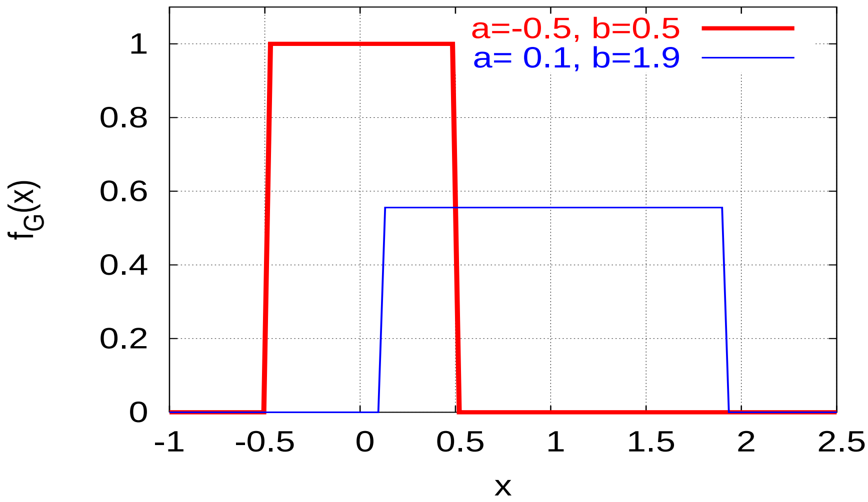
23. Einige kontinuierliche Verteilungen

Die Mathematisierung hat in einem Unternehmen vor allem den Vorzug, dass man sich viel genauer irren kann

23.1. Dichtefunktion der Gleichverteilung

$$f_G^{(a,b)}(x) = \begin{cases} \frac{1}{b-a} & \text{falls } a \le x \le b, \\ 0 & \text{sonst.} \end{cases}$$



Wo kommt die Gleichverteilung z.B. vor?

- Fehler beim Auf- oder Abrunden sowie beim Digitalisieren
- (Pseudo-) Zufallsgeneratoren von Programmen und Betriebssystemen.

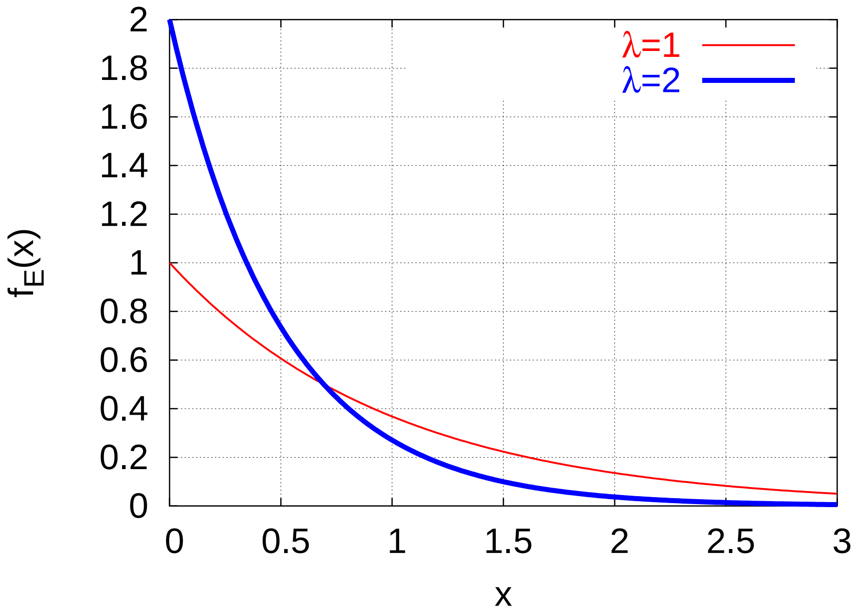
Aufgabe: Berechnen Sie für eine gleichverteilte Zufallsgröße $X \sim G(a,b)$

1. die Verteilungsfunktion $F_G(x)$,
2. den Erwartungswert $\mu = E(X)$ (Lösung: $\mu = \frac{a+b}{2}$),
3. und die Varianz $\sigma^2 = E(X-\mu)^2$ (Lösung: $\sigma^2 = \frac{1}{12}(b-a)^2$).

23.2 Exponentialverteilung

Dichtefunktion:

$$f_E^{(\lambda)}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } \; x \geq 0, \\ 0 & \text{sonst.} \end{cases}$$



Aufgabe: Berechnen Sie für eine exponentialverteilte Zufallsgröße $X \sim E(\lambda)$

1. die Verteilungsfunktion $F_E(x)$,
2. den Erwartungswert $\mu = E(X)$ (Lösung: $\mu = \frac{1}{\lambda}$),
3. und die Varianz $\sigma^2 = E(X - \mu)^2$ (Lösung: $\sigma^2 = \frac{1}{\lambda^2}$).

