

# Einführung in die induktive Statistik

Alois Pichler & Dana Uhlig

5. Oktober 2018

## Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>1</b>
<b>2</b>	<b>Zufallsgrößen</b>	<b>1</b>
2.1	Binomialverteilung . . . . .	7
2.2	Normalverteilung . . . . .	11
<b>3</b>	<b>Das Gesetz der großen Zahlen</b>	<b>16</b>
3.1	Grenzwertsatz von Moivre und Laplace: Approximation der Binomialverteilung durch die Normalverteilung . . . . .	18
3.2	Zentraler Grenzwertsatz . . . . .	21
<b>4</b>	<b>Induktive Statistik</b>	<b>22</b>
4.1	Punktschätzer . . . . .	22
4.2	Konfidenzintervalle . . . . .	27
4.2.1	Konfidenzintervalle für den Erwartungswert der Normalverteilung . . . . .	27
4.2.2	Konfidenzintervalle für die Binomialverteilung . . . . .	32
4.2.3	Einseitige Konfidenzintervalle . . . . .	35
4.3	Hypothesentests . . . . .	35
4.3.1	Gauß-Test zur Überprüfung des Erwartungswertes der Normalverteilung bei bekannter Varianz . . . . .	39
4.3.2	t-Tests zur Überprüfung des Erwartungswertes der Normalverteilung bei unbekannter Varianz . . . . .	43
4.3.3	Binomialtest für den Parameter $p$ der Binomialverteilung . . . . .	45
4.3.4	Tests für den Erwartungswert beliebiger Verteilungen . . . . .	46
<b>5</b>	<b>Gemischte Aufgaben</b>	<b>47</b>

## 1 Einführung

Ziel dieses Crashkurses ist, die Kenntnisse der beurteilenden bzw. induktiven Statistik aufzufrischen. Hierzu wiederholen wir zunächst in Kapitel 2 die wichtigsten Konzepte im Umgang Zufallsgrößen und das Gesetz der großen Zahlen in Kapitel 3, da diese als „Hauptwerkzeug“ beim Durchführen statistischer Tests oder der Berechnung von Konfidenzintervallen interpretiert werden können.

## 2 Zufallsgrößen

Die Modellierung des Zufalls geschieht mit Hilfe eines **Wahrscheinlichkeitsraumes**  $(\Omega, \mathcal{A}, P)$ , wobei  $\Omega$  die Menge aller Versuchsausgänge des entsprechenden Zufallsexperiments,  $\mathcal{A}$  das Ereignissystem ( $\sigma$ -Algebra) und  $P$  das Wahrscheinlichkeitsmaß bezeichnen.

Eine reelle **Zufallsvariable** bzw. **Zufallsgröße**  $X$  ist eine (messbare) Abbildung

$$X : \Omega \rightarrow \mathbb{R}, \tag{1}$$

die jedem Ausgang des Zufallsexperiments  $\omega \in \Omega$  eine reelle Zahl  $X(\omega) \in \mathbb{R}$  zuordnet, so dass für jedes  $x \in \mathbb{R}$

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A} \tag{2}$$

gilt.

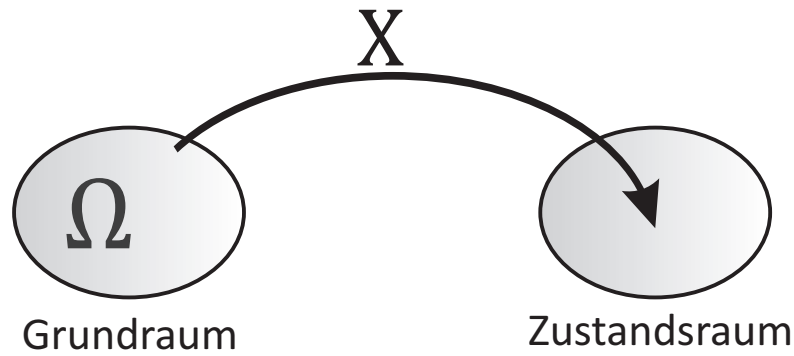


Abbildung 1: reelle Zufallsvariable: (messbare) Abbildung vom Ereignisraum auf den Zustandsraum  $\mathbb{R}$

Die Bedingung (2) der Messbarkeit sichert letztendlich, dass für jedes beliebige Intervall  $I \subset \mathbb{R}$  die Wahrscheinlichkeit  $P(X \in I)$  berechnet werden kann, da dann die entsprechenden Urbilder der Abbildung  $X^{-1}(I) = \{\omega : X(\omega) \in I\}$  Ereignisse sind, die zum Ereignissystem  $\mathcal{A}$  gehören.

**Beispiel 2.1: Glücksrad**

Kosten 1 mal Drehen:  
2,50€

Gewinnverteilung:

rot	grün	blau
5€	2€	1€

**Zufallsmodell**

**Wahrscheinlichkeitsraum**  $(\Omega, \mathcal{A}, P)$  mit  
**Ereignisraum**  $\Omega = \{\underbrace{\text{rot}}_{=\omega_1}, \underbrace{\text{grün}}_{=\omega_2}, \underbrace{\text{blau}}_{=\omega_3}\}$   
**Ereignissystem**  $\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \Omega\}$   
**Wahrscheinlichkeit**  $P$  mit  $P(\{\omega_1\}) = P(\{\omega_2\}) = \frac{1}{4}$  und  $P(\{\omega_3\}) = \frac{1}{2}$

**Zufallsgröße**  $X \stackrel{\Delta}{=} \text{Gewinn}$

$$X(\omega) = \begin{cases} 5, & \text{falls } \omega = \omega_1 \quad (\stackrel{\Delta}{=} \text{rot}) \\ 2, & \text{falls } \omega = \omega_2 \quad (\stackrel{\Delta}{=} \text{grün}) \\ 1, & \text{falls } \omega = \omega_3 \quad (\stackrel{\Delta}{=} \text{blau}) \end{cases}$$

**Rechnen mit Zufallsgrößen, Bsp für  $I = (-\infty, 2]$**

$P(X \leq 2) = P(\{\omega_2, \omega_3\}) = P(\{\omega_2\} \cup \{\omega_3\}) = P(\{\omega_2\}) + P(\{\omega_3\}) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$   
 $= 1 - P(\{\omega_1\}) = 1 - \frac{1}{4}$

Das zufällige Verhalten der Zufallsgröße  $X$  wird durch die Funktion  $F : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$F(x) = P(X \leq x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) \quad (3)$$

beschrieben. Die Funktion (3) wird als **Verteilungsfunktion** der Zufallsgröße  $X$  bezeichnet.

**Beispiel 2.2: Fortsetzung Glücksrad, Beispiel 2.1**

Für die Zufallsgröße  $X$ , die den Gewinn beim Drehen des Glücksrades aus Beispiel 2.1 beschreibt, gilt

$$P(X = x) = \begin{cases} 1/4, & \text{falls } x = 2 \text{ oder } x = 5 \\ 1/2, & \text{falls } x = 1 \\ 0, & \text{sonst.} \end{cases} \quad (4)$$

Der Wert der Verteilungsfunktion an der Stelle 3 lässt sich mit

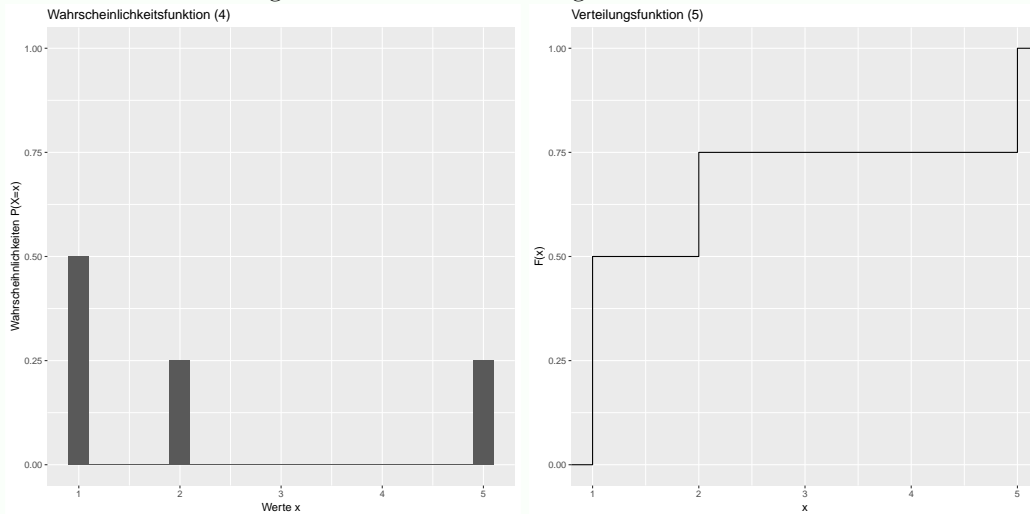
$$F(3) = P(X \leq 3) = P(X = 1) + P(X = 2) = 1/2 + 1/4 = 3/4$$

einfach berechnen und sagt aus, dass die Wahrscheinlichkeit beim Drehen des Glücksrades

höchstens 3€ zu gewinnen 75 Prozent beträgt. Allgemein lautet die Verteilungsfunktion des Gewinns

$$F(x) = \begin{cases} 0, & \text{falls } x < 1 \\ 1/2, & \text{falls } 1 \leq x < 2 \\ 3/4, & \text{falls } 2 \leq x < 5 \\ 1, & \text{falls } x \geq 5, \end{cases} \quad (5)$$

deren Verlauf in der folgenden rechten Grafik dargestellt ist.



Da die Zufallsgröße  $X$  nur die drei Werte 1, 2 oder 5 annimmt, wächst die Verteilungsfunktion nur in diesen Punkten genau um die Wahrscheinlichkeiten, mit denen diese Punkte angenommen werden. Die Sprunghöhen der Verteilungsfunktion (5) entsprechen also genau den Wahrscheinlichkeiten der einzelnen Werte der Zufallsgröße. Die komplette Information der Wahrscheinlichkeitsfunktion (4) ist also in der Verteilungsfunktion (5) enthalten.

Die Funktion (4), die jedem Wert  $x \in \mathbb{R}$  seine Wahrscheinlichkeit  $P(X = x)$  zuordnet, nennt man auch Wahrscheinlichkeitsfunktion.

Man unterscheidet je nach Wertebereich der Zufallsvariablen **diskrete** und **stetige** Zufallsgrößen. Der Wertebereich diskreter Zufallsvariablen ist abzählbar, während er für stetige Zufallsvariablen überabzählbar<sup>1</sup> ist. Die Zufallsgröße aus Beispiel 2.1 nimmt genau drei Werte an, so dass sie zur Klasse der diskreten Zufallsgrößen zählt. Allgemein charakterisiert die Wahrscheinlichkeitsfunktion  $x_i \mapsto p_i = P(X = x_i)$  jeder diskreten Zufallsgröße, die die abzählbaren Werte  $x_1, x_2, \dots$  annimmt, ebenso wie die Verteilungsfunktion

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p_i$$

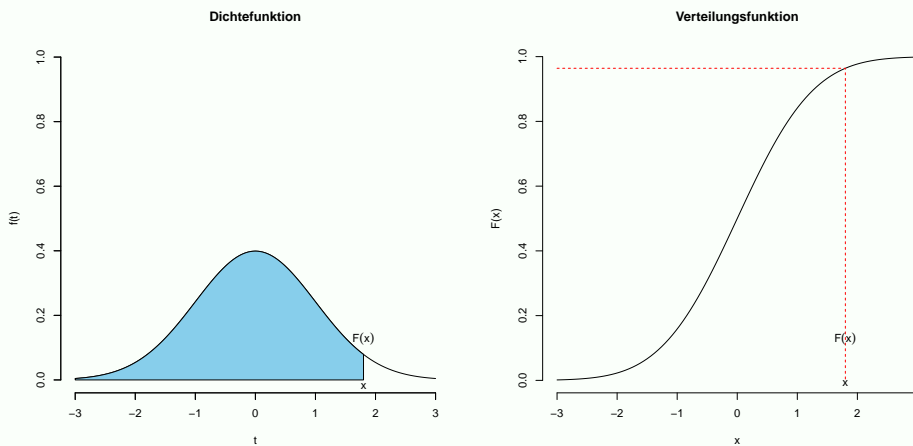
das komplette Zufallsverhalten. Für stetige Zufallsgrößen ist eine Betrachtung der Wahrscheinlichkeitsfunktion allerdings nicht ausreichend, um das Zufallsverhalten vollständig zu beschreiben. Typischerweise gilt für stetige Zufallsvariablen vielmehr sogar  $P(X = x) = 0$  für überabzählbar viele Punkte  $x \in \mathbb{R}$ . Falls für die stetige Zufallsgröße  $X$  mit Verteilungsfunktion  $F$  eine integrierbare Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  mit

$$F(x) = \int_{-\infty}^x f(t) dt, \quad (6)$$

existiert, gehört diese Zufallsvariable zur Klasse der absolut stetigen Zufallsgrößen und die Funktion  $f$  aus Gleichung (6) wird als **Dichtefunktion** bezeichnet. Beispiel 2.7 zeigt den Zusammenhang zwischen Dichte- und Verteilungsfunktion (absolut) stetiger Zufallsgrößen am Beispiel der Standardnormalverteilung.

<sup>1</sup>Es gelingt nicht alle Werte einer stetigen Zufallsvariablen durchzunummerieren. Misst die Zufallsvariable  $X$  beispielsweise die Entfernung zweier Punkte, so kann theoretisch jeder nicht negative Wert gemessen werden, so dass  $X \in \mathbb{R}_+$  gilt. Selbst wenn aus technischen Gründen nur Werte zwischen eins und zehn gemessen werden können, enthält das Intervall  $[1, 10]$  unendlich viele reelle Zahlen, die überabzählbar sind.

### Beispiel 2.3: Standardnormalverteilung



Die blaue Fläche unter der Dichtefunktion  $f$  bis zum Punkt  $x$  in der linken Grafik, genauer

$$\int_{-\infty}^x f(t) dt = F(x)$$

entspricht der Verteilungsfunktion im Punkt  $x$ , welche in der rechten Grafik abgetragen ist. Verschiebt man den Punkt  $x$  auf der Achse der Dichtefunktion nach links, wird die blaue Fläche kleiner. Verschiebt man in der rechten Grafik den Punkt  $x$  ebenso nach links, verringert sich der Wert der Verteilungsfunktion entsprechend. Wird der Punkt  $x$  hingegen nach rechts verschoben, wächst die blaue Fläche, so dass die Verteilungsfunktion ebenso größer wird. Der Anstieg der Verteilungsfunktion entspricht somit genau der Dichte, da die Umkehrung  $F'(x) = f(x)$  für alle Punkte  $x \in \mathbb{R}$  gilt.

Dichtefunktionen absolut stetiger Zufallsvariablen verallgemeinern das Konzept der Wahrscheinlichkeitsfunktion. Während für diskrete Zufallsgrößen mit Werten  $x_1, x_2, \dots$  die Wahrscheinlichkeit Werte aus dem Intervall  $(a, b]$  anzunehmen

$$P(a < X \leq b) = F(b) - F(a) = \sum_{a < x_i \leq b} p_i$$

gerade der Summe über alle Wahrscheinlichkeiten  $p_i = P(X = x_i)$  aller Werte  $x_i$ , die im Intervall  $(a, b]$  liegen, entspricht, gilt für absolut stetige Zufallsvariablen mit Dichtefunktion  $f$  hingegen

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t) dt.$$

Beispiel 2.4 veranschaulicht diesen Zusammenhang für die stetige Gleichverteilung.

### Beispiel 2.4: Gleichverteilung

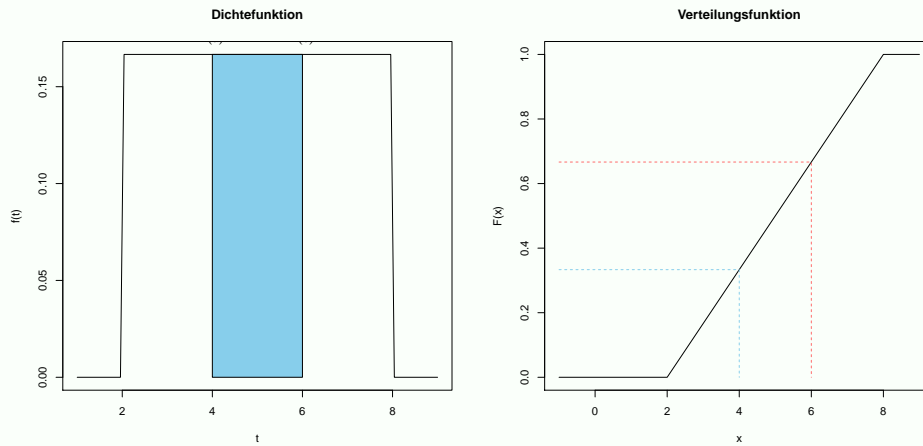
Die Zufallsgröße  $X$  mit Dichtefunktion

$$f(t) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq t \leq b \\ 0 & \text{für } t \notin [a, b] \end{cases} \quad (7)$$

heißt Gleichverteilung auf dem Intervall  $[a, b]$  ( $a < b$ ). Die Verteilungsfunktion steigt linear auf dem Intervall mit  $[a, b]$

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x \geq b \end{cases} \quad (8)$$

Die folgende Grafiken zeigen die Gleichverteilung auf dem Intervall  $[2, 8]$ , so dass die Dichte auf diesem Intervall konstant  $1/6$  ist.



Die blaue Fläche unter der Dichtefunktion  $f$  zwischen  $x = 4$  und  $x = 6$  entspricht der Wahrscheinlichkeit

$$P(4 \leq X \leq 6) = \int_4^6 f(t) dt = F(6) - F(4) = 0.6667 - 0.3333 = 0.3333.$$

Außerdem gilt für alle Wahrscheinlichkeiten der diskreten Zufallsvariablen  $0 \leq p_i \leq 1$  sowie  $\sum_i p_i = 1$  und in Analogie für die Dichtefunktion einer stetigen Zufallsvariablen  $f(t) \geq 0$  für alle Punkte  $t \in \mathbb{R}$  und  $\int_{-\infty}^{\infty} f(t) dt = 1$ . Allerdings unterliegen Dichtefunktionen nicht mehr der Einschränkung  $f(t) \leq 1$ , da diese nicht direkt Wahrscheinlichkeiten beschreiben, sondern die Fläche unter der Dichte Wahrscheinlichkeiten entsprechen. Für absolut stetige Zufallsgrößen gilt ferner für jedes  $x \in \mathbb{R}$  die Beziehung

$$P(X = x) = \int_x^x f(t) dt = F(x) - F(x) = 0,$$

so dass kein einzelner Punkt  $x \in \mathbb{R}$  mit  $P(X = x) > 0$  existiert.<sup>2</sup>

### Aufgabe 2.1

Schreiben Sie für die diskrete Zufallsvariable  $X \in \{0, 1, 2, \dots\}$  mit Hilfe der Verteilungsfunktion  $F$  die folgenden Wahrscheinlichkeiten

- (a)  $P(4 \leq X \leq 8)$
- (b)  $P(2 \leq X < 5)$
- (c)  $P(4 < X < 7)$
- (d)  $P(0 \leq X < 8)$

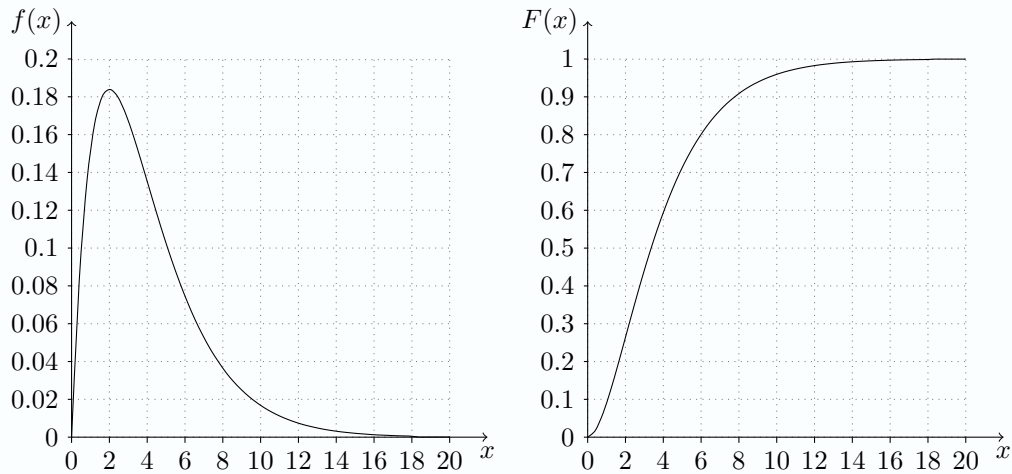
Lösung auf Seite ??

### Aufgabe 2.2

Die Dichte- und die Verteilungsfunktion einer stetigen Zufallsvariablen  $X$  sei gegeben durch:

$$f(x) = \begin{cases} \frac{1}{4}xe^{-x/2} & \text{für } x > 0 \\ 0 & \text{sonst,} \end{cases} \quad F(x) = \begin{cases} 1 - (1 + \frac{x}{2})e^{-x/2} & \text{für } x > 0 \\ 0 & \text{sonst.} \end{cases}$$

<sup>2</sup>Es gibt natürlich stetige Zufallsvariablen, für die einzelne Werte  $x_0 \in \mathbb{R}$  mit positiver Wahrscheinlichkeit  $P(X = x_0) > 0$  existieren können. Die Verteilungsfunktion springt dann in den singulären Punkt  $x_0$  um diese Wahrscheinlichkeit  $P(X = x_0) > 0$ . Für solche Zufallsvariablen existiert dann allerdings keine Dichtefunktion mit der Eigenschaft (6). Häufig wird die letzte Klasse der stetigen, aber nicht absolut stetigen Zufallsvariablen nicht weiter betrachtet und dann typischerweise die Klasse der absolut stetigen Zufallsvariablen einfach nur als stetige Zufallsvariablen bezeichnet.



1. Weisen Sie nach, dass die Funktion  $F$  tatsächlich Verteilungsfunktion der Dichte  $f$  ist.
2. Bestimmen Sie grafisch die Wahrscheinlichkeiten  $P(X \leq 2)$ ,  $P(X > 5)$  und  $P(2 \leq X \leq 5)$  mit Hilfe der Dichte- bzw. Verteilungsfunktion. Bestimmen Sie zusätzlich den exakten Wert.
3. Bestimmen Sie grafisch dasjenige  $x \in \mathbb{R}$ , so dass  $F(x) = p$  mit  $p = 0.25$ ,  $p = 0.5$  und  $p = 0.75$ . Interpretieren Sie die Ergebnisse inhaltlich.

Lösung auf Seite ??

Häufig betrachtet man neben der Verteilungsfunktion, welche das zufällige Verhalten von Zufallsvariablen vollständig beschreibt,<sup>3</sup> Kenngrößen wie Erwartungswert, Varianz, allgemeine höhere Momente oder Quantile, mit deren Hilfe die Verteilung einfacher charakterisiert werden kann. Der **Erwartungswert**

$$\mathbb{E}X = \int_{\Omega} X(\omega)P(d\omega) = \begin{cases} \sum_i x_i P(X = x_i), & \text{falls } X \text{ diskret mit Werten } x_1, x_2, \dots \\ \int_{-\infty}^{\infty} x f(x) dx, & \text{falls } X \text{ absolut stetig mit Dichte } f, \end{cases}$$

beschreibt beispielsweise die zentrale Lage und entspricht der durchschnittlichen Realisierung. Die **Varianz**

$$\text{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$$

hingegen gibt an, wie stark die einzelnen Realisierungen um den Erwartungswert schwanken. Mit Hilfe der allgemeinen **höheren Momente** ( $k \in \mathbb{N}$ )

$$\mathbb{E}X^k = \int_{\Omega} X(\omega)^k P(d\omega) = \begin{cases} \sum_i x_i^k P(X = x_i), & \text{falls } X \text{ diskrete Zufallsgröße mit Werten } x_1, x_2, \dots \\ \int_{-\infty}^{\infty} x^k f(x) dx, & \text{falls } X \text{ absolut stetige Zufallsgröße mit Dichte } f. \end{cases}$$

kann man die **Schiefe**  $\frac{\mathbb{E}(X - \mathbb{E}X)^3}{\sqrt{\text{Var}(X)}^3}$  und **Krümmung**  $\frac{\mathbb{E}(X - \mathbb{E}X)^4}{\text{Var}(X)^2}$  einer Verteilung berechnen. Ebenso gibt das **Quantil der Ordnung  $p$**

$$Q_p = F^-(p) := \min \{x \in \mathbb{R} : F(x) \geq p\} \quad (9)$$

wichtige Erkenntnisse über das typische Verhalten der betrachteten Zufallsgröße. Für die meisten stetigen Zufallsvariablen existiert die inverse Verteilungsfunktion  $F^{-1}$  und stimmt dann mit der verallgemeinerten Inversen  $F^-$  aus Gleichung (9) überein. Beispiel 2.5 stellt die verallgemeinerte Inverse einer diskreten Zufallsgröße dar.

### Beispiel 2.5: Quantile einer diskreten Zufallsgröße

Wir betrachten die diskrete Zufallsgröße, die die folgenden fünf Werte mit gleicher Wahrscheinlichkeit annimmt.

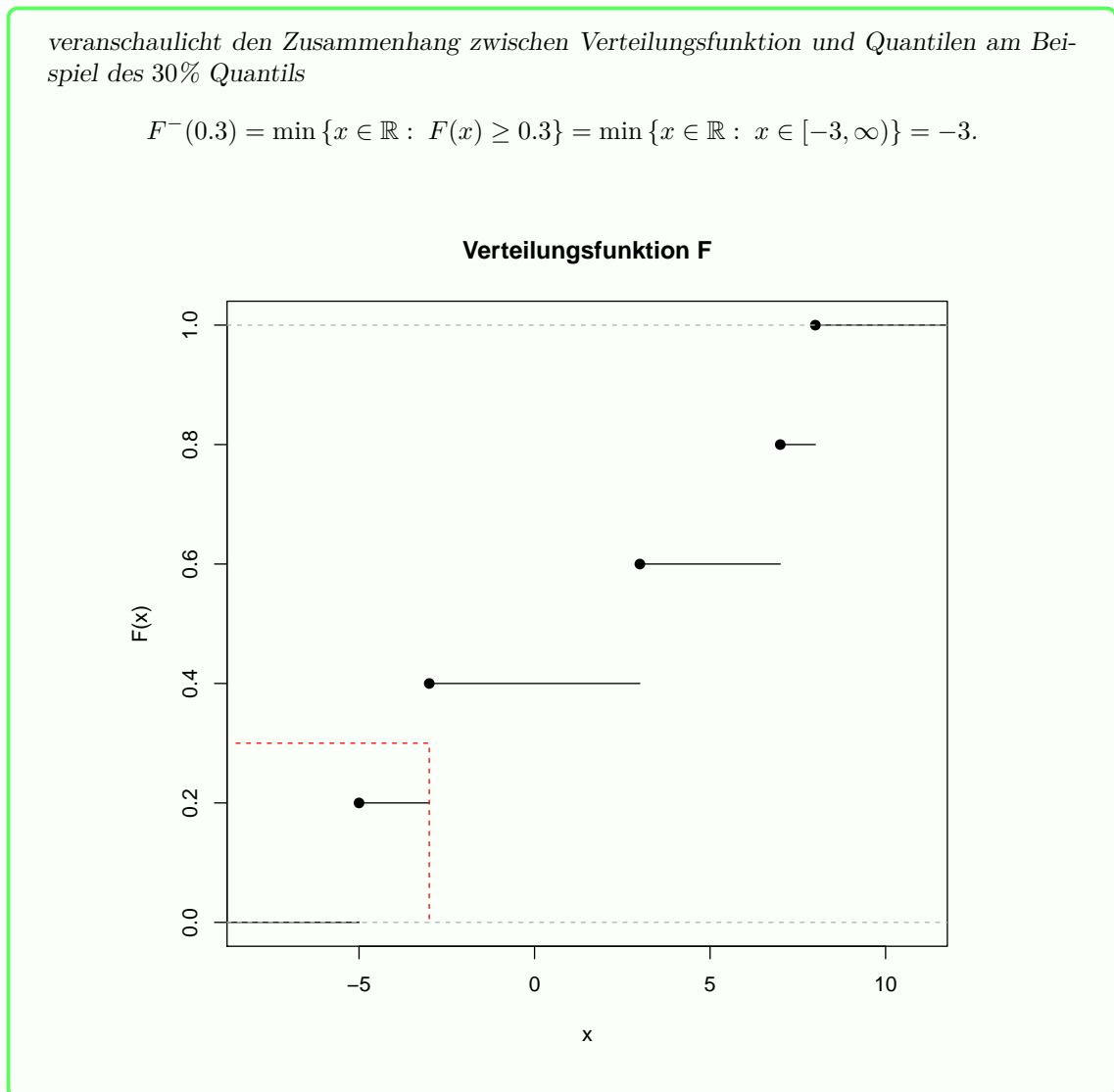
$x$	-5	-3	3	7	8
$P(X = x)$	1/5	1/5	1/5	1/5	1/5

Die folgende Grafik

<sup>3</sup>Natürlich sagt uns diese nicht voraus, welche Realisierung  $X(\omega)$  im nächsten Zug auftreten wird, allerdings können wir einschätzen wie wahrscheinlich es ist, dass die nächste Realisierung beispielsweise im Intervall  $(a, b]$  liegt, denn es gilt  $P(X \in (a, b]) = F(b) - F(a)$ . Siehe hierzu auch die Simulation in Beispiel 2.6 auf Seite 8.

veranschaulicht den Zusammenhang zwischen Verteilungsfunktion und Quantilen am Beispiel des 30% Quantils

$$F^{-}(0.3) = \min \{x \in \mathbb{R} : F(x) \geq 0.3\} = \min \{x \in \mathbb{R} : x \in [-3, \infty)\} = -3.$$



Im Folgenden sollen die Binomial- und die Normalverteilung näher besprochen werden.

