Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

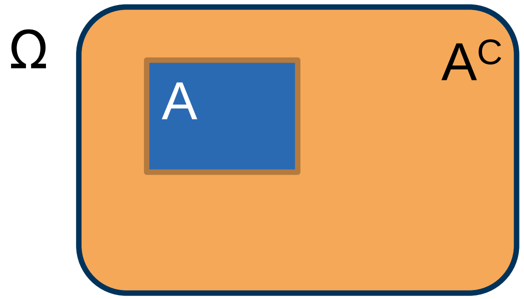
DMATH
Department of Mathematics

Mathematik IV: Statistik

für D-UWIS, D-ERDW, D-USYS und D-HEST – SS15

Repetition: W'keitsmodell

- Grundraum Ω: Alle möglichen Elementarereignisse ω_i
  - Bsp. Würfel: $\Omega=\{1, 2, 3, 4, 5, 6\}$, $\omega_4=4$
- Ereignisse A, B, C, …: Teilmenge des Grundraums, $A \subset \Omega$
  - Bsp. Würfel: $A=\{2, 4, 6\}$ "*gerade Zahlen*"
- Wahrscheinlichkeit P: Axiome von Kolmogorov
  1. $0 \leq P[A] \leq 1$
  2. $P[\Omega] = 1$
  3. $P[A\cup B] = P[A] + P[B]$, falls $A\cap B=\emptyset$
- W'keit berechnen
  - Summe der Elementarereignisse
  - Günstige/mögliche Fälle
  - Mengenoperationen / Venn Diagramm



$P[A^C] = P[\Omega] - P[A] = 1 - P[A]$


Magie

Geburtstagsparadox

Ein **Paradox(on)** (auch **Paradoxie**, Plural **Paradoxien** oder **Paradoxa**; von altgriechisch παράδοξον, von παρά *para* ‚neben', ‚außer', ‚daran vorbei' und δόξα *doxa*‚Meinung', ‚Ansicht') ist eine Aussage, die scheinbar[1] oder tatsächlich einen unauflösbaren Widerspruch enthält.

- Annahme: *Alle Geburtstage sind gleich wahrscheinlich*
- Wie gross ist die W'keit, dass in einer Gruppe von n Personen **mind.** 2 Personen am gleichen Tag Geburtstag haben?

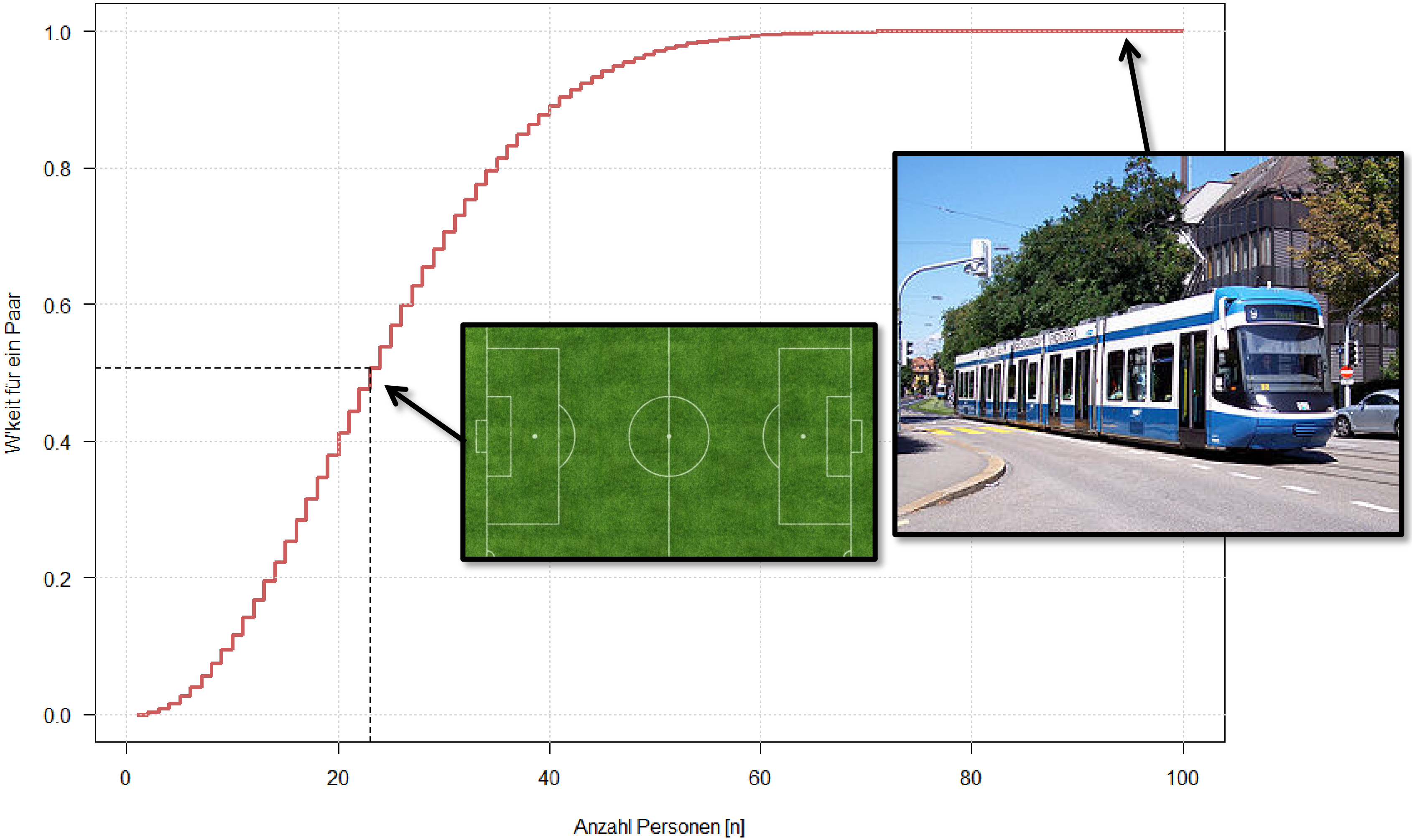
➢ Ereignis G:
"Mind. 2 Personen sind am gleichen Tag geboren"

$$P[G^C] = \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

- Bsp.: $n = 40$:

$$P[G] = 1 - P[G^C] = \frac{365 \times 364 \times \cdots \times 326}{365^n} \approx 0.89$$

Geburtstagsparadox



Abgezockt – Wollen Sie Informationen kaufen?

