Frank Klinker

Stochastik

Teil 6: Vom Bernoulli-Experiment zur Binomialverteilung

1 Bernoulli-Experiment und Bernoulli-Kette

1.1 Bernoulli-Experiment

Ein **Bernoulli-Experiment** ist eine Zufallsexperiment, dessen Ergebnismenge aus genau zwei Ergebnissen besteht: $\Omega = \{e, f\}$, also $f = \bar{e}$ und umgekehrt. Mit $P(e) = p$ und $P(f) = q = 1 - p$ nennt man p auch die **Erfolgswahrscheinlichkeit**.

Diese Begriffe erklären sich durch die folgenden oder ganz ähnliche Spielbeispiele:

Beispiel 1.
1. Ein typisches Bernoulli-Experiment ist der Wurf einer idealen Münze mit $p = q = \frac{1}{2}$.
2. Ein weiteres Standardbeispiel ist der Würfelwurf mit $e :$ "Werfen einer sechs" und $f :$ "Werfen keiner sechs" mit $p = \frac{1}{6}$.

1.2 Bernoulli-Kette

Unter einer **Bernoulli-Kette** versteht man ein mehrstufiges Zufallsexperiment, bei dem jede einzelne Stufe das gleiche Bernoulli-Experiment ist. Die Anzahl der Stufen nennt man die **Länge der Bernoulli-Kette**.

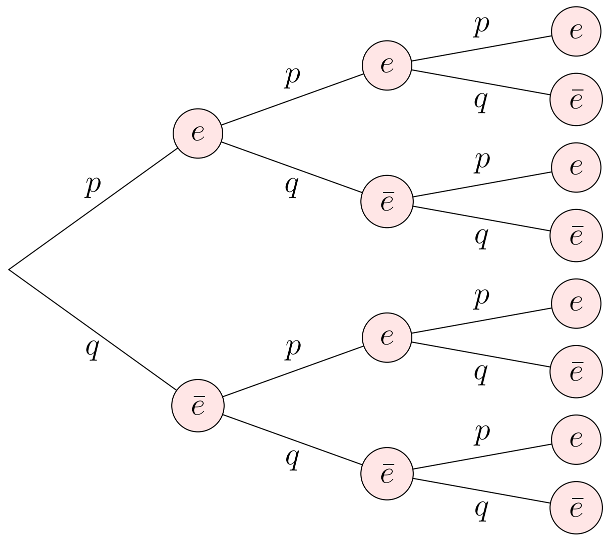
Eine Bernoulli-Kette ist also durch die Erfolgswahrscheinlichkeit p des Bernoulli-Experiments und der Länge n der Kette festgelegt.

Beispiel 2.
1. Hundertmal die gleiche Münze (oder den gleichen Würfel) werfen, ist eine Bernoulli-Kette der Länge $n = 100$ mit Erfolgswahrscheinlichkeit $p = \frac{1}{2}$ (oder $p = \frac{1}{6}$).
2. Aus der Produktion eines Bauteils, bei der mit der Wahrscheinlichkeit $p = 0{,}005$ ein Bauteil defekt ist, entnimmt man als Stichprobe zufällig 1500 Bauteile. Das ist eine Bernoulli-Kette der Länge $n = 1500$ mit der Erfolgswahrscheinlichkeit $p = 0{,}005$.

Bezeichnung 3. Wegen der sehr wichtigen Anwendung aus Beispiel 2.2 verwendet man oft die Begriffe Länge und Stichprobenumfang synonym. Ebenso spricht man statt von Erfolgswahrscheinlichkeit auch von Fehlerwahrscheinlichkeit.

Adresse: Eduard-Spranger-Berufskolleg, 59067 Hamm
E-Mail: mail@frank-klinker.de
Version: 7. Januar 2026

Da eine Bernoulli-Kette ein mehrstufiges Zufallsexperiment ist, kann man sie mit Hilfe eines Baumes darstellen (hier $n = 3$):



- Ein Ergebnis einer Bernoulli-Kette mit den Parametern n und p kann so durch einen vollständigen Pfad dieses Baumes beschrieben werden.
- Es reicht allerdings die einzelnen Ergebnisse der n Bernoulli-Experimente hintereinander zu schreiben. So lässt sich dann ein Ergebnis der Bernoulli-Kette mit einem n-Tupel identifizieren, dessen Einträge e und $\bar{e}$ sind.