## 2.1 Binomialverteilung

Eine diskrete Zufallsvariable  $X$ , die die Werte  $0, 1, \dots, n$  mit der Wahrscheinlichkeit

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n \quad (10)$$

heißt **binomialverteilt** mit Parametern  $n \in \mathbb{N}$  und  $p \in [0, 1]$ . Abkürzend schreibt man  $X \sim \text{Bin}(n, p)$ .

Die Binomialverteilung hat vielfältige Anwendungen und lässt sich mit dem Galton-Brett beschreiben. Dieses besteht aus  $n$  Ebenen wie in Abbildung 2 abgebildet. Von oben werden Kugeln eingeleitet, die sich in jeder Ebene jeweils unabhängig davon, was in der Ebene zuvor geschah, mit der Wahrscheinlichkeit  $p$  nach rechts und entsprechend mit der Wahrscheinlichkeit  $1-p$  nach links bewegen. Diese zufällige Bewegung einer Kugel auf der  $i$ -ten Ebene kann durch die unabhängigen Bernoulli verteilten Zufallsgrößen

$$X_i = \begin{cases} 1, & \text{Kugel nach rechts} \\ 0, & \text{Kugel nach links} \end{cases} \quad (11)$$

mit  $P(X_i = 1) = p$  modelliert werden. Dann entspricht die zufällige Summe

$$X = \sum_{i=1}^n X_i \quad (12)$$

genau der Anzahl der „Bewegungen nach rechts“. Landet eine Kugel also nach Durchlauf des Galtonbretts ganz links, nimmt  $X$  den Wert Null an. Landet die Kugel hingegen ganz rechts, gilt

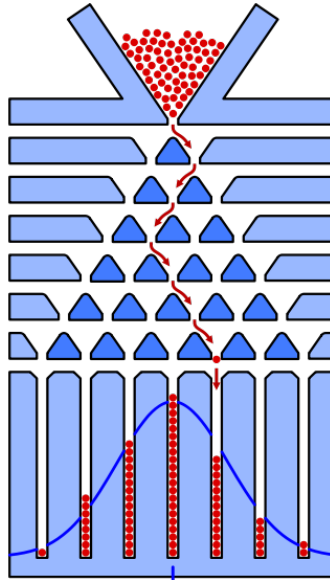


Abbildung 2: Galtonbrett

$X = n$ . Für  $X = k$  gibt es  $\binom{n}{k}$  verschiedene Pfade, die die Kugel durch das Galtonbrett nehmen kann. Da die Kugel sich in der  $i$ -ten Ebene unabhängig von der Bewegung in der vorherigen Ebene  $i - 1$  jeweils mit der Wahrscheinlichkeit  $p$  nach rechts und mit Wahrscheinlichkeit  $1 - p$  nach links bewegt, folgt somit

$$\begin{aligned} P(X = k) &= \binom{n}{k} P(\{X_1 = 1\} \cap \dots \cap \{X_k = 1\} \cap \{X_{k+1} = 0\} \cap \dots \cap \{X_n = 0\}) \\ &= \binom{n}{k} P(X_1 = 1) \cdot \dots \cdot P(X_{k+1} = 0) \cdot \dots \cdot P(X_n = 0) = \binom{n}{k} p^k (1 - p)^{n-k}, \end{aligned}$$

da für  $X = k$  die Kugel sich genau  $k$  mal nach rechts und  $n - k$  mal nach links bewegt.<sup>4</sup>

Das Prinzip des Galtonbretts lässt sich verallgemeinern. Jede Zufallsgröße  $X$ , die das Eintreten eines Ereignisses  $A$  mit  $P(A) = p$  bei  $n$  facher unabhängiger Wiederholung zählt, ist binomialverteilt. Tatsächlich ist die Zuordnung  $p_k = P(X = k)$  für  $k = 0, 1, \dots, n$  eine Wahrscheinlichkeitsfunktion, denn es gilt  $0 \leq p_k \leq 1$  sowie

$$\sum_{k=0}^n p_k = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1.$$

### Beispiel 2.6: 20 maliger Würfelwurf

Ein (idealer) Würfel wird  $n = 20$  mal geworfen. Wie hoch ist die Wahrscheinlichkeit, dass mindestens dreimal eine 6 gewürfelt wird? Zur Lösung betrachten wir die Zufallsvariable  $X$ , die die Anzahl der geworfenen Sechsen beim 20 maligen Würfelwurf zählt. Dann ist  $X \sim \text{Bin}(n = 20, p = 1/6)$  und es folgt

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) = 1 - F(2) = 1 - (P(X = 0) + P(X = 1) + P(X = 2)) \\ &= 1 - \left( \binom{20}{0} \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{20} + \binom{20}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{19} + \binom{20}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{18} \right) \\ &= 0.6713, \end{aligned}$$

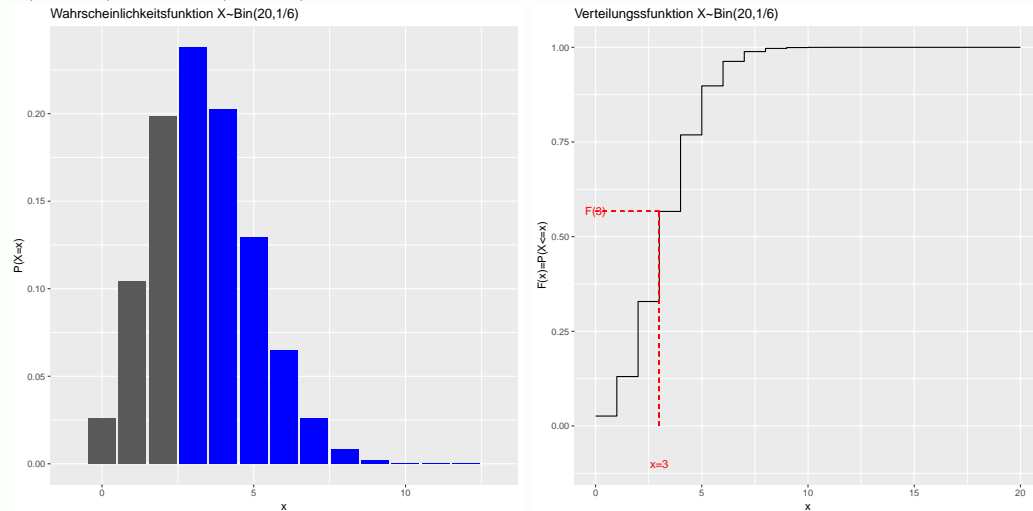
d.h. bei 20 Würfeln fällt mit 67%-iger Wahrscheinlichkeit mindestens dreimal eine 6. Die Berechnung der gesuchten Wahrscheinlichkeit

$$P(X \geq 3) = \sum_{k=3}^{20} P(X = k)$$

<sup>4</sup>In obiger Formel wurde unterstellt, dass die Kugel sich in der ersten  $k$  Ebenen nach rechts und den restlichen  $n - k$  Ebenen nach links bewegt hat. Natürlich spielt die Anordnung der Bewegungen nach rechts und links keine Rolle, da die Wahrscheinlichkeit für  $k$  Rechtsbewegungen und  $n - k$  Linksbewegungen immer  $p^k(1 - p)^{n-k}$  beträgt.



als Summe über alle Wahrscheinlichkeiten der Wahrscheinlichkeitsfunktion, für die  $X$  den Wert  $3, 4, \dots, 20$  (blau eingefärbt) annimmt, ist in der folgenden linken Grafik dargestellt. Die rechte Grafik zeigt hingegen die Berechnung  $P(X \geq 3) = 1 - F(2)$  mit Hilfe des Gegenereignisses  $\{X \geq 3\} = \{X < 3\} = \{X \leq 2\}$  sowie der Verteilungsfunktion  $F(2) = P(X \leq 2)$ . Die Wahrscheinlichkeit  $P(X \leq 2) = F(2) = P(X = 0) + P(X = 1) + P(X = 2)$  entspricht natürlich der Summe der grau eingefärbten Einzelwahrscheinlichkeiten  $P(X = 0)$ ,  $P(X = 1)$  sowie  $P(X = 2)$  in der linken Grafik.

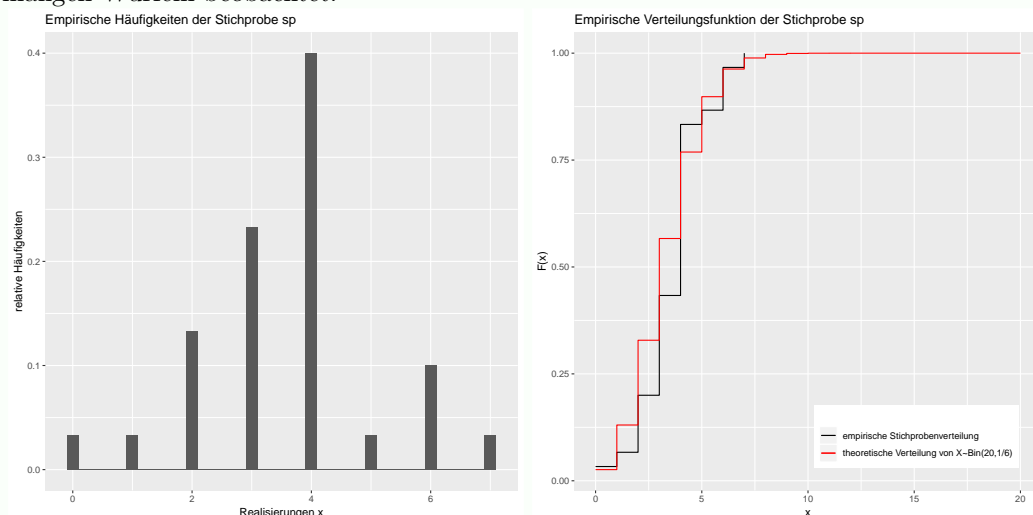


An dieser Stelle wollen wir nochmals betonen, dass die Verteilungsfunktion den kompletten Zufallscharakter beschreibt. Wenn wir jetzt 20 mal nacheinander mit einem fairen Würfel würfeln, dann beträgt die Wahrscheinlichkeit mindestens drei Sechsen zu beobachten gerade  $P(X \geq 3) = 0.6713$ . Das erlaubt natürlich keine Voraussage, dass bei unserem Experiment tatsächlich dieses Ereignis eintritt, sondern nur mit welcher Wahrscheinlichkeit. Würden wir das Experiment allerdings sehr oft wiederholen, so würde in etwas 67.1341 Prozent aller Experimente tatsächlich das Ereignis  $\{X \geq 3\}$  auftreten.

Zum Abschluss dieses Beispiels wollen wir dies demonstrieren, indem wir das Würfelexperiment zunächst 30 mal durchführen bzw. simulieren. Dabei erhalten wir die folgende Stichprobe

$$sp = (6, 2, 4, 4, 4, 4, 2, 4, 0, 3, 6, 3, 4, 4, 4, 7, 3, 4, 6, 1, 3, 3, 4, 2, 3, 5, 4, 3, 4, 2),$$

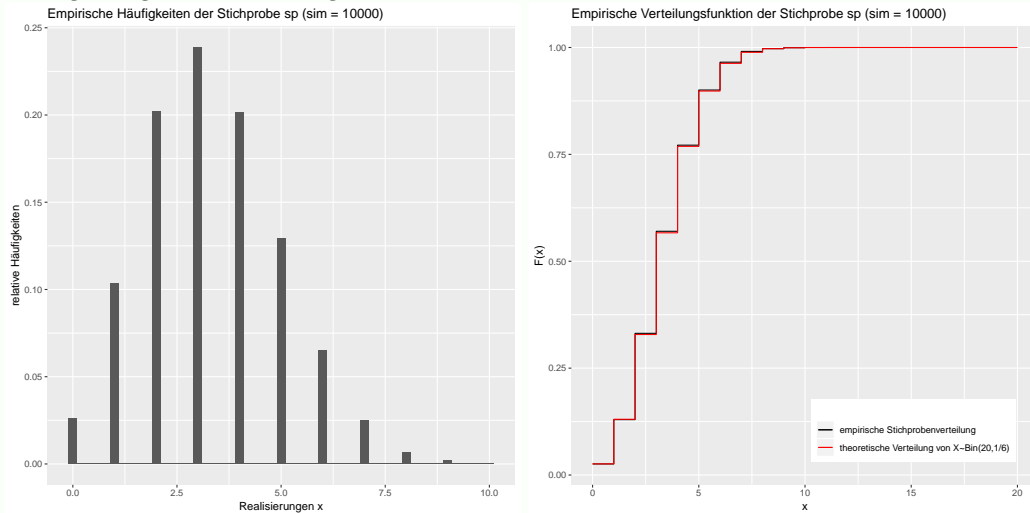
d.h. beim ersten Versuch haben wir genau 6 und beim zweiten genau 2 Sechsen beim 20 maligen Würfeln beobachtet.



Obige Grafiken zeigen die relativen und kumulierten Häufigkeiten dieser Simulation und stellen die empirische Wahrscheinlichkeits- sowie die empirische Verteilungsfunktion der Stichprobe dar. So haben wir in unserer Stichprobe 24 mal das Ereignis  $\{X \geq 3\}$  in 30 Durchgängen beobachtet, so dass die relative Häufigkeit  $24/30 = 0.8$  die tatsächliche Wahrscheinlichkeit  $P(X \geq 3) = 0.6713$  gut approximiert.

Mit Erhöhung der Simulationsanzahl auf 10000 verbessert sich auch die Güte dieser Approximation auf  $6686/10000 = 0.6686$ . Auch die Konvergenz der schwarzen empirischen Verteilungsfunktion gegen die rote theoretische Verteilungsfunktion ist dann zu beobachten, wie

die zugehörigen Grafiken zeigen.



Für den Erwartungswert der binomialverteilten Zufallsvariablen  $X \sim \text{Bin}(n, p)$  gilt

$$\begin{aligned} \mathbb{E}X &= \sum_{k=0}^n kP(X = k) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \frac{k n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n \frac{n(n-1)!}{(k-1)!(n-1-(k-1))!} p p^{k-1} (1-p)^{n-1-(k-1)} \\ &= np \sum_{j=0}^{n-1} \frac{n!}{j!(n-1-j)!} p^j (1-p)^{n-1-j} = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = np \end{aligned}$$

Diese etwas umständliche Berechnung lässt sich allerdings deutlich verkürzen, wenn wir die Darstellung (12) von  $X$  als Summe unabhängiger Bernoulli verteilter Null-Eins-Zufallsvariablen<sup>5</sup>

$$X = \sum_{i=1}^n X_i$$

mit  $P(X_i = 1) = p$  und  $P(X_i = 0) = 1 - p$  sowie die Additivität des Erwartungswertes nutzen, denn dann erhalten wir für den Erwartungswert

$$\mathbb{E}X = \mathbb{E} \left( \sum_{i=1}^n X_i \right) = \sum_{i=1}^n (\mathbb{E}X_i) = \sum_{i=1}^n p = np \quad (13)$$

Ebenso erhalten wir wegen der Unabhängigkeit von  $X_1, X_2, \dots, X_n$  für die Varianz von  $X \sim \text{Bin}(n, p)$

$$\text{Var} X = \text{Var} \left( \sum_{i=1}^n X_i \right) = \sum_{i=1}^n (\text{Var} X_i) = \sum_{i=1}^n p(1-p) = np(1-p). \quad (14)$$

Zum Abschluss wollen wir in Abbildung 3 noch kurz zeigen, wie mittels Geogebra<sup>6</sup> einfach Wahrscheinlichkeiten berechnet werden können, dass Zufallsvariablen Werte eines bestimmten Intervalles annehmen.

### Aufgabe 2.3

Im Schnitt gibt es in einem von vier Haushalten mindestens zwei Fernsehgeräte. In einer Siedlung befinden sich 12 Haushalte.  $X$  sei die Anzahl der Haushalte mit mehr als einem Fernsehgerät. Bestimmen Sie die Verteilungsfunktion und berechnen Sie die Wahrscheinlichkeit, dass in dieser Siedlung mindestens zwei Haushalte mehr als ein Fernsehgerät haben.

<sup>5</sup>Für den Erwartungswert der Bernoulli-Verteilung gilt  $\mathbb{E}X_i = 1 \cdot p + 0 \cdot (1-p) = p$  und für die Varianz  $\text{Var}(X_i) = \mathbb{E}X_i^2 - (\mathbb{E}X_i)^2 = 1^2 \cdot P(X_i = 1) + 0^2 \cdot P(X_i = 0) - p^2 = p - p^2 = p(1-p)$ .

<sup>6</sup>Unter Verwendung von Geogebra Classic im Unterprogramm Probability sind eine Vielzahl diskreter und stetiger Verteilungen implementiert.

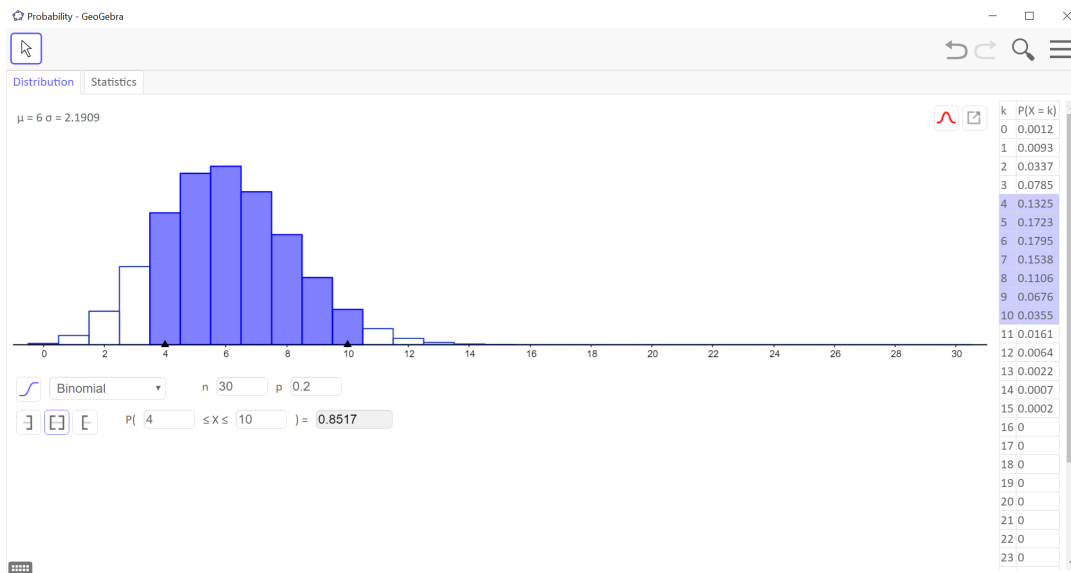


Abbildung 3: Berechnungen zur Binomialverteilung mit Geogebra

lichkeit, dass

- (a) höchstens die Hälfte der Haushalte maximal ein Fernsehgerät besitzt.
- (b) mindestens die Hälfte der Haushalte mehr als ein Fernsehgerät besitzt.
- (c) weniger als zwei Drittel der Haushalte maximal ein Fernsehgerät besitzt.
- (d) höchstens zwei Haushalte mehr als ein Fernsehgerät besitzen.

Lösung auf Seite ??

#### Aufgabe 2.4

Ein Multiple-Choice Test besteht aus 30 Aufgaben mit jeweils 5 Antworten, von denen jeweils nur eine richtig ist. Mit welcher Wahrscheinlichkeit wird man durch bloßes Raten folgende Anzahl von Aufgaben richtig beantworten?

- (a) mehr als 5 Aufgaben
- (b) mindestens 4 und höchstens 10 Aufgaben
- (c) weniger als 3 Aufgaben
- (d) genau 6 Aufgaben

Lösung auf Seite ??

#### Aufgabe 2.5

Einer Lieferung von Kondensatoren werden 50 Kondensatoren entnommen. Erfahrungsgemäß sind 15% der Kondensatoren schadhaft. Es beschreibe  $X$  die Anzahl schadhaft entnommener Kondensatoren. Bestimmen Sie den Erwartungswert und die Varianz und interpretieren Sie beide Größen!

Lösung auf Seite ??

## 2.2 Normalverteilung

Die stetige Zufallsgröße mit Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R} \quad (15)$$

für die Parameter  $\mu \in \mathbb{R}$  und  $\sigma \in \mathbb{R}_+$  mit  $\sigma > 0$  heißt **normalverteilt** und wird abkürzend mit  $X \sim \mathcal{N}(\mu, \sigma^2)$  bezeichnet.

## Beispiel 2.7: Standardnormalverteilung

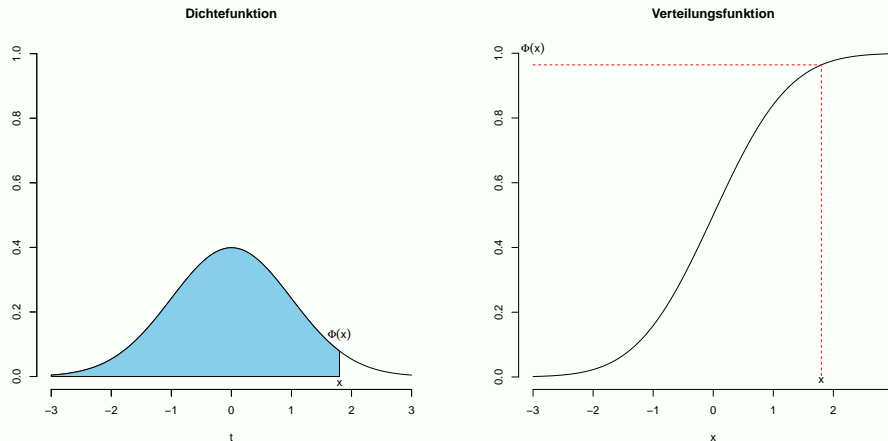
Für  $\mu = 0$  und  $\sigma = 1$  ergibt sich der Spezialfall der **Standardnormalverteilung**  $Z \sim \mathcal{N}(0, 1)$ , deren Dichtefunktion häufig mit  $\varphi$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (16)$$

und Verteilungsfunktion

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^z \varphi(t) dt \quad (17)$$

mit  $\Phi$  bezeichnet werden.



Die Verteilungsfunktion von  $X \sim \mathcal{N}(\mu, \sigma^2)$  ist formal definiert als

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \sigma e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt. \quad (18)$$

Gleichung (18) lässt sich allerdings nur numerisch auswerten, nicht aber analytisch. Daher kommt der Standardnormalverteilungsfunktion  $\Phi$  aus Gleichung (17) eine zentrale Bedeutung zu, da diese in Statistiksoftware und Taschenrechnern implementiert bzw. in Standardnormalverteilungstabellen für viele Werte  $z \in \mathbb{R}$  aufgelistet ist und sich die Verteilung von  $X \sim \mathcal{N}(\mu, \sigma^2)$  wegen  $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$  und

$$F(x) = P(X \leq x) = P(X - \mu \leq x - \mu) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Z \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

auf die Standardnormalverteilung zurückführen lässt.<sup>7</sup>

Erwartungswert und Varianz der Normalverteilung  $X \sim \mathcal{N}(\mu, \sigma^2)$  können mittels Variablensubstitution<sup>8</sup>

$$\begin{aligned} \mathbb{E}X &= \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \dots = \mu \\ \text{Var}(X) &= \int_{-\infty}^{\infty} (x - \mathbb{E}X)^2 f(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \dots = \sigma^2 \end{aligned}$$

berechnet werden, so dass der Parameter  $\mu \in \mathbb{R}$  dem Erwartungswert und  $\sigma > 0$  der Standardabweichung entspricht. Daher beeinflussen diese beiden Parameter auch den Verlauf der Dichtefunktion, wie Abbildung 4 verdeutlicht.

<sup>7</sup>Häufig sind die Werte  $\Phi(x)$  nur für positive  $x$  tabelliert, da für beliebiges  $x \in \mathbb{R}$  aufgrund der Symmetrie der Dichte einer Standardnormalverteilung immer  $\Phi(-x) = 1 - \Phi(x)$  gilt. Betrachten wir beispielsweise  $Z \sim \mathcal{N}(0, 1)$  so ist aufgrund der Symmetrie auch  $-Z \sim \mathcal{N}(0, 1)$  und es folgt  $\Phi(-x) = P(Z \leq x) = 1 - P(Z > -x) = 1 - P(-Z < x) = 1 - \Phi(x)$ .

<sup>8</sup>Hierfür nutzen wir  $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi} = \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx$ .

### Dichte der Normalverteilung

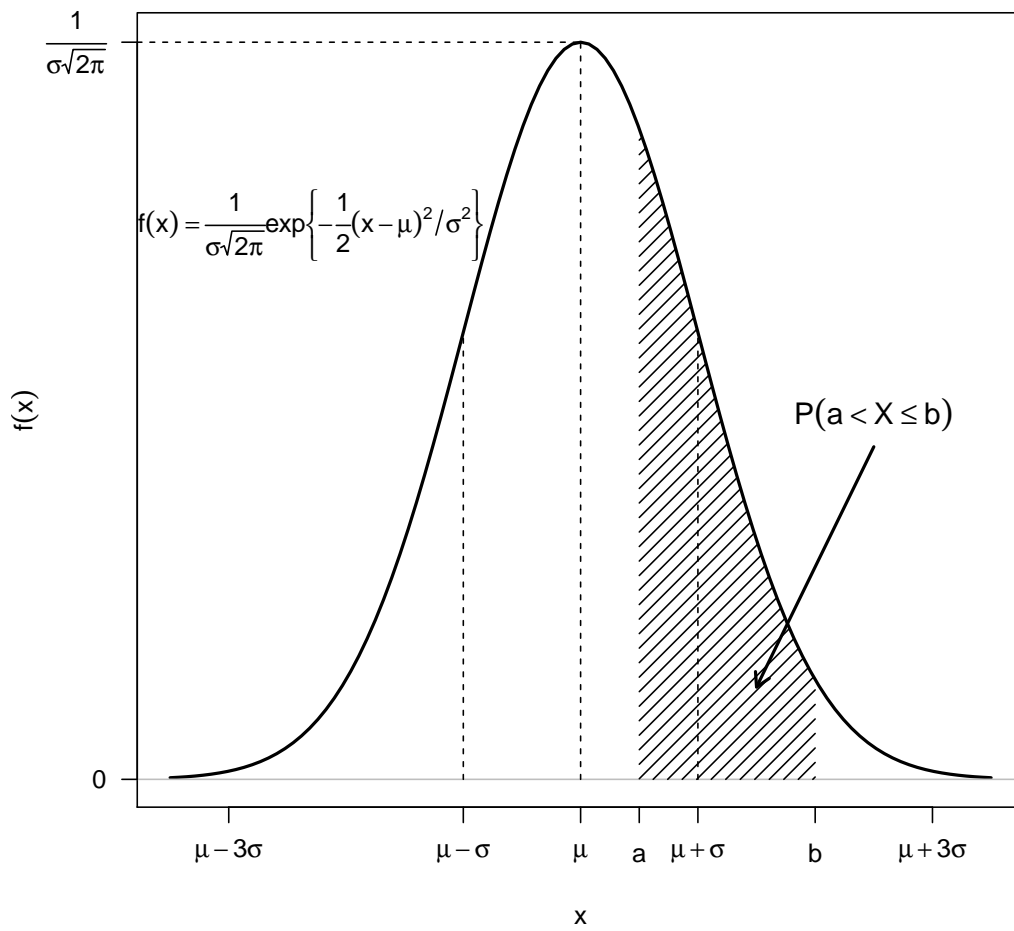


Abbildung 4: Einfluss der Parameter auf die Dichtefunktion der Normalverteilung

Das Maximum der glockenförmigen Dichte wird im Punkt  $x = \mu$  angenommen. Zudem geht durch diesen Punkt die Symmetrieachse der symmetrischen Dichte. Die Punkte  $x = \mu \pm \sigma$  entsprechen Wendestellen. Daher bestimmt  $\sigma^2$  den Breitenverlauf der Glockenkurve, d.h. ein großes  $\sigma$  liefert eine breite Glockenkurve, da die Werte der Zufallsgröße  $X$  stark um den Erwartungswert  $\mu$  schwanken. Ein kleines  $\sigma$  führt hingegen zu einer nadelförmigen Glockenkurve, so dass die Werte nur wenig um  $\mu$  streuen.

