Stochastik

Barth, Friedrich

München, [20]03

15. 6. Der zentrale Grenzwertsaatz und die Normalverteilung

urn:nbn:de:hbz:466:1-83580

15.6. Der zentrale Grenzwertsatz und die Normalverteilung

Jede nach $B(n; p)$ verteilte Zufallsgröße läßt sich – wie in **14.5** gezeigt – als Summe $\sum_{i=1}^{n} X_i$ von n unabhängigen Zufallsgrößen X_i schreiben, die alle nach $B(1; p)$ verteilt sind; dabei bedeutet $X_i :=$ »Anzahl der Treffer beim i-ten Versuch«. Nach Satz 205.1 bzw. 209.1 errechnet sich dann der Erwartungswert μ bzw. die Varianz σ^2 dieser Zufallsgröße zu $\mathscr{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathscr{E} X_i$ bzw. zu $\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var} X_i$. Mit diesen Ausdrücken für μ und σ^2 gewinnt der Integralgrenzwertsatz (Satz 294.1) folgende Gestalt:

$$\lim_{n\to\infty} P\left(-\infty < \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathscr{E} X_i}{\sqrt{\sum_{i=1}^{n} \operatorname{Var} X_i}} \leqq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{1}{2}t^2}\, \mathrm{d}t = \Phi(x)$$

Diese Beziehung besagt: Die standardisierte kumulative Verteilungsfunktion einer Summe von n unabhängigen nach $B(1; p)$ verteilten Zufallsgrößen ist für großes n annähernd gleich der *Gauß*schen Integralfunktion.

Die Voraussetzung, daß die einzelnen Summanden binomial verteilt sein müssen, ist eine sehr starke Forderung. Es lag nahe zu untersuchen, ob ein Integralgrenzwertsatz in der obigen Gestalt auch unter schwächeren Voraussetzungen über die Summanden gilt. Es zeigte sich, daß die Forderung, grob gesprochen, die »Streuung« jedes einzelnen Summanden müsse beschränkt sein, ausreicht. Dies ist der wesentliche Inhalt des 1920 von *Georg Pólya* (1887–1985) erstmals im Druck so genannten zentralen Grenzwertsatzes. Darin betrachtet man nicht mehr eine Summe aus endlich vielen Zufallsgrößen, sondern die Teilsummenfolgen $S_n := \sum_{i=1}^{n} X_i$ einer unendlichen Folge von unabhängigen Zufallsgrößen X_i $(i = 1, 2, \ldots)$.

Satz 301.1: Zentraler Grenzwertsatz.

X_i sei eine Folge von Zufallsgrößen $(i = 1, 2, \ldots)$. Endlich viele der X_i seien stets unabhängig. S_n^* sei die standardisierte Zufallsgröße zu $S_n := X_1 + \ldots + X_n$, also

$$S_n^* := \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mathscr{E} X_i}{\sqrt{\sum_{i=1}^{n} \operatorname{Var} X_i}}.$$

Falls es reelle Zahlen A, B, C gibt, so daß für alle i

$0 < A < \operatorname{Var} X_i < B$ und $\mathscr{E}(|X_i - \mathscr{E} X_i|^3) < C$ erfüllt ist, dann gilt

$$\lim_{n\to\infty} P(S_n^* \leqq x) = \Phi(x).$$

Der Integralgrenzwertsatz von *de Moivre* und *Laplace* erweist sich als Spezialfall des zentralen Grenzwertsatzes. Daß ein solch allgemeiner Satz gilt, wurde schon früh vermutet. 1810 bewies *Laplace* (1749–1827) einen zentralen Grenzwertsatz für gleichverteilte Zufallsgrößen*. 1887 stellte *Tschebyschow* (1821–1894) den allgemeinen zentralen Grenzwertsatz auf und beweist ihn, leider lückenhaft, für eine bestimmte Klasse von Zufallsgrößen**. Sein Schüler *Andrei Andrejewitsch Markow* (1856–1922)*** kann 1898 die Lücken schließen. 1901 gelingt *Tschebyschows* Schüler *Alexandr Michailowitsch Ljapunow* (1857–1918) sogar unter noch schwächeren Voraussetzungen der vollständige Beweis****. Moderne Arbeiten konnten dann die oben angegebenen Voraussetzungen über die Zufallsgrößen X_i noch weiter abschwächen. – Der Beweis dieses sehr tief liegenden Satzes übersteigt bei weitem unsere Möglichkeiten.
Der zentrale Grenzwertsatz macht verständlich, daß die standardisierte kumulative Verteilungsfunktion einer binomial verteilten Zufallsgröße für großes n durch die *Gauß*sche Integralfunktion Φ approximiert werden kann. Darüber hinaus offenbart er, warum bei so vielen empirisch gewonnenen Zufallsgrößen die standardisierte kumulative Verteilungsfunktion näherungsweise durch Φ ausgedrückt werden kann. Man kann nämlich annehmen, daß solche Zufallsgrößen sich als Summe einer großen Zahl voneinander unabhängiger Zufallsgrößen ergeben, deren Verteilungen alle ungefähr gleich streuen, wobei die einzelnen Summanden nur einen verschwindend kleinen Einfluß auf die Summe ausüben dürfen. Diese letztere Bedingung ist im wesentlichen die Einschränkung, die *Ljapunow* den X_i auferlegen mußte!
Wir verdeutlichen nun den zentralen Grenzwertsatz an einem

