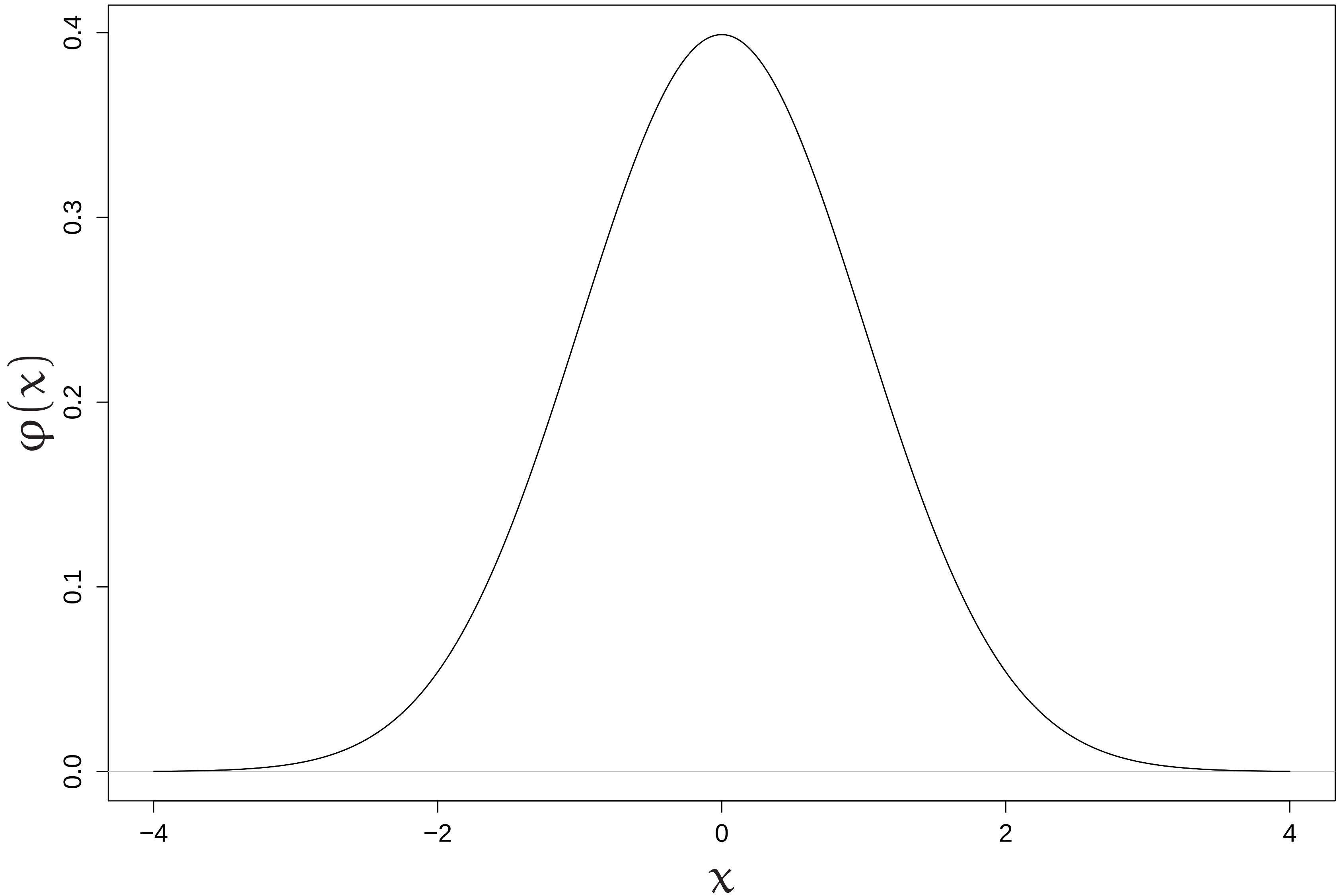
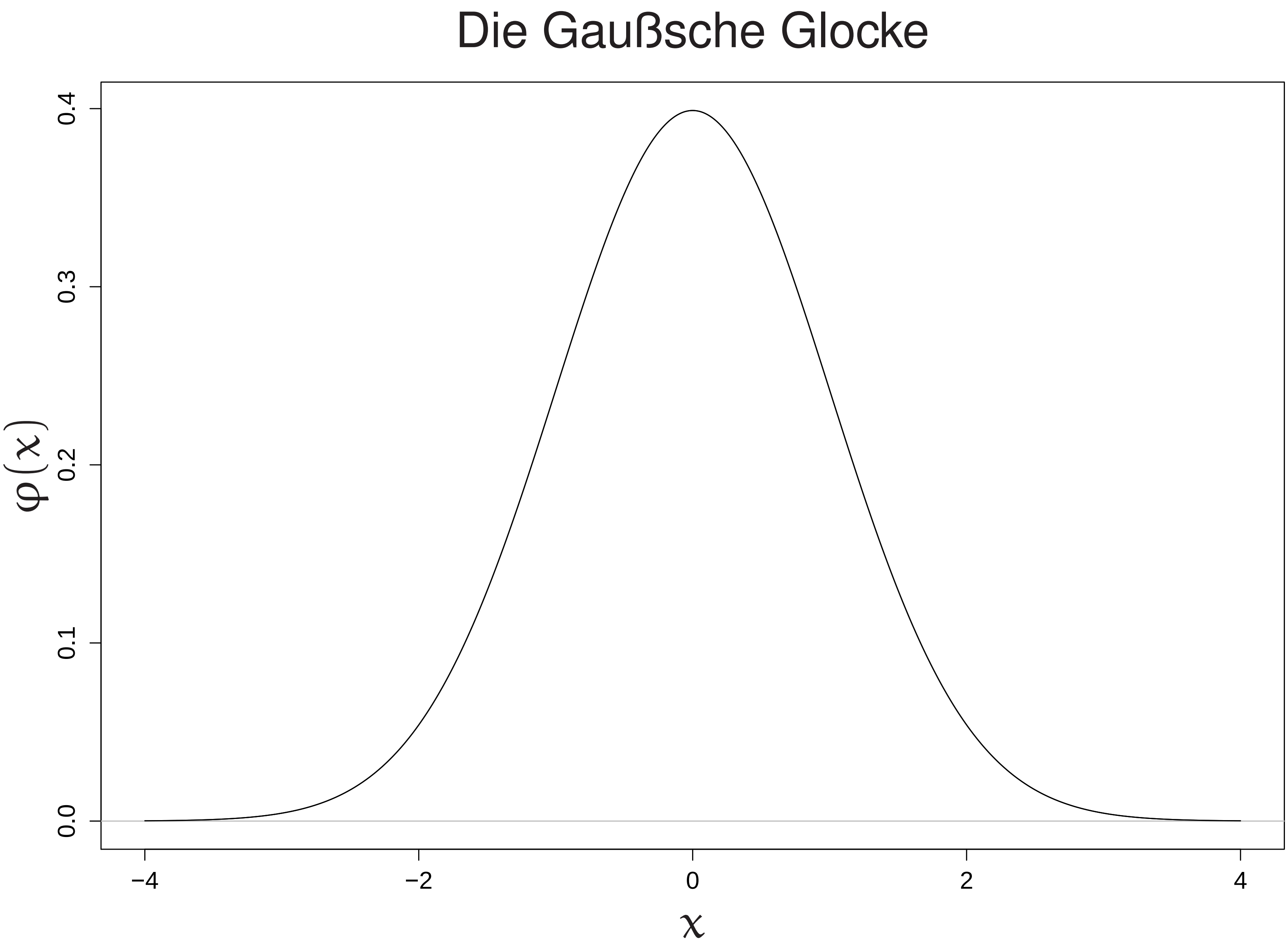
Vorlesung 7a

Der zentrale Grenzwertsatz

0. Wiederholung: Die Normalverteilung

Dichtefunktion φ der Standardnormalverteilung




Die Gaußsche Glocke
0.0
0.1
0.2
0.3
0.4
φ(x)
−4
−2
0
2
4
x

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 standard-normalverteilt:

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

Für standard-normalverteiltes Z gilt:

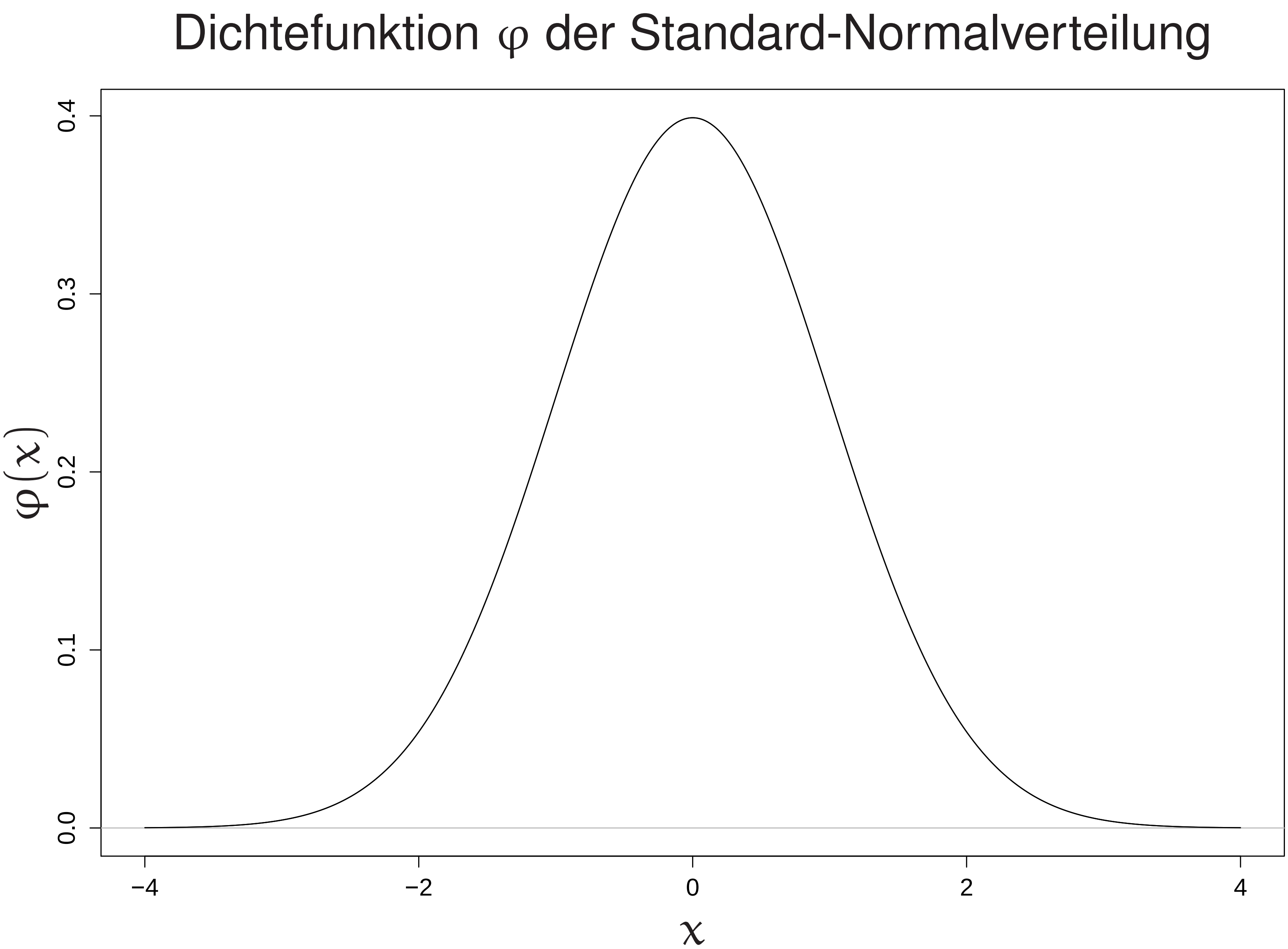
$$\mathbf{E}Z = 0 \qquad \sigma_Z = 1$$

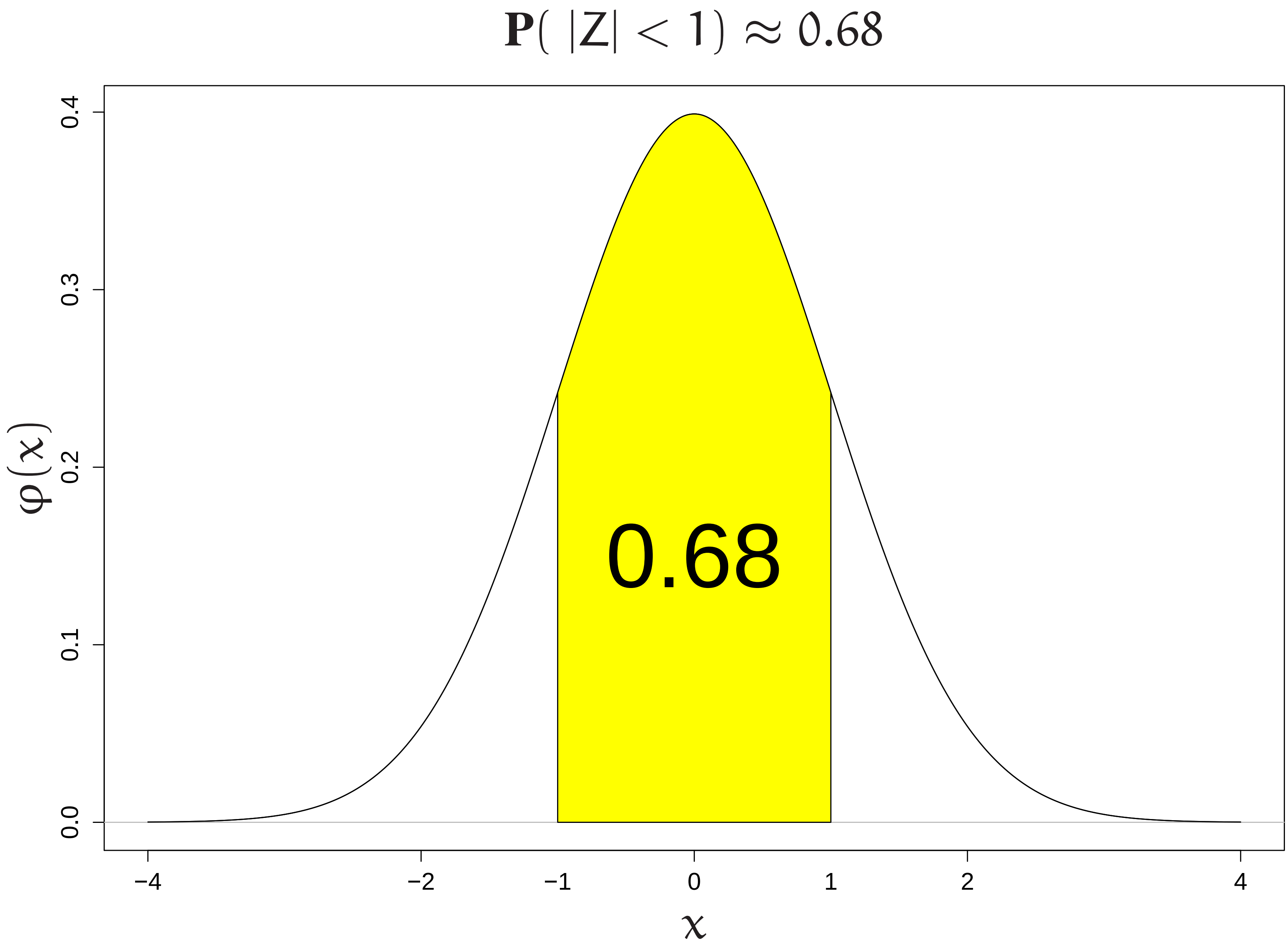
Allgemeine normalverteilte Zufallsvariable entstehen so:

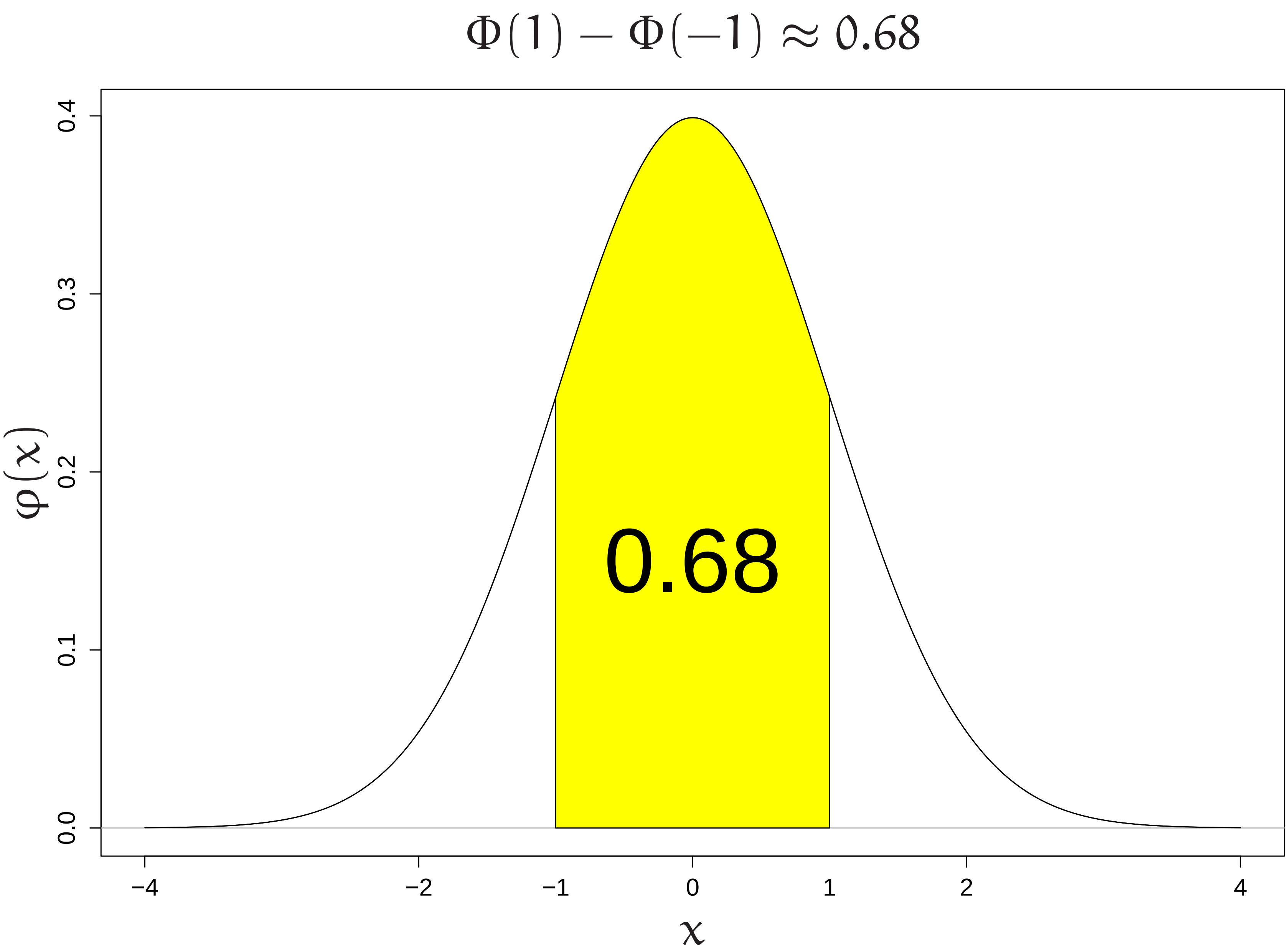
$$N = \mu + \sigma Z$$

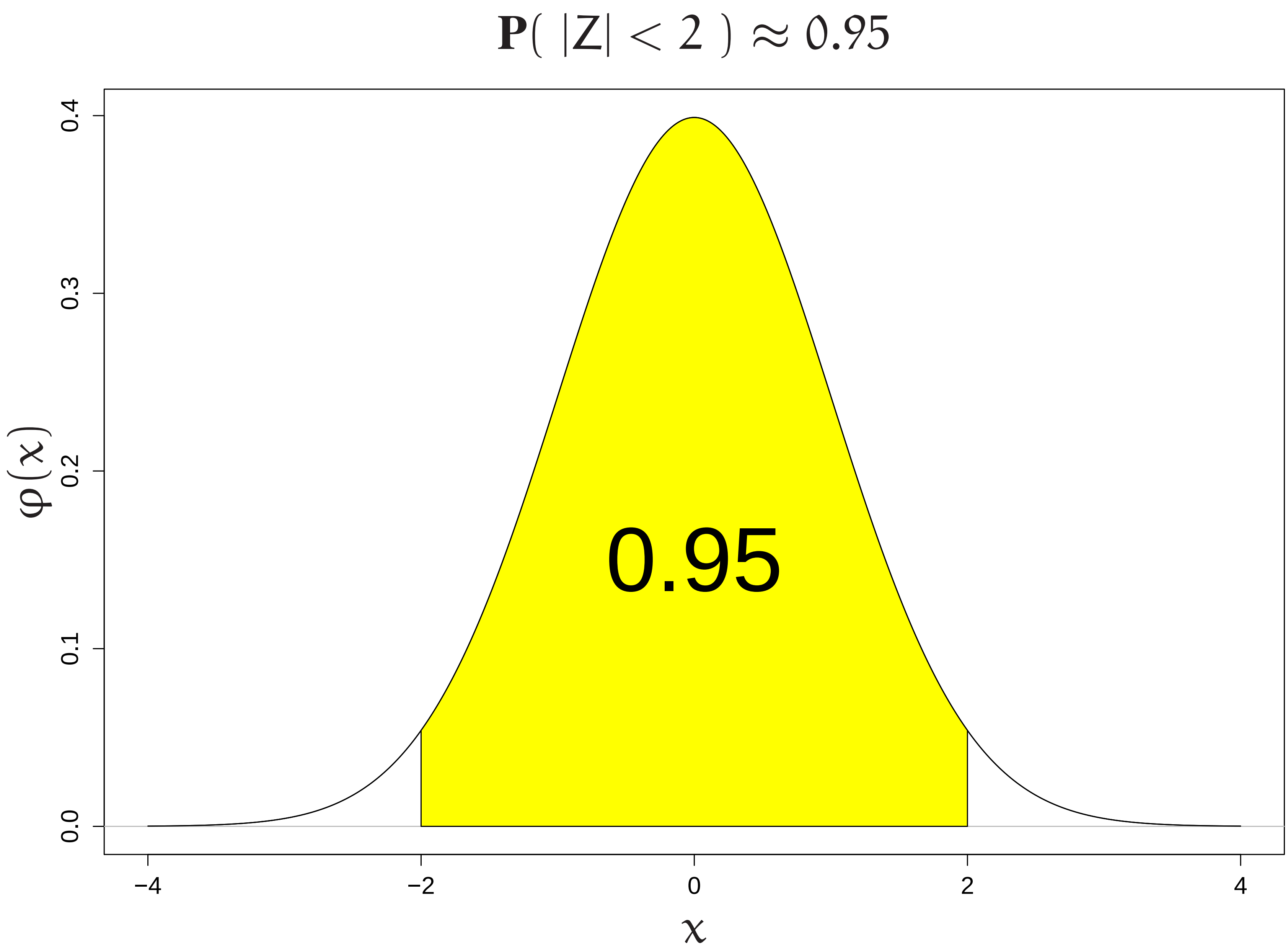
Für sie gilt:

$$\mathbf{E}N = \mu \qquad \sigma_N = \sigma$$

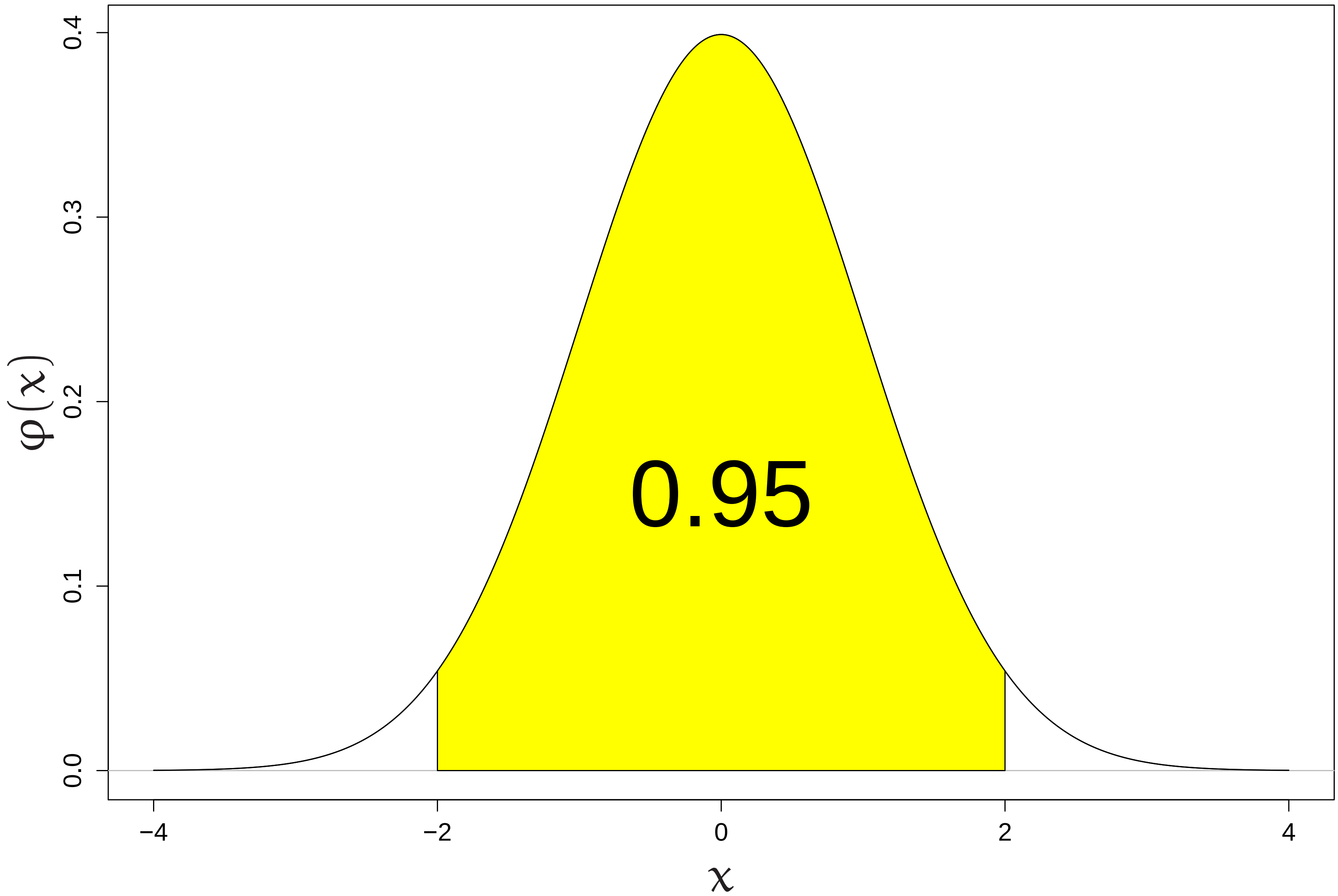

Dichtefunktion φ der Standard-Normalverteilung
φ(x)
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
x

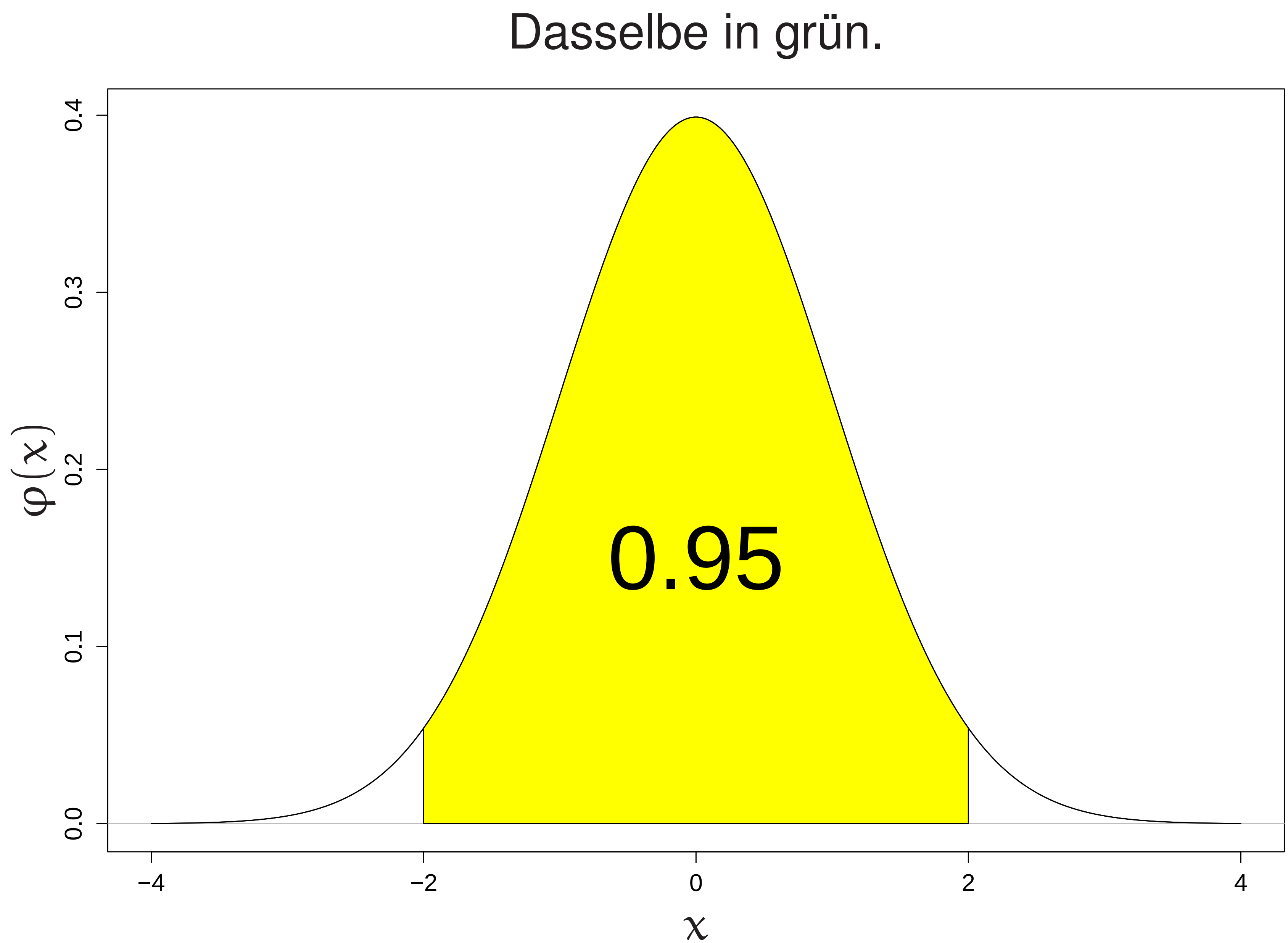

$\mathbf{P}(\ |Z| < 1) \approx 0.68$
0.68
0.0
0.1
0.2
0.3
0.4
−4
−2
−1
0
1
2
4
x
$\varphi(x)$

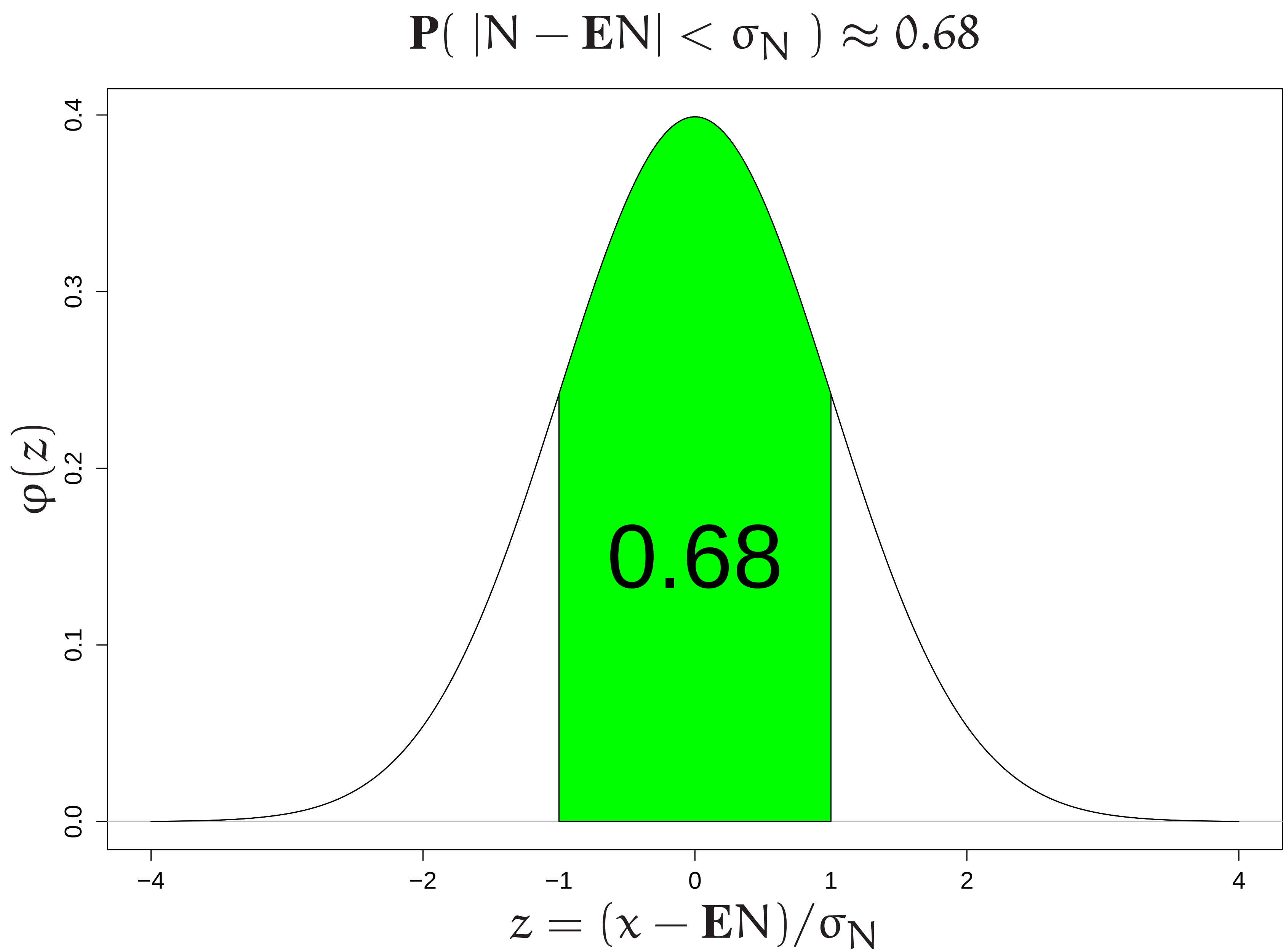

$\Phi(1) - \Phi(-1) \approx 0.68$
0.68
$\varphi(x)$
0.0
0.1
0.2
0.3
0.4
−4
−2
−1
0
1
2
4
x

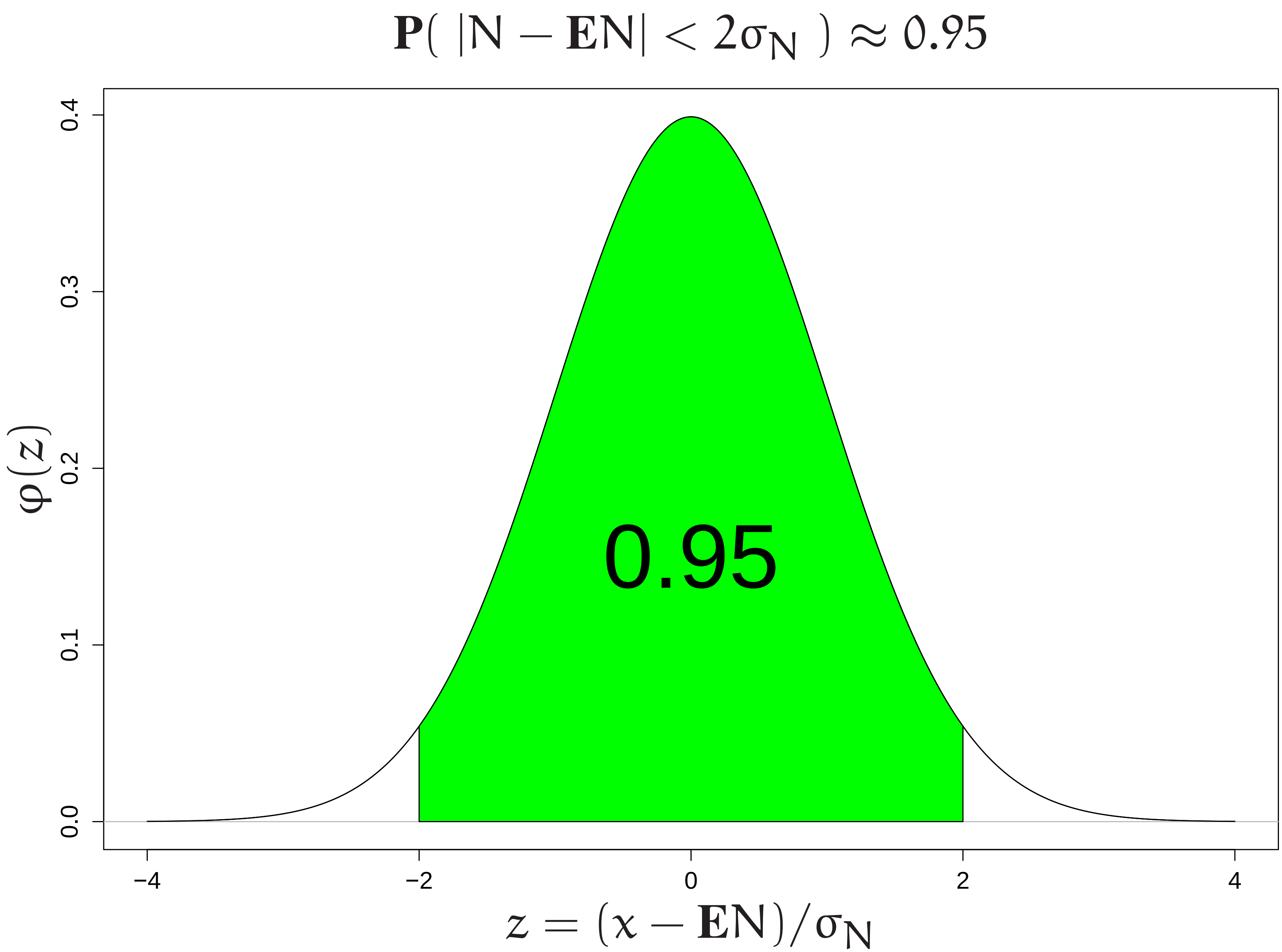

P(|Z| < 2) ≈ 0.95
0.95
φ(x)
x
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4

Und für allgemeine normalverteilte Zufallsgrößen $\mathcal{N}$?




Dasselbe in grün.
0.95
φ(x)
x
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4


$\mathbf{P}(\ |N - \mathbf{E}N| < \sigma_N\ ) \approx 0.68$
0.68
$\varphi(z)$
0.0
0.1
0.2
0.3
0.4
−4
−2
−1
0
1
2
4
$z = (x - \mathbf{E}N)/\sigma_N$



$\mathbf{P}(\ |N - \mathbf{E}N| < 2\sigma_N\ ) \approx 0.95$
0.95
$\varphi(z)$
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
$z = (x - \mathbf{E}N)/\sigma_N$

1. Unabhängige normalverteilte Zufallsvariable und deren Summen

Zur Erinnerung:

$(Z_1, \ldots, Z_n)$ heißt standard-normalverteilt im $\mathbb{R}^n$

$$:\Longleftrightarrow$$

$Z_1, \ldots, Z_n$ sind unabhängig und N(0,1)-verteilt

Aus der

Rotationssymmetrie der Standard-Normalverteilung im $\mathbb{R}^n$

hatten wir gefolgert:

Für Zahlen $\tau_1, \ldots, \tau_n$ mit $\tau_1^2 + \cdots + \tau_n^2 = 1$ gilt

$\tau_1 Z_1 + \cdots \tau_n Z_n$ ist N(0,1)-verteilt.

Eine wichtige Folgerung hieraus:

Die Summe von

unabhängigen, normalverteilten Zufallsvariablen

ist wieder normalverteilt.

