ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

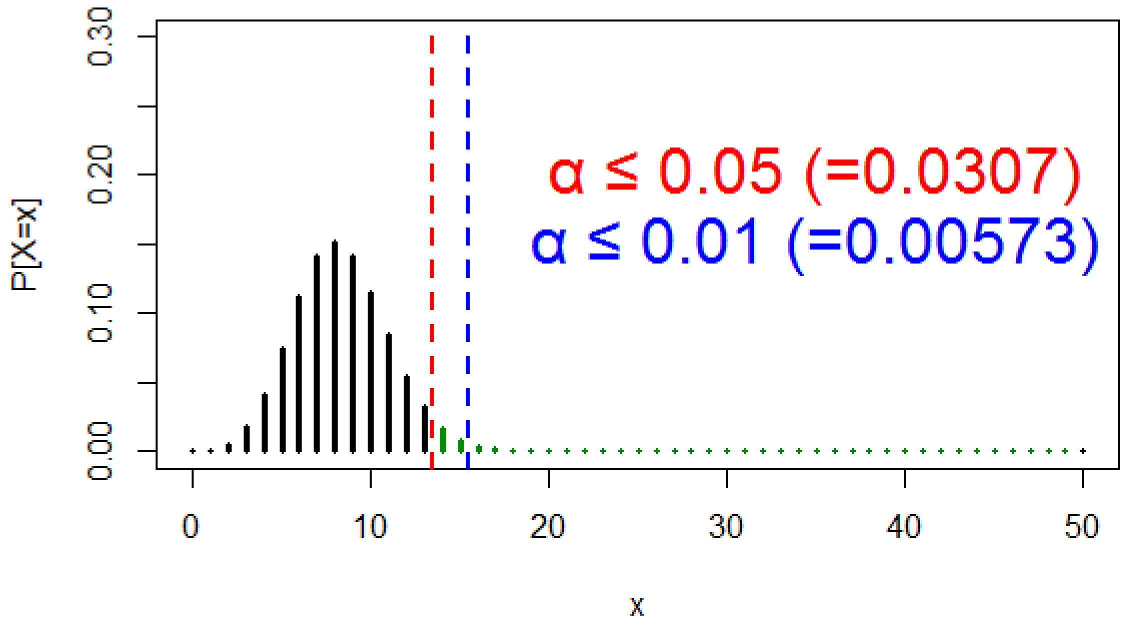
DMATH
Department of Mathematics

Zweiseitiger Binomialtest

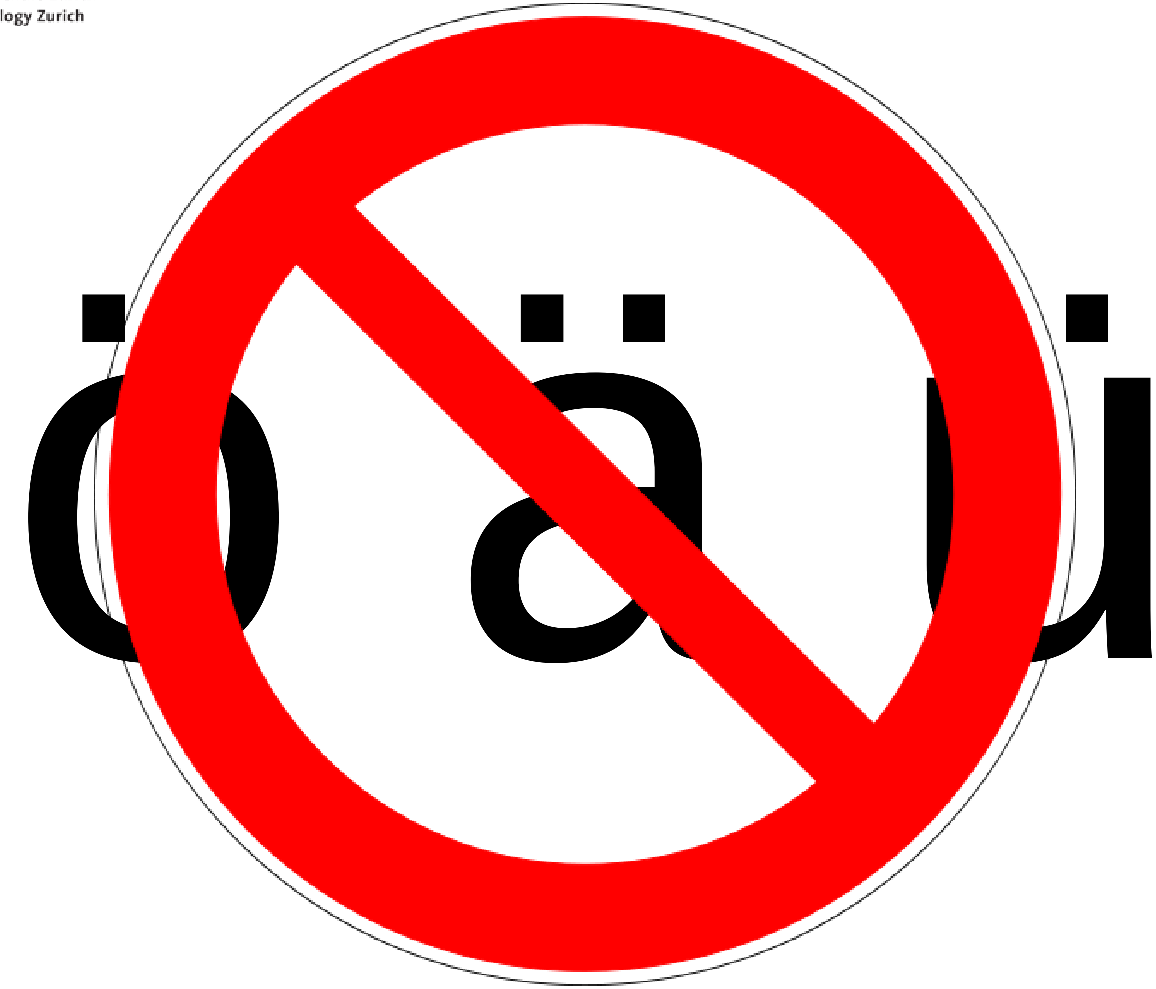
für D-UWIS, D-ERDW, D-USYS und D-HEST – SS15

