

Ein Gang durch die Beurteilende Statistik mit Quantilen

Gerd Hinrichs

GTR / CAS Zu den Standard-Inhalten von Statistik-Kursen gehören Verfahren der Beurteilenden Statistik, z.B. Intervallschätzer, Konfidenzintervalle und Hypothesentests. Die GTR- und CAS- Rechner von TI bieten in diesem Bereich sehr umfangreiche Hilfsmittel. Zugleich ermöglichen sie über Interpretationen zu diskutieren und andere als die üblichen (Binomial-, Normal-) Verteilungen zu behandeln (z.B. Poissonverteilung, geometrische Verteilung, χ^2 -Verteilung). Wenn man bei der Erarbeitung der Verfahren Zeit einspart, kann man auch Maximum-Likelihood-Schätzer im Unterricht thematisieren, die ein Analysis und Stochastik verbindendes Verfahren zur Ermittlung optimaler Punktschätzer darstellen (siehe z.B. Barth/Haller [1]).

Die für den Unterricht nützlichen Funktionen der Taschenrechner werden im Folgenden aufgezeigt.

1. Was sind Quantile?

Wahrscheinlichkeitsverteilungen lassen sich eindeutig beschreiben durch die Angabe ihrer Verteilungsfunktion

$$F(x) := P(X \leq x).$$

Im TI-84Plus¹ sind z.B. die Verteilungsfunktionen

- `normalcdf(l, x, μ , σ)` für die Normalverteilung,
- `binomcdf(n, p, x)` für die Binomialverteilung,
- `poissoncdf(μ , x)` für die Poissonverteilung und
- `geometcdf(p, x)` für die geometrische Verteilung

implementiert. Dabei sind μ , σ und p die üblichen Parameter der Verteilungen. Leider benötigt die Funktion `normalcdf(l, x, μ , σ)` den zusätzlichen Parameter l , der die linke Grenze des Integrals unter der Normalverteilungsdichte angibt; hier müsste für die Verteilungsfunktion $l = -\infty$ verwendet werden, in der Praxis reicht meistens eine hinreichend „große“ negative Zahl aus, z.B. $l = -100$. Quantile entsprechen dann quasi einer Umkehrung dieser Verteilungsfunktionen: x_p ist (für unsere Zwecke) ein **p-Quantil**, wenn gilt

$$x_p = \inf \{x : P(X \leq x) \geq p\}.$$

Ist die Verteilungsfunktion umkehrbar, so gilt $x_p = F^{-1}(p)$.

Bekannt ist wahrscheinlich das 0,5-Quantil einer Verteilung, der sogenannte Median. Häufig verwendet werden auch die linken und rechten Quartile, das 0,25- bzw. das 0,75-Quantil, z.B. bei Boxplots.

2. Wie bestimmt man Quantile?

Quantile spielen in der Beurteilenden Statistik eine wichtige Rolle. Früher musste man sie in Tabellen nachschlagen oder

den Lehrbüchern glauben; der Wert 1,96 der $1,96\sigma$ -Umgebungen bei den einschlägig bekannten Näherungsverfahren ist das gerundete 0,975-Quantil der Standardnormalverteilung ($\mu=0$, $\sigma=1$)!

Zur graphisch-numerischen Bestimmung mit dem TI-84 Plus gibt man einfach die Verteilungsfunktion als eine Funktion ein und die Wahrscheinlichkeit des gesuchten p -Quantils als eine andere. Im Beispiel wird das 0,95-Quantil der Binomialverteilung mit den Parametern $n=100$ und $p=1/6$ bestimmt:

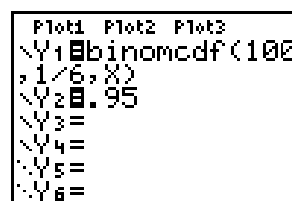


Abb. 1

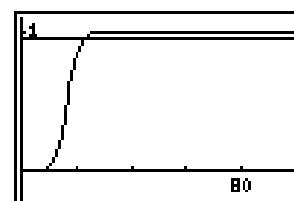


Abb. 2

Per numerischer Schnittpunktbestimmung erhält man schnell: $x_{0,95} \approx 23$. Auch wenn es eigentlich keine Schnittpunkte gibt (`binomcdf(100,1/6,23) \approx 0,962 \neq 0,95`), bestimmt der TI-84 Plus automatisch das 0,95-Quantil der Verteilung:

$$\inf \{x : \text{binomcdf}(100, 1/6, x) \geq 0,95\}.$$

3. Intervallschätzer (Schluss von der Gesamtheit auf die Stichprobe)

Eine Standard-Aufgabe lautet: „Es sei bekannt, dass in deutschen 4-Personen-Haushalten mit mittlerem Einkommen in 68,5% eine Geschirrspülmaschine vorhanden ist. Eine Stichprobe vom Umfang 720 wird durchgeführt.

