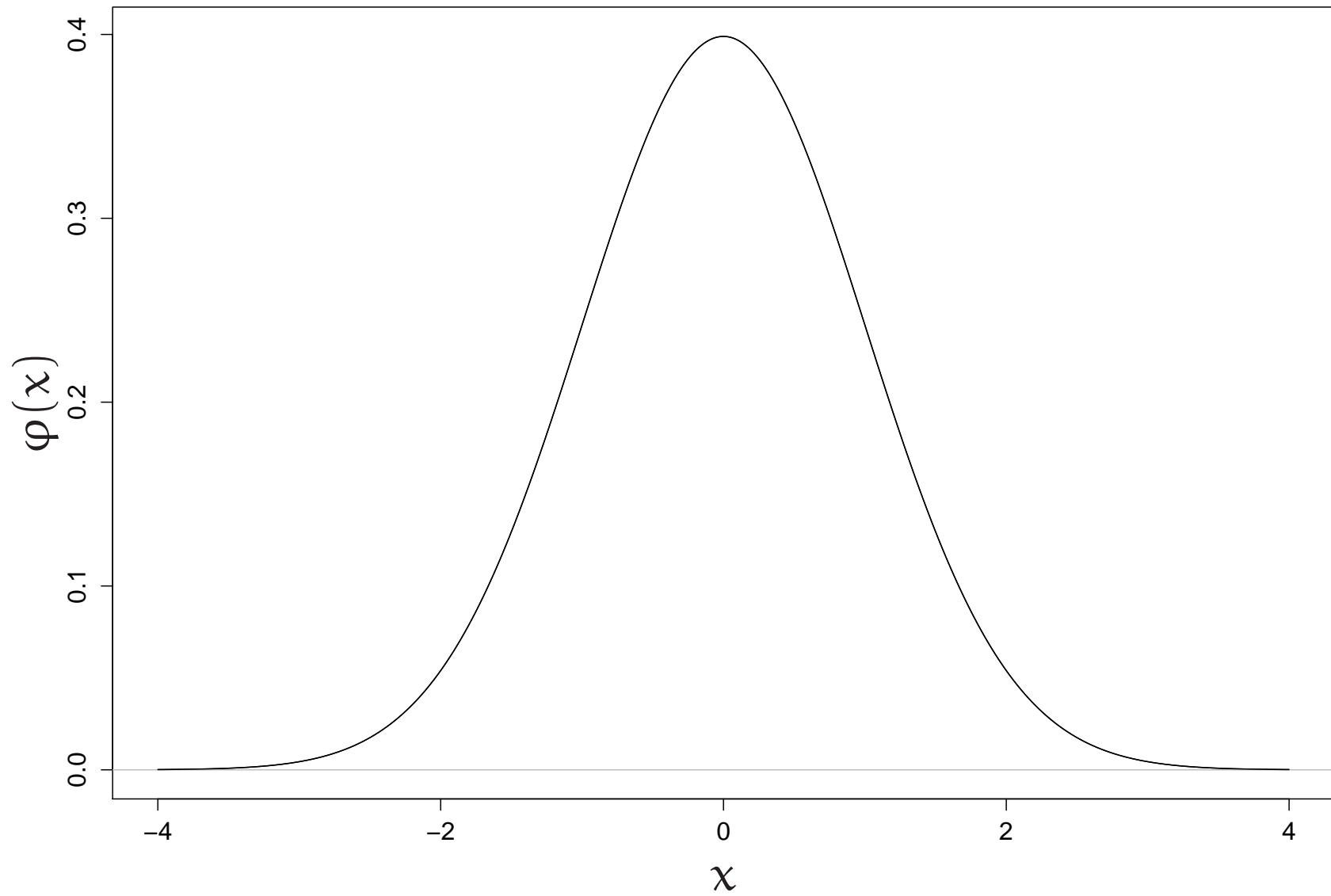


Vorlesung 6b

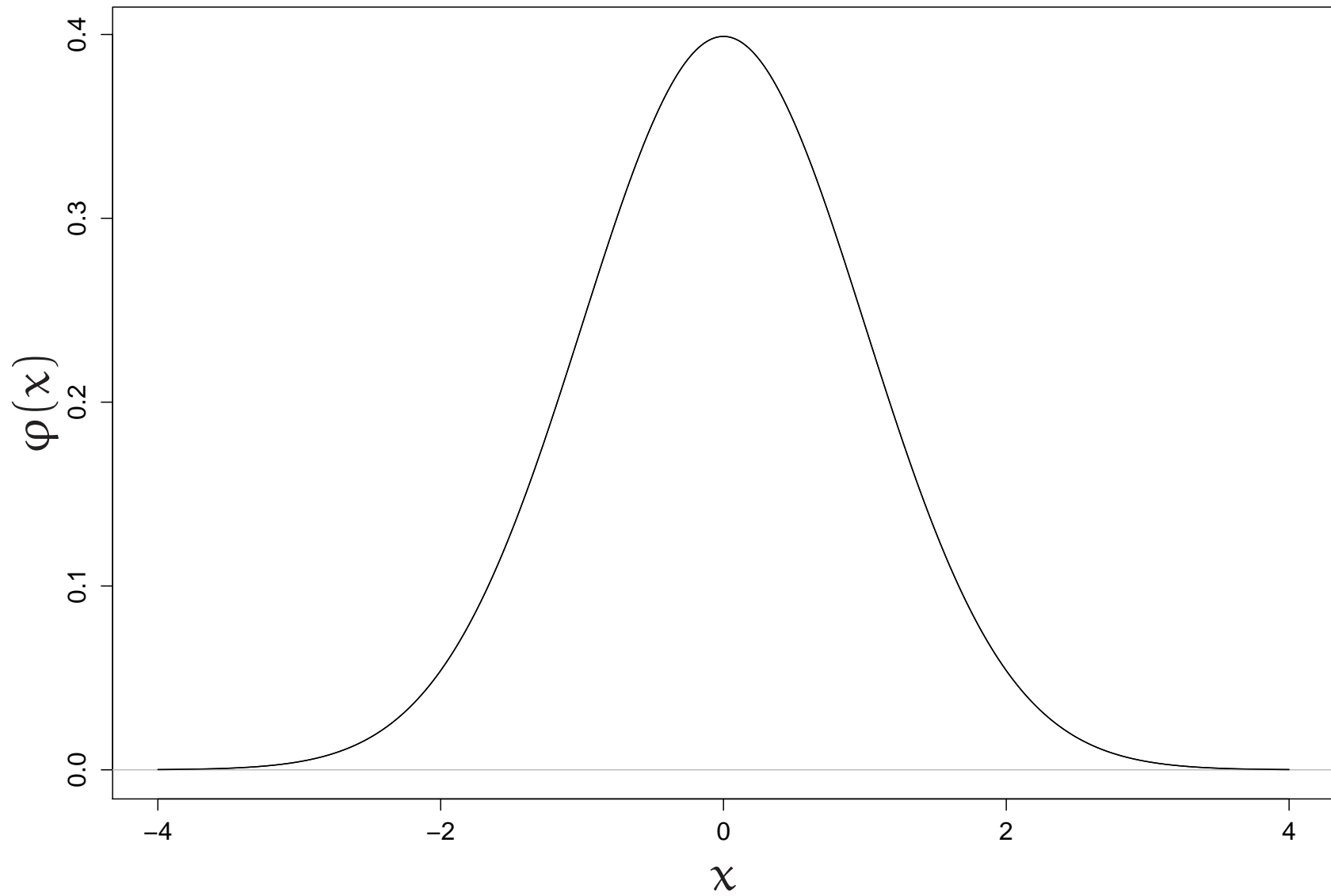
Der Zentrale Grenzwertsatz

Die Normalverteilung

Dichtefunktion φ der Standardnormalverteilung



Die Gaußsche Glocke



Die Standardnormalverteilung

Dichtefunktion:

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Verteilungsfunktion:

$$\Phi(a) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-x^2/2} dx$$

Z standardnormalverteilt:

$$\mathbf{P}(Z \leq a) = \Phi(a)$$

Die Standardnormalverteilung

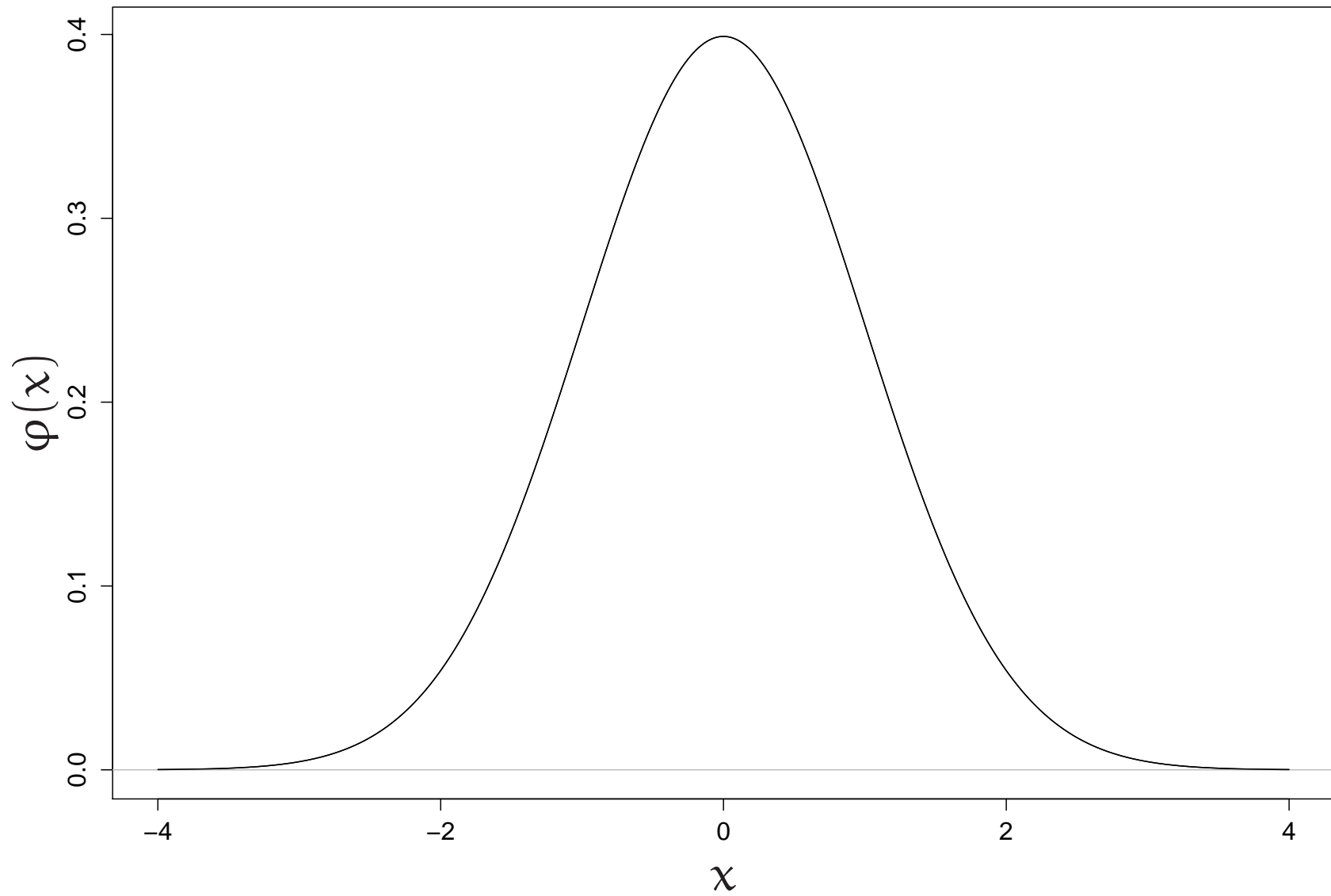
$$\mathbf{EZ} = 0 \quad \sigma_Z = 1$$

Die allgemeine Normalverteilung

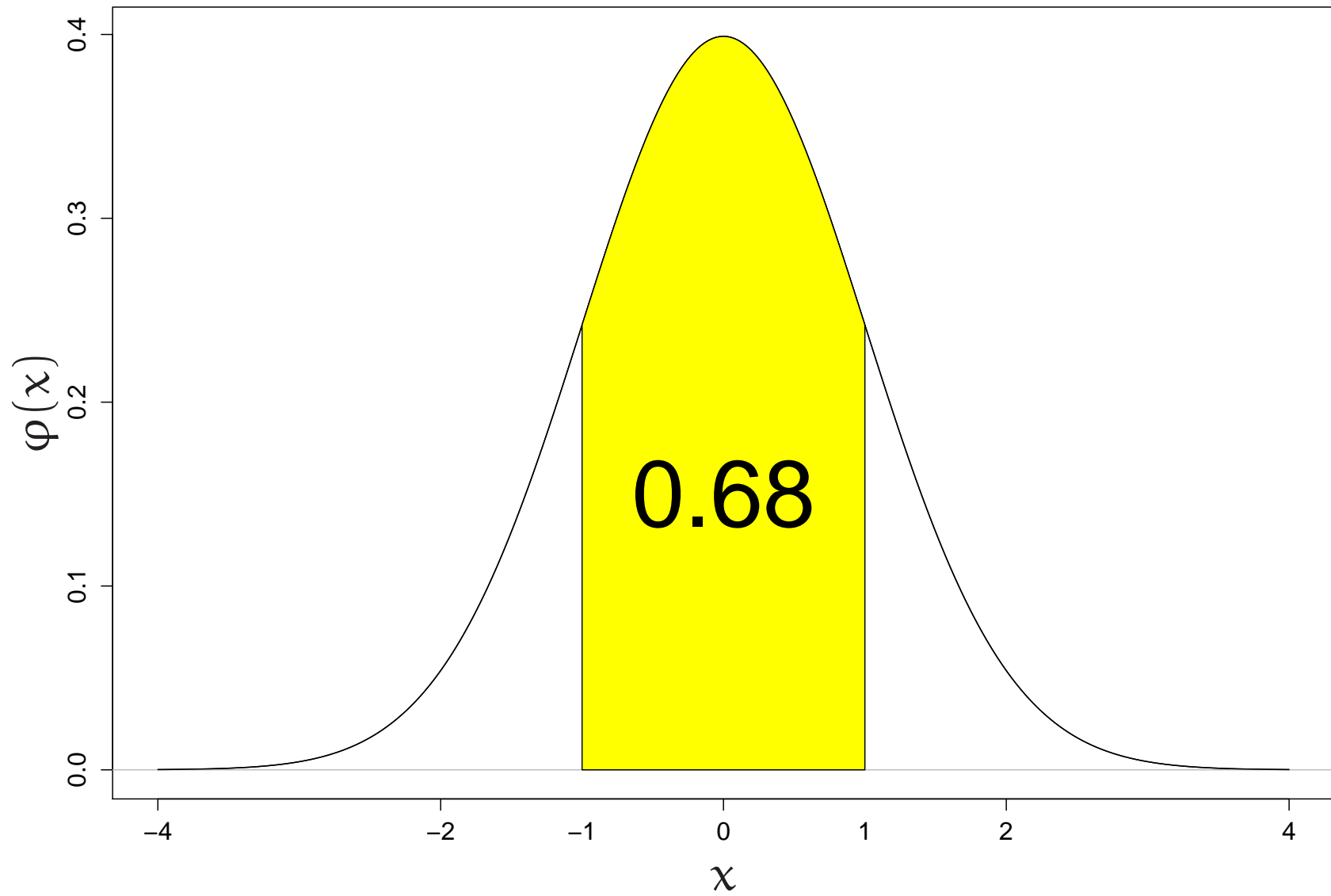
$$\mathbf{N} = \mu + \tau Z$$

$$\mathbf{EN} = \mu \quad \sigma_N = \tau$$

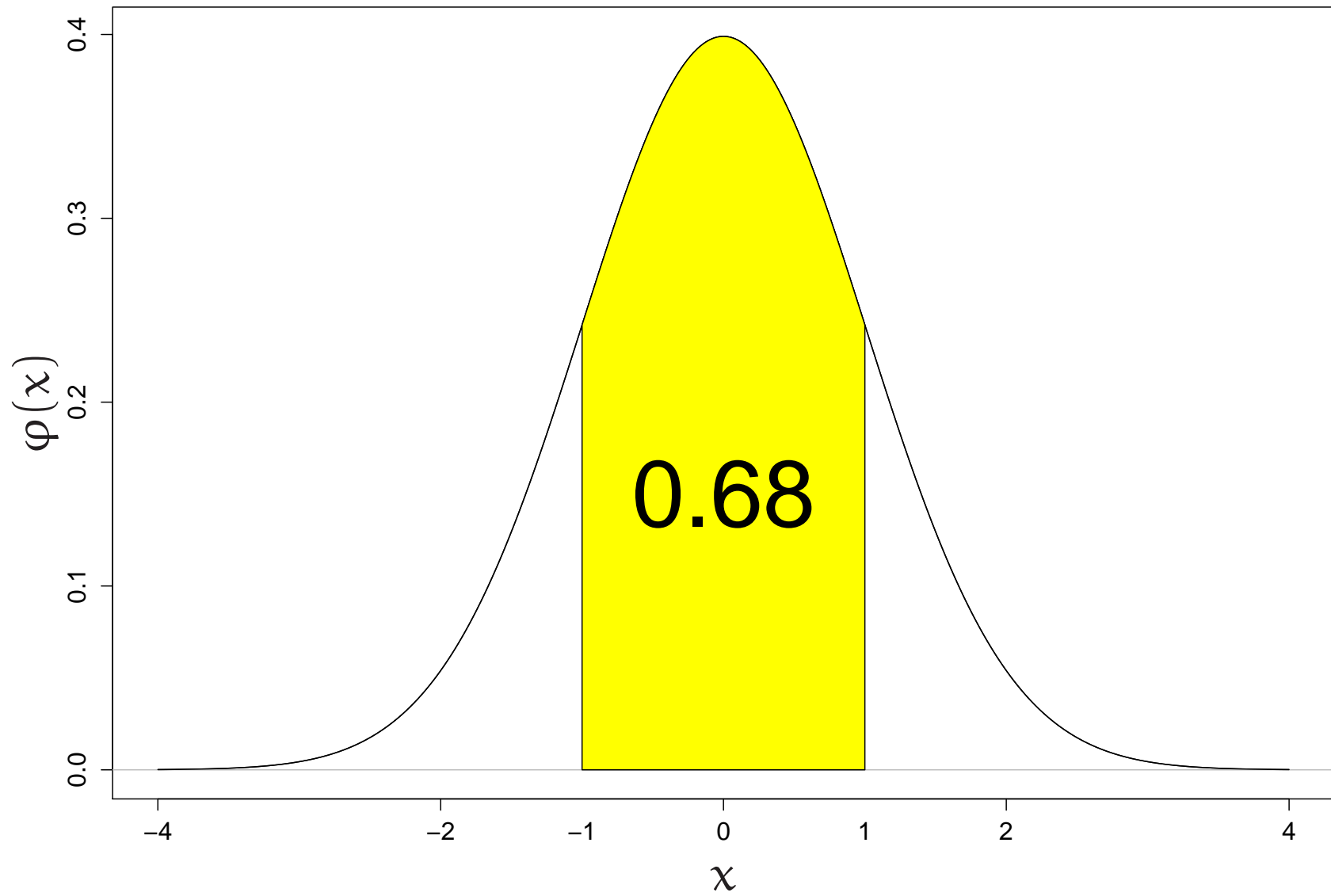
Dichtefunktion φ der Standardnormalverteilung



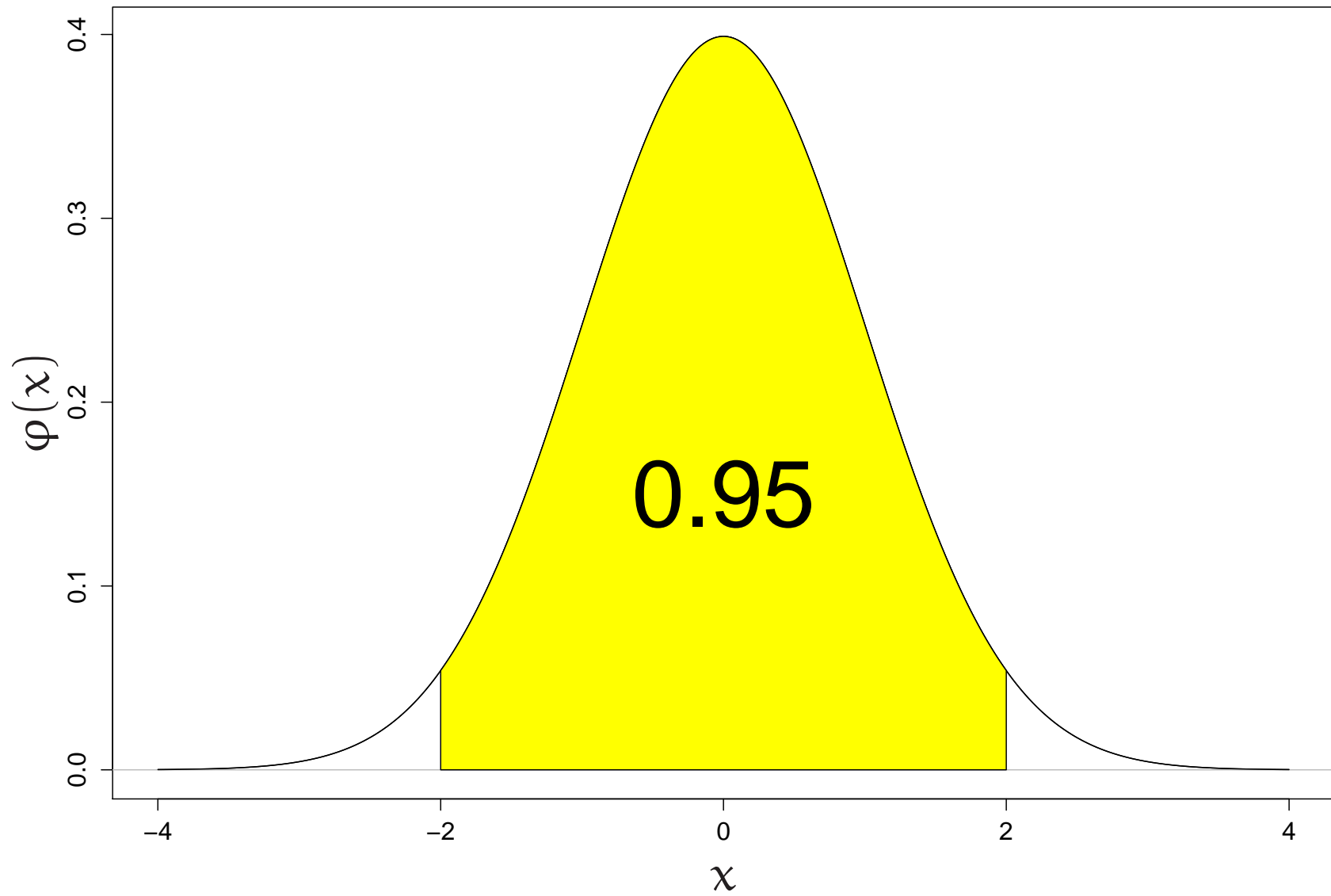
$$\mathbf{P}(|Z| < 1) \approx 0.68$$



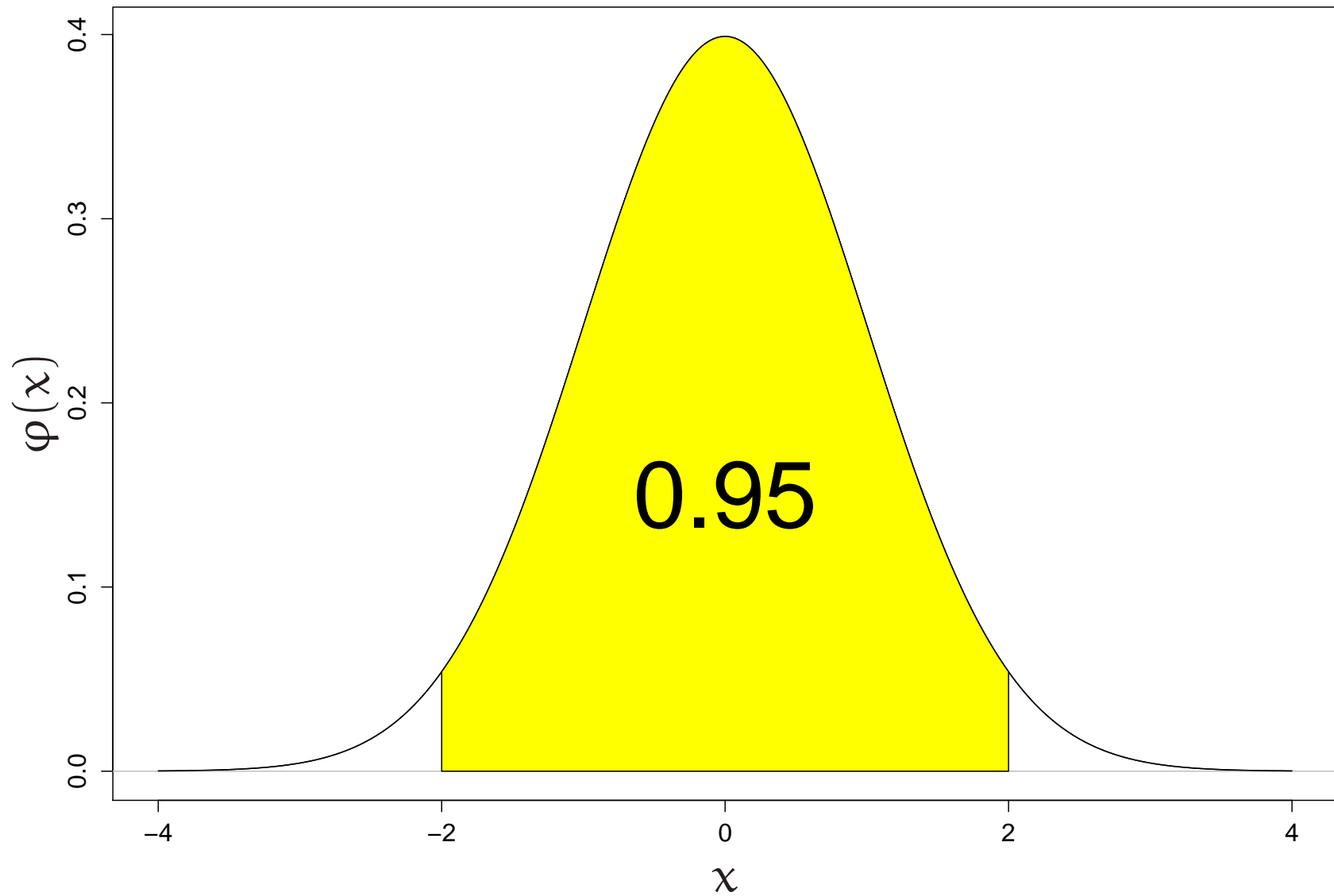
$$\Phi(1) - \Phi(-1) \approx 0.68$$



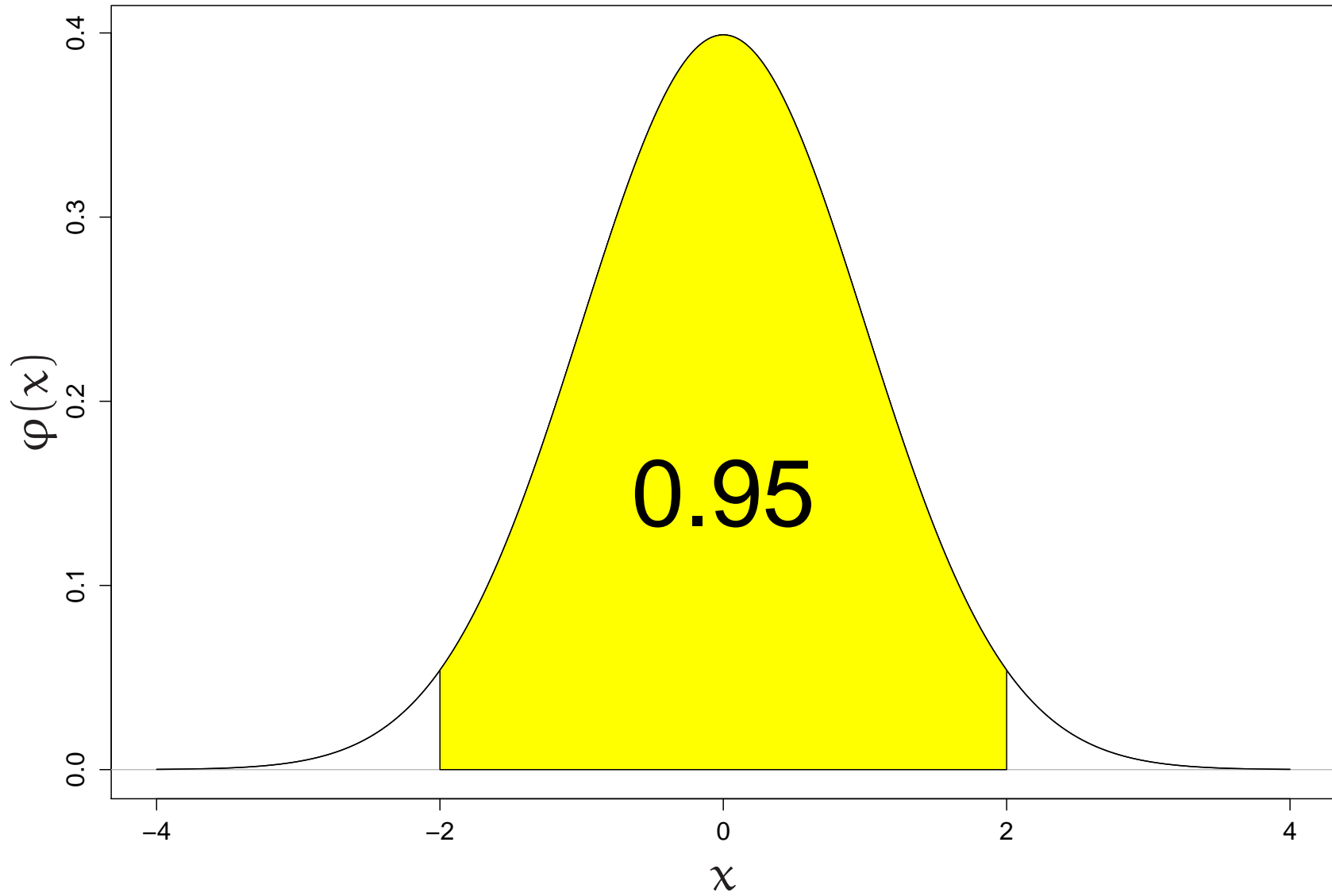
$$\mathbf{P}(|Z| < 2) \approx 0.95$$



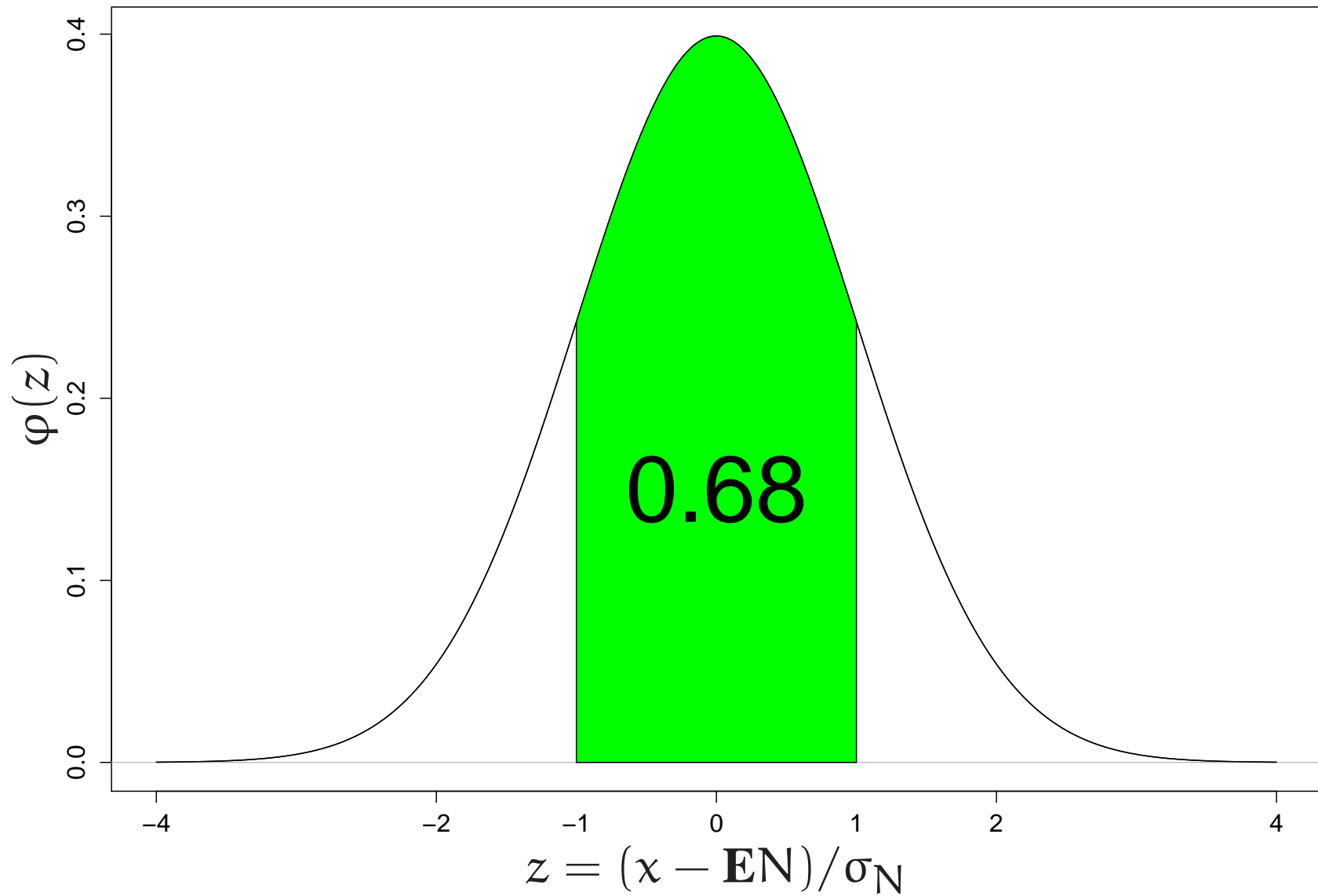
Und für allgemeine normalverteilte Zufallsgrößen N ?



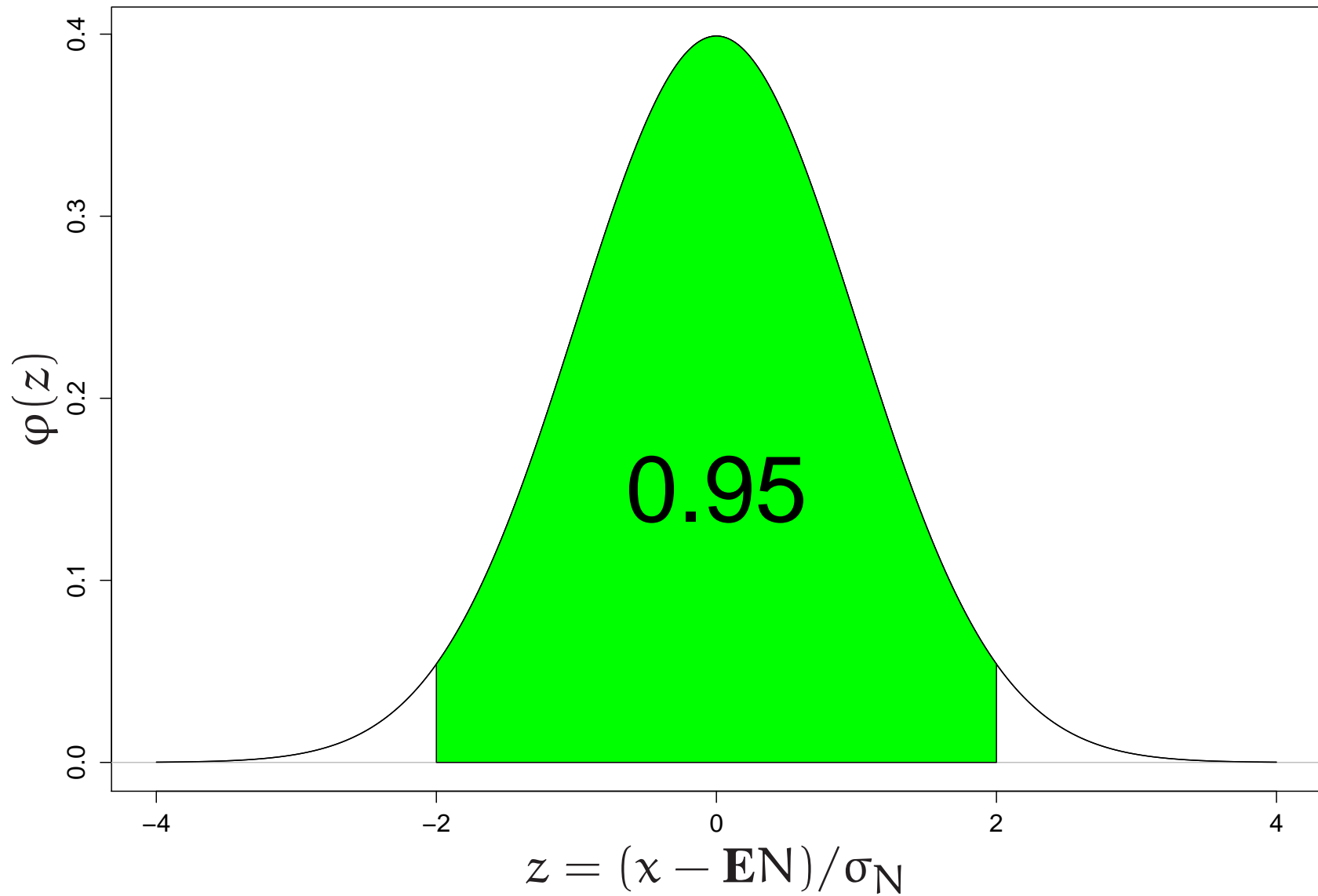
Dasselbe in grün.



$$\mathbf{P}(|N - \mathbf{E}N| < \sigma_N) \approx 0.68$$



$$\mathbf{P}(|\mathbf{N} - \mathbf{EN}| < 2\sigma_{\mathbf{N}}) \approx 0.95$$



Der
Zentrale
Grenzwertsatz

Die Summe
von vielen kleinen,
unabhängigen Summanden
ist annähernd
normalverteilt.

DER ZENTRALE GRENZWERTSATZ

Seien X_1, X_2, \dots

unabhängige, identisch verteilte Zufallsgrößen.

Sei

$$S_n := X_1 + \dots + X_n$$

und

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n}.$$

$$(\mathbf{E}Z_n = 0, \quad \sigma_{Z_n} = 1.)$$

Dann gilt:

Z_n

ist

asymptotisch
standardnormalverteilt.

Das heißt:

$$\mathbf{P}(Z_n \leq a) \rightarrow \Phi(a).$$

Zentraler Grenzwertsatz:

Die standardisierte Summe von unabhängigen,
identisch verteilten \mathbb{R} -wertigen Zufallsvariablen
mit endlicher Varianz
konvergiert in Verteilung
gegen eine standard-normalverteilte Zufallsvariable.

Formal:

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariable mit endlichem Erwartungswert μ und endlicher Varianz $\sigma^2 > 0$. Dann gilt für alle $\ell < r \in \mathbb{R}$

$$\mathbf{P} \left(\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \in [\ell, r] \right) \xrightarrow{n \rightarrow \infty} \mathbf{P}(Z \in [\ell, r]).$$

Dabei ist Z standard-normalverteilt.

Geschichte
des
Zentralen
Grenzwertsatzes

Abraham de Moivre



Der faire Münzwurf (1733)

Pierre-Simon Laplace



Allgemeine binomiale Zufallsgrößen (1812)

Pafnuty Lvovich Chebyshev



$$\mathbf{P}(|X - \mathbf{E}X| > k\sigma) \leq \frac{1}{k^2} \quad (1846)$$

Andrei Andreyevich Markov



Allgemeiner zentraler Grenzwertsatz

Aleksandr Mikhailovich Lyapunov



Noch allgemeiner (1906)

Nehmen wir an,
diese Herren hätten sich
auf ihre vielen anderen Interessen
beschränkt.

ZENTRALER GRENZWERTSATZ

Unbekannt.

Könnten wir ihn entdecken?

Wie kämen wir auf φ ?

Warum gerade $e^{-x^2/2}$?

BEISPIEL

Rundungsfehler

bei

Addition

In Wirklichkeit

$$\pi =$$

3.141592653589793238462643383279502884197169399375105...

