Vorlesung 5a

Die Varianz

1. Varianz und Standardabweichung:
Elementare Eigenschaften

(Buch S. 24)

X sei reellwertige Zufallsvariable

mit endlichem Erwartungswert μ.

Die **Varianz** von X ist definiert als

$$\mathbf{Var}[X] := \mathbf{E}[(X - \mu)^2],$$

die erwartete quadratische Abweichung

der Zufallsvariablen X von ihrem Erwartungswert μ.

Statt $\mathbf{Var}[X]$ schreiben wir auch

$$\mathbf{Var}\, X$$

oder

$$\sigma_X^2$$

oder (wenn klar ist, welche Zufallsvariable gemeint ist)

$$\sigma^2.$$

Wie ändert sich die Varianz,

wenn man X um eine Konstante verschiebt?

$$\mathbf{Var}[X + d] = \mathbf{E}[((X + d) - (\mu + d))^2] = \mathbf{Var}X$$

Und wenn man X mit einer Konstanten multipliziert

("skaliert")?

$$\mathbf{Var}[cX] = \mathbf{E}[(cX - c\mu)^2] = c^2\mathbf{Var}X$$

Die **Standardabweichung (Streuung)** von X

ist die Wurzel aus der Varianz:

$$\sigma := \sigma_X := \sqrt{\mathbf{Var} X}$$

$$= \sqrt{\mathbf{E}[(X-\mu)^2]}.$$

Sie gibt an, mit welcher „typischen Abweichung“ der Zufallsvariablen X von ihrem Erwartungswert man rechnen sollte.

Es gilt:

$$\sigma_{X+d} = \sigma_X,$$

$$\sigma_{cX} = c\,\sigma_X.$$

Man sagt: σ ist ein *Skalenparameter*.

Für Zufallsvariable mit endlichem Erwartungswert gilt:

$$\mathbf{Var}[X] = 0 \quad \Leftrightarrow \quad \mathbf{P}\big(X = \mathbf{E}[X]\big) = 1 \, .$$

Man sagt dann: X ist *fast sicher* konstant.

Die Äquivalenz sieht man aus der Gleichheit

$$\mathbf{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]$$

zusammen mit dem

Satz über die Positivität des Erwartungswertes.

Wie der Erwartungswert ist auch die Varianz von X durch die Verteilung von X bestimmt:

Hat X die Verteilungsgewichte $\rho(a), \quad a \in S \subseteq \mathbb{R}$ und Erwartungswert μ, so ist

$$\mathbf{Var}\, X = \sum_{a \in S} (a - \mu)^2 \, \rho(a) \; .$$

2. Einfache Beispiele

Beispiel 1:

Eine faire Münze wird dreimal geworfen.

$X = Z_1 + Z_2 + Z_3 \ldots$ Anzahl Köpfe

$$\mathbf{Var}\,[X]$$

$$= \frac{1}{8}(0 - \frac{3}{2})^2 + \frac{3}{8}(1 - \frac{3}{2})^2 + \frac{3}{8}(2 - \frac{3}{2})^2 + \frac{1}{8}(3 - \frac{3}{2})^2$$

$$= \frac{1}{8} \cdot \frac{9+3+3+9}{4} = 3 \cdot \frac{1}{4}$$

Beispiel 2:

Eine p-Münze wird einmal geworfen.

$$\mathbf{P}(Z = 1) = p, \quad \mathbf{P}(Z = 0) = q$$

$$\mathbf{Var}[Z]$$

$$= q(0-p)^2 + p(1-p)^2 = qp^2 + p^2$$

$$= pq(p+q) = pq.$$

Beispiel 3:

Anzahl der Erfolge beim zweimaligen p Münzwurf.

