Statistik für Informatiker, SS 2017

1.2 Bedingte Wahrscheinlichkeiten und mehrstufige Zufallsexperimente

1.2.1 Nochmal zur Unabhängigkeit
1.2.2 Faltung

Matthias Birkner

http://www.staff.uni-mainz.de/birkner/StatInfo17/

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

7.6.2017

Erinnerung. Zufallsvariablen $X_1, \ldots, X_n$ (X_i habe Wertebereich S_i) heißen (stochastisch) *unabhängig*, wenn für alle Ereignisse $\{X_i \in B_i\}$ gilt

$$P\big(X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n\big) = \prod_{i=1}^{n} P(X_i \in B_i) \qquad (1)$$

(vgl. Def. 1.13).

Dann ist die Pfadformel (vgl. Beob. 1.51) besonders „angenehm".

Erinnerung. Zufallsvariablen $X_1, \ldots, X_n$ (X_i habe Wertebereich S_i) heißen (stochastisch) *unabhängig*, wenn für alle Ereignisse $\{X_i \in B_i\}$ gilt

$$P\big(X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n\big) = \prod_{i=1}^{n} P(X_i \in B_i) \qquad (1)$$

(vgl. Def. 1.13).

Dann ist die Pfadformel (vgl. Beob. 1.51) besonders „angenehm".

Definition 1.53

Ereignisse $A_1, \ldots, A_n$ heißen *unabhängig*, wenn dies für ihre Indikatorvariablen $I_{A_1}, \ldots, I_{A_n}$ gilt.

Speziell: A und B unabhängige Ereignisse (mit $P(B) > 0$), so ist $P(A \cap B) = P(A)P(B)$ und somit $P(A \mid B) = P(A)$.

Bemerkung 1.54

Sind ZVn $X_1, \dots, X_n$ unabhängig, so auch

1. jede Teilfamilie $X_{i_1}, \dots, X_{i_k}$ (für $1 \le i_1 < \cdots < i_k \le n$) (wähle $B_i = S_i$ in (1) für $i \notin \{i_1, \dots, i_k\}$)
2. $f_1(X_1), f_2(X_2), \dots, f_n(X_n)$ für Funktionen $f_i : S_i \to S_i'$ (beachte $\{f_i(X_i) \in B_i'\} = \{X_i \in f_i^{-1}(B_i')\}$ in (1), vgl. Bsp. 1.4, 2.)

Bemerkung 1.54 (Fortsetzung)

3. In Def. 1.53 genügt es i.A. nicht, jeweils nur Paare auf Unabhängigkeit zu prüfen:

   Beispiel: Seien X_1, X_2, X_3 unabhängige faire Münzwürfe $P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$, $Y_1 = I_{\{X_1 = X_2\}}, Y_2 = I_{\{X_1 = X_3\}}, Y_3 = I_{\{X_2 = X_3\}}$. Dann sind jeweils Y_1 und Y_2, Y_1 und Y_3, Y_2 und Y_3 unabhängig, aber Y_1, Y_2, Y_3 zusammen *nicht*.

Bemerkung 1.54 (Fortsetzung)

3. In Def. 1.53 genügt es i.A. nicht, jeweils nur Paare auf Unabhängigkeit zu prüfen:

   Beispiel: Seien X_1, X_2, X_3 unabhängige faire Münzwürfe $P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$, $Y_1 = I_{\{X_1 = X_2\}}$, $Y_2 = I_{\{X_1 = X_3\}}$, $Y_3 = I_{\{X_2 = X_3\}}$. Dann sind jeweils Y_1 und Y_2, Y_1 und Y_3, Y_2 und Y_3 unabhängig, aber Y_1, Y_2, Y_3 zusammen *nicht*.

   (Es ist $P(Y_i = 1) = \frac{1}{2}$, z.B. $P(Y_1 = 1, Y_2 = 1) = P\big(\{X_1 = X_2 = X_3 = 1\} \cup \{X_1 = X_2 = X_3 = 0\}\big) = (1/2)^3 + (1/2)^3 = 1/4 = P(Y_1 = 1)P(Y_2 = 1)$, $P(Y_1 = 1, Y_2 = 0) = P\big(\{X_1 = X_2 = 1, X_3 = 0\} \cup \{X_1 = X_2 = 0, X_3 = 1\}\big) = (1/2)^3 + (1/2)^3 = 1/4 = P(Y_1 = 1)P(Y_2 = 0)$, etc.)

