Beweis des Satzes von de Moivre-Laplace

Paul Ruppen

23. Juni 2007

Der folgende Beweis ist eine Ausarbeitung eines Beweises in Ulrich Krengel, (1998), *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, Wiesbaden, Vieweg. Die Stirlingsche Formel wird dabei mitbewiesen. Die aus der Analysis bekannten Sätze $\int\limits_{-\infty}^{\infty} \exp(-\frac{x}{2})^2 dx = \sqrt{2\pi}$ und der *Satz von Taylor* werden hingegen beweislos übernommen. Ebenso wird der *Satz von Tschebyscheff* vorausgesetzt. Statt "$\lim\limits_{n\to\infty} A = b$" wird manchal "$A \to b$" geschrieben. Der Satz von de Moivre-Laplace ist ein Spezialfall des Zentralen Grenzwertsatzes. Der direkte Beweis des Satzes stellt allerdings eine Herleitung der Dichtefunktion der Standardnormalverteilung dar.

Der Satz von de Moivre-Laplace lautet:

Satz 1. *Für $p \in]0,1[$ und $X_n \sim B(n,p)$ für $n \in \mathbb{N}$ gilt:*

$$\lim_{n\to\infty} P(a \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b) = \Phi(b) - \Phi(a),$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist ("$X_n \sim B(n,p)$" bedeutet "X_n ist binomialverteilt mit den Parametern n und p").

Zuerst wird eine Formel für die näherungsweise Berechnung von Fakultäten hergeleitet. Dazu brauchen wir die folgenden zwei Sätze:

Satz 2. *Wenn für $x \in [a-0.5; a+0.5]$ $f(x) \geq 0$ und $f(x)$ konkav ist, dann ist*

$$f(a) \geq \int\limits_{a-0.5}^{a+0.5} f(x)dx.$$

Beweis. Es gilt, dass $f(a) = \int\limits_{a-0.5}^{a+0.5} f(a)dx = 1 \cdot f(a)$, d.h. $f(a)$ ist die Fläche des Rechteckes A auf der Basis $[a-0.5; a+0.5]$ und der Höhe $f(a)$. Jedes Trapez auf $[a-0.5; a+0.5]$, das aus dem Rechteck A entsteht, indem man die waagrechte Gerade g durch $f(a)$ um einen Winkel $\alpha \in]-90^0, 90^0[$ dreht und für $x \in [a-0.5; a+0.5]$ $g(x) \geq 0$, hat dieselbe Fläche wie A. Entsprechend hat das Trapez auf $[a-0.5; a+0.5]$, das oben durch die Tangente $g(x)$ in $f(a)$ begrenzt ist, die Fläche von A. Zudem gilt wegen der Konkavität von $f(x)$ in $[a-0.5; a+0.5]$ $g(x) \geq f(x)$ in $[a-0.5; a+0.5]$. Somit gilt $f(a) = \int\limits_{a-0.5}^{a+0.5} g(x)dx \geq \int\limits_{a-0.5}^{a+0.5} f(x)dx$. $\square$

Satz 3. *Wenn für $x \in [a-0.5; a+0.5]$ $f(x) \geq 0$ und $f(x)$ konkav ist, dann ist*

$$\int\limits_{a-0.5}^{a+0.5} f(x)dx \geq \frac{1}{2}(f(a-0.5)+f(a+0.5)).$$

Beweis. $\frac{1}{2}(f(a-0.5)+f(a+0.5))$ ist das arithmetische Mittel der Funktionswerte $f(a-0.5)$ und $f(a+0.5)$, wobei $\frac{1}{2}(f(a-0.5)+f(a+0.5))$ auf der Geraden g durch die Werte $f(a-0.5)$ und $f(a+0.5)$ liegt. Es gilt, dass die Fläche des Rechteckes auf der Basis $[a-0.5; a+0.5]$ mit der Höhe $\frac{1}{2}(f(a-0.5)+f(a+0.5))$ gleich dem Trapez auf dieser Basis ist, das oben durch g begrenzt wird. Zudem gilt wegen der Konkavität von f, dass $f(x) \geq g(x) \geq 0$ in $[a-0.5; a+0.5]$. Entsprechend gilt $\int\limits_{a-0.5}^{a+0.5} f(x)dx \geq \frac{1}{2}(f(a-0.5)+f(a+0.5))$ □

Definition 4. $a \sim b$ genau dann, wenn $\lim\limits_{n\to\infty} \frac{a}{b} = 1$. ◊

Satz 5. *Es gibt ein eindeutig bestimmtes $\tau \in \mathbb{R} > 0$, so dass $n! \sim \tau e^{-n} n^{(n+\frac{1}{2})}$.*

Beweis.