Im Rechner

$$\pi \leftarrow 3.14159265358979$$

MODELL

Zahl = Rechnerdarstellung + Rundungsfehler.

$$A = a^{[R]} + \varepsilon X \quad \varepsilon = 10^{-15}$$

X uniform verteilt auf $[-0.5, 0.5]$.

$$\sum_{i=1}^n A_i = ?$$

$$\sum_{i=1}^n A_i = \sum_{i=1}^n a_i^{[R]} + \varepsilon \sum_{i=1}^n X_i$$

Wie groß ist der Fehler?

$$\sum_{i=1}^n X_i \approx ?$$

Empirische Verteilung von

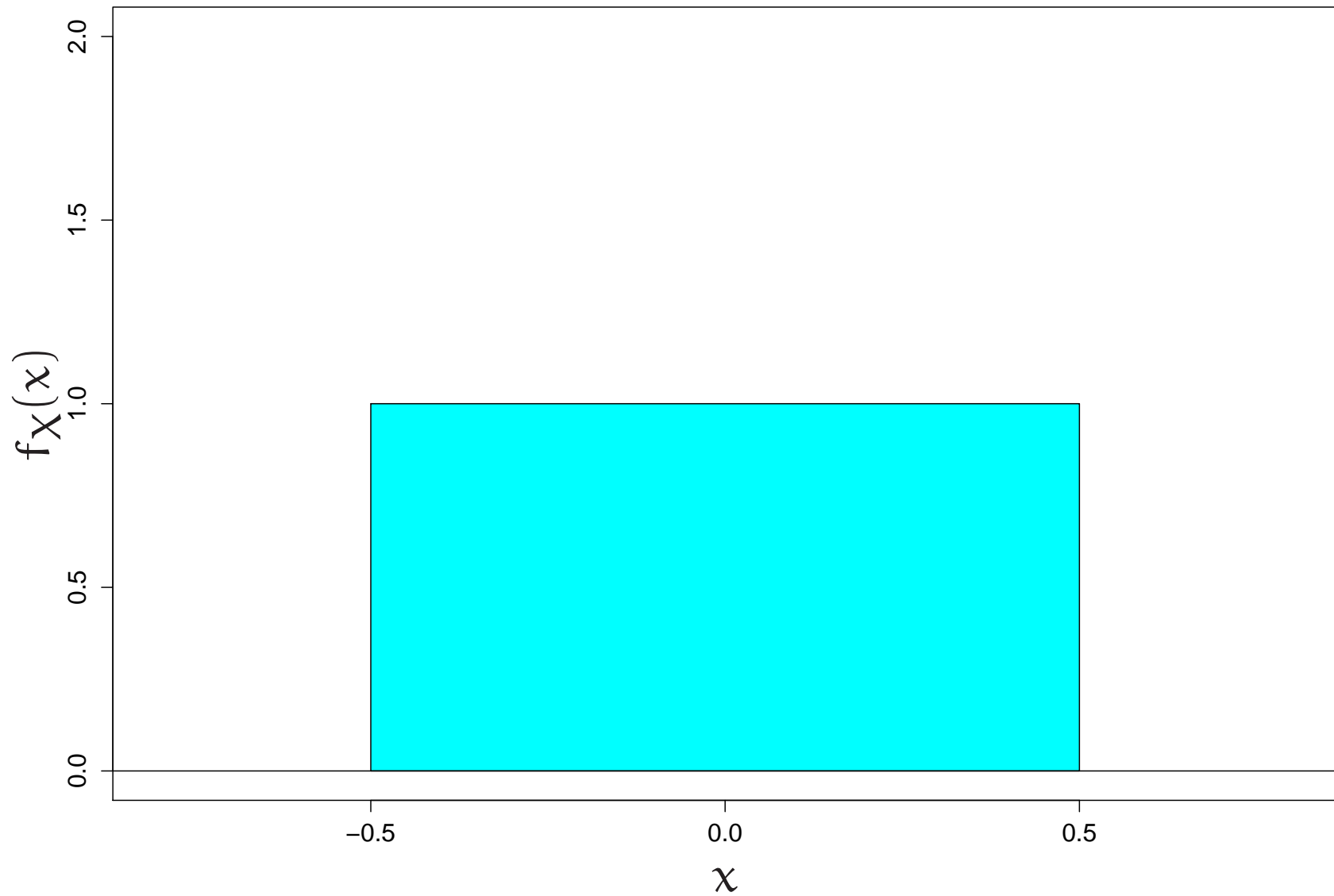
$$S_n := X_1 + \dots + X_n$$

100000 Simulationen

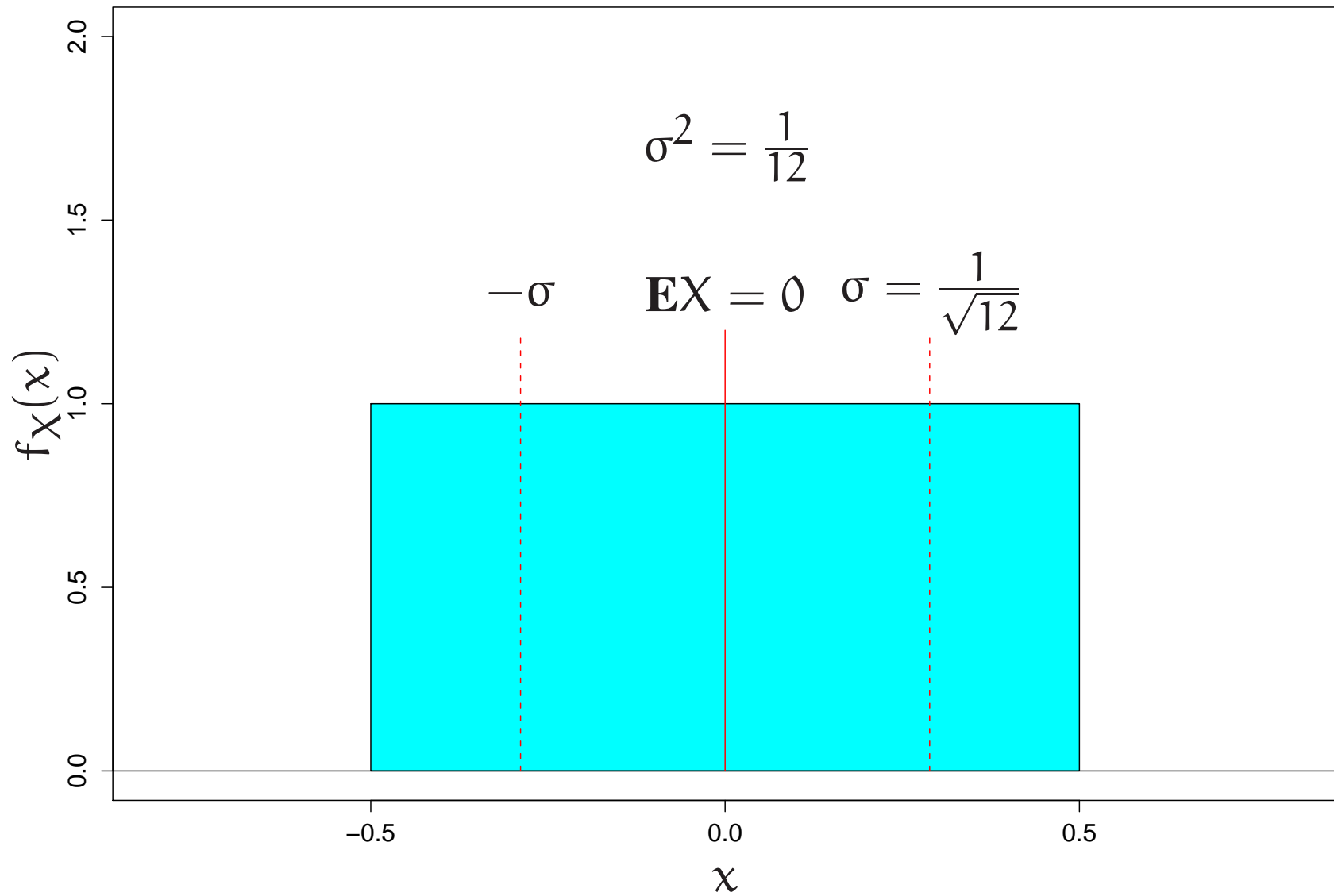
$$n = 1, 2, \dots, 10$$

$$n = 15, 20, \dots, 100$$

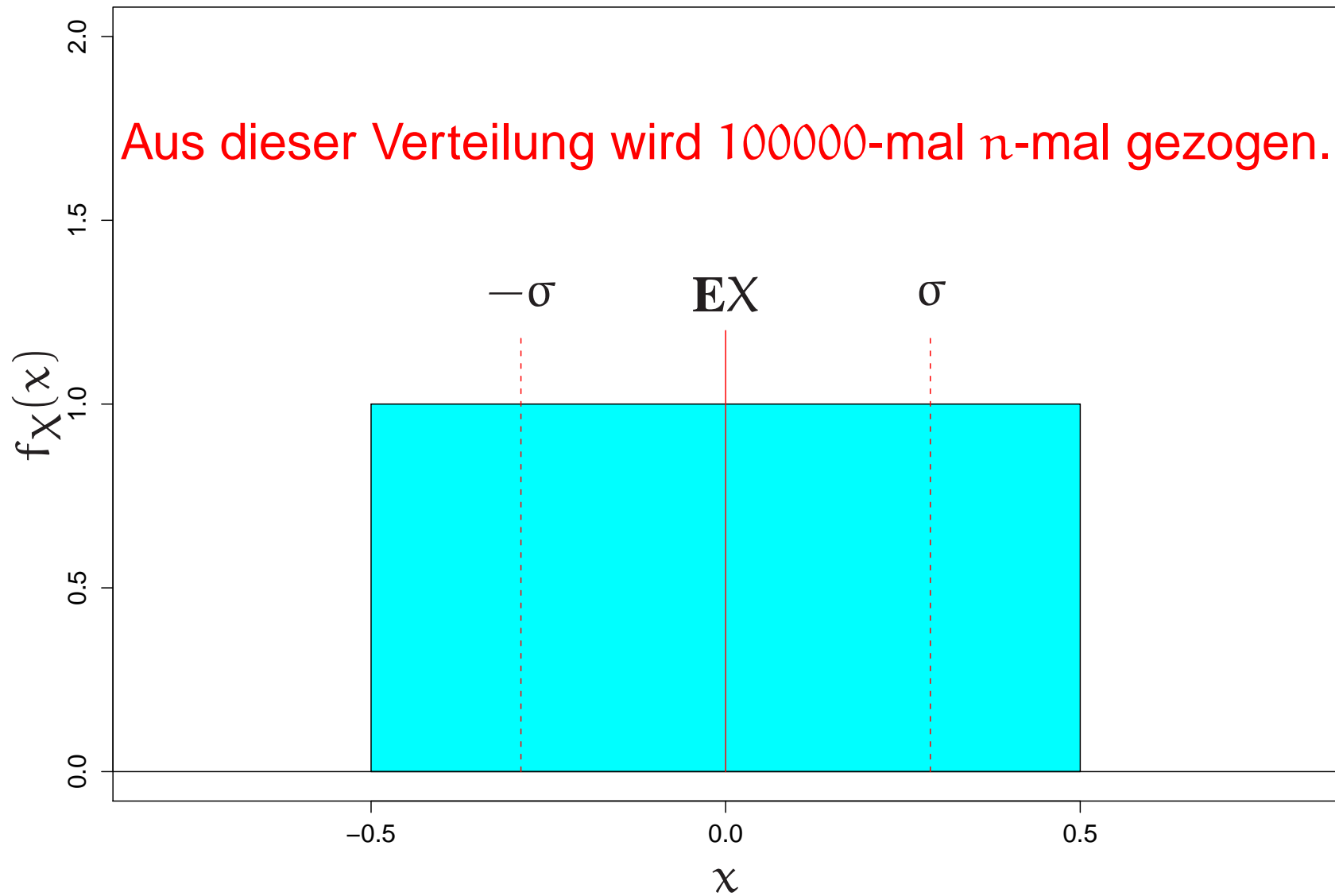
Dichtefunktion f_X der Verteilung von X



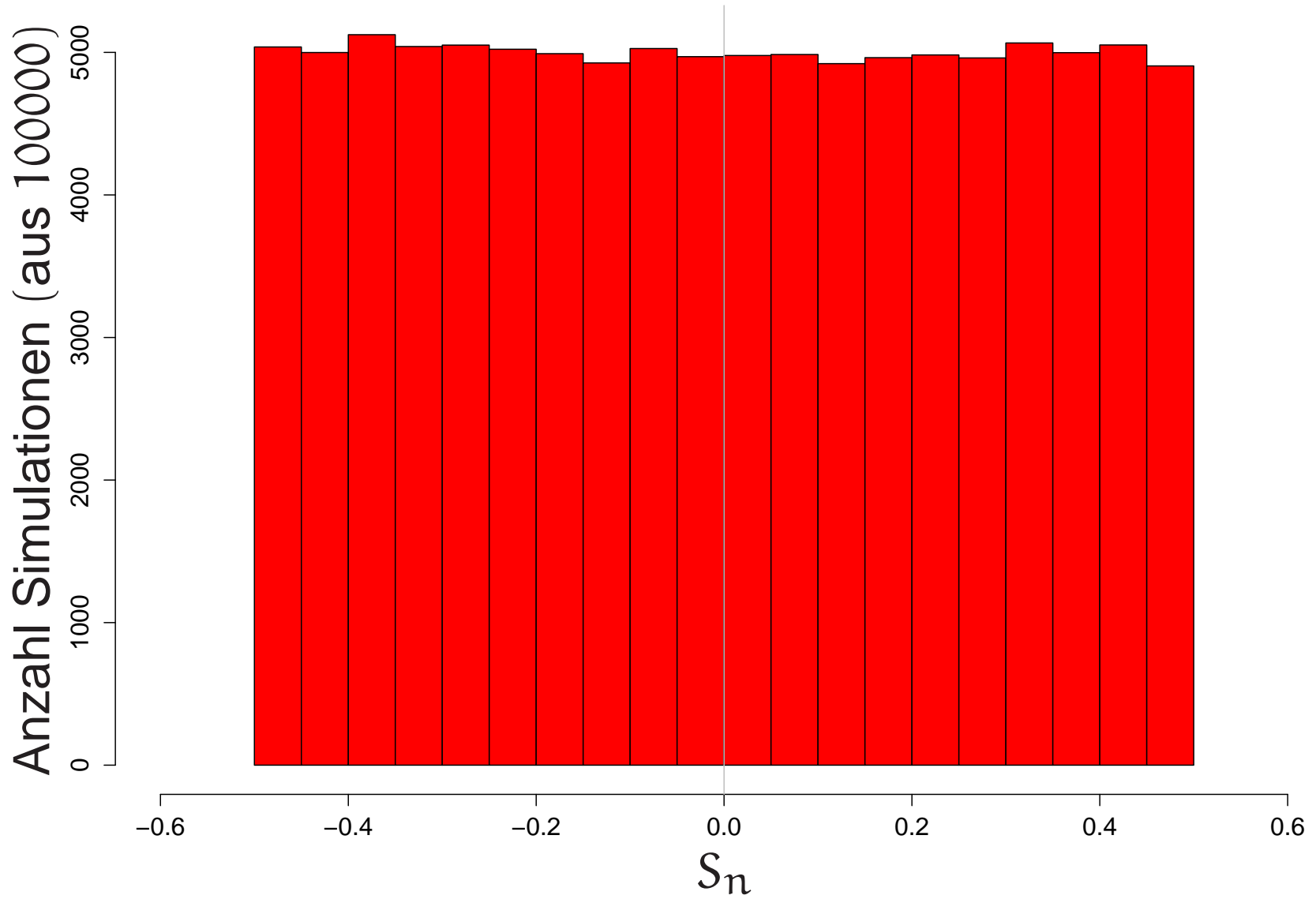
Dichtefunktion f_X der Verteilung von X



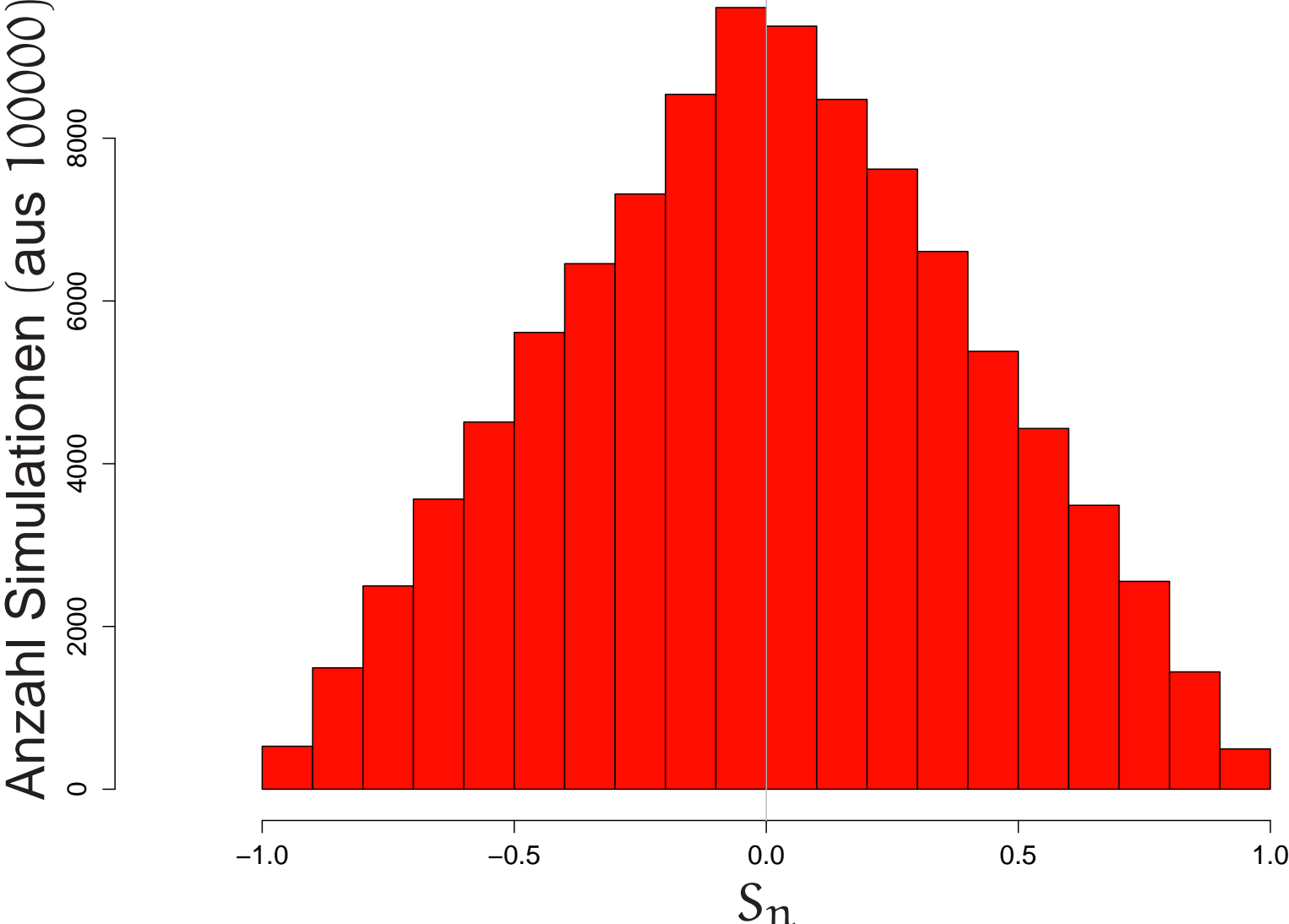
Dichtefunktion f_X der Verteilung von X



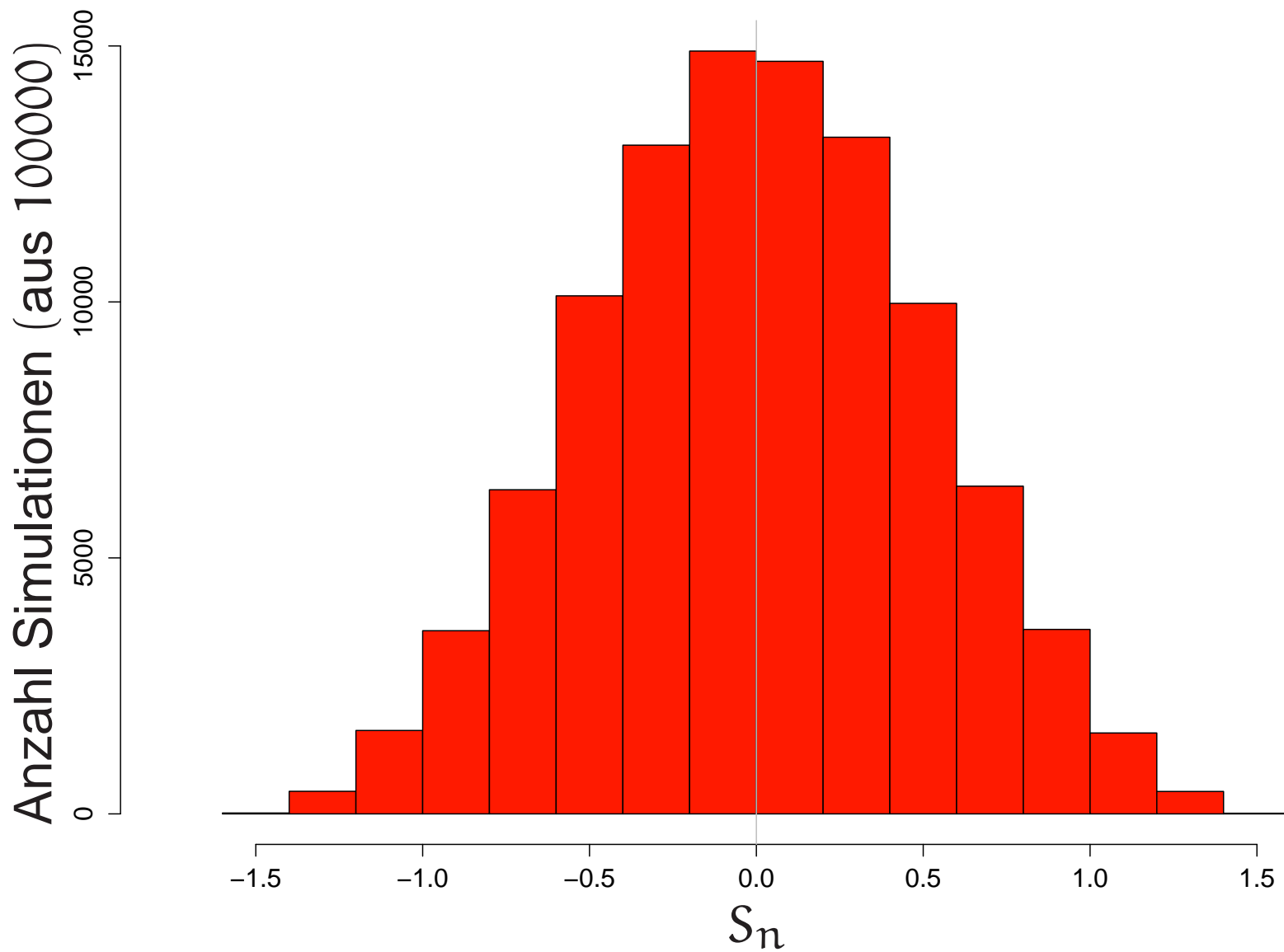
Verteilung von S_n ($n = 1$)



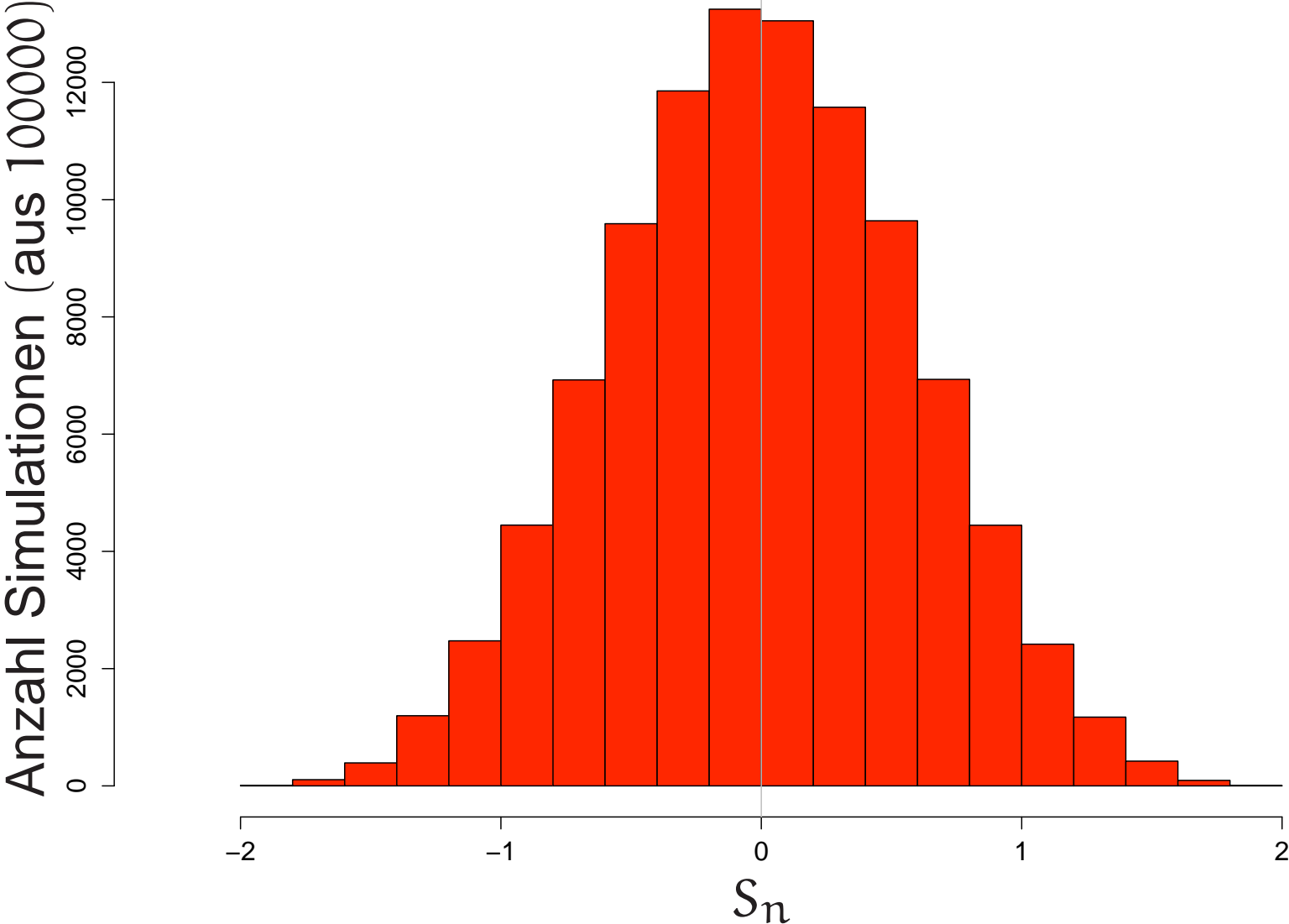
Verteilung von S_n ($n = 2$)



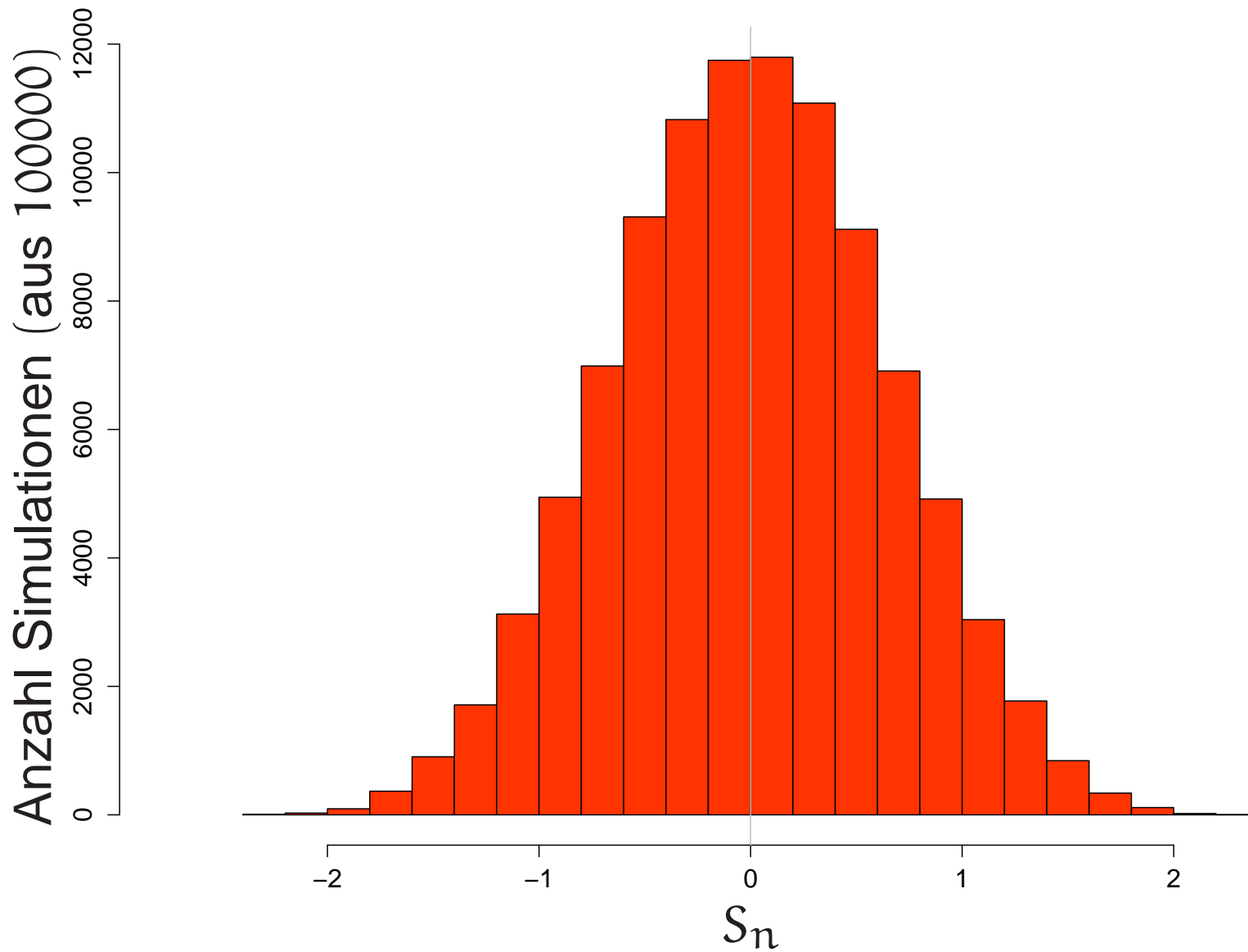
Verteilung von S_n ($n = 3$)



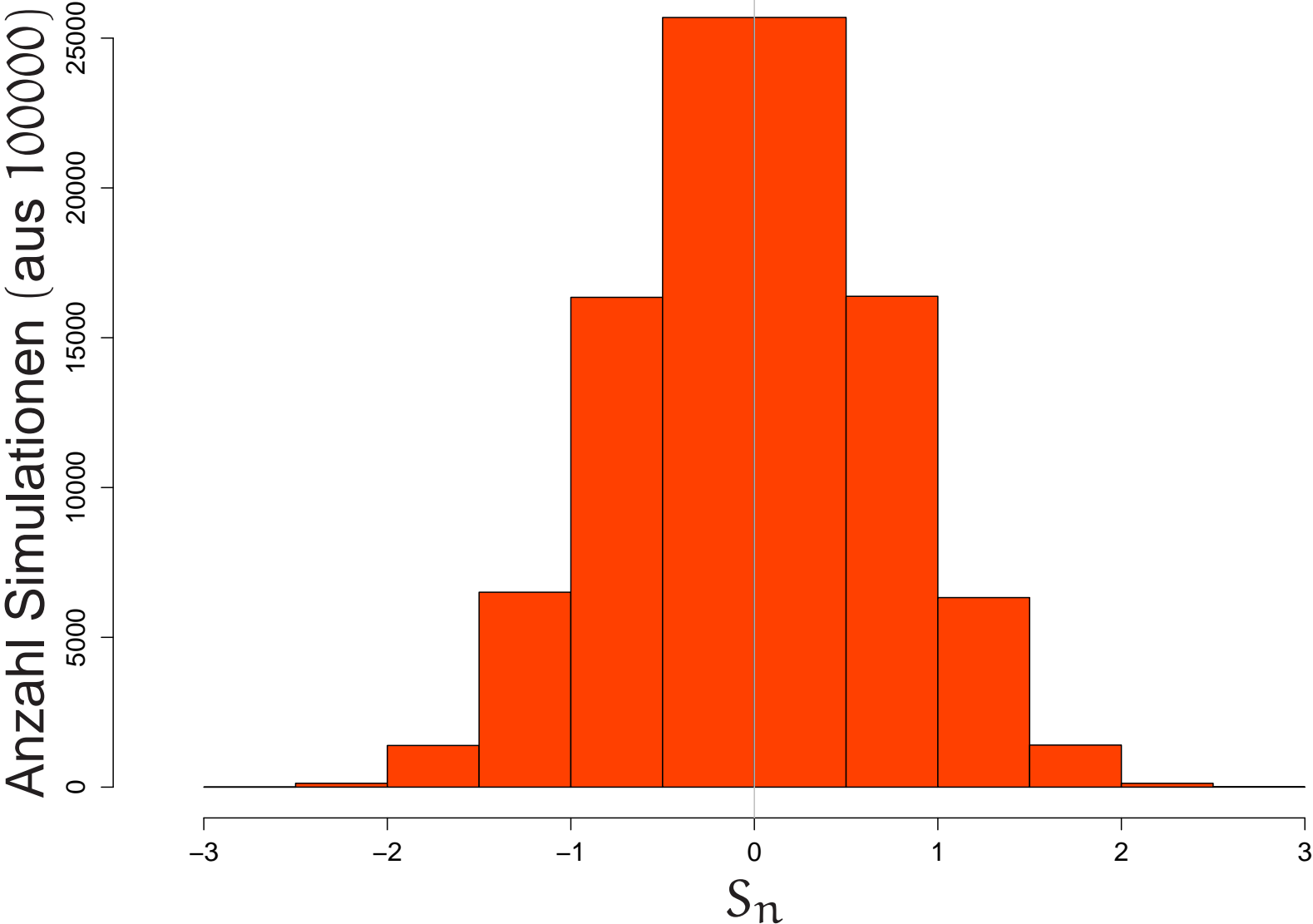
Verteilung von S_n ($n = 4$)



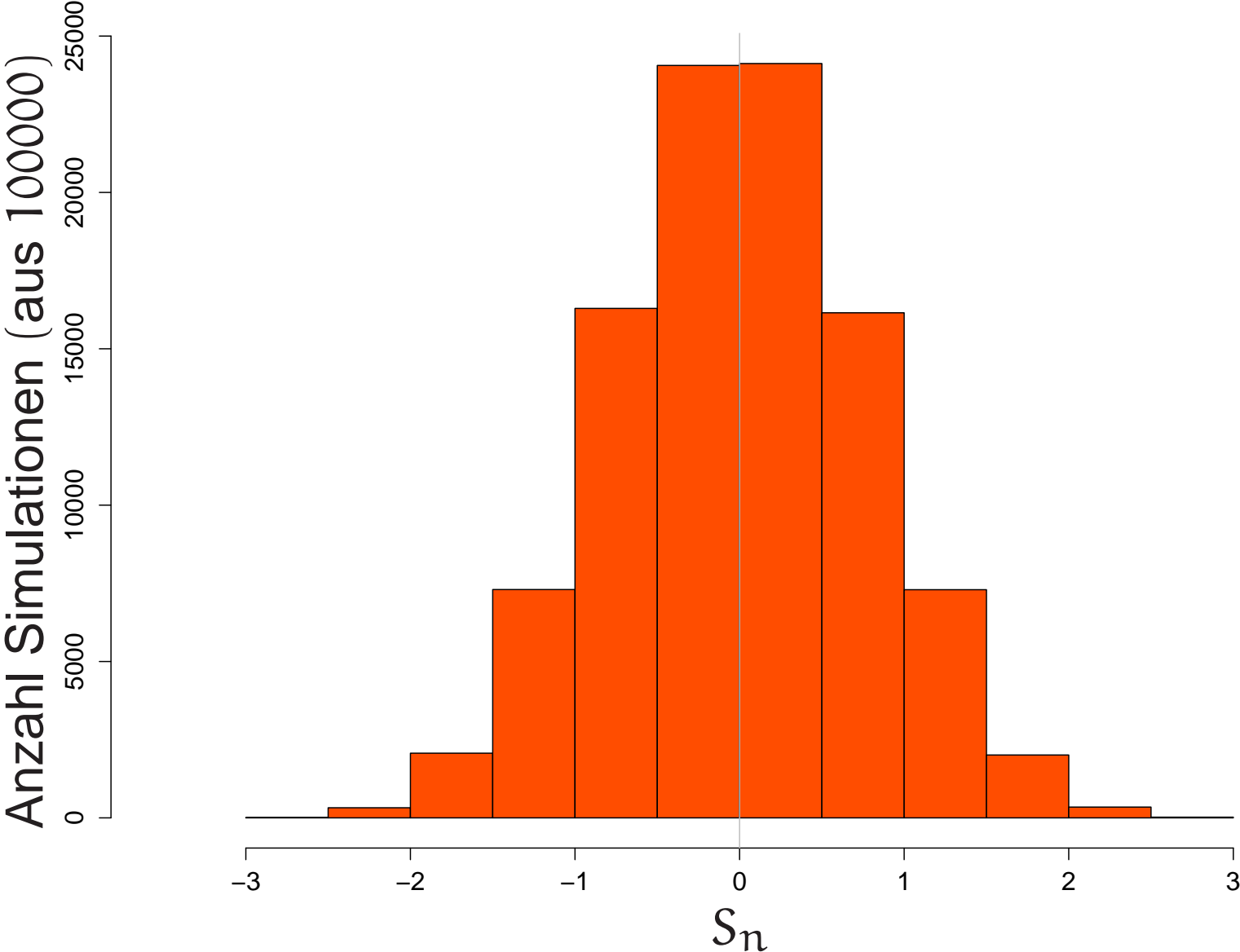
Verteilung von S_n ($n = 5$)



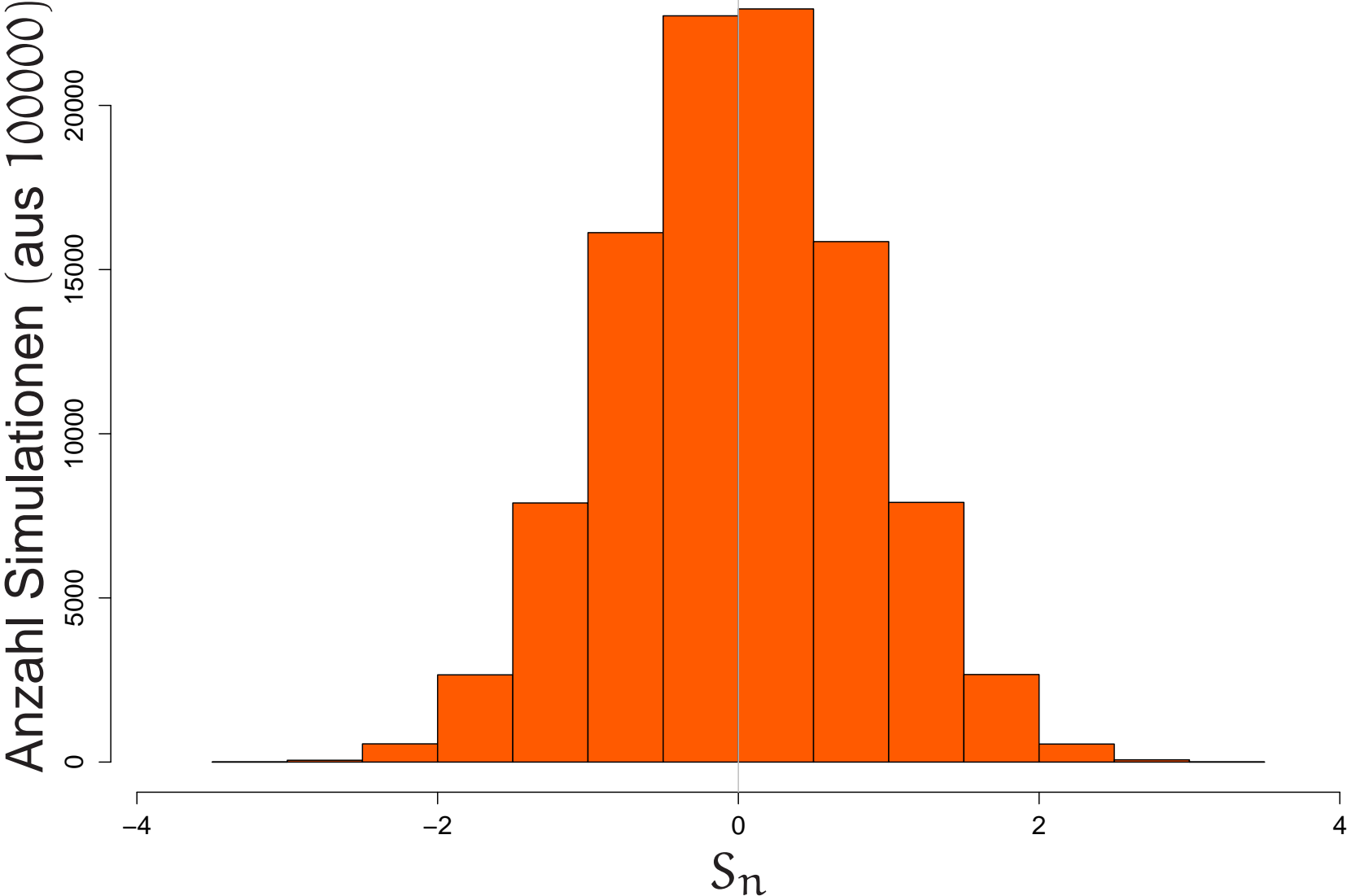
Verteilung von S_n ($n = 6$)



