Zusammenfassung
Wahrscheinlichkeit und Statistik

Autor:
André Gasser, gassera@student.ethz.ch

Datum:
Januar 2013

1 Kombinatorik

1.1 Permutation

n unterschiedlichen Kugeln:

$$P(n) = n!$$

n unterschiedliche Kugeln mit $n_1, n_2, ..., n_k$ gleichen Kugeln:

$$P(n, n_1, n_2, ..., n_k) = \frac{n!}{n_1!n_2!...n_k!}$$

Bemerkungen:

- Ermitteln der Anzahl Anordnungen.

1.2 Kombination

Mit Zurücklegen:

$$C_w(n,k) = \left(\begin{array}{c} n+k-1 \\ k \end{array} \right) = \frac{(n+k-1)!}{k!(n-1)!}$$

Ohne Zurücklegen:

$$C(n,k) = \left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!}$$

Bemerkungen:

- Reihenfolge spielt keine Rolle.

1.3 Variation

Mit Zurücklegen:

$$V_w(n,k) = n^k$$

Ohne Zurücklegen:

$$V(n,k) = \frac{n!}{(n-k)!}$$

Bemerkungen:

- Reihenfolge ist wesentlich.

2 Wahrscheinlichkeit

2.1 Allgemeine Rechenregeln

$$\begin{array}{rcl} P\left[A^C\right] & = & 1 - P[A] \\ P\left[\Omega\right] & = & P[A] + P\left[A^C\right] = 1 \\ P[\emptyset] & = & 0 \\ P\left[B^C|A\right] & = & 1 - P[B|A] \\ A \subseteq B & \Rightarrow & P[A] \leq P[B] \end{array}$$

2.2 DeMorgan'sche Gesetze

$$\begin{array}{rcl} \overline{A \cap B} & = & \overline{A} \cup \overline{B} \\ \overline{A \cup B} & = & \overline{A} \cap \overline{B} \end{array}$$

2.3 Additionssatz

Zur Berechnung der Wahrscheinlichkeit, dass Ereignis A oder B eintritt.

Allgemein:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

A,B disjunkt:

$$P[A \cup B] = P[A] + P[B]$$

Bemerkungen:

- Wenn die Ereignisse nicht offensichtlich disjunkt sind, die **erste Formel verwenden**!

2.4 Multiplikationssatz

Zur Berechnung der Wahrscheinlichkeit, dass die Ereignisse A und B eintreten.

$$P[A \cap B] = P[A] \cdot P[B|A]$$

Bemerkungen

- $P[A, B] = P[A \cap B]$.

2.5 Bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeit von B gegeben A entspricht der Wahrscheinlichkeit, dass B eintritt, wenn man schon weiss, dass A eingetreten ist. Es gilt:

$$P[B|A] = \frac{P[A \cap B]}{P[A]}$$

$$P[A \cap B] = P[A] \cdot P[B|A] = P[B] \cdot P[A|B]$$

Bemerkungen:

- Die **Pfadregel**, nach der Wahrscheinlichkeiten in einem **Wahrscheinlichkeitsbaum** multipliziert werden, um die Wahrscheinlichkeit eines Blattes zu erhalten, entspricht einer Verkettung bedingter Wahrscheinlichkeiten.

2.5.1 Satz der totalen Wahrscheinlichkeit

$$P[B] = \sum_{i=1}^{n} P[A_i] \cdot P[B|A_i]$$

2.5.2 Satz von Bayes

Zur Berechnung einer bestimmten Zwischenstation A_k in einem Ereignisbaum, wobei mehrere Ereignisse A_i zu Ereignis B führen.

$$P[A_k|B] = \frac{P[A_k] \cdot P[B|A_k]}{\sum\limits_{i=1}^{n} P[A_i] \cdot P[B|A_i]}$$

2.6 Eigenschaften von Ereignissen

2.6.1 Unabhängigkeit

Wenn zwischen zwei Ereignissen A und B kein kausaler Zusammenhang besteht (d.h. es gibt keine gemeinsamen Ursachen oder Ausschliessungen), dann sind sie *unabhängig* voneinander. In diesem Fall gilt:

$$P[A \cap B] = P[A] \cdot P[B]$$

3 Zufallsvariablen

3.1 Diskrete Zufallsvariablen

3.1.1 Gewichtsfunktion

Die Summe aller Gewichte ist immer 1 und die Werte immer im Intervall [0,1]:

$$\sum_{i=1}^{\infty} f(x) = 1$$

3.1.2 Verteilungsfunktion

3.2 Stetige Zufallsvariablen

3.2.1 Dichtefunktion

Die Fläche unter der Dichtefunktion ist immer 1 und die Werte immer im Intervall [0,1]:

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

Bemerkungen

- Durch Integrieren der Dichtefunktion erhält man die Verteilungsfunktion.

3.2.2 Verteilungsfunktion

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(u)du$$

Bemerkungen

- Durch Ableiten der Verteilungsfunktion erhält man die Dichtefunktion.

