Satz. *Seien $\mathcal{U}$ und $\mathcal{V}$ binomialverteilt mit $\mathcal{U} \sim Bin(m,p)$ und $\mathcal{V} \sim Bin(n,q)$. Dann sind äquivalent:*

- *$\mathcal{U} + \mathcal{V}$ ist binomialverteilt*
- *$p = q$.*

Beweis. Wenn $p = q$ ist, dann zählt $\mathcal{U}$ die Erfolge von m unabängigen Bernoulli-Experimenten mit Erfolgswahrscheinlichkeit p und $\mathcal{V}$ zählt die Erfolge von n solcher Experimente. Also zählt die Summe die Erfolge von $m+n$ solcher Experimente und ist deshalb $Bin(m,p)$-verteilt.

Wir zeigen nun, dass $\mathcal{U} + \mathcal{V}$ nicht binomialverteilt ist, wenn wir $p \neq q$ haben. Ohne Einschränkung können wir annehmen, dass $p < q$ ist. Die Varianzen von $\mathcal{U}$ und $\mathcal{V}$ sind $\sigma^2_{\mathcal{U}} = mp(1-p)$ und $\sigma^2_{\mathcal{V}} = nq(1-q)$.

Angenommen $\mathcal{W} = \mathcal{U} + \mathcal{V}$ wäre binomialverteilt. Dann wären die möglichen Werte von $\mathcal{W}$ die Zahlen zwischen 0 und $m+n$ und $\mathcal{W}$ würde die Anzahl der Erfolge von $m+n$ Bernoulli-Experimenten mit einer Wahrscheinlichkeit r zählen. Also wäre $\mathcal{W} \sim Bin(m+n, r)$.

Dann hätte $\mathcal{W}$ die Varianz $\sigma^2_{\mathcal{W}} = (m+n)r(1-r)$. Andererseits, weil $\mathcal{U}$ und $\mathcal{V}$ unabhängig sind, ist die Varianz von $\mathcal{W}$ gleich der Summe der Varianzen von $\mathcal{U}$ und $\mathcal{V}$, also $\sigma^2_{\mathcal{W}} = \sigma^2_{\mathcal{U}} + \sigma^2_{\mathcal{V}}$, und damit

$$(m+n)r(1-r) = mp(1-p) + nq(1-q)$$

also

$$r(1-r) = \frac{m}{m+n}p(1-p) + \frac{n}{m+n}q(1-q).$$

Sei f die Funktion $f\colon [0{,}1] \to \mathbb{R}$, $f(x) = x(1-x)$. Dann steht oben

$$f(r) = \frac{m}{m+n}f(p) + \frac{n}{m+n}f(q). \tag{1}$$

Wir wissen über die Erwartungswerte, dass

$$\begin{aligned} E[\mathcal{U}] &= mp \\ E[\mathcal{V}] &= nq \\ E[\mathcal{W}] &= (m+n)r. \end{aligned}$$

Daher ist

$$(m+n)r = E[\mathcal{W}] = E[\mathcal{U} + \mathcal{V}] = E[\mathcal{U}] + E[\mathcal{V}] = mp + nq,$$

also

$$r = \frac{m}{m+n}p + \frac{n}{m+n}q. \tag{2}$$

Die Gleichungen (1) und (2) haben eine gewisse Ähnlichkeit. Gleichung (2) sagt, dass r eine Konvexkombination von p und q ist, und Gleichung (1) sagt, dass $f(r)$ eine Konvexkombination von $f(p)$ und $f(q)$ ist mit den gleichen Koeffizienten.

Aus der linearen Algebra wissen wir: Die Punkte in der Ebene auf der Strecke zwischen den Punkten

$$\begin{bmatrix} p \\ f(p) \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} q \\ f(q) \end{bmatrix}$$

haben die Form

$$\alpha \begin{bmatrix} p \\ f(p) \end{bmatrix} + \beta \begin{bmatrix} q \\ f(q) \end{bmatrix},$$

mit $0 \leq \alpha, \beta \leq 1$ und $\alpha + \beta = 1$.

