

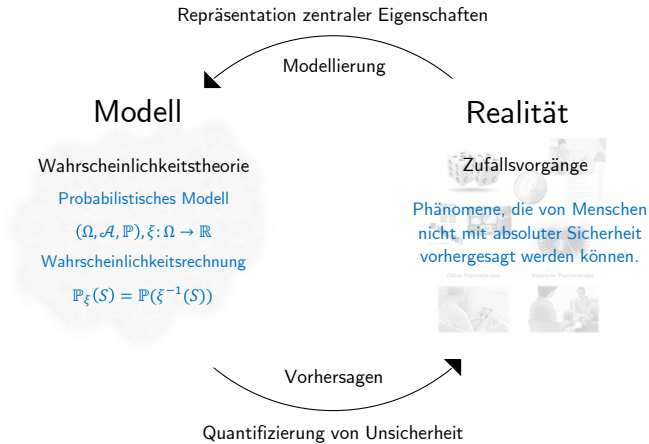


# Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2022/23

Prof. Dr. Dirk Ostwald

(6) Erwartungswert, Varianz, Kovarianz



Wir nehmen an, dass die BDI Score Fehler der Proband:innen Realisierungen  
**unabhängiger und identisch** normalverteilter Zufallsvariablen sind.

Modell

Wahrscheinlichkeitstheorie

$$y_{1j} = \mu_1 + \varepsilon_{1j}, \varepsilon_{1j} \sim N(0, \sigma^2), j = 1, \dots, n_1$$

$$y_{2j} = \mu_2 + \varepsilon_{2j}, \varepsilon_{2j} \sim N(0, \sigma^2), j = 1, \dots, n_2$$

$$\mathbb{E}(\varepsilon_{ij}) = 0, \mathbb{V}(\varepsilon_{ij}) = \sigma^2 \quad \forall i, j$$

$$\mathbb{C}(\varepsilon_{ij}, \varepsilon_{kl}) = 0 \quad \forall i \neq k, j \neq l$$

Modellierung

Realität

Zufallsvorgang

Online Psychotherapie      Klassische Psychotherapie



Klinische Studie zum Vergleich der Effekte von  
Face-to-Face und Online PT bei Depression

	Online	FS	FS
1	Online	-25	35
2	Online	14	20
3	Online	-25	29
4	Online	14	20
5	Online	-25	29
6	Online	-25	29
7	Online	-25	29
8	Online	14	20
9	Online	-25	29
10	Online	-25	29
11	Online	-25	29
12	Online	-25	29
13	Online	-25	29
14	Online	-25	29
15	Online	-25	29
16	Online	-25	29
17	Online	-25	29
18	Online	-25	29
19	Online	-25	29
20	Online	-25	29
21	Online	-25	29
22	Online	-25	29
23	Online	-25	29
24	Online	-25	29
25	Online	-25	29
26	Online	-25	29
27	Online	-25	29
28	Online	-25	29
29	Online	-25	29
30	Online	-25	29
31	Online	-25	29
32	Online	-25	29
33	Online	-25	29
34	Online	-25	29
35	Online	-25	29
36	Online	-25	29
37	Online	-25	29
38	Online	-25	29
39	Online	-25	29
40	Online	-25	29
41	Online	-25	29
42	Online	-25	29
43	Online	-25	29
44	Online	-25	29
45	Online	-25	29
46	Online	-25	29
47	Online	-25	29
48	Online	-25	29
49	Online	-25	29
50	Online	-25	29
51	Online	-25	29
52	Online	-25	29
53	Online	-25	29
54	Online	-25	29
55	Online	-25	29
56	Online	-25	29
57	Online	-25	29
58	Online	-25	29
59	Online	-25	29
60	Online	-25	29
61	Online	-25	29
62	Online	-25	29
63	Online	-25	29
64	Online	-25	29
65	Online	-25	29
66	Online	-25	29
67	Online	-25	29
68	Online	-25	29
69	Online	-25	29
70	Online	-25	29
71	Online	-25	29
72	Online	-25	29
73	Online	-25	29
74	Online	-25	29
75	Online	-25	29
76	Online	-25	29
77	Online	-25	29
78	Online	-25	29
79	Online	-25	29
80	Online	-25	29
81	Online	-25	29
82	Online	-25	29
83	Online	-25	29
84	Online	-25	29
85	Online	-25	29
86	Online	-25	29
87	Online	-25	29
88	Online	-25	29
89	Online	-25	29
90	Online	-25	29
91	Online	-25	29
92	Online	-25	29
93	Online	-25	29
94	Online	-25	29
95	Online	-25	29
96	Online	-25	29
97	Online	-25	29
98	Online	-25	29
99	Online	-25	29
100	Online	-25	29

Vorhersagen

---

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Stichprobenkovarianz und Stichprobenkorrelation

Selbstkontrollfragen

---

## **Erwartungswert**

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Stichprobenkovarianz und Stichprobenkorrelation

Selbstkontrollfragen

## Definition (Erwartungswert)

$(\Omega, \mathcal{A}, \mathbb{P})$  sei ein Wahrscheinlichkeitsraum und  $\xi$  sei eine Zufallsvariable. Dann ist der *Erwartungswert* von  $\xi$  definiert als

- $\mathbb{E}(\xi) := \sum_{x \in \mathcal{X}} x p_{\xi}(x)$ , wenn  $\xi : \Omega \rightarrow \mathcal{X}$  diskret mit WMF  $p_{\xi}$  und Ergebnisraum  $\mathcal{X}$  ist,
- $\mathbb{E}(\xi) := \int_{-\infty}^{\infty} x p_{\xi}(x) dx$ , wenn  $\xi : \Omega \rightarrow \mathbb{R}$  kontinuierlich mit WDF  $p_{\xi}$  ist.

Der Erwartungswert einer Zufallsvariable heißt *existent*, wenn er endlich ist.

## Bemerkungen

- Der Erwartungswert ist eine skalare Zusammenfassung einer Verteilung.
- Intuitiv ist  $\mathbb{E}(\xi) \approx \frac{1}{n} \sum_{i=1}^n \xi_i$  für eine große Zahl  $n$  von Kopien  $\xi_i$  von  $\xi$ .

## Beispiel (Erwartungswert einer Bernoulli Zufallsvariable)

Es sei  $\xi \sim \text{Bern}(\mu)$ . Dann gilt  $\mathbb{E}(\xi) = \mu$ .

