Bayes-Schätzer, Conjugate Priors

M.Gruber

SS 2009, KW 22

Bayes-Schätzung

- Verteilungsfamilie $\{f(. \mid \theta) \mid \theta \in \Theta\}$ bekannt, wahrer Parameter θ_0 unbekannt,
- A-Priori-Verteilung des wahren Parameters $f(\theta)$ intuitiv bekannt;
- Stichprobe, Experiment: $X_1, X_2, \ldots, X_n$, unabhängig, verteilt nach $f(. \mid \theta_0)$,
- A-Posteriori-Verteilung des wahren Parameters:

$$f(\theta \mid x_1, x_2, \ldots, x_n) = \frac{f(x_1, x_2, \ldots, x_n \mid \theta) f(\theta)}{\int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n \mid \theta) f(\theta)\, d\theta}.$$

Wechsel von A-Priori- zu A-Posteriori-Verteilung auf Grund von Stichprobendaten.

Beta-Verteilungen

Gamma-Funktion:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt \quad (x > 0)\,, \quad \Gamma(x+1) = x\Gamma(x)\,, \quad \Gamma(n) = (n-1)!$$

Betafunktion:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\, du \qquad (\alpha, \beta > 0)\,.$$

Betaverteilung:

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad (0 \le x \le 1)\,, \mathbf{E}X = \frac{\alpha}{\alpha+\beta}\,, \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}\,.$$

Dichten von Betaverteilungen

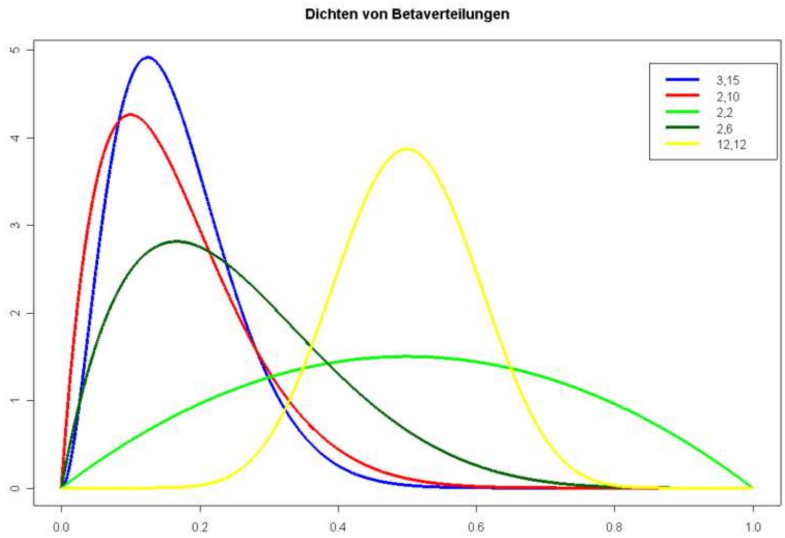

Abbildung 1: Quelle: `http://de.wikipedia.org/wiki/Betaverteilung`

Beispiel 1. **[conjugate prior für $X \sim \mathcal{B}(1,p)$]** *A-Priori-Wissen:* $\mathbf{E}p = 0.4$, $\mathrm{Var}(p) = 0.1$. *Die A-Priori-Verteilung f von p wird als Betaverteilung angenommen.*

α, β *passen wir an:* $\frac{\alpha}{\alpha+\beta} = 0.4$, $\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} = 0.1 \Rightarrow \alpha = \frac{14}{25}, \beta = \frac{21}{25}$.

$$\begin{aligned} f(p \mid x_1, \dots, x_n) &= \frac{f(x_1, \dots, x_n \mid p) f(p)}{\int_0^1 f(x_1, \dots, x_n \mid p) f(p)\, dp} \\ &= \frac{p^{s_n}(1-p)^{n-s_n} p^{\alpha-1}(1-p)^{\beta-1}}{\int_0^1 p^{s_n}(1-p)^{n-s_n} p^{\alpha-1}(1-p)^{\beta-1}\, dp} \\ &= \frac{\Gamma(\alpha + s_n + \beta + n - s_n)}{\Gamma(\alpha + s_n)\Gamma(\beta + n - s_n)} p^{(\alpha+s_n)-1}(1-p)^{(\beta+n-s_n)-1} \end{aligned}$$

A-Posteriori-Verteilung ist eine Beta-Verteilung mit $\alpha' = \alpha + s_n, \beta' = \beta + n - s_n$.

Für $n = 10$ und $s_n = 9$ ist z.B. $\alpha' = \alpha + 9$, $\beta' = \beta + 1$.

