Bestimmung von Vertrauensintervallen (Konfidenzintervallen) bei unbekannten Wahrscheinlichkeiten

Beispiel 1: Es soll untersucht werden, wie viele 18-jährige Erstwähler bei der nächsten Bundestagswahl wählen gehen werden. Dazu werden 600 Personen befragt, 493 von ihnen wollen wählen gehen.
p ist die bisher unbekannte Wahlbeteiligung der 18jährigen am Wahltag.
Sind die Hypothesen a) p = 0,79 bzw. b) p = 0,85 auf 5% Signifikanzniveau mit der Umfrage vereinbar?

a. $p = 0{,}79$; $n = 600$; $\mu = n\,p = 474$; $\sigma = 9{,}977$; Signifikanzniveau $\alpha = 0{,}05 \Rightarrow k = 1{,}96$
$474 - 1{,}96 \cdot 9{,}977 = 454{,}445$; $474 + 1{,}96 \cdot 9{,}977 = 493{,}555$
Annahmebereich [455 ; 493] (zur sicheren Seite gerundet)
493 liegt gerade noch im Annahmebereich (an der oberen Grenze), daher ist die Hypothese mit der Umfrage vereinbar.

b. $p = 0{,}85$; $n = 600$; $\mu = n\,p = 510$; $\sigma = 8{,}746$; Signifikanzniveau $\alpha = 0{,}05 \Rightarrow k = 1{,}96$
$510 - 1{,}96 \cdot 8{,}746 = 492{,}852$; $510 + 1{,}96 \cdot 8{,}746 = 527{,}142$
Annahmebereich [493 ; 527] (zur sicheren Seite gerundet)
493 liegt schon im Annahmebereich (an der unteren Grenze), daher ist die Hypothese mit der Umfrage vereinbar.

Offenbar sind Werte für p zwischen 0,79 und 0,85 mit der Hypothese vereinbar.
Diese Wahrscheinlichkeiten bilden ein Intervall, das **Vertrauensintervall** (Konfidenzintervall) für die Wahrscheinlichkeit p. Die Sicherheit (oder Vertrauenszahl, Vertrauensniveau) γ ergibt sich aus dem Signifikanzniveau α: $\gamma = 1 - \alpha$. Im Beispiel 1 ist $\gamma = 0{,}95$. Interpretation: In 95 von 100 Fällen liegt die wahre Wahrscheinlichkeit im Vertrauensintervall, in 5 von 100 Fällen aber auch nicht!

Bestimmen des Vertrauensintervalls

Beispiel 2: Es wird mit einem Quader „gewürfelt“, bei 481 Würfen erscheint 163 mal eine Vier. Wie groß ist die Wahrscheinlichkeit, mit diesem Quader eine Vier zu würfeln? ($\alpha = 5\%$)

$n = 481$, $X = 163$. Für die relative Häufigkeit der Vier ergibt sich damit $h_n = \frac{163}{481} \approx 0{,}3389$.
Damit p mit der gewürfelten relativen Häufigkeit verträglich ist, kann man es aus der folgenden Ungleichungskette bestimmen:

$$\mu - 1{,}96\sigma \le X \le \mu + 1{,}96\sigma$$

$$n \cdot p - 1{,}96\sqrt{n \cdot p \cdot (1-p)} \le X \le n \cdot p + 1{,}96\sqrt{n \cdot p \cdot (1-p)} \quad / : n$$

$$p - \frac{1{,}96}{n}\sqrt{n \cdot p \cdot (1-p)} \le \frac{X}{n} \le p + \frac{1{,}96}{n}\sqrt{n \cdot p \cdot (1-p)}$$

$$\left(\frac{X}{n} = h_n;\ \frac{1}{n}\ \textit{unter die Wurzel bringen} \to \frac{1}{n^2}\ \textit{und kürzen}\right)$$

$$p - 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}} \le h_n \le p + 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}} \quad (\textit{Ungleichung } *)$$