3.3 Gemeinsame Verteilungen

Gemeinsame Verteilungsfunktion: Die gemeinsame Verteilungsfunktion von n Zufallsvariablen $X_1, ..., X_n$ ist die Abbildung $F : \mathbb{R}^n \to [0, 1]$,

$$F(t_1, ..., t_n) := P[X_1 \leq t_1, ..., X_n \leq t_n]$$

Gemeinsame Gewichtsfunktion: Falls $X_1, ..., X_n$ *diskrete* Zufallsvariablen sind, ist ihre gemeinsame Gewichtsfunktion $p : \mathbb{R}^n \to [0, 1]$ definiert durch

$$p(x_1, ..., x_n) := P[X_1 = x_1, ..., X_n = x_n]$$

Gemeinsame Dichte: Seien $X_1, ..., X_n$ *stetige* Zufallsvariablen mit gemeinsamer Verteilungsfunktion $F(t_1, ..., t_n)$. Die Funktion $f : \mathbb{R}^n \to [0, \infty)$ heisst *gemeinsame Dichte* von $X_1, ..., X_n$, falls für alle $t_i \in \mathbb{R}$ gilt

$$F(t_1, ..., t_n) = \int_{-\infty}^{t_1} ... \int_{-\infty}^{t_n} f(x_1, ..., x_n)\ dx_n...dx_1$$

Randverteilung im diskreten Fall:

$$P[X = x] = \sum_{i=1}^{n} P[X = x, Y = y_i] = \sum_{i=1}^{n} p(x, y_i)$$

3.4 Funktionen diskreter Zufallsvariablen

3.4.1 Summe von Zufallsvariablen

...

3.4.2 Produkte von Zufallsvariablen

...

3.5 Funktionen stetiger Zufallsvariablen

$$\begin{aligned}
Y = aX + b &= F_Y(t) = P[Y \leq t] = P[aX + b \leq t] \\
&= P[X \leq \tfrac{t-b}{a}] = F_X(\tfrac{t-b}{a}) \\
Y = X^2 &= F_Y(t) = P[Y \leq t] = P[X^2 \leq t] \\
&= P[-\sqrt{t} \leq t \leq \sqrt{t}] = F_X(\sqrt{t}) - F_X(-\sqrt{t}) \\
Y = \tfrac{1}{X} &= F_Y(t) = P[Y \leq t] = P[\tfrac{1}{X} \leq t] \\
&= P[X \geq \tfrac{1}{t}] = 1 - P[X \leq \tfrac{1}{t}] = 1 - F_X(\tfrac{1}{t})
\end{aligned}$$

3.6 Chebyshev-Ungleichung

Die Chebyshev-Ungleichung liefert eine Abschätzung von Wahrscheinlichkeiten, auch wenn die genaue Verteilungsfunktion nicht bekannt ist. Es muss nur der Erwartungswert $E[X]$ und die Varianz $Var[X] < \infty$ einer Zufallsvariablen X bekannt sein, dann gilt für jedes $k > 0$:

$$P\left[|X - E[X]| \geq k\right] \leq \frac{Var[X]}{k^2}$$

3.7 Eigenschaften von Zufallsvariablen

3.7.1 Unabhängigkeit

Zwei Zufallsvariablen X und Y heissen unabhängig, wenn stets gilt

$$\begin{aligned}
F(x, y) &= F_X(x) \cdot F_Y(y) \\
f(x, y) &= f_X(x) \cdot F_Y(y)
\end{aligned}$$

wobei $F_X(x), F_X(y)$ die Verteilungsfunktionen und $f_X(x), f_Y(y)$ die Gewichts- bzw. Dichtefunktionen von X und Y sind.

3.7.2 Unkorreliert

Zwei Zufallsvariablen X und Y heissen *unkorreliert*, falls gilt $Cov(X, Y) = 0$.
Eine Menge von Zufallsvariablen $X_1, X_2, ..., X_n$ heissen *paarweise unkorreliert*, wenn alle Paare X_i, X_j mit $i \neq j$ unkorreliert sind.

3.8 Erwartungswert

Der Erwartungswert ist das langfristige Durchschnittsergebnis bei einem Zufallsexperiment mit vielen Wiederholungen.

Diskrete Verteilung:

$$E[X] = \sum_i x_i \cdot f(x_i)$$

Stetige Verteilung:

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

3.8.1 Additionssatz für Erwartungswerte

Der Erwartungswert einer aus n (diskreten oder stetigen) Zufallsvariablen $X_1, X_2, ..., X_n$ gebildeten Summe

$$Z = a_1X_1 + a_2X_2 + ... + a_nX_n$$

ist gleich der Summe der Erwartungswerte der einzelnen Zufallsvariablen:

$$E[Z] = a_1E[X_1] + a_2E[X_2] + ... + a_nE[X_n]$$

3.8.2 Multiplikationssatz für Erwartungswerte

Der Erwartungswert eines aus n *stochastisch unabhängigen* (diskreten oder stetigen) Zufallsvariablen $X_1, X_2, ..., X_n$ gebildeten Produkts

$$Z = X_1 \cdot X_2 \cdot ... \cdot X_n$$

ist gleich dem Produkt der Erwartungswerte der einzelnen Zufallsvariablen:

$$E[Z] = E[X_1 \cdot X_2 \cdot ... \cdot X_n] = E[X_1] \cdot E[X_2] \cdot ... \cdot E[X_n]$$

3.8.3 Formeln für 2. Moment

Dieses wird z.B. für die Momenten-Methode und zur Berechnung der Varianz benötigt.

