

# Warum Statistik?

Statistik 1 für Biol./Pharm. Wiss./HST – Herbst 2014



# Bei Risiken und Nebenwirkungen...

Liste Patienteninformation

Patienteninformation

## Dafalgan® 500 mg Tabletten

BRISTOL-MYERS SQUIBB

**AMZV**

**Was ist Dafalgan und wann wird es angewendet?**  
Dafalgan Tabletten enthalten den Wirkstoff Paracetamol, der schmerzlindernd und fiebersenkend wirkt.  
Dafalgan Tabletten werden angewendet zur kurzfristigen Behandlung von Kopfschmerzen, Zahnschmerzen, Schmerzen im Bereich von Gelenken und Bändern, Rückenschmerzen, Schmerzen während der Monatsblutung, Schmerzen nach Verletzungen (z.B. Sportverletzungen), Schmerzen bei Erkältungskrankheiten sowie gegen Fieber.

- 
- 
- 

### **Welche Nebenwirkungen kann Dafalgan haben?**

Folgende Nebenwirkungen können bei der Einnahme von Dafalgan auftreten:

In **selteneren Fällen** können Überempfindlichkeitsreaktionen wie Schwellungen der Haut und Schleimhäute oder Hautausschläge, Übelkeit, Atemnot oder Asthma auftreten. Treten Zeichen einer Überempfindlichkeit auf, so ist das Arzneimittel abzusetzen und der Arzt bzw. die Ärztin zu konsultieren.

Wenn Sie Nebenwirkungen bemerken, die hier nicht beschrieben sind, sollten Sie Ihren Arzt, Apotheker oder Drogisten bzw. Ihre Ärztin, Apothekerin oder Drogistin informieren.

# Häufigkeit von Nebenwirkungen

## Bundesinstitut für Arzneimittel und Medizinprodukte

Sehr häufig ( $\geq 1/10$ )

Häufig ( $\geq 1/100$  bis  $< 1/10$ )

Gelegentlich ( $\geq 1/1.000$  bis  $< 1/100$ )

Selten ( $\geq 1/10.000$  bis  $< 1/1.000$ )

Sehr selten ( $< 1/10.000$ )

Nicht bekannt (Häufigkeit auf Grundlage der verfügbaren Daten nicht abschätzbar)

Wa. für Nebenwirkung:

0.0001 bis 0.001

Klinische Studie: 10.000 Patienten, 7 Fälle von Nebenwirkungen

Ist es sicher, dass Wa. für Nebenwirkungen  $< 0.001$  ?

**Binomialtest:** Nein! 95% Vertrauensintervall für Wa. ist [0.00028; 0.0014]



“Gelegentlich”

# Pipettieren kann jeder !



Pipette auf 200 $\mu$ L eingestellt – wie genau ist das?

Gib einen Bereich an, in dem der  
Mittelwert von sehr vielen Messungen  
mit 95% Wa. liegt

Messung mit Wasser (Dichte 0.99807 g/mL):

