Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2024/25

Prof. Dr. Dirk Ostwald

(5) Erwartungswerte

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Selbstkontrollfragen

Definition (Erwartungswert)

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

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

Beispiel (Erwartungswert einer diskreten Zufallsvariable)

ξ sei eine diskrete Zufallsvariable, die Werte in $\mathcal{X} := \{-1, 0, 1\}$ annehme und deren WMF p gegeben sei durch

$$p(-1) = \frac{1}{4}, \quad p(0) = \frac{1}{2}, \quad p(1) = \frac{1}{4}. \tag{1}$$

Dann gilt

$$\mathbb{E}(\xi) = 0. \tag{2}$$

<u>Beweis</u>

ξ ist eine diskrete Zufallsvariable. Also gilt

$$\begin{aligned}
\mathbb{E}(\xi) &= \sum_{x \in \mathcal{X}} x\, p(x) \\
&= -1 \cdot p(-1) + 0 \cdot p(0) + 1 \cdot p(1) \\
&= -1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} \\
&= 0.
\end{aligned} \tag{3}$$

$\square$

Beispiel (Erwartungswert einer Bernoulli Zufallsvariable)

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

Beweis

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

$$
\begin{aligned}
\mathbb{E}(\xi) &= \sum_{x \in \{0,1\}} x \, \text{Bern}(x; \mu) \\
&= 0 \cdot \mu^0 (1-\mu)^{1-0} + 1 \cdot \mu^1 (1-\mu)^{1-1} \\
&= 1 \cdot \mu^1 (1-\mu)^0 \\
&= \mu.
\end{aligned} \tag{4}
$$

$\square$

Beispiel (Erwartungswert einer normalverteilten Zufallsvariable)

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

Bemerkungen

- Der Beweis dieser Tatsache ist mathematisch aufwändig.
- Für eine Skizze des Beweises verweisen wir auf den Anhang.

Theorem (Einige Erwartungswerte)

Zufallsvariable	$\mathbb{E}(\xi)$
$\xi \sim \text{Bern}(\mu)$	μ
$\xi \sim \text{Bin}(\mu, n)$	$n\mu$
$\xi \sim N(\mu, \sigma^2)$	μ
$\xi \sim G(\alpha, \beta)$	$\alpha\beta$
$\xi \sim \text{Beta}(\alpha, \beta)$	$\frac{\alpha}{\alpha+\beta}$
$\xi \sim U(a, b)$	$\frac{a+b}{2}$

Bemerkung

- Zufallsvariablen, deren WMF/WDFen von wenigen Parametern abhängen, heißen auch *parametrisch*.
- Erwartungswerte von parametrischen Zufallsvariablen sind typischerweise Funktionen der Parameter.
- Wir verzichten auf Beweise.

Definition (Erwartungswert einer Funktion einer Zufallsvariable)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum, ξ sei eine Zufallsvariable mit Ergebnisraum $\mathcal{X}$ und $f : \mathcal{X} \to \mathcal{Z}$ sei eine Funktion mit Zielmenge $\mathcal{Z}$. Dann ist der *Erwartungswert der Funktion* f *der Zufallsvariable* ξ definiert als

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

Bemerkung

- Der Erwartungswert einer Zufallsvariable ist der Spezialfall mit

$$f : \mathcal{X} \to \mathcal{Z}, x \mapsto f(x) := x. \tag{5}$$

Definition (Erwartungswert eines Zufallsvektors)

ξ sei ein n-dimensionaler Zufallvektor. Dann ist der *Erwartungwert* von ξ definiert als der n-dimensionale reelle Vektor

$$\mathbb{E}(\xi) := \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n) \end{pmatrix}. \tag{6}$$

Bemerkung

- Der Erwartungswert eines Zufallsvektors ist der Vektor der Erwartungswerte seiner Komponenten.

Definition (Erwartungswert einer Funktion eines Zufallsvektors)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum, ξ sei ein Zufallsvektor mit Ergebnisraum $\mathcal{X}$ und $f : \mathcal{X} \to \mathcal{Z}$ sei eine Funktion mit Zielmenge $\mathcal{Z}$. Dann ist der *Erwartungswert der Funktion f des Zufallsvektors ξ* definiert als

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

Bemerkung

- Die Definition ist anlog zur Definition des Erwartungswerts einer Funktion einer Zufallsvariable.

Theorem (Eigenschaften des Erwartungswerts)

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

$$\mathbb{E}(a\xi + b) = a\mathbb{E}(\xi) + b. \tag{7}$$

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

$$\mathbb{E}\left(\sum_{i=1}^n a_i \xi_i\right) = \sum_{i=1}^n a_i \mathbb{E}(\xi_i). \tag{8}$$

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

$$\mathbb{E}\left(\prod_{i=1}^n \xi_i\right) = \prod_{i=1}^n \mathbb{E}(\xi_i). \tag{9}$$

Bemerkung

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

Erwartungswert

<u>Beweis</u>

Eigenschaft (1) folgt aus den Linearitätseigenschaften von Summen und Integralen. Wir betrachten nur den Fall einer kontinuierlichen Zufallsvariable ξ mit WDF p genauer und definieren zunächst $\upsilon := a\xi + b$. Dann gilt

$$
\begin{aligned}
\mathbb{E}(\upsilon) &= \mathbb{E}(a\xi + b) \\
&= \int_{-\infty}^{\infty} (ax + b)p(x)\,dx \\
&= \int_{-\infty}^{\infty} axp(x) + bp(x)\,dx \\
&= a\int_{-\infty}^{\infty} xp(x)\,dx + b\int_{-\infty}^{\infty} p(x)\,dx \\
&= a\mathbb{E}(\xi) + b.
\end{aligned}
\tag{10}
$$

