

Wahrscheinlichkeitsrechnung und Statistik
im Biologie-Bachelorstudiengang der LMU
**Wiederholung zum Rechnen mit Zufallsvariablen,
Erwartungswerten und (Co-)Varianzen**

Dirk Metzler

24. März 2026

Inhaltsverzeichnis

1	Zufallsvariablen	1
1.1	Was sind Zufallsvariablen und Verteilungen?	1
1.2	Rechnen mit Zufallsvariablen	4
1.3	Stochastische (Un-)Abhängigkeit, bedingte Wahrscheinlichkeiten	5
2	Erwartungswerte	7
3	Varianz, Standardabweichung, Kovarianz und Korrelation	8

1 Zufallsvariablen

1.1 Was sind Zufallsvariablen und Verteilungen?

Nehmen wir an, in einer kleinen Population von $n = 100$ Individuen hat ein neutrales Allel A in der aktuellen Generation eine Häufigkeit von 0.32.

Wie wird die Häufigkeit X von A in der nächsten Generation sein?

Das können wir nicht genau vorhersagen, denn es hängt vom Zufall ab.

X ist eine **Zufallsvariable**.

Berechnen können wir aber z.B.:

$\mathbb{E}X$: den Erwartungswert von X ,

$\Pr(X = 0.32)$: die Wahrscheinlichkeit, dass X einen bestimmten Wert annimmt.

Allerdings müssen wir dazu die Population zunächst stochastisch modellieren, und das Ergebnis wird davon abhängen, wie wir das machen.

Beispiele für Zufallsvariablen

- Ergebnis eines Würfelwurfs
- Karte, die aus einem gemischten Kartenspiel gezogen wird
- Koordinatenpaar des Pfeils in einer Dart-Scheibe
- Größe eines zufällig gezogenen Individuums aus einer Population
- Mittelwert der Größe von 10 aus einer Population gezogenen Individuen
- Jede Statistik, d.h. aus Daten berechnete Werte

Als **Zufallsgröße oder Zufallsvariable** bezeichnet man das (Mess-)Ergebnis eines zufälligen Vorgangs.

Der **Wertebereich** \mathcal{S} (engl. state space) einer Zufallsgröße ist die Menge aller möglichen Werte.

Die **Verteilung einer Zufallsgröße** X weist jeder Menge $A \subseteq \mathcal{S}$ die **Wahrscheinlichkeit** $\Pr(X \in A)$ zu, dass X einen Wert in A annimmt.

Für Zufallsgrößen werden üblicherweise Großbuchstaben verwendet (z.B. X, Y, Z), für konkrete Werte Kleinbuchstaben.

Binomialverteilung

Sei X die Anzahl der Erfolge bei n unabhängigen Versuchen mit Erfolgswahrscheinlichkeit von jeweils p . Dann gilt für $k \in \{0, 1, \dots, n\}$

$$\Pr(X = k) = \binom{n}{k} p^k \cdot (1 - p)^{n-k}$$

und X heißt *binomialverteilt*, kurz:

$$X \sim \text{bin}(n, p).$$

Die Wahrscheinlichkeitsverteilung der Binomialverteilung weist jeder Teilmenge $U \subseteq \{0, 1, \dots, n\}$ die folgende Wahrscheinlichkeit zu:

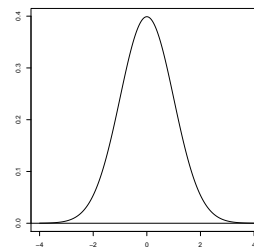
$$\Pr(X \in U) = \sum_{k \in U} \binom{n}{k} p^k \cdot (1 - p)^{n-k}$$

Sei X normalverteilt mit Mittelwert μ und Varianz σ^2 , kurz:

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

X hat dann die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$



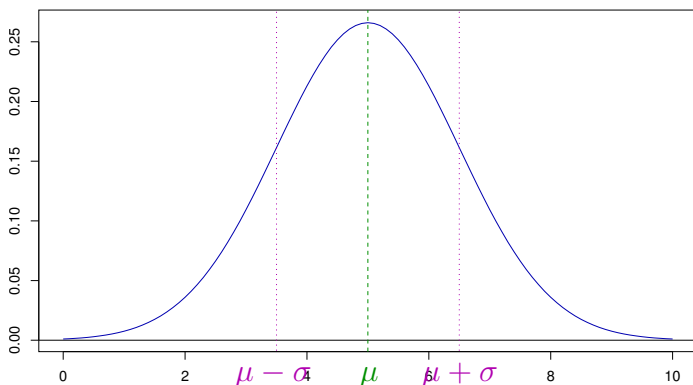
“Gauß-Glocke” mit $\mu = 0$
und $\sigma = 1$

Die Verteilung von X ordnet jeder messbaren¹ Teilmenge U von \mathbb{R} die Wahrscheinlichkeit

$$\Pr(X \in U) = \int_U \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

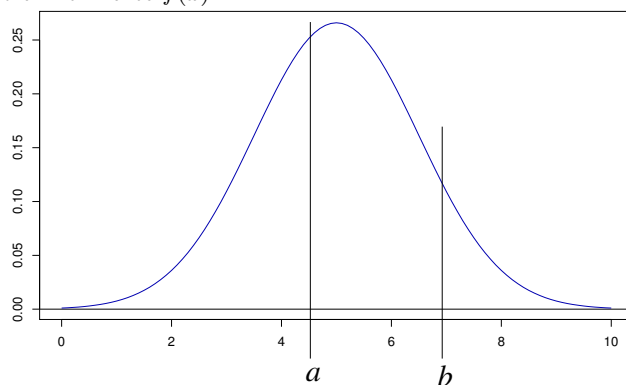
zu.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Dichten brauchen Integrale

Sei Z eine Zufallsvariable mit Dichte $f(x)$.



Dann gilt

$$\Pr(Z \in [a, b]) = \int_a^b f(x) dx.$$

Frage: Wie berechnet man $\Pr(Z = 5)$?

Antwort: Für jedes $x \in \mathbb{R}$ gilt $\Pr(Z = x) = 0$ (da Fläche der Breite 0)

Was wird dann aus $\mathbb{E}Z = \sum_{x \in \mathcal{S}} x \cdot \Pr(Z = x)$?

Bei einer kontinuierlichen Zufallsvariable mit Dichtefunktion f definiert man:

$$\mathbb{E}Z := \int_{-\infty}^{\infty} x \cdot f(x) dx$$

¹erforderliche Einschränkung, damit das Integral überhaupt existiert (mathematisch etwas komplizierter)

Die Normalverteilung in R

Die Normalverteilung hat in R das Kürzel `'norm'`.

Es gibt 4 R-Befehle:

- `dnorm()`: Dichte der Normalverteilung (**d**ensity)
- `rnorm()`: Ziehen einer Stichprobe (**r**andom **s**ample)
- `pnorm()`: Verteilungsfunktion der Normalverteilung (**p**robability)
- `qnorm()`: Quantilfunktion der Normalverteilung (**q**uantile)

1.2 Rechnen mit Zufallsvariablen

Zufallsvariablen, deren Werte Zahlen oder Vektoren sind, heißen auch Zufallszahlen bzw. -vektoren.

Mit ihnen kann man rechnen.

Rechnen mit Zufallsvariablen

Am Beispiel der Summe $W_1 + W_2$ zweier Würfelergebnisse W_1, W_2 haben wir gesehen, dass man mit Zufallsvariablen auch wie mit normalen Variablen rechnen kann. Auch $W_1 + W_2$ ist wieder eine Zufallsvariable und es gilt:

$$\begin{aligned}\Pr(W_1 + W_2 = k) &= \sum_{\{(x,y) : x+y=k\}} \Pr(W_1 = x, W_2 = y) \\ &= \sum_x \Pr(W_1 = x, W_2 = k - x)\end{aligned}$$

In Worten: $\Pr(W_1 + W_2 = k)$ ist die Summe der Wahrscheinlichkeiten aller Wertepaare für W_1 und W_2 , die in Summe k ergeben.