Repetition: Macht

- Quiz 5 – Aufgabe 6 und 7



  - Verwerfungsbereich mit $\alpha \leq 0.05$ ist $K = [14, 15, \ldots, 25]$
  - Verwerfungsbereich mit $\alpha \leq 0.01$ ist $K = [16, 17, \ldots, 25]$



Do not put umlauts into folder names ... ever!

Lernziele heute

- Zweiseitiger Binomialtest
- P-Wert
- Vertrauensintervall

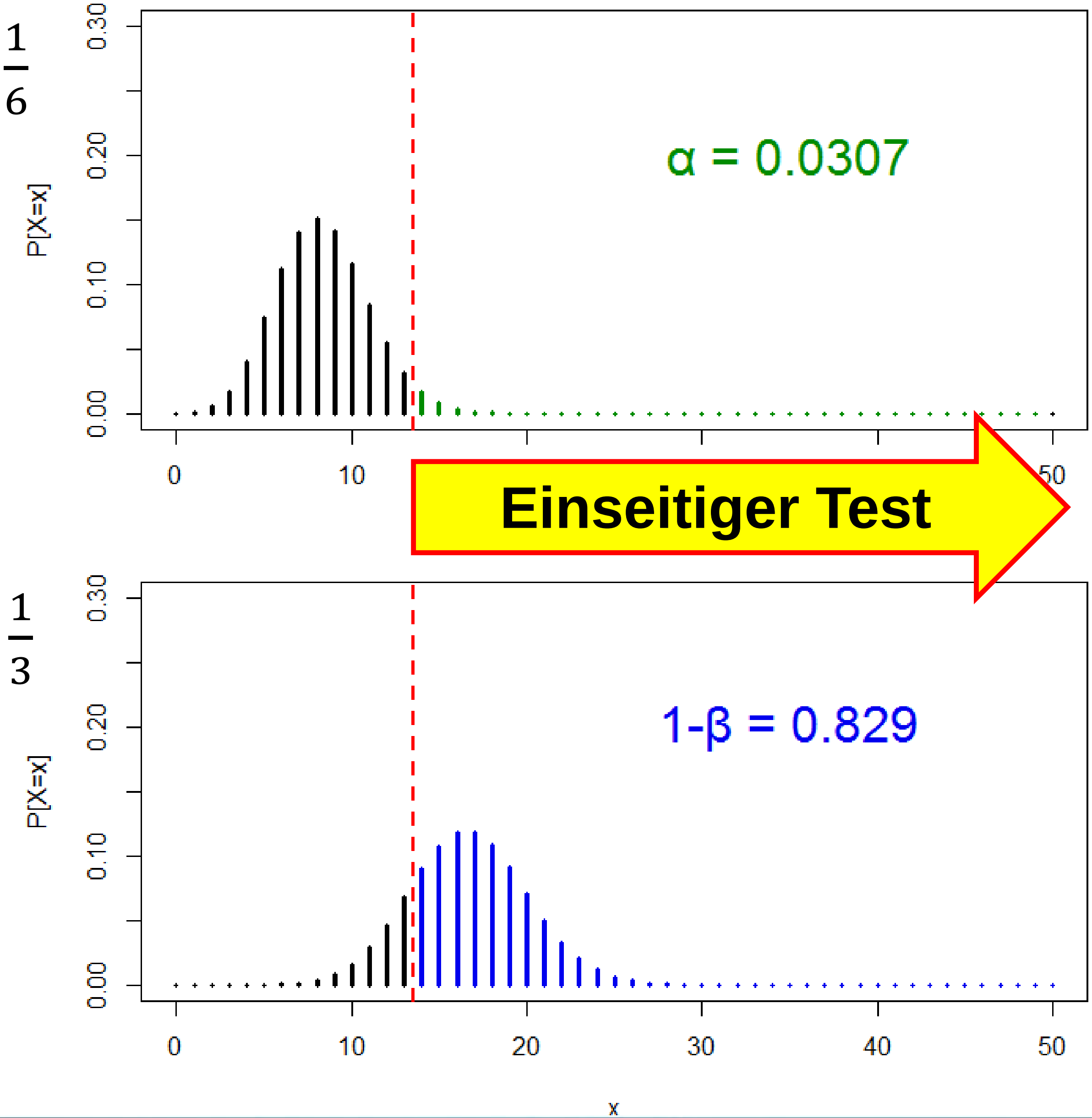
Hausaufgaben

- Skript: Kapitel 3.2.2 – 3.2.3 lessen
- Serie 6 lösen
- Quiz 6 bearbeiten

Wie war das nochmal mit dem Zauberwürfel?

