

Hochschule für Technik und Wirtschaft

DOKUMENTATION

Projektseminar

Optimierung und Unsicherheitsquantifizierung mit Bayesianischer Statistik und MCMC-Methoden

(Prof. Schwarzenberger)

Clemens Näther, s85426 Jakub Kliemann, s85515 Dokumentation Seite 1 von 16

Contents

1	Ein	leitung	2		
2	The 2.1	Grundlagen der bayesianischen Statistik und das Bayes'sche Theorem . 2.1.1 Einführung in die bayesianische Statistik	3 3		
	2.2 2.3 2.4	2.1.2 Das Bayes'sche Theorem und seine Bestandteile	3 4 6 7 8		
3	Pra 3.1 3.2 3.3 3.4 3.5	ktischer Teil Implementierung bayesianischer Modelle unter Verwendung in Python . Anwendung der Modelle auf verschiedene Datensätze Durchführung von MCMC-Simulationen	9 10 11 12 13		
4	Zusammenfassung und Ausblick 1				
5	Literaturverzeichnis				
6	Selbstständigkeitserklärung 1				

Dokumentation Seite 2 von 16

1 Einleitung

Dokumentation Seite 3 von 16

2 Theoretischer Teil

2.1 Grundlagen der bayesianischen Statistik und das Bayes'sche Theorem

2.1.1 Einführung in die bayesianische Statistik

Die bayesianische Statistik ist ein Zweig der Statistik. Sie unterscheidet sich im wesentlichen in der Interpretation der Wahrscheinlichkeit von der klassischen Statistik. Die klassische Statistik definiert die Wahrscheinlichkeit als die **relative Häufigkeit** in einem Zufallsexperiment [2, p. 2]. In der bayesianischen Statistik hingegen wird die Wahrscheinlichkeit als Grad des Glaubens respektiv als **Plausibilität** eines Ereignisses oder einer Aussage interpretiert [1, p. 1].

Kern der bayesianischen Statistik ist es Wissen über ein Ereignis zu verfeinern, sobald neue Informationen vorliegen. Dazu nutzt man hauptsächlich das **Bayes'sche Theorem**, welches erlaubt das Vorwissen (Prior) mit neuen Daten (Likelihood) zu kombinieren und daraus eine aktualisierte Wahrscheinlichkeit (Posterior) zu berechnen.

Mit Hilfe des Bayes'schen Theorems kann man unbekannte Parameter schätzen, ein Konfidenzintervall für diese Parameter angeben und Hypothesen prüfen. Die klassische Statistik benötigt hingegen dafür Testgrößen, weshalb die bayesianische Statistik als flexibler und intuitiver gilt. [1, p. 1].

Problem der bayesianischen Statistik ist jedoch, dass die Berechnung der Posterioriverteilung analytisch oft nicht möglich ist. Da es nun aber gute numerische Methoden wie die Markov Chain Monte Carlo (MCMC) Methoden gibt, findet die bayesianische Statistik immer mehr Anwendungen. So zum Beispiel in der Medizin oder für künstliche Intelligenzen. [2, p. 1].

2.1.2 Das Bayes'sche Theorem und seine Bestandteile

Das Bayes'sche Theorem ist ein fundamentales Konzept der bayesianischen Statistik. Es beschreibt, wie man vorhandenes Vorwissen durch neue Daten aktualisiert.

Die **Prioriverteilung** beschreibt die anfänglichen Annahmen oder das Vorwissen über einen Parameter oder ein Ereignis, bevor neue Daten berücksichtigt werden. Dabei "enthält die Priorverteilung eines Parameters θ , ausgedrückt durch f (θ), was man vor Auswertung der Stichprobe über θ weiß." [2, p. 90].

Als Priori-Wahrscheinlichkeit wird somit die Wahrscheinlichkeit P(A) bezeichnet.