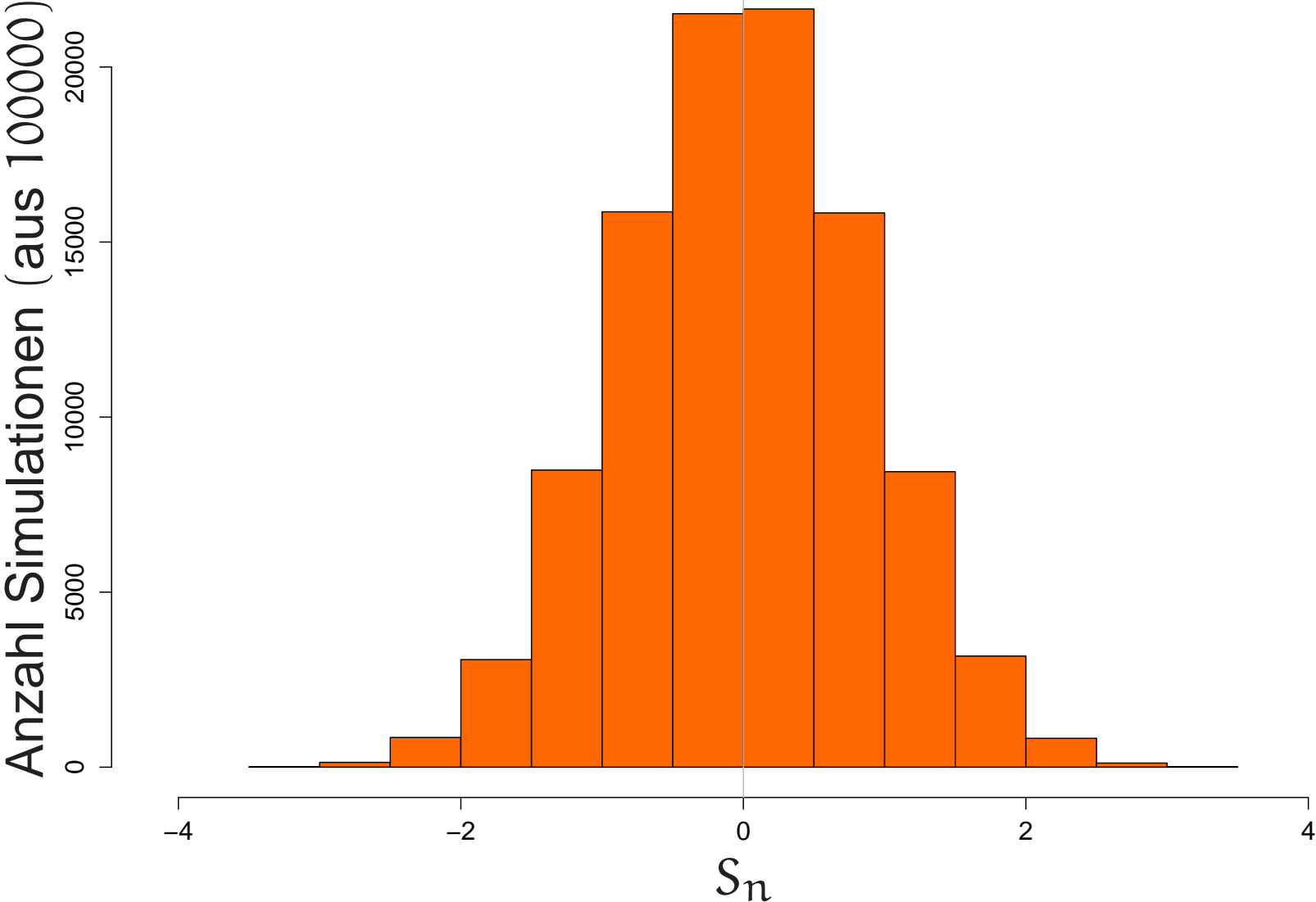
Verteilung von S_n ($n = 7$)



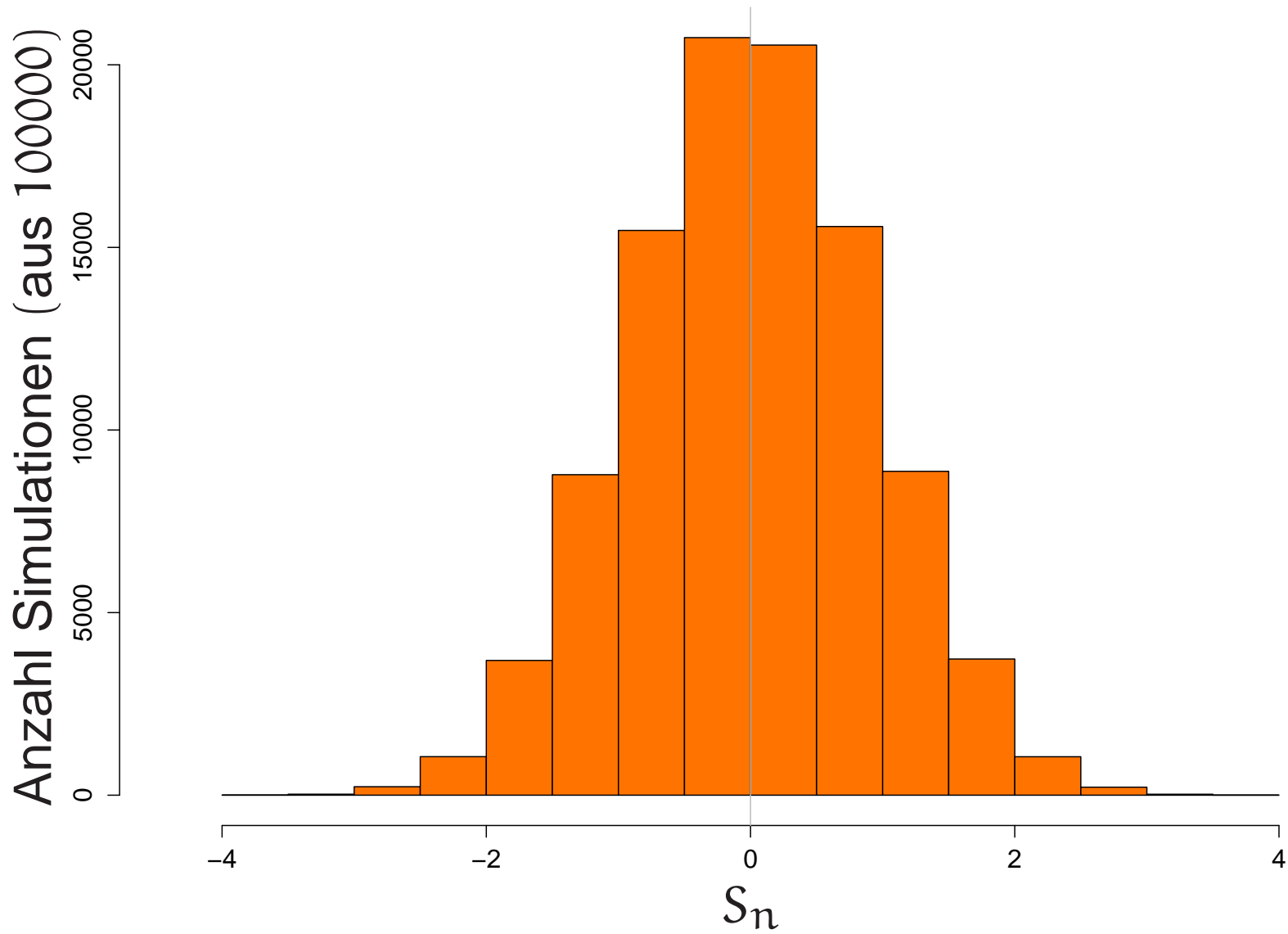
Verteilung von S_n ($n = 8$)



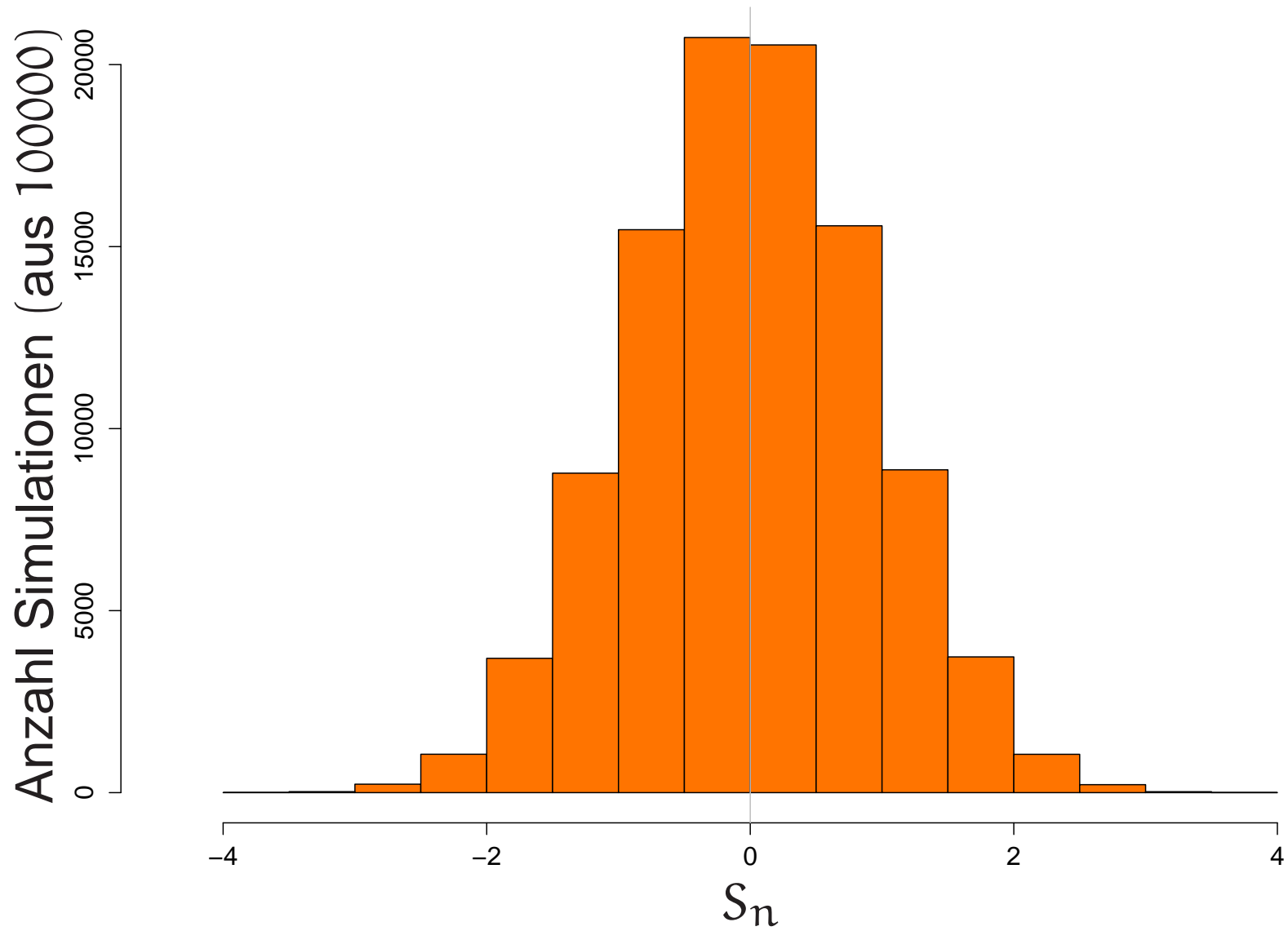
Verteilung von S_n ($n = 9$)



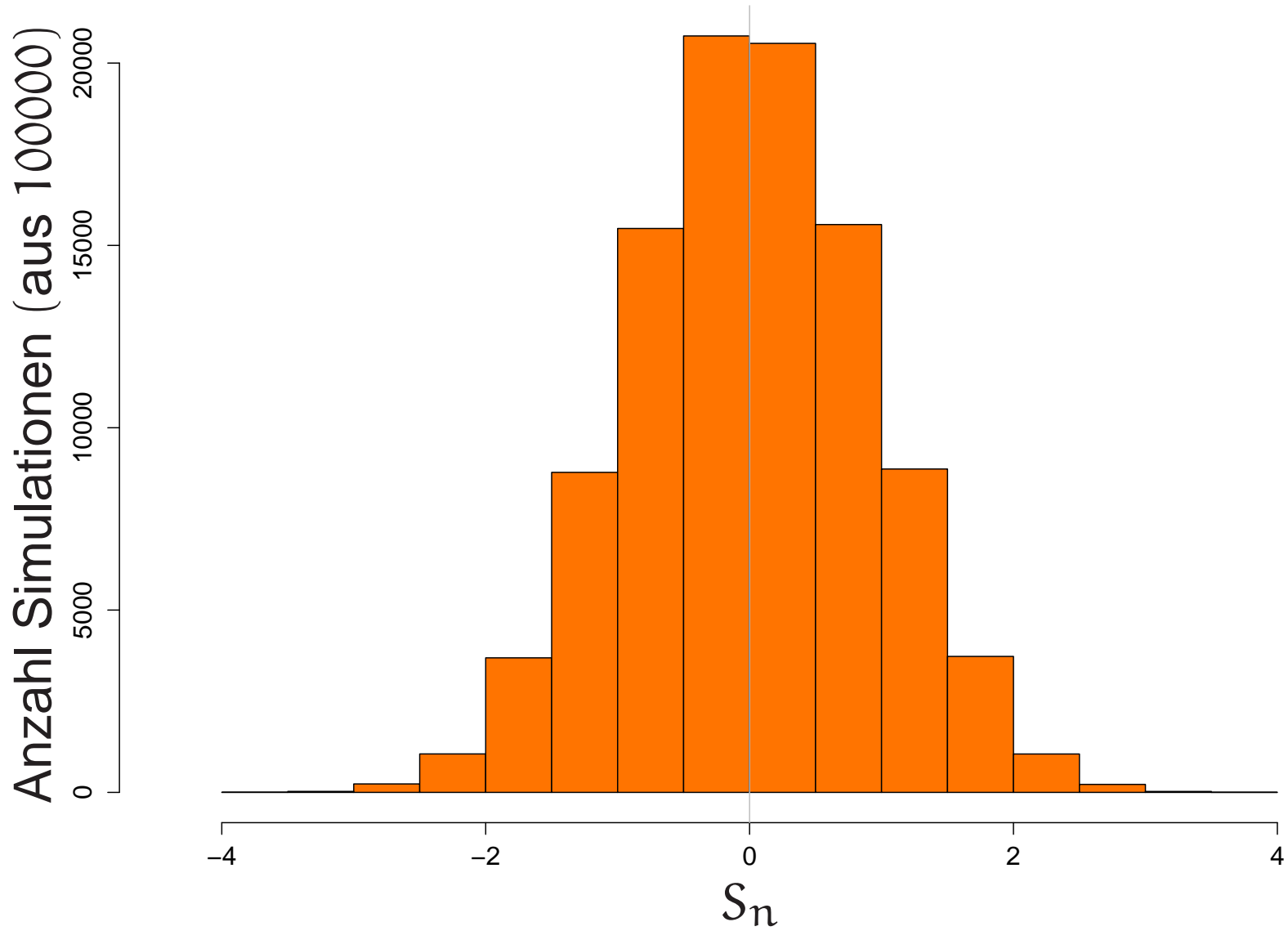
Verteilung von S_n ($n = 10$)



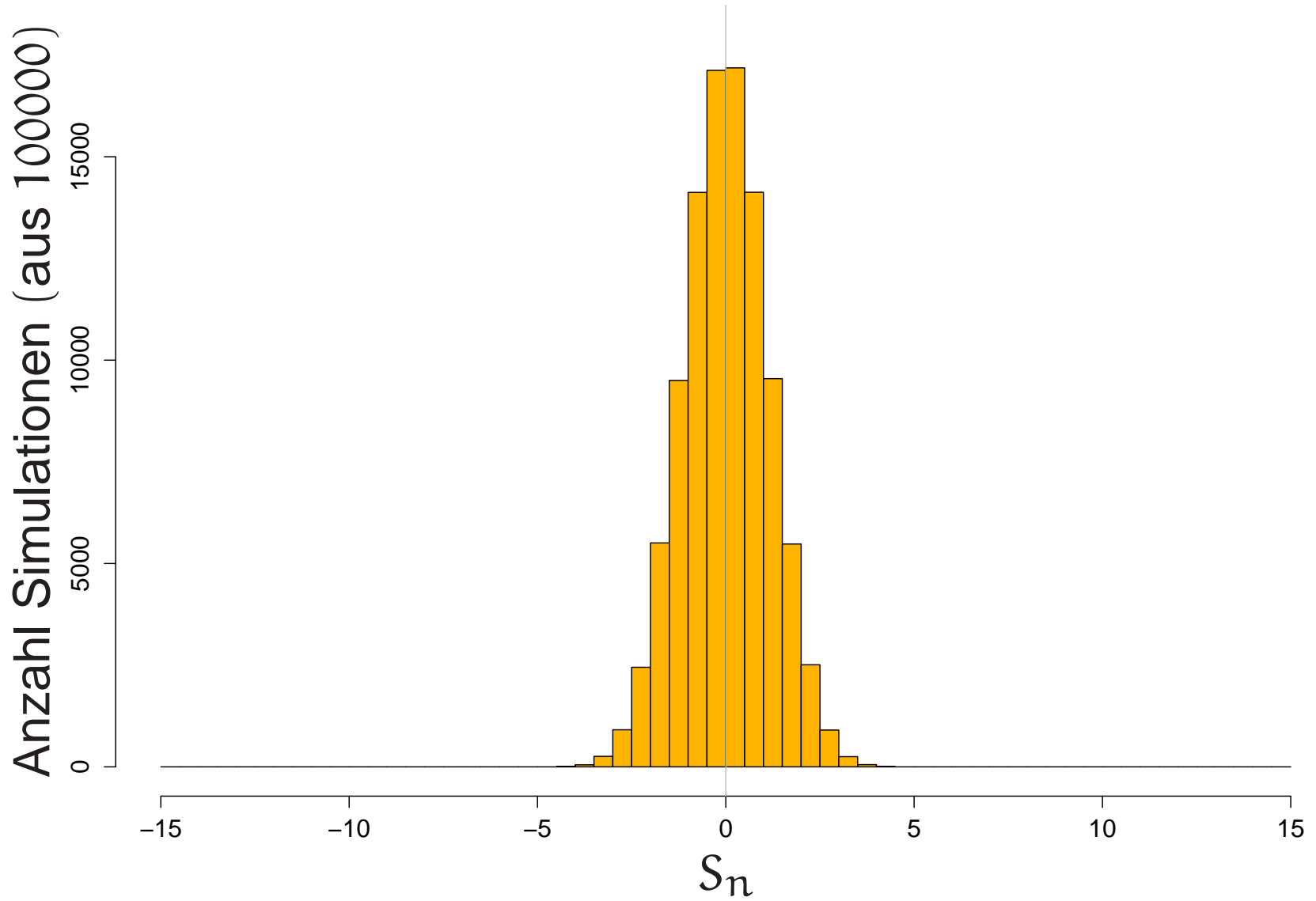
Bisher: dynamische Skalierung



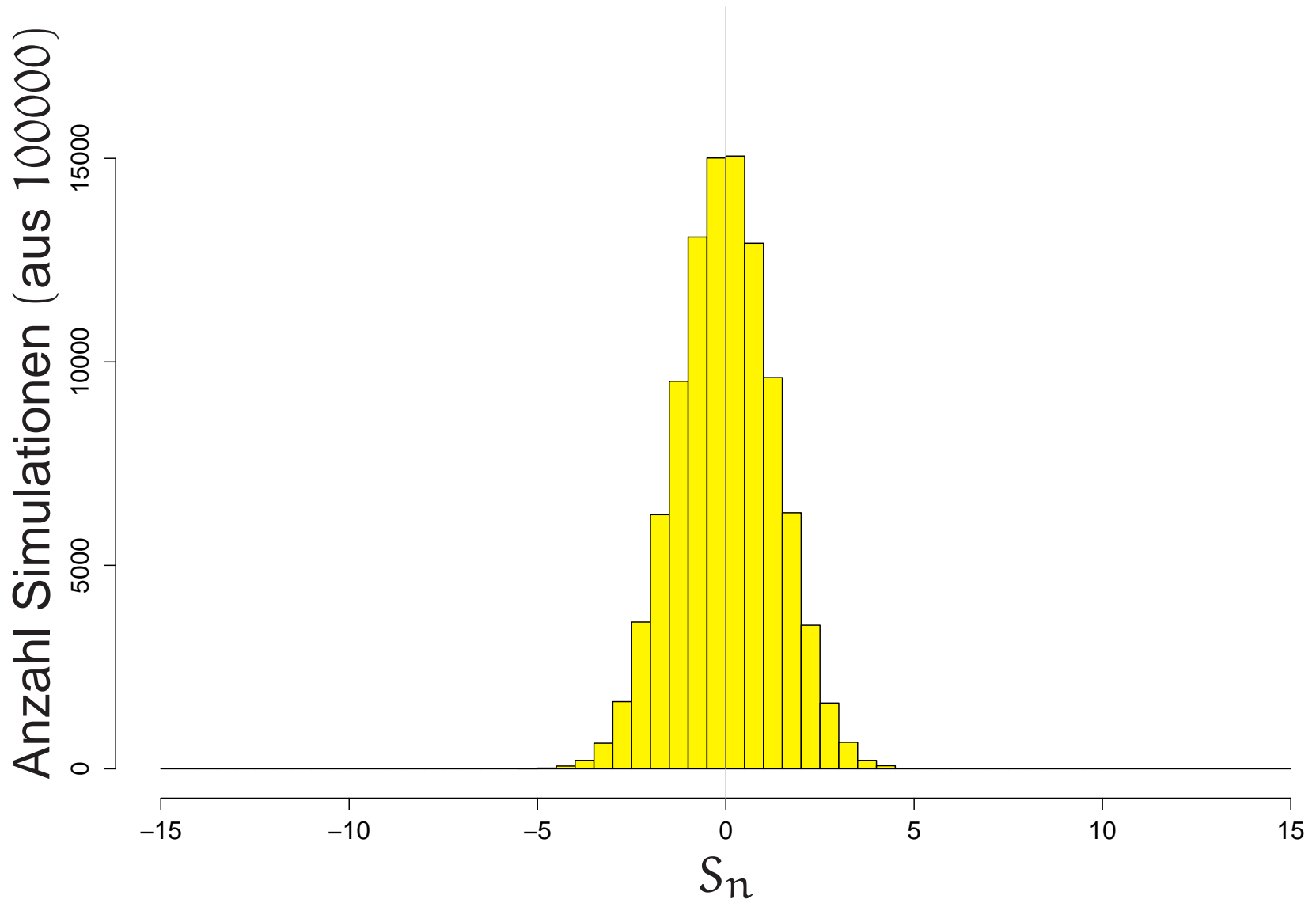
Jetzt: feste Skalierung



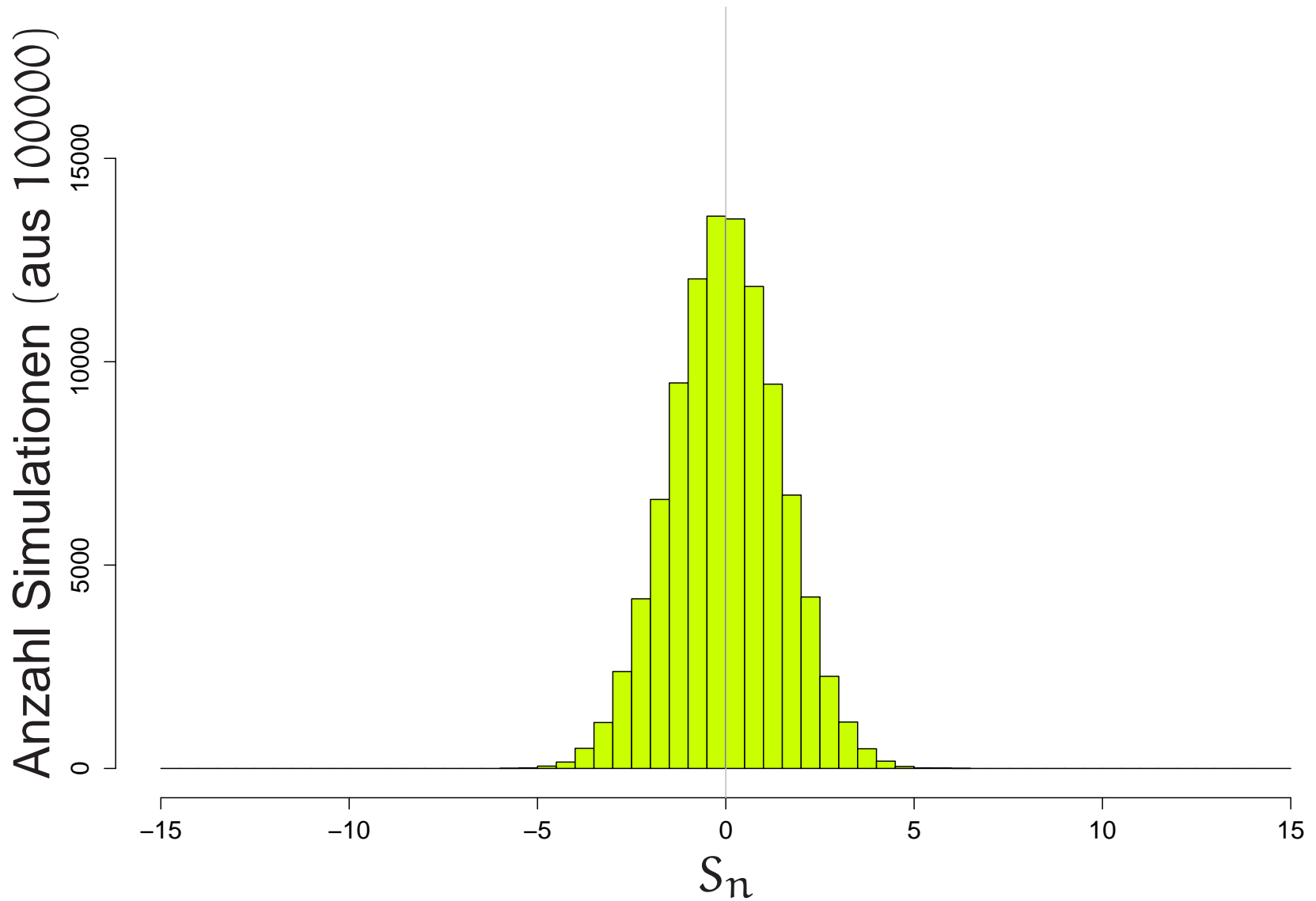
Verteilung von S_n ($n = 15$)



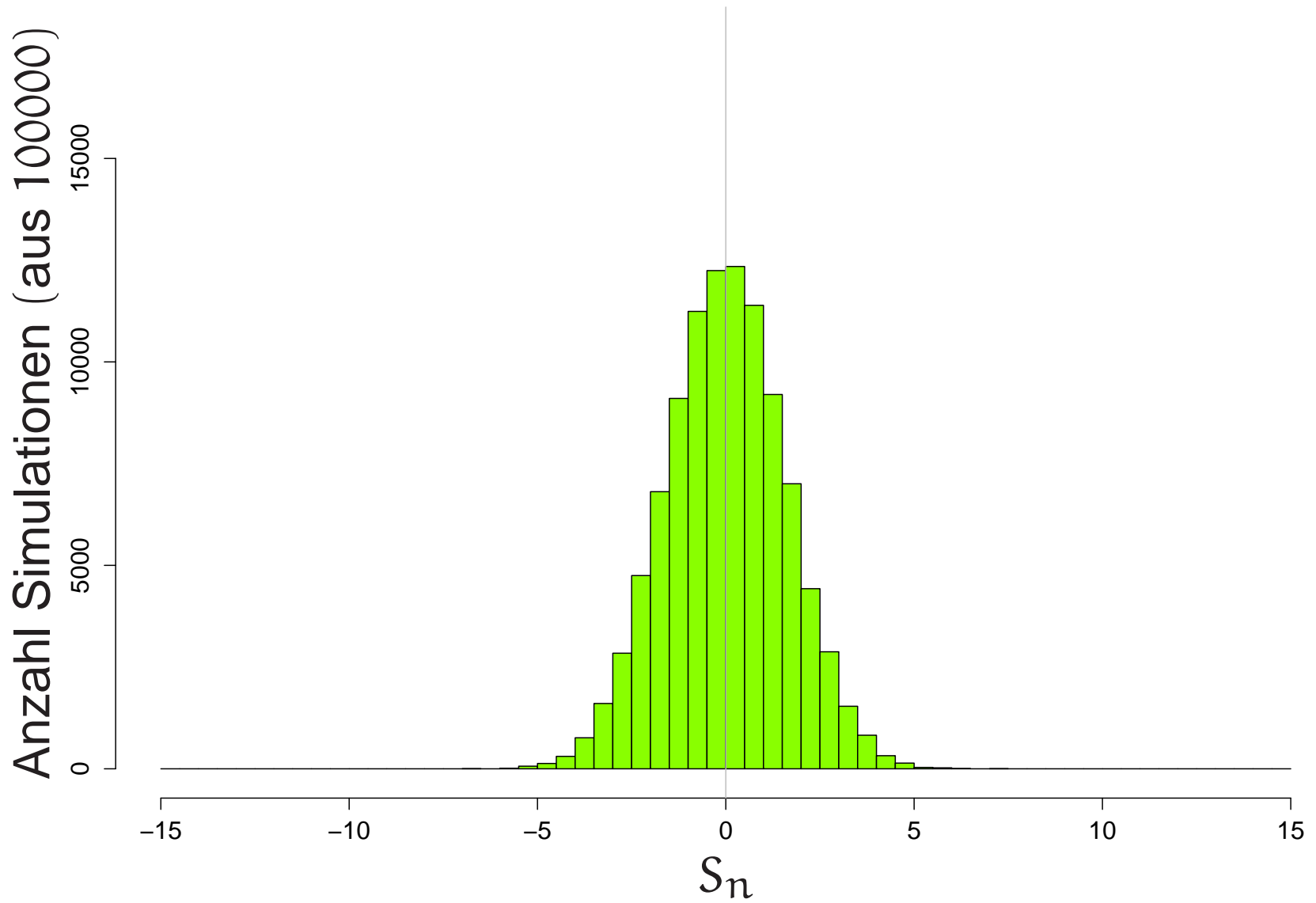
Verteilung von S_n ($n = 20$)



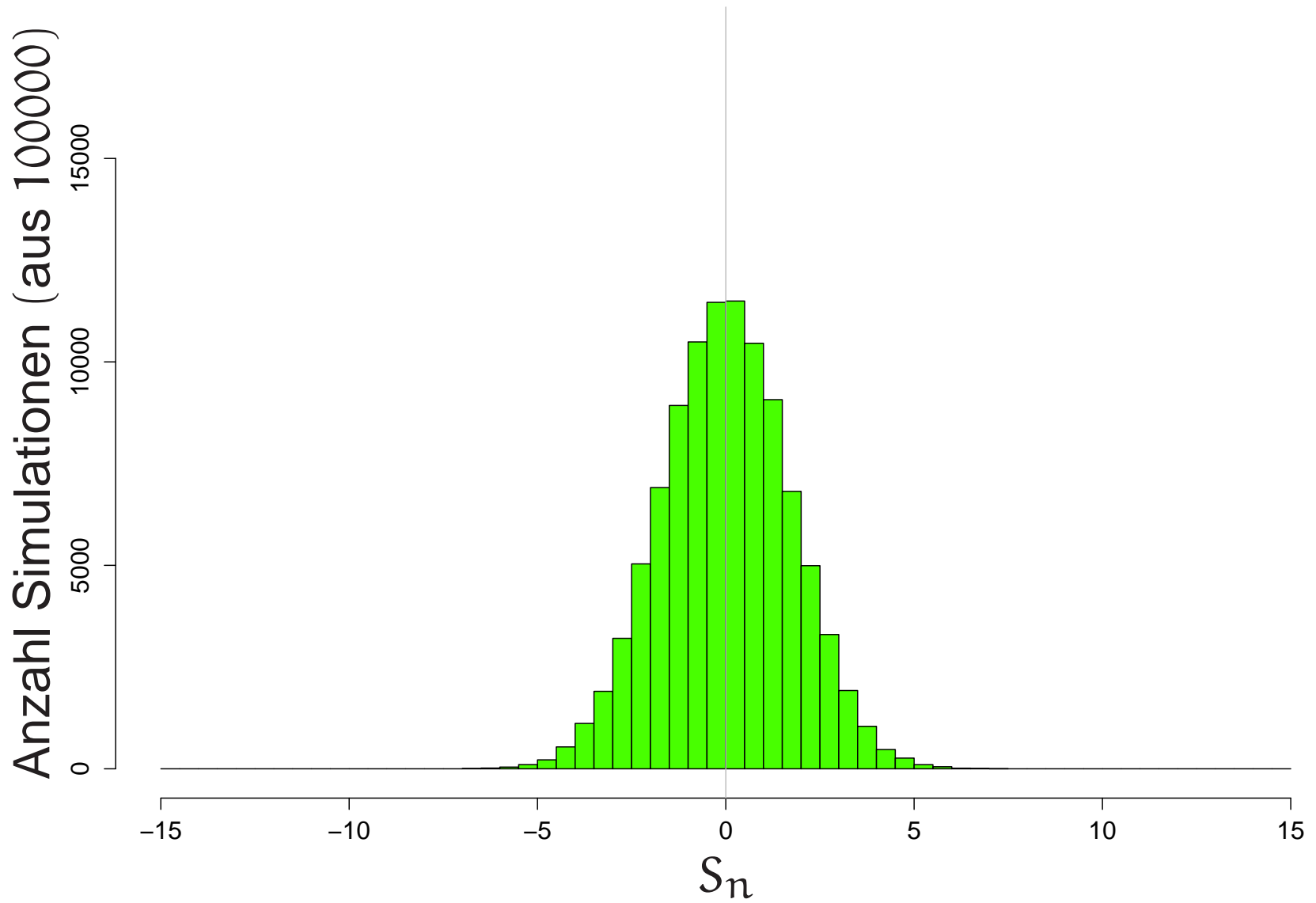
Verteilung von S_n ($n = 25$)



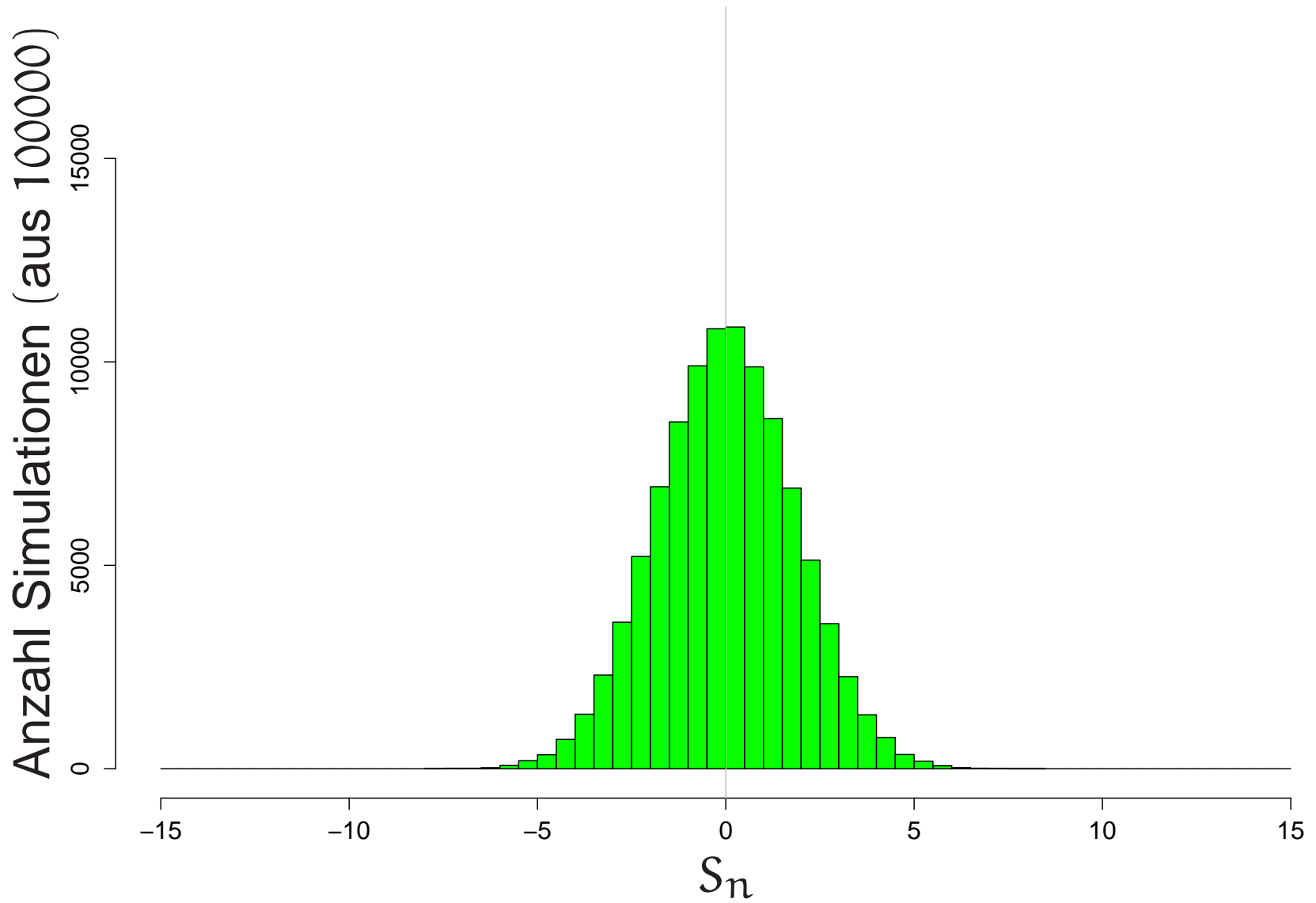
Verteilung von S_n ($n = 30$)