Erwartungswert

Beweis (fortgeführt)

Eigenschaft (2) folgt gleichfalls aus den Linearitätseigenschaften von Summen und Integralen. Wir betrachten nur den Fall von zwei kontinuierlichen Zufallsvariablen ξ_1 und ξ_2 mit bivariater WDF genauer. In diesem Fall gilt

$$
\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{2} a_i\xi_i\right) &= \mathbb{E}(a_1\xi_1 + a_2\xi_2) \\
&= \iint_{\mathbb{R}^2} (a_1x_1 + a_2x_2)p(x_1,x_2)\,dx_1\,dx_2 \\
&= \iint_{\mathbb{R}^2} a_1x_1p(x_1,x_2) + a_2x_2p(x_1,x_2)\,dx_1\,dx_2 \\
&= a_1 \iint_{\mathbb{R}^2} x_1p(x_1,x_2)\,dx_1\,dx_2 + a_2 \iint_{\mathbb{R}^2} x_2p(x_1,x_2)\,dx_1\,dx_2 \\
&= a_1 \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{\infty} p(x_1,x_2)\,dx_2\right) dx_1 + a_2 \int_{-\infty}^{\infty} x_2 \left(\int_{-\infty}^{\infty} p(x_1,x_2)\,dx_1\right) dx_2 \\
&= a_1 \int_{-\infty}^{\infty} x_1p_{\xi_1}(x_1)\,dx_1 + a_2 \int_{-\infty}^{\infty} x_2p_{\xi_2}(x_2)\,dx_2 \\
&= a_1\mathbb{E}(\xi_1) + a_2\mathbb{E}(\xi_2) \\
&= \sum_{i=1}^{2} a_i\mathbb{E}(\xi_i).
\end{aligned}
$$

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

Beweis (fortgeführt)

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

$$p(x_1, ..., x_n) = \prod_{i=1}^n p(x_i). \tag{11}$$

Weiterhin gilt also

$$\begin{aligned}
\mathbb{E}\left(\prod_{i=1}^n \xi_i\right) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{i=1}^n x_i\right) p(x_1, ..., x_n)\, dx_1 \dots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^n x_i \prod_{i=1}^n p(x_i)\, dx_1 \dots dx_n \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{i=1}^n x_i p(x_i)\, dx_1 \dots dx_n \\
&= \prod_{i=1}^n \int_{-\infty}^{\infty} x_i p(x_i)\, dx_i \\
&= \prod_{i=1}^n \mathbb{E}(\xi_i).
\end{aligned} \tag{12}$$

$\square$

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Selbstkontrollfragen

Definition (Varianz und Standardabweichung)

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

$$\mathbb{V}(\xi) := \mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right), \tag{13}$$

unter der Annahme, dass dieser Erwartungswert existiert. Die *Standardabweichung von* ξ ist definiert

$$\mathbb{S}(\xi) := \sqrt{\mathbb{V}(\xi)}. \tag{14}$$

Bemerkungen

- Die Varianz misst die Streuung (Breite) einer Verteilung im Sinne der *Chebyshev Ungleichung*.
- Quadration 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 $-\mathbb{E}(\ln p)$.

Beispiel (Varianz und Standardabweichung einer diskreten Zufallsvariable)

ξ sei eine diskrete Zufallsvariable, die Werte in $\mathcal{X} := \{-1, 0, 1\}$ annehme und deren WMF p gegeben sei durch

$$p(-1) = \frac{1}{4}, \quad p(0) = \frac{1}{2}, \quad p(1) = \frac{1}{4}. \tag{15}$$

Dann gelten

$$\mathbb{V}(\xi) = \frac{1}{2} \text{ und } \mathbb{S}(\xi) = \frac{1}{\sqrt{2}}. \tag{16}$$

<u>Beweis</u>

Es gilt

$$\begin{aligned}
\mathbb{V}(\xi) &= \mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) \\
&= \mathbb{E}\left((\xi - 0)^2\right) \\
&= \mathbb{E}\left(\xi^2\right) \\
&= \sum_{x \in \mathcal{X}} x^2 \, p(x) \\
&= (-1)^2 \cdot p(-1) + 0^2 \cdot p(0) + 1^2 \cdot p(1) \\
&= 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} \\
&= \frac{1}{2}.
\end{aligned} \tag{17}$$

Die Form der Standardabweichung folgt dann direkt. □

Beispiel (Varianz und Standardabweichung einer Bernoulli Zufallsvariable)

Es sei $\xi \sim \text{Bern}(\mu)$. Dann gelten

$$\mathbb{V}(\xi) = \mu(1-\mu) \text{ und } \mathbb{S}(\xi) = \sqrt{\mu(1-\mu)}. \tag{18}$$

Beweis

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

$$\begin{aligned}
\mathbb{V}(\xi) &= \mathbb{E}\left((\xi-\mu)^2\right) \\
&= \sum_{x \in \{0,1\}} (x-\mu)^2 \text{Bern}(x;\mu) \\
&= (0-\mu)^2 \mu^0 (1-\mu)^{1-0} + (1-\mu)^2 \mu^1 (1-\mu)^{1-1} \\
&= \mu^2(1-\mu) + (1-\mu)^2 \mu \\
&= \left(\mu^2 + (1-\mu)\mu\right)(1-\mu) \\
&= \left(\mu^2 + \mu - \mu^2\right)(1-\mu) \\
&= \mu(1-\mu).
\end{aligned} \tag{19}$$

Die Form der Standardabweichung folgt dann direkt. $\square$

Varianz und Standardabweichung

Beispiel (Varianz und Standardabweichung einer normalverteilten Zufallsvariable)

Es sei $\xi \sim N(\mu, \sigma^2)$. Dann gelten

$$\mathbb{V}(\xi) = \sigma^2 \text{ und } \mathbb{S}(\xi) = \sigma. \tag{20}$$

Bemerkungen

- Der Beweis dieser Tatsache ist mathematisch aufwändig.
- Für eine Skizze des Beweises verweisen wir auf den Anhang.