### Beweis

$\xi$  ist diskret mit  $\mathcal{X} = \{0, 1\}$ . Also gilt

$$\begin{aligned}\mathbb{E}(\xi) &= \sum_{x \in \{0, 1\}} x \text{Bern}(x; \mu) \\ &= 0 \cdot \mu^0 (1 - \mu)^{1-0} + 1 \cdot \mu^1 (1 - \mu)^{1-1} \\ &= 1 \cdot \mu^1 (1 - \mu)^0 \\ &= \mu.\end{aligned}\tag{1}$$

□



# Erwartungswert

## Beispiel (Erwartungswert einer normalverteilten Zufallsvariable)

Es sei  $\xi \sim N(\mu, \sigma^2)$ . Dann gilt  $\mathbb{E}(\xi) = \mu$ .

### Beweis

Wir halten zunächst ohne Beweis fest, dass

$$\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}. \quad (2)$$

Mit der Definition des Erwartungswerts für kontinuierliche Zufallsvariablen gilt

$$\mathbb{E}(\xi) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) dx. \quad (3)$$

Mit der Substitutionsregel

$$\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(x))g'(x) dx \quad (4)$$

und der Definition von

$$g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto g(x) := \sqrt{2\sigma^2}x + \mu \text{ with } g'(x) = \sqrt{2\sigma^2}, \quad (5)$$

gilt dann

$$\begin{aligned}\mathbb{E}(\xi) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx \\&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x + \mu) \exp\left(-\frac{1}{2\sigma^2}\left(\left(\sqrt{2\sigma^2}x + \mu\right) - \mu\right)^2\right) \sqrt{2\sigma^2} dx \\&= \frac{\sqrt{2\sigma^2}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x + \mu) \exp(-x^2) dx \\&= \frac{1}{\sqrt{\pi}} \left( \sqrt{2\sigma^2} \int_{-\infty}^{\infty} x \exp(-x^2) dx + \mu \int_{-\infty}^{\infty} \exp(-x^2) dx \right) \\&= \frac{1}{\sqrt{\pi}} \left( \sqrt{2\sigma^2} \int_{-\infty}^{\infty} x \exp(-x^2) dx + \mu\sqrt{\pi} \right)\end{aligned}\tag{6}$$

Eine Stammfunktion von  $x \exp(-x^2)$  ist  $-\frac{1}{2} \exp(-x^2)$ . Mit

$$\lim_{x \rightarrow -\infty} \exp(-x^2) = 0 \text{ und } \lim_{x \rightarrow \infty} \exp(-x^2) = 0\tag{7}$$

verschwindet der Integralterm und wir erhalten

$$\mathbb{E}(\xi) = \frac{1}{\sqrt{\pi}} (\mu\sqrt{\pi}) = \mu.\tag{8}$$

□

## Theorem (Eigenschaften des Erwartungswerts)

(1) (Linear-affine Transformation) Für eine Zufallsvariable  $\xi$  und  $a, b \in \mathbb{R}$  gilt

$$\mathbb{E}(a\xi + b) = a\mathbb{E}(\xi) + b. \quad (9)$$

(2) (Linearkombination) Für Zufallsvariablen  $\xi_1, \dots, \xi_n$  und  $a_1, \dots, a_n \in \mathbb{R}$  gilt

$$\mathbb{E}\left(\sum_{i=1}^n a_i \xi_i\right) = \sum_{i=1}^n a_i \mathbb{E}(\xi_i). \quad (10)$$

(3) (Faktorisierung bei Unabhängigkeit) Für unabhängige Zufallsvariablen  $\xi_1, \dots, \xi_n$  gilt

$$\mathbb{E}\left(\prod_{i=1}^n \xi_i\right) = \prod_{i=1}^n \mathbb{E}(\xi_i). \quad (11)$$

### Bemerkung

- Die genannten Eigenschaften sind oft nützlich zur Berechnung von Erwartungswerten.

## Beweis

Eigenschaft (1) folgt aus den Linearitätseigenschaften von Summen und Integralen. Wir betrachten nur den Fall einer kontinuierlichen Zufallsvariable  $\xi$  mit WDF  $p_\xi$  genauer. In diesem Fall gilt

$$\begin{aligned}\mathbb{E}(v) &= \mathbb{E}(a\xi + b) \\ &= \int_{-\infty}^{\infty} (ax + b)p_\xi(x) dx \\ &= \int_{-\infty}^{\infty} axp_\xi(x) + bp_\xi(x) dx \\ &= a \int_{-\infty}^{\infty} xp_\xi(x) dx + b \int_{-\infty}^{\infty} p_\xi(x) dx \\ &= a\mathbb{E}(\xi) + b.\end{aligned}\tag{12}$$

## Beweis (fortgeführt)

Eigenschaft (2) folgt gleichfalls aus den Linearitätseigenschaften von Summen und Integralen. Wir betrachten nur den Fall von zwei kontinuierlichen Zufallsvariablen  $\xi_1$  und  $\xi_2$  mit bivariater WDF  $p_{\xi_1, \xi_2}$  genauer. In diesem Fall gilt

$$\begin{aligned} & \mathbb{E} \left( \sum_{i=1}^2 a_i \xi_i \right) \\ &= \mathbb{E}(a_1 \xi_1 + a_2 \xi_2) \\ &= \iint_{\mathbb{R}^2} (a_1 x_1 + a_2 x_2) p_{\xi_1, \xi_2}(x_1, x_2) dx_1 dx_2 \\ &= \iint_{\mathbb{R}^2} a_1 x_1 p_{\xi_1, \xi_2}(x_1, x_2) + a_2 x_2 p_{\xi_1, \xi_2}(x_1, x_2) dx_1 dx_2 \\ &= a_1 \iint_{\mathbb{R}^2} x_1 p_{\xi_1, \xi_2}(x_1, x_2) dx_1 dx_2 + a_2 \iint_{\mathbb{R}^2} x_2 p_{\xi_1, \xi_2}(x_1, x_2) dx_1 dx_2 \end{aligned} \tag{13}$$

## Beweis (fortgeführt)

$$\begin{aligned} &= a_1 \int_{-\infty}^{\infty} x_1 \left( \int_{-\infty}^{\infty} p_{\xi_1, \xi_2}(x_1, x_2) dx_2 \right) dx_1 + a_2 \int_{-\infty}^{\infty} x_2 \left( \int_{-\infty}^{\infty} p_{\xi_1, \xi_2}(x_1, x_2) dx_1 \right) dx_2 \\ &= a_1 \int_{-\infty}^{\infty} x_1 p_{\xi_1}(x_1) dx_1 + a_2 \int_{-\infty}^{\infty} x_2 p_{\xi_2}(x_2) dx_2 \\ &= a_1 \mathbb{E}(\xi_1) + a_2 \mathbb{E}(\xi_2) \\ &= \sum_{i=1}^2 a_i \mathbb{E}(\xi_i). \end{aligned}$$

(14)

Ein Induktionsargument erlaubt dann die Generalisierung vom bivariaten zum  $n$ -variaten Fall.