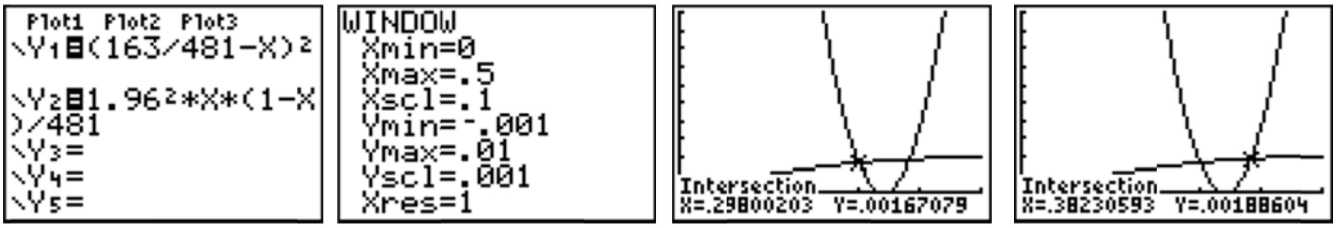
Alle Werte für p, die diese Beziehung erfüllen, definieren das Vertrauensintervall zu der relativen Häufigkeit h_n. Es enthält die unbekannte Wahrscheinlichkeit p mit einer Sicherheit von 95%. Für andere Sicherheiten (Vertrauenszahlen) muss k = 1,96 entsprechend angepasst werden.

Um eine **exakte Lösung für p** zu finden, subtrahiert man p und formt um in eine Betragsungleichung. Diese wird anschließend quadriert. Statt 1,96 wird allgemein k eingesetzt.

$$\left|h_n - p\right| \le k \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$

$$(h_n - p)^2 \le k^2 \, \frac{p \cdot (1-p)}{n}$$

Damit lässt sich die exakte Lösung zu Beispiel 2 mit dem GTR ermitteln (p wird durch X ersetzt):



also ist das Vertrauensintervall für p [0,2980 ; 0,3823].

Eine weitere Möglichkeit zur grafischen Bestimmung des exakten Vertrauensintervalls besteht in der Eingabe der Ungleichungskette (*)

$$p - 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}} \le h_n \le p + 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}}$$ als drei Funktionen Y1 ≤ Y2 ≤ Y3.

Mit den Eingaben p = X, n = 481, h = 163/481 aus Beispiel 2 erhält man folgende grafische Darstellung, wobei der Vorteil bei dieser Methode darin liegt, dass eine geeignete WINDOW-Einstellung nicht erst gesucht, sondern immer gleich gewählt werden kann (auf beiden Achsen von 0 bis 1).

Mit INTERSECT erhält man dann auf gewohnte Weise die Schnittpunkte der Horizontalen Y2 mit den beiden Graphen Y1 und Y3. Es ergibt sich das gleiche Vertrauensintervall [0,2980 ; 0,3823].

Eine gute **Näherungslösung** für p erhält man, wenn in der Ungleichung (*) oben im Wurzelterm p durch die relative Häufigkeit h_n ersetzt wird:

$$p - 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}} \le h_n \le p + 1{,}96\sqrt{\frac{p \cdot (1-p)}{n}} \qquad /\ p \textit{ im Wurzelterm durch } h_n \textit{ ersetzen}$$

$$p - 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} \le h_n \le p + 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}}$$

Linker Teil:

$$p - 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} \le h_n \qquad / + 1{,}96 \cdot \sqrt{\ }$$

$$p \le h_n + 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}}$$

Rechter Teil:

$$h_n \le p + 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} \qquad / - 1{,}96 \cdot \sqrt{\ }$$

$$h_n - 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} \le p$$

Insgesamt:

$$h_n - 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} \le p \le h_n + 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}}$$

Damit ist die **Näherung für das Vertrauensintervall**

$$\left[h_n - 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} ; h_n + 1{,}96\sqrt{\frac{h_n \cdot (1-h_n)}{n}} \right]$$

Für das Beispiel 2 ergibt sich mit $0{,}3389 \pm 1{,}96\sqrt{\frac{0{,}3389 * (1-0{,}3389)}{481}}$ das Intervall [0,2922 ; 0,3812], die Abweichung zu den oben ermittelten Werten ist gering.

Der Term $\sqrt{\frac{h_n \cdot (1-h_n)}{n}}$ kann durch Erweitern des Bruches mit n und teilweises Wurzelziehen umgeformt werden: $\sqrt{\frac{h_n \cdot (1-h_n)}{n}} = \sqrt{\frac{n \cdot h_n \cdot (1-h_n)}{n^2}} = \frac{\sqrt{n \cdot h_n \cdot (1-h_n)}}{n} = \frac{\sigma_n}{n}$, im Zähler steht die Standardabweichung von h_n. Setzt man noch $\varepsilon = k \cdot \frac{\sigma_n}{n}$, so lässt sich das Vertrauensintervall schreiben als $[h_n - \varepsilon ; h_n + \varepsilon]$.

