Statistik und empirische Methoden
Wahrscheinlichkeitsrechnung

Silke Bott

Sommersemester 2023

stetige Zufallsvariablen

Eine Variable oder ein Merkmal heißt *stetig*, falls zu zwei Werten $a < b$ auch jeder Zwischenwert im Intervall $[a, b]$ möglich ist. Falls die Werte von X als Ergebnis eines Zufallsvorgangs resultieren, wird X zu einer stetigen Zufallsvariablen. Wie lassen sich nun Wahrscheinlichkeiten für Ereignisse der Form $\{a \leq X \leq b\}$ festlegen? Für stetige Zufallsvariablen sind die x-Werte in $[a, b]$ nicht mehr abzählbar, sondern überabzählbar, so dass ein Aufsummieren nicht möglich ist.

Beispiel

Beim zufälligen Ziehen einer reellen Zahl aus dem Intervall $[0, 10]$ ist die Zufallsvariable X: *gezogene Zahl* stetig mit Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{10} & \text{für } 0 \leq x \leq 10 \\ 0 & \text{sonst} \end{cases}$$

stetige Zufallsvariablen

Definition

Eine Zufallsvariable X heißt **stetig**, wenn es eine integrierbare Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ gibt mit $f(x) \geq 0$ für alle $x \in \mathbb{R}$, so dass für alle $a \leq b$ gilt

$$p(X \in [a, b]) = \int_a^b f(x)\, dx$$

Die Funktion $f(x)$ heißt **Dichte** von X.

Bemerkung

Genau dann ist X eine stetige Zufallsvariable mit Dichte $f(x)$, wenn

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

stetige Zufallsvariablen

Eine typische Eigenschaft stetiger Zufallsvariablen ist, dass einzelne Werte immer mit Wahrscheinlichkeit 0 auftreten und erst bei Intervallen mit positiver Wahrscheinlichkeit zu rechnen ist.

Bemerkung

Ist X eine stetige Zufallsvariable, so gilt für jedes $a \in \mathbb{R}$:

$$p(X = a) = 0$$

Es ist nämlich $p(X = a) = p(X \in [a, a]) = \int_a^a f(x)dx = 0$

stetige Zufallsvariablen

Regel

Für die Verteilungsfunktion F und die Dichtefunktion f einer stetigen Zufallsvariable X gilt:

1. $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1$. *Insbesondere existiert dieses uneigentliche Integral.*
2. *F ist stetig und monoton wachsend mit Werten im Intervall* $[0,1]$
3. $\lim\limits_{x\to-\infty} F(x) = 0$ *und* $\lim\limits_{x\to\infty} F(x) = 1$.
4. *Ist x eine Stetigkeitsstelle von $f(x)$, so ist F differenzierbar in x mit $F'(x) = f(x)$.*
5. *Für $a, b \in \mathbb{R}$ mit $a \leq b$ gilt*

$$p(a \leq X \leq b) = F(b) - F(a), \qquad p(X \geq a) = 1 - F(a)$$

stetige Zufallsvariablen

Beispiel

Die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} \frac{1}{2} \cdot x & \text{für } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

könnte die Dichtefunktion einer Zufallsvariable X sein, denn hierfür gilt

1. $f(x)$ ist integrierbar.
2. $f(x) \geq 0$ für alle $x \in \mathbb{R}$.
3. $\int\limits_{-\infty}^{\infty} f(x)\,dx = \int\limits_{0}^{2} \frac{x}{2}\,dx = \left[\frac{x^2}{4}\right]_0^2 = 1.$

stetige Zufallsvariablen

Bemerkung

Ist $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine Funktion mit

1. $f(x)$ ist integrierbar.
2. $f(x) \geq 0$ für alle $x \in \mathbb{R}$.
3. $\int\limits_{-\infty}^{\infty} f(x)\,dx = 1.$

so gibt es eine stetige Zufallsvariable X, die $f(x)$ als Dichte hat.
Umgekehrt muss auch jede Dichtefunktion einer stetige Zufallsvariable diese Bedingungen erfüllen.

stetige Zufallsvariablen

Übung

Überprüfen Sie, ob $f : \mathbb{R} \longrightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} \frac{3}{2} \cdot x - \frac{3}{4} \cdot x^2 & \text{für } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

die Dichtefunktion einer stetigen Zufallsvariable X sein kann.

