

# Von der Normalverteilung zur Arcussinus-Verteilung

JÖRG MEYER, HAMELN

**Zusammenfassung:** Wenn man sich mehrere Graphen der Binomialverteilung „schräg“ besieht, erkennt man den Graphen der Arcussinus-Verteilung.

## 1 Einleitung

3D-Drucker liefern eine gute Möglichkeit, die Histogramme von Binomialverteilungen für unterschiedliche Werte  $p$  der Einzelerfolgs-Wahrscheinlichkeit zu visualisieren. Man muss nicht mehr die Graphen nacheinander besichtigen, sondern kann sie alle gleichzeitig in Augenschein nehmen.

Abb. 1 zeigt ein Foto einer solchen „Binomialfläche“, die ein Kollege hat herstellen lassen. Dabei ist die Binomialfläche aus den „Deckeln“ der Quader zusammengesetzt.

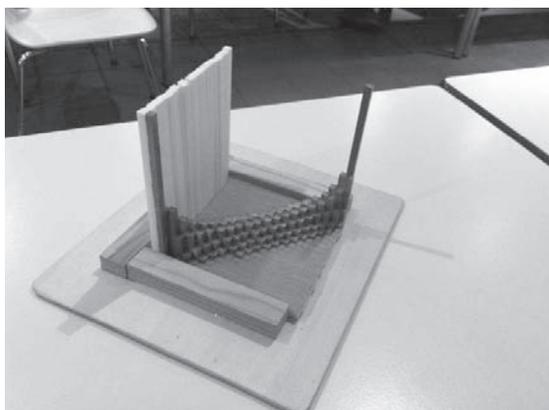


Abb. 1: Die Binomialfläche

Wenn man sich die Diagonale des 3D-Drucks besieht, drängt sich die Frage auf, durch welche Kurve der U-förmige Höhenzug beschrieben werden kann. Hierzu muss man sich Gedanken machen, wie der 3D-Druck hergestellt wird.

## 2 Die Schichten des Modells

Will man  $n + 1$  Histogramme darstellen, sollte  $p$  die Werte  $p = \frac{0}{n}, p = \frac{1}{n}, p = \frac{2}{n}, \dots, p = \frac{n}{n}$  annehmen. Zu jedem Wert von  $p$  gehört die Zuordnung  $k \mapsto B_{n,p}(k) := \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$  mit  $k = 0, \dots, n$ .

Abb. 2 zeigt einige „3D-Druck-Realisierungen“ dieser Zuordnungsvorschrift; Abb. 3 zeigt das gesamte Modell in einem Schrägbild. In beiden Abbildungen verläuft  $k = 0, \dots, n$  von links nach rechts und  $p = \frac{0}{n}, \dots, \frac{n}{n}$  von vorne nach hinten.

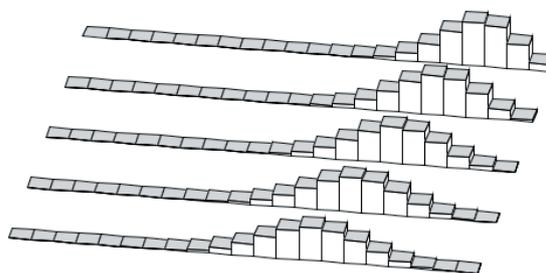


Abb. 2: Einige Schichten der Binomialfläche

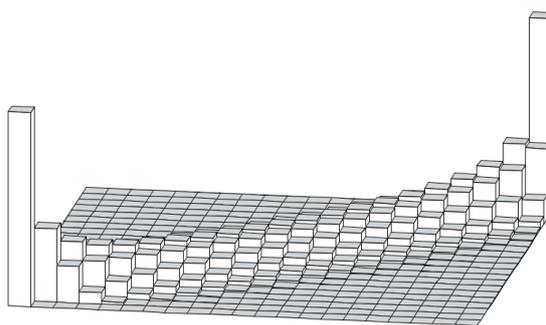


Abb. 3: Modell der Binomialfläche

Für verschiedene Werte von  $n$  sind die Modelle unterschiedlich breit. Um diesen Effekt zu vermeiden, sollte man die  $k$ -Werte durch  $n$  teilen, also die Zuordnung  $\frac{k}{n} \mapsto B_{n,p}(k)$  betrachten (dass die Zuordnung  $x \mapsto \binom{n}{x} \cdot p^{x \cdot n} \cdot (1-p)^{n-x \cdot n}$  re-normiert werden muss, um eine Verteilungsfunktion darzustellen, ist für unsere Zwecke nicht von Belang). Die zugehörigen Histogramme haben ihren Erwartungswert nicht an der Stelle  $n \cdot p$ , sondern (aufgrund der Stauchung der Rechtsachse) stets an der Stelle  $p$ .

## 3 Zur Normalapproximation

Mit  $q := 1 - p, \mu := n \cdot p$  und  $\sigma := \sqrt{n \cdot p \cdot q}$  lässt sich die Zuordnung  $k \mapsto B_{n,p}(k)$  für hinreichend große Werte von  $n$  approximieren durch  $x \mapsto \frac{1}{\sigma} \cdot \varphi\left(\frac{x - \mu}{\sigma}\right)$  mit  $\varphi(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-x^2/2}$  – dies ist die gewöhnliche Normalapproximation der Binomialverteilung. Falls die Ungleichung  $\sigma > 3$  gilt (Faustformel), liegt das 99,7 %-Intervall  $\mu \pm 3 \cdot \sigma$  der Normalverteilung vollständig innerhalb von  $[0; n]$  (Scheid 1992, S. 106). Abb. 4 zeigt beide Verteilungen; auf der Rechtsachse ist  $k \in \{0; 1; \dots; n\}$  bzw.  $x \in \mathbb{R}$  aufgetragen, und auf der Hochachse sind die zugehörigen Werte  $B_{n,p}(k)$  bzw.  $\frac{1}{\sigma} \cdot \varphi\left(\frac{x - \mu}{\sigma}\right)$  aufgetragen.

