11 Umfragen und Wahrscheinlichkeit

Einführung in die quantitativen Forschungsmethoden

Heute

- Wahrscheinlichkeit
- Umfragen und Wahrscheinlichkeit

Wahrscheinlichkeit

Das Problem

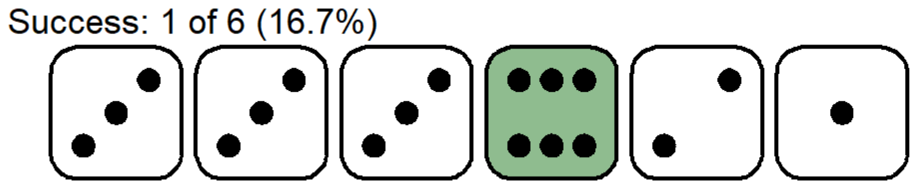
- **statistische Messwerte** - wie der Mittelwert - helfen, Daten zu verstehen
- **statistische Signifikanz**: Ist ein Phänomen - z.B. unser Messwert - ein Zufallsprodukt oder ein realer Effekt?
- **statistische Tests** helfen uns mit Prüfgrößen zu testen, ob ein Resultat auf einem festgelegten Niveau signifikant ist

Anwendungen

- Ist der Würfel gezinkt?
- Welche Gerste ist gut, um Bier zu brauen?
- Sind die Noten bei Prof. X besser als bei Professor Y?
- Ist der gemessene Effekt aussagekräftig?

Würfeln: Wer hat die meisten Augen?

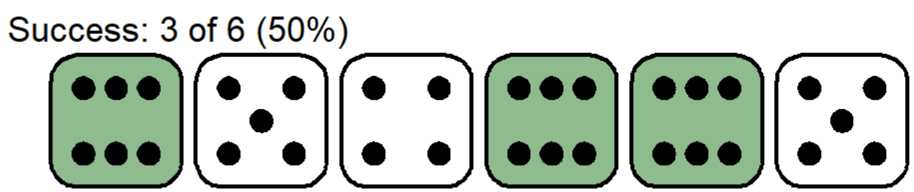
Würfelset 1



18

. . .

Würfelset 2



32

Erwartungswert

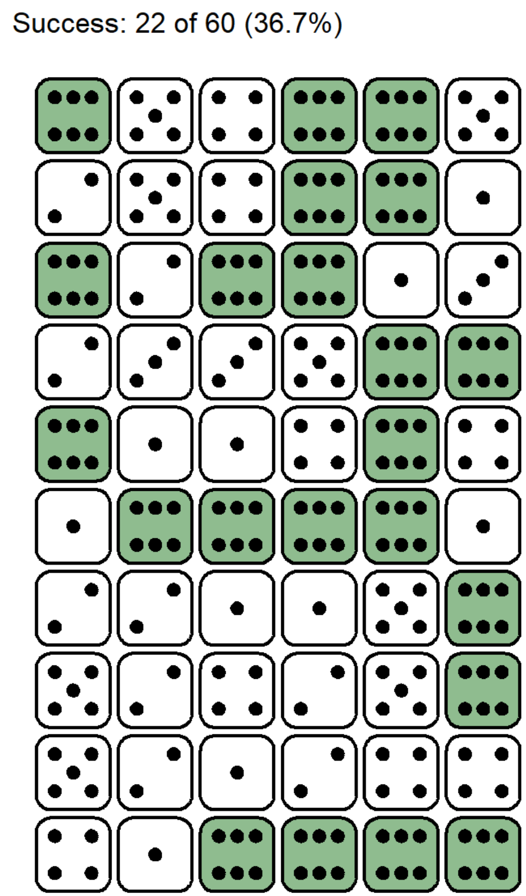
Was würden Sie erwarten? Welche Augensumme sollten wir typischerweise mit 6 Würfen erreichen?

