen 1.

Bei der Kugelbude sei G der Gewinn in Euro, also Auszahlung minus Einsatz. Bei Gold ist $G = 4 - 1 = 3$, bei Silber $G = 0$, bei Weiß $G = -1$:

Gewinn g in €	3	0	−1
$P(G = g)$	0,1	0,2	0,7

Beispiel 2
Drei Münzen werden geworfen, X ist die Anzahl der Wappen. Es gibt $2 \cdot 2 \cdot 2 = 8$ gleich wahrscheinliche Ergebnisse (Zählprinzip). Davon hat eines kein Wappen (ZZZ), drei haben ein Wappen (WZZ, ZWZ, ZZW), drei haben zwei Wappen und eines hat drei. Die Verteilung zeigt Abbildung 375.

k	0	1	2	3
$P(X = k)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

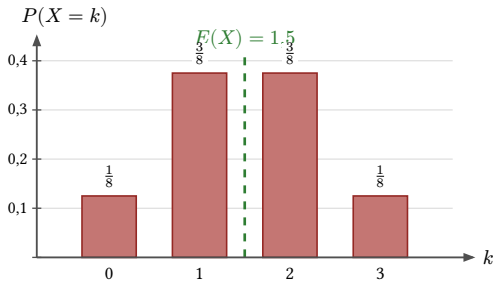


Abbildung 375: Wahrscheinlichkeitsverteilung der Anzahl X der Wappen bei drei Münzwürfen als Säulendiagramm. Die gestrichelte Linie markiert den Erwartungswert.

3.1 Der Erwartungswert

Was gewinnt man an der Kugelbude im Mittel? Man stellt sich vor, man spielt 1000-mal (die Kugel kommt jedes Mal zurück in den Beutel). Dann kommt ungefähr 100-mal Gold (+3 €), ungefähr 200-mal Silber (0 €) und ungefähr 700-mal Weiß (−1 €). Zusammen sind das etwa

$$100 \cdot 3 + 200 \cdot 0 + 700 \cdot (-1) = -400 \text{ €},$$

also $-400 : 1000 = -0{,}40$ € pro Spiel. Teilt man schon vorher durch 1000, sieht man, dass man einfach jeden Wert mit seiner Wahrscheinlichkeit multipliziert und addiert:

$$3 \cdot 0{,}1 + 0 \cdot 0{,}2 + (-1) \cdot 0{,}7 = -0{,}4.$$

Definition 3.2 (Erwartungswert)
Nimmt die Zufallsgröße X die Werte x_1, ..., x_n an, so heißt

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

der **Erwartungswert** von X. Man schreibt auch μ (griechisch „mü“). Auf lange Sicht ist der Mittelwert der beobachteten Werte ungefähr gleich $E(X)$ (Gesetz der großen Zahlen, M14).

Bemerkung
Mit dem Summenzeichen schreibt man kurz $E(X) = \sum_{i=1}^{n} x_i \cdot P(X = x_i)$.

Beispiel 3
Bei drei Münzwürfen ist $E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1{,}5$. Man erwartet im Mittel anderthalb Wappen, obwohl X nie den Wert 1,5 annimmt. Ebenso ist beim Wurf eines Tetraeders $E(X) = (1 + 2 + 3 + 4) \cdot \frac{1}{4} = 2{,}5$. Der Erwartungswert ist ein Mittelwert, kein Wert, den man „erwarten“ müsste.

Typischer Fehler
Falsch: $E(G) = \frac{3+0+(-1)}{3} = \frac{2}{3}$
Richtig: $E(G) = 3 \cdot 0{,}1 + 0 \cdot 0{,}2 + (-1) \cdot 0{,}7 = -0{,}4$

Die Werte sind nicht gleich wahrscheinlich. Jeder Wert muss mit seiner Wahrscheinlichkeit gewichtet werden; der einfache Mittelwert der Werte passt nur, wenn alle gleich wahrscheinlich sind.

3.2 Faire Spiele

Definition 3.3 (faires Spiel)
Ein Glücksspiel heißt **fair**, wenn der Erwartungswert des Gewinns (Auszahlung minus Einsatz) null ist. Gleichwertig: Die zu erwartende Auszahlung ist so groß wie der Einsatz.

Die Kugelbude ist nicht fair, denn $E(G) = -0{,}40$ € < 0: Auf lange Sicht verliert der Spieler pro Spiel 40 Cent, und die Klassenkasse freut sich.