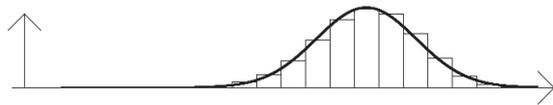


Abb. 4: Normalapproximation der Binomialverteilung

Geht man von der Zuordnung  $k \mapsto B_{n,p}(k)$  über zu  $\frac{k}{n} \mapsto B_{n,p}(k)$ , so wird die Rechtsachse um den Faktor  $n$  gestaucht (die  $x$ -Werte der Normalverteilung müssen also mit  $n$  multipliziert werden); dies ergibt (bis auf Re-Normierung) die Approximation

$$x \mapsto \frac{1}{\sigma} \cdot \varphi\left(\frac{n \cdot x - \mu}{\sigma}\right) = \frac{1}{\sigma} \cdot \varphi\left(\frac{n}{\sigma} \cdot (x - p)\right).$$

Verschiedene Werte von  $p$  führen zu der Approximationsvorschrift

$$(x, p) \mapsto \frac{1}{\sigma} \cdot \varphi\left(\frac{n}{\sigma} \cdot (x - p)\right);$$

dabei ist  $p \in (0; 1)$ , und  $x$  kann ebenfalls aus  $(0; 1)$  angenommen werden, wenn  $\sigma > 3$  gilt (s. o.). Damit lässt sich die Fläche für Werte von  $p$ , die bei hinreichend großem  $n$  nur wenig von den Rändern 0 und 1 entfernt zu sein brauchen, annähernd beschreiben.

#### 4 Zurück zur Diagonale

Die Diagonale der Fläche wird durch  $x = p$  beschrieben. Zur Schnittkurve gehört demnach die Zuordnung

$$x \mapsto \frac{1}{\sqrt{n \cdot x \cdot (1-x)}} \cdot \varphi(0) = \frac{1}{\sqrt{2 \cdot \pi \cdot n}} \cdot \frac{1}{\sqrt{x \cdot (1-x)}}.$$

Abb. 5 zeigt (analog zu Abb. 4) sowohl die Binomial- als auch die Normalzuordnung.

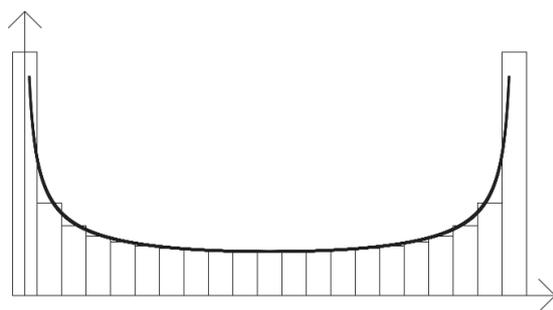


Abb. 5: Die Diagonalkurve

Bis auf Streckung handelt es sich bei der approximierenden Verteilung um die Dichte

$$g(x) = \frac{1}{\pi \cdot \sqrt{x \cdot (1-x)}} \quad (0 < x < 1)$$

der Arcussinus-Verteilung. Sie hat ihren Namen daher, dass die zugehörige kumulative Verteilungsfunktion für  $0 < t < 1$  durch

$$\int_0^t g(x) \cdot dx = \frac{2}{\pi} \cdot \arcsin(\sqrt{t})$$

gegeben ist.

#### 5 Zur Arcussinus-Verteilung

Die im Titel genannte Verteilung tritt auch bei Münzwürfen auf. So hat der Zeitpunkt des letzten Ausgleichs bei  $2 \cdot n$  Münzwürfen eine (diskrete) Arcussinus-Verteilung (Feller 1970, Meyer 2008, Henze 2013): Wirft man mit einer fairen Münze, die die Seiten 1 und  $-1$  hat, und summiert die Wurfresultate, so kann man bei  $2 \cdot n$  Würfeln fragen, wann zum letzten Mal die Summe 0 erreicht wurde. Aufgrund der Symmetrie der Arcussinus-Verteilung gelten dann zum Beispiel die folgenden sehr kontraintuitiven Resultate (nähere Angaben in der angegebenen Literatur):

Die Wahrscheinlichkeit, dass bei 100 Würfeln die Summe 0 zuletzt beim 2. Wurf erreicht wurde, ist genauso groß wie die Wahrscheinlichkeit, dass die Summe 0 zuletzt beim 98. Wurf erreicht wurde.

Mit Wahrscheinlichkeit  $\frac{1}{2}$  gab es bei 100 Würfeln in den letzten 50 Würfeln niemals die Summe 0.

#### Literatur

- Feller, W. (1970): An introduction to probability theory and its applications. 3rd edition. Volume I. New York usw.: John Wiley.
- Henze, N. (2013): Irrfahrten und verwandte Zufälle. Wiesbaden: Springer Spektrum.
- Meyer, J. (2008): Überraschungen beim Münzwurf. In: *Der Mathematikunterricht* Vol. 54 (1); S. 35–48.
- Scheid, H. (1992): Wahrscheinlichkeitsrechnung. Mannheim: BI Wissenschaftsverlag.

#### Anschrift des Verfassers

Jörg Meyer  
Schäfertrift 16  
31789 Hameln  
J.M.Meyer@t-online.de