$\mathcal{H}_0$richtig: $p_0 = \frac{1}{6}$

$\mathcal{H}_A$richtig: $p_A = \frac{1}{3}$




CASINO
PENNY SLOTS

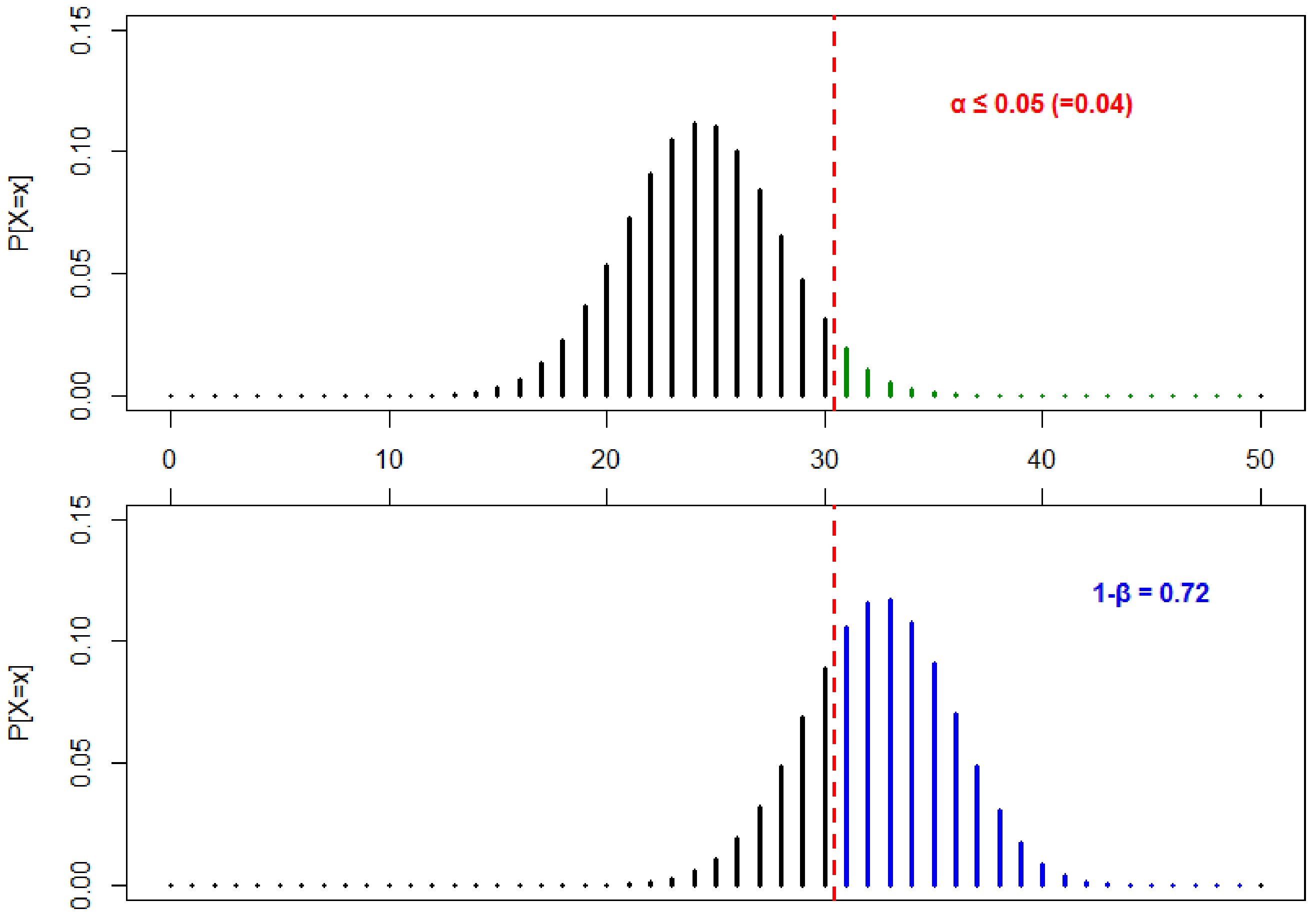
Roulette *(franz. Rädchen)*

Roulette *(franz. Rädchen)*

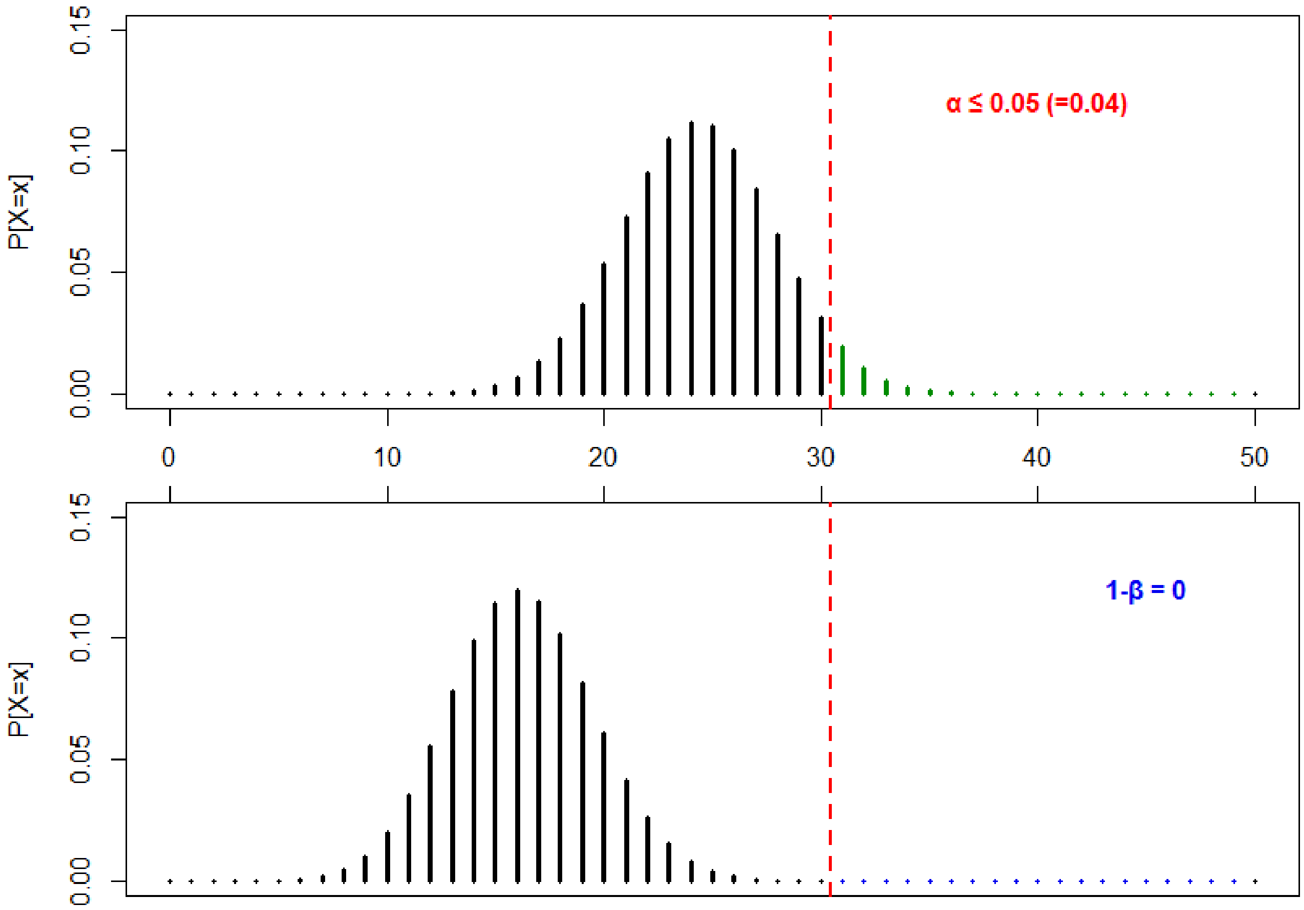
- 37 Felder (18 rote, 18 schwarze, 1 grünes)
- Verschiedene Wetttypen:
  - Straight up bet → $p = \frac{1}{37}$
  - Corner bet → $p = \frac{4}{37}$
  - Dozen bet → $p = \frac{12}{37}$
  - Even/odd bet → $p = \frac{18}{37}$
    (0 wird nicht als gerade gewertet)
  - Colour bet → $p = \frac{18}{37}$
- Wir spielen 50 mal rot,
  wenn die Roulette fair ist, dann gilt:

$$X \sim Bin(50, \frac{18}{37})$$

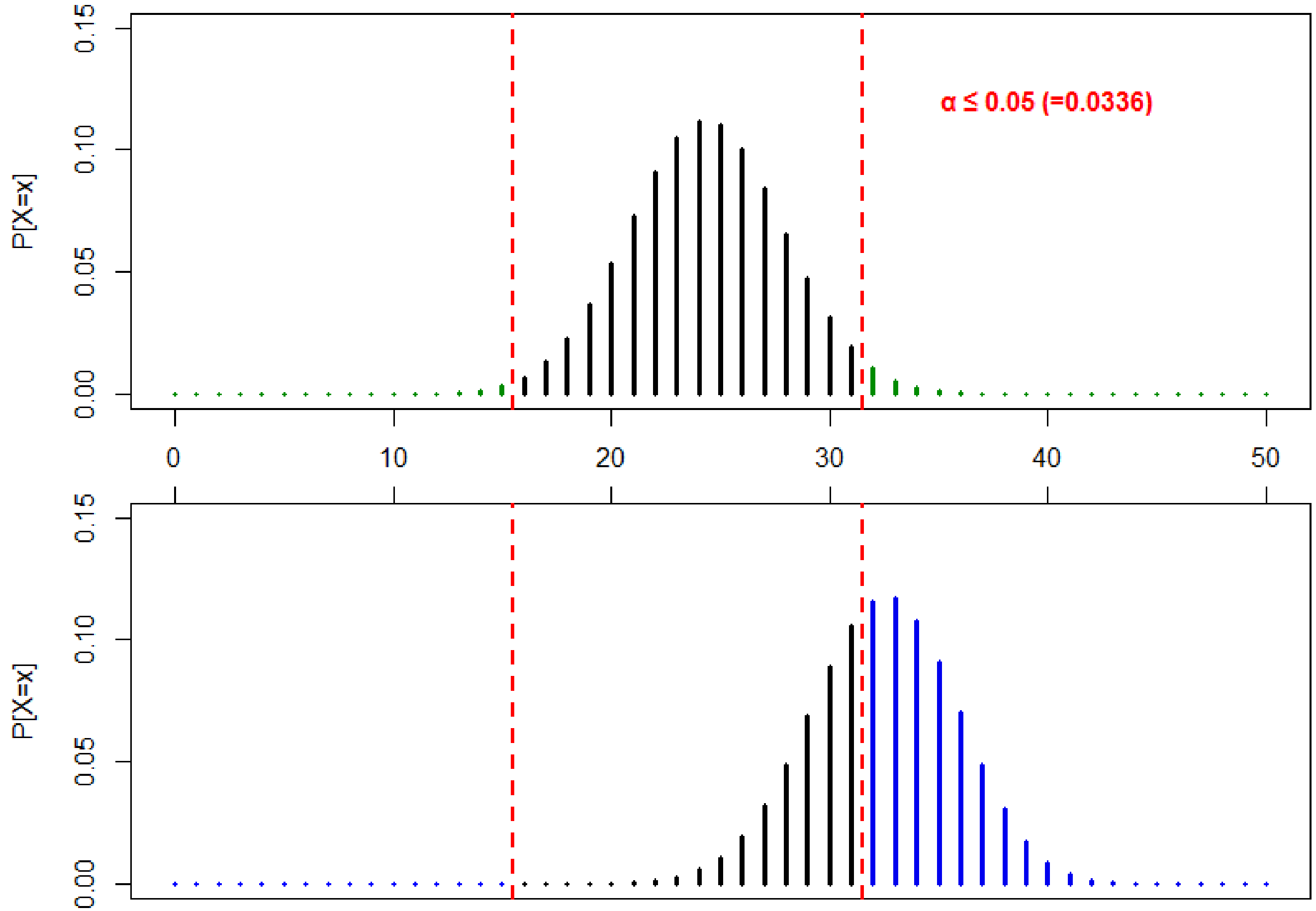
Einseitiger Test – zu viel rot



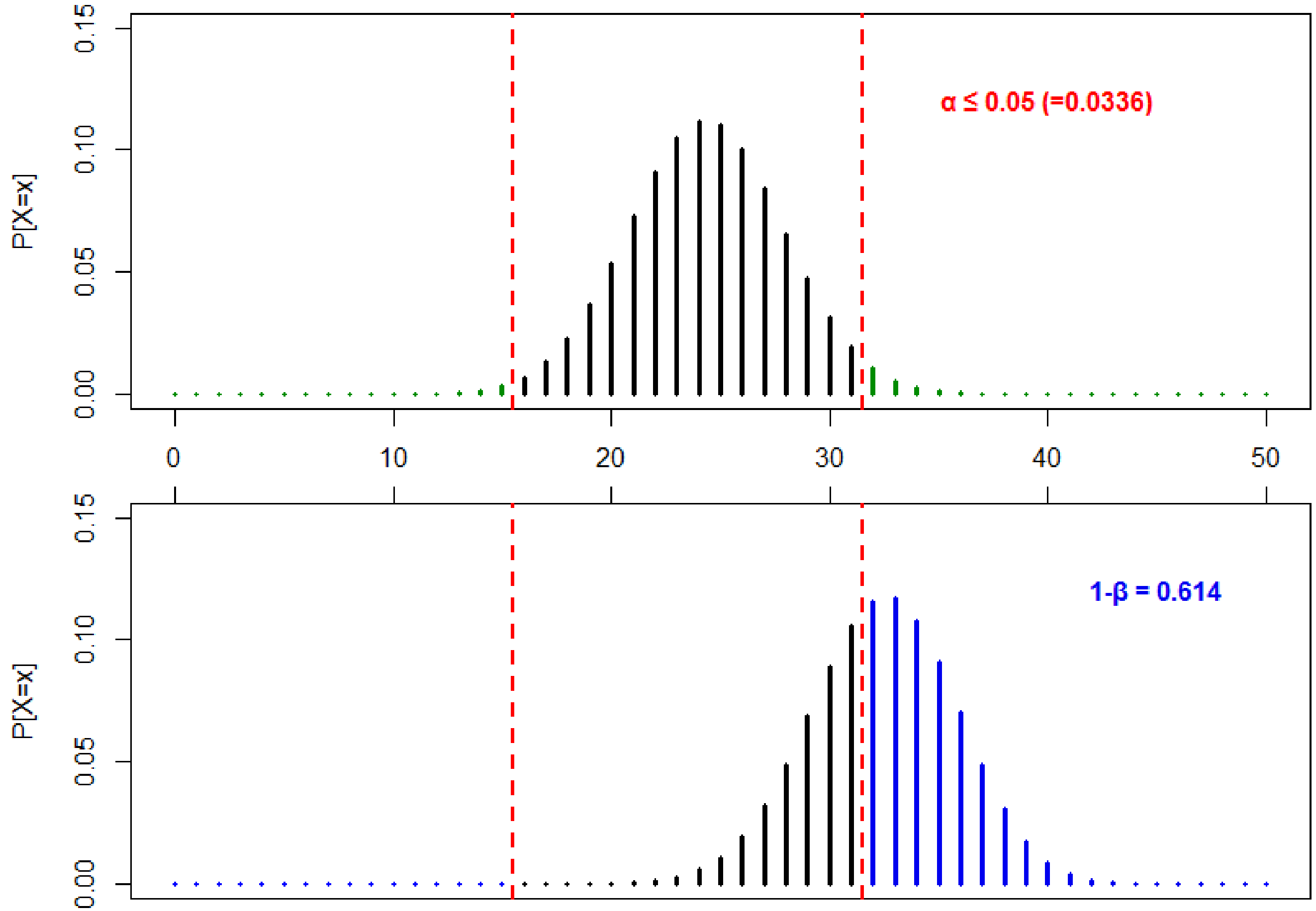
Einseitiger Test – zu viel schwarz



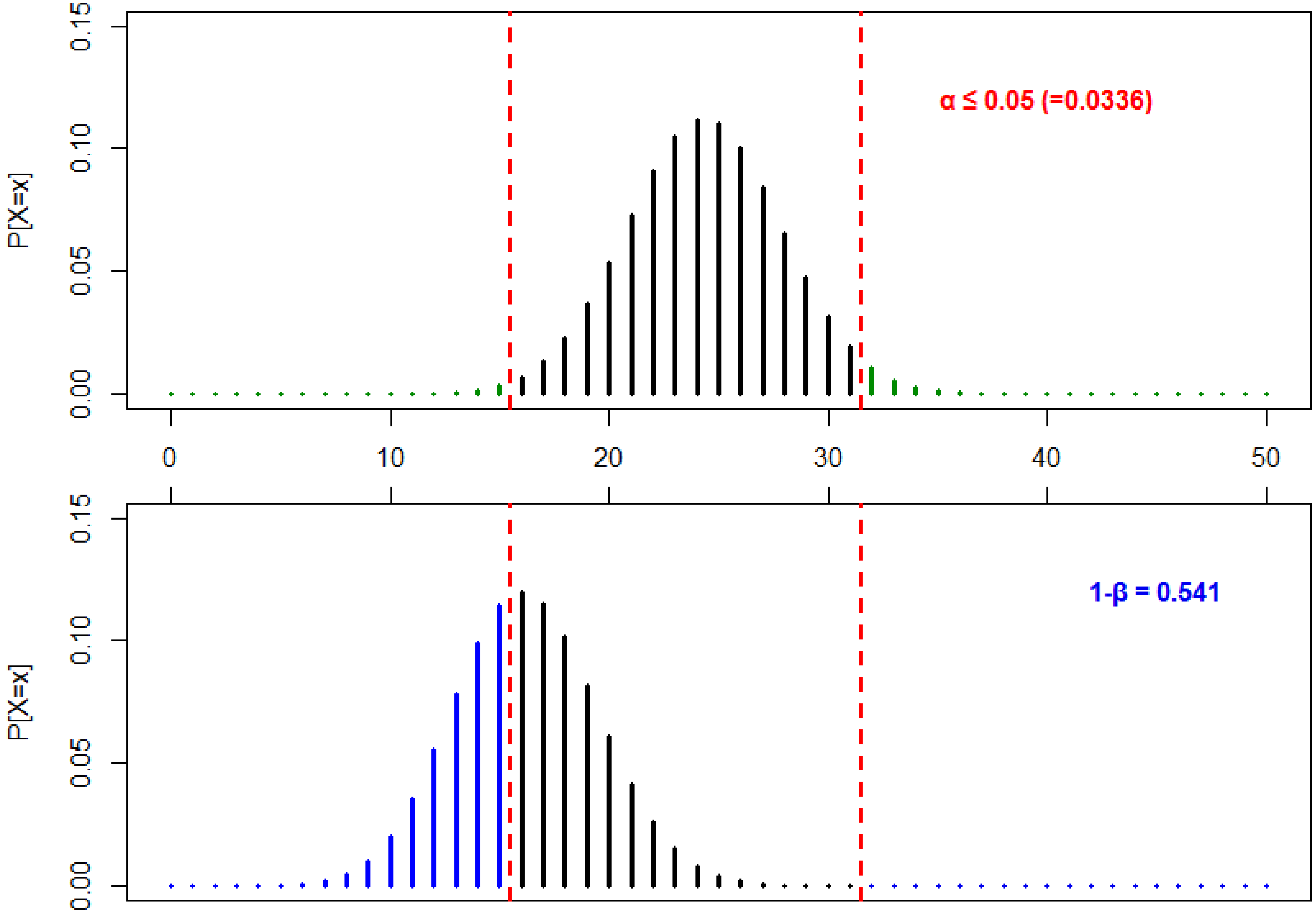
Zweiseitiger Test – zu viel rot



Zweiseitiger Test – zu viel rot



Zweiseitiger Test – zu viel schwarz



Zweiseitig versus Einseitig

Testart ($\mathcal{H}_0: \pi = \frac{18}{37}; \alpha = 0.05$)	**Macht**	
	$\pi_A = \frac{12}{37}$	$\pi_A = \frac{24}{37}$
Einseitig ($\mathcal{H}_A: \pi > \frac{18}{37}$)	0.00	0.72
Zweiseitig ($\mathcal{H}_A: \pi \neq \frac{18}{37}$)	0.54	0.61

Zweiseitig versus Einseitig

Einseitig

- Auf einer Seite blind
- Auf anderer Seite grosse Sehschärfe (grosse Macht)

Zweiseitig

- Sieht auf beiden Seiten
- Aber auf keiner besonders gut (kleine Macht)

Binomialtest – ganz formal…

1. Modell X
2. Hypothesen $\mathcal{H}_0$ und $\mathcal{H}_A$
3. Teststatistik T und Verteilung von T, falls $\mathcal{H}_0$ stimmt
4. Signifikanzniveau α
5. Verwerfungsbereich K der Teststatistik
6. Testentscheid

Binomialtest – Modell (1/6)

- X: Anzahl Erfolge bei n Versuchen
- $X \sim \mathrm{Bin}(n, \pi)$

Binomialtest – Hypothesen (2/6)

- Nullhypothese $\mathcal{H}_0: \pi = \pi_0$
- Alternativhypothese $\mathcal{H}_A$:
  - $\pi \neq \pi_0$ (zweiseitig)
  - $\pi > \pi_0$ (einseitig, nach oben)
  - $\pi < \pi_0$ (einseitig, nach unten)

Binomialtest – Teststatistik (3/6)

- T: Anzahl Treffer bei n Versuchen
- Verteilung von T, falls $\mathcal{H}_0$ stimmt:

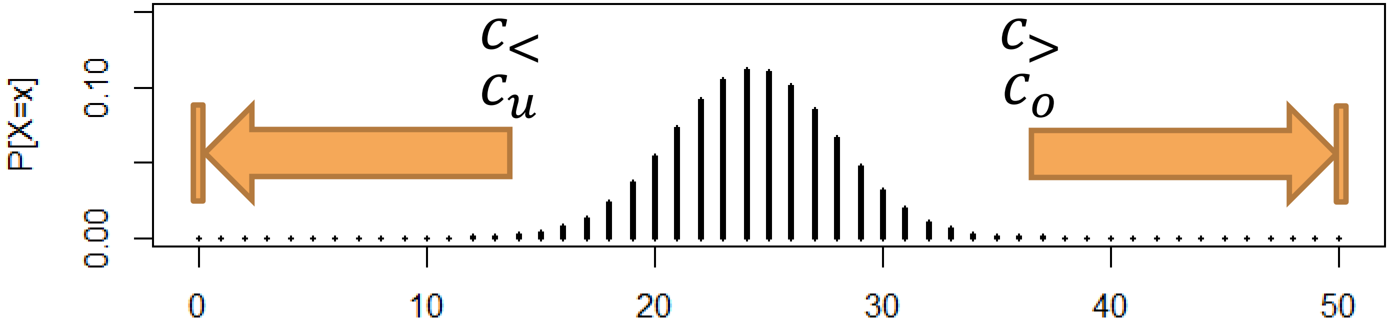
$$T \sim \mathrm{Bin}(n, \pi_0)$$

Binomialtest - Signifikanzniveau

- "*Wie unwahrscheinlich muss eine Beobachtung sein, damit wir die Nullhypothese verwerfen.*"
- Konvention, meist $\alpha = 0.05$

Binomialtest – Verwerfungsbereich (5/6)

- Form:
  - $K = [0, c_u] \cup [c_o, n]$, falls $\mathcal{H}_A\colon \pi \neq \pi_0$
  - $K = [c_>, n]$, falls $\mathcal{H}_A\colon \pi > \pi_0$
  - $K = [0, c_<]$, falls $\mathcal{H}_A\colon \pi < \pi_0$
- Grenzen (c's) werden bestimmt, sodass gilt:
  - $P[T \leq c_u] \approx \frac{\alpha}{2}, \quad P[T \geq c_o] \approx \frac{\alpha}{2}$
  - $P[T \geq c_>] \approx \alpha$
  - $P[T \leq c_<] \approx \alpha$



Binomialtest – Verwerfungsbereich (5/6)

- Form:
  - $K = [0, c_u] \cup [c_o, n]$, falls $\mathcal{H}_A: \pi \neq \pi_0$
  - $K = [c_>, n]$, falls $\mathcal{H}_A: \pi > \pi_0$
  - $K = [0, c_<]$, falls $\mathcal{H}_A: \pi < \pi_0$
- Normalapproximation für ein $\alpha = 0.05$
  - $c_u = n\pi_0 - 1.96\sqrt{n\pi_0(1-\pi_0)}$ abrunden
  - $c_o = n\pi_0 + 1.96\sqrt{n\pi_0(1-\pi_0)}$ aufrunden
  - $c_< = n\pi_0 + 1.64\sqrt{n\pi_0(1-\pi_0)}$ abrunden
  - $c_> = n\pi_0 - 1.64\sqrt{n\pi_0(1-\pi_0)}$ aufrunden
- …das klappt, wenn n gross und π nicht nahe bei 0 oder 1
  (genauere Faustregel im Skript)

Binomialtest – Testentscheid (6/6)

- Liegt der beobachtete Wert t von T in K?
  - Falls ja:
    $\mathcal{H}_0$ kann auf dem Signifikanzniveau α verworfen werden
  - Falls nein:
    $\mathcal{H}_0$ kann auf dem Signifikanzniveau α nicht verworfen werden

P-Wert

- Definition 1:
  - Der P-Wert ist das kleinste Signifikanzniveau, bei dem $\mathcal{H}_0$ gerade noch verworfen wird.
- Definition 2:
  - Der P-Wert ist die Wahrscheinlichkeit, diese Beobachtung oder eine noch extremere zu erhalten, wenn $\mathcal{H}_0$ wahr ist.

Def 1 & 2 sind äquivalent

P-Wert – Beispiel

W'keit unter $\mathcal{H}_0$

π ist die wahre W'keit

- Binomialtest, $n = 10$, $\mathcal{H}_0 : \pi_0 = 0.4$

t	0	1	2	3	4	5	6	7	8	9	10
$P[T = t]$	0.006	0.040	0.12	0.21	0.25	0.20	0.11	0.042	0.01	0.002	0.0001

- $\mathcal{H}_A : \boxed{\pi} > 0.4$
  Angenommen, $t = 9$ beobachtet:
  $p = P[T \geq t] = P[T = 9] + P[T = 10] \approx 0.002 + 0.0001 = 0.0021$
- $\mathcal{H}_A : \pi < 0.4$
  Angenommen, $t = 1$ beobachtet:
  $p = P[T \leq \mathrm{t}] = P[T = 0] + P[T = 1] \approx 0.006 + 0.04 = 0.046$
- $\mathcal{H}_A : \pi \neq 0.4$
  angenommen, $t = 1$ beobachtet:
  $$p = P[T \in \{0, 1, 8, 9, 10\}]$$
  $$\approx 0.006 + 0.04 + 0.01 + 0.002 + 0.0001 = 0.0581$$

Vertrauensintervall

Welche Werte von π sind mit den Daten vereinbar?

- Definition 1:
  - Die Werte von π, bei denen $\mathcal{H}_0$ nicht verworfen wird, sind ein $(1-\alpha)$-Vertrauensintervall (CI) für π.

- Bsp.:
  - $n = 50, x = 14 \Rightarrow 95\% - CI: [0.17, 0.41]$
    - Mit einem Signifikanzniveau von 5% sind alle diese π's mit den Daten vereinbar.

