4. Zufallsvariablen

4.1 Grundlagen: Diskrete & Stetige Zufallsvariablen

Diskrete Zufallsvariablen (ZVn)

- Ergebnisse von Zufallsvorgängen sind nicht notwendigerweise Zahlen
- Oft ist es aber hilfreich diese durch Zahlen zu repräsentieren, um mit ihnen rechnen zu können
- Beispiel: 4-maliger Wurf einer Münze

  $\Omega = \{\text{Wappen}, \text{Zahl}\}^4 = \{W, Z\}^4$

  $|\Omega| = 2^4 = 16$ z. B. $\omega = (W, Z, Z, W)$

Eine diskrete Zufallsvariable

Angenommen man interessiert sich für

$$X := \text{“Anzahl von Wappen”}$$

Dann nennt man X eine **Zufallsvariable** (ZV) mit reellen **Ausprägungen** bzw. **Realisierungen** $x \in \mathbb{R}$.
Man schreibt kurz $X = x$, wenn die Ausprägung x der ZV X eingetreten ist.

X ist also eine **Abbildung** von Ω nach $\mathbb{R}$.
Also $X : \Omega \to \mathbb{R}$

Die Menge der möglichen Ausprägungen $\{0,1,2,3,4\}$ heißt **Träger** $\mathcal{T}$ der ZV X.

Vorteile von Zufallsvariablen

1. Man kann mit X "rechnen":

$$\text{z. B. } P(X \leq a) \quad \text{oder} \quad P(X^2 > b)$$

   Oder: Welche Zahl erhalten wir "im Mittel"?

2. Ursprünglicher Wahrscheinlichkeitsraum (Ω, P) wird letztendlich nicht mehr benötigt, stattdessen Wahrscheinlichkeiten für Ausprägungen der ZV.

Definition

Eine ZV X heißt **diskret**, falls sie nur endliche oder abzählbar unendlich viele Werte $x_1, x_2, \ldots$ annehmen kann. Die Menge $\mathcal{T} = \{x_1, x_2, \ldots\}$ der möglichen Ausprägungen (d.h. alle x_i mit $P(X = x_i) > 0$) von X heißt **Träger** der ZV X.

Die **Wahrscheinlichkeitsfunktion** von X ist durch

$$f(x_i) = P(X = x_i) = P\left(\{\omega \in \Omega : X(\omega) = x_i\}\right)$$

für $x_i \in \mathbb{R}$ gegeben. Die Wahrscheinlichkeitsfunktion $f(x_i)$ heißt auch **Wahrscheinlichkeitsdichte**.

Folgerungen

1. Für $x \notin \mathcal{T}$ ist $f(x) = P\{\emptyset\} = 0$.

2. Als Funktion von $B \subset \mathbb{R}$ ist also

$$P(X \in B) = P\{\omega \in \Omega : X(\omega) \in B\}$$

   eine Wahrscheinlichkeitsverteilung auf $\mathbb{R}$. Man nennt diese die **Verteilung** der Zufallsvariable X. Diese wird durch die Abbildung X und die Wahrscheinlichkeitsverteilung P induziert.

Die Verteilungsfunktion

Die **Verteilungsfunktion** einer diskreten ZV ist definiert als

$$F(x) = P(X \leq x) = \sum_{i: x_i \leq x} f(x_i)$$

Kennt man also die **Wahrscheinlichkeitsfunktion** $f(x)$ für alle $x \in \mathcal{T}$, so kennt man auch die **Verteilungsfunktion** $F(x)$ (dies gilt auch umgekehrt).

Eigenschaften der Verteilungsfunktion

- $F(x)$ ist monoton wachsend ("Treppenfunktion")
- $F(x)$ ist stückweise konstant mit Sprungstellen an Werten x_i mit $f(x_i) > 0$, d.h. an allen Realisierungen $x_i \in \mathcal{T}$
- $\lim\limits_{x \to \infty} F(x) = 1$
- $\lim\limits_{x \to -\infty} F(x) = 0$

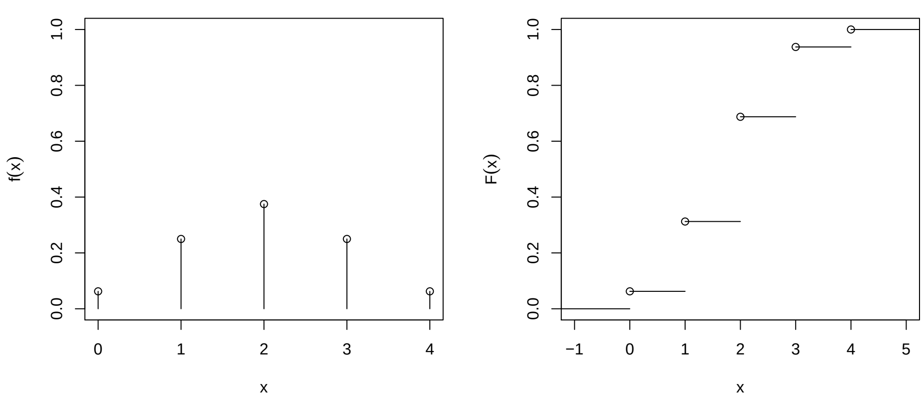
Beispiel: 4-maliger Wurf einer Münze

$X :=$ "Anzahl Kopf", $\mathcal{T} = \{0, 1, 2, 3, 4\}$

$$\begin{array}{rcl} f(0) & = & 1/16 \\ f(1) & = & 4/16 \\ f(2) & = & 6/16 \\ f(3) & = & 4/16 \\ f(4) & = & 1/16 \end{array} \quad \Rightarrow \quad F(x) = \begin{cases} 0 & : \quad x < 0 \\ 1/16 & : \quad 0 \le x < 1 \\ 5/16 & : \quad 1 \le x < 2 \\ 11/16 & : \quad 2 \le x < 3 \\ 15/16 & : \quad 3 \le x < 4 \\ 1 & : \quad x \ge 4 \end{cases}$$

Beachte: $f(x) = 0$ für alle $x \notin \mathcal{T}$

Wahrscheinlichkeits- und Verteilungsfunktion



Wahrscheinlichkeitsfunktion (links) und Verteilungsfunktion (rechts) für den viermaligen Münzwurf

Stetige Zufallsvariablen: Einleitung

Idee: Eine Zufallsvariable X ist **stetig**, falls ihr Träger eine *überabzählbare* Teilmenge der reellen Zahlen $\mathbb{R}$ ist.

Beispiel: Glücksrad mit stetigem Wertebereich $[0, 2\pi]$

Von Interesse ist die Zufallsvariable, die den *exakten* Winkel angibt, an dem das Glücksrad stehen bleibt.

Definition von stetigen Zufallsvariablen

Eine Zufallsvariable X heißt **stetig**, wenn es eine Funktion $f(x) \geq 0$ gibt, so dass sich die Verteilungsfunktion $F(x) = P(X \leq x)$ von X wie folgt darstellen lässt:

$$F(x) = \int_{-\infty}^{x} f(u)\,du$$

Die Funktion $f(x)$ heißt **Wahrscheinlichkeitsdichte** (kurz: **Dichte** oder **Dichtefunktion**) von X. Der Träger $\mathcal{T}$ von X ist die Menge aller Elemente $x \in \mathbb{R}$ für die $f(x) > 0$ gilt.
Beachte den Unterschied zu diskreten Zufallsvariablen! Hier gilt:

$$F(x) = \sum_{i: x_i \leq x} f(x_i)$$

Einige Folgerungen

$$\begin{aligned} P(X = x) &= 0 \quad \text{für alle } x \in \mathbb{R} \\ P(X \in [a, b]) &= \int_a^b f(x)\,dx \\ \int_{-\infty}^{+\infty} f(x)\,dx &= 1 \end{aligned}$$

Beispiel: Die stetige Gleichverteilung

Eine Zufallsvariable X heißt **stetig gleichverteilt** auf dem Intervall $[a, b]$, falls ihre Dichtefunktion die folgende Form hat:

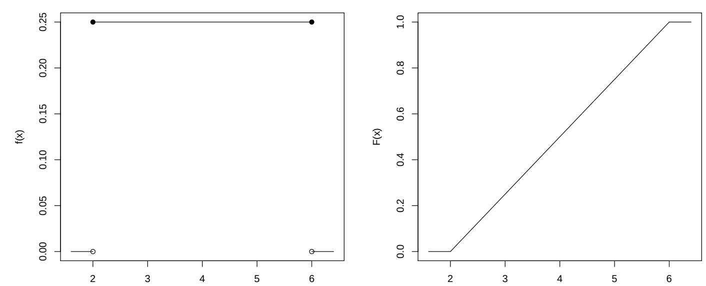
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

Der Träger von X ist also $\mathcal{T} = [a, b]$.
Die Verteilungsfunktion $F(x)$ von X ergibt sich zu

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

Man schreibt kurz: $X \sim \mathcal{U}(a, b)$

Beispiel: Die stetige Gleichverteilung II



Dichtefunktion (links) und Verteilungsfunktion (rechts) der stetigen Gleichverteilung für $a = 2$ und $b = 6$

Funktionen in R:

- `dunif(...)` berechnet die Dichtefunktion
- `punif(...)` berechnet die Verteilungsfunktion
- `qunif(...)` berechnet die Quantilsfunktion
- `runif(...)` erzeugt Zufallszahlen

Eigenschaften der Verteilungsfunktion

1. $\lim_{x \to -\infty} F(x) = 0$ und $\lim_{x \to \infty} F(x) = 1$
2. An allen Stetigkeitsstellen von $f(x)$ gilt: $F'(x) = f(x)$
3. $P(a \leq X \leq b) = F(b) - F(a)$
4. $P(X > a) = 1 - F(a)$
   etc.

Allgemeine Definition von stetigen ZVn

Frage: Für welche Mengen B ist die Aussage

$$P(X \in B) = \int_B f(x)dx$$

überhaupt sinnvoll?

Sei $\mathcal{F}$ die Mengenfamilie aller offenen Intervalle in $\mathbb{R}$. Dann gibt es eine sogenannte $\boldsymbol{\sigma}$-**Algebra** (eine spezielle Mengenfamilie) $\sigma(\mathcal{F})$, die $\mathcal{F}$ enthält.

σ-Algebren

Für eine σ-Algebra $\sigma(\mathcal{F})$ muss gelten:

1. $\emptyset$ und $\Omega \in \sigma(\mathcal{F})$
2. Für $A, B \in \sigma(\mathcal{F})$ ist auch $B \setminus A \in \sigma(\mathcal{F})$.
3. Für $A_1, A_2, \ldots \in \sigma(\mathcal{F})$ ist auch
   - $\bigcup_{n=1}^{\infty} A_n \in \sigma(\mathcal{F})$ und
   - $\bigcap_{n=1}^{\infty} A_n \in \sigma(\mathcal{F})$.

Axiome von Kolmogorov

Ein Wahrscheinlichkeitsmaß P auf Ω wird nun mittels $\sigma(\mathcal{F})$ definiert:
Für alle paarweise disjunkten Mengen $A_1, A_2, \ldots \in \sigma(\mathcal{F})$ soll gelten
(vgl. Axiom A3, Abschnitt 2.3):

$$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$$

Ferner müssen natürlich auch A1 und A2 erfüllt sein:

$$P(\emptyset) = 0$$
$$P(\Omega) = 1$$

Aus $P(\emptyset) = 0$ folgt, dass A1 gilt.
Begründung: Wenn $P(\emptyset) = 0 \Rightarrow P(A) \geq 0, \forall A \neq \emptyset$.