$$\mathbf{Var}[Z_1 + Z_2] = ?$$

Wir rechnen mit Zufallsvariablen:

$$\begin{aligned}
\mathbf{E}[(Z_1 + Z_2 - 2p)^2] &= \mathbf{E}[(Z_1 - p + Z_2 - p)^2] \\
&= \mathbf{E}[(Z_1 - p)^2 + (Z_2 - p)^2 + 2(Z_1 - p)(Z_2 - p)] \\
&= \mathbf{E}[(Z_1 - p)^2] + \mathbf{E}[(Z_2 - p)^2] + 2\mathbf{E}[(Z_1 - p)(Z_2 - p)] \\
&= \mathbf{Var}[Z_1] + \mathbf{Var}[Z_2] + 2\mathbf{E}[Z_1 - p]\,\mathbf{E}[Z_2 - p]
\end{aligned}$$

Beispiel 3:

Anzahl der Erfolge beim zweimaligen p Münzwurf.

$$\mathbf{Var}[Z_1 + Z_2] = ?$$

Wir rechnen mit Zufallsvariablen:

$$\mathbf{E}[(Z_1 + Z_2 - 2p)^2] = \mathbf{E}[(Z_1 - p + Z_2 - p)^2]$$

$$= \mathbf{E}[(Z_1 - p)^2 + (Z_2 - p)^2 + 2(Z_1 - p)(Z_2 - p)]$$

$$= \mathbf{E}[(Z_1 - p)^2] + \mathbf{E}[(Z_2 - p)^2] + 2\mathbf{E}[(Z_1 - p)(Z_2 - p)]$$

$$= \mathbf{Var}[Z_1] + \mathbf{Var}[Z_2] + 2\mathbf{E}[Z_1 - p]\,\mathbf{E}[Z_2 - p]$$

wegen der Unabhängigkeit von Z_1 und Z_2

Beispiel 3:

Anzahl der Erfolge beim zweimaligen p Münzwurf.

$$\mathbf{Var}[Z_1 + Z_2]$$

$$= \mathbf{E}[(Z_1 + Z_2 - 2p)^2] = \mathbf{E}[(Z_1 - p + Z_2 - p)^2]$$
$$= \mathbf{E}[(Z_1 - p)^2 + (Z_2 - p)^2 + 2(Z_1 - p)(Z_2 - p)]$$
$$= \mathbf{E}[(Z_1 - p)^2] + \mathbf{E}[(Z_2 - p)^2] + 2\mathbf{E}[(Z_1 - p)(Z_2 - p)]$$
$$= \mathbf{Var}[Z_1] + \mathbf{Var}[Z_2] \quad + \quad 0$$
$$= 2pq$$

3. Die Varianz der Binomialverteilung

(Buch S. 50 und S. 26)

$(Z_1, \ldots, Z_n)$ sei n-facher p-Münzwurf.

$$\mathbf{Var}[Z_1 + \cdots + Z_n] = ?$$

Gehen wir genauso vor wie im vorigen Beispiel, so finden wir

$$\mathbf{Var}[Z_1 + \cdots + Z_n]$$
$$= \mathbf{Var}[Z_1] + \cdots + \mathbf{Var}[Z_n] + 0.$$

Fazit: Die Varianz der Bin(n, p)-Verteilung ist npq.

4. Zwei Formeln für $\mathbf{Var}[X]$

Zwei (manchmal) hilfreiche Formeln:

$$\boxed{(1) \quad \mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2}$$

$$\boxed{(2) \quad \mathbf{Var}[X] = \mathbf{E}[X(X-1)] - (\mathbf{E}[X])^2 + \mathbf{E}[X]}$$

(2) folgt aus (1) durch Subtrahieren und Addieren von $\mathbf{E}[X]$

$$\boxed{\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2}$$

Denn:

$$\mathbf{E}[(X-\mu)^2] = \mathbf{E}[X^2 - 2\mu X + \mu^2]$$

$$= \mathbf{E}[X^2] - 2\mu\mathbf{E}[X] + \mu^2 = \mathbf{E}[X^2] - \mu^2$$

(wegen Linearität des Erwartungswertes)

5. Die Varianz der Poissonverteilung

(Buch S. 29)

Zur Erinnerung:

Die Poissonverteilung mit Parameter λ

entsteht als Grenzwert von Binomialverteilungen mit

$n \to \infty,\ p \to 0,\ np \to \lambda$.

