Bayesian updating

Frage: wie lassen sich die neuen Informationen in den Entscheidungsprozess integrieren?

Θ - states / Zustände $\quad$ D - decisions / Entscheidungen

$\mathcal{U} : \Theta \times D \to \mathbb{R}$ - Nutzenfunktion gegeben

$\theta \in \Theta$, $d \in D$

Zustand der Natur $\quad$ Entscheidung des Subjekts

Annahme: • $\pi : \Theta \to \mathbb{R}_+$

prior gegeben $\quad$ Wahrscheinlichkeitsdichte auf Zuständen, d.h. $\int_A \pi(\theta)\, d\theta$ Wahrscheinlichkeits-maß auf Θ.

$$\delta \in \arg\max_{d \in D} \int_\Theta \mathcal{U}(\theta, d) \cdot \pi(\theta)\, d\theta$$

"Maximierung des erwarteten Nutzens"

• X - Erfahrungen / Observations

Bayes estimator

$$\delta(x) \in \arg\max_{d \in D} \int_\Theta \mathcal{U}(\theta, d) \cdot \pi(\theta | x)\, d\theta$$

Bayes decision rule

$\delta : X \to D$

posterior $\quad$ Wahrscheinlichkeitsdichte auf Zuständen gegeben Erfahrung x

Bayesformel:

likelihood $\quad$ prior

Diskret: $$\underbrace{P(A|E)}_{\text{posterior}} = \frac{P(E|A) \cdot P(A)}{\underbrace{P(E)}_{\text{evidence}}}$$

W'dichte der Erfahrung x gegeben Zustand θ

Stetig: $$\pi(\theta | x) = \frac{f(x|\theta) \cdot \pi(\theta)}{\int_\Theta f(x|\theta) \cdot \pi(\theta)\, d\theta}$$

W'dichte der Erfahrung x

$$\int_\Theta f(x|\theta) \cdot \pi(\theta)\, d\theta \overset{!!}{=} m(x)$$

Utilities

$$\rho(d,x) := \int_\Theta \mathcal{U}(\theta,d)\,\pi(\theta|x)\,d\theta$$

posterior expected utility for $d \in D$, $x \in X$

$$R(\theta,\delta) := \int_X \mathcal{U}(\theta,\delta(x))\cdot f(x|\theta)\,dx$$

frequentist utility for $\theta \in \Theta$, $\delta: X \to D$

"average utility from δ gegeben state θ"

$$r(\delta) := E_\theta[R(\theta,\delta)] = \int_\Theta \Big[\int_X \mathcal{U}(\theta,\delta(x))\,f(x|\theta)\,dx\Big]\,\pi(\theta)\,d\theta$$

integrated utility

$$\underset{\text{Fubini}}{=} \int_X \Big[\int_\Theta \mathcal{U}(\theta,\delta(x))\,\underbrace{f(x|\theta)\cdot\pi(\theta)}_{\underset{\text{Bayes}}{=}\,\pi(\theta|x)\cdot m(x)}\,d\theta\Big]\,dx =$$

$$= \int_X \Big[\underbrace{\int_\Theta \mathcal{U}(\theta,\delta(x))\cdot\pi(\theta|x)\,d\theta}_{=\rho(\delta(x),x)}\Big]\,m(x)\,dx$$

$$\Rightarrow \quad r(\delta) = \int_X \rho(\delta(x),x)\cdot m(x)\,dx$$

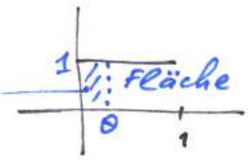
Also: $\boxed{\delta(x) \in \underset{\delta: X\to D}{\operatorname{argmax}}\ r(\delta)}$ Bayes estimator den integrierten Nutzen

Interpretation von Bayes estimator

Entscheidung:		Ansatz	Gewinn
$d_1 \nearrow$	Spielen "Wurf 0/1"	κ_1	$K_1\cdot\theta$
$d_2 \searrow$	auf z.B. 0	κ_2	$K_2\cdot\theta$

$\Leftarrow$ unbekannte Gewinnwahrscheinlichkeit

$$\underbrace{\mathcal{U}(\theta,d_1) = K_1\cdot\theta - \kappa_1,\quad \mathcal{U}(\theta,d_2) = K_2\cdot\theta - \kappa_2}_{\text{erwarteter Nutzen}}$$



