Thema: Konfidenzintervalle

Gertrud Aumayr, Christian Zöpfl

☒ TI-Nspire™ CAS

Schlagworte: Analyse von Stichproben, Konfidenzintervall, Vertrauensbereich

Unterrichtsmaterial:

Aufgabe/Arbeitsauftrag:

Beispiel 1: *Simulation einer Umfrage durch Würfeln*

Eine Umfrage bei der die Antworten *JA* und *NEIN* gegeben werden können, soll durch das Werfen eines Würfels simuliert werden. Jedes Mal, wenn die Augenzahl *sechs* gewürfelt wird, wird die Antwort *JA* angenommen, in allen andere Fällen die Antwort *NEIN*.

- ***a)*** Wir wählen eine Stichprobe von 600. Führe diese Simulation durch und berechne für dein Ergebnis der Häufigkeit in der Stichprobe eine 1,96 – Sigma Umgebung (95% - Umgebung).
- ***b)*** Verwende nun die am Rechner vorgegebe Möglichkeit zur Berechnung eines Konfidenzintervalls: menu → Statistics → Confidence Intervalls →1-Prop Z Intervall und vergleiche mit deinem vorigen Ergebnis.
- ***c)*** Stelle die in der Klasse erhaltenen Ergebnisse der Konfidenzintervalle im Heft oder beliebig viele Intervalle automatisch am Rechner dar. Vergleiche mit der in unserem Fall ausnahmsweise bekannten Wahrscheinlichkeit für die Antwort JA.
- ***d)*** Bei der Darstellung am Rechner kann das Problem kopiert werden und relativ leicht die Sicherheit von 95 % auf 90 % oder 60 % verändert werden. Interpretiere die Ergebnisse.
- ***e)*** Bei der Darstellung am Rechner kann das Problem erneut kopiert werden und relativ leicht der Stichprobenumfang verändert werden. Interpretiere die Ergebnisse.

Didaktischer Kommentar:

Zur Schätzung des unbekannten relativen Anteils p eines Merkmals in einer Grundgesamtheit wird eine Stichprobe von großem Umfang n untersucht. In dieser ergibt sich der Wert h für die relative Häufigkeit des untersuchten Merkmals. Die Menge aller Schätzwerte für p, deren zugehöriger γ-Schätzbereiche den in der Stichprobe beobachteten Wert h überdecken, heißt Konfidenzintervall mit der Sicherheit γ für den unbekannten relativen Anteil p.

D.h.: $p \in \gamma-\text{Konfidenzintervall} \Leftrightarrow p-z\cdot\sqrt{\frac{p\cdot(1-p)}{n}} \le h \le p+z\cdot\sqrt{\frac{p\cdot(1-p)}{n}}$

Näherung: Für große n ist der Wurzelausdruck sehr klein. Er ändert sich nicht sehr, wenn man den relativen Anteil p in der Grundgesamtheit durch die relative Häufigkeit in der Stichprobe ersetzt. Dadurch erhält man näherungsweise

$$p \in \gamma-\text{Konfidenzintervall} \Leftrightarrow p-z\cdot\sqrt{\frac{h\cdot(1-h)}{n}} \le h \le p+z\cdot\sqrt{\frac{h\cdot(1-h)}{n}}$$

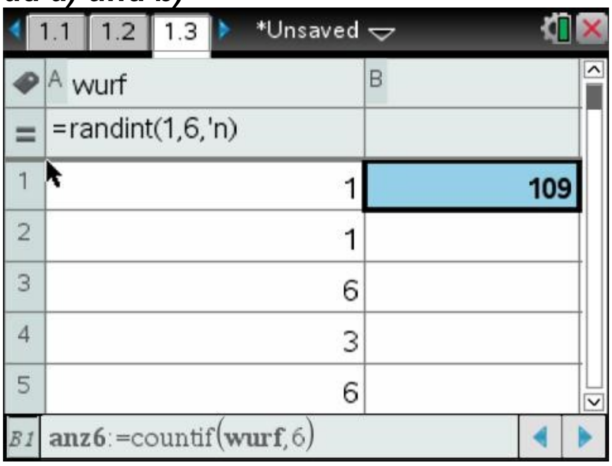
bzw. durch Umformen

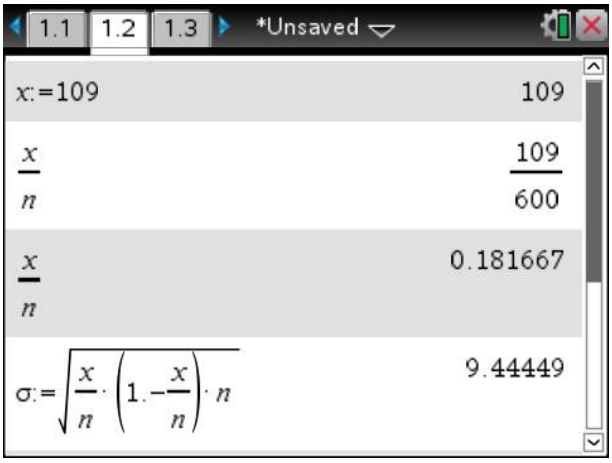
$$p \in \gamma - \text{Konfidenzintervall} \;\Leftrightarrow\; h - z \cdot \sqrt{\frac{h \cdot (1-h)}{n}} \le p \le h + z \cdot \sqrt{\frac{h \cdot (1-h)}{n}}$$

