3 Mehrdimensionale Zufallsvariablen, Verteilungsparameter

Zoltán Zomotor

Versionsstand: 23. März 2014, 22:07

Die nummerierten Felder bitte während der Vorlesung ausfüllen.

Inhaltsverzeichnis

1 Verteilungsparameter 1
- 1.1 Lageparameter . . . 1
- 1.2 Streuungsparameter . . . 5
- 1.3 Erwarungswert und Varianz wichtiger Verteilungen . . . 6
- 1.4 Weitere Ausagen über Erwartungswert und Varianz . . . 6
- 1.5 Kovarianz und Korrelation . . . 7
- 1.6 Weitere Aussagen über Kovarianz und Korrelation . . . 7
- 1.7 Übungsaufgaben . . . 8

2 Mehrdimensionale Zufallsvariablen 8
- 2.1 Mehrdimensionale diskrete Zufallsvariablen . . . 8
- 2.2 Mehrdimensionale stetige Zufallsvariablen . . . 9
- 2.3 Gesetz der großen Zahlen . . . 9
- 2.4 Zentraler Grenzwertsatz . . . 11
- 2.5 Übungsaufgaben . . . 12

1 Verteilungsparameter[1]

1.1 Lageparameter

Modus oder Modalwert x_{Mod}: (siehe auch [Pap11, Seite 486])

1

$f(x_{\text{Mod}}) \geqq f(x)$ für alle x

Im Allgemeinen *nicht* eindeutig, zum Beispiel Gleichverteilung.

[1] [Pap11, Teil II, Kapitel 5]

Beispiele:

- Normalverteilung: 2 $x_{\text{Mod}} = \mu$
- Diskrete Verteilung mit

x	0	1	2
$f(x)$	0.25	0.5	0.25

$\Rightarrow$ 3 $x_{\text{Mod}} = 1$

Median oder Zentralwert x_{Med}:

4 $F(x_{\text{Med}}) = \frac{1}{2}$ bzw. kleinstes x mit $F(x) > \frac{1}{2}$

Beispiele:

- Normalverteilung: 5 $x_{\text{Med}} = \mu$
- Diskrete Verteilung oben: 6 $F(0) = \frac{1}{4} < \frac{1}{2}$, $F(1) = \frac{3}{4} > \frac{1}{2} \Rightarrow x_{\text{Med}} = 1$

α-Fraktil x_α (wird auch Quantil, oder Perzentil genannt)

7 $P(X \leq x_\alpha) = F(x_\alpha) = \alpha$

Beispiel: $X \sim N(0,1)$, $Y \sim N(3,2)$

8

$$\begin{aligned} x_{0.975} &= 1.96 \\ x_{0.025} &= -x_{0.975} = -1.96 \\ y_{0.025} &= 2 \cdot x_{0.025} + 3 = -0.92 \end{aligned}$$

Hinweise:

- $x_{\text{Med}} =$ 9 $x_{0.5}$
- Falls x_α nicht vertafelt $\Rightarrow$ Interpolation:

10 $$x_\alpha = x_a + (x_b - x_a)\frac{\alpha - a}{b - a} \qquad \begin{pmatrix} a & \text{größte vertafelte Zahl} < \alpha \\ b & \text{kleinste vertafelte Zahl} > \alpha \end{pmatrix}$$

Beispiel: $X \sim N(0,1)$

11

$$x_{0.6} \approx 0.25 + (0.26 - 0.25) \cdot \frac{0.6 - 0.5987}{0.6026 - 0.5987} = 0.2533$$

Erwartungswert $\mathrm{E}(X)$ bzw. μ:

12

$$\mathrm{E}(X) = \begin{cases} \sum_i x_i f(x_i), & \text{falls } X \text{ diskret} \\ \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x & \text{falls } X \text{ stetig} \end{cases}$$

Beispiele:

- Normalverteilung: $\mathrm{E}(X) = \mu$
- Diskrete Verteilung oben:

13

$$\mathrm{E}(X) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Rechenregeln für den Erwartungswert

(R1) Ist f symmetrisch bezüglich a, so gilt 14 $\mathrm{E}(X) = a$

Beispiel: Gleichverteilung: 15 $\mathrm{E}(X) = \dfrac{a+b}{2}$

(R2) Ist $Y = g(X)$, so gilt

16

$$\mathrm{E}(Y) = \mathrm{E}(g(x)) = \begin{cases} \sum_i g(x_i) f(x_i), & \text{falls } X \text{ diskret} \\ \int_{-\infty}^{\infty} g(x) \cdot f(x)\,\mathrm{d}x & \text{falls } X \text{ stetig} \end{cases}$$

