

Spezielle Eigenschaften der Binomialverteilung

Wir unterscheiden:

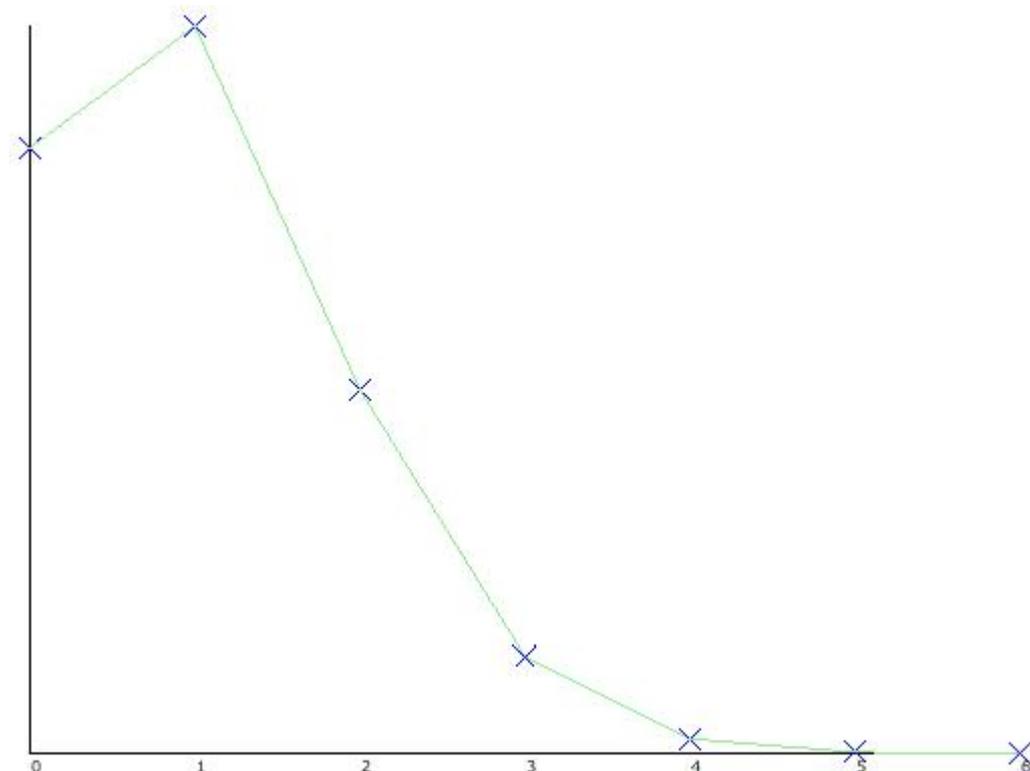
- 1) die Wahrscheinlichkeitsfunktion einer diskreten Variablen
- 2) die Verteilungsfunktion einer diskreten Variablen.

1) Die Wahrscheinlichkeitsfunktion einer diskreten Variablen haben wir im Falle der Binomialverteilung bereits kennengelernt. Sie ist nichts anderes als die Wahrscheinlichkeitsfunktion der Binomialverteilung, wie wir sie bereits kennengelernt haben. In grafischer Darstellung werden auf der x-Achse die verschiedenen k 's und auf der y-Achse die verschiedenen Wahrscheinlichkeiten aufgetragen. In dem bereits erwähnten Beispiel einer Binomialverteilung mit $n=6$ und $p=1/6$ erhalten wir folgende Darstellung:

$$p = \frac{1}{6}$$

$$n = 6$$

Binomialverteilung

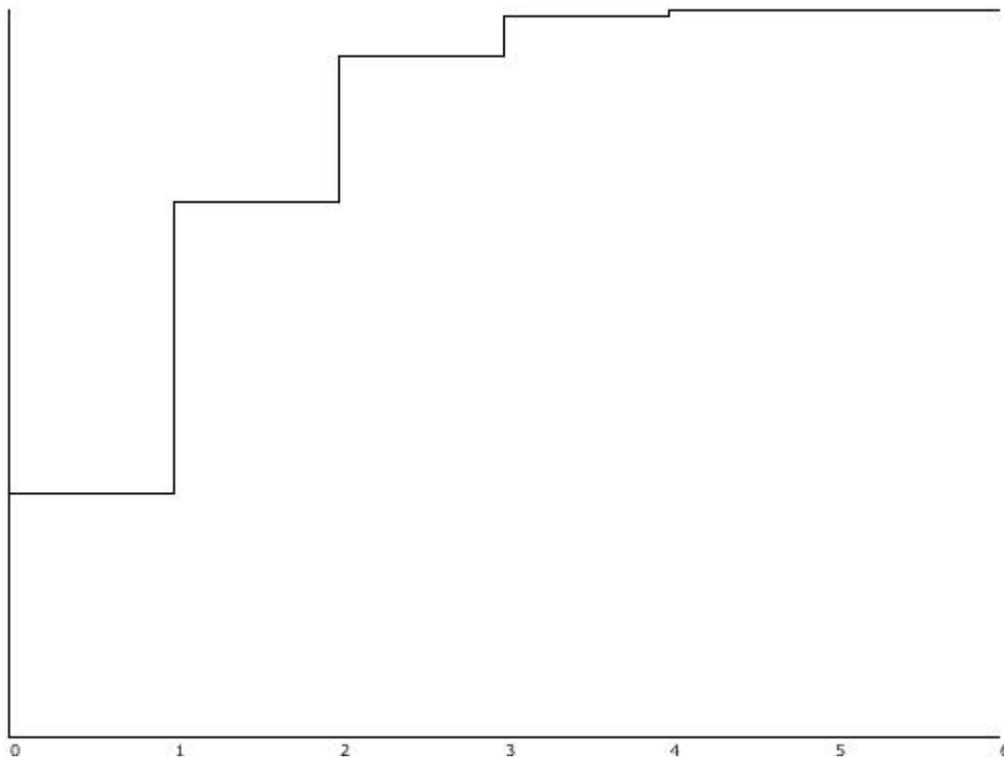


2) Die Verteilungsfunktion einer diskreten Zufallsvariable k erhalten wir einfach dadurch, dass wir von links beginnend die den verschiedenen k 's zugeordneten Wahrscheinlichkeiten aufkumulieren. Spielen wir diesen Fall für $n = 6$ und $p=1/6$ einmal durch:

Solange $k < 0$ ist, so ist die zugehörige Wahrscheinlichkeit gleich 0. Als nächster Wert kommt bei einer diskreten Zufallsvariablen nur der Wert $k = 0$ in Frage. Die zugehörige Wahrscheinlichkeit kennen wir bereits. Sie beträgt 0.334897977. Da es zwischen $k < 0$ und $k=0$ keine Zwischenwerte gibt, springt die Wahrscheinlichkeit von 0 auf 0.334897977. Der

nächste Wert von k ist gleich 1. Die Wahrscheinlichkeit der Binomialverteilung liefert uns hierfür den Wert $P = 0.401877572$. Wollen wir nun wissen, wie wahrscheinlich es ist, bei 6-maligem Würfeln 0mal oder 1-mal eine 6 zu bekommen, so müssen wir die beiden Wahrscheinlichkeiten nur summieren: Die Wahrscheinlichkeit dafür beträgt $0.334897977 + 0.401877572 = 0.736775549$. Wollen wir nun wissen, wie wahrscheinlich es ist, 0-mal oder 1-mal oder zweimal eine 6 zu bekommen, so ergibt sich die Summe: $0.334897977 + 0.401877572 + 0.200938786 = 0.937714335$. Da es sich bei den verschiedenen k 's um diskrete Werte handelt, erhalten wir in der graphische Darstellung eine sogenannte Treppenfunktion, die bis zum Wert von $p = 1$ ansteigt.

Treppenfunktion



Wir können auf diese Weise grafisch veranschaulichen, wie wahrscheinlich es ist, bei 6-maligem Würfeln 0mal oder 1mal oder 2mal oder 3mal oder 4mal eine 6 zu bekommen (die Wahrscheinlichkeit dafür beträgt in der Summe 0,999335562).

Wollen wir umgekehrt nun die Wahrscheinlichkeit wissen 5mal oder mehr eine 6 zu würfeln, so können wir auch die soeben berechnete Summe von 1 subtrahieren: die Wahrscheinlichkeit 5-mal oder mehr bei 6-maligem Würfeln eine 6 zu bekommen ist daher ungefähr $1 - 0,999335562 = 0.000664437$ - was nicht gerade umwerfend ist.