$$\ln n! = \sum_{i=1}^{n} \ln i$$

Da $\ln x$ konkav und monoton wachsend ist gilt für $k = 1, ..., n$

$$\int\limits_{k-\frac{1}{2}}^{k+\frac{1}{2}} \ln x dx \leq \ln k \leq \int\limits_{k}^{k+1} \ln x dx$$

(die erste Ungleichung gilt wegen Satz 2, die zweite gilt, weil $\ln k$ mit der Fläche des Rechtsecks über $[k, k+1]$ und der Höhe $\ln k$ identisch ist. Da $\ln x$ monoton steigend ist, gilt für $x \geq k$ $\ln x \geq \ln k$ und damit $\int\limits_{k}^{k+1} \ln x dx \geq \ln k$).
Summiert man über 1 bis n erhält man

$$\int\limits_{\frac{1}{2}}^{n+\frac{1}{2}} \ln x dx \leq \ln n! \leq \int\limits_{1}^{n+1} \ln x dx.$$

Da $x \ln x - x$ eine Stammfunktion von $\ln x$ ist (für $x > 0$), gilt

$$(x \ln x - x)|_{\frac{1}{2}}^{n+\frac{1}{2}} \leq \ln n! \leq (x \ln x - x)|_{1}^{n+1}. \tag{1}$$

Es ist

$$\begin{aligned}(x \ln x - x)|_{\frac{1}{2}}^{n+\frac{1}{2}} &= \left(n+\frac{1}{2}\right)\ln\left(n+\frac{1}{2}\right) - n - \frac{1}{2} - \left(\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\right) \\ &= \left(n+\frac{1}{2}\right)\ln\left(n+\frac{1}{2}\right) - n - \frac{1}{2}\ln\frac{1}{2}\end{aligned}$$

und

$$\begin{aligned}(x\ln x - x)|_1^{n+1} &= (n+1)\ln(n+1)-(n+1)-(-1)\\ &= (n+1)\ln(n+1)-n.\end{aligned}$$

Setzt man

$$d_n := \ln n! - \left(n+\frac{1}{2}\right)\ln n + n \quad (\text{für } n \geq 1),$$

so zeigt die Ungleichung (1) für $\ln n!$, dass

$$d_n \geq -\frac{1}{2}\ln\frac{1}{2}, \tag{2}$$

denn gemäss (1) gilt

$$\ln n! - \left(n+\frac{1}{2}\right)\ln(n+\frac{1}{2}) + n \geq -\frac{1}{2}\ln\frac{1}{2}$$

und da

$$\ln\left(n+\frac{1}{2}\right) > \ln(n)$$

gilt auch

$$\ln n! - \left(n+\frac{1}{2}\right)\ln(n) + n \geq -\frac{1}{2}\ln\frac{1}{2}.$$

Zudem gilt

$$d_n - d_{n+1} = \left(n+\frac{1}{2}\right)\ln\frac{n+1}{n} - 1,$$

denn
$d_n - d_{n+1} = \ln n! - \left(n+\frac{1}{2}\right)\ln n + n - \left(\ln(n+1)! - \left(n+\frac{3}{2}\right)\ln(n+1) + n + 1\right) =$
$\ln n! - \left(n+\frac{1}{2}\right)\ln n + n - \ln(n+1)! + \left(n+\frac{3}{2}\right)\ln(n+1) - n - 1 =$
$\ln n! - \left(n+\frac{1}{2}\right)\ln n - \ln n! - \ln(n+1) + \left(n+\frac{3}{2}\right)\ln(n+1) - 1 =$
$-\left(n+\frac{1}{2}\right)\ln n - \ln(n+1) + \left(n+\frac{3}{2}\right)\ln(n+1) - 1 =$
$-\left(n+\frac{1}{2}\right)\ln n - \ln(n+1)\left(1-\left(n+\frac{3}{2}\right)\right) - 1 =$
$-\left(n+\frac{1}{2}\right)\ln n - \ln(n+1)\left(-n-\frac{1}{2}\right) - 1 =$
$-\left(n+\frac{1}{2}\right)\ln n + \ln(n+1)\left(n+\frac{1}{2}\right) - 1 =$
$\left(n+\frac{1}{2}\right)\left(-\ln n + \ln(n+1)\right) - 1 =$
$\left(n+\frac{1}{2}\right)\ln\frac{n+1}{n} - 1.$

