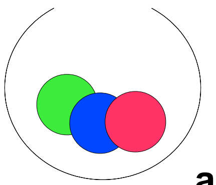
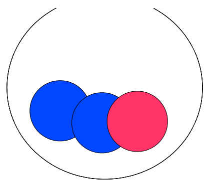
Kombinatorik

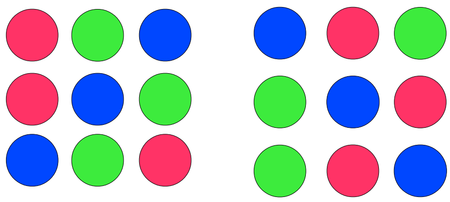
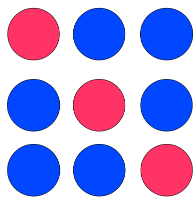
n Auswahl aus einer n Menge





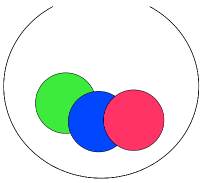
Vollständige Anordnung

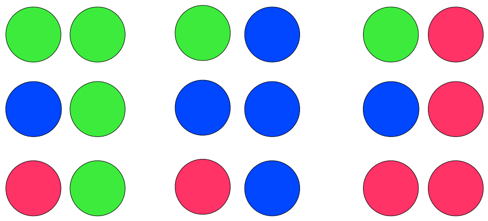
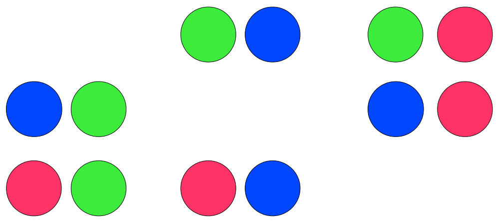
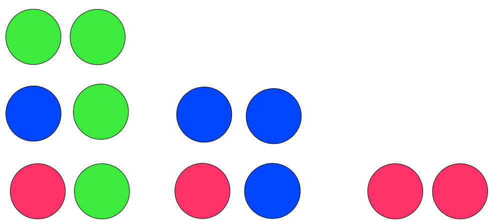
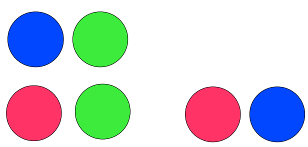
Permutation

alle unterscheidbar	nicht alle unterscheidbar
	
n! : 3! = 6	$\frac{n!}{k!}$: $\frac{3 \bullet 2 \bullet 1}{2 \bullet 1} = 3$
alle Positionen verschiedene Farben	vertauschte Positionen nicht unterscheidbar

Kombinatorik

k Auswahl aus einer n Menge

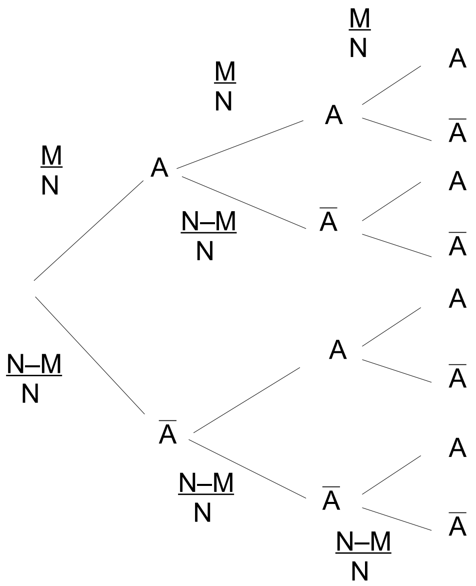


<table>
<tr><th></th><th>mit Zurücklegen</th><th>ohne Zurücklegen</th><th></th></tr>
<tr><td rowspan="3">Geordnete Stichprobe</td><td></td><td></td><td rowspan="3">Variation</td></tr>
<tr><td>n^k : $3^2 = 9$</td><td>$\frac{n!}{(n-k)!}$: $\frac{3 \cdot 2 \cdot 1}{1}$ = 3 2 1 = 6</td></tr>
<tr><td>alle Positionen besetzt</td><td>es fehlen die gleichfarbigen</td></tr>
<tr><td rowspan="3">Ungeordnete Stichprobe</td><td></td><td></td><td rowspan="3">Kombination</td></tr>
<tr><td>$\binom{n+k-1}{k}$: $\binom{3+2-1}{2} = 6$</td><td>$\binom{n}{k}$: $\binom{3}{2} = 3$</td></tr>
<tr><td>es fehlen die vertauschten</td><td>es fehlen die gleichfarbigen und die vertauschten</td></tr>
</table>

1. Mit Zurücklegen

Wahrscheinlichkeitsereignisse, die mit Zurücklegen auszuführen sind, haben bei jedem neuen Versuch wieder die die gleichen Verhältnisse, wie am Anfang. Damit gibt es für jeden Versuch eine feste Wahrscheinlichkeit, die sich mit der Anzahl der Versuche nicht ändert.

Betrachtet man Versuche „mit Wiederholung“, dann kann man zunächst den Baum erstellen.



Bekanntermaßen sind Ziehungen mit Zurücklegen stochastisch unabhängig. Damit sind die Wahrscheinlichkeiten an allen Bäumen gleich. Damit gilt für die Wahrscheinlichkeit von einer Stufe zur nächsten:

$P(X_1 = A) = M/N$ $\quad$ $P(X_1 = \bar{A}) = (N-M)/N$

$P(X_2 = A) = M/N$ $\quad$ $P(X_2 = \bar{A}) = (N-M)/N$

… $\quad$ …

$P(X_n = A) = M/N$ $\quad$ $P(X_n = \bar{A}) = (N-M)/N$

Allgemein kann man für die Wahrscheinlichkeit, dass nach k – Versuchen k mal das Ereignis A eingetreten ist und damit bei n Versuchen notwendigerweise n – k mal das Ereignis A eingetreten sein muss, folgende Wahrscheinlichkeiten ermitteln:

$$P(X_k = A) = \left(\frac{M}{N}\right)^k \qquad P(X_k = \bar{A}) = \left(\frac{N-M}{N}\right)^{n-k}$$

Die Wahrscheinlichkeiten sind an allen Bäumen gleich und sind nur abhängig von der internen Verteilung der Eigenschaften (z.B. Farben), aber nicht davon, wie viele Elemente gezogen werden. Die Anzahl der gezogenen Element findet nur in den Exponenten ihren Niederschlag.

1.1. Mit Zurücklegen, mit Reihenfolge

1.1.1. Zwei Unterscheidungsmerkmale

In einer Urne sind
N Kugeln
M rote Kugeln
N – M schwarze Kugeln

Es werden
n Kugeln gezogen
k Kugeln davon sollen rot sein

Für geordnete Stichproben ist im Baumdiagramm jeder Pfad einzeln zu betrachten.

Kombinatorik: „Variation mit Wiederholung“

Aus der Kombinatorik lässt sich folgende Formel herleiten:

Gesamtanzahl: N^n	rote Kugel: M^k	schwarze Kugel: $(N-M)^{n-k}$
Anzahl der Möglichkeiten • aus N Kugeln • n Kugeln • mit Wiederholung auszuwählen	Anzahl der Möglichkeiten • aus M roten Kugeln • k Kugeln • mit Wiederholung auszuwählen	Anzahl der Möglichkeiten • aus N – M roten Kugeln • n – k Kugeln • mit Wiederholung auszuwählen

da k rote Kugeln (nur) gezogen wurden, müssen n–k Kugeln schwarz sein:

$$\mathbf{P} = \frac{M^k \ (N-M)^{n-k}}{N^n} = \frac{M^k \ (N-M)^{n-k}}{N^k \ \ N^{n-k}} = p^k \ q^{n-k}$$

Dabei ist $\frac{M}{N}$ = p die Wahrscheinlichkeit aus N Kugeln M rote zu ziehen, und

$\frac{N-M}{N}$ = q die Wahrscheinlichkeit aus N – M schwarzen n – k schwarze zu ziehen.

1.1.2. Drei Unterscheidungsmerkmale

gezogene rote Kugeln : r Wahrscheinlichkeit rote Kugeln : p_r
gezogene weiße Kugeln : w Wahrscheinlichkeit weiße Kugeln : p_w
gezogene blaue Kugeln : b Wahrscheinlichkeit blaue Kugeln : p_b

r + w + b = n

Zuerst werden nur zwei Unterscheidungsmerkmale betrachtet, da dafür die Formel gesichert ist.
Es sollen r rote und n – r nicht rote gezogen werden
Damit werden insgesamt n Kugeln angeordnet, die die Anzahl der Möglichkeiten angibt.

$$P_{n;pr}(X = r) = p_r^{\,r} \ (1 - p_r)^{n-r} = p_r^{\,r} \ (p_w + p_b)^{w+b} = p_r^{\,r} \ (p_w + p_b)^w \ (p_w + p_b)^b$$

Aus n – r = w + b nicht rote Kugeln sollen w weiße und b blaue gezogen werden, b blaue sind damit auch b „nicht weiße“ also n – r – w von der verbleibenden Gesamtanzahl.
Bezeichnet man die Gesamtzahl der roten Kugeln mit R, die der weißen mit W und die blauen mit B, so ergibt sich die Wahrscheinlichkeit aus: p_r = R/(R+W+B) usw.
Für den Teil, dass jetzt nur noch weiße und blaue Kugeln aber keine roten mehr betrachtet werden, ändern sich die Wahrscheinlichkeiten:

$$\frac{W}{W+B} = \frac{W \quad : N}{(W+B) : N} = \frac{p_w}{p_w + p_b}$$ (nur an den weißen Kugeln demonstriert)

$$P_{n-r\,;\,pw}(X = k_w) = \left(\frac{p_w}{p_w + p_b}\right)^{w} \left(\frac{p_b}{p_w + p_b}\right)^{n-r-w} = \left(\frac{p_w}{p_w + p_b}\right)^{w} \left(\frac{p_b}{p_w + p_b}\right)^{b}$$

Da beide Ereignisse unabhängig sind, sind die Wahrscheinlichkeiten zu multiplizieren.

$$P_{n;pr;pw;pb}\,(X = (r,w,b)\,) = P_{n;pr}(X = r) \bullet P_{n;(1-p);(n-r)}(X = n\text{–}r) = p_r^{\ r}\,(p_w + p_b)^{w}\,(p_w + p_b)^{b}$$

$$= p_r^{\ r} \bullet (p_w + p_b)^{w}\,(p_w + p_b)^{b} \left(\frac{p_w}{p_w + p_b}\right)^{w} \left(\frac{p_b}{p_w + p_b}\right)^{b}$$

$$= p_r^{\ r}\ p_w^{\ w}\ p_b^{\ b}$$

1.1.3. Vier Unterscheidungsmerkmale

Es sind rote, grüne, blaue und weiße Kugeln vorhanden. Die Wahrscheinlichkeiten, jeweils eine davon zu ziehen, betragen p_R, p_G, p_B und p_W (die Summe beträgt 1) Die Wahrscheinlichkeit, genau r rote, g grüne, b blaue und w weiße (insgesamt n) MIT ZURÜCKLEGEN zu ziehen ist :

$$\boxed{\mathbf{P} = p_R^{\ r}\ p_G^{\ g}\ p_B^{\ b}\ p_W^{\ w}}$$

1.2. Mit Zurücklegen, ohne Reihenfolge

Dieser Fall kann nicht über die Kombinatorikformel hergeleitet werden, da die Anzahlbestimmung $\binom{n-k+1}{k}$ zu keiner Laplace – Wahrscheinlichkeit führt. Deshalb wird hier auf die vorhergehende Herleitung zurückgegriffen und zusätzlich die Anzahl der Pfade bestimmt.

1.2.1. Zwei Unterscheidungsmerkmale

In einer Urne sind
N Kugeln
M rote Kugeln
N – M schwarze Kugeln

Es werden
n Kugeln gezogen
k Kugeln davon sollen rot sein

Für ungeordnete Stichproben sind im Baumdiagramm die Pfade zusammenzufassen, die zur gleichen Auswahl führen

Wie viele Pfade gibt es in einem Baumdiagramm, die zur jeweils gleichen Ereignisanzahl führen. Dazu liefern die kombinatorischen Regeln den Weg:
Es ist die Permutation von n Elementen gesucht, bei der jeweils k und n – k Elemente sich nicht unterscheiden lassen. Damit handelt es sich um eine Permutation mit Wiederholung.

Die zugehörige Formel lautet: $\frac{n!}{k!\,(n-k)!}$

Diese Formel entspricht genau der Definition des Binomialkoeffizienten. Damit ergibt sich die Wahrscheinlichkeit für n–maliges Ziehen mit Wiederholung mit k positiven Ereignissen und n – k negativen Ereignissen über alle möglichen Zweige des Baumdiagramms:

$$\binom{n}{k}\left(\frac{M}{N}\right)^{k}\left(\frac{N-M}{N}\right)^{n-k} = \binom{n}{k} p^k q^{n-k}$$

Das ist genau die Formel der Binomialverteilung.

1.2.2. Drei Unterscheidungsmerkmale

gezogene rote Kugeln : r Wahrscheinlichkeit rote Kugeln : p_r
gezogene weiße Kugeln : w Wahrscheinlichkeit weiße Kugeln : p_w
gezogene blaue Kugeln : b Wahrscheinlichkeit blaue Kugeln : p_b

$r + w + b = n$

Es sollen r rote und n – r nicht rote gezogen werden
Damit werden insgesamt n Kugeln angeordnet, die die Anzahl der Möglichkeiten angibt.

$$P_{n;pr}(X = r) = \binom{n}{r} p_r^{\,r}\,(1-p_r)^{n-r} = \frac{n!}{(n-r)!\;r!}\,p_r^{\,r}\,(1-p_r)^{n-r} = \frac{n!}{(n-r)!\;r!}\,p_r^{\,r}\,(p_w+p_b)^{w+b} =$$

$$= \frac{n!}{(n-r)!\;r!}\,p_r^{\,r}\,(p_w+p_b)^{w}\,(p_w+p_b)^{b}$$

Aus n – r = w + b nicht rote Kugeln sollen w weiße und b blaue gezogen werden, b blaue sind damit auch b „nicht weiße" von der verbleibenden Gesamtanzahl.
Bezeichnet man die Gesamtzahl der roten Kugeln mit R, die der weißen mit W und die blauen mit B, so ergibt sich die Wahrscheinlichkeit aus: $p_r = R/(R+W+B)$ usw.
Für den Teil, dass jetzt nur noch weiße und blaue Kugeln aber keine roten mehr betrachtet werden, ändern sich die Wahrscheinlichkeiten:

$$\frac{W}{W+B} = \frac{W \;:N}{(W+B):N} = \frac{p_w}{p_w+p_b}$$ (nur an den weißen Kugeln demonstriert)

$$P_{n-r\,;\,pw}(X = k_w) = \binom{n-r}{w} \left(\frac{p_w}{p_w + p_b}\right)^w \left(\frac{p_b}{p_w + p_b}\right)^{n-r-w} = \frac{(n-r)!}{(n-r-w)!\ w!} \left(\frac{p_w}{p_w + p_b}\right)^w \left(\frac{p_b}{p_w + p_b}\right)^{n-r-w}$$

$$= \frac{(n-r)!}{b!\ w!} \left(\frac{p_w}{p_w + p_b}\right)^w \left(\frac{p_b}{p_w + p_b}\right)^b$$

Da beide Ereignisse unabhängig sind, sind die Wahrscheinlichkeiten zu multiplizieren.

$$= \frac{n!}{(n-r)!\ r!}\ p_r^{\ r}\ (p_w + p_b)^w\ (p_w + p_b)^b \bullet \frac{(n-r)!}{b!\ w!} \left(\frac{p_w}{p_w + p_b}\right)^w \left(\frac{p_b}{p_w + p_b}\right)^b$$

$$= \frac{n!}{(n-r)!\ r!}\ \frac{(n-r)!}{b!\ w!}\ p_r^{\ r}\ (p_w + p_b)^w\ (p_w + p_b)^b \left(\frac{p_w}{p_w + p_b}\right)^w \left(\frac{p_b}{p_w + p_b}\right)^b$$

$$= \frac{n!}{r!\ b!\ w!}\ p_r^{\ r}\ p_w^{\ w}\ p_b^{\ b}$$

1.2.3. Vier Unterscheidungsmerkmale

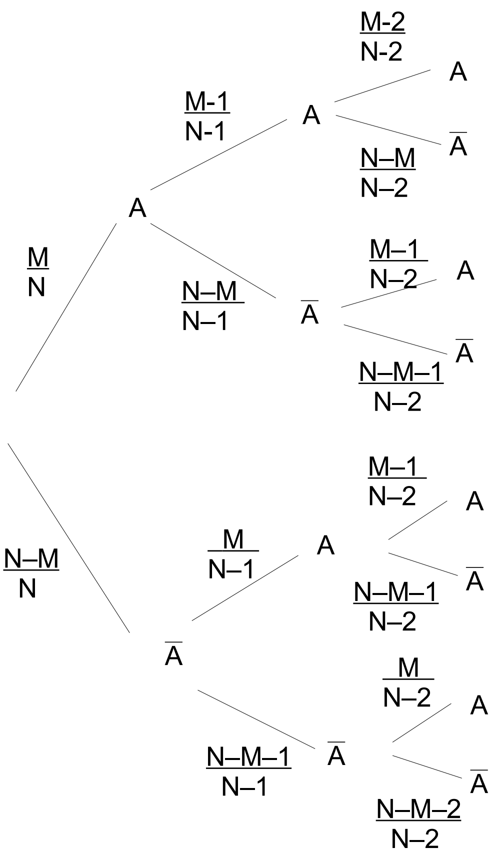
Es sind rote, grüne, blaue und weiße Kugeln vorhanden. Die Wahrscheinlichkeiten, jeweils eine davon zu ziehen, betragen p_R, p_G, p_B und p_W (die Summe beträgt 1) Die Wahrscheinlichkeit, genau r rote, g grüne, b blaue und w weiße (insgesamt n) MIT ZURÜCKLEGEN zu ziehen ist :

$$\mathbf{P} = \frac{n\,!}{r!\ g!\ b!\ w!}\ p_R^{\ r}\ p_G^{\ g}\ p_B^{\ b}\ p_W^{\ w}$$

2. Ohne Zurücklegen

Wahrscheinlichkeitsereignisse, die ohne Zurücklegen auszuführen sind, können nicht mit Wahrscheinlichkeiten arbeiten, da sich diese ständig ändern.
Solche Aufgaben können nur über die (aktuellen) Anzahlen berechnet werden.

Betrachtet man Versuche „ohne Wiederholung“, dann ergibt sich folgender Baum:



			$AA\overline{A}$	$A\overline{A}\,\overline{A}$
$\frac{M}{N}$	$\frac{M-1}{N-1}$	$\frac{M-2}{N-2}$		
$\frac{M}{N}$	$\frac{M-1}{N-1}$	$\frac{N-M}{N-2}$	$\frac{M\,(M-1)\,(N-M)}{N\,(N-1)\,(N-2)}$	
$\frac{M}{N}$	$\frac{N-M}{N-1}$	$\frac{M-1}{N-2}$	$\frac{M\,(M-1)\,(N-M)}{N\,(N-1)\,(N-2)}$	
$\frac{M}{N}$	$\frac{N-M}{N-1}$	$\frac{N-M-1}{N-2}$		$\frac{M\,(N-M)(N-M-1)}{N\,(N-1)\,(N-2)}$
$\frac{N-M}{N}$	$\frac{M}{N-1}$	$\frac{M-1}{N-2}$	$\frac{M\,(M-1)\,(N-M)}{N\,(N-1)\,(N-2)}$	
$\frac{N-M}{N}$	$\frac{M}{N-1}$	$\frac{N-M-1}{N-2}$		$\frac{M\,(N-M)(N-M-1)}{N\,(N-1)\,(N-2)}$
$\frac{N-M}{N}$	$\frac{N-M-1}{N-1}$	$\frac{M}{N-2}$		$\frac{M\,(N-M)(N-M-1)}{N\,(N-1)\,(N-2)}$
$\frac{N-M}{N}$	$\frac{N-M-1}{N-1}$	$\frac{N-M-2}{N-2}$		

Es ist zu erkennen, dass die Nenner in allen Fällen gleich sind. Außerdem sind die Wahrscheinlichkeiten für und für ebenfalls gleich, unabhängig davon, wann welches Ereignis eingetreten ist.