Beispiel 4
Welche Auszahlung für Gold macht die Kugelbude fair? Es sei a die Auszahlung für Gold (in Euro); für Silber gibt es weiterhin 1 €. Die zu erwartende Auszahlung muss gleich dem Einsatz sein:

$$a \cdot 0{,}1 + 1 \cdot 0{,}2 = 1 \quad \rightarrow \quad 0{,}1a = 0{,}8 \quad \rightarrow \quad a = 8.$$

Probe mit dem Gewinn: $7 \cdot 0{,}1 + 0 \cdot 0{,}2 + (-1) \cdot 0{,}7 = 0$. Bei 8 € Auszahlung für Gold ist das Spiel fair.

Echte Glücksspiele sind nie fair, sonst verdiente der Veranstalter nichts. Beim Lotto wird zum Beispiel nur etwa die Hälfte der Einsätze wieder ausgezahlt: Wer 1 € einsetzt, bekommt im Mittel etwa 50 Cent zurück.

Üben: **noctua** ↗ Ausprobieren: **Erwartungswert und Varianz** ↗

4 Bernoulli-Ketten, Binomialkoeffizient

G8 · Kl. 10 G9 · Kl. 11

Oft interessiert bei einem Zufallsexperiment nur, ob etwas Bestimmtes eintritt oder nicht: Treffer oder kein Treffer, richtig geraten oder falsch, Teil defekt oder in Ordnung. Wiederholt man ein solches Experiment mehrmals unter gleichen Bedingungen, entsteht eine Bernoulli-Kette, benannt nach dem Schweizer Mathematiker Jakob Bernoulli (1655–1705).

Definition 4.1 (Bernoulli-Experiment und Bernoulli-Kette)
Ein **Bernoulli-Experiment** ist ein Zufallsexperiment mit genau zwei Ergebnissen, **Treffer** (T) mit der Wahrscheinlichkeit p und **Niete** (N) mit der Wahrscheinlichkeit $q = 1 - p$. Wird es n-mal so wiederholt, dass die Durchgänge unabhängig sind und p gleich bleibt, entsteht eine **Bernoulli-Kette der Länge n**.

Was als Treffer gilt, legt man selbst fest; es muss nichts Erfreuliches sein. Einige Beispiele:

- 10-mal ein Oktaeder werfen, Treffer „8“: Bernoulli-Kette mit $n = 10$, $p = \frac{1}{8}$.
- 20 Fragen mit je vier Antworten raten, Treffer „richtig“: $n = 20$, $p = \frac{1}{4}$.
- 5-mal mit Zurücklegen aus einer Urne mit 3 roten und 7 weißen Kugeln ziehen, Treffer „rot“: $n = 5$, $p = 0{,}3$.
- Ziehen *ohne* Zurücklegen ist keine Bernoulli-Kette, weil sich p von Zug zu Zug ändert. Zieht man aber nur wenige aus einer sehr großen Menge (etwa 50 Personen aus einer Großstadt), ändert sich p kaum, und man rechnet näherungsweise mit einer Bernoulli-Kette.
- Werfen *bis zum ersten Treffer* ist keine Bernoulli-Kette der Länge n, weil die Zahl der Würfe nicht vorher feststeht.

4.1 Pfade mit gleich vielen Treffern

Abbildung 376 zeigt eine Bernoulli-Kette der Länge 3. Entlang eines Pfades multipliziert man für jeden Treffer p und für jede Niete q. Auf die Reihenfolge kommt es dabei nicht an: Alle Pfade mit genau k Treffern haben dieselbe Wahrscheinlichkeit $p^k \cdot q^{n-k}$.

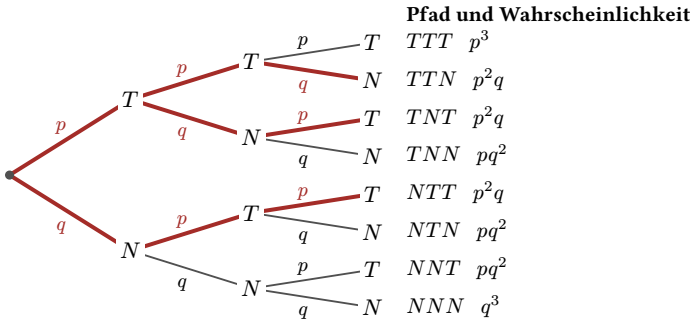


Abbildung 376: Bernoulli-Kette der Länge 3. Hervorgehoben sind die drei Pfade mit genau zwei Treffern; jeder hat die Wahrscheinlichkeit p^2q.

Genau zwei Treffer gibt es auf den drei Pfaden TTN, TNT und NTT. Also ist

$$P(\text{genau zwei Treffer}) = 3 \cdot p^2 \cdot q.$$

Wirft man zum Beispiel dreimal ein Tetraeder, so ist $P(\text{genau zweimal die } 4) = 3 \cdot \left(\frac{1}{4}\right)^2 \cdot \frac{3}{4} = \frac{9}{64} \approx 0{,}14$.
Für $n = 3$ kann man die Pfade noch abzählen. Bei $n = 10$ hat der Baum aber schon $2^{10} = 1024$ Pfade. Man braucht also eine Formel für die Anzahl der Pfade mit genau k Treffern.

