Universität Bielefeld

Fakultät für Mathematik
Prof. Dr. Barbara Gentz SS 2015

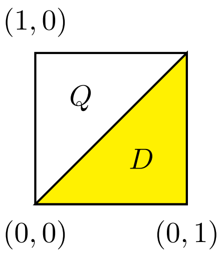
Vertiefung NWI: 12. Vorlesung zur Wahrscheinlichkeitstheorie

Mittwoch, 24.6.2015

12. Unabhängigkeit von Zufallsgrößen mit Dichten

12.1 Mehrdimensionale Dichten

Beispiel
Wahrscheinlichkeit, daß ein im Einheitsquadrat Q zufällig gewählter Punkt in dem Dreieck D liegt, das unterhalb der Winkelhalbierenden liegt:



$$Q = [0,1] \times [0,1] = \{(x,y) \in \mathbb{R}^2 \colon 0 \le x \le 1, 0 \le y \le 1\}$$
$$D = \{(x,y) \in Q \colon y \le x\}$$

Es muß offenbar $\mathbb{P}\{D\} = |D|/|Q| = 1/2$ gelten, wobei $|\cdot|$ den Inhalt (hier: den Flächeninhalt) einer Menge bezeichne.

Formelle Berechnung der Inhalte:

$$|Q| = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_Q(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \left(\int_0^1 1 \, \mathrm{d}x \right) \mathrm{d}y = \int_0^1 1 \, \mathrm{d}y = 1$$
$$|D| = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_D(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \left(\int_y^1 1 \, \mathrm{d}x \right) \mathrm{d}y = \int_0^1 (1-y) \, \mathrm{d}y = 1 - \frac{1}{2} = \frac{1}{2}$$
$$\mathbb{P}\{D\} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_D(x,y) \, \mathrm{d}x \, \mathrm{d}y}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_Q(x,y) \, \mathrm{d}x \, \mathrm{d}y} = \frac{1}{2}$$

Definition: *mehrdimensionale Dichte*

- $f:\mathbb{R}^n \to [0,\infty)$ heißt n-*dimensionale Dichte*, falls f integrierbar ist (vgl. Literatur zur Maß- und Integrationstheorie) mit

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1,\dots,x_n)\,\mathrm{d}x_1\dots\mathrm{d}x_n = 1$$

- Eine n-dimensionale Dichte f heißt *gemeinsame Dichte* der Zufallsgrößen $X_1,\dots,X_n$, wenn für alle $a_1,\dots,a_n \in \mathbb{R}$

$$(*) \qquad \mathbb{P}\{X_1 \le a_1,\dots,X_n \le a_n\} = \int_{-\infty}^{a_n} \dots \int_{-\infty}^{a_1} f(x_1,\dots,x_n)\,\mathrm{d}x_1\dots\mathrm{d}x_n$$

  gilt.

Bemerkungen

- Mit $C = (-\infty,a_1] \times \dots \times (-\infty,a_n]$ kann $(*)$ geschrieben werden als

$$(**) \qquad \mathbb{P}\{(X_1,\dots,X_n) \in C\} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} 1_C(x_1,\dots,x_n) f(x_1,\dots,x_n)\;\mathrm{d}x_1\dots\mathrm{d}x_n$$

  Dabei gilt $(**)$ auch für allgemeinere Mengen C, nämlich für alle *meßbaren* Mengen. Zu den meßbaren Mengen gehören insbesondere alle Mengen, die sich als höchstens abzählbare Vereinigungen und Schnitte von Rechtecken darstellen lassen (vgl. Literatur zur Maß- und Integrationstheorie).

  Abkürzend schreiben wir

$$\int_C f(x_1,\dots,x_n)\,\mathrm{d}x_1\dots\mathrm{d}x_n$$

  für die iterierten Integrale auf der rechten Seite von $(**)$.

