
Mathematik für Biologen 2

Malte Braack

Institut für Angewandte Mathematik, Universität Heidelberg

SoSe 2006

July 18, 2006

Vorlesung

Inhalt:

- Wahrscheinlichkeitsrechnung (Fortsetzung)
 - ▶ Zufallsvariablen (diskrete, stetige)
 - ▶ Erwartungswert, Varianz
 - ▶ Multivariate Zufallsvariablen
- Beurteilende Statistik
 - ▶ Schätzung von Wahrscheinlichkeiten
 - ▶ Konfidenzintervalle
 - ▶ Parametertests / Testen von Hypothesen
 - ▶ Korrelation und Regression
 - ▶ χ^2 -Unabhängigkeitstests
 - ▶ Anpassungstests
- Gewöhnliche Differentialgleichungen
- Funktionen mehrerer Veränderlicher

Übungen

- Teilnahmepflicht und Anwesenheitspflicht (Scheinkriterium)
- in Gruppen mit ca. 30 Personen
- Es stehen 2 Termine zur Verfügung:
Mo 11-13, Mo 14-16
- Ausgabe und Abgabe: jeweils Mittwochs vor der Vorlesung
- Bearbeitung in 1-3er Gruppen
- Korrektur durch die Tutoren (Reklamationen auch dort)
- Rückgabe in den Gruppenübungen
- Zulassungskriterium Klausur: 50% der Punkte und mündl. Teilnahme

Klausur

- Termin: Mi. 26. Juli 2006, 11:00 Uhr, INF 230, gHS
- Hier **keine** Gruppenarbeit.
- Personalausweis **und** Studentenausweis mitbringen.
- Wer gut bei den Übungen mitgearbeitet hat wird profitieren.

Literatur

- 1 E. Batschelet: Mathematik für Biologen, Springer Verlag, 1980.
- 2 H. Vogt: Grundkurs Mathematik für Biologen, Teubner, 1983.
- 3 A. Riede: Mathematik für Biologen.
- 4 H. Behncke: A. Riede: Mathematik für Biologen I, Vorlesungsskript
<http://www.mathematik.uni-osnabrueck.de/lehre/bio04/>.

Folien im Anschluß an die jeweilige Vorlesungsstunde:

<http://numerik.uni-hd.de/~braack/bio2>

Gruppe A

Catalina Filler, Mo 11-13 Uhr, INF 294, Rm -102

- Auer, Thomas
- Becker, Marco
- Gross, Marion-Jana
- Faltermann, Susanne
- Hangel, Christian
- Hruschka, Stefan
- Huber, Monika
- Jürgens, Maike Christine
- Kern Christian
- Kirsch, Patrick
- Landau, Anja
- Merkel, Mirko
- Mitterer, Christian
- Risch, Sophia
- Schütz, Christian
- Stauffer, Sarah
- Strompen, Jennifer
- Vernaldi, Saskia
- Weißer, Melanie
- Wolf, Jana

Gruppe B

Steffen Haschler, Mo 11-13 Uhr, INF 293, SR 215

- Albrecht, Christian
- Bohle, Dmitrij
- Brozy, Johannes
- Gärtner, Linda
- Goldschmidt, Melanie
- Herrmann, Sascha
- Evrard, Lionel
- Fülling, Christine
- Krannich, Jörg
- Lauinger, Linda
- Lee, Jennifer
- Mewes, Vassilios
- Okbay, Doruk
- Radiykov, Danijela
- Rechl, Max
- Vardeh, David
- Schad, Jan-Ulrich
- Schmitt, Axel
- Zugschwerdt, Petra

Gruppe C

Helke Hesse, Mo 14-16 Uhr, INF 293, SR 215

- Bernecker, Conny
- Diehl, Ulrike
- Dietz, Gerhard
- Egarter, Saskia
- Feger, Franziska Klara
- Hick, Meike
- Jöst, Hanna
- Kornfeld, Jörgen
- Krüger, Tanja
- Luzón, Javier
- Madzharova, Violeta
- Michling, Florian
- Rosenbaum-Feldbrügge, Kaja
- Senusta, Azize-Pinar
- Seyton, Sam
- Suara, Fabian
- Tietze, Daniela
- Weber, Dominik
- Weidner, Kerstin

Kap. 9: Wahrscheinlichkeitsrechnung (Fortsetzung)

Zufallsvariablen

Def.: Unter einer **Zufallsvariablen** X verstehen wir eine Abbildung $X : \Omega \rightarrow \mathbb{R}$, wobei $X(\omega)$ das sogenannte Merkmal ist. Wenn X nur diskrete Werte annehmen kann, so heißt $X : \Omega \rightarrow \{x_1, \dots, x_m\}$ **diskrete Zufallsvariable**.

Notation: $P(X = k)$ ist die Wahrscheinlichkeit, dass X den Wert $k \in \mathbb{R}$ annimmt.

Beispiel: 2-maliges Würfeln, $X : \Omega \rightarrow \{2, 3, \dots, 12\}$ beschreibt die Summe der Augenzahlen.

$$P(X = 3) = \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{18}$$

Beispiel: Die Körpergröße ist (streng genommen) keine *diskrete* Zufallsvariable.

Wahrscheinlichkeitsverteilung

Def.: Die Funktion $f : \mathbb{R} \rightarrow [0, 1]$, $f(x_k) = P(X = x_k)$ einer diskreten Zufallsvariablen X heißt **Wahrscheinlichkeitsverteilung** oder **Wahrscheinlichkeitsfunktion**.

Bem. Bei nicht diskreten Zufallsvariablen heisst die Wahrscheinlichkeitsverteilung auch **Wahrscheinlichkeitsdichte**.

Bem. Aufgrund der Normierung der Wahrscheinlichkeit gilt stets

$$\sum_{i=1}^m f(x_i) = 1.$$

Beispiel: 2-maliges Würfeln $\Omega = \{1, \dots, 6\}^2$. Zufallsvariable Augensumme:

$$\begin{aligned} f(2) &= P(X = 2) = \frac{1}{6^2}, & f(3) &= P(X = 3) = \frac{1}{18} \\ f(4) &= P(X = 4) = \frac{3}{36}, \dots & f(12) &= P(X = 12) = \frac{1}{36} \end{aligned}$$

Verteilungsfunktion

Def.: Ist f die Wahrscheinlichkeitsdichte einer diskreten Zufallsvariablen X , und $x_1 < \dots, x_m$, so ist die **Verteilungsfunktion** von X die Funktion

$$F : \{x_1, \dots, x_m\} \rightarrow [0, 1], \quad F(x_k) = \sum_{i=1}^k f(x_i).$$

$F(x_k)$ ist die Wahrscheinlichkeit, dass X mind. den Wert x_k annimmt.

Bem.

- F ist monoton wachsend, d.h. $F(x_{i+1}) \geq F(x_i)$ für $i = 1, \dots, m - 1$,
- $F(x_m) = 1$.

Bernoulli-Versuche

- Unter einem **Bernoulli-Versuch** oder **Bernoulli-Experiment** versteht man einen Zufallsversuch mit genau zwei möglichen Ergebnissen, “Erfolg” und “Mißerfolg”.
- Die **Bernoulli-Verteilung** ist die zugehörige Wahrscheinlichkeitsverteilung $P(X_{Bern})$ eines Bernoulli-Versuchs:

$$f(k) = P(X_{Bern} = k) = p^k(1 - p)^{1-k}$$

- Führt man einen Bernoulli-Versuch n -mal unter gleichen Bedingungen aus, so spricht man von einem **Bernoulli-Prozess** oder **Bernoulli-Kette**.
- Die Wahrscheinlichkeitsverteilung zur Wahrscheinlichkeitsvariablen X_{Bin} , die die Anzahl von Erfolgen summiert, ist die **Binomial-Verteilung**:

$$f(k) = P(X_{Bin} = k) = \binom{n}{k} p^k(1 - p)^{n-k}$$

Verteilungsfunktion einer Bernoulli-Kette

Für die Verteilungsfunktion eines Bernoulli-Prozesses gilt tatsächlich ($q = 1 - p$):

$$\begin{aligned} F(x_m) &= \sum_{i=1}^m f(x_i) = \sum_{k=0}^n f(k) \\ &= \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=0}^n (p + q)^n = 1^n = 1 \quad (\text{Binomischer Lehrsatz}) \end{aligned}$$

Erwartungswert

Analog zum arithmetischen Mittelwert bei Stichproben x_1, \dots, x_n mit Ausprägungen $a_1, \dots, a_m \in \mathbb{R}$,

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j = \sum_{i=1}^m h_i a_i \quad (h_i = \text{re. Häufigkeit zu Ausprägung } a_i)$$

betrachten wir:

Def.: Erwartungswert

$$E(X) = \sum_{i=1}^m P(X = x_i) x_i = \sum_{i=1}^m f(x_i) x_i$$

Der Erwartungswert wird auch häufig mit μ bezeichnet.

Bsp.: 1-maliges Würfeln

$$E(X) = \sum_{i=1}^m \frac{1}{6} i = 3.5$$

Geometrische Verteilung

Bernoulli-Prozess mit der Einzelwahrscheinlichkeit p für “Erfolg” und $q = 1 - p$ für “Mißerfolg”.

Frage: Wie lange dauert es bis man zum ersten mal Erfolg hat?

Zufallsvariable $X : \Omega \rightarrow \mathbb{N}$ beschreibe dies. z.B. $X((0, 0, 0, 1, \dots)) = 4$.

$$P(X_{geo} = k) = q^{k-1}p$$

Diese Wahrscheinlichkeitsverteilung heißt **geometrische Verteilung**.

Die geometrische Verteilung kommt häufig bei Warteprozessen vor.

Erwartungswert der geometrischen Verteilung

$$E(X_{geo}) = \sum_{k=1}^{\infty} P(X_{geo} = k)k = p \sum_{k=0}^{\infty} q^{k-1}k$$

Trick: $g_k(q) = q^k$, $g'_k(q) = kq^{k-1}$,

$$E(X_{geo}) = p \sum_{k=1}^{\infty} g'_k(q) = p \frac{\partial}{\partial q} \sum_{k=0}^{\infty} g_k(q)$$

Aufgrund der geometrischen Reihe gilt $\sum_{k=0}^{\infty} g_k(q) = 1/(1-q)$, also

$$E(X_{geo}) = p \frac{\partial}{\partial q} \left(\frac{1}{1-q} \right) = p \frac{1}{(1-q)^2} = \frac{1}{p}$$

Bsp.: Man muß im Schnitt 6-mal würfeln, um einen "Sechser" zu bekommen.

Varianz

Analog zur **Varianz** als Streuungsmaß bei Stichproben

$$\text{Var}(x) = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2 = \sum_{i=1}^m (a_i - \bar{x})^2 h_i$$

definieren wir:

Def.: **Varianz** einer diskreten Zufallsvariablen X als

$$\text{Var}(X) = \sum_{i=1}^m (x_i - E(X))^2 P(X = x_i)$$

Die **Standardabweichung** von X ist dann $\sigma(X) = \sqrt{\text{Var}(x)}$.

Summen und Produkte von Zufallsvariablen

Sind X, Y Zufallsvariablen und $c \in \mathbb{R}$, so versteht man unter $X + Y$, XY und cX ebenfalls Zufallsvariablen, z.B.

$$XY : \Omega \rightarrow \mathbb{R} : XY(\omega) = X(\omega) \cdot Y(\omega)$$

Es gilt:

- (i) $E(cX) = cE(X)$
- (ii) $E(X + Y) = E(X) + E(Y)$
- (iii) $Var(cX) = c^2 Var(X)$

Bei unabhängigen Zufallsvariablen X, Y gilt zudem:

- (iv) $E(XY) = E(X) \cdot E(Y)$
- (v) $Var(X + Y) = Var(X) + Var(Y)$

Beispiel zu Produkten von Zufallsvariablen

Szenario: Würfeln mit zwei Würfeln. X bezeichne die Augenzahl des ersten und Y die Augenzahl des zweiten Würfels. Da X und Y voneinander unabhängig sind, gilt für das Produkt der beiden Augenzahlen

$$E(XY) = 3,5^2 = 12,25 = E(X) \cdot E(Y)$$

Dagegen gilt für das Produkt der **ersten** Augenzahl:

$$\begin{aligned} E(X^2) &= \sum_{k=1}^m P(X = k)k^2 = \frac{1}{6}(1 + 2 + 4 + \dots + 36) = \frac{91}{6} \\ &\neq E(X)^2 = 12,25 \end{aligned}$$

Verschiebungssatz der Varianz

Es gilt:

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

Beweis: Mit $\mu = E(X)$ gilt:

$$\begin{aligned}\text{Var}(X) &= \sum_{i=1}^m (x_i - \mu)^2 P(X = x_i) \\ &= \sum_{i=1}^m (x_i^2 - 2x_i\mu + \mu^2) P(X = x_i) \\ &= \sum_{i=1}^m x_i^2 P(X = x_i) - 2\mu \sum_{i=1}^m x_i P(X = x_i) + \mu^2 \sum_{i=1}^m P(X = x_i) \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - E(X)^2\end{aligned}$$

Spezielle Varianzen

- 1 Die Varianz der Bernoulli-Verteilung folgt aus dem Verschiebungssatz

$$\begin{aligned} \text{Var}(X_{\text{Bern}}) &= E(X_{\text{Bern}}^2) - E(X_{\text{Bern}})^2 \\ &= E(X_{\text{Bern}}) - E(X_{\text{Bern}})^2 = p - p^2 = pq \end{aligned}$$

da $X_{\text{Bern}} = X_{\text{Bern}}^2$.

- 2 Die Varianz der Binomialverteilung ist

$$\text{Var}(X_{\text{Bin}}) = npq$$

- 3 Für die geometrischen Verteilung gilt:

$$\text{Var}(X_{\text{geo}}) = \frac{1-p}{p^2}$$

Binomialverteilung für $n \rightarrow \infty$

Szenario: Der radioaktive Zerfall von Atomen könnte durch die Binomialverteilung beschrieben werden: 1=Zerfall.

Selbst wenn man die Anzahl n an Atomen kennen würde, so läßt sich die Binomialverteilung aufgrund der Größe von n nicht auswerten.

$$P(X_{Bin}=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Frage: Was passiert im Grenzübergang $n \rightarrow \infty$, $p \rightarrow 0$?

Für $m = pn$ gilt:

$$\begin{aligned} P(X_{Bin}=k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{m^k}{k!} \underbrace{\frac{n!}{(n-k)! n^k q^k}}_{\rightarrow 1} \underbrace{\left(1 - \frac{m}{n}\right)^n}_{\rightarrow \exp(-m)} \rightarrow \frac{m^k}{k!} \exp(-m) \end{aligned}$$

Dieser Grenzwert wird **Poisson-Verteilung** genannt.

Poisson-Verteilung

Def.: Die **Poisson-Verteilung** zum Parameter $m = pn$ ist gegeben durch

$$P(X_{Poi} = k) = \frac{m^k}{k!} e^{-m}$$

Der Parameter m ist gerade der Erwartungswert der Binomialverteilung,
 $E(X_{Bin}) = m = pn$

Zufällig verteilte Objekte sind Poisson verteilt.