Je Wurf: $^1/_6 \times 1 + {}^1/_6 \times 2 + {}^1/_6 \times 3 + {}^1/_6 \times 4 + {}^1/_6 \times 5 + {}^1/_6 \times 6 = \mathbf{3.5}$

Summe: $6 \times 3.5 = \mathbf{21}$

→ Der **Erwartungswert** über viele Versuche hinweg ist 21

Mehr Experimente mit Würfelset 2



1	32
2	24
3	24
4	25
5	22
6	26
7	17
8	24
9	18
10	29

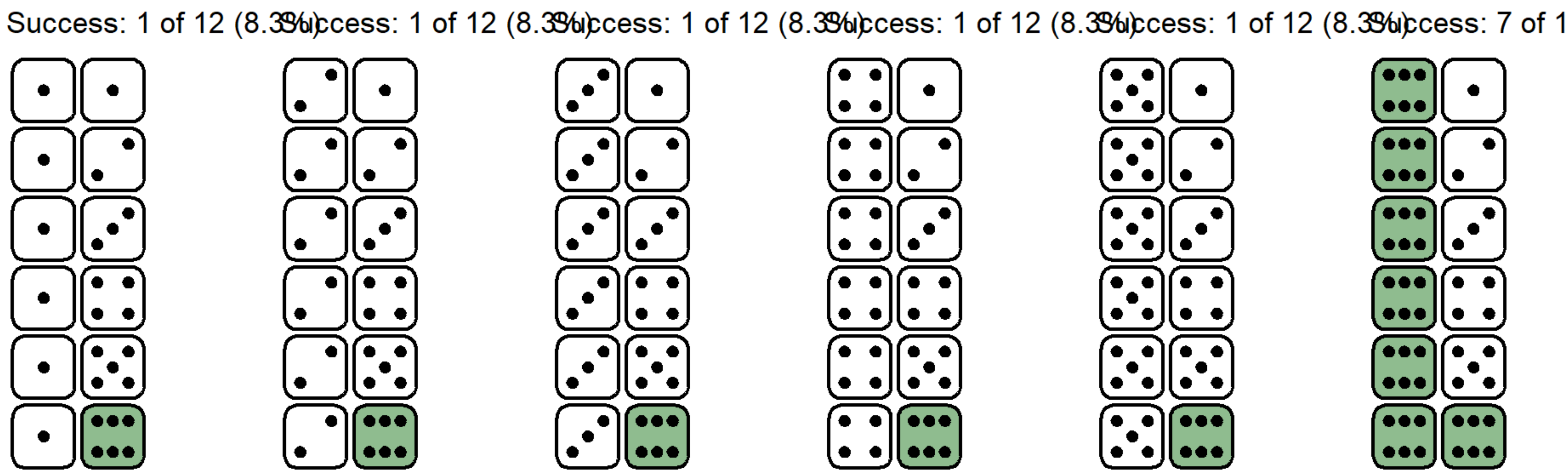
→ **Wie bestimmen wir die Wahrscheinlichkeit, dass die Würfel gezinkt sind?**

