Inhalt von Abschnitt 6

Bei der Durchführung von Zufallsexperimenten interessiert man sich in der Regel für zwei oder mehrere Variablen, die für dieselben Untersuchungseinheiten erfasst werden.

Es ist dann notwendig, die gemeinsame Verteilung dieser Merkmale zu betrachten. Im Folgenden wird zuerst das Konzept mehrdimensionaler Zufallsvariablen eingeführt und anschließend der zweidimensionale Fall für diskrete und stetige Zufallsvariablen eingehender betrachtet.

Ein wesentlicher Abschnitt gilt der Kovarianz und der Korrelation als Verteilungsparametern, die den Zusammenhang zwischen je zwei Zufallsvariablen charakterisieren. Wichtig ist außerdem das Konzept der bedingten Verteilung.

Zweidimensionale diskrete Zufallsvariablen

Definition 6.1 (Gemeinsame Wahrscheinlichkeitsfunktion)

Die Wahrscheinlichkeitsfunktion der bivariaten diskreten Zufallsvariable (X, Y) ist bestimmt durch

$$f(x,y) = \begin{cases} P(X = x, Y = y) & \text{für } (x,y) \in \mathcal{T}_X \times \mathcal{T}_Y, \\ 0 & \text{sonst,} \end{cases}$$

$\mathcal{T}_X \times \mathcal{T}_Y = \{(x_i, y_j) : x_i \in \mathcal{T}_X, y_j \in \mathcal{T}_Y\}$. Wir bezeichnen die Wahrscheinlichkeitsfunktion auch als (gemeinsame) diskrete Dichte oder (gemeinsame) Verteilung.

In der Wahrscheinlichkeitsfunktion ist die gesamte Information des Zufallsexperiments in Bezug auf die Merkmale X, Y enthalten. Besitzen X und Y jeweils nur endlich viele Ausprägungen, lässt sich die gemeinsame Wahrscheinlichkeitsfunktion übersichtlich in Kontingenztafeln zusammenfassen.
Besitze X die möglichen Ausprägungen $x_1, \ldots, x_I$ und Y die möglichen Ausprägungen $y_1, \ldots, y_J$. Bezeichne

$$p_{ij} = P(X = x_i, Y = y_j) = f(x_i, y_j).$$

- Kontingenztafel der Wahrscheinlichkeiten

$$\begin{array}{c|ccc|}
 & y_1 & \cdots & y_J \\
\hline
x_1 & p_{11} & \cdots & p_{1J} \\
\vdots & \vdots & \ddots & \vdots \\
x_I & p_{I1} & \cdots & p_{IJ} \\
\hline
\end{array}$$

Definition 6.2 (Randverteilungen)

Die Randverteilung von X ist gegeben durch

$$f_X(x) = P(X = x) = \sum_{\text{alle } y_j} f(x, y_j),$$

die Randverteilung von Y durch

$$f_Y(y) = P(Y = y) = \sum_{\text{alle } x_i} f(x_i, y).$$

Kontingenztafel der Wahrscheinlichkeiten

Die $(I \times J)$ Kontingenztafel der Wahrscheinlichkeiten hat die Form

$$\begin{array}{c|ccc|c}
 & y_1 & \cdots & y_J & \\
\hline
x_1 & p_{11} & \cdots & p_{1J} & p_{1\cdot} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
x_I & p_{I1} & \cdots & p_{IJ} & p_{I\cdot} \\
\hline
 & p_{\cdot 1} & \cdots & p_{\cdot J} & 1
\end{array}$$

Dabei bezeichnen für $i = 1, \ldots, I$, $j = 1, \ldots, J$,

$p_{ij} = P(X = x_i, Y = y_j)$ die Wahrscheinlichkeiten für (x_i, y_j),

$p_{i\cdot} = \sum_{j=1}^{J} p_{ij}$ die (Rand-) Wahrscheinlichkeiten für x_i,

$p_{\cdot j} = \sum_{i=1}^{I} p_{ij}$ die (Rand-) Wahrscheinlichkeiten für y_j.

