Prognoseintervalle - absolut und relativ

Auf dem Weg zu Konfidenzintervallen

Reimund Vehling

Ricarda-Huch-Schule Hannover, eA-Ma3L

05.05.2020

Binomialverteilungen - eine kurze Wiederholung

Ein Standardbeispiel für die Modellierung mit einer Binomialverteilung ist das n-fache, unabhängige Drehen eines Glücksrades und Zählen der Treffer (Trefferwahrscheinlichkeit p). Wenn die Zufallsgröße X die Anzahl der Treffer zählt, dann gilt für die Wahrscheinlichkeit, genau k Treffer bei n Drehungen zu erzielen:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

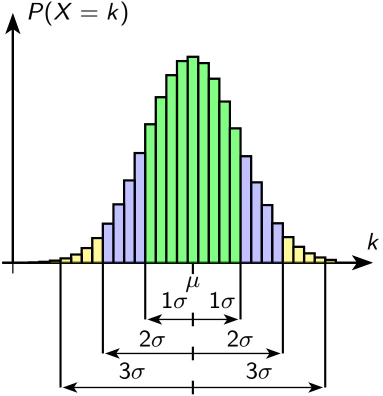
- $\binom{n}{k}$ gibt die Anzahl der Pfade in dem zugehörigen n-stufigen Baumdiagramm an, genau k Treffer und $n-k$ Nieten zu erzielen.
- Der Term $p^k \cdot (1-p)^{n-k}$ liefert die Wahrscheinlichkeit (1. Pfadregel) längs eines Pfades (k Treffer, $n-k$ Nieten).

Bei einer Binomialverteilung gibt es für den Erwartungswert μ und die Standardabweichung σ einfache Terme, ohne immer auf die Definition der beiden Kenngrößen zurückgehen zu müssen

$$\mu = n \cdot p; \quad \sigma = \sqrt{n \cdot p \cdot (1-p)}$$

Sigma-Umgebungen bei Binomialverteilungen

Umgebungen vom Erwartungswert μ, die als Vielfache der Standardabweichung σ angegeben werden, führen zu Sigma-Regeln. Die Wahrscheinlichkeiten sind fast unabhängig von n und p. Diese Regel liefert gute Näherungswerte, falls $\sigma > 3$ gilt. Die Näherungswerte werden umso besser, je größer σ ist.



k	$P(\mu - k \cdot \sigma \leq X \leq \mu + k \cdot \sigma)$
1,000...	68,3%
1,644...	90%
1,959...	95%
2,575...	99%
2,004...	95,5%
2,967...	99,7%

$$\left[\mu - k \cdot \sqrt{n \cdot p \cdot (1-p)};\ \mu + k \cdot \sqrt{n \cdot p \cdot (1-p)}\right]$$

Prognoseintervalle - absolut

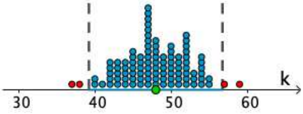
$k \cdot \sigma$-Umgebungen um μ

$k \cdot \sigma$-Umgebungen um $\mu = n \cdot p$ werden auch **Prognoseintervalle** genannt.

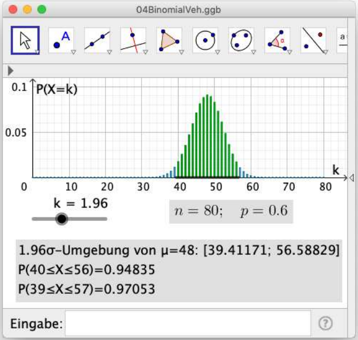
95%-Prognoseintervall für absolute Häufigkeiten H (1,96σ-Umgebung um μ):

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

Bsp.:
Ein Glücksrad wird 80-mal gedreht. Die Trefferwahrscheinlichkeit beträgt $p = 0{,}6$. Mit welchen Trefferanzahlen ist in 95 % aller Fälle zu rechnen, wenn ich den Zufallsversuch sehr oft durchführe?



100-mal eine Simulation durchgeführt und die Trefferanzahlen als Punkte dargestellt.



Prognoseintervalle verstehen mit Simulationen

Prognoseintervalle

Die zentrale Aussage von Umgebungswahrscheinlichkeiten der Form

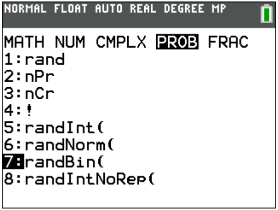
$$P(\mu - 2 \cdot \sigma \leq X \leq \mu - 2 \cdot \sigma) \approx 0{,}954$$

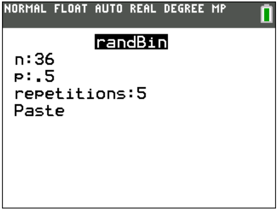
und den zugehörigen 95,4 % Prognoseintervallen $[\,\mu - 2 \cdot \sigma;\ \mu + 2 \cdot \sigma\,]$ kann man besser verstehen, wenn man Simulationen durchführt.

