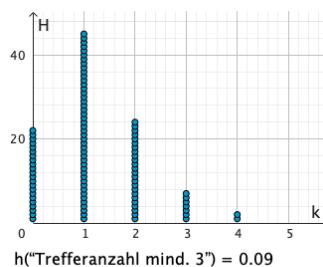
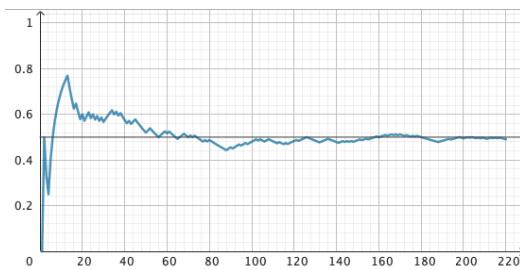


Sinnstiftender Einsatz von GeoGebra im Stochastikunterricht der Sek II

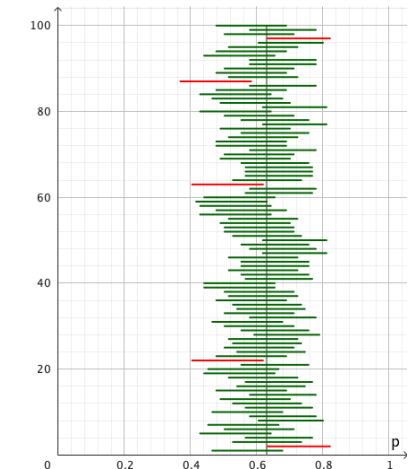
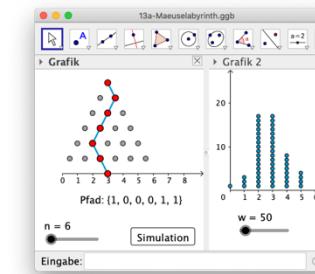
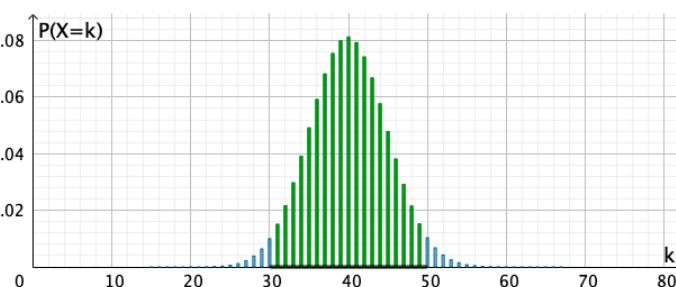
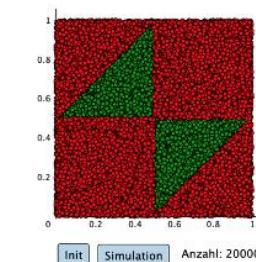
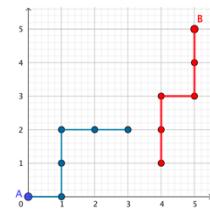
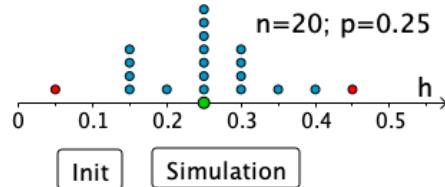
Einsatzmöglichkeiten von GeoGebra im Stochastikunterricht in der Sek II

Reimund Vehling

12.03.2025



Wert	Relative Häufigkeit
0	0.22
1	0.45
2	0.24
3	0.07
4	0.02



Sinnstiftender Einsatz von GeoGebra im Stochastikunterricht der Sek II

Einsatzmöglichkeiten von GeoGebra im Stochastikunterricht in der Sek II

Reimund Vehling

12.03.2025

Das Folgende habe ich angekündigt, mal sehen ...

- Es werden nur Grundkenntnisse in GeoGebra vorausgesetzt.
- Anhand von tragfähigen Beispielen wird der Einsatz von Simulationen, Berechnungen und grafischen Darstellungen zu Problemstellungen aus der Wahrscheinlichkeitsrechnung und der beurteilenden Statistik vorgestellt.
- Dabei wird besonderer Wert auf eine einfache Umsetzung gelegt.
- Ziel ist es, Lehrkräfte in die Lage zu versetzen, das Programm selbstständig und reflektiert im Unterricht einzusetzen.

Bevor es losgeht – eine Einordnung

Für einen gelingenden Stochastikunterricht stehen der Aufbau und die Förderung **stochastischer Grundvorstellungen** im Vordergrund.

GeoGebra hat dabei zunächst einmal eine dienende Funktion und sollte **nicht als Selbstzweck** eingesetzt werden, nur weil man damit schnell viele Simulationen und schöne Grafiken erstellen kann.

GeoGebra kann aber an bestimmten Stellen entscheidend den **Lernprozess unterstützen, um**

- Zufallsschwankungen zu untersuchen, anstatt sie wegzuwünschen,
- zu einer vertiefenden Begriffsbildung zu gelangen.
- zu erkennen, dass Wahrscheinlichkeitsberechnungen immer auf Modellannahmen basieren,
- Modell und Realität zu trennen.

Das reale Galtonbrett – mehr als eine Simulation



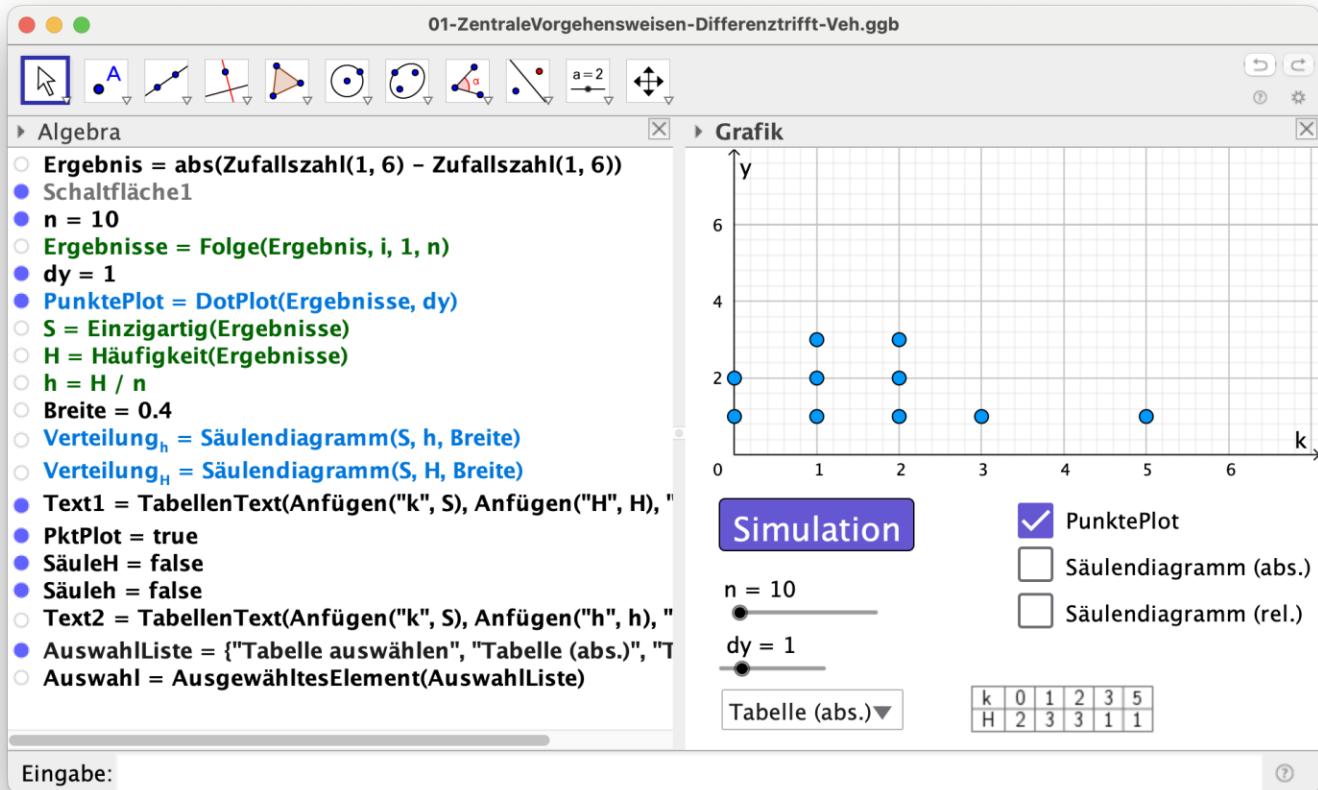
Vorzüge eines realen Galton-Brettes:

- mehrere Zufallsexperimente können betrachtet werden
- stochastisch wichtige Zusammenhänge werden erfahrbar und einsichtig: Vergrößerung, Verfeinerung, Stufung, Reduzierung
- einfache Durchführung und gut einsehbar
- Häufigkeitsverteilung sofort sichtbar
- schiefstellen, dann $p \neq 0,5$

Inhalte

- Zentrale Vorgehensweisen
- Einstiegsbeispiele: Differenz trifft; Geburtstagsproblem
- Untersuchung von Zufallsschwankungen und das $\frac{1}{\sqrt{n}}$ - Gesetz
- Binomialverteilung, zwei mögliche Einstiege
- Binomialverteilung, mit und ohne Simulation
- Erwartungswert und Varianz bei der Binomialverteilung
- Prognoseintervalle und das Konzept des Zweifelns
- Konfidenzintervalle – mit Simulationen zur Sicherheitswahrscheinlichkeit
- Normalverteilung – mit und ohne Simulationen
- Approximation von π ; der zerbrochene Stab und ein Dreieck

Zentrale Befehle



Folgebefehl – immer wenn es mehrere Simulationen gibt:

Folge(<Ausdruck>, <Variable>, <von>, <bis>)

Bsp. (ohne Zufall):

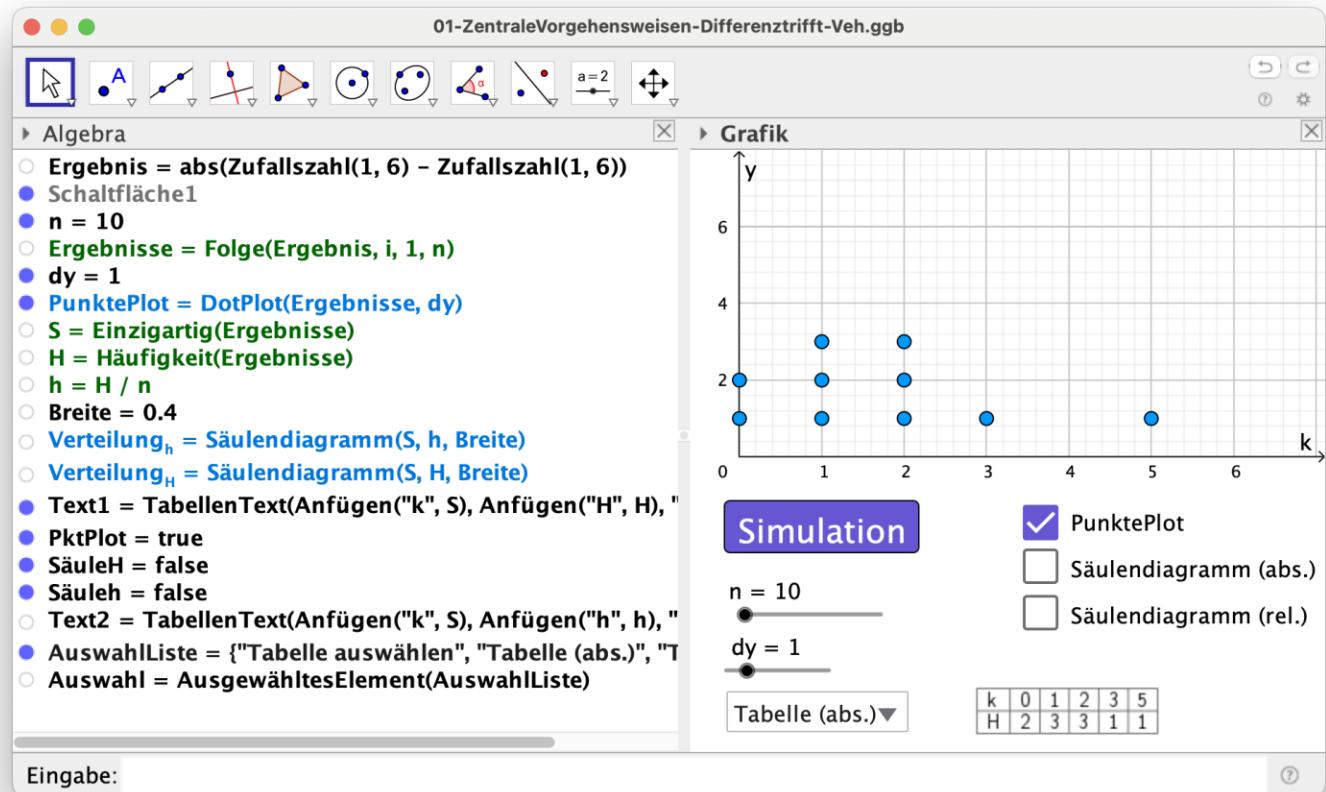
Eingabe: Folge(k^2, k, 1, 4)

Ausgabe in eine Liste: l1={1,4,9,16}

Ausgabe der Daten

- dotPlot(Daten, Skalierungsfaktor)
- Säulendiagramm(Daten, Häufigkeit, Breite)
 - Daten=Einzigartig(<Liste>)
 - Häufigkeit=Häufigkeit(<Liste>)

Zentrale Befehle



Zentrale Befehle

Schaltfläche

Beschriftung

Simulation

GeoGebra Skript

AktualisiereKonstruktion()

Grundeinstellungen	Text	Farbe	Position	Erweitert
Skripting				

Häufigkeitstabelle(Ergebnisse) Häufigkeitstabelle(Ergebnisse,1/n)

Wert	Häufigkeit	Wert	Relative Häufigkeit
0	2	0	0.2
1	4	1	0.4
2	1	2	0.1
4	1	4	0.1
5	2	5	0.2

Zentrale Befehle



GeoGebra

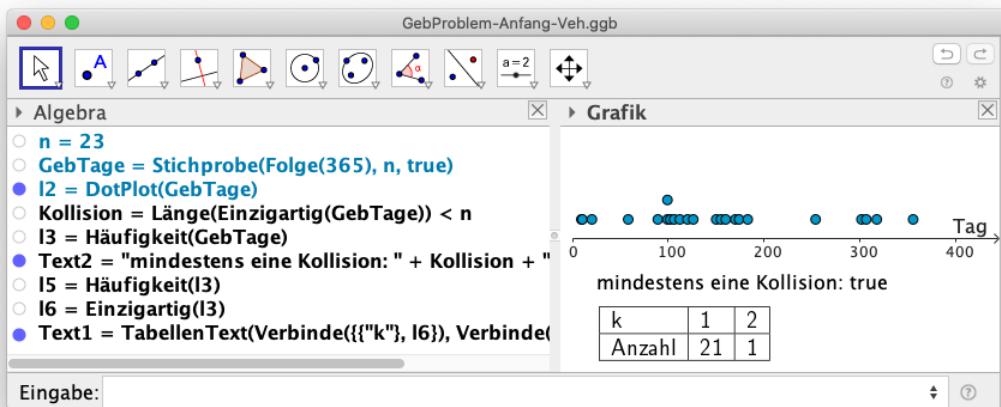
Wichtige GeoGebra-Befehle aus dem Bereich der Stochastik

Befehl	Erläuterungen
=nCr(n, k)	Berechnung von $\binom{n}{k}$
=Binomial(n, p) bzw. Binomial(n, p, false)	Säulendiagramm einer Binomialverteilung B_{np}
=Binomial(n, p, true)	Diagramm der kumulierten Binomialverteilung F_{np}
=Binomial(n, p, k, false)	$P(X = k)$
=Binomial(n, p, k, true)	$P(X \leq k)$
=Binomial(n, p, a .. b)	$P(a \leq X \leq b)$
=InversBinomial(n, P, a)	kleinste ganze Zahl k mit $P(X \leq k) \geq a$
=Normal(μ , σ , x, false)	Dichte $\varphi_{\mu, \sigma}$ der Normalverteilung (Gauß-Glocke)
=Normal(μ , σ , x, true)	zu $\varphi_{\mu, \sigma}$ gehörige Verteilungsfunktion $\Phi_{\mu, \sigma}$
=InversNormal(μ , σ , p)	Stelle x mit $P(X \leq x) = p$.
=Zufallszahl(a, b)	ganzahlige Zufallszahl x mit $a \leq x \leq b$ Münze: =Zufallszahl(0, 1); Würfel: =Zufallszahl(1, 6)
=ZufallszahlBinomialverteilt(n, p)	binomialverteilte Zufallszahl
=ZufallszahlBinomialverteilt(n, p) />	relative Trefferhäufigkeit bei n Bernoulli-Versuchen
=Folge(ZufallszahlBinomialverteilt(n, p), 1, k)	k binomialverteilte Zufallszahlen (Liste)
=ZufallszahlDiskret(L, L _j)	Zufallszahl: Ergebnisliste L, Wahrscheinlichkeitsliste L _j =ZufallszahlDiskret({1, 2, 3}, {0.5, 0.2, 0.3}): Glücksspiel mit drei Feldern und zugehörigen Wahrscheinlichkeiten
=ZufallszahlGleichverteilt(a, b)	gleichverteilte „dezimale“ Zufallszahl x mit $a \leq X \leq b$
=ZufallszahlGleichverteilt(a, b, n)	gleichverteilte „dezimale“ Zufallszahlen x mit $a \leq X \leq b$ (Liste)
=ZufallszahlNormalverteilt(μ , σ)	$\varphi_{\mu, \sigma}$ -normalverteilte Zufallszahl
=Folge(Zufallszahl(a, b)+Zufallszahl1(a, b), 1, 1, n)	n Summen zweier (ganzzahliger) Zufallszahlen zwischen a und b (Liste); für a = 1, b = 6: n Augensummen zweier Würfe!
=Summe(Folge(Zufallszahl(0, 1)==1, 1, 1, 10))	ersetzt den Befehl =ZufallszahlBinomial(10, 0.5)
=Stichprobe(L, n, true/false)	Aus der Liste L werden n Elemente mit (true)/ohne (false) Zurücklegen ausgewählt. Spezialfall: =Stichprobe({1, 2, 3}, 3, false) liefert eine Permutation der Zahlen 1, 2 und 3
=Mittelwert(L)	arithmetisches Mittel der Zahlenliste L
=stdevp(L)	Standardabweichung einer Grundgesamtheit L: Liste, „Division durch n“
=stdev(L)	Standardabweichung einer Stichprobe L: Liste, „Division durch n - 1“
=Säulendiagramm(L, b)	Säulendiagramm zu einer Urliste L; b: Balkenbreite
=Säulendiagramm(L _i , l _i , b)	Säulendiagramm zu einer Häufigkeitsverteilung mit L _i : Datenliste, l _i : Häufigkeitsliste; b: Balkenbreite
K-Klassen(l, n)	Die Daten aus L werden eingeteilt (l) in n Klassen, (ii) in Klassen der Breite b, Start bei a.
K-Klassen(l, a .. b)	Histogramm der Daten L gemäß Klasseneinteilung (s.o.)
=Histogramme(K, L)	
=Boxplot(y-Position, y-Breite, L, true/false)	Boxplot zur Datenliste L auf vorgegebener y-Position mit vorgegebener y Breite, mit/ohne Ausreißer
L=A1:A100	Tabellenkalkulationsbereiche werden zu Listen

Das Geburtstagsproblem – ein „Muss“ im Stochastikunterricht

Erste Idee einer Simulation – ein Dreizeiler

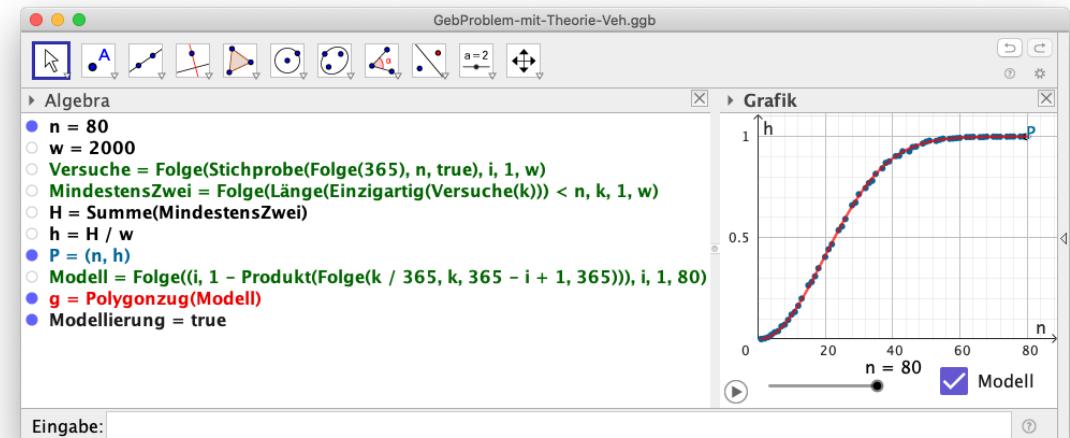
Nur zur Vorsicht: Zu Beginn Vermutungen einfordern und im Kurs (in Kursen) Daten sammeln und auswerten.



- Folge(365): Liste mit den natürlichen Zahlen von 1 bis 365
- Stichprobe(<Liste>, <Größe>, <Wiederholung möglich true|false>)
- DotPlot(<Liste>): Die Listenelemente werden als Punkte dargestellt. Bei Gleichheit, werden die zugehörigen Punkte übereinander gezeichnet.

[Geburtstagsproblem - Anfang](#)

Simulation als Animation - passt sehr gut zur Theorie



A: keine Kollision

$$\mathbb{P}(\bar{A}) = \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n}$$

[Geburtstagsproblem - Ende](#)

Das Geburtstagsproblem – Tiefbohrung

Modellierung: alle Tage sind gleichwahrscheinlich. Ist das gerechtfertigt?

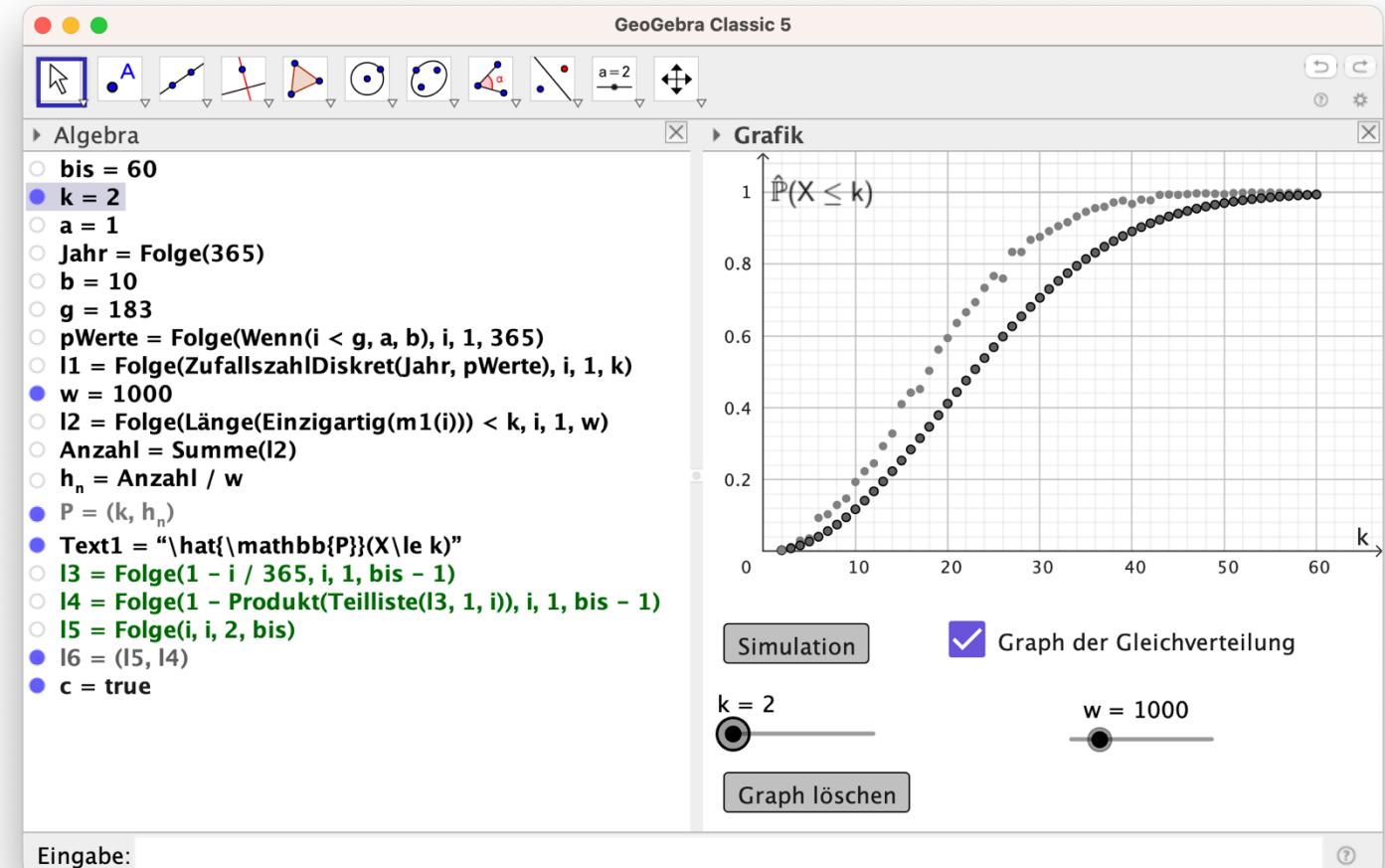
Was passiert, wenn der Ansatz „alle Tage im Jahr sind gleichwahrscheinlich“ entfällt?

Bsp.: Die ersten 182 Tage haben eine 10-mal kleinere „Geburtstags-Wahrscheinlichkeit“ als die restlichen Tage.

Ergebnis: Der zugehörige Graph liegt über dem Graphen mit dem Modell mit der Annahme der Gleichverteilung der Tage.

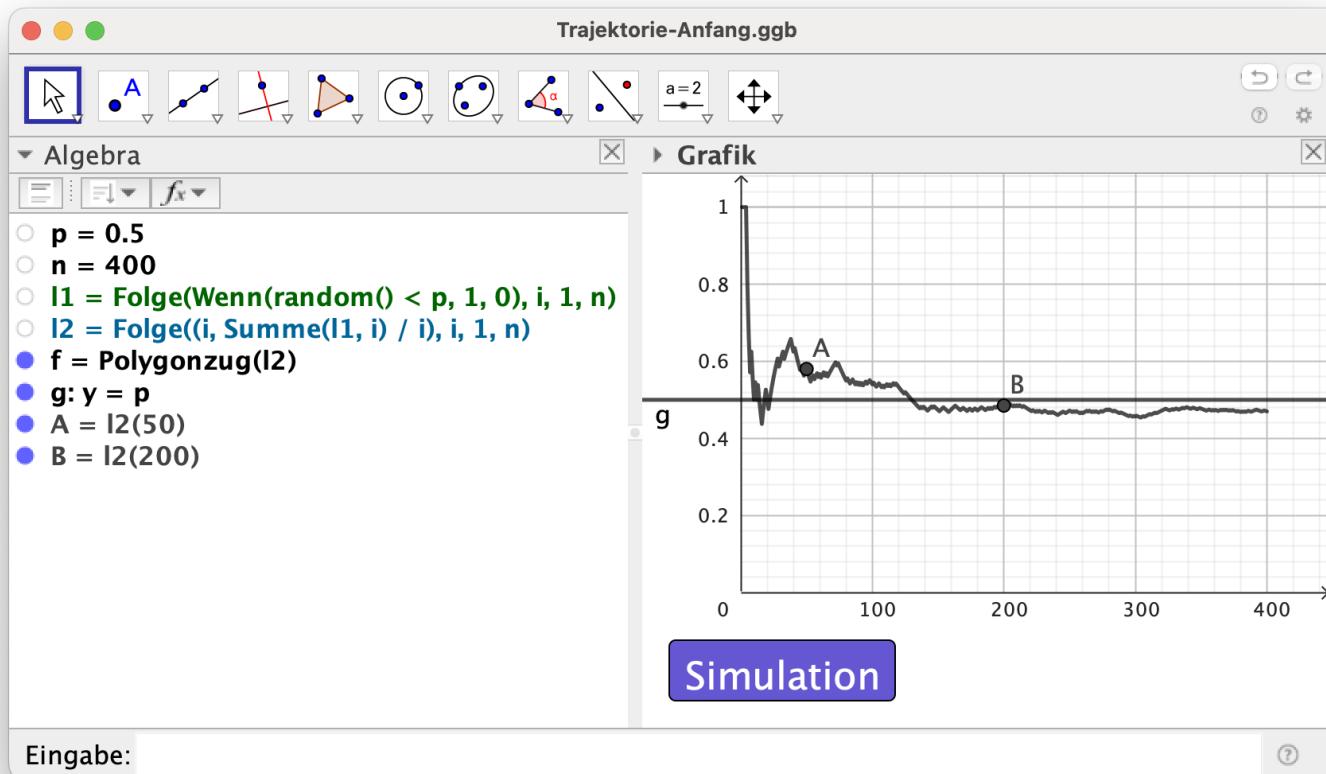
Forschungsfrage: Ist das immer so?