Die Normalverteilung nimmt unter den Verteilungen eine besondere Rolle ein. Einerseits ist die Summe beliebig normalverteilter unabhängiger Zufallsvariablen wieder normalverteilt.<sup>9</sup> Genauer gilt für beliebige normalverteilte Zufallsgrößen  $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$  für  $i = 1, 2, \dots, n$ , die vollständig unabhängig sind, dass die Summe

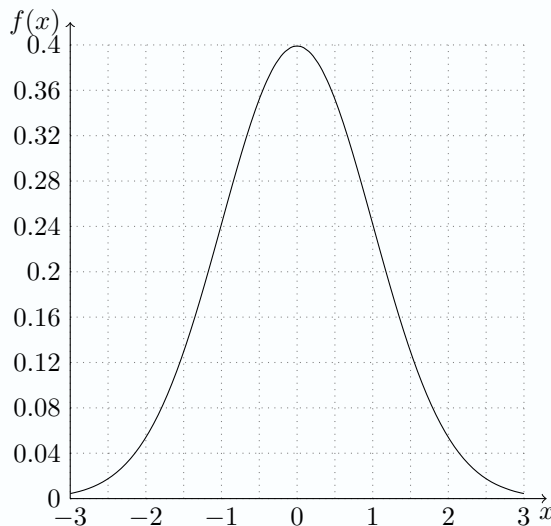
$$Z = \sum_{i=1}^n X_i \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

wieder normalverteilt ist. Diese Eigenschaft, dass unabhängige Zufallsvariablen einer Verteilungsklasse bezüglich der Addition invariant sind, kommt nur bei sehr wenigen Verteilungen vor. So ist beispielsweise die Summe  $Z = X + Y$  für unabhängige binomialverteilte Zufallsvariablen  $X \sim \text{Bin}(1, p)$  und  $Y \sim \text{Bin}(1, q)$  für  $p \neq q$  nicht binomialverteilt.<sup>10</sup> Andererseits kommt der Normalverteilung durch den zentralen Grenzwertsatz eine zentrale Bedeutung zu, da die Summe  $n$  unabhängiger und identisch, aber beliebig verteilter Zufallsvariablen mit endlichen Erwartungswert  $\mu = \mathbb{E}X_i$  ( $|\mu| < \infty$ ) und endlicher Varianz  $\sigma^2 = \text{Var}(X_i) < \infty$  für große  $n$  näherungsweise normalverteilt ist. Die Normalverteilung ist insbesondere dann für die Beschreibung von stochastischen Modellen geeignet, wenn sich die betrachtete Zufallsvariable als Summe einer großen Anzahl von unabhängigen Einflüssen ergibt. Beispielsweise können stetige Fehlergrößen oder Messfehler im Allgemeinen in guter Näherung als normalverteilt angenommen werden.

Abbildung 5 zeigt für  $X \sim \mathcal{N}(100, 3^2)$  die einfache Berechnung  $P(93 \leq X \leq 100) = F(100) - F(93) = \Phi\left(\frac{100-100}{3}\right) - \Phi\left(\frac{93-100}{3}\right) = \Phi(0) - \Phi(-2.3333) = 0.5 - 0.0098 = 0.4902$  mittels Geogebra.

### Aufgabe 2.6

Es sei  $Z \sim \mathcal{N}(0, 1)$  eine standardnormalverteilte Zufallsgröße. Schraffieren und berechnen Sie die folgenden Wahrscheinlichkeiten.



- (a)  $P(-2 < Z < 1)$
- (b)  $P(-1 < Z < 2)$
- (c)  $P(-1 < Z < 1)$
- (d)  $P(-2 < Z < 2)$
- (e)  $P(Z > -1)$
- (f)  $P(Z > 1)$

Lösung auf Seite ??

<sup>9</sup>Auch die Summe abhängiger normalverteilter Zufallsvariablen  $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$  und  $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$  ist wieder normalverteilt mit Erwartungswert  $\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y = \mu_X + \mu_Y$ . Allerdings berechnet sich dann die Varianz  $\text{Var}(X + Y) = \mathbb{E}(X - \mathbb{E}X + Y - \mathbb{E}Y)^2 = \text{Var}(X) + 2\mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)) + \text{Var}(Y) = \text{Var}(X) + 2\text{cov}(X, Y) + \text{Var}(Y)$  in Abhängigkeit von der Kovarianz  $\text{cov}(X, Y) = \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$  zwischen  $X$  und  $Y$ , die für unabhängige Zufallsgrößen wegen  $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}Y$  immer verschwindet.

<sup>10</sup>Allerdings bleibt die Binomialverteilung bezüglich der Summenbildung für binomialverteilte Zufallsgrößen  $X_i \sim \text{Bin}(m_i, p)$  erhalten, wenn alle unabhängigen Summanden die selbe Erfolgsquote  $p \in [0, 1]$  besitzen, da dann  $Z = \sum_{i=1}^n X_i \sim \text{Bin}\left(\sum_{i=1}^n m_i, p\right)$  gilt.

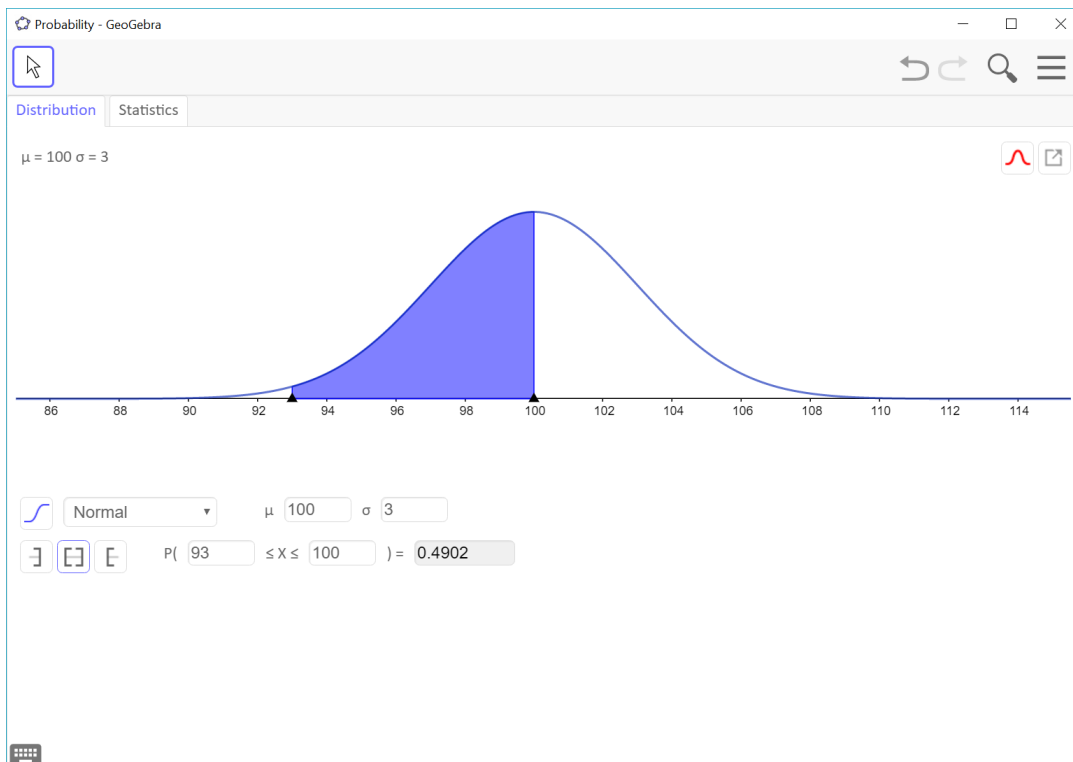


Abbildung 5: Berechnungen zur Normalverteilung mit Geogebra

### Beispiel 2.8: Sigma-Regel für die Normalverteilung

Das symmetrische Intervall

$$[\mu - k\sigma, \mu + k\sigma]$$

um den Erwartungswert  $\mu = \mathbb{E}X$  der normalverteilten Zufallsvariablen  $X \sim \mathcal{N}(\mu, \sigma^2)$  nennt man auch  $k$ -Sigma-Intervall. Die Wahrscheinlichkeit

$$\begin{aligned} p &= P(X \in [\mu - k\sigma, \mu + k\sigma]) = P(\mu - k\sigma \leq X \leq \mu + k\sigma) = P\left(-k \leq \frac{X - \mu}{\sigma} \leq k\right) \quad (19) \\ &= \Phi(k) - \Phi(-k) = 2\Phi(k) - 1, \end{aligned}$$

dass die Zufallsvariable Werte innerhalb dieses  $k$ -Sigma-Intervalls annimmt, hängt nur von  $k$  ab. Die folgende Tabelle zeigt die Wahrscheinlichkeiten aus Formel (19) für klassische Werte

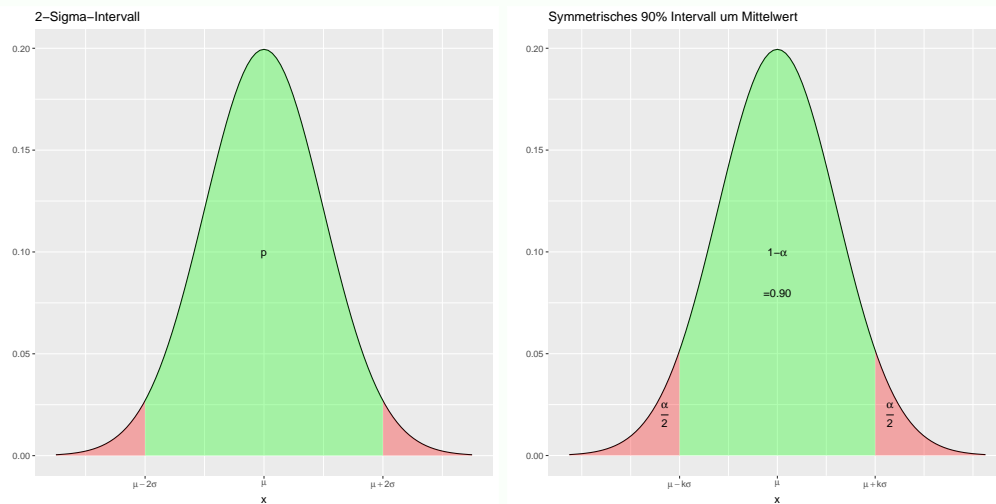
$k$	1	2	3	4
$p$	0.6827	0.9545	0.9973	0.9999

So ergibt sich für  $k = 2$ , dass die Wahrscheinlichkeit, dass sich die Werte der Normalverteilung um maximal  $2\sigma$  vom Erwartungswert entfernen, gerade 0.9545 beträgt. Dies ist in der linken Grafik dargestellt. Man kann allerdings auch sich die Wahrscheinlichkeit  $p = 1 - \alpha$  vorgeben<sup>a</sup> und daraus invers wegen

$$k = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

den Wert  $k \in \mathbb{R}_+$  ermitteln. Für typische  $\alpha$  Werte, die bei der Konstruktion von Konfidenzintervallen eine wichtige Rolle spielen, zeigt die folgende Tabelle die entsprechenden Werte für  $k$ . Die rechte Grafik zeigt das Intervall, dass sich für  $\alpha = 0.1$  ergibt.

$\alpha$	0.01	0.05	0.1	0.2
$k$	2.5758	1.96	1.6449	1.2816



<sup>a</sup>In der induktiven Statistik, die wir näher in Abschnitt 4 diskutieren werden, weist man typischerweise dem Intervall  $[\mu - k\sigma, \mu + k\sigma]$  die Wahrscheinlichkeit  $1 - \alpha$  zu, so dass die Wahrscheinlichkeit, die rot dargestellten Wahrscheinlichkeiten zusammen  $\alpha$  betragen.

### Aufgabe 2.7

Der Interelligenzquotient (IQ) ist normalverteilt mit  $\mu = 100$  und  $\sigma = 15$ .

- Welchen IQ muss man haben, um zu den intelligentesten 2% der Bevölkerung zu zählen?
- In welchem symmetrischen Bereich um den mittleren IQ liegen 80% der Bevölkerung?
- Ein einfacher Intelligenztest ist so ausgelegt, dass ihn 85% aller Personen bestehen. Welcher IQ ist dazu notwendig?

Lösung auf Seite ??

### Aufgabe 2.8

Die Auswertung eines Aufnahmetests ergibt, dass die erreichte Punktzahl in etwa normalverteilt mit Erwartungswert  $\mu = 80$  und Standardabweichung  $\sigma = 12$  ist.

- Wie viele Punkte muss man erreichen, um zu den besten 10% zu gehören?
- In welchem symmetrischen Intervall um den Erwartungswert liegen 95% aller Ergebnisse?
- Wie hoch ist die Punktzahl, die von 95% aller Antretenden mindestens erreicht wird?
- Es bewerben sich 2000 Personen, aber nur die besten 40 Bewerber können aufgenommen werden. Welche Punktzahl muss man dazu erreichen?

Lösung auf Seite ??

## 3 Das Gesetz der großen Zahlen

Ein Grundlage um typische Fragestellungen der beurteilenden Statistik bearbeiten zu können, bildet die Frage nach der Verteilung von  $Y_n = a_1X_1 + a_2X_2 + \dots + a_nX_n$  für identisch verteilte und unabhängige Zufallsvariablen  $X_1, X_2, \dots, X_n$  mit  $\mathbb{E}X_k = \mu$  und  $\text{Var}(X_k) = \sigma^2$ . Insbesondere der Fall  $a_1 = \dots = a_n = \frac{1}{n}$  spielt in der beurteilenden Statistik eine zentrale Rolle, da er auf das arithmetische Mittel

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$$

führt und dieses nach dem schwachen Gesetz der großen Zahlen gegen den Erwartungswert  $\mu$  konvergiert, falls  $n$  hinreichend groß gewählt wird. Somit kann man aus einer repräsentativen Stichprobe mit unabhängigen Beobachtungen durch die Bildung des Stichprobenmittels Rückschlüsse auf den Parameter  $\mu = \mathbb{E}X$  der beobachteten Zufallsvariablen ziehen. Dies werden wir später im Abschnitt 4 besprechen.

Egal wie die identisch verteilten unabhängigen Zufallsvariablen  $X_1, X_2, \dots, X_n$  verteilt sind, gilt



für den Erwartungswert des arithmetischen Mittels stets

$$\mathbb{E}\bar{X} = \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^n X_k\right) = \frac{1}{n}\sum_{k=1}^n \mathbb{E}(X_k) = \mu \quad (20)$$

und wegen der Unabhängigkeit<sup>11</sup> für die Varianz

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n}\sum_{k=1}^n X_k\right) = \frac{1}{n^2}\sum_{k=1}^n \text{Var}(X_k) = \frac{\sigma^2}{n} \xrightarrow{n \rightarrow \infty} 0. \quad (21)$$

Somit folgt aus der Tschebyscheff'schen Ungleichung für beliebiges  $\varepsilon > 0$  die Ungleichung

$$P(|\bar{X}_n - \mu| > \varepsilon) = P(|\bar{X}_n - \mathbb{E}\bar{X}_n| > \varepsilon) < \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

und folglich gilt für jedes  $n \in \mathbb{N}$

$$1 \geq P(|\bar{X}_n - \mu| \leq \varepsilon) > 1 - \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = 1 - \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 1,$$

womit letztendlich der Grenzfall  $n \rightarrow \infty$  dem schwachen Gesetz der großen Zahlen

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \leq \varepsilon) = 1$$

entspricht. Dieses besagt, dass die Wahrscheinlichkeit einer Abweichung größer als  $\varepsilon > 0$  zwischen arithmetischem Mittel und Erwartungswert für hinreichend große  $n$  gegen Null konvergiert.

Einen Spezialfall der schwachen Konvergenz der relativen Häufigkeit  $H_n(A)$  gegen die Wahrscheinlichkeit  $p = P(A)$  des Ereignisses  $A$ , erhält man durch  $n$  unabhängige Wiederholungen des Experimentes. Hierfür betrachten wir die Bernoulli verteilten Zufallsgrößen

$$X_i = \begin{cases} 1, & \text{wenn } A \text{ im } i\text{-ten Versuch eintritt} \\ 0, & \text{wenn } A \text{ im } i\text{-ten Versuch nicht eintritt,} \end{cases}$$

die unabhängig und identisch verteilt sind. Dann gilt  $\mu = \mathbb{E}X_i = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = P(A) = p$  und das arithmetische Mittel  $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$  entspricht der relativen Häufigkeit  $H_n(A)$  des Ereignisses  $A$ . Daher besagt das schwache Gesetz der großen Zahlen, dass für wachsende Versuchsanzahl  $n$  die relative Häufigkeit  $H_n(A) = \bar{X}$  schwach gegen die Wahrscheinlichkeit  $p = P(A)$ , genauer

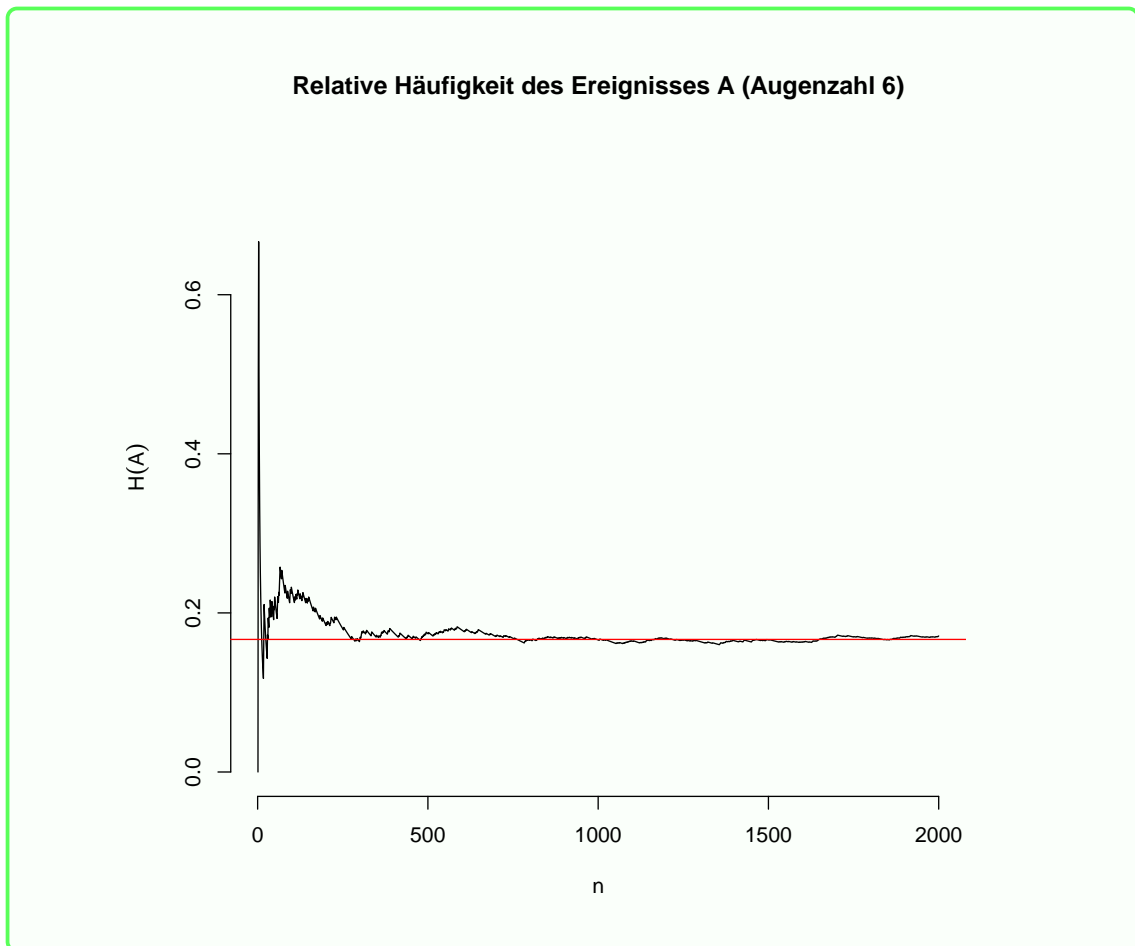
$$\lim_{n \rightarrow \infty} P(|H_n(A) - p| \leq \varepsilon) = 1$$

konvergiert.

### Beispiel 3.1: Konvergenz der relativen Häufigkeit beim Würfeln

Das schwache Gesetz der großen Zahlen gewährleistet, dass relative Häufigkeiten gegen die Wahrscheinlichkeiten konvergieren. Dies wollen wir anhand einer Simulation verdeutlichen und würfeln insgesamt  $n = 2000$  mal mit einem fairen Würfel und zählen wie oft das Ereignis  $A$ , dass eine Sechs gewürfelt wird, eintritt und berechnen daraus die relative Häufigkeit. Diese tragen wir in der folgenden Grafik gegen  $n$  ab, d.h. für  $n = 10$  werden bei der Berechnung der relativen Häufigkeit nur die ersten zehn Simulationen berücksichtigt, während für  $n = 2000$  alle Ergebnisse Berücksichtigung finden.

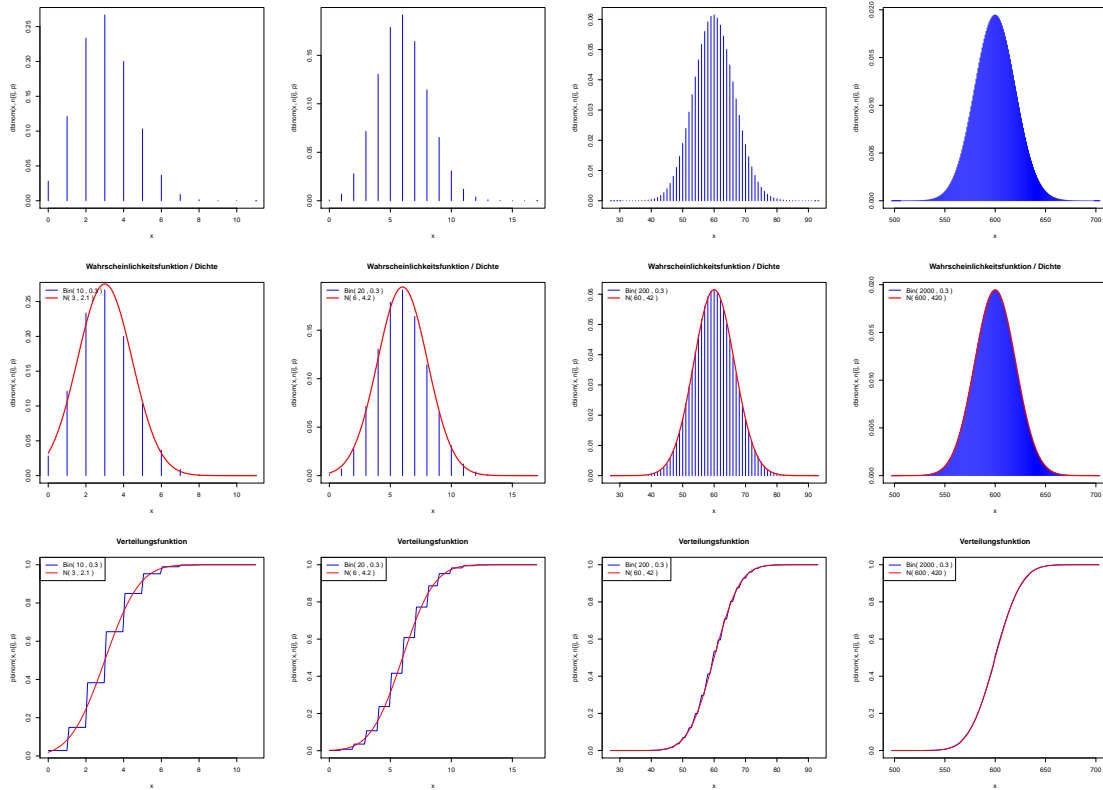
<sup>11</sup>Aus der Unabhängigkeit von  $X_k$  und  $X_j$  folgt  $\mathbb{E}(X_k X_j) = \mathbb{E}X_k \mathbb{E}X_j$  und somit  $\text{cov}(X_k, X_j) = \mathbb{E}((X_k - \mathbb{E}X_k)(X_j - \mathbb{E}X_j)) = \mathbb{E}(X_k X_j) - \mathbb{E}X_k \mathbb{E}X_j = 0$  für alle  $k \neq j$ , so dass die Varianz der Summe unabhängiger Zufallsvariablen  $\text{Var}(\sum_{i=1}^n X_i) = \sum_{j=1}^n \sum_{k=1}^n \text{cov}(X_j, X_k) = \sum_{j=1}^n \text{cov}(X_j, X_j) = \sum_{j=1}^n \text{Var}(X_j)$  die Summe aller Einzelvarianzen ist.



Beispiel 3.1 zeigt sehr deutlich, dass für größere  $n$  die relative Häufigkeit  $H_n(A)$  des Ereignisses  $A$  gegen die Wahrscheinlichkeit  $p = P(A) = 1/6$  konvergiert. Es ist allerdings auch zu erkennen, dass die relative Häufigkeit schwankt und selbst zufällig ist. Die Frage nach der Verteilung des arithmetischen Mittels beantwortet der zentrale Grenzwertsatz, den wir in Abschnitt 3.2 kurz etwas näher betrachten wollen. Der Grenzwertsatz von Moivre und Laplace kann als Spezialfall des zentralen Grenzwertsatzes betrachtet werden und beschreibt die Normalapproximation der Binomialverteilung für große  $n$ .

### 3.1 Grenzwertsatz von Moivre und Laplace: Approximation der Binomialverteilung durch die Normalverteilung

Der Grenzwertsatz von Moivre und Laplace, nachdem sich die Binomialverteilung von  $X \sim \text{Bin}(n, p)$  durch die Normalverteilung approximieren lässt, sofern der Parameter  $n$  hinreichend groß ist, stellt einen Spezialfall des zentralen Grenzwertsatzes dar und soll hier nur kurz anhand eines Beispiels demonstriert werden. Hierfür betrachten wir für  $p = 0.3$  und vier verschiedene Parameter  $n \in \{10, 20, 200, 2000\}$  jeweils die Zufallsgrößen  $X = X_1 + X_2 + \dots + X_n = n\bar{X} \sim \text{Bin}(n, p)$  als Summe Bernoulli verteilter Zufallsvariablen  $X_1, \dots, X_n$  (siehe Formel (11)) und stellen grafisch die Wahrscheinlichkeitsfunktion der entsprechenden Binomialverteilung und die Dichte der approximierten Normalverteilung sowie deren Verteilungsfunktionen dar. Historisch wurde diese Approximation 1730 von Abraham de Moivre für  $p = 1/2$  und im Jahre 1812 von Pierre-Simon Laplace für allgemeines  $0 < p < 1$  gezeigt, bevor dann später Lindberg und Lévy den zentralen Grenzwertsatz für allgemeine Verteilungen formulierten.



