Die Suche nach der Wahrheit: statistisches Testen

Vortrag – 10.11.2021

Referendariatsseminar

Dr. D. Uhlig
Professur Finanzmathematik
Fakultät für Mathematik

Stand 9. November 2021

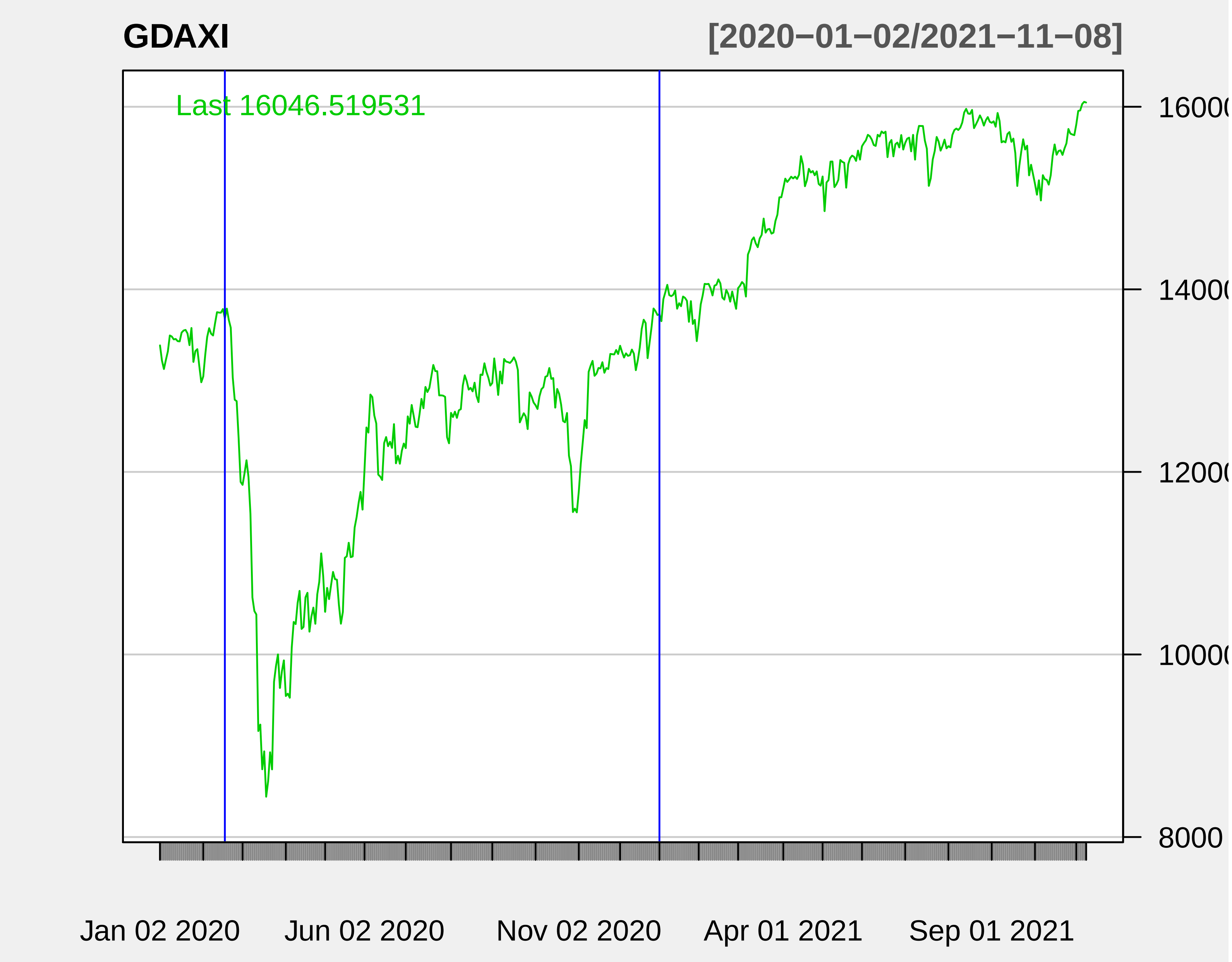
Aktuelles Thema

1. Was ist Statistik?
   - Was kann Statistik und wie?
   - Bereiche der Statistik
   - Schlussfolgerungen von der Stichprobe auf die Population

2. Statistische Tests

Die Suche nach der Wahrheit: Interessante Fragestel- lungen

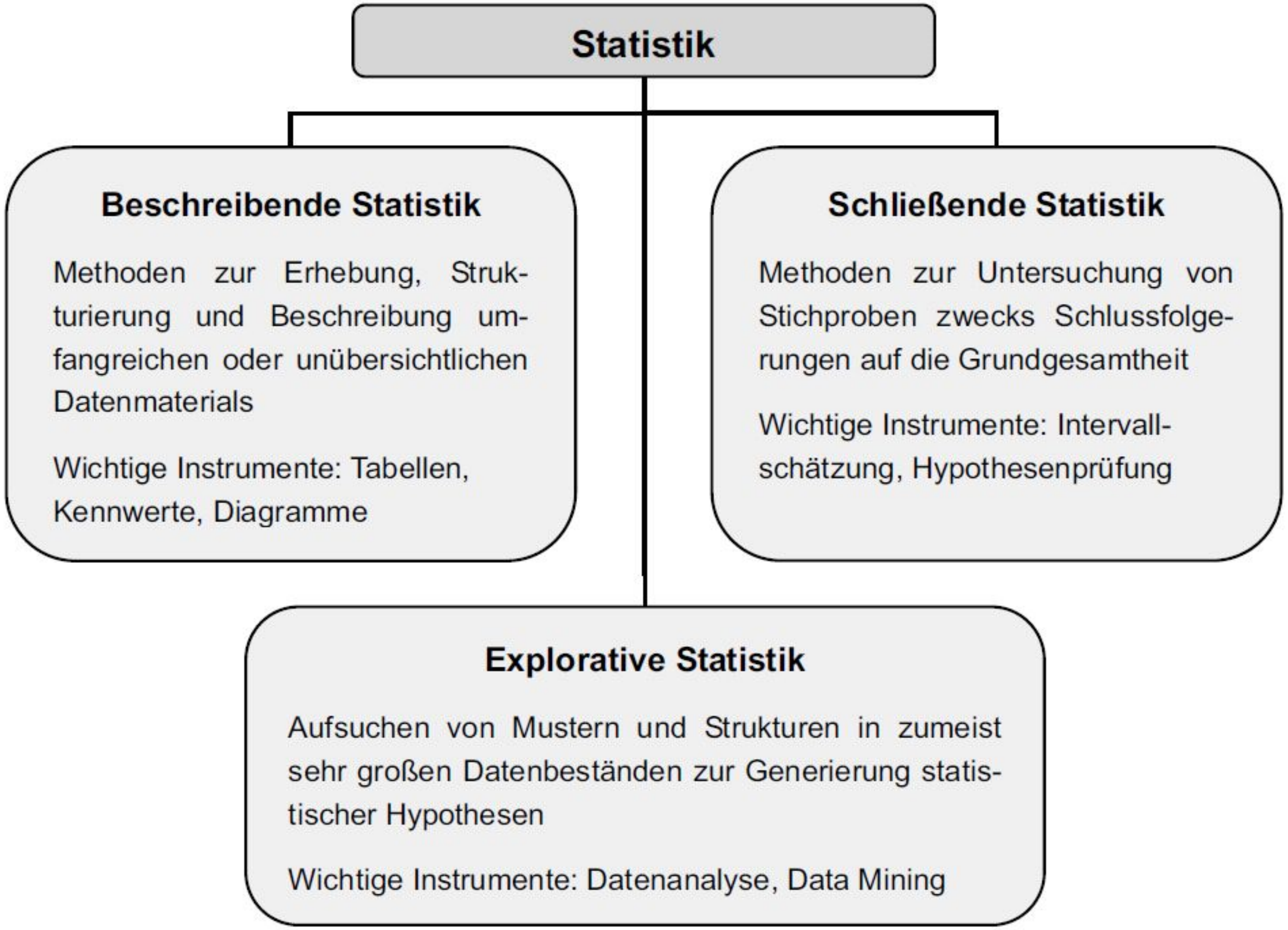
- Welche Mannschaft gewinnt die nächste Weltmeisterschaft?
- Bei welchem Stand schließt der DAX 2021 ab?
- Beeinflussen Politiker die Arbeitslosenquote?
- Ist ein Spiel fair?
- ...