Prior: $\pi(\theta) = 1\!\!1_{[0,1]}(\theta)$, d.h. $P(\Theta \le \theta) = \int_{-\infty}^{\theta} 1\!\!1_{[0,1]}(\theta)\,d\theta = \theta$

$$\int_\Theta \underbrace{\mathcal{U}(\theta,d_i)\cdot\pi(\theta)}_{\text{expected utility}}\,d\theta = \int_0^1 (K_i\cdot\theta - \kappa_i)\,d\theta = K_i\cdot\frac{\theta^2}{2} - \kappa_i\theta\,\Big|_0^1 = \frac{K_i}{2} - \kappa_i$$

$$d_1 \succ d_2 \iff \frac{K_1}{2} - \kappa_1 > \frac{K_2}{2} - \kappa_2$$

Entscheidungskriterium, vgl. $E(K_i\cdot\Theta - \kappa_i) = K_i\,E(\Theta) - \kappa_i = K_i\cdot\frac{1}{2} - \kappa_i$

Erfahrung: $f(x|\theta) \sim Bin(n, \theta)$ - Binomialverteilung mit n-Versuchen und Gewinnwahrscheinlichkeit θ.

d.h. $\mathbb{P}(\underbrace{X=x}_{}|\Theta=\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$.

Anzahl von "o"-Werten im Versuch

$$\underset{\overset{\shortparallel}{\mathbb{P}(X=x)}}{m(x)} = \int_\Theta f(x|\theta)\cdot\pi(\theta)\,d\theta = \int_0^1 \binom{n}{x}\theta^x(1-\theta)^{n-x}\,d\theta = \binom{n}{x}B(x+1, n-x+1).$$

Gamma-Funktion

$$\left[\ B(a,b) = \int_0^1 t^{a-1}\cdot(1-t)^{b-1}dt = \frac{\Gamma(a)\cdot\Gamma(b)}{\Gamma(a+b)},\quad \Gamma(n) := (n-1)!\ \right]$$

Beta-Funktion

Posterior

$$\pi(\theta|x) = \frac{f(x|\theta)\cdot\pi(\theta)}{\int_\Theta f(x|\theta)\pi(\theta)d\theta} = \frac{\cancel{\binom{n}{x}}\theta^x(1-\theta)^{n-x}}{\cancel{\binom{n}{x}}B(x+1, n-x+1)}\mathbb{1}_{[0,1]}(\theta)$$

posterior expected utility

$$\rho(d_i, x) = \int_\Theta \mathcal{U}(\theta, d_i)\cdot\pi(\theta|x)\,d\theta = \int_0^1 (K_i\cdot\theta - \kappa_i)\cdot\frac{\theta^x(1-\theta)^{n-x}}{B(x+1, n-x+1)}\,d\theta =$$

$$\underset{s.o.}{=} \frac{K_i}{B(x+1,n-x+1)}\cdot\underbrace{\int_0^1 \theta^{x+1}(1-\theta)^{n-x}d\theta}_{B(x+2,\ n-x+1)} - \kappa_i = K_i\cdot\frac{B(x+2, n-x+1)}{B(x+1, n-x+1)} - \kappa_i$$

$$= \left[\frac{\Gamma(x+2)\cdot\cancel{\Gamma(n-x+1)}}{\Gamma(n+3)}\cdot\frac{\Gamma(n+2)}{\Gamma(x+1)\cancel{\Gamma(n-x+1)}} = \frac{(x+1)!\cdot(n+1)!}{(n+2)!\ \ x!} = \frac{x+1}{n+2}\right] =$$

$$= K_i\cdot\frac{x+1}{n+2} - \kappa_i.$$

$$\Rightarrow \quad \underset{\text{Bayes decision}}{\delta(x)} = \begin{cases} d_1, & \text{falls}\quad K_1\frac{x+1}{n+2} - \kappa_1 > K_2\frac{x+1}{n+2} - \kappa_2 \\ d_2, & \text{sonst}\end{cases}$$

o.B.d.A. $\frac{K_1}{2} - \kappa_1 > \frac{K_2}{2} - \kappa_2 \Rightarrow d_1$ - optimal

