

# Schüler-SimuLab

# Kurs 3

**Kursreihe stochastische Simulationen**

**Erste stochastische Simulationen**

Stefan Hartmann

Forschungszentrum caesar

## Was ist eine stochastische Simulation?

### **Definition 1 (stochastische Simulation)**

*Eine **stochastische Simulation** ist eine Nachbildung eines stochastischen Systems auf einem Computer zwecks Untersuchung der Eigenschaften dieses Systems.*

Damit unterscheidet sich also eine stochastischen Simulation sowohl von einem

**echten empirischen (d.h. aus Beobachtungen und Erfahrungen gewonnenen) Experiment,**

bei dem man einen Vorgang in der Natur oder der Wirtschaft beobachtet und Daten erhebt, als auch von einer

**rein mathematisch-analytischen Untersuchung.**

Was genau sind die

**Unterschiede und Gemeinsamkeiten**

zu beiden?

## Gemeinsamkeiten

- mit einem echten empirischen Experiment:

**der empirische Ansatz** (d.h. das Zählen, Messen, etc.)

- mit einer rein mathematischen Untersuchung:

ein zu Grunde liegendes **mathematisches Modell** (als Abbild der Wirklichkeit).

## Beispiel: Lagerhaltung einer Firma

Die Lagerhaltung hängt einerseits ab von

- der Nachfrage
- der Lagerhaltungspolitik
- den Lieferfristen

Nachfrage und Lieferfristen werden stochastisch simuliert.

### Ziel:

**Optimierung der Lagerhaltungspolitik** im Hinblick auf die Kosten und auf die Wahrscheinlichkeit, dass ein verlangter Artikel im Lager ist

**Wieso macht man überhaupt**

Simulationen?

## Gründe für die Simulation

- Reduktion des Aufwandes an Zeit und Geld im Vergleich zu einem echten Experiment
- Reduktion der Komplexität eines Systems im Vergleich zu einer rein mathematisch-analytischen Untersuchung

**Unser Ziel:**

Approximative Berechnung von Wahrscheinlichkeiten und Flächeninhalten (geometrischen Wahrscheinlichkeiten),

deren direkte Berechnung schwierig ist

Umgekehrt:

Überprüfung und Plausibilisierung errechneter Wahrscheinlichkeiten/Flächeninhalte

## Ausgangslage:

- Elementarereignis  $\omega$ : Ausgang einer Messung, Beobachtung
- endlicher Ereignisraum  $\Omega$ : Gesamtheit der Elementarereignisse
- Ereignisse  $A$ : Teilmengen von  $\Omega$

Später:  $\Omega$  unendlich! (Dann: **Vorsicht bei den Ereignissen!**)

## Beispiel: Würfeln mit einem Würfel

- sechs mögliche Elementarereignisse 1, 2, 3, 4, 5, 6
- Ereignisraum  $\Omega = \{1, 2, 3, 4, 5, 6\}$
- mögliches Ereignis „gerade Augenzahl“:  $A = \{2, 4, 6\}$

## Spezielle Ereignisse:

- die leere Menge  $\emptyset$  (das **unmögliche Ereignis**)
- der ganze Ereignisraum  $\Omega$  (das **sichere Ereignis**)

## Potenzmenge

Sei  $\Omega$  eine Menge.

$\mathcal{P}(\Omega)$ : Menge aller Teilmengen von  $\Omega$

### Beispiel

$$\Omega = \{1, 2, 3\}$$

$$\mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \Omega\}$$

Wahrscheinlichkeit  $P$  als Abbildung  $\mathcal{P}(\Omega) \rightarrow [0, 1]$ :

Forderungen:

- I.  $0 \leq P(A) \leq 1$  für alle Ereignisse  $A$ ,
- II.  $P(\emptyset) = 0, P(\Omega) = 1$ ,
- III.  $P(A \cup B) = P(A) + P(B)$ , wenn  $A \cap B = \emptyset$ ,

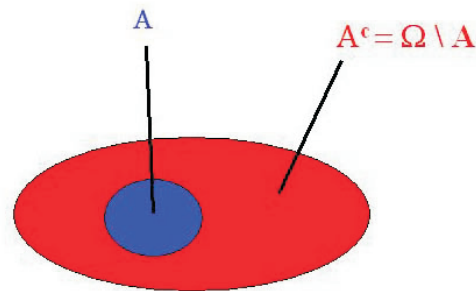
(Bei unendlichen Mengen wird die dritte Bedingung erweitert. Dort betrachtet man eine ganze Folge paarweise disjunkter Ereignisse. Zudem kann man dann  $P$  nicht auf ganz  $\mathcal{P}(\Omega)$  definieren.)

