Direkt schließende Statistik

Vorteile:
1) Mit der direkt schließenden Statistik können wir in die „richtige“ Richtung denken und schließen - also so, wie wir es im Alltag ohnehin tun.
2) Wir können jeder (!) Stichprobe eine Wahrscheinlichkeitsverteilung von Grundgesamtheiten zuweisen und müssen nicht auf viele Versuchsdurchführungen warten.
3) Wir machen keine Fehler wie z. B. beim Hypothesentest, weil wir einfach den verschiedenen Grundgesamtheiten Wahrscheinlichkeiten zuordnen, ohne bestimmte Grundgesamtheiten ein- oder auszuschließen.
4) Da das Gleichsetzen von relativer Häufigkeit und Wahrscheinlichkeit falsch ist, brauchen wir eine Methode für die Mittelstufe, mit der wir von der Stichprobe auf die Grundgesamtheit schließen können. Dass diese Methode dann logischer ist als der übliche Schluss von der Stichprobe auf die Grundgesamtheit, der normalerweise in der Oberstufe gemacht wird, ist ein zusätzlicher Vorteil.
5) Wir lassen die Denkweise, es gebe ein Gesetz des Ausgleichs, gar nicht erst aufkommen - also den Glauben, beim Münzwurf müsse nach mehrmaligem *Kopf*-Wurf nun mehrmalig ein *Zahl*-Wurf erfolgen, damit „es“ wieder ausgeglichen ist.

Die Methode

Der Anfang - Boxen und Kugeln

Beginnen wir mit 5 Boxen, die jeweils 4 Kugeln enthalten, welche entweder blau oder rot sind. Es seien alle möglichen Anteile roter Kugeln vertreten. (Wir könnten auch auf die Anteile blauer Kugeln achten, entscheiden uns aber aus reiner Willkür für die roten Kugeln.)

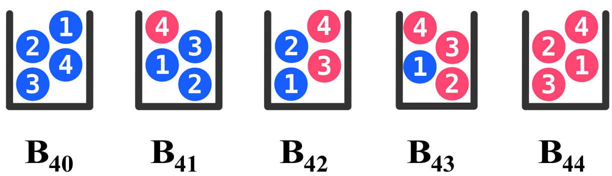


Abb. 1 Die B_4-Boxen

Angenommen, wir ziehen zufällig mit Zurücklegen 3 Kugeln aus einer der Boxen, ohne zu wissen, aus welcher. Dann können wir Vermutungen darüber anstellen, aus welcher Box die Stichprobe gezogen wurde. Wären z. B. eine blaue und 2 rote Kugeln gezogen worden, könnten wir sagen: „Wahrscheinlich ist aus B_3 gezogen worden. Es könnte zwar auch aus B_1 gezogen worden sein, das wäre aber unwahrscheinlich.“ Bestünde die Stichprobe aus 3 blauen Kugeln, könnten wir denken: „B_0 ist am wahrscheinlichsten und B_3 ist am unwahrscheinlichsten.“
Wir Menschen denken so.
Wir weisen Grundgesamtheiten Wahrscheinlichkeiten zu.

Wie das mathematisch funktioniert, soll im Folgenden eine kurze Gegenüberstellung von Wahrscheinlichkeitsrechnung und Statistik zeigen.

Mathematische Anmerkung Die Methode basiert auf dem Konzept der bedingten Wahrscheinlichkeit. Im obigen Beispiel können wir fragen, wie hoch die Wahrscheinlichkeit ist, dass eine bestimmte Stichprobe aus einer bestimmten Box gezogen wird.

Da das Verständnis der bedingten Wahrscheinlichkeit für Schüler schwierig sein kann, vereinfacht diese Methode den Prozess, indem sie sich zunächst auf bestimmte Anzahlen von Elementen in den Grundgesamtheiten beschränkt und damit genauso zugänglich wird wie die Laplace-Wahrscheinlichkeit.

Glücklicherweise können wir dieses Konzept mit Hilfe der Kombinatorik und der Integralrechnung auf Grundgesamtheiten jeder Größe erweitern, wobei die zugrunde liegende Idee immer dieselbe bleibt: Die Wahrscheinlichkeit einer Grundgesamtheit wird bestimmt, indem wir die Anzahl *bestimmter* Stichproben durch die Anzahl *aller* Stichproben teilen.

Nebenbei bemerkt: Wie schwierig das Verständnis bedingter Wahrscheinlichkeiten ist, kann man daran sehen, dass nur sehr wenige Studenten folgende Aufgabe lösen können: Es sind zwei schwarze und zwei weiße Kugeln in einer Box. Es wird zweimal ohne Zurücklegen gezogen. Wie groß ist die Wahrscheinlichkeit, dass beim ersten mal eine schwarze Kugel gezogen wird unter der Bedingung, dass beim zweiten mal eine weiße Kugel gezogen wird? (Es gibt durchaus Menschen, die diese Aufgabe für unlösbar halten.)

Wahrscheinlichkeit, relative Häufigkeit, Grundgesamtheitsanteil

In der Wahrscheinlichkeitsrechnung haben wir eine Grundgesamtheit (hier mit einem bestimmen **Grundgesamtheitsanteil** roter Kugeln) und wir ordnen den möglichen Stichproben (hier mit ihren unterschiedlichen **relativen Häufigkeiten** roter Kugeln) **Wahrscheinlichkeiten** zu.

Dargestellt ist das dreifache zufällige Ziehen aus der Grundgesamtheit $\mathbf{B_{53}}$ mit Zurücklegen und mit Reihenfolge. Die möglichen Stichproben mit gleichen relativen Häufigkeiten roter Kugeln werden zu Ereignissen zusammengefasst.

$\mathcal{P}((3;1))$ bezeichnet die Wahrscheinlichkeit, dass in einer Stichprobe vom Umfang 3 die relative Häufigkeit der roten Kugeln gleich $\frac{1}{3}$ ist.

Wir können aber auch aufgrund einer Stichprobe (mit einer bestimmten **relativen Häufigkeit** roter Kugeln), den Grundgesamtheiten (mit ihren unterschiedlichen **Grundgesamtheitsanteilen** roter Kugeln), aus denen die Stichprobe gezogen werden kann, **Wahrscheinlichkeiten** zuordnen. Das ist dann der Schluss von der Stichprobe auf die Grundgesamtheit.