Nebenrechnung

1 Es gilt

$$\lim_{x \rightarrow \infty} \left(1 - \frac{1}{x}\right)^{-x} = \lim_{x \rightarrow \infty} \left(\frac{x}{x-1}\right)^x = e$$

Mit $x = m/n$ folgt daher

$$\lim_{n \rightarrow \infty} \left(1 - \frac{m}{n}\right)^n \rightarrow e^{-m}$$

2

$$\begin{aligned} \frac{n!}{(n-k)! n^k q^k} &= \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{n^k \left(1 - \frac{m}{n}\right)^k} = \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{(n-m)^k} \\ &= \frac{n^k + \dots + n^{k-1} + \dots + \dots + n}{n^k + \dots + n^{k-1} + \dots} = \frac{1 + \dots + n^{-1} + \dots + \dots + n^{-k}}{1 + \dots + n^{-1} + \dots} \\ &\rightarrow 1 \quad (n \rightarrow \infty) \end{aligned}$$

Erwartungswert der Poisson-Verteilung

Die Erwartungswerte der Binomialverteilung und der Poisson-Verteilung sind identisch:

$$E(X_{Poi}) = E(X_{Bin}) = m = pn$$

Denn:

$$\begin{aligned} E(X_{Poi}) &= \sum_{k=1}^{\infty} P(X_{Poi} = k)k = \sum_{k=1}^{\infty} \frac{m^k}{(k-1)!} e^{-m} \\ &= me^{-m} \sum_{k=1}^{\infty} \frac{m^{k-1}}{(k-1)!} = me^{-m} \sum_{k=0}^{\infty} \frac{m^k}{k!} \\ &= me^{-m} e^m = m \end{aligned}$$

Varianz der Poisson-Verteilung

Die Varianz der Poisson-Verteilung berechnen wir über den Verschiebungssatz. Dazu benötigen wir:

$$\begin{aligned} E(X_{Poi}^2) &= \sum_{k=1}^{\infty} P(X_{Poi} = k)k^2 = me^{-m} \sum_{k=1}^{\infty} \frac{m^{k-1}}{(k-1)!} k \\ &= me^{-m} \sum_{k=0}^{\infty} \frac{m^k}{k!} (k+1) = m + me^{-m} \sum_{k=1}^{\infty} \frac{m^k}{(k-1)!} \\ &= m + m^2 e^{-m} \sum_{k=1}^{\infty} \frac{m^{k-1}}{(k-1)!} = m + m^2 \end{aligned}$$

Damit ergibt sich unmittelbar:

$$\text{Var}(X_{Poi}) = E(X_{Poi}^2) - E(X_{Poi})^2 = m + m^2 - m^2 = m$$

Fazit: Vergleichen wir dieses Ergebnis mit der Binomialverteilung, so erhalten wir den gleichen Erwartungswert und falls $p \ll 1$ auch nahezu die gleiche Varianz, $\text{Var}(X_{Bin}) = m(1-p)$.

Radioaktiver Zerfall

Als Anwendungsbeispiel der Poisson-Verteilung betrachten wir eine Messung von radioaktivem Zerfall mittels Geigerzähler.

Erwartete Anzahl von Impulsen per Zeiteinheit:

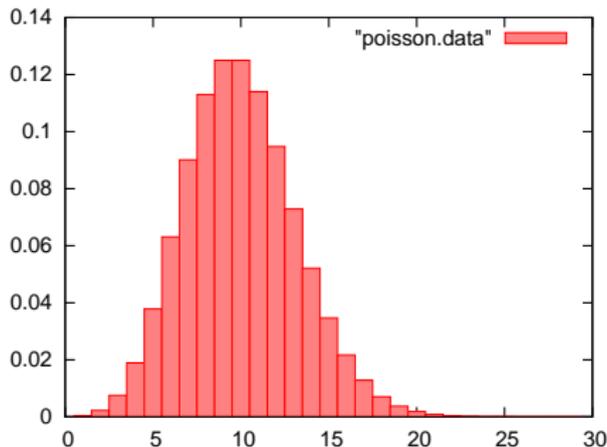
$$E(X_{Poi}) = m = np$$

Hier angenommen $m = 10$.

Frage: Wie groß ist die W., dass in einer Zeiteinheit k Impulse gemessen werden?

$$P(X_{Poi} = k) = \frac{m^k}{k!} e^{-m}$$

k	$P(X = k)$
8	11,3 %
10	12,5 %
12	9,5 %
20	0,2 %



Simulation einer Zeitreihe

- Erzeugung von Zufallszahlen $x_1, \dots, x_n \in [0, 1]$ per Computer.
- Jede Zufallszahl entspreche einem Atom, dass mit der W. p ($\ll 1$) zerfällt.
Also:

$$0 \leq x_k \leq p \quad : \quad k\text{-tes Atom zerfällt,}$$

$$p < x_k \leq 1 \quad : \quad k\text{-tes Atom stabil.}$$

- Wir zählen die Anzahl X der Zerfälle.
- Diesen Versuch führen wir häufiger aus und erwarten, dass X Poisson verteilt ist mit Erwartungswert $E(X_{Poi}) = np$.

Multinomialverteilung

n -maliges Wiederholen eines Einzelerperimentes mit N Möglichkeiten

- Ereignisraum: $\Omega = \{\omega_1, \dots, \omega_N\}^n$
- Spezialfall $N = 2 \rightarrow$ Binomialverteilung.
- Zufallsvariable

$$X : \Omega \rightarrow \mathbb{N}^N, \quad \omega \rightarrow X(\omega) = (x_1, \dots, x_N), \quad 0 \leq x_k \leq n$$

Ermittlung der Verteilungsfunktion:

- p_k = Wahrscheinlichkeit, dass im Einzelerperiment ω_k auftritt:

$$\sum_{k=1}^N p_k = 1$$

- W., dass die ersten n_1 Versuche ω_1 ergeben, die nachfolgenden n_2 das Ereignis ω_2 , etc:

$$p_1^{n_1} \cdot \dots \cdot p_N^{n_N}, \quad \sum_{k=1}^N n_k = n$$

Verteilungsfunktion der Multinomialverteilung

- Es gibt $\binom{n}{n_1}$ Möglichkeiten ω_1 n_1 -mal zu beobachten, $\binom{n - n_1}{n_2}$ Möglichkeiten ω_2 n_2 -mal zu beobachten, etc:
- Wie man sich schnell überlegt:

$$\binom{n}{n_1} \binom{n - n_1}{n_2} \cdots \binom{n - n_1 - \dots - n_{N-1}}{n_N} = \frac{n!}{n_1! \dots n_N!}$$

- Damit erhalten wir:

$$P(X_{multi} = (n_1, \dots, n_N)) = \frac{n!}{n_1! \dots n_N!} p_1^{n_1} \cdots p_N^{n_N}$$

Beispiel zur Multinomialverteilung

- 1 Genort mit zwei Allelen A und a.
- Ergibt $N = 3$ Genotypen AA, Aa und aa mit den Wahrscheinlichkeiten $1/4, 1/2, 1/4$
- Wie groß ist die W., dass bei n Individuen der Genotyp AA n_1 -mal auftritt, etc?
- Zahlenbeispiel: $n = 5, n_1 = 4, n_2 = 0, n_3 = 1$:

$$P(X_{multi} = (4, 0, 1)) = \frac{5!}{4!} \left(\frac{1}{4}\right)^4 \frac{1}{4} = \frac{5}{4^5} \approx 2\%$$

Kontinuierliche Wahrscheinlichkeitsverteilungen

- **Def.:** Unter einer **kontinuierliche Wahrscheinlichkeitsverteilung** versteht man eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}_0^+$ mit der Normierungseigenschaft

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

- Die zugehörige Verteilungsfunktion ($F' = f$) ist die Stammfunktion

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

- Die Wahrscheinlichkeiten berechnen sich dann mittels:

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

- Beispielsweise gilt:

$$P(X \in \mathbb{R}) = 1$$

$$P(X \leq b) = \int_{-\infty}^b f(x) dx$$

$$P(X = a) = \int_a^a f(x) dx = 0$$

Erwartungswert und Varianz kontinuierlicher Verteilungen

- Der Erwartungswert ist definiert als

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx$$

- und die Varianz entsprechend als

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx \\ &= E(X^2) - E(X)^2 \end{aligned}$$

Gleichverteilungen

Unter der Gleichverteilungen im Intervall $[a, b]$ versteht man

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

Bsp.: Drehen eines Kreisel, Endwinkel $\theta \in [0, 2\pi)$.

Standard-Normalverteilung

Unter der **Standard-Normalverteilung** $N(0, 1)$ versteht man die Wahrscheinlichkeitsdichte

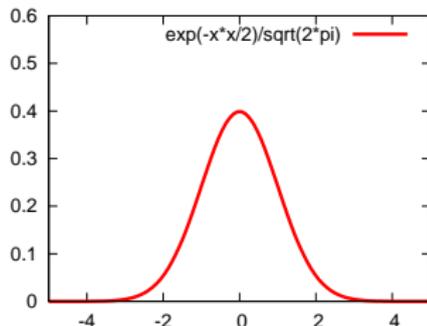
$$f(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$$

Ist eine Zufallsvariable X normalverteilt, so schreibt man $X \sim N(0, 1)$.

Eigenschaften:

- $0 < f(x) \leq \frac{1}{\sqrt{2\pi}}$
- symmetrisch zur y -Achse (gerade Funktion),
 $f(x) = f(-x)$
- Eine nicht triviale Rechnung ergibt:

$$\int_{\mathbb{R}} f(x) dx = 1$$



Erwartungswert und Varianz von $N(0, 1)$

- Der Erwartungswert von $N(0, 1)$ ist $E(X) = 0$, denn $xf(x)$ ist ungerade:

$$E(X_{N(0,1)}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \exp(-x^2/2) = 0$$

- Die Varianz ist

$$\begin{aligned} \text{Var}(X_{N(0,1)}) &= E(X_{N(0,1)}^2) - E(X_{N(0,1)})^2 = E(X_{N(0,1)}^2) \\ &= \int_{-\infty}^{\infty} x^2 f(x) = \dots = 1 \end{aligned}$$

- Die zugehörige Verteilungsfunktion lautet:

$$\Phi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt$$

Rechnen mit der Standard-Normalverteilung

Die Wahrscheinlichkeiten für eine normalverteilte Zufallsvariable X werden üblicherweise über Tabellen bestimmt:

z-Werte (Teil 1)	Teil 2 des z-Wertes (2. Nachkommastelle)							
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	...
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	...
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	...
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Bsp.:

$$P(X_{N(0,1)} \leq -2.83) = 0,23\%$$

$$P(X_{N(0,1)} \geq 2.71) = P(X_{N(0,1)} \leq -2.71) = 0,34\%$$

$$\begin{aligned} P(-2.94 \leq X_{N(0,1)} \leq -2.73) &= P(X_{N(0,1)} \leq -2.73) - P(X_{N(0,1)} < -2.94) \\ &= (0,32 - 0,16)\% \end{aligned}$$

Normalverteilungen

- Die Verallgemeinerung der Standard-Normalverteilung ist die Normalverteilung zu Parametern $\mu \in \mathbb{R}$, $\sigma^2 > 0$.
- Unter der **Normalverteilung** $N(\mu, \sigma^2)$ versteht man die Wahrscheinlichkeitsdichte

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

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

Eigenschaften der Normalverteilung:

- $0 < f(x) \leq \frac{1}{\sqrt{2\pi}\sigma}$
- symmetrisch zur Achse $x = \mu$:

$$f(\mu - x) = f(\mu + x)$$

- Es gilt ebenfalls $\int_{\mathbb{R}} f(x) dx = 1$

Erwartungswert und Varianz der Normalverteilungen

- Der Erwartungswert von $N(\mu, \sigma^2)$ ergibt sich durch Substitution $y = (x - \mu)/\sigma$:

$$\begin{aligned} E(X_{N(\mu, \sigma^2)}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\left(\frac{-(x - \mu)^2}{2\sigma^2}\right) dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (y + \mu) \exp\left(\frac{-y^2}{2}\right) dy \\ &= \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \exp\left(\frac{-y^2}{2}\right) dy}_{=E(Y_{N(0,1)})=0} + \mu \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(\frac{-y^2}{2}\right) dy}_{=1} \\ &= \mu \end{aligned}$$

- Die Varianz von $N(\mu, \sigma^2)$ ist

$$\text{Var}(X_{N(\mu, \sigma^2)}) = \sigma^2$$

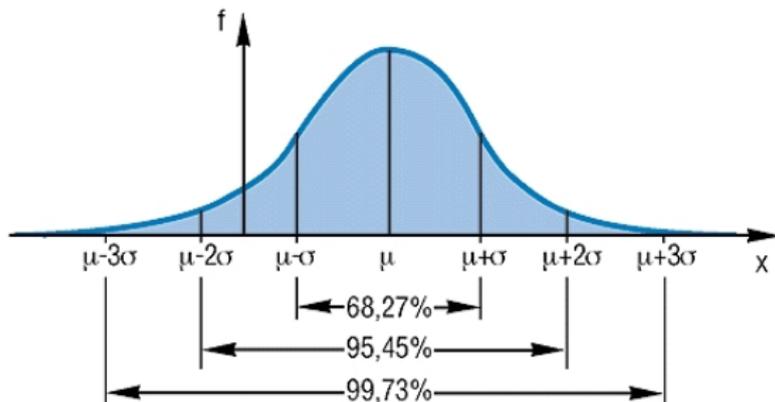
Zahlen über die Normalverteilung

Für eine normalverteilte Zufallsvariable X beträgt die Wahrscheinlichkeit, dass ein Ereigniss im Intervall $[\mu - k\sigma, \mu + k\sigma]$ liegt:

$$P(\mu - \sigma \leq X_{N(\mu, \sigma^2)} \leq \mu + \sigma) \approx 0.6827$$

$$P(\mu - 2\sigma \leq X_{N(\mu, \sigma^2)} \leq \mu + 2\sigma) \approx 0.9545$$

$$P(\mu - 3\sigma \leq X_{N(\mu, \sigma^2)} \leq \mu + 3\sigma) \approx 0.9973$$



Rechnen mit der Normalverteilung

- Eine Normalverteilung $N(\mu, \sigma^2)$ wird idR. in die Standard-Normalverteilung umgerechnet:

$$X \sim N(\mu, \sigma^2) \Leftrightarrow Y = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

- Damit ergibt sich:

$$P(X_{N(\mu, \sigma^2)} \leq a) = P(Y_{N(0, 1)} \leq (a - \mu)/\sigma)$$

- **Beispiel:** Die Körpergröße der ausgewachsenen HD Eichhörnchen sei $N(30 \text{ cm}, (4 \text{ cm})^2)$ -verteilt. Wie groß ist die W., dass ein Heidelberger Eichhörnchen größer als 35 cm ist?

$$\begin{aligned} P(X_{N(\mu, \sigma^2)} \geq 35 \text{ cm}) &= P(Y_{N(0, 1)} \geq 1.25) = 1 - P(Y_{N(0, 1)} < 1.25) \\ &= 1 - 0.8944 = 10.56\% \end{aligned}$$

- Bei gleichverteilten Ereignissen ist die Zufallsvariable

$$Y_{Poisson}^{[0,1]} = \text{Anzahl Ereignisse im Einheitszeitintervall } [0, 1]$$

gegeben durch die Poisson-Verteilung:

$$P(Y_{Poisson}^{[0,1]} = k) = \frac{\mu^k}{k!} e^{-\mu},$$

wobei μ den Erwartungswert für die Anzahl von Ereignissen in $[0, 1]$ darstellt.