$m_1=199.26$  mg,  $m_2=199.40$  mg, ...,  $m_{10}=199.29$  mg

**t-Test:** 95% Vertrauensintervall für wahre Einstellung  
[199.5  $\mu$ L; 199.7  $\mu$ L]

# Wie fit sind Sie?

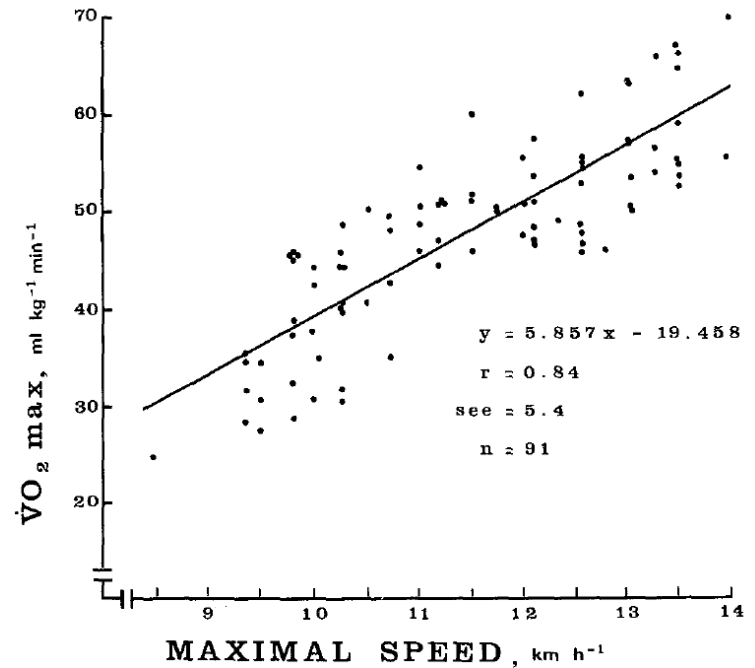
## Achtung, fertig, ..., Beep!

### A Maximal Multistage 20-m Shuttle Run Test to Predict $\dot{V}O_2 \text{ max}^*$

Luc A. Léger<sup>1</sup> and J. Lambert<sup>2</sup>

<sup>1</sup>Département d'éducation physique, Université de Montréal, CEPSUM, C.P. 6128, Succ. "A", Montréal (Québec), Canada, H3C 3J7

<sup>2</sup>Département de Médecine sociale et préventive, Université de Montréal, Canada



Lineare Regression:  $y = 5.857x - 19.458$

# Lernziele

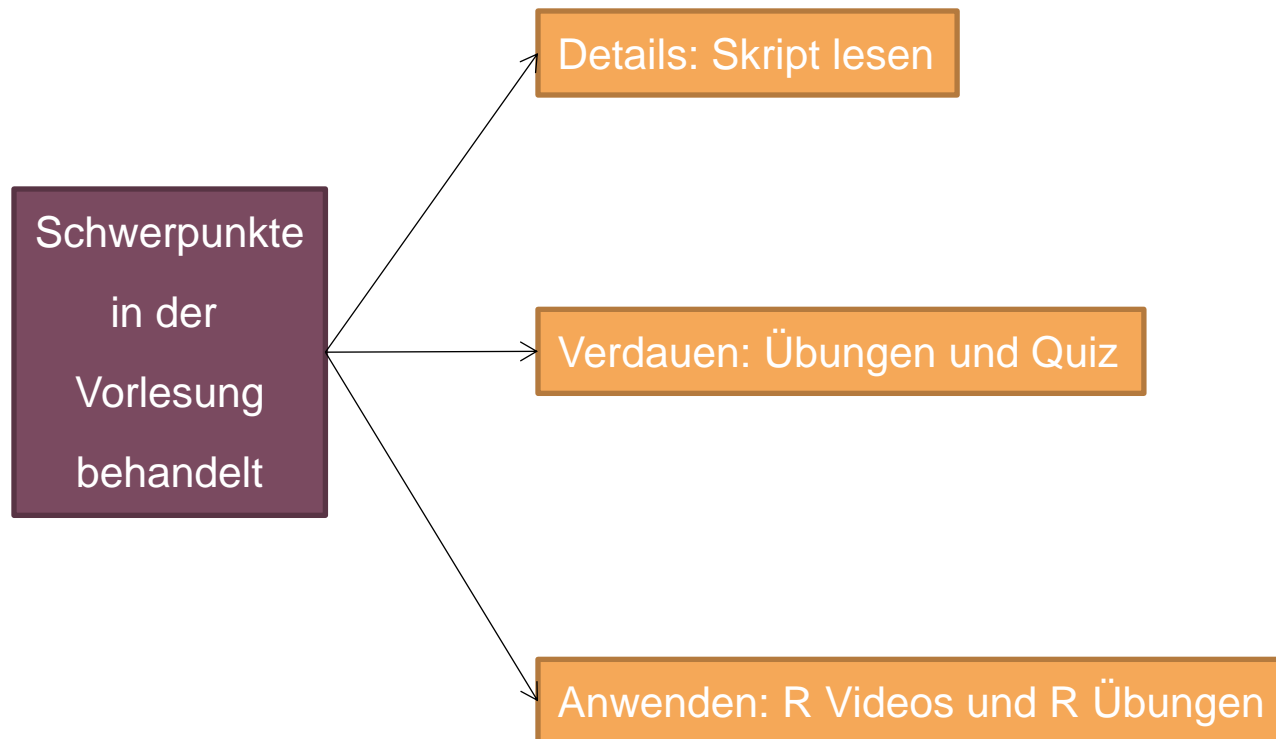
- Grundkonzepte der Wahrscheinlichkeit und Statistik
- Statistische Software “R” kennenlernen und anwenden
- Mit Ungenauigkeit umgehen:
  - Punktschätzung
  - Vertrauensintervall
  - Hypothesentest
- Konkret an den Beispielen:
  - Binomialtest
  - t-Test
  - Lineare Regression
- Fundament für Statistik 2 (Herbstsemester)