Zusammenfassung:
Gegeben: Stichprobenumfang: n, absolute Häufigkeit eines Merkmals: X, Vertrauenszahl: γ

Dann ist $h_n = \frac{X}{n}$ die relative Häufigkeit des Merkmals.

$\varepsilon = k \cdot \frac{\sigma_n}{n}$; wobei k von γ abhängt (95%: k=1,96; 99%: k=2,58) und $\sigma_n = \sqrt{n \cdot h_n \cdot (1-h_n)}$ ist.

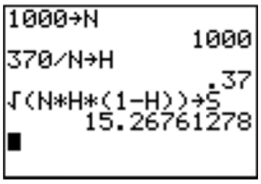
Damit ergibt sich als Näherung für das Vertrauensintervall $[h_n - \varepsilon ; h_n + \varepsilon]$.

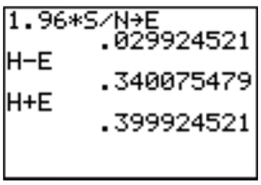
Je größer n ist, desto geringer ist die Länge des Vertrauensintervalls, da durch wachsendes n der Nenner von ε schneller wächst als der Zähler, denn dort steht n unter der Wurzel.

Je größer γ ist, desto länger ist das Vertrauensintervall, da ε proportional zu k ist.

Beispiel 3: Die Partei SGL gibt eine Umfrage in Auftrag. Eine repräsentative Stichprobe von 1000 Wählern ergibt 370 Wähler der SGL. Bestimmen Sie das 95%-Vertrauensintervall für den unbekannten Anteil p der SGL-Wähler in der Bevölkerung.

Näherungslösung:





$n = 1000$, gespeichert in N

$h_n = \frac{370}{1000} = 0{,}37$, gespeichert in H

$\sigma_n = \sqrt{1000 \cdot 0{,}37 \cdot (1-0{,}37)} = 15{,}2676$, gespeichert in S

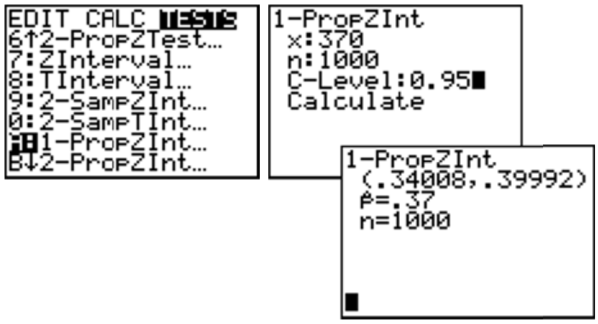
$\gamma = 0{,}95 \Rightarrow k = 1{,}96$

Damit ergibt sich $\varepsilon = k \cdot \frac{\sigma_n}{n} = 1{,}96 \cdot \frac{15{,}27}{1000} = 0{,}02992$, gespeichert in E

und das Vertrauensintervall

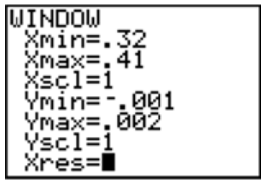
$[0{,}37 - 0{,}02992 ; 0{,}37 + 0{,}02992] = [0{,}34008 ; 0{,}39992]$

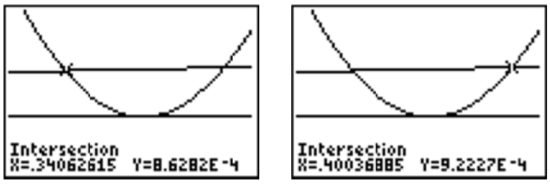
Schneller geht's mit dem Test
STAT→TESTS→A:1-propZInt…



Interpretation im Sachzusammenhang:
Die Partei SGL kann bei der Wahl mit einer Sicherheit von 95 Prozent ein Ergebnis zwischen 34,01% und 39,99% der Wählerstimmen erwarten (eine Spanne von fast 6%).

Die **exakte Lösung** verwendet $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ statt σ_n. Das Vertrauensintervall ergibt sich aus der Lösung der Ungleichung $(h_n - p)^2 \le k^2 \cdot \frac{\sigma^2}{n^2}$ bzw. $(h_n - p)^2 \le k^2 \cdot \frac{p \cdot (1-p)}{n}$ und würde das Intervall [0,34063…0,40037] ergeben (Lösung mit GTR; *p* durch *x* ersetzen).