23.2 Exponentialverteilung II: Beispiele

Typischerweise exponentialverteilt sind Wartezeiten und Abstände aller Art, falls diese reellwertig und unabhängig von der Vorgeschichte bzw. vom Vorgänger sind:

1. Lebensdauer von Geräten, Glühbirnen, Bauelementen
2. Zeitlicher Abstand zwischen zwei Telefonanrufen
3. Abfertigungsdauer eines Kunden an der Supermarktkasse
4. Räumliche und zeitliche Abstände der Kfz auf der Autobahn bei sehr geringer Verkehrsdichte.

Zusammenhang mit der Poissonverteilung

Gilt für die Zahl n der innerhalb eines Zeitraums T eintreffenden Ereignisse die Poissonverteilung, so ist der Abstand Δ zwischen zwei aufeinanderfolgenden Ereignissen exponentialverteilt:

$$n \sim \mathrm{Po}(\mu) \;\Leftrightarrow\; \Delta \sim E(\mu/T)$$

Die poissonverteilten Größen n zu obigen Beispielen sind:

Zahl der in einer Zeiteinheit (z.B. Stunde) ...

1. kaputtgegangenen Glühbirnen,
2. Telefonanrufe
3. abgefertigten Kunden
4. vorbeigefahrenen Autos.

23.2 Exponentialverteilung III: Aufgabe

Weitere Aufgaben: Vgl. Übungsaufgaben

Eine Detektoranlage auf einer dreispurigen Autobahn misst einen Verkehrsfluss von 900 Fahrzeugen pro Stunde (summiert über alle Spuren einer Richtung). Da dies pro Stunde und Spur nur 300 Kfz entspricht, ist das Verkehrsaufkommen gering und die Zahl der (auf einer beliebigen Spur) passierenden Fahrzeuge kann als poissonverteilt angenommen werden. Ihr Parameter μ werde durch obigen Fluss abgeschätzt.

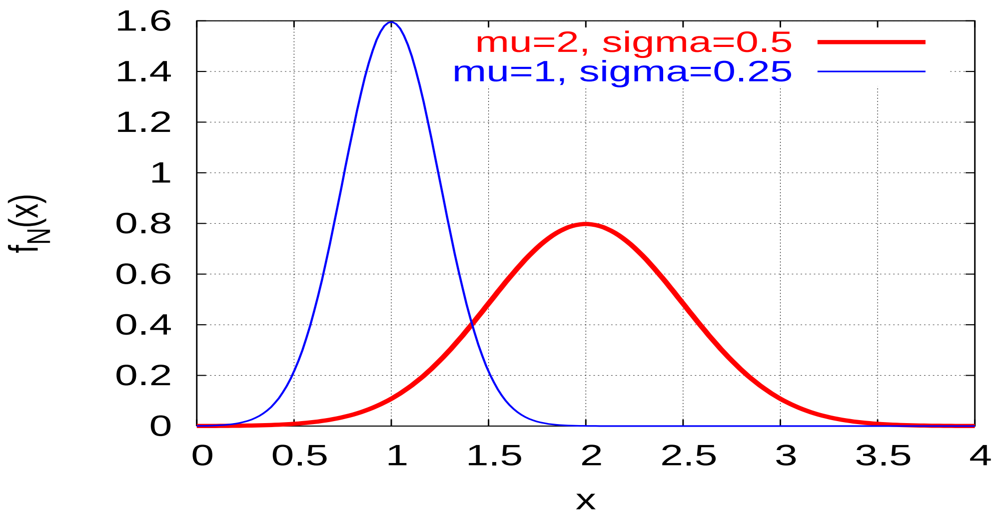
1. Mit welcher Wahrscheinlichkeit kommen in den nächsten 10 Sekunden weniger als 2 Fahrzeuge vorbei?
2. Geben Sie die Dichtefunktion der Zeitabstände zwischen zwei detektierten Fahrzeugen an.
3. Es herrsche nun dichter Verkehr (1800 Kfz/h/Spur) und es werde nur *eine* Spur betrachtet. Warum ist der Verkehr nun nicht mehr poissonverteilt?