4.2 Wichtige Verteilungen

Spezielle diskrete Verteilungen

Man unterscheidet nun bestimmte “gängige” Verteilungen, die häufig von weiteren Parametern abhängen.

Das einfachste Beispiel ist die **Bernoulli-Verteilung**. Eine Bernoulli-verteilte ZV kann nur die Werte 0 und 1 annehmen:

$$\begin{aligned} P(X=1) &= f(1) &= \pi \\ P(X=0) &= f(0) &= 1-\pi \end{aligned}$$

$\pi \in [0,1]$ ist der Parameter der Bernoulli-Verteilung.

Man schreibt kurz: $X \sim \mathcal{B}(\pi)$.

Die diskrete Gleichverteilung

Die allgemeine **diskrete Gleichverteilung** hat einen endlichen Träger $\mathcal{T} = \{x_1, x_2, \ldots, x_k\}$, wobei

$$P(X = x_i) = f(x_i) = \frac{1}{k}$$

mit $i = 1, \ldots, k$ gilt.

Häufig sind alle natürlichen Zahlen zwischen $a \in \mathbb{N}$ und $b \in \mathbb{N}$ Element des Trägers $\mathcal{T}$. Die Grenzen a und b sind dann die Parameter der **diskreten Gleichverteilung**.

Beispiel: Augenzahl beim fairen Würfelwurf

Die geometrische Verteilung

Ein Zufallsvorgang, bei dem mit Wahrscheinlichkeit π ein Ereignis A eintritt, wird unabhängig voneinander so oft wiederholt, bis zum ersten Mal A eintritt.

Sei X die ZV “Anzahl der Versuche bis zum ersten Mal A eintritt”. Dann ist $\mathcal{T} = \mathbb{N}$ und die Wahrscheinlichkeitsfunktion von X lautet:

$$f(x) = (1-\pi)^{x-1} \cdot \pi \qquad x = 1, 2, 3, \ldots$$

$\pi \in (0,1)$ ist der Parameter der geometrischen Verteilung.

Man schreibt kurz: $X \sim \mathcal{G}(\pi)$.

Eine Variante der geometrischen Verteilung

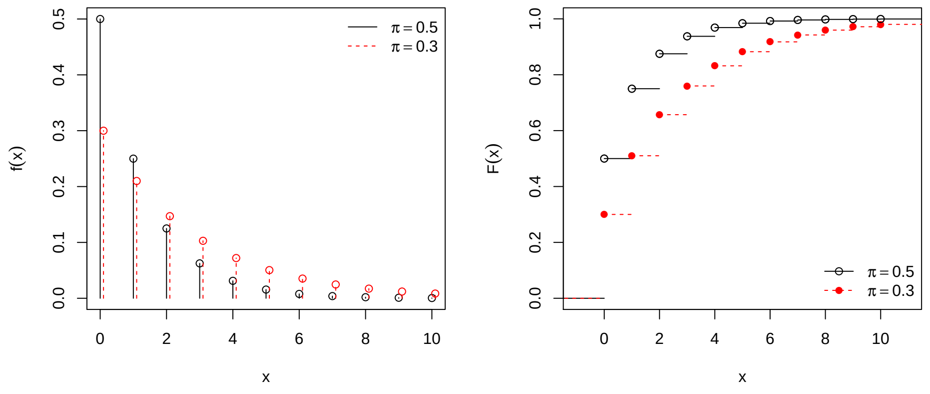
Sei $Y :=$ "Anzahl der Versuche *bevor* das erste mal A eintritt", d.h. $Y = X - 1$. Dann ist

$$\begin{aligned} \mathcal{T} &= \{0, 1, 2, \ldots\} \\ f(y) &= (1-\pi)^y \cdot \pi \end{aligned}$$

Für diese Form gibt es folgende Funktionen in **R**:

`dgeom()` berechnet Wahrscheinlichkeitsfunktion
`pgeom()` berechnet Verteilungsfunktion
`rgeom()` berechnet Zufallszahlen aus der geom. Verteilung

Wahrscheinlichkeits- und Verteilungsfunktion



Vergleich der geometrischen Wahrscheinlichkeitsfunktionen (links) und Verteilungsfunktionen (rechts) für die beiden Parameter $\pi = 0.3$ und $\pi = 0.5$

Die Binomialverteilung

Bei einer Folge von Bernoulli-Experimenten interessiert man sich häufig nur für die **Anzahl** $X := \sum_{i=1}^{n} X_i$, wie oft $X_i = 1$ aufgetreten ist.

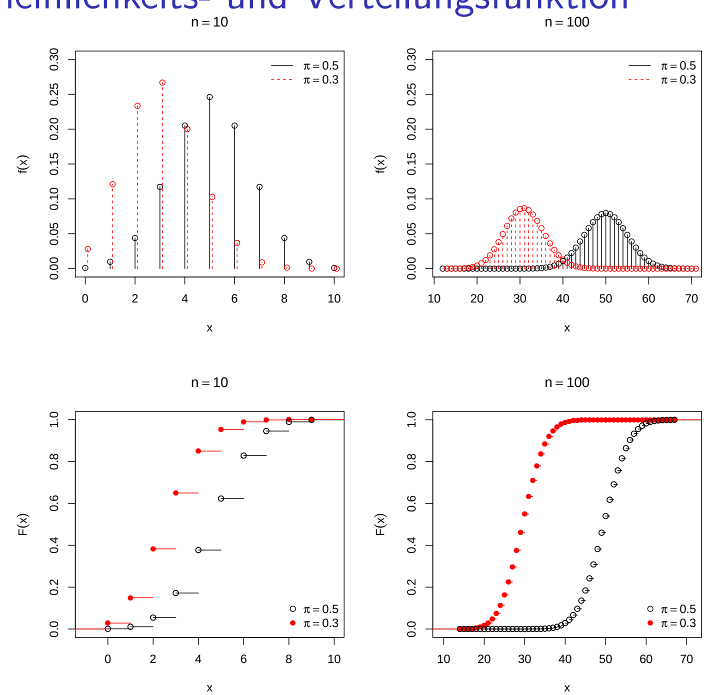
Diese ZV X heißt **binomialverteilt** mit Parametern $n \in \mathbb{N}$, $\pi \in [0, 1]$ und hat den Träger $\mathcal{T} = \{0, 1, \ldots, n\}$ sowie die Wahrscheinlichkeitsfunktion:

$$P(X = x) = f(x) = \binom{n}{x} \cdot \pi^x (1-\pi)^{n-x} \text{ für } x \in \mathcal{T}$$

Man schreibt kurz $X \sim \mathcal{B}(n, \pi)$ und es gilt $\mathcal{B}(1, \pi) = \mathcal{B}(\pi)$.

Funktionen in **R**:
`dbinom()`, `pbinom()`, `qbinom()`, `rbinom()`

Wahrscheinlichkeits- und Verteilungsfunktion



Beispiele

Das **Urnenmodell**: Zufälliges Ziehen **mit Zurücklegen** einer Stichprobe von n Kugeln aus einer Urne mit N Kugeln, darunter M markierte.
Sei X: "Anzahl der markierten Kugeln in der Stichprobe". Dann:

$$X \sim \mathcal{B}(n, M/N).$$

Die Poisson-Verteilung

Siméon Poisson [1781-1840]

Häufig gibt es zufällige Vorgänge, bei denen es keine natürliche obere Grenze für die Anzahl an Ereignissen gibt, z.B.:

- Die Anzahl an Telefonanrufen in einem "Call-Center" pro Stunde
- Die Anzahl der Tore in einem Bundesligaspiel
- Anzahl von Todesfällen durch Hufschlag in der Preußischen Armee (L. von Bortkiewicz, 1893)

Die einfachste Verteilung für solche Phänomene ist die Poisson-Verteilung.

Die Poisson-Verteilung II

Eine Zufallsvariable X folgt einer **Poisson-Verteilung**, wenn sie Träger $\mathcal{T} = \mathbb{N}_0$ und Wahrscheinlichkeitsfunktion

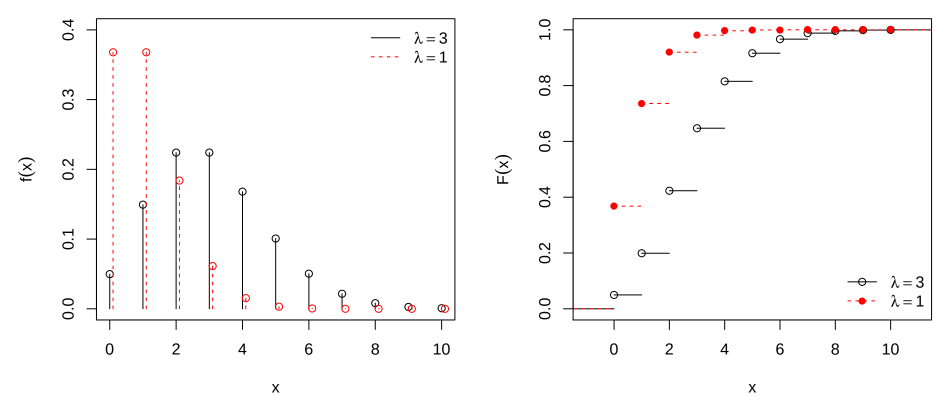
$$f(x) = \frac{\lambda^x}{x!} \cdot \exp(-\lambda)$$

hat. Der Parameter $\lambda \in \mathbb{R}^+$ ist die durchschnittliche **Rate** oder die **Intensität**, mit der die Ereignisse in dem zugrundeliegenden Zeitintervall auftreten.

Man schreibt kurz: $X \sim \mathcal{P}(\lambda)$

Funktionen in **R**: `{dpqr}pois`

Wahrscheinlichkeits- und Verteilungsfunktion



Vergleich der Wahrscheinlichkeitsfunktionen (links) und Verteilungsfunktionen (rechts) für eine poissonverteilte Zufallsvariable mit dem Parameter $\lambda = 1$ bzw. $\lambda = 3$

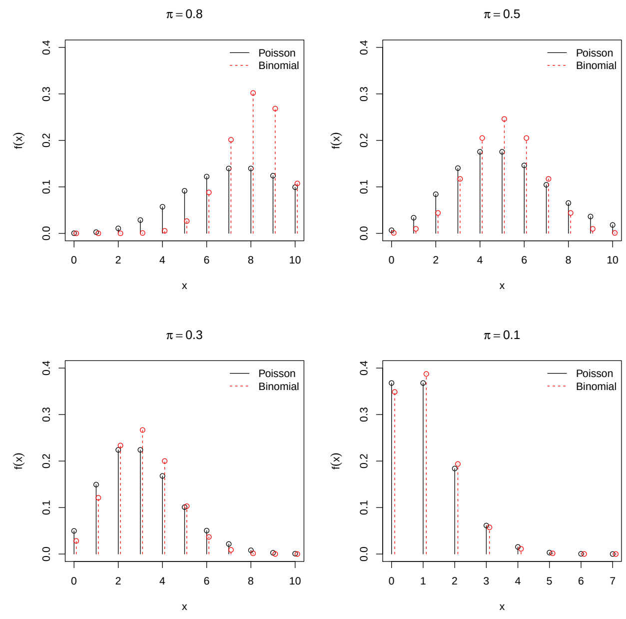
Approximation der Binomialverteilung

Die Binomialverteilung $B(n, \pi)$ kann für "großes n" und "kleines π" gut durch die Poisson-Verteilung mit $\lambda = n \cdot \pi$ approximiert werden.

