Überraschung beim Umgang mit der Normalverteilung

Hans Rudolf Schneebeli, Robert Märki

1. Motivation

Die Normalverteilung mit Erwartungswert μ und Streuung σ besitzt die Dichtefunktion

$$\rho(\mu,\sigma,x)=\frac{1}{\sigma\sqrt{2\pi}}\,exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\quad .$$

Daher treten in allen möglichen Anwendungen Integrale der folgenden Art auf:

$$\int_a^b \rho\left(\mu,\sigma,x\right)\,dx =: I\left(\mu,\sigma,a,b\right)$$

Bekanntlich ist die Dichtefunktion ρ der Normalverteilung nicht elementar integrierbar. Der Hauptsatz der Integralrechnung hilft also nicht weiter, um Wahrscheinlichkeiten als bestimmte Integrale mit der Dichte ρ zu berechnen. Es bleibt als Ausweg die numerische Integration, sofern die Integrationsgrenzen a und b konkret gegeben sind. Traditionell wurde die allgemeine Normalverteilung durch eine affine Koordinatentransformation standardisiert. Die Standardnormalverteilung besitzt die Dichte $\rho(0,1,x)$. Die zugehörige Verteilungsfunktion

$$\Phi\left(u\right)=\int_{-\infty}^{u}\rho\left(0,\,1,\,z\right)\,dz$$

wurde für ausgewählte Werte von u numerisch berechnet und tabelliert. In den Anwendungen wird im allgemeinen Fall die Transformation

$$u=\frac{x-\mu}{\sigma}$$

zur Standardisierung verwendet.

Taschenrechner machen leistungsfähige numerische Integrationsroutinen allgemein verfügbar. So stellt sich die Frage, ob die Standardnormalverteilung im Unterricht noch eine besondere Rolle spielen soll. Erlaubt uns die numerische Integration, Normalverteilung ohne Standardisierung zu behandeln? Die Antwort lautet: Im Prinzip ja, *aber ... .*

2. Erfahrungen

Die folgenden Beispiele beziehen sich auf die Implementation des Gauss-Kronrod-Verfahrens bei den Rechnern TI Voyage™200 (OS-Version 3.10) und TI-Nspire™. Mit dem Befehl

`nInt(ρ(μ,σ,x),x,a,b)` $=I\left(\mu,\sigma,a,b\right)$

wurden die Werte aus Tabelle 1 gefunden. Mit den gewählten a, b wäre $I\approx 1$ zu erwarten.

μ	σ	a	b	$I(\mu,\sigma,a,b)$
0	1	$-\infty$	∞	1
0	0.01	$-\infty$	∞	1
10	0.01	$-\infty$	∞	1
10	0.001	$-\infty$	∞	**0**
100	0.01	0	300	**0**
0	1	-5000	7000	**0**

Tabelle 1: Experimente mit den Gauss-Kronrod-Verfahren

3. Weitere Fehler, eine Analyse

Eine grosse Überraschung erlebt, wer das Integralzeichen des TI-Voyage™200 benutzt, um z.B. das Integral

$$\int_5^{5,1}\rho\left(5,\,0.1,\,x\right)\,dx$$

zu berechnen: Die Antwort lautet $5.405...\cdot 10^{457}$. Der Grund für das Verhalten der Software wird klar, wenn man versucht,

$$\int_5^{10}\rho\left(5,\,0.1,\,x\right)\,dx$$

mit dem CAS des Voyage™200 'formal' zu integrieren, statt die numerische Integration direkt aufzurufen. Das CAS entwickelt den Term

$$-\frac{\left(x-\mu\right)^2}{2\sigma^2}\qquad (*)$$

vor der numerischen Integration in eine Summe und benutzt dann die Darstellung

$$exp\left(-\frac{\left(x-5\right)^2}{2\cdot 0{,}01}\right)=exp\left(-1250\right)\cdot exp\left(500\cdot x-50\cdot x^2\right)$$
$$\approx 1{,}35...\cdot 10^{-543}\cdot exp\left(500\cdot x-50\cdot x^2\right)$$

Dieser numerische Sündenfall lässt sich mit der standardisierten Normalverteilung vermeiden. Würde das CAS die vom Benutzer eingegebene Standardisierung in $(*)$ für die numerische Integration verwenden, wäre das Beispiel harmlos. Das Beispiel zeigt einen Nachteil einer Software, die alleine nach syntaktischen Regeln Termumformungen vornimmt. Im Kontext mit der numerischen Integration ist die Umformung unsinnig. Sie zerstört die implizit formulierte Standardisierung - mit verheerenden Folgen für die nachgeschaltete numerische Integration. Wer von Anfang an mit der numerischen Integration `nInt(...)` arbeitet, findet hier jedoch das korrekte Ergebnis [$\approx 0{,}5$].

4. Folgerungen

'Formal exakte' Integration ist bei der Normalverteilung an sich sinnlos. Die numerische Integration ist in diesem Falle prinzipiell der richtige Ansatz. Aber bei gegebenen Werten für μ und σ kann es Werte für a und b geben, so dass das Gauss-Kronrod-Verfahren das Integral

$$I\left(\mu,\sigma,a,b\right)$$

nicht korrekt erfasst. Das ist eigentlich keine Überraschung, ist doch bekannt, dass die numerische Integration mit nur endlich vielen Abtastwerten des Integranden arbeitet. Wenn die Funktionswerte an den Abtastpunkten so nahe bei 0 liegen, dass das Verfahren sie zu 0 rundet und wenn die Dichte ρ im Intervall $[a,b]$ asymmetrisch verteilt und lokal konzentriert ist, so ist die blind angewandte numerische Integration nicht zuverlässig.

Bei der standardisierten Normalverteilung liefert die Methode von Gauss-Kronrod für alle praktischen Belange zuverlässige Ergebnisse, wenn die Integrationsgrenzen mit Rücksicht auf die σ- Regeln einsichtig gewählt werden.

Die *Standardisierung der Normalverteilung kann helfen*, dass jene Überlegungen gemacht werden, die nötig sind, damit das numerische Integrationsverfahren gute Näherungen liefert. Die Arbeit mit den Tabellen ist aber definitiv ein Anachronismus.

Autoren:
Robert Märki, Thun (Schweiz)
Gymnasium Schadau,
r.maerki@gymhmsschadau.ch

Dr. H.R. Schneebeli, Baden (Schweiz)
Aargauische Kantonsschule,
schneebe@othello.ch