Datenanalyse in der Physik

Vorlesung 4

Zentraler Grenzwertsatz und mehrdimensionale Wahrscheinlichkeitsdichteverteilungen

Prof. Dr. J. Mnich

DESY und Universität Hamburg

Der zentrale Grenzwertsatz

Der zentrale Grenzwertsatz ist der wichtigste mathematische Satz in der Statistik

Er begründet die wichtige Rolle der Gauß-Verteilung

Seien x_i $(i = 1, \ldots, n)$ unabhängige, kontinuierliche Zufallsvariablen, die beliebigen Wahrscheinlichkeitsdichteverteilungen mit Mittelwerten μ_i und Varianzen σ_i^2 folgen.

Im Limes $n \to \infty$ geht die Wahrscheinlichkeitsdichteverteilung für die Summe

$w = \sum_{i=1}^{n} x_i$ in eine Gauß-Verteilung mit Mittelwert

$\langle w \rangle = \mu = \sum_{i=1}^{n} \mu_i$ und Varianz

$\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$ über.

Beweis des zentralen Grenzwertsatzes findet sich z.B. im Buch von G. Cowan

Der zentrale Grenzwertsatz und systematische Fehler

Warum ist der zentrale Grenzwertsatz und die Gauß-Verteilung so wichtig in der Physik?

- Sehr häufig hängen Messungen von vielen Unsicherheiten ab
  Beiträge der einzelnen Unsicherheiten können durch Zufallsvariablen beschrieben werden
- Beispiel: Messung einer Länge mit einem Maßstab
  Ergebnis hängt von vielen (hoffentlich) kleinen Unsicherheiten ab
  optische Parallaxe, Eichung des Maßstabs, Temperatur, Zittern der Hand, . . .

- Auch wenn einzelne Quellen und ihre Beiträge unbekannt sind kann wegen des zentralen Grenzwertsatzes der Gesamtbeitrag der Unsicherheiten durch Gauß-Verteilung approximiert werden

  unabhängig von den pdf's der beitragenden Zufallsvariablen

Systematische Unsicherheiten, auch systematische Fehler genannt, können deshalb oft durch eine gaußverteilte Zufallsvariable beschrieben werden

Der zentrale Grenzwertsatz und systematische Fehler

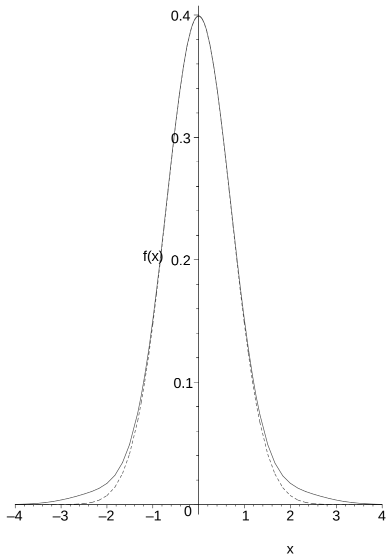
Aber Achtung:

Die Gauß-Verteilung ist nur eine Näherung, insbesondere außerhalb des Zentralbereichs (die Schwänze der Verteilung)
Beispiel:

- Beim Durchgang durch Materie erleiden hochenergetische, geladene Teilchen viele Stöße, die zu einer kleinen Ablenkung führen

  $\Longrightarrow$ Gauß-Verteilung des resultierenden Ablenkwinkels
- Selten kommt es zu Stößen, die zu einer großen Ablenkungen führen meist gar nicht, selten einmal, noch seltener zweimal

  $\Longrightarrow$ nicht gaußverteilter Beitrag zum Ablenkwinkel
- Resultierende Wahrscheinlichkeitsverteilung ist nur im Zentralbereich gut durch Gauß-Verteilung beschrieben



Der zentrale Grenzwertsatz

Bestimmung bzw. Abschätzung der Varianz der resultierenden Gauß-Verteilung, d.h. des systematischen Fehlers, ist oft das Hauptproblem einer Messung

Wichtige Anwendung des zentralen Grenzwertsatzes:

Standardabweichungen aus unabhängigen Fehlerquellen addieren sich quadratisch

$$\sigma = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$$

Beim Auslegen eines Experimentes gilt es, besonders die großen Beiträge zum Fehler zu beachten