Bemerkung 1.54 (Fortsetzung)

3. In Def. 1.53 genügt es i.A. nicht, jeweils nur Paare auf Unabhängigkeit zu prüfen:

   Beispiel: Seien X_1, X_2, X_3 unabhängige faire Münzwürfe $P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$, $Y_1 = I_{\{X_1 = X_2\}}$, $Y_2 = I_{\{X_1 = X_3\}}$, $Y_3 = I_{\{X_2 = X_3\}}$. Dann sind jeweils Y_1 und Y_2, Y_1 und Y_3, Y_2 und Y_3 unabhängig, aber Y_1, Y_2, Y_3 zusammen *nicht*.

   (Es ist $P(Y_i = 1) = \frac{1}{2}$, z.B. $P(Y_1 = 1, Y_2 = 1) = P\big(\{X_1 = X_2 = X_3 = 1\} \cup \{X_1 = X_2 = X_3 = 0\}\big) = (1/2)^3 + (1/2)^3 = 1/4 = P(Y_1 = 1)P(Y_2 = 1)$, $P(Y_1 = 1, Y_2 = 0) = P\big(\{X_1 = X_2 = 1, X_3 = 0\} \cup \{X_1 = X_2 = 0, X_3 = 1\}\big) = (1/2)^3 + (1/2)^3 = 1/4 = P(Y_1 = 1)P(Y_2 = 0)$, etc.)

   Man sagt dazu auch: Y_1, Y_2, Y_3 sind *paarweise unabhängig*, aber eben nicht unabhängig.

Faltung

Definition 1.55

X und Y unabhängige reellwertige ZVn, $X \sim \mu$, $Y \sim \nu$ (in einem gewissen Zufallsexperiment $\mathscr{X}$). Die Verteilung von $X + Y$ heißt die *Faltung* von μ und ν, geschrieben $\mu * \nu$:

$$(\mu * \nu)(B) = P(X + Y \in B), \quad B \subset \mathbb{R}$$

Faltung

Definition 1.55

X und Y unabhängige reellwertige ZVn, $X \sim \mu$, $Y \sim \nu$ (in einem gewissen Zufallsexperiment $\mathscr{X}$). Die Verteilung von $X + Y$ heißt die *Faltung* von μ und ν, geschrieben $\mu * \nu$:

$$(\mu * \nu)(B) = P(X + Y \in B), \quad B \subset \mathbb{R}$$

Bemerkung. $\mu * \nu = \nu * \mu$ (denn $X + Y = Y + X$).

Beobachtung 1.56 (Diskreter Fall)

Falls $\mu(\mathbb{Z}) = \nu(\mathbb{Z}) = 1$ *(d.h. X und Y haben Werte in* $\mathbb{Z}$*), so ist*

$$(\mu * \nu)(\{k\}) = P(X + Y = k) = \sum_{m\in\mathbb{Z}} P(X = m, Y = k - m)$$
$$= \sum_{m\in\mathbb{Z}} \mu(\{m\})\nu(\{k - m\}).$$

Im allg. diskreten Fall $P(X \in \{x_i, i \in \mathbb{N}\}, Y \in \{y_j, j \in \mathbb{N}]\}) = 1$ *muss man die „Doppelsumme" betrachten:*

$$P(X + Y = z) = \sum_{i,j\,:\,x_i+y_j=z} P(X = x_i)P(Y = y_j)$$

Beispiel 1.57

1. W_1, W_2 unabhängige 6-er Würfelwürfe, dann ist

$$S := W_1 + W_2 \sim \mathrm{Unif}_{\{1,2,\ldots,6\}} * \mathrm{Unif}_{\{1,2,\ldots,6\}}$$

mit

$$\begin{aligned} P(S=k) &= \sum_{m=\max\{k-6,1\}}^{\min\{k-1,6\}} P(W_1 = m) P(W_2 = k-m) \\ &= \frac{1}{36}\Big(\min\{k-1,6\} - \max\{k-6,1\} + 1\Big) \\ &= \frac{6-|7-k|}{36} \end{aligned}$$

für $k \in \{2, 3, \ldots, 12\}$

Beispiel 1.57 (Fortsetzung)

2. X, Y u.a., $\sim \mathrm{Ber}_p$, so ist $X + Y \sim \mathrm{Bin}_{2,p}$, d.h. $\mathrm{Ber}_p * \mathrm{Ber}_p = \mathrm{Bin}_{2,p}$.

