

# Wahrscheinlichkeitsrechnung und Statistik für Biologen

## **3. Grundlagen aus der Wahrscheinlichkeitstheorie**

Martin Hutzenthaler & Dirk Metzler

[http://evol.bio.lmu.de/\\_statgen](http://evol.bio.lmu.de/_statgen)

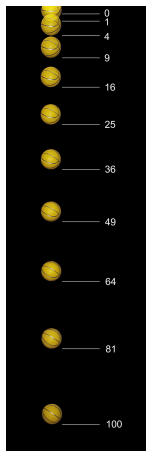
2. Mai 2012

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test

# Inhalt

- 1 **Deterministische und zufällige Vorgänge**
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test

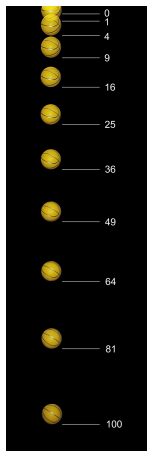
## Was können wir vorhersagen:



(c) by Michael Maggs

- Freier Fall: Falldauer eines Objektes bei gegebener Fallhöhe läßt sich vorhersagen (falls Luftwiderstand vernachlässigbar)

## Was können wir vorhersagen:



(c) by Michael Maggs

- Freier Fall: Falldauer eines Objektes bei gegebener Fallhöhe läßt sich vorhersagen (falls Luftwiderstand vernachlässigbar)

**Deterministische** Vorgänge laufen immer gleich ab. Aus Beobachtungen lassen sich künftige Versuche vorhersagen.

## Was können wir vorhersagen:

- **Würfelnwurf:** Das Ergebnis eines einzelnen Würfelnwurfes lässt sich nicht vorhersagen.



(c) public domain

## Was können wir vorhersagen:

- **Würfelnwurf:** Das Ergebnis eines einzelnen Würfelnwurfes lässt sich nicht vorhersagen.



(c) public domain

- **Wiederholter Würfelnwurf:**

Würfelt man 600 mal, so würde man gerne darauf wetten, dass die Anzahl an Einsern zwischen 75 und 125 liegt.

## Was können wir vorhersagen:

- **Würfelnwurf:** Das Ergebnis eines einzelnen Würfelwurfes lässt sich nicht vorhersagen.



(c) public domain

- **Wiederholter Würfelwurf:**

Würfelt man 600 mal, so würde man gerne darauf wetten, dass die Anzahl an Einsern zwischen 75 und 125 liegt.

Die genaue Anzahl lässt sich wieder nicht vorhersagen.



## Was können wir vorhersagen:

- **Würfeln:** Das Ergebnis eines einzelnen Würfelwurfes lässt sich nicht vorhersagen.



(c) public domain

- **Wiederholter Würfelwurf:**

Würfelt man 600 mal, so würde man gerne darauf wetten, dass die Anzahl an Einsern zwischen 75 und 125 liegt.

Die genaue Anzahl lässt sich wieder nicht vorhersagen.

Aber: **Eine Aussage über die Verteilung ist möglich**  
(die besser ist als reines Raten.)

Empirisch stellt man fest:

Bei Wiederholung eines Zufallsexperiments stabilisieren sich die relativen Häufigkeiten der möglichen Ergebnisse.

Empirisch stellt man fest:

Bei Wiederholung eines Zufallsexperiments stabilisieren sich die relativen Häufigkeiten der möglichen Ergebnisse.

Beispiel:

Beim Würfelwurf  
stabilisiert sich die relative Häufigkeit  
jeder der Zahlen  $\{1, 2, \dots, 6\}$  bei  $\frac{1}{6}$ .

Empirisch stellt man fest:

Bei Wiederholung eines Zufallsexperiments stabilisieren sich die relativen Häufigkeiten der möglichen Ergebnisse.

Beispiel:

Beim Würfelwurf stabilisiert sich die relative Häufigkeit jeder der Zahlen  $\{1, 2, \dots, 6\}$  bei  $\frac{1}{6}$ .

Fazit:

Das Ergebnis eines einzelnen zufälligen Vorgangs läßt sich nicht vorhersagen.  
Aber: Eine Aussage über die Verteilung ist möglich (die besser ist als reines Raten).

Abstraktionsschritt:

Verwende empirisch ermittelte Verteilung  
als Verteilung jedes Einzelexperimentes!

Abstraktionsschritt:

Verwende empirisch ermittelte Verteilung  
als Verteilung jedes Einzelexperiments!

Beispiel:

Wir nehmen an,  
daß bei einem einzelnen Würfelwurf  
jede der Zahlen  $\{1, 2, \dots, 6\}$   
die **Wahrscheinlichkeit**  $\frac{1}{6}$  hat.

# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung**
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test

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



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.

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.

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.

# Mengenschreibweise

Das Ereignis, dass  $X$  einen Wert in  $A$  annimmt, kann man mit geschweiften Klammern schreiben:

$$\{X \in A\}$$

Dies kann man interpretieren als die Menge aller Elementarereignisse, für die  $X$  einen Wert in  $A$  annimmt.

# Mengenschreibweise

Das Ereignis, dass  $X$  einen Wert in  $A$  annimmt, kann man mit geschweiften Klammern schreiben:

$$\{X \in A\}$$

Dies kann man interpretieren als die Menge aller Elementarereignisse, für die  $X$  einen Wert in  $A$  annimmt. Die Schnittmenge

$$\{X \in A\} \cap \{X \in B\} = \{X \in A, X \in B\}$$

ist dann das Ereignis, dass der Wert von  $X$  in  $A$  liegt **und** in  $B$  liegt.

# Mengenschreibweise

Das Ereignis, dass  $X$  einen Wert in  $A$  annimmt, kann man mit geschweiften Klammern schreiben:

$$\{X \in A\}$$

Dies kann man interpretieren als die Menge aller Elementarereignisse, für die  $X$  einen Wert in  $A$  annimmt. Die Schnittmenge

$$\{X \in A\} \cap \{X \in B\} = \{X \in A, X \in B\}$$

ist dann das Ereignis, dass der Wert von  $X$  in  $A$  liegt **und** in  $B$  liegt. Die Vereinigungsmenge

$$\{X \in A\} \cup \{X \in B\}$$

ist das Ereignis, dass der Wert von  $X$  in  $A$  liegt **oder** (auch) in  $B$  liegt.

# Mengenschreibweise

Das Ereignis, dass  $X$  einen Wert in  $A$  annimmt, kann man mit geschweiften Klammern schreiben:

$$\{X \in A\}$$

Dies kann man interpretieren als die Menge aller Elementarereignisse, für die  $X$  einen Wert in  $A$  annimmt. Die Schnittmenge

$$\{X \in A\} \cap \{X \in B\} = \{X \in A, X \in B\}$$

ist dann das Ereignis, dass der Wert von  $X$  in  $A$  liegt **und** in  $B$  liegt. Die Vereinigungsmenge

$$\{X \in A\} \cup \{X \in B\}$$

ist das Ereignis, dass der Wert von  $X$  in  $A$  liegt **oder** (auch) in  $B$  liegt. Bei Wahrscheinlichkeiten läßt man die Klammern oft weg:

$$\Pr(X \in A, X \in B) = \Pr(\{X \in A, X \in B\})$$

**Beispiel:** Würfelwurf  $W =$  Augenzahl des nächsten Würfelwurfs.

$$\mathcal{S} = \{1, 2, \dots, 6\}$$

$$\Pr(W = 1) = \dots = \Pr(W = 6) = \frac{1}{6}$$

$$(\Pr(W = x) = \frac{1}{6} \text{ für alle } x \in \{1, \dots, 6\})$$

Die Verteilung erhält man aus einer Symmetrieüberlegung  
oder aus einer langen Würfelreihe.



**Beispiel:** Würfelwurf  $W =$  Augenzahl des nächsten Wurfes.

$$\mathcal{S} = \{1, 2, \dots, 6\}$$

$$\Pr(W = 1) = \dots = \Pr(W = 6) = \frac{1}{6}$$

$$(\Pr(W = x) = \frac{1}{6} \text{ für alle } x \in \{1, \dots, 6\})$$

Die Verteilung erhält man aus einer Symmetrieüberlegung  
oder aus einer langen Würfelreihe.

**Beispiel:** Geschlecht  $X$  bei Neugeborenen.

$$\mathcal{S} = \{\text{„männlich“}, \text{„weiblich“}\}$$

Die Verteilung erhält man aus einer langen Beobachtungsreihe.

**Beispiel:** Würfelwurf  $W =$  Augenzahl des nächsten Wurfelwurfs.

$$\mathcal{S} = \{1, 2, \dots, 6\}$$

$$\Pr(W = 1) = \dots = \Pr(W = 6) = \frac{1}{6}$$

$$(\Pr(W = x) = \frac{1}{6} \text{ für alle } x \in \{1, \dots, 6\})$$

Die Verteilung erhält man aus einer Symmetrieüberlegung  
oder aus einer langen Würfelreihe.

**Beispiel:** Geschlecht  $X$  bei Neugeborenen.

$$\mathcal{S} = \{\text{„männlich“}, \text{„weiblich“}\}$$

Die Verteilung erhält man aus einer langen Beobachtungsreihe.

**Beispiel:** Körpergrößenverteilung in Deutschland.

Die Verteilung erhält man aus einer langen Messreihe.

# Rechenregeln:

**Beispiel** Würfelnwurf  $W$ :

$$\begin{aligned}\Pr(W \in \{2, 3\}) &= \frac{2}{6} = \frac{1}{6} + \frac{1}{6} \\ &= \Pr(W = 2) + \Pr(W = 3)\end{aligned}$$

$$\begin{aligned}\Pr(W \in \{1, 2\} \cup \{3, 4\}) &= \frac{4}{6} = \frac{2}{6} + \frac{2}{6} \\ &= \Pr(W \in \{1, 2\}) + \Pr(W \in \{3, 4\})\end{aligned}$$

Vorsicht:

$$\begin{aligned}\Pr(W \in \{2, 3\}) + \Pr(W \in \{3, 4\}) &= \frac{2}{6} + \frac{2}{6} = \frac{4}{6} \\ &\neq \Pr(W \in \{2, 3, 4\}) = \frac{3}{6}\end{aligned}$$

**Beispiel zweifacher Würfelwurf ( $W_1, W_2$ ):**

Sei  $W_1$  (bzw  $W_2$ ) die Augenzahl des ersten (bzw zweiten) Würfels.

$$\begin{aligned}\Pr(W_1 \in \{4\}, W_2 \in \{2, 3, 4\}) \\ &= \Pr((W_1, W_2) \in \{(4, 2), (4, 3), (4, 4)\}) \\ &= \frac{3}{36} = \frac{1}{6} \cdot \frac{3}{6} \\ &= \Pr(W_1 \in \{4\}) \cdot \Pr(W_2 \in \{2, 3, 4\})\end{aligned}$$

Allgemein:

$$\Pr(W_1 \in A, W_2 \in B) = \Pr(W_1 \in A) \cdot \Pr(W_2 \in B)$$

für alle Mengen  $A, B \subseteq \{1, 2, \dots, 6\}$

Sei  $S$  die Summe der Augenzahlen, d.h.  $S = W_1 + W_2$ .  
Was ist die Wahrscheinlichkeit, daß  $S = 5$  ist,  
wenn der erste Würfel die Augenzahl  $W_1 = 2$  zeigt?

$$\begin{aligned}\Pr(S = 5 | W_1 = 2) &\stackrel{!}{=} \Pr(W_2 = 3) \\ &= \frac{1}{6} = \frac{1/36}{1/6} = \frac{\Pr(S=5, W_1=2)}{\Pr(W_1=2)}\end{aligned}$$

Sei  $S$  die Summe der Augenzahlen, d.h.  $S = W_1 + W_2$ .  
 Was ist die Wahrscheinlichkeit, daß  $S = 5$  ist,  
 wenn der erste Würfel die Augenzahl  $W_1 = 2$  zeigt?

$$\begin{aligned} \Pr(S = 5 | W_1 = 2) &\stackrel{!}{=} \Pr(W_2 = 3) \\ &= \frac{1}{6} = \frac{1/36}{1/6} = \frac{\Pr(S=5, W_1=2)}{\Pr(W_1=2)} \end{aligned}$$

Was ist die Ws von  $S \in \{4, 5\}$  unter der Bedingung  $W_1 = 1$ ?

$$\begin{aligned} &\Pr(S \in \{4, 5\} | W_1 = 1) \\ &\stackrel{!}{=} \Pr(W_2 \in \{3, 4\}) \\ &= \frac{2}{6} = \frac{2/36}{1/6} = \frac{\Pr(W_2 \in \{3, 4\}, W_1 = 1)}{\Pr(W_1 = 1)} \\ &= \frac{\Pr(S \in \{4, 5\}, W_1 = 1)}{\Pr(W_1 = 1)} \end{aligned}$$

# Rechenregeln:

Seien  $X, Y$  Zufallsgrößen mit Wertebereich  $\mathcal{S}$ .