Beispiel: Ein Experiment habe zwei Fehlerquellen σ_1 und $\sigma_2 = 0.1\sigma_1$

$$\Longrightarrow \sigma = \sqrt{\sigma_1^2 + \sigma_2^2} = \sigma_1\sqrt{1 + (0.1)^2} \approx \sigma_1(1 + 0.005)$$

Fehler, Fehler, Fehler, …

Eine erste Zusammenfassung zum Thema Fehler (Unsicherheit) und dessen Bestimmung:

Wir betrachten Zählexperimente, also Experimente, deren Ausgang eine bestimmte Zahl von Ereignisse ist (radioaktive Zerfälle, Experimente an Beschleunigern etc.)

- Einmalige Messung mit dem Ausgang x_m
  Die zu Grunde liegende pdf ist eine Gauß-Verteilung (oder Poisson).
  In diesem Fall gilt $\sigma^2 = \mu$, wobei $\mu = \hat{x}$ der wahre Mittelwert der Gauß-Verteilung ist.
  Falls wir μ kennen, z.B aus anderen, genaueren Experimenten, einer Theorie etc., folgt das Ergebnis:

$$x_m \pm \sqrt{\mu}$$

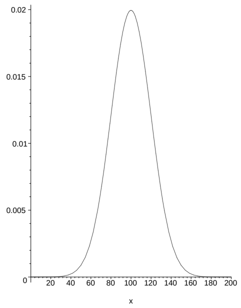
  $\sigma_m = \sqrt{\mu}$ ist der Fehler der Einzelmessung. In dem so definierten Intervall ist der wahre Wert $\hat{x}$ mit 68% Wahrscheinlichkeit (Konfidenz) enthalten. Falls man den wahren Wert nicht kennt, so nehme man x_m als beste Schätzung, also

$$x_m \pm \sqrt{x_m}$$

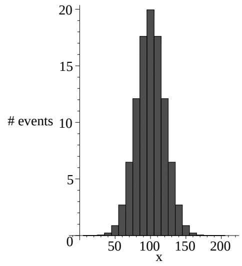
Fehler, Fehler, Fehler, . . .

- Wiederholte Messung (n-mal) einer Größe x, die gleicher pdf mit Varianz σ^2 unterliegt

  Mittelwert $x_m = \langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i$ Fehler auf Mittelwert $\sigma_m = \frac{\sigma}{\sqrt{n}}$



pdf mit Mittelwert μ und Varianz σ^2



n Messungen mit Mittelwert x_m und Varianz σ^2 und Fehler auf dem Mittelwert $\sigma_m = \frac{\sigma}{\sqrt{n}}$

- Beweis durch zentralen Grenzwertsatz:

  Betrachte Summe s der Messwerte $s = \sum_{i=1}^{n} x_i = n x_m$

  Die Varianz dieser Summe ist $\sigma_s^2 = \sum_{i=1}^{n} \sigma^2 = n\sigma^2$ oder $\sigma_s = \sqrt{n}\sigma$

  Daraus folgt sofort $\sigma_m = \frac{1}{n}\sigma_s = \frac{\sigma}{\sqrt{n}}$

  Mehr zu Fehlern folgt später

Mehrdimensionale Wahrscheinlichkeitsdichteverteilungen

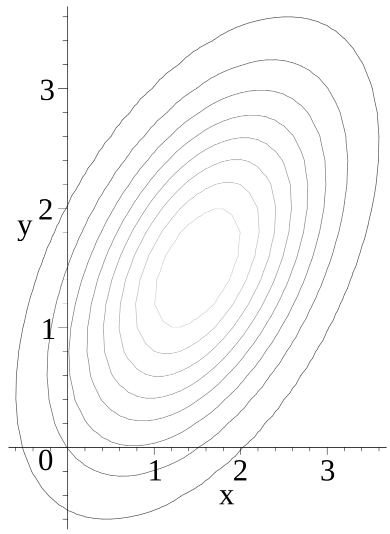
Häufig wird in Messungen mehr als eine Größe bestimmt, die entsprechend durch mehrere Zufallsvariablen beschrieben werden

$\Longrightarrow$ mehrdimensionale Wahrscheinlichkeitsdichteverteilungen

Wir betrachten zunächst den zweidimensionalen Fall, d.h. die pdf $f(x, y)$ hängt von den beiden kontinuierlichen Zufallsvariablen x und y ab

