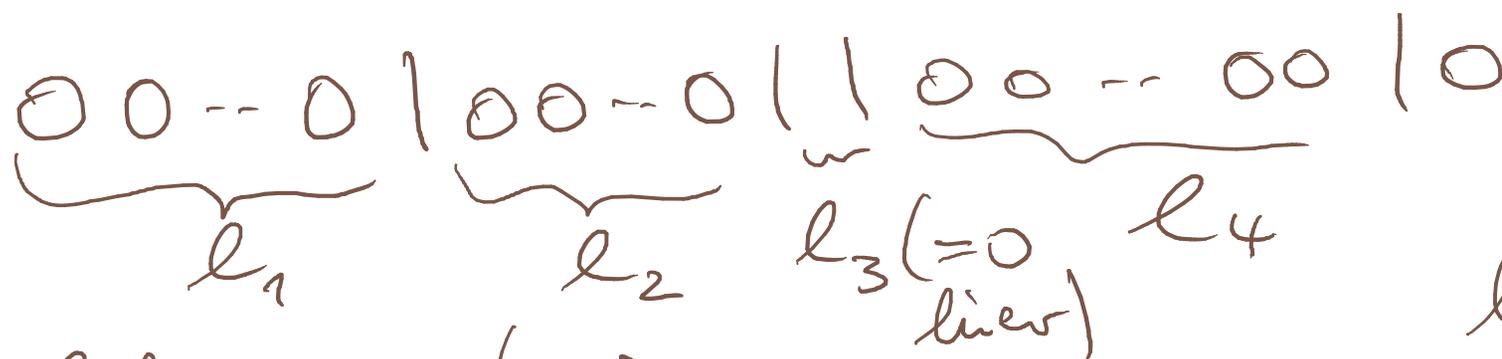


## Bemerkung 1.14. Es gilt

$$|W_4| = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

(Insbesondere ist die Verteilung auf  $W_4$  aus Beispiel 1.13, 4. nicht die uniforme.)

Zähltrick: lege  $k$  „Kugeln“ und  
 $n-1$  „Trennstäbe“ in eine Reihe



# solcher Muster:  
 $n-1+k$  Objekte ( $k$  Positionen)

Die uniforme Verteilung auf dem  $W_4$  aus Beispiel 1.13, 4. heißt auch die „Bose-Einstein-Verteilung“, die in Beispiel 1.13, 4. betrachtete Verteilung heißt die „Maxwell-Boltzmann-Verteilung“.

**Beispiel.** Eine Hörsaalreihe habe  $n$  Plätze, darauf nehmen  $m$  ( $\leq n/2$ ) Männer und  $n - m$  Frauen rein zufällig Platz.

Die Wahrscheinlichkeit, dass (in diesem Modell) keine zwei Männer nebeneinander sitzen

**Beispiel 1.15** (Hypergeometrische Verteilung). Eine Urne enthalte  $n$  Kugeln, davon  $s$  schwarze und  $w$  weiße ( $s + w = n$ ), ziehe  $k$ -mal ohne Zurücklegen,

$$\text{Hyp}_{s,w,k}(\{l\}) = \frac{\binom{s}{l} \binom{w}{k-l}}{\binom{s+w}{k}}, \quad l = 0, 1, \dots, k$$

ist die W'keit, genau  $l$  schwarze Kugeln zu ziehen.

( Nummeriere: Kugeln  $1, \dots, s$  Schwarz  
— " —  $s+1, s+2, \dots, s+w$  Weiß

$\binom{s}{l}$  Wahlen,  $l$  Kugeln mit Nummer  $\leq s$   
und  $\binom{w}{k-l}$  — " —  $k-l$  — " —  $> s$   
auszuwählen.

### Beispiel 1.16 ( $p$ -Münzwurf und Binomialverteilung).

1.  $S = \{0, 1\}$ ,  $\text{Ber}_p(\{1\}) = p = 1 - \text{Ber}_p(\{0\})$  mit einem  $p \in [0, 1]$   
 („Bernoulli-Verteilung“<sup>1</sup>)

2.  $n$ -facher  $p$ -Münzwurf (mit  $p \in [0, 1]$ ):  $S = \{0, 1\}^n$ ,

$$\text{Ber}_p^{\otimes n}(\{(x_1, \dots, x_n)\}) = p^{|\{i \leq n : x_i = 1\}|} (1-p)^{|\{i \leq n : x_i = 0\}|} \quad (p + 1 - p)^n = 1^n$$

Für  $X = (X_1, \dots, X_n) \sim \text{Ber}_p^{\otimes n}$  sagt man:  $X_1, \dots, X_n$  unabhängig und  
 $X_i \sim \text{Ber}_p$  für  $i = 1, 2, \dots, n$

3. Binomialverteilung (zum Parameter  $n$  und  $p$ ,  $n \in \mathbb{N}$ ,  $p \in [0, 1]$ ):

$$\text{Bin}_{n,p}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in S := \{0, 1, \dots, n\}$$

(dies ist die W'keit, beim  $n$ -fachen Münzwurf genau  $k$  Erfolge zu beobachten)

<sup>1</sup>nach Jakob Bernoulli, 1654–1705

**Beispiel 1.16, 3. mit Zufallsvariablen ausgesprochen:**

$$X = (X_1, \dots, X_n) \sim \text{Ber}_p^{\otimes n}, \quad Y := X_1 + X_2 + \dots + X_n,$$

so ist  $Y \sim \text{Bin}_{n,p}$ , denn für  $k \in \{0, 1, \dots, n\}$  und für jede Wahl von  $(x_1, \dots, x_n) \in \{0, 1\}^n$  mit  $x_1 + \dots + x_n = k$  ist

$$P(X_1 = x_1, \dots, X_n = x_n) = p^k (1-p)^{n-k}, \text{ somit mit}$$

$$B_k = \{(x_1, \dots, x_n) \in \{0, 1\}^n : x_1 + \dots + x_n = k\}$$

$$\begin{aligned} P(Y = k) &= P((X_1, \dots, X_n) \in B_k) \\ &= \sum_{(x_1, \dots, x_n) \in B_k} P((X_1, \dots, X_n) = (x_1, \dots, x_n)) \\ &= \sum_{(x_1, \dots, x_n) \in B_k} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k} \end{aligned}$$

da  $\#B_k = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} = \binom{n}{k}$ .

Insoweit ist dies eine Beispiel-Instanz zu Beobachtung 1.11.

**Beispiel 1.17 (Geometrische Verteilung).**  $p \in (0, 1)$ ,  $S = \mathbb{N}_0$ ,

$$\text{Geom}_p(\{j\}) = p(1-p)^j, \quad j \in \mathbb{N}_0$$

ist die W'keit, bei wiederholtem  $p$ -Münzwurf genau  $j$  Misserfolge vor dem ersten Erfolg zu beobachten.

Beachte: Manche Autoren betrachten die geometrische Verteilung auf  $\mathbb{N}$  (statt auf  $\mathbb{N}_0$ ), dann ist das Gewicht  $p(1-p)^{k-1}$  und die Interpretation „ $k$  Würfe (einschließlich) bis zum ersten Erfolg.“

beob.  $\therefore$

$$\sum_{j=0}^{\infty} p(1-p)^j = p \underbrace{\sum_{j=0}^{\infty} (1-p)^j}_{=1} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1$$

**Beispiel 1.18 (Multinomialverteilung).**  $s \in \{2, 3, \dots\}$ ,  $p_1, \dots, p_s \in [0, 1]$ ,  
 $p_1 + \dots + p_s = 1$ ,  $n \in \mathbb{N}$ ,  $S = \{(k_1, \dots, k_s) \in \mathbb{N}_0^s : k_1 + \dots + k_s = n\}$ ,

$$\text{Mult}_{n;p_1, \dots, p_s}(\{(k_1, \dots, k_s)\}) = \binom{n}{k_1, k_2, \dots, k_s} p_1^{k_1} p_2^{k_2} \dots p_s^{k_s}$$

Interpretation:  $n$  Züge mit Zurücklegen ohne Beachtung der Reihenfolge aus einer Urne mit  $s$  Kugeln ( $s$  verschiedene „Farben“), Farbe  $i$  wird mit W'keit  $p_i$  gezogen), obiges ist die W'keit, genau  $k_i$ -mal Farbe  $i$  zu ziehen für  $i = 1, 2, \dots, s$ .