- $0 \leq \Pr(X \in A) \leq 1$  für jede Teilmenge  $A \subseteq \mathcal{S}$
- $\Pr(X \in \mathcal{S}) = 1$
- Sind  $A, B \subseteq \mathcal{S}$  disjunkt, d.h.  $A \cap B = \emptyset$ ,

$$\Pr(X \in A \cup B) = \Pr(X \in A) + \Pr(X \in B)$$

- Allgemein gilt die **Einschluss-Ausschluss-Formel**

$$\Pr(X \in A \cup B) = \Pr(X \in A) + \Pr(X \in B) - \Pr(X \in A \cap B)$$

- **Bayes-Formel für die bedingte Wahrscheinlichkeit:**

Ws des Ereignisses  $\{Y \in B\}$  unter der Bedingung  $\{X \in A\}$

$$\Pr(Y \in B | X \in A) := \frac{\Pr(Y \in B, X \in A)}{\Pr(X \in A)}$$

„bedingte Ws von  $\{Y \in B\}$  gegeben  $\{X \in A\}$ “

- Bayes-Formel für die bedingte Wahrscheinlichkeit:

Ws des Ereignisses  $\{Y \in B\}$  unter der Bedingung  $\{X \in A\}$

$$\Pr(Y \in B | X \in A) := \frac{\Pr(Y \in B, X \in A)}{\Pr(X \in A)}$$

„bedingte Ws von  $\{Y \in B\}$  gegeben  $\{X \in A\}$ “



- Bayes-Formel für die bedingte Wahrscheinlichkeit:

Ws des Ereignisses  $\{Y \in B\}$  unter der Bedingung  $\{X \in A\}$

$$\Pr(Y \in B | X \in A) := \frac{\Pr(Y \in B, X \in A)}{\Pr(X \in A)}$$

„bedingte Ws von  $\{Y \in B\}$  gegeben  $\{X \in A\}$ “

Beachte:

$$\Pr(X \in A, Y \in B) = \Pr(X \in A) \cdot \Pr(Y \in B | X \in A)$$

Wir wollen

$$\Pr(X \in A, Y \in B) = \Pr(X \in A) \cdot \Pr(Y \in B \mid X \in A)$$

in Worten ausdrücken:

Wir wollen

$$\Pr(X \in A, Y \in B) = \Pr(X \in A) \cdot \Pr(Y \in B \mid X \in A)$$

in Worten ausdrücken:

Die Ws des Ereignisses  $\{X \in A, Y \in B\}$  läßt sich in zwei Schritten berechnen:

- Zunächst muss das Ereignis  $\{X \in A\}$  eintreten.
- Die Ws hiervon wird multipliziert mit der Ws von  $\{Y \in B\}$ , wenn man schon weiß, daß  $\{X \in A\}$  eintritt.

# Stochastische Unabhängigkeit

## Definition (stochastische Unabhängigkeit)

Zwei Zufallsgrößen  $X$  und  $Y$  heißen (stochastisch) unabhängig, wenn für alle Ereignisse  $\{X \in A\}$ ,  $\{Y \in B\}$  gilt

$$\Pr(X \in A, Y \in B) = \Pr(X \in A) \cdot \Pr(Y \in B)$$

# Stochastische Unabhängigkeit

## Definition (stochastische Unabhängigkeit)

Zwei Zufallsgrößen  $X$  und  $Y$  heißen (stochastisch) unabhängig, wenn für alle Ereignisse  $\{X \in A\}$ ,  $\{Y \in B\}$  gilt

$$\Pr(X \in A, Y \in B) = \Pr(X \in A) \cdot \Pr(Y \in B)$$

Beispiel:

- Werfen zweier Würfel:

$X$  = Augenzahl Würfel 1,  $Y$  = Augenzahl Würfel 2.

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

# Stochastische Unabhängigkeit

In der Praxis wendet man häufig Resultate an, die Unabhängigkeit einer Stichprobe voraussetzen.

# Stochastische Unabhängigkeit

In der Praxis wendet man häufig Resultate an, die Unabhängigkeit einer Stichprobe voraussetzen.

Beispiele:

- Für eine Studie wird eine zufällige Person in München und eine zufällige Person in Hamburg befragt. Die Antworten dürfen als unabhängig voneinander angenommen werden.

# Stochastische Unabhängigkeit

In der Praxis wendet man häufig Resultate an, die Unabhängigkeit einer Stichprobe voraussetzen.

Beispiele:

- Für eine Studie wird eine zufällige Person in München und eine zufällige Person in Hamburg befragt. Die Antworten dürfen als unabhängig voneinander angenommen werden.
- Befragt man zwei Schwestern oder nahe verwandte Personen (getrennt voneinander), so werden die Antworten nicht unabhängig voneinander sein.



# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung**
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test

# Ja/Nein-Zufallsgrößen

Das einfachsten Zufallsexperimente sind welche mit mit exakt zwei möglichen Werten.

Wir werden solche Zufallsgrößen als **Ja/Nein-Zufallsgrößen** bezeichnen.

# Ja/Nein-Zufallsgrößen

Das einfachsten Zufallsexperimente sind welche mit mit exakt zwei möglichen Werten.

Wir werden solche Zufallsgrößen als **Ja/Nein-Zufallsgrößen** bezeichnen.

Die zwei möglichen Ergebnisse werden üblicherweise mit 1 und 0 bezeichnet

# Ja/Nein-Zufallsgrößen

Das einfachsten Zufallsexperimente sind welche mit mit exakt zwei möglichen Werten.

Wir werden solche Zufallsgrößen als **Ja/Nein-Zufallsgrößen** bezeichnen.

Die zwei möglichen Ergebnisse werden üblicherweise mit 1 und 0 bezeichnet ,  
beziehungsweise als 'Erfolg' und 'Misserfolg'.

# Ja/Nein-Zufallsgrößen

Das einfachsten Zufallsexperimente sind welche mit mit exakt zwei möglichen Werten.

Wir werden solche Zufallsgrößen als **Ja/Nein-Zufallsgrößen** bezeichnen.

Die zwei möglichen Ergebnisse werden üblicherweise mit 1 und 0 bezeichnet ,  
beziehungsweise als 'Erfolg' und 'Misserfolg'.

**Ja/Nein-Zufallsgröße**  $X$ :

Zustandsraum  $\mathcal{S} = \{0, 1\}$ .

Verteilung:

$$\Pr(X = 1) = p$$

$$\Pr(X = 0) = 1 - p$$

Der Parameter  $p \in [0, 1]$  heißt **Erfolgswahrscheinlichkeit**.

# Ja/Nein Zufallsgrößen

Beispiele:

- Münzwurf: mögliche Werte sind „Kopf“ und „Zahl“.

# Ja/Nein Zufallsgrößen

Beispiele:

- Münzwurf: mögliche Werte sind „Kopf“ und „Zahl“.
- Hat die gesampelte Drosophila eine Mutation, die weiße Augen verursacht?  
Mögliche Antworten sind „Ja“ und „Nein“.

# Ja/Nein Zufallsgrößen

Beispiele:

- Münzwurf: mögliche Werte sind „Kopf“ und „Zahl“.
- Hat die gesampelte *Drosophila* eine Mutation, die weiße Augen verursacht?  
Mögliche Antworten sind „Ja“ und „Nein“.
- Das Geschlecht einer Person hat die möglichen Werte „männlich“ und „weiblich“.



Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

- 1 ...immer gelingt?

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

1 ...immer gelingt?

$$p \cdot p \cdot p \cdots p = p^n$$

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

1 ...immer gelingt?

$$p \cdot p \cdot p \cdots p = p^n$$

2 ...immer scheitert?

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

1 ...immer gelingt?

$$p \cdot p \cdot p \cdots p = p^n$$

2 ...immer scheitert?

$$(1 - p) \cdot (1 - p) \cdots (1 - p) = (1 - p)^n$$

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

- 1 ...immer gelingt?

$$p \cdot p \cdot p \cdots p = p^n$$

- 2 ...immer scheitert?

$$(1 - p) \cdot (1 - p) \cdots (1 - p) = (1 - p)^n$$

- 3 ...erst  $k$  mal gelingt und dann  $n - k$  mal scheitert?

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

- 1 ...immer gelingt?

$$p \cdot p \cdot p \cdots p = p^n$$

- 2 ...immer scheitert?

$$(1 - p) \cdot (1 - p) \cdots (1 - p) = (1 - p)^n$$

- 3 ...erst  $k$  mal gelingt und dann  $n - k$  mal scheitert?

$$p^k \cdot (1 - p)^{n-k}$$

- 4 ...insgesamt  $k$  mal gelingt und  $n - k$  mal scheitert?

Angenommen, ein Ja/Nein-Experiment (z.B. Münzwurf zeigt Kopf) mit Erfolgsws  $p$ , wird  $n$  mal *unabhängig* wiederholt.

Wie groß ist die Wahrscheinlichkeit, dass es...

- 1 ...immer gelingt?

$$p \cdot p \cdot p \cdots p = p^n$$

- 2 ...immer scheitert?

$$(1 - p) \cdot (1 - p) \cdots (1 - p) = (1 - p)^n$$

- 3 ...erst  $k$  mal gelingt und dann  $n - k$  mal scheitert?

$$p^k \cdot (1 - p)^{n-k}$$

- 4 ...insgesamt  $k$  mal gelingt und  $n - k$  mal scheitert?

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



## Erläuterung

$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$  ist die Anzahl der Möglichkeiten, die  $k$  Erfolge in die  $n$  Versuche einzusortieren.

# 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:

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

# Binomialverteilung

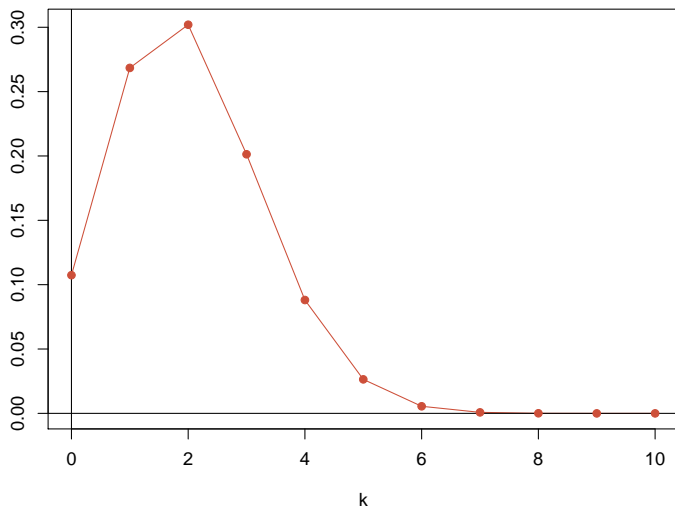
Sei  $X$  die Anzahl der Erfolge bei  $n$  unabhängigen Versuchen mit Erfolgswahrscheinlichkeit von jeweils  $p$ . Dann gilt für  $k \in \{0, 1, \dots, n\}$

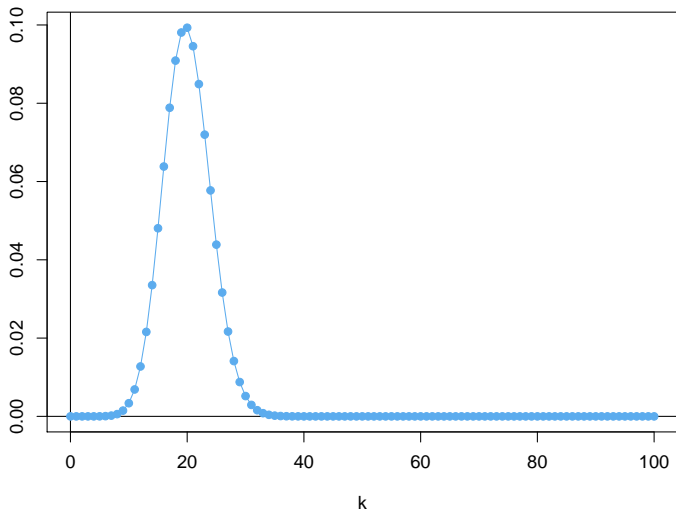
$$\Pr(X = k) = \binom{n}{k} p^k \cdot (1 - p)^{n-k}$$

und  $X$  heißt *binomialverteilt*, kurz:

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

Eine Ja/Nein-Zufallsgröße nennt man auch eine Bernoulli-Zufallsgröße.

probabilities of bin( $n=10, p=0.2$ )

probabilities of bin( $n=100, p=0.2$ )

# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert**
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test

**Beispiel:** In einer Klasse mit 30 Schülern gab es in der letzten Klausur 3x Note 1, 6x Note 2, 8x Note 3, 7x Note 4, 4x Note 5 und 2x Note 6.

**Beispiel:** In einer Klasse mit 30 Schülern gab es in der letzten Klausur 3x Note 1, 6x Note 2, 8x Note 3, 7x Note 4, 4x Note 5 und 2x Note 6. Die Durchschnittsnote war

$$\frac{1}{30} \left( 1+1+1 + \underbrace{2+\dots+2}_{6 \text{ mal}} + \underbrace{3+\dots+3}_{8 \text{ mal}} + \underbrace{4+\dots+4}_{7 \text{ mal}} + 5+5+5+5 + 6+6 \right) = \frac{1}{30} 99 = 3.3$$



**Beispiel:** In einer Klasse mit 30 Schülern gab es in der letzten Klausur 3x Note 1, 6x Note 2, 8x Note 3, 7x Note 4, 4x Note 5 und 2x Note 6. Die Durchschnittsnote war

$$\frac{1}{30} \left( 1+1+1 + \underbrace{2+\dots+2}_{6 \text{ mal}} + \underbrace{3+\dots+3}_{8 \text{ mal}} + \underbrace{4+\dots+4}_{7 \text{ mal}} + 5+5+5+5 + 6+6 \right) = \frac{1}{30} 99 = 3.3$$

Mit relativen Häufigkeiten:

$$1 \cdot \frac{3}{30} + 2 \cdot \frac{6}{30} + 3 \cdot \frac{8}{30} + 4 \cdot \frac{7}{30} + 5 \cdot \frac{4}{30} + 6 \cdot \frac{2}{30} = 3.3$$

**Beispiel:** In einer Klasse mit 30 Schülern gab es in der letzten Klausur 3x Note 1, 6x Note 2, 8x Note 3, 7x Note 4, 4x Note 5 und 2x Note 6. Die Durchschnittsnote war

$$\frac{1}{30} \left( 1+1+1 + \underbrace{2+\dots+2}_{6 \text{ mal}} + \underbrace{3+\dots+3}_{8 \text{ mal}} + \underbrace{4+\dots+4}_{7 \text{ mal}} + 5+5+5+5 + 6+6 \right) = \frac{1}{30} 99 = 3.3$$

Mit relativen Häufigkeiten:

$$1 \cdot \frac{3}{30} + 2 \cdot \frac{6}{30} + 3 \cdot \frac{8}{30} + 4 \cdot \frac{7}{30} + 5 \cdot \frac{4}{30} + 6 \cdot \frac{2}{30} = 3.3$$

**Merke:** Der Durchschnittswert ist die Summe über alle möglichen Werte gewichtet mit den relativen Häufigkeiten

**Beispiel:** In einer Klasse mit 30 Schülern gab es in der letzten Klausur 3x Note 1, 6x Note 2, 8x Note 3, 7x Note 4, 4x Note 5 und 2x Note 6. Die Durchschnittsnote war

$$\frac{1}{30} \left( 1+1+1 + \underbrace{2+\dots+2}_{6 \text{ mal}} + \underbrace{3+\dots+3}_{8 \text{ mal}} + \underbrace{4+\dots+4}_{7 \text{ mal}} + 5+5+5+5 + 6+6 \right) = \frac{1}{30} 99 = 3.3$$

Mit relativen Häufigkeiten:

$$1 \cdot \frac{3}{30} + 2 \cdot \frac{6}{30} + 3 \cdot \frac{8}{30} + 4 \cdot \frac{7}{30} + 5 \cdot \frac{4}{30} + 6 \cdot \frac{2}{30} = 3.3$$

**Merke:** Der Durchschnittswert ist die Summe über alle möglichen Werte gewichtet mit den relativen Häufigkeiten

Synonym zu Durchschnittswert ist das Wort **Erwartungswert**.