Beispiel einer zweidimensionalen pdf der Zufallsvariablen x und y

Gezeigt sind die Konturlinien für $f(x, y) = \mathrm{const}$



Zweidimensionale Verteilungen

- Die Wahrscheinlichkeit, ein Wertepaar im Intervall $a \leq x < b$ und $c \leq y < d$ zu finden, ist

$$P(a \leq x < b, c \leq y < d) = \int_a^b \int_c^d f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

- Die Normierungsbedingung lautet

$$\int\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$$

- Falls die pdf $f(x,y)$ als Produkt $f(x,y) = g(x) \cdot h(y)$ geschrieben werden kann, sind die Zufallsvariablen <u>unabhängig</u>

$$\begin{aligned} P(a \leq x < b, c \leq y < d) &= \int_a^b \int_c^d g(x)\,h(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_a^b g(x)\,\mathrm{d}x \cdot \int_c^d h(y)\,\mathrm{d}y \\ &= P(a \leq x < b) \cdot P(c \leq y < d) \end{aligned}$$

Mittelwert und Varianz

- Definition von Mittelwert und Varianz sind Verallgemeinerungen des eindimensionalen Falls:

$$\begin{aligned} \langle x \rangle &= E[x] = \int\int x\,f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ \langle y \rangle &= E[y] = \int\int y\,f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ \sigma_x^2 &= E[(x - \langle x \rangle)^2] = \int\int (x - \langle x \rangle)^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ \sigma_y^2 &= E[(y - \langle y \rangle)^2] = \int\int (y - \langle y \rangle)^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y \end{aligned}$$

Kovarianz und Korrelationskoeffizient

- Ein sehr wichtiger Begriff ist die Kovarianz

$$\begin{aligned}\sigma_{xy} &= \mathrm{cov}(x,y) = E[(x-\langle x\rangle)(y-\langle y\rangle)] \\ &= \int\!\!\int (x-\langle x\rangle)(y-\langle y\rangle) f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int\!\!\int x\,y f(x,y)\,\mathrm{d}x\,\mathrm{d}y + \langle x\rangle\langle y\rangle \int\!\!\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &\quad - \langle x\rangle \int\!\!\int y f(x,y)\,\mathrm{d}x\,\mathrm{d}y - \langle y\rangle \int\!\!\int x f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= E[x\,y] - \langle x\rangle\langle y\rangle\end{aligned}$$

- und der Korrelationskoeffizient

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x\,\sigma_y}$$

Kovarianz und Korrelationskoeffizient

- Für unahängige Variablen x, y ist die Kovarianz (Korrelation) Null
Aus $f(x,y) = g(x)\,h(y)$ und oBdA $\int g(x)\,\mathrm{d}x = \int h(y)\,\mathrm{d}y = 1$ folgt

$$\begin{aligned}\sigma_{xy} &= \int (x-\langle x\rangle)\,g(x)\,\mathrm{d}x \int (y-\langle y\rangle)\,h(y)\,\mathrm{d}y \\ &= \left(\int x\,g(x)\,\mathrm{d}x - \langle x\rangle \int g(x)\,\mathrm{d}x\right)\left(\int y\,h(y)\,\mathrm{d}y - \langle y\rangle \int h(y)\,\mathrm{d}y\right) \\ &= (\langle x\rangle - \langle x\rangle)\,(\langle y\rangle - \langle y\rangle) = 0\end{aligned}$$

oder äquivalent für den Korrelationskoeffizienten

$$\rho_{xy} = 0$$

Achtung: Die Umkehrung gilt nicht!
Aus $\sigma_{xy} = 0$ bzw. $\rho_{xy} = 0$ folgt nicht unbedingt, dass die Variablen unabhängig sind

Bedeutung der Kovarianz

Die Kovarianz ist der Erwartungswert der Funktion $(x - \langle x \rangle)(y - \langle y \rangle)$

