

Definition 101

Zwei kontinuierliche Zufallsvariablen X und Y heißen **unabhängig**, wenn

$$\Pr[X \leq x, Y \leq y] = \Pr[X \leq x] \cdot \Pr[Y \leq y]$$

für alle $x, y \in \mathbb{R}$ gilt.

Dies ist gleichbedeutend mit

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y) .$$

Differentiation ergibt

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) .$$

Für mehrere Zufallsvariablen X_1, \dots, X_n gilt analog: X_1, \dots, X_n sind genau dann unabhängig, wenn

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot \dots \cdot F_{X_n}(x_n)$$

bzw.

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot \dots \cdot f_{X_n}(x_n)$$

für alle $x_1, \dots, x_n \in \mathbb{R}$.

3.3 Warteprobleme mit der Exponentialverteilung

Warten auf mehrere Ereignisse

Satz 102

Die Zufallsvariablen X_1, \dots, X_n seien unabhängig und exponentialverteilt mit den Parametern $\lambda_1, \dots, \lambda_n$. Dann ist auch $X := \min\{X_1, \dots, X_n\}$ exponentialverteilt mit dem Parameter $\lambda_1 + \dots + \lambda_n$.

Beweis:

Der allgemeine Fall folgt mittels Induktion aus dem für $n = 2$. Für die Verteilungsfunktion F_X gilt:

$$\begin{aligned}1 - F_X(t) &= \Pr[X > t] = \Pr[\min\{X_1, X_2\} > t] \\&= \Pr[X_1 > t, X_2 > t] \\&= \Pr[X_1 > t] \cdot \Pr[X_2 > t] \\&= e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = e^{-(\lambda_1 + \lambda_2)t}.\end{aligned}$$



Anschaulich besagt Satz 102, dass sich die Raten addieren, wenn man auf das erste Eintreten eines Ereignisses aus mehreren unabhängigen Ereignissen wartet. Wenn beispielsweise ein Atom die Zerfallsrate λ besitzt, so erhalten wir bei n Atomen die Zerfallsrate $n\lambda$ (wie uns auch die Intuition sagt).

Poisson-Prozess

Wir hatten bei der Diskussion der geometrischen und der Poisson-Verteilung festgestellt:

Wenn der zeitliche Abstand der Treffer geometrisch verteilt ist, so ist ihre Anzahl in einer festen Zeitspanne binomialverteilt.

Im Grenzwert $n \rightarrow \infty$, wobei wir die Trefferwahrscheinlichkeit mit $p_n = \lambda/n$ ansetzen, konvergiert die geometrische Verteilung gegen die Exponentialverteilung und die Binomialverteilung gegen die Poisson-Verteilung. Im Grenzwert $n \rightarrow \infty$ erwarten wir deshalb die folgende Aussage:

Wenn man Ereignisse zählt, deren zeitlicher Abstand exponentialverteilt ist, so ist die Anzahl dieser Ereignisse in einer festen Zeitspanne Poisson-verteilt.

Seien $T_1, T_2 \dots$ unabhängige exponentialverteilte Zufallsvariablen mit Parameter λ . Die Zufallsvariable T_i modelliert die Zeit, die zwischen Treffer $i - 1$ und i vergeht.

Für den Zeitpunkt $t > 0$ definieren wir

$$X(t) := \max\{n \in \mathbb{N} \mid T_1 + \dots + T_n \leq t\}.$$

$X(t)$ gibt also an, wie viele Treffer sich bis zur Zeit t (von Zeit Null ab) ereignet haben. Es gilt:

Fakt 103

Seien T_1, T_2, \dots unabhängige Zufallsvariablen und sei $X(t)$ für $t > 0$ wie oben definiert. Dann gilt: $X(t)$ ist genau dann Poisson-verteilt mit Parameter $t\lambda$, wenn es sich bei T_1, T_2, \dots um exponentialverteilte Zufallsvariablen mit Parameter λ handelt.

