Wahrscheinlichkeitsrechnung und Statistik - Zusamenfassung

Patrick Pletscher

22. September 2003

1 Wahrscheinlichkeiten

1.1 Ereignisraum

Der *Ereignisraum* Ω umfasst alle möglichen Ausgänge eines Zufallsexpermiments.
Ein *Elementarereignis* Ω ist ein Element $\omega \in \Omega$
Ein Ereignis A ist eine Teilmenge von Ω, d.h. eine Kombination von Elementarereignissen $A \subset \Omega$

A^c (Komplement) ist das Ereignis, dass A nicht eintritt.

$\mathcal{A}$ ist die Klasse der beobachteten Ereignisse. Falls Ω endlich ist, dann ist $\mathcal{A}$ die Menge aller Teilmengen von Ω, d.h. die Potenzmenge.

1.2 Das Wahrscheinlichkeitsmass

$P : \mathcal{A} \to [0, 1]$
P(A)$\doteq$ 'die Wahrscheinlichkeit, dass A eintritt'

Axiome der Wahrscheinlichkeitstheorie

A1 $0 \leq P(A) \leq 1$ für alle $A \subset \Omega$

A2 $P(\Omega) = 1$

A3 Sei $A_1, A_2, \ldots$ eine Folge disjunkter Ereignisse, dann $P(\cup_{i=1}^{\infty}) = \sum_{i=1}^{\infty} P(A_i)$

Weitere Rechenregeln

$P(A^c) = 1 - P(A)$

$P(\emptyset) = 0$

$A \subset B \Rightarrow P(A) \leq P(B)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cup B) = 1 - (P(\bar{A}) + P(\bar{B}))$

und $\rightarrow$ Multiplikation

1.3 Berechnung von W'keiten in endlichen Räumen

$P(A) = \frac{|A|}{|\Omega|}$

Exkurs in die Kombinatorik

Permutationen ohne Zurücklegen
Aus n Objekten sind $k \leq n$ herauszugreifen, wobei die Reihenfolge eine Rolle spielen soll.

$\sharp$ Mögl. $= \frac{n!}{(n-k)!}$

Permutationen mit Zurücklegen
Gegeben sind n Objekte. Wieviele Folgen der Länge k können gebildet werden, falls jedes Objekt beliebig oft gewählt werden darf.

$\sharp$ Mögl. $= n^k$

Kombinationen ohne Zurücklegen
Gegeben eine Menge mit n Elementen. Wieviele Teilmengen mit $k \leq n$ Elementen kann man daraus bilden?

$\sharp$ Mögl. $= \frac{n!}{(n-k)!k!} = \binom{n}{k}$

Bsp: Lotto (6 aus 45)
$A = \{6 Richtige\} = \binom{45}{6}$
$B = \{4 Richtige\} = \binom{6}{4}\binom{39}{2}$

1.4 Bedingte W'keiten

Seien A,B Ereignisse, $P(A) > 0$
Def.: Die *bedingte W'keit*, dass A gegeben B eintritt, ist
$P(B|A) = \frac{P(A \cap B)}{P(A)}$

Multiplikationssatz
Sei $P(A) > 0$: Dann
$P(A \cap B) = P(B|A) \cdot P(A)$

Satz von der totalen W'keit
Gegeben seien Ereignisse $A_1, \ldots, A_n$ mit $P(A_i) > 0$.
Def.: Die A_i bilden eine Zerlegung von Ω, falls

1. $A_i \cap A_j = \emptyset$ für $i \neq j$, paarw. disj.

2. $\Omega = \bigcup_{i=1}^{n} A_i$

Dann gilt für $B \subset \Omega$ dass
$P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$

Satz von Bayes
$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^C) \cdot P(A^C)}$

Allgemeine Formel von Bayes
$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$

1.5 Unabhängigkeit

Seien $A, B \subset \Omega$ zwei Ereignisse.
Def. A und B sind *unabhängig*, falls $P(A \cap B) = P(A)P(B)$
Falls $P(A) > 0$:
A,B unabh. $\Leftrightarrow P(B|A) = P(B)$
Falls $P(B) > 0$:
A,B unabh. $\Leftrightarrow P(A|B) = P(A)$

Allgemeine Definition der Unabhängigkeit

n Ereignisse $A_1, \ldots, A_n$ heissen *unabhängig*, falls für *jede* Wahl von je m Ereignissen
$A_{k1}, \ldots, A_{km}, \{k_1, \ldots, k_m\} \subset \{1, \ldots, n\}$,
stets gilt:
$P(A_{k1} \cap \ldots \cap A_{km}) = P(A_{k1}) \cdot \ldots \cdot P(A_{km})$

2 Zufallsvariable

2.1 Begriff der ZV

Es sei Ω ein Ereignisraum. Eine ZV auf Ω ist eine Funktion $\Omega \to \mathbb{R}$.
Wir nennen eine ZV diskret, falls sie endliche oder abzählbar viele Werte annimmt.

2.2 Wahrscheinlichkeit- und Verteilungsfunktion

Def. Die W'keitsfunktion der diskreten ZV X ist die Funktion
$p(x) = \begin{cases} P(X = x) & \text{falls } x \in \omega \\ 0 & \text{sonst} \end{cases}$
ω: Wertebereich von X

Verteilungsfunktion F: $F(x) = P(X \leq x)$
Für diskrete ZV ist
$F(x) = \sum p(x_i)$
$x_i : x_i \in \omega, x_i \leq x$

Die Verteilungsfunktion ist:

- rechtsstetig
- $\lim_{x \to -\infty} F(X) = 0$, $\lim_{x \to +\infty} F(X) = 1$

2.3 Einige wichtige diskrete Verteilungen

1. Uniforme Verteilung (Gleichverteilung)
   $\omega = \{x_1, \ldots, x_n\}$
   $p(x_i) = P(X = x_i) = \frac{1}{n}$ Bsp: Würfeln mit einem Würfel