Wollen wir nun hingegen aber die Wahrscheinlichkeit berechnen, von 1000 Würfeln 800 mal oder mehr eine sechs zu bekommen, so wird die Berechnung - falls wir hierzu die Wahrscheinlichkeitsfunktion der Binomialverteilung verwenden - ziemlich aufwändig. Wir können uns aber zur Erleichterung unserer Rechenaufgabe eine bestimmte Eigenschaft der Binomialverteilung zunutze machen. Man kann nämlich zeigen, dass bei großem n (also bei oftmaliger Wiederholung unseres Bernoulli-Versuches) die Binomialverteilung in eine Normalverteilung übergeht. Weiters lässt sich zeigen, dass die Näherung der Binomialverteilung an die Normalverteilung auch bereits bei kleinerem n hinreichend gut ist,

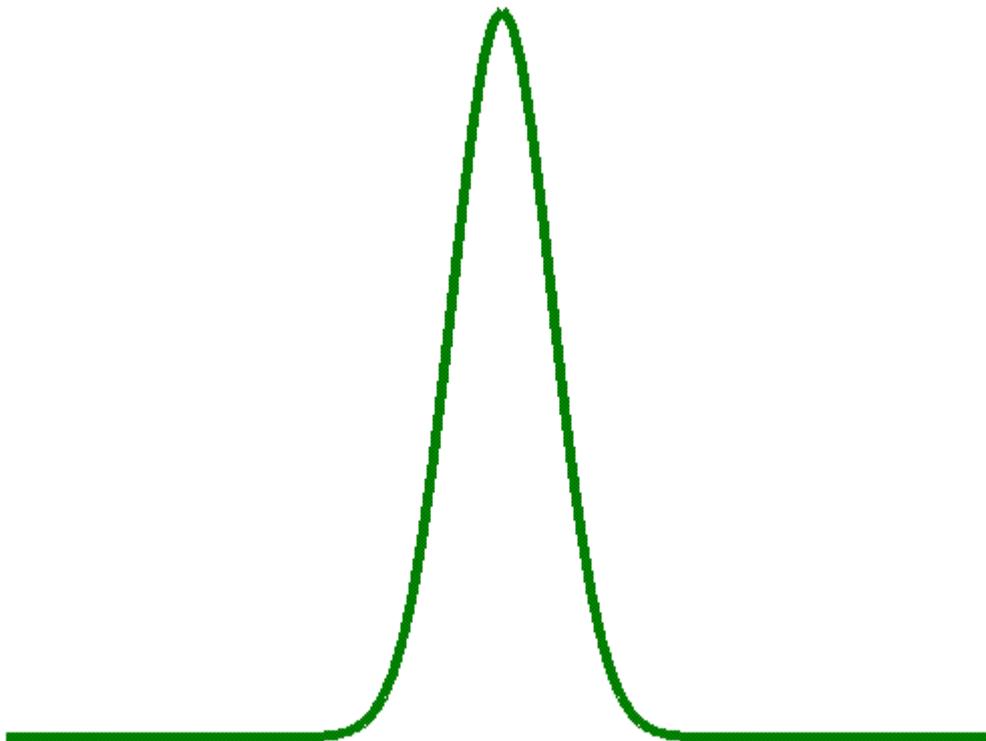
falls p (die Wahrscheinlichkeit für Erfolg) in der Nähe von 0.5 liegt. Ist p ungleich q , so ist die Binomialverteilung nicht symmetrisch und daher eine Approximation an die Normalverteilung bei kleinerem n schwieriger. Nach einer Faustregel von Sachs (1971) lässt sich auch eine Binomialverteilung, in der p ungleich q ist, in eine Normalverteilung überführen, wenn $n \cdot p \cdot q \geq 9$ ist.