Es gibt also insgesamt 2^n verschiedene Ergebnisse und die **Wahrscheinlichkeit eines Ergebnisses der Bernoulli-Kette** ist wegen der Pfadmultiplikationsregel

$$p^k q^{n-k} \tag{$*$}$$

wenn in dem Pfad (dem n-Tupel) genau k-mal ein e vorkommt. Damit ist insbesondere festgelegt, dass genau $(n-k)$-mal ein $\bar{e}$ in dem Pfad (dem n-Tupel) vorkommt.

2 Die Binomialverteilung

2.1 Die Wahrscheinlichkeitsverteilung einer Bernoulli=Kette

Im Zusammenhang mit Beispiel 2.2 – der aus der Produktion entnommenen Stichprobe –, kann man folgende exemplarische aber sehr natürliche Fragestellung formulieren:

Beispiel 4. "Mit welcher Wahrscheinlichkeit befinden sich unter den 1500 zufällig ausgewählten Bauteilen ...

a) ... genau 10 defekte Bauteile"
b) ... höchstens 10 defekte Bauteile"
c) ... mindestens 5 defekte Bauteile"
d) ... mindestens 5 und höchstens 10 defekte Bauteile"

Zur Berechnung dieser Wahrscheinlichkeiten nutzen wir eine Zufallsvariable, welche die Erfolge in einem Ergebnis der Bernoulli-Kette zählt:

$$X : \text{Ergebnis der Bernoulli-Kette} \mapsto \text{Anzahl der } e \text{ in dem Ergebnis}$$

Wir starten mit einem Beispiel, das die Idee sehr gut darstellt und das sich ohne Probleme verallgemeinern lässt.

Beispiel 5. Wir betrachten eine Bernoulli-Kette der Länge $n = 4$. Alle Ergebnisse sind damit (wir nutzen hier eine etwas kompaktere Schreibweise der Tupel)

$$\begin{gathered} eeee,\ eee\bar{e},\ ee\bar{e}e,\ e\bar{e}ee,\ \bar{e}eee,\ ee\bar{e}\bar{e},\ e\bar{e}\bar{e}e,\ \bar{e}\bar{e}ee, \\ e\bar{e}e\bar{e},\ \bar{e}e\bar{e}e,\ \bar{e}ee\bar{e},\ e\bar{e}\bar{e}\bar{e},\ \bar{e}e\bar{e}\bar{e},\ \bar{e}\bar{e}e\bar{e},\ \bar{e}\bar{e}\bar{e}e,\ \bar{e}\bar{e}\bar{e}\bar{e}\,. \end{gathered}$$

Hier kann X die Werte $k = 0, 1, 2, 3$ und 4 annehmen. Wenn wir sortieren, wie die Zuordnung der Ergebnisse zu diesen Werten ist, dann erhalten wir folgende Tabelle:

k	$X^{-1}(k)$	$\#X^{-1}(k)$
0	$\bar{e}\bar{e}\bar{e}\bar{e}$	$1 = \binom{4}{0}$
1	$e\bar{e}\bar{e}\bar{e}$, $\bar{e}e\bar{e}\bar{e}$, $\bar{e}\bar{e}e\bar{e}$, $\bar{e}\bar{e}\bar{e}e$	$4 = \binom{4}{1}$
2	$ee\bar{e}\bar{e}$, $e\bar{e}\bar{e}e$, $\bar{e}\bar{e}ee$, $e\bar{e}e\bar{e}$, $\bar{e}e\bar{e}e$, $\bar{e}ee\bar{e}$	$6 = \binom{4}{2}$
3	$eee\bar{e}$, $ee\bar{e}e$, $e\bar{e}ee$, $\bar{e}eee$	$4 = \binom{4}{3}$
4	$eeee$	$1 = \binom{4}{4}$