## Definition (Erwartungswert)

Sei  $X$  eine Zufallsvariable mit endlichem oder abzählbarem Wertebereich  $\mathcal{S} = \{x_1, x_2, x_3 \dots\} \subseteq \mathbb{R}$ .

## Definition (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

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

## Definition (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

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

Man schreibt auch  $\mu_X$  statt  $\mathbb{E}X$ .

## Definition (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

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

Ersetzt man in der Definition die Wahrscheinlichkeit durch relative Häufigkeiten, so erhält man die bekannte Formel

$$\text{Erwartungswert} = \frac{\text{Summe der Werte}}{\text{Anzahl der Werte}} :$$

## Definition (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

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

Ersetzt man in der Definition die Wahrscheinlichkeit durch relative Häufigkeiten, so erhält man die bekannte Formel

$$\text{Erwartungswert} = \frac{\text{Summe der Werte}}{\text{Anzahl der Werte}} :$$

Sei  $k_x$  die Häufigkeit des Wertes  $x$  in einer Gesamtheit der Größe  $n$ , so schreibt sich der Erwartungswert als

$$\mathbb{E}X = \sum_x x \cdot \frac{k_x}{n} = \frac{\sum_x x \cdot k_x}{n} = \frac{\text{Summe der Werte}}{\text{Anzahl der Werte}} .$$



## Definition (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

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

Beispiele:

- Sei  $X$  eine Ja/Nein Zufallsgröße mit Erfolgswahrscheinlichkeit  $p \in [0, 1]$ . Dann gilt

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

## Definition (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

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

Beispiele:

- Sei  $X$  eine Ja/Nein Zufallsgröße 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

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

## Definition (Erwartungswert)

Sei  $X$  eine Zufallsvariable mit endlichem oder abzählbarem Wertebereich  $\mathcal{S}$ . Sei  $f: \mathcal{S} \rightarrow \mathbb{R}$  eine Funktion.

## Definition (Erwartungswert)

Sei  $X$  eine Zufallsvariable mit endlichem oder abzählbarem Wertebereich  $\mathcal{S}$ . Sei  $f: \mathcal{S} \rightarrow \mathbb{R}$  eine Funktion. Dann ist der *Erwartungswert* von  $f(X)$  definiert durch

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

## Definition (Erwartungswert)

Sei  $X$  eine Zufallsvariable mit endlichem oder abzählbarem Wertebereich  $\mathcal{S}$ . Sei  $f: \mathcal{S} \rightarrow \mathbb{R}$  eine Funktion. Dann ist der *Erwartungswert* von  $f(X)$  definiert durch

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

Beispiel:

Sei  $W$  die Augenzahl bei einem Würfelwurf. Dann gilt

$$\begin{aligned} \mathbb{E}[W^2] &= 1^2 \cdot \Pr(W = 1) + 2^2 \cdot \Pr(W = 2) + \dots + 6^2 \cdot \Pr(W = 6) \\ &= 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + \dots + 6^2 \cdot \frac{1}{6} = 91 \cdot \frac{1}{6} \end{aligned}$$

# Rechnen mit Erwartungswerten

## Satz (Linearität des Erwartungswerts)

*Sind  $X$  und  $Y$  Zufallsvariablen mit Werten in  $\mathbb{R}$  und ist  $a \in \mathbb{R}$ , so gilt:*

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

# Rechnen mit Erwartungswerten

## Satz (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 (Nur für Unabhängige!)

Sind  $X$  und  $Y$  **stochastisch unabhängige** Zufallsvariablen mit Werten in  $\mathbb{R}$ , so gilt

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

# Rechnen mit Erwartungswerten

## Satz (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 (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:

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



Beweis der Linearität: Sei  $\mathcal{S}$  der Zustandsraum von  $X$  und  $Y$  und seien  $a, b \in \mathbb{R}$ .

$$\begin{aligned} & \mathbb{E}(a \cdot X + b \cdot Y) \\ &= \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} (a \cdot x + b \cdot y) \Pr(X = x, Y = y) \\ &= a \cdot \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} x \Pr(X = x, Y = y) + b \cdot \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} y \Pr(X = x, Y = y) \\ &= a \cdot \sum_{x \in \mathcal{S}} x \sum_{y \in \mathcal{S}} \Pr(X = x, Y = y) + b \cdot \sum_{y \in \mathcal{S}} y \sum_{x \in \mathcal{S}} \Pr(X = x, Y = y) \\ &= a \cdot \sum_{x \in \mathcal{S}} x \Pr(X = x) + b \cdot \sum_{y \in \mathcal{S}} y \Pr(Y = y) \\ &= a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y) \end{aligned}$$

Beweis der Produktformel: Sei  $\mathcal{S}$  der Zustandsraum von  $X$  und  $Y$  und seien  $X$  und  $Y$  (stochastisch) **unabhängig**.

$$\begin{aligned}\mathbb{E}(X \cdot Y) &= \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} (x \cdot y) \Pr(X = x, Y = y) \\ &= \sum_{x \in \mathcal{S}} \sum_{y \in \mathcal{S}} (x \cdot y) \Pr(X = x) \Pr(Y = y) \\ &= \sum_{x \in \mathcal{S}} x \Pr(X = x) \cdot \sum_{y \in \mathcal{S}} y \Pr(Y = y) \\ &= \mathbb{E}X \cdot \mathbb{E}Y.\end{aligned}$$

# Erwartungswert der Binomialverteilung

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

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

# 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.

# 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

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

# 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

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

# 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:

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

# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation**
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test



## Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer  $\mathbb{R}$ -wertigen Zufallsgröße  $X$  ist

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

## Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer  $\mathbb{R}$ -wertigen Zufallsgröße  $X$  ist

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

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

## Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer  $\mathbb{R}$ -wertigen Zufallsgröße  $X$  ist

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

$\sigma_X = \sqrt{\text{Var} X}$  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  $Y$ .

## Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer  $\mathbb{R}$ -wertigen Zufallsgröße  $X$  ist

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

$\sigma_X = \sqrt{\text{Var} X}$  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

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

ist die durchschnittliche quadrierte Abweichung vom Mittelwert.

## Die Varianz

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

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.

## Die Varianz

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

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.

$$\text{Cor}(X, Y) = 0.$$

# 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.  $\Pr(X = x_i) = \frac{1}{n}$ ), so gilt:

$$\mathbb{E}X = \sum_{i=1}^n x_i \Pr(X = x_i) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

und

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

# 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.  $\Pr(X = x_i) = \frac{1}{n}$ ), so gilt:

$$\mathbb{E}X = \sum_{i=1}^n x_i \Pr(X = x_i) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

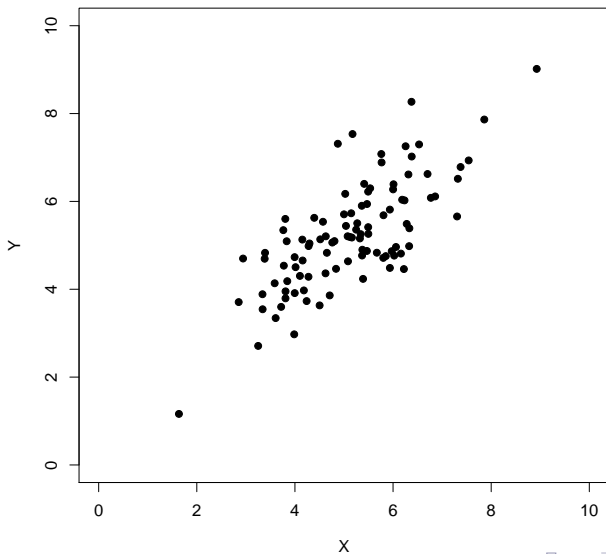
und

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

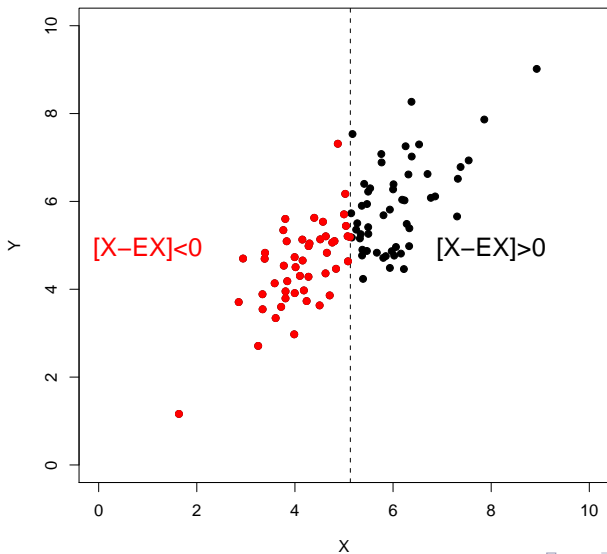
Sind  $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R} \times \mathbb{R}$  Daten und entsteht  $(X, Y)$  durch rein zufälliges Ziehen aus diesen Daten (d.h.  $\Pr((X, Y) = (x_i, y_i)) = \frac{1}{n}$ ), so gilt:

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

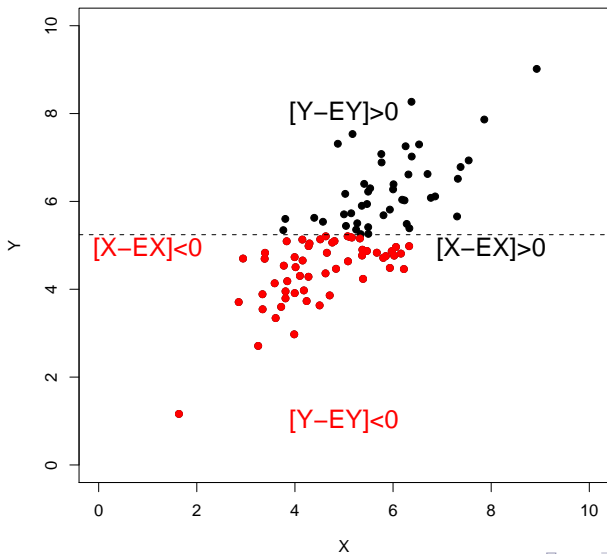
# Wieso $\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}X][Y - \mathbb{E}Y])$ ?



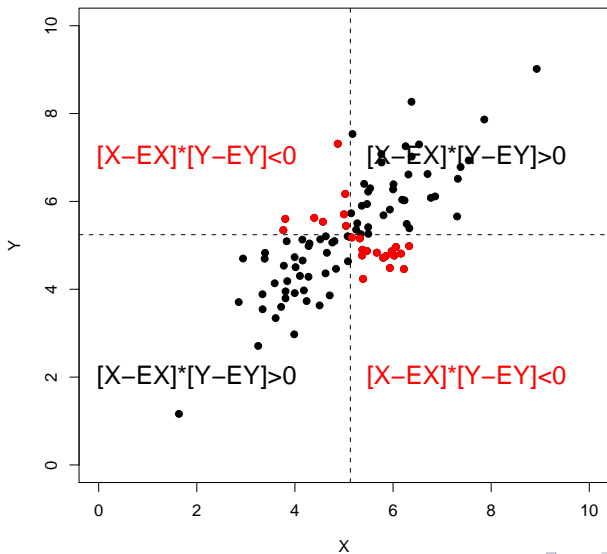
# Wieso $\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}X][Y - \mathbb{E}Y])$ ?



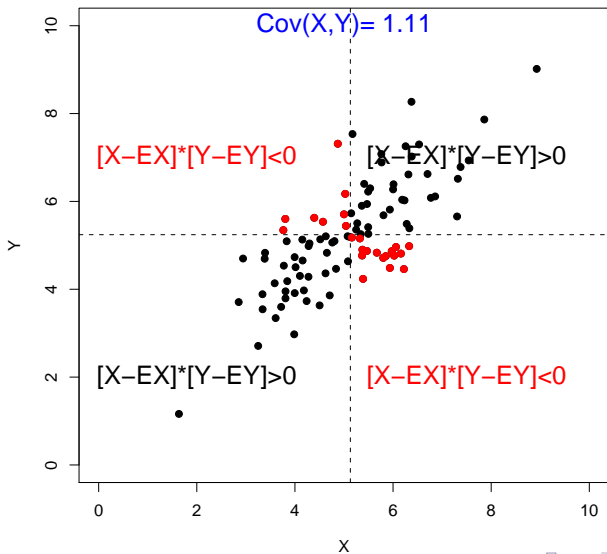
# Wieso $\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}X][Y - \mathbb{E}Y])$ ?