# Stellenwert im Studium

- Phys. – chem. Praktikum im 4. Semester
- In vielen Fachvorlesungen  
(Ökologie, Systembiologie, Genomics, Versuche, ...)
- Masterarbeit  
Es gibt auch einen statistischen Beratungsdienst der ETH,  
den Sie in Anspruch nehmen können:  
[beratung@stat.math.ethz.ch](mailto:beratung@stat.math.ethz.ch)
- Wissenschaftliche Paper: Sehr oft wird eine statistische  
Auswertung gemacht

# Aufbau der Vorlesung

- Wir folgen dem Skript (siehe dort Lernziele pro Kapitel)





# etutoR

- Vorlesung: Konzepte verstehen
- Praxis: Computer verwenden
- Eine der besten Programmiersprachen für Statistik ist 'R'
- etutoR: Begleitend zur Vorlesung werden wir Selbstlernvideos einsetzen, die Ihnen zeigen, wie man 'R' verwendet.
  
- Prüfung Statistik 1: R Output verstehen
- (Prüfung Statistik 2: R anwenden – in Planung)

# EduApp



Stundenplan überblicken

Hörsäle finden

Lehrveranstaltungen interaktiv gestalten

- Aktive Mitarbeit: Quiz und kurze Diskussion
- Jederzeit anonymes Feedback an
  - mich
  - Semestersprecher
- Bitte bringen Sie dafür ihr Smartphone oder Laptop in die VL mit und machen Sie sich vorgängig mit EduApp vertraut
- <http://www.eduapp.ethz.ch/>

# Administration 1/2

- Homepage <http://stat.ethz.ch/education/semesters/ss2014/bio>
- Skript: Auf Homepage
  
- Übungen: Auf Homepage;  
Aufgaben werden auf Wunsch korrigiert;  
geben Sie an, welche (sonst wird keine korrigiert)  
kein Testat
- Quiz: Ca. einmal pro Woche per mail
  
- Prüfung
  - 90 Minuten
  - Open-book
  - 4 Aufgaben  
(Binomialtest, t-Test, Regression, Vermischtes)
  - PVK: Von Studenten organisiert, nicht Teil der VL; Stoff gemäss VL

# Administration 2/2

- Fragen:
  - Fragen Sie in der Vorlesung
  - “Office hour” im Anschluss an die VL im Hörsaal
  - Semesterferien: 1-2 Fragestunden bei Assistenten
- Semstersprecher pro Studiengang? Bitte in Pause zu mir.
- Feedback der vergangenen Jahre zeigt: Lärm in der VL stört enorm.

**Helfen Sie bitte mit,  
damit alle  
gut zuhören  
können !**



# Hypothesentest: Ein Beispiel





661 Bilder

Packung

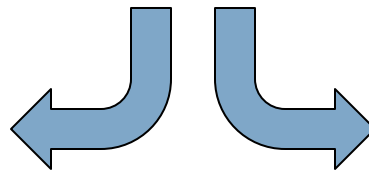


5 Bilder

Box



100 Packungen =  
500 Bilder





# Beobachtung von Vorjahren

- Ganze Box: Wenige doppelte Bilder
- Einzelne Packungen an verschiedenen Kiosks:  
Viele doppelte Bilder
- “Null”hypothese: Bilder werden zufällig verpackt  
(“Null”, weil kein System hinter dem Verpacken steckt)
- Alternativhypothese: Die Bilder werden systematisch verpackt, sodass man weniger doppelte hat
- Wie könnte man zwischen diesen beiden Hypothesen unterscheiden?