4.2 Der Binomialkoeffizient

Ein Pfad mit k Treffern ist dadurch festgelegt, auf welchen k der n Stufen die Treffer liegen. Man muss also zählen, wie viele Möglichkeiten es gibt, aus n Stufen k auszuwählen. Das geht mit dem Zählprinzip aus M14 und der Überlegung mit dem doppelten Zählen von den Handschlägen dort:

- Wählt man die Treffer-Stufen *nacheinander* aus, hat man für die erste n Möglichkeiten, für die zweite noch $n - 1$, ..., für die k-te noch $n - k + 1$. Das sind $n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$ Möglichkeiten.

- Dabei kommt jede Auswahl derselben k Stufen in allen ihren Reihenfolgen vor, also $k! = k \cdot (k-1) \cdot \ldots \cdot 1$-mal. Man muss deshalb durch $k!$ teilen.

Für $n = 5$ und $k = 2$ ergibt das $\frac{5 \cdot 4}{2 \cdot 1} = 10$ Pfade mit genau zwei Treffern.

Definition 4.2 (Binomialkoeffizient)
Die Anzahl der Möglichkeiten, aus n Dingen k auszuwählen, ohne auf die Reihenfolge zu achten, heißt **Binomialkoeffizient** $\binom{n}{k}$ (gelesen „n über k“). Sie ist zugleich die Anzahl der Pfade mit genau k Treffern in einer Bernoulli-Kette der Länge n.

Satz 4.1 (Berechnung des Binomialkoeffizienten)
Für natürliche Zahlen $0 \le k \le n$ gilt

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!} = \frac{n!}{k! \cdot (n-k)!},$$

wobei man $0! = 1$ festlegt. Insbesondere ist $\binom{n}{0} = \binom{n}{n} = 1$.

Begründung Die erste Form ist die Überlegung oben. Für die zweite erweitert man den Bruch mit $(n-k)! = (n-k) \cdot \ldots \cdot 2 \cdot 1$; dann steht im Zähler das Produkt aller Zahlen von n bis 1, also $n!$. Aus n Dingen keines auszuwählen geht auf genau eine Art, deshalb $\binom{n}{0} = 1$, und die Festlegung $0! = 1$ sorgt dafür, dass die Formel auch dann stimmt.

Beispiel 5
$\binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = \frac{720}{6} = 120$. Beim Lotto „6 aus 49“ gibt es

$$\binom{49}{6} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 13\,983\,816$$

mögliche Tipps. Auf dem Taschenrechner heißt die Taste meist nCr: Man tippt 49, nCr, 6.

Typischer Fehler
Falsch: $\binom{10}{3} = \frac{10}{3}$ oder $\binom{10}{3} = 10 \cdot 3 = 30$
Richtig: $\binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$
$\binom{n}{k}$ ist kein Bruch, auch wenn n oben und k unten steht, und es gibt keinen Bruchstrich in der Klammer.

4.3 Das Pascalsche Dreieck

Schreibt man die Binomialkoeffizienten zeilenweise auf, $\binom{0}{0}$ in die erste Zeile, $\binom{1}{0}$ und $\binom{1}{1}$ in die zweite und so weiter, entsteht das **Pascalsche Dreieck** (Abbildung 377). Jede Zahl ist die Summe der beiden Zahlen schräg über ihr.

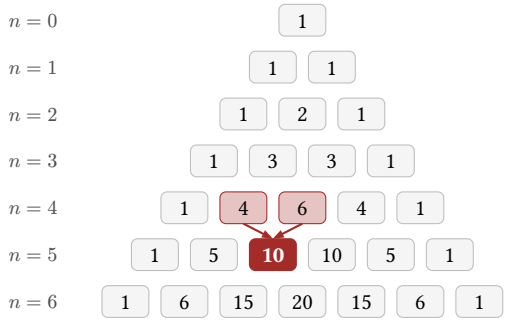


Abbildung 377: Pascalsches Dreieck bis $n = 6$. In Zeile n stehen $\binom{n}{0}$, $\binom{n}{1}$, …, $\binom{n}{n}$. Hervorgehoben: $\binom{4}{1} + \binom{4}{2} = 4 + 6 = 10 = \binom{5}{2}$.