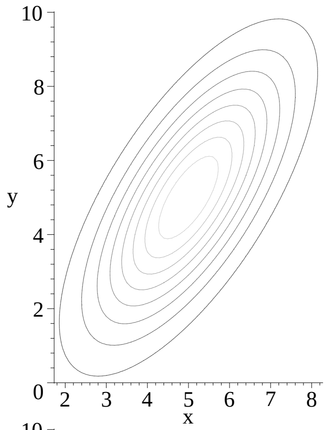
$\Longrightarrow$ Produkt der Abweichungen von den Mittelwerten

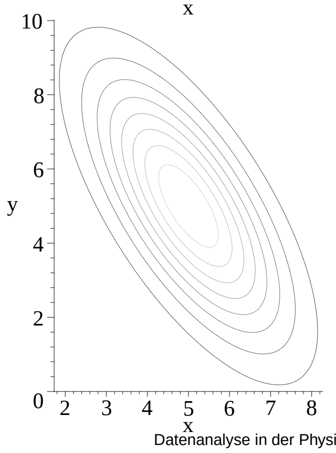
- Nehmen wir an, dass $\mathrm{cov}(x, y) > 0$ sei

  $\Longrightarrow$ wenn ein Wert $x > \langle x \rangle$ gefunden wird, ist die Wahrscheinlichkeit auch einen Wert $y > \langle y \rangle$ zu finden erhöht (und umgekehrt)

- Falls $\mathrm{cov}(x, y) < 0$ ist

  $\Longrightarrow$ wenn ein Wert $x > \langle x \rangle$ gefunden wird, ist die Wahrscheinlichkeit einen Wert $y > \langle y \rangle$ zu finden geringer (und umgekehrt)



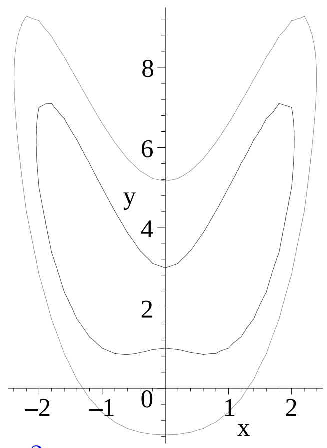


Der Korrelationskoeffizient

- Der Korrelationskoeffizient ist ein Maß für die Abhängigkeit zweier Zufallsvariablen

  Sinnvolle Größe wenn Konturen der pdf ungefähr elliptisch sind
  aber nicht unbedingt bei seltsamen Wahrscheinlichkeitsdichten!

  In dieser pdf sind die Zufallsvariablen x, y nicht unabhängig
  Auf Grund der Symmetrie ergibt sich aber trotzdem $\rho_{xy} = 0$



- Für den Korrelationskoeffizienten gilt

$$-1 \leq \rho_{xy} \leq +1 \qquad \text{oder} \qquad \rho_{xy}^2 \leq 1$$

$\rho_{xy} > 0$: die Zufallsvariablen sind korreliert
$\rho_{xy} < 0$: die Zufallsvariablen sind anti-korreliert

Die Extremfälle $\rho_{xy} = \pm 1$ werden angenommen, wenn ein linearer Zusammenhang zwischen den Variablen besteht, also $y = a + bx$

Beweise zum Korrelationskoeffizienten

Die nachfolgenden Rechnungen vereinfachen sich, wenn man folgende, leicht zu verifizierende Rechenregeln für Erwartungswerte benutzt ($a \in R$):

$$E[a\,h(x)] = a\,E[h(x)] \quad E[a + h(x)] = a + E[h(x)] \quad E[h(x) + j(x)] = E[h(x)] + E[j(x)]$$

Wir betrachten die transformierten Variablen

$$u = \frac{x - \langle x \rangle}{\sigma_x} \qquad \text{und} \qquad v = \frac{y - \langle y \rangle}{\sigma_y}$$

für die man leicht verifizieren kann, das gilt

$$\langle u \rangle = \langle v \rangle = 0 \qquad \text{und} \qquad \sigma_u^2 = \sigma_v^2 = 1$$

Nun berechnen wir die Varianz der Zufallsvariablen $u + v$ und $u - v$

$$\sigma_{u \pm v}^2 = E[(u \pm v)^2] = E[u^2] + E[v^2] \pm 2\,E[uv] = \sigma_u^2 + \sigma_v^2 \pm 2\,\mathrm{cov}(u,v) = 2\,(1 \pm \rho_{uv})$$