Beispiel: Die Voraussetzungen des zentralen Grenzwertsatzes sind z.B. erfüllt, wenn alle X_i gleichverteilt sind und endlichen Erwartungswert sowie endliche, von 0 verschiedene Varianz besitzen. In diesem Fall hängen die Größen $\operatorname{Var} X_i$ und $\mathscr{E}(|X_i - \mathscr{E} X_i|^3)$ nicht von i ab und sind daher trivialerweise beschränkt. Eine möglichst einfache Zufallsgröße, die sich als Summe solcher X_i darstellen läßt, gewinnen wir folgendermaßen.
Bei einem Laplace-Würfel werden die Augenzahlen wie folgt gewertet:
⚀ = ⚃ = 1, ⚁ = ⚄ = 2, ⚂ = ⚅ = 3. X_i bedeute den Augenwert beim i-ten Wurf. Die X_i sind gleichmäßig verteilt; für jeden Wert ist $p = \frac{1}{3}$. Ferner gilt $\mathscr{E} X_i = 2$ und $\operatorname{Var} X_i = \frac{2}{3}$. Die Zufallsgröße $S_n := \sum\limits_{i=1}^{n} X_i$ bedeutet die Summe der Augenwerte der ersten n Würfe. Es interessiert nun, wie gut sich die kumulative Verteilungsfunktion der zugehörigen standardisierten Zufallsgröße S_n^* mit wachsendem n durch Φ approximieren läßt.
Statt dessen kann man auch die Approximation der Dichtefunktion von S_n durch $\varphi_{\mu\sigma}$ mit wachsendem n untersuchen. Dazu müssen wir die Wahrscheinlichkeitsverteilungen der S_n aufstellen.

* *Mémoire sur les approximations des formules qui sont fonctions de très grands nombres, et sur leur application aux probabilités.*
** *Sur deux théorèmes relatifs aux probabilités* (Originalarbeit auf russisch).
*** Марков (Betonung auf dem a) – Siehe Seite 395.
**** Ляпунов (Betonung auf o) – *Nouvelle forme du théorème sur la limite de probabilité.* – Siehe Seite 395.

Für S_1 gilt $P(S_1 = k) = \frac{1}{3}$ für $k = 1, 2, 3$.

Für S_2 gilt $P(S_2 = k) = \sum_{j=1}^{3} P(X_1 = j \wedge X_2 = k - j)$ für $k = 2, 3, \dots, 6$.

Wegen der stochastischen Unabhängigkeit von X_1 und X_2 erhält man daraus

$$P(S_2 = k) = \sum_{j=1}^{3} P(X_1 = j) \cdot P(X_2 = k - j) \qquad \text{für } k = 2, \dots, 6.$$

Man nennt die rechts stehende Summe eine **Faltung** der Wahrscheinlichkeitsverteilungen von X_1 und X_2.

Wegen der stochastischen Unabhängigkeit von X_1, X_2 und X_3 erhält man die Wahrscheinlichkeitsverteilung von S_3 durch Faltung der Wahrscheinlichkeitsverteilungen von S_2 und X_3:

$$\begin{aligned} P(S_3 = k) &= \sum_{t=1}^{3} \sum_{s=1}^{3} P(X_1 = s \wedge X_2 = t \wedge X_3 = k - (s + t)) = \\ &= \sum_{t=1}^{3} \sum_{s=1}^{3} P(X_1 = s) \cdot P(X_2 = t) \cdot P(X_3 = k - (s + t)) = \\ &= \sum_{j=2}^{6} \left(\sum_{s=1}^{3} P(X_1 = s) \cdot P(X_2 = j - s) \right) \cdot P(X_3 = k - j) = \\ &= \sum_{j=2}^{6} P(S_2 = j) \cdot P(X_3 = k - j) \qquad \text{für } k = 3, \dots, 9, \end{aligned}$$

was auch anschaulich klar ist.

Für S_4 ergeben sich 2 Möglichkeiten der Faltung, nämlich entweder

$$P(S_4 = k) = \sum_{j=3}^{9} P(S_3 = j) \cdot P(X_4 = k - j) \qquad \text{für } k = 4, \dots, 12 \qquad \text{oder}$$

$$P(S_4 = k) = \sum_{j=2}^{6} P(S_2 = j) \cdot P(S_2 = k - j) \qquad \text{für } k = 4, \dots, 12.$$

Schließlich erhält man noch für die Wahrscheinlichkeitsverteilung von S_8

$$P(S_8 = k) = \sum_{j=4}^{12} P(S_4 = j) \cdot P(S_4 = k - j) \qquad \text{für } k = 8, \dots, 24.$$

Die numerische Auswertung der obigen Ausdrücke zeigt Tabelle 303.1.

Zur Bestimmung der approximierenden $\varphi_{\mu\sigma}$-Funktionen benötigen wir die jeweiligen Erwartungswerte $\mathscr{E}S_n$ und Standardabweichungen $\sigma_n := \sqrt{\operatorname{Var} S_n}$; ihre Werte sind in Tabelle 304.1 wiedergegeben. In Tabelle 305.1 sind schließlich die Werte $P(S_n = k)$ und $\varphi_{\mu\sigma}(k)$ einander gegenübergestellt.

k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$3 \cdot P(S_1 = k)$	1	1	1													
$9 \cdot P(S_2 = k)$		1	2	3	2	1										
$27 \cdot P(S_3 = k)$			1	3	6	7	6	3	1							
$81 \cdot P(S_4 = k)$				1	4	10	16	19	16	10	4	1				
$6561 \cdot P(S_8 = k)$								1	8	36	112	266	504	784	1016	1107