In der rechten, ergänzten Spalte steht die Anzahl der Ergebnisse, die auf k abgebildet werden. Dass es hier genau die Binomialkoeffizienten sind, ist kein Zufall:

$$\boxed{\binom{4}{k} \text{ ist die Anzahl an 4-Tupeln mit genau } k\text{-mal } e}$$

Damit und mit dem Ergebnis aus $(*)$ können wir in der Tabelle die Wahrscheinlichkeiten $P(X = k)$ problemlos ergänzen:

k	$X^{-1}(k)$	$\#X^{-1}(k)$	$P(X = k)$
0	$\bar{e}\bar{e}\bar{e}\bar{e}$	1	$q^4 = \binom{4}{0}p^0q^{4-0}$
1	$e\bar{e}\bar{e}\bar{e}$, $\bar{e}e\bar{e}\bar{e}$, $\bar{e}\bar{e}e\bar{e}$, $\bar{e}\bar{e}\bar{e}e$	4	$4pq^3 = \binom{4}{1}p^1q^{4-1}$
2	$ee\bar{e}\bar{e}$, $e\bar{e}\bar{e}e$, $\bar{e}\bar{e}ee$, $e\bar{e}e\bar{e}$, $\bar{e}e\bar{e}e$, $\bar{e}ee\bar{e}$	6	$6p^2q^2 = \binom{4}{2}p^2q^{4-2}$
3	$eee\bar{e}$, $ee\bar{e}e$, $e\bar{e}ee$, $\bar{e}eee$	4	$4p^3q = \binom{4}{3}p^3q^{4-3}$
1	$eeee$	4	$p^4 = \binom{4}{4}p^4q^{4-4}$

Die naheliegende Verallgemeinerung fassen wir wie folgt zusammen:

Binomialverteilung

In einer Bernoulli-Kette der Länge n und der Erfolgswahrscheinlichkeit p ist die Wahrscheinlichkeit für genau k Erfolge

$$B_{n,p}(k) = P(X = k) = \begin{cases} \binom{n}{k} p^k q^{n-k} & \text{für } k = 0, 1, \ldots, n \\ 0 & \text{sonst} \end{cases}$$

(dabei ist $q = 1 - p$).

$B_{n,p}$ heißt die **(Wahrscheinlichkeitsfunktion der) Binomialverteilung**.

Damit kann man auch weitere Wahrscheinlichkeiten einer (n, p)-verteilten Binomialverteilung berechnen:

$$\begin{aligned} P(X \leq k) =& P(X = 0) + P(X = 1) + \ldots + P(X = k - 1) + P(X = k) \\ =& B_{n,p}(0) + B_{n,p}(1) + \ldots + B_{n,p}(k - 1) + B_{n,p}(k) \\ =& \sum_{\ell=0}^{k} B_{n,p}(\ell) \end{aligned}$$

Diese summierte Version der Binomialverteilung nennt man auch die **kumulierte Binomialverteilung** und ist auf vielen Taschenrechner und CAS-Systemen implementiert. Wir schreiben dafür kurz $B_{n,p}^{\text{kum}}(k)$, also

$$\boxed{P(X \leq k) = B_{n,p}^{\text{kum}}(k) = \sum_{\ell=0}^{k} B_{n,p}(\ell)}$$

Damit lassen sich weitere Wahrscheinlichkeiten berechnen:

$$\begin{aligned} P(X < k) &= B_{n,p}^{\text{kum}}(k - 1) \\ P(X \geq k) &= 1 - B_{n,p}^{\text{kum}}(k - 1) \\ P(X > k) &= 1 - B_{n,p}^{\text{kum}}(k) \end{aligned}$$

und

$$\begin{aligned} P(k_1 \leq X \leq k_2) &= B_{n,p}^{\text{kum}}(k_2) - B_{n,p}^{\text{kum}}(k_1 - 1) \\ P(k_1 \leq X < k_2) &= B_{n,p}^{\text{kum}}(k_2 - 1) - B_{n,p}^{\text{kum}}(k_1 - 1) \\ P(k_1 < X \leq k_2) &= B_{n,p}^{\text{kum}}(k_2) - B_{n,p}^{\text{kum}}(k_1) \\ P(k_1 < X < k_2) &= B_{n,p}^{\text{kum}}(k_2 - 1) - B_{n,p}^{\text{kum}}(k_1) \end{aligned}$$