Weil nun für jede Varianz gelten muss $\sigma^2 \geq 0$ folgt $1 + \rho_{uv} \geq 0$ und $1 - \rho_{uv} \geq 0$, also

$$-1 \leq \rho_{uv} \leq +1 \qquad \text{oder} \qquad \rho_{uv}^2 \leq 1$$

Beweise zum Korrelationskoeffizienten

Es bleibt zu zeigen, dass $\rho_{uv} = \rho_{xy}$:

$$\rho_{uv} = \frac{\sigma_{uv}}{\sigma_u \sigma_v} = \frac{E[uv]}{1} = \frac{1}{\sigma_x \sigma_y} E[(x - \langle x \rangle)(y - \langle y \rangle)] = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} = \rho_{xy}$$

Wir betrachten den Extremfall $\rho_{xy} = \rho_{uv} = +1$. Daraus folgt $\sigma_{u-v}^2 = 0$ und wegen

$$0 = E[(u - v)^2] = \int\!\!\int (u - v)^2\, f(u,v)\, \mathrm{d}u\, \mathrm{d}v$$

ist das i.a. nur der Fall für $u - v = 0$.
Ausgedrückt in x und y heißt das

$$\frac{x - \langle x \rangle}{\sigma_x} = \frac{y - \langle y \rangle}{\sigma_y}$$

also ein linearer Zusammenhang der Form $y = ax + b$ mit $b > 0$
Entsprechen erhält man für $\rho_{xy} = -1$ diese Form mit $b < 0$

Zweidimensionale Gauß-Verteilung

Ein wichtiger Spezialfall liegt vor, wenn zwei gaußverteilte Zufallsvariablen vorliegen. Die zweidimensionale pdf lautet:

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

Die Mittelwerte und Varianzen sind

$$\langle x\rangle = \xi \quad \langle y\rangle = \eta \qquad E[(x-\langle x\rangle)^2] = \sigma_x^2 \quad E[(y-\langle y\rangle)^2] = \sigma_y^2$$

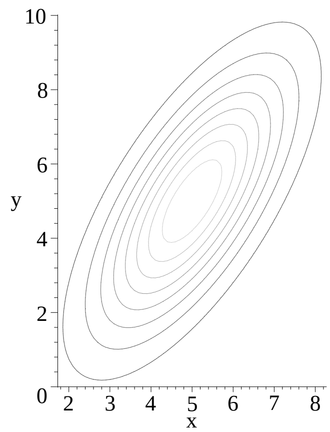
und der Korrelationskoeffizient ist

$$\rho_{xy} = \rho$$

Zweidimensionale Gauß-Verteilung

- Kurven gleicher Wahrscheinlichkeitsdichte (Konturen) sind Ellipsen in der (x,y)-Ebene

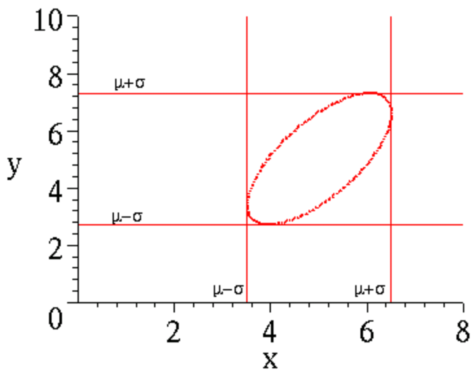
  Exponent in der pdf ist eine Ellipsengleichung!



- Wir betrachten die Kontur, bei der die Wahrscheinlichkeitsdichte auf $1/\sqrt{e}$ gefallen ist (die 1-σ-Kontur)

  Diese Ellipse passt genau in das Rechteck, das gegeben ist durch

$$\begin{array}{ccccc} \langle x\rangle - \sigma_x & \leq & x & \leq & \langle x\rangle + \sigma_x \\ \langle y\rangle - \sigma_y & \leq & y & \leq & \langle y\rangle + \sigma_y \end{array}$$



Beweis 1-σ-Kontur

OBdA setzen wir $\langle x\rangle = \langle y\rangle = 0$, d.h. das Maximum von $f(x,y)$ ist bei $x = y = 0$ und die 1-σ-Kontur ist definiert durch die Ellipsengleichung

$$\left(\frac{x}{\sigma_x}\right)^2 - 2\,\rho\,\frac{x}{\sigma_x}\,\frac{y}{\sigma_y} + \left(\frac{y}{\sigma_y}\right)^2 = 1 - \rho^2$$