# Hypothesentest

- Ich habe eine Box mit 500 Bildern gekauft. In eine leeres Album (661 mögliche Bilder) konnte ich 477 Bilder einkleben.
- Angenommen, die Nullhypothese stimmt:  
Ist es plausibel, dass ich dann 477 Bilder einkleben kann?
- Passen die Nullhypothese “zufällig verpackt” und die Beobachtung “477 Bilder eingeklebt” zusammen?

# Problem: Was ist “normal”?

- Wenn wir viel mehr Bilder als “normal” einkleben konnten, wurden die Bilder wohl nicht zufällig verpackt.
- Angenommen, die Nullhypothese stimmt (Bilder zufällig verpackt):

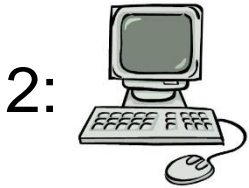
## Wie viele Bilder kann man normalerweise einkleben?

- Signifikanzniveau  $\alpha$ : Wie “abnormal” muss die Beobachtung sein, damit wir der Nullhypothese nicht mehr glauben?  
Z.B.:  $\alpha = 1/1.000.000$  ; wir lehnen die Nullhypothese ab, wenn wir etwas beobachten, das weniger wahrscheinlich als  $1/1.000.000$  ist.

