4.5.2015

VO Stat. Sch. k 2

I. Zufallsvektoren

Oft: Messen 2 (oder mehr) Merkmale gleichzeitig.
D.h. wir betrachten Zufallsvariablen X und Y, welche auf demselben Stichprobenraum Ω definiert sind.
Wir nennen dann $\begin{pmatrix} X \\ Y \end{pmatrix}$ einen 2-dimensionalen Zufallsvektor. Bsp: X ... Gewicht
Y ... Körpergröße einer Person

Wissen: Aus Kenntnis der Verteilungsfunktionen von X können alle Wahrscheinlichkeiten der Form $P(X \in A)$, A beliebig, bestimmt werden.

Wollen bestimmen:

$$P(70 \le X \le 93,\ 160 \le Y \le 171)$$

Kann man das aus Kenntnis der Verteilungsfunktionen von X bzw. Y bestimmen?
Nein: Es fehlt Information über die Kopplung (d.h. die Form der Abhängigkeit) von X und Y

Bsp (a) 2-facher Münzwurf (faire Münze)
Unabhängig $\checkmark$

X ... Ausgang 1. Wurf $P(X=0) = \frac{1}{2} = P(X=1)$
Y ... Ausgang 2. Wurf $P(Y=0) = \frac{1}{2} = P(Y=1)$

$P(X=1, Y=1) \overset{\text{(!!)}}{=} P(X=1) \cdot P(Y=1)$ — unabh.
$\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

4.5.2016

(b) Abhängigkeit

$P(X=0) = \frac{1}{2} = P(X=1)$

$P(Y=0) = \frac{1}{2} = P(Y=1)$

$P(X=1, Y=1)$

$\overset{''}{P(X=1) \cdot P(Y=1)}$

$\overset{''}{\frac{1}{2} \cdot 1 = \frac{1}{2}}$

$\Rightarrow$ Im Fall (a) und (b) sind Verteilungen von X und Y die gleichen!

Aber: $P(X=1, Y=1)$ ist in den Fällen (a) und (b) unterschiedlich!

($\rightarrow$ (Un)abhängigkeit)

Definition I.1

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ ein 2-dimensionaler Zufallsvektor. Dann heißt

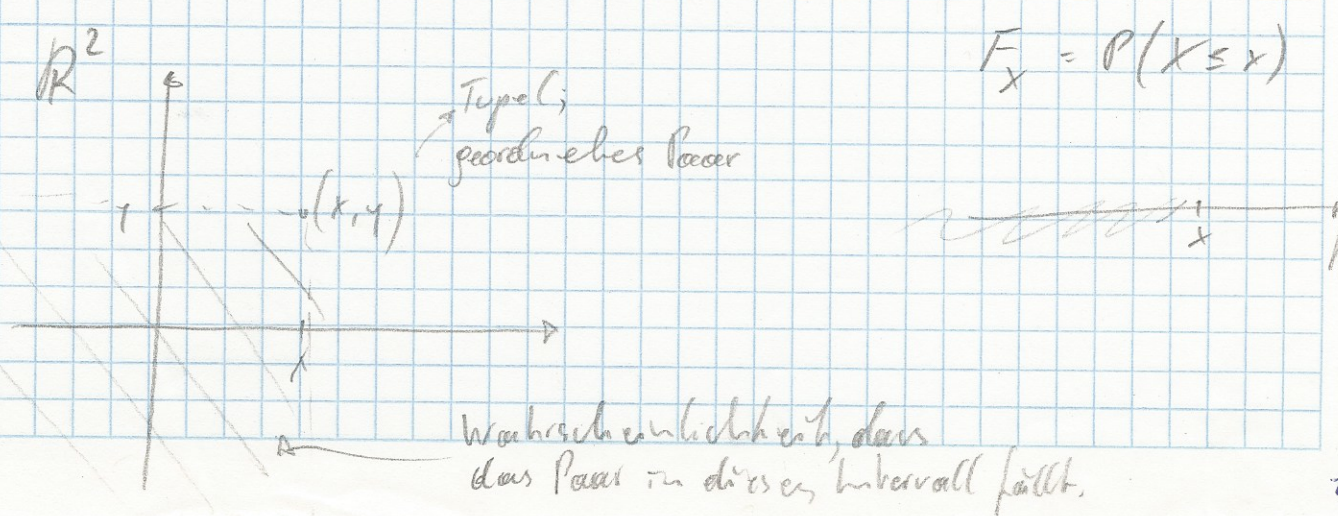
$$F(x,y) = P(X \leq x, Y \leq y) \quad \text{für } x \in \mathbb{R}, y \in \mathbb{R}$$

die Verteilungsfunktion von $\begin{pmatrix} X \\ Y \end{pmatrix}$

(auch: gemeinsame Verteilungsfunktion)

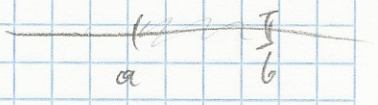
F_{XY}

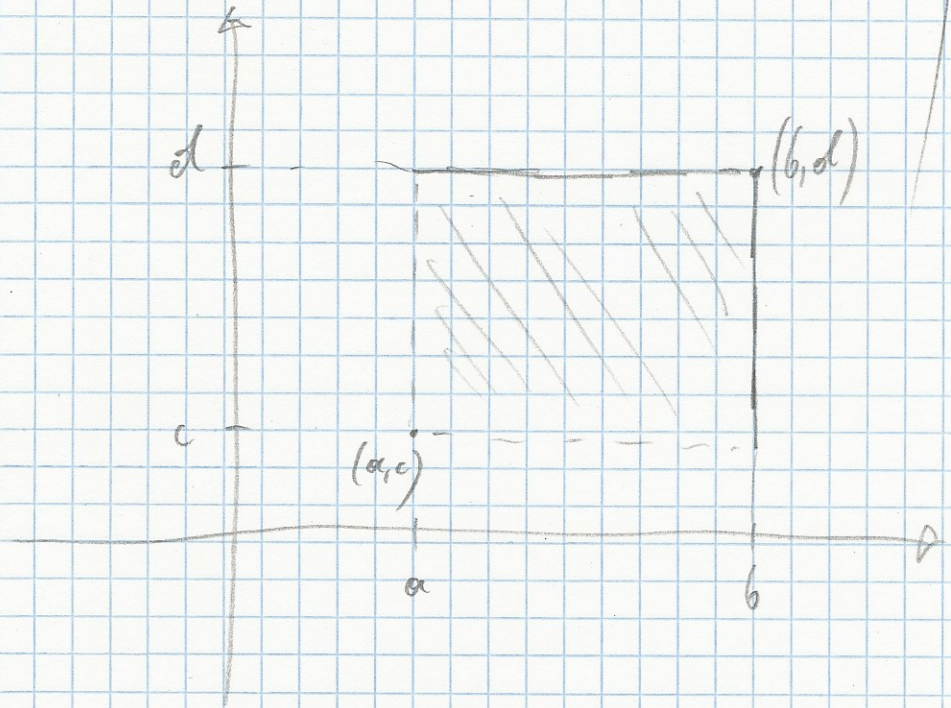
$F_X = P(X \leq x)$



Es g.t.: $P(a < X \leq b, c < Y \leq d)^*$

$$P(a < X \leq b) = P(X \leq b) - P(X < a)$$
$$= F_X(b) - F_X(a)$$





$$^*P(X \leq b, Y \leq d) - P(X \leq a, Y \leq d) - P(X \leq b, Y \leq c) + P(X \leq a, Y \leq c)$$

$$= F(b,d) - F(a,d) - F(b,c) + F(a,c)$$

Allgemein:

Jede Wahrscheinlichkeit der Form $P\left(\begin{pmatrix} X \\ Y \end{pmatrix} \in C\right)$ für $C \in \mathbb{R}^2$ kann aus Kenntnis der Verteilungsfunktion F_{XY} bestimmt werden.

I. 2.

(i) F_{xy} ist in jedem Argument monoton nicht fallend, d.h. für $x \le x'$ gilt

$F_{xy}(x,y) \le F_{xy}(x',y)$ für alle y

für $y \le y'$ $F_{xy}(x,y) \le F_{xy}(x,y')$ für alle x

$\lim_{x \to -\infty} F_{x,y}(x,y) = 0$ für beliebiges y

$\lim_{y \to -\infty} F_{x,y}(x,y) = 0$ für beliebiges x

2.

A_n Ereignisse in $\mathbb{R}$

Wahrscheinlichkeit des Durchschnitts = leere Menge = 0

(a) $A_n \supseteq A_{n+1} \quad \forall n \in \mathbb{N} \Rightarrow P(A_n) \xrightarrow[n \to \infty]{} P\left(\bigcap_{n=1}^{\infty} A_n\right)$

(b) $A_n \subseteq A_{n+1} \quad \forall n \in \mathbb{N} \Rightarrow P(A_n) \xrightarrow[n \to \infty]{} P\left(\bigcup_{n=1}^{\infty} A_n\right)$

$F_{X,Y}(\infty, y) = \lim_{x \to \infty} F_{X,Y}(x,y) = P(Y \le y) = F_Y(y)$ für jedes y

$\lim_{y \to \infty} F_{X,Y}(x,y) = P(X \le x) = F_X(x)$

$\lim_{x \to \infty} \lim_{y \to \infty} F_{X,Y}(x,y) = \lim_{x \to \infty} F_X(x) = 1$

$\lim_{y \to \infty} \lim_{x \to \infty} F_{X,Y}(x,y) = \lim_{y \to \infty} F_Y(y) = 1$

I.1 Diskrete Zufallsvektoren

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ ein Zufallsvektor, wobei X und Y diskrete Zufallsvariablen sind $\Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix}$ diskreter Zufallsvektor.

Die Zufallsvariable X nehmen Werte $X(\Omega) = \{x_1, x_2, \dots\} \subseteq \mathbb{R}$ an. $Y(\Omega) = \{y_1, y_2, \dots\}$

$X(\Omega)$ ist endlich oder abzählbar unendlich

$\Rightarrow$ $\begin{pmatrix} X \\ Y \end{pmatrix}$ nimmt Werte in $X(\Omega) \times Y(\Omega)$ an.

Def. I.3

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ diskreter Zufallsvektor $\Rightarrow$

$$p(x_i, y_i) = P(X = x_i, Y = y_i) \quad \text{für } x_i \in X(\Omega),\ y_i \in Y(\Omega)$$

(gemeinsame) Wahrscheinlichkeitsfunktion von $\begin{pmatrix} X \\ Y \end{pmatrix}$

Wenn notwendig: $p_{X,Y}(x_i, y_i)$

$$F_{X,Y}(x,y) = \sum_{\substack{x_i \le x \\ y_i \le y}} p_{X,Y}(x_i, y_i)$$

$F_X(x) = \sum_{x_i \le x} p(x_i)$

z.B. 1, 2, 3 — abzählbare Mengen

Bemerkung

$$F_{X,Y}(x,y) = P(X \le x, Y \le y) = P\left(\bigcup_{\substack{x_i \le x \\ y_i \le y}} \{X = x_i, Y = y_i\}\right)$$

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

$$= \sum_{\substack{x_i \le x \\ y_i \le y}} P(X = x_i, Y = y_i) = \sum_{\substack{x_i \le x \\ y_i \le y}} p_{X,Y}(x_i, y_i)$$

Bsp: I.5

3-facher Münzwurf, faire Münze, unabhängig

X ... Anzahl der Köpfe beim 1. Wurf

Y ... Gesamtanzahl der Köpfe bei allen drei Würfen

$X(\Omega) = \{0,1\}$

$Y(\Omega) = \{0,1,2,3\}$

$\Omega = \{(kkk),(kkz),(kzk),(zkk),(kzz),(zkz),(zzk),(zzz)\}$

X \ Y	0	1	2	3	p_X
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{4}{8}$
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{4}{8}$
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

p_Y

$p(\omega) = \frac{1}{8}$

$p_X(0) = P(X=0) = P(X=0, Y=0 \vee X=0, Y=1 \dots) = \frac{4}{8}$

$= p(0,0) + p(0,1) + \dots = \frac{1}{8} + \frac{2}{8} + \dots$

$p_X(1) = \frac{4}{8}$

$p_{X,Y}(0,0) = P(X=0, Y=0) = \frac{1}{8}$

$p_{X,Y}(0,1) = P(X=0, Y=1) = \frac{2}{8}$

$p_{X,Y}(0,2) = P(X=0, Y=2) = \frac{1}{8}$

$p_{X,Y}(0,3) = P(X=0, Y=3) = 0$

$p_Y(2) = P(Y=2) = P(X=0, Y=2 \vee X=1, Y=2)$

$= p_{X,Y}(0,2) + p_{X,Y}(1,2) = \frac{1}{8} + \frac{2}{8} = \frac{3}{8}$

$F_{X,Y}(1,2) = P(X \le 1, Y \le 2) =$

$= p(0,0) + p(0,1) + p(0,2) + (0,3) +$

$p(1,0) + p(1,1) + p(1,2) = \frac{7}{8}$

~~p_X~~

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ ein diskreter Zufallsvektor, dann

$$p_X(x_i) = \sum_{y_i \in Y(\Omega)} p_{X,Y}(x_i, y_i) \quad \forall\, x_i \in X(\Omega)$$

$$p_Y(y_i) = \sum_{x_i \in X(\Omega)} p_{X,Y}(x_i, y_i) \quad \forall\, y_i \in Y(\Omega)$$

auch: Randverteilungsfunktionen
Marginale Wahrscheinlichkeitsfunktionen

P.B. **Bemerk.**

$$p_X(x_i) = P(X = x_i) = P(X = x_i, Y = y_1 \vee X = x_i, Y = y_2, \dots)$$

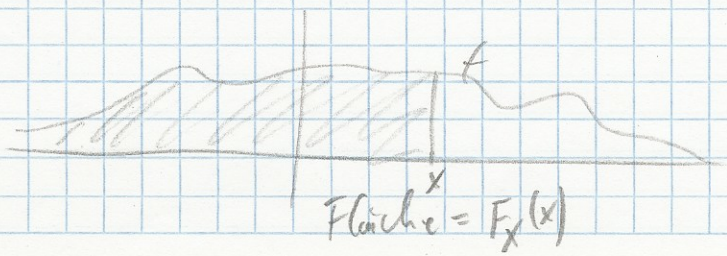
$$= P\Big(\bigcup_{y_j \in Y(\Omega)} \{X = x_i, Y = y_j\}\Big) = \sum_{y_j \in Y(\Omega)} P(X = x_i, Y = y_i)$$

$$= \sum_{y_j \in Y(\Omega)} p_{X,Y}(x_i, y_i)$$

X F_X

X ges.tal R.vable, falls es eine Funktion $f: \mathbb{R} \to \mathbb{R}$ gibt

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\, d\xi \quad \forall\, x \in \mathbb{R}$$



4.5.2015

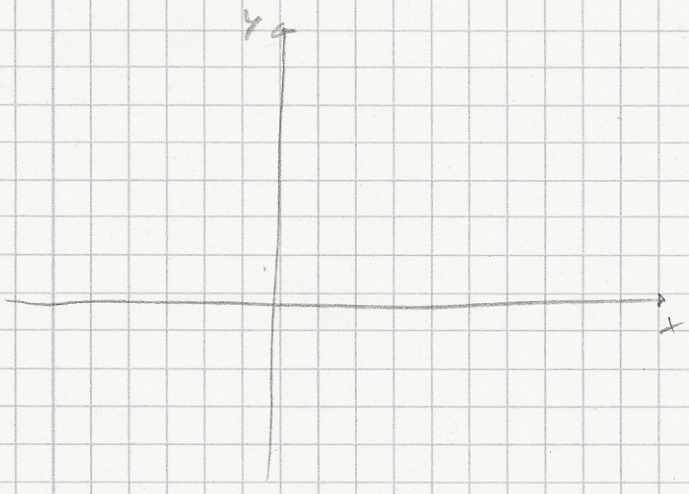
I. 2 Zufallsvektor u. d. Verteilung

Definition I. 7.

Zufallsvektor $\begin{pmatrix} X \\ Y \end{pmatrix}$ besitzt Dichte, wenn es eine Funktion $f: \mathbb{R}^2 \to \mathbb{R}$ gibt

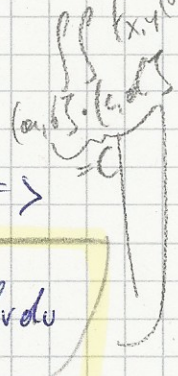
$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u,v)\,dv\,du \quad \forall x \in \mathbb{R}, y \in \mathbb{R}$$

f_{XY} ... gemeinsame Dichte



I. 8 Besitzt $\begin{pmatrix} X \\ Y \end{pmatrix}$ eine Dichte f_{XY} $\Rightarrow$

$$P(a < X \le b, c < Y \le d) = \int_a^b\int_c^d f_{XY}(u,v)\,dv\,du$$



$$P(a < X \le b, c < Y \le d) = F(b,d) - F(a,d) - F(b,c) + F(a,c) =$$

$$= \int_{-\infty}^{b}\int_{-\infty}^{d} f(u,v)\,dv\,du - \int_{-\infty}^{a}\int_{-\infty}^{d} f(u,v)\,dv\,du - \int_{-\infty}^{b}\int_{-\infty}^{c} f(u,v)\,dv\,du + \int_{-\infty}^{a}\int_{-\infty}^{c} f(u,v)\,dv\,du$$

$$= \int_{-\infty}^{b}\left[\int_{-\infty}^{d} f(u,v)\,dv - \int_{-\infty}^{c} f(u,v)\,dv\right]du + \int_{-\infty}^{a}\left[\int_{-\infty}^{c} f(u,v)\,dv + \int_{c}^{d} f(u,v)\,dv\right]du =$$

BRUNNEN

4.5.2015

$$\int_{-\infty}^{b} g(u)\,du - \int_{-\infty}^{b} h(u)\,du = \int_{-\infty}^{b} (g(u)-h(u))\,du$$

$$= \int_{-\infty}^{b} \left[\int_{c}^{d} f(u,v)\,dv \right] du - \int_{-\infty}^{a} \left[\int_{c}^{d} f(u,v)\,dv \right] du$$

$g(u)$ $\qquad$ $g(u)$

$$= \int_{a}^{b} g(u)\,du = \int_{a}^{b} \int_{c}^{d} f(u,v)\,dv\,du$$

I.9 Bes. Anz. $\begin{pmatrix} X \\ Y \end{pmatrix}$ ... R.vekt., so g.Ü.

„irgendeine Menge"

$$P\left(\begin{pmatrix} X \\ Y \end{pmatrix} \in C\right) = \iint_{C} f(u,v)\,dv\,du$$

z.B. $(a,b] \times (c,d]$

für C beliebige Teilmenge des $\mathbb{R}^2$

R.vekt. erfüllt immer:

N.B.

$$f_{XY}(u,v) \geq 0 \quad \forall\, u \in \mathbb{R}, v \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(u,v)\,dv\,du = 1$$

BRUNNEN

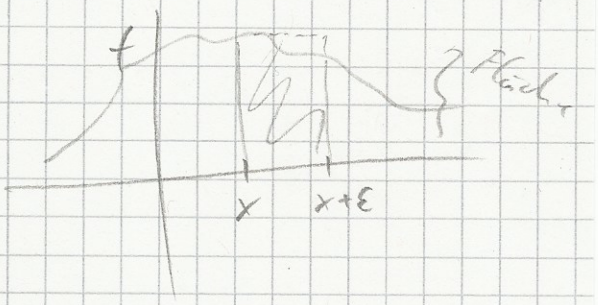
4.5.15

$p_X(x_i) = P(X = x_i)$

$p_{XY}(x_i, y_i) = P(X = x_i, Y = y_i)$

Interpretation der Dichte

Sei X Zufallsvariable mit Dichte f_X

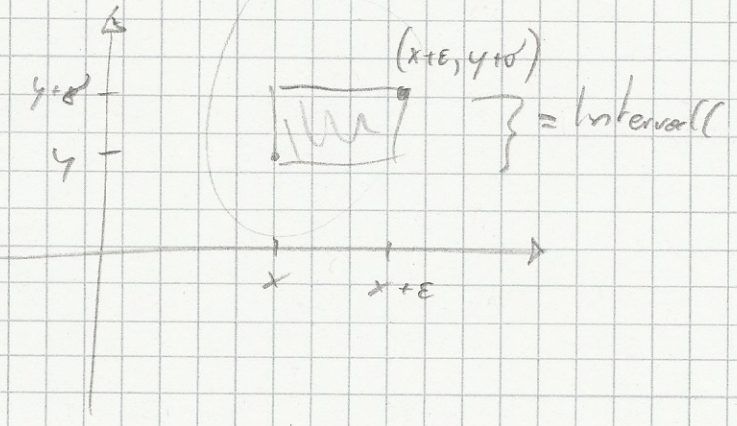


Bezeichnung: $P(X \in (x, x+\varepsilon])$ mit Wahrscheinlichkeit $\varepsilon > 0$

$$\int_x^{x+\varepsilon} f_X(u)\,du \approx \varepsilon \cdot f_X(x)$$

$P(X \in$ (kleines) Intervall mit linkem Endpunkt x ~~fällt~~ ist proportional mit $f_X(x)$

Faktor: Länge des Intervalls



$C = (x, x+\varepsilon] \times (y, y+\sigma]$

$x \in (x, x+\varepsilon]$ $y \in (y, y+\sigma]$

$$P(x < X \le x+\varepsilon, y < Y \le y+\sigma)$$

$$\int_x^{x+\varepsilon}\int_y^{y+\sigma} f_{XY}(u,v)\,dv\,du$$

$$\approx \varepsilon \cdot \sigma \cdot f_{XY}(x,y)$$

„Höhe"

BRUNNEN

I.10

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ ein Zufallsvektor m. V. Sei $F_{X,Y}$ 2x stetig partiell differenzierbar

$$\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial^2 F}{\partial x^2}, \frac{\partial^2 F}{\partial y^2}, \frac{\partial^2 F}{\partial x \partial y}$$

$$\Downarrow$$

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

Sei $f_{X,Y}$ gemeinsame Dichte von $\begin{pmatrix} X \\ Y \end{pmatrix}$

I.11

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,v)\,dv \quad \forall x \in \mathbb{R}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(u,y)\,du \quad \forall y \in \mathbb{R}$$

$\Rightarrow$ Randdichten

Beweisskizze

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y) = \lim_{y\to\infty} \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du \overset{!}{=} \int_{-\infty}^{x} \lim_{y\to\infty} \int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du$$

$$= \int_{-\infty}^{x} \underbrace{\left[\int_{-\infty}^{\infty} f_{X,Y}(u,v)\,dv\right]}_{g(u)} du = \int_{-\infty}^{x} g(u)\,du \quad \forall x \in \mathbb{R}$$

$\Rightarrow$ g ist Dichte f_X

4.5.2015

D.h. $f_X(u) = \int_{-\infty}^{\infty} f_{X,Y}(u,v)\,dv \quad \forall\, u \in \mathbb{R}$

Bemerkung II.4.2 (11)

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ Zufallsvektor mit Dichte $f_{X,Y}$,

dann:

$$P(a \le X \le b,\ c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(u,v)\,dv\,du$$

Abschn. II

5.5.2015

Beweis Wissen aus Prop I.9, dass

$$P(a \le X \le b,\ c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(u,v)\,dv\,du$$

Nun: $\{a < X < b,\ c < Y < d\} \subseteq \{a \le X \le b,\ c \le Y \le d\}$

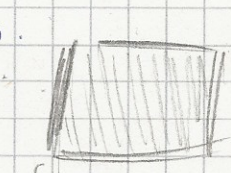
$$\subseteq \{a \le X \le b,\ c \le Y \le d\}$$

Es g.z.: Also zu zeigen:

$$P(a < X < b,\ c < Y < d) = P(a \le X \le b,\ c \le Y \le d)$$

$$= P\left(\begin{pmatrix} X \\ Y \end{pmatrix} \in \text{[shaded]}\right) = P\left(\begin{pmatrix} X \\ Y \end{pmatrix} \in \text{[shaded]}\right)$$

D.h.