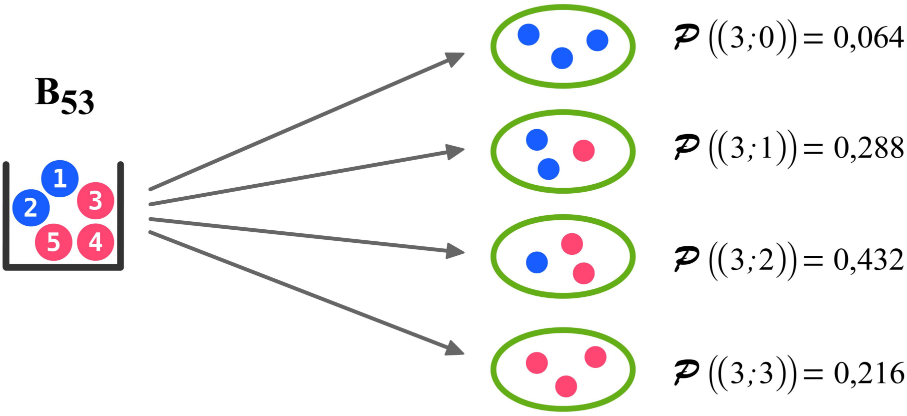


Abb. 2 Von der Grundgesamtheit zu Wahrscheinlichkeiten von Stichproben

$\mathcal{P}((5;2))$ bezeichnet die Wahrscheinlichkeit, dass in einer Grundgesamtheit vom Umfang 5 der Grundgesamtheitsanteil der roten Kugeln gleich $\frac{2}{5}$ ist.

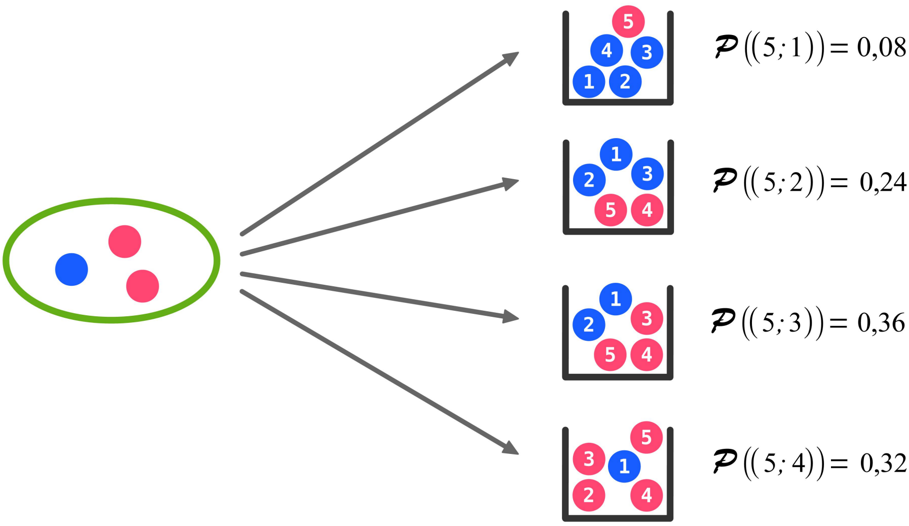


Abb. 3 Von der Stichprobe zu Wahrscheinlichkeiten von Grundgesamtheiten

Wahrscheinlichkeitsrechnung - Kugeln zählen

In der Wahrscheinlichkeitsrechnung haben wir eine **Grundgesamtheit** gegeben und berechnen die **Wahrscheinlichkeit**, eine Stichprobe mit einer bestimmten **relativen Häufigkeit** $\boldsymbol{h}$ zu ziehen, indem wir die Anzahl der Stichproben mit der relativen Häufigkeit $\boldsymbol{h}$ durch die Anzahl aller möglichen Stichproben (aus der gegebenen Grundgesamtheit) teilen.
Stichproben vom Umfang $\boldsymbol{n}$ mit $\boldsymbol{k}$ roten Kugeln nennen wir im Weiteren $(\boldsymbol{n};\boldsymbol{k})$-Stichproben.

Der Zufallsversuch ist das dreimalige Ziehen aus $\mathbf{B_{53}}$. Es wird mit Reihenfolge und mit Zurücklegen gezogen. Gezählt werden die roten Kugeln.
Die Anzahl der Kugeln in den Stichproben ist $\boldsymbol{n} = 3$.
Die Anzahlen der roten Kugeln in den möglichen Stichproben sind $\boldsymbol{k} = 0; 1; 2; 3$
Die Anzahl der Kugeln in der Grundgesamtheit ist $\boldsymbol{g} = 5$.
Die Anzahl der roten Kugeln in der Grundgesamtheit ist $\boldsymbol{r} = 3$.
Die Wahrscheinlichkeit, eine $(\boldsymbol{3};\boldsymbol{2})$-Stichprobe zu erhalten, falls aus der $(\boldsymbol{5};\boldsymbol{3})$-Grundgesamtheit gezogen wird, berechnen wir wie folgt:

$$\mathcal{P}((\mathbf{3};\mathbf{2})) = \frac{54}{8+36+54+27} = \frac{54}{125} = 0,432$$

54 ist dabei die Anzahl der möglichen $(\boldsymbol{3};\boldsymbol{2})$-Stichproben aus $\mathbf{B_{53}}$.
125 ist die Anzahl *aller* Stichproben vom Umfang 3 aus $\mathbf{B_{53}}$.

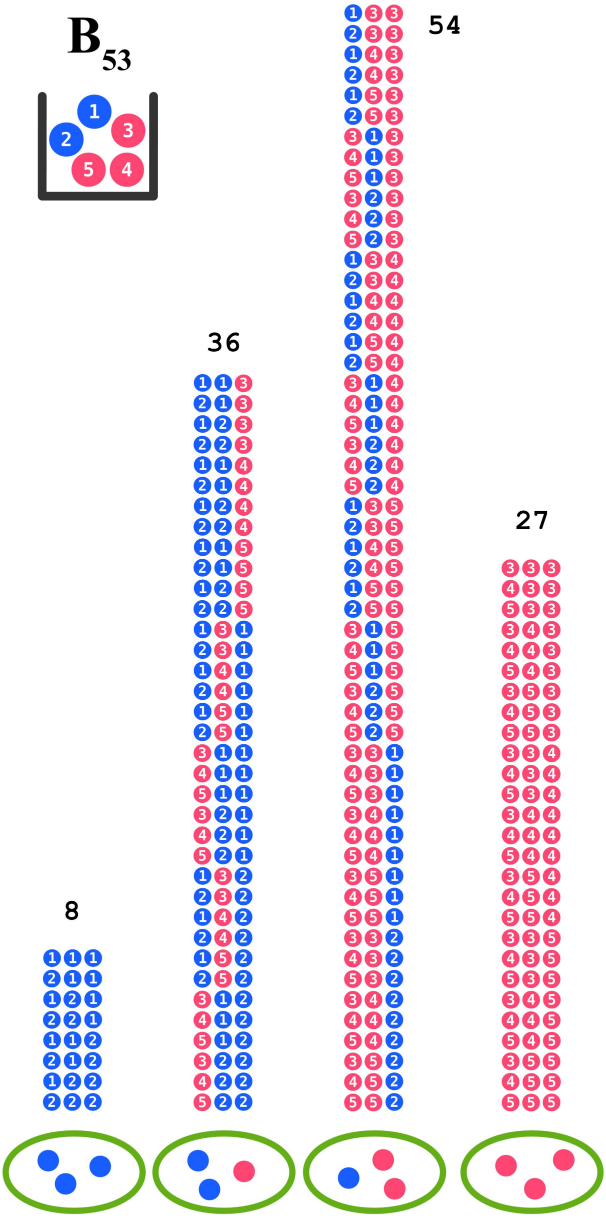


Abb. 4 Von der Grundgesamtheit zu Wahrscheinlichkeiten von Stichproben durch Abzählen