Es ist deutlich zu erkennen, dass die Normalapproximation für  $n = 200$  gut und für  $n = 2000$  sehr gut die Binomialverteilung beschreibt, während für  $n = 10$  und  $n = 20$  die Annäherungen der stetigen Normalverteilung an die diskrete Binomialverteilung noch recht grob sind. Der Grenzwertsatz macht nur für hinreichend große  $n$  Aussagen. Für praktische Anwendungen stellt sich natürlich die Frage, ob  $n$  bereits hinreichend groß ist. Als Faustregel gilt, dass unter Erfüllung der Forderung  $\text{Var}(X) = np(1 - p) > 9$  die Normalapproximation gute Näherungen liefert.

Zum Abschluss dieses Abschnittes wollen wir noch die sogenannte Stetigkeitskorrektur besprechen, die bei der Normalapproximation in der Regel bessere Ergebnisse liefert. Wie obige Grafik zeigt, ist gerade für kleiner  $n$  der Übergang von der treppenförmigen Verteilungsfunktion der Binomialverteilung zur stetigen Verteilungsfunktion der Normalverteilung teilweise recht grob. Somit weichen dann auch die Ergebnisse der Normalapproximation recht stark von dem exakten Ergebnis ab. Durch die Stetigkeitskorrektur versucht man die Approximation zu verbessern, indem die Grenzen leicht verschoben werden. Zur Illustration der Probleme beim Übergang von diskreten zu stetigen Zufallsvariablen betrachten wir für beliebige Intervallgrenzen  $a < b \in \mathbb{N}$  die folgende Wahrscheinlichkeiten

$$P(a \leq X \leq b) = P(a - 1 < X < b + 1) \tag{22}$$

Für die diskrete Zufallsvariable  $X \sim \text{Bin}(n, p)$  gilt tatsächlich obige Gleichung, während sie für eine stetige Zufallsvariable im Allgemeinen nicht wahr ist. Die Normalapproximation auf der linken und rechten Seite von (22) liefert daher verschiedene Ergebnisse

$$\begin{aligned} & \Phi\left(\frac{b - np}{\sqrt{np(1 - p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1 - p)}}\right) \\ & \neq \Phi\left(\frac{b + 1 - np}{\sqrt{np(1 - p)}}\right) - \Phi\left(\frac{a - 1 - np}{\sqrt{np(1 - p)}}\right). \end{aligned}$$

Mit der Stetigkeitskorrektur

$$\begin{aligned} P(a \leq X \leq b) &= P\left(a - \frac{1}{2} < X < b + \frac{1}{2}\right) \\ &\approx \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{np(1 - p)}}\right) - \Phi\left(\frac{a - \frac{1}{2} - np}{\sqrt{np(1 - p)}}\right) \end{aligned} \tag{23}$$

deutlich bessere Ergebnisse, wie das folgende Beispiel illustriert.

**Beispiel 3.2: Stetigkeitskorrektur bei der Normalapproximation von  $X \sim \text{Bin}(100, 0.25)$  für  $a = 15$  und  $b = 30$**

Die exakte Berechnung von

$$P(15 \leq X \leq 30) = \sum_{k=15}^{30} \binom{100}{k} 0.25^k 0.75^{100-k} = F(30) - F(14) = 0.8908$$

ist zwar leicht mit Statistiksoftware zu berechnen. Mit einem einfachen Taschenrechner ist die Berechnung allerdings recht umständlich, so dass die Normalapproximation in Verbindung mit der Standardnormalverteilungstabelle hier einen einfachen Weg zur Berechnung liefert. Die Faustregel  $np(1-p) = 18.75 > 9$  ist erfüllt, so dass die Normalapproximation der linken Seite von Gleichung (22)

$$P(15 \leq X \leq 30) \approx \Phi\left(\frac{30-25}{\sqrt{18.75}}\right) - \Phi\left(\frac{15-25}{\sqrt{18.75}}\right) = 0.8654$$

ebenso wie

$$\begin{aligned} P(15 \leq X \leq 30) &= P(14 < X \leq 30) = F(30) - F(14) \\ &\approx \Phi\left(\frac{30-25}{\sqrt{18.75}}\right) - \Phi\left(\frac{14-25}{\sqrt{18.75}}\right) = 0.8704 \end{aligned}$$

oder die rechte Seite von Gleichung (22)

$$\begin{aligned} P(15 \leq X \leq 30) &= P(14 < X < 31) \\ &\approx \Phi\left(\frac{31-25}{\sqrt{18.75}}\right) - \Phi\left(\frac{14-25}{\sqrt{18.75}}\right) = 0.9115 \end{aligned}$$

eine gute Näherung liefern. Die Stetigkeitskorrektur

$$P(15 \leq X \leq 30) = P(14.5 < X < 30.5) \approx \Phi\left(\frac{30.5-25}{\sqrt{18.75}}\right) - \Phi\left(\frac{14.5-25}{\sqrt{18.75}}\right) = 0.8903$$

liefert allerdings eine deutlich bessere Approximation.

**Aufgabe 3.1**

Ein Würfel wird

- (a) 42 mal,
- (b) 72 mal,
- (c) 114 mal

geworfen.  $X$  mit  $\mu = \mathbb{E}X$  und  $\sigma^2 = \text{Var}(X)$  gibt an, wie oft die Augenzahl 5 oder 6 gewürfelt wird. Berechnen sie die Wahrscheinlichkeit, dass  $X$  um maximal  $\sigma$  von  $\mu$  abweicht exakt und näherungsweise.

Lösung auf Seite ??

### Aufgabe 3.2

Schätzungen zufolge ist der Anteil der Mopedfahrer, die ein auffrisiertes Moped fahren, 40%.

- Die Polizei kontrolliert 120 Mopedfahrer. Mit welcher Wahrscheinlichkeit werden weniger als 40 mit einem auffrisierten Moped angetroffen?
- Bei einer weiteren Kontrolle sind von 100 Mopedfahrern 45 mit einem auffrisierten unterwegs. Wie groß ist die Wahrscheinlichkeit für eine so starke oder eine noch stärkere Abweichung<sup>a</sup> vom erwarteten Wert?

Lösung auf Seite ??

<sup>a</sup>Es interessieren an dieser Stelle nur Abweichungen nach oben, d.h. Werte über dem Erwartungswert.

### Aufgabe 3.3

Im letzten Jahr hatte die Computerfirma „Digitoll“ einen Marktanteil von 11 %.

- Berechnen Sie die Wahrscheinlichkeit, dass von 400 Computerkäufern sich mehr als 50 für die Marke „Digitoll“ entscheiden.
- Auf einer Computermesse, die dieses Jahr stattfindet, werden 8 000 Besucher erwartet, die einen Computer kaufen werden. Wie viele davon werden einen Computer der Marke „Digitoll“ kaufen? Bestimmen Sie dazu eine minimale bzw. maximale Anzahl symmetrisch um den Erwartungswert, so dass die tatsächliche Anzahl mit 95 %-iger Wahrscheinlichkeit zwischen Minimal- und Maximalwert liegt.

Lösung auf Seite ??

## 3.2 Zentraler Grenzwertsatz

Wir betrachten das arithmetische Mittel  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  von beliebig, aber identisch verteilten und unabhängigen Zufallsvariablen  $X_1, X_2, \dots, X_n$  mit  $\mu = \mathbb{E}X_i \in \mathbb{R}$  sowie  $0 < \text{Var}(X_i) = \sigma^2 < \infty$ . Dann konvergiert die Verteilung des standardisierten arithmetischen Mittels

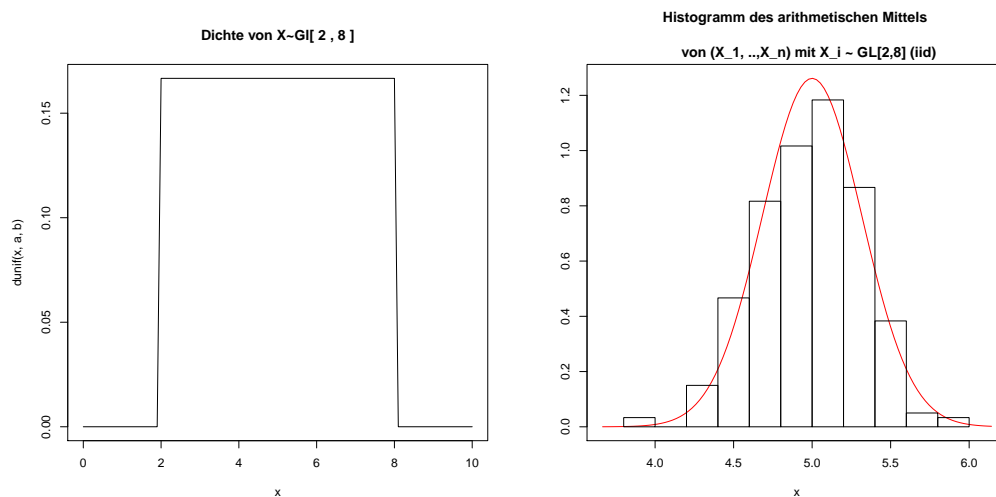
$$Z_n = \frac{\bar{X}_n - \mathbb{E}\bar{X}_n}{\sqrt{\text{Var}(\bar{X}_n)}} = \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}$$

für alle  $x \in \mathbb{R}$  gegen die Standardnormalverteilung, d.h. es gilt

$$\lim_{n \rightarrow \infty} P(Z_n \leq x) = \lim_{n \rightarrow \infty} P\left(\frac{\bar{X}_n - \mu}{\sigma} \sqrt{n} \leq x\right) = \Phi(x).$$

Für große Anzahl  $n$  ist somit das standardisierte arithmetische  $Z_n \approx \mathcal{N}(0, 1)$  näherungsweise standardnormalverteilt. Daher ist das arithmetische Mittel  $\bar{X}_n$  für große  $n$  näherungsweise  $\mathcal{N}(\mu, \frac{\sigma^2}{n})$  verteilt und die Summe  $X_1 + \dots + X_n$  für große  $n$  näherungsweise  $\mathcal{N}(n\mu, n\sigma^2)$  verteilt.

Hier wählen wir zur Illustration die Gleichverteilung und mitteln über  $n = 30$  unabhängige und gleichverteilte Zufallsgrößen. Die linke Grafik zeigt die Dichtefunktion der Gleichverteilung und die rechte Grafik die Verteilung des arithmetischen Mittels.



## 4 Induktive Statistik

Die induktive Statistik bezeichnet man auch als Inferenzstatistik, schließende oder beurteilende Statistik, deren Ziel es ist, Rückschlüsse von repräsentativen Stichproben auf Zufallsmodelle zu ziehen. Diese werden mit Hilfe von Wahrscheinlichkeitsräumen und Zufallsvariablen, die wir in Abschnitt 2 näher kennengelernt haben, modelliert. Mit Kenntnis des Zufallsmodells kann man dann Wahrscheinlichkeiten interessierender Ereignisse oder wichtige Kenngrößen der Verteilung berechnen, um Erkenntnisse über den Zufallscharakter zu gewinnen. Modelliert man beispielsweise die Einnahmen eines Unternehmens durch eine normalverteilte Zufallsvariable  $X \sim \mathcal{N}(\mu, \sigma^2)$ , so müssen die im Allgemeinen unbekannt Parameter  $\mu$  und  $\sigma$  aus den Daten einer Stichprobe geschätzt werden. Als Stichprobe  $x_1, \dots, x_n$  könnte man beispielsweise die Einnahmen der letzten Quartale oder die Einnahmen vergleichbarer Unternehmen heranziehen.

Allgemeines Ziel der schließenden Statistik ist der Rückschluss von der repräsentativen Stichprobe  $x_1, \dots, x_n$  auf die Verteilung der Zufallsgröße  $X$ . Diese konkrete Stichprobe  $(x_1, \dots, x_n)$  interpretiert man als Realisierung des Zufallsvektors

$$(X_1, X_2, \dots, X_n), \quad (24)$$

dessen Komponenten  $X_i$  unabhängige Zufallsgrößen sind, die identisch wie die Zufallsgröße  $X$ , über deren Verteilung wir Informationen aus der Stichprobe gewinnen wollen, verteilt sind. Den Zufallsvektor (24) bezeichnet man auch als **mathematische Stichprobe**. Die Annahme der unabhängigen und identischen Verteilung der Komponenten einer mathematischen Stichprobe ist wesentlich, da diese Repräsentativität und Unabhängigkeit von Stichproben gewährleisten. Betrachtet man ein Zufallsmodell mit bekannter Verteilung, die nur von Parametern abhängt, spricht man von parametrischer Statistik, andernfalls von nichtparametrischer Statistik. Im Folgenden sollen die Methoden der parametrischen Statistik besprochen werden, die Rückschlüsse über die unbekannt Parameter einer Zufallsvariablen zulassen.

### 4.1 Punktschätzer

In erster Linie ist man an einem konkreten Schätzwert für den Parameter interessiert, so dass man typischerweise von Punktschätzern spricht. Interessiert man sich beispielsweise für den unbekannt Parameter  $\mu$  eines normalverteilten Merkmals  $X \sim \mathcal{N}(\mu, \sigma^2)$ , so kann man wegen  $\mu = \mathbb{E}X$  und dem Gesetz der großen Zahlen aus der Stichprobe  $x_1, \dots, x_n$  aus dem Stichprobenmittel

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

den Parameter schätzen. Würde man eine andere Stichprobe  $\tilde{x}_1, \dots, \tilde{x}_n$  zur Schätzung verwenden, würde man entsprechend einen anderen Schätzwert für den Parameter  $\mu$  erhalten, so dass die Schätzer selbst zufällig sind. Genauer genommen sind die Schätzer als Funktionen der mathematischen Stichprobe (24) Zufallsvariablen. So gilt beispielsweise für obige Schätzfunktion des Parameters  $\mu$

$$\hat{\mu} = \hat{\mu}(X_1, \dots, X_n) = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i. \quad (25)$$

Setzt man in die Schätzfunktion (25) die konkrete Stichprobe  $(x_1, \dots, x_n)$  ein, so erhält man auch den konkreten Schätzwert, der typischerweise ebenso mit  $\hat{\mu}$  bezeichnet wird<sup>12</sup>.

### Beispiel 4.1: Alternative Schätzfunktionen

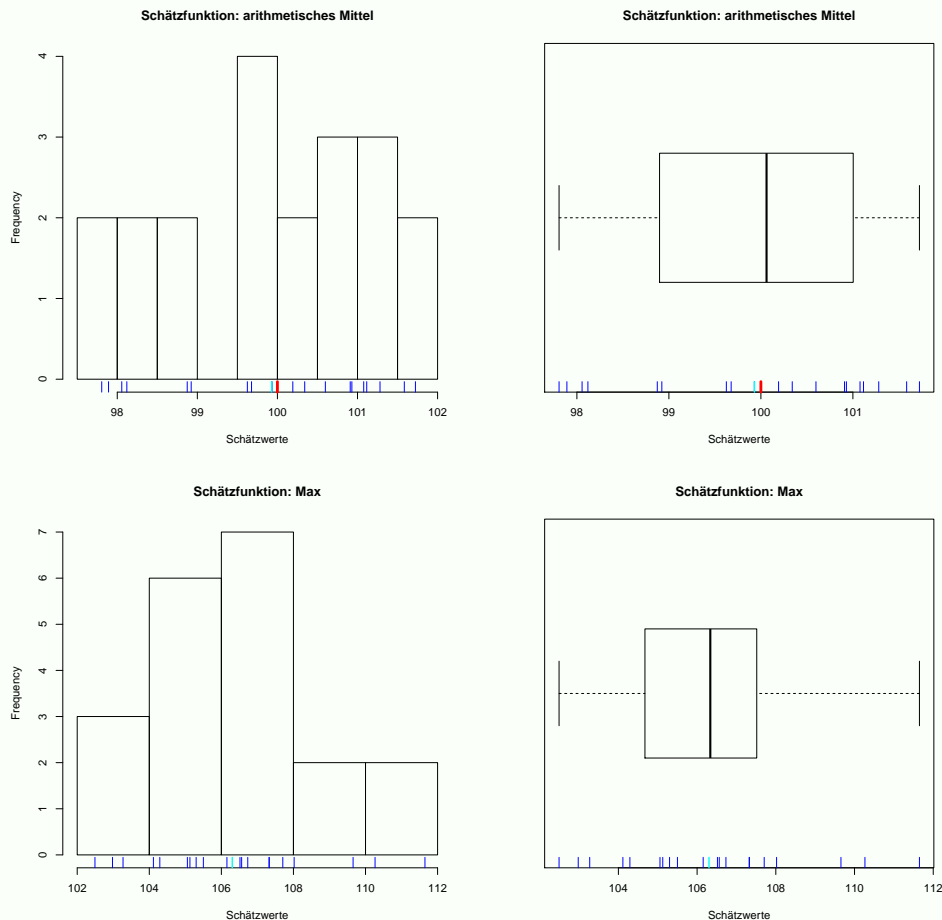
Zur Illustration ziehen wir insgesamt 20 Stichproben vom Umfang  $n = 10$  der Normalverteilung mit Parametern  $\mu = 100$  und  $\sigma^2 = 20^2$  und berechnen für jede dieser Stichproben den Schätzer

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$$

für den normalerweise unbekanntem Parameter  $\mu$ . Anschließend zeichnen wir für alle 20 berechneten arithmetischen Stichprobenmittel ein Histogramm sowie Boxplot der Schätzwerte und fügen in rot den wahren Parameter und in cyan das Mittel aller Schätzer hinzu. Neben dem arithmetischen Mittel betrachten wir noch die Schätzfunktion

$$\hat{\mu}_M = \max \{X_1, X_2, \dots, X_n\},$$

um die Güte der Schätzfunktionen zu vergleichen und beurteilen zu können.



Aus den 20 Simulationen ist klar zu erkennen, dass der Schätzer

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$$

im Mittel richtig schätzt, da der mittlere Schätzwert 99.93 mit dem wahren Parameter 100 fast übereinstimmt. Hingegen ist deutlich zu erkennen, dass die Schätzfunktion

$$\hat{\mu}_M = \max \{X_1, X_2, \dots, X_n\}$$

verzerrt schätzt, da die Schätzwerte im Mittel (106.3 > 100) deutlich zu groß sind.

<sup>12</sup>Somit hat die Bezeichnung  $\hat{\mu}$  eine Doppeldeutung. Allerdings ist üblicherweise aus dem Kontext heraus klar, ob der konkrete Schätzwert einer konkreten Stichprobe  $(x_1, \dots, x_n)$  oder der zufällige Schätzer als Funktion der mathematischen Stichprobe  $(X_1, \dots, X_n)$  gemeint ist.

Prinzipiell gibt es keine Einschränkungen bei der Wahl der Schätzfunktion  $\hat{\theta}$  für einen unbekanntem Parameter  $\theta$ , allerdings möchte man natürlich eine gewisse Qualität beim Schätzen gewährleisten. Wollen wir beispielsweise den Parameter  $\mu$  der Normalverteilung schätzen, können wir natürlich nicht erwarten, dass die Schätzwerte  $\hat{\mu}$  jeder Stichprobe  $(x_1, \dots, x_n)$  mit dem unbekanntem exakten Wert  $\mu$  übereinstimmen.

Um aber Vertrauen in die Schätzer zu haben, sollte der Schätzer möglichst **erwartungstreu** sein, also im Mittel richtig schätzen. Diese Forderung wird durch die Gleichung

$$\mathbb{E}\hat{\mu} = \mu$$

beschrieben. Eine systematische Abweichung wird als Bias

$$\text{bias}(\hat{\mu}) = \mathbb{E}\hat{\mu} - \mu$$

bezeichnet. Der oben diskutierte Schätzer (25) für den Parameter  $\mu$  ist wegen

$$\mathbb{E}\hat{\mu} = \mathbb{E}\bar{X}_n = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}X_i = \mu$$

tatsächlich erwartungstreu, schätzt also im Mittel korrekt, so dass der Bias  $\text{bias}(\hat{\mu}) = 0$  verschwindet. Bei dem ersten Schätzer aus Beispiel 4.1 haben wir die Erwartungstreue gesehen, während der zweite Schätzer einen positiven Bias aufwies, da die Schätzwerte im Mittel deutlich zu hoch geschätzt wurden. Neben der Erwartungstreue, sind weitere Eigenschaften zur Beurteilung der Güte eines Schätzers von Bedeutung. Von den Betrachtungen des arithmetischen Mittels  $\bar{X}$  in Abschnitt 3 ist überdies bekannt, dass die Varianz (21) wegen

$$\text{Var}(\hat{\mu}) = \text{Var}(\bar{X}) = \frac{\sigma^2}{n} \xrightarrow{n \rightarrow \infty} 0$$

für hinreichend große Stichproben immer kleiner wird. Die Varianz  $\text{Var}(\hat{\mu})$  beschreibt, wie stark die einzelnen Schätzwerte um den mittleren Schätzwert schwanken. Ist der Schätzer erwartungstreu, so beschreibt die Varianz des Schätzers entsprechend die Schwankung der Schätzwerte  $\hat{\theta}$  um den gesuchten Parameter  $\theta$ . Hat man zwei Schätzer  $\hat{\theta}_1$  und  $\hat{\theta}_2$  mit dem selben Schätzfehler ( $\text{bias}(\hat{\theta}_1) = \text{bias}(\hat{\theta}_2)$ ), so wählt man den mit der geringeren Schwankung.

#### Beispiel 4.2: Varianz der Schätzfunktionen

Für eine Stichprobe vom Umfang  $n$  können wir neben dem herkömmlichen Schätzer  $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$  für den Erwartungswert  $\mu = \mathbb{E}X$  auch den Schätzer

$$\hat{\mu}_m = \frac{1}{m} \sum_{i=1}^m X_i$$

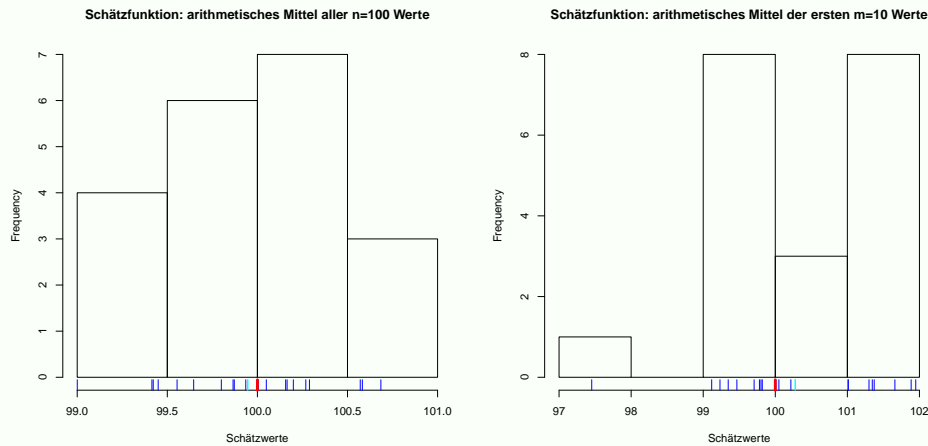
für ein beliebiges  $1 \leq m < n$  betrachten. Beide Schätzer sind erwartungstreu, allerdings ist wegen

$$\text{Var}(\hat{\mu}) = \frac{\sigma^2}{n} < \frac{\sigma^2}{m} = \text{Var}(\hat{\mu}_m)$$

mit  $\sigma^2 = \text{Var}(X)$  der herkömmliche Schätzer zu bevorzugen, da er eine kleinere Varianz aufweist.

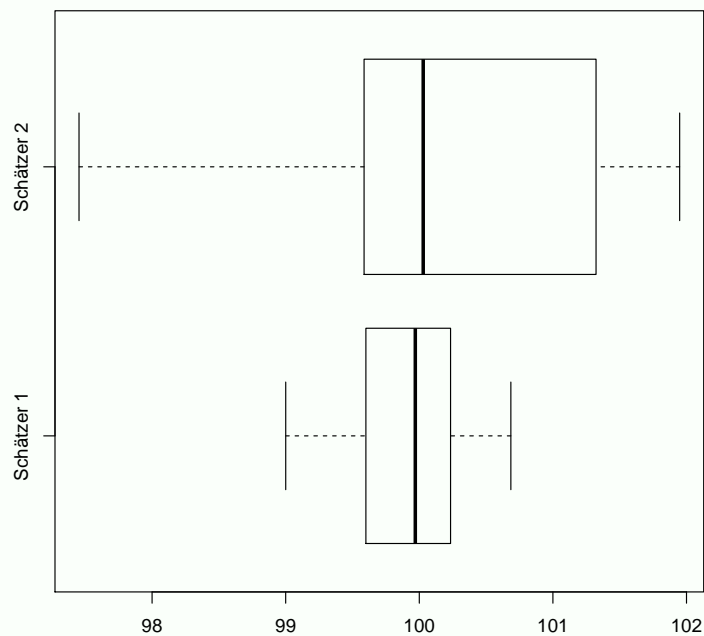
Eine kleine Simulation mit jeweils 20 Stichproben vom Umfang  $n = 100$  eines normalverteilten Merkmals  $X \sim \mathcal{N}(100, 16)$  soll dies demonstrieren. Wir bestimmen für beide Schätzer jeweils die Schätzwerte beider Schätzfunktionen  $\hat{\mu}$  und  $\hat{\mu}_{10}$  und stellen die Ergebnisse grafisch dar. Der zweite Schätzer mittelt anstatt über alle  $n = 100$  Stichprobenwerte jeweils nur die ersten  $m = 10$ .





Beide Schätzfunktionen schätzen im Mittel richtig, allerdings streuen die Schätzwerte<sup>a</sup> des zweiten Schätzers viel mehr als die Werte des ersten Schätzers. Der abschließende Boxplot veranschaulicht nochmals die Schwankung beider Schätzer und zeigt deutlich, dass die Varianz der ersten Schätzfunktion  $\hat{\mu} = \hat{\mu}_{100}$  kleiner ist, als die Varianz des zweiten Schätzers  $\hat{\mu}_{10}$ .

### Streuung beider (erwartungstreuen) Schätzer



<sup>a</sup>In dunkelblau sind die konkreten Schätzwerte auf der x-Achse abgetragen. In hellblau ist das Mittel aller Schätzer dargestellt, dass wegen der Erwartungstreue in beiden Fällen nah am exakten Wert  $\mu = 100$  liegt, der in rot abgetragen ist.

Bevor wir uns dem Schätzer für die Binomialverteilung widmen, wollen wir kurz motivieren, warum wir uns im nächsten Abschnitt 4.2 Konfidenzintervallen widmen. Dafür nehmen wir der Einfachheit halber an, dass das Merkmal  $X \sim \mathcal{N}(\mu, \sigma^2)$  für das wir mittels dem Punktschätzer

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

aus der Stichprobe  $(X_1, \dots, X_n)$  den unbekannt Parameter  $\mu$  schätzen. Aufgrund der Normalverteilung und Unabhängigkeit aller Stichprobenwerte  $X_i$  folgt  $\hat{\mu} = \bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$  und daher gilt  $P(\hat{\mu} = \mu) = 0$ . Letzteres besagt, dass die Wahrscheinlichkeit, dass der konkrete Schätzwert mit dem tatsächlichen Parameter  $\mu$  exakt übereinstimmt, Null ist. Aus diesem Grund betrachtet man neben Punktschätzern, die genau einen Punkt schätzen, Konfidenzintervalle, die den unbekannt Parameter enthalten. Wir werden dieses Konzept in Abschnitt 4.2 näher besprechen.