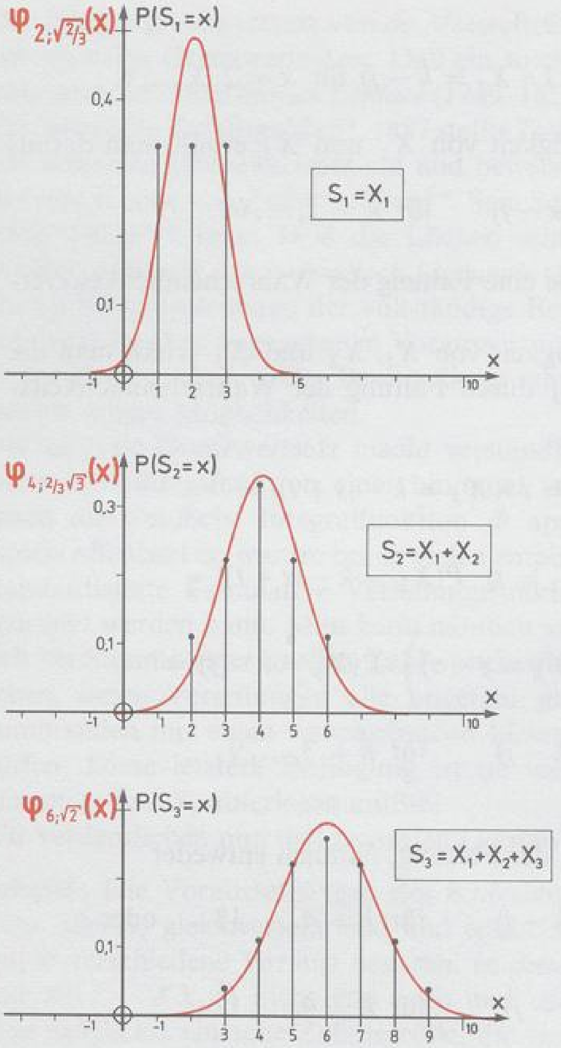
Tab. 303.1 Wahrscheinlichkeitsverteilungen der S_n. – Für S_8 ist symmetrisch zu ergänzen!

n	$\mathscr{E}S_n$	$\operatorname{Var} S_n$	$\sigma_n := \sqrt{\operatorname{Var} S_n}$
1	2	$\frac{2}{3}$	$\frac{1}{3}\sqrt{6} \approx 0{,}8165$
2	4	$\frac{4}{3}$	$\frac{2}{3}\sqrt{3} \approx 1{,}1546$
3	6	2	$\sqrt{2} \approx 1{,}4142$
4	8	$\frac{8}{3}$	$\frac{2}{3}\sqrt{6} \approx 1{,}6330$
8	16	$\frac{16}{3}$	$\frac{4}{3}\sqrt{3} \approx 2{,}3094$

Tab. 304.1 Erwartungswerte und Standardabweichungen der S_n

Figur 304.1 zeigt, wie die Stabdiagramme der Zufallsgrößen S_n mit wachsendem n immer besser durch die zugehörigen Dichtefunktionen $\varphi_{\mu\sigma}$ angenähert werden.

Die Konvergenz muß natürlich nicht in jedem Fall so gut sein wie in unserem Beispiel. In der Praxis wird man sehr häufig erst bei einer großen Anzahl n mit einer befriedigenden Approximation rechnen können. Dies trifft vor allem zu beim Messen einer physikalischen Größe, wenn man annimmt, daß nur viele zufällige und kleine Fehler sich addieren. Ebenso ist es bei der industriellen Herstellung von Massenartikeln. Man darf dabei von der Annahme ausgehen, daß die Abweichungen vom Sollwert nur bedingt sind durch viele Einzeleinwirkungen, die jede



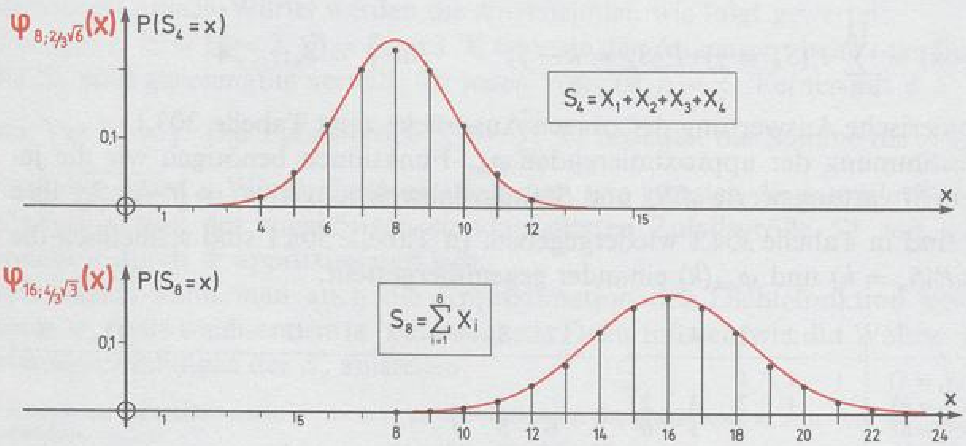


Fig. 304.1 Veranschaulichung des zentralen Grenzwertsatzes für gleichverteilte Zufallsgrößen X_i

k	1	2
$P(S_1=k)$	0,333	0,333
$\varphi_{2;\frac{1}{3}\sqrt{6}}$	0,231	0,489

k	2	3	4
$P(S_2=k)$	0,111	0,222	0,333
$\varphi_{4;\frac{2}{3}\sqrt{3}}$	0,077	0,237	0,345

k	3	4	5	6
$P(S_3=k)$	0,037	0,111	0,222	0,259
$\varphi_{6;\sqrt{2}}$	0,030	0,104	0,222	0,282

k	4	5	6	7	8
$P(S_4=k)$	0,012	0,049	0,123	0,198	0,235
$\varphi_{8;\frac{2}{3}\sqrt{6}}$	0,012	0,045	0,115	0,203	0,244

k	8	9	10	11	12	13	14	15	16
$P(S_8=k)$	0,00015	0,0012	0,0059	0,017	0,041	0,077	0,119	0,155	0,169
$\varphi_{16;\frac{4}{3}\sqrt{3}}$	0,00043	0,0018	0,0059	0,017	0,038	0,075	0,119	0,157	0,173