Theorem (Varianzverschiebungssatz)

Es sei ξ eine Zufallsvariable. Dann gilt

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

Beweis

Mit der Definition der Varianz und der Linearität des Erwartungswerts gilt

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

$\square$

Bemerkung

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

Theorem (Einige Erwartungswerte, Varianzen, Standardabweichungen)

Zufallsvariable	$\mathbb{E}(\xi)$	$\mathbb{V}(\xi)$	$\mathbb{S}(\xi)$
$\xi \sim \text{Bern}(\mu)$	μ	$\mu(1-\mu)$	$\sqrt{\mu(1-\mu)}$
$\xi \sim \text{Bin}(\mu, n)$	$n\mu$	$n\mu(1-\mu)$	$\sqrt{n\mu(1-\mu)}$
$\xi \sim N(\mu, \sigma^2)$	μ	σ^2	σ
$\xi \sim G(\alpha, \beta)$	$\alpha\beta$	$\alpha\beta^2$	$\sqrt{\alpha}\beta$
$\xi \sim \text{Beta}(\alpha, \beta)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$\sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}$
$\xi \sim U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\sqrt{\frac{(b-a)^2}{12}}$

Bemerkungen

- Varianzen parametrischer Zufallsvariablen sind typischerweise Funktionen der Parameter.
- Wir verzichten auf Beweise.

Theorem (Eigenschaften der Varianz)

(1) (Nichtnegativität) ξ sei eine Zufallsvariable. Dann gilt

$$\mathbb{V}(\xi) \geq 0. \tag{23}$$

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

$$\mathbb{V}(a\xi + b) = a^2\mathbb{V}(\xi) \text{ und } \mathbb{S}(a\xi + b) = |a|\mathbb{S}(\xi). \tag{24}$$

(3) (Linearkombination bei Unabhängigkeit) $\xi_1, ..., \xi_n$ seien unabhängige Zufallsvariable und es seien $a_1, ..., a_n \in \mathbb{R}$. Dann gilt

$$\mathbb{V}\left(\sum_{i=1}^n a_i\xi_i\right) = \sum_{i=1}^n a_i^2\mathbb{V}(\xi_i). \tag{25}$$

Varianz und Standardabweichung

Beweis

Hinsichtlich (1) betrachten wir den Fall einer kontinuierlichen Zufallsvariable. Dann gilt zunächst

$$(x - \mathbb{E}(\xi))^2 \geq 0 \text{ für alle } x \in \mathbb{R}. \tag{26}$$

Weiterhin gilt für die WDF p von ξ, dass

$$p(x) \geq 0 \text{ für alle } x \in \mathbb{R}. \tag{27}$$

Also folgt

$$p(x)(x - \mathbb{E}(\xi))^2 \geq 0 \text{ für alle } x \in \mathbb{R}. \tag{28}$$

Damit gilt dann aber

$$\mathbb{V}(\xi) = \int_{-\infty}^{\infty} p(x)(x - \mathbb{E}(\xi))^2 \, dx \geq 0, \tag{29}$$

denn das Riemann Integral jeder nicht-negativen Funktion ist nicht-negativ. Analog zeigt man die Nichtnegativität der Varianz bei diskreten Zufallsvariablen.

Varianz und Standardabweichung

Beweis (fortgeführt)

Um Eigenschaft (2) zu zeigen, definieren wir zunächst $\upsilon := a\xi + b$ und halten fest, dass $\mathbb{E}(\upsilon) = a\mathbb{E}(\xi) + b$. Für die Varianz von υ ergibt sich dann

$$
\begin{aligned}
\mathbb{V}(\upsilon) &= \mathbb{E}\left((\upsilon - \mathbb{E}(\upsilon))^2\right) \\
&= \mathbb{E}\left((a\xi + b - a\mathbb{E}(\xi) - b)^2\right) \\
&= \mathbb{E}\left((a\xi - a\mathbb{E}(\xi))^2\right) \\
&= \mathbb{E}\left((a(\xi - \mathbb{E}(\xi))^2\right) \\
&= \mathbb{E}\left(a^2(\xi - \mathbb{E}(\xi))^2\right) \\
&= a^2\mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) \\
&= a^2\mathbb{V}(\xi)
\end{aligned}
\tag{30}
$$

Wurzelziehen ergibt dann das Resultat für die Standardabweichung.