Zum Zufallsexperiment, das durch T_1, T_2, \dots definiert ist, erhalten wir für jeden Wert $t > 0$ eine Zufallsvariable $X(t)$. Hierbei können wir t als Zeit interpretieren und $X(t)$ als Verhalten des Experiments zur Zeit t . Eine solche Familie $(X(t))_{t>0}$ von Zufallsvariablen nennt man allgemein einen **stochastischen Prozess**. Der hier betrachtete Prozess, bei dem T_1, T_2, \dots unabhängige, exponentialverteilte Zufallsvariablen sind, heißt **Poisson-Prozess** und stellt ein fundamentales und zugleich praktisch sehr bedeutsames Beispiel für einen stochastischen Prozess dar.

Beispiel 104

Wir betrachten eine Menge von Jobs, die auf einem Prozessor sequentiell abgearbeitet werden. Die Laufzeiten der Jobs seien unabhängig und exponentialverteilt mit Parameter $\lambda = 1/30[1/s]$. Jeder Job benötigt also im Mittel $30s$.

Gemäß Fakt 103 ist die Anzahl von Jobs, die in einer Minute vollständig ausgeführt werden, Poisson-verteilt mit Parameter $t\lambda = 60 \cdot (1/30) = 2$.

Die Wahrscheinlichkeit, dass in einer Minute höchstens ein Job abgearbeitet wird, beträgt in diesem Fall ($t\lambda = 2$)

$$e^{-t\lambda} + t\lambda e^{-t\lambda} \approx 0,406 .$$

3.4 Summen von Zufallsvariablen

Satz 105

Seien X und Y unabhängige kontinuierliche Zufallsvariablen. Für die Dichte von $Z := X + Y$ gilt

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(z - x) \, dx.$$

Beweis:

Nach Definition der Verteilungsfunktion gilt

$$F_Z(t) = \Pr[Z \leq t] = \Pr[X + Y \leq t] = \int_{A(t)} f_{X,Y}(x, y) \, dx \, dy$$

wobei $A(t) = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq t\}$.

Beweis (Forts.):

Aus der Unabhängigkeit von X und Y folgt

$$\begin{aligned} F_Z(t) &= \int_{A(t)} f_X(x) \cdot f_Y(y) \, dx \, dy \\ &= \int_{-\infty}^{\infty} f_X(x) \cdot \left(\int_{-\infty}^{t-x} f_Y(y) \, dy \right) dx. \end{aligned}$$

Mittels der Substitution $z := x + y$, $dz = dy$ ergibt sich

$$\int_{-\infty}^{t-x} f_Y(y) \, dy = \int_{-\infty}^t f_Y(z - x) \, dz$$

und somit

$$F_Z(t) = \int_{-\infty}^t \left(\int_{-\infty}^{\infty} f_X(x) f_Y(z - x) \, dx \right) dz.$$



Satz 106 (Additivität der Normalverteilung)

Die Zufallsvariablen X_1, \dots, X_n seien unabhängig und normalverteilt mit den Parametern μ_i, σ_i ($1 \leq i \leq n$). Es gilt: Die Zufallsvariable

$$Z := a_1 X_1 + \dots + a_n X_n$$

ist normalverteilt mit Erwartungswert $\mu = a_1 \mu_1 + \dots + a_n \mu_n$ und Varianz $\sigma^2 = a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2$.

Beweis:

Wir beweisen zunächst den Fall $n = 2$ und $a_1 = a_2 = 1$. Nach Satz 105 gilt für $Z := X_1 + X_2$, dass

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{X_1}(z-y) \cdot f_{X_2}(y) \, dy \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \underbrace{\left(\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)}_{=:v}\right) \, dy. \end{aligned}$$

Beweis (Forts.):

Wir setzen

$$\begin{aligned}\mu &:= \mu_1 + \mu_2 \\ \sigma^2 &:= \sigma_1^2 + \sigma_2^2 \\ v_1 &:= (z - \mu)/\sigma \\ v_2^2 &:= v - v_1^2\end{aligned}$$

Damit ergibt sich unmittelbar

$$v_2^2 = \frac{(z - y - \mu_1)^2}{\sigma_1^2} + \frac{(y - \mu_2)^2}{\sigma_2^2} - \frac{(z - \mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2},$$

woraus wir

$$v_2 = \frac{y\sigma_1^2 - \mu_2\sigma_1^2 + y\sigma_2^2 - z\sigma_2^2 + \mu_1\sigma_2^2}{\sigma_1\sigma_2\sigma}$$

ermitteln.