Weil dann npq gegen λ konvergiert,

steht zu vermuten:

Die Varianz einer Pois(λ)-verteilten Zufallsvariablen X ist λ.

Beweis durch Rechnung:

$$\mathbf{E}[X(X-1)] = \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda}$$

$$= \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda} = \lambda^2.$$

Nach der obigen Formel (2) gilt:

$$\mathbf{Var}[X] = \mathbf{E}[X(X-1)] - (\mathbf{E}[X])^2 + \mathbf{E}[X] =$$

$$\lambda^2 - \lambda^2 + \lambda. \quad \square$$

6. Die Varianz einer Summe von ZV'en und die Kovarianz von zwei ZV'en

(Buch S. 60)

Beim zweifachen p-Münzwurf Z_1, Z_2

ergab sich aus der Unabhängigkeit der Z_i:

$$\mathbf{Var}[Z_1 + Z_2] = \mathbf{Var}[Z_1] + \mathbf{Var}[Z_2].$$

Wie “streuen” Summen von

nicht unabhängigen Zufallsgrößen?

Wie steht’s mit der

Varianz einer Summe von Zufallsvariablen?

$$\mathbf{Var}[X + Y| = \mathbf{E}[((X - \mu_X) + (Y - \mu_Y))^2]$$

$$= \mathbf{E}[(X - \mu_X)^2 + \mathbf{E}[(Y - \mu_Y)^2] + 2\mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$$