2. Bernoulli-Verteilung
   X nimmt nur die Werte 0 und 1 an
   $P(X = 1) = p$
   $P(X = 0) = 1 - p$
   $P(X) = p^x \cdot (1-p)^{1-x}$ $, x = 0, 1$
   $X \sim Be(p)$

3. Binomialverteilung
   X="Anzahl Erfolge bei n Versuchen"
   $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$
   $X \sim B(n, p)$
   Approx. durch Poisson
   Multinomialverteilung
   $P[X_1 = x_1, X_2 = x_2, X_3 = x_3] = \frac{n!}{x_1! x_2! x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$

4. Die geometrische Verteilung
   X="Anzahl der Versuche, bis ein Erfolg eintritt"
   $p(k) = P(X = k) = (1-p)^{k-1} p$

5. Die negativbinomiale Verteilung
   X="Anzahl Versuche, bis ich r-mal erfolgreich bin"
   $P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$

6. Die hypergeometrische Verteilung
   n Gegenstände in einer Urne, r vom Typ I und n-r vom Typ II, ich ziehe m davon (ohne Zurücklegen) $P(X = k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}$
   Bsp: Lotto, 6 aus 45, W'keit eines Vierers:
   n=45, r=6, m=6, k=4

7. Die Poisson-Verteilung
   X="Anzahl Ereignisse in einem Zeitintervall(Anrufe, gedruckte Files)"
   X ist Poisson-verteilt mit Parameter λ, falls
   $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, \ldots$
   Wertebereich $W = 1, \ldots, \infty$
   $\lambda = E[X] = n \cdot p$

2.4 Stetige Zufallsvariablen

Sind ZV, die Werte in einem Intervall W annehmen können.
z.B.: $W = \mathbb{R}$, $W = \mathbb{R}^+$, $W = [0, 1]$
Definition
Sei X eine ZV mit *Verteilungsfunktion oder W'keitsfunktion* $F(x) = P[X \leq x]$.
Falls es eine Funktion f(x) gibt, so dass

$F(x) = \int_{-\infty}^{x} f(y)dy$ für alle $x \in W$, dann heisst $f(x)$ die *Dichte* von X.

Eigenschaften

- $f(x) \geq 0$ für alle x
- $f(x)$ ist stetig oder stückweise stetig
- $\int\limits_{-\infty}^{\infty} f(x)dx = 1$, weil $\lim\limits_{x\to\infty} F(x) = 1$
- $a < b:\ P[a < X \leq b] = P[X \leq b] - P[X \leq a] = F(b) - F(a) = \int\limits_{a}^{b} f(x)dx$
- $P[X = a] = 0$

Satz
An allen Stellen, an denen $f(x)$ stetig ist, gilt:
$F'(x) = f(x)$

Beispiele

1. Gleichverteilung (Uniform Distribution): $X \sim U(0,1)$
   $f(x) = \begin{cases} 1 \text{ für } x \in [0,1] \\ 0 \text{ sonst} \end{cases}$
   $F(x) = \begin{cases} 0 \text{ für } x < 0 \\ x \text{ für } x \in [0,1] \\ 1 \text{ für } x > 1 \end{cases}$
2. Exponentialverteilung $X \sim Exp(\lambda)$
   $\lambda > 0$
   $f(x) = \begin{cases} \lambda e^{-\lambda x} \text{ für } x > 0 \\ 0 \text{ sonst} \end{cases}$
   $F(x) = \begin{cases} \int\limits_0^x f(y)dy = \int\limits_0^x \lambda e^{-\lambda x}dy = 1 - e^{-\lambda x} \text{ für } x \geq 0 \\ 0 \text{ für } x < 0 \end{cases}$
   Die Exp-Verteilung ist gedächnislos.
   Ein Ankunftsprozess, bei dem die Zeiten zwischen den Ankünften Exponentiell-verteilt ist, heisst Poisson-Prozess. λ heisst dann die (Ankunfts)Rate des Poisson-Prozess.
3. Die Normalverteilung / Gauss-Verteilung $X \sim N(\mu, \sigma^2)$
   $$f(x) = \frac{1}{\sqrt{2\Pi}\sigma} exp\{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\}$$
   $\mu =$ "Mean", Mittelwert, $-\infty < \mu < \infty$
   $\sigma =$ Standardabweichung, $\sigma > 0$
   $\sigma^2 =$ Varianz
   $F(x) = \int\limits_{-\infty}^{x} f(y)dy$: keine geschlossene Form, aber Tabellen
   Standardisierte NV:
   $\mu = 0$, $\sigma = 1$
   Wenn $X \sim N(\mu, \sigma^2)$ dann ist $\frac{x-\mu}{\sigma} \sim N(0,1)$
   $P[X \leq x'] = P[\frac{x-\mu}{\sigma} \leq \frac{x'-\mu}{\sigma}] = \Phi(\frac{x'-\mu}{\sigma})$
   $\Phi(-x) = 1 - \Phi$

2.5 Transformation von ZV

Sei X eine ZV mit Verteilungsfkt. $F_X(x)$ und Dichte $f_X(x)$.
Gesucht: Verteilung und Dichte von $Y = g(X)$

Satz
Wenn $X \sim N(\mu, \sigma^2)$ und $Y = aX + b$, dann gilt $Y \sim N(a\mu + b, a^2\sigma^2)$

Satz
Sei X stetig, mit Dichte $f_X(x)$, Verteilung $F_X(x)$.
Sei $Y = g(X)$, mit g diff'bar, streng monoton steigend/fallend auf einem Intervall I, wobei $f_X(x) = 0$ für $x \notin I$ ($I = \mathbb{R}$ ist zugelassen).
Dann ist die Dichte von Y
$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \cdot |\frac{d}{dy} g^{-1}(y)| \\ 0 \text{ sonst} \end{cases}$
wobei g^{-1} die Umkehrfkt. von g und $y \in \{g(x)|x \in I\}$

und die Verteilungsfkt. von Y ist
$F_Y(y) = \begin{cases} F_X(g^{-1}(y)) \text{ für g steigend} \\ 1 - F_X(g^{-1}(y)) \text{ für g fallend} \end{cases}$
$y \in \{g(x)|x \in I\}$