Diskrete Verteilung:

$$E[X^2] = \sum_i x_i^2 \cdot f(x_i)$$

Stetige Verteilung:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx$$

3.8.4 Weitere Rechenregeln

$$\begin{aligned}
E[a] &= a(\text{für a = const.}) \\
E[aX] &= a \cdot E[X](\text{für a = const.}) \\
E[aX+b] &= a \cdot E[X] + b \\
E[E[X]] &= E[X]
\end{aligned}$$

3.9 Varianz

Die Varianz ist ein Streuungsmass, also ein Mass für die Abweichung einer Zufallsvariable von ihrem Erwartungswert.

Diskrete Verteilung:

$$Var[X] = \sum_i (x_i - E[X])^2 \cdot f(x_i)$$

Stetige Verteilung:

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

Vereinfachte Berechnung:

$$Var[X] = E[X^2] - E[X]^2$$

3.9.1 Additionssatz für Varianzen

Die Varianz einer aus n *stochastisch unabhängigen* (diskreten oder stetigen) Zufallsvariablen $X_1, X_2, ..., X_n$ gebildeten Summe

$$Z = a_1X_1 + a_2X_2 + ... + a_nX_n$$

ist gleich der Summe der Varianzen der einzelnen Zufallsvariablen:

$$Var[Z] = a_1^2Var[X_1] + a_2^2Var[X_2] + ... + a_n^2Var[X_n]$$

3.9.2 Weitere Rechenregeln

$$\begin{aligned}
Var[aX+b] &= a^2 \cdot Var[X] \\
Var[aX+bY] &= a^2 \cdot Var[X] + b^2 \cdot Var[Y] + \\
& 2ab \cdot Cov[X,Y]
\end{aligned}$$

3.10 Kovarianz

Die Kovarianz ist ein Mass für den Zusammenhang zwischen zwei Zufallsvariablen X und Y bzw. der Streuung zwischen ihnen.

$$\begin{aligned}
Cov[X,Y] &= E[(X - E[X])(Y - E[Y])] \\
Cov[X,Y] &= E[XY] - E[X] \cdot E[Y] \\
Cov[X,Y] &= \tfrac{1}{2} \cdot (Var[X+Y] - Var[X] - Var[Y]) \\
Cov[X,Y] &= Cov[Y,X] \\
Cov[X,X] &= Var[X] \\
Cov[aX+b,Y] &= a \cdot Cov[X,Y] \\
Cov[X+Y,Z] &= Cov[X,Z] + Cov[Y,Z] \\
\forall a \in \mathbb{R} : Cov[X,a] &= 0 \\
\forall b \in \mathbb{R} : Cov[X,bY] &= b \cdot Cov[X,Y]
\end{aligned}$$

3.11 Standardabweichung

$$sd[X] = \sqrt{Var[X]}$$

4 Diskrete Verteilungen

4.1 Bernoulli-Verteilung

Verteilung eines Experiments mit zwei Ausgängen (Erfolg, Misserfolg) mit Erfolgsparameter p.

Notation:

$$X \sim Be(p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = p^x \cdot (1-p)^{1-x}$$

Erwartungswert:

$$E[X] = p$$

Varianz:

$$Var[X] = p(1-p)$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{np(1-p)}$$

4.2 Binomial-Verteilung

Beschreibt die Anzahl der Erfolge in einer Serie von *gleichartigen* und *unabhängigen* Versuchen, die jeweils genau zwei mögliche Ergebnisse haben ("Erfolg" oder "Misserfolg"). n ist die Anzahl der Versuche bzw. Wiederholungen, p ist die Wahrscheinlichkeit für einen Ërfolg".

Notation:

$$X \sim Bin(n, p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$$

Erwartungswert:

$$E[X] = np$$

Varianz:

$$Var[X] = np(1-p)$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{np(1-p)}$$

Bemerkungen:

- Die Binomialverteilung $Bin(n, p)$ darf für grosse n und kleine p näherungsweise durch die Poisson-Verteilung mit Parameter $\lambda = np$ ersetzt werden (Faustregel: $np < 10$ und $n > 1500p$).

4.3 Geometrische Verteilung

Wartezeit auf ersten Erfolg bei einer Folge von 0-1-Experimenten mit Erfolgsparameter p. X ist die Nummer des ersten erfolgreichen Experiments.

Notation:

$$X \sim Geom(p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = p(1-p)^{x-1}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = 1 - (1-p)^x$$

Erwartungswert:

$$E[X] = \frac{1}{p}$$

Varianz:

$$Var[X] = \frac{1-p}{p^2}$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{\frac{1-p}{p^2}}$$

4.4 Poisson-Verteilung

Ereignet sich in einem Intervall (z.B. in einer gewissen Zeit, auf einer gewissen Fläche, in einem gewissen Volumen, usw.) ein völlig zufällig auftretendes Ereignis im Schnitt μ mal (Erwartungswert), dann ist die Zufallsgrösse, welche die Häufigkeit des Ereignisses in diesem Intervall angibt, poissonverteilt mit Parameter $\lambda = \mu$.

Notation:

$$X \sim P(\lambda), X \sim Pois(\lambda)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = e^{-\lambda} \frac{\lambda^x}{x!}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = e^{-\lambda} \sum_{k=0}^{x} \frac{\lambda^k}{k!}$$

Erwartungswert:

$$E[X] = \lambda$$

Varianz:

$$Var[X] = \lambda$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{\lambda}$$

Bemerkungen

- Falls $Z = X + Y$ und $X \sim Pois(\lambda)$, $Y \sim Pois(\mu)$ dann gilt $Z \sim Pois(\lambda + \mu)$.