- Würfelspiel mit 1 Würfel:
  - gerade AZ: Sie gewinnen 10 Fr.
  - ungerade AZ: Sie verlieren 10 Fr.

- Bei jeder Runde, nach dem Wurf, aber vor dem Aufdecken dürfen Sie 50 Rp. zahlen um zu…
  - erfahren, ob die AZ ≤ 3 ist und…
  - entscheiden, ob Sie diese Runde mitspielen wollen.

- Lohnt es sich für diese Information zu zahlen?
  - Enthält das Ereignis «AZ ≤ 3» Informationen über das Ereignis «gerade AZ»?

Lernziele heute

- Unabhängigkeit: $P[A \cap \mathrm{B}] = P[A] \cdot P[B]$
- Bedingte Wahrscheinlichkeit: $P[A|B]$
- odds: $^{P[A]}/_{P[A^C]}$

Hausaufgaben

- Skript: Kapitel 2.2 – 2.3 lessen
- Serie 2 lösen
- Quiz 2 bearbeiten
- etutoR 1 anschauen

2.2 Unabhängigkeit

- A und B sind unabhängig, wenn das Auftreten von A die W'keit von B nicht beeinflusst

$$A \text{ und } B \text{ sind unabhängig} \Leftrightarrow P[A \cap B] = P[A] \cdot P[B]$$

- Beispiel: **Fairer Würfel**
  - Ereignis A: gerade Augenzahl (AZ), Ereignis B: $AZ \leq 3$
  - $\Omega = \{1, 2, 3, 4, 5, 6\}; A = \{2, 4, 6\}; B = \{1, 2, 3\}$
  - $A \cap B = \{2\}$

$\Rightarrow$ Laplace Modell: $P[A] = \frac{|A|}{|\Omega|} = \frac{3}{6} = \frac{1}{2} = 50\%$

  - $P[A \cap B] = \frac{1}{6}; P[A] \cdot P[B] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$\Rightarrow P[A] \cdot P[B] \neq P[A \cap B]$ und somit A und B abhängig!

Abgezockt!
0.67 Fr./Spiel

2.2 Unabhängigkeit

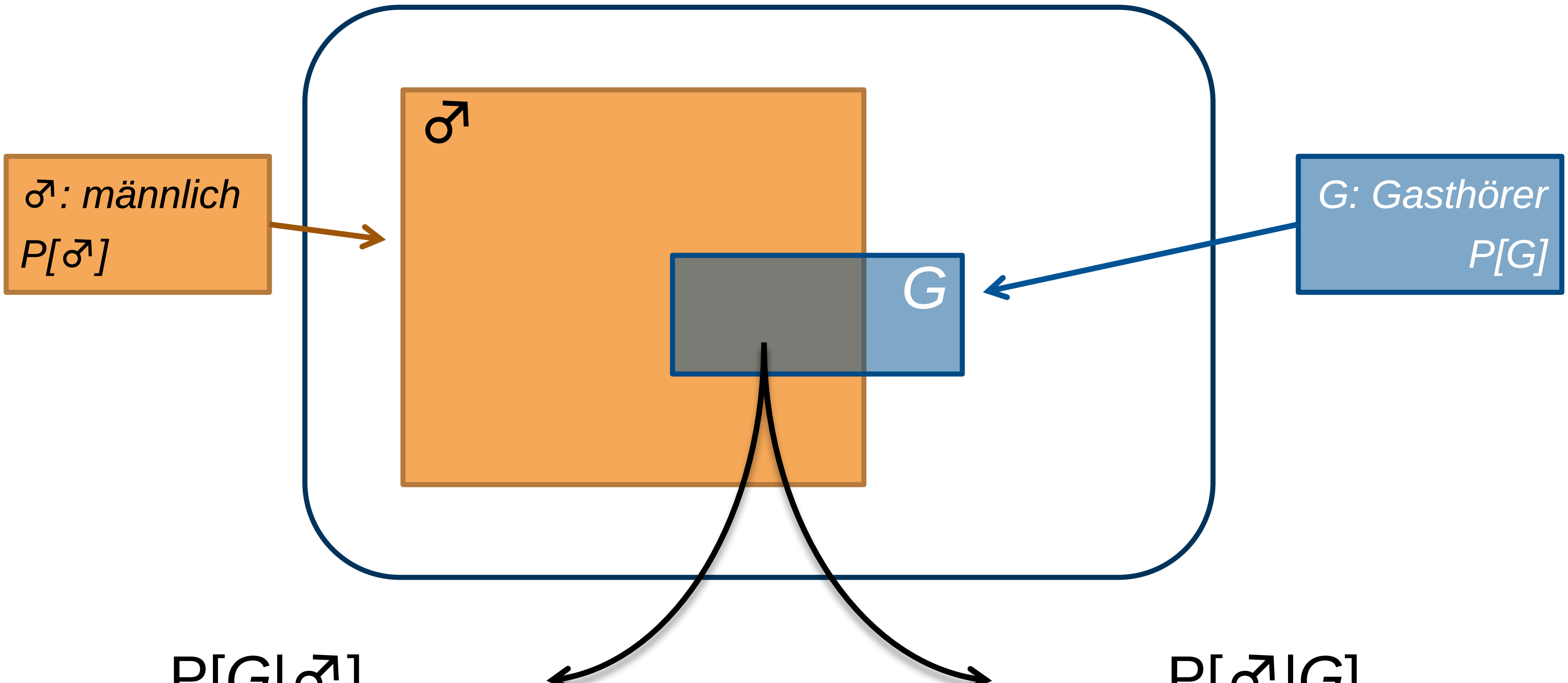
W'keit benutzt um Unabhängigkeit zu prüfen