Gemeinsame Verteilungsfunktion

Die Verallgemeinerung des Konzeptes der Verteilungsfunktion auf zwei Variablen führt zu einer zweidimensionalen Funktion. Die gemeinsame Verteilungsfunktion von X und Y ist gegeben durch

$$F(x,y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} f(x_i, y_j)$$

$$\Longrightarrow \quad F_X(x) = P(X \leq x) = F(x, \infty) = \sum_{x_i \leq x} \sum_{y_j} f(x_i, y_j)$$

$$F_Y(x) = P(Y \leq y) = F(\infty, y) = \sum_{x_i} \sum_{y_j \leq y} f(x_i, y_j).$$

Bedingte Verteilungen

Wenn man die gemeinsame Verteilung der ZVn X und Y kennt, kann man einfach ableiten, wie eine der ZVn verteilt ist, wenn man die Ausprägung der anderen kennt.

Unter der bedingten Verteilung von Y gegeben $X = x$, kurz $Y \mid X = x$, versteht man die Verteilung der Zufallsvariable Y, wenn bekannt ist, dass das Ereignis $\{X = x\}$ eingetreten ist.

Es gilt also bei festgehaltenem x zu bestimmen, mit welcher Wahrscheinlichkeit die Werte $y_1, y_2, \ldots$ unter dieser Voraussetzung auftreten.

Definition 6.3 (Bedingte Wahrscheinlichkeitsfunktion)

Die bedingte Wahrscheinlichkeitsfunktion von Y gegeben $X = x$ ist (für festes x und $f_X(x) > 0$) bestimmt durch

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

Für $f_X(x) = 0$ legt man $f_{Y|X=x}(y) = 0$ fest.

Die bedingte Wahrscheinlichkeitsfunktion von $X \mid Y = y$ ist analog definiert.

- $f_{Y|X=x}(y)$ ist bei festgelegtem x eine Dichte, d.h. $\sum_{y_j} f_{Y|X=x}(y_j) = 1$, analog für $f_{X|Y=y}(x)$.

Bedingter Erwartungswert und Varianz

Gegeben die bedingte W-Verteilung einer numerischen Variable, ist es natürlich, nach Erwartungswert und Varianz dieser Verteilung zu fragen:

Definition 6.4 (Bedingter Erwartungswert)

Die reelle Zahl

$$\mathbb{E}(Y \mid X = x) = \sum_{y \in \mathcal{T}_Y} y \cdot f_{Y|X=x}(y)$$

für $x \in \mathcal{T}_x$ heißt bedingter Erwartungswert von Y gegeben $X = x$.

Bedingter Erwartungswert und Varianz (Forts.)

Definition 6.5 (Bedingte Varianz)

Die reelle Zahl

$$\begin{aligned}\mathbb{V}(Y \mid X = x) &= \mathbb{E}([Y - \mathbb{E}(Y \mid X = x)]^2 \mid X = x) \\ &= \sum_{y \in \mathcal{T}_Y} [y - \mathbb{E}(Y \mid X = x)]^2 \cdot f_{Y \mid X = x}(y)\end{aligned}$$

für $x \in \mathcal{T}_x$ heißt bedingte Varianz von Y gegeben $X = x$.

Zweidimensionale stetige Zufallsvariablen

Definition 6.6 (Gemeinsame stetige Verteilung und Dichte zweier Zufallsvariablen)

Die Zufallsvariablen X und Y sind gemeinsam stetig verteilt, wenn es eine zweidimensionale Dichtefunktion $f(x, y) \geq 0$ gibt, so dass

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

Aus dieser Formel gewinnt man auch diejenige für die gemeinsame Verteilungsfunktion von X und Y: setze $a = -\infty, \; c = -\infty$.

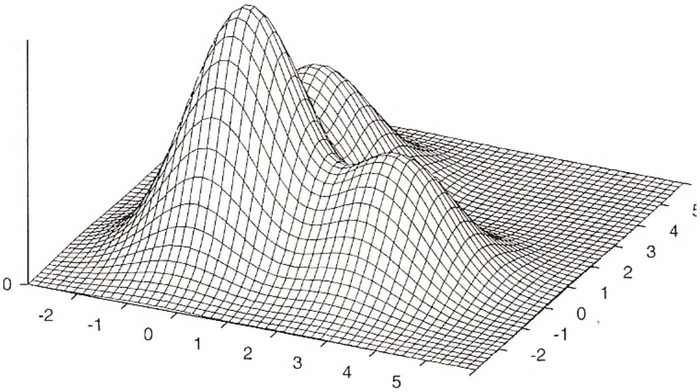


ABBILDUNG 8.2: Form einer zweidimensionalen Dichte $f(x, y)$

Figure: aus Fahrmeir et al.