Rechnen mit Zufallsvariablen

Sie können Zufallsvariablen auch mit Zahlen multiplizieren. Stellen Sie sich z.B. vor, dass Sie bei einem Brettspiel die "Speed-Karte" gezogen haben, die Ihnen erlaubt, immer das Doppelte $D = 2 \cdot W$ der gewürfelten Augenzahl W voranzuschreiten.

Dann ist $D = 2 \cdot W$ ebenfalls eine Zufallsvariable. Ihre möglichen Werte sind 2,4,6,8,10 und 12, und D nimmt jeden dieser Werte mit Wahrscheinlichkeit $\frac{1}{6}$ an.

Anwendung einer Funktion auf eine Zufallsvariable

Wenn man eine Funktion f auf eine Zufallsvariable X anwendet, erhält man eine Zufallsvariable $Y = f(X)$.

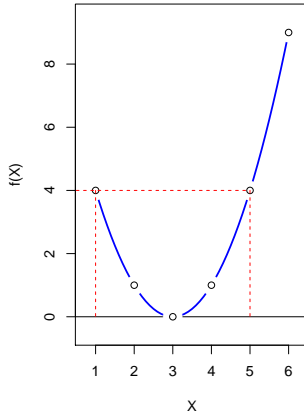
Als Beispiel stellen wir uns ein Spiel vor, bei dem man, wenn man das Würfelergebnis X hat, $f(X) = (X - 3)^2$ Felder vorrücken darf.

Dann ist die Wahrscheinlichkeit, dass man 4 Felder vorrücken darf, also dass die Zufallsvariable $f(X)$ den Wert 4 hat (also X den Wert 1 oder 5):

$$\Pr(f(X) = 4) = \Pr(X \in \{1, 5\}) = 1/3.$$

Um das mathematisch klar zu fassen, benötigen wir den Begriff **Urbild**: Für eine Menge M ist $f^{-1}(M) = \{x : f(x) \in M\}$ das **Urbild** von M unter f , also in Worten: die Menge aller x , die von f auf ein Element von M abgebildet werden.

Das Urbild einer Menge unter einer Funktion f



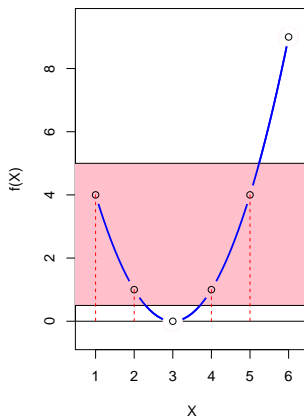
Die Funktion $f : x \mapsto (x - 3)^2$ für $x \in \{1, 2, 3, 4, 5, 6\}$ ist nicht umkehrbar. So ist $f^{-1}(4)$ nicht definierbar, denn es wäre ja auch $f(1) = 4 = f(5)$. Etwas anderes ist jedoch $f^{-1}(\{4\})$. Hier ist mit f^{-1} keine Umkehrfunktion gemeint, sondern die

Urbildfunktion, die auf Mengen operiert:

$$f^{-1}(\{4\}) = \{x : f(x) \in \{4\}\} = \{1, 5\}$$

Oder z.B.:

$$f^{-1}([0.5, 5]) = \{x : f(x) \in [0.5, 5]\} = \{1, 2, 4, 5\}$$



1.3 Stochastische (Un-)Abhängigkeit, bedingte Wahrscheinlichkeiten

Stochastische Unabhängigkeit

Definition 1 (stochastische Unabhängigkeit) Ereignisse U und V sind *(stochastisch) unabhängig*, wenn gilt

$$\Pr(U, V) = \Pr(U) \cdot \Pr(V).$$

Zwei Zufallsgrößen X und Y sind *(stochastisch) unabhängig*, wenn jedes Ereignis der Art $\{X \in A\}$ von jedem Ereignis der Art $\{Y \in B\}$ stochastisch unabhängig ist.