- Beispiel: **Fairen Münzwurf wiederholen**
  - Werfe Münze einmal: $\Omega = \{K, Z\}, P[K] = P[Z] = \frac{1}{2}$
  - Werfe Münze zweimal: $\Omega = \{KK, ZK, KZ, ZZ\}$
    $P[KK] = P[ZK] = P[KZ] = P[ZZ] = \frac{1}{4}$
  - …
  - Werfe Münze k-mal: $P[ZZZ \dots Z] = \left(\frac{1}{2}\right)^k$

Unabhängigkeit benutzt um W'keit einfacher zu berechnen

Bedingte Wahrscheinlichkeit

Ω: Studenten in dieser Vorlesung



P[*G*|♂]

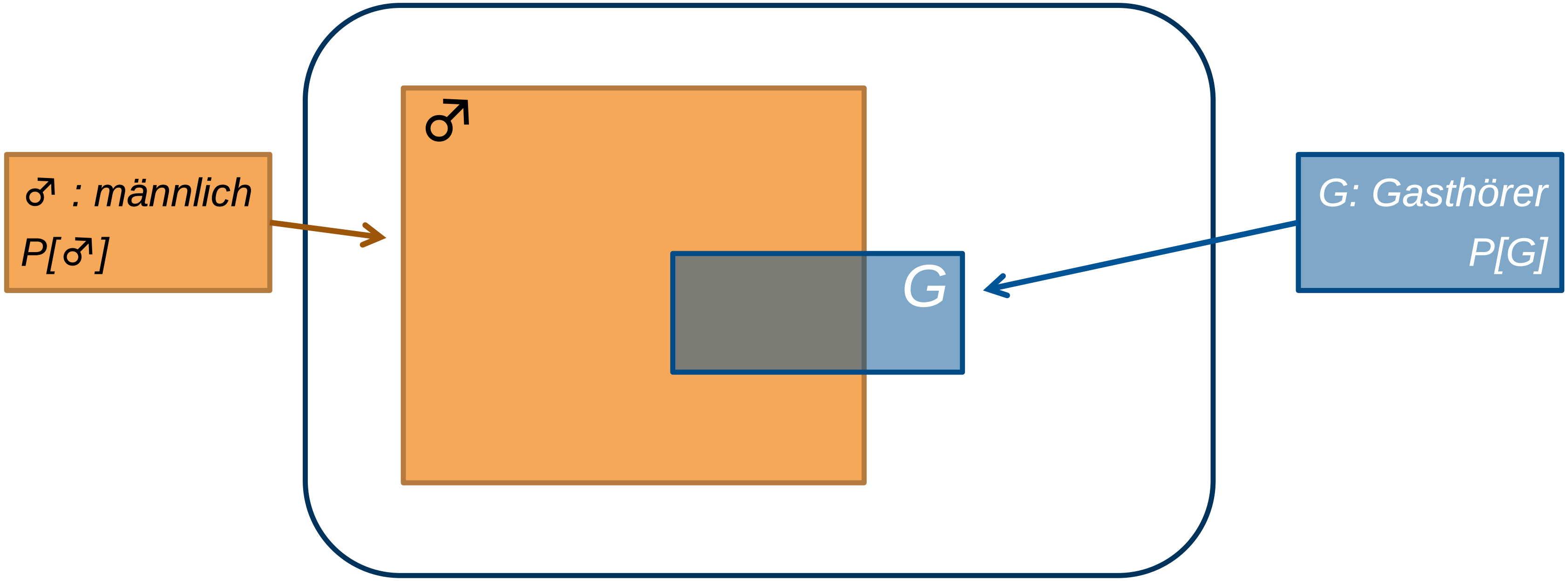
W'keit für einen Gasthörer,
wenn ein Mann gewählt wurde

P[♂|*G*]