23.3 Normalverteilung

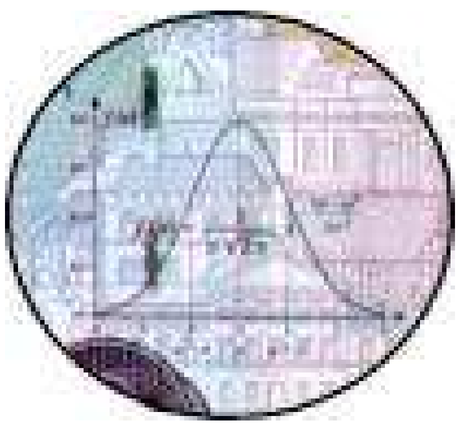
Die **Normalverteilung** oder auch **Gaußverteilung** ist die wichtigste stetige Verteilungsfunktion.

Die **Dichtefunktion** einer normalverteilten Zufallsvariablen $X \sim N(\mu, \sigma^2)$ ist gegeben durch

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



- Zwei Parameter $\mu = E(X)$ und $\sigma = \sqrt{E(x-\mu)^2}$
- Bereits im Jahre 1809 von Carl Friedrich Gauß entdeckt



23.3 Normalverteilung II

Warum ist die Normalverteilung so wichtig? Die Bedeutung der Gaußverteilung liegt darin, dass alle stetigen Zufallsgrößen, die aus vielen

- im Einzelbeitrag kleinen
- und unabhängigen

Größen resultieren, gaußverteilt sind: z.B. Messfehler, Größen, Alter etc. Dies folgt aus dem **Zentralen Grenzwertsatz**.

Berechnung der Verteilungsfunktion

Das die Normalverteilung definierende Integral ist nicht analytisch berechenbar, aber lässt sich einfach auf die sog. **standardisierte** Normalverteilung

$$\Phi(x) \equiv F_n^{(0,1)}(x)$$

transformieren. $\Phi(x)$ liegt für Argumente $x \geq 0$ häufig als $Tabelle$ vor. Wichtige Beziehungen dazu:

- Falls $X \sim N(\mu, \sigma^2)$ der Normalverteilung gehorcht, so ist folgende **transformierte** Zufallsvariable standardnormalverteilt:

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

- Es gilt: $\quad \Phi(-c) = 1 - \Phi(c)$.

- Damit lässt sich $F_N(x)$ für ein beliebiges Argument x durch $\Phi(z)$ mit positivem Argument ausdrücken:
  $P(X \leq x) = P(\sigma Z + \mu \leq x) = P(Z \leq \frac{x-\mu}{\sigma})$, also

$$F_N^{(\mu,\sigma^2)}(x) = \Phi(\frac{x - \mu}{\sigma})$$

23.3 Normalverteilung III: Aufgaben

1. Zeigen Sie, dass für beliebige Verteilungen $F_X(x)$ gilt:

$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

2. Zeigen Sie weiterhin, dass $Z = (X - \mu)/\sigma \sim N(0;1)$, wenn $X \sim N(\mu, \sigma^2)$

3. Die Höchstgeschwindigkeiten V von PKW's seien Gaußverteilt mit Erwartungswert 180 km/h und Varianz (20 km/h)^2.
   (a) Wieviel Prozent der PKW's haben Höchsgeschwindigkeiten von unter 190 km/h? Transformieren Sie die Verteilung auf eine Standardnormalverteilung, so dass Sie die entsprechende Tabelle verwenden können.
   (b) Wieviel Prozent der PKW's fahren schneller als 200 km/h?
   (c) Wieviel Prozent der PKW's fahren maximal zwischen 170 km/h und 200 km/h?
   (d) Berechnen Sie das 60% Quantil $v_{0.6}$ und das 25% Quantil $v_{0.4}$ der Höchstgeschwindigkeiten. Drücken Sie die Ergebnisse auch in Worten aus.

Weitere Aufgaben in den Übungen

Bemerkung: In der (Verkehrs-) Wirtschaft sind Zufallsgrößen häufig nicht die Summe, sondern das *Produkt* vieler kleiner zufälliger Einflüsse: Zum Beispiel sind Firmengrößen und Aktienkurse das Produkt der vergangenen Wachstumsfaktoren. Dann gilt die Normalverteilung für die *Logarithmen*.

23.4 Additionsgesetz und zentraler Grenzwertsatz

Additionsgesetz: Seien n voneinander unabhängige Zufallsvariable X_i, $i = 1, \cdots, n$ gegeben, deren (nicht notwendigerweise identische) Verteilungsfunktionen jeweils endliche Varianzen haben. Dann gelten für die Summe

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

die Beziehungen (siehe "Exkurse"):

$$\begin{aligned} E(Y_n) &= \textstyle\sum_{i=1}^{n} \mu_i = \mu, \\ V(Y_n) = E(Y_n - \mu)^2 &= \textstyle\sum_{i=1}^{n} \sigma_i^2 = \sigma^2. \end{aligned}$$

Auch für die Verteilungen *selbst* lassen sich Aussagen machen, z.B. ist die Summe zweier gleichverteilten Größen dreiecksverteilt. Näheres in den *Exkursen* zu diesem Kapitel.