### Beispiel 4.3: Flugüberbuchungen

Fluggesellschaften überbuchen häufig ihre angebotenen Flüge, d.h. sie verkaufen mehr Tickets, als tatsächlich Plätze vorhanden sind, um ihren Gewinn zu steigern. Hierfür nutzen sie die Tatsache aus, dass häufig gebuchte Flüge nicht genutzt werden. Ist dann tatsächlich ein Flug überbucht, werden Fluggästen attraktive Prämien geboten, wenn sie einen Flug später fliegen. Das Risiko der Überbuchung lässt sich für die Fluggesellschaften abschätzen. Für die  $m$  verkauften Flugtickets beschreibt die Zufallsgröße<sup>a</sup>  $X \sim \text{Bin}(m, p)$  die Anzahl der tatsächlich besetzten Plätze. Aus den Flugdaten der letzten  $n = 15$  Flüge ist die Anzahl der besetzten Plätze

(91, 92, 91, 89, 90, 82, 94, 92, 90, 95, 89, 94, 90, 91, 88)

bei jeweils  $m = 100$  verkauften Plätzen bekannt. Aus dieser Stichprobe soll der Parameter  $p$ , der die Rate der angetretenen Flüge beschreibt, geschätzt werden.

Da für die binomialverteilte Zufallsgröße  $X \sim \text{Bin}(m, p)$  für den Erwartungswert  $\mathbb{E}X = mp$  gilt, und uns  $m = 100$  bekannt ist, schätzen wir die Rate

$$\hat{p} = \frac{\bar{X}}{m}$$

mit Hilfe des arithmetischen Stichprobenmittels und erhalten für unsere konkrete Stichprobe

$$\hat{p} = \frac{1}{100} \cdot \frac{1358}{15} = 0.9053.$$

Angenommen die Fluggesellschaft setzt einen Flieger mit 95 Plätzen ein. Dann könnte man nun mittels  $\hat{p}$  die Wahrscheinlichkeit berechnen, dass der kommende Flug überbucht wird

$$P(X > 95) = 1 - P(X \leq 94) = 1 - F(94) = 0.0345.$$

<sup>a</sup>Gehen wir davon aus, dass jeder Passagier unabhängig von den anderen Passagieren entscheidet, ob der gebuchte Flug wahrgenommen wird, ist die Anzahl der tatsächlich besetzten Plätze binomialverteilt. Bei Flügen, die hauptsächlich von allein reisenden Geschäftsleuten gebucht werden, ist dies eine realistische Annahme. Werden die Flüge hingegen hauptsächlich von Familien gebucht, die gemeinsam verreisen, kann die Unabhängigkeitsannahme nicht mehr gehalten werden und die tatsächliche Verteilung der wahrgenommenen Flugplätze weicht von der Binomialverteilung ab.

Zum Abschluss wollen wir Schätzer für die Varianz  $\sigma^2 = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X - \mu)^2$  diskutieren. Um erwartungstreue Schätzer zu erhalten, muss man unterscheiden, ob der Parameter  $\mu = \mathbb{E}X$  bekannt oder unbekannt ist. Im ersten Fall erhält man mit

$$\hat{V} = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

wegen

$$\mathbb{E}\hat{V} = \frac{1}{n} \mathbb{E} \sum_{i=1}^n (X_i - \mu)^2 = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i - \mu)^2 = \sigma^2$$

eine erwartungstreue Schätzfunktion für  $\sigma^2$ . Ist  $\mu = \mathbb{E}X$  hingegen unbekannt, liegt es nahe diesen unbekanntem Wert durch den Schätzer  $\hat{\mu} = \bar{X}$  aus der Stichprobe zu ersetzen. Der damit erhaltene Schätzer für die Varianz  $\sigma^2$

$$\hat{V} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

ist allerdings wegen<sup>13</sup>

$$\mathbb{E}\hat{V} = \frac{1}{n} \sum_{i=1}^n \mathbb{E}X_i^2 - \mathbb{E}\bar{X}^2 = \sigma^2 + \mu^2 - (\mu^2 + \frac{\sigma^2}{n}) = \sigma^2 - \frac{\sigma^2}{n} = \sigma^2 \frac{n-1}{n}$$

nicht erwartungstreu. Der systematische Schätzfehler  $\text{bias}(\hat{V}) = -\sigma^2/n$  zeigt an, dass der Schätzer  $\hat{V}$  die Varianz  $\sigma^2$  im Mittel etwas zu niedrig schätzt, sich dieser Fehler allerdings mit zunehmender Stichprobengröße immer weiter verringert. Durch Skalierung

$$\hat{S}^2 = \frac{n}{n-1} \hat{V} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

<sup>13</sup>Wir nutzen  $\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$  sowie  $\text{Var}(X_i) = \sigma^2 = \mathbb{E}X_i^2 - \mu^2$  und  $\text{Var}(\bar{X}) = \sigma^2/n = \mathbb{E}\bar{X}^2 - \mu^2$ .

erhält man dann eine erwartungstreue Schätzfunktion für die Varianz bei unbekanntem  $\mu = \mathbb{E}X$ .

## 4.2 Konfidenzintervalle

Wie in Abschnitt 4.1 bereits besprochen, beträgt die Wahrscheinlichkeit Null, dass der Punktschätzer den exakten Parameter trifft. Anstatt also einen konkreten Schätzwert zu bestimmen, versucht man ein Intervall anzugeben, in dem man den unbekannt Parameter mit einer gewissen Sicherheit vermutet. Hierfür benötigt man allerdings eine Verteilung des zufälligen Schätzers. Entspricht beispielsweise das zu untersuchende Merkmal  $X \sim \mathcal{N}(\mu, \sigma^2)$  einer Normalverteilung, so ist auch der Schätzer  $\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$  der mathematischen Stichprobe  $(X_1, \dots, X_n)$  normalverteilt. In Abschnitt 4.2.1 werden wir daher zunächst Konfidenzintervalle für den Parameter  $\mu = \mathbb{E}X$  der Normalverteilung besprechen.

Für binomialverteilte Merkmale  $X \sim \text{Bin}(m, p)$  werden wir in Abschnitt 4.2.2 Konfidenzintervalle für den Parameter  $p$  betrachten. Allerdings basiert deren Berechnung auf der Binomialverteilung und ist recht umständlich durchzuführen, so dass man für große Stichproben mit Hilfe der Normalapproximation diese einfacher berechnen kann.

### 4.2.1 Konfidenzintervalle für den Erwartungswert der Normalverteilung

Entspricht das zu untersuchende Merkmal  $X \sim \mathcal{N}(\mu, \sigma^2)$  einer Normalverteilung mit bekannter Varianz  $\sigma^2$ , so ist auch der Schätzer  $\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$  der mathematischen Stichprobe  $(X_1, \dots, X_n)$  normalverteilt. Da  $\sigma$  bekannt ist, folgt die Standardisierung<sup>14</sup> einer Standardnormalverteilung

$$Z = \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Mit Hilfe der Quantile<sup>15</sup>  $z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$  sowie  $z_{1-\frac{\alpha}{2}}$  der Standardnormalverteilung, beschreibt das Intervall  $[-z_{1-\frac{\alpha}{2}}, z_{1-\frac{\alpha}{2}}]$  wegen

$$P(-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}) = \Phi(z_{1-\frac{\alpha}{2}}) - \Phi(-z_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} - (1 - (1 - \frac{\alpha}{2})) = 1 - \alpha$$

ein symmetrisches Intervall um die Null, dass mit der Wahrscheinlichkeit  $1 - \alpha$  die um die Null streuenden Werte der standardisierten Zufallsgröße  $Z \sim \mathcal{N}(0, 1)$  enthält.

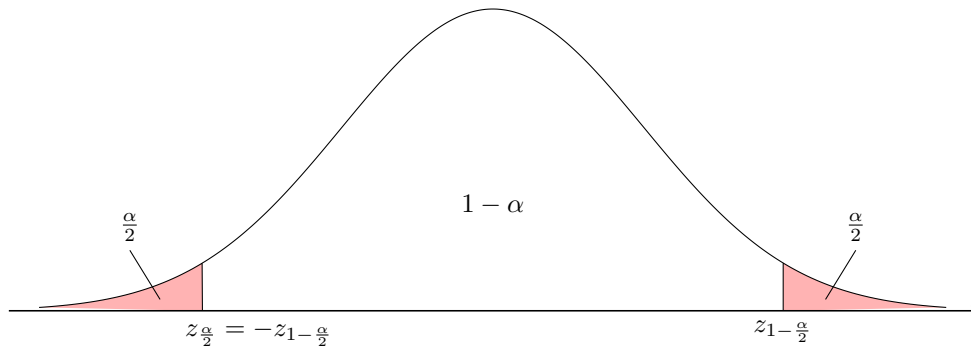


Abbildung 6: Quantile der Standardnormalverteilung zur Herleitung von Konfidenzbereichen

<sup>14</sup>In den meisten Anwendungen ist die Varianz  $\sigma^2 = \text{Var}(X)$  unbekannt, so dass man beim Standardisieren  $\sigma$  durch die empirische Standardabweichung

$$\hat{S} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

ersetzt. Dann ist die  $\sqrt{n} \frac{\bar{X} - \mu}{\hat{S}}$  nicht mehr standardnormalverteilt sondern folgt einer t-Verteilung mit  $n - 1$  Freiheitsgraden. Daher werden in dem Fall der unbekannt Varianz  $\sigma^2$  in den Formeln der Konfidenzintervalle die Quantile der Standardnormalverteilung durch die Quantile der t-Verteilung mit  $n - 1$  Freiheitsgraden ersetzt.

<sup>15</sup>Für das Quantil  $z_q = \Phi^{-1}(q)$  der Ordnung  $q$  der standardnormalverteilten Zufallsgröße  $Z \sim \mathcal{N}(0, 1)$  gilt  $\Phi(z_q) = P(Z \leq z_q) = \Phi(z_q) = \Phi(\Phi^{-1}(q)) = q$ .

Abbildung 6 veranschaulicht dieses Intervall. Stellen<sup>16</sup> wir nun dieses Intervall nach dem unbekanntem Parameter  $\mu$  um, erhalten wir

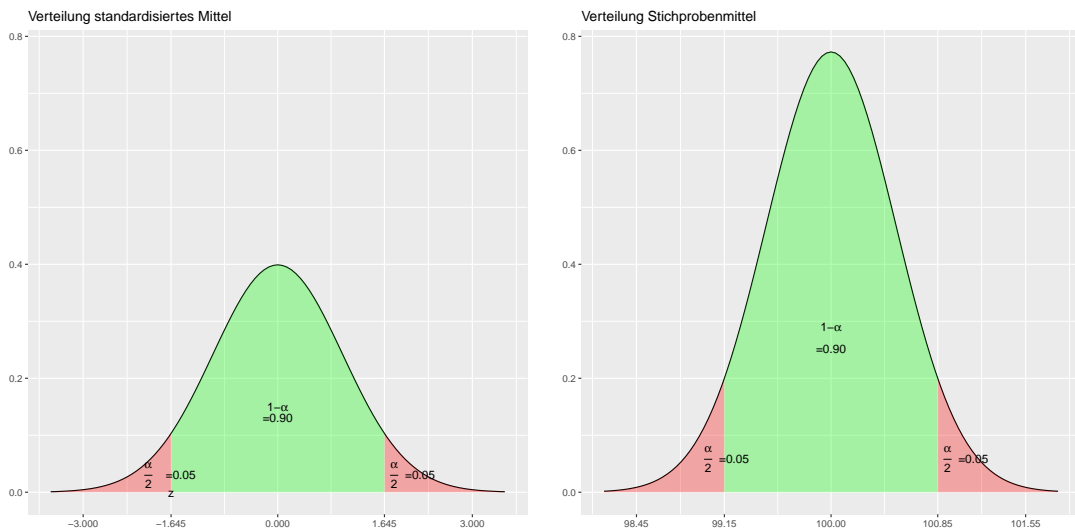
$$P\left(\bar{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha. \quad (26)$$

Dies bedeutet, dass die Wahrscheinlichkeit, dass sich der Parameter  $\mu$  im Intervall

$$\left[\bar{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right] \quad (27)$$

befindet, gerade  $1 - \alpha$  entspricht.

Letztendlich basiert das Konfidenzintervall (27) auf dem Konfidenzintervall  $[-z_{1-\frac{\alpha}{2}}, z_{1-\frac{\alpha}{2}}]$  der Standardnormalverteilung des standardisierten Mittels  $Z = \sqrt{n} \frac{\bar{X} - \mu}{\sigma}$ , welches wir nach dem unbekanntem Parameter  $\mu$  umgestellt haben. Da das standardisierte Mittel nur eine lineare Transformation darstellt, ist die Verteilung des Stichprobenmittels  $\bar{X} \sim \mathcal{N}(\mu, \sigma^2)$  lediglich eine einfache Skalierung und Verschiebung der Standardnormalverteilung. Für  $\alpha = 0.1$  wollen wir das für den Parameter  $\mu = 100$  des Merkmals  $X \sim \mathcal{N}(\mu, \sigma^2)$  mit bekanntem Parameter  $\sigma = 2$  für Stichproben vom Umfang  $n = 15$  grafisch illustrieren. Die linke Grafik zeigt die Dichte der Standardnormalverteilung und die entsprechenden Quantile  $z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}} = -z_{0.95} = -1.6449$  sowie  $z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.6449$  der Standardnormalverteilung, so dass die äußeren roten Flächen jeweils einer Wahrscheinlichkeit von  $\frac{\alpha}{2} = 0.05$  und die grüne Fläche der Wahrscheinlichkeit von  $1 - \alpha = 0.9$  entsprechen. Die rechte Grafik zeigt die transformierte Verteilung des Stichprobenmittels  $\bar{X} \sim \mathcal{N}(100, 4)$ , die wir natürlich hier nur zeichnen können, da der eigentlich unbekannt Parameter  $\mu = 100$  uns an dieser Stelle natürlich bekannt ist<sup>17</sup>.



Bei der Interpretation des Konfidenzintervalls (27) ist zu beachten, dass die obere und untere Grenze des Intervalls Zufallsvariablen sind. Ersetzt man das arithmetische Mittel  $\bar{X}$  der mathematischen Stichprobe  $(X_1, \dots, X_n)$  durch das arithmetische Mittel  $\bar{x}$  einer konkreten Stichprobe  $(x_1, \dots, x_n)$ , so erhält man eine konkrete Schätzung des Konfidenzintervalls.

#### Beispiel 4.4: Konfidenzintervall für den Parameter $\mu$ der Normalverteilung bei bekannter Varianz

Für das Merkmal  $X \sim \mathcal{N}(\mu, \sigma^2)$  mit  $\sigma^2 = 1$  sollen aus der Stichprobe  $(x_1, \dots, x_{10})$  der Parameter  $\mu$  bestimmt sowie ein Konfidenzintervall angegeben werden. Das arithmetische Mittel  $\bar{x} = 14.6$  der Stichprobe liefert den Punktschätzer  $\hat{\mu} = 14.6$ . Zur Bestimmung des 95%-igen Konfidenzintervalls wählen wir  $\alpha = 0.05$  und nutzen das entsprechende Quantile der Normalverteilung  $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.960$ . Somit erhalten wir als Schätzung der unteren Intervallgrenze  $\bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 14.6 - 1.96 \cdot \frac{1}{\sqrt{10}} = 13.98$  und für die obere Intervallgrenze

<sup>16</sup>Die Ungleichung  $-z_{1-\frac{\alpha}{2}} \leq Z \leq z_{1-\frac{\alpha}{2}}$  entspricht  $-z_{1-\frac{\alpha}{2}} \leq \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \leq z_{1-\frac{\alpha}{2}}$  und ist äquivalent zu  $-\bar{X} - z_{1-\frac{\alpha}{2}} + \frac{\sigma}{\sqrt{n}} \leq -\mu \leq -\bar{X} + z_{1-\frac{\alpha}{2}} + \frac{\sigma}{\sqrt{n}}$ .

<sup>17</sup>Somit können wir die linke Grafik immer zeichnen, die rechte aber nur, wenn wir den unbekannt Parameter  $\mu$ , den wir schätzen wollen, kennen. Beide Grafiken sind aber äquivalent, so dass mit Hilfe der linken Grafik ein Konfidenzintervall für  $\mu$  bestimmt werden kann.

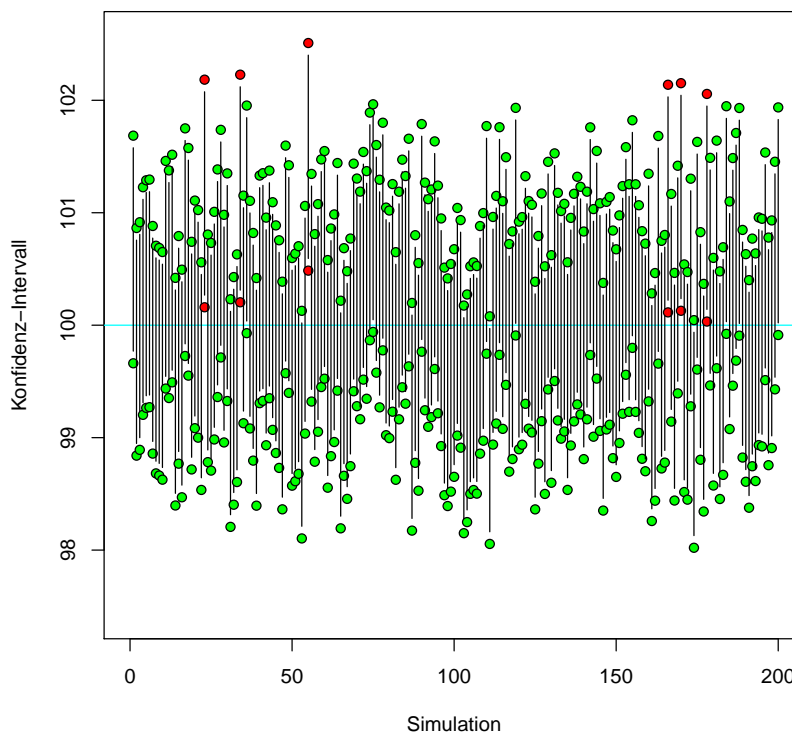
$\bar{x}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 14.6 + 1.96 \cdot \frac{1}{\sqrt{10}} = 15.22$ . Somit erhalten wir das 95%-ige Konfidenzintervall für den Parameter  $\mu$  basierend auf der konkreten vorliegenden Stichprobe

$$KI_{95\%}^{\mu} = [13.98, 15.22]$$

Das konkrete Intervall  $\left[ \bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$  enthält den tatsächlichen Parameter oder nicht, daher ist die richtige Interpretation, dass der exakte Parameter  $\mu$  in diesem konkreten Intervall  $\left[ \bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$  mit einer Wahrscheinlichkeit von  $1 - \alpha$  enthalten ist, falsch. Die Interpretation des konkreten Konfidenzintervalls ist daher nicht so trivial. Durch die Konstruktion des zufälligen Konfidenzintervalls ergibt sich aber folgende Deutung. Würden wir nicht nur eine, sondern sehr viele Stichproben auswerten, dann wird der tatsächliche Parameter in etwa  $(1 - \alpha) \cdot 100$  Prozent aller Intervalle enthalten sein. Dies bedeutet allerdings auch, dass die restlichen  $\alpha \cdot 100$  Prozent aller Intervalle den wahren Parameter  $\mu$  nicht enthalten. Typischerweise haben wir allerdings nur eine Stichprobe zur Verfügung, die entweder den Parameter enthält oder nicht. Nach der Konstruktionsweise von Konfidenzintervallen erhoffen wir natürlich, dass das berechnete Intervall dieser Stichprobe zur ersten Kategorie zählt, die den wahren Parameter tatsächlich enthalten.

Dies illustrieren wir anhand einer kleinen Simulation, bei der jeweils 200 Stichproben vom Umfang  $n = 15$  zum Merkmal  $X \sim \mathcal{N}(100, 4)$  simuliert werden. Somit ist hier der exakte Parameter  $\mu = 100$  bekannt. Aus den 200 simulierten Stichproben berechnen wir jeweils das zugehörige konkrete Konfidenzintervall und kontrollieren, ob diese jeweils den wahren Parameter enthalten. In der folgenden Grafik sind für  $\alpha = 0.05$  die Intervalle, die  $\mu$  enthalten grün und die restlichen Intervalle, die  $\mu$  nicht enthalten, rot eingefärbt. Wegen  $\alpha = 0.05$  erwarten wir, dass 10 aller 200 Intervalle  $\mu$  nicht enthalten.

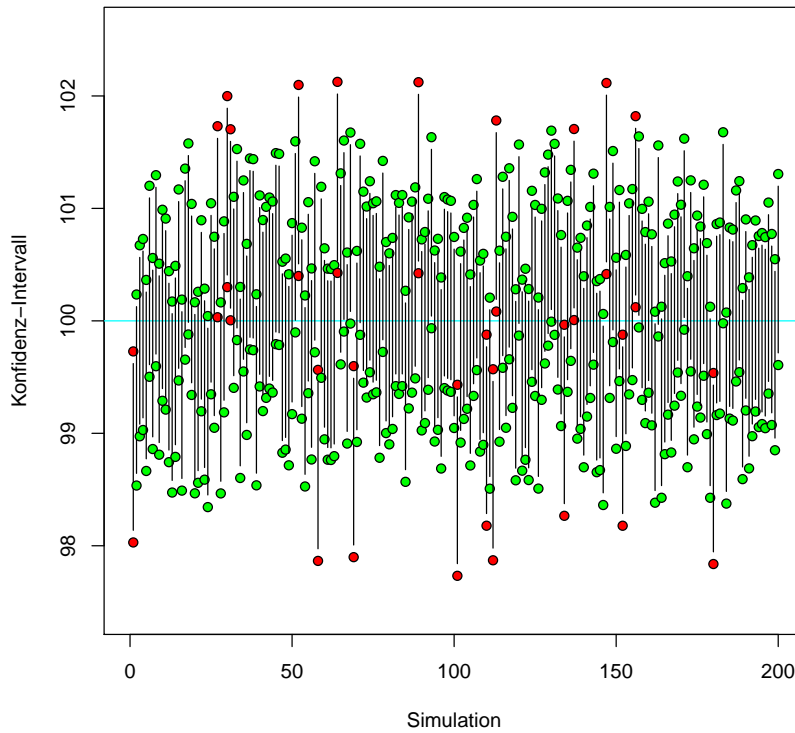
**KI alpha= 0.05 für anz= 200 verschiedene SP**



Tatsächlich überdecken 194 aller berechneten Konfidenzintervalle den wahren Parameter  $\mu$  und 6 nicht<sup>18</sup>. Erhöhen wir  $\alpha$  auf  $\alpha = 0.1$ , verringert sich das Konfidenzniveau von 0.95 auf 0.9. In dieser zweiten Simulation erwarten wir daher, dass 20 aller 200 berechneten Konfidenzintervalle  $\mu$  nicht enthalten.

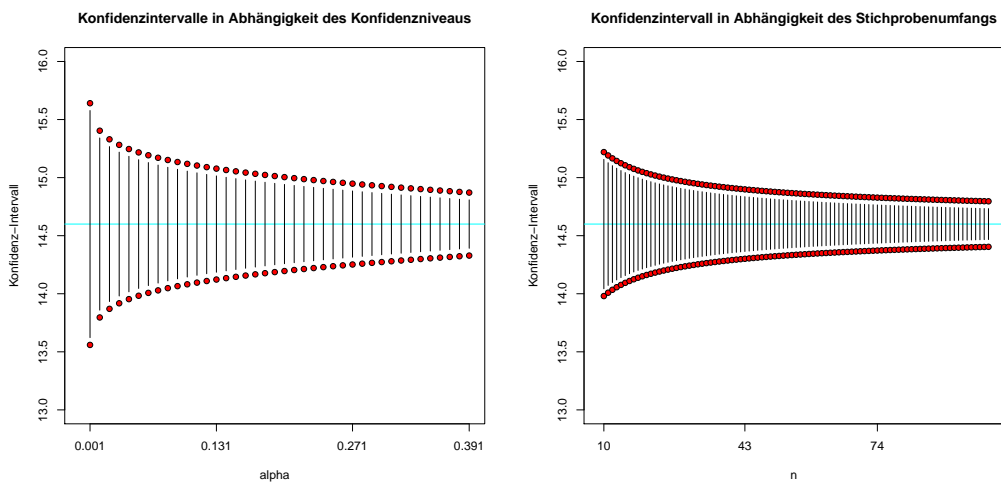
<sup>18</sup>Würden wir statt 200 deutlich mehr Stichproben untersuchen, würde sich nach dem Gesetz der großen Zahlen die Fehlerrate tatsächlich  $\alpha$  annähern.

KI alpha= 0.1 für anz= 200 verschiedene SP



Tatsächlich überdecken 19 Intervalle den wahren Parameter  $\mu$  nicht.

In einer weiteren Simulation wollen wir zeigen, dass mit wachsenden  $\alpha$  sowie mit steigender Stichprobengröße  $n$  sich die Breite der Konfidenzintervalle verringert.



**Aufgabe 4.1**

Eine Maschine schneidet Stahlbleche automatisch auf eine geforderte Länge zu. Man weiß, dass die tatsächlichen Längen minimale Abweichungen vom Normwert aufweisen können. Beobachtungen über einen längeren Zeitraum haben ergeben, dass die Standardabweichung für diesen Maschinentyp  $\sigma = 2.2$  (cm) beträgt. Es soll bei der derzeitigen Einstellung der Maschine die mittlere Länge der Stahlbleche geschätzt werden. Wie lautet das 95%ige Konfidenzintervall für den Erwartungswert  $\mu$  der Länge der Stahlbleche, wenn eine Stichprobe vom Umfang  $n = 40$  die mittlere Länge  $\bar{x} = 80.5$  (cm) ergeben hat? Gehen Sie davon aus, dass die Zufallsvariable  $X =$  „Länge eines Stahlblechs (in cm)“ normalverteilt ist.

*Lösung auf Seite ??*

In den meisten Anwendungen ist die Varianz  $\text{Var}(X) = \sigma^2$  unbekannt, so dass Formel (27) zur Bestimmung des Konfidenzintervalls für den Parameter  $\mu$  nicht verwendet werden kann. Ersetzt man

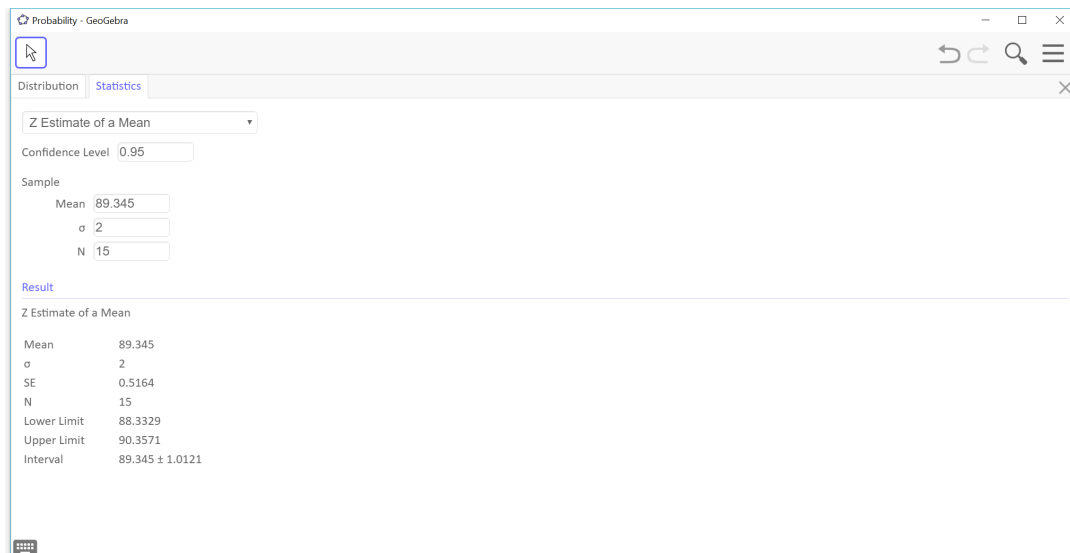


Abbildung 7: Berechnung Konfidenzintervall für Parameter  $\mu$  von  $X \sim \mathcal{N}(\mu, \sigma^2)$  für bekanntes  $\sigma$  mit Geogebra

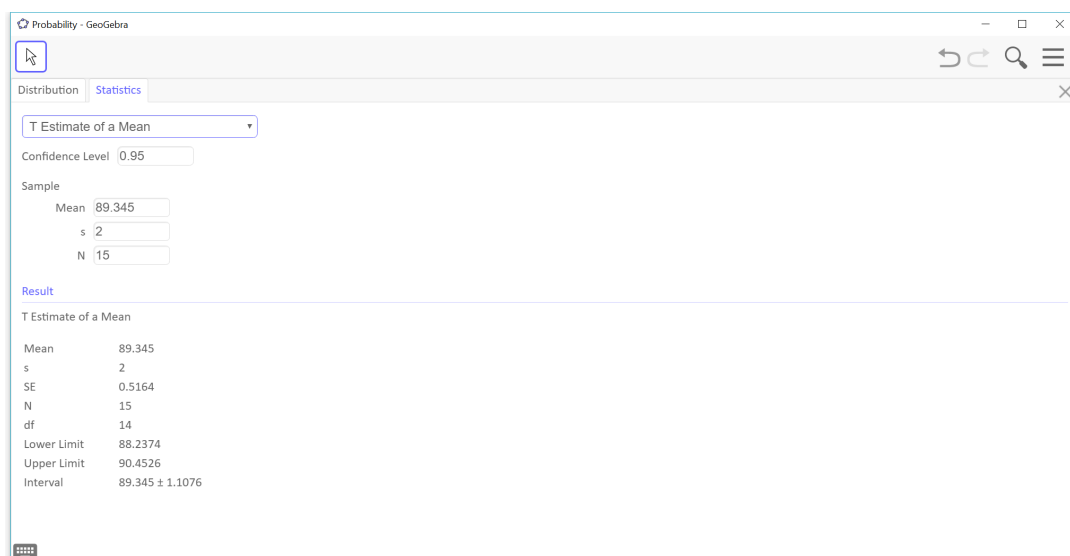


Abbildung 8: Berechnung Konfidenzintervall für Parameter  $\mu$  von  $X \sim \mathcal{N}(\mu, \sigma^2)$  für unbekanntes  $\sigma$  mit Geogebra

die unbekannte Varianz einfach durch die empirische Varianz  $\hat{S}^2$ , die  $\sigma^2$  erwartungstreu schätzt, ist die Stichprobengröße  $T = \sqrt{n} \frac{\bar{X} - \mu}{\hat{S}}$  nicht mehr standardnormalverteilt, sondern folgt einer t-Verteilung mit  $n - 1$  Freiheitsgraden. Daher lautet das Konfidenzintervall für den Parameter  $\mu$  des Merkmals  $X \sim \mathcal{N}(\mu, \sigma^2)$  im Falle der unbekanntenen Varianz einfach

$$\left[ \bar{X} - t_{1-\frac{\alpha}{2}, n-1} \frac{\hat{S}}{\sqrt{n}}, \bar{X} + t_{1-\frac{\alpha}{2}, n-1} \frac{\hat{S}}{\sqrt{n}} \right]. \quad (28)$$

Die Abbildungen 7 und 8 zeigen die Berechnung der Konfidenzintervalle des Parameters  $\mu$  mittels Geogebra.