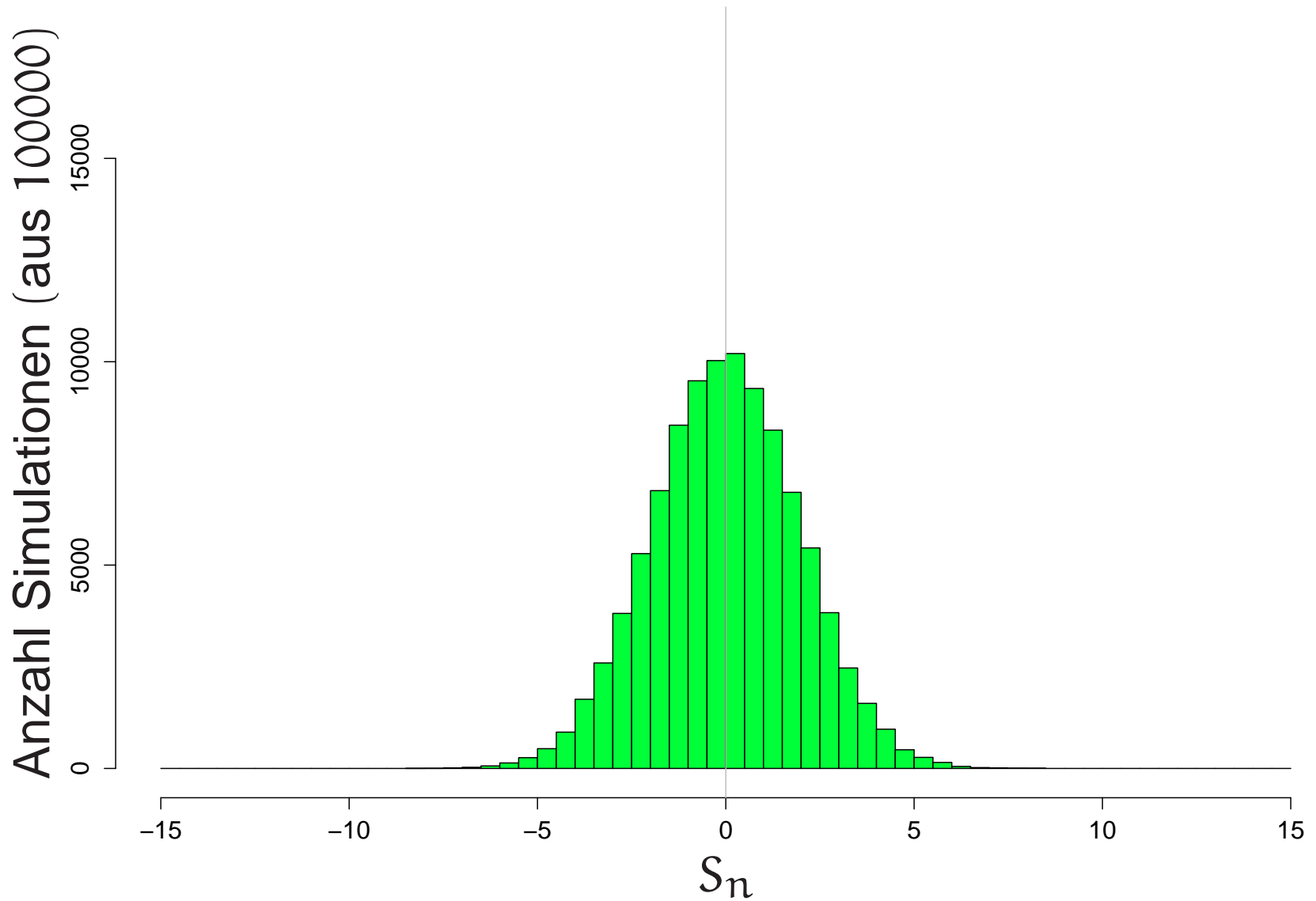
Verteilung von S_n ($n = 35$)



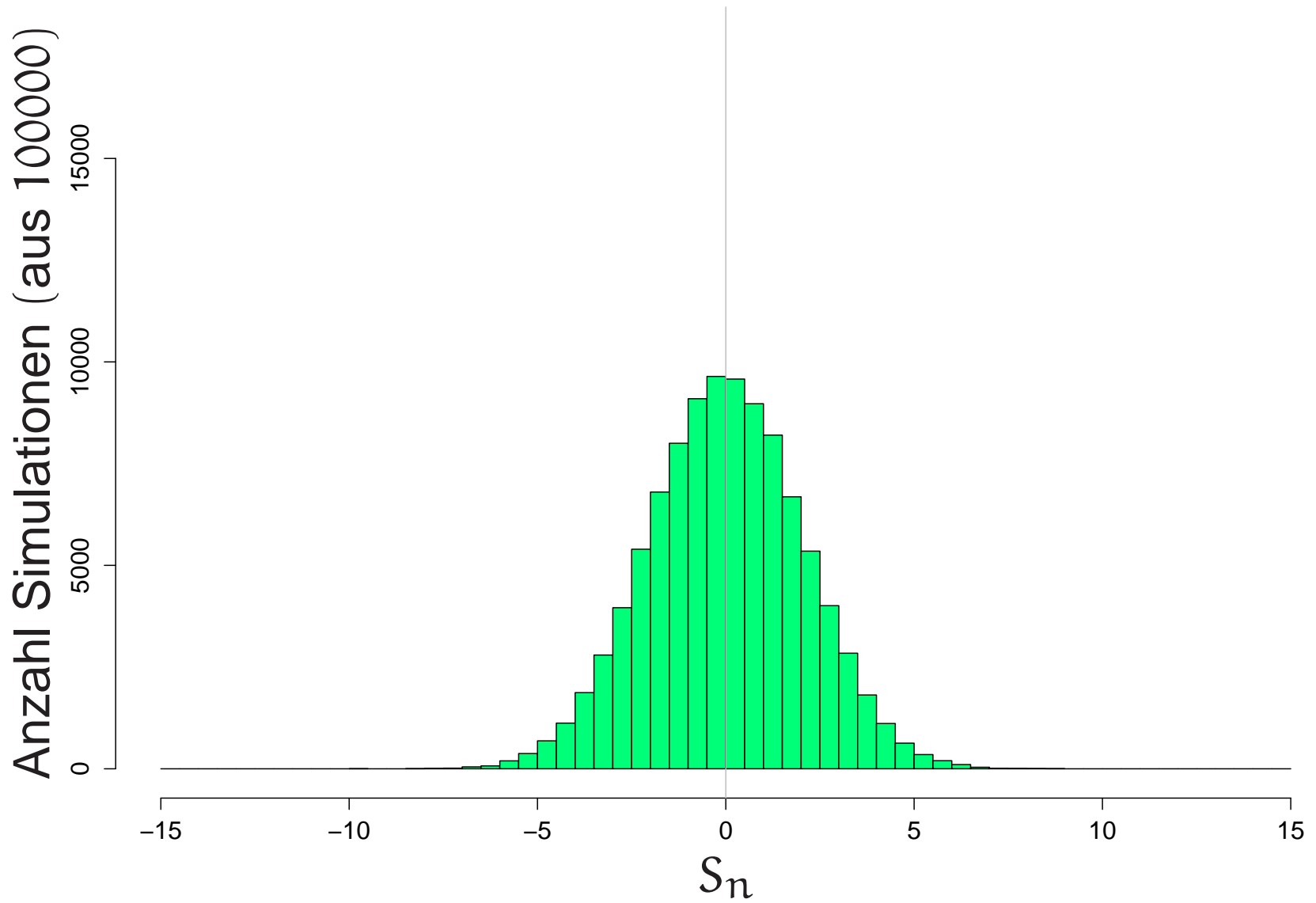
Verteilung von S_n ($n = 40$)



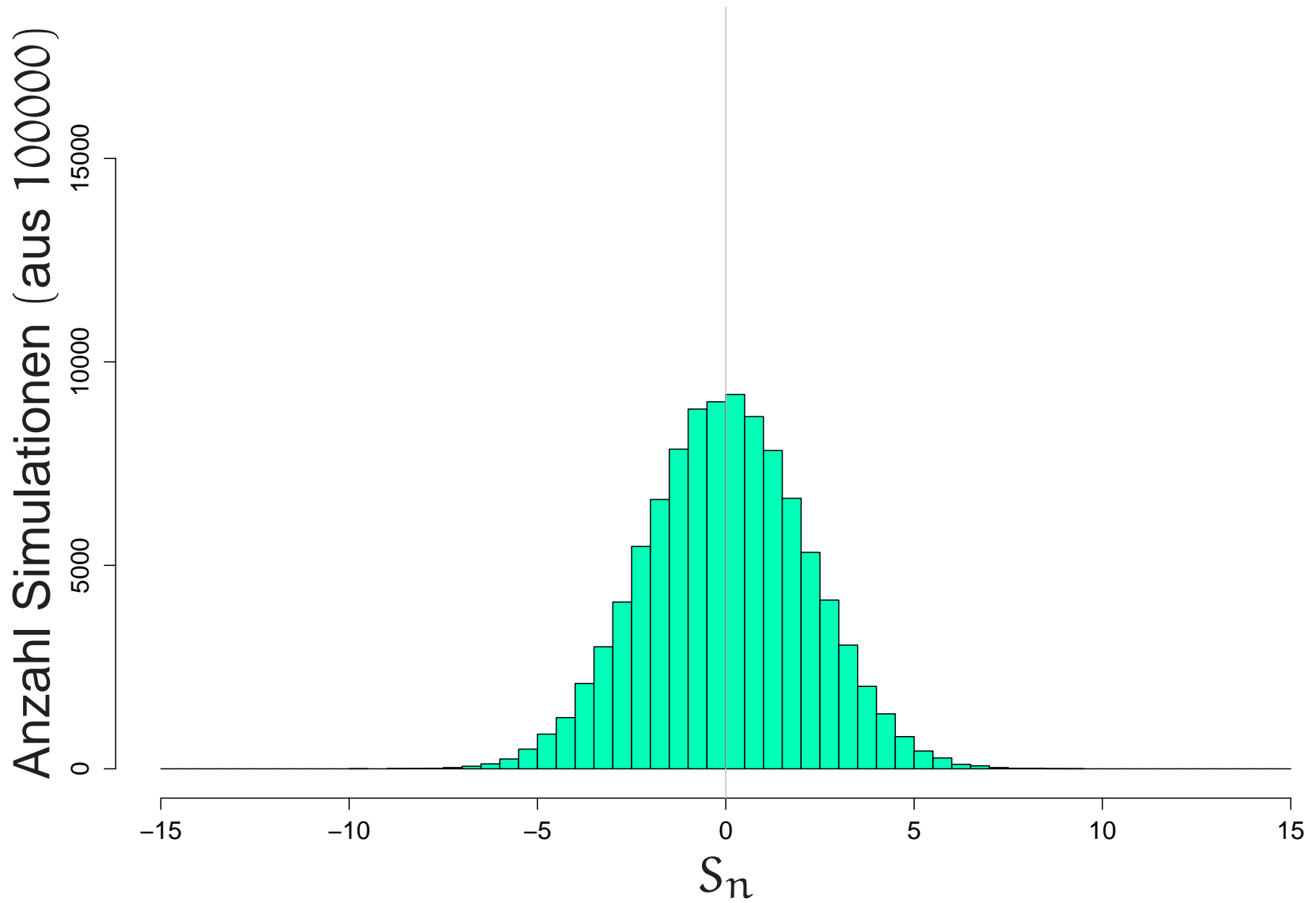
Verteilung von S_n ($n = 45$)



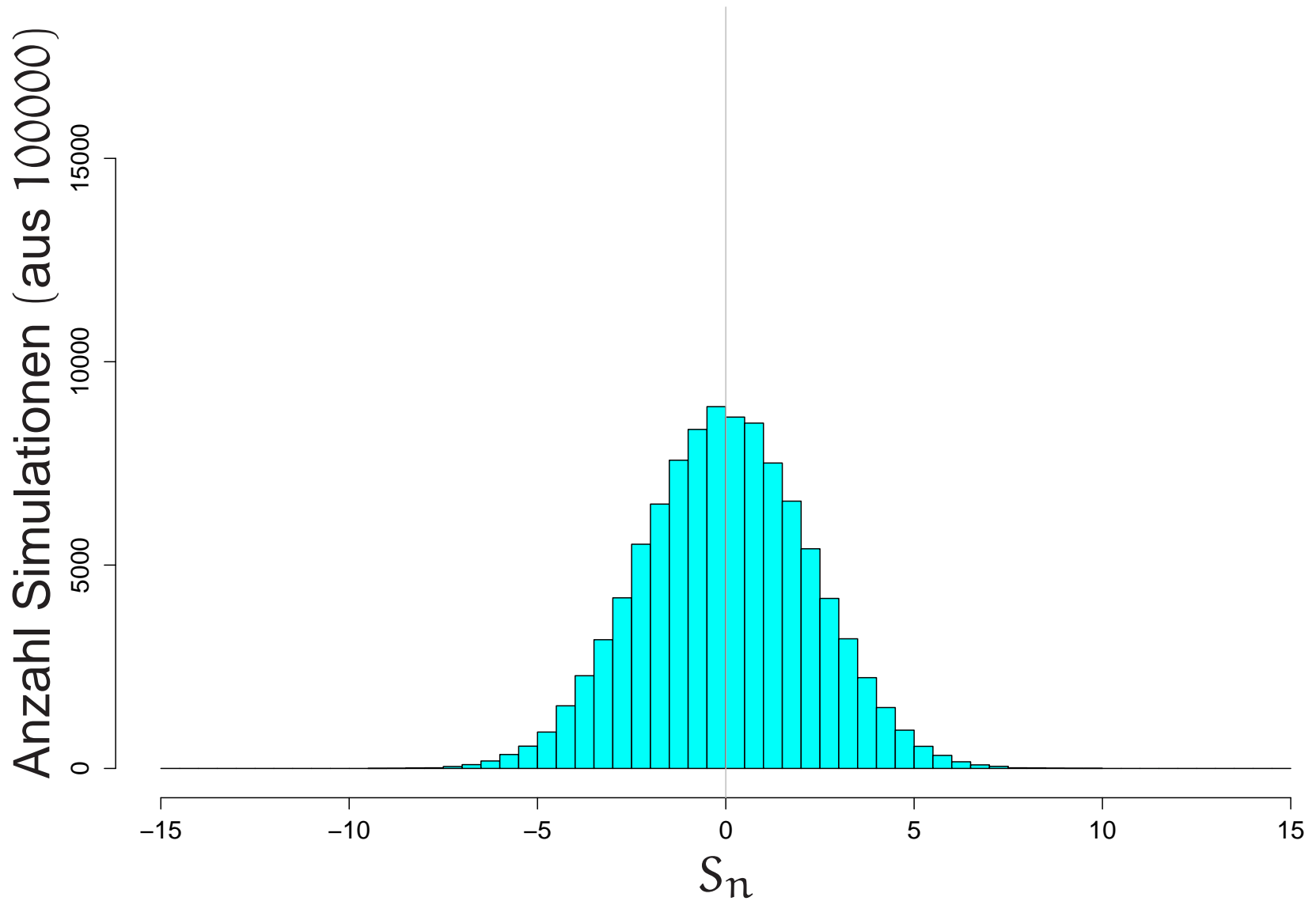
Verteilung von S_n ($n = 50$)



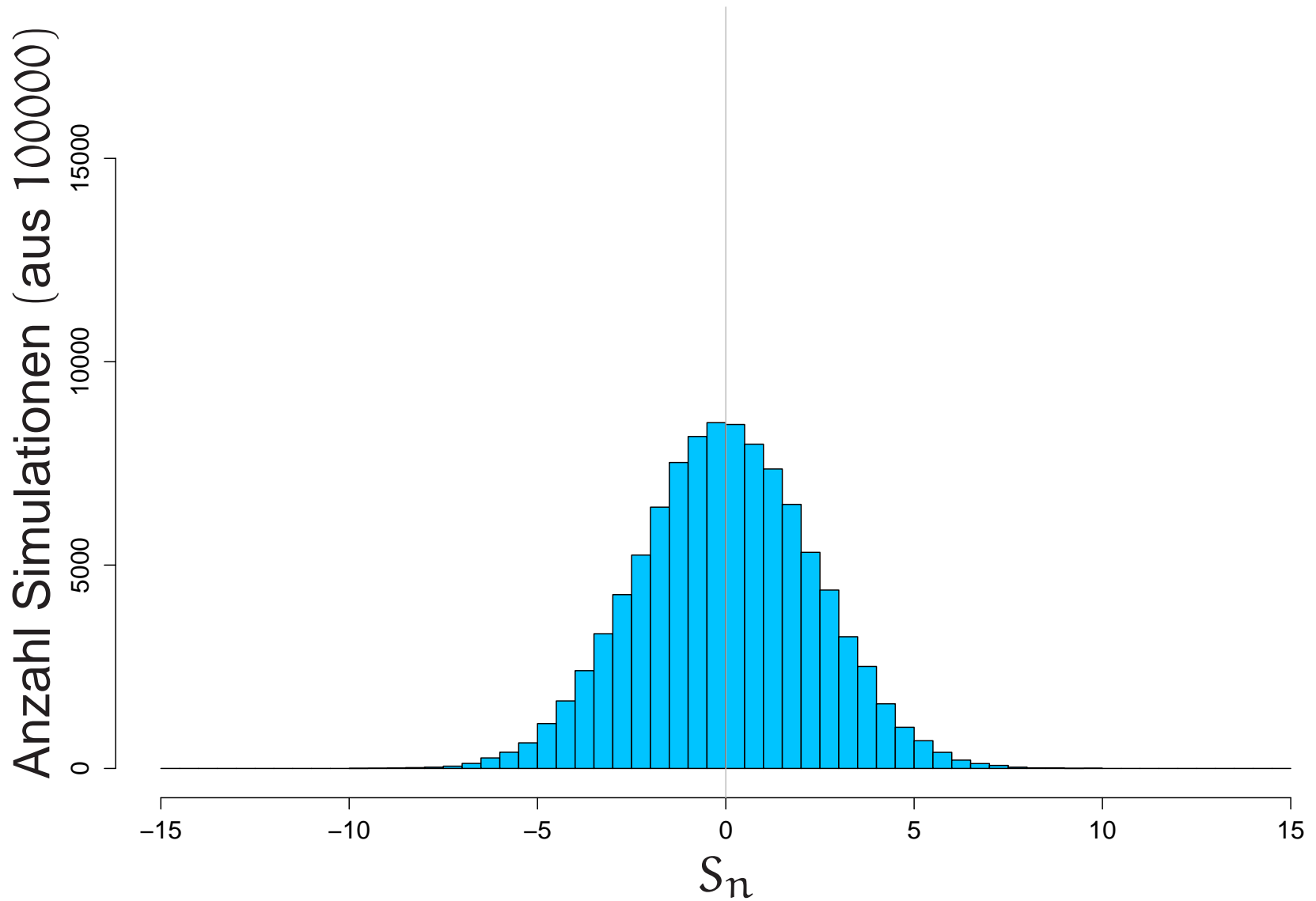
Verteilung von S_n ($n = 55$)



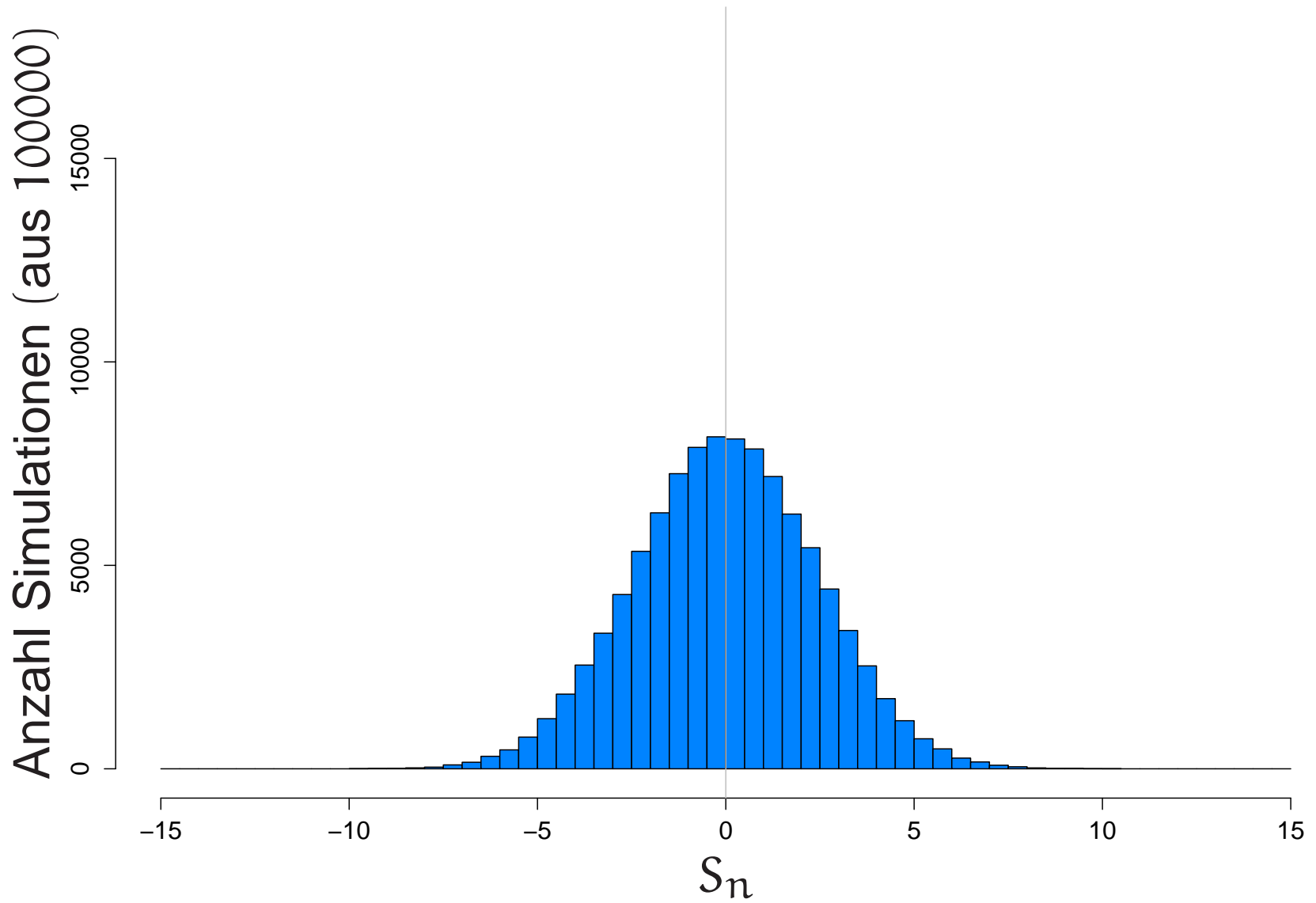
Verteilung von S_n ($n = 60$)



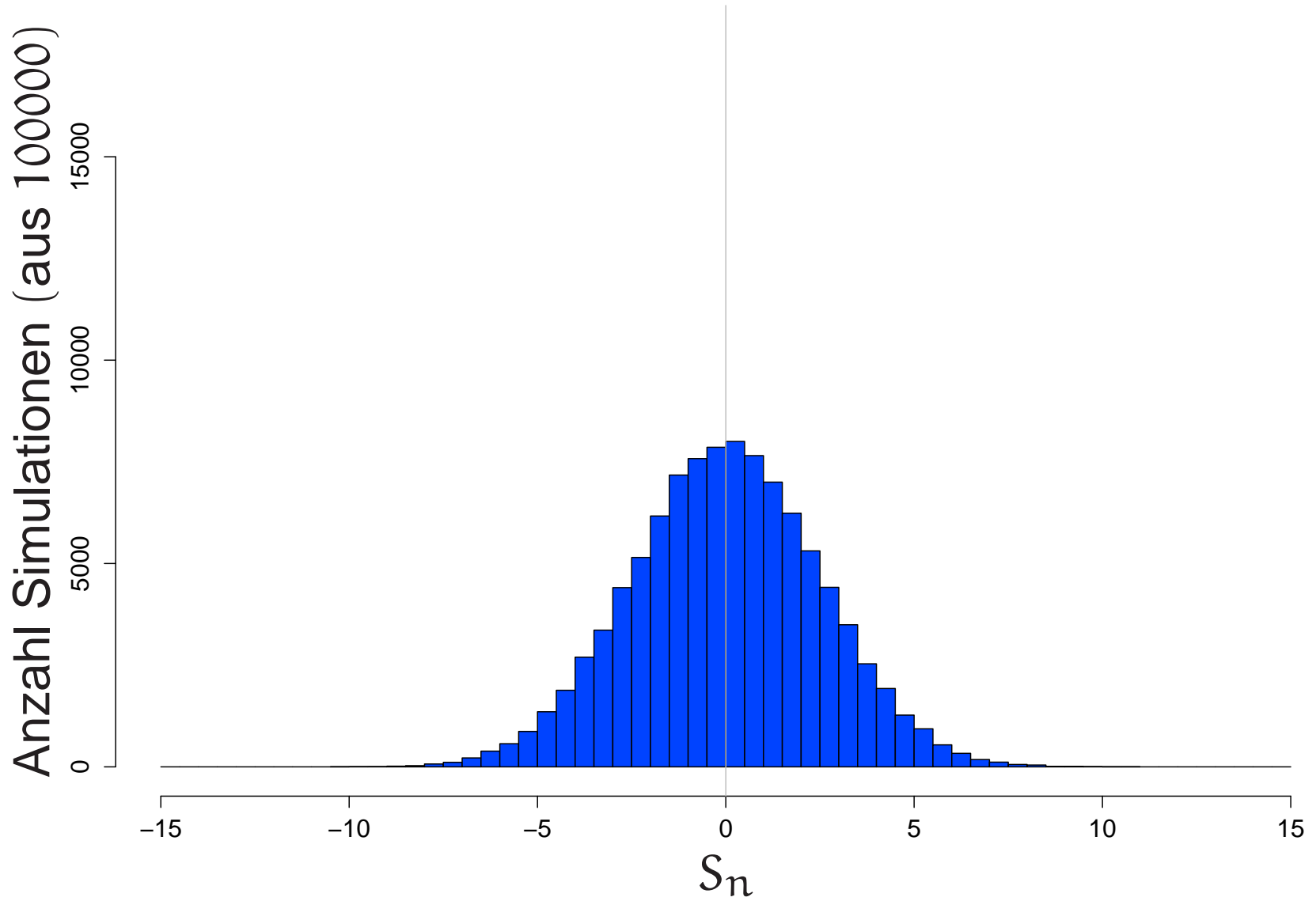
Verteilung von S_n ($n = 65$)



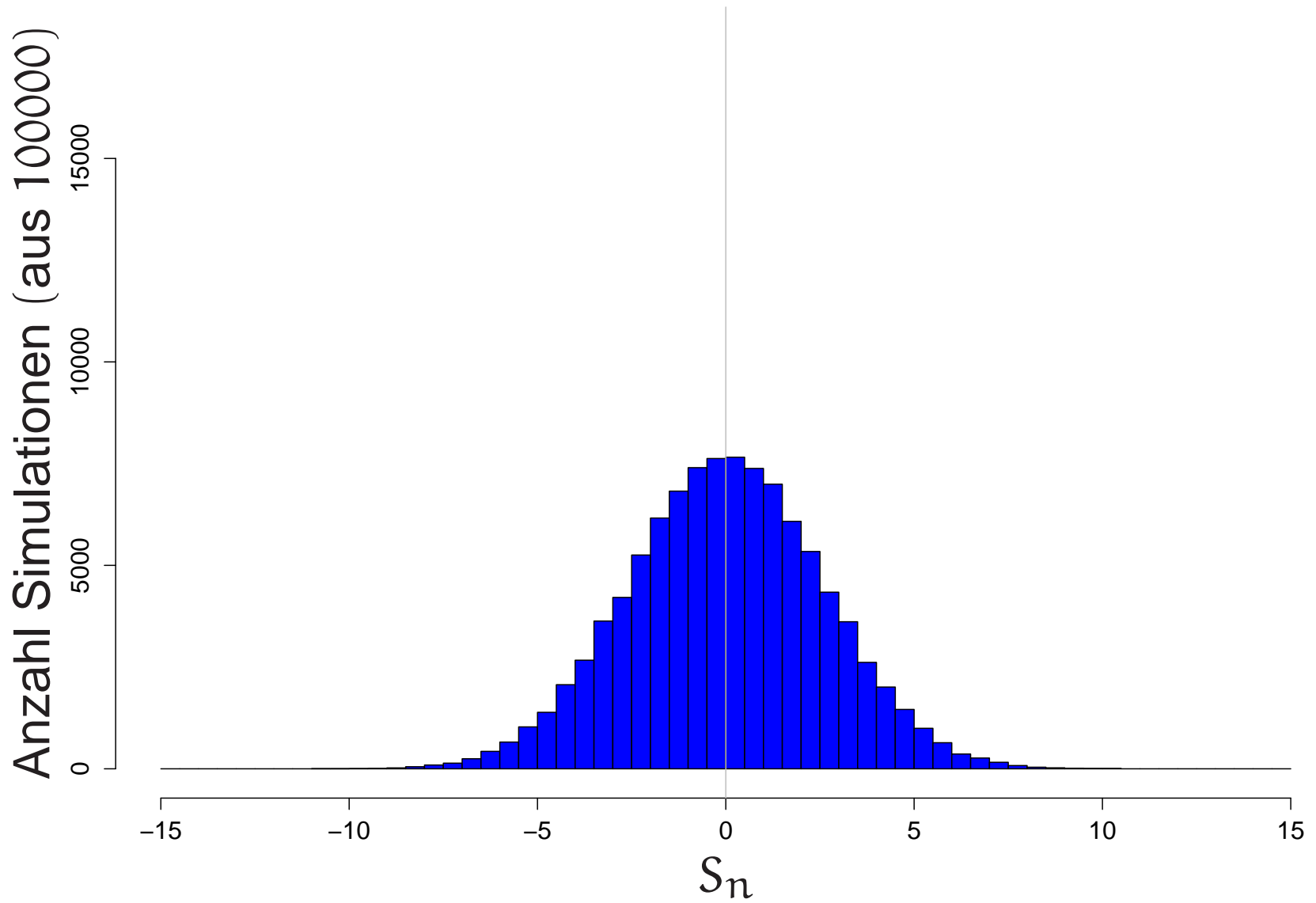
Verteilung von S_n ($n = 70$)



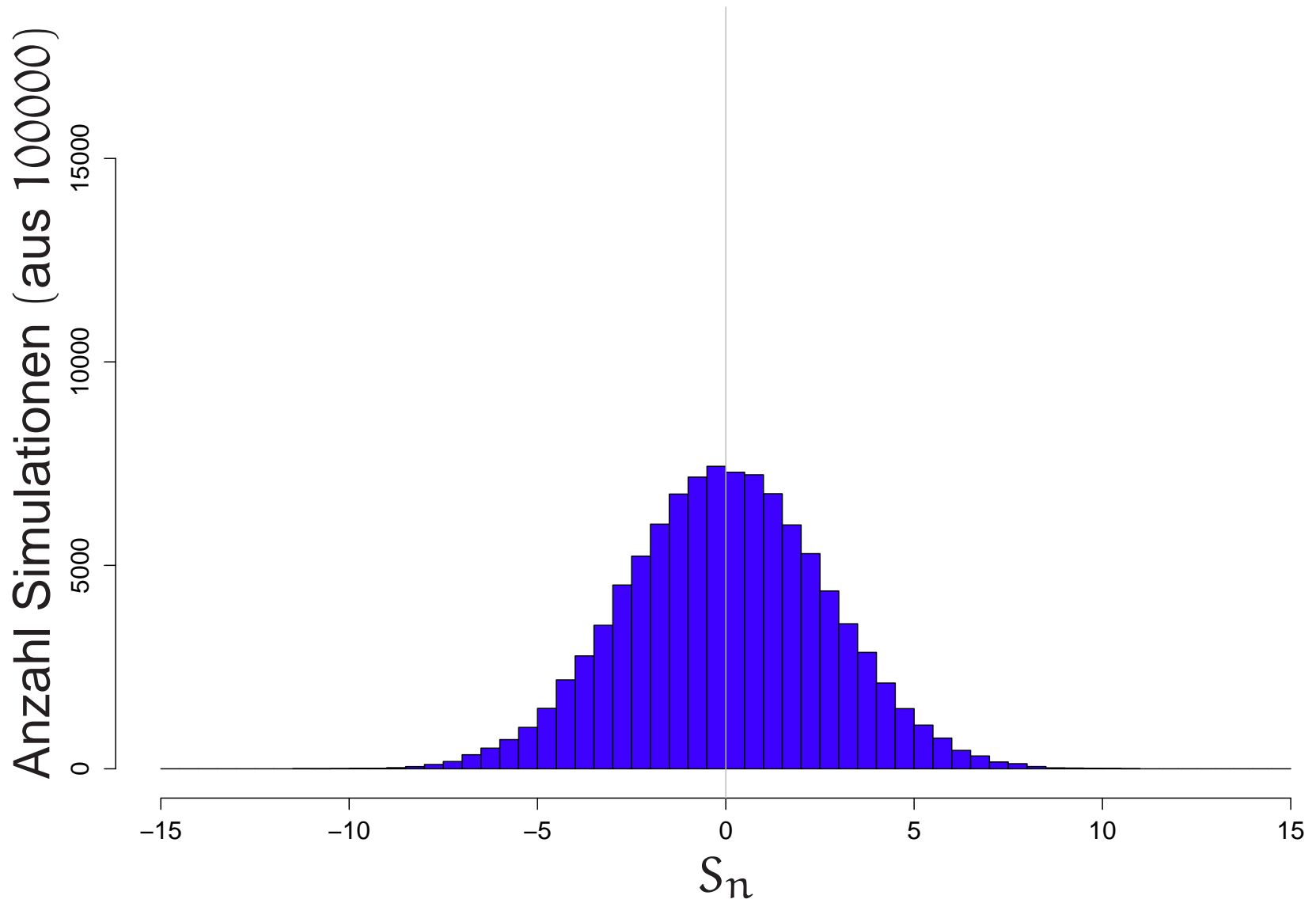
Verteilung von S_n ($n = 75$)



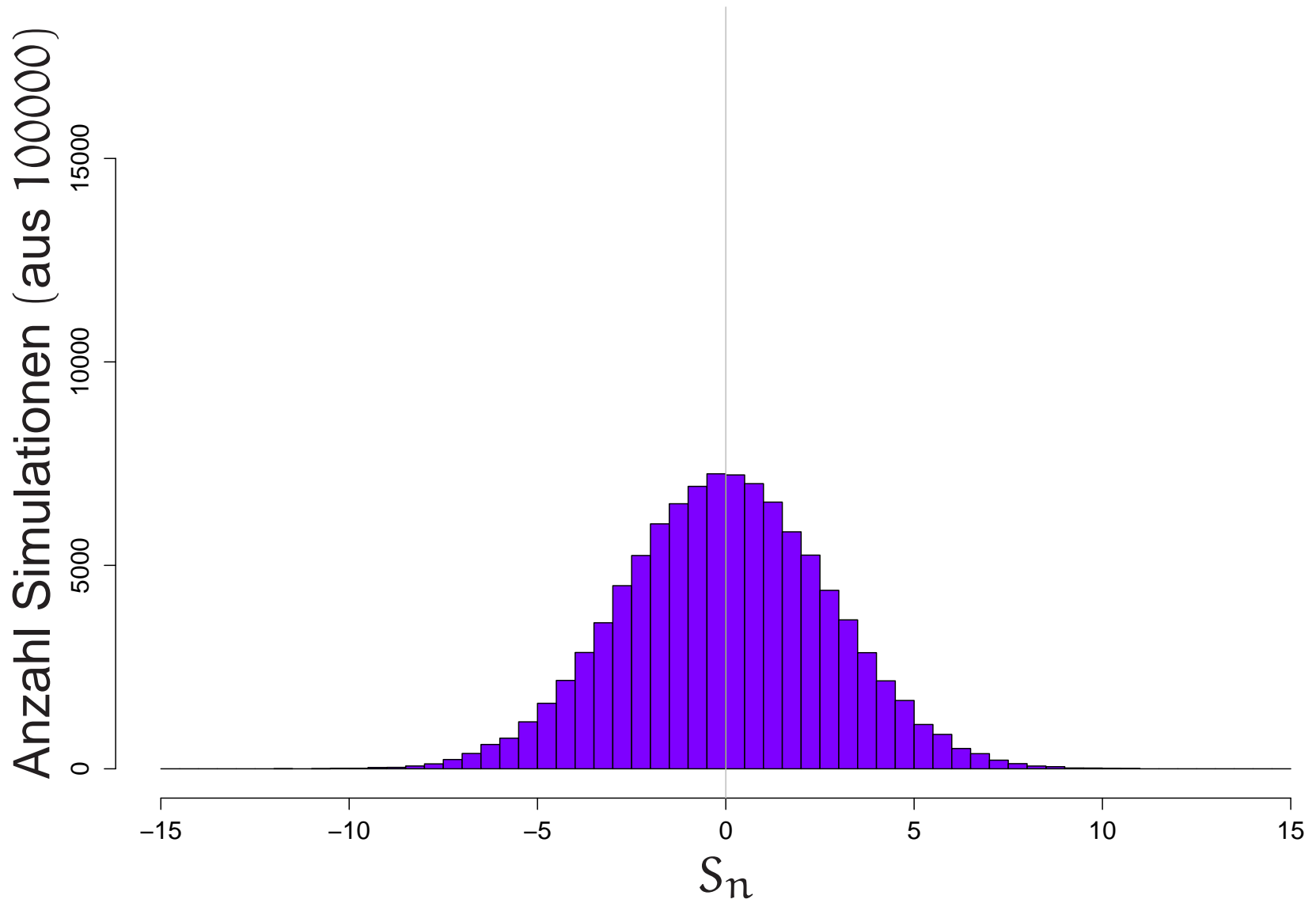
Verteilung von S_n ($n = 80$)



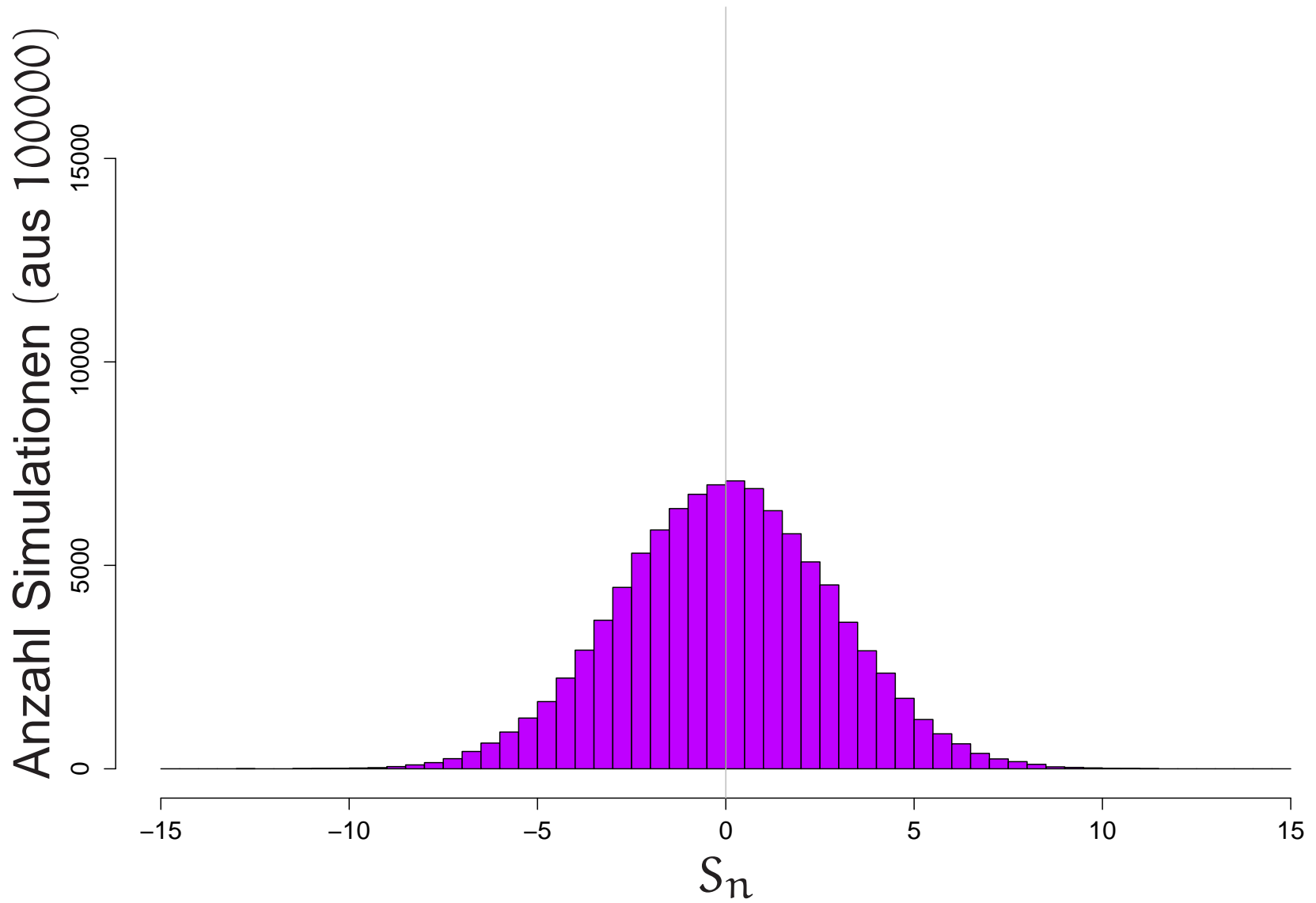
Verteilung von S_n ($n = 85$)