Schliesslich gilt

$$d_n - d_{n+1} = \left(n+\frac{1}{2}\right)\ln\frac{n+1}{n} - 1 = \int_n^{n+1}\ln x dx - \frac{1}{2}(\ln(n+1)+\ln n) \geq 0,$$

denn $\int\limits_n^{n+1} \ln x dx = (n+1)\ln(n+1) - (n+1) - (n \ln n - n)$
$= (n+1)\ln(n+1) - n - 1 - n\ln n + n$
$= (n+1)\ln(n+1) - 1 - n\ln n$
$= n\ln(n+1) + \ln(n+1) - 1 - n\ln n.$
Damit gilt dann
$\int\limits_n^{n+1} \ln x dx - \frac{1}{2}(\ln(n+1) + \ln n)$
$= n\ln(n+1) + \ln(n+1) - 1 - n\ln n - \frac{1}{2}\ln(n+1) - \frac{1}{2}\ln n$
$= n\ln(n+1) + \frac{1}{2}\ln(n+1) - 1 - n\ln n - \frac{1}{2}\ln n$
$= n\left(\ln(n+1 - \ln n)\right) + \frac{1}{2}\left(\ln(n+1) - \ln n\right) - 1$
$= n\ln\frac{n+1}{n} + \frac{1}{2}\ln\frac{n+1}{n} - 1$
$= \left(n + \frac{1}{2}\right)\ln\frac{n+1}{n} - 1.$

Auf Grund des Satzes 3 gilt $\int\limits_n^{n+1} \ln x dx \geq \frac{1}{2}(\ln(n+1) + \ln n)$ und damit

$$d_n - d_{n+1} = \int\limits_n^{n+1} \ln x dx - \frac{1}{2}(\ln(n+1) + \ln n) \geq 0.$$

Also ist (d_n) monoton fallend und laut (2) nach unten beschränkt. Somit konvergiert (d_n) gegen eine (eindeutig bestimmte) Konstante C mit

$$\lim_{n\to\infty} d_n =: C \geq -\frac{1}{2}\ln\frac{1}{2} \approx 0.346\,57 > 0.$$

Es gilt

$$e^{\lim\limits_{n\to\infty}\left(\ln n! - \left(n+\frac{1}{2}\right)\ln n + n\right)} = e^C = \lim_{n\to\infty} e^{\ln n! - \left(n+\frac{1}{2}\right)\ln n + n}$$

und mit

$$e^{\ln n! - \left(n+\frac{1}{2}\right)\ln n + n} = n! \cdot \left(e^{-\ln n)}\right)^{n+\frac{1}{2}} \cdot e^n = n! \cdot n^{-\left(n+\frac{1}{2}\right)} \cdot e^n = \frac{n!}{n^{n+\frac{1}{2}}e^{-n}}$$

gilt

$$e^C = \lim_{n\to\infty} \frac{n!}{n^{n+\frac{1}{2}}e^{-n}} > 0.$$

Mit

$$\tau := e^C$$

und durch Division der Gleichung durch τ erhält man:

$$1 = \lim_{n\to\infty} \frac{n!}{\tau n^{n+\frac{1}{2}}e^{-n}}$$

und damit

$$n! \sim \tau n^{n+\frac{1}{2}} e^{-n} \text{ mit } \tau > 0.$$

□

Satz 6. *Sei* $p \in]0,1[$; k_n *Wert einer binomialverteilten Zufallsvariable* $X_n \sim B(n,p)$; $x(n,k_n) := \frac{k_n - np}{\sigma_n}$ $(\sigma_n := \sqrt{np(1-p)} := \sqrt{npq})$ *und* $\frac{x(n,k_n)^3}{\sqrt{n}} \to 0$ *Dann gilt*

$$P(X_n = k_n) \sim \frac{1}{\tau \sigma_n} \exp(-\frac{1}{2} x(n,k_n)^2),$$

wobei $\tau > 0$ *die durch* $n! \sim \tau n^{n+\frac{1}{2}} e^{-n}$ *(Satz 3) eindeutig bestimmte reelle Zahl ist.*

Beweis. Aus $\frac{x(n,k_n)^3}{\sqrt{n}} \to 0$ folgt $\frac{k_n}{n} \to p$, da mit $t_n := \frac{k_n}{n}$ gilt:

$$\begin{aligned} x(n,k_n)^3 &= \frac{(k_n - np)^3}{npq} \\ &= \frac{(nt_n - np)^3}{npq} \\ &= n^3 \frac{(t_n - p)^3}{npq} \\ &= \frac{n^2}{pq} (t_n - p)^3 . \end{aligned} \tag{3}$$

Damit gilt dann laut Voraussetzung

$$\begin{aligned} \frac{x(n,k_n)^3}{\sqrt{n}} &= \frac{n^2}{\sqrt{npq}} (t_n - p)^3 \\ &= \frac{n\sqrt{n}}{pq} (t_n - p)^3 \to 0 \end{aligned}$$

