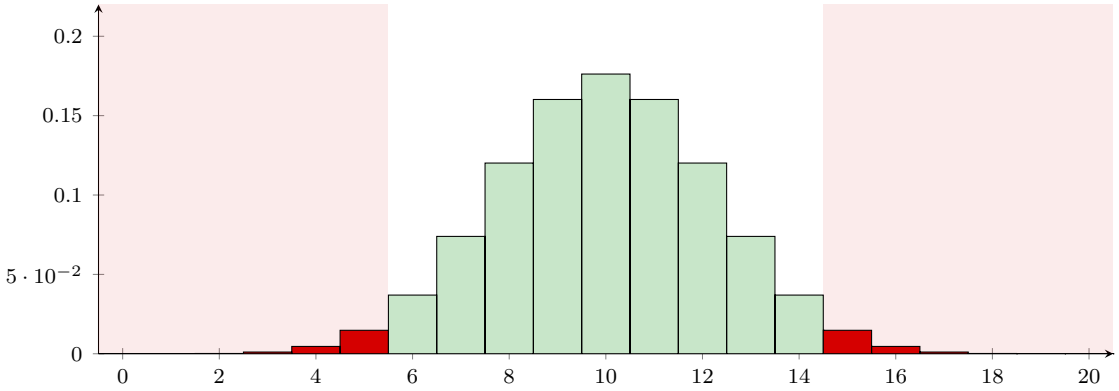
Beurteilende Statistik

Stefan Rothe

14.11.2026



Verwerfungsbereich V (rot) und Annahmebereich (grün)

Inhaltsverzeichnis

1 Statistischer Test **2**
2 Hypothesen **2**
3 Nullhypothese **3**
4 Gegenhypothese **3**
5 Fehler und Signifikanzniveau **3**
6 Verwerfungsbereich **4**
7 Auswertung des Tests **6**
8 Abschätzung des Verwerfungsbereichs **6**
9 Berechnung des Fehlers 1. Art **8**

© Gymnasium Kirchenfeld
Dieses Skript steht unter einer Creative Commons Attribution 4.0 International-Lizenz.
(CC BY 4.0)

1 Statistischer Test

Es gibt eine breite Auswahl von statistischen Tests. Hier werden ausschliesslich Signifikanztests für die Binomialverteilung betrachtet.

Beispiel. *Es soll wissenschaftlich überprüft werden, ob Testpersonen hellseherische Fähigkeiten haben. Dazu wird einer Person 20-mal die Rückseite einer zufällig gewählten Spielkarte gezeigt. Die Person muss die Spielfarbe der Karte (Herz, Schaufel, Kreuz oder Ecke) korrekt nennen.*

Eine Person kann die Farbe insgesamt 8-mal korrekt nennen. Hat diese Person nun hellseherische Fähigkeiten?

Eine solche Fragestellung kann mit einem statistischen Test beurteilt werden. Ein statistischer Test überprüft, ob ein beobachtetes Ergebnis eines Zufallsexperiments mit einer vermuteten Wahrscheinlichkeitsverteilung übereinstimmt.

Dazu muss zunächst die vermutete **Verteilung** angegeben werden.

Beispiel. *Jede korrekt genannte Kartenfarbe wird als Erfolg angeschaut. Wenn angenommen wird, dass die Erfolgswahrscheinlichkeit p immer gleich ist, stellt das Experiment eine Bernoullikette dar.*

Die Zufallsgrösse X stellt die Anzahl korrekt genannter Karten dar und ist binomialverteilt mit $n = 20$. Unter der Annahme, dass nur geraten wird, ist die Erfolgswahrscheinlichkeit $p = 0.25$.

2 Hypothesen

Ein statistischer Test stellt zwei Hypothesen auf und überprüft, ob eine davon verworfen werden kann. Bei Signifikanztests heissen diese **Nullhypothese** H_0 und **Gegenhypothese** H_1.

Der Test überprüft, ob die Nullhypothese verworfen werden kann. Wird sie verworfen, bedeutet das nicht, dass die Gegenhypothese bewiesen ist. Es bedeutet, dass die beobachteten Daten unter der Nullhypothese sehr unwahrscheinlich wären.

