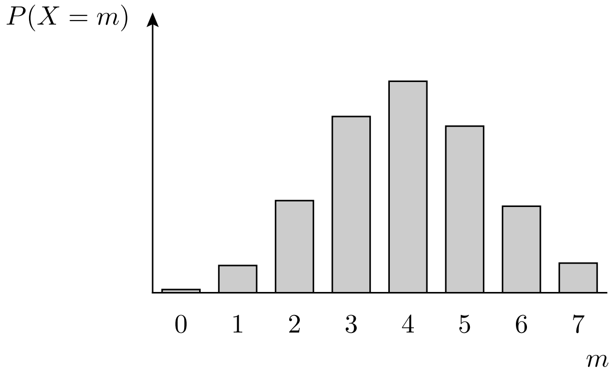
1.3.2.5 Hypergeometrische Verteilung

Als Referenzmodell dient die bereits bekannte Urne mit N Kugeln, von denen M Kugeln schwarz und $N-M$ Kugeln weiß sind. Wir ziehen ohne Zurücklegen n Kugeln, wobei unsere Zufallsgröße X die Anzahl der entnommenen schwarzen Kugeln ist. Dann gilt:

$$P(X=m)=\frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}},$$

wobei $\max(0, n-(N-M)) \leq m \leq \min(n, M)$ ist.

Definition 1.3.19. Eine Zufallgröße X mit der obigen Wahrscheinlichkeitsfunktion heißt *hypergeometrisch verteilt* und wir schreiben $X \sim \mathbf{H}(n, N, M)$.



Erwartungswert Der Erwartungswert für $X \sim \mathbf{H}(n, N, M)$ beträgt:

$$\mathbf{E}X = n\frac{M}{N}.$$

Varianz Die Varianz für $X \sim \mathbf{H}(n, N, M)$ beträgt:

$$\mathbf{D}^2X = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}.$$

1.3.3 Stetige Zufallsgrößen

Im Abschnitt über Wahrscheinlichkeitsräume haben wir bereits die Brenndauer einer Glühlampe und die Reichweite eines Fahrzeugs bei begrenztem Treibstoffvorrat als Beispiele für stetige Zufallsgrößen betrachtet. Da stetige Zufallsgrößen überabzählbar unendlich viele Werte besitzen und somit deren Werte ganze Intervalle der reellen Achse ausfüllen können, ist es nicht mehr möglich, die Wahrscheinlichkeit für jeden einzelnen Wert in einer Wahrscheinlichkeitsfunktion auszudrücken. Jedoch kann man mit Hilfe sogenannter Dichtefunktionen die Verteilung der Wahrscheinlichkeitsmasse auf der reellen Achse angeben und so die Wahrscheinlichkeit dafür charakterisieren, dass der Wert der Zufallsgröße in einem gegebenen Intervall liegt.

Definition 1.3.20. Eine Zufallsgröße X heißt *stetig*, wenn es eine integrierbare reelle Funktion f gibt, so dass für beliebige reelle Zahlen $a \le b$ gilt:

$$P(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x.$$

Die Funktion f heißt *Dichtefunktion* der Zufallsgröße X.

Eigenschaften von Dichtefunktionen

- $f(x) \ge 0 \quad$ für alle $x \in \mathbb{R}$,
- $\int\limits_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$

Das Integral ist dabei im Sinne von Riemann oder Lebesgue zu verstehen. Als Dichtefunktionen f treten vorzugsweise stetige und stückweise stetige Funktionen auf, die auch schwache Polstellen besitzen dürfen. Die Fläche unter dem Graphen von f bleibt mit dem Wert 1 jedoch stets endlich. Wegen

$$P(X = a) = P(a \le X \le a) = \int_a^a f(x)\,\mathrm{d}x = 0$$

ist die Wahrscheinlichkeit, dass X genau einen festen Wert annimmt, immer gleich Null.

Definition 1.3.21. Die durch

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

definierte reelle Funktion F heißt *Verteilungsfunktion* der stetigen Zufallsgröße X.

Eigenschaften der Verteilungsfunktion

- $\lim\limits_{x \to -\infty} F(x) = 0.$
- $\lim\limits_{x \to \infty} F(x) = 1.$
- $x_1 < x_2 \Rightarrow F(x_1) \le F(x_2)$, d.h. F ist monoton wachsend (nicht notwendigerweise streng).
- $P(a \le X \le b) = F(b) - F(a).$
- F ist stetig in allen Punkten $x \in \mathbb{R}$.
- Falls die Dichtefunktion f in x_0 stetig ist, so ist F in x_0 differenzierbar und es gilt $F'(x_0) = f(x_0)$.

1.3.3.1 Erwartungswert und Varianz

Definition 1.3.22. Der *Erwartungswert* einer stetigen Zufallsgröße X ist gegeben durch

$$\mathbf{E}X := \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x.$$

$\mathbf{E}X$ ist eine endliche Zahl, wenn gilt

$$\int_{-\infty}^{\infty} |x| f(x)\,\mathrm{d}x < \infty.$$

Definition 1.3.23. Sei X eine stetige Zufallsgröße. Der *Erwartungswert einer Funktion* $g(X)$ ist gegeben durch

$$\mathbf{E}g(X) := \int_{-\infty}^{\infty} g(x) f(x)\,\mathrm{d}x.$$

$\mathbf{E}g(X)$ ist eine endliche Zahl, wenn gilt

$$\int_{-\infty}^{\infty} |g(x)| f(x)\,\mathrm{d}x < \infty.$$

Definition 1.3.24. Die *Varianz* (oder *Streuung*) einer stetigen Zufallsgröße X ist wie im diskreten Fall definiert durch

$$\sigma^2 := \mathbf{D}^2 X := \mathbf{E}(X - \mathbf{E}X)^2.$$

Die folgenden Sätze aus Abschnitt 1.3.2.1 gelten auch für stetige Zufallsgrößen:

- Hilfssatz 1.3.8: $\mathbf{E}(aX + b) = a\mathbf{E}X + b$.
- Hilfssatz 1.3.10: $\mathbf{D}^2(aX + b) = a^2 \mathbf{D}^2 X$.
- Satz 1.3.11: $\mathbf{D}^2 X = \mathbf{E}X^2 - (\mathbf{E}X)^2$.
- Satz 1.3.15 (Tschebyscheff'sche Ungleichung): Für $\varepsilon > 0$ gilt:

$$P(|X - \mathbf{E}X| > \varepsilon) < \frac{\mathbf{D}^2 X}{\varepsilon^2}.$$

Beweis. Sei $M := \{x : |x - \mathbf{E}X| > \varepsilon\}$. Dann gilt:

$$\begin{aligned}\mathbf{D}^2 X &= \int_{-\infty}^{\infty} (x - \mathbf{E}X)^2 f(x)\,\mathrm{d}x \geq \int_M (x - \mathbf{E}X)^2 f(x)\,\mathrm{d}x \\ &> \varepsilon^2 \int_M f(x)\,\mathrm{d}x = \varepsilon^2 P(M) = \varepsilon^2 P(|x - \mathbf{E}X| > \varepsilon).\end{aligned}$$

