

Zusammenfassung: In Götz 2000 ist das Testen von Hypothesen vom Bayesianischen Standpunkt aus besprochen worden. Dabei hat die Binomialverteilung als Versuchsverteilung fungiert. Auch ein verteilungsfreies Testverfahren, der Vorzeichentest, ist dort Bayesianisch zur Sprache gekommen.

In dieser Arbeit soll jene Thematik der beurteilenden Statistik anhand von weniger im Stochastikunterricht verbreiteten Verteilungen, nämlich der Multinomial- und der Poisson-Verteilung (siehe auch Götz 1998), wiederum Bayesianisch dargestellt werden. Auch diesmal wird sich zeigen, dass Berechnungen, die sich in natürlicher Weise aus der Problemstellung ergeben, mit Hilfe von *DERIVE* nahezu mühelos bewältigt werden können. Ebenso wird die Erzeugung von Funktionsgraphen zur Verdeutlichung der stochastischen Situation gelingen. Die Normalverteilung wird in diesem Zusammenhang u. a. in Götz 2001 ausführlich diskutiert.

Schließlich wird noch kurz auf das (Bayesianische) Schätzen von Parametern, die zweite wichtige Thematik der beurteilenden Statistik in der Schule, eingegangen.

1 Vorbemerkung

Die grundlegende Idee des Bayesianischen Testens von Hypothesen ist u. a. in Götz 2000 und 2001, Meyer 1998 oder natürlich in Wickmann 1990 geschildert worden, sodass an dieser Stelle darauf verzichtet werden kann. Ein kritischer Vergleich der klassischen mit der Bayesianischen Methode ist in Wickmann 2001a und 2001b zu finden, und soll hier (wie auch in Götz 2000) unterbleiben. Stattdessen treten nun zwei Verteilungen in das Zentrum der Ausführungen, die eher am Rande des üblichen Stochastikunterrichts zu finden sind: die Multinomial- und die Poisson-Verteilung. Es sei schon an dieser Stelle erwähnt, dass das Erklären des Prinzips der Bayesianischen Herangehensweise *nicht* dieser beiden Verteilungen bedarf, sondern sehr gut z. B. mit Hilfe der Binomialverteilung möglich ist (Götz 2000). Für die Normalverteilung sei auf Götz 2001 verwiesen.

Warum also werden dann diese für den Mathematikunterricht eher „exotischen“ Verteilungen herangezogen? — Zum einen ist die Forderung, Funktionen in mehreren Veränderlichen

in den Unterricht zu integrieren, immer wieder erhoben worden (siehe z. B. Kirsch 1990, Klika 1985 und 1986 und Schweiger 1995). Ein aktueller in der Statistik angesiedelter Beitrag ist Strick 2000. Zum anderen bietet die Poisson-Verteilung viele Querverbindungen zur Analysis, wie wir noch sehen werden. Der für die Poisson-Verteilung typische Grenzübergang, der die Binomialverteilung in sie überführt, taucht auch hier wieder — leicht modifiziert — auf. Schließlich soll ein wenig auf die Frage im Anschluss an Götz 2000 eingegangen werden, was es denn sonst noch so gibt, also ein Blick über den Tellerrand gewagt werden, um so einen Beitrag zum Hintergrundwissen der Lehrenden zu liefern.

Das in Götz 2000 nicht zur Sprache gekommene Schätzen von Parametern wird zum Schluss noch auf Bayesianische Art und Weise vorgestellt (anhand der Poisson-Verteilung, eine Übertragung auf z. B. die Binomialverteilung ist leicht möglich).

In allen Fällen wird *DERIVE* bei Berechnungen und zum Plotten von Funktionsgraphen (zum Teil im 3D-Modus) herangezogen. Ersteres führen zu konkreten numerischen Ergebnissen, Letzteres dient zur Veranschaulichung der jeweiligen stochastischen Situation.

2 Die Multinomialverteilung als Versuchsverteilung

BEISPIEL 1: Ein Glücksrad mit drei (verdeckten) Sektoren wird $n = 20$ mal gedreht. Dabei ergeben sich folgende Häufigkeiten für die einzelnen Sektoren:

$$k_1 = 7 ; \quad k_2 = 9 ; \quad k_3 = 4 \quad (\text{Daten } D) .$$

Was kann daher über die Trefferwahrscheinlichkeiten p_1 , p_2 und p_3 für die einzelnen Sektoren ausgesagt werden?

Lösung: A priori setzen wir

$$\pi(p_1, p_2, p_3) = \text{const.} \quad \forall p_1, p_2, p_3 \in \mathbb{R}_0^+$$

und

$$p_1 + p_2 + p_3 = 1 .$$

Die Versuchsverteilung ist

$$\begin{aligned} P(X_1 = k_1, X_2 = k_2, X_3 = k_3) &= \\ &= \frac{n!}{k_1! \cdot k_2! \cdot k_3!} \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot p_3^{k_3} \end{aligned}$$

bzw.

$$\begin{aligned} P(X_1 = k_1, X_2 = k_2, X_3 = n - k_1 - k_2) &= \\ &= \frac{n!}{k_1! \cdot k_2! \cdot (n - k_1 - k_2)!} \cdot \\ &\cdot p_1^{k_1} \cdot p_2^{k_2} \cdot (1 - p_1 - p_2)^{n - k_1 - k_2}, \end{aligned}$$

wobei die Zufallsvariablen X_i ($i = 1, 2, 3$) die Treffer im i -ten Sektor zählen.