GDAXI
[2020−01−02/2021−11−08]
Last 16046.519531
16000
14000
12000
10000
8000
Jan 02 2020
Jun 02 2020
Nov 02 2020
Apr 01 2021
Sep 01 2021

Tradition vs. Innovation

- old-fashioned?
  - `https://trends.google.com/trends/explore?date=all&geo=DE&q=Statistik,Mathematik`
  - `https://trends.google.com/trends/explore?date=all&q=statistic,mathematics`
- modern
  - Big Data `https://trends.google.com/trends/explore?date=all&q=Big%20Data`
  - Data Science `https://trends.google.com/trends/explore?date=all&q=Data%20Science`

Bereiche der Statistik

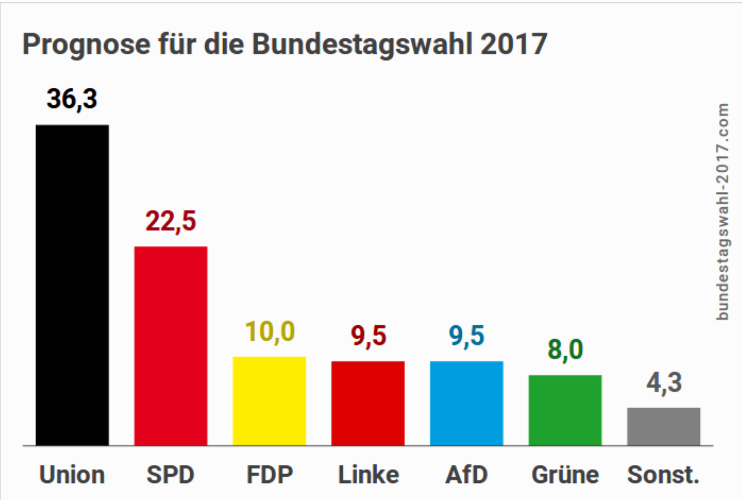


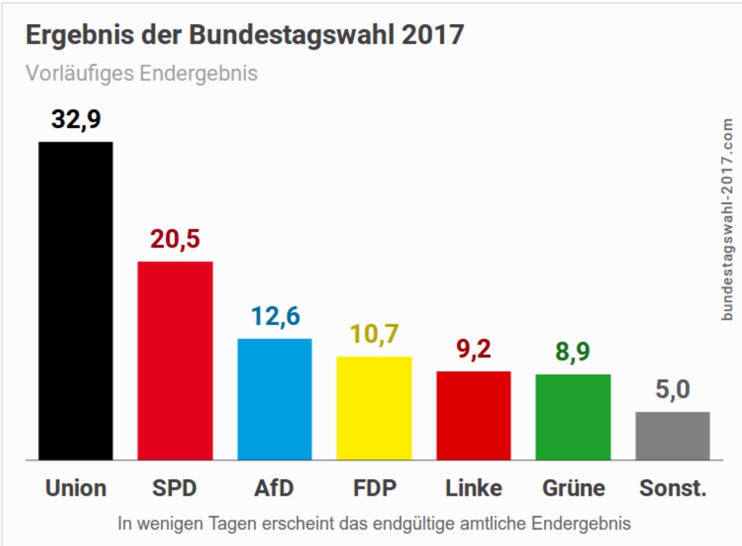
U. Bankhofer, J. Vogel, Datenanalyse und Statistik

Beispiel: Wahlprognose vs. Wahl

Frage: Ausgang der Wahl **vor** Wahltag?

- schlechtestes Vorhersage: keine Ahnung?
- zweitschlechteste Prognose: Wahlergebnisse der letzten Wahl 2013 $\curvearrowright$ zu ungenau
- Möglichkeit: Umfrage aller Wahlbeteiligten $\curvearrowright$ unmöglich, zu teuer
- Alternative: Sonntagsumfrage (Grafik: Forsa-Umfrage vom 19.09.2017 für Wahl am 24.09.2017) repräsentative Umfrage $\Rightarrow$ Schlussfolgerungen auf Gesamtpopulation $\triangleq$ induktive Statistik





Ausgangspunkt: die zu analysierende **konkrete Stichprobe** $(x_1, \ldots, x_n)$ wird als Realisierung des Zufallsvektors

$$\mathbf{X} = (X_1, X_2, \ldots, X_n)\,,$$

dessen Komponenten X_i Zufallsgrößen vom Typ i.i.d. und wie X verteilt sind, interpretiert

Stichprobenraum

Der Wertebereich $\mathcal{X}_n$ der mathematischen Stichprobe $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, d.h. die Menge aller möglichen Werte von $\mathbf{X}$, heißt **Stichprobenraum**.

Bemerkung: Die Annahme $X_1, \ldots, X_n$ sind vom Typ i.i.d. ist wesentlich

Unabhängigkeit: Beobachtungen beeinflussen sich nicht gegenseitig

identische Verteilung: es wird $n-$mal das gleiche Experiment durchgeführt und alle Merkmalsträger der Population haben die gleiche Chance in die Stichprobe aufgenommen zu werden

$\Rightarrow$ **repräsentative** Daten

Beispiel: Qualitätskontrolle

Einem Posten mit 100 000 Teilen wird eine konkrete Stichprobe vom Umfang $n = 100$ entnommen $\Rightarrow k = 7$ fehlerhafte Teile

1. Wie hoch ist der Ausschussanteil des gesamten Postens? $\Rightarrow$ Parameterschätzung
2. Der Hersteller behauptet der Ausschussanteil liegt unter 5%. Ist diese Behauptung zu halten, d.h. sind die beobachteten 7% nur zufallsbedingt oder muss man von einer höheren Ausschussrate ausgehen? $\Rightarrow$ Testverfahren

Aktuelles Thema

1. Was ist Statistik?

2. Statistische Tests
    - Vom Spiel zum Test
    - stetig vs. diskret
    - Allgemeines Vorgehen bei statistischen Tests
    - Gaußtest
    - t-Test
    - Parametertests für binomialverteilte Merkmale
    - Mittelwerttest für beliebig verteilte Merkmale

Würfelspiel: $W: 1,2,3,4,5 \mapsto 2€$, $6 \mapsto -10€$

Zusammenfassung nach 20 Würfen:

1	2	3	4	5	6
2	1	1	0	1	15

Frage 1: Spielen Sie nochmals mit mir?

A Ja, denn ich hatte einfach Pech, beim nächsten Spiel ist das Glück wieder auf meiner Seite...

B Nein, Glücksspiel ist generell nichts für mich...

C Nein, das war kein faires Spiel, hier stimmt etwas nicht...

D Ja, das Spiel ist fair, denn der Erwartungswert des Gewinns beträgt Null

`https://tweedback.de/`

Würfelspiel: $W\,:\;1,2,3,4,5\;\mapsto\;2$€, $6\;\mapsto\;-10$€

Zusammenfassung nach 20 Würfen:

1	2	3	4	5	6
2	1	1	0	1	15

Frage 2: Wie groß ist die Wahrscheinlichkeit, dass bei einem fairen Würfel aus Ihrer Sicht solch ein Ergebnis oder sogar ein schlechteres realisiert wird?

$X \triangleq$ Anzahl gewürfelter Sechsen

A Um das zu berechnen fehlen mir wichtige Verteilungsinformationen ...