Denn:

$$\sigma_1 Z_1 + \sigma_2 Z_2 = \sqrt{\sigma_1^2 + \sigma_2^2} \left(\frac{\sigma_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} Z_1 + \frac{\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} Z_2 \right)$$

ist $N(0, \sigma_1^2 + \sigma_2^2)$-verteilt, und

$$(\sigma_1 Z_1 + \mu_1) + (\sigma_2 Z_2 + \mu_2) = (\sigma_1 Z_1 + \sigma_2 Z_2) + (\mu_1 + \mu_2).$$

Insbesondere ergibt sich:

Die standardisierte Summe von unabhängigen, identisch normalverteilten Zufallsvariablen ist standard-normalverteilt

Mit anderen Worten: Sind $N_1, N_2, \ldots, N_n$ unabhängig und $N(\mu, \sigma^2)$-verteilt, dann ist

$$\frac{N_1 + \cdots + N_n - n\mu}{\sqrt{n\sigma^2}}$$ standard-normalverteilt.

2. Der Zentrale Grenzwertsatz: Die Botschaft

Der Zentrale Grenzwertsatz liefert

eine gewaltige Weiterung der vorigen Aussage

(asymptotisch für große n):

Zentraler Grenzwertsatz:

"Die standardisierte Summe von VIELEN

unabhängigen, identisch verteilten

nicht notwendig normalverteilten

$\mathbb{R}$-wertigen Zufallsvariablen

mit endlicher Varianz

ist annähernd standard-normalverteilt"

Formal:

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

$$\mathbf{P}\Big(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}} \in [c, d]\Big) \underset{n\to\infty}{\longrightarrow} \mathbf{P}(Z \in [c, d]).$$

Dabei ist Z standard-normalverteilt.

In Worten:

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

Ein (erster) Hinweis darauf, dass die standardisierte Summe von n unabhängigen, identisch verteilten $\mathbb{R}$-wertigen ZV'en etwas mit asymptotischer Normalität zu tun haben könnte:

Sei X eine Zufallsvariable mit $\mathbf{E}[X] = 0$ und $\mathbf{Var}[X] = 1$.

$$X_1, X_2, \ldots, \quad X_1', X_2', \ldots$$

seien unabhängige, identisch verteilte Kopien von X.

$$Z_n := \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n), \quad Z_n' := \frac{1}{\sqrt{n}}(X_1' + \cdots + X_n').$$

Dann ist $\frac{1}{\sqrt{2}}(Z_n + Z_n')$ so verteilt wie Z_{2n}.

Und für unabhängige, N(0,1)-verteilte Zufallsvariable Z, Z'

wissen wir:

$\frac{1}{\sqrt{2}}(Z + Z')$ ist so verteilt wie Z

weil ja, wie wir schon gesehen haben,

(Z, Z') rotationssymmetrisch verteilt ist!

Vielleicht ist für große n auch

(Z_n, Z_n') annähernd rotationssymmetrisch verteilt?

3. Zentraler Grenzwertsatz:

Meilensteine in seiner Geschichte

Abraham de Moivre:

Der faire Münzwurf (1733)

Pierre-Simon Laplace:

Allgemeine binomiale Zufallsgrößen (1812)

Pafnuty Lvovich Chebyshev:

Skizze eines Beweises für den allgemeinen Fall (1887)

Aleksandr Mikhailovich Lyapunov:

Noch allgemeiner (1901, 1906)

Andrei Andreyevich Markov:

weitere Verallgemeinerungen (~ 1910)

Nehmen wir an,

diese Herren hätten sich

auf ihre vielen anderen Interessen

beschränkt.

ZENTRALER GRENZWERTSATZ

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

e

Ein Ausflug mit Brooks Ferebee

4. Ein Beispiel: Summen von unabhängigen uniform verteilten Zufallsvariablen

Wir denken an

Rundungsfehler bei Addition

In Wirklichkeit

$$\pi =$$

$$3.14159265358979323846264338327950288419716939937510 5...$$

Im Rechner

$$\pi \leftarrow 3.14159265358979$$

MODELL

Zahl = Rechnerdarstellung + Rundungsfehler.

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

Annahme: 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?

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

Ein Beispiel:

$X_1, X_2, \ldots$ unabhängig

und uniform auf $[-0.5, 0.5]$ verteilt

Empirische Verteilung von

$$S_n := X_1 + ... + X_n$$

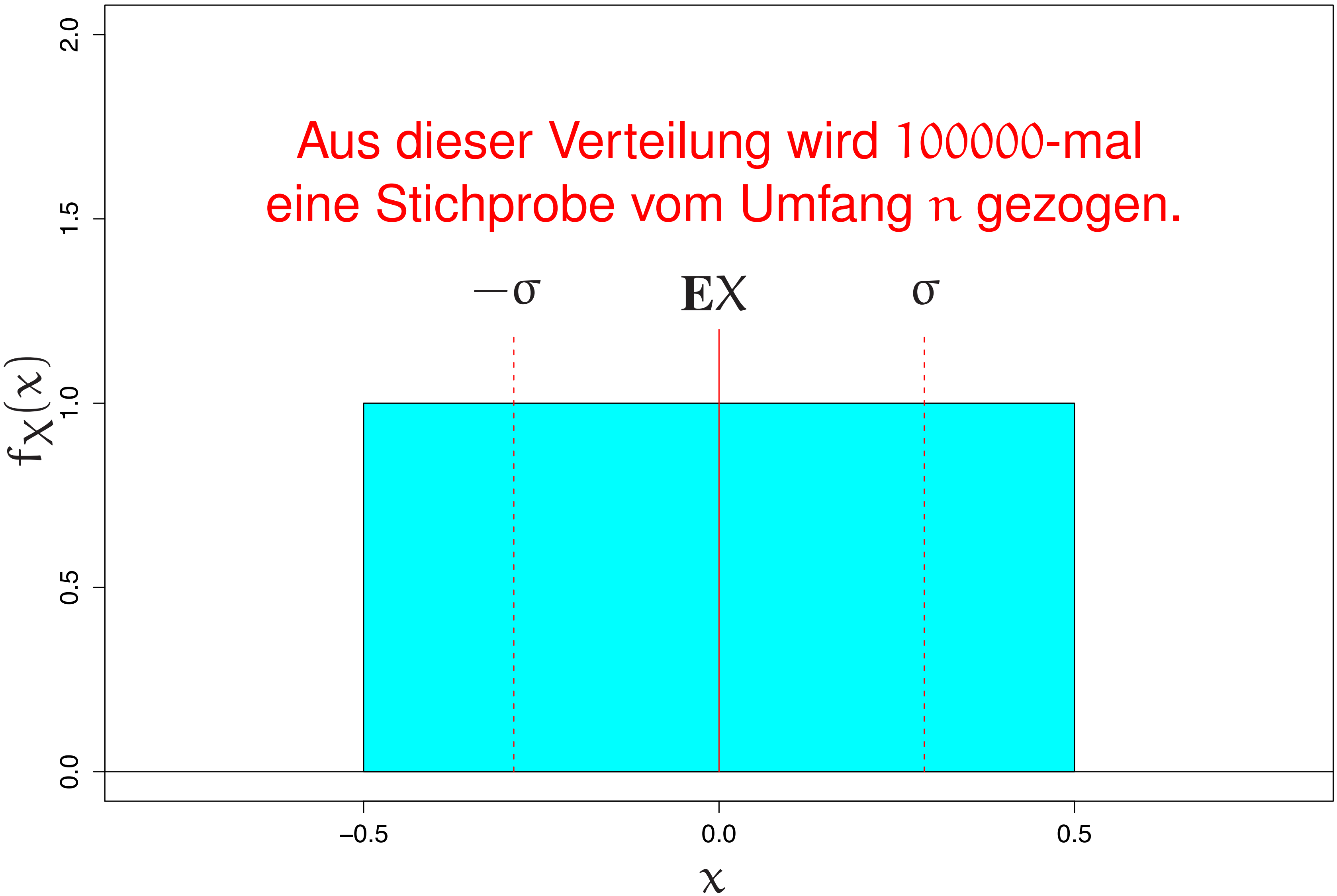
100000 Simulationen

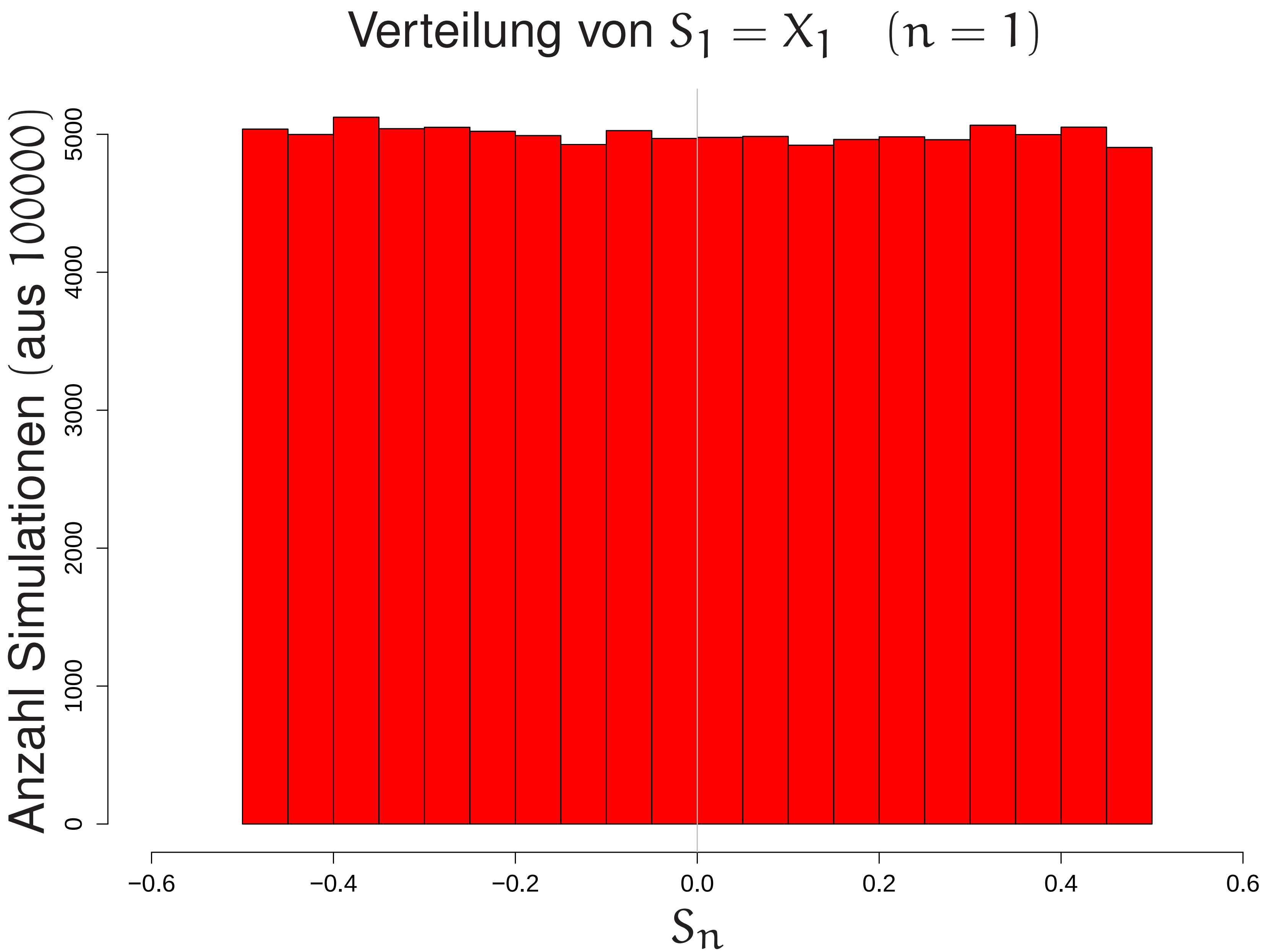
jeweils für

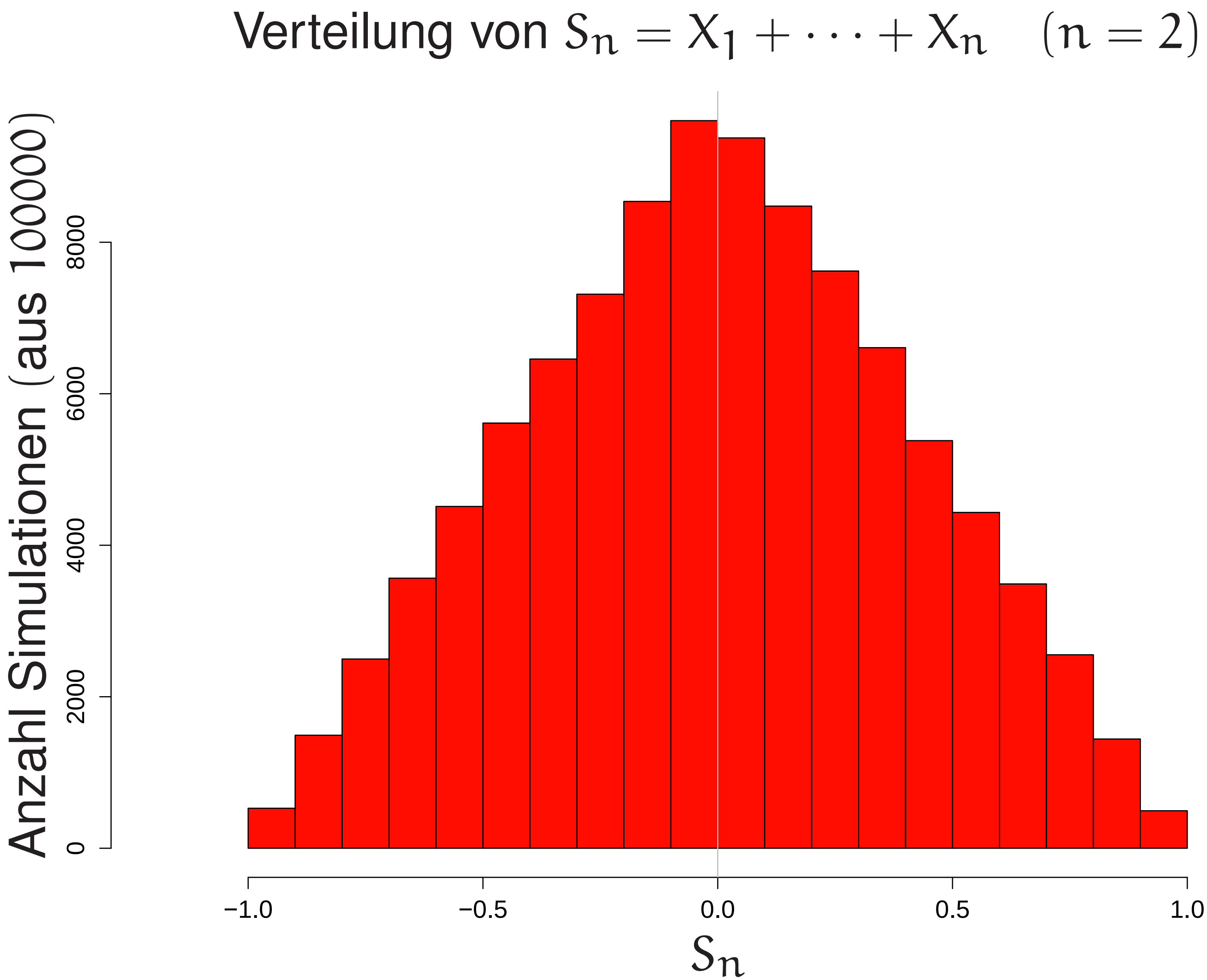
$$n = 1, 2, ..., 10$$

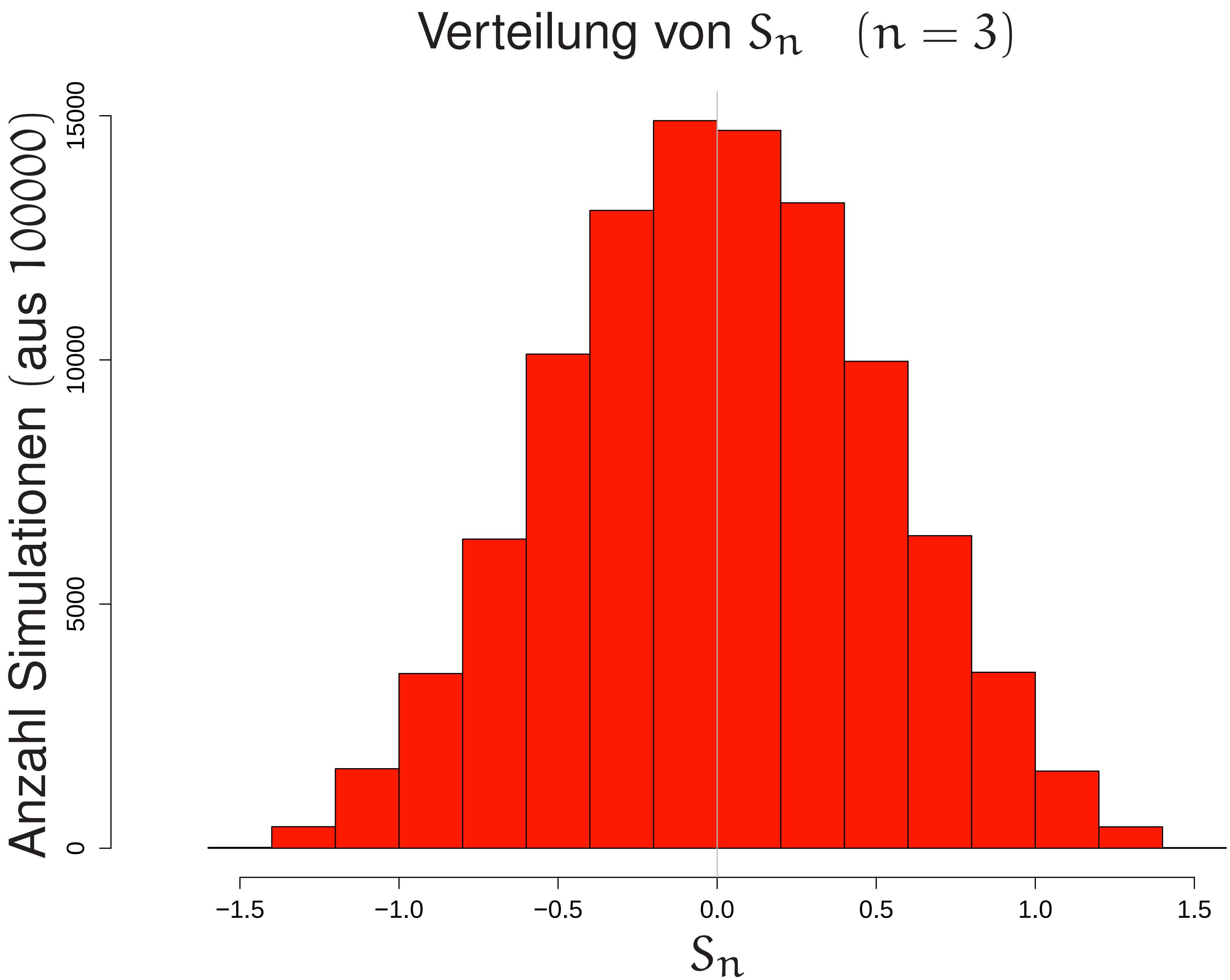
$$n = 15, 20, ..., 100$$

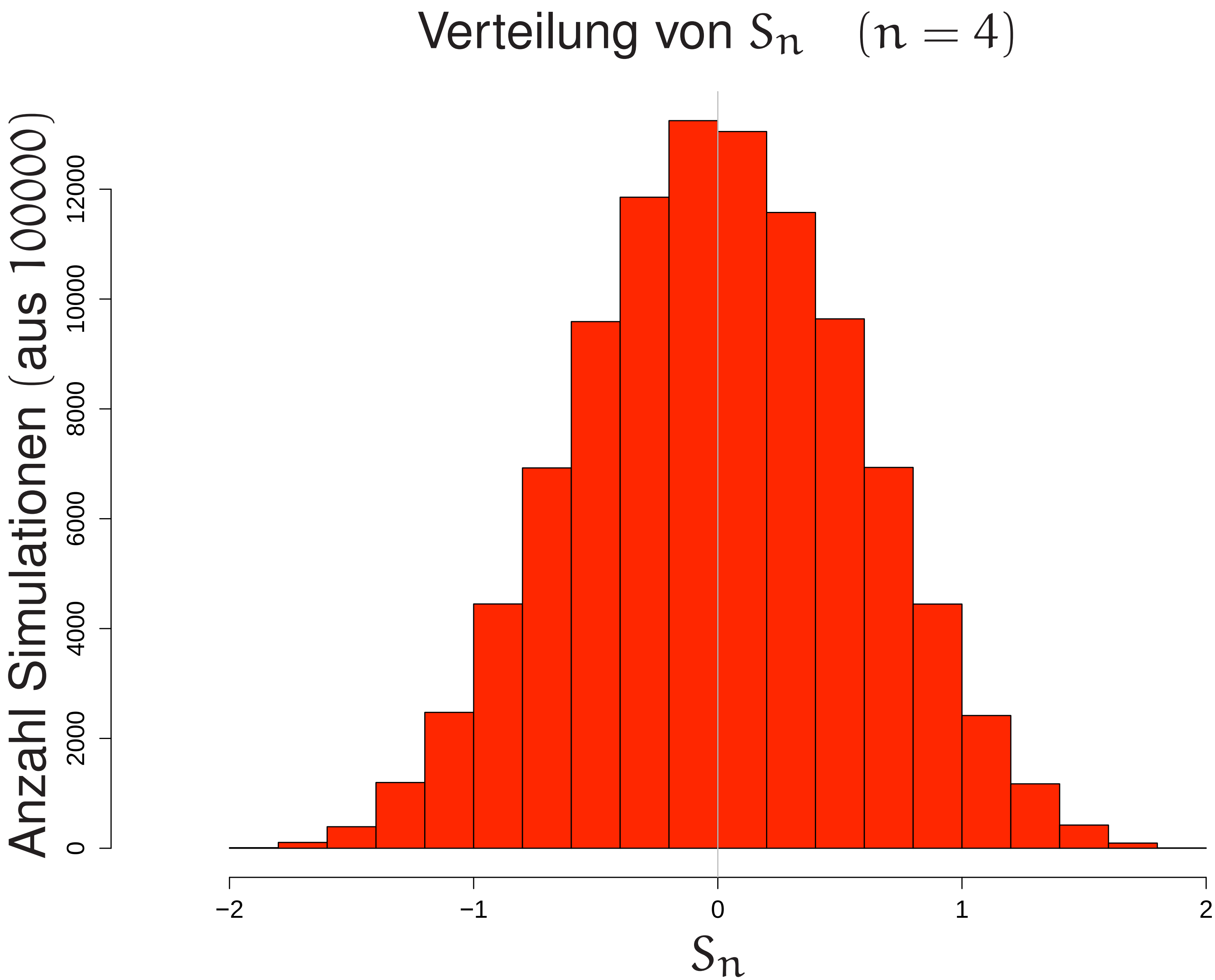
Dichtefunktion f_X der Verteilung von X



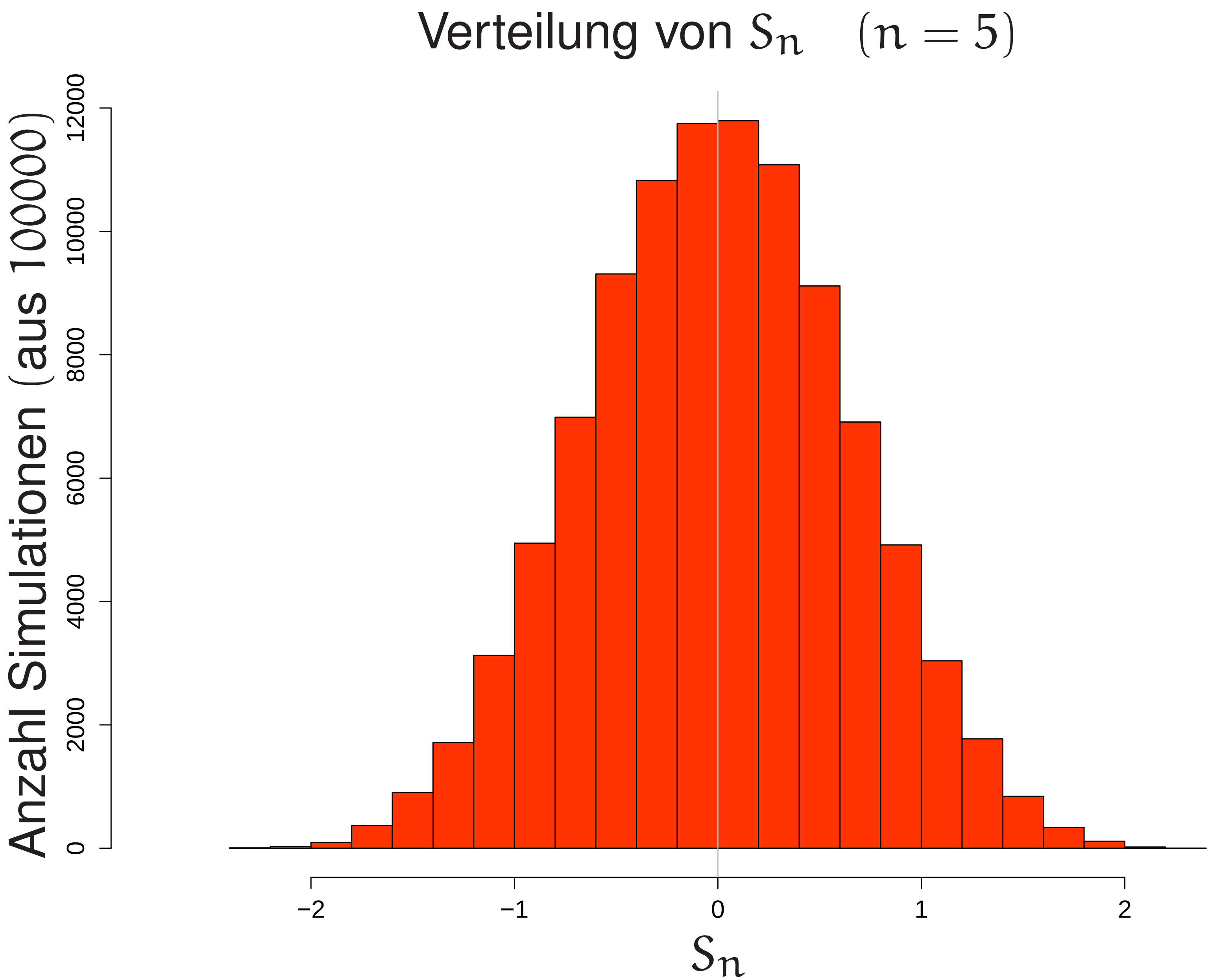

Verteilung von $S_1 = X_1 \quad (n = 1)$
Anzahl Simulationen (aus 100000)
0
1000
2000
3000
4000
5000
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
S_n


Verteilung von $S_n = X_1 + \cdots + X_n \quad (n = 2)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−1.0
−0.5
0.0
0.5
1.0
S_n

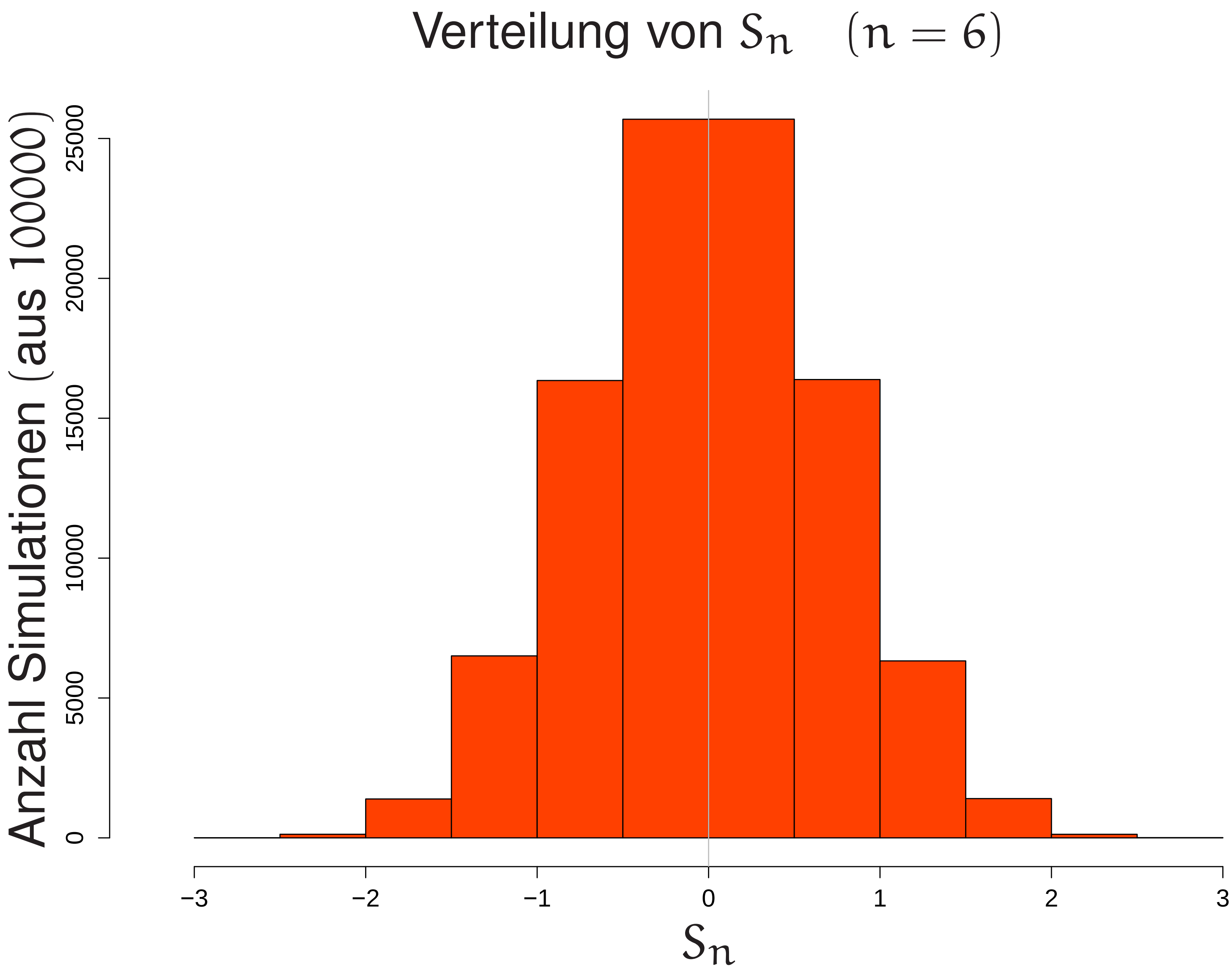

Verteilung von S_n $(n = 3)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
S_n



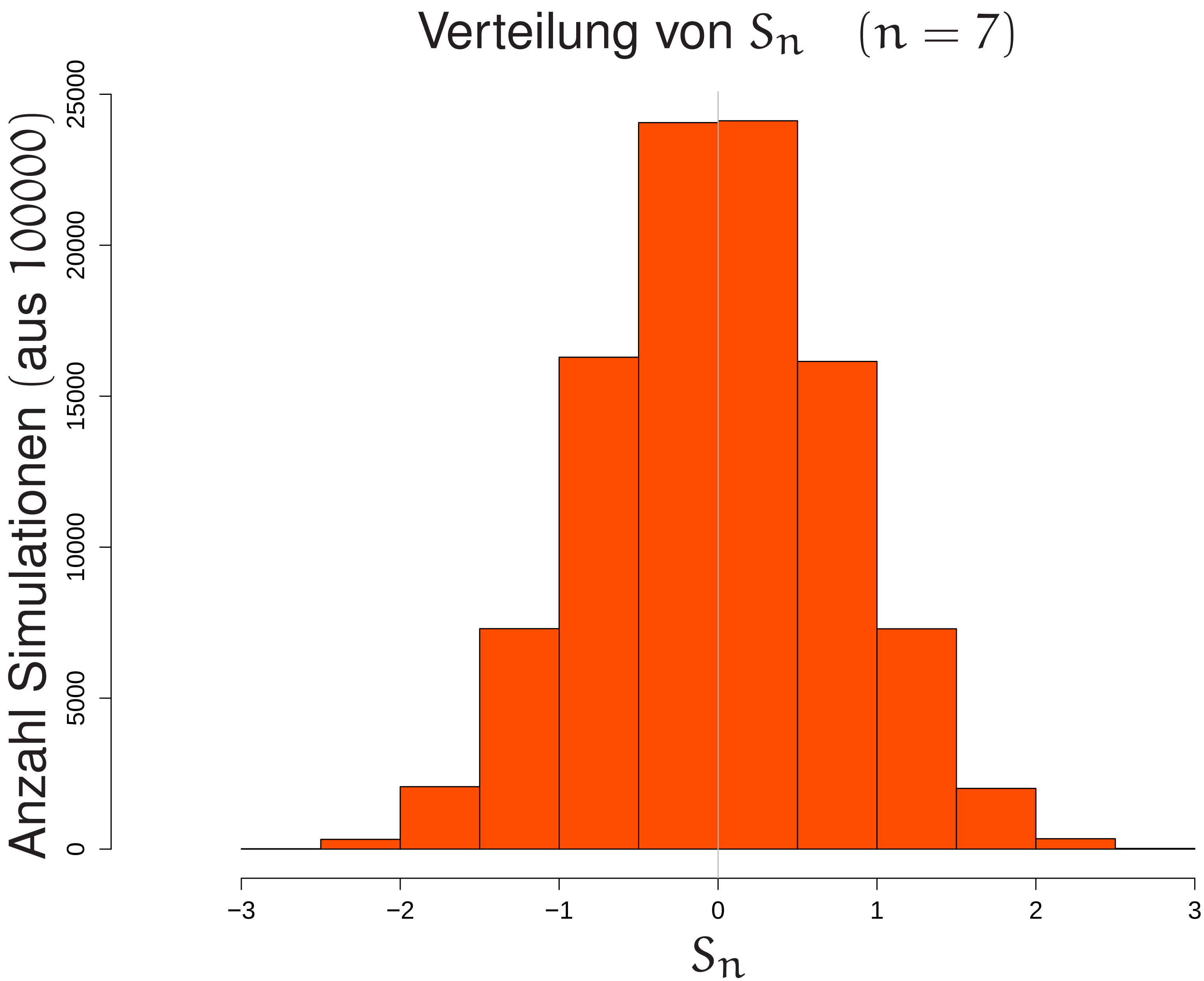
Verteilung von S_n $(n = 4)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
10000
12000
−2
−1
0
1
2
S_n



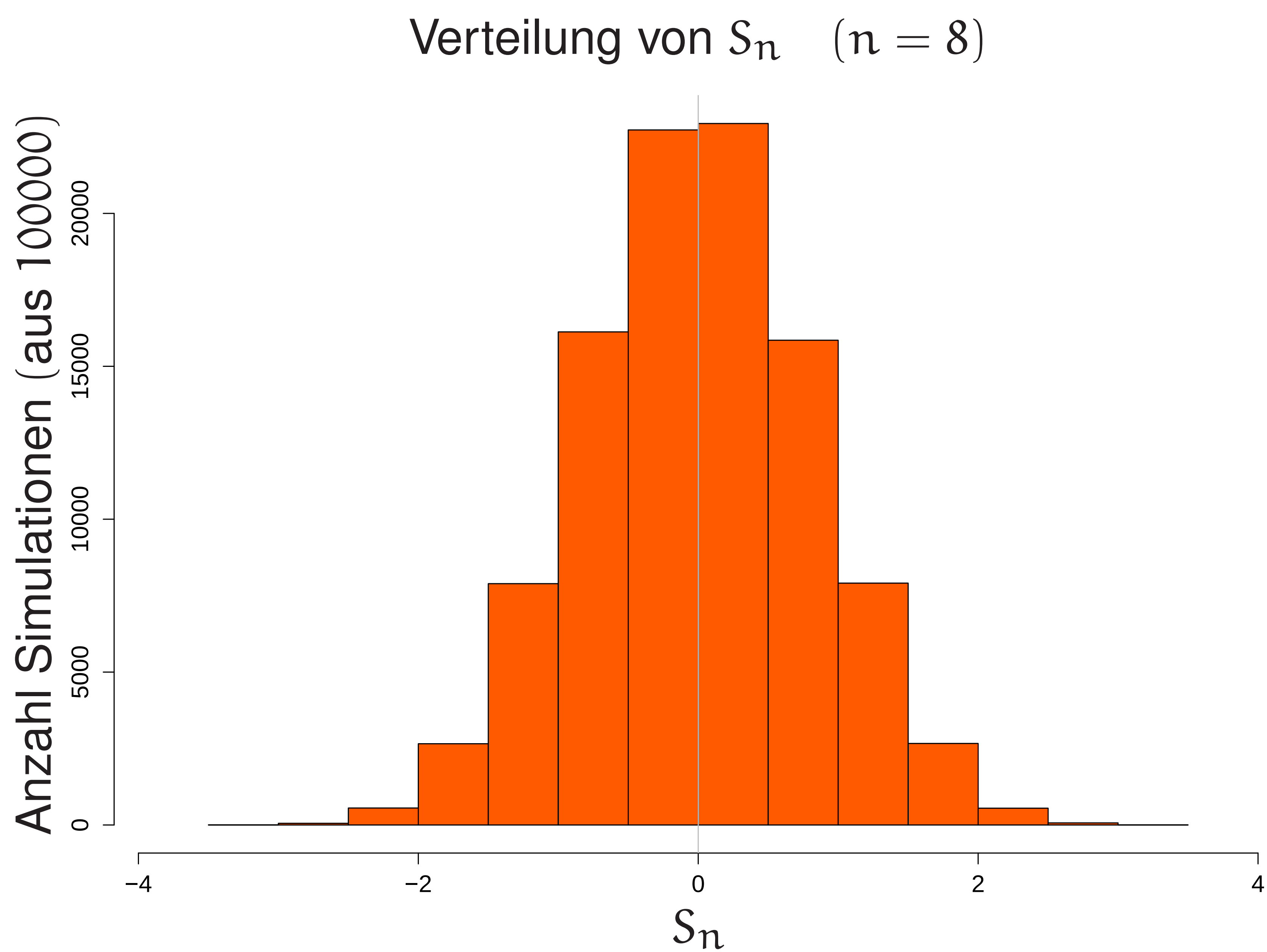
Verteilung von S_n $(n = 5)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
10000
12000
−2
−1
0
1
2
S_n



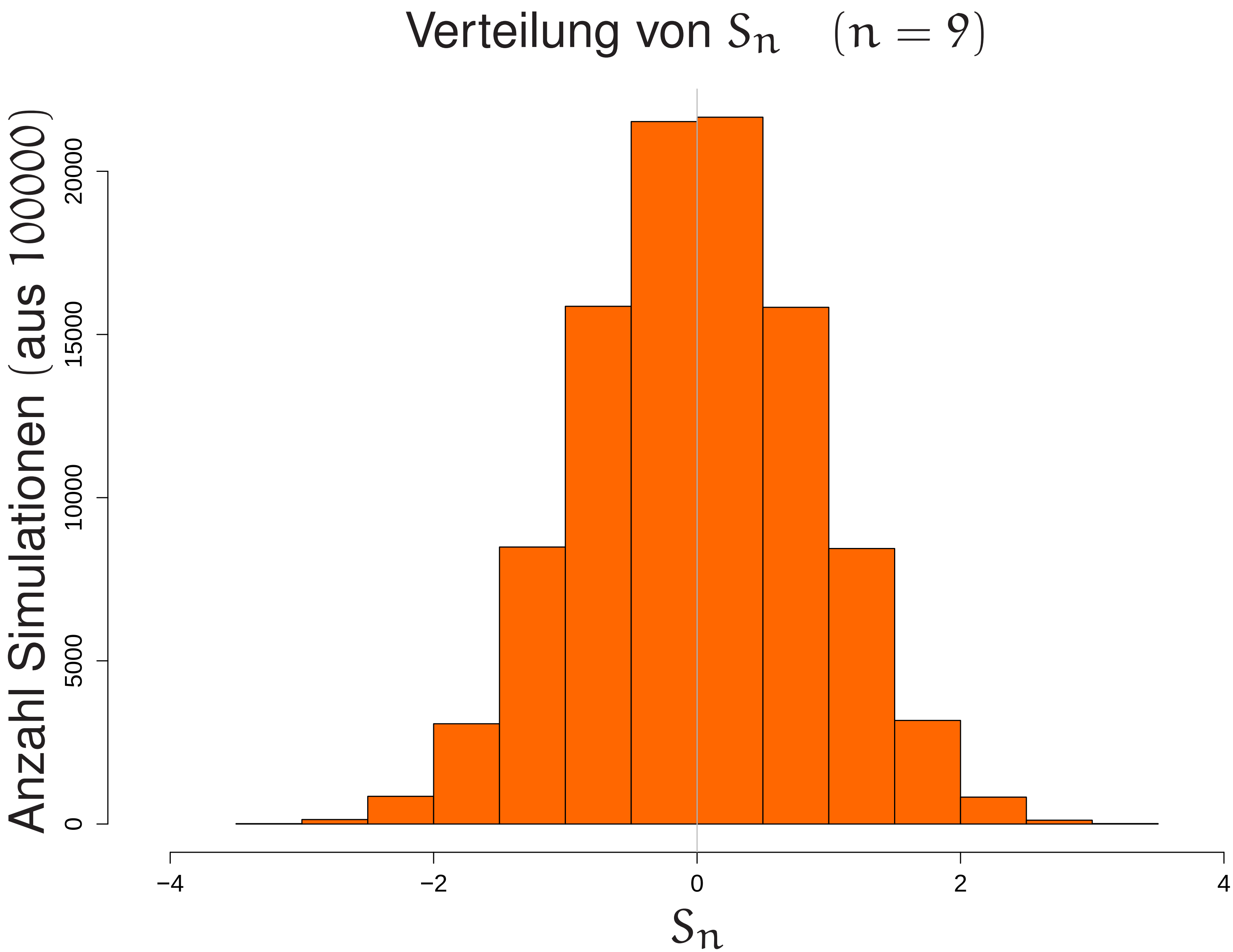
Verteilung von $S_n \quad (n = 6)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
25000
−3
−2
−1
0
1
2
3
S_n