Also ist a posteriori

$$\begin{aligned} \pi(p_1, p_2, p_3 | D) &= \\ &= \frac{P(D | p_1, p_2, p_3) \cdot \pi(p_1, p_2, p_3)}{P(D)} = \\ &= \frac{p_1^{k_1} \cdot p_2^{k_2} \cdot (1 - p_1 - p_2)^{n - k_1 - k_2}}{N} = \\ &= \frac{(n + 2)!}{k_1! \cdot k_2! \cdot k_3!} \cdot p_1^{k_1} \cdot p_2^{k_2} \cdot p_3^{k_3} \end{aligned}$$

mit

$$N = \int_0^1 \int_0^{1-p_1} p_1^{k_1} \cdot p_2^{k_2} \cdot (1 - p_1 - p_2)^{n - k_1 - k_2} dp_2 dp_1$$

(das ist die Dichte einer *multivariaten Beta-Verteilung*). Das Ergebnis folgt aus zwei Rechnungen, welche im Anhang angeführt sind.

Bei uns bedeutet dies konkret

$$\begin{aligned} \pi(p_1, p_2 | D) &= \\ &= \frac{22!}{7! \cdot 9! \cdot 4!} \cdot p_1^7 \cdot p_2^9 \cdot (1 - p_1 - p_2)^4. \end{aligned}$$

DERIVE zeigt uns den 3D-Funktionsgraphen: **Abbildung 1**.

Für $n = 60$ (also Erhöhung des Stichprobenumfangs) und

$$k_1 = 21; \quad k_2 = 27; \quad k_3 = 12 \quad (\text{Daten } D)$$

ist eine schärfere Einschätzung der Lage a posteriori möglich (**Abbildung 2** mittels *DERIVE*):

$$\begin{aligned} \pi(p_1, p_2 | D) &= \frac{62!}{21! \cdot 27! \cdot 12!} \cdot \\ &\cdot p_1^{21} \cdot p_2^{27} \cdot (1 - p_1 - p_2)^{12}. \end{aligned}$$

Aus diesen beiden Abbildungen ist weiters die Lage des Maximums abzulesen, nämlich an den Stellen der relativen Häufigkeiten, also hier bei $\frac{7}{20}$ und $\frac{9}{20}$. Die daraus resultierende Vermutung, dies ist immer so, kann mittels partieller Differentiation bestätigt werden.

Mit *DERIVE* können über diese Punktschätzungen hinaus (siehe auch Abschnitt 4) zweidimensionale Bereiche unter der A-posteriori-Dichtefunktion quantifiziert werden. Folgende Wahrscheinlichkeit ist beispielsweise ein solches Resultat (unter Verwendung der ursprünglichen A-posteriori-Dichtefunktion für $n = 20$), wobei wir $\frac{1}{4} \leq \frac{7}{20} \leq \frac{2}{5}$ und $\frac{1}{5} \leq \frac{9}{20} \leq \frac{1}{2}$ beachten:

$$\begin{aligned} P\left(\frac{1}{4} \leq p_1 \leq \frac{2}{5}; \frac{1}{5} \leq p_2 \leq \frac{1}{2}\right) &= \\ &= \int_{\frac{1}{4}}^{\frac{2}{5}} \int_{\frac{1}{5}}^{\frac{1}{2}} \frac{22!}{7! \cdot 9! \cdot 4!} \cdot p_1^7 \cdot p_2^9 \cdot \\ &\cdot (1 - p_1 - p_2)^4 dp_2 dp_1 = 0,392134. \end{aligned}$$

Daraus können wir auch

$$P\left(\frac{1}{10} \leq p_3 \leq \frac{11}{20}\right) = 0,392134$$

wegen

$$1 - \frac{1}{4} - \frac{1}{5} = \frac{11}{20} \quad \text{und} \quad 1 - \frac{2}{5} - \frac{1}{2} = \frac{1}{10}$$

schließen.

Interessieren wir uns nur z. B. für p_1 , dann liefert die *Randdichte* mittels *DERIVE*

$$\begin{aligned} P\left(p_1 \leq \frac{7}{20}\right) &= \\ &= \int_0^{\frac{7}{20}} \int_0^{1-p_1} \frac{22!}{7! \cdot 9! \cdot 4!} \cdot p_1^7 \cdot p_2^9 \cdot \\ &\cdot (1 - p_1 - p_2)^4 dp_2 dp_1 = 0,526440. \end{aligned}$$

3 Die Poisson-Verteilung als Versuchsverteilung

BEISPIEL 2 (siehe auch Götz 1998): Eine Telefonzentrale erhält im Laufe einer Stunde durchschnittlich 60 Anrufe. Wir zählen nun n Minuten lang die hereinkommenden Anrufe: Daten $D = (x_1, \dots, x_n)$. Die Zufallsvariable X zähle die Anrufe pro Minute, ihre n Kopien X_i , welche die Anrufe in der i -ten Minute zählen ($i = 1, \dots, n$), seien voneinander unabhängig.

Wir modellieren

$$P(X = x_i) = \frac{\lambda^{x_i}}{x_i!} \cdot e^{-\lambda} \quad i = 1, \dots, n.$$

Was kann aufgrund der Daten über den Parameter $\lambda \in \mathbb{R}^+$ ausgesagt werden?