Beispiel: X gleichverteilt in $[1,4]$

17

$$\mathrm{E}(X) = \frac{1+4}{2} = 2.5$$

$$\mathrm{E}(X^2) = \int_1^4 x^2 \cdot \frac{1}{4-1}\,\mathrm{d}x = \frac{1}{3}\left[\frac{1}{3}x^3\right]_1^4 = \frac{1}{9}(64-1) = \frac{63}{9} = 7$$

(R3) **Lineare Transformation:**

18

Spezialfall von (R2) mit $g(x) = a+bx$: $\mathrm{E}(a+bX) = a + b\cdot\mathrm{E}(X)$

(R4) **Summenbildung:**

19

$$\mathrm{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\mathrm{E}(X_i)$$

Beispiel: X gleichverteilt in $[0,10]$, $Y \sim N(1,1)$, $Z = X + 5Y$

20

$$\mathrm{E}(Z) = \begin{aligned} &\mathrm{E}(X+5Y) = \mathrm{E}(X) + \mathrm{E}(5Y) = \mathrm{E}(X) + 5\cdot\mathrm{E}(Y) \\ &= \frac{10+0}{2} + 5\cdot 1 = 10 \end{aligned}$$

(R5) **Unabhängigkeit:**

21

X, Y unabhängig $\Rightarrow$ $E(X\cdot Y) = \mathrm{E}(X)\cdot\mathrm{E}(Y)$

22

(R6) Ist stets $X \leqq Y$, so gilt $\mathrm{E}(X) \leqq \mathrm{E}(Y)$

(R7) **Jensensche Ungleichung:** In der Situation (R2) gilt:

23

$$\begin{aligned} \mathrm{E}(g(X)) &\geqq g(\mathrm{E}(X)), \quad \text{falls } g \text{ konvex} \\ \mathrm{E}(g(X)) &\leqq g(\mathrm{E}(X)), \quad \text{falls } g \text{ konkav} \end{aligned}$$

Beispiel: X gleichverteilt in $[1,4]$, $g(x) = x^2$ (vgl. (R2))

24

$g(x)$ ist konvex $\Rightarrow$

$$\mathrm{E}(g(X)) = \mathrm{E}(X^2) = 7 > 6.25 = 2.5^2 = (\mathrm{E}(X))^2 = g(\mathrm{E}(X))$$

1.2 Streuungsparameter

Varianz $\mathrm{Var}(X)$ oder σ^2: Erwartungswert der quadratischen Abweichung vom Erwartungswert:

25

$$\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}(X))^2 = \begin{cases} \sum_i (x_i - \mathrm{E}(X))^2 f(x_i), & \text{falls } X \text{ diskret} \\ \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 f(x)\,\mathrm{d}x & \text{falls } X \text{ stetig} \end{cases}$$

Standardabweichung $\mathrm{Sta}(X)$ oder σ: $\mathrm{Sta}(X) = \sigma = \sqrt{\mathrm{Var}(X)}$

Beispiel: Diskrete Verteilung oben:

26

$$\mathrm{Var}(X) = (0-1)^2 \cdot \frac{1}{4} + (1-1)^2 \cdot \frac{1}{2} + (2-1)^2 \frac{1}{4} = \frac{1}{2}$$

Rechenregeln für die Varianz

R1 **Verschiebungssatz:**

27 $\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$

Beispiel: Diskrete Verteilung oben:

28

$$\mathrm{E}(X^2) = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} = \frac{3}{2}$$

$$\Rightarrow \mathrm{E}(X^2) - (\mathrm{E}(X))^2 = \frac{3}{2} - 1^2 = \frac{1}{2} = \mathrm{Var}(X)$$

R2 **Lineare Transformation:**

29 $\mathrm{Var}(a + bX) = b^2 \cdot \mathrm{Var}(X)$

R3 **Summenbildung:**

30

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Voraussetzung: Unabhängigkeit der X_i!