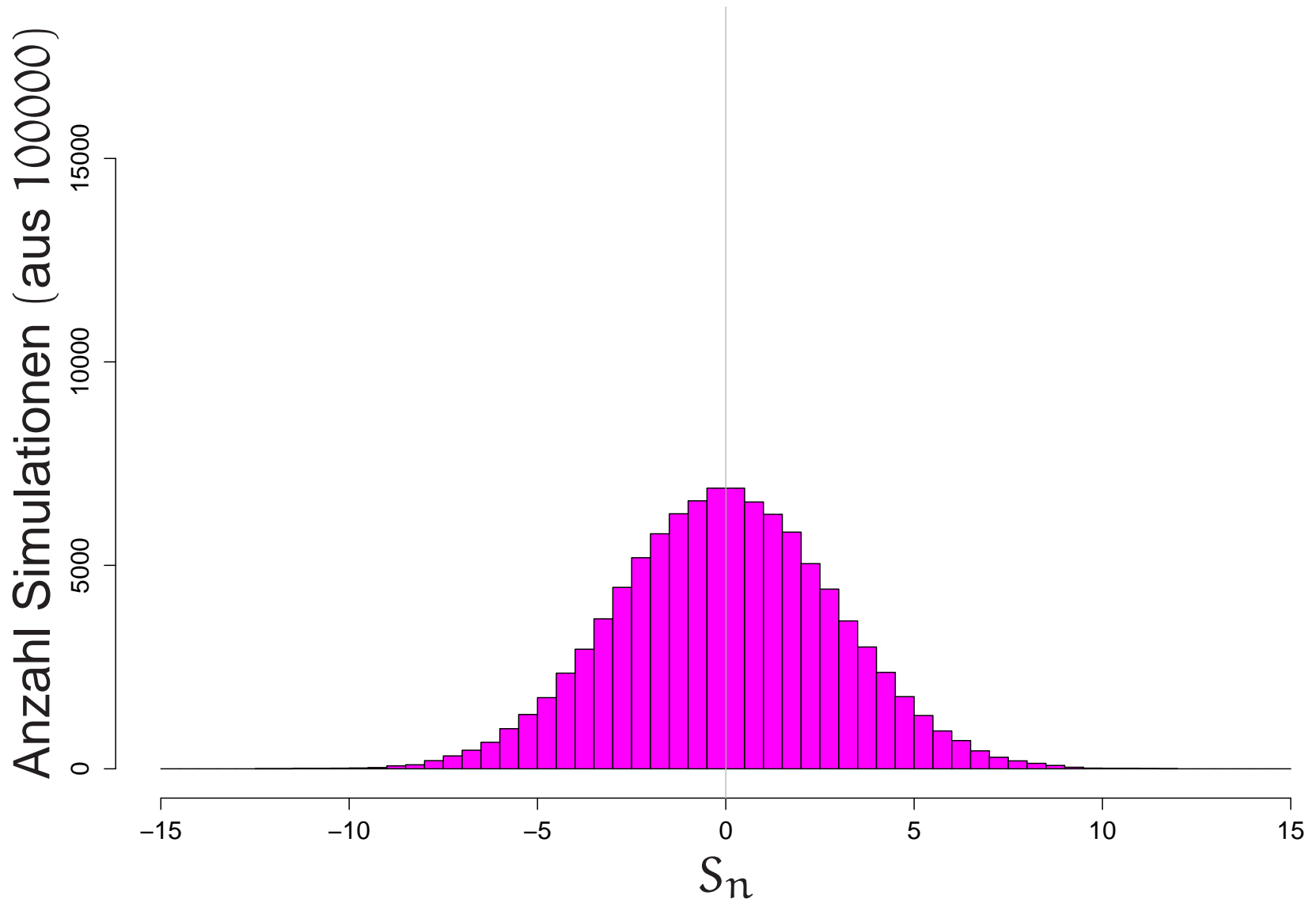
Verteilung von S_n ($n = 90$)



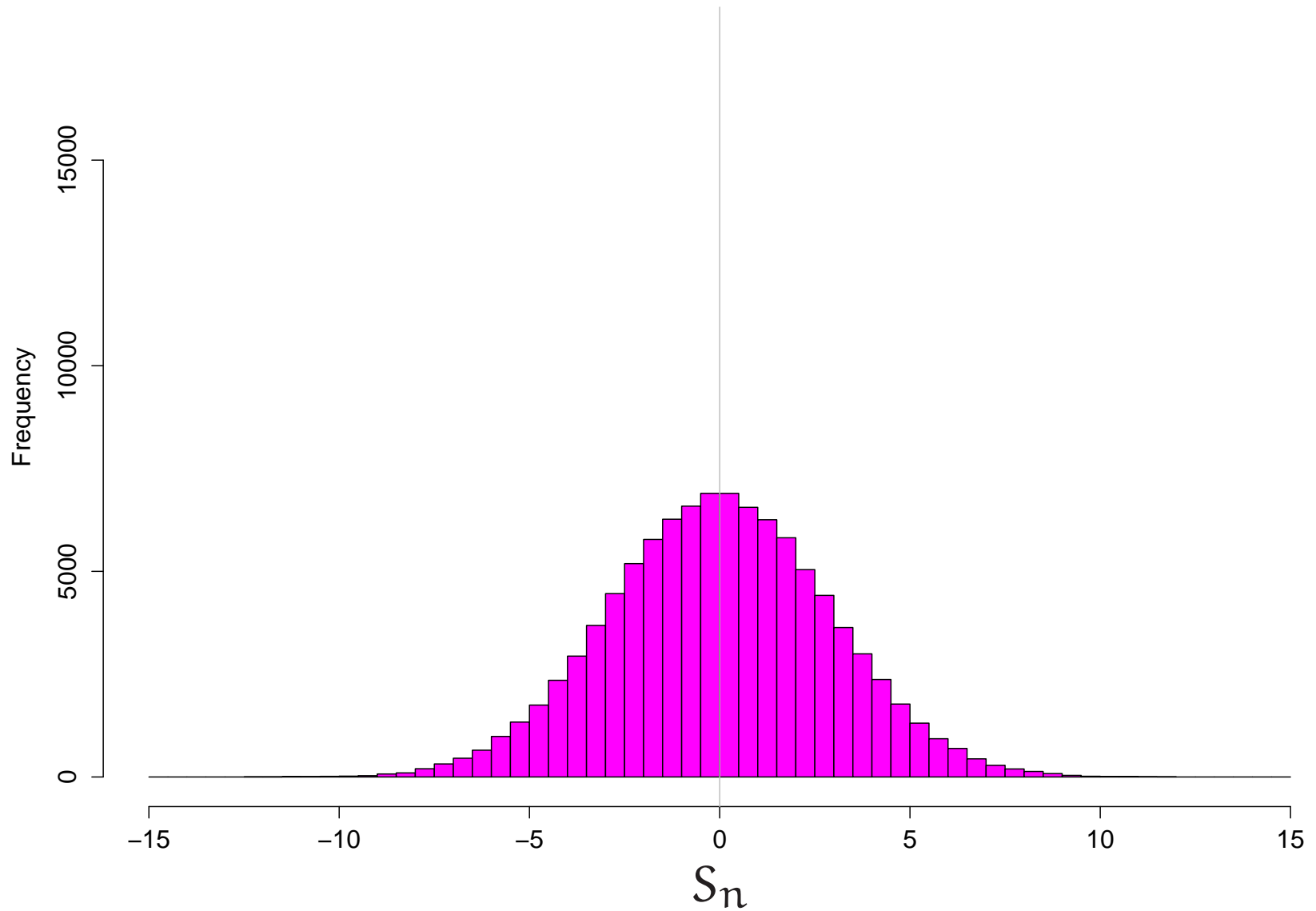
Verteilung von S_n ($n = 95$)



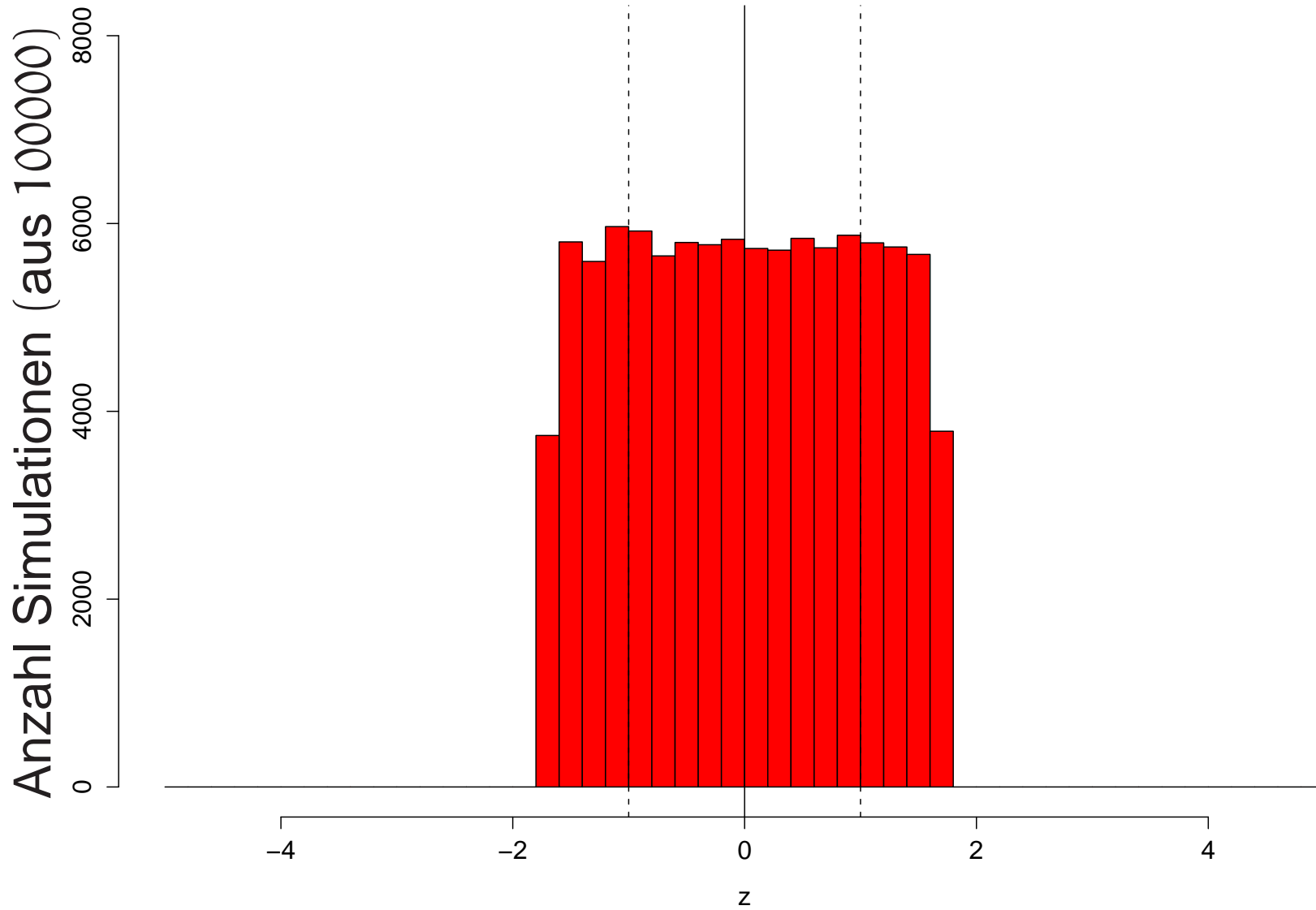
Verteilung von S_n ($n = 100$)



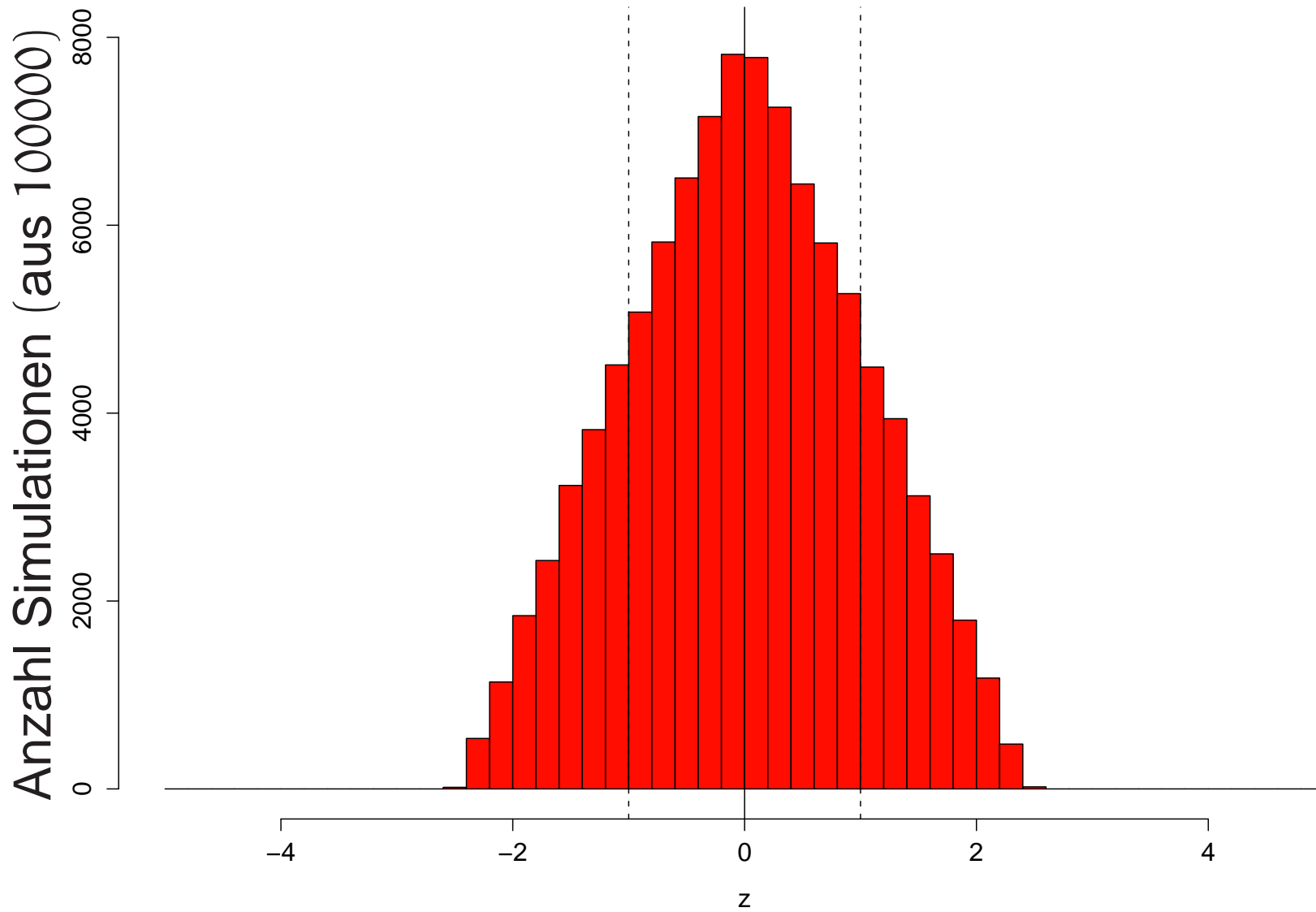
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n}$



Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 1)$

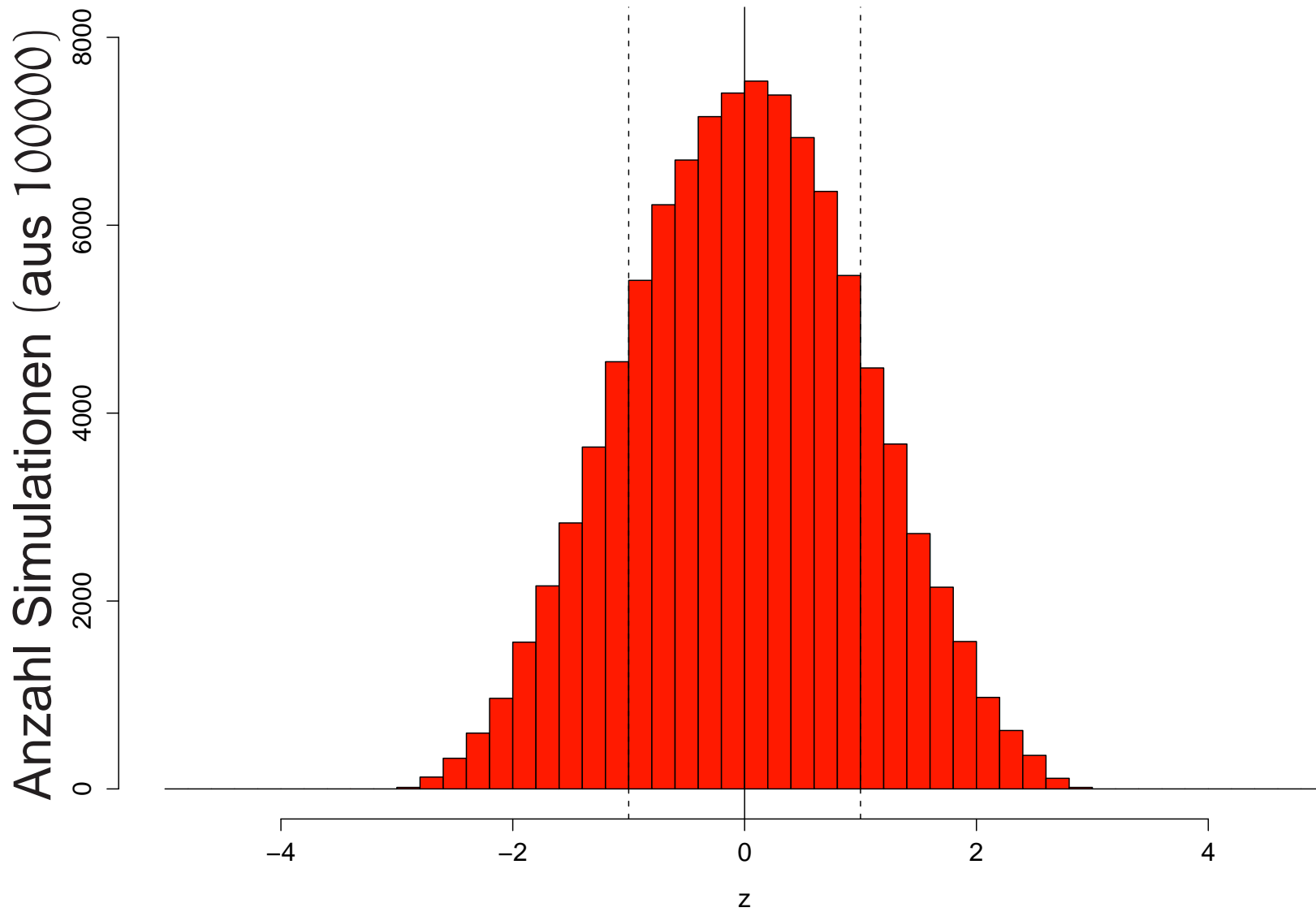


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 2)$

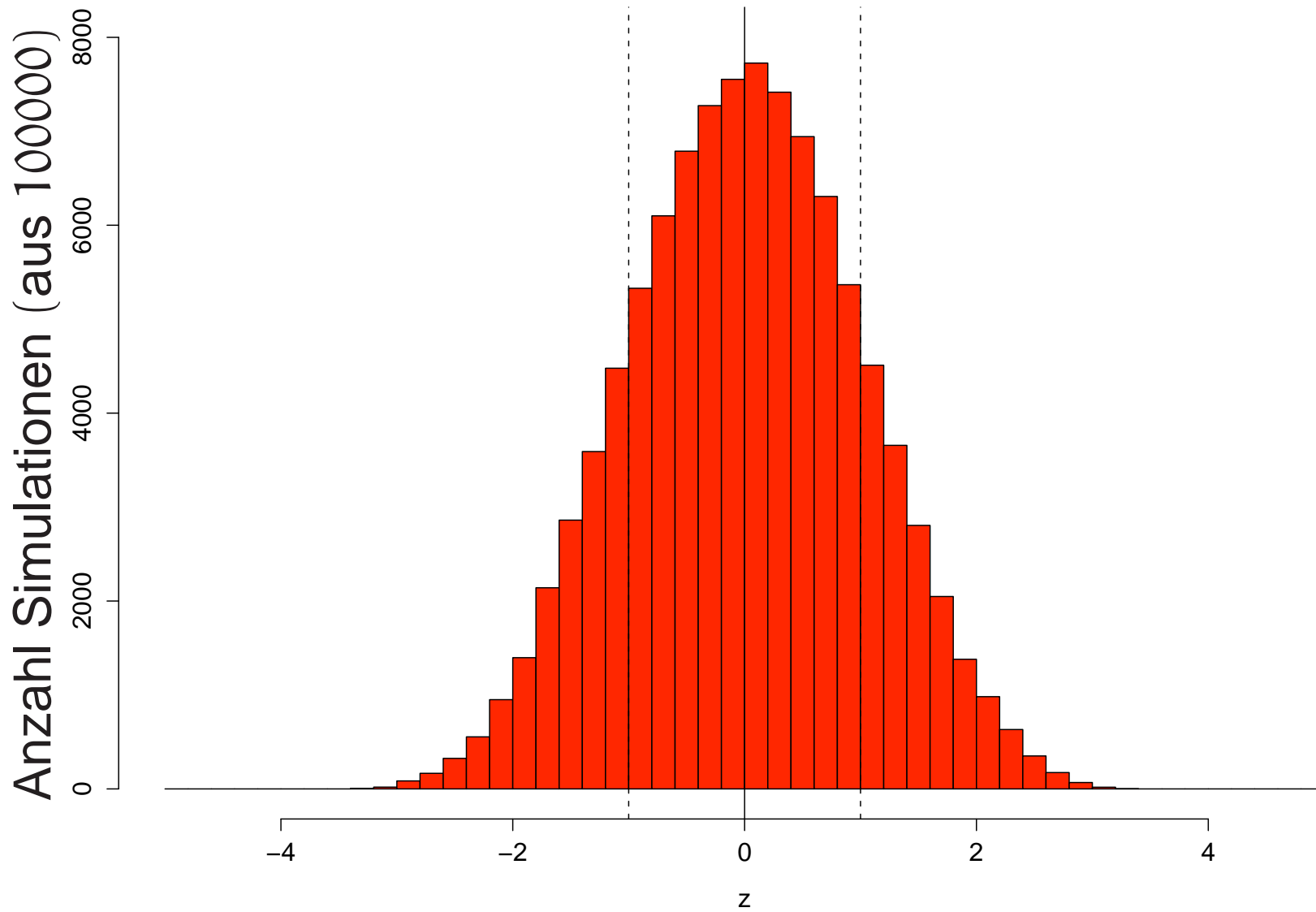


Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 3)$$

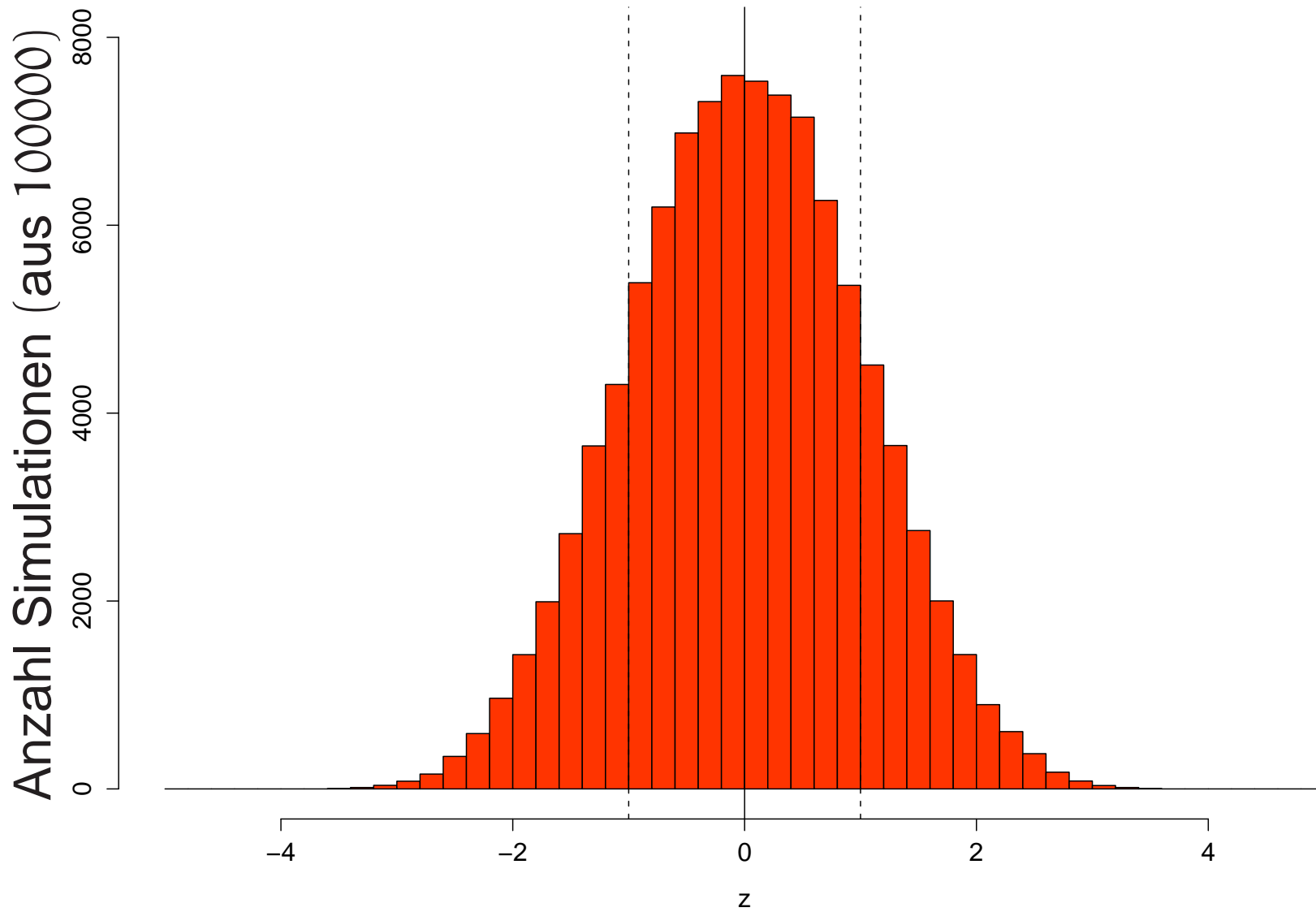


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 4)$



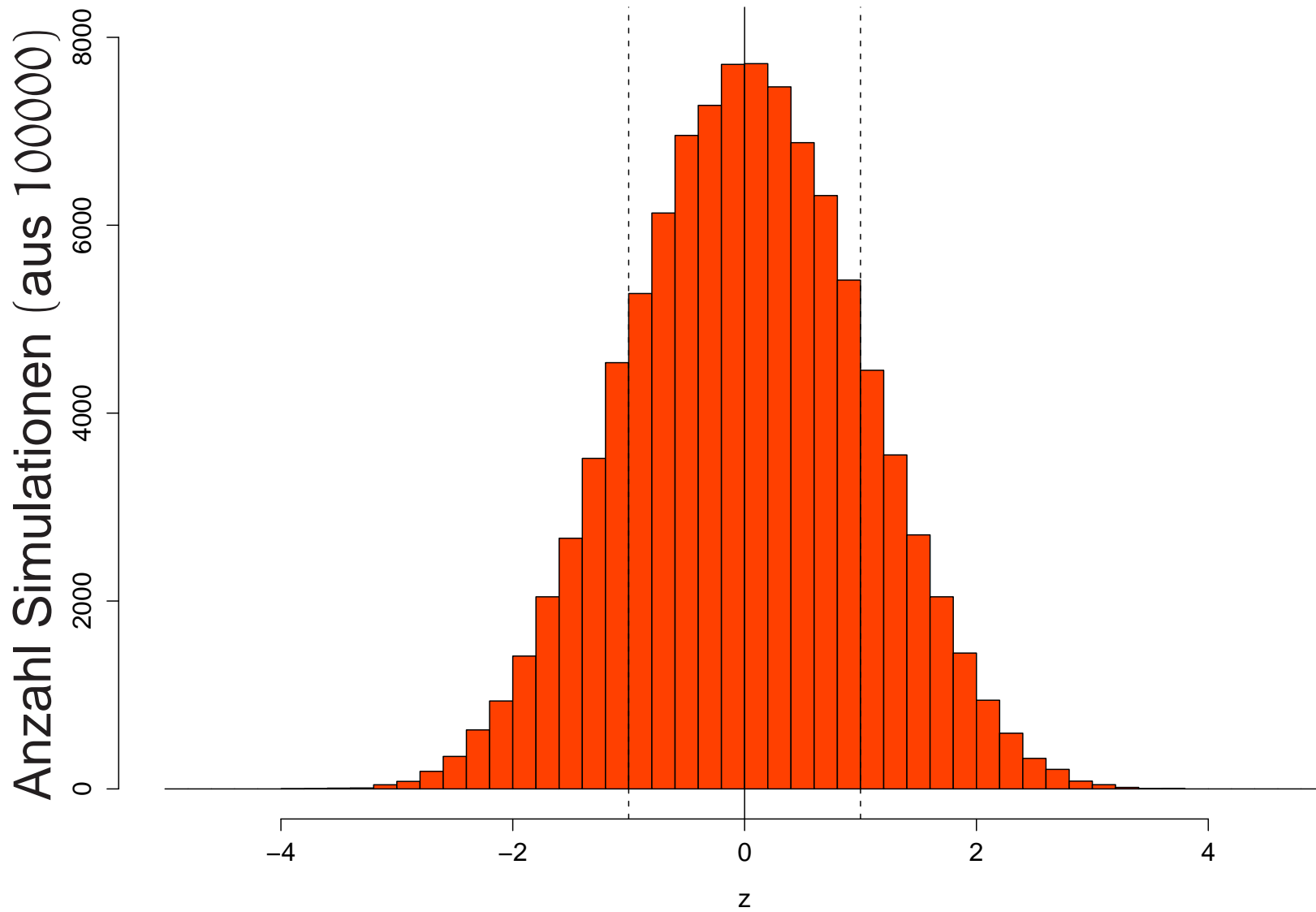
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 5)$$



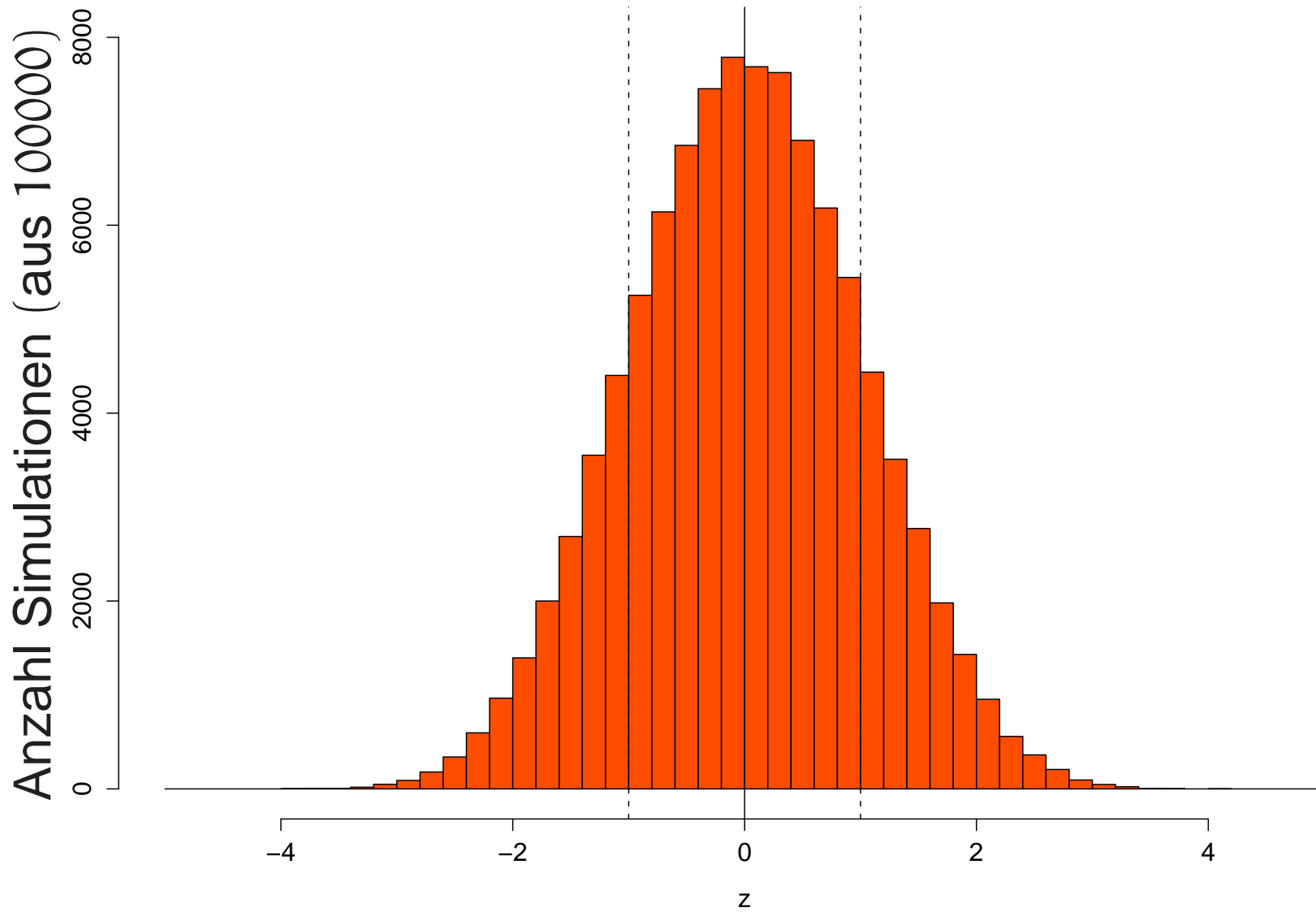
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 6)$$



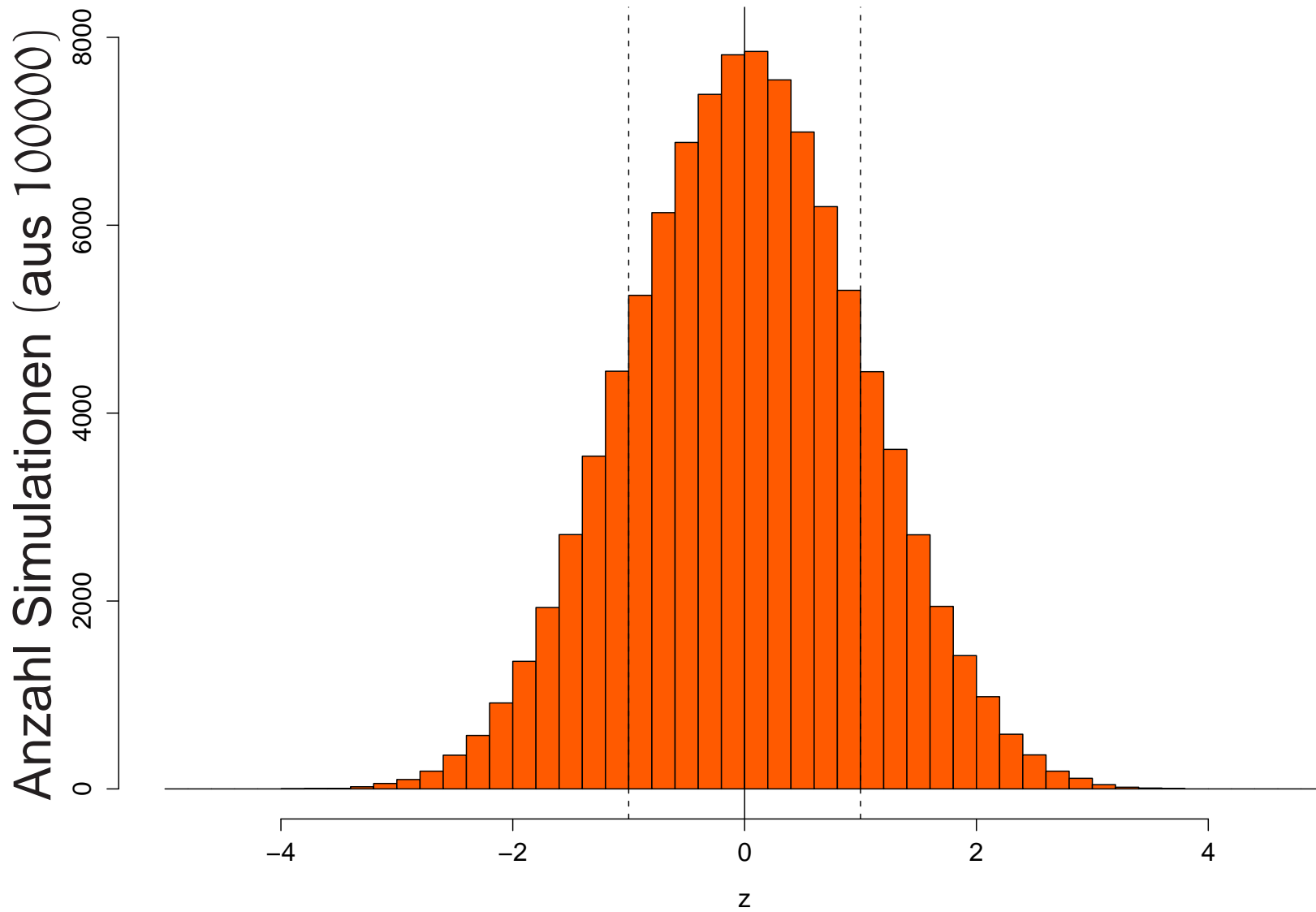
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 7)$$



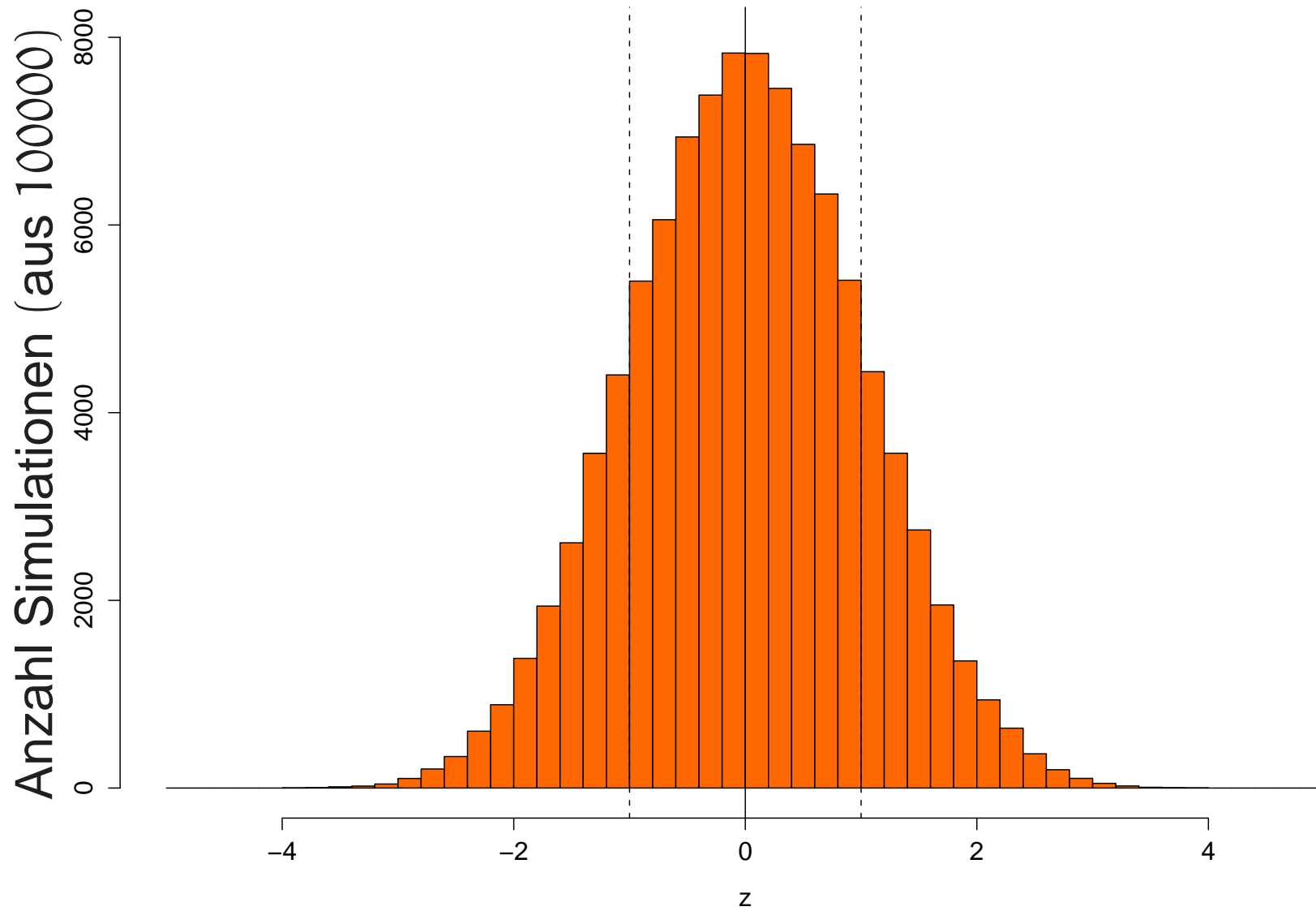
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 8)$$



