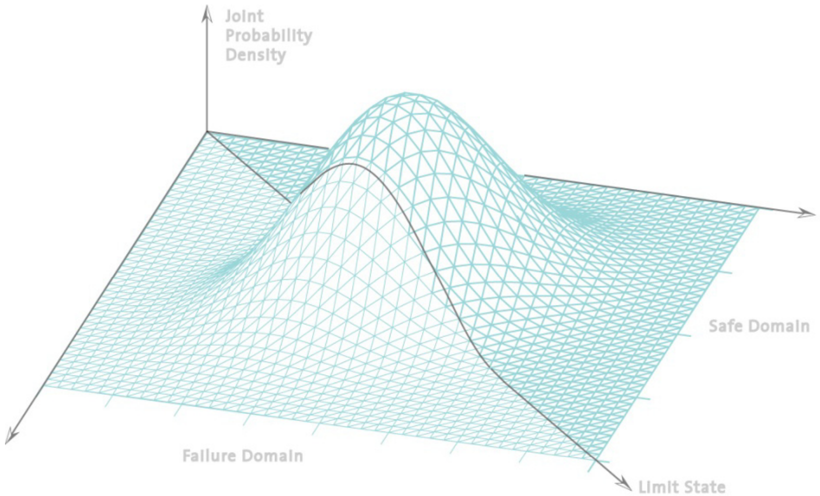


Statistik und Wahrscheinlichkeitsrechnung

5. Vorlesung

Prof. Dr. Michael Havbro Faber

Inhalt

- Modellierung von Unsicherheiten – Übersicht
- Zufallsvariablen
  - Eigenschaften des Erwartungswertoperators
  - Zufallsvektoren und Produktmomente
  - bedingte Verteilungen und bedingte Momente
  - Wahrscheinlichkeitsverteilung für die Summe zweier Zufallsvariablen
  - Wahrscheinlichkeitsverteilung für Funktionen von Zufallsvariablen

Modellierung von Unsicherheiten - Übersicht

- Zufallsvariablen und deren Charakteristika

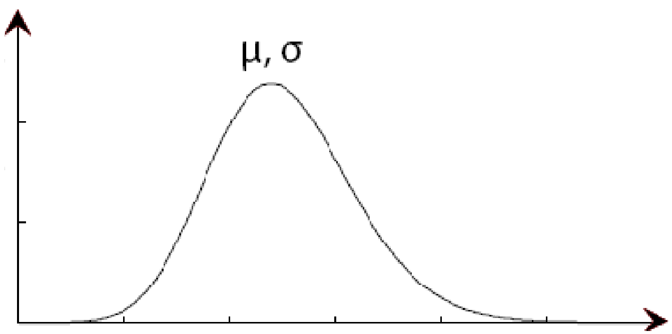
REALITÄT

UNSICHERHEIT

↓

DATEN / BEOBACHTUNGEN

MODELL



Zufallsvariablen

Zufallsvariablen

- Eigenschaften des Erwartungswertoperators

  Der Erwartungswertoperator ermöglicht die Berechnung des Erwartungswertes und der Varianz einer Zufallsvariablen.

  Wenn wir verstehen, wie der Erwartungswertoperator funktioniert, können wir den Erwartungswert und die Varianz von Funktionen von Zufallsvariablen berechnen.

  Dies dient im Besonderen der Analyse von Modellen im Ingenieurwesen mit einer oder mehreren Zufallsvariablen.

  BEISPIEL: Berechnung der Gesamtdauer eines Bauvorhabens als Funktion der Dauer der einzelnen Bauabschnitte.

Zufallsvariablen

- Eigenschaften des Erwartungswertoperators

Der Erwartungswertoperator besitzt folgende Eigenschaften:

$a,b,c = \text{Konstante}$

$X = \text{Zufallsvariable}$

$g_i(.) = \text{Funktion}$

$$E[c] = c$$

$$E[cX] = cE[X]$$

$$E[a + bX] = a + bE[X]$$

$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$$

Zufallsvariablen

- Eigenschaften des Varianzoperators

Die Varianz kann wie folgt beschrieben werden:

$$
\begin{aligned}
Var\left[X\right] &= E\left[(X-\mu_X)^2\right] \\
&= E\left[X^2+\mu_X^2-2\mu_X X\right] \\
&= \mu_X^2+E\left[X^2\right]-2\mu_X E\left[X\right] \\
&= \mu_X^2+E\left[X^2\right]-2\mu_X^2 = E\left[X^2\right]-\mu_X^2
\end{aligned}
$$