Tab. 305.1 Wahrscheinlichkeitsverteilungen der Zufallsgrößen S_n und approximierende Funktionswerte $\varphi_{\mu\sigma}(k)$. Alle Tabellen sind symmetrisch fortzusetzen.

für sich nur eine geringe »Streuung« besitzen und deren Einfluß auf den Sollwert jeweils verschwindend gering ist.

Aus diesem Grund haben viele in der Natur und Technik vorkommende Zufallsgrößen eine Verteilung, deren Dichtefunktion bzw. kumulative Verteilungsfunktion recht gut durch $\varphi_{\mu\sigma}$ bzw. $\Phi_{\mu\sigma}$ approximiert werden. Die entsprechenden standardisierten Funktionen werden dann durch φ bzw. Φ approximiert. Es liegt nahe zu vermuten, daß φ auch Dichtefunktion einer »Grenzzufallsgröße« sein wird. Dann wäre Φ die zugehörige kumulative Verteilungsfunktion. Da Φ alle Werte zwischen 0 und 1 kontinuierlich annimmt, müßte die »Grenzzufallsgröße« auch »kontinuierlich« sein, d.h., sie müßte als Wertemenge $\mathbb{R}$ haben. Ihr müßte daher ein überabzählbares Ω als Definitionsmenge zugrunde gelegt werden. Solche Zufallsgrößen kennen wir bis jetzt nicht. In einer erweiterten Theorie betrachtet man jedoch auch solche Zufallsgrößen. Man nennt sie **stetige Zufallsgrößen.** Beispiele hierfür sind Lebensdauern, Entfernung des Einschusses vom Mittelpunkt einer Zielscheibe usw. *Adolphe Quetelet* (1796–1874) und *Francis Galton* (1822–1911) haben gefunden, daß viele in der Natur vorkommende Größen Verteilungen besitzt, die sehr gut durch $\Phi_{\mu\sigma}$ dargestellt werden können. *Henri Poincaré* (1854–1912) nannte solche Verteilungen *normal**. Man definiert:

Definition 305.1: Eine stetige Zufallsgröße mit $\Phi_{\mu\sigma}$ als kumulativer Verteilungsfunktion heißt **normal verteilt.**

Bemerkungen:

1. In der Theorie der stetigen Zufallsgrößen muß man natürlich Erwartungswert und Standardabweichung neu definieren. Es treten dabei statt der Summen Integrale auf. Die Parameter μ und σ von $\Phi_{\mu\sigma}$ erweisen sich dann als Erwartungswert und Standardabweichung der Zufallsgröße, die $\Phi_{\mu\sigma}$ als kumulative Verteilungsfunktion hat. Nachweis in Aufgabe 316/**41**.

* *Calcul des Probabilités* – Vorlesungen während des 2. Semesters 1893–1894, herausgegeben 1896, Seite 76. – Siehe Seite 395.

2. Ist $\mu = 0$ und $\sigma = 1$, so heißt die zugehörige Zufallsgröße **standardisiert**; vielfach nennt man sie auch **normiert.** $\Phi_{0;1} = \Phi$ heißt daher **Standardnormalverteilung.**

Bei der Anwendung der Normalverteilung auf real vorkommende Zufallsgrößen, deren kumulative Verteilungsfunktion F näherungsweise gleich $\Phi_{\mu\sigma}$ ist, muß man zwei Fälle unterscheiden. Ist die Zufallsgröße X *stetig*, d.h., kann sie jeden Wert $x \in \mathbb{R}$ annehmen, dann wird man $F(x)$ durch $\Phi\left(\dfrac{x-\mu}{\sigma}\right)$ approximieren.

Nimmt dagegen eine Zufallsgröße nur *diskrete* Werte $k \in \mathbb{Z}$ an, dann wird man $F(k)$ durch $\Phi\left(\dfrac{k-\mu+0{,}5}{\sigma}\right)$ approximieren, wie bei der Approximation der Binomialverteilung. Durch den Summanden $\dfrac{0{,}5}{\sigma} = \dfrac{1}{2\sigma}$ im Argument von Φ wird berücksichtigt, daß man bei der Berechnung von $F(k)$ als letztes Rechteck das ganze Rechteck über k nehmen muß. (Vgl. Figur 306.1.) Man nennt den Summanden $\dfrac{1}{2\sigma}$ **Stetigkeitskorrektur.**