1.3 Erwarungswert und Varianz wichtiger Verteilungen

31

Verteilung von X	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
Binomialverteilung $B(n,p)$	np	$np(1-p)$
Hypergeom. Verteilung $Hyp(N,M,n)$	$n \cdot \frac{M}{N}$	$n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1}$
Poisson-Verteilung $P(\mu)$	μ	μ
Gleichverteilung in $[a,b]$ mit $a<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentialverteilung $Exp(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normalverteilung $N(\mu,\sigma)$	μ	σ^2

1.4 Weitere Ausagen über Erwartungswert und Varianz

1. Ist X beliebig verteilt mit $\mathrm{E}(X)=\mu$ und $\mathrm{Var}(X)=\sigma^2$, so besitzt

   32

   $$Y = \frac{X-\mu}{\sigma}$$

   den Erwartungswert 33 0 und die Varianz 34 1

2. Sind $X_1,\dots,X_n$ unabhängig mit $\mathrm{E}(X_i)=\mu$ und $\mathrm{Var}(X_i)=\sigma^2$, so besitzt das Stichprobenmittel

   35

   $$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

   den Erwartungswert 36 μ und die Varianz 37 $\frac{\sigma^2}{n}$

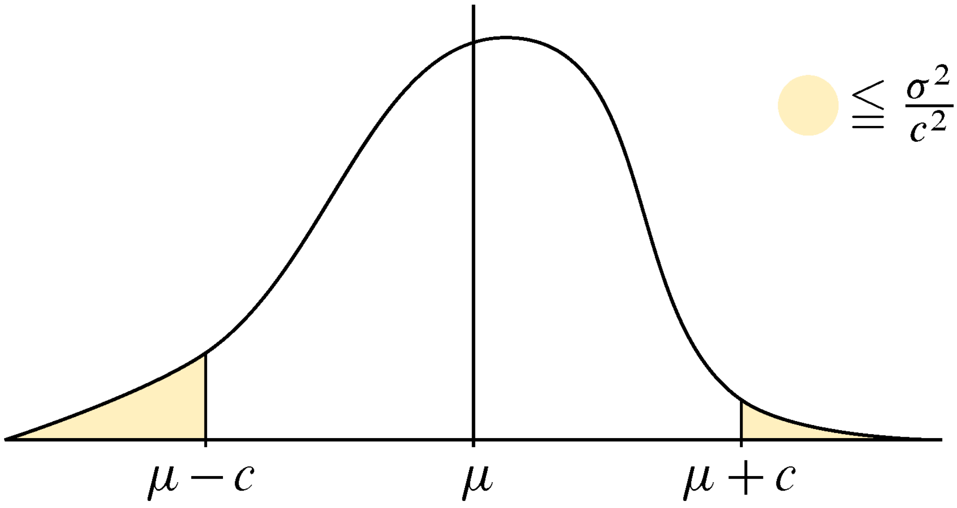
3. **Ungleichung von Tschebyscheff:**

   38

   $$P(|X-E(X)| \geqq c) \leqq \frac{\mathrm{Var}(X)}{c^2}$$

   wobei X beliebig verteilt sein darf.

39



1.5 Kovarianz und Korrelation

Kovarianz $\mathrm{Cov}(X, Y)$**:**

40

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathrm{E}[(X - \mathrm{E}(x)) \cdot (Y - \mathrm{E}(Y))] \\ &= \mathrm{E}(X \cdot Y) - \mathrm{E}(X) \cdot \mathrm{E}(Y) \quad \text{(Verschiebungssatz)}\end{aligned}$$

Korrelationskoeffizient $\varrho(X, Y)$**:**

41

$$\varrho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$$

Bemerkungen:

- $\varrho \in [-1, 1]$
- $|\varrho| = 1 \Leftrightarrow Y = a + bX$ mit $(b \neq 0$
- $\varrho = 0 \Leftrightarrow X, Y$ unkorreliert

1.6 Weitere Aussagen über Kovarianz und Korrelation

- Sind X, Y unabhängig, so gilt:

42 $\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y) \Rightarrow \mathrm{Cov}(X, Y) = 0 \Rightarrow \varrho(X, Y) = 0$

$\Rightarrow$ Aus Unabhängigkeit folgt Unkorreliertheit.

