
Beurteilung von Hypothesen in drei statistischen Paradigmata II

Dr. Dirk Hasenclever
IMISE, Leipzig

Hasenclever@IMISE.uni-Leipzig.de

Definition der Likelihood-Funktion

Gegeben Beobachtungen $y=(y_1,\dots,y_n)$
von Zufallsvariablen $Y=(Y_1,\dots,Y_n)$

Angenommen die Dichte von Y gehöre zu einer Familie von Verteilungsdichten $f(\theta,y)dy$ mit $\theta \in \Omega$ (Ω Parameterraum) mit unbekanntem Parameter θ .

Likelihood(θ , y)= $f_Y(\theta, y)$

Falls y_i unabhängig: Likelihood(θ , y)= $\prod_i f(\theta, y_i)$

Likelihood-Funktion im frequentistischen Paradigma

Maximum Likelihood Prinzip:

Der Parameter θ , für den die Daten am wahrscheinlichsten erscheinen (kontinuierlicher Fall: für den die Likelihood der Daten maximal ist) ist am plausibelsten.

Maximum Likelihood Schätzer

(Schätze θ durch Maximierung der Likelihoodfunktion) sind in der Regel asymptotisch normal verteilt.

Likelihood-Funktion im frequentistischen Paradigma

Test von Punkthypothesen:

Lemma von Neyman-Pearson:

Bester Ablehnungsbereich beim Test von Punkthypothesen wird durch Likelihood ratio $> c$ definiert.

(Maximale Power bei gegebener „Irrtumswahrscheinlichkeit“)

Law of Likelihood

Law of likelihood: If hypothesis A implies that the probability that a random variable X takes the value x is $p_A(x)$, while hypothesis B implies that the probability is $p_B(x)$, then the observation $X = x$ is evidence supporting A over B if and only if $p_A(x) > p_B(x)$, and the likelihood ratio, $p_A(x)/p_B(x)$, measures the strength of that evidence (Hacking, 1965).

Hypothesen-Vergleich im Likelihoodparadigma

- Die relative Evidenz (Plausibilität) von Punkthypothesen ($\theta = \theta_0$) wird durch die Likelihoodratio gemessen:

$$LR(\theta_1 \text{ versus } \theta_0) = \frac{pr(D | \theta_1)}{pr(D | \theta_0)}, \quad (D = \text{Daten})$$

- LRs können direkt interpretiert werden, als „das, was die Daten relativ θ_0 und θ_1 sagen“.

Kalibrierungsexperiment für Likelihoodratios I

- Wie bewertet man LRs?
- Gedankenexperiment:
 - ▶ Zwei äußerlich identische Urnen, eine enthält sehr viele nur weiße Kugeln, die zweite gleich viele aber 50% schwarze.
 - ▶ Wähle eine Urne zufällig aus und ziehe Kugeln und bewerte Evidenz zwischen den Hypothesen: „Alle weiß“ gegen „50:50“
 - ▶ (w, w, w) ist ziemlich deutliche Evidenz für „Alle weiß“
 - ▶ (w, w, w, w, w) ist starke Evidenz für „Alle weiß“

Kalibrierungsexperiment für Likelihoodratios II

- Es gilt:

- $LR(\text{„Alle weiß“ gegen „50:50“} | b \text{ wei\ss e Kugeln hintereinander})$
 $= 1 / 0.5^b = 2^b$

- Also lässt sich eine gegebene LR interpretieren als $b = \log_2(LR)$ weiße hintereinander Kugeln hintereinander im Kalibrierungsexperiment.

- > Intuitives Verständnis der LR

- Z.B.:

LR	4	8	10	16	20	32	50	100	1000
b	2	3	3,3	4	4,3	5	5,6	6,6	10

Hypothesen-Vergleich im Likelihoodparadigma II

■ Limitationen

- ▶ Likelihoodratios zwischen Mengen von Parametern (Hypothesen) sind nicht sinnvoll definierbar.
- ▶ Oft interessiert nur einen Komponente von $\theta=(\lambda,\phi)$.
 ϕ ist eine Störgröße („nuisance parameter“)
 - Es gibt verschiedenen Hilfskonstruktionen um Störparameter zu eliminieren:
z.B.: geschätzte Likelihood: ersetze ϕ durch den Maximum-Likelihoodschätzer $\hat{\phi}$ und analysiere die Likelihood als Funktion von λ allein.

Bayesianischer Hypothesenvergleich I

- Bayesianer ersetzen die Likelihoodratio durch den

Bayes-Faktor = posterior probability ratio

$$\text{BF}(\theta_1 \text{ versus } \theta_0) = \frac{\text{pr}(\theta_1 \mid D; \text{prior})}{\text{pr}(\theta_0 \mid D; \text{prior})}, \quad (D = \text{Daten})$$

- Zusammenhang mit Likelihood-Ratio:
posteriorOdds = LR * Prior Odds

Bayesianischer Hypothesenvergleich II

- Damit werden **auch Mengen von Hypothesen vergleichbar**:

Prior $\text{pr}(\theta)$ auf Parameterraum $\theta \in \Omega$.

Likelihoodfunktion : $\text{pr}(D | \theta)$ (D Daten)

$$\text{Posterior(dichte)} : \text{pr}(\theta | D; \text{prior}) = \frac{\text{pr}(D | \theta) * \text{pr}(\theta)}{\int_{\Omega} \text{pr}(D | \theta') * \text{pr}(\theta') d\theta'}$$

Posteriorwahrscheinlichkeit für Menge von Parametern $\Omega_1 \subset \Omega$:

$$\text{pr}(\theta \in \Omega_1 | D; \text{prior}) = \int_{\Omega_1} \text{pr}(\theta | D; \text{prior}) = \frac{\int_{\Omega_1} \text{pr}(D | \theta) * \text{pr}(\theta) d\theta}{\int_{\Omega} \text{pr}(D | \theta') * \text{pr}(\theta') d\theta'}$$