Beispiel:

- Werfen zweier Würfel: $X =$ Augenzahl Würfel 1, $Y =$ Augenzahl Würfel 2.

$$\Pr(X = 2, Y = 5) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \Pr(X = 2) \cdot \Pr(Y = 5)$$

“Stochastisch unabhängig” bedeutet nicht immer, dass zwei Zufallsexperimente ganz unabhängig voneinander durchgeführt wurden. Z.B. beim Ziehen einer Karte aus einem Spiel mit 32 Karten sei

Ereignis A : Die gezogene Karte ist ein Ass

Ereignis \heartsuit : Die gezogene Karte ist von der Farbe Herz

Dann gilt

$$\Pr(\heartsuit, A) = \frac{1}{32} = \frac{1}{4} \cdot \frac{1}{8} = \Pr(\heartsuit) \cdot \Pr(A).$$

Also sind die Ereignisse A und \heartsuit stochastisch unabhängig.

- **Bedingte Wahrscheinlichkeit:** Wahrscheinlichkeit des Ereignisses U unter der Bedingung V

$$\Pr(U|V) := \frac{\Pr(U, V)}{\Pr(V)}$$

„bedingte Wahrscheinlichkeit von U , gegeben V “

Daraus ergibt sich also auch:

$$\Pr(U, V) = \Pr(V) \cdot \Pr(U|V) = \Pr(U) \cdot \Pr(V|U)$$

A, B Ereignisse

Bedingte Wahrscheinlichkeit von A , gegeben B (sofern $\Pr(B) > 0$):

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

($A \cap B := A$ und B treten beide ein) „gegeben B “ bedeutet: wenn man schon weiß, dass B eintritt oder eingetreten ist[0.3cm]Satz von der totalen Wahrscheinlichkeit (mit $B^c := \{B \text{ tritt nicht ein}\}$):

$$\Pr(A) = \Pr(B) \Pr(A|B) + \Pr(B^c) \Pr(A|B^c)$$

Bayes-Formel:

$$\Pr(B|A) = \frac{\Pr(B) \Pr(A|B)}{\Pr(A)}$$

Beispiel medizinischer Test, etwa Mammographie:

A : Das Mammogramm zeigt Krebs an.

B : Die Patientin hat Krebs.

Die nicht bedingte Wahrscheinlichkeit $\Pr(B)$ heißt auch *a-priori*-Wahrscheinlichkeit für B , d.h. die Wahrscheinlichkeit, die man B zuordnet, *bevor* man die Daten A gesehen hat. In unserem Beispiel war 0.008 die Wahrscheinlichkeit, dass eine Vorsorgepatientin Brustkrebs hat.[0.5cm] Die bedingte Wahrscheinlichkeit $\Pr(B|A)$ heißt auch *a-posteriori*-Wahrscheinlichkeit für B . Das ist die Wahrscheinlichkeit, die man B zuweist, *nachdem* man die Daten A gesehen hat.

Die bedingte Wahrscheinlichkeit, dass die Patientin Krebs hat, gegeben, dass das Mammogramm dies anzeigt, ist:

$$\begin{aligned}\Pr(B|A) &= \frac{\Pr(B) \cdot \Pr(A|B)}{\Pr(A)} \\ &= \frac{\Pr(B) \cdot \Pr(A|B)}{\Pr(B) \cdot \Pr(A|B) + \Pr(B^C) \cdot \Pr(A|B^C)}\end{aligned}$$

Mit den Werten aus dem Beispiel aus der Vorlesung:

$$= \frac{0.008 \cdot 0.9}{0.008 \cdot 0.9 + 0.992 \cdot 0.07} \approx 0.0939$$

Bedingt darauf, dass das Mammogramm Krebs anzeigt, beträgt die Wahrscheinlichkeit, dass die Patientin Krebs hat, also lediglich 9.4%.