- Entsprechend ist die Zufallsvariable

$$Y_{Poisson}^{[0,t]} = \text{Anzahl Ereignisse im Zeitintervall } [0, t]$$

gegeben durch die Poisson-Verteilung:

$$P(Y_{Poisson}^{[0,t]} = k) = \frac{(\mu t)^k}{k!} e^{-\mu t},$$

- X_{exp} bezeichne nun die Zufallsvariable *“Zeitpunkt des ersten Ereignisses”*, z.B.:

$$P(X_{exp} > t) = P(Y_{Poisson}^{[0,t]} = 0) = e^{-\mu t}$$

Exponentialverteilung

- Wir betrachten die Wahrscheinlichkeit

$$P(X_{exp} > t) = e^{-\mu t}$$

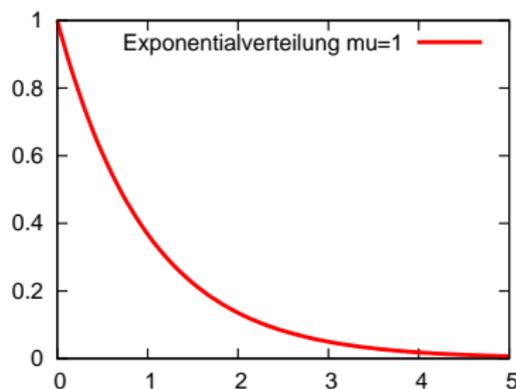
- Zugehörige **Verteilungsfunktion**:

$$F(t) = P(X_{exp} \leq t) = 1 - P(X_{exp} > t) = 1 - e^{-\mu t}$$

- Die zugehörige Wahrscheinlichkeitsdichte ist die **Exponentialverteilung**:

$$f(t) = F'(t) = \mu e^{-\mu t} \quad \text{für } t \geq 0.$$

- $X = X_{exp}$ heißt dann exponentialverteilt zum Parameter $\mu > 0$.



Erwartungswert und Varianz der Exponentialverteilung

- **Erwartungswert:**

$$E(X_{exp}) = \int_0^{\infty} t\mu e^{-\mu t} dt = \dots = \frac{1}{\mu}$$

- **Varianz:**

$$Var(X_{exp}) = E(X_{exp}^2) - E(X_{exp})^2 = \dots = \frac{1}{\mu^2}$$

- **Anwendungsbeispiel:** Die Exponentialverteilung wird häufig bei Warteprozessen eingesetzt, bei denen kein “Gedächtnis” existiert, z.B. elektronische Bauteile. Beispielsweise verhält sich die Lebensdauer einer Glühbirne annähernd exponential-verteilt:

Die Lebenserwartung einer alten Glühbirne ist $E(X_{exp}) = \frac{1}{\mu}$ und damit gleich derer einer neuen Glühbirne.

Anwendung der Exponentialverteilung: Überlebenszeiten 1

- **Szenario:** N_0 Lebewesen unter Nahrungsmangel (daher Alterungsprozeß irrelevant)
- μ = Sterberate, dh. im Zeitraum $[0, 1]$ sterben μN_0 Individuen.
- Ist X = Sterbezeit eines Individuums, so ist die zugehörige Wahrscheinlichkeitsdichte:

$$f(t) = \mu e^{-\mu t}$$

und die Verteilungsfunktion:

$$P(X \leq t) = F(t) = 1 - e^{-\mu t}$$

$$P(X > t) = 1 - P(X \leq t) = e^{-\mu t}$$

- Wie wird sich die Anzahl $N(t)$ an Individuen entwickeln?
- Gewöhnliche Differentialgleichung ($\dot{N} = \partial N / \partial t$):

$$\dot{N}(t) = -\mu N(t)$$

$$N(0) = N_0$$

- Lösung:

$$N(t) = N_0 e^{-\mu t}$$

Anwendung der Exponentialverteilung: Überlebenszeiten 2

Praxis: In der Praxis kennt man μ häufig nicht, sondern muß dies schätzen.

Beispielsweise misst man $N(t)$, $N(t - \Delta t)$, $N(t + \Delta t)$ und approximiert $\dot{N}(t)$ durch einen Differenzenquotienten:

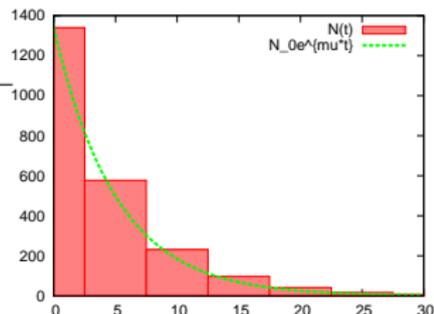
$$d(t) = \frac{1}{2\Delta t}(N(t + \Delta t) - N(t - \Delta t)) \approx \dot{N}(t)$$

$$\tilde{\mu}(t) = -\frac{d}{N(t)} \quad \text{Schätzung für } \mu$$

Beispiel:

t	0	5	10	15	20	25	30
$N(t)$	1340	577	231	97	42	17	7
$d(t)$		-111	-48	-18.9	-8	-3.5	
$\tilde{\mu}(t)$		0.19	0.21	0.19	0.19	0.21	

Der Mittelwert der $\tilde{\mu}(t)$ ergibt sich zu 0.1982.



Ungleichung von Tschebyscheff

Satz: Für eine beliebige ZV X mit $\mu = E(X)$ und $\text{Var}(X) = \sigma^2$ gilt:

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \quad \forall \epsilon > 0$$

Beweis: Dies sieht man einfach mittels:

$$\begin{aligned} \sigma^2 &= \int_{\mathbb{R}} (x - \mu)^2 f(x) dx \\ &\geq \int_{|x - \mu| > \epsilon} (x - \mu)^2 f(x) dx \\ &\geq \epsilon^2 \int_{|x - \mu| > \epsilon} f(x) dx \\ &\geq \epsilon^2 P(|X - \mu| > \epsilon) \end{aligned}$$

Für diskrete ZVn argumentiert man entsprechend.

Satz: Für eine beliebige ZV X mit $\mu = E(X)$ und $\text{Var}(X) = \sigma^2$ gilt

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

Beweis: Man wählt einfach in der Ungleichung von Tschebyscheff $\epsilon = k\sigma$:

$$\begin{aligned} P(|X - \mu| \leq k\sigma) &= 1 - P(|X - \mu| > \epsilon) \\ &\geq 1 - \frac{\sigma^2}{\epsilon^2} = 1 - \frac{1}{k^2} \end{aligned}$$

Beispiel:

$$P(|X - \mu| \leq 2\sigma) \geq 75\%$$

Für spezielle Verteilungen gelten u.U. größere Schranken, zB. für Normalverteilungen:

$$P(|X_{N(\mu, \sigma^2)} - \mu| \leq 2\sigma) \geq 95\%.$$

Zentraler Grenzwertsatz

Satz: (Satz von de Moivre–Laplace / Zentraler Grenzwertsatz)

Seien X_i , $i \in \mathbb{N}$ unabhängig und identisch verteilte ZVn mit $E(X_i) = \mu$ und $\text{Var}(X_i) = \sigma^2 > 0$. Dann gilt für die ZV $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$:

(a) $E(\bar{X}_n) = \mu$ und $\text{Var}(\bar{X}_n) = \sigma^2/n$.

(b) Die normalisierte ZV \bar{X}_n^* konvergiert gegen $N(0, 1)$.

$$\bar{X}_n^* = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

Bemerkungen:

zu (a):

$$\text{Var}(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\sigma^2}{n} \rightarrow 0 \quad (n \rightarrow \infty)$$

zu (b): \bar{X}_n ist näherungsweise normalverteilt, obgleich die X_i beliebig (aber gleich) verteilt sein können. Den Konvergenz-Begriff wollen wir in diesem Zusammenhang nicht näher spezifizieren.

Beispiel zum Zentralen Grenzwertsatz

- Eine Labormaschine fülle ein Reagenzglas mit $\mu = 1 \text{ g}$ und $\sigma = 0.12$. Nun werden $n = 36$ Proben in ein größeres Gefäß gefüllt. Mit welcher Wahrscheinlichkeit weicht die Endmenge um höchstens 1 g vom Zielwert $n\mu$ ab?
- X_i = Füllmenge vom i -ten Reagenzglas, $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$
- Annahme:

$$\bar{X}_n^* = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \sim N(0, 1)$$

- Wir erhalten:

$$\begin{aligned} P(|X_1 + \dots + X_n - n\mu| \leq 1) &= P(|\bar{X}_n - \mu| \leq \frac{1}{n}) \\ &= P(|\bar{X}_n^*| \leq \frac{1}{\sigma\sqrt{n}}) \\ &= P(|\bar{X}_n^*| \leq 1.389) \\ &= P(\bar{X}_n^* \leq 1.389) - P(\bar{X}_n^* \leq -1.389) \\ &\approx 0.9177 - 0.0823 \approx 83.54\% \end{aligned}$$

Gesetz der großen Zahlen

Satz: Seien X_1, X_2, \dots unabhängig verteilte ZVn mit gleichem Erwartungswert $E(X_i) = \mu$ und Varianz $Var(X_i) = \sigma^2 > 0$, für alle $i \in \mathbb{N}$. Dann gilt für die ZV $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$:

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \leq \epsilon) = 1 \quad \forall \epsilon > 0.$$

Beweis: Der Grund hierfür ist der, dass der Erwartungswert von \bar{X}_n ebenfalls μ ist, die Varianz von \bar{X}_n aber gegen Null geht, $Var(\bar{X}_n) \rightarrow 0$ für $n \rightarrow \infty$. Die Ungleichung von Tschebyscheff ergibt nun

$$\begin{aligned} P(|\bar{X}_n - \mu| \leq \epsilon) &= 1 - P(|\bar{X}_n - \mu| > \epsilon) \\ &\geq 1 - \underbrace{\frac{Var(\bar{X}_n)}{\epsilon^2}}_{\rightarrow 0} \end{aligned}$$

Kap. 10: Beurteilende Statistik

Schätzfunktionen

Situation: Wir haben n Stichproben $x_1, \dots, x_n \in \mathbb{R}$ und wollen hieraus *Maßzahlen* (zB. Erwartungswert oder Varianz) für die Grundgesamtheit, die durch eine ZV X modelliert werde, ermitteln.

Def. Unter einer **Schätzfunktion** verstehen wir eine Abbildung θ_n , die der Stichproben einen Schätzwert für eine Maßzahl κ zuordnet.

Beispiele:

- Schätzung von $E(X)$:

$$\theta_n(x_1, \dots, x_n) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \approx E(X)$$

- Schätzung von $Var(X)$ mittels empirischer Varianz $Var(x_1, \dots, x_n)$:

$$\theta_n(x_1, \dots, x_n) = Var(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \approx Var(X)$$

Erwartungstreue Schätzer

Beobachtung: Ist σ_n eine Schätzfunktion für eine Maßzahl κ zu Realisierungen x_1, \dots, x_n von ZVn X_i , so ist $\theta_n(X_1, \dots, X_n)$ wieder eine ZV.

Def. Einer Schätzfunktion θ_n für eine Maßzahl κ heißt **erwartungstreu**, wenn

$$E(\theta_n(X_1, \dots, X_n)) = \kappa$$

Beispiel:

- Schätzung von $E(X)$ ist erwartungstreu:

$$E(\theta_n(X_1, \dots, X_n)) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = E(X)$$

Konfidenzintervalle

Situation: Wir gehen aus von einer Schätzung von $E(X) = \mu$ mittels einer Schätzfunktion (also $\bar{x} \approx \mu$).

- Wie groß ist die Wahrscheinlichkeit, dass μ tatsächlich im Bereich $[\bar{x} - \epsilon, \bar{x} + \epsilon]$ zu gegebenem $\epsilon > 0$ liegt?
- Wie groß ist $\epsilon > 0$, so dass zu gegebenem $\alpha \in (0, 1)$:

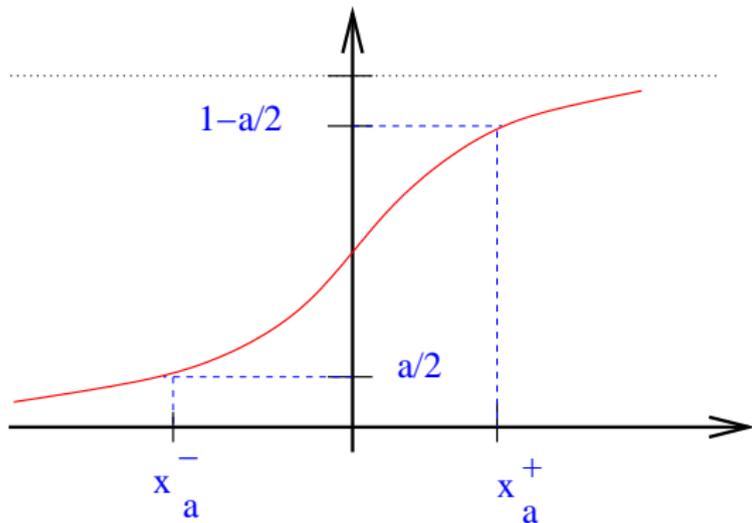
$$P(\bar{x} - \epsilon \leq \mu \leq \bar{x} + \epsilon) \geq 1 - \alpha$$

Definition: Zu $\alpha \in (0, 1)$ heißt ein Intervall $[x_{\alpha}^{-}, x_{\alpha}^{+}]$ α -Konfidenzintervall für X , wenn

$$P(x_{\alpha}^{-} \leq X \leq x_{\alpha}^{+}) = 1 - \alpha$$

Konfidenzintervalle für Standard-Normalverteilung

- 1 Bestimme x_{α}^{+} , so dass $\Phi(x_{\alpha}^{+}) = 1 - \frac{\alpha}{2}$.
- 2 Setze $x_{\alpha}^{-} = -x_{\alpha}^{+}$



Konfidenzintervalle für Normalverteilungen

Annahme: Varianz σ^2 sei bekannt.

