STOCHASTIK

Binomialverteilung

Teil 3

Schätzungen nach Stichproben
Hochrechnung von der Stichprobe auf das Endergebnis:
Vertrauensintervalle
Abhängigkeit der Genauigkeit vom Umfang der Stichprobe

Berücksichtigung der CAS-Rechner:

CASIO ClassPad 330
TI Nspire CAS

Datei Nr. 34 013

Stand: 26. Juli 2018

Friedrich W. Buckel

INTERNETBIBLIOTHEK FÜR SCHULMATHEMATIK

www.mathe-cd.de

Inhalt

§10 Hochrechnungen und Schätzungen

10.1 Schluss von der Gesamtheit auf die Stichprobe

Liegt einer Zufallsvariablen X in einem n-stufigen Experiment eine **Binomialverteilung** zugrunde, dann kann man den **Erwartungswert** durch $E(X) = np$ berechnen. Er sagt den Mittelwert voraus, den man erwarten kann. Der Zufall sorgt jedoch dafür, dass die Ergebnisse für X um diesen Wert schwanken. Diese Schwankungen oder Abweichungen vom Erwartungswert kann man mit der **Standardabweichung** $\sigma = \sqrt{np(1-p)}$ ganz gut abschätzen. Wenn $\sigma > 3$ ist, kann man die Verteilung der Werte von X rund um den Erwartungswert mit den sogenannten **Sigma-Intervallen** vorhersagen. Beispielsweise weiß man, dass dann X mit 95,5% Wahrscheinlichkeit um höchstens 2σ vom Erwartungswert abweicht. Es gibt sogar mehrere nützliche Sigma-Intervalle.

Das Folgende steht ausführlich in 34012 Sto BinVerteilung 2.

Zunächst nochmals ausführlich das $2\sigma-\text{Intervall}$:

X liegt zu etwa 95,5% Wahrscheinlichkeit in	$E-2\sigma \le X \le E+2\sigma$.
Das kann man so schreiben:	$P(E-2\sigma \le X \le E+2\sigma) \approx 0{,}955$
Oder in Form einer Betragsungleichung:	$P(\lvert X-E \rvert \le 2\sigma) \approx 0{,}955$

Dann das $3\sigma-\text{Intervall}$:

X liegt zu etwa 99,7% Wahrscheinlichkeit in	$E-3\sigma \le X \le E+3\sigma$.
Das kann man so schreiben:	$P(E-3\sigma \le X \le E+3\sigma) \approx 0{,}997$
Oder in Form einer Betragsungleichung:	$P(\lvert X-E \rvert \le 3\sigma) \approx 0{,}997$

Für viele Anwendungen sind „glattere" Prozentzahlen günstiger. Wir wollen uns jetzt an Hand einer Anwendungsaufgabe an solche Werte herantasten und beobachten, was uns dabei begegnet.

AUFGABE 1 für CAS-Rechner (wer keinen hat, sollte dennoch weiter lesen!)

Die Wahlbeteiligung an einer Parlamentswahl betrug 75 %. Das heißt, dass drei von vier Wahlberechtigten ihre Stimme abgegeben haben. Meinungsforschungsinstitute versuchen durch Befragen von Wählern, schon vor dem offiziellen Endergebnis herauszufinden, wie die Wahl ausgegangen „sein könnte". Sie befragen also Personen nach ihrem Wahlverhalten. Das Institut „Vorwissen" möchte in Berlin 2000 Personen befragen. *Das heißt aber nun nicht, dass alle 2000 angesprochenen auch zur Wahl gegangen sind.*

Berechne den Erwartungswert der Anzahl Wähler unter diesen 2000 und berechne die zu E symmetrischen **Intervalle, die zu 70%, 80% und 90% Wahrscheinlichkeit** gehören.