# Erwartungswert

## Beweis (fortgeführt)

Zu Eigenschaft (3) betrachten wir den Fall von  $n$  kontinuierlichen Zufallsvariablen mit gemeinsamer WDF  $p_{\xi_1, \dots, \xi_n}$ . Weil als  $\xi_1, \dots, \xi_n$  unabhängig vorausgesetzt sind, gilt

$$p_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \prod_{i=1}^n p_{\xi_i}(x_i). \quad (15)$$

Weiterhin gilt also

$$\begin{aligned} \mathbb{E} \left( \prod_{i=1}^n x_i \right) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left( \prod_{i=1}^n x_i \right) p_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^n x_i \prod_{i=1}^n p_{\xi_i}(x_i) dx_1 \dots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^n x_i p_{\xi_i}(x_i) dx_1 \dots dx_n \\ &= \prod_{i=1}^n \int_{-\infty}^{\infty} x_i p_{\xi_i}(x_i) dx_i = \prod_{i=1}^n \mathbb{E}(\xi_i). \end{aligned} \quad (16)$$

□

---

Erwartungswert

## **Varianz und Standardabweichung**

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Stichprobenkovarianz und Stichprobenkorrelation

Selbstkontrollfragen



## Definition (Varianz und Standardabweichung)

Es sei  $\xi$  eine Zufallsvariable mit Erwartungswert  $\mathbb{E}(\xi)$ . Die *Varianz von  $\xi$*  ist definiert als

$$\mathbb{V}(\xi) := \mathbb{E} \left( (\xi - \mathbb{E}(\xi))^2 \right), \quad (17)$$

unter der Annahme, dass dieser Erwartungswert existiert. Die *Standardabweichung von  $\xi$*  ist definiert

$$\mathbb{S}(\xi) := \sqrt{\mathbb{V}(\xi)}. \quad (18)$$

## Bemerkungen

- Die Varianz misst die Streuung (Breite) einer Verteilung.
- Quadratur ist nötig wegen  $\mathbb{E}(\xi - \mathbb{E}(\xi)) = \mathbb{E}(\xi) - \mathbb{E}(\xi) = 0$ .
- Ein alternatives Maß für die Streuung einer Verteilung ist  $\mathbb{E}(|\xi - \mathbb{E}(\xi)|)$ .
- Ein weiteres Maß für die Streuung einer Verteilung ist die Entropie  $-\mathbb{E}(\ln p_\xi)$ .

# Varianz und Standardabweichung

## Beispiel (Varianz einer Bernoulli Zufallsvariable)

Es sei  $\xi \sim \text{Bern}(\mu)$ . Dann ist die Varianz von  $\xi$  gegeben durch

$$\mathbb{V}(\xi) = \mu(1 - \mu). \quad (19)$$

### Beweis

$\xi$  ist eine diskrete Zufallsvariable und es gilt  $\mathbb{E}(\xi) = \mu$ . Also gilt

$$\begin{aligned} \mathbb{V}(\xi) &= \mathbb{E}\left((\xi - \mu)^2\right) \\ &= \sum_{x \in \{0,1\}} (x - \mu)^2 \text{Bern}(x; \mu) \\ &= (0 - \mu)^2 \mu^0 (1 - \mu)^{1-0} + (1 - \mu)^2 \mu^1 (1 - \mu)^{1-1} \\ &= \mu^2 (1 - \mu) + (1 - \mu)^2 \mu \\ &= \left(\mu^2 + (1 - \mu)\mu\right) (1 - \mu) \\ &= \left(\mu^2 + \mu - \mu^2\right) (1 - \mu) \\ &= \mu(1 - \mu). \end{aligned} \quad (20)$$

□

## Theorem (Varianzverschiebungssatz)

Es sei  $\xi$  eine Zufallsvariable. Dann gilt

$$\mathbb{V}(\xi) = \mathbb{E}(\xi^2) - \mathbb{E}(\xi)^2. \quad (21)$$

### Beweis

Mit der Definition der Varianz und der Linearität des Erwartungswerts gilt

$$\begin{aligned} \mathbb{V}(\xi) &= \mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) \\ &= \mathbb{E}\left(\xi^2 - 2\xi\mathbb{E}(\xi) + \mathbb{E}(\xi)^2\right) \\ &= \mathbb{E}(\xi^2) - 2\mathbb{E}(\xi)\mathbb{E}(\xi) + \mathbb{E}\left(\mathbb{E}(\xi)^2\right) \\ &= \mathbb{E}(\xi^2) - 2\mathbb{E}(\xi)^2 + \mathbb{E}(\xi)^2 \\ &= \mathbb{E}(\xi^2) - \mathbb{E}(\xi)^2. \end{aligned} \quad (22)$$

□

### Bemerkung

- Das Theorem ist nützlich, wenn  $\mathbb{E}(\xi^2)$  und  $\mathbb{E}(\xi)$  leicht zu berechnen sind.

# Varianz und Standardabweichung

Beispiel (Varianz einer normalverteilten Zufallsvariable)

Es sei  $\xi \sim N(\mu, \sigma^2)$ . Dann gilt  $\mathbb{V}(\xi) = \sigma^2$ .

Beweis

Wir halten zunächst fest, dass mit dem Varianzverschiebungssatz gilt, dass

$$\mathbb{V}(\xi) = \mathbb{E}(\xi^2) - \mathbb{E}(\xi)^2 = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} x^2 \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) dx - \mu^2 \quad (23)$$