Beweis (Forts.):

Damit folgt für die gesuchte Dichte

$$f_Z(z) = \frac{1}{2\pi \cdot \sigma_1 \cdot \sigma_2} \cdot \exp\left(-\frac{v_1^2}{2}\right) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{v_2^2}{2}\right) \mathrm{d}y.$$

Wir substituieren noch

$$t := v_2 \text{ und } \mathrm{d}t = \frac{\sigma}{\sigma_1 \sigma_2} \mathrm{d}y$$

und erhalten

$$f_Z(z) = \frac{1}{2\pi \cdot \sigma} \cdot \exp\left(-\frac{(z - \mu)^2}{2\sigma^2}\right) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t.$$

Mit Lemma 92 folgt, dass $f_Z(z) = \varphi(z; \mu, \sigma)$ ist.

Beweis (Forts.):

Daraus erhalten wir die Behauptung für $n = 2$, denn den Fall $Z := a_1X_1 + a_2X_2$ für beliebige Werte $a_1, a_2 \in \mathbb{R}$ können wir leicht mit Hilfe von Satz 93 auf den soeben bewiesenen Fall reduzieren. Durch Induktion kann die Aussage auf beliebige Werte $n \in \mathbb{N}$ verallgemeinert werden. □

3.5 Momenterzeugende Funktionen für kontinuierliche Zufallsvariablen

Für diskrete Zufallsvariablen X haben wir die momenterzeugende Funktion

$$M_X(s) = \mathbb{E}[e^{Xs}]$$

eingeführt. Diese Definition kann man unmittelbar auf kontinuierliche Zufallsvariablen übertragen. Die für $M_X(s)$ gezeigten Eigenschaften bleiben dabei erhalten.

Beispiel 107

Für eine auf $[a, b]$ gleichverteilte Zufallsvariable U gilt

$$\begin{aligned}M_U(t) &= \mathbb{E}[e^{tX}] = \int_a^b e^{tx} \cdot \frac{1}{b-a} dx \\&= \left[\frac{e^{tx}}{t(b-a)} \right]_a^b \\&= \frac{e^{tb} - e^{ta}}{t(b-a)}.\end{aligned}$$

Beispiel (Forts.)

Für eine standardnormalverteilte Zufallsvariable $N \sim \mathcal{N}(0, 1)$ gilt

$$\begin{aligned}M_N(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{t\xi} e^{-\xi^2/2} \mathrm{d}\xi \\&= e^{t^2/2} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(t-\xi)^2/2} \mathrm{d}\xi \\&= e^{t^2/2} .\end{aligned}$$

Beispiel (Forts.)

Daraus ergibt sich für $Y \sim \mathcal{N}(\mu, \sigma^2)$ wegen $\frac{Y-\mu}{\sigma} \sim \mathcal{N}(0, 1)$

$$\begin{aligned}M_Y(t) &= \mathbb{E}[e^{tY}] \\&= e^{t\mu} \cdot \mathbb{E}[e^{(t\sigma) \cdot \frac{Y-\mu}{\sigma}}] \\&= e^{t\mu} \cdot M_N(t\sigma) \\&= e^{t\mu + (t\sigma)^2/2} .\end{aligned}$$

Weiterer Beweis von Satz 106:

Beweis:

Gemäß dem vorhergehenden Beispiel gilt

$$M_{X_i}(t) = e^{t\mu_i + (t\sigma_i)^2/2}.$$

Wegen der Unabhängigkeit der X_i folgt

$$\begin{aligned} M_Z(t) &= \mathbb{E}[e^{t(a_1 X_1 + \dots + a_n X_n)}] = \prod_{i=1}^n \mathbb{E}[e^{(a_i t) X_i}] \\ &= \prod_{i=1}^n M_{X_i}(a_i t) \\ &= \prod_{i=1}^n e^{a_i t \mu_i + (a_i t \sigma_i)^2/2} \\ &= e^{t\mu + (t\sigma)^2/2}, \end{aligned}$$

mit $\mu = a_1 \mu_1 + \dots + a_n \mu_n$ und $\sigma^2 = a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2$. □

4. Zentraler Grenzwertsatz

Satz 108 (Zentraler Grenzwertsatz)

Die Zufallsvariablen X_1, \dots, X_n besitzen jeweils dieselbe Verteilung und seien unabhängig. Erwartungswert und Varianz von X_i existieren für $i = 1, \dots, n$ und seien mit μ bzw. σ^2 bezeichnet ($\sigma^2 > 0$).

Die Zufallsvariablen Y_n seien definiert durch $Y_n := X_1 + \dots + X_n$ für $n \geq 1$. Dann folgt, dass die Zufallsvariablen

$$Z_n := \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

asymptotisch standardnormalverteilt sind, also $Z_n \sim \mathcal{N}(0, 1)$ für $n \rightarrow \infty$.

Etwas formaler ausgedrückt gilt: Die Folge der zu Z_n gehörenden Verteilungsfunktionen F_n hat die Eigenschaft

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x) \text{ für alle } x \in \mathbb{R}.$$

Wir sagen dazu auch: Die Verteilung von Z_n **konvergiert** gegen die Standardnormalverteilung für $n \rightarrow \infty$.

Dieser Satz ist von großer Bedeutung für die Anwendung der Normalverteilung in der Statistik. Der Satz besagt, dass sich die Verteilung einer Summe **beliebiger** unabhängiger Zufallsvariablen (mit endlichem Erwartungswert und Varianz) der Normalverteilung umso mehr annähert, je mehr Zufallsvariablen an der Summe beteiligt sind.

Beweis:

Wir betrachten $X_i^* := (X_i - \mu)/\sigma$ für $i = 1, \dots, n$ mit $\mathbb{E}[X_i^*] = 0$ und $\text{Var}[X_i^*] = 1$.
Damit gilt (gemäß vorhergehendem **Beispiel**)

$$\begin{aligned} M_Z(t) &= \mathbb{E}[e^{tZ}] = \mathbb{E}[e^{t(X_1^* + \dots + X_n^*)/\sqrt{n}}] \\ &= M_{X_1^*}(t/\sqrt{n}) \cdot \dots \cdot M_{X_n^*}(t/\sqrt{n}). \end{aligned}$$

Für beliebiges i betrachten wir die Taylorentwicklung von $M_{X_i^*}(t) =: h(t)$ an der Stelle $t = 0$

$$h(t) = h(0) + h'(0) \cdot t + \frac{h''(0)}{2} \cdot t^2 + \mathcal{O}(t^3).$$

Aus der Linearität des Erwartungswerts folgt

$$h'(t) = \mathbb{E}[e^{tX_i^*} \cdot X_i^*] \text{ und } h''(t) = \mathbb{E}[e^{tX_i^*} \cdot (X_i^*)^2].$$

Beweis (Forts.):

Damit gilt

$$h'(0) = \mathbb{E}[X_i^*] = 0 \text{ und } h''(0) = \mathbb{E}[(X_i^*)^2] = \text{Var}[X] = 1.$$

Durch Einsetzen in die Taylorreihe folgt $h(t) = 1 + t^2/2 + \mathcal{O}(t^3)$, und wir können $M_Z(t)$ umschreiben zu

$$M_Z(t) = \left(1 + \frac{t^2}{2n} + \mathcal{O}\left(\frac{t^3}{n^{3/2}}\right)\right)^n \rightarrow e^{t^2/2} \text{ für } n \rightarrow \infty.$$

Aus der Konvergenz der momenterzeugenden Funktion folgt auch die Konvergenz der Verteilung. Damit ist Z asymptotisch normalverteilt.