4.5 Negativbinomiale Verteilung

Wartezeit auf den r-ten Erfolg bei einer Folge von unabhängigen 0-1-Experimenten mit Erfolgsparameter p.
Notation:

$$X \sim NB(r,p)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \sum_{i=r}^{x} \binom{i-1}{r-1} p^r (1-p)^{i-r}$$

Erwartungswert:

$$E[X] = \frac{r}{p}$$

Varianz:

$$Var[X] = \frac{r(1-p)}{p^2}$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{\frac{r(1-p)}{p^2}}$$

4.6 Hypergeometrische Verteilung

Beim Ziehen von m Gegenständen *ohne Zurücklegen* aus einer Urne mit n Gegenständen, wovon r vom Typ 1 und $n-r$ vom Typ 2 sind.

Notation:

$$X \sim HypGeom(n,m,r)$$

Gewichtsfunktion:

$$f(x) = P[X = x] = \frac{\binom{r}{x}\binom{n-r}{m-x}}{\binom{n}{m}}$$

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \sum_{k=max(0,m-n)}^{x} \frac{\binom{r}{k}\binom{n}{m-k}}{\binom{r+n}{m}}$$

Erwartungswert:

$$E[X] = \frac{mr}{n}$$

Varianz:

$$Var[X] = \frac{mr}{n^2(n-1)}(n-r)(n-m)$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{\frac{mr}{n^2(n-1)}(n-r)(n-m)}$$

5 Stetige Verteilungen

5.1 Gleichverteilung

Alle Ereignisse zwischen a und b sind gleich wahrscheinlich.

Notation:

$$X \sim U(a,b)$$

Dichtefunktion:

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

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \leq x \leq b \\ 1 & \text{für } x > b \end{cases}$$

Erwartungswert:

$$E[X] = \frac{b+a}{2}$$

Varianz:

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

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{\frac{(b-a)^2}{12}}$$

5.2 Exponentialverteilung

Die Exponentialverteilung ist ein Modell für Wartezeiten und Lebensdauern. Beispielanwendungen sind die Berechnung der Lebensdauer von Bauteilen und Zugverspätungen.

Notation:

$$X \sim Exp(\lambda)$$

Dichtefunktion:

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

Verteilungsfunktion:

$$F(x) = P[X \leq x] = \begin{cases} 1 - e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

Erwartungswert:

$$E[X] = \frac{1}{\lambda}$$

Varianz:

$$Var[X] = \frac{1}{\lambda^2}$$

Standardabweichung:

$$sd[X] = \sqrt{Var[X]} = \sqrt{\frac{1}{\lambda^2}}$$

Bemerkungen:

- Sind $X_1 \sim Exp(\lambda_1), ..., X_n \sim Exp(\lambda_n)$ *stochastisch unabhängig*, so ist $\min(X_1, ..., X_n) \sim Exp(\lambda_1 + ... + \lambda_n)$.
- Die Summe von n exponentialverteilten Zufallsvariablen mit *gleichem Parameter* λ ist gammaverteilt mit $\Gamma(n, \lambda)$.
- Erwartungswert und Standardabweichung sind gleich $(E[X] = sd[X])$.

5.3 Normalverteilung

Die besondere Bedeutung der Normalverteilung beruht unter anderem auf dem zentralen Grenzwertsatz, der besagt, dass eine Summe von n unabhängigen, identisch verteilten Zufallsvariablen mit endlicher Varianz im Grenzwert $n \to \infty$ normalverteilt ist.

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

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

Verteilungsfunktion:
$$F(x) = P[X \leq x] = \frac{1}{\sigma\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

Erwartungswert:
$$E[X] = \mu$$

Varianz:
$$Var[X] = \sigma^2$$

Standardabweichung:
$$sd[X] = \sqrt{Var[X]} = \sqrt{\sigma^2}$$

Transformation einer normalverteilten Zufallsvariable X in eine standardnormalverteilte Zufallsvariable Z:
$$Z = \frac{X - \mu}{\sigma}$$

5.4 Standardnormalverteilung

Notation:
$$X \sim N(0, 1)$$

Dichtefunktion:
$$\phi(u) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}}$$

Verteilungsfunktion:
$$\Phi(u) = P[X \leq u] = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{u} e^{-\frac{t^2}{2}} dt$$

Erwartungswert:
$$E[X] = 0$$

Varianz:
$$Var[X] = 1$$

Standardabweichung:
$$sd[X] = \sqrt{Var[X]} = \sqrt{1}$$

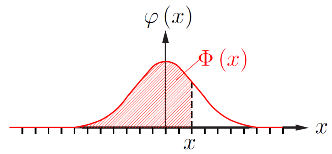
Bemerkungen:

- Die Standardnormalverteilung entspricht einer Normalverteilung mit den Parametern $\mu = 0$ und $\sigma = 1$.