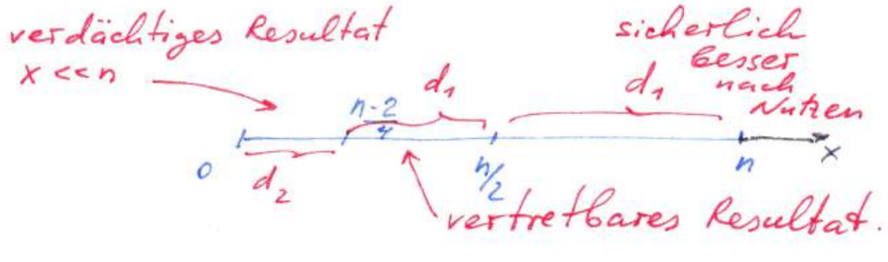
$\Leftrightarrow K_1 - K_2 > 2(\kappa_1 - \kappa_2) \underset{x \geq \frac{n}{2}\text{-gerade}}{\Rightarrow} K_1 - K_2 > \frac{n+2}{x+1}(\kappa_1 - \kappa_2) \Rightarrow \delta(x) = d_1$ für $x \geq \frac{n}{2}$

z.B. $K_1 = 10,\ \kappa_1 = 3\ ;\ K_2 = 6,\ \kappa_2 = 2.$

$\frac{10}{2} - 3 = 2 > \frac{6}{2} - 2 = 1 \Rightarrow d_1$ - optimal

$$K_1\frac{x+1}{n+2} - \kappa_1 > K_2\frac{x+1}{n+2} - \kappa_2 \Leftrightarrow 10\cdot\frac{x+1}{n+2} - 3 > 6\frac{x+1}{n+2} - 2 \Leftrightarrow \begin{cases} x \geq \frac{n-2}{4}: \delta(x) = d_1 & \text{optimal} \\ x < \frac{n-2}{4}: \delta(x) = d_2 \end{cases}$$

"decision switch"



Minimax

$$\delta_0 \in \arg\max_{\delta: X \to D} \underbrace{\inf_{\theta \in \Theta} R(\theta, \delta)}_{\text{worst-case}} = \int_X u(\theta, \delta(x)) \cdot f(x|\theta)\, dx$$ frequentist utility

minimax risk / estimator

Zusammenhang mit Bayes estimator:

$$r(\delta) = \mathbb{E}_\theta [R(\theta, \delta)] \geq \mathbb{E}_\theta \left[\inf_{\theta \in \Theta} R(\theta, \delta)\right] = \inf_{\theta \in \Theta} R(\theta, \delta)$$

$$\Rightarrow \underbrace{\max_{\delta: X \to D} r(\delta)}_{\text{Nutzen des Bayes estimators}} \geq \underbrace{\max_{\delta: X \to D} \inf_{\theta \in \Theta} R(\theta, \delta)}_{\text{Nutzen des minimax estimator}}$$ "pessimistischer Ansatz"

Admissibility

$\delta_0 : X \to D$ heißt admissible estimator, falls kein $\delta_1 : X \to D$ existiert mit $R(\theta, \delta_1) \geq R(\theta, \delta_0)\ \forall \theta$ "δ_1 dominiert δ_0"
und $R(\theta_0, \delta_1) > R(\theta_0, \delta_0)$ für mindestens ein $\theta_0 \in \Theta$.

Falls $\pi(\theta) > 0\ \forall \theta \in \Theta$, dann ist Bayes estimator $\delta: X \to D$ admissible.
und $R(\cdot, \delta)$ stetig

Ansonsten $\exists\, \delta_1 : X \to D$ mit $R(\theta, \delta_1) \geq R(\theta, \delta)\ \forall \theta$

$R(\theta_0, \delta_1) > R(\theta_0, \delta)$ für $\theta_0 \in \Theta$

R-stetig $\hookrightarrow R(\theta, \delta_1) > R(\theta, \delta)\ \forall \theta \in C \subseteq \Theta$ offen

Also: $r(\delta_1) = \underbrace{\mathbb{E}_\theta [R(\theta, \delta_1)] > \mathbb{E}_\theta [R(\theta, \delta)]}_{\geq}$

$$= \int_\Theta \underbrace{R(\theta, \delta)}_{< R(\theta, \delta_1) \text{ auf } C} \underbrace{\pi(\theta)}_{>0}\, d\theta$$

$$= \max_{\delta: X \to D} r(\delta) \quad \lightning$$