Vorlesung 4b

Vorlesung 4b

Die Varianz

X sei reellwertige Zufallsvariable

mit endlichem Erwartungswert μ.

X sei reellwertige Zufallsvariable

mit endlichem Erwartungswert μ.

Die **Varianz** von X ist definiert als

X sei reellwertige Zufallsvariable

mit endlichem Erwartungswert μ.

Die **Varianz** von X ist definiert als

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

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?

Wie ändert sich die Varianz,

wenn man X um eine Konstante verschiebt?

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

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”)?

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

Die **Standardabweichung (Streuung)** von X

ist die Wurzel aus der Varianz:

Die **Standardabweichung (Streuung)** von X

ist die Wurzel aus der Varianz:

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

Die **Standardabweichung (Streuung)** von X

ist die Wurzel aus der Varianz:

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

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

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:

Es gilt:

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

Es gilt:

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

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

Es gilt:

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

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

Man sagt: σ ist ein *Skalenparameter*.

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

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

Hat X die Verteilungsgewichte $\rho(a), \quad a \in S \subseteq \mathbb{R}$

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

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

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

Beispiel:

Eine faire Münze wird einmal geworfen.

Beispiel:

Eine faire Münze wird einmal geworfen.

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

$$\mathbf{Var}\ Z$$

Beispiel:

Eine faire Münze wird einmal geworfen.

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

$$\mathbf{Var}\ Z$$

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

Beispiel:

Eine faire Münze wird einmal geworfen.

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

$$\mathbf{Var}\ Z$$

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

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

Beispiel:

Eine faire Münze wird zweimal geworfen.

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

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

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

Beispiel:

Eine faire Münze wird dreimal geworfen.

$$\mathbf{Var}\,[Z_1 + Z_2 + Z_3]$$

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

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

Fairer Münzwurf:

$$\mathbf{Var}\, Z_1 = \frac{1}{4}$$

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

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

Zweifacher fairer Münzwurf:

a) $\mathbf{Var}[Z_1 + Z_2] = 2\mathbf{Var}[Z_1]$

Zweifacher fairer Münzwurf:

a) $\mathbf{Var}[Z_1 + Z_2] = 2\mathbf{Var}[Z_1]$

b) $\mathbf{Var}[Z_1 + Z_1] = \mathbf{Var}[2Z_1] = 4\mathbf{Var}[Z_1]$

Zweifacher fairer Münzwurf:

a) $\mathbf{Var}[Z_1 + Z_2] = 2\mathbf{Var}[Z_1]$

b) $\mathbf{Var}[Z_1 + Z_1] = \mathbf{Var}[2Z_1] = 4\mathbf{Var}[Z_1]$

c) $\mathbf{Var}[Z_1 + (1 - Z_1)] = \mathbf{Var}[1] = 0$

Zweifacher fairer Münzwurf:

a) $\mathbf{Var}[Z_1 + Z_2] = 2\mathbf{Var}[Z_1]$

b) $\mathbf{Var}[Z_1 + Z_1] = \mathbf{Var}[2Z_1] = 4\mathbf{Var}[Z_1]$

c) $\mathbf{Var}[Z_1 + (1 - Z_1)] = \mathbf{Var}[1] = 0$

In allen drei Beispielen haben die beiden Summanden

dieselbe Verteilung

aber jeweils unterschiedliche gemeinsame Verteilung.

Beispiel:

Eine p-Münze wird einmal geworfen.

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

Beispiel:

Eine p-Münze wird einmal geworfen.

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

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

Beispiel:

Eine p-Münze wird einmal geworfen.

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

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

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

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

Beispiel:

Eine p-Münze wird einmal geworfen.

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

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

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

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

Aufgabe:

Eine p-Münze wird n-mal geworfen.

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

Das möcht' ich jetzt nicht mehr zu Fuß ausrechnen müssen.

Aufgabe:

Eine p-Münze wird n-mal geworfen.

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

Das möcht' ich jetzt nicht mehr zu Fuß ausrechnen müssen.

Es gibt eine hilfreiche Formel.