Definition 6.7 (Randdichten)

Die Randdichte von X ist gegeben durch

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y,$$

die Randdichte von Y durch

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x.$$

Bedingte Verteilungen

Definition 6.8 (Bedingte Dichten)

Die bedingte Dichte von Y unter der Bedingung $X = x$, kurz $Y \mid X = x$, ist für festen Wert x und $f_X(x) \neq 0$ bestimmt durch

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

Für $f_X(x) = 0$ legt man $f_{Y|X=x}(y) = 0$ fest.
Die bedingte Dichte von $X \mid Y = y$ ist analog definiert.

- Für einen festen Wert $X = x$ ist $f_{Y|X=x}(y)$ wiederum eine stetige Dichte, d.h. $\int f_{Y|X=x}(y)\mathrm{d}y = 1$.

Bedingte Verteilungen (Forts.)

- Kennt man Randdichte f_X und bedingte Dichte $f_{Y|X=x}$, so lässt sich daraus die gemeinsame Dichte $f(x, y)$ berechnen via

$$f(x, y) = f_{Y|X=x}(y) \cdot f_X(x).$$

- Bsp. Regressionsanalysen: Man hat ein "Modell" für die Verteilung einer Responsevariable Y, gegeben Werte x einer Kovariable X, man hat also $f_{Y|X=x}(y)$ spezifiziert. Kennt man die Verteilung von X, so kann man die gemeinsame Verteilung von X und Y bestimmen.

Bedingter Erwartungswert und Varianz

Für stetige ZVn X und Y definiert man bedingten Erwartungswert und Varianz, gegeben einen konkreten Wert der anderen ZV, analog zum diskreten Fall.

Definition 6.9 (Bedingter Erwartungswert)

Seien X und Y stetige ZVn mit gemeinsamer Dichtefunktion. Dann heißt die Zahl

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{+\infty} y \cdot f_{Y|X=x}(y) \mathrm{d}y$$

bedingter Erwartungswert von Y gegeben $X = x$.
Der bedingte Erwartungswert von $X \mid Y = y$ ist analog definiert.

Bedingter Erwartungswert und Varianz (Forts.)

Definition 6.10 (Bedingte Varianz)

Die reelle Zahl

$$\begin{aligned}\mathbb{V}(Y \mid X = x) &= \mathbb{E}([Y - \mathbb{E}(Y \mid X = x)]^2 \mid X = x) \\ &= \int_{-\infty}^{+\infty} [y - \mathbb{E}(Y \mid X = x)]^2 \cdot f_{Y|X=x}(y)\mathrm{d}y\end{aligned}$$

heißt bedingte Varianz von $Y \mid X = x$.

- Beispiele für solche Settings sind
  - zweidimensionale Normalverteilung (s.u.)
  - lineare Regressionsmodelle

Unabhängigkeit von Zufallsvariablen

Für diskrete wie stetige ZVn definiert man:

Definition 6.11 (Unabhängigkeit von zwei Zufallsvariablen)

Die Zufallsvariablen X und Y heißen unabhängig, wenn für alle x und y gilt

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

Ansonsten heißen X und Y abhängig.

Unabhängigkeit von Zufallsvariablen (Forts.)

- Aus $f(x,y) = f_X(x) f_Y(y)$ folgt

$$\begin{aligned} P(X \leq x, Y \leq y) &= \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v) \mathrm{d}v \mathrm{d}u \\ &= \int_{-\infty}^{x} f_X(u) \mathrm{d}u \int_{-\infty}^{y} f_Y(v) \mathrm{d}v \\ &= P(X \leq x) \cdot P(Y \leq y). \end{aligned}$$

- Aus der Unabhängigkeit folgt gleichfalls, dass

$$f_{Y|X=x}(y) = f_Y(y) \text{ sowie } f_{X|Y=y}(x) = f_X(x).$$

Der Begriff der Unabhängigkeit von Zufallsvariablen lässt sich auf mehr als zwei Zufallsvariablen erweitern.