Der letzte Ausdruck kann nur gegen 0 gehen wenn $t_n \to p$ (und zwar muss t_n viel schneller gegen p gehen als dass $n\sqrt{n}$ gegen Unendlich).
Für k_n gilt weiter:

$$0 \leq k_n \leq n,$$

da k_n Wert einer binomialverteilten Zufallsvariable ist. Zudem gilt:

$$\lim_{n\to\infty} k_n = \infty,$$

denn es gilt $\frac{k_n}{n} \to p$. Gäbe es ein $\lim\limits_{n\to\infty} k_n = c \in \mathbb{R}$, so wäre $c \geq 0$, da $k_n \geq 0$, für $n \in \mathbb{N}$. Zudem wäre in diese Falle $\lim\limits_{n\to\infty} \frac{k_n}{n} = 0 \neq p \in]0,1[$.
Zudem gilt

$$\lim_{n\to\infty} (n - k_n) = \infty,$$

denn $\lim\limits_{n\to\infty}\left(1-\frac{k_n}{n}\right) = 1-p = \lim\limits_{n\to\infty}\left(\frac{n-k_n}{n}\right)$. Es gilt $n-k_n \geq 0$ für alle $n \in \mathbb{N}$. Wäre $\lim_{n\to\infty}(n-k_n) = c$ mit $0 \leq c < \infty$, so wäre $\lim\limits_{n\to\infty}\left(\frac{n-k_n}{n}\right) = 0 \neq 1-p \in]0,1[$. Damit gilt $\lim\limits_{n\to\infty}(n-k_n) = \infty$.
Da nun $P(X_n = k_n) = \binom{n}{k_n}p^{k_n}q^{n-k_n} = \frac{n!}{(n-k_n)!k_n!}p^{k_n}q^{n-k_n}$, erhält man mit Satz 3 und $k := k_n$

$$P(X_n = k) \sim \frac{1}{\tau}\sqrt{\frac{n}{k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}, \tag{4}$$

denn

$$\begin{aligned}
\frac{n!}{(n-k)!k!} &\sim \frac{\tau n^{n+\frac{1}{2}}e^{-n}}{\tau(n-k)^{n-k+\frac{1}{2}}e^{-(n-k)}\tau k^{k+\frac{1}{2}}e^{-k}} \\
&= \frac{n^{n+\frac{1}{2}}e^{-n}}{\tau(n-k)^{n-k+\frac{1}{2}}e^{-n}e^{k}k^{k+\frac{1}{2}}e^{-k}} \\
&= \frac{n^{n+\frac{1}{2}}}{\tau(n-k)^{n-k+\frac{1}{2}}k^{k+\frac{1}{2}}} \\
&= \frac{n^{n-k}n^k\sqrt{n}}{\tau(n-k)^{n-k}\sqrt{n-k}k^k\sqrt{k}} \\
&= \frac{1}{\tau}\sqrt{\frac{n}{k(n-k)}}\left(\frac{n}{k}\right)^k\left(\frac{n}{n-k}\right)^{n-k}.
\end{aligned}$$

Damit wird dann

$$\begin{aligned}
\frac{n!}{(n-k)!k!}p^kq^{n-k} &\sim \frac{1}{\tau}\sqrt{\frac{n}{k(n-k)}}\left(\frac{n}{k}\right)^k\left(\frac{n}{n-k}\right)^{n-k}p^kq^{n-k} \\
&= \frac{1}{\tau}\sqrt{\frac{n}{k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}.
\end{aligned} \tag{5}$$

Zudem gilt

$$\sqrt{\frac{n}{k(n-k)}} \sim \frac{1}{\sigma_n}, \tag{6}$$

denn einerseits gilt

$$k \sim np,$$

da mit $\frac{k}{n} \to p$ gilt: $\frac{k}{np} \to 1$ und damit $k \sim np$.
Andererseits gilt

$$n-k \sim nq,$$

da $\lim\limits_{n\to\infty}\frac{n-k}{nq} = \lim\limits_{n\to\infty}\frac{\frac{1}{n}(n-k)}{\frac{n(1-p)}{n}} = \lim\limits_{n\to\infty}\frac{1-\frac{k}{n}}{1-p} = \frac{1-p}{1-p} = 1$.
Damit ergibt sich

$$\sqrt{\frac{n}{k(n-k)}} \sim \sqrt{\frac{n}{np(nq)}} = \sqrt{\frac{1}{npq}} = \frac{1}{\sigma_n}.$$

Daher genügt es, das Grenzverhalten von

$$\chi(n,k) := \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \tag{7}$$

zu betrachten. Sei $t := t_n := \frac{k_n}{n} =: \frac{k}{n}$. Dann gilt $t \to p$ für $n \to \infty$.
Durch Logarithmieren erhält man:

$$\begin{aligned}
\ln \chi(n,k) &= k \ln \frac{np}{k} + (n-k) \ln \frac{nq}{n-k} \\
&= nt \ln \frac{np}{nt} + (n-nt) \ln \frac{nq}{n-nt} \\
&= n \left(t \ln \frac{p}{t} + (1-t) \ln \frac{q}{1-t} \right)
\end{aligned}$$

und damit

$$-\ln \chi(n,k) = n \left(t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{q} \right).$$