In wie vielen Haushalten dieses Typs wird man ein solches Gerät vorfinden?“ (vgl. z.B. [6])

Wenn man sich auf eine Sicherheitswahrscheinlichkeit (zumeist 95%) festgelegt hat, braucht man lediglich noch die Quantile der zugehörigen Verteilungsfunktion zu bestimmen und hat die „Lösung“: Es sollen die 5% unwahrscheinlichsten Anzahlen ausgeschlossen werden, bei denen die Stichprobe signifikant zu wenig oder zu viel Familien mit Geschirrspülmaschinen enthält. Also bestimmt man wie oben beschrieben das 0,025- und das 0,975-Quantil der Binomialverteilung mit den Parametern $n=720$ und $p=0,685$; von der Eignung der Annahme einer Binomialverteilung sollten Schüler sich natürlich überzeugen und auch die eben ausgeschlossenen 5% deuten. Der TI-84 Plus liefert $x_{0,025} \approx 469$ und $x_{0,975} \approx 517$. Die Anzahl der Haushalte mit Geschirrspüler wird also mit 95%iger Wahrscheinlichkeit zwischen 469 und 517 liegen.

¹ Alle Angaben hier beziehen sich auf den TI-83/84 Plus/SE; für die CAS-Rechner benötigt man die Applikation „Statistics with List Editor“ und die Befehle haben leicht veränderte Bezeichnungen.

4. Konfidenzintervalle (Schluss von der Stichprobe auf die Gesamtheit)

In der Praxis wird man sehr viel häufiger eine Stichprobe erheben und daraus auf die Gesamtheit schließen wollen.

Aufgabe: Bei einer Umfrage unter 836 Erwachsenen gaben 423 an, für Partei „Blau“ stimmen zu wollen. Die Politiker dieser Partei interessiert nun, ob sie schon mit der absoluten Mehrheit der Stimmen rechnen können oder ob sie ihren Werbeaufwand noch verstärken sollten.

Diesmal ist der Parameter p unbekannt. Geschätzt werden soll die zugrundeliegende Wahrscheinlichkeit. Hier hilft uns, dass die Funktion `binomcdf()` fest implementiert ist. Erneut legt man die Sicherheitswahrscheinlichkeit z.B. zu 0,95 fest. Man sucht dann die Wahrscheinlichkeit $p_{0,025}$, so dass der Anteil $\frac{423}{836}$ gerade noch im entsprechenden Intervallschätzer der Binomialverteilung mit den Parametern $n=836$ und $p=p_{0,025}$ liegt.

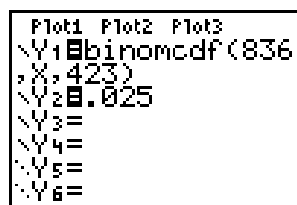


Abb. 3

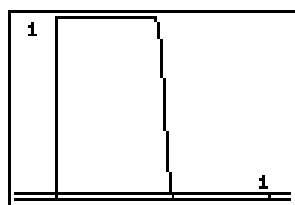


Abb. 4

Als Schnittstelle ergibt sich $p_{0,025} \approx 0,540$. Analog erhält man $p_{0,975} \approx 0,473$. Die zugrundeliegende Wahrscheinlichkeit wird also mit großer Wahrscheinlichkeit im Intervall $[0,473; 0,540]$ liegen; Partei „Blau“ kann sich der absoluten Mehrheit damit noch nicht hinreichend sicher sein.

Verständnisfördernd kann eine ausführliche Interpretation der Graphen sein. Schwächeren Schülerinnen und Schülern bereitet es enorme Schwierigkeiten, geeignete Fenstereinstellungen zu wählen; dies sollte insofern explizit thematisiert werden.

Man erkennt, dass auf diese Weise in der Beurteilenden Statistik leicht auf die Näherung durch die Normalverteilung verzichtet werden könnte. Meiner Meinung nach nimmt jene jedoch - auch historisch - eine derart wichtige Stellung ein, dass sie den Schülerinnen und Schülern nicht vorenthalten werden sollte. Verzichten kann man jedoch auf die in Lehrbüchern noch thematisierte Standardisierung von Zufallsvariablen

$$Z := \frac{X - \mu}{\sigma}$$

da die Normalverteilung mit beliebigem Erwartungswert μ und beliebiger Standardabweichung σ fest implementiert ist.