Lösung:

Die Funktion $f(x)$ erfüllt alle Bedingungen.

stetige Zufallsvariablen

Definition

Der **Erwartungswert** einer stetigen Zufallsvariable X ist

$$E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

falls dieses uneigentliche Integral existiert.
Die **Varianz** einer stetigen Zufallsvariable X mit Erwartungswert μ ist

$$\mathrm{Var}(X) = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\, dx$$

(falls dieses uneigentliche Integral existiert), und die **Standardabweichung** einer stetigen Zufallsvariabel X ist

$$\sigma = +\sqrt{\mathrm{Var}(X)}$$

stetige Zufallsvariablen

Regel

Für stetige Zufallsvariablen gilt:

- *Ist* $Y = a \cdot X + b$, *so ist*

$$E(Y) = a \cdot E(X) + b$$

- *Ist* $Z = X + Y$, *so ist*

$$E(Z) = E(X) + E(Y)$$

stetige Zufallsvariablen

Regel

Für stetige Zufallsvariablen gilt:

- *Ist $\mu = E(X)$, so gilt*

$$\mathrm{Var}(X) = E((X-\mu)^2) = E(X^2) - \mu^2$$

- *Ist $Y = a \cdot X + b$, so gilt*

$$\mathrm{Var}(Y) = a^2 \cdot \mathrm{Var}(X)$$

- *Sind X und Y unabhängige Zufallsvariablen, so gilt*

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

stetige Zufallsvariablen

Übung

Bestimmen Sie Erwartungswert und Varianz der stetigen Zufallsvariable X mit Dichte

$$f(x) = \begin{cases} \frac{1}{10} & \text{für } 0 \leq x \leq 10 \\ 0 & \text{sonst} \end{cases}$$

Lösung:

Es gilt

$$\begin{aligned} E(X) &= 5 \\ \mathrm{Var}(X) &= \frac{25}{3} \end{aligned}$$

stetige Zufallsvariablen

Beispiel

Eine stetige Zufallsvariable X heißt **gleichverteilt** auf dem Intervall $[a, b]$, wenn sie eine Dichte

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

mit

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

besitzt. Eine auf dem Intervall $[0, 1]$ gleichverteilte Zufallsvariable heißt auch **standardgleichverteilt**.

Ist X gleichverteilt auf dem Intervall $[a, b]$, so gilt

$$\begin{aligned} E(X) &= \frac{a+b}{2} \\ \mathrm{Var}(X) &= \frac{(b-a)^2}{12} \end{aligned}$$

stetige Zufallsvariablen

Beispiel

Eine stetige Zufallsvariable X heißt **exponentialverteilt** mit Parameter $\lambda > 0$, wenn sie die Dichte

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{für} \quad x \geq 0 \\ 0 & \text{für} \quad x < 0 \end{cases}$$

besitzt.

Ist X exponentialverteilt mit Parameter λ, so gilt

$$\begin{aligned} E(X) &= \lambda \cdot \int_0^\infty x \cdot e^{-\lambda \cdot x}\, dx &&= \frac{1}{\lambda} \\ \mathrm{Var}(X) &= \lambda \cdot \int_0^\infty \left(x - \frac{1}{\lambda}\right)^2 \cdot e^{-\lambda \cdot x}\, dx &&= \frac{1}{\lambda^2} \end{aligned}$$

stetige Zufallsvariablen

Analog zur *geometrischen Verteilung*: Wir betrachten die Wartezeit als stetige Zufallsvariable. Als Voraussetzung muss gelten, dass das System nicht altert. Also die Ausfallwahrscheinlichkeit ist unabhängig vom Alter gleich groß. Daraus ergibt sich die Gedächtnislosigkeit der Exponentialverteilung.
Die Exponentialverteilung steht auch in engem Zusammenhang mit der *Poisson-Verteilung*. Die Anzahl von Ereignissen in einem Zeitintervall ist genau dann $Po(\lambda)$-verteilt, wenn die Zeitdauern zwischen zwei aufeinanderfolgenden Ereignissen unabhängig und exponentialverteilt mit Parameter λ ist.

stetige Zufallsvariablen

Die Exponentialfunktion beschreibt häufig Lebensdauerprozesse von Objekten, bei denen die Wahrscheinlichkeit für das Ende ihrer Lebensdauer in einer gegebenen Zeitperiode fester Länge immer gleich groß ist, vorausgesetzt, sie erleben den Beginn einer Periode. Ein typisches Beispiel hierfür ist der radioaktive Zerfall. Die Wahrscheinlichkeit für den Zerfall eines Teilchens in der nächsten Sekunde ist immer gleich groß.