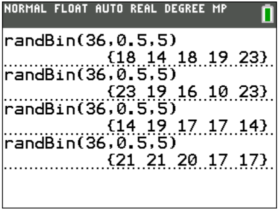
Beispiel: geg.: $X \sim \mathrm{Bin}(36; 0{,}5)$

ges.: 95,4 %-Prognoseintervall für die Trefferanzahlen

Simulationen mit $\texttt{randBin}(\texttt{n}, \texttt{p}, <\texttt{Wiederholungen}>)$







Prognoseintervalle verstehen mit Simulationen

Beispiel

geg.: $X \sim \mathrm{Bin}(36; 0{,}5)$

ges.: 95,4 %-Prognoseintervall für die Trefferanzahlen
Simulationen mit $\texttt{randBin}(\texttt{n}, \texttt{p}, <\texttt{Wiederholungen}>)$

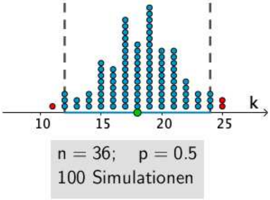
Lösung:

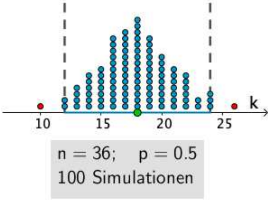
$\mu = 36 \cdot 0{,}5 = 18$; $\sigma = \sqrt{36 \cdot 0{,}5 \cdot 0{,}5} = 3$

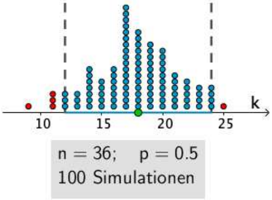
95,4 %-Prognoseintervall: $[\,18 - 2 \cdot 3;\ 18 + 2 \cdot 3\,] = [\,12;\ 24\,]$

$P(12 \leq X \leq 24) \approx 0{,}954$; exakt: $P(12 \leq X \leq 24) = 0{,}9711\ldots$

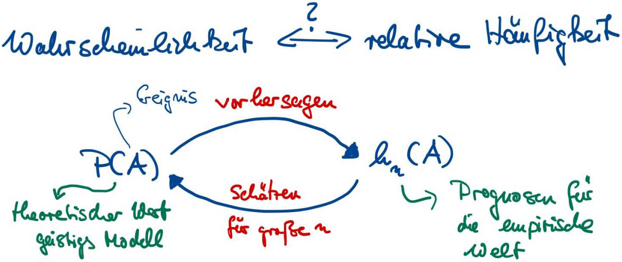
Drei Simulationen mit GeoGebra:







Einschub: Wahrscheinlichkeiten und relative Häufigkeiten



Wahrscheinlichkeiten sagen relative Häufigkeiten auf lange Sicht voraus. Sie sind gut gewählt, wenn die (durch Zufallseinflüsse schwankenden) relativen Häufigkeiten um die Wahrscheinlichkeit pendeln, also etwa gleich oft über wie unter der Wahrscheinlichkeit liegen.

Die Frage danach, was Wahrscheinlichkeiten sind, können wir nicht beantworten. Wir können aber Antworten (hoffentlich Sie auch) finden auf folgende Fragen:
Was kann ich mir darunter vorstellen? Was leisten sie? Wie erhalte ich sie? Wie kann ich damit rechnen?
Den Weg von einer relativen Häufigkeit zu einer Wahrscheinlichkeit (Schätzproblem) werden wir bald gehen. Dann kommen (endlich) die Konfidenzintervalle ins Spiel.

Prognoseintervalle - relativ

Prognoseintervalle für relative Häufigkeiten

Ein **95 %-Prognoseintervall** für relative Häufigkeiten $h = H/n$:

$$\left[p - 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}};\ p + 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \right] \tag{2}$$

Beispiel:
Wenn das Glücksrad mit der Trefferwahrscheinlichkeit von $p = 0{,}6$ nun 80-mal gedreht wird und genau 50-mal ein Treffer anzeigt wird, hat sich in 50 von 80 Fällen ein Treffer eingestellt, also in 62,5 % aller Fälle. Dividiert man also die absolute Trefferanzahl durch n - hier 80 - erhält man den relativen Anteil an Treffern: $\frac{50}{80} = 0{,}625 = 62{,}5\,\%$.

Wenn wir $p = 0{,}6$ nicht kennen würden, könnten wir im ersten Zugriff $h = 0{,}625$ als Schätzwert für die Trefferwahrscheinlichkeit p nehmen. Führen wir diesen Versuch sehr oft durch, erhalten wird ein Prognoseintervall für p. Wir können also eine Prognose (Schätzwerte) für p abgeben.