Beispiel. *Die beiden Hypothesen lauten umgangssprachlich:*

- *Nullhypothese: Die Testperson hat keine hellseherischen Fähigkeiten.*
- *Gegenhypothese: Die Testperson hat hellseherische Fähigkeiten.*

Falls der Test die Nullhypothese verwirft, bedeutet das nicht, dass hellseherische Fähigkeiten bewiesen sind. Es bedeutet, dass das Ergebnis mit der Nullhypothese kaum vereinbar ist.

Falls der Test die Nullhypothese nicht verwirft, liefern die Daten keinen Hinweis auf hellseherische Fähigkeiten. Das beweist aber nicht, dass keine solchen Fähigkeiten vorhanden sind.

3 Nullhypothese

Allgemein: Die Nullhypothese besagt, dass die Zufallsgrösse die vermutete Verteilung oder Eigenschaft hat.

Binomialverteilung: Die Nullhypothese H_0 besagt, dass die Erfolgswahrscheinlichkeit p einer angenommenen Wahrscheinlichkeit p_0 entspricht.

Beispiel. *Falls die Testperson keine hellseherischen Fähigkeiten hat, errät sie die Kartenfarbe zufällig mit der Erfolgswahrscheinlichkeit $p_0 = 0.25$. Die Nullhypothese lautet also*

$$H_0 \colon p = 0.25$$

4 Gegenhypothese

Allgemein: Bei einem statistischen Test wird die Nullhypothese immer einer Alternative gegenübergestellt. Diese wird Gegenhypothese H_1 genannt. Sie besagt, dass die Zufallsgrösse eine abweichende Verteilung oder Eigenschaft hat.

Binomialverteilung: Es gibt drei Möglichkeiten, die Gegenhypothese zu formulieren:

- $p > p_0$ (rechtsseitiger Test): Es wird vermutet, dass die tatsächliche Erfolgswahrscheinlichkeit p grösser als diejenige der Nullhypothese p_0 ist.
- $p < p_0$ (linksseitiger Test): Es wird vermutet, dass die tatsächliche Erfolgswahrscheinlichkeit p kleiner als diejenige der Nullhypothese p_0 ist.
- $p \neq p_0$ (zweiseitiger Test): Es wird vermutet, dass die tatsächliche Erfolgswahrscheinlichkeit p nicht gleich derjenigen der Nullhypothese p_0 ist.

Beispiel. *Wenn die Testperson hellseherische Fähigkeiten hat, so ist ihre Erfolgswahrscheinlichkeit grösser als 25 %. Es wird also ein rechtsseitiger Test benötigt, die Gegenhypothese lautet:*

$$H_1 \colon p > 0.25$$

5 Fehler und Signifikanzniveau

Mit einem statistischen Test kann niemals mit Sicherheit geurteilt werden, welche der beiden Hypothesen wahr ist. Es besteht immer die Möglichkeit, dass die Nullhypothese verworfen wird, obschon sie wahr ist. Es gibt zwei Möglichkeiten, sich zu irren:

	H_0 ist wahr	H_1 ist wahr
H_0 wird verworfen	Fehler 1. Art	richtiger Entscheid
H_0 wird nicht verworfen	richtiger Entscheid	Fehler 2. Art

Ein Fehler 1. Art bedeutet, dass die Nullhypothese fälschlicherweise verworfen wird. Ein Fehler 2. Art bedeutet, dass die Nullhypothese irrtümlicherweise nicht verworfen wird, obschon die Gegenhypothese wahr ist.

Das **Signifikanzniveau** α ist die maximal zulässige Wahrscheinlichkeit dafür, dass die Nullhypothese verworfen wird, obschon sie wahr ist (Fehler 1. Art).