- 1 Bestimme z_α^+ aus der Standard-Normalverteilung, $\Phi(z_\alpha^+) = 1 - \frac{\alpha}{2}$.
- 2 Skalieren um

$$x_\alpha^\pm = \bar{x} \pm \frac{\sigma}{\sqrt{n}} z_\alpha^+$$

Beispiel: Ernte von 100 Äpfeln mit $\bar{x} = 142 \text{ g}$ und $\text{Var}(x_i) = (10 \text{ g})^2$. Gesucht sei das 10%-Konfidenzintervall um \bar{x} :

$$\Phi(z_\alpha^+) = 1 - \frac{\alpha}{2} = 0.95 \Rightarrow z_\alpha^+ \approx 1.645$$

$$x_\alpha^\pm = \bar{x} \pm 1.645 \approx 143.645$$

Es gilt also

$$P(140.355 \text{ g} \leq \mu \leq 143.645 \text{ g}) \geq 90\%$$

Asymptotisch erwartungstreuer Schätzer für die Varianz

- Der Schätzer für die Varianz

$$\theta_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

ist nur asymptotisch erwartungstreu:

$$E(\theta_n(X_1, \dots, X_n)) \rightarrow \text{Var}(X) \quad (n \rightarrow \infty)$$

Erwartungstreuer Schätzer für die Varianz

- Erwartungstreuer Schätzer für die Varianz

$$\theta_n(x_1, \dots, x_n) = S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- Verschiebungssatz:

$$\theta_n(X_1, \dots, X_n) = \frac{n}{n-1} (\overline{X^2} - \bar{X}^2)$$

- Da $E(\overline{X^2}) = \text{Var}(\bar{X}) + E(\bar{X})^2 = \frac{1}{n} \text{Var}(X) + E(X)^2$ und $E(\bar{X}^2) = E(X^2)$:

$$\begin{aligned} E(\theta_n) &= \frac{n}{n-1} (E(X^2) - \frac{1}{n} \text{Var}(X) - E(X)^2) \\ &= \frac{n}{n-1} (\text{Var}(X) - \frac{1}{n} \text{Var}(X)) \\ &= \text{Var}(X) \end{aligned}$$

Konfidenzintervalle für $N(\mu, \sigma^2)$ bei unbekannter Varianz

Situation: Varianz σ^2 sei unbekannt. Gesucht sei ein Konfidenzintervall für μ .

- Wähle Schätzer für μ und σ^2 :

$$\mu \approx \bar{x} \quad \text{und} \quad \sigma^2 \approx S_n^2$$

- Die ZV \bar{X}_n ist $N(\mu, \frac{\sigma^2}{n})$ verteilt.
- Die standardisierte ZV \bar{X}_n^* ist $N(0, 1)$ verteilt.

$$\bar{X}_n^* = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \sim N(0, 1)$$

- Die “**Students t -Verteilung mit $n - 1$ Freiheitsgraden**” ist die Verteilung zum geschätzten Erwartungswert:

$$T_n = \frac{\sqrt{n}}{S_n}(\bar{X}_n - \bar{x}) \sim t_{n-1}$$

und ist eine Approximation an \bar{X}_n^* .

Students t -Verteilung (Gosset 1908)

$$T_n = \frac{\sqrt{n}}{S_n}(\overline{X}_n - \bar{x})$$

- Wahrscheinlichkeitsdichte von t_{n-1} :

$$f(x) = C_{n-1} \left(1 + \frac{x^2}{n-1} \right)^{-n/2}$$

C_{n-1} dient der Normierung.

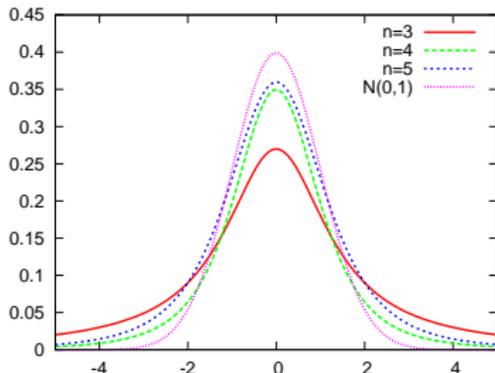
- symmetrisch bzgl. der $x=0$ -Achse und

$$E(T_n) = 0$$

$$\text{Var}(T_n) = \begin{cases} \frac{n-1}{n-3}, & \text{wenn } n > 3 \\ \infty, & \text{sonst} \end{cases}$$

- konvergiert für $n \rightarrow \infty$ gegen $N(0, 1)$:
- Die Varianz der Students t -Verteilung nimmt mit wachsendem n ab (wird aber nicht beliebig klein):

$$1 < \dots < \text{Var}(T_n) < \text{Var}(T_{n-1}) < \dots$$



- Ist $n > 30 \rightarrow$ Normalverteilung, sonst:
- Bestimmung des α -Konfidenzintervalls der t_{n-1} -Verteilung:

$$P(-t_\alpha \leq T_n \leq t_\alpha) = 1 - \alpha$$

mittels Tabelle.

- Skalierung:

$$x_\alpha = \frac{S_n}{\sqrt{n}} t_\alpha$$

- Konfidenzintervall des Erwartungswerts bei geschätzter Varianz:

$$P(\bar{x} - x_\alpha \leq X \leq \bar{x} + x_\alpha) = 1 - \alpha$$

Konfidenzintervalle für die Varianz bei $N(\mu, \sigma^2)$ (1)

- $X \sim N(\mu, \sigma^2)$ aber μ und σ^2 seien unbekannt.
- Gesucht ist zu gegebenem $0 < \alpha < 1$ ein Konfidenzintervall $[s_\alpha^-, s_\alpha^+]$, so dass

$$P(s_\alpha^- \leq \sigma^2 \leq s_\alpha^+) = 1 - \alpha$$

- Wir betrachten die ZV

$$\hat{S}_n^2 = \frac{n-1}{\sigma^2} S_n^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

und bezeichnen ihre Verteilung als χ_n^2 -Verteilung (sprich: *Chi-Quadrat-verteilt mit n Freiheitsgraden*).

- \hat{S}_n^2 ist eine Summe von n unabhängig und identisch verteilten $N(0, 1)$ ZVn.
- Die Wahrscheinlichkeitsdichte von χ_n^2 ist:

$$f_n(x) = \begin{cases} 0 & , \text{für } t \leq 0 \\ C_n x^{n-2} e^{-x/2} & , \text{sonst.} \end{cases}$$

mit $E(\chi_n^2) = n$ und $\text{Var}(\chi_n^2) = 2n$.

Konfidenzintervalle für die Varianz bei $N(\mu, \sigma^2)$ (2)

- α -Konfidenzintervall für \hat{S}_n^2 :

$$P(\hat{s}_\alpha^- \leq \hat{S}_n^2 \leq \hat{s}_\alpha^+) = 1 - \alpha$$

bzw.

$$F_{\chi_n^2}(\hat{s}_\alpha^-) = P(\hat{S}_n^2 < \hat{s}_\alpha^-) = \frac{\alpha}{2}$$

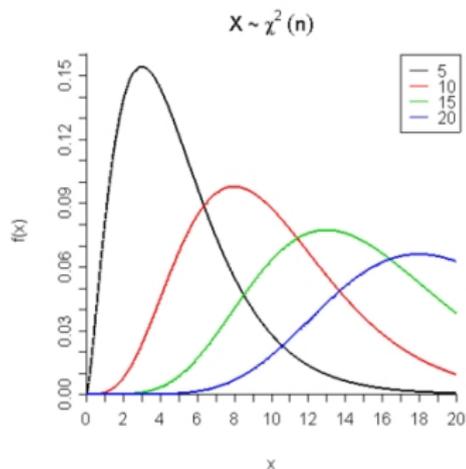
$$F_{\chi_n^2}(\hat{s}_\alpha^+) = P(\hat{S}_n^2 < \hat{s}_\alpha^+) = 1 - \frac{\alpha}{2}$$

- Skalierung:

$$s_\alpha^+ = \frac{n-1}{\hat{s}_\alpha^-} S_n^2$$

$$s_\alpha^- = \frac{n-1}{\hat{s}_\alpha^+} S_n^2$$

χ_n^2 -Wahrscheinlichkeitsdichten:



Beispiel zur χ_n^2 -Verteilung

Situation: Die Größe eines Insekts sei als (μ, σ^2) -normalverteilt angenommen. Die Parameter μ, σ^2 seien unbekannt. Eine Untersuchung an $n = 15$ Insekten zeigt eine empirische Varianz von $S_n = 2.5$.

Frage: Wie lautet ein 10%-Konfidenzintervall für σ^2 ?

Lösung:

- Bestimme \hat{s}_α^- , \hat{s}_α^+ , so dass

$$F_{\chi_n^2}(\hat{s}_\alpha^-) = \frac{\alpha}{2} \quad \text{und} \quad F_{\chi_n^2}(\hat{s}_\alpha^+) = 1 - \frac{\alpha}{2}$$

Hier ergibt sich: $\hat{s}_\alpha^+ = 25$ und $\hat{s}_\alpha^- = 7.26$.

- Nun gilt $P(\hat{s}_\alpha^- < \hat{S}_n^2 < \hat{s}_\alpha^+) = 1 - \alpha$.
- Setze

$$s_\alpha^+ = \frac{n-1}{\hat{s}_\alpha^-} S_n^2 = \frac{14}{7.26} \frac{5^2}{2^2} = 12.05$$

$$s_\alpha^- = \frac{n-1}{\hat{s}_\alpha^+} S_n^2 = \frac{14}{25} \frac{5^2}{2^2} = 3.5$$

Wir halten also mit 90%-iger Sicherheit:

$$\sigma^2 \in [3.5, 12.05]$$

Testen von Hypothesen

Beispiel: Eine Bier-Zapfanlage auf der WM 2006 soll genau $\mu_0 = 500 \text{ ml}$ Bier abfüllen. Zu prüfen ist die **Null-Hypothese**:

$$H_0 : E(X) = \mu_0$$

Frage: Wann ist eine Abweichung von \bar{x} von μ signifikant ?

Die **Gegenhypothese** kann von der Interessengruppe abhängen:

- **Linksseitige Fragestellung:** (Verband der Biertrinker)

$$H_1 : E(X) < \mu_0$$

- **Rechtsseitige Fragestellung:** (Bierproduzent)

$$H_1 : E(X) > \mu_0$$

- **Zweiseitige Fragestellung:** (Produzent der Abfüllanlage)

$$H_1 : E(X) \neq \mu_0$$

Merke: Die Gegenhypothese muß nicht die logische Negation der Null-Hypothese sein.

Beispiel: Testen von Hypothesen

- Wir nehmen an, dass $X \sim N(\mu, \sigma^2)$ mit μ und σ unbekannt sind.
- Stichprobendaten: Biermengen x_1, \dots, x_n mit $n = 10$, Fehlermaß sei $\alpha = 0.1$.
- Prüfgröße ist t_9 -verteilt:

$$T_n = \frac{\sqrt{n}}{S_n}(\bar{X}_n - \mu)$$

- Wähle Konfidenzintervall für T_n zu gegebenem Konfidenzniveau α :

$$P(-1.383 \leq T \leq 1.383) = 1 - \alpha = 0.9$$

- Reskaliertes Konfidenzintervall:

$$x_\alpha = 1.383 \frac{S_n}{\sqrt{n}}$$
$$\bar{x}_n \in [\mu - x_\alpha, \mu + x_\alpha]$$

- Beispielsweise muss für $S_n = 16$ und $\bar{x}_n = 492 \text{ ml}$ die Nullhypothese $H_0 : \mu = 500 \text{ ml}$ unter der Fehlergrenze α verworfen werden ($\mu - x_\alpha = 493$).

t-Test auf Lageunterschied (1)

- **Situation:** Zwei Gruppen von Laborratten mit unterschiedlicher Fütterung. Betrachtet wird die Gewichtszunahme.
- **Frage:** Unterscheidet sich die mittlere Gewichtszunahme?
- **Mathematisches Modell:** $X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, \sigma^2)$, also unterschiedliche Erwartungswerte bei gleicher (unbekannter) Varianz.
- Hypothesen:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

- Die ZV

$$Z := \bar{Y} - \bar{X} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i - \frac{1}{n_1} \sum_{i=1}^{n_1} X_i$$

ist normalverteilt mit:

$$E(Z) = \mu_2 - \mu_1$$

$$\text{Var}(Z) = \text{Var}(\bar{Y}) + \text{Var}(\bar{X}) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

t-Test auf Lageunterschied (2)

- Da σ nicht bekannt, verwende erwartungstreuen Schätzer für $\text{Var}(Z)$:

$$S_n^2 = \frac{1}{n-2} \left(\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \right), \quad n = n_1 + n_2$$

- Dann ist

$$T = \frac{Z}{S_n \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

t-verteilt mit $n - 2$ Freiheitsgraden.

- Bestimme α -Konfidenzintervall $I_\alpha = [-t_\alpha, t_\alpha]$.
- Reskaliere: Liegt $\bar{y} - \bar{x} \in [-z_\alpha, z_\alpha]$, mit

$$z_\alpha = t_\alpha S_n \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad ?$$

t-Test auf Lageunterschied bei verbundenen Stichproben

- **Situation:** Ein neues Futtermittel bei Milchkühen ergebe folgende Milchmengen jeweils vor und nach der Fütterung:

Kuh Nr.	1	2	3	4	5
vorher	5	3	4	6	5
nachher	4.5	6	7	4	5

- **Frage:** Taugt das Futtermittel etwas?
- Betrachte ZV $Z = \overline{Y} - \overline{X}$ die als normalverteilt angenommen wird mit unbekanntem μ und σ^2 .
- Nullhypothese:

$$H_0 : \mu = 0$$

- $Z^* = \sqrt{S_5^2/n}(Z - \bar{z})$ ist t_4 -verteilt. 10 %-Konfidenzniveau:

$$P(Z^* \in [-1.533, 1.533]) = 0.9$$

- Rückskalierung, $Z = \bar{z} + \sqrt{5}/S_5 Z^*$ mit $S_5 = \sqrt{4.95}$ und $\bar{z} = 1.5$, also $Z \approx 1.5 + Z^*$:

$$P(Z \in [-0.023, 3.033]) = 0.9$$

- Da $0 \in [-0.023, 3.033]$ wird die Nullhypothese akzeptiert.

Test auf Lageunterschied bei nicht normalverteilten Daten

Beispiel: Können Ratten lernen? Bei zwei Gruppen von Ratten wird die Anzahl von Tagen gemessen, bis ein gewisses Lernziel erreicht ist. Gruppe 1 **mit** Training, Gruppe 2 **ohne** Training.

X	110	53	70	51	88	90	77
Y	78	63	75	45	48	67	71

Mathematische Frage: Existiert ein $\delta \in \mathbb{R}$, so dass

$$P(X \leq x + \delta) = P(Y \leq x) \quad \forall x \in \mathbb{R} \quad ?$$

Gibt es solch ein δ mit $\delta > 0$, so interpretieren wir dies als Aussage, dass die Daten aus der Stichprobe X tendenziell um δ größer sind.

Problem: Was tun, wenn man **keine** Normalverteilung voraussetzen darf, da

- die Verteilung schief sein kann,
- Ausreisser dabei sein können, oder
- die Datenmenge zu gering ist?