[Geburtstagsproblem - nicht gleichverteilt](#)



Der Münzwurf – und das $1/\sqrt{n}$ - Gesetz

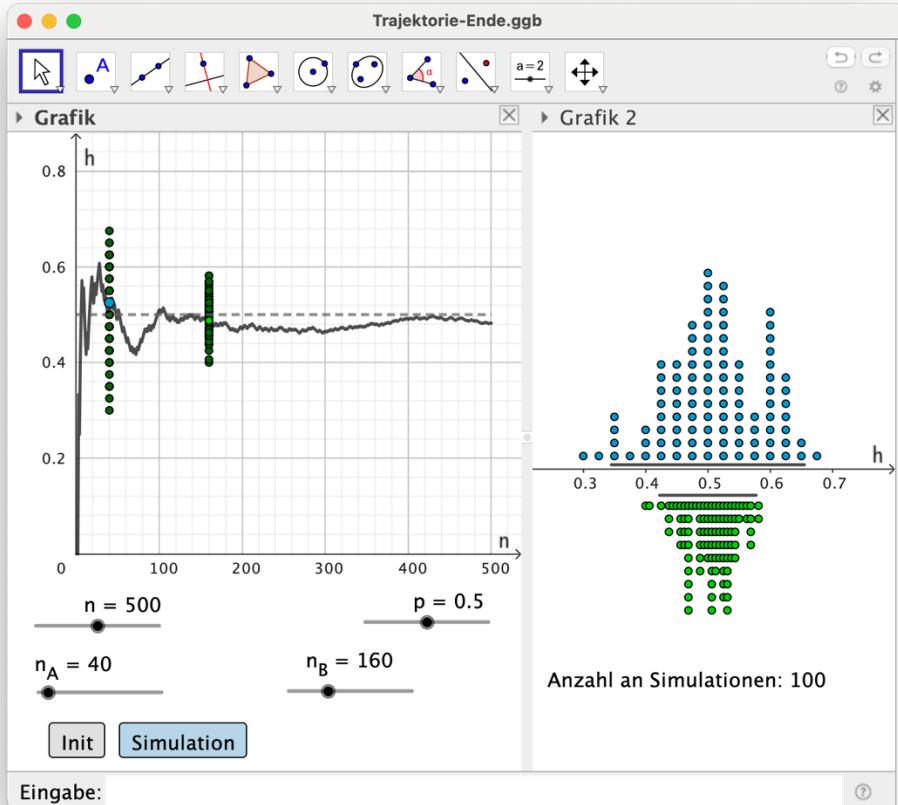
Eine einfache Simulation des Münzwurfs



[Trajektorie - Anfang](#)

Der Münzwurf – und das $1/\sqrt{n}$ - Gesetz

Trajektorie und zwei Punkteplots



```

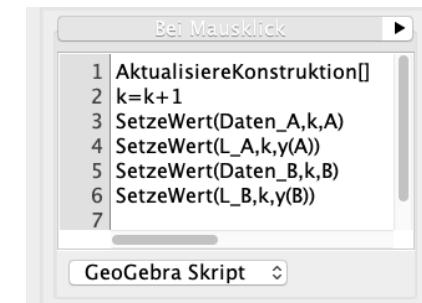
• p = 0.5
• n = 500
• Zufallsfolge = Folge(Wenn(random() < p, 1, 0), i, 1, n)
• relH = Folge((i, Summe(Zufallsfolge, i) / i), i, 1, n)
• f = Polygonzug(relH)
• k = 3
• LA = {0.58, 0.4, 0.4}
• LB = {0.51, 0.45, 0.49}
• dy = 0.1
• PunktePlotB: Spiegle(DotPlot(LB, dy), y = -0.9 dy)
• DatenA = {(40, 0.58), (40, 0.4), (40, 0.4)}
• DatenB = {(160, 0.51), (160, 0.45), (160, 0.49)}
• nA = 40
• nB = 160
• A = relH(nA)
• B = relH(nB)
• g = Strecke((0, p), (n, p))
• Text1 = "Anzahl an Simulationen: " + k + ""
• I1A = p - 1.96sqrt(p (1 - p) / nA)
• I1B = p + 1.96sqrt(p (1 - p) / nB)
• Q1 = (I1A, 0)
• Q2 = (I1B, 0)
• I2B = p + 1.96sqrt(p (1 - p) / nB)
• I2A = p - 1.96sqrt(p (1 - p) / nA)
• R1 = (I2A, 0)
• R2 = (I2B, 0)
• PunktePlotA = DotPlot(LA, dy)
• i = Strecke((x(Q1), dy / 3), (x(Q2), dy / 3))
• P1 = (p, dy / 3)
• P2 = (p, (-dy) / 3)
• h = Strecke((x(R1), -2 dy), (x(R2), -2 dy))

```

Init

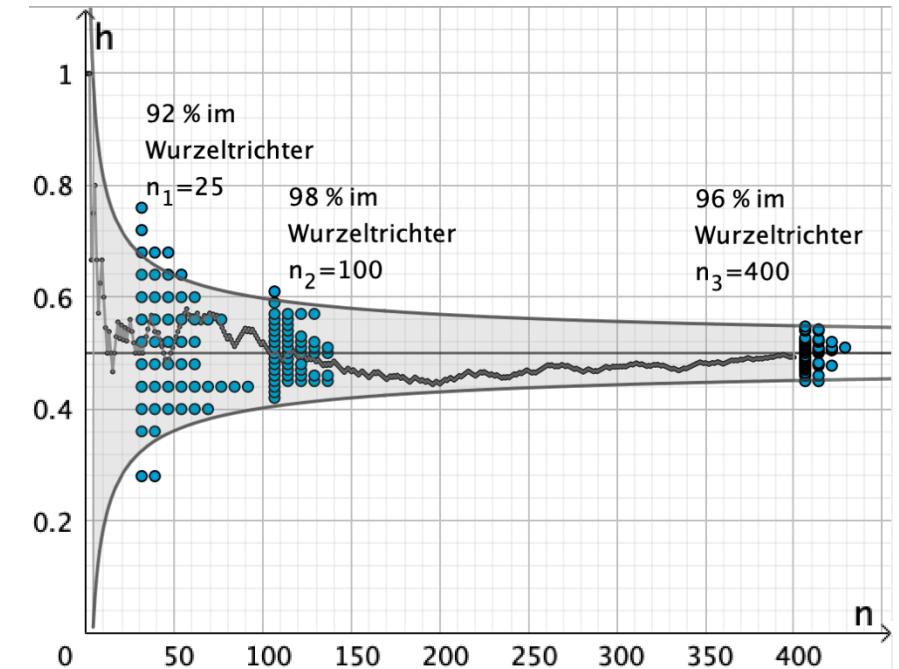
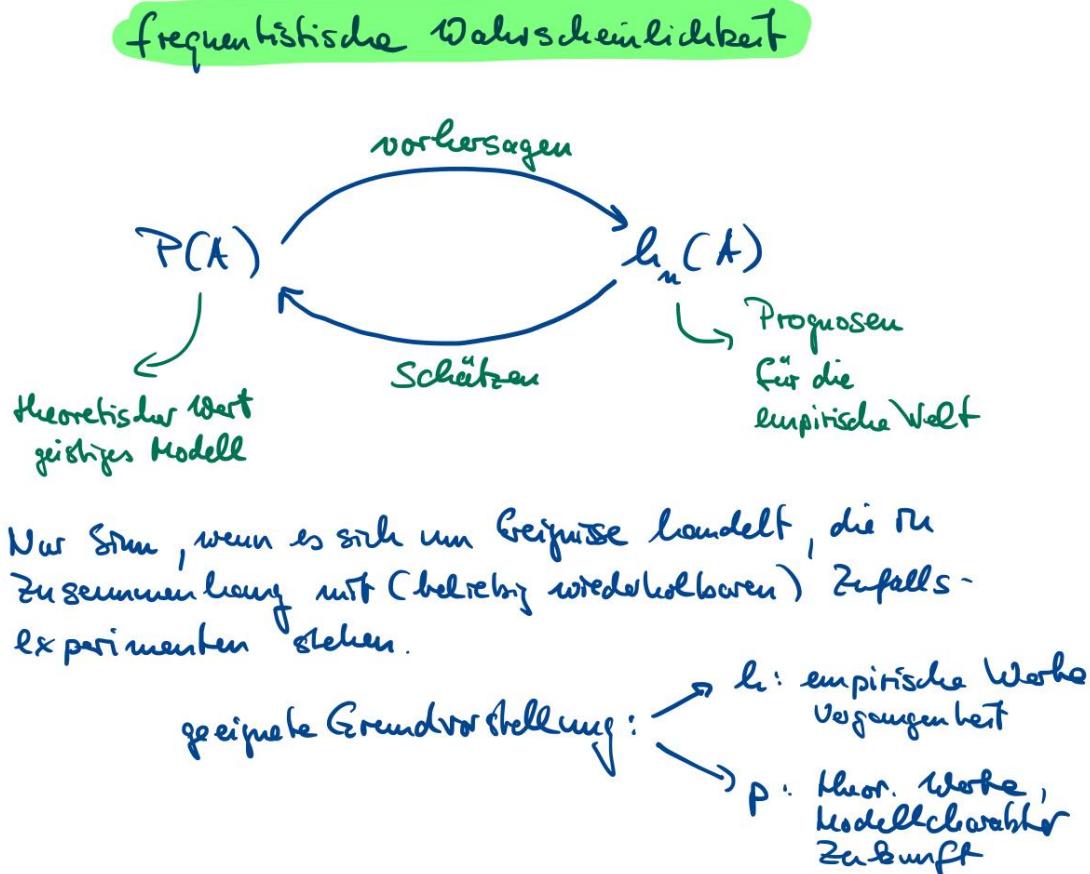


Simulation

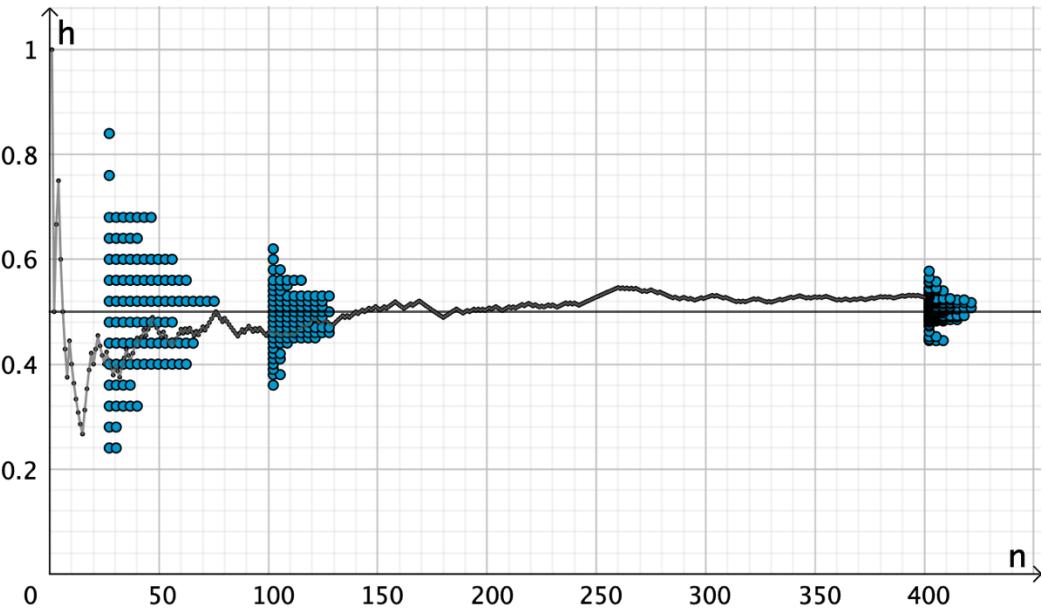


[Trajektorie - Ende](#)

Modell und Realität – und das $1/\sqrt{n}$ –Gesetz



Wahrscheinlichkeit und relative Häufigkeit



Zusatz von mir:

Häufigkeiten zeigen in die Vergangenheit,
Wahrscheinlichkeiten in die Zukunft.

... und das sagt Wolfgang Riemer:

(a) Wahrscheinlichkeiten sagen relative Häufigkeiten voraus. Sie sind gut **festgelegt**, wenn die (durch Zufallseinflüsse schwankenden) relativen Häufigkeiten gleichmäßig um die Wahrscheinlichkeit streuen, also etwa gleich oft über wie unter der Wahrscheinlichkeit liegen.

(b) Wahrscheinlichkeiten „leben“ - vor der Datenerhebung als **Modelle** der Wirklichkeit „im Kopf“. Sie spiegeln Symmetrien genau wider. **Relative Häufigkeiten** entstehen erst nach der Datenerhebung, sie „leben“ in der „**Realitätsebene**“. Sie spiegeln wegen der Zufallsschwankungen Symmetrien nur näherungsweise wider.

Wahrscheinlichkeit entsteht, wenn aus Erfahrung Erwartung wird

Zitat von Wolfgang Riemer

Bei Zufallsexperimenten kann man einzelne Ergebnisse nicht vorhersagen, man kann ihnen aber Wahrscheinlichkeiten zuordnen, die zusammen 100% ergeben.

Die Wahrscheinlichkeiten sind gut **gewählt**, wenn

- man Symmetrien beachtet,
- die relativen Häufigkeiten in der Nähe der Wahrscheinlichkeit liegen,
- die relativen Häufigkeiten bei Versuchswiederholungen zufällig um die Wahrscheinlichkeiten pendeln, mal etwas darüber liegen, mal etwas darunter.
- Die Wahrscheinlichkeiten drücken aus, wo man zukünftig in etwa die relativen Häufigkeiten erwartet.
- Je mehr Erfahrungen zugrunde liegen, desto größer das Vertrauen in die aufgestellten Wahrscheinlichkeiten.

Realitätsebene	Modellebene
$h_1 + \dots + h_n = 100\%$	$p_1 + \dots + p_n = 100\%$
relative Häufigkeiten	Wahrscheinlichkeiten
leben im Becher	leben im Kopf
schauen zurück	schauen nach vorne
schwanken zufällig	werden festgelegt, bezweifelt, verbessert
Im Fall von Teilsymmetrien	
ungefähr gleich	genau gleich
Mittelwert $\bar{x} = x_1 \cdot h_1 + \dots + x_n \cdot h_n$	Erwartungswert $\mu = x_1 \cdot p_1 + \dots + x_n \cdot p_n$
Standardabw. (empirisch) $s = \sqrt{\sum (x_i - \bar{x})^2 h_i}$	Standardabweichung $\sigma = \sqrt{\sum (x_i - \mu)^2 p_i}$

$$p_1 + \dots + p_n = 1$$

So finde ich es
(noch) besser - ist
aber Ansichtssache.

Abb. 2-4: Regelhefeintrag mit präziser Abgrenzung der Realitäts- von der Modellebene. Man beachte, dass die Wahrscheinlichkeiten als Modelle **gewählt** werden, so, wie das im Unterricht erlebt wurde.

Idee: W. Riemer, z. B. (leicht verändert) enthalten in: Riemer, W. (2023): Statistik unterrichten. Kallmeyer

Ein MC-Test

Ein TV-Sender möchte sich auch bei den beliebten Quiz-Shows etablieren. Da sich die Show durch ein hohes Bildungsniveau auszeichnen soll, müssen sich die Kandidaten über die Beantwortung eines Fragebogens qualifizieren. Folgende Fragen sollen beantwortet werden:

1. Wer legte 1954 den Grundstein für den Restaurant-Konzern McDonalds?

A	Ronald McDonald	B	Rick van Mac
C	Ray Kroc	D	Archie Macmillan

2. Welcher Stoff bringt ein Glühwürmchen zum Leuchten?

A	Methan	B	Luziferin
C	Osmarol	D	Glyzerin

Ein MC-Test – die Fragen 3 bis 5

3. In welchem Land erschienen 1840 die ersten Briefmarken der Welt?

A	Deutschland	B	Spanien
C	Frankreich	D	Großbritannien

4. Wer ist die Zwillingschwester des Apollon?

A	Artemis	B	Aphrodite
C	Alkmene	D	Leto

5. Welcher Philosoph sagte: "Das Schicksal mischt die Karten und wir spielen"?

A	Platon	B	Descartes
C	Sartre	D	Schopenhauer

Ein MC-Test (Lösung)

1. Wer legte 1954 den Grundstein für den Restaurant-Konzern McDonalds?

A	Ronald McDonald	B	Rick van Mac
C	Ray Kroc	D	Archie Macmillan

3. In welchem Land erschienen 1840 die ersten Briefmarken der Welt?

A	Deutschland	B	Spanien
C	Frankreich	D	Großbritannien

5. Welcher Philosoph sagte: "Das Schicksal mischt die Karten und wir spielen"?

A	Platon	B	Descartes
C	Satre	D	Schopenhauer

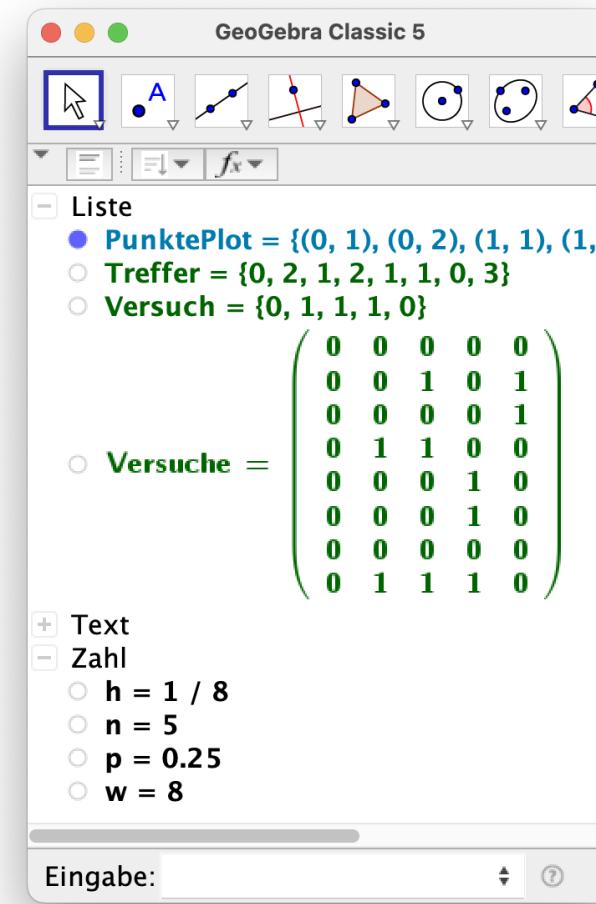
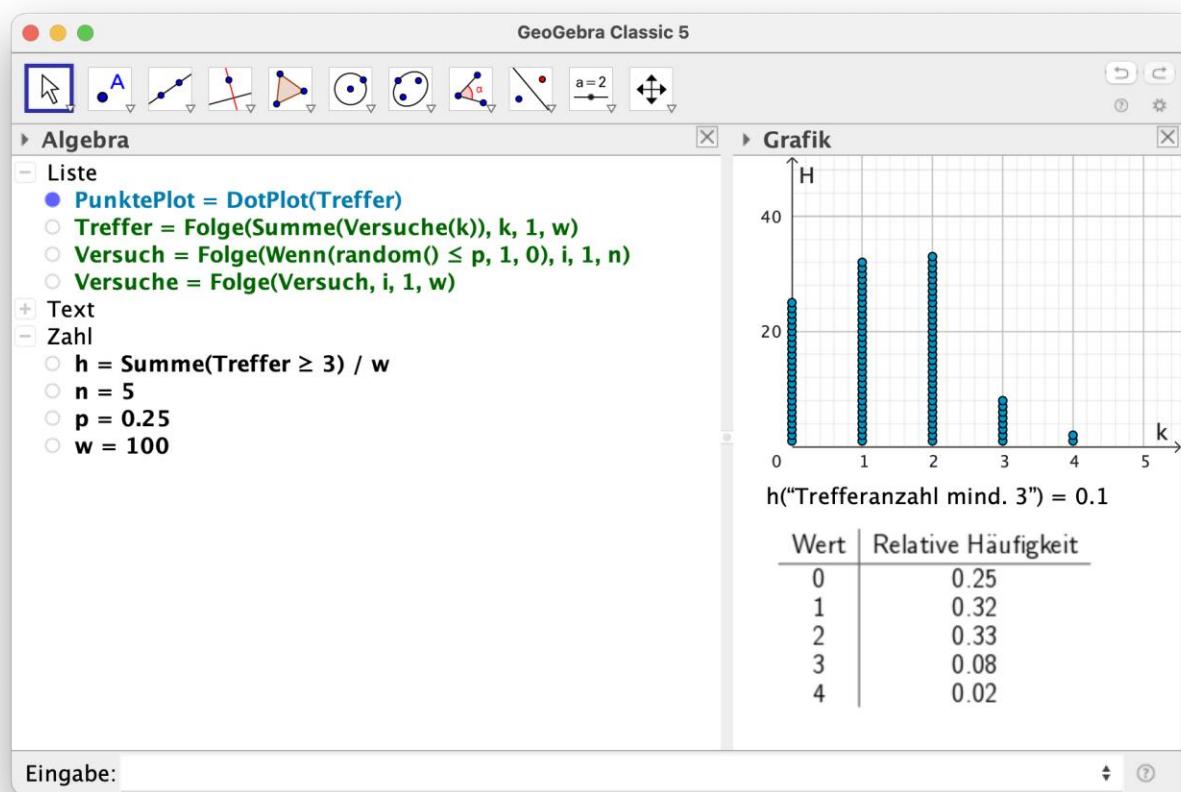
2. Welcher Stoff bringt ein Glühwürmchen zum Leuchten?

A	Methan	B	Luziferin
C	Osmarol	D	Glyzerin

4. Wer ist die Zwillingsschwester des Apollon?

A	Artemis	B	Aphrodite
C	Alkmene	D	Leto

Mit Simulationen zum Modell der Binomialverteilung



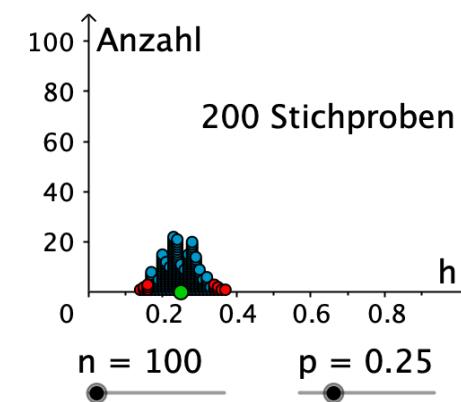
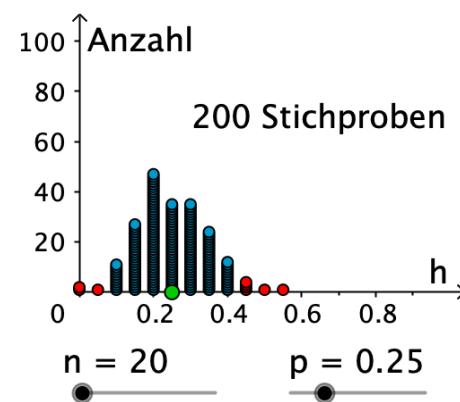
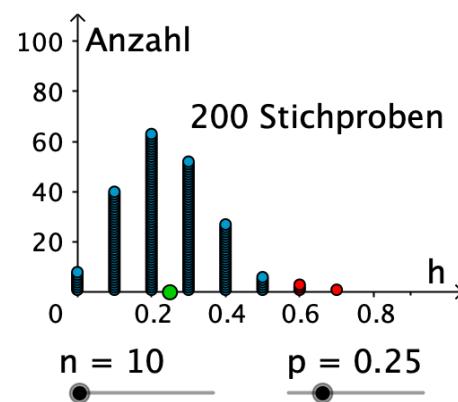
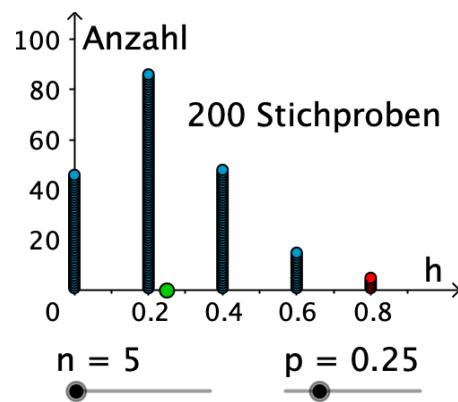
[MC-Simulation](#)

Ein MC-Test – Vertiefung

Ihr könnt zwischen einem Multiple-Choice-Test mit 5, 10, 20 oder 100 Fragen auswählen.

Bei jedem dieser Tests seid ihr qualifiziert, wenn mindestens 60% der betreffenden Fragen gelöst werden. Ihr könnt nur raten ($p=0,25$).

- Welchen der vier Tests würdet ihr auswählen?
- Begründet eure Entscheidung, ohne zu rechnen.



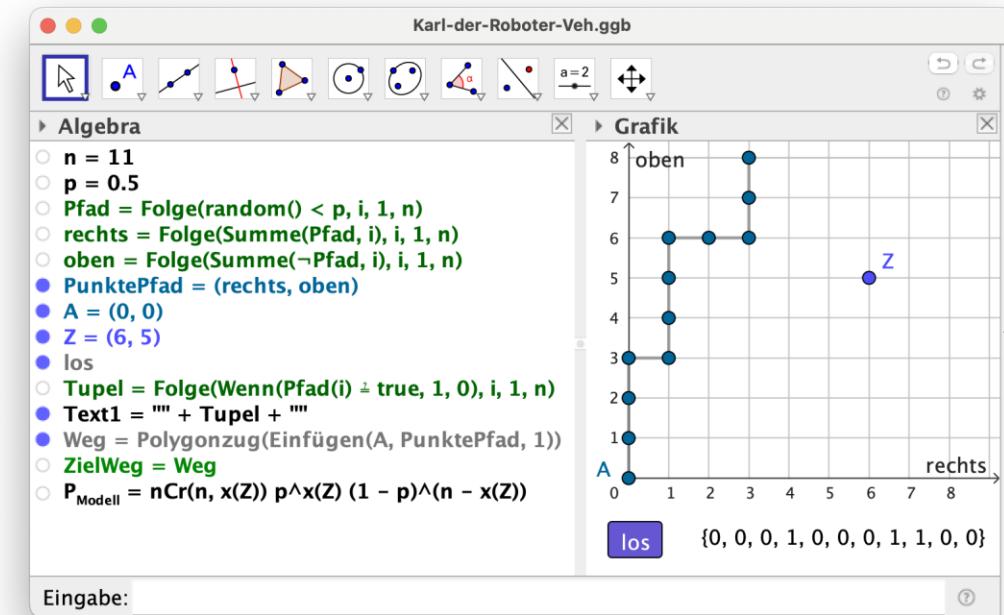
Und wieder grüßt das $1/\sqrt{n}$ - Gesetz, ungewöhnlich verpackt.