Mit der Definition der Kovarianz

$$\mathbf{Cov}[X, Y] := \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$$

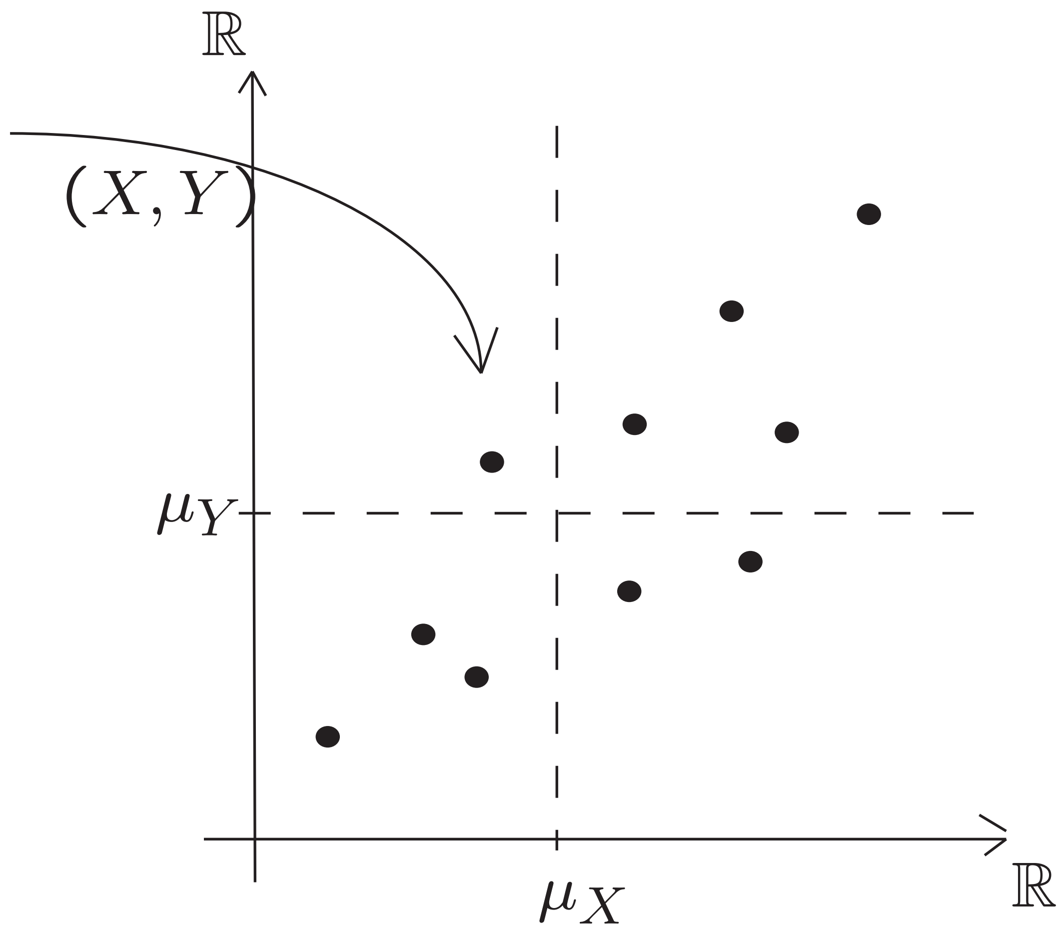
bekommen wir

$$\boxed{\mathbf{Var}[X + Y] = \mathbf{Var}\ X + \mathbf{Var}\ Y + 2\mathbf{Cov}[X, Y].}$$

Die Kovarianz

$$\mathbf{Cov}[X,Y] = \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

ist positiv, wenn X und Y die Tendenz haben,
gemeinsam über bzw. gemeinsam unter
ihrem Erwartungswert auszufallen.
(Größere Abweichungen fallen dabei mehr ins Gewicht.)


ℝ
(X, Y)
μY
μX
ℝ

Ist $\mathbf{Cov}[X, Y]$

$= 0$,	dann nennt man X, Y	unkorreliert
> 0,	$\ldots$	positiv korreliert
< 0,	$\ldots$	negativ korreliert.

Zum Spezialfall von Indikatorvatiablen I_{E_1}, I_{E_2}

siehe auch Abschnitt 6 in V4b.

Zwei weitere Spezialfälle:

$$Y = X:$$

$$\mathbf{Cov}[X,Y] = \mathbf{E}[(X-\mu_X)(X-\mu_X)] = \mathbf{Var}[X]$$

$$Y = -X:$$

$$\mathbf{Cov}[X,Y] = \mathbf{E}[(X-\mu_X)(-X+\mu_X)] = -\mathbf{Var}[X]$$

$$\boxed{\mathbf{Var}[X+Y] = \mathbf{Var}X + \mathbf{Var}Y + 2\mathbf{Cov}[X,Y].}$$

Ganz analog ergibt sich:

$$\mathbf{Var}[Z_1 + \cdots + Z_n]$$

$$= \mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_n + 2 \sum_{i<j} \mathbf{Cov}[Z_i, Z_j]$$

Eine nützliche Umformung von

$\mathbf{Cov}[X, Y] = \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$:

$$
\begin{aligned}
\mathbf{E}[(X - \mu_X)(Y - \mu_Y)] &= \mathbf{E}[XY - \mu_X Y - Y\mu_Y + \mu_X \mu_Y] \\
&= \mathbf{E}[XY] - \mu_X \mu_Y
\end{aligned}
$$

wegen der Linearität des Erwartungswertes.

Also:

$$
\boxed{\mathbf{Cov}[X, Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]}
$$

$$\boxed{\mathbf{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]}$$

Aus der Multiplikationsformel für den Erwartungswert sehen wir (noch einmal):

Unabhängige Zufallsvariable X und Y sind unkorreliert.

Wir halten fest:

Sind $X_1, \dots, X_n$ reellwertige Zufallsvariable

mit endlicher Varianz und

$$\mathbf{Cov}[X_i, X_j] = 0 \quad \text{für } i \neq j$$

(man sagt dafür auch: die X_i sind *paarweise unkorreliert*)

dann gilt:

$$\mathbf{Var}[X_1 + \cdots + X_n] = \mathbf{Var}\, X_1 + \cdots + \mathbf{Var}\, X_n$$

Und allgemein gilt:

$$\mathbf{Var}[Z_1 + \cdots + Z_n] =$$

$$\mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_n + 2 \sum_{i<j} \mathbf{Cov}[Z_i, Z_j]$$

Speziell ist für unabhängige Zufallsvariable

mit endlichen Varianzen

die Varianz der Summe gleich der Summe der Varianzen.