# Lösung: Computersimulation



350 Bilder eingeklebt



361 Bilder eingeklebt

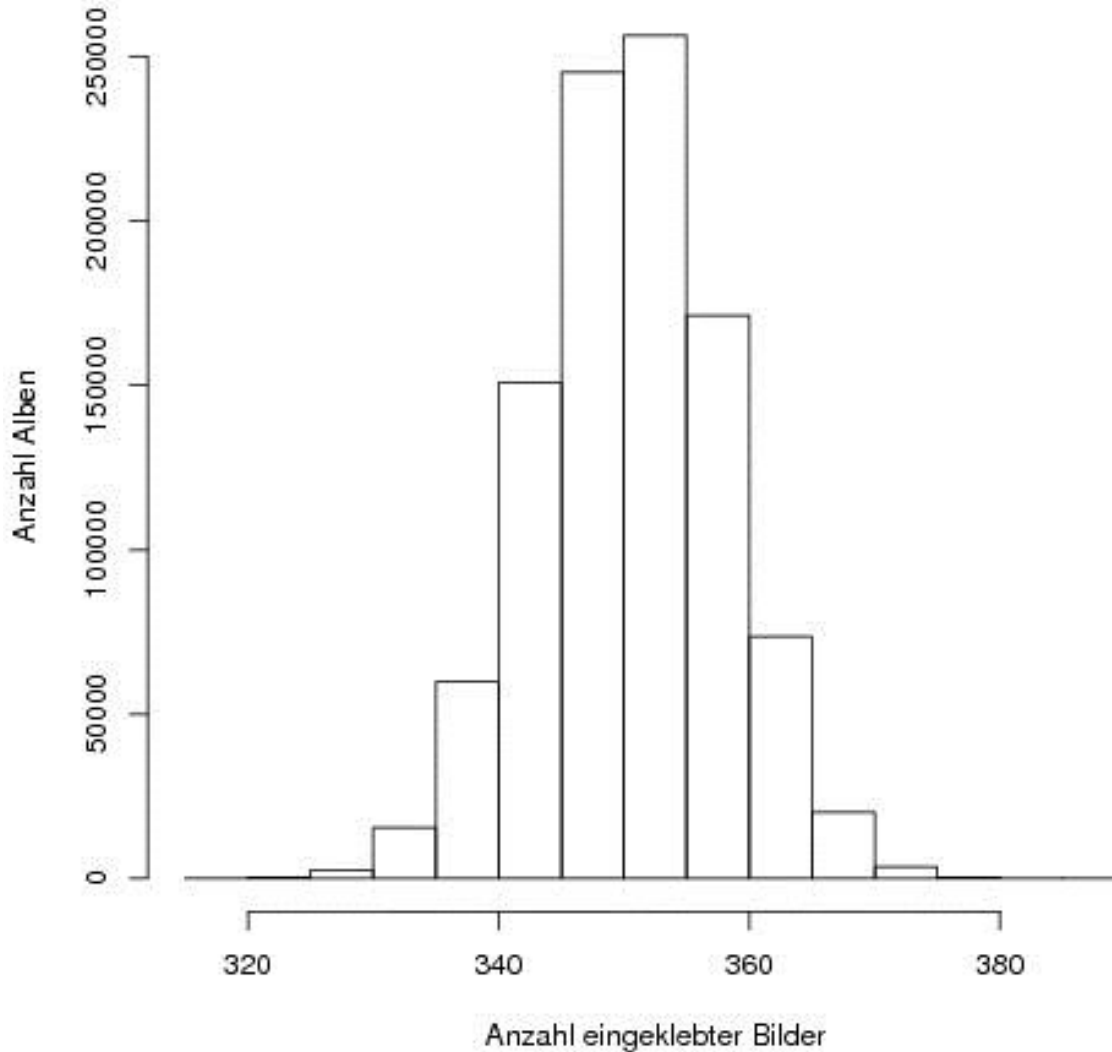
⋮



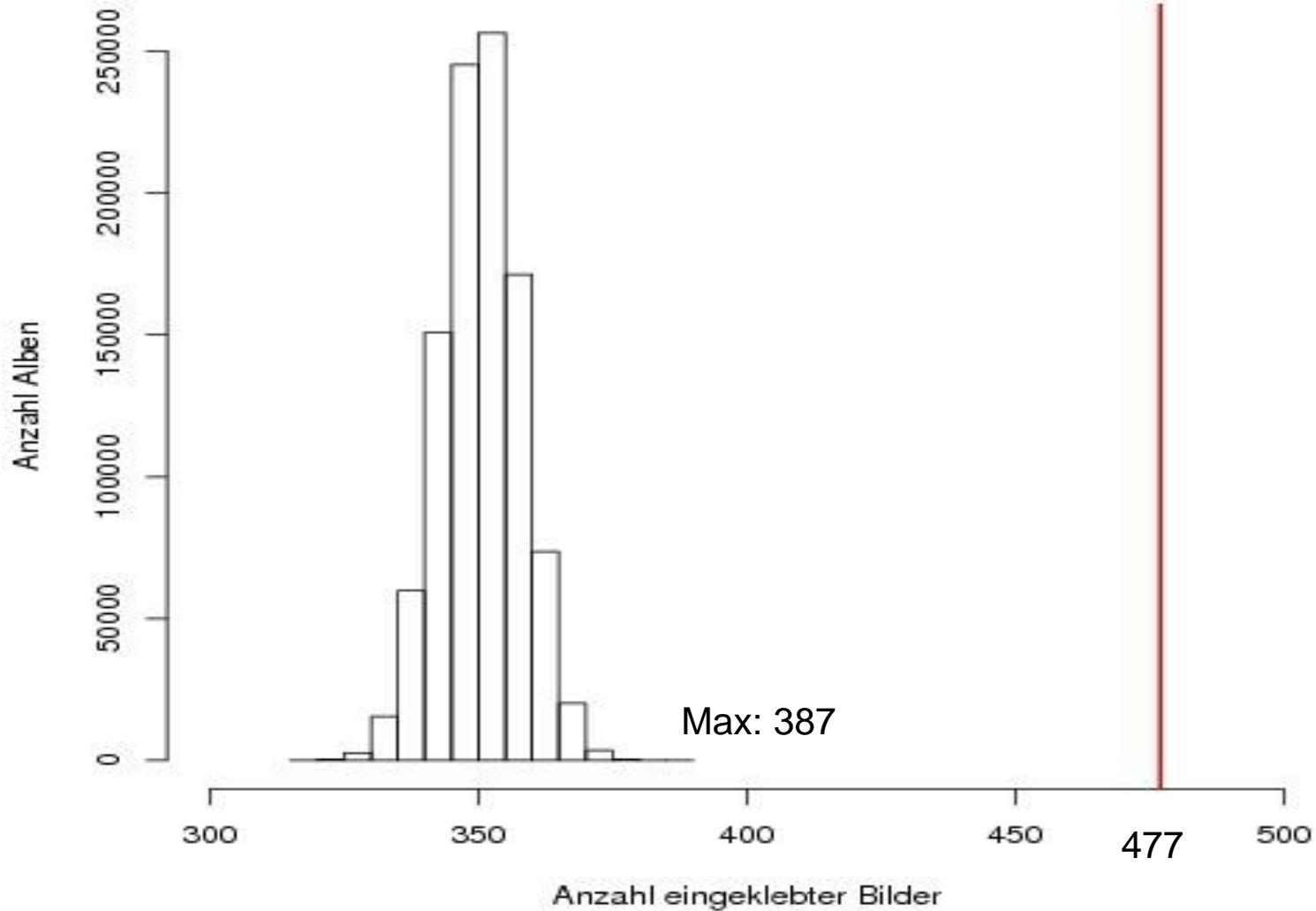
358 Bilder eingeklebt

# Ergebnis der Computersimulation

Computersimulation: Einkleben von Panini-Bildern

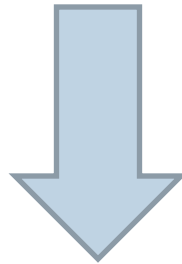


# Passt unsere Beobachtung zur Computersimulation?



