# Excelblatt vereinfacht Stochastik*

STEFAN BARTZ, MECKEL

***Zusammenfassung:*** *Mit einem einzigen Excel-Tabellenblatt können die Binomial-, Normal-, Poisson- und Hypergeometrische Funktionen[1] sehr übersichtlich dargestellt werden. Die Funktionswerte lassen sich so leicht ermitteln, dass viele Aufgaben exakter und einfacher gelöst werden können. Die ausgedehnten Verteilungstafeln für die Binomial- und Standardnormalverteilung sind nicht mehr notwendig. Prognose- und Vertrauensintervalle lassen sich – ohne σ-Regel – exakt bestimmen.*

## Vorteile für den Unterricht

### Flexiblere Aufgabengestaltung

Im Gegensatz zu den tabellierten Wahrscheinlichkeitswerten in den Schulbüchern, die lückenhaft und sehr begrenzt sind, können über das Excel-Tabellenblatt *alle* Werte der Binomial-, Normal-, Poisson- und Hypergeometrische Funktionen[1] abgerufen werden. Damit lassen sich Aufgaben lösen, die im bisherigen Stochastikunterricht nicht gelöst werden konnten. Man erhält mehr Freiheiten in der Aufgabengestaltung.

### Weniger Ablesefehler

Das mühselige und oft fehlerbehaftete Ablesen in den umfangreichen Verteilungstafeln der Binomial- und Standardnormalverteilung erübrigt sich.

### Größere Anschauung

Die Graphen der Verteilungen werden visualisiert. Ihre Eigenschaften treten so deutlicher hervor und können besser vermittelt werden. Man kann beispielsweise durch Scrollen an den Parameterwerten direkt erkennen, wie sich diese auf den Verlauf des Graphen auswirkt: So wandert der Graph bei der Binomialverteilung (Abb. 2 blau) beim Hochscrollen von n weiter nach rechts und die Glocke wird gleichzeitig flacher und breiter. Dagegen verändert μ bei der Normalverteilung (Abb. 2 orange) nicht die Glockenform, sondern nur ihre horizontale Lage.

### Unterschied zwischen Parametern und Variablen erfahrbar

Im Gegensatz zu den Parametern verändert sich der Graph durch eine Erhöhung der Variablen x nicht. Es wird lediglich der Flächeninhalt einer weiteren Säule in das Ergebnis mit einbezogen (Abb. 1). Über das Kontrollkästchen "Fläche anzeigen" wird sichtbar, bis zu welcher Stelle der Flächeninhalt bestimmt wird. Somit kann auch der Zusammenhang zwischen den Werten der oben bestimmten kumulierten $Bin_{n,p}(X \leq x)$ und der im Diagramm angezeigten nicht-kumulierten Binomialverteilung $Bin_{n,p}(X = x)$ veranschaulicht werden.

### Graphenvergleich und Graphenanpassung möglich

Die Graphen der verschiedenen Verteilungen können gleichzeitig eingeblendet und miteinander verglichen werden. So kann durch Scrollen an den μ- und σ-Werten gezeigt werden, unter welchen Bedingungen eine Normalverteilung eine vorgegebene Binomialverteilung approximieren kann (s. Abb. 2).

### Synchronisation möglich

Durch Klicken eines Kontrollkästchens können Normal- und Poisson-Verteilung mit der Binomialverteilung synchronisiert werden, d.h. man kann sehen, wie sich die Parameterwerte und die Graphen automatisch an diese anpassen. (Das ist z.B. bei älteren Excelversionen interessant. Excel2002 und älter Versionen können ab n = 1030 nicht mehr alle