Beispiel 6 (Fortsetzung von Beispiel 4)**.** Wir können nun die Wahrscheinlichkeiten aus dem Beispiel berechnen:

a) $P(X = 10) = B_{1500;\,0{,}005}(10) \approx 0{,}0859$ also etwa 8,59%.

b) $P(X \leq 10) = B^{\text{kum}}_{1500;\,0{,}005}(10) \approx 0{,}8628$ also etwa 86,59%.

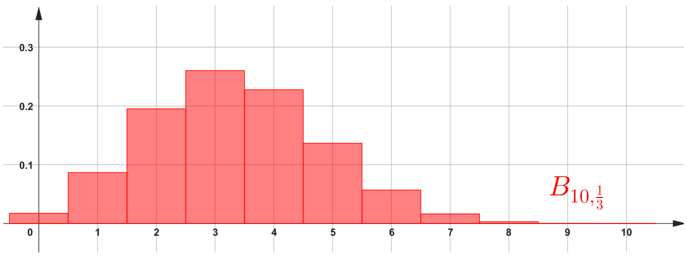
c) $P(X \geq 5) = 1 - B^{\text{kum}}_{1500;\,0{,}005}(4) \approx 1 - 0{,}1314 = 0{,}8686$ also etwa 86,86%.

d) $P(5 \leq X \leq 10) = B^{\text{kum}}_{1500;\,0{,}005}(10) - B^{\text{kum}}_{1500;\,0{,}005}(4) \approx 0{,}8659 - 0{,}1314 = 0{,}7345$ also etwa 73,45%.

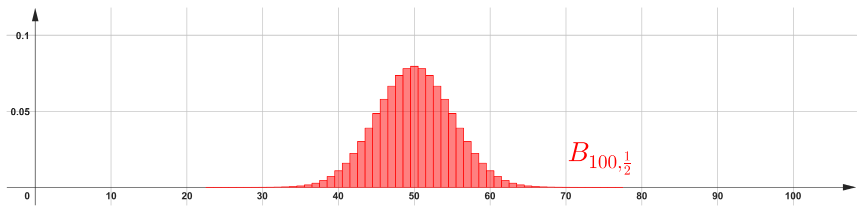
2.2 Die Graphen der Binomialverteilung

Wir sehen uns an, wie die Binomialverteilung $Bn, p(k)$ in Abhängigkeit der Variable k für unterschiedliche Parameter aussieht:

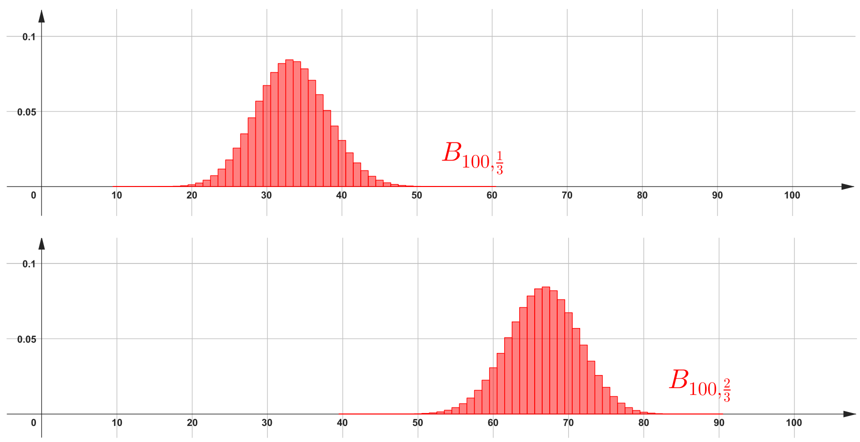
1. $n = 10, p = \frac{1}{3}$



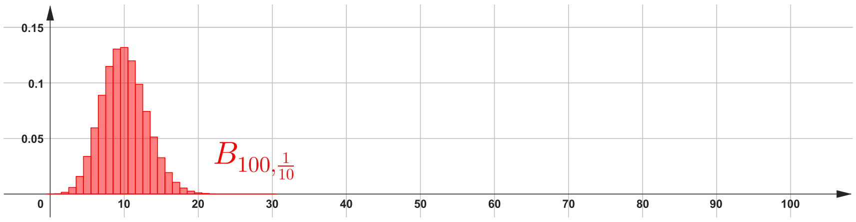
2. $n = 100, p = \frac{1}{2}$

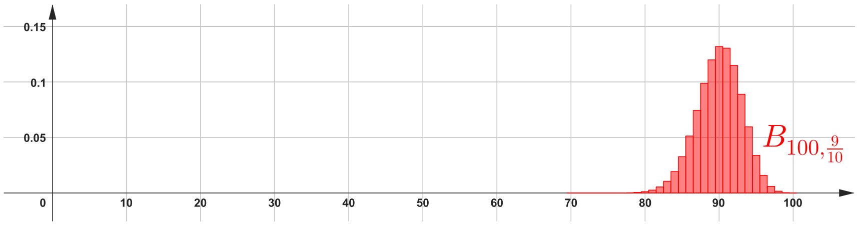


3. $n = 100, p = \frac{1}{3}$ und $n = 100, p = \frac{2}{3}$

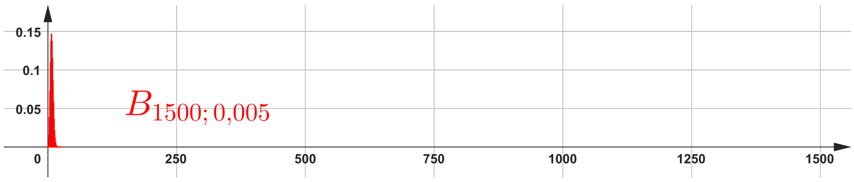


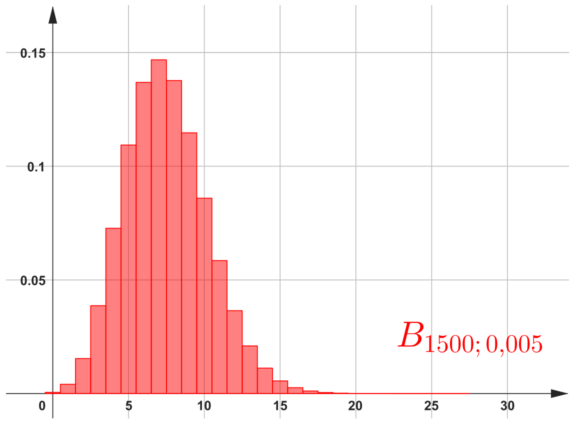
4. $n = 100, p = \frac{1}{10}$ und $n = 100, p = \frac{9}{10}$





5. $n = 1500, p = 0{,}005$ (einmal komplett und einmal nur der "Anfang")





Erste Beobachtungen:

- Ändert man p bei gleichem n, dann "verschiebt" sich der Graph.

  Genauer: die Graphen von $B_{n,p}$ und $B_{n,1-p}$ gehen durch Spiegelung an der Geraden $k = \frac{n}{2}$ ineinander über.
- Vor allem bei großem n wird deutlich, dass die Werte der Binomialverteilung "eng" um einen Wert konzentriert sind: es entsteht eine Glockenform.
- Ist p sehr klein (oder nahe 1), dann ist die Glocke sehr schmal.
- Insbesondere bei großem n ist die Glocke "symmetrisch" um ihr "Zentrum".
- Die Symmetrie der Glocke geht "kaputt" wenn p sehr klein (oder nahe 1) ist.

