§ 9 Stetige Zufallsvariablen

Erinnerung

Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Eine Zufallsvariable ist eine Abbildung

$$X: \Omega \longrightarrow W,$$

wobei $W \subseteq \mathbb{R}$. der Wertebereich / Stichprobenraum ist, so dass: $A = \{\omega \in \Omega : X(\omega) \leq a\}$ für $a \in \mathbb{R}$ muss $A \in \mathcal{A}$ erfüllen.

Definition

Sei X eine Zufallsvariable. Dann nennen wir X stetig, falls $X(\Omega) \subseteq W$ einen kontinuierlichen / stetigen Bereich (= Intervall) enthält.

Definition (Dichte).

Sei X eine stetige Zufallsvariable und sei $f: \mathbb{R} \to [0, \infty)$ eine integrierbare Funktion, so dass:

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

Dann nennen wir f eine Dichte (-funktion) von X.

Bemerkung Nicht alle stetigen Zufallsvariablen haben Dichten (z.B.: die Cantor-Verteilung hat keine). Im Folgenden beschränken wir uns auf Zufallsvariablen mit Dichten.

Warum brauchen wir Dichtefunktionen?

Sei X eine stetige Zufallsvariable. Angenommen $P(X=a) > 0$ für alle $a \in W$. Dann gilt: $1 = \sum_{a \in W} P(X=a) = \infty$. Widerspruch!

Wir können die Annahme nicht machen!

Stetige Zufallsvariablen können nur Intervallen positive Wahrscheinlichkeiten zuordnen. Diese Zuordnung passiert mit Hilfe der Dichte durch

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

Bemerkung:
- Da $P(a \le X \le b) \ge 0$ für <u>alle</u> $a, b \in \mathbb{R}$ gilt: $f(x) \ge 0$ für alle $x \in \mathbb{R}$.
- $\int_{-\infty}^{\infty} f(x)\,dx = P(-\infty \le X \le \infty) = 1$.

Wichtige stetige Verteilungen (mit Dichten)

1. Gleichverteilung auf dem Intervall $[a,b]$

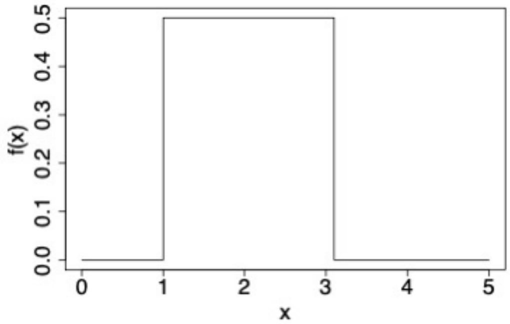
Eine Zufallsvariable X hat die Gleichverteilung auf $[a,b]$, wenn X die folgende Dichte hat:

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

Wir schreiben: $X \sim \text{Unif}([a,b])$

In R können wir $f(x)$ durch dunif(x, min=1, max=2) (für $a=1, b=2$).

Graph der Dichte von $X \sim Unif([1,3])$:



2. Normalverteilung

Sei X eine Zufallsvariable mit Dichte

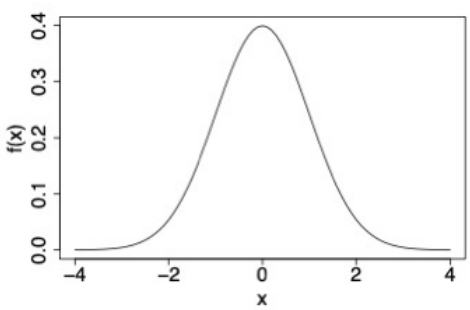
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

mit $\mu \in \mathbb{R}$, $\sigma^2 > 0$. Dann nennen wir X normalverteilt mit Parametern μ und σ^2. Wir schreiben $X \sim N(\mu, \sigma^2)$.

Falls $X \sim N(0,1)$, dann nennen wir X standardnormalverteilt.

In R: die Dichte der Normalverteilung wird mit dnorm(x, mean=μ, sd=σ). angegeben.

Graph der Dichte von $N(0,1)$:



3. Exponentialverteilung

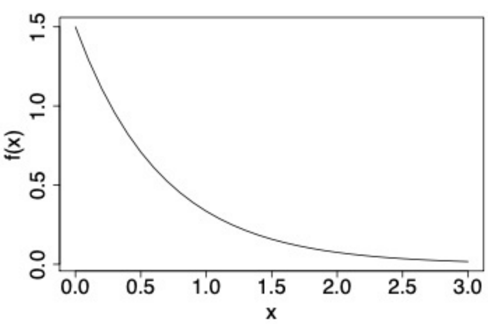
Sei X eine Zufallsvariable mit Dichte

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

wobei $\lambda > 0$ ein Parameter ist. Wir nennen X exponentialverteilt mit Parameter λ und schreiben $X \sim Exp(\lambda)$.

In R: Die Dichte wird durch dexp(x, rate = λ) angegeben.

Graph der Dichte von $Exp(1.5)$:



4. t-Verteilung

Seien $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ unabhängige Zufallsvariablen ("iid").

Wir definieren $\bar{X} := \frac{1}{n} \sum_{i=1}^{n} X_i$, $S^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

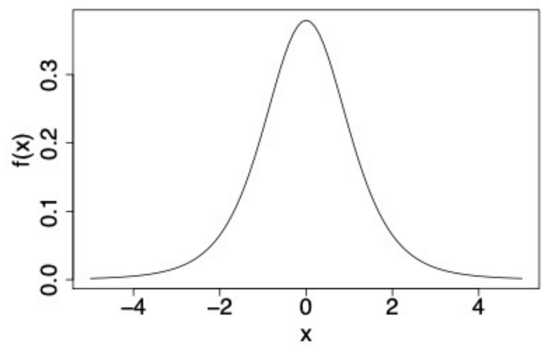
Die Zufallsvariable $Y := \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$ nennen wir t-verteilt. Wir schreiben $Y \sim t_{n-1}$.