Statistik - Kugeln zählen

In der Statistik haben wir eine Stichprobe mit einer bestimmten **relativen Häufigkeit $\boldsymbol{h}$** gegeben und berechnen die **Wahrscheinlichkeit**, diese Stichprobe aus einer bestimmten **Grundgesamtheit** gezogen zu haben, indem wir die Anzahl der Stichproben (mit der relativen Häufigkeit $\boldsymbol{h}$) aus dieser Grundgesamtheit durch die Anzahl aller Stichproben (mit der relativen Häufigkeit $\boldsymbol{h}$) aus allen möglichen Grundgesamtheiten teilen.

Grundgesamtheiten mit $\boldsymbol{g}$ Kugeln von denen $\boldsymbol{r}$ rot sind, nennen wir im Weiteren $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheiten.

Ein Beispiel:

Der Zufallsversuch ist das dreimalige Ziehen aus einer der Boxen $\mathbf{B_{50}}$ bis $\mathbf{B_{55}}$ (siehe Abb. 5). Es wird mit Reihenfolge und mit Zurücklegen gezogen. Gezählt werden die roten Kugeln.[1]

Die Anzahl der Kugeln in der Stichprobe ist $\boldsymbol{n} = 3$.

Die Anzahl der roten Kugeln in der Stichprobe ist $\boldsymbol{k} = 2$.

Die Anzahl der Kugeln in jeder Grundgesamtheit ist $\boldsymbol{g} = 5$.

Die Anzahlen der roten Kugeln in den Grundgesamtheiten ist $\boldsymbol{r} = 1;\ldots;5$.

Die Wahrscheinlichkeit, die gegebene $(\boldsymbol{3};\boldsymbol{2})$-Stichprobe aus der möglichen $(\boldsymbol{5};\boldsymbol{3})$-Grundgesamtheit gezogen zu haben, berechnen wir wie folgt:

$$\mathcal{P}((\mathbf{5};\mathbf{3})) = \frac{54}{12+36+54+48} = \frac{54}{150} = 0,36$$

54 ist dabei die Anzahl der $(\boldsymbol{3};\boldsymbol{2})$-Stichproben aus $\mathbf{B_{53}}$.

150 ist die Anzahl der $(\boldsymbol{3};\boldsymbol{2})$-Stichproben aus *allen* Grundgesamtheiten von $\mathbf{B_{50}}$ bis $\mathbf{B_{55}}$.

Die Wahrscheinlichkeit, die gegebene $(\boldsymbol{3};\boldsymbol{2})$-Stichprobe aus $\mathbf{B_{53}}$ gezogen zu haben, ist also ca. ein Drittel, weil ca. ein Drittel aller $(\boldsymbol{3};\boldsymbol{2})$-Stichproben aus $\mathbf{B_{53}}$ kommen.

[1] Erfreulicherweise ermöglicht uns die direkt schließende Statistik auch bei sehr geringen Stichprobenumfängen sinnvolle Schätzungen abzugeben.

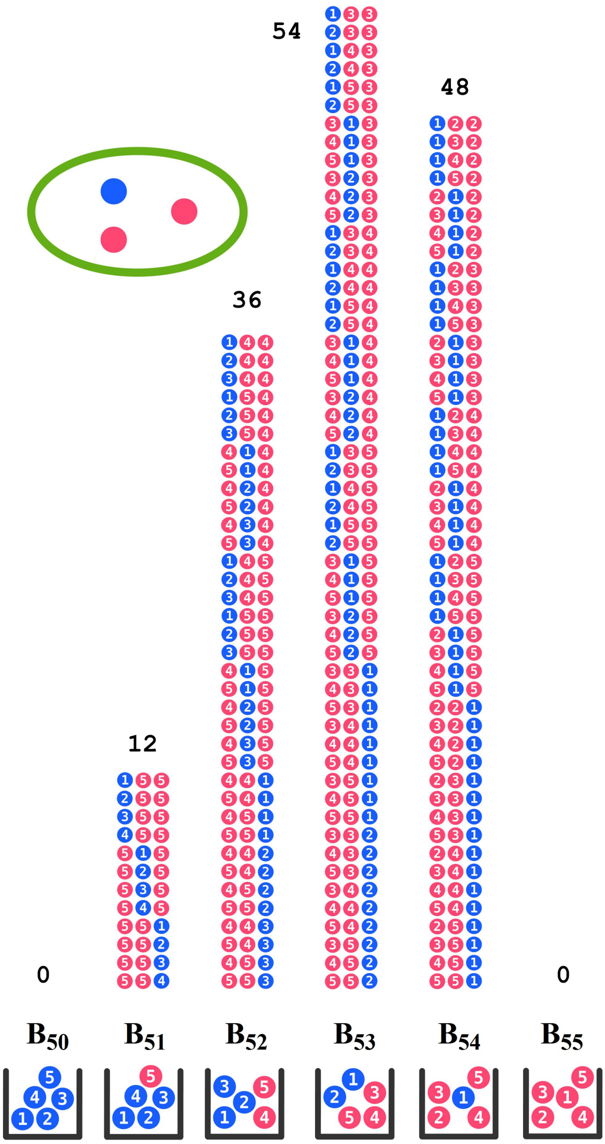
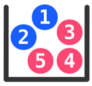


Abb. 5 Von der Stichprobe zu Wahrscheinlichkeiten von Grundgesamtheiten durch Abzählen

Wahrscheinlichkeitsrechnung - mit Kombinatorik

Haben wir für die Wahrscheinlichkeitsrechnung die Kombinatorik zur Verfügung, können wir mit größeren Grundgesamtheiten und größeren Stichproben umgehen. Die Rechnungen bleiben aber prinzipiell gleich. Wir werden also wieder bei gegebener **Grundgesamtheit** die **Wahrscheinlichkeit**, eine Stichprobe mit einer bestimmten **relativen Häufigkeit** zu ziehen, berechnen, indem wir die Anzahl der Stichproben mit dieser relativen Häufigkeit durch die Anzahl aller möglichen Stichproben aus der gegebenen Grundgesamtheit teilen.