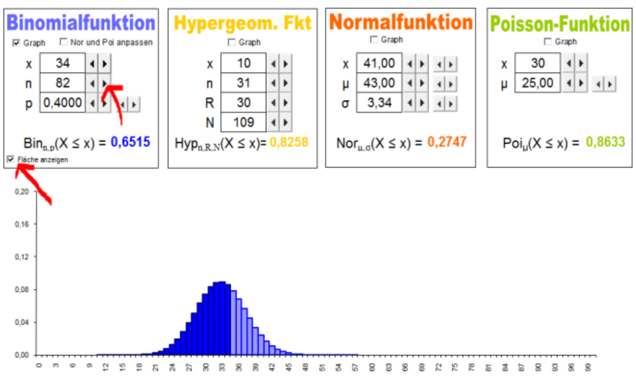



**Abb. 1**

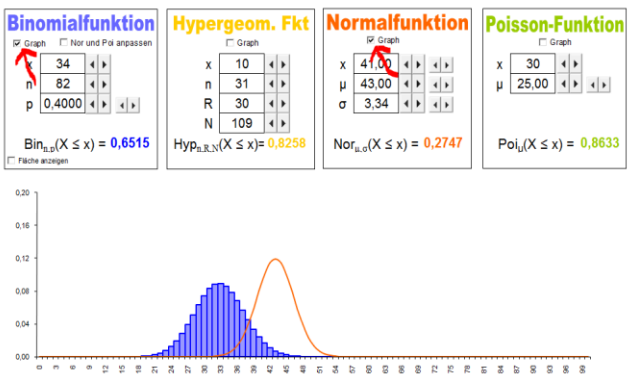



**Abb. 2**

* aus: Stochastik in der Schule 27(2007)2, S. 25–29

kumulierten Binomialfunktionswerte bestimmen und geben stattdessen "#####" aus.) Die gesuchten Wahrscheinlichkeiten können dann bei der angepassten kumulierten Normalverteilung abgelesen werden. Im unteren Beispiel beträgt sie 0,5363. Der σ-Wert von 203,10 (> 3) zeigt, dass es sich dabei um eine gute Näherung handelt.

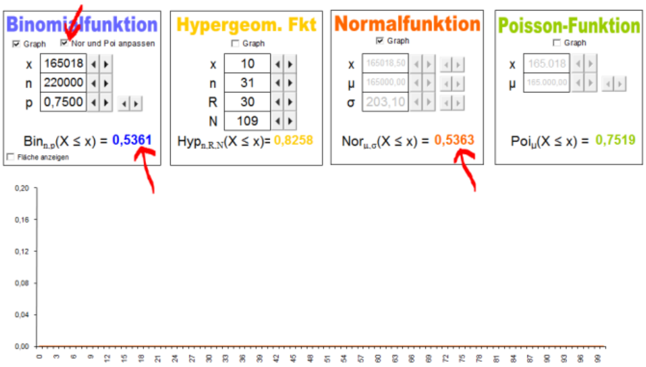



**Abb. 3**

## Transformation nicht mehr notwendig

Eine Transformation von der jeweiligen Normal- zur Standardnormalverteilung, die vielen Schülern schwer fällt und oft rezeptartig und ohne tieferes Verständnis vorgenommen wird, ist nicht mehr notwendig. Die entsprechenden Werte lassen sich nun direkt bestimmen.

**Beispiel 1: n gesucht**
Wie oft muss man einen gerechten Würfel mind. werfen, um mit einer Wk von mind. 95% mind. 2-mal eine "6" zu erhalten?

(1) E: $X \geq 2$
(2) binomialverteilt mit n = ? und $p = \frac{1}{6}$
(3) $P(E) \geq 0{,}95$
$\Rightarrow$ $Bin_{n;\,0,167}(X \geq 2) \geq 0{,}95$
$\Leftrightarrow$ $1 - Bin_{n;\,0,167}(X \leq 1) \geq 0{,}95$
$\Leftrightarrow$ $Bin_{n;\,0,167}(X \leq 1) \leq 0{,}05$
$\Rightarrow$ $n \in [27;\infty)$

**Beispiel 2: p gesucht (Schätzen)**
Man hat keinen Anhaltspunkt, wie groß die Trefferwahrscheinlichkeit p für eine "6" bei einem verbeulten Würfel ist. Deshalb wirft man ihn 50-mal und erhält 32-mal eine "6". Schätze p mit einem 95% Vertrauensintervall.