Für die Funktion $g(t) := t \ln \frac{t}{p} + (1-t) \ln \frac{1-t}{q}$ gilt:

$$\begin{aligned}
g(p) &= p \ln \frac{p}{p} + (1-p) \ln \frac{1-p}{1-p} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
g'(t) &= \ln \frac{t}{p} + t\frac{1}{t} - \ln \frac{1-t}{q} + (1-t)\left(\frac{-1}{1-t}\right) \\
&= \ln \frac{t}{p} + 1 - \ln \frac{1-t}{q} - 1 \\
&= \ln \frac{t}{p} - \ln \frac{1-t}{q}
\end{aligned}$$

$$\begin{aligned}
g'(p) &= \ln \frac{p}{p} - \ln \frac{1-p}{1-p} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
g''(t) &= \frac{1}{t} - \frac{-1}{1-t} \\
&= \frac{1}{t} + \frac{1}{1-t}
\end{aligned}$$

$$
\begin{aligned}
g''(p) &= \frac{1}{p} + \frac{1}{1-p} \\
&= \frac{1}{p} + \frac{1}{q} \\
&= \frac{q+p}{pq} \\
&= \frac{1-p+p}{pq} \\
&= \frac{1}{pq}
\end{aligned}
$$

Damit erhält in p die Taylorentwicklung

$$g(t) = g(p) + g'(p)(p-t) + \frac{1}{2}g''(p)(p-t)^2 + R_3(t)$$

die Form

$$g(t) = \frac{1}{2pq}\,(p-t)^2 + R_3(t),$$

wobei R_3 die Restfunktion ist. g ist dreimal ableitbar. Damit gibt es nach dem Satz von Taylor ein $\xi_t \in]p,t[$, so dass $R_3(t) = \frac{g'''(\xi_t)}{3!}(t-p)^3$. Somit gibt es eine Konstante $c > 0$ mit $c \in \mathbb{R}$, so dass in einer Umgebung U von p für $t \in U$ gilt

$$|R_3(t)| \leq c\,|p-t|^3\,.$$

Ein solches c kann z.B. bestimmt werden durch

$$c := \max\{c_t \mid c_t := \left|\frac{g'''(\xi_t)}{3!}\right|, t \in U\})$$

Aus der Voraussetzung $\lim\limits_{n\to\infty} \frac{x(n,k_n)^3}{\sqrt{n}} \to 0$ folgt wegen (3)

$$\lim_{n\to\infty} nc\,(p-t)^3 \to 0,$$

Da $|nR_3(t)| \leq nc\,|p-t|^3$, folgt daraus

$$\lim_{n\to\infty} (nR_3(t))) \to 0.$$

Für Taylorpolynome $g_p(t)$ von $f(t)$ in p gilt: $\lim\limits_{t\to p} g_p(t) = \lim\limits_{t\to p} f(x)$. Entsprechend gilt, da $t \to p$ für $n \to \infty$,

$$\lim_{n\to\infty} (-\ln\chi(n,k)) = \lim_{n\to\infty} \left(\frac{n}{2pq}\,(t-p)^2 + nR_3(t)\right)$$

und damit

$$\lim_{n\to\infty} (-\ln\chi(n,k)) = \lim_{n\to\infty} \left(\frac{n}{2pq}\,(t-p)^2\right)$$

sowie daraus

$$\lim_{n\to\infty}\left(-\ln\chi(n,k)-\frac{n}{2pq}\,(t-p)^2\right)=0.$$

Gemäss analoger Herleitung zu (3) gilt die Gleichung

$$\frac{1}{2}x(n,k)^2=\frac{n}{2pq}\,(t-p)^2\,.$$

und damit

$$\begin{aligned}
&\lim_{n\to\infty}\left(-\ln\chi(n,k)-\frac{n}{2pq}\,(t-p)^2\right)\\
=\;&\lim_{n\to\infty}\left(-\ln\chi(n,k)-\frac{1}{2}x(n,k)^2\right)\\
=\;&0.
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
1 &= \exp 0=\exp\left(\lim_{n\to\infty}\left(-\ln\chi(n,k)-\frac{1}{2}x(n,k)^2\right)\right)\\
&= \lim_{n\to\infty}\exp\left(-\ln\chi(n,k)-\frac{1}{2}x(n,k)^2\right)\\
&= \lim_{n\to\infty}\left(\frac{\exp\left(-\ln\chi(n,k)\right)}{\exp\left(\frac{1}{2}x(n,k)^2\right)}\right)\\
&= \lim_{n\to\infty}\left(\frac{-\chi(n,k)}{\exp\left(\frac{1}{2}x(n,k)^2\right)}\right)\\
&= \lim_{n\to\infty}\left(\frac{\chi(n,k)}{\exp\left(-\frac{1}{2}x(n,k)^2\right)}\right).
\end{aligned}$$