**Beispiel 1.19** (Poissonverteilung\*).  $\lambda \in (0, \infty)$ ,

$$\text{Poi}_\lambda(\{k\}) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N}_0$$

beob.:

$$\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$
$$= e^{-\lambda + \lambda} = e^0 = 1 \quad \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{= e^\lambda}$$

**Proposition 1.20** (Poissonapproximation der Binomialverteilung).

Seien  $p_n \in [0, 1]$  mit  $p_n \rightarrow 0$  und  $np_n \rightarrow \lambda \in (0, \infty)$  für  $n \rightarrow \infty$ , so gilt für jedes  $k \in \mathbb{N}_0$

$$\text{Bin}_{n,p_n}(\{k\}) \xrightarrow{n \rightarrow \infty} \text{Poi}_\lambda(\{k\}).$$

z.B.  $p_n = \frac{\lambda}{n}$

denk:

$$\binom{n}{k} p_n^k (1-p_n)^{n-k}$$

$$= \frac{n(n-1)\dots \cdot 2 \cdot 1}{k! (n-k)(n-k)\dots \cdot 2 \cdot 1}$$

$$= \frac{\underbrace{n(n-1)\dots \cdot (n-k+1)}_{n^k} \cdot \underbrace{k!}_{k!}}{k! \cdot \underbrace{n^k}_{n^k}}$$

$n \rightarrow \infty \downarrow 1$

$$e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= \underbrace{\binom{n}{k} p_n^k}_{\lambda^k} \underbrace{\left(1 - \frac{n-p_n}{n}\right)^n}_{e^{-\lambda}} \underbrace{(1-p_n)^{n-k}}_{1 = 1^k}$$

## Proposition 1.20:

$$np_n \rightarrow \lambda \in (0, \infty), \quad \text{so gilt} \quad \text{Bin}_{n,p_n}(k) \xrightarrow[n \rightarrow \infty]{} \text{Poi}_\lambda(k)$$

Prop. 1.20 motiviert, warum die Poissonverteilung oft in Anwendungssituationen vorkommt, in denen man viele unabhängige Ereignisse betrachtet, von denen jedes nur mit einer sehr kleinen W'keit eintritt.

Man denke etwa an Schadensfälle bei Versicherungen, Zerfallsergebnisse in einer Probe radioaktiven Materials oder an genetische Mutationen.

**Beispiel 1.21.** L. von Bortkewitsch berichtete in seinem Buch *Das Gesetz der kleinen Zahlen*, Teubner, 1898 verschiedene Datensätze, die gut zur Poissonverteilung passen.

Speziell in § 12, 4. („Die durch Schlag eines Pferdes im preußischen Heere getöteten“) werden für 20 Jahre (1875–1894) und 10 Armee-kops der preußischen Kavallerie, also insgesamt  $20 \cdot 10 = 200$  „Korps-jahre“ berichtet, in wievielen davon sich  $x$  Todesfälle durch Schlag eines Pferds ereigneten (Tabelle b) auf S. 25):

Ergebnis $x$	Anz. „Korpsjahre“
0	109
1	65
2	22
3	3
4	1
$\geq 5$	0

## Beispiel 1.21, Forts.

Angenommen, die Anzahl durch Schlag eines Pferdes während eines Jahres in einem Korps getöteter Soldaten wäre  $Poi_{\lambda}$ -verteilt mit  $\lambda = 0,61$ , so würden wir das Resultat  $x$  je  $200 \times Poi_{0,61}(x)$ -mal erwarten:

Ergebnis $x$	Anz. „Korpsjahre“	$200 \times Poi_{0,61}(x)$
0	109	108,67
1	65	66,29
2	22	20,22
3	3	4,11
4	1	0,63
$\geq 5$	0	0,08

Von Bortkewitsch, a.a.O., S. 25 schreibt: „Die Kongruenz der Theorie mit der Erfahrung lässt [...], wie man sieht, nichts zu wünschen übrig.“

## Beispiel 1.21, Forts. 2

Übrigens: Wie ist von Bortkewitsch auf  $\lambda = 0,61$  gekommen?

Die beobachtete „mittlere Anzahl Todesfälle pro Korpsjahr“ in den Daten ist

$$\widehat{\lambda} = \frac{109}{200} \cdot 0 + \frac{65}{200} \cdot 1 + \frac{22}{200} \cdot 2 + \frac{3}{200} \cdot 3 + \frac{1}{200} \cdot 4 + 0 = 0,61$$

und es ist auch

$$\sum_{x=0}^{\infty} x \text{Poi}_{\lambda}(x) = \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} = \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} e^{-\lambda} = \lambda$$

(„der Erwartungswert von  $\text{Poi}_{\lambda}$  ist  $\lambda$ “) und somit ist obiges der nahe-  
liegende „Momentenschätzer“ – wir werden darauf zurückkommen.

## Verteilung (Fall mit Dichte)

Zufallsvariablen mit Dichten sind ein kontinuierliches Analogon zu Zufallsvariablen mit Gewichten. In vielen Situationen ist eine Modellierung eines zufälligen Werts  $X$  als „allgemeine“ reelle Zahl angemessen, d.h. die Annahme, dass der Wertebereich  $S$  diskret ist, ist zu „eng“.

(Auch wenn man argumentieren könnte, dass die Menge der im Rechner mit gegebener Genauigkeit darstellbaren Werte prinzipiell diskret ist, ist es oft „unpraktisch“, sich immer auf eine konkrete Diskretisierung festlegen zu müssen.)

## Beispiel 1.22 (Approximation der Exponentialverteilung durch reskalierte geometrisch verteilte ZVn).

Sei  $W \sim \text{Geom}_p$  mit  $p \ll 1$ ,  $X := pW$  (hat Werte in  $p\mathbb{Z}_+ \subset \mathbb{R}$ ).

### Bsp. 1.22, Forts.

$$P(pW \geq x) = (1-p)^{\lceil x/p \rceil} \approx (1-p)^{x/p} \approx e^{-x}$$

$$\begin{array}{c} \nearrow \\ W \geq \frac{x}{p} \end{array}$$

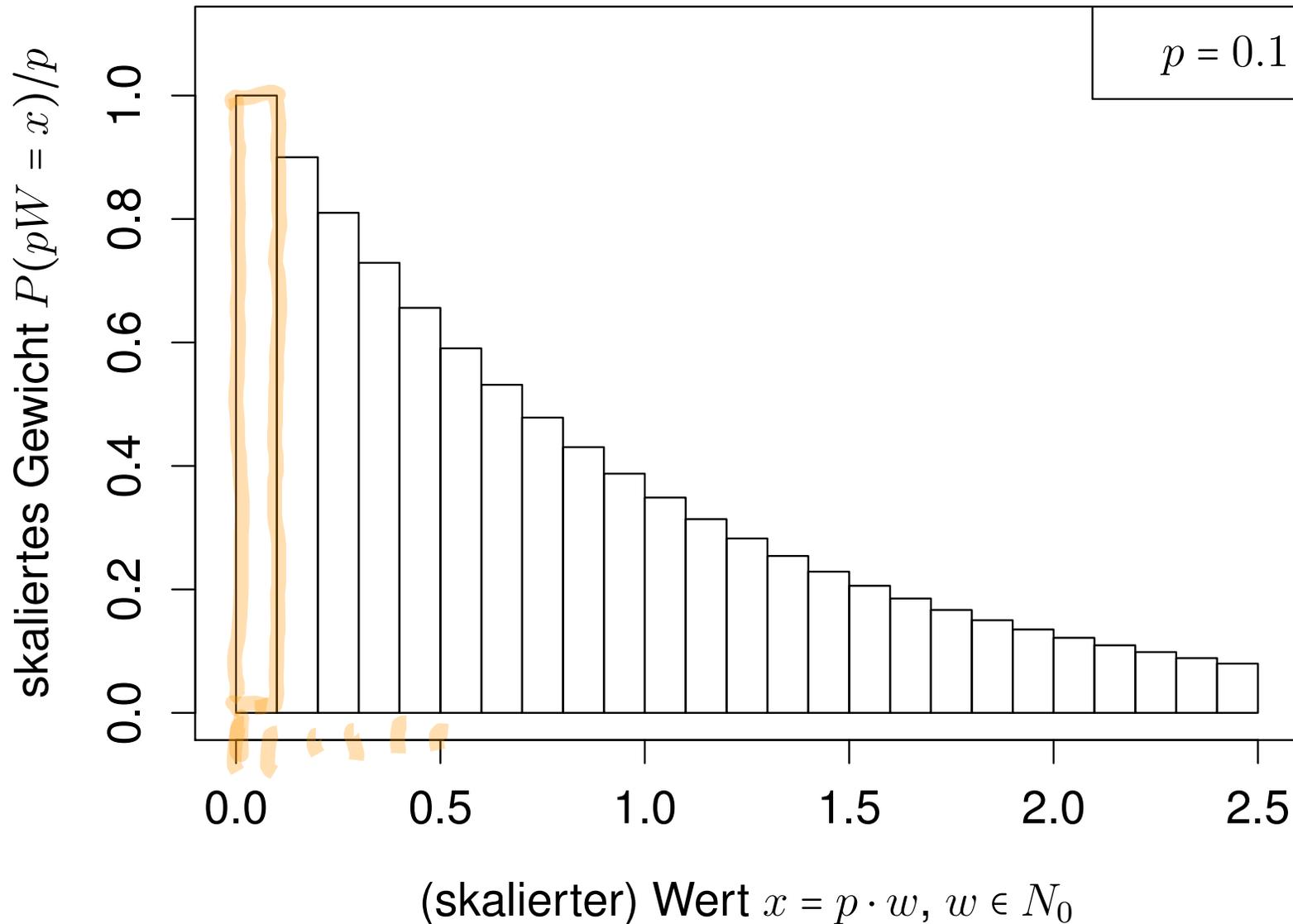
$$= \left( (1-p)^{\frac{1}{p}} \right)^x \xrightarrow{\quad} e^{-x}$$
$$n := \frac{1}{p} \quad \left(1 - \frac{1}{n}\right)^n \xrightarrow{\quad} e^{-1}$$

