

FACHARBEIT

in dem Fach

Leistungskurs Mathematik

(Schuljahre 1996 - 1998)

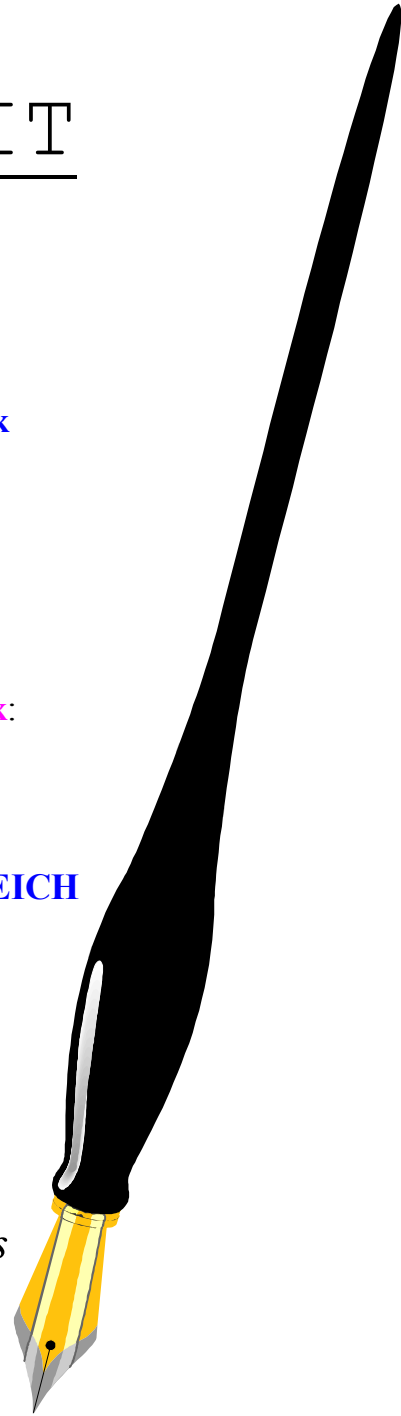
Thema:

Aus dem Fachgebiet **Stochastik**:

**VERTEILUNGEN:
ÜBERSICHT UND VERGLEICH**

Abgabetermin: 2. Februar 1998

Verfasser: *Paul Holleis*



INHALTSVERZEICHNIS

| | |
|--|-----------|
| VORWORT | 3 |
| I. GRUNDLAGEN | 5 |
| 1. Zufallsgröße..... | 5 |
| 2. Wahrscheinlichkeitsfunktion..... | 7 |
| 3. Verteilungsfunktion..... | 9 |
| 1. Diskrete Verteilungen..... | 10 |
| 2. Stetige Verteilungen | 12 |
| 3. Maßzahlen / Momente von Verteilungsfunktionen | 14 |
| a) Erwartungswert..... | 14 |
| b) Varianz (Streuung) | 17 |
| c) Standardabweichung..... | 19 |
| 4. Standardisierung | 20 |
| II. SPEZIELLE VERTEILUNGEN | 22 |
| 1. Binomialverteilung..... | 22 |
| 2. Hypergeometrische Verteilung..... | 26 |
| 3. Negativ-Binomialverteilung | 29 |
| 4. Geometrische Verteilung..... | 31 |
| 5. Poisson-Verteilung..... | 33 |
| 6. Normalverteilung | 36 |
| III. ZUSAMMENHÄNGE ZWISCHEN DEN VERTEILUNGEN (GRENZWERTSÄTZE), NÄHERUNG DER | 38 |
| 1. Binomialverteilung an die hypergeometrische Verteilung | 38 |
| 2. Poisson-Verteilung an die Binomialverteilung | 43 |
| 3. Poisson-Verteilung an die hypergeometrische Verteilung | 48 |
| 4. Normalverteilung an die Binomialverteilung | 51 |
| 5. Normalverteilung an die hypergeometrische Verteilung | 57 |
| 6. Normalverteilung an die Poisson-Verteilung..... | 59 |
| IV. ÜBERSICHT / ZUSAMMENFASSUNG | 62 |
| ANHANG | 63 |

VORWORT

Der Themenkreis „Stochastik“ enthält sowohl die Wahrscheinlichkeitsrechnung als auch den Bereich Statistik. Während letzterer sich hauptsächlich mit dem Testen von Hypothesen und der Auswertung von empirisch ermitteltem Zahlenmaterial beschäftigt, spielen in der Wahrscheinlichkeitsrechnung vor allem Verteilungen eine übergeordnete Rolle. Diese Facharbeit – „Verteilungen, Übersicht und Vergleich anhand von Beispielen“ – soll einen Überblick über die wichtigsten Verteilungen und ihre Zusammenhänge geben.

Bei den dargestellten Verteilungen wurde bewusst auf zu komplizierte Herleitungen und Beweise verzichtet. So wurden sie zum Beispiel nicht auf höhere Momente wie Schiefe und Exzess hin untersucht. Dies war nötig, um einen angemessenen Umfang der Arbeit zu gewährleisten. Außerdem sollten die Ausführungen grundsätzlich das Niveau des Mathematikunterrichts nicht überschreiten.

Bei der Auswahl der Beispiele wurde viel Wert auf deren Anschaulichkeit beziehungsweise praktische Anwendbarkeit gelegt, um auch einem Leser, der weniger Wert auf abstrakte mathematische Formulierungen und Berechnungen legt, einen Anreiz zu bieten, sich mit diesem Gebiet zu beschäftigen.

Dem Verfasser war es besonders wichtig, das Thema in einem abgeschlossenen Rahmen zu präsentieren, um es auch Lesern nahebringen zu können, die sich vorher nicht oder nur wenig mit der Materie beschäftigt haben. So ist es zu rechtfertigen, dass sich der erste Teil hauptsächlich mit den Grundlagen von Verteilungen beschäftigt, die bei entsprechenden Vorkenntnissen kurz überflogen oder ganz überblättert werden können. Prinzipiell sollten aber möglichst wenig Grundlagen vorausgesetzt werden müssen. Bedingt durch den beschränkten Umfang der Arbeit ließ es sich jedoch nicht vermeiden, auf eine detaillierte Ausführung bestimmter Bereiche zu verzichten und diese Kenntnisse als bekannt vorauszusetzen. Darunter fallen zum Beispiel der allgemeine Wahrscheinlichkeitsbegriff und grundlegende kombinatorische Fertigkeiten. Grundsätzlich sollte es aber auch für einen Neuling in diesem speziellen Bereich der Stochastik möglich sein, den wesentlichen Inhalt nachvollziehen zu können.

HINWEISE FÜR DEN BENUTZER:

Dem Leser, der sich hauptsächlich für die Ergebnisse dieser Arbeit interessiert, dürfte vor allem die groß angelegte **Übersicht** auf Seite 62 von Nutzen sein. Hier sind alle behandelten Verteilungen mit ihren Charakteristiken, Anwendungsbereichen und gegenseitigen Beziehungen plakativ aufgelistet.

Zur Verbesserung der Übersichtlichkeit und Vereinfachung der Lesbarkeit wurden bestimmte Textbereiche einheitlich farblich markiert beziehungsweise hervorgehoben:

Beispiele: Das Zufallsexperiment bzw. Schlagworte **grün**; grüner Strich am linken Rand;

Definitionen: Gesamte Definition rot umrahmt; der jeweils definierte Begriff **rot**;

Schlüsselbegriffe: Mit **blau** hervorgehoben; **wichtige Formeln blau umrahmt**;

Faustregeln: Der gesamte Absatz gestrichelt blau umrahmt; Bedingungen **blau**;

I. Grundlagen

Verteilungen sind spezielle Funktionen, die bestimmte Ereignisse und Experimente beschreiben. Die Verteilungsfunktion basiert auf der Wahrscheinlichkeitsfunktion. Beide sind auf der sogenannten Zufallsgröße X aufgebaut. Zum Vergleich verschiedener Verteilungen benötigt man zusätzlich Kenntnisse über den Erwartungswert, die Varianz und die Standardabweichung. Diese Grundlagen sollen nun zuerst einmal vorgestellt werden:

I. 1. Zufallsgröße

Hat man einen Versuch ausgeführt, bei dem es mehrere Ergebnisse gibt, hat jedes eine bestimmte Wahrscheinlichkeit. Man verknüpft diese beiden nun mit Hilfe einer Funktion. Um den Graph dieser Funktionen später auch zeichnen zu können, verwendet man für die Ereignisse nicht Buchstaben („A“, „B“, „E“, ...) oder Beschreibungen („erster Wurf ist eine Sechs“, ...), sondern Zahlen aus dem Bereich der reellen Zahlen.

Definition: Eine Funktion X heißt **Zufallsgröße**, wenn sie jedem möglichen Ergebnis ω eines Zufallsexperiments eine reelle Zahl zuordnet.

Die Definitionsmenge der Funktion D_X ist also die Ergebnismenge Ω des Zufallsexperiments, die Wertemenge W_X ist eine Teilmenge von \mathbb{R} .

$$X : \Omega \rightarrow W_X, \quad \omega \mapsto X(\omega)$$

(Die Funktion X bildet den Ergebnisraum Ω auf $W_X \subseteq \mathbb{R}$ ab; jedem Ergebnis ω wird eine Zahl $X(\omega)$ zugeordnet)

X ist also keine einfache Zahl, sondern eine Funktion. Die Funktionsgleichung ist in den allermeisten Fällen nur abschnittsweise definiert zu beschreiben.

Beispiel: Beim **Roulette** gibt es insgesamt 37 verschiedene Zahlen (0-36), die auf unterschiedlichste Arten zusammengefasst werden können. Die Zufallsgröße X ordnet diesen Ereignissen nun verschiedene Zahlen zu:

1. Beispiel:

| $\Omega = D_X$ | $W_X \in \mathbb{R}$ |
|----------------|----------------------|
| Rot | 1 |
| Schwarz | 2 |
| Null | 3 |

Die Funktion X fasst hier also mehrere Ereignisse zusammen, so dass nur noch genau die drei Ereignisse die von Interesse sind auftreten. Sie lautet:

$$X(\omega) = \begin{cases} 1 & \text{für } \omega \in \{\text{"rote Zahlen"}\} \\ 2 & \text{für } \omega \in \{\text{"schwarze Zahlen"}\} \\ 3 & \text{für } \omega = \{0\} \end{cases}$$

2. Beispiel:

| $\Omega = D_X$ | $W_X \in \mathbb{R}$ |
|----------------|----------------------|
| 1. Drittel | 1 |
| 2. Drittel | 2 |
| 3. Drittel | 3 |
| Null | 4 |

Wieder komprimiert die Funktion X alle Ereignisse auf vier Zahlen. Deshalb kann man die Funktion so beschreiben:

$$X(\omega) = \begin{cases} 1 & \text{für } \omega \in \{1,2,3,\dots,12\} \\ 2 & \text{für } \omega \in \{13,14,\dots,24\} \\ 3 & \text{für } \omega \in \{25,26,\dots,36\} \\ 4 & \text{für } \omega = \{0\} \end{cases}$$

I. 2. Wahrscheinlichkeitsfunktion

Wenn man nun den möglichen Ereignissen jeweils eine Zahl zugeordnet hat, kann man nach den Wahrscheinlichkeiten für diese Zahlen fragen.

Jede Zahl wird so wiederum mit ihrer Wahrscheinlichkeit verbunden. Diese Zuordnung geschieht mit Hilfe der Wahrscheinlichkeitsfunktion.

Definition: Ist X eine diskrete Zufallsgröße mit der Wertemenge $W_X = \{a_1, a_2, \dots, a_n, \dots\}$, so bezeichnet man die Funktion f , die jedem $x \in W_X$ die Wahrscheinlichkeit des Ereignisses $\{X = x\}$ zuordnet, als **Wahrscheinlichkeitsfunktion der Zufallsgröße X** :

$$f: W_X \rightarrow [0; 1], x \mapsto f(x), \text{ mit } f(x) = P(X = x)$$

(Die Funktion f bildet die Wertemenge von X (die Zahlen) auf Wahrscheinlichkeiten ab; jeder Zahl x wird eine Wahrscheinlichkeit $f(x)$ zugeordnet)

Um den Definitionsbereich der Funktion f zu vervollständigen, vereinbart man zusätzlich:

$$\text{für alle } x \notin W_X \text{ gilt } f(x) = 0$$

Dies ist verständlich, wenn man bedenkt, dass eine Zahl, die nicht in der Wertemenge von X ist, also keine zugeordneten Ereignisse hat, nie auftreten kann und somit ihre Wahrscheinlichkeit 0 ist.

Beispiele

In dem oben genannten **Roulette** lassen sich die Wahrscheinlichkeiten für die Zahlen in der Wertemenge der Zufallsgröße X leicht berechnen:

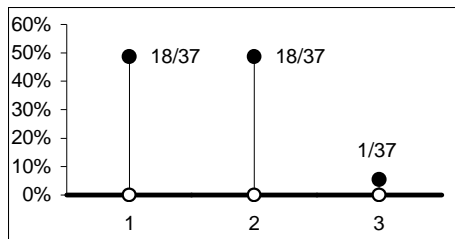
1. Beispiel:

$$P(1) = P(\text{„rot“}) = 18/37;$$

$$P(2) = P(\text{„schwarz“}) = 18/37;$$

$$P(3) = P(\text{„Null“}) = 1/37;$$

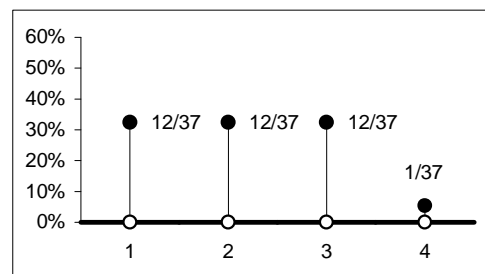
Damit lässt sich $f(x)$ zeichnen:



Die dick durchgezogene Linie auf der X-Achse deutet an, dass überall neben den gekennzeichneten Stellen (bei genau 1, 2 und 3) die Wahrscheinlichkeit gleich 0 ist.

2. Beispiel:

Fasst man mit der Zufallsgröße die Ereignisse 1-12 (**unteres Drittel**), 13-24 (**mittleres Drittel**), 25-36 (**oberes Drittel**) und 0 (**Null**) zusammen, erhält man folgende Wahrscheinlichkeitsfunktion:



$$P(1) = P(\text{„1-12“}) = 12/37;$$

$$P(2) = P(\text{„13-24“}) = 12/37;$$

$$P(3) = P(\text{„25-36“}) = 12/37;$$

$$P(4) = P(\text{„0“}) = 1/37;$$

Aus den Werten der Diagramme lässt sich eine wichtige Eigenschaft der Wahrscheinlichkeitsfunktion ablesen: Addiert man die Wahrscheinlichkeiten aller Zahlen, so erhält man 1.

Es gilt also stets: $\sum_{i=1}^n f(x_i) = 1$

Beweis: $\sum_{i=1}^n f(x_i) = f(x_1) + f(x_2) + \dots + f(x_n) = f(E_1 \cap E_2 \cap \dots \cap E_n) = f(\Omega) = 1$

I. 3. Verteilungsfunktion

Bei vielen Problemen in der Stochastik ist nicht nur die Kenntnis der Wahrscheinlichkeit dafür wichtig, dass eine Zufallsgröße X genau einen bestimmten Wert x ($\in \mathbb{R}$) annimmt, also das Ereignis $\{X = x\}$ eintritt, sondern auch dafür, dass sie in einem Intervall liegt oder einen Wert unter- beziehungsweise überschreitet.

Bei technischen Prozessen zum Beispiel ist die Wahrscheinlichkeit von Interesse, mit der eine zulässige Höchstbelastung überschritten wird, bei physikalischen Experimenten sind unter Umständen kritische Temperaturgrenzen einzuhalten.

Die Wahrscheinlichkeit, dass die Zufallsgröße X Werte unterhalb einer Grenze x annimmt, hängt nur von diesem x ab und ist somit eine Funktion von x . Diese Funktion nennt man Verteilungsfunktion der Zufallsgröße X und bezeichnet sie mit $F(x)$.

Der Zusammenhang zwischen der Wahrscheinlichkeits- und der Verteilungsfunktion liegt nahe: Fragt man nach der Wahrscheinlichkeit, dass ein Ereignis oder ein anderes auftritt, so zählt man die beiden Einzelwahrscheinlichkeiten einfach zusammen. Dies ist natürlich nur möglich, wenn diese Ereignisse disjunkt sind, das heißt $E1 \cap E2 = \{\}$. Da diese Eigenschaft aus der Definition der Zufallsgröße erfolgt (jedem Ereignis wird genau eine Zahl zugeordnet), gilt nach dem

Axiom III: Da $E1 \cap E2 = \{\}$ so gilt $P(E1 \cup E2) = P(E1) + P(E2)$.

Daraus folgt folgende Festlegung:

Definition: Eine Funktion $F(X)$ heißt (kumulative) Verteilungsfunktion der Zufallsgröße X , wenn sie die Wahrscheinlichkeit für alle Ereignisse $\{X \leq x\}$ wiedergibt:

$$F(x) = P(X \leq x) = \sum_{x^i \leq x} P(X = x^i) = \sum_{x^i \leq x} f(x^i)$$

Aus der weiter oben gewonnenen Eigenschaft, dass die Summe der Wahrscheinlichkeiten aller Ereignisse gleich 1 ist, erkennt man:

für $x \geq x_{\max}$ gilt $F(x) = 1$, mit x_{\max} als höchste Zahl, die die Zufallsgröße annehmen kann.

für $x \leq x_{\min}$ gilt $F(x) = 0$, mit x_{\min} als niedrigste Zahl, die die Zufallsgröße annehmen kann.

I. 3. 1. Diskrete Verteilungen

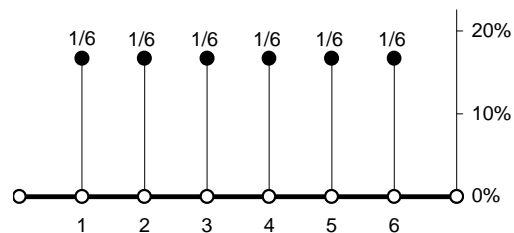
Bei den meisten Zufallsexperimenten (wie unter anderem würfeln oder Münzen werfen), nimmt die auftretende Zufallsgröße nur **endlich** (bzw. **abzählbar unendlich**) viele Werte an. Man nennt sie **diskret**.

Für eine **diskrete Zufallsgröße X** gilt:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} f(x_i)$$

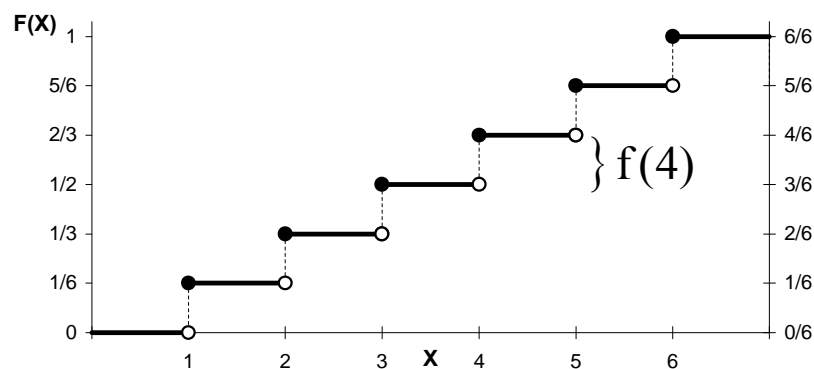
die Summe erstreckt sich über alle Werte $f(x_i)$, für die $X \leq x$ zutrifft

Beispiel: $X =$ Augenzahl beim Werfen eines 6-Seitigen Würfels



Da die sechs möglichen Ergebnisse alle gleichwahrscheinlich sind ($p = 1/6$), ergibt sich für die Wahrscheinlichkeitsfunktion das nebenstehende Bild. Die Wahrscheinlichkeiten anderer Werte von X als 1, 2, 3, 4, 5 und 6 sind immer null:

Aus diesen Wahrscheinlichkeiten lässt sich die zugehörige Verteilungsfunktion bilden:



Berechnung:

$$F(1) = P(X \leq 1) = P(1) = 1/6;$$

$$F(2) = P(X \leq 2) = P(1) + P(2) = 2/6$$

usw.

Werte außerhalb D_X : z.B.: $F(1,5) = P(X \leq 1.5) = P(X \leq 1) = P(1) = \frac{1}{6}$, da zwischen 1 und 1,5 keine Erhöhung der Wahrscheinlichkeit stattfindet, da die Zufallsvariable diese Werte nicht annehmen kann, d.h. es sind alle $P(1 < X \leq 1,5) = 0$ (unmögliche Ereignisse);

Für $x \geq 6$ ergibt sich immer $F(x) = 1$, die Wahrscheinlichkeit des sicheren Ereignisses, da alle möglichen Werte, die der Würfel (bzw. die zugehörige Zufallsgröße) annehmen kann, eingeschlossen sind.