Wir gehen jetzt von einer gegebenen Grundgesamtheit mit $\boldsymbol{g}$ Kugeln, von denen $\boldsymbol{r}$ rot sind, aus und suchen die Wahrscheinlichkeit einer möglichen Stichprobe vom Umfang $\boldsymbol{n}$ mit $\boldsymbol{k}$ roten Kugeln. Anders gesagt suchen wir die Wahrscheinlichkeit, aus einer gegebenen $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit eine mögliche $(\boldsymbol{n};\boldsymbol{k})$-Stichprobe ziehen zu können.



B$_{53}$

Schauen wir uns das an einem konkreten Beispiel an: Wir haben eine Grundgesamtheit mit $\boldsymbol{g} = 5$ Kugeln, von denen $\boldsymbol{r} = 3$ rot sind. Wir wollen eine Stichprobe vom Umfang $\boldsymbol{n} = 20$ ziehen und fragen nach der Wahrscheinlichkeit, mit der $\boldsymbol{k} = 11$ Kugeln der Stichprobe rot sind. Dabei ziehen wir mit Zurücklegen und mit Reihenfolge.
Wir rechnen zunächst die Anzahl aller möglichen Stichproben vom Umfang 20 mit 11 roten Kugeln aus.

$$\mathcal{S}((\boldsymbol{20};\boldsymbol{11})) = \binom{20}{11} \cdot 3^{11} \cdot (5-3)^{20-11}$$

Dann rechnen wir die Anzahl aller möglichen Stichproben vom Umfang 20 aus.

$$\sum_{i=0}^{20} \binom{20}{i} \cdot 3^{i} \cdot (5-3)^{20-i}$$

Der Quotient beider Zahlen ist die Wahrscheinlichkeit, aus einer gegebenen Grundgesamtheit mit $\boldsymbol{g} = 5$ Kugeln, von denen $\boldsymbol{r} = 3$ rot sind, eine Stichprobe vom Umfang $\boldsymbol{n} = 20$ mit $\boldsymbol{k} = 11$ roten Kugeln zu ziehen.

$$\mathcal{P}((\boldsymbol{20};\boldsymbol{11})) = \frac{\binom{20}{11} \cdot 3^{11} \cdot (5-3)^{20-11}}{\sum_{i=0}^{20} \binom{20}{i} \cdot 3^{i} \cdot (5-3)^{20-i}}$$

Berücksichtigen wir noch, dass gilt

$$\sum_{i=0}^{20} \binom{20}{i} \cdot 3^{i} \cdot (5-3)^{20-i} = 5^{20}$$

erhalten wir für die Wahrscheinlichkeit

$$\mathcal{P}((\boldsymbol{20};\boldsymbol{11})) = \frac{\binom{20}{11} \cdot 3^{11} \cdot (5-3)^{20-11}}{5^{20}}$$

Also:

$$\mathcal{P}((\boldsymbol{20};\boldsymbol{11})) = \frac{15\,233\,848\,381\,440}{5^{20}} \approx 0,1597$$

Dargestellt sind die Größenverhältnisse der Wahrscheinlichkeiten der möglichen Stichproben für $\boldsymbol{g} = 5$, $\boldsymbol{r} = 3$ und $\boldsymbol{n} = 20$.

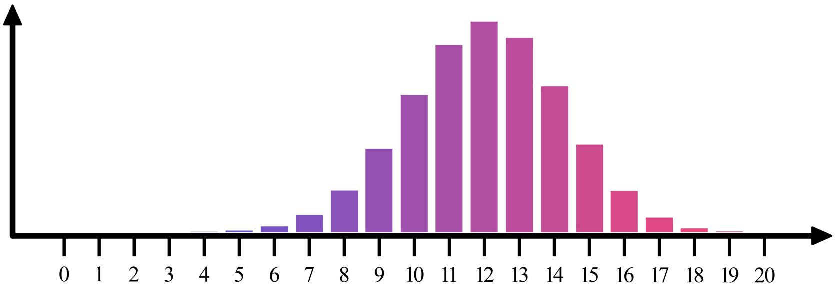


Abb. 6 Verteilung der Stichproben nach relativen Häufigkeiten als Säulendiagramm

Zur besseren Anschauung sind im Folgenden die Anzahlen der Stichproben geordnet nach den Anzahlen der roten Kugeln aufgelistet.[2]

k=0	**1 048 576**	k=7	**1 388 840 878 080**	k=14	**11 864 824 220 160**
k=1	**31 457 280**	k=8	**3 385 299 640 320**	k=15	**7 118 894 532 096**
k=2	**448 266 240**	k=9	**6 770 599 280 640**	k=16	**3 336 981 811 920**
k=3	**4 034 396 160**	k=10	**11 171 488 813 056**	k=17	**1 177 758 286 560**
k=4	**25 719 275 520**	k=11	**15 233 848 381 440**	k=18	**294 439 571 640**
k=5	**123 452 522 496**	k=12	**17 138 079 429 120**	k=19	**46 490 458 680**
k=6	**462 946 959 360**	k=13	**15 819 765 626 880**	k=20	**3 486 784 401**

Abb. 7 Verteilung der Stichproben nach relativen Häufigkeiten in absoluten Zahlen

Allgemein haben wir die Anzahl der möglichen $(\boldsymbol{n};\boldsymbol{k})$-Stichproben aus einer gegebenen $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit

$$\mathcal{S}\left((\boldsymbol{n};\boldsymbol{k})\right) = \binom{n}{k} \cdot r^{k} \cdot (g-r)^{n-k}$$

und die Anzahl aller Stichproben vom Umfang $\boldsymbol{n}$ aus einer $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit

$$\sum_{i=0}^{n} \binom{n}{i} \cdot r^{i} \cdot (g-r)^{n-i}$$

aus denen wir die Wahrscheinlichkeit errechnen können, aus einer gegebenen $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit eine mögliche $(\boldsymbol{n};\boldsymbol{k})$-Stichprobe zu ziehen.