# Schlussfolgerung

- Angenommen, die Bilder werden zufällig verpackt. Die Wa. 477 oder mehr Bilder einkleben zu können ist kleiner als ein Millionstel !



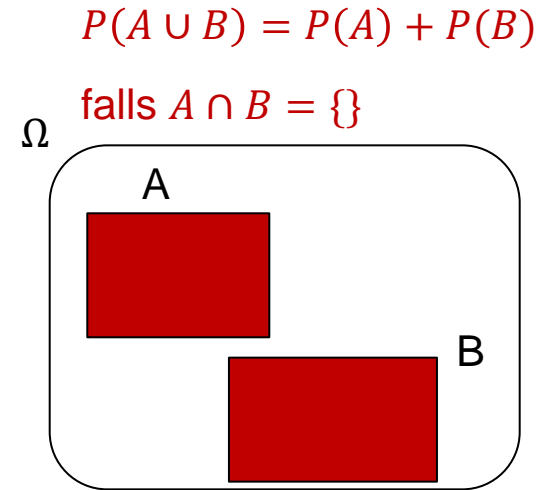
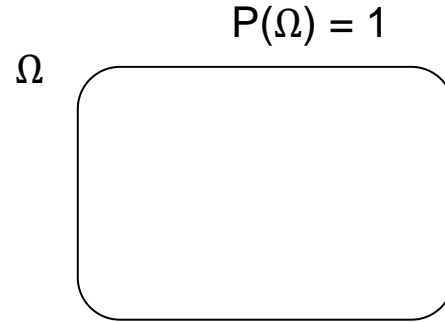
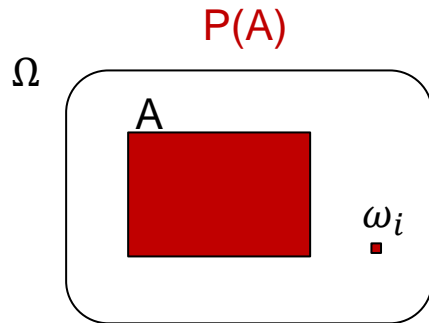
- Beobachtung und Simulation passen nicht zusammen:  
Die Bilder werden wohl NICHT zufällig eingepackt.