Da der TI-84 Plus nur Funktionsterme einer Variablen interpretieren kann, ist es mit obigem Verfahren nicht möglich, den notwendigen Stichprobenumfang abzuschätzen. Dazu kann man - wie auch bislang üblich - die Näherung durch die Nor-

malverteilung verwenden. Es gilt ja (unter geeigneten Bedingungen)

$$P\left(\left|\frac{X}{n} - p\right| \leq 1,96 \frac{\sigma}{n}\right) \approx 0,95.$$

Plant man eine Umfrage und möchte man die relativen Ergebnisse auf 1% genau schätzen, so geht man vom Ansatz $1,96 \frac{\sigma}{n} \leq 0,01$ aus. Durch Umformen erhält man

$$n(p) \geq \left(\frac{1,96}{0,01}\right)^2 \cdot p \cdot (1-p).$$

Die Funktion minimal erforderlicher Stichprobenumfänge kann man erneut zeichnen lassen und interpretieren.

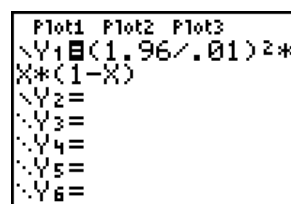


Abb. 5

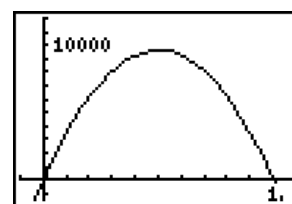


Abb. 6

Interessant ist es an dieser Stelle, realistische Erhebungen aus den Medien zu diskutieren: War der Stichprobenumfang so groß gewählt, dass die suggerierte Genauigkeit der Ergebnisse berechtigt ist?

5. Zweiseitige Hypothesentests

Die mathematische Behandlung zweiseitiger Hypothesentests mit einer Hypothese $H: p=p_0$ und einer Alternative $K: p \neq p_0$ lässt sich sehr leicht auf die obige Diskussion von Intervallschätzern zurückführen. Man geht von der Richtigkeit der Hypothese aus und bestimmt einen Annahmebereich A so, dass $P_{p_0}(X \in A) \geq 1-\alpha$ gilt (Beschränkung des Fehlers 1. Art), wobei α das Niveau des Tests ist. Man bestimmt also den Intervallschätzer wie oben mit dem $\frac{\alpha}{2}$ -Quantil und dem

$\left(1 - \frac{\alpha}{2}\right)$ -Quantil für den Annahmebereich. Dieses Verfahren funktioniert für die anderen implementierten Wahrscheinlichkeitsverteilungen genauso.

Spannender wird es, wenn man den Fehler 2. Art betrachten möchte. Dafür definiert man die Gütefunktion des Tests:

$$\beta(p) := P_p(X \in \bar{A}).$$

Der Fehler 2. Art ist dann für die zur Alternative gehörenden Wahrscheinlichkeiten p gegeben durch $1-\beta(p)$.

Aufgabe: Bei einem Glücksspielautomaten beträgt angeblich die Gewinnwahrscheinlichkeit $p=0,3$. Diese Angabe soll in 170 Spielrunden überprüft werden.

Der Annahmebereich des Tests ergibt sich bei einem Testniveau $\alpha=0,05$ aus dem 0,025-Quantil $x_{0,025}$ und dem 0,975-Quantil $x_{0,975}$ der Binomialverteilung mit den Parametern $n=170$ und $p_0=0,3$ zu $A=[40;63]$. Liegt also das Stichprobenergebnis nicht in A , sollte man die Hypothese verwerfen.

Den Fehler 2. Art kann man nun mit Hilfe des Rechners graphisch darstellen:

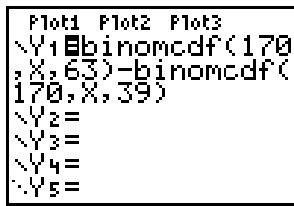


Abb. 7

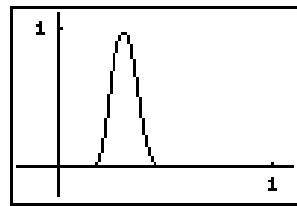


Abb. 8

Mit der TRACE-Funktion kann man auf dem Graphen entlang fahren und erkennen, dass die Wahrscheinlichkeit für einen Fehler 2. Art erst dann kleiner ist als z.B. 10%, wenn die wahre Wahrscheinlichkeit kleiner ist als 0,193 oder größer ist als 0,422; anderenfalls ist die Wahrscheinlichkeit für einen Fehler 2. Art nicht unbedeutend.