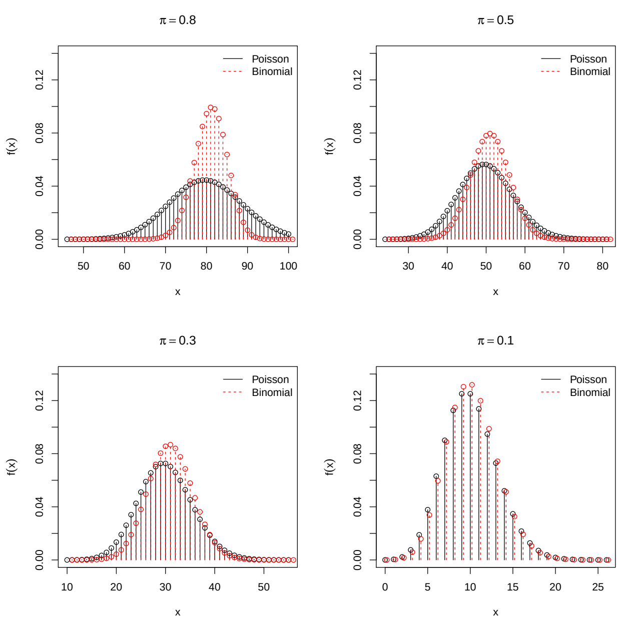
$$\mathcal{B}(n, \pi) \quad \approx \quad \mathcal{P}(\lambda = n \cdot \pi)$$

Je größer n ist und je kleiner π, desto besser ist die Approximation.

Vergleich von Binomial- und Poissonverteilung



Vergleich von Binomial- und Poissonverteilung



Wichtige stetige Verteilungen

Im Folgenden werden wir nun wichtige stetige Verteilungen kennenlernen. Stetige Verteilungen hängen wie diskrete Verteilungen von einem oder mehreren **Parametern** ab.
Zur Charakterisierung werden wir meist die **Dichtefunktion** und den **Träger** angeben.
Eine Verteilung haben wir schon kennengelernt, die **stetige Gleichverteilung** mit Parametern $a \in \mathbb{R}$ und $b \in \mathbb{R}$ $(a < b)$. Sie hat die Dichtefunktion

$$f(x) = \frac{1}{b - a}$$

und den Träger $\mathcal{T} = [a, b]$.

Die Exponentialverteilung

Eine stetige Zufallsvariable X mit positivem Träger $\mathcal{T} = \mathbb{R}_+$ heißt **exponentialverteilt** mit Parameter $\lambda \in \mathbb{R}_+$, wenn sie die Dichte

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

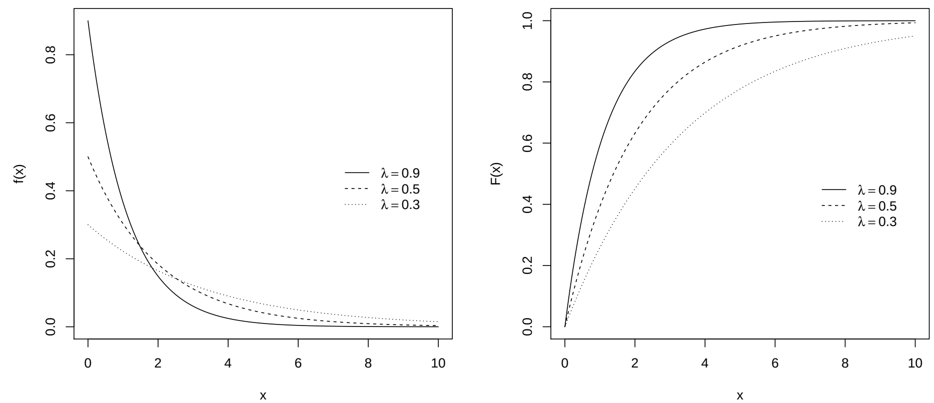
besitzt. Die Verteilungsfunktion ergibt sich zu

$$F(x) = \begin{cases} 1 - \exp(-\lambda x) & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

Notation: $X \sim \mathcal{E}(\lambda)$

Funktionen in `R`: `dexp()`, etc.

Die Exponentialverteilung II



Dichtefunktion (links) und Verteilungsfunktion (rechts) der Exponentialverteilung mit verschiedenen Raten

Die Gammaverteilung

Die Gammaverteilung ist eine Verallgemeinerung der Exponential-verteilung. Sie hat auch den positiven Träger $\mathcal{T} = \mathbb{R}_+$, aber einen Parameter mehr:

Eine stetige Zufallsvariable X heißt **gammaverteilt** mit Parametern $\alpha \in \mathbb{R}_+$ und $\beta \in \mathbb{R}_+$ (Notation: $X \sim \mathcal{G}(\alpha, \beta)$), falls sie die Dichte

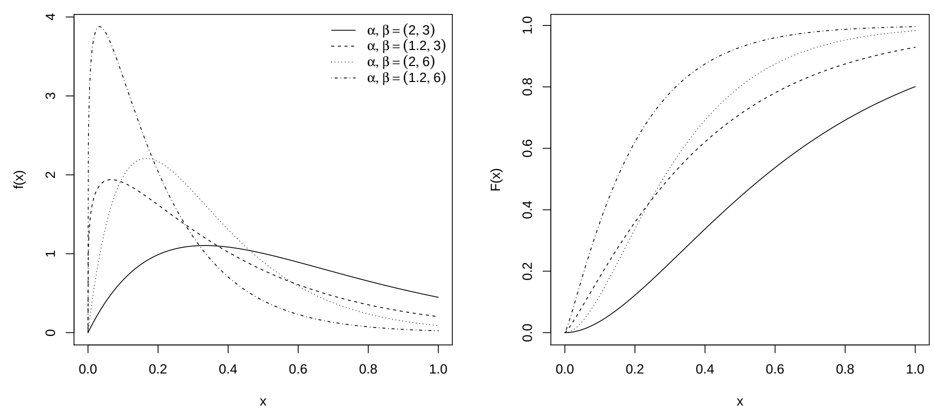
$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x) & \text{für } x > 0 \\ 0 & \text{sonst} \end{cases}$$

besitzt. Hierbei ist $\Gamma(\alpha)$ die **Gammafunktion**

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp(-x)\, dx$$

wobei $\Gamma(x+1) = x!$ für $x = 0, 1, 2, \ldots$ gilt.

Die Gammaverteilung II



Dichtefunktion (links) und Verteilungsfunktion (rechts) der Gammaverteilung mit verschiedenen Werten für α und β

Eigenschaften der Gammaverteilung

- Für $\alpha = 1$ ergibt sich die Exponentialverteilung mit Parameter $\lambda = \beta$.
- Für $\alpha = d/2$ mit $d \in \mathbb{N}$ und $\beta = \frac{1}{2}$ ergibt sich die sogenannte χ^2-**Verteilung** mit d Freiheitsgraden.

  Notation: $X \sim \chi^2(d)$
- Funktionen in `R`:
  - Gammaverteilung: `dgamma(x, shape =` α`, rate =` β`)`, etc.
  - χ^2-Verteilung: `dchisq(x, df = Freiheitsgrade)`, etc.

Wieso ist bei der Gammaverteilung $\int f(u)du = 1$?

Verwendung der Substitutionsregel:

$$\int \tilde{f}(g(x)) \cdot g'(x)\, dx = \int \tilde{f}(z)\, dz$$

mit $g(x) = \beta \cdot x$:

$$\begin{aligned} f(x) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x) \\ &= \frac{\beta}{\Gamma(\alpha)} g(x)^{\alpha-1} \exp(-g(x)) \\ &= \frac{1}{\Gamma(\alpha)} g(x)^{\alpha-1} \exp(-g(x)) \underbrace{\beta}_{g'(x)} \end{aligned}$$

Die Normalverteilung

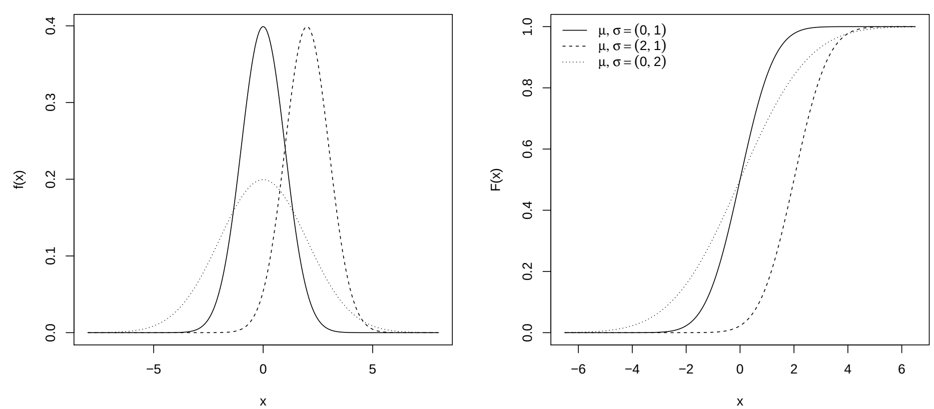
"Gaußsche Glockenkurve"

Eine Zufallsvariable X mit Träger $\mathcal{T} = \mathbb{R}$ und Parametern $\mu \in \mathbb{R}$ und $\sigma^2 \in \mathbb{R}_+$ heißt **normalverteilt**, falls sie die Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left(-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}\right) \quad \text{für } x \in \mathbb{R}$$

hat. Für $\mu = 0$ und $\sigma^2 = 1$ nennt man die Zufallsvariable **standardnormalverteilt**.
Man schreibt kurz: $X \sim \mathcal{N}(\mu, \sigma^2)$

Die Normalverteilung II



Dichtefunktion (links) und Verteilungsfunktion (rechts) der Normalverteilung mit verschiedenen Werten für μ und σ

Mehr zur Normalverteilung

Funktionen in R:

- `dnorm(...)` berechnet die Dichtefunktion
- `pnorm(...)` berechnet die Verteilungsfunktion
- `qnorm(...)` berechnet die Quantilsfunktion
- `rnorm(...)` erzeugt Zufallszahlen

Beachte:

$$F(x) = \int_{-\infty}^{x} f(u)\, du$$

ist nicht analytisch zugänglich (d.h. man findet keine Stammfunktion und braucht numerische Integration).

Wieso ist bei der Normalverteilung $\int f(u) du = 1$?

Man weiß aus der Analysis, dass für $a > 0$ gilt:

$$\int_{-\infty}^{\infty} \exp(-a^2 x^2)\, dx = \frac{\sqrt{\pi}}{a}$$

Ferner kann man leicht zeigen, dass

$$\int \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) dx = \int \exp\left(-\frac{x^2}{2\sigma^2}\right) dx$$

für alle $\mu \in \mathbb{R}$ gilt.