Differenziere nach x $\quad \left(y = y(x) \quad y' = \frac{\mathrm{d}y}{\mathrm{d}x}\right)$

$$\frac{2\,x}{\sigma_x^2} - \frac{2\,\rho}{\sigma_x\sigma_y}\,(y + x\,y') + \frac{2\,yy'}{\sigma_y^2} = 0$$

Extremwerte y_m von $y(x)$ werden erreicht bei $y'(x_m) = 0$

$$\frac{2\,x_m}{\sigma_x^2} - \frac{2\,\rho}{\sigma_x\sigma_y}\,y_m = 0 \qquad \Longrightarrow \qquad x_m = \rho\,y_m\,\frac{\sigma_x}{\sigma_y}$$

Einsetzen in obige Ellipsengleichung

$$\rho^2\left(\frac{y_m}{\sigma_y}\right)^2 - 2\,\rho^2\,\frac{y_m}{\sigma_y}\,\frac{y_m}{\sigma_y} + \left(\frac{y_m}{\sigma_y}\right)^2 = 1-\rho^2 \qquad \Longrightarrow \qquad \left(\frac{y_m}{\sigma_y}\right)^2\left[\rho^2 - 2\,\rho^2 + 1\right] = 1-\rho^2$$

$$\Longrightarrow \qquad y_m^2 = \sigma_x^2 \qquad \text{oder} \qquad y_m = \pm\sigma_y$$

Die Berechnung der maximalen x-Werte der Ellipse erfolgt analog

Integrierte zweidimensionale Gauß-Verteilung

- Erinnerung eindimensionale Gauß-Verteilung

  Wahrscheinlichkeit im Intervall $|x - \langle x\rangle| \leq 1\sigma$ ist 68%
  Wahrscheinlichkeit im Intervall $|x - \langle x\rangle| \leq 2\sigma$ ist 95% etc.

  An den Integrationsgrenzen ist die Wahrscheinlichkeitsdichte auf $e^{-1/2}$, e^{-2}, ... des Maximums gesunken

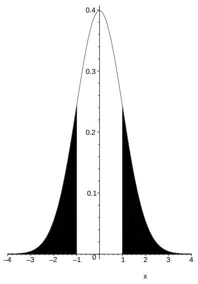
- Zweidimensionale Gauß-Verteilung

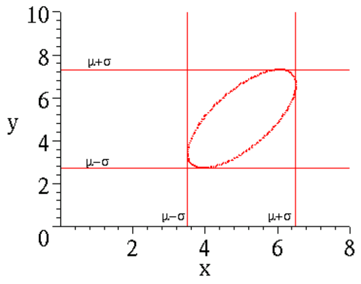
  Analog kann man integrierte Wahrscheinlichkeiten definieren durch Integrale über von Ellipsen umschlossene Flächen

  1-σ-Kontur, 2-σ-Kontur,... definiert über konstante Wahrscheinlichkeits- dichten bei $e^{-1/2}$, e^{-2}, ... des Maximums

Integrierte zweidimensionale Gauß-Verteilung

Vergleich integrierter Wahrscheinlichkeiten von ein- und zweidimensionalen Gauß-Verteilungen





	eindimensional	zweidimensional	
		Rechteck	k-σ-Konturen
k	$\lvert x - \langle x \rangle \rvert \leq k\,\sigma$	$\lvert x - \langle x \rangle \rvert \leq k\,\sigma_x$ $\lvert y - \langle y \rangle \rvert \leq k\,\sigma_y$	
1	0,68	$0,47\ (= 0,68^2)$	0,39
2	0,95	$0,91\ (= 0,95^2)$	0,63
		für $\rho = 0$	unabhängig von ρ

Integrierte zweidimensionale Gauß-Verteilung

- Die zweidimensionalen Wahrscheinlichkeiten für eine gleiche Zahl von σ's sind kleiner als im eindimensionalen Fall
  - Die Wahrscheinlichkeit, dass sich (x, y) im durch $\lvert x - \langle x \rangle \rvert \leq k\,\sigma_x$ und $\lvert y - \langle y \rangle \rvert \leq k\,\sigma_y$ befindet, ist das Produkt der Wahrscheinlichkeiten im eindimensionalen Fall (für $\rho = 0$).
  - Die Fläche der Ellipse ist kleiner als das Rechteck und deshalb auch der Wahrscheinlichkeitsinhalt