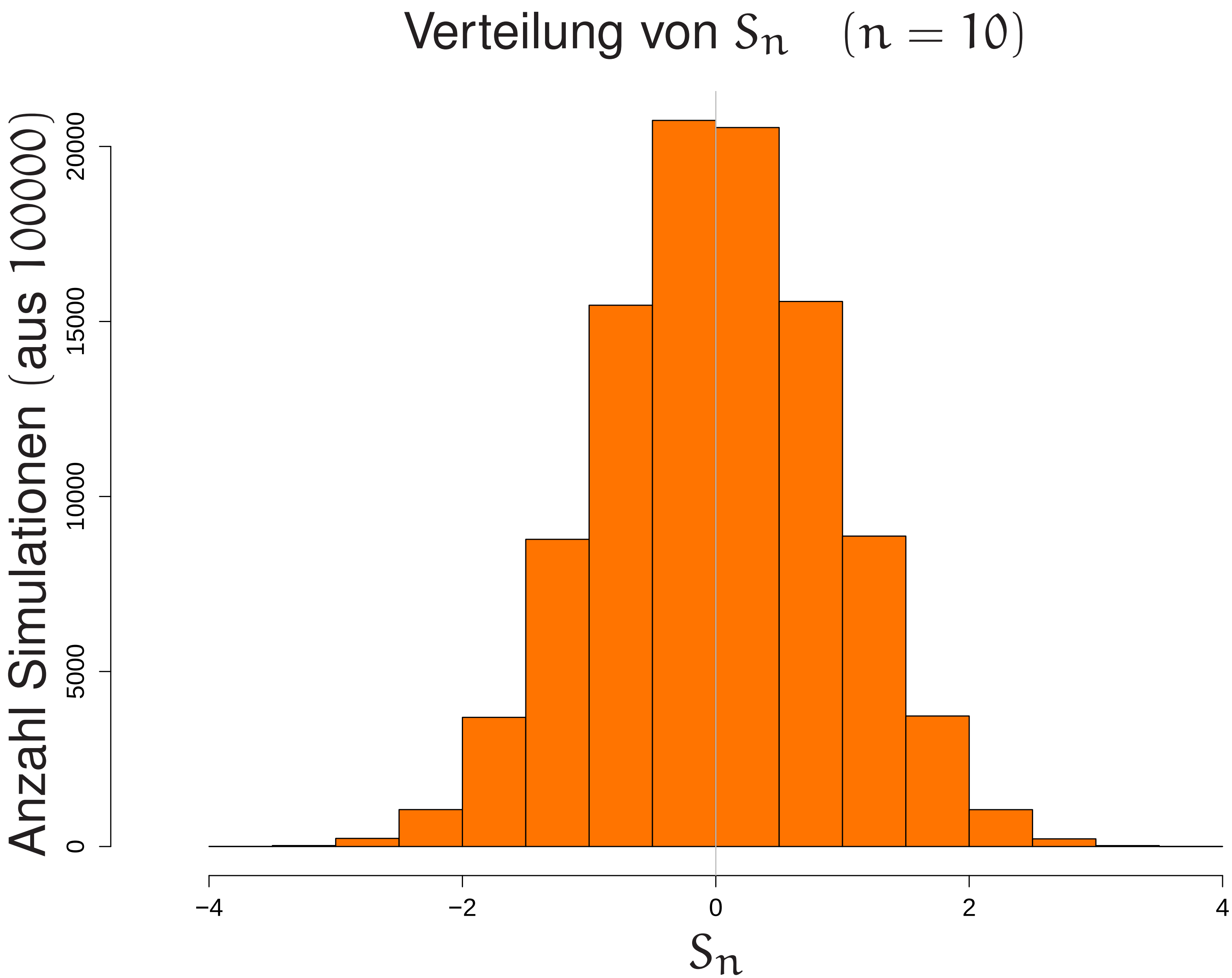
Verteilung von S_n $(n = 7)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
25000
−3
−2
−1
0
1
2
3
S_n



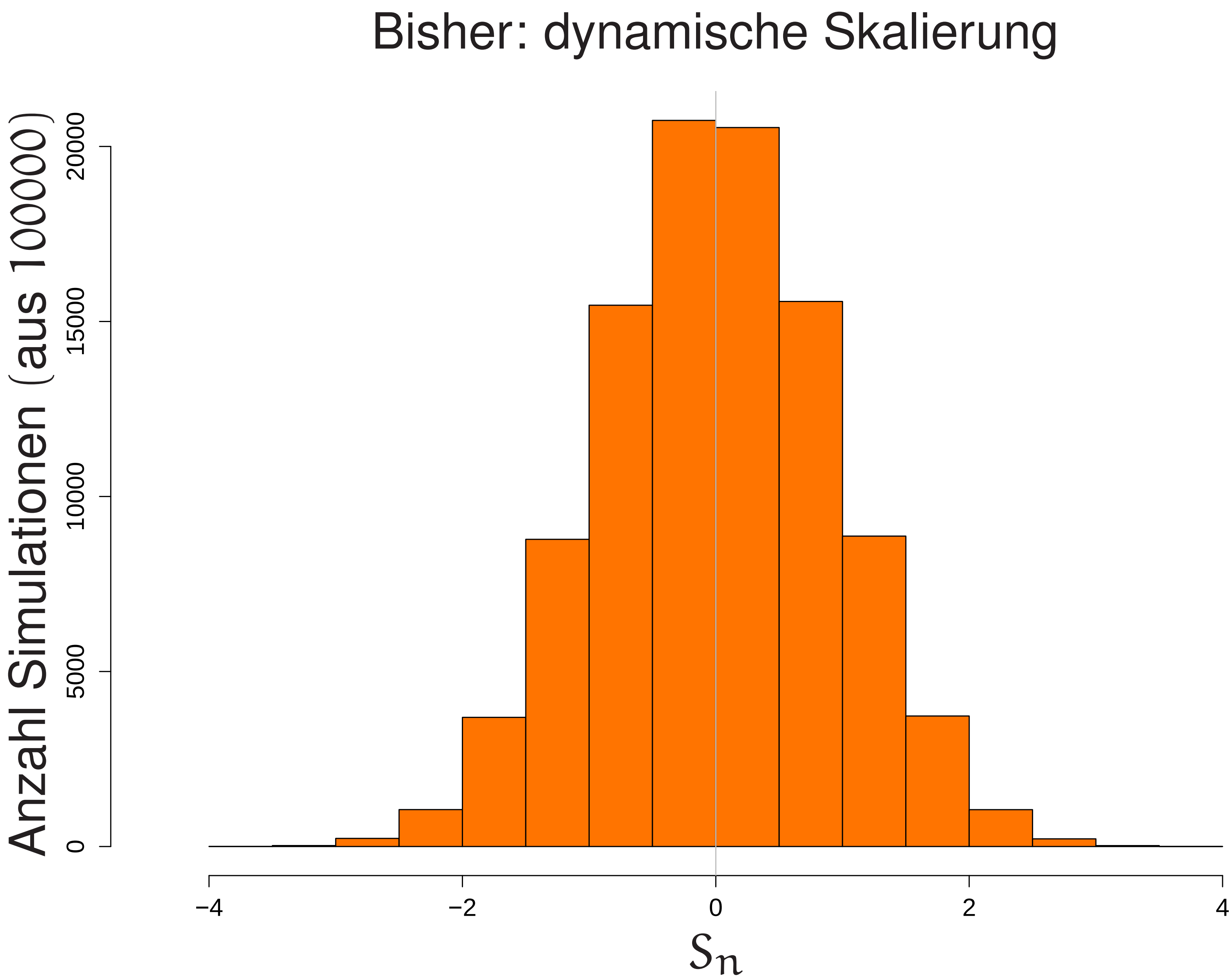
Verteilung von S_n $(n = 8)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
−4
−2
0
2
4
S_n



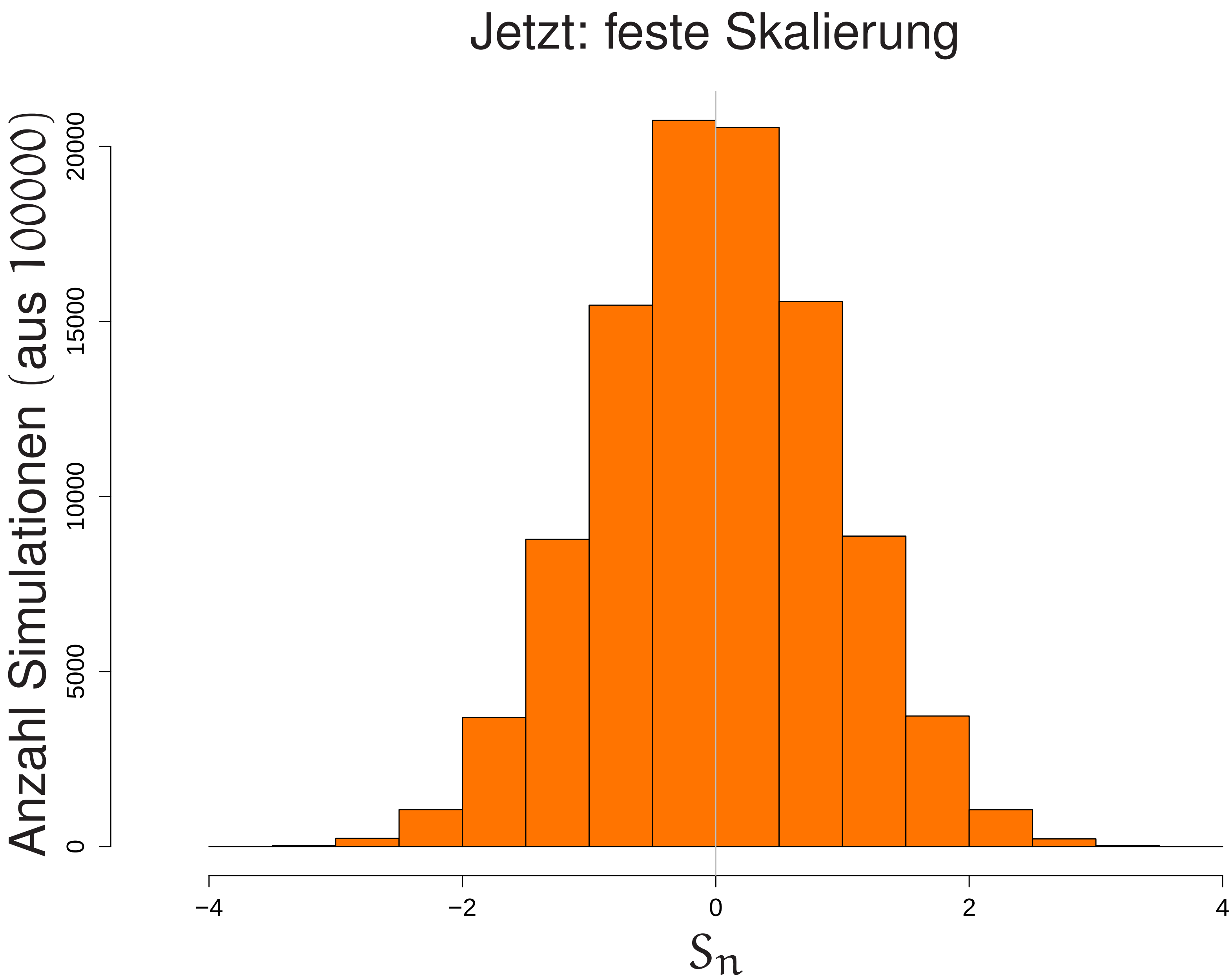
Verteilung von $S_n \quad (n = 9)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
−4
−2
0
2
4
S_n

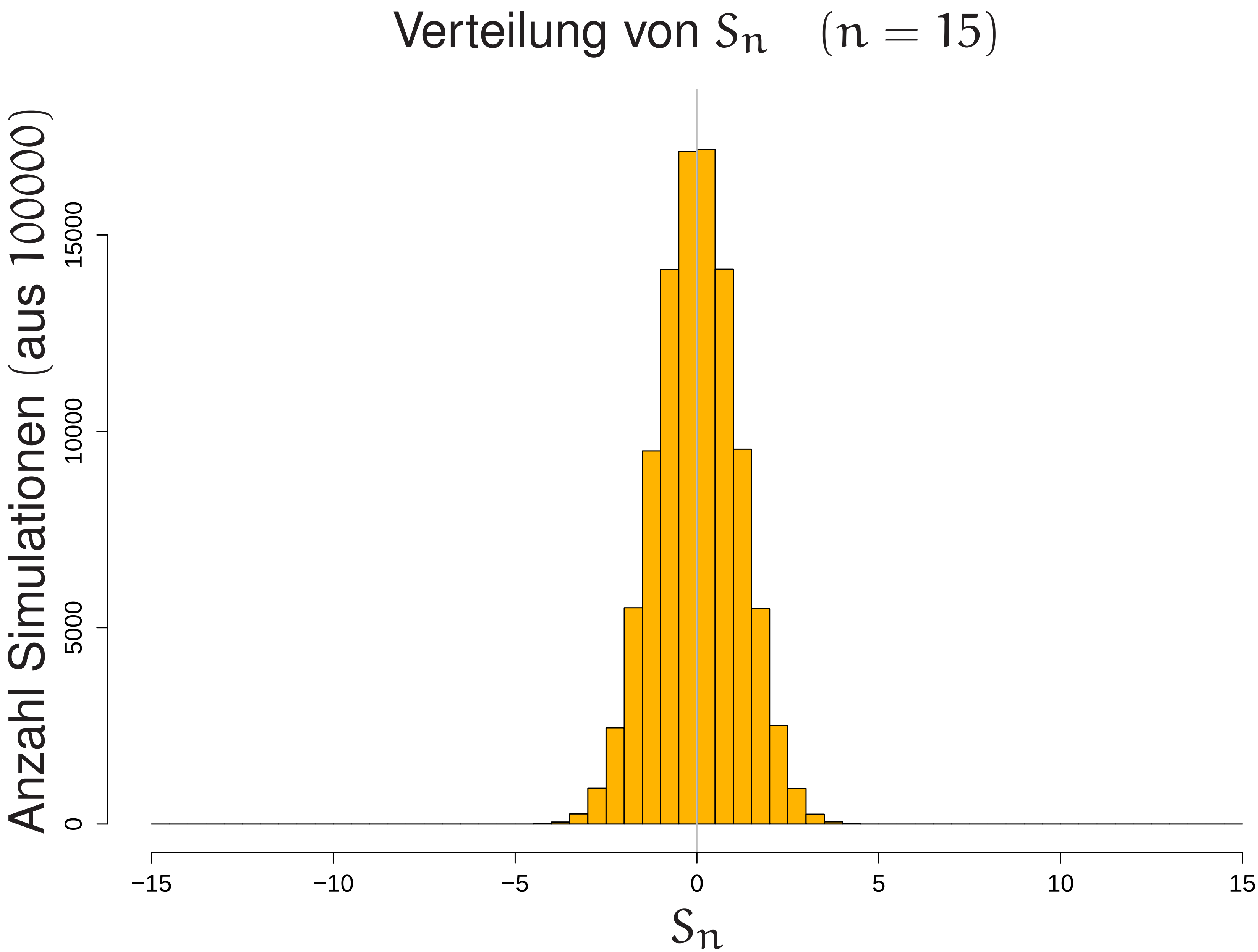


Verteilung von S_n $(n = 10)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
−4
−2
0
2
4
S_n

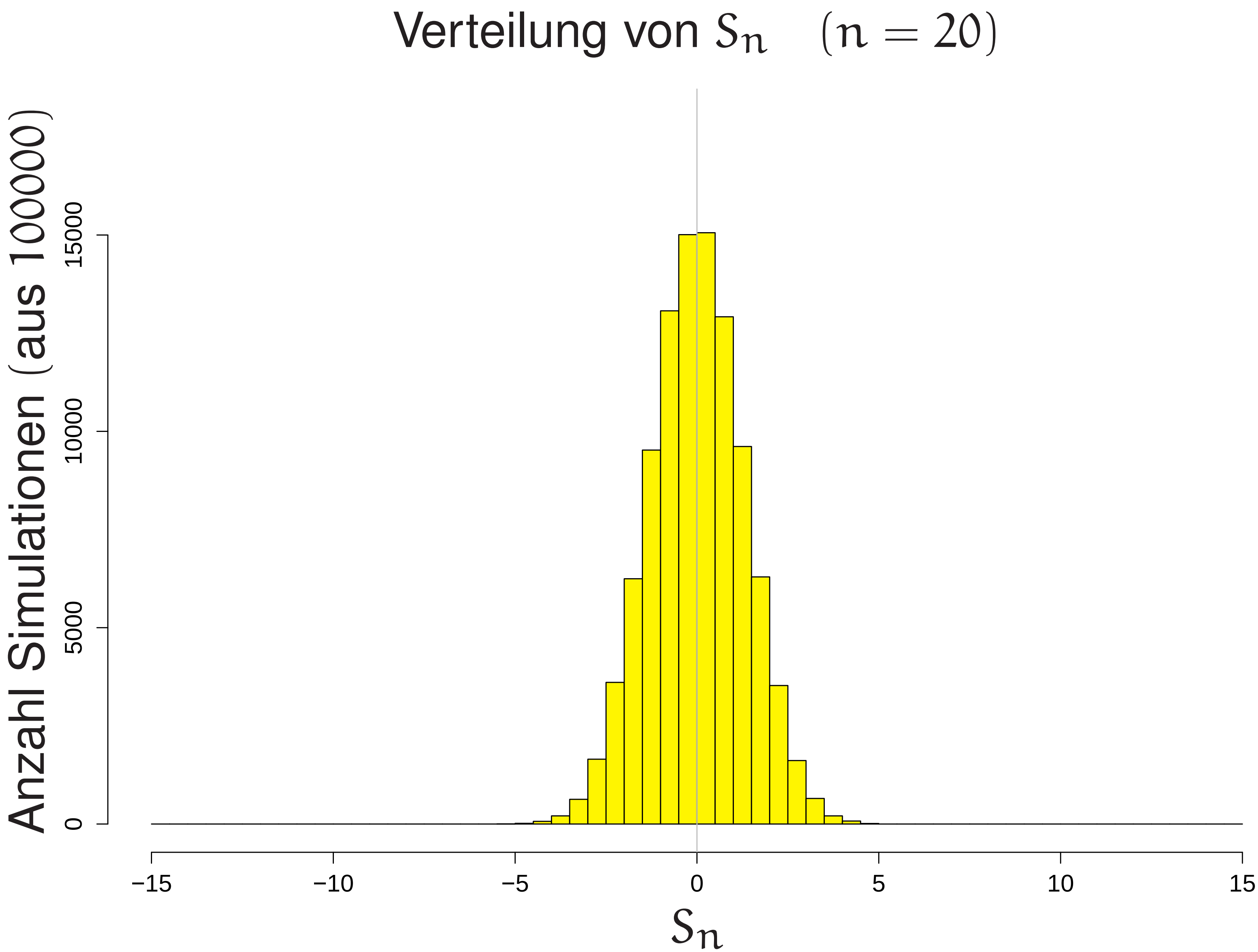


Bisher: dynamische Skalierung
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
−4
−2
0
2
4
S_n

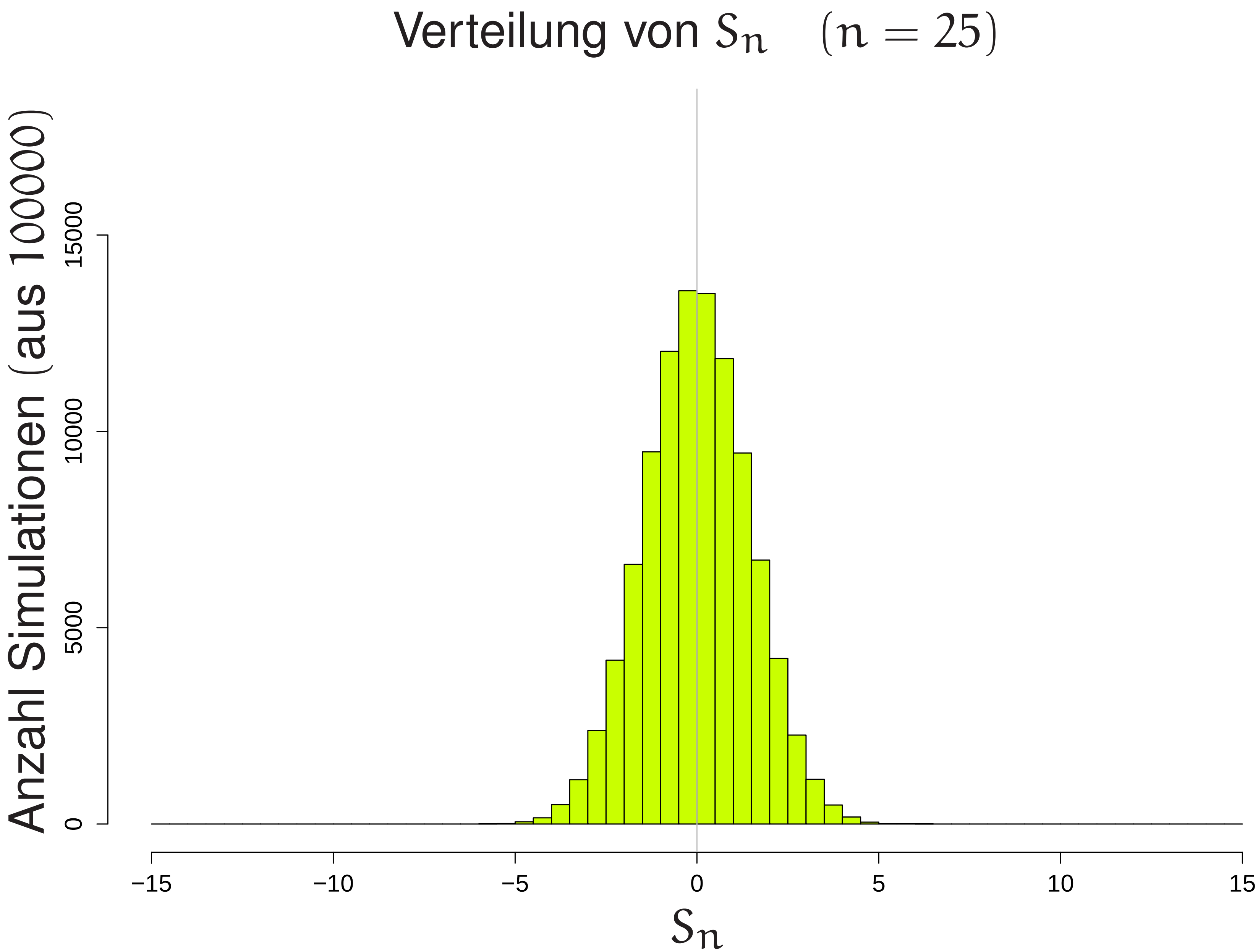

Jetzt: feste Skalierung
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
20000
−4
−2
0
2
4
S_n



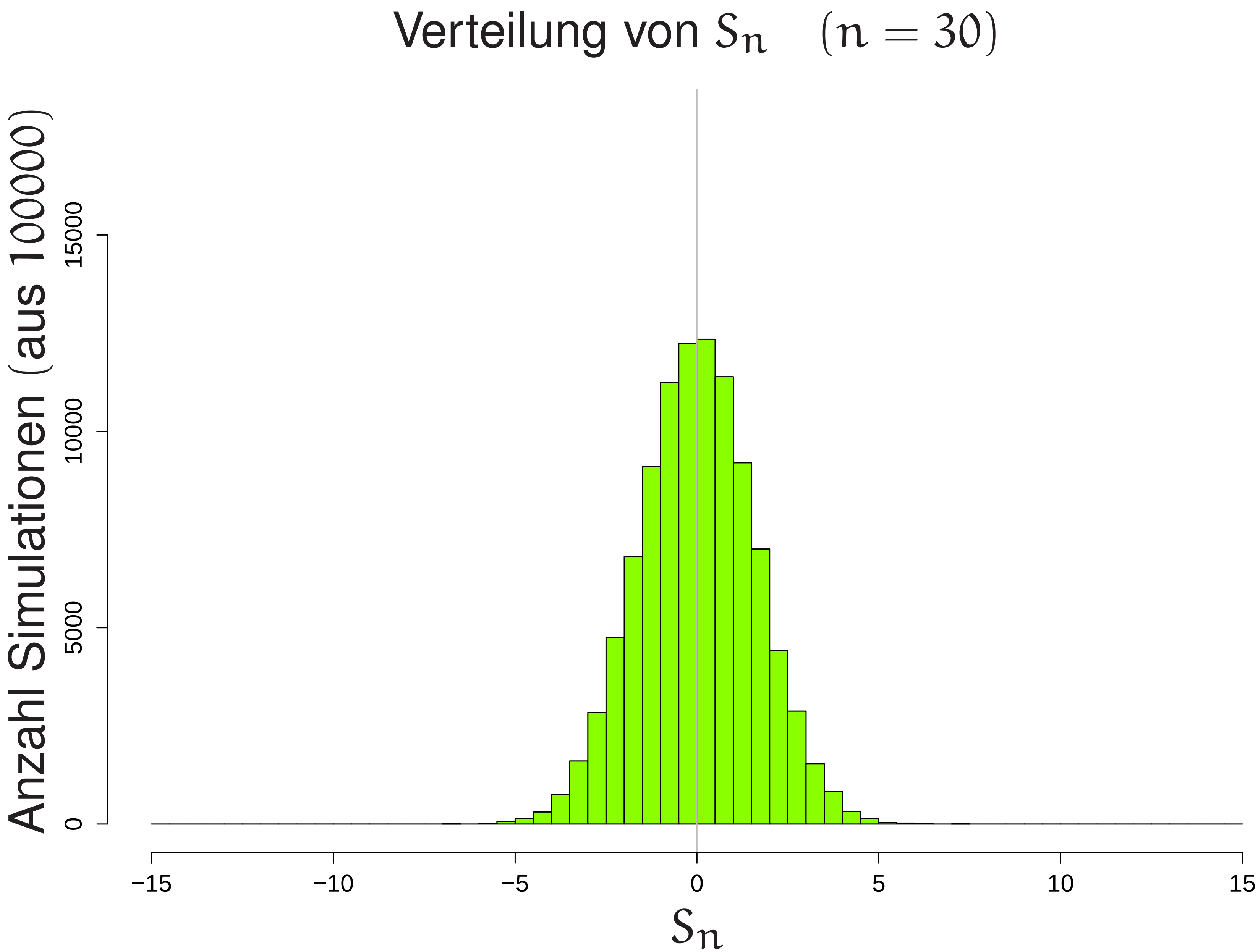
Verteilung von S_n $(n = 15)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n



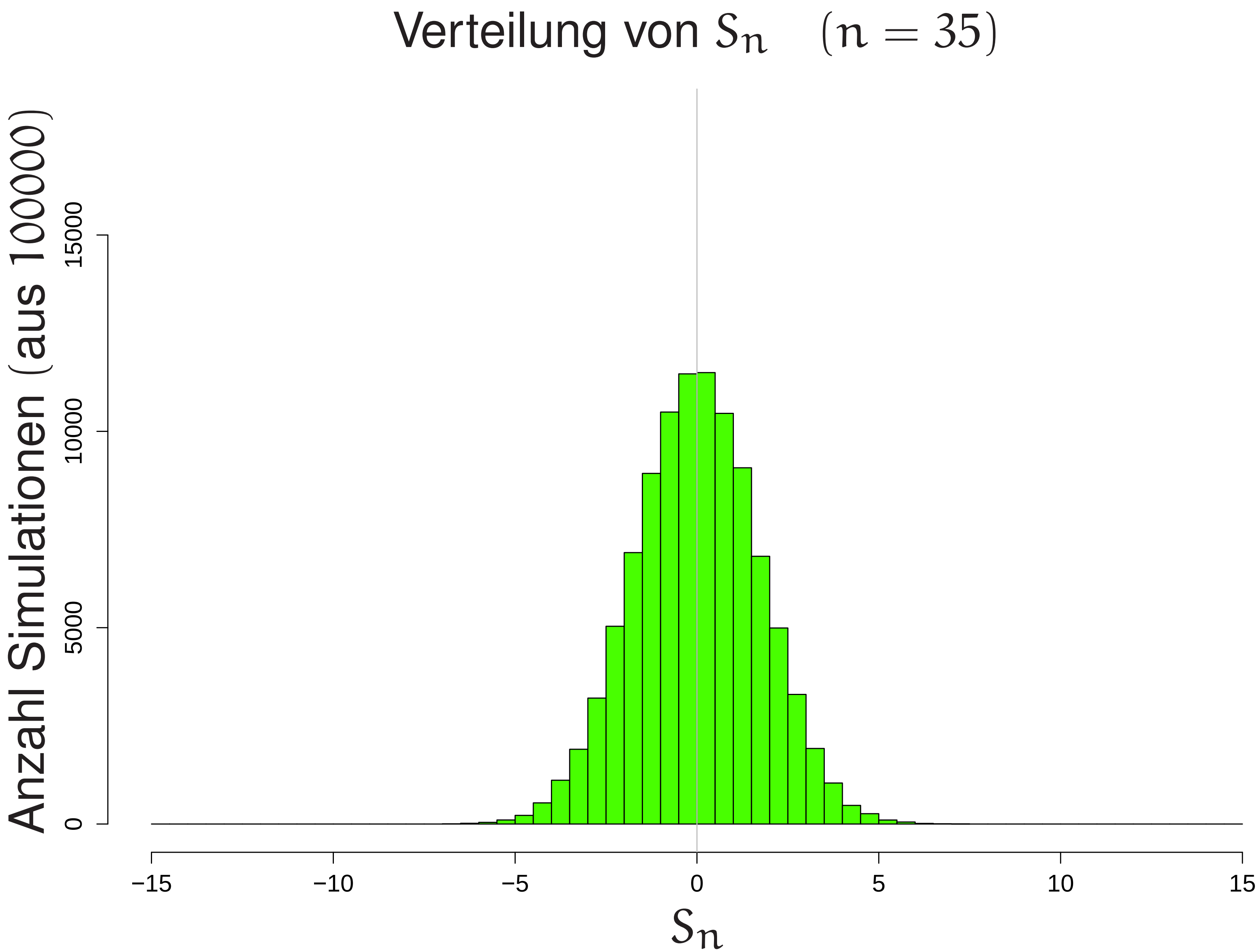
Verteilung von S_n $(n = 20)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n



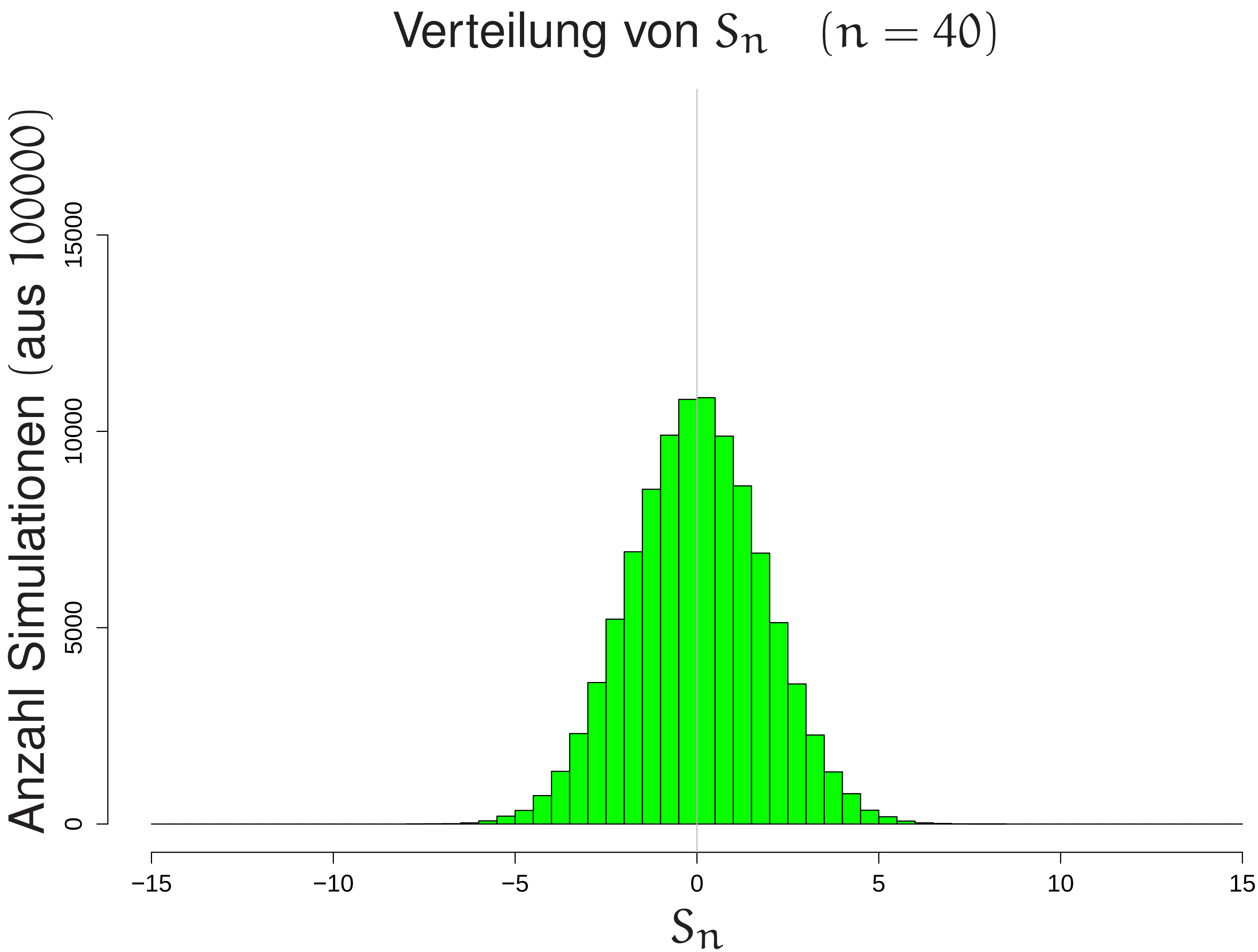
Verteilung von S_n $(n = 25)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

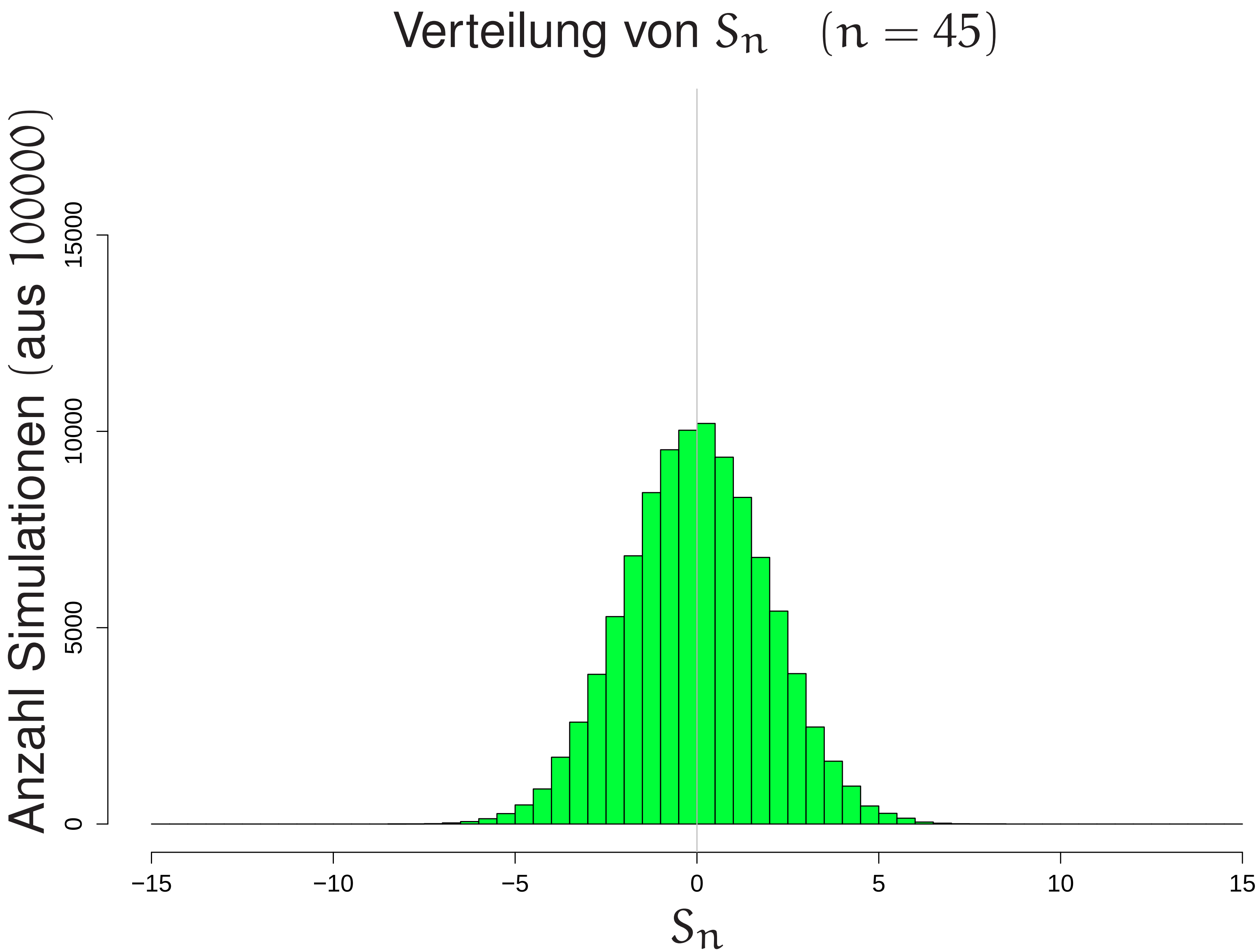


Verteilung von S_n $(n = 30)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