Beispiel

Caesium–137 ist ein radioaktives Material mit einer Halbwertszeit von 11 000 Tagen, dh. die Wahrscheinlichkeit dafür, dass ein Teilchen innerhalb der nächsten 11 000 Tage zerfällt ist exakt 0.50. Der Zerfallszeitpunkt eines zufällig ausgewählten Teilchens definiert also eine stetige, exponentialverteilte Zufallsvariable X mit Dichte

$$f(t) = \begin{cases} \lambda \cdot \mathrm{e}^{-\lambda \cdot t} & \text{für} \quad t \geq 0 \\ 0 & \text{für} \quad t < 0 \end{cases}$$

stetige Zufallsvariablen

Beispiel

Zur Bestimmung von λ:

$$\begin{aligned} 0.50 &= p(X \leq 11\,000) = F(11\,000) = \int\limits_0^{11000} \lambda \cdot \mathrm{e}^{-\lambda \cdot t}\, dt \\ &= \left[-e^{-\lambda \cdot x}\right]_0^{11000} = 1 - \mathrm{e}^{-11000\lambda} \end{aligned}$$

Daraus erhalten wir $\mathrm{e}^{-11000\lambda} = 0.50$ also (durch Logarithmieren) $-11\,000 \cdot \lambda = \ln(0.50) = -\ln(2)$ und damit

$$\lambda = \frac{\ln(2)}{11\,000}$$

(wobei λ in $\frac{1}{\text{Tage}}$ gemessen wird).

stetige Zufallsvariablen

Beispiel

Damit ist die Wahrscheinlichkeit dafür, dass ein gegebenes Teilchen innerhalb der nächsten 100 Tage zerfällt, genau

$$p = \int_0^{100} \frac{\ln(2)}{11\,000} \cdot \mathrm{e}^{-\frac{\ln(2)}{11\,000}\cdot t}\, dt = 1 - \mathrm{e}^{-\frac{\ln(2)}{11\,000}\cdot 100} = 0.0063$$

Da die Anzahl der radioaktiven Teilchen in einer Probe selbst in einer sehr kleinen) sehr hoch ist, kann man sehr verlässlich davon ausgehen, dass von einer Caesium–137–Probe tatsächlich innerhalb der nächsten 100 Tage exakt 0.63 % zerfallen.

stetige Zufallsvariablen

Übung

Die Ausfallwahrscheinlichkeit von Maschinen und Bauteilen (sofern sie immer gut gewartet und gepflegt werden) kann in guter Näherung durch eine Exponentialverteilung beschrieben werden.
Von einem Bauteil ist bekannt, dass es im Mittel nach 180 Tagen ausfällt. Bestimmen Sie die Dichtefunktion für die Zufallsvariable X: *Lebensdauer eines Bauteils* und die Wahrscheinlichkeit dafür, dass ein zufällig ausgewähltes Bauteil nach 200 Tagen noch funktioniert.

stetige Zufallsvariablen

Lösung:

Es gilt nach Angabe

$$\frac{1}{\lambda} = E(X) = 180$$

Also $\lambda = \frac{1}{180}$ (in $\frac{1}{\text{Tage}}$. Damit ist die Wahrscheinlichkeit dafür, dass ein zufällig ausgewähltes Teilchen in 200 Tagen noch funktioniert gegeben durch

$$p = 1 - p(X \leq 200) = 1 - \int_0^{200} \frac{1}{180} \cdot \mathrm{e}^{-\frac{1}{180} \cdot t} \, dt = \mathrm{e}^{-\frac{200}{180}} = 0.3292$$

Die Wahrscheinlichkeit dafür, dass ein Bauteil nach 180 Tagen noch funktioniert, ist dann übrigens $p = e^{-1} = 0.3679$ und nicht 0.50, wie man vielleicht erwarten würde.