Beweis (Forts.):

Die momenterzeugende Funktion existiert leider nicht bei allen Zufallsvariablen und unser Beweis ist deshalb unvollständig. Man umgeht dieses Problem, indem man statt der momenterzeugenden Funktion die so genannte **charakteristische Funktion** $\tilde{M}_X(t) = \mathbb{E}[e^{itX}]$ betrachtet. Für Details verweisen wir auf die einschlägige Literatur. □

Der Zentrale Grenzwertsatz hat die folgende intuitive Konsequenz:

Wenn eine Zufallsgröße durch lineare Kombination vieler unabhängiger, identisch verteilter Zufallsgrößen entsteht, so erhält man näherungsweise eine Normalverteilung.

Ein wichtiger Spezialfall des Zentralen Grenzwertsatzes besteht darin, dass die auftretenden Zufallsgrößen Bernoulli-verteilt sind.

Korollar 109 (Grenzwertsatz von de Moivre)

X_1, \dots, X_n seien unabhängige Bernoulli-verteilte Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p . Dann gilt für die Zufallsvariable H_n mit

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

für $n \geq 1$, dass die Verteilung der Zufallsvariablen

$$H_n^* := \frac{H_n - np}{\sqrt{np(1-p)}}$$

für $n \rightarrow \infty$ gegen die Standardnormalverteilung konvergiert.

Beweis:

Die Behauptung folgt unmittelbar aus dem Zentralen Grenzwertsatz, da $\mu = \frac{1}{n}\mathbb{E}[H_n] = p$ und $\sigma^2 = \frac{1}{n}\text{Var}[H_n] = p(1 - p)$. □

Bemerkung

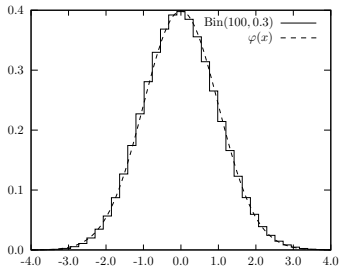
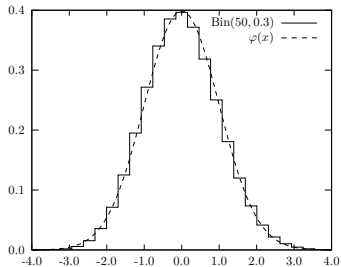
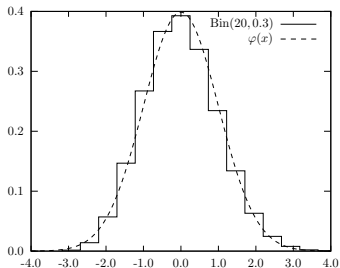
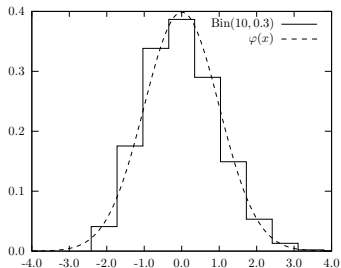
Wenn man X_1, \dots, X_n als Indikatorvariablen für das Eintreten eines Ereignisses A bei n unabhängigen Wiederholungen eines Experimentes interpretiert, dann gibt H_n die absolute Häufigkeit von A an.

4.1 Normalverteilung als Grenzwert der Binomialverteilung

Korollar 109 ermöglicht, die Normalverteilung als Grenzwert der Binomialverteilung aufzufassen. Die folgende Aussage ist eine Konsequenz von Korollar 109:

Korollar 110

Sei $H_n \sim \text{Bin}(n, p)$ eine binomialverteilte Zufallsvariable. Die Verteilung von H_n/n konvergiert gegen $\mathcal{N}(p, p(1-p)/n)$ für $n \rightarrow \infty$.



Vergleich von Binomial- und Normalverteilung

$\text{Bin}(n, 0.3)$ bei $0.3n$ zentriert, mit $\sqrt{0.3 \cdot 0.7n}$ horizontal gestaucht und vertikal gestreckt

Historisch gesehen entstand Korollar 109 vor Satz 108.