Mit der Substitutionsregel

$$\int_a^b f(g(x))g'(x) dx = \int_{g(a)}^{g(b)} f(x) dx \quad (24)$$

und der Definition von

$$g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sqrt{2\sigma^2}x + \mu, g(-\infty) := -\infty, g(\infty) := \infty, \text{ with } g'(x) = \sqrt{2\sigma^2}, \quad (25)$$

kann das Integral auf der rechten Seite von Gleichung (23) dann als

# Varianz und Standardabweichung

Beweis (fortgeführt)

$$\begin{aligned} & \int_{-\infty}^{\infty} x^2 \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx \\ &= \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x + \mu)^2 \exp\left(-\frac{1}{2\sigma^2}((\sqrt{2\sigma^2}x + \mu) - \mu)^2\right) \sqrt{2\sigma^2} dx \\ &= \sqrt{2\sigma^2} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x + \mu)^2 \exp\left(-\frac{2\sigma^2 x^2}{2\sigma^2}\right) dx \\ &= \sqrt{2\sigma^2} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x + \mu)^2 \exp(-x^2) dx. \end{aligned} \tag{26}$$

geschrieben werden. Also gilt

$$\begin{aligned} \mathbb{V}(\xi) &= \frac{\sqrt{2\sigma^2}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x + \mu)^2 \exp(-x^2) dx - \mu^2 \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x)^2 + 2\sqrt{2\sigma^2}x\mu + \mu^2 \exp(-x^2) dx - \mu^2 \\ &= \frac{1}{\sqrt{\pi}} \left( 2\sigma^2 \int_{-\infty}^{\infty} x^2 \exp(-x^2) dx + 2\sqrt{2\sigma^2}\mu \int_{-\infty}^{\infty} x \exp(-x^2) dx + \mu^2 \int_{-\infty}^{\infty} \exp(-x^2) dx \right) - \mu^2 \end{aligned}$$

# Varianz und Standardabweichung

## Beweis (fortgeführt)

Wir halten weiterhin ohne Beweis fest, dass

$$\int_{-\infty}^{\infty} x \exp(-x^2) dx = 0 \text{ und } \int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}. \quad (27)$$

Es ergibt sich dann

$$\begin{aligned} \mathbb{V}(\xi) &= \frac{1}{\sqrt{\pi}} \left( 2\sigma^2 \int_{-\infty}^{\infty} x^2 \exp(-x^2) dx + \mu^2 \sqrt{\pi} \right) - \mu^2 \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 \exp(-x^2) dx + \mu^2 - \mu^2 \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 \exp(-x^2) dx \end{aligned} \quad (28)$$

Mit der partiellen Integrationsregel

$$\int_a^b f'(x)g(x) dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x) dx \quad (29)$$

und der Definition von

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) := \exp(-x^2) \text{ with } f'(x) = -2 \exp(-x^2) \quad (30)$$

# Varianz und Standardabweichung

## Beweis (fortgeführt)

und

$$g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto g(x) := -\frac{1}{2}x \text{ with } g'(x) = -\frac{1}{2}, \quad (31)$$

so dass

$$f'(x)g(x) = -2 \exp(-x^2) \left( -\frac{1}{2}x \right) = x^2 \exp(-x^2), \quad (32)$$

gilt, ergibt sich dann

$$\begin{aligned} \mathbb{V}(\xi) &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 \exp(-x^2) dx \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \left( -\frac{1}{2}x \exp(-x^2) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \exp(-x^2) \left( -\frac{1}{2} \right) dx \right) \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \left( -\frac{1}{2}x \exp(-x^2) \Big|_{-\infty}^{\infty} + \frac{1}{2} \int_{-\infty}^{\infty} \exp(-x^2) dx \right), \end{aligned} \quad (33)$$

Aus  $\lim_{x \rightarrow \pm\infty} \exp(-x^2) = 0$  schließen wir, dass der erste Term in den Klammern auf der rechten Seite der obigen Gleichung gleich 0 ist. Schließlich ergibt sich

$$\mathbb{V}(\xi) = \frac{2\sigma^2}{\sqrt{\pi}} \left( \frac{1}{2} \int_{-\infty}^{\infty} \exp(-x^2) dx \right) = \frac{\sigma^2}{\sqrt{\pi}} \sqrt{\pi} = \sigma^2. \quad (34)$$

□

## Theorem (Eigenschaften der Varianz)

(1) (Linear-affine Transformation) Für eine Zufallsvariable  $\xi$  und  $a, b \in \mathbb{R}$  gelten

$$\mathbb{V}(a\xi + b) = a^2\mathbb{V}(\xi) \text{ und } \mathbb{S}(a\xi + b) = |a|\mathbb{S}(\xi). \quad (35)$$

(2) (Linearkombination bei Unabhängigkeit) Für unabhängige Zufallsvariablen  $\xi_1, \dots, \xi_n$  und  $a_1, \dots, a_n \in \mathbb{R}$  gilt

$$\mathbb{V}\left(\sum_{i=1}^n a_i \xi_i\right) = \sum_{i=1}^n a_i^2 \mathbb{V}(\xi_i). \quad (36)$$



# Varianz und Standardabweichung

## Beweis

Um Eigenschaft (1) zu zeigen, definieren wir zunächst  $v := a\xi + b$  und halten fest, dass  $\mathbb{E}(v) = a\mathbb{E}(\xi) + b$ . Für die Varianz von  $v$  ergibt sich dann

$$\begin{aligned}\mathbb{V}(v) &= \mathbb{E}\left((v - \mathbb{E}(v))^2\right) \\ &= \mathbb{E}\left((a\xi + b - a\mathbb{E}(\xi) - b)^2\right) \\ &= \mathbb{E}\left((a\xi - a\mathbb{E}(\xi))^2\right) \\ &= \mathbb{E}\left(a^2(\xi - \mathbb{E}(\xi))^2\right) \\ &= \mathbb{E}\left(a^2(\xi - \mathbb{E}(\xi))^2\right) \\ &= a^2\mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) \\ &= a^2\mathbb{V}(\xi)\end{aligned}\tag{37}$$

Wurzelziehen ergibt dann das Resultat für die Standardabweichung.

Für Eigenschaft (2) betrachten wir den Fall zweier unabhängiger Zufallsvariablen  $\xi_1$  und  $\xi_2$  genauer. Wir halten zunächst fest, dass in diesem Fall gilt, dass

$$\mathbb{E}(a_1\xi_1 + a_2\xi_2) = a_1\mathbb{E}(\xi_1) + a_2\mathbb{E}(\xi_2).\tag{38}$$