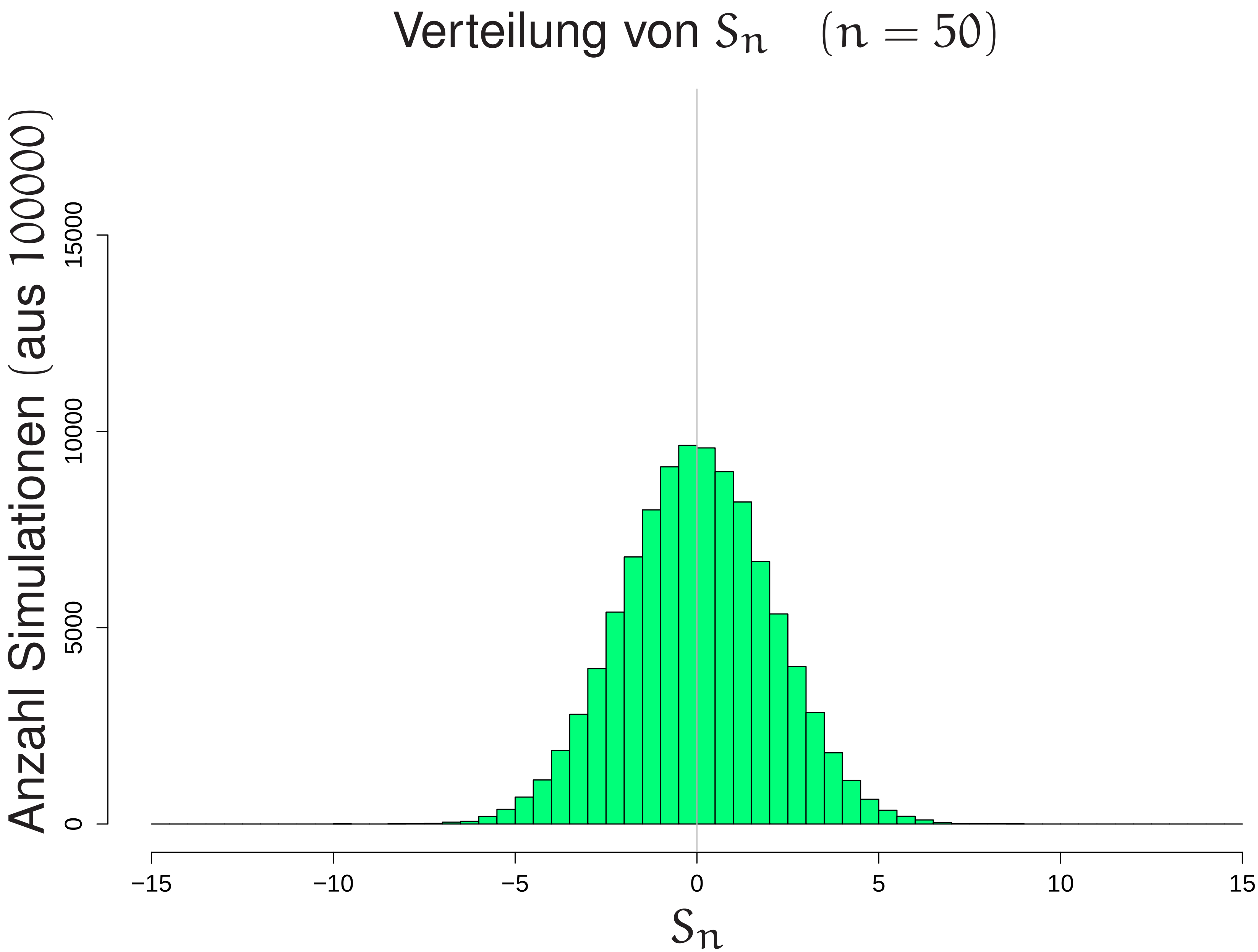


Verteilung von S_n $(n = 35)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

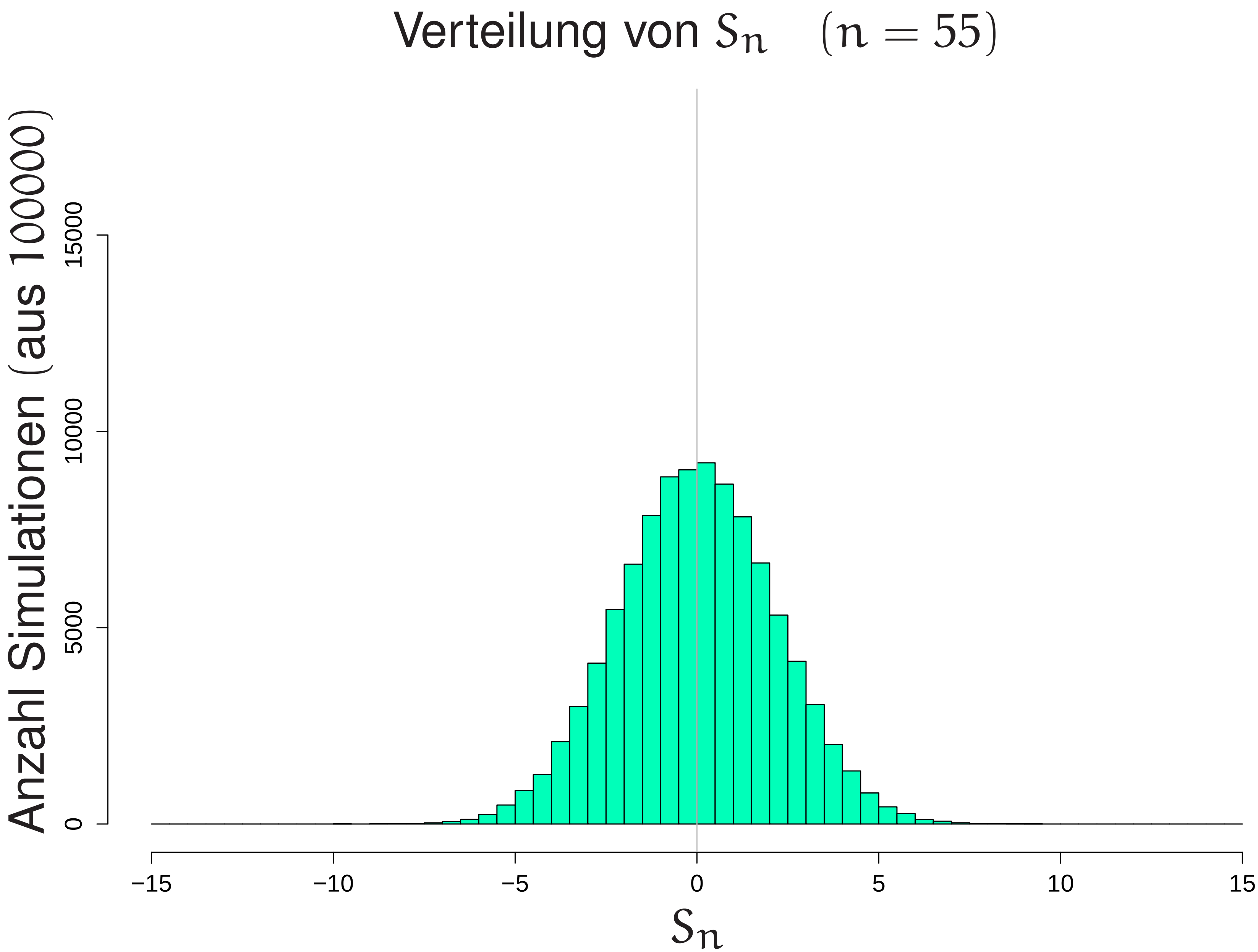

Verteilung von S_n $(n = 40)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n



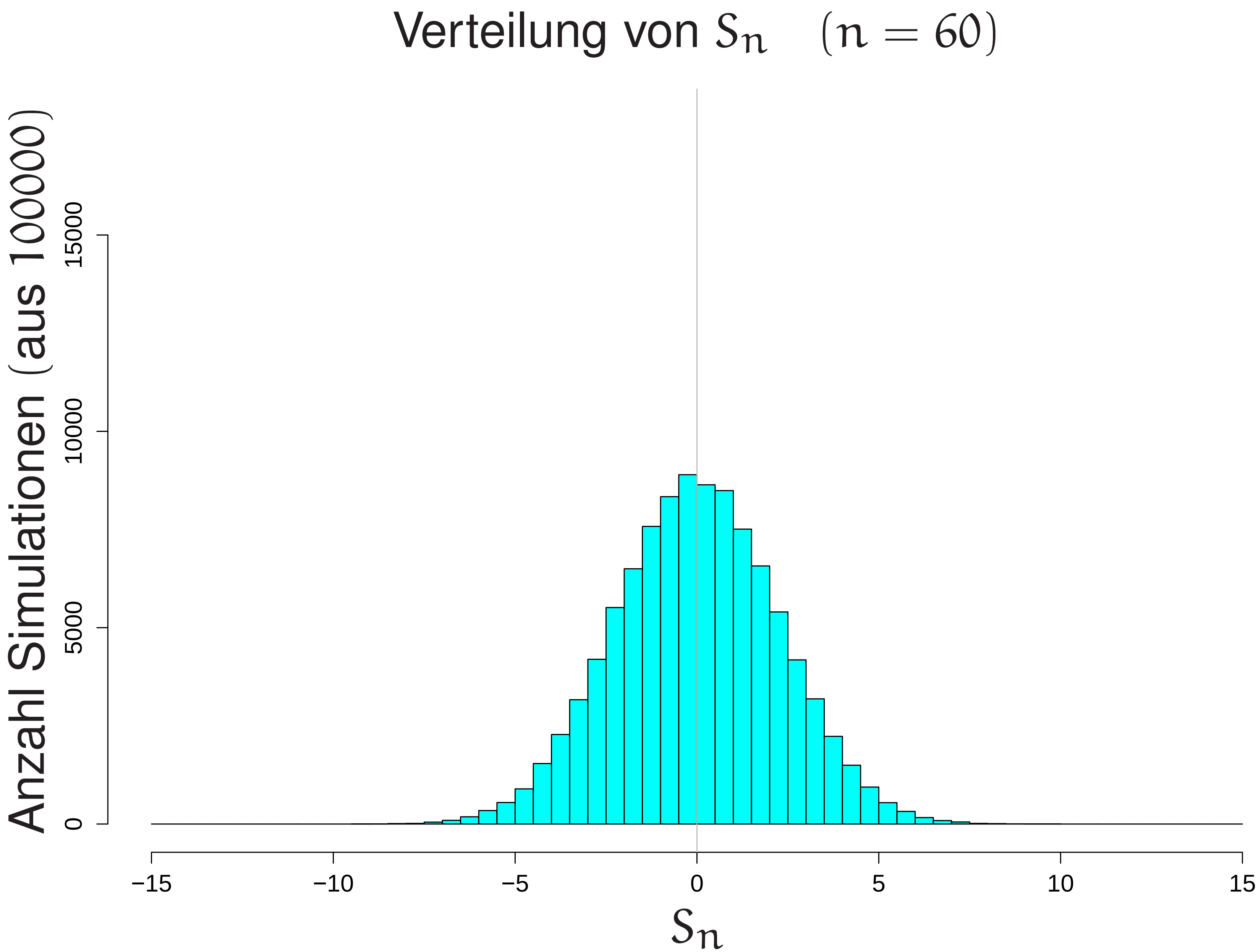
Verteilung von S_n $(n = 45)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n



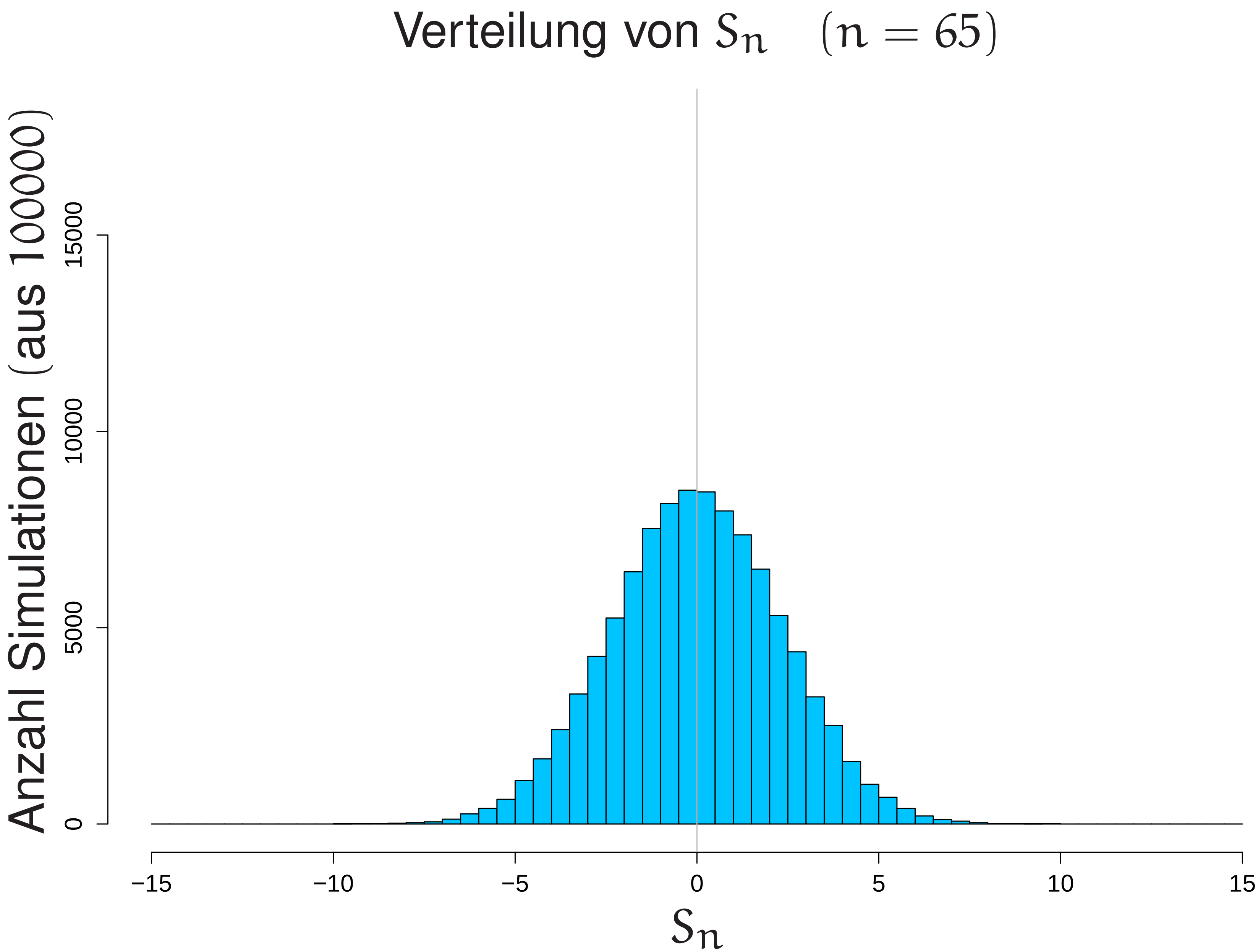
Verteilung von S_n $(n = 50)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n



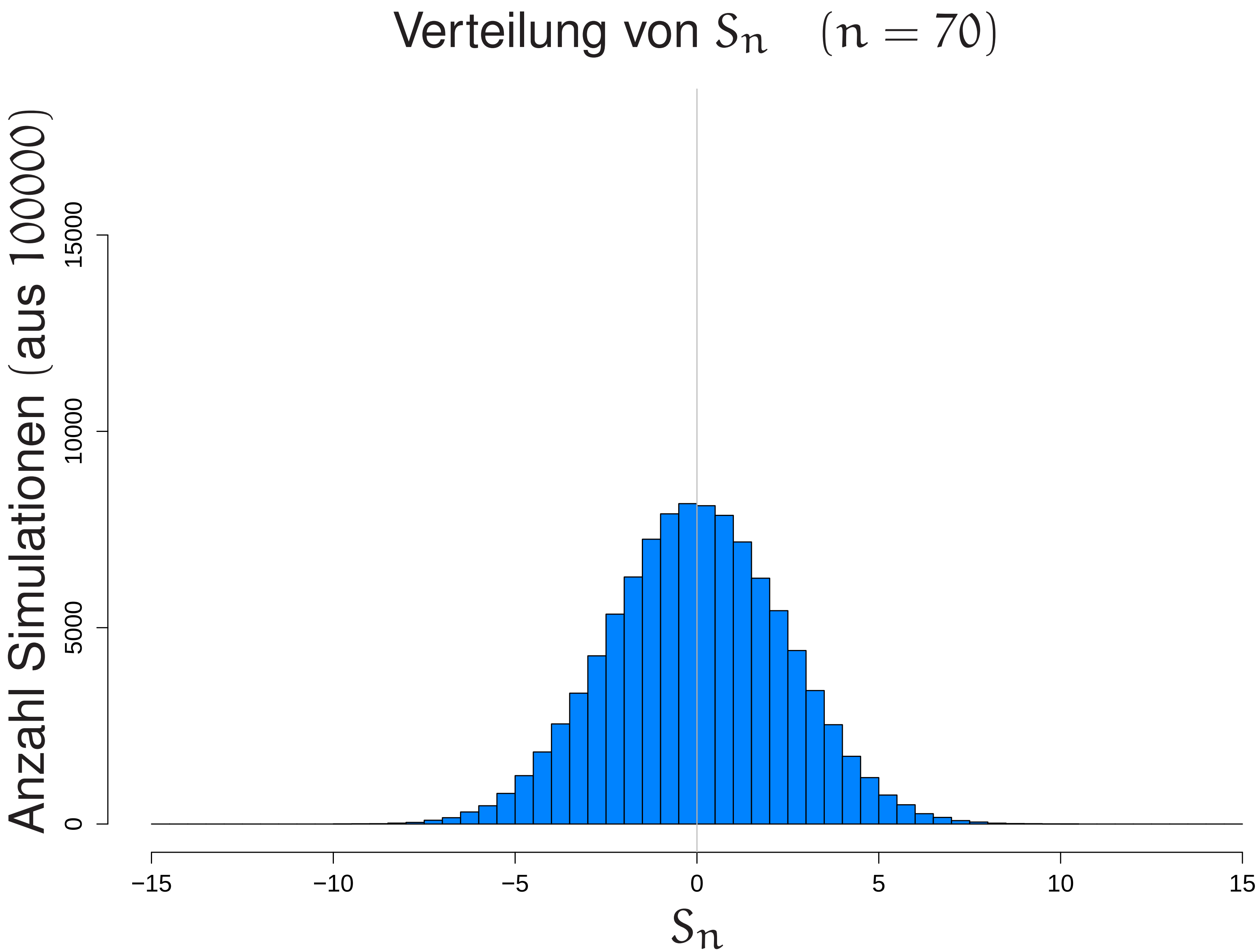
Verteilung von S_n $(n = 55)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

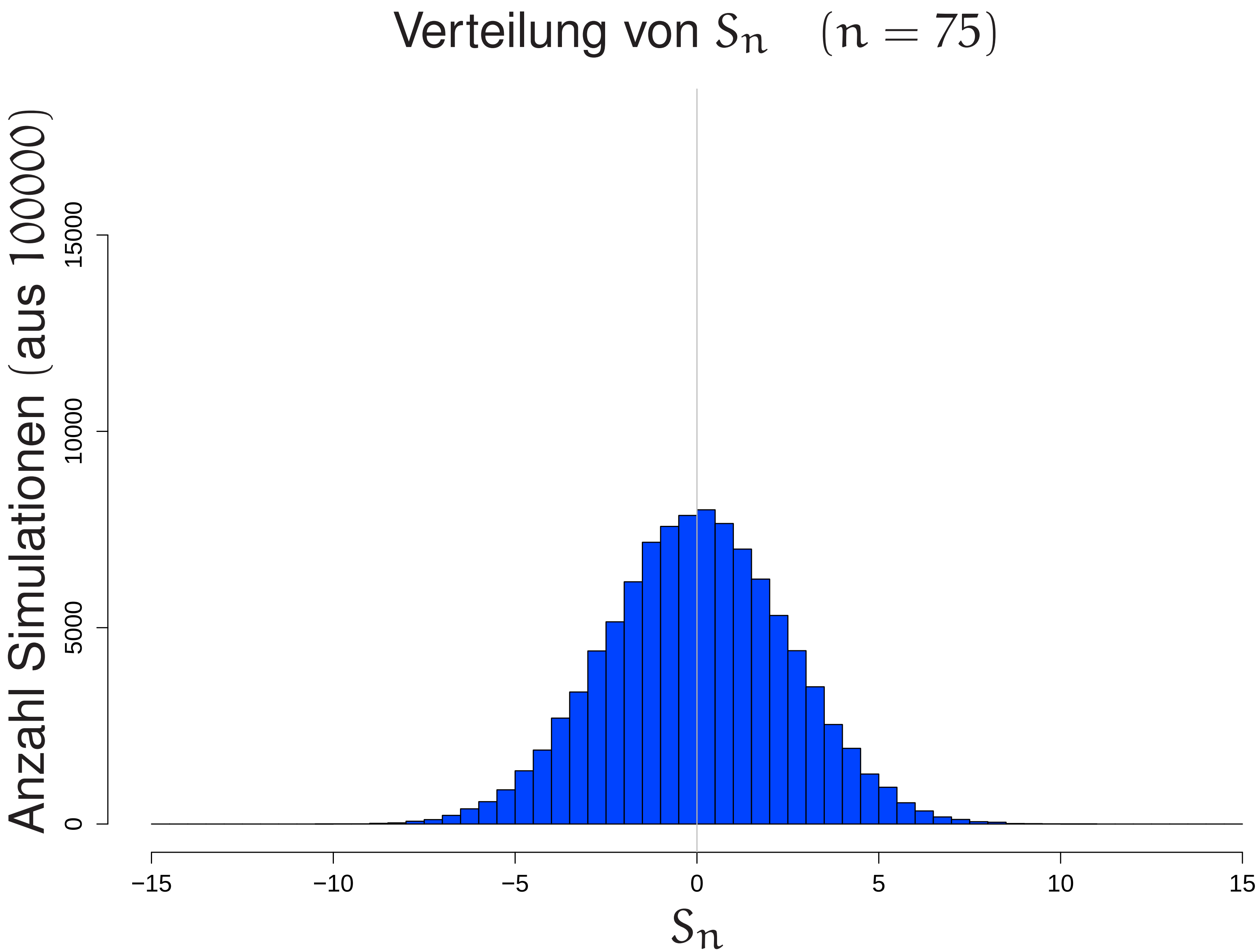


Verteilung von S_n $(n = 60)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

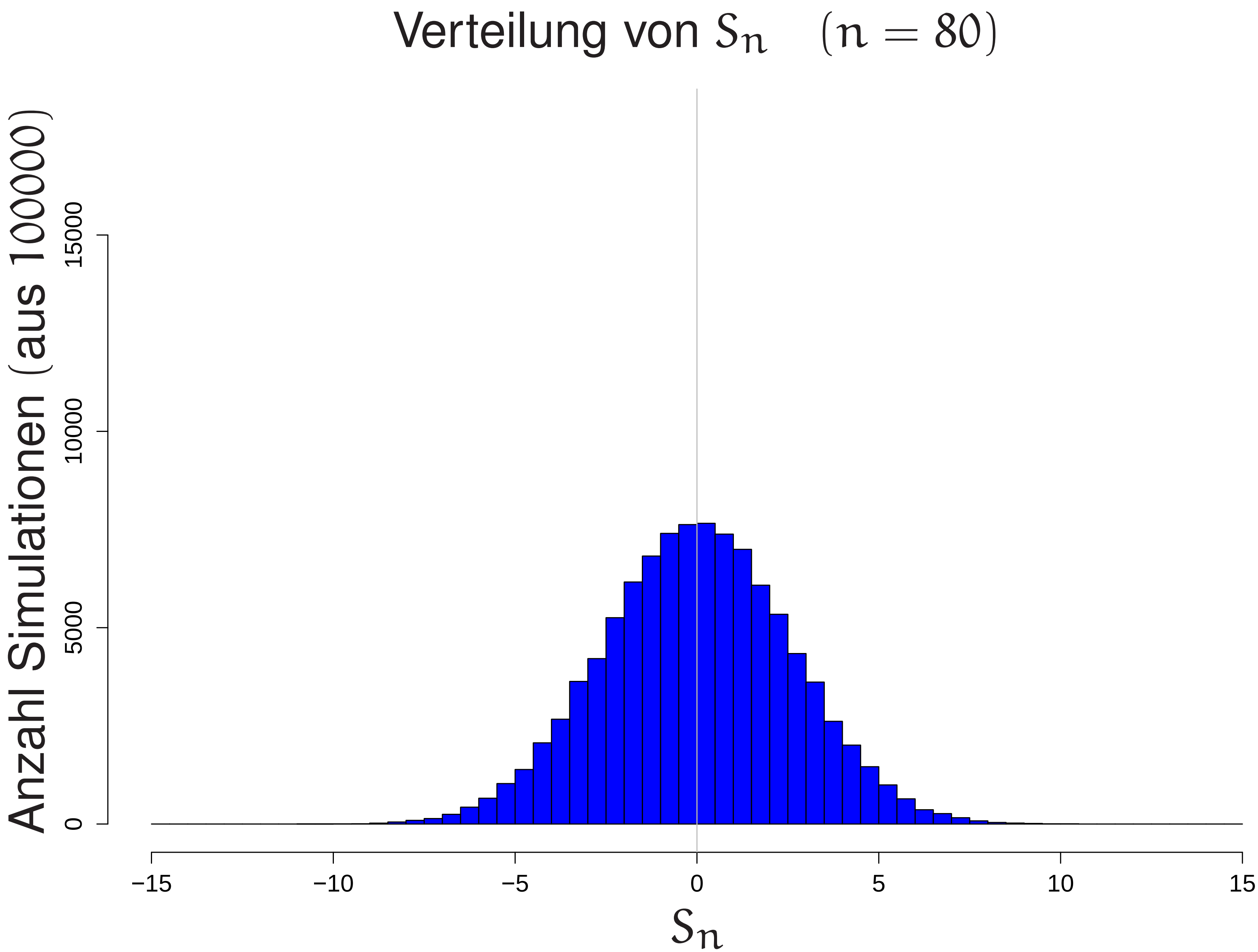


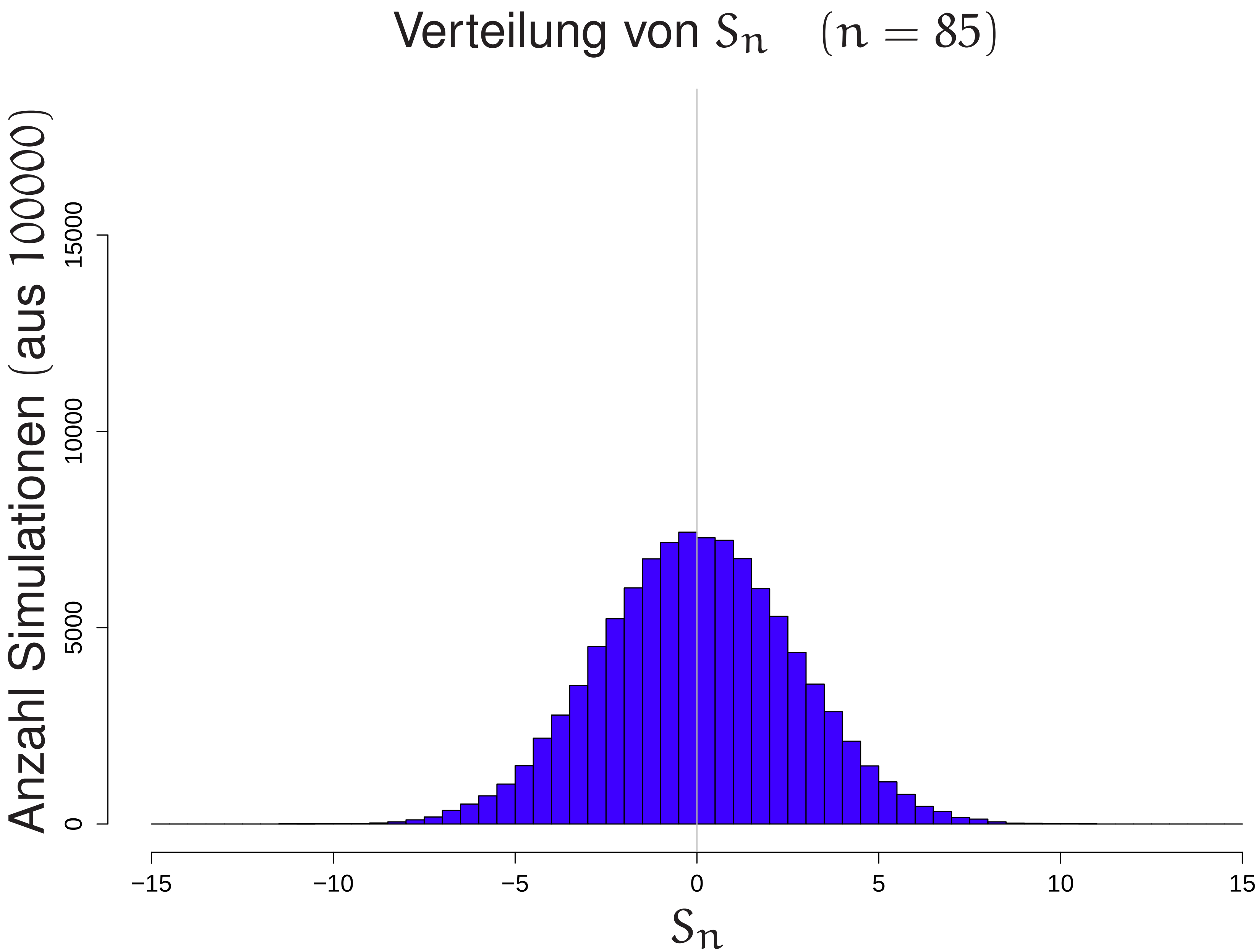
Verteilung von S_n $(n = 65)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

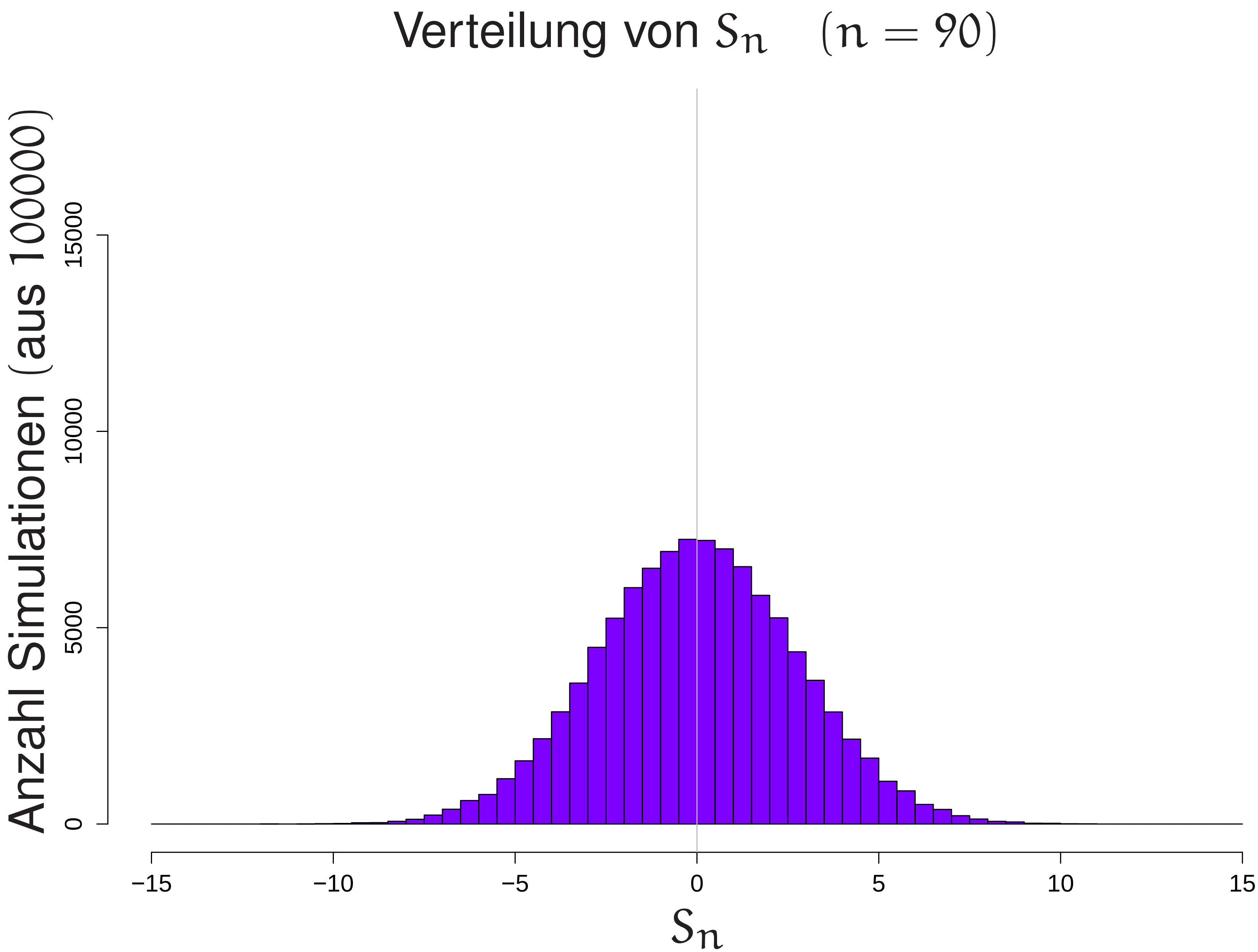

Verteilung von S_n $(n = 70)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

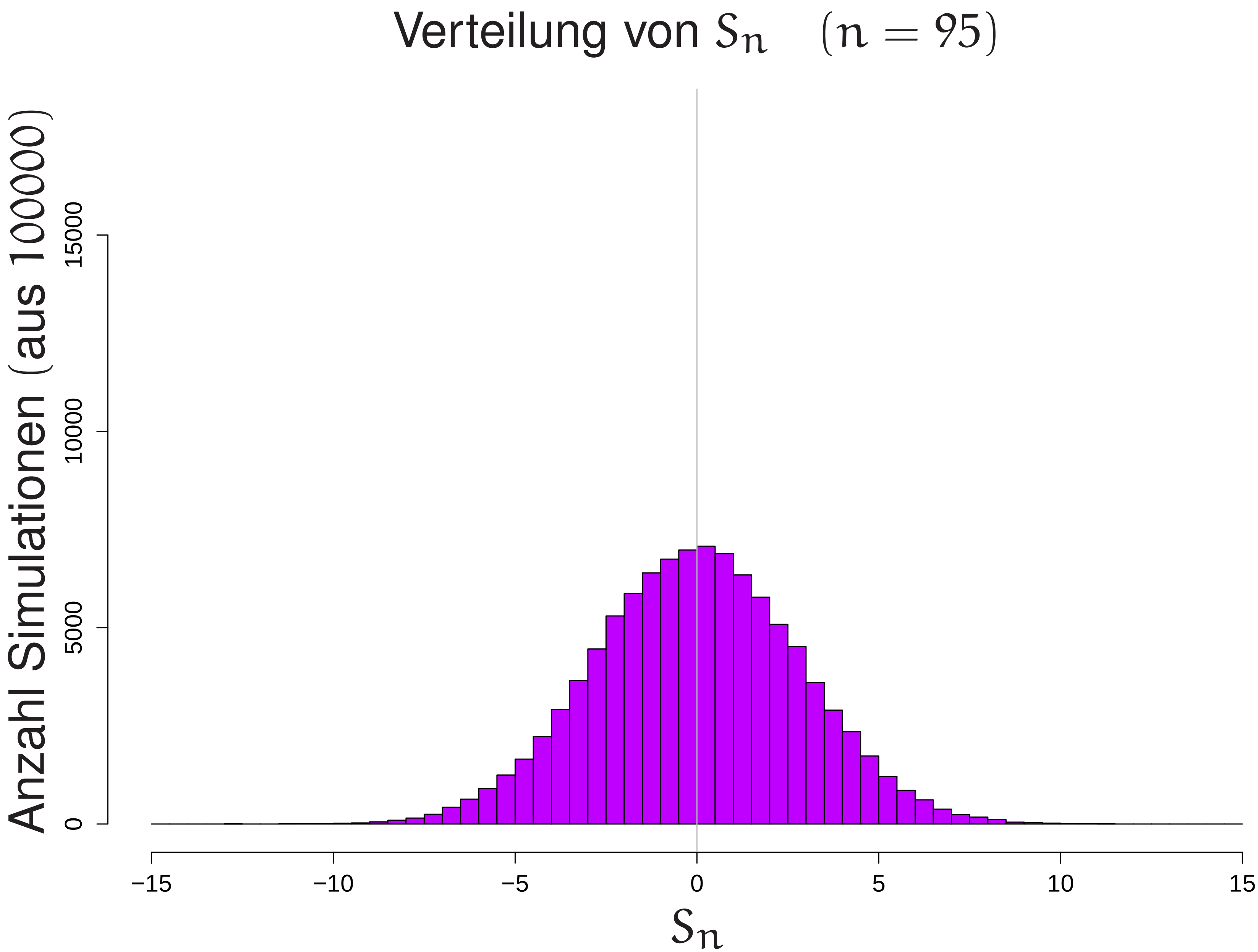


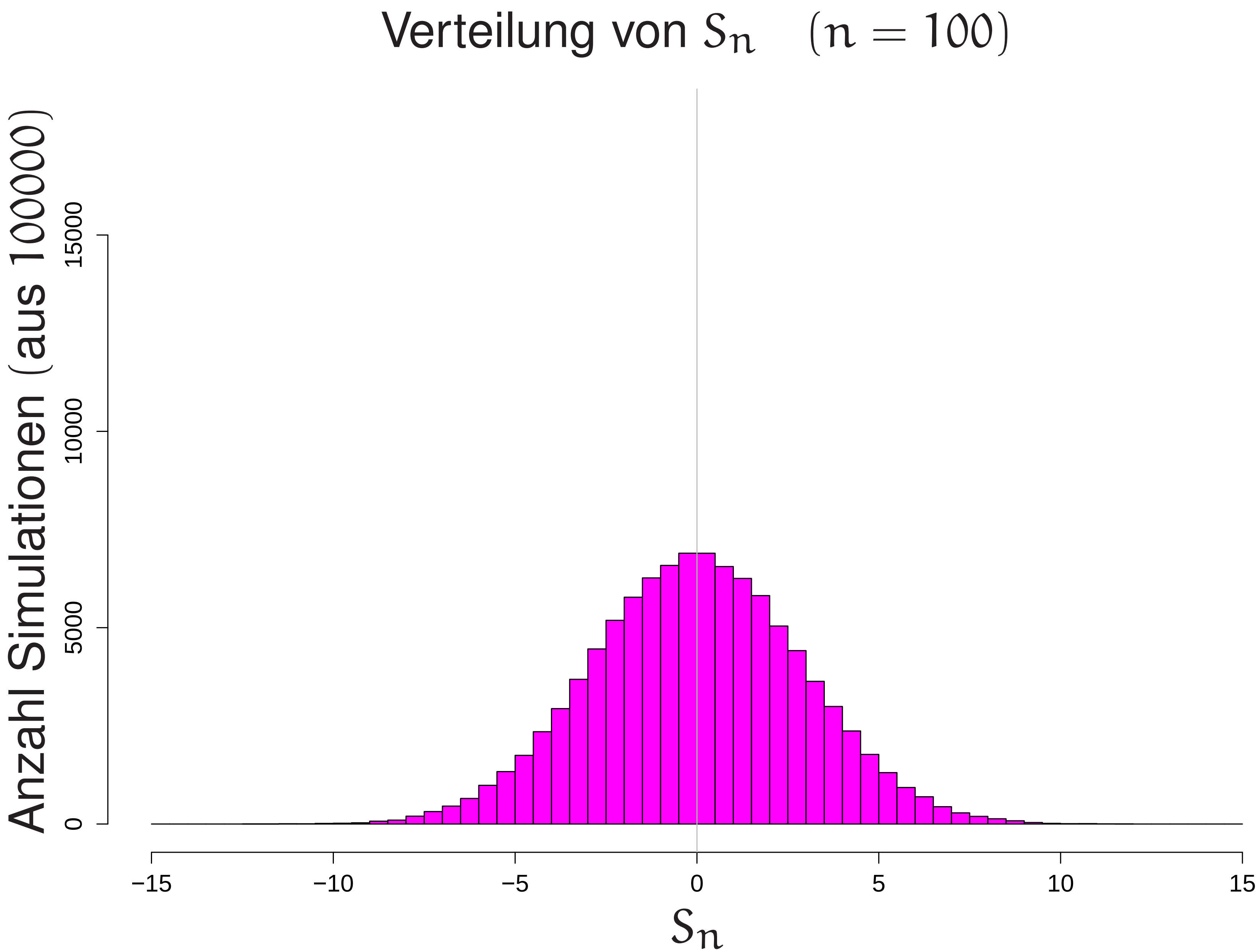
Verteilung von S_n $(n = 75)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n


Verteilung von S_n $(n = 80)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