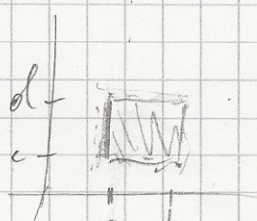


$$= \{(a,y) : c \le y \le d\} \to A_1$$

$$\cup \{(b,y) : c \le y \le d\} \to A_2$$

$$\cup \{(x,c) : a \le x \le b\} \to A_3$$

$$\cup \{(x,d) : a \le x \le b\} \to A_4$$



BRUNNEN

Beachte: W.ssen $P(a<X\le b, c<Y\le d) \le P(a\le X\le b, c\le Y\le d)$
weil $\{a<X\le b, c<Y\le d\} \subseteq \{a\le X\le b, c\le Y\le d\}$

$$P(a\le X\le b, c\le Y\le d) - P(a<X\le b, c<Y\le d) =$$

$$= P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in \boxed{}\right) - P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in \boxed{}\right)$$

$A\subset B \Rightarrow P(B|A) = P(B)-P(A)$

$$= P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in \boxed{} \setminus \boxed{}\right)$$

$$= P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in \boxed{\ \boxed{}}\right) = P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in A_1 \cup A_2 \cup A_3 \cup A_4\right) \le$$

$$\le P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in A_1\right) = P(X=a, c\le Y\le d) \le P(X=a)$$ W.ssen aus Prop I.11, dass ? eine Dichte bes. Ist.

$$= 5P(X=a) = 0$$

$$\Rightarrow 5P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in A_1\right) = 0$$

Analog: $0 \le P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in A_2\right) = P(X=b) = 0$

$$0 \le P\left(\begin{pmatrix}X\\Y\end{pmatrix} \in A_3\right) = P(a\le X\le b, Y\le c) \le 0 \quad P(Y\le c) = 0$$

$\le$ Y bes. Ist. Dichte

A_4 analog

Zusammenfassend: $0 \le P(a\le X\le b, c\le Y\le d) - P(a<X\le b, c<Y\le d) \le$

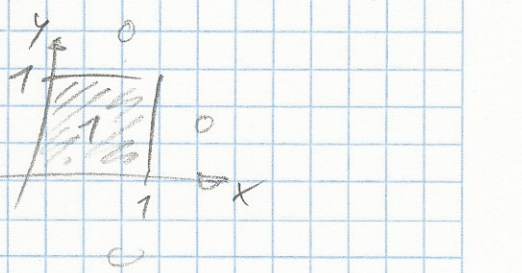
$$\le 0+0+0+0$$

Also sind die beiden Whs. gleich

Bsp I. 12 $f_{XY}(x,y) = \begin{cases} 1 & \text{für } 0 \le x \le 1 \\ 0 & \text{sonst} \end{cases}$

Randdichten

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \begin{cases} \int_{-\infty}^{\infty} 0\,dy & x<0 \vee x>1 \\ \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy & 0 \le x \le 1 \end{cases}$$



↳ gleichverteilt auf Einheitsquadrat $[0,1]^2$

$$\int_{-\infty}^{0} 0\,dy + \int_{0}^{1} 1\,dy + \int_{1}^{\infty} 0\,dy =$$

($\int_{-\infty}^{0} 0\,dy = 0$, $\int_{1}^{\infty} 0\,dy = 0$)

$$= y\Big|_0^1 = 1 - 0 = 1$$

Also $f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{sonst} \end{cases}$

d.h. X ist gleichverteilt auf $[0,1]$

Analog $f_Y(y) = \begin{cases} 1 & 0 \le y \le 1 \\ 0 & \text{sonst} \end{cases}$ d.h. Y ist gleichverteilt auf $[0,1]$

Bsp I 13

$$f_{XY}(x,y) = \begin{cases} 2x + 2y - 4xy & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{sonst} \end{cases}$$

Randdichten:

$f_X(x) = 0$ falls $x<0 \vee x>1$

Sei jetzt $0 \le x \le 1$

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_{-\infty}^{0} 0 + \int_{0}^{1} f_{XY}(x,y) + \int_{1}^{\infty} 0 = \int_{0}^{1} (2x + 2y - 4xy)\,dy =$$

$$= 2xy + \frac{2y^2}{2} - \frac{4xy^2}{2}\Big|_0^1 = 2x + 1 - 2x - (0+0+0) = 1$$

Also: $f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{sonst} \end{cases}$

$f_Y(y) = \begin{cases} 1 & 0 \le y \le 1 \\ 0 & \text{sonst} \end{cases}$

Bsp. I.14

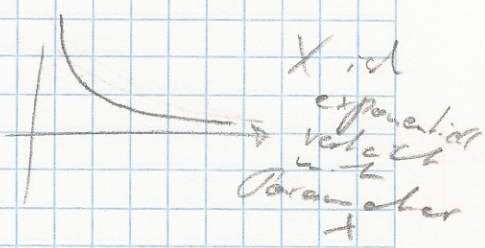
$$f_{XY}(x,y) = \begin{cases} \lambda^2 \cdot e^{-\lambda y} & 0 \le x \le y < \infty \\ 0 & \text{sonst} \end{cases} \quad (\lambda > 0)$$

$$f_X(x) = \int_x^\infty \lambda^2 \cdot e^{-\lambda y}\, dy = \lambda^2 \int_x^\infty e^{-\lambda y}\, dy =$$

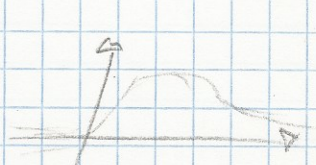


$$= \lambda^2 \cdot \left(-\frac{1}{\lambda} \cdot e^{-\lambda y}\Big|_x^\infty \right) = \lambda^2 \left(\frac{1}{\lambda} \cdot e^{-\lambda x} \right) =$$

$$= f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & x \ge 0 \\ 0 & \text{sonst} \end{cases}$$



$$f_Y(y) = \begin{cases} \lambda^2 y e^{-\lambda y} & y \ge 0 \\ 0 & y \le 0 \end{cases}$$



(Spezialfall der Gammaverteilung)

$$f_Y(y) = \int_0^y \lambda^2 \cdot e^{-\lambda y}\, dx = x\lambda^2 e^{-\lambda y}\Big|_0^y$$

$$= \lambda^2 y e^{-\lambda y} - 0 = \lambda^2 y e^{-\lambda y}$$

Bsp: I.15 Die bivariate Normalverteilung

Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ ein Zufallsvektor mit D.-dichte f_{XY}, welche die folgende Form hat:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho)^2}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right\}\right)$$

Parameter $\mu_X \in \mathbb{R}$, $\mu_Y \in \mathbb{R}$, $0 < \sigma_X < \infty$, $0 < \sigma_Y < \infty$,
$-1 \le \rho \le 1$

Beachte: $f_{XY}(x,y) \ge 0 \quad \forall x,y \in \mathbb{R}$

Man kann zeigen $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy\,dx = 1$ $\Big\}$ $\Rightarrow$ D.h. f_{XY} ist eine D.dichte

Wenn $\begin{pmatrix} X \\ Y \end{pmatrix}$ die oben angegebene D.dichte bez. hat, dann sagt man, dass $\begin{pmatrix} X \\ Y \end{pmatrix}$ bivariat normalverteilt ist.

Höhenschichtlinien sind Ellipsen

$$\{(x,y) : c = f_{XY}(x,y)\}$$

VO Stochastik 2

11.5.2015

$\begin{pmatrix} X \\ Y \end{pmatrix}$ m.t. Dichte

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho)^2}\left\{\frac{(x-\mu_x)}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y} - 2\rho\left[\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]\right\}\right)$$

Parameter $\mu_x \in \mathbb{R}$, $\mu_y \in \mathbb{R}$, $0 < \sigma_x < \infty$, $0 < \sigma_y < \infty$,
$-1 \leq \rho \leq 1$

Ist $\begin{pmatrix} X \\ Y \end{pmatrix}$ bivariat normalverteilt, dann gilt:

$$X \sim N(\mu_x, \sigma_x^2), \quad Y \sim N(\mu_y, \sigma_y^2)$$
$$\mu_x = E(X), \ \sigma_x^2 = \mathrm{Var}(X) \qquad \mu_y = E(Y), \ \sigma_y^2 = \mathrm{Var}(Y)$$

Warnung:

Aus $X \sim N$ und $Y \sim N$ folgt i.A. nicht, dass $\begin{pmatrix} X \\ Y \end{pmatrix}$ bivariat normalverteilt (s.u.) ist.

S.-od $X \sim N(\mu_x, \sigma_x^2)$ und $Y \sim N(\mu_y, \sigma_y^2)$ unabhängig, dann $\Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix}$ bivariat normalverteilt

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

(Unabhängigkeit)

11.5.2015

Bedeutung des Parameters ρ („rho")?

Behauptung: Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ bivariat normalverteilt

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

$$\rho\, \sigma_X \sigma_Y = \mathrm{Cov}(X,Y)$$

Beweis:

$$\mathrm{Cov}(X,Y) = E\big((X - \underbrace{E(X)}_{=\mu_X}) \cdot (Y - \underbrace{E(Y)}_{=\mu_Y})\big)$$

$$= E\big((X-\mu_X)(Y-\mu_Y)\big) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f_{XY}(x,y)\, dy\, dx$$

$u = \frac{x-\mu_X}{\sigma_X}$

$\frac{du}{dx} = \frac{1}{\sigma_X} \Rightarrow dx = du \cdot \sigma_X$

analog: $dy = dv \cdot \sigma_Y$

$$= \frac{1}{2\pi} \cdot \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) \exp\left(-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho \left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)\right\}\right) dy\, dx$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u \cdot v \cdot \exp\left(-\frac{1}{2(1-\rho^2)} \{u^2 + v^2 - 2\rho uv\}\right) dv\, du\, \sigma_X \sigma_Y$$

$$= \frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv \exp\left(-\frac{1}{2(1-\rho^2)} \{(v-\rho u)^2 + u^2(1-\rho^2)\}\right) dv\, du$$

$$= \frac{\sigma_X \sigma_Y}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} u \exp\left(\frac{-u^2}{2}\right) \cdot \left[\int_{-\infty}^{\infty} v \cdot \exp\left(-\frac{1}{2(1-\rho^2)}(v-\rho u)^2\right) dv\right] du$$

(innerer Integrand: Kern von v ab; Integral: Kern einer $N(\rho u, 1-\rho^2)$)

$$= \sigma_X \sigma_Y \int_{-\infty}^{\infty} u \cdot \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right) \left[\int_{-\infty}^{\infty} v \left(\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(v-\rho u)^2\right)\right) dv\right] du$$

(h: Kern von $N(0,1)$; g: Kern von $N(\rho u, 1-\rho^2)$; $\sqrt{\tau^2} = \tau$, v)

17.5.2013

$h(v)$ $\int_{-\infty}^{\infty} v\, g(v)\, dv$ = Erwartungswert der Zufallsvariable mit Dichte g

$= \sigma_x \sigma_y \int_{-\infty}^{\infty} u^2 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) du$ $\overset{!!}{=}$ ρ

$h(u)$ $\int_{-\infty}^{\infty} u^2 h(u)\, du = E(u^2)$ $U \sim N(0,1)$

$= \mathrm{Var}(u) + (E(u))^2$

$= \mathrm{Var}(u) = 1$

$= \sigma_x \sigma_y \rho$

I.3 Unabhängige Zufallsvariablen

Def I.16 Sei $(X,Y)'$ ein Zufallsvektor. Dann heißen X und Y unabhängig genau dann, wenn

$$P(X \in C, Y \in D) = P(X \in C) \cdot P(Y \in D)$$

für alle Teilmengen C und D von $\mathbb{R}$

(d.h. falls die Ereignisse $\{\omega \in \Omega : X(\omega) \in C\}$ und $\{\omega \in \Omega : Y(\omega) \in D\}$ unabhängig sind für jede Wahl von C und D)

11.5.15

Satz I.17

(a) X und Y unabhängig $\Leftrightarrow F_{XY}(x,y) = F_Y(y) \cdot F_X(x)$

$\forall x \in \mathbb{R}, y \in \mathbb{R}$

(b) Sei $(X,Y)'$ diskret:

X und Y unabhängig $\Leftrightarrow p_{XY}(x,y) = p_X(x) \cdot p_Y(y)$

$\forall x \in X(\Omega)$,
$y \in Y(\Omega)$

(c) $(X,Y)'$ bes. tzt Dichte f_{XY}

X und Y unabhängig $\Leftrightarrow f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$

für „alle" $x,y \in \mathbb{R}$
sofern Dichte existiert

Beweis (a) Seien $(X,Y)'$ unabhängig

$\Rightarrow$

$$F_{XY}(x,y) = P(X \le x, Y \le y) = P(\underbrace{X \in (-\infty, x]}_{X \in C}, \underbrace{Y \in (-\infty, y]}_{Y \in D}) =$$

Nach Definition I.16

$$= P(X \in (-\infty, x]) \cdot P(Y \in (-\infty, y]) = P(X \le x) \cdot P(Y \le y) =$$
$$= F_X(x) \cdot F_Y(y)$$

$\Leftarrow$ ohne Beweis

11.5.15

Beweis (b)

$\Rightarrow$ $p_{XY}(x_i, y_j) = P(X = x_i, Y = y_j) = P(X \in \underbrace{\{x_i\}}_{=C}, Y \in \underbrace{\{y_j\}}_{=D}) =$

nach Def. I.16 $= P(X = x_i) \cdot P(Y = y_j) = p_X(x_i) \cdot p_Y(y_j)$

für beliebige $x \in X(\Omega)$, $y \in Y(\Omega)$

$\Leftarrow$ Seien C, D beliebige Teilmengen von $\mathbb{R}$

$$P(X \in C, Y \in D) = P\Big(\bigcup_{\substack{x_i \in C \\ y_j \in D}} \{X = x_i, Y = y_j\}\Big) =$$

(abzählbare Vereinigung)

$P(A \cup B) = P(A) + P(B)$ bei Disjunktheit

$$= \sum_{\substack{x_i \in C \\ y_j \in D}} P(X = x_i, Y = y_j) =$$

Def. I.16

$$= \sum_{\substack{x_i \in C \\ y_j \in D}} p_{XY}(x_i, y_j) = \sum_{\substack{x_i \in C \\ y_j \in D}} p_X(x_i) p_Y(y_j) = \sum_{x_i \in C} \sum_{y_j \in D} p_X(x_i) p_Y(y_j) =$$

$$= \sum_{x_i \in C} \Big[p_X(x_i) \sum_{y_j \in D} p_Y(y_j) \Big] = \sum_{y_j \in D} p_Y(y_j) \sum_{x_i \in C} p_X(x_i) = P(Y \in D) \, P(X \in C)$$

für beliebige $C, D \subseteq \mathbb{R}$

(c) $\Rightarrow$ (Beweis nur unter der Voraussetzung, dass $F_{X,Y}$ 2x stetig partiell differenzierbar)

laut Prop. I.10 g.U. dann: $f_{XY}(x,y) = \frac{\partial^2 F_{XY}}{\partial x \, \partial y}$

Da X, Y unabhängig, wissen wir aus Teil (a)

$$F_{XY}(x,y) = F_X(x) \cdot F_Y(y)$$

D.h. $f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y}\left(F_X(x)F_Y(y)\right) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y} F_X(x)F_Y(y)\right) =$

$= \frac{\partial}{\partial x}\left(F_X(x)\frac{\partial F_Y(y)}{\partial y}\right) = \frac{\partial F_X(x)}{\partial x}\cdot\frac{\partial F_Y(y)}{\partial y} = f_X(x)\cdot f_Y(y)$

$\Leftarrow$

$f_{XY}(x,y) = f_X(x) f_Y(y)$

Nach Teil (a) genügt es zu zeigen:

$F_{XY}(x,y) = F_X(x)\cdot F_Y(y) \quad \forall x,y \in \mathbb{R}$

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(s,t)\,dt\,ds$$

$$= \int_{-\infty}^{x}\left[\int_{-\infty}^{y} f_X(s)\cdot f_Y(t)\,dt\right]ds$$

$$= \int_{-\infty}^{x}\left[f_X(s)\int_{-\infty}^{y} f_Y(t)\,dt\right]ds$$

$$= \int_{-\infty}^{x} f_X(s)\,ds \cdot \int_{-\infty}^{y} f_Y(t)\,dt = F_X(x)\cdot F_Y(y)$$

Bemerkung

X und Y unabhängig; $Var(X) < \infty$, $Var(Y) < \infty$

$\Rightarrow Cov(X,Y) = 0$

Beweis (im Fall einer Dichte):

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

Unabhängigkeit

$$0 = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}(x-E(x))(y-E(y))f_X(x)f_Y(y)\,dy\right]dx$$

$$= \int_{-\infty}^{\infty}\left[(x-E(x))f_X(x)\cdot\underbrace{\int_{-\infty}^{\infty}(y-E(y))f_Y(y)\,dy}_{c}\right]dx$$

$$= \int_{-\infty}^{\infty}(y-c)f_Y(y)\,dy = \int_{-\infty}^{\infty} y f_Y(y)\,dy - \int_{-\infty}^{\infty} c f_Y(y)\,dy$$

$$= \int_{-\infty}^{\infty} y f_Y(y)\,dy - c\int_{-\infty}^{\infty} f_Y(y)\,dy$$

$$= E(y) \qquad = E(y) - c$$

$$E(y) - E(y) = 0$$

$$= \int_{-\infty}^{\infty}(x-E(x))f_X(x)\cdot 0\,dx \Rightarrow \int_{-\infty}^{\infty} 0\,dx = \underline{0}$$

Bsp. I.8 Sei $(X,Y)'$ bivariat normalverteilt.