Gehen wir also im Folgenden von einer fiktiven Binomialverteilung aus, die sich hinreichend gut durch eine Normalverteilung approximieren lässt. Um zu verstehen, welche Eigenschaften eine solche an die Normalverteilung angenäherte Binomialverteilung hat, benötigen wir dreierlei:

1. Die Wahrscheinlichkeitsfunktion einer kontinuierlichen Variable
2. Die Verteilungsfunktion einer kontinuierlichen Variable
3. Die Wahrscheinlichkeitsfunktion der Binomialverteilung, angenähert an die Normalverteilung

1) Die Normalverteilung ist die Wahrscheinlichkeitsfunktion einer kontinuierlichen Variable

$$\frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$$



Diese Funktion ist in diesem Zusammenhang weniger von mathematischem Interesse. Wichtig dabei ist nur zu beachten, dass wir es mit der Wahrscheinlichkeitsfunktion einer kontinuierlichen Variablen zu tun haben. Da hier nicht einzelnen diskreten Werten - wie im Falle der Binomialverteilung - eine Wahrscheinlichkeit zugewiesen wird, sondern immer nur einem bestimmten Bereich von x -Werten, spricht man hier von einer sogenannten Wahrscheinlichkeitsdichtefunktion. Die Wahrscheinlichkeit eines einzelnen Wertes einer kontinuierlichen Variablen wäre gleich Null, da sich bei einer kontinuierlichen Variablen die Wahrscheinlichkeit nicht als einzelner Punktwert, sondern nur als Fläche angeben lässt. Rückt

diese Fläche gegen Null, so geht auch die entsprechende Wahrscheinlichkeit gegen Null. (Beispiel: Zwischen einer Reaktionszeit von 30,0001 und 30,0002 gibt es theoretisch unendlich viele Zwischenwerte; wäre die Wahrscheinlichkeit an einem bestimmten Punkt einer kontinuierlichen Variablen nicht gleich Null, so würden wir eine unendliche Summe und nicht 1 bekommen - was nicht der Fall sein kann).

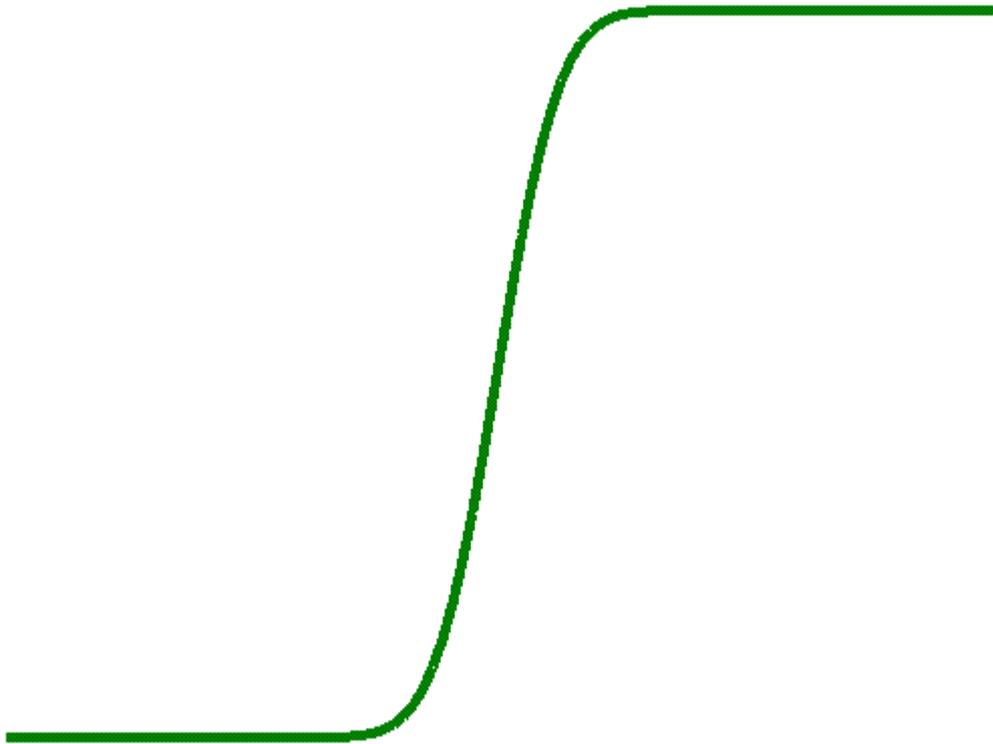
Wenn man alle Werte der Normalverteilung z-transformiert:

$$z = \frac{x_i - \bar{x}}{s}$$

so erhält man eine *Standardnormalverteilung*, mit einem Mittelwert von 0 und einer Standardabweichung von 1. Die Gesamtfläche unter dieser Standardnormalverteilung geht von $z = -\infty$ bis $z = +\infty$ und ist gleich 1.