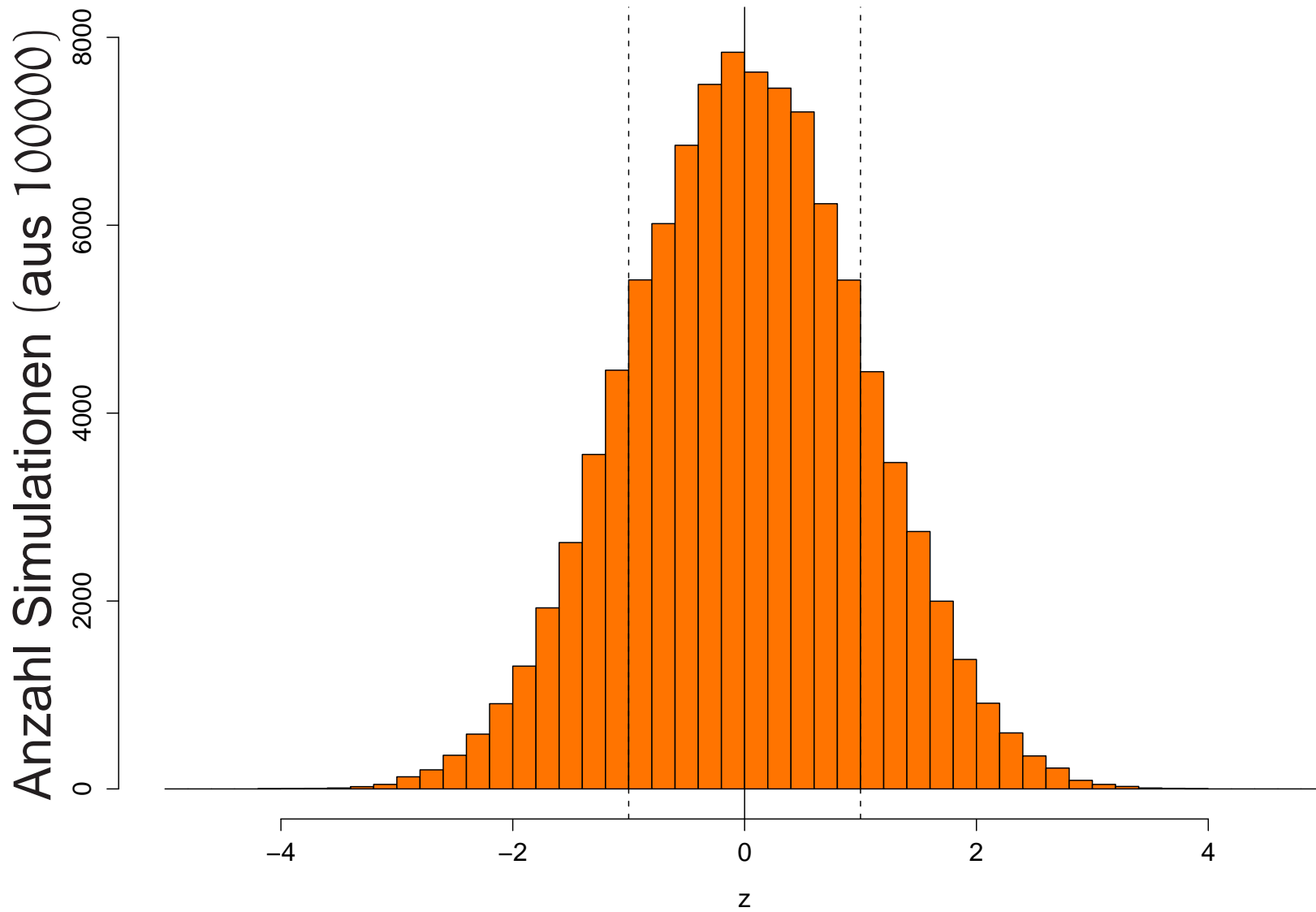
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 9)$$



Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 10)$$



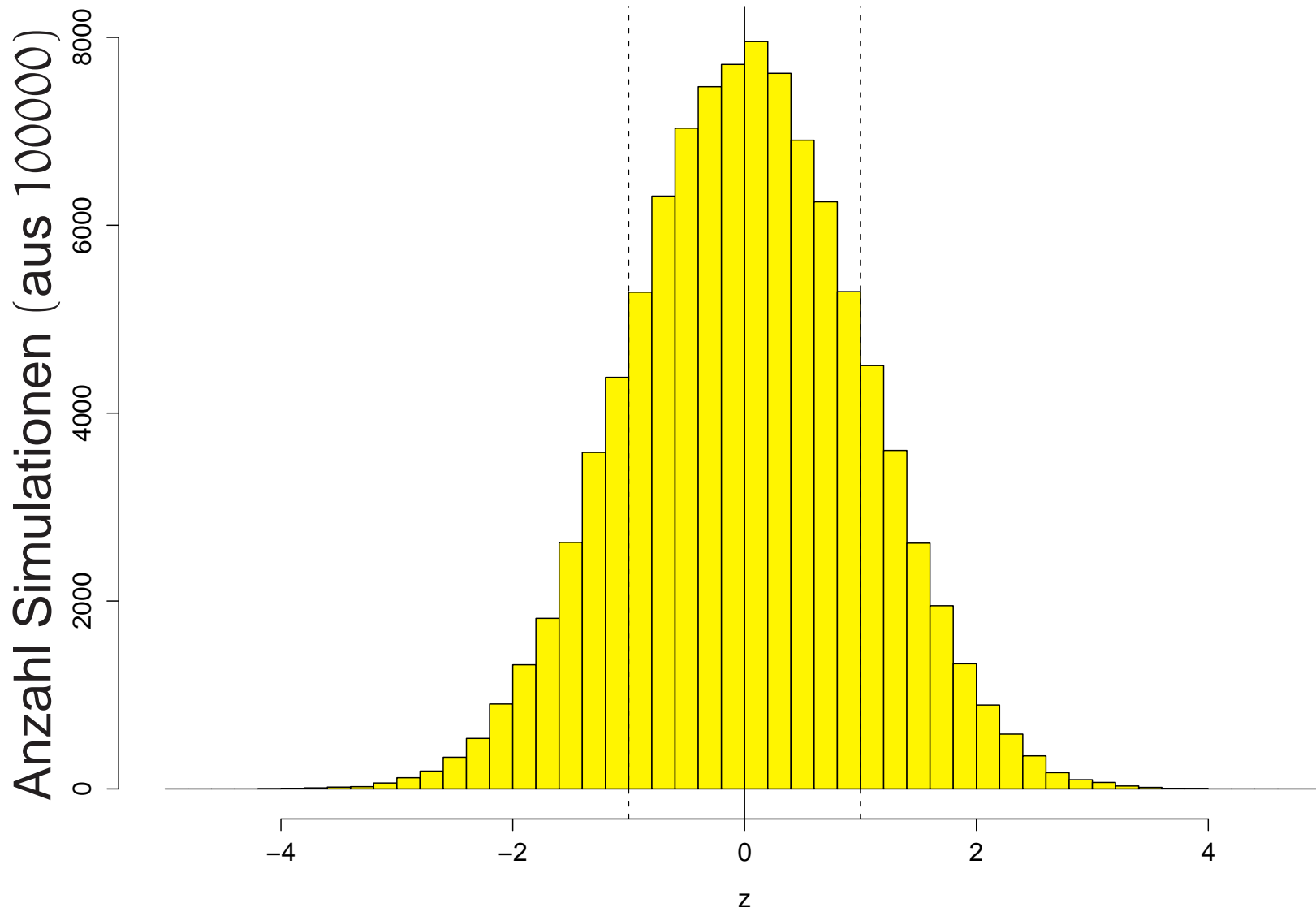
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 15)$$



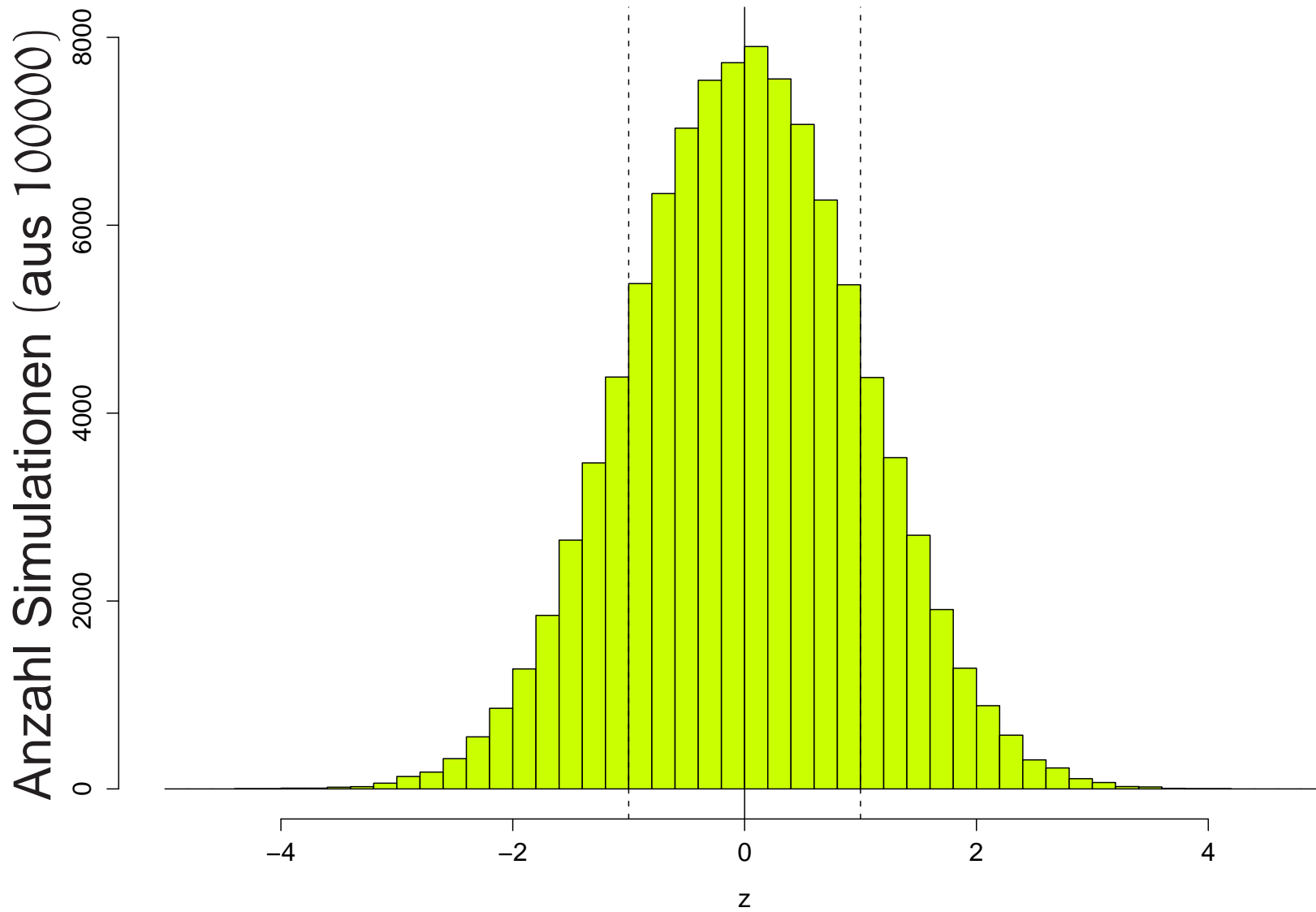
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 20)$$



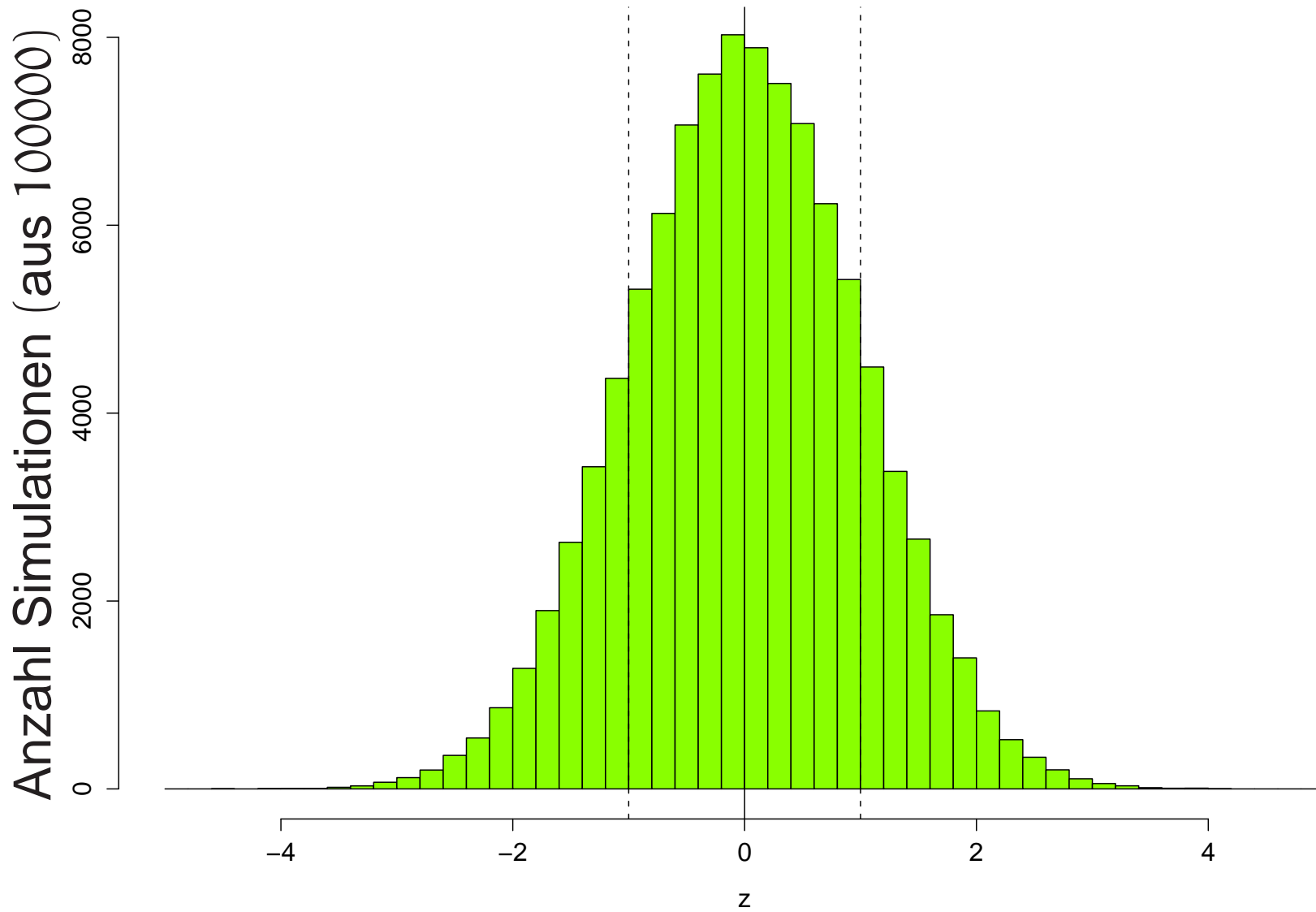
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 25)$$



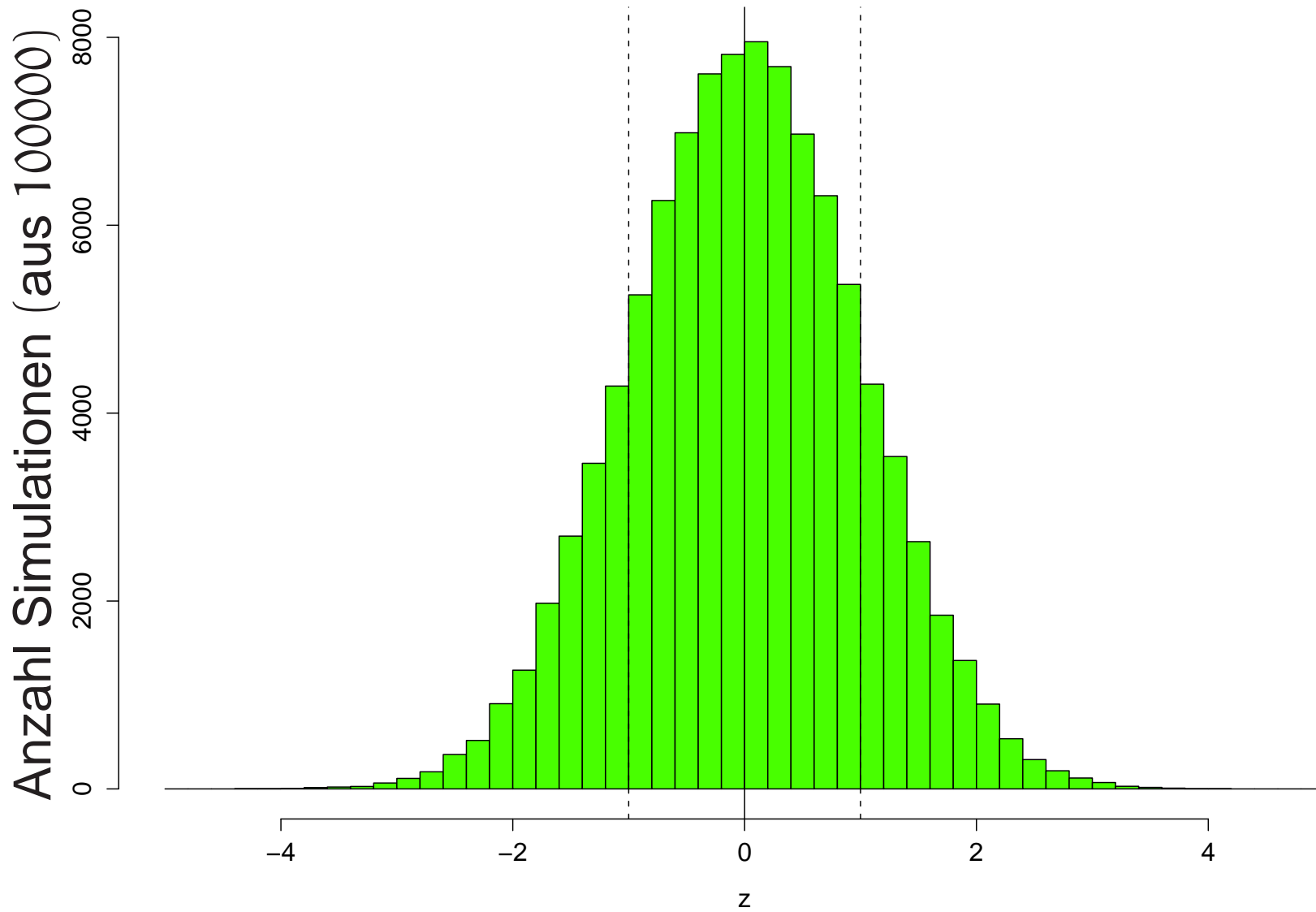
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 30)$$



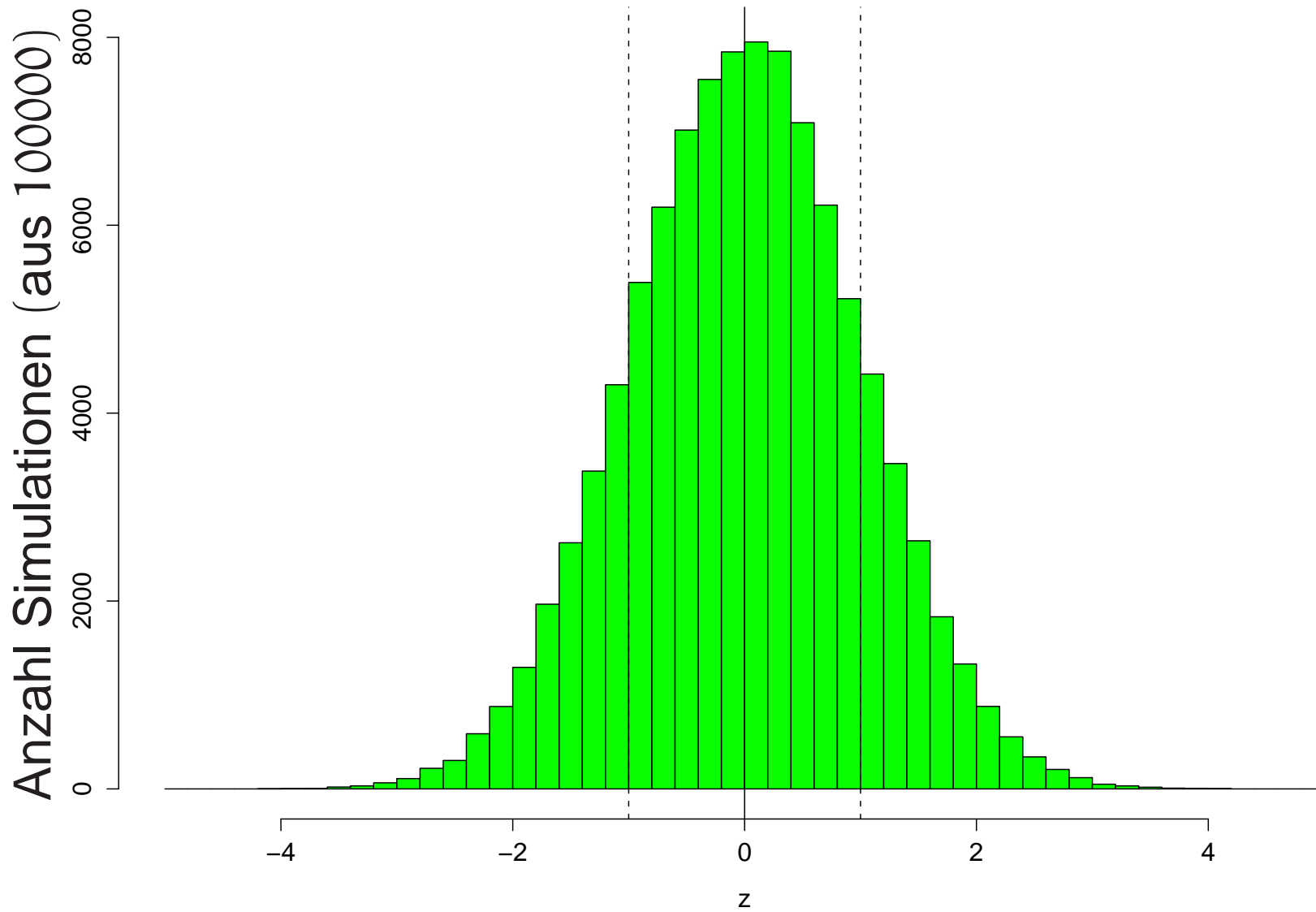
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 35)$$



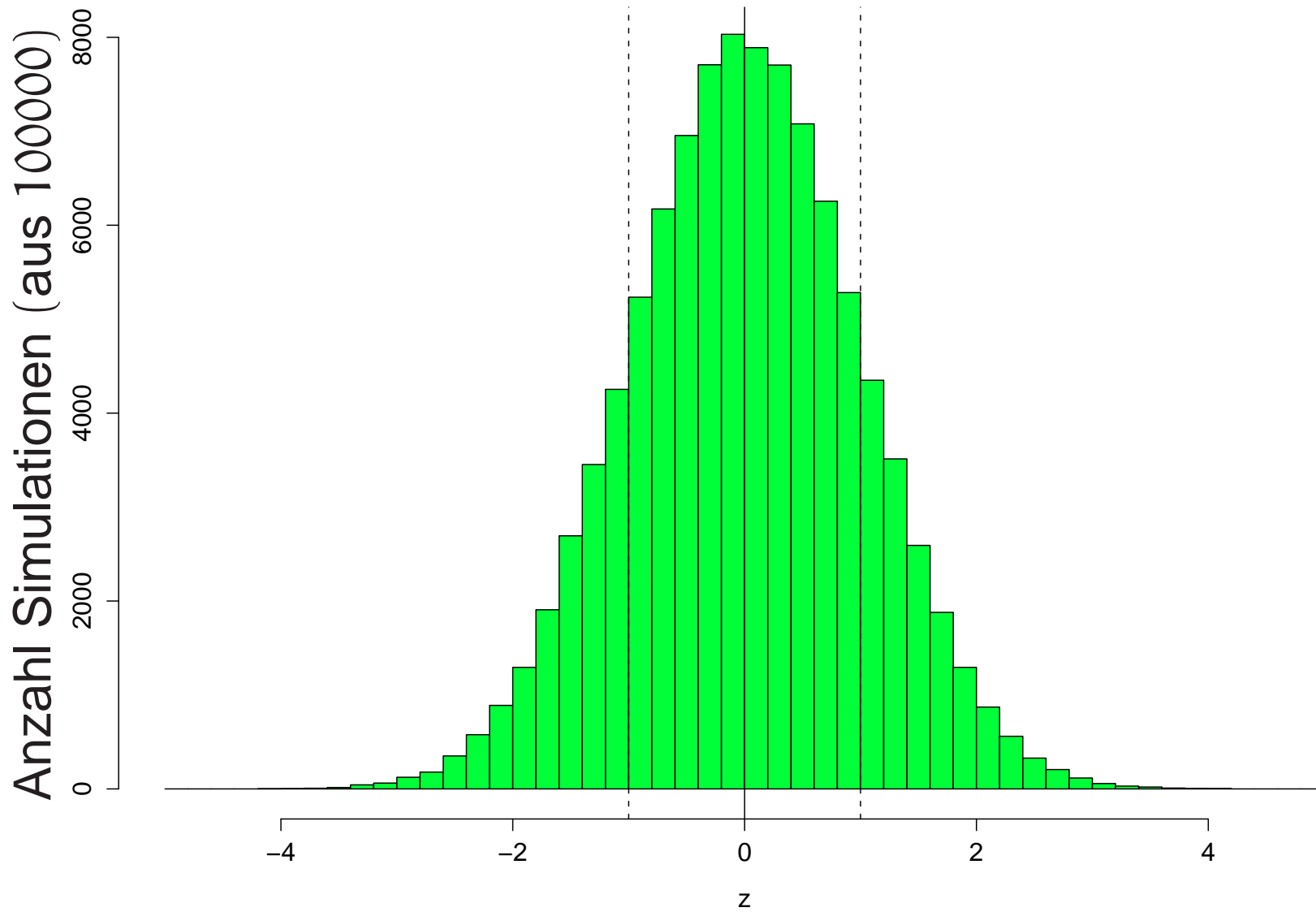
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 40)$$



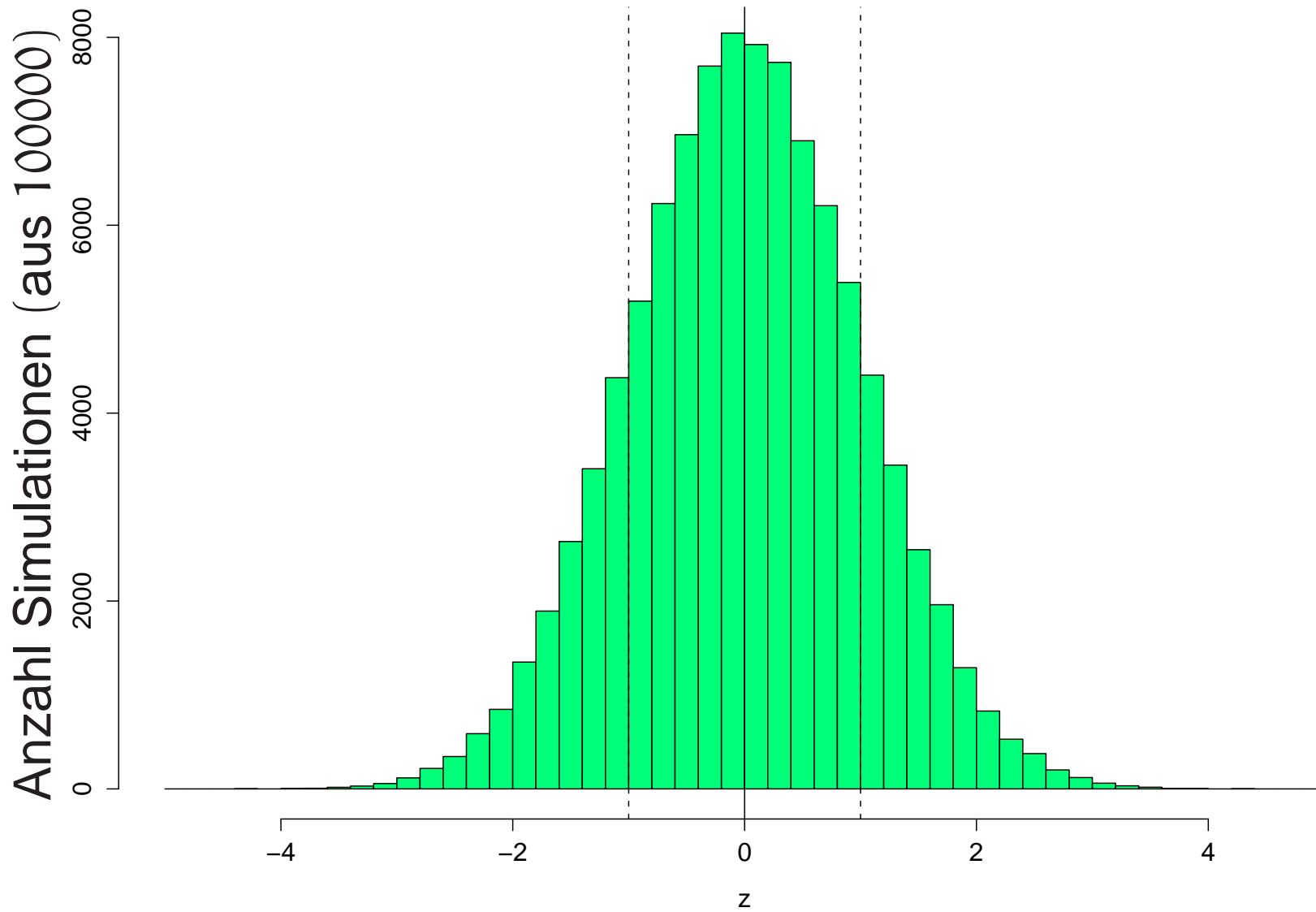
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 45)$$



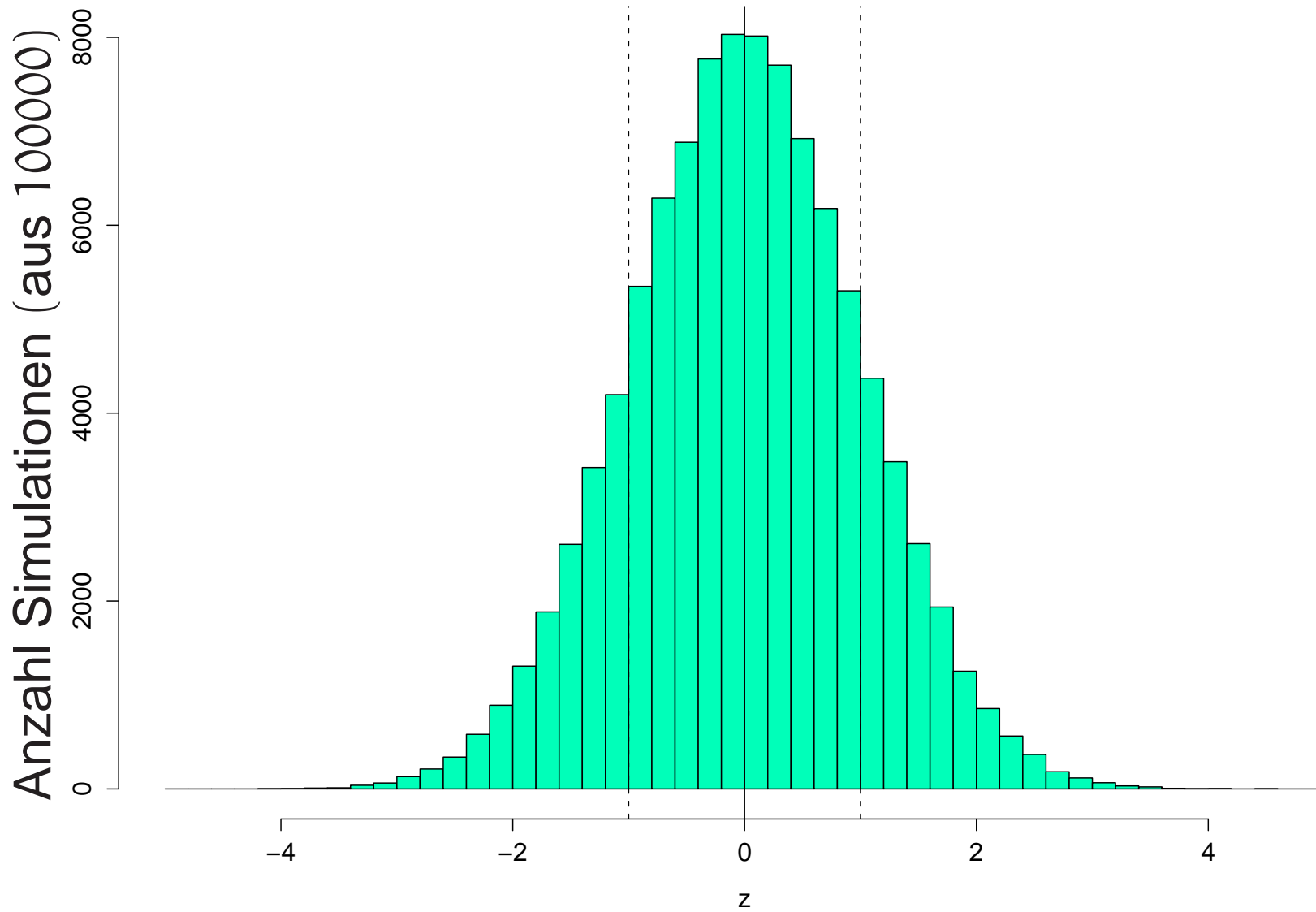
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 50)$$



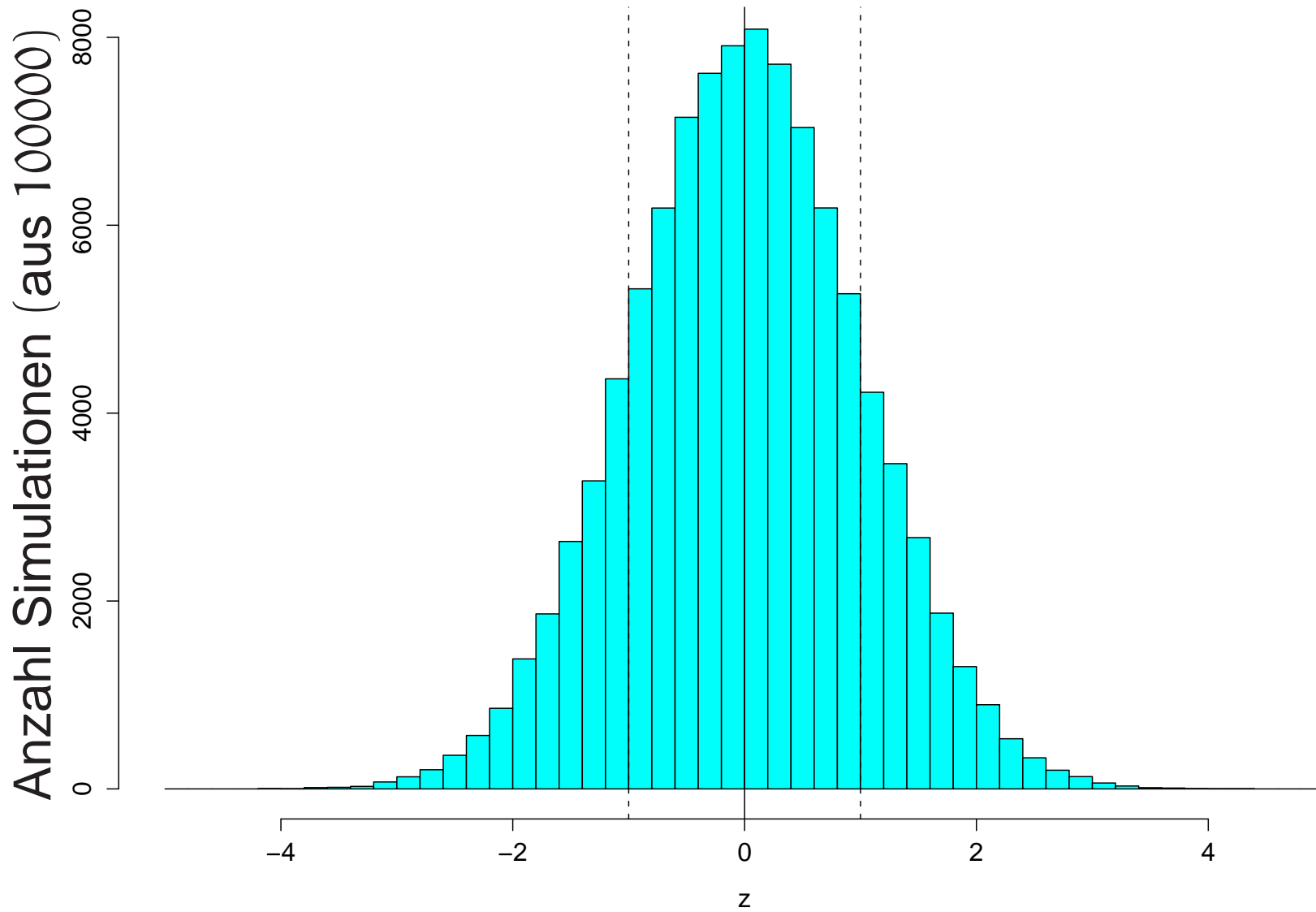
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 55)$$



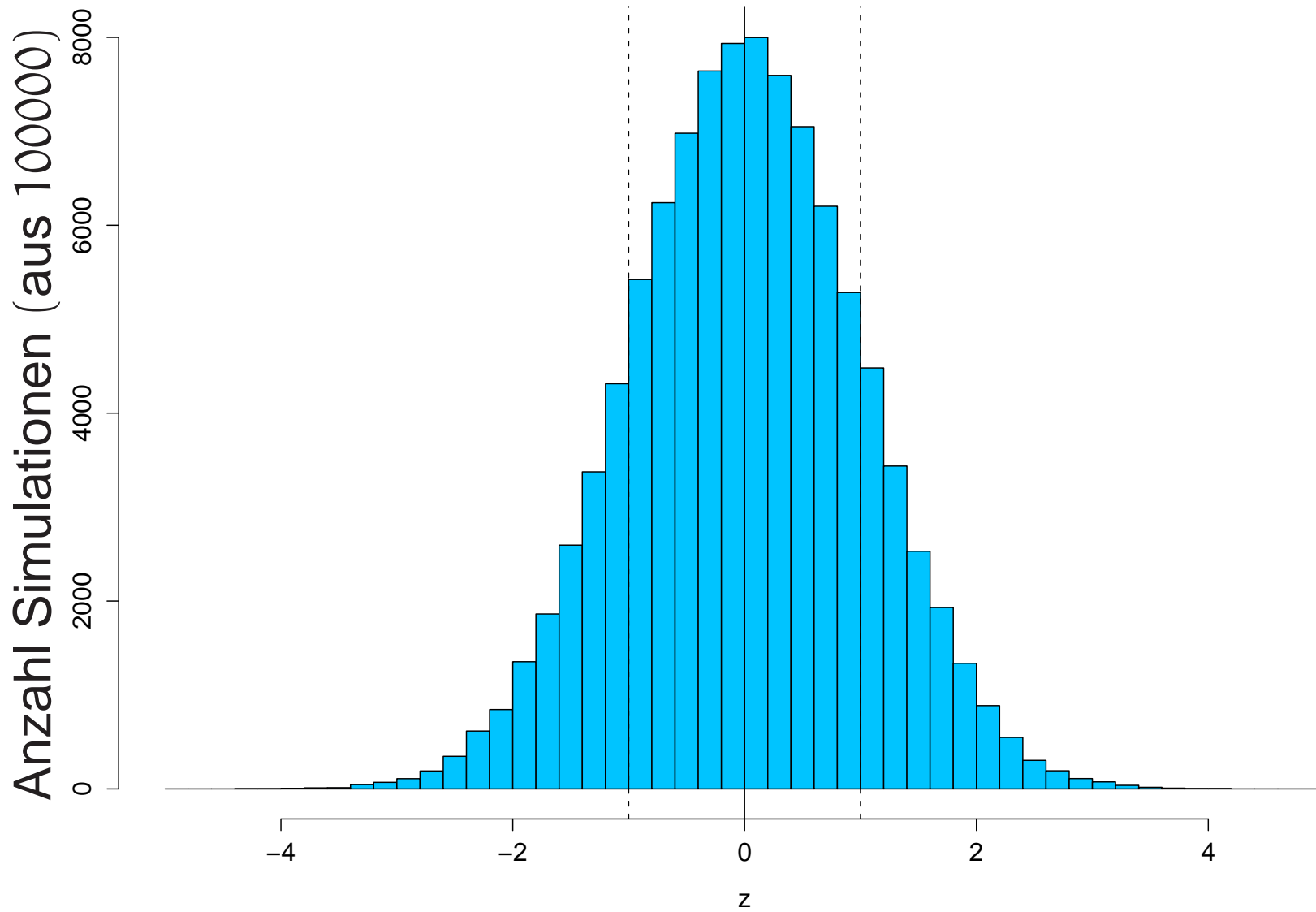
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 60)$$