→ Frage: **Stammen die Würfe (nicht) aus einer Population mit Mittelwert 21?**

Gedankenexperiment: Mögliche Würfe

Vereinfachung: Population zwei Würfel

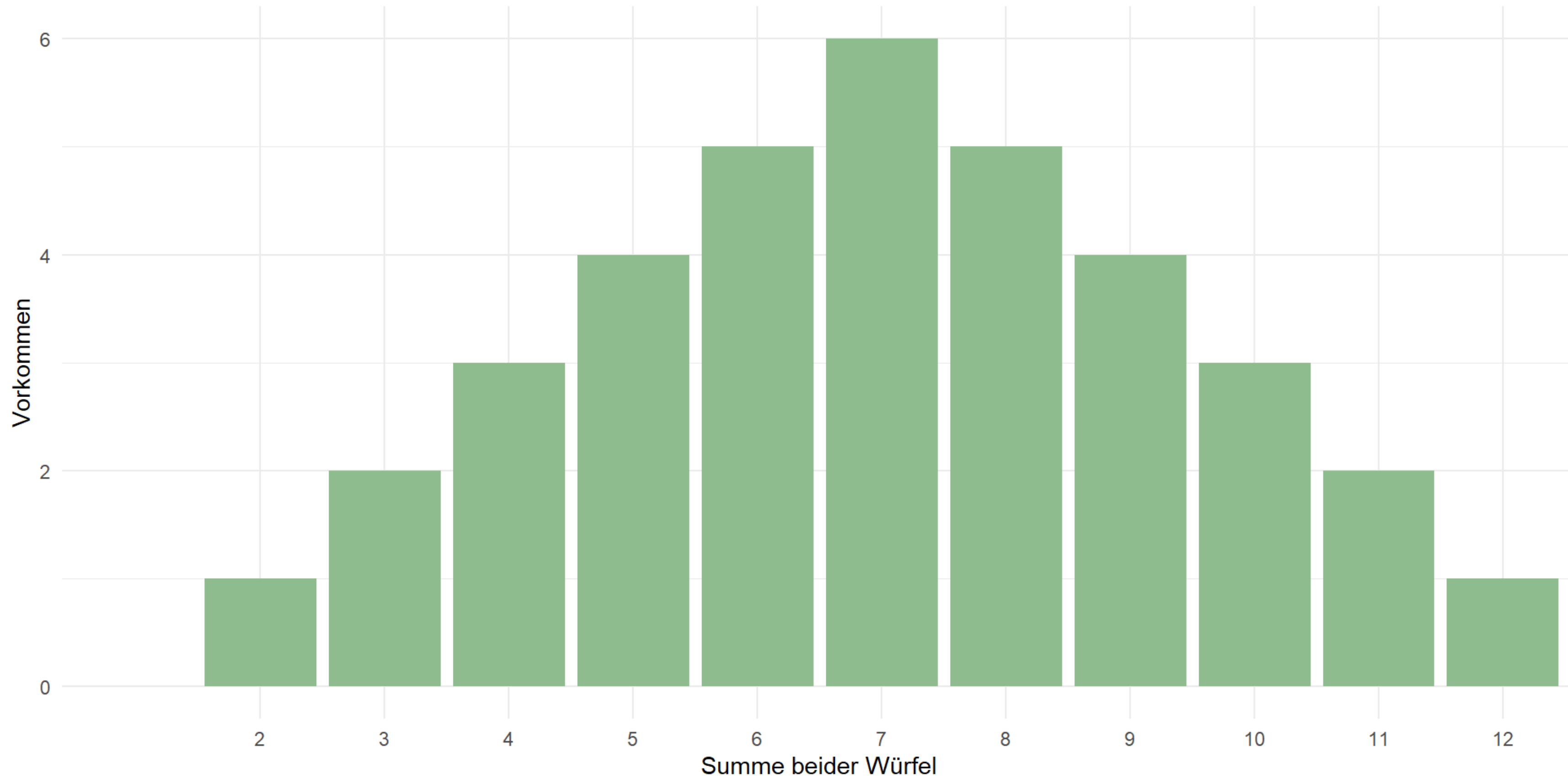
...



Gedankenexperiment: Mögliche Würfe

Verteilung der Augensummen (2 Würfel)

