

# Vorlesung 5a

## Varianz und Kovarianz

# 1. Varianz und Standardabweichung: Elementare Eigenschaften

(Buch S. 24)

$X$  sei reellwertige Zufallsvariable  
mit endlichem Erwartungswert  $\mu$ .

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  $\mu$ .

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

$\text{Var}X$

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:

$$\begin{aligned}\sigma &:= \sigma_X := \sqrt{\text{Var}X} \\ &= \sqrt{\mathbf{E}[(X - \mu)^2]}.\end{aligned}$$

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:  $\sigma$  ist ein *Skalenparameter*.

Für Zufallsvariable mit endlichem Erwartungswert gilt:

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

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

Die Äquivalenz sieht man aus  $\mathbf{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]$   
und aus 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  $\rho(a)$ ,  $a \in S \subseteq \mathbb{R}$   
und Erwartungswert  $\mu$ , so ist

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

## 2. Einfache Beispiele

Beispiel:

Eine faire Münze wird einmal geworfen.

$$\mathbf{P}(Z = 1) = \mathbf{P}(Z = 0) = \frac{1}{2}.$$

$\mathbf{Var} Z$

$$= \frac{1}{2}\left(0 - \frac{1}{2}\right)^2 + \frac{1}{2}\left(1 - \frac{1}{2}\right)^2 = \frac{1}{4}$$

$$\text{Oder: } \left(Z - \frac{1}{2}\right)^2 \equiv \frac{1}{4}, \quad \mathbf{Var} Z = \frac{1}{4}.$$

Beispiel:

Eine faire Münze wird zweimal geworfen.

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

$$= \frac{1}{4}(0 - 1)^2 + \frac{1}{2}(1 - 1)^2 + \frac{1}{4}(2 - 1)^2$$

$$= 2 \cdot \frac{1}{4}$$

Beispiel:

Eine faire Münze wird dreimal geworfen.

$$\text{Var} [Z_1 + Z_2 + Z_3]$$

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

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

Fairer Münzwurf:

$$\text{Var } Z_1 = \frac{1}{4}$$

$$\text{Var } [Z_1 + Z_2] = 2 \cdot \frac{1}{4}$$

$$\text{Var } [Z_1 + Z_2 + Z_3] = 3 \cdot \frac{1}{4}$$

Beispiel:

Eine  $p$ -Münze wird einmal geworfen.

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

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

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

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

# 3. Die Varianz der Binomialverteilung

(Buch S. 26)



Aufgabe:

Eine  $p$ -Münze wird  $n$ -mal geworfen.

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

Anders gefragt: Was ist die Varianz der  $\text{Bin}(n, p)$ -Verteilung?

Erst einmal gibt es eine hilfreiche Formel.

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

Denn:

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

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

(wegen Linearität des Erwartungswertes)

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

Beispiel: Sei  $X$  eine  $\text{Bin}(n, p)$ -verteilte Zufallsvariable

$$\text{Var } X = ?$$

Man rechnet nach (z.B. wie im Buch auf S. 26):

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

Also folgt:

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

$$\text{Var}[X] = npq$$

Wir stellen fest:

Für einen  $p$ -Münzwurf  $Z_1, \dots, Z_n$  ist

$$\text{Var}[Z_1 + \dots + Z_n] = npq,$$

also

$$\text{Var}[Z_1 + \dots + Z_n] = \text{Var}[Z_1] + \dots + \text{Var}[Z_n].$$

Welche allgemeine Struktur steckt hier dahinter?

Wie “streuen” Summen und Mittelwerte von Zufallsgrößen?

Wie steht’s mit der

**Varianz einer Summe von Zufallsvariablen?**

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

(Buch S. 60)

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

Mit der Definition der **Kovarianz**

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

bekommen wir

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

## Die Kovarianz

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

ist positiv, wenn  $X$  und  $Y$  die Tendenz haben,  
gemeinsam über bzw. gemeinsam unter  
ihrem Erwartungswert zu sein.

(Größere Abweichungen fallen dabei mehr ins Gewicht.)

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

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



Zwei Spezialfälle:

$$Y = X :$$

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

$$Y = -X :$$

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

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

Ganz analog ergibt sich:

$$\begin{aligned} & \text{Var}[Z_1 + \cdots + Z_n] \\ &= \text{Var} Z_1 + \cdots + \text{Var} Z_n + 2 \sum_{i < j} \text{Cov}[Z_i, Z_j] \end{aligned}$$

Was ist  $\text{Cov}[Z_i, Z_j]$  beim Münzwurf?