[MC-Simulation; n variabel](#)

Binomialverteilung und die Suche nach dem Ziel

Roboter Karl beginnt an Punkt A und hat genau **11 Schritte** zur Verfügung, um das Ziel Z zu erreichen. Er kann dabei entweder nach **rechts** oder nach **oben** gehen. Diese Entscheidungen trifft er rein zufällig und unabhängig von vorherigen Schritten, wobei jede Möglichkeit die gleiche Wahrscheinlichkeit hat.

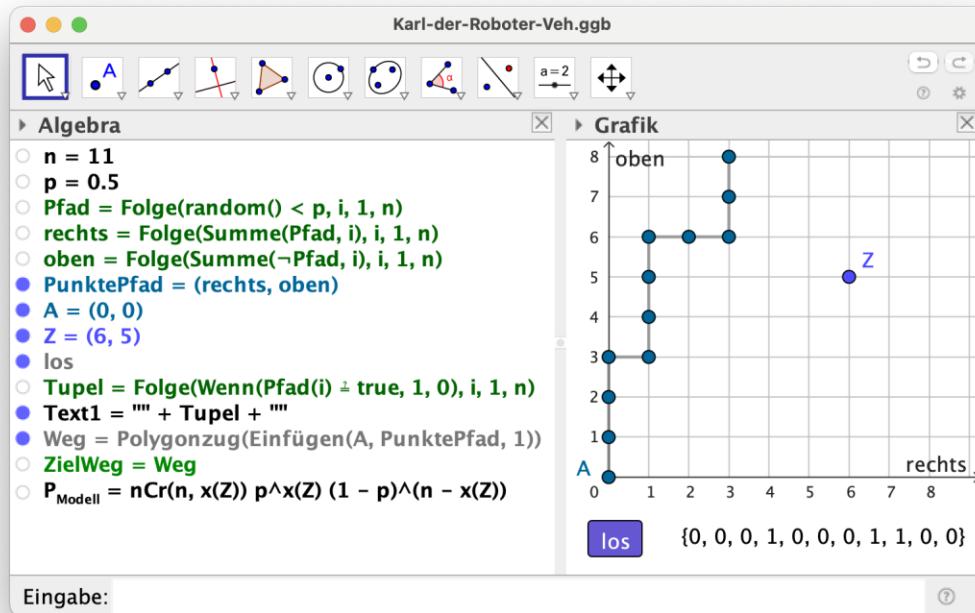
Führen Sie eine Simulation durch, um die **Wahrscheinlichkeit abzuschätzen**, dass Roboter Karl das Ziel in genau 11 Schritten erreicht.



[Karl der Roboter](#)

Binomialverteilung und die Suche nach dem Ziel

- $n = 11$
- $p = 0.5$
- Pfad = {true, true, true, false, false, true, rechts = {1, 2, 3, 3, 3, 4, 4, 5, 5, 6, 6}}
- oben = {0, 0, 0, 1, 2, 2, 3, 3, 4, 4, 5}
- PunktePfad = {(1, 0), (2, 0), (3, 0), (3, 1), A = (0, 0)}
- Z = (6, 5)
- los
- Tupel = {1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0}
- Text1 = "{1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0}"
- Weg = 11
- ZielWeg = 11
- P_{Modell} = 0.226



Modell:

- p konstant; nur zwei mögliche Ausgänge (rechts/oben); Versuche 11-mal wiederholt; die 11 Versuche unabhängig voneinander
- Darstellung als 11-Tupel: (0,0,1,1,1,0,1,0,0,1,1); Z wird erreicht, da 6-mal rechts (1) und 5-mal links (0) abgebogen wurde
- Nur diejenigen Tupel mit genau sechs Einsen (und 5 Nullen) gelangen zum Ziel.

Das gelingt nur dann, wenn von den 11 Plätzen genau 6 Plätze ausgewählt werden. Das geht auf $\binom{11}{6}$ verschiedene Arten (s. Lotto).

Damit folgt: $\mathbb{P}(\text{"Z erreicht"}) = \binom{11}{6} \cdot p^6 \cdot (1 - p)^5; p = 0.5 \Rightarrow \mathbb{P}(\text{"Z erreicht"}) = \binom{11}{6} \cdot 0.5^{11} = 462 \cdot 0.5^{11} = 0,22558 \dots$

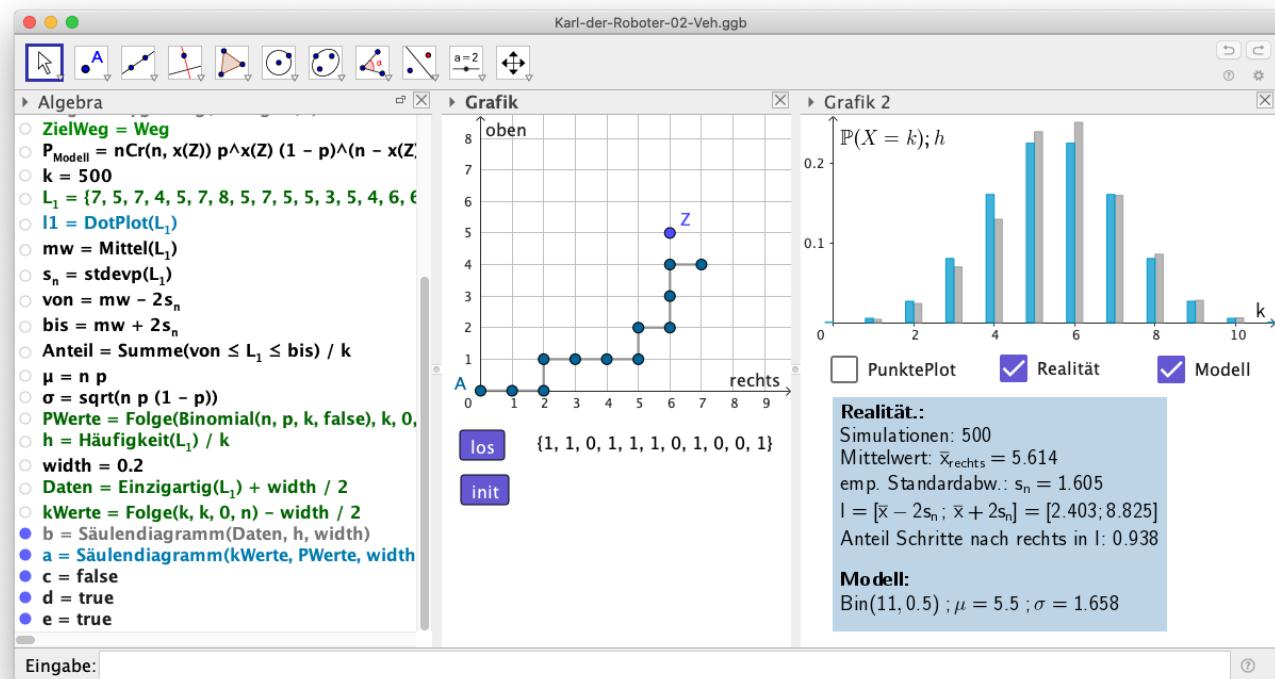
Binomialverteilung und die Suche nach dem Ziel

Wo landet Karl in ca. 95% aller Fälle?

Realität: 95%-PI mit Daten: $[\bar{x} - 2s_n; \bar{x} + 2s_n]$

Modell: 95%-PI mit $X \sim \text{Bin}(n, p)$: $[\mu - 2\sigma_n; \mu + 2\sigma_n]$

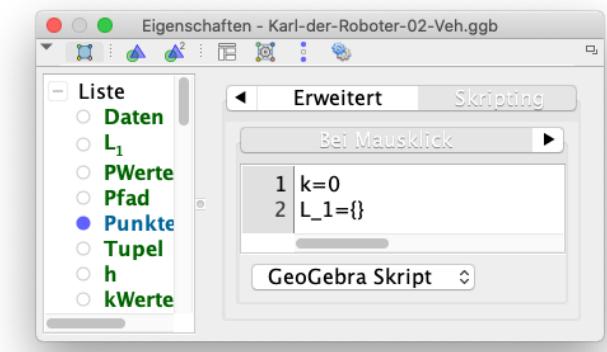
Ein fruchtbare Zusammenspiel von Realität und Modell



Schaltfläche: los



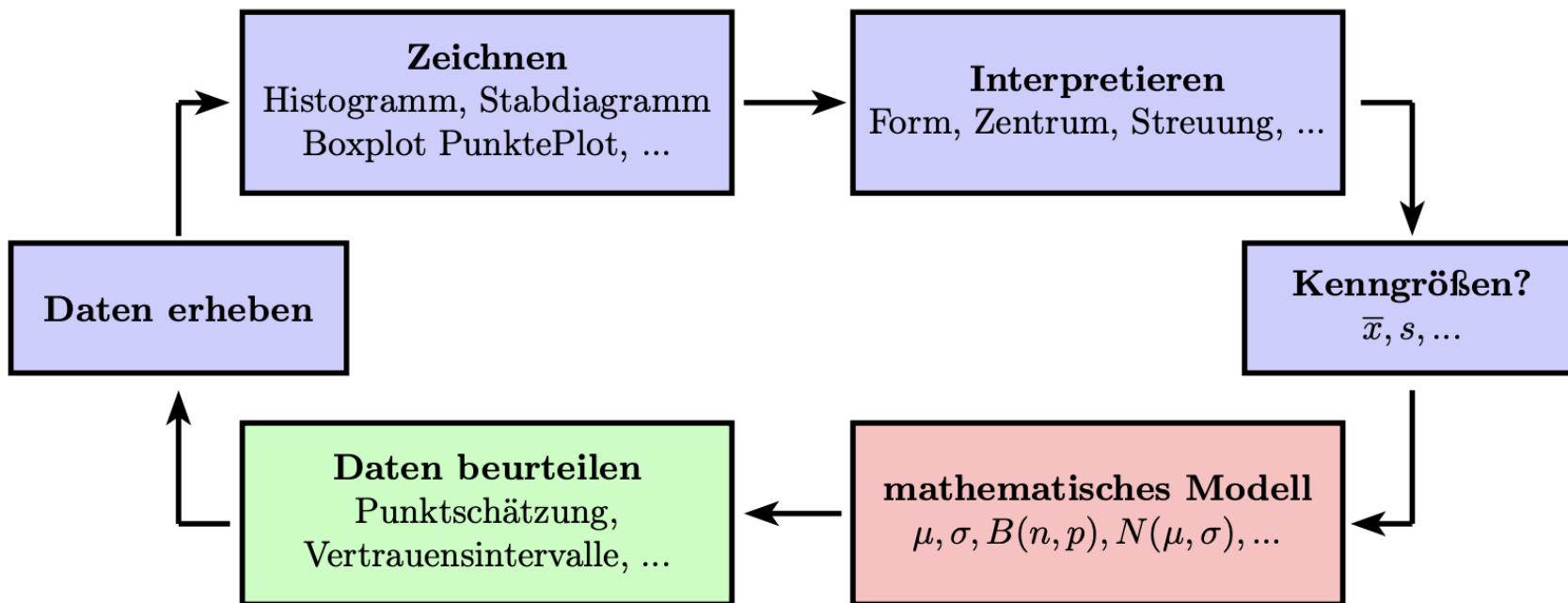
Schaltfläche: init



Karl der Roboter - Modell u. Realität

Nur zusammen sind sie stark.

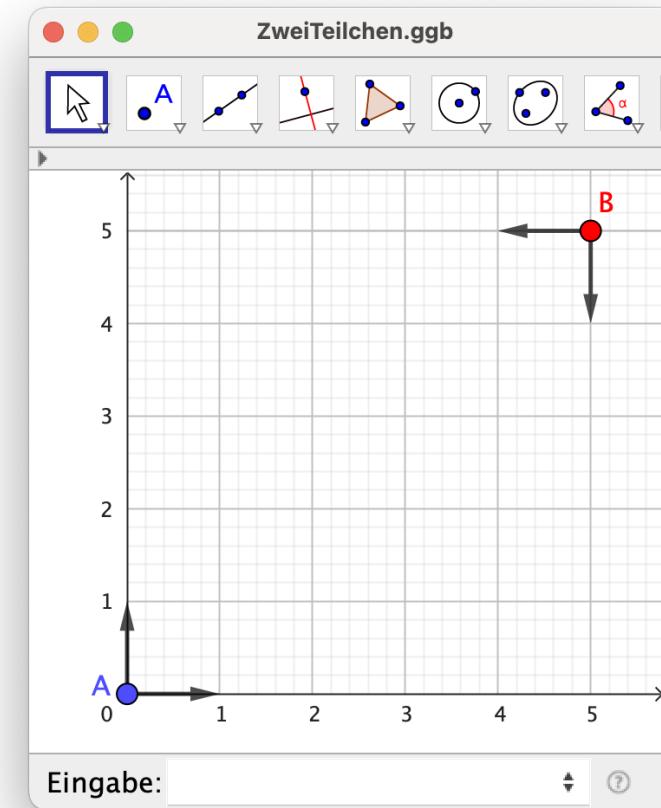
Beschreibende Statistik, Wahrscheinlichkeitsrechnung, Beurteilende Statistik



Binomialverteilung und die Suche nach dem Treffen

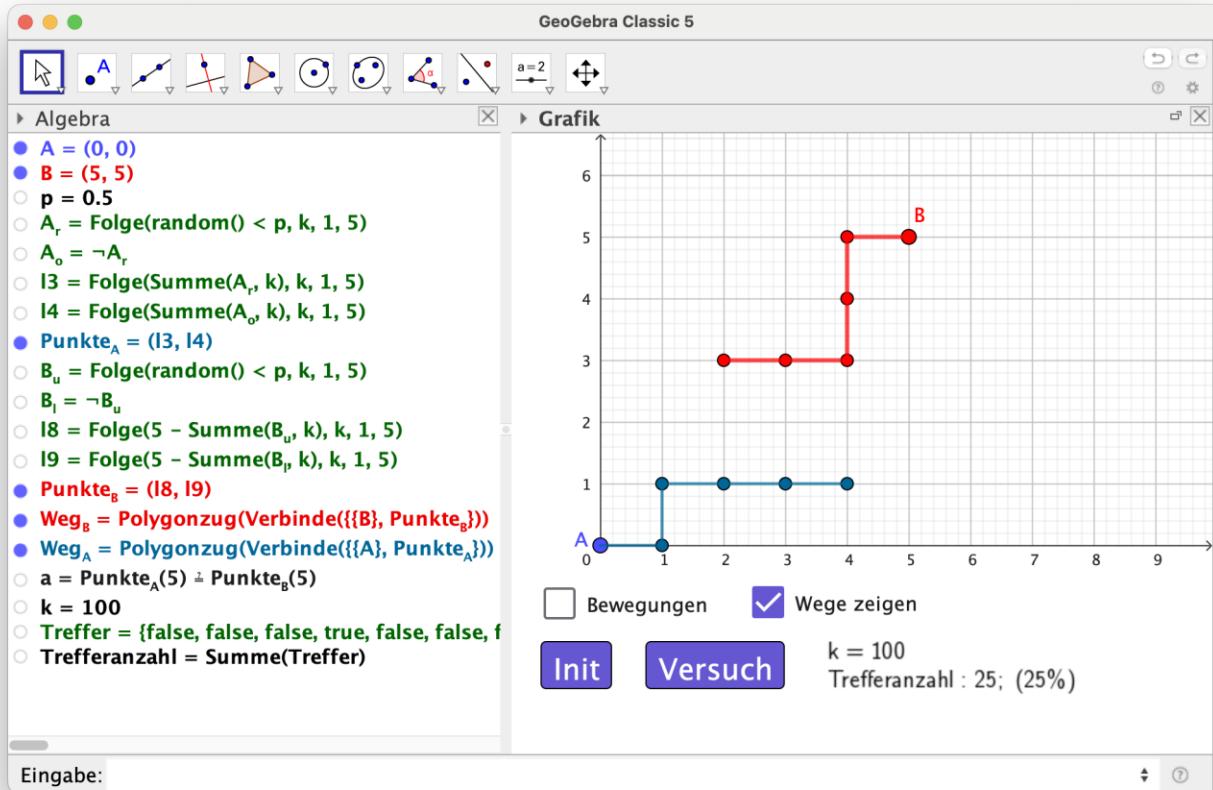
In dem abgebildeten Gitter bewegen sich die Teilchen A und B pro Zeiteinheit jeweils um einen Gitterpunkt. Teilchen A geht nach oben oder nach rechts, Teilchen B geht nach unten oder nach links, jeweils mit Wahrscheinlichkeit 0,5. Sie starten beide zur Zeit 0.

- Ist es möglich, dass sich A und B treffen?
- Nach wie vielen Schritten ist spätestens ein Treffen möglich?
- Simulieren Sie die Bewegung der Teilchen 20-mal, und ermitteln Sie aufgrund Ihrer Beobachtungen einen Schätzwert für die Wahrscheinlichkeit des Treffens beider Teilchen.



Entnommen aus: Stochastik. Mathematik Sekundarstufe II. S. 170. Volk und Wissen. 2001

Binomialverteilung und die Suche nach einem Treffen



[Treffen zweier Teilchen](#)

Teilchen A:

X zählt die Anzahl der Schritte nach **rechts**
 $X \sim \text{Bin}(5; 0.5)$

Teilchen B:

Y zählt die Anzahl der Schritte nach **links**
 $Y \sim \text{Bin}(5; 0.5)$

Modellierung: A und B bewegen sich **unabhängig** voneinander

E: A und B treffen sich nach 5 Schritten

$$P(E) = \sum_{k=0}^{5} P(X = k) \cdot P(Y = 5 - k)$$

$$P(E) = \frac{1}{2^{10}} \cdot \sum_{k=0}^{5} \binom{5}{k} \cdot \binom{5}{5-k} = \frac{252}{1024}$$

Oder so:

$$P(E) = \frac{1^2 + 5^5 + 10^2 + 10^2 + 5^2 + 1}{2^{10}} \approx 0.246$$

Binomialverteilung – Überbuchungen

Eine Ferienwohnanlage hat 80 Wohnungen. Erfahrungsgemäß werden 15% der Buchungen wieder storniert. Der Besitzer nimmt für die Pfingstferien 95 Buchungen an.

- Berechnen Sie die Wahrscheinlichkeit, dass zu viele Buchungen angenommen wurden.
- Bestimmen Sie die Anzahl an Buchungen, die der Besitzer für die Pfingstferien annehmen sollte, damit die Wahrscheinlichkeit, dass zu viele Buchungen angenommen wurden, kleiner als 5% ist.

Entnommen aus: Lambacher Schweizer. Mathematik Qualifikationsphase. Leistungskurs. S: 286. Niedersachsen. 2018

Dieser Aufgabentyp findet sich in vielen Schulbüchern und auch in Abituraufgaben.

Vorschlag: Wie sinnvoll ist es eigentlich, mit einer Binomialverteilung zu modellieren? Nennen Sie Gründe, die dagegensprechen.

Man kann natürlich ChatGPT 4o fragen – hier wird sofort mit $\text{Bin}(95; 0.85)$ modelliert.

Aufgabe (a): Wahrscheinlichkeit der Überbuchung

Der Besitzer nimmt 95 Buchungen an, es gibt aber nur 80 Wohnungen. Jede Buchung wird mit einer Wahrscheinlichkeit von 15% storniert, das heißt, die Wahrscheinlichkeit, dass eine Buchung nicht storniert wird, beträgt:

$$p = 1 - 0.15 = 0.85$$

Die Anzahl der tatsächlich ankommenden Gäste ist binomialverteilt:

$$X \sim B(n = 95, p = 0.85)$$

Gesucht ist die Wahrscheinlichkeit, dass mehr als 80 Gäste erscheinen:

$$P(X > 80) = 1 - P(X \leq 80)$$

Die Berechnung erfolgt über die kumulative Verteilungsfunktion der Binomialverteilung. Das ergibt:

$$P(X > 80) = 1 - P(X \leq 80) \approx 1 - 0.458 = 0.542$$

Die Wahrscheinlichkeit, dass zu viele Gäste erscheinen, beträgt also ca. 54,2%.

Binomialverteilung – Überbuchungen

Eine Ferienwohninanlage hat 80 Wohnungen.

Erfahrungsgemäß werden 15% der Buchungen wieder storniert. Der Besitzer nimmt für die Pfingstferien 95 Buchungen an.

- Berechnen Sie die Wahrscheinlichkeit, dass zu viele Buchungen angenommen wurden.
- Bestimmen Sie die Anzahl an Buchungen, die der Besitzer für die Pfingstferien annehmen sollte, damit die Wahrscheinlichkeit, dass zu viele Buchungen angenommen wurden, kleiner als 5% ist.

Entnommen aus: Lambacher Schweizer. Mathematik Qualifikationsphase. Leistungskurs. S: 286. Niedersachsen. 2018

... und hier die Lösung von ChatGPT 4o zu b)

Aufgabe (b): Maximale Anzahl an Buchungen mit Überbuchungswahrscheinlichkeit < 5%

Wir suchen eine Anzahl n von Buchungen, sodass die Wahrscheinlichkeit einer Überbuchung unter 5% liegt:

$$P(X > 80) < 0.05$$

Dabei gilt weiterhin:

$$X \sim B(n, p = 0.85)$$

Wir berechnen nun für verschiedene Werte von n , bis die Bedingung erfüllt ist:

- Für $n = 88$ ergibt sich:

$$P(X > 80) = 1 - P(X \leq 80) \approx 0.048$$

- Für $n = 89$ ergibt sich:

$$P(X > 80) \approx 0.057$$

Da $P(X > 80) < 0.05$ für $n = 88$ gerade noch erfüllt ist, aber für $n = 89$ überschritten wird, lautet die optimale Anzahl der angenommenen Buchungen:

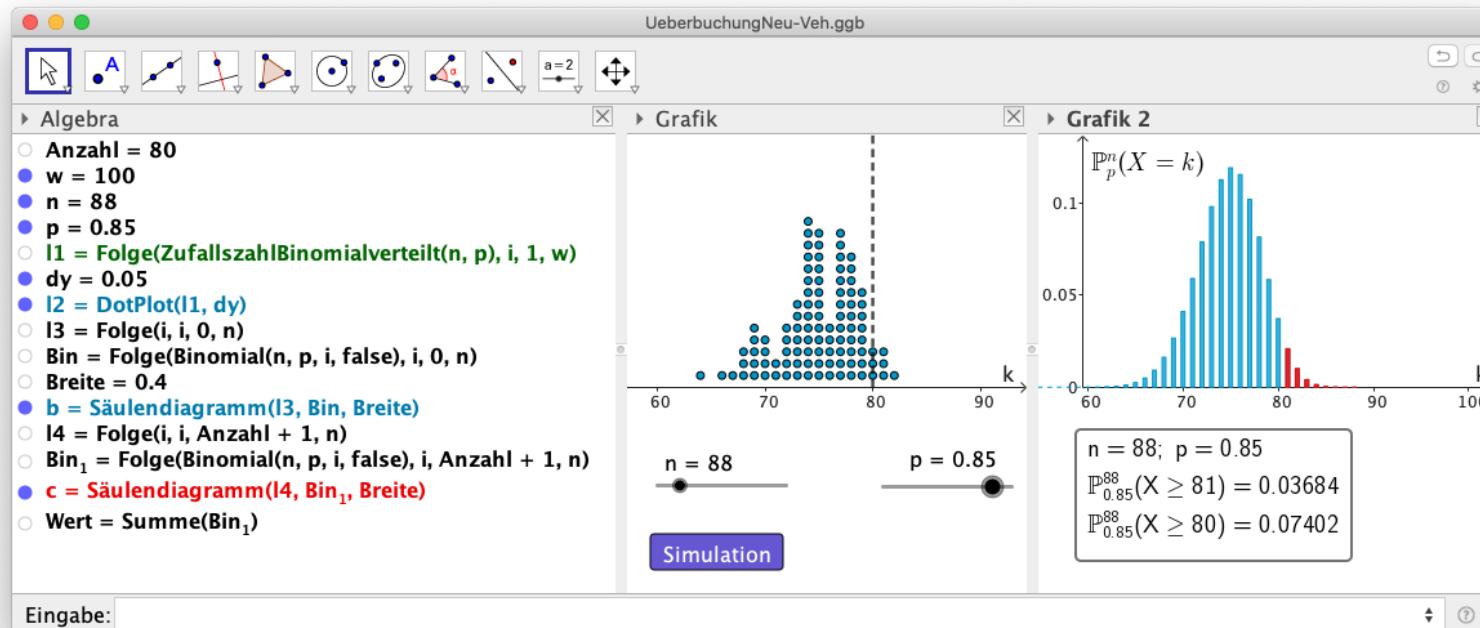
$$n = 88$$

Das bedeutet: Der Besitzer sollte maximal 88 Buchungen annehmen, damit die Wahrscheinlichkeit einer Überbuchung unter 5% bleibt.

Binomialverteilung – Überbuchungen

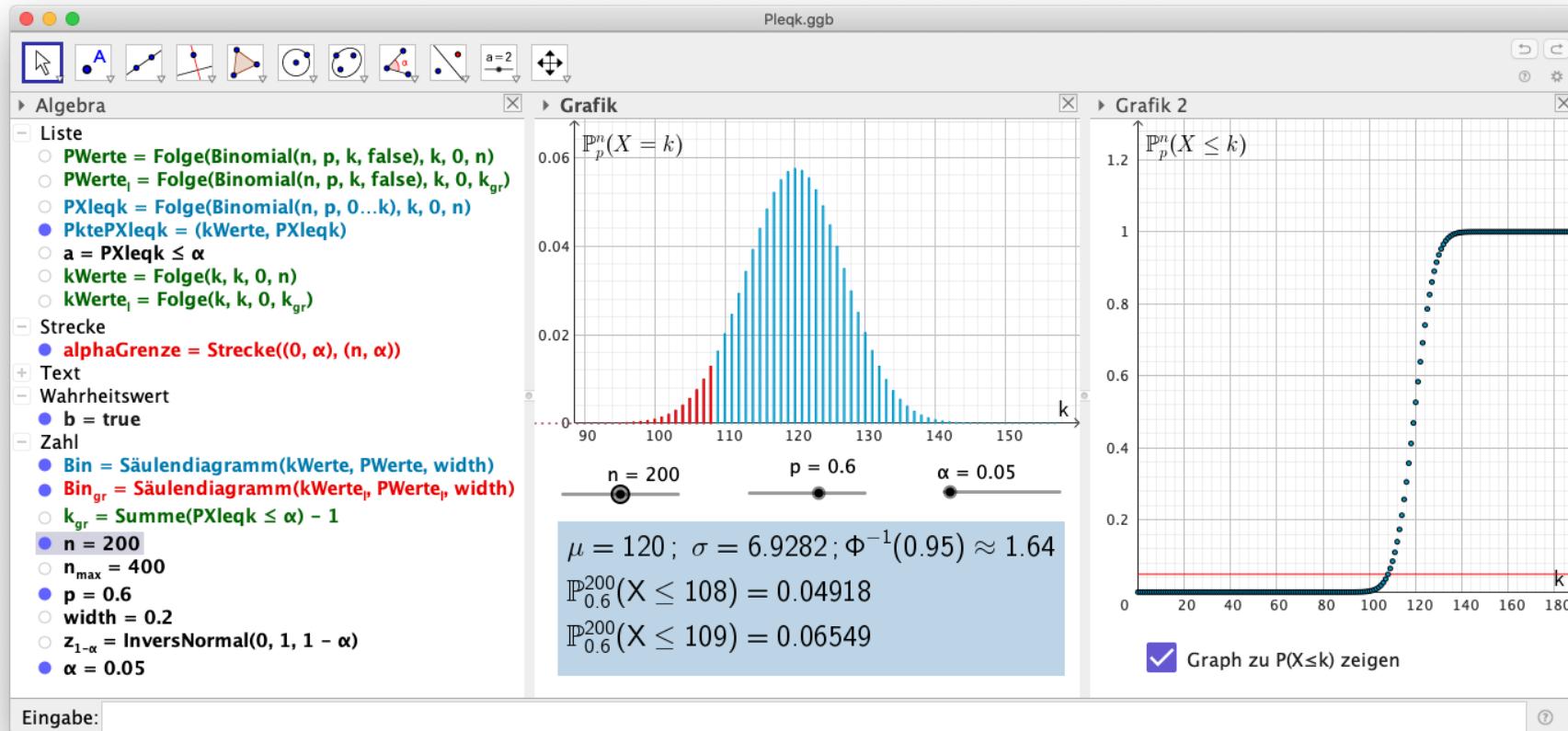
Eine Ferienwohnanlage hat 80 Wohnungen. Erfahrungsgemäß werden 15% der Buchungen wieder storniert. Der Besitzer nimmt für die Pfingstferien 95 Buchungen an.