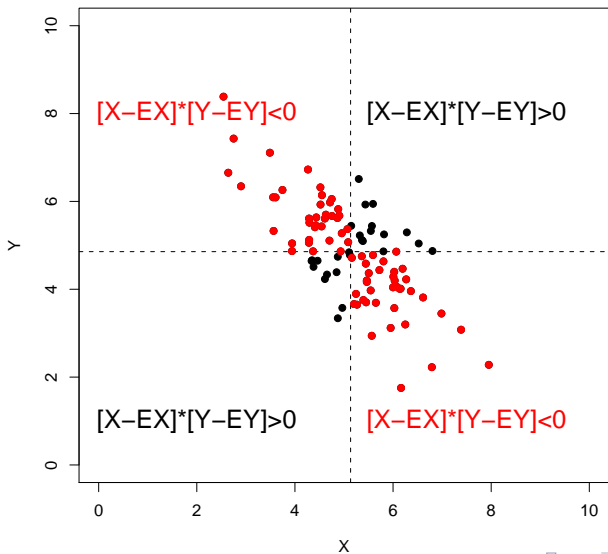
# Wieso $\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}X][Y - \mathbb{E}Y])$ ?



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

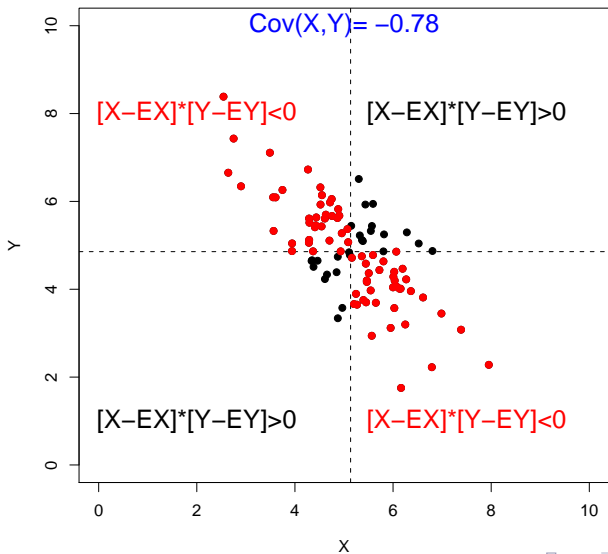


# Wieso $\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}X][Y - \mathbb{E}Y])$ ?

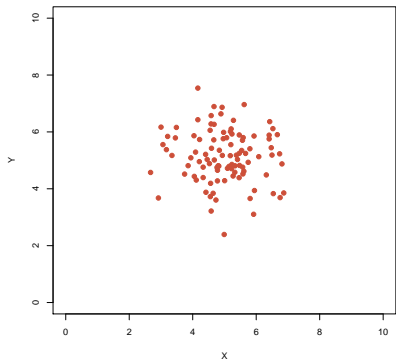




# Wieso $\text{Cov}(X, Y) = \mathbb{E}([X - \mathbb{E}X][Y - \mathbb{E}Y])$ ?

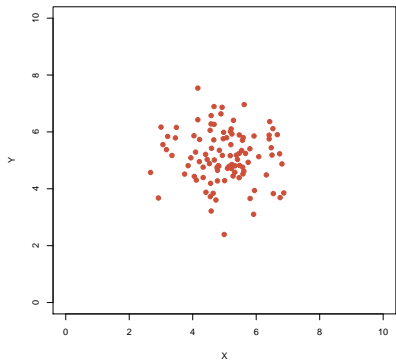


$$\sigma_X = 0.95, \sigma_Y = 0.92$$



$$\sigma_X = 0.95, \sigma_Y = 0.92$$

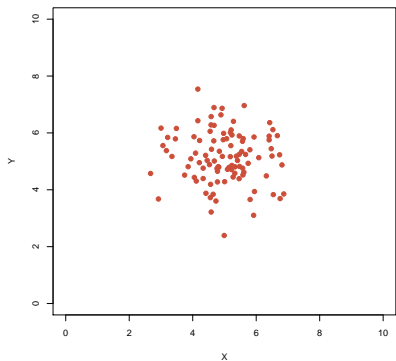
$$\text{Cov}(X, Y) = -0.06$$



$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

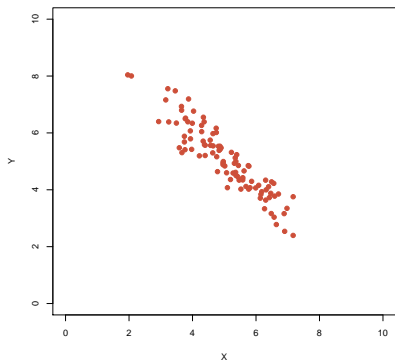
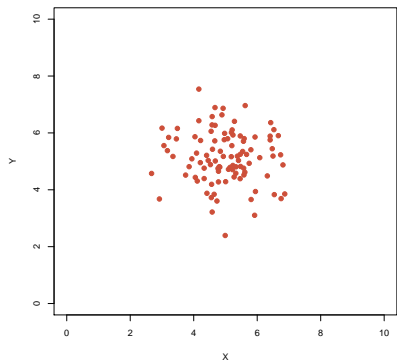


$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$



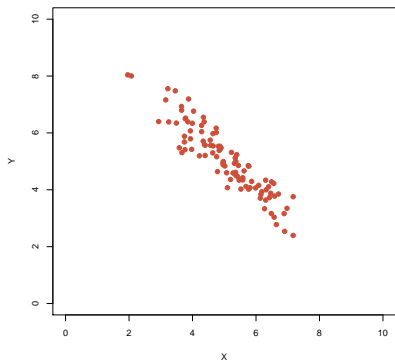
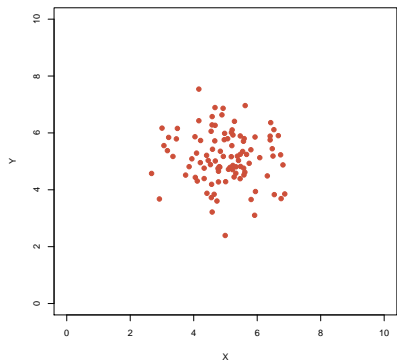
$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$



$$\sigma_X = 0.95, \sigma_Y = 0.92$$

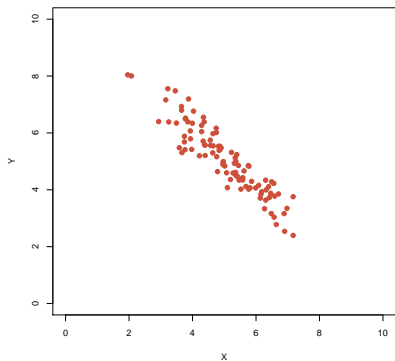
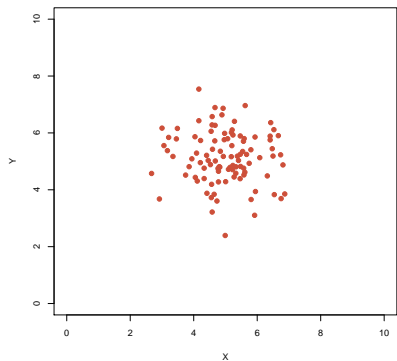
$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

$$\text{Cor}(X, Y) = -0.92$$

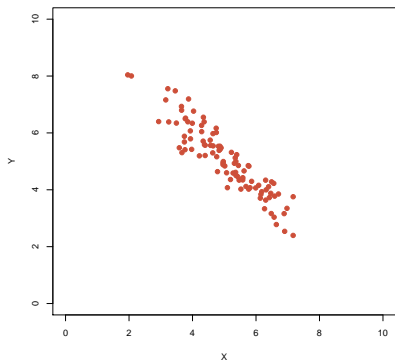
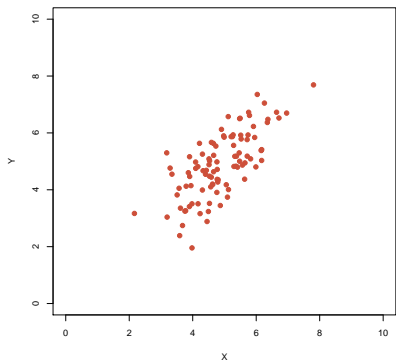


$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

$$\text{Cor}(X, Y) = -0.92$$





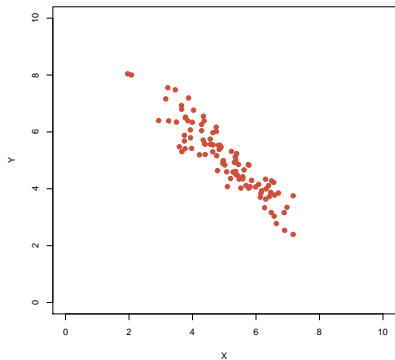
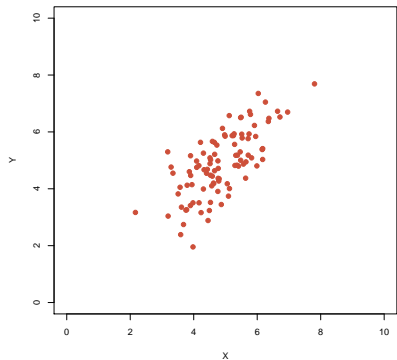
$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

$$\text{Cor}(X, Y) = -0.92$$



$$\sigma_X = 1.14, \sigma_Y = 0.78$$

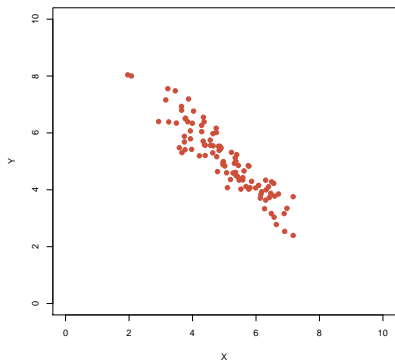
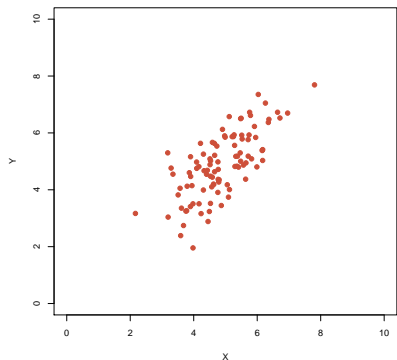
$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

$$\text{Cor}(X, Y) = -0.92$$

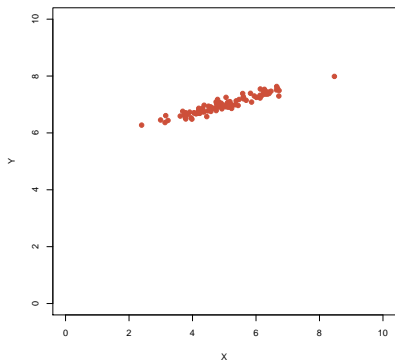
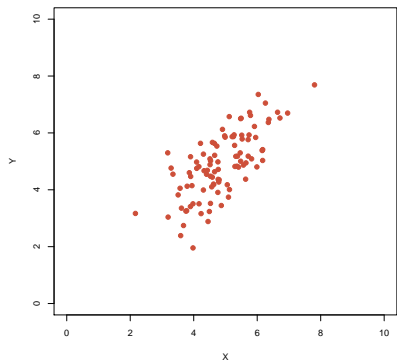


$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$



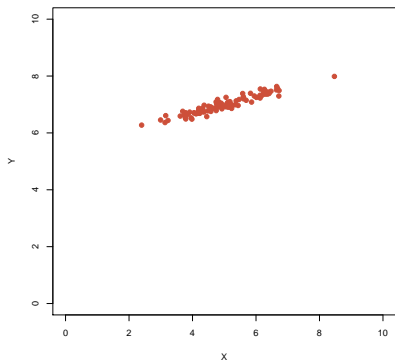
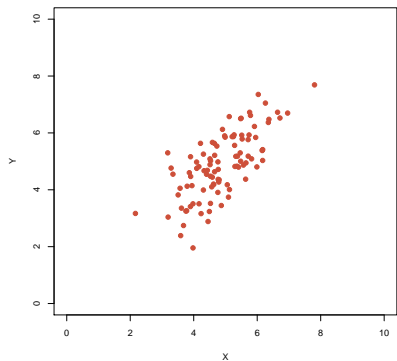
$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$



$$\sigma_X = 1.14, \sigma_Y = 0.78$$

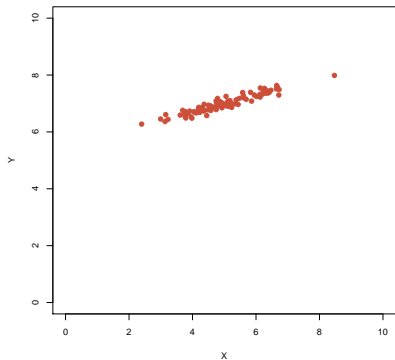
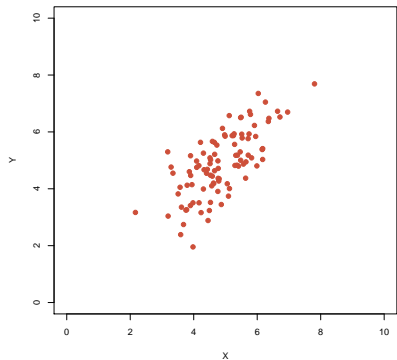
$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$

$$\text{Cor}(X, Y) = 0.95$$

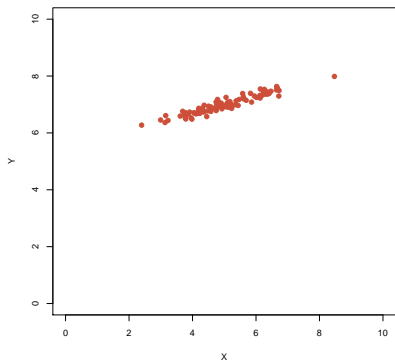
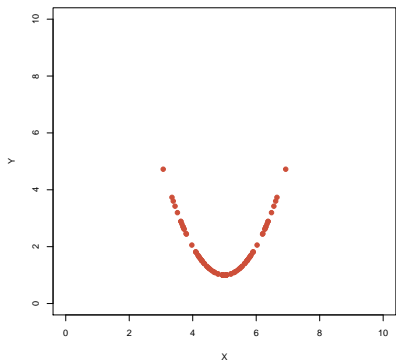


$$\sigma_X = 0.91, \sigma_Y = 0.88$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$

