

Seminarvortrag zu 'Philosophischen Grundlagen der Statistik'

## Kritik an Bayes-Inferenz

Malte Schierholz

München, den 17. Dezember 2010

Bayesian methods constitute a radically different way of doing science.  
(Brian Dennis)

# Was ist bayesianische Inferenz?

## Bayes Theorem

$$f(\theta|x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n|\theta) \cdot f(\theta)}{f(x_1, \dots, x_n)}$$

## Beispielsweise

$$X_1, \dots, X_n|\theta \stackrel{iid}{\sim} Po(\theta)$$

$$\theta \sim Gamma(a_0, b_0)$$

$$\Rightarrow \theta|x_1, \dots, x_n \sim Gamma(a_0 + \sum x_i, b_0 + n)$$

## Bezeichnungen:

Posteriori-Dichte  $f(\theta|x_1, \dots, x_n)$

Likelihood  $f(x_1, \dots, x_n|\theta)$

Priori-Dichte  $f(\theta)$

## Widersprüchliche Ergebnisse in einem Beispiel

Gegeben:  $X \sim N(\mu, \sigma^2)$ , wodraus sieben Mal unabhängig gezogen wird mit  $\bar{x} = 1$  und  $s = 1$ .

- 1 Wenn  $\sigma$  unbekannt ist, lautet das 95%-Konfidenzintervall:

$$\bar{x} \pm 2.447(s/\sqrt{n}) = 1 \pm 0.92$$

Dies ist auch das bayesianische Kreditabilitätsintervall, wenn  $f(\mu, \sigma) \propto 1/\sigma$  priori-Dichte ist.

- 2 Wenn  $\sigma = 1.5$  bekannt ist, lautet das 95%-Konfidenzintervall:

$$\bar{x} \pm 1.96(1.5/\sqrt{n}) = 1 \pm 1.11$$

Dies gleicht ebenfalls dem bayesianischen Kreditabilitätsintervall.

- 3 Wenn  $\sigma > 1.5$  bekannt ist, lässt sich kein exaktes Konfidenzintervall berechnen. Das Kreditabilitätsintervall lautet:

$$1 \pm 1.45$$

## Problematische Interpretationen des Beispiels

- Kreditibilitätsintervalle besser interpretierbar als Konfidenzintervalle, aber priori-Dichte notwendig.
- Wie kann Intervall (1) kleiner als (2) sein, obwohl bei (2) mehr Informationen vorhanden sind?
- Für  $\sigma > 1.5$  ist (2) ein zu schmales Konfidenzintervall, da es auf einer bekannten Unterschätzung von  $\sigma$  beruht.
- (3) hängt von der Wahl der Priori-Dichte ab
- Wo kommt die Angabe  $\sigma = 1.5$  in (2) her? ML-Schätzer eines vorherigen Experiments? Gäbe es stattdessen eine Posteriori-Dichte der vorherigen Experimente, so könnte man diese mit der Messung  $s = 1$  verknüpfen und so genauere Ergebnisse erhalten.

(vgl. Little, S. 3)

# Was ist bayesianische Inferenz?

## Frequentisten

- Annahme: Der unbekannte Parameter kann mit einer Statistik geschätzt werden. (Asymptotische) Eigenschaften zeigen die Güte des Schätzers (vgl. Gelman, S. 467).
- Methoden: Punktschätzer mit bewertbaren Eigenschaften, Tests, p-Werte, Konfidenzintervalle
- Wesentliches Ziel: Erklärung. Welche Parameter beeinflussen die abhängige Variable signifikant? (Dennis, S. 1097)

# Was ist bayesianische Inferenz?

## Bayesianer

- Annahme: Wahrscheinlichkeiten werden als subjektiver Glaubensgrad interpretiert (Dennis, S. 1097).
- Methoden: Aus Likelihood und Priori-Dichte wird eine Posteriori-Dichte errechnet, die für weitere Wahrscheinlichkeitsaussagen genutzt wird.
- Ziel: subjektives Wissen über Parameter datengestützt verbessern um Vorhersagen zu optimieren (Gelman, S. 467)