Die **Posterioriverteilung** beschreibt das Wissen über einen Parameter oder ein Ereignis, nachdem alle vorhandenen Daten berücksichtigt wurden. Durch die neuen Daten, meist einer Stichprobe, wird die anfängliche Annahme, die durch die Prioriverteilung ausgedrückt wird, aktualisiert. Dies führt zu einer neuen Verteilung die widerspiegelt, wie wahrscheinlich verschiedene Werte des Parameters auf Grundlage sowohl des Vorwissens als auch der neuen Informationen sind. [2, p. 109]

Die Posteriori-Wahrscheinlichkeit wird somit als P(A|B) bezeichnet.

Dokumentation Seite 4 von 16

Die **Likelihood-Funktion** enthält die Informationen, die die Daten über den Parameter oder das Ereignis liefern. Dabei beschreibt die Likelihood die Informationen aus den neuen Daten, die zur Aktualisierung der Prioriverteilung beitragen. [2, p. 88] Die Likelihood-Wahrscheinlichkeit wird somit als P(B|A) bezeichnet.

Die Wahrscheinlichkeit P(B) wird als Normierungskonstante bezeichnet. Sie sorgt dafür, dass die Posterioriverteilung korrekt normiert ist, das heißt, dass die Summe der Wahrscheinlichkeiten aller möglichen Werte des Parameters 1 ergibt. [2, p. 109]

Das Bayes'sche Theorem lässt sich somit wie folgt darstellen:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{1}$$

Alternativ kann das Bayes-Theorem ohne die Normierungskonstante P(B) in folgender proportionaler Form dargestellt werden [3, p. 15]:

$$P(A|B) \propto P(B|A) \cdot P(A)$$
 (2)

2.1.3 Beispiele und praktische Anwendungen

Beispiel 1: m gleichgeformte Kugeln, unter denen sich k rote Kugeln und m-k schwarze Kugeln befinden. Eine Kugel wird zufällig gezogen. Die Wahrscheinlichkeit, dass die gezogene Kugel rot ist, beträgt

$$P(A) = \frac{k}{m} = p \tag{3}$$

Der Versuch wird erweitert, sodass n-mal eine Kugel mit Zurücklegen gezogen wird. Die Wahrscheinlichkeit, dass x-mal eine rote Kugel bie n-maligem Ziehen gezogen wird, beträgt

$$P(x|n,p) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \tag{4}$$

Sei nun p unbekannt. Dieses p ist nun zu schätzen. Die Binomialverteilung wird nun als Likelihood-Funktion verwendet:

$$P(n,x|p) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \tag{5}$$

wobei $0 \le p \ge 1$. Als Prioridichte wird die Gleichverteilung verwendet, da es keine Informationen über p gibt.

$$P(p) = \begin{cases} 1, & \text{für } 0 \le p \le 1\\ 0, & \text{sonst} \end{cases}$$
 (6)

Da der Term $\binom{n}{k}$ konstant ist, kann dieser bei der Berechnung der Posterioridichte vernachlässigt werden. Die Posterioridichte ergibt sich somit zu:

$$P(p|n,x) \propto p^x \cdot (1-p)^{n-x} \tag{7}$$

Dokumentation Seite 5 von 16

Vergleicht man dies mit der Dichtefunktion der Beta-Verteilung, so erkennt man dass die Posterioridichte einer Beta-Verteilung entspricht.

$$P(p|n,x) = \frac{(n+1)!}{x! \cdot (n-x)!} \cdot p^x \cdot (1-p)^{n-x} = \frac{\Gamma(n+1)}{\Gamma(x+1) \cdot \Gamma(n-x+1)} \cdot p^x \cdot (1-p)^{n-x}$$
(8)

Somit suche nach Maximum der Posterioridichte, um den Schätzer für p zu finden.

$$\frac{d}{dp}P(p|n,x) = xp^{x-1} \cdot (1-p)^{n-x} - (n-x)p^x \cdot (1-p)^{n-x-1} = 0$$
(9)

$$\Rightarrow xp^{x-1} \cdot (1-p)^{n-x} = (n-x)p^x \cdot (1-p)^{n-x-1} \tag{10}$$

Vereinfacht ergibt sich:

$$\Rightarrow x(1-p) = (n-x)p \tag{11}$$

$$\Rightarrow \frac{x}{p} - \frac{n-x}{1-p} = 0 \tag{12}$$

$$\Rightarrow p = \frac{x}{n} \tag{13}$$

Der Schätzer für p ist somit der relative Anteil der roten Kugeln an der Gesamtanzahl der Kugeln.