Verallgemeinerung auf mehrdimensionale pdf

Verallgemeinerung des Falles zweier kontinuierlicher Zufallsvariablen (x, y) auf n Zufallsvariablen $x_1, x_2, \ldots, x_n$

Hier ist es praktisch Vektoren und Matrizen einzuführen

- Die Zufallsvariablen bilden einen n-komponentigen Spaltenvektor

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

- Die Wahrscheinlichkeitsdichteverteilung ist eine Funktion dieses Vektors

$$f(\vec{x}) = f(x_1, x_2, \ldots, x_n)$$

- Die Wahrscheinlichkeitsdichteverteilung ist natürlich korrekt normiert

$$\int \ldots \int f(\vec{x})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n = 1$$

Mehrdimensionale pdf: Mittelwert und Varianz

- Definition der Mittelwerte

$$\langle x_i \rangle = E[x_i] = \int \ldots \int x_i\, f(\vec{x})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Die n Mittelwerte $\mu_i = \langle x_i \rangle$ können wieder als Vektor aufgefasst werden:

$$\vec{\mu} = \begin{pmatrix} \langle x_1 \rangle \\ \langle x_2 \rangle \\ \vdots \\ \langle x_n \rangle \end{pmatrix}$$

- Definition der Varianzen

$$\sigma_i^2 = \int \ldots \int (x_i - \langle x_i \rangle)^2\, f(\vec{x})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Kovarianzmatrix

Als Verallgemeinerung der Kovarianz zweier Zufallsvariablen erhält man die Kovarianzmatrix

$$V = V(\vec{x}) = E\left[(\vec{x} - \langle\vec{x}\rangle)\,(\vec{x} - \langle\vec{x}\rangle)^{\mathrm{T}}\right] = E\left[\begin{pmatrix} x_1 - \langle x_1\rangle \\ \vdots \\ x_n - \langle x_n\rangle \end{pmatrix} (x_1 - \langle x_1\rangle, \ldots, x_n - \langle x_n\rangle)\right]$$

$(\vec{x} - \langle\vec{x}\rangle)^{\mathrm{T}}$ ist der Zeilenvektor zu $(\vec{x} - \langle\vec{x}\rangle)$, d.h. die transponierte $n \times 1$ Matrix

In Komponentenschreibweise

- Diagonalelemente der Matrix V

$$V_{ii} = E\left[(x_i - \langle x_i\rangle)\,(x_i - \langle x_i\rangle)\right] = E\left[(x_i - \langle x_i\rangle)^2\right] = \sigma_i^2$$

$\Longrightarrow$ Varianz der i-ten Komponente von $\vec{x}$ (bzw. der Zufallsvariablen x_i)

- Nicht-Diagonalelemente der Matrix V

$$V_{ij} = E\left[(x_i - \langle x_i\rangle)\,(x_j - \langle x_j\rangle)\right] = \mathrm{cov}(x_i, x_j) = \sigma_{ij}$$

$\Longrightarrow$ Kovarianz der Zufallsvariablen x_i und x_j

- Die Kovarianzmatrix ist eine symmetrische $n \times n$-Matrix, denn offenbar gilt

$$V_{ij} = V_{ji} \qquad \text{bzw.} \qquad \sigma_{ij} = \sigma_{ji}$$

Kovarianz- und Korrelationsmatrix

Ausgeschrieben lautet die Kovarianzmatrix:

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \cdots & \sigma_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \sigma_{3n} & \cdots & \sigma_n^2 \end{pmatrix}$$

Entsprechend definiert man die Elemente C_{ij} der Korrelationsmatrix

$$C_{ij} = \rho_{ij} = \frac{V_{ij}}{\sigma_i\,\sigma_j}$$

$$C = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \rho_{23} & \cdots & \rho_{2n} \\ \rho_{13} & \rho_{23} & 1 & \cdots & \rho_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1n} & \rho_{2n} & \rho_{3n} & \cdots & 1 \end{pmatrix}$$