Die Dichte von $Y \sim t_n$:

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\ \Gamma(\frac{n}{2})} \left(1 + \frac{x^2}{2}\right)^{-\frac{n+1}{2}}$$

wobei $\Gamma(z) := \int_0^\infty w^{z-1} e^{-w}\, dw$

Graph der Dichte von t_5:



<u>Bemerkung</u> für $n \to \infty$ nähert sich t_n an $N(0,1)$ an.

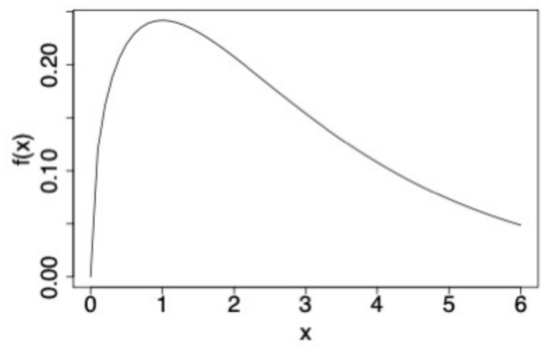
In R: Die Dichte wird durch $t(x, df=n)$ angegeben.

<u>5. Chi-Quadrat-Verteilung</u>

Seien $X_1, \ldots, X_n \sim N(0,1)$ unabhängige Zufallsvariable ("iid").
Die Zufallsvariable $Y := X_1^2 + \ldots + X_n^2$ hat die Chi-Quadrat-Verteilung mit n Freiheitsgraden. Wir schreiben $Y \sim \chi_n^2$.
In R: Die Dichte von $Y \sim \chi_n^2$ wird durch chisq $(x, df=n)$.
Die Dichte von $Y \sim \chi_n^2$ ist:

$$f(x) = \begin{cases} 0, & \text{falls } x < 0 \\ \dfrac{1}{\sqrt{2}^{\,n} \cdot \Gamma(\frac{n}{2})} \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}}, & \text{falls } x \geq 0 \end{cases}$$

Graph der Dichte von χ_3^2:



<u>Satz</u> (Summen unabhängiger normalverteilter Zufallsvariablen)

- Seien $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ unabhängige Zufallsvariablen ("iid").
  Dann gilt für $Y = X_1 + \dots + X_n$, dass $Y \sim N(n \cdot \mu, n \cdot \sigma^2)$.
- Seien $X_1, \dots, X_n$ unabhängige Zufallsvariablen mit $X_i \sim N(\mu_i, \sigma_i^2)$.
  Dann gilt für $Y = X_1 + \dots + X_n$, dass $Y \sim N(\mu_1 + \dots + \mu_n, \sigma_1^2 + \dots + \sigma_n^2)$

<u>Erinnerung</u>:

Sei X eine Zufallsvariable. Dann definieren wir

$$F(a) := P(X \leq a) = P(\{\omega \in \Omega \mid X(\omega) \leq a\})$$

$F(a)$ heißt (kumulative) Verteilungsfunktion von X.

Für eine stetige Zufallsvariable X mit Dichte f und Verteilungsfunktion F gilt also:

$$F(a) = \int_{-\infty}^{a} f(x)\, dx.$$

Definition (Erwartungswert, Varianz).

Sei X eine stetige Zufallsvariable mit Dichte $f(x)$. Dann ist

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

der Erwartungswert von X und

$$Var(X) := \mathbb{E}\,(X - \mathbb{E}(X))^2$$

ist die Varianz von X.

Satz

Sei $X \sim N(\mu, \sigma^2)$. Dann gilt: $\mathbb{E}(X) = \mu$, $Var(X) = \sigma^2$.

Es gelten die gleichen Regeln wie für endliche/diskrete Zufallsvariablen:

Seien X, Y stetige Zufallsvariablen.

(1) $\mathbb{E}(aX + b) = a \cdot \mathbb{E}(X) + b$, für $a, b \in \mathbb{R}$.

(2) $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

(3) Falls X und Y unabhängig sind, gilt: $\mathbb{E}(X \cdot Y) = \mathbb{E}(X)\,\mathbb{E}(Y)$

(4) $Var(X) \geq 0$

(5) $Var(aX + b) = a^2 \cdot Var(X)$ für $a, b \in \mathbb{R}$.

(6) Falls X und Y unabhängig sind, gilt: $Var(X+Y) = Var(X) + Var(Y)$.

(7) $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$

(8) Tschebyscheff-Ungleichung: $P(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \dfrac{Var(X)}{\varepsilon^2}$

Satz (Standardisieren)

Sei X eine Zufallsvariable mit $\mu := \mathbb{E}(X)$ und $\sigma^2 := Var(X)$, so dass $-\infty < \mu < \infty$ und $0 < \sigma^2 < \infty$. Dann gilt für $Y := \frac{X-\mu}{\sigma}$, dass

$$\mathbb{E}(Y) = 0 \quad \text{und} \quad Var(Y) = 1.$$

Wir nennen Y die Standardisierung von X.

Falls $X \sim N(\mu, \sigma^2)$, dann gilt für die Standisierung Y von X:

$$Y \sim N(0,1).$$

(deswegen nennen wir $N(0,1)$ die Standardnormalverteilung).