Beispiel. *Der Fehler 1. Art ist es, bei einer Testperson ohne hellseherische Fähigkeiten die Hypothese zu verwerfen, dass sie keine solchen Fähigkeiten hat, also irrtümlicherweise bei einer «normalen» Person übernatürliche Fähigkeiten zu vermuten.*

Die Wahrscheinlichkeit für einen solchen Fehler soll maximal 5 % *betragen, also wird das Signifikanzniveau* $\alpha = 0.05$ *festgelegt.*

6 Verwerfungsbereich

Der Verwerfungsbereich ist die Menge der Werte für X, für welche die Nullhypothese mit dem gewählten Signifikanzniveau verworfen werden kann.

Um den Verwerfungsbereich zu bestimmen, wird je nach Art des Tests unterschiedlich vorgegangen.

6.1 Linksseitiger Test

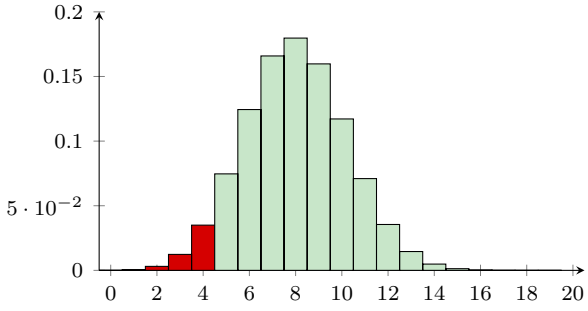
Gesucht wird der grösste mögliche Wert k für die Zufallsgrösse X, für welchen $P(X \leq k) \leq \alpha$ gilt. Der Verwerfungsbereich ist dann $V = \{0;\ \ldots;\ k\}$.

Beispiel. *Der Verwerfungsbereich des folgenden Tests ist zu bestimmen:*

$$H_0\colon p = 0.4 \qquad H_1\colon p < 0.4 \qquad n = 20 \qquad \alpha = 10\,\%$$

Vorgehen:

$$\begin{aligned} P(X \leq 0) &= 0.00004 \\ P(X \leq 1) &= 0.00052 \\ P(X \leq 2) &= 0.00361 \\ P(X \leq 3) &= 0.01596 \\ P(X \leq 4) &= 0.05095 \\ P(X \leq 5) &= 0.12560 \\ \Longrightarrow \quad V &= \{0;\ \ldots;\ 4\} \end{aligned}$$



Bei $k = 5$ *wird die Schwelle von* α *überschritten, also ist 4 die grösste Zahl im Verwerfungsbereich.*

Obschon die Wahrscheinlichkeit $P(X = 5)$*, welche durch den ersten grünen Balken dargestellt wird, noch kleiner als 0.1 ist, liegt die kumulative Wahrscheinlichkeit aller* X *von 0 bis 5 schon knapp darüber.*

6.2 Rechtsseitiger Test

Beim rechtsseitigen Test muss auf der rechten Seite der Verteilung eine Wahrscheinlichkeit von höchstens α übrigbleiben. Da aber nur linksseitige kumulative Wahrscheinlichkeiten berechnet werden können, wird das Gegenereignis betrachtet. Wir suchen die kleinste Zahl k, sodass $P(X \leq k) \geq 1 - \alpha$ gilt.

$$P(X \leq k) \geq 1 - \alpha \quad \Longleftrightarrow \quad P(X \geq k + 1) \leq \alpha$$

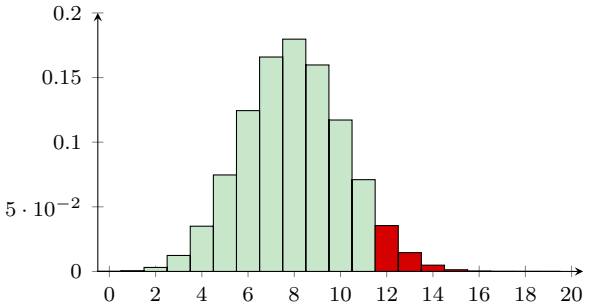
Der Verwerfungsbereich ist dann $V = \{k + 1;\ \ldots;\ n\}$.