X,Y unabhängig $\Leftrightarrow \rho = 0$ $\left(\Leftrightarrow Corr(X,Y)=0 \Leftrightarrow Cov(X,Y)=0\right)$

X,Y unabhängig, dann $Cov(X,Y)=0 \Rightarrow \rho = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = 0$

Umgekehrt: Sei $\rho = 0$

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left(-\frac{1}{2}\left\{\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right\}\right) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)\cdot\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left(-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{1}{2\sigma_X^2}(x-\mu_X)^2\right)}_{\text{Randdichte von } X}\cdot\underbrace{\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left(-\frac{1}{2\sigma_Y^2}(y-\mu_Y)^2\right)}_{\text{Randdichte von } Y} =$$

$$= f_X(x)\cdot f_Y(y)$$

Satz I.19

Seien X und Y unabhängig. Sei $Z = g(X)$ und $W = h(Y)$ $\Rightarrow$ Z und W unabhängig

Beweis:

$$P(Z \in C, W \in D) = P(g(X) \in C, h(Y) \in D)$$

$$= P(X \in \underbrace{g^{-1}(C)}_{=A}, Y \in \underbrace{h^{-1}(D)}_{=B}) = P(X \in A, Y \in B)$$

Unabhängigkeit

$$= P(X \in A) P(Y \in B) = P(X \in g^{-1}(C)) \cdot P(Y \in h^{-1}(D))$$

$$= P(g(X) \in C) \cdot P(h(Y) \in D) = P(Z \in C, \cancel{W \in D}) \cdot P(W \in D)$$

für beliebige Mengen C, D

I.4 Erweiterung auf n-dimensionale Zufallsvektoren

Definition I.20: Sei $Z = (Z_1, \dots, Z_n)'$ ein n-dimensionaler Zufallsvektor, d.h. $Z_1, \dots, Z_n$ sind reellwertige Zufallsvariable, welche auf ein und demselben W.-Raum Ω definiert sind. Dann ist die Verteilungsfunktion von Z gegeben durch

$$F_Z(z) = F_Z(z_1, \dots, z_n) = P(Z_1 \le z_1, \dots, Z_n \le z_n)$$

wobei $z = (z_1, \dots, z_n)$ ist und $z_i \in \mathbb{R}$.

Bemerkung I.21: (i) Wir schreiben statt F_Z auch $F_{Z_1, \dots, Z_n}$ und sprechen dann von F_Z als der gemeinsamen Verteilungsfunktion von $Z_1, \dots, Z_n$.

~~(ii) Wir schreiben Zufallsvektoren Z und Zeilenvektoren als Spaltenvektoren. Dies ist immer kontextabhängig~~

(iii) Jede W. der Form $P(Z \in C)$ läßt sich aus Kenntnis von F_Z bestimmen.

Proposition I.22: Sei Z ein n-dimensionaler Zufallsvektor. Dann gilt:

(i) $F_Z(z_1, \dots, z_n)$ ist in jedem Argument monoton nichtfallend.

(ii) $\lim_{z_i \to -\infty} F_Z(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n) = 0$ für jedes i und beliebige feste z_j, $j \neq i$, $z_j \in \mathbb{R}$.

(iii) $\lim_{z_i \to \infty} F_Z(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n)$ existiert für jedes i und beliebige feste z_j, $j \neq i$, $z_j \in \mathbb{R}$.

Wir schreiben dafür $F_Z(z_1, \dots, z_{i-1}, \infty, z_{i+1}, \dots, z_n)$.

(iv) Rand (iii) gilt analog wenn mehrere z_i's $\to \infty$.

Bemerkung I. 22: (i) Die eindimensionalen Randverteilungen $F_{Z_i}(z_i)$ sind gegeben durch

$$F_{Z_i}(z_i) = F_Z(\infty, \dots, \infty, z_i, \infty, \dots, \infty).$$

(ii) Die zweidimensionalen Randverteilungen $F_{Z_i, Z_j}(z_i, z_j)$ sind gegeben durch

$$F_{Z_i, Z_j}(z_i, z_j) = F_Z(\infty, \dots, \infty, z_i, \infty, \dots, \infty, z_j, \infty, \dots, \infty).$$

Und so weiter...

Definition I. 23: Sei Z ein diskreter n-dimensionaler Zufallsvektor, d.h. jede Komponente Z_i von Z ist eine diskrete Zufallsvariable. Die Zufallsvariable Z_i nehme Werte in $\tilde{Z}_i \subseteq \mathbb{R}$ an, wobei $\tilde{Z}_i = \{z_i^{(1)}, z_i^{(2)}, \dots\}$ eine endliche oder abzählbare Menge ist. Die W.-Funktion von Z ist

$$p_Z(z) = P(Z = z) \quad \text{für } z \in \tilde{Z}_1 \times \dots \times \tilde{Z}_n,$$

d.h. $p_Z(z_1^{(i_1)}, \dots, z_n^{(i_n)}) = P(Z_1 = z_1^{(i_1)}, \dots, Z_n = z_n^{(i_n)})$.

Wir schreiben auch $p_{Z_1, \dots, Z_n}$ für p_Z und sprechen von der gemeinsamen W.-Funktion von $Z_1, \dots, Z_n$.

Proposition I. 24: Ist Z ein diskreter Zufallsvektor so gilt für $z = (z_1, \dots, z_n)$ mit $z_i \in \mathbb{R}$:

$$F_Z(z) = \sum_{\substack{\bar{z} \in \tilde{Z}_1 \times \dots \times \tilde{Z}_n \\ \bar{z}_1 \le z_1, \dots, \bar{z}_n \le z_n}} p_Z(\bar{z})$$

$F_{XY}(x,y) = \sum_{\substack{x_i \in X(\Omega) \\ y_j \in Y(\Omega)}} p_{XY}(x_i, y_j)$ $\Big\} = (x_i, y_j) \in X(\Omega) \times Y(\Omega)$

Bemerkung I.25: Für die eindimensionalen Rand-V.-Funktionen gilt:

Für $z_i \in \tilde{Z}_i$ haben wir

$$p_{Z_i}(z_i) = \sum_{\substack{z_1 \in \tilde{Z}_1 \\ z_2 \in \tilde{Z}_2 \\ \vdots \\ z_{i-1} \in \tilde{Z}_{i-1} \\ z_{i+1} \in \tilde{Z}_{i+1} \\ \vdots \\ z_n \in \tilde{Z}_n}} p_Z(z_1, z_2, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n)$$

„fest"

Analog für die 2-, 3-, ... dimensionalen Randverteilungen.

Wir wenden uns jetzt Zufallsvektoren, welche eine Dichte besitzen, zu.

Definition I.26: Sei Z ein n-dimensionaler Zufallsvektor mit Verteilungsfunktion F_Z. Falls es eine Funktion $f_Z : \mathbb{R}^n \to \mathbb{R}$ gibt mit

$$F_Z(z) = \int_{-\infty}^{z_n} \int_{-\infty}^{z_{n-1}} \cdots \int_{-\infty}^{z_1} f_Z(\xi_1, \dots, \xi_n)\, d\xi_1\, d\xi_2 \dots d\xi_n$$

für alle $z = (z_1, \dots, z_n)$, so heißt f_Z die Dichte von Z.

Bemerkung I.27: (i) Es gilt dann $P(Z \in C) = \int \cdots \int_C f_Z(\xi)\, d\xi_1 \dots d\xi_n$.

(ii) Es gilt dann $f_Z(z) \geq 0$ für alle z und $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_Z(\xi)\, d\xi_1 \dots d\xi_n = 1$.

(iii) Die 1-dimensionalen Randdichten erhält man wie folgt:

$$f_{Z_i}(z_i) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-1 \text{ fach}} f_Z(\xi_1, \dots, \xi_{i-1}, z_i, \xi_{i+1}, \dots, \xi_n)\, d\xi_1 \dots d\xi_{i-1}\, d\xi_{i+1} \dots d\xi_n$$

siehe 2-dimensionalen Fall

Analog für 2-, 3-, ... dimensionale Randdichten.

Wir verallgemeinern nun noch das Konzept der Unabhängigkeit auf endlich viele Zufallsvariablen.

Definition I.27a: Seien $Z_1, \dots, Z_n$ Zufallsvariablen auf einem gemeinsamen W.-Raum Ω. Die Variablen $Z_1, \dots, Z_n$ sind unabhängig falls

$$P(Z_1 \in A_1, Z_2 \in A_2, \dots, Z_n \in A_n) = \prod_{i=1}^{n} P(Z_i \in A_i)$$

(Produkt)

für alle $A_1, \dots, A_n$.

Es gilt: (i) $Z_1, \dots, Z_n$ unabhängig $\Leftrightarrow$

$$\Leftrightarrow F_{Z_1,\dots,Z_n}(z_1, \dots, z_n) = \prod_{i=1}^{n} F_{Z_i}(z_i) \quad \forall z_1, \dots, z_n \in \mathbb{R}$$

(ii) Falls eine gemeinsame Dichte existiert gilt:

$$Z_1, \dots, Z_n \text{ unabhängig} \Leftrightarrow f_{Z_1,\dots,Z_n}(z_1, \dots, z_n) = \prod_{i=1}^{n} f_{Z_i}(z_i) \quad \forall z_1, \dots, z_n \in \mathbb{R}$$

(iii) Im diskreten Fall gilt: $Z_1, \dots, Z_n$ unabhängig genau dann, wenn

$$p_{Z_1,\dots,Z_n}(z_1, \dots, z_n) = \prod_{i=1}^{n} p_{Z_i}(z_i) \quad \forall z_1 \in \mathcal{Z}_1, \dots, z_n \in \mathcal{Z}_n$$

(iv) $Z_1, \dots, Z_n$ unabhängig $\Rightarrow h_1(Z_1), \dots, h_n(Z_n)$ unabhängig.

VO Stat. f. I.-A. 2

12.5.2015

$Z_1, \dots, Z_n$ unabhängig

$$\Rightarrow F_{\underline{Z}}(z_1, \dots, z_n) = P(Z_1 \le z_1, Z_2 \le z_2, \dots, Z_n \le z_n)$$

$$= P(\underbrace{Z_1 \in (-\infty, z_1]}_{A_1}, \underbrace{Z_2 \in (-\infty, z_2]}_{A_2}, \dots, \underbrace{Z_n \in (-\infty, z_n]}_{A_3})$$

$$\overset{!}{=} \prod_{i=1}^{n} P(Z_i \in (-\infty, z_i])$$

$$= \prod_{i=1}^{n} P(Z_i \le z_i) = \prod_{i=1}^{n} F_{Z_i}(z_i)$$

$Z_1, \dots, Z_n$ unabhängig $\Rightarrow h_1(Z_1), \dots, h_n(Z_n)$ unabhängig

Bew: $P(h_1(Z_1) \in A_1, \dots, h_n(Z_n) \in A_n) = P(Z_1 \in \underbrace{h_1^{-1}(A_1)}_{B_1}, \dots, Z_n \in \underbrace{h_n^{-1}(A_n)}_{B_n}) =$

$$= P(Z_1 \in B_1, \dots, Z_n \in B_n) \overset{!}{=} P(Z_1 \in B_1) \cdot \dots \cdot P(Z_n \in B_n) =$$

$$= P(Z_1 \in h_1^{-1}(A_1)) \cdot \dots \cdot P(Z_n \in h_n^{-1}(A_n)) = P(h_1(Z_1) \in A_1) \cdot \dots \cdot P(h_n(Z_n) \in A_n)$$

für beliebige $A_1, \dots, A_n$

Bemerkung:

$F(x,y)$ F 2x stetig, partiell differenzierbar

$$\Rightarrow f_{XY} = \frac{\partial^2}{\partial x \partial y} F(x,y)$$

Proposition:

Z n-dimensionaler Zufallsvektor mit F_Z
Ist F n-mal stetig partiell differenzierbar, dann gilt für die Dichte f_Z:

$$f_Z(z) = \frac{\partial^n}{\partial z_1 \partial z_2 \dots \partial z_n} F_Z(z)$$

I.5 Erwartungswert eines Zufallsvektors

Definition I.28 Sei $Z = (Z_1, \dots, Z_n)'$ ein n-dimensionaler Zufallsvektor. Dann ist der Erwartungswert von Z gegeben durch:

$$E(Z) = \begin{bmatrix} EZ_1 \\ EZ_2 \\ \vdots \\ EZ_n \end{bmatrix}$$

sofern EZ_i für $i = 1, \dots, n$ wohldefiniert, d.h. sofern $E|Z_i| < \infty$ $\forall i = 1, \dots, n$

Zur Erinnerung:
X Zufallsvariable (diskret)

$$EX = \sum_{x_i \in \mathcal{X}(\Omega)} x_i p_X(x_i)$$

(falls $\sum_{x_i \in \mathcal{X}(\Omega)} |x_i| p_X(x_i) < \infty$)

Bsp I.29 Sei $\begin{pmatrix} X \\ Y \end{pmatrix}$ bivariat normalverteilt mit $\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho$
(d.h. $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$, $n = 2$)

$$E\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} EX \\ EY \end{bmatrix} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}$$

Bemerkung I.30

(i) Angenommen z bez. zu f_z. Dann bez. zu jede Komponente z_i auch eine Dichte f_{z_i}.

$$E(z) = \begin{bmatrix} E(z_1) \\ E(z_2) \\ \vdots \\ E(z_n) \end{bmatrix} = \begin{bmatrix} \int_{-\infty}^{\infty} z_1 f_{z_1}(z_1)\,dz_1 \\ \int_{-\infty}^{\infty} z_2 f_{z_2}(z_2)\,dz_2 \\ \vdots \\ \int_{-\infty}^{\infty} z_n f_{z_n}(z_n)\,dz_n \end{bmatrix} = \begin{bmatrix} \int_{-\infty}^{\infty} z_1 \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_z(z_1, z_2, \dots, z_n)\,dz_2 \cdots dz_n \right] dz_1 \\ \int_{-\infty}^{\infty} z_2 \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_z(z_1, z_2, \dots, z_n)\,dz_1\,dz_3 \cdots dz_n \right] dz_2 \\ \vdots \\ \int_{-\infty}^{\infty} z_n \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_z(z_1, \dots, z_n)\,dz_1 \cdots dz_{n-1} \right] dz_n \end{bmatrix}$$

(Variabel fest)

$$= \begin{bmatrix} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} z_1 f_z(z_1, \dots, z_n)\,dz_1 \cdots dz_n \\ \vdots \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} z_n f_z(z_1, \dots, z_n)\,dz_1 \cdots dz_n \end{bmatrix}$$

(gemeinsame Dichte)

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} f_z(z_1, \dots, z_n)\,dz_1 \cdots dz_n$$

(Vektor $\begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$!)

$$f_{z_1}(z_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_z(z_1, z_2, \dots, z_n)\,dz_2 \cdots dz_n$$

Diskreter Fall:

$$E(z) = \begin{bmatrix} E(z_1) \\ \vdots \\ E(z_n) \end{bmatrix} = \begin{bmatrix} \sum_{z_1 \in z(\Omega)_1} z_1 p_{z_1}(z_1) \\ \vdots \\ \sum_{z_n \in z(\Omega)_n} z_n p_{z_n}(z_n) \end{bmatrix} = \begin{bmatrix} \sum_{z_1 \in z(\Omega)_1} z_1 \left[\sum_{\substack{z_2 \in z(\Omega)_2 \\ \vdots \\ z_n \in z(\Omega)_n}} p_z(z_1, \dots, z_n) \right] \end{bmatrix}$$

$$p_{z_1}(z_1) = \sum_{\substack{z_2 \in z(\Omega)_2 \\ \vdots \\ z_n \in z(\Omega)_n}} p_z(z_1, \dots, z_n)$$

$$= \begin{bmatrix} \sum_{z \in z(\Omega)_1 \times \dots \times z(\Omega)_n} z_1 p_z(z_1, \dots, z_n) \\ \vdots \end{bmatrix} = \sum_{\substack{z_1 \in z(\Omega)_1 \\ \vdots \\ z_n \in z(\Omega)_n}} \begin{matrix} \text{Vektor} \\ z \end{matrix} \cdot p_z(z)$$

VO Statistik 2

(Abschrift von eigenen Unterlagen)

18.5.2015

$Z = (Z_1, \dots, Z_n)'$ Zufallsvektor

$$E(Z) = \begin{bmatrix} E(Z_1) \\ E(Z_2) \\ \vdots \\ E(Z_n) \end{bmatrix} \quad (\text{falls } E|Z_i| < \infty \ \forall i = 1, \dots, n)$$

Falls Z eine Dichte f_Z besitzt

$$E(Z) = \int_{\mathbb{R}^n} z f_Z(z) dz = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{n\text{-fach}} z f_Z(z) dz = \begin{bmatrix} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} z_1 f_Z(z_1, \dots, z_n) dz_1, \dots, dz_n \\ \vdots \\ \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} z_n f_Z(z_1, \dots, z_n) dz_1, \dots, dz_n \end{bmatrix}$$

($n \times 1$)

Falls Z diskret

$$E(Z) = \sum_{z \in \mathbb{Z}_1 \times \dots \times \mathbb{Z}_n} z p_Z(z) = \begin{bmatrix} \sum_{z \in \mathbb{Z}_1 \times \dots \times \mathbb{Z}_n} z_1 p_Z(z) \\ \vdots \\ \sum_{z \in \mathbb{Z}_1 \times \dots \times \mathbb{Z}_n} z_n p_Z(z) \end{bmatrix}$$

($n \times 1$)

Bsp: $n = 3$ Dichte f_Z $Z = (Z_1, Z_2, Z_3)'$

$W = Z_1^2 + \sin(Z_2) \cdot e^{Z_3}$ / $E(W) = ?$

oder: $W = Z_1^2 + Z_2^2 + Z_3^2$

oder: $W = Z_1 \cdot Z_2 \cdot Z_3$

$g(Z)$

Ang. W besitzt Dichte

$$E(W) = \int_{-\infty}^{\infty} w \cdot f_W(w) dw$$

Dafür benötigt man Kenntnis oder Dichte f_W

$$E(g(z)) = E(w) = \int_{\mathbb{R}^n} g(z) \cdot f_z(z)\, dz$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 \cdot f_x(x)\, dx$$

$V = X^2$ m.t Dichte f_V

$$E(X^2) = E(V) = \int v \cdot f_V(v)\, dv$$

Satz I.31 Sei Z ein $n \times 1$ Zufallsvektor und sei $g: \mathbb{R}^n \to \mathbb{R}$ eine Funktion; Sei $W = g(Z)$

(a) Sei Z diskret m.t W.-Funktion p_Z, so g. Ü.:

$$E(W) = E(g(Z)) = \sum_{z \in Z_1 \times \dots \times Z_n} g(z) \cdot p_Z(z) \quad \text{sofern} \sum_{z \in Z_1 \times \dots \times Z_n} |g(z)| p_Z(z) < \infty$$

W. ist selbst diskret m.t W.-Funktion p_W

$$E(W) = \sum_{w \in W(\Omega)} w \cdot p_W(w)$$

"

$E(g(Z))$

(b) Besitzt Z eine Dichte f_Z, so g. Ü.:

$$E(W) = E(g(Z)) = \int_{\mathbb{R}^n} g(z) f_Z(z)\, dz$$

$$= \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{n\text{-fach}} g(z_1, \dots, z_n) f_Z(z_1, \dots, z_n)\, dz_1, \dots, dz_n$$

sofern $\int_{\mathbb{R}} |g(z)| f_Z(z)\, dz < \infty$

Beweis (für (a)):

Sei $z(i)$, $i=1,\dots$ eine Aufzählung der Elemente in $Z_1 \times \dots \times Z_n$

$$\sum_{z \in Z_1 \times \dots \times Z_n} g(z) p_Z(z) = \sum_i g(z(i)) p_Z(z(i)) = w_1 \cdot \sum_{g(z(i))=w_1} p_Z(z(i)) + w_2 \cdot \sum_{g(z(i))=w_2} p_Z(z(i)) + \dots + w_s \sum_{g(z(i))=w_s} p_Z(z(i)) =$$

$g(z(i))$ durchläuft alle möglichen Werte von $w_1, \dots, w_s$

$$= w_1 \cdot P(W=w_1) + w_2 \cdot P(W=w_2) + \dots + w_s \cdot P(W=w_s) =$$

$$= \sum_s w_s p_W(w_s) = E(W)$$

$$p_W(W=w_1) = P(g(Z)=w_1) =$$

$$= P\left(\bigcup_{g(z(i))=w_1} \{Z=z(i)\} \right) = \sum_{g(z(i))=w_1} P(Z=z(i)) =$$

$$= \sum_{g(z(i))=w_1} p_Z(z(i))$$

Bsp: Z ist exponentialverteilt mit Parameter $\lambda > 0$

D.h. Dichte von Z ist:

$$f_Z(z) = \begin{cases} \lambda e^{-\lambda z} & \text{für } z > 0 \\ 0 & z \le 0 \end{cases}$$

Variante I

Bestimme f_W von W

→ berechne $E(W) = \int_{-\infty}^{\infty} w f_W(w) \, dw$

Variante II

$$E(W) = E(Z^2) = \int_{-\infty}^{\infty} z^2 \cdot f_Z(z) = \int_{-\infty}^{0} z^2 f_Z(z) dz + \int_0^{\infty} z^2 f_Z(z) dz = \lambda \int_0^{\infty} z^2 \cdot e^{-\lambda z} dz$$

$f(z) = z^2 \quad f'(z) = 2z$

$g(z) = \frac{e^{-\lambda z}}{-\lambda} \quad g'(z) = e^{-\lambda z}$

$$\int f(z)\cdot g'(z) = f(z)\cdot g(z) - \int f'(z) g(z)$$

$$= \underbrace{\lambda z^2 \frac{e^{-\lambda z}}{-\lambda}\Big|_0^\infty}_{=0} - 2\lambda \int z \frac{e^{-\lambda z}}{-\lambda}\, dz$$

$$= 2\int_0^\infty z\cdot e^{-\lambda z}\, dz = 2\cdot\left[\underbrace{z\cdot\frac{e^{-\lambda z}}{-\lambda}\Big|_0^\infty}_{=0} - \int_0^\infty \frac{e^{-\lambda z}}{-\lambda}\, dz\right] =$$