Lognormale Verteilung
$X \sim N(\mu, \sigma^2)$, $Y = e^x$
$g(x) = e^x,\ g^{-1}(y) = \ln y,\ I = \mathbb{R}$

Satz
Sei $U \sim U(0,1)$.
Sei $F(x)$ eine stetige, streng monoton wachsende Verteilungsfunktion. Setze
$X := F^{-1}(U)$
Dann gilt:
$P[X \leq x] = F(x)$

Definition

- Der Wert $F^{-1}(p)$ (für $p \in (0,1)$) heisst das p-Quantil der Verteilung F
- $F^{-1}(0.5)$ heisst der Median der Verteilung F.

3 Gemeinsame Verteilung mehrerer ZV

Seien $X_1, \ldots, X_n$ Zufallsvariablen.

Dann ist $F(x_1, \ldots, x_n) = P[X_1 \leq x_1, \ldots, X_n \leq x_n]$

3.1 Stetige ZV

Falls wir $F(x_1, \ldots, x_n)$ folgendermassen darstellen können $F(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} f(y_1, \ldots, y_n) dy_n \ldots dy_1$ dann heisst $f(x_1 \ldots x_n)$ die Dichte von $(X_1 \ldots X_n)$

Eigenschaften

- $P[(X_1 \ldots X_n) \in A] = \iint\limits_A f(\vec{x}) d\vec{x} = \int \ldots \int\limits_{(x_1 \ldots x_n) \in A} f(x_1, \ldots, x_n) dx_1 \ldots dx_1$
- $\int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_n \ldots dx_1 = 1$
- $f(x_1, \ldots, x_n = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F(x_1 \ldots x_n)$

3.2 Randverteilungen

Gegeben sei die gemeinsame Verteilung von X und Y:$F(x, y)$. Die Randverteilung von X ist:
$F_X(x) = P[X \leq x] = P[X \leq x;\ Y \in (-\infty, \infty)] = \lim\limits_{y \to \infty} F(x, y)$

Diskrete ZV: $Y \in \{y_1, y_2, \ldots\}$
Die W'keitsfkt. der Randverteilung von X ist
$p_X(x) = \sum\limits_j p(x, y_j)$

Stetige ZV: Die Dichte der RV von X ist
$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} \int\limits_{-\infty}^{x} [\int\limits_{-\infty}^{\infty} f(x', y') dy'] dx' = \int\limits_{-\infty}^{\infty} f(x, y') dy'$

Beispiel
$f(x, y) = \begin{cases} 6(x-y) \text{ für } 0 \leq y \leq x \leq 1 \\ 0 \text{ sonst} \end{cases}$
$f_X(x) = \int\limits_0^x f(x, y) dy$
$f_Y(y) = \int\limits_y^1 f(x, y) dx$

3.3 Unabhängigkeit von ZV

Die ZV $X_1, \ldots, X_n$ sind unabhängig, falls $F(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \ldots \cdot F_{X_n}(x_n)$ für alle $(x_1 \ldots x_n) \in \mathbb{R}^n$

1. Diskreter Fall: $(X_1 \ldots X_n)$ unabh. $\Leftrightarrow p(x_1 \ldots x_n) = P_{X_1} \cdot \ldots \cdot P_{X_n}(x_n)$
2. Stetiger Fall: $f(x_1 \ldots x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n)$

X,Y unabhängig $(F(x, y) = F_X(x) \cdot F_Y(y)) \Leftrightarrow f(x, y) = f_X(x) \cdot f_Y(y)$

Wichtige Mehrdimensionale Verteilungen

1. die mehrdimensionale Normalverteilung (stetige Verteilung)
   Dichte:
   $f(x, y) = \frac{1}{2\Pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \cdot$
   $exp\{-\frac{1}{2(1-\rho^2)[(\frac{x-\mu_X}{\sigma_X})^2 + (\frac{y-\mu_Y}{\sigma_Y})^2 - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X \sigma_Y}]}\}$
   $-1 < \rho < +1$ Korrelation zw. X und Y.
   Randverteilung
   $f_X(x) = \frac{1}{\sigma_X \sqrt{2\Pi}} exp(-\frac{1}{2}(\frac{x-\mu_X}{\sigma_X})^2)$
   Unabhängigkeit von X und Y, genau dann, wenn $\rho = 0$
2. Die Multinomiale Verteilung (diskret)
   Es werden n unabhängige Experimente durchgeführt. Es gibt jeweils r mögliche Ergebnisse mit W'keit $p_1, \ldots, p_r$. $\sum_{i=1}^n p_i = 1$.
   Sei N_i die Anzahl der Ergebnisse "i".
   Die Vtlg. von $[N_1, \ldots, N_r]$ heisst Multinomialvtlg.
   $p(N_1 = n_1, \ldots, N_r = n_r) = \binom{n}{n_1 \cdots n_r} p_1^{n_1} \cdots p_r^{n_r}$
   $\binom{n}{n_1 \cdots n_r} = \frac{n!}{n_1! \cdots n_r!}$: Multinomialkoeffizient
   RV von N_i:
   $p_{N_i}(n_i) = P[N_i = n_i] = \binom{n}{n_i} p_i^{n_i} (1-p_i)^{n-n_i}$
   $p(n_1 \ldots n_r) \neq \prod\limits_{i=1}^{r} p_{N_i}(n_i) \Leftrightarrow$ keine Unabh.
3. Mehrdim. Verteilungen mit U(0,1)-RV
   $F(x, y) = exp(-[(-\ln x)^\beta + (-\ln y)^\beta]^{\frac{1}{\beta}}$
   $x, y \in]0, 1], \beta \geq 1$
   RV: $F_X(x) = F(x, y = 1) = x$: U(0,1)-Vrtlg. (Y genauso)

3.4 Bedingte Verteilungen

Diskrete ZV

X,Y ZV diskret mit gemeinsamer W'keitsfkt. p(x,y)
Def. Die bedingte W'keitsfkt. von X unter der Bedingung, dass Y=y ist, ist $p_{X|Y}(x|y) = P[X = x|Y = y] = P[X = x, Y = y]/P[Y = y]$

Bem:

- $\sum\limits_i p_{X|Y}(x_i, y) = \frac{\sum\limits_i p(x_i, y)}{p_Y(y)} = 1$
- Wenn $p_{X|Y}(x|y) = p_X(x)$ für alle x,y, dann sind X und Y unabh.