Die Verteilungsfunktionen diskreter Zufallsgrößen sind also **Treppenfunktionen**. An den Stellen $X = x_i$ haben sie **Sprungstellen** der Höhe $f(x_i)$. Der **Wertebereich** geht von **0** (unmögliches Ereignis, bis $-\infty$) bis **1** (sicheres Ereignis, bis $+\infty$).

Wenn X (abzählbar) unendlich viele Werte annimmt, so muss die Summe trotzdem einem Grenzwert zustreben, da selbst bei Summation über alle x_i der Wert die Wahrscheinlichkeit 1 nicht überschreiten kann / darf.

Aus der Verteilungsfunktion lassen sich nun auf einfache Weise Wahrscheinlichkeiten für folgende Ereignisse ablesen:

$$P(X \leq a) = F(a)$$

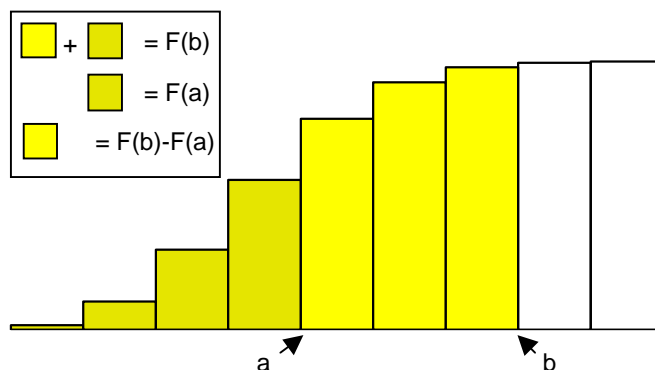
, nach Definition;

$$P(X > a) = 1 - F(a)$$

, da dies das Gegenereignis zu $P(X \leq a)$ ist;

$$P(a < X \leq b) = F(b) - F(a)$$

Die letzte Formel kann man sich leicht an folgendem Diagramm veranschaulichen:



I. 3. 2. Stetige Verteilungen

Viele Experimente in der Praxis sind nicht so einfach einzuteilen.

Der **Zeiger einer Drehscheibe** zum Beispiel kann jede reelle Zahl zwischen 0 und 2π (0° bis 360°) annehmen.

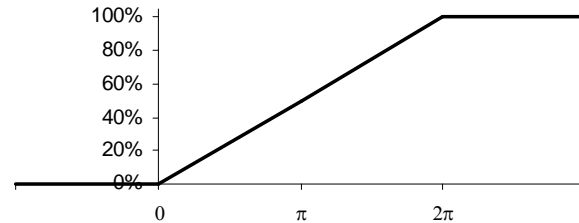
Die zugehörige Zufallsgröße ist hier **überabzählbar unendlich**, man nennt sie **stetig**.

Einige Eigenschaften der diskreten Zufallsgröße lassen sich auch bei der stetigen anwenden.

Es ist auch hier möglich, die Verteilungsfunktion $F(x) = P(X \leq x)$ zu bilden.

Für das Beispiel „**X = Richtung des Zeigers**“ ergibt sich folgende Funktion:

$$F(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ \frac{1}{2\pi} x & \text{für } 0 < x \leq 2\pi \\ 1 & \text{für } x > 2\pi \end{cases}$$



Für eine **stetige Zufallsgröße X** gilt:

$$F(x) = P(X \leq x) = \int_0^x f(v) dv$$

Hier ist die Verteilungsfunktion keine Treppenfunktion mehr. Sie ist im gesamten Definitionsbereich $D = \mathbb{R}$ **stetig**.

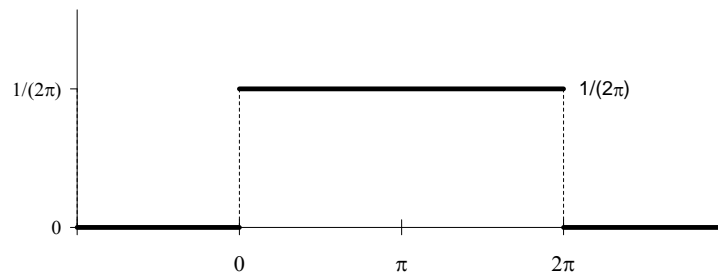
Bis auf die beiden Knickstellen ist sie auch **differenzierbar**. Die Ableitung

$$F'(x) = f(x) \quad (\text{'groß F Strich von x ist klein f von x'})$$

macht eine Aussage darüber, wie stark die Wahrscheinlichkeiten anwachsen. Ein großer Anstieg an einer Stelle x bedeutet aber, dass sehr viele Versuchsergebnisse in dem Bereich um x zu erwarten sind.

Daher nennt man die **Ableitung der Verteilungsfunktion** einer stetigen Zufallsgröße auch **Dichtefunktion**: $f(x) = F'(x)$

Die Dichtefunktion des Beispiels „ $X = \text{Richtung des Zeigers}$ “ hat folgende Gestalt:



Die gemeinsame Verwendung des Buchstabens f für die Dichtefunktion und die Wahrscheinlichkeitsfunktion hat folgenden Grund: Wie bereits festgestellt wurde, ist die **Sprunghöhe** einer diskreten Verteilungsfunktion genau $f(x_i)$. Die Wahrscheinlichkeitsfunktion legt also ebenfalls die „**Steigung**“ ihrer Verteilungsfunktion fest. In Formeln ausgedrückt erkennt man den Zusammenhang sehr schnell:

$$F(x) = \int_{-\infty}^x f(v) dv \quad \text{bei **stetigen** Zufallsgrößen und}$$

$$F(x) = \sum_{x_i \leq x} f(x_i) \quad \text{bei **diskreten** Zufallsgrößen.}$$

Beide Funktionen geben die **Gesamtwahrscheinlichkeit bis zur Stelle x** , also $P(X \leq x)$ an.

Wie schon bei diskreten, so kann man auch bei stetigen Zufallsgrößen Wahrscheinlichkeiten direkt ablesen. Die Begründungen wären analog zum diskreten Fall zu führen. Deshalb sind sie hier weggelassen und nur die Ergebnisse festgehalten:

$$P(X > a) = 1 - F(a) = \int_a^{\infty} f(x) dx \qquad P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f(x) dx = 0; \qquad \text{(im stetigen Fall lassen sich die Gleichheitszeichen beliebig verteilen)}$$

Bei der letzten Formel erkennt man einen Unterschied zwischen der Wahrscheinlichkeitsfunktion und der Dichtefunktion: Wenn a ein unmögliches Ereignis ist, dann weiß man, dass die Wahrscheinlichkeit $f(a) = 0$ sein muss. Bei einer diskreten Zufallsgröße kann man außerdem aus $f(a) = 0$ darauf schließen, dass a ein unmögliches Ereignis ist. Bei stetigen Verteilungen ist dies nicht der Fall, da $f(a)$ für jede Zahl gleich 0 ist. Das Ereignis selber ist aber trotzdem nicht unbedingt unmöglich. Dies liegt daran, dass sich die Wahrscheinlichkeit auf unendlich viele mögliche Ereignisse aufteilt. Für jeden einzelnen Punkt bleibt dann nur noch unendlich wenig ($\neq 0$) übrig.

I. 3. 3. Maßzahlen / Momente von Verteilungsfunktionen

Um eine Verteilung zu beschreiben verwendet man sogenannte **Momente**. Sie sind **Maßzahlen**, die den **Verlauf von Verteilungen** genauer beschreiben. Die drei wichtigsten sind **Erwartungswert**, **Varianz** und **Standardabweichung**.

I. 3. 3. a) Erwartungswert

Der Erwartungswert $E(X)$ (meist einfach μ genannt) macht eine Aussage über die Lage der Wahrscheinlichkeitsfunktion bzw. Dichte im Koordinatensystem, also darüber, welchen Wert die Zufallsgröße X „im Mittel“ annimmt. Deshalb bezeichnet man ihn auch als **Lageparameter**.

Statistischer / mathematischer Mittelwert

Das **arithmetische Mittel** von n Zahlen x_1, x_2, \dots, x_n lässt sich bekanntlich folgendermaßen berechnen:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad [\text{z.B.: } 1, 2, 2, 2, 3: (1+2+2+2+3)/5 = 10/5 = 2]$$

Geht man davon aus, dass die Zahlen x_i untereinander verschieden sind (was bei Betrachtungen mit Zufallsgrößen ja gegeben ist), so muss man die (absolute) Häufigkeit $H(x)$ ihres Auftretens berücksichtigen. Man spricht vom **gewogenen Mittel**:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k H(x_i) \cdot x_i \quad [(1, 2, 2, 2, 3: (1 \cdot 1 + 3 \cdot 2 + 1 \cdot 3)/5 = 10/5 = 2]$$

(k ist die Anzahl der untereinander verschiedenen Zahlen aus den n vorhandenen).

Nachdem $\frac{1}{n} H(x_i)$ gleichbedeutend mit der relativen Häufigkeit $h(x_i)$ ist, ergibt sich damit für

den Mittelwert: $\bar{x} = \sum_{i=1}^k h(x_i) \cdot x_i$

Der Mittelwert ist also die **Summe der auftretenden verschiedenen Werte von x , gewogen jeweils mit ihrer relativen Häufigkeit**.

Stochastischer Erwartungswert

Die **relativen Häufigkeiten** $h(x_i)$ sind ja (bei vielen Durchführungen des Experiments) Näherungswerte für die Wahrscheinlichkeit $f(x_i)$ (\rightarrow *Gesetz der Großen Zahlen*). Deshalb definiert man für die Behandlung von Zufallsgrößen einen entsprechenden Mittelwert und nennt ihn Erwartungswert $E(X)$:

Ist X eine **diskrete** Zufallsgröße mit der Wertemenge $W_X = \{x_1, x_2, \dots, x_n\}$ und den Wahrscheinlichkeiten $f(x_1), f(x_2), \dots, f(x_n)$, so heißt

$$E(X) = \mu = \sum_{i=1}^n x_i \cdot f(x_i) \quad \text{Erwartungswert der Zufallsgröße } X^1.$$

Diese Definition lässt sich leicht auf stetige Zufallsgrößen übertragen. Für den Erwartungswert verwendet man hier statt der Wahrscheinlichkeitsfunktion die Dichtefunktion und ersetzt die Summe durch ein Integral:

Ist X eine **stetige** Zufallsgröße mit der Dichte f , so heißt

$$E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x) dx \quad \text{Erwartungswert der Zufallsgröße } X^1.$$

Der Erwartungswert ist eine reelle Zahl, die keineswegs in der Definitionsmenge der Zufallsgröße enthalten sein muss.

So gilt zum Beispiel bei dem **Wurf mit einem sechsseitigen L-Würfel**:

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \underline{\underline{3,5}}.$$

Für die Zufallsgröße X gelten folgende **Sätze**:

- Sind alle Werte x_i **gleichwahrscheinlich**, so ist $E(X)$ gleich dem **arithmetischen Mittel** der Werte x_i (so war es bei der obigen Berechnung von $E(X)$ für einen Würfel).
- Sind die Wahrscheinlichkeiten **symmetrisch** um eine Zahl c angeordnet, so ist:
 $E(X) = E(c) = c$ (so ist es bei der Augensumme zweier Würfel: $E(X) = 7$).

¹ Den Erwartungswert bezeichnet man mitunter auch als Moment 1. Ordnung

- $E(c \cdot X) = c \cdot E(X)$; [Bew.: $E(c \cdot X) = \sum_{i=1}^n c \cdot x_i \cdot P(x_i) = c \cdot \sum_{i=1}^n x_i \cdot P(x_i) = c \cdot E(X)$]
- $E(X + c) = E(X) + c$; [Bew.: $E(X + c) = \sum_{i=1}^n (x_i + c) \cdot P(x_i) = \sum_{i=1}^n x_i \cdot P(x_i) + \sum_{i=1}^n c \cdot P(x_i) = E(X) + c \cdot \sum_{i=1}^n P(x_i) = E(X) + c \cdot 1 = E(X) + c$]

Die Beweise für stetige Zufallsgrößen laufen analog.

$$\Rightarrow E(a \cdot X + b) = a \cdot E(X) + b$$

Im Beispiel „ $X = \text{Winkel des Zeigers einer Drehscheibe}$ “ ist f symmetrisch zu $x = \pi$ und damit ist $E(X) = \pi$. Man kann den Erwartungswert aber auch über die Formel ausrechnen:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^{2\pi} x \cdot \frac{1}{2\pi} dx = \left[\frac{1}{2\pi} \cdot \frac{x^2}{2} \right]_0^{2\pi} = \frac{4\pi^2}{4\pi} = \pi.$$

I. 3. 3. b) Varianz (Streuung)

Für die Beurteilung einer Verteilung ist nicht nur der Mittelwert interessant, sondern auch die „Streuung“, also die **Abweichungen von diesem Mittelwert**.

Diese Tatsache kann durch ein einfaches Beispiel der Statistik veranschaulicht werden:

Bei einem **Test von je 40 Schülern** ergaben sich folgende **Notenverteilungen**:

| | | | | | | | | | | | | | |
|---------|----|---|---|---|---|---|---------|---|---|----|---|---|---|
| Noten: | 1 | 2 | 3 | 4 | 5 | 6 | Noten: | 1 | 2 | 3 | 4 | 5 | 6 |
| Anzahl: | 12 | 8 | 3 | 8 | 3 | 6 | Anzahl: | 5 | 8 | 15 | 7 | 4 | 1 |

Möchte man nun beide Ergebnisse vergleichen, so rechnet man bei beiden zweckmäßigerweise den Mittelwert aus:

$$\bar{x}_1 = \frac{12 \cdot 1 + 8 \cdot 2 + 3 \cdot 3 + 8 \cdot 4 + 3 \cdot 5 + 6 \cdot 6}{12 + 8 + 3 + 8 + 3 + 6} = \frac{120}{40} = 3; \quad \bar{x}_2 = \frac{5 \cdot 1 + \dots + 1 \cdot 6}{40} = \frac{120}{40} = 3;$$

Beide Mittelwerte sind also gleich. Trotzdem sind die Verteilungen sehr unterschiedlich. Die meisten Noten der ersten Verteilung liegen viel weiter von \bar{x} entfernt, als bei der zweiten. Deshalb benötigt man zur genaueren Analyse einen Wert für die „Streuung“ um den Mittelwert. In der Wahrscheinlichkeitsrechnung stellt sich das gleiche Problem mit der Abweichung um den Erwartungswert.

Die Abweichung eines Messwertes x_i vom Erwartungswert μ ist durch die Differenz $x_i - \mu$ gegeben. Naheliegender wäre also, als Maß der mittleren Abweichung den Mittelwert all dieser

Differenzen zu nehmen: $\frac{1}{n} \sum_{i=1}^n (x_i - \mu)$. Dies ist jedoch nicht möglich, da gilt:

$$\text{aus } \mu = (x_1 + x_2 + \dots + x_n)/n \text{ folgt } (x_1 + x_2 + \dots + x_n) = n \cdot \mu$$

$$\sum_{i=1}^n (x_i - \mu) = (x_1 - \mu) + (x_2 - \mu) + \dots + (x_n - \mu) = (x_1 + x_2 + \dots + x_n) - n \cdot \mu = n \cdot \mu - n \cdot \mu = 0$$

Wie schon der Name Mittelwert andeutet, heben sich alle positiven und negativen Abweichungen von μ gegenseitig auf.

(Nach den früher bewiesenen Sätzen gilt ja auch: $E(X - \mu) = E(X) - E(\mu) = \mu - \mu = 0$.)

Da für die Abweichung weniger die Richtung als die absolute Größe von Bedeutung ist, ist der nächste logische Schritt, den absoluten Betrag der Differenz, also $|x_i - \mu|$ zu verwenden. Um das umständliche Rechnen mit Beträgen zu vermeiden, verwendet man statt dem Betrag das Quadrat und definiert:

Ist X eine Zufallsgröße mit dem Erwartungswert $E(X) = \mu$, so heißt die reelle Zahl

(im **diskreten** Fall) $\sigma^2 = V(X) = E((X - \mu)^2) = \sum_i (x_i - \mu)^2 \cdot f(x_i)$

(im **stetigen** Fall) $\sigma^2 = V(X) = E((X - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx$

Varianz² von X .

[Eine Varianz von 0 kommt also nur dann vor, wenn eine sogenannte Einpunktverteilung vorliegt, also die Zufallsgröße nur den Wert μ annimmt.]

Aufgrund der Eigenschaft der **Varianz**, den *Verlauf der Verteilung* zu beschreiben, bezeichnet man sie als **Formparameter** im Gegensatz zu dem **Lageparameter Erwartungswert**.

Die Wahl des Quadrats hat außerdem den Vorteil, dass stärkere Abweichungen vom Erwartungswert auch mehr ins Gewicht fallen als kleinere.

Für die Varianz gelten folgende **Formeln** (auf die etwas längeren Beweise wird hier zugunsten der Übersichtlichkeit verzichtet):

Wichtig ist vor allem der folgende Spezialfall des **Verschiebungssatzes** für $a = 0$:

- $V(X) = E((X - a)^2) - (E(X) - a)^2 \Rightarrow V(X) = E(X^2) - E(X)^2 = E(X^2) - \mu^2$;

(„Varianz ist der Erwartungswert der Quadrate minus dem Quadrat des Erwartungswerts“)

- $V(a \cdot X + b) = a^2 \cdot V(X)$;

- $V(c) = 0$ [Bew.: $V(c) = E(c^2) - E(c)^2 = c^2 - c^2 = 0$

oder: $V(c) = E[c - E(c)]^2 = E(c - c) = E(0) = 0$]

² Andere Bezeichnungen für die Varianz sind: Streuungsquadrat, mittlere quadratische Abweichung, $D^2(X)$ (vom englischen ‘deviation’), Dispersion, zentrales Moment 2. Ordnung

I. 3. 3. c) Standardabweichung

Da bei vielen praktischen Anwendungen **Elemente eines Größenbereichs** vorkommen (also Werte mit einer Benennung), verwendet man auch die Wurzel aus der Varianz, die dann die gleiche Einheit hat, wie die Größen selbst.

Man nennt $\sigma = \sqrt{\text{Var}(X)}$ die **Standardabweichung**³ einer Verteilung.

Bei dem vorher behandelten **Notenbeispiel** erkennt man nun die unterschiedlichen Verteilungen an den Werten für die Standardabweichung σ :

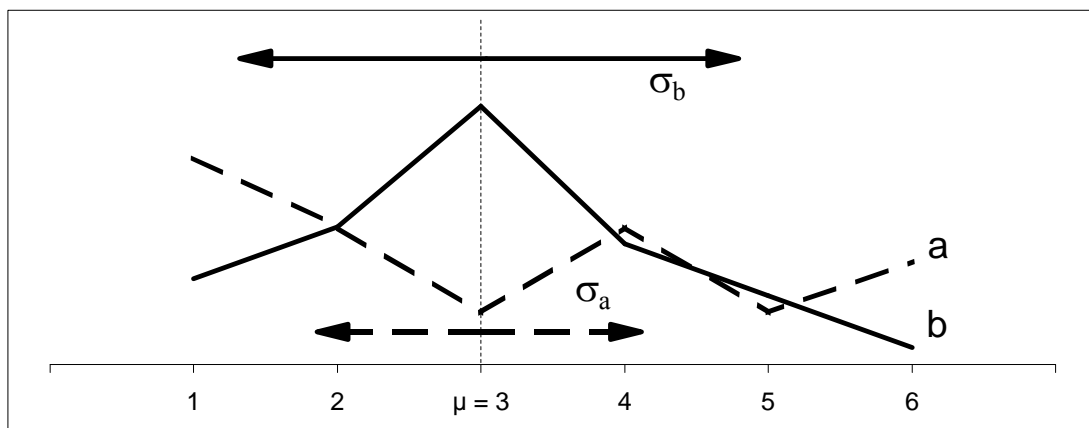
$$1) (1-3)^2 \cdot \frac{12}{40} + (2-3)^2 \cdot \frac{8}{40} + (3-3)^2 \cdot \frac{3}{40} + (4-3)^2 \cdot \frac{8}{40} + (5-3)^2 \cdot \frac{3}{40} + (6-3)^2 \cdot \frac{6}{40} =$$

$$= (4 \cdot 12 + 1 \cdot 8 + 0 \cdot 3 + 1 \cdot 8 + 4 \cdot 3 + 9 \cdot 6) / 40 = 130 / 40 = 3,25 \Rightarrow \underline{\underline{\sigma = 1,80}}$$

$$2) (1-3)^2 \cdot \frac{5}{40} + (2-3)^2 \cdot \frac{8}{40} + (3-3)^2 \cdot \frac{15}{40} + (4-3)^2 \cdot \frac{7}{40} + (5-3)^2 \cdot \frac{4}{40} + (6-3)^2 \cdot \frac{1}{40} =$$

$$= (4 \cdot 5 + 1 \cdot 8 + 0 \cdot 15 + 1 \cdot 7 + 4 \cdot 4 + 9 \cdot 1) / 40 = 60 / 40 = 1,5 \Rightarrow \underline{\underline{\sigma = 1,22}}$$

Im folgenden Diagramm sind zu dieser Notenverteilung die jeweiligen Standardabweichungen und der Mittelwert eingetragen. Mit Hilfe der Varianz erkennt man sofort die unterschiedliche Verteilung der beiden Notenergebnisse trotz der jeweils gleichen Durchschnittsnote:



³ Die Standardabweichung wird oft auch als Streuung oder $D(X)$ bezeichnet

I. 3. 4. Standardisierung

Um unterschiedliche Verteilungen gut miteinander **vergleichen** zu können, müssen sie bestimmte Bedingungen erfüllen.