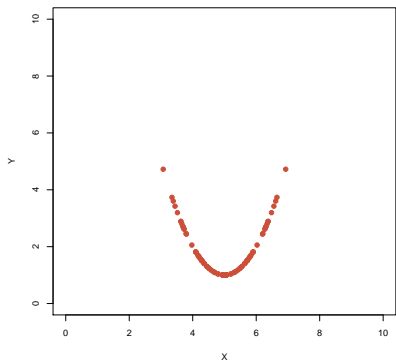
$$\text{Cov}(X, Y) = 0.32$$

$$\text{Cor}(X, Y) = 0.95$$



$$\sigma_X = 0.91, \sigma_Y = 0.88$$

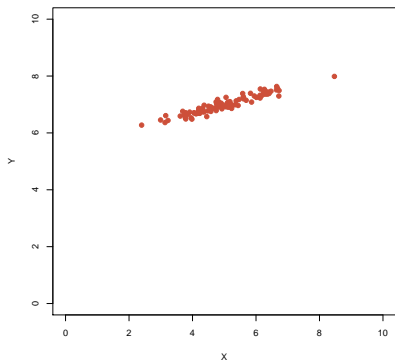
$$\text{Cov}(X, Y) = 0$$



$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$

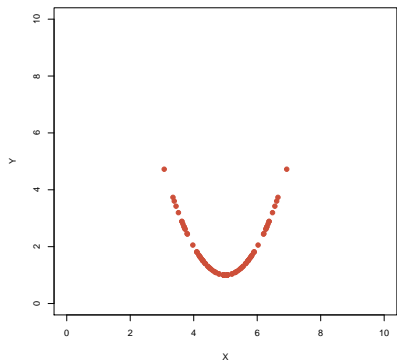
$$\text{Cor}(X, Y) = 0.95$$



$$\sigma_X = 0.91, \sigma_Y = 0.88$$

$$\text{Cov}(X, Y) = 0$$

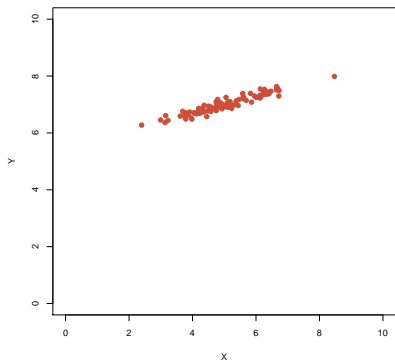
$$\text{Cor}(X, Y) = 0$$



$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$

$$\text{Cor}(X, Y) = 0.95$$





# Rechenregeln für Varianzen

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

- $\text{Var}X = \text{Cov}(X, X)$

# Rechenregeln für Varianzen

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

- $\text{Var}X = \text{Cov}(X, X)$
- $\text{Var}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$

# Rechenregeln für Varianzen

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

- $\text{Var}X = \text{Cov}(X, X)$
- $\text{Var}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$
- $\text{Var}(a \cdot X) = a^2 \cdot \text{Var}X$

# Rechenregeln für Varianzen

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

- $\text{Var}X = \text{Cov}(X, X)$
- $\text{Var}X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$
- $\text{Var}(a \cdot X) = a^2 \cdot \text{Var}X$
- $\text{Var}(X + Y) = \text{Var}X + \text{Var}Y + 2 \cdot \text{Cov}(X, Y)$

# Rechenregeln für Varianzen

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

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

# 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:

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

# Rechenregeln für Kovarianzen

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

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

# Rechenregeln für Kovarianzen

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

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



# Rechenregeln für Kovarianzen

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

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

# Rechenregeln für Kovarianzen

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

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

# Rechenregeln für Kovarianzen

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

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

# Rechenregeln für Kovarianzen

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

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

# Rechenregeln für Kovarianzen

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

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

Die letzten drei Regeln beschreiben die Bilinearität der Kovarianz.

# Rechenregeln für die Korrelation

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

- $-1 \leq \text{Cor}(X, Y) \leq 1$
- $\text{Cor}(X, Y) = \text{Cor}(Y, X)$
- $\text{Cor}(X, Y) = \text{Cov}(X/\sigma_X, Y/\sigma_Y)$

# 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  $Y = a \cdot X + b$

# 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  $Y = a \cdot X + b$



Mit diesen Rechenregeln können wir nun endlich beweisen:

## Satz

Sind  $X_1, X_2, \dots, X_n$  unabhängige  $\mathbb{R}$ -wertige Zufallsgrößen mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ , 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.

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

Mit diesen Rechenregeln können wir nun endlich beweisen:

## Satz

Sind  $X_1, X_2, \dots, X_n$  unabhängige  $\mathbb{R}$ -wertige Zufallsgrößen mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ , 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  $\frac{\sigma}{\sqrt{n}}$  ist ein Schätzer der Standardabweichung  $\sigma_{\bar{X}}$  des Stichprobenmittels  $\bar{X}$  der Stichprobe  $(X_1, X_2, \dots, X_n)$ .

Mit diesen Rechenregeln können wir nun endlich beweisen:

## Satz

Sind  $X_1, X_2, \dots, X_n$  unabhängige  $\mathbb{R}$ -wertige Zufallsgrößen mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ , 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  $\sigma$  der Grundgesamtheit.

**Beweis:** Linearität des Erwartungswertes impliziert

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

**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

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

# Bernoulli-Verteilung

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

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

und Varianz

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

# 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  $\Pr(Y = 1) = p$  und  $\Pr(Y = 0) = (1 - p)$  folgt

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

# 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  $\Pr(Y = 1) = p$  und  $\Pr(Y = 0) = (1 - p)$  folgt

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

Varianz:

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



# 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:

$$\text{Var } X =$$

# 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:

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

# 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:

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

# 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 (Erwartungswert und Varianz der Binomialverteilung)

*Ist  $X$  binomialverteilt mit Parametern  $(n, p)$ , so gilt:*

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

*und*

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

# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel**
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test

In



E.N. Moriyama (2003) Codon Usage

*Encyclopedia of the human genome*, Macmillan Publishers Ltd.

werden u.a. 9497 menschliche Gene auf “Codon Bias” untersucht.

In diesen Genen wird die Aminosäure Prolin 16710 mal durch das Codon CCT und 18895 mal durch das Codon CCC codiert.

In



E.N. Moriyama (2003) Codon Usage

*Encyclopedia of the human genome*, Macmillan Publishers Ltd.

werden u.a. 9497 menschliche Gene auf “Codon Bias” untersucht.

In diesen Genen wird die Aminosäure Prolin 16710 mal durch das Codon CCT und 18895 mal durch das Codon CCC codiert. Ist es nur vom reinen Zufall abhängig, welches Codon verwendet wird?



In

 E.N. Moriyama (2003) Codon Usage

*Encyclopedia of the human genome*, Macmillan Publishers Ltd.

werden u.a. 9497 menschliche Gene auf “Codon Bias” untersucht.

In diesen Genen wird die Aminosäure Prolin 16710 mal durch das Codon CCT und 18895 mal durch das Codon CCC codiert. Ist es nur vom reinen Zufall abhängig, welches Codon verwendet wird?

Dann wäre die Anzahl  $X$  der CCC binomialverteilt mit  $p = \frac{1}{2}$  und  $n = 16710 + 18895 = 35605$ .

Angenommen die Anzahl  $X$  ( $= 18895$ ) der CCC ist binomialverteilt mit  $p = \frac{1}{2}$  und  $n = 16710 + 18895 = 35605$ .

Angenommen die Anzahl  $X$  ( $= 18895$ ) der CCC ist binomialverteilt mit  $p = \frac{1}{2}$  und  $n = 16710 + 18895 = 35605$ .

$$\mathbb{E}X = n \cdot p = 17802.5$$

Angenommen die Anzahl  $X$  ( $= 18895$ ) der CCC ist binomialverteilt mit  $p = \frac{1}{2}$  und  $n = 16710 + 18895 = 35605$ .

$$\mathbb{E}X = n \cdot p = 17802.5$$

$$\sigma_X = \sqrt{n \cdot p \cdot (1 - p)} \approx 94.34$$

Angenommen die Anzahl  $X$  ( $= 18895$ ) der CCC ist binomialverteilt mit  $p = \frac{1}{2}$  und  $n = 16710 + 18895 = 35605$ .

$$\mathbb{E}X = n \cdot p = 17802.5$$

$$\sigma_X = \sqrt{n \cdot p \cdot (1 - p)} \approx 94.34$$

$$18895 - 17802.5 = 1092.5 \approx 11.6 \cdot \sigma_X$$

Angenommen die Anzahl  $X$  ( $= 18895$ ) der CCC ist binomialverteilt mit  $p = \frac{1}{2}$  und  $n = 16710 + 18895 = 35605$ .

$$\mathbb{E}X = n \cdot p = 17802.5$$

$$\sigma_X = \sqrt{n \cdot p \cdot (1 - p)} \approx 94.34$$

$$18895 - 17802.5 = 1092.5 \approx 11.6 \cdot \sigma_X$$

Sieht das nach Zufall aus?

Die Frage ist:

Wie groß ist die Wahrscheinlichkeit  
einer Abweichung vom Erwartungswert  
von mindestens  $\approx 11.6 \cdot \sigma_X$ , wenn alles Zufall ist?

Die Frage ist:

Wie groß ist die Wahrscheinlichkeit  
einer Abweichung vom Erwartungswert  
von mindestens  $\approx 11.6 \cdot \sigma_X$ , wenn alles Zufall ist?

Wir müssen also

$$\Pr(|X - \mathbb{E}X| \geq 11.6\sigma_X)$$

berechnen.



Die Frage ist:

Wie groß ist die Wahrscheinlichkeit  
einer Abweichung vom Erwartungswert  
von mindestens  $\approx 11.6 \cdot \sigma_X$ , wenn alles Zufall ist?

Wir müssen also

$$\Pr(|X - \mathbb{E}X| \geq 11.6\sigma_X)$$

berechnen.

Das Problem bei der Binomialverteilung ist:  $\binom{n}{k}$  exakt zu berechnen,  
ist für große  $n$  sehr aufwändig. Deshalb:

Die Frage ist:

Wie groß ist die Wahrscheinlichkeit  
einer Abweichung vom Erwartungswert  
von mindestens  $\approx 11.6 \cdot \sigma_X$ , wenn alles Zufall ist?

Wir müssen also

$$\Pr(|X - \mathbb{E}X| \geq 11.6\sigma_X)$$

berechnen.

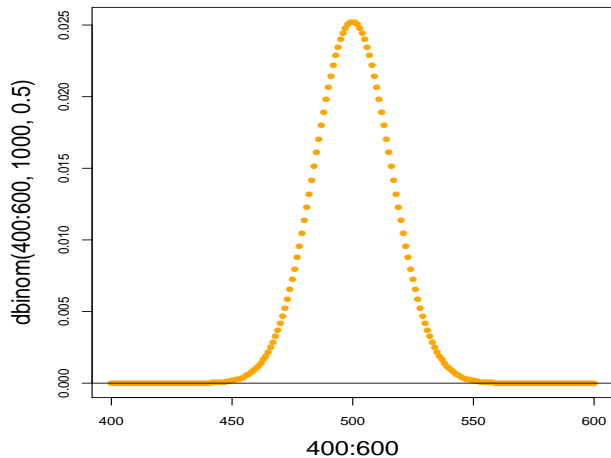
Das Problem bei der Binomialverteilung ist:  $\binom{n}{k}$  exakt zu berechnen,  
ist für große  $n$  sehr aufwändig. Deshalb:

Die Binomialverteilung wird oft  
durch andere Verteilungen approximiert.

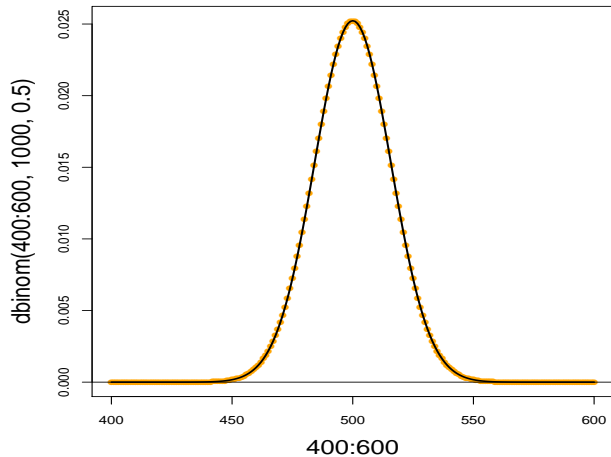
# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung**
- 8 Normalapproximation
- 9 Der z-Test

Die Binomialverteilung mit großer Versuchszahl  $n$  sieht aus wie die Normalverteilung:



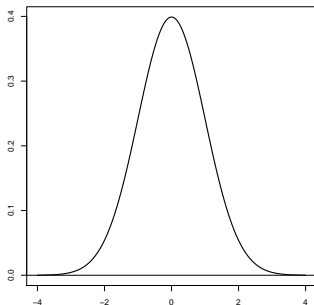
Die Binomialverteilung mit großer Versuchszahl  $n$  sieht aus wie die Normalverteilung:



# Dichte der Standardnormalverteilung

Eine Zufallsvariable  $Z$  mit der Dichte

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



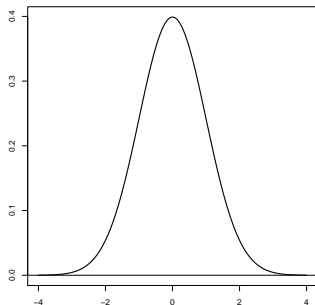
“Gauß-Glocke”

heißt *standardnormalverteilt*.