Zufallsvariablen

- Eigenschaften des Varianzoperators

Darüber hinaus gilt:

$$Var[c] = 0$$
$$Var[cX] = c^2 Var[X]$$
$$Var[a + bX] = b^2 Var[X]$$

$$E[c] = c$$
$$E[cX] = cE[X]$$
$$E[a + bX] = a + bE[X]$$
$$E[g_1(X) + g_2(X)] =$$
$$E[g_1(X)] + E[g_2(X)]$$

Zufallsvariablen

- Eigenschaften des Erwartungswertoperators

Anhand des Ergebnisses

$$Var[X] = E\left[(X-\mu_X)^2\right] = E\left[X^2 + \mu_X^2 - 2\mu_X X\right] = E\left[X^2\right] - \mu_X^2$$

ist erkennbar, dass grundsätzlich gilt:

$$E[g(X)] \neq g(E[X])$$

für konvexe / konkave Funktionen – JENSEN'S Ungleichheit!!

Gleichheit gilt ausschließlich für lineare Funktionen!

Zufallsvariablen

- Zufallsvektoren

Wir beschäftigen uns häufig mit Modellen, die nicht nur eine, sondern mehrere Zufallsvariablen beinhalten.

Diese Zufallsvariablen können in einem Vektor vereint werden.

Generell sind die Bestandteile eines Vektors voneinander abhängig.
BEISPIEL: Regenwetter und Wasserpegel.

Folglich ist es notwendig, dass wir probabilistische Modelle erstellen, welche diese Abhängigkeit berücksichtigen.
Dies kann anhand von multivariaten kumulativen Verteilungsfunktionen bewerkstelligt werden.

Zufallsvariablen

- Zufallsvektoren

Wir betrachten nun nicht mehr nur eine diskrete Zufallsvariable, sondern einen Vektor von mehreren diskreten Zufallsvariablen; hier z.B. zwei:

$$\mathbf{Z} = (X, Y)^T$$

Die bivariate Wahrscheinlichkeitsdichtefunktion ist:

$$p_{\mathbf{Z}}(\mathbf{z}) = p_{X,Y}(x, y) = P\big((X = x) \cap (Y = y)\big)$$

Die bivariate kumulative Verteilungsfunktion ist:

$$P_{\mathbf{Z}}(\mathbf{z}) = P_{X,Y}(x, y) = P\big((X \le x) \cap (Y \le y)\big) = \sum_{x_i \le x} \sum_{y_i \le y} p_{X,Y}(x_i, y_i)$$