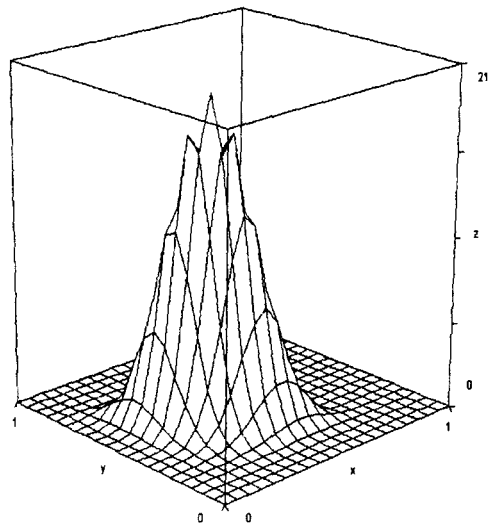


Abbildung 1: Graph der A-posteriori-Dichtefunktion für $n = 20$

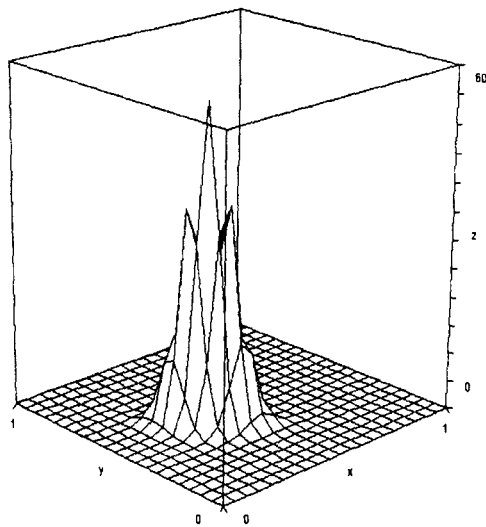


Abbildung 2: Graph der A-posteriori-Dichtefunktion für $n = 60$

Lösung: Wegen $E(X) = \lambda \approx 1$ setzen wir a priori

$$\pi(\lambda) = e^{-\lambda} \quad \lambda \in \mathbb{R}^+.$$

[Es ist $E(\lambda) = \int_0^\infty \lambda \cdot e^{-\lambda} d\lambda = -\lambda \cdot e^{-\lambda} \Big|_0^\infty + \int_0^\infty e^{-\lambda} d\lambda = -e^{-\lambda} \Big|_0^\infty = 1$. Natürlich gibt es noch andere Dichtefunktionen mit Erwartungswert eins. Tatsächlich sind es vor allem technische Gründe, die uns zur Exponentialverteilung greifen lassen, wie wir noch sehen werden. Inhaltlich könnten wir argumentieren, dass die Poisson-Verteilung auch „Verteilung seltener Ereignisse“ genannt wird, siehe z. B. Reichel/Hanisch/Müller 1992, S. 183, und dass daher die Wahrscheinlichkeit für große Parameterwerte gering sein wird. Dies spiegelt die gewählte Exponentialverteilung gut wider.]

A posteriori ist dann mit

$$P(D|\lambda) = \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{x_n}}{x_n!} \cdot e^{-\lambda}$$

und

$$\sum_{i=1}^n x_i = m \in \mathbb{N}$$

$$\begin{aligned} \pi(\lambda|D) &= \frac{P(D|\lambda) \cdot \pi(\lambda)}{P(D)} = \\ &= \frac{e^{-n \cdot \lambda} \cdot \left(\prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \right) \cdot e^{-\lambda}}{\int_0^\infty e^{-n \cdot \lambda} \cdot \left(\prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} \right) \cdot e^{-\lambda} d\lambda} = \\ &= \frac{e^{-(n+1) \cdot \lambda} \cdot \lambda^m}{\int_0^\infty e^{-(n+1) \cdot \lambda} \cdot \lambda^m d\lambda} = \\ &= \frac{e^{-\lambda \cdot (1+n)} \cdot \lambda^m \cdot (1+n)^{m+1}}{m!}. \end{aligned}$$

Die letzte Umformung wird im Anhang erklärt.

Erhalten wir also in n Minuten m Anrufe, so ist $\pi(\lambda|D)$ wie eben berechnet. ($\sum_{i=1}^n X_i$ ist eine *suffiziente* Statistik, das heißt, für die A-posteriori-Einschätzung der Lage kommt es nicht auf die einzelnen Daten an, sondern nur auf ihre Summe.)

Allgemein lautet die Dichte der sogenannten *Gamma-Verteilung* mit $a, b \in \mathbb{N}$ ($b > 0$)

$$f(x) = \begin{cases} \frac{1}{a!} \cdot x^a \cdot b^{a+1} \cdot e^{-b \cdot x}, & x > 0, \\ 0 & \text{sonst.} \end{cases}$$

Wegen

a priori	a posteriori
$a = 0$	$a = \sum_{i=1}^n x_i = 0 + \sum_{i=1}^n x_i$
$b = 1$	$b = 1 + n$

erkennen wir, dass sowohl die A-priori als auch die A-posteriori-Dichtefunktion zur Familie der Gamma-Verteilungen gehören. Sie unterscheiden sich voneinander nur durch die Parameter a und b , welche immer mehr von den Daten alleine bestimmt werden. Wir sagen: „Die Poisson-Verteilung ist zur Gamma-Verteilung konjugiert.“ Diese Eigenschaft ist sehr praktisch und tritt auch bei anderen Verteilungspaaren auf (z. B. die Binomial- und die Beta-Verteilung, siehe Götz 2000). Eine (neuerliche) Datenerhebung bringt „nur“ eine Änderung der Parameter der A-posteriori-Dichtefunktion mit sich, der Verteilungstyp bleibt gleich.