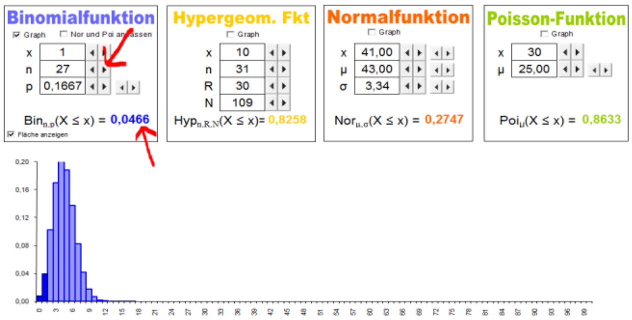
(1) E: X =32
(2) binomialverteilt mit n = 50 und p = ?
(3) Annahme: E liegt im 95%igen Hauptstreubereich
$\Rightarrow$ $0{,}025 \leq Bin_{50;\,p}(X \leq 32) \leq 0{,}975$
$\Rightarrow$ $p \in [0{,}5123;\, 0{,}7708]$
Wenn man davon ausgeht, dass sich das erzielte Stichprobenergebnis innerhalb des HSB (≥95%) befindet, muss p zwischen 51,23% und 77,08% liegen.

## Näherungsverfahren erübrigen sich

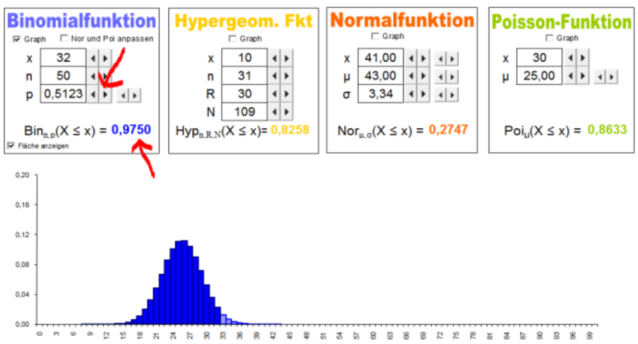
Bisher wurde die Poisson-Verteilung verwendet, um die Binomialfunktionswerte bei sehr kleinem p und großem n annähernd berechnen zu können. Excel kann die *relevanten* Werte jedoch direkt ermitteln (auch für n ≥ 1030). Ebenso wurden alle Prognose- und Vertrauensintervalle von nicht-normalverteilten Zufallsvariablen näherungsweise mit der σ-Regel bestimmt. Auch diese können nun exakt abgerufen werden.

## Umkehraufgaben leichter lösbar

Das Excelblatt lässt sich nicht nur in den Kapiteln Binomial-, Hypergeometrische- und Normalverteilung einsetzen, sondern leistet gerade bei weiterführenden Themen wie Schätzen und Testen von Hypothesen sehr gute Dienste. Bei diesen Umkehraufgaben sind Wahrscheinlichkeiten vorgegeben und x, n oder p gesucht. Diese Variablen- bzw. Parameterwerte lassen sich nun experimentell ermitteln, indem man z.B. den p-Wert über den jeweiligen Scrollbutton so lange verändert, bis die geforderte Wahrscheinlichkeit erreicht ist. Drei Beispielaufgaben sollen das verdeutlichen:




Im Excel-Tabellenblatt stellt man bei der Binomialfunktion p=1/6 und x=1 ein. Beim Verändern des n-Wertes über den Scrollbutton erkennt man dann, dass n ≥ 27 werden muss, wenn man eine Wahrscheinlichkeit von kleiner gleich 0,05 erreichen will.



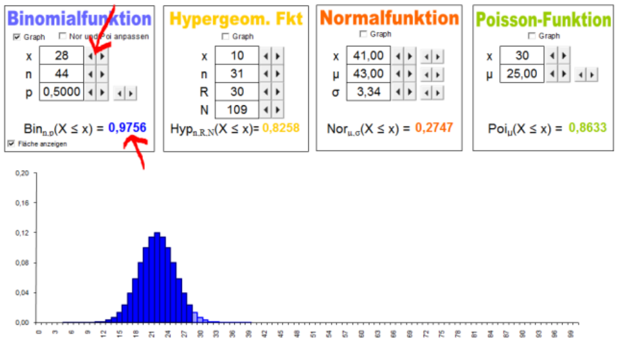

Hier erhält man die Lösung mit Hilfe der Scrollbuttons von p.