Stetige ZV

Seien X,Y stetige ZV mit gem. Dichte $f(x, y)$
Def. Die bedingte Dichte von Y, gegeben X=x, ist $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$, wenn $0 < f_X(x) < \infty$ sont $f_{Y|X}(y|x) = 0$

Bem:

- $\int\limits_{-\infty}^{\infty} f_{Y|X}(y|x)dy = 1$
- $f_{Y|X}(y|x) = f_Y(y)$ für alle x,y $\Leftrightarrow$ X und Y sind unabh.

3.5 Funktionen von ZV

X,Y ZV mit bekannter gem. Vtlg.

Was ist die Vtlg. der Summe X+Y?

1. Diskreter Fall
   $Z = X + Y$
   $p_Z(z) = \sum\limits_i p(x_i, z - x_i)$,
   Falls X,Y unabhängig:
   $p_Z(z) = \sum\limits_i p_X(x_i)p_Y(z - x_i)$ (Faltung von X,Y)
2. stetiger Fall
   X,Y Dichtefkt. f(x,y) sind gegeben. Z=X+Y
   $f_Z(z) = \int\limits_{-\infty}^{+\infty} f(x, z - x)dx$
   Falls X,Y unabhängig:
   $f_Z(z) = \int\limits_{-\infty}^{+\infty} f_X(x)f_Y(z - x)dx$

4 Erwartungswert

4.1 Definition und Eigenschaften

Diskrete ZV
X sei diskrete ZV mit W'keitsfunktion p(x). Dann ist der EW von X def. durch $E[X] = \sum\limits_i x_i p(x_i)$ falls die Reihe absolut konvergiert.

1. Bernoulli Verteilung ($X \sim Be(p)$)
   $E[X] = 1\cdot p + 0\cdot (1 - p) = p$
2. Geometrische Verteilung ($X \sim Nb(1, p)$)
   $E[X] = \frac{1}{p}$
3. Poisson Verteilung ($X \sim P_\lambda(x)$)
   $E[X] = \lambda \sum\limits_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}$

Stetige ZV
X sei stetig verteilt mit Dichtefkt. f_X. Dann ist $E[X] = \int\limits_{-\infty}^{+\infty} x f_X(x)dx$, falls das Integral absolut konvergiert.

1. Uniforme Verteilung ($X \sim U(0, 1)$)
   $E[X] = \int\limits_0^1 x dx = [\frac{1}{2}x^2]_0^1 = \frac{1}{2}$
2. Normale Verteilung ($X \sim N(\mu, \sigma^2)$)
   $E[X] = \mu$
3. Cauchy Verteilung
   $f(x) = \frac{1}{\Pi}\frac{1}{1+x^2}$, $-\infty < x < +\infty$
   $E[X] = +\infty$

4.2 Funktionen von ZV

Satz
Sei $Y = g(X)$. Falls X diskret mit W'keitsfkt. p(x) ist, dann ist $E[g(X)] = \sum\limits_i g(x_i)p_x(x_i)$, falls die Summe absolut konvergiert.
Falls X stetig verteilt verteilt ist mit Dichtefkt. f, dann ist $E[g(X)] = \int\limits_{-\infty}^{+\infty} g(x)f(x)dx$ falls das Integral konvergiert.

Satz
Seien $X_1, \ldots, X_n, Y$ ZV mit $Y = g(X_1, \ldots, X_n)$. Falls $X_1, \ldots, X_n$ diskret sind mit $p(x_1, \ldots, x_n)$, dann $E[Y] = \sum\limits_{X_1} \ldots \sum\limits_{X_n} g(x_1, \ldots, x_n)p(x_1, \ldots, x_n)$
stetiger Fall:
$E[Y] = \int\limits_{-\infty}^{+\infty} g(x_1, \ldots, x_n)f(x_1, \ldots, x_n)dx_1 \ldots dx_n$
falls das Integral absolut konvergiert.

Korollar zum Satz
Seien X,Y unabh. ZV
Dann $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$

4.3 Linearkombinationen von ZV

Satz
Der E'wert ist ein linearer Operator, d.h. $X_1, \ldots, X_n$ seien ZV mit E'werten $E[X_1], \ldots, E[X_n]$. Sei $Y = a_i + \sum\limits_{i=1}^{n} b_i X_i$. Dann ist $E[Y] = a_i + \sum\limits_{i=1}^{n} b_i E[X_i]$.