# Zusammenfassung: Hypothesentest

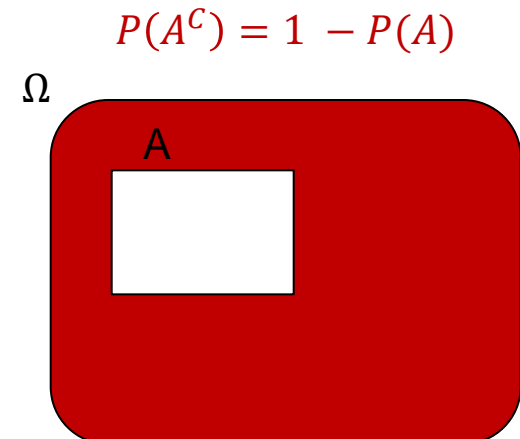
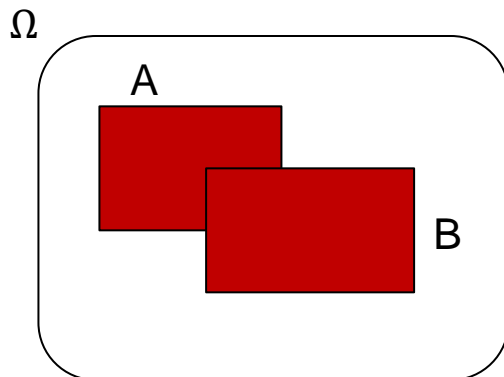
1. Modell: Ziehen 500 Bilder mit Zurücklegen aus 661 Bildern
2. Nullhypothese: "Panini-Bilder in Kiste zufällig eingepackt"  
Alternative: "Systematisch eingepackt, sodass weniger Doppelte"
3. Teststatistik: Anz. Bilder, die man in eine leeres Album einkleben kann, wenn man eine Kiste mit 500 Bildern hat  
Verteilung der Teststatistik, wenn Nullhypothese stimmt:  
Computersimulation
4. Signifikanzniveau  $\alpha = 1/1.000.000$
5. Verwerfungsbereich der Teststatistik:  
Computer beobachtet bei 1 Mio Simulationen nie mehr als 387 eingeklebte Bilder  
Verwerfungsbereich:  $K = \{388, 389, \dots, 500\}$
6. Testentscheid: Der beobachtete Wert (477) liegt im Verwerfungsbereich der Teststatistik. Daher wird die Nullhypothese auf dem Signifikanzniveau  $1/1.000.000$  verworfen.



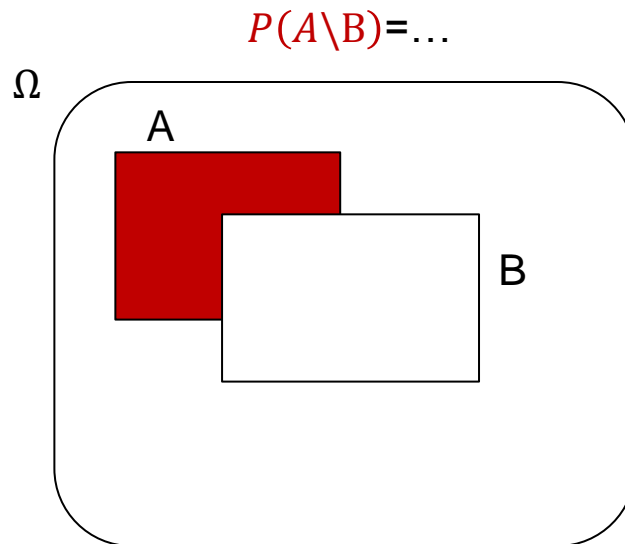
# Visualisierung mit Venn Diagrammen: Wa. entspricht Fläche



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



# Knobelaufgabe



“A ohne B”

1.  $P(A) - P(B)$
2.  $P(A) + P(B)$
3.  $P(A) - P(A \cap B)$
4.  $P(A) + P(B) - P(A \cap B)$