

Bayessche Statistik

statistik-online.ch

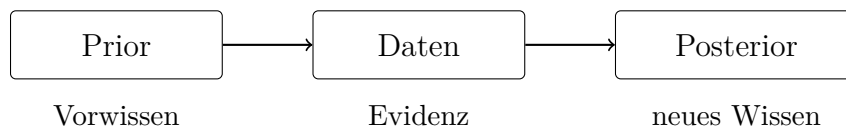
Die **bayessche Statistik** beschreibt, wie man bestehendes Wissen mit neuen Daten kombiniert. Vor der Datenerhebung gibt es eine Annahme. Nach den Daten wird diese Annahme aktualisiert.

1 Grundidee

Bayessche Statistik bedeutet: Wissen wird durch Daten aktualisiert.

Typische Fragen sind:

- Wie wahrscheinlich ist eine Hypothese nach den beobachteten Daten?
- Wie stark verändern neue Daten unsere vorherige Einschätzung?
- Wie kann Vorwissen in eine Analyse einbezogen werden?
- Welche Werte sind nach den Daten besonders plausibel?



2 Drei zentrale Begriffe

Begriff	Bedeutung
Prior	Vorwissen oder Annahme vor den Daten.
Likelihood	Wie gut passen die beobachteten Daten zu einer bestimmten Annahme?
Posterior	Aktualisierte Einschätzung nach Einbezug der Daten.

Kurz gesagt:

$$\text{Posterior} \propto \text{Likelihood} \cdot \text{Prior}$$

$\propto$ bedeutet **proportional zu**.

3 Prior

Der **Prior** $p(\theta)$ ist die Verteilung eines Parameters vor Berücksichtigung der aktuellen Daten. Er kodiert externe Information, theoretische Annahmen oder frühere Evidenz.

Prior-Typ	Bedeutung
Schwacher Prior	Breite Verteilung, wenig Information. Die Daten dominieren schneller.
Starker Prior	Konzentrierte Verteilung, viel Information. Der Posterior wird stärker stabilisiert.
Neutraler Prior	Prior mit möglichst geringer inhaltlicher Festlegung.
Fachlicher Prior	Prior auf Basis von Theorie, früheren Studien oder Expertenwissen.

4 Likelihood

Die **Likelihood** ist die Wahrscheinlichkeit der beobachteten Daten als Funktion des Parameters.

$$L(\theta) = p(\text{Daten} \mid \theta)$$

Die Daten sind dabei fix, θ variiert. Die Likelihood ist keine Wahrscheinlichkeitsverteilung über θ .

Sie beantwortet nicht die Frage:

Wie wahrscheinlich ist θ ?

sondern:

Wie plausibel sind die beobachteten Daten, wenn θ gilt?

Beispiel: Bei 14 Klicks aus 20 Beobachtungen ist $L(\theta)$ für Werte nahe $\theta = 0.70$ grösser als für $\theta = 0.10$.

5 Satz von Bayes

Der Satz von Bayes lautet:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Symbol	Bedeutung
$P(A)$	Prior: Wahrscheinlichkeit von A vor der neuen Information
$P(B \mid A)$	Likelihood: Wahrscheinlichkeit von B , falls A gilt
$P(B)$	Gesamtwahrscheinlichkeit der beobachteten Information B
$P(A \mid B)$	Posterior: Wahrscheinlichkeit von A nach Beobachtung von B

6 Normalisierung

Die Kurzform

$$\text{Posterior} \propto \text{Likelihood} \cdot \text{Prior}$$

liefert unnormalisierte Posterior-Gewichte. Die Normalisierung erfolgt durch die Evidenz $P(D)$.

Für zwei Hypothesen H_1 und H_2 :

$$P(H_1 | D) = \frac{P(D | H_1)P(H_1)}{P(D | H_1)P(H_1) + P(D | H_2)P(H_2)}$$

Der Nenner ist die Summe der unnormalisierten Gewichte. Er skaliert die Posterior-Wahrscheinlichkeiten so, dass sie zusammen 1 ergeben.