Bei Ziehungen ohne Zurücklegen ändert sich an jedem Baum die Wahrscheinlichkeit, da mit jedem Ziehen die Anzahl der Kugeln sich reduziert, aber auch die Farbzusammensetzungen ändern sich ständig.

Zwei Ziehungen

Bei beliebigen N und M ergeben sich für die Anzahlen in Zähler und Nenner folgende Brüche:

Zweimal Ereignis A:

$$P(X_1=A) = \frac{M}{N} \qquad P(X_2=A \mid X_1=A) = \frac{M-1}{N-1}$$

$$\frac{M\,(M-1)}{N\,(N-1)} = \frac{\frac{M!}{(M-2)!}}{\frac{N!}{(N-2)!}} = \frac{\frac{M!}{(M-2)!\,2!}}{\frac{N!}{(N-2)!\,2!}} = \frac{\binom{M}{2}}{\binom{N}{2}}$$

Das erste Mal Ereignis A, das zweite Mal Ereignis $\overline{A}$:

$$P(X_1=A) = \frac{M}{N} \qquad P(X_2=\overline{A} \mid X_1=A) = \frac{N-M}{N-1}$$

$$\frac{M\;(N-M)}{N\;(N-1)} = \frac{\frac{M!}{(M-1)!}\;\frac{(N-M)!}{(N-M-1)!}}{\frac{N!}{(N-2)!}} = \frac{\frac{M!}{(M-1)!\,1!}\;\frac{(N-M)!}{(N-M-1)!\,1!}}{\frac{N!}{(N-2)!\,2!}}\;\frac{1!\,1!}{2!} = \frac{\binom{M}{1}\binom{N-M}{1}}{\binom{N}{2}\;\binom{2}{1}}$$

Das erste Mal Ereignis $\bar{A}$, das zweite Mal Ereignis A :

$$P(X_1 = \bar{A}) = \frac{N-M}{N}$$

$$P(X_2 = A \,|\, X_1 = \bar{A}) = \frac{M}{N-1}$$

$$\frac{M\ (N-M)}{N\ (N-1)}$$

es entsteht der gleiche Quotient wie im vorhergehenden Fall, deshalb gibt es die gleiche Umrechnung.

$$\frac{\binom{M}{1}\binom{N-M}{1}}{\binom{N}{2}\binom{2}{1}}$$

Zweimal Ereignis $\bar{A}$:

$$P(X_1 = \bar{A}) = \frac{N-M}{N}$$

$$P(X_2 = \bar{A} \,|\, X_1 = \bar{A}) = \frac{N-M-1}{N-1}$$

$$\frac{(N-M)\ (N-M-1)}{N\ (N-1)} = \frac{\frac{(N-M)!}{(N-M-2)!}}{\frac{N!}{(N-2)!}} = \frac{\frac{(N-M)!}{(M-2)!\ 2}}{\frac{N!}{(N-2)!\ 2}} = \frac{\binom{N-M}{2}}{\binom{N}{2}}$$

Drei Ziehungen

Dreimal Ereignis A:

$$P(X_1 = A) = \frac{M}{N}$$

$$P(X_2 = A \,|\, X_1 = A) = \frac{M-1}{N-1}$$

$$P(X_3 = A \,|\, X_2 = A \,|\, X_1 = A) = \frac{M-2}{N-2}$$

$$\frac{M\ (M-1)\ (M-2)}{N\ (N-1)\ (N-2)} = \frac{\frac{M!}{(M-3)!}}{\frac{N!}{(N-3)!}} = \frac{\frac{M!}{(M-2)!\ 3!}}{\frac{N!}{(N-2)!\ 3!}} = \frac{\binom{M}{3}}{\binom{N}{3}}$$

Das erste Mal Ereignis A, das zweite und dritte Mal Ereignis $\bar{A}$:

$$P(X_1 = A) = \frac{M}{N}$$

$$P(X_2 = \bar{A} \,|\, X_1 = A) = \frac{N-M}{N-1}$$

$$P(X_3 = \bar{A} \,|\, X_2 = \bar{A} \,|\, X_1 = A) = \frac{N-M-1}{N-2}$$

$$\frac{M\ \ (N-M)\ (N-M-1)}{N\ \ (N-1)\ \ (N-2)} = \frac{\frac{M!}{(M-1)!}\ \frac{(N-M)!}{(N-M-2)!}}{\frac{N!}{(N-3)!}} =$$

$$\frac{\frac{M!}{(M-1)!\ 1!}\ \frac{(N-M)!}{(N-M-2)!\ 2!}}{\frac{N!}{(N-3)!\ 3!}}\ \frac{1!\ 2!}{3!} = \frac{\frac{M!}{(M-1)!\ 1!}\ \frac{(N-M)!}{(N-M-2)!\ 2!}}{\frac{N!}{(N-3)!\ 3!}}\ \frac{1}{\frac{3!}{1!\ 2!}} = \frac{\binom{M}{1}\binom{N-M}{2}}{\binom{N}{3}\binom{3}{1}}$$

Rechnung mit Doppelbrüchen

Der Zähler wird durch 3! dividiert, der Nenner aber nur durch 2! (1! ist 1 und damit uninteressant). Deshalb muss der Nenner noch durch 3 dividiert werden, damit die Brüche übereinstimmen.

Das erste und das zweite Mal Ereignis A, das dritte Mal Ereignis $\bar{A}$:

$$P(X_1 = A) = \frac{M}{N}$$

$$P(X_2 = A \,|\, X_1 = A) = \frac{M-1}{N-1}$$

$$P(X_3 = \bar{A} \,|\, X_2 = A \,|\, X_1 = A) = \frac{N-M}{N-2}$$

$$\frac{M\ \ (M-1)\ (N-M)}{N\ \ (N-1)\ \ (N-2)} = \frac{\frac{M!}{(M-2)!}\ \frac{(N-M)!}{(N-M-1)!}}{\frac{N!}{(N-3)!}} =$$

$$\frac{\frac{M!}{(M-2)!\ 2!}\ \frac{(N-M)!}{(N-M-1)!\ 1!}}{\frac{N!}{(N-3)!\ 3!}}\ \frac{1!\ 2!}{3!} = \frac{\frac{M!}{(M-2)!\ 2!}\ \frac{(N-M)!}{(N-M-1)!\ 1!}}{\frac{N!}{(N-3)!\ 3!}}\ \frac{1}{\frac{3!}{2!\ 1!}} = \frac{\binom{M}{2}\binom{N-M}{1}}{\binom{N}{3}\binom{3}{2}}$$

Der zweite Binomialkoeffizient entsteht aus dem gleichen Grund wie oben.

Dreimal Ereignis $\bar{A}$:

$$P(X_1=\bar{A}) = \frac{N-M}{N}$$

$$P(X_2=\bar{A} \mid X_1=\bar{A}) = \frac{N-M-1}{N-1}$$

$$P(X_3=\bar{A} \mid X_2=\bar{A} \mid X_1=\bar{A}) = \frac{N-M-2}{N-2}$$

$$\frac{(N-M)\,(N-M-1)\,(N-M-2)}{N\,(N-1)\,(N-2)} = \frac{\frac{(N-M)!}{(N-M-3)!}}{\frac{N!}{(N-3)!}} = \frac{\frac{(N-M)!}{(N-M-3)!\,3!}}{\frac{N!}{(N-2)!\,3!}} = \frac{\binom{N-M}{3}}{\binom{N}{3}}$$

Berücksichtigt man weiterhin, dass alle Binomialkoeffizienten, deren untere Zahl Null ist, den Wert 1 besitzen, dann kann man folgende Ergänzung vornehmen:

3xA : $\dfrac{\binom{M}{3}\binom{N-M}{0}}{\binom{N}{3}}$ 2xA ; 1x $\bar{A}$: $\dfrac{\binom{M}{2}\binom{N-M}{1}}{\binom{N}{3}\binom{3}{2}}$ 1xA ; 2x $\bar{A}$: $\dfrac{\binom{M}{1}\binom{N-M}{2}}{\binom{N}{3}\binom{3}{1}}$ 3x $\bar{A}$: $\dfrac{\binom{M}{0}\binom{N-M}{3}}{\binom{N}{3}}$

n Ziehungen

Aus den ersten k Ereignissen A ergibt sich:

$$\frac{M}{N}\;\frac{M-1}{N-1}\;\frac{(M-2)}{N-2}\;\ldots\;\frac{(M-k+1)}{(N-k+1)}$$

Aus den folgenden n–k Ereignissen $\bar{A}$ ergibt sich:

$$\frac{N-M}{N-k}\;\frac{N-M-1}{N-k-1}\;\frac{N-M-2}{N-k-2}\;\ldots\;\frac{N-M-(n-k-1)}{N-(n-1)}$$

Zähler und Nenner der beiden Ausdrücke sind Produkte von aufeinanderfolgenden natürlichen Zahlen. Solche Produkte kann man als Fakultäten schreiben. Aber die Fakultäten sind nicht vollständig, da die Produkte nicht bis zum Faktor 1 fallen, sondern an unterschiedlichen Stellen vorher enden. Deshalb müssen die Werte ausgeglichen werden:

1 Zähler: $\dfrac{M!}{(M-k)!}$ 2. Zähler: $\dfrac{(N-M)!}{(N-M-(n-k))!}$ Nenner: $\dfrac{N!}{(N-n)!}$

$$\boxed{\frac{\frac{M!}{(M-k)!}\;\frac{(N-M)!}{(N-M-(n-k))!}}{\frac{N!}{(N-n)!}}}$$

Wahrscheinlichkeit eines Pfades beim „Ziehen ohne Zurücklegen“

Zähler und Nenner kann man in Binomialkoeffizienten umschreiben, wobei auch da Fakultätsausdrücke fehlen:

$$\frac{M!}{(M-k)!} = \frac{M!}{(M-k)!}\;\frac{k!}{k!} = \binom{M}{k}\,k! \qquad \frac{(N-M)!}{(N-M-(n-k))!} = \frac{(N-M)!}{(N-M-(n-k))!}\;\frac{(n-k)!}{(n-k)!} = \binom{N-M}{n-k}\,(n-k)!$$

$$\frac{N!}{(N-n)!} = \frac{N!}{(N-n)!}\;\frac{n!}{n!} = \binom{N}{n}\,n!$$

Mit dieser Umformung lassen sich die Ausdrücke in folgende Formel umwandeln:

$$\frac{\binom{M}{k}\,k!\,\binom{N-M}{n-k}\,(n-k)!}{\binom{N}{n}\,n!} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}\;\frac{n!}{k!\,(n-k)!}} = \boxed{\frac{\binom{s}{k}\binom{N-M}{n-k}}{\binom{N}{n}\binom{n}{k}}}$$

Für ungeordnete Stichproben gibt es im Baumdiagramm zu jedem Ereignis mehrere Pfade.

Das ist die Wahrscheinlichkeit **eines** Pfades bei:

- N Kugeln davon
- M Kugeln von der gesuchten Farbe
- N – M Kugeln von anderer Farbe
- mit n Ziehung ohne Zurücklegen
- bei der k Kugeln der gesuchten Farbe gezogen werden.

Die Anzahl aller Pfade ergibt sich aus $\binom{n}{k}$

2.1. Ohne Zurücklegen, mit Reihenfolge

2.1.1. Zwei Unterscheidungsmerkmale

In einer Urne sind
N Kugeln
M rote Kugeln
N – M schwarze Kugeln

Es werden
n Kugeln gezogen
k Kugeln davon sollen rot sein

Für geordnete Stichproben gibt es im Baumdiagramm zu jedem Ereignis nur einen Pfad.

Kombinatorik: „Variation ohne Wiederholung“ $\frac{N!}{(N-n)!}$ N nPr (N – n)

Aus der Kombinatorik lässt sich folgende Formel herleiten:

Gesamtanzahl: $\frac{N!}{(N-n)!}$	rote Kugel: $\frac{M!}{(M-k)!}$	schwarze Kugel: $\frac{(N-M)!}{(N-M-(n-k))!}$
Anzahl der Möglichkeiten • aus N Kugeln • n Kugeln • mit Wiederholung auszuwählen	Anzahl der Möglichkeiten • aus M roten Kugeln • k Kugeln • mit Wiederholung auszuwählen	Anzahl der Möglichkeiten • aus N – M schwarzen Kugeln • n – k Kugeln • mit Wiederholung auszuwählen

da k rote Kugeln (nur) gezogen wurden, müssen n–k Kugeln schwarz sein:

$$\mathbf{P} = \frac{\frac{M!}{(M-k)!} \quad \frac{(N-M)!}{(N-M-n+k)!}}{\frac{N!}{(N-n)!}}$$

GTR

M nPr (M–k) • (N – M) nPr (N – M – n – k)
N nPr (N – n)

$$\frac{M!}{(M-k)!} = \binom{M}{M-k} k! = \binom{M}{k} k!$$

$$\frac{(N-M)!}{(N-M-n+k)!} = \binom{N-M}{N-M-n+k} (n-k)! = \binom{N-M}{n-k} (n-k)!$$

$$\frac{N!}{(N-n)!} = \binom{N}{N-n} n! = \binom{N}{n} n!$$

$$\frac{\binom{M}{k} k! \binom{N-M}{n-k} (n-k)!}{\binom{N}{n} n!} = \boxed{\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \cdot \frac{1}{\binom{n}{k}}}$$

GTR

M nCr k • (N – M) nCr (n – k)
N nCr n n nCr k

Das ist Hypergeometrische Verteilung (ohne Zurücklegen) dividiert durch die Anzahl der Pfade, da wegen der Reihenfolge jeder Pfad einzeln betrachtet werden muss.

2.1.2. Drei Unterscheidungsmerkmale

gezogene rote Kugeln : r — Anzahl rote Kugeln : R
gezogene weiße Kugeln : w — Anzahl weiße Kugeln : W
gezogene blaue Kugeln : b — Anzahl blaue Kugeln : B

$r + w + b = n$ $\qquad$ $R + W + B = N$

Gesamtanzahl: $\frac{N!}{(N-n)!}$ $\qquad$ rote Kugeln: $\frac{R!}{(R-r)!}$ $\qquad$ nicht rote Kugeln: $\frac{(N-R)!}{(N-R-(n-r))!}$

Formel zur Unterscheidung von zwei Farben. Zunächst wird nur unterschieden zwischen „rot“ und „nicht rot“. Das ist eine Unterscheidung in zwei Farben.

$$\mathbf{P} = \frac{\frac{R!}{(R-r)!} \quad \frac{(N-R)!}{(N-R-n+r)!}}{\frac{N!}{(N-n)!}}$$

$$\frac{R!}{(R-r)!} = \frac{R! \quad r!}{(R-r)! \quad r!} = \binom{R}{R-r} r! = \binom{R}{r} r! \qquad \frac{(N-R)!}{(N-R-n+k)!} = \binom{N-R}{N-R-n+r}(n-r)! = \binom{N-R}{n-r}(n-r)!$$

$$\frac{N!}{(N-n)!} = \binom{N}{N-n} n! = \binom{N}{n} n!$$

Aus $N - R = W + B$ nicht rote Kugeln sollen w weiße und b blaue gezogen werden, b blaue sind damit auch b „nicht weiße“ von der verbleibenden Gesamtanzahl. Damit ist die Anzahl der „nicht roten“ Kugeln wieder in zwei Farben getrennt.

$$\frac{\binom{W}{w} w! \binom{N-R-W}{n-r-w}(n-r-w)!}{\binom{N-R}{n-r}(n-r)!} = \frac{\binom{W}{w} w! \binom{B}{b} b!}{\binom{N-R}{n-r}(n-r)!} \qquad \frac{\frac{W!}{(W-w)!} \quad \frac{B!}{(B-b)!}}{\frac{(W+B)!}{(W+B-(w+b))!}} = \frac{\frac{W!}{(W-w)!} \quad \frac{B!}{(B-b)!}}{\frac{(N-R)!}{(N-R-n+r)!}}$$

Die Multiplikation mit dem Faktor für die roten Kugeln liefert folgende Formel:

$$\frac{\binom{R}{r}\binom{N-R}{n-r} r! \ (n-r)!}{\binom{N}{n} \quad n!} \cdot \frac{\binom{W}{w}\binom{B}{b} w! \ b!}{\binom{N-R}{n-r}(n-r)!} \qquad \frac{\frac{R!}{(R-r)!} \quad \frac{(N-R)!}{(N-R-n+r)!}}{\frac{N!}{(N-n)!}} \cdot \frac{\frac{W!}{(W-w)!} \quad \frac{B!}{(B-b)!}}{\frac{(N-R)!}{(N-R-n+r)!}}$$

$$= \boxed{\frac{\binom{R}{r}\binom{W}{w}\binom{B}{b} \ r! \ w! \ b!}{\binom{N}{n} \quad n!}} \qquad = \boxed{\frac{\frac{R!}{(R-r)!} \quad \frac{W!}{(W-w)!} \quad \frac{B!}{(B-b)!}}{\frac{N!}{(N-n)!}}}$$

2.1.3. Vier Unterscheidungsmerkmale

Sind R,G,B,W rote, grüne, blaue und weiße Kugeln (zusammen N) vorhanden und wir ziehen r, g, b, w rote, grüne, blaue und weiße (insgesamt n) OHNE ZURÜCKLEGEN, so ist die Wahrscheinlichkeit dafür

$$\boxed{\mathbf{P} = \frac{\binom{R}{r}\binom{G}{g}\binom{B}{b}\binom{W}{w}}{\binom{N}{n}} \cdot \frac{r! \ g! \ b! \ w!}{n!} = \frac{\frac{R!}{(R-r)!} \quad \frac{G!}{(G-g)!} \quad \frac{B!}{(B-b)!} \quad \frac{W!}{(W-w)!}}{\frac{N!}{(N-n)!}}}$$

2.2. Ohne Zurücklegen, ohne Reihenfolge

2.2.1. Zwei Unterscheidungsmerkmale

In einer Urne sind
N Kugeln
M rote Kugeln
N – M schwarze Kugeln

Es werden
n Kugeln gezogen
k Kugeln davon sollen rot sein

Gesamtanzahl: $\binom{N}{n}$	rote Kugel: $\binom{M}{k}$	schwarze Kugel: $\binom{N-M}{n-k}$
Anzahl der Möglichkeiten • aus N Kugeln • n Kugeln • mit Wiederholung auszuwählen	Anzahl der Möglichkeiten • aus M roten Kugeln • k Kugeln • mit Wiederholung auszuwählen	Anzahl der Möglichkeiten • aus N – M roten Kugeln • n – k Kugeln • mit Wiederholung auszuwählen

$$\mathbf{P} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

GTR

M nCr k • (N – M) nCr (n – k)
N nCr n

Hypergeometrische Verteilung

Die Anzahl aller Pfade ist die gleiche, wie bei der Binomialverteilung: $\binom{n}{k}$

Die Pfadanzahl ist in der Formel bereits enthalten und darf nicht zusätzlich multipliziert werden.