Satz 4.2 (Eigenschaften des Binomialkoeffizienten)
Für $0 \le k < n$ gilt

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \quad \text{und} \quad \binom{n}{k} = \binom{n}{n-k}.$$

Begründung Man zählt die Pfade mit $k+1$ Treffern in einer Bernoulli-Kette der Länge $n+1$. Ist die letzte Stufe ein Treffer, müssen unter den ersten n Stufen genau k Treffer liegen: $\binom{n}{k}$ Pfade. Ist die letzte Stufe eine Niete, liegen alle $k+1$ Treffer unter den ersten n Stufen: $\binom{n}{k+1}$ Pfade. Zusammen sind es $\binom{n+1}{k+1}$.
Die zweite Gleichung: Wer festlegt, wo die k Treffer liegen, legt damit auch fest, wo die $n-k$ Nieten liegen. Deshalb ist jede Zeile des Dreiecks symmetrisch.

Das **Galtonbrett** aus M14 macht die Zahlen sichtbar (Abbildung 378): Eine Kugel fällt durch mehrere Reihen von Nägeln und wird an jedem Nagel nach links oder rechts abgelenkt. Jeder Weg ist ein Pfad einer Bernoulli-Kette mit Treffer „rechts“; im Fach k landen alle Wege mit genau k Rechts-Ablenkungen. Die Zahl der Wege zu jedem Nagel ist eine Zahl des Pascalschen Dreiecks.

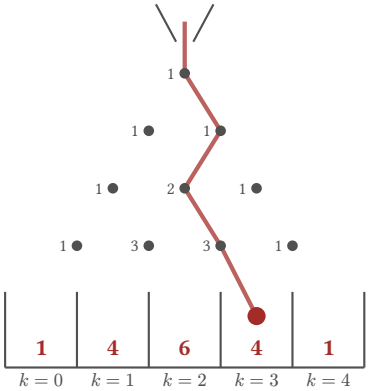


Abbildung 378: Galtonbrett mit vier Nagelreihen. Links neben jedem Nagel steht die Zahl der Wege dorthin, in den Fächern die Zahl der Wege ins Fach: 1, 4, 6, 4, 1. Eingezeichnet ist der Weg rechts, links, rechts, rechts ins Fach $k = 3$.

Bemerkung
Der Name „Binomialkoeffizient“ kommt von den binomischen Formeln (M10): Die Zahlen einer Zeile des Dreiecks sind die Vorfaktoren beim Ausmultiplizieren von $(a+b)^n$. So ist $(a+b)^2 = 1a^2 + 2ab + 1b^2$ und $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$, passend zu den Zeilen 1, 2, 1 und 1, 3, 3, 1.

Üben: **noctua** ↗ Ausprobieren: **Kombinatorik** ↗ · **Galtonbrett und Binomialverteilung** ↗

5 Binomialverteilung

G8 · Kl. 10 G9 · Kl. 11

Ein Test besteht aus 10 Fragen mit je vier Antworten, von denen genau eine richtig ist. Tim hat nicht gelernt und kreuzt bei jeder Frage zufällig an. Es sei X die Anzahl seiner richtigen Antworten. Das ist eine Bernoulli-Kette mit $n = 10$ und $p = \frac{1}{4} = 0{,}25$, und X zählt die Treffer. Wie wahrscheinlich sind genau drei richtige Antworten?

Satz 5.1 (Bernoulli-Formel)
Ist X die Anzahl der Treffer in einer Bernoulli-Kette der Länge n mit der Trefferwahrscheinlichkeit p, so gilt für $k = 0, 1, \ldots, n$:

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

Begründung Es gibt $\binom{n}{k}$ Pfade mit genau k Treffern, und jeder hat die Wahrscheinlichkeit $p^k \cdot (1-p)^{n-k}$. Nach der Summenregel addiert man diese gleichen Wahrscheinlichkeiten, also multipliziert man mit der Anzahl.

Definition 5.1 (Binomialverteilung)
Eine Zufallsgröße X mit den Werten $0, 1, \ldots, n$ und den Wahrscheinlichkeiten der Bernoulli-Formel heißt **binomialverteilt mit den Parametern n und p**. Man schreibt dafür kurz $B_{n;p}$, für Tims Test also $B_{10;0{,}25}$.

Beispiel 6
Für Tim ist

$$P(X = 3) = \binom{10}{3} \cdot 0{,}25^3 \cdot 0{,}75^7 = 120 \cdot 0{,}015625 \cdot 0{,}13348\ldots \approx 0{,}2503.$$

Die Wahrscheinlichkeit, dass er keine einzige Frage richtig hat, ist $P(X = 0) = 0{,}75^{10} \approx 0{,}0563$.