Wilcoxon Regressionstest

- Ordne die Stichproben x_1, \dots, x_{n_1} und y_1, \dots, y_{n_2} in aufsteigender Reihenfolge an:

Wert	45	48	51	53	63	67	70	71	75	77	78	88	90	110
Gruppe	Y	Y	X	X	Y	Y	X	Y	Y	X	Y	X	X	X
Rang	1	2	3	4	5	6	7	8	9	10	11	12	13	14

- Bilde die Summe der Rangzahlen:

$$T_X = \sum_{i=1}^{n_1} R_{x_i}, \quad T_Y = \sum_{i=1}^{n_2} R_{y_i}$$

- Es gilt: $T_X + T_Y = \frac{1}{2}n(n+1)$ mit $n = n_1 + n_2$
- Wenn $\delta = 0$ liefern kombinatorische Betrachtungen:

$$E(T_X) = \frac{n_1}{n}(T_X + T_Y) = \frac{1}{2}n_1(n+1) =: \mu$$
$$\text{Var}(T_X) = \frac{1}{12}n_1n_2(n+1) =: \sigma^2$$

- Ist n_1 oder n_2 hinreichend groß (≥ 10), so kann man annehmen:

$$T_X \sim N(\mu, \sigma^2) \quad \text{bzw.} \quad Z = (T - \mu)/\sigma \sim N(0, 1)$$

- Teste Z auf Standard-Normalverteilung. Hier gilt $T_x = 63$, also $P(T_x \geq 63) = 1 - P(Z \leq 1.34) = 1 - 0.9099 \approx 9\%$

Regression

Die **Regressionsrechnung** beschäftigt sich mit dem funktionalen Zusammenhang einer Variablen von einer anderen.

Beispiele:

- Wie hängt die Körpergröße einer Person von der des Vaters ab?
- Wie verhält sich im Mittel der Blutdruck einer Person als Funktion des Alters?

Es wird jeweils ein funktionaler Zusammenhang vorausgesetzt. Die Regression versucht nun diese Funktion näher zu bestimmen.

Korrelation

Die **Korrelation** beschäftigt sich mit der Frage, ob überhaupt ein (funktionaler) Zusammenhang zweier Variablen vorliegt.

Beispiele:

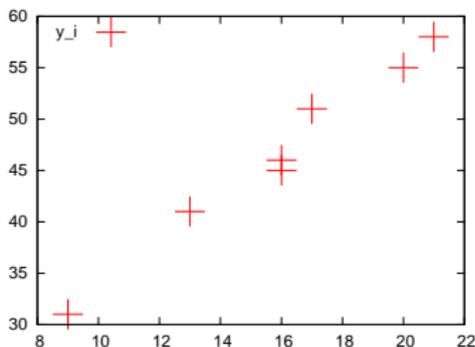
- Wie verhält sich die Anzahl von Neuzulassungen von PKWs in Abhängigkeit vom Bruttonsozialprodukt?
- Wie verhält sich die Konzentration eines bestimmten Hormons im Blut im Vergleich zur Herzfrequenz.

Es ist also zunächst fraglich, ob eine Abhängigkeit besteht. Die Korrelation versucht nun hierauf eine quantitative Antwort zu geben.

Lineare Korrelation

Beispiel: Gegeben seien folgende Daten von Futtermenge x und Körpergröße y an n Lebewesen:

x_i	y_i
13	41
16	46
9	31
17	51
20	55
16	45
21	58



Frage: Können wir von einem linearen Zusammenhang ausgehen ?

$$f(x) = ax + b$$

Empirischen Kovarianz

Definition: Unter der **empirischen Kovarianz** verstehen wir die Größe:

$$\text{Cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Eigenschaften:

- Ausmultiplizieren ergibt:

$$\text{Cov}(x, y) = \overline{xy} - \bar{x} \cdot \bar{y}$$

- Eine unmittelbare Konsequenz ist:

$$\text{Cov}(x, x) = \overline{x^2} - \bar{x}^2 = \text{Var}(x)$$

- Man kann zeigen, dass gilt:

$$|\text{Cov}(x, y)| \leq \sqrt{\text{Var}(x) \text{Var}(y)}$$

Korrelationskoeffizient

Definition: Unter dem **empirischen Korrelationskoeffizienten** (nach Pearson) versteht man die Größe:

$$\varrho(x, y) = \text{Cor}(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x) \text{Var}(y)}}$$

Wegen $\text{Cov}(x, y)^2 \leq \text{Var}(x) \text{Var}(y)$ gilt:

$$-1 \leq \varrho(x, y) \leq 1$$

Satz: Gilt für die x_i und y_i ein (exakter) linearer Zusammenhang, so gilt

$$\varrho(x, y) \in \{-1, 1\}$$

Praxis: In der Praxis wird $\varrho(x, y)$ als Indikator für einen linearen Zusammenhang verwendet:

- Ist $\varrho(x, y)$ nahe bei -1 oder bei $+1$ so kann von einem linearen Zusammenhang ausgegangen werden.
- Im Fall von $\rho \sim 0$ ist kein linearer Zusammenhang erkennbar.
- Das Vorzeichen von $\varrho(x, y)$ entspricht dem Vorzeichen vom linearen Koeffizienten a .

Beispiel Korrelationskoeffizient

Gegeben sind die Werte x_i, y_i , alle anderen werden berechnet:

i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	13	41	169	1681	533
2	16	46	256	2116	736
3	9	31	81	961	279
4	17	51	289	2601	867
5	20	55	400	3025	1100
6	16	45	256	2025	720
7	21	58	441	3364	1218
Summe	112	327	1892	15773	5453
Mittelwert	16	46.7	270.3	2253.3	779

$$\text{Var}(x) = \overline{x^2} - \bar{x}^2 = 270.3 - 16^2 = 14.3$$

$$\text{Var}(y) = \overline{y^2} - \bar{y}^2 = 2253.3 - 46.7^2 = 72.4$$

$$\text{Cov}(x, y) = \overline{xy} - \bar{x} \cdot \bar{y} = 779 - 16 \cdot 46.7 = 31.8$$

$$\rho = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x) \text{Var}(y)}} = \frac{31.8}{\sqrt{14.3 \cdot 72.4}} = 0.988$$

Lineare Korrelation bei Zufallsvariablen

Bei zwei ZVn X und Y läßt sich der Korrelationskoeffizient entsprechend definieren:

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

Die Kovarianz ist hierbei:

$$\text{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right)$$

Eigenschaften:

- $-1 \leq \text{Cor}(X, Y) \leq 1$
- $\text{Cor}(X, X) = 1$ und $\text{Cor}(X, -X) = -1$
- Sind X und Y unabhängig voneinander, so gilt:

$$\text{Cor}(X, Y) = 0$$

denn $\text{Cor}(X, Y) = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$ und $E(XY) = E(X)E(Y)$

Lineare Regression

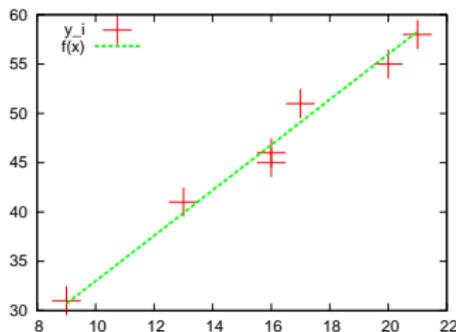
Beispiel: Wir prognostizieren einen linearen Zusammenhang zwischen Futtermenge x und Körpergröße y . Untersuchungen an n Lebewesen liefert nebenstehende Tabelle.

x_i	y_i
13	41
16	46
9	31
17	51
20	55
16	45
21	58

Wir wollen den linearen Zusammenhang

$$f(x) = ax + b$$

so konstruieren, dass die mittlere quadratische Abweichung minimal wird:



$$\sum_{i=1}^n (y_i - f(x_i))^2 \rightarrow \text{Min!} \quad (\text{"least-squares fit"})$$

Die Koeffizienten a und b heißen **Regressionskoeffizienten**.

Bestimmung der Regressionskoeffizienten

Die "optimale" Gerade

$$f(x) = ax + b$$

im Sinne der "kleinsten Quadrate" ist charakterisiert durch:

$$a = \frac{\text{Cov}(x, y)}{\text{Var}(x)} \quad (\text{Steigung})$$

$$b = \bar{y} - a\bar{x} \quad (\text{Konstante})$$

Unser Beispiel: $\bar{x} = 16$, $\bar{y} \approx 46.71$, $\text{Var}(x) = 14.29$, $\text{Cov}(x, y) = 31.57$:

$$a = 2.21 \quad , \quad b = 11.36$$

Begründung zur linearen Korrelation

Existiert solch ein linearer Zusammenhang, so gilt nach der Regressionsrechnung:

$$\begin{aligned}y_i &= f(x_i) = ax_i + \bar{y} - a\bar{x} = a(x_i - \bar{x}) + \bar{y} \\ &= \frac{\text{Cov}(x, y)}{\text{Var}(x)}(x_i - \bar{x}) + \bar{y}\end{aligned}$$

Und somit

$$\frac{\text{Cov}(x, y)^2}{\text{Var}(x)^2}(x_i - \bar{x})^2 = (y_i - \bar{y})^2$$

Summation über i ergibt

$$\frac{\text{Cov}(x, y)^2}{\text{Var}(x)^2} \text{Var}(x) = \text{Var}(y)$$

bzw.

$$\rho(x, y)^2 = \frac{\text{Cov}(x, y)^2}{\text{Var}(x) \text{Var}(y)} = 1$$

Konfidenzintervalle für Regressionskoeffizienten

Frage: Wie verlässlich sind die ermittelten Regressionskoeffizienten a und b ?

- Seien \hat{a} und \hat{b} die tatsächlichen Koeffizienten.
- Residuen $\epsilon_i = y_i - f(x_i)$
- Erwartungstreuer Schätzer für deren Varianz:

$$S_n^2 = \frac{1}{n-2} \sum_{i=1}^n \epsilon_i^2$$

- Dann ist die ZV

$$T = (\hat{a} - a) \sqrt{\frac{n \operatorname{Var}(x)}{S_n^2}}$$

t -verteilt mit $n - 2$ Parametern.

- Bestimmung des Konfidenzintervalls $[-t_\alpha, t_\alpha]$ für T .
- Reskalierung:

$$I_\alpha = [a - x_\alpha, a + x_\alpha] \quad \text{mit } x_\alpha = t_\alpha \sqrt{\frac{S_n^2}{n \operatorname{Var}(x)}}$$

ist das entsprechende Konfidenzintervall für \hat{a} :

$$P(\hat{a} \in I_\alpha) = 1 - \alpha$$

Beispiel: Konfidenzintervalle für Regressionskoeffizienten

- Zunächst berechnen wir die Residuen:

x_i	y_i	$f(x_i)$	ϵ_i
13	41	40.09	-0.09
16	46	46.72	-0.72
9	31	31.25	-0.25
17	51	48.93	2.07
20	55	55.56	-0.56
16	45	46.72	-1.72
21	58	57.77	0.23

- In unserem Beispiel ergibt sich das 10%-Intervall für T aus der t_5 Verteilung:
 $t_\alpha = 1.476$.
- Schätzer für die Varianz: $S_n^2 = 1.64$
- Reskalierung:

$$x_\alpha = t_\alpha \sqrt{\frac{S_n^2}{n \operatorname{Var}(x)}} = 1.476 \sqrt{\frac{1.64}{7 \cdot 14.29}} = 0.19$$

- 10%-Konfidenzintervall

$$P(2.02 \leq \hat{a} \leq 2.4) = 0.9$$

Nichtlineare Regression

- Wenn ein nichtlinearer Zusammenhang angenommen wird, z.B.

$$y = B \exp(Ax)$$

so wird zunächst auf ein lineares Regressionsmodell transformiert:

$$u = \ln(y) = \ln(A) + Bx$$

- Aus den Werten y_i werden also zunächst $u_i = \ln(y_i)$ berechnet.
- Führe lineare Regression mit den Wertepaaren (x_i, u_i) durch:

$$u = ax + b$$

- (Näherungsweise) Exponentieller Zusammenhang:

$$y = \exp(ax + b) = B \exp(Ax)$$

mit $B = \exp(b)$ und $A = a$.

Kontingenztafeln und χ^2 -Unabhängigkeitstest

- Zusammenhang zwischen Rauchen der Mutter und Kindstod?

	Raucherin	Nichtraucherin	Σ
Kind gestorben	246	264	510
Kind lebendig	8160	10710	18870
Σ	8406	10974	19380

- Beobachtungen von Merkmalen mit möglichen **Ausprägungen 0 und 1**:
Kindstod x_i

$$\bar{x} = \frac{510}{19380}, \quad \text{Var}(x) = \overline{x^2} - \bar{x}^2 = \bar{x}(1 - \bar{x}) \approx 0.02562$$

Raucherin y_i ,

$$\bar{y} = \frac{8406}{19380}, \quad \text{Var}(y) = \overline{y^2} - \bar{y}^2 = \bar{y}(1 - \bar{y}) \approx 0.2456$$

Covarianz

$$\text{Cov}(x, y) = \overline{xy} - \bar{x}\bar{y} = \frac{10710}{19380} - \frac{510 \cdot 8406}{19380^2} = 0.01141$$

- Korrelationskoeffizient nach Pearson:

$$\rho(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x) \text{Var}(y)}} = \frac{0.01141}{0.07933} = 0.1439$$

- Der Φ -Kontingenzkoeffizient $\Phi(x, y) \in [0, 1]$ wird auch häufig als Maß für den Zusammenhang zwischen zwei betrachteten Größen betrachtet:

$$\Phi(x, y) = \rho(x, y)^2 = 2.1 \cdot 10^{-4}$$

- Ein kleiner Wert für Φ läßt darauf schließen, dass **kein** Zusammenhang besteht.
- **Frage:** Läßt Φ lediglich auf einen **linearen** Zusammenhang schließen ?

Kap. 11: Gewöhnliche Differentialgleichungen

- **Beispiel:** Population in Abhängigkeit der Zeit $t \geq 0$:

Zahl von Individuen: $y = y(t)$

Sterberate: $s = s(t)$

Geburtsrate: $g = g(t)$

Nettozuwanderungsrate: $b = b(t)$

- Zeitliche Änderung:

$$\frac{\partial y(t)}{\partial t} = g(t)y(t) - s(t)y(t) + b(t)$$

- Die Kurzschreibweise einer solchen **gewöhnlichen Differentialgleichung** mit Anfangsbedingungen y_0 lautet:

$$\begin{aligned}y' &= (g - s)y + b \\ y(0) &= y_0\end{aligned}$$

- Die Größe $r = g - s$ heißt **Wachstumsrate**.

- **Fragen:**

- ▶ Gibt es eine Lösung $y = y(t)$?
- ▶ Ist die Lösung eindeutig?
- ▶ Wie lautet (lauten) die Lösung(en)?

