Grundbegriffe der Statistik

- Datenmenge
- Mittelwert

  → Durchschnitt oder arithmetisches Mittel
- Varianz

  →Streuung um den Mittelwert

  $$\mathbf{var}(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$$
- Standardabweichung

  → Wurzel aus der Varianz, mittlerer quadratischer Fehler der Einzelwerte
- Kovarianz

  →Maßzahl für den Zusammenhang zweier Werte X und Y

Grundbegriffe der Stochastik

- Statistik ←→ Stochastik
- Ergebnisse (Ω)
- Ereignisraum (A)
- Wahrscheinlichkeit (P)

→ weitere Darstellungsformen:

- – Verteilungsfunktion ($F(x)$)
- – Wahrscheinlichkeitsdichte ($f(x)$)

Grundlagen der Stochastik (I)

- Die Axiome von Kolmogorov:

(1) Axiom: $P(A) \geq 0$ für jedes Ereignis A

(2) Axiom: $P(\Omega) = 1$

(3) Axiom: $P(A_1 \cup A_2) = P(A_1) + P(A_2)$

für A_1 und A_2 so, dass $A_1 \cap A_2 = \emptyset$

Grundlagen der Stochastik (II)

- Bedingte Wahrscheinlichkeit

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{falls} \quad P(B) > 0$$

- Stochastische Unabhängigkeit

$$\begin{aligned} & P(A \cap B) &=& P(A)P(B) \\ \Leftrightarrow \quad & P(A|B) &=& P(A), \quad \text{falls } P(B) > 0 \\ \Leftrightarrow \quad & P(B|A) &=& P(B), \quad \text{falls } P(A) > 0 \end{aligned}$$

Rechenregeln

$P(A) \leq 1 \quad \forall A \in \Omega$

$P(\emptyset) = 0$

Aus $A \subset B$ folgt $P(A) \leq P(B)$

$P(\bar{A}) = 1 - P(A)$ → Gegenereignis

(Allgemeiner Additionssatz):

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Die Matrix in der Stochastik

- Datenmatrix einer Messreihe mit N Messungen von p Parametern

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \cdots & \ddots & \cdots \\ x_{p1} & \cdots & x_{pN} \end{pmatrix}$$

- Kovarianzmatrix mit Zusammenhängen der Daten
  {1 … N} x {1 … p}

Einfache Verteilungen

- Diskrete Gleichvertteilung (Laplace) $P(A) := \frac{|A|}{|\Omega|}, \quad A \subset \Omega,$
  → Würfeln, Günstige / Mögliche Ergebnisse

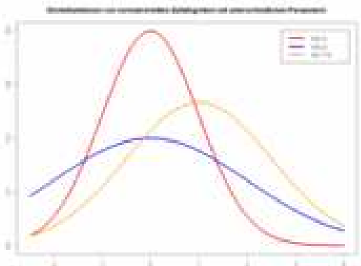
- Geometrische Verteilung $P(\{k\}) := p \cdot (1-p)^{k-1}, \quad k \in \mathbb{N}$
  → ∞ - facher Münzwurf, erster Kopf

- Binomial-Verteilung $P(\{k\}) := \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \Omega = \{0, ..., n\}$
  → n Münzwürfe, k Anzahl für „Kopf"

Multivariante Gauß Verteilung

Gauß'sche Normalverteilung als Dichtefuntion mit σ als Std.-Abweichung und μ als Erwartungswert:

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



„Gauß'sche Glocke"

N(·;*μ*;∑) bezeichnet die Multivariante Gauß Verteilung mit Vektor *μ* als Erwartungswert und Kovarianzmatrix ∑. *n* hat hier die Dimension von *o*

$$\mathcal{N}(o;\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} e^{-\frac{1}{2}(o-\mu)'\Sigma^{-1}(o-\mu)}$$

Maximum-Likelihood-Methode

- Induktive Stochastik
- Parameterschätzung
- Gegeben: Daten (meist Stichprobe, (x1…xn)), Annahmen über die Verteilung der Daten (*f(x)*)
- Likelihood-Funktion: $$L(q) = \prod_{i=1}^{n} f_{X_i}(x_i; q)$$
- Prüfung bei welchem Parameter *q* die Daten am wahrscheinlichsten sind (max-Funktion) ergibt die Maximum-Likelihood-Schätzung.
- Mathematisch: Erste Ableitung nach *q* bilden und gleich Null setzen

$$l(q) = \sum_{k=1}^{n} ln(f_{X_i}(x_i; q))$$

Formelsammlung, Einführung und Quellen

- http://www.statistik.lmu.de/institut/ag/biostat/teaching/stochbio2004/formelsammlung2003.pdf
- http://barolo.ipc.uni-tuebingen.de/pharma/index.html
- Prof. Dr. Rost, Daniel, 2001, Analysis II (für Informatiker), München
- http://www.uni-konstanz.de/FuF/wiwi/heiler/os/index.html
- http://de.wikipedia.org/wiki/Maximum_Likelihood
- http://de.wikipedia.org/wiki/Normalverteilung