6. Einseitige Hypothesentests

Diese lassen sich sehr ähnlich behandeln, man benötigt jedoch nur ein Quantil für den Annahmehbereich.

Aufgabe: An einer Schule fallen durchschnittlich 11% der Mathematik-Klassenarbeiten schlechter als ausreichend aus. Ein neuer Kollege hat nun im letzten Schuljahr aber 43 seiner insgesamt 290 Mathematikarbeiten schlechter als ausreichend beurteilt. Ist er ein besonders „scharfer Hund“?

Zu testen ist $H: p \leq 0,11$ gegen $K: p > 0,11$. Man bestimmt für ein Niveau von $\alpha = 0,05$ das 0,95-Quantil der Binomialverteilung mit den Parametern $n=290$ und $p=0,11$ zu $x_{0,95} \approx 41$. Die Hypothese H , dass er nicht ungewöhnlich viele schlechte Noten gibt, ist also zu verwerfen, wenn das Stichprobenergebnis außerhalb $A=[0, 41]$ liegt; dieses ist der Fall: Die Hypothese muss verworfen werden, der neue Kollege scheint Klassenarbeiten tatsächlich recht streng zu benoten. Er hat in

$$\frac{43}{290} \approx 0,148$$

Fällen eine schlechtere Note als ausreichend gegeben. Bestimmt man jedoch die Wahrscheinlichkeit für einen Fehler 2. Art für diese relative Häufigkeit, $\text{binomcdf}(290; 0,148; 41) \approx 0,414$, so sieht man, dass die Irrtumswahrscheinlichkeit dieser Aussage noch relativ groß ist. Für genauere Aussagen müsste die Stichprobe vergrößert werden.

Über die geschickte Hypothesenwahl bei einseitigen Tests kann man sehr gut diskutieren - insbesondere dann, wenn wirtschaftliche Argumente einbezogen werden können.

Auch die Testverfahren lassen sich analog mit anderen implementierten Wahrscheinlichkeitsverteilungen behandeln.²

7. Ausblick

² Eine Übersicht über wichtige Wahrscheinlichkeitsverteilungen mit deren Kenngrößen und Anwendungsmöglichkeiten kann vom Autor per E-Mail angefordert werden.

Die Rechner von TI haben auch feste Module für Konfidenzintervalle und Hypothesentests implementiert. Jene können jedoch nur als Black-Box eingesetzt werden und geben einen P -Wert an, so dass die Test-Module flexibel zu verschiedenen Testniveaus verwendet werden können. Der P -Wert ist ein Maß (aber keine Wahrscheinlichkeit) für die Verträglichkeit von Stichprobenausfall und Hypothese. Ist der P -Wert $\leq \alpha$, so wird die Hypothese verworfen.³

Unter anderem ist auch ein Test-Modul zum χ^2 -Test vorhanden. Dieses ist nach der Behandlung der Hypothesentests im Unterricht sehr gut einsetzbar, da mit dem χ^2 -Test die Unabhängigkeit mehrerer Zufallsvariablen sowie die Annahme der Normalverteilung von Zufallsvariablen untersucht werden können. Beide Annahmen setzt man im Allgemeinen unkritisch voraus.⁴

Literatur:

- [1] F. Barth, R. Haller: *Stochastik Leistungskurs*; Ehrenwirth Verlag, München 1985
- [2] R. Puscher (Hrsg.): *PROST - Problemorientierte Stochastik, Stochastik-Sammlung 1*; MUED e.V., Appelhülsen 2003
- [3] L. Sachs: *Angewandte Statistik*; Springer-Verlag, Berlin 2002
- [4] W. Stahel: *Statistische Datenanalyse*; Vieweg Verlag, Braunschweig 2000
- [5] A. Warmeling: *PROST - Problemorientierte Stochastik, Stochastik-Sammlung 2*, MUED e.V., Appelhülsen 2002
- [6] H. Griesel u.a.(Hrsg.): *Elemente der Mathematik, Leistungskurs Stochastik*, Schroedel Verlag, Hannover 2003

Autor:

Gerd Hinrichs
Weddigenstraße 9a
26603 Aurich
E-Mail: gerd_hinrichs@web.de

³ Näheres zum P -Wert findet man bei Stahel [4].

⁴ Näheres zum χ^2 -Test findet man bei Stahel [4] und im Schulbuch *Elemente der Mathematik* [6] oder sehr ausführlich auch bei Sachs [3].

