

Verschiedene Binomialverteilungen und ihre Bedeutung für die Inferenzstatistik

In dem einführenden Dokument zur Inferenzstatistik (ein „quickly prototyping“) wurde die Frage gestellt: Können wir bei einem Spieler, der bei beispielsweise 1000 Versuchen hintereinander einen 6er würfelt, noch länger an einen Zufall glauben? Wegen der Unwahrscheinlichkeit der Unschuldsbehauptung haben wir diese verworfen - die Wahrscheinlichkeit, 1000 mal hintereinander eine 6 zu würfeln, beträgt $(1/6)^{1000}$ - und in dem Beispiel dem Spieler den weiteren Zugang zum Spielbetrieb verwehrt. Das Beispiel verfolgte freilich nur eine didaktische Zielsetzung: es sollte vorbereitend erläutert werden, dass die Unschuldsbehauptung (wir haben sie auch die Nullhypothese genannt) dann verworfen wird, wenn die für sie berechnete Wahrscheinlichkeit sehr gering ist.

Tatsächlich würden Sie im vorliegenden Falle eines Falschspielers nach anderen Kriterien vorgehen. Stellen Sie sich vor, Sie würden dem Spieler bei seiner Tätigkeit zuschauen. Würde er Ihnen auch dann verdächtig vorkommen, wenn er von 1000 Würfeln 999-mal eine sechs würfelt? Oder nur 900-mal? Oder 850-mal? Vermutlich würden Sie in all den vier genannten Fällen einen Verdacht haben. Verdächtig ist der Spieler, wenn er also beispielsweise bei 1000-mal Würfeln 1000-mal, aber auch wenn er 999-mal, 900-mal und auch 850-mal eine sechs würfelt. Nicht die Einzelwahrscheinlichkeit dafür, gerade 1000-mal hintereinander eine sechs zu würfeln, ist für Sie ausschlaggebend, sondern die Wahrscheinlichkeit eines ganz Intervalls von verschiedenen Münzwürfen. Sie wollen beispielsweise wissen, wie wahrscheinlich ist es, 800-mal, 801-mal usw. bis 1000-mal bei 1000 Würfeln eine sechs zu bekommen. Woran Sie also interessiert sind, ist die Frage, 800-mal *oder mehr* eine 6 zu würfeln. Die Wahrscheinlichkeit eines derartigen Intervalls bestimmen Sie mit Hilfe der Wahrscheinlichkeitsfunktion der Binomialverteilung. Sie wird auch als Überschreitungswahrscheinlichkeit bezeichnet.

Betrachten wir hierzu zwei kleine Beispiele von Binomialverteilungen:

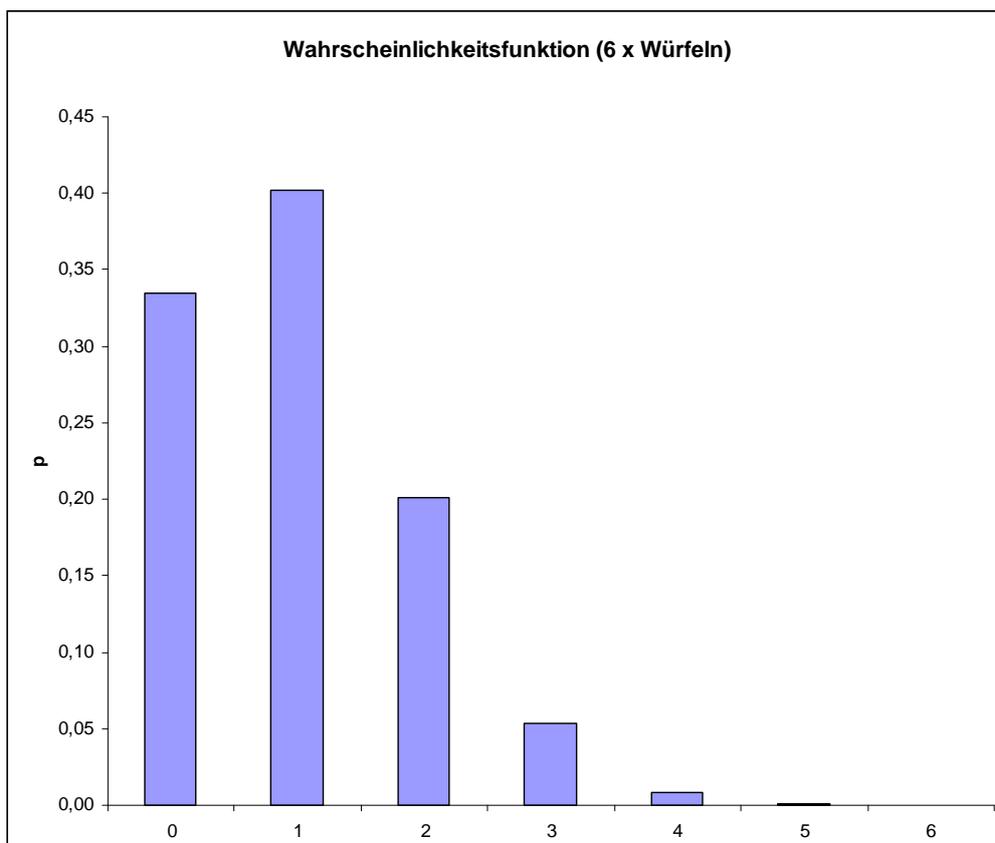
- 1) Die Wahrscheinlichkeitsfunktion der Binomialverteilung, bei 6-maligem Würfeln gerade k mal eine 6 zu bekommen
- 2) bei 4 Münzwürfen k mal Wappen zu bekommen