W'keit für einen Mann,
wenn ein Gasthörer gewählt wurde

Bedingte Wahrscheinlichkeit

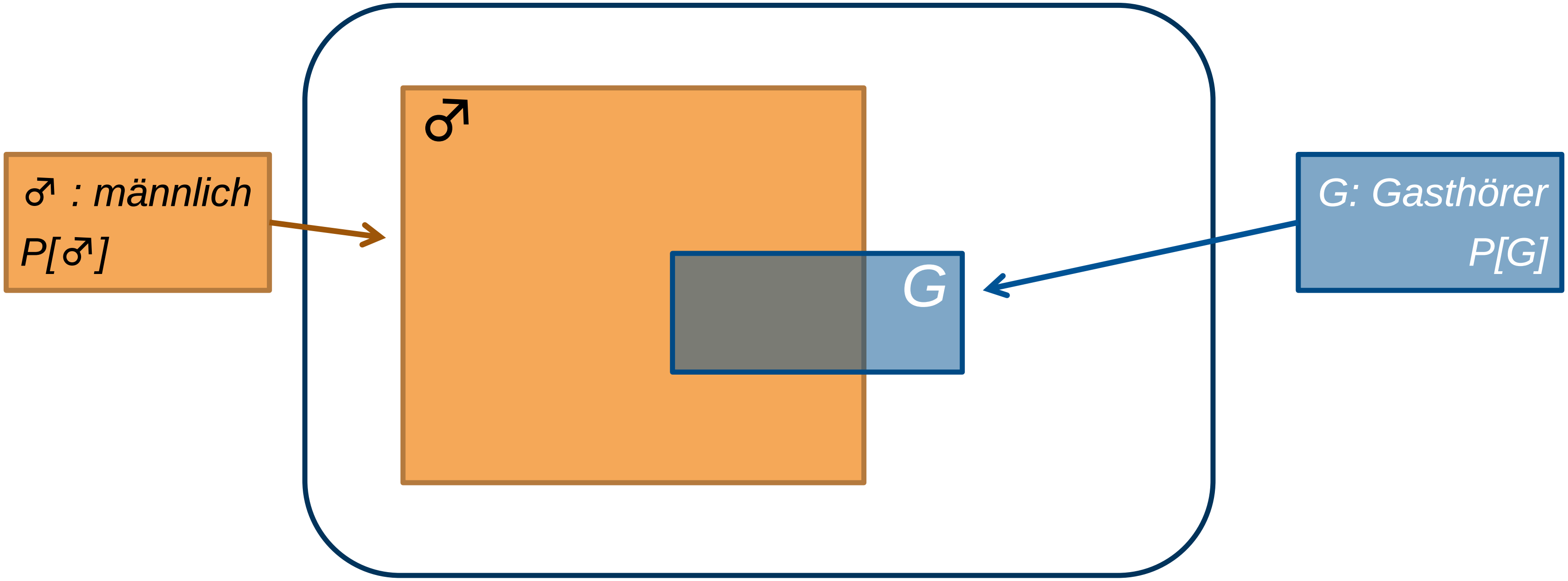
Ω: Studenten in dieser Vorlesung

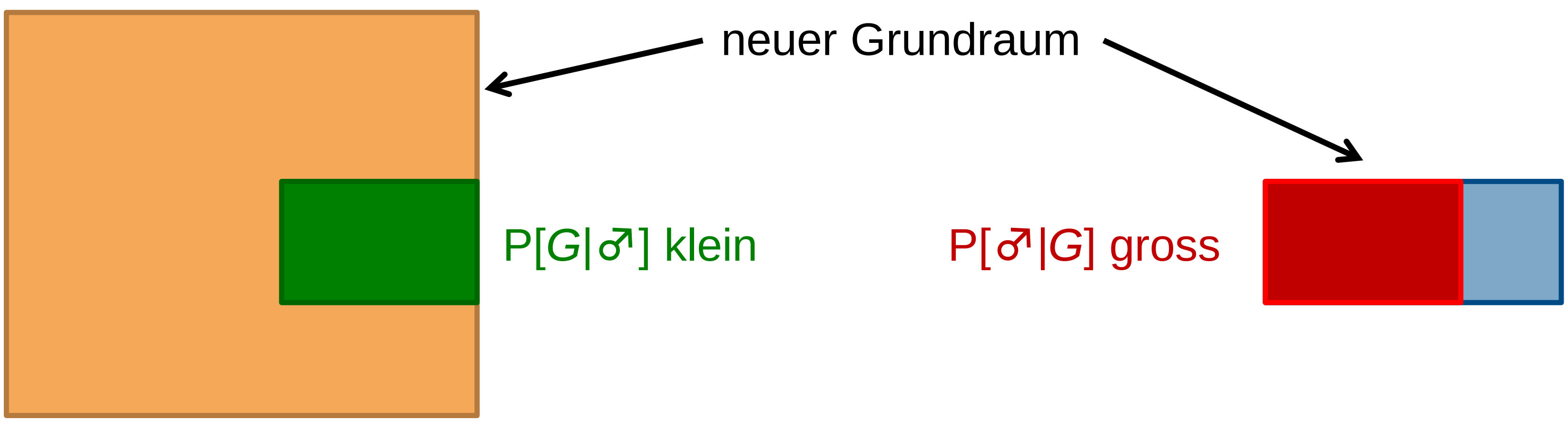


Welche Aussage ist richtig?

1. P[♂|*G*] = P[*G*|♂]
2. P[♂|*G*] > P[*G*|♂]
3. P[♂|*G*] < P[*G*|♂]

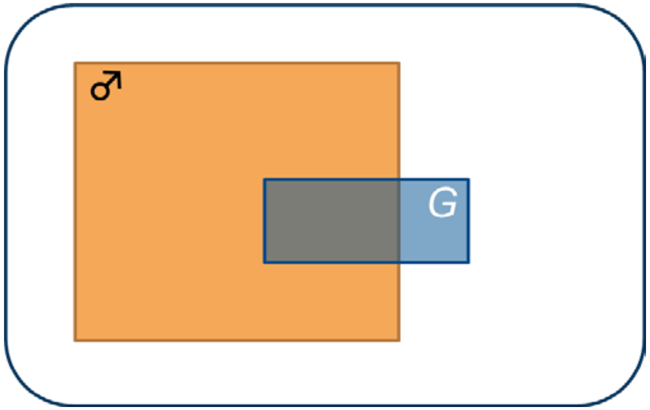
Ω: Studenten in dieser Vorlesung





2.3 Bedingte Wahrscheinlichkeit

- Was ist die W'keit für einen Gasthörer G, wenn wir uns nur die ♂ anschauen?
  - W'keit G gegeben ♂
  - $P[G|♂] := \frac{P[G \cap ♂]}{P[♂]}$



ACHTUNG: $P[A|B] \neq P[B|A]$ im Allgemeinen!