Damit gilt unter der Bedingung $\lim_{n\to\infty}\left(\frac{x(n,k)^3}{\sqrt{n}}\right)=0$

$$\chi(n,k)\sim\exp\left(-\frac{1}{2}x(n,k)^2\right),$$

woraus mit (4), (5), (6) und (7) folgt:

$$P(X_n=k_n)\sim\frac{1}{\tau\sigma_n}\exp\left(-\frac{1}{2}x(n,k)^2\right).$$

□

Satz 7. *Wenn $X_n\sim B(n,p)$, dann*

$$\lim_{n\to\infty}P(a\leq\frac{X_n-np}{\sigma_n}\leq b)=\frac{1}{\tau}\int_a^b\exp\left(-\frac{1}{2}x^2\right)dx,$$

wobei $\tau > 0$ die durch $n! \sim \tau n^{n+\frac{1}{2}} e^{-n}$ (Satz 3) eindeutig bestimmte reelle Zahl ist. $x(n, k_n) := \frac{k_n - np}{\sigma_n}$ ($\sigma_n := \sqrt{np(1-p)} := \sqrt{npq}$).

Beweis. Es gilt:

$$\begin{aligned} a &\leq \frac{X_n - np}{\sigma_n} \leq b \\ &\Longleftrightarrow a\sigma_n + np \leq X_n \leq b\sigma_n + np \end{aligned}$$

und mit

$$\begin{aligned} \alpha_n &: = \min\{x \mid x \geq a\sigma_n + np \text{ und } x \in \mathbb{Z}\} \\ \beta_n &: = \max\{x \mid x \leq b\sigma_n + np \text{ und } x \in \mathbb{Z}\} \end{aligned}$$

gilt

$$\begin{aligned} \{y &\mid a \leq \frac{X_n(y) - np}{\sigma_n} \leq b\} \\ &= \{y \mid \alpha_n \leq X_n(y) \leq \beta_n\} \\ \text{für } y &\in \{0,1\}^n \text{ und } X_n(y) = \sum_{i=1}^{n} y_i. \end{aligned}$$

Zudem gilt damit:

$$\begin{aligned} |x(n, \alpha_n) - a| &\leq \frac{1}{\sigma_n} \\ |x(n, \beta_n) - b| &\leq \frac{1}{\sigma_n}, \end{aligned}$$

denn gemäss Definition von α_n gilt z.B.:

$$|\alpha_n - a\sigma_n - np| < 1$$

und damit mittels Division durch σ_n

$$\begin{aligned} &\frac{1}{\sigma_n} |\alpha_n - a\sigma_n - np| \\ = &\left| \frac{\alpha_n - np}{\sigma_n} - a \right| \\ = &|x(n, \alpha_n) - a| \\ \leq &\frac{1}{\sigma_n}. \end{aligned}$$

Damit gilt

$$\lim_{n \to \infty} |x(n, \alpha_n) - a| = 0$$

und

$$\lim_{n \to \infty} x(n, \alpha_n) = a$$

und entsprechend

$$\begin{aligned}\frac{x(n,\alpha_n)^3}{\sqrt{n}} &\rightarrow 0\\ \frac{x(n,\beta_n)^3}{\sqrt{n}} &\rightarrow 0\end{aligned}$$

sowie

$$\frac{x(n,k)^3}{\sqrt{n}} \rightarrow 0$$

für $\alpha_n \leq k \leq \beta_n$, denn $|k-np| \leq |\alpha_n - np|$ und $|k-np| \leq |\beta_n - np|$.
Womit für $\alpha_n \leq k \leq \beta_n$ die Konsequenz von Satz 6 gilt:

$$\binom{n}{k} p^k q^{n-k} \sim \frac{1}{\tau\sigma_n} \exp\left(-\frac{1}{2}x(n,k)^2\right)$$

Damit gibt es eine Folge $\varepsilon_n \rightarrow 0$ so dass

$$1-\varepsilon_n \leq \frac{\binom{n}{k}p^k q^{n-k}}{\frac{1}{\tau\sigma_n}\exp\left(-\frac{1}{2}x(n,k)^2\right)} \leq 1+\varepsilon_n$$

oder äquivalent

$$\begin{aligned}(1-\varepsilon_n)\frac{1}{\tau\sigma_n}\exp\left(-\frac{1}{2}x(n,k)^2\right) &\leq \binom{n}{k}p^k q^{n-k}\\ &\leq (1+\varepsilon_n)\frac{1}{\tau\sigma_n}\exp\left(-\frac{1}{2}x(n,k)^2\right).\end{aligned} \tag{8}$$