Varianz und Standardabweichung

Beweis (fortgeführt)

Für Eigenschaft (3) betrachten wir den Fall zweier unabhängiger Zufallsvariablen ξ_1 und ξ_2 genauer. Wir halten zunächst fest, dass in diesem Fall gilt, dass

$$\mathbb{E}\left(a_1\xi_1 + a_2\xi_2\right) = a_1\mathbb{E}(\xi_1) + a_2\mathbb{E}(\xi_2). \tag{31}$$

Es ergibt sich also

$$\begin{aligned}
\mathbb{V}\left(\sum_{i=1}^{2} a_i\xi_i\right) & = \mathbb{V}(a_1\xi_1 + a_2\xi_2) \\
& = \mathbb{E}\left(\left(a_1\xi_1 + a_2\xi_2 - \mathbb{E}\left(a_1\xi_1 + a_2\xi_2\right)\right)^2\right) \\
& = \mathbb{E}\left(\left(a_1\xi_1 + a_2\xi_2 - a_1\mathbb{E}(\xi_1) - a_2\mathbb{E}(\xi_2)\right)^2\right) \\
& = \mathbb{E}\left(\left(a_1\xi_1 - a_1\mathbb{E}(\xi_1) + a_2\xi_2 - a_2\mathbb{E}(\xi_2)\right)^2\right) \\
& = \mathbb{E}\left(\left((a_1(\xi_1 - \mathbb{E}(\xi_1)) + (a_2(\xi_2 - \mathbb{E}(\xi_2))\right)^2\right) \\
& = \mathbb{E}\left((a_1(\xi_1 - \mathbb{E}(\xi_1)))^2 + 2(a_1(\xi_1 - \mathbb{E}(\xi_1))(a_2(\xi_2 - \mathbb{E}(\xi_2)) + (a_2(\xi_2 - \mathbb{E}(\xi_2)))^2\right) \\
& = \mathbb{E}\left((a_1^2(\xi_1 - \mathbb{E}(\xi_1))^2 + 2a_1a_2(\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2)) + a_2^2(\xi_2 - \mathbb{E}(\xi_2))^2\right) \\
& = a_1^2\mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))^2\right) + 2a_1a_2\mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))\right) + a_2^2\mathbb{E}\left((\xi_2 - \mathbb{E}(\xi_2))^2\right) \\
& = a_1^2\mathbb{V}(\xi_1) + 2a_1a_2\mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))\right) + a_2^2\mathbb{V}(\xi_2) \\
& = \sum_{i=1}^{2} a_i^2\mathbb{V}(\xi_i) + 2a_1a_2\mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))\right).
\end{aligned} \tag{32}$$

Beweis (fortgeführt)

Weil ξ_1 und ξ_2 unabhängig sind, ergibt sich mit den Eigenschaften des Erwartungswerts für unabhängige Zufallsvariablen, dass

$$
\begin{aligned}
\mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))(\xi_2 - \mathbb{E}(\xi_2))\right) &= \mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))\right)\mathbb{E}\left((\xi_2 - \mathbb{E}(\xi_2))\right) \\
&= (\mathbb{E}(\xi_1) - \mathbb{E}(\xi_1))(\mathbb{E}(\xi_2) - \mathbb{E}(\xi_2)) \\
&= 0
\end{aligned} \tag{33}
$$

ist. Damit folgt also

$$
\mathbb{V}\left(\sum_{i=1}^{2} a_i\xi_i\right) = \sum_{i=1}^{2} a_i^2\mathbb{V}(\xi_i). \tag{34}
$$

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

$\square$

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Selbstkontrollfragen

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Definition (Stichprobenmittel, -varianz, -standardabweichung)

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

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

$$\bar{\xi} := \frac{1}{n}\sum_{i=1}^{n}\xi_i. \tag{35}$$

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

$$S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2. \tag{36}$$

- Die *Stichprobenstandardabweichung* ist definiert als

$$S := \sqrt{S^2}. \tag{37}$$

Bemerkungen

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

Beispiel (Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung)

- Es seien $\xi_1, ..., \xi_{10} \sim N(1,2)$.
- 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

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{6.88}{10} = 0.68. \tag{38}$$

- Die Stichprobenvarianzrealisation ist

$$s^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - \bar{x})^2 = \frac{1}{9}\sum_{i=1}^{10}(x_i - 0.68)^2 = \frac{25.37}{9} = 2.82. \tag{39}$$

- Die Stichprobenstandardabweichungrealisation ist

$$s = \sqrt{s^2} = \sqrt{2.82} = 1.68. \tag{40}$$

Zum Sinn der (Stichproben)Standardabweichung

- Offenbar sind (Stichproben)Varianz und (Stichproben)Standardabweichung äquivalent.
- Die Varianz ist historisch eng mit dem Exponentialargument der Normalverteilung verwoben.
- Die Varianz ist historisch eng mit den Quadratsummen der Varianzanalyse verwoben.
- Die Varianz misst Variabilität allerdings im Sinne quadratiert Einheiten, z.B.

$$s^2 = \frac{1}{9}\sum_{i=1}^{10}\left(x_i \text{ m} - 0.68 \text{ m}\right)^2 = \frac{1}{9}\cdot 25.37 \text{ m}^2 = 2.82 \text{ m}^2. \tag{41}$$

- Das Maß für die Variabilität von Längen in Meter hätte hier die Einheit einer Fläche.
- Wurzelziehen korrigiert die Einheiten dann zurück auf die Einheit der Daten, z.B.

$$s = \sqrt{2.82 \text{ m}^2} = 1.68 \text{ m}. \tag{42}$$

- In welchem Sinn die Varianz Variabilität misst, ergibt sich durch die *Chebyshev Ungleichung*.

Zur Prävention von Verwechselungen

- Der *Erwartungswert* $\mathbb{E}(\xi)$, die *Varianz* $\mathbb{V}(\xi)$ und die Standardabweichung $\mathbb{S}(\xi)$ sind Kennzahlen von Zufallsvariablen und werden basierend auf der Verteilung (WMF/WDF) einer Zufallsvariable bestimmt.