Bayes-Schätzer

Definition. *Bayes-Schätzer $\theta(X_1, \dots, X_n)$ eines unbekannten Parameters θ_0 ist der Erwartungswert der A-Posteriori-Verteilung.*

Beispiel. *(Fortsetzung von Beispiel 1)*

$$p(x_1, \dots, x_n) = \frac{\alpha + s_n}{\alpha + s_n + \beta + n - s_n} = \frac{\alpha/n + \bar{x}}{\alpha/n + \beta/n + 1} \longrightarrow \bar{x}$$

Für $n \to \infty$ schwindet der Einfluss des A-Priori-Wissens.

Gamma-Verteilungen

- Dichtefunktion:

$$f(x \mid k, \theta) = x^{k-1} \frac{e^{-x/\theta}}{\theta^k \, \Gamma(k)} \quad \text{for} \quad x > 0$$

mit Formparameter $k > 0$ Skalierungsparameter $\theta > 0$.

- Alternative Darstellung mit Formparameter $\alpha = k$ und inversem Skalierungsparameter $\beta = 1/\theta$:

$$g(x \mid \alpha, \beta) = x^{\alpha-1} \frac{\beta^\alpha \, e^{-\beta \, x}}{\Gamma(\alpha)} \quad \text{for} \quad x > \theta$$

- Erwartungswert $k\theta$ bzw. α/β, Varianz $k\theta^2$ bzw. α/β^2.

Dichten von Gammaverteilungen ($\alpha = p, \beta = \frac{1}{b}$)

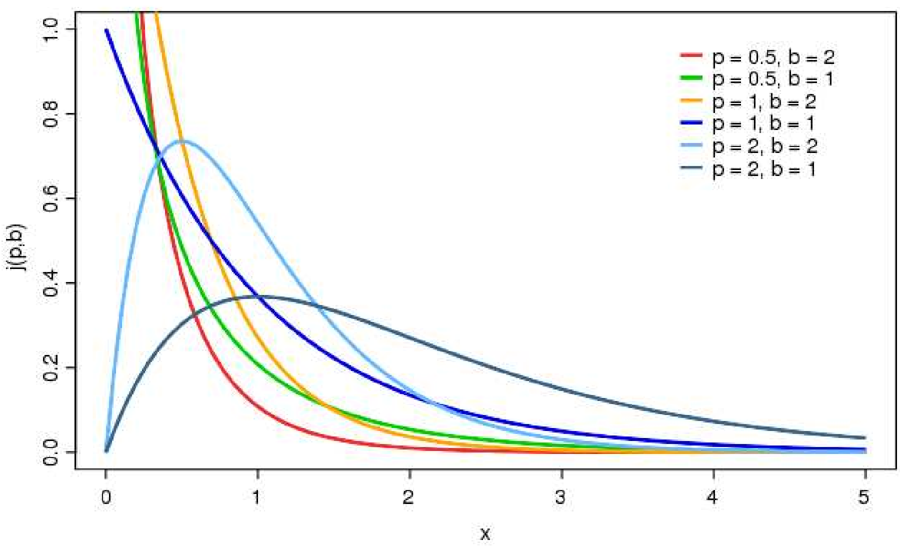

Abbildung 2: Quelle: `http://de.wikipedia.org/wiki/Gammaverteilung`

Beispiel 2. [conjugate prior für $X \sim \Pi(\lambda)$]

$$f(x \mid \lambda) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$$

$$f(x_1, \dots, x_n \mid \lambda) = \prod_{1 \le i \le n} \left(\frac{\lambda^{x_i}}{x_i!} \cdot e^{-\lambda} \right) = \frac{\lambda^{\sum x_i}}{\prod x_i!} \cdot e^{-n\lambda}$$

Für beliebige A-Priori-Verteilung f erhalten wir als A-Posteriori-Verteilung

$$f(\lambda \mid x_1, \dots, x_n) = \frac{\frac{\lambda^{\sum x_i}}{\prod x_i!} \cdot e^{-n\lambda} f(\lambda)}{\int \frac{\lambda^{\sum x_i}}{\prod x_i!} \cdot e^{-n\lambda} f(\lambda)\, d\lambda} \sim \lambda^{\sum x_i} \cdot e^{-n\lambda} f(\lambda)$$

Für die Gammaverteilung $g(x \mid \alpha, \beta)$ als A-Priori-Verteilung mit

$$g(\lambda \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

erhalten wir

$$\begin{aligned} f(\lambda \mid x_1, \dots, x_n) &\sim \lambda^{\sum x_i} e^{-n\lambda} g(\lambda \mid \alpha, \beta) \\ &\sim \lambda^{\sum x_i} e^{-n\lambda} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \\ &\sim \lambda^{\sum x_i + \alpha - 1} e^{-(n+\beta)\lambda} \end{aligned}$$

und damit wieder eine Gammaverteilung: $g(x \mid \alpha + \sum x_i, \beta + n)$.

Bayes-Schätzer $\lambda(x_1, \dots, x_n)$ ist somit $\frac{\alpha + \sum x_i}{n + \beta}$.