stetige Zufallsvariablen

Die **Normalverteilung** ist die *bekannteste und wichtigste* Verteilung. Zwei wesentliche Gründe dafür sind: In vielen Anwendungen lässt sich die empirische Verteilung von Daten, die zu einem stetigen oder feinabstufigen diskreten (quasi-stetig), Merkmal X erhoben werden, durch eine Normalverteilung ausreichend gut approximieren. Die glockenförmige Gestalt dieser Dichte ist insbesondere dann ein gutes Modell für die Verteilung einer Variablen X, wenn diese durch das Zusammenwirken einer größeren Zahl von zufälligen Einflüssen entsteht, etwa bei physikalischen Größen wie Gewicht, Länge, Volumen, Punktezahlen in Tests usw.
Für die induktive Statistik noch entscheidender ist, dass sich viele andere Verteilungen, insbesondere solche, die sich bei Schätz- und Testprozeduren mit größerem Stichprobenumfang auftreten, durch die Normalverteilung gut approximieren lassen.

stetige Zufallsvariablen

Definition (Die Normalverteilung)

Eine stetige Zufallsvariable X heißt **normalverteilt** mit den Parametern μ und σ^2, wenn sie die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

bestitzt. Sie heißt **standardnormalverteilt**, wenn $\mu = 0$ und $\sigma^2 = 1$.

Bemerkung

Ist X normalverteilt mit Parametern μ und σ^2, so schreiben wir hierfür

$$X \sim N(\mu, \sigma^2)$$

stetige Zufallsvariablen

Bemerkung

Die Dichte der Standardnormalverteilung ist gegeben durch

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

Bemerkung

Ist $X \sim N(\mu, \sigma^2)$, so gilt

$$\begin{aligned} E(X) &= \mu \\ \mathrm{Var}(X) &= \sigma^2 \end{aligned}$$

stetige Zufallsvariablen

Bemerkung

Die Verteilungsfunktion von X ist gegeben durch

$$F(X) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

Die Verteilungsfunktion einer standardnormalverteilten Zufallsvariable ist gegeben durch

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

Diese Integrale lassen sich nicht mehr in einer geschlossenen Form schreiben. Die Werte von $\Phi(x)$ finden sich jedoch in Tabellen (in jedem Buch zur Statistik).

stetige Zufallsvariablen

Bemerkung

ist $X \sim N(\mu, \sigma^2)$, so gilt

- $p(\mu - \sigma \leq X \leq \mu + \sigma) = 0.6827$.
- $p(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9545$.
- $p(\mu - 0.675 \cdot \sigma \leq X \leq \mu + 0.675 \cdot \sigma) = 0.50$.
- $p(\mu - 1.645 \cdot \sigma \leq X \leq \mu + 1.645 \cdot \sigma) = 0.90$.

Bemerkung

Ist $X \sim N(\mu, \sigma^2)$, so hat die Dichtefunktion $f(x)$ von X Wendepunkte in $\mu - \sigma$ und $\mu + \sigma$ und das sind die beiden einzigen Wendepunkte von $f(x)$.

stetige Zufallsvariablen

Regel

Ist $X \sim N(\mu, \sigma^2)$, so ist die transformierte Zufallsvariable

$$Y = \frac{X - \mu}{\sigma}$$

standardnormalverteilt.
Ist $X \sim N(\mu, \sigma^2)$, so ist die transformierte Zufallsvariable

$$Y = a \cdot X + b$$

normalverteilt mit Parametern $a \cdot \mu + b$ und $a^2 \cdot \sigma^2$.
Sind $X_1 \sim N(\mu_1, \sigma_1^2)$ und $X_2 \sim N(\mu_2, \sigma_2^2)$ zwei unabhängige, normalverteilte Zufallsvariablen, so ist auch $X_1 + X_2$ normalverteilt und es gilt

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

stetige Zufallsvariablen

Beispiel

Die Größe X deutscher Männer ist (in guter Näherung) normalverteilt mit einem Mittelwert von 180.0 cm und einer Standardabweichung von 5.0 cm. Die Wahrscheinlichkeit dafür, dass ein zufällig ausgewählter deutscher Mann zwischen 1.74 m und 1.84 cm groß ist, berechnet sich wie folgt: Nach den Angaben ist $X \sim N(180, 25)$, also ist