□

1.3.3.2 Gleichverteilung

Als erste stetige Wahrscheinlichkeitsverteilung betrachten wir die recht einfache Gleichverteilung. Wir nennen eine Zufallsgröße X gleichverteilt auf dem Intervall $[a,b]$, wenn X nur Werte aus dem Intervall annehmen kann und diese gleichwahrscheinlich über das Intervall verteilt sind. Für die Dichtefunktion ergibt sich also:

$$f(x) = \begin{cases} c, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases},$$

wobei $c = const$ eine Konstante ist. Wegen

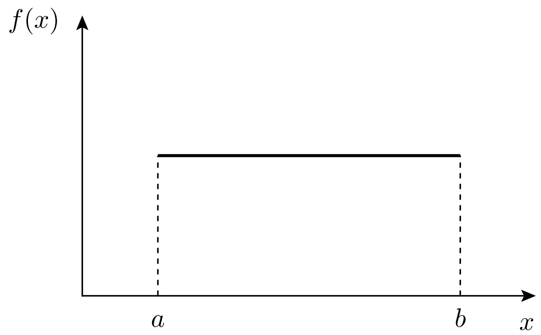
$$c(b-a) = \int\limits_a^b c \,\mathrm{d}x = \int\limits_a^b f(x)\,\mathrm{d}x = \int\limits_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

erhalten wir $c = \frac{1}{b-a}$.

Definition 1.3.25. Ein stetige Zufallsgröße X mit der Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}$$

heißt *gleichverteilt* mit den beiden Parametern a und b.



Die Verteilungsfunktion F nimmt offensichtlich für $x < a$ den Wert 0 und für $x > b$ den Wert 1 an. Für $a \leq x \leq b$ ergibt sich:

$$F(x) = \int\limits_{-\infty}^{x} f(t)\,\mathrm{d}t = \int\limits_a^x \frac{1}{b-a}\,\mathrm{d}t = \left.\frac{t}{b-a}\right|_a^x = \frac{x-a}{b-a},$$

also ist

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}.$$

Erwartungswert Für den Erwartungswert einer gleichverteilten stetigen Zufallsgröße X erhalten wir:

$$\mathbf{E}X = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \int_a^b \frac{x}{b-a}\,\mathrm{d}x = \left.\frac{x^2}{2(b-a)}\right|_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

Zur Berechnung der Varianz benötigen wir noch $\mathbf{E}X^2$.

$$\mathbf{E}X^2 = \int_a^b \frac{x^2}{b-a}\,\mathrm{d}x = \left.\frac{x^3}{3(b-a)}\right|_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}.$$

Varianz Für die Varianz einer gleichverteilten stetigen Zufallsgröße X erhalten wir also:

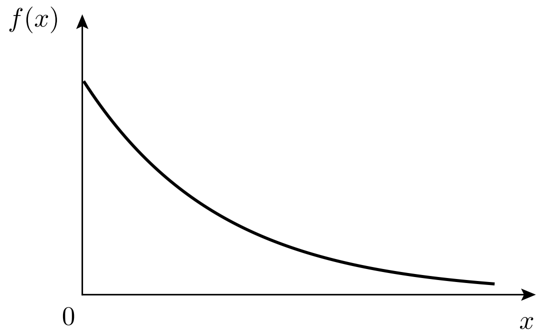
$$\mathbf{D}^2X = \mathbf{E}X^2 - (\mathbf{E}X)^2 = \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{a^2-2ab+b^2}{12} = \frac{(a-b)^2}{12}.$$

1.3.3.3 Exponentialverteilung

Definition 1.3.26. Besitzt eine stetige Zufallsgröße X die Dichtefunktion

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \lambda e^{-\lambda x}, & x > 0 \end{cases},$$

so nennen wir X *exponentialverteilt* mit dem Parameter $\lambda > 0$ und schreiben $X \sim \mathbf{Ex}(\lambda)$.



Aus der Dichtefunktion f erhält man die Verteilungsfunktion

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\lambda x}, & x > 0 \end{cases}.$$

Zusammenhang zwischen Exponential- und Poisson-Verteilung Im Unterabschnitt über die Poisson-Verteilung haben wir als Modell eine Telefonzentrale betrachtet, wobei die Zufallsgröße X_t die Anzahl der Anrufe in einem Zeitintervall der Länge t beschrieb. X_t war Poisson-verteilt mit dem Parameter μ. Dabei gab μ die durchschnittliche Anrufanzahl pro Zeiteinheit an. Dieses Modell können wir auch nutzen, um die Exponentialverteilung zu veranschaulichen. Betrachten wir als Zufallsgröße T die Länge des Zeitintervalls zwischen zwei eingehenden Anrufen, so ist T exponentialverteilt mit demselben Parameter μ wie bei der Poisson-Verteilung.

Beispiel. In einer Telefonzentrale kommen im Mittel 20 Anrufe pro Stunde an. Gesucht ist die Wahrscheinlichkeit, dass zwischen zwei Anrufen 3 bis 6 Minuten vergehen. Rechnen wir in Minuten, so ist $\mu = \frac{20}{60} = \frac{1}{3}$. Wir erhalten dann:

$$P(3 \leq T \leq 6) = F(6) - F(3) = \left(1 - e^{-6\mu}\right) - \left(1 - e^{-3\mu}\right) = e^{-1} - e^{-2} = 0{,}2325.$$

Exponentialverteilung als Lebensdauerverteilung Wartezeiten, Reparaturzeiten und die Lebensdauer von Bauelementen können als exponentialverteilt angenommen werden. Wie die folgende Überlegung zeigt, muss dabei jedoch beachtet werden, dass keine Alterungseffekte modelliert werden können: Für $X \sim \mathbf{Ex}(\lambda)$ gilt

$$\begin{aligned} P(X \leq x_0 + x | X \geq x_0) &= \frac{P(x_0 \leq X \leq x_0 + x)}{P(X \geq x_0)} = \frac{F(x_0 + x) - F(x_0)}{1 - F(x_0)} \\ &= \frac{\left(1 - e^{-\lambda(x_0+x)}\right) - \left(1 - e^{-\lambda x_0}\right)}{1 - (1 - e^{-\lambda x_0})} = \frac{e^{-\lambda x_0} - e^{-\lambda(x_0+x)}}{e^{-\lambda x_0}} \\ &= 1 - e^{-\lambda x} = P(X \leq x), \end{aligned}$$

d.h. wenn wir als Zufallsgröße X die Lebensdauer eines Bauelements betrachten, so ist die Wahrscheinlichkeit, dass das Bauelement innerhalb einer Zeitdauer x eine Störung aufweist, unabhängig davon, ob es bereits über eine Zeitdauer x_0 in Betrieb war oder ob es neu ist.