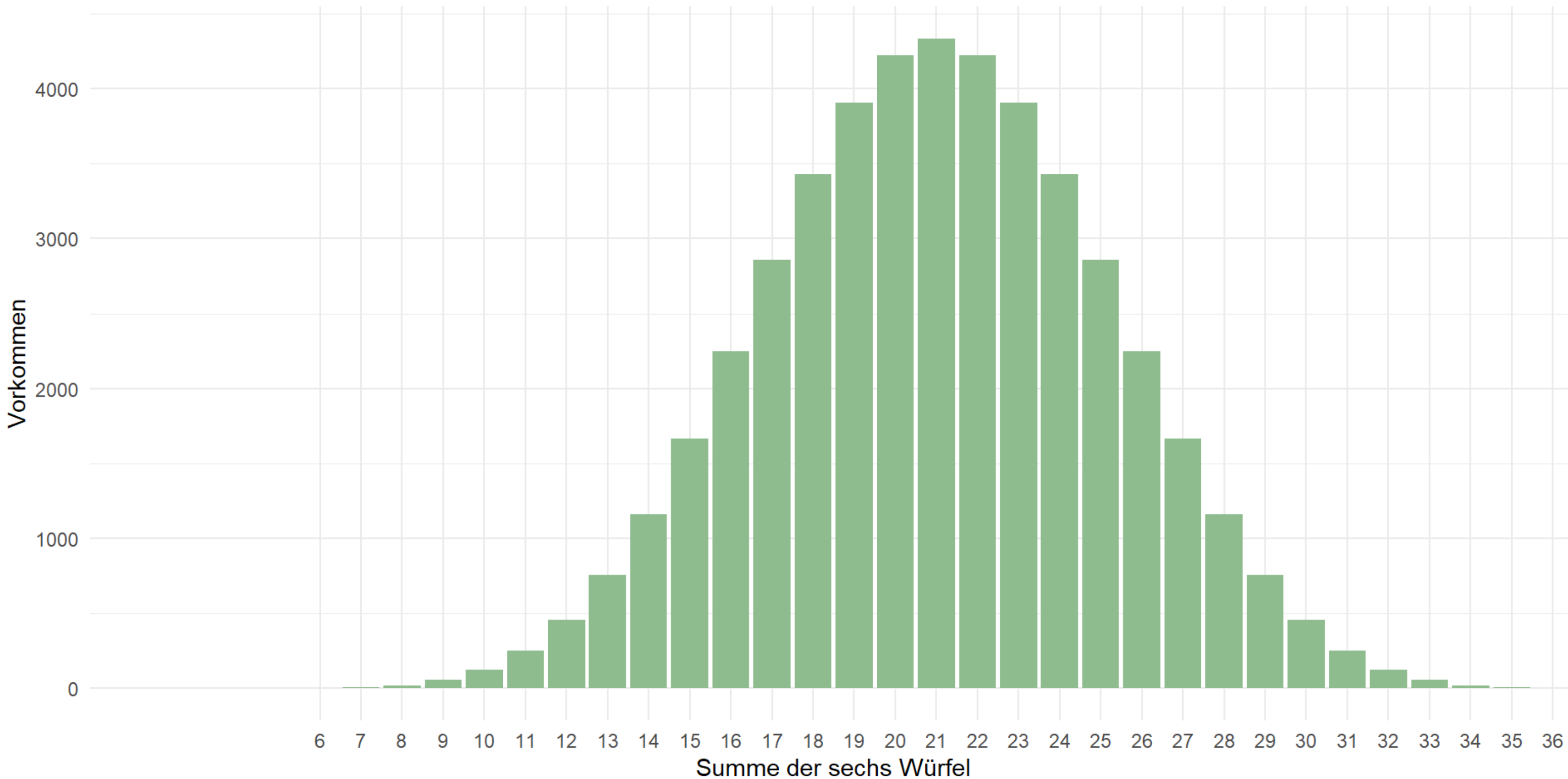
. . .



