Übungen mit dem Applet „Zentraler Grenzwertsatz“

1 Statistischer Hintergrund

1.1 Zentraler Grenzwertsatz

Der „Zentrale Grenzwertsatz“ besagt, dass die Verteilung des Mittelwerts von n unabhängigen Zufallszahlen aus einer beliebigen Verteilung mit endlichem Mittelwert μ und endlicher Standardabweichung σ sich mit zunehmendem n immer mehr einer Normalverteilung mit Mittelwert μ und Standardabweichung $\sigma/\sqrt{n}$ annähert.

Der Zentrale Grenzwertsatz ist der Grund für die Bedeutung der Normalverteilung in der Statistik: Häufig ist man an der Verteilung von Mittelwerten interessiert, und auch die Zufallsstreuung eines Einzelwertes ist häufig die Summe vieler unabhängiger Einzelbeiträge und damit näherungsweise normalverteilt.

1.2 Beispiel Würfeln

Dieses Applet visualisiert den Zentralen Grenzwertsatz am Beispiel des Würfelns:

Beim Würfeln mit **einem** Würfel tritt jede der Augenzahlen 1 bis 6 mit derselben Wahrscheinlichkeit 1/6 auf (und nur diese Ergebnisse sind möglich, Zwischenwerte wie z.B. die Augenzahl 1,4 sind unmöglich). Die Verteilung der möglichen Augenzahlen ist damit sicher nicht normalverteilt – bei einer Normalverteilung könnten beliebige Ergebnisse auftreten und die Verteilung hätte eine Glockenform.

Der Mittelwert der Verteilung der Augenzahl ist

$$\mu = \frac{1}{6} \cdot \left(1 + 2 + 3 + 4 + 5 + 6\right) = 3{,}5 .$$

Die Standardabweichung der Verteilung (nicht einer Stichprobe!) der Augenzahl ist

$$\sigma = \sqrt{\frac{1}{6} \cdot \left((1-3{,}5)^2 + (2-3{,}5)^2 + (3-3{,}5)^2 + (4-3{,}5)^2 + (5-3{,}5)^2 + (6-3{,}5)^2\right)} = \sqrt{\frac{35}{12}} .$$

Der Zentrale Grenzwertsatz besagt damit:

Die Verteilung der mittleren Augenzahl von n Würfeln nähert sich mit zunehmendem n immer mehr einer Normalverteilung mit dem Mittelwert $\mu = 3{,}5$ und der Standardabweichung $\sigma = \sqrt{\frac{35}{12n}}$.

Hinweis:
Die mittlere Augenzahl von n Würfeln kann die Werte 1, (1+1/n), (1+2/n), …, 6 annehmen, d.h. sie liegt immer zwischen 1 und 6, die möglichen Werte liegen mit zunehmendem n jedoch immer dichter beieinander. Mit 3 Würfeln sind z.B. die Werte 1, 4/3, 5/3, 2, …, 17/3, 6 möglich.

1.3 Wahrscheinlichkeit und relative Häufigkeit

Die Wahrscheinlichkeit für eine bestimmte mittlere Augenzahl ergibt sich direkt aus der Anzahl möglicher Ergebnisse, die diese mittlere Augenzahl ergeben.

Beispiel: n = 2 Würfel

Insgesamt sind 6*6 = 36 verschiedene Würfelergebnisse möglich. Dabei erhält man
1-mal die mittlere Augenzahl 1 (Kombination 1+1) ⇒ Wahrscheinlichkeit 1/36
2-mal die mittlere Augenzahl 3/2 (Kombinationen 1+2 und 2+1) ⇒ Wahrscheinlichkeit 2/36
3-mal die mittlere Augenzahl 2 (Kombinationen 1+3, 2+2 und 3+1) ⇒ Wahrsch. 3/36
usw.

Diese **Wahrscheinlichkeiten** wurden durch **theoretische Überlegungen** ermittelt, man nennt sie daher **mathematische Wahrscheinlichkeit**. Man verwendet Kombinatorik und die Annahme, dass die Würfel fair sind (d.h. dass jedes Ergebnis zwischen 1 und 6 mit derselben Wahrscheinlichkeit 1/6 auftritt).

Würfelt man nun wirklich, so kann man zählen, wie häufig eine bestimmte mittlere Augenzahl aufgetreten ist. Bei 100 Würfen mit jeweils 2 Würfeln könnte z.B. 4-mal die mittlere Augenzahl 1, 5-mal die mittlere Augenzahl 3/2, 9-mal die mittlere Augenzahl 2, usw. aufgetreten sein. Man sagt dann: Bei diesem Versuch betrug die **relative Häufigkeit** für die

mittlere Augenzahl 1	4/100=4%	
mittlere Augenzahl 3/2	5/100=5%	
mittlere Augenzahl 2	9/100=9%	usw.