2.2.2. Drei Unterscheidungsmerkmale

gezogene rote Kugeln : r
gezogene weiße Kugeln : w
gezogene blaue Kugeln : b

Anzahl rote Kugeln : R
Anzahl weiße Kugeln : W
Anzahl blaue Kugeln : B

$r + w + b = n$

$R + W + B = N$

Es sollen r rote und n – r nicht rote gezogen werden
Damit werden insgesamt n Kugeln angeordnet, die die Anzahl der Möglichkeiten angibt.

$$\frac{\binom{R}{r}\binom{N-R}{n-r}}{\binom{N}{n}}$$

Aus N – R = W + B nicht rote Kugeln sollen w weiße und b blaue gezogen werden,
b blaue sind damit auch b „nicht weiße“ von der verbleibenden Gesamtanzahl.

$$\frac{\binom{W}{w}\binom{N-R-W}{n-r-w}}{\binom{N-R}{n-r}} = \frac{\binom{W}{w}\binom{B}{b}}{\binom{N-R}{n-r}}$$

Die Multiplikation der beiden Wahrscheinlichkeiten liefert folgende Formel:

$$\frac{\binom{R}{r}\binom{N-R}{n-r}}{\binom{N}{n}} \cdot \frac{\binom{W}{w}\binom{B}{b}}{\binom{N-R}{n-r}} = \frac{\binom{R}{r}\binom{W}{w}\binom{B}{b}}{\binom{N}{n}}$$

2.2.3. Vier Unterscheidungsmerkmale

Sind R,G,B,W rote, grüne, blaue und weiße Kugeln (zusammen N) vorhanden und wir ziehen r, g, b, w rote, grüne, blaue und weiße (insgesamt n) OHNE ZURÜCKLEGEN, so ist die Wahrscheinlichkeit dafür

$$\boxed{\mathbf{P} = \frac{\binom{R}{r}\binom{G}{g}\binom{B}{b}\binom{W}{w}}{\binom{N}{n}}}$$

$$\frac{\binom{R}{r}\binom{G}{g}\binom{B}{b}\binom{W}{w}}{\binom{N}{n}} \qquad \frac{\frac{R!}{(R-r)!\;r!}\;\frac{G!}{(G-g)!\;g!}\;\frac{B!}{(B-b)!\;b!}\;\frac{W!}{(W-w)!\;w!}}{\frac{N!}{(N-n)!\;n!}}$$

$$\frac{R!\quad G!\quad B!\quad W!}{(R-r)!\;(G-g)!\;(B-b)!\;(W-w)!} \qquad \frac{(N-n)!}{N!} \qquad \frac{n!}{r!\;g!\;b!\;w!}$$

$$\frac{R!}{(R-r)!} \qquad \frac{1\,2\,3\ldots(R-r)\,(R-r+1)\ldots R}{1\,2\,3\ldots(R-r)} \qquad (R-r+1)\ldots R \qquad \frac{1\,2\,3\ldots(N-n)}{1\,2\,3\ldots(N-n)\,(N-n+1)\ldots N}$$

$$\frac{G!}{(G-g)!} \qquad \frac{1\,2\,3\ldots(G-g)\,(G-g+1)\ldots G}{1\,2\,3\ldots(G-g)} \qquad (G-g+1)\ldots G$$

$$\frac{B!}{(B-b)!} \qquad \frac{1\,2\,3\ldots(B-b)\,(B-b+1)\ldots B}{1\,2\,3\ldots(B-b)} \qquad (B-b+1)\ldots B$$

$$\frac{W!}{(W-w)!} \qquad \frac{1\,2\,3\ldots(W-w)\,(W-w+1)\ldots W}{1\,2\,3\ldots(W-w)} \qquad (W-w+1)\ldots W$$

$$\frac{R\ldots(R-r+1)\qquad B\ \ldots(B-b+1)\qquad G\ \ldots(G-g+1)\qquad W\ \ldots(W-w+1)}{N\ldots(N-r+1)\quad (N-r)\ldots(N-r-b+1)\quad (N-r-b)\ldots(N-r-b-g+1)\quad (N-r-b-g)\ldots(N-r-b-g-w+1)}$$

2.3. Beziehungen zwischen den Verteilungsfunktionen

In einer Urne befinden sich N = 13 Kugeln: S= 6 schwarze, R = 4 rote und G= 3 grüne Kugeln.
Wir ziehen n = 9 mal ohne Zurücklegen.
Wie groß ist die Wahrscheinlichkeit, dass dabei
s = 4 mal die schwarze Kugel, r= 3 mal die rote Kugel und g= 2 mal die grüne Kugel kommt ?

Kombinatorik: „Kombination ohne Wiederholung“ „ohne Reihenfolge ohne Wiederholung“	**Hypergeometrische Verteilung.**

$$\frac{\binom{S}{s}\binom{R}{r}\binom{G}{g}}{\binom{N}{n}} \qquad \frac{\binom{6}{4}\binom{4}{3}\binom{3}{2}}{\binom{13}{9}} = \frac{\frac{6!}{2!\;4!}\;\frac{4!}{1!\;3!}\;\frac{3!}{1!\;2!}}{\frac{13!}{4!\;9!}} = \frac{6!}{2!\;4!}\;\frac{4!}{1!\;3!}\;\frac{3!}{1!\;2!}\;\frac{4!\;9!}{13!} = \frac{36}{134}$$

im Taschenrechner : $\dfrac{6 \text{ nCr } 4 \;\bullet\; 4 \text{ nCr } 3 \;\bullet\; 3 \text{ nCr } 2}{13 \text{ nCr } 9}$

Aus der Permutation folgt: 9 Kugeln lassen sich in 9! Möglichkeiten anordnen.

davon sind :	4 schwarze nicht unterscheidbar	4!
davon sind :	3 rote nicht unterscheidbar	3!
davon sind :	2 grüne nicht unterscheidbar	2!

und es ergeben sich insgesamt $\dfrac{9!}{4!\;3!\;2!}$ verschiedene Anordnungsmöglichkeiten = Pfade.

Diese vier Faktoren treten in der Hypergeometrischen Verteilung an verschiedenen Stelle auf.
Sie sind jeweils der Bestandteil des Binomialkoeffizienten, der die Anzahl der ausgewählten Kugeln angibt.

$$\frac{6!}{2!\;\mathbf{4!}}\;\frac{4!}{1!\;\mathbf{3!}}\;\frac{3!}{1!\;\mathbf{2!}}\;\frac{4!\;\mathbf{9!}}{13!}$$

Fasst man die Fakultäten der Pfadfaktoren zusammen ergibt sich:

$$\frac{6!}{2!}\;\frac{4!}{1!}\;\frac{3!}{1!}\;\frac{4!}{13!} \qquad \frac{9!}{4!\;3!\;2!}$$

Der restliche Ausdruck der nicht mehr die Pfadanzahl erhält ergibt in ausgeschriebener Form:

$$\frac{6\;5\;4\;3\;\cancel{2\;1}}{\cancel{2\;1}}\;\frac{4\;3\;2\;1}{1}\;\frac{3\;2\;1}{1}\;\frac{\cancel{4\;3\;2\;1}}{13\;12\;11\;10\;.\;.\;.\;5\;\cancel{4\;3\;2\;1}}$$

$$\frac{6\;\;5\;\;4\;\;3}{13\;12\;11\;10}\;\frac{4\;\;3\;\;2}{9\;\;8\;\;7}\;\frac{3\;\;2}{6\;\;5}$$

Da es auf die Reihenfolge nicht ankommt, kann man die Wahrscheinlichkeit für eine Anordnung ermitteln und dann mit der Anzahl der Pfade multiplizieren.

$$\overset{\text{s s s s}}{\frac{6\;\;5\;\;4\;\;3}{13\;12\;11\;10}}\;\overset{\text{r r r}}{\frac{4\;\;3\;\;2}{9\;\;8\;\;7}}\;\overset{\text{g g}}{\frac{3\;\;2}{6\;\;5}} \qquad \overset{\text{Anzahl der Pfade}}{\frac{9!}{4!\;3!\;2!}}$$

Für die 1. schwarze Kugel gibt es : 6 Möglichkeiten von 13
Für die 2. schwarze Kugel gibt es : 5 Möglichkeiten von 12
Für die 3. schwarze Kugel gibt es : 4 Möglichkeiten von 11
Für die 4. schwarze Kugel gibt es : 3 Möglichkeiten von 10

Für die 1. rote Kugel gibt es : 4 Möglichkeiten von 9
Für die 2. rote Kugel gibt es : 3 Möglichkeiten von 8
Für die 3. rote Kugel gibt es : 2 Möglichkeiten von 7

Für die 1. grüne Kugel gibt es : 3 Möglichkeiten von 6
Für die 2. grüne Kugel gibt es : 2 Möglichkeiten von 5

Die Anzahl der Pfade ist identisch mit der Anzahl der Pfade in der verallgemeinerten Binomialverteilung.

Kombinatorik: „Variation ohne Wiederholung“
„mit Reihenfolge ohne Wiederholung“ **(kein Name)**

Wahrscheinlichkeit eines Pfades

$$\frac{\overset{s}{6}\ \overset{s}{5}\ \overset{s}{4}\ \overset{s}{3}\quad \overset{r}{4}\ \overset{r}{3}\ \overset{r}{2}\quad \overset{g}{3}\ \overset{g}{2}}{13\ 12\ 11\ 10\quad 9\ 8\ 7\quad 6\ 5}$$

$$\frac{\frac{6!}{(6-4)!}\ \frac{4!}{(4-3)!}\ \frac{3!}{(3-2)!}}{\frac{13!}{(13-9)!}}$$

Anzahl der Pfade

$$\frac{9!}{4!\ 3!\ 2!} = \frac{9!}{4!\ 5!}\ \frac{5!}{2!\ 3!} = \frac{n!}{s!\ r!\ g!}$$

$$\binom{9}{5}\binom{5}{3} \qquad \binom{n}{s}\binom{n-s}{r}$$

$$\frac{\frac{S!}{(S-s)!}\ \frac{R!}{(R-r)!}\ \frac{G!}{(G-g)!}}{\frac{N!}{(N-n)!}} \qquad \frac{n!}{s!\ r!\ g!}$$

Das ist die Wahrscheinlichkeit für ***einen Pfad*** (= mit Reihenfolge) mit s schwarzen, r roten und g grünen Kugeln
multipliziert mit der Anzahl aller Pfade ergibt das wieder die Wahrscheinlichkeit ohne Zurücklegen und ohne Reihenfolge

im Taschenrechner : $\frac{6\ nPr\ 4\ \bullet\ 4\ nPr\ 3\ \bullet\ 3\ nPr\ 2}{13\ nPr\ 9}$ 9 nCr 5 • 5 nCr 3

Inwieweit diese Formel einfacher ist, als die ursprüngliche Formel der hypergeometrischen Verteilung sei mal dahingestellt. Da das Produkt der hinteren beiden Binomialkoeffizienten die Anzahl der Pfade darstellt, ergibt der erste Ausdruck mit den nPr die **Wahrscheinlichkeit jeweils eines Pfades**.Dafür ist der erste Ausdruck gut zu gebrauchen, an Stelle der Division der Hypergeometrischen Verteilung durch die Anzahl der Pfade.

Kombinatorik: „Variation mit Wiederholung“
„mit Reihenfolge mit Wiederholung“ **(kein Name)**

Für die Wahrscheinlichkeit der schwarzen Kugeln erhält man bei „mit zurücklegen“:

$$\frac{S}{S+R+G} = \frac{S}{N}$$

werden 4 schwarze von 13 Kugeln gezogen erhält man für jede einzelne Kugel $\frac{6}{13}$

und damit für die 4 schwarzen Kugeln insgesamt $\left(\frac{6}{13}\right)^4$ Kombinatorik: „Variation mit Wiederholung“

$$\frac{6}{13}\ \frac{6}{13}\ \frac{6}{13}\ \frac{6}{13} = \frac{6^4}{13^4} = \left(\frac{S}{N}\right)^s$$

Damit würde sich bei Zurücklegen und der entsprechenden Anzahl Ziehungen als Wahrscheinlichkeit für einen Pfad ergeben:

$$= \left(\frac{S}{N}\right)^s \left(\frac{R}{N}\right)^r \left(\frac{G}{N}\right)^g$$

Kombinatorik: „Kombination mit Wiederholung“
„ohne Reihenfolge mit Wiederholung“ **Binomialverteilung**

multipliziert mit der Anzahl der Pfade $\frac{n!}{s!\ r!\ g!}$ ergibt sich die Formel der Binomialverteilung

Da aber ohne Zurücklegen gezogen wird ändern sich die Faktoren in Zähler und Nenner, da jeweils eine Kugel weniger vorhanden ist und auch eine schwarze Kugel weniger vorhanden ist.

Kombinatorik: „Variation ohne Wiederholung“
„mit Reihenfolge ohne Wiederholung“ **(kein Name)**

$$\frac{6 \quad 5 \quad 4 \quad 3}{13 \quad 12 \quad 11 \quad 10} \quad \frac{4 \quad 3 \quad 2}{9 \quad 8 \quad 7} \quad \frac{3 \quad 2}{6 \quad 5}$$

$$\frac{6 \text{ nPr } 4 \quad \bullet \quad 4 \text{ nPr } 3 \quad \bullet \quad 3 \text{ nPr } 2}{13 \text{ nPr } 9 \quad \bullet \quad 9 \text{ nPr } 6 \quad \bullet \quad 6 \text{ nPr } 4}$$

$$\frac{\dfrac{6! \quad \bullet \quad 4! \quad \bullet \quad 3!}{(6\text{-}4)! \bullet (4\text{-}3)! \bullet (3\text{-}2)!}}{\dfrac{13! \quad \bullet \quad 9! \quad \bullet \quad 6!}{(13\text{-}9)! \bullet (9\text{-}6)! \bullet (6\text{-}4)!}} \qquad \frac{\dfrac{S!}{(S-s)!} \quad \dfrac{R!}{(R-r)!} \quad \dfrac{G!}{(G-g)!}}{\dfrac{N!}{(N-n)!} \quad \dfrac{(N\text{-}n)!}{(N-n-S)!} \quad \dfrac{(N\text{-}n\text{-}r)!}{(N-n-S-R)!}}$$

N = 13 S= 6 schwarze, R = 4 rote G = 3 grüne Kugeln.
n = 9 s = 4 schwarze Kugel, r= 3 rote Kugel g = 2 grüne Kugel

Variation bedeutet „mit Beachtung der Reihenfolge“. Damit ist die Wahrscheinlichkeit, die bei „Variation“ angegeben wird immer die Wahrscheinlichkeit eines Pfades. Arbeitet man „ohne Beachtung der Reihenfolge“ werden alle Pfade mit den gleichen Farb – (Eigenschafts)zusammen- stellungen zu einem Ergebnis zusammengefasst. Deshalb ist dann die Wahrscheinlichkeit der „Variation“ mit der Anzahl der Pfade zu multiplizieren und man erhält die Wahrscheinlichkeit der „Kombination“ (ohne Beachtung der Reihenfolge).

Die Anzahl der Pfade berechnet sich aus der Permutation: Es werden insgesamt 9 Elemente gezogen, deshalb gibt es 9! verschiedene Anordnungen.

Innerhalb dieser Anordnung lassen sich die Vertauschungen der 4 schwarzen Kugeln nicht unterscheiden, nach außen wirken sie wie eine Anordnung. $\frac{9!}{4!}$

Innerhalb dieser Anordnung lassen sich aber auch die Vertauschungen der 3 roten Kugeln nicht unterscheiden, nach außen wirken sie wie eine Anordnung. $\frac{9!}{4!\,3!}$

ebenso für die zwei grünen Kugeln, die gezogen werden sollen: $\frac{9!}{4!\,3!\,2!}$

Kombinatorik: „Permutation mit Wiederholung“

$$\frac{9!}{4!\;3!\;2!} = \frac{9!}{4!\;5!}\;\frac{5!}{2!\;3!} = \frac{n\,!}{s!\;r!\;g!} = \binom{9}{5}\binom{5}{3} = \binom{n}{s}\binom{n\text{-}s}{r}$$

$$\frac{\dfrac{S!}{(S-s)!} \quad \dfrac{R!}{(R-r)!} \quad \dfrac{G!}{(G-g)!}}{\dfrac{N!}{(N-n)!} \quad \dfrac{(N\text{-}n)!}{(N-n-S)!} \quad \dfrac{(N\text{-}n\text{-}r)!}{(N-n-S-R)!}} \qquad \frac{n\,!}{s!\;r!\;g!}$$

<u>im Taschenrechner :</u> $\frac{6 \text{ nPr } 4 \quad \bullet \quad 4 \text{ nPr } 3 \quad \bullet \quad 3 \text{ nPr } 2}{13 \text{ nPr } 9 \quad \bullet \quad 9 \text{ nPr } 6 \quad \bullet \quad 6 \text{ nPr } 4}$ 9 nCr 5 • 5 nCr 3

2.4. Rechenschema für alle Fälle

						Zeilenprodukt	Gesamtanzahl	
Eigenschaft	1	2	3	4	5			
Gesamtanzahl	6	8	4	0	0		18	
gezogene Anzahl	2	3	2	0	0		7	
Wahrscheinlichkeit	$\frac{6}{18}$	$\frac{8}{18}$	$\frac{4}{18}$					
Wahrscheinlichkeit hoch Anzahl	$\left(\frac{6}{18}\right)^2$	$\left(\frac{8}{18}\right)^3$	$\left(\frac{4}{18}\right)^2$			0,000482		
Binomialkoeffizient	$\binom{6}{2}$	$\binom{8}{3}$	$\binom{4}{2}$			5040	31 824	$\binom{18}{7}$
Fakultät	2!	3!	2!			24	5 040	7!
obere Faktorielle $\frac{n!}{(n-k)!}$	$\frac{6!}{(6-2)!}$	$\frac{8!}{(8-3)!}$	$\frac{4!}{(4-2)!}$			120 960	160 392 960	$\frac{18!}{(18-7)!}$

<u>Welche Werte werden an welcher Stelle gebraucht ?</u>

ohne Reihenfolge ; mit Zurücklegen	$\frac{5\,040}{24} \cdot 0{,}000482$	Binomialverteilung
mit Reihenfolge ; mit Zurücklegen	0,000482	ein Pfad der Binomialverteilung
ohne Reihenfolge ; ohne Zurücklegen	$\frac{5\,040}{31\,824}$	Hypergeometrische Verteilung
mit Reihenfolge ; ohne Zurücklegen	$\frac{120\,960}{160\,392\,960}$	ein Pfad der Hypergeometrische Verteilung
	$= \frac{24}{5\,040} \cdot \frac{5\,040}{31\,824}$	Hypergeometrische Verteilung dividiert durch die Anzahl der Pfade

3. Binomialverteilung

Aufgaben dieser Art behandeln die sogenannte **Binomialverteilung** von Wahrscheinlichkeiten. Zur Lösung wird das **Urnen-Modell** verwendet.

In einem Gefäß (Urne) befindet sich eine bestimmte Anzahl von Kugeln. Die Gesamtzahl der Kugeln ist die Menge **n**. Jede einzelne Kugel ist Element der Menge **n**. Jede Kugel besitzt ein bestimmtes Merkmal. Aus der Anzahl der Kugeln (Elemente) mit demselben Merkmal ergibt sich die Häufigkeit und daraus wiederum die Wahrscheinlichkeit für das Auftreten des Merkmals.