Die sogenannte **Standardisierung** transformiert eine Zufallsgröße so, dass sie den Erwartungswert $E(X) = 0$ und die Standardabweichung $\sigma = 1$ hat. Diese neue Zufallsgröße nennt man meist Z^4 .

$\Rightarrow Z := \frac{X - E(X)}{\sigma} = \frac{X - \mu}{\sigma}$ heißt die zu X gehörende **standardisierte Zufallsgröße**.

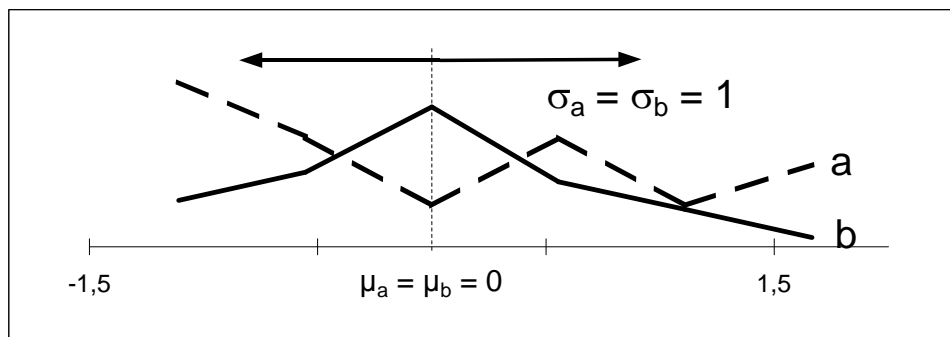
Bei ihr gilt: $E(Z) = 0$ und $\sigma = 1$.

Unter Verwendung der bewiesenen Formeln für den Erwartungswert und die Standardabweichung lässt sich folgender Beweis führen (man berechnet den Erwartungswert und die Standardabweichung für beliebige Zufallsgrößen X):

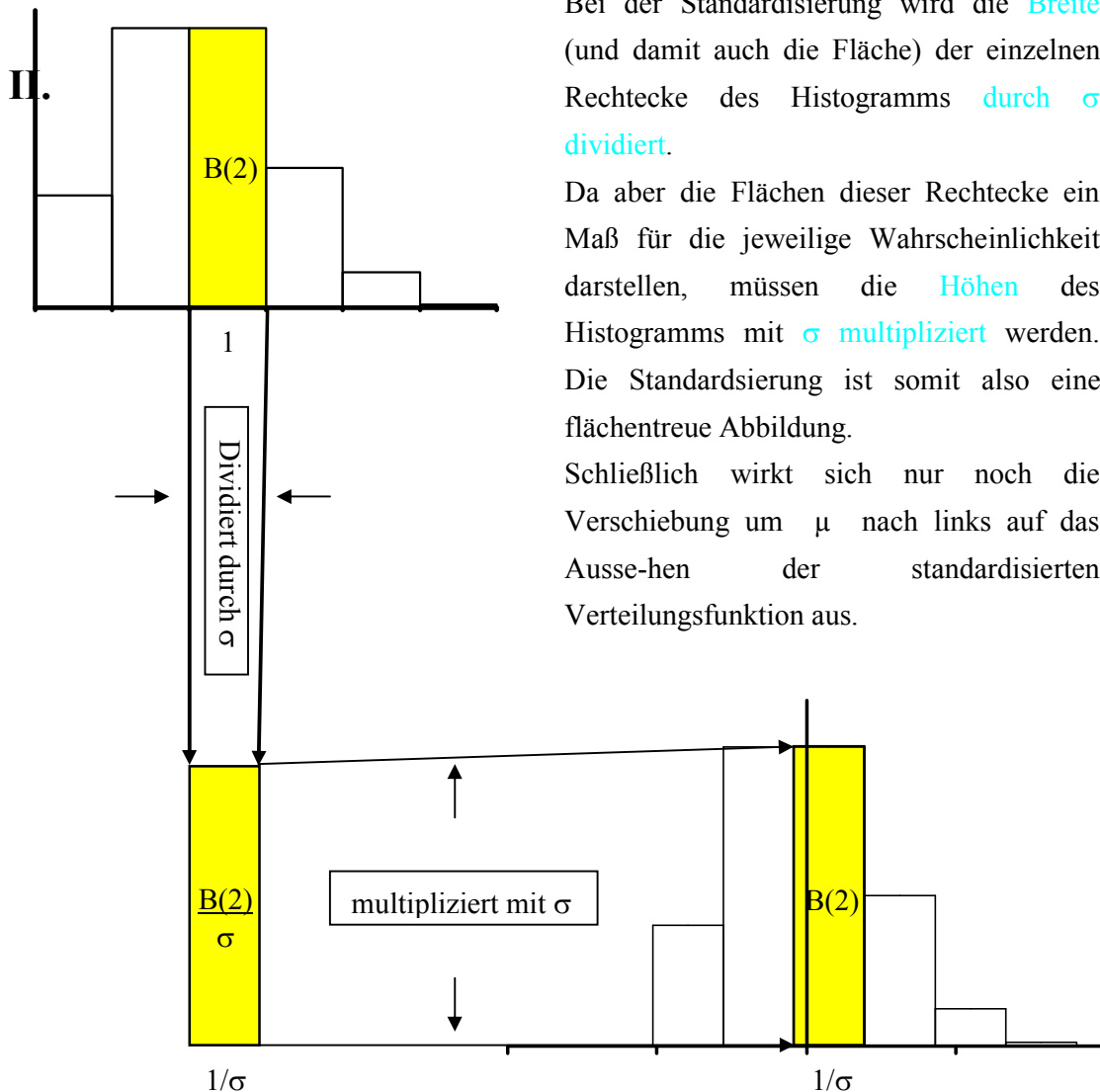
$$E(Z) = E\left(\frac{X - E(X)}{\sigma}\right) = E\left(\frac{1}{\sigma} \cdot (X - E(X))\right) = \frac{1}{\sigma} \cdot (E(X) - E(E(X))) = \frac{1}{\sigma} \cdot (E(X) - E(X)) = 0$$

$$\sigma(Z) = \sigma\left(\frac{X - E(X)}{\sigma(X)}\right) = \sigma\left(\frac{X}{\sigma(X)} - \frac{E(X)}{\sigma(X)}\right) = \sigma\left(\frac{1}{\sigma(X)} \cdot X\right) = \sigma\left(\frac{1}{\sigma(X)}\right) \cdot \sigma(X) = \frac{\sigma(X)}{\sigma(X)} = 1$$

An der Standardisierung der beiden Notenverteilungen von vorher erkennt man, dass die Verteilungen trotz gleichem $E(X)$ und σ ihre charakteristischen Eigenschaften behalten haben:



Anschaulich lässt sich der Vorgang der **Standardisierung** zum Beispiel gut bei der später genauer eingeführten Binomialverteilung zeigen. Bei der folgenden Wahrscheinlichkeitsfunktion (mit völlig willkürlich gewählten Parametern) ist ein Balkenstück des Graphen hervorgehoben:



Bei der Standardisierung wird die **Breite** (und damit auch die Fläche) der einzelnen Rechtecke des Histogramms **durch σ dividiert**.

Da aber die Flächen dieser Rechtecke ein Maß für die jeweilige Wahrscheinlichkeit darstellen, müssen die **Höhen** des Histogramms mit σ **multipliziert** werden. Die Standardisierung ist somit also eine flächentreue Abbildung.

Schließlich wirkt sich nur noch die Verschiebung um μ nach links auf das Aussehen der standardisierten Verteilungsfunktion aus.

Die Breite des Flächenstück $B(2)$ wurde im ersten Schritt von 1 auf $1/\sigma$ verkürzt. Deshalb wird die Höhe von $B(2)$ mit σ auf $\sigma \cdot B(2)$ vergrößert.

Die Höhen der standardisierten Verteilungsfunktion sind also $\sigma \cdot B(k)$.

⁴ Auch die Bezeichnung U ist üblich (vom englischen 'unit'). Wegen der Division durch σ ist U immer dimensionslos.

Spezielle Verteilungen

Nach diesen allgemeinen Betrachtungen über Wahrscheinlichkeits-, Dichte- und Verteilungsfunktionen sollen nun die *wichtigsten besonderen Verteilungen der Stochastik vorgestellt werden*. Sie haben alle eine große praktische Bedeutung, auf die in den jeweiligen Abschnitten noch besonders eingegangen wird. Die Vorstellung dieser Verteilungen ist deshalb so wichtig, da nur so *einheitliche Schreibweisen und Bezeichnungen* garantiert werden können, um danach keine Probleme bei dem Vergleich der Verteilungen zu haben.

II. 1. Binomialverteilung

Die wohl am häufigsten eingeführte Verteilung beruht auf den sogenannten **Bernoulli-Experimenten**⁵. Diese zeichnen sich dadurch aus, dass sie (zumindest theoretisch) beliebig oft wiederholbar sind, und dass sich trotzdem die Wahrscheinlichkeit für die möglichen Ereignisse nicht ändert. Das heißt, die Durchführungen müssen voneinander **unabhängig** sein. Das wohl wichtigste Kriterium ist aber, dass (nur) **genau zwei einander ausschließende Ereignisse** eintreten können.

Dies scheint eine strenge Einschränkung zu sein, ist jedoch bei einer Vielzahl von Ereignissen der Fall. Beim Wurf einer Münze unterscheidet man zum Beispiel meist zwischen Kopf und Zahl, bei Losen zwischen Gewinn und Niete. Selbst Experimente mit mehreren möglichen Ergebnissen kann man geeignet einteilen: Bei einem Würfel kann man so unter anderem nur die Ereignisse „gerade“ und „ungerade“ oder „größer 3“ und „kleiner gleich 3“ betrachten.

⁵ Jakob BERNOULLI, Basel 1654 - 1705; seit 1687 Professor der Mathematik in Basel; trug u.a. entscheidend zur Einführung der Infinitesimalrechnung bei; in seiner *Ars conjectandi* förderte er die Wahrscheinlichkeitsrechnung (Bernoullische Zahlen und Gesetz der großen Zahlen)

Herleitung: Die Wahrscheinlichkeit für das Ereignis E sei p, für das Gegenereignis E' sei sie q = 1 - p. Mit n wird die Anzahl der Durchführungen, mit k die Zahl der günstig ausgegangenen Bernoulli-Experimente bezeichnet.

Die Wahrscheinlichkeit, dass k mal das Ereignis E (und damit n-k mal E') in einer bestimmten Reihenfolge eintritt, ist somit $p^k \cdot q^{n-k}$. Es gibt aber $\binom{n}{k}$ Möglichkeiten um k mal E und n-k mal E' anzuordnen.

Deshalb gilt insgesamt: $P(X = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$.

⇒ Die Wahrscheinlichkeit, dass bei n Durchführungen eines Bernoulli-Experiments das Ereignis E mit der Wahrscheinlichkeit $P(E) = p$ genau k-mal eintritt, ist mit der **Wahrscheinlichkeitsfunktion der Binomialverteilung** auszurechnen:

$$B(n;p;k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad \text{mit } n \in \mathbb{N}, k \in \{0, 1, 2, \dots, n\} \text{ und } q = 1 - p.$$

Nach der allgemeinen Definition von Verteilungsfunktionen lautet die **(kumulative) Binomialverteilungsfunktion** (wobei $n \in \mathbb{N}$ und $p \in [0; 1]$):

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ \sum_{0 \leq k \leq x} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} & \text{für } 0 \leq x < n \\ 1 & \text{für } x \geq n \end{cases}$$

Die Eigenschaft jeder Verteilungsfunktion, dass $F(n) = 1$ sein muss, ist erfüllt: Nach dem binomischen Lehrsatz gilt nämlich:

$$F(n) = \sum_{k=0}^n \binom{n}{k} p^k \cdot q^{n-k} = (p + q)^n \quad (\text{daher auch der Name}). \text{ Aus } q = 1 - p \text{ folgt } p + q = 1 \text{ und damit } F(n) = 1^n = 1.$$

Eine Bernoulli-Kette der Länge n lässt sich durch die Summe von n sogenannten zweipunktverteilten Zufallsgrößen X_i darstellen. Diese können nur die zwei Werte 0 und 1 mit den Wahrscheinlichkeiten p und q annehmen. Sie haben den Erwartungswert $E(X_i) = 1 \cdot p + 0 \cdot q = p$, und die Varianz $V(X_i) = q^2 \cdot p + p^2 \cdot q = p \cdot q \cdot (q + p) = p \cdot q$. Mit den Formeln für die Summe aus mehreren unabhängigen Zufallsgrößen lässt sich berechnen:

$$E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = n \cdot E(X_i) = n \cdot p;$$

$$V(X) = V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n) = n \cdot V(X_i) = n \cdot p \cdot q;$$

Beispiel: Angenommen, dass bei **Flugzeugen die Motoren** unabhängig voneinander mit der Wahrscheinlichkeit $q = 1 - p$ **ausfallen**. So ein Flugzeug kann sein Ziel dann noch erreichen, wenn wenigstens die Hälfte seiner Motoren noch funktionieren. Welche Maschine sollte man vorziehen, einmotorig, mit zwei oder mit drei Motoren? Hängt diese Entscheidung von q ab?

$X =$ Anzahl der funktionierenden Motoren bei einem Flug

- **1-motorig:** $P(\text{„erfolgreicher Flug“}) = P(X = 1) = \underline{p}$

- **2-motorig:** $P(\text{„erfolgreicher Flug“}) = P(X = 1) = P(X = 1) + P(X = 2) =$
 $= 1 - P(X = 0) = 1 - \binom{2}{0} \cdot p^0 q^2 = \underline{1 - q^2}$

- **3-motorig:** $P(\text{„erfolgreicher Flug“}) = P(X = 2) = P(X = 2) + P(X = 3) =$
 $= \binom{3}{2} p^2 q^1 + \binom{3}{3} p^3 q^0 = 3p^2 q + p^3 = 3p^2(1 - p) + p^3 = 3p^2 - 3p^3 + p^3 = \underline{3p^2 - 2p^3}$

Zum *Vergleich*, welche Art Flugzeug nun wann die beste ist, prüft man:

- $P(\text{„Flugzeug mit 1 Motor kommt an“}) \leftarrow P(\text{„Flugzeug mit 2 Motoren kommt an“})?$

$$1 - q^2 \leftarrow 1 - q; \quad q^2 \leftarrow q \quad \Rightarrow \quad \text{für } 0 < q < 1 \text{ gilt } 1 - q^2 > p;$$

Bei $q = 0$ bzw. $q = 1$ ist es also egal, welches Flugzeug man verwendet. In allen anderen Fällen empfiehlt es sich aber die zweimotorige Maschine zu wählen.

- $P(\text{„Flugzeug mit 2 Motoren kommt an“}) \leftarrow P(\text{„Flugzeug mit 3 Motoren kommt an“})?$

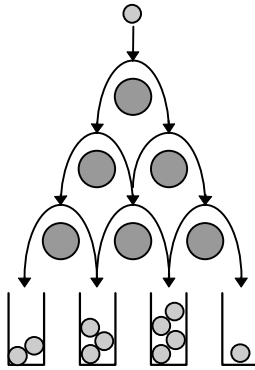
$$1 - q^2 \leftarrow 3p^2 - 2p^3; \quad \text{Linke Seite: } 1 - (1 - p)^2 = 1 - (1 - 2p + p^2) = 2p - p^2$$

$$2p^3 - 3p^2 + 2p - p^2 \leftarrow 0; \quad 2p^3 - 4p^2 + 2p \leftarrow 0; \quad p^2 - 2p + 1 \leftarrow 0; \quad (p - 1)^2 \leftarrow 0;$$

Ignoriert man die Trivialfälle $p = 0$ und $p = 1$ (bei denen es keine Rolle spielt, welche Flugzeugkonfiguration man verwendet, da man immer abstürzt bzw. ankommt), so erkennt man, dass es keine p gibt, für das die dreimotorige sicherer als der Flieger mit nur zwei Motoren ist.

Muss man sich also zwischen Flugzeugen mit einem, zwei oder drei Motoren entscheiden, so sollte man hier das doppelt bestückte Modell wählen.

Wie die Rechnung zeigt, ist eine Maschine mit **vier** Motoren nur dann sicherer als eine Zweimotorige, wenn die Ausfallwahrscheinlichkeit q der einzelnen Motoren unter $1/3$ liegt.



Besonders anschaulich lässt sich die Verteilung, die auf Bernoulli-Experimenten beruht, mit einem sogenannten **Galton-Brett**⁶ zeigen:

Hier treffen **Kugeln** auf gleichmäßig angeordnete Hindernisse (zum Beispiel halb eingeschlagene **Nägel**). Der Raum unterhalb dieser Anordnung ist in gleich breite Kammern aufgeteilt, in die die Kugeln fallen. In jeder Nagelebene ist die Wahrscheinlichkeit, dass die Kugel nach rechts (oder links) hinunterfällt 0,5 (50 Prozent). Deshalb sind die Kugeln in den Auffanggefäßen im Idealfall binomial verteilt (mit den Parametern $p = 0,5$ und $n = \text{Anzahl der Nägel}$).

Auch Verteilungen mit Trefferwahrscheinlichkeiten ungleich 0,5 lassen sich auf diese Art und Weise realisieren: Man ordnet mehrere **Wasserbehälter** untereinander an. Jeder dieser Behälter besitzt **zwei Abläufe**, einen nach rechts und einen nach links. Das Verhältnis der Querschnitte dieser beiden Rohre bestimmt die „Wahrscheinlichkeit“, mit der das Wasser in die jeweiligen Richtungen abfließt. Ist der linke Abfluss zum Beispiel doppelt so groß wie der rechte, so fließt das Wasser mit einer Wahrscheinlichkeit von $\frac{2}{3}$ in den nächsten Behälter nach links (beziehungsweise jedes Atom bewegt sich mit dieser Wahrscheinlichkeit in diese Richtung).

⁶ Sir Francis GALTON, England 1822 - 1911; er war tätig im Bereich der Geographie, Meteorologie, Vererbungslehre und erarbeitete (um sein angesammeltes statistisches Material zu verarbeiten) die Korrelationsrechnung.

II. 2. Hypergeometrische Verteilung

Obwohl im stochastischen Umgang mit Wahrscheinlichkeiten die Binomialverteilung am häufigsten angewandt wird, sind die meisten Zufallsexperimente hypergeometrisch verteilt. Die hypergeometrische Verteilung gibt nämlich die **Wahrscheinlichkeit für Versuche ohne Zurücklegen** wieder. Sie ist daher unter anderem für alle **Stichproben** geeignet, bei denen durch Untersuchung einiger Teile **auf den Gesamtanteil** von Merkmalen in einer größeren Menge **geschlossen** werden soll. Hierbei werden die untersuchten Teile natürlich **gleichzeitig gezogen** beziehungsweise **nicht** mehr **zurückgegeben**. Trotzdem wird häufig die Binomialverteilung zur Berechnung angewandt, da sie – wie in Punkt III.1. gezeigt wird – für größere Mengen eine gute Näherung darstellt.

Herleitung: Es soll die Wahrscheinlichkeit ermittelt werden, dass bei einer Stichprobe genau k Stücke defekt sind (beziehungsweise allgemein ein besonderes Merkmal besitzen), wenn n -mal (ohne Zurücklegen) aus einer Menge N mit insgesamt K defekten Stücken gezogen wird.

Zuerst gibt es ‚ K über k ‘ Möglichkeiten, aus K defekten Teilen k herauszugreifen und die Anzahl der Kombinationen der $n - k$ herausgegriffenen, nicht defekten Stücke innerhalb der Restmenge von $N - K$ Teilen errechnet sich durch ‚ $(N - K)$ über $(n - k)$ ‘. Insgesamt ergeben sich also bis jetzt $\binom{K}{k} \cdot \binom{N - K}{n - k}$ Möglichkeiten. Stichproben vom Umfang n aus einer Menge von N Elementen zu ziehen sind auf ‚ N über n ‘ Arten möglich. Daraus ergibt sich nun die Wahrscheinlichkeit von $\binom{K}{k} \cdot \binom{N - K}{n - k} / \binom{N}{n}$.