Erwartungswert Der Erwartungswert für $X \sim \mathbf{Ex}(\lambda)$ beträgt:

$$\mathbf{E}X = \int_0^\infty x\lambda e^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda}\int_0^\infty te^{-t}\,\mathrm{d}t = \frac{1}{\lambda}\left(t\left(-e^{-t}\right)\Big|_0^\infty + \int_0^\infty e^{-t}\,\mathrm{d}t\right) = \frac{1}{\lambda}.$$

In ähnlicher Weise berechnet man $\mathbf{E}X^2 = \frac{2}{\lambda^2}$.

Varianz Die Varianz für $X \sim \mathbf{Ex}(\lambda)$ beträgt:

$$\mathbf{D}^2X = \mathbf{E}X^2 - (\mathbf{E}X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Beispiel. Als Zufallsgröße X betrachten wir die Zeitdauer für eine PKW-Inspektion in einer Werkstatt. Im Mittel dauert eine Inspektion 2 Stunden. Wie groß ist die Wahrscheinlichkeit, dass eine Inspektion länger als 3 Stunden dauert? Als Einheit für unsere Berechnung wählen wir Stunden und es sei $X \sim \mathbf{Ex}(\lambda)$. Somit erhalten wir aus $\mathbf{E}X = 2$ den Parameter $\lambda = \frac{1}{2}$. Es

ergibt sich

$$P(X > 3) = P(X \geq 3) = P(3 \leq X < \infty) = F(\infty) - F(3) = 1 - \left(1 - e^{-3\lambda}\right) = e^{-\frac{3}{2}} = 0{,}223,$$

d.h. in durchschnittlich 22,3 % aller Fälle dauert die Inspektion länger als 3 Stunden.

1.3.3.4 Normalverteilung

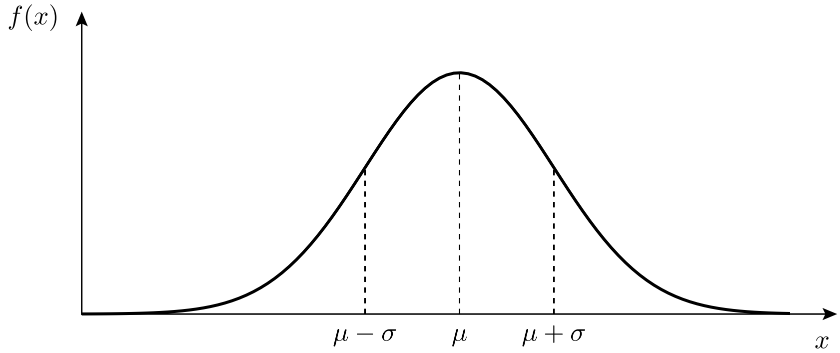
Die Normalverteilung (oder auch GAUSS'sche Verteilung) ist die wichtigste stetige Verteilung, da sie in der Praxis eine Vielzahl von Anwendungen hat.

Definition 1.3.27. Besitzt eine stetige Zufallsgröße X die Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}},$$

so nennen wir X *normalverteilt* mit den Parametern μ und σ^2 $(\sigma > 0)$ und schreiben $X \sim \mathbf{N}(\mu, \sigma^2)$.

Interpretation der Parameter Die Dichtefunktion der Normalverteilung wird aufgrund ihrer Form als *Glockenkurve* bezeichnet. Glockenkurven sind symmetrische, eingipfelige Kurven mit Wendestellen bei $x = \mu \pm \sigma$ und einem auf der Symmetrieachse liegenden Maximum (Top der Glocke) von $\frac{1}{\sqrt{2\pi}\sigma}$. Wir nennen $\mu \in \mathbb{R}$ den *Lageparameter*, da μ die Lage der Symmetrieachse angibt, und $\sigma^2 > 0$ den *Formparameter*, da σ^2 den Breitenverlauf der Glockenkurve festlegt. Bei großem σ ist die Glockenkurve breit gezogen, bei kleinem σ ist sie nadelförmig.



Verteilungsfunktion Die Verteilungsfunktion F einer normalverteilten Zufallsgröße ist gegeben durch

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t.$$

F ist nicht durch einen geschlossenen analytischen Ausdruck darstellbar. Die Funktionswerte müssen mittels nummerischer Integration oder durch andere Techniken näherungsweise bestimmt werden. Weiter unten werden wir sehen, dass es genügt, die Werte der Verteilungsfunktion für $\mu = 0$ und $\sigma = 1$ zu kennen. Diese sind in Tabellen erfasst.

Erwartungswert und Varianz Für $X \sim \mathbf{N}(\mu, \sigma^2)$ ist der Erwartungswert $\mathbf{E}X = \mu$ und die Varianz beträgt $\mathbf{D}^2 X = \sigma^2$.

Standardisierung einer Zufallsgröße Die lineare Transformation

$$Y := \frac{X - \mathbf{E}X}{\sqrt{\mathbf{D}^2 X}}$$

einer Zufallsgröße X heißt *Standardisierung* von X. Aufgrund der Linearität dieser Transformation besitzt Y die gleiche Verteilungsart wie X. Für den Erwartungswert von Y erhalten wir

$$\mathbf{E}Y = \mathbf{E}\frac{X - \mathbf{E}X}{\sqrt{\mathbf{D}^2 X}} = \frac{1}{\mathbf{D}^2 X}(\mathbf{E}X - \mathbf{E}X) = 0$$

und die Varianz beträgt

$$\mathbf{D}^2 Y = \mathbf{E}Y^2 = \mathbf{E}\frac{(X - \mathbf{E}X)^2}{D^2 X} = \frac{1}{\mathbf{D}^2 X}\mathbf{E}(X - \mathbf{E}X)^2 = \frac{1}{\mathbf{D}^2 X}\mathbf{D}^2 X = 1.$$

Standardisierung einer normalverteilten Zufallsgröße Wenden wir das beschriebene Standardisierungsverfahren auf eine Zufallsgröße $X \sim \mathbf{N}(\mu, \sigma^2)$ an, so erhalten wir die entsprechende *standardisiert normalverteilte* Zufallsgröße $Y \sim \mathbf{N}(0, 1)$ mit $Y = \frac{X-\mu}{\sigma}$. Als Dichtefunktion der standardisierten Normalverteilung ergibt sich

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

und somit ist die Verteilungsfunktion

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, \mathrm{d}t.$$

Für $x \geq 0$ sind die Funktionswerte von Φ tabelliert. Für $x < 0$ nutzt man den aus der Symmetrie der Glockenkurve resultierenden Zusammenhang $\Phi(-x) = 1 - \Phi(x)$.

Berechnung von Wahrscheinlichkeiten In der Praxis müssen oft Wahrscheinlichkeiten des Typs $P(a \leq X \leq b)$ mit einer Zufallsgröße $X \sim \mathbf{N}(\mu, \sigma^2)$ berechnet werden. Durch Ausnutzung der Standardisierung einer Zufallsgröße führt man solche Berechnungen auf die Berechnung einer Differenz zweier Werte der Verteilungsfunktion Φ der standardisierten Normalverteilung zurück, da diese Werte in Tabellen erfasst sind:

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} \leq Y \leq \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

Anwendung normalverteilter Zufallsgrößen Stetige Fehlergrößen (Messfehler usw.) können im Allgemeinen in guter Näherung als normalverteilt angenommen werden. Die Normalverteilung ist insbesondere dann für die Beschreibung von stochastischen Modellen geeignet, wenn

sich die betrachtete Zufallsgröße als Summe einer großen Anzahl von unabhängigen Einflüssen ergibt (z.B. als Summe zahlreicher kleiner Fehler oder Störungen).