2 Erwartungswerte

Definition 2 (Erwartungswert) Sei X eine Zufallsvariable mit endlichem oder abzählbarem Wertebereich $\mathcal{S} = \{x_1, x_2, x_3, \dots\} \subseteq \mathbb{R}$. Dann ist der *Erwartungswert* von X definiert durch

$$\mathbb{E}X = \sum_{x \in \mathcal{S}} x \cdot \Pr(X = x)$$

Bei einer kontinuierlichen Zufallsvariable Z mit Dichtefunktion f lautet die Definition:

$$\mathbb{E}Z := \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Beispiele

- Sei X eine 0-1-Zufallsvariable mit Erfolgswahrscheinlichkeit $p \in [0, 1]$. Dann gilt

$$\mathbb{E}X = 1 \cdot \Pr(X = 1) + 0 \cdot \Pr(X = 0) = \Pr(X = 1) = p$$

- Sei W die Augenzahl bei einem Würfelwurf. Dann gilt

$$\begin{aligned}\mathbb{E}W &= 1 \cdot \Pr(W = 1) + 2 \cdot \Pr(W = 2) + \dots + 6 \cdot \Pr(W = 6) \\ &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = 21 \frac{1}{6} = 3.5\end{aligned}$$

- Ist k_x die Häufigkeit des Wertes x in einer Gesamtheit der Größe n , aus der rein zufällig gezogen wird, so folgt:

$$\mathbb{E}X = \sum_x x \cdot \frac{k_x}{n} = \frac{\sum_x x \cdot k_x}{n}$$

Rechnen mit Erwartungswerten

Satz 1 (Linearität des Erwartungswerts) Sind X und Y Zufallsvariablen mit Werten in \mathbb{R} und ist $a \in \mathbb{R}$, so gilt:

- $\mathbb{E}(a \cdot X) = a \cdot \mathbb{E}X$
- $\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y$

Satz 2 (Nur für Unabhängige!) Sind X und Y *stochastisch unabhängige* Zufallsvariablen mit Werten in \mathbb{R} , so gilt

- $\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y$.

Im allgemeinen gilt $\mathbb{E}(X \cdot Y) \neq \mathbb{E}X \cdot \mathbb{E}Y$. Beispiel:

$$\mathbb{E}(W \cdot W) = \frac{91}{6} = 15.167 > 12.25 = 3.5 \cdot 3.5 = \mathbb{E}W \cdot \mathbb{E}W$$

Erwartungswert der Binomialverteilung

Seien Y_1, Y_2, \dots, Y_n die Indikatorvariablen der n unabhängigen Versuche d.h.

$$Y_i = \begin{cases} 1 & \text{falls der } i\text{-te Versuch gelingt} \\ 0 & \text{falls der } i\text{-te Versuch scheitert} \end{cases}$$

Dann ist $X = Y_1 + \dots + Y_n$ binomialverteilt mit Parametern (n, p) , wobei p die Erfolgswahrscheinlichkeit der Versuche ist.

Wegen der Linearität des Erwartungswerts gilt

$$\begin{aligned} \mathbb{E}X &= \mathbb{E}(Y_1 + \dots + Y_n) \\ &= \mathbb{E}Y_1 + \dots + \mathbb{E}Y_n \\ &= p + \dots + p = np \end{aligned}$$

Wir halten fest:

$$X \sim \text{bin}(n, p) \Rightarrow \mathbb{E}X = n \cdot p$$

3 Varianz, Standardabweichung, Kovarianz und Korrelation

Definition 3 (Varianz, Kovarianz und Korrelation) Die *Varianz* einer \mathbb{R} -wertigen Zufallsgröße X ist

$$\text{Var}X = \sigma_X^2 = \mathbb{E}[(X - \mathbb{E}X)^2].$$

$\sigma_X = \sqrt{\text{Var}X}$ ist die *Standardabweichung*.

Ist Y eine weitere reellwertige Zufallsvariable, so ist

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X) \cdot (Y - \mathbb{E}Y)]$$

die *Kovarianz* von X und Y .

Die *Korrelation* von X und Y ist

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}.$$

Die Varianz

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

ist die durchschnittliche quadrierte Abweichung vom Mittelwert.