Beispiel 1

Befinden sich im Gefäß 36 schwarze und 12 weiße Kugeln, dann treten mit einer Wahrscheinlichkeit von 0,75 oder 75 % schwarze und mit 0,25 oder 25 % weiße Kugeln auf.

Bei diesen Aufgaben gibt es immer 2 mögliche Merkmale bzw. deren Wahrscheinlichkeiten. Sollten mehr als 2 Merkmale auftreten, dann wird im Urnen-Modell ein einzelnes Merkmal für sich genommen und die restlichen Merkmale zu einem zweiten Merkmal zusammengefasst.

Die Wahrscheinlichkeiten für beide Merkmale werden mit **p** und **q** bezeichnet, wobei **q** auch **Gegenwahrscheinlichkeit** heißt. Daher gilt:

Formel 1

$$\boxed{\begin{array}{l} p=1-q \\ q=1-p \end{array}}$$

Wahrscheinlichkeiten haben immer einen Wert zwischen 0 und 1 oder in Prozent ausgedrückt zwischen 0 % und 100 %. Es genügt also **p** zu berechnen, die Differenz zu 1 ist dann die Gegenwahrscheinlichkeit **q**. Mindestens eine der beiden Wahrscheinlichkeiten ist bei diesen Aufgaben immer gegeben, sonst können sie nicht gelöst werden. Wenn mehr als 2 Merkmale vorhanden sind, werden alle außer dem **p** - Merkmal zur Gegenwahrscheinlichkeit **q** der übrigen Merkmale zusammengefasst.

Mit dem Gefäß-Modell können viele praktische Aufgaben gelöst werden. Dabei wird ein tatsächlicher Sachverhalt für das Gefäß-Modell vereinfacht. Man nennt das auch: **abstrahieren**. Man muss immer feststellen, welche tatsächlichen Gegenstände die Gesamtmenge **n** darstellen und welche Merkmale für **p** und **q** stehen. Im weiteren Sinne wird statt Merkmal auch der Begriff **Ereignis** verwendet. Das bedeutet soviel wie: ein bestimmtes Merkmal trifft zu oder ereignet sich. So werden Merkmale und Ereignisse sachlich gleichbedeutend verwendet.

Zur Lösung der Aufgaben wird gedanklich folgendes Experiment (Versuch) durchgeführt. Aus der Gesamtmenge **n** (aus dem Gefäß) wird eine festgelegte Anzahl Kugeln herausgenommen und es wird überprüft, welches Merkmal **p** oder **q** diese Kugeln tragen. Entsprechend kann man auch sagen, es wird geprüft, welche Ereignisse **p** oder **q** beim Herausnehmen eingetreten sind. Auf diese Weise werden Sachverhalte und Vorgänge der Wirklichkeit am Urnen-Modell vollzogen oder **simuliert**. Deshalb ist ein Modell auch eine Simulation der Wirklichkeit.

Auf welche Weise kann überhaupt eine Kugel aus dem Gefäß genommen werden?

a) einzeln (mit / ohne Beachtung der Reihenfolge)
b) mehrere auf einmal

Außerdem gibt es noch zwei Möglichkeiten, wenn der Versuch mehrmals durchgeführt wird. Dann kann man die herausgenommenen Kugeln bei jedem Versuch

c) wieder in das Gefäß zurücklegen
d) nicht wieder zurücklegen.

Jede dieser Arten a), b), c), d) stehen für die Simulation verschiedener Wahrscheinlichkeiten und man muss bei der Lösung einer Aufgabe prüfen, welche Art der tatsächlichen Situation entspricht.

Schaut man sich die hier gestellten Aufgaben an, erkennt man, dass **Binomialverteilung**

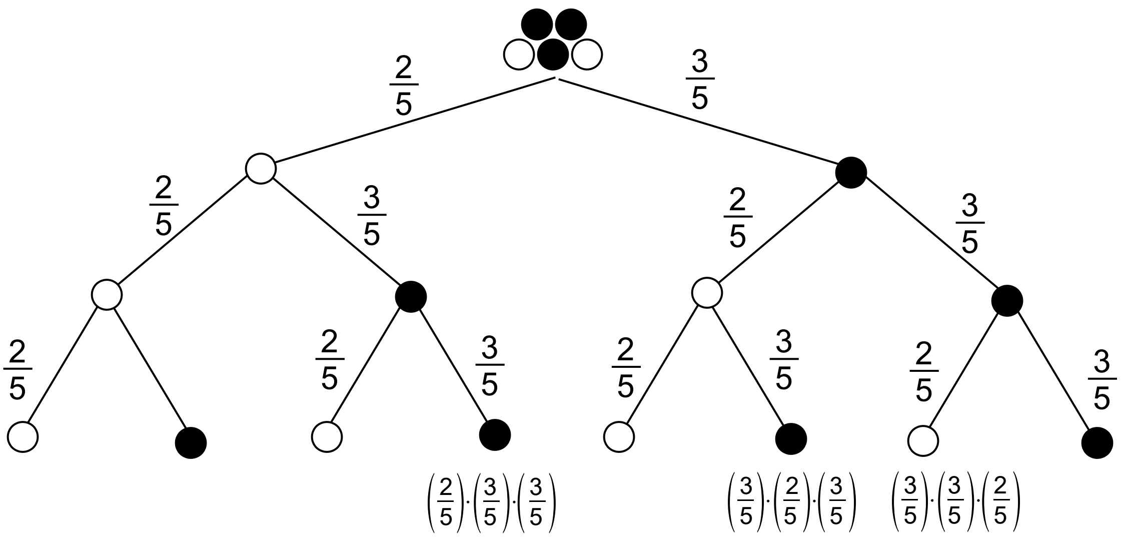
- **Versuche mit Zurücklegen** der Kugeln und
- **ohne Beachtung der Reihenfolge** ist.

So wird bei Aufgabe 01 sozusagen aus dem Gefäß "Sitzung" jedesmal neu die Anwesenheit geprüft. Auch bei den anderen Teilaufgaben handelt es sich um Versuche mit Zurücklegen.

Alles wäre ganz einfach, wenn man aus dem Gefäß eine Anzahl Kugeln zieht und aus deren Merkmalen ihre Wahrscheinlichkeiten bestimmen könnte. Das eigentliche Problem, das die Berechnung schwierig macht, ist die Tatsache, dass die Kugeln mit gleichen Merkmalen auf eine bestimmte Art im Gefäß **verteilt** sind, und gerade diese Verteilung muss gefunden werden.

Angenommen, in einem Gefäß befinden sich 5 Kugeln, davon 3 schwarze und 2 weiße. Bei jedem Versuch wird eine Kugel herausgenommen, ihr Merkmal (schwarz / weiß) geprüft und die Kugel wieder zurückgelegt. Man führt 3 solche Versuche durch. Wie groß ist dabei die Wahrscheinlichkeit, dass man genau 2 schwarze Kugeln gezogen hat? Den Ablauf der 3 Versuche macht man sich mit einem Baumdiagramm deutlich.

Beispiel 2



Mit der **Pfadregel** ermittelt man die Wahrscheinlichkeiten, die man zusammenfassen kann: $\left(\frac{2}{5}\right)\cdot\left(\frac{3}{5}\right)^2$

Das entspricht dem allgemeinen Term für die Wahrscheinlichkeit eines Pfades:

Formel 2 $\boxed{p^k \cdot q^{n-k}}$

wobei **n** die Gesamtmenge der Kugeln ist und **k** die Anzahl der gezogenen Kugeln mit dem Merkmal **p**.

Damit hat man aber die gefragte Wahrscheinlichkeit (genau 2 schwarze Kugeln bei 3 Versuchen) noch nicht gefunden. Warum nicht? Genau betrachtet, ist es bei jedem Versuch egal, welche der 3 schwarzen Kugeln man zieht. Das wirkt sich aber auf die Wahrscheinlichkeit aus. Man muss also zusätzlich noch fragen: wie viele Möglichkeiten gibt es, aus einer Menge von 3 schwarzen Kugeln 2 schwarze zu ziehen. Diese Frage kann man auch umformulieren: Auf wie viele Arten lassen sich 2 schwarze Kugeln auf 3 (freie) Plätze verteilen? Diese Rechnung wird mit dem **Binomial-Koeffizienten** ausgeführt.

Formel 3 $\boxed{\binom{n}{k}=\frac{n!}{(n-k)!k!}}$

n ist dabei die Anzahl der freien Plätze, **k** die Anzahl der gezogenen Kugeln.

Bemerkung: Aus der Formel wird ersichtlich, dass sowohl im Zähler wie im Nenner **Permutationen** stehen. Das sollte man vor allem beachten bei der Bewertung des **n** . Hier handelt es sich nicht mehr um das ursprüngliche **n** = Gesamtmenge der Kugeln im Gefäß, sondern bereits um **n** = k - Permutationen aus dieser Gesamtmenge. Statt der einfachen Gesamtmenge aller Elemente wird hier in der Formel für **n** die Anzahl der durchgeführten Versuche genommen. Die Anzahl der Versuche weicht aber in der Regel vom Wert der Gesamtmenge aller Elemente ab. Leider wird dieser Unterschied in der Schreibweise wenig deutlich gemacht.

Nun kann man die gefragte Wahrscheinlichkeit ermitteln.

Formel 4 $\boxed{P(X=k)=\binom{n}{k}p^k\cdot q^{n-k}}$

Jacob BERNOULLI (1654 - 1705) hat diese Formel gefunden, weshalb man auch bei solchen Aufgaben manchmal von BERNOULLI-Kette oder bei der Binomialverteilung von BERNOULLI-Formel spricht.

Für das Beispiel gilt $P(X=2)=\binom{3}{2}\cdot\left(\frac{3}{5}\right)^2\cdot\left(\frac{2}{5}\right)^1=0{,}432$

Mit einer Wahrscheinlichkeit von 43,2 % werden bei drei Versuchen 2 schwarze Kugeln gezogen.

Für die Berechnung des Binomial-Koeffizienten gibt es am Taschenrechner meistens eine eigene Taste, die z.B. das Symbol: **nCr** hat.

Auch **k** muss in jeder Aufgabe vorgegeben sein. Dieses k stellt nicht einfach nur eine Zahl oder Anzahl von Elementen dar, sondern im weiteren Sinne handelt es sich sozusagen um ein **k - Ereignis** mit dem **p - Merkmal**. Oft bezeichnet man das p - Merkmal als "Treffer" und das q - Ereignis als "Niete", also als "gewonnen" oder "verloren". Deshalb spricht man auch von **k Treffern** und **p** ist dann die **Trefferwahrscheinlichkeit**.

Mit der Formel 4 können eine Reihe von Aufgaben zur Binomialverteilung gelöst werden. Die gegebenen Werte sollen noch einmal mit den gängigen Bezeichnungen zusammengefasst werden.

n = Anzahl der durchgeführten, unabhängigen Versuche
k = Anzahl der Treffer
p = Trefferwahrscheinlichkeit
q = Gegenwahrscheinlichkeit

Der Wert für **n** kann theoretisch eine beliebige ganze positive Zahl > 1 annehmen. Für große **n** wird die Formel 4 allerdings unbrauchbar. Die Wahrscheinlichkeiten **p** und **q** haben stets Werte zwischen 0 und 1 und zwar so, dass beide Werte zusammen = 1 sind.

3.1. Anzahl positiver Ereignisse

Für die Werte von **k** gibt es bei den Aufgaben zur Binomialverteilung prinzipiell drei Möglichkeiten.
Folgende 3 Wahrscheinlichkeiten können gefragt sein:

a) für genau **k** Treffer
b) für mindestens **k** Treffer (= wenigstens **k** Treffer)
c) für höchstens **k** Treffer

Dafür kann man sich eine Folge (Kette) von **n** Versuchen vorstellen. Nun kann man die möglichen Trefferzahlen abzählen. Angenommen **n** = 10. Bei diesen 10 Versuchen ist es immerhin möglich, dass man in keinem der 10 Versuche einen Treffer hat. Ebenso ist es möglich, dass man 10 Treffer hat, nämlich in jedem Versuch einen. Die Wahrscheinlichkeit dafür würde man Glück nennen. Natürlich sind zwischen 0 und 10 alle ganzen Zahlen als Treffer möglich. Die Übersicht über alle möglichen Treffer bei 10 Versuchen bringt man durch den Binomial-Koeffizienten zum Ausdruck. Dabei entspricht der k - Wert den möglichen Treffern.

$$\binom{10}{0} \quad \binom{10}{1} \quad \binom{10}{2} \quad \binom{10}{3} \quad \binom{10}{4} \quad \binom{10}{5} \quad \binom{10}{6} \quad \binom{10}{7} \quad \binom{10}{8} \quad \binom{10}{9} \quad \binom{10}{10}$$

Jeder einzelne Binomial-Koeffizient kann nun in die Formel 4 eingesetzt werden. Wie man sieht, sind es 11 Binomial-Koeffizienten, obwohl es 10 Versuche sind. Das liegt daran, dass auch der Wert k = 0 mit aufgenommen werden muss. Es sind deshalb immer n + 1 Koeffizienten. Das entspricht auch der Ordnung im PASCALschen Dreieck, das auch mit dem Binom $(a + b)^{\mathbf{0}}$ und damit mit Zeile 0 beginnt, sodass die Anzahl der Koeffizienten = Zeilennummer + 1 beträgt.

Man sieht auch, dass für jede Trefferzahl eine andere Wahrscheinlichkeit gilt. Denn erstens ergeben sich für einige Binomialkoeffizienten verschiedene Werte. Und zweitens muss der Koeffizient noch jeweils mit dem Term $p^k \cdot q^{n-k}$ multipliziert werden, der für jedes einzelne **k** einen anderen Wert liefert. Auf diese Weise kommt man der Verteilung der Wahrscheinlichkeiten auf die Spur.

X = Ereignis
Y = Gegenereignis
n = Länge der Bernoulli-Kette (Anzahl der Versuche)
p = Wahrscheinlichkeit für X
q = Wahrscheinlichkeit für Y
k,m = Anzahl der Treffer

3.1.1. Genau k – Treffer

Wenn man nun die Aufgabe lösen will, die Trefferwahrscheinlichkeit für das Ereignis X mit der Wahrscheinlichkeit p **genau k Treffer** zu ermitteln, dann muss man aus der Reihe der Binomial-Koeffizienten den mit dem entsprechenden **k** - Wert herausnehmen und nach der Formel 4 berechnen.

Beispiel 3

Trefferwahrscheinlichkeit für genau 4 Treffer: P(X = 4)

$$\binom{10}{0} \quad \binom{10}{1} \quad \binom{10}{2} \quad \binom{10}{3} \quad \underset{k=4}{\binom{10}{4}} \quad \binom{10}{5} \quad \binom{10}{6} \quad \binom{10}{7} \quad \binom{10}{8} \quad \binom{10}{9} \quad \binom{10}{10}$$

Ereignis Gegenereignis

$$P_{10;p}(X=4) = \binom{10}{4} \cdot p^4 \cdot q^{(10-4)} = \binom{10}{6} \cdot q^6 \cdot p^4 = P_{10;q}(Y=6)$$

Formel 4

$$P_{n;p}(X=k) = \binom{n}{k} \cdot p^k \cdot q^{(n-k)} = \binom{n}{n-k} \cdot q^{(n-k)} \cdot p^k = P_{n;q}(Y=n-k)$$

Die auftretenden Faktoren sind aber die gleichen, wie beim Gegenereignis Y, das mit der Wahrscheinlichkeit q auftritt und dann 6 mal eintritt. Wenn bei 10 Ereignissen 4x X eintritt, dann muss 6x Y eintreten, sonst gibt es keine 10 Versuche.
Die Gleichheit der Binomialkoeffizienten lässt sich mit der Definition der Binomialkoeffizienten nachweisen. Um sich die etwas aufwendige Berechnung zu ersparen existieren für bestimmte n Werte Tabellen, bzw. kann diese Berechnung auch jeder bessere GTR ausführen. Dafür ist üblicherweise eine Funktion **binompdf** implementiert.

3.1.2. Höchstens k – Treffer

Die Trefferwahrscheinlichkeit aus Beispiel 2 für höchstens 4 Treffer ergibt sich aus allen Fällen, wo genau 4 Treffer oder weniger erzielt werden. Höchstens 4 bedeutet dabei: 4 oder weniger, also k ≤ 4 (kleiner / gleich 4).

$$\binom{10}{0} \quad \binom{10}{1} \quad \binom{10}{2} \quad \binom{10}{3} \quad \binom{10}{4} \quad \binom{10}{5} \quad \binom{10}{6} \quad \binom{10}{7} \quad \binom{10}{8} \quad \binom{10}{9} \quad \binom{10}{10}$$

k = 0 k = 1 k = 2 k = 3 k = 4

Auch diese Einzelwerte werden addiert (kumuliert) und ergeben den Wert der Binomial-**Verteilung**.

$$P_{10;p}(X \le 4) = \binom{10}{0} \cdot p^0 \cdot q^{10} + \binom{10}{1} \cdot p^1 \cdot p^9 + \binom{10}{2} \cdot p^2 \cdot q^8 + \binom{10}{3} \cdot p^3 \cdot q^7 + \binom{10}{4} \cdot p^4 \cdot q^6$$

Für das Gegenereignis vertauschen sich die Wahrscheinlichkeiten von p und q und damit auch die Anzahl k der erfolgreichen Treffer. Wenn das Ereignis X k erfolgreiche Treffer hat, muss das Ereignis Y n – k erfolgreiche Treffer haben. Wenn die Anzahl der Treffer für X nur 4 oder weniger sind, müssen es für Y 6 oder mehr sein.

$$P_{10;p}(Y \ge 6) = \binom{10}{6} \cdot q^6 \cdot p^4 + \binom{10}{7} \cdot q^7 \cdot p^3 + \binom{10}{8} \cdot q^8 \cdot p^2 + \binom{10}{9} \cdot q^9 \cdot p^1 + \binom{10}{10} \cdot q^{10} \cdot p^0$$

Die Gleichheit der Binomialkoeffizienten folgt wieder aus der Symmetrieeigenschaft der Binomialkoeffizienten.

Für die Berechnung der Summe von 0 bis k gibt es keine geschlossene Summenformel. Man muss jeden Wert einzeln berechnen und dann addieren. Deshalb stehen für diese Berechnung Tabellen zur Verfügung, bzw. erledigt das jeder bessere GTR selbständig. Die dazu implementiere Funktion hat die Bezeichnung **binomcdf**.