# Warum werden meistens frequentistische Inferenzkonzepte verwendet?

## Inhaltsübersicht

- Wesentliche Konzepte zur Datensammlung und zum Experimentdesign wurden von Nicht-Bayesianern entwickelt (Efron, S. 2).
- Frequentisten liefern leicht anwendbare und weitgehend automatisierte Verfahren. Bayesianer hingegen nicht (Efron, S. 1):
  - Wie sieht ein plausibles und berechenbares bayesianisches Modell aus?
  - Welche priori-Dichte beschreibt den Parameter adäquat?
- Frequentistische Methoden gelten als wissenschaftlich objektiv (Efron, S. 4).

## Komplizierte bayesianische Modellbildung

- Insbesondere in komplexen Problemen ist es schwierig, ein passendes Modell (Likelihood und Priori-Dichte) zu finden. Warum gerade jenes Modell und ist es vertrauenswürdig? (Little, S. 6)
- unterschiedliche Priori-Dichten liefern unterschiedliche Posteriori-Dichten. Anwender wollen aber eine klare Antwort und keine Sensitivitätsanalyse durchführen. (Little, S. 6)
- Im hierarchischen Modell lässt sich annehmen, dass die Daten nicht identisch verteilt sind, also:

$$X_i \sim F(\theta_i)$$

Modellformulierung wird schwieriger, denn welche Daten folgen identischen Verteilungen?



## Konjugierte Priori-Dichten

- In konjugierten Bayes-Analysen folgen Priori- und Posteriori-Dichten der gleichen Verteilung (siehe Beispiel am Anfang)
- Nur mit konjugierten Dichten und Mischungen derselben lässt sich rechnen. (Kadane, 456)
- Problematik: Die Priori-Dichte soll das subjektive Vorwissen abbilden. Lässt sich dies alleine mit konjugierten Formen erreichen? (Gelman, S. 470)

## Konjugierte Priori-Dichten

Problematik: Die Priori-Dichte soll das subjektive Vorwissen abbilden. Lässt sich dies alleine mit konjugierten Formen erreichen?

Unproblematisch, da

- Mithilfe von computerintensiven Verfahren lassen sich sehr flexible Priori-Dichten verwenden. Dennoch bleibt ein Spannungsfeld zwischen Realismus und Berechenbarkeit (Gelman, S. 470)
- Der Priori-Dichte kann so wenig Gewicht erhalten, dass sie kaum Einfluss auf die Posteriori-Dichte hat (Sensitivitätsanalyse).
- Da es sich ohnehin nur um ein Modell handelt, sind minimale Abweichungen von der Wirklichkeit irrelevant. (Kadane, S. 456)

## Objektive Bayesianer

- Ziel: kein Vorwissen in die Priori-Dichte aufnehmen.
- Beispiele:

$$X \sim B(1, \pi) \text{ mit } f(\pi) = 1 \text{ für } 0 \leq \pi \leq 1$$

$$X \sim N(\mu, \sigma^2) \text{ mit } f(\mu, \sigma) \propto 1/\sigma \text{ oder } f = \text{konstant?}$$

- keine Theorie vorhanden, was objektive bayesianische Analysen sind (Efron, S. 4)
- Maximum-Likelihood-Schätzer lässt sich durch Bayes-Statistik begründen (vgl. Little, S. 2):

$$f(\theta) = \text{konstant} \Leftrightarrow \max(f(\theta|x)) = \hat{\theta}_{ML}$$

ML-Schätzer nur für Teilmenge der wissenschaftlichen Analysen ohne Vorwissen geeignet

## Objektive Bayesianer

Vorteile:

- Parameter hat eine Verteilung, so dass Wahrscheinlichkeitsaussagen möglich werden
- (approximativ gleiche) Konfidenzintervalle lassen sich intuitiv interpretieren: 'Parameter liegt mit 95% Wahrscheinlichkeit im Intervall' anstatt 'In 95% der Experimente überdeckt das Intervall den wahren Parameter'

(Bernardo, S. 452)

## Subjektive Bayesianer

- Ziel: Vorwissen möglichst genau über die Priori-Dichte in die Auswertung aufnehmen. Ergebnisse werden so genauer.