Beispiel. *Der Verwerfungsbereich des folgenden Tests ist zu bestimmen:*

$$H_0\colon p = 0.4 \qquad H_1\colon p > 0.4 \qquad n = 20 \qquad \alpha = 10\,\%$$

Vorgehen:

$$\begin{aligned} P(X \leq 9) &= 0.7553 \\ P(X \leq 10) &= 0.8725 \\ P(X \leq 11) &= 0.9435 \\ P(X \leq 12) &= 0.9790 \\ \Longrightarrow \quad V &= \{12;\ \ldots;\ 20\} \end{aligned}$$



Bei $k = 11$ liegt die kumulative Wahrscheinlichkeit mit $P(X \leq 11) = 0.9435$ erstmals über der Schwelle $1 - \alpha = 0.9$, während $P(X \leq 10) = 0.8725$ noch darunter liegt. Da 11 zum Gegenereignis des Verwerfungsbereichs gehört, ist 12 dessen kleinste Zahl.

6.3 Zweiseitiger Test

Beim zweiseitigen Test wird das Signifikanzniveau α gleichmässig auf beide Seiten aufgeteilt. Die beiden Seiten werden wie ein links- bzw. rechtsseitiger Test behandelt. Am Schluss werden die beiden Verwerfungsbereiche zusammengefasst.

Es werden also der grösste Wert k_L und der kleinste Wert k_R gesucht mit

$$P(X \leq k_L) \leq \frac{\alpha}{2} \qquad\qquad P(X \leq k_R) \geq 1 - \frac{\alpha}{2}$$

Dann ist der kombinierte Verwerfungsbereich $V = \{0;\ \ldots;\ k_L\} \cup \{k_R + 1;\ \ldots;\ n\}$.

Beispiel. *Der Verwerfungsbereich des folgenden Tests ist zu bestimmen:*

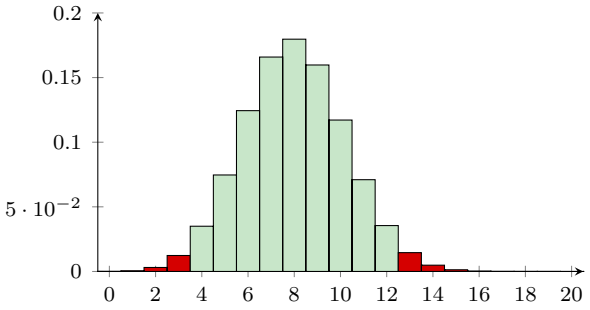
$$H_0\colon p = 0.4 \qquad H_1\colon p \neq 0.4 \qquad n = 20 \qquad \alpha = 10\,\%$$

Vorgehen:

$$P(X \leq 3) = 0.01596$$
$$P(X \leq 4) = 0.05095$$
$$\Longrightarrow \quad V_L = \{0;\ \ldots;\ 3\}$$

$$P(X \leq 11) = 0.9435$$
$$P(X \leq 12) = 0.9790$$
$$\Longrightarrow \quad V_R = \{13;\ \ldots;\ 20\}$$

$V = V_L \cup V_R = \{0;\ \ldots;\ 3\} \cup \{13;\ \ldots;\ 20\}$



7 Auswertung des Tests

Wenn der Verwerfungsbereich ermittelt worden ist, steht der Test zur Durchführung bereit. Nun wird das Experiment n-mal durchgeführt und die Anzahl Erfolge x gezählt.

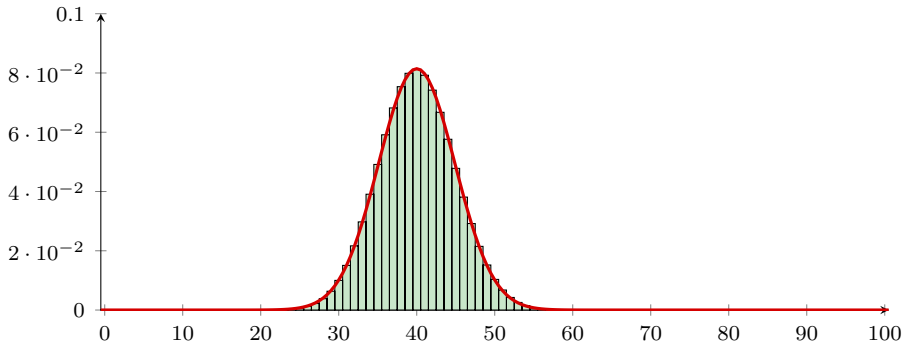
Danach wird überprüft, ob die Anzahl Erfolge x im Verwerfungsbereich liegt.