5.4.1 Tabelle ablesen

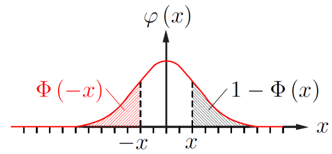
$\Phi(x)$ bedeutet, dass man zu einem Wert x die Wahrscheinlichkeit in der Tabelle auslesen will. $\Phi^{-1}(x)$ bedeutet, dass für eine Wahrscheinlichkeit der Tabellenwert ermittelt werden soll.

Positive Werte: $\Phi(x)$



Bsp: $\Phi(1.25) = 0.8944$

Negative Werte: $\Phi(-x) = 1 - \Phi(x)$



Bsp: $\Phi(-1.25) = 1 - \Phi(1.25) = 1 - 0.8944$

Für nicht tabellierte Werte:
$$\Phi^{-1}(0.03) = -\Phi^{-1}(0.97) = -1.88$$

Regeln zum Auslesen der Werte:

- Ist ein gesuchter Wert c nicht tabelliert, so wird derjenige Wert a oder b genommen, der näher bei c liegt.
- Sind zwei Werte a und b gleich weit vom gesuchten Wert c entfernt, so wird der Mittelwert von a und b verwendet.

6 Grenzwertsätze

In vielen Situationen taucht die Summe von *vielen gleichartigen Zufallsvariablen* auf. Wir möchten wissen, wie sich diese Summe etwa verhält, und untersuchen deshalb ihre Asymptotik, wenn die Anzahl der Summanden gegen unendlich geht.
Für die folgenden Sätze definieren wir einige Grössen:
$S_n = \sum_{i=1}^{n} X_i$ und $\overline{X}_n = \frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$

6.1 Schwaches Gesetz der grossen Zahlen

Seien $X_1, X_2, ..., X_n$ *unabhängige* oder *paarweise unkorrelierte* Zufallsvariablen mit gleichem Erwartungswert $E[X_i] = \mu$ und gleicher Varianz $Var[X_i] = \sigma^2$. Sei $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Dann gilt:

$$\forall \epsilon > 0 : \lim_{n\to\infty} P\left[|\overline{X}_n - \mu| > \epsilon\right] = 0$$

Bemerkungen:

- Für hinreichend grosse n konvergiert der Mittelwert $\overline{X}_n$ gegen den Erwartungswert μ.
- Der Satz funktioniert nicht, falls der Erwartungswert oder die Varianz nicht definiert sind.

6.2 Starkes Gesetz der grossen Zahlen

Seien $X_1, ..., X_n$ *unabhängige* Zufallsvariablen mit *gleicher Verteilung* (i.i.d) und Erwartungswert $E[X_i] = \mu$. Sei $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Dann gilt:

$$P\left[\lim_{n\to\infty} \overline{X}_n = \mu\right] = 1$$

Bemerkungen:

- Für hinreichend grosse n konvergiert der Mittelwert $\overline{X}_n$ gegen den Erwartungswert μ.

6.3 Zentraler Grenzwertsatz

$X_1, X_2, X_3, ..., X_n, ...$ seien *stochastisch unabhängige* Zufallsvariablen, die alle der *gleichen* Verteilungsfunktion mit Erwartungswert μ und Varianz σ^2 genügen. Dann konvergiert die Verteilungsfunktion $F_Z(u)$ der standardisierten Zufallsvariablen

$$Z_n = \frac{(X_1 + X_2 + ... + X_n) - n\mu}{\sqrt{n}\sigma}$$

im Grenzfall $n \to \infty$ gegen die Verteilungsfunktion $\Phi(t)$ der Standardnormalverteilung:

$$\lim_{n\to\infty} F_Z(u) = \Phi(u) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^u e^{-\frac{t^2}{2}} dt$$

Bemerkungen:

- Für ein hinreichend grosses n ist $Z_n = X_1 + X_2 + ... + X_n$ annähernd *normalverteilt* mit Erwartungswert $E[Z_n] = n\mu$ und Varianz $Var[Z_n] = n\sigma^2$.

7 Statistik

7.1 Schätzer

Schätzer sind Funktionen von Zufallsvariablen und somit selbst wieder Zufallsvariablen. Sie verfügen deshalb über einen Erwartungswert und eine Varianz.

7.1.1 Erwartungstreuer Schätzer

Ein Schätzer T ist *erwartungstreu* wenn der Erwartungswert des Schätzers gleich dem zu schätzenden Parameter ϑ ist:

$$E[T] = \vartheta$$

7.1.2 Konsistenter Schätzer

Ein Schätzer T ist *konsistent* wenn er mit zunehmendem Stichprobenumfang n gegen den gesuchten Parameter ϑ konvergiert:

$$\lim_{n\to\infty} P_\vartheta\left[|T^{(n)} - \vartheta| > \epsilon\right] = 0 \text{ (für jedes } \epsilon > 0)$$

7.1.3 Momenten-Methode

Sei $X_1, X_2, ..., X_n$ eine Stichprobe vom Umfang n mit gegebener Verteilung t. Die Parameter von t seien unbekannt. Mit der Momenten-Methode können diese geschätzt werden. Der Momentenschätzer ist i.d.R. nicht erwartungstreu.

Vorgehen:

1. 1. Moment berechnen: $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
2. 2. Moment berechnen: $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$
3. Nimm die Formel für den Erwartungswert der geg. Verteilung t und setze sie gleich $\overline{x}$
4. Nimm die Formel für die Varianz der geg. Verteilung t und setze sie gleich s^2
5. Löse das Gleichungssystem auf. Du erhälst die gesuchten Parameter.