- Berechnen Sie die Wahrscheinlichkeit, dass zu viele Buchungen angenommen wurden.
- Bestimmen Sie die Anzahl an Buchungen, die der Besitzer für die Pfingstferien annehmen sollte, damit die Wahrscheinlichkeit, dass zu viele Buchungen angenommen wurden, kleiner als 5% ist.



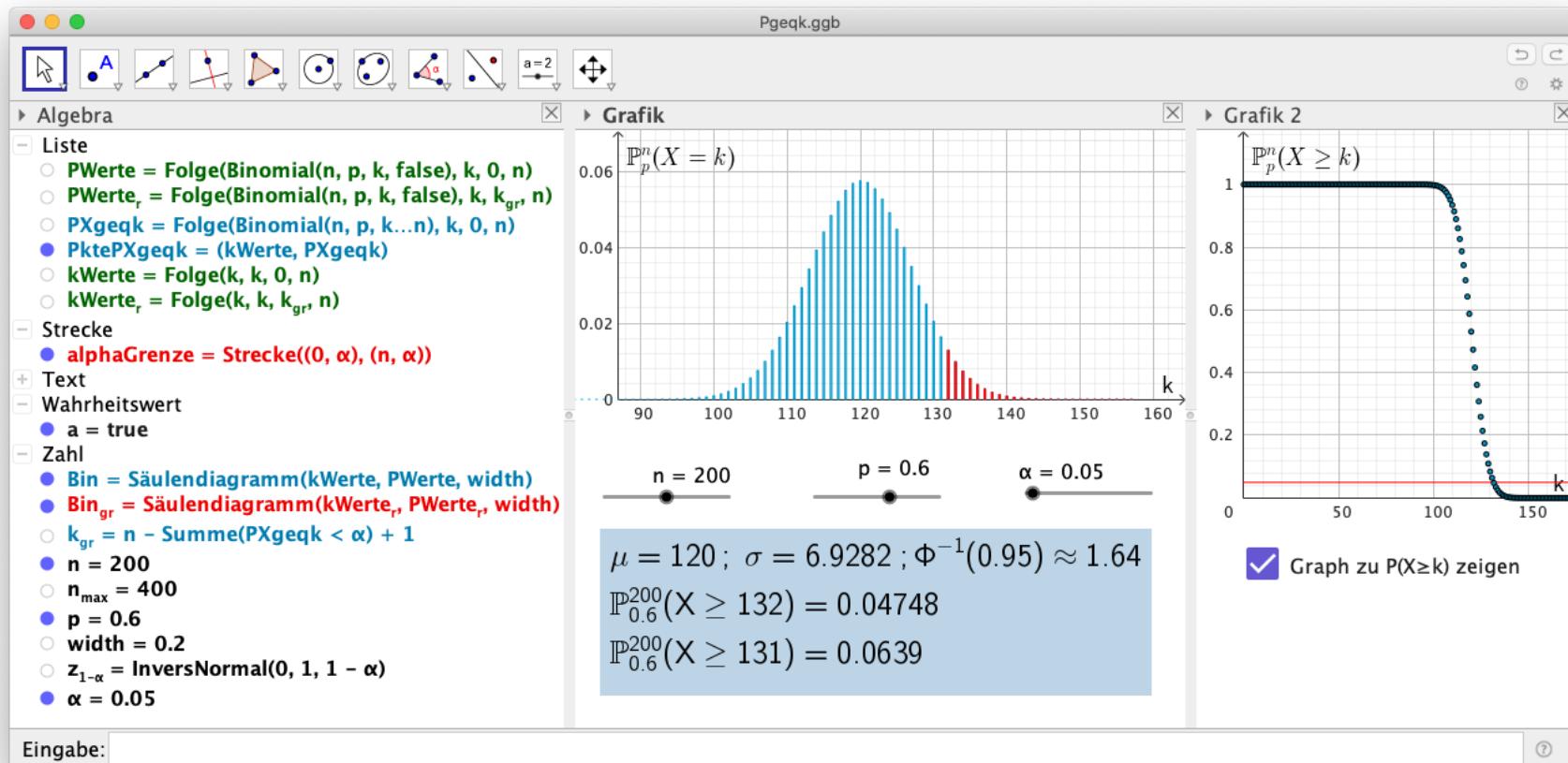
Überbuchungen

Binomialverteilung – ges. k_{max} mit $P_p^n(X \leq k_{max}) < \alpha$



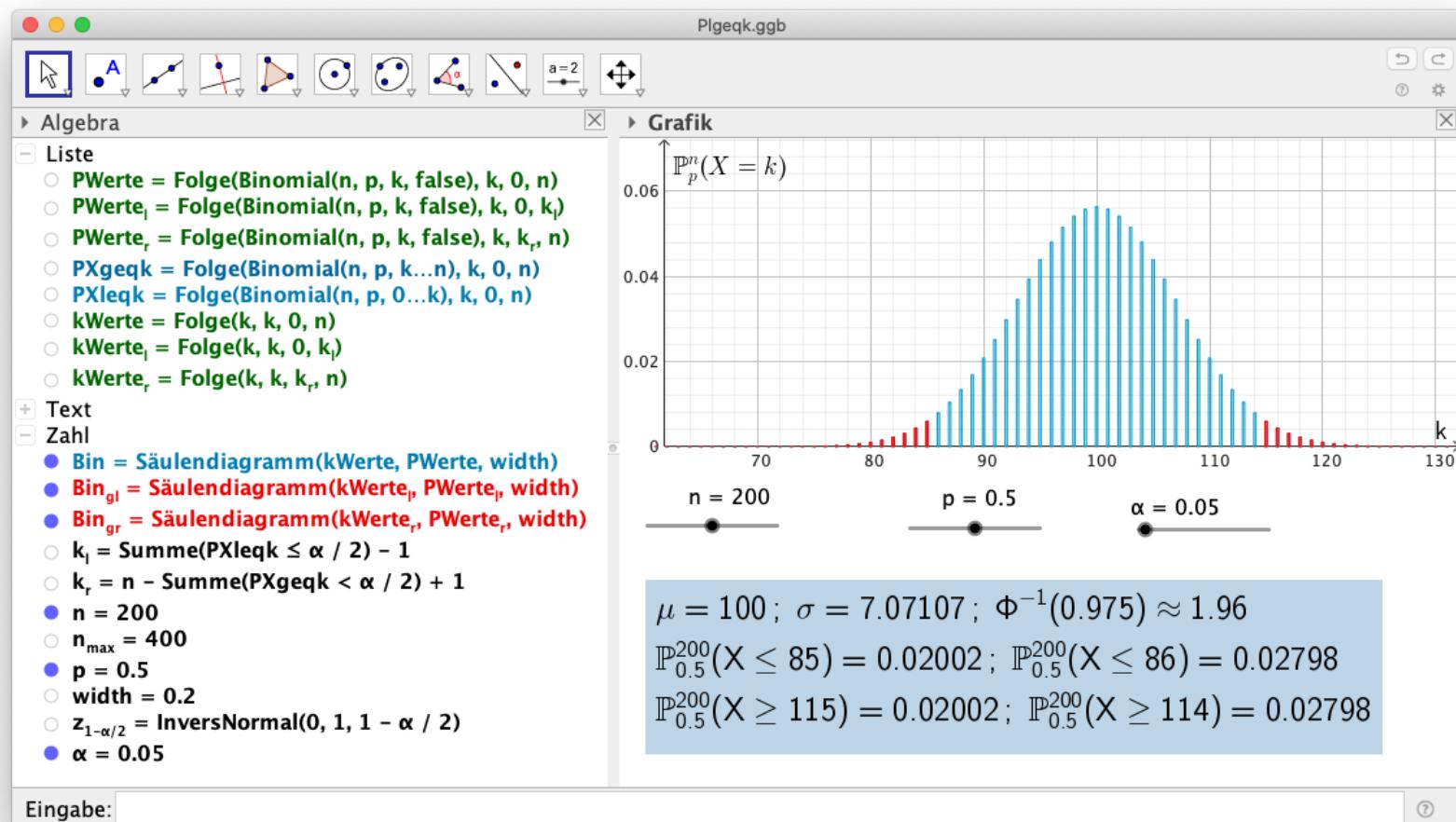
$P(X \leq k)$

Binomialverteilung – ges. k_{min} mit $P_p^n(X \geq k_{min}) < \alpha$



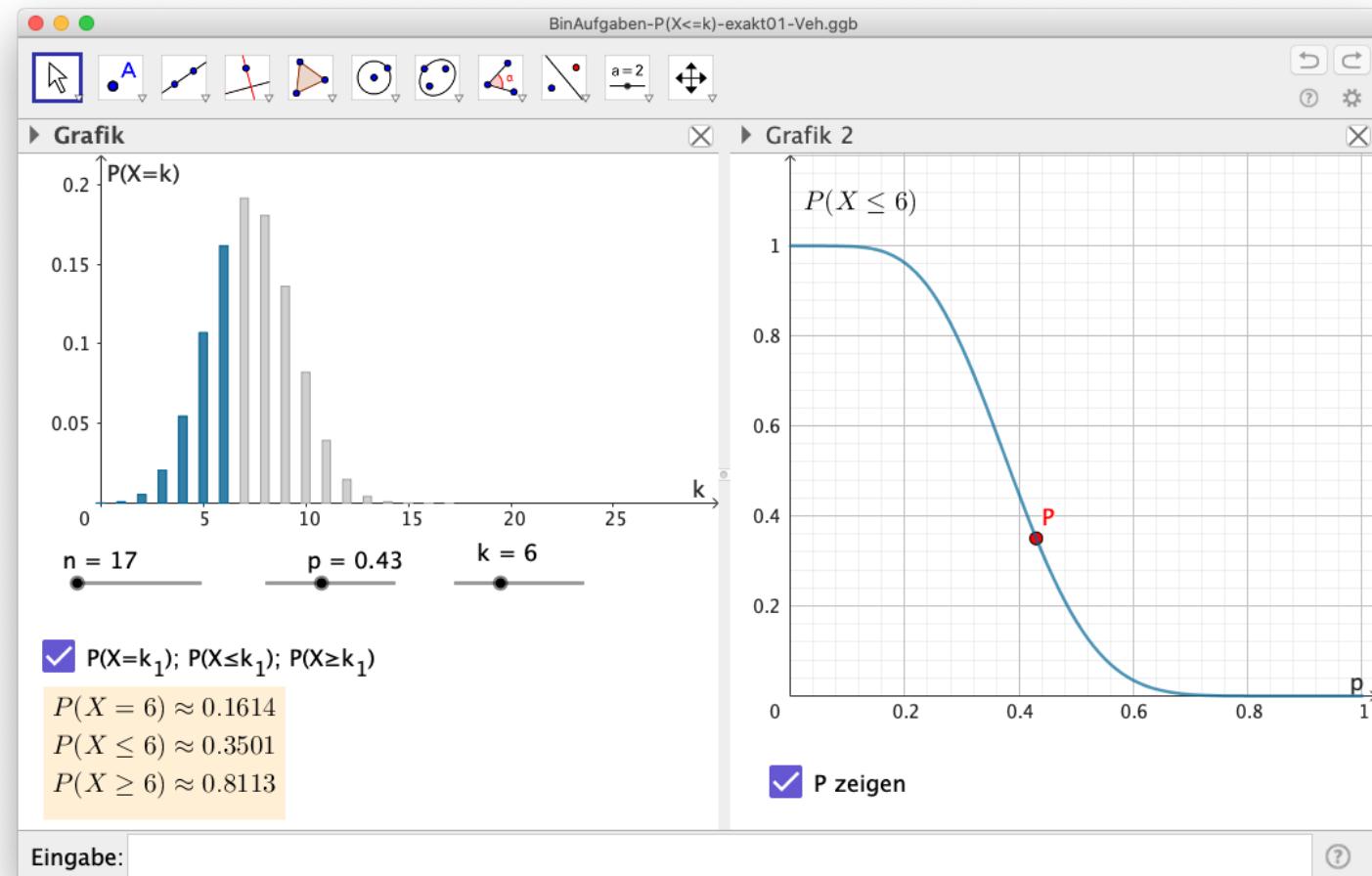
$P(X \geq k)$

Binomialverteilung – σ -Umgebungen berechnen

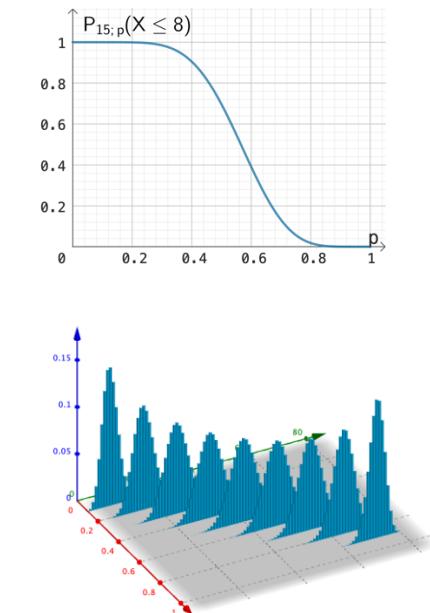
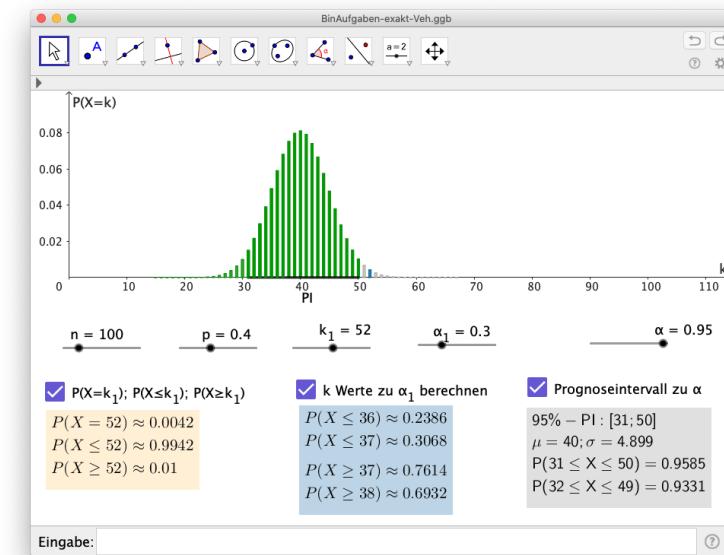
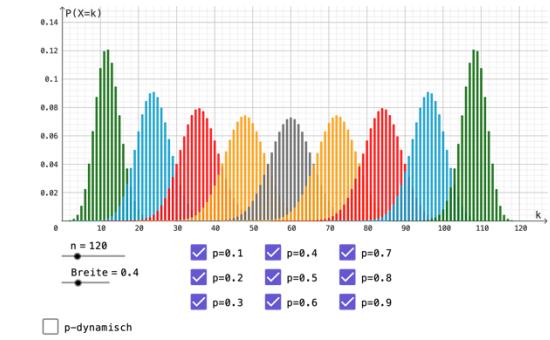
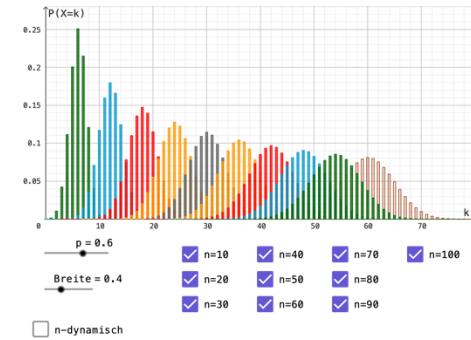
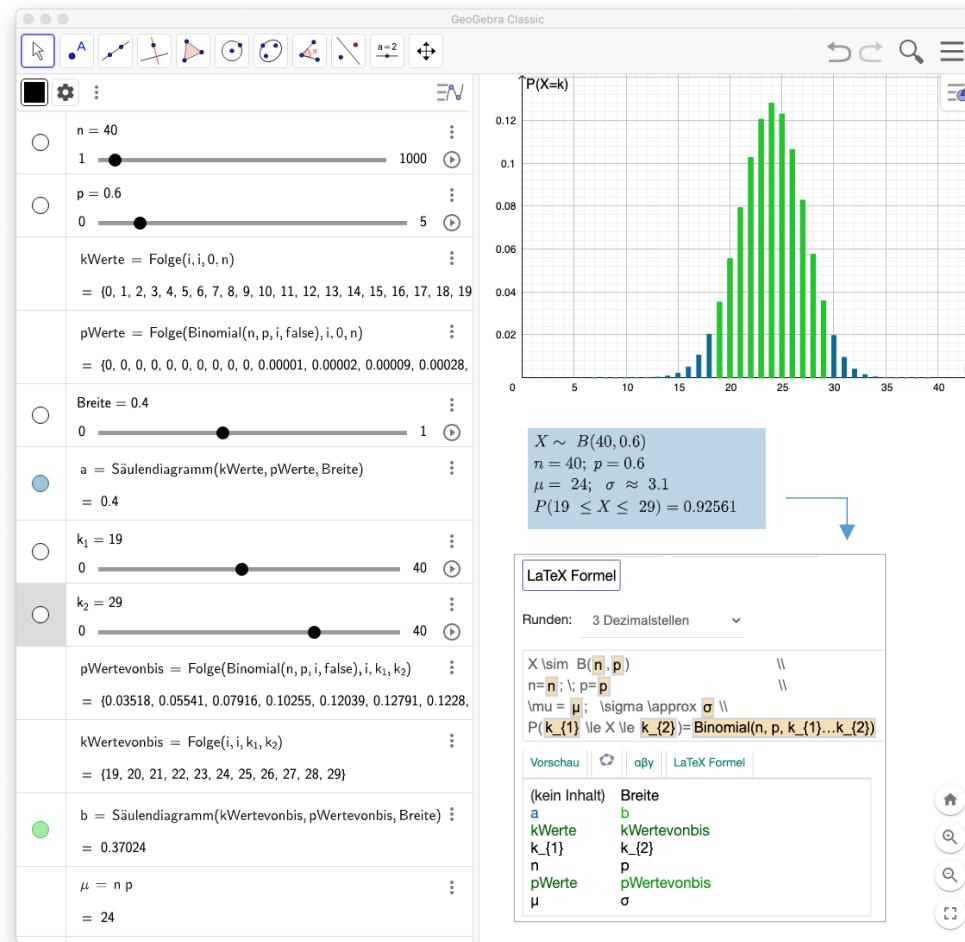