Die Betaverteilung

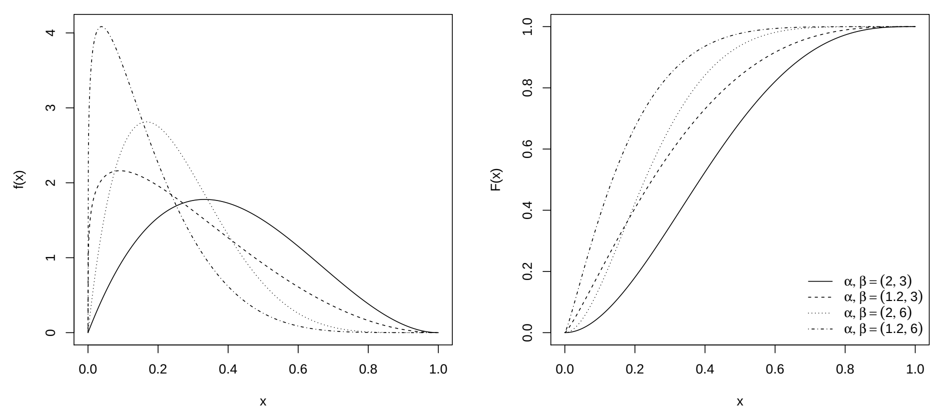
Eine Zufallsvariable X mit Träger $\mathcal{T} = (0, 1)$ und Parametern $\alpha \in \mathbb{R}_+$ und $\beta \in \mathbb{R}_+$ heißt **betaverteilt** ($X \sim \mathcal{B}e(\alpha, \beta)$), falls sie die Dichtefunktion

$$f(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{für } 0 < x < 1 \\ 0 & \text{sonst} \end{cases}$$

besitzt. Hierbei ist die **Betafunktion** $B(\alpha, \beta)$ gerade so definiert, dass die Dichtefunktion die Normierungseigenschaft $\int_0^1 f(x)\, dx = 1$ besitzt:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx$$

Die Betaverteilung II



Dichtefunktion (links) und Verteilungsfunktion (rechts) der Betaverteilung mit verschiedenen Werten für α und β

Die Betaverteilung III

Beachte: Für $\alpha = \beta = 1$ erhält man die Gleichverteilung auf dem Intervall $[0, 1]$.

Funktionen in `R`:

`dbeta(...)` berechnet Dichtefunktion
`pbeta(...)` berechnet Verteilungsfunktion
`qbeta(...)` berechnet Quantilsfunktion
`rbeta(...)` erzeugt Zufallszahlen

4.3 Erwartungswerte und Varianzen

Erwartungswerte und Varianzen – Einleitung

Zur Charakterisierung von Verteilungen unterscheidet man

- **Lagemaße**
  - Erwartungswert ("mittlerer Wert")
  - Modus
  - Median
- und **Streuungsmaße**
  - Varianz und Standardabweichung
  - mittlere absolute Abweichung (MAD)

Am einfachsten mathematisch zu handhaben sind Erwartungswerte und Varianzen, da diese immer eindeutig und meist leicht zu berechnen sind.

Der Modus von Zufallsvariablen

Ein **Modus** einer Zufallsvariable X ist ein Wert x_{Mod}, für den für alle $x \in T$ gilt:

$$f(x_{Mod}) \geq f(x)$$

Der Modus ist nicht notwendigerweise eindeutig, noch muss er existieren.

Beispiele

1. Der Modus der Betaverteilung ist für $\alpha > 1$ und $\beta > 1$ eindeutig gleich
$$x_{Mod} = \frac{\alpha - 1}{\alpha + \beta - 2}$$
2. Der Modus der Exponentialverteilung ist gleich Null.
3. Der Modus der Normalverteilung ist μ.
4. Der Modus der Gammaverteilung ist für $\alpha > 1$ eindeutig gleich
$$x_{Mod} = \frac{\alpha - 1}{\beta}$$
Für $\alpha < 1$ existieren keine Modi.

Der Erwartungswert einer diskreten ZV

Der **Erwartungswert** $\mathrm{E}(X) = \mathrm{E}X$ einer diskreten ZV X mit Träger $\mathcal{T}$ ist definiert als

$$\mathrm{E}(X) = \sum_{x \in \mathcal{T}} x \cdot P(X = x) = \sum_{x \in \mathcal{T}} x \cdot f(x)$$

wenn diese Summe absolut konvergent ist.

Beachte: Man könnte die Summe auch über alle $x \in \mathbb{R}$ laufen lassen.

Beispiele für Erwartungswerte:

- Bernoulli-Verteilung: Für $X \sim \mathcal{B}(\pi)$ ist $\mathrm{E}(X) = \pi$.
- Poisson-Verteilung: Für $X \sim \mathcal{P}(\lambda)$ ist $\mathrm{E}(X) = \lambda$.

Der Erwartungswert einer stetigen ZV

Der **Erwartungswert** einer stetigen Zufallsvariable X ist definiert als

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

unter der Voraussetzung, dass die Funktion $x \cdot f(x)$ **absolut integrierbar** ist:

$$\int_{-\infty}^{\infty} |xf(x)|\, dx = \int_{-\infty}^{\infty} |x| f(x)\, dx < \infty$$

Andernfalls sagt man, der Erwartungswert von X existiert nicht bzw. ist unendlich.

Eigenschaften des Erwartungswertes

1. Sei $X = a$ mit Wahrscheinlichkeit 1 (**deterministische ZV**). Dann gilt:
$$\mathrm{E}(X) = a$$
2. Ist $f(x)$ symmetrisch um einen Punkt c, d.h.
$$f(c - x) = f(c + x) \quad \forall x \in \mathcal{T}$$
dann ist $\mathrm{E}(X) = c$.
3. Seien $a, b \in \mathbb{R}$ und X, Y beliebige ZV. Dann gilt:
$$\mathrm{E}(a \cdot X + b \cdot Y) = a \cdot \mathrm{E}(X) + b \cdot \mathrm{E}(Y)$$
"*Linearität des Erwartungswertes*"
4. Für beliebige $a, b \in \mathbb{R}$ gilt daher
$$\mathrm{E}(aX + b) = a \cdot \mathrm{E}(X) + b$$

Eigenschaften des Erwartungswertes II

Allgemeiner gilt dann natürlich auch für beliebige $a_1, \ldots, a_n \in \mathbb{R}$ und beliebige ZV $X_1, \ldots, X_n$:

$$\mathrm{E}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i \cdot \mathrm{E}(X_i)$$

Daher gilt für $X \sim \mathcal{B}(n, \pi)$: $\mathrm{E}(X) = n\pi$

Transformationsregel für Erwartungswerte

- Sei X diskrete ZV und $g(x)$ eine reelle Funktion. Dann gilt für $Y = g(X)$:

$$\mathrm{E}(Y) = \mathrm{E}(g(X)) = \sum_{x \in \mathcal{T}} g(x)\, f(x)$$

- Im stetigen Fall:
$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$ für eine beliebige Funktion $g: \mathbb{R} \to \mathbb{R}$

Beachte: Im Allgemeinen gilt nicht: $\mathrm{E}(g(X)) = g(\mathrm{E}(X))$!

Erwartungswert von ZVn mit Träger $\mathbb{N}$

Hat X den Träger $\mathcal{T} = \mathbb{N}$, so gilt:

$$\mathrm{E}(X) = \sum_{k=1}^{\infty} P(X \geq k)$$

Anwendung: Erwartungswert der geometrischen Verteilung: Ist $X \sim G(\pi)$ so gilt:

$$\mathrm{E}(X) = \frac{1}{\pi}$$

Beispiel

Sei X eine ZV mit folgender Wahrscheinlichkeitsfunktion

$$f(x) = \begin{cases} 1/4 & \text{für} \quad x = -2 \\ 1/8 & \text{für} \quad x = -1 \\ 1/4 & \text{für} \quad x = 1 \\ 3/8 & \text{für} \quad x = 3 \end{cases}$$

Berechne den Erwartungswert von $\mathrm{E}(X^2)$

Varianzen und Standardabweichungen

Die **Varianz** Var(X) (auch V(X)) einer diskreten ZV ist definiert als:

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

"*Erwartete quadratische Abweichung vom Erwartungswert*"
Die **Varianz** einer stetigen Zufallsvariable definiert man analog:

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

mit $\mu = \mathrm{E}(X)$.

Beachte: Auch die Varianz kann nicht existieren, d.h. unendlich sein. Existiert der Erwartungswert nicht, so existiert auch die Varianz nicht.

Eigenschaften von Varianzen

1. Zur einfacheren Berechnung kann man häufig den **Verschiebungssatz** verwenden:

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

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

3. Sind X und Y unabhängig, so gilt:

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

Die Standardabweichung

Die **Standardabweichung** einer ZV X ist definiert als die Wurzel aus der Varianz:

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

Im Gegensatz zur Varianz gilt für die Standardabweichung:

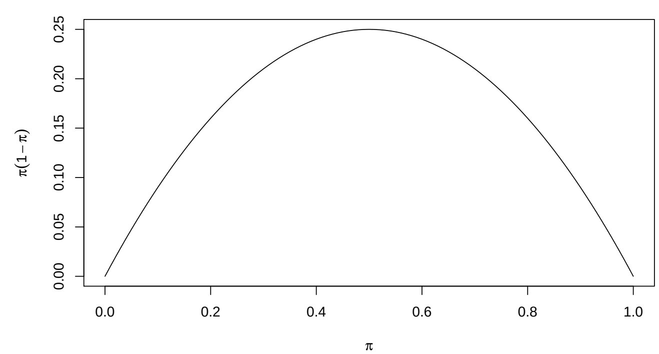
$$\sigma(aX + b) = |a| \cdot \sigma(X)$$

Bemerkung:
Die mittlere absolute Abweichung $\mathrm{E}(|X - \mathrm{E}X|)$ erscheint als Streuungsmaß intuitiver, ist aber deutlich schwerer mathematisch zu handhaben.

Beispiel: Binomialverteilung

Ist $X \sim \mathcal{B}(n, \pi)$, so gilt für die Varianz:

$$\mathrm{Var}(X) = n \cdot \pi \cdot (1 - \pi)$$



Varianz der Bernoulliverteilung als Funktion von π

Beispiel: Varianz der stetigen Gleichverteilung

Zunächst gilt

$$E(X^2) = \frac{1}{3} \cdot \frac{b^3 - a^3}{b - a}$$

und mit dem Verschiebungssatz ergibt sich:

$$\mathrm{Var}(X) = EX^2 - (EX)^2 = \frac{(b-a)^2}{12}$$

Die Varianz wächst also *quadratisch* und die Standardabweichung somit *linear* mit der Breite $b - a$ des Trägers.