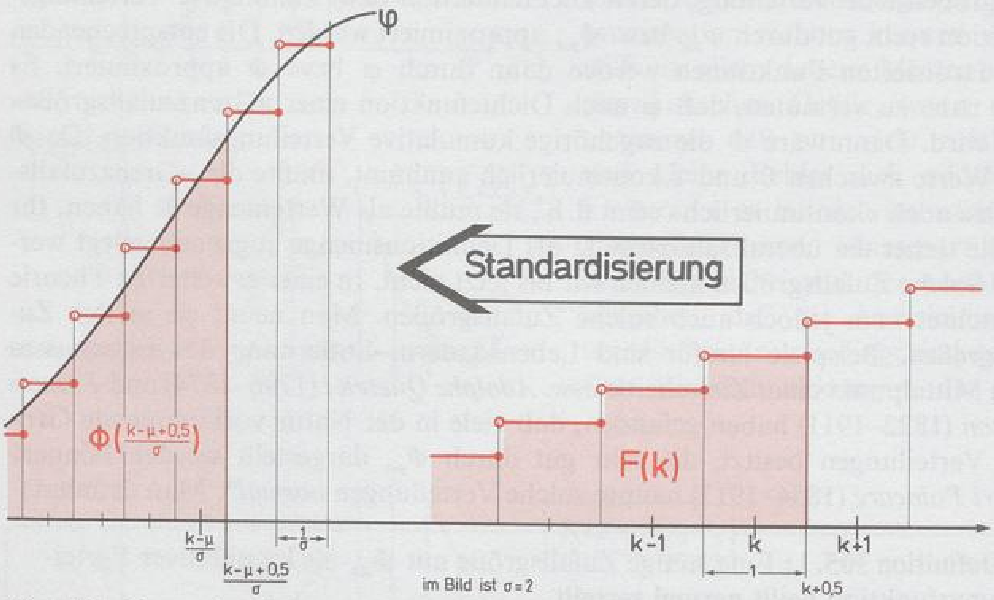


Fig. 306.1 Veranschaulichung der Stetigkeitskorrektur. Rechts ist ein Ausschnitt der Dichtefunktion gezeichnet, links die zugehörige standardisierte Dichtefunktion mit der approximierenden φ-Funktion.

Nimmt die Zufallsgröße zwar diskrete, aber nicht ganzzahlige Werte an, dann kann man durch geeignete Wahl der Einheiten erreichen, daß die Werte ganzzahlig werden, und dann wieder die Stetigkeitskorrektur anwenden.

Beispiel: Mit einer Maschine werden Stifte hergestellt. Die Länge X der Stifte läßt sich als Zufallsgröße auffassen, die annähernd normal verteilt ist. Ihr Mittelwert sei $\mu = 10$ mm, ihre Standardabweichung $\sigma = 0{,}02$ mm. Ein Stift muß

mehr als 9,97 mm lang sein, damit er brauchbar ist. Wie groß ist die Wahrscheinlichkeit für einen zu kurzen Stift?
Nehmen wir zunächst an, X sei stetig verteilt, d.h., daß jede Länge vorkommen kann. Dann gilt:

$$P(X \leqq 9{,}97) = F(9{,}97) \approx \Phi\left(\tfrac{9{,}97-10}{0{,}02}\right) = \Phi(-1{,}5) = 1 - \Phi(1{,}5) = 1 - 0{,}93318 =$$
$$= 0{,}06682 \approx 6{,}7\%.$$

Gehen wir aber davon aus, daß die Meßgenauigkeit 0,01 mm beträgt, dann nimmt die Zufallsgröße nur diskrete Werte an, die in der Einheit 0,01 mm ganzzahlig sind. In diesem Fall ist es sinnvoller, die Stetigkeitskorrektur zu verwenden:

$$P(X \leqq 997) = F(997) \approx \Phi\left(\tfrac{997-1000+0{,}5}{2}\right) = \Phi(-1{,}25) = 1 - \Phi(1{,}25) =$$
$$= 1 - 0{,}89434 = 0{,}10566 \approx 10{,}6\%.$$

Für stetige Zufallsgrößen hat es keinen Sinn, eine Wahrscheinlichkeitsverteilung zu betrachten, da ja die Wahrscheinlichkeit, daß die stetige Zufallsgröße einen bestimmten Wert annimmt, für jeden Wert 0 ist. Es hat nur einen Sinn, danach zu fragen, mit welcher Wahrscheinlichkeit die stetige Zufallsgröße Werte aus einem bestimmten Intervall annimmt. Dabei ist es belanglos, ob man abgeschlossene oder offene Intervalle betrachtet, weil ja $P(X = x) = 0$ gilt. Also ist z.B. $P(X \leqq x) = P(X < x)$.

Für beliebige Zufallsgrößen konnten wir auf Seite 185 die Wahrscheinlichkeit $P(|X-\mu| < t\sigma)$ mit Hilfe der *Bienaymé-Tschebyschow*-Ungleichung abschätzen zu

$$P(|X-\mu| < t\sigma) \geqq 1 - \frac{1}{t^2}.$$

Für normal verteilte Zufallsgrößen können wir nun genauere Werte für diese Wahrscheinlichkeiten berechnen. Man findet sie in Abhängigkeit von t tabellarisiert in den *Stochastik-Tabellen* auf Seite 45 als σ-Bereichstabelle. Insbesondere erhält man

Satz 307.1: Ist X eine normal verteilte Zufallsgröße, so gilt auf Promille gerundet

$$P(|X-\mu| < \sigma) = 68{,}3\%$$
$$P(|X-\mu| < 2\sigma) = 95{,}5\%$$
$$P(|X-\mu| < 3\sigma) = 99{,}7\%.$$

Figur 307.1 veranschaulicht diesen Satz.