$f(z) = z \quad f'(z) = 1$

$g(z) = \frac{e^{-\lambda z}}{-\lambda} \quad g'(z) = e^{-\lambda z}$

$$= \frac{2}{\lambda}\int_0^\infty e^{-\lambda z}\, dz = \frac{2}{\lambda}\,\frac{e^{-\lambda z}}{-\lambda}\Big|_0^\infty = \frac{2}{\lambda}\left(0 - \left(\frac{1}{-\lambda}\right)\right) = \frac{2}{\lambda}\cdot\frac{1}{\lambda} = \frac{2}{\lambda^2}$$

$$\Rightarrow E(Z^2) = \frac{2}{\lambda^2}$$

Korollar I.32

Sei $\mathbf{Z}$ ein $n\times 1$ Zufallsvektor mit $E|Z_i| < \infty$ für $i = 1,\dots,n$

$$\underbrace{E(\alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_n Z_n + \alpha_0)}_{= E(\alpha' Z + \alpha_0)} = \underbrace{E(Z_1)\alpha_1 + \alpha_2 E(Z_2) + \dots + \alpha_n E(Z_n) + \alpha_0}_{= \alpha' E(Z) + \alpha_0}$$

wobei $\alpha_0, \dots, \alpha_n \in \mathbb{R}$

$\alpha \in \mathbb{R}^n$

Beweis: f_Z Dichte von Z

Setze $W = g(Z) = \alpha_1 Z_1 + \dots + \alpha_n Z_n + \alpha_0$ (d.h. $g(z) = \sum_{i=1}^n \alpha_i z_i + \alpha_0$)

Dann:

$$E(a_1 Z_1 + \dots + a_n Z_n + a_0) = E(g(Z)) \overset{\text{Satz I 3.1}}{=} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(z_1,\dots,z_n) f_Z(z_1,\dots,z_n)\, dz_1,\dots,dz_n =$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (a_1 z_1 + \dots + a_n z_n + a_0) f_Z(z_1,\dots,z_n)\, dz_1,\dots,dz_n =$$

$$= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Big[a_1 z_1 f_Z(z_1,\dots,z_n) + a_2 z_2 f_Z(z_1,\dots,z_n) + \dots + a_n z_n f_Z(z_1,\dots,z_n) + a_0 f_Z(z_1,\dots,z_n) \Big] dz_1,\dots,dz_n =$$

$$= a_1 \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} z_1 f_Z(z_1,\dots,z_n)\, dz_1,\dots,dz_n + \dots + a_n \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} z_n f_Z(z_1,\dots,z_n)\, dz_1,\dots,dz_n + a_0 \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_Z(z_1,\dots,z_n)\, dz_1,\dots,dz_n}_{=1} =$$

$$= a_1 E(Z_1) + a_2 E(Z_2) + \dots + a_n E(Z_n) + a_0$$

Kürzer: $a_1 Z_1 + a_2 Z_2 + \dots + a_n Z_n = a' Z$ mit $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$

Lineare Transformation

$Z \to W = a' Z + a_0$ (Skalar)

$$a_1 E(Z_1) + \dots + a_n E(Z_n) = a' E(Z)$$

$$Z \to \begin{bmatrix} W \\ U \\ V \end{bmatrix} = \begin{bmatrix} a' Z + a_0 \\ b' Z + b_0 \\ c' Z + c_0 \end{bmatrix} = \begin{bmatrix} a' Z \\ b' Z \\ c' Z \end{bmatrix} + \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix}$$

$$\begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} Z + \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix}$$

$3 \times n$ $\quad n \times 1$ $\quad 3 \times 1$

$\Rightarrow \boxed{DZ + d}$

$3 \times n$ Matrix $\quad 3 \times 1$ Vektor

Korollar I.34

Sei Z ein $n \times 1$ Zufallsvektor mit $E|Z_i| < \infty$ für $i = 1, \dots, n$
Sei D eine $q \times n$ Matrix und d ein $q \times 1$ Vektor
(wobei D und d nicht zufällig)

Dann:

$$E(DZ + d) = DE(Z) + d$$

(DZ: $q \times n$ mal $n \times 1$ $= q \times 1$; $DE(Z)$: $q \times n$ mal $n \times 1$ $= q \times 1$)

$$E\left(\begin{bmatrix} W \\ V \\ U \end{bmatrix}\right) = \begin{bmatrix} EW \\ EV \\ EU \end{bmatrix} = \begin{bmatrix} E(a'Z + a_0) \\ E(b'Z + b_0) \\ E(c'Z + c_0) \end{bmatrix} = \begin{bmatrix} a'E(Z) + a_0 \\ b'E(Z) + b_0 \\ c'E(Z) + c_0 \end{bmatrix} = \begin{bmatrix} a'E(Z) \\ b'E(Z) \\ c'E(Z) \end{bmatrix} + \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix} =$$

$E(DZ + d)$

mit $D = \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix}$ und $d = \begin{bmatrix} a_0 \\ b_0 \\ c_0 \end{bmatrix}$

$$= \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} E(Z) + \begin{pmatrix} a_0 \\ b_0 \\ c_0 \end{pmatrix} = DE(Z) + d$$

Beweis

Es genügt zu zeigen, dass die i-te Komponente der linken Seite gleich der i-ten Komponente der rechten Seite ist.

Linke Seite:

$E(i\text{-te Komponente von } DZ + d) = E(i\text{-te Zeile von } DZ + d^{(i)}) =$

Schreibe $D = [d_{ij}]_{\substack{i=1,\dots,q \\ j=1,\dots,n}}$ (Korollar I.32)

$= E([d_{i1}, d_{i2}, \dots, d_{in}]Z + d^{(i)}) = [d_{i1}, d_{i2}, \dots, d_{in}]E(Z) + d^{(i)}$

$= (i\text{-te Zeile von } D \times E(Z)) + d^{(i)} = i\text{-te Komponente von } DE(Z) + d$

Satz I 35 Seien $Z_1, \dots, Z_n$ unabhängige Zufallsvariablen auf gemeinsamen Stichprobenraum
Seien $h_i : \mathbb{R} \to \mathbb{R}$ Funktionen, sodass $E(|h_i(Z_i)|) < \infty$ für $i = 1, \dots, n$

Dann:

$$E(h_1(Z_1) h_2(Z_2) \dots h_n(Z_n)) = E(h_1(Z_1)) E(h_2(Z_2)) \dots E(h_n(Z_n))$$

Beweis:

Nur für $n = 2$ Z diskret

$$E(\underbrace{h_1(Z_1) h_2(Z_2)}_{= g(Z)}) \overset{\text{Satz I 31}}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(z_1, z_2) f_{(Z_1, Z_2)'}(z_1, z_2) \, dz_1 \, dz_2 =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(z_1) h_2(z_2) f_{(Z_1, Z_2)'}(z_1, z_2) \, dz_1 \, dz_2 \overset{!}{=} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} h_1(z_1) h_2(z_2) f_{Z_1}(z_1) f_{Z_2}(z_2) \, dz_1 \right) dz_2 =$$

$$= \int_{-\infty}^{\infty} h_2(z_2) f_{Z_2}(z_2) \left(\int_{-\infty}^{\infty} h_1(z_1) f_{Z_1}(z_1) \, dz_1 \right) dz_2 = \int_{-\infty}^{\infty} h_1(z_1) f_{Z_1}(z_1) \, dz_1 \cdot \int_{-\infty}^{\infty} h_2(z_2) f_{Z_2}(z_2) \, dz_2 =$$

(das Integral $\int_{-\infty}^{\infty} h_1(z_1) f_{Z_1}(z_1) \, dz_1$ hängt nicht von z_2 ab)

$$= E(h_1(Z_1)) \cdot E(h_2(Z_2))$$

Korollar I 36 Seien $Z_1, \dots, Z_n$ unabhängig mit $E|Z_i| < \infty$ für $i = 1, \dots, n$

Dann:

(i) $E(Z_i Z_j) = E(Z_i) \cdot E(Z_j)$ $\quad i \neq j$

(ii) $\mathrm{Cov}(Z_i, Z_j) = 0$ $\quad$ falls $\quad i \neq j$

Beweis:

(i) Wende Satz I 35 auf Zufallsvektor $\begin{pmatrix} Z_i \\ Z_j \end{pmatrix}$ und $h_1 = h_2 = z$ an

(ii) $\mathrm{Cov}(Z_i, Z_j) = E\left[(Z_i - E(Z_i))(Z_j - E(Z_j))\right]$

$= E\left[Z_i Z_j - Z_i E(Z_j) - Z_j E(Z_i) + E(Z_i) E(Z_j)\right] =$

$= E(Z_i Z_j) - E(Z_i E(Z_j)) - E(Z_j E(Z_i)) + E(E(Z_i) E(Z_j)) =$ (Konstant)

$= E(Z_i Z_j) - E(Z_i) E(Z_j) - E(Z_i) E(Z_j) + E(Z_i) E(Z_j) =$

$= E(Z_i Z_j) - E(Z_i) E(Z_j) \overset{\text{Unabhängigkeit + Teil (i)}}{=} E(Z_i) E(Z_j) - E(Z_i) E(Z_j) = 0$

19.5.15

VO Statistik 2

I.6 Varianz-Kovarianz Matrix

Erinnerung: X Zufallsvariable

$E(X)$ ... 1. Moment

$E(X^2)$ ... 2. (nichtzentrales) Moment

$$\mathrm{Var}(X) = E\left[(X - E(X))^2\right] = \underbrace{E(X^2) - (E(X))^2}_{\text{2. zentrales Moment}}$$

$$= \underbrace{E\left[(X - E(X))(X - E(X))'\right]}$$

Sei $Z = (Z_1, \dots, Z_n)'$ ein Zufallsvektor

$$E(Z) = \begin{bmatrix} E(Z_1) \\ \vdots \\ E(Z_n) \end{bmatrix} \quad \text{1. Moment}$$

$\mathrm{Var}(Z_1), \dots, \mathrm{Var}(Z_n)$

$\mathrm{Cov}(Z_1, Z_2), \mathrm{Cov}(Z_2, Z_3), \dots$ } 2. (zentrale) Momente

Falls $\mathrm{Var}(Z_i) < \infty$ für $i = 1, \dots, n$, so heißt

$$VC(Z) = \begin{bmatrix} \overbrace{\mathrm{Var}(Z_1)}^{\text{“}} & \mathrm{Cov}(Z_1, Z_2) & \mathrm{Cov}(Z_1, Z_3) & \dots & \mathrm{Cov}(Z_1, Z_n) \\ \mathrm{Cov}(Z_2, Z_1) & \mathrm{Cov}(Z_2, Z_2) & \dots & & \mathrm{Cov}(Z_2, Z_n) \\ \vdots & & & & \\ \mathrm{Cov}(Z_n, Z_1) & \mathrm{Cov}(Z_n, Z_2) & \dots & & \underbrace{\mathrm{Var}(Z_n)}_{\text{“}} \end{bmatrix}_{n \times n} = \left[\mathrm{Cov}(Z_i, Z_j)\right]_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$$

die Varianz-Kovarianz Matrix von Z

Bemerkung

(i) $VC(z)$ ist immer symmetrisch, d.h. $VC(z)' = VC(z)$

(ii) $VC(z)$ ist eine Diagonalmatrix genau dann, wenn z_i und z_j unkorreliert sind; $\forall i \neq j$

Definition I 36

Seien W_{ij} für $i = 1, \dots, m$ und $j = 1, \dots, n$ Zufallsvariablen, alle definiert auf ein und demselben Stichprobenraum

Dann heißt die $m \times n$ Matrix $W = [W_{ij}]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ Zufallsmatrix (bzw. zufällige Matrix)

Der Erwartungswert $E(W)$ der Zufallsmatrix W ist definiert als die $m \times n$ Matrix (nicht zufällig!) $E(W) = [E(W_{ij})]_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$

Satz I 37:

Sei z ein $n \times 1$ Zufallsvektor mit $Var(z_i) < \infty$ für alle $i = 1, \dots, n$

Dann:

$$VC(z) = E\underbrace{\left[\underset{n\times 1}{(z - E(z))}\underset{1\times n}{(z - E(z))'}\right]}_{= n \times n \text{ Zufallsmatrix}}$$

Beweis:

$$E\left((z - E(z))(z - E(z))'\right) = E\left[\begin{bmatrix} z_1 - E(z_1) \\ z_2 - E(z_2) \\ \vdots \\ z_n - E(z_n) \end{bmatrix}_{n\times 1} \underset{1\times n}{\left[z_1 - E(z_1), z_2 - E(z_2), \dots, z_n - E(z_n)\right]}\right] =$$

$$= E\begin{bmatrix} (Z_1-E(Z_1))(Z_1-E(Z_1)), & (Z_1-E(Z_1))(Z_2-E(Z_2)), & \dots, & (Z_1-E(Z_1))(Z_n-E(Z_n)) \\ (Z_2-E(Z_2))(Z_1-E(Z_1)), & (Z_2-E(Z_2))(Z_2-E(Z_2)), & \dots, & (Z_2-E(Z_2))(Z_n-E(Z_n)) \\ \vdots & & & \\ (Z_n-E(Z_n))(Z_1-E(Z_1)), & (Z_n-E(Z_n))(Z_2-E(Z_2)), & \dots, & (Z_n-E(Z_n))(Z_n-E(Z_n)) \end{bmatrix} =$$

$E[(Z_1-E(Z_1))(Z_1-E(Z_1))] = E[(Z_1-E(Z_1))^2] = Var(Z_1)$

$E[(Z_1-E(Z_1))(Z_2-E(Z_2))] = Cov(Z_1,Z_2)$

$$= \begin{bmatrix} E[(Z_1-E(Z_1))(Z_1-E(Z_1))], & E[(Z_1-E(Z_1))(Z_2-E(Z_2))], & \dots, & E[(Z_1-E(Z_1))(Z_n-E(Z_n))] \\ \vdots & & & \\ E[(Z_n-E(Z_n))(Z_1-E(Z_1))], & E[(Z_n-E(Z_n))(Z_2-E(Z_2))], & \dots, & E[(Z_n-E(Z_n))(Z_n-E(Z_n))] \end{bmatrix} =$$

$$= VC(Z)$$

Proposition I 38 Sei W eine $m \times n$ Zufallsmatrix mit $E|W_{ij}| < \infty \ \forall i,j$
Sei A eine $s \times m$ Matrix und B eine $n \times r$ Matrix
(beide nicht zufällig!)

Dann: $E(AWB) = AE(W)B$

$s \times m \quad m \times n \quad n \times r$

$s \times n$

$s \times r$

Beweis:

Es genügt Gleichheit des (i,j)-ten Elements auf beiden Seiten zu zeigen

(i,j)-te Element der linken Seite $= E\big((i,j)\text{-te Element von } \underbrace{(AW)}_{=C}B\big)$

$\rightarrow$ inneres Produkt der i-ten Zeile von (AW) mit j-ten Spalte von B,

d.h.

$$\sum_{k=1}^{n} (AW)_{ik} \cdot b_{kj} = \sum_{k=1}^{n}\sum_{l=1}^{m} a_{il} W_{lk} \cdot b_{kj} \quad \ldots (i,j)\text{-te Element von } AWB$$

$$(AW)_{ik} = \sum_{l=1}^{m} a_{il} W_{lk}$$

$$E\left(\sum_{k=1}^{n}\sum_{l=1}^{m} a_{il} W_{lk} \cdot b_{kj}\right) = \sum_{k=1}^{n}\sum_{l=1}^{m} E\left(a_{il} W_{lk} b_{kj}\right) = \sum_{k=1}^{n}\sum_{l=1}^{m} a_{il} E(W_{lk}) b_{kj}$$

$= (i,j)$-te Element der rechten Seite

Bsp: $z = \underset{2\times 1}{[z_1, z_2]'}$

$\mathrm{Var}(z_1+z_2) = \mathrm{Var}(z_1) + \mathrm{Var}(z_2) + 2\,\mathrm{Cov}(z_1, z_2)$

$\mathrm{Var}(\alpha z_1 + \beta z_2) = \alpha^2 \mathrm{Var}(z_1) + \beta^2 \mathrm{Var}(z_2) + 2\alpha\beta\,\mathrm{Cov}(z_1, z_2)$

Korollar I 39 Sei z ein $n \times 1$ Zufallsvektor mit $\mathrm{Var}(z_i) < \infty$
Sei A eine $m \times n$ Matrix und a ein $m \times 1$ Vektor
(A, a n.v.d. zufällig)

Dann:

$$VC(Az + a) = A\,VC(z)\cdot A'$$

Beweis: Satz I 37

$$VC(Az+\alpha) = E\Big[\big((Az+\alpha)-E(Az+\alpha)\big)\big((Az+\alpha)-E(Az+\alpha)\big)'\Big] =$$

$$= E\Big[(Az+\alpha-AE(z)-\alpha)(Az+\alpha-AE(z)-\alpha)'\Big] = E\Big[\big(A(z-E(z))\big)\big(A(z-E(z))\big)'\Big] =$$

$(AB)' = B'A'$

$$= E\Big[A(z-E(z))(z-E(z))'A'\Big] \overset{\text{Prop I 38}}{=} AE\Big[(z-E(z))(z-E(z))'\Big]A' =$$

Satz I 37

$$= A \cdot VC(z) \cdot A'$$

Bsp: $A = \underset{1\times 2}{[\alpha, \beta]}$ $\quad Az+\alpha = [\alpha,\beta]z = \alpha z_1 + \beta z_2$

$$Var(\alpha z_1 + \beta z_2) = VC(\alpha z_1 + \beta z_2) = VC(Az+\alpha) = AVC(z)A' = [\alpha,\beta]VC(z)\begin{bmatrix}\alpha\\ \beta\end{bmatrix} =$$

$$[\alpha,\beta]\underset{2\times 2}{\begin{bmatrix} Var(z_1), & Cov(z_1,z_2) \\ Cov(z_2,z_1), & Var(z_2)\end{bmatrix}}\underset{2\times 1}{\begin{bmatrix}\alpha\\ \beta\end{bmatrix}} = \underset{1\times 2}{[\alpha,\beta]}\underset{2\times 1}{\begin{bmatrix}\alpha Var(z_1) + \beta Cov(z_1,z_2) \\ \alpha Cov(z_1,z_2) + \beta Var(z_1)\end{bmatrix}} =$$

$$= \alpha^2 Var(z_1) + \alpha\beta Cov(z_1,z_2) + \alpha\beta Cov(z_1,z_2) + \beta^2 Var(z_2) =$$

$$= \alpha^2 Var(z_1) + 2\alpha\beta Cov(z_1,z_2) + \beta^2 Var(z_2)$$

19.5.2015

I 7 Die Multivariate Normalverteilung

Zur Erinnerung: $Z = (Z_1, Z_2)'$ heißt bivariat normalverteilt, falls die Dichte von Z die Form

$$f_Z(z_1, z_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{z_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{z_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{z_1-\mu_1}{\sigma_1}\right)\left(\frac{z_2-\mu_2}{\sigma_2}\right)\right)\right\}$$

besitzt, wobei $\mu_i \in \mathbb{R}$, $0 < \sigma_i < \infty$ für $i = 1, 2$, $-1 \le \rho \le 1$

$E(Z_1) = \mu_1$, $E(Z_2) = \mu_2$, $\mathrm{Var}(Z_1) = \sigma_1^2$, $\mathrm{Var}(Z_2) = \sigma_2^2$, $\mathrm{Corr}(Z_1, Z_2) = \rho$

$\mathrm{Cov}(Z_1, Z_2) = \rho\sigma_1\sigma_2$

Definition: $\mu = \begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} \in \mathbb{R}^2$ Berechne: $E(Z) = \mu$

$$\underset{2\times 2}{\Sigma} = \begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}$$ Berechne: $VC(Z) = \begin{bmatrix}\mathrm{Var}(Z_1) & \mathrm{Cov}(Z_1, Z_2)\\ \mathrm{Cov}(Z_1, Z_2) & \mathrm{Var}(Z_2)\end{bmatrix} =$

$$= \begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix} = \Sigma$$

Determinante $|\Sigma| = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = \sigma_1^2\sigma_2^2(1-\rho^2)$

$$\Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{bmatrix} = \frac{1}{1-\rho^2}\begin{bmatrix}\frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2}\\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2}\end{bmatrix}$$

Nun p.d. $Z = (Z_1, Z_2)'$ $n = 2$

$$f_Z(z) = \frac{1}{2\pi}(\det\Sigma)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(z-\mu)'\Sigma^{-1}(z-\mu)\right]$$

$$(z-\mu)'\Sigma^{-1}(z-\mu) = \underset{1\times 2}{(z_1-\mu_1, z_2-\mu_2)} \underset{2\times 2}{\begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} z_1-\mu_1 \\ z_2-\mu_2 \end{bmatrix}} \cdot \frac{1}{1-\rho^2} =$$

$$= \frac{1}{1-\rho^2} \underset{1\times 2}{(z_1-\mu_1, z_2-\mu_2)} \underset{2\times 1}{\begin{bmatrix} \frac{z_1-\mu_1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2}(z_2-\mu_2) \\ \frac{-\rho}{\sigma_1\sigma_2}(z_1-\mu_1) & \frac{z_2-\mu_2}{\sigma_2^2} \end{bmatrix}} =$$

$$= \frac{1}{1-\rho^2}\left(\frac{(z_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho}{\sigma_1\sigma_2}(z_1-\mu_1)(z_2-\mu_2) + \frac{(z_2-\mu_2)^2}{\sigma_2^2} \right) =$$

$$= \frac{1}{1-\rho^2}\left(\left(\frac{z_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{z_2-\mu_2}{\sigma_2}\right)^2 - 2\rho \left(\frac{z_1-\mu_1}{\sigma_1}\right)\left(\frac{z_2-\mu_2}{\sigma_2}\right) \right)$$

Definition I.40 Sei Z ein $n\times 1$ Zufallsvektor. Z multivariat normalverteilt, falls es einen Vektor $\mu \in \mathbb{R}^n$ und eine symmetrische und positiv definite $n\times n$ Matrix Σ gibt, so dass die Dichte von Z die Form

$$f_Z(z) = (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(z-\mu)'\Sigma^{-1}(z-\mu) \right]$$

besitzt.