Allgemein gilt

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

Allgemein gilt

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

Denn:

Allgemein gilt

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

Denn:

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

Allgemein gilt

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

Denn:

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

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

Allgemein gilt

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

Denn:

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

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

(wegen Linearität des Erwartungswertes)

Beispiel: Einfacher p-Münwurf Z:

$$\mathbf{E}[Z^2] = \mathbf{E}[Z] = p$$

Beispiel: Einfacher p-Münwurf Z:

$$\mathbf{E}[Z^2] = \mathbf{E}[Z] = p$$

$$\mathbf{E}[Z^2] - \mathbf{E}[Z]^2 = p - p^2 = p(1-p) = pq.$$

Beispiel: Sei X eine Bin(n, p)-verteilte Zufallsvariable

Beispiel: Sei X eine Bin(n, p)-verteilte Zufallsvariable

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

Beispiel: Sei X eine Bin(n, p)-verteilte Zufallsvariable

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

$$\mathbf{E}[X(X-1)] = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k q^{n-k}$$

Beispiel: Sei X eine Bin(n, p)-verteilte Zufallsvariable

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

$$\mathbf{E}[X(X-1)] = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k q^{n-k}$$

$$= n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} q^{(n-2)-(k-2)}$$

Beispiel: Sei X eine Bin(n, p)-verteilte Zufallsvariable

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

$$\mathbf{E}[X(X-1)] = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k q^{n-k}$$

$$= n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} q^{(n-2)-(k-2)}$$

$$= n(n-1)p^2 \,.$$

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

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

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

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

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

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

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

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

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

$$= n(n-1)p^2 + np - (np)^2 = np - np^2 = npq.$$

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

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

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

$$= n(n-1)p^2 + np - (np)^2 = np - np^2 = npq.$$

Fazit:

Die Varianz einer Bin(n, p)-verteilten Zufallsvariablen X ist

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

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

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

$$= n(n-1)p^2 + np - (np)^2 = np - np^2 = npq.$$

Fazit:

Die Varianz einer Bin(n, p)-verteilten Zufallsvariablen X ist

$$\boxed{\mathbf{Var} X = npq.}$$

Die Varianz einer Poissonverteilten Zufallsvariablen:

Die Varianz einer Poissonverteilten Zufallsvariablen:

Zur Erinnerung:

Die Varianz einer Poissonverteilten Zufallsvariablen:

Zur Erinnerung:

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

Die Varianz einer Poissonverteilten Zufallsvariablen:

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,

Die Varianz einer Poissonverteilten Zufallsvariablen:

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 Poissonverteilten Zufallsvariablen:

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:

Beweis durch Rechnung:

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

Beweis durch Rechnung:

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

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

Beweis durch Rechnung:

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

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

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

Beweis durch Rechnung:

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

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

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

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

$Z_1, \ldots, Z_n \quad p$-Münzwurf,

$Z_1, \ldots, Z_n$ p-Münzwurf,

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

$$Z_1, \ldots, Z_n \quad p\text{-Münzwurf},$$

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

$$\mathbf{Var}[Z_i] = pq,$$

$Z_1, \ldots, Z_n \quad p$-Münzwurf,

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

$$\mathbf{Var}[Z_i] = pq,$$

$$\mathbf{Var}[X] = npq.$$

$$Z_1, \ldots, Z_n \quad p\text{-Münzwurf},$$

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

$$\mathbf{Var}[Z_i] = pq,$$

$$\mathbf{Var}[X] = npq.$$

Wir haben das über die Biomialgewichte ausgerechnet.

$$Z_1, \ldots, Z_n \quad p\text{-Münzwurf},$$

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

$$\mathbf{Var}[Z_i] = pq,$$

$$\mathbf{Var}[X] = npq.$$

Wir haben das über die Biomialgewichte ausgerechnet.

Kann man es auch direkt sehen?

$Z_1, \ldots, Z_n \quad p$-Münzwurf,

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

$$\mathbf{Var}[Z_i] = pq,$$

$$\mathbf{Var}[X] = npq.$$

Wir haben das über die Biomialgewichte ausgerechnet.