**Beispiel 3: x gesucht (Testen)**
Annika behauptet, dass sie nur am Geschmack erkennt, ob der Tee mit entkalktem oder nicht entkalktem Wasser hergestellt wurde. Bei 44 Versuchen stimmt ihre Angabe in 29 Fällen. Ist damit bewiesen, dass sie einen sehr sensiblen Geschmacksinn hat (Signifikanzniveau 5%)?

(1) E: X= 29; $p_0 = 0{,}5$ $p_{spek} \neq 0{,}5$ (→ beidseitig)
(2) binomialverteilt mit n = 44 und $p_0 = 0{,}5$
(3) Liegt E noch im HSB von $Bin_{44;p_0}$?

$$0{,}025 \le Bin_{44;\,0{,}5}(X \le x) \le 0{,}975$$
$$\Rightarrow \quad x \in [16; 28]$$

Da das Stichprobenergebnis E außerhalb des HSB (95%) liegt, kann $p_0$ abgelehnt und somit $p_{spek}$ bestätigt werden; d.h. man kann davon ausgehen, dass Annika nicht nur rät.




$p_{spek}$ beschreibt die spektakuläre neue Hypothese, die man verifizieren will; $p_0$ die ihr entgegenstehende Nullhypothese. Im Excelblatt wird n = 44 und p = 0,5 eingestellt und x durch Scrollen ermittelt.

## Vergleich alter und neuer Lösungswege

Demgegenüber waren die bisherigen Lösungsverfahren dieser Umkehraufgaben oft so aufwendig, dass der eigentliche Lösungsgedanke leicht aus den Augen verloren wurde. Die angewandten Näherungen waren manchmal so grob, dass sie zu falschen Entscheidungen führten. Folgende Gegenüberstellung veranschaulicht das anhand der Beispielaufgaben 2 und 3:

| **Bisheriger Lösungsweg** (mit σ-Regel) | **Neuer Lösungsweg** (mit Excelblatt) |
|---|---|
| ges.: p, so dass $0{,}025 \le Bin_{50;\,p}(X \le 32) \le 0{,}975$<br>nun: X ist in guter Näherung normalverteilt \|σ > 3; σ-Regel<br>$\Rightarrow P(\mu - 1{,}96\sigma \le X \le \mu + 1{,}96\sigma) \approx 0{,}95$ \| Quantilentbl.<br>$\Rightarrow np - 1{,}96\sqrt{np(1-p)} \le X \le np + 1{,}96\sqrt{np(1-p)}$ \|:n \| $h := \frac{x}{n}$<br>$\Leftrightarrow p - 1{,}96\sqrt{\frac{p(1-p)}{n}} \le h \le p + 1{,}96\sqrt{\frac{p(1-p)}{n}}$ \|-p<br>$\Leftrightarrow -1{,}96\sqrt{\frac{p(1-p)}{n}} \le h - p \le 1{,}96\sqrt{\frac{p(1-p)}{n}}$ \|hoch 2<br>$\Rightarrow (h-p)^2 \le 1{,}96^2 \cdot \frac{p(1-p)}{n}$<br>$\Leftrightarrow (n+1{,}96^2)p^2 - (2hn + 1{,}96^2)p + nh^2 \le 0$ \|quad. Ergänz.<br>$\Leftrightarrow p \lessgtr \frac{2hn + 1{,}96^2 \pm 1{,}96\sqrt{1{,}96^2 + 4nh(1-h)}}{2(n+1{,}96^2)}$<br>$\Leftrightarrow p \lessgtr \frac{h + \frac{1{,}96^2}{2n} \pm 1{,}96\sqrt{\frac{1{,}96^2}{4n^2} + \frac{h(1-h)}{n}}}{1 + \frac{1{,}96^2}{n}}$ \|abschätzen[1]<br>$\approx> h - 1{,}96\sqrt{\frac{h(1-h)}{n}} \le p \le h + 1{,}96\sqrt{\frac{h(1-h)}{n}}$<br>$\Rightarrow \underline{0{,}5070} \le p \le 0{,}7730$ | $0{,}025 \le Bin_{50;\,p}(X \le 32) \le 0{,}975$<br>$\Rightarrow p \in [\underline{0{,}5123}; 0{,}7708]$ |
| ges. x, so dass $0{,}025 \le Bin_{44;\,0{,}5}(X \le x) \le 0{,}975$<br>nun: X ist in guter Näherung normalverteilt \|σ > 3<br>$\Rightarrow Nor_{\mu,\sigma}(\mu - 1{,}96\sigma \le X \le \mu + 1{,}96\sigma) \ge 0{,}95$ \|σ-Regel<br>$\Rightarrow Nor_{22;\,3{,}32}(15{,}499 \le X \le 28{,}501) \ge 0{,}95$ \|nach außen<br>$\approx> Nor_{22;\,3{,}32}(14{,}5 \le X \le 29{,}5) \ge 0{,}95$ \|Stetigk.Kor.<br>$\approx> Bin_{44;\,0{,}5}(15 \le X \le 29) \ge 0{,}95$<br>$\Rightarrow x \in \underline{[15; 29]}$ | $0{,}025 \le Bin_{44;\,0{,}5}(X \le x) \le 0{,}975$<br>$\Rightarrow x \in \underline{[16; 28]}$ |