Bemerkung I.41

$f_Z(z) > 0$ gilt immer. Man kann zeigen, dass

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_Z(z_1,\ldots,z_n)\,dz_1\ldots dz_n = 1 \Rightarrow f_Z \text{ tatsächlich eine Dichte}$$

Bsp: Seien $Z_1, \dots, Z_n$ unabhängig und

$Z_i \sim N(\mu_i, \sigma_i^2)$

Sei $Z = (Z_1, \dots, Z_n)'$

Dichte f_Z ?

$$f_Z(z) = \prod_{i=1}^{n} f_{Z_i}(z_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}(z_i - \mu_i)^2\right) =$$

$$= (2\pi)^{-\frac{n}{2}} \frac{1}{\sigma_1 \sigma_2 \cdots \sigma_n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{(z_i - \mu_i)^2}{\sigma_i^2}\right) =$$

Setze $\mu = (\mu_1, \dots, \mu_n)'$ $= (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(z-\mu)' \Sigma^{-1} (z-\mu)\right)$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}$$

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n^2} \end{bmatrix}$$

$\Rightarrow$ Z multivariat normalverteilt mit Parameter μ und Σ

VO Statistik 2

1.6.2015

$$\underset{n\times 1}{z} \sim N_n(\mu, \Sigma) \quad \Leftrightarrow \quad f_z(z) = (2\pi)^{-\frac{n}{2}} (\det\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(z-\mu)'\Sigma^{-1}(z-\mu)\right)$$

$\mu \in \mathbb{R}^n$, $n\times n$

? $\sum_{i=1}^{n} (z_i - \mu_i)(z_j - \mu_j)\cdot \frac{1}{\sigma_{ij}}$?

Bsp. I.41 Seien $z_i \sim N(\mu_i, \sigma_{ii})$ für $i = 1, \dots, n$. Weiters seien $z_1, \dots, z_n$ unabhängig

Sei $z = (z_1, \dots, z_n)'$, $\mu = (\mu_1, \dots, \mu_n)'$, $\Sigma = \begin{pmatrix} \sigma_{11} & & 0 \\ & \ddots & \\ 0 & & \sigma_{nn} \end{pmatrix}$

$$f_z(z) = \prod_{i=1}^{n} f_{z_i}(z_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{ii}}} \exp\left(-\frac{1}{2\sigma_{ii}}(z_i - \mu_i)^2\right) =$$

$z_i \sim N$

$$= (2\pi)^{-\frac{n}{2}} \frac{1}{\sqrt{\sigma_{11}\cdots\sigma_{nn}}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{(z_i-\mu_i)^2}{\sigma_{ii}}\right)$$

Setze $(z_1, \dots, z_n)' = z$

$$= (2\pi)^{-\frac{n}{2}} (\det\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(z-\mu)'\Sigma^{-1}(z-\mu)\right)$$

$$= \underbrace{(z_1-\mu_1, \dots, z_n-\mu_n)}_{1\times n} \underbrace{\begin{pmatrix} \frac{1}{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_{nn}} \end{pmatrix}}_{n\times n} \underbrace{\begin{pmatrix} z_1-\mu_1 \\ \vdots \\ z_n-\mu_n \end{pmatrix}}_{n\times 1}$$

$$= \underbrace{(z_1-\mu_1, \dots, z_n-\mu_n)}_{1\times n} \begin{pmatrix} \frac{z_1-\mu_1}{\sigma_{11}} \\ \frac{z_2-\mu_2}{\sigma_{22}} \\ \vdots \\ n\times 1 \end{pmatrix}$$

$$\sum_{i=1}^{n} \frac{(z_i-\mu_i)^2}{\sigma_{ii}}$$

Bemerkung I.42

(i) Man kann Beispiele von $Z_i \sim N(\mu_i, \sigma_{ii})$ finden, welche nicht unabhängig sind, sodass $Z = (Z_1, \dots, Z_n)'$ nicht multivariat normalverteilt ist.

(ii) Ist jedoch $Z \sim N_n(\mu, \Sigma)$ ein $n \times 1$ Zufallsvektor, dann ist automatisch jedes Z_i (univariat) normalverteilt.

Satz I.43

(ohne Beweis) Sei $Z \sim N_n(\mu, \Sigma)$, dann gilt:

$$\mu = E(Z), \quad \Sigma = VC(Z)$$

Satz I.44 (Reproduktionseigenschaft der Normalverteilung)

N.B.

(ohne Beweis) Sei $\underset{n\times 1}{Z} \sim N_n(\mu, \Sigma)$, dann gilt:

(a) $\underbrace{a'Z + \gamma}_{a_1 Z_1 + \dots + a_n Z_n + \gamma} \sim N(a'\mu + \gamma, a'\Sigma a) \quad \forall\, a \in \mathbb{R}^n \setminus \{0\}, \forall\, \gamma \in \mathbb{R}$

(b) $\underbrace{\underbrace{AZ}_{m\times n\ n\times 1} + \underbrace{a}_{m\times 1}}_{m\times 1} \sim N_m(A\mu + a, A\Sigma A') \quad \forall\, m \times n$ Matrizen A, welche $\operatorname{rang}(A) = m$ haben, $\forall\, a \in \mathbb{R}^m$

$U_{m\times 1}$ Zufallsvektor

Bemerkung: (i) Die Form der Parameter ist klar:

Wissen: $E(a'Z + \gamma) = a'E(Z) + \gamma = a'\mu + \gamma$

$VC(a'Z + \gamma) = VC(a'Z) = a'VC(Z)a = a'\Sigma a$

Ebenso: $E(AZ + a) = AE(Z) + a = A\mu + a$

$VC(AZ + a) = VC(AZ) = AVC(Z)A' = A\Sigma A'$

(ii) Teil (a) ist Spezialfall von Teil (b):

Setze: $m = 1$, $A = \alpha'$, $a = \mu$

Beachte: $\text{rang}(A) = \text{rang}(\alpha') = \text{rang}(\alpha) = 1$ genau dann, wenn $\alpha \neq 0$

(iii) Warum Annahme $\alpha \neq 0$ in Teil (a)?

Wenn $\alpha = 0$, dann ist $\alpha' Z + \mu = \mu$

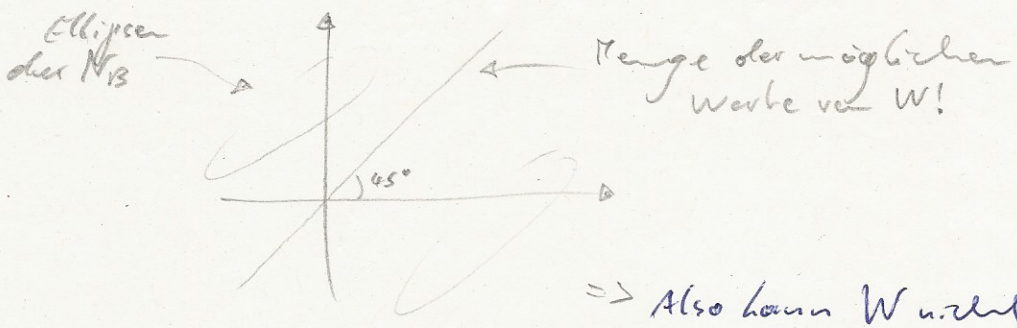
Beachte: $\text{Var}(\alpha' Z + \mu) = 0$!

Also kann $\alpha' Z + \mu = \mu$ nicht in unserem Sinne normalverteilt sein!

(iv) Warum Annahme $\text{rang}(A) = m$ in Teil (b)?

Bsp: $n = 2$, $m = 2$ $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $\text{rang}(A) = 1 < 2 = m$; $a = 0$

$$AZ = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} Z_1 + Z_2 \\ Z_1 + Z_2 \end{bmatrix} = W$$



=> Also kann W nicht bivariat normalverteilt sein

Korollar I.45 Sei $Z \sim N_n(\mu, \Sigma)$, dann g.d.l für jedes $i = 1, \dots, n$

$$Z_i \sim N(\mu_i, \sigma_{ii}),$$

wobei μ_i das i-te Element von μ ist und σ_{ii} das i-te Diagonalelement von Σ

Beweis: Betrachte $Z_i = e_i' Z$, wobei e_i der i-te ? Standardbasisvektor im $\mathbb{R}^n$ ist

$e_1 = (1,0,0,\dots,0)$
$e_2 = (0,1,0,\dots,0)$
$\vdots$
$e_n = (0,0,0,\dots,1)$

Setze $\mu = 0$

Nach Satz I.45 (a) folgt:

$$Z_i \sim N(e_i'\mu, e_i'\Sigma e_i) = N(\mu_i, \sigma_{ii})$$

i-te Spalte von Σ: $\begin{pmatrix} \sigma_{1i} \\ \sigma_{2i} \\ \vdots \\ \sigma_{ni} \end{pmatrix}$

Bemerkung I.46

(i) Sind $Z_i \sim N(\mu_i, \sigma_{ii})$ für $i = 1,\dots,n$, dann ist $Z = (Z_1,\dots,Z_n)'$ nicht notwendigerweise multivariat normalverteilt!

(Sogar, wenn Z_i unkorreliert)

(ii) Sind $Z_i \sim N(\mu_i, \sigma_{ii})$ für $i = 1,\dots,n$, dann ist

$$\sum_{i=1}^{n} \alpha_i Z_i + \gamma$$

im Allgemeinen nicht normalverteilt

Satz I.44'

Sei $Z \sim N_n(\mu, \Sigma)$, dann gilt:

Z_i's sind unabhängig $\Leftrightarrow$ Z_i's unkorreliert $\Leftrightarrow$ Σ Diagonalmatrix

Beweis:

Die 2. Äquivalenz ist trivial, da $\Sigma = VC(Z)$

Zur ersten Äquivalenz:

Die Implikation $\Rightarrow$ ist klar, da die Varianzen endlich sind und somit Unabhängigkeit Unkorreliertheit impliziert. (→ Korollar I.35)

zur Umkehrung:

Seien Z_i's unkorreliert, dann $\Sigma = \begin{pmatrix} \sigma_{11} & & 0 \\ & \sigma_{22} & \\ 0 & & \sigma_{nn} \end{pmatrix} \Rightarrow \det\Sigma = \sigma_{11} \cdot \ldots \cdot \sigma_{nn}$

Betrachte:

$$f_Z(z) = (2\pi)^{-\frac{n}{2}} (\det\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(z-\mu)' \Sigma^{-1} (z-\mu)\right) =$$

$$= (2\pi)^{-\frac{n}{2}} \frac{1}{\sqrt{\sigma_{11} \cdot \ldots \cdot \sigma_{nn}}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{(z_i - \mu_i)^2}{\sigma_{ii}}\right)$$

$$= \prod_{i=1}^{n} \frac{1}{2\pi \sqrt{\sigma_{ii}}} \exp\left(-\frac{1}{2\sigma_{ii}} (z_i - \mu_i)^2\right)$$

$$= \prod_{i=1}^{n} f_{Z_i}(i)$$ (Erinnerung: Wissen aus Kor. I.45, dass $Z_i \sim N(\mu_i, \sigma_{ii})$)

$\Rightarrow$ das heißt (nach Def. I.27 a), dass Z_i unabhängig sind

QUIZ: Seien $Z_i \sim N(\mu_i, \sigma_{ii})$ für $i = 1, \ldots, n$

(i) Seien Z_i unkorreliert. Folgt, dass Z_i unabhängig sind? NEIN!

(ii) – " – Folgt, dass $Z = (Z_1, \ldots, Z_n)$ multinormal? NEIN!

(iii) – " – Folgt, dass $\alpha' Z + \gamma$ normalverteilt? NEIN!
$\forall \alpha \neq 0, \forall \gamma$

Bsp: I 45': Seien $Z_i \sim N(\nu, \sigma^2)$ $\nu \in \mathbb{R}$, $0 < \sigma^2 < \infty$ (« $N_\nu[N\sigma]$ »)

Z_i unabhängig

Z_i's Zufallsstichprobe aus einer $N(\nu, \sigma^2)$-Verteilung

Schätzer: $\bar{Z} = \hat{\nu} = \frac{1}{n}\sum_{i=1}^{n} Z_i$, dann g.Ch:

$$\bar{Z} \sim N\left(\nu, \frac{\sigma^2}{n}\right)$$

Beweis: Wissen, dass $Z = (Z_1, \dots, Z_n)'$ multivariat normalverteilt ist

$$Z \sim N\left(\underbrace{\begin{pmatrix}\nu\\ \nu\\ \vdots\\ \nu\end{pmatrix}}_{n\times 1}, \underbrace{\begin{pmatrix}\sigma^2 & & 0\\ & \ddots & \\ 0 & & \sigma^2\end{pmatrix}}_{n\times n}\right) =: \Sigma$$

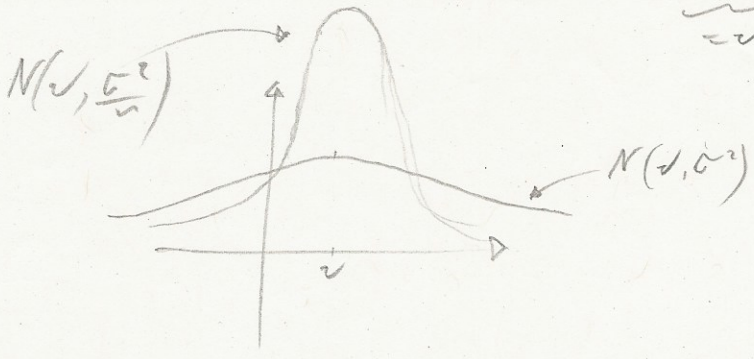
Satz I.44

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} 1 \cdot Z_i = \underbrace{\frac{1}{n}\underbrace{(1,1,\dots,1)}_{1\times n}}_{\alpha'} \underbrace{\begin{pmatrix}Z_1\\ Z_2\\ \vdots\\ Z_n\end{pmatrix}}_{n\times 1 \text{ multinormal}} \sim N\left(\frac{1}{n}\underbrace{(1,1,\dots,1)}_{1\times n}\underbrace{\begin{pmatrix}\nu\\ \nu\\ \vdots\\ \nu\end{pmatrix}}_{n\times 1}, \underbrace{\frac{1}{n}\underbrace{(1,1,\dots,1)}_{1\times n}\Sigma\underbrace{\begin{pmatrix}1\\ 1\\ \vdots\\ 1\end{pmatrix}}_{n\times 1}\frac{1}{n}}_{=\frac{n\sigma^2}{n^2}}\right)$$

$\alpha \neq 0$

$\nu + \nu + \dots + \nu = n\nu$, $\frac{1}{n} \cdot n\nu = \nu$

$$= N\left(\nu, \frac{\sigma^2}{n}\right)$$



I.8 Chi-Quadrat, t- und F-Verteilung

Definition I.46':

Seien $Z_i, i=1,\dots,n$ standardnormalverteilte Zufallsvariablen, welche unabhängig sind. Dann heißt die Verteilung von

$$X = \sum_{i=1}^{n} Z_i^2$$

Chi-Quadrat Verteilung mit n Freiheitsgraden (FG)

schr.: $X \sim \chi^2_n$

Bemerkung I.47': Man kann zeigen, dass die Dichte von X die Form hat:

$$f_X(x) = \begin{cases} c_n \cdot x^{\frac{n}{2}-1} \cdot e^{-\frac{x}{2}} & \text{für } x > 0 \\ 0 & \text{für } x \le 0 \end{cases} \quad \text{wobei } c_n = 2^{-\frac{n}{2}} \left(\Gamma\left(\tfrac{n}{2}\right)\right)^{-1}, \text{ wobei}$$

$$\Gamma\left(\tfrac{n}{2}\right) = \int_0^\infty v^{\frac{n}{2}-1} \cdot e^{-v}\, dv$$

Gammafunktion

$Z_1, \dots, Z_n$ Zufallsstichprobe $N(\nu, \sigma^2)$

Schätzung von σ^2? $\sigma^2 = E((Z_i - \nu)^2)$

$\nu = E(Z_i)$

$\hat{\nu} = \frac{1}{n}\sum_{i=1}^{n} Z_i$

„Schätzer" $\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (Z_i - \nu)^2$

← kann als Schätzer nur in einem äußerst unwahrscheinlichen Fall, in dem ν bekannt und σ^2 unbekannt ist, verwendet werden

W.'sen: $Z_i - \nu \sim N(0, \sigma^2)$

$\frac{Z_i - \nu}{\sigma} \sim N(0,1)$

Daher:

$$\tilde{\sigma}^2 = \left(\frac{1}{n}\sum_{i=1}^{n} \left(\frac{Z_i - \nu}{\sigma}\right)^2\right)\sigma^2 = \frac{\sigma^2}{n} \cdot \sum_{i=1}^{n} W_i, \quad \text{wobei } W_i \sim N(0,1) \text{ und } W_i \text{ unabhängig}$$

$$\Rightarrow \tilde{\sigma}^2 \cdot \frac{n}{\sigma^2} \sim \chi^2_n$$

VO Statistik 2

2.6.2015

Satz I.49': Zufallsstichprobe Z_i aus $N(\mu, \sigma^2)$

Schätzer $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (Z_i - \bar{Z})^2$

verzerrt

homogrieder (unverzerrter) Schätzer

$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2$

$\sum_{i=1}^{n} (Z_i - \bar{Z}) = \sum Z_i - \sum \bar{Z}$

$= \sum Z_i - n\bar{Z} = 0$

$\sum Z_i$

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

Ohne Beweis

Satz I.50 Zufallsstichprobe Z_i aus $N(\mu, \sigma^2)$

Dann sind: $\bar{Z}$ und $\hat{\sigma}^2$ unabhängig

$g(Z_1, \dots, Z_n)$ $\quad$ $h(Z_1, \dots, Z_n)$

Ohne Beweis

Definition I.51'

Sei $Z \sim N(0,1)$ und sei $S \sim \chi^2_q$, wobei Z und S unabhängig sind. Die Zufallsvariable T

$$T = \frac{Z}{\sqrt{\frac{S}{q}}}$$

folgt dann einer sogenannten (Student)-t-Verteilung mit q Freiheitsgraden.