7.1.4 Maximum-Likelihood-Methode

Methode zur systematischen Gewinnung von Schätzfunktionen.

Vorgehen:

1. Likelihood-Funktion $L(x_1, ..., x_n, \vartheta) = \prod_{i=1}^{n} f(x_i, \vartheta)$ aufstellen
2. Log-Likelihood-Funktion $\log L(x_1, ..., x_n, \vartheta)$ aufstellen (logarithmieren von L)
3. $\log L$ mit Hilfe elementarer Logarithmenregeln möglichst vereinfachen
4. $\log L$ nach jedem unbekannten Parameter partiell ableiten
5. Partielle Ableitungen $= 0$ setzen
6. Gleichungssystem nach Parametern auflösen

7.2 Tests

7.2.1 Wichtige Begriffe

- **Nullhypothese** H_0**:** Die zu prüfende Annahme.
- **Alternativhypothese** H_A: Alternative Annahme, falls H_0 verworfen werden muss.
- **Signifikanzniveau** α: Ist die Wahrscheinlichkeit, dass die Nullhypothese verworfen werden muss (auch Wahrscheinlichkeit für Fehler 1. Art).
- **Teststatistik T:** Test- oder Prüfwert.
- **P-Wert:** Das kleinste Niveau, auf dem der Test die Nullhypothese noch verwirft.
- **Fehler 1. Art:** H_0 wird verworfen, obwohl sie richtig wäre.
- **Fehler 2. Art:** H_0 wird beibehalten, obwohl H_A stimmt.

7.2.2 Allgemeines Vorgehen

1. Wahl des Modells
2. Formulieren von H_0 und H_A.
3. Signifikanzniveau α festlegen.
4. Teststatistik T T festlegen.
5. Bestimmen der kritischen Grenzen.
6. Konkreter Wert für Teststatistik T berechnen.
7. Testentscheidung: H_0 beibehalten oder verwerfen?

7.2.3 z-Test

Der z-Test ist ein Test für den Erwartungswert bei *bekannter* Varianz σ^2. Es seien also $X_1, X_2, ..., X_n \sim N(\vartheta, \sigma^2)$ *i.i.d.*. Wir wollen die Nullhypothese $\vartheta = \vartheta_0$ testen.

Mittelwert $\overline{X}$:	$\overline{X} = \frac{X_1+X_2+...+X_n}{n}$
Teststatistik T:	$T = \frac{\overline{X}-\vartheta_0}{\sigma/\sqrt{n}} \sim N(0,1)$

Kritischer Bereich:
Den kritischen Bereich bestimmt man abhängig von der Alternativhypothese H_A:

- $\vartheta_A > \vartheta_0$ einseitig: $K = (c_>, \infty)$ mit
$$c_> = (1-\alpha)\text{-Quantil} = \Phi^{-1}(1-\alpha)$$
- $\vartheta_A < \vartheta_0$ einseitig: $K = (-\infty, c_<)$ mit
$$c_< = \alpha\text{-Quantil} = \Phi^{-1}(\alpha)$$
- $\vartheta_A \neq \vartheta_0$ zweiseitig: $K = (-\infty, c_{\neq}) \cup (c_{\neq}, \infty)$ mit
$$c_{\neq} = \left(1-\frac{\alpha}{2}\right)\text{-Quantil} = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$$

7.2.4 t-Test

Der t-Test ist ein Test für den Erwartungswert bei *unbekannter Varianz* σ^2. Es seien also $X_1, X_2, ..., X_n \sim N(\vartheta, \sigma^2)$*i.i.d.*. Wir wollen die Nullhypothese $\vartheta = \vartheta_0$ testen.

Mittelwert $\overline{X}$:	$\overline{X} = \frac{X_1+X_2+...+X_n}{n}$
Schätzfunktion für S^2:	$S^2 = \frac{1}{n-1} \cdot \sum_{i=0}^{n} (X_i - \overline{X})^2$
Teststatistik T:	$T = \frac{\overline{X}-\vartheta_0}{S/\sqrt{n}} \sim t_{n-1}$

Die Teststatistik T ist **t-verteilt** mit $n-1$ **Freiheitsgraden**. Der kritische Bereich lässt sich analog zum z-Test mit den Quantilen der t_{n-1}-Verteilung bestimmen.

7.2.5 Ungepaarter Zweistichproben-z-Test

Seien $X_1, X_2, ..., X_n \sim N(\mu_X, \sigma^2)$ und $Y_1, Y_2, ..., Y_n \sim N(\mu_Y, \sigma^2)$ zwei Stichproben mit $m = n$ oder $m \neq n$. Die Erwartungswerte μ_X und μ_Y seien unbekannt, die Varianz σ^2 sei *bekannt*. Der Test kann wie ein normaler z-Test durchgeführt werden, als Teststatistik T wird jedoch folgende Formel verwendet:

$$T = \frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0,1)$$

7.2.6 Ungepaarter Zweichstichproben-t-Test

Seien $X_1, X_2, ..., X_n \sim N(\mu_X, \sigma^2)$ und $Y_1, Y_2, ..., Y_n \sim N(\mu_Y, \sigma^2)$ zwei Stichproben mit $m = n$ oder $m \neq n$. Die Erwartungswerte μ_X und μ_Y seien unbekannt, die Varianz σ^2 sei ebenfalls *unbekannt*. Der Test kann wie ein normaler t-Test durchgeführt werden, als Teststatistik T wird jedoch folgende Formel verwendet:

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2$$

$$S_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(Y_j - \overline{Y}_m\right)^2$$

$$S^2 = \frac{1}{m+n-2}\left((n-1)S_X^2 + (m-1)S_Y^2\right)$$

$$T = \frac{\left(\overline{X}_n - \overline{Y}_m\right) - (\mu_X - \mu_Y)}{S\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t_{n+m-2}$$

7.3 Konfidenzbereiche

Ein Konfidenzbereich gibt ein Intervall an, in dem sich ein gesuchter Parameter mit sehr hoher Wahrscheinlichkeit befindet.

8 Differentialrechnung

8.1 Faktorregel

Ein konstanter Faktor bleibt beim Differenzieren erhalten:

$$y = C \cdot f(x) \Rightarrow y' = C \cdot f'(x)$$

8.2 Summenregel

Bei einer endlichen Summe von Funktionen darf gliedweise differenziert werden:

$$y = f_1(x)+f_2(x)+...+f_n(x) \Rightarrow y' = f_1'(x)+f_2'(x)+...+f_n'(x)$$

8.3 Produktregel

Die Ableitung einer in Produktform $y = u(x) \cdot v(x)$ darstellbaren Funktion erhält man nach folgender Produktregel:

$$y' = u'(x) \cdot v(x) + u(x) \cdot v'(x) = u'v + uv'$$

8.4 Quotientenregel

Die Ableitung einer Funktion, die als Quotient zweier Funktionen $u(x)$ und $v(x)$in der Form $y = \frac{u(x)}{v(x)}$ darstellbar ist, erhält man nach der Quotientenregel:

$$y' = \frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{v^2(x)}$$

8.5 Kettenregel

Die Ableitung einer zusammengesetzten (verketteten) Funktion $y = F(u(x)) = f(x)$ erhält man als Produkt aus äusserer und innerer Ableitung:

$$y' = \frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

8.6 Partielle Differentiation

Summanden, die keine Variable beinhalten nach der abgeleitet wird, fallen WEG!

8.7 Wichtige elementare Ableitungen

$$\begin{array}{lll}
f(x) = c & \rightarrow & f'(x) = 0 \\
f(x) = x^n & \rightarrow & f'(x) = n \cdot x^{n-1} \\
f(x) = \sqrt{x} & \rightarrow & f'(x) = \frac{1}{2\sqrt{x}} \\
f(x) = e^x & \rightarrow & f'(x) = e^x \\
f(x) = a^x & \rightarrow & f'(x) = (\ln a) \cdot a^x \\
f(x) = \ln x & \rightarrow & f'(x) = \frac{1}{x} \\
f(x) = \log_a x & \rightarrow & f'(x) = \frac{1}{(\ln a) \cdot x}
\end{array}$$

9 Integralrechnung

9.1 Faktorregel

Ein konstanter Faktor darf vor das Integral gezogen werden:

$$\int_a^b C \cdot f(x)dx = C \cdot \int_a^b f(x)dx$$

9.2 Summenregel

Eine endliche Summe von Funktionen darf giedweise integriert werden:

$$\int_a^b (f_1(x) + f_2(x) + ... + f_n(x))\, dx =$$

$$\int_a^b f_1(x)dx + \int_a^b f_1(x)dx + ... + \int_a^b f_1(x)dx$$

9.3 Vertauschungsregel

Vertauschen der beiden Integrationsgrenzen bewirk einen Vorzeichenwechsel des Integrals:

$$\int_a^b f(x)dx = -\int_b^a f(x)dx$$

9.4 Gleiche Integrationsgrenzen

Fallen die Integrationsgrenzen zusammen ($a = b$), so ist der Integralwert gleich Null:

$$\int_a^b f(x)dx = 0$$

9.5 Partielle Integration

$$\int f(x)dx = \int u(x) \cdot v'(x)dx = u(x) \cdot v(x) - \int u'(x) \cdot v(x)dx$$

Vorgehen:

1. Integrand in $u(x)$ und $v'(x)$ zerlegen.
2. $u(x)$ ableiten, $v'(x)$ integrieren.
3. Formel aufschreiben und lösen.

9.6 Wichtige Stammintegrale

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C \qquad \int a^x dx = \frac{a^x}{\ln a} + C$$