Definition: Die Wahrscheinlichkeit, dass sich in einer Stichprobe vom Umfang n aus einer Menge mit N Elementen, von denen K eine bestimmte Eigenschaft haben, genau k Elemente mit dieser Eigenschaft befinden ist **hypergeometrisch verteilt**. Es gilt:

$$P(X = k) = H(N; K; n; k) = \frac{\binom{K}{k} \cdot \binom{N - K}{n - k}}{\binom{N}{n}}$$

Die Berechnung der **Momente** der hypergeometrischen Verteilung ist nicht so einfach wie bei der Binomialverteilung, da hier die einzelnen Versuche **nicht unabhängig** sind. Man muss also berechnen:

$$E(X) = \sum_{k=0}^{\min(n;K)} k \cdot H(N;K;n;k) = \sum_{k=0}^{\min(n;K)} k \cdot \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}} = \sum_{k=1}^{\min(n;K)} k \cdot \frac{\frac{K}{k} \cdot \binom{K-1}{k-1} \cdot \binom{N-K}{n-k}}{\frac{N}{n} \cdot \binom{N-1}{n-1}} =$$

$$\left[\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N}{n} \cdot \frac{(N-1)!}{(n-1)![(N-1)-(n-1)]!} = \frac{N}{n} \cdot \binom{N-1}{n-1}; \quad \text{analog: } \binom{K}{k} = \frac{K}{k} \cdot \binom{K-1}{k-1} \right]$$

$$= \sum_{m=k-1}^{\min(n-1;K-1)} n \cdot \frac{K}{N} \cdot \frac{\binom{K-1}{m} \cdot \binom{(N-1)-(K-1)}{n-1-m}}{\binom{N-1}{n-1}} = n \cdot \frac{K}{N} \cdot \sum_{m=0}^{\min(n-1;K-1)} H(N-1;K-1;n-1;m) =$$

$$= n \cdot \frac{K}{N} \cdot \underbrace{\sum_{m=0}^{\min(n-1;K-1)} F_H(m)}_{=1} = \underline{\underline{n \cdot \frac{K}{N}}}$$

Der Erwartungswert der hypergeometrischen Verteilung ist daher: $E(X) = n \cdot \frac{K}{N} = n \cdot p$

Die Varianz auszurechnen erfordert noch etwas mehr Aufwand. Man verwendet als Ansatz $V(X) = E(X^2) - E(X)^2$ und erhält schließlich: $V(X) = npq \cdot$ ‚Korrekturglied‘

Die Varianz der hypergeometrischen Verteilung ist: $V(X) = n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N-n}{N-1}$

Die hypergeometrische Verteilung ist zu verwenden, wenn man mehrere Elemente gleichzeitig (= ohne Zurücklegen) zieht. So zum Beispiel bei vielen **Kartenspielen**:

Beispiel: Es lassen sich auf diese Weise beispielsweise schnell folgende Wahrscheinlichkeiten beim **Schafkopf**[en] ermitteln:

$$\begin{aligned} P(\text{„alle 4 Ober und 4 Unter in einer Hand“}) &= \frac{\binom{8}{8} \cdot \binom{24}{0}}{\binom{32}{8}} = \frac{1}{\binom{32}{8}} = \underline{\underline{9,5 \cdot 10^{-8}}} \\ &= P(\text{„8 bestimmte Karten in der Hand“}) = \end{aligned}$$

$$\Rightarrow n = 8; N = 32; K = 8; k = 8;$$

Die Chancen stehen also 1 : 10518300, also **1 : 10⁷**;

$$\begin{aligned} P(\text{„alle 4 Ober“}) = P(\text{„4 bestimmte Karten“}) &= \frac{\binom{4}{4} \cdot \binom{28}{4}}{\binom{32}{8}} = \frac{20475}{\binom{32}{8}} = \underline{\underline{1,9 \cdot 10^{-3}}} \end{aligned}$$

$$\Rightarrow n = 8; N = 32; K = 4; k = 4;$$

Die Chancen stehen also **1 : 514**;

$$P(\text{„6 oder mehr Trümpfe für ein beliebiges Solo“}) = P(\text{„6, 7 oder 8 aus 14 möglichen“}) =$$

$$\Rightarrow n = 8; N = 32; K = 14; k \geq 6; \quad = P(6) + P(7) + P(8) =$$

$$= \left[\binom{14}{6} \cdot \binom{18}{2} + \binom{14}{7} \cdot \binom{18}{1} + \binom{14}{8} \cdot \binom{18}{0} \right] / \binom{32}{8} = \frac{459459 + 61776 + 3003}{10518300} = \underline{\underline{0,050}}$$

Die Chancen stehen somit **1 : 20**;

II. 3. Negativ-Binomialverteilung

Der Negativ-Binomialverteilung liegt zwar wie bei der Binomialverteilung ein **Bernoulli-Experiment** zu Grunde, aber mit der Einschränkung, dass (nur) so oft **wiederholt** wird, **bis genau r-mal** ‘Erfolg’ eingetreten ist. Die Zufallsgröße X bezeichnet dann die Anzahl der benötigten Versuche ($r = 1$ entspricht also zum Beispiel der Fragestellung: ‘Wie oft muss man etwas versuchen, bis es [einmal] klappt?’).

Da r Erfolge eintreten müssen, kann X die Werte $r, r+1, r+2, \dots$ annehmen. Die Versuchsreihe kann theoretisch auch endlos fortgesetzt werden müssen.

Um zur Wahrscheinlichkeitsfunktion zu kommen, überlegt man:

Der letzte Versuch muss erfolgreich ausgegangen sein. Er darf daher nicht berücksichtigt werden.

Bei k Versuchen gibt es genau $\binom{k-1}{r-1}$ Möglichkeiten, die restlichen $r-1$ nötigen Ausführungen zu verteilen.

Wie bei der Bernoulli-Kette gezeigt wurde, ist die Wahrscheinlichkeit, dass so eine Versuchsreihe mit genau r Erfolgen und damit $k-r$ Fehlschlägen auftritt, gleich $p^r \cdot q^{k-r}$.

Es ergibt sich bis jetzt also insgesamt: $P(X = k) = \binom{k-1}{r-1} \cdot p^r \cdot q^{k-r}$ wobei $k \geq r$.

Interessant ist nun hauptsächlich die Anzahl der Versuche, die *nicht erfolgreich* ausgegangen sind. Da aber immer mindestens r Versuche durchgeführt werden müssen, berechnet man zweckmäßigerweise die Verteilung der Zufallsgröße $Y = X - r$. Für die Anzahl der Versuche muss jetzt ein anderer Buchstabe verwendet werden (hier m statt k).

Das bedeutet: $P(Y = m) = P(X - r = m) = P(X = m+r) = \binom{m+r-1}{r-1} \cdot p^r \cdot q^{m+r-r}$.

Da $\binom{r+m-1}{r-1} = \frac{(r+m-1)!}{(r-1)!(r+m-1-r+1)!} = \frac{(r+m-1)!}{(r-1)! \cdot m!} = \binom{r+m-1}{m}$ ist, gilt:

Definition: Eine **negativ-binomialverteilte Zufallsgröße** liegt dann vor, wenn für die Wahrscheinlichkeitsfunktion gilt: $P(X = m) = \binom{r+m-1}{m} \cdot p^r \cdot q^m$

Diese Schreibweise lässt sich formal noch umformulieren zu⁷:

$$P(Y = m) = \binom{-r}{m} \cdot p^r \cdot (-q)^m. \text{ Daher auch der Name } \textit{negativ-binomial}.$$

Die Negativ-Binomialverteilung gibt also die Wahrscheinlichkeit wieder, dass vor dem r-ten Erfolg genau m Misserfolge liegen.

Bei der Herleitung des **Erwartungswerts** und der **Varianz** bedient man sich desselben Tricks wie schon bei der Binomialverteilung: Man teilt die Zufallsgröße Y auf und berechnet:

$$E(Y) = E(X - r) = E(X) - E(r) = E(X_1) + E(X_2) + \dots + E(X_r) - E(r) = r \cdot E(X_i) - r$$

Die einzelnen X_i sind geometrisch verteilt, das heißt es gilt: $E(X_i) = 1/p$.

$$E(Y) = r \cdot \frac{1}{p} - r = r \cdot \left(\frac{1}{p} - 1 \right) = r \cdot \left(\frac{1-p}{p} \right) = \underline{\underline{r \cdot \frac{q}{p}}}$$

Für die Varianz folgt entsprechend: $V(Y) = V(X_1) + \dots + V(X_r) - V(r) = r \cdot V(X_i) = r \cdot \frac{q}{p^2}$

$$E(Y) = r \cdot \frac{q}{p} \qquad V(Y) = r \cdot \frac{q}{p^2}$$

Nun soll diese Wahrscheinlichkeitsberechnung mit Hilfe der Negativ-Binomialverteilung noch auf ein kurzes Beispiel angewandt werden:

Beispiel: Ein Schüler macht seine Hausaufgaben mit einer Wahrscheinlichkeit von 0,4. Wie groß ist nun die Wahrscheinlichkeit, dass er am Ende einer Woche (also am fünften Tag) zum dritten Mal *vorbereitet* in die Schule geht?

Y = Anzahl der Tage, an denen er keine Aufgaben bearbeitet hat

Die Zufallsgröße X ist dann negativ-binomial verteilt und es gilt: $P(\text{„Am 5. Tag zum 3. Mal vorbereitet“}) = P(\text{„Zweimal nicht vorbereitet“}) = P(Y = 2)$;

Die Parameter sind: **m = 2; p = 0,4; q = 1 - p = 0,6; r = 3;**

$$P(Y = 2) = \binom{3+2-1}{2} \cdot 0,4^3 \cdot 0,6^2 = 0,13824 = \underline{\underline{13,8\%}}$$

⁷ $\binom{r+m-1}{m} = \frac{(r+m-1)!}{m!(r-1)!} = \frac{(r+m-1) \cdot \dots \cdot (r+1) \cdot r}{m!} = (-1)^m \cdot \frac{\overbrace{(-r-m+1) \cdot (-r-m+2) \cdot \dots \cdot (-r-1) \cdot (-r)}^{m \text{ Faktoren}}}{m!} = (-1)^m \cdot \frac{(-r)!}{m! \cdot (-r-m)!} = (-1)^m \cdot \binom{-r}{m}$

Das $(-1)^m$ zieht man in q^m hinein und es ergibt sich schließlich: $P(Y = m) = \binom{-r}{m} \cdot p^r \cdot (-q)^m$.

II. 4. Geometrische Verteilung

Ein **Bernoulli-Experiment** (bei dem also nur Erfolg oder Misserfolg möglich ist), wird solange durchgeführt, **bis zum ersten Mal „Erfolg“** auftritt. Ist dies beim k -ten Versuch der Fall, so müssen die ersten $k - 1$ Versuche Misserfolge gewesen sein:

$$\underbrace{M M M \dots M}_{(k-1)\text{-mal}} \quad \underbrace{E}_{k\text{-ter Versuch}}$$

Die Zufallsgröße X bezeichnet die Anzahl der benötigten Versuche.

Kürzt man noch ab: $P(\text{„Erfolg“}) =: p$ und $P(\text{„Misserfolg“}) = 1 - p =: q$, so gilt:

Eine Zufallsgröße nennt man **geometrisch verteilt**, wenn für ihre Wahrscheinlichkeitsfunktion gilt:

$$P(X = k) = q^{k-1} \cdot p \quad \text{für } k \in \mathbb{N}$$

Ihr Name rührt daher, dass die Wahrscheinlichkeiten $q^{k-1} \cdot p$ Glieder einer *geometrischen Reihe* sind.

Die geometrische Verteilung ist ein **Spezialfall der** bereits eingeführten **Negativ-Binomialverteilung**. Dort wurde die Wahrscheinlichkeit der Anzahl der Versuche ausgerechnet, die nötig sind, um genau r Erfolge gehabt zu haben: $P(X = k) = \binom{k-1}{r-1} \cdot p^r \cdot q^{k-r}$. Setzt man r nun gleich 1, so erhält man genau die geometrische Verteilung:

$$P(X = k) = \binom{k-1}{0} \cdot p^1 \cdot q^{k-1} = q^{k-1} \cdot p.$$

$$\text{Es gilt hierbei: } P(X \geq k) = \sum_{v=k}^{\infty} q^{v-1} \cdot p = p \cdot q^k \cdot \sum_{v=k}^{\infty} q^{v-1-k} = p \cdot q^k \cdot \sum_{v=0}^{\infty} q^{v-1} =$$

unendliche geometrische Reihe:

$$s_{\infty} = \frac{a_1}{1-q}; a_1 = q^{0-1} = q^{-1}; q = q$$

$$= p \cdot q^k \cdot \frac{q^{-1}}{1-q} = p \cdot q^k \cdot \frac{q^{-1}}{p} = \underline{\underline{q^{k-1}}}$$

Bei der geometrischen Verteilung gilt: $P(X \geq k) = q^{k-1}$.

Daraus lässt sich berechnen: $q^{m-1} = P(X \geq m) = P(X > m - 1) = 1 - P(X \leq m - 1)$

Setzt man $k = m - 1$, so erhält man: $P(X \leq k) = 1 - q^k$

Die Momente der geometrischen Verteilung lassen sich aus der Definition des **Erwartungswerts** und den Formeln zur *Reihenentwicklung* berechnen. Außerdem

differenziert man und verwendet die Beziehung $\frac{d}{dq}(q^k) = k \cdot q^{k-1}$:

$$E(X) = \sum_{k=1}^{\infty} k \cdot q^{k-1} \cdot p = p \cdot \sum_{k=1}^{\infty} \frac{d}{dq} q^k = p \cdot \frac{d}{dq} \left(\sum_{k=1}^{\infty} q^k \right) = p \cdot \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \cdot \frac{1 \cdot (1-q) - 1 \cdot (-1)}{(1-q)^2} =$$

$$= p \cdot \frac{p+q}{p^2} = p \cdot \frac{1}{p^2} = \frac{1}{p}$$

unendliche geometrische Reihe (FS S.51):

$$s_{\infty} = \frac{a_1}{1-q}; a_1 = q^1 = q; q = q;$$

Um zur **Varianz** zu gelangen, muss man zweimal ableiten und erhält schließlich nach ganz ähnlicher Rechnung:

$$V(X) = \frac{q}{p^2} \quad E(X) = \frac{1}{p}$$

Beispiel: Angenommen, ein **Lehrer** hätte sich zu Beginn seiner Laufzeit **20 verschiedene Schulaufgaben** zu einem Stoffbereich **vorbereitet**. Jetzt wählt er jedes Jahr eine davon zufällig aus und legt sie dann wieder zu seiner Sammlung zurück. Wie hoch ist die Wahrscheinlichkeit, dass er die erste Schulaufgabe, die er ausgewählt hat,

- nach genau 21 Jahren
- nach frühestens 21 Jahren
- innerhalb der nächsten 4 Jahre

noch einmal schreibt?

Die Parameter der Verteilung, die hier zugrunde liegt, sind: $p = 1/20$; $q = 19/20$;

$$a) P(X = 21) = q^{k-1} \cdot p = \left(\frac{19}{20}\right)^{20} \cdot \frac{1}{20} = \left(\frac{19^{20}}{20^{21}}\right) = 0,0179 = \underline{\underline{1,8\%}}$$

$$b) P(X \geq 21) = q^{k-1} = \left(\frac{19}{20}\right)^{20} = 0,35849 = \underline{\underline{35,8\%}}$$

$$c) P(X \leq 4) = P(X < 5) = 1 - P(X \geq 5) = 1 - \left(\frac{19}{20}\right)^4 = 0,18549 = \underline{\underline{18,5\%}}$$

$$\text{Oder: } P(X \leq 4) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = q^0 p + q^1 p + q^2 p + q^3 p =$$

$$= \left[1 + \frac{19}{20} + \left(\frac{19}{20}\right)^2 + \left(\frac{19}{20}\right)^3 \right] \cdot \frac{1}{20} = \underline{\underline{18,5\%}} \quad \checkmark$$

II. 5. Poisson-Verteilung

Da für große Zahlen die Fakultäten in der Formel für die binomial verteilten Wahrscheinlichkeiten sehr schnell Schwierigkeiten bereiteten, suchte man nach einer *geeigneten Näherung*. Daher wird die Poisson-Verteilung hier auch aus der Binomialverteilung **hergeleitet**:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k} = \frac{n!}{k!(n-k)!} \cdot p^k (1-p)^{n-k} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k!} \cdot \frac{p^k \cdot n^k}{n^k} \cdot \frac{(1-p)^n}{(1-p)^k}$$

Im letzten Schritt wurde mit n^k erweitert und aufgeteilt: $(1-p)^{n-k} = \frac{(1-p)^n}{(1-p)^k}$

$$P(X = k) = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{n^k \cdot (1-p)^k} \cdot \frac{(p \cdot n)^k}{k!} \cdot (1-p)^n = \frac{1 \cdot \left(1 - \frac{1}{n}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{n}\right)}{\left(1 - \frac{\mu}{n}\right)^k} \cdot \frac{\mu^k}{k!} \cdot \left(1 - \frac{\mu}{n}\right)^n$$

Zuerst wurden nur die einzelnen Faktoren sinnvoll vertauscht und $p^k \cdot n^k$ in $(p \cdot n)^k$ zusammengefasst. Statt $p \cdot n$ lässt sich auch μ schreiben (der Erwartungswert der Binomialverteilung ist ja gerade $n \cdot p$) und ebenso kann man p durch $\frac{\mu}{n}$ ersetzen.

Gleichzeitig wurde noch der große Bruch durch n geteilt, um eine Grenzwertbestimmung durchführen zu können.

Nun kann man daran gehen, den Grenzwert der Binomialverteilung für $n \rightarrow \infty$ zu bestimmen:

$$\text{Laut Formelsammlung Seite 55 gilt: } \lim_{v \rightarrow \infty} \left(1 + \frac{k}{v}\right)^v = e^k \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{n}\right)^n = e^{-\mu}$$

Alle Brüche mit n im Nenner streben gegen 0:

$$\lim_{n \rightarrow \infty} B(n; p; k) = \frac{1 \cdot 1 \cdot \dots \cdot 1}{(1-0)^k} \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} = \underline{\underline{e^{-\mu} \cdot \frac{\mu^k}{k!}}}$$

Definition: Man spricht von einer **Poisson-verteilten Zufallsgröße** mit Erwartungswert μ , wenn für die Wahrscheinlichkeitsfunktion gilt: $P(X = k) = e^{-\mu} \cdot \frac{\mu^k}{k!}$

Die Poisson-Verteilung wird hauptsächlich dann verwendet, wenn Wahrscheinlichkeiten von Ereignissen berechnet werden sollen, die genau k -mal innerhalb einer vorgegebenen Zeitspanne eintreffen. Also zum Beispiel bei *Unfällen*, *Selbstmord*, aber auch bei *Druckfehlern in einem Buch*, bei *Rosinen in einem Kuchen* oder *Sterne im Kosmos*.

Da für konstanten Erwartungswert $\mu = n \cdot p$ die Trefferwahrscheinlichkeit p gegen 0 streben muss, wenn n gegen unendlich geht, muss p also sehr klein sein, um eine gute Annäherung an die Binomialverteilung gewährleisten zu können (siehe auch Punkt III.2.).

Deshalb nennt man die Poisson-Verteilung häufig auch **Verteilung der seltenen Ereignisse**.

In diesem Zusammenhang werden in vielen Büchern über dieses Thema nun als Beispiel die Versuchsergebnisse von Geiger und Rutherford erwähnt. Die beiden Physiker beobachteten **Zerfallsvorgänge radioaktiver Präparate** und registrierten die Anzahl der Zerfälle in gewissen Zeitabschnitten. Die genauere Beschreibung der Experimente und ihrer Ergebnisse würde aber einen unverhältnismäßig großen Raum einnehmen, so dass hier nur auf diese Versuche hingewiesen werden soll⁸.