ad 1) $n = 6$; $p = 1/6$; $q = 5/6$

$$P(k|6) = \binom{6}{k} \cdot \left(\frac{1}{6}\right)^k \cdot \left(\frac{5}{6}\right)^{n-k}$$

0	$\binom{6}{0} \cdot \left(\frac{1}{6}\right)^0 \cdot \left(\frac{5}{6}\right)^6 = 0.334897977$
1	$\binom{6}{1} \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{5}{6}\right)^5 = 0.401877572$
2	$\binom{6}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^4 = 0.200938786$
3	$\binom{6}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^3 = 0.053583676$
4	$\binom{6}{4} \cdot \left(\frac{1}{6}\right)^4 \cdot \left(\frac{5}{6}\right)^2 = 0.008037551$
5	$\binom{6}{5} \cdot \left(\frac{1}{6}\right)^5 \cdot \left(\frac{5}{6}\right)^1 = 0.000643004$
6	$\binom{6}{6} \cdot \left(\frac{1}{6}\right)^6 \cdot \left(\frac{5}{6}\right)^0 = 0.000021433$

Die Summe dieser Einzelwahrscheinlichkeiten ist, wie man leicht berechnen kann, 1!



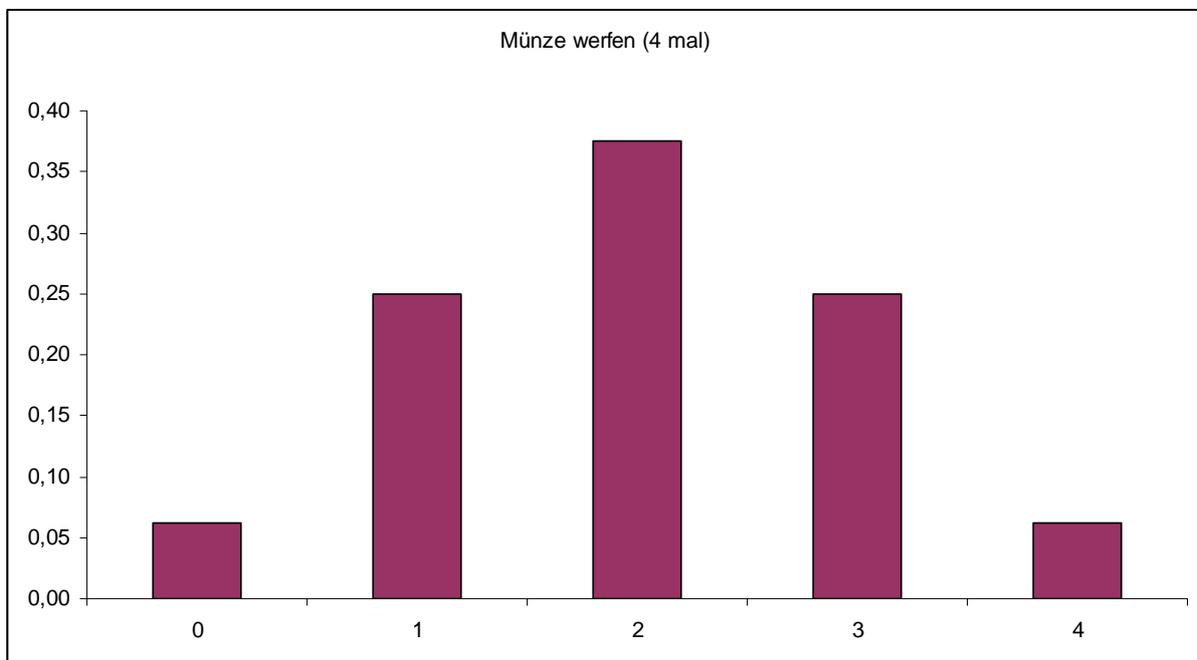
ad 2) bei 4 Münzwürfen k mal Wappen zu bekommen