Damit entsteht die Formel der summierten Binomialverteilung:

Formel 5 (summierte Binomialverteilung von 0 bis k)

$$P_{n;p}(X \le k) = \sum_{m=0}^{k} \binom{n}{m} \cdot p^m \cdot q^{(n-m)}$$

3.1.3. Mindestens k – Treffer

Will man nun die Trefferwahrscheinlichkeit für mindestens 4 Treffer ermitteln, kann man aus der Reihe der Binomialkoeffizienten erkennen, in welchen Fällen mindestens 4 Treffer erreicht werden. Mindestens 4 bedeutet dabei: 4 oder mehr, also k ≥ 4 (größer / gleich 4).

$$\binom{10}{0} \quad \binom{10}{1} \quad \binom{10}{2} \quad \binom{10}{3} \quad \binom{10}{4} \quad \binom{10}{5} \quad \binom{10}{6} \quad \binom{10}{7} \quad \binom{10}{8} \quad \binom{10}{9} \quad \binom{10}{10}$$

k = 4 k = 5 k = 6 k = 7 k = 8 k = 9 k = 10

In jedem dieser Fälle hat man mindestens 4 Treffer erzielt. Nun muss man für jeden Fall P(X = k) mit der Formel 4 berechnen und **die Einzel-Ergebnisse addieren**. Weil es sich um eine Addition von Einzelwerten handelt, spricht man auch von einer **kumulativen** Rechnung. (kumulativ = lateinisch: angehäuft) Diese Binomial-Verteilung kann in Tabellen abgelesen werden, wo die Einzelrechnungen bereits kumulativ in einem Wert zusammengefasst sind.

$$P_{10;p}(X\geq 4)=\binom{10}{4}\cdot p^4\cdot q^6+\binom{10}{5}\cdot p^5\cdot p^5+\binom{10}{6}\cdot p^6\cdot q^4+\ldots+\binom{10}{10}\cdot p^{10}\cdot q^0$$

Für diese Summe ergeben sich 7 Summanden von k = 4 bis k = 10 . Was bedeutet das für das Gegenereignis:

$$\binom{10}{0}\cdot q^0\cdot p^{10}+\ldots+\binom{10}{4}\cdot q^4\cdot p^6+\binom{10}{5}\cdot q^5\cdot q^5+\binom{10}{6}\cdot q^6\cdot q^4=P_{10;q}(Y\leq 6)$$

Die Anordnung in aufsteigenden Potenzen von q. Auch hier handelt es sich um 7 Summanden mit den gleichen Potenzen der Wahrscheinlichkeiten, die Änderungen in den Binomialkoeffizienten sind aus der Symmetrie der Binomialkoeffizienten nachweisbar.

Für diese Art der Summierung existieren keine Tabellen und lassen sich auch nicht direkt mit dem GTR durchführen. Wenn man für das Ereignis X die Summe aller Wahrscheinlichkeiten ab k addieren will, reicht es aus, die Summe aller Wahrscheinlichkeiten bis k – 1 von 1 zu subtrahieren. Damit erreicht man wieder eine Summe von 0 bis k – 1, die mit der oben angegebenen Formel errechnen kann. Deshalb gilt für diese Berechnung:

$$\boxed{P_{10;p}(X\geq 4)=1-P_{10;p}(X<4)=1-P_{10;p}(X\leq 3)}$$

Formel 6 (summierte Binomialverteilung von k bis n)

$$\boxed{P_{n;p}(X\geq k)=\sum_{m=k}^{n}\binom{n}{m}\cdot p^m\cdot q^{(n-m)}=1-\sum_{m=0}^{k-1}\binom{n}{m}\cdot p^m\cdot q^{(n-m)}=1-P_{n;p}(X\leq k-1)}$$

oder man benutzt die summierte Wahrscheinlichkeit für das Gegenereignis:

$$\boxed{P_{10;p}(X\geq 4) \quad = \quad P_{10;q}(Y\leq 6)}$$

Formel 7 (summierte Binomialverteilung des Gegenereignisses von 0 bis n – k)

$$\boxed{P_{n;p}(X\geq k)=\sum_{m=k}^{n}\binom{n}{m}\cdot p^m\cdot q^{(n-m)} \quad = \quad \sum_{m=0}^{n-k}\binom{n}{m}\cdot q^m\cdot p^{(n-m)}=P_{n;q}(Y\leq n-k)}$$

Wichtig ist dabei zu beachten, dass sich nicht nur die Grenzen von k verändern, sondern auch die Reihenfolge der Wahrscheinlichkeiten. Für das Gegenereignis Y ist die Wahrscheinlichkeit q diejenige, die das Eintreffen eines positiven Ereignisses angibt.

Wenn das Ereignis X	**mindestens k mal**	auftritt
kann das Gegenereignis Y	**höchstens n – k**	mal auftreten

3.2. Beispiele zur Binomialverteilung

Die drei typischen Trefferwahrscheinlichkeiten genau k Treffer / höchstens k Treffer / mindestens k Treffer sollen an einem weiteren Beispiel verdeutlicht werden.

Bei einem Würfelspiel wird 10 mal hintereinander mit zwei Würfeln gleichzeitig geworfen.

3.2.1. Wie groß ist die Wahrscheinlichkeit, *genau* 3 Treffer zu erzielen ?

Zuerst muss die Wahrscheinlichkeit für **p** und **q** ermittelt werden. Das **p - Merkmal** kann man so formulieren: beide Würfel ergeben zusammen die Augenzahl 7. Das **q - Merkmal** lautet dann: beide Würfel ergeben zusammen eine Augenzahl ungleich 7. Wie viele mögliche Augenzahlen (und damit deren Summe) gibt es bei zwei Würfeln? (Dies berechnet man als die Anzahl der **k** - Tupel aus einer **n** - Menge; das typische Beispiel sind Zahlenschloss-Kombinationen.) Es gibt genau 36 Kombinationen von jeweils 6 Augenzahlen zweier Würfel. Nun stellt man fest, wie viele der 36 Möglichkeiten die Augenzahl 7 ergeben. Auch dafür gibt es Verfahren, aber das sicherste ist das Abzählen. Es gibt genau 6 Möglichkeiten (1 / 6) (2 / 5) (3 / 4) (4 / 3) (5 / 2) (6 / 1). Nun kann man mit einer der wichtigsten Formeln der Wahrscheinlichkeits-Rechnung die Wahrscheinlichkeit **p** ermitteln.

$$\mathbf{p} = \frac{\text{Anzahl der günstigen Möglichkeiten}}{\text{Anzahl aller Möglichkeiten}}$$

$$p = \frac{6}{36} = \frac{1}{6}; \quad q = 1 - p = \frac{30}{36} = \frac{5}{6}$$

Nun sind alle Werte für die **Formel 4** bekannt und können eingesetzt werden.

	Anwendung von **Formel 4**
	binompdf(10 , 1/6 , 3)

Die Wahrscheinlichkeit, bei 10 Versuchen genau 3 mal die Augenzahl 7 zu werfen, beträgt ca. 15,5 %.

3.2.2. Wie groß ist die Wahrscheinlichkeit, *höchstens* 3 Treffer zu erzielen?

Hier müssen die Wahrscheinlichkeiten aller möglichen Treffer summiert werden, die zwischen k = 0 und k = 3 liegen.	Oder es findet die **Formel 5** Verwendung $P_{n;p}(X \le k) = \sum_{m=0}^{k} \binom{n}{m} \cdot p^m \cdot q^{(n-m)}$
	binomcdf(10 , 1/6 , 3)
Mit einer Wahrscheinlichkeit von ca. 93 % wirft man höchstens 3 mal die Augenzahl 7.	

3.2.3. Wie groß ist die Wahrscheinlichkeit, *mindestens* 3 Treffer zu erzielen?

3.2.3.1. Wahrscheinlichkeit, dass das Ereignis X mindestens 3 mal auftritt

Hier muss für alle in Frage kommenden **k** die Trefferwahrscheinlichkeit mit Formel 4 berechnet und die Einzelergebnisse addiert werden.

Hier müssen die Wahrscheinlichkeiten aller möglichen Treffer summiert werden, die zwischen k = 3 und k = 10 liegen.	Oder es findet die **Formel 6** Verwendung $P_{n;p}(X \geq k) = \sum_{m=k}^{n} \binom{n}{m} \cdot p^m \cdot q^{(n-m)} = 1 - \sum_{m=0}^{k-1} \binom{n}{m} \cdot p^m \cdot q^{(n-m)}$
	1 – binomcdf(10 , 1/6 , 2)
Mit einer Wahrscheinlichkeit von ca. 22,4 % würfelt man mindestens 3 mal die Augenzahl 7.	

Man sollte sich die Einzelergebnisse genauer anschauen und sehen, wie die Einzelwahrscheinlichkeiten verteilt sind. Man sieht auch, dass es sehr viel wahrscheinlicher ist, mindestens 3 mal die Augenzahl 7 zu treffen als genau 1 mal.

Die Berechnung der mindestens - Wahrscheinlichkeit kann wie man sieht umfangreich werden und es können sich leicht Fehler einschleichen. Man kann diese Wahrscheinlichkeit aber auch leichter berechnen, wenn man folgende Überlegung anstellt. Für die Berechnung der Trefferwahrscheinlichkeit: mindestens 3 mal die Augenzahl 7 werden von allen Binomialkoeffizienten diejenigen mit k = 3 bis k = 10 berechnet. Diese Wahrscheinlichkeit entspricht der Wahrscheinlichkeit **p**. Alle Wahrscheinlichkeiten der einzelnen Binomialkoeffizienten zusammen ergeben immer **1**. Damit ergeben die Binomialkoeffizienten, die bei der mindestens - Wahrscheinlichkeit übrig bleiben, die Gegenwahrscheinlichkeit **q**.

$$\overbrace{\underbrace{\binom{10}{0}\;\binom{10}{1}\;\binom{10}{2}}_{\mathbf{q}\;=\;1-p}\;\;\underbrace{\binom{10}{3}\;\binom{10}{4}\;\binom{10}{5}\;\binom{10}{6}\;\binom{10}{7}\;\binom{10}{8}\;\binom{10}{9}\;\binom{10}{10}}_{\mathbf{p}\;=\;\text{mindestens 3 mal Augenzahl 7}}}^{\text{Summe} = \mathbf{1}}$$

3.2.3.2. Wahrscheinlichkeit, dass das Gegenereignis Y höchstens 2 mal auftritt

Nun gilt nach Formel 1 ebenfalls $p = 1 - q$
Da sich **q** in diesem Beispiel leichter berechnen lässt, ist diese Umformung von Vorteil.

Hier müssen die Wahrscheinlichkeiten des Gegenereignisses aller möglichen Treffer summiert werden, die zwischen k = 0 und k = 2 liegen.	Oder es findet die **Formel 7** Verwendung
	binomcdf(10 , 5/6 , 2)

Wie man sieht, kommt man mit weniger Rechnerei zum selben Ergebnis.

3.3. Wahrscheinlichkeiten von Ereignissen und Gegenereignissen

X = Ereignis
Y = Gegenereignis
n = Länge der Bernoulli-Kette (Anzahl der Versuche)
p = Wahrscheinlichkeit für X
q = Wahrscheinlichkeit für Y
k,m = Anzahl der Treffer

Die Zusammenhänge zwischen den Wahrscheinlichkeiten für das Ereignis X und den Wahrscheinlichkeiten für das Ereignis Y sollen hier noch einmal zusammenhängend dargestellt werden.

Formel 4

$$\boxed{P_{n;p}(X=k)=P_{n;q}(Y=n-k)}$$

Die Wahrscheinlichkeit bei n Versuchen
das Ereignis X mit der Wahrscheinlichkeit p **genau** k mal zu erreichen
ist gleich groß, wie
das Gegenereignis Y mit der Wahrscheinlichkeit q **genau** n – k mal zu erreichen

---

Formel 5

$$\boxed{P_{n;p}(X\leqslant k)=P_{n;q}(Y\geq n-k)}$$

Die Wahrscheinlichkeit bei n Versuchen
das Ereignis X mit der Wahrscheinlichkeit p **höchstens** k mal zu erreichen
ist gleich groß, wie
das Gegenereignis Y mit der Wahrscheinlichkeit q **mindestens** n – k mal zu erreichen

---

Formel 6

$$\boxed{P_{n;p}(X\geqslant k)=P_{n;q}(Y\leqslant n-k)} \quad \boxed{=1-P_{n;p}(X\leqslant k-1)}$$

Die Wahrscheinlichkeit bei n Versuchen
das Ereignis X mit der Wahrscheinlichkeit p **mindestens** k mal zu erreichen
ist gleich groß, wie
das Gegenereignis Y mit der Wahrscheinlichkeit q **höchstens** n – k mal zu erreichen

---

$$\boxed{P_{n;p}(X>k)=P_{n;q}(Y<n-k)} \quad \boxed{=1-P_{n;p}(X\leqslant k)}$$

Die Wahrscheinlichkeit bei n Versuchen
das Ereignis X mit der Wahrscheinlichkeit p **mehr als** k mal zu erreichen
ist gleich groß, wie
das Gegenereignis Y mit der Wahrscheinlichkeit q **weniger als** n – k mal zu erreichen

---

$$\boxed{P_{n;p}(k\leqslant X\leqslant m)=P_{n;q}(n-m\leqslant Y\leqslant n-k)}$$

Die Wahrscheinlichkeit bei n Versuchen
das Ereignis X mit der Wahrscheinlichkeit p **mindestens** k und **höchstens** m mal zu erreichen
ist gleich groß, wie
das Gegenereignis Y mit der Wahrscheinlichkeit q **mindestens** n – m und **höchstens** n – k mal zu erreichen.
Die Grenzen k und m beim Ereignis X sind beim Ereignis Y vertauscht !

3.4. Modellieren mit der Binomialverteilung - Aufgabentypen

Typ 1: Berechnung von Wahrscheinlichkeiten P **(n, p und k gegeben)**

Bsp.: Ein Tierarzt behandelt 10 kranke Tiere mit einem Medikament, das nach Angaben des Herstellers in 80 % aller Anwendungen zur Heilung führt.

X: Tier wird geheilt

n = 10; p = 0,8

gesucht: P

Einzelne Wahrscheinlichkeit bei n Versuchen

Mit welcher Wahrscheinlichkeit werden ***genau*** 6 von 10 Tieren geheilt? Kapitel 3.1

$$P_{0,8}^{10}(X=6)=\binom{10}{6}\cdot 0,8^6\cdot 0,2^4=0,08808=8,8\%$$

zur Berechnung mit dem GTR ist folgender Aufruf zu benutzen:

binompdf(10 , 0.8 , 6)

Mit welcher Wahrscheinlichkeit werden ***höchstens*** 3 von 10 Tieren geheilt? Kapitel 3.2

$$P_{0,8}^{10}(X\leq 3)=P(X=0)+P(X=1)+P(X=2)+P(X=3)$$

$$=\binom{10}{0}\cdot 0,8^0\cdot 0,2^{10}+\binom{10}{1}\cdot 0,8^1\cdot 0,2^9+\binom{10}{2}\cdot 0,8^2\cdot 0,2^8+\binom{10}{3}\cdot 0,8^3\cdot 0,2^7=37,58\%$$

zur Berechnung mit dem GTR ist folgender Aufruf zu benutzen:

binomcdf(10 , 0.8 , 3)

Mit welcher Wahrscheinlichkeit werden ***mindestens*** 9 von 10 Tieren geheilt? Kapitel 3.3

$$P_{0,8}^{10}(X\geq 9)=P(X=9)+P(X=10)=\binom{10}{9}\cdot 0,8^9\cdot 0,2^1+\binom{10}{10}\cdot 0,8^{10}\cdot 0,2^0=37,58\%$$

zur Berechnung mit dem GTR ist folgender Aufruf zu benutzen:

1 − binomcdf(10 , 0.8 , 8)

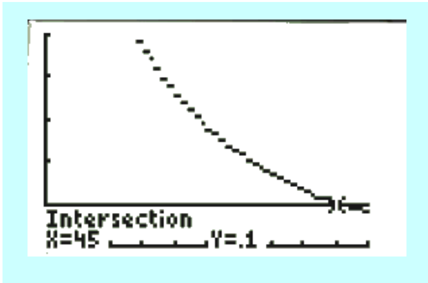
Es ist darauf zu achten, dass der Wert für k um 1 niedriger sein muss, als der „mindestens" Wert angibt.

Die Funktion der Binomialverteilung ist im allgemeinen nicht dafür geschaffen die verschiedenen Parameter als Funktionsvariable zu benutzen. Die werte für n und k müssen immer ganzzahlig sein, sonst lassen sich keine Wahrscheinlichkeiten berechnen. Die Binomialverteilung gehört zu den „diskreten Verteilungen, dh. Verteilungen, die keine reellwertigen k zulassen. Man kann keine Wahrscheinlichkeit für k=3,23 berechnen. Damit entstehen als Funktionen sogenannte Treppenfunktionen, die mit jeder neuen ganzen Zahl eine Stufe höher steigt. Diese Tatsache wirkt sich natürlich auch die Frage nach Schnittpunkten aus. Die GTR arbeiten so, dass als Lösung für k die nächste größere ganzen Zahl angegeben, es wird also das k ermittelt, bei dem das erste Mal der vorgegebene Wert überschritten wird. Die nächsten Anwendungen benutzen die Binomialverteilung als Funktion ihrer jeweils drei Parameter.

Typ 2: Berechnung der Länge einer Bernoulli-Kette n – Versuchsanzahl ist Variable – p und P gegeben, k = 0

Bsp.: Man rechnet mit 5 % Schwarzfahrern auf einer bestimmten Buslinie. Wie viele Fahrgäste muss der Kontrolleur mindestens nach ihrem Fahrschein fragen, bis er mit einer Wahrscheinlichkeit von mindestens 90 % **mindestens** **einen** Schwarzfahrer ertappt hat ?

(Die angegebene Berechnung mit der Hand funktioniert nur, wenn „mindestens ein“ Element gesucht wird. Würde die Aufgabenstellung auf „mindestens zwei“ Elemente lauten, kann keine manuelle Berechnung mehr erfolgen, da man eine Binomialverteilung nicht nach n auflösen kann. Hier kann nur der GTR weiter helfen.)

manuelle Berechnung	Berechnung mit dem GTR
X: Fahrgast ist Schwarzfahrer p = 0,05; P ≥ 0,90 **Gesucht: n** A: Es müssen **mind. 45** Fahrgäste kontrolliert werden.	Y1:binompdf(int(X) , 0.05 , 0) Y2:0,1  Fenstereinstellung: Xmin = 0 ; Xmax = 50 Xscl = 5 Ymin = 0 ; Ymax = 0.5 Yscl = 0.1

4% der männlichen Bevölkerung sind farbenblind. Wie groß muss eine Gruppe von Männern **mindestens** sein, damit mit **mindestens** 90% iger Wahrscheinlichkeit **mindestens**
(daher haben diese Aufgaben die Bezeichnung „3 mal mindestens Aufgaben“ erhalten)
a) einer aus der Gruppe farbenblind ist
b) fünf aus der Gruppe farbenblind sind
In diesem Fall sind nicht alle Parameter für die Verteilungsfunktion bekannt: p ja; n nein
Dafür ist die Wahrscheinlichkeit bekannt, die herauskommen soll. Mathematisch gesehen ist die Gleichung nach n aufzulösen, aber genau das geht bei der Formel nicht. Man müsste in der Binomialverteilung das n durch eine Variable ersetzen, um eine Funktion der Variablen n zu erhalten. Binomialverteilungen im GTR sind auf das n sehr empfindlich, n muss unbedingt eine ganze Zahl sein (!), sonst streikt die Funktion.
Der gesuchte Wert n ist in diesen Fällen ein minimaler Wert. Bei größeren n ist die Wahrscheinlichkeit größer, dass das Ereignis eintritt. Deshalb werden hier „mindestens“ Werte ermittelt. „Mindestens“ ist aber eine obere Summe von einem festen k bis zur oberen Grenze n. Dazu gibt es keine geeignete Funktion, deshalb muss die Funktion binomcdf auf das Gegenereignis umgestellt werden. Damit tritt diese Funktion nur in Verbindung mit „1 –„ auf. Die obere Grenze muss dann genau um 1 tiefer sein, als die untere Grenze bei „mindestens“ für „mindestens einer“ lässt sich die Aufgabe noch von Hand rechnen:
Hier wird wieder die Funktion binomcdf benötigt, die als Verteilungsfunktion der Binomialverteilung mit einem großen F angegeben wird.

$$F_{n,p} = \binom{n}{0}\cdot p^0\cdot q^n + \underbrace{\binom{n}{1}\cdot p^1\cdot q^{n-1} + \ldots + \binom{n}{n-1}\cdot p^{n-1}\cdot q^1 + \binom{n}{n}\cdot p^n\cdot q^0}_{\text{mindestens einer}}$$

Damit führt die Fragestellung nach P(X = 0) zum **ersten Element der Summe** „Mindestens“ einer führt zur Aufgabenstellung 1 – „ keiner“, da eine Summe von 1 bis n nicht direkt berechnet werden kann.

$$1-\binom{n}{0}\cdot p^0\cdot(1-p)^n = 1-1\cdot 0{,}04^0\cdot 0{,}96^n$$

Da der Wert des Binomialkoeffizienten = 1 (alle Binomialkoeffizienten „ … über 0“ führen zu einem Wert =1) und die Potenz von p dazu führt, dass p^0 =1 ebenfalls nicht geschrieben werden braucht, führt der Ausdruck zu einer Potenz der Gegenwahrscheinlichkeit.
Die Ungleichung stellt eine untere Grenze für n dar, was auch sachlich richtig ist, denn je größer n, desto größer muss die Wahrscheinlichkeit sein. Außerdem ist der Wert immer auf die nächste ganze Zahl aufzurunden, da eine Abrundung eine Reduzierung der Wahrscheinlichkeit bewirken würde
Diese Gleichung läßt sich noch über den Logarithmus nach n auflösen. Da beim Auflösen der Gleichung durch einen Logarithmus dividiert werden muss, der kleiner als 0 ist (alle Logarithmen zwischen 0 und 1 sind negativ) dreht sich das Ungleichheitszeichen bei der Division um.

$$\frac{\ln(0{,}1)}{\ln(0{,}96)} \leqslant n$$

$$\frac{-2{,}3026}{-0{,}0408} = 56{,}4 \leqslant n \quad \Rightarrow \quad n \geq 56{,}4 \text{ mindestens } n = 57$$

Diese Gleichung läßt sich noch über den Logarithmus nach n auflösen. Da beim Auflösen der Gleichung durch einen Logarithmus dividiert werden muss, der kleiner als 0 ist (alle Logarithmen zwischen 0 und 1 sind negativ) dreht sich das Ungleichheitszeichen bei der Division um.