Für den Fall $p = 1/2$ wurde Korollar 109 bereits von Abraham de Moivre (1667–1754) bewiesen. De Moivre war gebürtiger Franzose, musste jedoch aufgrund seines protestantischen Glaubens nach England fliehen. Dort wurde er unter anderem Mitglied der Royal Society, erhielt jedoch niemals eine eigene Professur.

Die allgemeine Formulierung von Korollar 109 geht auf Pierre Simon Laplace (1749–1827) zurück. Allerdings vermutet man, dass die Lösung des allgemeinen Falls $p \neq 1/2$ bereits de Moivre bekannt war.

4.2 Elementarer Beweis des Grenzwertsatzes von de Moivre für $p = 1/2$

Wir betrachten die Wahrscheinlichkeit $\Pr[a \leq H_{2n}^* \leq b]$ für $p = 1/2$ und $a, b \in \mathbb{R}$ mit $a \leq b$. Wenn die Verteilung von H_{2n}^* , wie in Korollar 109 angegeben, gegen $\mathcal{N}(0, 1)$ konvergiert, so sollte $\Pr[a \leq H_{2n}^* \leq b] \approx \int_a^b \varphi(t) \, dt$ für genügend große n gelten. Wir schreiben $f(n) \sim_{\infty} g(n)$ für $\lim_{n \rightarrow \infty} f(n)/g(n) = 1$, wollen also zeigen:

$$\Pr[a \leq H_{2n}^* \leq b] \sim_{\infty} \int_a^b \varphi(t) \, dt.$$

Da für $H_{2n} \sim \text{Bin}(2n, 1/2)$ gilt, dass $\mathbb{E}[H_{2n}] = n$ und $\text{Var}[H_{2n}] = n/2$ ist, erhalten wir

$$H_{2n}^* = \frac{H_{2n} - n}{\sqrt{n/2}},$$

und es folgt

$$\begin{aligned}\Pr[a \leq H_{2n}^* \leq b] &= \Pr[n + a\sqrt{n/2} \leq H_{2n} \leq n + b\sqrt{n/2}] \\ &= \sum_{i \in I_n} \Pr[H_{2n} = n + i]\end{aligned}$$

für $I_n := \{z \in \mathbb{Z} \mid a\sqrt{n/2} \leq z \leq b\sqrt{n/2}\}$. Damit ist

$$\Pr[a \leq H_{2n}^* \leq b] = \sum_{i \in I_n} \underbrace{\binom{2n}{n+i} \cdot \left(\frac{1}{2}\right)^{2n}}_{=: p_{n,i}}.$$

Es gilt

$$\max_i p_{n,i} \leq p_n^* := \binom{2n}{n} \cdot \left(\frac{1}{2}\right)^{2n} = \frac{(2n)!}{(n!)^2} \cdot \left(\frac{1}{2}\right)^{2n},$$

und mit der Stirling'schen Approximation für $n!$

$$p_n^* \sim_{\infty} \frac{(2n)^{2n} \cdot e^{-2n} \cdot \sqrt{2\pi \cdot 2n}}{(n^n \cdot e^{-n} \cdot \sqrt{2\pi n})^2} \cdot \left(\frac{1}{2}\right)^{2n} = \frac{1}{\sqrt{\pi n}}.$$

Ersetzen wir nun die $p_{n,i}$ durch p_n^* , so entsteht dabei ein Fehler, den wir mit $q_{n,i} := \frac{p_{n,i}}{p_n^*}$ bezeichnen.

Für $i > 0$ gilt

$$\begin{aligned} q_{n,i} &= \frac{\binom{2n}{n+i} \cdot \left(\frac{1}{2}\right)^{2n}}{\binom{2n}{n} \cdot \left(\frac{1}{2}\right)^{2n}} = \frac{(2n)! \cdot n! \cdot n!}{(n+i)! \cdot (n-i)! \cdot (2n)!} \\ &= \frac{\prod_{j=0}^{i-1} (n-j)}{\prod_{j=1}^i (n+j)} = \prod_{j=1}^i \frac{n-j+1}{n+j} = \prod_{j=1}^i \left(1 - \frac{2j-1}{n+j}\right). \end{aligned}$$

Wegen der Symmetrie der Binomialkoeffizienten gilt $q_{n,-i} = q_{n,i}$, womit auch der Fall $i < 0$ abgehandelt ist.