Die Korrelation

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

liegt immer im Intervall $[-1, 1]$. Die Variablen X und Y sind

- **positiv korreliert**, wenn X und Y tendenziell entweder beide überdurchschnittlich große Werte oder beide unterdurchschnittlich große Werte annehmen.
- **negativ korreliert**, wenn X und Y tendenziell auf verschiedenen Seiten ihrer Erwartungswerte liegen.

Sind X und Y unabhängig, so sind sie auch **unkorreliert**, d.h. $\text{Cor}(X, Y) = 0$.

Bernoulli-Verteilung

Eine Bernoulli-verteilte Zufallsvariable Y mit Erfolgsws $p \in [0, 1]$ hat Erwartungswert

$$\mathbb{E}Y = p$$

und Varianz

$$\text{Var} Y = p \cdot (1 - p)$$

Beweis: Aus $\text{Pr}(Y = 1) = p$ und $\text{Pr}(Y = 0) = (1 - p)$ folgt

$$\mathbb{E}Y = 1 \cdot p + 0 \cdot (1 - p) = p.$$

Varianz:

$$\begin{aligned} \text{Var} Y &= \mathbb{E}(Y^2) - (\mathbb{E}Y)^2 \\ &= 1^2 \cdot p + 0^2 \cdot (1 - p) - p^2 = p \cdot (1 - p) \end{aligned}$$

Beispiel: Würfel

Varianz des Würfelergebnisses W :

$$\begin{aligned} \text{Var}(W) &= \mathbb{E}[(W - \mathbb{E}W)^2] \\ &= \mathbb{E}[(W - 3.5)^2] \\ &= (1 - 3.5)^2 \cdot \frac{1}{6} + (2 - 3.5)^2 \cdot \frac{1}{6} + \dots + (6 - 3.5)^2 \cdot \frac{1}{6} \\ &= \frac{17.5}{6} \\ &= 2.91667 \end{aligned}$$

Beispiel: Die empirische Verteilung

Sind $x_1, \dots, x_n \in \mathbb{R}$ Daten und entsteht X durch rein zufälliges Ziehen aus diesen Daten (d.h. $\text{Pr}(X = x) = \frac{n_x}{n}$, wobei $n_x = |\{i : x_i = x\}|$ die Anzahl der x_i ist mit $x_i = x$), so gilt:

$$\mathbb{E}X = \sum_x x \cdot \frac{n_x}{n} = \sum_{i=1}^n x_i \cdot \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

und

$$\text{Var } X = \mathbb{E}[(X - \mathbb{E}X)^2] = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Sind $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}$ Daten und entsteht (X, Y) durch rein zufälliges Ziehen aus diesen Daten (d.h. $\Pr((X, Y) = (x, y)) = \frac{|\{i : (x_i, y_i) = (x, y)\}|}{n}$), so gilt:

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Definition 4 (Korrigierte Stichprobenvarianz, -standardabweichung und -kovarianz) Ist x_1, x_2, \dots, x_n eine Stichprobe der Zufallsgrößen X , schätzen wir die Varianzen σ_X^2 und die Standardabweichung σ_X von X in der Regel durch

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{und} \quad s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2},$$

und bei gepaarten Stichproben $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ die Kovarianz $\text{Cov}(X, Y)$ durch

$$\text{cov}(x, y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

sowie die Korrelation $\text{Cor}(X, Y)$ durch

$$\text{cor}(x, y) = \frac{\text{cov}(x, y)}{s_x \cdot s_y}.$$

Rechenregeln für Kovarianzen

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X) \cdot (Y - \mathbb{E}Y)]$$

- Sind X und Y unabhängig, so folgt $\text{Cov}(X, Y) = 0$ (die Umkehrung gilt nicht!)
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}X \cdot \mathbb{E}Y$
- $\text{Cov}(a \cdot X, Y) = a \cdot \text{Cov}(X, Y) = \text{Cov}(X, a \cdot Y)$
- $\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$
- $\text{Cov}(X, Z + Y) = \text{Cov}(X, Z) + \text{Cov}(X, Y)$

Die letzten drei Regeln beschreiben die Bilinearität der Kovarianz.