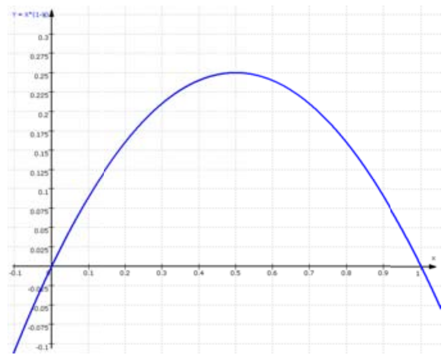
Soll der **Stichprobenumfang** für ein vorgegebenes Vertrauensintervall bestimmt werden, muss bedacht werden, dass ε die halbe Länge, den Radius des Intervalls, angibt (Formelsammlung S. 43; Konfidenzintervalle). Dieser Term wird nach n aufgelöst:

$$k \cdot \frac{\sqrt{n \cdot h_n \cdot (1-h_n)}}{n} = \varepsilon \quad | \, ()^2$$

$$k^2 \cdot \frac{n \cdot h_n \cdot (1-h_n)}{n^2} = \varepsilon^2 \quad | \, n \; kürzen$$

$$k^2 \cdot \frac{h_n \cdot (1-h_n)}{n} = \varepsilon^2 \quad | \cdot n, : \varepsilon^2$$

$$k^2 \cdot \frac{h_n \cdot (1-h_n)}{\varepsilon^2} = n$$



Da $h_n \cdot (1-h_n) \; stets \; \le \frac{1}{4}$ ist (der Graph ist eine nach unten geöffnete Parabel mit S(0,5|0,25), s. Abb. rechts), kann der Zähler des Bruches durch $\frac{1}{4}$ abgeschätzt werden:

$k^2 \cdot \frac{\frac{1}{4}}{\varepsilon^2} = n \Leftrightarrow \left(\frac{k}{2\varepsilon}\right)^2 = n$. Für alle größeren n ist der Radius des Intervalls geringer als ε.

Wenn man einen Näherungswert für h_n hat, erreicht man damit statt mit $\frac{1}{4}$ eine genauere Abschätzung:

$$n > k^2 \cdot \frac{h_n \cdot (1-h_n)}{\varepsilon^2} \Leftrightarrow n > \left(\frac{k}{\varepsilon}\right)^2 \cdot h_n \cdot (1-h_n)$$

Beispiel: Zum Vertrauensniveau $\gamma = 0{,}95$ und einer gewünschten Länge 0,04 (4%) des Vertrauensintervalls soll die Anzahl der zu befragenden Personen bestimmt werden:

$$k = 1{,}96,\; \varepsilon = \frac{0{,}04}{2} = 0{,}02 \;\Rightarrow\; n > \left(\frac{k}{2\varepsilon}\right)^2 = \left(\frac{1{,}96}{0{,}04}\right)^2 = 2401$$

Es müssen mindestens 2402 Personen befragt werden, um das Wahlergebnis mit einer Genauigkeit von 4% vorherzusagen (bei einer Sicherheit von 95%).
Falls man z.B. bereits weiß, dass h_n deutlich unter 15% liegt, rechnet man mit $0{,}15 \cdot (1 - 0{,}15)$.

$k = 1{,}96\,,\ \varepsilon = \dfrac{0{,}04}{2} = 0{,}02 \ \Rightarrow\ n > \left(\dfrac{k}{\varepsilon}\right)^2 \cdot 0{,}15 \cdot (1 - 0{,}15) = \left(\dfrac{1{,}96}{0{,}02}\right)^2 \cdot 0{,}15 \cdot 0{,}85 = 1224{,}51$

Man braucht also dann nur noch 1225 Personen zu befragen.

Zusammenfassung:

- Wenn die Länge des Vertrauensintervalls bei unbekannter relativer Häufigkeit höchstens $l = 2\varepsilon$ betragen soll, wählt man den Stichprobenumfang
  $n > \left(\frac{k}{l}\right)^2$
  bzw. $n > \left(\frac{k}{2\varepsilon}\right)^2$.
- Wenn h_n bekannt ist (z.B. aus einer kleinen Probebefragung), wählt man
  $n > \left(\dfrac{k}{\varepsilon}\right)^2 \cdot h_n \cdot (1 - h_n)$

Zusammenstellung: Wolfgang Hanisch www.w-hanisch.de/mathe,
Peter Bollmann,
Ralf Oldenburg www.oldenburg-ralf.de