Aus diesem Grund haben wir auch a priori die Exponentialverteilung gewählt. Ihr verdanken wir die geschlossene Darstellung der A-posteriori-Verteilung. Eine „gleichmäßige Verteilung“ a priori für den Parameter λ würde vielleicht weniger willkürlich wirken, und sie führt ebenfalls a posteriori zur Gamma-Verteilung für λ . Allerdings stellt sich hier das Problem, dass $\lambda \in \mathbb{R}^+$ gilt, d. h. dass der Parameterraum unbeschränkt ist. Daher ist die gleichmäßige Dichte darauf eine uneigentliche. Die Wahl anderer Dichten a priori lässt i. Allg. keine geschlossene Auswertung der auftretenden Integrale zur Berechnung der A-posteriori-Verteilung zu. Letztere muss dann ausschließlich numerisch geschehen. Insofern stellt die Wahl der Exponentialverteilung für λ a priori einen Kompromiss dar: numerische Probleme treten auf, jedoch ist gleichzeitig eine gewisse Struktur gegeben.

Sehen wir uns nun den Spezialfall an, dass in n Minuten $m = n$ Anrufe hereinkommen. (Dies entspricht ja der eingangs erwähnten Beobachtung, dass in einer Stunde durchschnittlich 60 Anrufe zu verzeichnen sind.)

DERIVE liefert für $n = m = 10$

$$\begin{aligned} P(\lambda \leq 1) &= \int_0^1 \frac{e^{-11 \cdot \lambda} \cdot \lambda^{10} \cdot 11^{11}}{10!} d\lambda = \\ &= 0,540111. \end{aligned}$$

Für $n = m = 1000$ ergibt sich

$$\begin{aligned} P(\lambda \leq 1) &= \\ &= \int_0^1 \frac{e^{-1001 \cdot \lambda} \cdot \lambda^{1000} \cdot 1001^{1001}}{1000!} d\lambda = \\ &= 0,504183. \end{aligned}$$

Es zeigt sich also, dass die ursprüngliche Asymmetrie der A-priori-Verteilung bezüglich dem Wert $\lambda = 1$ [es ist $P(\lambda \leq 1) = \int_0^1 e^{-\lambda} d\lambda = 0,632121$] aufgrund der erhobenen Daten mit wachsendem Stichprobenumfang zugunsten von $P(\lambda \leq 1) = P(\lambda > 1) \approx 0,5$ a posteriori verschwindet.

In jedem Fall (i. e. $n = m$) ist mit *DERIVE*

$$E(\lambda) = \int_0^\infty \lambda \cdot \pi(\lambda|D) d\lambda = 1.$$

Eine Rechnung im Anhang bestätigt dieses Ergebnis.

Ist $n = 20$ und $m = 10$, so ist

$$\pi(\lambda|D) = \frac{e^{-21 \cdot \lambda} \cdot \lambda^{10} \cdot 21^{11}}{10!}.$$

Den Graphen zeichnet *DERIVE* (**Abbildung 3**, es ist der breitere und dementsprechend niedrigere) und den Erwartungswert berechnet *DERIVE* zu

$$E(\lambda) = 0,523808.$$

Eine Erhöhung des Stichprobenumfangs macht uns wieder sicherer:

Für $n = 40$ und $m = 20$ ist

$$\pi(\lambda|D) = \frac{e^{-41 \cdot \lambda} \cdot \lambda^{20} \cdot 41^{21}}{20!}.$$

Wir gewinnen mittels *DERIVE* den zugehörigen Graphen, der ebenfalls in **Abbildung 3** zu sehen ist. Er ist deutlich schlanker und dementsprechend höher als der für $n = 20$ und $m = 10$.

Diesen durch *DERIVE* gewonnenen Eindruck können wir auch theoretisch absichern:

Der Erwartungswert ergibt sich allgemein zu

$$E(\lambda) = \frac{m+1}{n+1},$$

und die Varianz ist

$$D^2(\lambda) = \frac{m+1}{(n+1)^2},$$

wie die zugehörigen Rechnungen im Anhang zeigen.

Wenn nun das Verhältnis $\frac{m}{n}$ (i. e. die Anzahl der Anrufe pro Minute) auch für große n konstant bleibt, so folgt

$$\begin{aligned} \lim_{\substack{n \rightarrow \infty \\ \frac{m}{n} \text{ fest}}} D^2(\lambda) &= \\ &= \lim_{\substack{n \rightarrow \infty \\ \frac{m}{n} \text{ fest}}} \frac{\frac{m}{n} + \frac{1}{n}}{\left(\frac{n+1}{n}\right)^2} = \lim_{\substack{n \rightarrow \infty \\ \frac{m}{n} \text{ fest}}} \frac{1}{n} \cdot \frac{m}{n} = 0. \end{aligned}$$

Dieser Grenzübergang ist typisch für die Poisson-Verteilung: Bekanntlich konvergiert eine Binomialverteilung mit den Parametern n und p_n , wenn wir den Grenzübergang

$$n \rightarrow \infty \quad \text{und} \quad \lim_{n \rightarrow \infty} n \cdot p_n = \lambda \in \mathbb{R}^+$$

betrachten, gegen eine Poisson-Verteilung mit dem Parameter λ . Halten wir uns weiters vor Augen, dass einerseits der Erwartungswert der Binomialverteilung gerade $n \cdot p_n$ ist, andererseits der Erwartungswert λ der Poisson-Verteilung nach der Maximum-Likelihood-Methode durch den Durchschnittswert $\frac{m}{n}$ (i. e. das arithmetische Mittel der Daten) geschätzt wird, so ist der obige Grenzübergang in diesem Zusammenhang nicht überraschend.