7 Beispiel: Medizinischer Test

Eine Krankheit kommt bei 1% der Personen vor. Ein Test ist positiv. Der Test erkennt kranke Personen mit 95% Wahrscheinlichkeit. Bei gesunden Personen ist der Test in 10% der Fälle fälschlicherweise positiv.

Gesucht ist:

$$P(\text{krank} | \text{positiv})$$

Gegeben:

$$P(\text{krank}) = 0.01$$

$$P(\text{positiv} | \text{krank}) = 0.95$$

$$P(\text{positiv} | \text{gesund}) = 0.10$$

Zuerst berechnen wir die Gesamtwahrscheinlichkeit für einen positiven Test:

$$\begin{aligned} P(\text{positiv}) &= P(\text{positiv} | \text{krank})P(\text{krank}) + P(\text{positiv} | \text{gesund})P(\text{gesund}) \\ &= 0.95 \cdot 0.01 + 0.10 \cdot 0.99 = 0.1085 \end{aligned}$$

Dann:

$$P(\text{krank} | \text{positiv}) = \frac{0.95 \cdot 0.01}{0.1085} = 0.0876$$

Antwort: Trotz positivem Test beträgt die Wahrscheinlichkeit für die Krankheit nur etwa 8.8%. Der Grund ist, dass die Krankheit selten ist.

8 Bayes-Faktor

Der **Bayes-Faktor** ist das Verhältnis der Evidenz der Daten unter zwei Hypothesen oder Modellen.

$$BF_{10} = \frac{P(D \mid H_1)}{P(D \mid H_0)}$$

Bei zusammengesetzten Modellen ist $P(D \mid H)$ die marginale Likelihood des Modells.

Symbol	Bedeutung
D	beobachtete Daten
H_0	Nullhypothese oder Referenzhypothese
H_1	Alternativhypothese
BF_{10}	Evidenz für H_1 relativ zu H_0

Interpretation:

Bayes-Faktor	Bedeutung
$BF_{10} = 1$	gleiche Evidenz für beide Hypothesen.
$BF_{10} > 1$	Evidenz zugunsten von H_1 .
$BF_{10} < 1$	Evidenz zugunsten von H_0 .
$BF_{10} = 5$	fünffache Evidenz für H_1 relativ zu H_0 .

Beispiel:

$$P(D \mid H_1) = 0.30, \quad P(D \mid H_0) = 0.05$$

$$BF_{10} = \frac{0.30}{0.05} = 6$$

Interpretation: Die beobachteten Daten liefern sechsmal mehr Evidenz für H_1 als für H_0 .

9 Prior Odds, Bayes-Faktor und Posterior Odds

Der Bayes-Faktor aktualisiert die **Odds** zweier Hypothesen.

$$\text{Prior Odds} = \frac{P(H_1)}{P(H_0)}$$

$$\text{Posterior Odds} = \frac{P(H_1 | D)}{P(H_0 | D)}$$

$$\text{Posterior Odds} = \text{Prior Odds} \cdot BF_{10}$$

Interpretation: Der Bayes-Faktor ist der Multiplikator, mit dem die Prior Odds in Posterior Odds umgerechnet werden.

Beispiel:

$$\text{Prior Odds} = 1$$

Das bedeutet: H_0 und H_1 haben vor den Daten gleiche Odds.

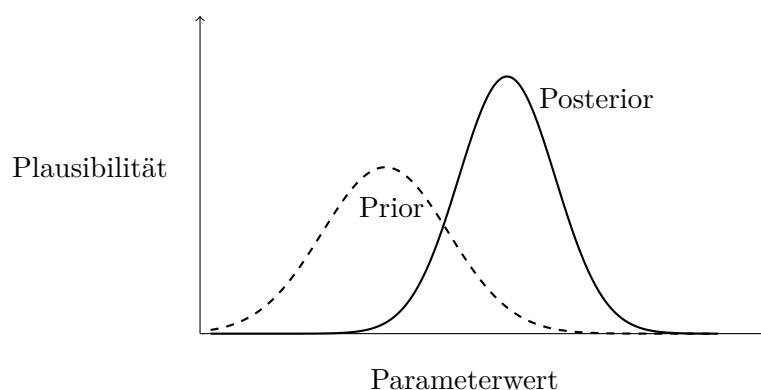
$$BF_{10} = 6$$

$$\text{Posterior Odds} = 1 \cdot 6 = 6$$

Antwort: Die Posterior Odds betragen 6 : 1 zugunsten von H_1 .