- Für nichtnegative (meßbare) Integranden ist die Reihenfolge der Integration egal und es gilt

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1,\dots,x_n)\,\mathrm{d}x_1\dots\mathrm{d}x_n = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1,\dots,x_n)\,\mathrm{d}x_{\pi_1}\dots\mathrm{d}x_{\pi_n}$$

  für jede Permutation $(\pi_1,\dots,\pi_n)$ von $(1,\dots,n)$.
  Überzeugen Sie sich davon anhand des obigen Beispiels:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_D(x,y)\;\mathrm{d}y\,\mathrm{d}x = \int_0^1 \left(\int_0^x 1\;\mathrm{d}y\right)\mathrm{d}x = \int_0^1 x\;\mathrm{d}x = \frac{1}{2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_D(x,y)\;\mathrm{d}x\,\mathrm{d}y$$

Beispiel: *Uniforme Verteilung*

- Für eine (meßbare) Menge $C \subset \mathbb{R}^2$ gibt das Doppelintegral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_C(x,y)\ \mathrm{d}x\,\mathrm{d}y = \int_C 1\ \mathrm{d}x\,\mathrm{d}y$$

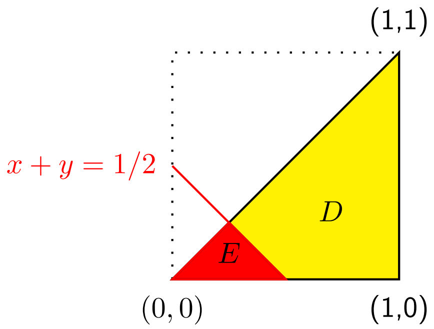
den Flächeninhalt $|C|$.

- Für eine (meßbare) Menge $C \subset \mathbb{R}^n$ ist

$$f(x_1,\ldots,x_n) = \frac{1}{|C|} 1_C(x_1,\ldots,x_n)$$

die *Dichte der uniformen Verteilung auf* C.

- Sei $C = D$ das Dreieck aus dem obigen Beispiel. Wie groß ist die Ws., daß die Summe $X+Y$ der Koordinaten (X,Y) eines zufällig gewählten Punkts in D kleiner als 1/2 ist?
Sei $E = \{(x,y) \in D\colon x+y < 1/2\}$.
Raten Sie die Ws. mit Hilfe einer Skizze:



$$\mathbb{P}\{(X,Y)\in E\} = \frac{|E|}{|D|} = \frac{2\cdot\frac{1}{2}\cdot(\frac{1}{4})^2}{\frac{1}{2}} = \frac{1}{8}$$

- Das Berechnen gemäß Definition erfordert eine Vorbereitung:

$$\begin{aligned}(x,y)\in E = E\cap D &\iff 0\le y\le x\le 1 \text{ und } x+y<\tfrac{1}{2}\\ &\iff x\in[0,\tfrac{1}{2}) \text{ und } 0\le y\le x \text{ und } y<\tfrac{1}{2}-x\end{aligned}$$

Somit ist

$$\begin{aligned}\mathbb{P}\{(X,Y)\in E\} &= \int_E f(x,y)\ \mathrm{d}x\,\mathrm{d}y && \text{mit } f = \frac{1}{|D|}1_D(x,y)\\ &= \frac{1}{|D|}\int_{E\cap D} 1\ \mathrm{d}y\,\mathrm{d}x && \text{(Vertauschen von } \mathrm{d}x \text{ und } \mathrm{d}y)\\ &= \frac{1}{|D|}\int_0^{1/2}\Big(\int_0^{\min\{x,1/2-x\}} 1\ \mathrm{d}y\Big)\,\mathrm{d}x\\ &= \frac{1}{|D|}\int_0^{1/2}\min\{x,1/2-x\}\ \mathrm{d}x\\ &= \frac{1}{|D|}\Big(\int_0^{1/4} x\ \mathrm{d}x + \int_{1/4}^{1/2}(1/2-x)\ \mathrm{d}x\Big) = \frac{1}{|D|}\,2\cdot\frac{x^2}{2}\Big|_{x=0}^{1/4} = \frac{1}{8}\end{aligned}$$