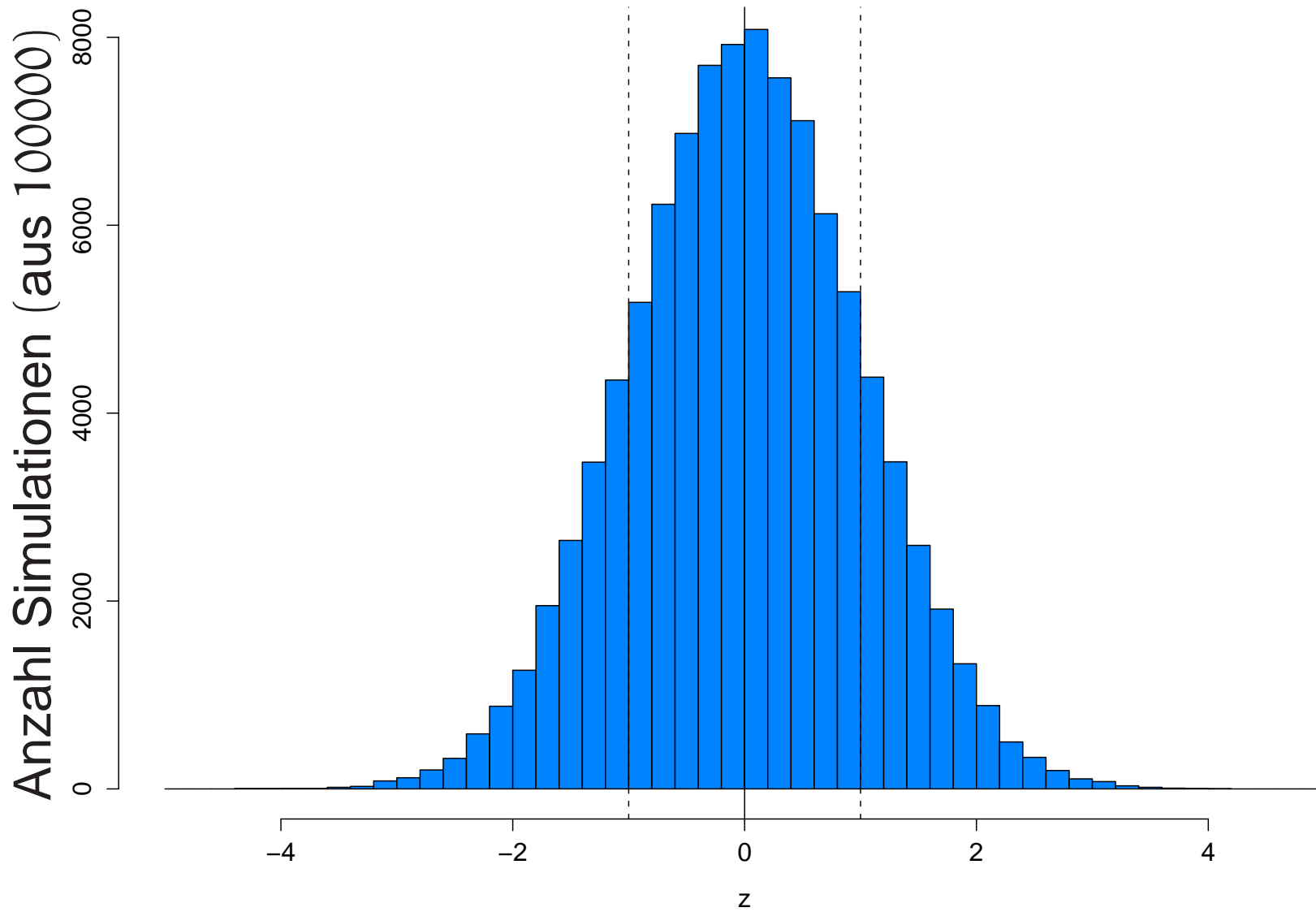
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 65)$$



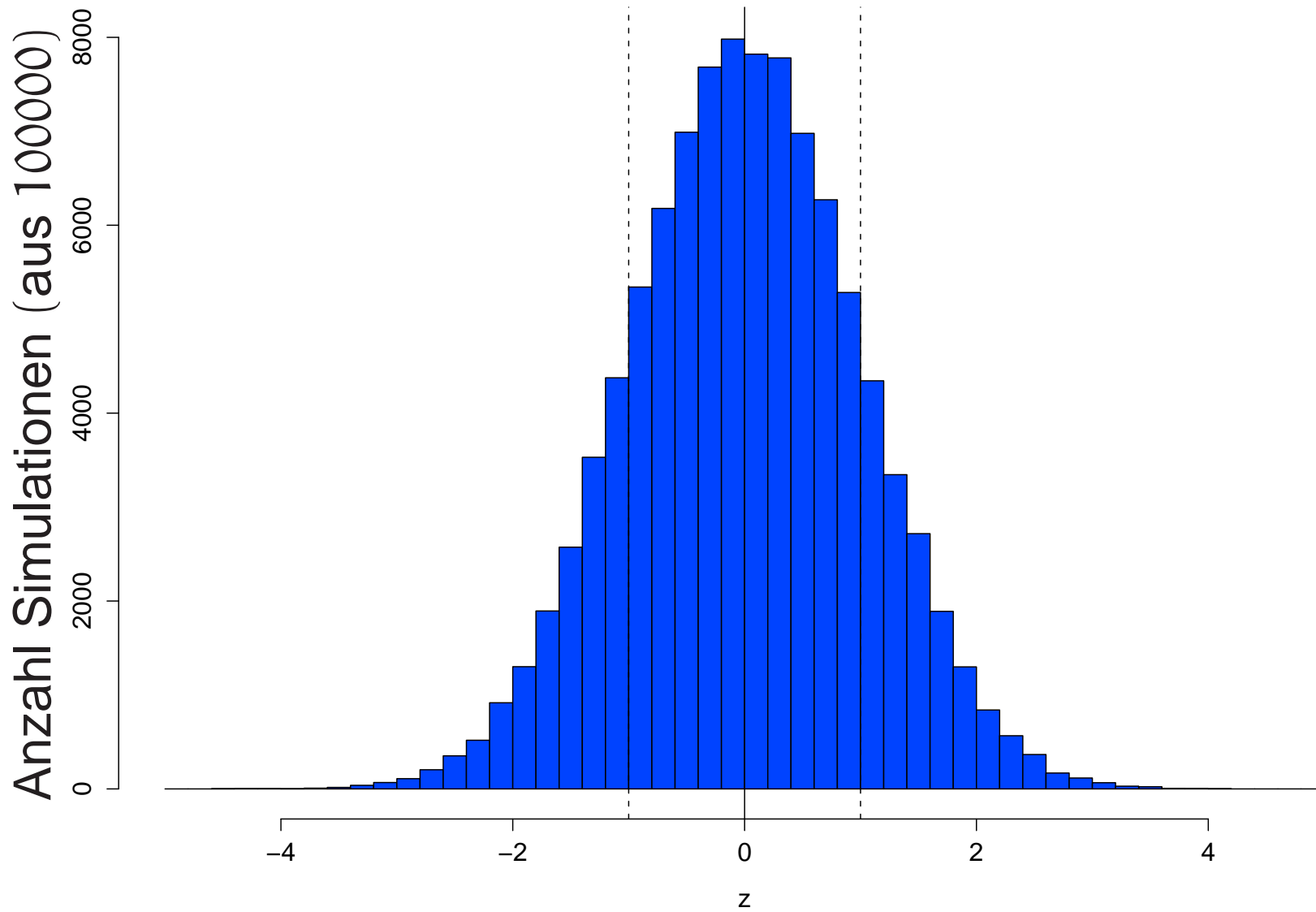
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 70)$$



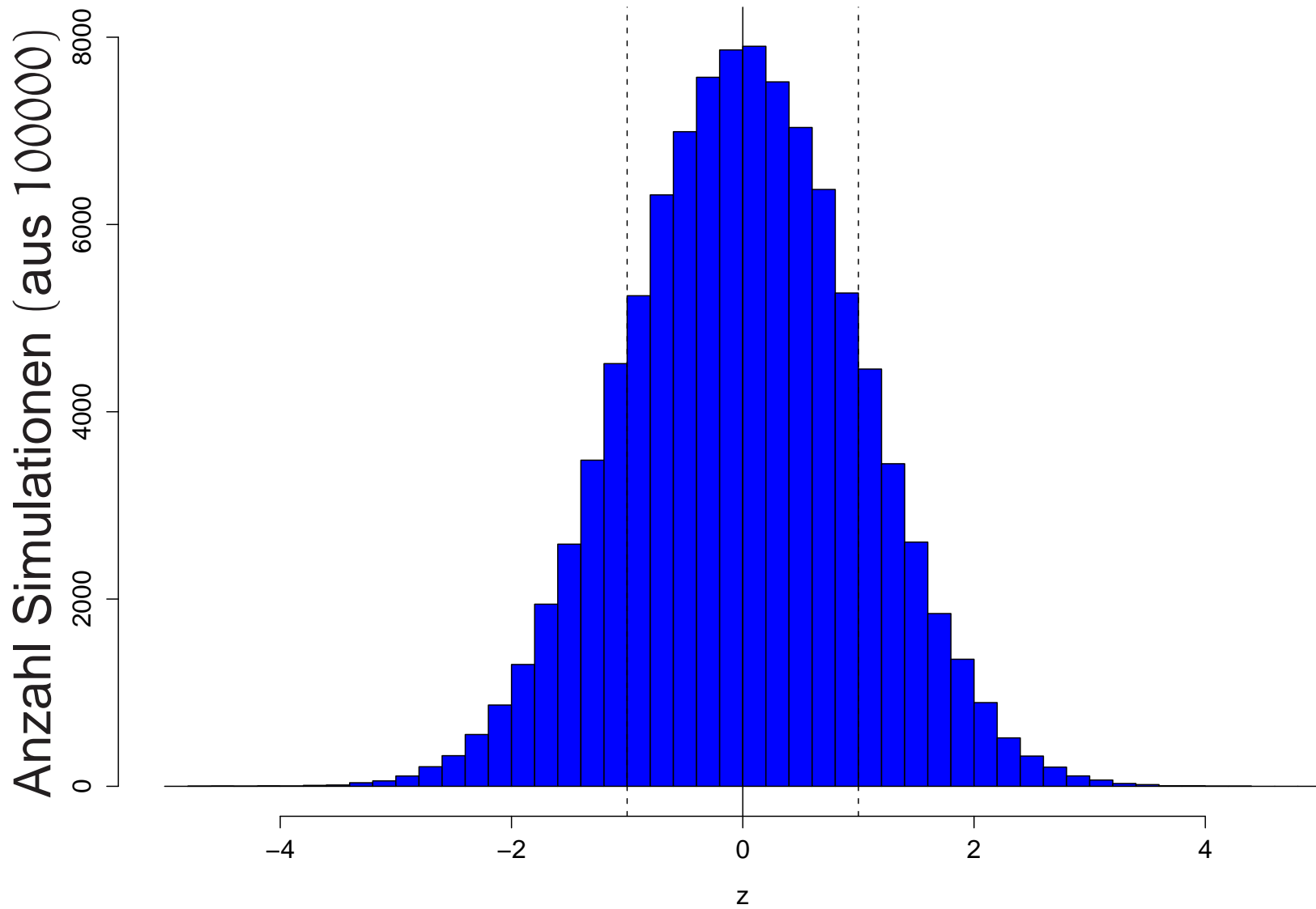
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 75)$$



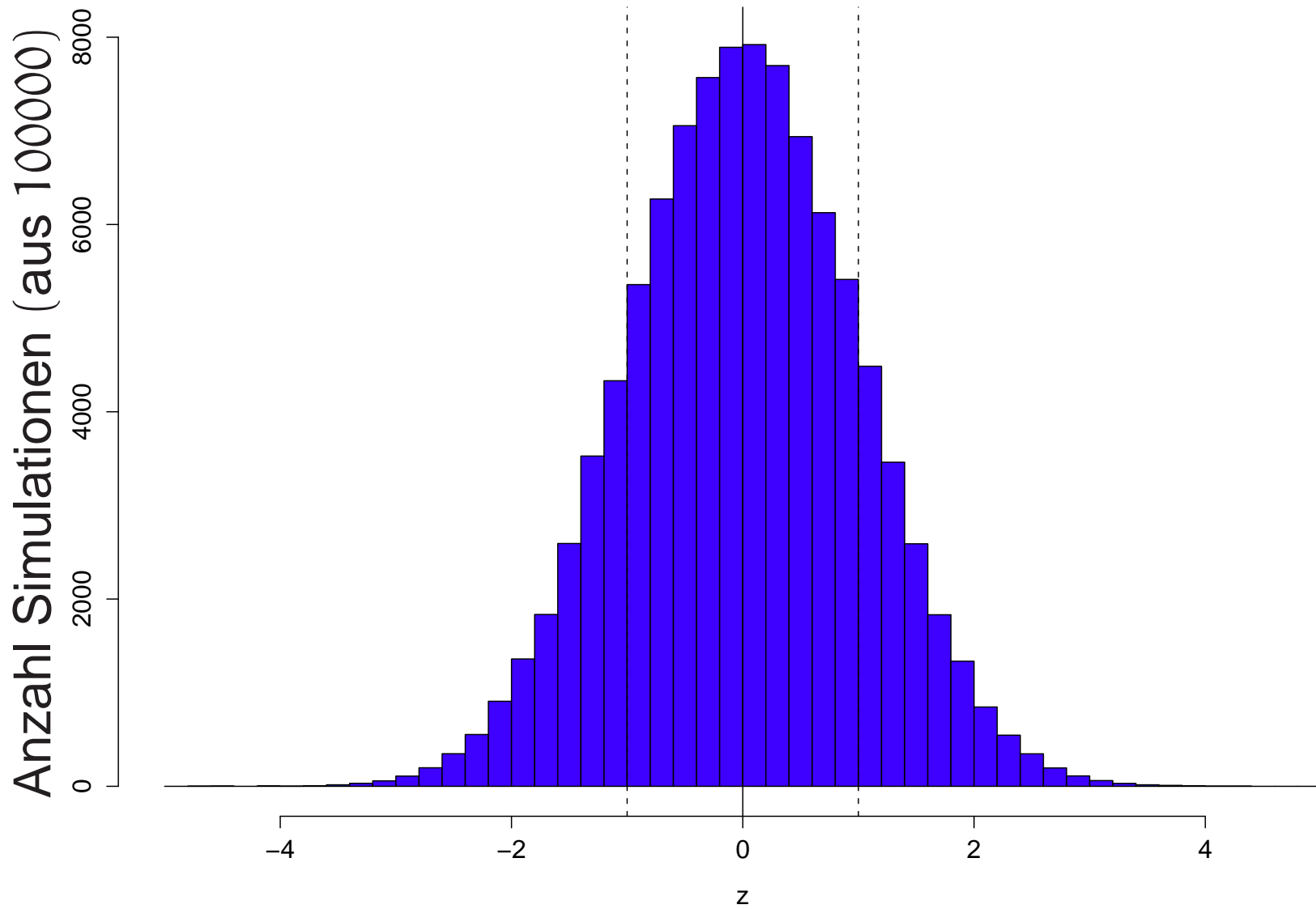
Standardisierung:

$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 80)$$

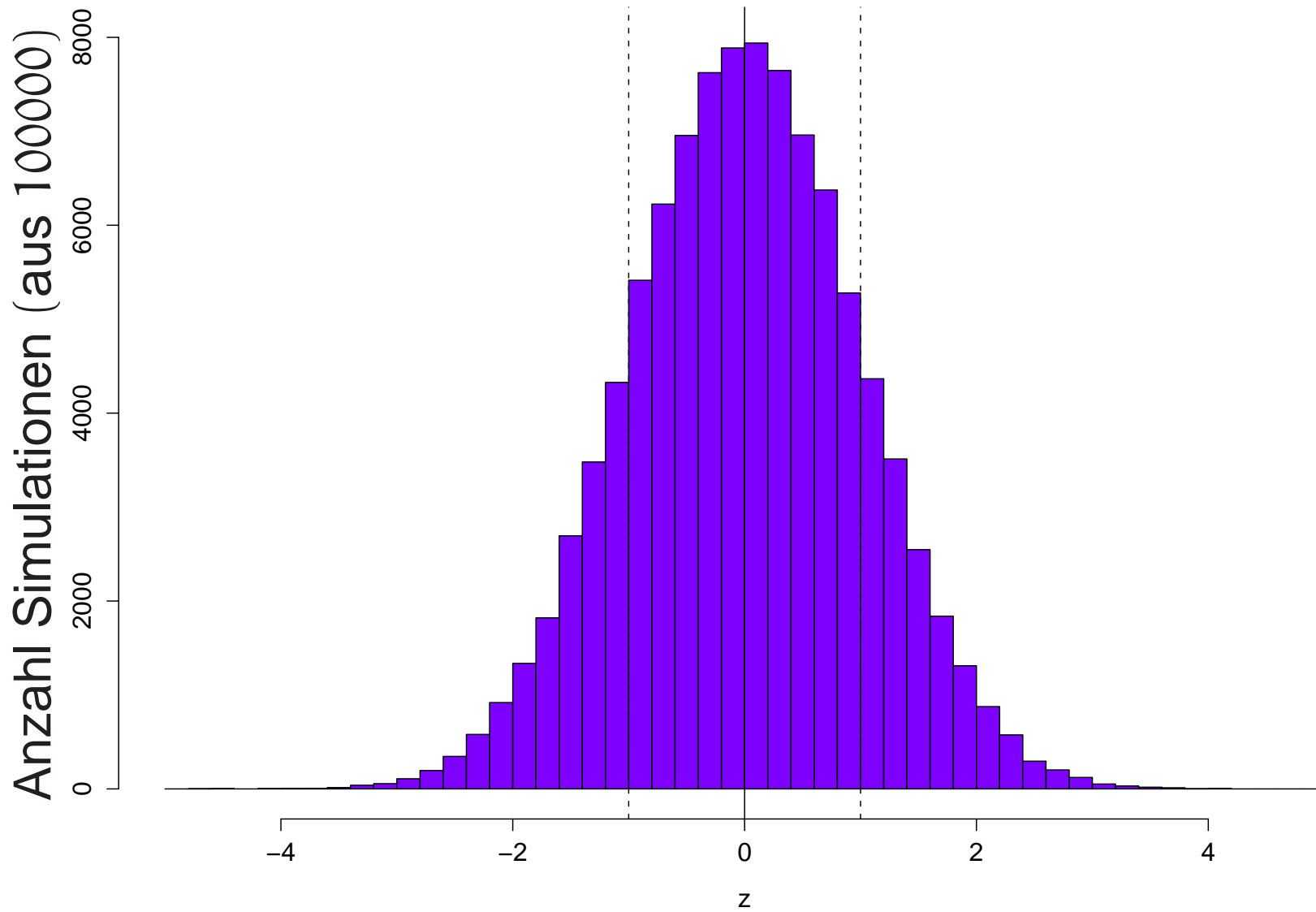


Standardisierung:

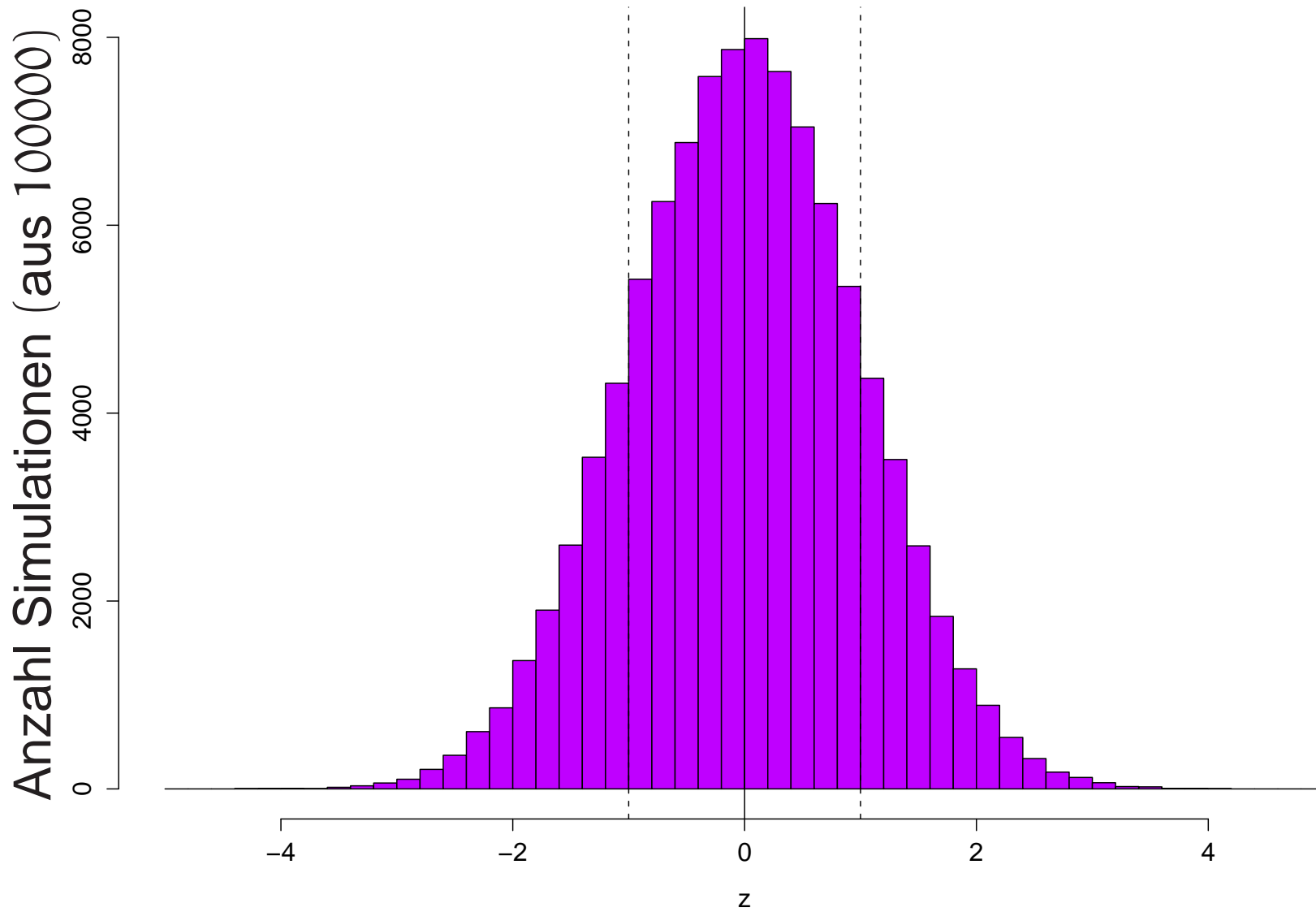
$$Z_n := (S_n - \mathbf{E}S_n) / \sigma_{S_n} \quad (n = 85)$$



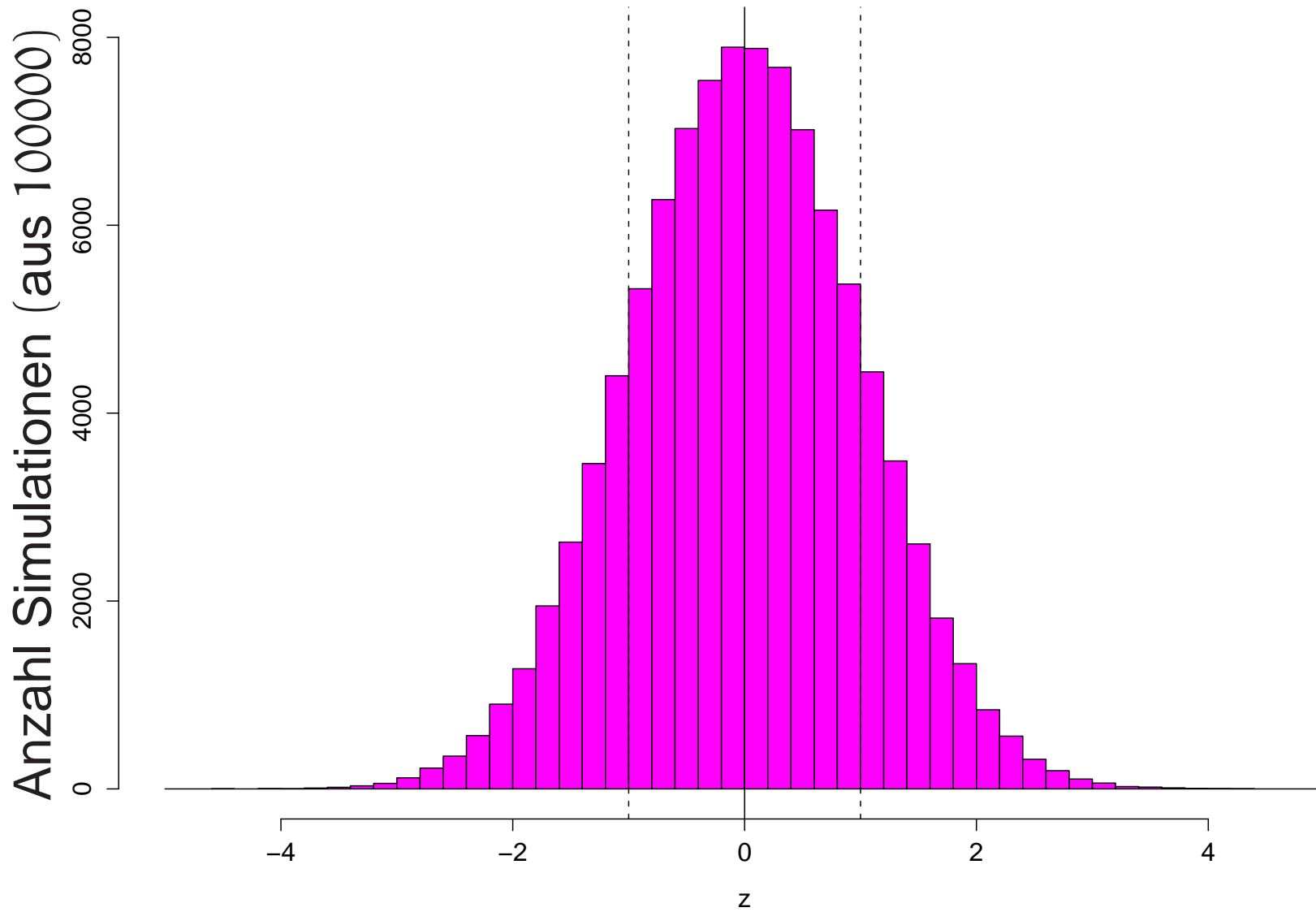
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n}$ ($n = 90$)



Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n}$ ($n = 95$)



Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n}$ ($n = 100$)



Die Verteilung von Z_n
scheint zu konvergieren.

Welche Form
hat die Grenzverteilung?

Die Verteilung von

Z_{100}

ist glockenförmig.

Welche Glocke?

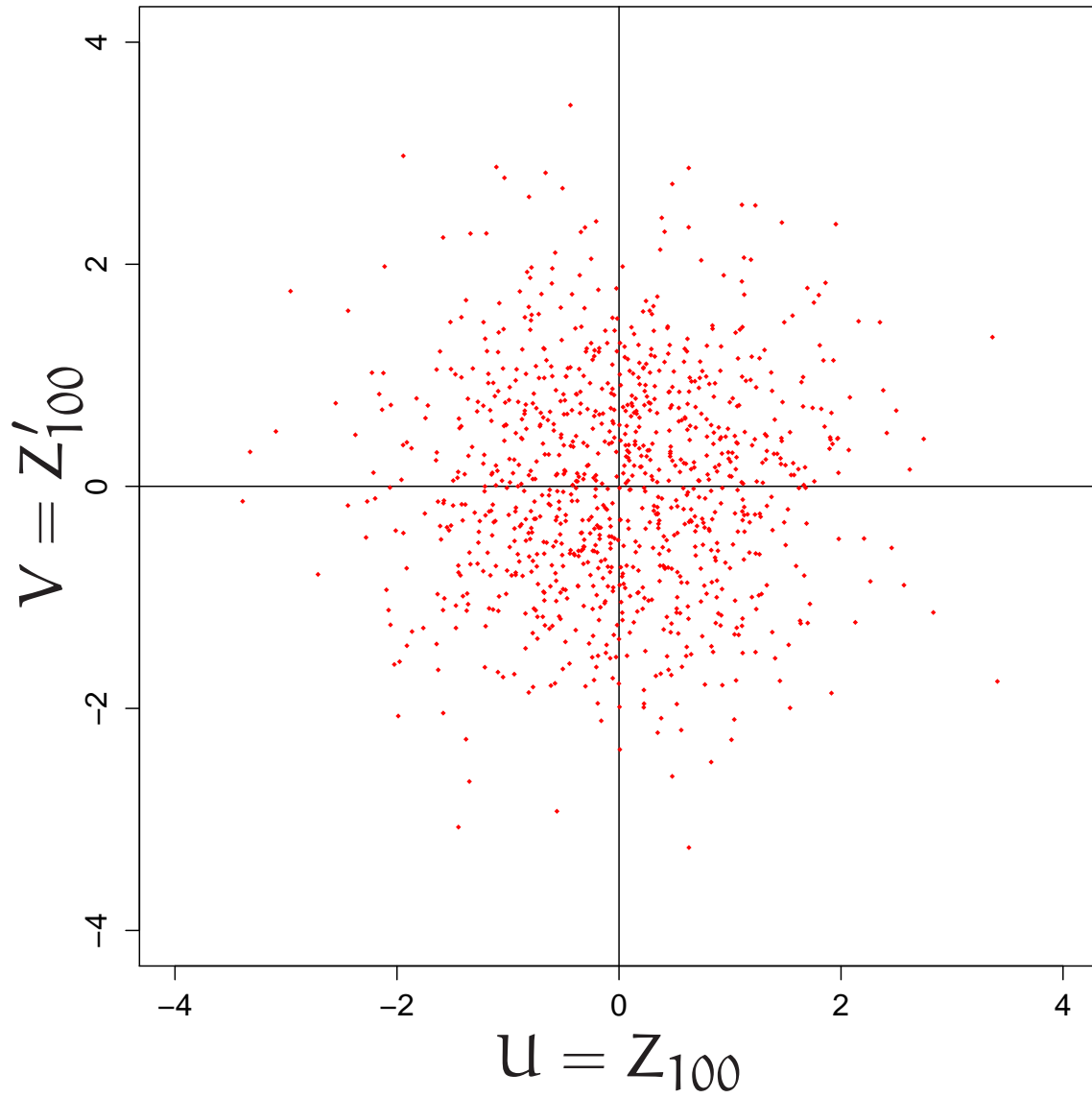
Glücklicher Einfall:

Zwei unabhängige Kopien

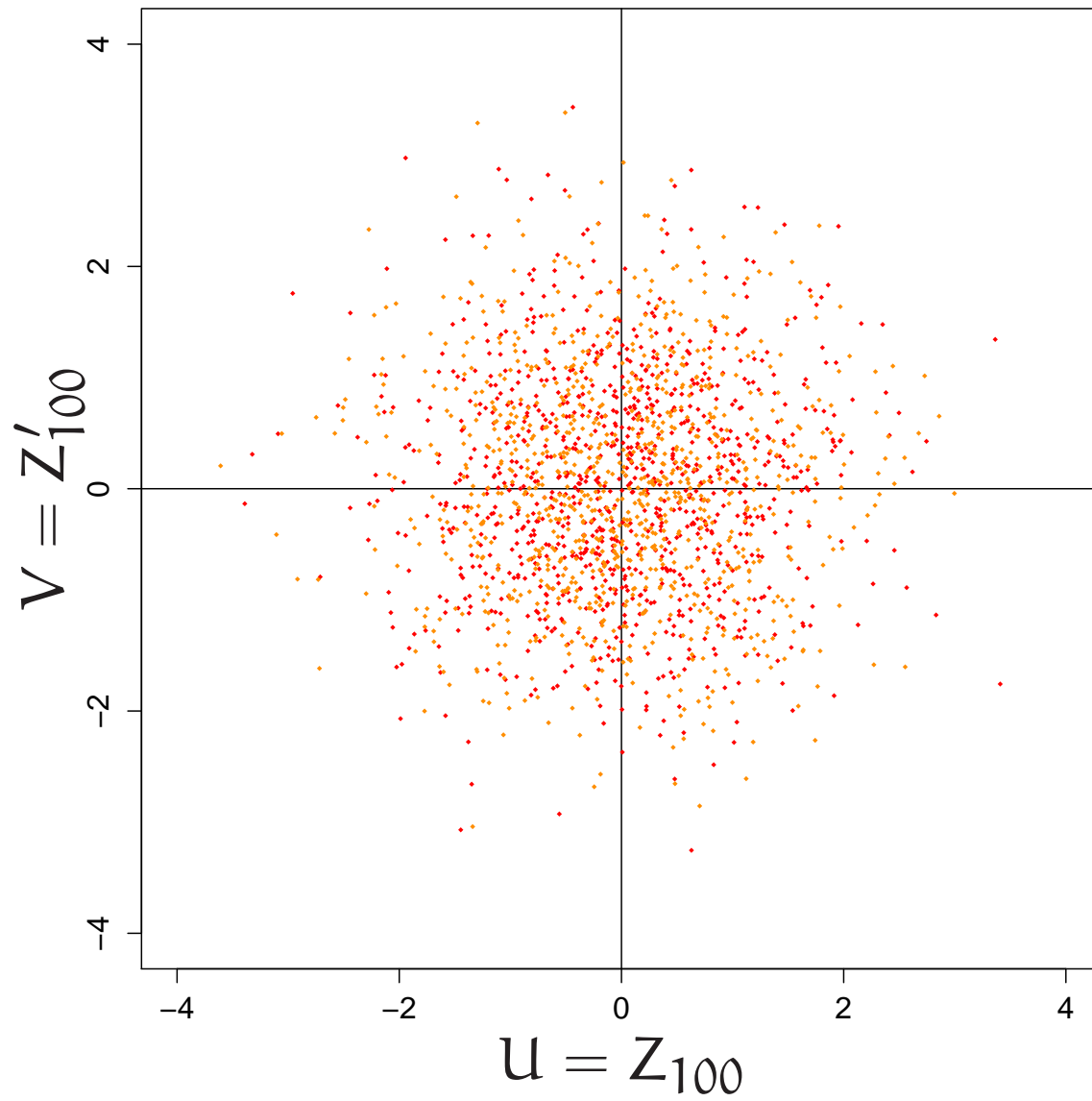
$$(U, V) = (Z_{100}, Z'_{100})$$

Wie sieht die gemeinsame Verteilung
von U und V aus?