$\frac{1}{\sigma_n}\exp\left(-\frac{1}{2}x(n,k)^2\right)$ entspricht für jedes $\alpha_n \leq k \leq \beta_n$ der Fläche des Rechtecks auf

$$\left[\frac{k-np-0.5}{\sigma_n};\frac{k-np+0.5}{\sigma_n}\right],$$

denn dieses Intervall weist die Länge $\frac{1}{\sigma_n}$ auf:

$$\begin{aligned} & \frac{k-np+0.5}{\sigma_n} - \frac{k-np-0.5}{\sigma_n}\\ =\ & \frac{k-np+0.5-k+np+0.5}{\sigma_n}\\ =\ & \frac{1}{\sigma_n}\end{aligned}$$

Zudem beträgt die Höhe dieser Rechtecke $\exp\left(-\frac{1}{2}x(n,k)^2\right)$, wobei $x(n,k)$ jeweils die Basis

$[\frac{k-np-0.5}{\sigma_n}; \frac{k-np+0.5}{\sigma_n}]$ halbiert, denn

$$\begin{aligned}
& \frac{1}{2}\left(\frac{k-np-0.5}{\sigma_n}+\frac{k-np+0.5}{\sigma_n}\right) \\
= & \frac{1}{2}\left(\frac{k-np-0.5+k-np+0.5}{\sigma_n}\right) \\
= & \frac{1}{2}\left(\frac{2k-2np}{\sigma_n}\right) \\
= & \frac{k-np}{\sigma_n} \\
= & x(n,k).
\end{aligned}$$

Damit ist

$$R_n := \sum_{k=\alpha_n}^{\beta_n} \frac{1}{\sigma_n} \exp\left(-\frac{1}{2}x(n,k)^2\right) \tag{9}$$

eine Riemannsumme auf dem Intervall $[x(n,\alpha_n-0.5); x(n,\beta_n+0.5)]$ und es gilt mit der Ungleichung (8)

$$(1-\varepsilon_n)\frac{1}{\tau}R_n \leq \sum_{i=\alpha_n}^{\beta_n} \binom{n}{k} p^k q^{n-k} \leq (1+\varepsilon_n)\frac{1}{\tau}R_n$$

oder äquivalent

$$(1-\varepsilon_n)\frac{1}{\tau}R_n \leq P(a \leq \frac{X_n-np}{\sigma_n} \leq b) \leq (1+\varepsilon_n)\frac{1}{\tau}R_n. \tag{10}$$

Für $n \to \infty$ geht R_n von (9) gegen

$$\int_a^b \exp\left(-\frac{1}{2}x^2\right) dx,$$

denn $\exp\left(-\frac{1}{2}x^2\right)$ ist stetig sowie beschränkt und es gilt

$$\begin{aligned}
x(n,\beta_n+0.5) & \to b \\
x(n,\alpha_n-0.5) & \to a,
\end{aligned}$$

da auf Grund der Definition von α_n und β_n es reelle Zahlen $0 \leq m, l < 1$ gibt, so dass

$$\alpha_n = a\sigma_n + np + m$$

$$\beta_n = b\sigma_n + np + l.$$

Damit gilt dann z.B.:

$$\begin{aligned}
& x(n, \beta_n + 0.5) \\
= \quad & \frac{\beta_n + 0.5 - np}{\sigma_n} \\
= \quad & \frac{b\sigma_n + np + l + 0.5 - np}{\sigma_n} \\
= \quad & \frac{b\sigma_n + l + 0.5}{\sigma_n} \\
= \quad & b + \frac{l}{\sigma_n} + \frac{0.5}{\sigma_n}
\end{aligned}$$

und

$$\begin{aligned}
& \lim_{n\to\infty} \left(b + \frac{l}{\sigma_n} + \frac{0.5}{\sigma_n} \right) \\
= \quad & b.
\end{aligned}$$

Durch Grenzwertbildung $n \to \infty$ auf

$$(1 - \varepsilon_n)\frac{1}{\tau}R_n \leq P(a \leq \frac{X_n - np}{\sigma_n} \leq b) \leq (1 - \varepsilon_n)\frac{1}{\tau}R_n$$

(s. Ungleichung (10)) folgt damit (da $\lim_{n\to\infty}(1 - \varepsilon_n) = 1$)

$$\frac{1}{\tau}\int_a^b \exp\left(-\frac{1}{2}x^2\right) dx \leq \lim_{n\to\infty} P(a \leq \frac{X_n - np}{\sigma_n} \leq b) \leq \frac{1}{\tau}\int_a^b \exp\left(-\frac{1}{2}x^2\right) dx$$

und damit folgt, was zu beweisen war,

$$\lim_{n\to\infty} P(a \leq \frac{X_n - np}{\sigma_n} \leq b) = \frac{1}{\tau}\int_a^b \exp\left(-\frac{1}{2}x^2\right) dx.$$