Der Weg von absolut zu relativ

Memo 1: $\mu = n \cdot p$; $\sigma = \sqrt{np(1-p)}$

Memo 2: Für $n > 0$ gilt: $n = \sqrt{n^2}$; $\quad \frac{1}{n} = \sqrt{\frac{1}{n^2}}$

$$[\,\mu - 1{,}96 \cdot \sigma;\; \mu + 1{,}96 \cdot \sigma\,]$$

$$\left[\, n \cdot p - 1{,}96 \cdot \sqrt{n \cdot p \cdot (1-p)};\; n \cdot p + 1{,}96 \cdot \sqrt{n \cdot p \cdot (1-p)} \,\right] \quad \Big| : n \text{ oder } \cdot \frac{1}{n}$$

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

$$\left[\, p - 1{,}96 \cdot \sqrt{\frac{n \cdot p \cdot (1-p)}{n^2}};\; p + 1{,}96 \cdot \sqrt{\frac{n \cdot p \cdot (1-p)}{n^2}} \,\right]$$

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

Aufgaben

Aufgabe 1

In einer Urne befinden sich 400 grüne und 600 rote Kugeln. Es werden 80 Kugeln mit Zurücklegen gezogen. Ein Treffer wird genau dann erzielt, wenn eine grüne Kugel gezogen wird. Die Zufallsgröße X, die die Anzahl der grünen Kugeln in der 80-er Stichprobe zählt, kann als binomialverteilt angenommen werden.
Ermitteln Sie ein um p symmetrisches Intervall, in dem ca. 95 % der relativen Häufigkeiten h liegen.

Lösung

Aufgabe 1

In einer Urne befinden sich 400 grüne und 600 rote Kugeln. Es werden 80 Kugeln mit Zurücklegen gezogen. Ein Treffer wird genau dann erzielt, wenn eine grüne Kugel gezogen wird. Die Zufallsgröße X, die die Anzahl der grünen Kugeln in der 80-er Stichprobe zählt, kann als binomialverteilt angenommen werden.
Ermitteln Sie ein um p symmetrisches Intervall, in dem ca. 95 % der relativen Häufigkeiten h liegen.

Das gesuchte Intervall ist das 95 %-Prognoseintervall um p. Es gilt

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

Mit $p = \frac{400}{1000} = 0{,}4$ und $n = 80$ (Stichprobenumfang) ergibt sich das um p symmetrische Intervall, in das ca. 95 % aller relativen Häufigkeiten liegen: $[0{,}2926\ldots; 0{,}5073\ldots]$.

Interpretationen der Lösung

Das gesuchte Intervall ist das 95 %-Prognoseintervall um p. Es gilt

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

Mit $p = \frac{400}{1000} = 0{,}4$ und $n = 80$ (Stichprobenumfang) ergibt sich das um p symmetrische Intervall, in das ca. 95 % aller relativen Häufigkeiten liegen: $[0{,}2926\ldots; 0{,}5073\ldots]$.

Mögliche Interpretationen:

Interpretationen der Lösung

Das gesuchte Intervall ist das 95 %-Prognoseintervall um p. Es gilt

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

Mit $p = \frac{400}{1000} = 0{,}4$ und $n = 80$ (Stichprobenumfang) ergibt sich das um p symmetrische Intervall, in das ca. 95 % aller relativen Häufigkeiten liegen: $[0{,}2926\ldots; 0{,}5073\ldots]$.

Mögliche Interpretationen:

- Wenn ich das Zufallsexperiment oft durchführe, erwarte ich in ca. 95 % aller Fälle (meine Prognose), mindestens 0,2926... und höchstens 0,5073... als relative Häufigkeit für den relativen Anteil an grünen Kugeln in der 80-er Stichprobe zu erhalten.
- Wenn ich das Zufallsexperiment 100-mal durchführe, erwarte ich in ca. 5 Fällen, dass ich kleinere h-Werte als 0,2926... oder größere h-Werte also 0,5073... erhalte.
- In dem um p symmetrischen Intervall liegen ca. 95 % aller Stichprobenergebnisse, wenn ich das Zufallsexperiment oft durchführe.

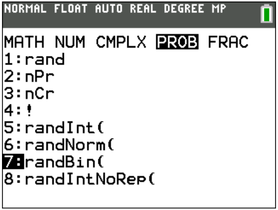
Simulationen mit dem TI-84+ mithilfe des randBin-Befehls

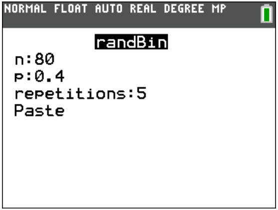
Aufgabe 1

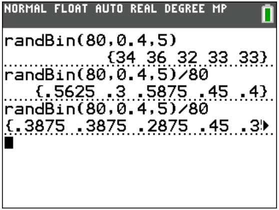
In einer Urne befinden sich 400 grüne und 600 rote Kugeln. Es werden 80 Kugeln mit Zurücklegen gezogen. Ein Treffer wird genau dann erzielt, wenn eine grüne Kugel gezogen wird. Die Zufallsgröße X, die die Anzahl der grünen Kugeln in der 80-er Stichprobe zählt, kann als binomialverteilt angenommen werden.

Ermitteln Sie ein um p symmetrisches Intervall, in dem ca. 95 % der relativen Häufigkeiten h liegen.