[2] Es ist für Schüler sehr wichtig, sich mit den tatsächlichen Anzahlen der möglichen Stichproben vertraut zu machen. Z. B. ist die Anzahl der Stichproben, in der keine roten Kugeln vorkommen, größer als 1 Million. Das sind also die Stichproben, von denen im Schulunterricht gerne behauptet wird, es gebe davon zu wenige, als dass sie gezogen werden könnten.

$$\mathcal{P}((\boldsymbol{n};\boldsymbol{k})) = \frac{\binom{n}{k} \cdot r^{\,k} \cdot (g-r)^{n-k}}{\sum_{i=0}^{n} \binom{n}{i} \cdot r^{\,i} \cdot (g-r)^{n-i}}$$

Beachten wir noch die Gleichung

$$\sum_{i=0}^{n} \binom{n}{i} \cdot r^{\,i} \cdot (g-r)^{n-i} = g^n$$

können wir schreiben

$$\mathcal{P}((\boldsymbol{n};\boldsymbol{k})) = \frac{\binom{n}{k} \cdot r^{\,k} \cdot (g-r)^{n-k}}{g^n}$$

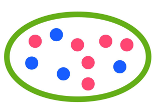
Nach einer kleinen Umformung sieht diese Rechnung dann auch so aus, wie wir es von der Binomialverteilung her gewohnt sind.

$$\mathcal{P}((\boldsymbol{n};\boldsymbol{k})) = \binom{n}{k} \cdot \left(\frac{r}{g}\right)^{k} \cdot \left(1-\frac{r}{g}\right)^{n-k}$$

Statistik - mit Kombinatorik

Haben wir für die direkt schließende Statistik die Kombinatorik zur Verfügung, können wir ebenfalls mit größeren Stichproben und größeren Grundgesamtheiten umgehen. Auch hier bleiben die Rechnungen prinzipiell gleich. Wir werden also wieder bei gegebener Stichprobe mit einer bestimmten **relativen Häufigkeit** die **Wahrscheinlichkeit** einer **Grundgesamtheit** bestimmen, indem wir die Anzahl der Stichproben aus dieser Grundgesamtheit durch die Anzahl aller möglichen Stichproben teilen.

Wir gehen jetzt von einer Stichprobe vom Umfang $\boldsymbol{n}$ mit $\boldsymbol{k}$ roten Kugeln aus und suchen die Wahrscheinlichkeit, dass diese $(\boldsymbol{n};\boldsymbol{k})$-Stichprobe aus einer Grundgesamtheit mit g Kugeln, von denen $\boldsymbol{r}$ rot sind, gezogen wurde. Anders gesagt suchen wir die Wahrscheinlichkeit, mit der wir die gegebene $(\boldsymbol{n};\boldsymbol{k})$-Stichprobe aus der $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit (und nicht aus einer anderen Grundgesamtheit) gezogen haben. Wir weisen also der $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit eine Wahrscheinlichkeit zu.



Schauen wir uns das an einem konkreten Beispiel an: Wir haben eine Stichprobe vom Umfang $\boldsymbol{n} = 10$ mit $\boldsymbol{k} = 6$ roten Kugeln und Grundgesamtheiten mit jeweils $\boldsymbol{g} = 20$ Kugeln gegeben.

Wir fragen nach der Wahrscheinlichkeit, mit der die $(\boldsymbol{10};\boldsymbol{6})$-Stichprobe aus der Grundgesamtheit mit $\boldsymbol{r} = 13$ roten Kugeln gezogen wurde. Wir haben mit Zurücklegen und mit Reihenfolge gezogen.

Wir rechnen zunächst die Anzahl aller möglichen $(\boldsymbol{10};\boldsymbol{6})$-Stichproben aus, die aus einer Grundgesamtheit mit 20 Kugeln, von denen 13 rot sind, gezogen werden können.

$$\mathcal{S}((\boldsymbol{20};\boldsymbol{13})) = \binom{10}{6} \cdot 13^6 \cdot (20-13)^{10-6}$$

Dann rechnen wir die Anzahl aller möglichen $(\boldsymbol{10};\boldsymbol{6})$-Stichproben aus, die aus Grundgesamtheiten mit 20 Kugeln gezogen werden können.

$$\sum_{i=0}^{20} \binom{10}{6} \cdot i^6 \cdot (20-i)^{10-6}$$

Der Quotient beider Zahlen ist die Wahrscheinlichkeit, eine gegebene Stichprobe vom Umfang $\boldsymbol{n} = 10$ mit $\boldsymbol{k} = 6$ roten Kugeln aus einer Grundgesamtheit mit $\boldsymbol{g} = 20$ Kugeln, von denen $\boldsymbol{r} = 13$ rot sind, zu ziehen.

$$\mathcal{P}((\boldsymbol{20};\boldsymbol{13})) = \frac{\binom{10}{6} \cdot 13^6 \cdot (20-13)^{10-6}}{\sum_{i=0}^{20} \binom{10}{6} \cdot i^6 \cdot (20-i)^{10-6}}$$

Konkret haben wir dann:

$$\mathcal{P}((\boldsymbol{20};\boldsymbol{13})) = \frac{2\,433\,725\,365\,890}{18\,618\,197\,650\,500} \approx 0,1307 \tag{1}$$

Wir sind also sehr ähnlich wie in der Wahrscheinlichkeitsrechnung vorgegangen. Dort haben wir die Wahrscheinlichkeit einer möglichen Stichprobe aus einer gegebenen Grundgesamtheit berechnet; hier berechnen wir die Wahrscheinlichkeit einer möglichen Grundgesamtheit zu einer gegebenen Stichprobe.

Stellen wir die Wahrscheinlichkeiten für jede mögliche Grundgesamtheit mit $\boldsymbol{g} =$ 20 Kugeln (bei gegebener $(\boldsymbol{10};\boldsymbol{6})$-Stichprobe) als Säulendiagramm dar, ergibt sich ein ähnliches Bild wie im Fall der Binomialverteilung. Das passt zu dem, was wir intuitiv ohnehin erwarten: Eine Stichprobe mit einer bestimmten relativen Häufigkeit roter Kugeln haben wir wahrscheinlich aus einer Grundgesamtheit mit einem ähnlichen Anteil roter Kugeln gezogen. Wir können diese Stichprobe auch aus einer sehr unähnlichen Grundgesamtheit gezogen haben. Das wäre aber unwahrscheinlich.

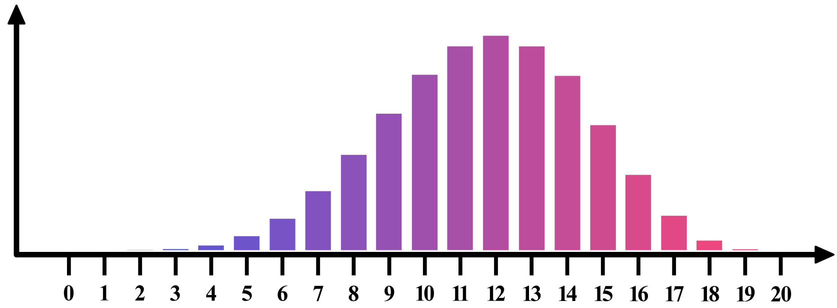


Abb. 8 Verteilung der Grundgesamtheiten nach Grundgesamtheitsanteilen als Säulendiagramm