Als **Beispiel** für die Poisson-Verteilung soll hier vielmehr die Auswertung einer Statistik dienen: Am **Karls-Gymnasium Bad Reichenhall** befinden sich 878 Schüler⁹. Nachdem die **Geburtstage aller Gymnasiasten** erfasst und sortiert wurden, konnte gezählt werden, wieviele Tage es gibt, an denen genau k Schüler gleichzeitig Geburtstag haben. Die Auswertung dieser Tabellen ergab letztendlich folgende Verteilung:

| | | | | | | | | | | |
|------------------------------------|----|----|----|----|----|----|---|---|---|---|
| k Schüler gleichzeitig Geburtstag: | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Anzahl der jeweiligen Tage: | 47 | 79 | 90 | 60 | 42 | 29 | 5 | 9 | 3 | 1 |

Der **Mittelwert**, der sich aus dieser Verteilung ergibt, ist:

$$\bar{x} = \frac{0 \cdot 47 + 1 \cdot 79 + 2 \cdot 90 + 3 \cdot 60 + 4 \cdot 42 + 5 \cdot 29 + 6 \cdot 5 + 7 \cdot 9 + 8 \cdot 3 + 1 \cdot 9}{47 + 79 + 90 + 60 + 42 + 29 + 5 + 9 + 3} = \frac{878}{365} = \underline{\underline{2,41}}$$

Nun betrachtet man die Zufallsgröße $X = \text{Anzahl der Tage, an denen } k \text{ Schüler gleichzeitig Geburtstag haben}$ und setzt die Wahrscheinlichkeit p , dass jemand an einem bestimmten Tag Geburtstag hat gleich $1/365$. Man erhält somit einen **Erwartungswert** ($n = 878$) von $\mu = E(X)$

⁸ Eine ausführliche Beschreibung dieses Versuchs findet man unter anderem in: Barth · Bergold · Haller, Stochastik 2, Ehrenwirth

$= \frac{878}{365} = \underline{2,41}$. Damit lassen sich die einzelnen Wahrscheinlichkeiten exakt mit Hilfe der **Binomialverteilung** ausrechnen. Da diese Werte sehr schnell recht unhandlich werden, bietet sich die **Poisson-Verteilung** an. In der folgenden Tabelle sind die relativen Häufigkeiten der Tabellenauswertung, die Wahrscheinlichkeiten der Binomialverteilung und die Näherungswerte der Poisson-Näherung jeweils in Prozent wiedergegeben:

| k | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-------------------------|------|------|------|------|------|-----|-----|-----|-----|-----|
| Binomial-V: | 9,0 | 21,7 | 26,1 | 21,0 | 12,6 | 6,0 | 2,4 | 0,8 | 0,2 | 0,1 |
| Poisson-V: | 9,0 | 21,7 | 26,1 | 20,9 | 12,6 | 6,1 | 2,4 | 0,8 | 0,3 | 0,1 |
| rel. Häufigkeit: | 12,9 | 21,6 | 24,7 | 16,4 | 11,5 | 7,9 | 1,4 | 2,5 | 0,8 | 0,3 |

Man erkennt den erstaunlich guten Zusammenhang zwischen den theoretisch ermittelten und den empirisch gewonnenen Daten.

Obwohl n (also die Anzahl der Schüler) nicht gerade extrem groß ist (zum Beispiel gemessen an der Anzahl von Atomen in einem Präparat), hat die Poisson-Verteilung in Fällen wie diesen durchaus ihre Berechtigung.

Bei der Herleitung der Poisson-Verteilung wurde bereits die Beziehung $\mu = n \cdot p$ angewandt. Deshalb ist es naheliegend, dass dieses μ in der Poisson-Verteilung auch ihren **Erwartungswert** darstellt. Dies lässt sich schnell mit Hilfe der allgemeinen Definition des Erwartungswerts nachrechnen:

$$E(X) = \sum_{k=0}^{\infty} k \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} = \sum_{k=1}^{\infty} k \cdot \frac{\mu \cdot \mu^{k-1}}{k \cdot (k-1)!} \cdot e^{-\mu} = \mu \cdot \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} \cdot e^{-\mu} = \mu \cdot \underbrace{\sum_{m=0}^{\infty} \frac{\mu^m}{m!} e^{-\mu}}_{=F_{\mu}(\infty)=1} = \mu$$

Die **Varianz** berechnet man am einfachsten über den bekannten Ansatz: $V(X) = E(X^2) - E(X)^2$

Einfach ist $E(X)^2 = \mu^2$; Schwieriger ist $E(X^2)$ zu erhalten:

Man ermittelt zuerst den Wert für $E(X^2) - E(X) = E(X^2 - X) = E(X \cdot (X-1))$:

$$E(X \cdot (X-1)) = \sum_{k=0}^{\infty} k(k-1) \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} = \mu^2 \cdot \sum_{k=2}^{\infty} \frac{\mu^{k-2}}{(k-2)!} \cdot e^{-\mu} = \mu^2 \cdot \underbrace{\sum_{m=0}^{\infty} \frac{\mu^m}{m!} e^{-\mu}}_{=F_{\mu}(\infty)=1} = \mu^2$$

Es wurde also bis jetzt gezeigt: $E(X^2) - E(X) = \mu^2$; $\Rightarrow E(X^2) = \mu^2 + E(X) = \mu^2 + \mu$

Nun kann man schließlich die Varianz ausrechnen: $V(X) = E(X^2) - E(X)^2 = \mu^2 + \mu - \mu^2 = \underline{\underline{\mu}}$

Erwartungswert und Varianz stimmen hier überein: $E(X) = \mu$ $V(X) = \mu$

⁹ Die Zahlen wurden am 13. Januar 1998 aktualisiert. Eine differenziertere Tabelle findet man im Anhang.

II. 6. Normalverteilung

Betrachtet man die Graphen von (standardisierten) Binomialverteilungen für wachsendes n , so stellt man fest, dass sie (für beliebiges p) gegen eine gemeinsame Grenzfunktion zu streben scheinen. Diese Verteilung, an die sich die binomial verteilten Wahrscheinlichkeiten annähern, wurde von den französischen Mathematikern *de Moivre*¹⁰ und *Laplace*¹¹ gefunden. Da die exakte Herleitung sehr umfangreich und kompliziert ist, soll hier nur auf den Beweis der Approximation in Kapitel III.4. hingewiesen werden. Das Ergebnis ist im sogenannten **lokalen Grenzwertsatz von de Moivre – Laplace** festgehalten:

Für eine **binomial verteilte Zufallsgröße X** gilt näherungsweise:

$$P(X = k) = B(n; p; k) \approx \varphi_{\mu\sigma}(k) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{k-\mu}{\sigma} \right)^2}$$

mit $k \in \mathbb{N}_0$ und $\mu = n \cdot p$, $\sigma = \sqrt{n \cdot p \cdot (1-p)}$

Für eine bereits standardisierte binomialverteilte Zufallsgröße Z gilt einfacher:

$$P(Z = z) \approx \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2}$$

Der Graph der Funktion hat auch den Namen **Gaußsche Kurve**, nach dem deutschen Mathematiker C.F. Gauß, der sie bei der Suche nach einem Verteilungsgesetz für zufällige Beobachtungsfehler entdeckte. Aufgrund ihrer Form wird sie häufig auch **Glockenkurve** genannt.

Bei ihr handelt es sich nicht mehr um eine diskrete Funktion. Nach der Definition stetiger Verteilungen legt man fest:

Die kumulative Standardnormalverteilung lautet: $\Phi_{0,1}(x) = \Phi(x) = \int_{-\infty}^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$

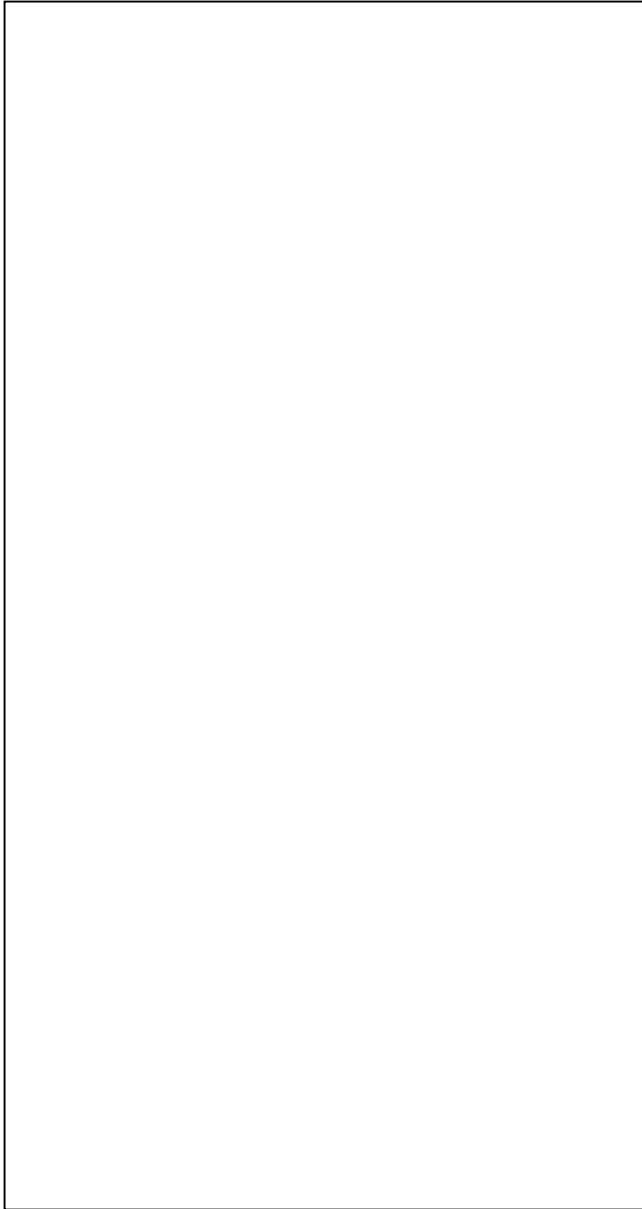
Daraus folgt der **Integralgrenzwertsatz von de Moivre und Laplace**: Für eine **nach $B(n; p)$ verteilte** Zufallsgröße X gilt:

$$P\left(a < \frac{X - np}{\sqrt{npq}} \leq b \right) \approx \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \cdot \int_a^b e^{-\frac{1}{2}u^2} du$$

¹⁰ Abraham de Moivre, Frankreich 1667 – 1754; Er arbeitete im Bereich der Differentialrechnung und gilt als Mitbegründer der Trigonometrie des Komplexen; der Begriff der erzeugenden Funktion und die analytischen Grundlagen der Wahrscheinlichkeitsrechnung gehen auf ihn zurück.

¹¹ Pierre Simon Marquis de Laplace, französischer Mathematiker, Physiker und Astronom, 1749 – 1827; neben Beiträgen zu physikalischen und astronomischen Bereichen entwickelte er systematisch die Wahrscheinlichkeitsrechnung.

Hier sind nun die **Graphen** der **Dichtefunktion** der standardisierten Normalverteilung und der **kumulativen Verteilungsfunktion** (ebenfalls mit $\mu = 0$ und $\sigma = 1$) gezeichnet:



Grundsätzlich gelten für die Gaußsche Funktion natürlich die Eigenschaften aller stetigen Verteilungen. Die Standardnormalverteilung erlaubt aber auch grundsätzliche Abschätzungen über Wahrscheinlichkeiten.

Man berechnet die Wahrscheinlichkeit von Abweichungen vom Mittelwert um weniger als $k \cdot \sigma$ und erhält zum Beispiel:

$$\begin{aligned} P(\mu - 3\sigma < X < \mu + 3\sigma) &= \\ &= \Phi\left(\frac{\mu + 3\sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - 3\sigma - \mu}{\sigma}\right) = \\ &= \Phi(3) - \Phi(-3) = 2\Phi(3) - 1 = \underline{\underline{0,99730}} \end{aligned}$$

Analog erhält man die eingezeichneten Werte für σ und 2σ :

$$P(|X - \mu| < \sigma) = \underline{\underline{0,68268}}$$

$$P(|X - \mu| < 2\sigma) = \underline{\underline{0,95450}}$$

$$P(|X - \mu| < 3\sigma) = \underline{\underline{0,99730}}$$

| Daten der Dichtefunktion: | Daten der Verteilungsfunktion: |
|--|--|
| Maximum: bei $(0; 1/\sqrt{2\pi} = 0,399)$ | Wendepunkt bei $(0; 0,5)$ |
| Wendepunkte: $(-\sigma = -1; 1/\sqrt{2\pi e} = 0,242)$ $(\sigma = 1; 1/\sqrt{2\pi e} = 0,242)$ | Grenzwerte: $\lim_{x \rightarrow -\infty} \Phi(x) = 0$; $\lim_{x \rightarrow +\infty} \Phi(x) = 1$ |

III. Zusammenhänge zwischen den Verteilungen (Grenzwertsätze)

III. 1. Näherung der Binomialverteilung an die hypergeometrische Verteilung

Obwohl die hypergeometrische Verteilung auf anderen Voraussetzungen beruht als die Binomialverteilung (die einzelnen Durchführungen sind **nicht unabhängig!**), scheinen die Experimente durchaus Ähnlichkeiten aufzuweisen (es werden zum Beispiel nur **zwei mögliche Ereignisse** betrachtet). Es ist also gerechtfertigt, sich näher mit der Verwandtschaft dieser Verteilungen zu beschäftigen.

Anhand eines **Beispiels** soll der Zusammenhang der Verteilungen in der praktischen Anwendung gezeigt werden:

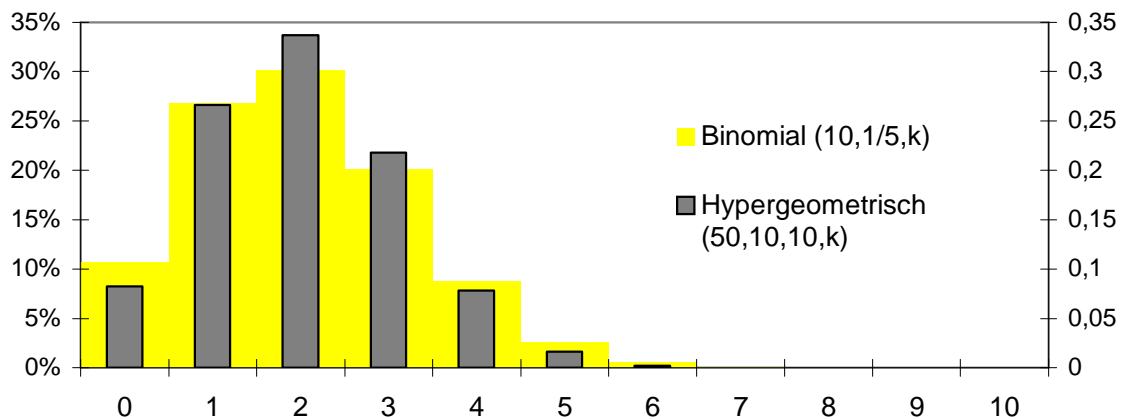
Zwei **Glückautomaten** H und B (**H**ypergeometrisch und **B**inomial) enthalten jeweils $N = 50$ Kugeln. Davon sind $K = 10$ golden und damit $N - K = 40$ schwarz. Durch Einwurf eines Einsatzes bekommt man 10 Kugeln. Die Größe des **Gewinns** richtet sich nach der **Anzahl der gezogenen goldenen Kugeln**.

Bei dem Automaten H fallen die 10 Kugeln auf einmal heraus (= **ohne Zurücklegen**).

Bei B bekommt man erst eine Kugel. Die nächste Kugel erhält man erst, wenn man die gezogene Kugel wieder eingefüllt hat (= **mit Zurücklegen**).

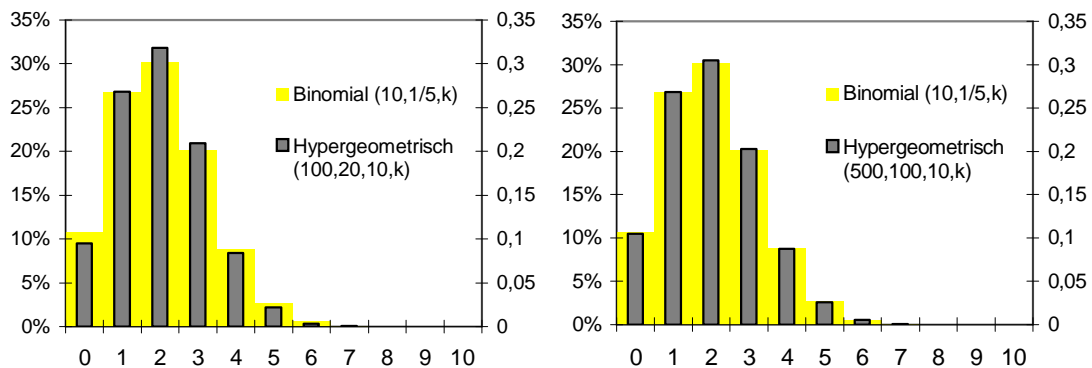
Die Wahrscheinlichkeiten sind **binomial** bzw. **hypergeometrisch** verteilt:

$$p_H(k) = H(N; K; n; k) = \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}} \quad \text{und} \quad p_B(k) = B(n; p; k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k};$$



Man erkennt, dass sich die beiden Verteilungen nicht sehr stark unterscheiden. Der Unterschied liegt ja lediglich darin, dass sich bei der hypergeometrischen Verteilung die Wahrscheinlichkeit nach jedem Zug ein wenig ändert. Sind jedoch sehr viele Kugeln vorhanden (d.h. **ist N sehr groß**) und man nimmt nur wenige heraus (d.h. **ist n klein gegen N**), so ändert sich die Wahrscheinlichkeit nach jedem Zug sehr wenig. Deshalb kann man die hypergeometrische Verteilung durch die Binomialverteilung ersetzen, wenn gilt, dass n gegenüber den Werten N , K und $N - K$ klein ist:

Erhöht man die Anzahl der vorhandenen Kugeln auf 100 beziehungsweise 500 (wobei immer noch $\frac{1}{5}$ der Kugeln golden sind) so erhält man folgende Diagramme, bei denen man die Annäherung noch deutlicher erkennen kann:



Neben dieser eher anschaulichen Erklärung für die Näherung kann man die Ähnlichkeit auch an den jeweiligen **Momenten** erkennen:

Während der **Erwartungswert** sogar **exakt übereinstimmt** ($E(X) = n \cdot p$), so unterscheidet sich die **Varianz** nur um den **Faktor** $\frac{(N - n)}{(N - 1)}$. Aber dieser geht ja für $N \gg n$ gegen 1 und damit auch die Varianz der hypergeometrischen Verteilung gegen $\sigma^2_{\text{Binomial}} = n \cdot p \cdot q$.

Die **Güte der Approximation** erkennt man an den folgenden zwei **Beispielen**:

- 1) Bei der üblichen **Lottoziehung „6 aus 49“** (ohne Besonderheiten wie Zusatzzahl oder ähnlichem) handelt es sich natürlich um eine „Ziehung ohne Zurücklegen“, da man keine Zahl zweimal ankreuzen kann. Deshalb muss man, um ein exaktes Ergebnis zu erhalten, die Formel für Wahrscheinlichkeiten der **hypergeometrischen Verteilung** anwenden.

Um die Wahrscheinlichkeit für **genau drei Treffer** zu erhalten, setzt man ein:

Da der **relative Fehler** nur noch **0,26 %** beträgt, erweist sich die Ersetzung von $H(N; K; n)$ durch $B(n; K/N)$ als durchaus brauchbar.

⇒ Man erkennt, dass sich Binomial- und hypergeometrische Verteilung um so ähnlicher werden, je größer N (also die Gesamtmenge) ist und die Anzahl der Ziehungen n relativ dazu klein ist!

Als **Faustregel** für eine im Allgemeinen ausreichend gute Annäherung gilt:

$$N \geq 2000 \text{ und } n/N \leq 0,1 \text{ (bzw. } n \leq 200);$$

außerdem sollte $p = K/N$ in der Nähe von 0,5 liegen und $k \ll n$ sein;

Man kann den **Grenzübergang** von der hypergeometrischen zur Binomialverteilung auch **allgemein zeigen**:

$$H\left(N; n; \frac{K}{N}\right) = \binom{K}{k} \cdot \binom{N-K}{n-k} / \binom{N}{n} = \frac{K!}{k!(K-k)!} \cdot \frac{(N-K)!}{(n-k)!(N-K-n+k)!} \cdot \frac{n!(N-n)!}{N!};$$

$n!$, $k!$ und $(n-k)!$ zieht man vor, schreibt die Fakultäten ausführlich und kürzt.

So wird zum Beispiel aus

$$\frac{K!}{(K-k)!} = \frac{K \cdot (K-1) \cdot (K-2) \cdot \dots \cdot (K-k+1) \cdot (K-k) \cdot (K-k-1) \cdot \dots \cdot 3 \cdot 2 \cdot 1}{(K-k) \cdot (K-k-1) \cdot \dots \cdot 3 \cdot 2 \cdot 1} = \underline{\underline{K \cdot (K-1) \cdot (K-2) \cdot \dots \cdot (K-k+1)}}$$

Somit erhält man: $H\left(N; n; \frac{K}{N}\right) =$

$$= \frac{n!}{k!(n-k)!} \cdot \frac{K \cdot (K-1) \cdot \dots \cdot (K-k+1) \cdot (N-K) \cdot (N-K-1) \cdot \dots \cdot (N-K-n+k+1)}{N \cdot (N-1) \cdot \dots \cdot (N-n+1)} =$$

Nach Teilung von Zähler und Nenner durch N gehen alle neuen Brüche der Form $1/N, 2/N, \dots, (n-k-1)/N$ usw. für $N \rightarrow \infty$ gegen 0 und können vernachlässigt werden:

$$= \binom{n}{k} \cdot \frac{\frac{K}{N} \cdot \left(\frac{K}{N} - \frac{1}{N}\right) \cdot \dots \cdot \left(\frac{K}{N} - \frac{k-1}{N}\right) \cdot \left(1 - \frac{K}{N}\right) \cdot \left(1 - \frac{K}{N} - \frac{1}{N}\right) \cdot \dots \cdot \left(1 - \frac{K}{N} - \frac{n-k-1}{N}\right)}{1 \cdot \left(1 - \frac{1}{N}\right) \cdot \dots \cdot \left(1 - \frac{n-1}{N}\right)} =$$

Schreibt man nun noch $p = K/N$ so ergibt sich:

$$= \binom{n}{k} \cdot \frac{\overbrace{p \cdot p \cdot \dots \cdot p}^{k\text{Faktoren}} \cdot \overbrace{(1-p) \cdot (1-p) \cdot \dots \cdot (1-p)}^{n-k\text{Faktoren}}}{1 \cdot 1 \cdot \dots \cdot 1} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = B(n;p) = B\left(n; \frac{K}{N}\right). \quad \blacksquare$$

Streng genommen wurde hier nur gezeigt, dass sich die Wahrscheinlichkeitsfunktionen der beiden Verteilungen aneinander nähern, aber es gilt ebenfalls, dass man die zugehörigen **Verteilungsfunktionen** für **große N** und **kleine n** untereinander annähernd ersetzen kann.