B $P(X \geq 15) = \sum_{k=15}^{20} \binom{20}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{20-k} = 1.412014e-08$

C $P(X = 15) = \binom{20}{15} \left(\frac{1}{6}\right)^{15} \left(\frac{5}{6}\right)^5 = 1.325161e-08$

D $P(X \geq 15) = 1 - F(14) = 1 - \sum_{k=0}^{14} \binom{20}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{20-k} = 1.412014e-08$

Würfelspiel: $W\,:\,1,2,3,4,5 \mapsto 2€$, $6 \mapsto -10€$

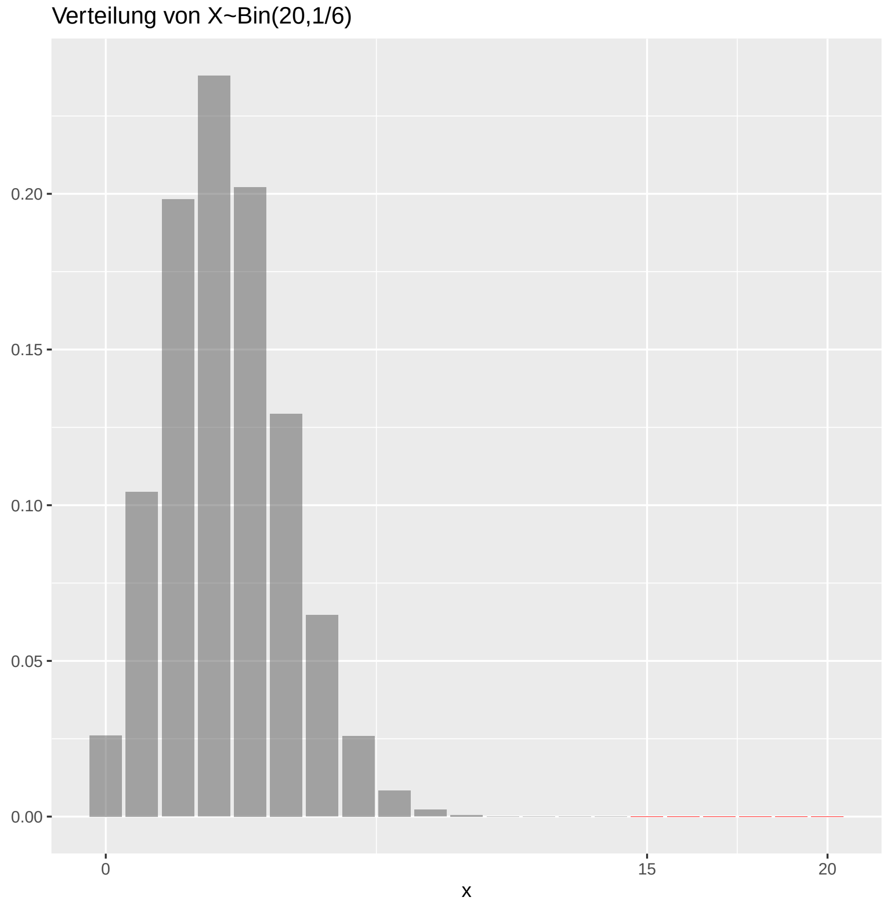
Zusammenfassung nach 20 Würfen:

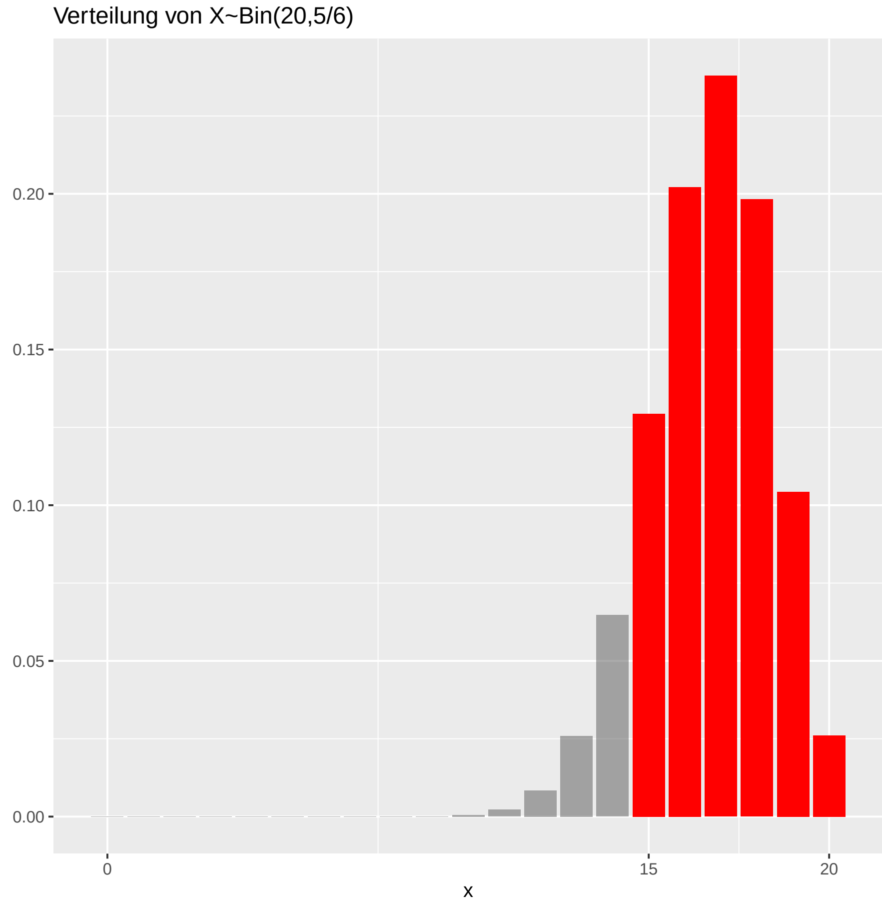
1	2	3	4	5	6
2	1	1	0	1	15

Frage 3: Berechnen Sie nochmals diese Wahrscheinlichkeit, falls bei dem verwendeten Würfel die Wahrscheinlichkeit für eine Sechs bei $5/6$ liegt.

$X \triangleq$ Anzahl gewürfelter Sechsen

- A Um das zu berechnen fehlen mir wichtige Verteilungsinformationen ...
- B $P(X \geq 15) = \sum_{k=15}^{20} \binom{20}{k} \left(\frac{5}{6}\right)^k \left(\frac{1}{6}\right)^{20-k} = 0.8981595$
- C $P(X = 15) = \binom{20}{15} \left(\frac{5}{6}\right)^{15} \left(\frac{1}{6}\right)^5 = 0.1294103$
- D $P(X \geq 15) = 1 - F(14) = 1 - \sum_{k=0}^{14} \binom{20}{k} \left(\frac{5}{6}\right)^k \left(\frac{1}{6}\right)^{20-k} = 0.8981595$