Brustkrebs

- Die W'keit, dass eine 40-jährige Frau Brustkrebs hat ist 1%

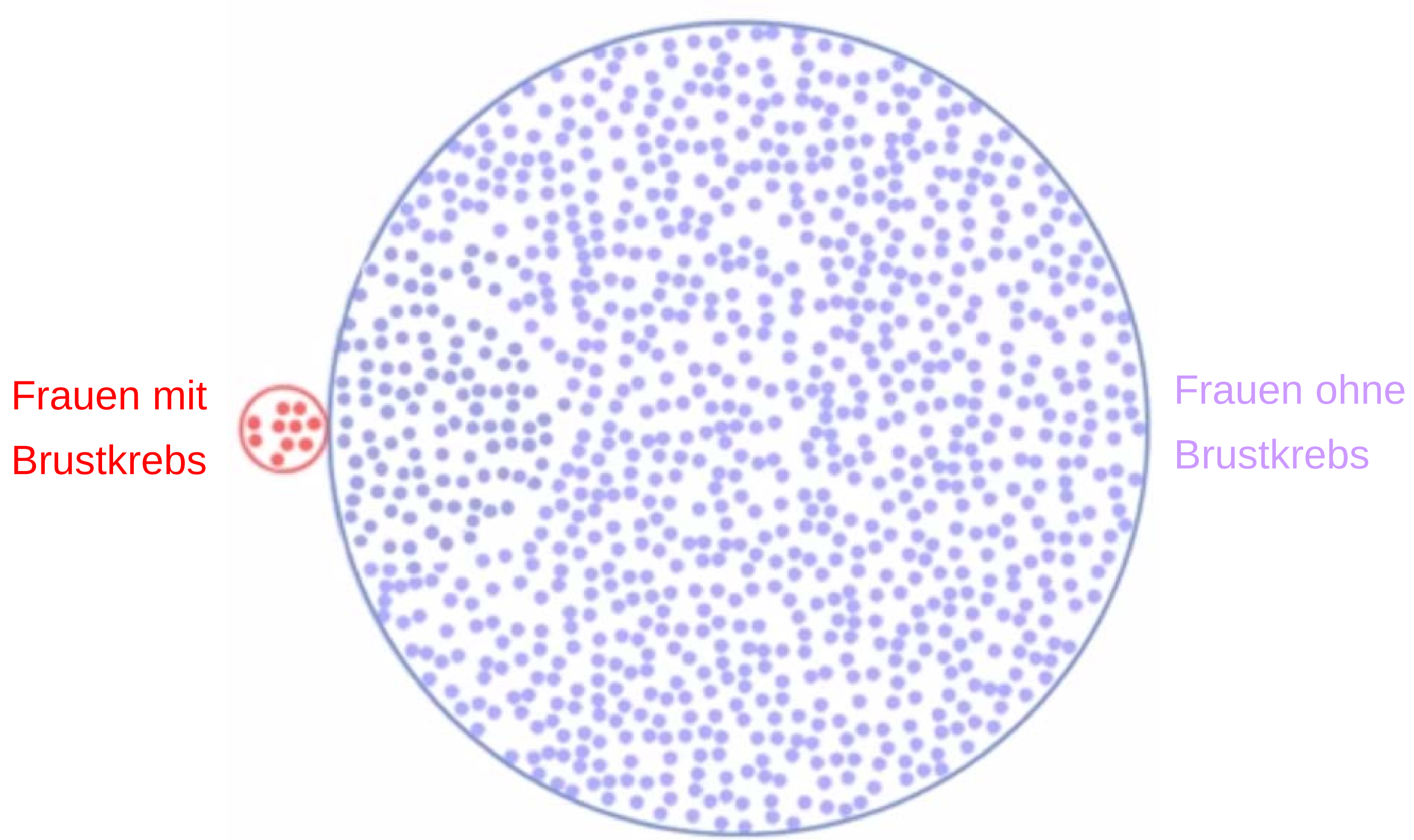
- Mammographie *positiv*, falls krank: 80%
- Mammographie *negativ*, falls gesund: 90%
  - i.e. Mammographie *positiv*, falls gesund: 10%

- Wenn Sie (40, ♀) nun eine Mammographie durchführen und das Testresultat ist *positiv*. Wie wahrscheinlich ist es, dass Sie Brustkrebs haben?