Interpretation z.B.:

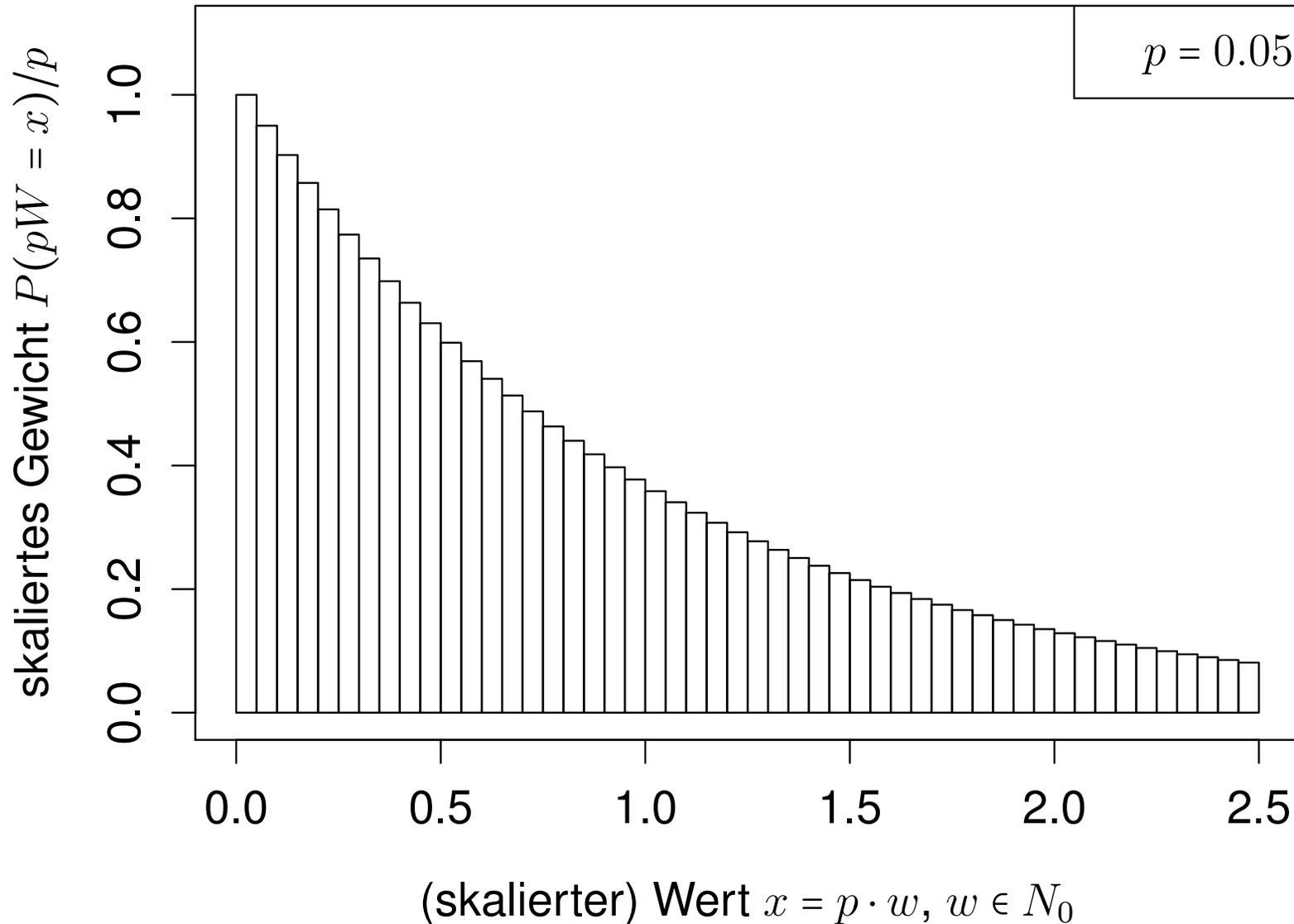
via Wartezeiten auf sehr fein zeitdiskretisiertem Gitter

Frage:

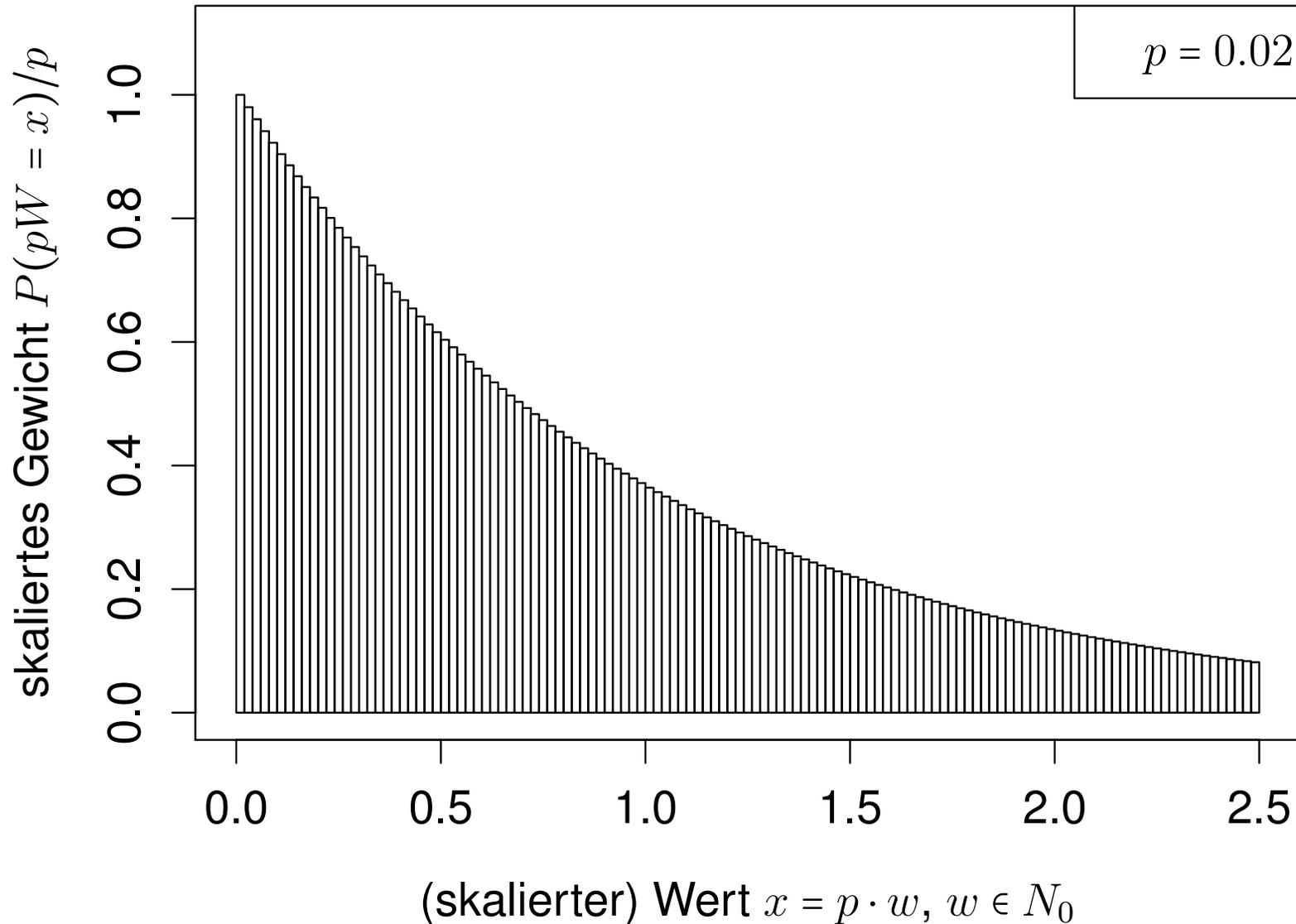
Gibt es eine reellwertige ZV  $X$ , für die obiges  $(P(X \geq x) = e^{-x})$  als Identität gilt?