# Varianz und Standardabweichung

## Beweis (fortgeführt)

Es ergibt sich also

$$\begin{aligned} & \mathbb{V} \left( \sum_{i=1}^2 a_i \xi_i \right) \\ &= \mathbb{V}(a_1 \xi_1 + a_2 \xi_2) \\ &= \mathbb{E} \left( (a_1 \xi_1 + a_2 \xi_2 - \mathbb{E}(a_1 \xi_1 + a_2 \xi_2))^2 \right) \\ &= \mathbb{E} \left( (a_1 \xi_1 + a_2 \xi_2 - a_1 \mathbb{E}(\xi_1) - a_2 \mathbb{E}(\xi_2))^2 \right) \\ &= \mathbb{E} \left( (a_1 \xi_1 - a_1 \mathbb{E}(\xi_1) + a_2 \xi_2 - a_2 \mathbb{E}(\xi_2))^2 \right) \\ &= \mathbb{E} \left( ((a_1(\xi_1 - \mathbb{E}(\xi_1)) + (a_2(\xi_2 - \mathbb{E}(\xi_2))))^2 \right) \\ &= \mathbb{E} \left( (a_1(\xi_1 - \mathbb{E}(\xi_1)))^2 - 2(a_1(\xi_1 - \mathbb{E}(\xi_1))(a_2(\xi_2 - \mathbb{E}(\xi_2))) + (a_2(\xi_2 - \mathbb{E}(\xi_2)))^2 \right) \\ &= \mathbb{E} \left( (a_1^2(\xi_1 - \mathbb{E}(\xi_1))^2 - 2a_1a_2(\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2)) + a_2^2(\xi_2 - \mathbb{E}(\xi_2))^2 \right) \\ &= a_1^2 \mathbb{E} \left( (\xi_1 - \mathbb{E}(\xi_1))^2 \right) - 2a_1a_2 \mathbb{E}((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))) + a_2^2 \mathbb{E}((\xi_2 - \mathbb{E}(\xi_2))^2) \\ &= a_1^2 \mathbb{V}(\xi_1) - 2a_1a_2 \mathbb{E}((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))) + a_2^2 \mathbb{V}(\xi_2) \\ &= \sum_{i=1}^2 a_i^2 \mathbb{V}(\xi_i) - 2a_1a_2 \mathbb{E}((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))) \end{aligned}$$

## Beweis (fortgeführt)

Weil  $\xi_1$  und  $\xi_2$  unabhängig sind, ergibt sich mit den Eigenschaften des Erwartungswerts für unabhängige Zufallsvariablen, dass

$$\begin{aligned}\mathbb{E}((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))) &= \mathbb{E}((\xi_1 - \mathbb{E}(\xi_1))) \mathbb{E}((\xi_2 - \mathbb{E}(\xi_2))) \\ &= (\mathbb{E}(\xi_1) - \mathbb{E}(\xi_1))(\mathbb{E}(\xi_2) - \mathbb{E}(\xi_2)) \\ &= 0\end{aligned}\tag{39}$$

ist. Damit folgt also

$$\mathbb{V}\left(\sum_{i=1}^2 a_i \xi_i\right) = \sum_{i=1}^2 a_i^2 \mathbb{V}(\xi_i).\tag{40}$$

Ein Induktionsargument erlaubt dann die Generalisierung vom bivariaten zum  $n$ -variaten Fall.

□

---

Erwartungswert

Varianz und Standardabweichung

**Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung**

Kovarianz und Korrelation

Stichprobenkovarianz und Stichprobenkorrelation

Selbstkontrollfragen

## Definition (Stichprobenmittel, -varianz, -standardabweichung)

$\xi_1, \dots, \xi_n$  seien Zufallsvariablen. Dann nennt man  $\xi_1, \dots, \xi_n$  auch eine *Stichprobe*.

- Das *Stichprobenmittel* von  $\xi_1, \dots, \xi_n$  ist definiert als der arithmetische Mittelwert

$$\bar{\xi}_n := \frac{1}{n} \sum_{i=1}^n \xi_i. \quad (41)$$

- Die *Stichprobenvarianz* von  $\xi_1, \dots, \xi_n$  ist definiert als

$$S_n^2 := \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi}_n)^2. \quad (42)$$

- Die *Stichprobenstandardabweichung* ist definiert als

$$S_n := \sqrt{S_n^2}. \quad (43)$$

## Bemerkungen

- $\mathbb{E}(\xi)$ ,  $\mathbb{V}(\xi)$ , und  $\mathbb{S}(\xi)$  sind Kennzahlen einer Zufallsvariable  $\xi$ .
- $\bar{\xi}_n$ ,  $S_n^2$ , und  $S_n$  sind Kennzahlen einer Stichprobe  $\xi_1, \dots, \xi_n$ .
- $\bar{\xi}_n$ ,  $S_n^2$ , und  $S_n$  sind Zufallsvariablen, ihre Realisationen werden mit  $\bar{x}_n$ ,  $s_n^2$ , und  $s_n$  bezeichnet.

## Beispiel (Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung)

- Es seien  $\xi_1, \dots, \xi_{10} \sim N(1, 2)$ .
- Wir nehmen die folgenden Realisationen an

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
0.54	1.01	-3.28	0.35	2.75	-0.51	2.32	1.49	0.96	1.25

- Die Stichprobenmittelrealisation ist

$$\bar{x}_{10} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{6.88}{10} = 0.68. \quad (44)$$

- Die Stichprobenvarianzrealisation ist

$$s_{10}^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x}_{10})^2 = \frac{1}{9} \sum_{i=1}^{10} (x_i - 0.68)^2 = \frac{25.37}{9} = 2.82. \quad (45)$$

- Die Stichprobenstandardabweichungrealisation ist

$$s_{10} = \sqrt{s_{10}^2} = \sqrt{2.82} = 1.68. \quad (46)$$

---

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

**Kovarianz und Korrelation**

Stichprobenkovarianz und Stichprobenkorrelation

Selbstkontrollfragen

## Definition (Kovarianz und Korrelation)

Die *Kovarianz* zweier Zufallsvariablen  $\xi$  und  $v$  ist definiert als

$$\mathbb{C}(\xi, v) := \mathbb{E}((\xi - \mathbb{E}(\xi))(v - \mathbb{E}(v))). \quad (47)$$

Die *Korrelation* zweier Zufallsvariablen  $\xi$  und  $v$  ist definiert als

$$\rho(\xi, v) := \frac{\mathbb{C}(\xi, v)}{\sqrt{\mathbb{V}(\xi)}\sqrt{\mathbb{V}(v)}} = \frac{\mathbb{C}(\xi, v)}{\mathbb{S}(\xi)\mathbb{S}(v)}. \quad (48)$$

### Bemerkungen

- Die Kovarianz von  $\xi$  mit sich selbst ist die Varianz von  $\xi$ ,

$$\mathbb{C}(\xi, \xi) = \mathbb{E}((\xi - \mathbb{E}(\xi))^2) = \mathbb{V}(\xi). \quad (49)$$

- $\rho(\xi, v)$  wird auch *Korrelationskoeffizient* von  $\xi$  und  $v$  genannt.
- Wenn  $\rho(\xi, v) = 0$  ist, werden  $\xi$  und  $v$  *unkorreliert* genannt.
- Wir zeigen später mit der Cauchy-Schwarz Ungleichung, dass  $-1 \leq \rho(\xi, v) \leq 1$ .