□

Um den Satz von de Moivre-Laplace zu beweisen, muss noch gezeigt werden, dass $\tau = \sqrt{2\pi}$. Ohne Beweis halten wir dazu den folgenden aus der Analysis bekannten Satz fest:

Satz 8. $\int\limits_{-\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx = \sqrt{2\pi}$

Nun können wir übergehen zu

Satz 9. *Das eindeutig bestimmte $\tau \in \mathbb{R} > 0$, so dass $n! \sim \tau e^{-n} n^{(n+\frac{1}{2})}$, ist mit $\sqrt{2\pi}$ identisch, d.h. es gilt die Stirlingsche Formel*

$$n! \sim \sqrt{2\pi} e^{-n} n^{(n+\frac{1}{2})}.$$

Beweis. Sei $\eta > 0$ beliebig klein. Dann gibt es ein b so dass $\frac{1}{b^2} < \eta$ (je kleiner η desto grösser muss b sein). Wählen wir für jedes η ein solches b, so erhält man eine Funktion

$$\begin{aligned} \beta &: \mathbb{R}^+\backslash\{0\} \to \mathbb{R}^+\backslash\{0\} \\ \beta(\eta) &: = b. \end{aligned}$$

Gemäss Tschebyscheffscher Ungleichung gilt

$$P\left(\left|\frac{X_n - np}{\sqrt{np(1-p)}}\right| > b\right) \leq \frac{1}{b^2} < \eta, \tag{11}$$

da $E\left(\frac{X_n - np}{\sqrt{np(1-p)}}\right) = 0$ und $V\left(\frac{X_n - np}{\sqrt{np(1-p)}}\right) = 1$ (z-standardisierte Zufallsvariable!).
Sei nun $a := -b$ und α_n, β_n wie im Beweis von Satz 7 definiert. Wie dort gezeigt gilt $x(n, \alpha_n) \to -b$ und $x(n, \beta_n) \to b$ sowie

$$P(\alpha_n \leq X_n \leq \beta_n) = P(-b \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b) \to \frac{1}{\tau}\int_{-b}^{b} \exp\left(-\frac{1}{2}x^2\right) dx.$$

Da nun laut (11)

$$P\left(-b \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b\right) \geq 1 - \eta$$

für $n \in \mathbb{N}$ gilt und damit

$$\lim_{n\to\infty} P\left(-b \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b\right) \geq 1 - \eta,$$

gilt

$$\frac{1}{\tau}\int_{-b}^{b} \exp\left(-\frac{1}{2}x^2\right) dx \geq 1 - \eta.$$

Zudem gilt auf Grund der Wahrscheinlichkeitstheorie

$$\lim_{n\to\infty} P\left(-b \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b\right) \leq 1$$

und damit

$$\frac{1}{\tau}\int_{-b}^{b} \exp\left(-\frac{1}{2}x^2\right) dx \leq 1.$$

Entsprechend gilt

$$(1 - \eta) \leq \frac{1}{\tau}\int_{-\beta(\eta)}^{\beta(\eta)} \exp\left(-\frac{1}{2}x^2\right) dx \leq 1 \tag{12}$$

für alle $\eta > 0$ und

$$\lim_{\eta \to 0} (1 - \eta) \leq \lim_{\eta \to 0} \frac{1}{\tau} \int\limits_{-\beta(\eta)}^{\beta(\eta)} \exp\left(-\frac{1}{2}x^2\right) dx \leq \lim_{\eta \to 0} 1.$$

Nun gilt $\lim\limits_{\eta \to 0} \beta(\eta) = \infty$, wobei $\int\limits_{\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx$ laut Satz 8 definiert ist. Damit folgt aus (12) durch Grenzwertbildung

$$1 \leq \int\limits_{\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx \leq 1$$

und damit

$$\frac{1}{\tau} \int\limits_{\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx = 1.$$

Da nun gemäss Satz 8

$$\int\limits_{\infty}^{\infty} \exp\left(-\frac{1}{2}x^2\right) dx = \sqrt{2\pi}$$

gilt

$$\tau = \sqrt{2\pi}$$

□

Der Satz von de Moivre-Laplace folgt nun aus Satz 7, Satz 9 sowie der

Definition 10. $\Phi(t) := \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} \exp\left(-\frac{1}{2}x^2\right) dx$ (= Verteilungsfunktion der Standardnormalverteilung) $\Diamond$

Für $p \in]0,1[$ und $X_n \sim B(n,p)$ gilt:

$$\lim_{n \to \infty} P(a \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq b) = \Phi(b) - \Phi(a).$$