x **liegt im Verwerfungsbereich.** Das bedeutet, dass das Experiment Daten liefert, die ausgehend von der Nullhypothese H_0 als sehr unwahrscheinlich erscheinen. Deshalb wird in diesem Fall die Nullhypothese **verworfen**, die Gegenhypothese H_1 kann als plausibel betrachtet werden. Dabei ist zu beachten, dass man sich bei dieser Entscheidung mit einer Wahrscheinlichkeit von maximal α irren kann (Fehler 1. Art).

x **liegt nicht im Verwerfungsbereich.** Das bedeutet, dass die Daten ausgehend von der Nullhypothese H_0 plausibel erscheinen. Somit kann die Nullhypothese **nicht verworfen** werden, und die Gegenhypothese H_1 wird durch die Daten nicht gestützt.

8 Abschätzung des Verwerfungsbereichs

Nach dem Satz von Moivre-Laplace konvergiert die standardisierte Zufallsgrösse $(X - np)/\sigma$ bei festem p für $n \to \infty$ gegen die Standardnormalverteilung. Das bedeutet, dass für grosse Stichproben die Binomialverteilung durch die Normalverteilung angenähert werden kann. Der Satz kann angewendet werden, wenn die Standardabweichung der Binomialverteilung grösser als drei ist.



Laplace-Bedingung. Eine Binomialverteilung mit n Versuchen und einer Erfolgswahrscheinlichkeit von p kann durch die Normalverteilung angenähert werden, falls

$$\sigma = \sqrt{np(1-p)} > 3$$

Falls die Laplace-Bedingung erfüllt ist, kann mit Hilfe der Normalverteilung die Schranke k für die gewünschte kumulative Wahrscheinlichkeit φ näherungsweise berechnet werden. Zunächst wird das reelle Quantil q der Normalverteilung mit

$$P(X \leq q) \approx \varphi$$

bestimmt und daraus das ganzzahlige k abgeleitet.

Für das Quantil müssen zunächst Erwartungswert und Standardabweichung der Binomialverteilung berechnet werden:

$$\mu = np \qquad \sigma = \sqrt{np(1-p)}$$

Die Schranke k für eine kumulative Wahrscheinlichkeit φ kann wie folgt mit der Normalverteilung abgeschätzt werden:

TI-30X Plus: Schranke für kumulative Binomialverteilung abschätzen

1. Mit 2nd data das Stochastikmenü öffnen.
2. Mit ▶ das Untermenü **DISTR** auswählen.
3. Mit 2×▼ enter den Menüpunkt **invNormal** auswählen.
4. Gewünschte kumulative Wahrscheinlichkeit φ bei **area** eingeben.
5. Erwartungswert μ bei **mean=mu** eingeben.
6. Standardabweichung σ bei **sigma** eingeben.
7. Mit ▼ enter **CALC** auswählen.

Beispiel. *Der Verwerfungsbereich des folgenden Tests soll bestimmt werden:*

$$H_0\colon p = 0.4 \qquad H_1\colon p \neq 0.4 \qquad n = 100 \qquad \alpha = 10\,\%$$

Erwartungswert und Standardabweichung für diese Verteilung lauten:

$$\mu = 100 \cdot 0.4 = 40 \qquad \sigma = \sqrt{100 \cdot 0.4 \cdot 0.6} \approx 4.8990$$

Für den linken Verwerfungsbereich wird das grösste k_L mit $P(X \leq k_L) \leq \frac{\alpha}{2} = 0.05$ gesucht.