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 180}{5} \sim N(1, 0)$$

Damit gilt

$$\begin{aligned} 174 \leq X \leq 184 \quad &\iff \quad \tfrac{174-180}{5} \leq Z \leq \tfrac{184-180}{5} \\ &\iff \quad -1.20 \leq Z \leq 0.80 \end{aligned}$$

stetige Zufallsvariablen

Beispiel

Damit ist

$$\begin{aligned} p(174 \leq X \leq 184) &= p(-1.20 \leq Z \leq 0.80) \\ &= \Phi(0.80) - \Phi(-1.20) \\ &= \Phi(0.80) - (1 - \Phi(1.20)) \\ &= \Phi(0.80) + \Phi(1.20) - 1 \\ &= 0.8849 + 0.7881 - 1 \\ &= 0.6730 \end{aligned}$$

stetige Zufallsvariablen

Beispiel

In welchem symmetrischen Intervall um den Mittelwert 180 cm liegt die Größe von 80 % der deutschen Männer?
Gesucht ist hier ein $a > 0$ so, dass $p(180 - a \leq X \leq 180 + a) = 0.80$. Wir führen das wieder auf die Standardnormalverteilung zurück indem wir wie in a) ausnutzen, dass

$$\begin{aligned} 180 - a \leq X \leq 180 + a \quad &\iff \quad \tfrac{180-a-180}{5} \leq Z \leq \tfrac{180+a-180}{5} \\ &\iff \quad -\tfrac{a}{5} \leq Z \leq \tfrac{a}{5} \end{aligned}$$

stetige Zufallsvariablen

Beispiel

Damit erhalten wir die Bedingung

$$\begin{aligned}
0.80 &= p(180 - a \le X \le 180 + a) \\
&= p\left(-\tfrac{a}{5} \le Z \le \tfrac{a}{5}\right) \\
&= \Phi\left(\tfrac{a}{5}\right) - \Phi\left(-\tfrac{a}{5}\right) \\
&= \Phi\left(\tfrac{a}{5}\right) - \left(1 - \Phi\left(\tfrac{a}{5}\right)\right) \\
&= 2 \cdot \Phi\left(\tfrac{a}{5}\right) - 1
\end{aligned}$$

Daraus erhalten wir

$$\Phi\left(\frac{a}{5}\right) = 0.90$$

und aus der Tabelle der Standardnormalverteilung lesen wir ab, dass $\frac{a}{5} = 1.28$, also $a = 6.40$. Das gesuchte Intervall ist also $[173.6, 186.4]$.

stetige Zufallsvariablen

Übung

Vom Durchmesser einer Frucht ist bekannt, dass er normalverteilt mit Mittelwert 14.00 cm ist, und dass 85.00 % aller Früchte einen Durchmesser von höchstens 15.50 cm haben.
Wie groß ist die Standardabweichung für den Durchmesser dieser Frucht?

Lösung:

Es gilt

$$\sigma = \frac{1.50}{1.035} = 1.45$$

stetige Zufallsvariablen

Definition

Eine Zufallsvariable Z heißt **Chi–Quadrat–verteilt** (χ^2–verteilt) mit n Freiheitsgraden, wenn es stochastisch unabhängige und standardnormalverteilte Zufallsvariablen $X_1, \ldots, X_n$ gibt mit

$$Z = X_1^2 + \cdots + X_n^2$$

Bemerkung

Ist Z eine χ^2–verteilte Zufallsvariable mit n Freiheitsgraden, so schreiben wir

$$Z \sim \chi^2(n)$$

Es gilt

$$\begin{aligned} E(Z) &= n \\ \mathrm{Var}(Z) &= 2n \end{aligned}$$

stetige Zufallsvariablen

Definition

Eine Zufallsvariable T heißt t–**verteilt** mit n Freiheitsgraden, wenn es eine standardnormalverteilte Zufallsvariable X und eine χ^2–verteilte Zufallsvariable Z mit n Freiheitsgraden gibt, die stochastisch unabhängig sind, mit

$$T = \frac{X}{\sqrt{\frac{1}{n} \cdot Z}}$$

Bemerkung

Ist T eine t–verteilte Zufallsvariable mit n Freiheitsgraden, so schreiben wir $T \sim t(n)$ und sprechen auch von eine **Student–Verteilung** mit n–Freiheitsgraden.

stetige Zufallsvariablen

Bemerkung

Ist $X \sim t(n)$, so gilt

$$E(X) = 0$$

Außerdem gilt für $n > 2$:

$$\mathrm{Var}(X) = \frac{n}{n-2}$$

Für $n = 1$ und $n = 2$ existiert die Varianz von $t(n)$ nicht.