Beispiel 2: Beispiel 1 wird erweitert. Es wird nun eine zweite Stichprobe gezogen. Stichprobe 1: $n_1 = 10$, $x_1 = 4$, Stichprobe 2: $n_2 = 20$, $x_2 = 6$. Daten sind unabhängig voneinander. Die Daten (Posterioriverteilung) der 1. Stichprobe dienen nun als Prioridichte für die 2. Stichprobe. Man erhält somit:

$$P(p|n_1, x_1, n_2, x_2) \propto P(p|n_1, x_1) \cdot P(p|n_2, x_2)$$
(14)

Dabei ist die Prioridichte identisch zur Posterioridichte der 1. Stichprobe, siehe Gleichung (7). Die Posterioridichte der 2. Stichprobe ergibt sich somit zu:

$$P(p|n_1, x_1, n_2, x_2) \propto p^{x_1} \cdot (1-p)^{n_1-x_1} \cdot p^{x_2} \cdot (1-p)^{n_2-x_2}$$
(15)

Oder mithilfe der Beta-Verteilung:

$$P(p|n_1, x_1, n_2, x_2) = \frac{\Gamma(n_1 + n_2 + 1)}{\Gamma(x_1 + x_2 + 1) \cdot \Gamma(n_1 + n_2 - x_1 - x_2 + 1)} \cdot p^{x_1 + x_2} \cdot (1 - p)^{n_1 + n_2 - x_1 - x_2}$$
(16)

Die Daten der 1. und 2. Stichprobe könnnen somit kombiniert werden. Für die Daten $n_1 = 10$, $x_1 = 4$, $n_2 = 20$, $x_2 = 6$:

$$P(p|10, 4, 20, 6) = 931395465p^{10} \cdot (1-p)^{20}$$
(17)

Dokumentation Seite 6 von 16

2.2	Binomiale Verteilu	ng und	deren	bayesianische	Interpreta-
	tion				

Dokumentation Seite 7 von 16



Dokumentation Seite 8 von 16

2.4	Konvergenzkriterien und Diagnosewerkzeuge für MCMC-
	Simulationen

Dokumentation Seite 9 von 16

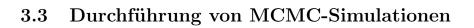
3 Praktischer Teil

 ${\bf 3.1} \quad {\bf Implementierung \ bayesian is cher \ Modelle \ unter \ Verwendung \ in \ Python}$

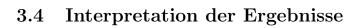
Dokumentation Seite 10 von 16



Dokumentation Seite 11 von 16



Dokumentation Seite 12 von 16



Dokumentation Seite 13 von 16



Dokumentation Seite 14 von 16



Dokumentation Seite 15 von 16

5 Literaturverzeichnis

References

[1] Karl-Rudolf Koch. Einführung in die Bayes-Statistik. Berlin [u.a.]: Springer, 2000. ISBN: 3540666702. URL: https://katalog.slub-dresden.de/id/0-306244284.

- [2] Wolfgang Tschirk. Statistik: Klassisch oder Bayes zwei Wege im Vergleich. Berlin , , © 2014. ISBN: 3642543847. URL: https://katalog.slub-dresden.de/id/0-160866449X.
- [3] Dieter Wickmann. Bayes-Statistik Einsicht gewinnen und entscheiden bei Unsicherheit. Mannheim: BI-Wiss.-Verl., 1990. ISBN: 3411146710. URL: https://katalog.slub-dresden.de/id/0-276492471.

Dokumentation Seite 16 von 16

