

# Hypothesentest für gebackene Würfel

$$p(\text{😊} | \text{🎲}) = 1$$

Ein Unterrichtsentwurf aus der Reihe **Probably Fun – Games to teach Statistics**

von Dr. Kristian Rother ([www.academis.eu/probably\\_fun/](http://www.academis.eu/probably_fun/))

Nutzbar unter den Bedingungen der Creative Commons Attribution Share-Alike License 4.0

## Unterrichtsziel

Die Teilnehmer prüfen mit einem  $\chi^2$ -Test, ob ein Würfel ausbalanciert ist.

## Zeit

90 Minuten

## Begriffe

- Nullhypothese
- alternative Hypothese
- statistisches Testverfahren
- Chi-Quadrat-Test
- Signifikanzniveau (alpha)
- p-Wert
- predictive power



Nutzbar unter den Bedingungen  
der CC-BY-SA 4.0 Lizenz

$$p(\text{😊} | \text{🎲}) = 1$$

## Würfel backen



*Würfel mit unterschiedlichen Backzeiten*

Dies ist ein didaktisches “Rezept” der besonderen Art. Kurz: du bäckst einige W6, um sie zu manipulieren. Zur Vorbereitung des Unterrichts benötigst du eine Packung mit sechsseitigen Würfeln. Ich empfehle normale Würfel ohne Glitzer oder anderen Schnickschnack, da diese im Ofen manchmal aufschäumen.

Zum Backen gehe folgendermaßen vor:

- heizen Sie den Backofen auf 120 Grad Celsius vor
- schiebe einen einzelnen sechsseitigen Würfel auf einem Stück Alufolie in den Ofen. Lege die 6 nach oben (wir wollen ja auch etwas von unserer Mühe haben).
- starte eine Stoppuhr
- stelle eine Schüssel mit Eiswasser bereit
- Nach 10-20 Minuten nimm den Würfel heraus und wirf ihn in das Eiswasser
- Wiederhole den Vorgang mit einem anderen Würfel und variiere die Backzeit oder Temperatur

Wenn die Temperatur zu hoch ist, verwandelt sich der Würfel in ein amorphes Etwas. Ist sie zu niedrig, passiert nichts. Stimmen die Bedingungen jedoch, sackt das Material etwas nach unten, so dass die oben liegende Zahl häufiger fällt. Die Temperatur in den meisten Öfen ist leider nicht sehr stabil. Wenn du ein Ofenthermometer oder ein Physikkabor zur Verfügung hast, hilft das, diesen Teil wissenschaftlich präzise durchzuführen. Es geht aber auch gut ohne.



Nutzbar unter den Bedingungen  
der CC-BY-SA 4.0 Lizenz

$$p(\text{😊} | \text{🎲}) = 1$$

Die Würfel dürfen übrigens ruhig als manipuliert erkennbar sein. Wir wollen ja Mathe machen und nicht schummeln, oder?

### Anmerkung:

Du kannst die Würfel auch gerne anders verunstalten: Ecken absägen oder Bleigewichte hineinoperieren ist hier erlaubt. Wenn du einen enthusiastischen Kurs hast, kannst du die Backerei auch an die Teilnehmer delegieren.

## Unterrichtsablauf

Hast du erst einmal manipulierte Würfel, könnt ihr im Unterricht die Frage bearbeiten, ob sich denn beweisen lässt, dass ein Würfel manipuliert ist. Hier kommt der Hypothesentest ins Spiel. Stelle zunächst den Ablauf des Tests vor (siehe Tabelle). Achte penibel darauf, dass die Signifikanzschwelle festgelegt ist, bevor ihr anfangt Daten zu sammeln. Sobald die Teilnehmer anfangen zu würfeln wird es auch etwas lauter. Wichtig ist auch, mindestens zwei Kontrollen mitzuführen: ein imaginärer Würfel der perfekt ausbalanciert ist, sowie ein zweiter der nur Sechsen würfelt. So schützt ihr euch davor, euch an blöden Fehlern zu verhaspeln. Das Review mit ChatGPT dient übrigens dem gleichen Zweck.

Ich habe in dieser Lektion eine Menge komplizierter Dinge ausgelassen: andere Tests, Fehlertypen, Vorhersagekraft, denn dies war für die Teilnehmer die erste Begegnung mit Hypothesenteststesten. Stattdessen habe ich einige Zeit damit verbracht, die Methode mit umfangreichen Warnungen zu versehen. Wir haben Beispiele für p-Hacking und die Reproduzierbarkeitskrise besprochen. Meistens schicke ich eine Stunde über Erkenntnistheorie und die wissenschaftliche Methode voraus, um die Frage "Wie können wir überhaupt etwas wissen?" bei den Teilnehmern zu verankern.

Schritt	Aktivität	Zeit
1.	Erkläre das Rezept zum Durchführen eines Hypothesentests (die weiteren Schritte)	10'
2.	Verteile Würfel an die Teilnehmer und lasse sie in Paaren 120 Mal würfeln und dabei eine Strichliste über die Ergebnisse führen	15'
3.	Formuliere eine falsifizierbare Nullhypothese ("der W6 ist ausgewogen")	5'
4.	Wähle eine Teststatistik (in diesem Fall Chi-Quadrat)	5'
5.	Lege die Signifikanzschwelle alpha fest (z.b. 90%)	5'
6.	Schlage den kritischen Bereich mit Hilfe der Anzahl Freiheitsgrade (hier n=5)	5'









Nutzbar unter den Bedingungen  
der CC-BY-SA 4.0 Lizenz

$$p(\text{😊} | \text{🎲}) = 1$$

7.	Berechne die Teststatistik aus den Messergebnissen	10'
8.	Prüfe ob die Teststatistik im kritischen Bereich liegt	5'
9.	Akzeptiere oder verwirfe die Nullhypothese	5'
10.	Gib die Messwerte in ChatGPT oder ein anderes Sprachmodell ein, um deine Berechnung zu überprüfen	5'

## Ergebnisse

Ein Tafelbild mit den Würfelerggebnissen und Chi-Quadrat-Werten. Ein Würfel wurde vom Test als manipuliert erkannt, die anderen müßten länger backen oder es müßte öfter gewürfelt werden.

	unbaked		baked			=E		
	Adi	Seb	Dar	Tho	Ber	Kvi	C1	C2
E=20								
	14 <sub>36</sub>	25	25 <sub>25</sub>	18	13	20 <sub>0</sub>	20	0
	18 <sub>4</sub>	21	21 <sub>1</sub>	16	27	7 <sub>169</sub>	20	0
	13 <sub>49</sub>	19	17 <sub>9</sub>	21	21	17 <sub>9</sub>	20	0
	17 <sub>9</sub>	20	21 <sub>1</sub>	20	23	32 <sub>144</sub>	20	0
	30 <sub>100</sub>	22	16 <sub>16</sub>	28	21	28 <sub>64</sub>	20	0
	17 <sub>9</sub>	13	20 <sub>52</sub>	17	15	15 <sub>25</sub>	20	100
						<u>493</u>	0	420
$\chi^2$	10.75		2.6			20.65		
p-value:	0.966		0.76			0.0009	1.	<0.00001
	>0.05		$H_0$ holds			<0.05	✓	✓
	$H_0$ holds					$H_0$ rejected		