Bei jedem Versuch erhält man andere Werte – die Ergebnisse unterliegen dem Zufall und können deutlich von der mathematischen Wahrscheinlichkeit abweichen. Führt man jedoch immer mehr Würfe durch, so verändern sich die relativen Häufigkeiten immer weniger, sie gehen gegen einen Grenzwert. Man sagt: Wenn die Anzahl der Würfe gegen unendlich geht, geht die relative Häufigkeit in die **empirische Wahrscheinlichkeit** über. Wenn die Würfel fair waren, stimmen mathematische und empirische Wahrscheinlichkeit überein.

1.4 Verallgemeinerung: m-seitiger Würfel

Bei einem m-seitigen Würfel tritt jede der Augenzahlen 1 bis m mit derselben Wahrscheinlichkeit 1/m auf. Ein zweiseitiger Würfel entspricht einer Münze mit den Werten 1 und 2 (statt Wappen und Zahl), ein 20-seitiger Würfel wird z.B. bei Magic verwendet. Wie beim normalen, 6-seitigen Würfel nähert sich die Verteilung der mittleren Augenzahl von n Würfeln mit zunehmendem n immer mehr einer Normalverteilung mit Mittelwert $\mu = \frac{m+1}{2}$ und Standardabweichung $\sigma = \sqrt{\frac{m^2-1}{12n}}$.

2 Visualisierungen mit dem Applet

2.1 Zentraler Grenzwertsatz

Beginnen Sie mit einem Würfel – jedes Ergebnis zwischen 1 und 6 ist gleich wahrscheinlich (Rechteckverteilung).

Erhöhen Sie nun mit  die Anzahl Würfel und beobachten Sie, wie die Stufen immer schmäler werden (da immer mehr Zwischenwerte möglich sind) und die Form sich immer mehr einer Glockenform annähert. Da hier die Wahrscheinlichkeit für jede mögliche mittlere Augenzahl einzeln aufgetragen ist, werden diese Einzelwahrscheinlichkeiten mit zunehmender Anzahl Würfel immer niedriger, da immer mehr verschiedene Werte für die Augenzahl möglich werden und die Summe 100% bleibt. Blenden Sie nun mit  die geeignet skalierte Normalverteilung ein – entsprechend dem Zentralen Grenzwertsatz stimmen die beiden Kurven praktisch überein (je größer n, desto kleiner wird der Unterschied).

Beobachten Sie, wie mit zunehmendem n die Standardabweichung (in der Liste rechts) immer kleiner und die Glocke immer schmäler wird. Der Mittelwert bleibt unverändert.

Erniedrigen Sie mit nun die Anzahl der Würfel wieder und beobachten Sie, wie die Abweichungen von der Normalverteilung wieder größer werden, bis bei n = 1 wieder die ursprüngliche Rechteckform hergestellt ist.

2.2 Wahrscheinlichkeit und relative Häufigkeit

Schalten Sie mit die Normalverteilung wieder aus und simulieren Sie mit die voreingestellte Anzahl Würfe (10 000). Die relativen Häufigkeiten stimmen näherungsweise mit den berechneten Wahrscheinlichkeiten überein. Durch wiederholtes Würfeln erhöhen Sie die Anzahl „Aufsummierte Würfe“ und die relativen Häufigkeiten nähern sich immer mehr den Wahrscheinlichkeiten. Die empirische und die mathematische Wahrscheinlichkeit stimmen überein.

Wiederholen Sie die Simulationen für verschiedene Anzahl Würfel n – empirische und mathematische Wahrscheinlichkeiten stimmen überein, die Gesetzmäßigkeiten des Zufalls sind berechenbar.

Reduzieren Sie nun mit die Anzahl Würfe pro Durchlauf z.B. auf 100 und beobachten Sie, dass die Abweichungen zwischen der relativen Häufigkeit und der Wahrscheinlichkeit sehr groß sind und auch mit zusätzlichen Simulationen nur sehr langsam kleiner werden. Die Gesetzmäßigkeiten des Zufalls werden erst bei einer ausreichend großer Anzahl von Versuchen deutlich. Bei einer kleinen Anzahl von Versuchen unterschätzen wir die Wirkung des Zufalls intuitiv häufig – Statistik schützt uns hier vor übereilten Schlüssen.

2.3 m-seitiger Würfel

Als zusätzliche Übung können Sie mit bzw. die Anzahl der Seiten je Würfel erhöhen bzw. erniedrigen. Auch hier sind die Wahrscheinlichkeiten berechenbar und die Rechteckform bei einem Würfel geht mit zunehmendem n in die Normalverteilung über.

Beachten Sie die geänderte Skalierung der x-Achse.

3 Verallgemeinerung

In dieser Simulation wurde gezeigt, wie sich die Verteilung der Mittelwerte von n Werten aus einer Rechteckverteilung mit zunehmendem n der Normalverteilung annähert. Auch die Mittelwerte von Zufallszahlen aus anderen Verteilungen nähern sich der Normalverteilung an, wenn die in 1.1 genannten Voraussetzungen erfüllt sind. Das Würfeln ist nur ein leicht zugängliches Beispiel.