- https://www.youtube.com/watch?v=D8VZqxcu0I0

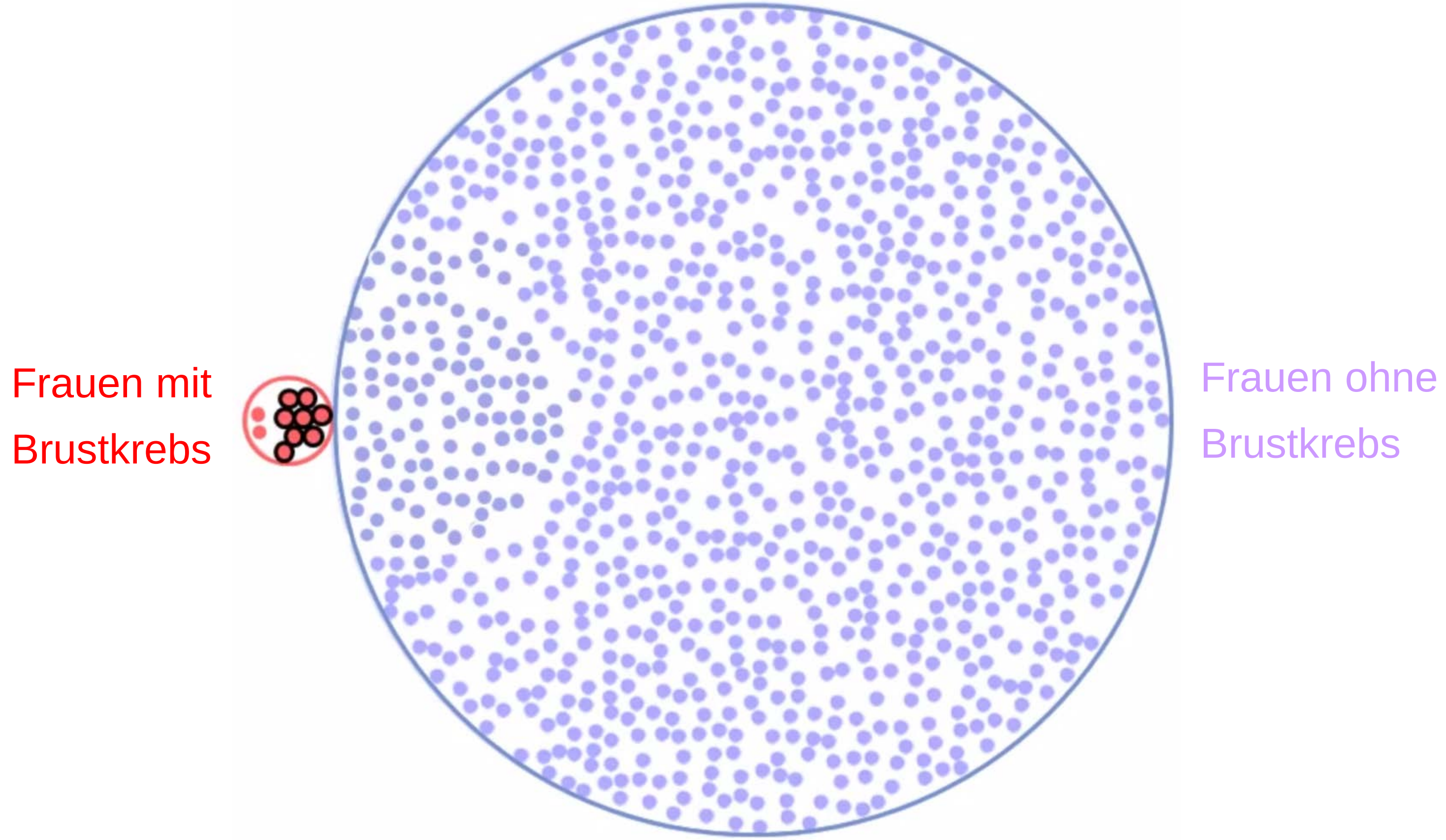
Brustkrebs

- Die W'keit, dass eine 40-jährige Frau Brustkrebs hat ist 1%



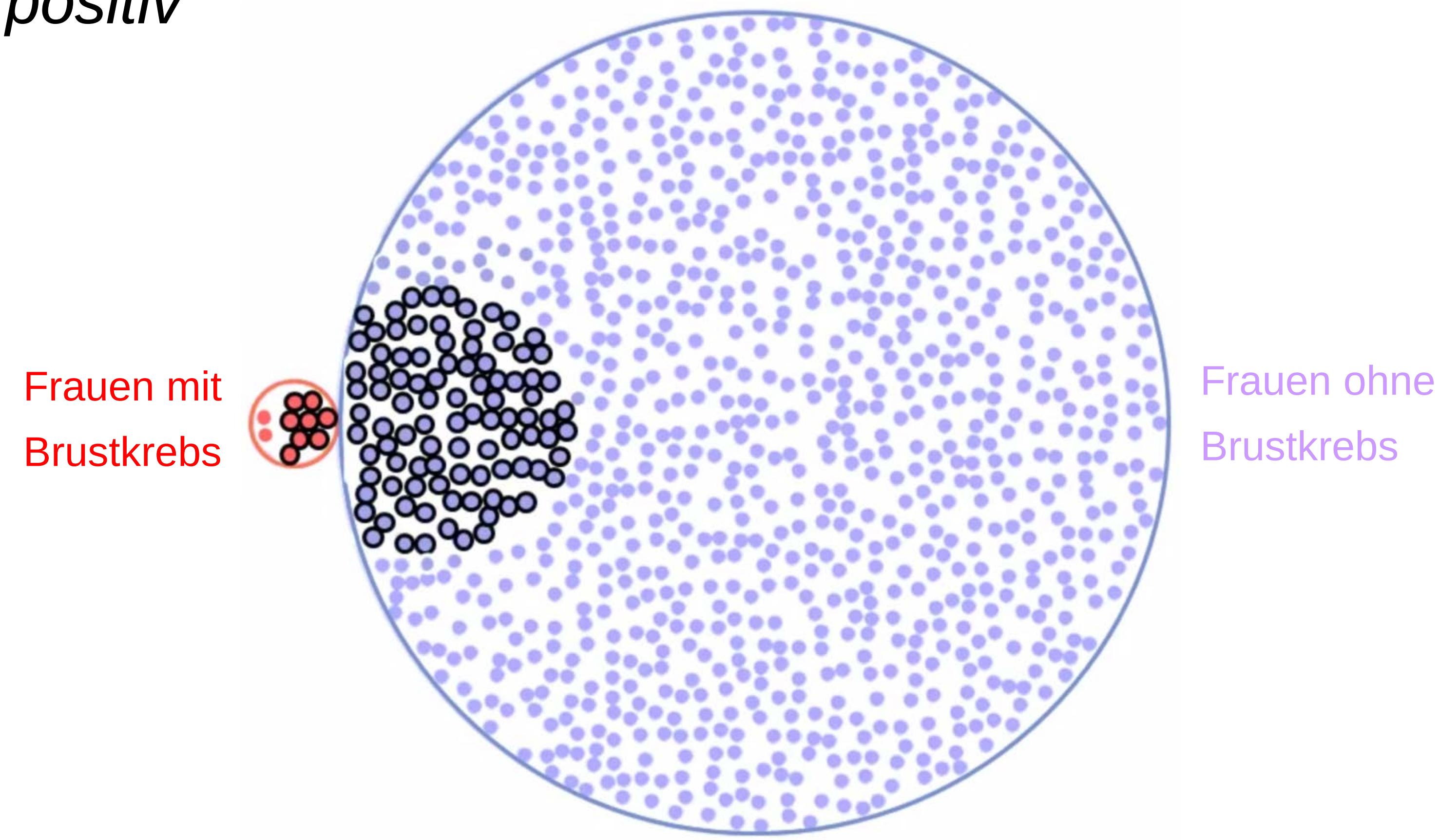
Brustkrebs

- Bei 80% der kranken Frauen fällt der Test *positiv* aus



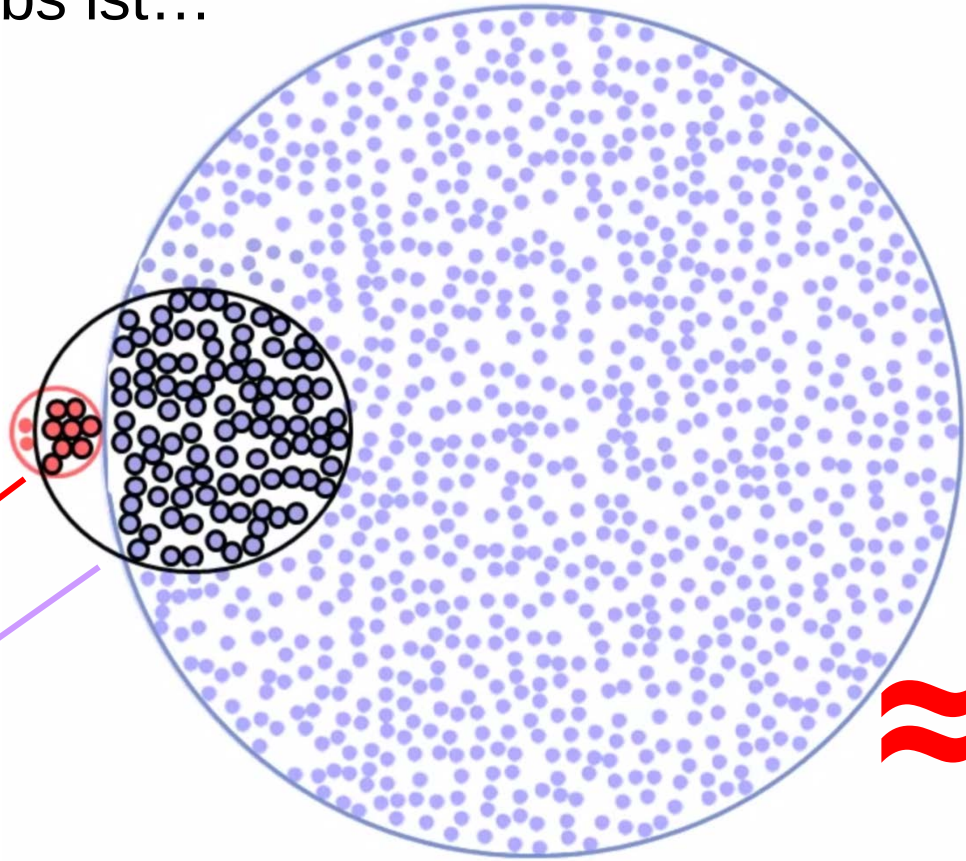
Brustkrebs

- 10% der Mammographien bei den **gesunden** Frauen sind *positiv*



Brustkrebs

- Sie (40, ♀) haben ein positives Testresultat. Die W'keit für Brustkrebs ist…



$$\frac{8}{8+99} \approx 7.5\%$$

Satz von Bayes

- Zusammenhang zwischen $P[A|B]$ und $P[B|A]$:

$$P[A|B] \cdot P[B] = P[A \cap B] = P[B|A] \cdot P[A]$$

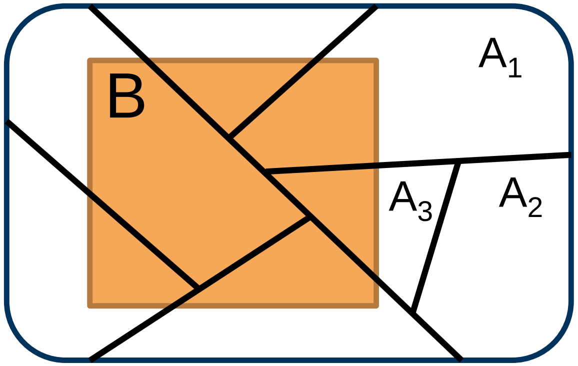
- daraus folgt der Satz von Bayes:

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B|A]P[A]}{P[B]}$$

Satz der Totalen Wahrscheinlichkeit