Es gilt also für eine **hypergeometrisch verteilte Zufallsgröße X** näherungsweise:

$$P(X = k) \approx B\left(n; n \cdot \frac{K}{N}; k\right) = \binom{n}{k} \cdot \left(\frac{K}{N}\right)^k \cdot \left(1 - \frac{K}{N}\right)^{n-k}$$

$$P(X \leq k) \approx F_{n; n \cdot \frac{K}{N}}(k) = \sum_{i=0}^k \binom{n}{i} \cdot p^i \cdot q^{n-i} = \sum_{i=0}^k \binom{n}{i} \cdot \left(\frac{K}{N}\right)^i \cdot \left(1 - \frac{K}{N}\right)^{n-i}$$

Allgemeine Fehlerabschätzungen sind schwierig, aber es lässt sich allgemein sagen, dass die Näherung meistens dann gut anwendbar ist, wenn $N \geq 2000$ und $\frac{n}{N} \leq 0,1$ gilt. Außerdem sollte die Trefferwahrscheinlichkeit **p** in der Nähe von **0,5** liegen. Da sich die Erfolgswahrscheinlichkeit kaum ändert, wenn nur wenig Kugeln gezogen werden sollen, trägt ein **kleines k** (gegenüber n) zu einer guten Approximation bei.

III. 2. Näherung der **Poisson-Verteilung** an die **Binomialverteilung**

In Kapitel II.5. wurde die **Poisson-Verteilung** bereits als **Näherung** für die oftmals sehr aufwendig zu berechnende **Binomialverteilung** eingeführt.

Um ein Kriterium aufstellen zu können, wann der **Approximationsfehler ausreichend klein** ist, werden die jeweiligen Verteilungen auf die folgenden **Beispiele** angewendet.

- 1) Ein typisches Experiment, auf das die Binomialverteilung zutrifft, ist der **mehrfache Wurf einer L-Münze**.

Geht man vom Idealfall aus, so ist die Wahrscheinlichkeit für eine der beiden Seiten gleich 0,5. Ist nun die Wahrscheinlichkeit zu errechnen, mit der bei zehn Würfeln genau k-mal ‘Kopf’ auftritt, so gilt exakt nach der **Binomialverteilung** für $k = 2$ und $k = 5$:

$$P_B(2) = \binom{10}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1 - \frac{1}{2}\right)^{10-2} = \binom{10}{2} \cdot \left(\frac{1}{2}\right)^{10} = \underline{\underline{4,39\%}}; \quad P_B(5) = \binom{10}{5} \cdot \left(\frac{1}{2}\right)^{10} = \underline{\underline{24,6\%}}$$

Die **Poisson-Näherung** ergibt (wobei $\mu = n \cdot p = 10 \cdot \frac{1}{2} = 5$ ist):

$$P_P(2) = e^{-5} \cdot \frac{5^2}{2!} = \underline{\underline{8,42\%}}; \quad P_P(5) = e^{-5} \cdot \frac{5^5}{5!} = \underline{\underline{17,6\%}}$$

Daraus ergibt sich ein **relativer Fehler** von **48 %** ($k = 2$) beziehungsweise **29 %** ($k = 5$).

⇒ Unter diesen Bedingungen stellt die Poisson-Verteilung also eine zu grobe Näherung dar. Da außerdem die kleinen Werte selbst ohne Taschenrechner keine Probleme bereiten, wird man die Poisson-Wahrscheinlichkeiten hier nicht anwenden.

- 2) Anders stellt sich das Problem im zweiten Beispiel dar: Ein **europäischer Roulette-Tisch** hat in der Regel 37 verschiedene Zahlen (0-36), auf die gesetzt werden kann. Die Wahrscheinlichkeit, dass eine bestimmte Nummer fällt, ist also $\frac{1}{37} \approx 2,70\%$. Lässt man die Kugel 111-mal rollen, so erwartet man $E(X) = n \cdot p = \frac{111}{37} = 3$ -mal im Durchschnitt jede Zahl. Wie hoch ist nun die Wahrscheinlichkeit, dass jeweils genau k-mal eine bestimmte Zahl fällt?

Berechnet wird jeweils der exakte Wert der **Binomialverteilung** und die Werte der **Poisson-Näherung**:

| k | 0 | 1 | 2 | 3 | 4 | 5 | 6 | < 7 | 20 |
|--------------------------------------|----------|----------|----------|----------|----------|----------|----------|---------------|-----------|
| $P_B(k;111;1/37)$: | 4,78 | 14,73 | 22,50 | 22,71 | 17,03 | 10,13 | 4,97 | 96,85 | 1,9E-09 |
| $P_P(k;3)$: | 4,98 | 14,94 | 22,40 | 22,40 | 16,80 | 10,08 | 5,04 | 96,65 | 7,1E-09 |
| Fehler: | 4,04 | 1,38 | 0,45 | 1,36 | 1,36 | 0,44 | 1,42 | 0,21 | 73,16 |

Der **relative Fehler** für sinnvolle Werte von k liegt also fast immer **unter 2 %**. Hier ist die Poisson-Approximation sinnvoll.

Besonders geeignet für die Anwendung der Poisson-Verteilung sind Vorgänge bei denen **Atome** im Spiel sind, **da ihre Anzahl extrem groß** ist. Ein gutes Beispiel wären die bereits bei der Herleitung der Poisson-Verteilung (Punkt II.5.) erwähnten **Experimente von Geiger und Rutherford**, da die Anzahl n bei Betrachtungen mit Atomen meist sehr groß ist und bestimmte Ereignisse nur selten eintreten. Aus Platzgründen muss aber auch hier auf Sekundärliteratur verwiesen werden (siehe Fußnote 8 auf Seite 34).

Dem erwähnten Zerfallsexperiment und den beiden Beispielen könnte man nun entnehmen, dass die Poisson-Approximation für **größere n** und **kleinere p** besser wird. Außerdem sollte sich **k nicht zu weit vom Erwartungswert entfernen**. Diese Vermutungen sollen jetzt durch eine allgemeine Berechnung des Fehlers, den man bei Anwendung der Poisson-Verteilung macht, bestätigt werden:

Es zeigt sich, dass es für die Vereinfachung des Terms günstig ist, nicht die Differenz, sondern den Quotienten von $B(k; n; p)$ und $P(k; \mu)$ zu bilden und ihn zu logarithmieren. Für die Erfolgswahrscheinlichkeit kann man $p = \mu/n$ schreiben, so dass der Fehler δ nur noch von n und k abhängt: [Die Berechnung ist auf der linken Seite durchgeführt, rechts befinden sich die zu jedem Schritt nötigen Erläuterungen.]

$$\delta = \ln \frac{B(k; n; \frac{\mu}{n})}{P(k; \mu)} = \ln \frac{\frac{n!}{k!(n-k)!} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k}}{e^{-\mu} \frac{\mu^k}{k!}} =$$

$$= \ln \left[\frac{n(n-1) \dots (n-k+1)}{k!} \cdot \frac{\mu^k}{n^k} \cdot \left(1 - \frac{\mu}{n}\right)^{n-k} \cdot e^\mu \cdot \frac{k!}{\mu^k} \right] =$$

$$= \ln \left[\frac{n}{n} \left(\frac{n-1}{n}\right) \dots \left(\frac{n-k+1}{n}\right) \cdot \left(1 - \frac{\mu}{n}\right)^{n-k} \cdot e^\mu \right] =$$

$$= \ln \left[\underbrace{\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)}_1 \right] + \underbrace{(n-k) \cdot \ln \left(1 - \frac{\mu}{n}\right)}_2 + \underbrace{\mu}_{\frac{\mu}{3}}$$

$$\frac{n!}{(n-k)!} = n(n-1) \dots (n-k+1)$$

$$\left(\frac{\mu}{n}\right)^k = \frac{\mu^k}{n^k}; \frac{1}{e^{-\mu} \frac{\mu^k}{k!}} = e^\mu \frac{k!}{\mu^k}$$

$k!, \mu^k$ gekürzt; $n(n-1) \dots (n-k+1)$ sind genau k Faktoren, jeder wird durch n geteilt $\Rightarrow n^k$ fällt weg;

$\ln(a \cdot b) = \ln(a) + \ln(b)$; $\ln(e^\mu) = \mu$;

$$\ln \left(1 - \frac{\mu}{n}\right)^{n-k} = (n-k) \cdot \ln \left(1 - \frac{\mu}{n}\right)$$

Der Übersicht halber werden diese drei Summanden nun kurzfristig einzeln behandelt:

(1) $= \ln \left(1 - \frac{1}{n}\right) + \ln \left(1 - \frac{2}{n}\right) + \dots + \ln \left(1 - \frac{k-1}{n}\right) = \sum_{j=1}^{k-1} \ln \left(1 - \frac{j}{n}\right)$. Für die weitere Vereinfachung

bietet sich eine Reihenentwicklung des Logarithmus an (Formelsammlung Seite 52):

$\ln(1 + y) = y - \frac{1}{2}y^2 + \frac{1}{3}y^3 \pm \dots$ für $|x| \leq 1$. Für das vorliegende Problem setzt man $y = -x$:

$\ln(1 - x) = -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 \pm \dots = -x \left(1 + \frac{1}{2}x + \frac{1}{3}x^2 + \dots\right)$.

Eine für diese Fehlerabschätzung ausreichend gute Näherung erhält man, wenn man die Reihe nach dem dritten Glied abbricht.

Zur Berechnung der obigen Summe nimmt man $x = j/n$ und erhält so:

(1) $= \sum_{j=1}^{k-1} \ln \left(1 - \frac{j}{n}\right) = \sum_{j=1}^{k-1} \left\{ -\frac{j}{n} \left[1 + \frac{1}{2} \frac{j}{n} + \frac{1}{3} \left(\frac{j}{n}\right)^2 \right] \right\} = - \sum_{j=1}^{k-1} \left\{ \frac{j}{n} \left[1 + \frac{1}{2} \frac{j}{n} + \frac{1}{3} \left(\frac{j}{n}\right)^2 \right] \right\}$

(2) $= (n-k) \cdot \ln \left(1 - \frac{\mu}{n}\right) = -(n-k) \frac{\mu}{n} \left[1 + \frac{1}{2} \frac{\mu}{n} + \frac{1}{3} \left(\frac{\mu}{n}\right)^2 \right]$ erhält man analog.

Der Ausdruck für den Fehler lautet mittlerweile also:

$$\delta = -\sum_{j=1}^{k-1} \left\{ \frac{j}{n} \left[1 + \frac{1}{2} \frac{j}{n} + \frac{1}{3} \left(\frac{j}{n} \right)^2 \right] \right\} - (n-k) \frac{\mu}{n} \left[1 + \frac{1}{2} \frac{\mu}{n} + \frac{1}{3} \left(\frac{\mu}{n} \right)^2 \right] + \mu.$$

Definitionsgemäß ($P(k;\mu) = \lim_{n \rightarrow \infty} B(k;p;n)$) kann man n als groß voraussetzen, wenn die Näherung überhaupt einen Sinn machen soll. Deshalb kann man die Terme mit n^2 und n^3 im Nenner vernachlässigen. So erhält man immer noch eine ziemlich gute Näherung für δ :

Zieht man in der ersten Summe j/n in die Klammer und sortiert die Terme, so ergibt sich:

$$\delta = -\sum_{j=1}^{k-1} \frac{j}{n} - \underbrace{\frac{1}{2} \sum_{j=1}^{k-1} \frac{j^2}{n^2} - \frac{1}{3} \sum_{j=1}^{k-1} \frac{j^3}{n^3}}_{\text{fällt weg}} - n \frac{\mu}{n} \left(1 + \frac{1}{2} \frac{\mu}{n} + \frac{1}{3} \frac{\mu^2}{n^2} \right) + k \frac{\mu}{n} \left(1 + \frac{1}{2} \frac{\mu}{n} + \frac{1}{3} \frac{\mu^2}{n^2} \right) + \mu =$$

$$= -\sum_{j=1}^{k-1} \frac{j}{n} - \mu \left(1 + \frac{1}{2} \frac{\mu}{n} \right) + k \frac{\mu}{n} + \mu =$$

$$= -\frac{k(k-1)}{2n} - \mu - \frac{\mu^2}{2n} + k \frac{\mu}{n} + \mu =$$

$$= \frac{1}{2n} [-k(k-1) - \mu^2 + 2k\mu] = \frac{1}{2n} (-k^2 + k - \mu^2 + 2k\mu) = \frac{1}{2n} [-(k^2 - 2k\mu + \mu^2) + k] =$$

$$= \frac{1}{2n} (k - (k - \mu)^2)$$

$$\text{arithmetische Reihe: } \sum_{j=1}^{k-1} \frac{j}{n} = s_n = \frac{n}{2} (a_1 + a_n);$$

$$\text{wobei } n = k-1; a_1 = 1/n \text{ und } a_n = (k-1)/n;$$

$$\Rightarrow s_n = \frac{k-1}{2} \left(\frac{1}{n} + \frac{k-1}{n} \right) = \frac{k-1}{2} \cdot \frac{k}{n} = \frac{k(k-1)}{2n}.$$

Es lässt sich nun zeigen (auf diese Ausführungen wurde hier aber verzichtet), dass der errechnete Fehler $\delta = \ln[B(k; p; n) / P(k; \mu)]$ bei kleinen Abweichungen sehr gut dem relativen Fehler der beiden Werte entspricht.

Wie an der folgenden Tabelle zu erkennen ist (es wurden die Werte des zweiten Beispiels (Roulette) übernommen), ist diese Fehlerabschätzung nur für kleine Abweichungen als Korrekturfaktor brauchbar. Sie genügt aber, um zum Beispiel schon im Voraus bestimmen zu können, ob im konkreten Fall die Poisson-Verteilung eine geeignete Approximation an die Binomialverteilung darstellt:

| k | 0 | 1 | 2 | 3 | 4 | 5 | 6 | < 7 | 20 |
|-----------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|------|---------------|
| Fehler: | 4,04 | 1,38 | 0,45 | 1,36 | 1,36 | 0,44 | 1,42 | 0,21 | 73,16 |
| Ln(P/B): | 4,13 | 1,39 | 0,45 | 1,37 | 1,37 | 0,44 | 1,43 | 0,21 | 131,51 |
| (k-(k-μ)²)/2n: | 4,05 | 1,35 | 0,45 | 1,35 | 1,35 | 0,45 | 1,35 | | 121,17 |

Aus der erhaltenen Formel lässt sich bereits folgern, dass der Fehler in der Tat um so geringer wird, je **kleiner** n im Nenner wird.

Da aber $p = \mu/n$ ist, so erkennt man sofort, dass die Poisson-Näherung gut anwendbar ist, wenn **p klein** ist (Verteilung der seltenen Ereignisse!).

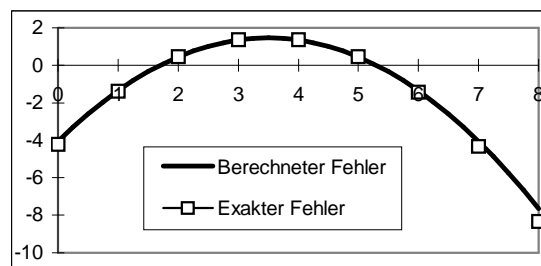
Setzt man $\delta = 0$, so erhält man nach kurzer Rechnung: $k_{1/2} = (\mu + 1/2) \pm \sqrt{\mu + 1/4}$.

Dies bedeutet, dass der Fehler dann besonders gering ist, wenn **k in der Nähe** von (nicht unbedingt genau auf) dem **Erwartungswert** liegt. Dies zeigt auch die folgende Grafik:

Sie stellt den exakten relativen (in Prozent) (ausgerechnet durch

$$100 \cdot B(k; p; n) / P(k; \mu)$$

und den durch die Formel angenäherten Fehler dar ($\mu = 3$):



[Das Maximum dieser Funktion liegt bei $\frac{1}{2n} \left(\mu + \frac{1}{4} \right)$. Berechnet man nun die Werte von k , für die der Betrag des relativen Fehlers unter diesem Wert liegt, so erhält man als Lösungen: $k_{1/2} = (\mu + 1/2) \pm \sqrt{2\mu + 1}$. Das bedeutet, dass der relative Fehler kleiner als $\frac{1}{2n} (\mu + 1/4)$ ist, wenn k von $\mu + 1/2$ um weniger als $\sqrt{2\mu + 1}$ abweicht.]

Zusammenfassend lässt sich daher schreiben:

Für eine **binomial verteilte Zufallsgröße X** gilt, da $\mu = n \cdot p$ ist:

$$P(X = k) = B\left(n; \frac{\mu}{n}; k\right) \approx P_{\mu}(k) = e^{-\mu} \cdot \frac{(\mu)^k}{k!} \quad \text{bzw.} \quad P(X \leq k) \approx F_{\mu}(k) = \sum_{i=0}^k e^{-\mu} \cdot \frac{(\mu)^i}{i!}$$

Aus den hier dargestellten Überlegungen lässt sich nun folgern: Die Approximation ist um so besser, je größer n in Relation zu μ und $|k - \mu|$ ist. Dies impliziert ein kleines $p = \mu/n$ und ein k , dass möglichst in der Nähe des Erwartungswerts sein soll.

Für eine angemessene Näherung sollte mindestens gelten: **p ≤ 0,1** und **n ≥ 30**.

Für **p ≤ 0,01** und **n ≥ 100** zeigen Tabellen, dass der maximale Fehler unter 2 ‰ liegt.

III. 3. **Näherung der Poisson-Verteilung an die hypergeometrische Verteilung**

Dass die Poisson-Approximation als Näherung der hypergeometrischen Verteilung verwendet werden kann, kann man bereits aus der Tatsache schließen, dass sich die **hypergeometrische** an die **Binomialverteilung** und sich diese wiederum an die **Poisson-Verteilung** annähert. Deshalb wird hier der Zusammenhang hauptsächlich auf die Brauchbarkeit der Ersetzung hin untersucht.

Da sich die Wahrscheinlichkeiten von hypergeometrisch beziehungsweise binomial verteilten Zufallsgrößen für große N nicht mehr wesentlich unterscheiden (siehe III.1.) wird man bei der Poisson-Näherung von vornherein nur dann gute Ergebnisse erwarten, wenn N **groß** und die Trefferwahrscheinlichkeit p **klein** ist.

Zur Überprüfung dieser Vermutung betrachten wir folgende **Beispiele**:

- 1) Verwendet man das Experiment des **Glücksautomaten** aus Kapitel III.1 (10 goldene Kugeln und 40 weitere; 10 werden gleichzeitig (d.h. nacheinander ohne Zurücklegen) gezogen), so erhält man folgende **hypergeometrische verteilten** Wahrscheinlichkeiten (die jeweiligen **Poisson-Näherungen** mit $\mu = n \cdot M/N = 2$ sind bereits mit angegeben):

| k | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------------------------|------|------|------|------|------|------|------|------|------|-------|-------|
| Hypergeometr. V: | 8,25 | 26,6 | 33,7 | 21,8 | 7,85 | 1,61 | 0,19 | 0,01 | 0,00 | 0,00 | 0,00 |
| Poisson-V: | 13,5 | 27,1 | 27,1 | 18,0 | 9,02 | 3,61 | 1,20 | 0,34 | 0,09 | 0,02 | 0,00 |
| Rel. Fehler: | 39,0 | 1,65 | 19,7 | 17,1 | 13,0 | 55,3 | 84,5 | 96,6 | 99,6 | 100,0 | 100,0 |

Es scheint hier die Anzahl der Kugeln zu gering zu sein, so dass die Näherung viel zu ungenau ist, um sie hier anwenden zu dürfen.

- 2) Bei einer **Produktion von Bauteilen** sind, aus technischen Gründen, erfahrungsgemäß im Mittel 1 % **fehlerhaft**. Eine Firma kauft nun 1000 Stück aus dieser Produktion und möchte die Angaben des Verkäufers prüfen. Da die Überprüfung der Bauteile kostenaufwendig ist, entnimmt der Kontrolleur der Lieferung zufällig 100 Stück (d.h. „Ziehen ohne Zurücklegen“) und will den Posten ablehnen, wenn mehr als die angegebenen 1%, also mehr als *ein* Stück nicht funktioniert.