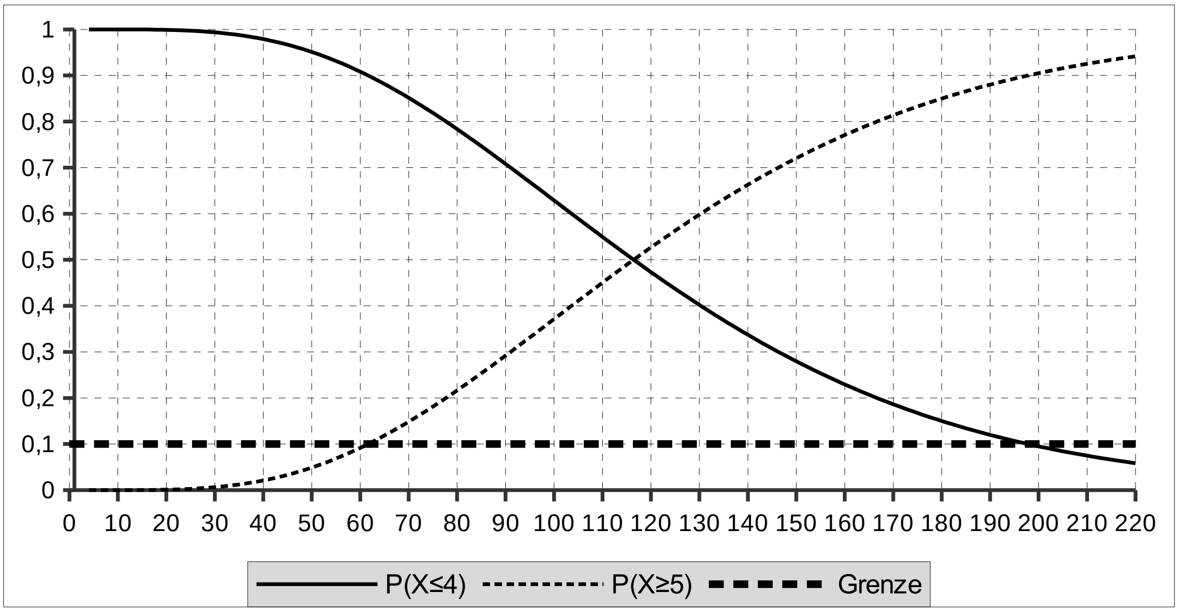
$$\frac{-2{,}3026}{-0{,}0408} = 56{,}4 \leqslant n$$

Die Ungleichung stellt eine untere Grenze für n dar, was auch sachlich richtig ist, denn je größer n, desto größer muss die Wahrscheinlichkeit sein. Außerdem ist der Wert immer auf die nächste ganze Zahl aufzurunden, da eine Abrundung eine Reduzierung der Wahrscheinlichkeit bewirken würde

$\Rightarrow$ n ≥ 56,4 mindestens n = 57

für „mindestens fünf“ (alles, was mehr als einer ist) lässt sich die Aufgabe nur noch mit GTR rechnen:
Hier wird wieder die Funktion binomcdf benötigt, die als Verteilungsfunktion der Binomialverteilung mit einem großen F angegeben wird.

$$F_{n,p} = \binom{n}{0} \cdot p^0 \cdot q^n + \binom{n}{1} \cdot p^1 \cdot q^{n-1} + \ldots + \underbrace{\binom{n}{5} \cdot p^5 \cdot q^{n-5} + \ldots + \binom{n}{n-1} \cdot p^{n-1} \cdot q^1 + \binom{n}{n} \cdot p^n \cdot q^0}_{\text{mindestens fünf}}$$



$P_{n;0{,}04}(X \geq 5) = 1 - P_{n;0{,}04}(X \leq 4) = 1 - \text{binomcdf}(n;\ 0{,}04\ ;\ 4) \geq 0{,}9$
In diesen beiden Kurven sind der Werte der Binomialverteilung dargestellt in Abhängigkeit von n. In diesem speziellen Fall sind für k = 5 die untere Summenwahrscheinlichkeit P(X ≤ 4) als durchgehende Linie und P(x ≥ 5) als gestrichelte Linie dargestellt. Durch Umstellen der Formel wurde die Aufgabe so umgewandelt, dass kein „ 1 – „ mehr auftritt, dafür geht es nicht mehr um eine mindest Wahrscheinlichkeit, sondern um eine höchste Wahrscheinlichkeit. Damit ist die durchgehende Kurve zu betrachten.

Bei dieser Funktion handelt es sich um eine monoton fallende Funktion. Das bedeutet,

- wenn man eine Summenwahrscheinlichkeit bestimmt hat und die ist größer als die angegebene Grenzwahrscheinlichkeit, muss man den n Wert erhöhen.
- Ist die berechnete Wahrscheinlichkeit kleiner als die Summenwahrscheinlichkeit, muss man den n Wert reduzieren.

Verhalten der Funktionswerte

Im unteren Bereich sind die k Wert der einzelnen Funktionen bei gleicher Wahrscheinlichkeit p angegeben.
Für niedrigen p Wert sind die Kurven weit auseinander gezogen.
Die n Werte sind viel größer als die k Werte.
Für die n Werte kann man etwa bei n = 6 * k beginnen.

	k	n
blau	5	44
rot	15	103
braun	25	160

Für p = 0,2 ist 1/0,2 = 5 und damit 5 * k ein guter Startwert.

Im unteren Bereich der Grafik sind die zugehörigen k Werte bei gleicher Wahrscheinlichkeit p angegeben.

Für ein hohes p sind die Kurven sehr schmal, was dazu führt, dass die n Werte nicht viel höher sind als die k Werte. Man kann eventuell mit einem Wert von
n = 1,5 k beginnen.

	k	n
blau	5	14
rot	15	26
braun	25	42

Für p = 0,7 ist 1/0,7 = 1,43 und damit n = 1,5 k ein guter Startwert.

Die Berechnung sollte mit einem n = k : p begonnen werden.

Mit wachsenden Werten von n reduziert sich die Wahrscheinlichkeit, dh.:

- Liegt die berechnete Wahrscheinlichkeit noch über der Grenzwahrscheinlichkeit, ist der Wert von n zu erhöhen.

- Liegt die berechnete Wahrscheinlichkeit unter der Grenzwahrscheinlichkeit, ist der Wert von n zu reduzieren.

Für den ersten Parameter in der bjnomcdf Funktion kann man keine Variable setzen. Deshalb muss man eine Tabelle erzeugen mit den möglich Werten für n.
Geht man jetzt der Empfehlung nach, mit welchem n soll man beginnen, so beginnt man mit n = 5/0.04 = 125

binomcdf(125;0,04;4) = 0,4369

Die Berechnete Wahrscheinlichkeit liegt über der Grenzwahrscheinlichkeit deshalb ist der Wert von n zu erhöhen. Üblicherweise halbiert man jetzt das Intervall bis zur oberen Grenze. Da es diese obere Grenze nicht gibt sollte man vielleicht die Hälfte von 125 dem Wert von 125 hinzuaddieren, aber auf eine ganze Zahl runden:

binomcdf(l88; 0,04 ; 4) = 0,1255

Dieser Wert liegt immer noch über der Grenzwahrscheinlichkeit von 0.1. also ist der Wert von n weiter zu erhöhen. Die Halbierung des Intervalls bis 125 führte zu dem Wert 63. deshalb sollte man jetzt den Wert 62 oder 64 halbieren und zu 188 addieren. 188 + 31 = 219

binomcdf(219;0,04;4) = 0.06

Dieser Wert liegt unter der Grenzwahrscheinlichkeit von 0.1. also ist n wieder zu reduzieren. In diesem Fall könnte man wieder auf die Hälfte des Intervalls gehen und z.B zu 188 nicht 31. sondern nur 16 addieren, also wieder die Hälfte von 31.

binomcdf(204;0,04;4) = 0,086

Der Wert liegt nur wenig unter 0.1. deshalb die Intervalle kleiner wählen, etwa in 5-er Abschnitten, bis man wieder über die Grenzwahrscheinlichkeit kommt. Irgendwann im 1-er Intervallen arbeiten.

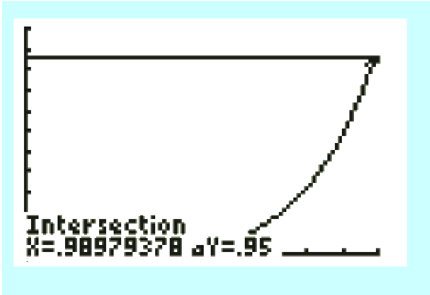
n	P(X=4)	P(X≤ 4)
193	0,0640	0,1119
194	0,0627	0,1093
195	0,0615	0,1068
196	0,0602	0,1044
197	**0,0590**	**0,1020**
198	**0,0578**	**0,0996**
199	0,0566	0,0973

Damit die Wahrscheinlichkeit größer als 90% ist, muss die Anzahl der Versuche mindestens 198 sein.

<u>Typ 3:</u> Berechnung der Trefferwahrscheinlichkeit p – <u>Wahrscheinlichkeit ist Variable</u> – n und P gegeben, k = 0

Bsp.: Ein Gerät besteht aus 5 Bauteilen, die *unabhängig voneinander* mit der gleichen Funktionswahrscheinlichkeit arbeiten. Fällt ein Bauteil aus, so arbeitet das Gerät nicht mehr. Welche Funktionswahrscheinlichkeit müssen die Bauteile haben, wenn das Gerät mit einer Wahrscheinlichkeit von mehr als 95 % funktionieren soll ?

(Die Berechnung stützt sich auf die Voraussetzung „unabhängig voneinander“, da in diesem Fall der Multiplikationssatz ohne bedingte Wahrscheinlichkeit gilt. In diesem Fall ist die bedingte Wahrscheinlichkeit jeweils die Wahrscheinlichkeit für das Eintreten des Ereignisses selbst.)

manuelle Berechnung	Berechnung mit dem GTR
X: Bauteil funktioniert n = 5; P > 0,95 **Gesucht p ;** $P_p^5(X=0) = \binom{5}{0}\cdot p^5\cdot(1-p)^0 \geq 0{,}95$ $p^5 \geq 0{,}95 \quad \vert \sqrt[5]{\ }$ $p \geq \sqrt[5]{0{,}95} \approx 98{,}98\,\%$	Y1:binompdf(5 ,X, 0) Y2:0,95  Fenstereinstellung: Xmin = 0 ; Xmax = 1 Xscl = 0.1 Ymin = 0 ; Ymax = 1.1 Yscl = 0.1

Jedes Bauteil einer Produktionsserie fällt mit der Wahrscheinlichkeit p aus. Die Bauteile werden unabhängig voneinander produziert Wie groß darf p höchstens sein, damit mit einer Wahrscheinlichkeit von höchstens 80% höchstens 10 von 100 Bauteilen ausfallen.

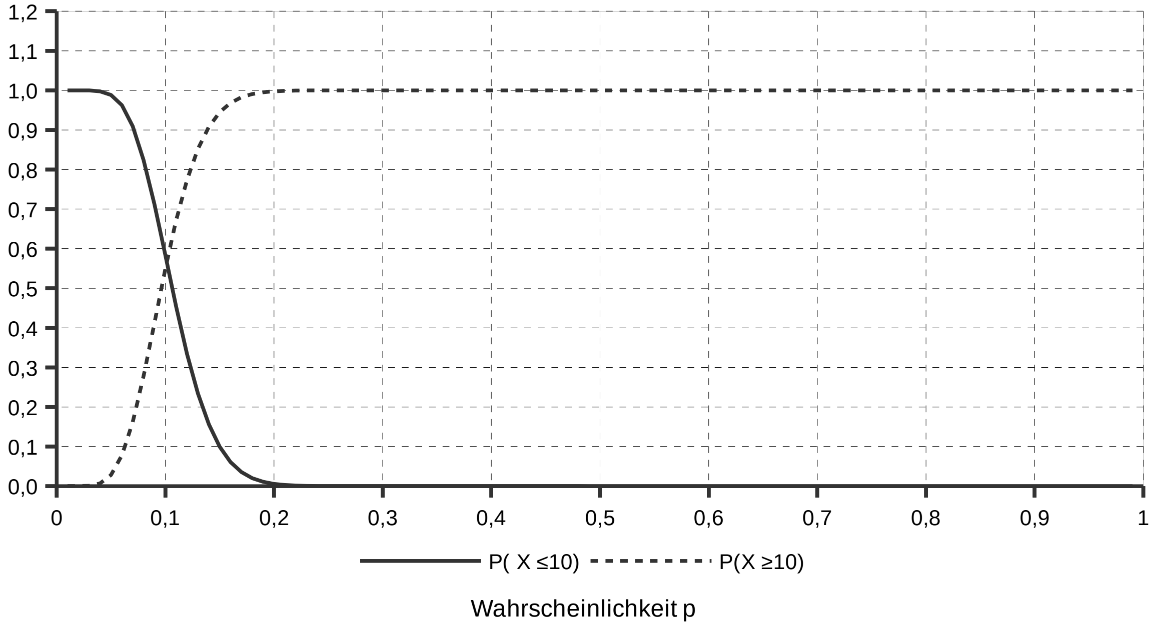
In diesem Fall sind nicht alle Parameter für die Verteilungsfunktion bekannt: n ja; k ja; p nein
Dafür ist die Wahrscheinlichkeit bekannt, die herauskommen soll. Mathematisch gesehen ist die Gleichung nach p aufzulösen, aber genau das geht bei der Formel nicht Dazu kommt noch, dass das notwendige q ebenfalls 1 - p ist Man kann den Wert von p durch eine Variable ersetzen, da p eine reelle Zahl ist, die Frage ist, ob der GTR das ermöglicht und ob man einen y - Wert eingeben kann und der zugehörige x Wert ermittelt wird.

Die gesuchte Wahrscheinlichkeit ist eine obere Grenze der Wahrscheinlichkeiten, ist die Wahrscheinlichkeit für den Ausfall eines Bauteil kleiner, dann ist natürlich auch die gesamte Ausfallquote kleiner. Deshalb wird hier mit„höchstens" gearbeitet, aber das liefert genau die Funktion binomcdf .Deshalb gibt es keinen Wechsel zum Gegenereignis.

$$F_{100,p} = \binom{100}{0}\cdot p^0\cdot q^{100} + \binom{100}{1}\cdot p^1\cdot q^{99} + \ldots + \binom{100}{10}\cdot p^{10}\cdot q^{90} \leq 0{,}8$$

$P_{100;p}$ (X ≤ 10) = binomcdf (100; p ; 10) ≤ 0,8

Verhalten der Funktionswerte



In diesen beiden Kurven sind der Werte der Binomialverteilung dargestellt in Abhängigkeit von p. In diesem speziellen Fall sind für n = 100 und k = 10 die untere Summenwahrscheinlichkeit P(X ≤ 10) als durchgehende Linie und P(X ≥ 10) als gestrichelte Linie dargestellt.
Falls eine obere Summenwahrscheinlichkeit gesucht ist, dann wird die Berechnung so umgestellt, dass es auf eine untere Summenwahrscheinlichkeit führt, da obere Summenwahrscheinlichkeiten nicht berechnet werden können. Am Ende ist es immer die untere Summe, es ändert sich nur ob eine oberer Grenzwert, oder ein unterer Grenzwert gegeben ist; ob eine vorgegebene Wahrscheinlichkeit mindestens erreicht werden muss, oder ob die Wahrscheinlichkeit höchstens erreicht werden darf.

Bei dieser Funktion handelt es sich um eine monoton fallende Funktion.

	k	p
blau	5	0,09
rot	15	0,21
braun	25	0,32

Für k = 5 ist p = 0,09. Für n = 100 ist k / n = 1/20 = 0,05.
Für k = 15 ist p = 0,21. Für n = 100 ist k / n = 0,15.
Für k = 25 ist p = 0,32. Für n = 100 ist k / n = 0,25
Alle Werte liegen unter der gesuchten Wahrscheinlichkeit, man könnte die Werte leicht aufrunden.

	k	p
blau	5	0,22
rot	15	0,49
braun	25	0,73

Für k = 5 ist p = 0,22. Für n = 40 ist k / n = 0,13.
Für k = 15 ist p = 0,49. Für n = 40 ist k / n = 0,38.
Für k = 25 ist p = 0,73. Für n = 40 ist k / n = 0,63
Alle Werte liegen unter der gesuchten Wahrscheinlichkeit, aber die Werte sind gute Startwerte für die Berechnung.

Die Berechnung sollte mit einem p = k : n begonnen werden.

Mit wachsenden Werten von p reduziert sich die Wahrscheinlichkeit, dh.:

- Liegt die berechnete Wahrscheinlichkeit unter der Grenzwahrscheinlichkeit, ist der Wert von p zu reduzieren.
- Liegt die berechnete Wahrscheinlichkeit über der Grenzwahrscheinlichkeit, ist der Wert von p zu erhöhen.

Geht man jetzt der Empfehlung nach, mit welchem p soll man beginnen, so beginnt man mit p = 10/100 = 0,1

$P_{100;0,1}$ (X ≤ 10) = bjnomcdf (100 ; 0.1 ; 10) = 0,5832

Dieser Wert ist noch weit von dem gesuchten Wert 0.8 entfernt. Die Wahrscheinlichkeit ist unter der Grenzwahrscheinlichkeit, deshalb ist die Wahrscheinlichkeit zu reduzieren. In der numerischen Mathematik hat sich bewährt das Intervall immerzu halbieren. In diesem Fall also mit p = 0.05 weiter zu rechnen:

$P_{100\,;\,0,05}$ (X ≤ 10) = bjnomcdf(100; 0.05 ; 10) = 0,9885

Diese Wahrscheinlichkeit ist über der Grenzwahrscheinlichkeit, deshalb ist die Wahrscheinlichkeit zu erhöhen auf die Hälfte zwischen 0.1 und 0.05. also p = 0.075

$P_{100\,;\,0,075}$ (X ≤ 10) = bjnomcdf(100; 0.075 ; 10) = 0,8707

Wenn der GTR für p eine Liste akzeptiert, ist die Arbeit mit Liste am schnellsten. In einer Liste sollte der Wert k/ n etwa in der Mitte der Listenwerte liegen. In diesem Fall bei p = 0,1 sollte man mit der Liste bei 0,05 beginnen und dann Schrittweiten von 0,01 angeben, als obere Grenze 0,15 benutzen.