- Angenommen wir haben eine Partitionierung $A_i, i = 1, \ldots, n$ von Ω, dann gilt:

$$P[B] = \sum_{i=1}^{n} P[A_i \cap B] = \sum_{i=1}^{n} P[B|A_i] P[A_i]$$



➢ Berechne die Fläche von B ohne das Puzzle zusammenzusetzen

Brustkrebs - Berechnung

- Ereignis K: krank, Ereignis T: Test *positiv*
- Geg.: $P[K] = 0.01, P[T|K] = 0.8, P[T|K^C] = 0.1$
- Ges.: $P[K|T]$
- Satz der totalen Wahrscheinlichkeit:

$$P[T] = P[T|K]P[K] + P[T|K^C]P[K^C]$$

  - $P[T] = 0.8 \cdot 0.01 + 0.1 \cdot 0.99 = 0.107$
- Satz von Bayes:

$$P[K|T] = \frac{P[T|K]P[K]}{P[T]}$$

  - $P[K|T] = \frac{0.8 \cdot 0.01}{0.107} \approx 0.0748$

≈ 7.5%

Odds statt Wahrscheinlichkeiten

- Rauchen und Lungenkrebs
  - Doll and Hill, British Med J, 1950, 739-748

	Lungenkrebs (L; case)	Kontrolle (control)	Zeilensumme
Raucher (R)	688	650	**1338**
Nicht Raucher	21	59	**80**
Spaltensumme	**709**	**709**	**1418**