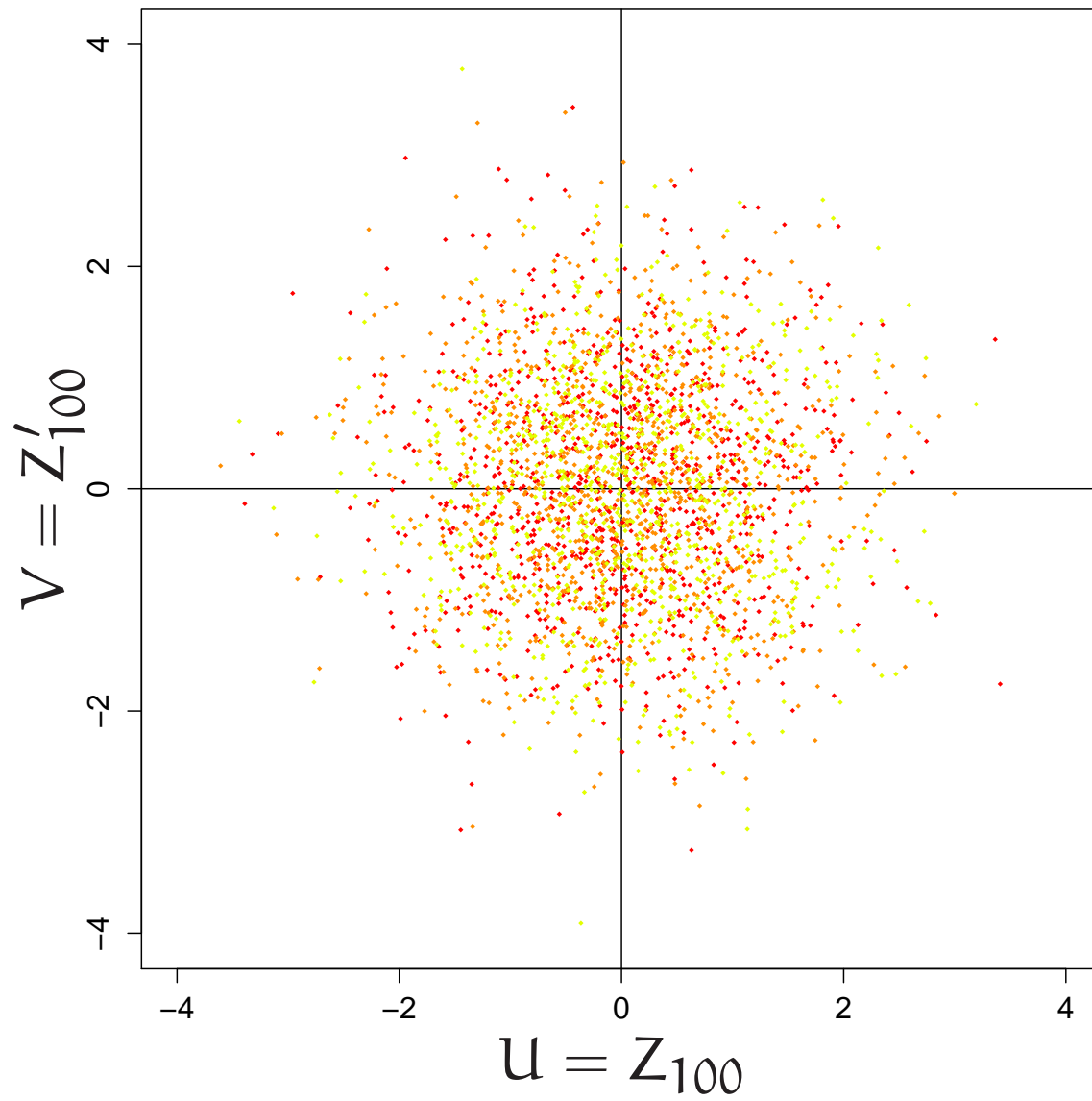
1000 Simulationen



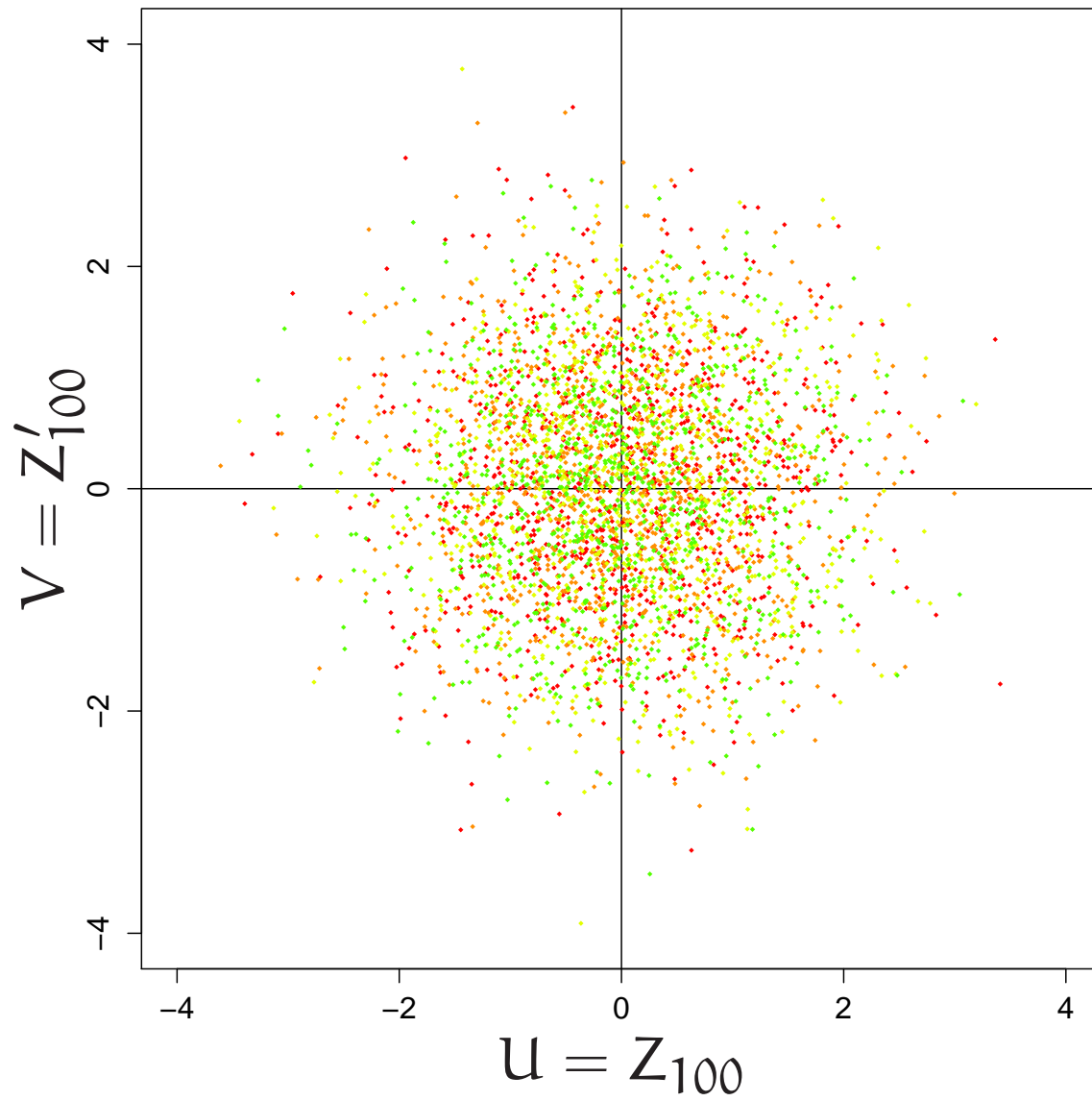
2000 Simulationen



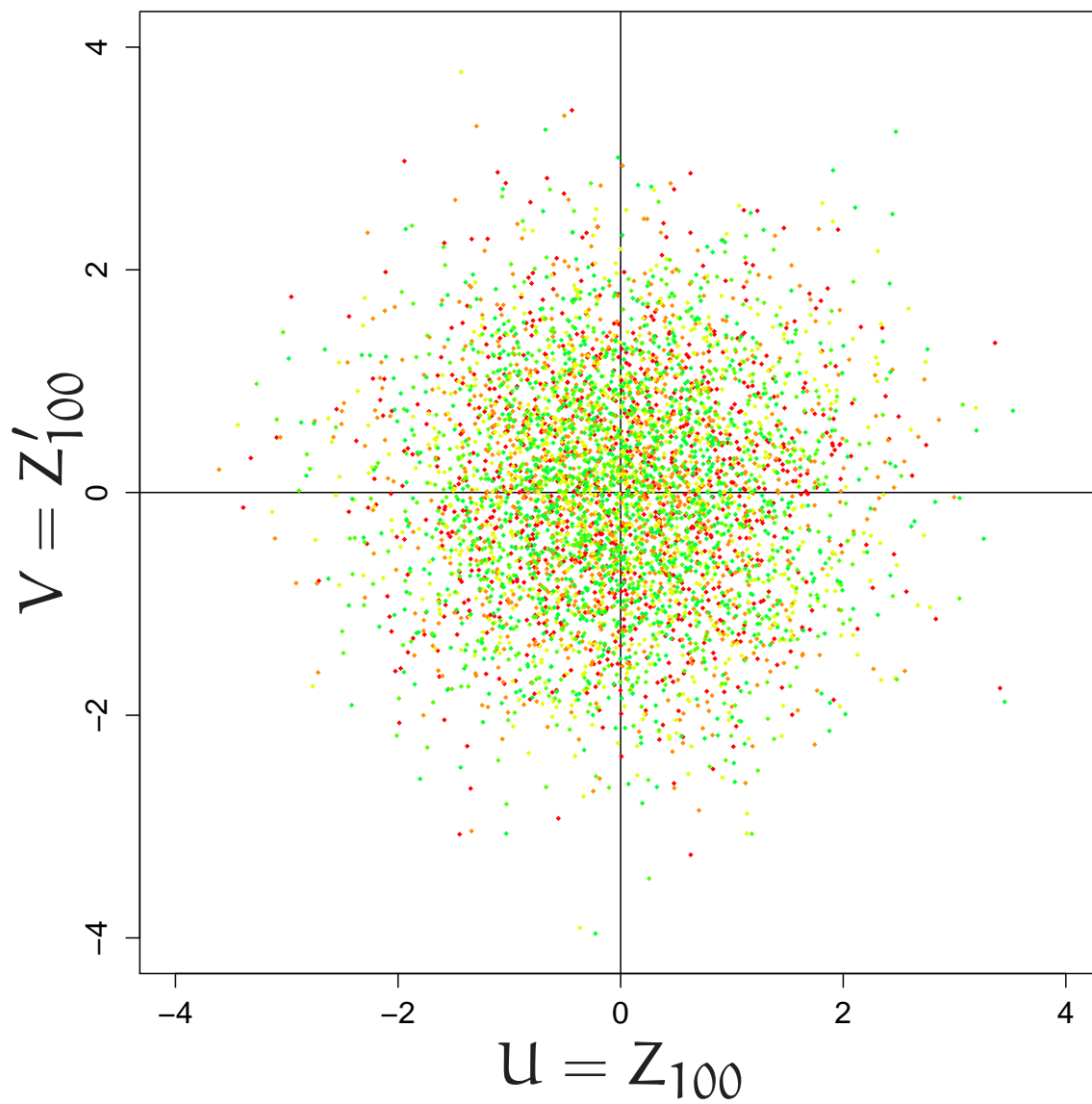
3000 Simulationen



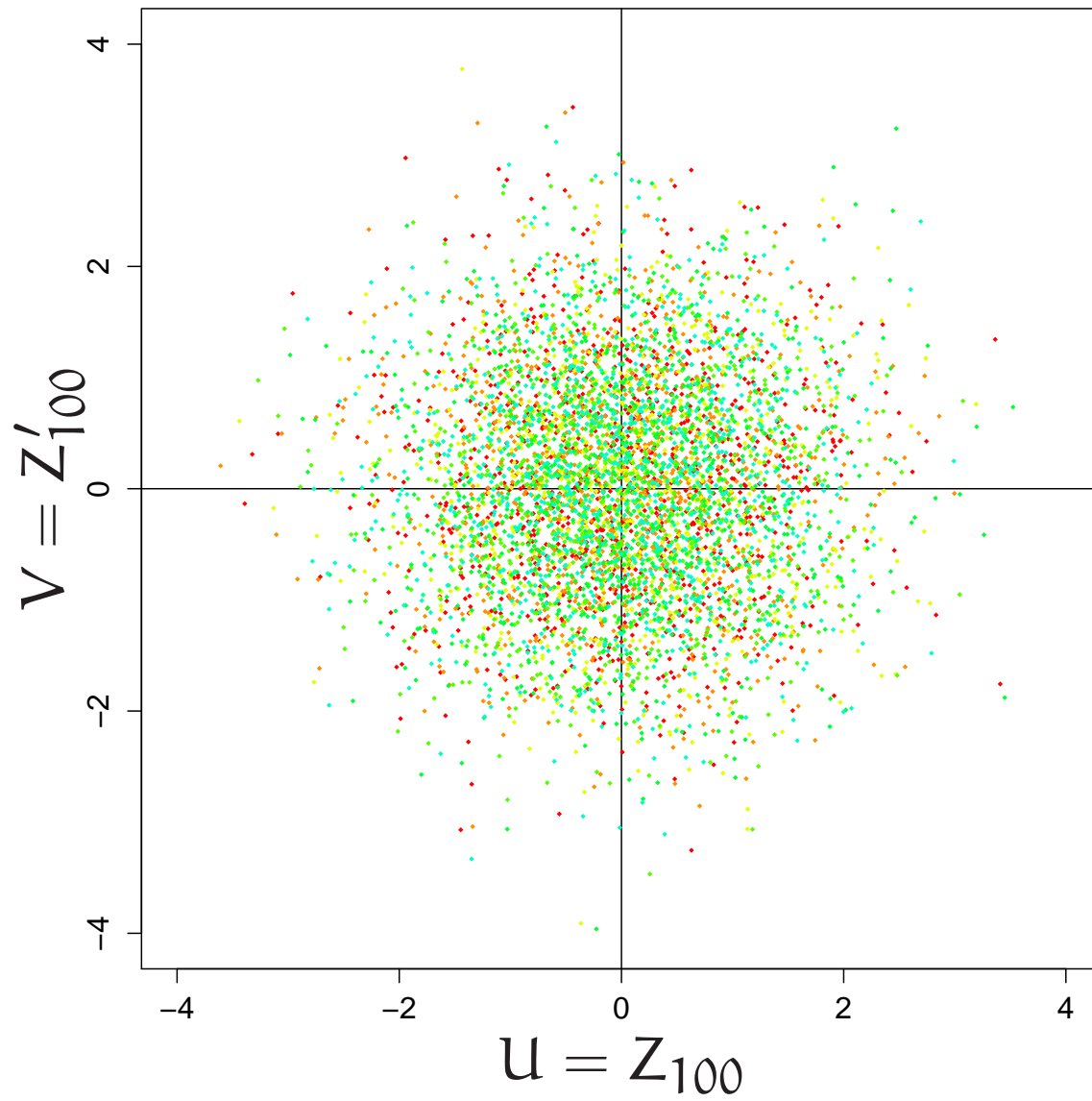
4000 Simulationen



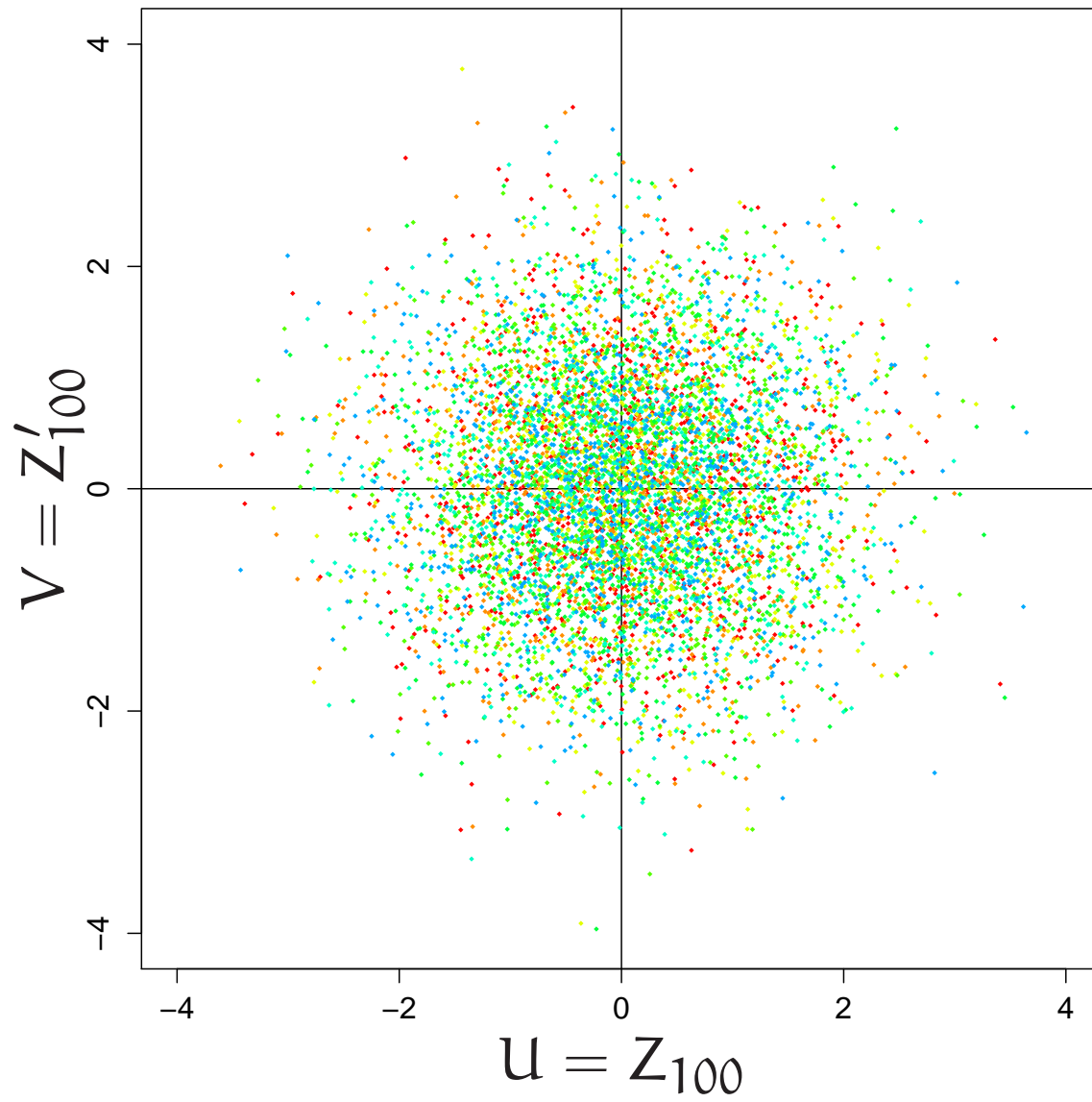
5000 Simulationen



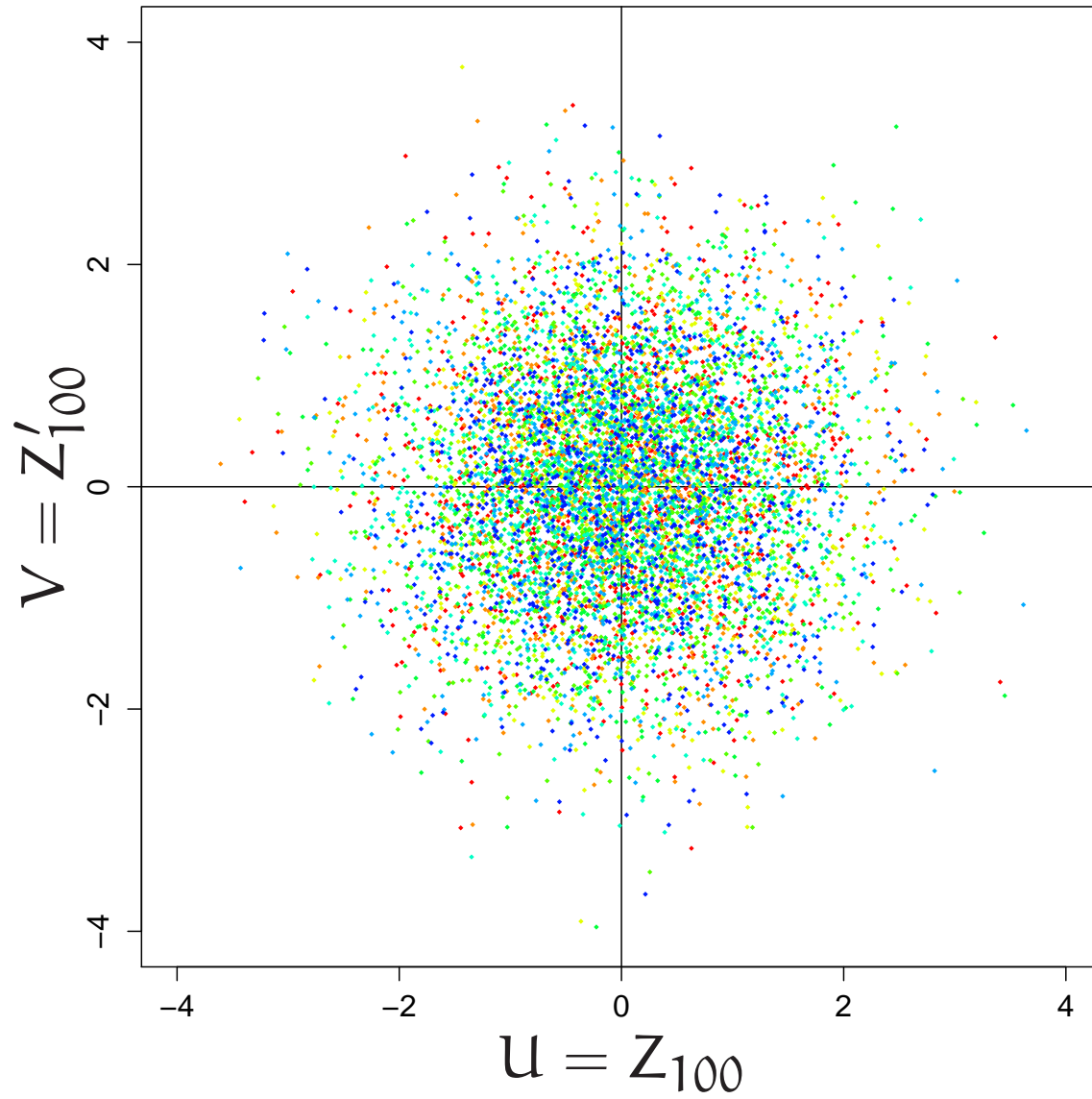
6000 Simulationen



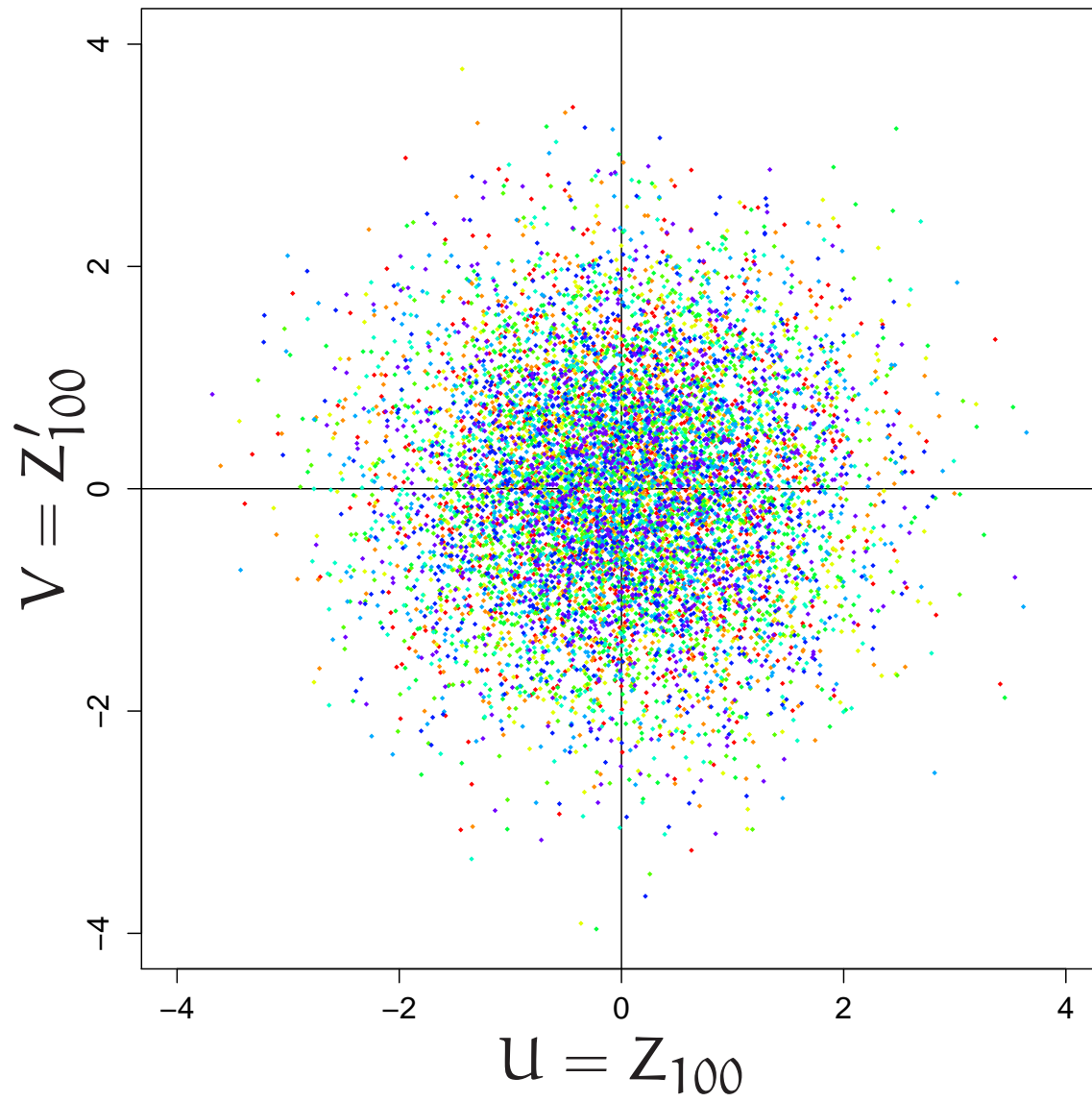
7000 Simulationen



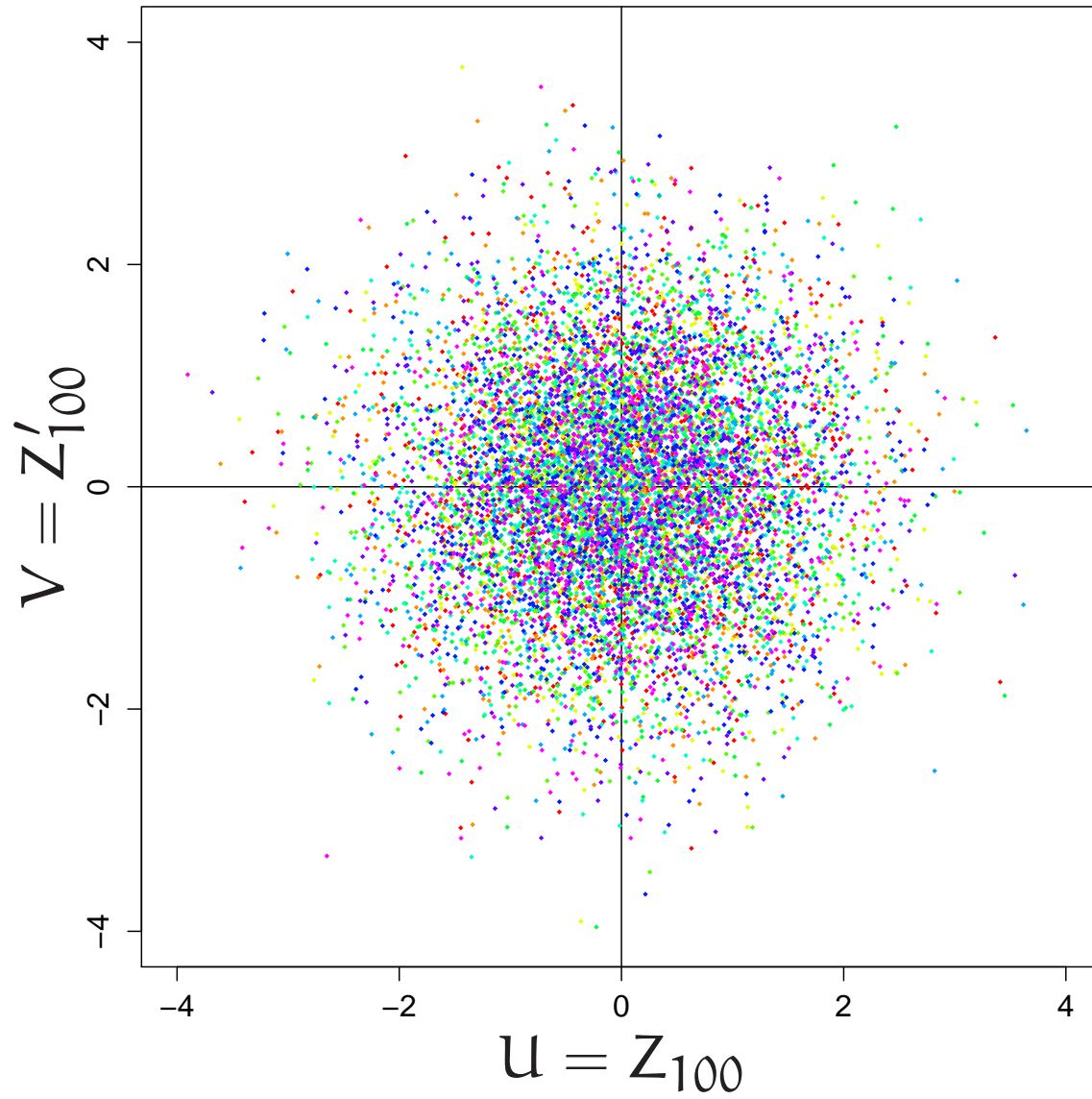
8000 Simulationen



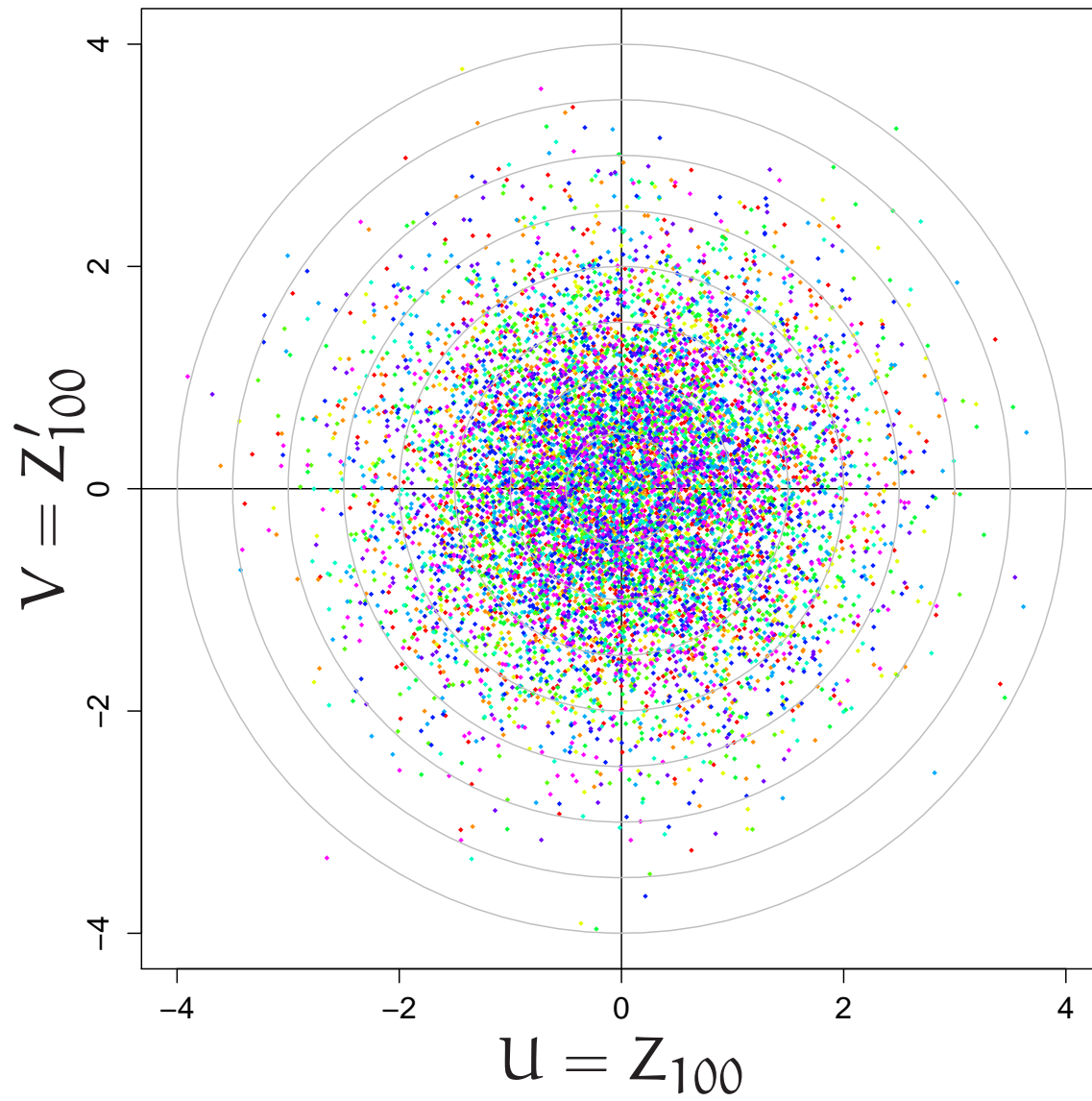
9000 Simulationen



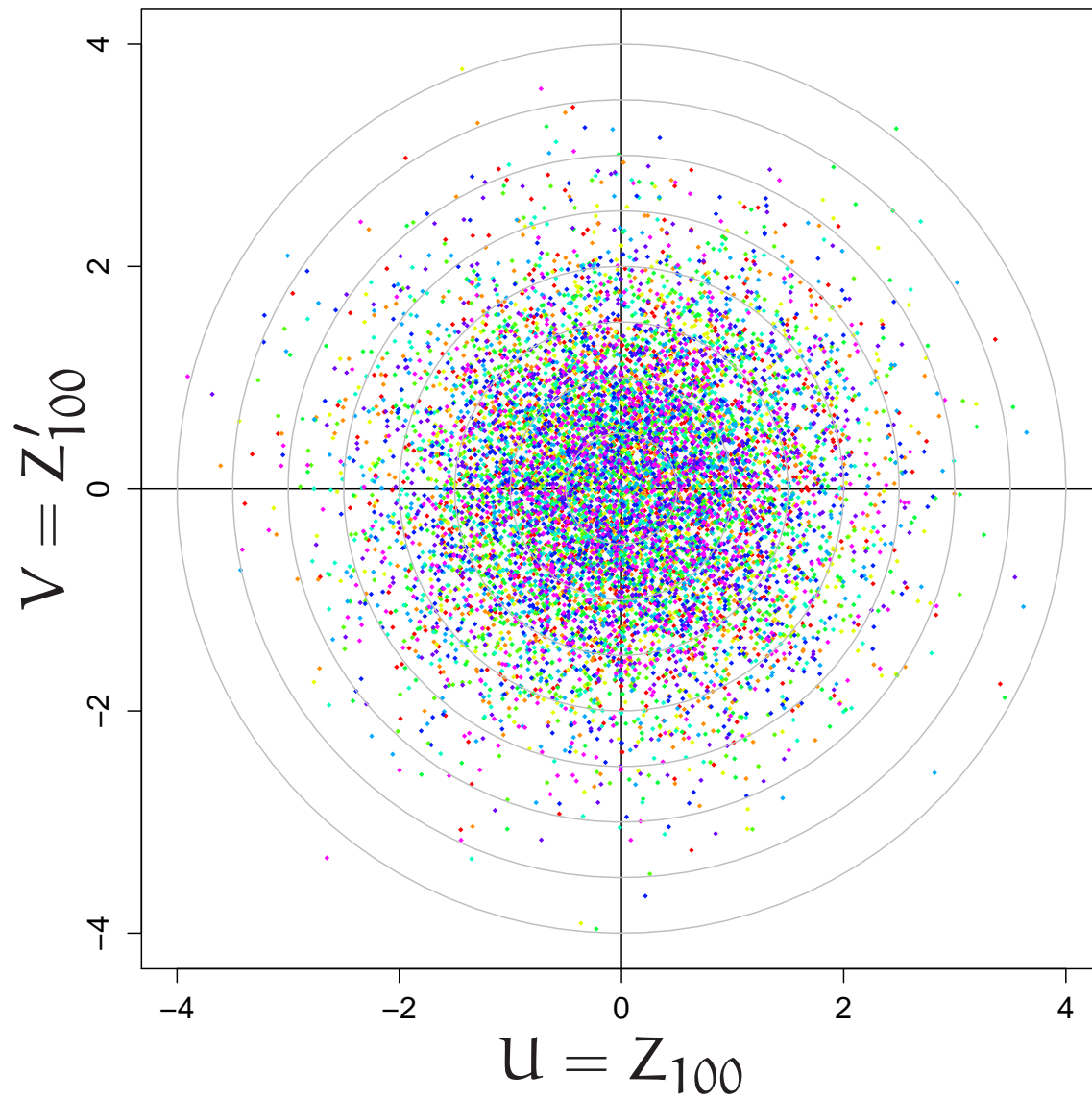
10000 Simulationen



10000 Simulationen



Die Verteilung von (U, V) ist rotationssymmetrisch!



Behauptung:

Aus U und V unabhängig

und

Verteilung von (U, V) rotationssymmetrisch

folgt,

dass U und V normalverteilt sind:

$$f_U(x) = f_V(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}$$

Denn:

U, V unabhängig bedeutet:

$$f_{(U,V)}(a, b) = f_U(a)f_V(b)$$

$f_{(U,V)}$ *rotationssymmetrisch* heißt: es existiert ein g mit

$$f_{(U,V)}(a, b) = g(r) \quad r := \sqrt{a^2 + b^2}$$

Mit $f_U = f_V =: h$ folgt

$$g(r) = h(a)h(b)$$

$$g(r) = h(a) h(b)$$

$$g(r) = c \cdot h(r) \quad (c := h(0))$$

$$c h(\sqrt{a^2 + b^2}) = h(a) h(b)$$

Eine Lösung:

$$h(x) = e^{-x^2}$$

$$e^{-(a^2+b^2)} = e^{-a^2} e^{-b^2}$$

Allgemeine Lösung
(logarithmieren; nach a ableiten)

$$h(x) = k_1 e^{-k_2 x^2}$$

Nebenbedingungen:

$$\mathbf{P}(U \in \mathbb{R}) = \int h(x) dx = 1$$

$$\sigma_U^2 = \int x^2 h(x) dx = 1$$

$$h(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

FAZIT

Der Zentrale Grenzwertsatz

lässt sich erraten

(in konkreten Fällen,

mit etwas Glück).

Der Münzwurf passt in den Zentralen Grenzwertsatz:

Seien X_1, X_2, \dots unabhängige und identisch verteilte Zufallsvariable mit endlichem Erwartungswert μ und endlicher Varianz $\sigma^2 > 0$. Dann gilt für alle $\ell < r \in \mathbb{R}$

$$\mathbf{P} \left(\frac{X_1 + \dots + X_n - n\mu}{\sqrt{n\sigma^2}} \in [\ell, r] \right) \xrightarrow{n \rightarrow \infty} \mathbf{P}(Z \in [\ell, r]).$$

Dabei ist Z standard-normalverteilt.

Mit $\mu = p$ und $\sigma^2 = pq$ ergibt sich der alte Satz von de Moivre und Laplace.

Das passt zur Approximation der Binomialgewichte

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{\sigma} \varphi\left(\frac{k - \mu}{\sigma}\right) \approx \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right) dx,$$

mit $\mu := np$, $\sigma := \sqrt{npq}$

$$\mathbf{P}(k_1 \leq X \leq k_2) = \int_{k_1-\frac{1}{2}}^{k_2+\frac{1}{2}} \frac{1}{\sigma} \varphi\left(\frac{x - \mu}{\sigma}\right) dx = \int_{(k_1-\frac{1}{2}-\mu)/\sigma}^{(k_2+\frac{1}{2}-\mu)/\sigma} \varphi(z) dz$$

Bei unserer Entdeckungsreise zum Zentralen Grenzwertsatz hatten wir die (Beinahe-) Rotationssymmetrie der Verteilung von (Z_n, Z'_n) aus den Simulationen abgelesen.

Diese lässt sich aber auch konzeptionell begründen, wenn man die Existenz (irgend-) einer Grenzverteilung im ZGS unterstellt.

Aufgabe: Entdecken Sie die Rotationssymmetrie der Grenzverteilung von (Z_n, Z'_n) (und damit die Gaußsche Glockenkurve) nochmal, indem Sie (ohne Rückgriff auf den ZGS) Folgendes begründen:

Falls für unabhängige, identisch verteilte X_1, X_2, \dots

$$\text{mit } \mathbf{E}[X_i] = 0, \mathbf{E}[X_i^2] = 1$$

die Folge $\frac{1}{\sqrt{n}}(X_1 + \dots + X_n)$ in Verteilung konvergiert,
dann hat die Grenzverteilung ν die folgende Eigenschaft:

Sind Z und Z' unabhängig mit Verteilung ν ,
und sind α und β positive Zahlen mit $\alpha^2 + \beta^2 = 1$,
dann hat auch $\alpha Z + \beta Z'$ die Verteilung ν .

Hier ist noch einmal die (im ZGS präzierte) Botschaft der Stunde:

Summen (und Mittelwerte) von vielen unabhängigen,
identisch verteilten ZV mit endlicher Varianz
sind annähernd normalverteilt.

Diese Aussage bleibt übrigens auch
unter schwächeren Bedingungen bestehen,
sowohl was die Unabhängigkeit,
als auch was die identische Verteiltheit betrifft.

Für unabhängige, identisch verteilte ZV'e X_1, X_2, \dots
mit Erwartungswert μ und endlicher Varianz σ^2
bekommen wir aus dem Zentralen Grenzwertsatz:

Für große n ist der Mittelwert $\frac{X_1 + \dots + X_n}{n}$
annähernd normalverteilt

mit Erwartungswert μ und Standardabweichung $\frac{\sigma}{\sqrt{n}}$.

Daraus folgt insbesondere, dass für große n
die Verteilung von $\frac{X_1 + \dots + X_n}{n}$

in der Nähe von μ konzentriert ist:

$$\mathbf{P} \left(\left| \frac{X_1 + \dots + X_n}{n} - \mu \right| \geq \varepsilon \right) \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

Das Schwache Gesetz der großen Zahlen

$$\mathbf{P} \left(\left| \frac{X_1 + \dots + X_n}{n} - \mu \right| \geq \varepsilon \right) \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

gilt sogar für jede Folge von paarweise unkorrelierten Zufallsvariablen mit ein- und demselben Erwartungswert μ und ein-und derselben Varianz σ^2 .

Dahinter steckt,
dass der Erwartungswert von $\frac{X_1 + \dots + X_n}{n}$ gleich μ ist
und seine Standardabweichung klein wird, nämlich $\frac{\sigma}{\sqrt{n}}$.

Den (simplen) Beweis des
Schwachen Gesetzes der Großen Zahlen
bereiten wir vor durch zwei einfache Ungleichungen:
die Markov-Ungleichung und die Chebyshev-Ungleichung.

Markov-Ungleichung:

Für jede Zufallsvariable $X \geq 0$ und jedes $\varepsilon > 0$ gilt

$$\mathbf{P}(X \geq \varepsilon) \leq \frac{1}{\varepsilon} \mathbf{E}[X].$$

Beweis: Es gilt

$$\varepsilon \mathbf{I}_{\{X \geq \varepsilon\}} = \varepsilon \mathbf{1}_{[\varepsilon, \infty]}(X) \leq X.$$

Mit der Monotonie des Erwartungswertes

folgt die Behauptung. \square

Chebyshev-Ungleichung:

Für eine reellwertige Zufallsvariable X mit endlichem Erwartungswert gilt für beliebiges $\varepsilon > 0$

$$\mathbf{P}(|X - \mathbf{E}[X]| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \cdot \mathbf{Var}[X]$$

Beweis:

Markov-Ungleichung angewandt auf $Y = (X - \mathbf{E}[X])^2$:

$$\begin{aligned} \mathbf{P}(|X - \mathbf{E}[X]| \geq \varepsilon) &= \mathbf{P}(Y \geq \varepsilon^2) \\ &\leq \varepsilon^{-2} \mathbf{E}[Y] = \varepsilon^{-2} \mathbf{Var}[X] . \end{aligned}$$

□

Schwaches Gesetz der Großen Zahlen

(Buch S. 74)

Die Zufallsvariablen X_1, X_2, \dots seien reellwertig, identisch verteilt mit endlichem Erwartungswert μ und endlicher Varianz, und sie seien unkorreliert.

Dann gilt für alle $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{X_1 + \dots + X_n}{n} - \mu \right| \geq \varepsilon \right) = 0 .$$

Beweis: Gemäß Voraussetzung gilt

$$\mathbf{E}\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{\mathbf{E}[X_1] + \cdots + \mathbf{E}[X_n]}{n} = \mu ,$$
$$\mathbf{Var}\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{\mathbf{Var}[X_1] + \cdots + \mathbf{Var}[X_n]}{n^2} = \frac{\mathbf{Var}[X_1]}{n} .$$

Die Chebyshev-Ungleichung, angewandt auf

$(X_1 + \cdots + X_n)/n$ ergibt

$$\mathbf{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\mathbf{Var}[X_1]}{\varepsilon^2 n} .$$

Die Behauptung folgt nun mit $n \rightarrow \infty$. \square

Das Schwache Gesetz der Großen Zahlen wurde von Jacob Bernoulli im Münzwurfmodell entdeckt.