Satz 1.3.28 (Additionssatz). *Seien $X_i \sim \mathbf{N}(\mu_i, \sigma_i^2)$ für $i = 1, 2, \ldots, n$ vollständig unabhängige normalverteilte Zufallsgrößen. Dann gilt*

$$Z := \sum_{i=1}^{n} X_i \sim \mathbf{N}\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right),$$

d.h. die Summe Z ist wieder eine normalverteilte Zufallsgröße.

Beispiel. Der Kern eines Transformators bestehe aus 25 Blechen und 24 zwischen diesen Blechen liegenden Isolierschichten. Für die Dicken (in Millimeter) X_i der Bleche und Y_j der Isolierschichten gelte $X_i \sim \mathbf{N}(0{,}8; 0{,}04^2)$ und $Y_j \sim \mathbf{N}(0{,}2; 0{,}03^2)$. Uns interessieren die folgenden beiden Fragen:

1. Wie groß ist die Wahrscheinlichkeit, dass zwei Bleche und eine Isolierschicht zusammen dicker als 1,85 mm sind?
2. Die Spulenöffnung sei 25,3 mm breit. Wie groß ist die Wahrscheinlichkeit, dass der Kern zu dick ist?

Wir wissen aus dem vorhergehenden Satz, dass $Z := X_1 + Y_1 + X_2 \sim \mathbf{N}(1{,}8; 0{,}0041)$ ist. Somit erhalten wir als Antwort auf Frage 1:

$$P(Z > 1{,}85) = P\left(\frac{Z - 1{,}8}{\sqrt{0{,}0041}} \geq \frac{1{,}85 - 1{,}8}{\sqrt{0{,}0041}}\right) = 1 - \Phi(0{,}7809) = 0{,}2174.$$

Mit $Z := \sum\limits_{i=1}^{25} X_i + \sum\limits_{j=1}^{24} Y_j \sim \mathbf{N}(24{,}8; 0{,}0616)$ ergibt sich für Frage 2:

$$P(Z > 25{,}3) = P\left(\frac{Z - 24{,}8}{\sqrt{0{,}0616}} \geq \frac{25{,}3 - 24{,}8}{\sqrt{0{,}0616}}\right) = 1 - \Phi(2{,}015) = 0{,}022.$$

1.3.3.5 Schiefe und Exzess

Wir betrachten neben der Varianz $\sigma^2 = \mathbf{D}^2 X$, d.h. neben dem zweiten zentralen Moment μ_2 (siehe Definition 1.3.12), nun auch die dritten und vierten zentralen Momente $\mu_3 = \mathbf{E}(X - \mathbf{E}X)^3$ und $\mu_4 = \mathbf{E}(X - \mathbf{E}X)^4$ einer Zufallsgröße X.

Definition 1.3.29. Sei X eine Zufallsgröße. Dann heißen

$$\gamma_1 := \frac{\mu_3}{\sigma^3} = \frac{\mu^3}{\left(\sqrt{\mu_2}\right)^3} \qquad \text{und} \qquad \gamma_2 := \frac{\mu_4}{\sigma^4} - 3$$

Schiefe von X und *Exzess* von X.

Die Schiefe γ_1 ist ein Maß für die Asymmetrie der Verteilung, also für die Abweichung des Verhaltens der Zufallsgröße X von dem einer symmetrischen Verteilung. Da bei einer symmetrischen stetigen Zufallsgröße X für alle $x \in \mathbb{R}$ $f(\mu - x) = f(\mu + x)$ gilt, wobei $x = \mu = \mathbf{E}X$

die Symmetrieachse der (symmetrischen) Dichtefunktion f ist, erhalten wir

$$\mu_3 = \mathbf{E}(X - \mathbf{E}X)^3 = \int\limits_{-\infty}^{\infty} (x-\mu)^3 f(x)\,\mathrm{d}x = \int\limits_{-\infty}^{\mu} (x-\mu)^3 f(x)\,\mathrm{d}x + \int\limits_{\mu}^{\infty} (x-\mu)^3 f(x)\,\mathrm{d}x$$

$$= -\int\limits_{0}^{\infty} x^3 f(\mu - x)\,\mathrm{d}x + \int\limits_{0}^{\infty} x^3 f(\mu + x)\,\mathrm{d}x = 0$$

und somit ist die Schiefe γ_1 einer symmetrischen Zufallsgröße gleich Null.

Der Exzess γ_2 ist ein Maß für die Abweichung der Zufallsgröße X von der Normalverteilung. Den Quotienten $\frac{\mu_4}{\sigma^4}$ nennt man *Wölbung*. Der Exzess ist also die um 3 verminderte Wölbung. Wie wir weiter unten sehen werden, gilt für eine normalverteilte Zufallsgröße $\mu_4 = 3\sigma^4$. Somit ist der Exzess einer normalverteilten Zufallsgröße gleich Null.

Satz 1.3.30. *Schiefe und Exzess einer Zufallsgröße X bleiben bei Standardisierung unverändert, d.h. mit $Y := \frac{X-\mu}{\sigma}$ gilt $\gamma_1(X) = \gamma_1(Y)$ und $\gamma_2(X) = \gamma_2(Y)$.*

Beweis. Da Y eine standardisierte Zufallsgröße ist, gilt $\mathbf{E}Y = 0$ und $D^2Y = 1$. Somit erhalten wir

$$\gamma_1(Y) = \frac{\mathbf{E}(Y - \mathbf{E}Y)^3}{\left(\sqrt{\mathbf{D}^2 Y}\right)^3} = \mathbf{E}Y^3 = \mathbf{E}\left(\frac{X-\mu}{\sigma}\right)^3 = \frac{\mathbf{E}(X - \mathbf{E}X)^3}{\sigma^3} = \frac{\mu_3}{\sigma^3} = \gamma_1(X).$$

Der Beweis für γ_2 erfolgt analog. □

Satz 1.3.31. *Existieren für eine Zufallsgröße X die ersten vier zentralen Momente, so gilt*

$$\gamma_2 \geq \gamma_1^2 - 2.$$

Satz 1.3.32. *Sei $X \sim \mathbf{N}(\mu, \sigma^2)$ eine normalverteilte Zufallsgröße. Dann gilt für $k = 1, 2, \ldots$*

$$\mu_{2k-1} = 0, \qquad \mu_{2k} = \frac{(2k)!}{2^k k!}\sigma^{2k}$$

und $\gamma_1 = \gamma_2 = 0$.

1.3.3.6 Die charakteristische Funktion

Zur Charakterisierung der Verteilung einer Zufallsgröße X kann neben der Verteilungsfunktion $F(x)$ auch die (komplexwertige) charakteristische Funktion $\varphi_X(t)$ verwendet werden. Wir betrachten dies für stetige Zufallsgrößen X.