Zentraler Grenzwertsatz

Mit steigendem n konvergiert die Summe Y der (nicht notwendigerweise alle derselben Verteilung gehorchenden) Zufallsvariablen X_i gegen die Normalverteilung:

$$Y_n \overset{n\to\infty}{\sim} N(\mu, \sigma^2).$$

Faustregel: Die Näherung $Y_n \sim N(\mu, \sigma^2)$ ist akzeptabel. falls für alle Einzelvarianzen gilt: $\sigma_i^2 \leq \sigma^2/30$.

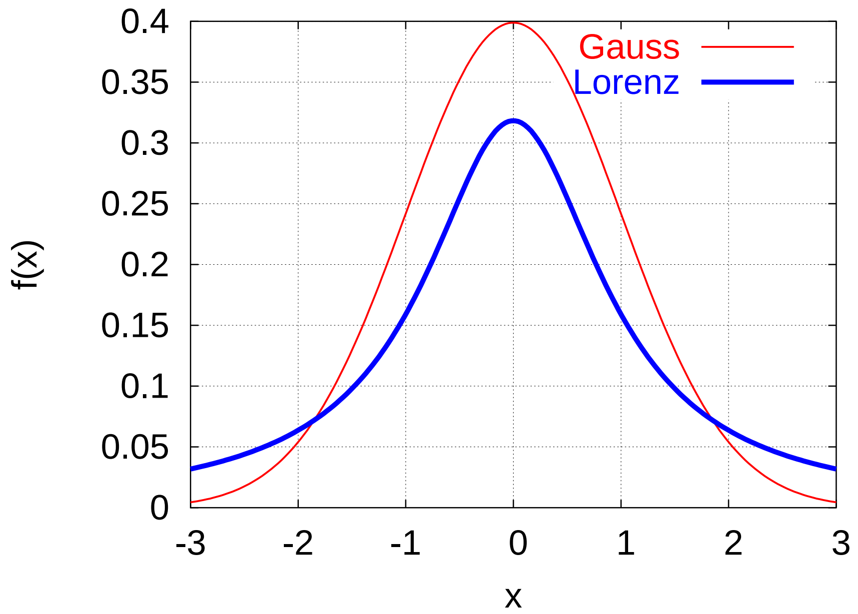
23.4 Zentraler Grenzwertsatz II: Bemerkungen

- Viele Zufallsprozesse der Natur und Technik lassen sich durch eine Zufallsgröße Y darstellen, die eine Summe vieler kleiner Einzelbeiträge ist und die obige Bedingungen erfüllt (z.B. auch die Fehlerfortpflanzung bei Landvermessungen, wo Gauß "seine" Verteilung entdeckte...). Daher die "zentrale" Bedeutung!