$$\int \frac{1}{x} dx = \ln|x| + C \qquad \int e^x dx = e^x + C$$

9.7 Weitere Integrale

$$
\begin{array}{rcll}
\int a\ dx & = & ax & \\
\int x^a\ dx & = & \frac{1}{a+1}x^{a+1}, & a \neq -1 \\
\int (ax+b)^c\ dx & = & \frac{1}{a(c+1)}(ax+b)^{c+1}, & c \neq -1 \\
\int \frac{1}{x}\ dx & = & \log|x|, & x \neq 0 \\
\int \frac{1}{ax+b}\ dx & = & \frac{1}{a}\log|ax+b| & \\
\int \frac{1}{x^2+a^2}\ dx & = & \frac{1}{a}\arctan\frac{x}{a} & \\
\int e^{ax}\ dx & = & \frac{1}{a}e^{ax} & \\
\int xe^{ax}\ dx & = & \frac{e^{ax}}{a^2}(ax-1) & \\
\int x^2e^{ax}\ dx & = & e^{ax}\left(\frac{x^2}{a}-\frac{2x}{a^2}+\frac{2}{a^3}\right) & \\
\int \log|x|\ dx & = & x(\log|x|-1) & \\
\int \log_a|x|\ dx & = & x(\log_a|x|-\log_a e) & \\
\int x^a\log x\ dx & = & \frac{x^{a+1}}{a+1}\left(\log x-\frac{1}{a+1}\right), & a\neq -1, x>0 \\
\int \frac{1}{x}\log x\ dx & = & \frac{1}{2}\log^2 x, & x>0
\end{array}
$$

10 Verschiedenes

Binomialkoeffizient
Der Binomialkoeffizient

$$\binom{n}{k}$$

gibt für $n, k \in \mathbb{N}$ an, **wie viele Möglichkeiten es gibt, k Objekte aus n Objekten auszuwählen**. Damit gibt der Binomialkoeffizient an, wie viele k-elementige Teilmengen aus einer n-elementigen Menge gebildet werden können (gesprochen: "k aus n" oder "k tief n").

Für $k = 0$ ist der Binomialkoeffizient 1:

$$\binom{n}{0} = 1$$

Für $k = 1$ ist der Binomialkoeffizient n:

$$\binom{n}{1} = n$$

Für $k > n$ ist der Binomialkoeffizient stets 0.

Gammafunktion

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt = (x-1)!$$

Diverse Summenformeln

$$\sum_{k=0}^{\infty}\frac{1}{k!} = e$$

$$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^\lambda$$

$$\sum_{k=1}^{\infty}\frac{\lambda^k}{k!} = e^\lambda - 1$$

$$\sum_{i=0}^{n} i = \sum_{i=1}^{n} i = \frac{n(n-1)}{2}$$

$$\sum_{i=1}^{n} i^2 = \frac{n\cdot(n+1)\cdot(2n+1)}{6}$$

$$\sum_{i=0}^{\infty} a\cdot p^i = \sum_{i=1}^{\infty} a\cdot p^{i-1} = \frac{a}{1-p}$$

Produkteformeln

$$\prod_{k=m}^{n} a\cdot x_i = a^{n-m+1}\cdot\prod_{k=m}^{n} x_i$$

Logarithmenregeln

$$
\begin{array}{rcl}
\log(uv) & = & \log(u)+\log(v) \\
\log\left(\frac{u}{v}\right) & = & \log(u)-\log(v) \\
\log_b(r) & = & \frac{\log_a(r)}{\log_a(b)} \\
\log_a(u^k) & = & k\cdot\log_a(u) \\
\log_a(\sqrt[n]{u}) & = & \log_a(u^{\frac{1}{n}}) = \frac{1}{n}\cdot\log_a(u) \\
\log_a(a^b) & = & b \\
\log 1 & = & 0 \\
\log_a(a) & = & 1 \\
\ln e & = & 1
\end{array}
$$

11 Beispiele

11.1 t-Test

Ein Waschmittelhersteller bringt 5kg-Packungen in den Umlauf. Die Konsumentenschutzorganisation kauft 25 Packungen. Es ergibt sich ein Mittel von $\overline{X} = 4.9kg$ und eine empirische Stichprobenvarianz $S^2 = 0.1kg^2$. Die einzelnen Gewichte seine durch unabhängige $N(\mu, \sigma^2)$ Zufallsvariablen beschrieben.

1. Wie lauten die Hypothesen H_0 und H_A?

2. Wie ist $(\overline{X} - 5)/(S/5))$ verteilt unter H_0?

3. Führen Sie den t-Test auf dem 5

4. Berechnen Sie das 95%-Vertrauensintervall für den in 3. beschriebenen Test.

Lösung:

1. $H_0 : \mu = 5, H_1 : \mu < 5$

2. Verteilung: $\sim t_{5^2-1} = t_{24}$

3. $T_1 = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} = -1.581$. Wir ermitteln das t-Quantil für $n - 1 = 24$ Freiheitsgrade und $t^{-1}(\alpha) = t^{-1}(0.05) = t^{-1}(0.95) = 1.711$. Somit erhalten wir eine kritische Grenze von $c = -1.711$. Weil $T_1 = -1.581 > c$ ist, wird die Nullhypothese nicht verworfen.

4. ???

11.2 Erwartungstreuer Schätzer

Wir prüfen, ob der Schätzer $T = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ erwartungstreu ist.

$$E[T] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=0}^{n} E[X_i] = \frac{1}{n}\cdot n\cdot\mu = \mu$$

Der Schätzer T ist erwartungstreu.

11.3 Verteilung der Summe zweier normalerverteilter ZV

Seien $X \sim N(\mu_1, \sigma_1^2)$ und $Y \sim N(\mu_2, \sigma_2^2)$. Wie ist $Z = 1 + aX + bY$ verteilt?

Lösung:

$$\begin{aligned} Z &\sim N(1 + aE[X] + bE[Y], a^2Var[X] + b^2Var[Y]) \\ &\sim N(1 + a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2) \end{aligned}$$