Eine nützliche Umformung von  
 $\text{Cov}[X, Y] = \mathbf{E}[(X - \mu_X)(Y - \mu_Y)]:$

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

wegen der Linearität des Erwartungswertes.

Also:

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

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

Aus der Produktformel für den Erwartungswert (V 4b) folgt:

Unabhängige Zufallsvariable  $X$  und  $Y$  sind unkorreliert.

Also ist für unabhängige Zufallsvariable  
mit endlichen Varianzen

die Varianz der Summe gleich der Summe der Varianzen.

Bei einem  $p$ -Münzwurf  $Z_1, \dots, Z_n$  gilt:

Die  $Z_i$  sind unabhängig und haben jeweils Varianz  $pq$ .

Also ist  $\text{Var}[Z_1 + \dots + Z_n] = npq$ .

Damit haben wir noch einmal bewiesen:

Die Varianz einer  $\text{Bin}(n, p)$ -verteilten Zufallsvariablen ist  $npq$ .

Wir halten fest:

Sind  $X_1, \dots, X_n$  reellwertige Zufallsvariable  
mit endlicher Varianz und

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

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

dann gilt:

$$\text{Var}[X_1 + \dots + X_n] = \text{Var } X_1 + \dots + \text{Var } X_n$$

Und allgemein gilt:

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

Speziell ist für unabhängige Zufallsvariable  
mit endlichen Varianzen

die Varianz der Summe gleich der Summe der Varianzen.

## 5. Das $\sqrt{n}$ -Gesetz:

folgt aus der Additivität der Varianz unabhängiger ZV'er:

Seien  $X_1, \dots, X_n$  unabhängig  
und identisch verteilt mit Varianz  $\sigma^2$ .

Dann gilt für die Varianz  
des Mittelwerts  $M_n := \frac{1}{n}(X_1 + \dots + X_n)$ :

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

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

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



## 6. Die Varianz der hypergeometrischen Verteilung

(Buch S. 32 und S. 61)

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

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.

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

Zur Erinnerung:

Mit  $g := r + b$  ist

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

$X$  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

... den Wert 1 annimmt, falls die  $i$ -te gezogene Kugel rot ist,  
... und sonst den Wert 0.

Man sagt dafür auch:

$Z_i$  ist die *Indikatorvariable*

(kurz: der *Indikator*)

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

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

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

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

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

Und wie stehts mit der Varianz von  $X$ ?

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

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

Sei  $g = r + b$  die Gesamtanzahl der Kugeln,  
 $p := \frac{r}{g}$  der Anteil der roten Kugeln in der Urne,

$$q := 1 - p.$$

$$\text{Var } Z_i = pq.$$

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

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

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

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

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

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

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

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



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

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

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

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

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

Fazit:

Die Varianz von  $\text{Hyp}(n, g, pg)$  ist

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

## 7. Die Varianz der Poissonverteilung

(Buch S. 29)

Zur Erinnerung:

Die Poissonverteilung mit Parameter  $\lambda$  entsteht als Grenzwert von Binomialverteilungen mit  $n \rightarrow \infty$ ,  $p \rightarrow 0$ ,  $np \rightarrow \lambda$ .

Weil dann  $npq$  gegen  $\lambda$  konvergiert,  
steht zu vermuten:

Die Varianz einer  $\text{Pois}(\lambda)$ -verteilten Zufallsvariablen  $X$  ist  $\lambda$ .

Beweis durch Rechnung:

$$\begin{aligned}\mathbf{E}[X(X - 1)] &= \sum_{k=0}^{\infty} k(k - 1) \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k - 2)!} e^{-\lambda} = \lambda^2.\end{aligned}$$

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

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

## 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  $\mu$ .

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.  $\square$



## 9. Das Schwache Gesetz der Großen Zahlen

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

Seien  $X_1, \dots, X_n$  unabhängig  
(oder zumindest paarweise unkorreliert)  
und identisch verteilt mit Erwartungswert  $\mu$  und  
endlicher Varianz. Dann gilt für die Mittelwerte

$$M_n := \frac{1}{n}(X_1 + \dots + X_n):$$
$$\mathbf{P}(|M_n - \mu| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbf{Var}[M_n] \rightarrow 0.$$

# 9. Zusammenfassung

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

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

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  $\text{Bin}(n, p)$  ist  $npq$ .

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

Die Varianz einer  $\text{Poisson}(\lambda)$ -verteilten Zufallsvariablen  
ist so groß wie ihr Erwartungswert,  
nämlich  $\lambda$ .

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

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

Speziell für Indikatorvariable:

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