Analog zum Fall zweier ZVn lassen sich die ZVn $X_1, \ldots, X_n$ als unabhängig verstehen, wenn die bedingte Verteilung jeder dieser ZVn, gegeben die Ausprägungen der übrigen $n-1$ ZVn, nicht von diesen Ausprägungen abhängt.

Dies führt zur einfachen Darstellung der gemeinsamen Dichte als Produkt der Randdichten.

Definition 6.12 (Unabhängigkeit von Zufallsvariablen)

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen unabhängig, wenn für alle $x_1, \ldots, x_n$ gilt

$$P(X_1 \le x_1, \ldots, X_n \le x_n) = P(X_1 \le x_1) \cdot \ldots \cdot P(X_n \le x_n).$$

Äquivalent dazu ist die Produktbedingung

$$f(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n),$$

wobei $f(x_1, \ldots, x_n)$ die gemeinsame Dichte von $X_1, \ldots, X_n$ und $f_{X_i}(x_i)$ die Dichte der Zufallsvariable X_i bezeichnen $(i = 1, \ldots, n)$.

Kovarianz und Korrelation

Definition 6.13 (Kovarianz)

Die Kovarianz der Zufallsvariablen X und Y ist bestimmt durch

$$\mathrm{Cov}(X,Y) = \mathbb{E}([X - E(X)][Y - E(Y)])$$

$$= \begin{cases} \displaystyle\sum_i \sum_j f(x_i, y_j)(x_i - \mathbb{E}(X))(y_j - \mathbb{E}(Y)) \\ \qquad X \text{ und } Y \text{ diskret,} \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)(x - \mathbb{E}(X))(y - \mathbb{E}(Y))\mathrm{d}x\mathrm{d}y \\ \qquad X \text{ und } Y \text{ stetig.} \end{cases}$$

- Verschiebungssatz

$$\mathrm{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

- Symmetrie

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

- Lineare Transformation
  Die Kovarianz der transformierten Zufallsvariablen $\widetilde{X} = a_X X + b_X$, $\widetilde{Y} = a_Y Y + b_Y$ ist bestimmt durch

$$\mathrm{Cov}(\widetilde{X}, \widetilde{Y}) = a_X a_Y \,\mathrm{Cov}(X, Y).$$

Definition 6.14 (Korrelationskoeffizient)

Der Korrelationskoeffizient ist definiert als

$$\rho = \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)}} = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y}.$$

- Sein Wertebereich ist

$$-1 \leq \rho(X,Y) \leq 1.$$

- Es lässt sich zeigen, dass
  $|\rho(X,Y)| = 1$ genau dann gilt, wenn Y eine lineare Transformation von X ist, d.h. $Y = aX + b$ für Konstanten a, b gilt. Wenn $a > 0$ ist, gilt $\rho(X,Y) = 1$, und wenn $a < 0$ ist, gilt $\rho(X,Y) = -1$.

Definition 6.15 (Unkorreliertheit)

Zwei Zufallsvariablen X und Y heißen unkorreliert, wenn gilt

$$\rho(X,Y) = 0.$$

Wenn $\rho(X,Y) \neq 0$ gilt, heißen sie korreliert.

Satz 6.16 (Unabhängigkeit und Korrelation)

Sind zwei Zufallsvariablen unabhängig, so sind sie auch unkorreliert, d.h. es gilt $\rho(X,Y) = 0$.

- Kovarianz und Korrelation sind Maße für den linearen Zusammenhang von Zufallsvariablen.

- Varianz der Summe zweier Zufallszahlen

Für die Varianz einer Summe von ZVn gilt

$$\mathbb{V}(X_1 + X_2) = \mathbb{V}(X_1) + \mathbb{V}(X_2) + 2\,\mathrm{Cov}(X_1, X_2).$$

Für n ZVn $X_1, \ldots, X_n$ lässt sich jeweils paarweise die Kovarianz $\mathrm{Cov}(X_i, X_j)$ bzw. die Korrelation $\rho(X_i, X_j)$ betrachten.