Mit $p = \frac{400}{1000} = 0{,}4$ und $n = 80$ (Stichprobenumfang) ergibt sich das um p symmetrische Intervall, in das ca. 95 % aller relativen Häufigkeiten liegen: $[0{,}2926...; 0{,}5073...]$.







Aufgabe 2

Gegeben ist die Wahrscheinlichkeit $p = 0{,}4$.
Ermitteln Sie für die Stichprobenumfänge 50, 100, 200 und 5000 jeweils das 95 %-Prognoseintervall für p.

Lösung:

n	95 % -Prognoseintervall	Länge
50		
100		
200		
5000		

Aufgaben

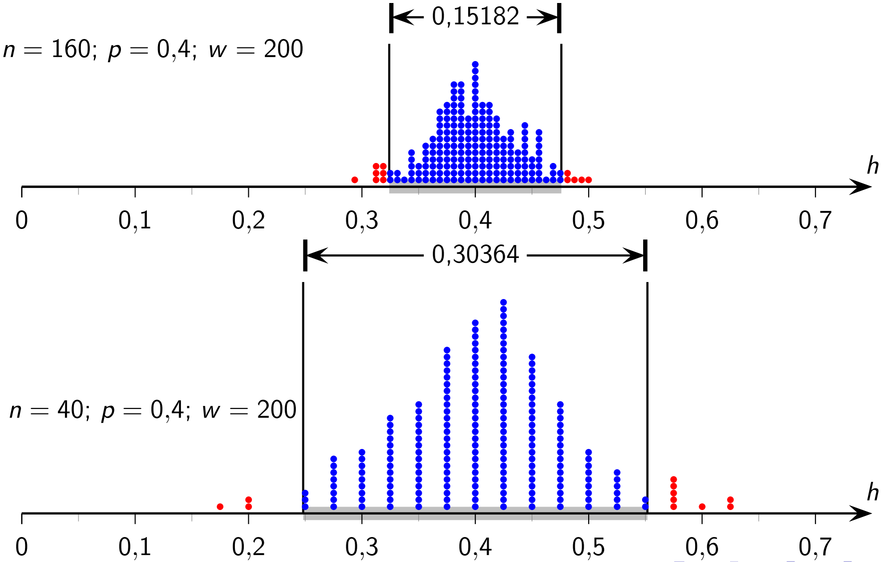
Aufgabe 2

Gegeben ist die Wahrscheinlichkeit $p = 0{,}4$.
Ermitteln Sie für die Stichprobenumfänge 50, 100, 200 und 5000 jeweils das 95 %-Prognoseintervall für p.

Lösung:

n	95 % -Prognoseintervall	Länge
50	[0,2642...; 0,5357...]	0,2715...
100	[0,3039...; 0,4960...]	0,1920...
200	[0,3321...; 0,4678...]	0,1357...
5000	[0,3864...; 0,4135...]	0,0271...

Mit Simulationen zum $1/\sqrt{n}$ -Gesetz



Das $1/\sqrt{n}$ -Gesetz

Prognoseintervalle werden mit wachsendem n immer kleiner, p wird sozusagen immer enger eingeschnürt. Das schauen wir uns nun genauer an:
95 %-Prognoseintervall:

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

Die Länge dieses Intervalls ergibt sich durch Subtraktion der beiden Grenzen:

$$2 \cdot 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}}$$

$$2 \cdot 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} = \frac{1}{\sqrt{n}} \cdot \underbrace{2 \cdot 1{,}96 \cdot \sqrt{p \cdot (1-p)}}_{Konstante} = \frac{1}{\sqrt{n}} \cdot Konstante.$$

Man spricht deshalb vom $1/\sqrt{n}$-Gesetz.
Die Stichprobengröße vervierfachen bringt (nur) eine Halbierung der Intervalllänge.

Machen Sie sich diesen Sachverhalt an Beispielen und danach mit $n_1 = n$ und $n_2 = 4n_1$ klar!

Aufgabe 3

Von einem (relativen) 95 %-Prognoseintervall ist das Folgende bekannt:
Die linke Intervallgrenze hat den Wert 0,65 und für den Stichprobenumfang n gilt: $n = 400$.
Ermitteln Sie das zugehörige p und die rechte Intervallgrenze.

Lösung:

Der Ansatz $p - 1{,}96 \cdot \sqrt{\frac{p\cdot(1-p)}{400}} = 0{,}65$ führt auf die Lösung 0,6951... Hierbei wurde die Gleichung grafisch-numerisch durch Schnittpunktbestimmung gelöst.
Somit gilt $p = 0{,}6951...$. Für die rechte Intervallgrenze ergibt sich mit $p + 1{,}96 \cdot \sqrt{\frac{p\cdot(1-p)}{n}}$ der Wert 0,7402...