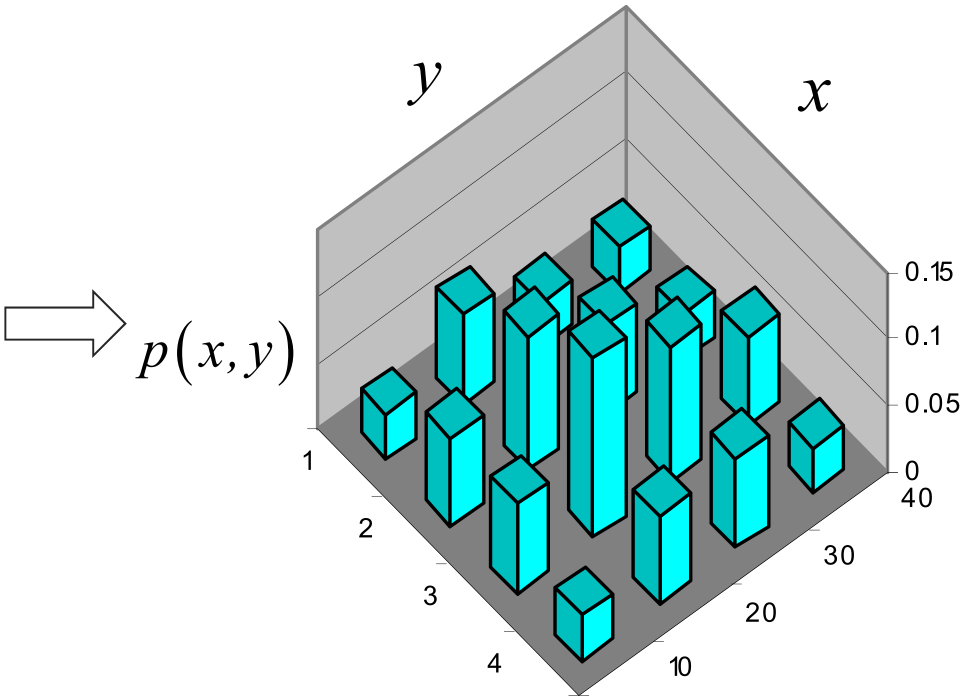
Zufallsvariablen

- Zufallsvektoren

x,y	p(x,y)
1,10	0.033
1,20	0.067
1,30	0.033
1,40	0.033
2,10	0.067
2,20	0.100
2,30	0.067
2,40	0.033
3,10	0.067
3,20	0.133
3,30	0.100
3,40	0.067
4,10	0.033
4,20	0.067
4,30	0.067
4,40	0.033

$\Sigma=1$

Gegeben sei eine zweidimensionale diskrete Wahrscheinlichkeitsdichtefunktion.



Zufallsvariablen

- Zufallsvektoren

Die marginale Wahrscheinlichkeitsdichtefunktion einer diskreten Zufallsvariablen X ist wie folgt definiert:

$$p_X(x) \equiv P(X = x) = \sum_{alle\ y_i} p_{X,Y}(x, y_i)$$

Zufallsvariablen

- Zufallsvektoren

Die marginale Wahrscheinlichkeitsverteilung einer diskreten Zufallsvariablen X ist wie folgt definiert:

$$p_X(x) \equiv P(X = x) = \sum_{alle\ y_i} p_{X,Y}(x, y_i)$$

marginale Wahrscheinlichkeitsdichte

$$P_X(x) \equiv P(X \le x) = \sum_{x_i \le x} p_X(x_i)$$

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

marginale Wahrscheinlichkeits-verteilung

Zufallsvariablen

- Zufallsvektoren

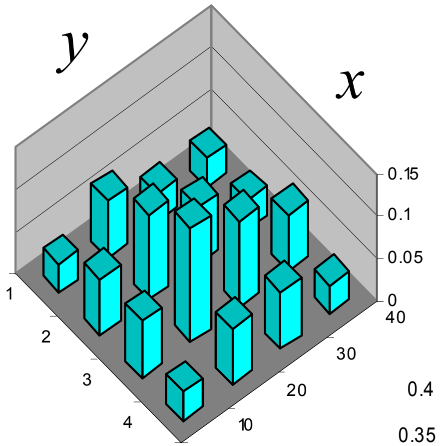
x,y	p(x,y)	
1,10	0.033	
1,20	0.067	
1,30	0.033	
1,40	0.033	Σ=0.17
2,10	0.067	
2,20	0.100	
2,30	0.067	
2,40	0.033	Σ=0.27
3,10	0.067	
3,20	0.133	
3,30	0.100	
3,40	0.067	Σ=0.37
4,10	0.033	
4,20	0.067	
4,30	0.067	
4,40	0.033	Σ=0.20

Σ=1

Gegeben sei eine zweidimensionale diskrete Wahrscheinlichkeitsdichtefunktion:

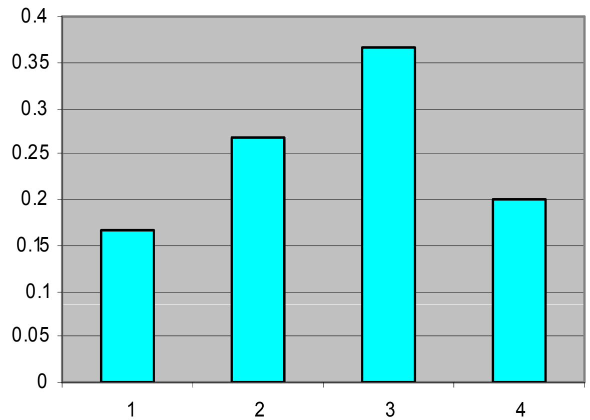
y x

$p(x,y)$



Diskrete multivariate Dichte

$p_x(x)$



x

marginale Dichte für Zufallsvariable X

Zufallsvariablen

- Zufallsvektoren

Wir betrachten nun nicht mehr nur eine kontinuierliche Zufallsvariable, sondern einen Vektor von mehreren kontinuierlichen Zufallsvariablen; hier z.B. zwei:

$$\mathbf{Z} = (X, Y)^T$$

Die bivariate kumulative Verteilungsfunktion ist:

$$F_{\mathbf{Z}}(\mathbf{z}) = F_{X,Y}(x, y) = P(X \leq x \cap Y \leq y)$$

Die bivariate Wahrscheinlichkeitsdichtefunktion ist:

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

Zufallsvariablen

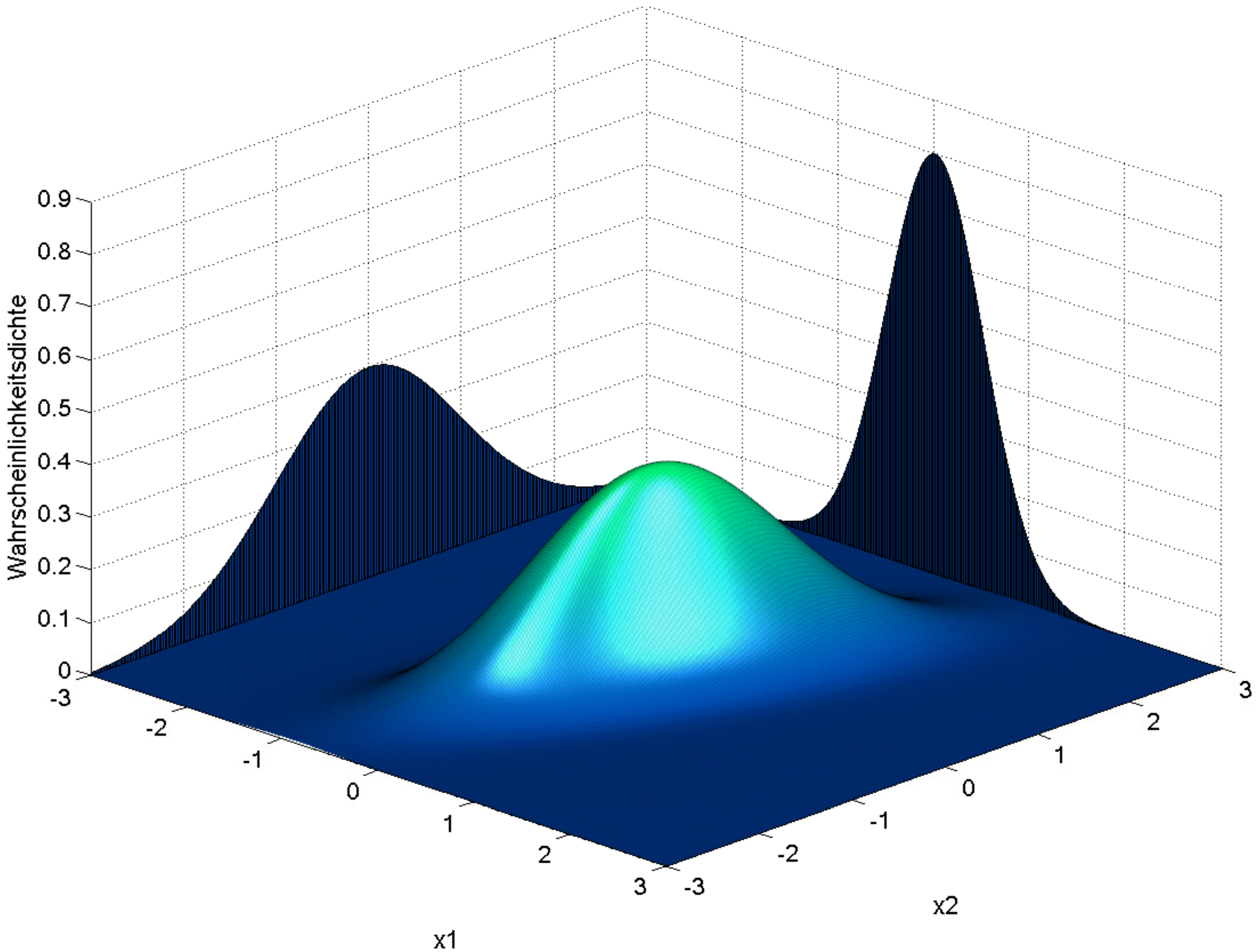
- Zufallsvektoren

Die marginale Wahrscheinlichkeitsdichtefunktion einer kontinuierlichen Zufallsvariablen X in einem bivariaten Zufallsvektor (X,Y) ist wie folgt definiert:

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

Zufallsvariablen

- Bivariate Wahrscheinlichkeitsdichtefunktion



Zufallsvariablen

- Zufallsvektoren und multivariate Momente

Die Kovarianz zwischen dem i-ten und j-ten Element eines Zufallsvektors kontinuierlicher Zufallsvariablen heisst zentrales multivariates Moment.