[1] Bei großem n lassen sich die Terme $\frac{1{,}96^2}{n}$, $\frac{1{,}96^2}{2n}$ bzw. $\frac{1{,}96^2}{4n^2}$ gegenüber von 1, h bzw. $\frac{h(1-h)}{n}$ vernachlässigen.

## Schlussbemerkung

- Dieses neu entwickelte Excel-Tabellenblatt befreit den Stochastikunterricht von aufwendigen und ungenauen Verfahren. Sie werden ersetzt durch ein vielfältig anwendbares experimentell-numerisches Vorgehen, das für Schüler leichter zu verstehen und anzuwenden ist. Neue Möglichkeiten eröffnen sich bezüglich der Aufgabengestaltung und der Vermittlung der Inhalte. Es wird Freiraum geschaffen, um weiterführende Themen in Angriff nehmen zu können (z.B. Markow-Ketten, Chi-Quadrat-Test).
- Auch graphikfähige TR, CAS oder Simulationsprogramme könnten den klassischen Stochastiklehrgang mit Tabellen, Transformationen und σ-Regel als Näherung durch ein experimentell-numerisches Vorgehen ersetzen. Dazu bedarf es aber eines ganz anderen Ansatzes.
- Der Autor verzichte seit Jahren auf tabellierte Wahrscheinlichkeitswerte im Schulunterricht und arbeite nur noch mit dem vorgestellten Excelblatt. Fast alle Schüler können mittlerweile zu Hause auf Excel zugreifen. Notfalls können sie für die Hausaufgaben die Rechner der Schulbibliothek nutzen. In Kurs- und Abiturarbeiten werden 4 Notebooks mit der Exceldatei bereitgestellt, über die die jeweils benötigten Werte abgerufen werden können. Dabei lösen die Schüler die Aufgaben so weit wie möglich im Heft ohne Notebook. Die gewünschten Werte können dann gegen Ende – für alle Aufgaben zusammen – ermittelt werden.
- Das Tabellenblatt ist ausbaubar, so dass es auch im Hochschulbereich verwendet werden kann. Einige stochastische Funktionen wie die $\chi^2$-, t- und Fisher-Verteilung sind aus diesem Grunde bereits eingearbeitet worden.
- Unter www.stefanbartz.de/materialien.htm ist die vorgestellte Exceldatei "stochastik.xls" als Freedownload abrufbar.

## Anmerkungen

[1] Aus didaktischen Gründen werden die Begriffe "Binomialfunktion" für $Bin_{n,p}(X=x)$ und "kumulierte Binomialfunktion" für $Bin_{n,p}(X \leq x)$, statt der für Schüler verwirrenden Begriffe "Binomialverteilung" und "Verteilungsfunktion der Binomialverteilung" verwendet.
Die neu eingeführte Schreibweise knüpft an die der Analysis an: Funktionsname mit 3 Buchstaben, x als Variablenname und in Klammern, Scharparameter als Index, also $Bin_{n,p}(x)$ analog zu $sin(x)$ bzw. $f_t(x)$. Entsprechend wird bei den übrigen Verteilungen $Hyp_{n,R,N}(x)$, $Poi_{\mu}(x)$ und $Nor_{\mu,\sigma}(x)$ verfahren.