## Beispiel (Kovarianz und Korrelation zweier diskreter Zufallsvariablen)

Es sei  $\xi := (\xi_1, \xi_2)$  ein Zufallsvektor mit WMF  $p_{\xi_1, \xi_2}$  definiert durch

$p_{\xi_1, \xi_2}(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$p_{\xi_1}(x_1)$
$x_1 = 1$	0.10	0.05	0.15	0.30
$x_1 = 2$	0.60	0.05	0.05	0.70
$p_{\xi_2}(x_2)$	0.70	0.10	0.20	

$\xi_1, \xi_2$  sind also zwei Zufallsvariablen mit einer definierten bivariaten Verteilung. Um  $\mathbb{C}(\xi_1, \xi_2)$  und  $\rho(\xi_1, \xi_2)$  zu berechnen, halten wir zunächst fest, dass

$$\mathbb{E}(\xi_1) = \sum_{x_1=1}^2 x_1 p_{\xi_1}(x_1) = 1 \cdot 0.3 + 2 \cdot 0.7 = 1.7 \quad (50)$$

und

$$\mathbb{E}(\xi_2) = \sum_{x_2=1}^3 x_2 p_{\xi_2}(x_2) = 1 \cdot 0.7 + 2 \cdot 0.1 + 3 \cdot 0.2 = 1.5. \quad (51)$$

Mit der Definition der Kovarianz von  $\xi_1$  und  $\xi_2$ , gilt dann

$$\begin{aligned}C(\xi_1, \xi_2) &= \mathbb{E}((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))) \\&= \sum_{x_1=1}^2 \sum_{x_2=1}^3 (x_1 - \mathbb{E}(\xi_1))(x_2 - \mathbb{E}(\xi_2))p_{\xi_1, \xi_2}(x_1, x_2) \\&= \sum_{x_1=1}^2 \sum_{x_2=1}^3 (x_1 - 1.7)(x_2 - 1.5)p_{\xi_1, \xi_2}(x_1, x_2) \\&= \sum_{x_1=1}^2 (x_1 - 1.7)(1 - 1.5)p_{\xi_1, \xi_2}(x_1, 1) \\&\quad + (x_1 - 1.7)(2 - 1.5)p_{\xi_1, \xi_2}(x_1, 2) \\&\quad + (x_1 - 1.7)(3 - 1.5)p_{\xi_1, \xi_2}(x_1, 3) \\&= (1 - 1.7)(1 - 1.5)p_{\xi_1, \xi_2}(1, 1) + (1 - 1.7)(2 - 1.5)p_{\xi_1, \xi_2}(1, 2) + (1 - 1.7)(3 - 1.5)p_{\xi_1, \xi_2}(1, 3) \\&\quad + (2 - 1.7)(1 - 1.5)p_{\xi_1, \xi_2}(2, 1) + (2 - 1.7)(2 - 1.5)p_{\xi_1, \xi_2}(2, 2) + (2 - 1.7)(3 - 1.5)p_{\xi_1, \xi_2}(2, 3) \\&= (-0.7) \cdot (-0.5) \cdot 0.10 \quad + (-0.7) \cdot 0.5 \cdot 0.05 \quad + (-0.7) \cdot 1.5 \cdot 0.15 \\&\quad + 0.3 \cdot (-0.5) \cdot 0.60 \quad + 0.3 \cdot 0.5 \cdot 0.05 \quad + 0.3 \cdot 1.5 \cdot 0.05 \\&= 0.035 - 0.0175 - 0.1575 - 0.09 + 0.0075 + 0.0225 \\&= -0.2.\end{aligned}$$

## Theorem (Kovarianzverschiebungssatz)

$\xi$  und  $v$  seien Zufallsvariablen. Dann gilt

$$\mathbb{C}(\xi, v) = \mathbb{E}(\xi v) - \mathbb{E}(\xi)\mathbb{E}(v). \quad (52)$$

### Beweis

Mit der Definition der Kovarianz gilt

$$\begin{aligned} \mathbb{C}(\xi, v) &= \mathbb{E}((\xi - \mathbb{E}(\xi))(v - \mathbb{E}(v))) \\ &= \mathbb{E}(Xv - X\mathbb{E}(v) - \mathbb{E}(\xi)v + \mathbb{E}(\xi)\mathbb{E}(v)) \\ &= \mathbb{E}(\xi v) - \mathbb{E}(\xi)\mathbb{E}(v) - \mathbb{E}(\xi)\mathbb{E}(v) + \mathbb{E}(\xi)\mathbb{E}(v) \\ &= \mathbb{E}(\xi v) - \mathbb{E}(\xi)\mathbb{E}(v). \end{aligned} \quad (53)$$

□

### Bemerkungen

- Das Theorem ist nützlich, wenn  $\mathbb{E}(\xi v)$ ,  $\mathbb{E}(\xi)$ , und  $\mathbb{E}(v)$  leicht zu berechnen sind.
- Für  $v = \xi$  erhalten wir  $\mathbb{V}(\xi) = \mathbb{E}(\xi^2) - \mathbb{E}(\xi)^2$ .

## Theorem (Korrelation und Unabhängigkeit)

$\xi$  und  $v$  seien zwei Zufallsvariablen. Wenn  $\xi$  und  $v$  unabhängig sind, dann ist  $\mathbb{C}(\xi, v) = 0$  und  $\xi$  und  $v$  sind unkorreliert. Ist dagegen  $\mathbb{C}(\xi, v) = 0$  und sind  $\xi$  und  $v$  somit unkorreliert, dann sind  $\xi$  und  $v$  nicht notwendigerweise unabhängig.

### Beweis

Wir zeigen zunächst, dass aus der Unabhängigkeit von  $\xi$  und  $v$   $\mathbb{C}(\xi, v) = 0$  folgt. Hierzu halten wir zunächst fest, dass für unabhängige Zufallsvariablen gilt, dass

$$\mathbb{E}(\xi v) = \mathbb{E}(\xi)\mathbb{E}(v). \quad (54)$$

Mit dem Kovarianzverschiebungssatz folgt dann

$$\mathbb{C}(\xi, v) = \mathbb{E}(\xi v) - \mathbb{E}(\xi)\mathbb{E}(v) = \mathbb{E}(\xi)\mathbb{E}(v) - \mathbb{E}(\xi)\mathbb{E}(v) = 0. \quad (55)$$

Mit der Definition des Korrelationskoeffizienten folgt dann unmittelbar, dass  $\rho(\xi, v) = 0$  und  $\xi$  und  $v$  somit unkorreliert sind.