4.4 Varianz und Standardabweichung

Definition
X sei eine ZV mit E'wert E[X]. Dann heisst $var(X) = E[(X - E[X])^2]$ die *Varianz* von X (falls $var(X) < +\infty$). Es gilt aber auch $var(X) = E[X^2] - (E[X])^2$, was meist einfacher zu berechnen ist. Es gilt: $var(X) > 0$!
$\sigma(X) = \sqrt{var(X)}$ heisst *Standardabweichung.*

X diskret:
$var(X) = \sum\limits_i (x_i - \mu)^2 p(x_i)$, $\mu = E[X]$

X stetig:
$var(X) = \int\limits_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx$, $\mu = E[X]$

$E[X^2] = E[X]^2 - var(X)$

Satz
Sei X eine ZV mit $var(X) < +\infty$, $a, b \in \mathbb{R}$. Dann ist $var(a + bX) = b^2 var(X)$

1. Bernoulli-Verteilung $X \sim Be(p)$

$E[X] = p$
$var(X) = p(1-p)$

2. Binomial Verteilung $X \sim B(n,p)$
$E[X] = np$
$var(X) = np(1-p)$

3. Normalverteilung $X \sim N(\mu, \sigma)$
$E[X] = \mu$
$var(x) = \sigma^2$

4. Uniforme Verteilung $U \sim U(0,1)$
$E[X] = \frac{1}{2}$
$var(U) = \frac{1}{12}$

5. $X \sim Poisson(\mu)$
$E[X] = \mu$
$var(X) = \mu$

6. Geometrische Verteilung
$E[X] = \frac{1}{p}$
$var(X) = \frac{1-p}{p^2}$

7. Exponential Verteilung
$E[X] = \frac{1}{\lambda}$
$var(X) = \frac{1}{\lambda^2}$

Gamma Funktion

$\alpha > 0 \ \Gamma(\alpha) = \int\limits_0^\infty u^{\alpha-1} e^{-u} du$
$\alpha \in \mathbb{R}^+ : \ \Gamma(\alpha) = (\alpha-1)!$
$\alpha > 0 : \ \Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$
$\Gamma(\frac{1}{2}) = \sqrt{\Pi}$, $\Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\Pi}$

4.5 Kovarianz und Korrelation

$\mu_X = E[X]$, $\mu_Y = E[Y]$
$cov(X,Y) = E[(x-\mu_X)(Y-\mu_Y)]$
$cov(X,Y) = E[XY] - \mu_X \mu_Y$

$E[XY] = \iint xyf(x,y)dxdy$ falls gem. Dichtefkt.
$E[XY] = E[X] \cdot E[Y]$ falls unabh.

Rechenregeln

1. X,Y unabh. $\Rightarrow$ $cov(X,Y) = 0$
Die Umkehrung gilt aber nicht.
2. $cov(X,X) = var(X)$
3. $cov(aX,Y) = a \cdot cov(X,Y)$
4. $cov(X+Y,Z) = cov(X,Z) + cov(Y,Z)$
5. $cov(X+Y,Z+W) = cov(X,Z) + cov(X,W) + cov(Y,Z) + cov(Y,W)$
6. $cov(X,a) = 0$ für $a \in \mathbb{R}$
7. $cov(a + \sum\limits_{i=1}^n b_i X_i, c + \sum\limits_{j=1}^m d_j Y_j) = \sum\limits_{i=1}^n \sum\limits_{j=1}^m b_i d_j cov(X_i, Y_j)$
8. $var(a + \sum\limits_{i=1}^n b_i X_i) = \sum\limits_{i=1}^n \sum\limits_{j=1}^n b_i b_j cov(X_i, X_j)$
9. $var(X+Y) = var(X) + var(Y) + 2cov(X,Y)$
10. Falls $X_1, \ldots, X_n$ unabh.:
$var(\sum\limits_{i=1}^n X_i) = \sum\limits_{i=1}^n var(X_i)$
im Gegensatz dazu gilt immer: $E[\sum\limits_{i=1}^n x_i] = \sum\limits_{i=1}^n E[X_i]$

X, Y seien ZV mit endlichen Varianzen. Dann heisst $P_{XY} = \rho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$
die *Korrelation*.

Eigenschaften

- $\rho(X,Y) \in [-1,1]$
- falls $\rho(X,Y) = \pm 1$ dann $P[Y = aX + b] = 1$ für $a, b \in \mathbb{R}, b \neq 0$
- $\rho(X,Y) \approx \pm 1 \Rightarrow$ X und Y sind stark linear abhängig
- $\rho(X,Y) \approx 0 \Rightarrow$ X und Y sind schwach linear abhängig

5 Grenzwertsätze

Seien $X_1, \ldots, X_n$ unabh. ZV, mit $E[X_i] = \mu$ und $var(X_i) = \sigma^2 < +\infty$.
Dann gilt:
$\lim\limits_{n\to\infty} P(\mid \bar{X}_n - \mu \mid \geq \epsilon) = 0 \ \forall \epsilon > 0$

5.1 Zentraler Grenzwertsatz

$X_1, X_2, \ldots$ Folge identisch verteilter, unabh. ZV mit $E[X_i] = \mu$, $var(X_i) = \sigma^2 < \infty$
$s_n = \sum\limits_{i=1}^n x_i$; Standardisierung: $\frac{s_n - \mu n}{\sqrt{n}\sigma}$
Standardisierung: $U = \frac{X-\mu}{\sigma} \sim N(0,1)$

Zentraler Grenzwertsatz

$\lim\limits_{n\to\infty} P(\frac{s_n - \mu n}{\sqrt{n}\sigma} \leq x) = \Phi(x)$, Φ ist die Vertfkt. der $N(0,1)$-Verteilung. Kann auch so geschrieben werden:
$\lim\limits_{n\to\infty} P(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq x) \sim N(0,1)$

Beispiel Gesucht ist x, so dass $P[\sum\limits_{i=1}^{120} I_i > x] = 0.95$
Aus dem ZGS folgt:
$0.05 = 1 - P[\sum\limits_{i=1}^{120} I_i > x] = P[\sum\limits_{i=1}^{120} I_i \leq x] = P[\frac{\sum\limits_{i=1}^{120} I_i - 120E[I]}{\sqrt{120\,Var(I)}} \leq \frac{x - 120E[I]}{\sqrt{120\,Var(I)}}] \approx \Phi(\frac{x-120E[I]}{\sqrt{120Var(I)}})$

Monte-Carlo-Integration

$j(f) = \int_0^1 f(x)dx$ sei numerisch zu berechnen

Generiere unabh. auf $[0,1]$ gleichverteilte ZV. $U_1, \ldots, U_n$ und berechne $f(\bar{U}) = \frac{1}{n}\sum_{i=1}^{n} f(u_i)$.