Definition 1.3.33. Sei X eine stetige Zufallsgröße. Dann heißt

$$\begin{aligned} \varphi_X(t) &:= \mathbf{E}e^{itX} \qquad (t \in \mathbb{R}) \\ &= \int\limits_{-\infty}^{\infty} e^{itx} f(x) dx \end{aligned}$$

charakteristische Funktion von X. Dabei bezeichnet $f(x)$ die Dichtefunktion von X.

Bemerkung. a) Aus der trigonometrischen Darstellung einer komplexen Zahl folgt

$$\varphi_X(t) = \mathbf{E}(\cos tX + i\sin tX) = \underbrace{\mathbf{E}\cos tX}_{\text{Realteil}} + i\underbrace{\mathbf{E}\sin tX}_{\text{Imaginärteil}}$$

$$\Longrightarrow |\varphi_X(t)| \leq \int_{-\infty}^{\infty} \underbrace{|e^{itx}|}_{=\sqrt{\cos^2\varphi+\sin^2\varphi}=1} f(x)dx = \$ \int_{-\infty}^{\infty} f(x)dx = 1 \qquad \forall t \in \mathbb{R}$$

b) $\varphi_X(t) = \overline{\varphi_X(-t)} = \mathbf{E}\cos tX + i\mathbf{E}\sin tX$

c) Sei $Y = aX + b$. Dann ist $\varphi_Y(t) = \mathbf{E}e^{it(aX+b)} = e^{itb}\mathbf{E}e^{itaX} = e^{itb}\varphi_X(at)$
Speziell bei der Standardisierung: $b = -\frac{\mu}{\sigma},\ a = \frac{1}{\sigma} :\Rightarrow \varphi_Y(t) = e^{-\frac{\mu it}{\sigma}}\varphi_X(\frac{t}{\sigma}$

d) $\varphi_X(t)$ ist eine gleichmäßig stetige Funktion, d. h. es gilt

$$|\varphi_X(t) - \varphi_X(t')| < \varepsilon, \text{ sobald } |t - t'| < \delta(\varepsilon)$$

Beispiel. Sei $X \sim \mathbf{N}(0,1)$. Unter Benutzung des komplexen Integrals $\int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}dx = \sqrt{2\pi}$ gilt

$$\begin{aligned}
\varphi_X(t) &= \int_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{itx-\frac{x^2}{2}}dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}e^{-\frac{t^2}{2}}dx \\
&= e^{-\frac{t^2}{2}}
\end{aligned}$$

Daraus berechnet man die charakteristische Funktion für $\tilde{X} \sim \mathbf{N}(\mu,\sigma),\ \tilde{X} = \sigma X + \mu$:

$$\begin{aligned}
\varphi_{\tilde{X}}(t) &= e^{it\mu}\varphi_X(\sigma t) \\
&= e^{it\mu}e^{-\frac{\sigma^2t^2}{2}} \\
&= e^{it\mu-\frac{\sigma^2t^2}{2}} \qquad \text{(reellwertig für } \mu\text{)}
\end{aligned}$$

Die charakteristische Funktion wird zudem zur Berechnung von Momenten genutzt:

$$\begin{aligned}
\varphi_X(t) &= \int_{-\infty}^{\infty} e^{itx}f(x)dx \\
\varphi_X'(t) &= \int_{-\infty}^{\infty} ixe^{itx}f(x)dx \\
\varphi_X'(0) &= i\int_{-\infty}^{\infty} xf(x)dx = i\mathbf{E}X \qquad \Rightarrow m_1 = \mathbf{E}X = \frac{\varphi_X'(0)}{i}
\end{aligned}$$

Analog folgt:

$$\left[m_k = EX^k = \frac{\varphi_X^{(k)}(0)}{i^k}\right] \qquad (k = 1, 2, \ldots)$$

Bemerkung.

$$\mathbf{D}^2X = m_2 - m_1^2 = \frac{\varphi_X''(0)}{i^2} - \left(\varphi_X'(0)\right)^2$$
$$\left[\mathbf{D}^2X = -\varphi_X''(0) + (\varphi_X'(0))^2\right]$$

Beispiel. Sei $X \sim \mathbf{N}(\mu, \sigma)$ eine normalverteilte Zufallsgröße. Die zugehörige charakteristische Funktion ist

$$\varphi_X(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$$

und die erste Ableitung ist

$$\varphi_X'(t) = (i\mu - \sigma^2 t)e^{it\mu - \frac{\sigma^2 t^2}{2}}.$$

Mit obiger Formel berechnet sich der Erwartungswert durch

$$\mathbf{E}X = \frac{\varphi_X'(0)}{i} = \frac{i\mu}{i} = \mu$$

sowieso analog die Varianz als $\mathbf{D}^2X = \sigma^2$.

Die charakteristische Funktion ist auch interessant für Summen von Zufallsgrößen:

Satz 1.3.34. *Seien X und Y stochastisch unabhängige Zufallsgrößen mit den charakteristischen Funktionen $\varphi_X(t)$ und $\varphi_Y(t)$. Dann gilt für die charakteristische Funktion der Zufallsgröße $Z = X + Y$*

$$\varphi_Z(t) = \varphi_X(t)\varphi_Y(t).$$

Beispiel. Seien $X \sim \mathbf{N}(\mu_X, \sigma_X^2)$ und $Y \sim \mathbf{N}(\mu_Y, \sigma_Y^2)$ normalverteilte Zufallsgrößen mit den charakteristischen Funktionen $\varphi_X(t) = e^{i\mu_X t - \frac{\sigma_X^2 t^2}{2}}$ und $\varphi_Y(t) = e^{i\mu_Y t - \frac{\sigma_Y^2 t^2}{2}}$. Dann folgt

$$\begin{aligned} \varphi_Z(t) &= e^{i(\mu_X + \mu_Y)t - \frac{(\sigma_X^2 + \sigma_Y^2)t^2}{2}} \\ \Rightarrow Z &\sim \mathbf{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2). \end{aligned}$$

Satz 1.3.35. *Existieren alle Momente, so gilt*

$$\varphi_X(t) = 1 + \sum_{k=1}^{\infty} \frac{m_k}{k!}(it)^k = \sum_{k=0}^{\infty} \frac{\varphi_X^{(k)}(0)}{k!} t^k,$$

falls die charakteristische Funktion in $t_0 = 0$ in eine Potenzreihe entwickelt werden kann.