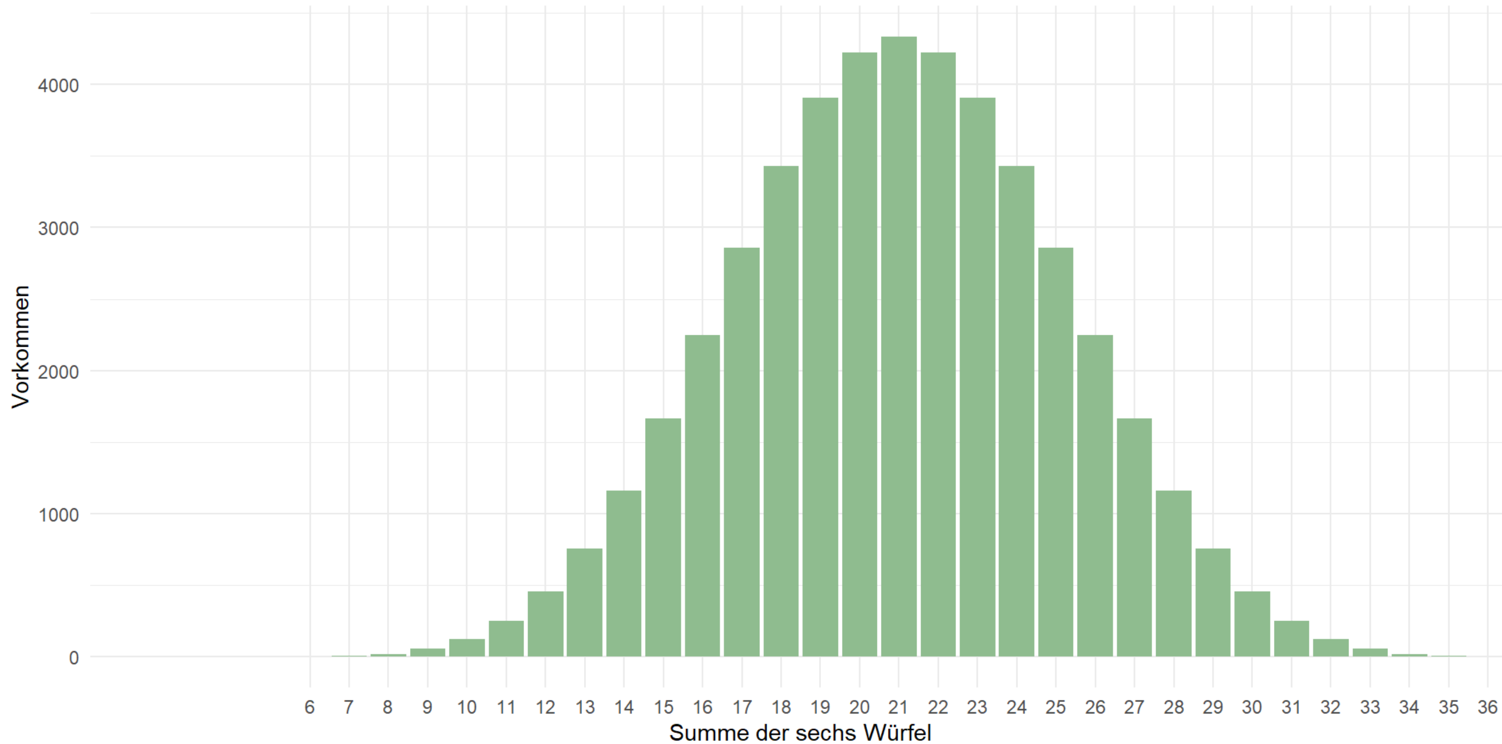
Gedankenexperiment: Mögliche Würfe

Verteilung der Augensummen (6 Würfel)

...