Aufgaben

Aufgabe 4 - ohne Lösung

Von einem 95 %-Prognoseintervall ist das Folgende bekannt:
Die rechte Intervallgrenze hat den Wert 0,75 und für den Stichprobenumfang n gilt: $n = 200$.
Ermitteln Sie das zugehörige p und die linke Intervallgrenze.

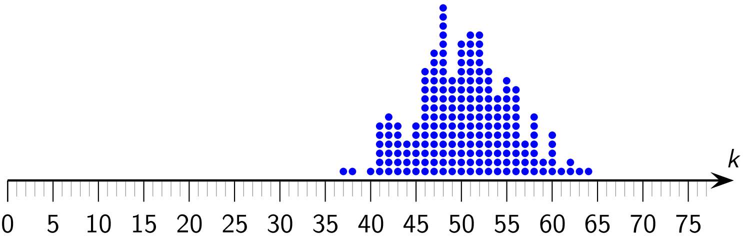
Das Konzept des Zweifelns und Prognoseintervalle

Ist die Münze gezinkt?

Eine Münze wird 100-mal geworfen. Dabei fällt 38-mal Kopf. Besteht ein Grund, an der Annahme $p_{\text{Kopf}} = 0{,}5$ zu zweifeln?

Eine Simulation kann helfen

Die folgende Abbildung zeigt das Ergebnis einer Simulation eines 100-fachen Münzwurfs mit $p_{\text{Kopf}} = 0{,}5$. Jeder Punkt symbolisiert das Ergebnis eines 100-fachen Münzwurfs. Insgesamt wurde das Experiment 200-mal wiederholt. Es werden also 200 Punkte dargestellt. Man erkennt z.B.: 6-mal wurde bei 100 Würfen 45-mal Kopf geworfen.



Das Konzept des Zweifelns und Prognoseintervalle

Vereinbarung

Wir zweifeln an der Annahme $p_{\text{Kopf}} = 0{,}5$, falls die Trefferanzahl von 38 (absolute Häufigkeit) nicht mehr in der 2σ-Umgebung des Erwartungswertes $\mu = 50$ liegt.

Unter der Voraussetzung, dass $p_{\text{Kopf}} = 0{,}5$ gilt, passiert in ca. 5 % aller Fälle ein Ereignis, das mehr als 2σ vom Erwartungswert entfernt ist. Man macht also (nur) einen Fehler von 5 %, wenn man sich irrt, also fälschlicherweise annimmt, die Münze sei gezinkt, obwohl sie es nicht ist.
Mehr ist nicht drin. Wir können nur mit Sicherheitswahrscheinlichkeiten argumentieren. Wenn die Stichprobe vom Umfang 100 eine Trefferanzahl in der 2σ-Umgebung ergeben hätte, kann man nur sagen, dass aufgrund dieses Stichprobenergebnisses auf der Basis der Sicherheitswahrscheinlichkeit von ca. 95 % kein Grund daran besteht, an $p_{\text{Kopf}} = 0{,}5$ zu zweifeln.
Ob für die Münze tatsächlich $p_{\text{Kopf}} = 0{,}5$ gilt, können wir niemals beweisen. Wahrscheinlichkeiten sind vom Menschen gesetzte Modelle, die in der Realität nie ganz richtig sind, nur „besser oder schlechter“. Wir können aber mit Testgrößen (hier 2σ-Umgebungen) in der Modellwelt arbeiten, mit denen wir die Modellgüte beurteilen. Zufallsschwankungen müssen wir hinnehmen. Sie sind nichts Negatives!

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	
0,72	
0,74	[0,6438...; 0,8361...]
0,76	
0,78	[0,6664...; 0,8535...]

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	[0,4307...; 0,6492...]
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	
0,72	
0,74	[0,6438...; 0,8361...]
0,76	
0,78	[0,6664...; 0,8535...]

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	[0,4307...; 0,6492...]
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	[0,5136...; 0,7263...]
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	
0,72	
0,74	[0,6438...; 0,8361...]
0,76	
0,78	[0,6664...; 0,8535...]

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	[0,4307...; 0,6492...]
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	[0,5136...; 0,7263...]
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	[0,5995...; 0,8004...]
0,72	
0,74	[0,6438...; 0,8361...]
0,76	
0,78	[0,6664...; 0,8535...]

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	[0,4307...; 0,6492...]
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	[0,5136...; 0,7263...]
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	[0,5995...; 0,8004...]
0,72	[0,6216...; 0,8183...]
0,74	[0,6438...; 0,8361...]
0,76	
0,78	[0,6664...; 0,8535...]

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	[0,4307...; 0,6492...]
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	[0,5136...; 0,7263...]
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	[0,5995...; 0,8004...]
0,72	[0,6216...; 0,8183...]
0,74	[0,6438...; 0,8361...]
0,76	[0,6664...; 0,8535...]
0,78	[0,6664...; 0,8535...]