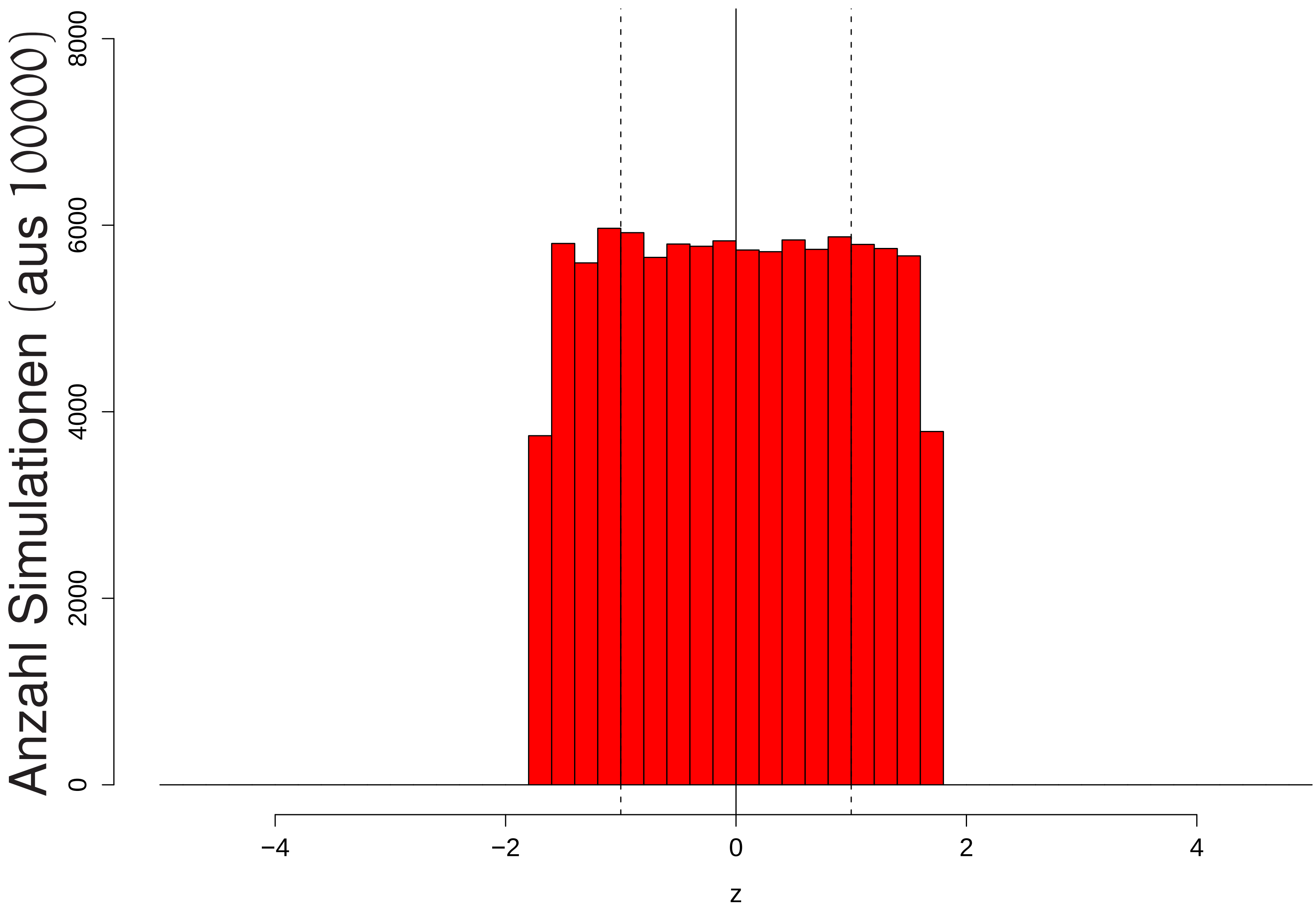

Verteilung von S_n $(n = 85)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

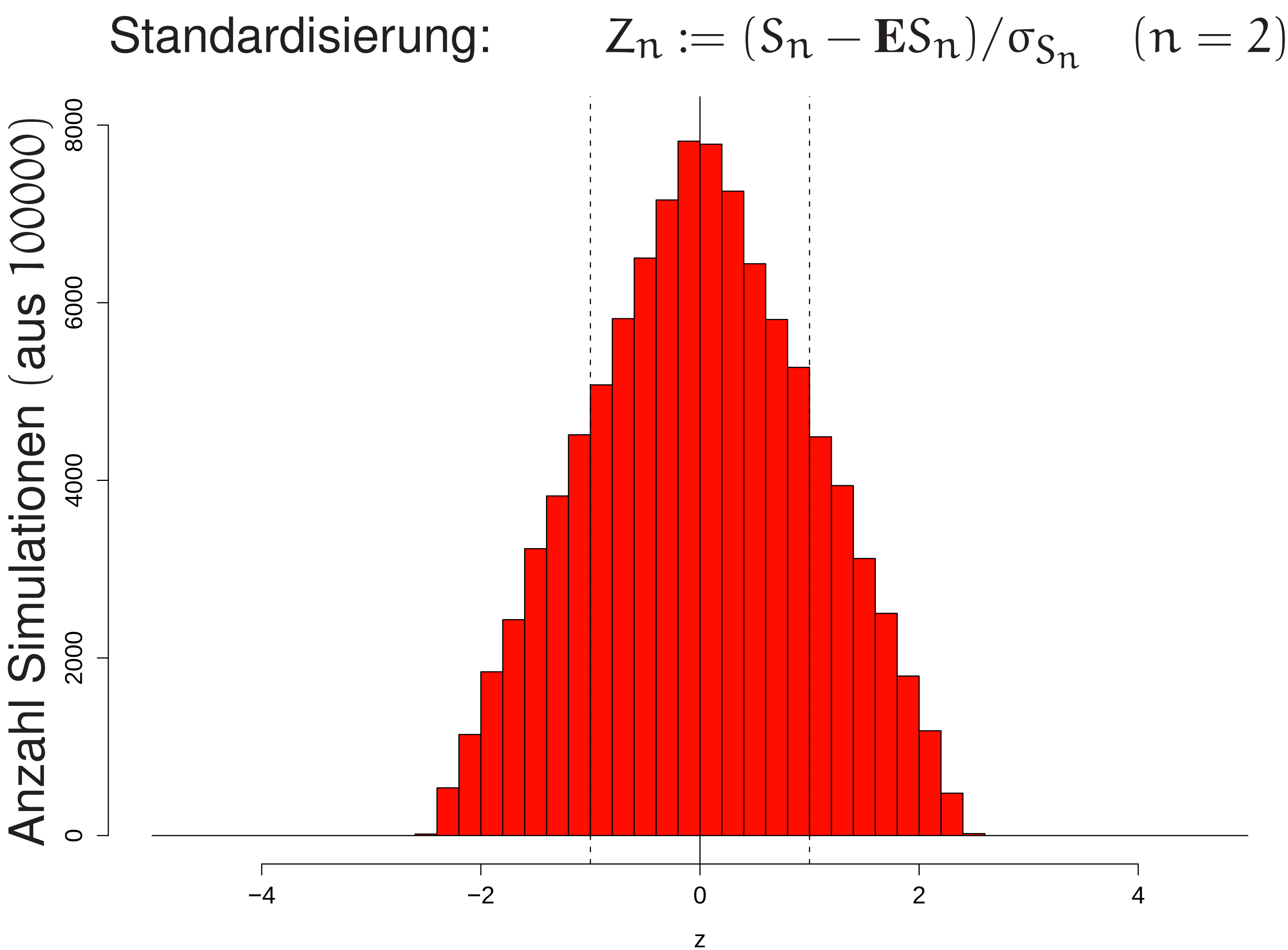

Verteilung von S_n $(n = 90)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

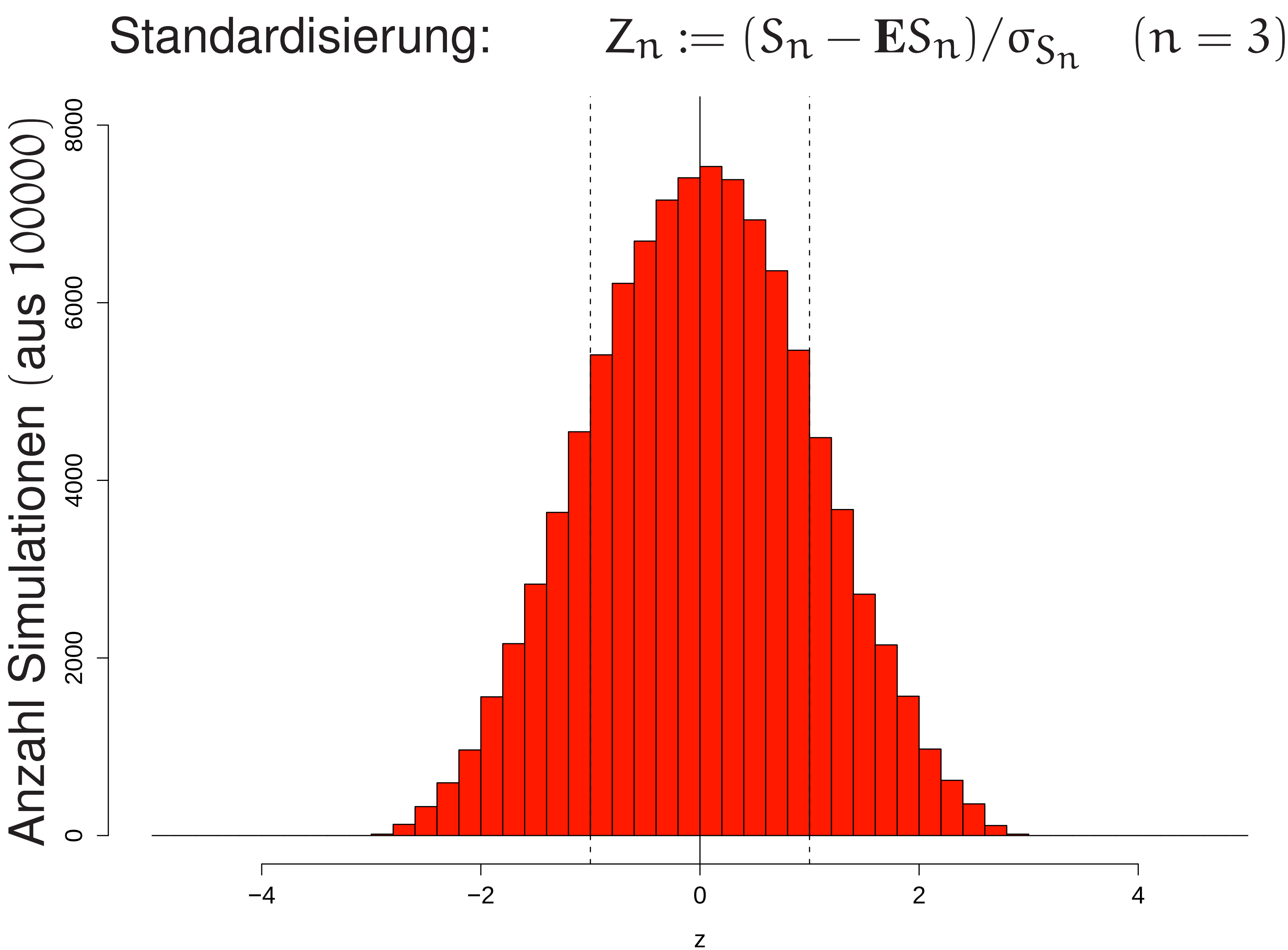

Verteilung von S_n $(n = 95)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n


Verteilung von S_n $(n = 100)$
Anzahl Simulationen (aus 100000)
0
5000
10000
15000
−15
−10
−5
0
5
10
15
S_n

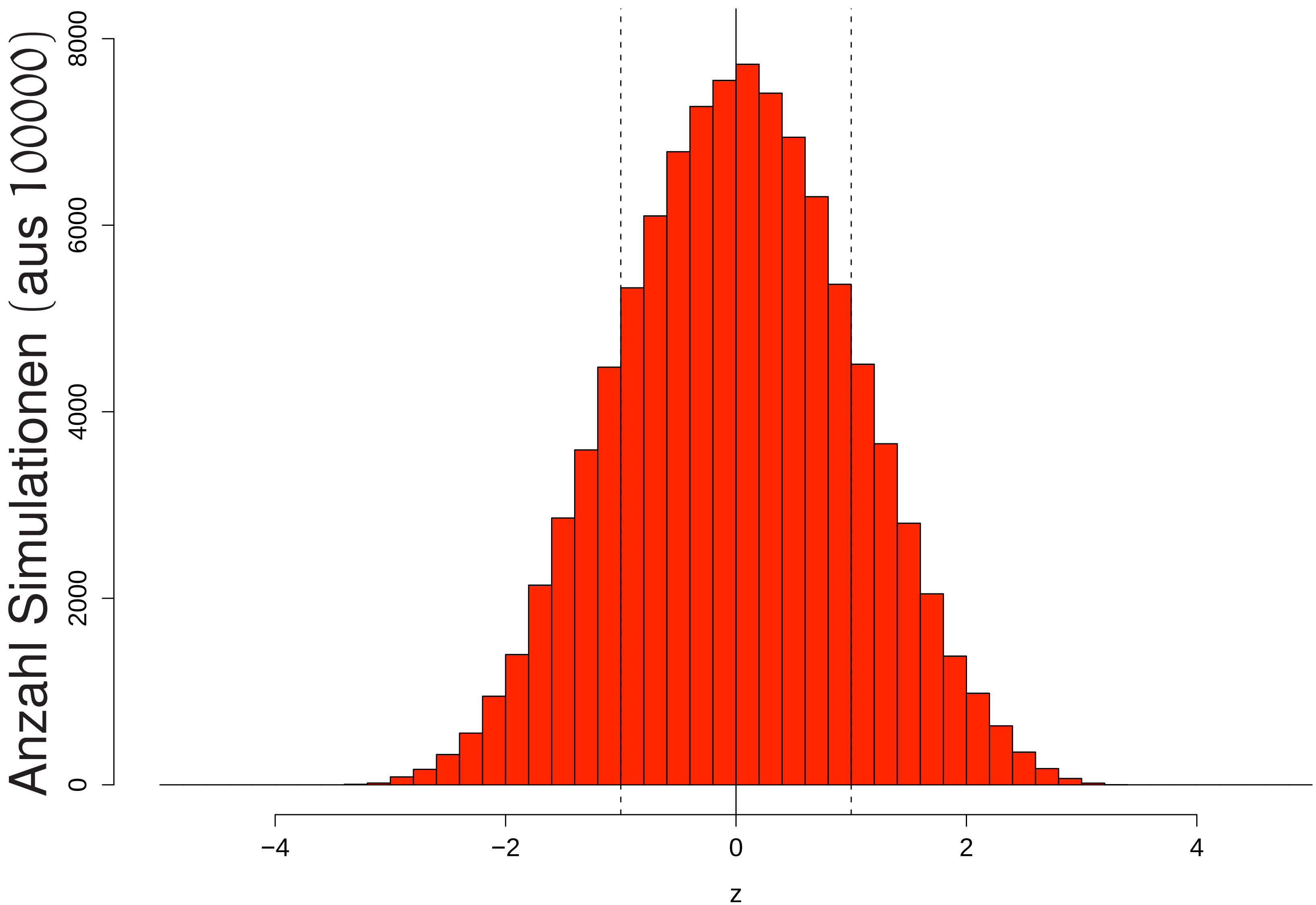
Standardisierung: $\quad Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 1)$

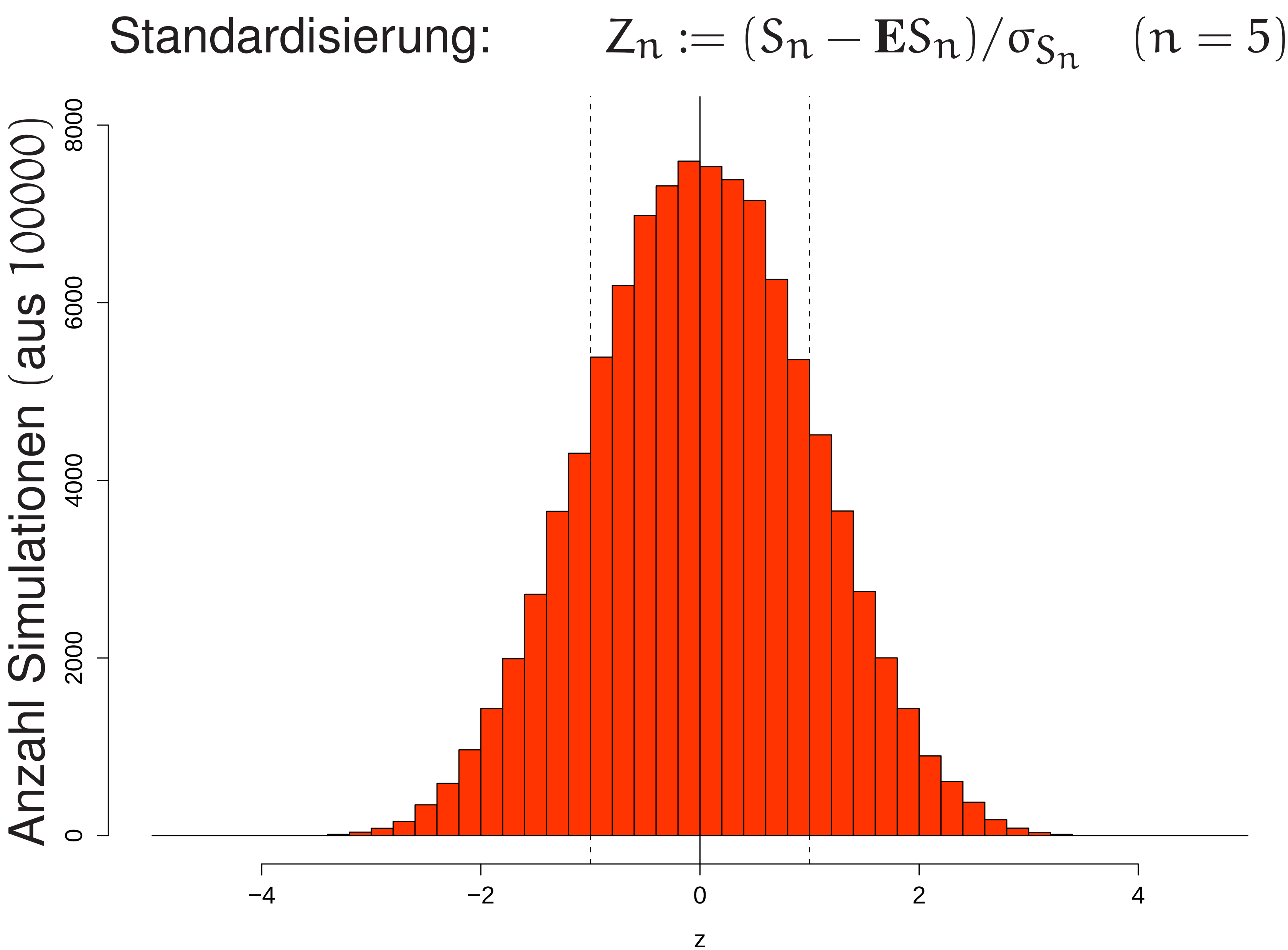


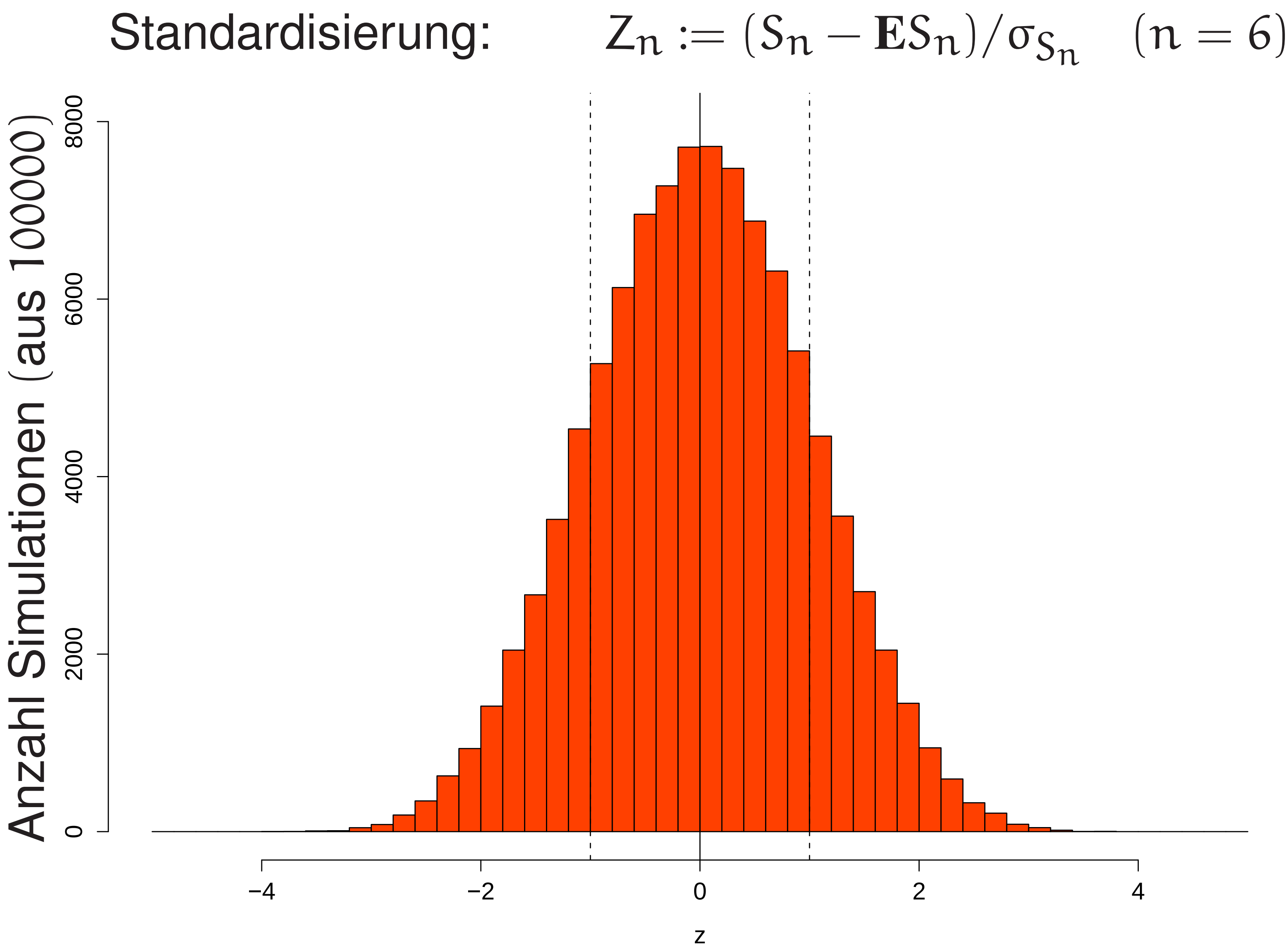

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 2)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

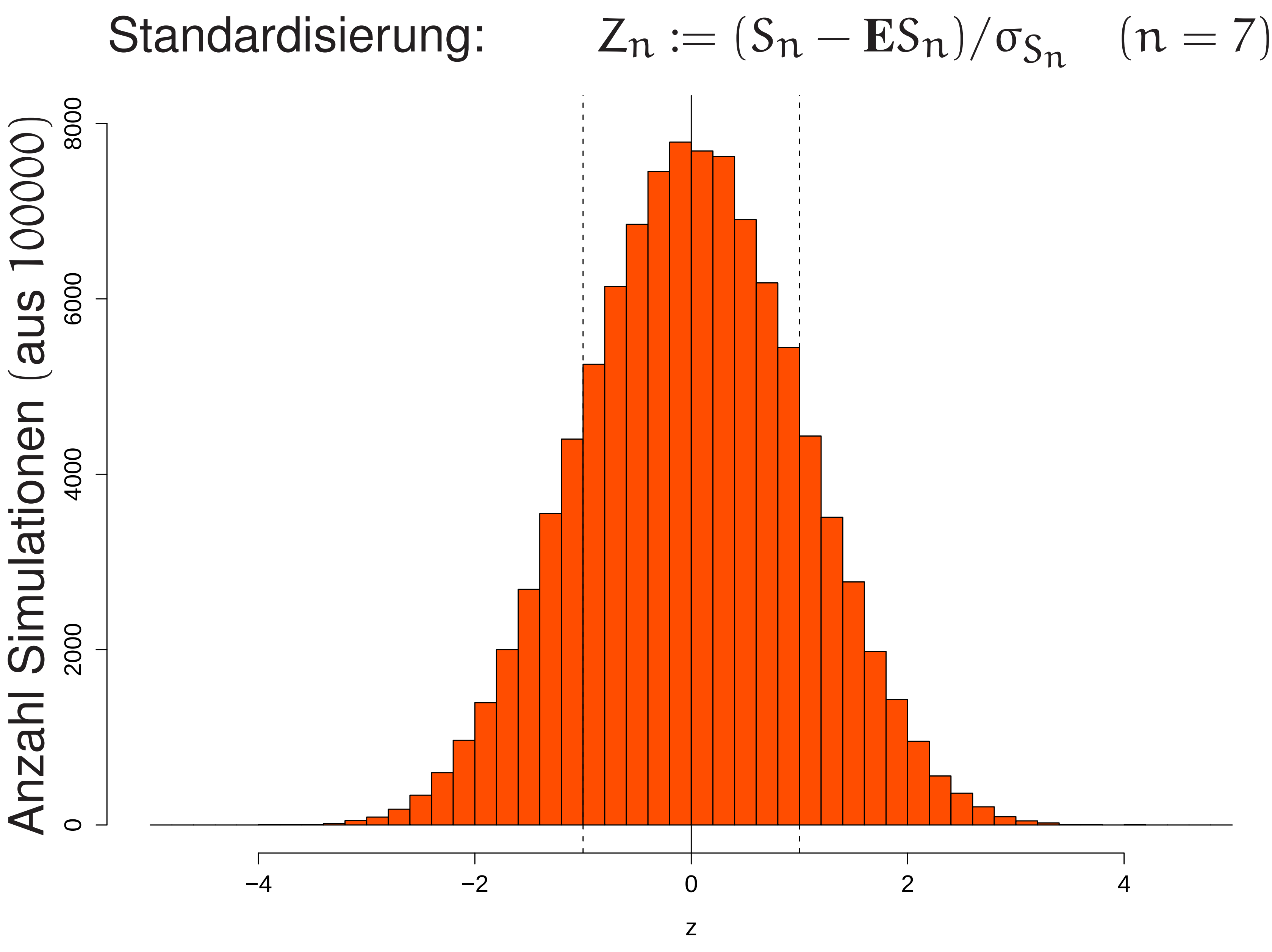

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 3)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

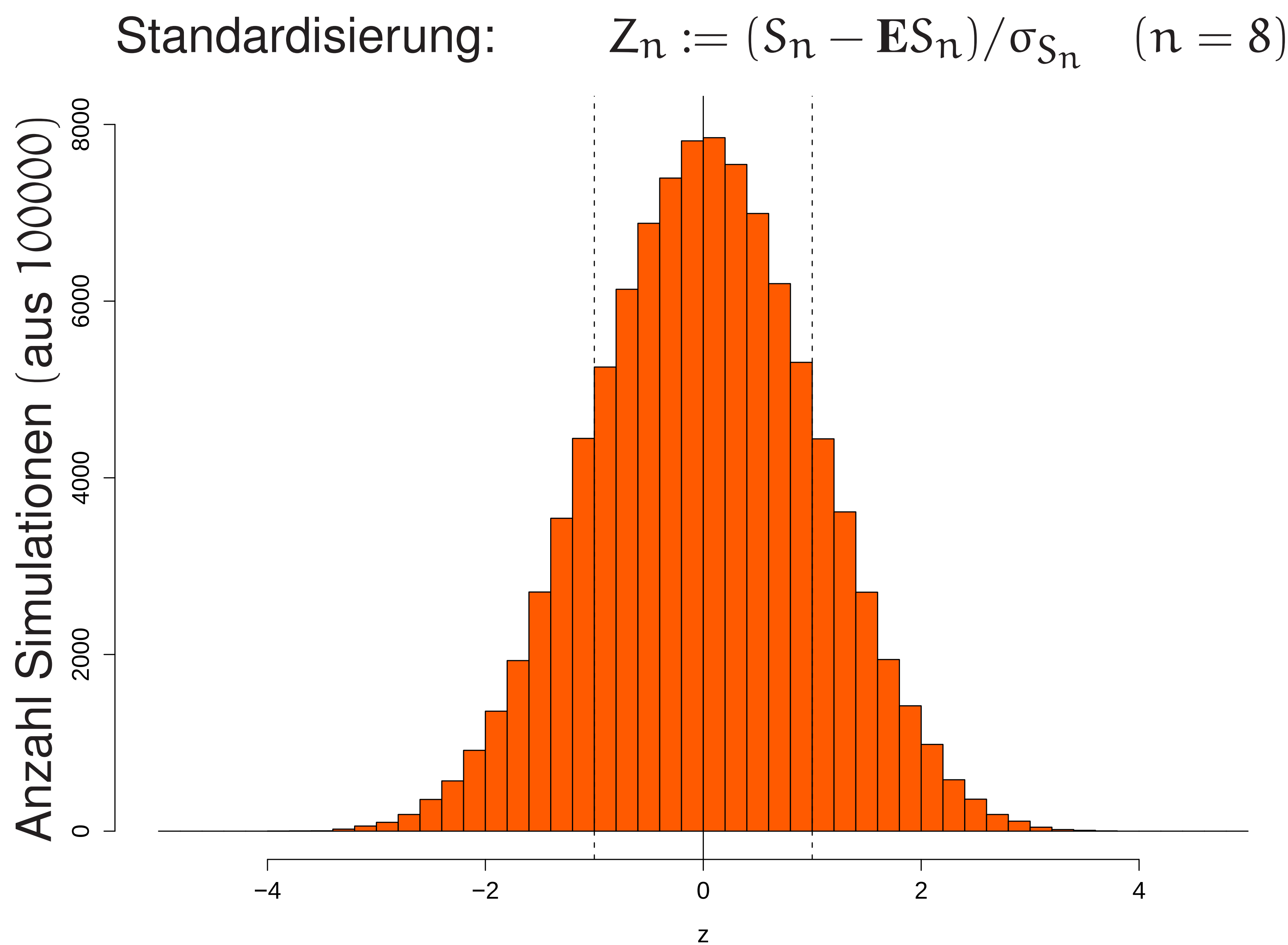
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 4)$



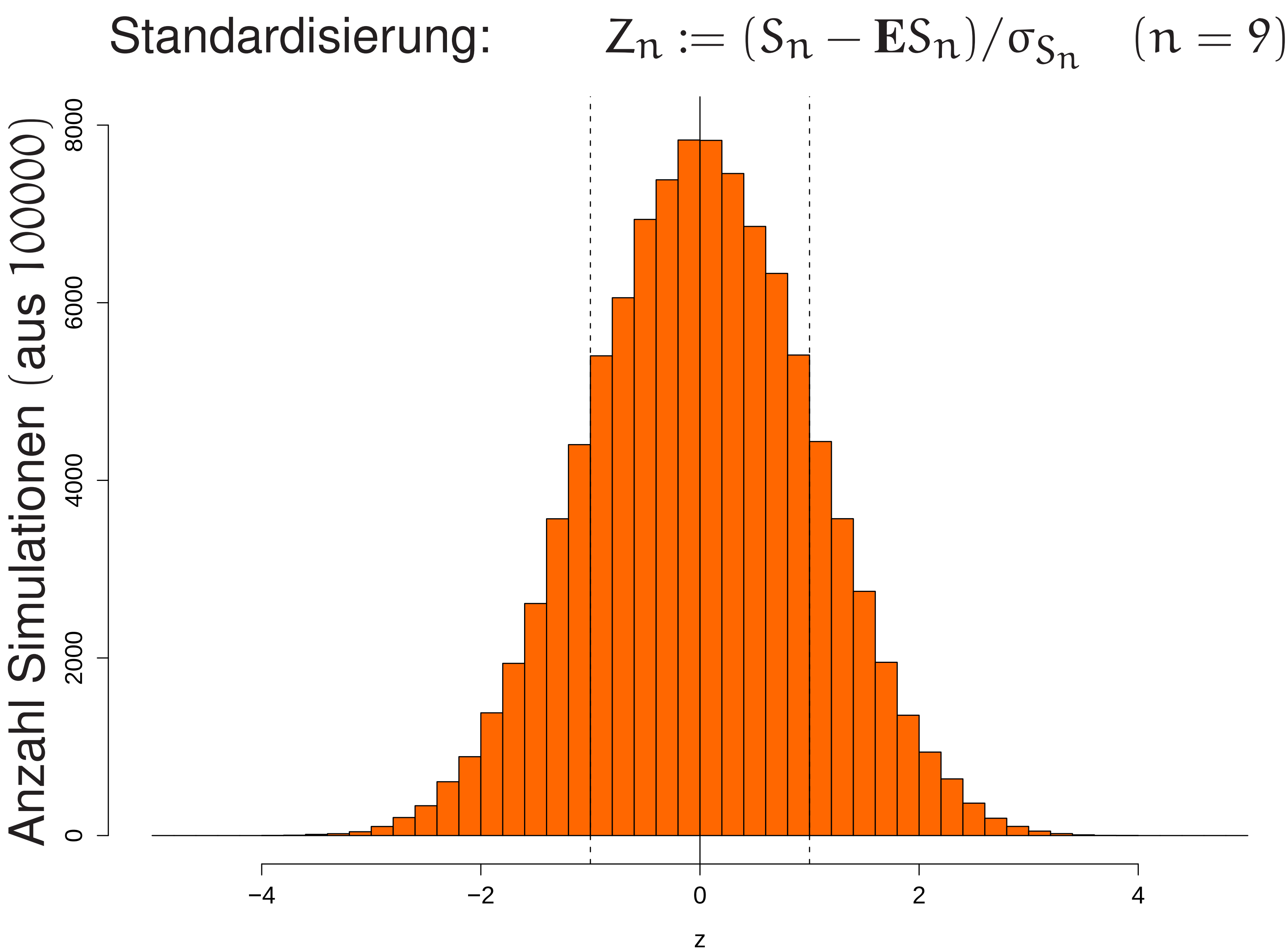

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 5)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 6)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

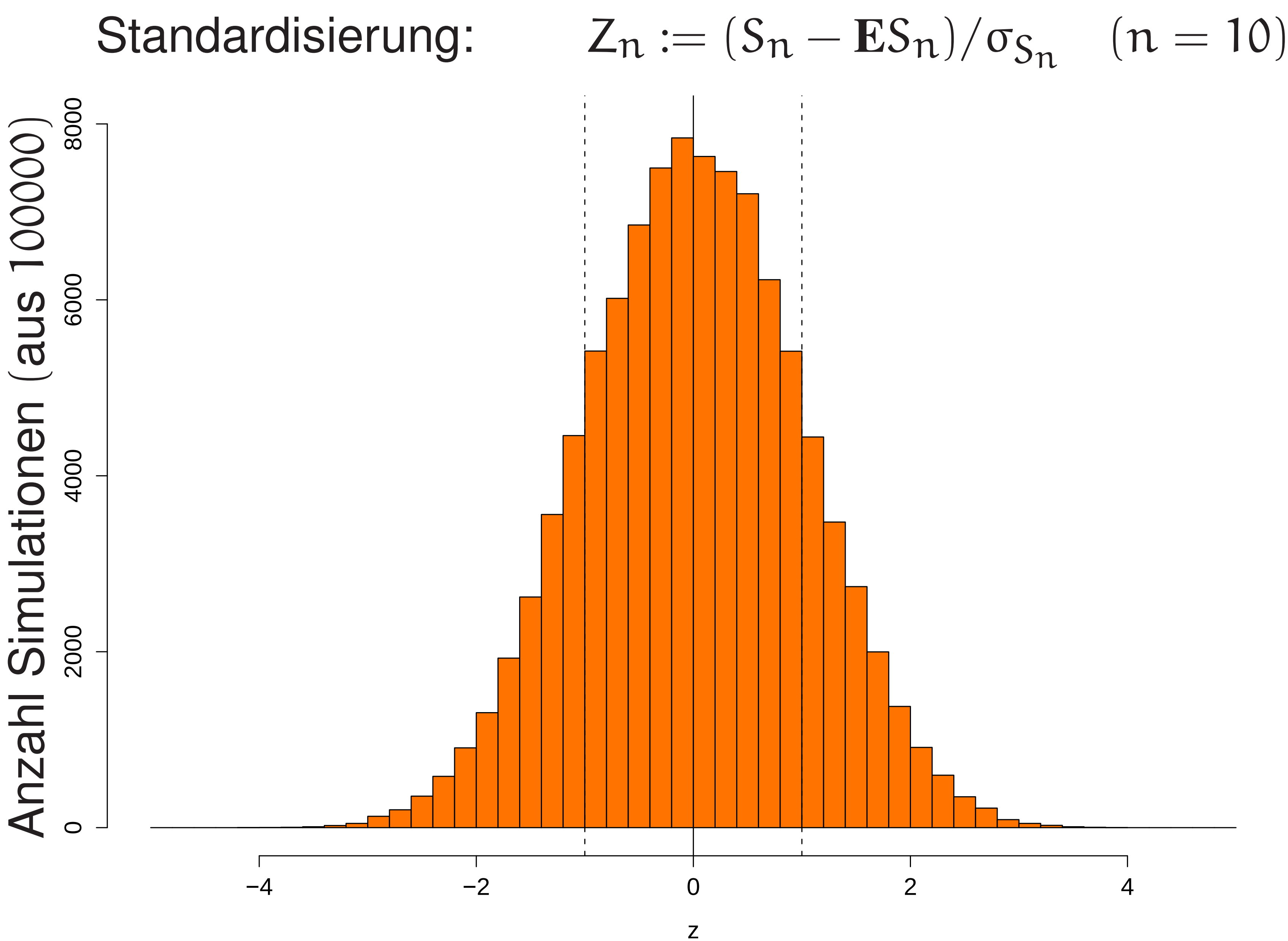

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 7)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z



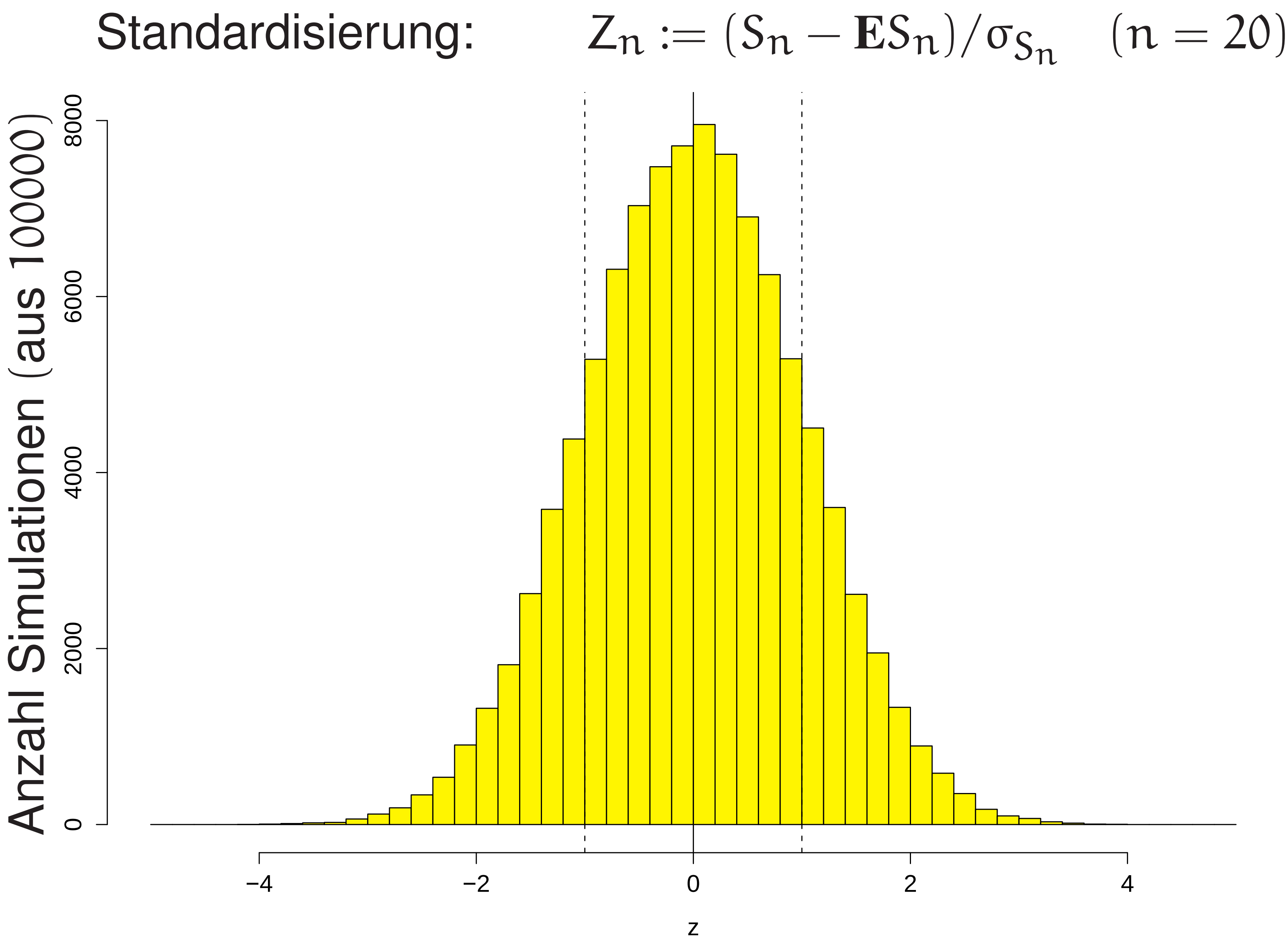
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 8)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