Bemerkung. Die charakteristische Funktion ist die Fourriertransformierte der Dichtefunktion. Die Rücktransformation ist möglich:

$$f_X(x) \quad = \quad \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt.$$

1.4 Das Gesetz der großen Zahlen und Grenzverteilungssätze

In vielen Anwendungen, vor allem in der mathematischen Statistik, treten Folgen von Zufallsgrößen $X_1, X_2, \dots, X_n$ und deren Linearkombinationen

$$Y_n := a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

auf. Dabei gilt nach Hilfssatz 1.3.14

$$\mathbf{E} Y_n = \sum_{i=1}^{n} a_i \mathbf{E} X_i$$

und, falls $X_1, \dots, X_n$ vollständig unabhängig sind,

$$\mathbf{D}^2 Y_n = \sum_{i=1}^{n} a_i^2 \mathbf{D}^2 X_i.$$

Definition 1.4.1. Die Zufallsgrößen $X_1, \dots, X_n$ heißen *unabhängig und identisch verteilt* oder *vom Typ i.i.d.* (von „independent and identically distributed“), wenn sie vollständig unabhängig sind, identische Verteilungen aufweisen und die Erwartungswerte und Streuungen existieren. Es gilt also

$$\mathbf{E} X_1 = \cdots = \mathbf{E} X_n =: \mu \in \mathbb{R}, \qquad \mathbf{D}^2 X_1 = \cdots = \mathbf{D}^2 X_n =: \sigma^2 < \infty.$$

Sind $X_1, \dots, X_n$ Zufallsgrößen vom Typ i.i.d. und ist

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

ihr arithmetisches Mittel, so gilt

$$\mathbf{E} \bar{X}_n = \sum_{i=1}^{n} \frac{1}{n} \mu = \mu, \qquad \mathbf{D}^2 \bar{X}_n = \sum_{i=1}^{n} \frac{1}{n^2} \sigma^2 = \frac{\sigma^2}{n}. \tag{1.4.1}$$

1.4.1 Das Gesetz der großen Zahlen

Satz 1.4.2 (schwaches Gesetz der großen Zahlen)**.** *Sind $X_1, \dots, X_n$ Zufallsgrößen vom Typ i.i.d. und ist $\mu = \mathbf{E} \bar{X}_n = \mathbf{E} X_i$ deren einheitlicher Erwartungswert, so gilt für alle $\varepsilon > 0$*

$$\lim_{n \to \infty} P\left(|\bar{X}_n - \mu| \leq \varepsilon\right) = 1,$$

d.h. das arithmetische Mittel $\bar{X}_n$ konvergiert für wachsendes n im Sinne der Wahrscheinlichkeit gegen den einheitlichen Erwartungswert der Zufallsgrößen $X_1, \ldots, X_n$.

Beweis. Mit $\sigma^2 = \mathbf{D}^2 X_1 = \cdots = \mathbf{D}^2 X_n$ erhalten wir aus der TSCHEBYSCHEFF'schen Ungleichung und den Formeln (1.4.1):

$$P\left(|\bar{X}_n - \mu| \leq \varepsilon\right) = 1 - P\left(|\bar{X}_n - \mathbf{E}\bar{X}_n| > \varepsilon\right) > 1 - \frac{\mathbf{D}^2 \bar{X}_n}{\varepsilon^2} = 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

Für n gegen unendlich ergibt sich also:

$$1 \geq \lim_{n\to\infty} P\left(|\bar{X}_n - \mu| \leq \varepsilon\right) \geq \lim_{n\to\infty} \left(1 - \frac{\sigma^2}{n\varepsilon^2}\right) = 1. \qquad \square$$

Am Ende von Abschnitt 1.3.2.1 haben wir die Aussage des obigen Satzes bereits verwendet, um zu zeigen, dass die relative Häufigkeit $H_n(A) = \bar{X}_n$ eines Ereignisses A für wachsendes n (Versuchsanzahl) gegen die Wahrscheinlichkeit $p = P(A)$ strebt. Dabei waren $X_1, \ldots, X_n$ mit

$$X_i = \begin{cases} 1, & \text{wenn } A \text{ im } i\text{-ten Versuch eintritt} \\ 0, & \text{wenn } A \text{ im } i\text{-ten Versuch nicht eintritt} \end{cases}$$

Zufallsgrößen vom Typ i.i.d. und es galt $\mu = \mathbf{E}\bar{X}_n = p$ und $\sigma^2 = D^2\bar{X}_n = \frac{p(1-p)}{n}$. Also ist nach obigem Satz

$$\lim_{n\to\infty} P(|H_n(A) - p| \leq \varepsilon) = 1.$$

Beispiel. Uns interessiert, wie viele Versuche zu einer Zufallssituation durchgeführt werden müssen, damit mit einer Wahrscheinlichkeit von mindestens 95 % die relative Häufigkeit $H_n(A)$ und die Wahrscheinlichkeit $p = P(A)$ eines Ereignisses A bis zwei Stellen nach dem Komma übereinstimmen. Wir setzen also $\varepsilon = 0{,}005$ und erhalten mit $p(1-p) = -(p - \frac{1}{2})^2 + \frac{1}{4} \leq \frac{1}{4}$ analog zum obigen Beweis:

$$P(|H_n(A) - p| \leq \varepsilon) > 1 - \frac{\mathbf{D}^2(H_n(A))}{\varepsilon^2} = 1 - \frac{p(1-p)}{n\varepsilon^2} \geq 1 - \frac{1}{4n\varepsilon^2} = 1 - \frac{10000}{n}.$$

Wenn für ein n die Gleichung $1 - \frac{10000}{n} = 0{,}95$ erfüllt ist, so können wir also sicher sein, dass $P(|H_n(A) - p| \leq \varepsilon) \geq 0{,}95$ gilt. Wir erhalten somit als Lösung $n = 200000$ (oder größer). Wie wir weiter unten sehen werden, ist diese Abschätzung sehr grob.

1.4.2 Grenzverteilungssätze

1.4.2.1 Zentraler Grenzverteilungssatz

Das schwache Gesetz der großen Zahlen liefert nur eine Aussage über den stochastischen Grenzwert des arithmetischen Mittels $\bar{X}_n$ einer Folge von Zufallsgrößen $X_1, \ldots, X_n$. In vielen Anwendungen wird aber auch die Grenzverteilung des standardisierten arithmetischen Mittels

$$\bar{Y}_n = \frac{\bar{X}_n - \mathbf{E}\bar{X}_n}{\sqrt{\mathbf{D}^2 \bar{X}_n}} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n}.$$

benötigt.

Satz 1.4.3 (Zentraler Grenzverteilungssatz von LINDEBERG/LEVY). *Sei $\bar{X}_n$ das arithmetische Mittel einer Folge $X_1, \dots, X_n$ von Zufallsgrößen vom Typ i.i.d., $\mu = \mathbf{E}X_i \in \mathbb{R}$ und $0 < \sigma^2 = \mathbf{D}^2 X < \infty$. Weiter sei $F_n(x)$ die Verteilungsfunktion der Zufallsgröße $\bar{Y}_n = \frac{\bar{X}_n - \mu}{\sigma}\sqrt{n}$, d.h.*

$$F_n(x) = P\left(\frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} < x\right) = P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sqrt{n\sigma^2}}\right)$$

für alle $x \in \mathbb{R}$. Dann gilt für alle $x \in \mathbb{R}$

$$\lim_{n\to\infty} F_n(x) = \Phi(x),$$

wobei Φ die Verteilungsfunktion der standardisierten Normalverteilung bezeichnet. Die Standardisierung $\bar{Y}_n$ von $\bar{X}_n$ ist also asymptotisch $\mathbf{N}(0,1)$-verteilt (Schreibweise: $\bar{X}_n \approx \mathbf{N}(0,1)$).