$$C_{X_iX_j} = E\left[(X_i - \mu_{X_i})(X_j - \mu_{X_j})\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(x_i - \mu_{X_i}\right)\left(x_j - \mu_{X_j}\right) f_{X_iX_j}\left(x_i, x_j\right) dx_i dx_j$$

Daraus ist unter der Bedingung $i=j$ die Varianz für X_i ersichtlich:

$$C_{X_iX_i} = Var\left[X_i\right]$$

Der Korrelationskoeffizient errechnet sich wie folgt:

$$\rho_{X_iX_j} = \frac{C_{X_iX_j}}{\sigma_{X_i}\sigma_{X_j}} \qquad \rho_{X_iX_i} = 1$$

Zufallsvariablen

- Zufallsvektoren und multivariate Momente

Der Erwartungswert und die Varianz einer linearen Funktion

$$Y = a_0 + \sum_{i=1}^{n} a_i X_i$$

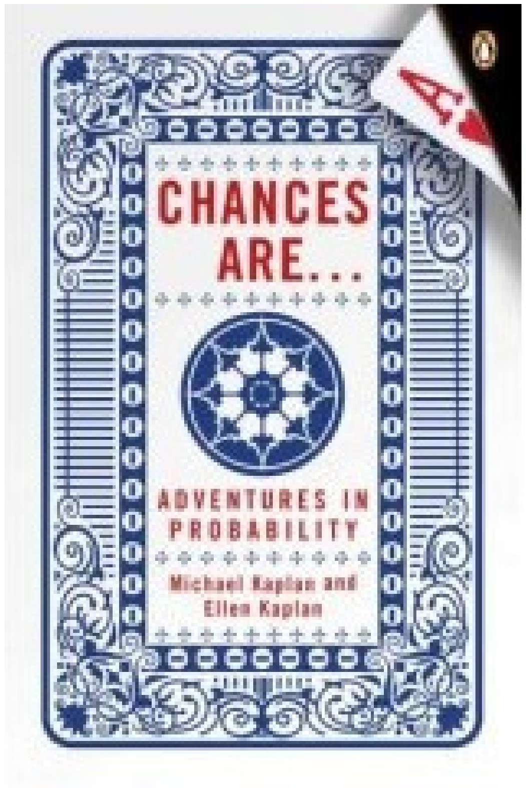
werden bestimmt durch:

$$E[Y] = a_0 + \sum_{i=1}^{n} a_i E[X_i]$$

$$Var[Y] = \sum_{i=1}^{n} a_i^2 Var[X_i] + 2\left(\sum_{\substack{i,j=1 \\ i \neq j}}^{n} a_i a_j C_{X_i X_j} \right)$$