- Das *Stichprobenmittel* $\bar{\xi}$, die *Stichprobenvarianz* S^2 und die *Stichprobenstandardabweichung* S sind Kennzahlen von Stichproben und werden basierend auf Stichproben, insbesondere ihrer Realisierungen, berechnet.

- Der *Erwartungswertparameter* μ und der *Varianzparameter* σ^2 sind Parameter der Normalverteilung und werden im Rahmen der probabilistischen Inferenz definiert oder geschätzt.

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Selbstkontrollfragen

Definition (Kovarianz und Korrelation)

Die *Kovarianz* zweier Zufallsvariablen ξ und υ ist definiert als

$$\mathbb{C}(\xi, \upsilon) := \mathbb{E}\left((\xi - \mathbb{E}(\xi))\,(\upsilon - \mathbb{E}(\upsilon))\right). \tag{43}$$

Die *Korrelation* zweier Zufallsvariablen ξ und υ ist definiert als

$$\rho(\xi, \upsilon) := \frac{\mathbb{C}(\xi, \upsilon)}{\sqrt{\mathbb{V}(\xi)}\sqrt{\mathbb{V}(\upsilon)}} = \frac{\mathbb{C}(\xi, \upsilon)}{\mathbb{S}(\xi)\mathbb{S}(\upsilon)}. \tag{44}$$

Bemerkungen

- Die Kovarianz von ξ mit sich selbst ist die Varianz von ξ,

$$\mathbb{C}(\xi, \xi) = \mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) = \mathbb{V}(\xi). \tag{45}$$

- $\rho(\xi, \upsilon)$ wird auch *Korrelationskoeffizient* von ξ und υ genannt.
- Wenn $\rho(\xi, \upsilon) = 0$ ist, werden ξ und υ *unkorreliert* genannt.
- Wir zeigen später mit der Cauchy-Schwarz Ungleichung, dass $-1 \le \rho(\xi, \upsilon) \le 1$.
- Die Begriffe der Stichprobenkovarianz und der Stichprobenkorrelation führen wir erst an späterer Stelle ein.

Beispiel (Kovarianz und Korrelation zweier diskreter Zufallsvariablen)

Es sei $\zeta := (\xi, \upsilon)$ ein Zufallsvektor mit WMF p definiert durch

$p(x,y)$	$y=1$	$y=2$	$y=3$	$p(x)$
$x=1$	0.10	0.05	0.15	0.30
$x=2$	0.60	0.05	0.05	0.70
$p(y)$	0.70	0.10	0.20	

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

$$\mathbb{E}(\xi) = \sum_{x=1}^{2} xp(x) = 1 \cdot 0.3 + 2 \cdot 0.7 = 1.7 \tag{46}$$

und

$$\mathbb{E}(\upsilon) = \sum_{y=1}^{3} yp(y) = 1 \cdot 0.7 + 2 \cdot 0.1 + 3 \cdot 0.2 = 1.5. \tag{47}$$

Mit der Definition der Kovarianz von ξ und υ, gilt dann

$$
\begin{aligned}
\mathbb{C}(\xi,\upsilon) &= \mathbb{E}((\xi-\mathbb{E}(\xi))(\upsilon-\mathbb{E}(\upsilon))) \\
&= \sum_{x=1}^{2}\sum_{y=1}^{3}(x-\mathbb{E}(\xi))(y-\mathbb{E}(\upsilon))p(x,y) \\
&= \sum_{x=1}^{2}\sum_{y=1}^{3}(x-1.7)(y-1.5)p(x,y) \\
&= \sum_{x=1}^{2}(x-1.7)(1-1.5)p(x,1) \\
&\qquad + (x-1.7)(2-1.5)p(x,2) \\
&\qquad + (x-1.7)(3-1.5)p(x,3) \\
&= \quad (1-1.7)(1-1.5)p(1,1)+(1-1.7)(2-1.5)p(1,2)+(1-1.7)(3-1.5)p(1,3) \\
&\quad + (2-1.7)(1-1.5)p(2,1)+(2-1.7)(2-1.5)p(2,2)+(2-1.7)(3-1.5)p(2,3) \\
&= \quad (-0.7)\cdot(-0.5)\cdot 0.10+(-0.7)\cdot 0.5\cdot 0.05+(-0.7)\cdot 1.5\cdot 0.15 \\
&\quad + 0.3\cdot(-0.5)\cdot 0.60+0.3\cdot 0.5\cdot 0.05+0.3\cdot 1.5\cdot 0.05 \\
&= 0.035-0.0175-0.1575-0.09+0.0075+0.0225 \\
&= -0.2.
\end{aligned}
\tag{48}
$$

Theorem (Symmetrie von Kovarianz und Korrelation)

ξ und υ seien zwei Zufallsvariablen. Dann gelten

$$\mathbb{C}(\xi,\upsilon) = \mathbb{C}(\upsilon,\xi) \text{ und } \rho(\xi,\upsilon) = \rho(\upsilon,\xi) \tag{49}$$

Beweis

Mit der Kommutativität der Multiplikation gelten

$$\mathbb{C}(\xi,\upsilon) = \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\upsilon - \mathbb{E}(\upsilon))\right) = \mathbb{E}\left((\upsilon - \mathbb{E}(\upsilon))(\xi - \mathbb{E}(\xi))\right) = \mathbb{C}(\upsilon,\xi) \tag{50}$$

und

$$\rho(\xi,\upsilon) = \frac{\mathbb{C}(\xi,\upsilon)}{\mathbb{S}(\xi)\mathbb{S}(\upsilon)} = \frac{\mathbb{C}(\upsilon,\xi)}{\mathbb{S}(\upsilon)\mathbb{S}(\xi)} = \rho(\upsilon,\xi) \tag{51}$$

Bemerkungen

- Man sagt auch, dass Kovarianz und Korrelation “nicht gerichtet” sind.