Anwendung Das arithmetische Mittel einer Folge von Zufallsgrößen kann also in guter Näherung als normalverteilt angenommen werden. Somit motiviert der zentrale Grenzverteilungssatz die Annahme, dass eine durch Überlagerung zahlreicher unabhängiger Einzeleinflüsse entstehende Zufallsgröße (z.B. Messfehler) als normalverteilt aufgefasst werden kann.

1.4.2.2 Grenzverteilungssatz von MOIVRE/LAPLACE

Wir betrachten nun einen Spezialfall des zentralen Grenzverteilungssatzes. Zu einer Zufallssituation werden n Versuche durchgeführt, wobei die Zufallsgrößen X_i angeben, ob das Ereignis A im i-ten Versuch eingetreten ist ($X_i = 1$) oder nicht ($X_i = 0$). Wir setzen $p = P(A) = P(X_i = 1)$ und $Y_n = X_1 + \cdots + X_n$. Die Zufallsgröße Y_n gibt also an, wie oft das Ereignis A bei n Versuchen eingetreten ist. Es gilt $Y_n \sim \mathbf{B}(n,p)$ und somit $\mathbf{E}Y_n = np$ und $\mathbf{D}^2 Y_n = np(1-p)$. Für genügend große n erhalten wir mit $\bar{X}_n = \frac{X_1 + \cdots + X_n}{n} = \frac{Y_n}{n}$ aus dem zentralen Grenzverteilungssatz die Beziehung

$$\mathbf{N}(0,1) \approx \frac{\bar{X}_n - \mathbf{E}\bar{X}_n}{\sqrt{\mathbf{D}^2 \bar{X}_n}}\sqrt{n} = \frac{\bar{X}_n - p}{\sqrt{p(1-p)}}\sqrt{n} = \frac{Y_n - np}{\sqrt{np(1-p)}} = \frac{Y_n - \mathbf{E}Y_n}{\sqrt{D^2 Y_n}},$$

d.h. $Y_n \approx \mathbf{N}(np, np(1-p))$. Damit haben wir die Aussage des Grenzverteilungssatzes von MOIVRE/LAPLACE hergeleitet.

Satz 1.4.4 (Grenzverteilungssatz von MOIVRE/LAPLACE). *Sei $X \sim \mathbf{B}(n,p)$ eine binomialverteilte Zufallsgröße und $F_n(x)$ die Verteilungsfunktion der standardisierten Zufallsgröße $Y = \frac{X - np}{\sqrt{np(1-p)}}$. Dann gilt für alle $x \in \mathbb{R}$*

$$\lim_{n\to\infty} F_n(x) = \Phi(x),$$

wobei Φ die Verteilungsfunktion der standardisierten Normalverteilung bezeichnet. Die binomialverteilte Zufallsgröße X ist also asymptotisch $N(np, np(1-p))$ verteilt.

Faustregel Der Grenzverteilungssatz von MOIVRE/LAPLACE ist in guter Näherung anwendbar, wenn $np(1-p) > 9$ gilt. Selbst für $np(1-p) > 4$ erhält man noch eine brauchbare Näherung.

Anwendung Die Berechnung von Wahrscheinlichkeiten ist bei binomialverteilten Zufallsgrößen extrem rechenaufwändig. Mit Hilfe des obigen Satzes kann man diese aufwändigen Rechnungen auf die einfacher handhabbare Normalverteilung zurückführen.

Beispiel. Nun haben wir eine weitere Möglichkeit zur Lösung des Problems aus dem vorhergehenden Beispiel. Gesucht war die Versuchsanzahl n, die benötigt wird, damit mit einer Wahrscheinlichkeit von mindestens 95 % die relative Häufigkeit $H_n(A)$ und die Wahrscheinlichkeit $p = P(A)$ eines Ereignisses A bis zwei Stellen nach dem Komma übereinstimmen. Wir suchen also ein n, so dass $P(|H_n(A) - p| \leq \varepsilon) \geq 0{,}95$ gilt, wobei $\varepsilon = 0{,}005$ ist. Unter Verwendung des Grenzverteilungssatzes von Moivre/Laplace und den Bezeichnungen aus dessen Herleitung erhalten wir zunächst:

$$\begin{aligned}
P(|H_n(A) - p| \leq \varepsilon) &= P(|\bar{X}_n - p| \leq \varepsilon) = P(|Y_n - np| \leq n\varepsilon) \\
&= P\left(\left|\frac{Y_n - np}{\sqrt{np(1-p)}}\right| \leq \frac{n\varepsilon}{\sqrt{np(1-p)}}\right) \\
&= P\left(\frac{-n\varepsilon}{\sqrt{np(1-p)}} \leq \frac{Y_n - np}{\sqrt{np(1-p)}} \leq \frac{n\varepsilon}{\sqrt{np(1-p)}}\right) \\
&= \Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) - \Phi\left(\frac{-\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) = 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) - 1.
\end{aligned}$$

Es muss also $2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) - 1 \geq 0{,}95$ gelten, d.h. $\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}}\right) \geq 0{,}975$. Dies ist äquivalent zu $\frac{\sqrt{n}\varepsilon}{\sqrt{p(1-p)}} \geq 1{,}96$. Da $p(1-p) = -(p - \frac{1}{2})^2 + \frac{1}{4} \leq \frac{1}{4}$, ist letztere Beziehung erfüllt, wenn $2\sqrt{n}\varepsilon \geq 1{,}96$ gilt. Als Ergebnis erhalten wir also $n = 38416$ (oder größer). Diese Abschätzung ist deutlich besser als die vorhergehende, die wir mit Hilfe der Tschebyscheff'schen Ungleichung erhalten hatten.

Methode der Stetigkeitskorrektur Mit der hier vorgestellten Methode erhält man bessere numerische Ergebnisse bei Verwendung des Grenzverteilungssatzes von Moivre/Laplace zur näherungsweisen Berechnung von Wahrscheinlichkeiten der Form $P(a \leq Y_n \leq b)$ einer binomialverteilten Zufallsgröße $Y_n \sim \mathbf{B}(n,p)$, wobei a und b positive ganze Zahlen sind. Es gilt also $Y_n \approx \mathbf{N}(np, np(1-p))$. Die Idee der Methode der Stetigkeitskorrektur ist, die Grenze a um $\frac{1}{2}$ zu verringern und b um $\frac{1}{2}$ zu erhöhen. Wir erhalten also:

$$P(a \leq Y_n \leq b) \approx P(a - \tfrac{1}{2} \leq Y_n \leq b + \tfrac{1}{2}) = \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - \frac{1}{2} - np}{\sqrt{np(1-p)}}\right).$$