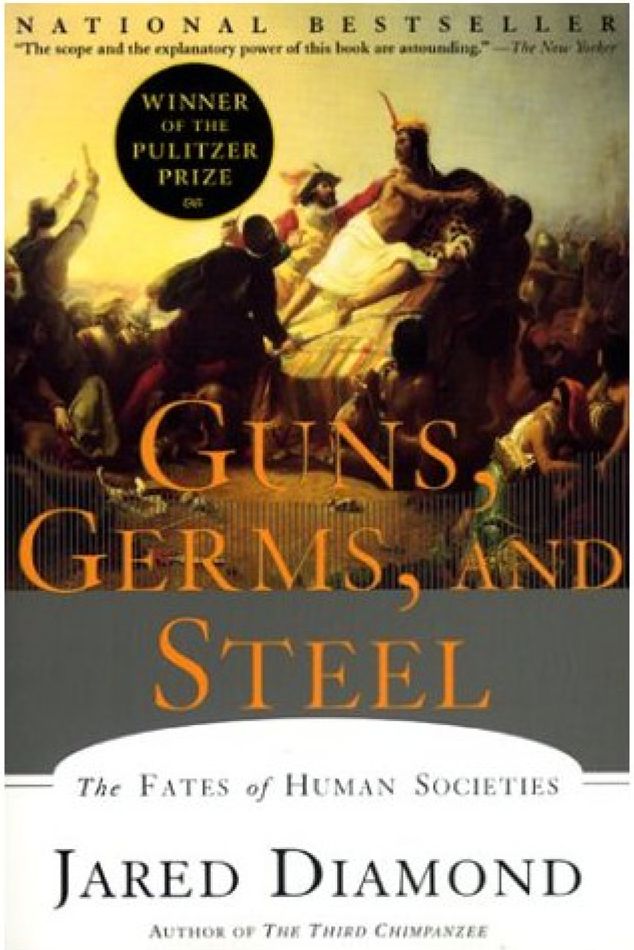
Literaturempfehlungen

- **Thematik:**
  Geschichte /Philosophie der Wahrscheinlichkeitstheorie und Statistik
- **Titel:**
  Chances are… Adventures in Probability
- **Autoren:**
  Michael und Ellen Kaplan, 2006



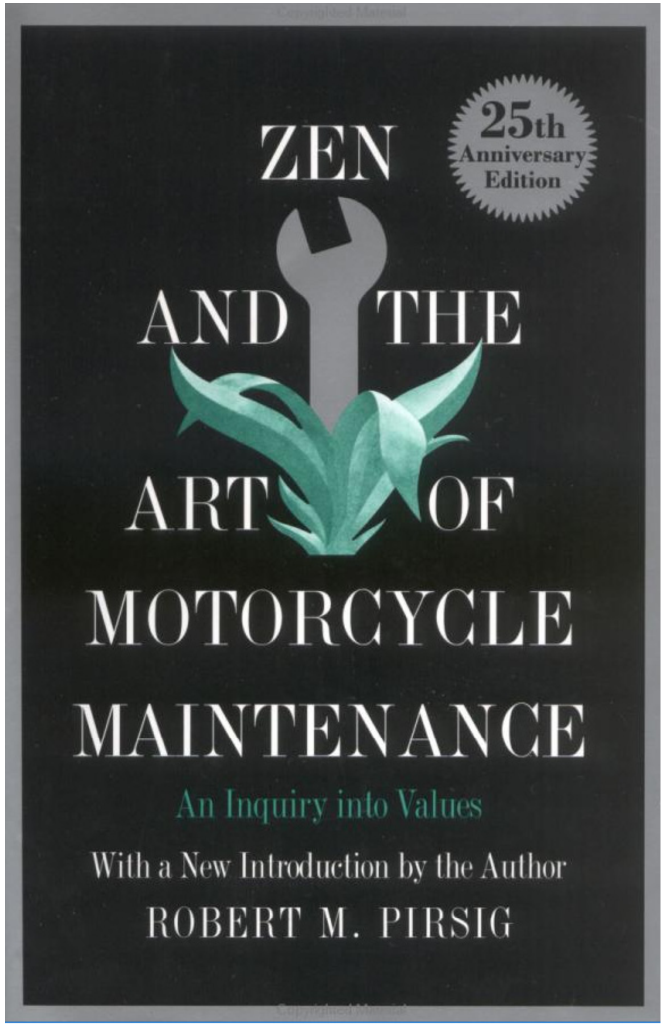
Literaturempfehlungen

- **Thematik:**
  Evolution von Zivilisationen
- **Titel:**
  Guns, Germs & Steel
- **Autor:**
  Jared Diamond, 1997



Literaturempfehlungen

- **Thematik:**
  Wissenschaftstheorie
- **Titel:**
  Zen and the Art of Motorcycle Maintenance
- **Autor:**
  Robert M. Pirsig, 1984



Zufallsvariablen

- bedingte Verteilungen und bedingte Momente

Nehmen wir an, Informationen über die Zufallsvariablen, die ein Ereignis definieren, sind bekannt.