4 Ein Schätzproblem

Wir wollen nun die A-posteriori-Dichte zur Bereichsschätzung heranziehen:

BEISPIEL 2 hat uns für $n = 20$ und $m = 10$

$$\pi(\lambda|D) = \frac{e^{-21 \cdot \lambda} \cdot \lambda^{10} \cdot 21^{11}}{10!} \quad \lambda \in \mathbb{R}^+$$

gebracht. Wir suchen nun ein Intervall $[\lambda_1^{\min}, \lambda_2^{\min}]$ kürzester Länge, welches für gegebenes $\gamma \in (0, 1)$

$$P(\lambda_1^{\min} \leq \lambda \leq \lambda_2^{\min}) = \int_{\lambda_1^{\min}}^{\lambda_2^{\min}} \pi(\lambda|D) d\lambda = \gamma$$

erfüllt. Ein solches Intervall heißt *HPD-Bereich* (höchster A-posteriori-Dichte-Bereich). In ihm liegt der Parameter λ mit Wahrscheinlichkeit γ .

Es gilt offensichtlich in unserem Fall $\pi(\lambda_1^{\min}|D) = \pi(\lambda_2^{\min}|D)$, wenn wir uns **Abbildung 3** vor Augen halten. Gehen wir von einem Intervall aus, dessen Grenzen nicht dieselben Funktionswerte aufweisen und welches die Wahrscheinlichkeitsmasse γ umfasst. Dann können wir durch Verschieben in Richtung des Maximums von $\pi(\lambda|D)$ immer ein Intervall finden, welches kürzer ist als das ursprüngliche und trotzdem dieselbe Masse γ enthält (siehe auch Götz 2001, S. 153 f.).

Die Bedingung

$$\pi(\lambda_1^{\min}|D) = \pi(\lambda_2^{\min}|D)$$

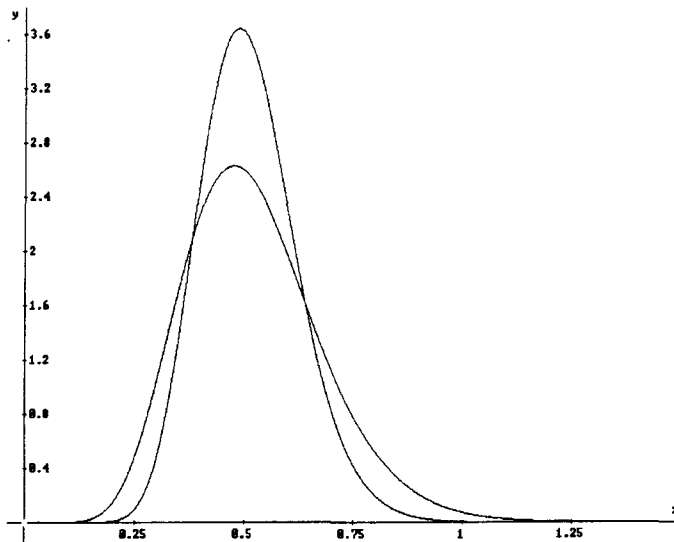


Abbildung 3: Graphen der A-posteriori-Dichtefunktionen für $n = 20$ und $m = 10$ bzw. $n = 40$ und $m = 20$

ist der Schlüssel zur näherungsweise Bestimmung des gesuchten HPD-Bereichs. Dabei tasten wir uns schrittweise heran: Die in **Abbildung 4** sichtbare waagrechte Linie wird nach oben oder unten verschoben, ihre Schnittpunkte (genauer: die x -Koordinaten davon) mit $\pi(\lambda|D)$ sind Näherungswerte für die gesuchten Grenzen. Das Ergebnis der Integration von einer Schnittstelle zur anderen der A-posteriori-Dichte zeigt uns, in welche Richtung die waagrechte Linie für den nächsten Schritt verschoben werden muss: Ist der erhaltene Wert kleiner als γ (hier: 0,95), dann wird nach unten verschoben, ist er zu groß, dann nach oben.

Abbildung 4 entsteht mit Hilfe von *DERIVE*, ebenso gelingt die Berechnung von

$$\lambda_1^{\min} = 0,237093 \quad \text{und} \quad \lambda_2^{\min} = 0,838729 .$$

Für $n = 40$ und $m = 20$ haben wir

$$\pi(\lambda|D) = \frac{e^{-41 \cdot \lambda} \cdot \lambda^{20} \cdot 41^{21}}{20!} \quad \lambda \in \mathbb{R}^+$$

erhalten (siehe auch **Abbildung 3**). Der entsprechende HPD-Bereich ist nun natürlich kürzer als eben:

$$\lambda_1^{\min} = 0,303398 \quad \text{und} \quad \lambda_2^{\min} = 0,735059$$

mittels *DERIVE*. Ein letztes Mal können wir uns davon überzeugen, dass eine Erhöhung des Stichprobenumfangs ein Mehr an Sicherheit mit sich bringt.