Man nennt  $P$  ein **Wahrscheinlichkeitsmaß**.

## Gruppenaufgabe:

Leite aus diesen Forderungen die Aussage her, dass für das **Komplement**  $A^c := \Omega \setminus A := \{\omega \in \Omega : \omega \notin A\}$  eines Ereignisses  $A$  die folgende Aussage gilt:

$$P(A^c) = 1 - P(A).$$



Komplement einer Menge

## Generelles Vorgehen bei stochastischen Simulation

- führe  $n$  Simulationsschritte durch
- lasse den Computer bei jedem Schritt bestimmen, ob  $A$  eingetroffen ist oder nicht
- bestimme die relative Häufigkeit dafür, dass  $A$  eingetreten ist:

$$h_n(A) = \frac{\text{Anzahl der Simulationsschritte, in denen } A \text{ eingetreten ist}}{n}.$$

- $h_n(A)$  approximiert  $P(A)$ :

$$h_n(A) \approx P(A) \quad \text{für große } n$$

Monte-Carlo-Simulation

# Das Galtonbrett



Das Galton-Brett in Excel

## Gruppenaufgabe:

Bevor du die Simulation durchführst: Glaubst du, dass in der ganz untersten Reihe nach einiger Zeit in allen Feldern ungefähr gleich viele Kugeln ankommen oder wird es Unterschiede geben? Wo werden tendenziell die meisten Kugeln ankommen?

Male dir mal ein Galton-Brett mit nur vier Reihen auf und fasse das (falls du so etwas kennst!) als Baumdiagramm mit der Elementarwahrscheinlichkeit  $p = \frac{1}{2}$  auf. Berechne jetzt die Wahrscheinlichkeit, mit der die Kugeln in den drei Kästchen der dritten Reihe ankommen.

⇒ *am Smartboard erklären!*

## Gruppenaufgabe:

Führe die Simulation jetzt ein paar Mal durch. Vergleiche dabei die relativen Häufigkeiten, die sich aus dem Simulationsexperiment ergeben, mit den gerade angegebenen theoretischen Wahrscheinlichkeiten. Dies kannst du auch in dem Excel-Blatt mit Hilfe des Diagramms optisch verfolgen. Ändere jetzt mal den Wert  $p$ , also den Wert für die Wahrscheinlichkeit, mit der die Kugel nach links rollt. Wie ändern sich jetzt die theoretischen Wahrscheinlichkeiten? Wird das auch in der Simulation deutlich?

## Gruppenaufgabe:

Wir versuchen nun zusammen, ein solches Galton-Brett mit elementaren Tabellenfunktionen selber kurz in Excel zu erstellen.

## Beispiel: Multiple-Choice-Test

- 12 Fragen mit jeweils 4 Antwortmöglichkeiten
- genau eine ist richtig
- man kreuzt nach Zufall an

**Ähnliche Situation wie im Galton-Brett!**

## Gruppenaufgabe:

Nehmen wir an man besteht den Test, wenn man mindestens die Hälfte der Fragen richtig beantwortet hat. Bestimme mit der stochastischen Simulation des Galton-Bretts approximativ (also näherungsweise) die Wahrscheinlichkeit für dieses Ereignis. Lohnt es sich also zu raten? ;-)

**Ziel:**

Erhalte einen Näherungswert für die Kreiszahl  $\pi$ !

**Methode:**

Monte-Carlo-Simulation

## „geometrische“ Wahrscheinlichkeiten:

Zunächst im Eindimensionalen:

$I$  : endliches Intervall, zum Beispiel  $I = [0, 2]$ .

$|I|$  : Länge von  $I$  (z.B.  $|[0, 2]| = 2$ .)