Bemerkung

Für $n > 30$ ist die $t(n)$–Verteilung sehr ähnlich zur Standardnormalverteilung und kann durch die Standardnormalverteilung ersetzt werden.

stetige Zufallsvariablen

Definition

Eine Zufallsvariable F heißt **Fisher–verteilt** oder **F–verteilt** mit Freiheitsgraden m und n, wenn es stochastisch unabhängige χ^2–verteilte Zufallsvariablen $X \sim \chi^2(m)$ und $Y \sim \chi^2(n)$ gibt mit

$$F = \frac{\frac{1}{m} \cdot X}{\frac{1}{n} \cdot Y}$$

Bemerkung

Ist F eine Fisher–verteilte Zufallsvariable mit Freiheitsgraden m und n, so schreiben wir $F \sim F(m, n)$. Es gilt

$$\begin{aligned} E(F) &= \tfrac{n}{n-2} && \text{für} \quad n \geq 3 \\ \mathrm{Var}(F) &= \tfrac{2n^2 \cdot (n+m-2)}{m \cdot (n-4) \cdot (n-2)^2} && \text{für} \quad n \geq 5 \end{aligned}$$

stetige Zufallsvariablen

Übung

Sei X eine beliebige stetige Zufallsvariable mit Dichte $f(x)$ und Verteilungsfunktion $F(x)$. Sind die folgenden Aussagen richtig oder unter Umständen falsch?

- $f(x) \leq 1$ für alle x.
- $F(x) \leq 1$ für alle x.
- $\int_x^\infty f(t)dt = 1 - F(x)$.
- Ist $x_i < x_j$ so ist $F(x_i) \leq F(x_j)$.

stetige Zufallsvariablen

Lösung:

- $f(x) \leq 1$ ist falsch. Betrachte als Gegenbeispiel die stetige Gleichverteilung zwischen $a = 0$ und $b = 0.1$. Hier gilt

$$f(x) = \begin{cases} 10 & \text{für} \quad 0 \leq x \leq 0.1 \\ 0 & \text{sonst.} \end{cases}$$

- $F(x) \leq 1$ ist nach Definition richtig.
- $\int_x^\infty f(t)dt = 1 - F(x)$ ist richtig, denn

$$\int_x^\infty f(t)dt = P(X \geq x) = 1 - P(X \leq x) = 1 - F(x).$$

- $F(x_i) \leq F(x_j)$ ist richtig, da F monoton wachsend ist.

stetige Zufallsvariablen

Übung

Sei X eine zum Parameter λ exponentialverteilte Zufallsvariable. Zeigen Sie die *Gedächtnislosigkeit* der Exponentialverteilung, d.h. dass

$$P(X \leq x | X > s) = P(X \leq x - s)$$

mit $x, s \in \mathbb{R}$ mit $s < x$ gilt.

stetige Zufallsvariablen

Lösung:

Für $s < x$ gilt mit der Verteilungsfunktion der Exponentialverteilung

$$\begin{aligned} P(X \leq x | X > s) &= \frac{P(s<X\leq x)}{P(X>s)} \\ &= \frac{P(X\leq x)-P(X\leq s)}{P(X>s)} \\ &= \frac{1-e^{\lambda x}-1+e^{-\lambda s}}{1-1+e^{-\lambda(x-s)}} \\ &= 1-e^{-\lambda(x-s)} \\ &= P(X \leq x-s). \end{aligned}$$

stetige Zufallsvariablen

Übung

In einem Institut der DHBW Mannheim ist der einzige Farbkopierer ausgefallen. Über die Zeit X (in Stunden), die ein Techniker benötigt, um den Fotokopierer zu reparieren, ist bekannt, dass diese einer Exponentialverteilung mit Parameter $\lambda = 3$ folgt.
Berechnen Sie die Wahrscheinlichkeit, dass der Techniker

- höchstens eine Viertelstunde,
- zwischen 0.5 und 0.75 Stunden,
- mehr als 1 Stunde

für die Reparatur benötigt.