# Kovarianz und Korrelation

## Beweis (fortgeführt)

Wir zeigen nun durch Angabe eines Beispiels, dass die Kovarianz von abhängigen Zufallsvariablen  $\xi$  und  $v$  null sein kann.

Zu diesem Zweck betrachten wir den Fall zweier diskreter Zufallsvariablen  $\xi$  und  $v$  mit Ergebnisräumen  $\mathcal{X} = \{-1, 0, 1\}$  und  $v = \{0, 1\}$ , marginaler WMF von  $\xi$  gegeben durch  $p_\xi(x) := 1/3$  für  $x \in \mathcal{X}$  und der Definition  $v := \xi^2$ .

Wir halten dann zunächst fest, dass

$$\mathbb{E}(\xi) = \sum_{x \in \mathcal{X}} x p_\xi(x) = -1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0 \quad (56)$$

und

$$\mathbb{E}(\xi v) = \mathbb{E}(\xi \xi^2) = \mathbb{E}(\xi^3) = \sum_{x \in \mathcal{X}} x^3 p_\xi(x) = -1^3 \cdot \frac{1}{3} + 0^3 \cdot \frac{1}{3} + 1^3 \cdot \frac{1}{3} = 0. \quad (57)$$

Mit dem Kovarianzverschiebungssatz ergibt sich dann

$$\mathbb{C}(\xi, v) = \mathbb{E}(\xi v) - \mathbb{E}(\xi)\mathbb{E}(v) = \mathbb{E}(\xi^3) - \mathbb{E}(\xi)\mathbb{E}(v) = 0 - 0 \cdot \mathbb{E}(v) = 0. \quad (58)$$

Die Kovarianz von  $\xi$  und  $v$  ist also null. Wie unten gezeigt faktorisiert die gemeinsame WMF von  $\xi$  und  $v$  jedoch nicht, und somit sind  $\xi$  und  $v$  nicht unabhängig.

## Beweis (fortgeführt)

Die Definition von  $v := \xi^2$  impliziert die folgende bedingte WMF

$p_{v \xi}(y x)$	$x = -1$	$x = 0$	$x = 1$
$y = 0$	0	1	0
$y = 1$	1	0	1

Die marginale WMF  $p_\xi$  und die bedingte WMF  $p_{v|\xi}$  implizieren die gemeinsame WMF

$p_{\xi,v}(x,y)$	$x = -1$	$x = 0$	$x = 1$	$p_v(y)$
$y = 0$	0	1/3	0	1/3
$y = 1$	1/3	0	1/3	2/3
$p_\xi(x)$	1/3	1/3	1/3	

Es gilt also zum Beispiel

$$p_{\xi,v}(-1, 0) = 0 \neq \frac{1}{9} = \frac{1}{3} \cdot \frac{1}{3} = p_\xi(-1)p_v(0) \quad (59)$$

und damit sind  $\xi$  und  $v$  nicht unabhängig.

□

## Theorem (Varianzen von Summen und Differenzen von Zufallsvariablen)

$\xi$  und  $v$  seien zwei Zufallsvariablen und es seien  $a, b, c \in \mathbb{R}$ . Dann gilt

$$\mathbb{V}(a\xi + bv + c) = a^2\mathbb{V}(\xi) + b^2\mathbb{V}(v) + 2ab\mathbb{C}(\xi, v). \quad (60)$$

Speziell gelten

$$\mathbb{V}(X + v) = \mathbb{V}(\xi) + \mathbb{V}(v) + 2\mathbb{C}(\xi, v) \quad (61)$$

und

$$\mathbb{V}(X - v) = \mathbb{V}(\xi) + \mathbb{V}(v) - 2\mathbb{C}(\xi, v) \quad (62)$$

### Bemerkungen

- Varianzen von Zufallsvariablen addieren sich nicht einfach.
- Die Varianz der Summe zweier Zufallsvariablen hängt von ihrer Kovarianz ab.

## Beweis

Wir halten zunächst fest, dass

$$\mathbb{E}(a\xi + bv + c) = a\mathbb{E}(\xi) + b\mathbb{E}(v) + c. \quad (63)$$

Es ergibt sich also

$$\begin{aligned} & \mathbb{V}(a\xi + bv + c) \\ &= \mathbb{E} \left( (a\xi + bv + c - a\mathbb{E}(\xi) - b\mathbb{E}(v) - c)^2 \right) \\ &= \mathbb{E} \left( (a(X - \mathbb{E}(\xi)) + b(v - \mathbb{E}(v)))^2 \right) \\ &= \mathbb{E} \left( a^2(\xi - \mathbb{E}(\xi))^2 + b^2(v - \mathbb{E}(v))^2 + 2ab(\xi - \mathbb{E}(\xi))(v - \mathbb{E}(v)) \right) \\ &= a^2\mathbb{E} \left( (\xi - \mathbb{E}(\xi))^2 \right) + b^2\mathbb{E} \left( (v - \mathbb{E}(v))^2 \right) + 2ab\mathbb{E}((\xi - \mathbb{E}(\xi))(v - \mathbb{E}(v))) \\ &= a^2\mathbb{V}(\xi) + b^2\mathbb{V}(v) + 2ab\mathbb{C}(\xi, v) \end{aligned} \quad (64)$$

Die Spezialfälle folgen dann direkt mit  $a := b := 1$  und  $a := 1, b := -1$ , respektive.

□



---

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

**Stichprobenkovarianz und Stichprobenkorrelation**

Selbstkontrollfragen

## Definition (Stichprobenkovarianz und Stichprobenkorrelation)

$(\xi_1, v_1), \dots, (\xi_n, v_n)$  seien zweidimensionale Zufallsvektoren.

- Das Stichprobenmittel der  $(\xi_1, v_1), \dots, (\xi_n, v_n)$  ist definiert als

$$\overline{(\xi, v)}_n := (\bar{\xi}_n, \bar{v}_n) = \left( \frac{1}{n} \sum_{i=1}^n \xi_i, \frac{1}{n} \sum_{i=1}^n v_i \right). \quad (65)$$

- Die Stichprobenkovarianz der  $(\xi_1, v_1), \dots, (\xi_n, v_n)$  ist definiert als

$$C_n := \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi}_n)(v_i - \bar{v}_n). \quad (66)$$

- Der Stichprobenkorrelationskoeffizient der  $(\xi_1, v_1), \dots, (\xi_n, v_n)$  ist definiert als

$$R_n := \frac{C_n}{S_{\xi,n} S_{v,n}}, \quad (67)$$

wobei  $S_{\xi,n}$  und  $S_{v,n}$  die Stichprobenstandardabweichungen von  $\xi_1, \dots, \xi_n$  und  $v_1, \dots, v_n$ , bezeichnen.