Typ 4: Bestimmung der Versuchszahl k – Anzahl positiver Ereignisse ist Variable – n, p und P gegeben

Bsp.: Ein Großhändler versorgt 10 Geschäfte, von denen jedes Bestellungen mit einer Wahrscheinlichkeit von p = 0,4 aufgibt.

a) Wie viel Bestellungen laufen mit größter Wahrscheinlichkeit ein ?
Welche Anzahl an Bestellungen ist am wahrscheinlichsten ?

b) Mit welcher Wahrscheinlichkeit weicht die Zahl der Bestellungen um höchstens 1 vom wahrscheinlichsten Wert ab?

X: Anzahl der Bestellungen

n = 10; p = 0,4

Gesucht: k

a) Diese Frage beantwortet der Erwartungswert. Der Erwartungswert einer Binomialverteilung beträgt immer
E(X) = n • p = 10 • 0,4 = 4. Die Wahrscheinlichkeit für dieses k beträgt:

$$P_{0,4}^{10}(X=4)=\binom{10}{4}\cdot 0,4^4\cdot 0,6^6=0,2508$$ = 0,25

binompdf(10 , 0.4 , 4)

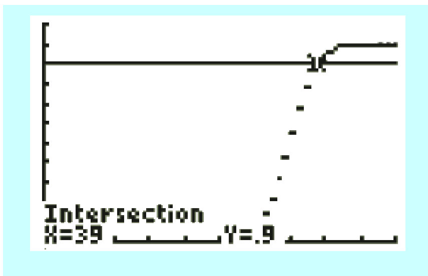
b)

$$P_{0,4}^{10}(3\le X\le 5)=P(X\le 5)-P(X\le 2)$$

$$=\binom{10}{3}\cdot 0,4^3\cdot 0,6^7+\binom{10}{4}\cdot 0,4^4\cdot 0,6^6+\binom{10}{5}\cdot 0,4^5\cdot 0,6^5=66,66\%$$

binomcdf(10 , 0.4 , 5)−binomcdf(10 , 0.4 , 2)

Bsp.: Nach einem Maschinenschaden sind erfahrungsgemäß 70 % der Teile Ausschuss. Es werden 50 Teile beliebig entnommen. Mit wie vielen Teilen muss man mindestens rechnen, wenn man ein Risiko von höchstens 10 % eingehen möchte?

manuelle Berechnung	Berechnung mit dem GTR
X: Anzahl der schlechten Teile n = 50; p = 0,7; P(X ≥ k) ≤ 0,10 Gesucht: k ≤ 0,10 1 – P (X ≤ k – 1) ≤ 0,10 ≥ 0,9 Ablesen aus der Tabelle: k – 1 ≥ 39 → k ≥ 40	Y1:binompdf(50 , 0.7 , int(X)) Y2:0,9  Fenstereinstellung: Xmin = 0 ; Xmax = 50 Xscl = 5 Ymin = 0 ; Ymax = 1.1 Yscl = 0.1

Bei rechten Grenzwerten von k muss man den Wert immer um 1 höher setzen, da mit der oberen Grenze nur k – 1 bestimmt werden kann.

3.5. Wartezeitaufgaben

3.5.1. Wahrscheinlichkeiten für den 1. Treffer

1. Wahrscheinlichkeit, dass der 1. Treffer genau im k Versuch eintritt

$$\boxed{P_{n;p}(\text{k Nieten, dann 1. Treffer}) = (1-p)^{k-1} \cdot p}$$

Ein Laplace-Würfel werde 10 mal geworfen.
Wie groß ist die Wahrscheinlichkeit, dass **die die erste 6 genau** beim 10. Wurf auftritt?

P(6 das erste Mal beim 10. Wurf) = $\left(\frac{5}{6}\right)^9 \cdot \frac{1}{6}$

2. Wahrscheinlichkeit, dass der 1. Treffer beim k-ten Versuch oder später eintritt

$$\boxed{P_{n;p}(\text{kein Treffer vor dem k} - \text{ten Versuch}) = (1-p)^{k-1}}$$

Ein Laplace-Würfel werde solange geworfen, bis eine 6 erscheint.
Wie groß ist die Wahrscheinlichkeit, dass **die erste 6 frühestens** beim 4. Wurf auftritt?

P(6 frühestens beim 4. Wurf) = $\left(\frac{5}{6}\right)^3$

3. Wahrscheinlichkeit, dass der 1. Treffer spätestens im k-ten Versuch eintritt.

$$\boxed{P_{n;p}(\text{höchstens k Nieten, dann 1. Treffer}) = 1-(1-p)^k}$$

Ein Laplace-Würfel werde solange geworfen, bis eine 6 erscheint.
Wie groß ist die Wahrscheinlichkeit, dass **die erste 6 spätestens** beim 8. Wurf auftritt?

P(erste 6 spätestens beim 8. Wurf) = $1 - \left(1 - \frac{1}{6}\right)^8 = 1 - \left(\frac{5}{6}\right)^8$

3.5.2. Wahrscheinlichkeiten für den k – ten Treffer genau

4. Wahrscheinlichkeit, dass dem 2. Treffer genau k Versuche (k – 1 Nieten und 1 Treffer) vorausgehen?

$$\boxed{P_{k;p}(X=2;\bar{X}=k) = \binom{k}{1} \cdot p^1 \cdot (1-p)^{k-1} \cdot p = k \cdot p^2 \cdot (1-p)^{k-1}}$$

Ein Laplace-Würfel werde solange geworfen, bis die zweite 6 erscheint.
Wie groß ist die Wahrscheinlichkeit dafür, dass dies *beim 10.Wurf* geschieht?

P(zweite 6 beim 10.Wurf) = $9 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^8$

5. Wahrscheinlichkeit, dass der k-te Treffer genau im n Versuch (k – 1 Treffer und n – (k -1) Nieten) eintritt ?

$$\boxed{P(X=k;\bar{X}=n-(k-1)) = \binom{n-1}{k-1} p^{k-1} \cdot (1-p)^{n-(k-1)} \cdot p = \binom{n-1}{k-1} p^k \cdot (1-p)^{n-(k-1)}}$$

6. Wahrscheinlichkeit, dass der 1. Treffer im m – ten und der k – te Treffer im n-ten Versuch eintritt ?

$$P_{n;p}()=\binom{n-m-1}{k-2}p^k\cdot(1-p)^{n-k}$$

3.5.3.Wahrscheinlichkeiten für den k – ten Treffer frühestens oder spätestens

7. Wahrscheinlichkeit, dass der k-te Treffer frühestens beim n-ten Versuch eintritt ?

$$P_{n;p}(X\leqslant k-1)=\sum_{i=0}^{k-1}\binom{n-1}{i}\cdot p^i\cdot(1-p)^{(n-i-1)}$$

Ein Laplace-Würfel werde solange geworfen, bis die zweite 6 erscheint.
Wie groß ist die Wahrscheinlichkeit dafür, dass dies *frühestens* beim 10.Wurf geschieht?

P(zweite 6 frühestens beim 10. Wurf) = $P_{n;p}(X\leqslant k-1)=\sum_{i=0}^{1}\binom{9}{i}\cdot p^i\cdot(1-p)^{(10-i-1)}=\left(\frac{5}{6}\right)^9+9\cdot\left(\frac{5}{6}\right)^8\cdot\left(\frac{1}{6}\right)^1$

8. Wahrscheinlichkeit, dass der k-te Treffer spätestens beim n-ten Versuch eintritt ?

$$P_{n;p}()=1-P_{n;p}(X\leqslant k-1)=1-\sum_{i=0}^{k-1}\binom{n}{i}\cdot p^i\cdot(1-p)^{(n-1)}$$

Ein Laplace-Würfel werde solange geworfen, bis die zweite 6 erscheint.
Wie groß ist die Wahrscheinlichkeit dafür, dass dies *spätestens beim 10. Wurf* geschieht?

P(zweite 6 spätestens beim 10.Wurf) =

$$P(X=k;\bar{X}\leq n-(k-1))=1-\sum_{i=0}^{1}\binom{10}{i}\cdot p^i\cdot(1-p)^{(10-i-1)}=1-\left(\frac{5}{6}\right)^{10}-10\cdot\left(\frac{5}{6}\right)^9\cdot\left(\frac{1}{6}\right)^1$$

4. Hypergeometrische Verteilung

Die Binomialverteilung behandelt Aufgabenstellungen „mit Wiederholung“ und „ohne Reihenfolge“. In den ersten Kapiteln dieses Dokuments wurde gezeigt, dass die Bedingung „ohne Reihenfolge“ problemlos in eine Bedingung „mit Reihenfolge“ umgewandelt werden kann, indem man durch die Anzahl der Pfade dividiert, die insgesamt die gleiche Anzahl positiver Ereignisse haben. Etwas anders sieht es aus, wenn man von „mit Wiederholung“ auf „ohne Wiederholung übergeht. Da zu soll das gleiche Beispiel, wie in Kapitel 3 betrachtet werden.

In einem Gefäß (Urne) befindet sich eine bestimmte Anzahl von Kugeln. Die Gesamtzahl der Kugeln ist die Menge **n**. Jede einzelne Kugel ist Element der Menge **n**. Jede Kugel besitzt ein bestimmtes Merkmal. Aus der Anzahl der Kugeln (Elemente) mit demselben Merkmal ergibt sich die Häufigkeit und daraus wiederum die Wahrscheinlichkeit für das Auftreten des Merkmals.

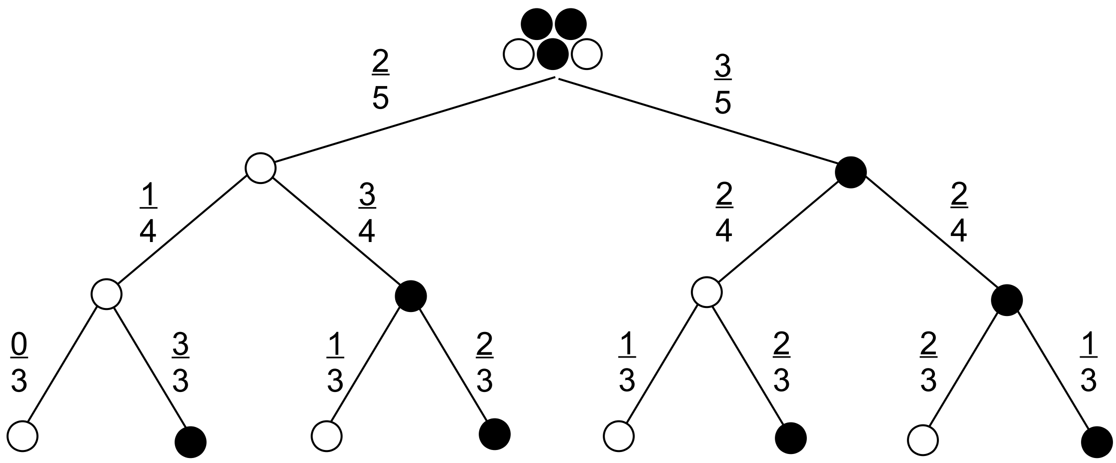
Beispiel 1

In einem Gefäß befinden sich 24 schwarze und 12 weiße Kugeln,

Bei diesen Aufgaben gibt es wieder 2 mögliche Merkmale bzw. deren Wahrscheinlichkeiten. Für die Betrachtung „mit Zurücklegen“ befinden sich in der Urne immer wieder 36 Kugeln mit der gleichen Farbzusammensetzung. Zieht man aus der Urne „ohne Wiederholung“, dann bedeutet das nicht nur, dass bei jeder neuen Ziehung eine Kugel weniger in der Urne ist, sondern auch, dass sich die Farbzusammensetzung der Kugeln in der Urne ständig ändert. Damit macht es keinen Sinn mit Wahrscheinlichkeiten p und q zu rechnen, da diese sich nach jeder Ziehung wieder ändern. Es macht nur Sinn, mit den jeweiligen Anzahlen die Wahrscheinlichkeit immer wieder neu zu berechnen. Die Herleitung der Formeln wurden am Anfang des Dokuments erläutert.

Angenommen, in einem Gefäß befinden sich 5 Kugeln, davon 3 schwarze und 2 weiße. Bei jedem Versuch wird eine Kugel herausgenommen, ihr Merkmal (schwarz / weiß) geprüft und die Kugel wieder zurückgelegt. Man führt 3 solche Versuche durch. Wie groß ist dabei die Wahrscheinlichkeit, dass man genau 2 schwarze Kugeln gezogen hat? Den Ablauf der 3 Versuche macht man sich mit einem Baumdiagramm deutlich.

Beispiel 2



$$\left(\frac{2}{5}\right)\cdot\left(\frac{3}{4}\right)\cdot\left(\frac{2}{3}\right) \qquad \left(\frac{3}{5}\right)\cdot\left(\frac{2}{4}\right)\cdot\left(\frac{2}{3}\right) \qquad \left(\frac{3}{5}\right)\cdot\left(\frac{2}{4}\right)\cdot\left(\frac{2}{3}\right)$$

Am Baumdiagramm sieht man, dass sich bei jeder neuen Ziehung der Nenner in der Wahrscheinlichkeit ändert, nämlich genau um 1, da 1 Kugel weniger in der Urne ist. Der Zähler ändert sich je nachdem, welche Farbe gezogen wurde, einmal bei der Anzahl der weißen Kugeln und einmal bei der Anzahl der schwarzen Kugeln. Betrachtet man allerdings wieder die Ereignisse „ 2 schwarze und 1 weiße“, dann ergibt sich in allen Fällen die gleiche Wahrscheinlichkeit. Die Nenner sind sowieso gleich, da sich diese mit jeder Ziehung um 1 reduzieren, gleichgültig, welche Farbe gezogen wurde. Betrachtet man die Zähler, stellt man fest, die Faktoren der Zähler sind auch gleich, nur in der Reihenfolge ändern sie sich. Damit ist auch die Wahrscheinlichkeit beim „Ziehen ohne Wiederholung“ nur Abhängig von der gewählten Anzahl der positiven Ereignisse. Bei gleicher Anzahl hat jeder Zweig im Baum die gleiche Wahrscheinlichkeit. Es genügt also auch hier, die Wahrscheinlichkeit eines Zweiges zu berechnen und dann mit der Anzahl der Zweige zu multiplizieren.

Dieser Zusammenhang wurde bereits in den einführenden Kapiteln diese Dokuments dargelegt.

Die Anzahl der Pfade entspricht genau der Anzahl, die auch bei der Binomialverteilung auftreten. Damit gilt auch hier für die Anzahl der Pfade:

Formel 3 $$\binom{n}{k}=\frac{n!}{(n-k)!k!}$$

Nur die Wahrscheinlichkeit eines Pfades ist anders zu bestimmen. Für 2 schwarze und 1 weiße stehen im Zähler auf alle Fälle einmal die Anzahl der vorhandenen schwarzen s = 3 und die Anzahl der vorhandenen weißen w = 2, denn jede Farbe wird mindestens einmal gezogen. Für die zweite schwarze Kugel sind jetzt nicht mehr die Gesamtanzahl der schwarzen verfügbar, sondern eine Kugel weniger, da für die zweite schwarze schon einmal eine schwarze gezogen sein muss, also s – 1 = 2. Damit setzt sich die Pfadwahrscheinlichkeit folgendermaßen zusammen:

$$\frac{S\cdot(S-1)\ \cdot\ W}{N\cdot(N-1)\cdot(N-2)}=\frac{3\cdot 2\ \cdot\ 2}{5\cdot 4\cdot 3}$$

Betrachtet man die Faktoren für schwarz und für die Gesamtanzahl stellt man fest, dass sie ein Stück eines Fakultätsausdrucks sind, und zwar der obere Teil der Berechnung einer Fakultät. Wenn man also den Ausdruck S! durch einen geeigneten Ausdruck dividiert, so dass S (S-1) herauskommt, kann man den Teil den Zählers durch einen Wahrscheinlichkeitsausdruck darstellen.

$$S\cdot(S-1)=\frac{S!}{(S-2)!}$$

Die Zahl 2, die im Nenner subtrahiert werden muss, entspricht genau der Anzahl der gezogenen schwarzen Kugeln. Analog kann man den Ausdruck für die Gesamtanzahl im Nenner umformen:

$$N\cdot(N-1)\cdot(N-2)=\frac{N!}{(N-3)!}$$

Damit entspricht die Zahl, die subtrahiert werden muss wieder der Anzahl der insgesamt gezogenen Kugel. Für die *Wahrscheinlichkeit eines Pfades* entsteht daraus die Wahrscheinlichkeit

Formel 2' $$\frac{\frac{S!}{(S-s)!}\cdot\frac{W!}{(W-w)!}}{\frac{N!}{(N-n)!}}$$

Diese Formel ist bereits bei der Herleitung der Wahrscheinlichkeit benutzt worden. Es handelt sich dabei um die Kombinatorikformel für „mit Reihenfolge, ohne Wiederholung“. Für einen einzelnen Pfad ist die Wahrscheinlichkeit natürlich das gleiche, wie „mit Reihenfolge“, da in einem Pfad nur eine Reihenfolge existiert. „Ohne Wiederholung“ resultiert daraus, dass die Kugeln nicht wieder zurückgelegt werden. Lässt man die Reihenfolge fallen, dh., es werden alle Pfade betrachtet, die das gleiche Endergebnis liefern entsteht die Hypergeometrische Verteilung.

Formel 4' $$P(X=s)=P(Y=w)=\frac{\binom{S}{s}\cdot\binom{W}{w}}{\binom{N}{n}}$$

Da mit der Festlegung der Anzahl der schwarzen Kugeln bei n Ziehung auch gleichzeitig die Anzahl der weißen Kugeln festliegt, kann man damit beide Wahrscheinlichkeiten berechnen.

4.1. Anzahl positiver Ereignisse

Für die Werte von **k** gibt es bei den Aufgaben ohne Wiederholung ebenfalls drei Möglichkeiten.
Folgende 3 Wahrscheinlichkeiten können gefragt sein:

a) für genau **k** Treffer
b) für mindestens **k** Treffer (= wenigstens **k** Treffer)
c) für höchstens **k** Treffer

Dafür kann man sich eine Folge (Kette) von **n** Versuchen vorstellen. Nun kann man die möglichen Trefferzahlen abzählen. Angenommen **n** = 10. Bei diesen 10 Versuchen ist es immerhin möglich, dass man in keinem der 10 Versuche einen Treffer hat. Ebenso ist es möglich, dass man 10 Treffer hat, nämlich in jedem Versuch einen. Die Wahrscheinlichkeit dafür würde man Glück nennen. Natürlich sind zwischen 0 und 10 alle ganzen Zahlen als Treffer möglich. Man sieht auch, dass für jede Trefferzahl eine andere Wahrscheinlichkeit gilt. Auf diese Weise kommt man der Verteilung der Wahrscheinlichkeiten auf die Spur.