Natürlich kann man auch Konfidenzintervalle für die Varianz  $\sigma^2$  eines normalverteilten Merkmals konstruieren. Dies soll aber hier nicht weiter betrachtet werden.

#### Aufgabe 4.2

Zur Bestimmung der Entfernung zwischen zwei Punkten wurden 25 Messungen durchgeführt und der Stichprobenmittelwert und die Stichprobenvarianz wie folgt ermittelt:

$$\bar{x} = 24.325 \text{ km}, \quad s^2 = 196 \text{ km}^2.$$

Geben Sie ein Konfidenzintervall für den Erwartungswert der Entfernung zum Konfidenz-

niveau  $1 - \alpha = 0.99$  an.

Lösung auf Seite ??

### Aufgabe 4.3

Bei einer TÜV-Prüfstelle wurde der CO-Gehalt von PKW-Abgasen, der als normalverteilt betrachtet werden kann, gemessen. Eine Stichprobe vom Umfang  $n = 30$  ergab folgende Schätzwerte für  $\mu$  und  $\sigma$  (in %):

$$\bar{x} = 3.22 \text{ und } s = 0.2025.$$

Berechnen Sie zum Konfidenzniveau  $1 - \alpha = 0.9$  ein Konfidenzintervall für den mittleren CO-Gehalt an.

Lösung auf Seite ??

### Aufgabe 4.4

Bei einer Arbeitsstudie seien folgende Werte für die Zeitdauer der Justierung eines Gerätes gemessen worden:

Minuten	3	4	5	6	8	9
Häufigkeiten	2	4	7	12	4	1

- Geben Sie erwartungstreue Schätzungen für den Erwartungswert und die Varianz der Zeitdauer an.
- Bestimmen Sie ein zweiseitiges Konfidenzintervall zum Konfidenzniveau  $1 - \alpha = 0.9$  für den Erwartungswert der Zeitdauer.
- Diskutieren Sie die notwendigen Verteilungsvoraussetzungen.

Lösung auf Seite ??

#### 4.2.2 Konfidenzintervalle für die Binomialverteilung

ist. Auch für den Parameter  $p$  der binomialverteilten Zufallsvariablen  $X \sim \text{Bin}(m, p)$  lässt sich basierend auf der unabhängigen Stichprobe  $(X_1, \dots, X_n)$  ein Konfidenzintervall bestimmen. Hierzu betrachten wir die Summe aller Stichprobenwerte, die mit dem selben Parameter  $p$  binomialverteilt

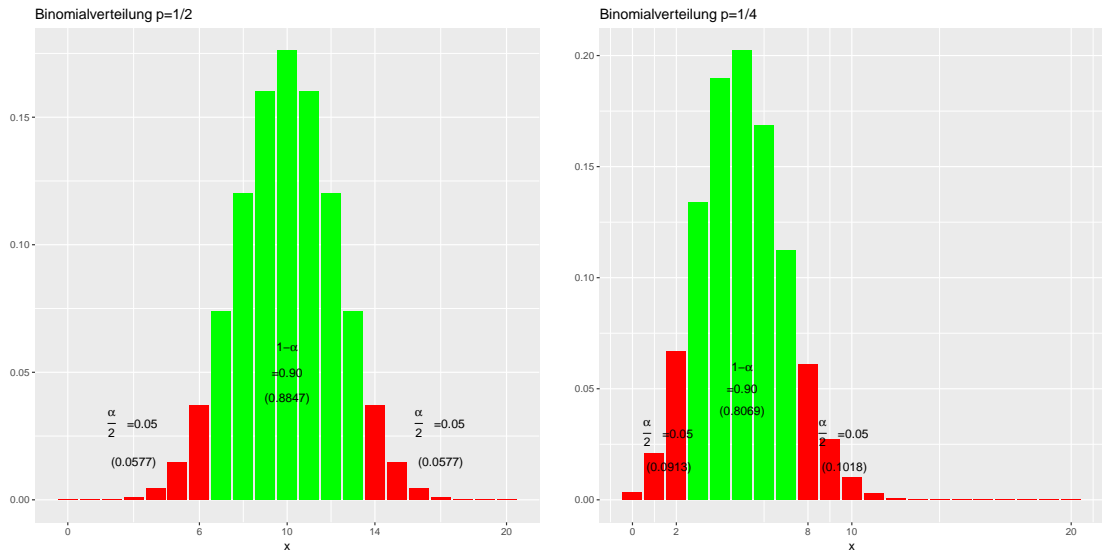
$$Y = \sum_{i=1}^n X_i = n\bar{X} \sim \text{Bin}(nm, p)$$

ist und konstruieren wieder ein symmetrisches Intervall  $[U, O]$ , das mit einer vorgegebenen Wahrscheinlichkeit  $1 - \alpha = P(p \in [U, O])$  den unbekanntem Parameter  $p$  enthält. Dieses gesuchte Intervall basiert, analog zur Herleitung des Konfidenzintervalls für den Erwartungswert der Normalverteilung auf Seite 28, auf einem symmetrischen Intervall  $[u, o]$  mit

$$P(u \leq Y \leq o) = F(o) - F(u - 1) = \sum_{j=u}^o P(Y = j) = \sum_{j=u}^o \binom{nm}{j} p^j (1-p)^{nm-j} = 1 - \alpha$$

um den Erwartungswert  $\mathbb{E}Y = nmp$  von  $Y \sim \text{Bin}(nm, p)$ . Allerdings haben wir es jetzt mit einer diskreten statt einer stetigen Verteilung zu tun. Dies hat zwei unangenehme Konsequenzen. Zum einen können wir hier nicht mehr die Quantile  $u$  und  $o$  exakt bestimmen, so dass  $P(Y < u) = P(Y > o) = \frac{\alpha}{2}$  gilt, da die Verteilungsfunktion der Binomialverteilung sprunghaft wächst und nicht zwingend den Wert  $\frac{\alpha}{2}$  annehmen muss. Dies ist in den beiden folgenden Grafiken dargestellt.





Die grünen Flächen sollen jeweils der Wahrscheinlichkeit  $1 - \alpha = 0.9$  entsprechen. Allerdings können wir die Quantile  $u$  bzw.  $o$  sinnvollerweise nur ganzzahlig wählen, so dass die tatsächliche Wahrscheinlichkeit nicht exakt  $1 - \alpha$  trifft, sondern in der linken Grafik  $0.8847$  und in der rechten Grafik sogar nur  $0.8069$  beträgt. Das weiteren können wir obiges Intervall auch nicht mehr so einfach nach  $p$  umformen. Anzumerken ist, dass beide Grafiken nur aus Illustrationszwecken unter der Kenntnis  $p = 1/2$  bzw.  $p = 1/4$  erstellt wurden, was uns natürlich bei der Bestimmung des Konfidenzintervalles nicht bekannt ist.

Das prinzipielle Vorgehen bei der Konstruktion eines Konfidenzintervalls unterscheidet sich aber nicht von dem Vorgehen, wie wir es bei der Normalverteilung kennengelernt haben. Eine Möglichkeit zur Konstruktion eines Konfidenzintervalles ist die Verwendung der Normalapproximation für große Stichproben, die wir in Abschnitt 3.1 kennengelernt haben. Diese besagt, dass für großes  $\tilde{n} = nm$  die Binomialverteilung sich sehr gut durch die Normalverteilung

$$Y \approx \mathcal{N}(nmp, nmp(1-p))$$

approximieren lässt. Folglich sind auch das Stichprobenmittel  $\bar{X} = \frac{Y}{n} \approx \mathcal{N}(mp, \frac{mp(1-p)}{n})$  und der Punktschätzer  $\hat{p} = \frac{\bar{X}}{m} = \frac{Y}{nm} \approx \mathcal{N}(p, \frac{p(1-p)}{nm})$  näherungsweise normalverteilt. Mit Hilfe der Sigma-Regel<sup>19</sup> der Normalverteilung

$$P\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\bar{X} - mp}{\sqrt{mp(1-p)}}\sqrt{n} \leq z_{1-\frac{\alpha}{2}}\right) = P\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\hat{p} - p}{\sqrt{p(1-p)}}\sqrt{nm} \leq z_{1-\frac{\alpha}{2}}\right) \approx 1 - \alpha$$

kann das asymptotische Konfidenzintervall  $[U, O]$  für den Parameter  $p$  durch Umstellung<sup>20</sup> der Ungleichungen

$$-z_{1-\frac{\alpha}{2}} \leq \frac{\bar{X} - mp}{\sqrt{mp(1-p)}}\sqrt{n} \leq z_{1-\frac{\alpha}{2}} \Leftrightarrow -z_{1-\frac{\alpha}{2}} \leq \frac{\hat{p} - p}{\sqrt{p(1-p)}}\sqrt{nm} \leq z_{1-\frac{\alpha}{2}} \quad (29)$$

nach dem unbekanntem Parameter  $p$  ein Konfidenzintervall mit der unteren Grenze

$$\begin{aligned} U &= \frac{1}{1 + \frac{1}{mn}z_{1-\frac{\alpha}{2}}^2} \left[ \frac{\bar{X}}{m} + \frac{z_{1-\frac{\alpha}{2}}^2}{2nm} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{nm}} \sqrt{\frac{\bar{X}}{m} \left(1 - \frac{\bar{X}}{m}\right) + \frac{z_{1-\frac{\alpha}{2}}^2}{4nm}} \right] \\ &= \frac{1}{1 + \frac{1}{mn}z_{1-\frac{\alpha}{2}}^2} \left[ \hat{p} + \frac{z_{1-\frac{\alpha}{2}}^2}{2nm} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{nm}} \sqrt{\hat{p}(1-\hat{p}) + \frac{z_{1-\frac{\alpha}{2}}^2}{4nm}} \right] \end{aligned}$$

und der oberen Grenze

$$\begin{aligned} O &= \frac{1}{1 + \frac{1}{mn}z_{1-\frac{\alpha}{2}}^2} \left[ \frac{\bar{X}}{m} + \frac{z_{1-\frac{\alpha}{2}}^2}{2nm} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{nm}} \sqrt{\frac{\bar{X}}{m} \left(1 - \frac{\bar{X}}{m}\right) + \frac{z_{1-\frac{\alpha}{2}}^2}{4nm}} \right] \\ &= \frac{1}{1 + \frac{1}{mn}z_{1-\frac{\alpha}{2}}^2} \left[ \hat{p} + \frac{z_{1-\frac{\alpha}{2}}^2}{2nm} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{nm}} \sqrt{\hat{p}(1-\hat{p}) + \frac{z_{1-\frac{\alpha}{2}}^2}{4nm}} \right] \end{aligned}$$

<sup>19</sup>Siehe Beispiel 2.8.

<sup>20</sup>Hierfür ist das Lösen einer quadratischen Gleichung nötig.

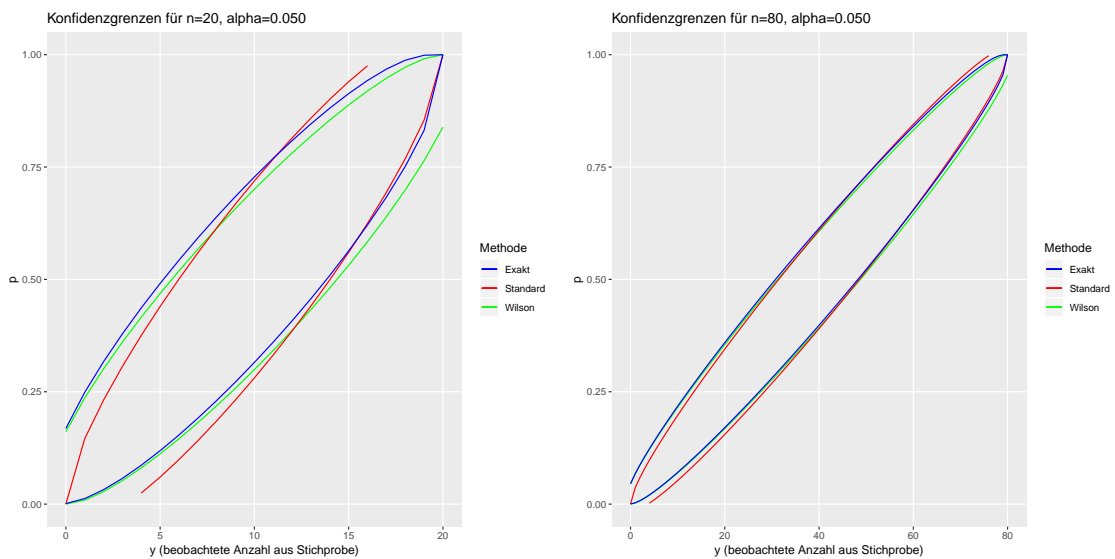
konstruiert werden. Diese Methode bezeichnet man auch als Wilson-Methode, benannt nach Wilson, der erstmalig dieses asymptotische Konfidenzintervall angab. Ein viel einfacher zu konstruierendes, aber auch unzuverlässigeres, Konfidenzintervall, das sogenannte Standard-Intervall

$$\left[ \hat{p} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{nm}} \sqrt{\hat{p}(1-\hat{p})}, \hat{p} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{nm}} \sqrt{\hat{p}(1-\hat{p})} \right]$$

erhält man ebenso aus der Normalapproximation, indem das Intervall (29) nach dem Schätzer  $\hat{p}$  umgestellt

$$p - z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{nm}} \leq \hat{p} \leq p + z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{nm}}$$

und in den Grenzen der unbekannt Parameter  $p$  einfach durch den Schätzer  $\hat{p}$  ersetzt wird. Mittels der Clopper-Pearson-Methode, basierend auf dem Zusammenhang zwischen der Binomial- und der Beta-Binomialverteilung, können exakte Intervallgrenzen hergeleitet werden. Die folgende Grafik zeigt für alle möglichen konkreten Stichprobenergebnisse  $y = \sum_{i=1}^n x_i \in \{0, 1, \dots, n\}$  und den Spezialfall  $m = 1$  die Intervallgrenzen. Es ist deutlich zu erkennen, dass für kleinere Stichproben insbesondere für sehr kleine oder große Anteilswerte<sup>21</sup> das in rot abgetragene Standard-Intervall deutlich von dem in grün dargestellten Wilson-Intervall bzw. dem in blau dargestellten exakten Intervall abweicht. Die rechte Grafik zeigt, dass sich für größere Stichproben alle drei Intervalle deutlich annähern.



#### Aufgabe 4.5

Zur Schätzung des Ausschussanteiles eines umfangreichen Lieferpostens werde diesem eine Stichprobe von 200 Teilen entnommen. Dabei seien 190 einwandfreie Teile festgestellt worden.

- Geben Sie eine Schätzung für den Ausschussanteil an.
- Wie groß kann maximal die Varianz der Schätzung sein?
- Berechnen Sie ein Konfidenzintervall für den Ausschussanteil zum Konfidenzniveau 0.95 und interpretieren Sie das Ergebnis.

Lösung auf Seite ??

#### Aufgabe 4.6

Eine Partei P möchte ihren Wähleranteil durch eine Umfrage feststellen. Von 1500 befragten Wahlberechtigten gaben 390 an, diese Partei zu wählen. Geben Sie die Konfidenzintervalle zum Konfidenzniveau

- 90%
- 95%
- 99%

an. Vergleichen Sie die Ergebnisse und interpretieren Sie diese.

Lösung auf Seite ??

<sup>21</sup>Für sehr kleine  $p \approx 0$  erwarten wir nur kleine Stichprobensummen  $y$  und für sehr große  $p \approx 1$  entsprechend sehr große  $y$ , so dass dann auch  $\hat{p} \approx 0$  oder  $\hat{p} \approx 1$  beobachtet werden.

### Aufgabe 4.7

Bei der Überprüfung eines Abfüllautomaten für 0.5-Liter-Flaschen ergaben sich folgende Werte für die tatsächlich abgefüllte Menge (Angaben in  $\text{cm}^3$ ):

499.5, 501.8, 501.0, 502.5, 501.0

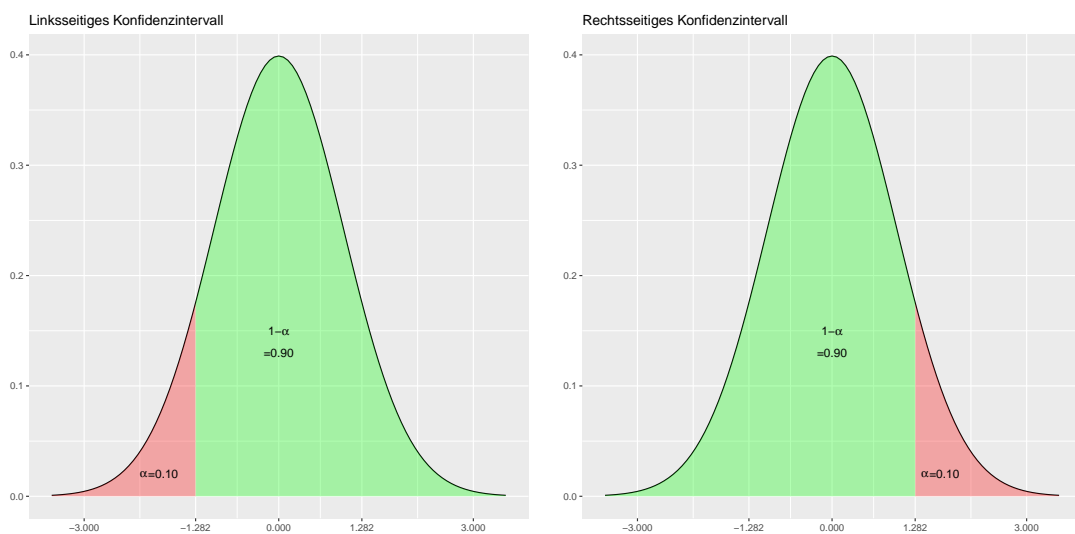
504.1, 497.9, 503.3, 498.7, 500.2

- Geben Sie erwartungstreue Schätzfunktionen und entsprechende Schätzwerte für den Mittelwert und die Varianz der Abfüllmenge an.
- Bestimmen Sie ein 90%iges Konfidenzintervall für den Erwartungswert.
- Welche Voraussetzung bzgl. der Verteilung müssen gegeben sein?

Lösung auf Seite ??

### 4.2.3 Einseitige Konfidenzintervalle

Hin und wieder ist nur eine Grenze des Konfidenzintervalles von Interesse. Dann konstruiert man einseitige Konfidenzintervalle, indem die uninteressante Grenze auf einen natürlichen Wert gelegt wird und statt dem Quantil der Ordnung  $\frac{\alpha}{2}$  verwendet man das Quantil der Ordnung  $\alpha$ , wie die folgende Grafik für den Parameter  $\mu$  der Normalverteilung veranschaulicht.



### 4.3 Hypothesentests

Beim parametrischen statistischen Testen werden Hypothesen über den unbekannt Parameter einer Verteilung untersucht. Als einführendes Beispiel betrachten wir den Ausgang eines Würfelspiels bei dem man 2€ gewinnt, wenn keine Sechs fällt und der Gegenspieler 10€ erhält, falls eine Sechs fällt. Nach 20 Spielen wurden folgende Ergebnisse

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

beim Würfel beobachtet. Demnach haben wir insgesamt 140€ verloren. Die zentrale Frage lautet, ob unser Spielpartner einen fairen Würfel verwendet hat oder wir einfach nur viel Pech hatten. Um diese Frage näher zu untersuchen betrachten wir die Bernoullische Zufallsgröße  $X \sim \text{Bin}(1, p)$  mit

$$X_i = \begin{cases} 1, & \text{falls im } i\text{-ten Wurf eine Sechs gewürfelt wird} \\ 0, & \text{falls im } i\text{-ten Wurf keine Sechs gewürfelt wird,} \end{cases}$$

wobei wir aktuell keine Informationen über den Parameter  $p$  haben. Allerdings haben wir die konkrete Stichprobe vom Umfang  $n = 20$  mit  $t = \sum_{i=1}^{20} x_i = 15$  vorliegen, anhand der wir Rückschlüsse auf  $p$  ziehen möchten.

Da uns interessiert, ob tatsächlich mit einem fairen Würfel gewürfelt wurde, untersuchen wir die Hypothese, dass der unbekannt Parameter  $p$  den Werte  $p_0 = \frac{1}{6}$  annimmt und betrachten hierfür die Summe  $T = \sum_{i=1}^n X_i \sim \text{Bin}(n, p)$ , welche die Anzahl der gewürfelten Sechsen zählt. Wenn die

Hypothese  $H_0 : p = p_0 = \frac{1}{6}$  tatsächlich gilt, dann können wir einfach die Wahrscheinlichkeit berechnen

$$P(T = 15) = \binom{20}{15} \left(\frac{1}{6}\right)^{15} \left(\frac{5}{6}\right)^5 = 1.3252 \times 10^{-8},$$

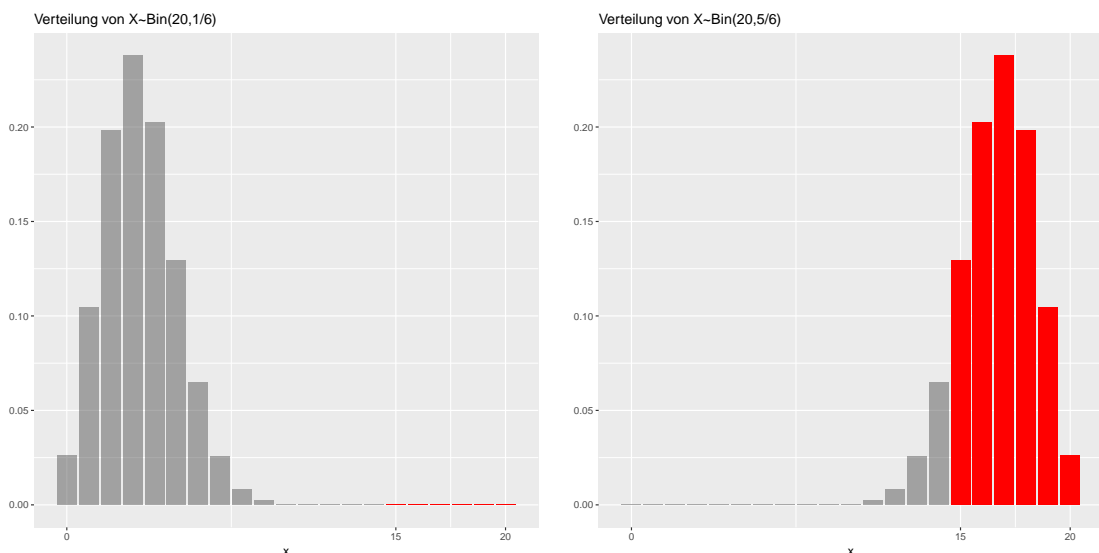
das beim Würfeln mit einem fairen Würfel exakt 15 mal die Sechs fällt. Diese Wahrscheinlichkeit ist sehr klein, allerdings sprechen nicht nur genau 15 gewürfelte Sechsen gegen die Hypothese  $H_0$ , sondern es würden natürlich erst recht 16 oder mehr gewürfelte Sechsen dagegen sprechen. Diese Wahrscheinlichkeit, mit einem fairen Würfel solch ein Ergebnis oder sogar noch ein schlechteres zu erhalten

$$P(T \geq 15) = \sum_{i=15}^{20} \binom{20}{i} \left(\frac{1}{6}\right)^i \left(\frac{5}{6}\right)^{20-i} = 1.412 \times 10^{-8}, \quad (30)$$

ist sehr klein. Tatsächlich gibt es jetzt zwei Möglichkeiten. Entweder wurde mit einem fairen Würfel gewürfelt und wir haben ein sehr unwahrscheinliches Ereignis beobachtet und demnach beim Glücksspiel einfach Pech gehabt. Andererseits könnte unser Gegenspieler natürlich auch einen gezinkten Würfel benutzt haben. Um dies zu veranschaulichen berechnen wir für den Parameter  $p = p_1 = \frac{5}{6}$  obige Wahrscheinlichkeit des Stichprobenergebnisses

$$P(T \geq 15) = \sum_{i=15}^{20} \binom{20}{i} \left(\frac{5}{6}\right)^i \left(\frac{1}{6}\right)^{20-i} = 0.8982.$$

Wäre der Würfel also so getrimmt, dass in fünf von sechs Fällen eine Sechs fällt, wäre obiges Ergebnis überhaupt nicht mehr unwahrscheinlich oder unglaublich. Die folgenden Grafiken veranschaulichen die beiden unterschiedlichen Verteilungen, die sich durch die Wahl von  $p_0$  beziehungsweise  $p_1$  ergeben. In rot sind die Wahrscheinlichkeiten, dass 15 oder mehr Sechsen beobachtet werden, dargestellt.



Prinzipiell kann es also zwei Gründe für die sehr kleine Wahrscheinlichkeit (30) geben. Entweder stimmt die Hypothese  $H_0 : p = p_0 = \frac{1}{6}$  nicht, oder es ist ein sehr unwahrscheinliches Ereignis eingetreten<sup>22</sup>. Dieses einführende Beispiel führt auf einen Binomialtest, den wir in Abschnitt 4.3.3 nochmals genauer besprechen werden.