10 Bayessche Parameterschätzung

In der bayesschen Statistik werden unbekannte Parameter selbst als unsicher betrachtet. Statt nur einen einzelnen Schätzwert zu berechnen, beschreibt man eine Verteilung plausibler Parameterwerte.



$$p(\theta | \text{Daten}) = \frac{p(\text{Daten} | \theta) \cdot p(\theta)}{p(\text{Daten})}$$

Ausdruck	Bedeutung
θ	unbekannter Parameter
$p(\theta)$	Prior-Verteilung
$p(\text{Daten} \mid \theta)$	Likelihood
$p(\theta \mid \text{Daten})$	Posterior-Verteilung

11 Beispiel: Anteil schätzen

Eine Website testet eine neue Schaltfläche. Von 20 Personen klicken 14 auf die Schaltfläche. Gesucht ist der unbekannte Klickanteil θ .

Eine bayessche Modellierung trennt Prior, Daten und Posterior:

- Prior: Annahme über θ vor den Daten.
- Daten: 14 Klicks bei 20 Beobachtungen.
- Posterior: aktualisierte Verteilung von θ .

$$\hat{\theta} = \frac{14}{20} = 0.70$$

$\hat{\theta} = 0.70$ ist eine Punktschätzung. Bayessch relevant ist die Posterior-Verteilung von θ , also die Unsicherheit über den Anteil nach Beobachtung der Daten.

12 Credible Interval

Ein **Credible Interval** ist ein bayessches Intervall für einen unbekannten Parameter.

Beispiel:

$$P(0.50 \leq \theta \leq 0.86 \mid \text{Daten}) = 0.95$$

Interpretation: Nach den Daten liegt der Parameter θ mit einer Wahrscheinlichkeit von 95% zwischen 0.50 und 0.86.

Das ist anders als beim klassischen Konfidenzintervall. Dort ist der Parameter fix, und das Intervall ist zufällig. Beim Credible Interval beschreibt die Wahrscheinlichkeit direkt die Unsicherheit über den Parameter.

13 Bayessch und klassisch im Vergleich

Frage	Klassisch	Bayessch
Parameter	fix, aber unbekannt	unsicher, durch Verteilung beschrieben
Vorwissen	meist nicht direkt im Modell	als Prior ausdrückbar
Ergebnis	Schätzwert, Test, Konfidenzintervall	Posterior, Credible Interval
Interpretation	langfristige Häufigkeit	Wahrscheinlichkeit nach den Daten

14 Wichtige Hinweise

- Der Prior sollte fachlich begründet und transparent angegeben werden.
- Bei wenig Daten kann der Prior einen grossen Einfluss haben.
- Bei vielen Daten wird der Einfluss des Priors oft kleiner.
- Bayessche Ergebnisse sind besonders gut geeignet, um Unsicherheit direkt zu beschreiben.

15 Zusammenfassung

Begriff	Kurzbeschreibung
Prior	Einschätzung vor den Daten
Likelihood	Passung der Daten zu einer Annahme
Posterior	aktualisierte Einschätzung nach den Daten
Bayes-Theorem	Regel zur Aktualisierung von Wahrscheinlichkeiten
Normalisierung	skaliert den unnormalisierten Posterior zu einer Verteilung
Bayes-Faktor	Verhältnis der Evidenz zweier Hypothesen oder Modelle
Credible Interval	bayessches Intervall für plausible Parameterwerte

16 Aufgaben

Aufgabe 1: Begriffe

Ordnen Sie die Begriffe zu:

- Prior
 - Likelihood
 - Posterior
- a) Aktualisierte Einschätzung nach den Daten.
- b) Vorwissen oder Annahme vor den Daten.
- c) Wie gut passen die Daten zu einer bestimmten Annahme?