# Dichte der Standardnormalverteilung

Eine Zufallsvariable  $Z$  mit der Dichte

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



“Gauß-Glocke”

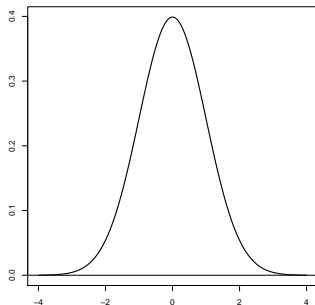
kurz:  $Z \sim \mathcal{N}(0, 1)$

heißt *standardnormalverteilt*.

# Dichte der Standardnormalverteilung

Eine Zufallsvariable  $Z$  mit der Dichte

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



“Gauß-Glocke”

kurz:  $Z \sim \mathcal{N}(0, 1)$

$$\mathbb{E}Z = 0$$

$$\text{Var } Z = 1$$

heißt *standardnormalverteilt*.



Ist  $Z \mathcal{N}(0, 1)$ -verteilt, so ist  $X = \sigma \cdot Z + \mu$  normalverteilt mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ , kurz:

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

Ist  $Z \mathcal{N}(0, 1)$ -verteilt, so ist  $X = \sigma \cdot Z + \mu$  normalverteilt mit Mittelwert  $\mu$  und Varianz  $\sigma^2$ , kurz:

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

$X$  hat dann die Dichte

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

# Merkregeln

Ist  $Z \sim \mathcal{N}(\mu, \sigma^2)$ , so gilt:

- $\Pr(|Z - \mu| > \sigma) \approx 33\%$

# Merkregeln

Ist  $Z \sim \mathcal{N}(\mu, \sigma^2)$ , so gilt:

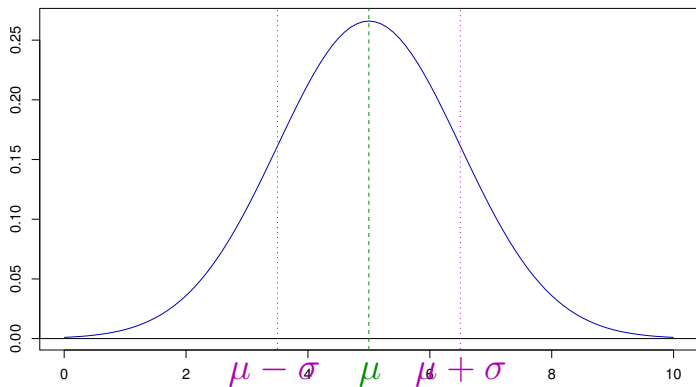
- $\Pr(|Z - \mu| > \sigma) \approx 33\%$
- $\Pr(|Z - \mu| > 1.96 \cdot \sigma) \approx 5\%$

# Merkregeln

Ist  $Z \sim \mathcal{N}(\mu, \sigma^2)$ , so gilt:

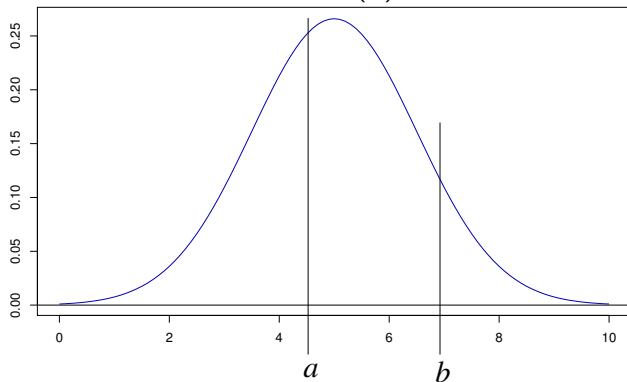
- $\Pr(|Z - \mu| > \sigma) \approx 33\%$
- $\Pr(|Z - \mu| > 1.96 \cdot \sigma) \approx 5\%$
- $\Pr(|Z - \mu| > 3 \cdot \sigma) \approx 0.3\%$

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



# Dichten brauchen Integrale

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

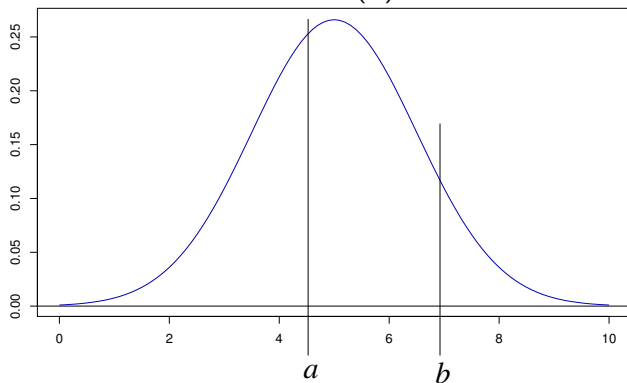


Dann gilt

$$\Pr(Z \in [a, b]) =$$

# 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  $\Pr(Z = 5)$ ?

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

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

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

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:

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

# Die Normalverteilung in $R$

Die Normalverteilung hat in  $R$  das Kürzel '*norm*'.

# Die Normalverteilung in $R$

Die Normalverteilung hat in  $R$  das Kürzel '*norm*'.

Es gibt 4  $R$ -Befehle:

`dnorm()`: Dichte der Normalverteilung (**d**ensity)

# 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 sample)

# 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 sample)

**pnorm()**: Verteilungsfunktion der Normalverteilung (**p**robability)



# Die Normalverteilung in *R*

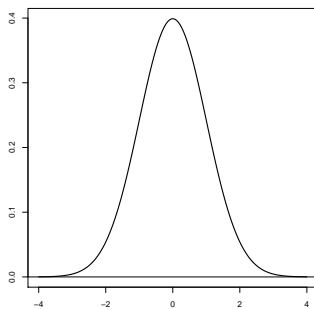
Die Normalverteilung hat in *R* das Kürzel '*norm*'.

Es gibt 4 *R*-Befehle:

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

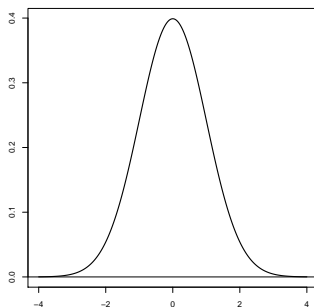
**Beispiel:** Dichte der Standardnormalverteilung:

```
> plot(dnorm,from=-4,to=4)
```



**Beispiel:** Dichte der Standardnormalverteilung:

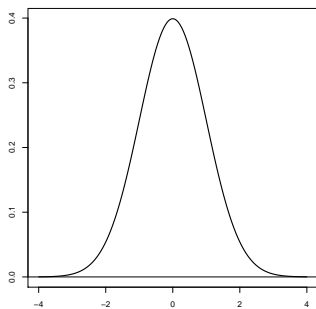
```
> plot(dnorm,from=-4,to=4)
```



```
> dnorm(0)  
[1] 0.3989423
```

**Beispiel:** Dichte der Standardnormalverteilung:

```
> plot(dnorm,from=-4,to=4)
```



```
> dnorm(0)
```

```
[1] 0.3989423
```

```
> dnorm(0,mean=1,sd=2)
```

```
[1] 0.1760327
```

## **Beispiel:** Ziehen einer Stichprobe

**Beispiel:** Ziehen einer Stichprobe

Ziehen einer Stichprobe der Länge 6 aus einer Standardnormalverteilung:

```
> rnorm(6)
[1] -1.24777899 0.03288728 0.19222813 0.81642692
-0.62607324 -1.09273888
```

**Beispiel:** Ziehen einer Stichprobe

Ziehen einer Stichprobe der Länge 6 aus einer Standardnormalverteilung:

```
> rnorm(6)
[1] -1.24777899 0.03288728 0.19222813 0.81642692
-0.62607324 -1.09273888
```

Ziehen einer Stichprobe der Länge 7 aus einer Normalverteilung mit Mittelwert 5 und Standardabweichung 3:

```
> rnorm(7,mean=5,sd=3)
[1] 2.7618897 6.3224503 10.8453280 -0.9829688 5.6143127
0.6431437 8.123570
```

**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$ , also standardnormalverteilt.

$\Pr(Z < a)$  berechnet man in R mit `pnorm(a)`



**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$ , also standardnormalverteilt.

$\Pr(Z < a)$  berechnet man in R mit `pnorm(a)`

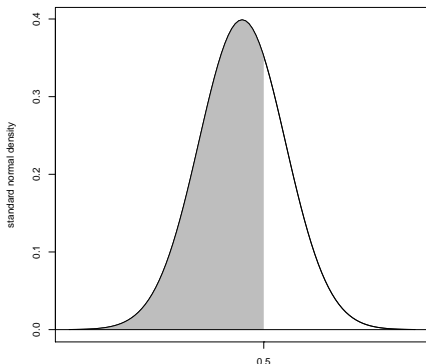
```
> pnorm(0.5) [1] 0.6914625
```

**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$ , also standardnormalverteilt.

$\Pr(Z < a)$  berechnet man in R mit `pnorm(a)`

```
> pnorm(0.5) [1] 0.6914625
```



**Beispiel:** Berechnung von Wahrscheinlichkeiten:  
Sei  $Z \sim \mathcal{N}(\mu = 5, \sigma^2 = 2.25)$ .

**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 5, \sigma^2 = 2.25)$ .

Berechnung von  $\Pr(Z \in [3, 4])$ :

**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 5, \sigma^2 = 2.25)$ .

Berechnung von  $\Pr(Z \in [3, 4])$ :

$$\Pr(Z \in [3, 4]) = \Pr(Z < 4) - \Pr(Z < 3)$$

**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 5, \sigma^2 = 2.25)$ .

Berechnung von  $\Pr(Z \in [3, 4])$ :

$$\Pr(Z \in [3, 4]) = \Pr(Z < 4) - \Pr(Z < 3)$$

```
> pnorm(4, mean=5, sd=1.5) - pnorm(3, mean=5, sd=1.5)
```

**Beispiel:** Berechnung von Wahrscheinlichkeiten:

Sei  $Z \sim \mathcal{N}(\mu = 5, \sigma^2 = 2.25)$ .

Berechnung von  $\Pr(Z \in [3, 4])$ :

$$\Pr(Z \in [3, 4]) = \Pr(Z < 4) - \Pr(Z < 3)$$

```
> pnorm(4, mean=5, sd=1.5) - pnorm(3, mean=5, sd=1.5)
[1] 0.1612813
```

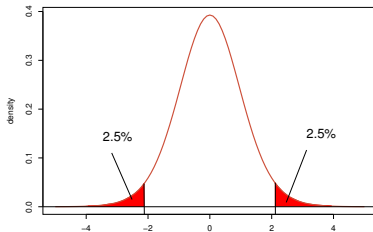
**Beispiel:** Berechnung von Quantilen:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$  standardnormalverteilt. Für welchen Wert  $z$  gilt  $\Pr(|Z| > z) = 5\%$ ?



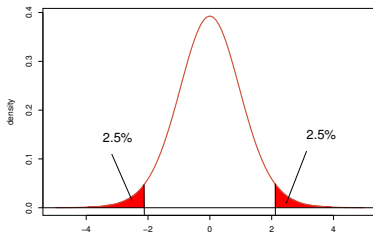
**Beispiel:** Berechnung von Quantilen:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$  standardnormalverteilt. Für welchen Wert  $z$  gilt  $\Pr(|Z| > z) = 5\%$ ?



**Beispiel:** Berechnung von Quantilen:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$  standardnormalverteilt. Für welchen Wert  $z$  gilt  $\Pr(|Z| > z) = 5\%$ ?



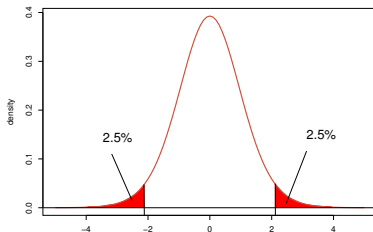
Wegen der Symmetrie bzgl der y-Achse gilt

$$\Pr(|Z| > z) = \Pr(Z < -z) + \Pr(Z > z) = 2 \cdot \Pr(Z < -z)$$

Finde also  $z > 0$ , so dass  $\Pr(Z < -z) = 2.5\%$ .

**Beispiel:** Berechnung von Quantilen:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$  standardnormalverteilt. Für welchen Wert  $z$  gilt  $\Pr(|Z| > z) = 5\%$ ?



Wegen der Symmetrie bzgl der y-Achse gilt

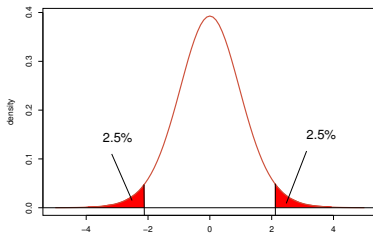
$$\Pr(|Z| > z) = \Pr(Z < -z) + \Pr(Z > z) = 2 \cdot \Pr(Z < -z)$$

Finde also  $z > 0$ , so dass  $\Pr(Z < -z) = 2.5\%$ .

```
> qnorm(0.025, mean=0, sd=1)
```

**Beispiel:** Berechnung von Quantilen:

Sei  $Z \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$  standardnormalverteilt. Für welchen Wert  $z$  gilt  $\Pr(|Z| > z) = 5\%$ ?