Beim statistischen Testen gibt man sich allgemein vorab eine Irrtumswahrscheinlichkeit  $\alpha$  vor und vergleicht diese mit der Wahrscheinlichkeit des Stichprobenergebnisses, welches unter der Hypothese  $H_0$  berechnet wurde. Ist diese kleiner als die vorgegebene Irrtumswahrscheinlichkeit, so vermutet man den zweiten Grund und glaubt nicht mehr daran rein zufällig ein sehr seltenes Ereignis beobachtet zu haben. Daher verwirft man die Nullhypothese und spricht in diesem Fall von einem signifikanten Ergebnis, da die Wahrscheinlichkeit einer Fehlentscheidung sich mit  $\alpha$  abschätzen lässt. Anderenfalls hat der Test keine Einwände gegen die Nullhypothese, was nicht als Nachweis dieser zu verstehen ist. Bei obigen Beispiel haben wir insgesamt 15 gewürfelte Sechsen beobachtet und die Wahrscheinlichkeit (30) berechnet, mit einem fairen Würfel 15 oder noch mehr Sechsen bei 20 Würfeln zu erhalten. Werten wir also unser Beispiel mit einer Irrtumswahrscheinlichkeit  $\alpha = 0.05$  aus, kommen wir zu dem signifikanten Schluss, dass kein fairer Würfel verwendet wurde.

<sup>22</sup>Jeder der Lotto spielt, hofft auf das Eintreten eines solch unwahrscheinlichen Ereignisses.

Um allgemein die Wahrscheinlichkeiten des Stichprobenergebnisses berechnen zu können, benötigt man die Verteilung der Stichprobenfunktion. Wie auch bei den Konfidenzintervallen, ist diese bekannt, wenn Aussagen über die Parameter der Normal- oder Binomialverteilung betrachtet werden.

Bevor wir spezielle Tests näher betrachten, wollen wir das allgemeine Vorgehen beim statistischen Testen beschreiben. Ausgangspunkt stellt die sogenannte Nullhypothese  $H_0 : F = F_0$  dar, dass die unbekannte Verteilung  $F$  der hypothetischen Verteilung  $F_0$  entspricht. In der parametrischen Statistik spezifizieren nur noch die unbekannt Parameter die hypothetische Verteilung  $F_0$ . Ist beispielsweise das Merkmal  $X \sim \mathcal{N}(\mu, \sigma^2)$  normalverteilt bei bekannter Varianz  $\sigma^2$ , so wird  $F_0$  nur noch durch den hypothetischen Parameter  $\mu_0$  beschrieben. Bei parametrischen Tests formuliert man die Nullhypothese daher als Aussage über den unbekannt Parameter, also beispielsweise  $H_0 : \mu = \mu_0$ . Aus der mathematischen Stichprobe  $(X_1, \dots, X_n)$  berechnet man dann eine Testgröße  $T = T(X_1, \dots, X_n)$ , welche die Unterschiede der hypothetischen Verteilung  $F_0$  zur tatsächlichen Verteilung  $F$  widerspiegelt. Mit Hilfe der Verteilung der Testgröße, die sich unter der hypothetischen Verteilung  $F_0$  ergibt, kann die Wahrscheinlichkeit berechnet werden, dass sich unter Gültigkeit der Nullhypothese tatsächlich ein schlechteres Stichprobenergebnis ergibt, als in der konkreten Stichprobe beobachtet  $t = T(x_1, \dots, x_n)$  wurde. Die Berechnung dieses sogenannten p-Wertes hängt von der konkreten Fragestellung ab. Sprechen nur zu große Testwerte gegen die Nullhypothese berechnet man  $P(T \geq t)$  und spricht in diesem Fall von einem rechtsseitigen Test, so wie wir das bei der einführenden Frage, ob ein fairer Würfel beim Spiel verwendet wurde, gesehen haben. Sprechen hingegen zu kleine Werte der Testgröße  $T$  gegen die Nullhypothese, entspricht der p-Wert  $P(T \leq t)$  und man spricht von einem linksseitigen Test. Von einem zweiseitigen Test spricht man hingegen, wenn sowohl zu große als auch zu kleine Testwerte gegen die Nullhypothese sprechen. Dann berechnet man den p-Wert  $P(|T| \geq |t|) = P(T \leq -|t|) + P(T \geq |t|)$ . Entsprechend der konkreten Fragestellung, die immer vor der Datenanalyse und Durchführung der statistischen Tests formuliert werden sollten, wird bei parametrischen Tests die Alternativhypothese  $H_1 : F \neq F_0$  genauer formuliert. Bei obigen Test zum Überprüfen, ob ein gezinkter Würfel verwendet wurde, sprechen nur zu viele geworfene Sechsen für die Alternativhypothese. Eine hohe beobachtete Anzahl gewürfelter Sechsen kann einerseits rein zufällig, durch Auftreten eines seltenen Ereignisses auftreten oder dadurch entstehen, dass in Wirklichkeit ein Würfel mit  $p = p_1 > \frac{1}{6}$  verwendet wurde, so dass die Hypothesen bei diesem Beispiel die Hypothesen<sup>23</sup>  $H_0 : p = 1/6$  gegen  $H_1 : p > 1/6$  mit dem p-Wert  $P(T \geq t)$  untersucht wurden. Ist dieser p-Wert, kleiner als die vorgegebene Irrtumswahrscheinlichkeit  $\alpha$ , so wird die Nullhypothese  $H_0$  verworfen und es gilt der signifikante Nachweis der Alternativhypothese  $H_1 : F \neq F_0$ .

Zur Berechnung des konkreten p-Wertes verwendet man spezielle Statistiksoftware, die die üblichen Testverteilungen implementiert hat. Die in Schulen typischerweise verwendeten Taschenrechner mit Computeralgebrasystemen stellen ebenso wie die Software Geogebra die üblichen Testverteilungen zur Verfügung. Möchte man statistische Tests ohne Statistiksoftware nur mit Hilfe tabellierter Quantile der Testverteilungen berechnen, konstruiert man alternativ für jeden Test den sogenannte kritischen Bereiche  $K^*$ . Dieser kritische Bereich wird je nach Hypothesen links-, rechts- oder zweiseitig mit Hilfe der Quantile der Testverteilung derart konstruiert, dass zu gegebenen  $\alpha \in (0, 1)$

$$P(T \in K^*) = \alpha$$

gilt. Für die konkrete Stichprobe  $(x_1, \dots, x_n)$  wird dann wiederum der konkrete Testwert  $t = T(x_1, \dots, x_n)$  berechnet und geprüft, ob dieser im kritischen Bereich  $K^*$  liegt. Falls  $t \in K^*$  ist dies äquivalent dazu, dass der entsprechende p-Wert des Stichprobenwertes kleiner als die Irrtumswahrscheinlichkeit  $\alpha$  ist. Dies führt zur Ablehnung der Nullhypothese  $H_0$  und damit verbunden dem signifikanten Nachweis der Alternativhypothese  $H_1$ . Liegt hingegen  $t \notin K^*$ , so hat der durchgeführte Test keine Einwände gegen die Nullhypothese. Die Äquivalenz dieser beiden Vorgehen verdeutlichen wir nochmals genauer in der Grafik auf Seite 40 für den t-Test.

Prinzipiell untersuchen wir Thesen über die uns unbekannt Realität, die entweder wahr oder falsch sind. Da wir die Realität jedoch nicht kennen nutzen wir den statistischen Test zur objektiven Einschätzung. Je nach p-Wert und vorgegebener Irrtumswahrscheinlichkeit, entscheidet sich der Test entweder für die Alternativhypothese  $H_1$  oder hat keine Einwände gegen die Nullhypothese  $H_0$ . Falls in Wirklichkeit tatsächlich die Nullhypothese gilt, und der Test sich auch für diese entscheidet, hat der Test natürlich die richtige Einschätzung getroffen. Es ist allerdings auch möglich, dass der Test die Nullhypothese verwirft und statt dessen die Alternativhypothese  $H_1$  annimmt, obwohl in Wirklichkeit die Nullhypothese richtig ist. Diesen Fehler nennt man auch  **$\alpha$ -Fehler** oder **Fehler 1. Art**, da der Test ja so konstruiert ist, dass die Wahrscheinlichkeit dieser Fehlentscheidung

<sup>23</sup>Häufig findet man für einseitige Tests auch die Schreibweise  $H_0 : p \leq 1/6$  gegen  $H_1 : p > 1/6$  um die Komplementarität beider Hypothesen zu verdeutlichen.

genau der vorgegebenen Irrtumswahrscheinlichkeit  $\alpha$  entspricht. Ebenso kann in Wirklichkeit natürlich die Alternativhypothese  $H_1$  gelten. Wenn sich der Test dann für die Alternativhypothese entscheidet, hat der Test wiederum die Realität richtig eingestuft. Gilt in Wahrheit allerdings die Alternativhypothese und der Test entscheidet sich fälschlicherweise für die Nullhypothese, spricht man vom **Fehler 2. Art** oder  $\beta$ -Fehler. Die Bestimmung des  $\beta$ -Fehlers ist sehr viel komplizierter und hängt mit der Power bzw. der Teststärke eines Tests zusammen. In Abschnitt 4.3.1 auf Seite 42 werden wir ein Beispiel für den rechtsseitigen Gauß-Test illustrieren. Nach Durchführung eines statistischen Tests, befindet man sich entweder in der ersten oder zweiten Spalte der Tabelle 1. Da uns die Realität unbekannt ist, wissen wir allerdings nicht, ob wir uns in der ersten oder zweiten Zeile befinden. Aus diesem Grund spricht man auch nur von einem statistisch signifikanten Ergebnis, wenn der Test die Nullhypothese verwirft und sich statt dessen für die Alternativhypothese entscheidet. Dieser Fall entspricht der zweiten Spalte und falls tatsächlich die Nullhypothese gilt, kennt man den  $\alpha$ -Fehler und spricht deshalb auch von einem signifikanten Nachweis der Alternativhypothese. Das Wort „signifikant“ beinhaltet also immer die Möglichkeit des  $\alpha$ -Fehlers, der aber durch die Wahl der Irrtumswahrscheinlichkeit  $\alpha$  gesteuert wird. Die Nullhypothese kann man hingegen aufgrund der Konstruktion des Tests niemals statistisch nachweisen.

	Test entscheidet sich für $H_0$	Test entscheidet sich für $H_1$
$H_0$ ist richtig	✓	Fehler 1. Art – $\alpha$ -Fehler
$H_1$ ist richtig	Fehler 2. Art – $\beta$ -Fehler	✓

Tabelle 1: Entscheidungstabelle beim statistischen Testen

Das statistische Testen kann man als Analogie zu einem Gerichtsverfahren sehen, bei dem sich die in Tabelle 2 dargestellten folgenden vier Fälle ergeben.

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

Tabelle 2: Entscheidungstabelle beim

Ebenso wie beim Gerichtsverfahren ohne Zeugen ein Angeklagter niemals seine Unschuld beweisen kann, ist es beim statistischen Testen nicht möglich, die Gültigkeit der Nullhypothese nachzuweisen.

#### Aufgabe 4.8

In einem Schultyp müssen durchschnittlich 8% der Schüler das Schuljahr wiederholen. Die Schulbehörde trifft einige Maßnahmen zur individuellen Förderung leistungsschwacher Schüler und möchte den Erfolg dieser Maßnahme durch einen statistischen Test überprüfen. Dazu ist geplant 150 Schüler repräsentativ auszuwählen und zu beobachten. Ermitteln Sie den Kritischen Bereich  $K^*$  dieses Tests und erklären Sie seine Bedeutung in diesem Zusammenhang.

*Lösung auf Seite ??*

#### Aufgabe 4.9

Es ist bekannt, dass 35% aller Personen, die ein Fahrzeug lenken, ohne Freisprecheinrichtung telefonieren. Nachdem die Polizei durch Aufklärungskampagnen und verstärkte Kontrollen versucht hat diese Situation zu verbessern, werden 780 Personen kontrolliert. 240 von ihnen telefonierten nach wie vor ohne Freisprecheinrichtung mit dem Handy beim Fahren. Die Verantwortlichen behaupten nun, dass diese Maßnahmen wirksam waren.

- Erklären Sie, warum und auf welche Art dieser Test versagen kann.
- Geben Sie die Nullhypothese, die Gegenhypothese, die Testgröße und deren Verteilung sowie den kritischen Bereich an und entscheiden Sie, ob es sich um einen links- oder rechtsseitigen Test handelt.

*Lösung auf Seite ??*

#### Aufgabe 4.10

Eine Maschine produziert Werkstücke mit einem Ausschussanteil von 8%. Ein statistischer Test soll überprüfen, ob dieser Anteil größer geworden ist. Man untersucht dazu 50 Werkstücke, von denen maximal 6 schadhaft sein dürfen.

- Beschreiben Sie diesen Test und formulieren Sie die Hypothesen.
- Wie sieht der kritische Bereich aus und was bedeutet er?

Lösung auf Seite ??

#### Aufgabe 4.11

Eine Maschine stanzt Werkstücke in großen Serien. Der Hersteller der Maschine behauptet, dass der Ausschussanteil 5% beträgt. Der potentielle Käufer einer Maschine glaubt ihm nur, wenn bei einer Probeserie von 180 Stück nicht mehr als 10 schadhafte Werkstücke auftreten. Um welchen Test handelt es sich? Berechnen Sie die Wahrscheinlichkeit, fälschlicherweise die Maschine nicht zu kaufen.

Lösung auf Seite ??

#### Aufgabe 4.12

Ein Nachhilfeeinstitut wirbt mit der Aussage, dass 95% der von ihnen betreuten Schüler die Abschlussprüfung schaffen. Mit welchem Parametertest kann man diese Behauptung untersuchen? Formulieren Sie die passenden Hypothesen. Was bedeutet es aus Sicht der Kunden, wenn die Fehler 1. Art oder 2. Art möglichst klein sind?

Lösung auf Seite ??

#### Aufgabe 4.13

Frau Müller glaubt, dass in ihrer Klasse mit 25 Schülern die Mathematikhausaufgaben zu 80% ordnungsgemäß durchgeführt werden. Ein Kollege zweifelt und glaubt ihr nur, wenn bei der nächsten Kontrolle der gesamten Klasse mindestens 17 Hefte in Ordnung sind. Wie groß ist die Wahrscheinlichkeit, dass Frau Müller widerlegt wird, obwohl ihre Annahme richtig ist?

Lösung auf Seite ??

### 4.3.1 Gauß-Test zur Überprüfung des Erwartungswertes der Normalverteilung bei bekannter Varianz

Bei einem einfachen zweiseitigen Test über den unbekannt Parameter  $\mu = \mu_0$  des normalverteilten Merkmals  $X \sim \mathcal{N}(\mu, \sigma^2)$  bei bekannter Varianz  $\sigma^2$  werden die Hypothesen  $H_0 : \mu = \mu_0$  gegen  $H_1 : \mu \neq \mu_0$  überprüft. Als Stichprobengröße betrachtet man das standardisierte arithmetische Mittel. Unter der Nullhypothese ist das arithmetische Mittel  $\bar{X} \sim \mathcal{N}(\mu_0, \frac{\sigma^2}{n})$  und daher

$$T = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim \mathcal{N}(0, 1)$$

standardnormalverteilt. Aus diesem Grund wird dieser Test auch als **Gauß-Test** bezeichnet<sup>24</sup>. Wenn die Nullhypothese  $H_0$  tatsächlich gilt, streuen die Werte des Stichprobenmittels  $\bar{X}$  um  $\mu_0$  und die Werte der Testgröße  $T$  um die Null. Ist hingegen die Alternativhypothese wahr, wird das Stichprobenmittel stark von  $\mu_0$  abweichen. Daher sprechen sowohl zu große, als auch zu kleine Werte der Testgröße gegen die Nullhypothese.

Somit kann einerseits für den konkreten Stichprobentestwert  $t = \frac{\bar{x} - \mu_0}{s} \sqrt{n}$  der konkreten Stichprobe  $x_1, \dots, x_n$  der p-Wert

$$p = P(|T| > |t|) = P(T > |t|) + P(T < -|t|)$$

berechnet werden. Alternativ konstruiert man den zweiseitigen kritischen Bereich

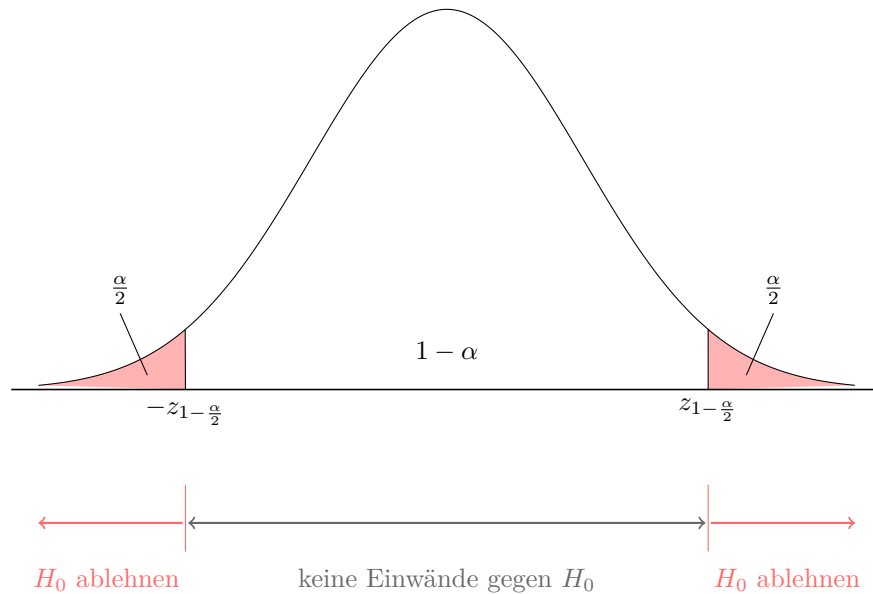
$$K^* = (-\infty, -z_{1-\frac{\alpha}{2}}) \cap (z_{1-\frac{\alpha}{2}}, \infty),$$

<sup>24</sup>Häufig wird das standardisierte Mittel mit  $Z$  bezeichnet, so dass wegen  $Z = T = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim \mathcal{N}(0, 1)$  dieser Test auch **Z-Test** genannt wird.

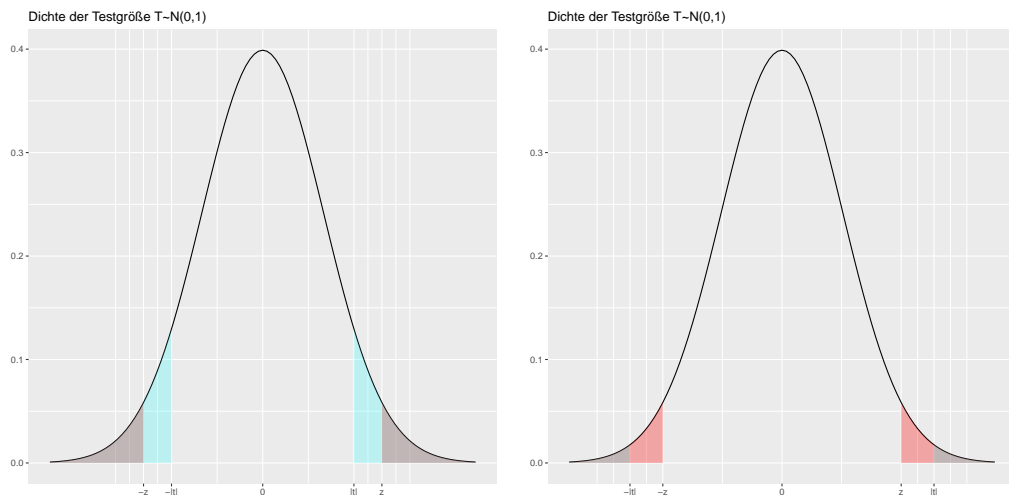
ähnlich wie bei der Konstruktion von Konfidenzintervallen, mit Hilfe der Sigma-Intervalle, wegen

$$1 - \alpha = P(T \notin K^*) = P(-z_{1-\frac{\alpha}{2}} \leq T \leq z_{1-\frac{\alpha}{2}}).$$

Der kritische Bereich entspricht somit dem Ablehnungsbereich der Nullhypothese und dessen Komplement dem Annahmehereich, wie in der folgende Abbildung dargestellt.



Ist der p-Wert kleiner als die vorgegebene Irrtumswahrscheinlichkeit  $\alpha$  wird die Nullhypothese abgelehnt. Dieser Fall ist äquivalent dazu, dass  $t \in K^*$ , d.h. der konkrete Testwert sich im kritischen Bereich befindet. Die beiden folgenden Grafiken illustrieren die Äquivalenz zwischen der Auswertung des p-Wertes und dem Vergleich, ob sich die konkrete Testgröße im kritischen Bereich befindet, für einen zweiseitigen Test. Dabei sind die Irrtumswahrscheinlichkeit  $\alpha = 0.05$  und der damit verbundene kritische Bereich  $K^* = (-\infty, -z) \cup (z, \infty)$  rot und der p-Wert  $P(|T| \geq |t|) = P(T \geq |t|) + P(T \leq -|t|)$  in blau dargestellt. Die linke Grafik entspricht dem Fall  $t \notin K^*$ . Es ist leicht erkennbar, dass der p-Wert größer als die Irrtumswahrscheinlichkeit ist und somit der Test keine Einwände gegen die Nullhypothese äußert. Der Fall  $t \in K^*$  ist in der rechten Grafik dargestellt, der äquivalent dazu ist, dass der p-Wert kleiner als die Irrtumswahrscheinlichkeit ist.



Bei einseitigen Tests ist man nur an einseitigen Fragestellungen interessiert. Dies wirkt sich auf die Konstruktion des kritischen Bereichs bzw. die Berechnung des p-Wertes aus. Wird beispielsweise die Nullhypothese  $H_0 : \mu = \mu_0$  gegen die Alternativhypothese  $H_1 : \mu > \mu_0$  untersucht, sprechen wegen  $\bar{X}_n - \mu_0 > 0$  nur zu große Werte der Testgröße gegen die Nullhypothese<sup>25</sup> und der kritische

<sup>25</sup>Die Nullhypothese des rechtsseitigen Tests wird auch häufig  $H_0 : \mu \leq \mu_0$  anstatt  $H_0 : \mu = \mu_0$  formuliert. Hiermit wird die Komplementarität der Alternativhypothese zur Nullhypothese besser verdeutlicht. Allerdings wird der p-Wert immer mit der hypothetischen Verteilung  $X \sim \mathcal{N}(\mu_0, \sigma^2)$  berechnet und die Darstellung  $H_0 : \mu \leq \mu_0$  kann man als eine Reihe von Tests mit Nullhypothese  $H_0 : \mu = \tilde{\mu}_0$  mit  $\tilde{\mu}_0 \leq \mu_0$  und damit verbundener Hypothetischer Verteilungen  $X \sim \mathcal{N}(\tilde{\mu}_0, \sigma^2)$  auffassen. Aus Monotoniegründen dominiert allerdings der Test mit  $H_0 : \mu = \mu_0$  diese Reihe.



Bereich  $K^* = \{t \in \mathbb{R} : t > z_{1-\alpha}\} = (z_{1-\alpha}, \infty)$  konzentriert sich wegen

$$P(T > z_{1-\alpha}) = 1 - \Phi(z_{1-\alpha}) = 1 - (1 - \alpha) = \alpha$$

auf die rechte Seite, so dass man auch von rechtsseitigen Tests spricht. Demnach wird bei einem linksseitigen Test die Nullhypothese<sup>26</sup>  $H_0 : \mu = \mu_0$  gegen die Alternativhypothese  $H_1 : \mu < \mu_0$  untersucht. Folglich sprechen wegen  $\bar{X}_n - \mu_0 < 0$  nur zu kleine Werte der Testgröße gegen die Nullhypothese  $H_0$  und wegen

$$P(T < z_\alpha) = \Phi(z_\alpha) = \alpha$$

konzentriert sich der kritische Bereich

$$K^* = \{t \in \mathbb{R} : t < z_\alpha\} = \{t \in \mathbb{R} : t < -z_{1-\alpha}\} = (-\infty, -z_{1-\alpha})$$

auf die linke Seite der Verteilung.

#### Beispiel 4.5: Untersuchung des Abfüllgewichts

Um zu kontrollieren, ob das Abfüllgewicht einer 100g Tafel Schokolade tatsächlich zu gering ist, wurde die Stichprobe

$$(100, 97, 101, 96, 98, 102, 96, 100, 101, 98)$$

erhoben. Aus dieser ergeben sich das Stichprobenmittel  $\bar{x} = 98.9$  und Standardabweichung  $s = 2.183$ . Unter der Annahme  $X \sim \mathcal{N}(\mu, 2^2)$  untersuchen wir bei 5%-iger Irrtumswahrscheinlichkeit die Nullhypothese  $H_0 : \mu \geq 100 = \mu_0$  gegen die Alternativhypothese  $H_1 : \mu < 100 = \mu_0$ , da wir letztere statistisch nachweisen wollen. Für die konkrete Testgröße ergibt sich  $t = \sqrt{n} \frac{\bar{x} - \mu_0}{s} = \sqrt{10} \frac{98.9 - 100}{2} = -1.739$ . Für  $\alpha = 0.05$  berechnen wir das Quantil der Standardnormalverteilung  $z_\alpha = -z_{1-\alpha} = -z_{0.95} = -1.645$  und erhalten den kritischen Bereich  $K^* = \{t \in \mathbb{R} : t < -1.645\} = (-\infty, -1.645)$ . Da der Testwert  $t = -1.739 \in K^*$  im kritischen Bereich liegt, lehnen wir die Nullhypothese  $H_0$  ab und haben somit den signifikanten Nachweis der Alternativhypothese  $H_1$ . Diese besagt, dass das Abfüllgewicht der Schokoladentafeln bei einer Irrtumswahrscheinlichkeit von 5% signifikant kleiner als 100g ist. Die Berechnung des p-Wertes

$$p = P(T < t) = P(T < -1.739) = \Phi(-1.739) = 0.041 < 0.05 = \alpha$$

hätte ebenso zur Verwerfung der Nullhypothese geführt.

Hätten wir bei obiger Fragestellung statt dessen den rechtsseitigen Test mit den Hypothesen  $H_0 : \mu \leq \mu_0 = 100$  gegen  $H_1 : \mu > \mu_0 = 100$  betrachtet, würde bei selber Irrtumswahrscheinlichkeit  $\alpha = 0.05$  wegen  $t = -1.739 \notin K^* = (1.645, \infty)$  gelten. Somit hätte der Test keine Einwände gegen die Nullhypothese, dass das Abfüllgewicht weniger als 100g beträgt. Dies ist allerdings keine signifikante Aussage und daher weniger wertvoll. Aus diesem Grunde sollte man die bei einseitigen Parametertests immer die Aussage, die man gerne statistisch nachweisen möchte, die Alternativhypothese entsprechend dieser Aussage wählen.

Abbildung 9 zeigt die Durchführung des Gauß-Tests mittels Geogebra.

Bevor im nächsten Abschnitt das Vorgehen beim Testen von Hypothesen über den Erwartungswert  $\mu = \mathbb{E}X$  des normalverteilten Merkmals  $X \sim \mathcal{N}(\mu, \sigma^2)$  im Falle der unbekanntem Varianz  $\text{Var}(X) = \sigma^2$  besprechen, wollen wir nochmals kurz die Wahrscheinlichkeiten der Fehlentscheidungen beim parametrischen Testen, wie wir sie auf Seite 37 besprochen haben, grafisch veranschaulichen. Hierfür betrachten wir den rechtsseitigen Test mit den Hypothesen  $H_0 : \mu = \mu_0$  gegen  $H_1 : \mu > \mu_0$ . Unter der Nullhypothese ist das Stichprobenmittel  $\bar{X} \sim \mathcal{N}(\mu_0, \frac{\sigma^2}{n})$ . Um den Fehler 2. Art berechnen zu können, benötigt man allerdings einen konkreten Parameter  $\mu_1$ , so dass unter der Alternativhypothese  $\bar{X} \sim \mathcal{N}(\mu_1, \frac{\sigma^2}{n})$  gilt. Dann können wir auch die Wahrscheinlichkeit  $P(T \in K^* | H_1)$  berechnen, dass die Nullhypothese nicht verworfen wird, obwohl in Wirklichkeit die Alternativhypothese  $H_1$  gilt. In der folgenden Grafik ist der  $\alpha$ -Fehler rot und der  $\beta$ -Fehler unter dem hypothetischen Parameter  $\mu_1$  der Alternativhypothese blau dargestellt.

<sup>26</sup>Anstatt  $H_0 : \mu = \mu_0$  wird die Nullhypothese des linksseitigen Tests, analog zum rechtsseitigen Test, häufig mittels der Ungleichung  $H_0 : \mu \geq \mu_0$  formuliert.