Erwartungswerte und Varianzen stetiger Verteilungen

Name	Symbol	E(X)	Var(X)
Gleichverteilung	$X \sim \mathcal{U}(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentialverteilung	$X \sim \mathcal{E}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gammaverteilung	$X \sim \mathcal{G}(\alpha,\beta)$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Normalverteilung	$X \sim \mathcal{N}(\mu,\sigma^2)$	μ	σ^2
Betaverteilung	$X \sim \mathcal{B}e(\alpha,\beta)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\cdot\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Ungleichung von Tschebyscheff

Als Maß für die Streuung einer Verteilung ist die Varianz
bzw. Standardabweichung einer ZV X schwer direkt zu interpretieren.
Es gilt aber zumindest folgende Ungleichung:

$$P(|X - E(X)| \geq c) \leq \frac{\mathrm{Var}(X)}{c^2}$$

Beispiel:

Sei $E(X)$ beliebig und $\mathrm{Var}(X) = 1$. Dann ist

$$P(|X - E(X)| \geq 1) \leq 1$$

$$P(|X - E(X)| \geq 2) \leq \frac{1}{4}$$

$$P(|X - E(X)| \geq 3) \leq \frac{1}{9}$$

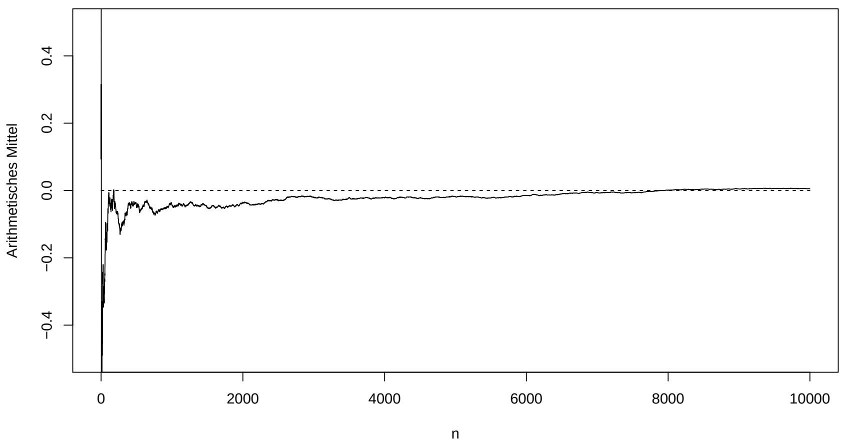
Das Gesetz der großen Zahlen

Das Gesetz der großen Zahlen ist eine Aussage über das **arithmetische Mittel**

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

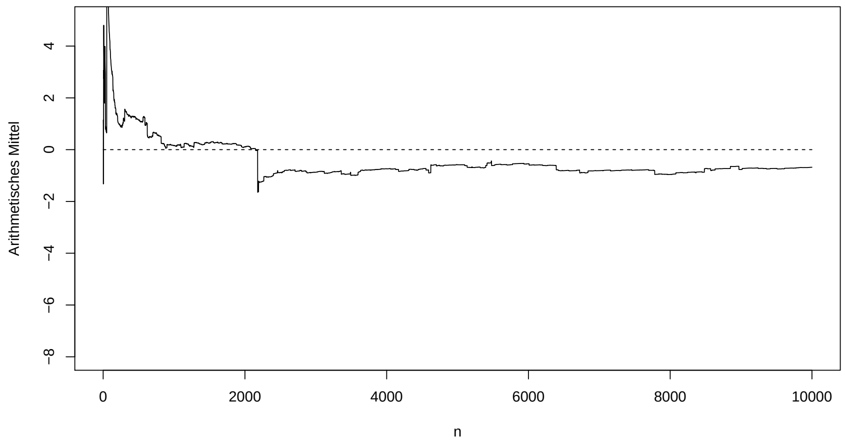
für $n \to \infty$, wobei die X_i ($i = 1, \ldots, n$) unabhängig und identisch verteilte Zufallsvariablen (engl.: *iid* = "independent and identically distributed") aus einer Verteilung mit Erwartungswert μ und Varianz σ^2 sind.
Es gilt: $E(\bar{X}_n) = \mu$ und $\mathrm{Var}(\bar{X}_n) = \frac{1}{n}\sigma^2$
Daher folgt sofort für $n \to \infty$: $\bar{X}_n \to \mu$ und $\mathrm{Var}(\bar{X}_n) \to 0$
$\Rightarrow$ Das arithmetische Mittel konvergiert gegen den Erwartungswert.
Dies funktioniert nicht bei der Cauchy-Verteilung!

Beispiel: Normalverteilung



Arithmetisches Mittel für 10000 standardnormalverteilte ZV

Beispiel: Cauchyverteilung



Arithmetisches Mittel für 10000 cauchyverteilte ZV

Stichprobenvarianz

Seien X_i ($i = 1, \dots, n$) wieder iid Zufallsvariablen aus einer Verteilung mit Erwartungswert μ und Varianz σ^2. Dann gilt für

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2,$$

dass $E(S^2) = \sigma^2$ und für $n \to \infty$: $S^2 \to \sigma^2$

Wenn μ unbekannt ist und durch $\bar{X}_n$ ersetzt werden muss ist allerdings $E(S^2) \neq \sigma^2$. Für $S^{2*} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ gilt die Gleichheit wieder. Mehr dazu in Kapitel 5.

Der Transformationssatz für Dichten

Sei X eine stetige Zufallsvariable mit Dichte $f_X(x)$. Betrachte nun die Zufallsvariable $Y = g(X)$, wobei z.B. $Y = \exp(X)$, $Y = X^2, \dots$
Frage: Wie lautet die Dichte $f_Y(y)$ von Y?
Für eine streng monotone und differenzierbare Funktion g gilt der **Transformationssatz für Dichten**:
Beweis über die Verteilungsfunktion $F_Y(y)$ von Y in der Vorlesung.

Beispiel: Das Quadrat einer Standardnormalverteilung

Wie lautet die Dichte von $Y = X^2$, falls $X \sim \mathcal{N}(0,1)$?
Betrachte zunächst $Z = |X|$. Z hat offensichtlich die Dichte

$$f(z) = \frac{2}{\sqrt{2\pi}} \exp(-\frac{1}{2} z^2) \quad \text{für } z > 0 \text{ und } 0 \text{ sonst}$$

Nun ist $X^2 = Y = Z^2 = g(Z)$ und g monoton wachsend auf dem Wertebereich $\mathbb{R}^+$. Es ergibt sich ($y = z^2 \Leftrightarrow z = \sqrt{y}$)

$$f(y) \quad = \quad \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} \cdot \exp(-\frac{1}{2} y)$$

Dies entspricht der Dichte einer $\mathcal{G}(0.5, 0.5)$, also einer χ^2-Verteilung mit 1 Freiheitsgrad: $Y = X^2 \sim \chi^2_1$

Beispiel: Erzeugung exponentialverteilter Zufallsvariablen

Betrachte $X \sim \mathcal{U}[0,1]$ und $Y = -\log(X)$, also $g(x) = -\log(X)$. Die Umkehrfunktion und deren Ableitung lauten:

$$g^{-1}(y) = \exp(-y) \qquad \frac{dg^{-1}(y)}{dy} = -\exp(-y)$$

Durch Anwendung des Transformationssatzes für Dichten erhält man:

$$f_Y(y) = 1 \cdot |-\exp(-y)| = \exp(-y)$$

Es gilt: $Y \sim \mathcal{E}(\lambda = 1)$! Dies ist also eine einfache Art, exponentialverteilte Zufallsvariablen zu erzeugen!
Allgemeiner liefert $Y = -\frac{1}{\lambda}\log(x)$ Zufallszahlen aus einer Exponentialverteilung mit Parameter λ : $\quad Y \sim \mathcal{E}(\lambda)$

Die Inversions-Methode

Allgemeiner kann man die **Inversions-Methode** zur Erzeugung von Zufallszahlen aus einer beliebigen stetigen Verteilung mit Verteilungsfunktion $F(x)$ verwenden:

- Erzeuge stetig gleichverteilte Zufallsvariablen $U_1, \ldots, U_n$ auf dem Intervall $[0,1]$.
- Dann sind
$$X_i = F^{-1}(U_i), \quad i = 1, \ldots, n$$
Zufallszahlen aus der gewünschten Verteilung.

Beweis: Die Dichte von X_i ergibt sich zu:

$$f_X(x) = \underbrace{f_U(F(x))}_{=1} \cdot \underbrace{F'(x)}_{f(x)} = f(x)$$

Beispiel: Zufallszahlen aus der Cauchy-Verteilung

Dichte- und Verteilungsfunktion der Cauchy-Verteilung sind:

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2} \quad \text{und} \quad F(x) = \frac{1}{2} + \frac{\arctan(x)}{\pi}$$

Die inverse Verteilungsfunktion ist somit:

$$F^{-1}(y) = \tan\left[\pi\left(y - \frac{1}{2}\right)\right]$$

Zufallszahlen aus der Cauchy-Verteilung lassen sich also leicht erzeugen, indem man $U_1, \ldots, U_n$ aus $\sim \mathcal{U}[0,1]$ erzeugt und $\tan\left[\pi\left(U_i - \frac{1}{2}\right)\right]$ berechnet.

Der zentrale Grenzwertsatz

- Aussage, dass - unter Regularitätsbedingungen - das arithmetische Mittel, geeignet standardisiert,
  von *beliebigen* unabhängig und identisch verteilten Zufallsvariablen gegen die Standardnormalverteilung konvergiert.
- Begründet die zentrale Rolle der Normalverteilung in der Stochastik.

Zunächst müssen wir noch *standardisierte* Zufallsvariablen definieren.

Standardisierte Zufallsvariablen

Jede Zufallsvariable X mit endlichem Erwartungswert $\mu = \mathrm{E}(X)$ und endlicher Varianz $\sigma^2 = \mathrm{Var}(X)$ kann man derart linear transformieren, dass sie Erwartungswert 0 und Varianz 1 besitzt:

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

Dann gilt:

$$\begin{aligned} \mathrm{E}(\tilde{X}) &= \frac{1}{\sigma}(\mathrm{E}(X) - \mu) = 0 \\ \mathrm{Var}(\tilde{X}) &= \frac{1}{\sigma^2}\mathrm{Var}(X) \quad = 1 \end{aligned}$$

Standardisierung von Summen von *iid* ZVn

Betrachte *iid* Zufallsvariablen $X_1, X_2, \ldots, X_n$ mit endlichem Erwartungswert $\mu = \mathrm{E}(X_i)$ und endlicher Varianz $\sigma^2 = \mathrm{Var}(X_i)$.
Für die Summe $Y_n = X_1 + X_2 + \ldots + X_n$ gilt offensichtlich:

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

Für die standardisierte Summe

$$Z_n = \frac{Y_n - n\mu}{\sqrt{n} \cdot \sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma}$$

gilt somit $\mathrm{E}(Z_n) = 0$ und $\mathrm{Var}(Z_n) = 1$.

Der zentrale Grenzwertsatz (ZGWS)

Die Verteilungsfunktion $F_n(z)$ von Z_n konvergiert für $n \to \infty$ an jeder Stelle $z \in \mathbb{R}$ gegen die Verteilungsfunktion $\Phi(z)$ der Standardnormalverteilung.

Schreibweise: $F_n(z) \to \Phi(z)$ für $n \to \infty$ und alle $z \in \mathbb{R}$
bzw. kurz $Z_n \overset{a}{\sim} \mathcal{N}(0, 1)$ (a = "asymptotisch")

In der Praxis kann man also die Verteilung von Z_n für große n gut durch eine Standardnormalverteilung approximieren.

Bemerkungen

- der ZGWS gilt sowohl für stetige als auch für diskrete ZV X_i
- X_i kann beliebig "schiefe" Verteilungen haben, z.B.

$$X_i \sim \mathcal{E}(\lambda)$$

- Die Standardisierung ist nicht notwendig zur Formulierung des ZGWS. Alternativ kann man auch direkt $Y_n = X_1 + \ldots + X_n$ betrachten. Dann gilt

$$Y_n \overset{a}{\sim} \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$$

Beispiel: Bernoulliverteilung

Seien $X_i \sim \mathcal{B}(\pi)$, $i = 1, \ldots, n$ und unabhängig.
Dann ist $Y_n = \sum_{i=1}^n X_i \sim \mathcal{B}(n, \pi)$ und asymptotisch gilt:

$$\frac{Y_n - n \cdot \pi}{\sqrt{n \cdot \pi(1-\pi)}} \overset{a}{\sim} \mathcal{N}(0,1)$$

bzw.

$$Y_n \overset{a}{\sim} \mathcal{N}(n \cdot \pi, n \cdot \pi(1-\pi))$$

4.4 Stochastische Unabhängigkeit

Unabhängigkeit von Zufallsvariablen

Betrachte zwei diskrete ZVn X und Y auf dem Wahrscheinlichkeitsraum (Ω, P)
mit Träger $\mathcal{T}_X = \{x_1, x_2, \ldots\}$ und $\mathcal{T}_Y = \{y_1, y_2, \ldots\}$ und Wahrscheinlichkeitsfunktionen $f_X(x)$ und $f_Y(y)$.