- Rechenregeln für die Kovarianz:

43

$$\begin{aligned} \mathrm{Cov}(X,X) &= \mathrm{Var}(X) \\ \mathrm{Cov}(X,Y) &= \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(a+bX, c+dY) &= b \cdot d \cdot \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X+Y,Z) &= \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z) \end{aligned}$$

- Summe zweier Zufallsvariablen:

44

$\mathrm{Var}(X+Y) =$ $\mathrm{Var}\,X + \mathrm{Var}\,Y + 2 \cdot \mathrm{Cov}(X,Y)$

- Differenz zweier Zufallsvariablen:

45

$\mathrm{Var}(X-Y) =$ $\mathrm{Var}\,X + \mathrm{Var}\,Y - 2 \cdot \mathrm{Cov}(X,Y)$

1.7 Übungsaufgaben

[Pap11, Zu Abschnitt 5, Seite 459ff]

2 Mehrdimensionale Zufallsvariablen[2]

2.1 Mehrdimensionale diskrete Zufallsvariablen

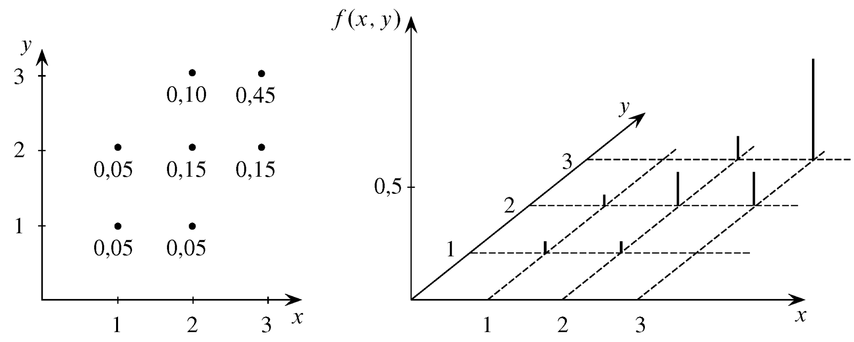
- $(X_1, \ldots, X_n)$ heißt *diskret,* falls der Wertebereich endlich ist.
- Gemeinsame Wahrscheinlichkeitsfunktion:

46

$f(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n)$

- Darstellungsmöglichkeiten: Tabelle, Streuungsdiagramm, Stabdiagramm.

47



[2] [Pap11, Teil II, Kapitel 7]

2.2 Mehrdimensionale stetige Zufallsvariablen

- $(X_1, \ldots, X_n)$ heißt *stetig,* wenn es eine Funktion $f(x_1, \ldots, x_n)$ gibt, so dass die Verteilungsfunktion F folgende Gestalt hat:

48
$$F(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_n} f(t_1, \ldots, t_n) \, \mathrm{d}t_n \ldots \mathrm{d}t_1$$

- $f(x_1, \ldots, x_n)$ heißt *gemeinsame Dichte* von $(X_1, \ldots, X_n)$
- Wir beschäftigen uns im Weiteren vor allem mit zweidimensionalen Zufallsvariablen (X, Y) oder (X_1, X_2)
- Randverteilungsfunktion: 49 $F_1(x) = F(x, \infty),\ F_2(y) = F(\infty, y)$
- Rand-Wahrscheinlichkeitsdichtefunktion:

50
$$f_1(x) = \int\limits_{-\infty}^{\infty} f(x, y) \, \mathrm{d}y, \qquad f_2(y) = \int\limits_{-\infty}^{\infty} f(x, y) \, \mathrm{d}x$$

- Bedingte Wahrscheinlichkeitsfunktion:

51
$$f_1(x|y) = \frac{f(x, y)}{f_2(y)} \text{ bzw. } f_2(y|x) = \frac{f(x, y)}{f_1(x)}$$

- Unabhängigkeit: $(X_1, \ldots, X_n)$ sind genau dann unabhängig, wenn gilt:

52
$$f(x_1, \ldots, x_n) = f_1(x_1) \cdot \ldots \cdot f_n(x_n)$$

2.3 Gesetz der großen Zahlen

- Gegeben: Zufallsvariablen $X_1, \ldots, X_n$ mit den Eigenschaften:

53
  - unabhängig und identisch verteilt („i.i.d.“)
  - $\mathrm{E}(X_i) = \mu$
  - $\mathrm{Var}(X_i) = \sigma^2$

- Gesucht: Verhalten von

54
$$\sum_{i=1}^{n} X_i \qquad \text{oder} \qquad \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

wenn n laufend erhöht wird.