Zur besseren Anschauung sind im Folgenden die Anzahlen der Stichproben geordnet nach den Anzahlen der roten Kugeln aufgelistet.[3]

r=0	**0**	r=7	**705 636 348 690**	r=14	**2 049 238 517 760**
r=1	**27 367 410**	r=8	**1 141 521 776 640**	r=15	**1 495 019 531 250**
r=2	**1 410 877 440**	r=9	**1 633 973 813 010**	r=16	**901 943 132 160**
r=3	**12 786 229 890**	r=10	**2 100 000 000 000**	r=17	**410 580 048 690**
r=4	**56 371 445 760**	r=11	**2 440 874 461 410**	r=18	**114 281 072 640**
r=5	**166 113 281 250**	r=12	**2 568 423 997 440**	r=19	**9 879 635 010**
r=6	**376 390 748 160**	r=13	**2 433 725 365 890**	r=20	**0**

Abb. 9 Verteilung der Grundgesamtheiten nach Grundgesamtheitsanteilen in absoluten Zahlen

Allgemein haben wir die Anzahl der möglichen $(\boldsymbol{n};\boldsymbol{k})$-Stichproben aus einer gegebenen $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit

$$\mathcal{S}((\boldsymbol{n};\boldsymbol{k})) = \binom{n}{k} \cdot r^{k} \cdot (g-r)^{n-k} \tag{2}$$

und die Anzahl aller $(\boldsymbol{n};\boldsymbol{k})$-Stichproben aus allen Grundgesamtheiten mit $\boldsymbol{g}$ Elementen,

$$\sum_{i=0}^{g} \binom{n}{k} \cdot i^{k} \cdot (g-i)^{n-k} \tag{3}$$

aus denen wir die Wahrscheinlichkeit errechnen können, eine gegebene $(\boldsymbol{n};\boldsymbol{k})$-Stichprobe aus einer möglichen $(\boldsymbol{g};\boldsymbol{r})$-Grundgesamtheit gezogen zu haben.

[3] Auch in diesem Fall sind die tatsächlichen Anzahlen der möglichen Stichproben aus den verschiedenen Grundgesamtheiten interessant. Z. B. ist die Anzahl der Stichproben vom Umfang 10 mit 6 roten Kugeln, die aus einer Grundgesamtheit aus 20 Kugeln von denen nur eine einzige rot ist, entnommen werden können, größer als 27 Millionen. Vielleicht ist es also doch nicht unmöglich, eine $(\boldsymbol{10};\boldsymbol{6})$-Stichprobe aus einer $(\boldsymbol{20};\boldsymbol{1})$-Grundgesamtheit zu ziehen.

$$\mathcal{P}((\boldsymbol{g};\boldsymbol{r})) = \frac{\binom{n}{k} \cdot r^k \cdot (g-r)^{n-k}}{\sum_{i=0}^{g} \binom{n}{k} \cdot i^k \cdot (g-i)^{n-k}} \tag{4}$$

Stetige Verteilungen

Die Normalverteilung in der Wahrscheinlichkeitsrechnung

Werden in der Wahrscheinlichkeitsrechnung die Zahlen bei größeren Stichprobenumfängen zu groß, bestimmen wir Wahrscheinlichkeiten von Stichproben mit der Dichtefunktion der Normalverteilung.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{5}$$

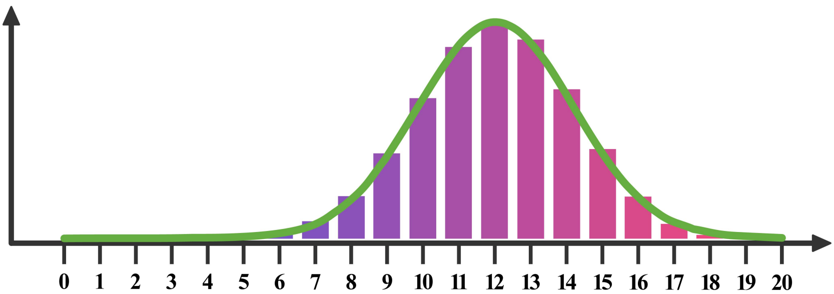


Abb. 10 Von der diskreten Verteilung zur Normalverteilung

Die Likelihood-Funktion in der Statistik

Wie wir in Abb. 9 gesehen haben, werden die Anzahlen der Stichproben schon bei dem relativ geringen Stichprobenumfang von $\boldsymbol{n} = 10$ und kleinen Grundgesamtheiten von nur $\boldsymbol{g} = 20$ Elementen doch sehr groß und damit unhandlich. Haben wir nicht nur 20 Elemente in den möglichen Grundgesamtheiten, sondern - wie in der Praxis üblich -, viele Millionen Elemente, werden die Zahlen für den schulischen Mathematikunterricht zu groß.

Wollen wir uns also in der Statistik bei der Anzahl der Grundgesamtheiten, aus denen wir eine gegebene Stichprobe gezogen haben können, nicht einschränken, können wir mit Hilfe der Likelihood-Funktion die Wahrscheinlichkeiten beliebiger Grundgesamtheiten bestimmen.

In diesem Fall sieht der Funktionsterm der Likelihood-Funktion fast so aus wie die Bernoulli-Formel - nur ist in der Bernoulli-Formel die Anzahl der Erfolge $\boldsymbol{k}$ die Variable, während die Funktionsvariable $\boldsymbol{x}$ der Likelihood-Funktion für die Erfolgswahrscheinlichkeit (bzw. den Grundgesamtheitsanteil) steht.

$$\mathscr{L}_{n;k}(x) = \binom{n}{k} \cdot x^k \cdot (1-x)^{n-k}$$

Das Verfahren der Verwendung der Likelihood-Funktion zur Bestimmung der Wahrscheinlichkeiten möglicher Grundgesamtheiten soll hier nicht formal hergeleitet, sondern nur anschaulich motiviert werden. Allerdings ist Anschauung in diesem Fall so sinnfällig und überzeugend, dass in bestimmten schulischen Kontexten auf die Herleitung verzichtet werden kann.

Anschauliche „Herleitung“ Teilen wir die Säulen des Säulendiagramms in Abb. 8 durch die Summe aller Stichproben aus allen möglichen Grundgesamtheiten - also durch den Nenner von Gleichung (4) - liegen die oberen Seitenmitten der Säulen auf dem Graphen der Likelihood-Funktion.

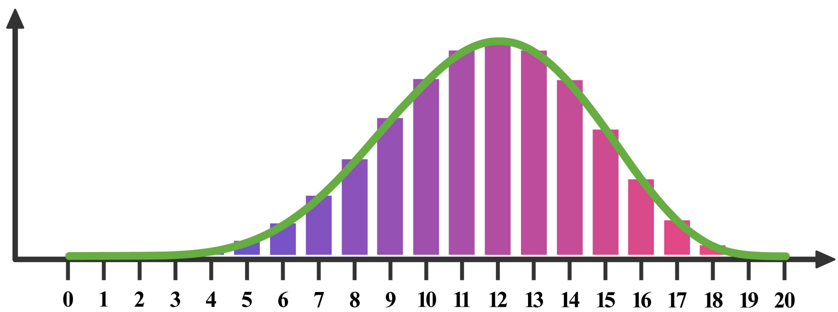


Abb. 11 Von der diskreten Verteilung zur Likelihood-Funktion

Wenn wir die Säulen so verbreitern, dass die Lücken gefüllt sind, können wir auch mit Flächen(-maßzahlen) rechnen. Wie in Abb. 12 dargestellt, kommen wir zum gleichen Ergebnis wie in Gleichung (1), wenn wir die blaue Fläche durch die Summe aus blauer und roter Fläche teilen.