1. Lösung mit TI-Nspire CAS

Erwartungswert für die Zahl X der Wähler: $E(X) = n \cdot p = 2000 \cdot 0{,}75 = 1500$

Zugehörende Standardabweichung: $\sigma = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{2000 \cdot 0{,}75 \cdot 0{,}25} \approx 19{,}36$

Ein zu E symmetrisches Intervall wird durch diese Doppelungleichung beschrieben:

$$E - x \le X \le E + x$$

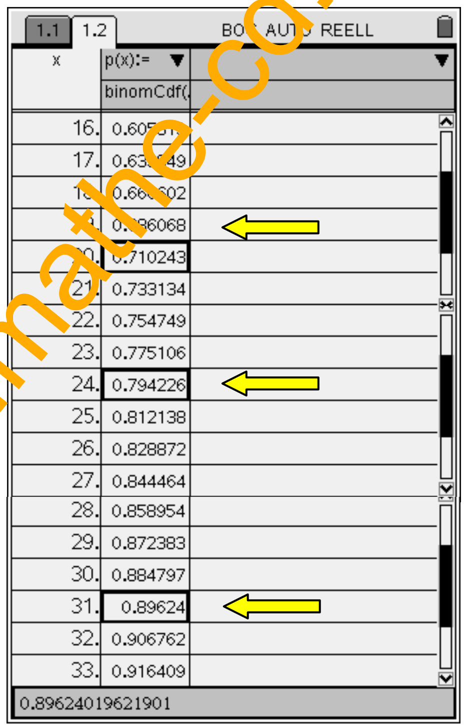
bzw. $1500 - x \le X \le 1500 + x$

Gesucht sind drei Werte für x so, dass diese drei Bedingungen erfüllt sind:

(1) $P(1500 - x \le X \le 1500 + x) \approx 0{,}7$

(2) $P(1500 - x \le X \le 1500 + x) \approx 0{,}8$

(3) $P(1500 - x \le X \le 1500 + x) \approx 0{,}9$



Ich habe mit dem CAS-Rechner **TI Nspire** diese Funktion definiert (siehe rechts):

$$p(x) = P(1500 - x \le X \le 1500 + x)$$

Mit der Verteilungsfunktion der Binomialverteilung kann man eine Wertetafel dieser Funktion erstellen.

Für x = 19, das ist ungefähr σ, bekommen wir:

$$P(1481 \le X \le 1519) \approx \boxed{70\%}$$

Für x = 24, das ist etwa $1{,}25\,\sigma$, bekommen wir

$$P(1476 \le X \le 1524) \approx \boxed{80\%}$$

Für x = 31, das sind etwa $1{,}6\,\sigma$, bekommen wir $P(1469 \le X \le 1531) \approx \boxed{90\%}$.

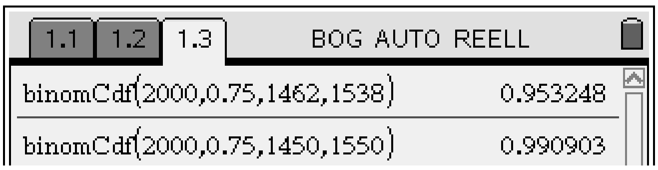
Weitere wichtige Werte sind:

Für $x = 1{,}96\,\sigma \approx 38$ erhält man 95% und für $x = 2{,}58\,\sigma \approx 50$ sogar 99% der Werte.

Für $\boxed{x = 1{,}96\,\sigma} \approx 38$ erhält man $1500 - 38 \le X \le 1500 + 38 \Leftrightarrow 1462 \le X \le 1538$. Das sind 95,3 %.

für $\boxed{x = 2{,}58\,\sigma} \approx 50$ erhält man $1500 - 50 \le X \le 1500 + 50 \Leftrightarrow 1450 \le X \le 1550$, Das sind $99{,}1\,\%$.

Kontrollieren wir das mit unserem Rechner:



Auf den nächsten Seiten folgen zwei Berechnungsmethoden dazu mit ***CASIO ClassPad****.*