Theorem (Kovarianzverschiebungssatz)

ξ und υ seien Zufallsvariablen. Dann gilt

$$\mathbb{C}(\xi, \upsilon) = \mathbb{E}(\xi\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon). \tag{52}$$

Beweis

Mit der Definition der Kovarianz gilt

$$\begin{aligned}
\mathbb{C}(\xi, \upsilon) &= \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\upsilon - \mathbb{E}(\upsilon))\right) \\
&= \mathbb{E}\left(\xi\upsilon - \xi\mathbb{E}(\upsilon) - \mathbb{E}(\xi)\upsilon + \mathbb{E}(\xi)\mathbb{E}(\upsilon)\right) \\
&= \mathbb{E}(\xi\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon) + \mathbb{E}(\xi)\mathbb{E}(\upsilon) \\
&= \mathbb{E}(\xi\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon).
\end{aligned} \tag{53}$$

$\square$

Bemerkungen

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

Theorem (Varianzen von Summen und Differenzen von Zufallsvariablen)

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

$$\mathbb{V}(a\xi + b\upsilon + c) = a^2\mathbb{V}(\xi) + b^2\mathbb{V}(\upsilon) + 2ab\mathbb{C}(\xi, \upsilon). \tag{54}$$

Speziell gelten

$$\mathbb{V}(\xi + \upsilon) = \mathbb{V}(\xi) + \mathbb{V}(\upsilon) + 2\mathbb{C}(\xi, \upsilon) \tag{55}$$

und

$$\mathbb{V}(\xi - \upsilon) = \mathbb{V}(\xi) + \mathbb{V}(\upsilon) - 2\mathbb{C}(\xi, \upsilon) \tag{56}$$

Bemerkungen

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

Kovarianz und Korrelation

Beweis

Wir halten zunächst fest, dass

$$\mathbb{E}(a\xi + b\upsilon + c) = a\mathbb{E}(\xi) + b\mathbb{E}(\upsilon) + c. \tag{57}$$

Es ergibt sich also

$$\begin{aligned}
&\mathbb{V}(a\xi + b\upsilon + c) \\
&= \mathbb{E}\left((a\xi + b\upsilon + c - a\mathbb{E}(\xi) - b\mathbb{E}(\upsilon) - c)^2\right) \\
&= \mathbb{E}\left((a(\xi - \mathbb{E}(\xi)) + b(\upsilon - \mathbb{E}(\upsilon)))^2\right) \\
&= \mathbb{E}\left(a^2(\xi - \mathbb{E}(\xi))^2 + 2ab(\xi - \mathbb{E}(\xi))(\upsilon - \mathbb{E}(\upsilon))) + b^2(\upsilon - \mathbb{E}(\upsilon))^2\right) \\
&= a^2\mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) + b^2\mathbb{E}\left((\upsilon - \mathbb{E}(\upsilon))^2\right) + 2ab\mathbb{E}\left((\xi - \mathbb{E}(\xi))(\upsilon - \mathbb{E}(\upsilon)))\right) \\
&= a^2\mathbb{V}(\xi) + b^2\mathbb{V}(\upsilon) + 2ab\mathbb{C}(\xi, \upsilon)
\end{aligned} \tag{58}$$

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

$\square$

Theorem (Korrelation und Unabhängigkeit)

ξ und υ seien zwei Zufallsvariablen. Dann gelten:

(1) Wenn ξ und υ unabhängig sind, dann ist $\mathbb{C}(\xi, \upsilon) = 0$ und ξ und υ sind unkorreliert.

(2) Wenn $\mathbb{C}(\xi, \upsilon) = 0$ ist, also ξ und υ unkorreliert sind, dann sind ξ und υ nicht notwendigerweise unabhängig.

Beweis

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

$$\mathbb{E}(\xi\upsilon) = \mathbb{E}(\xi)\mathbb{E}(\upsilon). \tag{59}$$

Mit dem Kovarianzverschiebungssatz folgt dann

$$\mathbb{C}(\xi, \upsilon) = \mathbb{E}(\xi\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon) = \mathbb{E}(\xi)\mathbb{E}(\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon) = 0. \tag{60}$$

Mit der Definition des Korrelationskoeffizienten folgt dann, dass $\rho(\xi, \upsilon) = 0$ ist und ξ und υ unkorreliert sind.

Beweis (fortgeführt)

(2) Wir zeigen nun durch Angabe eines Beispiels, dass die Kovarianz von abhängigen Zufallsvariablen ξ und υ null sein kann. Zu diesem Zweck betrachten wir den Fall zweier diskreter Zufallsvariablen ξ und υ mit Ergebnisräumen $\mathcal{X} = \{-1, 0, 1\}$ und $\mathcal{Y} = \{0, 1\}$, marginaler WMF von ξ gegeben durch $p(x) := \frac{1}{3}$ für alle $x \in \mathcal{X}$ und der Definition von

$$\upsilon := \xi^2. \tag{61}$$

Die Definition von υ impliziert dann die folgende bedingte WMF

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

Die marginale WMF p und die bedingte WMF $p_{\upsilon\mid\xi}$ implizieren dann die gemeinsame WMF

$p(x, y)$	$x = -1$	$x = 0$	$x = 1$	$p_\upsilon(y)$
$y = 0$	0	$1/3$	0	$1/3$
$y = 1$	$1/3$	0	$1/3$	$2/3$
$p(x)$	$1/3$	$1/3$	$1/3$	

Es gilt also zum Beispiel

$$p(-1, 0) = 0 \neq \frac{1}{9} = \frac{1}{3} \cdot \frac{1}{3} = p(-1)p(0). \tag{62}$$

Die gemeinsame WMF von ξ und υ faktorisiert also nicht und ξ und υ sind nicht unabhängig.