Lineare gewöhnliche Differentialgleichungen 1. Ordnung

- Unter einer **linearen gewöhnlichen Differentialgleichung 1. Ordnung** im Raum \mathbb{R} versteht man eine Gleichung der Form:

$$\begin{aligned}y'(t) &= a(t)y(t) + b(t) \quad \text{für } t > 0 \\y(0) &= y_0\end{aligned}$$

- Ist $b = 0$, so spricht man von einer **homogenen** DGL.
- Eine homogenen, lineare DGL Gleichung 1. Ordnung mit konstanten Koeffizienten hat die eindeutige Lösung:

$$y(t) = y_0 \exp(at)$$

- Für $a > 0$ hat man ein exponentielles Wachstum.

Logistische Gleichung

- Einfluß von Wettbewerb, Kämpfe, begrenzte Nahrungsmittel führen auf einen zusätzlichen Einfluß proportional zu y^2 :

$$\begin{aligned}y' &= (g - s)y - ky^2 + b \\ y(0) &= y_0\end{aligned}$$

Diese Gleichung ist nicht mehr linear und heißt **logistische Gleichung**.

- Im homogenen Fall ($b = 0$) und bei t unabhängigen Koeffizienten g, s läßt sich die in dimensionslose Größen umschreiben:

$$\begin{aligned}z' &= z(1 - z) \\ z(0) &= z_0\end{aligned}$$

mit der Variablen $z = z(\tau)$ mittels $\tau = (g - s)t$ und $z = yk/(g - s)$ sowie $z_0 = y_0k/(g - s)$.

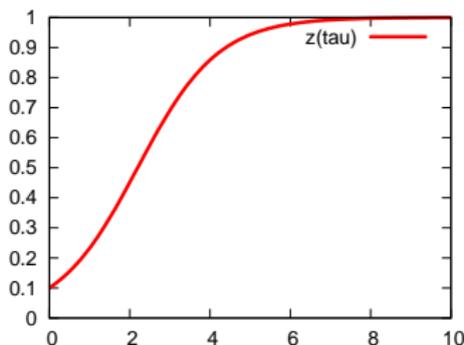
Lösung der logistischen Gleichung

- Man kann zeigen, dass sich die Lösung der logistischen Gleichung ergibt zu

$$z(\tau) = \frac{z_0}{z_0 + e^{-\tau}(1 - z_0)}$$

- Die Funktion y erhält man mittels:

$$y(t) = \frac{r}{k} z(rt)$$



Ermittlung der Wachstumsrate bei logistischen Gleichungen

Situation: Man besitzt experimentelle Daten und möchte hieraus die Wachstumsrate r ermitteln.

t_i	0	9	12	15	18	21	25	27	29	33	36	39
y_i	22	39	105	152	225	390	499	547	618	791	877	938

- Umskalierung:

$$w(t) = \ln\left(\frac{z(t)}{1-z(t)}\right), \quad w_0 = w(0)$$

- Nun gilt die lineare Gleichung:

$$w(t) = rt + w_0,$$

denn mit $\sigma = z_0 + e^{-\tau}(1 - z_0)$

$$\begin{aligned}w(t) &= \ln(z) - \ln(1-z) \\ &= \ln(z_0) - \ln(\sigma) - \ln(e^{-\tau}(1-z_0)) + \ln(\sigma) \\ &= \ln(z_0) - \ln(e^{-\tau}(1-z_0)) \\ &= \ln(z_0/(1-z_0)) + rt\end{aligned}$$

- Ersetze k/r durch eine Schätzung κ und berechne als Approximation an die w_i die Größen \tilde{w}_i :

$$\tilde{w}_i = \ln \left(\frac{\kappa y_i}{1 - \kappa y_i} \right)$$

- zB, für $\kappa = 1/1000$:

t_i	0	9	12	15	18	21	25	27	29	33
y_i	22	39	105	152	225	390	499	547	618	791
\tilde{w}_i	0.0225	0.0406	0.1173	0.179	0.290	0.639	...			

- Berechnung der linearen Regressionskoeffizienten a und b :

$$at_i + b \approx \tilde{w}_i$$

- Setze Wachstumsrate $r = a$ und $k = r\kappa$.

Abbau von Drogen und Medikamenten

Gleichung im Intervall $I = [0, T]$:

Der Abbau von Drogen und Medikamenten im Körper ist proportional zur Konzentration. Dies läßt sich durch eine gewöhnliche Differentialgleichung 1. Ordnung beschreiben:

$$y'(t) = -\alpha y(t), \quad y(0) = y_0 \quad \text{mit } \alpha > 0.$$

Die Lösung hiervon ist einfach:

$$y(t) = y_0 e^{-\alpha t}$$

Der Parameter α hängt zusammen mit der Halbwertszeit T^* :

$$\exp(-\alpha T^*) = \frac{1}{2} \iff T^* = \frac{\ln 0.5}{-\alpha}$$

Ein großes α ergibt also eine kurze Halbwertszeit (schneller Abbau).

Gleichung im allgemeinen Intervall I :

Sind die Anfangswerte y_0 zu einem anderen Zeitpunkt gegeben, $I = [t_0, T]$:

$$y'(t) = -\alpha y(t), \quad y(t_0) = y_0$$

so ergibt sich die Lösung einfach durch Verschiebung:

$$y(t) = y_0 e^{-\alpha(t-t_0)}$$

Trennung der Variablen

Eine nichtlineare Differentialgl. 1. Ordnung im Intervall $I = [0, T]$ kann man i.a. nicht analytisch lösen:

$$y' = f(y), \quad y(0) = y_0$$

Es gibt aber ein paar Techniken, die in manchen Fällen funktionieren:

- Angenommen, die Differentialgleichung ist von der Form:

$$y'(t) = f(t) \cdot g(y), \quad y(0) = y_0$$

- Sofern $g(y) \neq 0$ für alle $y \in I$, ergibt formale Umformung:

$$\begin{aligned} \frac{dy(t)}{g(y)} &= f(t)dt \\ \int \frac{dy(t)}{g(y)} &= \int f(t)dt + c \end{aligned}$$

Diese Gleichung kann man nun versuchen nach y aufzulösen und c gemäß der Anfangsbedingungen y_0 wählen.

- Falls $g(y) = 0$ für ein $y \in I$, so muss dieser Fall gesondert betrachtet werden.

Beispiel 1 (Trennung der Variablen)

Lineare Differentialgleichung 1. Ordnung mit variablen Koeffizienten:

$$y'(t) = (t^2 - t)y(t), \quad y(0) = y_0 \neq 0$$

- Hier gilt also $g(y) = y$ und $f(t) = t^2 - t$, sofern $y \neq 0$.

$$\ln(y) = \int \frac{1}{y} dy = \int (t^2 - t) dt + c = \frac{1}{3}t^3 - \frac{1}{2}t^2 + c$$

- Also

$$y(t) = \exp\left(\frac{1}{3}t^3 - \frac{1}{2}t^2\right) \cdot \exp(c)$$

- Den Koeffizienten c erhält man aufgrund der Anfangsbedingung:

$$c = \ln(y_0)$$

- Also insgesamt:

$$y(t) = y_0 \exp\left(\frac{1}{3}t^3 - \frac{1}{2}t^2\right)$$

Beispiel 2 (Trennung der Variablen)

Nichtlineare Differentialgleichung 1. Ordnung mit variablen Koeffizienten:

$$y'(t) = \frac{3t^2}{2y(t)}, \quad y(0) = y_0$$

- Hier gilt also $g(y) = \frac{1}{2y}$ und $f(t) = 3t^2$, sofern $y \neq 0$.

$$y^2 = \int 2y \, dy = 3 \int t^2 \, dt + c = t^3 + c$$

- Also haben wir die zwei Lösungen

$$y(t) = \pm \sqrt{t^3 + c}$$

wobei $c \geq 0$ gelten muss.

- Den Koeffizienten c erhält man aufgrund der Anfangsbedingung:

$$y_0 = y(0) = \pm \sqrt{c}$$

- Insgesamt ist die Lösung abhängig vom Vorzeichen des Anfangswertes y_0 :

$$y_0 > 0 \Rightarrow y(t) = \sqrt{t^3 + y_0^2}$$

$$y_0 < 0 \Rightarrow y(t) = -\sqrt{t^3 + y_0^2}$$

Lineare DGL 1. Ordnung mit variablen Koeffizienten

Die Methode der Trennung der Variablen liefert uns die Lösung von lineare Differentialgleichung mit variablen Koeffizienten:

$$y'(t) = a(t)y(t), \quad y(0) = y_0 \neq 0$$

Die Lösung lautet:

$$y(t) = y_0 \exp\left(\int_0^t a(x)dx\right)$$

Begründung:

- Trennung der Variablen ergibt (mit einem noch zu bestimmenden c):

$$\ln(y) = \int \frac{1}{y} dy = \int_0^t a(x)dx + c$$

- Unter Berücksichtigung der Anfangswerte y_0 ergibt sich obige Lösung.

Bemerkung:

Im Fall von konstanten Koeffizienten $a(t) = a$ erhalten wir die Lösung wie zuvor, da

$$\int_0^t a dx = at$$

Inhomogene lineare DGL 1. Ordnung

$$y'(t) = a(t)y(t) + b(t), \quad y(0) = y_0$$

- Zur Lösung der inhomogenen Gleichung betrachten wir zunächst die Lösung $w(t)$ einer zugehörigen homogenen Gleichung, also

$$w(t)' = a(t)w(t), \quad w(0) = 1$$

- Dann ist die Lösung der inhomogenen Gleichung eindeutig gegeben durch:

$$y(t) = w(t) \left(y_0 + \int_0^t \frac{b(x)}{w(x)} dx \right)$$

- **Variation der Konstanten** Man sieht dies mittels des Ansatzes $y(t) = w(t)c(t)$ mit geeigneter Funktion $c(t)$:

$$y' = w'c + wc' = awc + wc' = ay + wc'$$

- Es muss also gelten $c' = b/w$ bzw. $c(t) = y_0 + \int_0^t b(x)/w(x) dx$.
- w ist gegeben durch $w(t) = \exp(\int_0^t a(x) dx)$.

Beispiel einer inhomogenen linearen DGL 1. Ordnung

Beispiel:

$$y' = 2ty + \exp(t^2), \quad y(0) = y_0$$

- Lösung der homogenen Gleichung $w' = 2tw$:

$$w(t) = \exp\left(\int_0^t 2x \, dx\right) = \exp(t^2)$$

- Die "Konstante" ergibt sich zu

$$\begin{aligned} c(t) &= y_0 + \int_0^t \frac{\exp(x^2)}{w(x)} dx = y_0 + \int_0^t \exp(x^2) \exp(-x^2) dx \\ &= y_0 + \int_0^t e^0 dx = y_0 + t \end{aligned}$$

- Also insgesamt:

$$y(t) = (y_0 + t) \exp(t^2)$$

Biologisches Beispiel

Situation: In einer Zelle sei s die Konzentration einer gewissen chemischen Substanz und s^* sei die Konzentration der Umgebung.

Frage: Wie ändert sich s durch Diffusion durch die Zellmembran?

Model:

- Der Stofftransport Δm in einem Zeitintervall Δt ist proportional zu Oberfläche F der Zelle, zu Δt und dem Konzentrationsunterschied $s - s^*$:

$$\Delta m = \alpha F \Delta t (s - s^*)$$

- Die Konzentrationsänderung in der Zelle ist dann umgekehrt proportional zum Volumen der Zelle:

$$\Delta s = -\frac{\Delta m}{V}$$

- Wir erhalten also

$$\frac{\Delta s}{\Delta t} = -\alpha \frac{F}{V} (s - s^*)$$

- Der Grenzübergang $\Delta t \rightarrow 0$ ergibt die lineare inhomogene DGL:

$$\frac{\partial s(t)}{\partial t} = s'(t) = -\alpha \frac{F}{V} (s(t) - s^*)$$

Beispiel: Stimulation von Nervenzellen (1)

- **Reiz / Schmerz** falls

$$\frac{\text{Konzentration an **stimulierenden** Ionen}}{\text{Konzentration an **hemmenden** Ionen}} = \frac{k_1(t)}{k_2(t)} \geq \sigma$$

- Bei konstanten Stromfluß I durch Zelle und konstanten Umgebungskonzentrationen k_i^* :

$$k_i(t)' = b_i I - q_i(k_1(t) - k_i^*)$$

- Lösungen:

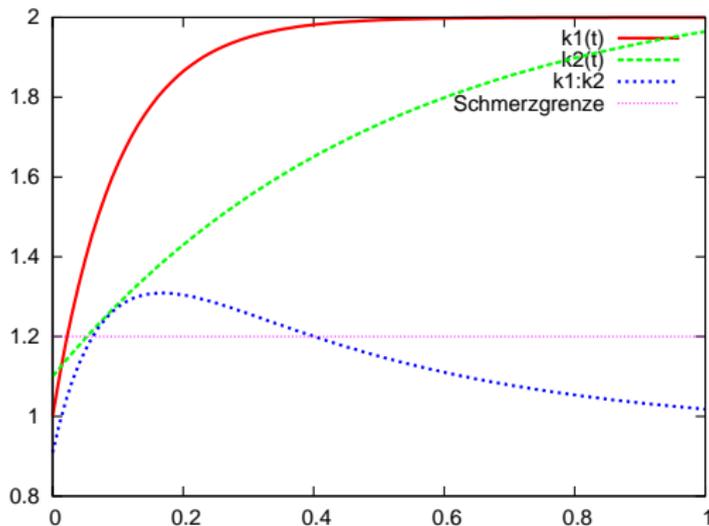
$$k_i(t) = \alpha_i + k_i^* + c_i e^{-q_i t}, \quad \alpha_i = \frac{b_i I}{q_i}$$

mit Konstanten c_i zur Erreichung der jeweiligen Anfangswerte.

- Speziell bei $k_i(0) = k_i^*$:

$$k_i(t) = k_i^* + \alpha_i(1 - e^{-q_i t})$$

Beispiel: Stimulation von Nervenzellen (2)



- Rote Kurve: stimulierenden Ionen k_1
- Grüne Kurve: hemmenden Ionen k_2
- Blaue Kurve: Verhältnis k_1/k_2
- Pinke Linie: Schmerzgrenze

Sterbegesetz nach Gompertz (1)

Das Wachstum von **Tumoren** lässt sich häufig besser durch die DGL von Gompertz beschreiben:

$$y'(t) = -ry(t) \ln\left(\frac{y(t)}{K}\right), \quad y(0) = y_0$$



Der zusätzliche Faktor $-\ln(y/K)$ führt hierbei

- für $y \ll K$: zu einem noch schnelleren Wachstum
- für $y \approx K$: zu einer gewissen Sättigung.