Wie hoch ist die Wahrscheinlichkeit, dass er die Bauteile zurückschickt, obwohl der vom Hersteller angegebene Prozentsatz stimmt?

Exakt ist die Anzahl der defekten Stücke **hypergeometrisch verteilt**:

$$\begin{aligned} P(\text{„mehr als 1 defekt“}) &= 1 - P(\text{„keines defekt“}) - P(\text{„genau 1 defekt“}) = \\ &= 1 - H(0; 100; 10; 1000) - H(1; 100; 10; 1000) = \\ &= 1 - \binom{10}{0} \cdot \binom{990}{100} / \binom{1000}{100} - \binom{10}{1} \cdot \binom{990}{99} / \binom{1000}{100} = 1 - 0,347 - 0,390 = \underline{\underline{26,37\%}} \end{aligned}$$

Der Händler wird also im Schnitt in **26,4 %** aller Fälle die Lieferungen zurücksenden, obwohl sie eigentlich noch den Bedingungen entsprechen.

Da der Rechenaufwand hier sehr hoch ist, ist es zweckmäßig, eine einfacher zu berechnende Näherung wie die **Poisson-Verteilung** heranzuziehen:

Der Parameter μ der Verteilung ist ja der Erwartungswert. Er lässt sich leicht berechnen, da gilt: $\mu = n \cdot p = 100 \cdot 0,01 = 1$. Man erwartet also durchschnittlich ein defektes Teil in der Auswahl.

Nun wird analog zu oben die Wahrscheinlichkeit $P(\text{„mehr als 1 Teil ist defekt“})$ berechnet:

$$\begin{aligned} P(\text{„2 oder mehr Teile sind defekt“}) &= 1 - P(\text{„keines defekt“}) - P(\text{„genau eines defekt“}) = \\ &= 1 - e^{-1} \cdot \frac{1^0}{0!} - e^{-1} \cdot \frac{1^1}{1!} = 1 - \frac{1}{e} - \frac{1}{e} = 1 - \frac{2}{e} = \underline{\underline{26,42\%}}. \end{aligned}$$

Nach dieser wesentlich einfacheren Rechnung wird er ebenfalls etwa **26,4 %** der korrekten Lieferungen fälschlich abweisen.

Der **relative Fehler** beträgt lediglich **2,0 ‰**.

Bei dieser Parameterkonstellation (großes N , kleines $p = K/N$) stellt die Poisson-Verteilung also eine gute Näherung der hypergeometrischen Verteilung dar.

Für eine allgemeine Abschätzung des relativen Fehlers ist natürlich die Formel, die bei der Approximation Poisson- \rightarrow Binomialverteilung hergeleitet wurde, verwendbar. P_B und P_H kommen ja einander für größere Werte von N beliebig nahe.

Der Wert des relativen Fehlers liegt also ungefähr bei $\frac{1}{2n}(k - (k - \mu)^2)$

Es lässt sich somit für große N und kleine p nähern:

Für eine **hypergeometrisch verteilte Zufallsgrösse X** gilt näherungsweise, da $\mu = n \cdot \frac{K}{N}$ ist:

$$P(X = k) \approx P_{\mu}(k) = P_{n, \frac{K}{N}}(k) = e^{-n \cdot \frac{K}{N}} \cdot \frac{(n \cdot \frac{K}{N})^k}{k!}$$

beziehungsweise

$$P(X \leq k) \approx F_{\mu}(k) = F_{n, \frac{K}{N}}(k) = \sum_{i=0}^k e^{-n \cdot \frac{K}{N}} \cdot \frac{(n \cdot \frac{K}{N})^i}{i!}$$

Die Poisson-Verteilung ist ursprünglich als Näherung der Binomialverteilung für große n eingeführt worden. Daraus lässt sich folgern, dass eine gute Approximation grundsätzlich erst für **höhere Werte von n** zu erwarten ist.

Die beiden Verteilungen sind sich aber auch erst dann besonders ähnlich, wenn die **Trefferwahrscheinlichkeit p relativ klein** ist (die Poisson-Näherung ist ja die Verteilung der seltenen Ereignisse).

Deshalb sollte auch der Wert von k in einem Bereich liegen, der **wesentlich kleiner als n** ist. Da n groß und p klein sein muss, sollte man einen **mittelgroßen Erwartungswert** voraussetzen, will man eine brauchbare Näherung erhalten.

Die Tatsache, dass sich die hypergeometrische Verteilung und die Binomialverteilung erst für große N nicht mehr wesentlich unterscheiden, legt nahe, dass dieses N **eine große Zahl** sein sollte. Außerdem darf man nicht zu viele Versuche durchführen, da sich sonst die Trefferwahrscheinlichkeit zu stark ändert. Das bedeutet, dass n **wesentlich kleiner als N** sein muss.

Die hypergeometrische Verteilung lässt sich im Allgemeinen genau genug von der Poisson-Verteilung annähern, wenn in etwa gilt:

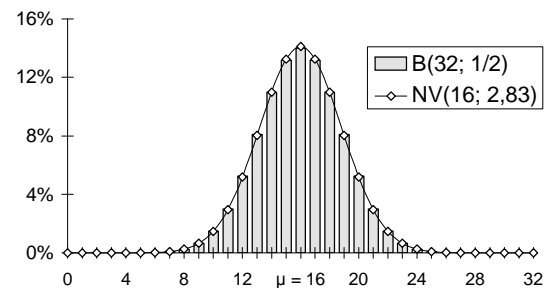
$$p = K/N \leq 0,1 \text{ und } n \leq 0,1 \cdot N$$

III. 4. Näherung der Normalverteilung an die Binomialverteilung

Die Bedeutung der bisher rein theoretisch definierten **Normalverteilung** ergibt sich aus ihrer Eigenschaft, dass sich (unter bestimmten, sehr wenig einschränkenden Bedingungen) **alle möglichen Verteilungen an sie annähern**. Dies ist im sogenannten **Zentralen Grenzwertsatz** enthalten.

Die Annäherung speziell von Binomialverteilungen an die Normalverteilung soll im Folgenden näher untersucht werden:

Zeichnet man den Graphen der Binomialverteilung für $p = q = \frac{1}{2}$ und $n = 32$, so erkennt man an dem entstehenden **Häufigkeitspolygon** eine große Ähnlichkeit der ebenfalls eingetragenen Dichtefunktion der Normalverteilung:



Für die vorliegenden Werte lassen sich in der Zeichnung bereits keine Abweichungen mehr zwischen den Säulen (das heißt der **Binomialverteilung** mit $n = 32$ und $p = \frac{1}{2}$) und der durchgezogenen Linie (Werte der **Normalverteilung** mit $\mu = 32 \cdot \frac{1}{2} = 16$ und $\sigma = \sqrt{32 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 2\sqrt{2}$) erkennen.

Nun gilt es also allgemein zu untersuchen, ob jede beliebige Binomialverteilung mit wachsender Stichproben- / Versuchsanzahl n gegen eine Normalverteilung strebt.

Da bei größer werdendem n auch μ und σ gegen ∞ streben (das heißt, der **Verlauf wird immer flacher** und der Graph verschiebt sich im Koordinatensystem immer weiter **nach rechts**), während die **Standardnormalverteilung stets symmetrisch** zur Ordinatenachse bleibt, ist ein Grenzübergang für $n \rightarrow \infty$ sicher nicht direkt möglich. Damit man nicht mit der allgemeinen Normalverteilung arbeiten muss, sollte die Binomialverteilung also

standardisiert vorliegen. Um den *Erwartungswert* 0 und die *Standardabweichung* 1 zu erhalten, substituiert man und führt die neue **standardisierte Zufallsgröße** $Z = \frac{X - \mu}{\sigma}$ ein.

Es soll also jetzt diese Behauptung bewiesen werden: $\lim_{n \rightarrow \infty} \varphi_n(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}z^2}$

Dies bezeichnet man als den **Lokalen Grenzwertsatz von De Moivre und LaPlace**.

Da sich der Beweis über mehrere Schritte hinzieht, sind im gelb hinterlegten Bereich immer die Ergebnisse der links durchgeführten Berechnungen festgehalten. Somit kann man sich schnell einen Überblick über die Beweisführung verschaffen, ohne die einzelnen Rechnungen im Detail nachvollziehen zu müssen.

Hier und auf den nächsten Seiten steht eigentlich der Beweis, der in der Datei „III4BEW.DOC“ enthalten ist!

Da sich die Wahrscheinlichkeitsfunktion der Binomialverteilung und die Dichtefunktion der allgemeinen Normalverteilung aneinander annähern, besteht dieser Zusammenhang auch für die jeweiligen kumulativen Verteilungsfunktionen (der Wertebereich liegt ja bei beiden zwischen 0 und 1). Wie schon in Kapitel II.6. gezeigt kann man also binomialverteilte Wahrscheinlichkeiten folgendermaßen approximieren:

Ist X eine binomial verteilte Zufallsgröße, so gilt näherungsweise:

$$P(X \leq k) \approx \Phi\left(\frac{k - np + \frac{1}{2}}{\sqrt{npq}}\right)$$

$$B(n; p; k) \approx \frac{1}{\sigma} \cdot \Phi\left(\frac{k - \mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi \cdot npq}} \cdot e^{-\frac{(k - np)^2}{2 \cdot npq}}$$

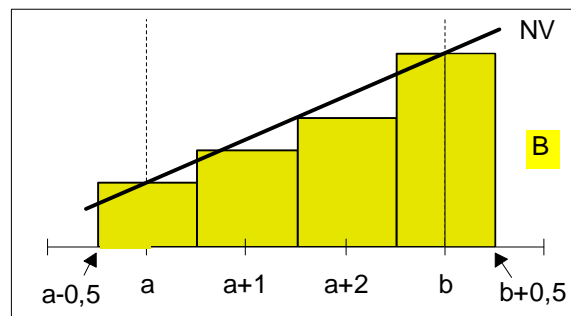
oder: $B(n; p; k) \approx \Phi\left(\frac{k - np + 0,5}{\sqrt{npq}}\right) - \Phi\left(\frac{k - np - 0,5}{\sqrt{npq}}\right)$

Auf die Bedeutung der Summanden $\pm \frac{1}{2}$ wird im Folgenden noch genauer eingegangen:

Die Näherung $P(a \leq X \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - np}{\sqrt{npq}}\right)$ lässt sich besonders für kleine Werte von x beziehungsweise $x - \mu$ oft erheblich verbessern (vor allem, wenn $a = b$ ist), wenn man statt dessen berechnet: $P(a \leq X \leq b) \approx \Phi\left(\frac{b + 0,5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - 0,5 - np}{\sqrt{npq}}\right)$. Diese sogenannte **Kontinuitäts- oder Stetigkeitskorrektur** lässt sich folgendermaßen rechtfertigen:

Die Wahrscheinlichkeit $P(a \leq X \leq b)$ entspricht der Fläche unter der **Wahrscheinlichkeitsfunktion der Binomialverteilung**, beziehungsweise annähernd der unter der **Dichtefunktion der Normalverteilung**.

Die erste Näherung verwendet die Fläche unter der *Dichtefunktion* (schwarze Linie) zwischen den Werten a und b als Maß für die Wahrscheinlichkeit. Sie ist hier mit *waagrechten Strichen* schraffiert:



Exakt wäre aber der *gesamte* (gelb hinterlegte) *Flächeninhalt* der vier Rechtecke.

Die Näherung mit der Korrektur

nimmt noch zusätzlich die (senkrecht schraffierte) Fläche unter der *Dichtefunktion* bis zu den Grenzen $a - 0,5$ und $b + 0,5$ hinzu und verbessert so im Allgemeinen die Approximation.

Definitionsgemäß kann man nur für **große n** gute Näherungen erwarten. Außerdem ist die Binomialverteilung nur dann ausreichend genau normal verteilt, wenn sie in etwa **symmetrisch** ist ($\Rightarrow p \approx 1/2$). Dies schlägt sich auch in der **Faustregel** nieder: $\sigma = \sqrt{n \cdot p \cdot q} > 9$.

Zur Verdeutlichung der Güte der Approximation sollen folgende zwei **Beispiele** dienen:

1) Ein neuer **Schachcomputer fordert den Weltmeister heraus**. Experten schätzen, dass die **Chance zu gewinnen**, bei der Maschine bei **60 %** liegt. Es wird nun ein Turnier über zehn Spiele angesetzt. Wie hoch ist die Wahrscheinlichkeit, dass der Mensch unterliegt, also der Computer bei mindestens sechs Spielen als Sieger hervorgeht (Remis werden wiederholt)?

- Die **binomial** verteilte Wahrscheinlichkeit lässt sich folgendermaßen korrekt berechnen:

$$P(X \geq 6) = 1 - P(X \leq 5) = 1 - F_{10,0,6}(5) = 0,63310 = \underline{\underline{63,3 \%}}$$

- Zur Berechnung durch die Normalverteilung benötigt man Erwartungswert und Standardabweichung: $\mu = 10 \cdot 0,6 = 6$; $\sigma = \sqrt{10 \cdot 0,6 \cdot 0,4} = \frac{2}{5} \sqrt{15}$. Somit erhält man:

$$P(X \geq 6) = 1 - P(X \leq 5) = 1 - \Phi\left(\frac{5-6}{\frac{2}{5}\sqrt{15}}\right) = 1 - \Phi\left(-\frac{\sqrt{15}}{6}\right) = 1 - \Phi(0,6455) = 0,74070 = \underline{\underline{74,1\%}}$$

Der Fehler, den man begeht, beläuft sich auf **17 %**. Laut der Faustregel hätte man hier die Näherung durch die Normalverteilung auch gar nicht anwenden dürfen: $npq = 2,4 < 9$.

Hier zeigt sich aber die Bedeutung der Stetigkeitskorrektur. Das Ergebnis weicht jetzt nur noch um **1 %** vom Idealwert ab:

$$P(X \leq 6) = 1 - P(X \leq 5) = 1 - \Phi\left(\frac{5-6+0,5}{\frac{2}{5}\sqrt{15}}\right) = 1 - \Phi\left(-\frac{\sqrt{15}}{12}\right) = 1 - \Phi(0,3228) = 0,62656 = \underline{\underline{62,7\%}}$$

- 2) Gesucht ist die Wahrscheinlichkeit dafür, dass bei **30 Würfeln mit einem L-Würfel** die (absolute) Häufigkeit der auftretenden Sechsen um **höchstens eins vom erwarteten Mittelwert** abweicht. Man fragt also nach $P(4 \leq X \leq 6)$, da $\mu = np = 30/6 = 5$ ist:

- Exakte Berechnung mit Hilfe der Binomialverteilung:

$$P(4 \leq X \leq 6) = P(3 < X \leq 6) = F_{30;1/6}(6) - F_{30;1/6}(3) = 0,7765 - 0,2396 = 0,53692 = \underline{\underline{53,7\%}}$$

- Näherung durch die Normalverteilung:

$$\mathbf{E(X)} \text{ und } \mathbf{Standardabweichung} \text{ ergeben sich zu: } \mu = 30 \cdot \frac{1}{6} = 5; \quad \sigma = \sqrt{10 \cdot \frac{1}{6} \cdot \frac{5}{6}} = \frac{5}{6} \sqrt{6}$$

$$P(4 \leq X \leq 6) = P(3 < X \leq 6) = \Phi\left(\frac{6-5}{\frac{5}{6}\sqrt{6}}\right) - \Phi\left(\frac{3-5}{\frac{5}{6}\sqrt{6}}\right) = \Phi\left(\frac{\sqrt{6}}{5}\right) - 1 + \Phi\left(\frac{2\sqrt{6}}{5}\right) = \underline{\underline{52,4\%}}$$

Obwohl die Faustregel von $npq > 9$ nicht erfüllt ist ($npq \approx 4,67$), ergibt sich ein relativer Fehler von nur **2,3 %**. Die Stetigkeitskorrektur verbessert das Ergebnis noch:

$$P(3 \leq X \leq 6) = 0,53757 = \underline{\underline{53,8\%}}$$

Der Fehler beträgt nur noch **0,1 %**. Je höher n ist, desto bessere Ergebnisse erwartet man. In der Tat differieren die Werte von $P(9 \leq X \leq 11)$ bei $n = 60$ (also $npq \approx 8,33$) nur noch um **1,2 %** (ohne Kontinuitätskorrektur).

III. 5. Näherung der Normalverteilung an die hypergeometrische Verteilung

Es wurde bereits gezeigt, dass die hypergeometrische eng mit der Binomialverteilung zusammenhängt (siehe Punkt III.1.). Da letztere mit Hilfe der Gaußschen Glockenkurve angenähert werden kann, ist es augenscheinlich, dass auch Zufallsexperimente, die auf dem Modell „Ziehen *mit Zurücklegen*“ basieren, durch die Normalverteilung approximiert werden können. Diese Tatsache lässt sich unter anderem durch den **Zentralen Grenzwertsatz** belegen.

Der exakte Beweis ist aber kompliziert und umfangreich, deshalb soll an dieser Stelle auf die Näherung nur durch die Berechnung von **Beispielen** eingegangen werden:

1) Die Wahrscheinlichkeitsrechnung hat zu Beginn vor allem im Bereich von **Glücksspielen** eine große Rolle gespielt. So ein klassisches Beispiel soll jetzt durchgerechnet werden. Eines der berühmtesten Kartenspiele ist das **Pokern**. Um nicht auf Details der verschiedenen Varianten eingehen zu müssen, sollen hier nur Wahrscheinlichkeiten untersucht werden, mit denen man eine bestimmte Kombination von Karten beim (ersten) Austeilen in die Hand bekommt.

Wie hoch ist nun die Wahrscheinlichkeit, dass man genau **drei der vier** im Spiel enthaltenen **Asse** erhält (jeder Spieler bekommt fünf Karten)?

- Es wird natürlich nicht zurückgelegt und die Zufallsgröße $X = \text{Anzahl der Asse auf der Hand}$ ist **hypergeometrisch** verteilt. Die Wahrscheinlichkeit ist deshalb gegeben durch die Parameter: $N = 52$; $K = 4$; $n = 5$; $k = 3$:

$$P(X = 3) = \frac{\binom{4}{3} \cdot \binom{52-4}{5-3}}{\binom{52}{5}} = \frac{4 \cdot \binom{48}{2}}{\binom{52}{5}} = 0,17 \% \quad \Rightarrow \underline{1:576}$$

- Wendet man die **Normalverteilung** auf dieses Problem an, so errechnet man zuerst:

$$\mu = n \cdot p = 5 \cdot \frac{4}{52} = \frac{5}{13}; \quad \sigma = \sqrt{n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N-n}{N-1}} = \sqrt{5 \cdot \frac{4}{52} \cdot \frac{48}{52} \cdot \frac{52-5}{52-1}} = \sqrt{\frac{5 \cdot 4 \cdot 47}{13 \cdot 13 \cdot 51}} = \sqrt{\frac{940}{2873}} \approx \mathbf{0,572}$$

Jetzt gilt (erfordert **Kontinuitätskorrektur**, da sich sonst die Wahrscheinlichkeit 0 ergibt):

$$P(X = 3) = \Phi\left(\frac{3 + \frac{1}{2} - \frac{5}{13}}{0,572}\right) - \Phi\left(\frac{3 - \frac{1}{2} - \frac{5}{13}}{0,572}\right) = 0,0109 \% = 0,01 \% \quad \Rightarrow \underline{1:9200}$$

Diese Abschätzung ist mit einem **Fehler** von über **93 %** behaftet und deshalb nutzlos.

Besser gelingt die Näherung im nächsten Beispiel:

2) Jedes der 5000 Mitglieder einer Bücherei hat seine eigene Karteikarte. 40 Prozent der Leser sind Frauen. Man zieht nun (ohne Zurücklegen) acht dieser Karten. Wie hoch ist die Wahrscheinlichkeit, dass höchstens sieben Frauen darunter sind (oder, gleichbedeutend damit, dass mindestens ein männliches Mitglied gezogen wird)?

- Diese Zufallsgröße $X = \text{Anzahl der gezogenen Frauen}$ ist hypergeometrisch verteilt. Die Parameter sind: $N = 5000$; $K = 2000$; $n = 8$; $k \leq 7$

Daraus ergibt sich eine exakte Wahrscheinlichkeit von: $P(k \leq 7) =$

$$= P(k = 0) + P(k = 1) + \dots + P(k = 7) = \sum_{k=0}^7 \frac{\binom{2000}{k} \cdot \binom{3000}{8-k}}{\binom{5000}{8}} = \frac{1}{\binom{5000}{8}} \cdot \sum_{k=0}^7 \binom{2000}{k} \cdot \binom{3000}{8-k} =$$

$$= 1,67\% + 8,95\% + 20,9\% + 27,89\% + 23,24\% + 12,38\% + 4,12\% + 0,78\% = \underline{\underline{99,94\%}}$$

- Um die Normalverteilung als Näherung anwenden zu können, muss man zuerst die Größen μ und σ bestimmen: $\mu = n \cdot p = 8 \cdot 0,4 = 3,2$; $\sigma = \sqrt{8 \cdot 0,4 \cdot 0,6 \cdot \frac{4992}{4999}} \approx 1,385$

$$P(k \leq 7) = \Phi\left(\frac{7 - 3,2}{1,385}\right) = \Phi(2,744) = 0,99697 = \underline{\underline{99,70\%}}$$

Bei Anwendung dieser Näherung liegt der Fehler demnach unter **0,24 %**.