Gedankenexperiment: Mögliche Würfe

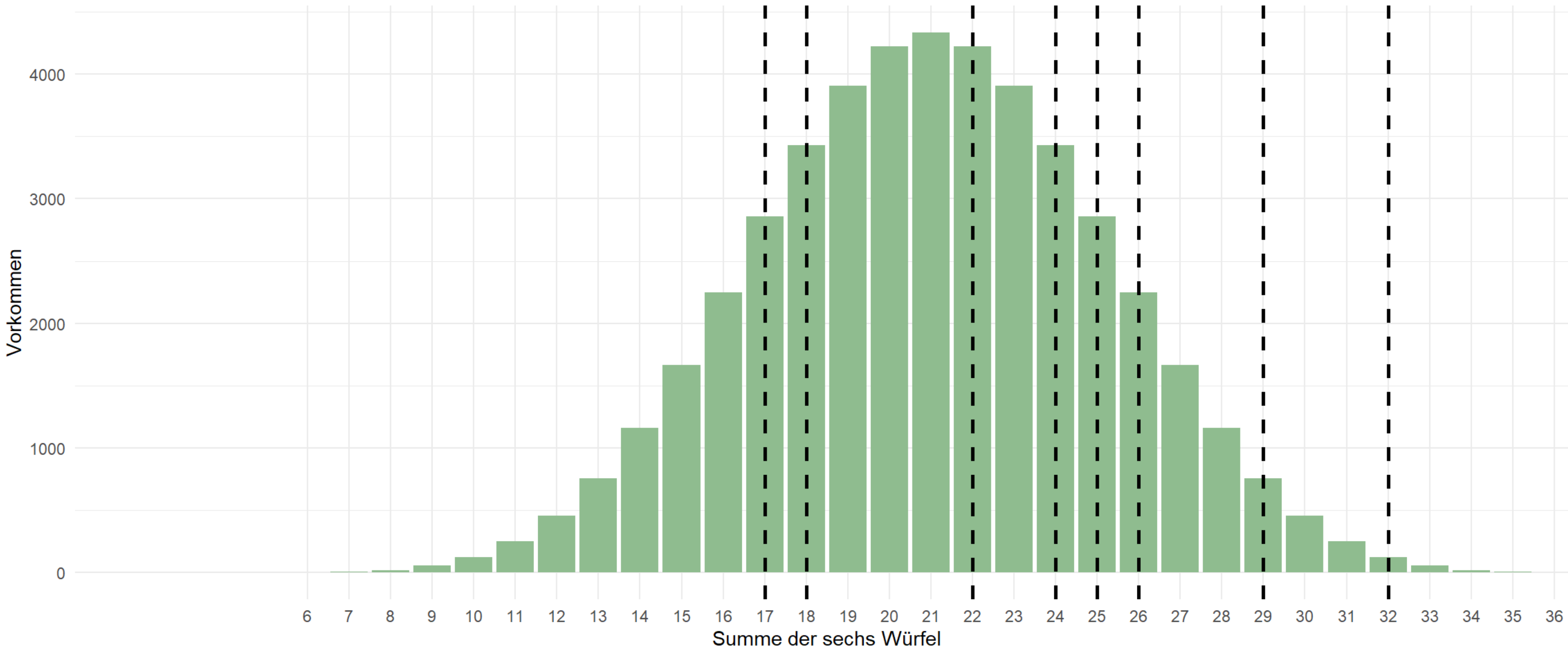


→ Würfeln: Ziehen einer Stichprobe aus der Population möglicher Summen

→ Teststatistiken können helfen zu schätzen ob Abweichung zufällig ist

→ Mit welcher Wahrscheinlichkeit stammt die Würfelsumme aus einer Grundgesamtheit mit der angenommenen Verteilung?

Gedankenexperiment: Mögliche Würfe



→ Würfeln als Ziehen von Stichproben

Umfragen und Wahrscheinlichkeit

→ Wie übertragen wir das auf Umfragen?

Population und Stichprobe

- typischerweise sind wir an einem sogenannten **Parameter** der Bevölkerung interessiert
  - z.B.: Anteil der Unterstützer:innen von Kandidat:in A an Wählenden
  - aber: im Gegensatz zum Würfeln kennen wir diesen Wert nicht!
- wir haben aber nur Zugang zu einer **Statistik** aus einer kleineren Stichprobe
  - z.B. Anteil der Unterstützer:innen unter den Wählenden in der Umfrage
- zwischen beiden Werten gibt es einen Unterschied, weil die Stichprobe ‘Rauschen’ / Verzerrungen enthält
  - *Stichprobenvariabilität* als Ursache

Population und Stichprobe

Stichprobenvariabilität: Beim Ziehen mehrerer Stichproben hat jede Stichprobe leicht andere Beobachtungen und damit andere statistische Messwerte

Bei kleineren Stichproben unterscheiden sich die Messwerte stärker, bei großen weniger stark

→ statistische Kennwerte sind ebenfalls verteilt und wir können über unsere Werte als Teil dieser Verteilung nachdenken

Population und Stichprobe

Zwei statistische Gesetze:

Gesetz der großen Zahlen: Die relative Häufigkeit eines Wertes näher sich der theoretischen Verteilung, wenn man viele Stichproben zieht

Zentraler Grenzwertsatz: Annäherung der Verteilung der Stichproben-Mittelwerte an *Normalverteilung*