12.2 Randdichten

Bemerkung: *von der gemeinsamen Dichte zur Randdichte*
Aus der gemeinsamen Dichte von $X_1, \dots, X_n$ läßt sich die Dichte eines jeden X_k berechnen:
Sei dazu $C = \mathbb{R}^{k-1} \times (-\infty, a] \times \mathbb{R}^{n-k}$, $a \in \mathbb{R}$ beliebig. Dann ist

$$\begin{aligned}
\mathbb{P}\{X_k \le a\} &= \mathbb{P}\{(X_1, \dots, X_n) \in C\} \\
&= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{a} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n)\ \mathrm{d}x_1 \dots \mathrm{d}x_n \\
&= \int_{-\infty}^{a} \left(\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n)\ \mathrm{d}x_1 \dots \mathrm{d}x_{k-1}\, \mathrm{d}x_{k+1} \dots \mathrm{d}x_n \right) \mathrm{d}x_k \\
&= \int_{-\infty}^{a} f_k(x)\ \mathrm{d}x
\end{aligned}$$

mit

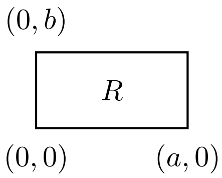
$$(***) \qquad f_k(x) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n)\ \mathrm{d}x_1 \dots \mathrm{d}x_{k-1}\, \mathrm{d}x_{k+1} \dots \mathrm{d}x_n$$

Somit ist f_k eine Dichte für X_k.

Definition: *k-te Randdichte*
Sei f die gemeinsame Dichte von $X_1, \dots, X_n$. Dann heißt f_k, wie in $(***)$ definiert, die k-te Randdichte von f.

Beispiele

- *Uniforme Verteilung auf einem Rechteck $R = [0, a] \times [0, b]$*



  Ist (X, Y) ein zufälliger Punkt aus R, so haben die Koordinaten X und Y die gemeinsame Dichte

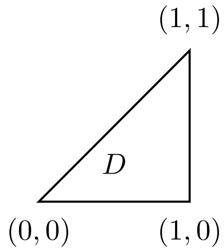
$$f(x, y) = \frac{1}{ab} 1_R(x, y) = \frac{1}{a} 1_{[0,a]}(x) \cdot \frac{1}{b} 1_{[0,b]}(y)\ .$$

  Es folgt für die Randdichten f_X von X und f_Y von Y durch “Rausintegrieren” von y bzw. x:

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x, y)\ \mathrm{d}y = \int_0^b \frac{1}{ab} 1_{[0,a]}(x)\ \mathrm{d}y = \frac{1}{a} 1_{[0,a]}(x) \\
f_Y(y) &= \int_{-\infty}^{\infty} f(x, y)\ \mathrm{d}x = \int_0^a \frac{1}{ab} 1_{[0,b]}(y)\ \mathrm{d}x = \frac{1}{b} 1_{[0,b]}(y)
\end{aligned}$$

  Wir haben in diesem Fall $f(x, y) = f_X(x) \cdot f_Y(y)$.

- *Uniforme Verteilung auf dem Dreieck D*



Ist (X, Y) ein zufälliger Punkt aus D, so haben die Koordinaten X und Y die gemeinsame Dichte

$$f(x,y) = \frac{1}{|D|} 1_D(x,y)$$

Es folgt für die Randdichten f_X von X und f_Y von Y:

$$\begin{aligned}
f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ \mathrm{d}y &= \begin{cases} \displaystyle\int_{-\infty}^{\infty} \frac{1}{|D|} 1_{[0,x]}(y)\ \mathrm{d}y & \text{für } x \in [0,1] \\ 0 & \text{sonst} \end{cases} \\
&= 1_{[0,1]}(x) \int_0^x \frac{1}{|D|}\ \mathrm{d}y = 2x 1_{[0,1]}(x) \\
f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\ \mathrm{d}x &= \int_y^1 \frac{1}{|D|} 1_{[0,1]}(y)\ \mathrm{d}x = 2(1-y) 1_{[0,1]}(y)
\end{aligned}$$

In diesem Fall gilt $f(x,y)] \neq f_X(x) \cdot f_Y(y)$.

12.3 Unabhängigkeit

Definition: *Unabhängigkeit von Zufallsgrößen*
$X_1, \ldots, X_n$ heißen *unabhängig*, wenn

$$\mathbb{P}\{X_1 \leq a_1, \ldots, X_n \leq a_n\} = \mathbb{P}\{X_1 \leq a_1\} \ldots \mathbb{P}\{X_n \leq a_n\}$$

für alle $a_1, \ldots, a_n \in \mathbb{R}$ gilt.

Bemerkung
Für diskrete Zufallsgrößen ist diese Definition äquivalent zur bekannten Definition.