### Bemerkungen

- $\overline{(\xi, v)}_n$ ,  $C_n$ , und  $R_n$  sind Zufallsvariablen
- $\overline{(x, y)}_n$ ,  $c_n$ , und  $r_n$  bezeichnen ihre Realisierungen

# Stichprobenkovarianz und Stichprobenkorrelation

## Beispiel (Stichprobenkovarianz und Stichprobenkorrelation)

- Es seien  $(\xi_1, v_1), \dots, (\xi_{10}, v_{10})$  zweidimensionale Zufallsvariablen.
- Wir nehmen folgende Realisierungen an

$(x_1, y_1)$	$(x_2, y_2)$	$(x_3, y_3)$	$(x_4, y_4)$	$(x_5, y_5)$	$(x_6, y_6)$	$(x_7, y_7)$	$(x_8, y_8)$	$(x_9, y_9)$	$(x_{10}, y_{10})$
(0.8, -0.7)	(1.1, 1.6)	(-0.8, 1.1)	(-0.2, 0.1)	(1.1, 0.4)	(0.5, 1.5)	(1.3, -1.2)	(1.8, 0.6)	(0.4, 0.2)	(1.5, -1.0)

- Die Stichprobenmittelrealisation ergibt sich zu

$$\overline{(x, y)}_{10} = (\bar{x}_{10}, \bar{y}_{10}) = \left( \frac{1}{10} \sum_{i=1}^{10} x_i, \frac{1}{10} \sum_{i=1}^{10} y_i \right) = (0.75, 0.26). \quad (68)$$

- Die Stichprobenstandardabweichungrealisationen ergeben sich zu

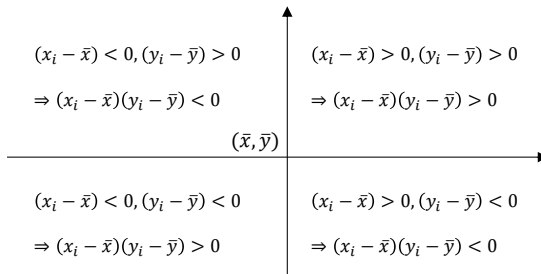
$$s_{\xi, n} = \sqrt{\frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x}_{10})^2} = 0.79 \text{ und } s_{v, n} = \sqrt{\frac{1}{9} \sum_{i=1}^{10} (y_i - \bar{y}_{10})^2} = 0.99. \quad (69)$$

- Die Stichprobenkovarianz- und Stichprobenkorrelationsrealisationen ergeben sich zu

$$c_n = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)(y_i - \bar{y}_n) = -0.26 \text{ und } r_n = \frac{c_n}{s_{x, n} s_{y, n}} = -0.33. \quad (70)$$

# Stichprobenkovarianz und Stichprobenkorrelation

## Zur Mechanik der Kovariationsterme



Häufige richtungsgleiche Abweichung der  $x_i$  und  $y_i$  von ihren Mittelwerten  $\Rightarrow$  Positive Korrelation

Häufige richtungsungleiche Abweichung der  $x_i$  und  $y_i$  von ihren Mittelwerten  $\Rightarrow$  Negative Korrelation

Keine häufigen richtungsgleichen oder -entgegengesetzten Abweichungen  $\Rightarrow$  Keine Korrelation

Eine anwendungsorientierte Diskussion der Stichprobenkorrelation und ihren engen Bezug zur einfachen linearen Regression gibt es im Rahmen der Einführung des Allgemeinen Linearen Modells, siehe [Regression](#) und [Korrelation](#).

---

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Stichprobenkovarianz und Stichprobenkorrelation

**Selbstkontrollfragen**

# Selbstkontrollfragen

---

1. Definieren und interpretieren Sie den Erwartungswert einer Zufallsvariable.
2. Berechnen Sie den Erwartungswert einer Bernoulli Zufallsvariable.
3. Nennen Sie drei Eigenschaften des Erwartungswerts.
4. Definieren und interpretieren Sie die Varianz einer Zufallsvariable.
5. Berechnen Sie die Varianz einer Bernoulli Zufallsvariable.
6. Drücken Sie  $\mathbb{E}(\xi^2)$  mithilfe der Varianz und des Erwartungswerts von  $\xi$  aus.
7. Was ist  $\mathbb{V}(a\xi)$  für konstantes  $a \in \mathbb{R}$ ?
8. Definieren Sie die Kovarianz und Korrelation zweier Zufallsvariablen  $\xi$  und  $\nu$ .
9. Geben Sie das Theorem zur Varianz von Linearkombinationen von Zufallsvariablen bei Unabhängigkeit wieder.
10. Definieren Sie den Begriff der Stichprobe.
11. Definieren Sie den Begriff des Stichprobenmittels.
12. Definieren Sie Stichprobenvarianz und Stichprobenstandardabweichung.

13. Erläutern Sie die Unterschiede zwischen dem Erwartungswertparameter, dem Erwartungswert und dem Stichprobenmittel von normalverteilten Zufallsvariablen.
14. Definieren Sie die Kovarianz und die Korrelation zweier Zufallsvariablen.
15. Schreiben Sie die Kovarianz zweier Zufallsvariablen mithilfe von Erwartungswerten.
16. Geben Sie das Theorem zur Korrelation und Unabhängigkeit zweier Zufallsvariablen wieder.
17. Was ist die Varianz der Summe zweier Zufallsvariablen bei Unabhängigkeit?
18. Was ist die Varianz der Summe zweier Zufallsvariablen im Allgemeinen?
19. Definieren Sie das Stichprobenmittel für eine Stichprobe zweidimensionaler Zufallsvektoren.
20. Definieren Sie die Stichprobenkovarianz einer Stichprobe von zweidimensionaler Zufallsvektoren.
21. Wann ergeben sich für die Stichprobenkovarianz hohe positive oder hohe negative Werte?
22. Wann ergeben sich für die Stichprobenkovarianzwerte nahe Null?
23. Definieren Sie den Stichprobenkorrelationskoeffizienten.