Bemerkung I.52' Man kann zeigen, dass T eine Dichte folgender Form besitzt:

(i) $f_{T_q}(t) = \frac{\Gamma\left(\frac{q-1}{2}\right)}{\sqrt{q\pi}\,\Gamma\left(\frac{q}{2}\right)} \cdot \left(1+\frac{t^2}{q}\right)^{-\frac{(q+1)}{2}}$ für $t \in \mathbb{R}$

„Skaliert Masse, die drauße liegt"

(ii) Für $q \to \infty$ strebt $f_{T_q}(t,q)$ gegen die Standardnormalverteilung

Bsp: I.53' Zufallsstichprobe Z_i aus $N(\nu, \sigma^2)$

Wollen testen: $H_0: \nu = \nu_0$ $H_1: \nu \neq \nu_0$

Unter H_0:
$\bar{Z} - \nu_0 \sim N(0, \frac{\sigma^2}{n})$
$\frac{\bar{Z}-\nu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$

Wenn σ^2 und ν unbekannt:

$$t = \frac{\bar{Z}-\nu_0}{\sqrt{\frac{\hat{\sigma}^2}{n}}} = \frac{\sqrt{n}\,\bar{Z}-\nu_0}{\sqrt{\hat{\sigma}^2}}$$

→ Was ist die Verteilung von t angenommen H_0 stimmt?

$$t = \frac{\sqrt{n}\left(\bar{Z}-\nu_0\right)/\sigma}{\sqrt{\hat{\sigma}^2}/\sigma} = \frac{\left(\bar{Z}-\nu_0\right)/\frac{\sigma}{\sqrt{n}}}{\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}}} = \frac{\text{Zähler}}{\text{Nenner}}$$

Zähler $\sim N(0,1)$

Nenner $= \sqrt{\frac{(n-1)\hat{\sigma}^2}{\sigma^2}/(n-1)}$

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

Zähler unabhängig nach Satz I.50'; der Zähler ist eine affine Transformation von $\bar{Z}$ und Funktionen von unabhängigen ZV unabhängig bleiben

Nenner ist eine Funktion von $\hat{\sigma}^2$

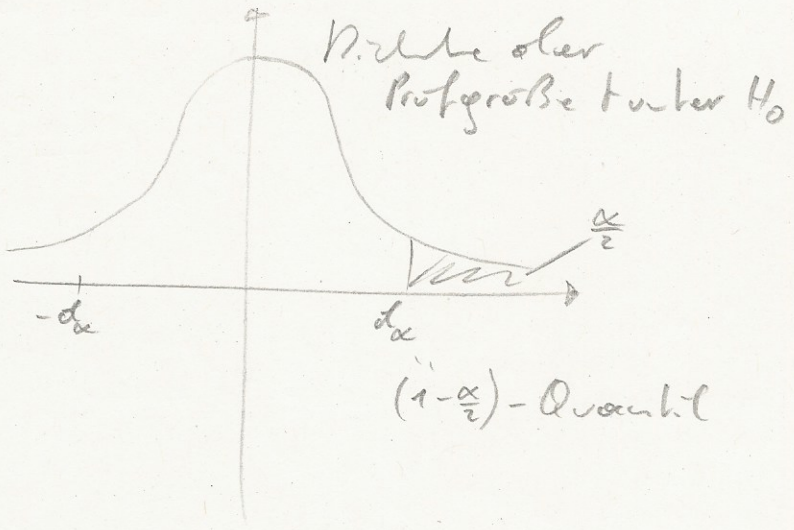
Also folgt nach Definition I.51', dass die Teststatistik t unter der Nullhypothese H_0 einer t-Verteilung mit n-1 FG folgt. Z

2.6.2015

Testprozedere

1. Wähle Signifikanzniveau α $(0 < \alpha < 1)$
2. Bestimme das $(1-\frac{\alpha}{2})$-Quantil d_α der Student-t-Verteilung mit $n-1$ Freiheitsgraden
3. Verwerfe H_0 falls $|t| \geq d_\alpha$

„t-Test"



Definition I.54'

Seien S_1 und S_2 unabhängig mit $S_1 \sim \chi^2_{q_1}$, $S_2 \sim \chi^2_{q_2}$

Die Zufallsvariable

$$F = \frac{S_1/q_1}{S_2/q_2}$$

hat eine F-Verteilung mit Freiheitsgraden q_1 und q_2

$$F \sim F_{q_1, q_2}$$

Bemerkung I.55

(i) F besitzt eine Dichte

$$f(x) = \begin{cases} c(q_1, q_2) \cdot x^{\frac{q_1}{2}-1} \left(1 + \frac{q_1}{q_2} x\right)^{-\frac{(q_1+q_2)}{2}} & \text{für } x > 0 \\ 0 & \text{sonst} \end{cases}$$

$$c(q_1, q_2) = \frac{1}{B\left(\frac{q_1}{2}, \frac{q_2}{2}\right)} \left(\frac{q_1}{q_2}\right)^{\frac{q_1}{2}} \quad \text{wobei} \quad B(\alpha, \beta) = \int_0^1 v^{\alpha-1}(1-v)^{\beta-1} dv$$

„Beta-Funktion"

(ii) $F_{q_1,q_2} \neq F_{q_2,q_1}$ (im Allgemeinen)

(iii) Ist $T \sim t_q$, dann ist $T^2 \sim F_{1,q}$

$$t^2 = \frac{"\chi^2_1/1"}{"\chi^2_{n-1}/n-1"}$$

(iv) Die F-Verteilung tritt als Verteilung von gewissen Testsstatistiken, etwa im Regressionsmodell, auf.

II. Bedingte Verteilungen

II.1 Diskreter Fall

Sei $(X,Y)'$ ein bivariater diskreter Zufallsvektor. Sei $X(\Omega)$ die Menge der möglichen Werte von X und sei $Y(\Omega) = \mathcal{Y}$ die Menge der möglichen Werte von Y. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $p_X(x_i) > 0$ für alle $x_i \in \mathcal{X}$ und auch $p_Y(y_j) > 0$ für alle $y_j \in \mathcal{Y}$ → siehe $P(A|B) = \frac{P(A \cap B)}{P(B)}$ ←

Def. II.1: Die bedingte W.-Funktion von Y gegeben $X = x_i \in \mathcal{X}$ ist definiert:

$$p_{Y|X}(y_j|x_i) = P(Y = y_j | X = x_i) = \frac{P(Y = y_j \wedge X = x_i)}{P(X = x_i)} =$$

$$= \frac{p_{X,Y}(x_i, y_j)}{p_X(x_i)} \quad \text{für beliebiges } x_i \in \mathcal{X}, y_j \in \mathcal{Y}$$

Satz II.2 Die bedingte W-Funktion erfüllt:

(i) $p_{Y|X}(y_j|x_i) \geq 0 \quad \forall x_i \in \mathcal{X}, y_j \in \mathcal{Y}$

(ii) $\sum_{y_j \in \mathcal{Y}} p_{Y|X}(y_j|x_i) = 1 \quad \forall x_i \in \mathcal{X}$

Beweis:

(i) ✓

(ii) $\sum_{y_j \in \mathcal{Y}} p_{Y|X}(y_j|x_i) = \sum_{y_j \in \mathcal{Y}} \frac{p(Y=y_j, X=x_i)}{p(X=x_i)} = \frac{1}{p(X=x_i)} \sum_{y_j \in \mathcal{Y}} p(Y=y_j, X=x_i) =$

$p(\{\omega \in \Omega : Y(\omega)=y_j, X(\omega)=x_i\})$ (disjunkt)

$= \frac{1}{p(X=x_i)} p\left(\bigcup_{y_j \in \mathcal{Y}} \{\omega \in \Omega : Y(\omega)=y_j, X(\omega)=x_i\}\right)$

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

$= \frac{1}{p(X=x_i)} p(\{\omega \in \Omega : X(\omega)=x_i\}) = \frac{1}{p(X=x_i)} p(X=x_i) = 1$

VO Stochastik 2

8.6.2015

$(X,Y)'$ diskreter Zufallsvektor

$$p_{Y|X}(y_j|x_i) = P(Y=y_j|X=x_i) = \frac{P(Y=y_j, X=x_i)}{P(X=x_i)} = \frac{p_{X,Y}(x_i,y_j)}{p_X(x_i)}$$

„Bedingte W-Funktion von Y gegeben $X=x_i$"

- $p_{Y|X}(y_j|x_i) \geq 0 \quad \forall x_i \in \mathcal{X}, \forall y_j \in \mathcal{Y}$
- $\sum_{y_j \in \mathcal{Y}} p_{Y|X}(y_j|x_i) = 1 \quad \forall x_i \in \mathcal{X}$

Definition II.2 Sei $(X,Y)'$ diskreter Zufallsvektor. Dann heißt

$$F_{Y|X}(y|x_i) = P(Y \leq y | X=x_i) = \frac{P(Y \leq y, X=x_i)}{P(X=x_i)}$$

die bedingte Verteilungsfunktion von Y gegeben $X=x_i$

Proposition II.4 Es gilt für jedes $x_i \in \mathcal{X}$

$$F_{Y|X}(y|x_i) = \sum_{\substack{y_j \leq y \\ y_j \in \mathcal{Y}}} p_{Y|X}(y_j|x_i)$$

Beweis: $F_{Y|X}(y|x_i) = \frac{P(Y \leq y, X=x_i)}{P(X=x_i)} =$

disjunkt $= \frac{P\left(\bigcup_{\substack{y_j \leq y \\ y_j \in \mathcal{Y}}} \{Y=y_j, X=x_i\}\right)}{P(X=x_i)} = \sum_{\substack{y_j \leq y \\ y_j \in \mathcal{Y}}} \frac{P(Y=y_j, X=x_i)}{P(X=x_i)}$

$$F_X(x) = \sum_{\substack{x_i \leq x \\ x_i \in \mathcal{X}}} p_X(x_i)$$

$$\| \quad P(X \leq x)$$

$$\| \quad P\left(\bigcup_{\substack{x_i \leq x \\ x_i \in \mathcal{X}}} \{X=x_i\}\right)$$

8.5.15

$$= \sum_{\substack{y_j \le y \\ y_j \in \mathcal{Y}}} \frac{P(Y=y_j, X=x_i)}{P(X=x_i)} = \sum_{\substack{y_j \le y \\ y_j \in \mathcal{Y}}} p_{Y|X}(y_j|x_i)$$

Satz II.5

X, Y sind unabhängig $\Leftrightarrow p_{Y|X}(y_j|x_i) = p_Y(y_j) \quad \forall y_j \in \mathcal{Y}, \forall x_i \in \mathcal{X}$

$\Leftrightarrow p_{X|Y}(x_i|y_j) = p_X(x_i) \quad \forall x_i \in \mathcal{X}, \forall y_j \in \mathcal{Y}$

Beweis:

$\Rightarrow$ $p_{Y|X}(y_j|x_i) = P(Y=y_j|X=x_i) = \frac{P(Y=y_j, X=x_i)}{P(X=x_i)} =$

$\overset{!}{=} \frac{P(Y=y_j) \cdot P(X=x_i)}{P(X=x_i)} = P(Y=y_j) = p_Y(y_j)$

$\Leftarrow$ Lat Annahme: $p_{Y|X}(y_j|x_i) = p_Y(y_j) \quad \forall x_i \in \mathcal{X}, \forall y_j \in \mathcal{Y}$

d.h. $\frac{p_{X,Y}(x_i, y_j)}{p_X(x_i)} = p_Y(y_j)$ — " —

d.h. $p_{X,Y}(x_i, y_j) = p_X(x_i) p_Y(y_j) \Rightarrow X, Y$ unabhängig nach Satz I.17

8.5.2025

Bsp II.6 3-facher Münzwurf, unabhängig

X ... Anzahl der Köpfe beim 1. Wurf

Y ... Gesamtanzahl der Köpfe bei allen 3 Würfen

$|\Omega| = 8$

X \ Y	0	1	2	3	
0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{4}{8}$
1	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{4}{8}$
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

$$P_{Y|X}(0|0) = \frac{P_{X,Y}(0,0)}{P_X(0)} = \frac{\frac{1}{8}}{\frac{4}{8}} = \frac{1}{4}$$

$$P_{Y|X}(0|1) = \frac{P_{X,Y}(1,0)}{P_X(1)} = \frac{0}{\frac{4}{8}} = 0$$

$P_{Y|X}(1|0) = \frac{1}{2}$ ✓, $P_{Y|X}(2|0) = \frac{1}{4}$ ✓ $\quad \frac{\frac{2}{8}}{\frac{4}{8}} = \frac{16}{32} = \frac{1}{2}$

$P_{Y|X}(1|1) = \frac{1}{3}$ $\quad P_{Y|X}(2|1) = \frac{2}{3}$ $\quad P_{Y|X}(3|0) = 0$ ✓

$\frac{\frac{1}{8}}{\frac{4}{8}} = \frac{8}{32} = \frac{1}{4}$

II.2 Stetiger Fall

Bsp: (i) $(X,Y)'$ ... X Körpergröße, Y Gewicht

(ii) X ... Einkommen eines Haushalts/Monat

Y ... Konsumausgaben -//-

(Randdichte von X = Einkommensverteilung)

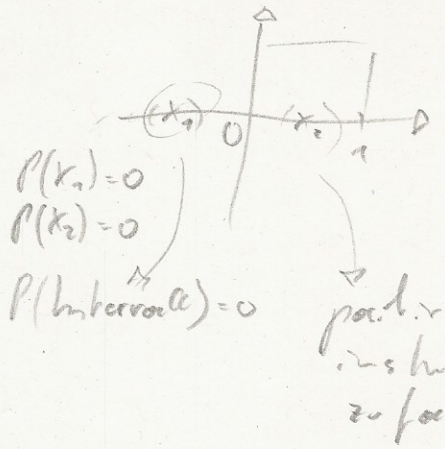
Definition II.7 Sei $(X,Y)'$ ein Zufallsvektor mit Dichte f_{XY}. Dann heißt

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

die bedingte Dichte von Y gegeben $X = x$ falls $f_X(x) > 0$

8.5.2015

X ... gleichverteilt $[0,1]$



Bemerkung II.8 Es g.Ü.:

(i) $f_{Y|X}(y|x) \geq 0 \quad \forall y, \forall x$ m.A. $f_X(x) > 0$

(ii) $\int_{-\infty}^{\infty} f_{Y|X}(y|x)\,dy = 1 \quad \forall x$ m.A. $f_X(x) > 0$

Beweis:

(i) trivial

(ii) $\int_{-\infty}^{\infty} f_{Y|X}(y|x)\,dy = \int_{-\infty}^{\infty} \frac{f_{XY}(x,y)}{f_X(x)}\,dy = \frac{1}{f_X(x)} \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \frac{1}{f_X(x)} f_X(x) = 1$

D.h. Für jedes x m.A. $f_X(x) > 0$
besitzt $f_{Y|X}(y|x)$ alle Eigenschaften einer Dichtefunktion

Erinnerung: Z ... Zufallsvariable m.A. Dichte g

$$P(Z \in A) = \int_A g(v)\,dv$$

$$F(z) = P(Z \leq z) = \int_{-\infty}^{z} g(v)\,dv = \int_A g(v)\,dv$$

$A = (-\infty, z]$

8.5.2015

Wir definieren:

$$P(Y \in A \mid X = x) := \int_A f_{Y|X}(y|x)\,dy \quad \text{sofern } f_X(x) > 0$$

(Für „alle" $A \subseteq \mathbb{R}$)

Setzt man $A = (-\infty, y]$ in obige Definition ein, erhält man:

$$F_{Y|X}(y|x) := P(Y \le y \mid X = x) = \int_{-\infty}^{y} f_{Y|X}(u|x)\,du \quad \text{sofern } f_X(x) > 0$$

$F_{Y|X}(y|x)$ heißt bedingte Verteilungsfunktion von Y gegeben $X = x$.

Satz II.9 Sei $(X,Y)'$ ein Zufallsvektor mit Dichte f_{XY}. Dann gilt:

X, Y unabhängig $\Leftrightarrow f_{Y|X}(y|x) = f_Y(y)$ für alle $y \in \mathbb{R}$ und alle x mit $f_X(x) > 0$

Beweis:

$\Rightarrow$ X, Y unabhängig impliziert nach Satz I.17, dass

$$f_{XY}(x,y) = f_X(x) f_Y(y) \quad \forall x \in \mathbb{R}, \forall y \in \mathbb{R}$$

Betrachte $x \in \mathbb{R}$ mit $f_X(x) > 0$. Dann folgt:

$$\frac{f_{XY}(x,y)}{f_X(x)} = f_Y(y) \quad \forall y \in \mathbb{R}, \forall x \text{ mit Dichte } f_X(x) > 0$$

d.h. $f_{Y|X}(y|x) = f_Y(y)$

8.5.2015

$\Leftarrow$ W.zeige $f_{Y|X}(y|x) = f_Y(y) \quad \forall y \in \mathbb{R}$, $\forall x$ m.A. D.dichte $f_X(x) > 0$

D.h. $\frac{f_{X,Y}(x,y)}{f_X(x)} = f_Y(y)$

d.h. $f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \forall y \in \mathbb{R}$, $\forall x$ m.A. D.dichte $f_X(x) > 0$

$(*)$

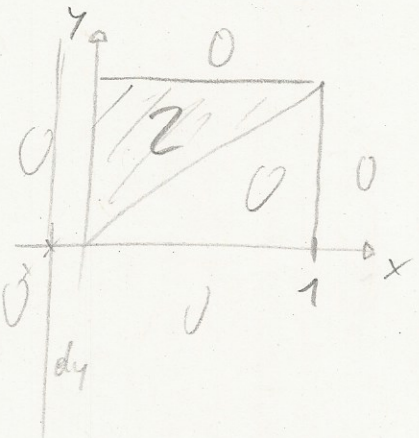
Falls es uns gelingt, $(*)$ für alle y und alle x m.A. $f_X(x) = 0$ nachzuweisen, dann g.d.h. $(*)$ für alle y und alle x und somit sind X, Y unabhängig nach Satz I.17

M.a.W.: Müssen zeigen, dass $f_X(x)$ impliziert $f_{X,Y}(x,y) = 0 \quad \forall y$

Beachte: $0 = f_X(x) = \int_{-\infty}^{\infty} \underbrace{f_{X,Y}(x,y)}_{\geq 0} dy \Rightarrow f_{X,Y}(x,y) = 0 \quad \forall y$

Bsp. II.10

$$f_{X,Y}(x,y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{sonst} \end{cases}$$



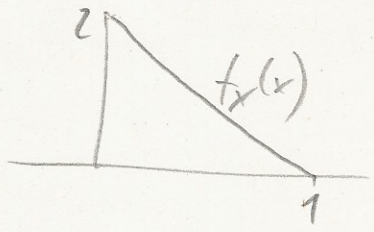
Bestimme $f_{Y|X}$:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \begin{cases} 0 \text{ falls } x \leq 0 \text{ oder } x \geq 1 \\ \int_{-\infty}^{x} \underbrace{f_{X,Y}(x,y)}_{=0} dy + \int_{x}^{1} \underbrace{f_{X,Y}(x,y)}_{=2} dy + \int_{1}^{\infty} \underbrace{f_{X,Y}(x,y)}_{=0} dy \end{cases} =$$

$$2\int_x^1 1\,dy = 2y\Big|_y^1 = 2(1-x)$$

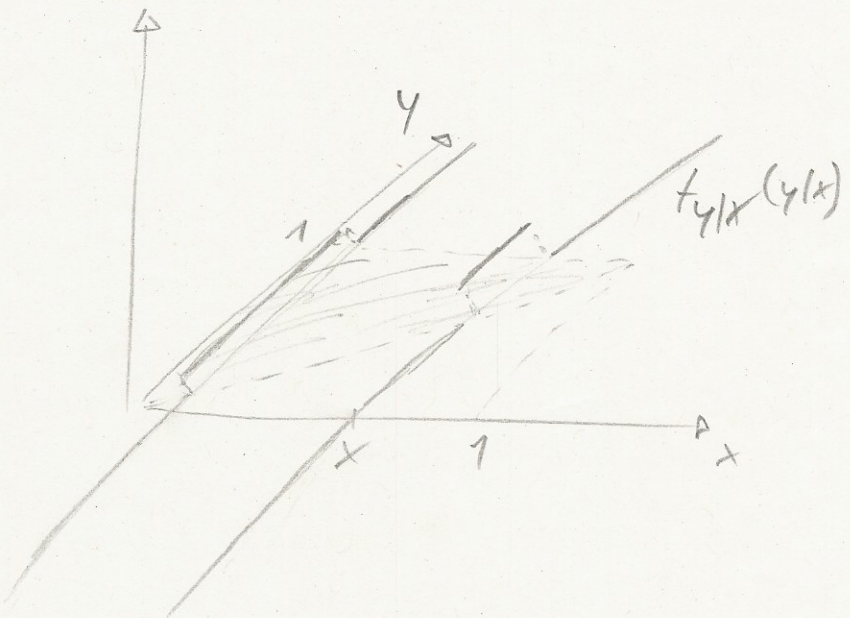
8.5.2015

$$= \begin{cases} 0 & \text{falls } x \le 0 \text{ oder } x \ge 1 \\ 2(1-x) & \text{falls } 0 < x < 1 \end{cases}$$



Für $0 < x < 1$:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} \frac{2}{2(1-x)} = \frac{1}{1-x} & \text{falls } 0 < x < y < 1 \\ 0 & \text{sonst} \end{cases}$$



! $f_{Y|X}$ ist eine Gleichverteilung auf dem Intervall $(x,1)$

Bsp II.11 Bivariate Normalverteilung

$(X,Y)' \sim N(\mu, \Sigma)$ wobei $\mu = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}$, $\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$

Gemeinsame Dichte:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho\frac{(x-\mu_X)}{\sigma_X}\frac{(y-\mu_Y)}{\sigma_Y}\right\}\right)$$

Wissen: $X \sim N(\mu_X, \sigma_X^2)$ $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\frac{(x-\mu_X)^2}{\sigma_X^2}\right)$ $f_X(x) > 0 \ \forall x \in \mathbb{R}$

$f_{Y|X}(y|x) = ?$ VO Stat. 1.12 9.5.2025

$$f_{Y|X} = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{1}{2(1-\rho^2)}\left\{v^2+v^2-2\rho uv\right\}+\frac{1}{2}u^2\right) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left\{v^2+v^2-2\rho uv-(1-\rho^2)u^2\right\}\right) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left\{\rho^2u^2+v^2-2\rho uv\right\}\right) =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\underbrace{(\rho u-v)^2}_{=(v-\rho u)^2}\right) =$$