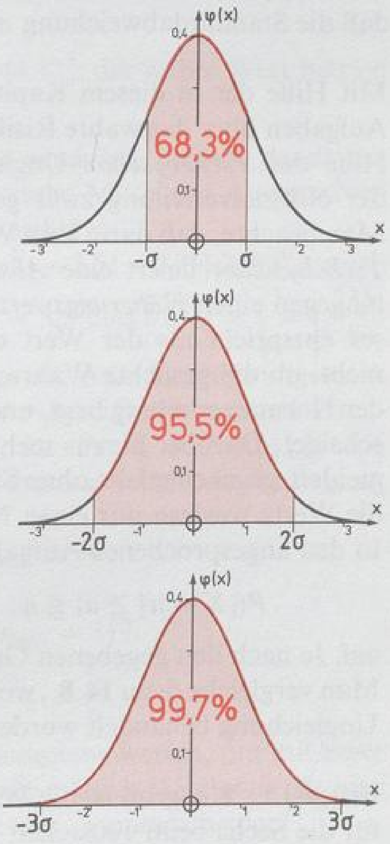


Fig. 307.1 Illustration zu Satz 307.1

Beweis: Unter Verwendung von Satz 300.1 erhält man

$$\begin{aligned} P(|X-\mu|<\sigma) &= P(\mu-\sigma<X<\mu+\sigma) = \\ &= \Phi_{\mu\sigma}(\mu+\sigma)-\Phi_{\mu\sigma}(\mu-\sigma) = \\ &= \Phi\left(\frac{\mu+\sigma-\mu}{\sigma}\right)-\Phi\left(\frac{\mu-\sigma-\mu}{\sigma}\right) = \\ &= \Phi(1)-\Phi(-1) = \\ &= 2\Phi(1)-1 = \\ &= 2\cdot 0{,}84134-1 = \\ &= 0{,}68268\,. \end{aligned}$$

Analog erhält man

$$P(|X-\mu|<2\sigma) = 2\Phi(2)-1 = 0{,}95450 \qquad \text{und}$$
$$P(|X-\mu|<3\sigma) = 2\Phi(3)-1 = 0{,}99730\,.$$

Auf Grund des Integralgrenzwertsatzes 294.1 gilt Satz 307.1 auch für Zufallsgrößen, die binomial verteilt sind, wenn n groß ist.
Übrigens sind die Werte aus Satz 307.1 bis auf Ungenauigkeiten im Promillebereich gerade die von *de Moivre* 1733 angegebenen Abschätzungen. (Siehe Seite 277.) Darüber hinaus bestätigt Satz 307.1 die richtige Erkenntnis *de Moivres*, daß die Standardabweichung σ die »Abschätzung reguliert«.

Mit Hilfe der in diesem Kapitel gewonnenen Erkenntnisse ist es nun möglich, Aufgaben über das wahre Risiko, die wir in **14.8.** unter Umständen nur grob mit Hilfe der *Tschebyschow*-Ungleichung bewältigen konnten, unter Verwendung der Normalverteilung zwar genauer, dafür aber nur näherungsweise zu lösen. Man beachte, daß darin kein Widerspruch liegt! Die Ungleichung von *Bienaymé-Tschebyschow* liefert eine *Abschätzung*, die oft grob ist, die Normalverteilung hingegen einen *Näherungswert*, der der gesuchten Wahrscheinlichkeit meist besser entspricht als der Wert der *Tschebyschow*-Abschätzung; wir wissen aber nicht, ob die gesuchte Wahrscheinlichkeit unter oder über dem Näherungswert der Normalverteilung liegt, und schon gar nicht, um wieviel sie sich davon unterscheidet. Darüber hinaus rechnet man, um lästige Fallunterscheidungen zu vermeiden, gegebenenfalls ohne Stetigkeitskorrektur, da man sich ja bewußt ist, daß die Werte sowieso nur einen Näherungscharakter haben.
In den angesprochenen Aufgaben treten Ungleichungen vom Typ

$$P(|X-\mu|\geqq a)\leqq\eta \qquad \text{bzw.} \qquad P(|H_n-p|\geqq\varepsilon)\leqq\eta$$

auf. Je nach den gegebenen Größen erhält man verschiedene Problemstellungen. Man vergleiche dazu **14.8.**, wo die nachfolgenden Beispiele mit der *Tschebyschow*-Ungleichung behandelt wurden.

Beispiel 1: Wie groß ist die Wahrscheinlichkeit dafür, daß die relative Häufigkeit für die Sechs beim 100fachen Wurf eines L-Würfels um weniger als 0,05 von der Wahrscheinlichkeit für eine Sechs abweicht?

Lösung: Gesucht ist der kleinste Wert für η, so daß $P(|H_{100} - \frac{1}{6}| < 0{,}05) \geqq 1 - \eta$. Dazu nähern wir $P(|X - \frac{100}{6}| < 5)$ durch die *Gauß*sche Integralfunktion Φ an. Berücksichtigen wir, daß X nur ganzzahlige Werte annehmen kann, dann erhalten wir für den letzten Ausdruck die Form $P(12 \leqq X \leqq 21)$. Mit $\sigma = \frac{5}{3}\sqrt{5}$ gewinnen wir die Näherung

$$P(12 \leqq X \leqq 21) \approx \Phi\left(\frac{21 - \frac{100}{6} + \frac{1}{2}}{\frac{5}{3}\sqrt{5}}\right) - \Phi\left(\frac{11 - \frac{100}{6} + \frac{1}{2}}{\frac{5}{3}\sqrt{5}}\right) =$$
$$= \Phi(1{,}2969) - \Phi(-1{,}3864) =$$
$$= 0{,}90266 - 1 + 0{,}91717 =$$
$$= 0{,}81983\,.$$

Rechnet man hingegen bequemlichkeitshalber so, als sei die Trefferanzahl X normal verteilt, dann erhält man unter Ausnutzung der Symmetrie von $\varphi_{\mu\sigma}$ (vgl. Figur 309.1)

$$P(|X - \tfrac{100}{6}| < 5) \approx$$
$$\approx 1 - 2\Phi\left(\frac{(\frac{100}{6} - 5) - \frac{100}{6}}{\frac{5}{3}\sqrt{5}}\right) =$$
$$= 2\Phi(\tfrac{3}{5}\sqrt{5}) - 1 = 0{,}82026\,.$$