2) Aus der Standardnormalverteilung erhält man nun die *Verteilungsfunktion* einer kontinuierlichen Variablen, indem man von links beginnend die Wahrscheinlichkeiten für die verschiedenen z-Werte aufkumuliert. Dies erfolgt ganz ähnlich wie bereits bei der Binomialverteilung. Wir beginnen von ganz links (also von $z = -\infty$) und berechnen die Wahrscheinlichkeit bis zu einem gewünschten z-Wert. Da wir es nun aber nicht mit einer diskreten Zufallsvariablen, also auch nicht mit Einzelwahrscheinlichkeiten zu tun haben, tritt an die Stelle einer Summe der Flächenanteil von $z = -\infty$ bis zu dem gewünschten z-Wert. Dieser Flächenanteil wird durch das Integral der Wahrscheinlichkeitsdichtefunktion berechnet.

Statt einer Treppenfunktion erhalten wir auf diese Weise eine S-Kurve, die bei $y = 0$ beginnt, dann bis in die Nähe des Maximums der Normalverteilung ansteigt (dieses Maximum ist der Mittelwert der Normalverteilung), um sich dann rasch an 1 zu nähern. (Es handelt sich hierbei um eine ähnliche Funktion wie die bereits besprochene Summenprozentkurve). Diese Verteilungsfunktion ist auch bekannt unter dem Namen *Queteletsche Kurve*.



Um sich die Bedeutung dieser Kurve von Quetelet an einem anschaulichen Vergleich zu verdeutlichen, gehe ich kurz auf eine Anekdote des Statistikers Van der Waerden ein, der bezüglich dieser Kurve folgendes berichtet hat:

"Lebhaft erinnere ich mich noch, wie mein Vater mich als Knaben an den Rand der Stadt führte, wo am Ufer die Weiden standen und mich 100 Weidenblätter willkürlich pflücken ließ. Nach Aussonderung der beschädigten Spitzen blieben noch 89 unversehrte Blätter übrig, die wir dann zu Hause, nach abnehmender Größe geordnet, wie Soldaten in Reih und Glied stellten. Dann zog mein Vater durch die Spitzen eine gebogene Linie und sagte: 'Dies ist die Kurve von QUETELET. Aus ihr siehst du, wie die Mittelmäßigen immer die große Mehrheit bilden und nur wenige nach oben und unten zurückbleiben.'" (zitiert nach Meschkowski, 115)

3) Die an eine Normalverteilung approximierte Binomialverteilung hat einen Mittelwert von $\mu = n \cdot p$ und eine Varianz von $s^2 = n \cdot p \cdot q$.

Der Mittelwert ist zugleich das Maximum dieser symmetrischen und glockenförmigen Kurve. Was das bedeutet, kann man sich unschwer am Beispiel von 1000maligem Würfeln überlegen. Da in diesem Beispiel das n sehr groß ist, lässt sich diese Binomialverteilung gut durch eine Normalverteilung approximieren. Diese Verteilung erreicht ihr Maximum bei $n \cdot p = 1000 \cdot 1/6 = 166.66666$. Es ist daher diejenige relative Häufigkeit am wahrscheinlichsten, die gleich p ($= 1/6$) ist. Weiters kann man zeigen, dass die Wahrscheinlichkeit, einen Wert in der Nähe dieses Maximums rein per Zufall zu bekommen, sehr groß ist. (Dieser wichtige Umstand der Wahrscheinlichkeitstheorie wird auch als das sogenannte Bernoullische Gesetz der großen Zahlen bezeichnet.) Die an die Normalverteilung approximierte Binomialverteilung lässt sich durch folgende Gleichung ausdrücken:

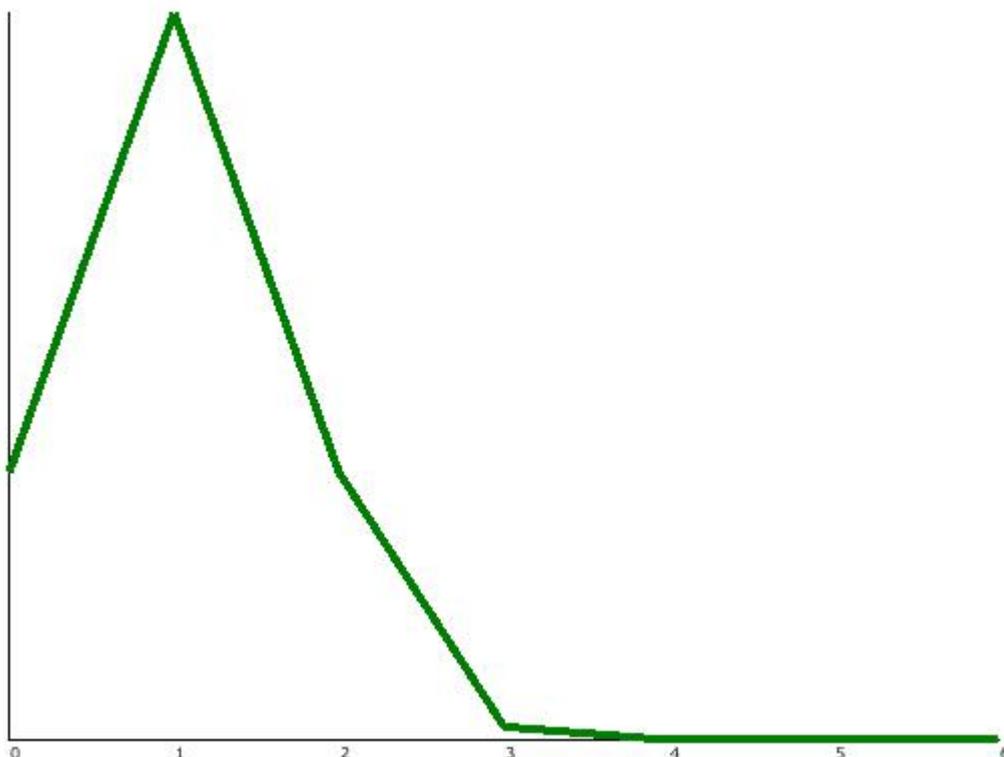
$$\frac{1}{\sqrt{2 \cdot \pi \cdot n \cdot p \cdot q}} \cdot e^{-\frac{(x-n \cdot p)^2}{2 \cdot n \cdot p \cdot q}}$$

Dies entspricht der Formel der Normalverteilung, wenn wir in diese nur den Mittelwert von $n \cdot p$ und die Varianz von $n \cdot p \cdot q$ einsetzen.

Was diese zunächst theoretischen Ausführungen bedeuten, versteht man am besten vor dem Hintergrund unserer Ausgangsfrage: Wie können wir unter Nutzung der Approximation der Binomialverteilung an die Normalverteilung auf bequemere Art und Weise die Wahrscheinlichkeit dafür zu berechnen, von 1000maligem Würfeln 800 mal oder mehr eine 6 zu bekommen.

Dazu müssen wir zunächst von folgender Überlegung ausgehen: Bisher haben wir die einzelnen möglichen k 's unserer Binomialverteilung als Werte einer diskreten Zufallsvariablen behandelt. Angenähert an eine Normalverteilung, müssen wir nun aber davon ausgehen, dass diesen verschiedenen diskreten k 's eine kontinuierliche Zufallsvariable zugrunde liegt. Die einzelnen k 's sind dazu nur die Klassenmitten unserer kontinuierlichen Variablen. Einen ähnlichen Vorgang haben wir bereits in der deskriptiven Statistik kennengelernt. Eine kontinuierliche Variable wurde durch Klassenbildung in eine diskrete Variable umgewandelt und bei der graphischen Darstellung der Klassenhäufigkeiten wurden auf der x-Achse die Klassenmitten aufgetragen.