Anstatt der einfachen Summe $X_1 + \cdots + X_n$ betrachte man allgemeiner die gewichtete Summe $X = a_1 X_1 + \cdots + a_n X_n$, dann gilt

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}([X - \mathbb{E}(X)]^2) \\ &= \sum_{i=1}^{n} a_i^2\,\mathbb{V}(X_i) + 2\sum_{i<j} a_i a_j\,\mathrm{Cov}(X_i, X_j).\end{aligned}$$

Sind alle Variablen unkorreliert, d.h. gilt $\rho(X_i, X_j) = 0$ und damit $\mathrm{Cov}(X_i, X_j) = 0$, erhält man die einfache Form

$$\mathbb{V}(X) = \sum_{i=1}^{n} a_i^2 \, \mathbb{V}(X_i).$$

Satz 6.17 (Erwartungswert und Varianz von Linearkombinationen)

Die gewichtete Summe $X = a_1 X_1 + \dots + a_n X_n$ der Zufallsvariablen $X_1, \dots, X_n$ besitzt den Erwartungswert

$$\mathbb{E}(X) = a_1\,\mathbb{E}(X_1) + \dots + a_n\,\mathbb{E}(X_n)$$

und die Varianz

$$\begin{aligned}\mathbb{V}(X) =& a_1^2\,\mathbb{V}(X_1) + \dots + a_n^2\,\mathbb{V}(X_n) \\ & + 2a_1a_2\,\mathrm{Cov}(X_1, X_2) + 2a_1a_3\,\mathrm{Cov}(X_1, X_3) + \dots \\ =& \sum_{i=1}^{n} a_i^2\,\mathbb{V}(X_i) + 2\sum_{i<j} a_i a_j\,\mathrm{Cov}(X_i, X_j).\end{aligned}$$

Odds Ratio

Ein anderes wichtiges Assoziationsmaß, speziell für binäre ZVn, ist das Odds Ratio. Es wird über die bedingten Verteilungen definiert.

Definition 6.18 (Odds, Odds Ratio)

Für eine Wahrscheinlichkeit $0 < \pi < 1$ ist $\mathrm{Odds}(\pi)$ definiert als

$$\mathrm{Odds}(\pi) = \pi/(1-\pi).$$

Für binäre ZVn X und Y ist das Odds Ratio ψ_{YX} definiert als

$$\psi_{YX} = \frac{\mathrm{Odds}(P(Y=1 \mid X=1))}{\mathrm{Odds}(P(Y=1 \mid X=0))}$$

$$= \frac{P(Y=1 \mid X=1)}{P(Y=0 \mid X=1)} \Bigg/ \frac{P(Y=1 \mid X=0)}{P(Y=0 \mid X=0)}.$$

Satz 6.19

Seien X und Y binäre ZVn. Dann sind äquivalent:

(a) *X und Y sind stochastisch unabhängig,*

(b) $P(Y = 1 \mid X = 1) = P(Y = 1 \mid X = 0)$,

(c) *Für das Odds Ratio gilt: $\psi_{YX} = 1$.*

- Das Odds Ratio spielt vor allem in der Epidemiologie eine große Rolle.

Definition 6.20 (Zweidimensionale Normalverteilung)

Die Zufallsvariablen X und Y heißen gemeinsam normalverteilt, wenn die Dichte bestimmt ist durch

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{Q}{2(1-\rho^2)}\right\},$$

wobei

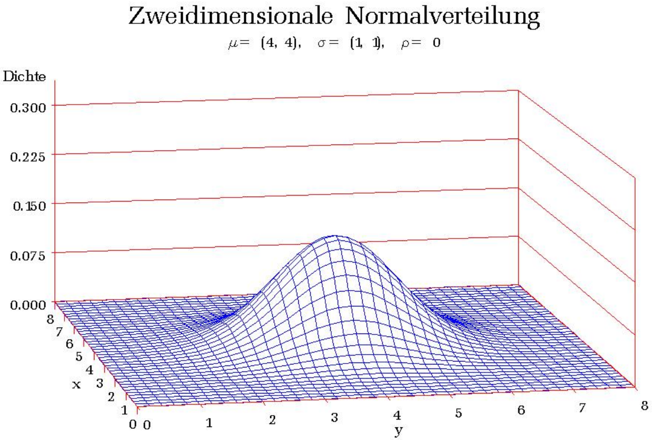
$$Q = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2,$$

$\mu_X = \mathbb{E}(X)$, $\mu_Y = \mathbb{E}(Y)$ und $\sigma_X^2 = \mathbb{V}(X)$, $\sigma_Y^2 = \mathbb{V}(Y)$.