Man macht sich leicht klar, dass $1 - 1/x \leq \ln x \leq x - 1$ für $x > 0$ gilt. Damit schließen wir, dass

$$\begin{aligned} \ln \left(\prod_{j=1}^i \left(1 - \frac{2j-1}{n+j} \right) \right) &= \sum_{j=1}^i \ln \left(1 - \frac{2j-1}{n+j} \right) \\ &\leq - \sum_{j=1}^i \frac{2j-1}{n+j} \leq - \sum_{j=1}^i \frac{2j-1}{n+i} \\ &= - \frac{i(i+1) - i}{n+i} = - \frac{i^2}{n} + \frac{i^3}{n(n+i)} \\ &= - \frac{i^2}{n} + \mathcal{O} \left(\frac{1}{\sqrt{n}} \right), \end{aligned}$$

da $i = \mathcal{O}(\sqrt{n})$ für $i \in I_n$.

Ebenso erhalten wir

$$\begin{aligned}\ln \left(\prod_{j=1}^i \left(1 - \frac{2j-1}{n+j} \right) \right) &\geq \sum_{j=1}^i \left(1 - \left(1 - \frac{2j-1}{n+j} \right)^{-1} \right) \\ &= \sum_{j=1}^i \frac{-2j+1}{n-j+1} \geq - \sum_{j=1}^i \frac{2j-1}{n-i} \\ &= -\frac{i^2}{n-i} = -\frac{i^2}{n} - \mathcal{O} \left(\frac{1}{\sqrt{n}} \right).\end{aligned}$$

Zusammen haben wir

$$e^{-\frac{i^2}{n-i}} = e^{-\frac{i^2}{n} - \mathcal{O} \left(\frac{1}{\sqrt{n}} \right)} \leq q_{n,i} \leq e^{-\frac{i^2}{n} + \mathcal{O} \left(\frac{1}{\sqrt{n}} \right)}$$

Wegen $e^{\pm \mathcal{O}(1/\sqrt{n})} = 1 \pm o(1)$ folgt daraus $q_{n,i} \sim_{\infty} e^{-i^2/n}$.

Damit schätzen wir nun $\Pr[a \leq H_{2n}^* \leq b]$ weiter ab:

$$\Pr[a \leq H_{2n}^* \leq b] = \sum_{i \in I_n} p_n^* \cdot q_{n,i} \sim_{\infty} \underbrace{\frac{1}{\sqrt{\pi n}} \cdot \sum_{i \in I_n} e^{-i^2/n}}_{=: S_n}.$$

Mit $\delta := \sqrt{2/n}$ können wir die Summe S_n umschreiben zu

$$S_n = \frac{1}{\sqrt{2\pi}} \cdot \sum_{i \in I_n} \delta e^{-(i\delta)^2 \cdot \frac{1}{2}}.$$

Diese Summe entspricht einer Näherung für $\int_a^b \varphi(t) \, dt = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} \, dt$ durch Aufteilung der integrierten Fläche in Balken der Breite δ . Für $n \rightarrow \infty$ konvergiert die Fläche der Balken gegen das Integral, d. h. $S_n \sim_{\infty} \int_a^b \varphi(t) \, dt$.

q. e. d.

4.3 Verschiedene Approximationen der Binomialverteilung

Sei $H_n \sim \text{Bin}(n, p)$ eine binomialverteilte Zufallsvariable mit der Verteilungsfunktion F_n . Für $n \rightarrow \infty$ gilt

$$\begin{aligned} F_n(t) &= \Pr[H_n/n \leq t/n] \\ &\rightarrow \Phi\left(\frac{t/n - p}{\sqrt{p(1-p)/n}}\right) = \Phi\left(\frac{t - np}{\sqrt{p(1-p)n}}\right). \end{aligned}$$

Wir können F_n somit für große n durch Φ approximieren. Diese Approximation ist in der Praxis deshalb von Bedeutung, da die Auswertung der Verteilungsfunktion der Binomialverteilung für große n sehr aufwändig ist, während für die Berechnung der Normalverteilung effiziente numerische Methoden vorliegen.