$A$ : Teilintervall von  $I$

Definiere:

$$P(A) := \frac{|A|}{|I|}$$

Im Zweidimensionalen:

$G$ : Gebiet

$|G|$ : Flächeninhalt von  $G$

$A$ : Teilmenge von  $G$  mit existierendem Flächeninhalt (!)

Definiere:

$$P(A) := \frac{|A|}{|G|}$$

Jeder Kurvenzug (zum Beispiel ein Geradenstück) besitzt dann die Wahrscheinlichkeit Null, da solche Kurvenzüge den Flächeninhalt Null besitzen.

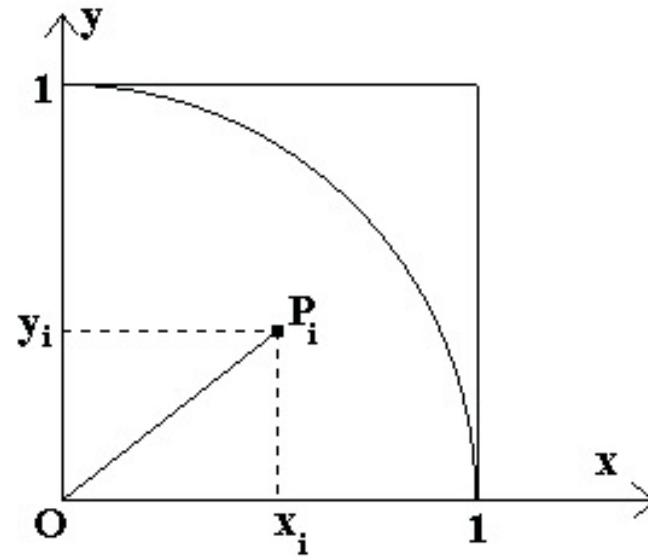
### **Achtung:**

Daraus folgt nicht, dass kein Punkt auf diesem Kurvenzug eintreten kann. Die Menge der Punkte auf einem solchen Kurvenzug ist aber im Vergleich zur Gesamtmenge  $G$  vernachlässigbar. Daher wird das Versuchsergebnis höchst selten (d.h. fast nie, also mit keiner positiven Wahrscheinlichkeit) auf diesem Kurvenzug liegen.

⇒ *Arbeitsblatt 1*

Wir betrachten das Einheitsquadrat

$$G = [0, 1]^2 = [0, 1] \times [0, 1] = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$



Viertelkreis  $A$  im Einheitsquadrat  $G$

Wir haben:  $|G| = 1$

Fläche eines Kreises mit Radius  $r$ :

$$\text{Fläche}_{\text{Kreis}} = \pi \cdot r^2.$$

Also:

$$|A| = \frac{1}{4} \pi r^2 = \frac{\pi}{4}$$

$$\Rightarrow P(A) = \frac{|A|}{|G|} = \frac{\pi}{4}$$

## Prinzip:

- I. Wir erzeugen  $n$  Zufallswerte aus  $[0, 1] \times [0, 1]$ , also Paare von Zufallszahlen.
- II. Liegen  $k$  der  $n$  zufälligen Paare in  $A$  liegen, so bilden wir die relative Häufigkeit:

$$h_n(A) = \frac{k}{n}.$$

- III. Wir schätzen (für große  $n$ ):

$$\frac{\pi}{4} = P(A) \approx h_n(A) = \frac{k}{n} \quad \Rightarrow \quad \pi \approx 4 \cdot \frac{k}{n}$$