→ Aus diesen Gesetzen können wir mithilfe von Statistik sogenannte Konfidenzintervalle (confidence intervals) berechnen, in denen sich ein Parameter wahrscheinlich (z.B. mit 95% Wahrscheinlichkeit) befindet

Was sie daraus mitnehmen

- Unsere Umfragen sind Stichproben der Bevölkerung
- die von uns berechneten Kennwerte (Regressionskoeffizienten, Mittelwerte etc.) kommen aus einer Verteilung der in verschiedenen Stichproben messbaren Werte
  - → Konfidenzintervalle
- wir können die Werte entsprechend interpretieren

Beispiel: Regressionen

- Aus dem berechneten Effekt und seinem Standardfehler können wir berechnen ob Effekte auf einem bestimmten Niveau *statistisch signifikant* sind
- Nullhypothese: kein Effekt
  - im Fall einer Regression: kein Effekt der Variablen / statistisch nicht anders als null
  - im Fall eines Mittelwerts z.B. testen auf signifikante Gruppendifferenzen

Beispiel: Regression

```
Call:
lm(formula = ccrdprs ~ agea, data = ess8)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.9155 -1.6290  0.3512  2.1635  4.8947 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  6.0637563  0.0376406  161.10   <2e-16 ***
agea        -0.0098809  0.0007187  -13.75   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.719 on 41794 degrees of freedom
  (2591 Beobachtungen als fehlend gelöscht)
Multiple R-squared:  0.004502,  Adjusted R-squared:  0.004478 
F-statistic:   189 on 1 and 41794 DF,  p-value: < 2.2e-16
```

→ neben der **Schätzung des Koeffizienten** erhalten wir auch **Standardfehler** (*hier*: Std. Error) und **statistische Signifikanz** (anhand der t-Verteilung, *hier* aus PR(>|t|))

Längere Erklärung: Llaudet & Imai Kapitel zu Wahrscheinlichkeit & Signifikanz

Beispiel: Regression

```
Call:
lm(formula = ccrdprs ~ agea, data = ess8)

Residuals:
    Min      1Q  Median      3Q     Max
-5.9155 -1.6290  0.3512  2.1635  4.8947

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.0637563  0.0376406  161.10   <2e-16 ***
agea        -0.0098809  0.0007187  -13.75   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.719 on 41794 degrees of freedom
  (2591 Beobachtungen als fehlend gelöscht)
Multiple R-squared:  0.004502,	Adjusted R-squared:  0.004478
F-statistic:   189 on 1 and 41794 DF,  p-value: < 2.2e-16
```

Ablesen von Signifikanz auf vorher bestimmtem Signifikanzniveau

→ hier: statistisch signifikanter Effekt von Alter

Interpretationshilfe Regressionsoutput in R

Beispiel: Regression

↔ Gegenstück: nur manche Länder sind signifikant verschieden von Österreich (=Basis)