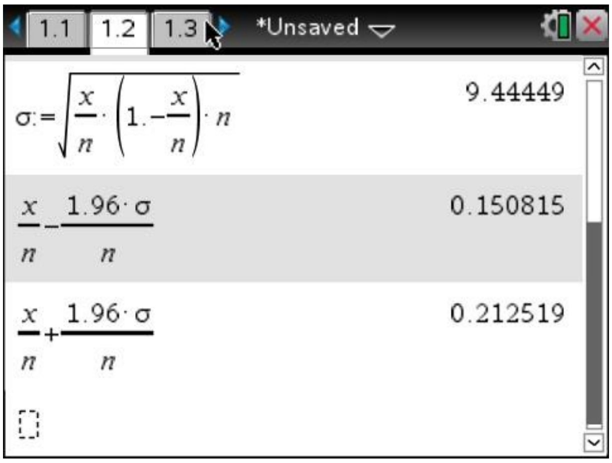
Somit ist das Intervall $\left[h - z \cdot \sqrt{\frac{h \cdot (1-h)}{n}} ; h + z \cdot \sqrt{\frac{h \cdot (1-h)}{n}} \right]$ (zumindest näherungsweise) das gesuchte γ – Konfidenzintervall. Bei der Verwendung von *1-Prop Z Intervall* wird diese Näherung verwendet.

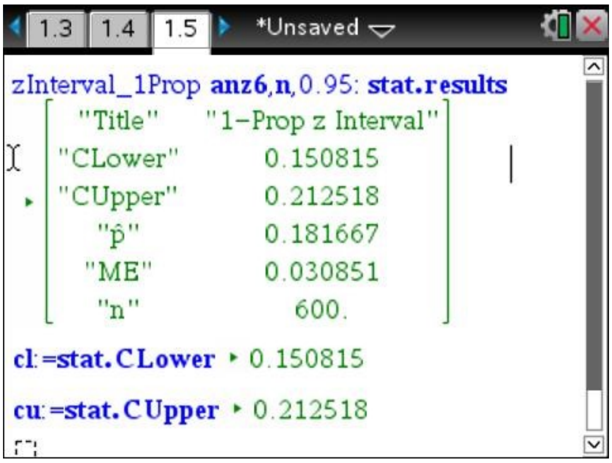
Vorschlag zur Umsetzung:

ad a) und b)



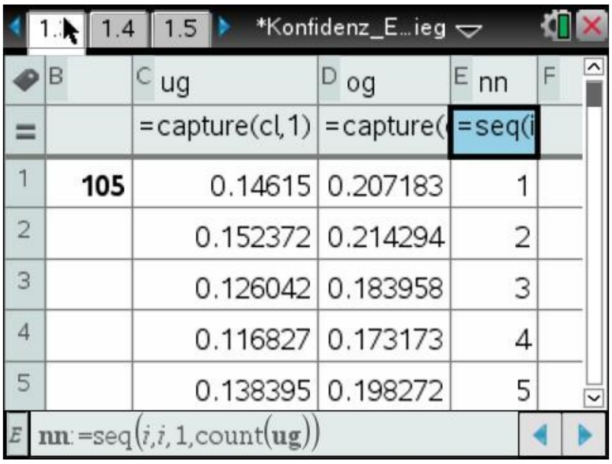






ad c) bis e)

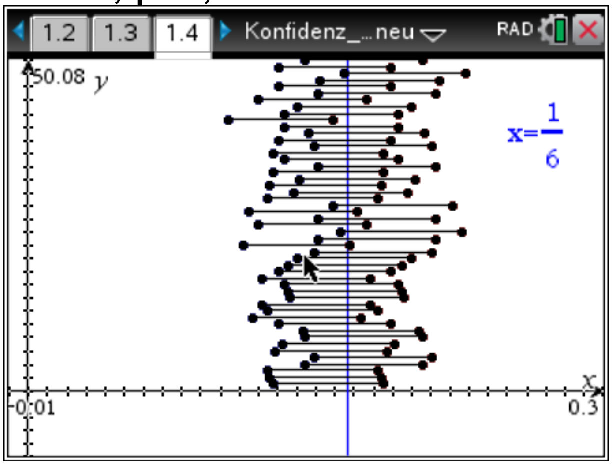
Hier wurde das Konfidenzintervall in einer Notes – Applikation berechnet und die untere bzw. obere Intervallgrenze abgespeichert in cl und cu.
Diese Daten können mit Hilfe von capture(cl,1) bzw. capture (cu,1) in der Tabellenkalkulation mitgeschrieben werden.
Bei jeder neuen Simulation, die man mit ctrl r erhält, werden die Intervallgrenzen mitgeschrieben und anschließend in einem Graphikfenster als Scatterplot dargestellt.