Beispiel 1.57 (Fortsetzung)

2. X, Y u.a., $\sim \mathrm{Ber}_p$, so ist $X + Y \sim \mathrm{Bin}_{2,p}$, d.h. $\mathrm{Ber}_p * \mathrm{Ber}_p = \mathrm{Bin}_{2,p}$.
3. (Binomialfamilie) $X_1, X_2, \ldots, X_n$ u.a., $\sim \mathrm{Ber}_p$, so ist $X_1 + X_2 + \cdots + X_n \sim \mathrm{Bin}_{n,p}$, d.h.

$$\mathrm{Ber}_p^{*n} = \underbrace{\mathrm{Ber}_p * \mathrm{Ber}_p * \cdots * \mathrm{Ber}_p}_{n\text{-mal}} = \mathrm{Bin}_{n,p}.$$

Insbes.

$$\mathrm{Bin}_{n_1,p} * \mathrm{Bin}_{n_2,p} = \mathrm{Bin}_{n_1+n_2,p} \quad \text{für } p \in [0,1], n_1, n_2 \in \mathbb{N},$$

die Binomialverteilungen bilden (für festes p) eine *Faltungsfamilie*.

Beispiel 1.57 (Fortsetzung)

2. X, Y u.a., $\sim \mathrm{Ber}_p$, so ist $X + Y \sim \mathrm{Bin}_{2,p}$, d.h. $\mathrm{Ber}_p * \mathrm{Ber}_p = \mathrm{Bin}_{2,p}$.
3. (Binomialfamilie) $X_1, X_2, \dots, X_n$ u.a., $\sim \mathrm{Ber}_p$, so ist $X_1 + X_2 + \cdots + X_n \sim \mathrm{Bin}_{n,p}$, d.h.

$$\mathrm{Ber}_p^{*n} = \underbrace{\mathrm{Ber}_p * \mathrm{Ber}_p * \cdots * \mathrm{Ber}_p}_{n\text{-mal}} = \mathrm{Bin}_{n,p}.$$

Insbes.

$$\mathrm{Bin}_{n_1,p} * \mathrm{Bin}_{n_2,p} = \mathrm{Bin}_{n_1+n_2,p} \quad \text{für } p \in [0,1], n_1, n_2 \in \mathbb{N},$$

die Binomialverteilungen bilden (für festes p) eine *Faltungsfamilie*.

(Schreibe $S_1 := X_1 + \cdots + X_{n_1} \sim \mathrm{Bin}_{n_1,p}$, $S_2 := X_{n_1+1} + X_{n_1+2} + \cdots + X_{n_1+n_2} \sim \mathrm{Bin}_{n_2,p}$, so ist $S_1 + S_2 = X_1 + \cdots + X_{n_1+n_2} \sim \mathrm{Bin}_{n_1+n_2,p}$.)

Beispiel 1.57 (Fortsetzung)

4. (Poissonfamilie) Für $\alpha, \beta > 0$ ist $\mathrm{Poi}_\alpha * \mathrm{Poi}_\beta = \mathrm{Poi}_{\alpha+\beta}$, denn

$$\begin{aligned}\sum_{m=0}^{k} e^{-\alpha}\frac{\alpha^m}{m!}\cdot e^{-\beta}\frac{\beta^{k-m}}{(k-m)!} &= e^{-(\alpha+\beta)}\frac{1}{k!}\sum_{m=0}^{k}\binom{k}{m}\alpha^m\beta^{k-m}\\ &= e^{-(\alpha+\beta)}\frac{(\alpha+\beta)^k}{k!}\\ &= \mathrm{Poi}_{\alpha+\beta}(\{k\}), \quad k\in\mathbb{N}_0.\end{aligned}$$

Auch die Poissonverteilungen bilden eine Faltungsfamilie.

Beobachtung 1.58 (Faltung von Dichten)

X, Y u.a. reellwertige ZVn mit Dichte f_X bzw. f_Y, so hat $X+Y$ die Dichte

$$(f_X * f_Y)(z) := \int_{\mathbb{R}} f_X(x) f_Y(z-x)\, dx, \quad z \in \mathbb{R}.$$

Beobachtung 1.58 (Faltung von Dichten)

X, Y u.a. reellwertige ZVn mit Dichte f_X bzw. f_Y, so hat $X + Y$ die Dichte

$$(f_X * f_Y)(z) := \int_{\mathbb{R}} f_X(x) f_Y(z-x)\, dx, \quad z \in \mathbb{R}.$$

Es ist nämlich

$$\begin{aligned} P(X+Y \le w) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{1}_{\{x+y \le w\}} f_X(x) f_Y(y)\, dy dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{1}_{\{x+z-x \le w\}} f_X(x) f_Y(z-x)\, dz dx \\ &= \int_{-\infty}^{\infty} \mathbf{1}_{\{z \le w\}} \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx\, dz \\ &= \int_{-\infty}^{w} (f_X * f_Y)(z)\, dz \end{aligned}$$

wobei wir in der 2. Zeile $y = z - x$ substituiert haben.