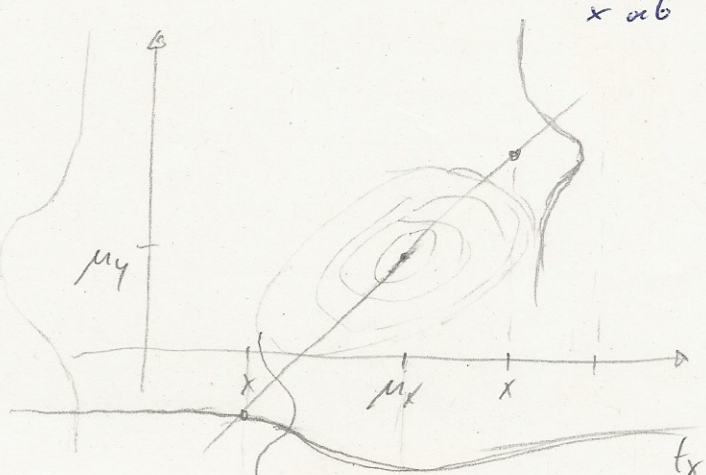
$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left(\frac{y-\mu_Y}{\sigma_Y}-\rho\frac{x-\mu_X}{\sigma_X}\right)^2\right) =$$

$$= \frac{1}{\underbrace{\sqrt{2\pi}(\sigma_Y\sqrt{1-\rho^2})}_{c}} \exp\left(-\frac{1}{\underbrace{2(1-\rho^2)\sigma_Y^2}_{c^2}}\left(y-\underbrace{\left(\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)}_{v(x)}\right)^2\right) =$$

$$= \frac{1}{\sqrt{2\pi}c} \exp\left(-\frac{1}{2c^2}(y-v(x))^2\right)$$

Bedingte Verteilung von $Y|X=x$ ist

$$N\Big(\underbrace{\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)}_{\text{hängt von } x \text{ ab}},\ \underbrace{\sigma_Y^2(1-\rho^2)}_{\text{hängt nicht von } x \text{ ab}}\Big)$$



II.3 Bedingte Erwartung

Definition II.12 Sei $(X,Y)'$ ein Zufallsvektor m.d. $E|Y|<\infty$. Dann ist der bedingte Erwartungswert $E(Y|X=x)$ wie folgt definiert:

(a) Ang. $(X,Y)'$ bes.d. Dichte $f_{X,Y}$. Dann $E(Y|X=x) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|x)\,dy$ sofern $f_X(x) > 0$

(b) $(X,Y)'$ diskret. Dann

$$E(Y|X=x) = \sum_{y_i \in \mathcal{Y}} y_i f_{Y|X}(y_i|x) \quad \forall x \in \mathcal{X}$$

Bemerkung:

(i) $E(Y|X=x)$ ist eine Funktion von x, d.h. $E(Y|X=x) = h(x)$

(ii) Sei $(X,Y)'$ ein Zufallsvektor wie in Definition II.12. Wir betrachten $g(X,Y)$, wobei $g(X,Y)$ eine reellwertige Funktion ist.
Setze $Z = g(X,Y)$. Dann ist $E(g(X,Y)|X=x) = E(Z|X=x)$, d.h. Definition II.12 angewendet auf das Paar (X,Z) (vorausgesetzt $(X,Z)'$ bes.d. Dichte oder ist diskret)

Satz II.13 (Ohne Beweis) Sei $(X,Y)'$ wie in Definition II.12. Dann g.d.:

$$E(a(x)\cdot Y + b(x)|X=x) = a(x)E(Y|X=x) + b(x)$$

wobei a, b reellwertige Funktionen von x sind

Def. II.14 Sei $(X,Y)'$ ein Zufallsvektor mit $E(Y^2)<\infty$. Weiters bestehe $(X,Y)'$ entweder aus diskreten oder aus stetigen (wobei X, Y sowie in Satz II.7)

Dann ist die bedingte Varianz von Y gegeben $X=x$ definiert als:

$$\mathrm{Var}(Y|X=x) = E\left((Y-E(Y|X=x))^2 \,|\, X=x\right)$$

Warnung: $\mathrm{Var}(Y|X=x) \neq E\left((Y-E(Y))^2|X=x\right)$ i.A.

Satz II.15 Sei $(X,Y)'$ wie in Def II.14. Dann g.Ü:

~~Dann g.Ü:~~

$$\mathrm{Var}(Y|X=x) = E(Y^2|X=x) - \left(E(Y|X=x)\right)^2$$

Beweis: $\mathrm{Var}(Y|X=x) = E\left((Y-E(Y|X=x))^2|X=x\right) =$

$$= E\Big[Y^2 - 2Y\underbrace{E(Y|X=x)}_{=g(x)} + \underbrace{\left(E(Y|X=x)\right)^2}_{=g(x)}|X=x\Big] =$$

$$= E(Y^2|X=x) - 2E(Y|X=x)E(Y|X=x) + \left(E(Y|X=x)\right)^2 =$$

$$= E(Y^2|X=x) - 2\left(E(Y|X=x)\right)^2 + \left(E(Y|X=x)\right)^2 =$$

$$= E(Y^2|X=x) - \left(E(Y|X=x)\right)^2$$

Bsp II.16

$$f_{XY}(x,y) = \begin{cases} 2 & \text{für } 0<x<y<1 \\ 0 & \text{sonst} \end{cases}$$

Für $0<x<1$

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

$$f_X(x) = \int_x^1 2\,dy = 2y\Big|_x^1 = 2-2x = 2(1-x)$$

$$E(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy =$$

$$= \underbrace{\int_{-\infty}^{x} y f_{Y|X}(y|x)\,dy}_{=0} + \int_x^1 \dots + \underbrace{\int_1^{\infty} \dots}_{=0} =$$

$$= \int_x^1 y \cdot \frac{1}{1-x}\,dy = \frac{1}{1-x}\int_x^1 y\,dy =$$

$$= \frac{1}{1-x}\,\frac{y^2}{2}\Big|_x^1 = \frac{1}{1-x}\,\frac{1}{2}(1-x^2) = \frac{1+x}{2}$$

$\left(\frac{1}{2} - \frac{1}{2}x^2\right) = \frac{1}{2}(1-x^2)$

$(1-x)(1+x)$

$$E(Y^2|X=x) = \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x)\,dy = \int_x^1 y^2 \frac{1}{1-x}\,dy = \frac{1}{1-x}\,\frac{y^3}{3}\Big|_x^1 = \frac{1}{(1-x)}\,\frac{1}{3}(1-x^3) =$$

$a^3 - b^3 = (a-b)(a^2+ab+b^2)$

$$= \frac{1}{3}(1+x+x^2)$$

$$\mathrm{Var}(Y|X=x) = \frac{1}{3}(1+x+x^2) - \left(\frac{1-x}{2}\right)^2$$

← hängt von x ab

Bsp II 17

(X,Y) bivariat normalverteilt

$E(Y|X=x) = \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)$ (Lineare Funktion von x)

$\mathrm{Var}(Y|X=x) = \sigma_Y^2(1-\rho^2)$ (hängt nicht von x ab)

$$N\left(\mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X),\ \sigma_Y^2(1-\rho^2)\right)$$

Bisher: $E(Y|X=x)$ für festes x

Beachte: $E(Y|X=x) = h(x)$

Oft betrachtet man $h(X)$

↳ Zufallsvariable

Analog: $\mathrm{Var}(Y|X=x) = g(x)$

Für $g(X)$ schreibt man $\mathrm{Var}(Y|X)$

Satz II.18 (Glättungseigenschaft) N.B.

Sei $(X,Y)'$ ein Zufallsvektor mit $E|Y| < \infty$. Dann gilt:

$$E(E(Y|X)) = E(Y)$$

Beweis: Fall, dass $(X,Y)'$ D. mit dichte bes. Zsh:

$E(E(Y|X)) = E(h(X))$ wobei $h(x) = E(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy$

Also: $E(E(Y|X)) = E(h(X)) = \int_{-\infty}^{\infty} h(x) f_X(x)\,dx = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy \right] f_X(x)\,dx =$

$P(A|B)\cdot P(B) = P(A \cap B)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{Y|X}(y|x)\cdot f_X(x)\,dy\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x,y)\,dy\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x,y)\,dx\,dy =$$

$$= \int_{-\infty}^{\infty} y \underbrace{\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx}_{\text{= Randdichte von } Y}\,dy = \int_{-\infty}^{\infty} y\, f_Y(y)\,dy = E(Y)$$

$f_Y(y)$

9.6.15

Bemerkung II.19

$$E(\mathrm{Var}(Y|X)) \neq \mathrm{Var}(Y) \quad \text{i.A.}$$

Beachte: $E[E[(Y-E(Y|X))^2|X]]$

$+$

$E((Y-E(Y))^2) = E[E((Y-E(Y))^2|X)]$

↑ Glättungseigenschaft

Satz II.20 Sei $(X,Y)'$ ein Zufallsvektor mit $E|Y|<\infty$

Dann gilt:

$$E(\alpha(X)+\beta(X)Y|X) = \alpha(X)+\beta(X)E(Y|X)$$

Beweis: W.ssen: $E(\alpha(x)+\beta(x)Y|X=x)$ $g(x) = \alpha(x)+\beta(x)\underbrace{E(Y|X=x)}_{=h(x)} =$

$= \alpha(x)+\beta(x)h(x)$ (Satz II 13)

Nach Def II 18

$E(\alpha(X)+\beta(X)Y|X) = g(X) = \alpha(X)+\beta(X)\underbrace{h(X)}_{=E(Y|X)} =$

$= \alpha(X)+\beta(X)E(Y|X)$

15.6.2015

VO Stochastik 2

III Momentenmethode und Maximum Likelihood

III.) Momentenmethode

Bsp. III.1: $Z_i \sim N(\mu, 1)$ unabhängig, $i = 1, \dots, n$

Beachte: Parameter $\mu = E(Z_i)$ = Moment

ersetze durch Stichprobenmittel

Schätzer: $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} Z_i$

Bsp III.2: $Z_i \sim N(\mu, \sigma^2)$ unabhängig

i.i.d $N(\mu, \sigma^2)$

Beachte:
$$\mu = E(Z_i)$$
$$\sigma^2 = E\left(\left(Z_i - E(Z_i)\right)^2\right) = E(Z_i^2) - \underbrace{\left(E(Z_i)\right)^2}_{=\mu}$$

Äquivalent:
$$\mu = E(Z_i) \quad \text{- erstes Moment}$$
$$\sigma^2 + \mu^2 = E(Z_i^2) \quad \text{- zweites Moment}$$

Stichprobenanalogon:

$$\left.\begin{aligned} \hat{\mu} &= \frac{1}{n} \sum_{i=1}^{n} Z_i = \overline{Z} \\ \hat{\sigma}^2 + \hat{\mu}^2 &\overset{!}{=} \frac{1}{n} \sum_{i=1}^{n} Z_i^2 = \overline{Z^2} \end{aligned}\right\} \Rightarrow \hat{\mu} = \overline{Z}$$

$$\sigma^2 = \overline{Z^2} - \hat{\mu}^2 = \overline{Z^2} - \left(\overline{Z}\right)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left(Z_i - \overline{Z}\right)$$

Bsp: III. 3 Z_i i.i.d $N(\mu, \sigma^2)$

Beachte: $E(Z_i^3) = \mu^3 + 3\mu\sigma^2$

$E(Z_i^2) = \sigma^2 + \mu^2$

Betrachte: $\mu^3 + 3\mu\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i^3$

$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i^2 = \overline{Z^2}$

NR: $E(Z_i^3) = E\left((Z_i - \mu + \mu)^3\right) = E(Z_i - \mu)^3$ ($=0$)

$+ 3E\left(Z_i - \mu\right)^2 \cdot \mu + 3E(Z_i - \mu)\mu^2$ ($=\sigma^2$, $=0$)

$+ \mu^3$

$E(Z_i - \mu)^3 = \int_{-\infty}^{\infty} (z-\mu)^3 \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(z-\mu)^2\right)$ ($=0$)

z

! Substitution: $\hat{\mu}^3 + 3\hat{\mu}\left(\overline{Z^2} - \hat{\mu}^2\right) = \overline{Z^3}$

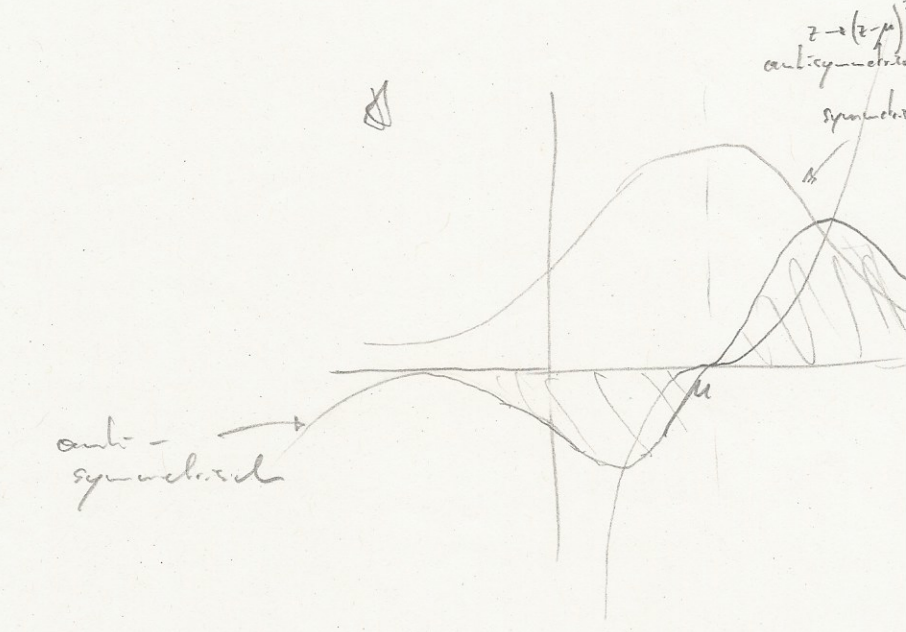
(Üb: Gleichung 3. Grades besitzt eine reelle Lösung)

Gleichung 3. Grades für $\hat{\mu}$ → $\hat{\mu}$

Dann: $\hat{\sigma}^2 = \overline{Z^2} - \hat{\mu}^2$

Gleichsetzen des theoretischen und empirischen Moments $\bar{s}$-ter Ordnung:

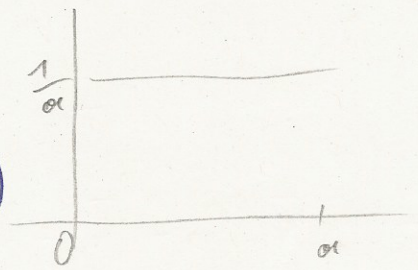
$E(X^{\bar{s}}) = \frac{1}{n}\sum_{i=1}^{n} Z_i^{\bar{s}}$



Bsp III Q: Z_i i.i.d $U([0, \alpha])$, $\alpha > 0$

Möglichkeit 1: $E(Z_i) = \int_0^{\alpha} z \cdot \frac{1}{\alpha} dz = \frac{1}{\alpha}\frac{z^2}{2}\Big|_0^{\alpha} = \frac{1}{\alpha}\left(\frac{\alpha^2}{2} - 0\right)$

$= \frac{\alpha}{2}$



d.h. $\frac{\alpha}{2} = E(Z_i)$

$\Rightarrow \frac{\hat{\alpha}}{2} = \frac{1}{n}\sum_{i=1}^{n} Z_i \quad = \hat{\alpha} = 2 \cdot \overline{Z}$

Möglichkeit 2: $E(Z_i^2) = \int_0^a z^2 \cdot \frac{1}{a}\, dz = \frac{1}{a}\frac{z^3}{3}\Big|_0^a = \frac{1}{a}\frac{a^3}{3} = \frac{a^2}{3}$

$$\frac{a^2}{3} = E(Z_i^2)$$

$$\Rightarrow \frac{\hat{a}^2}{3} = \frac{1}{n}\sum_{i=1}^{n} Z_i^2$$

$$\hat{a} = \sqrt{3 \cdot \bar{Z}^2}$$

? W.i.e ist das, wenn man das p-te Moment verwendet?

Bsp. III. 5: $Z_i = \varepsilon_i + b\varepsilon_{i-1}$, $b \in \mathbb{R}$ unbekannt, $E\varepsilon_i = 0$, $E\varepsilon_i^2 = \sigma^2$ unbekannt

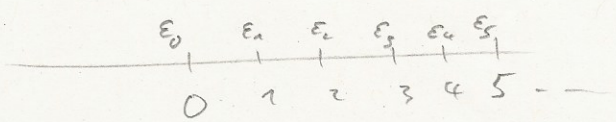
$E\varepsilon_i \varepsilon_j = 0$ für $i \neq j$

Daten $z_1, \dots, z_n$

(Moving average Modell der Ordnung 1)

$$E(Z_i) = E(\varepsilon_i + b\varepsilon_{i-1}) =$$

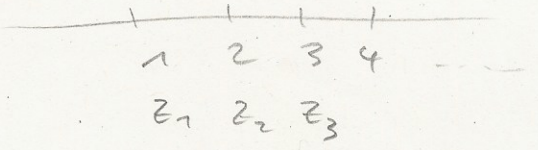
$$= E(\varepsilon_i) + bE(\varepsilon_{i-1}) = 0$$



$$Var(Z_i) = E(Z_i^2) = E\left((\varepsilon_i + b\varepsilon_{i-1})^2\right) =$$

$$= E\left(\varepsilon_i^2 + 2b\varepsilon_i\varepsilon_{i-1} + b^2\varepsilon_{i-1}^2\right) = \sigma^2 + 2b\underbrace{E(\varepsilon_i\varepsilon_{i-1})}_{=0} + b^2\sigma^2$$

$$= \sigma^2(1 + b^2)$$



$$Cov(Z_i, Z_{i-1}) = E(Z_i Z_{i-1}) = E\left((\varepsilon_i + b\varepsilon_{i-1})(\varepsilon_{i-1} + b\varepsilon_{i-2})\right) =$$

$$= E(\varepsilon_i\varepsilon_{i-1} + b\varepsilon_i\varepsilon_{i-2} + b\varepsilon_{i-1}^2 + b^2\varepsilon_{i-1}\varepsilon_{i-2}) = \underbrace{E(\varepsilon_i\varepsilon_{i-1})}_{=0} + b\underbrace{E(\varepsilon_i\varepsilon_{i-2})}_{=0} +$$

$$= \sigma^2 b \neq 0 \text{ !} \quad \text{Daher } Z_i \text{ sind korreliert!} \quad bE(\varepsilon_{i-1}^2) + b^2\underbrace{E(\varepsilon_{i-1}\varepsilon_{i-2})}_{=0}$$

σ^2

$$\mathrm{Cov}(z_i, z_{i-2}) = E(z_i z_{i-2}) = E\big((\varepsilon_i + b\varepsilon_{i-1})(\varepsilon_{i-2} + b\varepsilon_{i-3})\big) = 0$$

$$\mathrm{Cov}(z_i, z_{i-3}) = 0$$

$$\mathrm{Cov}(z_i, z_{i-n}) = 0 \text{ etc.}$$

Unbekannte Parameter: σ^2, b

$$\sigma^2(1+b^2) = E(z_i^2)$$

$$\sigma^2 b = E(z_i z_{i-1})$$

1. $\hat{\sigma}^2(1+\hat{b}^2) = \frac{1}{n}\sum_{i=1}^{n} z_i^2$

2. $\hat{\sigma}^2 \hat{b} = \frac{1}{n}\sum_{i=2}^{n} z_i z_{i-1}$

alternativ: $\frac{1}{n-1}$

$$\overset{2.}{\underset{1.}{:}} \frac{\hat{b}}{1+\hat{b}^2} = \frac{\frac{1}{n}\sum_{i=2}^{n} z_i z_{i-1}}{\frac{1}{n}\sum_{i=1}^{n} z_i^2} = A$$

$$\Rightarrow A\hat{b}^2 + A - \hat{b} = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{\hat{b}} \cdot \frac{1}{n}\sum_{i=2}^{n} z_i z_{i-1}$$

Momentenmethode (für i.i.d. Fall)

(dürfen $q \times 1$ Zufallsvektoren sein)

$z_1, \dots, z_n$ i.i.d. nach einer Verteilung, welche von einem Parameter (vektor) θ, $\theta \in \Theta \subseteq \mathbb{R}^p$, abhängt.

Gegeben ist eine Funktion $g: \mathbb{R}^q \to \mathbb{R}^p$, sodass $Eg(z_i)$ existiert.