Rechnet man alle Wahrscheinlichkeiten aus, erhält man die Verteilung in Tabelle 67. Als **Histogramm** gezeichnet (Abbildung 379), steht über jedem k eine Säule der Breite 1, deren Höhe und damit auch Fläche $P(X = k)$ ist.

k	0	1	2	3	4	5	6	7	8	9	10
$P(X = k)$	0,0563	0,1877	0,2816	0,2503	0,1460	0,0584	0,0162	0,0031	0,0004	0,0000	0,0000
$P(X \le k)$	0,0563	0,2440	0,5256	0,7759	0,9219	0,9803	0,9965	0,9996	1,0000	1,0000	1,0000

Tabelle 67: Binomialverteilung mit $n = 10$ und $p = 0{,}25$, auf vier Stellen gerundet. Ab $k = 9$ sind die Wahrscheinlichkeiten so klein, dass sie gerundet 0 ergeben.

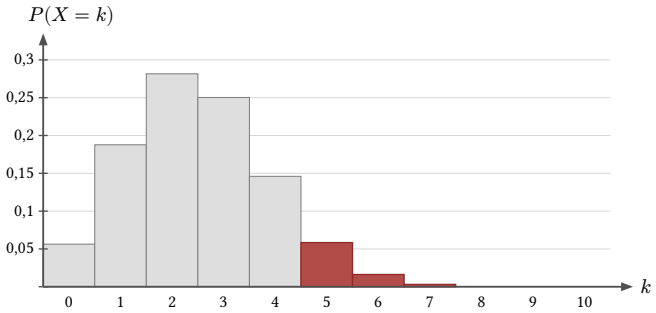


Abbildung 379: Histogramm der Binomialverteilung mit $n = 10$ und $p = 0{,}25$. Markiert sind die Säulen für „mindestens 5 richtige“.

5.1 Kumulierte Wahrscheinlichkeiten

Zum Bestehen braucht man mindestens 5 richtige Antworten. Die Wahrscheinlichkeit dafür ist die Summe der markierten Säulen in Abbildung 379: $P(X \ge 5) = P(X = 5) + P(X = 6) + \ldots + P(X = 10)$. Schneller geht es über das Gegenereignis „höchstens 4“:

$$P(X \ge 5) = 1 - P(X \le 4) \approx 1 - 0{,}9219 = 0{,}0781.$$

Durch reines Raten besteht Tim also nur mit etwa 8 % Wahrscheinlichkeit.

Wahrscheinlichkeiten der Form $P(X \leq k)$ heißen **kumuliert** (aufsummiert). Alle anderen Fragen lassen sich auf sie zurückführen; dabei muss man die Sprechweisen genau übersetzen (Tabelle 68).

In Worten	Ereignis	Berechnung mit $P(X \leq \ldots)$
höchstens k Treffer	$X \leq k$	$P(X \leq k)$
weniger als k Treffer	$X \leq k-1$	$P(X \leq k-1)$
mindestens k Treffer	$X \geq k$	$1 - P(X \leq k-1)$
mehr als k Treffer	$X \geq k+1$	$1 - P(X \leq k)$
mindestens a und höchstens b Treffer	$a \leq X \leq b$	$P(X \leq b) - P(X \leq a-1)$

Tabelle 68: Sprechweisen und ihre Übersetzung in kumulierte Wahrscheinlichkeiten.

Beispiel 7

Wie wahrscheinlich hat Tim mindestens 2 und höchstens 4 Fragen richtig? Mit Tabelle 67:

$$P(2 \leq X \leq 4) = P(X \leq 4) - P(X \leq 1) \approx 0{,}9219 - 0{,}2440 = 0{,}6779.$$

Bemerkung

Viele Taschenrechner haben Befehle für die Binomialverteilung, oft **binompdf** für $P(X = k)$ und **binomcdf** für die kumulierte Wahrscheinlichkeit $P(X \leq k)$, jeweils mit den Eingaben n, p und k. Die Namen unterscheiden sich je nach Modell; meist stehen die Befehle im Menü für Statistik oder Verteilungen.

Typischer Fehler

Falsch: $P(X \geq 5) = 1 - P(X \leq 5)$
Richtig: $P(X \geq 5) = 1 - P(X \leq 4)$

Das Gegenereignis von „mindestens 5“ ist „höchstens 4“. Die 5 selbst gehört zu „mindestens 5“ und darf nicht abgezogen werden.

5.2 Wie die Parameter das Histogramm formen

Ändert man p bei festem n, wandert der Gipfel (Abbildung 380): Bei kleinem p liegt er links, bei großem rechts. Für $p = 0{,}5$ ist das Histogramm symmetrisch, und die Histogramme zu p und $1 - p$ sind Spiegelbilder voneinander, denn Treffer und Nieten tauschen dann die Rollen.