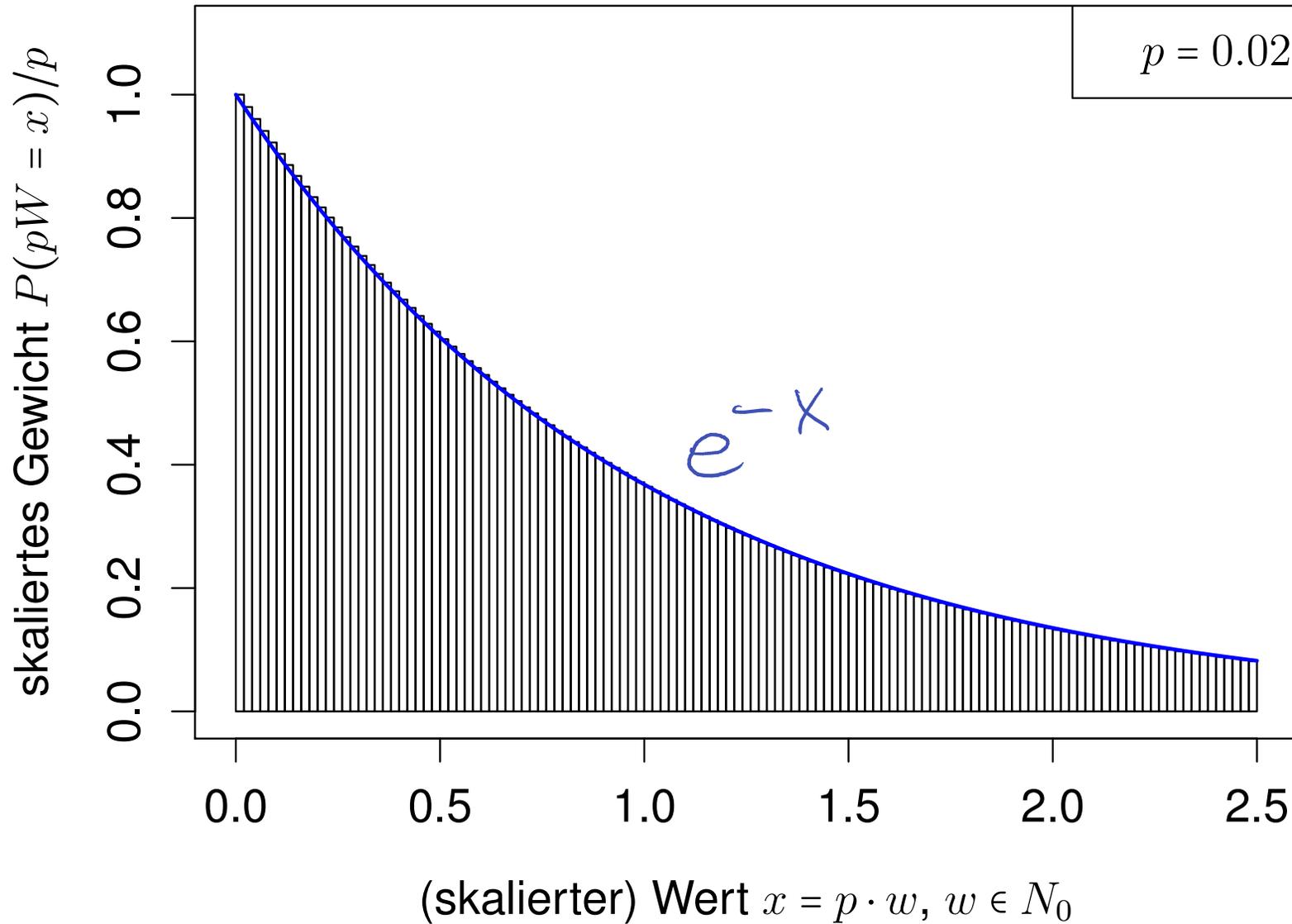
Die Höhe der Balken ist so skaliert, dass die Fläche des Balkens bei  $x = pw$  gerade  $P(pW = x) = p \cdot P(pW = x)/p$  entspricht.



Die Höhe der Balken ist so skaliert, dass die Fläche des Balkens bei  $x = pw$  gerade  $P(pW = x) = p \cdot P(pW = x)/p$  entspricht.

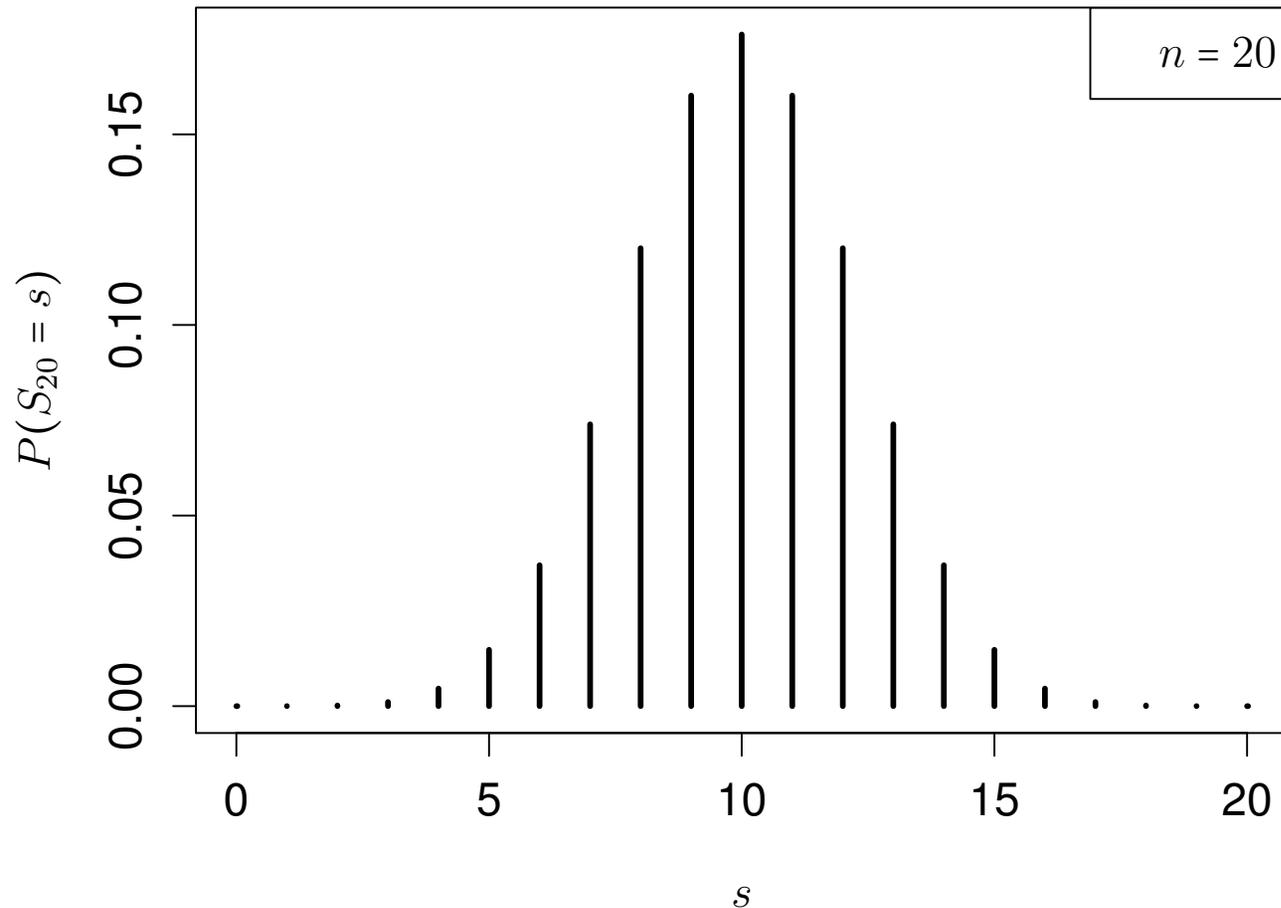


Die Höhe der Balken ist so skaliert, dass die Fläche des Balkens bei  $x = pw$  gerade  $P(pW = x) = p \cdot P(pW = x)/p$  entspricht.