### ausgedachte Priori-Dichten?

- Durch die Wahl entsprechender Priori-Dichten lassen sich Vorurteile bestätigen und unbeliebte Ergebnisse vermeiden (Gelman, S. 447 als Aprilscherz)
- (Politische) Auftraggeber können durch die Vorgabe einer Priori-Verteilung die (politisch) gewünschten Ergebnisse erhalten. Das soll Wissenschaft sein? (Dennis, S. 1099)

# Subjektive Bayesianer

## Informative Priori-Dichten

- 'Informative' Priori-Dichte kann aus vorher erhobenen Daten oder durch Kreuzvalidierung erhalten werden (Gelman, S. 469)
- Autor einer Studie muss die Modellannahmen (=Likelihood und Priori-Dichte) verteidigen. Wenn Zweifel bleiben, ist die Studie wertlos. (Kadane, S. 455)
- Wenn die Daten der Priori-Vorstellung widersprechen, muss das gesamte Modell re-evaluiert werden (Gelman, S. 469)

# Was ist Objektivität?

## Allgemeine Definition

- Subjektivität = Wahrnehmungen des Subjekts, keine Allgemeingültigkeit
- Objektivität = von Wahrnehmung unabhängiges 'Sein'

vgl. Internet Encyclopedia of Philosophy: Objectivity

# Was ist Objektivität?

## in Bezug auf Urteil/Überzeugung

- Objektive Urteile basieren auf Argumenten, die jedes rational denkende Wesen überzeugen
- Entsprechend basieren subjektive Urteile auf Argumenten, die nur einzelne Menschen überzeugen

Wenn viele subjektive Urteile übereinstimmen, wird es dann ein objektives Urteil?

vgl. Internet Encyclopedia of Philosophy: Objectivity



# Objektivität in der Statistik

Frequentisten gelten als objektiv, da sie Analysen nur auf dem Experiment aufbauen und kein Vorwissen benutzen.

## Widerspruch

- Experimentdesign hängt immer zwangsläufig stark von Vorwissen und Annahmen ab
- Statistische Modelle sind immer Vereinfachungen der Wirklichkeit und müssen daher subjektiv ausgewählt werden
- Da statistische Tests immer fehlerhafte Ergebnisse liefern können, hat man danach auch kein objektiv richtiges Ergebnis

## Objektivität in der Statistik

- Mathematische Methodik ist immer objektiv richtig oder falsch (Kadane, S. 455)
- Statistische Methoden liefern nie die objektive Wahrheit (Kadane, S. 455)

Bayesianisches Verständnis von Objektivität:

- Resultate hängen nur von Modellannahmen (Likelihood und Priori-Dichte) und Daten ab (Bernardo, S. 451)
- Jede Wissenschaft besteht bloß aus subjektiven Prozeduren, die aber objektiv getestet werden können (Gelman, S. 469)

## Falsifizierbarkeit

- Durch Tests und p-Werte können in der frequentistischen Statistik Annahmen fast sicher falsifiziert werden.
- In bayesianischen Analysen ist Falsifizierung nicht vorgesehen (Dennis, S. 1100). Die Aussage  $P(\theta < c) = 0.025$  falsifiziert nichts, da  $\theta$  kleiner als  $c$  sein kann.
- Lassen Wissenschaftler sich ihr wichtiges Werkzeug Falsifizierbarkeit nehmen?