Sei $var(f(U_i))$ endlich.

Nach dem Gesetz der grossen Zahlen gilt:

$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(U_i) = E[f(U_i)] = \int_0^1 dx$

$E[g(x)] = \int_{-\infty}^{\infty} g(x)\cdot f_x(x)dx$, wobei $f_x(x)$ die Dichte von x ist.

5.2 Normalapproximation der Binomialverteilung

Falls gilt $np(1-p) > 9$ so kann $B(n;p)$ durch $N(\mu = np; \sigma = \sqrt{np(1-p)})$ approximiert werden, sonst falls $np \leq 10$ und $n \geq 1500p$ durch $Poiss(\lambda = np)$.

Beispiel

VB auf α Niveau für Binomialvrtlg.

$\frac{X-np}{\sqrt{np(1-p)}} \sim N(0,1)$

$\Rightarrow \{X > \Phi^{-1}(1-\alpha)\cdot \sqrt{n\hat{p}(1-\hat{p})} + n\hat{p}\}$

6 Statistik

Definition

Eine *Stichprobe* vom Umfang n ist eine Folge $X_1, \ldots, X_n$ von unabh., ident. verteilten ZV. Eine Statistik ist eine ZV $g(X_1, \ldots, X_n)$, wobei $g: \mathbb{R}^n \to \mathbb{R}$

6.1 empirischer Mittelwert und empirische Varianz

Für eine Stichprobe $X_1, \ldots, X_n$, $n \geq 2$

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

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

Satz

Falls die X_i EW μ und var σ^2 haben, so ist $E[\bar{X}] = \mu$, $var(\bar{X}) = \frac{\sigma^2}{n}$. $\bar{X}$ ist eine ZV, μ eine Zahl. Man sagt, $\bar{X}$ ist ein Schätzer von μ.

Satz

Falls die X_i $N(\mu, \sigma^2)$-verteilt sind, so ist $\bar{X}$ $N(\mu, \frac{\sigma^2}{n})$-verteilt.

Satz

Sei $var(X_i) = \sigma^2$. Dann $E[S^2] = \sigma^2$. Man sagt, der Schätzer S^2 habe keinen Bias.

6.2 χ^2 Verteilung

Satz

Falls $X \sim N(0,1)$, dann ist $X^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$. Diese Verteilung nennt man χ_1^2-Verteilung. Lies: chi-Quadrat mit 1 Freiheitsgrad.

Satz

Seien $X_1, \ldots, X_n$ unabh. ident. verteilt, $X_i \sim \chi_1^2$. Dann ist $V = X_1 + \ldots + X_n$ χ_n^2-verteilt. χ_n^2 ist eine $\Gamma(\frac{1}{2}, \frac{1}{2})$-Verteilung. D.h., die Dichte ist

$f(t) = \frac{1}{\Gamma(\frac{n}{2})2^{n/2}} t^{\frac{n}{2}-1} e^{-\frac{t}{2}}$ $t \geq 0$

Falls $V \sim \chi_n^2$, dann $E[V] = n$, $var(V) = 2n$

Satz

Seien $X_1, \ldots, X_n$ eine Stichprobe von $N(\mu, \sigma^2)$-verteilten ZV. Dann ist $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$

6.3 t Verteilung

Satz

Sei $X_1, \ldots, X_n$ eine Stichprobe aus einer $N(\mu, \sigma^2)$-Population. Dann ist $t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$ t_{n-1} verteilt. Dichte der t-Verteilung: $f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\Pi n}\Gamma(\frac{1}{2})}\cdot(1 + \frac{x^2}{n})^{\frac{-n+1}{2}}$.

6.4 F Verteilung

U und V unabh. χ^2 ZV mit m bzw. n Freiheitsgraden, so wird die Verteilung:

$$W = \frac{U/m}{V/n}$$

als F Verteilung mit m und n Freiheitsgraden bezeichnet, geschrieben $F_{m,n}$

7 Konfidenzintervalle ...

7.1 ... für unbekannte Mittelwert μ einer Normalverteilung bei bekannter Varianz σ^2

Vertrauensniveau $(1-\alpha)$ wählen (z.B. 0.95)

Eine Stichprobe

1. Verteilung ist gleich $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
2. Konfidenzintervall
   a) *zweiseitig*
   $P[q_{\frac{\alpha}{2}} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq q_{1-\frac{\alpha}{2}}] = 1-\alpha$
   $\Rightarrow [\bar{X} - q_{1-\frac{\alpha}{2}}\cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + q_{\frac{\alpha}{2}}]$

b) *nach oben*
$P[\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq q_{1-\alpha}] = 1-\alpha$

c) *nach unten*
$P[q_\alpha \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}] = 1-\alpha$

Zwei Stichproben (Differenz der Mittelwerte)

gleich wie eine Stichprobe, aber mit folgender Verteilung:

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{m}+\frac{\sigma_Y^2}{n}}} \sim N(0,1)$$

7.2 ... für unbekannten Mittelwert μ einer Normalverteilung bei unbekannter Varianz σ^2

Vertrauensniveau $(1-\alpha)$ wählen (z.B. 0.95)

Eine Stichprobe

1. Verteilung ist gleich $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$
2. Konfidenzintervall

   a) *zweiseitig*
   $P[t_{n-1,\frac{\alpha}{2}} \leq T \leq t_{n-1,1-\frac{\alpha}{2}}] = 1-\alpha$
   $\Rightarrow [\bar{X_n} - t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X_n} + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}]$

   b) *nach oben*
   $P[T \leq t_{n-1,1-\alpha}] = 1-\alpha$

   c) *nach unten*
   $P[t_{n-1,\alpha} \leq T] = 1-\alpha$

Zwei Stichproben (Differenz der Mittelwerte)

gleich wie eine Stichprobe, aber mit folgender Verteilung:

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{S_P\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t_{m+n-2}$$

wobei S_p:

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

die gepoolte Varianz ist

falls m=n und Stichproben nicht unbedingt unabh.:
$D_i = X_i - Y_i$
$\frac{\bar{D}-(\mu_X-\mu_Y)}{S_D/\sqrt{n}} \sim t_{n-1}$

7.3 ... für unbekannte Varianz σ^2 einer Normalverteilung

Vertrauensniveau $(1-\alpha)$ wählen (z.B. 0.95)

Eine Stichprobe

$$\frac{S^2}{\sigma^2}(n-1) \sim \chi^2_{n-1}$$

Zwei Stichproben

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F_{(m-1,n-1)}$$

Wobei m (n) Anz. Experimente für ZV X (Y).