- Bekannt: siehe Lücken 35, 36 und 37:

55

$$\mathrm{E}(\bar{X}_n) = \mu$$
$$\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$$

- Ungleichung von Tschebyscheff (Lücke 38)

56
$$P(|X - E(X)| \geqq c) \leqq \frac{\mathrm{Var}(X)}{c^2}$$

57

angewandt auf $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

58

ergibt $P(|\bar{X}_n - \mu| \geqq c) \leqq \frac{\sigma^2}{nc^2}$

Für $n \to \infty$ folgt dann das *Gesetz der großen Zahlen:*

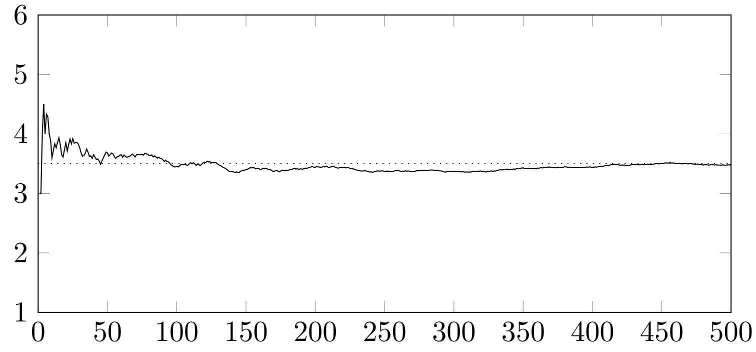
59

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| \geqq c) = 0 \quad \text{bzw.}$$
$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| \leqq c) = 1$$

- Beispiel: Simulation 500 mal Würfeln

60

$\mathrm{E}(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$



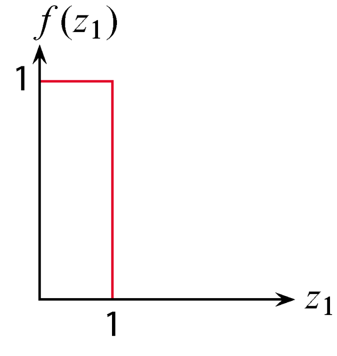
2.4 Zentraler Grenzwertsatz

- Summen von i.i.d. Zufallsvariablen sind für große n näherungsweise 61 normalverteilt

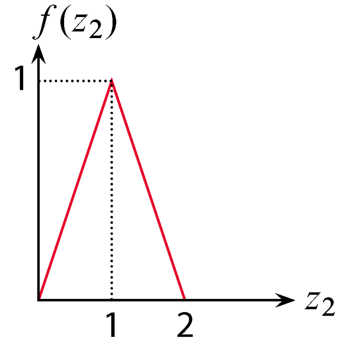
- Beispiel: X_1, X_2, X_3 in $\begin{bmatrix}0 & 1\end{bmatrix}$ gleichverteilt

  $Z_1 = X_1$ $\quad$ $Z_2 = X_1 + X_2$ $\quad$ $Z_3 = X_1 + X_2 + X_3$

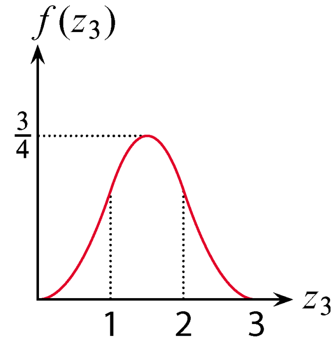
62



63



64



$\Rightarrow$ Zentraler Grenzwertsatz:

65

$$\sum_{i=1}^{n} X_i \sim N(n\mu, \sigma\sqrt{n})$$

- Approximationsvoraussetzungen:
  - Im Allgemeinen: 66 $n > 30$
  - Spezialfall Binomialverteilung: 67 $np \geq 5$ und $n(1-p) \geq 5$

- Gauß-Statistik:

68

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma \cdot \frac{1}{\sqrt{n}}} = \frac{\bar{X} - \mu}{\sigma} \cdot \sqrt{n} \sim N(0,1)$$

Beispiel:

Gegeben: $X_1, \ldots, X_{12}$ gleichverteilt in $[0,1]$

Gesucht: approximative Verteilung von 69 $Y = \sum_{i=1}^{12} X_i - 6$

Mit den Lücken 18, 19, 29 und 30 folgt

70

$$
\begin{aligned}
\mathrm{E}(Y) &= \sum_{i=1}^{12} \mathrm{E}(X_i) - 6 = 12 \cdot \frac{1}{2} - 6 = 0 \\
\mathrm{Var}(Y) &= \sum_{i=1}^{12} \mathrm{Var}(X_i) - 6 = 12 \cdot \frac{1}{2} = 1 \\
&\Rightarrow Y \sim N(0,1)
\end{aligned}
$$

Lücke 65 $n > 30$ schlechte Approximation!

2.5 Übungsaufgaben

[Pap11, Zu Abschnitt 7, Seite 467ff]