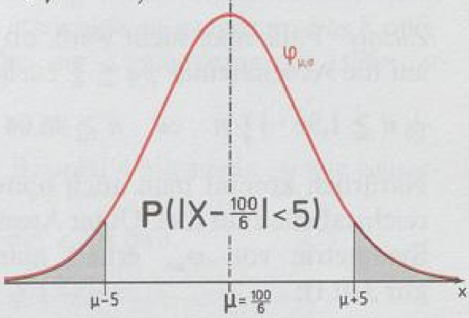


Fig. 309.1

$$P(|X - \tfrac{100}{6}| < 5) \approx 1 - 2\Phi\left(\frac{(\frac{100}{6} - 5) - \frac{100}{6}}{\frac{5}{3}\sqrt{5}}\right)$$

Die *Tschebyschow*-Ungleichung lieferte seinerzeit 44,4%; der wahre Wert beträgt hingegen 0,82210. (Vgl. Seite 253.)

Beispiel 2: Wie oft muß ein L-Würfel mindestens geworfen werden, damit mit einer Sicherheit von mindestens 60% das arithmetische Mittel der Augenzahlen um weniger als 0,25 vom Erwartungswert 3,5 abweicht?

Lösung: Gesucht ist ein kleinstes n, so daß $P(|\bar{X}_n - 3{,}5| < 0{,}25) \geqq 60\%$ wird. Auf Grund von Satz 212.1 und Satz 212.2 gilt $\mathscr{E}\bar{X}_n = 3{,}5$ und $\sigma(\bar{X}_n) = \frac{1}{\sqrt{n}}\sqrt{\frac{35}{12}}$.

Nimmt man wegen des zentralen Grenzwertsatzes an, daß $\bar{X}_n$ näherungsweise normal verteilt ist, dann kann man der Tabelle der σ-Bereiche bei normalverteilten Zufallsgrößen (*Stochastik-Tabellen*, Seite 45) für $P(|\bar{X}_n - 3{,}5| < \frac{1}{4}) \geqq 60\%$ den Wert $t \geqq 0{,}8416$ entnehmen. Das ergibt dann mit $t\sigma = 0{,}25$ die Bedingung

$$0{,}8416 \cdot \frac{1}{\sqrt{n}}\sqrt{\frac{35}{12}} \leqq 0{,}25 \quad \Leftrightarrow \quad n \geqq 33{,}05\ldots \Rightarrow n_{\min} = 34\,.$$

Die *Tschebyschow*-Ungleichung lieferte seinerzeit $n_{\min} = 117$.

Eine weitere Aufgabe desselben Typs enthält

Beispiel 2a: Wie oft muß man einen L-Würfel mindestens werfen, um mit einer Sicherheit von mindestens 95% zu erreichen, daß die relative Häufigkeit für eine Sechs sich von ihrer Wahrscheinlichkeit um höchstens $\frac{1}{10}$ unterscheidet?

Lösung: Gesucht ist zu $\varepsilon = \frac{1}{10}$ ein kleinstes n, so daß $P(|H_n - p| \leqq \varepsilon) \geqq 95\%$, d.h.

$P(|X-np| \leqq n\varepsilon) \geqq 95\%$ wird. Wollte man diese Aufgabe so genau wie möglich lösen, dann müßte man vorher schon wissen, ob $np \pm n\varepsilon$ ganzzahlig werden oder nicht. Will man hier weiterkommen, so bleibt also nur anzunehmen, daß die Zufallsgröße $X :=$ Anzahl der Treffer annähernd normal verteilt ist. Aus der σ-Bereichstabelle erhält man dann mit $P(|X-\mu| \leqq \frac{1}{10}n) \geqq 95\%$ für t die Bedingung $t \geqq 1{,}9600$. Somit

$$\tfrac{1}{10}n \geqq 1{,}96\sqrt{\tfrac{5}{36}n} \quad \Leftrightarrow \quad n \geqq 53{,}3\ldots \Rightarrow n_{\min} = 54\,.$$

Zusatz: Falls man nicht weiß, ob es sich um einen L-Würfel handelt, muß man auf die Abschätzung $pq \leqq \frac{1}{4}$ zurückgreifen. Es ergibt sich dann

$$\tfrac{1}{10}n \geqq 1{,}96 \cdot \tfrac{1}{2}\sqrt{n} \quad \Leftrightarrow \quad n \geqq 96{,}04 \Rightarrow n_{\min} = 97\,.$$

Natürlich kommt man auch ohne die σ-Bereichstabelle zum Ziel. Unter Ausnutzung der Symmetrie von $\varphi_{\mu\sigma}$ erhält man (vgl. Figur 310.1):

$$\Phi\left(\frac{\mu+n\varepsilon-\mu}{\sigma}\right) \geqq 97{,}5\%;$$ für einen L-Würfel also

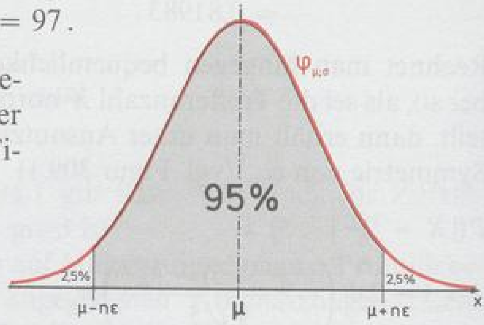


Fig. 310.1 $P(|X-\mu| \leqq n\varepsilon) = 95\%$.

$$\Phi\left(\frac{3\sqrt{n}}{5\sqrt{5}}\right) \geqq 0{,}975 \Leftrightarrow \frac{3\sqrt{n}}{5\sqrt{5}} \geqq 1{,}9600 \Rightarrow n_{\min} = 54\,.$$

Auch die beiden Fragestellungen hinsichtlich der gesuchten Intervalle können nun unter Umständen genauer behandelt werden.