$$P(a <= X <= b)$$

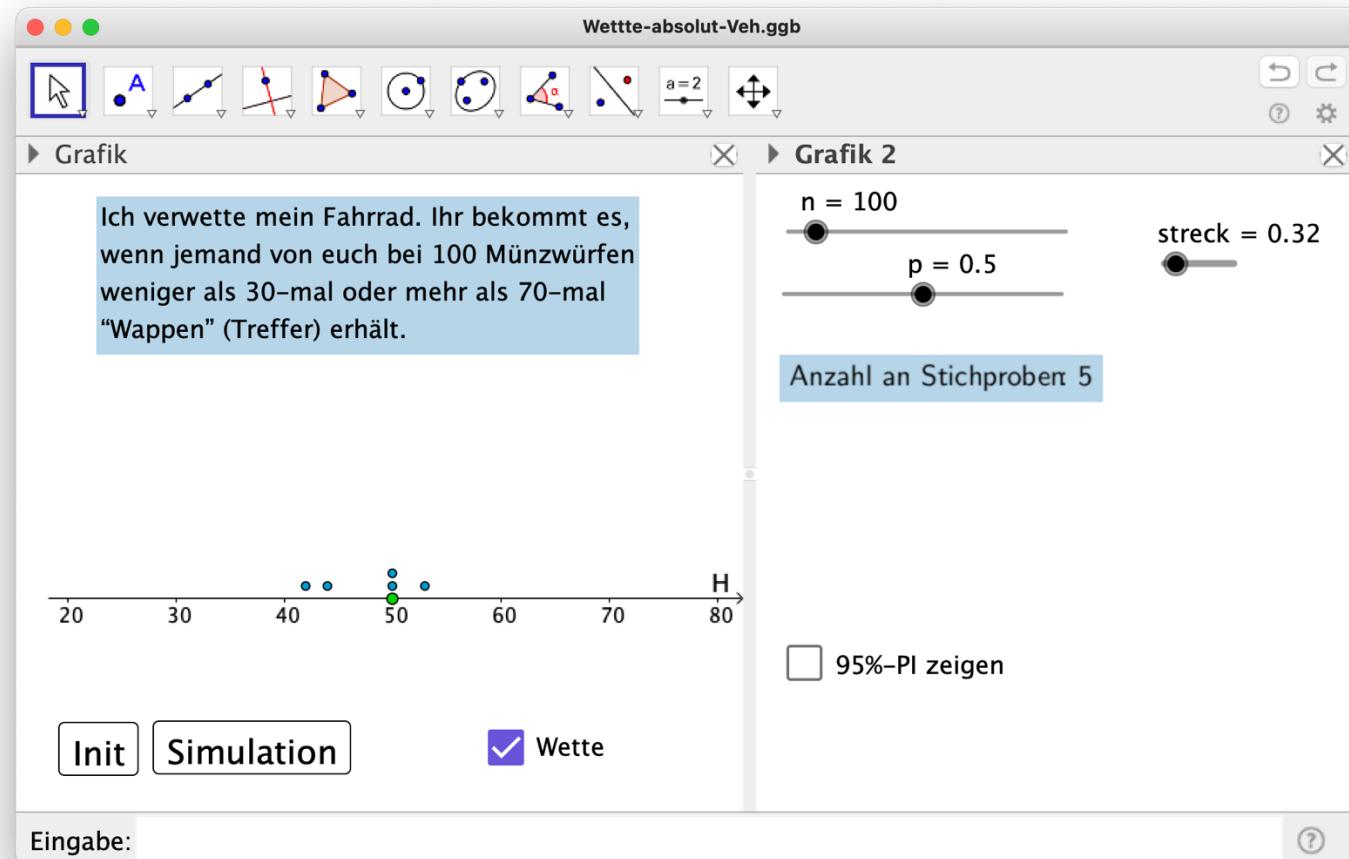
Binomialverteilung – einmal anders



Zentrale Vorgehensweisen

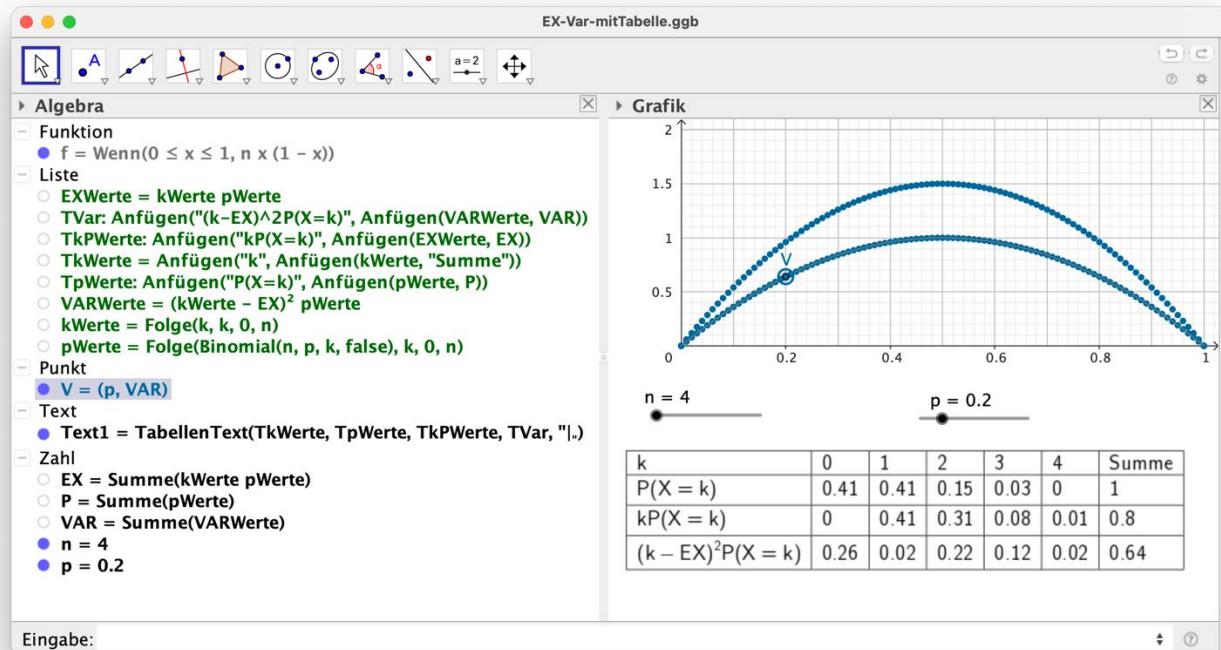


Binomialverteilung – σ -Sigma-Umgebungen mit einer Wette

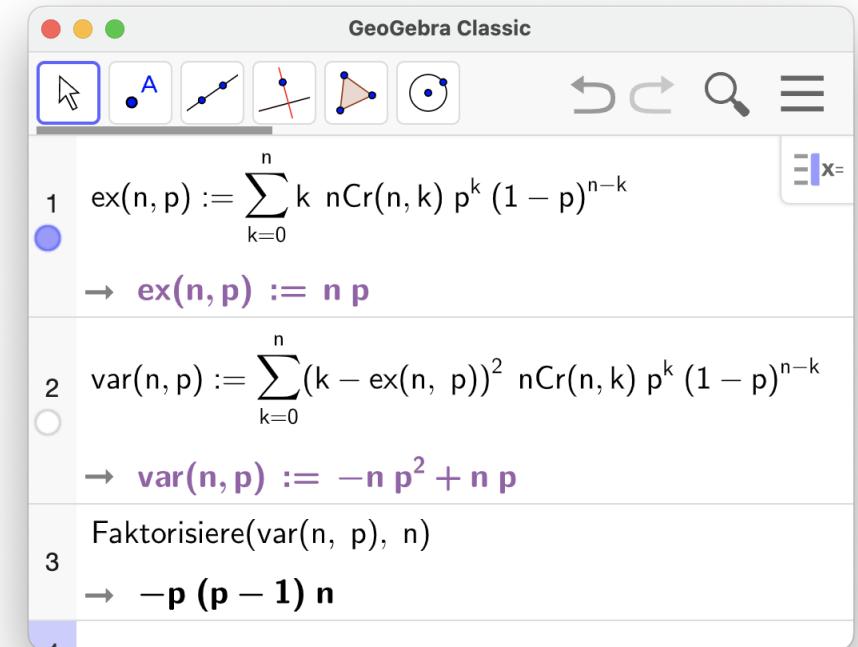


Wette - absolut

Erwartungswert und Varianz einer Binomialverteilung

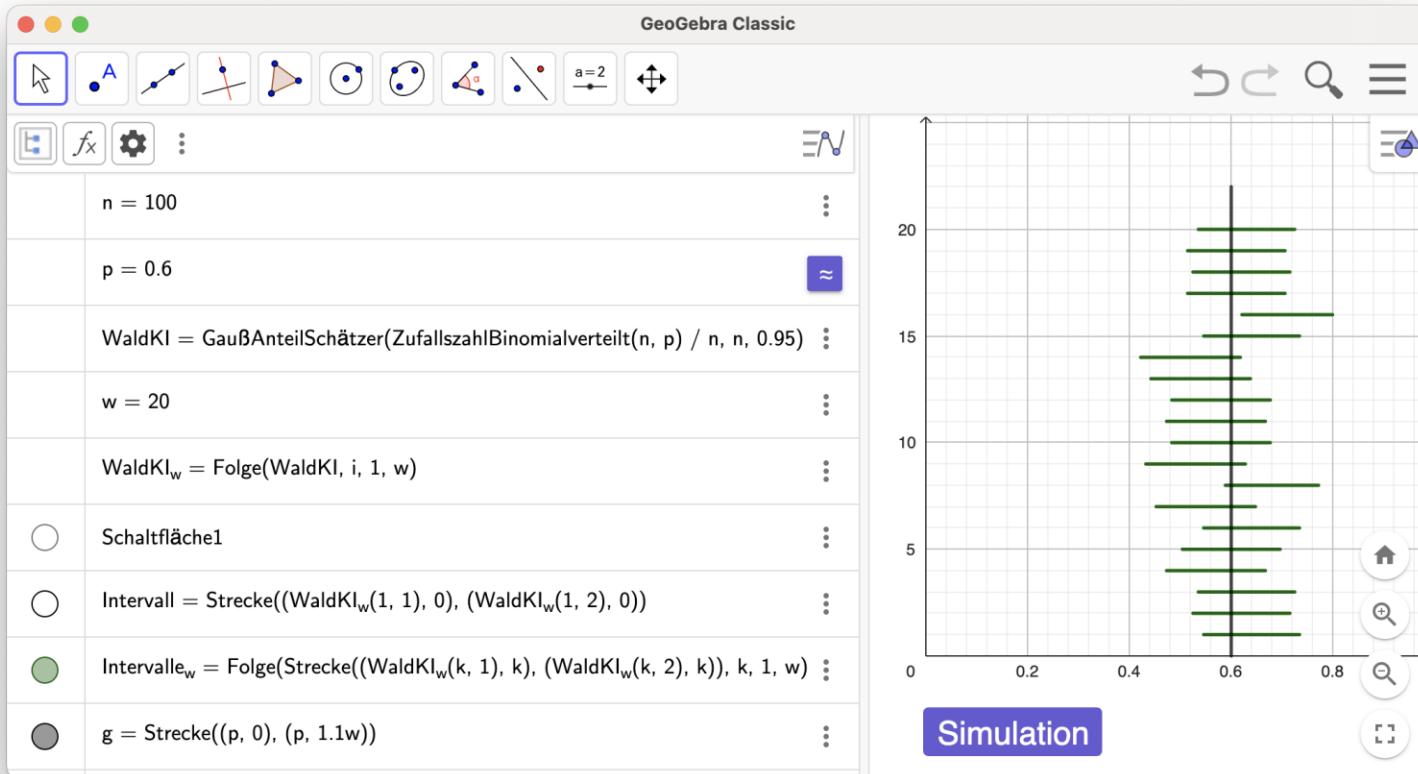


EX und VX mit Tabelle



EX und VX mit CAS

WALD-Konfidenzintervalle – eine einfache Simulation



$$s_n := \sqrt{h(1 - h)/n}$$

$$\text{WaldKI} = [h - 1.96s_n; h + 1.96s_n]$$

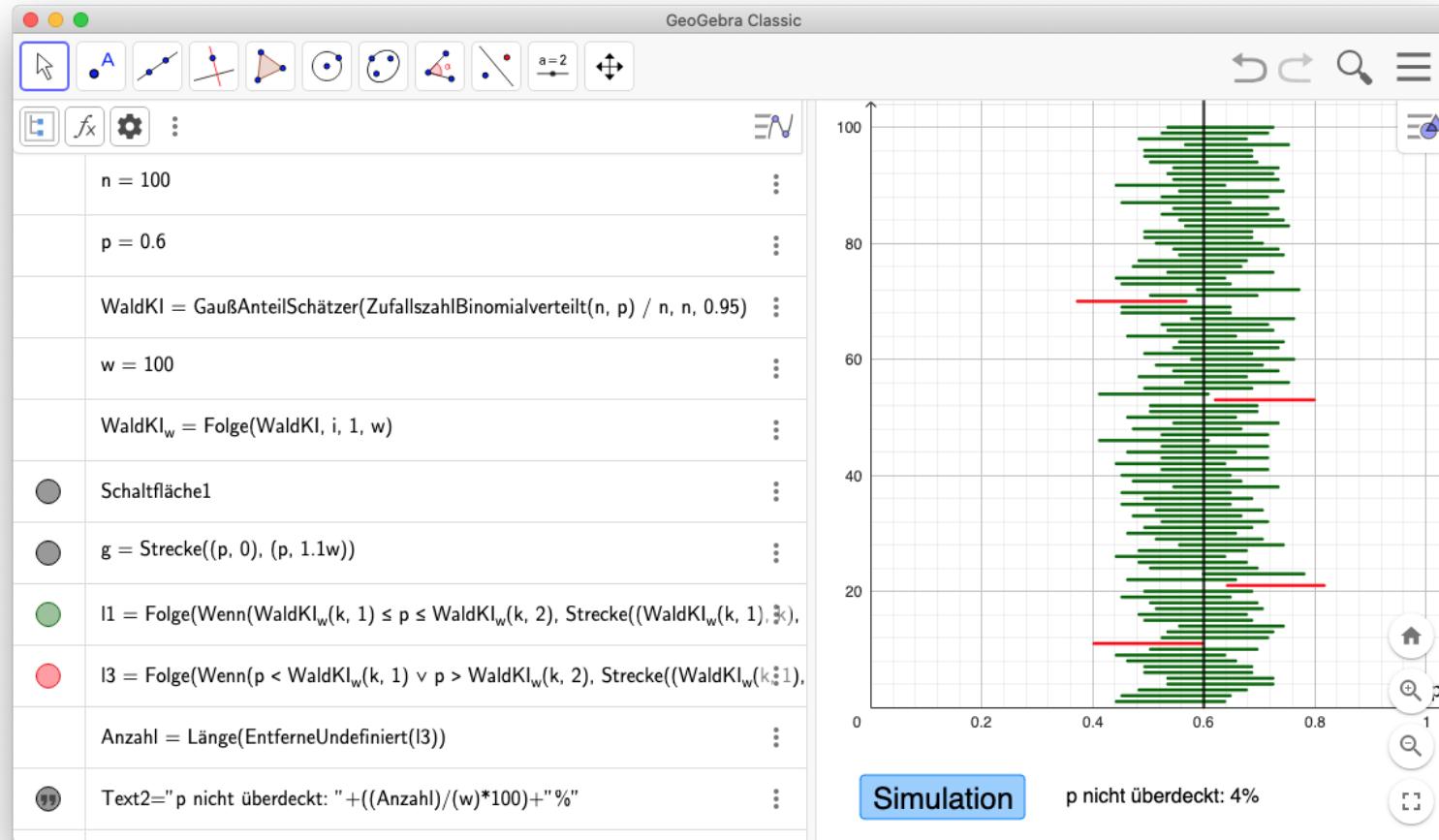
Das muss vermittelt werden:

Das Intervall ist selbst nach der Berechnung nicht mehr zufällig, sondern es lässt nur die Methode eine probabilistische Aussage zu.

Wenn man viele Stichproben zieht und für jede ein Konfidenzintervall berechnet, kann man anschaulich zeigen, dass einige Intervalle den wahren Wert enthalten und andere nicht – aber wir wissen nicht, ob unser spezielles Intervall „glücklich“ war (den für immer unbekannten Wert p überdeckt) oder nicht.

Ein hilfreiches Konzept ist es, **Konfidenzintervalle als Zufallsvariablen** zu betrachten: Die Grenzen des Intervalls sind Realisationen der beiden Zufallsvariablen, aber sobald wir unser Intervall ausgerechnet haben, ist es fix. Das macht den Unterschied zwischen der „Methodenwahrscheinlichkeit“ und der „posteriori Unsicherheit“ (hoffentlich) klarer.

WALD-Konfidenzintervalle – komplexe Simulation

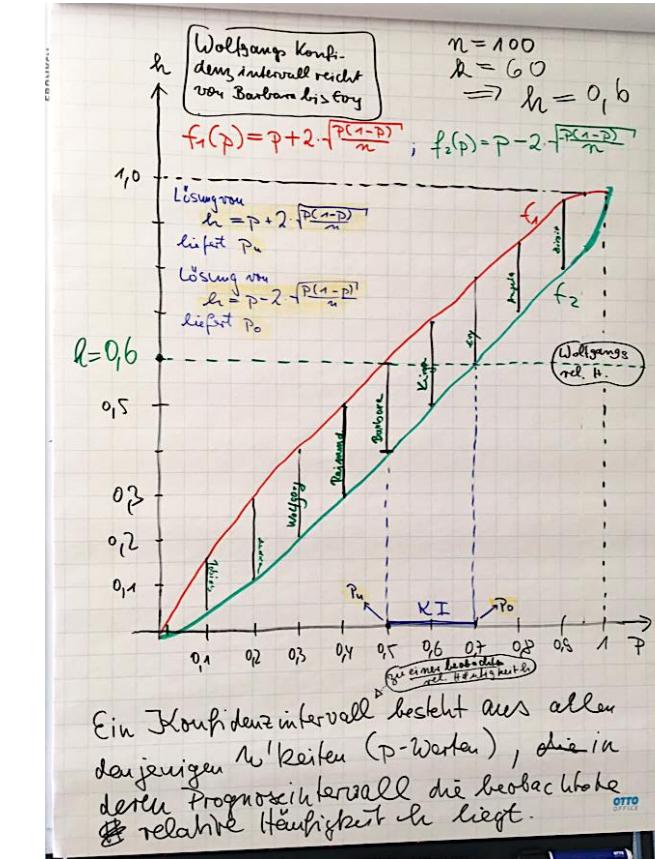
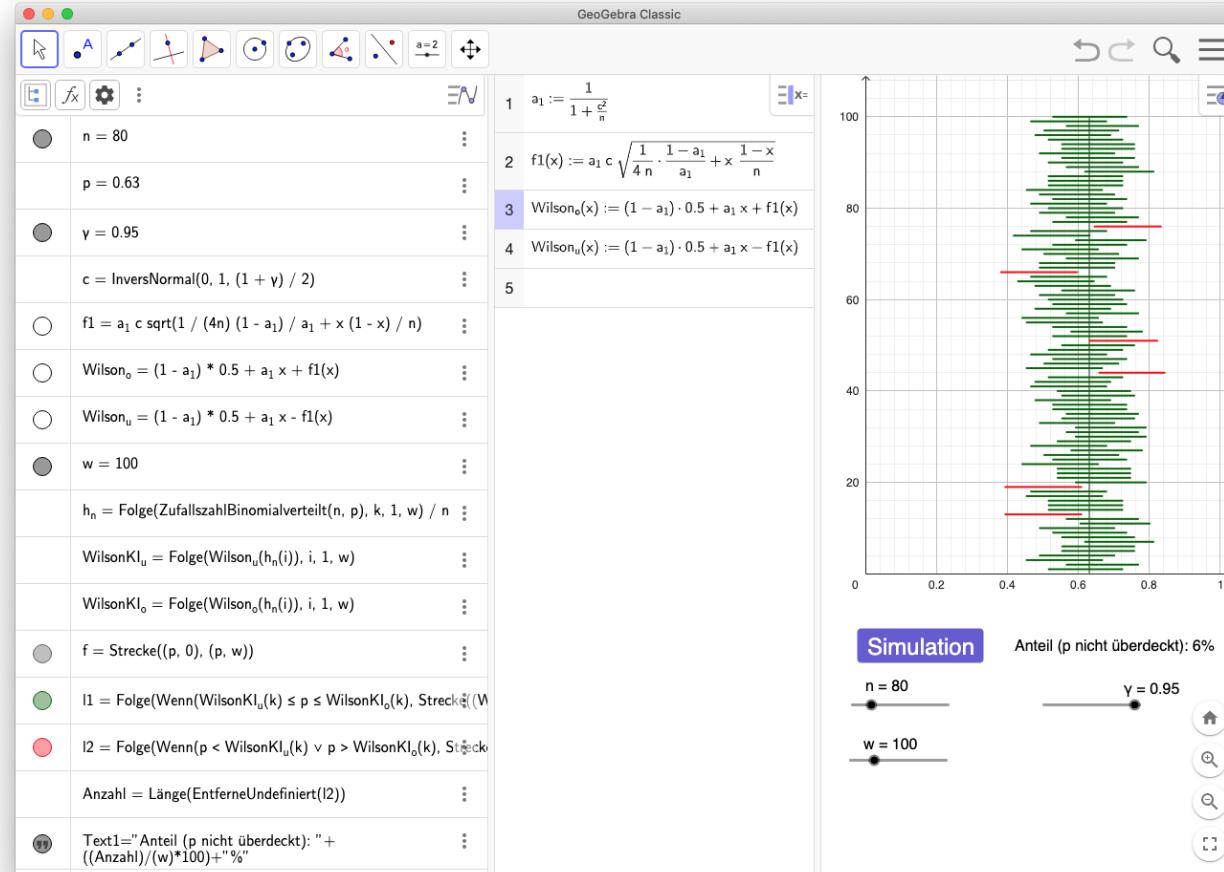


$$s_n := \sqrt{h(1 - h)/n}$$

$$\text{WaldKI} = [h - 1.96s_n; h + 1.96s_n]$$

[WALD-KI-Simulationen](#)

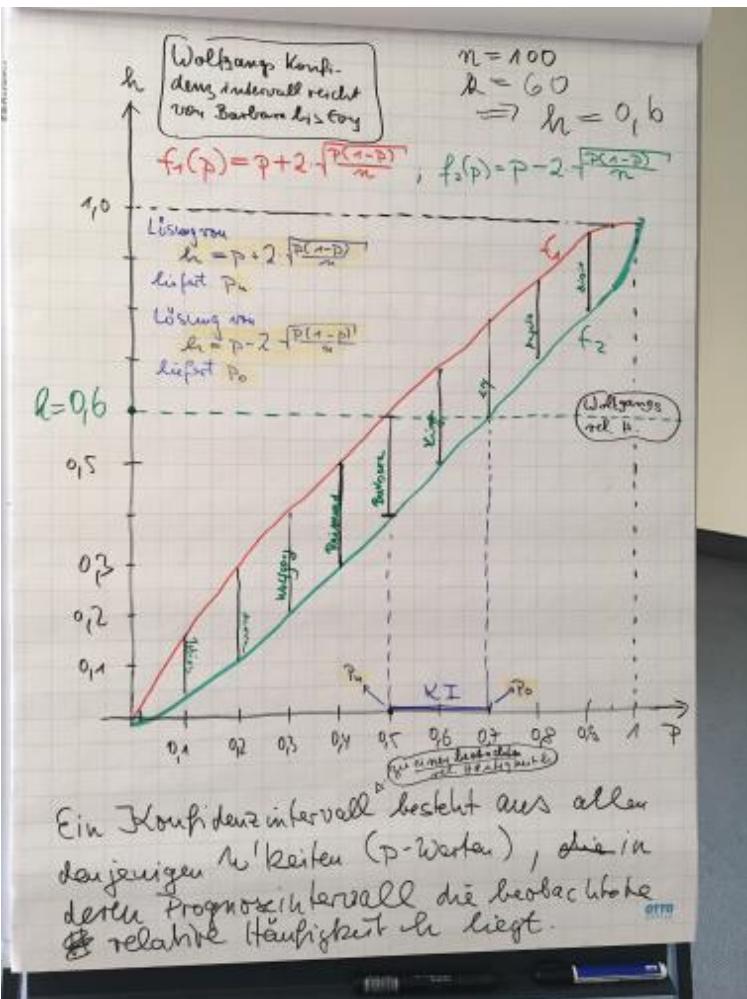
Konfidenzintervalle – eine Simulation mit WILSON-Intervallen



WILSON-KI-Simulationen

Das Bild stammt aus einer denkwürdigen Fortbildung in Mellingen.

Konfidenzintervalle – sehr knapp erläutert



Eine Folie aus einem Vortrag in Thüringen (s. Material):

Methode I - WILSON-Konfidenzintervall

Sichtweise von p aus:

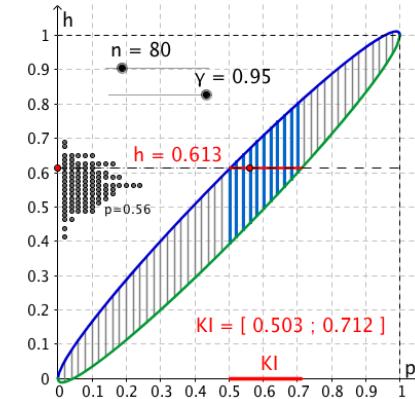
Zu unserem ersten Ansatz ist die folgende Herangehensweise gleichwertig:

Ein Konfidenzintervall besteht aus allen p^0 s, die h im Prognoseintervall haben.

Die Grafik verdeutlicht dies für die Ermittlung des 95%-Konfidenzintervalls. Zu einigen Werten für p werden die zugehörigen 95%-Prognoseintervalle als vertikale Strecken gezeichnet.

Die Horizontale zu h schneidet diese Intervalle. Dadurch ergeben sich (näherungsweise) die beiden Grenzen des Konfidenzintervalls $[p_u; p_o]$. Ermittlung durch Schnittpunktbestimmung:

$$h = p_o - 2 \cdot \frac{p_o \cdot (1 - p_o)}{n}; h = p_u + 2 \cdot \frac{p_u \cdot (1 - p_u)}{n}$$



Konfidenzintervalle – aus meinem (ehemaligen) Unterricht

95% - Prognoseintervalle

Von der Grundgesamtheit zur Stichprobe

p gegeben ; n : Ergebnis (Realisation) einer Zufallsvariablen

$$95\%-PI: \left[p - 1,96 \sqrt{\frac{p(1-p)}{n}}, p + 1,96 \sqrt{\frac{p(1-p)}{n}} \right]$$

Interpretation:

- ① Das Stichprobenergebnis n liegt mit 95%-iger Wahrscheinlichkeit im 95%-PI.
- ② Würden viele Stichproben (n immer gleich) durchgeführt, so erwartet man in ca. 95% aller Fälle, dass n im PI liegt.
Von ca. 5% aller Werte für n liegen dann außerhalb.

Bsp:

Bei der Landtagswahl 2017 in NDS hat die SPD 36,8% aller Stimmen erhalten.
Es soll nach der Wahl eine (Zufalls)-Stichprobe ($n = 400$) unter den Personen durchgeführt werden, die gewählt haben.

Errechnen Sie das zugehörige 95%-PI.

Lösung: 95%-PI: $[0,3001\ldots; 0,41785\ldots]$

ausführliche Interpretation:

Die W'keit, dass das Stichprobenergebnis im 95%-PI liegt, beträgt 95%.

95% - Vertrauensintervalle

Von der Stichprobe zur Grundgesamtheit

p unbekannt – das nicht das Ergebnis einer Zufallsvariablen „fest“

Die Grenzen des 95%-VI sind die Ergebnisse einer Zufallsvariablen.

Kont.: Die Intervallgrenzen sind zufällig, p ist „nur“ unbekannt, aber nicht zufällig.

Deshalb SIE: Ein PI hat „feste“ Intervallgrenzen.

Definition eines 95%-VI:

Alle p -Werte gehören zum 95%-VI, die das Stichprobenergebnis n im 95%-PI ließen.

Interpretation:

- ① Der unbekannte Wert für p wird mit 95%-W'keit von 95%-VI „überdeckt“.
- ② Würde man viele 95%-VI erstellen, dann würde man in ca. 95% aller Fälle erwarten, dass p vom 95%-VI „überdeckt“ wird.

Bsp:

Vor der Landtagswahl 2017 in NDS wird eine (Zufalls)-Stichprobe ($n = 400$) erhoben. Hierbei geben 133 Personen an, die SPD zu wählen.

Errechnen Sie das zugehörige 95%-VI!

Lösung: 95%-VI (WILSEN): $[0,3024\ldots; 0,3954\ldots]$

ausführliche Interpretation: Der unbekannte Wert p wird mit 95%-W'keit von VI überdeckt. Mehr kann man nicht sagen!

Normalverteilung - Einstieg

1 Auf dem Weg zur Normalverteilung

Ich habe vor einiger Zeit bei einem italienischen Lebensmittelhändler mehrere 1 kg-Tüten Mehl für leckeren Pizzateig gekauft. Als Mathelehrer habe ich zuhause gleich nachgemessen.

Unter der Voraussetzung, dass die Verpackung immer gleich viel wiegt, konnte ich beim Gewicht trotzdem Abweichungen feststellen.

Abweichungen vom Sollwert (hier 1 kg) sind völlig normal, wenn sie etwa gleichermaßen in beide Richtungen auftreten. Das war auch hier der Fall. Falls nicht, hätte ich mich beschweren können. Es gibt Tabellen mit Abweichungen, die man akzeptieren muss.

Verpackungsmaschinen sind komplexe Maschinen mit vielen Bauteilen, die im Zusammenspiel eine bestimmte Menge an Mehl in eine Packung schütten. Jedes Teil bewirkt eine (hoffentlich) kleine Abweichung, entweder nach oben oder nach unten. Alle Abweichungen zusammen ergeben die Gesamtabweichung. Der Hersteller möchte natürlich vorher wissen, welcher Anteil an Mehltüten ein bestimmtes Gewicht unter- bzw. überschreitet.

Und genau hier kommt die Mathematik mit einer überaus wichtigen stetigen Wahrscheinlichkeitsverteilung ins Spiel: Die Normal- oder Gauß-Verteilung (nach Carl Friedrich Gauß). Sie beschreibt Verteilungen, die durch additive Überlagerung einer großen Anzahl von unabhängigen Einflüssen entstehen.

Die Abweichungen der Messwerte vieler natur-, wirtschafts- und ingenieurwissenschaftlicher Vorgänge vom Erwartungswert lassen sich mit dem Modell der Normalverteilung oft in guter Näherung beschreiben.



Auswertung der Simulation mit den TI 84+

Summe	Häufigkeit	
-4	1	1
-3	1111	4
-2	LHT LHT	6
-1	LHT LHT LHT	8
0	LHT LHT LHT LHT	12
1	LHT LHT LHT LHT	12
2	LHT LHT	4
3	LHT	3
4		0
Σ	50	

Summe	Häufigkeit	
-4	1111	2
-3	LHT	4
-2	LHT LHT	10
-1	LHT LHT LHT	10
0	LHT LHT LHT	10
1	LHT LHT LHT LHT	11
2	LHT LHT LHT	10
3	LHT	2
4		1
Σ	50	

Summe	Häufigkeit	
-4	1	1
-3	1111	1
-2	LHT	3
-1	LHT LHT LHT	15
0	LHT LHT LHT LHT	12
1	LHT LHT LHT	11
2	LHT	5
3		0
4	11	2
Σ	50	

Summe	Häufigkeit	
-4	1111	0
-3	LHT LHT LHT	4
-2	LHT LHT LHT LHT	14
-1	LHT LHT LHT LHT	13
0	LHT LHT LHT	12
1	LHT	5
2	LHT	2
3		0
4		0
Σ	50	

Summe	Häufigkeit	Anteil
-4	2	0,010
-3	11	0,055
-2	27	0,135
-1	46	0,230
0	46	0,230
1	39	0,195
2	21	0,105
3	5	0,025
4	3	0,015
Σ	200	1

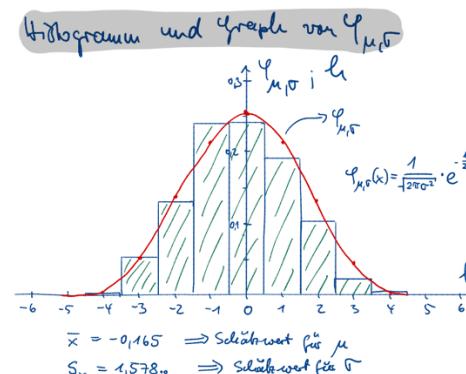
Kenngrößen ($n=200$)

$$\bar{x} = -0,165$$

$$s_x = 1,578$$

↳ Schätzwert für μ

$\varphi_{\mu, \sigma}$: Normalverteilung mit den Parametern $\mu = -0,165$ und $\sigma = 1,578$



Normalverteilung - Einstieg

Von der Realität zum Modell

Wir haben mithilfe des Befehls `sum(randInt(-1,1,4))` die Summe von vier Zufallszahlen aus der Menge $\{-1,0,1\}$ erzeugt. Insgesamt haben wir das 652-mal durchgeführt. Die Ergebnisse sind in der folgenden Tabelle zusammengestellt:

Summe	-4	-3	-2	-1	0	1	2	3	4
Anzahl	8	32	83	140	148	141	66	28	6
Anteil	0,0123	0,0491	0,1273	0,2147	0,2270	0,2163	0,1012	0,0429	0,0092

Kenngrößen: $n = 652$; $\bar{x} = -0,0813$; $s_{n-1} = 1,5856$ (gerundete Werte)

Wichtig:

Wir haben die diskreten Werte $-4, \dots, 4$ mithilfe einer Normalverteilung beschrieben. Deshalb sollte man in diesem Fall von einer linken Grenze 0,5 subtrahieren und zu einer rechten Grenze 0,5 addieren, um die Balken an den Rändern auch zu berücksichtigen.

Wenn wir aber mit einer Zufallsgröße X modellieren, die als normalverteilt angenommen wird, entfällt dieses Vorgehen.

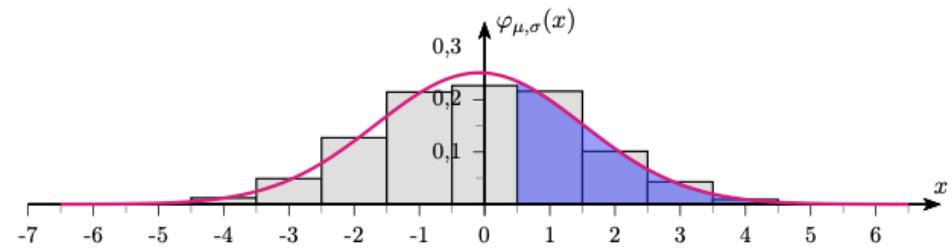
Hätten wir z. B. mit Bauteilen simuliert, die jeden Wert von -4 bis 4 annehmen können, hätten wir die linke bzw. rechte Grenze nicht um 0,5 verringert bzw. erhöht.