stetige Zufallsvariablen

Lösung:

Sei X die Zeit in Stunden, die benötigt wird, um den Fotokopierer zu reparieren mit

$$X \sim Ex(3)$$

. Dann ist

$$f(x) = \begin{cases} 3e^{-3x} & \text{für} \quad x \geq 0 \\ 0 & \text{sonst,} \end{cases}$$

und

$$F(x) = \begin{cases} 1 - e^{-3x} & \text{für} \quad x \geq 0 \\ 0 & \text{sonst.} \end{cases}$$

stetige Zufallsvariablen

Lösung:

Damit ergeben sich folgende Wahrscheinlichkeiten

- $P(X \leq 0.25) = F(0.25) = 1 - e^{-3 \cdot 0.25} = 1 - 0.4724 = 0.5276.$
- $$\begin{aligned} P(0.5 < X \leq 0.75) &= F(0.75) - F(0.5) \\ &= 1 - e^{-3 \cdot 0.75} - (1 - e^{-3 \cdot 0.5})) \\ &= e^{-1.5} - e^{-2.25} \\ &= 0.2231 - 0.1054 \\ &= 0.1177. \end{aligned}$$
- $P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - (1 - e^{-3 \cdot 1}) = 0.0498.$

stetige Zufallsvariablen

Übung

In einer Klinik wird eine Studie zum Gesundheitszustand von Frühgeburten durchgeführt. Das Geburtsgewicht X eines in der 28ten Schwangerschaftswoche geborenen Kinder wird als normalverteilte Zufallsvariable mit Erwartungswert 1000g und Standardabweichung 50g angenommen.

- Wie groß ist die Wahrscheinlichkeit, dass ein in der 28ten Schwangerschaftswoche geborenes Kind ein Gewicht zwischen 982 und 1050g hat?
- Geben Sie ein um den Erwartungswert symmetrisches Intervall an, in dem mit einer Wahrscheinlichkeit von 95% das Geburtsgewicht liegt.

stetige Zufallsvariablen

Lösung:

Den Angaben entnimmt man, dass für das Geburtsgewicht $X \sim N(1000, 50^2)$ gilt.

- Die gesuchte Wahrscheinlichkeit lässt sich nach Standarisierung über die Verteilungsfunktion der Standardnormalverteilung bestimmen als

$$
\begin{aligned}
P(982 \leq X \leq 1050) &= P(X \leq 1050) - P(X \leq 982) \\
&= P\left(\frac{X-1000}{50} \leq 1\right) - P\left(\frac{X-1000}{50} \leq -0.36\right) \\
&= \Phi(1) + \Phi(0.36) - 1 \\
&= 0.8413 + 0.6406 - 1 \\
&= 0.48190.
\end{aligned}
$$

stetige Zufallsvariablen

Lösung:

- Gesucht ist hier ein $a > 0$ so, dass $p(1000 - a \leq X \leq 1000 + a) = 0.95$.

$$
\begin{aligned}
0.95 &= p(1000 - a \leq X \leq 1000 + a) \\
&= p\left(\frac{1000-a-1000}{50} \leq Z \leq \frac{1000+a-1000}{50}\right) \\
&= p\left(-\frac{a}{50} \leq Z \leq \frac{a}{50}\right) \\
&= \Phi\left(\frac{a}{50}\right) - \Phi\left(-\frac{a}{50}\right) \\
&= \Phi\left(\frac{a}{50}\right) - \left(1 - \Phi\left(\frac{a}{50}\right)\right) \\
&= 2 \cdot \Phi\left(\frac{a}{50}\right) - 1
\end{aligned}
$$

Daraus erhalten wir

$$\Phi\left(\frac{a}{50}\right) = 0.975$$

und aus der Tabelle der Standardnormalverteilung lesen wir ab, dass $\frac{a}{50} = 1.96$, also $a = 98$. Das gesuchte Intervall ist also $[902, 1098]$.

stetige Zufallsvariablen

Übung

Ein genormter Leistungstest sei normalverteilt mit $\mu = 150$, $\sigma = 36$.

