Applications of Martingale Methods
Karlin & Taylor, *A First Course in Stochastic Processes, pp 297–325*

BAYESIAN ESTIMATION

Let $\theta \in \Theta$ be an unknown parameter governing the probability distribution of IID observable random variables X_i. If the distribution of each X_i has a density function $f(x|\theta)$, the joint density for n observations is the so-called *likelihood function*:

$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta).$$

Some statisticians would recommend "estimating" θ by that value $\hat{\theta}_n$ that maximizes the likelihood or, equivalently, its logarithm $\ell(\theta) = \sum_{i=1}^{n} \log\big(f(X_i|\theta)\big)$; others, notably Bayesians, would describe uncertainty about θ using a probability distribution $\pi(d\theta)$ on a suitable BF of events in Θ, and would describe uncertainty about θ *after* seeing the data $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ using the *posterior distribution*,

$$\pi(d\theta|\mathcal{X}_n) \;=\; \frac{\pi(d\theta)\prod_{i=1}^{n} f(X_i|\theta)}{\int_\Theta \pi(d\dot\theta)\prod_{i=1}^{n} f(X_i|\dot\theta)} \;=\; c\,\pi(d\theta)\,L(\theta)$$

or, if a point estimate is required, its mean

$$\begin{aligned}
\bar\theta_n &= \mathsf{E}[\theta|\mathcal{X}_n] \\
&= \frac{\int_\Theta \theta\pi(d\theta)\prod_{i=1}^{n} f(X_i|\theta)}{\int_\Theta \pi(d\dot\theta)\prod_{i=1}^{n} f(X_i|\dot\theta)} \\
&= c\int_\Theta \theta\,\pi(d\theta)\,L(\theta)
\end{aligned}$$

Quite generally this estimate $\bar\theta_n$ constitutes a martingale:

$$\begin{aligned}
\mathsf{E}\big[\bar\theta_{n+1} \mid \mathcal{X}_n\big] &= \int_{\mathcal{X}} \Big[\bar\theta_{n+1}\Big]\Big[f(X_{n+1} \mid \mathcal{X}_n)\Big]\,dX_n \\
&= \int_{\mathcal{X}} \Big[\bar\theta_{n+1}\Big]\Big[\int_\Theta f(X_{n+1} \mid \theta)\,\pi(d\theta|\mathcal{X}_n)\Big]\,dX_n \\
&= \int_{\mathcal{X}} \Big[\frac{\int_\Theta \theta\pi(d\theta)\prod_{i=1}^{n+1} f(X_i|\theta)}{\int_\Theta \pi(d\dot\theta)\prod_{i=1}^{n+1} f(X_i|\dot\theta)}\Big]\Big[\int_\Theta f(X_{n+1} \mid \dot\theta)\,\frac{\pi(d\dot\theta)\prod_{i=1}^{n} f(X_i|\dot\theta)}{\int_\Theta \pi(d\ddot\theta)\prod_{i=1}^{n} f(X_i|\ddot\theta)}\Big]\,dX_n \\
&= \int_{\mathcal{X}} \Big[\frac{\int_\Theta \theta\pi(d\theta)\prod_{i=1}^{n+1} f(X_i|\theta)}{1}\Big]\Big[\frac{1}{\int_\Theta \pi(d\ddot\theta)\prod_{i=1}^{n} f(X_i|\ddot\theta)}\Big]\,dX_n \\
&= \Big[\frac{\int_\Theta \theta\pi(d\theta)\prod_{i=1}^{n} f(X_i|\theta)}{1}\Big]\Big[\frac{1}{\int_\Theta \pi(d\ddot\theta)\prod_{i=1}^{n} f(X_i|\ddot\theta)}\Big] \\
&= \Big[\frac{\int_\Theta \theta\pi(d\theta)\prod_{i=1}^{n} f(X_i|\theta)}{\int_\Theta \pi(d\ddot\theta)\prod_{i=1}^{n} f(X_i|\ddot\theta)}\Big] \\
&= \bar\theta_n
\end{aligned}$$

If we can verify the conditions for the Martingale Convergence Theorem, *e.g.*, $\mathsf{E}[|\bar\theta_n|] \le M$ for some $M < \infty$ and all n, and possibly that the $\bar\theta_n$ are Uniformly Integrable, then we will know that there exists a random variable $\bar\theta_\infty$ such that $\bar\theta_n \to \bar\theta_\infty$ *a.s.* as $n \to \infty$, *i.e.*, that the Bayes estimates converge. The estimates are called *consistent* if $\bar\theta_n \to \theta$ *a.s.*, *i.e.*, if the $\bar\theta_n$ converge and moreover $\bar\theta_\infty$ is *a.s.* equal to the constant θ.