## Aufgabe 1:

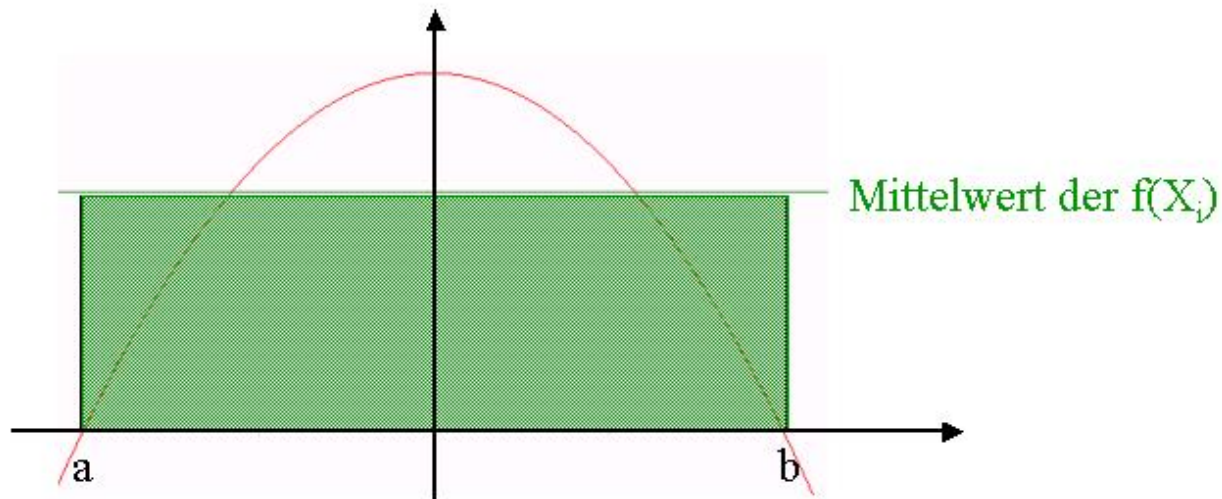
Versuche nun in einem Excel-Blatt (und mit Hilfe der Zufallszahlen von Excel) einen Näherungswert für  $\pi$  zu bestimmen, indem du ein kleines stochastisches Simulationsprogramm schreibst und dann  $n = 10000$  Simulationsschritte durchführt.

## Aufgabe 2:

Berechne mit einer Monte-Carlo-Simulation näherungsweise den Flächeninhalt zwischen dem Graphen der Funktion  $f(x) = -\frac{1}{4}x^2 + 1$  und der  $x$ -Achse. Betrachte dazu die Obermenge  $G = [-2, 2] \times [0, 1]$  und überlege dir, wann ein Punkt aus  $G$  zwischen dem Graphen und der  $x$ -Achse liegt. Eine Skizze ist hier sehr hilfreich! Beachte, dass du Zufallszahlen aus  $[0, 1]$  zunächst auf  $[-2, 2]$  transformieren musst. Ist nun  $X$  eine Zufallszahl aus  $[-2, 2]$  und  $Y$  eine Zufallszahl aus  $[0, 1]$ , dann musst du  $f(X)$  und  $Y$  vergleichen.

## Zusatzaufgabe:

Bilde jetzt mal den Mittelwert der  $f(X_i)$ , wobei die  $X_i$  die Zufallszahlen aus  $[-2, 2]$  und multipliziere den Wert wiederum mit 4.



## Die Entenjagd

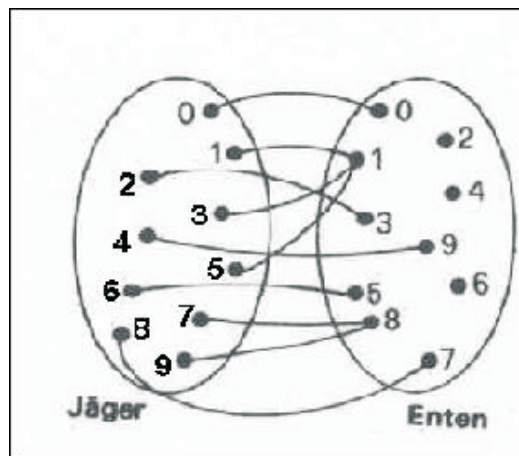
Zehn treffsichere Jäger schießen auf zehn Enten (ein Schuss, eventuell schießen mehrere auf die gleiche). Wie viele Enten überleben im Durchschnitt?

### Gruppenaufgabe:

Bevor du weiterliest: Gib erst einmal eine Schätzung ab! Was vermutest du? Wie viele Enten überleben im Durchschnitt? Jetzt überlege dir: Wie könnte man dieses Experiment mit dem Computer simulieren? Hast du Ideen? Diskutiert in der Gruppe darüber, wie man das anstellen könnte.