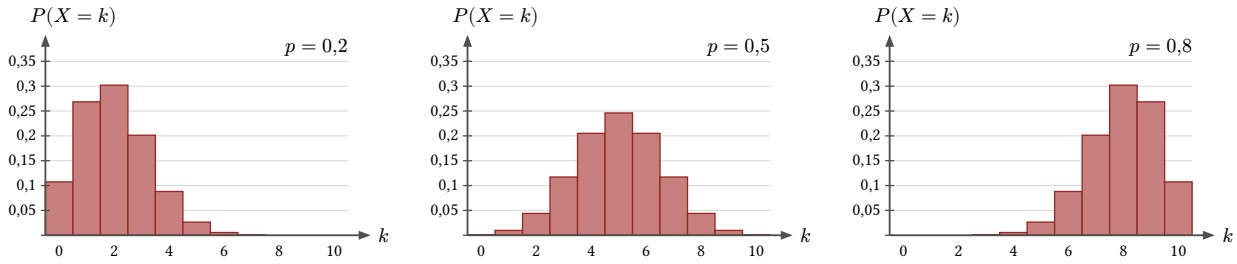


Abbildung 380: Binomialverteilungen mit $n = 10$ und verschiedenen p. Die Histogramme zu $p = 0{,}2$ und $p = 0{,}8$ sind Spiegelbilder.

Vergrößert man n bei festem p, rückt der Gipfel nach rechts, und das Histogramm wird breiter und flacher (Abbildung 381): Bei mehr Versuchen gibt es mehr mögliche Trefferzahlen, auf die sich die Gesamtwahrscheinlichkeit 1 verteilt.

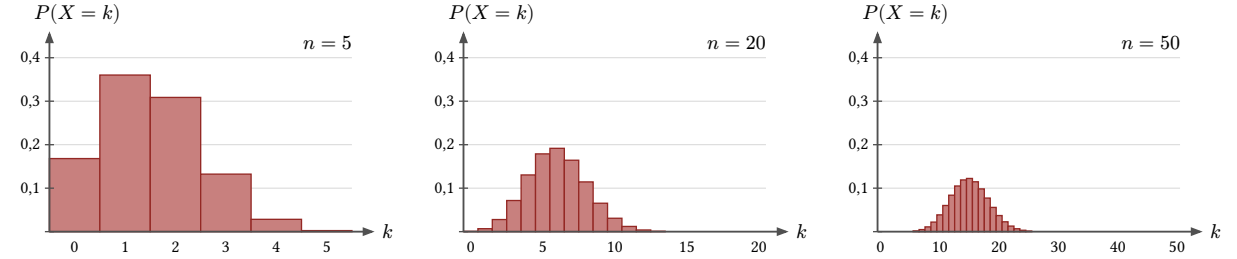


Abbildung 381: Binomialverteilungen mit $p = 0{,}3$ und verschiedenen n. Der Gipfel liegt jeweils in der Nähe von $n \cdot p$, also bei 1,5, 6 und 15.

5.3 Wie oft muss man es versuchen?

Beispiel 8

Eine Basketballspielerin trifft einen Dreier mit der Wahrscheinlichkeit 0,35. Wie viele Dreier muss sie mindestens werfen, damit sie mit mindestens 90 % Wahrscheinlichkeit mindestens einmal trifft?

Bei n Würfen ist X die Anzahl der Treffer, binomialverteilt mit n und $p = 0{,}35$ (man nimmt an, dass die Würfe unabhängig sind). Über das Gegenereignis „kein Treffer“ ist $P(X \geq 1) = 1 - P(X = 0) = 1 - 0{,}65^n$. Gesucht ist das kleinste n mit

$$1 - 0{,}65^n \geq 0{,}9, \quad \text{also} \quad 0{,}65^n \leq 0{,}1.$$

Systematisches Probieren mit dem Taschenrechner: $0{,}65^5 \approx 0{,}116$ ist noch zu groß, $0{,}65^6 \approx 0{,}075$ ist kleiner als $0{,}1$. Sie muss mindestens 6 Dreier werfen.

Mit dem Logarithmus (M18) kann man die Ungleichung auch direkt lösen: $n \geq \frac{\lg(0{,}1)}{\lg(0{,}65)} \approx 5{,}3$. Weil $\lg(0{,}65)$ negativ ist, dreht sich beim Dividieren das Ungleichheitszeichen um.

Vorgehen

a) Prüfen, ob eine Bernoulli-Kette vorliegt: zwei Ergebnisse, feste Länge n, gleichbleibendes p, unabhängige Durchgänge.
b) Treffer festlegen, n und p angeben, X als Anzahl der Treffer erklären.
c) Das gefragte Ereignis als Bedingung für X schreiben ($X = k$, $X \leq k$, $X \geq k$ …) und auf $P(X = k)$ oder $P(X \leq k)$ zurückführen.
d) Mit der Bernoulli-Formel oder dem Taschenrechner rechnen und einen Antwortsatz schreiben.