5 Resümee

Die Verwendung von *DERIVE* erlaubt es, die jeweilige A-posteriori-Einschätzung numerisch und graphisch zu bewerten.

Im Falle der *Multinomialverteilung* gelingt so eine Verallgemeinerung der Binomialverteilung, die in **BEISPIEL 1** vorkommende zweidimensionale A-posteriori-Dichtefunktion kann mittels *DERIVE* als 3D-Plot dargestellt werden, so dass hier eine (seltene) Auseinandersetzung mit *mehrdimensionalen* geometrischen Objekten in einem Koordinatensystem im Mathematikunterricht passieren kann. Analytische Untermauerungen wie die Berechnung des Normierungsfaktors (Mehrfachintegrale) oder die Lage des Maximums (partielle Differentiation) gehören in diesem Zusammenhang nicht in den Unterricht, aber zum Hintergrundwissen der Lehrenden.

Ähnliches gilt für die *Poisson-Verteilung*, die das Thema „uneigentliche Integrale“ (die Gamma-Funktion in dieser Darstellung) birgt. Die Auswertung derselben ergibt für natürliche Zahlen n (ohne Null) die Fakultät von $n - 1$. *Alleine dieses* Ergebnis benötigen wir für unsere Anwendungen, weil die Poisson-Verteilung nur auf den natürlichen Zahlen definiert ist.

Das vorgestellte *Schätzproblem* wird in sehr einfacher Weise mit *DERIVE* einer approximativen Lösung zugeführt. Dabei geht es weniger um das tatsächliche näherungsweise zu bestimmende Ergebnis als um den Weg dorthin. Die

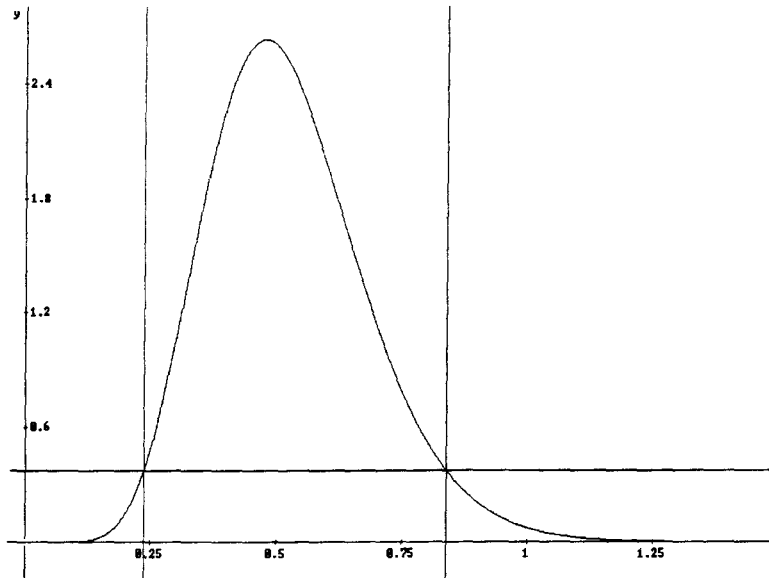


Abbildung 4: HPD-Bereich für λ ($n = 20$ und $m = 10$)

Planung des nächsten Iterationsschrittes (zum nächst„besseren“ Intervall) erfordert ein gründliches Erfassen der analytischen Situation, welche zwei Bedingungen mit sich bringt. Man lernt so die zugrundeliegende Funktion wirklich sehr genau kennen.

Insgesamt sollte hier anhand zweier eher am Rande eines „durchschnittlichen“ Stochastikcurriculums angesiedelten Verteilungen die Tragfähigkeit des Bayesianischen Konzepts gezeigt werden (in Ergänzung zu Götz 2000).

Das breite Anwendungsspektrum von *DERIVE* unterstützt dieses Vorhaben.

Anhang

Zur Multinomialverteilung

Erstens ist

$$\int_0^c x^n \cdot (c-x)^m dx = \frac{n! \cdot m!}{(n+m+1)!} \cdot c^{n+m+1}$$

$$\forall n, m \in \mathbb{N} \quad \text{und} \quad c \in \mathbb{R}.$$

Wir begründen mittels partieller Integration und erhalten

$$\begin{aligned} \int_0^c \underbrace{x^n}_{f'} \cdot \underbrace{(c-x)^m}_g dx &= \\ &= \frac{x^{n+1}}{n+1} \cdot (c-x)^m \Big|_0^c + \end{aligned}$$

$$\begin{aligned} &+ \int_0^c \frac{x^{n+1}}{n+1} \cdot m \cdot (c-x)^{m-1} dx = \\ &= \frac{m}{n+1} \cdot \left\{ \left[\frac{x^{n+2}}{n+2} \cdot (c-x)^{m-1} \right]_0^c + \right. \\ &\quad \left. + \int_0^c \frac{x^{n+2}}{n+2} \cdot (m-1) \cdot (c-x)^{m-2} dx \right\} = \\ &= \frac{m \cdot (m-1)}{(n+1) \cdot (n+2)} \cdot \int_0^c x^{n+2} \cdot (c-x)^{m-2} dx = \dots = \\ &= \frac{m \cdot (m-1) \cdot \dots \cdot 1}{(n+1) \cdot (n+2) \cdot \dots \cdot (n+m)} \cdot \int_0^c x^{n+m} dx = \\ &= \frac{m!}{(n+1) \cdot \dots \cdot (n+m)} \cdot \left[\frac{x^{n+m+1}}{n+m+1} \right]_0^c = \\ &= \frac{m!}{(n+1) \cdot \dots \cdot (n+m+1)} \cdot c^{n+m+1} = \\ &= \frac{m! \cdot n!}{(n+m+1)!} \cdot c^{n+m+1}. \end{aligned}$$