Die Funktion

$$f_{X,Y}(x,y) = P(X = x \text{ und } Y = y) = P(X = x, Y = y)$$

heißt **gemeinsame Wahrscheinlichkeitsfunktion** der zwei Zufallsvariablen X und Y.

Unter Unabhängigkeit kann man diese zurückführen auf die beiden Wahrscheinlichkeitsfunktionen $f_X(x)$ und $f_Y(y)$.

Die gemeinsame Verteilung zweier ZVn

Die **gemeinsame Verteilungsfunktion** zweier Zufallsvariablen X und Y ist die Funktion

$$F(x,y) = P(X \leq x \text{ und } Y \leq y)$$

Alternativ kann man die gemeinsame Verteilung von X und Y auch über deren **gemeinsame Dichtefunktion** $f(x,y)$ definieren, wobei

$$F(x,y) = \int_{v=-\infty}^{y} \int_{u=-\infty}^{x} f(u,v)\,du\,dv$$

für alle $x, y \in \mathbb{R}$ für stetige ZVn und

$$F(x,y) = \sum_{\{u \in \mathcal{T}_X | u \leq x\}} \sum_{\{v \in \mathcal{T}_Y | v \leq y\}} f(u,v)$$

für alle $x \in \mathcal{T}_X$, $y \in \mathcal{T}_Y$ für diskrete ZVn gelten muss.

Eigenschaften von $f(x,y)$ und $F(x,y)$

Für $f(x,y)$ stetig gilt:

$$\frac{d^2 F(x,y)}{dx\,dy} = f(x,y)$$

$f(x,y)$ ist normiert:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1$$

Bemerkung: Manchmal schreiben wir explizit $F_{X,Y}$ für F und $f_{X,Y}$ für f.

Randverteilungen einer Funktion

Die Dichten der Randverteilungen sind gegeben durch:

$$f_X(x) = \sum_y f_{X,Y}(x,y)\ ,\ f_Y(y) = \sum_x f_{X,Y}(x,y) \text{ im diskreten Fall und}$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy\ ,\ f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx \text{ im stetigen Fall}$$

Die **gemeinsame Verteilung** von X und Y enthält i. A. mehr Information als in den **Randverteilungen** von X und Y steckt!

Unabhängigkeit – Definition

ZVn X, Y heißen **unabhängig**, genau dann, wenn

$$\begin{aligned} F_{X,Y}(x,y) &= F_X(x) \cdot F_Y(y) \\ \text{bzw.} \quad f_{X,Y}(x,y) &= f_X(x) \cdot f_Y(y) \\ \text{bzw.} \quad P(X = x, Y = y) &= P(X = x) \cdot P(Y = y) \end{aligned}$$

$\forall x \in \mathcal{T}_X$ und $\forall y \in \mathcal{T}_Y$ im diskreten Fall, $\forall x, y \in \mathbb{R}$ im stetigen Fall.

Allgemein: ZVn $X_1, X_2, \ldots, X_n$ heißen **unabhängig**, falls

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

für alle $x_1, x_2, \ldots, x_n$ aus den entsprechenden Trägern gilt.

Beispiel: Bernoulli-Kette

Sind $X_1, X_2, \ldots, X_n$ Bernoulli-verteilt mit Parameter π und unabhängig, so heißt $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ **Bernoulli-Folge**.

Beispiel:

$n = 3$, $\pi = \frac{1}{6}$; wegen der Unabhängigkeit gilt z. B.

$$P(X_1 = 1, X_2 = 0, X_3 = 0) = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 = \frac{25}{216}$$

Die hypergeometrische Verteilung

Häufig wird jedoch im Urnenmodell **ohne Zurücklegen** gezogen, d. h. die "Auswahlwahrscheinlichkeiten" ändern sich von Ziehung zu Ziehung (Beispiel: Kartenspiele).

Die Verteilung von X (Anzahl der markierten Kugeln) nennt man dann **hypergeometrisch**. Sie hat den Träger

$$\mathcal{T} = \{\max(0, n - (N - M)), \ldots, \min(n, M)\}$$

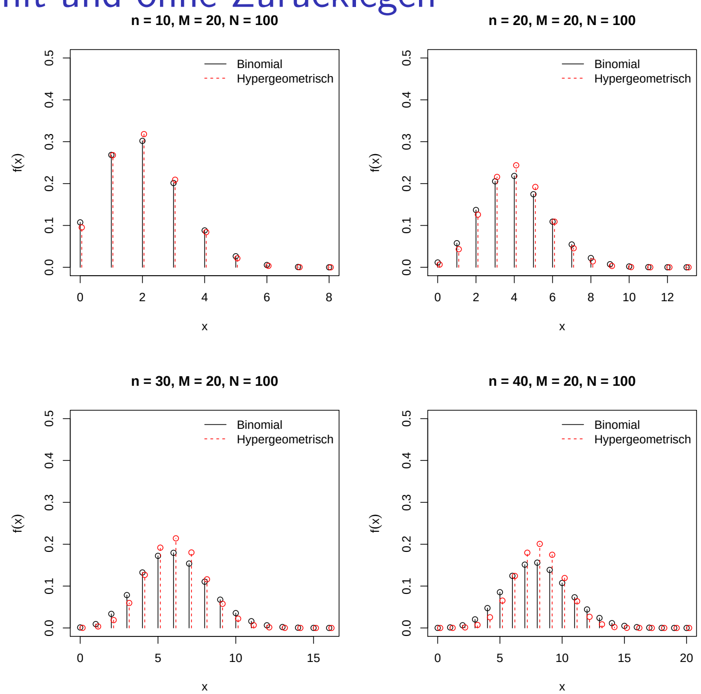
und die Wahrscheinlichkeitsfunktion

$$f(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \qquad \text{für } x \in \mathcal{T}.$$

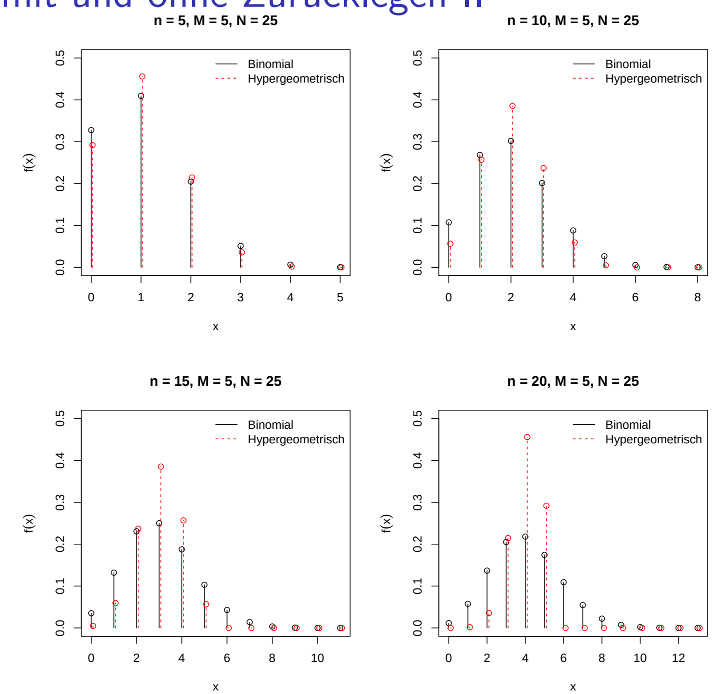
Man schreibt kurz: $X \sim \mathcal{H}(n, N, M)$

Funktionen in **R**: `{dpqr}hyper()`

Ziehen mit und ohne Zurücklegen



Ziehen mit und ohne Zurücklegen II



Approximation der hypergeometrischen Verteilung

Für N "groß" und n "klein" läßt sich die hypergeometrische Verteilung gut durch die Binomialverteilung approximieren:

$$\mathcal{H}(n, M, N) \quad \approx \quad \mathcal{B}\left(n, \pi = \frac{M}{N}\right)$$

Beispiel: Capture-Recapture-Experiment

Frage: Wie viele Fische schwimmen in einem See?

Idee : Fange M Fische, markiere diese und wirf sie dann wieder (lebendig) in den See zurück. Später werden n Fische gefangen.

Die ZV $X :=$ "Anzahl der markierten Fische" ist idealerweise hypergeometrisch verteilt mit Parametern N, M und n:

$$X \sim \mathcal{H}(n, N, M)$$

Beispiel: Capture-Recapture-Experiment II

Statistisches Problem: Wie groß ist N?

Naiver Ansatz zur Konstruktion eines **Schätzers** $\hat{N}$ von N:

$$\frac{N}{M} \approx \frac{n}{x} \quad \Rightarrow \quad \hat{N} \approx \frac{n}{x} \cdot M$$

Probleme:

- Im Allgemeinen $\hat{N} \notin \mathbb{N}$
- Keine Angaben über die Genauigkeit der Schätzung

Später mehr dazu.

Beispiel

Ein Lehrer bittet seine Schüler, eine (faire) Münze zweimal zu werfen, und das Ergebnis ("Kopf" $= 0$, "Zahl" $= 1$) für jeden Wurf zu notieren. Sei X das Ergebnis des ersten Wurfes und Y das Ergebnis des zweiten Wurfes.

- Ein gewissenhafter Schüler folgt genau den Anweisungen des Lehrers und notiert das Ergebnis X_G und Y_G. Ein fauler Schüler wirft nur eine Münze und notiert das erzielte Ergebnis zweimal: X_F und Y_F.
- Berechne die gemeinsame Wahrscheinlichkeitsfunktion von (X_G, Y_G) und von (X_F, Y_F).

4.5 Bedingte Wahrscheinlichkeiten

Bedingte Verteilungen von diskreten ZVn

Die **bedingte Verteilungsfunktion** und **bedingte Wahrscheinlichkeitsfunktion** von X, gegeben $Y = y$, sind definiert für alle y mit $P(Y = y) > 0$:

$$
\begin{aligned}
F_{X|Y}(x|y) &= P(X \leq x|Y = y) = \frac{P(X \leq x, Y = y)}{P(Y = y)} \\
f_{X|Y}(x|y) &= P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \\
&= \frac{f_{X,Y}(x, y)}{f_Y(y)}
\end{aligned}
$$

Folgerungen

Es gilt immer:

$$
\begin{aligned}
f_{X,Y}(x, y) &= f_{X|Y}(x|y) \cdot f_Y(y) \\
&= f_{Y|X}(y|x) \cdot f_X(x)
\end{aligned}
$$

Es folgt: X und Y sind genau dann unabhängig wenn

$$
\begin{aligned}
f_{X|Y}(x|y) &= f_X(x) \\
\text{oder} \quad f_{Y|X}(y|x) &= f_Y(y)
\end{aligned}
$$

für alle x und y gilt.