Aufgaben

Prognoseintervalle für verschiedene Werte von p berechnen

In den beiden Tabellen sind für $n = 80$ für verschiedene Werte von p die zugehörigen 95 %-Prognoseintervalle dargestellt.
Ermitteln Sie die fehlenden Intervalle.
Eine Stichprobe vom Umfang $n = 80$ liefert das Stichprobenergebnis $h = 0{,}65$.
Markieren Sie die Intervalle, in denen h liegt.

p	95 % -Prognoseintervall
0,52	[0,4105...; 0,6294...]
0,54	[0,4307...; 0,6492...]
0,56	[0,4512...; 0,6687...]
0,58	[0,4718...; 0,6881...]
0,60	[0,4926...; 0,7073...]
0,62	[0,5136...; 0,7263...]
0,64	[0,5348...; 0,7451...]

p	95 % -Prognoseintervall
0,66	[0,5561...; 0,7638...]
0,68	[0,5777...; 0,7822...]
0,70	[0,5995...; 0,8004...]
0,72	[0,6216...; 0,8183...]
0,74	[0,6438...; 0,8361...]
0,76	[0,6664...; 0,8535...]
0,78	[0,6664...; 0,8535...]

Ein Weg zu Konfidenzintervallen

Bürgermeisterwahl

Einen Monat vor der Bürgermeisterwahl will die Kandidatin Frau Schmidt ihre Chancen schätzen.
Frau Schmidt hofft auf die absolute Mehrheit. Bei einer Befragung von 260 Wählerinnen und Wählern sprachen sich 104 für Frau Schmidt als Bürgermeisterin aus.
Nehmen Sie **Stellung** zu den folgenden Aussagen. Sie können Ihre Aussagen auch mit Wahrscheinlichkeiten und Prognoseintervallen unterstützen.

(1)

Frau Schmidt bekommt bei der Wahl 40 % der Stimmen.

(2)

Es ist sehr wahrscheinlich, dass Frau Schmidt zwischen 35 % und 45 % erreichen wird.

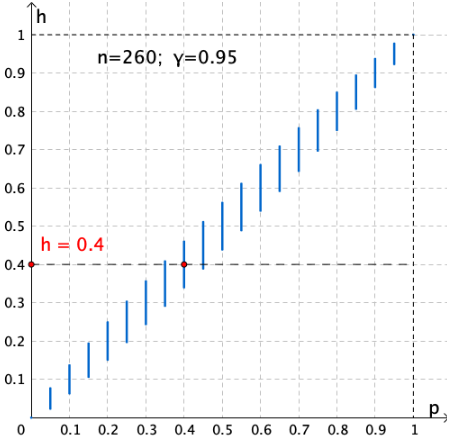
(3)

Die Wahl könnte auch mit 50 % der Stimmen für Frau Schmidt ausgehen.

Ein Weg zu Konfidenzintervallen

Bürgermeisterwahl

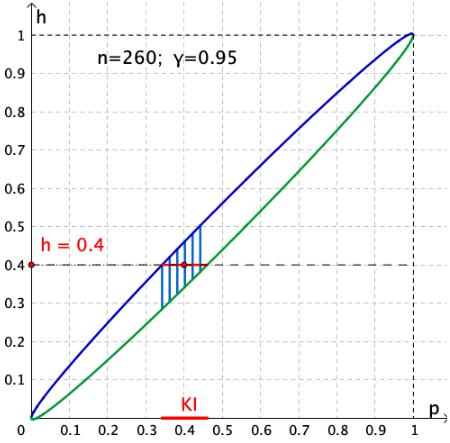
Frau Schmidt hofft auf die absolute Mehrheit. Bei einer Befragung von 260 Wählerinnen und Wählern sprachen sich 104 (40 %) für Frau Schmidt als Bürgermeisterin aus.



Ein Weg zu Konfidenzintervallen

Bürgermeisterwahl

Frau Schmidt hofft auf die absolute Mehrheit. Bei einer Befragung von 260 Wählerinnen und Wählern sprachen sich 104 (40 %) für Frau Schmidt als Bürgermeisterin aus.

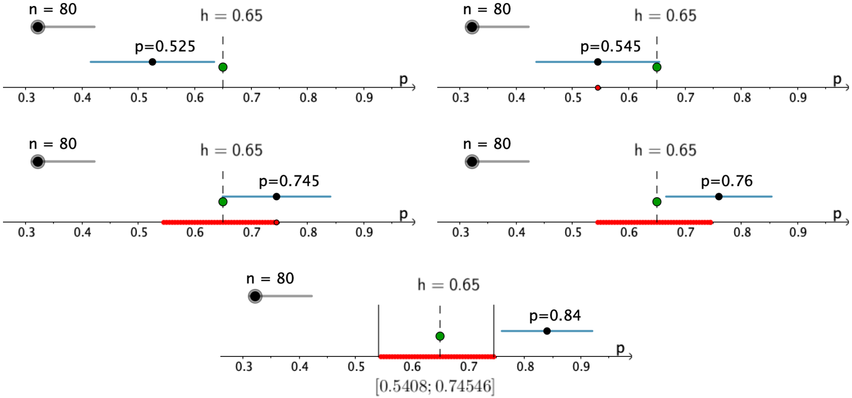


KI = [0.342 ; 0.461]

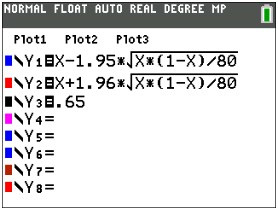
Konfidenzintervalle - nun mit $n = 80$ und $h = 0{,}65$

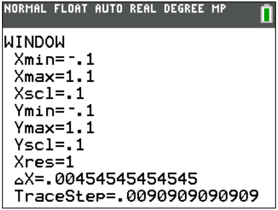
Was ist ein 95 %-Konfidenzintervall?