BEISPIEL: Für die Berechnung der Wahrscheinlichkeit, dass sich ein bestimmtes Projekt verzögert, wird angenommen, dass ein Teilabschnitt die geplante Laufzeit um 50% überschreitet.

Zufallsvariablen

- bedingte Verteilungen und bedingte Momente

Die bedingte Wahrscheinlichkeitsdichtefunktion für die Zufallsvariable X_1 bei gegebenem Realisation der Zufallsvariable X_2 ist:

$$f_{X_1|X_2}(x_1|x_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}$$

Wenn X_1 und X_2 voneinander unabhängig sind, gilt:

$$f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)$$

Zufallsvariablen

- bedingte Verteilungen und bedingte Momente

Die bedingte kumulative Verteilungsfunktion kann mittels Integration errechnet werden:

$$F_{X_1|X_2}(x_1|x_2) = \frac{\int_{-\infty}^{x_1} f_{X_1,X_2}(z,x_2)dz}{f_{X_2}(x_2)}$$

Zufallsvariablen

- bedingte Verteilungen und bedingte Momente

Die unbedingte kumulative Verteilungsfunktion für die Zufallsvariable X_1 kann anhand des Theorems der totalen Wahrscheinlichkeit aus der bedingten kumulativen Verteilungsfunktion abgeleitet werden:

$$F_{X_1}(x_1) = \int_{-\infty}^{\infty} F_{X_1|X_2}(x_1|x_2) f_{X_2}(x_2) dx_2$$

Der bedingte Erwartungswert ist wie folgt definiert:

$$\mu_{X_1|X_2} = E\left[X_1 \middle| X_2 = x_2\right] = \int_{-\infty}^{\infty} x_1 f_{X_1|X_2}\left(x_1 \middle| x_2\right) dx_1$$

Zufallsvariablen

Häufig ist man an Berechnungen der Wahrscheinlichkeiten für Funktionen von Zufallsvariablen interessiert.

Funktionen sind hilfreich, um Ereignisse, die von Interesse sind, zu beschreiben – sie dienen als Ingenieurmodelle.

Ein einfacher Fall ist die Summe zweier Zufallsvariablen – dafür ist es hilfreich, die kumulative Verteilungsfunktion zu bestimmen.

Zufallsvariablen

- Die kumulative Verteilungsfunktion für die Summe zweier Zufallsvariablen

Wir setzten die Summe $Y = X_1 + X_2$ und $f_{X_1,X_2}(x_1,x_2)$ voraus.

Somit können wir zuerst die Dichtefunktion für $Y = x_1 + X_2$ bestimmen, unter der Annahme, dass X_1 gegeben ist, z.B. durch

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)} \qquad f_{Y|X_1}(y|x_1) = f_{X_2|X_1}(y - x_1|x_1)$$

Wir erhalten: $f_{Y,X_1}(y,x_1) = f_{X_2|X_1}(y-x_1|x_1) f_{X_1}(x_1) = f_{X_2,X_1}(y-x_1,x_1)$

Zufallsvariablen

- Die kumulative Verteilungsfunktion für die Summe zweier Zufallsvariablen

Die marginale Wahrscheinlichkeitsdichtefunktion für Y kann nun durch Integration über X_1 ermittelt werden:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2,X_1}(y - x_1, x_1)\,dx_1$$

Falls X_1 und X_2 voneinander unabhängig sind, erhält man das sogenannte Faltungsintegral:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2}(y - x_1) f_{X_1}(x_1)\,dx_1$$

Zufallsvariablen

- Die kumulative Verteilungsfunktion für Funktionen von Zufallsvariablen

Es soll die kumulative Verteilungsfunktion für eine Funktion von Zufallsvariablen z.B. $Y = g(X)$ berechnet werden.