7.4 ... für unbekannten Parameter p einer Binomialverteilung

1. Verteilung ist gleich $Z = \frac{n\hat{p}-np}{\sqrt{n\hat{p}(1-\hat{p})}} \sim N(0,1)$, wobei $\hat{p} = \frac{k}{n}$ k= Anz. Erfolge bei n Versuchen.
2. Konfidenzintervall $[\hat{p} - \frac{q_{1-\frac{\alpha}{2}}}{n}\sqrt{n\hat{p}(1-\hat{p})} \leq p \leq \hat{p} - \frac{q_{\frac{\alpha}{2}}}{n}\sqrt{n\hat{p}(1-\hat{p})}]$

8 Schätztheorie

Ein Schätzer ist *erwartungstreu*, falls $E[\hat{\theta}] = \theta$.

8.1 Maximum Likelihood

Sei $X_1, \dots, X_n$ eine Stichprobe des Umfangs n einer Dichte $f(x,\theta)$, dann ist die gemeinsame Dichte von $(X_1, \dots, X_n)$ die Likelihood-Funktion.

- X diskret:
  $L(\Theta) = \prod_{i=1}^{n} P(X_i = x_i)$
- X stetig:
  $L(\Theta) = \prod_{i=1}^{n} f(x_i)$

Wähle Θ so dass die Realisierungen $X_1, \dots, X_n$ am wahrscheinlichsten sind. Um die Berechnung zu vereinfachen: logarithmiere, so ergibt sich aus dem Produkt eine Summe. $l(\Theta) = \log L(\Theta)$. Differenziere danach und setze gleich 0. $\frac{\partial}{\partial\Theta} l(\Theta) \stackrel{!}{=} 0$. Löse danach nach $\Theta \Rightarrow \hat{\Theta}_{ML}$

8.2 Momentenmethode

Berechne für $X \sim F(\Theta)$:

- $E[X]$ hängt von Θ ab (da $f(x)$ oder $P(X = x_i)$ von Θ abhängen).
  $E[X] = \int_{-\infty}^{\infty} x f(x) dx$ oder $E[X] = \sum_{i=1}^{n} x_i P(X = x_i)$
- $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Danach setze $E[X] = \bar{X}$ und löse nach $\Theta \Rightarrow \hat{\Theta}_{MM}$

p-tes Moment: Setze auch noch $E[X^2] = \bar{X^2} \dots E[X^p] = \bar{X^p}$ $\Rightarrow$ p-Gleichungen, löse nach Θ

9 Testen von Hypothesen

9.1 Neyman-Pearson Paradigma

Nullhypothese H_0:
Die zu zeigende Aussage, meist also $\mu = \mu_0$

Alternative H_A:
Was gilt, falls H_A nicht gilt:
$\mu \neq \mu_0$, $\mu > \mu_0$, $\mu < \mu_0$

Fehler 1. Art (α):
Verwerfung von H_0, obwohl H_0 richtig. $\alpha = P(Fehler\ 1.Art)$

Fehler 2. Art (β):
Keine Ablehnung von H_0, obwohl H_0 falsch. $\beta = P(Fehler\ 2.Art)$.

Macht Macht $= 1 - \beta$. W'keit, dass H_0 verworfen wird, wenn es tatsächlich falsch ist.

Man versucht α möglichst klein, und $1 - \beta$ möglichst gross zu wählen. Dazu wird α fixiert. Danach konstruiert man dazu einen Test mit möglichst grosser Macht.

Beispiel Macht, wenn μ_X, σ_X gleich wirklicher Mittelwert bzw. Varianz und G die zuvor ausgerechnete Grenze des VB's ist:
$Macht = 1 - P(Fehler\ 2.\ Art) = 1 - P(X \leq G) = 1 - P(\frac{X-\mu_x}{\sigma_X} \leq \frac{G-\mu_x}{\sigma_x}) = 1 - \Phi(\frac{G-\mu_x}{\sigma_x})$

9.2 Neyman-Pearson Lemma

Likelihood ratio $= \frac{f_0(x)}{f_A(x)}$

$$\rho_L = \begin{cases} 1 \text{ wenn } \frac{f_0(x)}{f_A(x)} < K_\alpha \\ 0 \text{ wenn } \frac{f_0(x)}{f_A(x)} > K_\alpha \end{cases}$$

K_α muss so gewählt sein, dass $E_0[\rho_L] = \alpha$

Der Likelihood-Test ρ_L ist der mächtigste Test unter den Tests ρ^* mit Signifikanzlevel $\alpha^* \leq \alpha$

9.3 Der t-Test

Mittelwert bei unbekanntem σ

Modellannahmen:

- $X_i \sim N(\mu, \sigma)$
- X_i unabhängig
- μ, σ unbekannt

Wir wollen testen, ob $\mu = \mu_0$ oder Alternative.