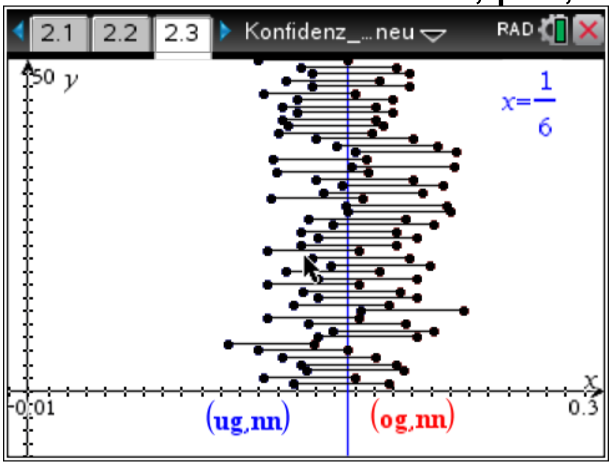
	B	C ug	D og	E nn	F
=		=capture(cl,1)	=capture(	=seq(i	
1	105	0.14615	0.207183	1	
2		0.152372	0.214294	2	
3		0.126042	0.183958	3	
4		0.116827	0.173173	4	
5		0.138395	0.198272	5	

und hier die zugehörigen Ergebnisse für…

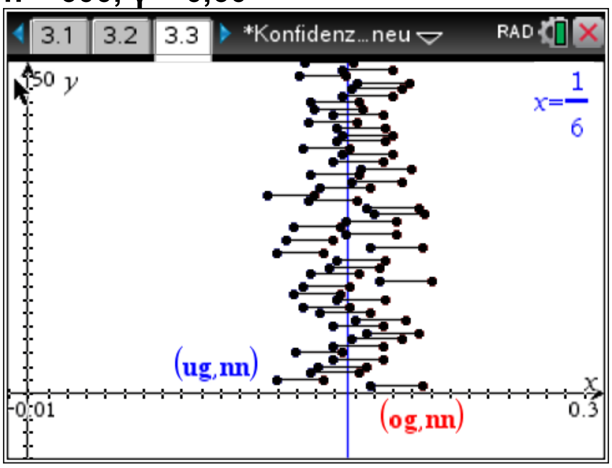
n = 600, γ = 0,95



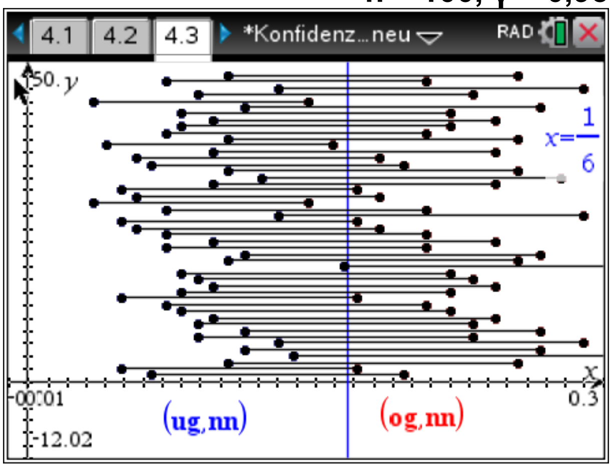
n = 600, γ = 0,90



n = 600, γ = 0,60



n = 100, γ = 0,95



Die Darstellungen zeigen die Zusammenhänge zwischen Vertrauensbereich, Stichprobenumfang und Irrtumswahrscheinlichkeit. Geringe Irrtumswahrscheinlichkeiten bedingen einen größeren Vertrauensbereich bei gleichbleibender Stichprobengröße. Umgekehrt wird zwar die Länge des Vertrauensbereichs kürzer, wenn die Irrtumswahrscheinlichkeit erhöht wird, aber dafür liegt dann eben auch der tatsächliche Wert öfter außerhalb des berechneten Vertrauensbereichs, ein Ergebnis, dass in der Praxis eher vermieden werden sollte. Zu kleine Stichprobenumfänge führen zu absurd großen Vertrauensbereichen.