Kovarianz und Korrelation

Beweis (fortgeführt)

Allerdings folgt mit

$$\mathbb{E}(\xi) = \sum_{x\in\mathcal{X}} xp(x) = -1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0 \tag{63}$$

und

$$\mathbb{E}(\xi\upsilon) = \mathbb{E}(\xi\xi^2) = \mathbb{E}(\xi^3) = \sum_{x\in\mathcal{X}} x^3 p(x) = -1^3 \cdot \frac{1}{3} + 0^3 \cdot \frac{1}{3} + 1^3 \cdot \frac{1}{3} = 0, \tag{64}$$

sowie dem Kovarianzverschiebungssatz, dass

$$\mathbb{C}(\xi,\upsilon) = \mathbb{E}(\xi\upsilon) - \mathbb{E}(\xi)\mathbb{E}(\upsilon) = \mathbb{E}(\xi^3) - \mathbb{E}(\xi)\mathbb{E}(\upsilon) = 0 - 0 \cdot \mathbb{E}(\upsilon) = 0. \tag{65}$$

Die Kovarianz von ξ und υ ist also null und ξ und υ damit unkorreliert, obwohl ξ und υ nicht unabhängig sind.

$\square$

Erwartungswert

Varianz und Standardabweichung

Stichprobenmittel, Stichprobenvarianz, Stichprobenstandardabweichung

Kovarianz und Korrelation

Selbstkontrollfragen

Selbstkontrollfragen

1. Geben Sie die Definition des Erwartungswerts einer Zufallsvariable wieder.
2. Geben Sie die Interpretation der Erwartungswerts einer Zufallsvariable wieder
3. Berechnen Sie den Erwartungswert einer Bernoulli Zufallsvariable.
4. Geben Sie das Theorem zu den Eigenschaften des Erwartungswerts wieder.
5. Geben Sie die Definition der Varianz und der Standardabweichung einer Zufallsvariable wieder.
6. Geben Sie die Interpretation der Varianz einer Zufallsvariable wieder.
7. Berechnen Sie die Varianz einer Bernoulli Zufallsvariable.
8. Geben Sie das Theorem zum Varianzverschiebungssatz wieder.
9. Geben Sie das Theorem zu den Eigenschaften der Varianz wieder.
10. Geben Sie die Definition des Begriffs einer Stichprobe wieder.
11. Geben Sie die Definitionen von Stichprobenmittel, -varianz und -standardabweichung wieder.
12. Geben Sie die Definition von Kovarianz und Korrelation zweier Zufallsvariablen wieder.
13. Geben Sie das Theorem zum Kovarianzverschiebungssatz wieder.
14. Geben Sie das Theorem zu Varianzen von Summen und Differenzen von Zufallsvariablen wieder.
15. Geben Sie das Theorem zur Korrelation und Unabhängigkeit zweier Zufallsvariablen wieder.

Selbstkontrollfragen

1. Siehe Definition (Erwartungswert).
2. Siehe Bemerkungen zu Definition (Erwartungswert).
3. Siehe Beweis zu Beispiel (Erwarungswert einer Bernoulli Zufallsvariable).
4. Siehe Theorem (Eigenschaften des Erwartungswerts).
5. Siehe Definition (Varianz und Standardabweichung).
6. Siehe Bemerkungen zu Definition (Varianz und Standardabweichung).
7. Siehe Beweis zu Beispiel (Varianz einer Bernoulli Zufallsvariable).
8. Siehe Theorem (Varianzverschiebungssatz).
9. Siehe Theorem (Eigenschaften der Varianz).
10. Siehe Definition (Stichprobenmittel, -varianz, -standardabweichung).
11. Siehe Definition (Stichprobenmittel, -varianz, -standardabweichung).
12. Siehe Definition (Kovarianz und Korrelation).
13. Siehe Theorem (Kovarianzverschiebungssatz).
14. Siehe Theorem (Varianzen von Summen und Differenzen von Zufallsvariablen).
15. Siehe Theorem (Korrelation und Unabhängigkeit).

Anhang

Skizze der Bestimmung des Erwartungswerts einer normalverteilten Zufallsvariable

Wir halten zunächst ohne Beweis fest, dass

$$\int_{-\infty}^{\infty} \exp\left(-x^2\right)\, dx = \sqrt{\pi}. \tag{66}$$

Mit der Definition des Erwartungswerts für kontinuierliche Zufallsvariablen gilt

$$\mathbb{E}(\xi) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)\, dx. \tag{67}$$