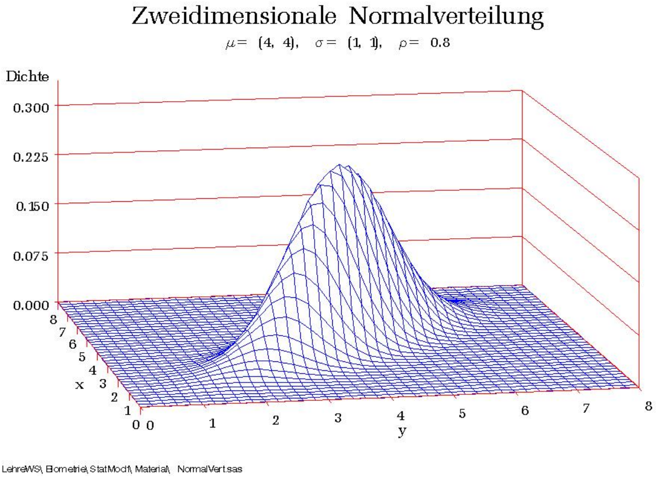
- Durch die quadratische Form Q werden Ellipsen $Q \equiv c$ mit Mittelpunkt (μ_X, μ_Y) als Höhenlinie $Q = c$ definiert. Ihre Form hängt von σ_X, σ_Y und ρ ab.
- Für $\rho = 0$ sind die Achsen parallel zu den Koordinatenachsen; für $\sigma_X > \sigma_Y$ ist die größere Achse parallel zur x-Achse. Für $\rho \neq 0$ sind die Achsen nicht mehr parallel zu den Koordinatenachsen.
- Für $\rho = 0$ folgt, dass $f(x, y) = f(x) \cdot f(y)$ mit $f(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)$, $f(y)$ entsprechend, d.h. X und Y sind unabhängig.

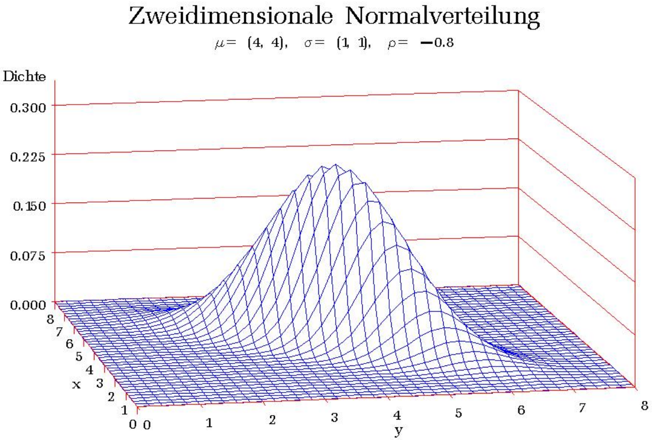
Satz 6.21 (Unabhängigkeit und Korrelation bei normalverteilten Zufallsvariablen)

Für gemeinsam normalverteilte Zufallsvariablen X und Y gilt: X und Y sind unabhängig genau dann, wenn sie unkorreliert sind.



LehreWS\ Biometrie\ StatMod1\ Material\ NormalVert.sas


Zweidimensionale Normalverteilung
$\mu =$ (4, 4), $\sigma =$ (1, 1), $\rho =$ 0.8
Dichte
0.300
0.225
0.150
0.075
0.000
x
0
1
2
3
4
5
6
7
8
y
0
1
2
3
4
5
6
7
8
LehreWS\ Biometrie\ StatModI\ Material\ NormalVert.sas



LehreWS\ Biometrie\ StatMod1\ Material\ NormalVert.sas

Randverteilungen

- Sind X und Y gemeinsam normalverteilt mit oben angegebener Dichte, so sind die Randverteilungen jeweils normal mit

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2),$$

$$Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2).$$

Bedingte Verteilung von Y gegeben $X = x$

Ist (X, Y) zweidimensional normalverteilt, so ist $Y|X = x$ normalverteilt mit Dichte

$$f_{Y|X=x}(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y \sqrt{1-\rho^2}} \cdot \exp\left\{ -\frac{1}{2\sigma_Y^2(1-\rho^2)} \left[y - \mu_Y - \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) \right]^2 \right\}.$$