Das beobachtete "Glockenzentrum" und das "Streuen um dieses Zentrum", lässt sich gut mit Hilfe der Lage- und Streumaße erklären, genauer mit dem Erwartungswert und der Standardabweichung.

3 Erwartungswert, Varianz und Standardabweichung der Binomialverteilung

3.1 Die Ergebnisse

Erwartungswert der Binomialverteilung
Sind n und p die Parameter der Binomialverteilung, also $P(X = k) = B_{n,p}(k)$, dann ist

$$E(B_{n,p}) = n \cdot p$$

der Erwetungswert.

Varianz und Standardabweichung der Binomialverteilung
Sind n und p die Parameter der Binomialverteilung, also $P(X = k) = B_{n,p}(k)$, dann sind

$$V(B_{n,p}) = n \cdot p \cdot q$$

die Varianz und

$$\sigma(B_{n,p}) = \sqrt{n \cdot p \cdot q}$$

die Standardabweichung (dabei ist $q = 1 - p$).

3.2 Begründungen für die Formeln aus 3.1

Wir werden die folgenden Formeln nutzen:

1) Der Erwartungswert einer Zufallsvariablen, die nur endlich viele Werte $\{x_1, \ldots, x_n\}$ annimmt ist durch

$$E(X) = \sum_{i=1}^{n} x_i P(X = x_i)$$

gegeben. In unserem Fall durchlaufen die Werte x_i die Zahlen $0, 1, \ldots, n$, sodass $E(B_{n,p}) = \sum_{k=0}^{n} k \cdot B_{n,p}(k)$, also

$$E(B_{n,p}) = \sum_{k=0} k \binom{n}{k} p^k q^{k-1} \,.$$

Wir müssen also die Summe auf der rechten Seite berechnen.

2) Die Varianz einer Zufallsvariablen, die nur endlich viele Werte $\{x_1, \ldots, x_n\}$ annimmt ist durch

$$V(X) = \sum_{i=1}^{n} \big(x_i - E(X)\big)^2 P(X = x_i)$$

gegeben. Es gilt außerdem die Formel $V(X) = E(X^2) - E(X)^2$, die wir nutzen werden. Mit der allgemeinen Formel $E(X^2) = \sum_{i=1}^{n} x_i^2 P(X = x_i)$ gibt das bei uns mit $E(B_{n,p}) = np$

$$V(B_{n,p}) = \sum_{k=0} k^2 \binom{n}{k} p^k q^{k-1} - (np)^2 \,.$$

Wir müssen also die rechte Seite berechnen.

3) Bei der Untersuchung der Binomialkoeffizienten sind wir auf die für alle Werte p und q gültige Formel

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$$

gestoßen.[1] Betrachtet man beide Seiten als Funktionen mit der Variablen p, dann kann man beide Seiten nach p ableiten. Das gibt dann

$$n(p+q)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} p^{k-1} q^{n-k} \tag{$*_1$}$$

und

$$n(n-1)(p+q)^{n-2} = \sum_{k=0}^{n} k(k-1) \binom{n}{k} p^{k-2} q^{n-k} \tag{$*_2$}$$

3.2.1 Begründung für den Erwartungswert

Wir nehmen Formel ($*_1$) aus 3) und multiplizieren beide Seiten mit p. Das gibt

$$\begin{aligned} np(p+q)^{n-1} &= p \sum_{k=0}^{n} k \binom{n}{k} p^{k-1} q^{n-k} \\ &= \sum_{k=0}^{n} k \binom{n}{k} p p^{k-1} q^{n-k} \\ &= \sum_{k=0}^{n} k \binom{n}{k} p^{k} q^{n-k} \end{aligned}$$

Wenn wir jetzt noch beachten, dass auf der linken Seite $p+q=1$ ist, dann erhalten wir

$$np = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} \,.$$

Das ist mit 1) genau das Ergebnis für den Erwartungswert $E(B_{n,p})$.

[1] Diese gibt uns mit $p+q=1$ die wichtige Vollständigkeitsformel $\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = 1$.