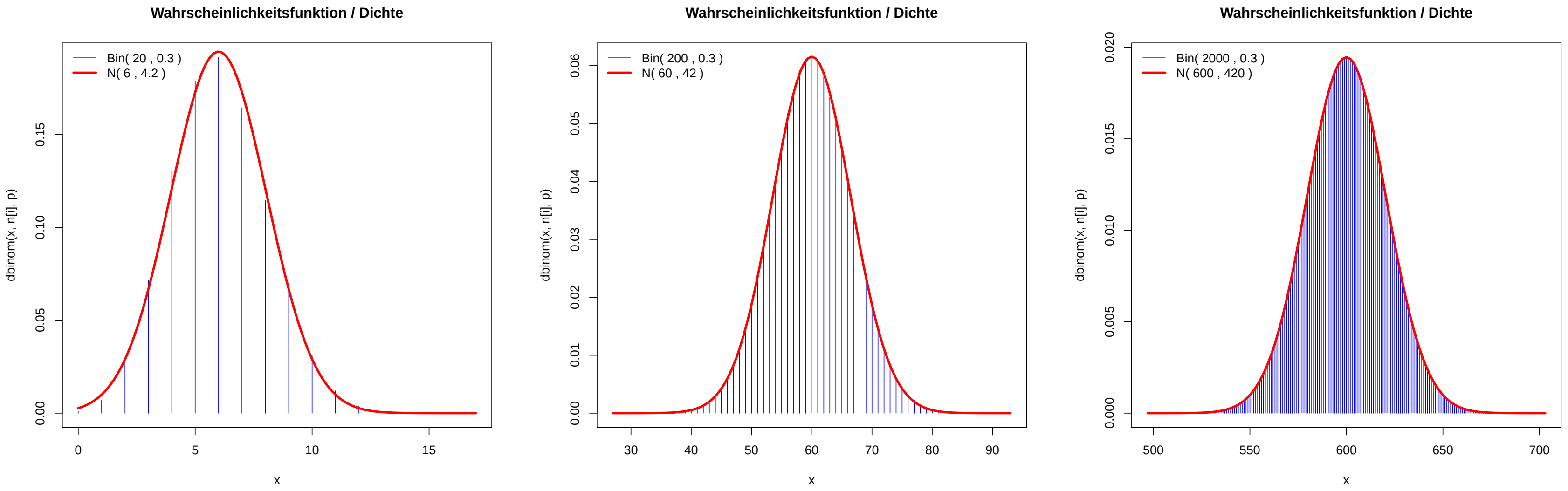


Tipp: `https://www.geogebra.org/classic#probability`

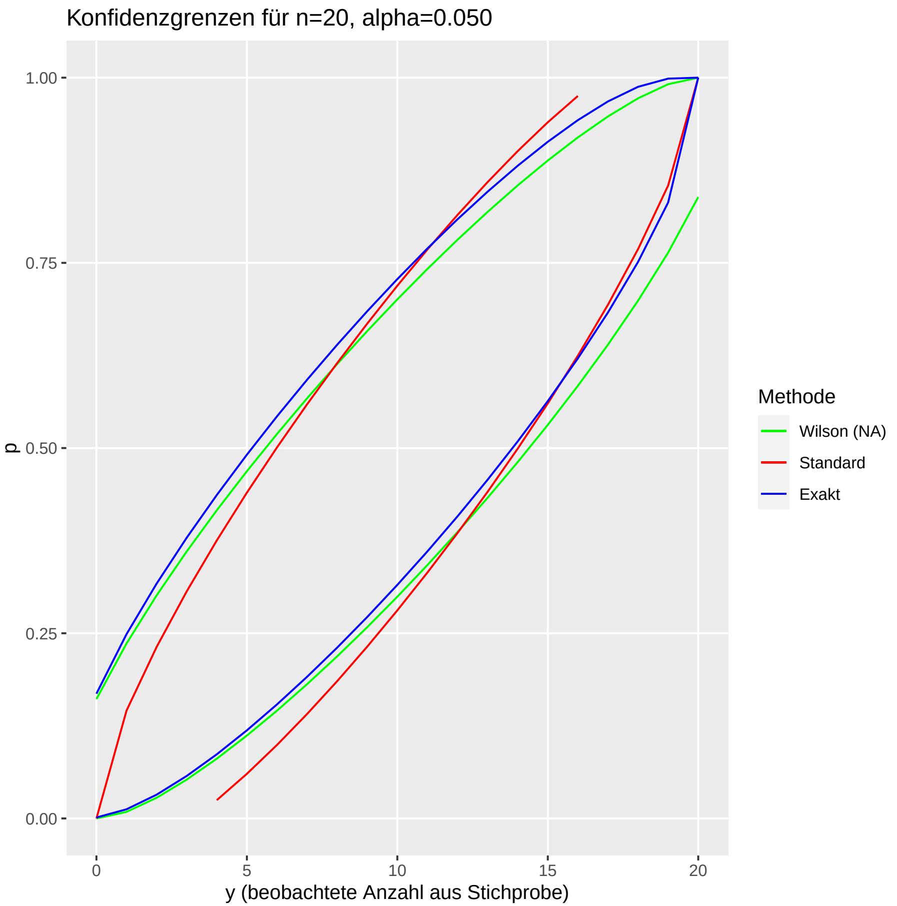
Entscheidungsmöglichkeiten

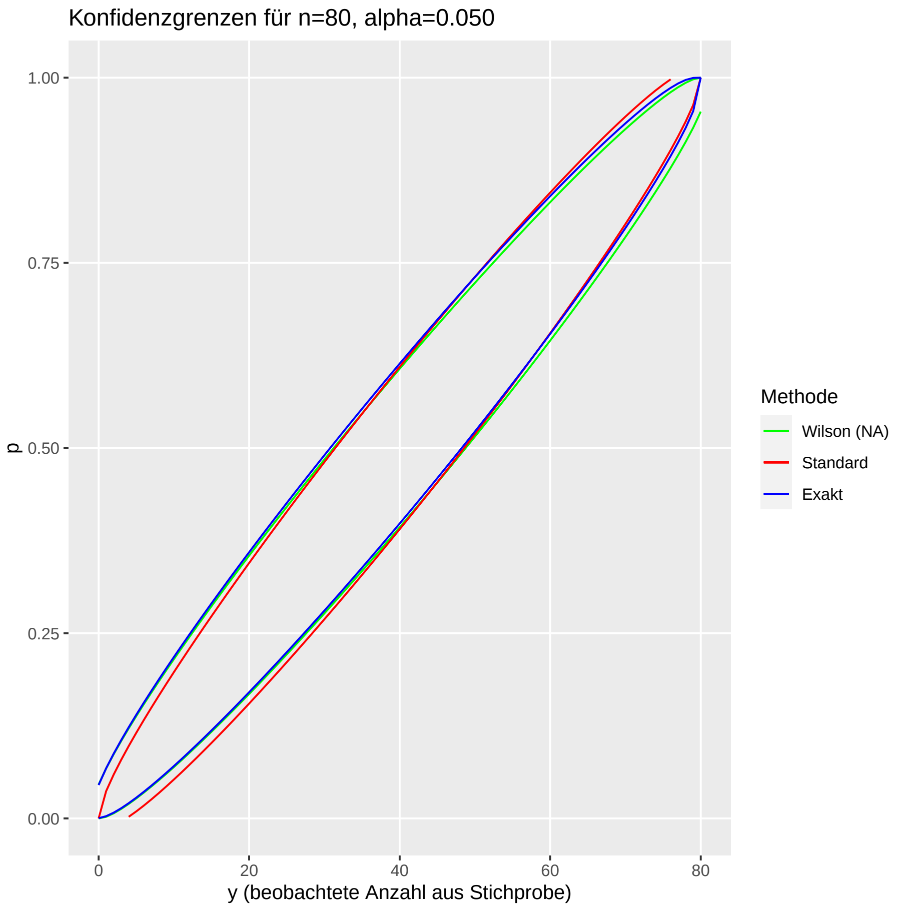
	weiterspielen (fairer Würfel)	aufhören (unfairer Würfel)
fairer Würfel	✓	↯
unfairer Würfel	↯	✓

Normalapproximation der Binomalverteilung



Exkurs: Konfidenzintervalle für Parameter p von $X \sim \mathrm{Bin}(n, p)$





Statistische Tests: Prüfen statistischer Hypothesen

- Ziel: Hypothese über die tatsächliche Verteilung eines Merkmals X einer Grundsgesamtheit anhand der Daten einer Stichprobe zu verwerfen oder als möglich anzusehen
- Parametertest: Verteilung $F(x, \theta) = P(X \leq x|\theta)$ ist bekannt, Aussagen über unbekannten Parameter θ werden überprüft
- parameterfreie Tests: Aussagen über die Art der Verteilung werden überprüft
- Untersuchung, ob es signifikante Abweichungen von der Hypothese gibt $\Rightarrow$ Signifikanztests

Durchführung statistischer Tests

Gegeben: Merkmal X mit Verteilungsfunktion F, Stichprobe $\mathbf{X} = (X_1, \dots, X_n)$, konkrete Stichprobe $\mathbf{x} = (x_1, \dots, x_n)$