Beispiel 3: In welchem Intervall um $p = \frac{1}{6}$ liegt bei 100maligem Werfen eines L-Würfels die relative Häufigkeit für die Augenzahl 6 mit einer Mindestwahrscheinlichkeit von 60%?

Lösung: Gesucht ist ein kleinstes ε, so daß $P(|H_{100} - \frac{1}{6}| < \varepsilon) \geqq 60\%$ wird. Wir formen um zu $P(|X - \frac{100}{6}| < 100\,\varepsilon) \geqq 0{,}6$ und nehmen zur Vereinfachung an, daß X angenähert normal verteilt ist. Für 60% erhalten wir aus der σ-Bereichstabelle $t \geqq 0{,}8416$. Also

$$100\,\varepsilon = t\sigma \geqq 0{,}8416 \cdot \tfrac{5}{3}\sqrt{5} \Rightarrow \varepsilon \geqq 0{,}03136\ldots$$

Bedenkt man wieder, daß H_{100} nur Hundertstelwerte aus $[0; 1]$ annehmen kann, so erhält man: Mit einer Wahrscheinlichkeit von mindestens 60% liegen die Werte von H_{100} (»Sechs«) im Intervall $[0{,}14; 0{,}19]$.
Die *Tschebyschow*-Ungleichung lieferte seinerzeit mit $\varepsilon \geqq 0{,}0589\ldots$ das Intervall $[0{,}11; 0{,}22]$.

Beispiel 4: Für welches Intervall kann man mit einer Sicherheit von mindestens 90% schließen, daß es die Wahrscheinlichkeit für eine Sechs enthält, wenn sich bei 100 Würfen $h_{100}(\{6\}) = 0{,}18$ ergeben hat?

Lösung: Gesucht ist für das Konfidenzintervall $]H_n - \varepsilon; H_n + \varepsilon[$ ein ε, so daß

$$P(|H_{100}(\{6\}) - p| < \varepsilon) \geqq 90\% \quad \Leftrightarrow \quad P(|X - 100\,p| < 100\,\varepsilon) \geqq 90\%$$

erfüllt ist; dabei gibt X die Anzahl der Treffer an. Der σ-Bereichstabelle entnimmt man $t \geqq 1{,}6449$. Also

$$100\,\varepsilon = t\sigma \geqq 1{,}6449\sqrt{100\,p(1-p)} \Leftrightarrow \varepsilon \geqq 0{,}16449\sqrt{p(1-p)}.$$

Da $p(1-p) \leqq \frac{1}{4}$, ist die eingangs gestellte Bedingung sicherlich erfüllt, wenn man $\varepsilon = 0{,}0823$ wählt. Es ergibt sich also ein etwa halb so großes grobes Konfidenzintervall $I(0{,}18) \subset]0{,}0977; 0{,}2623[$ wie bei der Abschätzung mit Hilfe der *Tschebyschow*-Ungleichung.

Wir behandeln nun die Aufgabenstellung von Beispiel 4 allgemein so wie seinerzeit auf Seite 258.
Gesucht ist also zum Konfidenzniveau $1 - \eta$ ein ε, so daß

$$P(|H_n - p| < \varepsilon) \geqq 1 - \eta \Leftrightarrow P(|X - np| < n\varepsilon) \geqq 1 - \eta$$

erfüllt ist. Dazu entnimmt man der σ-Bereichstabelle den zur Sicherheit $1 - \eta$ gehörenden t-Wert, der natürlich von η abhängt. Bezeichnen wir ihn mit $t(\eta)$, dann gilt

$$n\varepsilon \geqq t(\eta)\sqrt{np(1-p)} \Leftrightarrow \varepsilon \geqq \frac{t(\eta)\sqrt{p(1-p)}}{\sqrt{n}}.$$

Man erhält daraus zu jedem Wert h_n von H_n

– **das grobe Konfidenzintervall** $I(h_n)$, wenn man $p(1-p)$ durch $\frac{1}{4}$ abschätzt, also

$$I(h_n) = \left]h_n - \frac{t(\eta)}{2\sqrt{n}};\ h_n + \frac{t(\eta)}{2\sqrt{n}}\right[;$$

– **das Näherungskonfidenzintervall** $\tilde{I}(h_n)$, wenn man p durch die relative Häufigkeit h_n annähert, also

$$\tilde{I}(h_n) = \left]h_n - \frac{t(\eta)\sqrt{h_n(1-h_n)}}{\sqrt{n}};\ h_n + \frac{t(\eta)\sqrt{h_n(1-h_n)}}{\sqrt{n}}\right[;$$

– **das echte Konfidenzintervall** $\mathfrak{J}(h_n)$, dessen Grenzen die Lösungen der quadratischen Gleichung

$$|h_n - p| = \frac{t(\eta)\sqrt{p(1-p)}}{\sqrt{n}} \quad \text{sind.}$$

Eine leichte Rechnung liefert mit den Werten von Beispiel 4

$$\tilde{I}(0{,}18) \subset]0{,}116; 0{,}244[\text{ bzw. } \mathfrak{J}(0{,}18) \subset]0{,}125; 0{,}252[,$$

also wiederum kleinere Werte als die Abschätzung durch die Ungleichung von *Bienaymé-Tschebyschow*.