- Skizzieren Sie die Dichte dieser Verteilung.
- Zeichnen Sie jeweils die folgenden Wahrscheinlichkeiten als Fläche unter der Dichte ein, und berechnen Sie die Wahrscheinlichkeiten, Werte zu erreichen, die kleiner sind als 140, nicht im Bereich von 114 bis 190 liegen, größer sind als 175, kleiner als 200 und größer als 130 sind.
- Der Leistungstest wird nun an 49 Personen unabhängig voneinander durchgeführt. Wie Wahrscheinlich ist es, einen Mittelwert kleiner als 140 zu beobachten?
  Vergleichen Sie Ihr Ergebnis mit dem obigen. Wie erklären Sie sich den Unterschied?

stetige Zufallsvariablen

Lösung:

Sei $X=$ die Punktezahl des Leistungstests mit $X \sim N(150, 36^2)$.

- Dann gilt:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi\sigma}} \cdot exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) \\ &= \frac{1}{\sqrt{2\pi 36}} \cdot exp\left(-\frac{1}{2} \cdot \frac{(x-150)^2}{36^2}\right). \end{aligned}$$

Daraus folgt: $f(150) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{36} = 0.0111.$
Für $x = 150 + 36 = \mu + \sigma$ erhält man:

$$\begin{aligned} f(150+36) &= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{36} exp\left(-\frac{1}{2} \cdot \frac{(150+36-150)^2}{36^2}\right) \\ &= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{36} e^{-1/2} = 0.0067. \end{aligned}$$

stetige Zufallsvariablen

Lösung:

- $$\begin{aligned} P(X < 140) &= P\left(\frac{X-150}{36} < \frac{140-150}{36}\right) \\ &= P(Z < -0,28) \\ &= 1 - P(Z < 0,28) \\ &= 1 - \Phi(0,28) \\ &= 1 - 0.6103 = 0.3897. \end{aligned}$$

$$\begin{aligned} P(X < 114 \vee X > 190) &= P(X < 114) + P(X > 190) \\ &= \Phi\left(\frac{114-150}{36}\right) + 1 - \Phi\left(\frac{190-150}{36}\right) \\ &= \Phi(-1) + 1 - \Phi(1.11) \\ &= 1 - \Phi(1) + 1 - \Phi(1.11) \\ &= 2 - 0.8413 - 0.8665 = 0.2922. \end{aligned}$$

stetige Zufallsvariablen

Lösung:

- $$\begin{aligned} P(X > 175) &= 1 - P(X \leq 175) \\ &= 1 - \Phi\left(\frac{175-150}{36}\right) \\ &= 1 - \Phi(0.69) \\ &= 1 - \Phi(0.69) \\ &= 1 - 0.7549 \\ &= 0.2451. \end{aligned}$$

$$\begin{aligned} P(130 < X < 200) &= \Phi\left(\frac{200-150}{36}\right) - \Phi\left(\frac{130-150}{36}\right) \\ &= \Phi(1.39) - \Phi(-0.56) \\ &= \Phi(1.39) - 1 + \Phi(0.56) \\ &= 0.9177 - 1 + 0.7123 \\ &= 0.63. \end{aligned}$$

stetige Zufallsvariablen

Lösung:

- Der Leistungstest wird an $n = 49$ Personen unabhängig durchgeführt, d.h. $X_1, \dots, X_{49}$= Punktezahlen für jede der 49 Personen mit $X_i \sim N(150, 36^2)$. Betrachte $\overline{X} = \frac{1}{49}\sum_{i=1}^{49} X_i$. Es gilt: $\overline{X} \sim N(150, \frac{36^2}{49})$, woraus folgt:

$$\begin{aligned} P(\overline{X} < 140) &= \Phi\left(\frac{140-150}{36/7}\right) = \Phi(-1.94) \\ &= 1 - \Phi(1.94) = 1 - 0.9738 = 0.0262. \end{aligned}$$

  Die Wahrscheinlichkeit, dass $\overline{X}$ Werte kleiner als 140 annimmt, ist mit 0.0262 wesentlich kleiner als die 0.3897 für jedes einzelne X_i. Das liegt daran, dass $\overline{X}$ eine kleinere Varianz besitzt ($\frac{\sigma^2}{n}$) als $X_i(\sigma^2)$. Das bedeutet, dass sich die Wahrscheinlichkeitsmasse stärker um μ konzentriert und dadurch extreme Werte mit geringerer Wahrscheinlichkeit auftreten.