$n = 4$; $p=1/2$; $q=1/2$

$$P(k|4) = \binom{4}{k} \cdot \left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{2}\right)^{n-k}$$

0	$\binom{4}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^4 = 0.0625$
1	$\binom{4}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^3 = 0.25$
2	$\binom{4}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^2 = 0.375$
3	$\binom{4}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^1 = 0.25$
4	$\binom{4}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^0 = 0.0625$

Summe der P = 1!



Die letzten beiden Beispiele waren Beispiele für Binomialverteilungen mit verschiedenem n und verschiedenem p . Die allgemeine Formel

$$\binom{n}{k} \cdot p^k \cdot q^{n-k}$$

beschreibt also eine ganze Familie von Binomialverteilungen, abhängig von den Parametern n und p .

Wie berechnet man nun die Überschreitungswahrscheinlichkeit?

Wollen wir z.B. wissen, wie wahrscheinlich es ist, bei insgesamt 6-maligem Würfeln 5-mal oder mehr eine 6 zu bekommen, so lässt sich dies mit Hilfe der oben angegebenen Binomialverteilung (1. Beispiel) so berechnen: Wie zählen die Einzelwahrscheinlichkeiten für 5mal und 6mal eine 6 zu bekommen zusammen, also: $0.000643004 + 0.000021433 = 0,000664437$

Das gleiche Ergebnis (bis auf Rundungsfehler) bekommen wir, indem wir die Einzelwahrscheinlichkeiten für 1mal, 2mal, 3mal und 4mal eine 6 zu würfeln aufaddieren (das ergibt in der Summe $0,999335562$) und die erhaltene Summe dann von 1 (man beachte: die Summe aller Einzelwahrscheinlichkeiten beträgt 1) abziehen, nämlich: $1 - 0,999335562 = 0,000664438$

Diese zwei verschiedenen Berechnungsarten der Überschreitungswahrscheinlichkeit werden wir bei Binomialverteilungen mit einem hohen n benötigen.

Wollen wir die Wahrscheinlichkeit dafür berechnen, bei 1000-maligem Würfeln 800 mal mehr eine 6 zu bekommen, so wird diese Berechnung mit Hilfe der Wahrscheinlichkeitsfunktion der Binomialverteilung ziemlich aufwändig:

$$\binom{1000}{800} \cdot \left(\frac{1}{6}\right)^{800} \cdot \left(\frac{5}{6}\right)^{200} + \binom{1000}{801} \cdot \left(\frac{1}{6}\right)^{801} \cdot \left(\frac{5}{6}\right)^{199} + \binom{1000}{802} \cdot \left(\frac{1}{6}\right)^{802} \cdot \left(\frac{5}{6}\right)^{198} + \dots$$

$$\binom{1000}{1000} \cdot \left(\frac{1}{6}\right)^{1000} \cdot \left(\frac{5}{6}\right)^0$$

Was in einem solchen Falle zu tun ist, wird verständlicher, wenn wir uns zunächst bestimmte Eigenschaften der Binomialverteilung bei großem n näher ansehen.