Üben: **noctua** ↗ Ausprobieren: **Galtonbrett und Binomialverteilung** ↗

6 Erwartungswert und Standardabweichung

G8 · Kl. 10 G9 · Kl. 11

Wie viele Fragen hat Tim im Mittel richtig, wenn er alle 10 rät? Die Antwort liegt nahe: Bei jeder Frage trifft er mit Wahrscheinlichkeit $\frac{1}{4}$, also im Mittel bei einem Viertel der Fragen, das sind 2,5.

Satz 6.1 (Erwartungswert der Binomialverteilung)
Ist X binomialverteilt mit den Parametern n und p, so gilt

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

Begründung Für $n = 3$ rechnet man es mit der Definition nach. Mit $q = 1 - p$ ist

$$E(X) = 0 \cdot q^3 + 1 \cdot 3pq^2 + 2 \cdot 3p^2q + 3 \cdot p^3 = 3p \cdot (q^2 + 2pq + p^2)$$
$$= 3p \cdot (q + p)^2 = 3p \cdot 1^2 = 3p.$$

Dabei half die erste binomische Formel. Allgemein trägt jeder der n Durchgänge im Mittel p Treffer bei, zusammen also $n \cdot p$; den Beweis für alle n führt man in der Oberstufe.

Bei Tim ist $\mu = 10 \cdot 0{,}25 = 2{,}5$, wirft man ein Oktaeder 400-mal, so ist die erwartete Zahl der Achten $400 \cdot \frac{1}{8} = 50$. Der Erwartungswert sagt aber nichts darüber, wie weit die tatsächliche Trefferzahl davon abweichen kann.

6.1 Streuung: Varianz und Standardabweichung

Zwei Glücksräder haben je drei gleich große Felder. Beim ersten gewinnt man 4, 5 oder 6 Punkte, beim zweiten 0, 5 oder 10 Punkte. Beide Zufallsgrößen X und Y haben den Erwartungswert 5, doch beim zweiten Rad streuen die Werte viel stärker (Abbildung 382).

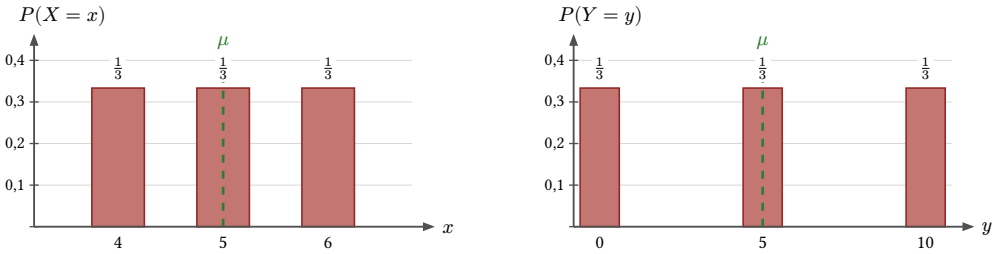


Abbildung 382: Zwei Zufallsgrößen mit demselben Erwartungswert $\mu = 5$ (gestrichelt), aber verschieden starker Streuung.

Als Maß für die Streuung nimmt man den Mittelwert der *quadrierten* Abweichungen vom Erwartungswert. Ohne das Quadrieren würden sich positive und negative Abweichungen gegenseitig aufheben.

Definition 6.1 (Varianz und Standardabweichung)
Nimmt X die Werte $x_1, \ldots, x_n$ an und ist $\mu = E(X)$, so heißt

$$V(X) = (x_1 - \mu)^2 \cdot P(X = x_1) + \ldots + (x_n - \mu)^2 \cdot P(X = x_n)$$

die **Varianz** von X und $\sigma = \sqrt{V(X)}$ (griechisch „sigma“) die **Standardabweichung**. Die Standardabweichung hat dieselbe Einheit wie X.

Beispiel 9
$V(X) = (4-5)^2 \cdot \frac{1}{3} + (5-5)^2 \cdot \frac{1}{3} + (6-5)^2 \cdot \frac{1}{3} = \frac{2}{3}$, also $\sigma_X = \sqrt{\frac{2}{3}} \approx 0{,}82$ Punkte.
$V(Y) = (0-5)^2 \cdot \frac{1}{3} + 0 + (10-5)^2 \cdot \frac{1}{3} = \frac{50}{3}$, also $\sigma_Y = \sqrt{\frac{50}{3}} \approx 4{,}08$ Punkte.
Die Standardabweichung ist beim zweiten Rad fünfmal so groß.

Satz 6.2 (Standardabweichung der Binomialverteilung)
Ist X binomialverteilt mit den Parametern n und p, so gilt

$$V(X) = n \cdot p \cdot (1 - p) \quad \text{und} \quad \sigma = \sqrt{n \cdot p \cdot (1 - p)}.$$

Begründung Für $n = 1$ nimmt X nur die Werte 0 (mit $1 - p$) und 1 (mit p) an, und es ist $\mu = p$. Dann gilt

$$V(X) = (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p = p(1-p) \cdot (p + 1 - p) = p(1-p).$$

Bei n unabhängigen Durchgängen addieren sich die Varianzen, das ergibt $n \cdot p \cdot (1-p)$. Diesen letzten Schritt beweist man in der Oberstufe.