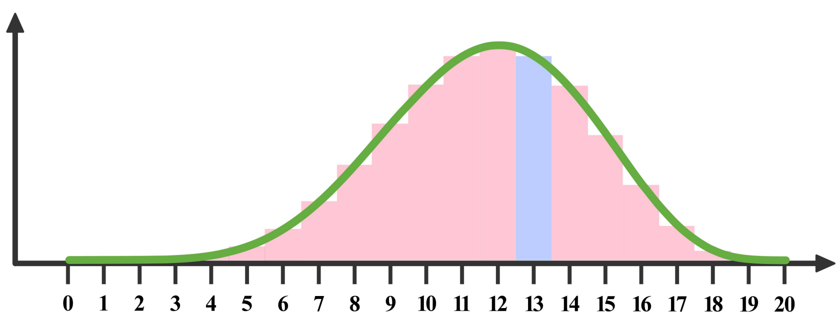


Abb. 12 Die Anzahlen entsprechen Flächen

Und jetzt kommt der entscheidende Punkt: Flächen können wir auch mit Hilfe der Integralrechnung bestimmen. Teilen wir dann die blaue Fläche unter dem Graphen der

Likelihood-Funktion durch die gesamte Fläche unter dem Graphen der Likelihood-Funktion in den Grenzen von 0 bis 1, haben wir eine sehr gute Näherung des Wertes aus Gleichung (1), nämlich $\mathcal{P}((\boldsymbol{20};\boldsymbol{13})) \approx 0,1301$.

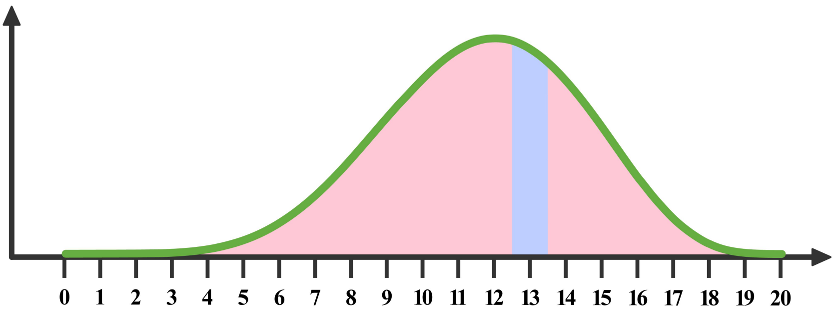


Abb. 13 Die Anzahlen entsprechen Flächen

Die Likelihood-Funktion und die Wahrscheinlichkeit von Intervallen Zu jeder Stichprobe gibt es genau eine Likelihood-Funktion. Die Likelihood-Funktion zur $(\boldsymbol{10};\boldsymbol{6})$-Stichprobe bezeichnen wir mit $\mathscr{L}_{10;6}(x)$.

Die Wahrscheinlichkeit, eine gegebene $(\boldsymbol{10};\boldsymbol{6})$-Stichprobe aus einer Grundgesamtheit mit einem möglichen Grundgesamtheitsanteil zwischen z. B. $0,6$ und $0,8$ zu ziehen, ist dann gleich dem Quotienten aus der markierten Fläche und der gesamten Fläche unter der Kurve zwischen 0 und 1.

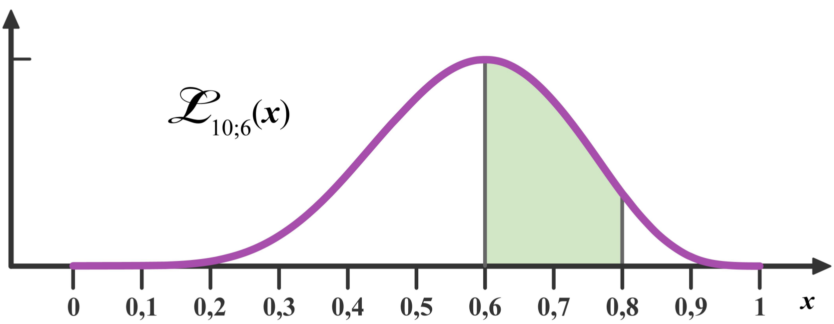


Abb. 14 Wahrscheinlichkeit eines beliebigen Intervalls

Um die Wahrscheinlichkeit konkret zu berechnen, bestimmen wir zunächst die Fläche zwischen Graph und x-Achse in den Grenzen von 0 bis 1.

$$\int_0^1 \mathscr{L}_{10;6}(x)dx = \int_0^1 \binom{10}{6} x^6 (1-x)^{10-6} dx = \frac{1}{11}$$

Das mit 11 multiplizierte bestimmte Integral in den Grenzen von $0,6$ bis $0,8$ ergibt dann die gesuchte Wahrscheinlichkeit.

$$\begin{aligned}\boldsymbol{\mathcal{P}}((\boldsymbol{0{,}6};\boldsymbol{0{,}8})) &= 11 \int_{0,6}^{0,8} \mathscr{L}_{10;6}(x)dx \\ &= 11 \int_{0,6}^{0,8} \binom{10}{6} x^6 (1-x)^{10-6} dx \approx 0,417\end{aligned}$$

Likelihood-Funktion und die Genauigkeit der Schätzung Schätzen wir ausgehend von einer Stichprobe den Grundgesamtheitsanteil mit der Likelihood-Funktion, passiert genau das, was wir Menschen intuitiv von solchen Schätzungen erwarten. Z. B.: Je gröber die Schätzung ist, desto wahrscheinlicher ist sie.

$$\begin{aligned}\boldsymbol{\mathcal{P}}((\boldsymbol{0{,}5};\boldsymbol{0{,}7})) &= 11 \int_{0,5}^{0,7} \mathscr{L}_{10;6}(x)dx \\ &= 11 \int_{0,5}^{0,7} \binom{10}{6} x^6 (1-x)^{10-6} dx \approx 0,515\end{aligned}$$