Approximation unserer Binomialverteilung durch eine Normalverteilung



800- mal eine sechs zu würfeln, um auf unser Ausgangsbeispiel zurückzukommen, lässt sich als *Klassenmitte des Intervalls* 799,5 bis 800,5 interpretieren. 800-mal oder mehr eine sechs zu würfeln beginnt also bei der *Klassenuntergrenze* 799,5. Nehmen wir mal an, die Verteilung lässt sich gut durch eine Normalverteilung repräsentieren, so können wir die

untere Klassengrenze z-transformieren. Dies erfolgt nach der Formel (man beachte: $n \cdot p$ entspricht dem Mittelwert und $n \cdot p \cdot q$ der Varianz der Binomialverteilung):

k 800 x = Klassenuntergrenze = 799,5
n 1000 Mittelwert = $n \cdot p = 166,666666666667$
 Varianz = $n \cdot p \cdot q = 138,888888888889$
p $\frac{1}{6}$ Standardabweichung = $\sqrt{\text{Varianz}} = 11,7851130197758$
 Der z-Wert = (x - Mittelwert) / Standardabweichung = 53,6976889633064
 Die Überschreitungswahrscheinlichkeit p geht gegen Null

Der Fläche von z = -unendlich bis zu 53.7 in einer Standardnormalverteilung entspricht die Gesamtwahrscheinlichkeit in einer Binomialverteilung, weniger als 800-mal eine 6 zu würfeln. (d.h.: 0-mal, 1-mal, 2-mal, 3-mal, ... 700-mal, ..., 799-mal)

Da die Standardnormalverteilung normiert ist - mit einem Mittelwert von 0 und einer Standardabweichung von 1 - können wir diese Wahrscheinlichkeit in einer Tabelle nachschlagen.

Man beachte: Normalverteilung und Standardnormalverteilung haben gleiche Flächenanteile!

In unserem extremen Beispiel ist indes unser z-Wert so groß, dass er in den meisten Tabellen nicht mehr angegeben ist. Die kumulierte Wahrscheinlichkeit, bis zu 800-mal eine 6 zu würfeln (bei 1000 Gesamtversuchen) liegt bei 0.999999999 usw., also nahezu bei 1. Die Wahrscheinlichkeit, 800-mal oder mehr eine 6 zu bekommen, ist dann die gesuchte Überschreitungswahrscheinlichkeit, also $1 - 0.99999999$. Diese Wahrscheinlichkeit liegt nahezu bei Null. Wir können also getrost einen Spieler, der 800-mal oder mehr bei 1000 Würfeln eine 6 bekommt, als Falschspieler bezeichnen bzw. die Unschuldsbehauptung verwerfen.

Was nun in dem konkreten Beispiel vorgeführt wurde, lässt sich allgemein so ausdrücken: Um die Überschreitungswahrscheinlichkeit für ein gegebenes k einer Binomialverteilung zu berechnen, verwendet man bei einem großen n folgende Approximationsformel:

$$z = \frac{k - 0.5 - n \cdot p}{\sqrt{n \cdot p \cdot q}}$$

Zu diesem z-Wert wird in der Standardnormalverteilungstabelle die Wahrscheinlichkeit aufgesucht. Die Überschreitungswahrscheinlichkeit ist nun nichts anderes als 1 minus diese nachgeschlagene Wahrscheinlichkeit.

Dabei gehen wir von einer doppelten Entsprechung aus:

- 1) Der Wahrscheinlichkeit bis zu 800 eine 6 zu würfeln, ist die aufkumulierte Wahrscheinlichkeit der Binomialverteilung für 0-mal, 1-mal, 2-mal, ... bis zu 799-mal eine 6 zu würfeln.
- 2) Diese Wahrscheinlichkeit ist gleich dem Flächenanteil in einer Normalverteilung mit einem Mittelwert von $800 \cdot (1/6)$ und einer Varianz von $800 \cdot (1/6) \cdot (5/6)$ von -unendlich bis zu 799,5.
- 3) Diese Wahrscheinlichkeit wiederum ist gleich dem Flächenanteil in einer Standardnormalverteilung von -unendlich bis zu dem z-transformierten Wert von 799,5 also bis zu 53, was sich aus der z-Tabelle ablesen lässt.

Nur wegen dieser doppelten Entsprechung macht es einen Sinn, die Wahrscheinlichkeit über die z-Tabelle zu bestimmen!