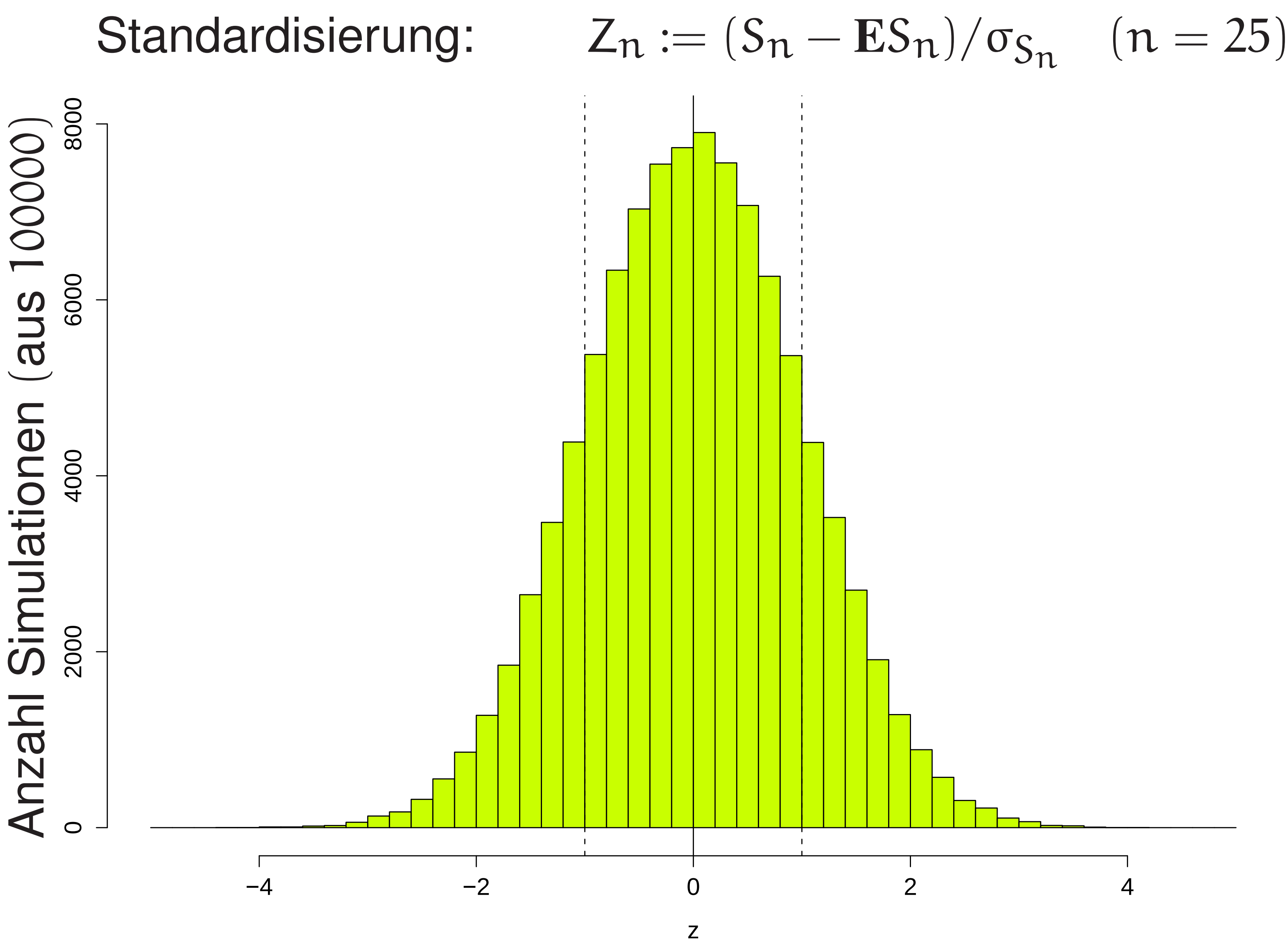


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 9)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

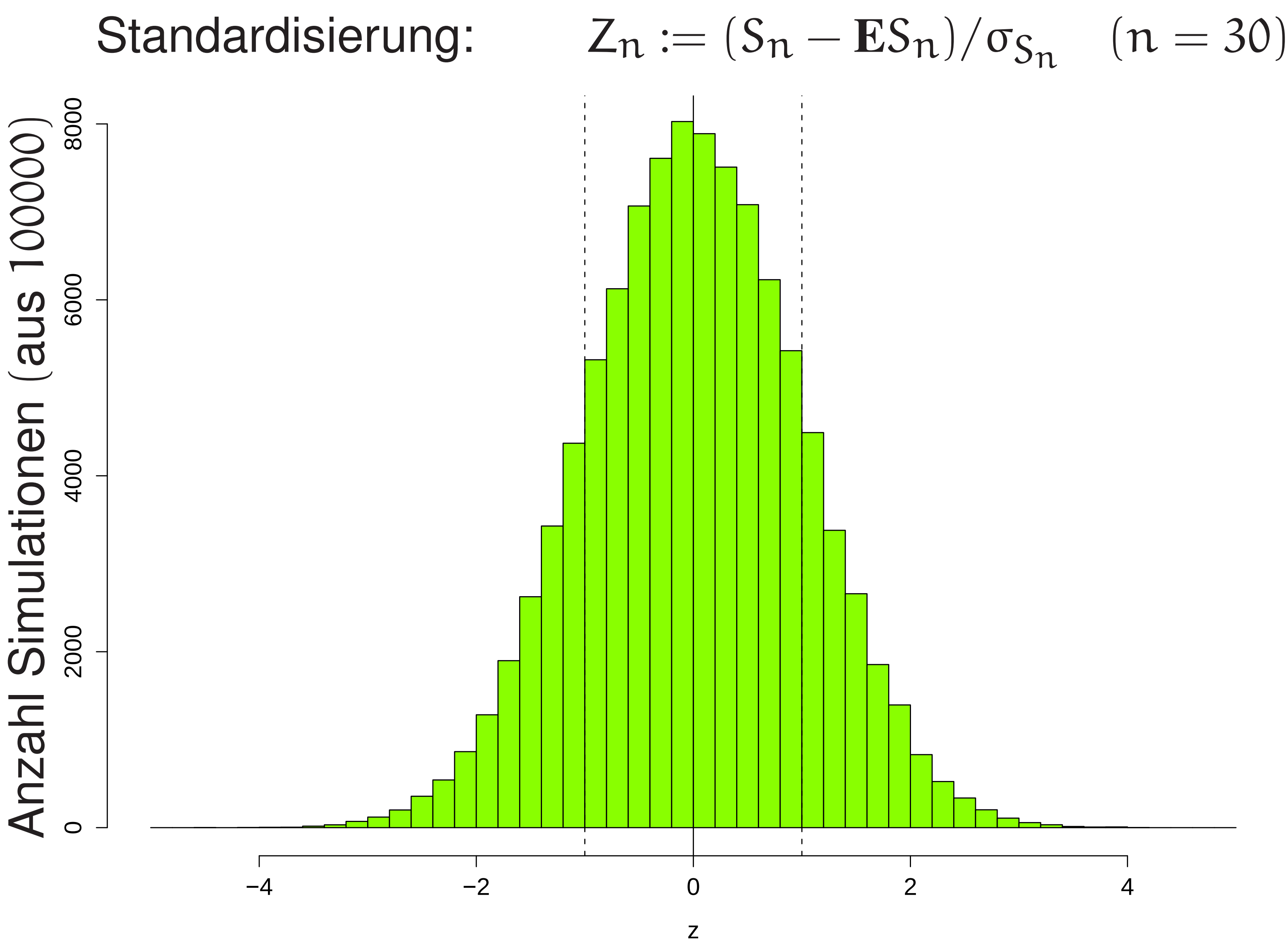

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 10)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

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


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 20)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

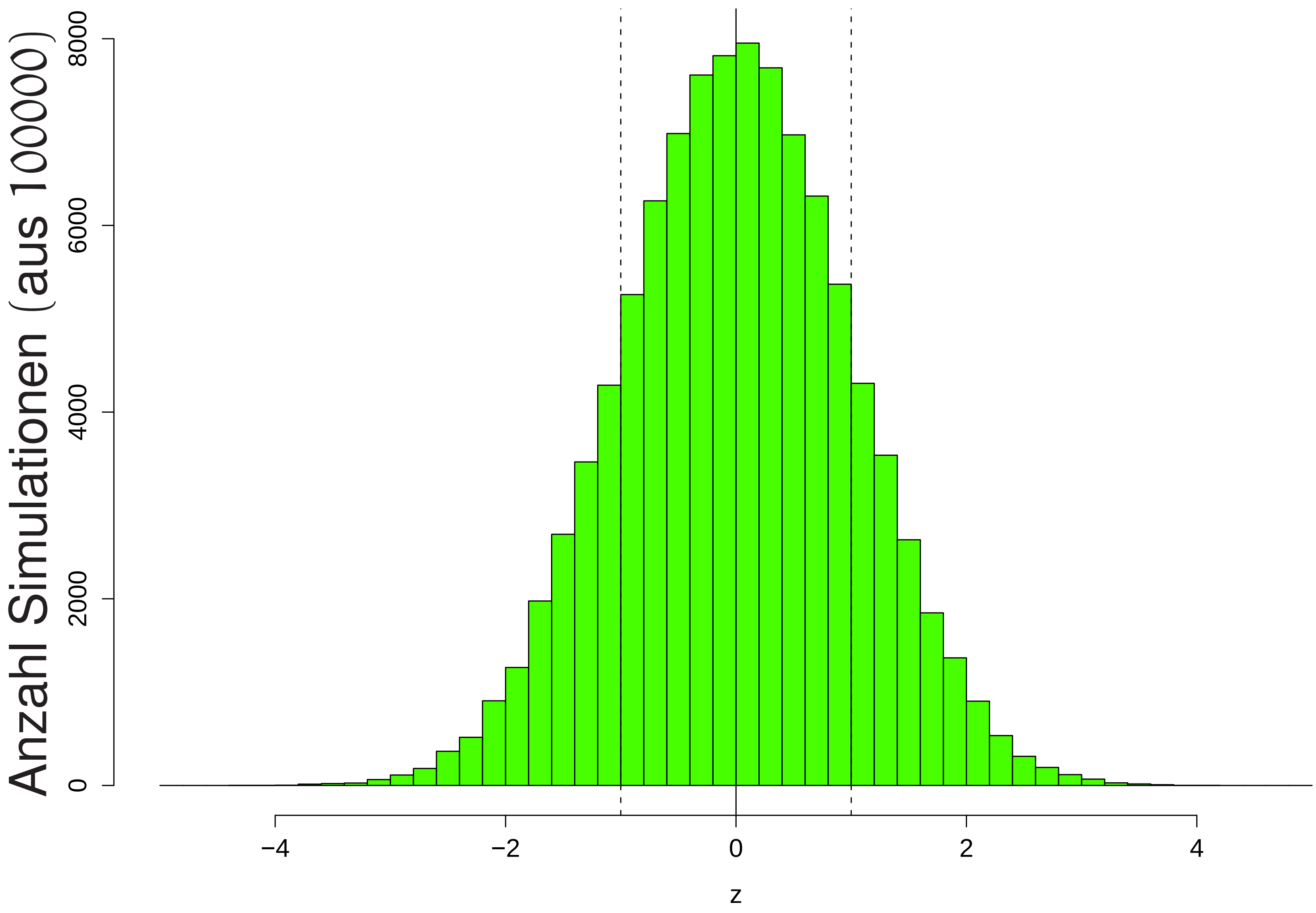


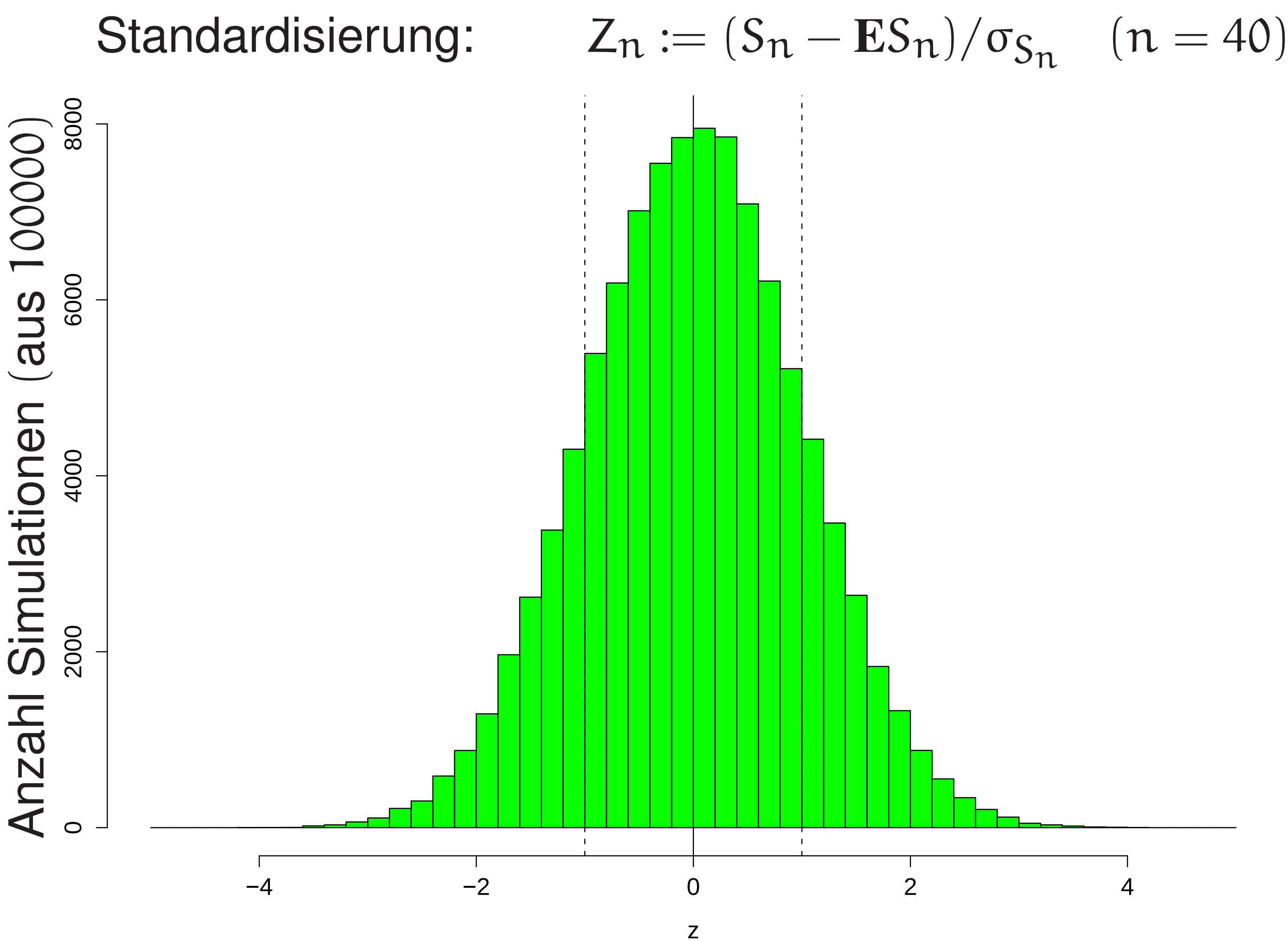
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 25)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z



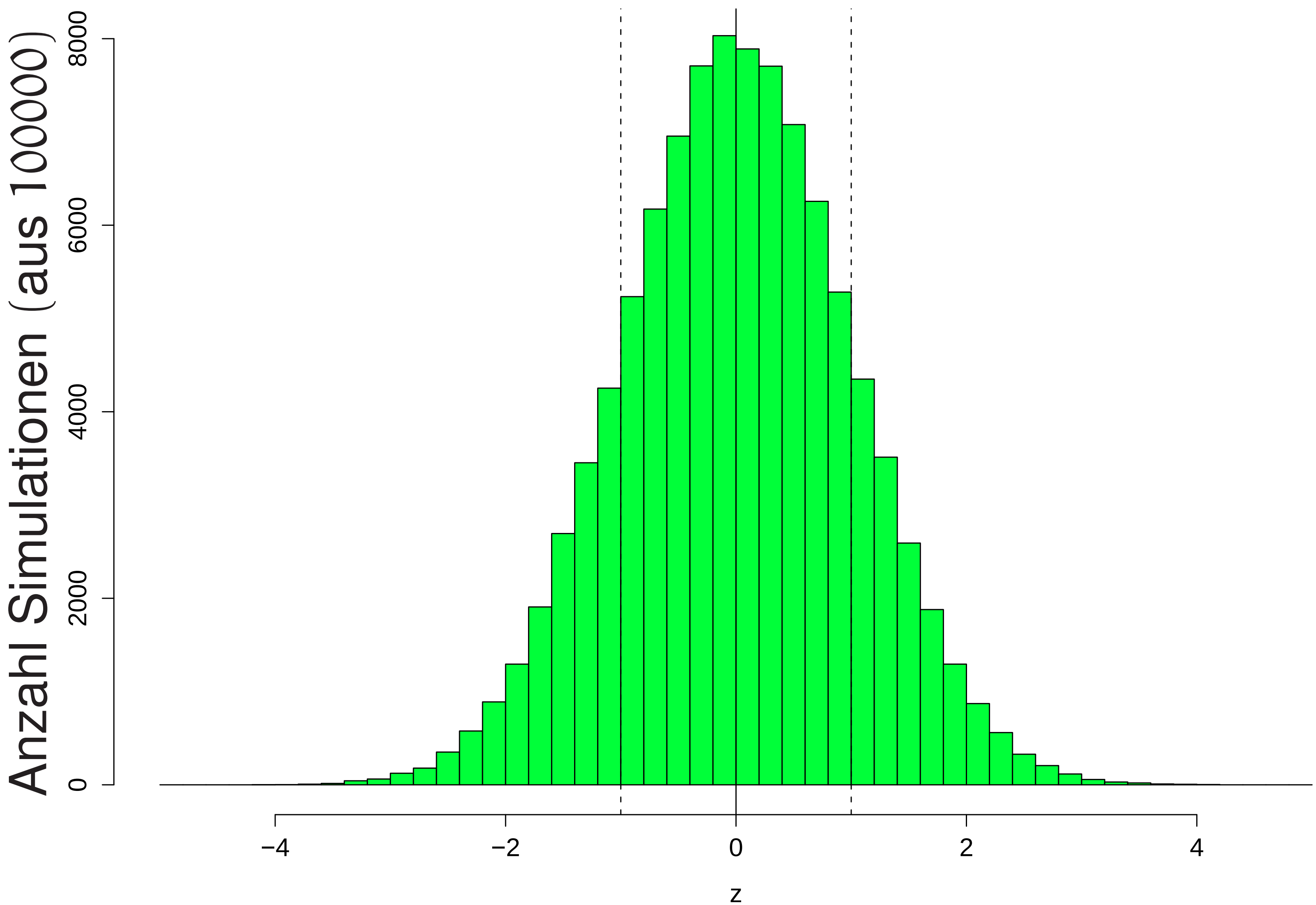
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 30)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

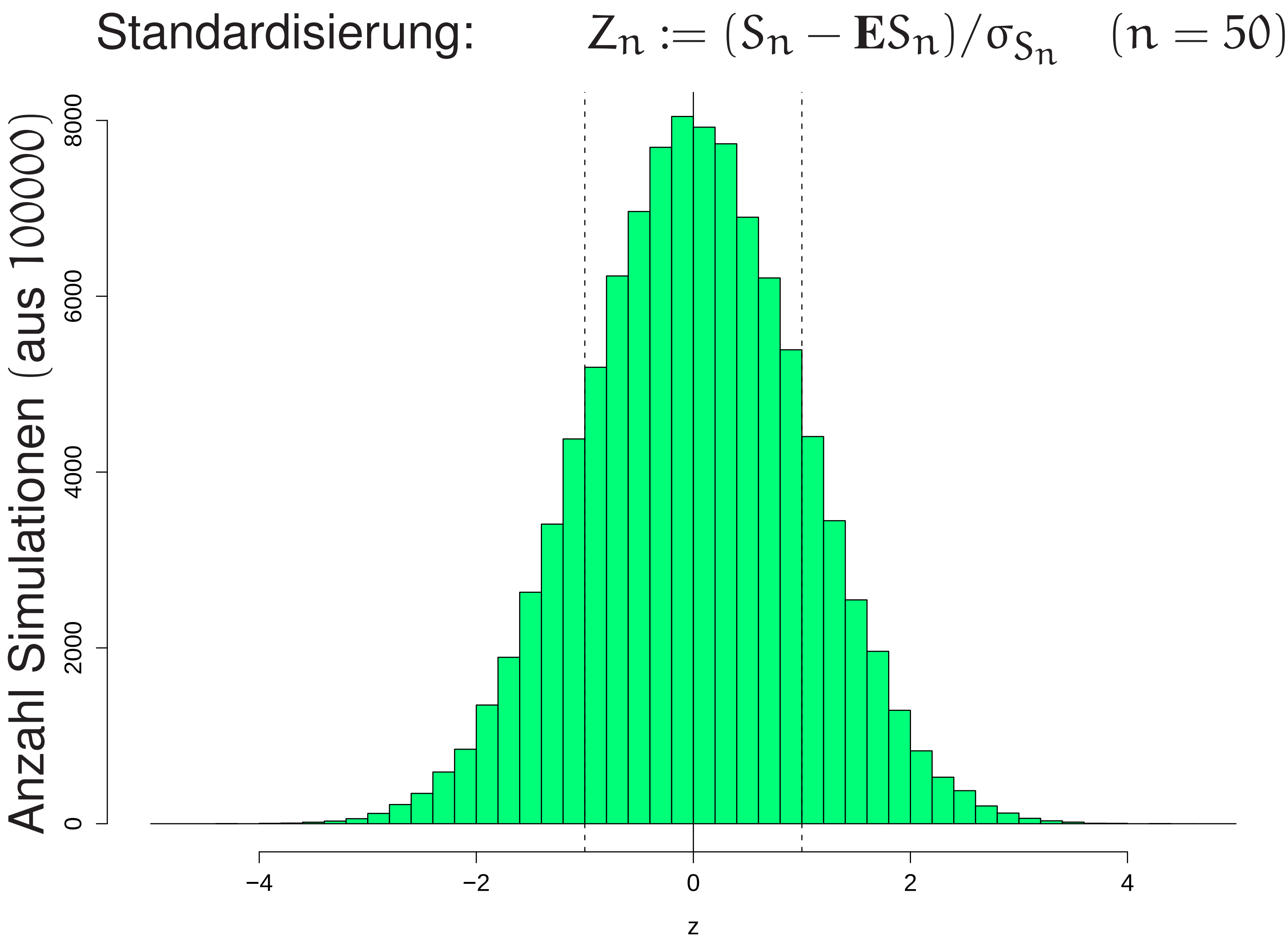
Standardisierung: $\quad Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 35)$

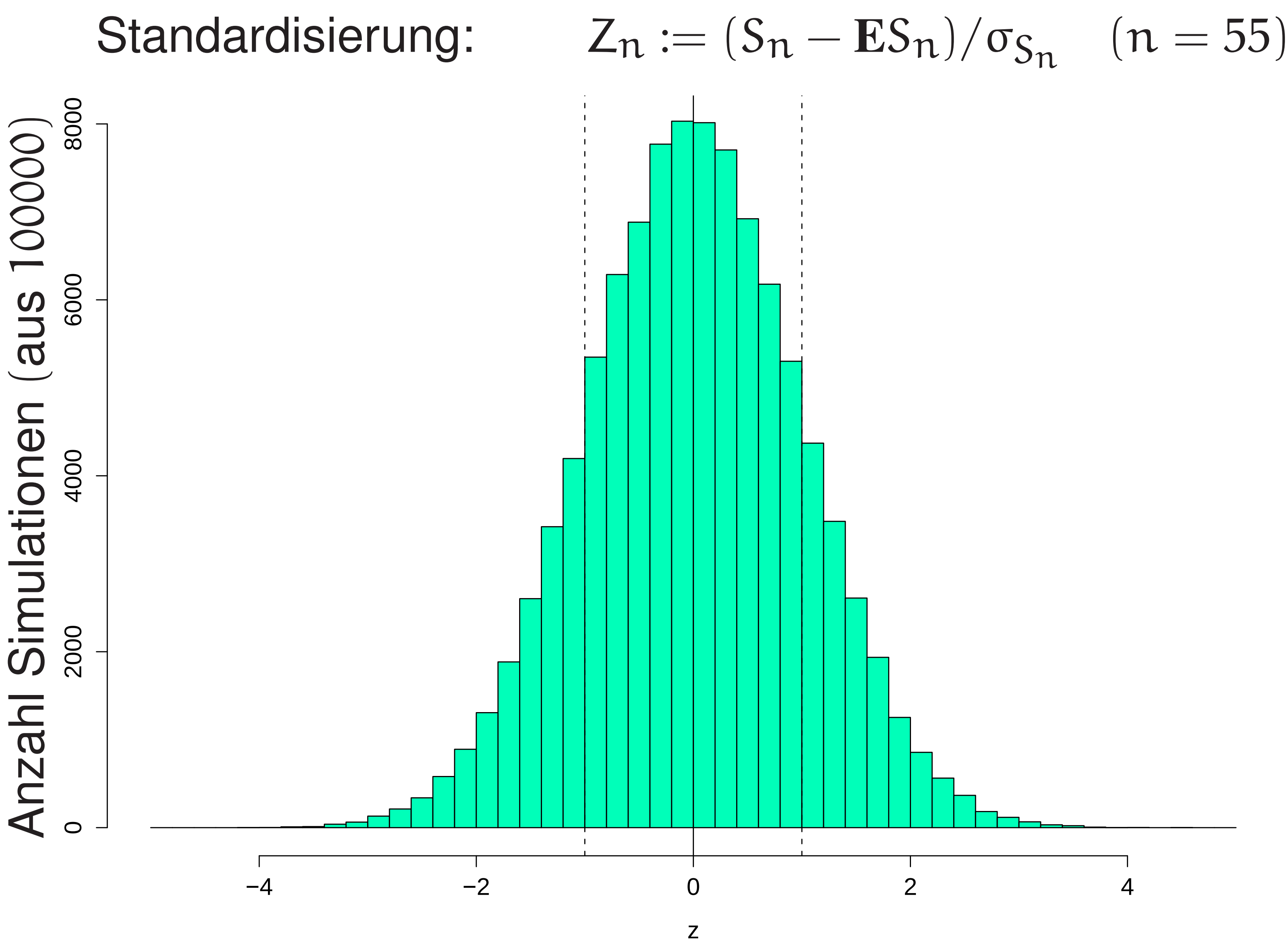



Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 40)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

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

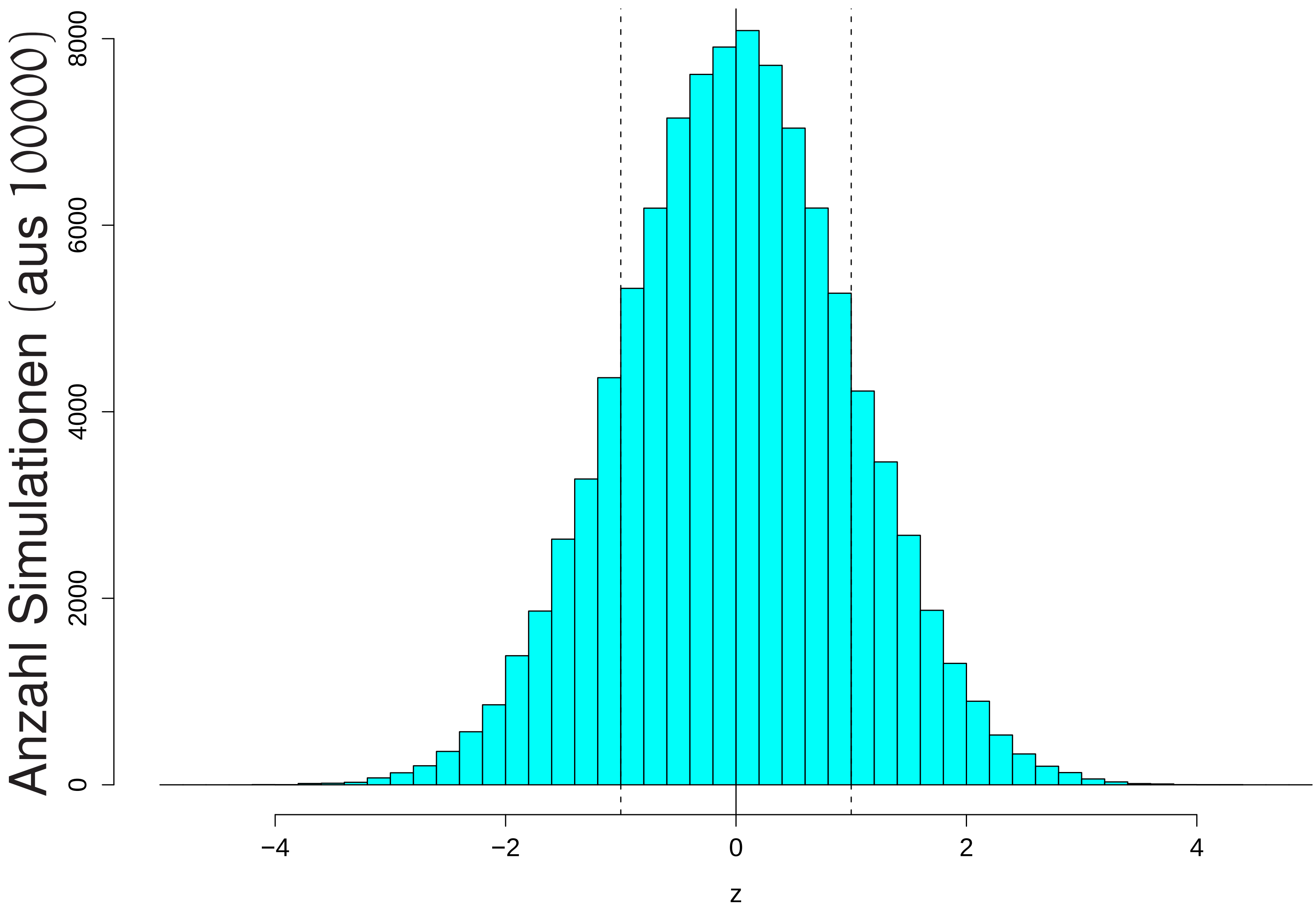


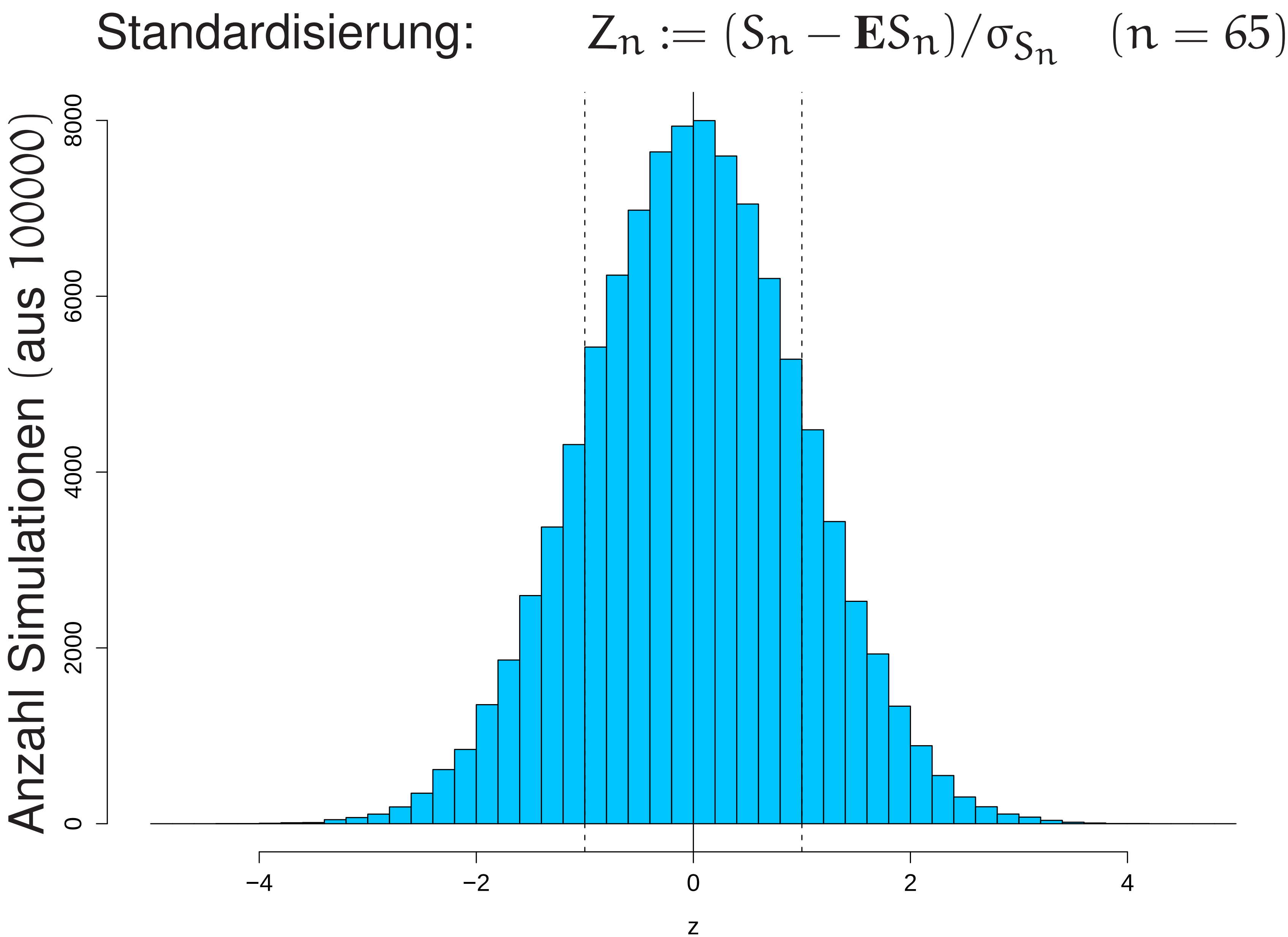

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 50)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

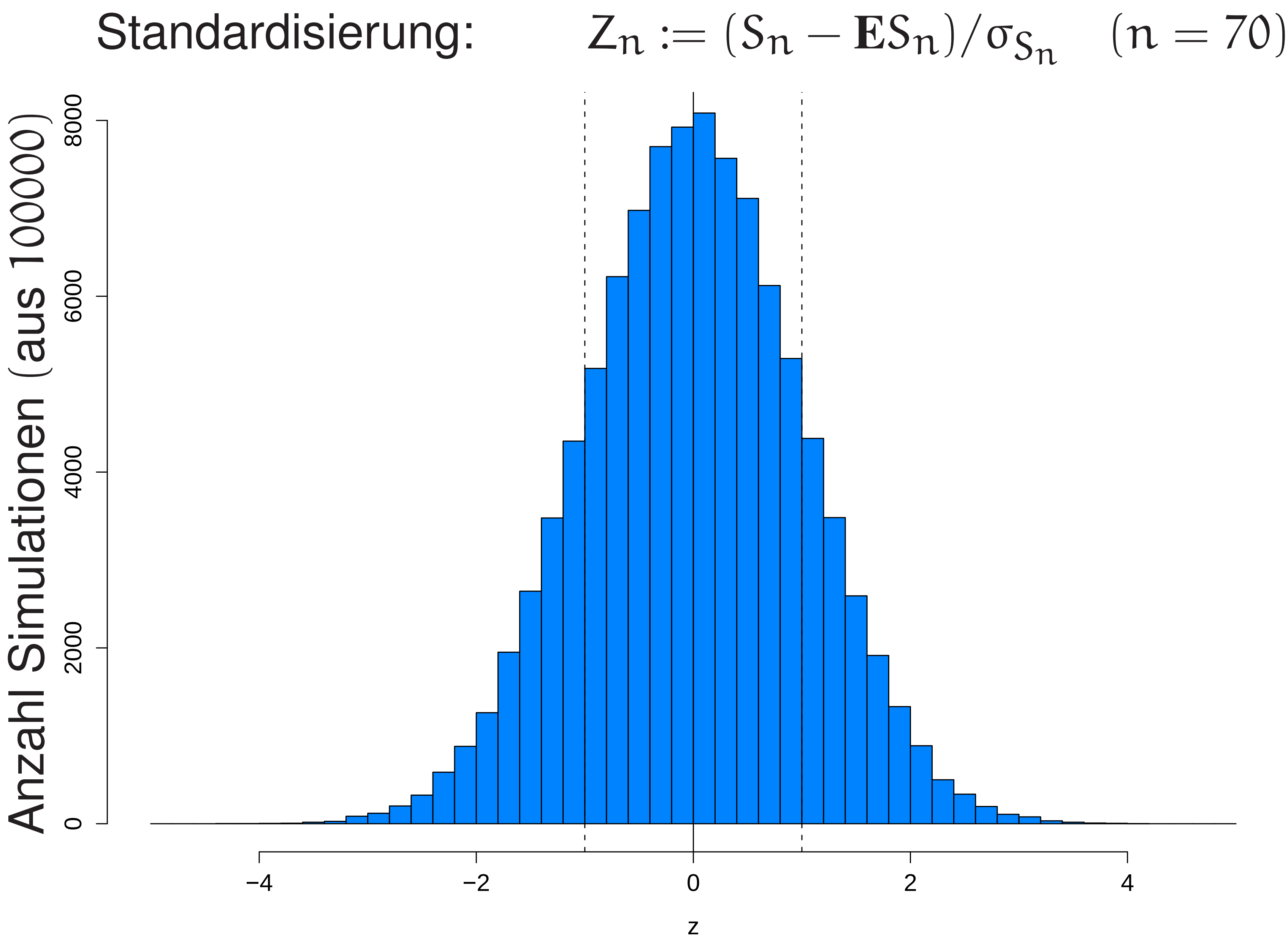


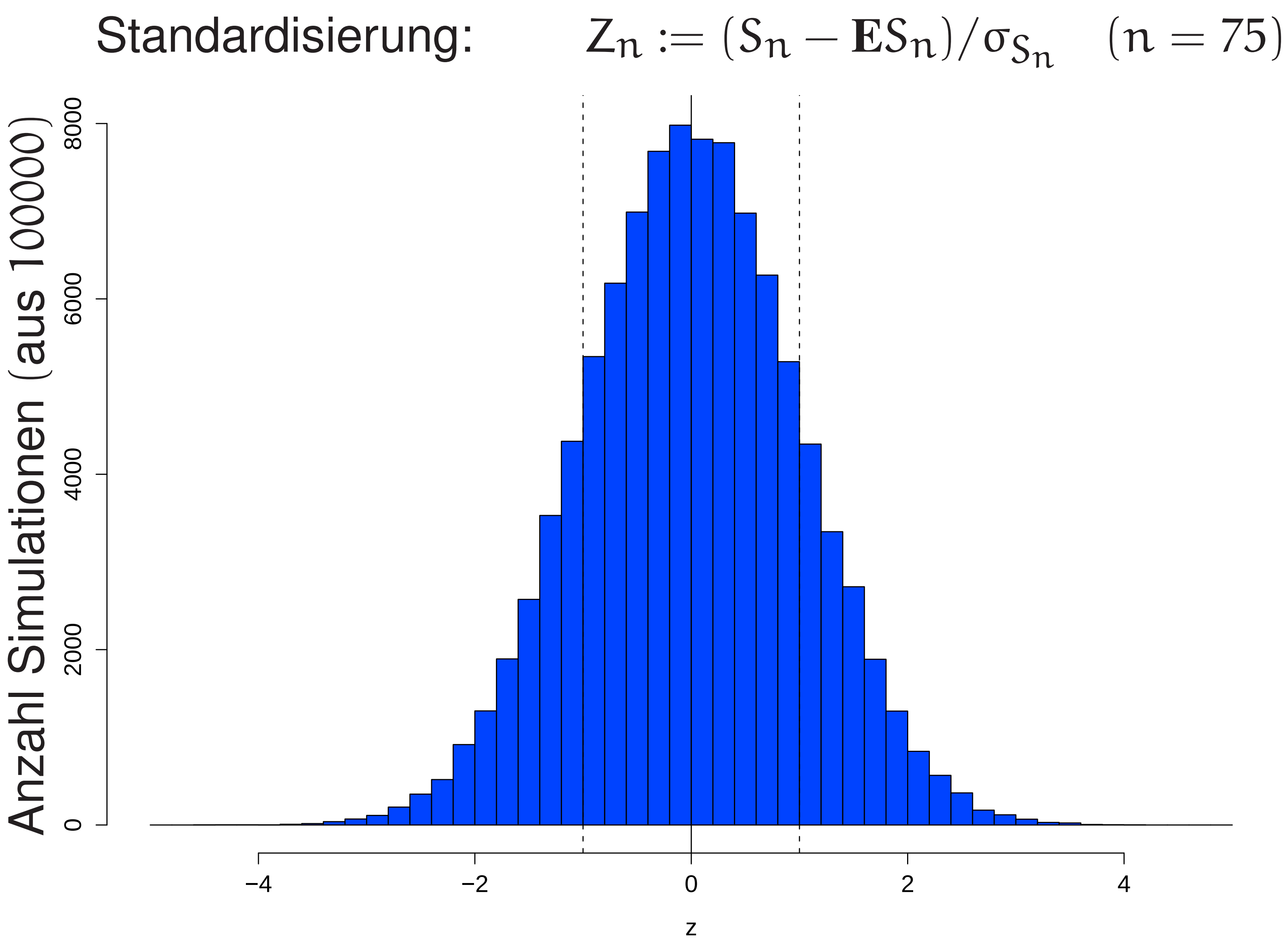
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 55)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