Beispiel. Wir wollen nun die Verbesserung der Ergebnisse durch die Methode der Stetigkeitskorrektur an konkreten Zahlen demonstrieren. Sei dazu $Y_n \sim \mathbf{B}(100; 0{,}25)$, $a = 15$ und $b = 30$. Da $np(1-p) = 18{,}75 > 9$ gilt, ist der Grenzverteilungssatz von Moivre/Laplace in guter Näherung anwendbar. Es gilt also $Y_n \approx \mathbf{N}(25; 18{,}75)$. Als exaktes, aber sehr rechenaufwändiges Ergebnis erhalten wir:

$$P(15 \leq Y_n \leq 30) = \sum_{k=15}^{30} 0{,}25^k 0{,}75^{100-k} = 0{,}8908.$$

Ohne Stetigkeitskorrektur liefert die standardisierte Normalverteilung:

$$P(15 \leq Y_n \leq 30) \approx \Phi\left(\frac{15-25}{\sqrt{18{,}75}}\right) - \Phi\left(\frac{30-25}{\sqrt{18{,}75}}\right) = 0{,}8645.$$

Unter Verwendung der Methode der Stetigkeitskorrektur erhalten wir:

$$P(14{,}5 \leq Y_n \leq 30{,}5) \approx \Phi\left(\frac{14{,}5-25}{\sqrt{18{,}75}}\right) - \Phi\left(\frac{30{,}5-25}{\sqrt{18{,}75}}\right) = 0{,}8903.$$

Wir sehen also deutlich, dass durch die Methode der Stetigkeitskorrektur eine bessere Näherung erreicht wird.

1.4.2.3 Grenzverteilungssatz von Poisson

Wir haben bereits zur Herleitung der Wahrscheinlichkeitsfunktion der Poisson-Verteilung einen Zusammenhang zwischen Binomial- und Poisson-Verteilung hergestellt. Allgemeiner erhalten wir den folgenden Satz.

Satz 1.4.5 (Grenzverteilungssatz von Poisson)**.** *Gegeben sei eine Folge von binomialverteilten Zufallsgrößen $Y_n \sim \mathbf{B}(n, p_n)$ mit $p_n \to 0$ und $np_n \to \lambda > 0$ für $n \to \infty$. Dann gilt*

$$\lim_{n\to\infty} P(Y_n = k) = \pi_\lambda(k),$$

d.h. die Zufallsgrößen Y_n sind asymptotisch Poisson*-verteilt mit dem Parameter λ.*

Beweis. Es gilt

$$\begin{aligned}
\lim_{n\to\infty} P(Y_n = k) &= \lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} \\
&= \lim_{n\to\infty} \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\to 1} \underbrace{\frac{(np_n)^k}{k!}}_{\to \frac{\lambda^k}{k!}} \underbrace{\left(1-\frac{np_n}{n}\right)^n}_{e^{-\lambda}} \underbrace{\frac{1}{(1-p_n)^k}}_{\to 1} \\
&= \frac{\lambda^k}{k!} e^{-\lambda} = \pi_\lambda(k).
\end{aligned}$$

□

Faustregel Den Grenzverteilungssatz von Poisson kann man ohne Bedenken anwenden, wenn die Ungleichungen $np \leq 10$ und $1500p \leq n$ erfüllt sind.

Anwendung Wie der Grenzverteilungssatz von Moivre/Laplace, so verringert auch der Grenzverteilungssatz von Poisson den Rechenaufwand bei einer binomialverteilten Zufallsgröße erheblich. Sollte der Parameter p der Binomialverteilung zu klein sein, um mit dem Grenzverteilungssatz von Moivre/Laplace eine ausreichend gute Näherung zu erzielen, so kann der Grenzverteilungssatz von Poisson genutzt werden.

Beispiel. Für eine binomialverteilte Zufallsgröße $X \sim \mathbf{B}(100; 0{,}01)$ soll die Wahrscheinlichkeit $P(2 \leq X \leq 10)$ berechnet werden. Wegen $np(1-p) = 0{,}99 < 4$ sollte der Grenzverteilungssatz von Moivre/Laplace nicht genutzt werden. Jedoch liefert der Grenzverteilungssatz von

POISSON eine gute Näherung, da $np = 1 \leq 10$ und $1500p = 15 \leq 100 = n$. Wir erhalten somit als Näherung

$$P(2 \leq X \leq 10) = \sum_{k=2}^{10} \pi_1(k) = 0{,}264241.$$

Exakte Rechnung ergibt

$$P(2 \leq X \leq 10) = \sum_{k=2}^{10} \binom{100}{k} 0{,}01^k 0{,}99^{100-k} = 0{,}264238.$$

1.4.2.4 Bemerkung zur hypergeometrischen Verteilung

Es lassen sich auch Grenzverteilungssätze für eine hypergeometrisch verteilte Zufallsgröße $X \sim \mathbf{H}(n, N, M)$ formulieren. Wenn z.B. $N \to \infty$, $M \to \infty$ und $\frac{M}{N} \to p$ für $n \to \infty$ gilt, so nutzt man die Näherung

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \approx \binom{n}{k} p^k (1-p)^{n-k},$$

d.h man ersetzt die hypergeometrische Verteilung näherungsweise durch eine entsprechende Binomialverteilung. Es gibt jedoch keine handhabbaren Faustregeln für die Nutzung dieser Approximation. Um die so erhaltene Binomialverteilung auszuwerten, kann man dann den Grenzverteilungssatz von MOIVRE/LAPLACE oder den Grenzverteilungssatz von POISSON verwenden.

1.5 Mehrdimensionale Verteilungen

Bisher wurden ausschließlich reellwertige, d.h. eindimensionale, Zufallsgrößen betrachtet. Wir haben also stets nur ein Merkmal des zu beobachtenden Objekts bei unseren Untersuchungen berücksichtigt. Häufig sind aber bei praktischen Modellierungsproblemen mehrere Merkmale der Beobachtungsobjekte gleichermaßen von Interesse. Somit benötigen wir mehrdimensionale Zufallsgrößen.

1.5.1 Einführung

Definition 1.5.1. Sei $\underline{X} = (X_1, X_2, \ldots, X_n)$ eine Zusammenfassung von n Zufallsgrößen $X_1, X_2, \ldots, X_n$. Dann heißt das Objekt $\underline{X}$ *n-dimensionale Zufallsgröße* oder *zufälliger Vektor*.

Definition 1.5.2. Die durch

$$F(x_1, \ldots, x_n) := P(X_1 < x_1, \ldots, X_n < x_n)$$

für alle Vektoren $(x_1, \ldots, x_n) \in \mathbb{R}^n$ definierte reelle Funktion F heißt *Verteilungsfunktion des zufälligen Vektors* $\underline{X} = (X_1, \ldots, X_n)$.

Wir werden im Folgenden nicht immer den allgemeinen Fall $\underline{X} = (X_1, \ldots, X_n)$ betrachten, sondern uns auf $n = 2$, d.h. $\underline{X} = (X, Y)$, beschränken, wenn dies das Verständnis erleichtert. Ein Großteil der Aussagen gilt dann analog für allgemeines n.