Wegen ${}^{m}/_{m+n} + {}^{n}/_{m+n} = {}^{m+n}/_{m+n} = 1$ erfüllen auch $\alpha = {}^{m}/_{m+n}$ und $\beta = {}^{n}/_{m+n}$ die Bedingung $0 \leq \alpha, \beta \leq 1$ und $\alpha + \beta = 1$. Wegen Gleichung (2) ist $r = \alpha p + \beta q$ und wegen Gleichung (1) ist $f(r) = \alpha f(p) + \beta f(q)$.

Geometrisch bedeutet dies, dass der Punkt

$$\begin{bmatrix} r \\ f(r) \end{bmatrix} = \begin{bmatrix} \alpha p + \beta q \\ \alpha f(p) + \beta f(q) \end{bmatrix}$$

auf der Geraden durch die Punkte $(p, f(p)$ und $(q, f(q))$ liegt.

Andererseits ist $f''(x) = \frac{d^2}{dx^2}\, x(1-x) = -2$. Die Funktion f ist also konkav. Sind p, q mit $p < q$ Werte im Definitionsintervall von f, und ist $g_{p,q}$ die Gerade durch $(p, f(p))$ und $(q, f(q))$ (d.h., die Sekante durch f für p und q), dann liegen im Intervall (p, q) die Funktionswerte von f, da f konkav ist, über der Geraden $g_{p,q}$. Das heißt $g_{p,q}(r) < f(r)$ für alle $r \in (p, q)$.

Wir haben also den Widerspruch hergeleitet, dass einerseits wegen der Gesetze für Erwartungswert und Varianz der Punkt $(r, f(r))$ auf der Geraden durch $(p, f(q))$ und $(q, f(q))$ liegen muss, dies andererseits wegen der Konkavität von f nicht möglich ist.

Abbildung 2 zeigt die Funktion f und die Sekante durch die Punkte $(p, f(p)$ und $(q, f(q))$ für $p = 0{,}1$ und $q = 0{,}8$. Sie stammt aus einer Visualisierung) auf DESMOS, die f zeigt und in der man p, $q \in [0{,}1]$ beliebig wählen kann und dann die entsprechende Sekante durch den Graphen von f sieht. $\square$

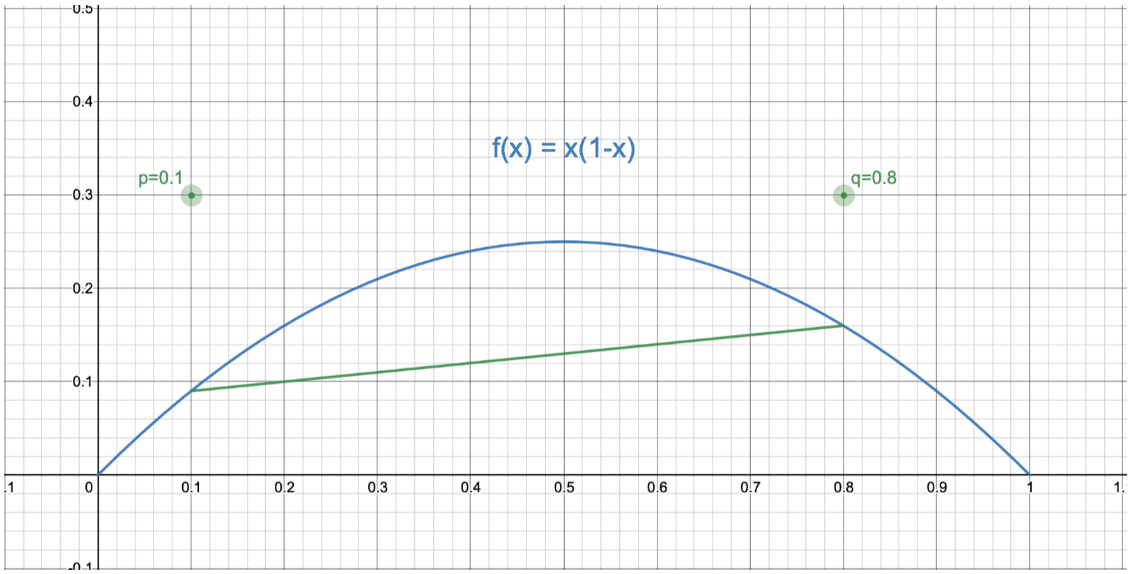

Abbildung 2: Sekante durch den Graphen von $f(x) = x(1-x)$.