Die Funktion $\varphi_{\mu,\sigma}$ mit

$$\varphi_{\mu,\sigma}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-0,5 \cdot (x-\mu)^2 / \sigma^2}, x \in \mathbb{R},$$

beschreibt unsere Daten gut (s. Abbildung). Um unsere Frage (s. o.) zu beantworten, muss man Flächeninhalte bestimmen, also integrieren. Damit auch der Balken über der Eins berücksichtigt wird, muss man als linke Integrationsgrenze 0,5 nehmen. Das ist etwas verwirrend. Auch die rechte Grenze ∞ ist sicherlich ungewohnt. Das ist der Modellierung geschuldet. Damit man mit der Funktion $\varphi_{\mu,\sigma}$ Wahrscheinlichkeiten durch Flächenbestimmung erhält, muss die gesamte Fläche von $-\infty$ bis $+\infty$ den Wert 1 haben. Darum benötigt man übrigens den Vorfaktor $1/\sqrt{2\pi\sigma^2}$. Man sollte immer Realität und Modell sowie die Grenzen des Modells im Hinterkopf haben.

$$P(X \geq 0,5) := \int_{0,5}^{\infty} \varphi_{\mu,\sigma}(x) dx = 0,3570 \dots \text{ (Modellwert)} \quad \text{Realwert (s. o.): } \frac{241}{652} = 0,3696 \dots$$



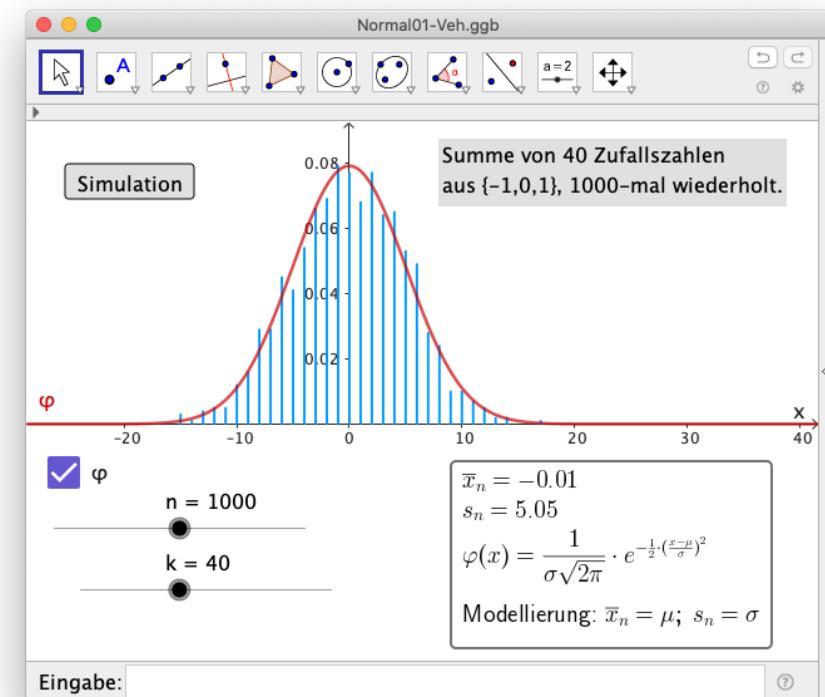
Normalverteilung – Einstieg mit GeoGebra

1.1 Simulation und Modellierung

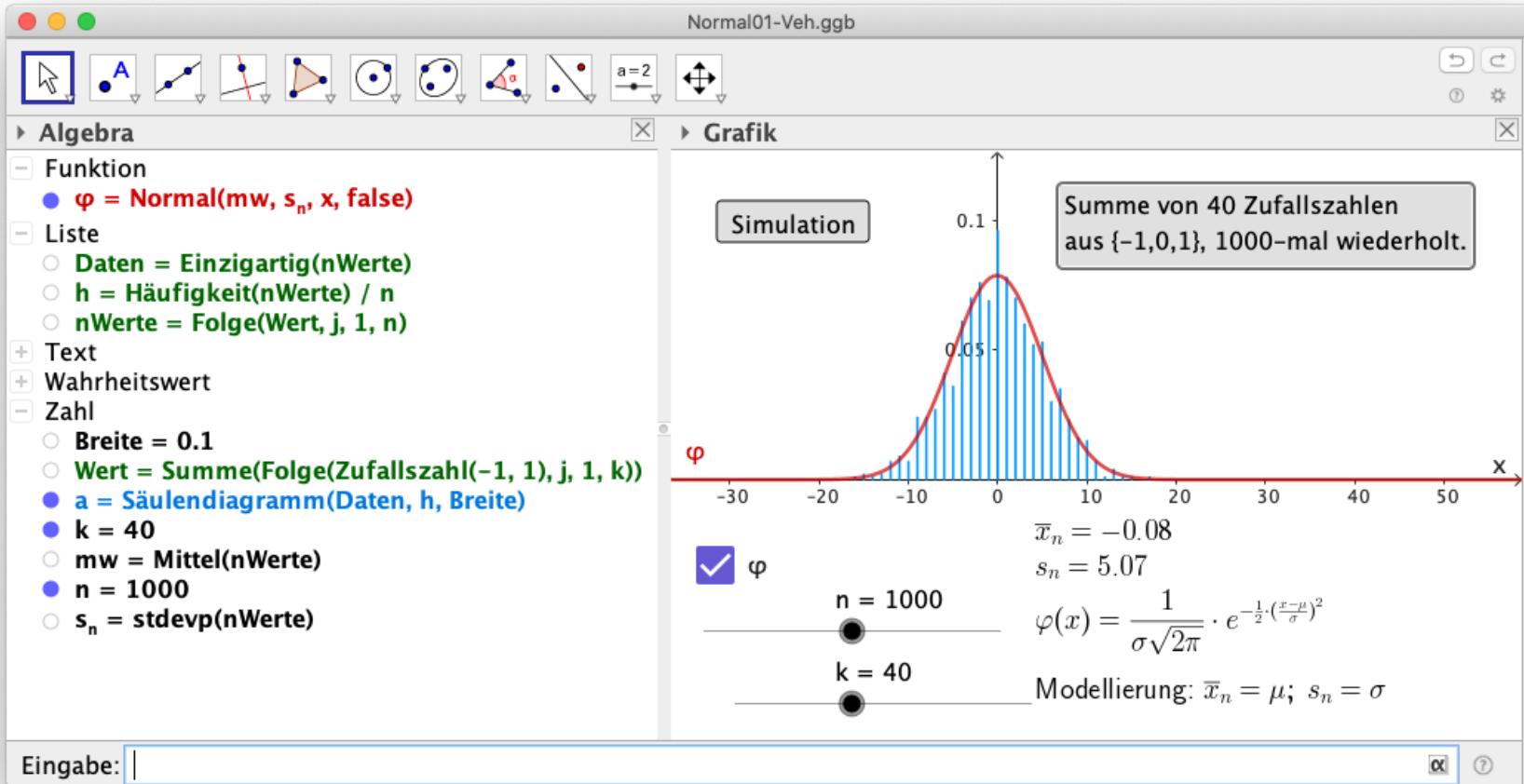
Im Folgenden werden wir uns mit einer einfachen Modellierung dieser wichtigen Verteilung nähern. Wir gehen davon aus, dass das Füllgewicht bei der Verpackungsmaschine nur von vier Bauteilen abhängt, die unabhängig voneinander entweder exakt funktionieren, oder aber eine Abweichung von $\pm 1\text{ g}$ bewirken. Dabei soll das Auftreten dieser drei Fälle gleichwahrscheinlich sein.

Ziel ist es, eine (stetige) Wahrscheinlichkeitsverteilung zu finden, mit der man dann z. B. die folgenden Fragen beantworten kann:

- Wie groß ist die Wahrscheinlichkeit, eine Packung zu erhalten, die 2 g vom Normgewicht abweicht?
 - Welcher Anteil an Tüten wird bei einer Lieferung von 1000 Stück mit 95 %-Wahrscheinlichkeit ein Gewicht von mindestens 501 g aufweisen?
1. Zuerst sollen Sie Häufigkeitsverteilungen mithilfe von Simulationen erzeugen, um solche Fragen zu beantworten. Natürlich nur mit Wahrscheinlichkeitsangaben.
 2. Um mit dem Modell der Normalverteilung Wahrscheinlichkeiten zu bestimmen, benötigt man nur zwei Kenngrößen, den Erwartungswert und die Standardabweichung der zugehörigen Zufallsgröße X . Damit ist eine Normalverteilung eindeutig bestimmt. Simulationen liefern Mittelwerte und empirische Standardabweichungen. Mit diesen Kenngrößen werden Schätzungen für die Modellgrößen $E(X)$, $V(X)$ und $\sigma(X)$ vorgenommen.
 3. Zum Schluss werden Sie mit einem theoretischen Modell diese drei Kenngrößen bestimmen und mit den Schätzgrößen vergleichen.



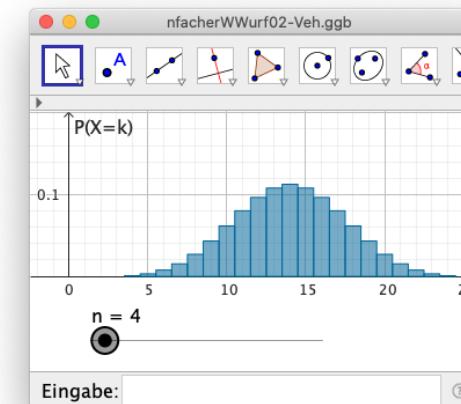
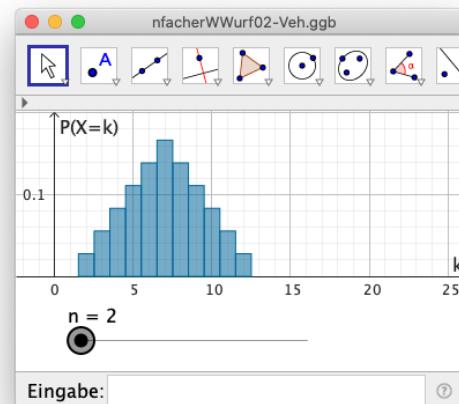
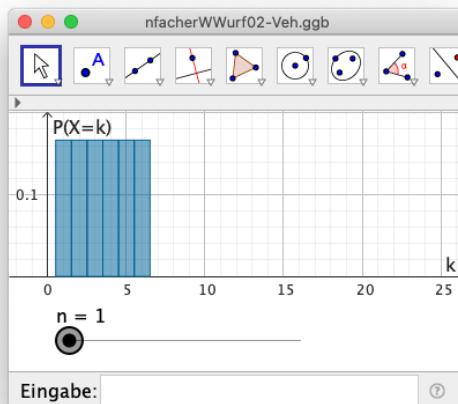
Normalverteilung – Einstieg mit GeoGebra



Normalverteilung – Einstieg mit n-fachem Würfelwurf

Ein fairer Würfel wird n mal in unabhängiger Folge geworfen. Sei S_n die Augensumme aus allen Würfeln.

$$P(S_n = k) = ? \quad (k \in \{n, n+1, \dots, 6n\})$$



Sinnvoller CAS-Einsatz – erzeugende Funktionen in der Schule

Ausmultiplizieren liefert $P(S_n = k) = \frac{\text{Koeffizient vor } x^k}{6^k}$.

WIESO?

n-facher Würfelwurf

f(x):=Summe(x^(j),j,1,k)^n
→ f(x) := $(x^6 + x^5 + x^4 + x^3 + x^2 + x)^4$

Multipliziere(f(x))
→ $x^{12} + 2x^{11} + 3x^{10} + 4x^9 + 5x^8 + 6x^7 + 5x^6 + 4x^5 + 3x^4 + 2x^3 + x^2$

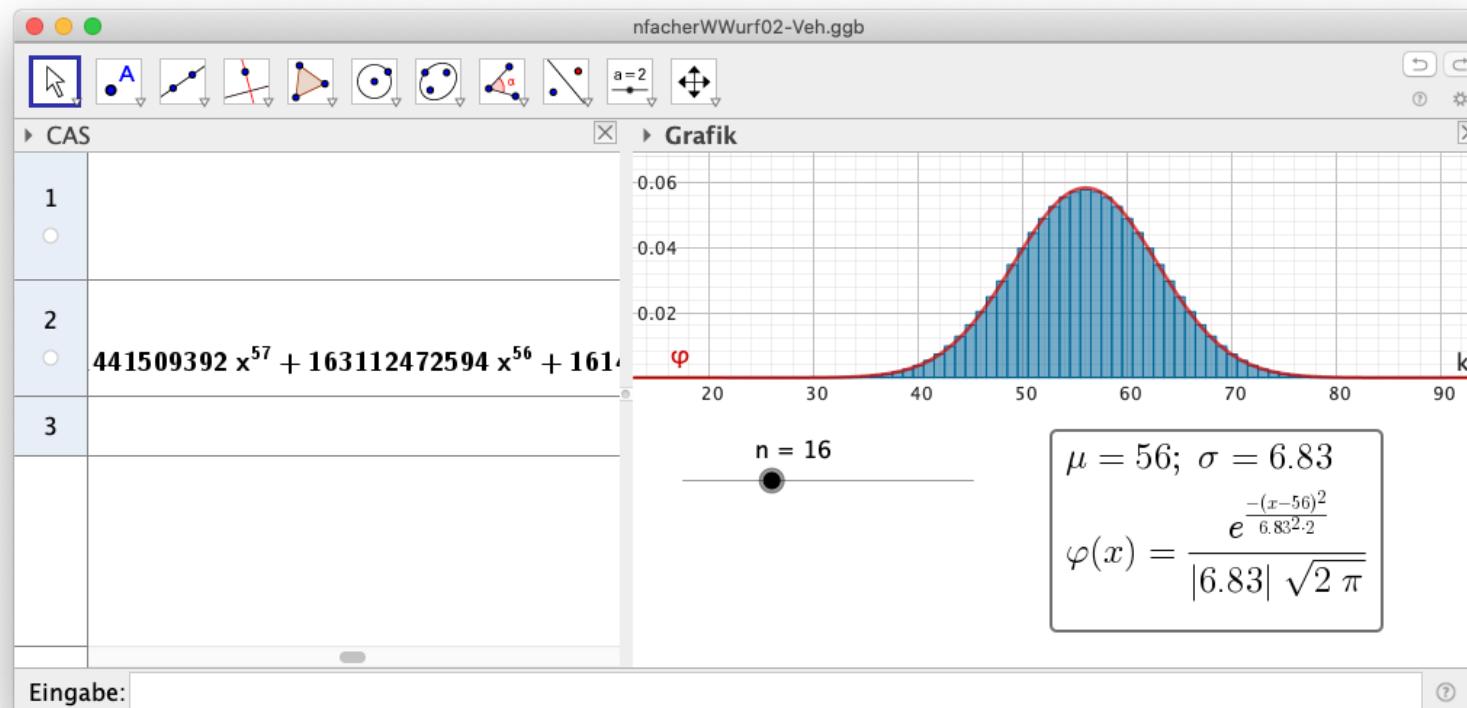
Koeffizienten(f)
→ {1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1, 0, 0}

Eingabe:

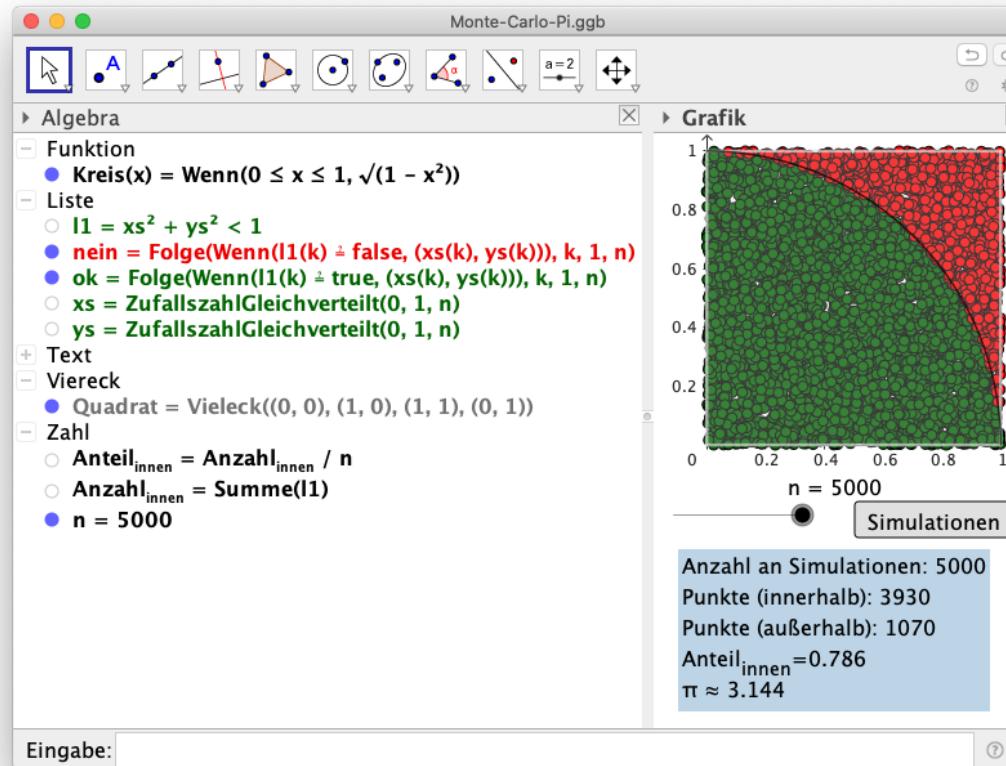
Normalverteilung als Approximation beim n-fachen Würfelwurf

Ein fairer Würfel wird n mal in unabhängiger Folge geworfen. Sei S_n die Augensumme aus allen Würfen.

$$P(S_{16} = 56) = \frac{163.112.472.584}{6^{16}}$$



Eine Näherung für π mit einer Simulation



Monte-Carlo-Simulation

Der Zufall gibt sein Muster preis

Der zerbrochene Stab und ein Dreieck

Ein 1m-Stab wird an zwei zufällig ausgewählten Stellen z_1 und z_2 geschnitten.
Gesucht ist die Wahrscheinlichkeit, dass sich aus den drei Teilstücken ein Dreieck bilden lässt?

1. Fall: $0 < z_1 < z_2 < \frac{1}{2}$ (z.B. d.h. $z_1 < z_2$)

geht nicht
grün + rot < blau

2. Fall: $z_1 < \frac{1}{2} \wedge z_2 > \frac{1}{2} \wedge z_2 - z_1 < \frac{1}{2}$
 $(z_2 - z_1 < \frac{1}{2})$
 $\frac{1}{2} < z_1 < z_2$, sonst kein Δ möglich

✓

3. Fall: $\frac{1}{2} < z_1 < z_2 < 1$

geht nicht
grün > rot + blau

Zwei Wege

Bleibt u. Papiro → Simulation →

zentrale Idee: Darstellungswechsel
 z_1, z_2 als Koordinaten eines Punktes identifizieren $v(z_1, z_2)$

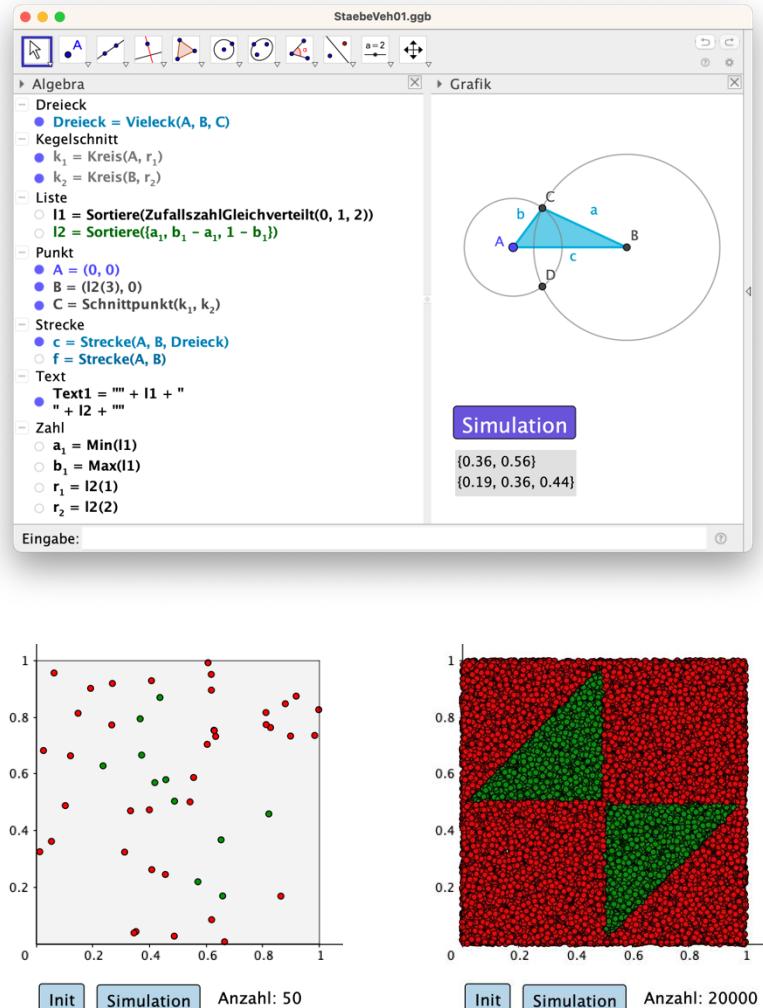
$z_1 < \frac{1}{2} \wedge z_2 > \frac{1}{2} \wedge z_2 = z_1 + \frac{1}{2}$
 $z_1 < \frac{1}{2} \wedge z_2 > \frac{1}{2} \wedge z_2 = z_1 - \frac{1}{2}$
 $z_1 - z_2 > \frac{1}{2}$

$\Rightarrow z_1 < \frac{1}{2} \wedge z_2 < \frac{1}{2}$

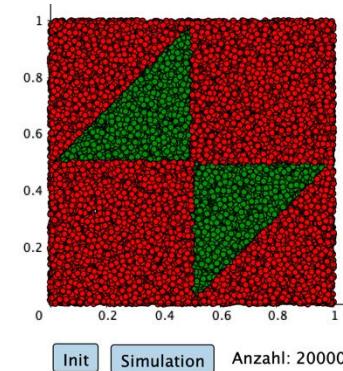
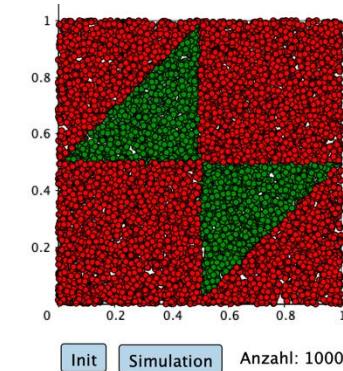
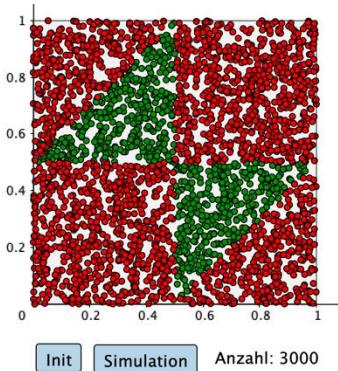
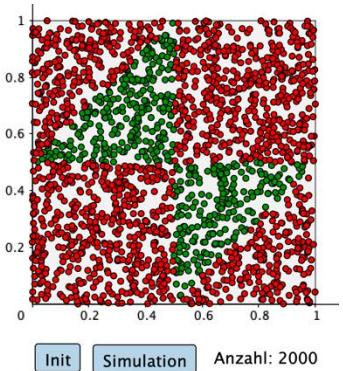
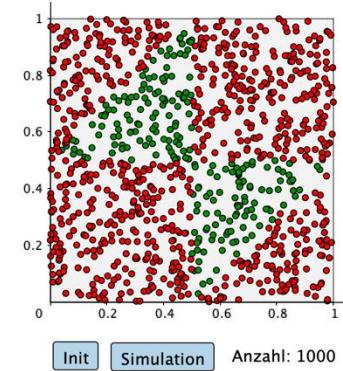
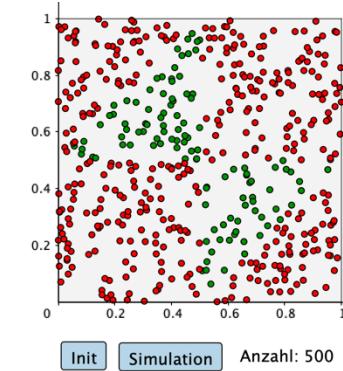
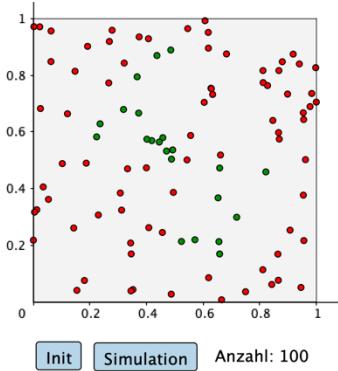
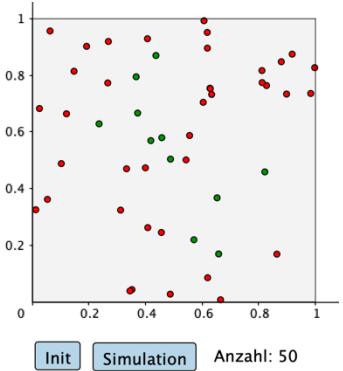
II: $z_1 < \frac{1}{2} \wedge z_2 > \frac{1}{2}$
 $z_1 - z_2 < \frac{1}{2}$

I: $z_1 - z_2 > \frac{1}{2}$

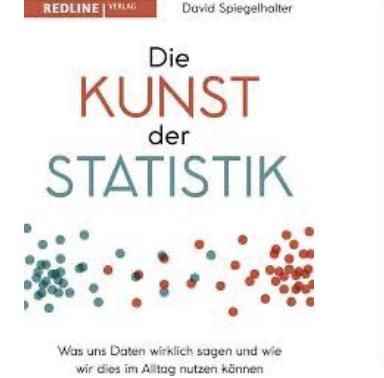
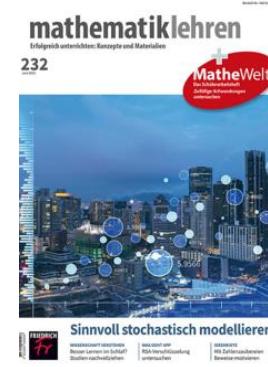
Monte-Carlo-Simulation



Der Zufall gibt sein Muster preis



Literatur – eine kleine Auswahl

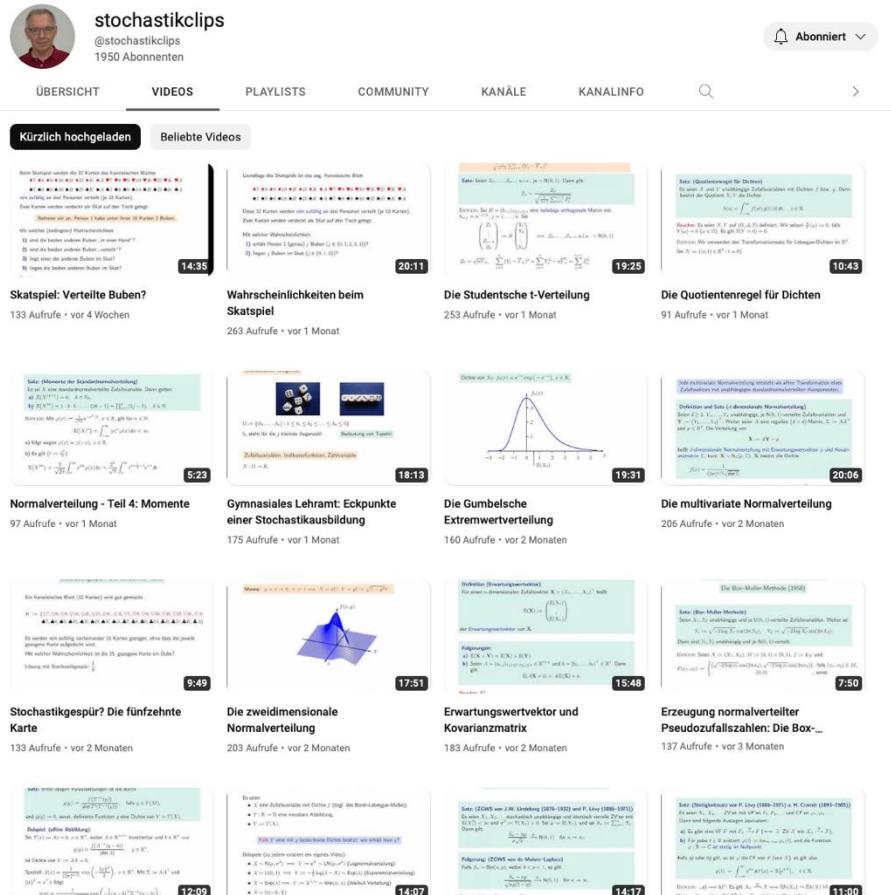


Paradigmen

- Pflege einen passenden Wahrscheinlichkeitsbegriff
- Trenne Modell und Realität messerscharf und konsequent
- Untersuche Zufallsschwankungen statt sie wegzuwünschen
- Stelle authentische Probleme ins Zentrum
- Nutze den „didaktischen Dreischritt“ Spekulieren - Experimentieren - Reflektieren

Zwei interessante Adressen im Internet

9. Stochastik	Deckblatt	Roter Faden	In Ni nicht verlangt	21, 22, 23, 24, 25, 26, 27, 34, 35, 36, 37, 38, 39, 41, 42, 47
1. Wahrscheinlichkeitsrechnung	Einführung	Kurzfassung		
2. Aufgaben,	Wahrscheinlichkeit			
3. Lösungen der Aufgaben				
4. Aufgaben zur Kombinatorik	Fakultät, Binomialkoeffizient Sitzordnungen			
5. Pascalsches Dreieck				
6. Formeln, k Elemente auf n Plätze,	Aufgaben zur Kombinatorik			
7. Binomialkoeffizient				
8. Gegenwahrscheinlichkeit				
9. Pfadwahrscheinlichkeiten				
10. Einfache Aufgaben				
11. Bedingte Wahrscheinlichkeit	Folien	Video		
12. Erwartungswert Altersstruktur	Folien			
13. Standardabweichung				
14. Aufgaben				
15. Bernoulli-Kette, "mindestens ein Treffer"-Aufgabe, p-Wert	Folien			
16. Binomialverteilung, Histogramme				
17. Stichprobe, mit und ohne Zurücklegen,	Hypergeometrische Verteilung			
18. Unabhängigkeit Version 2	Version 3	Folien		
19. Erwartungswert der Binomialverteilung				
20. Varianz der Binomialverteilung, $V(X+Y) = V(X) + V(Y)$				
21. Sigma-Umgebungen	Folien			
22. Permutationstest				
23. Testen von Hypothesen mit Sigma-Umgebungen				
24. Alternativtest				
25. Hypothesentest, Fehler 1. und 2. Art, Ergänzungen				
26. Ablehnungsbereich mit Tabellen und GTR				
27. Signifikanztest, Gütekriterium				
28. Tschebyschow-Ungleichung, Gesetz der großen Zahlen				
29. Standardisierung				
30. Normalverteilung, Standardnormalverteilung				Folien Video
31. Konfidenzintervall Stichprobenumfang, Prognose- und Vertrauensintervall, Weiteres				Folien Video Wald Video Wilson Prognoseintervall
32. Daten analysieren				
33. Axiomensystem von Kolmogorow, Unabhängigkeit, Vier-Felder-Tafel, Mengenoperationen				
34. Unabhängigkeit zweier Merkmale				
35. Irrfahrt				
36. Absorbierende Markow-Ketten, Mason-Regel				
37. Absorbierende Markow-Ketten, Verweilzeiten				
38. Absorbierende Markow-Ketten, Beispiel				
39. Absorbierende Markow-Ketten, Alternative				
40. Erwartungstreue				
41. Schätzgrößen für Erwartungswert und Varianz				
42. Simulation Monte-Carlo-Methode				
43. Poisson-Verteilung				
44. Erwartungswert einer stetigen Zufallsvariablen				
45. Stetige Verteilungen, Exponentialverteilung				
46. Stetige und diskrete Zufallsvariablen				
47. Stochastik-Aufgaben Glücksrad-Aufgaben				
48. Regression, Korrelation				
49. Zuverlässigkeit eines Systems				
50. Wissens-Überprüfung Stochastik				
51. interaktive Merkhilfe Stochastik				
52. Test				Stochastik Folien



<http://groolfs.de>

<https://www.youtube.com/@stochastikclips/videos>

Das war es.