Für den t-Test gilt:
$T = \sqrt{n}\frac{\bar{X}-\mu_0}{S}$

Unter H_0: $T \sim t_{n-1}$

Verwerfungsbereich:

- $H_A : \mu \neq \mu_0$
  $VB = \{\mid T \mid \geq t_{n-1,1-\frac{\alpha}{2}}\}$
- $H_A : \mu > \mu_0$
  $VB = \{T > t_{n-1,1-\alpha}\}$
- $H_A : \mu < \mu_0$
  $VB = \{T < t_{n-1,\alpha}\}$

Falls T bzw. $|T|$ in VB $\Rightarrow$ H_0 verwerfen, sonst H_0 beibehalten.

2-Stichproben t-Test

$T = \frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{S_P\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t_{m+n-2}$. Wobei $S_p^2 = \frac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}$ die gepoolte Varianz ist.

9.4 Der z-Test

Mittelwert bei bekanntem σ

Modellannahmen:

- $X_i \sim N(\mu, \sigma)$
- X_i unabhängig
- μ unbekannt, σ bekannt

Wir wollen testen, ob $\mu = \mu_0$ oder Alternative.

Verteilung:
$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

Verwerfungsbereich:

- $H_A : \mu \neq \mu_0$
  $VB = \{|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}| > q_{1-\frac{\alpha}{2}}\}$
- $H_A : \mu > \mu_0$
  $VB = \{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > q_{1-\alpha}\}$
- $H_A : \mu < \mu_0$
  $VB = \{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < q_\alpha\}$

9.5 Likelihood-Ratio Tests für Multinomialverteilungen

m Zellen, n Beobachtungen $\Rightarrow$ Histogramm

$-2\ln\Lambda = 2\sum_{i=1}^{m} x_i \ln(\frac{x_i}{E_i})$ ist χ^2-verteilt mit $m-1-k$ Freiheitsgraden, wobei k die Anz. freier Parameter in H_0 ist.

χ^2 Anpassungstest

- Modellannahme: $X \sim F$, F irgend eine Verteilungsfkt.
- Nullhypothese $H_0 : F = Pois(\lambda)$
- Alternative $H_A : F \neq Pois(\lambda)$
- Die Teststatistik ist gegeben durch die χ^2 Teststatistik: man bildet die qudrierten Differenzen zwischen den beobachteten Häufigkeiten ($Beob_i$) und den erwarteten Häufigkeiten (Erw_i), man teilt durch die erwarteten Häufigkeiten (Erw_i) und summiert über alle möglichen Klassen.
  n= Anz. Klassen
  $\chi^2 = \sum\limits_{i=0}^{n} \frac{(Beob_i - Erw_i)^2}{Erw_i}$
  Unter H_0 ist die Teststatistik χ^2-verteilt mit f Freiheitsgraden, wobei
  f= Anz. Klassen - 1 - Anzahl geschätzter Parameter
- Entscheidung:
  VB=$\{\chi^2 > \chi^2_{f;p}\}$

Vorzeichen Test

Modellannahme:	D_i unabhängig mit Median $\tilde{m}$
Nullhypothese:	$\tilde{m} = 0$
Alternative:	$\tilde{m} \neq 0$
Teststatistik:	$T = \sum\limits_{i=1}^{n} 1\vert_{\{D_i > 0\}}$
Unter H_0:	$T \sim Bin(n, p)$

Beispiel für $T = 7$ und $n = 10$

$2P[T \geq 7] = 2 \sum\limits_{i=1}^{10} P[T = i] = 0.34$

Da $0.34 > 0.05$ wird H_0 beibehalten auf 5% Test

10 Methode der kleinsten Quadrate und lineare Regression

Beobachtungen:
$\{(x_i, y_i) \mid i = 1 \dots n\}$
y = die abhängige Variabel, "zu erklärende Variabel"
x = die unabhängige Variabel, "erklärende Variabel"

Ansatz:
Der Zusammenhang zw. x und y ist linear, d.h $y = \beta_0 + \beta_1 x$

Abweichung (Residuum) des i-ten Punktes:
$e_i = y_i - (\beta_0 + \beta_1 x_i)$

Ergebnis:

$$\hat{\beta}_0 = \frac{(\sum_{i=1}^{n} x_i^2)(\sum_{i=1}^{n} y_i) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i y_i)}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{n} (x_i - \bar{X})^2}$$

10.1 Statistisches Modell

$y_i = \beta_0 + \beta_1 x_i + e_i$
e_i: Beobachtungsfehler

Voraussetzungen:

1. e_i sind unabhängig
2. $E[e_i] = 0$
3. $var(e_i) = \sigma^2$
4. x_i sind fest

Varianz

$$var(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^{n} x_i^2}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$var(\hat{\beta}_1) = \frac{n\sigma^2}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{X})^2}$$

Korrelation

$$cov(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\sigma^2 \sum_{i=1}^{n} x_i}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

Bemerkungen

- $\hat{e_i} := y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$
- $S^2 = \frac{\sum_{i=1}^{n} (\hat{e_i})^2}{n-2}$ ist bias-freier Schätzer für σ^2
- Setze S^2 ein in Formel für $var(\hat{\beta}_0)$ bzw. $var(\hat{\beta}_1)$ und erhalte: $s^2_{\hat{\beta}_0}$ bzw. $s^2_{\hat{\beta}_1}$.
- Wenn e_i normalverteilt sind, dann sind $\hat{\beta}_0$, $\hat{\beta}_1$ auch normalverteilt, und $\frac{\hat{\beta}_0 - \beta_0}{s_{\hat{\beta}_0}}$ und $\frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}}$ sind t-verteilt mit n-2 Freiheitsgraden.

Korrelation und Regression

$S_{XX} = \frac{1}{n} \sum\limits_{i=1}^{n} (x_i - \bar{X})^2$ "Varianz von X"

$S_{YY} = \frac{1}{n} \sum\limits_{i=1}^{n} (y_i - \bar{Y})^2$ "Varianz von Y"

$S_{XY} = \frac{1}{n} \sum\limits_{i=1}^{n} (x_i - \bar{Y})(y_i - \bar{Y})$ "Kovarianz von X und Y"

Der Korrelationskoeffizient von X und Y ist:
$r = \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}}$