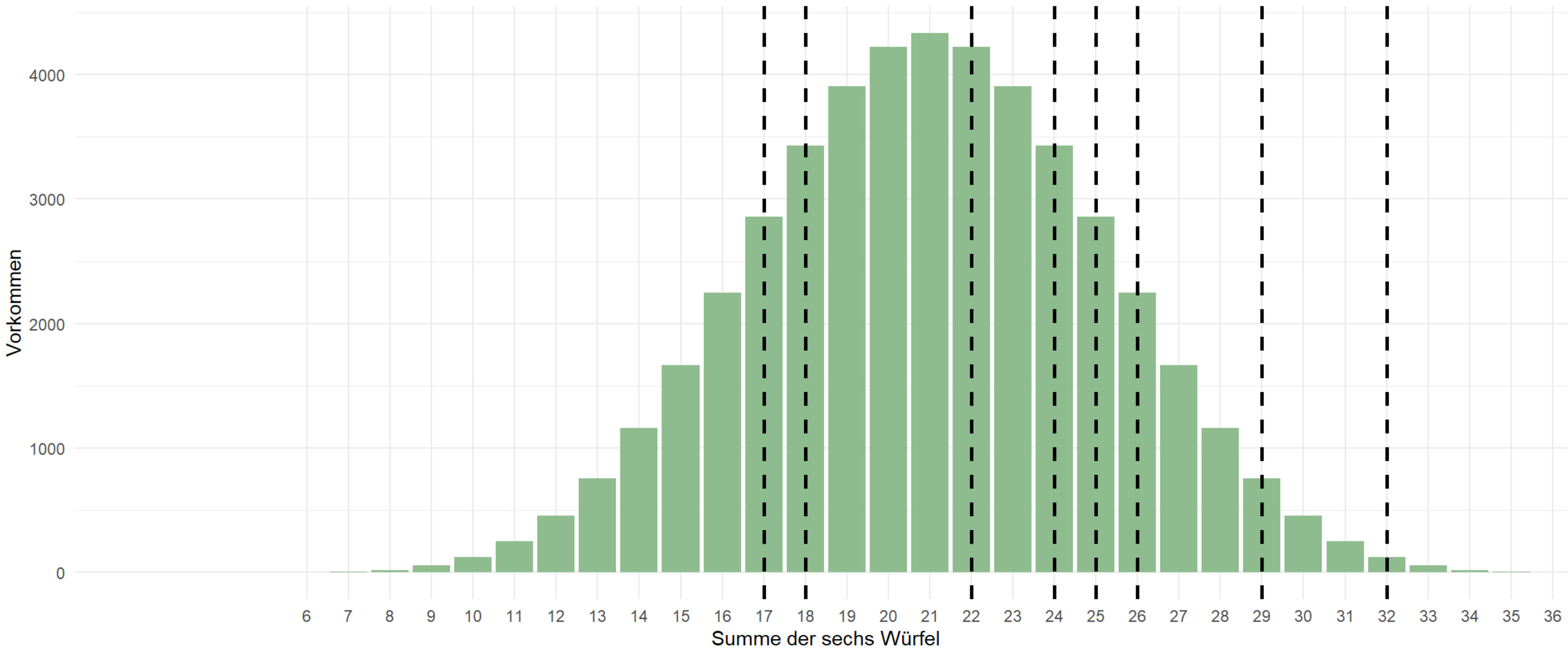
```
Call:
lm(formula = ccrdprs ~ cntry, data = ess8)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.9167 -1.5457  0.3607  1.7393  6.6552 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.91850    0.05842 101.305  < 2e-16 ***
cntryBE      0.07006    0.08457   0.828 0.407476    
cntryCH      0.94288    0.08837  10.670  < 2e-16 ***
cntryCZ     -2.57373    0.08051 -31.968  < 2e-16 ***
cntryDE      0.66932    0.07586   8.823  < 2e-16 ***
cntryEE     -1.66438    0.08230 -20.223  < 2e-16 ***
cntryES      0.01734    0.08346   0.208 0.835420    
cntryFI      0.62725    0.08287   7.569 3.84e-14 ***
cntryFR      0.99817    0.08145  12.255  < 2e-16 ***
cntryGB     -0.03923    0.08272  -0.474 0.635320    
cntryHU     -1.59293    0.08742 -18.222  < 2e-16 ***
cntryIE     -0.27916    0.07640  -3.654 0.000259 ***
```

Was sie daraus mitnehmen

- statistische Signifikanz ist nicht zwingend gleichbedeutend wie substanzielle Signifikanz
  - z.B. hängt statistische Signifikanz auch von der Größe der Stichprobe ab
- hierfür sollten wir vor allem die Relevanz des gezeigten Phänomens betrachten
  - z.B. Größe des Effekt
- → Statistik braucht unsere Interpretation

Was sie daraus mitnehmen

…und die Würfel waren natürlich gezinkt!



Nächste Woche: Hackathon

- kommen Sie gewappnet mit all Ihren Fragen
- Gelegenheit zum Fragen stellen, weiterarbeiten, diskutieren
- Vorbereitung für Präsentationen

Danke

…für ein spannendes Semester mit viel Beteiligung!