**Ich hoffe, es war nicht
(ganz) umsonst.**

**Viel Freude und Erfolg beim
Unterrichten.**

Konfidenzintervalle – aus meinem (ehemaligen) Unterricht

Methode I - WILSON-Konfidenzintervall

Sichtweise von p aus:

Zu unserem ersten Ansatz ist die folgende Herangehensweise gleichwertig:

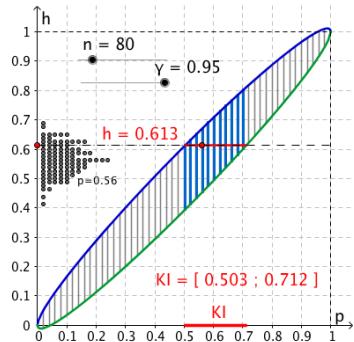
Ein Konfidenzintervall besteht aus allen p^0 s, die h im Prognoseintervall haben.

Die Grafik verdeutlicht dies für die Ermittlung des 95%-Konfidenzintervalls. Zu einigen Werten für p werden die zugehörigen 95%-Prognoseintervalle als vertikale Strecken gezeichnet.

Die Horizontale zu h schneidet diese Intervalle. Dadurch ergeben sich (näherungsweise) die beiden Grenzen des Konfidenzintervalls $[p_u; p_o]$.

Ermittlung durch Schnittpunktbestimmung:

$$h = p_o - 2 \cdot \frac{p_o \cdot (1 - p_o)}{n} ; h = p_u + 2 \cdot \frac{p_u \cdot (1 - p_u)}{n}$$



Interpretation eines Konfidenzintervalls

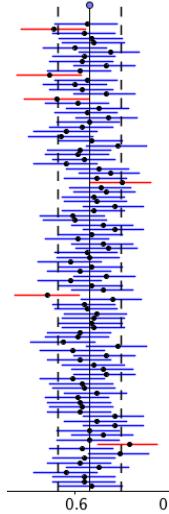
Noch einmal: Nachdem die Daten gesammelt und ein Konfidenzintervall berechnet wurde, liegt p entweder in diesem Intervall oder nicht.

Sie dürfen nicht sagen, dass Sie zu 95% sicher sind, dass p im Intervall enthalten ist. Sie können jedoch zu 95% dem Vorgang Vertrauen schenken, mit dem Sie die Daten erheben und das 95%-Konfidenzintervall ermitteln. Sie wissen, dass damit Intervalle resultieren, die p in ca. 95% der Fälle enthalten werden.

Mehr ist nicht drin - leider. Sie können auch daneben liegen. In etwa 5% aller Fälle wird dies eintreten.

Die nebenstehende Abbildung zeigt eine Simulation von 100 Intervallen ($\gamma = 95\%$, $n = 200$). Eigentlich ist p ja unbekannt, hier wird mit $p = 0,63$ simuliert. Insgesamt überdecken 6 Intervalle den Anteil p nicht.

Frage: Was passiert, wenn die Werte von γ und n verändert werden?



Normalverteilung – Aufgaben mit CAS/GTR

Beispiel:

Die Körpergröße des Menschen kann für ein Geschlecht betrachtet als normalverteilt angesehen werden (Modellierung - keine Realität). Laut einer Statistik des sozio-ökonomischen Panels (SOEP) aus dem Jahr 2006 liegt der Erwartungswert μ der Durchschnittsgröße bei Frauen in Deutschland bei 165,4 cm, die Standardabweichung σ bei 4,5 cm.

Damit können nun Bereichswahrscheinlichkeiten berechnet und als Anteile interpretiert werden, etwas so:

Die Zufallsgröße X beschreibt die Größe (in cm) von Frauen in Deutschland. Es soll mit $\mu = 165,4$ und $\sigma = 4,5$ modelliert werden.

$$P(X \geq 175) = 0,0164\dots \text{ [TR: } \text{normalcdf}(175, 1E99, 165.4, 4.5)]$$

Ca. 1,6 % aller Frauen in Deutschland sind mindestens 1,75 m groß.

Es gilt $\int_a^{\infty} f(x)dx = 0$. In der Modellwelt ist also $P(X = 175) = 0$. Damit ist $P(X > 175) = 0,0164\dots$ auch richtig. In Worten: Ca. 1,6 % aller Frauen in Deutschland sind größer als 1,75 m.

$$P(156,4 \leq X \leq 174,4) = 0,9544\dots \text{ [TR: } \text{normalcdf}(156.4, 174.4, 165.4, 4.5)]$$

Ca. 95,4 % aller Frauen in Deutschland sind mindestens 156,4 cm und höchstens 174,4 cm groß.

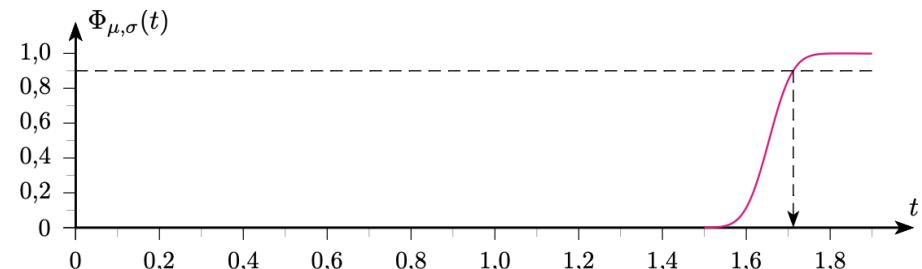
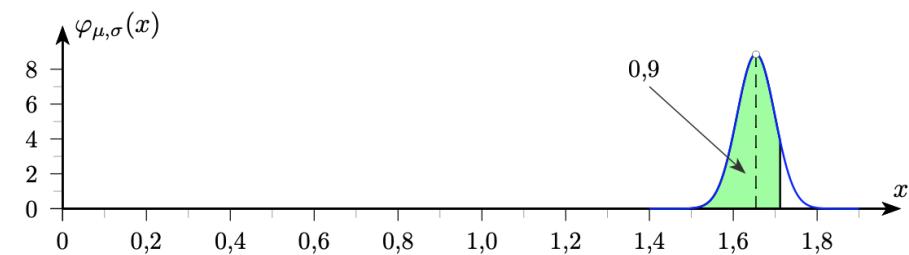
Das ist gerade die 2σ -Umgebung vom Erwartungswert 165,4 cm.

Gesucht wird k mit $P(X \leq k) = 0,9$.

$$\text{Ergebnis: } k = 171,1669\dots \text{ [TR: } \text{invnorm}(0.9, 165.4, 4.5)]$$

Oder durch Schnittpunktbestimmung: $y_1 = \text{normalcdf}(-1E99, X, 165.4, 4.5)$; $y_2 = 0.9$

Die beiden Grafiken veranschaulichen beiden Lösungswege. Es wurde als Einheit Meter statt Zentimeter benutzt.



Normalverteilung – Werte schätzen ohne Rechner

Noch einmal:

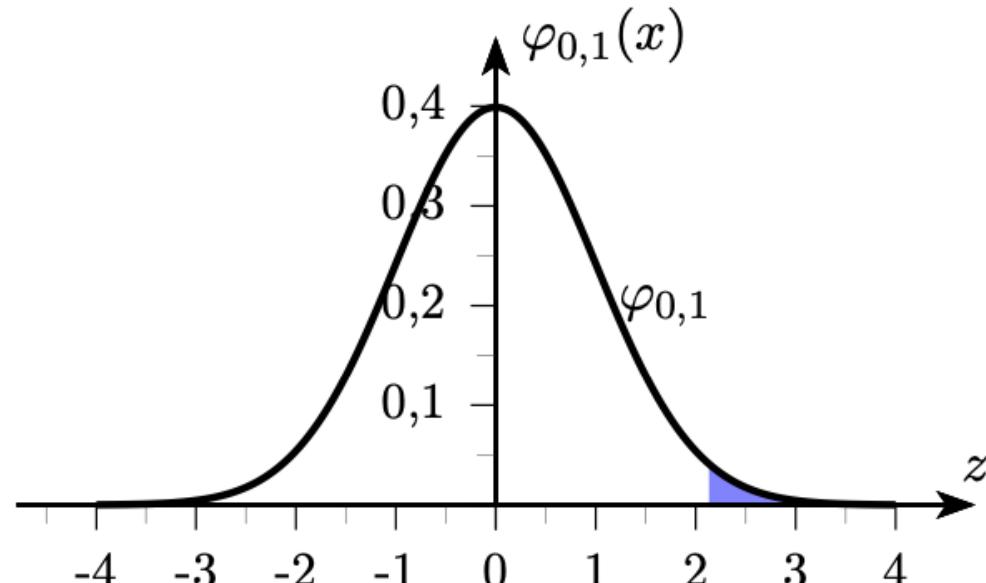
Die Zufallsgröße X beschreibt die Größe (in cm) von Frauen in Deutschland. Es soll mit $\mu = 165,4$ und $\sigma = 4,5$ modelliert werden.
ges.: $P(X \geq 175)$

Die Standardisierung hat Vorteile

$$Z = \frac{X - \mu}{\sigma}$$

$$\text{Neue Grenze: } z_1 = \frac{175 - 165,4}{4,5} \approx 2,13$$

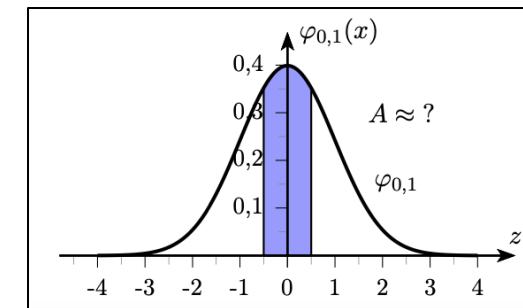
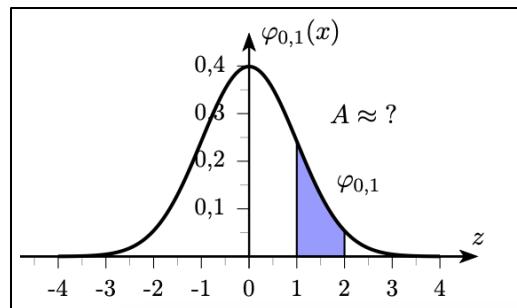
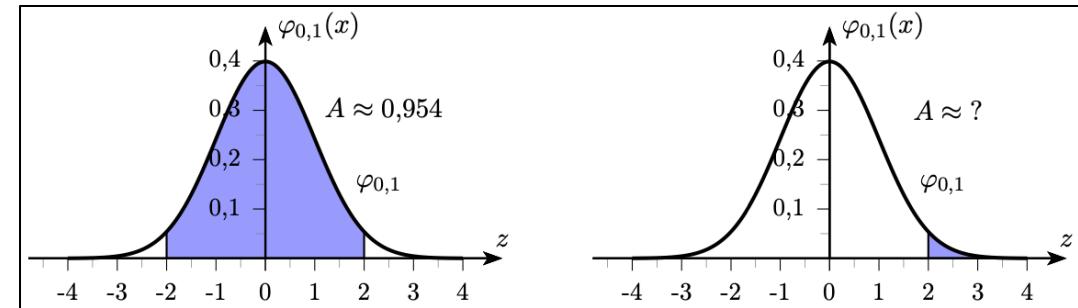
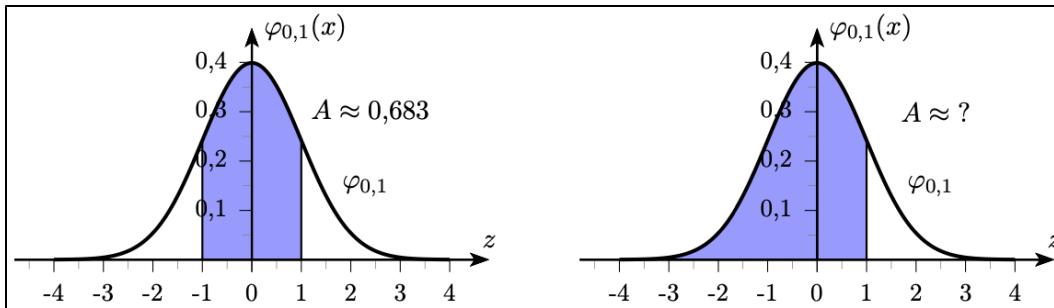
$$P(X \geq 175) = P(Z \geq 2,13) \approx 0,016$$



Normalverteilung – Werte schätzen

Aufgabe: Wahrscheinlichkeiten ohne Rechnereinsatz

Bestimmen Sie ohne Rechnereinsatz die Flächeninhalte und somit die gesuchten Bereichswahrscheinlichkeiten.



Normalverteilung – Aufgaben

Aufgabe 1: Kenngrößen gesucht

In Abbildung 1 sind die Graphen dreier Normalverteilungen dargestellt. Der Hoch- und die beiden Wendepunkte sind jeweils markiert. Ermitteln Sie mit diesen Angaben jeweils mögliche Werte für μ und σ .

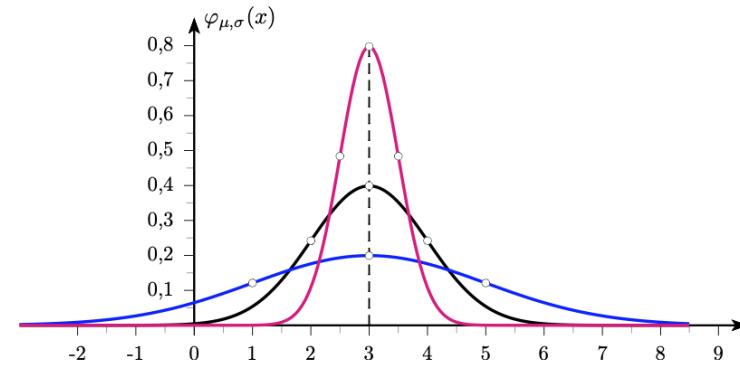


Abbildung 1: μ und σ gesucht

Aufgabe 2: Graphen zuordnen

In Abbildung 2 sind die drei Graphen der Verteilungsfunktionen zu den Dichtefunktionen aus Aufgabe 1 dargestellt.
Ordnen Sie den Graphen die jeweilige Dichtefunktion zu.

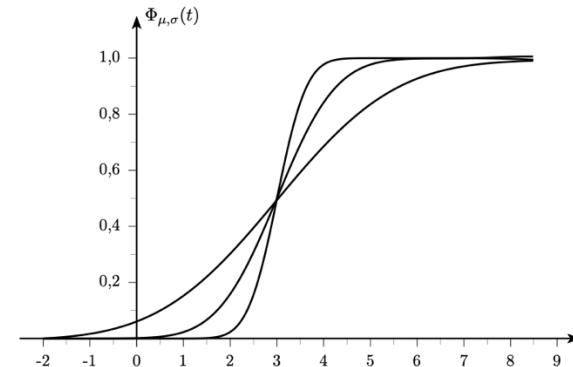


Abbildung 2: Zuordnen - wer ist wer?

Normalverteilung – Grundaufgaben

Aufgabe 4: Grundaufgabe - P gesucht

Gegeben ist die normalverteilte Zufallsgröße X mit $\mu = 20$ und $\sigma = 3$.

Bestimmen Sie die folgenden Wahrscheinlichkeiten.

- | | | |
|---------------------------|---------------------------|-------------------|
| a) $P(X \leq 18)$ | b) $P(X = 18)$ | c) $P(X \geq 22)$ |
| d) $P(17 \leq X \leq 23)$ | e) $P(14 \leq X \leq 26)$ | f) $P(X \leq 20)$ |

Aufgabe 5: Grundaufgabe - μ gesucht

Erstellen Sie zuerst zu jeder Problemstellung eine aussagekräftige Skizze.

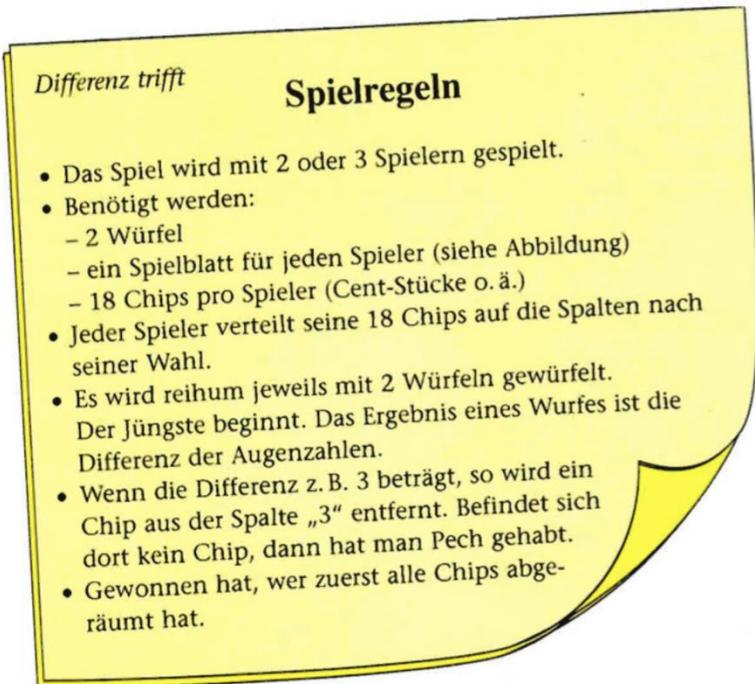
1. Gegeben ist die normalverteilte Zufallsgröße X mit $\sigma = 5$ und $P(X \leq 70) = 0,1586$
Ermitteln Sie den Wert von μ auf eine Nachkommastelle genau.
2. Gegeben ist die normalverteilte Zufallsgröße X mit $\sigma = 3$ und $P(X \geq 60) = 0,254$
Ermitteln Sie den Wert von μ auf eine Nachkommastelle genau.

Aufgabe 6: Grundaufgabe - σ gesucht

Erstellen Sie zuerst zu jeder Problemstellung eine aussagekräftige Skizze.

1. Gegeben ist die normalverteilte Zufallsgröße X mit $\mu = 50$ und $P(X \leq 53) = 0,658$
Ermitteln Sie den Wert der zugehörigen Standardabweichung σ auf eine Nachkommastelle genau.
2. Gegeben ist die normalverteilte Zufallsgröße X mit $\mu = 150$ und $P(X \geq 155) = 0,254$
Ermitteln Sie den Wert der zugehörigen Standardabweichung σ auf eine Nachkommastelle genau.

Differenz trifft – eine mögliche Einführung und zwei konkurrierende Modelle



Entnommen aus:

Lergenmüller, A., Schmidt, G. (2005): Mathematik Neue Wege NDS, Schroedel. S. 103

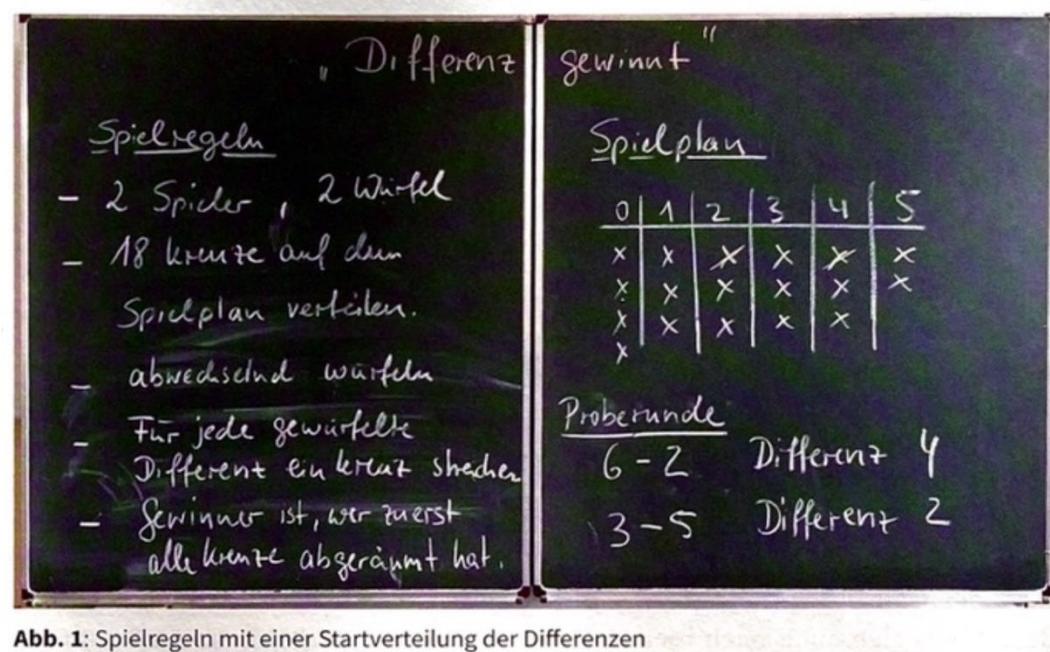


Abb. 1: Spielregeln mit einer Startverteilung der Differenzen

Entnommen aus:

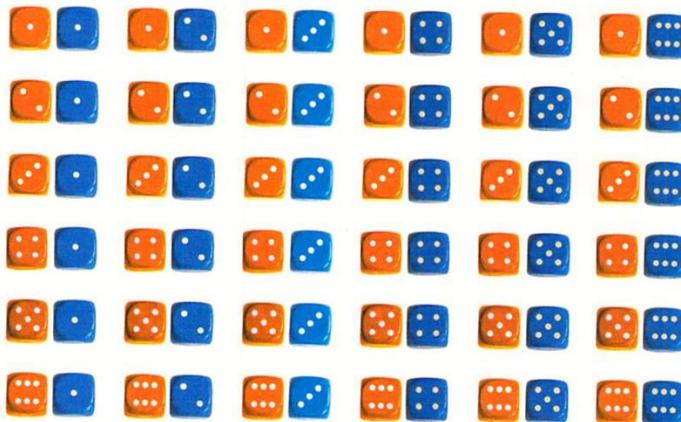
Prömmel, A., Vehling, R. (2022): Die Differenz 5 kommt ja nur einmal vor!. In: ml 232. S 32-35. Friedrich-Verlag

Differenz trifft – eine mögliche Einführung und zwei konkurrierende Modelle

Kommt die Differenz 5 einmal oder zweimal vor?

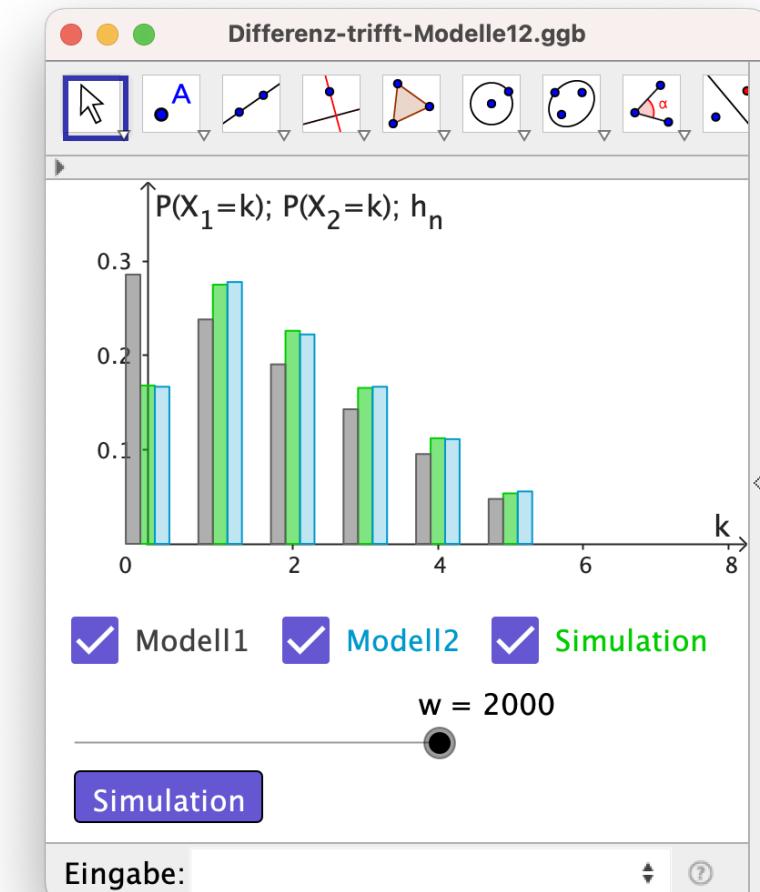
Modell 1: nicht unterscheidbar – 21 Ergebnisse

Modell 2: unterscheidbar – 36 Ergebnisse



Färben oder simulieren?
Soll doch die Natur
entscheiden, welches
Modell geeigneter ist.