⇒ *Arbeitsblatt 2*

Man erzeugt 10 Zufallsziffern von 0 bis 9. Die fehlenden Ziffern sind die Nummern der überlebenden Enten. Die folgende Abbildung zeigt ein von der Ziffernfolge 0131915878 gesteuertes Experiment:



Veranschaulichung des Ziffernblocks 0131915878 bei der Entenjagd

### Aufgabe 3:

Nehmen wir an, wir haben das Experiment simulativ durchgeführt und nach 5 Simulationsschritten die folgenden Blöcke erzeugt:

9416956090 ,  
7506682320 ,  
4628819077 ,  
7592050242 ,  
4233819129 .

Wie viele Enten haben in diesen 5 Simulationsschritten überlebt?  
Wie lautet somit der Mittelwert an überlebenden Enten?

### **Gruppenaufgabe:**

Schau dir das Excel-Programm zur Entenjagd an und versuche nachzuvollziehen, wie es arbeitet. Führe die Simulation nun für verschiedene Anzahlen durch und notiere dir die Werte.

Der theoretische Mittelwert (genauer: der Erwartungswert) für die Anzahl der überlebenden Enten beträgt übrigens:

$$m = 10 \cdot \left(\frac{9}{10}\right)^{10} \approx 3,4868.$$

Im Mittel überleben also ungefähr drei oder vier Enten.

# Eine Kuriosität: Die Würfelschlange



### **Gruppenaufgabe:**

Führt jetzt die Simulation ein paar Mal durch, um diese Beobachtung zu bestätigen. Kannst du dir eine Erklärung für dieses Phänomen vorstellen? Diskutiert diese Frage untereinander.

## Eine Erklärung des Phänomens



Die türkis markierten Würfel gehören zum Pfad mit dem ersten Würfel als Startwürfel. Damit gehören auch der dritte und vierte Würfel zum Pfad. Startet man vom fünften oder sechsten Würfel, landet man ebenfalls auf dem Pfad (nämlich beim achten Würfel). Nur der zweite Würfel hat einen anderen Endwürfel (nämlich den siebten Würfel, danach geht es nicht mehr weiter).

## Eine Abschätzung für die Wahrscheinlichkeit am gleichen Würfel auszukommen

bei 60 Würfeln:

$$1 - \left(\frac{5}{6}\right)^{10} \approx 0,84.$$

Dies erscheint nicht so hoch, wie es die Simulation erwarten lässt, aber man muss natürlich bedenken, dass wir hier nur eine **Abschätzung** vorgenommen haben und die theoretisch exakte Wahrscheinlichkeit für das Erreichen des Pfades sehr viel höher ist!

# Das Ziegenproblem



die amerikanische Spielshow „Let's make a deal“

## Gruppenaufgabe:

Spielt das Spiel jetzt ein paar Mal am Computer. Was fällt euch auf? Was meint ihr, auch aufgrund eurer Spielerfahrung: Ist es besser für den Kandidaten zu wechseln oder sollte er besser bei seiner ersten Wahl bleiben? Oder ist es völlig egal, ob er sich umentscheidet oder nicht, weil die Wahrscheinlichkeit auf den Autogewinn eh in beiden Fällen gleich hoch ist? Diskutiert ruhig darüber! Zumeist sind Diskussionen über dieses Problem recht emotional. ;-)

In der Kolumne „Ask Marilyn“ des amerikanischen Wochenmagazins „Parade“ erklärte die Journalistin Marilyn vos Savant:

**Es ist besser zu wechseln, da sich die Gewinnchancen dann verdoppeln!**



Marilyn vos Savant

Das war die Geburtsstunde des Ziegenproblems im Jahre 1991!

## Die Begründung von Marilyn

Die Wahrscheinlichkeit, dass sich das Auto hinter der Tür 1 befindet, ist  $\frac{1}{3}$ . Die Wahrscheinlichkeit, dass es sich hinter **einer der beiden anderen** Türen befindet, ist  $\frac{2}{3}$ . Der Moderator öffnet nun eine der beiden Türen, aber **nicht zufällig**, denn er weiß ja, wo das Auto steht. Die Wahrscheinlichkeit, dass es eine der beiden zu Beginn nicht gewählten Türen war, bleibt aber  $\frac{2}{3}$ . Es steht nun nur halt für den Außenstehenden fest, welche der beiden Türen die Wahrscheinlichkeit von  $\frac{2}{3}$  trägt.