## Wissenschaftstheorien

### Kritischer Rationalismus nach Karl Popper ("Logik der Forschung" erschien 1934)

Bei wissenschaftlichen Theorien lässt sich nie die Richtigkeit zeigen. Das Induktionsproblem ist nicht lösbar. Jedoch lassen sich Theorien durch Beobachtungen widerlegen, was den wissenschaftlichen Fortschritt ausmacht.

### Logischer Empirismus nach Rudolf Carnap ("Logical Foundations of Probability" erschien 1950)

Wissenschaftliche Ergebnisse werden mit Wahrscheinlichkeiten versehen. Aus den erhobenen Daten lässt sich der Grad der Bestätigung einer Theorie berechnen.

(vgl. Zynda)

## Wissenschaftsverständnis

- Die meiste wissenschaftliche und statistische Praxis besteht aktuell aus vielen kleinen Schritten, die mithilfe von Hypothesentests die möglichen Theorien immer weiter einschränken. Das Wissen wird so immer genauer. Notwendig sind hierfür einfache statistische Methoden, die von den Anwendungswissenschaften leicht verwendet werden können (vgl. Chernoff, S. 5)
- Gegenentwurf von S. James Press und Judith M. Tanur: Große wissenschaftliche Persönlichkeiten haben mithilfe ihrer subjektiven Vorstellungen anerkannte Theorien entwickelt. Dieser Prozess sollte durch Statistik nicht behindert werden, sondern könnte durch bayesianische Statistik beschleunigt werden (Hox).

## Verwendete Literatur

- Dennis, B (1996): "Discussion: Should Ecologists Become Bayesians?", *Ecological Applications*, Vol. 6, No. 4, pp.1095-1103
- Efron, B (1986): "Why isn't everyone a Bayesian?" *The American Statistician*, Vol. 40, No. 1  
Comments by Chernoff, Lindley, Morris, Press and Smith in the same journal were also used in my work.
- Gelman, A (2008): "Objections to Bayesian statistics", *Bayesian Analysis*, Vol. 3, No. 3, pp. 445-450  
Comments by Bernardo, Kadane, Senn, Wasserman and again Gelman in the same journal were also used in my work.
- Hox, J (2002): "Book Review: The Subjectivity of Scientists and the Bayesian Approach." *Journal of Official Statistics*, Vol. 18, No. 2

## Verwendete Literatur

- Dwayne H. Mulder (2004): "Objectivity", *Internet Encyclopedia of Philosophy*,  
<http://www.iep.utm.edu/objectiv/>
- Little, R. (2006): "Calibrated Bayes: A Bayes/Frequentist Roadmap", *The American Statistician*, Vol. 60, No. 3
- Wagner, H. (2010): "Einführung in die Bayes-Statistik",  
*Vorlesung an der LMU im WS 2010*
- Seite Wissenschaftstheorie und ähnliche. In: *Wikipedia, Die freie Enzyklopädie*. Bearbeitungsstand: 2. Dezember 2010.  
URL: <http://de.wikipedia.org/w/index.php?title=Wissenschaftstheorie&oldid=82204543>
- Zynda, L. (1994): "Carnap vs. Popper", *Lyle Zynda's Lectures on the Philosophy of Science*,  
[http://www.soc.iastate.edu/sapp/phil\\_sci\\_lecture08.html](http://www.soc.iastate.edu/sapp/phil_sci_lecture08.html)

# Fazit

## Stärken bayesianischer Methoden

- Mathematische Generalisierung des ML-Schätzers
- Lernen mithilfe vorheriger Experimente möglich
- Schätzverfahren werden flexibilisiert und Vorwissen beachtet.

## Schwächen bayesianischer Methoden

- Adäquate bayesianische Modellformulierung ist kompliziert.
- Es fehlt eine Theorie, wie man objektiv bayesianisch modelliert.
- Frequentistische Methoden haben ihre Zuverlässigkeit durch ihre zahlreichen Anwendungen bereits bewiesen - bayesianische nicht. (Dennis, S. 1100)