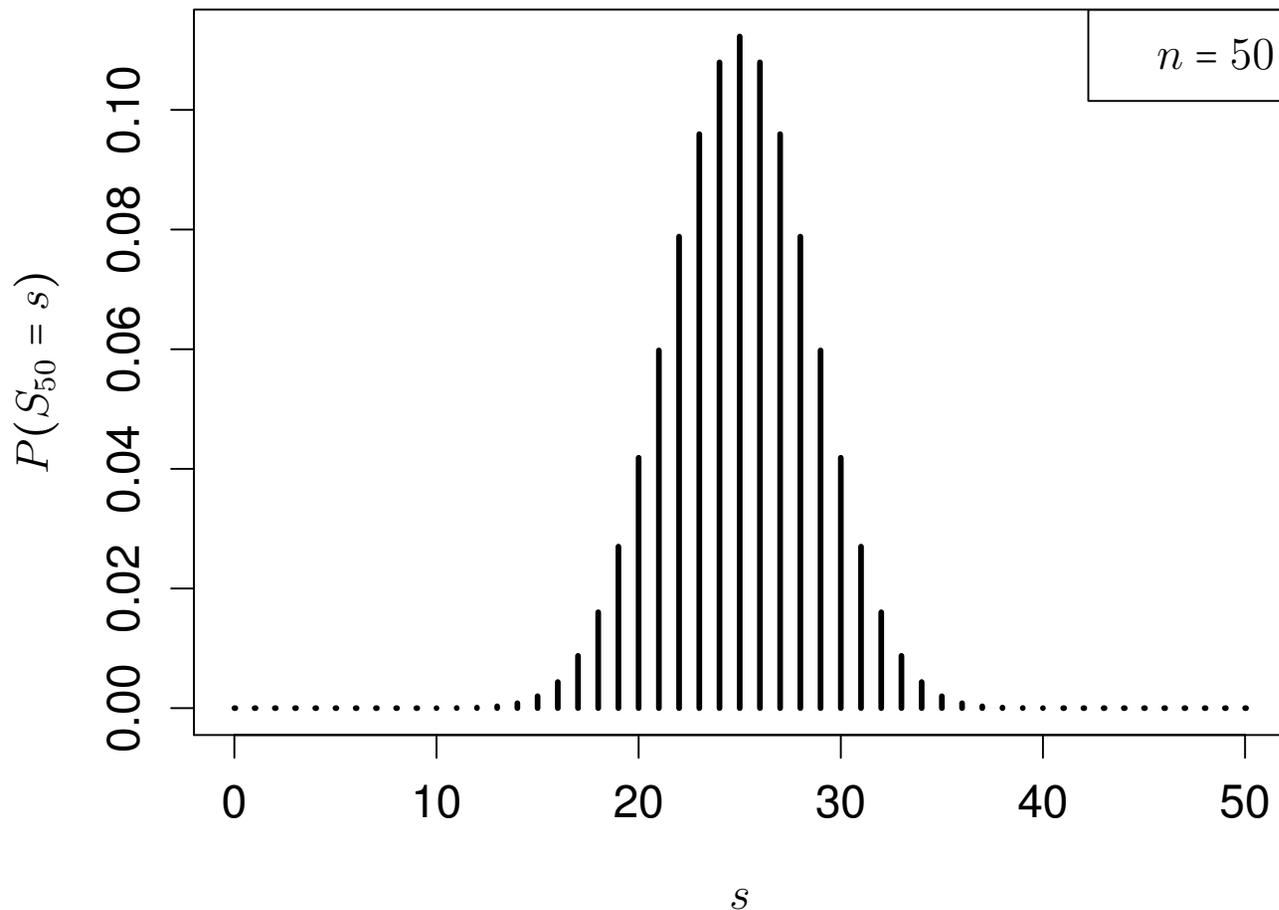


Die Höhe der Balken ist so skaliert, dass die Fläche des Balkens bei  $x = pw$  gerade  $P(pW = x) = p \cdot P(pW = x)/p$  entspricht. (Die blaue Kurve ist  $e^{-x}$ .)

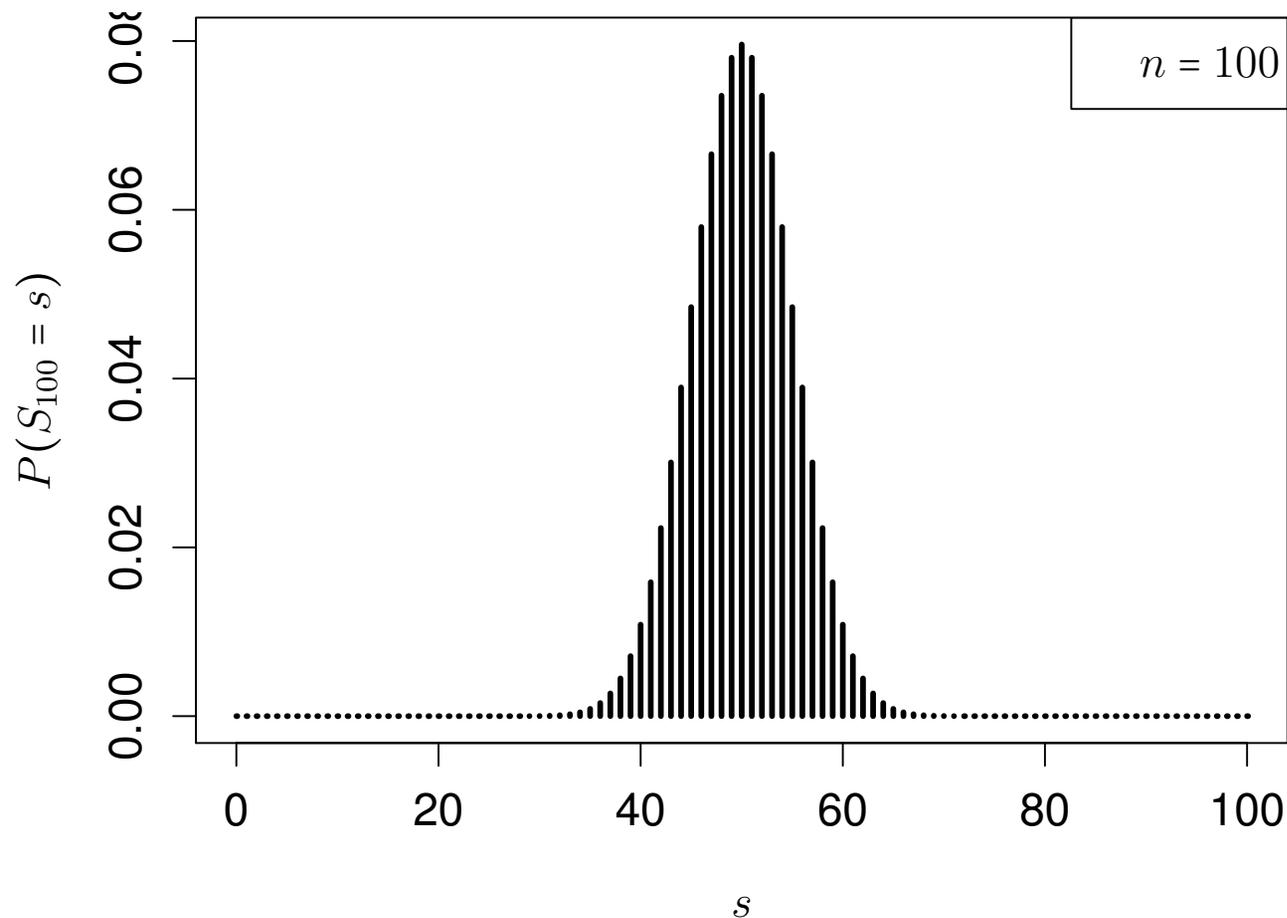
**Beispiel 1.23** (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn). Betrachten wir für  $S_n \sim \text{Bin}_{n,1/2}$  (und verschiedene  $n$ ) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert“.



**Beispiel 1.23** (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn). Betrachten wir für  $S_n \sim \text{Bin}_{n,1/2}$  (und verschiedene  $n$ ) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert“.



**Beispiel 1.23** (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn). Betrachten wir für  $S_n \sim \text{Bin}_{n,1/2}$  (und verschiedene  $n$ ) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert“.



**Beispiel 1.23** (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn). Betrachten wir für  $S_n \sim \text{Bin}_{n,1/2}$  (und verschiedene  $n$ ) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert“.

Zentrieren und stauchen wir:

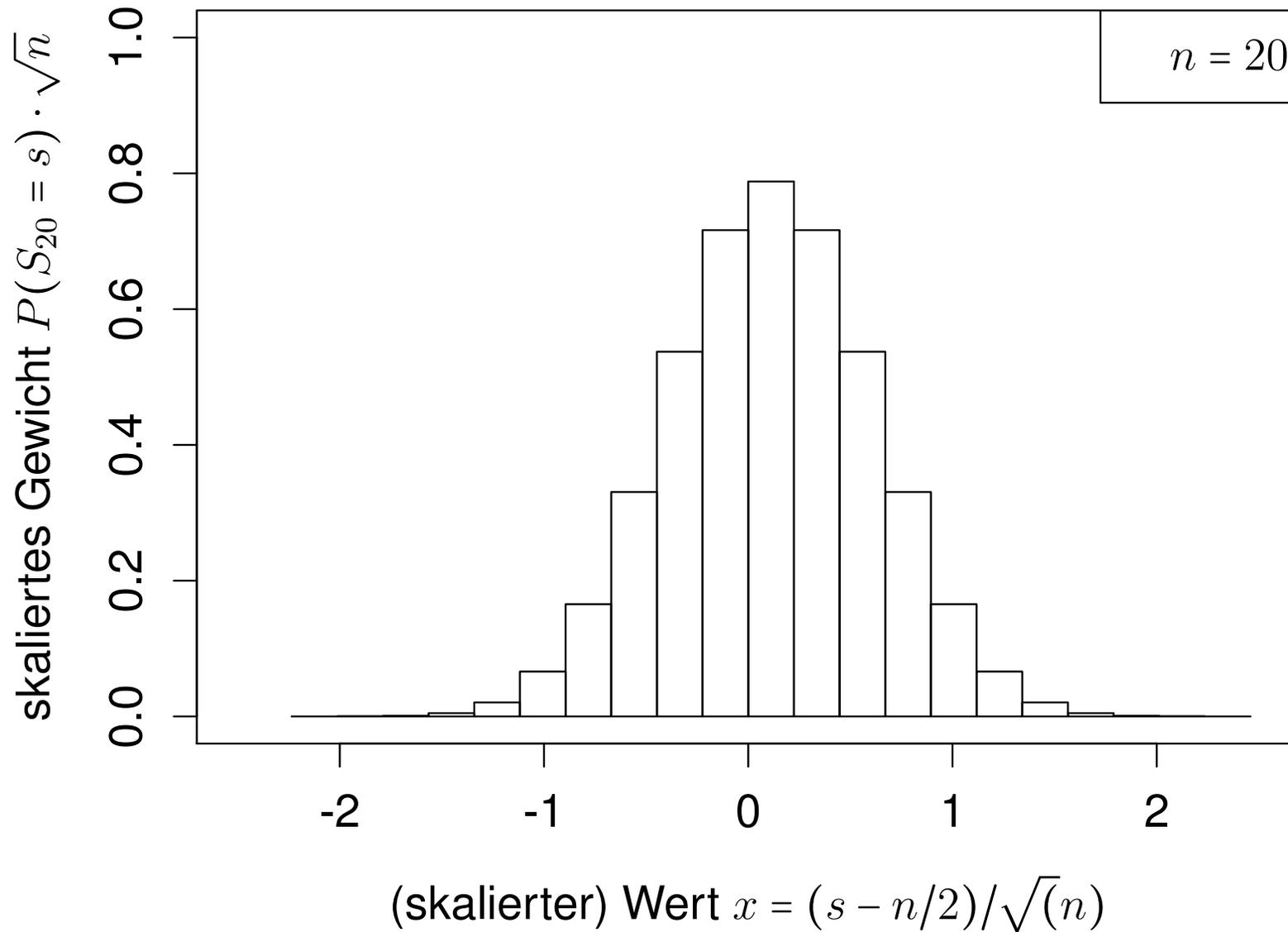
$$X_n := \frac{S_n - n/2}{\sqrt{n}}$$

und skalieren die „Balken“ so, dass die Fläche des Balkens bei  $(s - n/2)/\sqrt{n}$  gerade

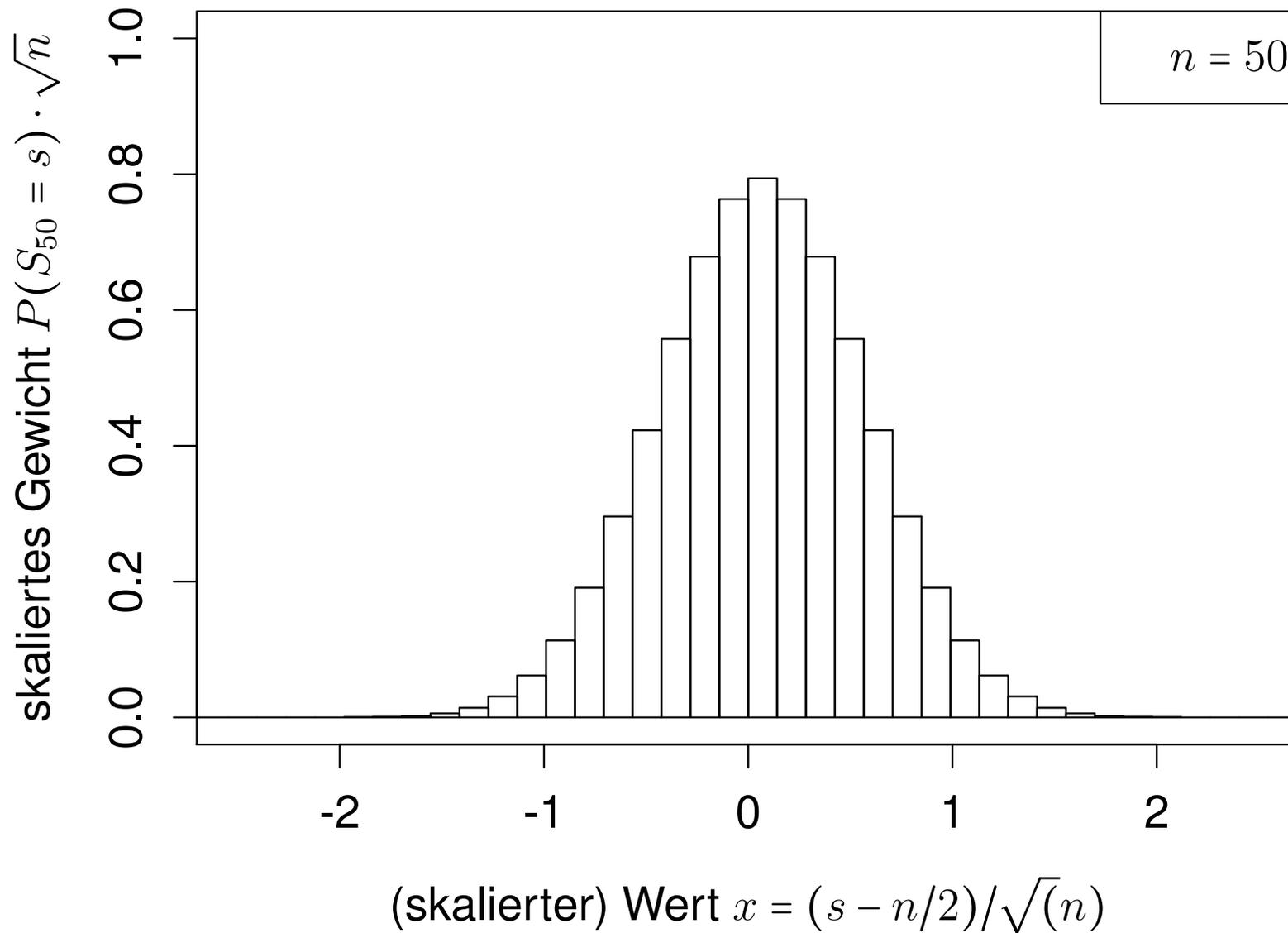
$$P(S_n = s) = \frac{1}{\sqrt{n}} \times \sqrt{n} P(S_n = s)$$

entspricht.