## Vielleicht einsichtigere Begründung

Der Standhafte gewinnt genau dann das Auto, wenn sich dieses hinter der ursprünglich gewählten Tür befindet, und die Wahrscheinlichkeit dafür ist  $\frac{1}{3}$ . Ein Wechsler gewinnt hingegen das Auto genau dann, wenn er sich zuerst für eine „Ziegentür“ entschieden hat (die Wahrscheinlichkeit hierfür ist  $p = \frac{2}{3}$  und in diesem Fall beim Wechseln ja durch das Öffnen der zweiten Ziegentür durch den Moderator automatisch auf die „Autotür“ geführt wird.

Das Problem ist also, dass die Chancen nach dem Öffnen der Tür durch den Moderator eben nicht 50:50 sind, sondern davon abhängen, ob man sich zu Beginn für die richtige Tür oder aber die falsche Tür entschieden hat.

Gibt es auch eine  
mathematischere Herleitung  
mit Methoden der  
Wahrscheinlichkeitsrechnung?

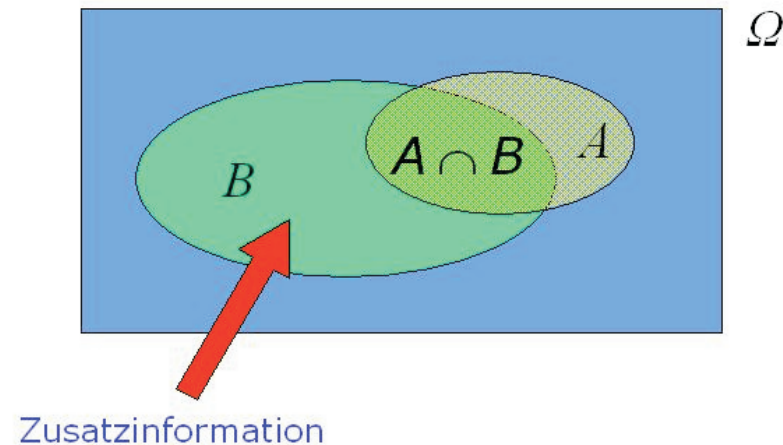
⇒ *Arbeitsblatt 3*

## **Definition 2 (bedingte Wahrscheinlichkeiten: $P(A|B)$ )**

*Die bedingte Wahrscheinlichkeit  $P(A|B)$  des Ereignisses  $A$  bei gegebenem Ereignis  $B$  mit  $P(B) > 0$  ist definiert als Quotient*

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

Man kann  $P(A|B)$  als die Wahrscheinlichkeit dafür ansehen, dass das Ereignis  $A$  eintritt, wenn man schon weiß, dass das Ereignis  $B$  eingetroffen ist, also als Wahrscheinlichkeit von  $A$  unter der **Zusatzinformation**  $B$ :