Das reelle Quantil der Normalverteilung für $\varphi = 0.05$ ist $q_L \approx 31.94$. Mit der Stetigkeitskorrektur – die 0.5 trägt der Ganzzahligkeit von X Rechnung – wird auf $k_L = \lfloor q_L - 0.5 \rfloor = 31$ abgerundet. Dies wird mit der kumulativen Binomialverteilung überprüft:

$$P(X \leq 31) = 0.0398$$
$$P(X \leq 32) = 0.0615$$

Damit ist der linke Verwerfungsbereich $V_L = \{0;\ \ldots;\ 31\}$.

Für den rechten Verwerfungsbereich ist $\varphi = 0.95$. Das reelle Quantil ist $q_R \approx 48.06$. Mit der Stetigkeitskorrektur wird auf $k_R = \lceil q_R - 0.5 \rceil = 48$ aufgerundet. Dies wird mit der kumulativen Binomialverteilung überprüft:

$$P(X \leq 48) = 0.9577$$
$$P(X \leq 47) = 0.9362$$

Hier ist der Wert 48 der erste, bei welchem die kumulative Wahrscheinlichkeit über 95 % liegt. Also ist der rechte Verwerfungsbereich $V_R = \{49;\ \ldots;\ 100\}$.

Der gesamte Verwerfungsbereich für diesen Test lautet somit

$$V = V_L \cup V_R = \{0;\ \ldots;\ 31\} \cup \{49;\ \ldots;\ 100\}$$

9 Berechnung des Fehlers 1. Art

Das Signifikanzniveau α legt nur die maximal zulässige Wahrscheinlichkeit für einen Fehler 1. Art fest. Die tatsächliche Wahrscheinlichkeit hängt vom Verwerfungsbereich ab und wird unter der Nullhypothese berechnet:

$$P(\text{Fehler 1. Art}) = P(X \in V \mid H_0) = \sum_{k \in V} \binom{n}{k} p_0^k (1 - p_0)^{n-k}$$

Da der Verwerfungsbereich aus den Randbereichen der Verteilung besteht, genügt die Berechnung kumulativer Wahrscheinlichkeiten:

- Linksseitiger Test ($V = \{0;\ \ldots;\ k\}$): $P(\text{Fehler 1. Art}) = P(X \leq k)$.
- Rechtsseitiger Test ($V = \{k+1;\ \ldots;\ n\}$): $P(\text{Fehler 1. Art}) = P(X \geq k+1) = 1 - P(X \leq k)$.
- Zweiseitiger Test ($V = V_L \cup V_R$): Die Wahrscheinlichkeiten der beiden Verwerfungsbereiche werden addiert.

Beispiel. *Für den linksseitigen Test aus Abschnitt 6.1 gilt* $n = 20$*,* $p_0 = 0.4$*,* $\alpha = 10\,\%$ *und* $V = \{0;\ \ldots;\ 4\}$*. Die tatsächliche Wahrscheinlichkeit für einen Fehler 1. Art ist*

$$P(X \leq 4) = 0.05095 \approx 5.1\,\%.$$

Sie liegt deutlich unter dem Signifikanzniveau von 10 %*, weil* $P(X \leq 5) = 0.1256$ *die Schwelle bereits überschreitet und der Verwerfungsbereich deshalb nicht grösser werden kann.*

Beispiel. *Für den zweiseitigen Test aus Abschnitt 6.3 gilt* $n = 20$*,* $p_0 = 0.4$*,* $\alpha = 10\,\%$ *und* $V = \{0;\ \ldots;\ 3\} \cup \{13;\ \ldots;\ 20\}$*. Damit ist*

$$\begin{aligned} P(\textit{Fehler 1. Art}) &= P(X \leq 3) + P(X \geq 13) \\ &= 0.01596 + (1 - 0.9790) \\ &= 0.03699 \approx 3.7\,\%. \end{aligned}$$

Die tatsächliche Wahrscheinlichkeit liegt häufig deutlich unter α, weil die Binomialverteilung diskret ist: Die kumulierte Wahrscheinlichkeit nimmt sprungweise zu und überschreitet die Schranke α oft erst mit einem grösseren Sprung. Mit wachsender Kettenlänge n werden diese Sprünge kleiner, und die tatsächliche Wahrscheinlichkeit nähert sich dem Signifikanzniveau an.