1. Aufstellen einer Hypothese $H_0 : F = F_0$ (**Nullhypothese**, $F_0 \triangleq$ hypothetische Verteilung)
   Bsp: der Parameter μ von $X \sim \mathcal{N}(\mu, \sigma^2)$ hat den Wert $\mu = \mu_0$ (Parametertest)
2. Konstruktion einer **Testgröße**

$$T = T(\mathbf{X}),$$

   die Unterschiede zwischen der hypothetischen Verteilung F_0 und der tatsächlichen Verteilung F widerspiegelt und deren Verteilung bekannt ist, falls H_0 richtig ist
   Bsp. Würfelspiel (Anzahl 6er): $T = \sum_{i=1}^{n} X_i \sim \mathrm{Bin}(n, p)$

3. Konstruktion eines **kritischen Bereichs** $K^\star$ derart, dass zu gegebenen $\alpha \in (0,1)$ gilt:

$$P\left(T \in K^\star\right) = P_{H_0}\left(T \in K^\star\right) = \alpha$$

4. für eine konkrete Stichprobe $\mathbf{x} = (x_1, \ldots, x_n)$ wird der konkrete Testwert $t = T(\mathbf{x})$ berechnet
5. Entscheidungsregel:
   - Fall $t \in K^\star$: Ablehnen der Nullhypothese H_0 (Test ist **signifikant**)
   - Fall $t \notin K^\star$: auf der Basis des durchgeführten Tests ist nichts gegen die aufgestellte Hypothese einzuwenden (Test ist **nicht signifikant**)

Bemerkungen

- Bei Parametertests wird ein gewisser Verteilungstyp vorausgesetzt und H_0 beschränkt sich auf die Wahl des Parameters $\theta = \theta_0$
- Die Konstruktion des kritischen Bereichs $K^\star$ ist von der **Alternativhypothese** H_1 abhängig, bei einem Parametertest für einen Parameter θ können beispielsweise
  - $H_0 : \theta = \theta_0$, $H_1 : \theta \neq \theta_0$
  - $H_0 : \theta \leq \theta_0$, $H_1 : \theta > \theta_0$
  - $H_0 : \theta \geq \theta_0$, $H_1 : \theta < \theta_0$

  betrachtet werden
- Die hypothetische Verteilung kann eine aus gewissen Erfahrungen vermutete Verteilung sein.
- Ein signifikanter Nachweis kann stets nur für Behauptungen der Alternativhypothese H_1 geführt werden.
- Testgrößen zu konstruieren ist im Allgemeinen schwierig (Verteilung von T muss bekannt sein, um $P(T \in K^\star)$ zu berechnen)

Bemerkungen

- die Wahrscheinlichkeit $\alpha = P\left(T \in K^{\star} \middle| H_0 \text{ ist richtig}\right)$ zur Konstruktion des kritischen Bereichs heißt **Irrtumswahrscheinlichkeit** und ist „klein“ zu wählen
  (typische Werte: $\alpha = 0.1$, $\alpha = 0.05$, $\alpha = 0.01$ oder $\alpha = 0.001$)
- Da das Ergebnis eines Parametertests nur auf Stichproben beruht, können die zwei folgenden Fehler auftreten:

Fehler 1. Art: Die Hypothese H_0 ist richtig, wird aber auf Grund der Stichprobe abgelehnt. Die Wahrscheinlichkeit für diesen Fehler beträgt α ($\triangleq$ Irrtumswahrscheinlichkeit).

Fehler 2. Art: Die Hypothese H_0 ist falsch, wird aber nicht abgelehnt, da die Stichprobe für H_0 spricht. Die Wahrscheinlichkeit für das Auftreten dieses Fehlers ($\beta-$Fehler) ist im Allgemeinen unbekannt.

Fehlentscheidungen beim Testen

	Test entscheidet sich für H_0	Test entscheidet sich für H_1
H_0 ist richtig	✓	Fehler 1. Art – α-Fehler
H_1 ist richtig	Fehler 2. Art – β-Fehler	✓

Analogie Gerichtsverfahren:

	Richter: Freispruch	Richter: Verurteilung
Angeklagter unschuldig	✓	Fehler (Fehlverurteilung)
Angeklagter schuldig	Fehler ($\triangleq$ Freispruch aus Mangel an Beweisen)	✓

Fehlentscheidungen beim Testen

	Test entscheidet sich für H_0	Test entscheidet sich für H_1
H_0 ist richtig	✓	Fehler 1. Art – α-Fehler
H_1 ist richtig	Fehler 2. Art – β-Fehler	✓

Analogie Corona-Test:

	Test: negativ	Test: positiv
Patient nicht infiziert	✓	Fehler (falsch-positiv)
Patient infiziert	Fehler (falsch-negativ)	✓

Gaußtest: Hypothesen über μ bei bekanntem σ^2 für $X \sim \mathcal{N}(\mu, \sigma^2)$