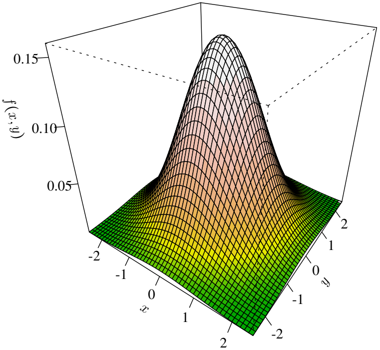
Beispiel 1.59 (Die Normalverteilungen bilden eine Faltungsfamilie)

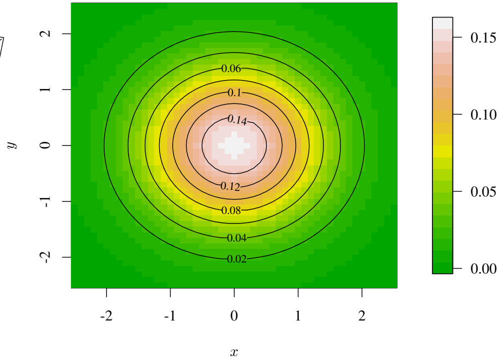
Es gilt

$$\mathcal{N}_{\mu_1,\sigma_1^2} * \mathcal{N}_{\mu_2,\sigma_2^2} = \mathcal{N}_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2} \quad \text{für } \mu_1, \mu_2 \in \mathbb{R},\ \sigma_1, \sigma_2 > 0$$

Erinnerung. Die 2-dimensionale Standard-Normalverteilung hat Dichte

$$f(x, y) = \frac{1}{2\pi} \exp\left(-\tfrac{1}{2}(x^2 + y^2)\right)$$





Zum Beweis von Bsp. 1.59:

Seien $a, b \in (0, 1)$ mit $a^2 + b^2 = 1$, so ist die 2×2-Matrix

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{orthogonal}$$

Dies ist eine Drehmatrix, wir könnten $a = \cos(\varphi), b = \sin(\varphi)$ für ein geeign. $\varphi \in [-\pi, \pi)$ schreiben, und

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & -ab + ba \\ -ba + ab & a^2 + b^2 \end{pmatrix}$$

Zum Beweis von Bsp. 1.59:

Seien $a, b \in (0,1)$ mit $a^2 + b^2 = 1$, so ist die 2×2-Matrix

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{orthogonal}$$

Dies ist eine Drehmatrix, wir könnten $a = \cos(\varphi), b = \sin(\varphi)$ für ein geeign. $\varphi \in [-\pi, \pi)$ schreiben, und

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & -ab + ba \\ -ba + ab & a^2 + b^2 \end{pmatrix}$$

Seien Z_1, Z_2 u.a., $\sim \mathcal{N}_{0,1}$, dann haben nach Beob. 1.42

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} aZ_1 + bZ_2 \\ -bZ_1 + aZ_2 \end{pmatrix}$$

dieselbe Verteilung

Zum Beweis von Bsp. 1.59:

Seien $a, b \in (0,1)$ mit $a^2 + b^2 = 1$, so ist die 2×2-Matrix

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{orthogonal}$$

Dies ist eine Drehmatrix, wir könnten $a = \cos(\varphi), b = \sin(\varphi)$ für ein geeign. $\varphi \in [-\pi, \pi)$ schreiben, und

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & -ab + ba \\ -ba + ab & a^2 + b^2 \end{pmatrix}$$

Seien Z_1, Z_2 u.a., $\sim \mathcal{N}_{0,1}$, dann haben nach Beob. 1.42

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} aZ_1 + bZ_2 \\ -bZ_1 + aZ_2 \end{pmatrix}$$

dieselbe Verteilung, d.h. auch $aZ_1 + bZ_2$ und $-bZ_1 + aZ_2$ sind u.i.v., $\sim \mathcal{N}_{0,1}$, insbesondere ist $aZ_1 + bZ_2$ standard-normalverteilt.

Zum Beweis von Bsp. 1.59:

Setzen wir $a := \frac{\sigma_1}{\sqrt{\sigma_1^2+\sigma_2^2}}$, $b := \frac{\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$, so finden wir:

$X_1 := \sigma_1 Z_1 \sim \mathcal{N}_{0,\sigma_1^2}$, $X_2 := \sigma_2 Z_2 \sim \mathcal{N}_{0,\sigma_2^2}$ (und X_1, X_2 sind u.a.),

$$\frac{X_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} + \frac{X_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} = aZ_1 + bZ_2 \sim \mathcal{N}_{0,1},$$

also gilt $X_1 + X_2 \sim \mathcal{N}_{0,\sigma_1^2+\sigma_2^2}$.

(Man kann – anstelle von Beob. 1.42 – in diesem Fall auch das Faltungsintegral explizit ausrechnen, vgl. Notizen)