Rechenregeln für die Korrelation

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

- $-1 \leq \text{Cor}(X, Y) \leq 1$
- $\text{Cor}(X, Y) = \text{Cor}(Y, X)$
- $\text{Cor}(X, Y) = \text{Cov}(X/\sigma_X, Y/\sigma_Y)$
- $\text{Cor}(X, Y) = 1$ genau dann wenn Y eine wachsende, affin-lineare Funktion von X ist, d.h. falls es $a > 0$ und $b \in \mathbb{R}$ gibt, so dass $Y = a \cdot X + b$
- $\text{Cor}(X, Y) = -1$ genau dann wenn Y eine fallende, affin-lineare Funktion von X ist, d.h. falls es $a < 0$ und $b \in \mathbb{R}$ gibt, so dass $Y = a \cdot X + b$

Rechenregeln für Varianzen

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

- $\text{Var}X = \text{Cov}(X, X)$
- $\text{Var}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$
- $\text{Var}(a \cdot X) = a^2 \cdot \text{Var}X$
- $\text{Var}(X + Y) = \text{Var}X + \text{Var}Y + 2 \cdot \text{Cov}(X, Y)$
- $\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \cdot \sum_{j=1}^n \sum_{i=1}^{j-1} \text{Cov}(X_i, X_j)$
- Sind (X, Y) stochastisch unabhängig, so folgt:

$$\text{Var}(X + Y) = \text{Var}X + \text{Var}Y$$

(Beweise an der Tafel; ergeben sich aus den Rechenregeln für Kovarianzen)

Standardabweichung einer Summe

Angenommen, es sind bekannt:

σ_X Standardabweichung von X

σ_Y Standardabweichung von Y

Wie berechnet man die Standardabweichung σ_{X+Y} von $X + Y$?

Antwort: Man benötigt zusätzlich $\text{Cov}(X, Y)$ oder $\text{Cor}(X, Y)$. Sei z.B. $\text{Cor}(X, Y)$ gegeben.

Dann verwenden wir die folgenden Gleichungen:

$$\begin{aligned}\text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y) \\ \text{Cov}(X, Y) &= \text{Cor}(X, Y) \cdot \sigma_X \cdot \sigma_Y\end{aligned}$$

Rechne also erst in Varianzen um und am Ende $\sigma_{X+Y} = \sqrt{\text{Var}(X + Y)}$. Der Rechenweg als Herleitung einer Formel zusammengefasst:

$$\begin{aligned}\sigma_{X+Y} &= \sqrt{\text{Var}(X + Y)} \\ &= \sqrt{\text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y)} \\ &= \sqrt{\sigma_X^2 + \sigma_Y^2 + 2 \cdot \text{Cor}(X, Y) \cdot \sigma_X \cdot \sigma_Y}\end{aligned}$$

Binomialverteilung

Seien nun Y_1, \dots, Y_n unabhängig und Bernoulli-verteilt mit Erfolgsws p . Dann gilt

$$\sum_{i=1}^n Y_i =: X \sim \text{bin}(n, p)$$

und es folgt:

$$\text{Var } X = \text{Var} \left(\sum_{i=1}^n Y_i \right) = \sum_{i=1}^n \text{Var } Y_i = n \cdot p \cdot (1 - p)$$

Binomialverteilung

Satz 3 (Erwartungswert und Varianz der Binomialverteilung) Ist X binomialverteilt mit Parametern (n, p) , so gilt:

$$\mathbb{E}X = n \cdot p$$

und

$$\text{Var } X = n \cdot p \cdot (1 - p)$$

Beispiel: Genetische Drift

In einer haploiden Population von n Individuen habe Allel A die Häufigkeit p . Wir gehen wieder davon aus, dass jedes der n Individuen der nächsten Generation unabhängig von den anderen mit W'keit p Allel A bekommt. Die absolute Häufigkeit K von A in der nächsten Generation ist damit (n, p) -binomialverteilt.