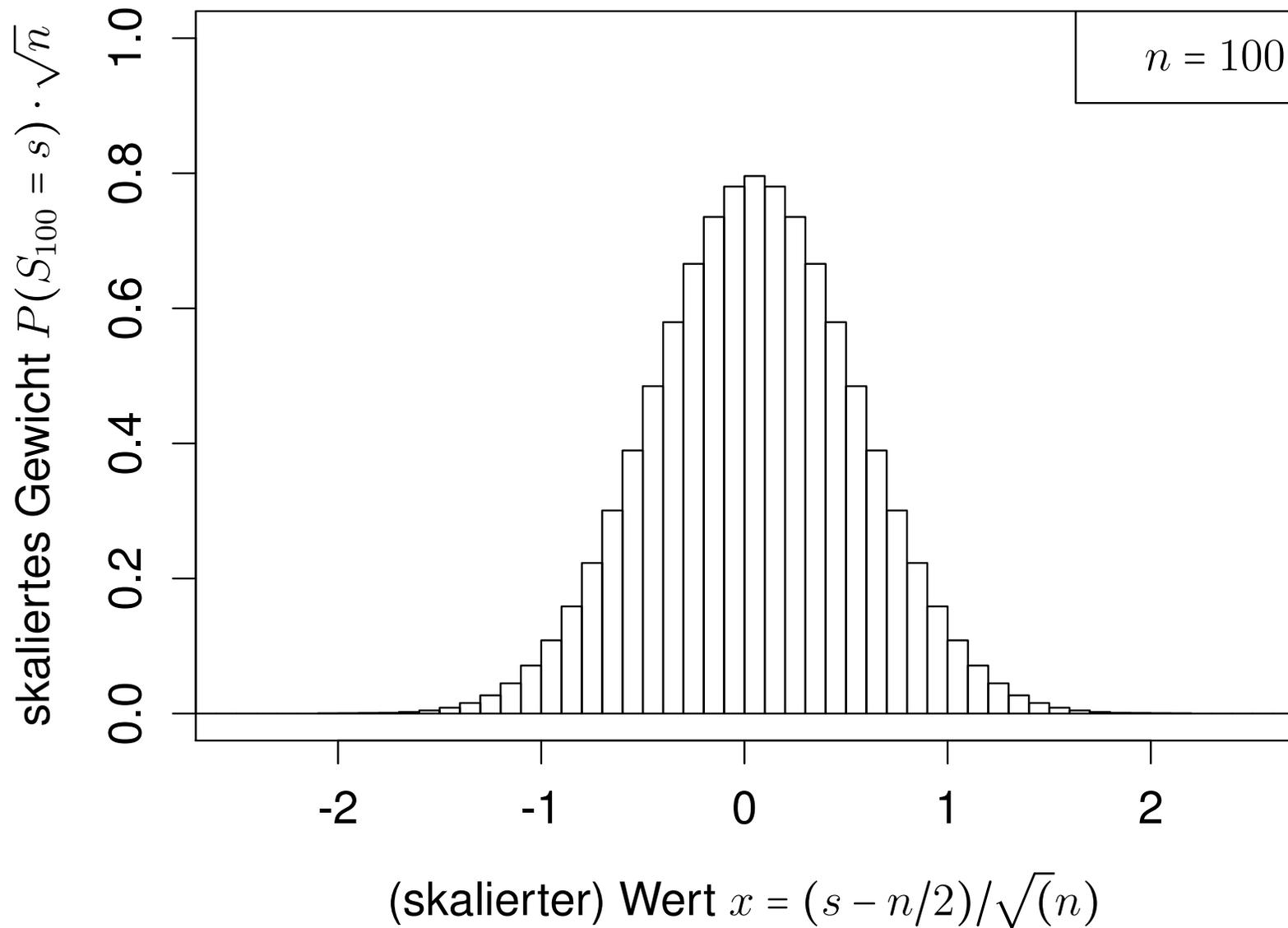
## Bsp. 1.23: Illustration, 2



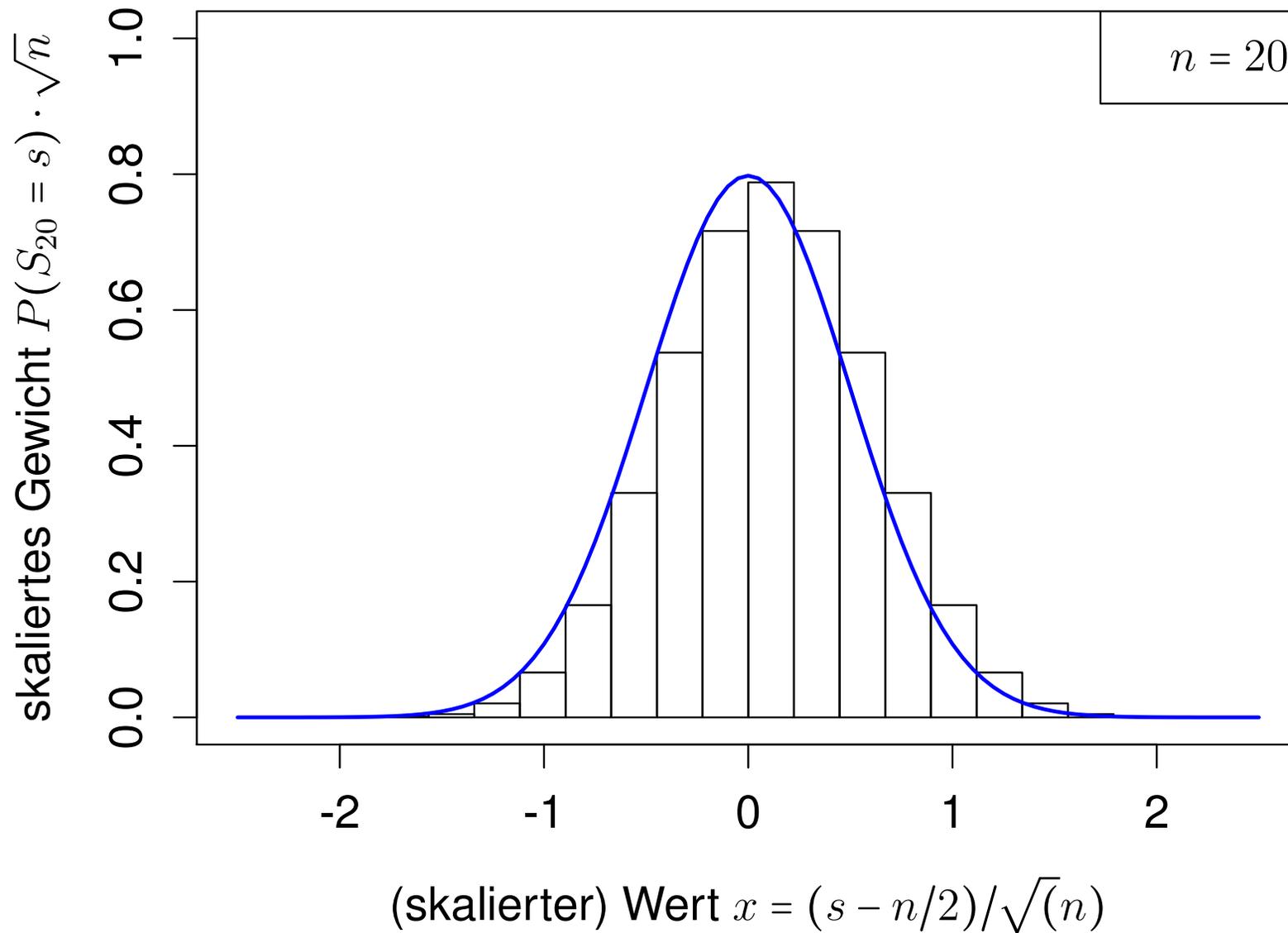
## Bsp. 1.23: Illustration, 2



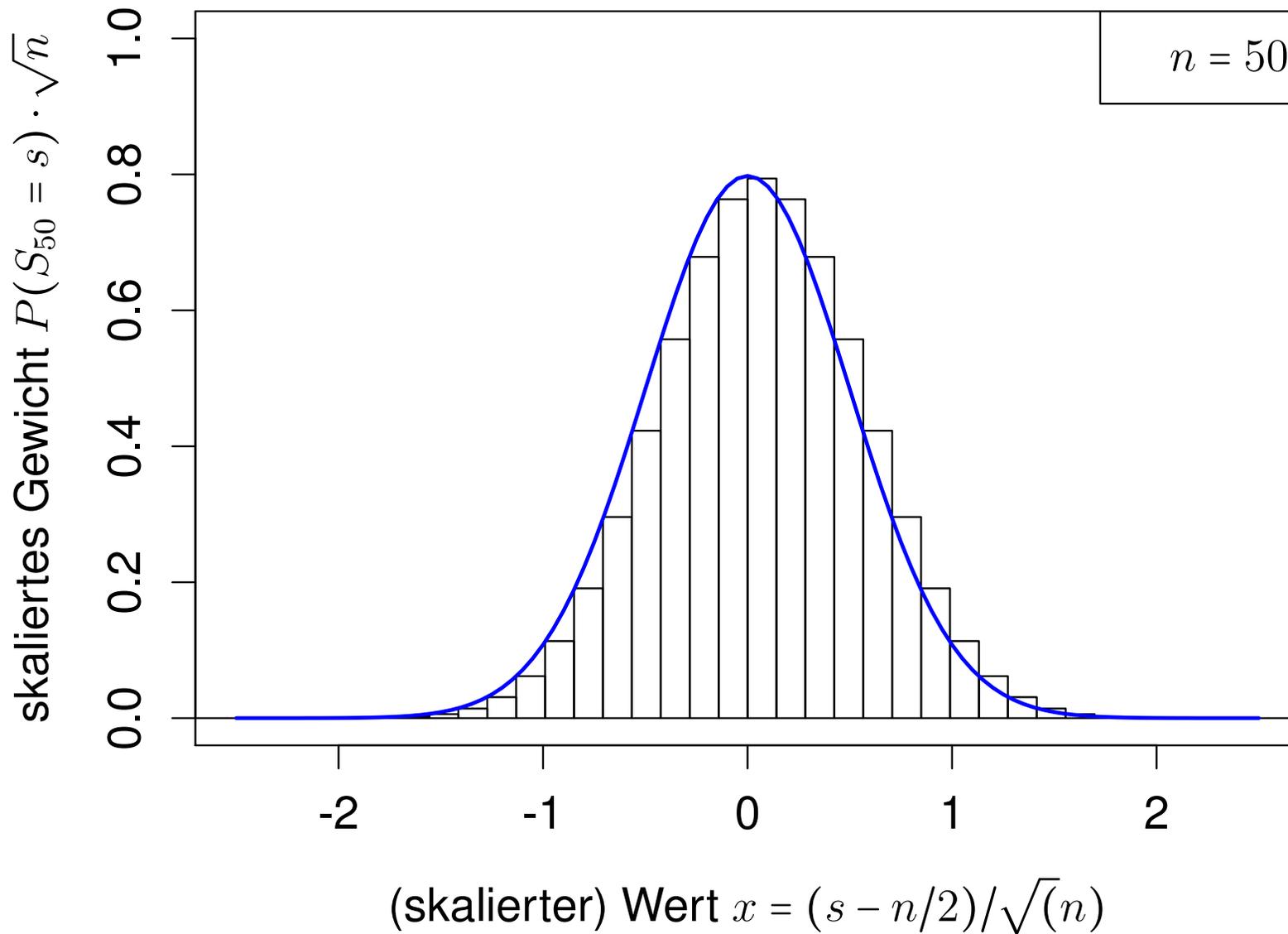
## Bsp. 1.23: Illustration, 2



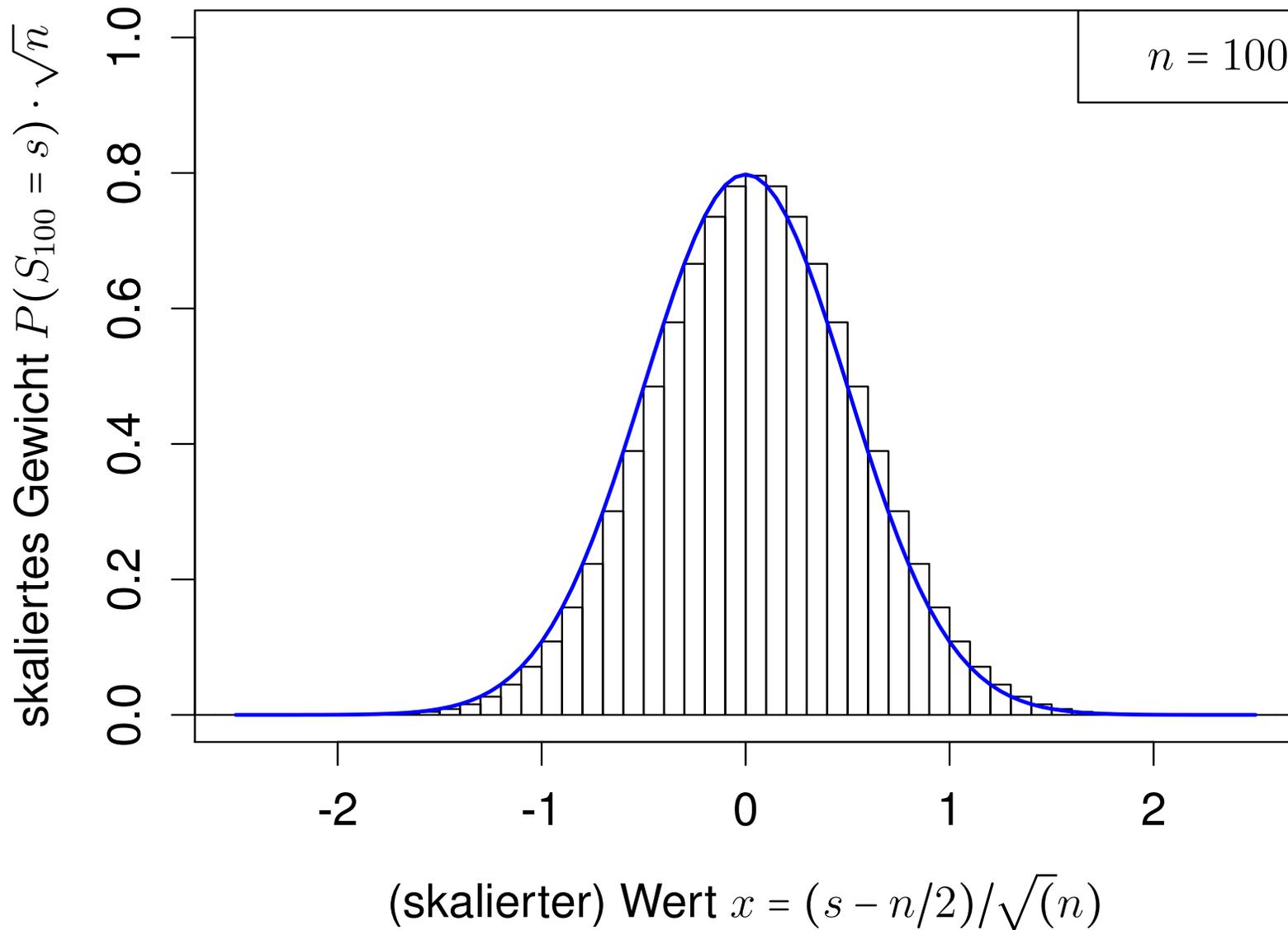
## Bsp. 1.23: Illustration, 2



## Bsp. 1.23: Illustration, 2



## Bsp. 1.23: Illustration, 2



**Definition 1.24.** Sei  $X$  eine Zufallsvariable mit Werten in einem Intervall  $S = [a, b] \subset \mathbb{R}$  mit  $-\infty \leq a < b \leq \infty$  (im Fall  $a > -\infty, b = \infty$  meinen wir  $S = [a, \infty)$ , etc.) und sei  $f : S \rightarrow [0, \infty)$  eine integrierbare Funktion mit

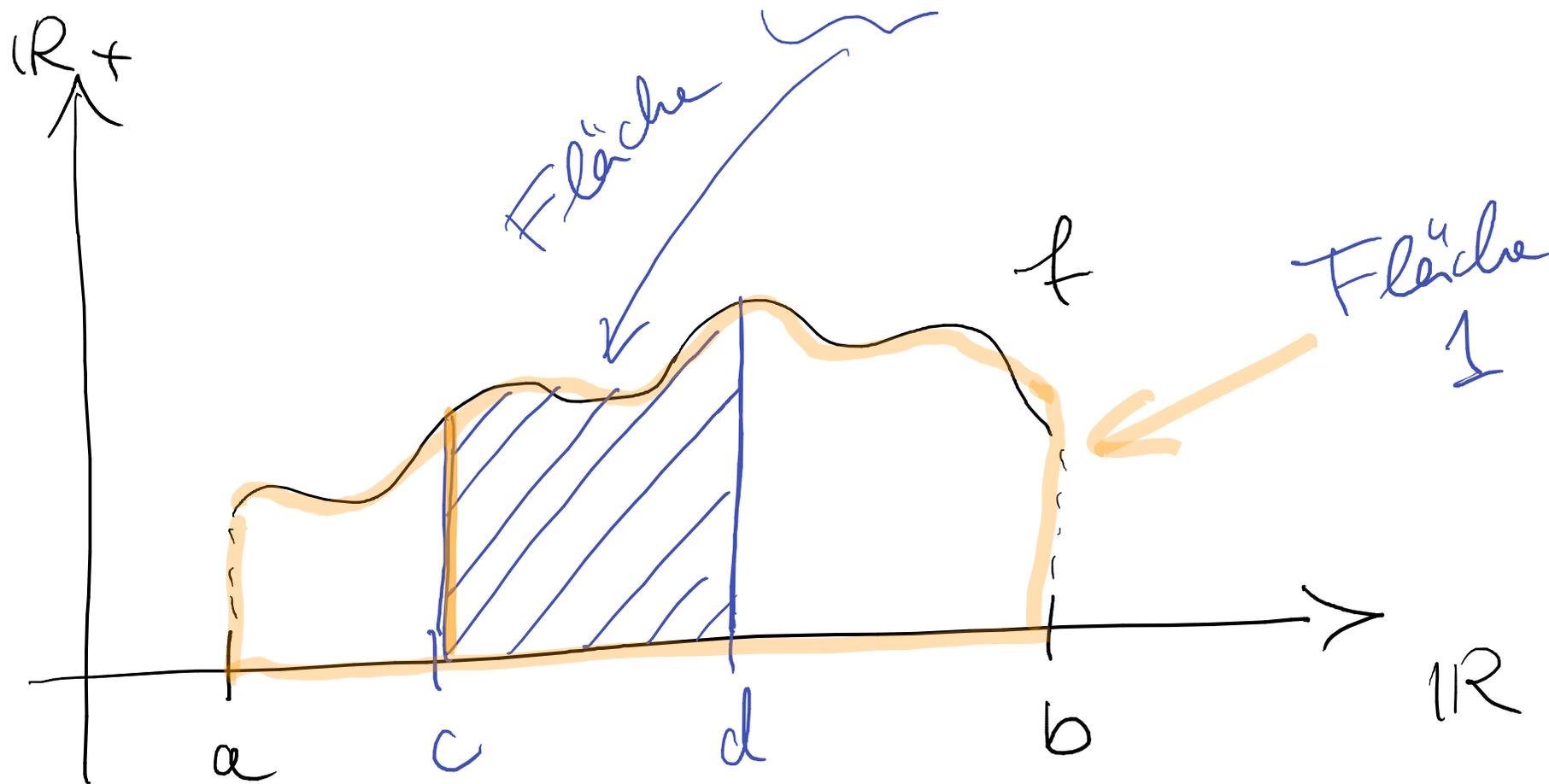
$$\int_a^b f(x) dx = 1.$$

$X$  besitzt die *Dichte* (auch: *Wahrscheinlichkeitsdichte*)  $f$ , wenn gilt

$$P(X \in [c, d]) = \int_c^d f(x) dx \quad \text{für jedes Teilintervall } [c, d] \subset S.$$

Wir notieren oft auch  $f_X$  für die Dichte einer ZV  $X$  (um den Bezug zu  $X$  zu betonen, speziell wenn wir mehrere ZVn zugleich ins Auge fassen).