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

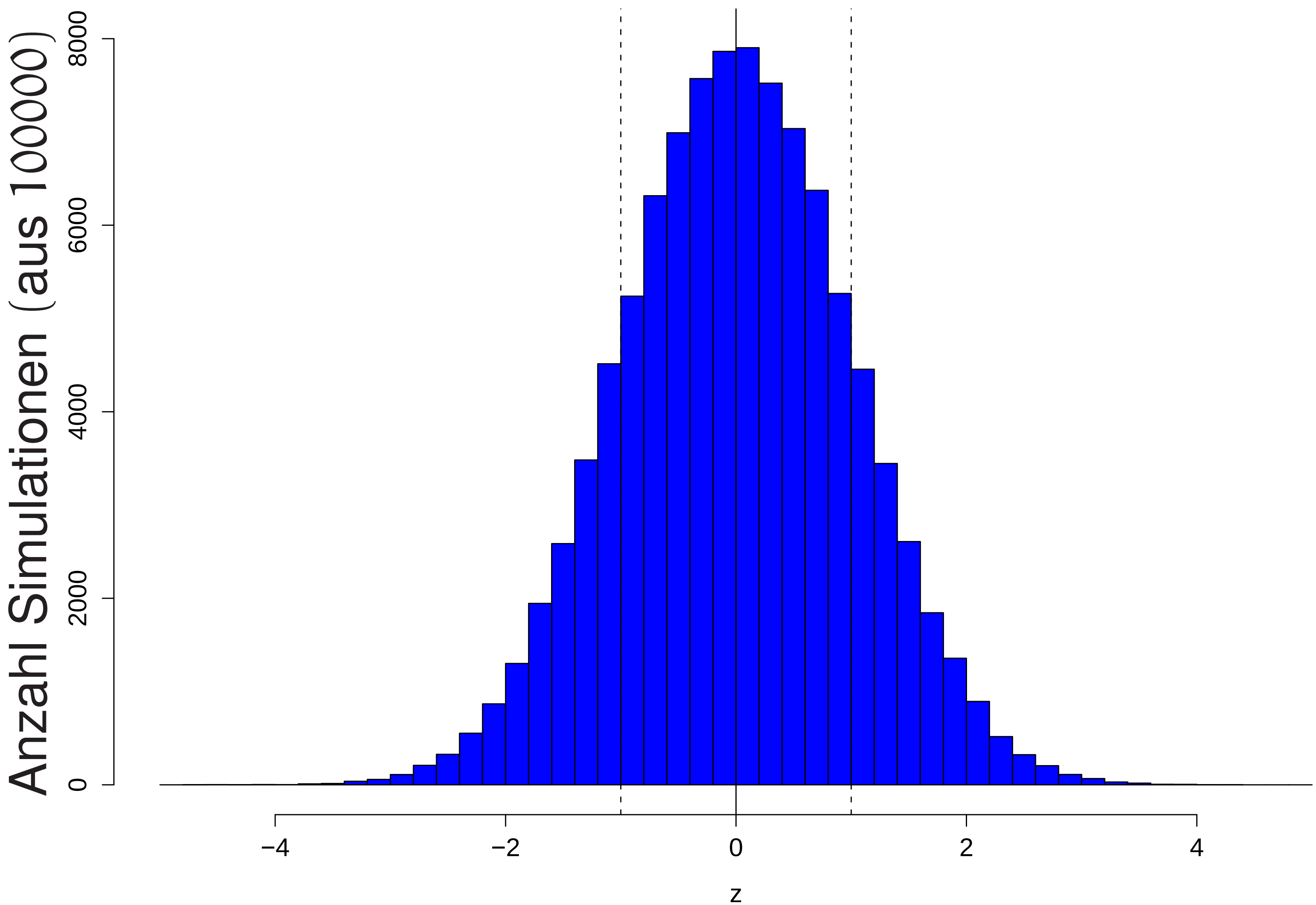


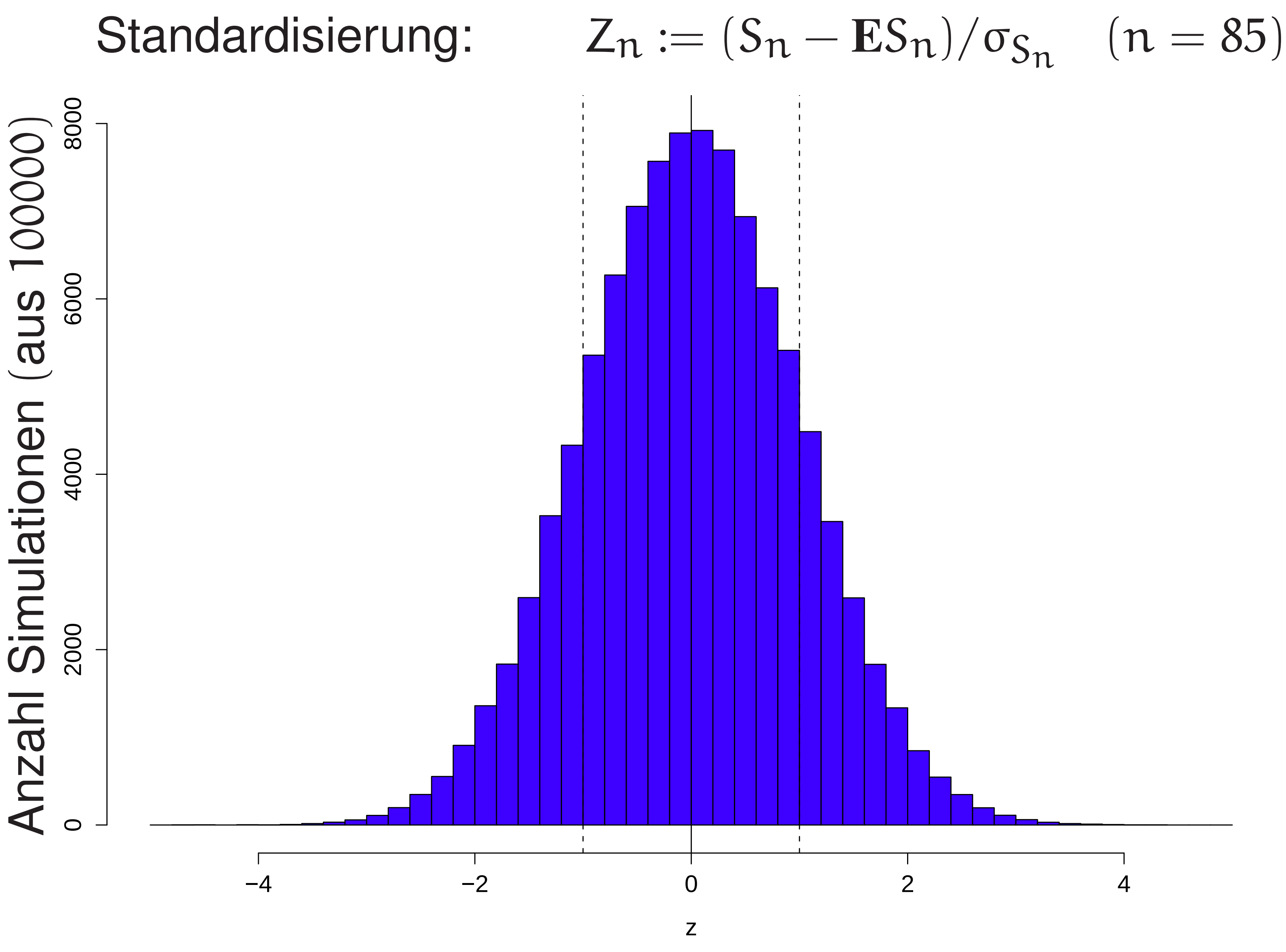

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 65)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

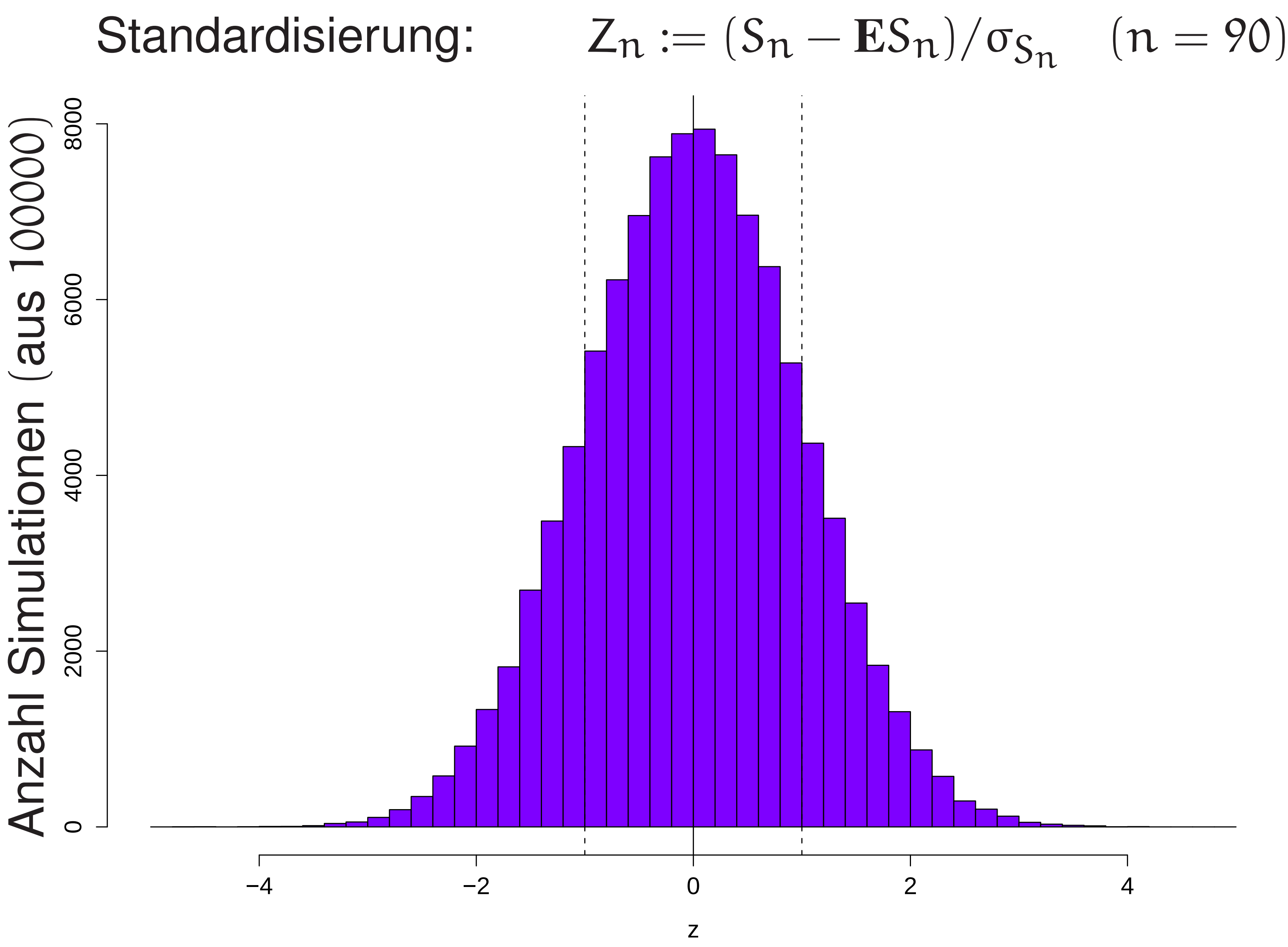

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 70)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

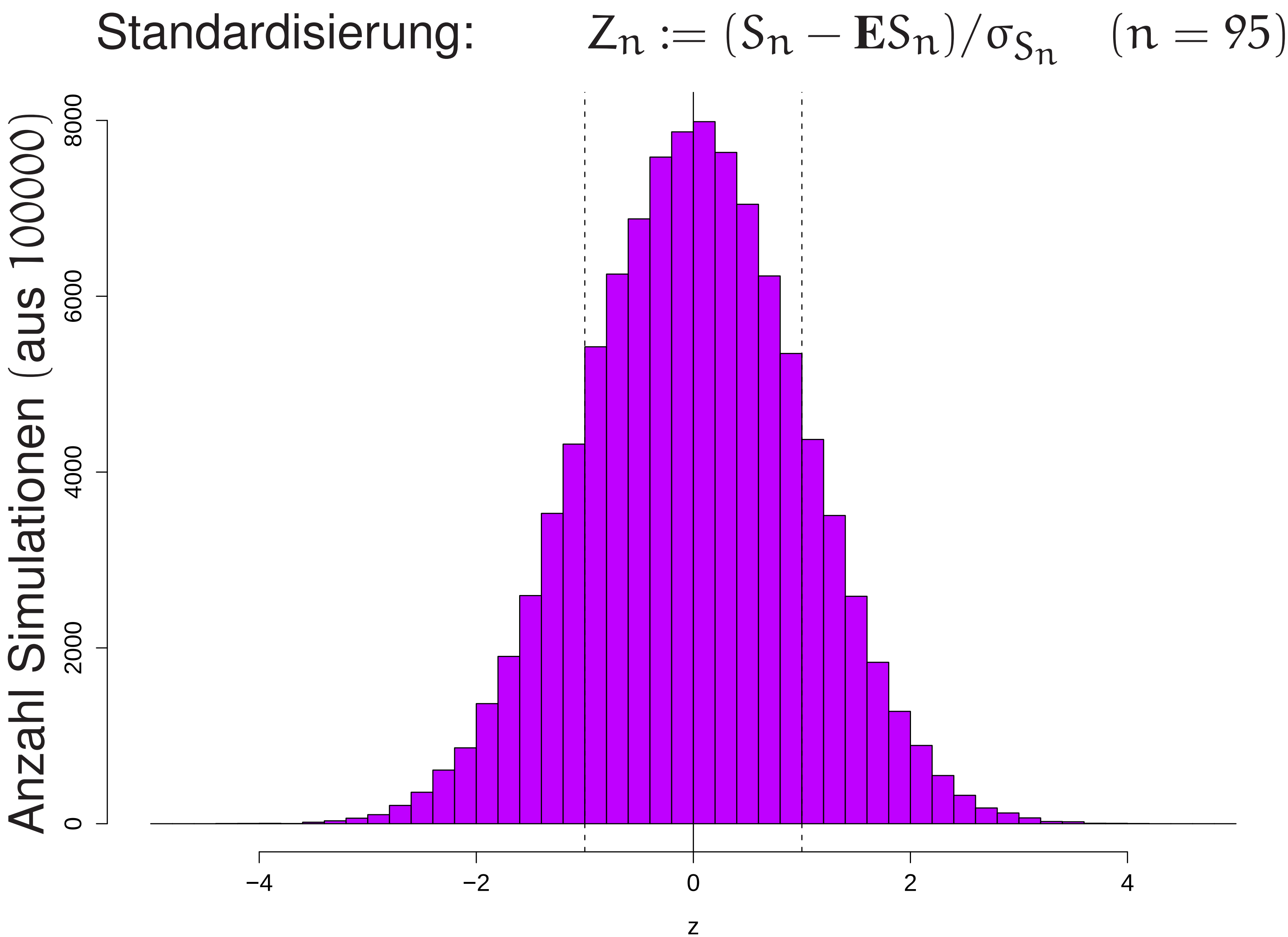

Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 75)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

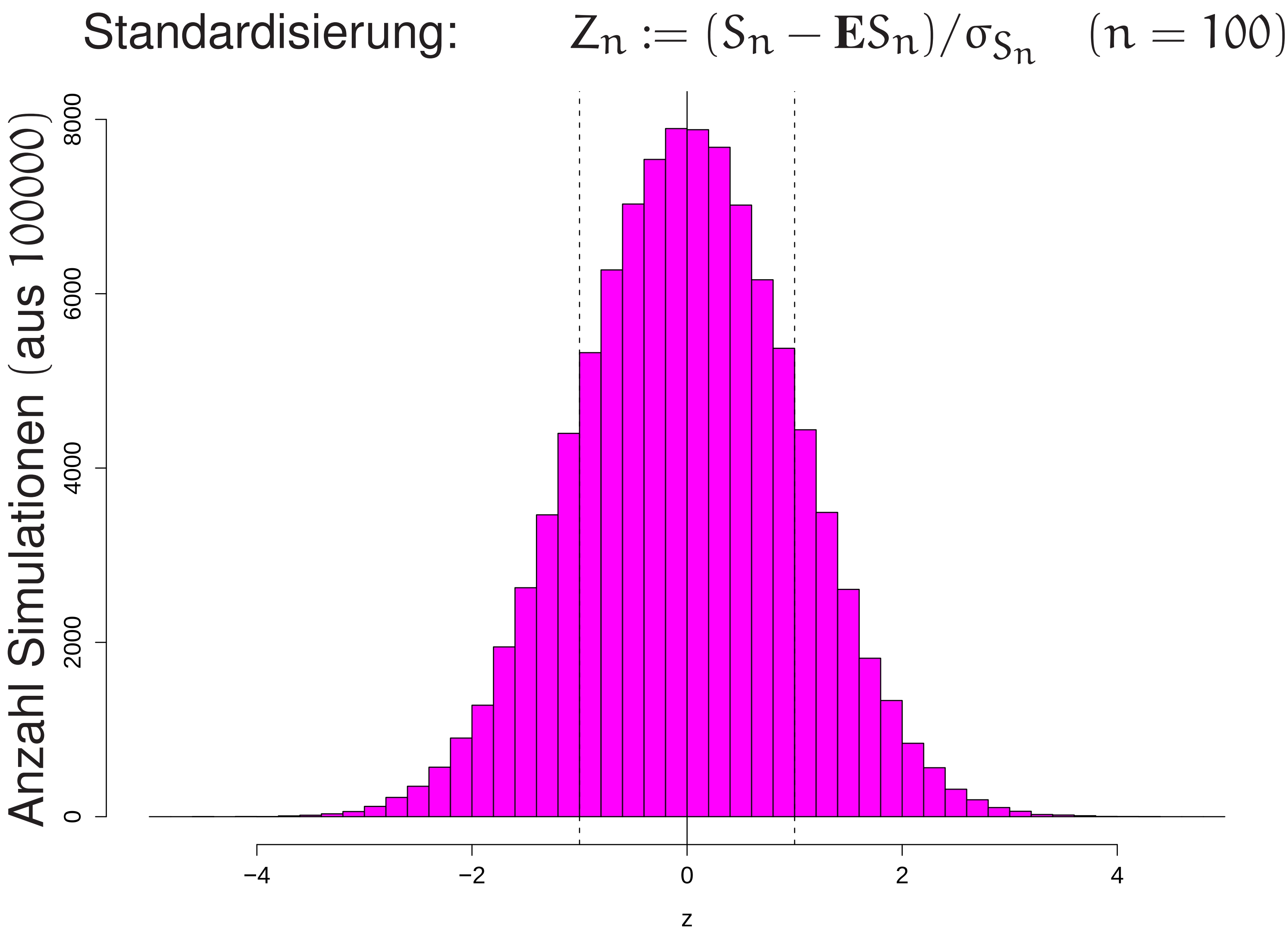
Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 80)$




Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 85)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 90)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z


Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 95)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z



Standardisierung: $Z_n := (S_n - \mathbf{E}S_n)/\sigma_{S_n} \quad (n = 100)$
Anzahl Simulationen (aus 100000)
0
2000
4000
6000
8000
−4
−2
0
2
4
z

Die Verteilung von Z_n

scheint zu konvergieren.

Die Verteilung von Z_n

scheint zu konvergieren.

Welche Form

hat die Grenzverteilung?

Die Verteilung von

Z_{100}

ist glockenförmig.

Die Verteilung von

$$Z_{100}$$

ist glockenförmig.

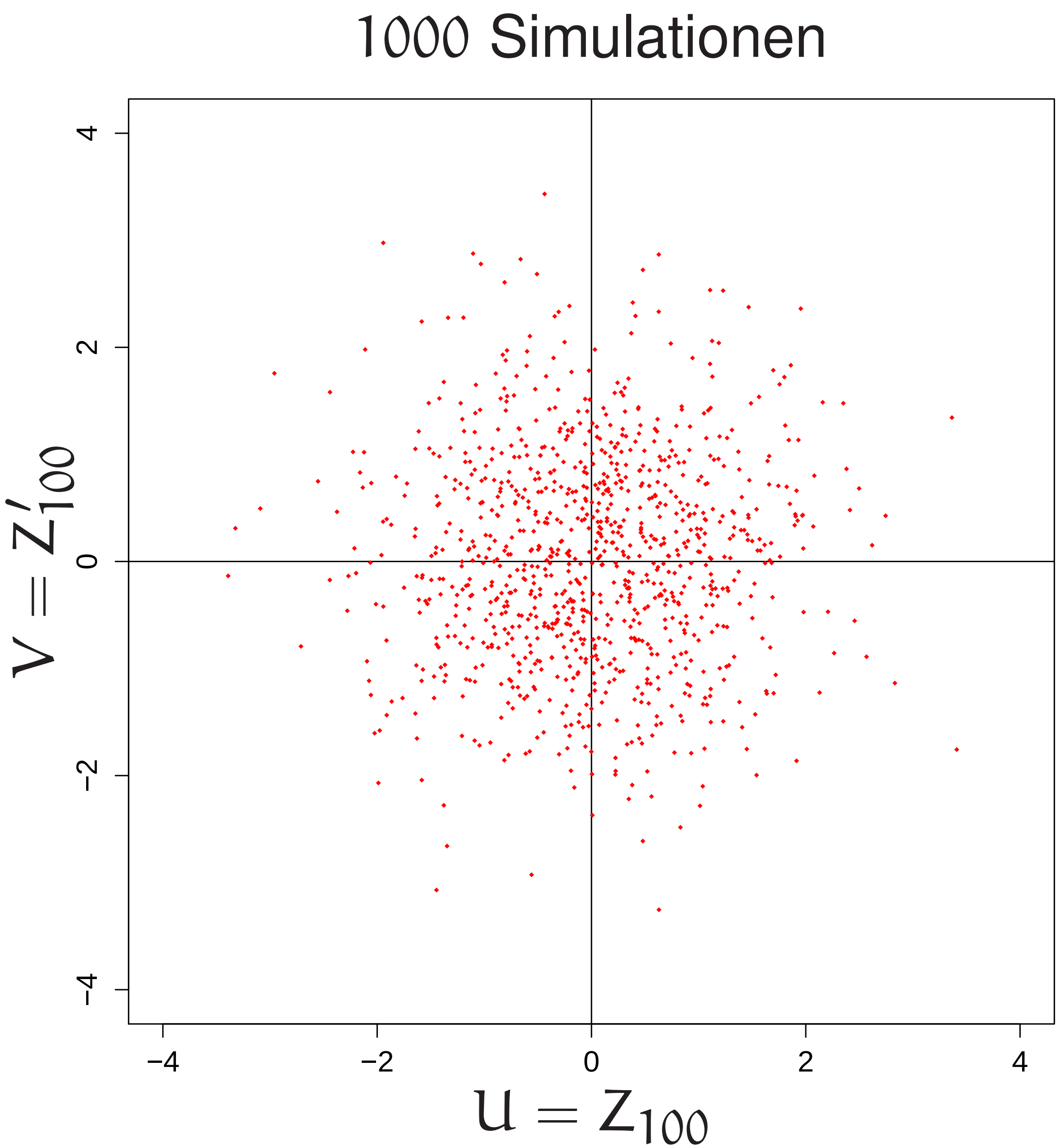
Um welche Glockenkurve handelt es sich genau?

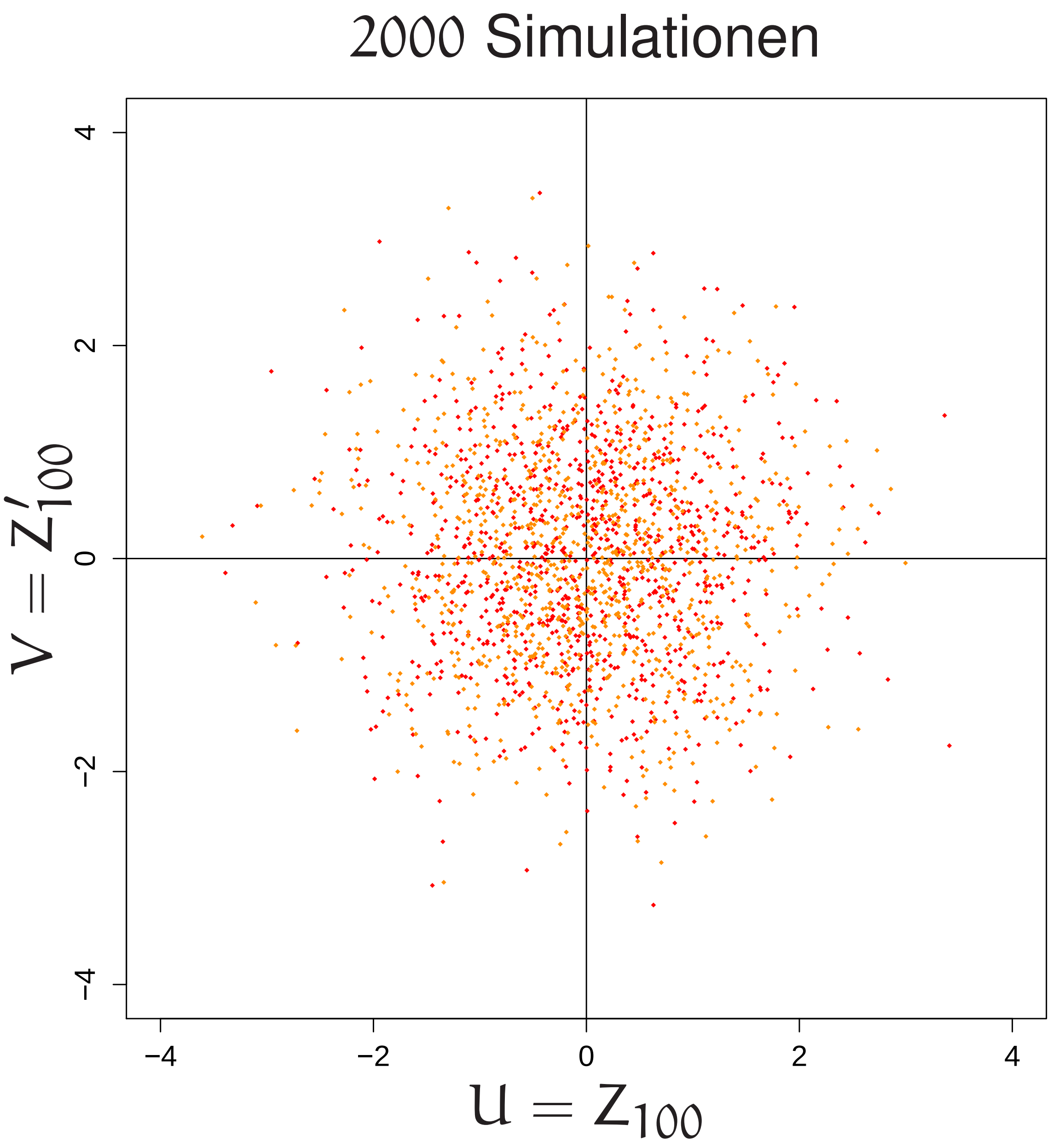
Glücklicher Einfall:

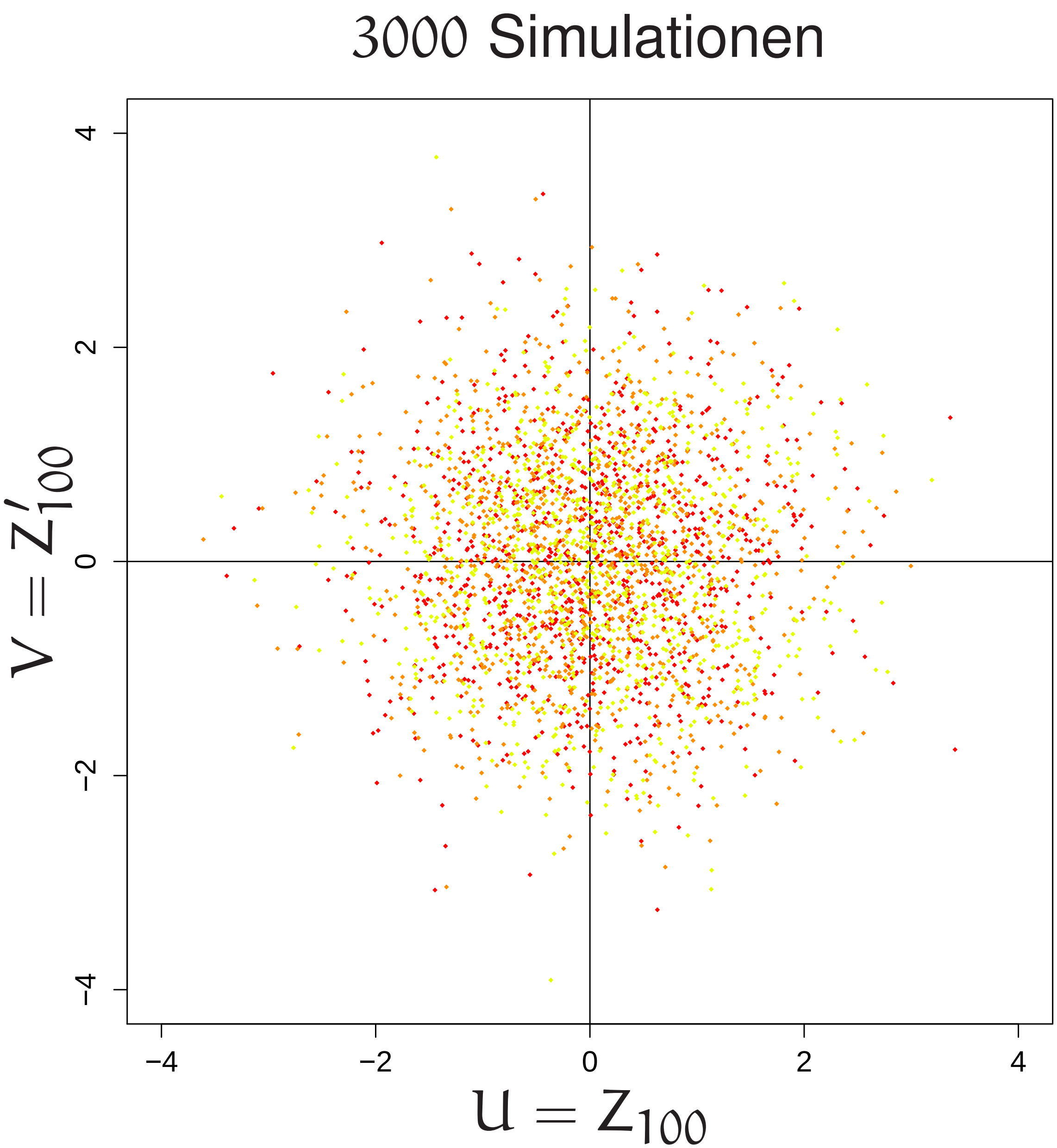
Nimm zwei unabhängige Kopien

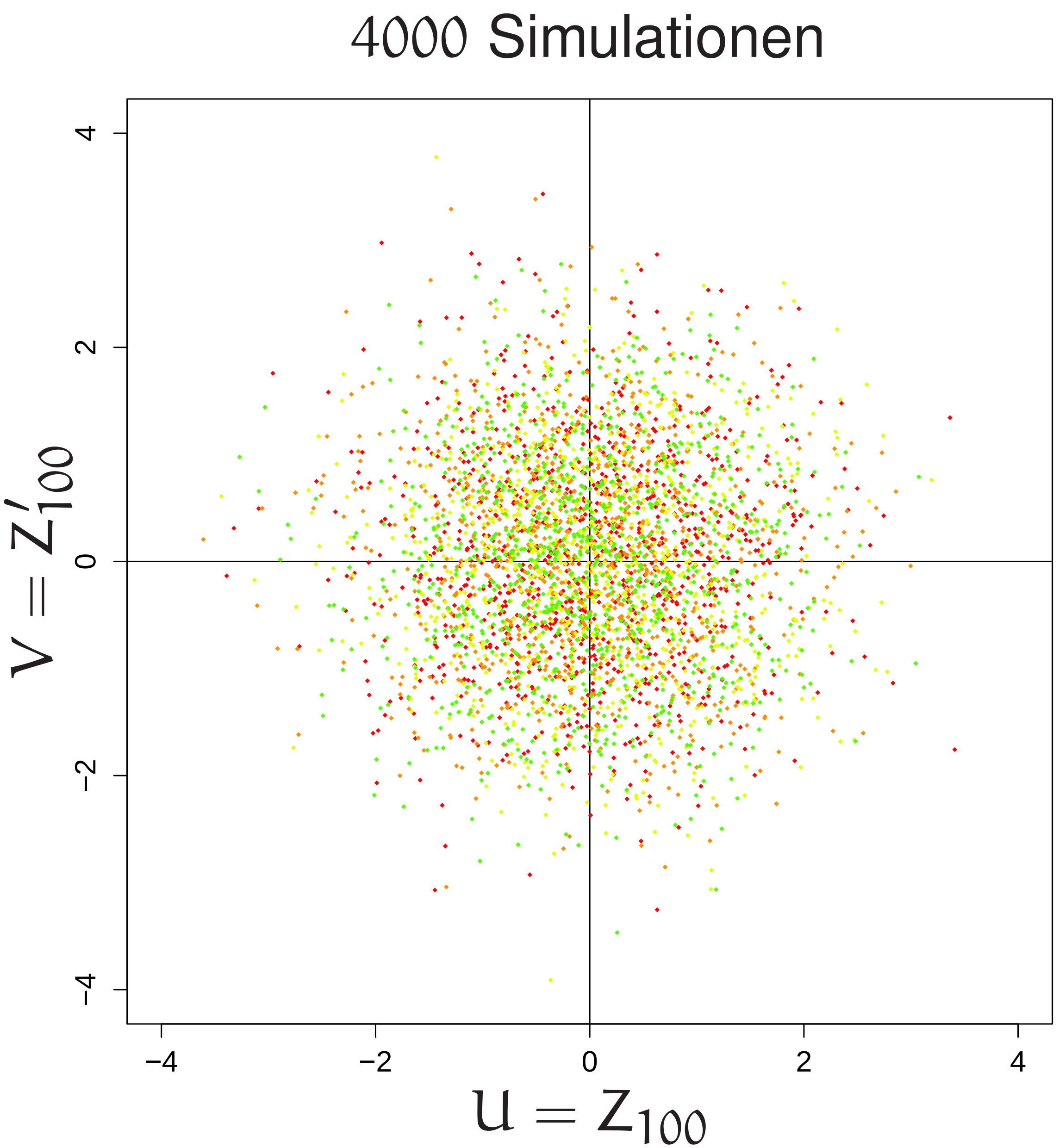
$$(U, V) := (Z_{100}, Z'_{100})$$

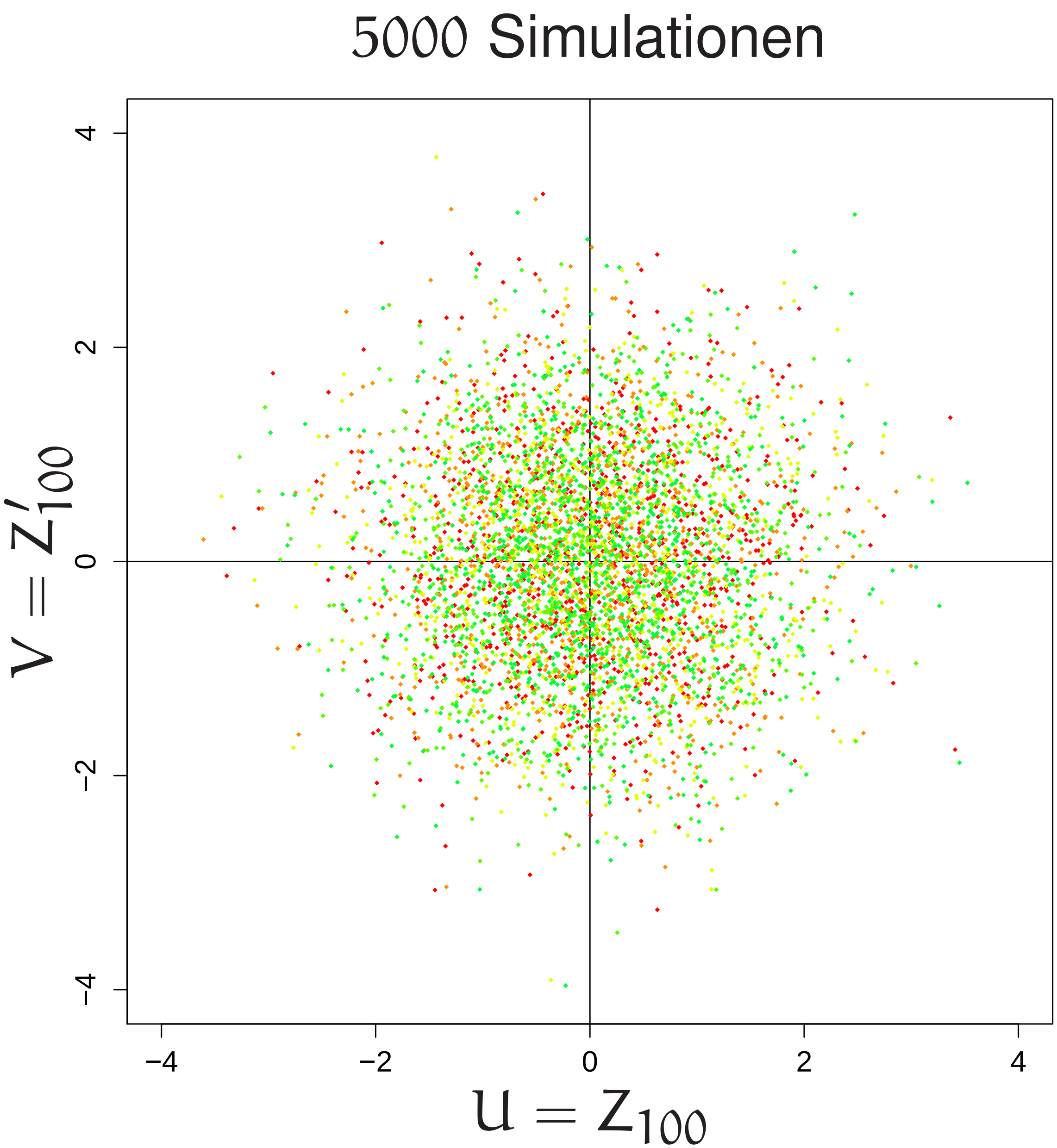
Wie sieht die gemeinsame Verteilung von U und V aus?