Mit **Stetigkeitskorrektur** ergibt sich für $P(k \leq 7) = \Phi(3,105) = \underline{\underline{99,90\%}}$, wobei der Fehler jetzt sogar unter **0,03 %** liegt.

Für eine hypergeometrisch verteilte Zufallsgröße X gilt näherungsweise:

$$P(X = k) \approx \Phi\left(\frac{k - \mu + 0,5}{\sigma}\right) - \Phi\left(\frac{k - \mu - 0,5}{\sigma}\right)$$

$$P(X \leq k) \approx \Phi_{\frac{n \cdot \frac{M}{N} \cdot \frac{M}{N} \cdot (1 - \frac{M}{N}) \cdot \frac{N-n}{N-1}}{\frac{N-n}{N-1}}}(k) = \Phi\left(\frac{k - n \cdot \frac{M}{N} + \frac{1}{2}}{\sqrt{n \cdot \frac{M}{N} \cdot (1 - \frac{M}{N}) \cdot \frac{N-n}{N-1}}}\right)$$

Wobei $+ \frac{1}{2}$ die **Stetigkeitskorrektur** ist.

$$\text{Faustregel: } \sigma = n \cdot p \cdot (1 - p) > 9; \quad \frac{n}{N} \leq 0,1$$

Bemerkung zu der Formel:

Bei der hypergeometrischen Verteilung gilt ja: $\mu = n \cdot p = n \cdot \frac{M}{N}$ und $\sigma^2 = npq \frac{N-n}{N-1}$.

III. 6. Näherung der Normalverteilung an die Poisson-Verteilung

Möchte man die Annäherung der Poisson- an die Normalverteilung untersuchen, so muss man berücksichtigen, dass beide reine mathematisch-theoretisch definierte **Näherungsverteilungen** darstellen. **Es gibt weder in der Theorie noch in der Praxis Größen**, die **genau normal** beziehungsweise **100%-ig Poisson-verteilt** sind.

Deshalb wird in diesem Abschnitt hauptsächlich darauf eingegangen, wie geeignet diese beiden Verteilungen sind, um die exakten Wahrscheinlichkeiten der Binomial- oder der hypergeometrischen Verteilung anzunähern.

Analog zum bisherigen Verfahren werden die Beziehungen mit Hilfe zweier **Beispiele** verdeutlicht:

1) Am **Karlslymnasium Bad Reichenhall** befinden sich insgesamt **878 Schüler**¹². Bei einer leicht idealisierten Annahme einer Wahrscheinlichkeit von $1/365$, dass jemand an einem bestimmten Tag im Jahr Geburtstag hat, wird nach der Wahrscheinlichkeit gefragt, dass **genau sechs Schüler** (zum Beispiel) **an Weihnachten geboren sind**.

Die Werte der Binomial- (= 2,4 %) und der Poisson-Verteilung (= 2,4 %) sowie die **relative Häufigkeit** der Tage, an denen sechs Schüler gleichzeitig Geburtstag haben (= 1,4 %) lassen sich aus der Tabelle auf Seite 35 entnehmen. Die Berechnungen sind hier aber trotzdem aufgeführt und die Ergebnisse erneut festgehalten:

- Die exakte Berechnung erfordert die **Binomialverteilung** mit den Parametern **n = 878**;

$$p = 1/365 \quad \text{und} \quad k = 6: \quad P_B = \binom{878}{6} \cdot \left(\frac{1}{365}\right)^6 \cdot \left(\frac{364}{365}\right)^{872} = 0,02418 = \underline{\underline{2,42 \%}}$$

- Die **Poisson-Näherung** benötigt den Erwartungswert (auszurechnen wie bei binomial verteilten Zufallsgrößen): **$\mu = n \cdot p = 2,41$** ;

$$P_P = e^{-\mu} \cdot \frac{\mu^k}{k!} = e^{-2,41} \cdot \frac{2,41^6}{6!} = 0,0243 = \underline{\underline{2,43 \%}}$$

- Die **Normalverteilung** hat folgende Momente: **$\mu = n \cdot p = 2,41$** ; **$\sigma = \sqrt{n \cdot p \cdot q} = 1,55$** ;

¹² Stand 13. Januar 1998; siehe auch Tabelle im Anhang;

$$P_N = \Phi\left(\frac{6,5 - 2,41}{1,55}\right) - \Phi\left(\frac{5,5 - 2,41}{1,55}\right) = \Phi(2,64) - \Phi(2,00) = \underline{\underline{1,88\%}}$$

Die Normalverteilung (Fehler: über 22 %) stellt hier eine schlechtere Näherung als die Poisson-Approximation (Fehler: 0,40 %) für die Binomialverteilung dar.

2) Ein Zöllner weiß, dass sich unter den nächsten 40 Passanten zwei Schmuggler befinden. Da er nicht alle vierzig Personen kontrollieren kann, aber mindestens einen von ihnen überführen möchte, wählt er n der vierzig Menschen aus. Wie viele muss er untersuchen, um mit einer Wahrscheinlichkeit von mindestens 90 % einen oder beide Täter zu ertappen?

- Es liegt eine hypergeometrische Verteilung vor, mit $N = 40$; $K = 2$; $k \geq 1$;

$$0,9 = P_H(X \geq 1) = 1 - P(X = 0) = 1 - H(40; 2; n; 0) = 1 - \binom{2}{0} \cdot \binom{40-2}{n-0} / \binom{40}{n}$$

$$\Rightarrow \frac{(40-n) \cdot (39-n)}{40 \cdot 39} = 0,1 \quad \text{Die Auswertung dieser quadratischen Gleichung ergibt als den sinnvollen Wert: } \underline{\underline{n = 27}} \quad (n_2 = 52)$$

- Die Binomialverteilung (mit $p = 2/40 = 0,05$) ergibt: $0,9 = P_B = 1 - \binom{n}{0} \cdot 0,05^n \cdot 0,95^{n-0}$;

Daraus folgt: $\underline{\underline{n}} = \ln(0,1)/\ln(0,95) = 44,89 \approx \underline{\underline{45}}$. Es müssten also alle durchsucht werden.

- Der zur Poisson-Verteilung gehörende Erwartungswert ist $\mu = n \cdot p = n \cdot 2/40 = 0,05 \cdot n$

Der Ansatz lautet demnach: $0,9 = P_P = 1 - e^{-0,05 \cdot n} \cdot \frac{(0,05 \cdot n)^0}{0!} = e^{-0,05 \cdot n}$. Durch

Logarithmieren erhält man für $n > 46,05$, das heißt, $\underline{\underline{n}} = \underline{\underline{47}}$. Dies entspricht einer Totalkontrolle.

- Die Normalverteilung ($\mu = 0,05 \cdot n$ und $\sigma = \sqrt{n \cdot \frac{K}{N} \cdot \left(1 - \frac{K}{N}\right) \cdot \frac{N-n}{N-1}} = \sqrt{n \cdot \frac{2}{40} \cdot \frac{38}{40} \cdot \frac{40-n}{39}} = \sqrt{0,0012 \cdot n \cdot (40-n)}$) nähert die Anzahl folgendermaßen:

$$0,9 = P_N = 1 - \Phi\left(\frac{0 - 0,05 \cdot n + 0,5}{\sqrt{0,0012 \cdot n \cdot (40-n)}}\right). \text{ Daraus folgt: } \frac{0 - 0,05 \cdot n + 0,5}{\sqrt{0,0012 \cdot n \cdot (40-n)}} = -1,281.$$

Löst man diese Gleichung, so erhält man $n = 26,7$. Es müssen damit nur $\underline{\underline{n}} = \underline{\underline{27}}$ Leute kontrolliert werden. Vergleicht man nun die verschiedenen Näherungen, die im zweiten Beispiel auftreten, so kommt man zu folgendem Ergebnis:

Das *schlechteste Resultat* erhält man mit der **Poisson-Verteilung** (**Fehler: über 42 %**). Auch die **Binomialverteilung** bringt *keine befriedigende Approximation* mit sich (**Fehler: etwa 40 %**). Allein durch Anwendung der **Normalverteilung** erhält man einen *brauchbaren Wert* (in diesem Fall sogar das **exakte Ergebnis**).

Wie man sehen konnte, sind die Voraussetzungen, die erfüllt sein müssen, damit man die Poisson-Verteilung als annähernd normal verteilt ansehen kann, nicht einfach zu fassen.

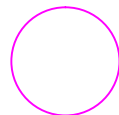
Trotzdem ist in den meisten Fällen folgende Bedingung für eine gute Näherung ausreichend:

$$\sigma = \mu > 9$$

Bei großem Parameter μ gilt für eine **Poisson-verteilte Zufallsgröße X** näherungsweise:

$$P(X = k) = \Phi\left(\frac{k - \mu + 0,5}{\sqrt{\mu}}\right) - \Phi\left(\frac{k - \mu - 0,5}{\sqrt{\mu}}\right)$$

$$P(X \leq k) = \Phi\left(\frac{k - \mu + 0,5}{\sqrt{\mu}}\right)$$



**Hier befinden sich eigentlich die großen Übersichten
über die besprochenen Verteilungen.**

**Dieses Blatt konnte aber nur mit Hilfe eines
Kopierers hergestellt werden.**

**Die verwendeten Daten und Teile befinden sich in
den Dateien ,IVA3.DOC‘ und ,IV.DOC‘.**

ANHANG

An dem jeweiligen Tag haben genau k Schüler Geburtstag.

Tage, an denen kein Gymnasiast Geburtstag hat, sind weggelassen.

| Tag | k | Tag | k | Tag | k | Tag | k | Tag | k | Tag | k |
|---------|---|---------|---|---------|---|---------|---|---------|---|---------|---|
| 01. Jan | 1 | 05. Mär | 1 | 17. Mai | 3 | 23. Jul | 4 | 23. Sep | 2 | 29. Nov | 5 |
| 02. Jan | 3 | 06. Mär | 4 | 18. Mai | 4 | 24. Jul | 5 | 25. Sep | 2 | 30. Nov | 5 |
| 03. Jan | 7 | 07. Mär | 1 | 19. Mai | 3 | 25. Jul | 3 | 26. Sep | 1 | 01. Dez | 3 |
| 04. Jan | 7 | 08. Mär | 2 | 20. Mai | 2 | 26. Jul | 1 | 27. Sep | 3 | 02. Dez | 2 |
| 05. Jan | 2 | 09. Mär | 3 | 21. Mai | 4 | 27. Jul | 9 | 28. Sep | 1 | 03. Dez | 1 |
| 06. Jan | 4 | 10. Mär | 2 | 22. Mai | 4 | 28. Jul | 5 | 30. Sep | 4 | 04. Dez | 3 |
| 07. Jan | 2 | 11. Mär | 3 | 23. Mai | 1 | 29. Jul | 3 | 03. Okt | 2 | 05. Dez | 1 |
| 08. Jan | 7 | 12. Mär | 1 | 25. Mai | 1 | 30. Jul | 6 | 04. Okt | 2 | 06. Dez | 2 |
| 09. Jan | 3 | 13. Mär | 1 | 26. Mai | 1 | 31. Jul | 3 | 06. Okt | 1 | 07. Dez | 1 |
| 10. Jan | 5 | 15. Mär | 4 | 27. Mai | 2 | 01. Aug | 1 | 08. Okt | 1 | 08. Dez | 3 |
| 13. Jan | 2 | 16. Mär | 5 | 29. Mai | 2 | 02. Aug | 1 | 09. Okt | 1 | 09. Dez | 8 |
| 14. Jan | 3 | 17. Mär | 3 | 31. Mai | 1 | 03. Aug | 2 | 10. Okt | 3 | 10. Dez | 2 |
| 15. Jan | 3 | 18. Mär | 3 | 01. Jun | 1 | 04. Aug | 2 | 11. Okt | 4 | 11. Dez | 2 |
| 16. Jan | 6 | 19. Mär | 1 | 02. Jun | 1 | 05. Aug | 1 | 12. Okt | 2 | 12. Dez | 2 |
| 17. Jan | 3 | 21. Mär | 2 | 03. Jun | 4 | 07. Aug | 2 | 13. Okt | 7 | 13. Dez | 4 |
| 18. Jan | 2 | 22. Mär | 4 | 05. Jun | 6 | 08. Aug | 3 | 14. Okt | 1 | 14. Dez | 5 |
| 19. Jan | 1 | 24. Mär | 3 | 07. Jun | 2 | 09. Aug | 3 | 15. Okt | 2 | 15. Dez | 4 |
| 20. Jan | 1 | 25. Mär | 1 | 08. Jun | 2 | 10. Aug | 1 | 16. Okt | 2 | 16. Dez | 2 |
| 21. Jan | 2 | 27. Mär | 3 | 09. Jun | 2 | 11. Aug | 2 | 17. Okt | 1 | 17. Dez | 2 |
| 22. Jan | 2 | 28. Mär | 2 | 10. Jun | 1 | 12. Aug | 4 | 18. Okt | 3 | 18. Dez | 7 |
| 23. Jan | 1 | 29. Mär | 2 | 11. Jun | 4 | 13. Aug | 1 | 19. Okt | 3 | 19. Dez | 2 |
| 24. Jan | 1 | 30. Mär | 3 | 12. Jun | 4 | 14. Aug | 2 | 20. Okt | 3 | 20. Dez | 5 |
| 25. Jan | 4 | 31. Mär | 5 | 13. Jun | 5 | 15. Aug | 4 | 21. Okt | 3 | 21. Dez | 5 |
| 26. Jan | 2 | 01. Apr | 2 | 14. Jun | 3 | 16. Aug | 1 | 22. Okt | 1 | 22. Dez | 4 |
| 27. Jan | 1 | 02. Apr | 2 | 15. Jun | 2 | 17. Aug | 1 | 23. Okt | 4 | 23. Dez | 2 |
| 28. Jan | 1 | 03. Apr | 2 | 16. Jun | 1 | 18. Aug | 2 | 24. Okt | 2 | 24. Dez | 3 |
| 30. Jan | 7 | 04. Apr | 3 | 17. Jun | 2 | 19. Aug | 3 | 25. Okt | 1 | 25. Dez | 3 |
| 31. Jan | 5 | 08. Apr | 1 | 18. Jun | 2 | 20. Aug | 1 | 26. Okt | 3 | 26. Dez | 1 |
| 01. Feb | 2 | 09. Apr | 7 | 19. Jun | 1 | 21. Aug | 4 | 27. Okt | 2 | 27. Dez | 1 |
| 02. Feb | 2 | 10. Apr | 3 | 21. Jun | 1 | 22. Aug | 3 | 28. Okt | 1 | 28. Dez | 2 |
| 03. Feb | 3 | 11. Apr | 5 | 22. Jun | 1 | 23. Aug | 3 | 29. Okt | 5 | 29. Dez | 4 |
| 05. Feb | 2 | 12. Apr | 1 | 23. Jun | 4 | 24. Aug | 3 | 30. Okt | 2 | 30. Dez | 3 |
| 07. Feb | 3 | 13. Apr | 1 | 24. Jun | 1 | 26. Aug | 2 | 31. Okt | 1 | 31. Dez | 6 |
| 08. Feb | 4 | 15. Apr | 2 | 25. Jun | 3 | 28. Aug | 5 | 01. Nov | 2 | | |
| 09. Feb | 4 | 16. Apr | 3 | 26. Jun | 2 | 30. Aug | 1 | 02. Nov | 5 | | |
| 10. Feb | 3 | 17. Apr | 3 | 27. Jun | 2 | 31. Aug | 1 | 03. Nov | 1 | | |
| 12. Feb | 2 | 18. Apr | 5 | 28. Jun | 1 | 01. Sep | 1 | 04. Nov | 4 | | |
| 13. Feb | 3 | 19. Apr | 1 | 29. Jun | 4 | 02. Sep | 3 | 06. Nov | 4 | | |
| 14. Feb | 2 | 21. Apr | 5 | 02. Jul | 1 | 03. Sep | 4 | 07. Nov | 1 | | |
| 15. Feb | 3 | 22. Apr | 1 | 03. Jul | 3 | 04. Sep | 2 | 08. Nov | 2 | | |
| 16. Feb | 1 | 23. Apr | 3 | 04. Jul | 2 | 05. Sep | 3 | 09. Nov | 5 | | |
| 17. Feb | 2 | 25. Apr | 4 | 05. Jul | 2 | 06. Sep | 5 | 10. Nov | 4 | | |
| 18. Feb | 3 | 26. Apr | 2 | 07. Jul | 6 | 07. Sep | 3 | 11. Nov | 1 | | |
| 19. Feb | 1 | 27. Apr | 2 | 08. Jul | 5 | 09. Sep | 3 | 12. Nov | 3 | | |
| 20. Feb | 2 | 28. Apr | 4 | 09. Jul | 4 | 10. Sep | 4 | 13. Nov | 5 | | |
| 21. Feb | 1 | 01. Mai | 2 | 10. Jul | 5 | 11. Sep | 2 | 14. Nov | 2 | | |
| 22. Feb | 2 | 02. Mai | 1 | 11. Jul | 2 | 12. Sep | 4 | 15. Nov | 2 | | |
| 23. Feb | 1 | 03. Mai | 7 | 12. Jul | 3 | 13. Sep | 5 | 16. Nov | 1 | | |
| 24. Feb | 2 | 05. Mai | 1 | 13. Jul | 1 | 14. Sep | 8 | 17. Nov | 3 | | |
| 25. Feb | 4 | 06. Mai | 2 | 14. Jul | 2 | 15. Sep | 4 | 18. Nov | 4 | | |
| 26. Feb | 2 | 07. Mai | 5 | 15. Jul | 4 | 16. Sep | 1 | 19. Nov | 2 | | |
| 27. Feb | 2 | 08. Mai | 1 | 16. Jul | 1 | 17. Sep | 2 | 20. Nov | 1 | | |
| 28. Feb | 2 | 10. Mai | 2 | 17. Jul | 2 | 18. Sep | 3 | 21. Nov | 2 | | |
| 01. Mär | 2 | 12. Mai | 7 | 19. Jul | 2 | 19. Sep | 1 | 24. Nov | 4 | | |
| 02. Mär | 5 | 13. Mai | 3 | 20. Jul | 2 | 20. Sep | 1 | 25. Nov | 2 | | |
| 03. Mär | 5 | 14. Mai | 4 | 21. Jul | 5 | 21. Sep | 1 | 26. Nov | 3 | | |
| 04. Mär | 4 | 16. Mai | 5 | 22. Jul | 2 | 22. Sep | 8 | 28. Nov | 1 | | |

Verzeichnis der zur Anfertigung dieser Arbeit herangezogenen Bücher:

- „Schülerduden Die Mathematik II“; Dudenverlag; 3. Auflage, Mannheim 1991
- J. Pfanzagl „Allgemeine Methodenlehre der Statistik II“;
Walter de Gruyter, 2. Auflage, Berlin 1966
- F. Reinhardt / H. Soeder „dtv-Atlas zur Mathematik“, Band 2;
Deutscher Taschenbuchverlag, 9. Auflage, München 1994
- H. Meschkowski „Wahrscheinlichkeitsrechnung“; Bibliographisches Institut AG, Mannheim 1968
- Stöckner „Taschenbuch mathematischer Formeln und moderner Verfahren“, Seiten 656 - 663
- J. Laub „Lehrbuch der Mathematik für die Oberstufe der allgemeinbildenden höheren Schulen“,
4. Band; Hölder-Pichler-Tempsky, Wien 1981
- L. Schmetterer „Einführung in die mathematische Statistik“; Springer-Verlag, Wien 1956
- Barth · Bergold · Haller „Stochastik 1“; Franz Ehrenwirth Verlag, München 1976
- Barth · Bergold · Haller „Stochastik 2“; Franz Ehrenwirth Verlag, München 1974
- Heller · Lindenberg · Nuske · Schriever „Wahrscheinlichkeitsrechnung“, Teil 1;
Birkhäuser Verlag, Basel 1979
- Heller · Lindenberg · Nuske · Schriever „Wahrscheinlichkeitsrechnung“, Teil 2;
Birkhäuser Verlag, Basel 1979
- H. Malle „Mathematik für Techniker“, Band 2, Seiten 114 - 149
- William Feller „An Introduction to Probability Theory and Its Applications“, Volume 1;
Wiley & Sons, New York, Seiten 172 - 192
- Heigl / Feurpfeil „Stochastik, Leistungskurs Mathematik“;
Bayerischer Schulbuch-Verlag, 2. Auflage, München 1976

Ich erkläre hiermit, dass ich die Facharbeit ohne fremde Hilfe angefertigt und nur die im Literaturverzeichnis angeführten Quellen und Hilfsmittel benützt habe.

Bad Reichenhall 2.2.98
..... , den

Ort Datum Unterschrift des Schülers