Kann man es auch direkt sehen?

Wie steht's mit der

Varianz einer Summe von Zufallsvariablen?

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

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

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

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

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

Mit der Definition der Covarianz

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

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

Mit der Definition der Covarianz

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

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

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

Mit der Definition der Covarianz

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

bekommen wir

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

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

Mit der Definition der Covarianz

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

bekommen wir

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

Die Covarianz

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

Die Covarianz

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

ist positiv, wenn X und Y die Tendenz haben,

gemeinsam über bzw. gemeinsam unter

ihrem Erwartungswert zu sein.

Die Covarianz

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

ist positiv, wenn X und Y die Tendenz haben,

gemeinsam über bzw. gemeinsam unter

ihrem Erwartungswert zu sein.

(Größere Abweichungen fallen dabei mehr ins Gewicht.)

Die Covarianz

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

ist positiv, wenn X und Y die Tendenz haben,

gemeinsam über bzw. gemeinsam unter

ihrem Erwartungswert zu sein.

(Größere Abweichungen fallen dabei mehr ins Gewicht.)

Zwei Spezialfälle:

Zwei Spezialfälle:

$$Y = X:$$

Zwei Spezialfälle:

$$Y = X:$$

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

Zwei Spezialfälle:

$$Y = X:$$

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

$$Y = -X:$$

Zwei Spezialfälle:

$$Y = X:$$

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

$$Y = -X:$$

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

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

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

Ganz analog ergibt sich:

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

Ganz analog ergibt sich:

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

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

Ganz analog ergibt sich:

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

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

Eine nützliche Umformung von

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

Eine nützliche Umformung von

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

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

Eine nützliche Umformung von

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

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

Beim p-Münzwurf:

Eine nützliche Umformung von

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

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

Beim p-Münzwurf:

$$\mathbf{E}[Z_i Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) = p^2,$$

Eine nützliche Umformung von

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

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

Beim p-Münzwurf:

$$\mathbf{E}[Z_i Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) = p^2,$$

$$\mathbf{E}[Z_i]\mathbf{E}[Z_j] = \mathbf{P}(Z_i = 1)\mathbf{P}(Z_j = 1) = p^2$$

Eine nützliche Umformung von

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

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

Beim p-Münzwurf:

$$\mathbf{E}[Z_i Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) = p^2,$$

$$\mathbf{E}[Z_i]\mathbf{E}[Z_j] = \mathbf{P}(Z_i = 1)\mathbf{P}(Z_j = 1) = p^2$$

Also ist beim p-Münzwurf $\quad \mathbf{Cov}[Z_i, Z_j] = 0$.

Damit haben wir noch einmal bewiesen:

Damit haben wir noch einmal bewiesen:

Die Varianz einer Bin(n, p)-verteilten Zufallsvariablen ist $n\,pq$.

Wir halten fest:

Wir halten fest:

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

Wir halten fest:

Sind $X_1, \ldots, X_n$ reellwertige Zufallsvariable

mit endlicher Varianz und

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

Wir halten fest:

Sind $X_1, \ldots, X_n$ 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*)

Wir halten fest:

Sind $X_1, \ldots, X_n$ 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:

Wir halten fest:

Sind $X_1, \ldots, X_n$ reellwertige Zufallsvariable

mit endlicher Varianz und

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

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

dann gilt:

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

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

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

Beispiel:

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

Beispiel:

Die Anzahl der “Erfolge” beim Ziehen ohne Zurücklegen –

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

Beispiel:

Die Anzahl der “Erfolge” beim Ziehen ohne Zurücklegen –

Die hypergeometrische Verteilung.

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

Beispiel:

Die Anzahl der “Erfolge” beim Ziehen ohne Zurücklegen –

Die hypergeometrische Verteilung.

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

Zur Erinnerung:

Mit $g := r + b$ ist

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

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 hypergeometrsich verteilt mit Parametern n, g und r.

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 hypergeometrsich verteilt mit Parametern n, g und r.