Zweitens berechnen wir damit den Wert N

des Integrals im Nenner, nämlich

$$\begin{aligned}
 & \int_0^1 \int_0^{1-p_1} p_1^{k_1} \cdot p_2^{k_2} \cdot (1-p_1-p_2)^{n-k_1-k_2} dp_2 dp_1 = \\
 &= \int_0^1 p_1^{k_1} \cdot \left[\int_0^{1-p_1} p_2^{k_2} \cdot (1-p_1-p_2)^{n-k_1-k_2} dp_2 \right] dp_1 = \\
 &= \int_0^1 p_1^{k_1} \cdot \frac{k_2! \cdot (n-k_1-k_2)!}{(n-k_1+1)!} \cdot (1-p_1)^{n-k_1+1} dp_1 = \\
 &= \frac{k_2! \cdot (n-k_1-k_2)!}{(n-k_1+1)!} \cdot \int_0^1 p_1^{k_1} \cdot (1-p_1)^{n+1-k_1} dp_1 = \\
 &= \frac{k_2! \cdot (n-k_1-k_2)!}{(n-k_1+1)!} \cdot \frac{k_1! \cdot (n+1-k_1)!}{(n+2)!} = \\
 &= \frac{k_1! \cdot k_2! \cdot k_3!}{(n+2)!}.
 \end{aligned}$$

Zur Poisson-Verteilung

Das Integral im Nenner berechnen wir diesmal mit der Substitution $\lambda \cdot (1+n) = \mu$, $\frac{d\mu}{d\lambda} = 1+n$ und $d\lambda = \frac{d\mu}{1+n}$, $\lambda = \frac{\mu}{1+n}$ so:

$$\begin{aligned}
 & \int_0^\infty e^{-\lambda \cdot (1+n)} \cdot \lambda^m d\lambda = \\
 &= \int_0^\infty e^{-\mu} \cdot \frac{\mu^m}{(1+n)^m} \cdot \frac{d\mu}{1+n} = \\
 &= \frac{1}{(1+n)^{m+1}} \cdot \int_0^\infty e^{-\mu} \cdot \mu^m d\mu = \\
 &= \frac{\Gamma(m+1)}{(1+n)^{m+1}}.
 \end{aligned}$$

Für die letzte Umformung rufen wir die Definition der *Gamma-Funktion* in Erinnerung:

$$\Gamma(p) := \int_0^\infty x^{p-1} \cdot e^{-x} dx \quad (p > 0).$$

Dieses uneigentliche Integral existiert genau dann, wenn $p > 0$ ist. Zum Beweis siehe Heuser 1986, S. 488. Für uns ist nun folgende Eigenschaft der Gamma-Funktion $\forall n \in \mathbb{N} \setminus \{0\}$ wichtig:

$$\Gamma(n) = (n-1)!,$$

wie man mittels partieller Integration sehen kann. Die in Rede stehende Situation bringt

nur Argumente der Gamma-Funktion aus den natürlichen Zahlen ohne Null. Daher kann sie durch den elementaren Begriff der Fakultät ersetzt werden.

Der Erwartungswert von $\pi(\lambda|D)$ ist

$$\begin{aligned}
 E(\lambda) &= \int_0^\infty \lambda \cdot \pi(\lambda|D) d\lambda = \\
 &= \int_0^\infty \frac{e^{-\lambda \cdot (1+n)} \cdot \lambda^{m+1} \cdot (n+1)^{m+1}}{m!} d\lambda = \\
 &= \frac{(n+1)^{m+1}}{m!} \cdot \int_0^\infty \lambda^{m+1} \cdot e^{-\lambda \cdot (1+n)} d\lambda = \\
 &= \frac{(n+1)^{m+1}}{m!} \cdot \int_0^\infty \frac{\mu^{m+1}}{(1+n)^{m+1}} \cdot e^{-\mu} \cdot \frac{d\mu}{1+n} = \\
 &= \frac{1}{m! \cdot (1+n)} \cdot \Gamma(m+2) = \\
 &= \frac{(m+1)!}{m! \cdot (1+n)} = \frac{m+1}{n+1}.
 \end{aligned}$$

(Dabei haben wir dieselbe Substitution wie vorher verwendet.)

Im Falle $m = n$ ist der Erwartungswert von $\pi(\lambda|D)$ daher eins.