$X$  hat Dichte  $f$ , so ist  $P(X \in [c, d]) = \int_c^d f(x) dx$



## Interpretation der Dichte:

$X$  ZV mit Dichte  $f_X$ , für  $x \in \mathbb{R}$  und kleines  $\delta > 0$  ist

$$P(X \in [x, x + \delta]) = \int_x^{x+\delta} f_X(a) da \approx \delta f_X(x)$$

(wörtlich zumindest für Stetigkeitsstellen  $x$  von  $f_X$ ), **also**

$$f_X(x) = \lim_{\delta \downarrow 0} \frac{1}{\delta} P(X \in [x, x + \delta])$$

Man formuliert dies gelegentlich auch mit „infinitesimalen Größen“ als

$$P(X \in dx) = f_X(x) dx$$

(Dieser suggestive Ausdruck erhält einen Sinn im Sinne der „Standard-Analysis“, wenn man auf beiden Seiten  $x$  über ein Intervall  $[c, d]$  integriert, dann erhält man Def. 1.24).

## Bemerkung.

Für eine ZV  $X$  mit Dichte  $f_X$  ist es – im Gegensatz zum Fall mit Gewichten – nicht besonders sinnvoll, nach der Wahrscheinlichkeit von Ereignissen  $\{X = x\}$  für feste Punkte  $x \in \mathbb{R}$  zu fragen.

$$\left( \text{denn } P(X=x) = \int_x^x f_X(a) da = 0 \right)$$