3.2.2 Begründung für Varianz und Standardabweichung

Wir nehmen Formel ($*_2$) aus 3) und multiplizieren beide Seiten mit p^2. Das gibt

$$\begin{aligned} np^2(n-1)(p+q)^{n-2} &= p^2 \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^{k-2} q^{n-k} \\ &= \sum_{k=0}^{n} (k^2-k)\binom{n}{k} p^2 p^{k-2} q^{n-k} \\ &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} - \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} \end{aligned}$$

Die zweite Summe auf der rechten Seite ist wegen Abschnitt 3.2.1 gerade np. Wir bringen diese Summe von der rechten auf die linke Seite und beachten auf der linken Seite wieder, dass $p+q=1$ ist. Das gibt

$$np^2(n-1) + np = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} \,.$$

Wegen $np^2(n-1)+np = n^2p^2 - np^2 + np = (np)^2 + np(1-p) = (np)^2 + npq$ schließlich nachdem man noch $(np)^2$ nach rechts gebracht hat

$$npq = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k q^{n-k} - (np)^2 \,.$$

Das ist mit 2) genau das Ergebnis für die Varianz $V(B_{n,p})$.

Weil die Standardabweichung die Quadratwurzel aus der Varianz ist erhält man damit

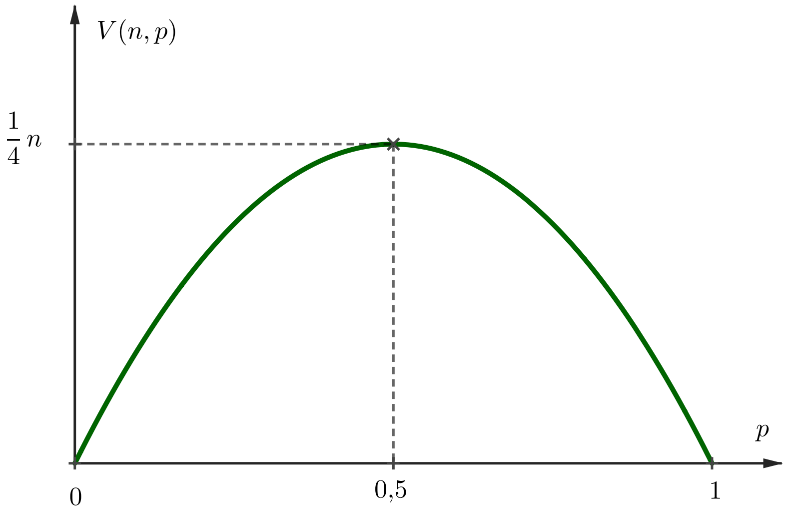
$$\sigma(B_{n,p}) = \sqrt{npq}$$

3.3 Zum Verhalten von $E(B_{n,p})$ und $V(B_{n,p})$ bei änderung der Parameter

Wir wollen uns ansehen, wie sich Erwartungswert und Varianz ändern, wenn man die Parameter n und/oder p ändert. Dazu schreiben wir kürzer $E = E(B_{n,p})$ und $V = V(B_{n,p})$ also $E = np$ und $V = npq = np(1-p)$

- E und V sind beide proportional zu n, sodass sie gleichmässig mit n wachsen und fallen.
- E ist auch proportional zu p, sodass E auch gleichmäßig mit p wächst und fällt.
- V hängt quadratisch von p ab, sodass sich eine Parabel ergibt. Diese hat Nullstellen

bei $p = 0$ und $p = 1$ und ein Maximum bei $p = 0{,}5$ mit dann $V = \frac{n}{4}$:



Dass V achsensymmetrisch zur Geraden $p = 0{,}5$ ist und dort ein Extremum hat, folgt allein schon aus der Tatsache, dass V symmetrisch von p und $q = 1 - p$ abhängt. Das ist also unabhängig davon, dass eine Parabel vorliegt. Dass dieses Extremum ein Maximum ist folgt allerdings aus der genaueren Abhängigkeit.