Die Lösung lautet:

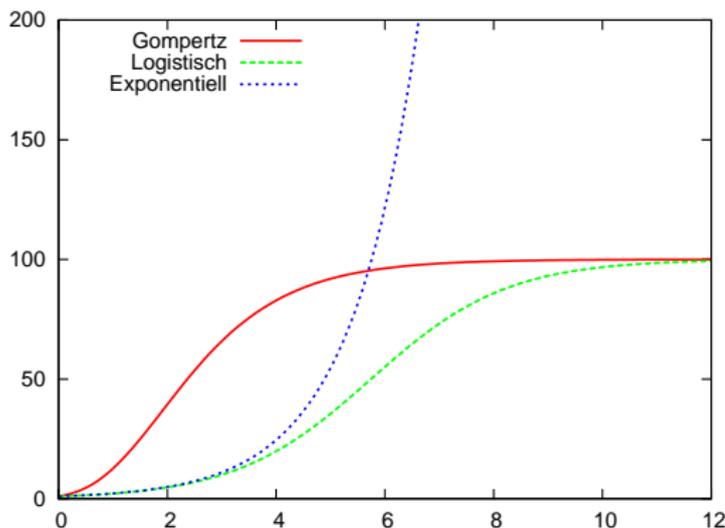
$$y(t) = K \exp(\beta \exp(-rt))$$

mit $\beta = \ln(y_0/K)$.

Sterbegesetz nach Gompertz (2)

Vergleich zwischen Lösungen der Gleichung von Gompertz, der logistischen Gleichung und reinem exponentiellem Wachstum:

- Bei kleinen Werten wächst Gompertz schneller als logistisch und exponentiell.
- Die Sättigung setzt bei Gompertz aber schneller ein.
- Die Lösung der linearen Gleichung wächst gegen ∞ .



- Wärmeleitung kann durch eine **lineare, inhomogene DGL 1. Ordnung** beschrieben werden. Die zeitliche Änderung der Temperatur ist proportional zur Oberfläche A und zur Differenz der Umgebungstemperatur: $T - T^*$

$$T'(x) = -kA \cdot (T - T^*), \quad T(0) = T_0$$

- Lösung der homogenen Gleichung $w' = -kAw$: $w(t) = \exp(-kAt)$
- Lösungsansatz: $T(t) = w(t)c(t)$ mit

$$\begin{aligned} c(t) &= T_0 + \int_0^t \frac{kAT^*}{w(x)} dx = T_0 + kAT^* \int_0^t e^{kAx} dx \\ &= T_0 + T^* \int_0^{kAt} e^y dy = T_0 + T^*(e^{kAt} - 1) \end{aligned}$$

- Die Lösung lautet daher:

$$\begin{aligned} T(t) &= e^{-kAt}(T_0 + T^*(e^{kAt} - 1)) \\ &= T^*(1 - e^{-kAt}) + T_0 e^{-kAt} \\ &= T^* + (T_0 - T^*)e^{-kAt} \end{aligned}$$

Freier Fall mit Reibung

- Ohne Reibung ist die zeitliche Änderung der Geschwindigkeit v eines Objektes im freien Fall ist gerade die (Erd-) Beschleunigung: $v' = g$.
- Durch den Effekt der Reibung wird die Geschwindigkeit allerdings abgebremst:

$$v'(t) = g - \gamma v^n$$

mit Reibungskonstante $\gamma \geq 0$ und Exponent $n \geq 1$.

- Maximalgeschwindigkeit:

$$v_\infty = \sqrt[n]{g/\gamma}$$

- $n = 1$: v klein, ausgedehnte Körper mit geringer Dichte, zähe Flüssigkeit:

$$v(t) = v_\infty(1 - e^{-\gamma t})$$

- $n = 2$: v groß, hohe Dichte

Verschmutzung eines Sees

- $u(t)$ = Verschmutzungskonzentration eines Sees mit Volumen V .
- u^* = Verschmutzungskonzentration des einfließenden Wassers mit Rate r .
- k = biologische Abbaurrate der Verschmutzung.
- **Annahme:** Regenmenge = verdunstete Menge, unmittelbare Durchmischung.
- Lineares Modell:

$$u'(t) = (u^* - u(t))\frac{r}{V} - ku(t), \quad u(0) = u_0$$

bzw. mit $\alpha = r/V + k$ und $\beta = u^*r/V$:

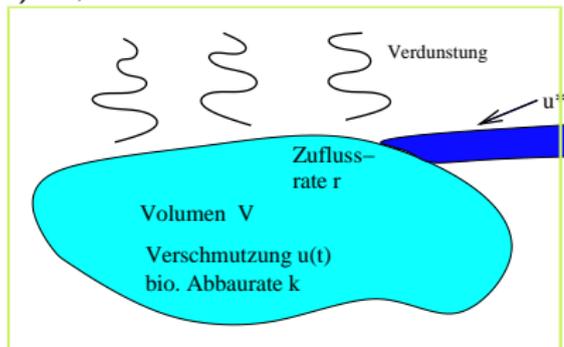
$$u'(t) = -\alpha u(t) + \beta$$

- **Lösung:**

$$u(t) = \frac{\beta}{\alpha} + e^{-\alpha t}\left(u_0 - \frac{\beta}{\alpha}\right)$$

also langfristige Schadstoffkonzentration

$$\lim_{t \rightarrow \infty} u(t) = \frac{\beta}{\alpha} = u^*(1 + Vk/r)^{-1}$$



Existenz und Eindeutigkeit von Lösungen von DGL

- **Lineare** Differentialgleichungen der Form

$$u'(t) = a(t)u(t) + b(t), \quad u(0) = u_0$$

haben stets eine **eindeutige** Lösung. Diese erhält man mit Hilfe der Trennung der Variablen:

$$w(t) = \exp\left(\int_0^t a(x)dx\right)$$

$$u(t) = w(t) \left(u_0 + \int_0^t \frac{b(x)}{w(x)} dx \right)$$

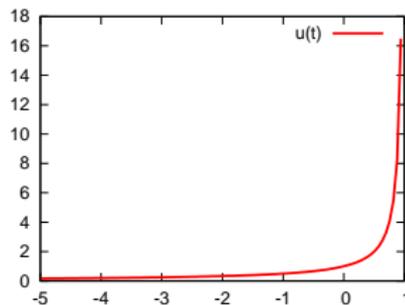
- **Nichtlineare** Differentialgleichungen haben im allgemeinen keine globale Lösung.

Beispiel:

$$u'(t) = u(t)^2, \quad u(0) = u_0 > 0$$

hat nur eine Lösung im Intervall $(-\infty, 1/u_0)$:

$$u(t) = \frac{u_0}{1 - u_0 t}$$



Biologisches Beispiel: Glukose Toleranztest

G	=	Blutzuckerspiegel mit Gleichgewichtszustand G^*
H	=	Hormonspiegel mit Gleichgewichtszustand H^*
g	= $G - G^*$	= Abweichung des Blutzuckerspiegels
h	= $H - H^*$	= Abweichung des Hormonspiegels
I	=	Rate extern zugeführter Glukose Dosis

Näherungsmodell für GTT:

$$\begin{aligned}g'(t) &= -ag(t) - bh(t) + I(t) \\h'(t) &= cg(t) - dh(t)\end{aligned}$$

$a, b, c, d > 0$, große Werte für $a, b, c, d \rightarrow$ schnelle Reaktion des Körpers
Indikator für **Diabetes**:

$$ad + bc$$

Typische Reaktionszeit des Körpers \sim einige Std.
Zufuhr von $I(t) \sim$ einige Minuten.

Frage: Wie löse ich ein solches **System** von gewöhnlichen linearen DGL?

Systeme von Differentialgleichungen

Wir betrachten zunächst das folgende lineare inhomogene System:

$$\begin{aligned}u_1'(t) &= a_{11}u_1(t) + a_{12}u_2(t) + b_1, & u_1(0) &= u_1^0 \\u_2'(t) &= a_{21}u_1(t) + a_{22}u_2(t) + b_2, & u_2(0) &= u_2^0\end{aligned}$$

oder kurz in Matrixschreibweise mit $u = (u_1, u_2, \dots, u_n)^T$:

$$u'(t) = Au(t) + b(t), \quad u(0) = u^0$$

Auch hier ergibt sich die Lösung durch Kombination von den Lösungen des **homogenen** Systems und dem inhomogenen Anteil $b(t)$.

Matrixexponentialfunktion

Zur Lösung von Systemen von Differentialgleichungen verwenden wir die Matrixexponentialfunktion:

Definition: Ist $A \in \mathbb{R}^{n \times n}$, so versteht man unter der **Matrixexponentialfunktion** die $n \times n$ -Matrix

$$\exp(A) = e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Wie bei den reellen Zahlen ist diese unendliche Reihe immer konvergent.

Beispiel für die Matrixexponentialfunktion

- Als Beispiel für die Matrixexponentialfunktion betrachten wir

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \quad \text{mit } a = 2.$$

- So gilt für das Produkt A^2 und allgemein A^k :

$$A^2 = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} = \begin{pmatrix} a^2 & 2a \\ 0 & a^2 \end{pmatrix}$$

$$A^k = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}^k = \begin{pmatrix} a^k & ka^{k-1} \\ 0 & a^k \end{pmatrix}$$

- Hieraus ergibt sich nun:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} a^k & ka^{k-1} \\ 0 & a^k \end{pmatrix} = \begin{pmatrix} b & c \\ 0 & b \end{pmatrix}$$

mit $b = \sum_{k=0}^{\infty} \frac{1}{k!} a^k = e^a$ und $c = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} a^{k-1} = e^a$.

- Wir erhalten damit:

$$\exp(A) = e^2 \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Ableitung der Matrixexponentialfunktion

Die Matrix-wertige Funktion

$$t \rightarrow \exp(tA)$$

ist stetig differenzierbar und es gilt

$$\frac{d}{dt} \exp(tA) = \lim_{h \rightarrow 0} \frac{1}{h} (\exp((t+h)A) - \exp(tA)) = A \exp(tA) \in \mathbb{R}^{n \times n}$$

Beispiel: Wir betrachten $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Dann gilt (siehe Übungsaufgabe):

$$\exp(tA) = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

Damit erhalten wir:

$$\frac{d}{dt} \exp(tA) = A \exp(tA) = e^{2t} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = e^{2t} \begin{pmatrix} 2 & 2t+1 \\ 0 & 2 \end{pmatrix}$$

Lösung eines homogenen linearen Systems von DGL

Die Lösung eines homogenen linearen Systems von Differentialgleichungen der Form

$$u'(t) = Au(t), \quad u(0) = u^0$$

mit einer von t -unabhängigen Matrix $A \in \mathbb{R}^{n \times n}$ ergibt sich zu

$$u(t) = \exp(tA)u^0$$

denn $u(0) = \exp(0A)u^0 = u^0$ und

$$u'(t) = \frac{d}{dt}(\exp(tA))u^0 = A \exp(tA)u^0 = Au(t)$$

Beispiel zur Lösung eines homogenen Systems von DGL

- Das System

$$\begin{aligned}u_1'(t) &= 2u_1(t) + u_2(t), & u_1(0) &= u_1^0 \equiv 1 \\u_2'(t) &= 2u_2(t), & u_2(0) &= u_2^0 \equiv 1\end{aligned}$$

kann mit der bereits betrachteten Matrix $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ beschrieben werden.

- Damit ergibt sich die Lösung zu

$$\begin{aligned}\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} &= \exp(tA) \begin{pmatrix} u_1^0 \\ u_2^0 \end{pmatrix} = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_1^0 \\ u_2^0 \end{pmatrix} \\ &= e^{2t} \begin{pmatrix} u_1^0 + tu_2^0 \\ u_2^0 \end{pmatrix} = e^{2t} \begin{pmatrix} 1 + t \\ 1 \end{pmatrix}\end{aligned}$$

- Also

$$u_1(t) = (1 + t)e^{2t} \quad \text{und} \quad u_2(t) = e^{2t}$$

Lösung eines inhomogenen linearen Systems von DGL

Die Lösung eines inhomogenen linearen Systems von Differentialgleichungen der Form

$$u'(t) = Au(t) + b(t), \quad u(0) = u^0$$

mit einer von t -unabhängigen Matrix $A \in \mathbb{R}^{n \times n}$ ergibt sich analog zum skalaren Fall zu

$$u'(t) = \exp(tA) \left(u^0 + \int_0^t \exp(-\tau A) b(\tau) d\tau \right)$$

Auch hier ist die Lösung stets wohldefiniert und eindeutig.

Bemerkungen:

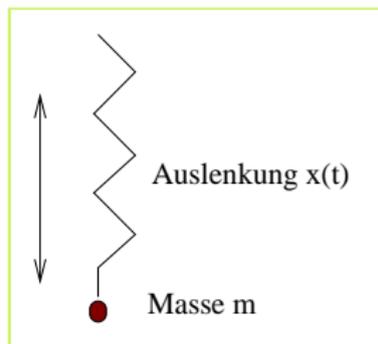
- Man prüfe einmal nach, dass dies tatsächlich eine Lösung ist. Die Eindeutigkeit ergibt sich unmittelbar aus der Eindeutigkeit des homogenen Systems.
- Wir werden im Rahmen dieser Vorlesung nicht näher auf diese inhomogenen linearen Systems von DGL eingehen.

Lineare DGL 2. Ordnung

Als Paradebeispiel einer Differentialgleichung 2. Ordnung eignet sich die Bewegung eines Gegenstandes (Massenpunkt) an einer Feder:

$$mx''(t) = -kx(t).$$

Hierbei bezeichnet m die Masse des Körpers, $x(t)$ die Position zum Zeitpunkt t und k die Federsteifigkeit.



- Die Rückstellkraft hat entgegengesetztes Vorzeichen zur Auslenkung.
- Hierdurch wird eine ungedämpfte harmonische Schwingung beschrieben.
- Wir benötigen noch weitere Anfangsbedingungen, z.B. für $x(0)$ und $x'(0)$.
- Im Fall einer Bewegung im zähen Medium kommt ein Reibungsterm hinzu:

$$mx''(t) = -rx'(t) - kx(t) \quad \text{mit } r > 0.$$

- Im Fall einer Kraft $K(t)$ von außen:

$$mx''(t) = -rx'(t) - kx(t) + K(t).$$

Ungedämpfte harmonische Schwingung

Wir betrachten die ungedämpfte harmonische Schwingung mit Anfangswerten für die Auslenkung und die Geschwindigkeit

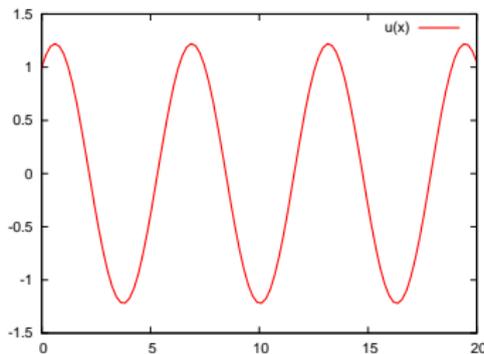
$$u''(t) + bu(t) = 0$$

$$u(0) = u_0$$

$$u'(0) = u'_0$$

mit $b > 0$. Die Lösung hierzu lautet:

$$u(t) = u_0 \cos(\sqrt{b}t) + \frac{u'_0}{\sqrt{b}} \sin(\sqrt{b}t)$$



Räuber-Beute Modell

- In der Biologie ist das System von linearen DGL zur Beschreibung von Räubern x und ihrer Beute y sehr populär (Konstanten $c, d, e, g > 0$):

$$y' = d - ex$$

$$x' = c(y - g)$$

- Dieses System kann einfach in ein skalares System 2. Ordnung überführt werden. Hierzu differenzieren wir die erste Gleichung nach t :

$$y'' = -ex'$$

und setzen die zweite Gleichung ein:

$$y'' = -ec(y - g)$$

- Das Räuber-Beute Modell ist somit auch beschreibbar durch eine ungedämpfte harmonische Schwingung mit "Rückstellkraft" $a = ec$ und "externer Kraft" $b = -ecg$:

$$y'' + ay + b = 0$$

Überführung in ein System 1. Ordnung

Ganz allgemein kann eine skalare, nichtlineare DGL n -ter Ordnung

$$y^{(n)} = f(t, y, y', \dots, y^{(n-1)})$$

mit den Anfangswerten $y(0) = y_0, y'(0) = y'_0, \dots, y^{(n-1)}(0) = y_0^{(n-1)}$ stets überführt werden in ein System 1. Ordnung:

$$\begin{aligned}y_1' &= y_2 \\y_2' &= y_3 \\&\vdots \\y_n' &= f(t, y_1, \dots, y_n)\end{aligned}$$

Ist f linear, so existiert immer genau eine Lösung.