Die Varianz erhalten wir mit

$$\begin{aligned}
 E(\lambda^2) &= \int_0^\infty \lambda^2 \cdot \pi(\lambda|D) d\lambda = \\
 &= \int_0^\infty \frac{e^{-\lambda \cdot (1+n)} \cdot \lambda^{m+2} \cdot (n+1)^{m+1}}{m!} d\lambda = \\
 &= \frac{(n+1)^{m+1}}{m!} \cdot \int_0^\infty \lambda^{m+2} \cdot e^{-\lambda \cdot (1+n)} d\lambda = \\
 &= \frac{(n+1)^{m+1}}{m!} \cdot \int_0^\infty \frac{\mu^{m+2}}{(1+n)^{m+2}} \cdot e^{-\mu} \cdot \frac{d\mu}{1+n} = \\
 &= \frac{1}{m! \cdot (1+n)^2} \cdot \Gamma(m+3) = \\
 &= \frac{(m+2)!}{m! \cdot (1+n)^2} = \frac{(m+2) \cdot (m+1)}{(n+1)^2}
 \end{aligned}$$

und dem Verschiebungssatz zu

$$\begin{aligned}
 D^2(\lambda) &= \\
 &= \frac{(m+2) \cdot (m+1)}{(n+1)^2} - \frac{(m+1)^2}{(n+1)^2} = \\
 &= \frac{m^2 + 3 \cdot m + 2 - m^2 - 2 \cdot m - 1}{(n+1)^2} = \\
 &= \frac{m+1}{(n+1)^2}.
 \end{aligned}$$

Literatur

- Borovcnik, M., Engel, J. und Wickmann, D. (Hgb., 2001): Anregungen zum Stochastikunterricht: Bericht von zwei Arbeitskreistagungen des Arbeitskreises „Stochastik in der Schule“ in der Gesellschaft für Didaktik der Mathematik e.V. vom 29./30. Oktober 1999 und 10.–12. November 2000 in Berlin. Verlag Franzbecker, Hildesheim, Berlin.
- Götz, S. (1998): Bayes-Statistik mit der Poisson-Verteilung. Erschienen in: Beiträge zum Mathematikunterricht 1998. Verlag Franzbecker, Hildesheim, Berlin (S. 224–227).
- Götz, S. (2000): Bayes-Statistik mit *DERIVE*. Erschienen in: Stochastik in der Schule **20**, Heft 3 (S. 2–14).
- Götz, S. (2001): Klassische und Bayesianische Behandlung von Stochastikaufgaben aus österreichischen Schulbüchern. Erschienen in: Borovcnik/Engel/Wickmann 2001 (S. 147–162).
- Heuser, H. (1986): Lehrbuch der Analysis, Teil 1. B. G. Teubner, Stuttgart (4. Auflage).
- Humenberger, H. (1996): Längen- und Winkelmessungen bei der Höhenbestimmung von Türmen. Optimierung und Fehlerbetrachtungen. Erschienen in: *mathematica didactica* **19**, Band 2 (S. 43–57).
- Kirsch, A. (1990): Funktionen zweier Veränderlicher in einem anwendungsorientierten Mathematikunterricht. Erschienen in: Runck, P. und Schlöglmann, W. (Hgb.): Vorträge bei der Fortbildungstagung für Mathematiklehrer der AHS und BHS **19**, Universität Linz (S. 44–61).
- Klika, M. (1985): Funktionen in mehreren Veränderlichen — kein Thema für den Mathematikunterricht? Erschienen in: Beiträge zum Mathematikunterricht 1985. Verlag B. Franzbecker, Didaktischer Dienst, Bad Salzdetfurth (S. 176–180).
- Klika, M. (1986): Zeichnen und zeichnen lassen: Funktionen von zwei Variablen. Erschienen in: *mathematiklehren* **14** (S. 61–63).
- Meyer, M. u. D. (1998): Hypothesentests nach Bayes. Entscheidungen für Hypothesen mit der Bayes-Formel. Erschienen in: Der Mathematikunterricht **44**, Heft 1 (S. 50–61).
- Reichel, H.-C., Hanisch, G. und Müller, R. (1992): Wahrscheinlichkeitsrechnung und Statistik. Mathematik für Schule und Praxis, Band 1 (herausgegeben von H.-C. Reichel). Verlag Hölder-Pichler-Tempsky, Wien (3. Auflage).
- Schweiger, F. (1995): Funktionen in mehreren Variablen — Aschenputtel der Schulmathematik. Erschienen in: Didaktik-Reihe der Österreichischen Mathematischen Gesellschaft, Heft 24 (S. 21–34).
- Strick, H. K. (2000): Einführung in die Multinomialverteilung und Chi-Quadrat-Anpassungstest mit Tabellenkalkulation. Erschienen in: Praxis der Mathematik **42**, Heft 5 (Schwerpunktthemenheft 23: Computer-Einsatz V) (S. 193–197).
- Wickmann, D. (1990): Bayes-Statistik. Einsicht gewinnen und entscheiden bei Unsicherheit. Mathematische Texte, Band 4 (herausgegeben von N. Knoche und H. Scheid). BI Wissenschaftsverlag, Mannheim/Wien/Zürich.
- Wickmann, D. (2001a): Der Theorieneintopf ist zu beseitigen. *Ereignis- und Zustandswahrscheinlichkeit* — Versuch einer Klärung des Wahrscheinlichkeitsbegriffs zum Zweck einer Methodenbereinigung. Erschienen in: Borovcnik/Engel/Wickmann 2001 (S. 123–131).
- Wickmann, D. (2001b): Inferenzstatistik ohne Signifikanztest. Vorschlag, den Signifikanztest im gymnasialen Unterricht nicht mehr zu verwenden. Erschienen in: Borovcnik/Engel/Wickmann 2001 (S. 133–138).

Anschriften des Verfassers

Mag. Dr. Stefan Götz
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Wien
Stefan.Goetz@univie.ac.at
Akademisches Gymnasium Wien I
Beethovenplatz 1
A-1010 Wien