- $P[L|R] \approx \frac{688}{1338} = 0.51; P[L|R^c] \approx \frac{21}{80} = 0.26$
- Alternative zu W'keit → odds *(Wettverhältnis)*
- Beispiel:
  - $odds$(würfle eine 6) $= \frac{1}{6} : \frac{5}{6} = \ 1 \ : \ 5 \ = \ 0.2$

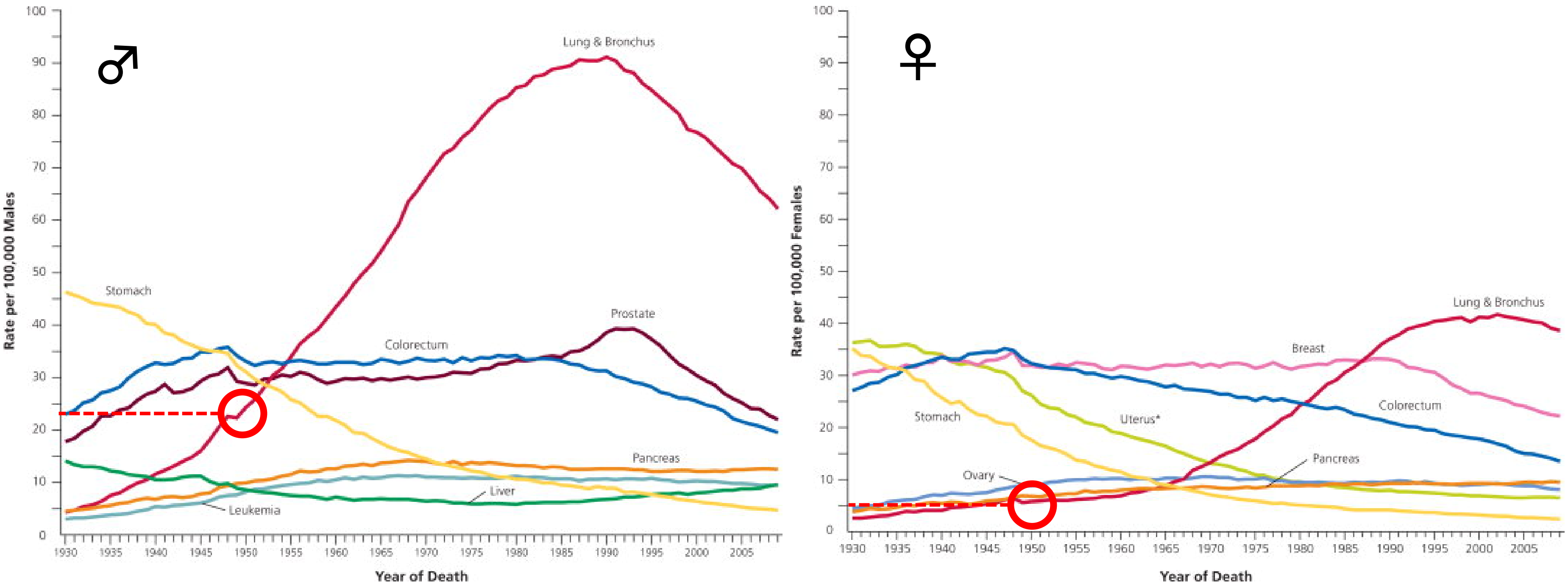
Odds und Odds Ratio (OR)

	Lungenkrebs (L; case)	Kontrolle (control)	Summe
Raucher (R)	688	650	**1338**
Nicht Raucher	21	59	**80**
Summe	**709**	**709**	**1418**

- Wettverhältnis für Lungenkrebs, wenn man raucht
  - $odds(L|R) = \frac{P[L|R]}{P[L^c|R]} = \frac{P[L|R]}{1-P[L|R]} \approx \frac{0.51}{0.49} = 1.04$
- Wettverhältnis für Lungenkrebs, wenn man **nicht** raucht
  - $odds(L|R^C) = \frac{P[L|R^c]}{P[L^c|R^c]} \approx 0.35$
- Das Verhältnis der odds ist aussagekräftig:
  - $OR = \frac{odds(L|R)}{odds(L|R^c)} = \frac{1.04}{0.35} = \boxed{2.97}$

"Die odds an Lungenkrebs zu erkranken sind für Raucher ca. 3 mal grösser."

…das war 1950!



- Aktuelle Zahlen unter http://www.cdc.gov/cancer/lung/basic_info/risk_factors.htm
- *“People who smoke are 15 to 30 times more likely to get lung cancer or die from lung cancer than people who do not smoke.”*
- *“Cigarette smoking is the number one risk factor for lung cancer. In the United States, cigarette smoking causes about 90% of lung cancers.”*

Knew that we ventured

on such dangerous seas

That if we wrought out life

'twas ten to one

—William Shakespeare, *Henry IV, Part II,*

Act I, Scene 1 lines 181–2.

Zusammenfassung

- Lernziele
  - Unabhängigkeit: 50 Rp. zahlen, um zu wissen, ob $AZ \leq 3$
  - Bedingte Wahrscheinlichkeit: Mammographie, wenn man kein Brustkrebs hat
  - odds: 15 bis 30 mal höhere odds für Lungenkrebs als Raucher

- Hausaufgaben
  - Skript: Kapitel 2.2 – 2.3 lessen
  - Serie 2 lösen
  - Quiz 2 bearbeiten
  - etutoR 1 anschauen