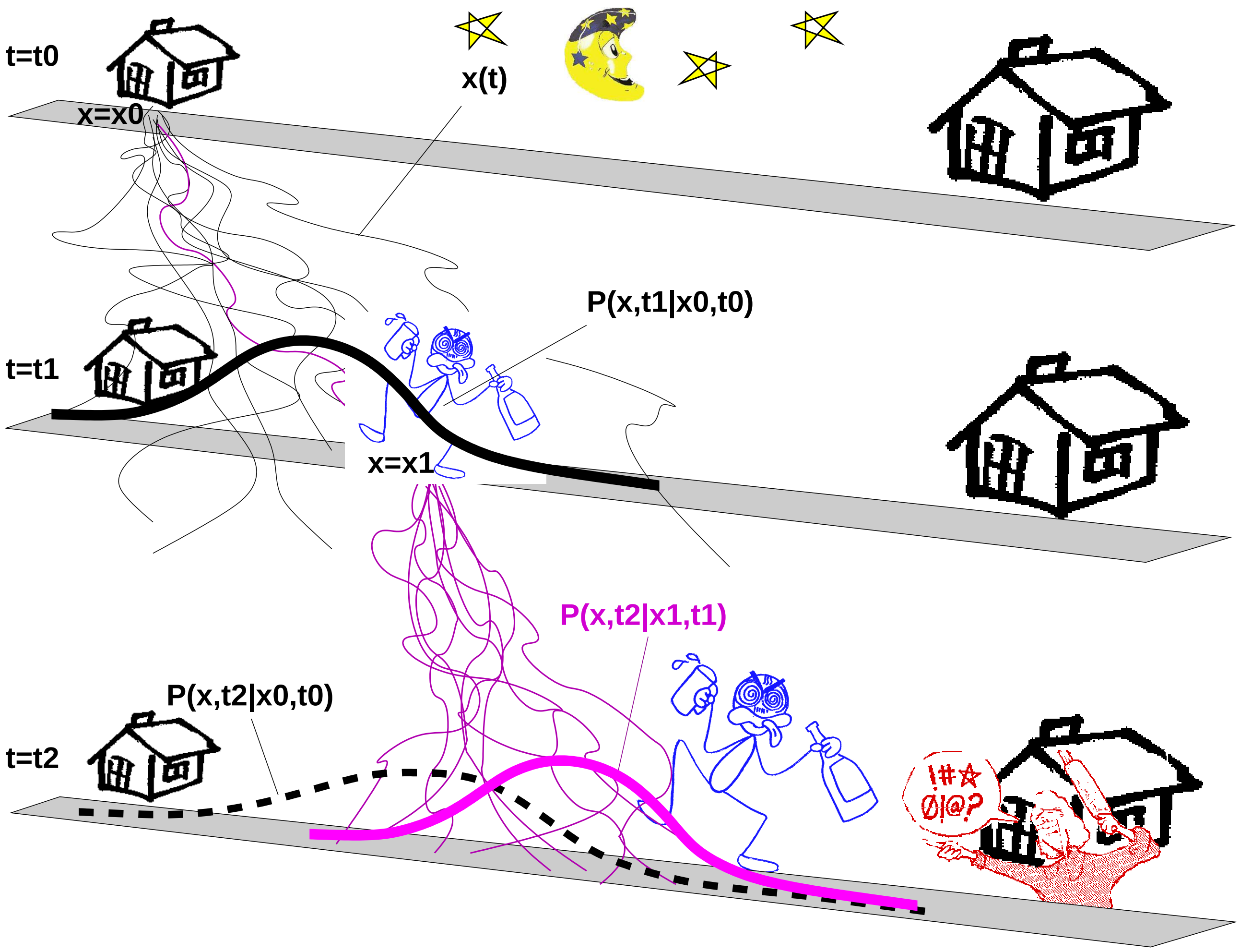
- In Verkehr und Wirtschaft wirken hingegen die Einzelbeiträge oft *multiplikativ* statt additiv, z.B. bei Börsenkursen, Firmengrößen, Einkommen (verdeutlichen Sie sich dies!). Dann gilt der zentrale Grenzwertsatz für die *Logarithmen*. Nach Transformation bekommt man die (hier nicht behandelte) **Lognormalverteilung**.

- Der zentraler Grenzwertsatz ist verblüffend und sehr weitreichend. Er gilt insbesondere auch für unsymmetrische (!) Einzelverteilungen, für stetige und diskrete Verteilungen (auch gemischt!), sowie für **alle** bisher behandelten Verteilungen.

- Es gibt aber "harmlos" aussehende Verteilungen wie die Lorenzverteilung mit der Dichte $f_L(x) = 1/(\pi(1 + x^2))$, für die er nicht gilt. Grund: Die Varianz der Lorenzverteilung ist unendlich!



Der “Drunkard’s Walk”



23.4 Zentraler Grenzwertsatz IV: Weitere Beispiele

(a) Auch die Binomialverteilung $B(n, \theta)$ geht für große n in die Normalverteilung über. Warum? Wie groß sind μ und σ?

(b) Übliche Zufallsgeneratoren erzeugen Realisierungen von (0,1)-gleichverteilten (Pseudo-) Zufallsvariablen $X \sim G(0, 1)$. Wie könnte man mit dem zentralen Grenzwertsatz daraus eine (μ, σ^2)-Normalverteilung simulieren? (Fällt Ihnen eine andere Methode ein, mit der man aus der Gleichverteilung beliebige Verteilungen erzeugen kann?)

(c) An der Salatbar einer Mensa kostet der Salat 1 € /100g. Der Salat wird gewogen und der Betrag zwecks leichterer Bezahlbarkeit auf ein Vielfaches von 50 Cent auf- oder abgerundet. Wie hoch ist das Risiko, dass der Student nach 192 maligem Salatessen durch die Rundung einen Nachteil von mehr als 3 € hat, wenn er das Salatgewicht nicht vorher abschätzt?