Wegen der Symmetrie bzgl der y-Achse gilt

$$\Pr(|Z| > z) = \Pr(Z < -z) + \Pr(Z > z) = 2 \cdot \Pr(Z < -z)$$

Finde also  $z > 0$ , so dass  $\Pr(Z < -z) = 2.5\%$ .

```
> qnorm(0.025, mean=0, sd=1)
```

```
[1] -1.959964
```

Antwort:  $z \approx 1.96$ , also knapp 2 Standardabweichungen

# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation**
- 9 Der z-Test

# Normalapproximation

Für große  $n$  und  $p$ , die nicht zu nahe bei 0 oder 1 liegen, kann man die Binomialverteilung durch die Normalverteilung mit dem entsprechenden Erwartungswert und der entsprechenden Varianz approximieren:

# Normalapproximation

Für große  $n$  und  $p$ , die nicht zu nahe bei 0 oder 1 liegen, kann man die Binomialverteilung durch die Normalverteilung mit dem entsprechenden Erwartungswert und der entsprechenden Varianz approximieren:

Ist  $X \sim \text{bin}(n, p)$  und  $Z \sim \mathcal{N}(\mu = \quad, \sigma^2 = \quad)$ , so gilt

$$\Pr(X \in [a, b]) \approx \Pr(Z \in [a, b])$$

# Normalapproximation

Für große  $n$  und  $p$ , die nicht zu nahe bei 0 oder 1 liegen, kann man die Binomialverteilung durch die Normalverteilung mit dem entsprechenden Erwartungswert und der entsprechenden Varianz approximieren:

Ist  $X \sim \text{bin}(n, p)$  und  $Z \sim \mathcal{N}(\mu = n \cdot p, \sigma^2 = \quad)$ , so gilt

$$\Pr(X \in [a, b]) \approx \Pr(Z \in [a, b])$$



# Normalapproximation

Für große  $n$  und  $p$ , die nicht zu nahe bei 0 oder 1 liegen, kann man die Binomialverteilung durch die Normalverteilung mit dem entsprechenden Erwartungswert und der entsprechenden Varianz approximieren:

Ist  $X \sim \text{bin}(n, p)$  und  $Z \sim \mathcal{N}(\mu = n \cdot p, \sigma^2 = n \cdot p \cdot (1 - p))$ , so gilt

$$\Pr(X \in [a, b]) \approx \Pr(Z \in [a, b])$$

# Normalapproximation

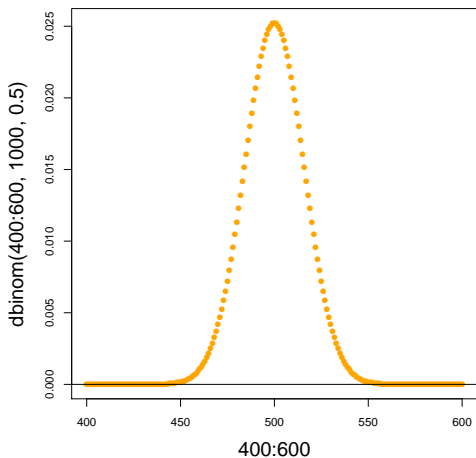
Für große  $n$  und  $p$ , die nicht zu nahe bei 0 oder 1 liegen, kann man die Binomialverteilung durch die Normalverteilung mit dem entsprechenden Erwartungswert und der entsprechenden Varianz approximieren:

Ist  $X \sim \text{bin}(n, p)$  und  $Z \sim \mathcal{N}(\mu = n \cdot p, \sigma^2 = n \cdot p \cdot (1 - p))$ , so gilt

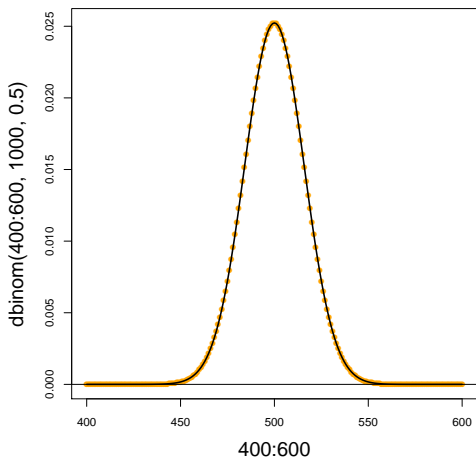
$$\Pr(X \in [a, b]) \approx \Pr(Z \in [a, b])$$

(eine Faustregel: für den Hausgebrauch meist okay, wenn  $n \cdot p \cdot (1 - p) \geq 9$ )

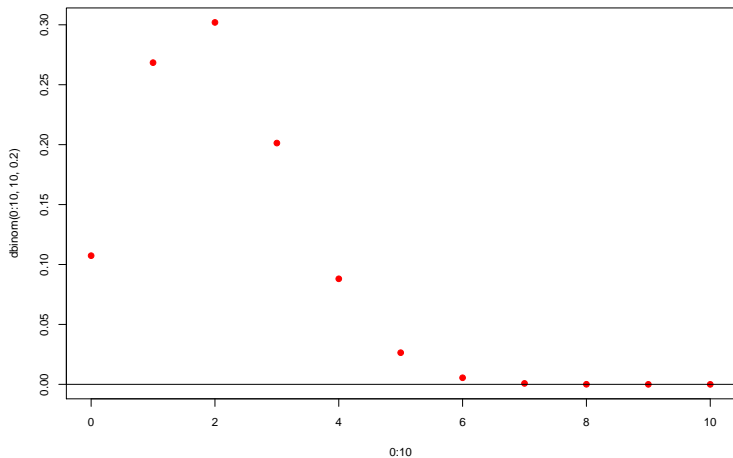
$$n = 1000, p = 0.5, n \cdot p \cdot (1 - p) = 250$$



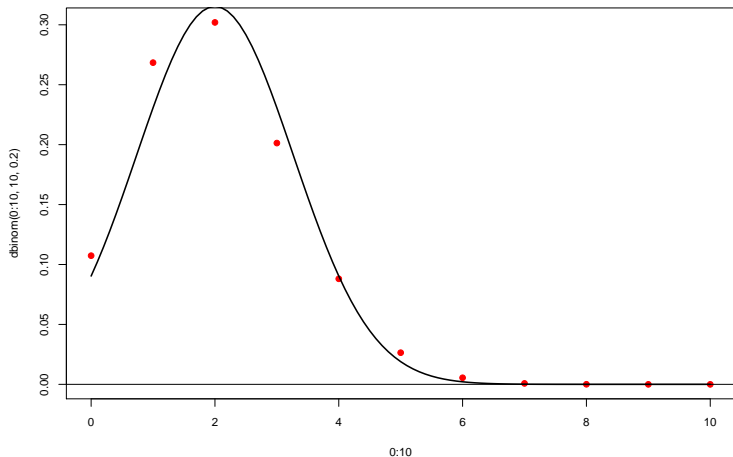
$$n = 1000, p = 0.5, n \cdot p \cdot (1 - p) = 250$$



$$n = 10, p = 0.2, n \cdot p \cdot (1 - p) = 1.6$$



$$n = 10, p = 0.2, n \cdot p \cdot (1 - p) = 1.6$$



# Zentraler Grenzwertsatz

Ein anderer Ausdruck für *Normalapproximation* ist  
**Zentraler Grenzwertsatz.**

# Zentraler Grenzwertsatz

Ein anderer Ausdruck für *Normalapproximation* ist  
**Zentraler Grenzwertsatz.**

Der zentrale Grenzwertsatz besagt,  
dass die Verteilung von Summen  
**unabhängiger und identisch verteilter**  
Zufallsvariablen in etwa  
die Normalverteilung ist.



## Theorem (Zentraler Grenzwertsatz)

Die  $\mathbb{R}$ -wertigen Zufallsgrößen  $X_1, X_2, \dots$  seien unabhängig und identisch verteilt mit endlicher Varianz  $0 < \text{Var } X_i < \infty$ . Sei außerdem

$$Z_n := X_1 + X_2 + \dots + X_n$$

die Summe der ersten  $n$  Variablen.

## Theorem (Zentraler Grenzwertsatz)

Die  $\mathbb{R}$ -wertigen Zufallsgrößen  $X_1, X_2, \dots$  seien unabhängig und identisch verteilt mit endlicher Varianz  $0 < \text{Var } X_i < \infty$ . Sei außerdem

$$Z_n := X_1 + X_2 + \dots + X_n$$

die Summe der ersten  $n$  Variablen. Dann ist die zentrierte und reskalierte Summe im Limes  $n \rightarrow \infty$  standardnormalverteilt, d.h.

$$\frac{Z_n - \mathbb{E}Z_n}{\sqrt{\text{Var } Z_n}} \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$$

bei  $n \rightarrow \infty$ .

## Theorem (Zentraler Grenzwertsatz)

Die  $\mathbb{R}$ -wertigen Zufallsgrößen  $X_1, X_2, \dots$  seien unabhängig und identisch verteilt mit endlicher Varianz  $0 < \text{Var } X_i < \infty$ . Sei außerdem

$$Z_n := X_1 + X_2 + \dots + X_n$$

die Summe der ersten  $n$  Variablen. Dann ist die zentrierte und reskalierte Summe im Limes  $n \rightarrow \infty$  standardnormalverteilt, d.h.

$$\frac{Z_n - \mathbb{E}Z_n}{\sqrt{\text{Var } Z_n}} \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$$

bei  $n \rightarrow \infty$ . Formal: Es gilt für alle  $-\infty \leq a < b \leq \infty$

$$\lim_{n \rightarrow \infty} \Pr \left( a \leq \frac{Z_n - \mathbb{E}Z_n}{\sqrt{\text{Var } Z_n}} \leq b \right) = \Pr(a \leq Z \leq b),$$

wobei  $Z$  eine standardnormalverteilte Zufallsvariable ist.

Anders formuliert: Für große  $n$  gilt

$$Z_n \sim \mathcal{N}(\mu = \mathbb{E}Z_n, \sigma^2 = \text{Var } Z_n)$$

Anders formuliert: Für große  $n$  gilt

$$Z_n \sim \mathcal{N}(\mu = \mathbb{E}Z_n, \sigma^2 = \text{Var } Z_n)$$

Die Voraussetzungen „unabhängig“ und „identisch verteilt“ lassen sich noch deutlich abschwächen.

Anders formuliert: Für große  $n$  gilt

$$Z_n \sim \mathcal{N}(\mu = \mathbb{E}Z_n, \sigma^2 = \text{Var } Z_n)$$

Die Voraussetzungen „unabhängig“ und „identisch verteilt“ lassen sich noch deutlich abschwächen.

Für den Hausgebrauch:

Anders formuliert: Für große  $n$  gilt

$$Z_n \sim \mathcal{N}(\mu = \mathbb{E}Z_n, \sigma^2 = \text{Var } Z_n)$$

Die Voraussetzungen „unabhängig“ und „identisch verteilt“ lassen sich noch deutlich abschwächen.

Für den Hausgebrauch:

Ist  $Y$  das Resultat von vielen kleinen Beiträgen, die großteils unabhängig voneinander sind, so ist  $Y$  in etwa normalverteilt, d.h.

$$Y \sim \mathcal{N}(\mu = \mathbb{E}Y, \sigma^2 = \text{Var } Y)$$

# Inhalt

- 1 Deterministische und zufällige Vorgänge
- 2 Zufallsvariablen und Verteilung
- 3 Die Binomialverteilung
- 4 Erwartungswert
- 5 Varianz und Korrelation
- 6 Ein Anwendungsbeispiel
- 7 Die Normalverteilung
- 8 Normalapproximation
- 9 Der z-Test**



Zurück zu dem Beispiel mit den  
Prolin-Codons im menschlichen Genom.

Zurück zu dem Beispiel mit den  
Prolin-Codons im menschlichen Genom.

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Zurück zu dem Beispiel mit den  
Prolin-Codons im menschlichen Genom.

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Frage: Kann dies Zufall sein?

Zurück zu dem Beispiel mit den  
Prolin-Codons im menschlichen Genom.

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Frage: Kann dies Zufall sein?

Wir meinen: **Nein.**

Zurück zu dem Beispiel mit den Prolin-Codons im menschlichen Genom.

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Frage: Kann dies Zufall sein?

Wir meinen: **Nein.**

Die Skeptiker sagen:

„Nur Zufall.“

## Die Hypothese

Reiner Zufall

**Kein** Unterschied

Die Hypothese

Reiner Zufall

Kein Unterschied

nennt man die

Nullhypothese.

Die Hypothese

Reiner Zufall

Kein Unterschied

nennt man die

Nullhypothese.

Um die Skeptiker zu überzeugen,

müssen wir die

Nullhypothese entkräften



## Die Hypothese

Reiner Zufall  
**Kein** Unterschied

nennt man die  
**Null**hypothese.

Um die Skeptiker zu überzeugen,  
müssen wir die  
**Nullhypothese entkräften**  
d.h. zeigen, dass unter der Nullhypothese  
die Beobachtung sehr unwahrscheinlich ist.

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Unter der Nullhypothese „alles nur Zufall“  
ist die Anzahl  $X$  der CCT  $\text{bin}(n, p)$ -verteilt  
mit  $n = 35605$  und  $p = 0.5$ .

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Unter der Nullhypothese „alles nur Zufall“  
ist die Anzahl  $X$  der CCT  $\text{bin}(n, p)$ -verteilt  
mit  $n = 35605$  und  $p = 0.5$ .