X = Ereignis
Y = Gegenereignis
N = Anzahl der gesamten Elemente
S = Anzahl der positiven Ereignisse für X
W = Anzahl der positiven Ereignisse für Y
s = Anzahl der Treffer für S
w = Anzahl der Treffer für W
n = Anzahl der gezogenen Elemente

4.1.1. Genau k – Treffer

Wenn man nun die Aufgabe lösen will, die Trefferwahrscheinlichkeit für das Ereignis X mit der Wahrscheinlichkeit p **genau k Treffer** zu ermitteln, dann muss man aus der Reihe der Binomial-Koeffizienten den mit dem entsprechenden **k** - Wert herausnehmen und nach der Formel 4 berechnen.

Beispiel 1
N = Anzahl der gesamten Elemente : 36
n = Anzahl der gezogenen Elemente : 10

S = Anzahl der positiven Ereignisse für X : 24
s = Anzahl der gezogenen positiven Ereignisse für X : 4

W = Anzahl der positiven Ereignisse für Gegenereignis Y : 12
w = Anzahl der gezogenen positiven Ereignisse für Gegenereignis Y : 6

Die Hypergeometrische Verteilung ist von mehr Parametern abhängig, als die Binomialverteilung. Leider ist bei der Angabe der Bezeichnung die Reihenfolge nicht immer die gleiche. Zur Bezeichnung der Verteilung wird hier folgende Reihenfolge verwendet:

1. Parameter : Gesamtzahl der Elemente
2. Parameter : Gesamtzahl der gezogenen Elemente
3. Parameter : Gesamtzahl der positiven Ereignisse für X (beim Gegenereignis für Y)
   Variable in der Klammer : Anzahl der gezogenen positiven Ereignisse

$$P_{N;n;S}\ (s)$$

Trefferwahrscheinlichkeit für genau 4 Treffer: P(X = 4) oder P(Y = 6)

Ereignis Gegenereignis

$$P_{36\,;\,10;\,24}(X=4) = \frac{\binom{24}{4}\cdot\binom{12}{6}}{\binom{36}{10}} = P_{36\,;\,10;\,12}(Y=6)$$

Formel 3'

$$P_{N;n;S}(X=s) = \frac{\binom{S}{s}\cdot\binom{N-S}{n-s}}{\binom{N}{n}} = \frac{\binom{N-W}{n-w}\cdot\binom{W}{w}}{\binom{N}{n}} = P_{N;n;W}(Y=w) = P_{N;n;W}(Y=n-s)$$

Die auftretenden Faktoren sind die gleichen, wie beim Gegenereignis Y. Wenn es S Elemente mit der Eigenschaft X gibt, dann muß es auch N – S = W Elemente mit der Eigenschaft Y geben. Damit ist S = N – W, aber auch W = N – S. Das gleiche trifft für die gezogenen Elemente zu. Wenn bei n gezogenen Elementen s Elemente die Eigenschaft S haben, dann müssen n – s = w Elemente die Eigenschaft W haben, da es keine anderen Eigenschaften gibt.
Das Problem bei der Berechnung der Wahrscheinlichkeiten besteht darin, dass in keinem der gängigen GTR die Hypergeometrische Verteilung implementiert ist, so dass die Berechnung immer elementar über die Binomialkoeffizienten erfolgen muss.

4.1.2. Höchstens k – Treffer

Die Trefferwahrscheinlichkeit aus Beispiel 1 für höchstens 4 Treffer ergibt sich aus allen Fällen, wo genau 4 Treffer oder weniger erzielt werden. Höchstens 4 bedeutet dabei: 4 oder weniger, also s ≤ 4 (kleiner / gleich 4).

Auch diese Einzelwerte werden addiert (kumuliert) und ergeben den Wert der Hypergeometrischen-Verteilung.

$$\begin{array}{cccccc} & s=0 & s=1 & s=2 & s=3 & s=4 \end{array}$$

$$P_{36;10,24}(X\le 4) = \frac{\binom{24}{0}\cdot\binom{12}{10}}{\binom{36}{10}} + \frac{\binom{24}{1}\cdot\binom{12}{9}}{\binom{36}{10}} \quad \frac{\binom{24}{2}\cdot\binom{12}{8}}{\binom{36}{10}} + \frac{\binom{24}{3}\cdot\binom{12}{7}}{\binom{36}{10}} + \frac{\binom{24}{4}\cdot\binom{12}{6}}{\binom{36}{10}} = P_{36;10,12}(Y\ge 6)$$

Für das Gegenereignis vertauscht sich nur die Reihenfolge der Faktoren im Zähler. Wenn das Ereignis X s erfolgreiche Treffer hat, muss das Ereignis Y n – s = w erfolgreiche Treffer haben. Wenn die Anzahl der Treffer für X nur 4 oder weniger sind, müssen es für Y 6 oder mehr sein.

Für die Berechnung der Summe von 0 bis k gibt es keine geschlossene Summenformel. Man muss jeden Wert einzeln berechnen und dann addieren.

Damit entsteht die Formel der summierten Hypergeometrischen Verteilung:

Formel 5 (summierte Hypergeometrische Verteilung von 0 bis k)

$$P_{N;n;S}(X\le k) = \frac{\sum_{m=0}^{k}\binom{S}{m}\cdot\binom{N-S}{n-m}}{\binom{N}{n}}$$

3.1.3. Mindestens k – Treffer

Will man nun die Trefferwahrscheinlichkeit für mindestens 4 Treffer ermitteln, kann man aus der Reihe der Binomialkoeffizienten erkennen, in welchen Fällen mindestens 4 Treffer erreicht werden. Mindestens 4 bedeutet dabei: 4 oder mehr, also k ≥ 4 (größer / gleich 4).

Nun muss man für jeden Fall P(X = k) mit der Formel 4' berechnen und **die Einzel-Ergebnisse addieren**. Weil es sich um eine Addition von Einzelwerten handelt, spricht man auch von einer **kumulativen** Rechnung.
(kumulativ = lateinisch: angehäuft)

s = 4 s = 5 s = 6 . . . . s = 10

$$P_{36;10,24}(X\geq 4)=\frac{\binom{24}{4}\cdot\binom{12}{6}}{\binom{36}{10}}+\frac{\binom{24}{5}\cdot\binom{12}{5}}{\binom{36}{10}}\;\frac{\binom{24}{6}\cdot\binom{12}{4}}{\binom{36}{10}}+\ldots+\frac{\binom{24}{10}\cdot\binom{12}{0}}{\binom{36}{10}}=P_{36;10,12}(Y\leq 6)$$

Für diese Summe ergeben sich 7 Summanden von s = 4 bis s = 10 .

Für diese Art der Summierung existieren keine Tabellen und lassen sich auch nicht direkt mit dem GTR durchführen. Da für die summierte hypergeometrische Verteilung im GTR keine Formel zur Verfügung steht kann man auch hier nur direkt die einzelnen Summanden berechnen. Wenn bei der Berechnung des Gegenereignisses erheblich weniger Summanden entstehen, wie hier zum Beispiel, sollte man überlegen, ob man zur Berechnung des Gegenereignisses übergeht. Die Summierung der unteren Einzelwahrscheinlichkeiten führen hier nur zu 4 einzelnen Summanden, statt zu 7.

$$\boxed{P_{36;10;24}(X\geq 4)=1-P_{36;10;24}(X<4)=1-P_{36;10;24}(X\leq 3)}$$

Formel 6 (summierte hypergeometrische Verteilung von k bis n)

$$P_{N;n;S}(X\geq s)=\frac{\sum\limits_{m=s}^{n}\binom{S}{m}\cdot\binom{N-S}{n-m}}{\binom{N}{n}}=1-\frac{\sum\limits_{m=0}^{s-1}\binom{S}{m}\cdot\binom{N-S}{n-m}}{\binom{N}{n}}=1-P_{N;n;S}(X\leq s-1)$$

oder man benutzt die summierte Wahrscheinlichkeit für das Gegenereignis:

$$\boxed{P_{36;10;24}(X\geq 4)\quad=\quad P_{36;10;12}(Y\leq 6)}$$

Formel 7 (summierte hypergeometrische Verteilung des Gegenereignisses von 0 bis n – k)

$$P_{N;n;S}(X\geq s)=\frac{\sum\limits_{m=s}^{n}\binom{S}{m}\cdot\binom{N-S}{n-m}}{\binom{N}{n}}=\frac{\sum\limits_{m=0}^{n-s}\binom{W}{m}\cdot\binom{N-W}{n-m}}{\binom{N}{n}}=P_{N;n;W}(Y\leq n-s)$$

Wichtig ist dabei zu beachten, dass sich nicht nur die Grenzen von m verändern, sondern auch die Reihenfolge der Wahrscheinlichkeiten. Tritt für das ursprüngliche Ereignis X das positive Ereignis **mindestens s** mal ein, dann tritt für das Gegenereignis Y das positive Ereignis **höchstens n – s** mal ein..

Wenn das Ereignis X	**mindestens s mal**	auftritt
kann das Gegenereignis Y	**höchstens n – s mal**	auftreten

4.2. Beispiele zur hypergeometrischen Verteilung

Die drei typischen Trefferwahrscheinlichkeiten genau k Treffer / höchstens k Treffer / mindestens k Treffer sollen an einem weiteren Beispiel verdeutlicht werden.

Eine Urne enthält 6 schwarze und 4 weiße Kugeln und es werden 5 Kugeln gezogen.

In den gängigen Taschenrechnern ist eine Hypergeometrische Verteilung nicht implementiert, so dass alle Werte über die Binomialkoeffizienten einzeln zu berechnen sind.

4.2.1. Wie groß ist die Wahrscheinlichkeit, *genau* 3 Treffer zu erzielen ?

Damit sind alle Werte für die **Formel 4'** bekannt und können eingesetzt werden.

	Anwendung von **Formel 4'**
	$\frac{\mathrm{nCr}(6,3)\cdot\mathrm{nCr}(4,2)}{\mathrm{nCr}(10,5)}$

Die Wahrscheinlichkeit, bei 5 Kugeln genau 3 weiße Kugeln zu bekommen ist 0,4687 .

4.2.2. Wie groß ist die Wahrscheinlichkeit, *höchstens* 3 Treffer zu erzielen?

Hier müssen die Wahrscheinlichkeiten aller möglichen Treffer summiert werden, die zwischen k = 0 und k = 3 liegen.	Es findet die **Formel 5'** Verwendung
	$\frac{\mathrm{nCr}(6,1)\cdot\mathrm{nCr}(4,4)}{\mathrm{nCr}(10,5)}$ $\frac{\mathrm{nCr}(6,2)\cdot\mathrm{nCr}(4,3)}{\mathrm{nCr}(10,5)}$ $\frac{\mathrm{nCr}(6,3)\cdot\mathrm{nCr}(4,2)}{\mathrm{nCr}(10,5)}$
Eine Wahrscheinlichkeit für 0 schwarze Kugeln kann es nicht geben, da es nur 4 weiße Kugeln gibt, muss mindestens 1 schwarze dabei sein.	
Mit einer Wahrscheinlichkeit von 0,73811 erhält man höchstens 3 schwarze Kugeln.	

4.2.3. Wie groß ist die Wahrscheinlichkeit, *mindestens* 3 Treffer zu erzielen?

Hier muss für alle in Frage kommenden **k** die Trefferwahrscheinlichkeit mit Formel 4 berechnet und die Einzelergebnisse addiert werden.

Hier müssen die Wahrscheinlichkeiten aller möglichen Treffer summiert werden, die zwischen k = 3 und k = 5 liegen.	Oder es findet die **Formel 6** Verwendung
	$\frac{nCr(6,3)\cdot nCr(4,2)}{nCr(10,5)}$ $\frac{nCr(6,4)\cdot nCr(4,1)}{nCr(10,5)}$ $\frac{nCr(6,5)\cdot nCr(4,0)}{nCr(10,5)}$
Mit einer Wahrscheinlichkeit von 0,73811 erhält man höchstens 3 schwarze Kugeln.	

Dass diese Wahrscheinlichkeit mit der vorherigen gleich ist, ist Zufall.

Vertauscht man die beiden Binomialkoeffizienten im Zähler in ihrer Reihenfolge, ergibt sich die Wahrscheinlichkeit des Gegenereignisses $P(Y \leq 2)$: Wie groß ist die Wahrscheinlichkeit, dass bei 5 Kugeln **höchsten** 2 weiße Kugeln gezogen werden. Damit werden 0, 1 oder 2 weiße Kugeln gezogen. Das sind die Zahlen im zweiten Binomialkoeffizienten, bei dem als obere Zahl eine 4 steht, da es 4 weiße Kugeln gibt.

5. Ziehen ohne zurücklegen mit Beachtung der Reihenfolge

Für diese Verteilung gibt es keinen gesonderten Namen, wird wohl nicht so oft gebraucht. Bei der Binomialverteilung, Ziehen mit Zurücklegen, hat man gesehen, dass nur der Binomialkoeffizient vor den Wahrscheinlichkeiten entfallen muss und schon hat man die Wahrscheinlichkeit für einen Pfad. In der hypergeometrischen Verteilung ist die Anzahl der Pfade mit in der Formel enthalten und nicht als extra Faktor vorhanden. Damit kann man auch nicht so leicht auf einen Pfad zurückgreifen. In der Herleitung der hypergeometrischen Verteilung am Anfang des Dokuments und zu Beginn des Kapitels 4 wurde auf diese Formeln verwiesen.

Formel 2'

$$\frac{\frac{S!}{(S-s)!} \cdot \frac{W!}{(W-w)!}}{\frac{N!}{(N-n)!}}$$

Es handelt sich dabei um die Kombinatorikformel für „mit Reihenfolge, ohne Wiederholung“. Für einen einzelnen Pfad ist die Wahrscheinlichkeit natürlich das gleiche, wie „mit Reihenfolge“, da in einem Pfad nur eine Reihenfolge existiert. „Ohne Wiederholung“ resultiert daraus, dass die Kugeln nicht wieder zurückgelegt werden.

Es soll jetzt das einführende Beispiel zur hypergeometrischen Verteilung unter „Beachtung der Reihenfolge“ untersucht werden.

Wenn man nun die Aufgabe lösen will, die Trefferwahrscheinlichkeit für das Ereignis X mit der Wahrscheinlichkeit p **genau k Treffer** zu ermitteln, dann muss man nach der Formel 2' berechnen.

Beispiel 1

Befinden sich im Gefäß 24 schwarze und 12 weiße Kugeln. Es wird 10 mal gezogen und es soll die Wahrscheinlichkeit bestimmt werden, 6 schwarze und 4 weiße in einer bestimmten Reihenfolge zu ziehen.

N = Anzahl der gesamten Elemente : 36
n = Anzahl der gezogenen Elemente : 10

S = Anzahl der positiven Ereignisse für X : 24
s = Anzahl der gezogenen positiven Ereignisse für X : 4

W = Anzahl der positiven Ereignisse für Gegenereignis Y : 12
w = Anzahl der gezogenen positiven Ereignisse für Gegenereignis Y : 6

Diese Verteilung ist ebenfalls von mehreren Parametern abhängig. Zur Bezeichnung der Verteilung wird hier folgende Reihenfolge verwendet:

1. Parameter : Gesamtzahl der Elemente
2. Parameter : Gesamtzahl der gezogenen Elemente
3. Parameter : Gesamtzahl der positiven Ereignisse für X (beim Gegenereignis für Y)
   Variable in der Klammer : Anzahl der gezogenen positiven Ereignisse

$$P_{N;n;S}\,(s)$$

Zunächst soll die Wahrscheinlichkeit über das Zählprinzip durch elementare Rechnung bestimmt werden. Von den 36 vorhandenen Kugeln werden 10 gezogen. Es sollen die Kugeln mit der Eigenschaft X durch x und die mit der Eigenschaft Y durch y bezeichnet werden. Gesucht ist die Wahrscheinlichkeit für folgende Reihenfolge:

x , x , y , y , x , y , y , y , y , x

Es existieren 24 Kugeln mit der Eigenschaft X und 12 Kugeln mit der Eigenschaft Y. Damit ergibt sich für die Wahrscheinlichkeit in der angegebenen Reihenfolge:

$$P_{36,10,24}(X=4)=\frac{24}{36}\cdot\frac{23}{35}\cdot\frac{12}{34}\cdot\frac{11}{33}\cdot\frac{22}{32}\cdot\frac{10}{31}\cdot\frac{9}{30}\cdot\frac{8}{29}\cdot\frac{7}{28}\cdot\frac{21}{27}$$

Da es sich bei den Brüchen um Faktoren handelt, kann die Reihenfolge der Faktoren beliebig geändert werden, ohne, dass sich der Wert des Bruches verändert. Es werden deshalb die Faktoren, die die Eigenschaft X darstellen und die Faktoren, die die Eigenschaft Y darstellen zusammengezogen. Im Nenner bleibt die absteigende Reihenfolge erhalten.

$$P_{36,10,24}(X=4)=\frac{24}{36}\cdot\frac{23}{35}\cdot\frac{22}{34}\cdot\frac{21}{33}\cdot\frac{12}{32}\cdot\frac{11}{31}\cdot\frac{10}{30}\cdot\frac{9}{29}\cdot\frac{8}{28}\cdot\frac{7}{27}$$

Durch das Vertauschen der Faktoren folgt im Sinne der Wahrscheinlichkeitsrechnung der Schluß, man muß die Wahrscheinlichkeit nicht in der geforderten Reihenfolge bestimmen, sondern man erhält die gleiche Wahrscheinlichkeit, wenn man die Ereignisse nach der Anzahl ihres Auftretens sortiert. Pfade mit der gleichen Anzahl von Ereignissen haben unabhängig von ihrer Reihenfolge immer die gleiche Wahrscheinlichkeit.

Aus der Formel 2' ergibt sich folgende Berechnung:

Trefferwahrscheinlichkeit für genau 5 Treffer: P(X = 4) oder P(Y = 6)

Ereignis Gegenereignis

$$P_{36;10;24}(X=4) = \frac{\frac{24!}{(24-4)!}\cdot\frac{12!}{(12-6)!}}{\frac{36!}{(36-10)!}}$$

$$P_{36;10;24}(X=4) = \frac{\frac{24\cdot23\cdot22\cdot21\cdot\ldots\cdot2\cdot1}{20\cdot19\cdot18\cdot17\cdot\ldots\cdot2\cdot1}\cdot\frac{12\cdot11\cdot10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{6\cdot5\cdot4\cdot3\cdot2\cdot1}}{\frac{36\cdot35\cdot34\cdot33\cdot\ldots\cdot2\cdot1}{26\cdot25\cdot24\cdot\ldots\cdot2\cdot1}}$$

$$P_{36;10;24}(X=4) = \frac{24\cdot23\cdot22\cdot21\cdot12\cdot11\cdot10\cdot9\cdot8\cdot7}{36\cdot35\cdot34\cdot33\cdot\ldots\cdot28\cdot27}$$

Diese Formel entspricht aber genau der Formel, die durch das elementare Berechnen ebenfalls entstanden ist.

Mit dem GTR lassen sich die Ausdrücke $\frac{n!}{(n-k)!}$ über nPr berechnen. Dabei ist als erste Zahl n und als zweite Zahl k einzugeben, nicht n – k vorher berechnen.

$$\frac{\text{nPr}(24{,}4)\cdot\text{nPr}(12{,}6)}{\text{nPr}(36{,}10)}$$

Auf die weiteren Berechnungen für die unteren Summen und die oberen Summen wird hier verzichtet, da sie entsprechend der hypergeometrischen Verteilung unter Benutzung der geänderten Formel durchzuführen sind. Alle Berechnungen sind mit dem GTR einzeln durchzuführen, eine nutzbare Verteilungsfunktion gibt es dafür nicht.