Setze $h(\theta) = Eg(z_i)$,

(Beachte: h ist eine Funktion von $\Theta \subseteq \mathbb{R}^p$ nach $\mathbb{R}^p$)

dann heißt ein Schätzer $\hat{\theta}$ Momenten-schätzer falls er erfüllt:

$$h(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} g(z_i)$$

(p Gleichungen in p Unbekannten)

$N(\mu, \sigma^2)$

$\theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$, $\Theta = \mathbb{R} \times \mathbb{R}_+ \subseteq \mathbb{R}^2$

$p = 2$

$$g(z) = \begin{pmatrix} z \\ z^2 \end{pmatrix} = h(\mu, \sigma^2)$$

$$Eg(z_i) = E\begin{pmatrix} z_i \\ z_i^2 \end{pmatrix} = \begin{pmatrix} Ez_i \\ Ez_i^2 \end{pmatrix} = \begin{pmatrix} \mu \\ \mu^2 + \sigma^2 \end{pmatrix}$$

$\mu = E(z_i)$

$\sigma^2 + \mu^2 = E(z_i^2)$

Bsp: Z_i i.i.d. $N(\mu, \sigma^2)$

Bsp III.2

$\hat{\mu} = \bar{Z}$

$\hat{\mu}^2 + \hat{\sigma}^2 = \overline{Z^2}$

1+2. Moment

$p = 2$

2 unbekannte Parameter 2 Gleichungen

Bsp. III 3:

$\hat{\mu}^3 + 3\hat{\mu}\hat{\sigma}^2 = \overline{Z^3}$

$\hat{\sigma}^2 + \hat{\mu}^2 = \overline{Z^2}$

2+3. Moment

$p = 2$

Wissen:

3 Glg. in 2 Unbekannte

$\mu = E(Z_i)$

$\mu^2 + \sigma^2 = E(Z_i^2)$

$\mu^3 + 3\mu\sigma^2 = E(Z_i^3)$

$g(z) = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$

Man könnte versuchen, Schätzer $\hat{\mu}, \tilde{\sigma}^2$ wie folgt zu bestimmen:

$\hat{\mu} = \bar{Z}$

$\hat{\mu}^2 + \hat{\sigma}^2 = \overline{Z^2}$

$\hat{\mu}^3 + 3\hat{\mu}\hat{\sigma}^2 = \overline{Z^3}$

} Diese 3 Glg lassen sich i.A. nicht lösen

Rückkehr zu Bsp III.2:

$\mu = E(Z_i)$

$\sigma^2 + \mu^2 = E(Z_i^2)$

$\hat{\mu} = \bar{Z}$

$\hat{\sigma}^2 + \hat{\mu}^2 = \overline{Z^2}$

$\Leftrightarrow \begin{cases} \hat{\mu} - \bar{Z} = 0 \\ \hat{\sigma}^2 + \hat{\mu}^2 - \overline{Z^2} = 0 \end{cases} \Leftrightarrow (\underbrace{\hat{\mu} - \bar{Z}}_{\geq 0})^2 + (\underbrace{\hat{\sigma}^2 + \hat{\mu}^2 - \overline{Z^2}}_{\geq 0})$

$\Leftrightarrow \hat{\mu}$ und $\hat{\sigma}^2$ lösen das Minimierungsproblem:

$(\hat{\mu} - \bar{Z})^2 + (\hat{\sigma}^2 + \hat{\mu}^2 + \overline{Z^2})^2 \to \min$

Gehe zu: $\mu = E(Z_i) = 0$, $\sigma^2 + \mu^2 = E(Z_i^2) = 0$, $\mu^3 + 3\mu\sigma^2 = E(Z_i^3) = 0$

bzw. $\hat{\mu} = \bar{Z} = 0$, $\hat{\sigma}^2 + \hat{\mu}^2 = \overline{Z^2} = 0$, $\hat{\mu}^3 + 3\hat{\mu}\hat{\sigma}^2 = \overline{Z^3} = 0$

: A. nicht lösbar

Kann aber $\hat{\mu}$ und $\hat{\sigma}^2$ definieren als Lösung des Minimierungsproblems:

$$\left(\hat{\mu} - \bar{z}\right)^2 + \left(\hat{\sigma}^2 + \hat{\mu}^2 - \overline{z^2}\right)^2 + \left(\hat{\mu}^3 + 3\hat{\mu}\hat{\sigma}^2 - \overline{z^3}\right)^2 \to \min$$

Dieses $\hat{\mu}, \hat{\sigma}^2$ heißt: Generalized method of moments „Schätzer" („GMM-Schätzer")

(inneres Produkt)

Allgemein:

$$\begin{pmatrix} \hat{\mu} - \bar{z} \\ \hat{\sigma}^2 + \hat{\mu}^2 - \overline{z^2} \\ \hat{\mu}^3 + 3\hat{\mu}\hat{\sigma}^2 - \overline{z^3} \end{pmatrix}' W \left(\hat{\mu} - \bar{z}, \hat{\sigma}^2 + \hat{\mu}^2 - \overline{z^2}, \hat{\mu}^3 + 3\hat{\mu}\hat{\sigma}^2 - \overline{z^3}\right)' \to \min$$

wobei W eine pos. def. Gewichtsmatrix

III.2 Maximum Likelihood

Bsp. III.8 z_i i.i.d. Poisson mit Parameter λ, $\lambda > 0$ (sichere Ereignisse, die unabhängig auftreten)

D.h.: z_i nimmt Werte $0, 1, 2, 3, \ldots$ mit Whs

$$P(z_i = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$ (Man kann zeigen: $E(z) = \lambda$)

$z_1, z_2, \ldots, z_n$ (d.h. $z_1, \ldots, z_n$ Realisierungen von $z_1, \ldots, z_n$) „Gedankenstart"

$$P(z_1 = z_1, z_2 = z_2, \ldots, z_n = z_n) = P(z_1 = z_1) \cdot P(z_2 = z_2) \cdot \ldots \cdot P(z_n = z_n) =$$

$$= e^{-\lambda} \frac{\lambda^{z_1}}{z_1!} \cdot e^{-\lambda} \frac{\lambda^{z_2}}{z_2!} \cdots e^{-\lambda} \frac{\lambda^{z_n}}{z_n!} = e^{-n\lambda} \lambda^{z_1 + \ldots + z_n} \cdot \frac{1}{z_1! z_2! \ldots z_n!} = L_n(\lambda) =$$

$$= L_n(\lambda; z_1, \ldots, z_n)$$

(λ: unbekannt; $z_1, \ldots, z_n$: bekannt)

L_n — „Likelihood-Funktion der Daten $z_1, \ldots, z_n$"

Schätze λ durch $\hat{\lambda}$, welches $L_n(\lambda)$ maximiert

$L_n(\lambda_2) = 0{,}53$

λ so, dass die Whs, dass die Daten auftreten, groß ist

$\max_{\lambda>0} L_n(\lambda)$ Äquivalent: $\max_{\lambda>0} \underbrace{\log L_n(\lambda)}_{l_n(\lambda)}$

$$l_n(\lambda) = -n\lambda + (z_1 + \dots + z_n)\log\lambda - \log(z_1! \dots z_n!)$$

$$\frac{\partial l_n(\lambda)}{\partial \lambda} = -n + (z_1 + \dots + z_n)\frac{1}{\lambda}$$

(First order condition):

FOC

$$-n + (z_1 + \dots + z_n)\frac{1}{\lambda} = 0$$

bzw. $\hat{\lambda}_{ML} = \frac{1}{n}\sum_{i=1}^{n} z_i$

SOC: $-\frac{(z_1 + \dots z_n)}{\hat{\lambda}^2} < 0$ d.h. Max

Beachte: $\hat{\lambda}_{ML}$ ist ein Beispiel für einen Maximum-Schätzer

16.6.2015

VO Statistik 2

z_i i.i.d. Poisson-verteilt mit $\lambda > 0$

Daher $z_1, \dots, z_n$

$$P(z_1 = z_1, \dots, z_n = z_n) \overset{\text{Unabhängigkeit}}{=} P(z_1 = z_1) \cdots P(z_n = z_n) = e^{-\lambda} \frac{\lambda^{z_1}}{z_1!} \cdots e^{-\lambda} \frac{\lambda^{z_n}}{z_n!} =$$

$$= e^{-n\lambda} \frac{\lambda^{z_1 + \dots + z_n}}{z_1! \cdots z_n!} = L_n(\lambda) = L_n(\lambda; z_1, \dots, z_n)$$

$L_n(\lambda)$ = Likelihood Funktion der Daten $z_1, \dots, z_n$

Bestimme $\hat{\lambda}$ so, dass $L_n(\hat{\lambda}) = \max_{\lambda > 0} L_n(\lambda)$

$\hat{\lambda}_{ML}$ ... Maximum Likelihood Schätzer

$$\dots \hat{\lambda}_{ML} = \frac{1}{n} \sum_{i=1}^{n} z_i$$

Definition III. 9) Seien $z_1, \dots, z_n$ i.i.d. absolutstetige Zufallsvektoren (Zufallsvariablen), wobei die W.-Funktion von z_i gegeben ist durch $p(z, \theta)$

Hier ist θ ein Parameter(vektor) mit $\theta \in \Theta$

Gegeben sind Daten, d.h. Beobachtungen $z_1, \dots, z_n$ von $z_1, \dots, z_n$. Dann heißt

$$L_n(\theta) = L_n(\theta; z_1, \dots, z_n) \overset{\text{i.i.d.}}{=} \prod_{i=1}^{n} p(z_i, \theta)$$

die Likelihoodfunktion der Daten $z_1, \dots, z_n$

Ein Schätzer $\hat{\theta}$, welcher erfüllt

$$L_n(\hat{\theta}) = \max_{\theta \in \Theta} L_n(\theta),$$

heißt Maximum Likelihood Schätzer

(Schätzer: $\hat{\theta}_{ML}$)

Bemerkung III. 10)

(i) $L_n(\theta)$ gemeinsame W.-Funktion von $z_1, \dots, z_n$ ausgewertet bei den Beobachtungen $z_1, \dots, z_n$ ~~und~~ betrachtet als Funktion von θ

(ii) Ist $L_n(\theta) > 0$, dann heißt $\ell_n(\theta) = \log L_n(\theta)$ die Log-Likelihood-Funktion

Bsp 3.11 z_i i.i.d. Bernoulli-verteilt mit Parameter $p \in [0,1]$ (z_i nimmt die Werte 0 bzw. 1 mit Whs $1-p$ bzw. p an)

Daten: $z_1, \dots, z_n$ (entweder = 0 oder 1)

$$L_n(p) = \prod_{i=1}^n p(z_i = z_i) = \prod_{i=1}^n p^{z_i}(1-p)^{1-z_i}$$

$$= \begin{cases} p \text{ falls } z_i = 1 \\ 1-p \text{ falls } z_i = 0 \end{cases}$$

$$= p^{z_i}(1-p)^{1-z_i}$$

Fall 1: $z_1 = \dots = z_n = 0$:

$$L_n(p) = \prod_{i=1}^n (1-p) = (1-p)^n \Rightarrow \hat{p}_{ML} = 0$$

Fall 2: $z_1 = \dots z_n = 1$

$$L_n(p) = p^n \Rightarrow \hat{p}_{ML} = 1$$

Fall 3: es g.Ü. mindestens ein $z_i = 1$ und ein $z_j = 0$

$$\left(\Leftrightarrow 0 < \frac{1}{n}\sum_{i=1}^n z_i < 1\right)$$

$$L_n(p) = \prod_{i=1}^n p^{z_i}(1-p)^{1-z_i}$$

Beachte: $L_n(0) = 0$ da für jenes i mit $z_i = 1$ $p^{z_i} = 0$ für $p = 0$

$L_n(1) = 0$ da für jenes j mit $z_j = 0$ $(1-p)^{z_j}$ für $p = 1$

Für $0 < p < 1$: $L_n(p) > 0$

Also folgt:

$$\hat{p}_{ML} \in (0,1)$$

Betrachten nur mehr $p \in (0,1)$. Da dann $L_n(p) > 0$ $\rightarrow$ Log-Likelihood-Funktion

$$l_n(p) = \sum_{i=1}^{n} \left[z_i \log p + ((1-z_i)(\log(1-p)) \right]$$

$$\frac{dl_n}{dp} = \sum_{i=1}^{n} \left(\frac{z_i}{p} - \frac{1-z_i}{1-p} \right) = \frac{\sum_{i=1}^{n} z_i}{p} - \frac{n - \sum_{i=1}^{n} z_i}{1-p}$$

FOC: $$\frac{\sum z_i}{\hat{p}_{ML}} = \frac{n - \sum z_i}{1-\hat{p}_{ML}}$$

$$(1-\hat{p}_{ML}) \sum z_i = n\hat{p}_{ML} - \hat{p}_{ML} \sum z_i$$

$$\sum z_i - \hat{p}_{ML} \sum z_i = n\hat{p}_{ML} - \hat{p}_{ML} \sum z_i$$

$$\hat{p}_{ML} = \frac{1}{n} \sum_{i=1}^{n} z_i$$ = relative Häufigkeit der "Erfolge"

SOC:

$$-\frac{\sum z_i}{p^2} - \frac{n - \sum z_i}{(1-p)^2} = -\left(\underbrace{\frac{\sum z_i}{\hat{p}_{ML}}}_{>0} + \underbrace{\frac{n - \sum z_i}{(1-\hat{p}_{ML})^2}}_{>0} \right) < 0 \Rightarrow \text{Max}$$

Bem: $E z_i = 1 \cdot p + 0 \cdot (1-p) = p$

Momentenschätzer:

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} z_i$$

$z_1, \dots, z_n$ i.i.d gemäß Dichte $f(z,\theta)$, $\theta \in \Theta$ (z.B. $z_1, \dots, z_n$ i.i.d $N(\mu, \sigma^2)$, $\theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}$, $\Theta = \mathbb{R} \times (0,\infty)$)

Daher: $z_1, \dots, z_n$

$P(z_1 = z_1, \dots, z_n = z_n) = 0$ für jeden Wert von θ !!

Sei $\varepsilon > 0$

Betrachten:

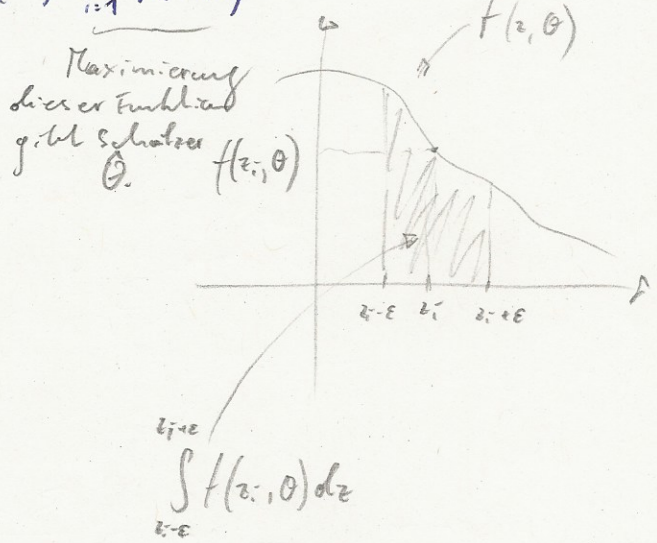
$$P(z_1 \in (z_1-\varepsilon, z_1+\varepsilon), z_2 \in (z_2-\varepsilon, z_2+\varepsilon), \dots, z_n \in (z_n-\varepsilon, z_n+\varepsilon)) =$$

$$= \prod_{i=1}^n P(z_i \in (z_i-\varepsilon, z_i+\varepsilon)) = \prod_{i=1}^n \int_{z_i-\varepsilon}^{z_i+\varepsilon} f(z,\theta)\,dz \approx$$

(ähnlich, falsch ;-)

$$\approx \prod_{i=1}^n 2\varepsilon f(z_i,\theta) = (2\varepsilon)^n \prod_{i=1}^n f(z_i,\theta)$$

Maximierung dieser Funktion gibt Schätzer $\hat{\theta}$.



$f(z)$

$P(z \in C) = \int_C f(z)\,dz$

$P(z \in (a,b)) = \int_a^b f(z)\,dz$

Definition III.12: Seien $z_1, \dots, z_n$ Zufallsvektoren (oder -variable) gegeben seien Beobachtungen $z_1, z_2, \dots, z_n$ von $z_1, \dots, z_n$ welche i.i.d sind mit Dichte $f(z,\theta)$, (Dichte der Dichte eines Vektors)

Dann heißt:

$$L_n(\theta) = L_n(\theta; z_1, z_2, \dots, z_n) \overset{\text{i.i.d}}{=} \prod_{i=1}^n f(z_i,\theta)$$

wobei $\theta \in \Theta$ ein Parameter ist.

die Likelihood-Funktion über Daten $z_1, \dots, z_n$

Ein Schätzer, welcher erfüllt:

$$L_n(\hat{\theta}) = \max_{\theta \in \Theta} L_n(\theta)$$

heißt Maximum-Likelihood-Schätzer ($\hat{\theta}_{ML}$)

Bemerkung III 13.

(i) $L_n(\theta)$ gemeinsame Dichte von $z_1, \dots, z_n$ evaluiert an den Daten $z_1, \dots, z_n$ betrachtet als Funktion von θ.

(ii) Ist $L_n(\theta) > 0$, dann heißt $l_n(\theta) = \log L_n(\theta)$ die Log-Likelihood Funktion der Daten $z_1, \dots, z_n$

Bsp. III 13': $z_i \dots \sim N(\mu, \sigma^2)$ (σ^2: „Einheit" Repräsentant für „Varianz") $\left(\theta = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \Theta = \mathbb{R} \times (0, \infty)\right)$

$$L_n(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(z_i - \mu)^2\right) > 0$$

$$l_n(\mu, \sigma^2) = \sum_{i=1}^{n} \left(-\frac{1}{2}\log 2\pi - \frac{1}{2}\log \sigma^2 - \frac{1}{2\sigma^2}(z_i - \mu)^2\right)$$

$$= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(z_i - \mu)^2$$

$$\frac{\partial l_n}{\partial (\sigma^2)} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2}\frac{1}{\sigma^4}\sum_{i=1}^{n}(z_i - \mu)^2$$

$$\frac{\partial l_n}{\partial \mu} = +\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(z_i - \mu)(+1) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(z_i - \mu)$$

FOC: $-\frac{n}{2}\frac{1}{\hat\sigma^2_{ML}} + \frac{1}{2}\frac{1}{\hat\sigma^4_{ML}}\sum_{i=1}^n (z_i-\hat\mu_{ML})^2 = 0 \;\Rightarrow\; \hat\sigma^2_{ML} = \frac{1}{n}\sum_{i=1}^n (z_i-\hat\mu_{ML})^2 = \frac{1}{n}\sum_{i=1}^n (z_i-\bar z)^2$

$\frac{1}{\hat\sigma^2_{ML}}\sum_{i=1}^n (z_i-\hat\mu_{ML}) = 0 \quad \Rightarrow \sum_{i=1}^n (z_i-\hat\mu_{ML}) = 0 \Rightarrow \hat\mu_{ML} = \frac{1}{n}\sum_{i=1}^n z_i = \bar z$

Bem.: Wissen: $E\hat\mu_{ML} = \mu$

$E\hat\sigma^2_{ML} = \frac{n-1}{n}\sigma^2 < \sigma^2$

$E\left(\hat\sigma^2_{ML}\right) = E\left(\frac{1}{n}\sum_{i=1}^n (z_i-\bar z)^2\right)$

Also: ML-Schätzer können verzerrt sein.

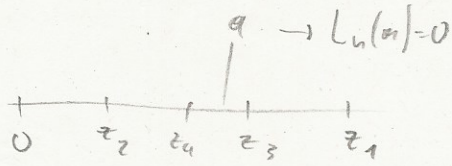
SOC: Hesse Matrix negativ definit

$$\begin{pmatrix} \frac{\partial^2 L_n}{(\partial(\sigma^2))^2} & \frac{\partial^2 L_n}{\partial\sigma\,\partial\mu} \\ & \frac{\partial^2 L_n}{(\partial\mu)^2} \end{pmatrix}$$

Bsp III.14: z_i i.i.d. $U([0,\alpha])$, $\alpha > 0$ $\quad(\theta = \alpha;\ \Theta \in (0,\infty))$

$L_n(\alpha) = \prod_{i=1}^n \frac{1}{\alpha}\mathbb{1}_{[0,\alpha]}(z_i)$ — Indikatorfunktion

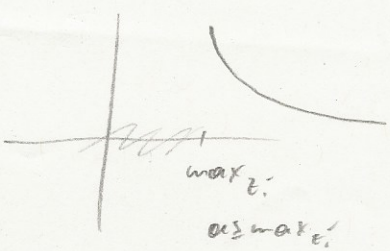
Daten $z_1, \dots, z_n$



Beachte: Ist $\alpha < \max_{i=1,\dots,n} z_i$, dann

$\Rightarrow L_n(\alpha) = 0$

Ist $\alpha \geq \max_{i=1,\dots,n} z_i$, dann ist $L_n(\alpha) = \frac{1}{\alpha^n} > 0$



$\Rightarrow \hat\alpha_{ML} = \max_{i=1,\dots,n} z_i$

Vgl.:

$\hat\alpha_{MOMENTEN} = 2\bar z$

Bsp III 15

z_i i.i.d $N(\mu, \Sigma)$ (d.h. z_i ... $q \times 1$ Zufallsvektoren)

→ Vektoren der Länge q

$$\hat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} z_i = \bar{z}$$

$$\hat{\Sigma}_{ML} = \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})(z_i - \bar{z})'$$

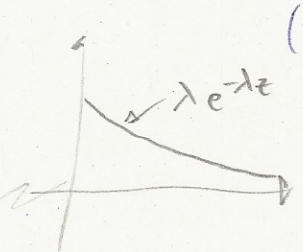
Ohne Rechnung

Bsp III 16

z_i i.i.d. exponentialverteilt mit $\lambda > 0$

Daten: $z_1, \dots, z_n$ $(z_i > 0)$

$$L_n(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda z_i} > 0 \text{ für } \lambda > 0$$



$$l_n(\lambda) = \sum_{i=1}^{n}(\log\lambda - \lambda z_i)$$

$$= n\log\lambda - \lambda\sum_{i=1}^{n} z_i$$

$$\frac{\partial l_n}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} z_i$$

FOC: $\frac{n}{\hat{\lambda}_{ML}} = \sum_{i=1}^{n} z_i \Rightarrow \hat{\lambda}_{ML} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} z_i} = \frac{1}{\bar{z}}$

SOC: $-\frac{n}{\hat{\lambda}^2_{ML}} < 0 \rightarrow \max$

$E(z_i) = \frac{1}{\lambda}$

Momentenmethode $\hat{\lambda} = \frac{1}{\bar{z}}$