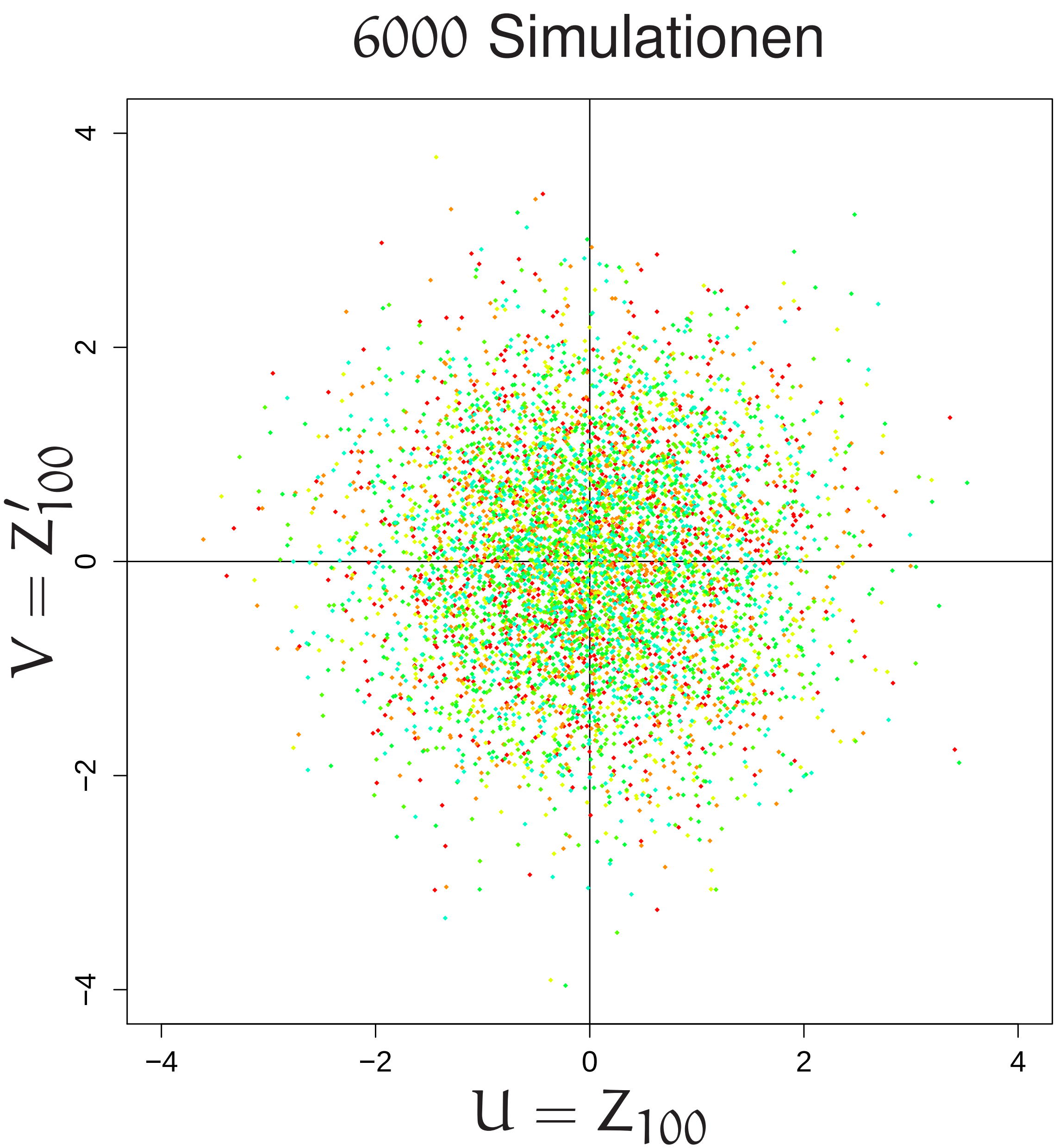

1000 Simulationen
4
2
0
−2
−4
−4
−2
0
2
4
$V = Z'_{100}$
$U = Z_{100}$

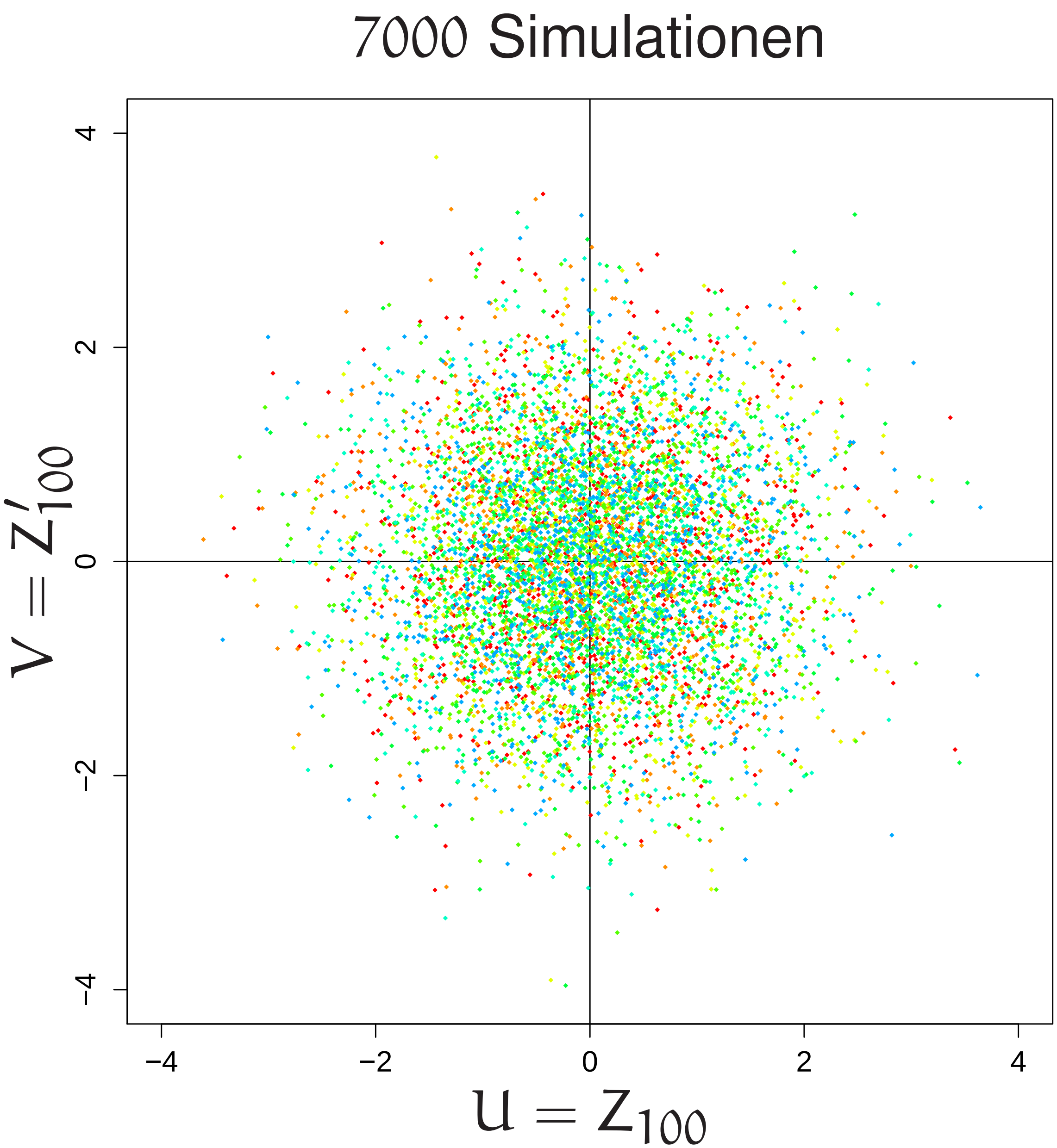

2000 Simulationen
$V = Z'_{100}$
$U = Z_{100}$
−4
−2
0
2
4

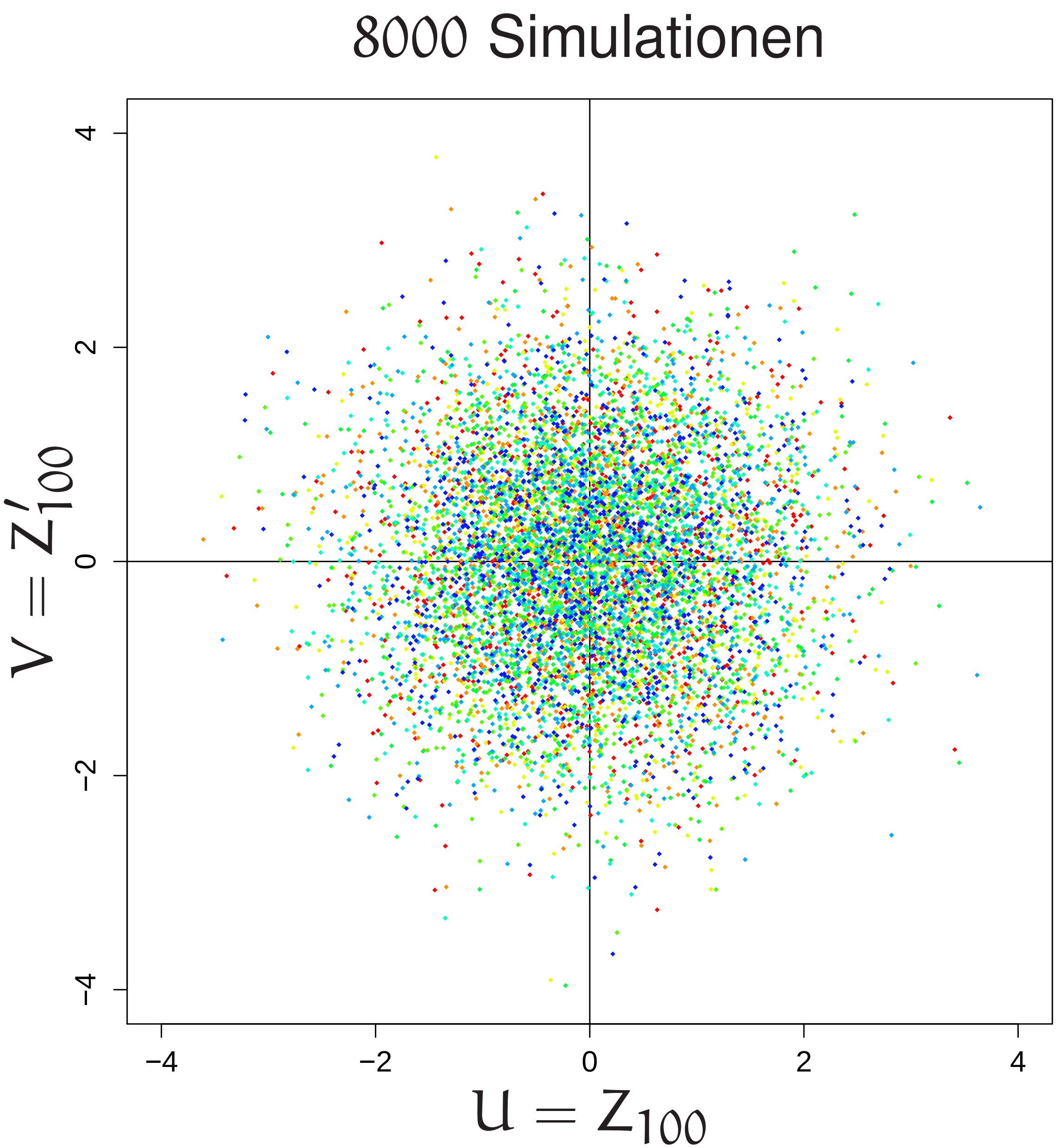

3000 Simulationen
4
2
0
−2
−4
−4
−2
0
2
4
$V = Z'_{100}$
$U = Z_{100}$

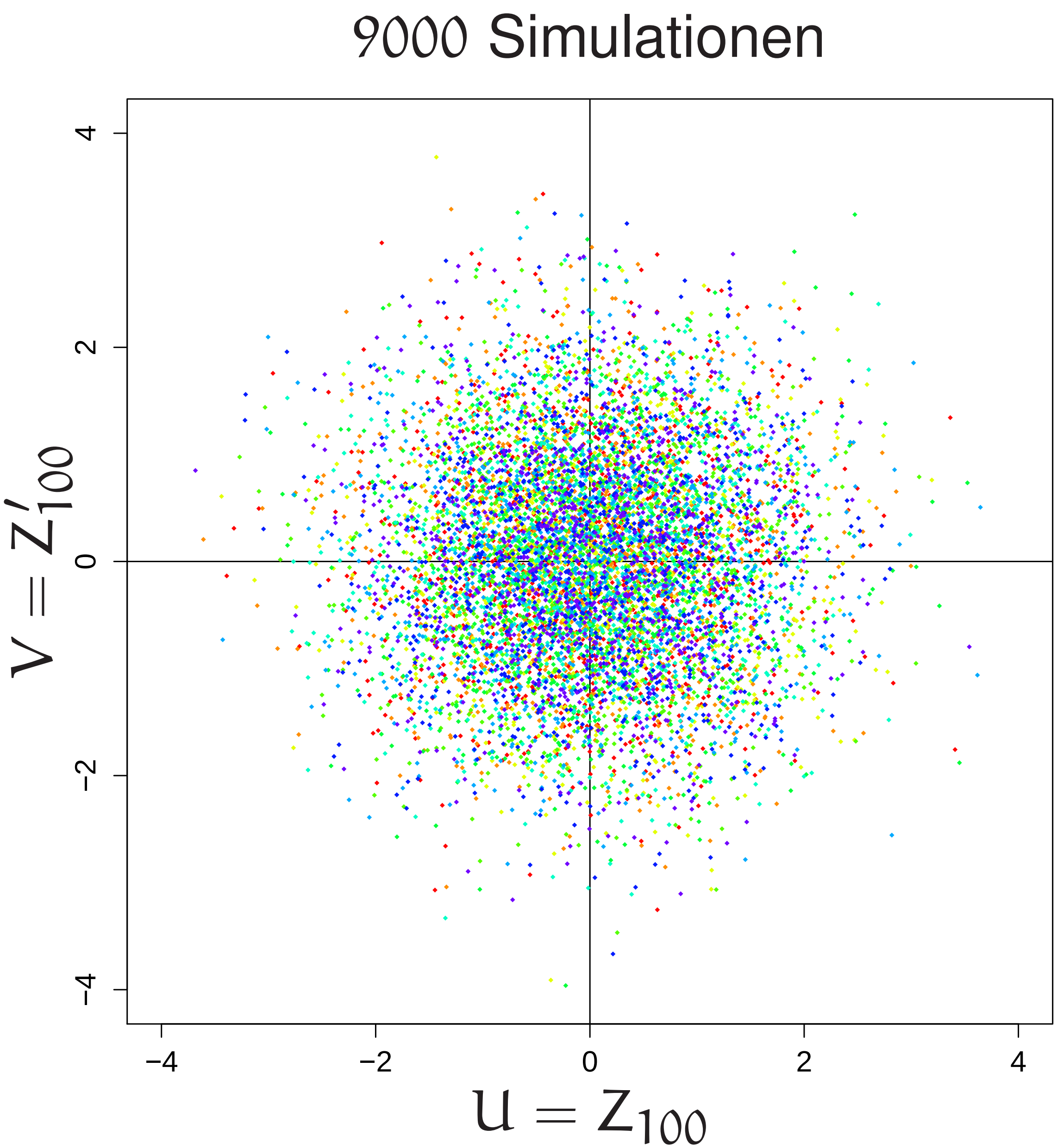

4000 Simulationen
4
2
0
−2
−4
−4
−2
0
2
4
$V = Z'_{100}$
$U = Z_{100}$

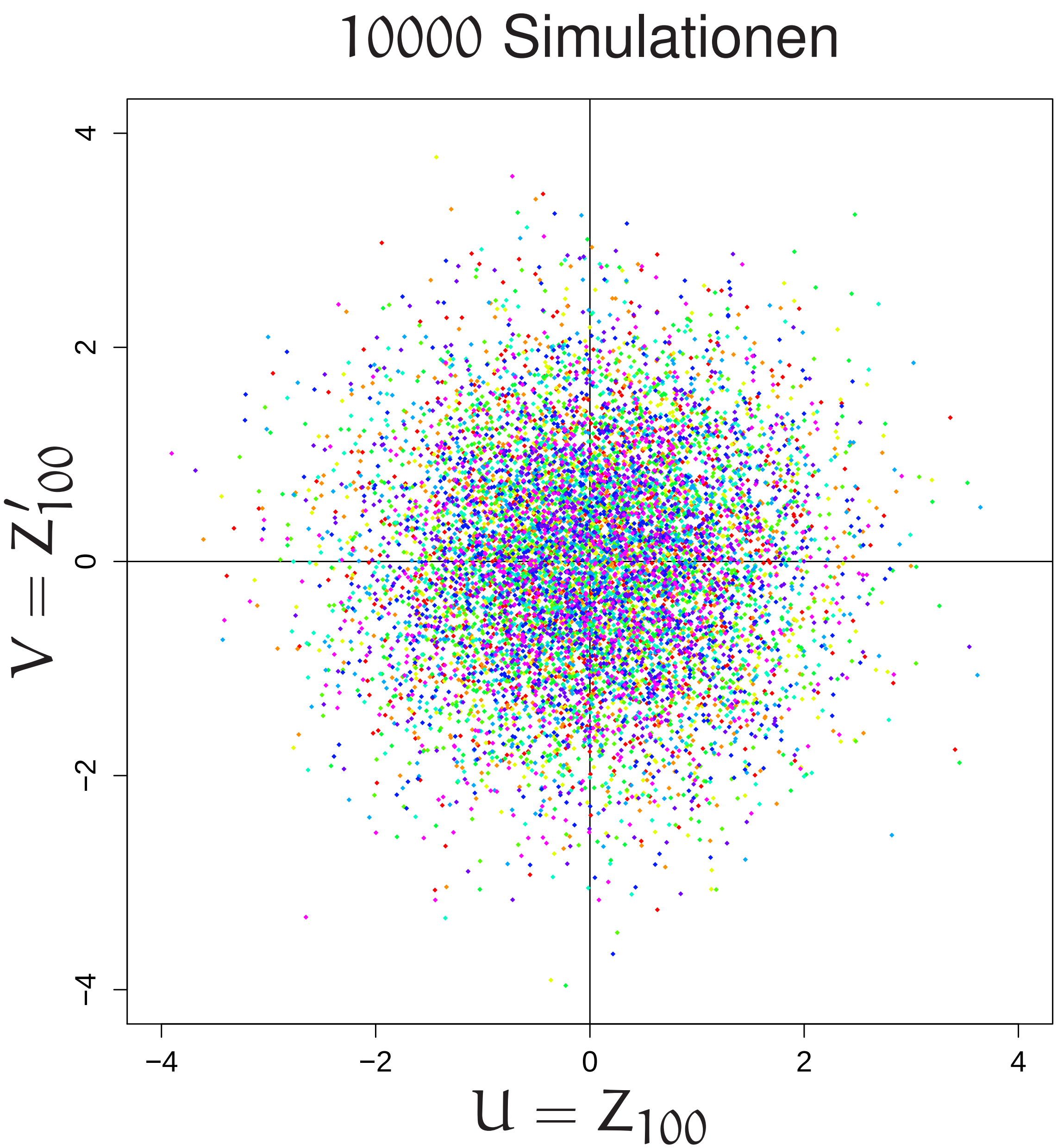

5000 Simulationen
$V = Z'_{100}$
$U = Z_{100}$
−4
−2
0
2
4


6000 Simulationen
$V = Z'_{100}$
$U = Z_{100}$
−4
−2
0
2
4

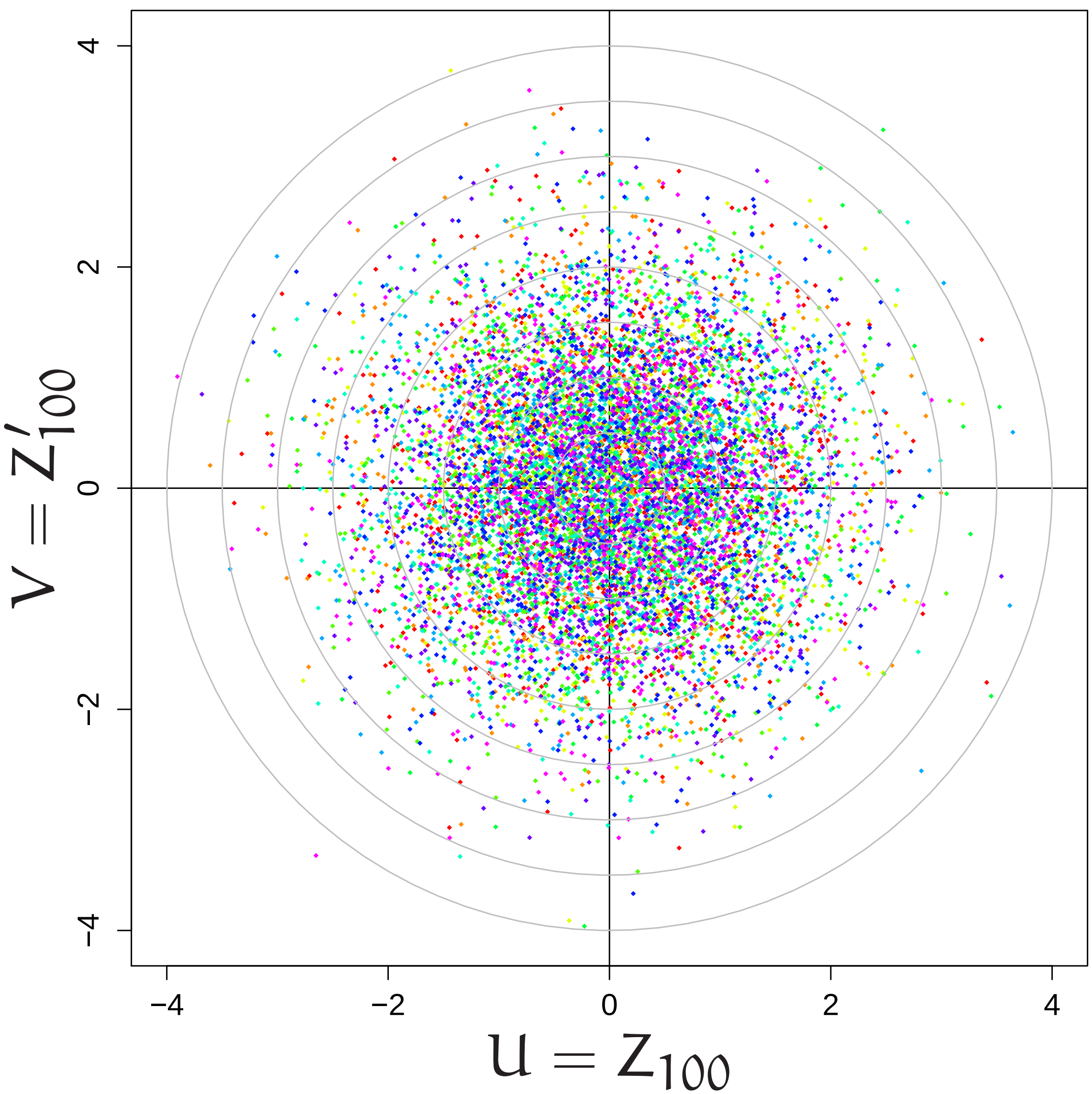

7000 Simulationen
4
2
0
−2
−4
−4
−2
0
2
4
$V = Z'_{100}$
$U = Z_{100}$


8000 Simulationen
$V = Z'_{100}$
$U = Z_{100}$
−4
−2
0
2
4

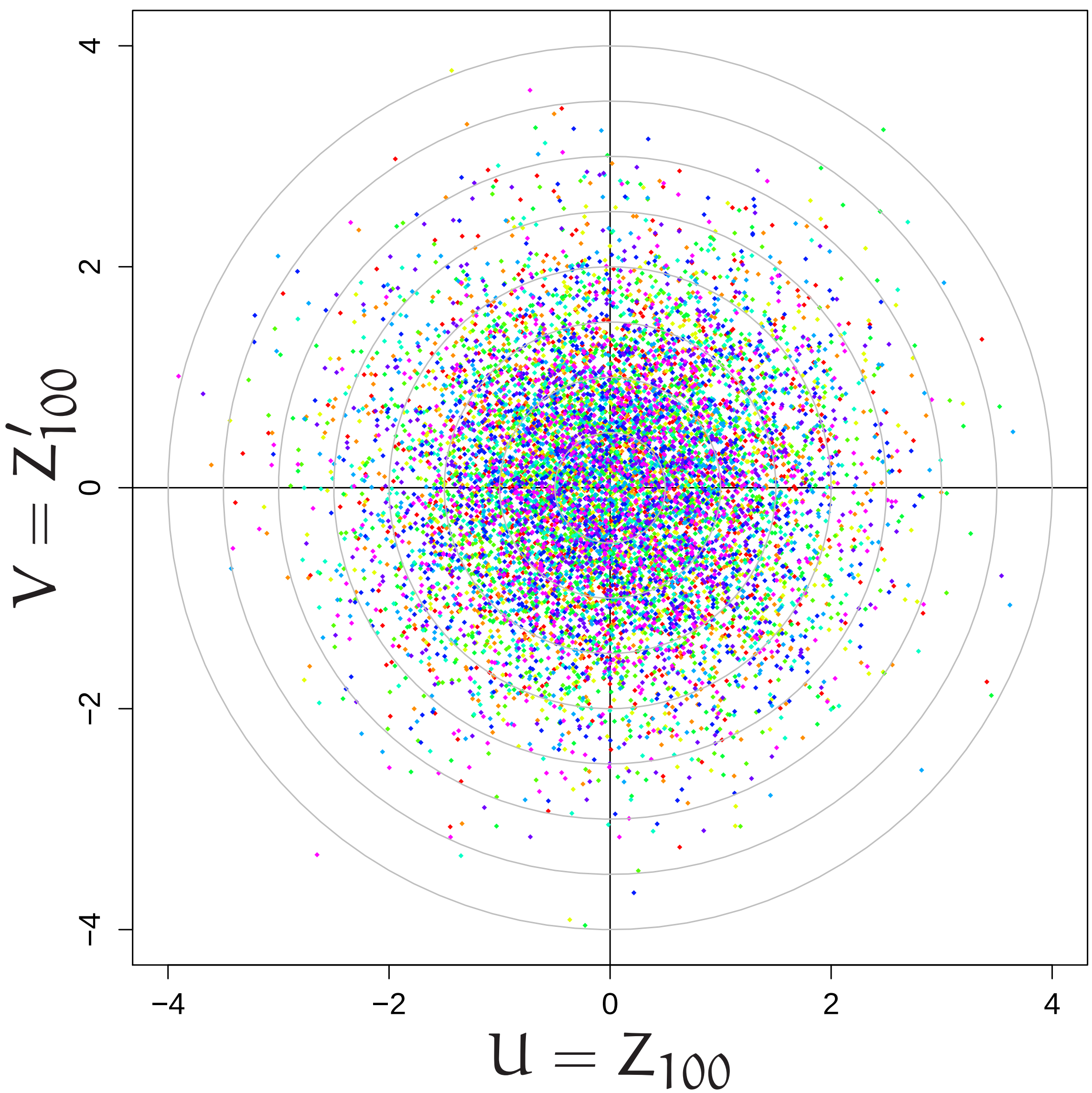

9000 Simulationen
4
2
0
−2
−4
$V = Z'_{100}$
$U = Z_{100}$


10000 Simulationen
4
2
0
−2
−4
$V = Z'_{100}$
−4
−2
0
2
4
$U = Z_{100}$

10000 Simulationen



Die Verteilung von (U, V) ist annähernd rotationssymmetrisch!



5. Eine Charakterisierieung der zweidimensionalen Standardnormalverteilung

Behauptung:

Aus "U und V unabhängig und identisch verteilt'

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

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

Mit $f_U = f_V =: h$ folgt

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

$$h(a)\,h(b) = g(r), \qquad r := \sqrt{a^2 + b^2}$$

Die zwei Paare (a, b) und $(0, \sqrt{a^2 + b^2})$ haben dasselbe r.

Also:

$$h(a)\,h(b) = h(0)\,h(\sqrt{a^2 + b^2})$$

Eine Lösung hiervon:

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

Denn

$$e^{-a^2}e^{-b^2} = 1 \cdot e^{-(a^2+b^2)}$$

$$h(a)\,h(b) = h(0)\,h(\sqrt{a^2+b^2})$$

Wie sieht die allgemeine Lösung aus?

$w(u) := h(\sqrt{u})$, $u \geq 0$, erfüllt

$$w(a^2)w(b^2) = w(0)w(a^2+b^2)$$

$$w(u)w(v) = k_0 w(u+v), \quad u, v \geq 0$$

hat als allgemeine Lösung

$$w(u) = k_0 e^{-k_1 u}$$

$$h(a) = w(a^2) = k_0 e^{-k_1 a^2}$$

FAZIT

Der Zentrale Grenzwertsatz

lässt sich erraten

(in konkreten Fällen,

mit etwas Glück).

Hier ist noch einmal die (im ZGS präzisierte) 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.

Eine Botschaft zum Mitnehmen ins Leben:

“Die Summe von vielen kleinen,

annähernd unabhängigen Zufallsvariablen

ist annähernd normalverteilt.”

6. Ein Beweis des klassischen Zentralen Grenzwertsatzes

Der Zentraler Grenzwertsatz in seiner klassischen Form:

Die standardisierte Summe von unabhängigen,

identisch verteilten $\mathbb{R}$-wertigen Zufallsvariablen

mit endlicher Varianz

konvergiert in Verteilung

gegen eine standard-normalverteilte Zufallsvariable.

Formal:

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

$$\mathbf{P}\Big(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}} \in [c, d]\Big) \underset{n\to\infty}{\longrightarrow} \mathbf{P}(Z \in [c, d]).$$

Dabei ist Z standard-normalverteilt.

$X_1, X_2, \ldots$ seien identisch verteilte rellwertige Zufallsvariable mit endlicher Varianz.

Ohne Einschränkung können wir annehmen:

$$\mathbf{E}[X_i] = 0, \quad \mathbf{Var} X_i = 1.$$

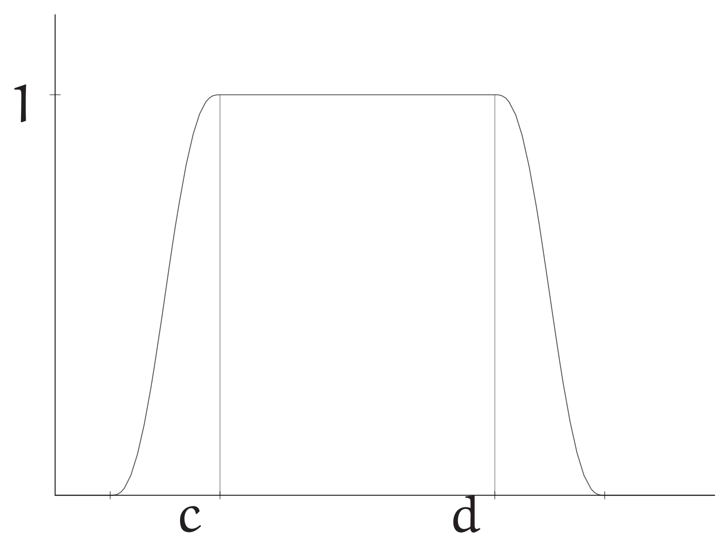
(Denn sonst gehen wir einfach zu den standardisierten Zufallsvariablen $\frac{X_i - \mu}{\sigma}$ über.)

Die Behauptung ist dann:

$$\mathbf{E}\Big[\mathbf{1}_{[c,d]}\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big)\Big] \underset{n\to\infty}{\longrightarrow} \mathbf{E}[\mathbf{1}_{[c,d]}(Z)]$$

mit standard-normalverteiltem Z.

Weil man Indikatorfunktionen $\mathbf{1}_{[c,d]}$
durch “glatte” Funktionen h approximieren kann, reicht es
(wie man zeigen kann, mehr dazu später),
diese Konvergenz nur für solche h zu beweisen.



Lemma

(Buch S. 78)

Sei $h : \mathbb{R} \to \mathbb{R}$ dreimal stetig differenzierbar und seien h', h'' und h''' beschränkt. Dann gilt

$$\mathbf{E}\Big[h\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big)\Big] \underset{n\to\infty}{\longrightarrow} \mathbf{E}[h(Z)]$$

Zum Beweis des Lemmas: *

Die Hauptidee besteht darin, eine Folge von unabhängigen standard-normalverteilten Zufallsvariablen $(Z_1, Z_2, \dots)$ ins Spiel zu bringen, die zusammen mit $(X_1, X_2, \dots)$ ein zufälliges Paar von Folgen bilden.

Dabei seien alle $Z_1, Z_2, \dots, X_1, X_2, \dots$ unabhängig.

Wir wissen schon, dass gilt:

$$\mathbf{E}[h(Z)] = \mathbf{E}\Big[h\Big(\frac{Z_1 + \cdots + Z_n}{\sqrt{n}}\Big)\Big].$$

*nach einer Idee von J. Lindeberg, * 1876, † 1932

Außerdem ergibt sich

mit der Linearität des Erwartungswertes:

$$\mathbf{E}\Big[h\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big)\Big] - \mathbf{E}\Big[h\Big(\frac{Z_1 + \cdots + Z_n}{\sqrt{n}}\Big)\Big]$$

$$= \mathbf{E}\Big[h\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big) - h\Big(\frac{Z_1 + \cdots + Z_n}{\sqrt{n}}\Big)\Big].$$

Es reicht also zu zeigen, dass letzteres

für $n \to \infty$ gegen Null konvergiert.

Eine clevere Idee ist es jetzt, die Differenz als Teleskopsumme darzustellen:

$$h\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)-h\left(\frac{Z_1+\cdots+Z_n}{\sqrt{n}}\right)$$

$$=\sum_{i=1}^{n}\Bigg(h\left(\frac{X_1+\cdots+X_{i-1}+X_i+Z_{i+1}+\cdots+Z_n}{\sqrt{n}}\right)$$

$$-h\left(\frac{X_1+\cdots+X_{i-1}+Z_i+Z_{i+1}+\cdots+Z_n}{\sqrt{n}}\right)\Bigg)$$

Taylorentwicklung ergibt:

$$= \sum_{i=1}^{n} \frac{X_i - Z_i}{\sqrt{n}} h'\left(\frac{X_1 + \cdots + X_{i-1} + Z_{i+1} + \cdots + Z_n}{\sqrt{n}}\right)$$

$$+ \sum_{i=1}^{n} \frac{X_i^2 - Z_i^2}{2n} h''\left(\frac{X_1 + \cdots + X_{i-1} + Z_{i+1} + \cdots + Z_n}{\sqrt{n}}\right)$$

$$+ \sum_{i=1}^{n} \left(\frac{X_i^3}{6n^{3/2}} h'''(Y_i) - \frac{Z_i^3}{6n^{3/2}} h'''(\tilde{Y}_i)\right)$$

mit passenden Zwischenstellen Y_i, $\tilde{Y}_i$.

Wir nehmen hier der Einfachheit halber an:

$$\mathbf{E}[|X_1|^3] < \infty.$$

(Der Fall ohne diese Zusatzbedingung

ist im Buch S. 78 abgehandelt.)

Ist C eine obere Schranke von $|h'''|$, so folgt

$$\Big|\mathbf{E}\Big[h\big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\big)-h\big(\frac{Z_1+\cdots+Z_n}{\sqrt{n}}\big)\Big]\Big| \\ \le n\frac{1}{6n^{3/2}}(\mathbf{E}[|X_1|^3]+\mathbf{E}[|Z_1|^3])C \longrightarrow 0,$$

denn die Erw. werte der ersten beiden Summen sind Null

wegen $\mathbf{E}[X_1]=\mathbf{E}[Z_1]=0$ und $\mathbf{E}[X_1^2]=\mathbf{E}[Z_1^2]=1$,

zusammen mit der Unabhängigkeit der X_i, Z_i

und der Produktformel

für die Erwartungswerte unabhängiger Zufallsvariabler. $\square$

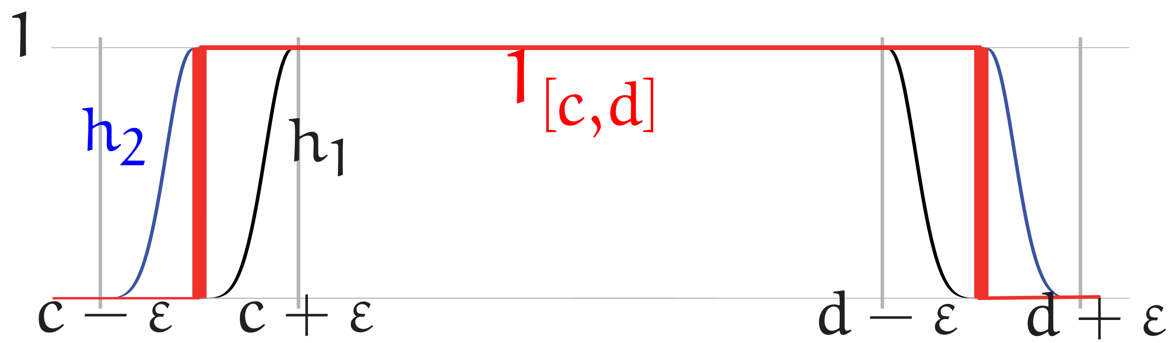
Wir folgern jetzt die Ausage des Zentralen Grenzwertsatzes

aus dem Lemma,

uind zwar mittels der vesrprochenen Approximation

der Indikatorfunktionen $\mathbf{1}_{[c,d]}$

durch hinreichend glatte Funktionen h:



$$\mathbf{1}_{[c+\varepsilon,d-\varepsilon]} \le h_1 \le \mathbf{1}_{[c,d]} \le h_2 \le \mathbf{1}_{[c-\varepsilon,d+\varepsilon]}$$

$$\begin{aligned}
&\mathbf{P}(c+\varepsilon \le Z \le d-\varepsilon) \\
&\le \mathbf{E}[h_1(Z)] = \lim_{n\to\infty} \mathbf{E}\left[h_1\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)\right] \\
&\le \liminf_{n\to\infty} \mathbf{P}\left(c \le \frac{X_1+\cdots+X_n}{\sqrt{n}} \le d\right) \\
&\le \limsup_{n\to\infty} \mathbf{P}\left(c \le \frac{X_1+\cdots+X_n}{\sqrt{n}} \le d\right) \\
&\le \lim_{n\to\infty} \mathbf{E}\left[h_2\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)\right] = \mathbf{E}[h_2(Z)] \\
&\le \mathbf{P}(c-\varepsilon \le Z \le d+\varepsilon)
\end{aligned}$$

$$\mathbf{P}(c+\varepsilon \le Z \le d-\varepsilon)$$

$$\le \liminf_{n\to\infty} \mathbf{P}\left(c \le \frac{X_1+\cdots+X_n}{\sqrt{n}} \le d\right)$$
$$\le \limsup_{n\to\infty} \mathbf{P}\left(c \le \frac{X_1+\cdots+X_n}{\sqrt{n}} \le d\right)$$

$$\le \mathbf{P}(c-\varepsilon \le Z \le d+\varepsilon).$$

Da Z eine Dichte besitzt, gilt

$\mathbf{P}(c \pm \varepsilon \le Z \le d \mp \varepsilon) \to \mathbf{P}(c \le Z \le d)$ für $\varepsilon \to 0$,

und die Behauptung folgt. $\square$

7. Ein Beweis der Stirling-Formel

Vorbemerkung

Die Aussage aus dem Lemma

$$\mathbf{E}\Big[h\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big)\Big] \underset{n\to\infty}{\longrightarrow} \mathbf{E}[h(Z)]$$

überträgt sich auf alle Funktionen h, die sich

(ähnlich wie dort die Funktionen $\mathbf{1}_{[c,d]}$)

durch “glatte” Funktionen approximieren lassen.

Ein Beispiel ist die Funktion $a \mapsto a^+ := \max(0, a)$:

Für unabhängige, identisch verteilte $X_1, X_2, \dots$ mit Erwartungswert 1 und Varianz 1 und $N(0,1)$-verteiltes Z gilt:

$$\mathbf{E}\Big[\Big(\frac{X_1 + \cdots + X_n - n}{\sqrt{n}}\Big)^+\Big] \to \mathbf{E}[Z^+]$$

Wählen wir $X_1, X_2, \ldots$ als Poissonverteilt zum Parameter 1, so ist (Übung!)

$X_1 + \cdots + X_n$ Poissonverteilt zum Parameter n,

und folglich

$$\mathbf{E}\Big[\Big(\frac{X_1 + \cdots + X_n - n}{\sqrt{n}}\Big)^+\Big] = \frac{1}{\sqrt{n}} \sum_{k=n+1}^{\infty} (k-n)e^{-n}\frac{n^k}{k!}$$

$$= \frac{e^{-n}}{\sqrt{n}} \sum_{k=n+1}^{\infty} \Big(\frac{n^k}{(k-1)!} - \frac{n^{k+1}}{k!}\Big)$$

$$= \frac{e^{-n}}{\sqrt{n}}\frac{n^{n+1}}{n!}\ .$$

Andererseits gilt:

$$\mathbf{E}[Z^+] = \frac{1}{\sqrt{2\pi}} \int_0^\infty z e^{-z^2/2}\, dz = \frac{1}{\sqrt{2\pi}} \left[e^{-z^2/2}\right]_0^\infty = \frac{1}{\sqrt{2\pi}} \,.$$

Insgesamt haben wir gefunden

$$\frac{e^{-n}}{\sqrt{n}} \frac{n^{n+1}}{n!} \to \frac{1}{\sqrt{2\pi}}$$

und damit eine Herleitung der Stirling-Formel

aus dem Zentralen Grenzwertsatz gewonnen. $\square$

8. Rückblick und Einordnung:

Münzwurf und Zentraler Grenzwertsatz

Der Münzwurf passt in den Zentralen Grenzwertsatz:

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

$$\mathbf{P}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}} \in [c, d]\right) \underset{n\to\infty}{\longrightarrow} \mathbf{P}(Z \in [c, d]).$$

Dabei ist Z standard-normalverteilt.

Ist X_i eine Bernoullifolge (mit $\mu := p$ und $\sigma^2 := pq$), so ergibt sich der alte Satz von de Moivre und Laplace.