Vertrauensintervall - Normalapproximation

- Unter gewissen Umständen kann man mit der Normalverteilung ein CI approximieren
  - Normalapproximation:

$$\frac{x}{n} \pm 1.96 \cdot \sqrt{\frac{x}{n}\left(1-\frac{x}{n}\right)} \cdot \frac{1}{\sqrt{n}}$$

  - Bsp. $n = 50, x = 14 \Rightarrow 95\% - CI: [0.17, 0.41]$

$$\frac{14}{50} \pm 1.96 \cdot \sqrt{\frac{14}{50}\left(1-\frac{14}{50}\right) \cdot \frac{1}{\sqrt{50}}} = 0.28 \pm 0.12$$

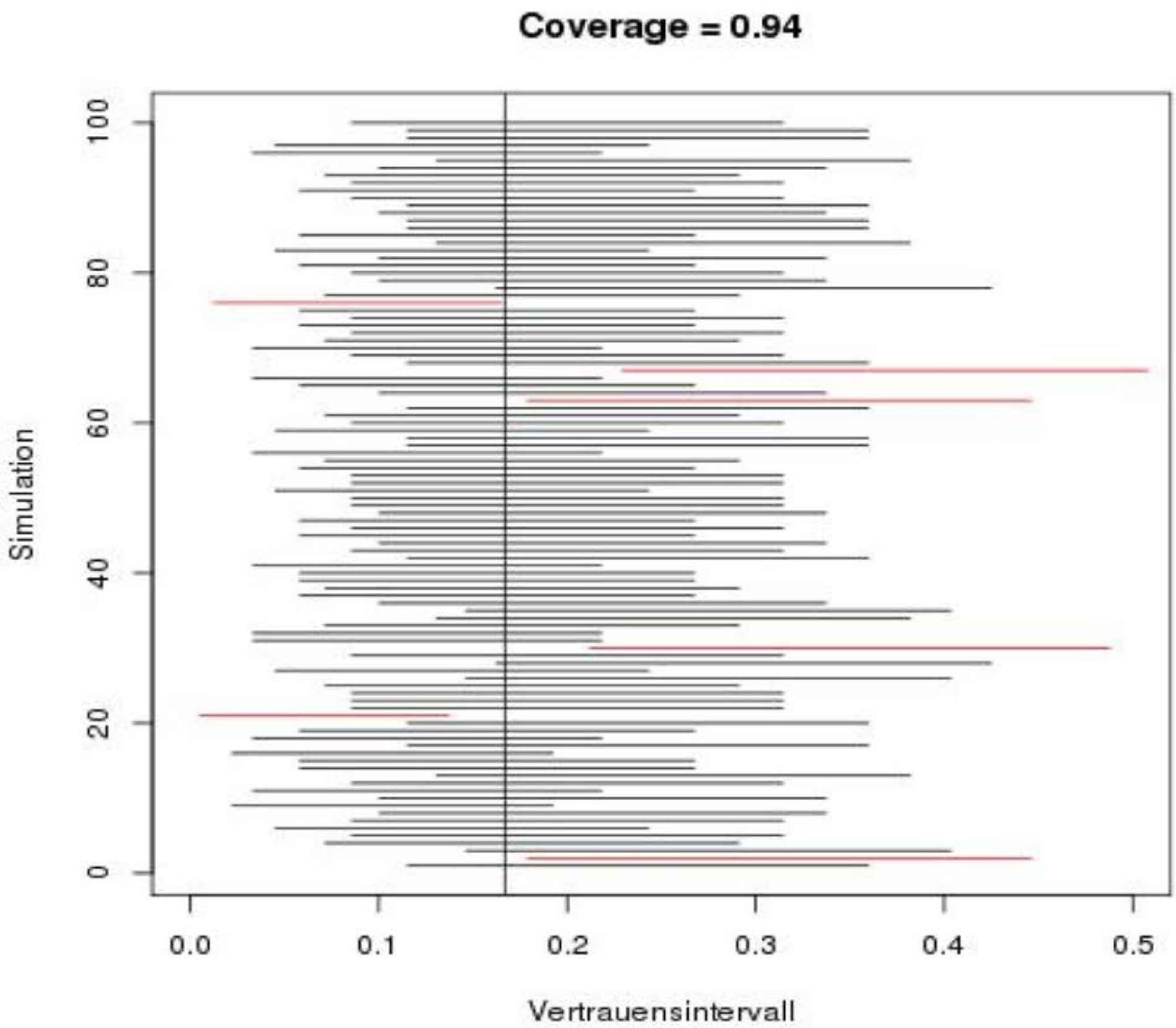
$$\Rightarrow [0.16, 0.40]$$

Vertrauensintervall

Welche Werte von π sind mit den Daten vereinbar?

- Definition 1:
  - Die Werte von π, bei denen $\mathcal{H}_0$ nicht verworfen wird, sind ein $(1-\alpha)$-Vertrauensintervall (CI) für π.

- Definition 2:
  - Ein $(1-\alpha)$-Vertrauensintervall enthält den wahren Parameter π mit Wahrscheinlichkeit $(1-\alpha)$.


Coverage = 0.94
Simulation
0
20
40
60
80
100
0.0
0.1
0.2
0.3
0.4
0.5
Vertrauensintervall

Zusammenfassung

- Zweiseitiger Binomialtest – einäugige Piraten mit Fernrohr
- P-Wert – diese oder eine noch extremere Beobachtung
- Vertrauensintervall – trifft den wahren Parameter mit Wahrscheinlichkeit $(1 - \alpha)$

Hausaufgaben

- Skript: Kapitel 3.2.2 – 3.2.3 lessen
- Serie 6 lösen
- Quiz 6 bearbeiten