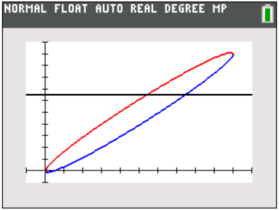
Alle Werte für p, in deren 95 %-Prognoseintervall das Stichprobenergebnis h liegt, sind „statistisch verträglich“ mit h.
Diese Werte bilden das **95 %-Konfidenzintervall**.

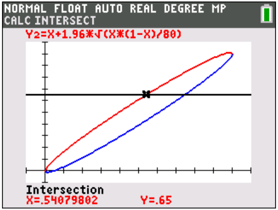


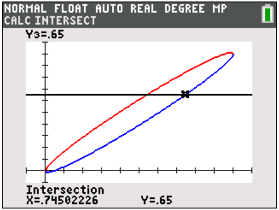
Berechnungen von Konfidenzintervallen mit dem TI-84+





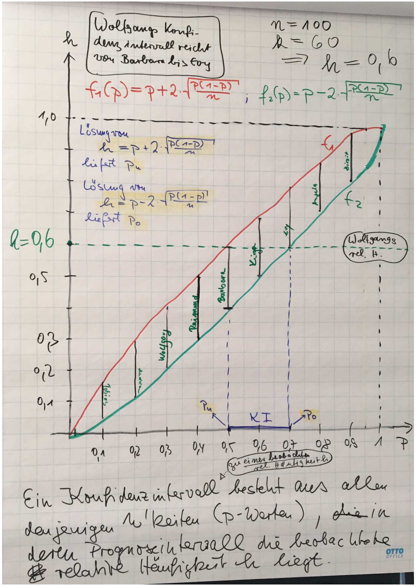






$$h = p + 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \Rightarrow p_u \, ; \quad h = p - 1{,}96 \cdot \sqrt{\frac{p \cdot (1-p)}{n}} \Rightarrow p_o$$

Prognoseintervalle liefern ein Konfidenzintervall

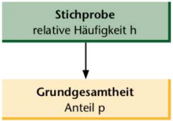


Konfidenzintervalle

Konfidenzintervalle für Anteile

Situation
Der Anteil p einer Merkmalsausprägung in der Grundgesamtheit ist unbekannt.
In einer Stichprobe mit dem Umfang n tritt die Merkmalsausprägung mit der relativen Häufigkeit h auf. Mit dem Ergebnis der Stichprobe möchte man auf die Gesamtheit schließen.

Schließen von der Stichprobe auf die Grundgesamtheit

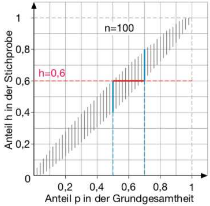


Fragen
Aus welcher Grundgesamtheit könnte das Stichprobenergebnis stammen?
Mit welchem Anteil p in der Grundgesamtheit ist das Stichprobenergebnis „statistisch verträglich"? Welche Anteile p kommen infrage?

Entscheidung
Alle Werte für p, in deren 95%-Prognoseintervall das Stichprobenergebnis h liegt, sind „statistisch verträglich" mit h. Diese Werte bilden das **95%-Konfidenzintervall**.

Bestimmung des Konfidenzintervalls
Das Diagramm bezieht sich auf eine Stichprobe mit dem Umfang 100, in der die relative Häufigkeit der Merkmalsausprägung h = 0,6 festgestellt wurde. Diese ist eingezeichnet, ebenso die 95%-Prognoseintervalle für verschiedene (angenommene) Werte von p.

Man kann ablesen: Für p ≈ 0,5 liegt h gerade noch in dem betreffenden 95%-Prognoseintervall, ebenso für p ≈ 0,7. Das **95%-Konfidenzintervall** ist etwa **[0,5; 0,7]**.



Interpretation
Mit einer Sicherheitswahrscheinlichkeit von 95% überdeckt das Konfidenzintervall [0,5; 0,7] den wahren Anteil p.

Aufgaben zum Schätzen mit Konfidenzintervallen

Aufgabe 1: Bürgermeisterwahl

Die Wahl zum Bürgermeister einer Großstadt gewinnt derjenige Kandidat, der mehr als 50 % der Stimmen erhält. Vor der Bürgermeisterwahl wurden Umfragen zum Wahlausgang durchgeführt. In einer Umfrage gaben 424 von 800 Personen an, den amtierenden Bürgermeister wählen zu wollen.