Beispiel

$f_{X,Y}(x, y)$	$Y = -1$	$Y = 0$	$Y = 2$
$X = 1$	$\frac{1}{18}$	$\frac{3}{18}$	$\frac{2}{18}$
$X = 2$	$\frac{2}{18}$	0	$\frac{3}{18}$
$X = 3$	0	$\frac{4}{18}$	$\frac{3}{18}$

Man berechne die Randverteilungen $f_X(x)$ und $f_Y(y)$, die bedingten Verteilungen $f_{X|Y}(x|y)$ und $f_{Y|X}(y|x)$ und untersuche X und Y auf Unabhängigkeit.

Beispiel II

- Betrachte zwei unabhängige ZV X und Y, die beide Poisson-verteilt sind mit Parameter λ bzw. μ.
- Definiere $Z = X + Y$.
- Man zeige: Die bedingte Verteilung von $X|Z = z$ ist binomial mit Parametern $n = z$ und $\pi = \lambda/(\lambda + \mu)$:

$$X|Z = z \sim \mathcal{B}(z, \pi = \lambda/(\lambda + \mu))$$

Bedingte Verteilungen von stetige ZVn

Betrachte die Zufallsvariablen X und Y mit gemeinsamer Dichte $f_{X,Y}(x,y)$. Wir interessieren uns für die bedingte Verteilung von X gegeben $Y = y$.

Problem: Es gilt $P(Y = y) = 0$ für alle y. Daher ist

$$P(X \leq x | Y = y) = \frac{P(X \leq x \quad \text{und} \quad Y = y)}{P(Y = y)}$$

nicht definiert.

Bedingte Verteilungs- und Dichtefunktion

Man geht nun anders vor und betrachtet

$$\begin{aligned}
& P(X \leq x | y \leq Y \leq y + dy) \\
= & \frac{P(X \leq x \quad \text{und} \quad y \leq Y \leq y + \delta)}{P(y \leq Y \leq y + dy)} \\
\approx & \frac{\int_{-\infty}^{x} f_{X,Y}(u,y)\, \delta \, du}{f_Y(y)\, \delta} \\
= & \int_{-\infty}^{x} \underbrace{\frac{f_{X,Y}(u,y)}{f_Y(y)}}_{(*)} du
\end{aligned}$$

$(*)$ Dichtefunktion der bedingten Verteilung von X geg. $Y = y$

Bedingte Verteilungs- und Dichtefunktion II

Man definiert nun:

Die **bedingte Verteilungsfunktion** von X, gegeben $Y = y$ ist

$$F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f_{X,Y}(u,y)}{f_Y(y)}\, du$$

für alle y mit $f_Y(y) > 0$. Die **bedingte Dichte** von X, gegeben $Y = y$ ist somit

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Beispiel

Betrachten wir wieder die Zufallsvariablen X und Y mit gemeinsamer Dichte

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{x} & \text{für} \quad 0 \leq y \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Für die bedingten Dichten von Y, gegeben $X = x$ ergibt sich:

$$f_{Y|X}(y|x) \quad = \quad \begin{cases} \frac{1}{x} & \text{für} \quad 0 \leq y \leq x \\ 0 & \text{sonst} \end{cases}$$

d.h. $Y|X = x$ ist gleichverteilt auf $[0, x]$.

Beispiel II

Für die bedingten Dichten von X, gegeben $Y = y$ ergibt sich:

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{\frac{1}{x}}{\log(\frac{1}{y})} \quad \text{für} \quad y \leq x \leq 1 \\ &= \begin{cases} -1/(x\log(y)) & \text{für} \quad y \leq x \leq 1 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

Simulation über bedingte Verteilungen

Bedingte Verteilungen sind sehr nützlich zum Simulieren aus gemeinsamen Verteilungen. Wegen

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) \cdot f_Y(y) \tag{1}$$

kann man zunächst eine Zufallsvariable $Y = y$ aus der Randverteilung $f_Y(y)$ ziehen, und dann bedingt auf $Y = y$ eine Zufallszahl aus der bedingten Verteilung $f_{X|Y}(x|y)$ ziehen. Oder andersherum:

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \cdot f_X(x) \tag{2}$$

Im Beispiel ist Version (2) einfacher zu implementieren.

Beispiel: Standardnormalverteilung

Angenommen X und Y sind bivariat standardnormalverteilt. Dann ist

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{\frac{1}{2\pi}\frac{1}{\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}\frac{1}{(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right)}{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}y^2\right)} \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}\frac{(x-\rho y)^2}{(1-\rho^2)}\right) \end{aligned}$$

also $X|Y = y \sim N(\rho \cdot y, 1-\rho^2)$

Analog erhält man $Y|X = x \sim N(\rho \cdot x, 1-\rho^2)$
$\rightarrow$ Simulation aus der bivariaten Standardnormalverteilung

Verteilung einer diskreten und einer stetigen ZV

Das Konzept von gemeinsamer und bedingter Verteilung lässt sich problemlos auch auf zwei Zufallsvariablen verallgemeinern, von denen eine diskret und eine stetig ist.
Wir illustrieren dies hier an einem Beispiel:

$X \sim \mathcal{B}e(\alpha, \beta)$

$Y|X \sim \mathcal{B}(n, \pi = X)$

Beispiel: Die gemeinsame Verteilung

Die gemeinsame Verteilung von X und Y ist

$$\begin{aligned} f(x,y) &= f(y|x) \cdot f(x) \\ &= \binom{n}{y} x^y (1-x)^{n-y} \cdot \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1} \\ &= \frac{1}{B(\alpha,\beta)} \binom{n}{y} x^{y+\alpha-1} (1-x)^{n-y+\beta-1} \end{aligned}$$

für $x \in [0,1]$ und $y \in \{0, 1, \ldots, n\}$.

Beispiel: Die bedingte Verteilung von $X|Y = y$

Für die bedingte Dichte $f(x|y)$ folgt:

$$f(x|y) = \frac{f(x,y)}{f(y)} \overset{(*)}{\propto} x^{y+\alpha-1}(1-x)^{n-y+\beta-1}$$

also: $X|Y = y \sim \mathcal{B}e(\alpha + y, \beta + n - y)$
Bei $(*)$ haben wir ausgenützt, dass der Nenner $f(y)$ in

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

nicht von x abhängt, also für $Y = y$ konstant ist.

Beispiel: Die Randverteilung von Y

Damit folgt für $f(y) = f(x,y)/f(x|y)$:

$$\begin{aligned} f(y) &= \binom{n}{y} \frac{B(y+\alpha, n-y+\beta)}{B(\alpha,\beta)} \\ &= \underbrace{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+n)}}_{\text{hängt nicht von y ab}} \binom{n}{y} \Gamma(\alpha+y)\Gamma(\beta+n-y) \end{aligned}$$

für $y = 0, \ldots, n$.
Diese Verteilung heißt "**Beta-Binomialverteilung**" und hat die Parameter α und β: $X \sim \mathcal{BB}e(\alpha, \beta)$

Allgemeine Zufallsvektoren

Allgemeiner kann man natürlich auch einen Zufallsvektor $\mathbf{X} = (X_1, \ldots, X_n)$ der Dimension n betrachten.

Dieser hat dann die Wahrscheinlichkeitsfunktion

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

und die Randverteilungen

$$f_{X_i}(x_i) = \sum_{x_j : j \neq i} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

Die Trinomialverteilung

Ein Experiment, bei dem ein von drei möglichen Ereignissen mit Wahrscheinlichkeit π_1, π_2 und π_3 $(\pi_1 + \pi_2 + \pi_3 = 1)$ auftritt, wird unabhängig voneinander n-mal wiederholt. Sei $\mathbf{X}$ ein drei-dimensionaler Zufallsvektor, dessen i-te Komponente angibt, wie oft das i-te Ereignis eingetreten ist.

Beispiel:
In einer Population mit Häufigkeiten π_1, π_2 und π_3 der Genotypen *aa*, *ab* und *bb* wird eine Stichprobe vom Umfang n gezogen. Die Anzahlen X_1, X_2 und X_3 der drei Genotypen ist dann **trinomialverteilt**.

Die Trinomialverteilung II

Ein drei-dimensionaler diskreter Zufallsvektor $\mathbf{X}$ heißt **trinomialverteilt**, falls er Träger

$$\mathcal{T} = \{\mathbf{x} = (x_1, x_2, x_3) : x_i \in \{0, 1, \ldots, n\} \text{ und } x_1 + x_2 + x_3 = n\}$$

und Wahrscheinlichkeitsfunktion

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \frac{n!}{x_1! x_2! x_3!} \pi_1^{x_1} \pi_2^{x_2} \pi_3^{x_3}$$

besitzt.

Man schreibt kurz: $\mathbf{X} \sim \mathcal{M}_3(n, \boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3))$
Hierbei steht $\mathcal{M}_3$ für **M**ultinomialverteilung der Dimension 3.

Man kann zeigen, dass für die Randverteilungen gilt: $X_i \sim \mathcal{B}(n, \pi_i)$

Transformationsregel für $E(g(X, Y))$

Seien X und Y zwei ZV mit gemeinsamer Wahrscheinlichkeitsfunktion $f_{XY}(x, y)$. Sei $g(x, y)$ eine reellwertige Funktion.
Dann gilt für $Z = g(X, Y)$

$$\mathrm{E}(Z) = \mathrm{E}(g(X, Y)) = \sum_x \sum_y g(x, y)\, f_{X,Y}(x, y).$$

falls X und Y diskret und analog für X und Y stetig:

$$\mathrm{E}(Z) = \mathrm{E}(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)\, f_{X,Y}(x, y) dx dy.$$

Speziell gilt daher:

$$\mathrm{E}(X \cdot Y) = \sum_x \sum_y x \cdot y \cdot f_{X,Y}(x, y)$$

bzw.

$$\mathrm{E}(X \cdot Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f_{X,Y}(x, y) dx dy.$$

4.6 Kovarianzen und Korrelationen

Kovarianzen und Korrelationen – Einleitung

Als Maße für die lineare stochastische Abhängigkeit von zwei diskreten oder zwei stetigen ZVn X und Y definiert man die **Kovarianz**

$$\mathrm{Cov}(X,Y) = \mathrm{E}[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))]$$

und die **Korrelation**

$$\rho = \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)} \cdot \sqrt{\mathrm{Var}(Y)}}$$

unter der Voraussetzung, dass $\mathrm{Var}(X) > 0$ und $\mathrm{Var}(Y) > 0$ gilt.
Beachte: $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

Kovarianzen und Korrelationen II

Seien (X_i, Y_i) $(i = 1, \ldots, n)$ iid Zufallsvariablen aus Verteilungen mit Erwartungswerten μ_X, μ_Y und existierenden Varianzen. Dann gilt für

$$S^2_{XY} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_X)(Y_i - \mu_Y),$$

dass $E(S^2_{XY}) = \mathrm{Cov}(X,Y)$ und für $n \to \infty$: $S^2_{XY} \to \mathrm{Cov}(X,Y)$

S^2_{XY} wird **empirische Kovarianz** genannt.

Der Verschiebungssatz für die Kovarianz

Es gilt zudem:

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\cdot\mathrm{E}(Y)$$

Beachte:
$\mathrm{E}(XY)$ kann mit Transformationssatz für Erwartungswerte leicht über die gemeinsame Wahrscheinlichkeitsfunktion $f_{XY}(x,y)$ von X und Y berechnet werden.