Erwartungswert und Varianz kann man direkt

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 hypergeometrsich verteilt mit Parametern n, g und r.

Erwartungswert und Varianz kann man direkt

über die Verteilungsgewichte ausrechnen (siehe Buch S. 30).

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 hypergeometrsich verteilt mit Parametern n, g und r.

Erwartungswert und Varianz kann man direkt

über die Verteilungsgewichte ausrechnen (siehe Buch S. 30).

Es geht auch eleganter (vgl Buch S. 48/49):

Wir betrachten dazu die Zufallsvariable Z_i, die

Wir betrachten dazu die Zufallsvariable Z_i, die

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

Wir betrachten dazu die Zufallsvariable Z_i, die

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

$\ldots$ und sonst den Wert 0.

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:

Wir betrachten dazu die Zufallsvariable Z_i, die

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

$\ldots$ und sonst den Wert 0.

Man sagt dafür auch:

Z_i ist die *Indikatorvariable*

(kurz: der *Indikator*)

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Und wie stehts mit der Varianz von X?

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

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

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

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

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

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

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

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

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

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

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

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

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

$q := 1 - p$.

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

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

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

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

$$q := 1 - p.$$

$$\mathbf{Var}\, Z_i = ?$$

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

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

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

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

$$q := 1 - p.$$

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

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

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

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

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

$$q := 1 - p.$$

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

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

$$\mathbf{Cov}[Z_i, Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) - \mathbf{P}(Z_i = 1)\mathbf{P}(Z_j = 1)$$

$$\mathbf{Cov}[Z_i, Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) - \mathbf{P}(Z_i = 1)\mathbf{P}(Z_j = 1)$$

$$= \frac{r(r-1)}{g(g-1)} - \left(\frac{r}{g}\right)^2$$

$$\mathbf{Cov}[Z_i, Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) - \mathbf{P}(Z_i = 1)\mathbf{P}(Z_j = 1)$$

$$= \frac{r(r-1)}{g(g-1)} - \left(\frac{r}{g}\right)^2$$

$$= \frac{r}{g}\left(\frac{r-1}{g-1} - \frac{r}{g}\right) = \frac{r}{g} - \frac{-g+r}{(g-1)g} = -\frac{1}{g-1}\frac{r(g-r)}{g^2}$$

$$\mathbf{Cov}[Z_i, Z_j] = \mathbf{P}(Z_i = 1, Z_j = 1) - \mathbf{P}(Z_i = 1)\mathbf{P}(Z_j = 1)$$

$$= \frac{r(r-1)}{g(g-1)} - \left(\frac{r}{g}\right)^2$$

$$= \frac{r}{g}\left(\frac{r-1}{g-1} - \frac{r}{g}\right) = \frac{r}{g} - \frac{-g+r}{(g-1)g} = -\frac{1}{g-1}\frac{r(g-r)}{g^2}$$

$$= -\frac{1}{g-1}\,pq.$$

$$\mathbf{Cov}[Z_i, Z_j] = -\frac{1}{g-1}\, pq$$

Bekommt man diese Formel noch eleganter?

Wir ziehen in Gedanken, bis die Urne leer ist.

Wir ziehen in Gedanken, bis die Urne leer ist.

Dann ist

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

Wir ziehen in Gedanken, bis die Urne leer ist.

Dann ist

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

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

Wir ziehen in Gedanken, bis die Urne leer ist.

Dann ist

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

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

Wir ziehen in Gedanken, bis die Urne leer ist.

Dann ist

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

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

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

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

Wir ziehen in Gedanken, bis die Urne leer ist.

Dann ist

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Fazit:

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

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

Zusammenfassung:

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

Zusammenfassung:

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

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

Zusammenfassung:

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

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

Die Varianz einer Summe von unkorrelierten ZV'en

ist gleich der Summe der Varianzen,

Zusammenfassung:

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

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

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

Die Varianz von Bin(n, p) ist npq

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

Die Varianz einer Poisson(λ)-verteilten Zufallsvariablen

ist so groß wie ihr Erwartungswert,

nämlich λ.