Mit der Substitutionsregel

$$\int_{g(a)}^{g(b)} f(x)\, dx = \int_a^b f(g(x))g'(x)\, dx \tag{68}$$

und der Definition von

$$g : \mathbb{R} \to \mathbb{R}, x \mapsto g(x) := \sqrt{2\sigma^2}x + \mu \text{ with } g'(x) = \sqrt{2\sigma^2}, \tag{69}$$

gilt dann

$$
\begin{aligned}
\mathbb{E}(\xi) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x+\mu) \exp\left(-\frac{1}{2\sigma^2}\left(\left(\sqrt{2\sigma^2}x+\mu\right)-\mu\right)^2\right) \sqrt{2\sigma^2}\, dx \\
&= \frac{\sqrt{2\sigma^2}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x+\mu) \exp\left(-x^2\right) dx \\
&= \frac{1}{\sqrt{\pi}} \left(\sqrt{2\sigma^2} \int_{-\infty}^{\infty} x \exp\left(-x^2\right) dx + \mu \int_{-\infty}^{\infty} \exp\left(-x^2\right) dx\right) \\
&= \frac{1}{\sqrt{\pi}} \left(\sqrt{2\sigma^2} \int_{-\infty}^{\infty} x \exp\left(-x^2\right) dx + \mu\sqrt{\pi}\right)
\end{aligned}
\tag{70}
$$

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

$$
\lim_{x\to-\infty} \exp\left(-x^2\right) = 0 \text{ und } \lim_{x\to\infty} \exp\left(-x^2\right) = 0
\tag{71}
$$

verschwindet der Integralterm und wir erhalten

$$
\mathbb{E}(\xi) = \frac{1}{\sqrt{\pi}} \left(\mu\sqrt{\pi}\right) = \mu.
\tag{72}
$$

Skizze der Bestimmung der Varianz einer normalverteilten Zufallsvariable

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

$$\mathbb{V}(\xi) = \mathbb{E}(\xi^2) - \mathbb{E}(\xi)^2 = \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty} x^2 \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx - \mu^2 \tag{73}$$

Mit der Substitutionsregel

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(x)\,dx \tag{74}$$

und der Definition von

$$g : \mathbb{R} \to \mathbb{R}, x \mapsto \sqrt{2\sigma^2}x + \mu, g(-\infty) := -\infty, g(\infty) := \infty, \text{ with } g'(x) = \sqrt{2\sigma^2}, \tag{75}$$

kann das Integral auf der rechten Seite von Gleichung (73) dann als

$$
\begin{aligned}
\int_{-\infty}^{\infty} x^2 \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx &= \int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x+\mu)^2 \exp\left(-\frac{1}{2\sigma^2}((\sqrt{2\sigma^2}x+\mu)-\mu)^2\right)\sqrt{2\sigma^2}\,dx \\
&= \sqrt{2\sigma^2}\int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x+\mu)^2 \exp\left(-\frac{2\sigma^2 x^2}{2\sigma^2}\right) dx \\
&= \sqrt{2\sigma^2}\int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x+\mu)^2 \exp\left(-x^2\right) dx.
\end{aligned} \tag{76}
$$

geschrieben werden. Also gilt

$$
\begin{aligned}
\mathbb{V}(\xi) &= \frac{\sqrt{2\sigma^2}}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x+\mu)^2 \exp\left(-x^2\right) dx - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} (\sqrt{2\sigma^2}x)^2 + 2\sqrt{2\sigma^2}x\mu + \mu^2)\exp\left(-x^2\right) dx - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}\left(2\sigma^2\int_{-\infty}^{\infty} x^2 \exp\left(-x^2\right) dx + 2\sqrt{2\sigma^2}\mu\int_{-\infty}^{\infty} x\exp\left(-x^2\right) dx + \mu^2\int_{-\infty}^{\infty}\exp\left(-x^2\right) dx\right) - \mu^2
\end{aligned} \tag{77}
$$

Wir halten weiterhin ohne Beweis fest, dass

$$\int_{-\infty}^{\infty} x \exp\left(-x^2\right)\, dx = 0 \text{ und } \int_{-\infty}^{\infty} \exp\left(-x^2\right)\, dx = \sqrt{\pi}. \tag{78}$$

Es ergibt sich dann

$$\begin{aligned}
\mathbb{V}(\xi) &= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} x^2 \exp\left(-x^2\right)\, dx + \mu^2\sqrt{\pi}\right) - \mu^2 \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 \exp\left(-x^2\right)\, dx + \mu^2 - \mu^2 \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 \exp\left(-x^2\right)\, dx
\end{aligned} \tag{79}$$

Mit der partiellen Integrationsregel

$$\int_a^b f'(x)g(x)\, dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x)\, dx \tag{80}$$

und der Definition von

$$f : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) := \exp(-x^2) \text{ with } f'(x) = -2\exp(-x^2) \tag{81}$$

und

$$g : \mathbb{R} \to \mathbb{R}, x \mapsto g(x) := -\frac{1}{2}x \text{ with } g'(x) = -\frac{1}{2}, \tag{82}$$

so dass

$$f'(x)g(x) = -2\exp(-x^2)\left(-\frac{1}{2}x\right) = x^2\exp(-x^2), \tag{83}$$

gilt, ergibt sich dann

$$\begin{aligned}
\mathbb{V}(\xi) &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 \exp\left(-x^2\right) dx \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \left(-\frac{1}{2}x\exp(-x^2)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \exp\left(-x^2\right)\left(-\frac{1}{2}\right) dx\right) \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \left(-\frac{1}{2}x\exp(-x^2)|_{-\infty}^{\infty} + \frac{1}{2}\int_{-\infty}^{\infty} \exp\left(-x^2\right) dx\right),
\end{aligned} \tag{84}$$

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

$$\mathbb{V}(\xi) = \frac{2\sigma^2}{\sqrt{\pi}} \left(\frac{1}{2}\int_{-\infty}^{\infty} \exp\left(-x^2\right) dx\right) = \frac{\sigma^2}{\sqrt{\pi}}\sqrt{\pi} = \sigma^2. \tag{85}$$