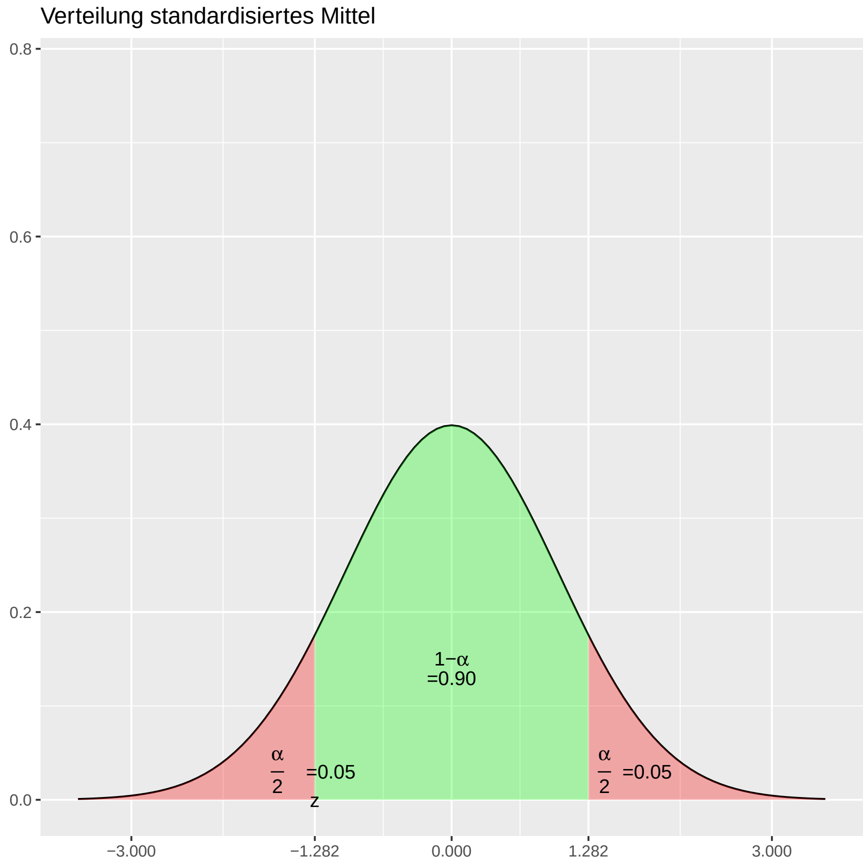
1. Nullhypothese H_0: $\mu = \mu_0$,

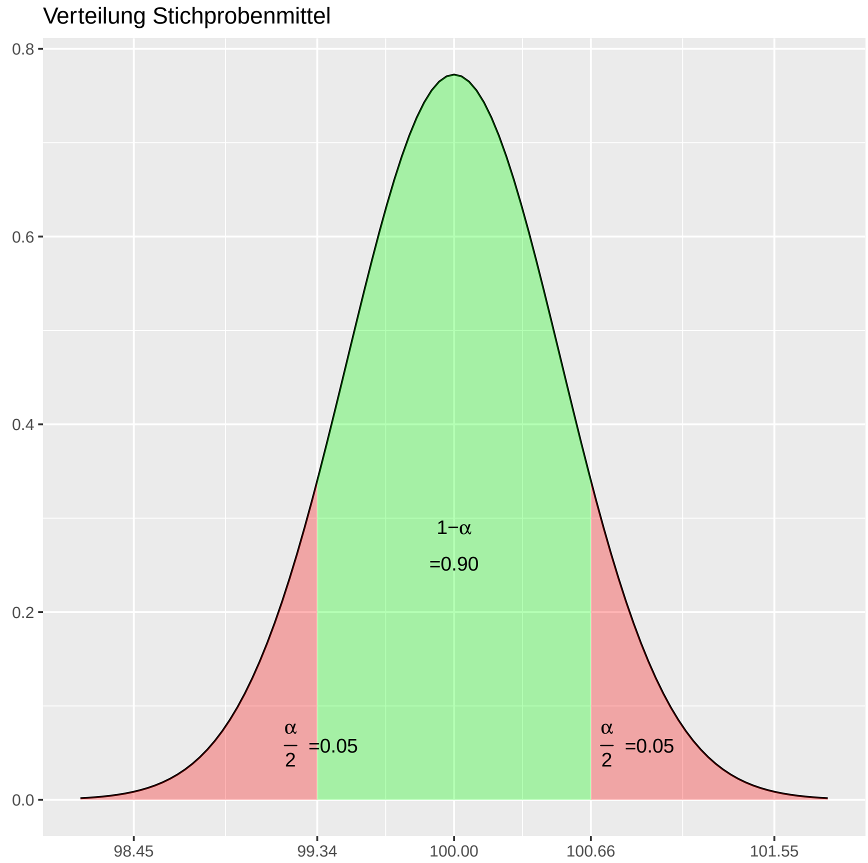
   Alternativhypothese $H_1 : \ \mu \neq \mu_0$
2. Konstruktion Testgröße:
   - falls tatsächliche Verteilung von hypothetischer Verteilung abweicht, unterscheiden sich das arithmetische Mittel $\overline{X}_n$ und der hypothetische Erwartungswert μ_0 deutlich
   - wissen $\overline{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i \sim \mathcal{N}(\mu_0, \frac{\sigma^2}{n})$, wenn H_0 stimmt

   betrachten daher als Testgröße

$$T(\mathbf{X}) = \frac{\overline{X}_n - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} = \sqrt{n}\frac{\overline{X}_n - \mu_0}{\sigma} \sim \mathcal{N}(0,1)$$

$$H_0 : \mu = \mu_0 = 100$$


Verteilung standardisiertes Mittel
0.8
0.6
0.4
0.2
0.0
1−α
=0.90
α/2 =0.05
α/2 =0.05
z
−3.000
−1.282
0.000
1.282
3.000


Verteilung Stichprobenmittel
0.8
0.6
0.4
0.2
0.0
1−α
=0.90
α/2 =0.05
α/2 =0.05
98.45
99.34
100.00
100.66
101.55

Gaußtest: Hypothesen über μ bei bekanntem σ^2

3. Kritischer Bereich:
   - sowohl positive als auch negative Abweichungen zwischen $\overline{X}_n$ und μ_0 sind kritisch
   - zu kleine und zu große Werte der Testgröße sprechen gegen die Nullhypothese $H_0 \Rightarrow$ zweiseitige Fragestellung
   - bestimmen zu gegebenen $\alpha \in (0,1)$ Quantil der Standardnormalverteilung $z_{1-\frac{\alpha}{2}}$ der Ordnung $1-\frac{\alpha}{2}$, so dass

$$P\left(|T(\mathbf{X})| < z_{1-\frac{\alpha}{2}}\right) = P\left(-z_{1-\frac{\alpha}{2}} < T(\mathbf{X}) < z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

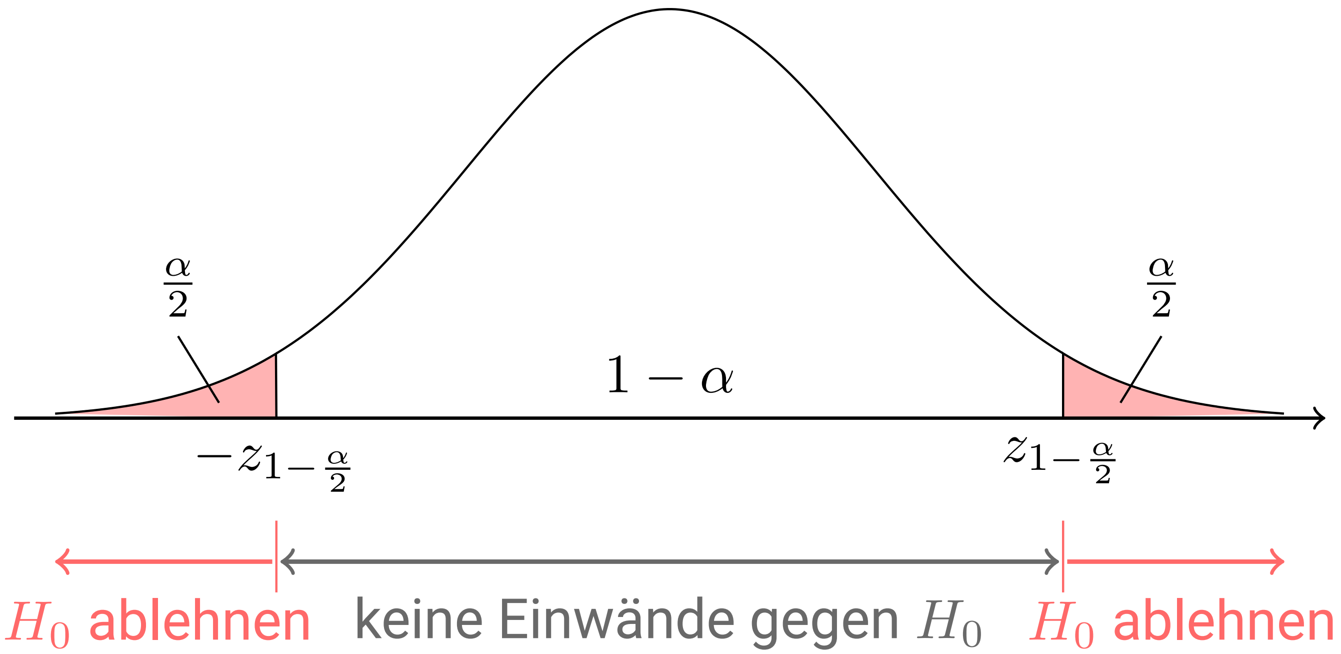
   und somit $K^\star = \left\{t \in \mathbb{R} : \ |t| > z_{1-\frac{\alpha}{2}}\right\}$

4. Entscheidungsregel: liegt konkrete Testwert $t = T(\mathbf{x})$ der konkreten Stichprobe $\mathbf{x}$ im kritischen Bereich?

   falls $t \in K^\star$ Ablehnen von H_0, d.h. der Erwartungswert μ unterscheidet sich signifikant von μ_0

   falls $t \notin K^\star$ Test hat keine Einwände gegen H_0, d.h. der durchgführte Test hat keine Einwände gegen $\mu = \mu_0$

Grafische Veranschaulichung



Hinweis: Formulierung alternativer Hypothesen (einseitige Fragestellungen) wirkt sich lediglich auf die Konstruktion des Kritischen Bereichs aus

$H_0: \ \mu \leq \mu_0$, $H_1: \ \mu > \mu_0$ (rechtsseitig)

Kritisch für H_0 sind nur positive Abweichungen $\overline{X}_n - \mu_0 > 0$, d.h. zu große Werte von T sprechen gegen H_0:

$$P\left(T(\mathbf{X}) > z_{1-\alpha}\right) = 1 - \Phi(z_{1-\alpha}) = 1 - (1-\alpha) = \alpha$$