Dann folgt für die relative Häufigkeit $X = K/n$ in der nächsten Generation:

$$\begin{aligned} \text{Var}(X) &= \text{Var}(K/n) = \text{Var}(K)/n^2 = n \cdot p \cdot (1 - p)/n^2 \\ &= \frac{p \cdot (1 - p)}{n} \end{aligned}$$

Mit den Rechenregeln konnten wir auch beweisen:

Satz 4 Sind X_1, X_2, \dots, X_n unabhängige \mathbb{R} -wertige Zufallsgrößen mit Mittelwert μ und Varianz σ^2 , so gilt für $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$:

$$\mathbb{E}\bar{X} = \mu$$

und

$$\text{Var } \bar{X} = \frac{1}{n} \sigma^2,$$

d.h.

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Insbesondere: Der Standardfehler $\frac{s}{\sqrt{n}}$ ist ein Schätzer der Standardabweichung $\sigma_{\bar{X}}$ des Stichprobenmittels \bar{X} der Stichprobe (X_1, X_2, \dots, X_n) .

Die Stichproben-Standardabweichung s ist ein Schätzer der Standardabweichung σ der Grundgesamtheit.

Beweis: Linearität des Erwartungswertes impliziert

$$\begin{aligned}\mathbb{E}\bar{X} &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \mu = \mu.\end{aligned}$$

Die Unabhängigkeit der X_i vereinfacht die Varianz zu

$$\begin{aligned}\text{Var } \bar{X} &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{1}{n} \sigma^2\end{aligned}$$

Beispiel: Erwartungstreue der korrigierten Stichprobenvarianz

Mit etwas mehr Mühe folgte auch:

Für X_1, X_2, \dots, X_n unabhängig identisch verteilt mit Varianz σ^2 ist die korrigierte Stichprobenvarianz

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

ein erwartungstreuer Schätzer für σ^2 , das heißt:

$$\mathbb{E}(s^2) = \sigma^2$$

(Aber die Erwartungstreue gilt nicht für s , sondern i.A. gilt $\mathbb{E}s < \sigma$).

Beispiel: Standardfehler der Steigung einer Regressionsgeraden

$$y_i = a + b \cdot x_i + \varepsilon_i \quad \text{mit } \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

nicht zufällig: a, b, x_i, σ^2 zufällig: ε_i, y_i

$$\text{Var}(y_i) = \text{Var}(a + b \cdot x_i + \varepsilon_i) = \text{Var}(\varepsilon_i) = \sigma^2$$

und y_1, y_2, \dots, y_n sind stochastisch unabhängig.

$$\hat{b} = \frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

$$\begin{aligned}\text{Var}(\hat{b}) &= \text{Var}\left(\frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}\right) = \frac{\text{Var}(\sum_i y_i (x_i - \bar{x}))}{(\sum_i (x_i - \bar{x})^2)^2} = \frac{\sum_i \text{Var}(y_i) (x_i - \bar{x})^2}{(\sum_i (x_i - \bar{x})^2)^2} \\ &= \sigma^2 \cdot \frac{\sum_i (x_i - \bar{x})^2}{(\sum_i (x_i - \bar{x})^2)^2} = \sigma^2 \left/ \sum_i (x_i - \bar{x})^2\right.\end{aligned}$$

Standardfehler der Steigung einer Regressionsgeraden

Tatsächlich ist \hat{b} Normalverteilt mit Mittelwert b und

$$\text{Var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Problem: Wir kennen σ^2 nicht. Wir schätzen σ^2 mit Hilfe der beobachteten Residuenvarianz durch

$$s^2 := \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}$$

Zu beachten ist, dass durch $n - 2$ geteilt wird. Das hat damit zu tun, dass zwei Modellparameter a und b bereits geschätzt wurden, und somit 2 Freiheitsgrade verloren gegangen sind.

$$\text{Var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Schätze σ^2 durch

$$s^2 = \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}$$

Dann ist

$$\frac{\hat{b} - b}{s / \sqrt{\sum_i (x_i - \bar{x})^2}}$$

Student- t -verteilt mit $n - 2$ Freiheitsgraden und wir können den t -Test anwenden, um die Nullhypothese $b = 0$ zu testen.