Hier irrte schon Leibniz.

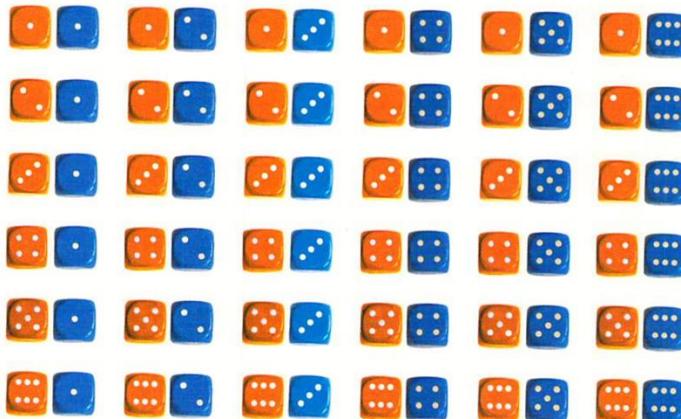


Differenz trifft – eine mögliche Einführung und zwei konkurrierende Modelle

Kommt die Differenz 5 einmal oder zweimal vor?

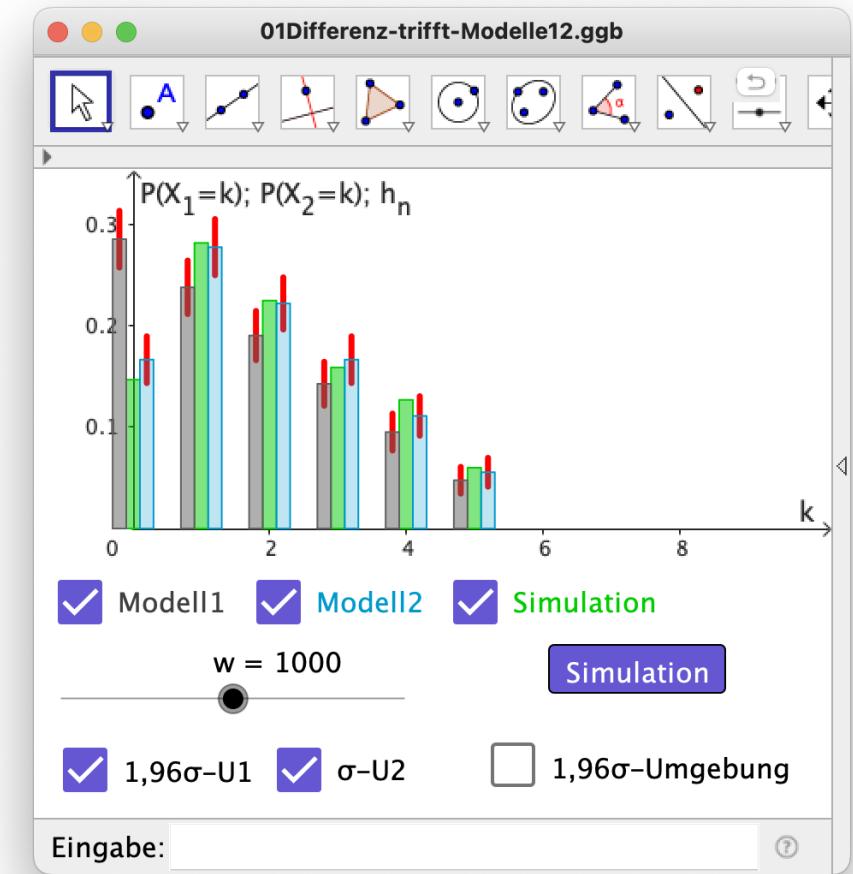
Modell 1: nicht unterscheidbar – 21 Ergebnisse

Modell 2: unterscheidbar – 36 Ergebnisse



Färben oder simulieren?
Soll doch die Natur
entscheiden, welches
Modell geeigneter ist.

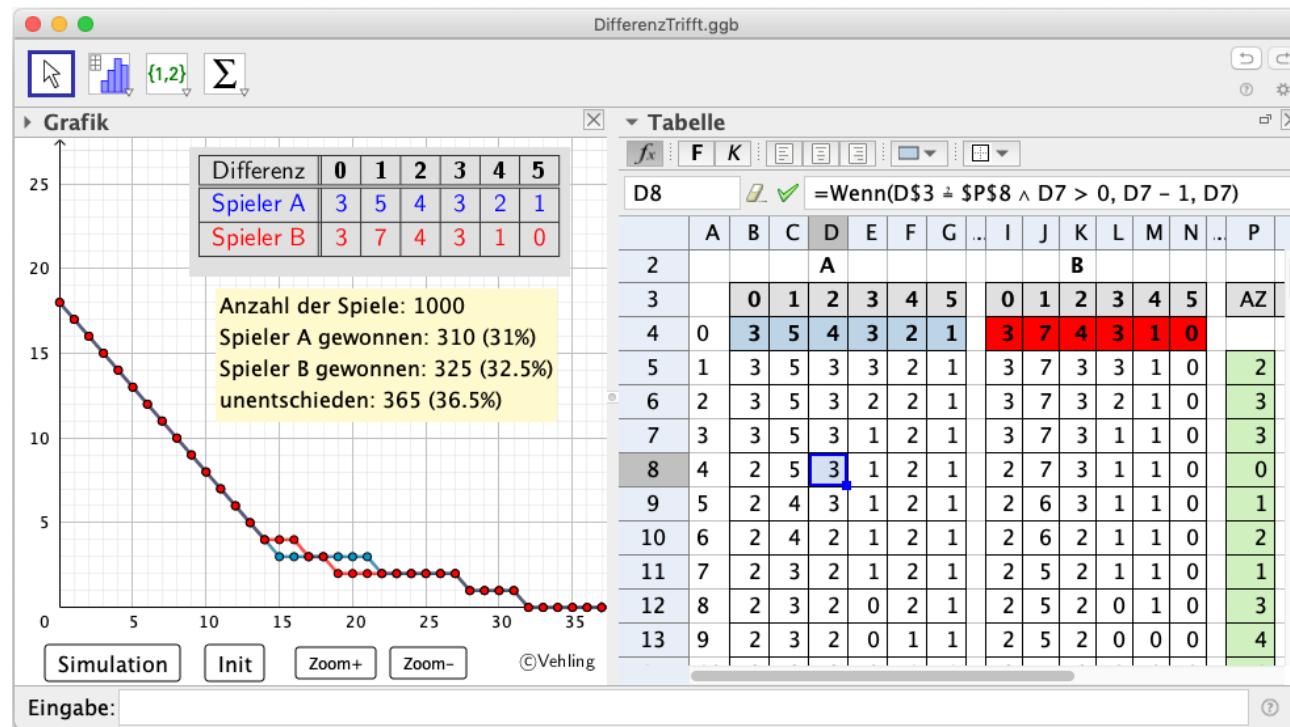
Hier irrte schon Leibniz.



Differenz trifft – Vorsicht bei vorschnellen Folgerungen

Vorsicht: Wenn es um Strategien geht, wird es deutlich komplizierter.

Das Verteilen der Kreuze mithilfe der zugehörigen Wahrscheinlichkeitsverteilung (Proportionalstrategie) ist auf lange Sicht nicht immer die beste Strategie.



Die Würfel von Efron und Baumdiagramme

1 „Höhere Augenzahl gewinnt!“
Für das folgende Spiel benötigt man zwei Würfel. Beklebe jede Seite mit Papier und schreibe die entsprechenden Punktzahlen auf die Würfel.

(A)  (B) 

Gespielt wird zu zweit. Jeder der Spieler erhält einen Würfel. Dann wird gegeneinander gewürfelt. Gewonnen hat der Spieler, der die höhere Augenzahl erzielt.
a) Wenn du die Wahl hättest, welchen der beiden Würfel würdest du wählen?
b) Spiele das Spiel zehnmal. Notiere, mit welchem Würfel gewonnen wurde. Übertrage dazu folgende Tabelle ins Heft.

Entnommen aus: Neue Wege Bd 8. S. 134. Westermann Schroedel. 2015

Wolfgang Riemers und meine Thesen:

- Stochastik lebt von Experimenten
- Experimente sollen echte Fragen beantworten
 - Spekulieren / Hypothesen formulieren
 - emotionale Beteiligung, Motivation, Staunen
 - mit Hypothesen Prognosen machen
 - Prognosen mit Versuchsergebnissen vergleichen
 - Hypothesen verwerfen oder beibehalten oder verändern
- Wahrscheinlichkeitsrechnung, beschreibende Statistik und beurteilende Statistik wachsen zusammen

Aufzeichnungen aus meinem Uni-Didaktik-Seminar

heterogene Zufallsexperimente mit Baumdiagrammen

möglicher Eintrag: „Höhere Anzahl gewinnt!“

Muttertierschützer

1. Spiel vorstellen: • eine Arbeitsblatt, frei
• Motivplan der beiden Würfel zeigen

2. SuS stellen **Varianztypen** auf - mit (Ketten) Beispielen (Bei solchen werden einige Möglichkeiten ausblenden (Antizipationswerte))

 - Käf., Min., 12-8, Schmett., Kombinationsraum
An Tafel Schmetterlinge (Strichliste) +   $\frac{1}{16}$ 11/11 4.SW.

3. Was können wir machen? → Rechnerisch, Simulation spielen

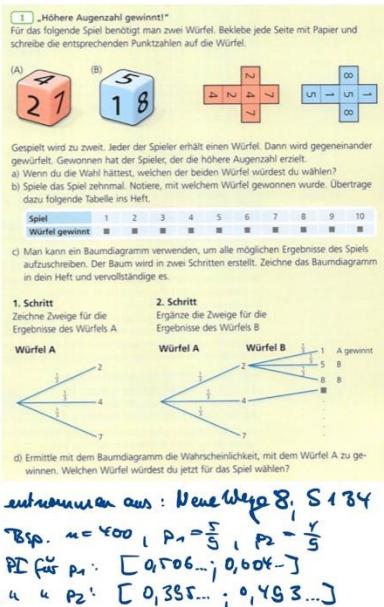
4. **Recht Spalten**, 2 SuS je 30-50x

 - L: welche vorher die Drosophila-Rubrabelle benutzt haben
SuS nutzen ihre Ergebnisse in einer Tabelle
Diskussion der Ergebnisse, Gruppenarbeit (Konsensus)
 - L: Nutzt Ihr eine Klammer (s.o., Strichliste)? fassung
 - n = 400 „reicht“ schon → Gruppen zusammen fassen

5. Können wir das Ergebnis der Gruppen auch schwärfen
Begründen?

Punkt Kombinationen notieren lassen

Ziel: Da ist mehrere Würfel, alle spät



6. Einführung des BD als Strukturierungshilfe

Sie können nicht auf ein Standardprogramm kommen, also muss der L. „anleiten“ gestufte Einführung, z.B. so

 roter Würfel

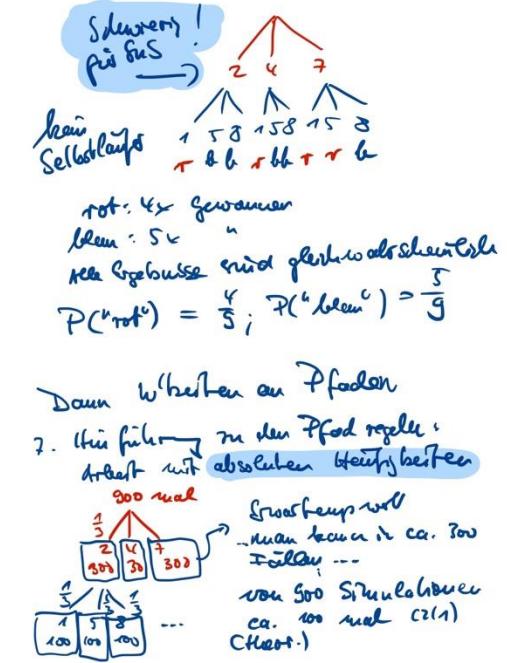
 158 blauer Würfel

158 Wie kann man die Zusammenbringe?
Villerlei weitere Hilfen geben

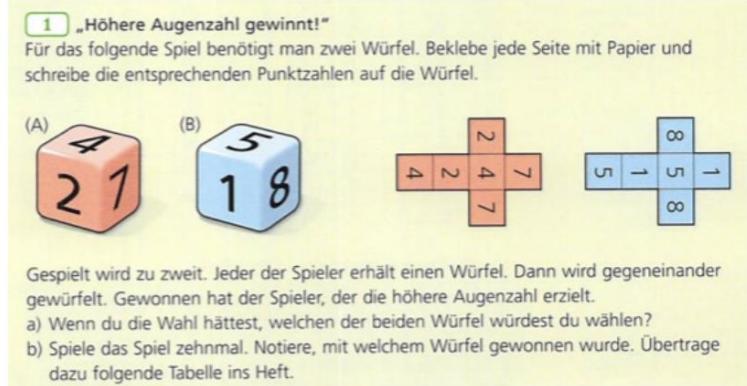
not
blue

¹¹⁸ Sich vervollständigen des RD und schließen: 6 7 2 4 7

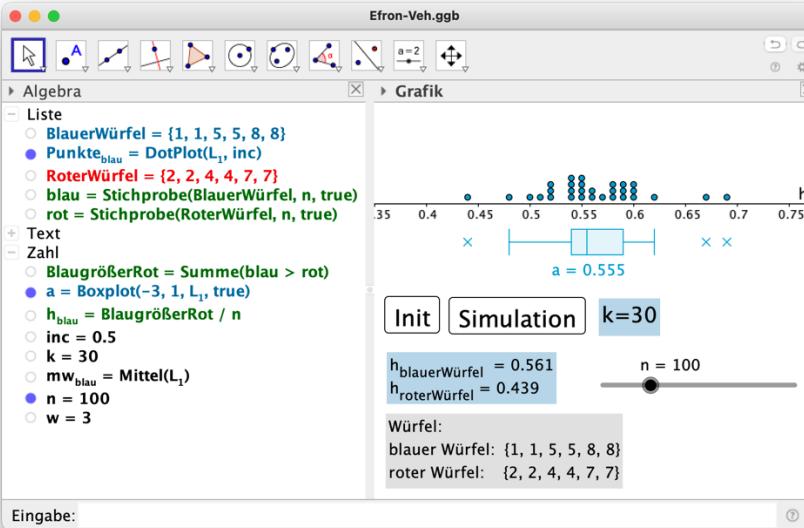
auch möglich - und gut
aber: höchstens 3 Stufen



Die Würfel von Efron – eine Simulation



Entnommen aus: Neue Wege Bd 8. S. 134. Schroedel. 2015



Liste

- BlauerWürfel = {1, 1, 5, 5, 8, 8}
- Punkte_{blau} = DotPlot(L₁, inc)
- RoterWürfel = {2, 2, 4, 4, 7, 7}
- blau = Stichprobe(BlauerWürfel, n, true)
- rot = Stichprobe(RoterWürfel, n, true)

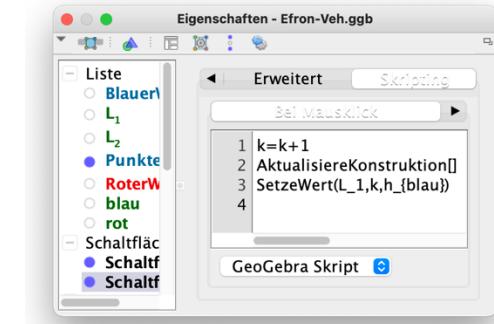
Text

○ BlaugrÖßerRot = Summe(blau > rot)

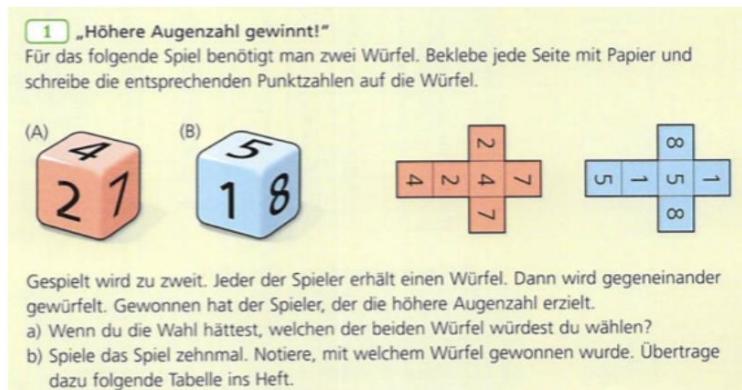
Zahl

- a = Boxplot(-3, 1, L₁, true)
- h_{blau} = BlaugrÖßerRot / n
- inc = 0.5
- k = 30
- mw_{blau} = Mittel(L₁)
- n = 100
- w = 3

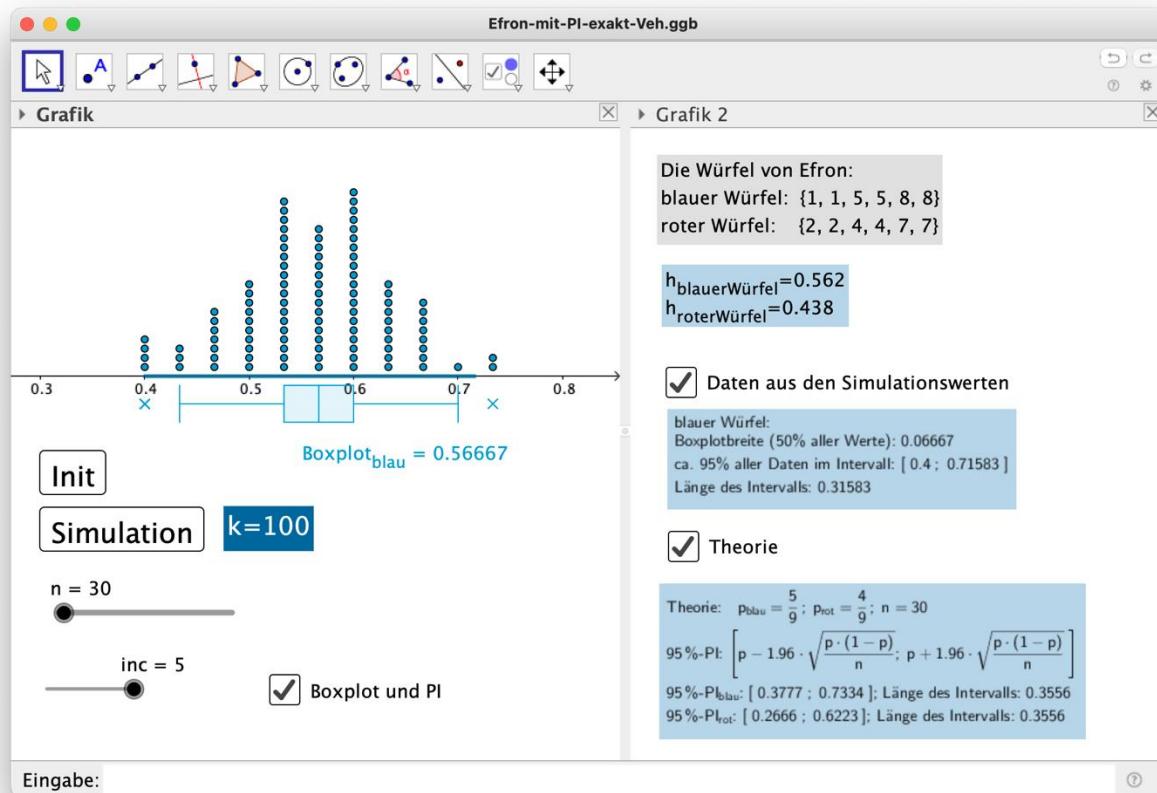
Eingabe:



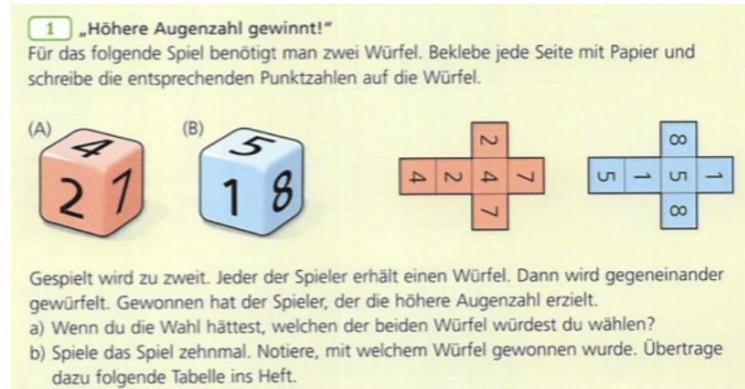
Die Würfel von Efron – eine Simulation



Entnommen aus: Neue Wege Bd 8. S. 134. Schroedel. 2015

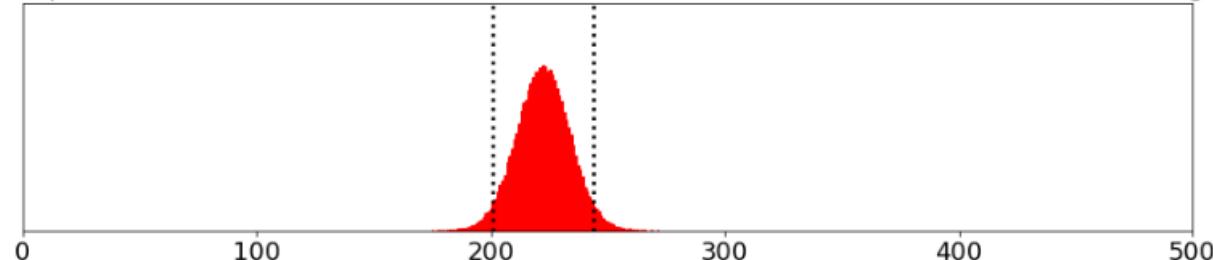


Die Würfel von Efron – eine Simulation mit Python

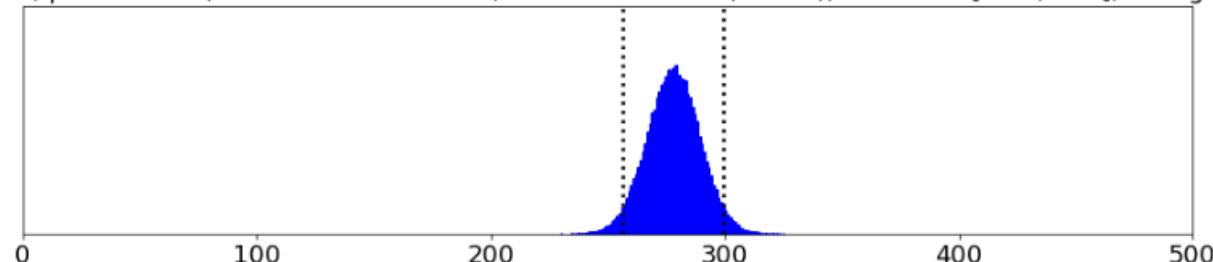


Entnommen aus: Neue Wege Bd 8. S. 134. Schroedel. 2015

n=500, p=0.444445, 100000 Simulationen, 5348 ausserhalb (5.35%), 95%-PI: [201 , 243], Länge: 42



n=500, p=0.555556, 100000 Simulationen, 5368 ausserhalb (5.37%), 95%-PI: [257 , 299], Länge: 42



Minute schätzen – Hypothesentest einmal anders

Mein toller LK im Matheunterricht ...



Aus Erfahrung lernen – kann man das „beweisen“?

Hypothese: Beim 2. Versuch, eine Minute zu schätzen, wird man besser.

Experiment:

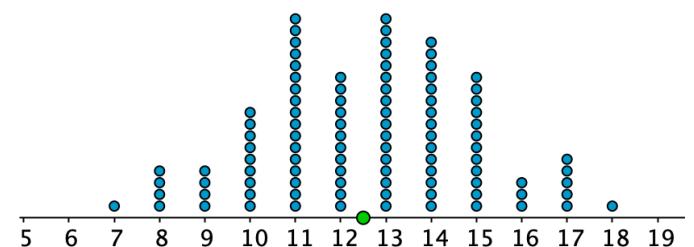
Jede Person legt mit geschlossenen Augen den Kopf auf die Bank. Nach einer Minute wird die Abweichung von einer Minute notiert. Danach noch einmal.

Daten sammeln:

Nr.	I	II	(+/-)
1	7	5	+
2	13	16	-
3	8	1	+
4	12	10	+
...

Simulation:

(n=25; p=0,5; 100-mal)



Gut gemeint – aber Vorsicht

Würfeexperiment einmal anders

Jede/r von euch wählt eine Zahl von 2 bis 12.

Ihr würfelt abwechselnd mit zwei Würfeln und addiert nach jedem Wurf die Augenzahlen.

Ist die Summe der Augenzahlen genau gleich der gewählten Zahl, erhält der/die Spieler/in diese Summe als Punkte gutgeschrieben.

Weicht die Summe der Augenzahlen um 1 von der gewählten Zahl ab, erhält man eine um 1 kleinere Zahl als die Summe der Augenzahlen gutgeschrieben.

Wer zuerst mindestens 30 erreicht, hat gewonnen.

Welche Zahl sollte man wählen?

gewählte Zahl:	2	3	4	5	6	7	8	9	10	11	12
vorausberechnete Punktezahl:	6	16	32	54	82	102	108	98	82	60	32

Tabelle 5: Vorausberechnete Punktezahlen für die zu Spielbeginn gewählte Zahl bei 36 Würfen

Entnommen aus: Hauer-Tippelt, P.: Angemessene Grundvorstellung zu Wahrscheinlichkeit und Zufall entwickeln– Vorschläge für den Stochastikunterricht

Schüler-Aktivität	Beitrag zum Aufbau tragfähiger Grundvorstellungen
Experimentelle Durchführung des Spiels Mehrmalige Entscheidung für eine Zahl	Erfahrungen zum Begriff Zufall Auseinandersetzen mit „Wahrscheinlichkeit“. Welcher Wahrscheinlichkeitsbegriff ist hier passend bzw. nützlich? Erkennen, dass subjektive Wahrscheinlichkeit unpassend ist.
Vorausberechnung	Verbindung von intuitiven Einsichten mit theoretischen Überlegungen. Den Aufbau des erwarteten Werts erfassen. Vorbereitung des Erwartungswertes für diskrete Zufallsvariablen.
Analyse: Theoretisches Ergebnis versus experimentelles Ergebnis	Die Rolle des Zufalls beschreiben können. Die Eignung der vorausberechneten Punktezahlen zur Vorhersage einschätzen können. Für das empirische Gesetz der großen Zahlen eine sichere intuitive Grundlage schaffen.

Tabelle 6: Spiel „Würfeexperiment einmal anders“: Beiträge zu angemessenen Grundvorstellungen

Entscheidung:

Man sollte die Zahl 8 nehmen.



Gut gemeint – aber Vorsicht

Simulation mit Python

```
Anzahl an Spielen: 1000000
Summe: 11
Augensumme Spieler 1: 7; Anteil gewonnen: 0.436794
Augensumme Spieler 2: 8; Anteil gewonnen: 0.420981
Unentschieden: 142225
Durchschnittliche Anzahl der Runden pro Spiel: 3.422736
```

```
Anzahl an Spielen: 1000000
Summe: 30
Augensumme Spieler 1: 7; Anteil gewonnen: 0.398052
Augensumme Spieler 2: 8; Anteil gewonnen: 0.526474
Unentschieden: 75474
Durchschnittliche Anzahl der Runden pro Spiel: 9.036793
```

```
Anzahl an Spielen: 1000000
Summe: 17
Augensumme Spieler 1: 7; Anteil gewonnen: 0.480097
Augensumme Spieler 2: 8; Anteil gewonnen: 0.412404
Unentschieden: 107499
Durchschnittliche Anzahl der Runden pro Spiel: 5.358919
```

Simulation mit GeoGebra