Kritischer Bereich: $K^\star = \{t \in \mathbb{R}: \ t > z_{1-\alpha}\}$

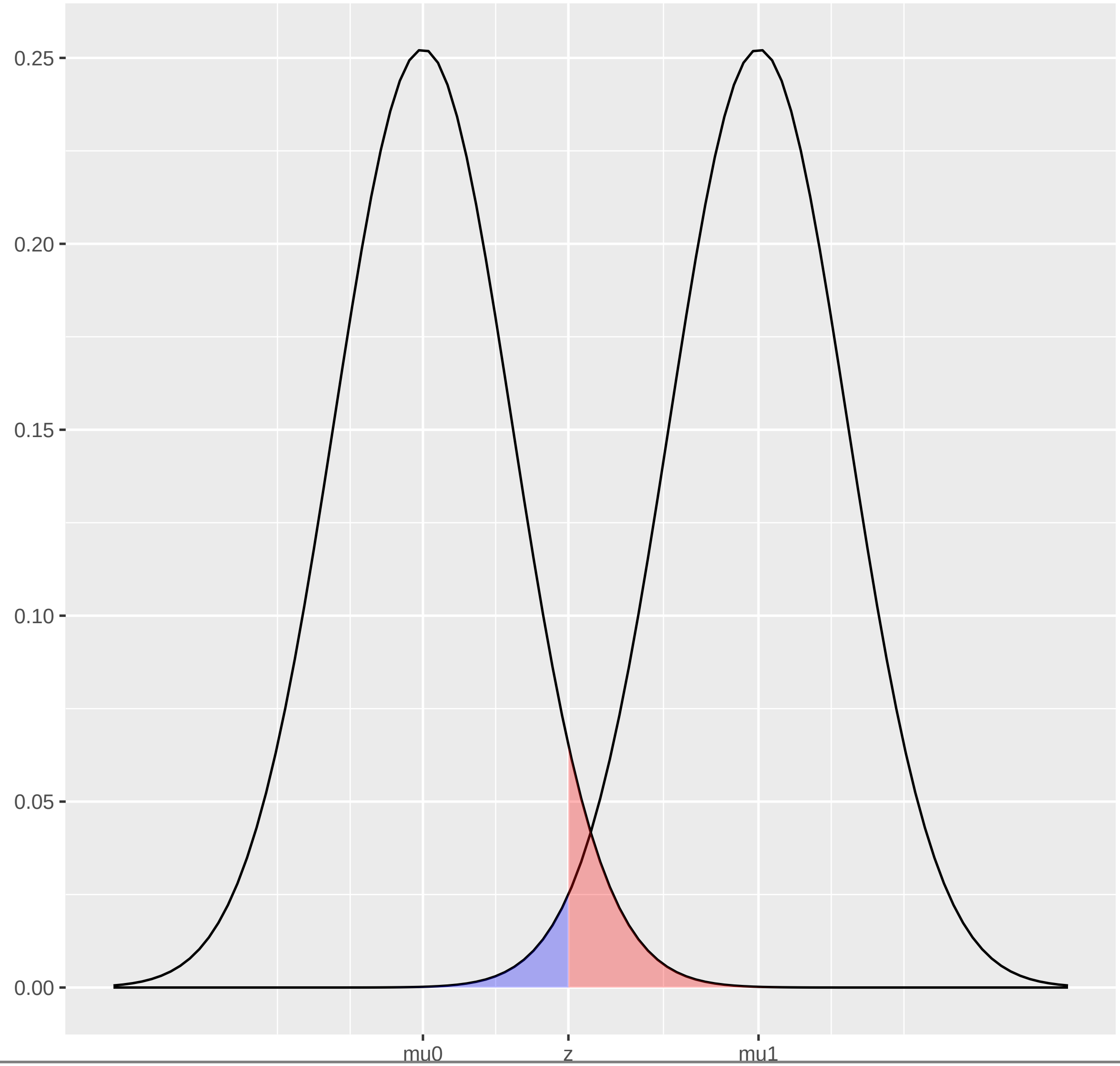
$H_0: \ \mu \geq \mu_0$, $H_1: \ \mu < \mu_0$ (linksseitig)

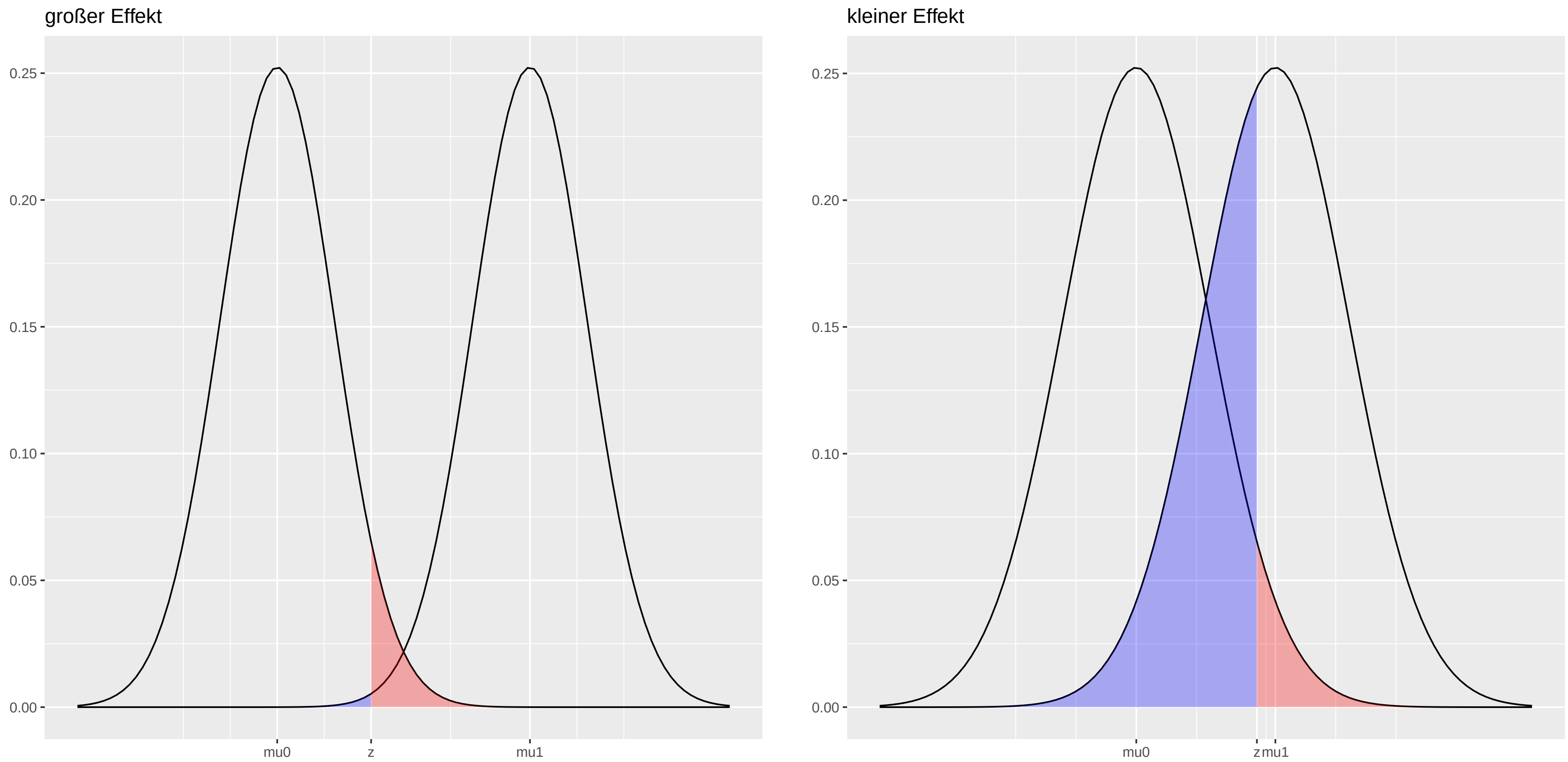
Kritisch für H_0 sind nur negative Abweichungen $\overline{X}_n - \mu_0 < 0$, d.h. zu kleine Werte von T sprechen gegen H_0:

$$P\left(T(\mathbf{X}) < z_{\alpha}\right) = \Phi(z_{\alpha}) = \alpha$$

Kritischer Bereich: $K^\star = \{t \in \mathbb{R}: \ t < z_{\alpha}\} = \{t \in \mathbb{R}: \ t < -z_{1-\alpha}\}$