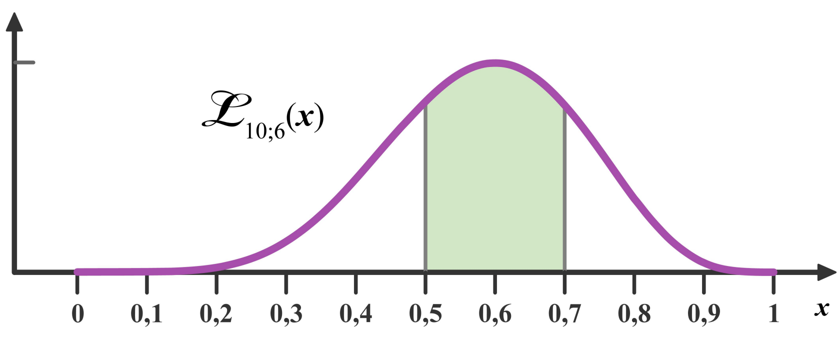


Abb. 15 Wahrscheinlichkeit eines kleinen Intervalls

$$\boldsymbol{\mathcal{P}}((\boldsymbol{0{,}4};\boldsymbol{0{,}8})) = 11 \int_{0{,}4}^{0{,}8} \mathscr{L}_{10;6}(x)dx$$

$$= 11 \int_{0{,}4}^{0{,}8} \binom{10}{6} x^6 (1-x)^{10-6} dx \approx 0,850$$

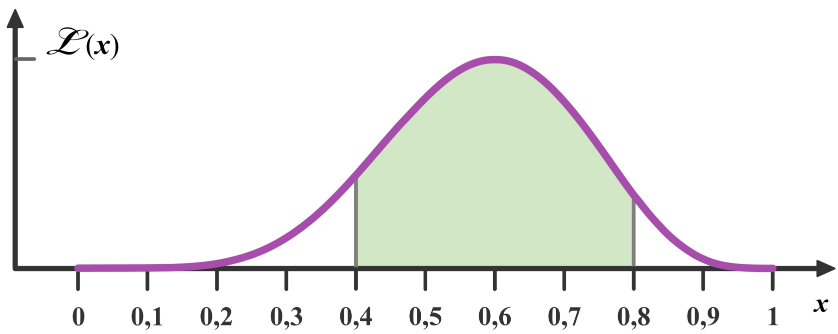


Abb. 16 Wahrscheinlichkeit eines großen Intervalls

Likelihood-Funktion und Stichprobenumfang Je größer der Stichprobenumfang ist, desto genauer ist die Schätzung. Dargestellt ist jeweils für die Stichprobenumfänge 10, 100 und 400 das Intervall, in dem mit ca. 90%iger Wahrscheinlichkeit der Grundgesamtheitsanteil liegt.

$$\boldsymbol{\mathcal{P}}((\boldsymbol{0{,}374};\boldsymbol{0{,}826})) = 11 \int_{0{,}374}^{0{,}826} \mathscr{L}_{10;6}(x)dx$$

$$= 11 \int_{0{,}374}^{0{,}826} \binom{10}{6} x^6 (1-x)^{10-6} dx \approx 0,900003$$

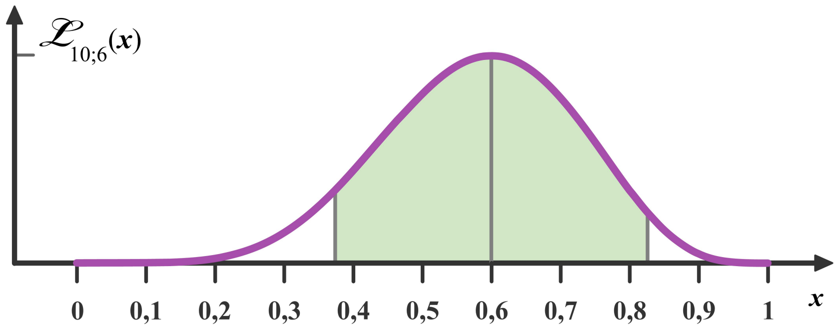


Abb. 17 Intervall, in dem bei gegebener (***10***;***6***)-Stichprobe mit 90%iger Wahrscheinlichkeit die Grundgesamtheit liegt

$$\mathcal{P}((\boldsymbol{0{,}520};\boldsymbol{0{,}680})) = 101 \int_{0{,}520}^{0{,}680} \mathscr{L}_{100;60}(x)dx$$
$$= 101 \int_{0{,}520}^{0{,}680} \binom{10}{6} x^{60} (1-x)^{100-60} dx \approx 0{,}902025$$

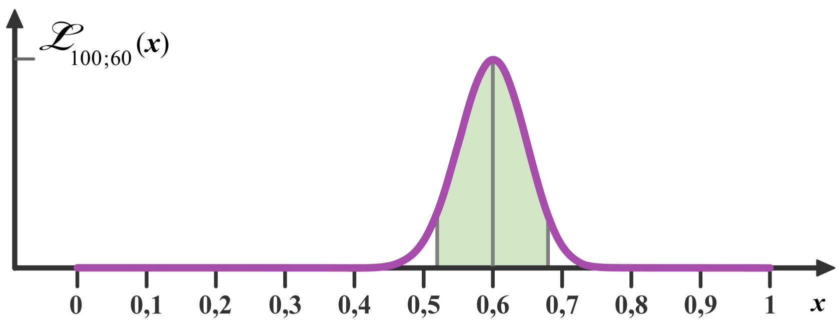


Abb. 18 Intervall, in dem bei gegebener (***100***;***60***)-Stichprobe mit 90%iger Wahrscheinlichkeit die Grundgesamtheit liegt

$$\mathcal{P}((\boldsymbol{0{,}5598};\boldsymbol{0{,}6402})) = 401 \int_{0{,}5598}^{0{,}6402} \mathscr{L}_{400;240}(x)dx$$
$$= 401 \int_{0{,}5598}^{0{,}6402} \binom{400}{240} x^{240} (1-x)^{400-240} dx \approx 0{,}900364$$

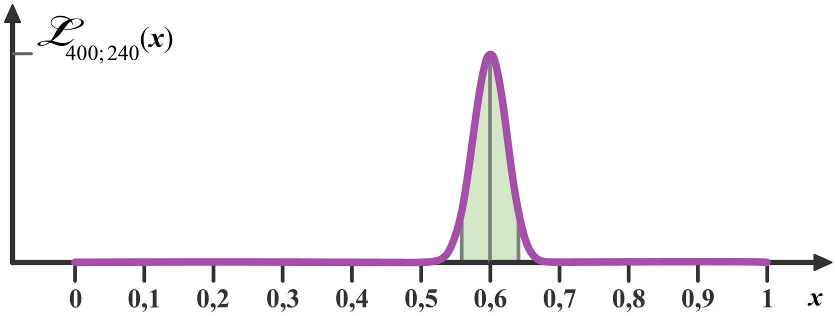


Abb. 19 Intervall, in dem bei gegebener (***400***;***240***)-Stichprobe mit 90%iger Wahrscheinlichkeit die Grundgesamtheit liegt