Satz
Gegeben seien Zufallsgrößen $X_1, \ldots, X_n$, die jeweils eine Dichte f_i besitzen. Dann gilt:

$$X_1, \ldots, X_n \text{ unabhängig} \qquad \Longleftrightarrow \qquad \begin{array}{l} (x_1, \ldots, x_n) \mapsto f_1(x_1) \cdot \cdots \cdot f_n(x_n) \\ \text{gemeinsame Dichte von } X_1, \ldots, X_n \end{array}$$

Bemerkung
Die Existenz einer gemeinsamen Dichte wird nicht vorausgesetzt. Der Satz sagt also insbesondere:

- Wenn die Zufallsgrößen Dichten haben und unabhängig sind, dann existiert die gemeinsame Dichte und hat Produktform.
- Wenn die gemeinsame Dichte existiert und Produktform hat, dann sind die Zufallsgrößen unabhängig.

Wenn wir unabhängige Zufallsexperimente modellieren wollen, tun wir das also, indem wir für die gemeinsame Dichte die Produktform wählen.

Beispiele

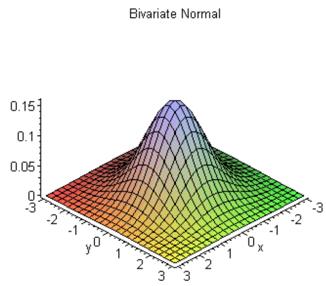
- Wir haben oben bereits gesehen, daß die gemeinsame Dichte der Koordinaten eines zufälligen Punkts im Rechteck R Produktform hat. Also sind die Koordinaten unabhängig.
- Wir haben auch gesehen, daß die Koordinaten eines zufälligen Punkts im Dreieck D nicht unabhängig sein können.

Erläutern Sie anschaulich den Grund für das unterschiedliche Verhalten.

Beispiel: *Mehrdimensionale Normalverteilung*

- Sind die Koordinaten X und Y unabhängig und standardnormalverteilt, so hat der zufällig gewählte Punkt (X, Y) die Dichte

$$f(x,y) = \varphi(x)\varphi(y) = \frac{1}{(\sqrt{2\pi})^2}\,\mathrm{e}^{-(x^2+y^2)/2}$$



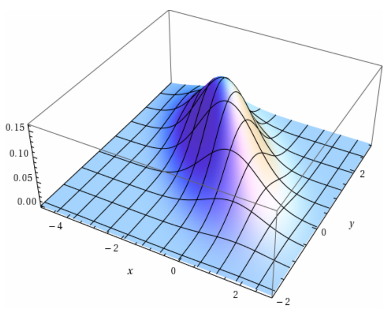
- Entsprechend ergibt sich für unabhängige standardnormalverteilte $X_1, \dots, X_n$ als gemeinsame Dichte die Dichte der n-dimensionale Normalverteilung zu

$$f(x) = \frac{1}{(\sqrt{2\pi})^n}\,\mathrm{e}^{-\|x\|^2/2} \qquad \text{für } x \in \mathbb{R}^n$$

- Die allgemeine, nicht-degenerierte n-dimensionale Normalverteilung ist gegeben durch ihre Dichte

$$f(x) = \frac{1}{\sqrt{(2\pi \det \Sigma)^n}}\,\mathrm{e}^{-\langle (x-\mu), \Sigma^{-1}(x-\mu)\rangle/2}$$

für $x \in \mathbb{R}^n$. Hier sind $\mu \in \mathbb{R}^n$, Σ eine positiv definite, symmetrische $(n \times n)$-Matrix und $\det \Sigma$ die Determinante von Σ.



$\mu = (-1, 1)$, $\Sigma = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$

Ist f die gemeinsame Dichte der Zufallsgrößen $X_1, \dots, X_n$, so können wir μ als Vektor der Erwartungswerte $(\mathbb{E}[X_1], \dots, \mathbb{E}[X_n])$ identifizieren, und die Elemente von Σ sind die Varianzen bzw. Kovarianzen $\Sigma_{i,j} = \mathrm{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]$.

- Die Randverteilungen sind wieder Normalverteilungen.

Satz

Die Zufallsgrößen X und Y seien unabhängig, mit existierenden Erwartungswerten.

- $\mathbb{E}[X \cdot Y]$ existiert, und es gilt $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.
- Falls die Varianzen von X, Y existieren, so gilt $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.