Dabei ist die Wahrscheinlichkeitsverteilungsfunktion von X gegeben als $F_X(x)$.

Falls $g(x)$ monoton ansteigt und der Zusammenhang eineindeutig ist, kann die Realisation von Y nur dann kleiner als y_0 sein, wenn auch die Realisation von X kleiner ist als x_0. Dabei gilt $x_0 = g^{-1}(y_0)$ und es folgt:

$$F_Y(y) = P(Y \leq y) = P(X \leq g^{-1}(y))$$

Die kumulative Verteilungsfunktion für Y ist also auf folgende Weise definiert:

$$F_Y(y) = F_X(g^{-1}(y))$$

Zufallsvariablen

- Die kumulative Verteilungsfunktion für Funktionen von Zufallsvariablen

ausgehend von $F_Y(y) = F_X(g^{-1}(y))$

erhält man $f_Y(y) = \dfrac{dF_X(g^{-1}(y))}{dy}$

$$f_Y(y) = \frac{dg^{-1}(y)}{dy} f_X(g^{-1}(y)) \quad \Rightarrow \quad f_Y(y) = \frac{dx}{dy} f_X(x)$$

Zufallsvariablen

- Die kumulative Verteilungsfunktion für Funktionen von Zufallsvariablen

Falls $g(x)$ monoton fallend ist, kann die Realisation von Y nur dann kleiner als y_0 sein, wenn die Realisation von X größer als x_0 ist. In diesem Fall muss das Vorzeichen vertauscht werden:

$$F_Y(y) = -F_X(g^{-1}(y))$$

man erhält:

$$f_Y(y) = -\frac{dx}{dy} f_X(x)$$

Grundsätzlich folgt daraus für monoton ansteigende oder fallende Funktionen:

$$f_Y(y) = \left|\frac{dx}{dy}\right| f_X(x)$$

Zufallsvariablen

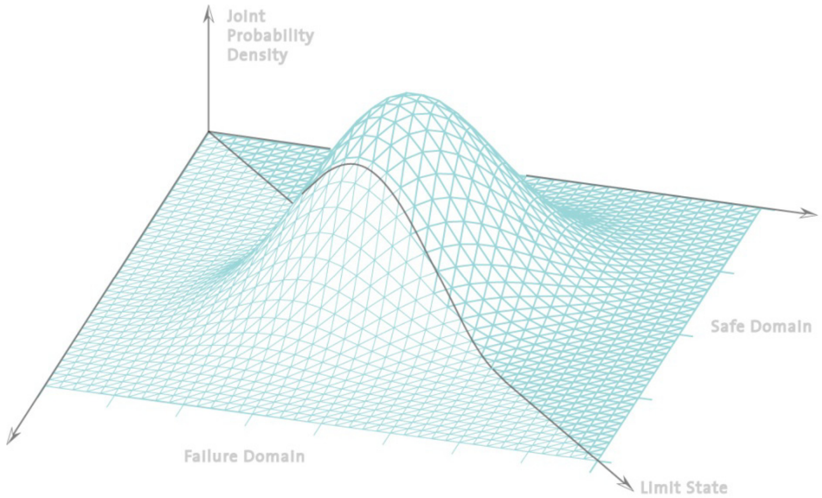
- Die kumulative Verteilungsfunktion für Funktionen von Zufallsvariablen

Falls die Elemente eines Zufallsvektors $\mathbf{Y}=(Y_1,Y_2,..Y_n)^T$ als eineindeutige Darstellung der monoton steigenden oder fallenden Funktionen $g_i,\ i=1,2,..n$ der Elemente des Zufallsvektors $\mathbf{X}=(X_1,X_2,..X_n)^T$ angenommen werden können,

$$Y_i = g_i(\mathbf{X})$$

ergibt sich: $f_{\mathbf{Y}}(\mathbf{y})=|\mathbf{J}|\,f_{\mathbf{X}}(\mathbf{x})$

Mit $|\mathbf{J}|$ als Betrag der Determinante von $\mathbf{J}=\begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots\cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots\cdots & \frac{\partial x_n}{\partial_n} \end{bmatrix}$



Statistik und Wahrscheinlichkeitsrechnung

Prof. Dr. Michael Havbro Faber