Beispiel 111

Wenn man die Wahrscheinlichkeit berechnen möchte, mit der bei 10^6 Würfeln mit einem idealen Würfel mehr als 500500-mal eine gerade Augenzahl fällt, so muss man eigentlich folgenden Term auswerten:

$$T := \sum_{i=5,005 \cdot 10^5}^{10^6} \binom{10^6}{i} \left(\frac{1}{2}\right)^{10^6}.$$

Dies ist numerisch kaum effizient möglich.

Die numerische Integration der Dichte φ der Normalverteilung ist hingegen relativ einfach. Auch andere Approximationen der Verteilung Φ , beispielsweise durch Polynome, sind bekannt. Entsprechende Funktionen werden in zahlreichen Softwarebibliotheken als „black box“ angeboten.

Beispiel

Mit der Approximation durch die Normalverteilung erhalten wir

$$\begin{aligned} T &\approx 1 - \Phi\left(\frac{5,005 \cdot 10^5 - 5 \cdot 10^5}{\sqrt{2,5 \cdot 10^5}}\right) \\ &= 1 - \Phi\left(\frac{5 \cdot 10^2}{5 \cdot 10^2}\right) \\ &= 1 - \Phi(1) \approx 0,1573 . \end{aligned}$$

Bei der Approximation der Binomialverteilung mit Hilfe von Korollar 109 führt man oft noch eine so genannte **Stetigkeitskorrektur** durch. Zur Berechnung von $\Pr[X \leq x]$ für $X \sim \text{Bin}(n, p)$ setzt man

$$\Pr[X \leq x] \approx \Phi \left(\frac{x + 0,5 - np}{\sqrt{np(1-p)}} \right)$$

statt

$$\Pr[X \leq x] \approx \Phi \left(\frac{x - np}{\sqrt{np(1-p)}} \right)$$

an.

Der Korrekturterm läßt sich in der **Histogramm-Darstellung** der Binomialverteilung veranschaulichen. Die Binomialverteilung wird dort durch Balken angegeben, deren Fläche in etwa der Fläche unterhalb der Dichte φ von $\mathcal{N}(0, 1)$ entspricht. Wenn man die Fläche der Balken mit „ $X \leq x$ “ durch das Integral von φ approximieren möchte, so sollte man bis zum Ende des Balkens für „ $X = x$ “ integrieren und nicht nur bis zur Mitte. Dafür sorgt der Korrekturterm 0,5.

Approximationen für die Binomialverteilung

- **Approximation durch die Poisson-Verteilung:** $\text{Bin}(n, p)$ wird approximiert durch $\text{Po}(np)$. Diese Approximation funktioniert sehr gut für seltene Ereignisse, d. h. wenn np sehr klein gegenüber n ist. Als Faustregel fordert man $n \geq 30$ und $p \leq 0,05$.
- **Approximation durch die Chernoff-Schranken:** Bei der Berechnung der **tails** der Binomialverteilung liefern diese Ungleichungen meist sehr gute Ergebnisse. Ihre Stärke liegt darin, dass es sich bei den Schranken nicht um Approximationen, sondern um echte Abschätzungen handelt. Dies ist vor allem dann wichtig, wenn man nicht nur numerische Näherungen erhalten möchte, sondern allgemeine Aussagen über die Wahrscheinlichkeit von Ereignissen beweisen möchte.

- **Approximation durch die Normalverteilung:** Als Faustregel sagt man, dass die Verteilungsfunktion $F_n(t)$ von $\text{Bin}(n, p)$ durch

$$F_n(t) \approx \Phi((t - np)/\sqrt{p(1-p)n})$$

approximiert werden kann, wenn $np \geq 5$ und $n(1-p) \geq 5$ gilt.