Aufgabe 2: Satz von Bayes

Gegeben:

$$P(A) = 0.20, \quad P(B \mid A) = 0.80, \quad P(B) = 0.40$$

Berechnen Sie $P(A \mid B)$.

Aufgabe 3: Medizinischer Test

Eine Krankheit kommt bei 2% der Personen vor. Ein Test erkennt kranke Personen mit 90% Wahrscheinlichkeit. Bei gesunden Personen ist der Test in 5% der Fälle falsch positiv.

Berechnen Sie $P(\text{krank} \mid \text{positiv})$.

Aufgabe 4: Credible Interval

Ein bayessches Modell ergibt:

$$P(12 \leq \mu \leq 18 \mid \text{Daten}) = 0.95$$

Interpretieren Sie dieses Intervall in Worten.

Aufgabe 5: Bayes-Faktor

Gegeben:

$$P(D \mid H_1) = 0.24, \quad P(D \mid H_0) = 0.08$$

Berechnen und interpretieren Sie BF_{10} .

Aufgabe 6: Posterior Odds

Die Prior Odds für H_1 gegenüber H_0 betragen 2. Der Bayes-Faktor beträgt 4.

Berechnen Sie die Posterior Odds.

Aufgabe 7: Multiple Choice

Kreuzen Sie alle richtigen Aussagen an.

- ☐ Der Posterior kombiniert Prior und Daten.
- ☐ Der Prior ist immer objektiv und hat nie Einfluss auf das Ergebnis.
- ☐ Ein Credible Interval kann direkt als Wahrscheinlichkeit für den Parameter interpretiert werden.
- ☐ Bei sehr vielen Daten wird der Einfluss des Priors oft kleiner.
- ☐ $BF_{10} = 3$ bedeutet dreimal mehr Evidenz für H_1 als für H_0 .

17 Lösungen

Lösung 1

- Prior: b
- Likelihood: c
- Posterior: a

Lösung 2

$$\begin{aligned} P(A | B) &= \frac{P(B | A) \cdot P(A)}{P(B)} \\ &= \frac{0.80 \cdot 0.20}{0.40} = \frac{0.16}{0.40} = 0.40 \end{aligned}$$

Antwort: $P(A | B) = 0.40$.

Lösung 3

$$P(\text{krank}) = 0.02, \quad P(\text{gesund}) = 0.98$$

$$P(\text{positiv}) = 0.90 \cdot 0.02 + 0.05 \cdot 0.98$$

$$= 0.018 + 0.049 = 0.067$$

$$P(\text{krank} | \text{positiv}) = \frac{0.90 \cdot 0.02}{0.067} = \frac{0.018}{0.067} = 0.269$$

Antwort: Die Wahrscheinlichkeit beträgt etwa 26.9%.

Lösung 4

Nach den Daten liegt μ mit einer Wahrscheinlichkeit von 95% zwischen 12 und 18.

Lösung 5

$$BF_{10} = \frac{P(D | H_1)}{P(D | H_0)} = \frac{0.24}{0.08} = 3$$

Antwort: Die Daten liefern dreimal mehr Evidenz für H_1 als für H_0 .

Lösung 6

$$\text{Posterior Odds} = \text{Prior Odds} \cdot BF_{10} = 2 \cdot 4 = 8$$

Antwort: Die Posterior Odds betragen 8 : 1 zugunsten von H_1 .

Lösung 7

- ✓ Richtig – Der Posterior kombiniert Prior und Daten.
- ✗ Falsch – Der Prior kann Einfluss auf das Ergebnis haben.
- ✓ Richtig – Ein Credible Interval erlaubt diese direkte Interpretation.
- ✓ Richtig – Bei vielen Daten dominiert häufig die Information aus den Daten.
- ✓ Richtig – $BF_{10} = 3$ bedeutet dreimal mehr Evidenz für H_1 als für H_0 .