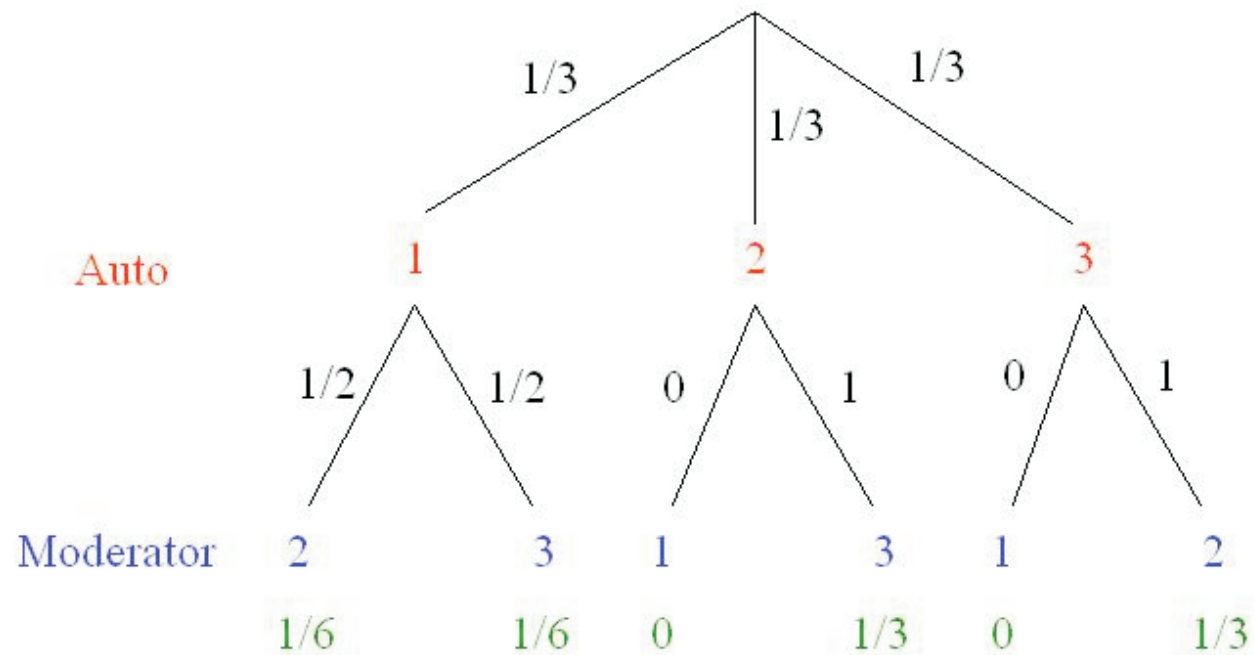
bedingte Wahrscheinlichkeiten als Zusatzinformationen

## Aufgabe 4:

- (a) Wie groß ist die Wahrscheinlichkeit dafür, dass man eine Zahl gewürfelt hat, die kleiner als 4 ist, wenn man bereits weiß, dass man eine Primzahl gewürfelt hat?
- (b) Wie groß ist die Wahrscheinlichkeit dafür, dass man weder eine 3 noch eine 5 gewürfelt hat unter der Voraussetzung, dass man keine 4 gewürfelt hat?

⇒ *Arbeitsblatt 4*

Entscheidet sich der Kandidat zunächst für die Tür Nummer 1, dann ergibt sich das folgende Baumdiagramm:



Baumdiagramm für das Ziegenproblem

Wir entscheiden uns für Tür 1, der Moderator öffnet Tür 3.

**Welcher Wert ist höher:**

$P(\text{das Auto befindet sich hinter Tür 2} \mid \text{der Moderator öffnet Tür 3})$

oder

$P(\text{das Auto befindet sich hinter Tür 1} \mid \text{der Moderator öffnet Tür 3})$  ?

## Zusammenfassung:

Die bedingte Wahrscheinlichkeit (unter der Bedingung, dass der Moderator die Tür 3 geöffnet hat) dafür, dass sich das Auto hinter der verbliebenen dritten Tür befindet, ist mit  $p = \frac{2}{3}$  doppelt so hoch wie die bedingte Wahrscheinlichkeit dafür, dass sich das Auto hinter der ersten Türwahl verbrigt.

**Daher ist es besser zu wechseln. Ein Wechsel verdoppelt die Gewinnchancen!**

## Aufgabe 5:

Wir spielen das Ziegenspiel jetzt einmal mit vier Türen. Hinter einer der vier Türen befindet sich ein Auto, hinter drei der vier Türen eine Ziege. Nehmen wir weiter an du hast dich für die erste Tür entschieden und der Moderator öffnet die zweite Tür. Ist es dann besser zur dritten (oder vierten) Tür zu wechseln oder bleibst du besser bei deiner ersten Wahl? Oder ist es völlig egal, da die Gewinnchancen eh die gleichen sind? Male dir wie oben ein Baumdiagramm auf und berechne die bedingten Wahrscheinlichkeiten. Wie sieht das Ganze wohl mit 100 Türen und einer Ziege aus?

## Gruppenaufgabe:

Wir versuchen uns nun zusammen am Smartboard einen Algorithmus für die Programmierung der Simulation zu überlegen.