Beispiel "revisited" – diskreter Fall

$f_{X,Y}(x,y)$	$Y=-1$	$Y=0$	$Y=2$	$f_X(x)$
$X=1$	$\frac{1}{18}$	$\frac{3}{18}$	$\frac{2}{18}$	$\frac{6}{18}$
$X=2$	$\frac{2}{18}$	0	$\frac{3}{18}$	$\frac{5}{18}$
$X=3$	0	$\frac{4}{18}$	$\frac{3}{18}$	$\frac{7}{18}$
$f_Y(y)$	$\frac{3}{18}$	$\frac{7}{18}$	$\frac{8}{18}$	

Es ergibt sich:

$$\left.\begin{aligned} \mathrm{E}(XY) &= \tfrac{29}{18} \\ \mathrm{E}(X) &= \tfrac{37}{18} \\ \mathrm{E}(Y) &= \tfrac{13}{18} \end{aligned}\right\} \quad \begin{aligned} \mathrm{Cov}(X,Y) &= \tfrac{29}{18} - \tfrac{37}{18}\cdot\tfrac{13}{18} = \tfrac{41}{324} \\ \rho &= \tfrac{41}{\sqrt{107413}} = 0.125 \end{aligned}$$

Beispiel II – stetiger Fall

Betrachte

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{x} & \text{für} \quad 0 \le y \le x \le 1 \\ 0 & \text{sonst} \end{cases}$$

Die Randverteilungen von X und Y ergeben sich zu

$$\begin{aligned} f_X(x) &= \int_0^x \frac{1}{x}\,dy = 1 && \text{für} \quad 0 \le x \le 1 \\ f_Y(y) &= \int_y^1 \frac{1}{x}\,dx = \log(1/y) && \text{für} \quad 0 \le y \le 1 \end{aligned}$$

Beispiel II – stetiger Fall

Man überprüft leicht, dass $\int_0^1 f(x)\,dx = 1$ und $\int_0^1 f(y)\,dy = 1$ und daher

$$\int\int f(x,y)\,dy\,dx = 1$$

Die Korrelation zwischen X und Y ergibt sich zu $\rho(X,Y) \approx 0.65$.

(Details in Vorlesung)

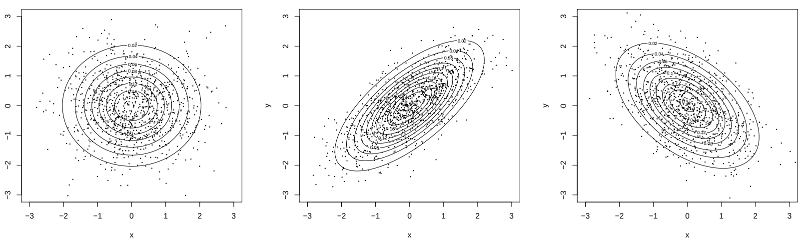
Die bivariate Standardnormalverteilung

Die bivariate ("zweidimensionale") Standardnormalverteilung mit Parameter ρ ($|\rho| < 1$) hat die Dichtefunktion

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right)$$

- Die Randverteilungen von X und Y sind (für jedes ρ) standard-normalverteilt.
- Die Korrelation zwischen X und Y ist gleich ρ.
- Aus Unkorreliertheit von X und Y folgt hier auch die Unabhängigkeit von X und Y.

Höhenlinien der bivariaten Standardnormalverteilung



Höhenlinien der Dichtefunktion der bivariaten Standardnormalverteilung mit jeweils 500 Stichproben für $\rho = 0$ (links), $\rho = 0.7$ (Mitte) und $\rho = -0.5$ (rechts)

Die bivariate Normalverteilung

Die allgemeine bivariate Normalverteilung erhält man durch die linearen Transformationen einer bivariaten Standardnormalverteilung:

$$\begin{aligned} X &\rightarrow \mu_X + \sigma_X \cdot X \\ Y &\rightarrow \mu_Y + \sigma_Y \cdot Y \end{aligned}$$

Insgesamt fünf Parameter: $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho$

Unkorreliertheit

X und Y heißen **unkorreliert**, wenn

$$\text{Cov}(X, Y) = 0 \quad \text{bzw.} \quad \rho(X, Y) = 0$$

d.h. wenn

$$\text{E}(X \cdot Y) = \text{E}(X) \cdot \text{E}(Y)$$

gilt.

Beachte: Aus Unabhängigkeit folgt Unkorreliertheit aber der Umkehrschluss gilt im Allgemeinen nicht!

Man sagt:

X und Y sind **positiv/negativ korreliert** falls

$$\rho(X, Y) > 0 \quad \text{bzw.} \quad \rho(X, Y) < 0$$

Beispiel

Seien $X \sim \mathcal{B}(\pi = \frac{1}{2})$ und $Y \sim \mathcal{B}(\pi = \frac{1}{2})$ unabhängig.
Betrachte

$$Z_1 = X + Y = \begin{cases} 0 & \text{mit Wkeit } \frac{1}{4} \\ 1 & \text{mit Wkeit } \frac{1}{2} \\ 2 & \text{mit Wkeit } \frac{1}{4} \end{cases}$$

$$Z_2 = X - Y = \begin{cases} -1 & \text{mit Wkeit } \frac{1}{4} \\ 0 & \text{mit Wkeit } \frac{1}{2} \\ 1 & \text{mit Wkeit } \frac{1}{4} \end{cases}$$

Dann sind Z_1 und Z_2 zwar unkorreliert aber nicht unabhängig!

Eigenschaften von Korrelationen

Für alle ZVn X und Y gilt:

$$-1 \leq \rho(X, Y) \leq 1$$

$|\rho(X, Y)| = 1$ gilt genau dann, wenn perfekte lineare Abhängigkeit zwischen X und Y besteht:

$$Y = a + b \cdot X \qquad \text{für bestimmte } a, b \in \mathbb{R} \text{ mit } b \neq 0$$

Lineare Transformationen

Seien $a, b, c, d \in \mathbb{R}$ mit $b \cdot d > 0$ und X, Y beliebige ZVn. Dann gilt:

$$\text{Cov}(a + bX, c + dY) = b \cdot d \cdot \text{Cov}(X, Y)$$

Daher gilt:

$$\rho(a + bX, c + dY) \quad = \quad \rho(X, Y)$$

d.h. die Korrelation ist **invariant** bzgl. linearer Transformationen

Die Varianz der Summe von zwei ZVn

Seien X und Y beliebige ZVn. Dann gilt für $X + Y$:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y)$$

Daher gilt speziell für unabhängige X und Y:

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

4.7 Quantile

Quantile – diskrete ZVn

Sei X eine diskrete ZV mit Verteilungsfunktion $F(x)$, $x \in \mathbb{R}$.

Sei $p \in [0, 1]$. Jeder Wert x, für den

$$\begin{aligned} P(X \leq x) &\geq p \text{ und} \\ P(X \geq x) &\geq 1 - p \end{aligned}$$

gilt, heisst p-**Quantil** x_p der Verteilung von X.

Problem: Definition nicht immer eindeutig.

Quantile – diskrete ZVn II

Um Eindeutigkeit zu erreichen, definiert man:

Das p-**Quantil** x_p der Verteilung von X ist definiert als der **kleinste** Wert x für den $F(x) \geq p$ gilt.

Somit gilt $P(X \leq x) = F(x_p) \geq p$ und daher "$x_p = F^{-1}(p)$" ("Inversion der Verteilungsfunktion").

Speziell nennt man das 0.5-Quantil den **Median** x_{med} der Verteilung.

Quantile – stetige ZVn

Wir nehmen an, dass der Träger der stetigen Zufallsvariable X ein Intervall ist. Somit ist die Umkehrfunktion $F^{-1}(p)$ der Verteilungsfunktion $F(x)$ von X eindeutig definiert.
Das **p-Quantil** der Verteilung von X ist definiert als der Wert x_p für den $F(x) = p$ gilt. Somit gilt $x_p = F^{-1}(p)$. Ist $f(x)$ symmetrisch um einen

Punkt c, so ist $x_{Med} = c$.

Beispiel: Bei einer normalverteilten Zufallsvariable $X \sim \mathcal{N}(\mu, \sigma^2)$ ist
$x_{Med} = \mu$.

Interpretation vom Median und von Quantilen

Beim Median x_{med} gilt:

$$\begin{aligned} P(X \leq x_{\text{med}}) &\geq 0.5 \\ P(X \geq x_{\text{med}}) &\geq 0.5 \end{aligned}$$

Allgemeiner gilt für das p-Quantil x_p:

$$\begin{aligned} P(X \leq x_p) &\geq p \\ P(X \geq x_p) &\geq 1 - p \end{aligned}$$

Numerische Bestimmung von Quantilen

Bei Kenntnis der Verteilungsfunktion $F(x)$ und endlichem Träger $\mathcal{T}$:

1. Berechne $F(x)$ für alle $x \in \mathcal{T}$.
2. Dann gilt für das p-Quantil x_p:

$$x_p = \min\{x : F(x) \geq p\}$$

3. in **R**: verwende Funktion `which`

Nicht so einfach, siehe Dokumentation `?quantile`.

Beispiel: geometrische Verteilung

Funktion `qgeom()` berechnet Quantilfunktion der geometrischen Verteilung. *Achtung R berechnet die Variante von Folie 13!*

0.95-Quantil $x_{0.95}$: In 95% aller Fälle muss man maximal $x_{0.95}$ Versuche unternehmen, *bevor* zum ersten mal A eintritt.

π	$x_{0.95}$	x_{med}
0.01	298	68
0.1	28	6
0.5	4	0
0.9	1	0
0.99	0	0

4.8 Faltungen

Faltungen – Einleitung

Sind X und Y unabhängige ZV mit Wahrscheinlichkeitsfunktionen $f_X(x)$ und $f_Y(y)$, so gilt für die Summe $Z = X + Y$:

$$\begin{aligned}
P(X+Y=z) &= \sum_x P(X=x, x+Y=z) \\
&= \sum_x P(X=x, Y=z-x) \\
&\overset{\text{unabh.}}{=} \sum_x P(X=x)\cdot P(Y=z-x) \\
&= \sum_x f_X(x)\cdot f_Y(z-x)
\end{aligned}$$

Faltungen II

Man nennt die Wahrscheinlichkeitsverteilung von $Z = X + Y$

$$P(X+Y=z) = \sum_x f_X(x) \cdot f_Y(z-x) = \sum_y f_X(z-y) \cdot f_Y(y)$$

die **Faltung** von X und Y.

Beispiel:
Ist $X \sim \mathcal{P}(\lambda_1)$ und $Y \sim \mathcal{P}(\lambda_2)$ unabhängig, so ist die Faltung von X und Y wieder Poisson-verteilt mit Parameter $\lambda_1 + \lambda_2$:

$$X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$$

Faltung von zwei geometrischen ZV

Seien $X \sim \mathcal{G}(\pi)$ und $Y \sim \mathcal{G}(\pi)$ unabhängig.

Berechne Träger und Wahrscheinlichkeitsfunktion der Summe $Z = X + Y$.

Wie kann man Z interpretieren?

Die negative Binomialverteilung

Betrachte die Summe von n unabhängigen geometrischen ZV $X_1, \ldots, X_n$:
$X = X_1 + \ldots + X_n$.

Dann hat X eine **negative Binomialverteilung** mit Parameter $n \in \mathbb{N}$ und $\pi \in (0,1)$ und Wahrscheinlichkeitsfunktion

$$f(x) = \binom{x-1}{n-1} \pi^n (1-\pi)^{x-n} \text{ für } x = n, n+1, \ldots$$

Funktionen in R:
`dnbinom(...)`, `pnbinom(...)`, `qnbinom(...)`, `rnbinom(...)`

Beachte: Unterschiedliche Definition in R! Träger immer gleich $\mathbb{N}_0$