Kap. 12: Funktionen mehrerer Veränderlicher

Funktionen mehrerer Veränderlicher

Ist $\Omega \subset \mathbb{R}^n$, so ist eine Funktion $f : \Omega \rightarrow \mathbb{R}$ eine Funktion mehrerer Veränderlicher:

$$z = f(x_1, \dots, x_n)$$

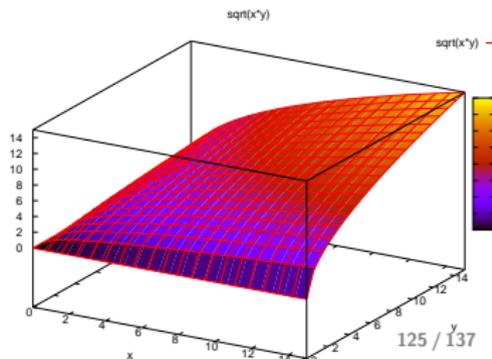
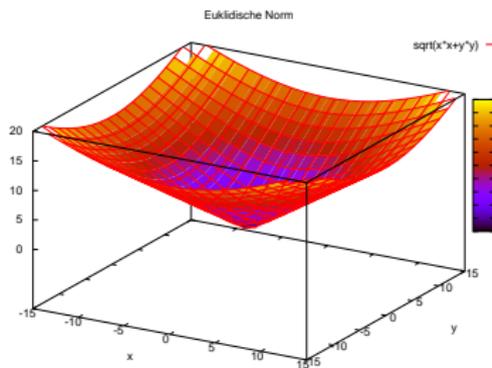
Beispiele:

- Für $n = 2$ und $\Omega = (\mathbb{R}_0^+)^2$:

$$f(x, y) = \sqrt{xy}$$

- Die euklidische Norm im \mathbb{R}^n :

$$f(x) = \|x\| = \sqrt{x_1^2 + \dots + x_n^2}$$



Partielle Ableitungen

Definition: Unter den **partiellen Ableitungen** einer Funktion $f(x, y)$ versteht man die Grenzwerte

$$\frac{\partial f(x)}{\partial x} = \partial_x f(x) = \lim_{h \rightarrow 0} (f(x + h, y) - f(x))$$

$$\frac{\partial f(x)}{\partial y} = \partial_y f(x) = \lim_{h \rightarrow 0} (f(x, y + h) - f(x))$$

sofern diese existieren. Im allgemeinen für $f : \Omega \rightarrow \mathbb{R}$ mit $\Omega \subset \mathbb{R}^n$:

$$\frac{\partial f(x)}{\partial x_i} = \lim_{h \rightarrow 0} (f(x + h e_i) - f(x))$$

wobei e_i den i -ten Einheitsvektor bezeichnet.

Beispiel:

Die euklidische Norm $f(x) = \|x\|$ ist für $x \neq 0$ partiell differenzierbar:

$$\frac{\partial f(x)}{\partial x_i} = \frac{x_i}{\|x\|}$$

Im Nullpunkt ist diese Funktion nicht partiell differenzierbar.

Partiell differenzierbar $\not\Rightarrow$ Stetigkeit

Im Gegensatz zum eindimensionalen Fall folgt aus der partiellen Differenzierbarkeit **nicht** notwendigerweise die Stetigkeit.

Sind hingegen die partiellen Ableitungen stetig, so ist auch die Funktion selbst stetig.

Beispiel: Partiiell differenzierbar \nrightarrow Stetigkeit

$F : \mathbb{R} \rightarrow \mathbb{R}$, definiert als $F(0) = 0$ und für $x \neq 0$ durch

$$F(x) = \frac{x_1 \cdot \dots \cdot x_n}{\|x\|^{2n}}.$$

- F ist auf ganz \mathbb{R}^n partiell differenzierbar, zB. für $x \neq 0$:

$$\frac{\partial F(x)}{\partial x_1} = \frac{x_2 \cdot \dots \cdot x_n}{\|x\|^{2n}} + x_1 \cdot \dots \cdot x_n \frac{\partial}{\partial x_1} \left(\frac{1}{\|x\|^{2n}} \right)$$

Da $\frac{\partial}{\partial x_1} (\|x\|^{-2n}) = -2nx_1 \|x\|^{-(2n+2)}$ folgt

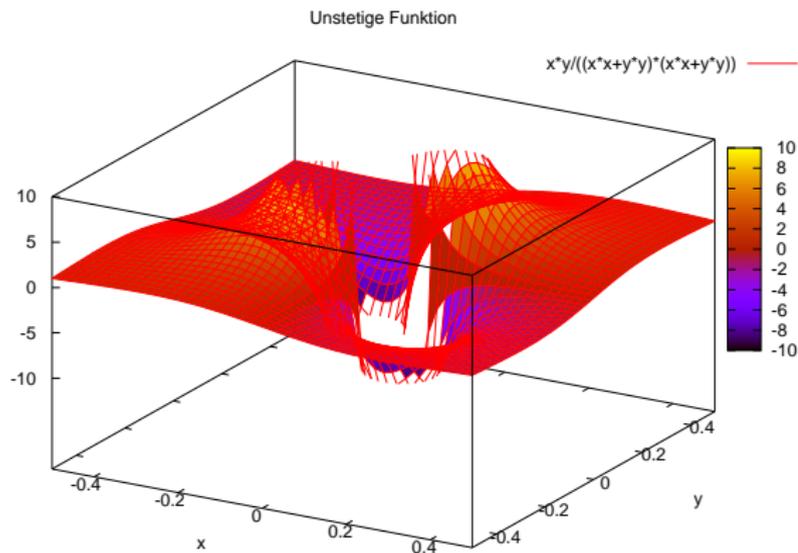
$$\frac{\partial F(x)}{\partial x_1} = \frac{x_2 \cdot \dots \cdot x_n}{\|x\|^{2n}} (1 - 2nx_1^2 \|x\|^{-2})$$

Und im Nullpunkt wegen $F(he_i) = 0$ für beliebiges h :

$$\partial_{x_i} F(0) = \lim_{h \rightarrow 0} (F(he_i) - F(0)) = 0.$$

- F ist aber im Nullpunkt nicht stetig.

Beispiel: Partiiell differenzierbar aber unstetig



$$F(x) = \frac{x_1 \cdot \dots \cdot x_n}{\|x\|^{2n}}$$

Gradienten

Definition: Unter dem **Gradienten** einer partiell differenzierbaren Funktion $f : \Omega \rightarrow \mathbb{R}$ versteht man den Vektor

$$\nabla f(x) = \begin{pmatrix} \partial_{x_1} f(x) \\ \vdots \\ \partial_{x_n} f(x) \end{pmatrix}$$

Beispiel: Der Gradient der euklidischen Norm $f(x) = \|x\|$ ergibt sich zu

$$\nabla f(x) = \frac{1}{f(x)} x = \frac{1}{f(x)} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Lokale Extrema

Definition: Unter einem **lokalen Minimum** einer Funktion $f : \Omega \rightarrow \mathbb{R}$ versteht man einen Punkt $x \in \Omega$, so dass

$$f(x) \leq f(y) \quad \forall y \in \mathcal{U}(x),$$

wobei $\mathcal{U}(x) \subset \Omega$ eine beliebig kleine “Umgebung” von x ist. Entsprechend ist ein **lokales Maximum** ein Punkt $x \in \Omega$ mit

$$f(x) \geq f(y) \quad \forall y \in \mathcal{U}(x).$$

Ein **lokales Extremum** ist ein lokales Minimum oder Maximum.

Beispiel: Die euklidische Norm $f(x) = \|x\|$ hat im Nullpunkt ein lokales Minimum.

Notwendige Bedingung lokaler Extrema

Satz: Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}$ partiell differenzierbar in Ω . Eine notwendige Bedingung eines lokalen Extremums im Punkt $x \in \Omega$ von f ist

$$\nabla f(x) = 0.$$

Beispiele:

- Paraboloid:

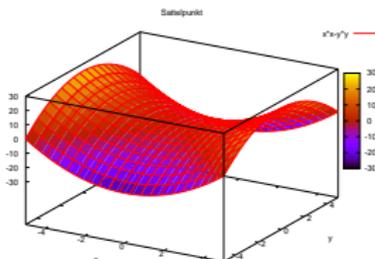
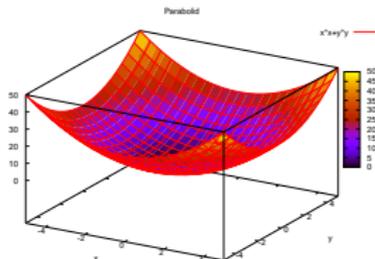
$$f(x, y) = x^2 + y^2, \quad \nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

Gradient verschwindet im Nullpunkt, $\nabla f(0, 0) = 0$.
Hier liegt ein lokales Minimum vor.

- Sattelpunkt:

$$f(x, y) = x^2 - y^2, \quad \nabla f = \begin{pmatrix} 2x \\ -2y \end{pmatrix}$$

Gradient verschwindet ebenfalls im Nullpunkt,
 $\nabla f(0, 0) = 0$. Hier liegt aber **kein** lokales
Extremum, sondern ein sogenannter **Sattelpunkt**.



Höhere Ableitungen

Analog zum eindimensionalen Fall sind höhere Ableitungen definiert, zB.:

$$f(x, y) = x^2 y^2$$
$$\frac{1}{\partial x} \left(\frac{\partial f(x, y)}{\partial x} \right) = \frac{\partial^2 f(x, y)}{\partial x^2} = 2y^2$$

sowie gemischte Ableitungen:

$$\frac{1}{\partial y} \left(\frac{\partial f(x, y)}{\partial x} \right) = 4xy$$
$$\frac{1}{\partial x} \left(\frac{\partial f(x, y)}{\partial y} \right) = 4xy$$

Es gilt allgemein bei **stetigen partiellen Ableitungen**, dass die Ableitungsreihenfolge egal ist:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{1}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{1}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x}$$

Kettenregel

$$x : \mathbb{R} \rightarrow \mathbb{R}^n, \quad x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$
$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

Satz: x sei differenzierbar im Punkt $t \in \mathbb{R}$ und f sei im Punkt $x(t) \in \mathbb{R}^n$ stetig partiell differenzierbar. Dann ist auch die Abbildung $f \circ x$ im Punkt t differenzierbar mit der Ableitung:

$$\frac{d}{dt} f(x(t)) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x(t)) \cdot \frac{\partial x_i}{\partial t}(t) = (\nabla f(x))^\top x'(t)$$

Wichtig ist hierbei also, dass die partiellen Ableitungen im Punkt $x(t)$ allesamt stetig sind.

Beispiel für die Kettenregel

- Sei

$$\begin{aligned}x(t) &= \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix} \\ f(x) &= \|x\|\end{aligned}$$

- Dann gilt

$$\begin{aligned}\frac{d}{dt} f(x(t)) &= \frac{\partial f}{\partial x_1}(x(t)) \cdot \frac{\partial x_1}{\partial t}(t) + \frac{\partial f}{\partial x_2}(x(t)) \cdot \frac{\partial x_2}{\partial t}(t) \\ &= \frac{x_1}{\|x\|}(-\cos(t)) + \frac{x_2}{\|x\|} \sin(t) \\ &= \frac{1}{\|x\|}(-\sin(t) \cos(t) + \cos(t) \sin(t)) = 0\end{aligned}$$

- Dies erwartet man auch, denn $x(t)$ beschreibt eine Kreisbahn und $f(x)$ den Abstand des Punktes zum Nullpunkt. Dieser bleibt aber stets gleich 1.

Biologisches Beispiel für die Kettenregel

- **Situation:** Östrogengehalt c im Futtermittel habe einen (positiven) Effekt auf das mittlere Gewicht eines Huhns:

$$w(c) = \alpha_0 + \alpha_1 c - \alpha_2 c^2, \quad \alpha_0, \alpha_1, \alpha_2 > 0.$$

- Andererseits nehme die Fleischqualität ab:

$$q(c) = q_0 - q_1 c^2, \quad q_0, q_1 > 0.$$

- Enderlös sei $P(w, q) = w \cdot q$.

- **Frage:** Ist eine Erlössteigerung möglich durch kleine Änderungen des Östrogengehalts?

$$\begin{aligned} \frac{d}{dc} P(w(c), q(c)) &= \frac{\partial P}{\partial w} \cdot \frac{\partial w}{\partial c} + \frac{\partial P}{\partial q} \cdot \frac{\partial q}{\partial c} \\ &= q(\alpha_1 - 2\alpha_2 c) + w(-2q_1 c) \\ &= q\alpha_1 - c(q\alpha_2 + w2q_1) \end{aligned}$$

- Im speziellen Fall $\alpha_0 = 100, \alpha_1 = 1, \alpha_2 = 0.01, q_0 = 28, q_1 = 0.01, c = 20$:

$$w(c) = 116 \quad q(c) = 24 \quad \frac{dP}{dc}(w(c), q(c)) = -27.2$$

Eine Verminderung des Östrogengehalts verspricht also eine Erlössteigerung.

Klausurrelevanter Stoff

Insbesondere solle man sich anschauen:

- **Begriffe:** Erwartungswert, Varianz einer ZV
- **Diskrete Verteilungen:** Binomial-Verteilung, geometrische Verteilung, Poisson Verteilung, Multinomial-Verteilung mit jeweils einem Anwendungsbeispiel
- **Kontinuierliche Verteilungen:** Gleichvtlg., Normalvtlg., Exponentialvtlg.
- Ungleichung von Tschebyscheff
- Zentraler Grenzwertsatz: $Var(\bar{X}) = \sigma^2/n$
- Sinngemäß: Gesetz der großen Zahlen
- Konfidenzintervalle, erwartungstreuer Schätzer für die Varianz.
- t -, χ^2 -Verteilung
- Regression, Korrelation
- Inhomogene lineare DGL 1. Ordnung mit variablen Koeffizienten

Irrelevant für Klausur: Wilcoxon Regressionstest, Konfidenzintervalle für Regressionskoeffizienten, logistische DGL