α und β−Fehler beim rechtsseitigen Gaußtest




großer Effekt
0.25
0.20
0.15
0.10
0.05
0.00
mu0
z
mu1
kleiner Effekt
0.25
0.20
0.15
0.10
0.05
0.00
mu0
z
mu1

Beispiel

$X \triangleq$ Abfüllgewicht einer Tafel Schokolade ($100g$)

Stichprobe $\mathbf{x} = (100,\ 97,\ 101,\ 96,\ 98,\ 102,\ 96,\ 100,\ 101,\ 98)$

ergibt $\overline{x}_{10} = 98.9$ und $s_{10} = 2.183$ Frage: Ist das Füllgewicht systematisch zu gering? Annahme: $X \sim \mathcal{N}(\mu, \sigma^2)$ mit $\sigma = 2$, $\alpha = 0.05$

Test:

- $H_0 : \ \mu \geq 100 = \mu_0$, $H_1 : \ \mu < 100 = \mu_0$
- Testgröße: $t = T(\mathbf{x}) = \sqrt{n}\frac{\overline{x}_n - \mu_0}{\sigma} = \sqrt{10}\frac{98.9-100}{2} = -1.739$
- Quantil der Standardnormalverteilung: $z_\alpha = -z_{1-\alpha} = -z_{0.95} = -1.645$
- Kritischer Bereich: $K^\star = \{t \in \mathbb{R} : \ t < -1.645\} = (-\infty, -1.645)$
- Auswertung: $t = -1.739 \in K^\star \Rightarrow$ lehnen H_0 ab, d.h. das Abfüllgewicht der Schokoladentafeln ist bei einer Irrtumswahrscheinlichkeit von 5% signifikant kleiner als $100g$

Was kommt in H_0 und was in H_1 beim einseitigen Testen?

Fortsetzung Beispiel Abfüllgewicht einer Tafel Schokolade

- **Untersuchen jetzt** $H_0: \ \mu \leq 100 = \mu_0$, $H_1: \ \mu > 100 = \mu_0$
- **identische** Testgröße: $t = T(\mathbf{x}) = \sqrt{n}\frac{\overline{x}_n - \mu_0}{\sigma} = \sqrt{10}\frac{98.9-100}{2} = -1.739$
- Quantil der Standardnormalverteilung: $z_{1-\alpha} = z_{0.95} = 1.645$
- Kritischer Bereich (**jetzt auf der rechten Seite**):
  $K^\star = \{t \in \mathbb{R}: \ t \geq 1.645\} = (1.645, \infty)$
- Auswertung: $t = -1.739 \notin K^\star \Rightarrow$ keine Einwände gegen H_0, d.h. Test hat bei 5%-tiger Irrtumswahrscheinlichkeit keine Einwände dagegen, dass das Abfüllgewicht der Schokoladentafeln kleiner als $100g$ ist $\triangleq$ **keine statistisch starke Aussage**
- **kein signifikanter** Nachweis, dass Abfüllgewicht über $100g$ liegt (diese Frage war allerdings vorab auch nicht von Interesse)

t-Test: Hypothesen über μ bei unbekanntem σ^2 für $X \sim \mathcal{N}(\mu, \sigma^2)$

1. Hypothesen H_0: $\mu = \mu_0$, $H_1 : \ \mu \neq \mu_0$

   bei einseitigen Fragestellungen:
   $H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$ bzw. $H_0 : \mu \geq \mu_0$, $H_1 : \mu < \mu_0$

2. Testgröße: ersetzen σ^2 durch $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$

$$T(\mathbf{X}) = \sqrt{n}\frac{\overline{X}_n - \mu_0}{S_n}$$

   falls H_0 stimmt, ist Testgröße $t-$verteilt mit $n-1$ Freiheitsgraden, d.h. $T(\mathbf{X}) \sim t_{n-1}$

Prüfen einer Wahrscheinlichkeit: Parametertest für binomialverteilte Merkmale

Merkmal $X \sim \mathrm{Bin}(1, p)$, d.h.

$$X = \begin{cases} 1\,, & \text{mit Wkt. } p \\ 0\,, & \text{mit Wkt. } 1-p \end{cases}$$

konkrete Stichprobe $\mathbf{x} = (x_1, \ldots, x_n)$ soll Aufschluss über den Parameter p geben

1. betrachten hier beispielhaft die einseitige Hypothese: $H_0 : \; p \leq p_0$, $H_1 : \; p > p_0$
2. Testgröße: $T(\mathbf{X}) = \sum\limits_{i=1}^{n} X_i \sim \mathrm{Bin}(n, p_0)$ falls H_0 stimmt
3. große Werte von T sprechen gegen H_0: zur Bestimmung des kritischen Bereichs ermitteln wir daher ein $k_{1-\alpha} \in \{0, 1, 2, \ldots, n\}$, so dass

$$P\left(T(\mathbf{X}) \leq k_{1-\alpha}\right) \geq 1-\alpha \quad \Leftrightarrow \quad P\left(T(\mathbf{X}) > k_{1-\alpha}\right) \leq \alpha$$

Prüfen einer Wahrscheinlichkeit: Parametertest für binomialverteilte Merkmale

4. Kritischer Bereich: $K^\star = \{t \in \{0, 1, \ldots, n\} : \ t > k_{1-\alpha}\}$

   bei alternativen Fragestellungen:

   für $H_0 : p \geq p_0$ ergibt sich $K^\star = \{t \in \{0, 1, \ldots, n\} : \ t < k_\alpha\}$

   für $H_0 : p = p_0$ ergibt sich
   $K^\star = \left\{t \in \{0, 1, \ldots, n\} : \ t < k_{\frac{\alpha}{2}} \text{ oder } t > k_{1-\frac{\alpha}{2}}\right\}$

5. Entscheidungsregel: liegt konkreter Testwert $t = T(\mathbf{x})$ der konkreten Stichprobe $\mathbf{x}$ im kritischen Bereich?

   falls $t \in K^\star$ Ablehnen von H_0, d.h. die Wahrscheinlichkeit p ist signifikant größer als p_0

   falls $t \notin K^\star$ Test hat keine Einwände gegen H_0, d.h. der durchgeführte Test hat keine Einwände gegen $p \leq p_0$

Prüfen einer Wahrscheinlichkeit: Parametertest für binomialverteilte Merkmale, asymptotisches Vorgehen

1. für große Stichproben ($n \to \infty$) ist Testgröße asymptotisch normalverteilt, falls H_0 richtig ist:

$$T(\mathbf{X}) = \sum_{i=1}^{n} X_i \approx \mathcal{N}(np_0\,,\, np_0(1-p_0))$$

2. für hinreichend große n (vgl. Faustregel zu Grenzwertsatz von Moivre / Laplace) können wir Testgröße

$$T(\mathbf{X}) = \frac{\sum\limits_{i=1}^{n} X_i - np_0}{\sqrt{np_0(1-p_0)}} \approx \mathcal{N}(0,1)$$

   verwenden
3. Kritischer Bereich für $H_0 : p = p_0$: $K^\star = \left\{t \in \mathbb{R} : \; |t| > z_{1-\frac{\alpha}{2}}\right\}$ (analog zu Fall 1 beim Test des Erwartungswerts eines normalverteilten Merkmals)

- für große n ist wegen dem Zentralen Grenzwertsatz das arithmetische Mittel näherungsweise normalverteilt, d.h.

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \approx \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

- somit können Gauss-Test und $t-$Test für große Stichproben auch für beliebig verteilte Merkmale X durchgeführt werden
- die $t-$Verteilung mit n Freiheitsgraden kann für $n \geq 30$ durch die Standardnormalverteilung approximiert werden