Beispiel 10

Für Tim ist $\sigma = \sqrt{10 \cdot 0{,}25 \cdot 0{,}75} = \sqrt{1{,}875} \approx 1{,}37$. Für die Zahl der Achten bei 400 Oktaederwürfen ist

$$\sigma = \sqrt{400 \cdot \frac{1}{8} \cdot \frac{7}{8}} = \sqrt{43{,}75} \approx 6{,}61.$$

Typischer Fehler

Falsch: $\sigma = n \cdot p \cdot (1-p) = 10 \cdot 0{,}25 \cdot 0{,}75 = 1{,}875$

Richtig: $\sigma = \sqrt{1{,}875} \approx 1{,}37$

$np(1-p)$ ist die Varianz. Die Standardabweichung ist ihre Wurzel.

6.2 Was σ über die Verteilung sagt

Abbildung 383 zeigt die Binomialverteilung mit $n = 50$ und $p = 0{,}4$. Ihr Erwartungswert ist $\mu = 20$, ihre Standardabweichung $\sigma = \sqrt{50 \cdot 0{,}4 \cdot 0{,}6} = \sqrt{12} \approx 3{,}46$. Im Bereich von $\mu - \sigma \approx 16{,}5$ bis $\mu + \sigma \approx 23{,}5$ liegen die Trefferzahlen 17 bis 23. Mit dem Taschenrechner ergibt sich

$$P(17 \le X \le 23) = P(X \le 23) - P(X \le 16) \approx 0{,}688.$$

Im doppelt so breiten Bereich von $\mu - 2\sigma \approx 13{,}1$ bis $\mu + 2\sigma \approx 26{,}9$ ist $P(14 \le X \le 26) \approx 0{,}941$.

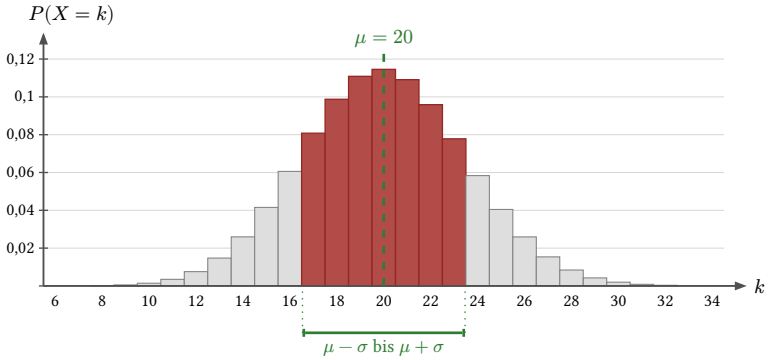


Abbildung 383: Binomialverteilung mit $n = 50$ und $p = 0{,}4$ (gezeichnet von $k = 6$ bis $k = 34$, außerhalb sind die Säulen kaum sichtbar). Markiert sind die Trefferzahlen zwischen $\mu - \sigma$ und $\mu + \sigma$; zusammen haben sie die Wahrscheinlichkeit 0,688.

Merksatz

Faustregel: Ist $\sigma > 3$, so liegt die Trefferzahl einer Binomialverteilung mit etwa 68 % Wahrscheinlichkeit zwischen $\mu - \sigma$ und $\mu + \sigma$ und mit etwa 95 % Wahrscheinlichkeit zwischen $\mu - 2\sigma$ und $\mu + 2\sigma$. Genaue Werte liefert der Taschenrechner.

Beispiel 11

Wie viele Preise braucht die Losbude? Beim Schulfest werden 200 Lose verkauft; jedes gewinnt mit der Wahrscheinlichkeit 0,15. Die Zahl X der Gewinnlose ist binomialverteilt mit $\mu = 200 \cdot 0{,}15 = 30$ und $\sigma = \sqrt{200 \cdot 0{,}15 \cdot 0{,}85} = \sqrt{25{,}5} \approx 5{,}05$. Nach der Faustregel reichen $\mu + 2\sigma \approx 40$ Preise fast immer. Genauer: $P(X \le 40) \approx 0{,}978$. Mit 40 Preisen geht die Losbude also nur in etwa 2 % der Fälle leer aus; mit 35 Preisen wäre es wegen $P(X \le 35) \approx 0{,}861$ schon in etwa jedem siebten Fall so.

Üben: **noctua** ↗ Ausprobieren: **Erwartungswert und Varianz** ↗ · **Galtonbrett und Binomialverteilung** ↗