Es ist

$$\mathbb{E}(Y \mid X = x) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X)$$
$$\mathbb{V}(Y \mid X = x) = \sigma_Y^2(1 - \rho^2).$$

Zweidimensionale gemischte Zufallsvariablen

Beispiele für bivariate ZVn mit einer diskreten und einer stetigen Komponente, sind

- (Geschlecht, Lebensdauer)
  Lebensdauer ist eine stetige ZV; die Lebensdauern für die zwei Geschlechter sind bekanntlich verschieden verteilt;

- Poisson-Gamma-Mischmodell (s.u.)
  Hier hat man eine $Po(\lambda)-$verteilte Zählvariable Y, der steuernde Parameter λ entstammt seinerseits einer Verteilung. Als Beispiel: Verteilung kariöser Zähne in österreichischen Kindern;

- Eine Situation wie in der unten abgebildeten drei-gipfeligen Verteilung: eine solche Situation könnte entstehen, wenn sich die Bevölkerung aus drei Subpopulationen zusammensetzt und die

Zweidimensionale gemischte Zufallsvariablen (Forts.)

Erwartungswerte des untersuchten Merkmals X in den Subpopulationen deutlich verschieden sind.

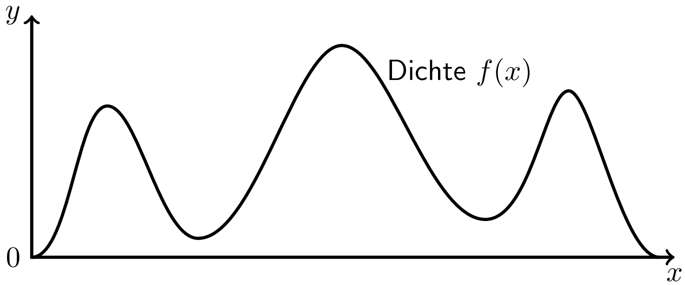


Figure: Drei-modale Wahrscheinlichkeitsdichte

Poisson-Gamma-Mischmodell

Satz 6.22 (Poisson-Gamma-Mischmodell)

Sei $\mathcal{L}(Y \mid \Lambda = \lambda) = Po(\lambda)$ und sei der Parameter Λ selbst $\Gamma(\nu, \alpha)-$verteilt. Dann gilt für die Randverteilung Y:
Y ist negativ binomialverteilt mit

$$Y \sim NB(\nu, p = \tfrac{\nu}{\nu+\mu}) \text{ mit } \mu = \tfrac{\nu}{\alpha}.$$

Insbesondere ist also $\mathbb{E}(Y) = \mu = \frac{\nu}{\alpha}$, das ist der E-Wert der mischenden Gammaverteilung, und $\mathbb{V}(Y) = \mu + \frac{\mu^2}{\nu}$.

Bedingte Erwartung und Varianz

Indem man in obiger Definition von bedingtem Erwartungswert und bedingter Varianz die Realisierung unbestimmt lässt, also 'X' statt '$X = x$' setzt, definiert man die Zufallsvariablen $\mathbb{E}(Y \mid X)$ und $\mathbb{V}(Y \mid X)$.

Definition 6.23 (Bedingte Erwartung und Varianz)

Seien X und Y ZVn mit gemeinsamer Dichte f_{XY}.Dann heißen die ZVn

$$\mathbb{E}(Y \mid X) \text{ und } \mathbb{V}(Y \mid X)$$

bedingte Erwartung und bedingte Varianz von Y gegeben X.

Satz 6.24

Unter obigen Voraussetzungen gilt:

(a) *(Satz vom iterierten Erwartungswert)*

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X)).$$

(b) *(Varianzzerlegung)*

$$\mathbb{V}(Y) = \mathbb{E}(\mathbb{V}(Y \mid X)) + \mathbb{V}(\mathbb{E}(Y \mid X)).$$

- Damit kann z. B. im stetigen Fall $\mathbb{E}(Y)$ berechnet werden via $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X)) = \int \mathbb{E}(Y \mid X = x) f_X(x) dx$.