7. Die Varianz der hypergeometrischen Verteilung

(Buch S. 32 und S. 61)

$$\mathbf{Var}[Z_1 + \cdots + Z_n] =$$

$$\mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_n + 2 \sum_{i<j} \mathbf{Cov}[Z_i, Z_j]$$

Beispiel:

Die Anzahl der “Erfolge” beim Ziehen ohne Zurücklegen.

In einer Urne sind r rote und b blaue Kugeln.

Es wird n-mal ohne Zurücklegen gezogen.

X := Anzahl der gezogenen roten Kugeln.

$\mathbf{Var}[X] =$?

Zur Erinnerung:

Mit $g := r + b$ ist

$$\mathbf{P}(X = k) = \frac{\binom{r}{k}\binom{b}{n-k}}{\binom{g}{n}}, \quad k = 0, \ldots, r.$$

X heißt hypergeometrisch verteilt mit Parametern n, g und r.

Erwartungswert und Varianz kann man direkt

über die Verteilungsgewichte ausrechnen (siehe Buch S. 32).

Es geht auch eleganter (vgl Buch S. 50/51):

Wir betrachten dazu die Zufallsvariable Z_i, die

$\dots$den Wert 1 annimmt, falls die i-te gezogene Kugel rot ist,

$\dots$und sonst den Wert 0.

Man sagt dafür auch:

Z_i ist die *Indikatorvariable*

(kurz: der *Indikator*)

des Ereignisses $\{i\text{-te gezogene Kugel rot}\}$.

$$X := Z_1 + \cdots + Z_n$$

$$\mathbf{E}[Z_i] = \mathbf{P}(Z_i = 1) = p, \quad \text{mit}$$

$p := \dfrac{r}{g}$ der Anteil der roten Kugeln in der Urne.

Also: $\mathbf{E}[X] = np$.

Und wie stehts mit der Varianz von X?

$$X := Z_1 + \cdots + Z_n$$

$$\mathbf{Var}\, X = \mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_n + 2 \sum_{1 \le i < j \le n} \mathbf{Cov}[Z_i, Z_j]$$

Sei $g = r + b$ die Gesamtanzahl der Kugeln,

$p := \dfrac{r}{g}$ der Anteil der roten Kugeln in der Urne,

$$q := 1 - p.$$

$$\mathbf{Var}\, Z_i = pq.$$

$$\mathbf{Cov}[Z_i, Z_j] = ?$$

Ein eleganter Weg zur Berechnung von $\mathbf{Cov}[Z_i, Z_j]$:

Wir ziehen in Gedanken, bis die Urne leer ist

(d.h. wir setzen $n = g$.)

Wir ziehen in Gedanken, bis die Urne leer ist.

Dann ist

$$Z_1 + \cdots + Z_g = r,$$

also $\mathbf{Var}[Z_1 + \cdots + Z_g] = 0$.

$$0 = \mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_g + 2 \sum_{1 \le i < j \le g} \mathbf{Cov}[Z_i, Z_j], \quad \text{d.h.}$$

$$0 = gpq + g(g-1)\mathbf{Cov}[Z_1, Z_2], \quad \text{d.h.}$$

$$\mathbf{Cov}[Z_1, Z_2] = -\frac{1}{g-1} pq$$

$$X = Z_1 + \cdots + Z_n$$

$$\mathbf{Var}\, X = \mathbf{Var}\, Z_1 + \cdots + \mathbf{Var}\, Z_n + 2 \sum_{1 \le i < j \le n} \mathbf{Cov}[Z_i, Z_j]$$

$$= n\mathbf{Var}\, Z_1 + n(n-1)\mathbf{Cov}[Z_1, Z_2]$$

$$= npq - n(n-1)\frac{1}{g-1}pq$$

$$= npq\left(1 - \frac{n-1}{g-1}\right) = npq\,\frac{g-n}{g-1}. \quad \square$$

Fazit:

Die Varianz von Hyp(n, g, pg) ist

$$npq\,\frac{g-n}{g-1}.$$

8. Die Ungleichung von Chebyshev

(Buch S. 74)

Es geht um die anschauliche Botschaft

“Je weniger eine reellwertige Zufallsvariable streut,
mit um so größerer Wahrscheinlichkeit
fällt sie nahe zu ihrem Erwartungswert aus.”

Quantifiziert wird das durch die

Ungleichung von Chebyshev:

Y sei eine reellwertige Zufallsvariable

mit endlichem Erwartungswert μ.

Dann gilt für alle $\varepsilon > 0$:

$$\mathbf{P}(|Y-\mu| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\mathbf{Var}[Y]$$

Beweis:

Mit $X := |Y-\mu|$ ist die Behauptung äquivalent zu

$$\mathbf{P}(X^2 \geq \varepsilon^2)) \leq \frac{1}{\varepsilon^2}\mathbf{E}[X^2]$$

Das aber folgt aus der Ungleichung von Markov. $\Box$

9. Das $\sqrt{n}$-Gesetz:

folgt aus der Additivität der Varianz unabhängiger ZV'er:

Seien $X_1, \ldots, X_n$ unabhängig

und identisch verteilt mit Varianz σ^2.

Dann gilt für die Varianz

des Mittelwerts $M_n := \frac{1}{n}(X_1 + \cdots + X_n)$:

$$\mathbf{Var}[M_n] = \frac{1}{n^2} \cdot n\sigma^2 = \frac{1}{n}\sigma^2.$$

Man hat somit das berühmte $\sqrt{n}$-Gesetz:

$$\sigma_{M_n} = \frac{1}{\sqrt{n}}\sigma.$$

10. Das Schwache Gesetz der Großen Zahlen

ist eine unmittelbare Folgerung aus dem $\sqrt{n}$-Gesetz zusammen mit der Ungleichung von Chebyshev:

Seien $X_1, X_2, \ldots$ unabhängig (oder zumindest paarweise unkolleliert) und identisch verteilt mit Erwartungswert μ und endlicher Varianz. Dann gilt für die Mittelwerte $M_n := \frac{1}{n}(X_1 + \cdots + X_n)$:

$$\mathbf{P}(|M_n - \mu| \geq \varepsilon) \leq \frac{1}{\varepsilon^2}\mathbf{Var}[M_n] \to 0 \quad \text{für } n \to \infty.$$

11. Zusammenfassung

$$\mathbf{Var}[X] := \mathbf{E}[(X-\mu)^2]$$

$$\mathbf{Var}[X+Y] = \mathbf{Var}[X] + \mathbf{Var}[Y] + 2\mathbf{Cov}[X,Y]$$

Die Varianz einer Summe von unkorrelierten ZV’en
ist gleich der Summe der Varianzen,
die Varianz einer Summe von negativ korrelierten ZV’en
ist kleiner als die Summe der Varianzen.

Die Varianz von Bin(n, p) ist npq.

Die Varianz von Hyp (n, g, pg) ist $npq\,\dfrac{g-n}{g-1}$.

Die Varianz einer Poisson(λ)-verteilten Zufallsvariablen

ist so groß wie ihr Erwartungswert,

nämlich λ.

Ungleichung von Chebyshev:

$$\mathbf{P}(|Y-\mu| \geq \varepsilon\,\sigma_Y)) \leq \frac{1}{\varepsilon^2}$$

$$\mathbf{Cov}[X, Y] := \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]$$
$$= \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$$

Speziell für Indikatorvariable:

$$\mathbf{Cov}[I_{E_1}, I_{E_2}]$$
$$= \mathbf{P}(E_1 \cap E_2) - \mathbf{P}(E_1)\mathbf{P}(E_2).$$