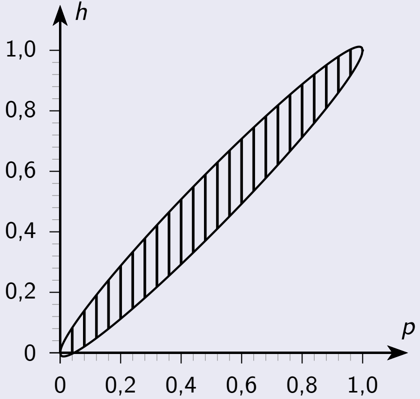
Entscheiden Sie mithilfe eines Konfidenzintervalls zur Sicherheitswahrscheinlichkeit von 95 %, ob der amtierende Bürgermeister mit einem Wahlsieg rechnen kann.

Aufgaben zum Schätzen mit Konfidenzintervallen

Aufgabe 2: Konfidenzintervalle ermitteln mithilfe der Ellipse

Die Abbildung zeigt für $n = 80$ einige 95 %-Prognoseintervalle und die zugehörige Ellipse.

Bestimmen Sie mithilfe der Abbildung für die Stichprobenergebnisse $h_1 = 0{,}8$, $h_2 = 0{,}5$ und $h_3 = 0{,}2$ die zugehörigen 95 %-Konfidenzintervalle.



Aufgaben zum Schätzen mit Konfidenzintervallen

Aufgabe 3: Frühstück an Morgen?

Von 152 Schülern einer Schule gaben 66 an, morgens zu Hause nicht mehr zu frühstücken.
Bestimmen Sie das 95 %-Vertrauensintervall für den unbekannten Anteil an Schülern der Schule, die morgens kein Frühstück zu sich nehmen.

Aufgabe 4: Lohnt sich der CD-Verkauf?

Bei einem Rockkonzert werden 250 Besuchern zufällig ausgewählt und befragt, ob sie am Ausgang eine CD der Gruppe kaufen werden. 27 der Befragten antworteten mit ja. Insgesamt waren 9543 Besucher bei dem Rockkonzert.
Schätzen Sie bei einer Sicherheitswahrscheinlichkeit von 95 %, mit welchem Minimal – und Maximalumsatz an CDs zu rechnen ist.

Aufg. 5: Interpretation der Ergebnisse bei vorgegebener Gesamtanzahl

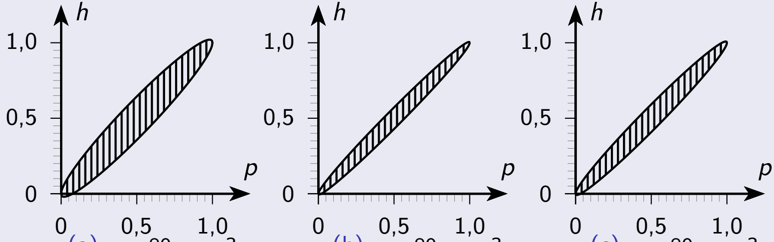
Für eine neue Attraktion im Zoo benötigt die Stadt von den Bürgern zusätzlich noch eine Spende von 250 000 €. Um die Spendenbereitschaft zu testen, wird eine Umfrage unter 800 Haushalten durchgeführt. Es wurde danach gefragt, ob ein Haushalt 20 € spenden würde. Die Auszählung ergab, dass 102 Haushalte von 800 Haushalten den Betrag von 20 € spenden würden. Die Stadt hat insgesamt 98 750 Haushalte.

Untersuchen Sie, ob man mit einer Sicherheit von 95 % davon ausgehen kann, den Betrag von 250 000 € zu erhalten?

Aufgabe 6: Zuordnung finden, γ verschieden

Die drei Grafiken in den Abbildungen zeigen für $n = 80$ jeweils einige 95 %-Prognoseintervalle zur Bestimmung eines Konfidenzintervalls. Hierbei sind die Sicherheitswahrscheinlichkeiten γ verschieden: $\gamma_1 = 0{,}8$; $\gamma_2 = 0{,}9$ und $\gamma_3 = 0{,}99$.

Ordnen Sie begründet den drei verschiedenen Werten für γ jeweils die zugehörige Abbildung **zu**.



(a) $n = 80$; $\gamma = ?$ (b) $n = 80$; $\gamma = ?$ (c) $n = 80$; $\gamma = ?$

Abbildung: Zuordnung finden

Aufgaben zum Schätzen mit Konfidenzintervallen

Aufg. 7: Grenze gesucht

Bei einem Stichprobenumfang von $n = 100$ ergibt sich das folgende 95 %-Vertrauensintervall: $[0{,}420;\ a]$.

1. **Ermitteln** Sie mithilfe der Abbildung einen Näherungswert für a.
2. **Berechnen** Sie den Wert für a.
3. **Angenommen**, 0,420 ist die rechte Grenze eines 95 %-Konfidenzintervalls. **Welche** Auswirkungen hat das auf die Ermittlung der linken Grenze?