Normalapproximation:  $X$  ist ungefähr  $\mathcal{N}(\mu, \sigma^2)$ -verteilt mit

$$\mu = n \cdot p = 17802.5$$

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Unter der Nullhypothese „alles nur Zufall“  
ist die Anzahl  $X$  der CCT  $\text{bin}(n, p)$ -verteilt  
mit  $n = 35605$  und  $p = 0.5$ .

Normalapproximation:  $X$  ist ungefähr  $\mathcal{N}(\mu, \sigma^2)$ -verteilt mit

$$\mu = n \cdot p = 17802.5 \approx 17800$$

und

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)} = 94.34$$

CCT kommt  $k = 16710$  mal vor  
CCC kommt  $n - k = 18895$  mal vor

Unter der Nullhypothese „alles nur Zufall“  
ist die Anzahl  $X$  der CCT  $\text{bin}(n, p)$ -verteilt  
mit  $n = 35605$  und  $p = 0.5$ .

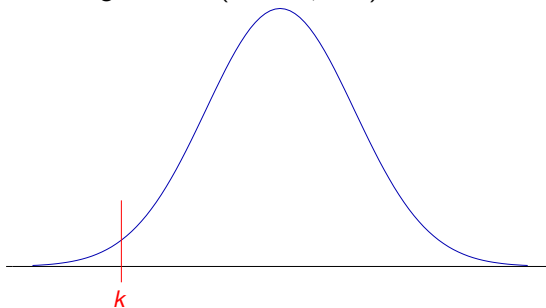
Normalapproximation:  $X$  ist ungefähr  $\mathcal{N}(\mu, \sigma^2)$ -verteilt mit

$$\mu = n \cdot p = 17802.5 \approx 17800$$

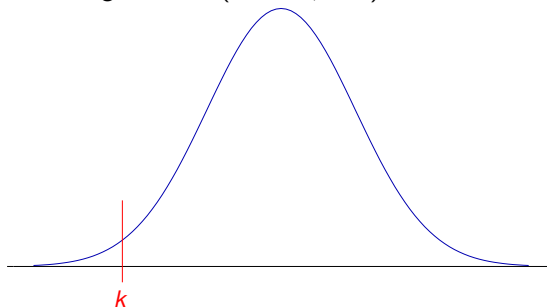
und

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)} = 94.34 \approx 95$$

Frage: Ist es plausibel, dass eine Größe  $X$ , die den Wert  $k = 18895$  angenommen hat, ungefähr  $\mathcal{N}(17800, 95^2)$ -verteilt ist?



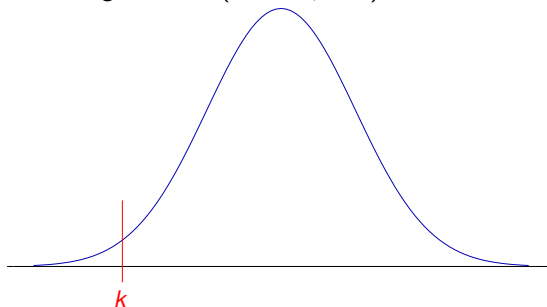
Frage: Ist es plausibel, dass eine Größe  $X$ , die den Wert  $k = 18895$  angenommen hat, ungefähr  $\mathcal{N}(17800, 95^2)$ -verteilt ist?



Wenn diese Nullhypothese  $H_0$  gilt, dann folgt

$$\Pr(X = 17800) =$$

Frage: Ist es plausibel, dass eine Größe  $X$ , die den Wert  $k = 18895$  angenommen hat, ungefähr  $\mathcal{N}(17800, 95^2)$ -verteilt ist?

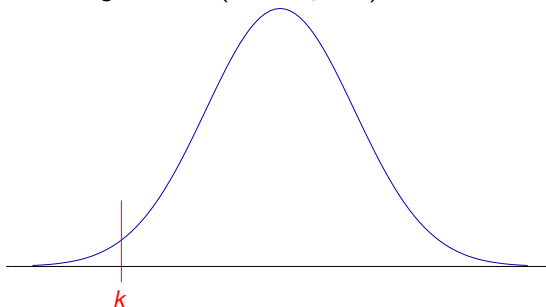


Wenn diese Nullhypothese  $H_0$  gilt, dann folgt

$$\Pr(X = 17800) = 0$$



Frage: Ist es plausibel, dass eine Größe  $X$ , die den Wert  $k = 18895$  angenommen hat, ungefähr  $\mathcal{N}(17800, 95^2)$ -verteilt ist?

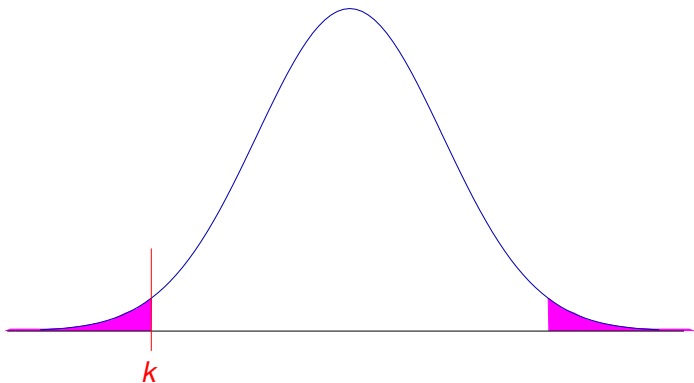


Wenn diese Nullhypothese  $H_0$  gilt, dann folgt

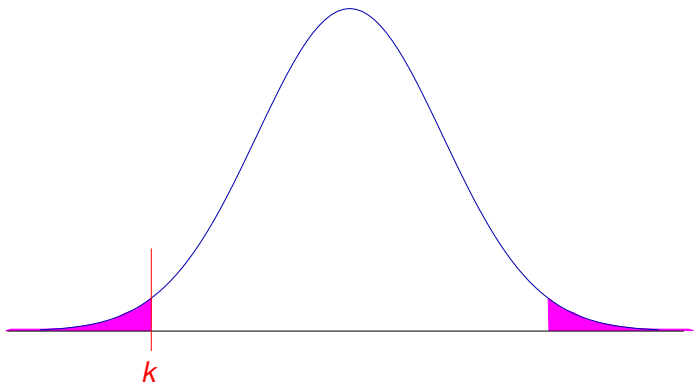
$$\Pr(X = 17800) = 0$$

Aber das bedeutet nichts, denn  $\Pr(X = k) = 0$  gilt für jeden Wert  $k$ !

Entscheidend ist die Wahrscheinlichkeit,  
dass  $X$  (unter Annahme der  $H_0$ ) einen  
mindestens so extremen Wert wie  $k$  annimmt:

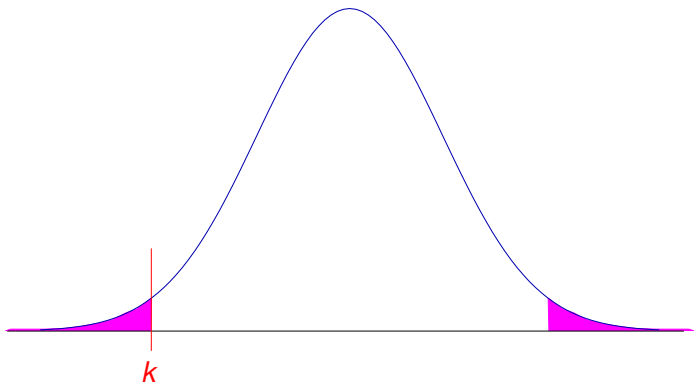


Entscheidend ist die Wahrscheinlichkeit,  
dass  $X$  (unter Annahme der  $H_0$ ) einen  
mindestens so extremen Wert wie  $k$  annimmt:



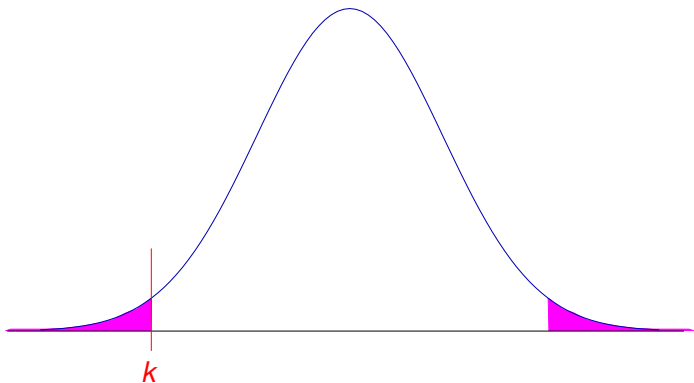
$$\Pr(|X - \mu| \geq |k - \mu|)$$

Entscheidend ist die Wahrscheinlichkeit,  
dass  $X$  (unter Annahme der  $H_0$ ) einen  
mindestens so extremen Wert wie  $k$  annimmt:



$$\Pr(|X - \mu| \geq |k - \mu|) = \Pr(|X - \mu| \geq 1092.5)$$

Entscheidend ist die Wahrscheinlichkeit,  
dass  $X$  (unter Annahme der  $H_0$ ) einen  
mindestens so extremen Wert wie  $k$  annimmt:



$$\Pr(|X - \mu| \geq |k - \mu|) = \Pr(|X - \mu| \geq 1092.5) \approx \Pr(|X - \mu| \geq 11.6 \cdot \sigma)$$

Wir wissen bereits:

$$Pr(|X - \mu| \geq 3 \cdot \sigma) \approx 0.003$$

Wir wissen bereits:

$$Pr(|X - \mu| \geq 3 \cdot \sigma) \approx 0.003 \quad (\text{siehe Merkgeln!})$$

Wir wissen bereits:

$$\Pr(|X - \mu| \geq 3 \cdot \sigma) \approx 0.003 \quad (\text{siehe Merkgeln!})$$

Also muss  $\Pr(|X - \mu| \geq 11.6 \cdot \sigma)$  extrem klein sein.



Wir wissen bereits:

$$\Pr(|X - \mu| \geq 3 \cdot \sigma) \approx 0.003 \quad (\text{siehe Merkgeln!})$$

Also muss  $\Pr(|X - \mu| \geq 11.6 \cdot \sigma)$  extrem klein sein.

In der Tat:

```
> 2 * pnorm(18895,mean=17800,sd=95,lower.tail=FALSE)
[1] 9.721555e-31
```

Wir wissen bereits:

$$\Pr(|X - \mu| \geq 3 \cdot \sigma) \approx 0.003 \quad (\text{siehe Merkgeln!})$$

Also muss  $\Pr(|X - \mu| \geq 11.6 \cdot \sigma)$  extrem klein sein.

In der Tat:

```
> 2 * pnorm(18895,mean=17800,sd=95,lower.tail=FALSE)
[1] 9.721555e-31
```

Ohne Normalapproximation:

```
> pbinom(16710,size=35605,p=0.5) +
+ pbinom(18895-1,size=35605,p=0.5,lower.tail=FALSE)
[1] 5.329252e-31
```

Wir können also argumentieren,  
dass eine derartig starke Abweichung vom Erwartungswert  
nur durch einen extremen Zufall zu erklären ist.

Wir werden also die  
Nullhypothese “alles nur Zufall” verwerfen  
und nach alternativen Erklärungen suchen,  
etwa unterschiedliche Effizienz von CCC und CCT  
oder unterschiedliche Verfügbarkeit von C und T.

# Zusammenfassung z-Test

**Nullhypothese  $H_0$**  (möchte man meistens verwerfen): der beobachtete Wert  $x$  kommt aus einer Normalverteilung mit Mittelwert  $\mu$  und **bekannter** Varianz  $\sigma^2$ .

# Zusammenfassung z-Test

**Nullhypothese  $H_0$**  (möchte man meistens verwerfen): der beobachtete Wert  $x$  kommt aus einer Normalverteilung mit Mittelwert  $\mu$  und **bekannter** Varianz  $\sigma^2$ .

**p-Wert**  $= \Pr(|X - \mu| \geq |x - \mu|)$ , wobei  $X \sim \mathcal{N}(\mu, \sigma^2)$ , also die Wahrscheinlichkeit einer *mindestens* so großen Abweichung wie der beobachteten.

# Zusammenfassung z-Test

**Nullhypothese  $H_0$**  (möchte man meistens verwerfen): der beobachtete Wert  $x$  kommt aus einer Normalverteilung mit Mittelwert  $\mu$  und **bekannter** Varianz  $\sigma^2$ .

**$p$ -Wert**  $= \Pr(|X - \mu| \geq |x - \mu|)$ , wobei  $X \sim \mathcal{N}(\mu, \sigma^2)$ , also die Wahrscheinlichkeit einer *mindestens* so großen Abweichung wie der beobachteten.

**Signifikanzniveau  $\alpha$**  : oft 0.05. Wenn der  $p$ -Wert kleiner ist als  $\alpha$ , verwerfen wir die Nullhypothese auf dem Signifikanzniveau  $\alpha$  und suchen nach einer alternativen Erklärung.

# Grenzen des z-Tests

Der z-Test kann nur angewendet werden, wenn die Varianz der Normalverteilung bekannt ist oder zumindest in der Nullhypothese als bekannt angenommen wird.

# Grenzen des z-Tests

Der z-Test kann nur angewendet werden, wenn die Varianz der Normalverteilung bekannt ist oder zumindest in der Nullhypothese als bekannt angenommen wird.

Das ist meistens nicht der Fall, wenn die Normalverteilung beim statistischen Testen verwendet wird.



# Grenzen des z-Tests

Der z-Test kann nur angewendet werden, wenn die Varianz der Normalverteilung bekannt ist oder zumindest in der Nullhypothese als bekannt angenommen wird.

Das ist meistens nicht der Fall, wenn die Normalverteilung beim statistischen Testen verwendet wird.

Meistens wird die Varianz aus den Daten geschätzt. Dann muss statt dem z-Test der berühmte

**t-Test**

angewendet werden.