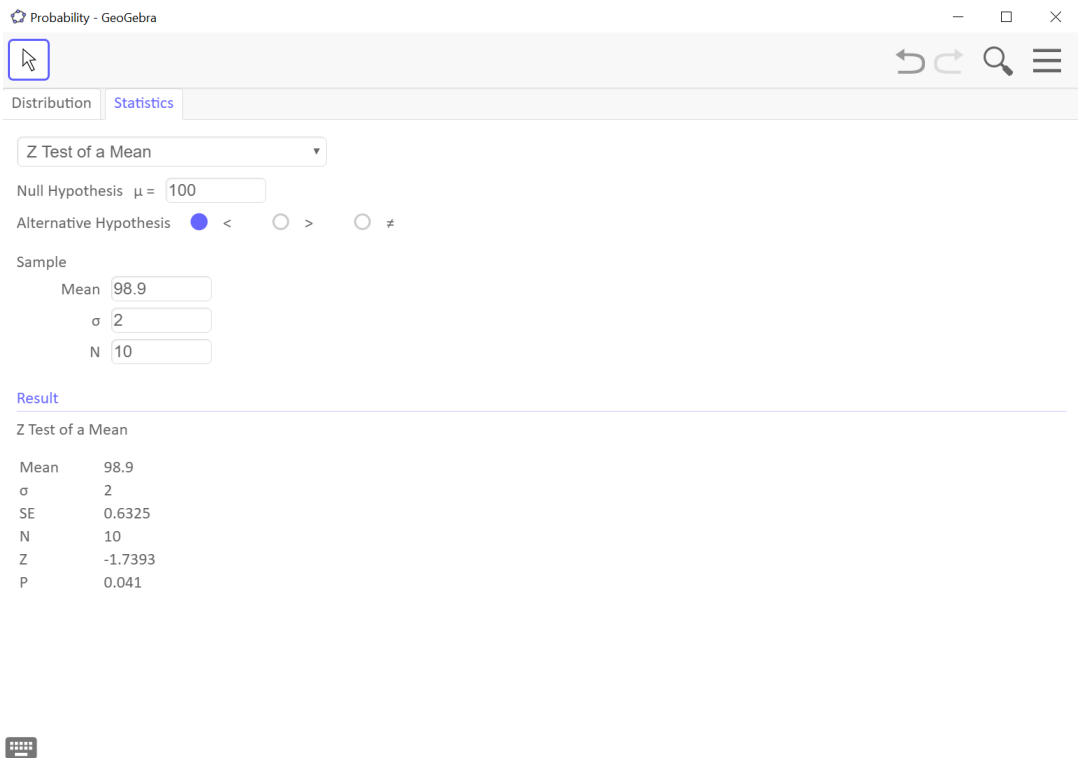
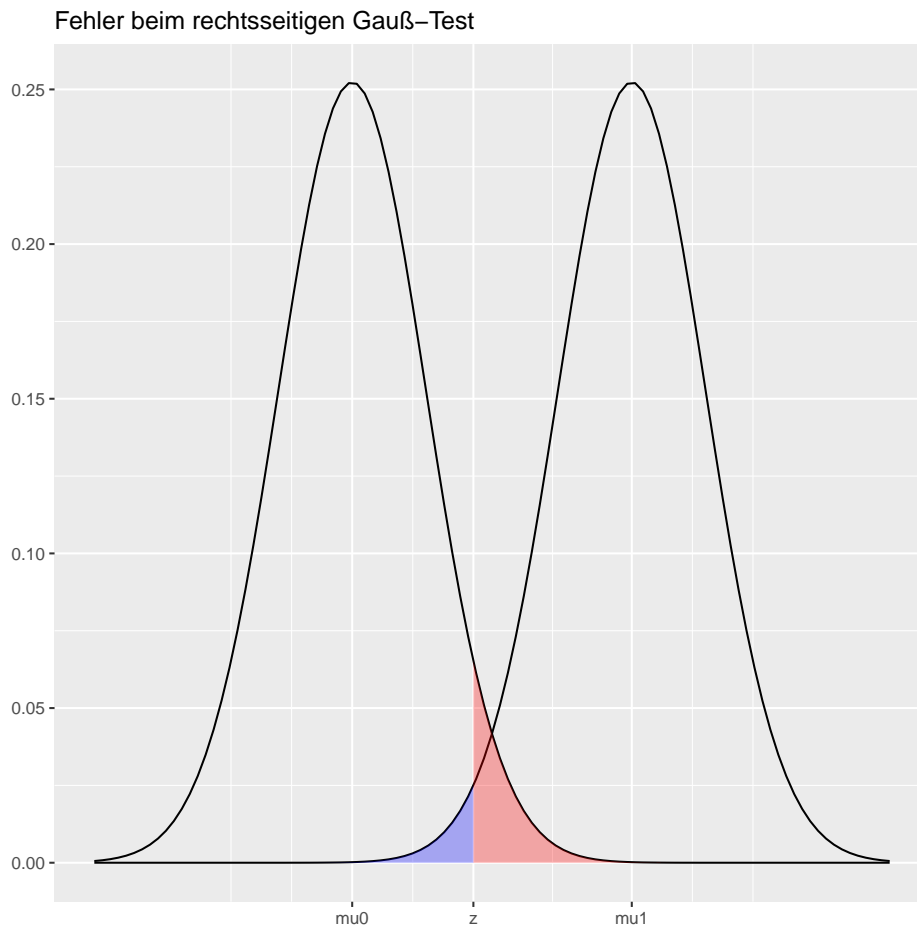
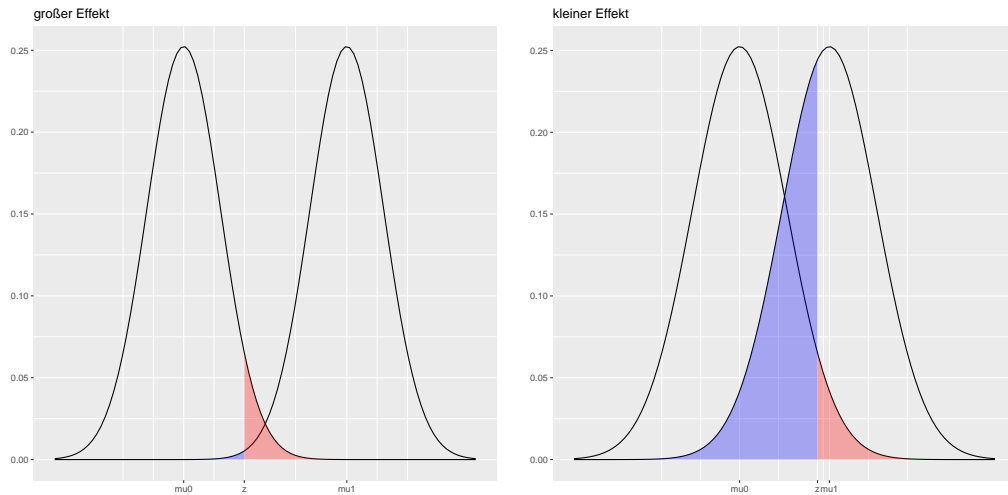


Abbildung 9: Durchführung eines Gauß-Tests mit Geogebra

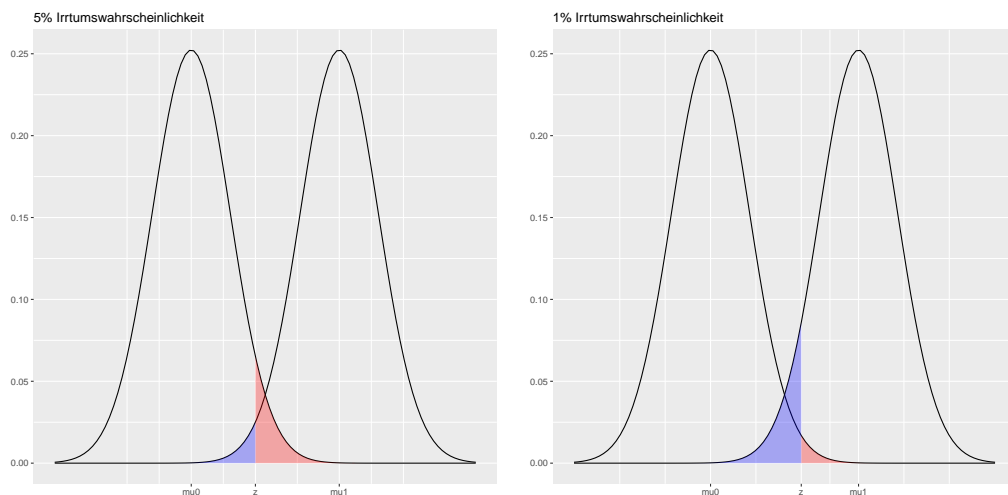


Die Differenz zwischen den beiden hypothetischen Mittelwerten  $\mu_0$  und  $\mu_1$  nennt man auch Effekt. Ist dieser Effekt also sehr klein, liegen  $\mu_0$  und  $\mu_1$  sehr nah beieinander und der Test hat Schwierigkeiten klar zwischen  $H_0$  und  $H_1$  zu trennen. Entsprechend ist der  $\beta$ -Fehler in diesem Fall sehr groß, wie die folgende rechte Grafik zeigt. Für sehr große Effekte liegen hingegen  $\mu_0$  und  $\mu_1$  sehr weit auseinander und der Test kann klar zwischen  $H_0$  und  $H_1$  trennen. Die linke Grafik verdeutlicht,

dass in diesem Fall der blau eingefärbte  $\beta$ -Fehler sehr klein wird.



Mit der letzten Grafik wird auch verständlich, dass sich die Wahl von der Irrtumswahrscheinlichkeit  $\alpha$  automatisch auf den  $\beta$ -Fehler niederschlägt. Verringert man  $\alpha$ , um eine fälschliche Ablehnung der Nullhypothese zu vermeiden, wächst dadurch der Fehler 2. Art, wie in der letzten Grafik veranschaulicht.



#### 4.3.2 t-Tests zur Überprüfung des Erwartungswertes der Normalverteilung bei unbekannter Varianz

Ist der Parameter  $\sigma$  unbekannt, wird die unbekannte Varianz  $\sigma^2$  durch die empirische Stichprobenvarianz  $\hat{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  ersetzt. Damit erhält man die Testgröße

$$T = \frac{\bar{X} - \mu_0}{\hat{S}} \sqrt{n},$$

die einer t-Verteilung mit  $n - 1$  Freiheitsgraden genügt und spricht daher auch von einem **t-Test**. Die Konstruktion des kritischen Bereichs  $K^* = \{t \in \mathbb{R} : |t| > t_{n-1, 1-\frac{\alpha}{2}}\}$  oder die Berechnung des p-Wertes  $p = P(|T| \geq |t|)$  basieren statt der Standardnormalverteilung nun auf der t-Verteilung mit  $n - 1$  Freiheitsgraden. Das Vorgehen unterscheidet sich allerdings nicht zum Gauß-Test. Die folgenden beiden Grafiken illustrieren die Ähnlichkeit der beiden Testverteilungen in Abhängigkeit von der Stichprobengröße  $n$  für  $\alpha = 0.05$ . In der linken Grafik ist deutlich zu erkennen, dass der rot dargestellte Verwerfungsbereich der Nullhypothese des t-Tests für kleine Stichproben aufgrund der zusätzlichen Unsicherheit durch die Unkenntnis der Varianz  $\sigma^2$  sich im Vergleich zum Gauß-Test verkleinert, da die t-Verteilung im Vergleich zur Standardnormalverteilung breitere Ränder besitzt. Für  $n = 30$  ist die Schätzung von  $\sigma^2$  aus den Daten allerdings schon deutlich zuverlässiger, so dass sich der Verwerfungsbereich des t-Tests entsprechend nur noch geringfügig von dem des Gauß-Tests unterscheidet, wie in der rechten Grafik zu sehen ist. Ein Grund hierfür ist sicherlich in der Tatsache zu sehen, dass sich für  $n \geq 30$  die t-Verteilung kaum von der Standardnormalverteilung unterscheidet und folglich auch die entsprechenden Quantile sehr nahe beieinander liegen.

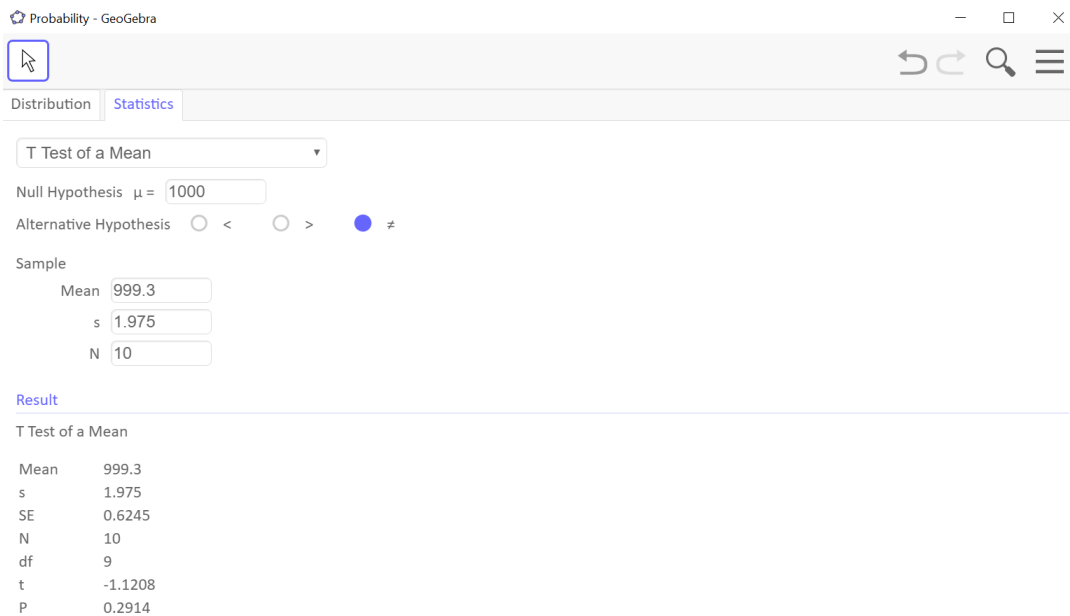
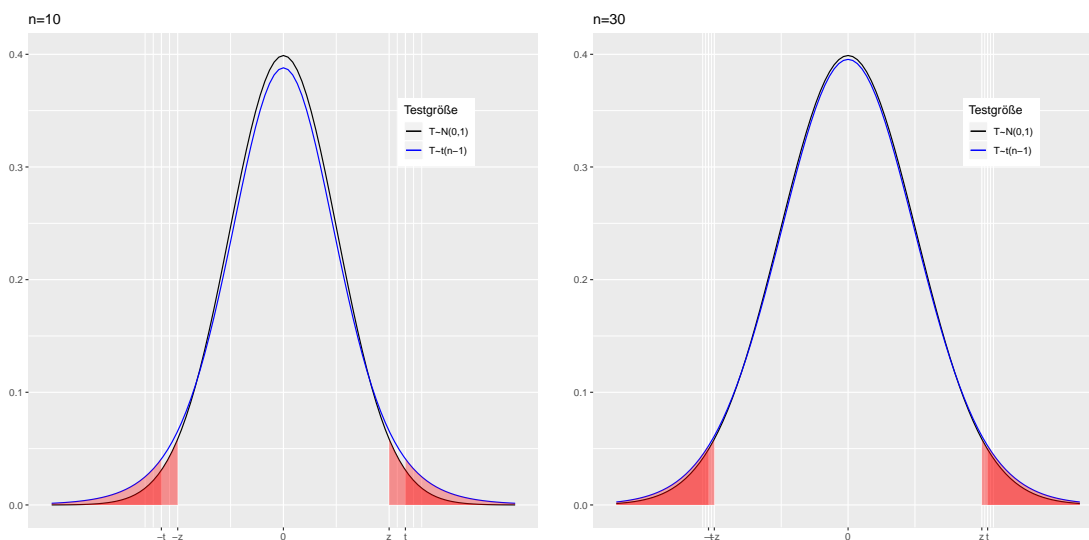


Abbildung 10: Durchführung eines t-Tests mit GeoGebra



#### Beispiel 4.6: Qualitätsbeurteilung eines neuen Streckenmessgeräts

Bei der Messung einer 1 km lange Referenzstrecke wurden mit dem neuen Testgerät die folgenden Werte (in Meter) ermittelt

(998.0, 1001.0, 1003.0, 1000.5, 999.0, 997.5, 1000.0, 999.5, 996.0, 998.5).

Daraus ergeben sich die empirischen Kenngrößen  $\bar{x} = 999.3(m)$ ,  $s^2 = 3.9(m^2)$  und  $s = 1.975(m)$ . Unter der Annahme, dass die „gemessene Länge“  $X \sim \mathcal{N}(\mu, \sigma^2)$  einer Normalverteilung folgt, wollen wir mit 5%–iger Irrtumswahrscheinlichkeit untersuchen, ob das Messgerät im Mittel richtig misst. Hierfür stellen wir die zweiseitigen Hypothese  $H_0 : \mu = \mu_0 = 1000(m)$  gegen Alternativhypothese  $H_1 : \mu \neq \mu_0$  auf, da sowohl zu kleine als auch zu große Messwerte unerwünscht sind. Typischerweise ist uns die Varianz  $\sigma^2$  unbekannt, so dass wir einen zweiseitigen t-Test mit kritischen Bereich

$$K^* = \{t : |t| > t_{n-1, 1-\frac{\alpha}{2}} = t_{9, 0.975} = 2.262\} = (-\infty, -2.262) \cup (2.262, \infty)$$

durchführen. Aus den Stichprobenwerten berechnen wir den konkreten Testwert

$$t = T(x_1, \dots, x_{10}) = \sqrt{n} \frac{\bar{x} - \mu_0}{s} = \sqrt{10} \frac{999.3 - 1000}{1.975} = -1.12 \notin K^*.$$

Da dieser nicht im kritischen Bereich liegt, hat der Test bei einer Irrtumswahrscheinlichkeit von 5% keine Einwände gegen die Behauptung, dass das Messgerät im Mittel korrekt arbeitet<sup>a</sup>.

<sup>a</sup>Analog gilt für den p-Wert  $p = P(|T| > |t|) = P(T < -1.12) + P(T > 1.12) = 0.2917 > 0.05 = \alpha$ . Diesen kann man auch mittels Geogebra einfach berechnen, wie Abbildung 10 zeigt.

Zum Abschluss dieses Abschnittes sei vermerkt, dass es natürlich auch Tests und Konfidenzintervalle für die Varianz  $\sigma^2$  des normalverteilten Merkmals  $X \sim \mathcal{N}(\mu, \sigma^2)$  gibt. Die zugehörigen Testgrößen führen auf die Chiquadratverteilung, sollen aber hier nicht weiter besprochen werden.

### 4.3.3 Binomialtest für den Parameter $p$ der Binomialverteilung

Um Hypothesen über den Parameter  $p$  des Merkmals  $X \sim \text{Bin}(m, p)$  mittels der Stichprobe  $(X_1, \dots, X_n)$  zu untersuchen, wird wie bei der Konstruktion von Konfidenzintervallen die binomialverteilte Testgröße

$$T = \sum_{i=1}^n X_i \sim \text{Bin}(nm, p)$$

betrachtet. Bei dem einführenden Beispiel auf Seite 35 sind wir der Frage nachgegangen, ob unserer Gegenspieler einen fairen oder gezinkten Würfel verwendet hat. Nach 20 Würfeln wurden insgesamt 15 Sechsen registriert, so dass wir sehr viel Geld bei diesem Glücksspiel verloren haben. Auch wenn bisher noch nicht erwähnt, haben wir letztendlich einen rechtsseitigen Binomialtest mit den Hypothesen  $H_0 : p = 1/6$  gegen  $H_1 : p > 1/6$  durchgeführt. Der p-Wert berechnet sich für den konkreten Wert  $t = \sum_{i=1}^n x_i$  mittels

$$P(T \geq t) = \sum_{k=t}^{nm} \binom{nm}{k} p_0^k (1-p_0)^{nm-k}.$$

Das prinzipielle Vorgehen bei Binomialtests entspricht der in Abschnitt 4.3 beschriebenen Methode. Ebenso die Konstruktion des kritischen Bereiches. Da die Testverteilung hier aber diskret und nicht stetig ist, ergeben sich, wie bei der Konstruktion von Konfidenzintervallen, Schwierigkeiten die Quantile zu bestimmen. Wird beispielsweise die einseitige Hypothese  $H_0 : p = p_0$  gegen  $H_1 : p > p_0$  untersucht, sprechen zu große Werte von  $T$  gegen die Nullhypothese, so dass der kritische Bereich mit Hilfe des Quantils  $k_{1-\alpha} \in \{0, 1, 2, \dots, n\}$ , mit

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

gebildet wird. Für große Stichprobenumfänge  $n$  oder große Parameter  $m$  kann die Normalapproximation  $T \approx \mathcal{N}(nmp_0, nmp_0(1-p_0))$  des Grenzwertsatzes von Moivre und Laplace genutzt werden, so dass dann die Quantile des kritischen Bereiches auf den Standardnormalverteilungsquantilen basieren<sup>27</sup>. Hierzu betrachtet man die standardisierte Testgröße

$$Z = \frac{T - nmp_0}{\sqrt{nmp_0(1-p_0)}} = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}} \sqrt{nm} \approx \mathcal{N}(0, 1)$$

und konstruiert analog zum Gauß-Test den kritischen Bereich. Für den rechtsseitigen Test gilt beispielsweise  $\alpha = P(Z > z_{1-\alpha})$ , so dass wir daraus den kritischen Bereich

$$K^* = \{t : t \geq z_{1-\alpha} \sqrt{nmp_0(1-p_0)} + nmp_0\}$$

für die Testgröße  $T$  ableiten können.

<sup>27</sup>Die Varianz  $\text{Var}(T) = nmp_0(1-p_0)$  ist unbekannt, so dass die Standardisierung nicht auf eine Standardnormalverteilung, sondern t-Verteilung führt. Allerdings muss  $nm$  groß sein für die Normalapproximation und für große Freiheitsgrade unterscheiden sich die t-Verteilung und die Standardnormalverteilung nur unwesentlich.

#### Aufgabe 4.14

Von einem Studenten wurde erwartet, dass er 80% eines bestimmten Wissensgebietes beherrscht. Bei einer Prüfung, in der ihm 20 voneinander unabhängige und gleich schwere Fragen gestellt wurden, konnte er allerdings nur 13 Fragen richtig beantworten.

Widerspricht dieses Prüfungsergebnis bei einer vorgegebenen Irrtumswahrscheinlichkeit  $\alpha = 0.05$  der Anfangserwartung?

Gilt die getroffene Aussage auch dann, wenn unter gleichen Bedingungen von 40 Fragen 26 richtig beantwortet werden?

Lösung auf Seite ??

#### Aufgabe 4.15

Übergewicht ist bekanntlich ein Risikofaktor für die Gesundheit. Nach einer älteren Studie sind 15% der Bevölkerung davon betroffen. Bei einer neuen Studie wurden 500 Personen untersucht, von denen 430 normalgewichtig waren. Kann man aufgrund dieser Untersuchung darauf schließen, dass sich der Anteil Übergewichtiger verringert hat?

Lösung auf Seite ??

#### Aufgabe 4.16

Im letzten Jahr hatte die Computerfirma „Digitoll“ einen Marktanteil von 11%. Nach einer Werbekampagne möchte der Manager wissen, ob zusätzliche Marktanteile gewonnen wurden. Hierzu wurden 300 Personen befragt, die im letzten Monat einen Computer gekauft haben. Darunter haben sich 40 für das Modell von „Digitoll“ entschieden. Entscheiden Sie mit einem statistischen Test, ob die Marketingkampagne erfolgreich war.

Lösung auf Seite ??

#### Aufgabe 4.17

Ein Unternehmen weiß, dass seine Produkte im vergangenen Geschäftsjahr einen Bekanntheitsgrad von 40% hatten. Bei einer Umfrage im aktuellen Geschäftsjahr kannten von 478 Personen nur 182 die Produkte des Unternehmens. Daraufhin diskutiert die Unternehmensleitung das Werbebudget zu erhöhen. Beraten Sie das Management mit Hilfe eines Parametertests mit 5% Irrtumswahrscheinlichkeit.

Lösung auf Seite ??

#### Aufgabe 4.18

Der Anteil der Personen, die ein Handy besitzen, beträgt 80%. Eine repräsentative Stichprobe ergibt, dass von 420 befragten Akademikern 350 ein Handy besitzen. Daher vermutet man, dass in der Gruppe der Personen mit Hochschulabschluss der Anteil der Handybesitzer über 80% liegt. Erklären Sie, um welchen Test es sich handelt und führen Sie diesen mit einer Irrtumswahrscheinlichkeit von 5% aus.

Lösung auf Seite ??

#### 4.3.4 Tests für den Erwartungswert beliebiger Verteilungen

Allgemein können für große Stichprobenumfänge  $n$  Aussagen über den unbekanntem Erwartungswert  $\mu = EX$  Hypothesen untersucht werden, da nach dem Zentralen Grenzwertsatz das arithmetische Mittel näherungsweise normalverteilt

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

ist. Somit können je nach Kenntnis über die Varianz  $\sigma^2$  der Gauß-Test oder t-Test für beliebig verteilte Zufallsvariablen  $X$  durchgeführt werden.

## 5 Gemischte Aufgaben

### Aufgabe 5.1

Die Masse der Packunen des Waschmittels „Riesenweiß“ sei normalverteilt mit Erwartungswert 3200 Gramm und Standardabweichung 100 Gramm.

- Berechnen Sie die Wahrscheinlichkeit, dass eine Packung weniger als 3150 Gramm Masse besitzt.
- Der Produzent möchte garantieren, dass höchstens 9% aller Packungen um mehr als  $c$  Gramm vom Erwartungswert abweichen. Wie muss  $c$  gewählt werden?
- Der Produzent möchte eine neue Abfüllmaschine kaufen, um sicherzustellen, dass bei unverändertem Erwartungswert höchstens 4% aller Packungen eine Masse von mehr als 3350 Gramm besitzen. Welche Standardabweichung muss die neue Maschine haben?

Lösung auf Seite ??

### Aufgabe 5.2

Analysieren Sie die folgenden Fragestellungen mit Hilfe geeigneter statistischer Testverfahren. Als Irrtumswahrscheinlichkeit soll  $\alpha = 0.05$  gewählt werden.

- In einer Umfrage während einer Vorlesung gaben 55% der männlichen Studenten an, dass Sie in Sport die Note sehr gut hatten. Von 104 weiblichen Studenten hatten 35 die Note sehr gut. Lässt sich statistisch absichern, dass der Anteil unter den weiblichen Studenten mit der Note sehr gut kleiner als 55% ist?
- Was ändert sich in (a), wenn 52 von 104 weiblichen Studenten angaben, in Sport die Note sehr gut erhalten zu haben?
- In der letzten Klausur hatten 47% der männlichen Studenten die Note gut oder sehr gut. Bei den weiblichen Studenten waren es 64 von 106. Lässt sich statistisch absichern, dass der Anteil unter den weiblichen Studenten mit der Note gut oder sehr gut größer als 47% ist?
- In Deutsch hatten 50% der männlichen Studenten die Note gut oder sehr gut. Bei den weiblichen Studenten 68 von 106. Lässt sich statistisch absichern, dass der Anteil unter den weiblichen Studenten mit der Note gut oder sehr gut verschieden von 50% ist?

Lösung auf Seite ??

### Aufgabe 5.3

Eine Fahrschule behauptet, dass 25% aller Fahrschüler beim ersten Versuch durch die Fahrprüfung fallen. Berechnen Sie die Wahrscheinlichkeit, dass bei 25%-tiger Durchfallquote dann bei einer Gruppe von 50 Fahrschülern

- höchstens 14 durchfallen,
- genau 37 durchkommen,
- höchstens 18 und mindestens 8 durchfallen.
- Wie groß ist die Irrtumswahrscheinlichkeit  $\alpha$ , wenn obige Hypothese anhand von 50 zufällig ausgewählten Fahrschülern getestet wird? Die Hypothese wird dabei verworfen, wenn mehr als 16 Personen durchfallen.

Lösung auf Seite ??