MA5208: Introduction to Bayesian Analysis

Satvik Saha, `19MS154` May 15, 2024

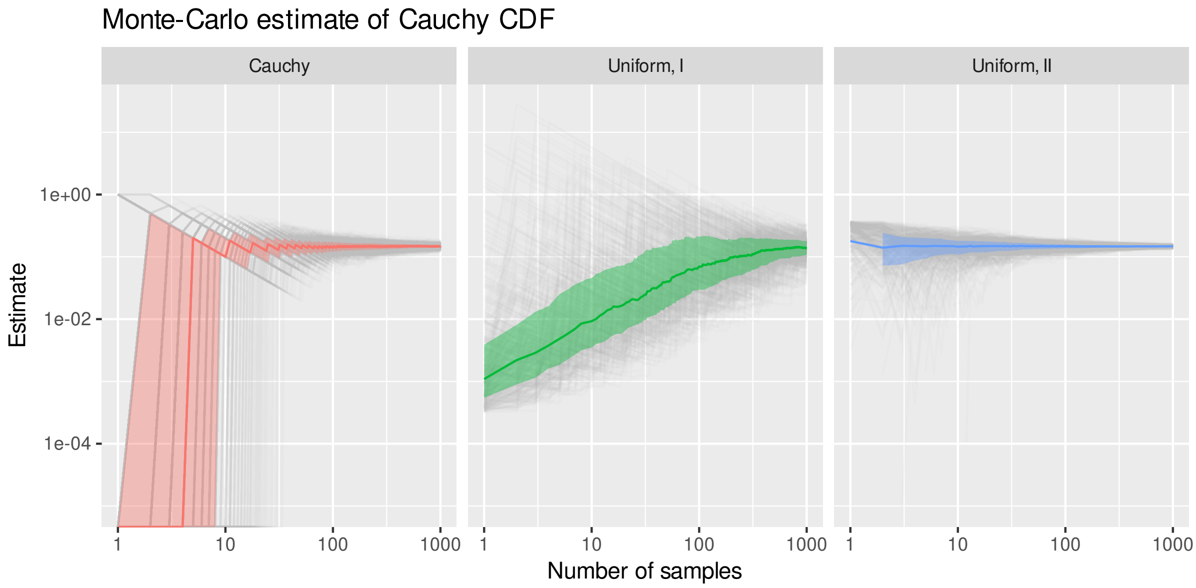
Problem 1

We estimate
$$I = \int_2^\infty \frac{1}{\pi(1+x^2)}\,dx$$
via Monte-Carlo methods. First, note that this is simply $\mathbb{E}[\mathbf{1}_{[2,\infty)}(X)]$ where $X \sim \text{Cauchy}(0,1)$. Thus, we sample $X_1, \dots, X_n \overset{\text{iid}}{\sim} \text{Cauchy}(0,1)$, and estimate
$$\hat{I}_{\text{Cauchy}} \;=\; \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{[2,\infty)}(X_i) \;\xrightarrow{p}\; I.$$

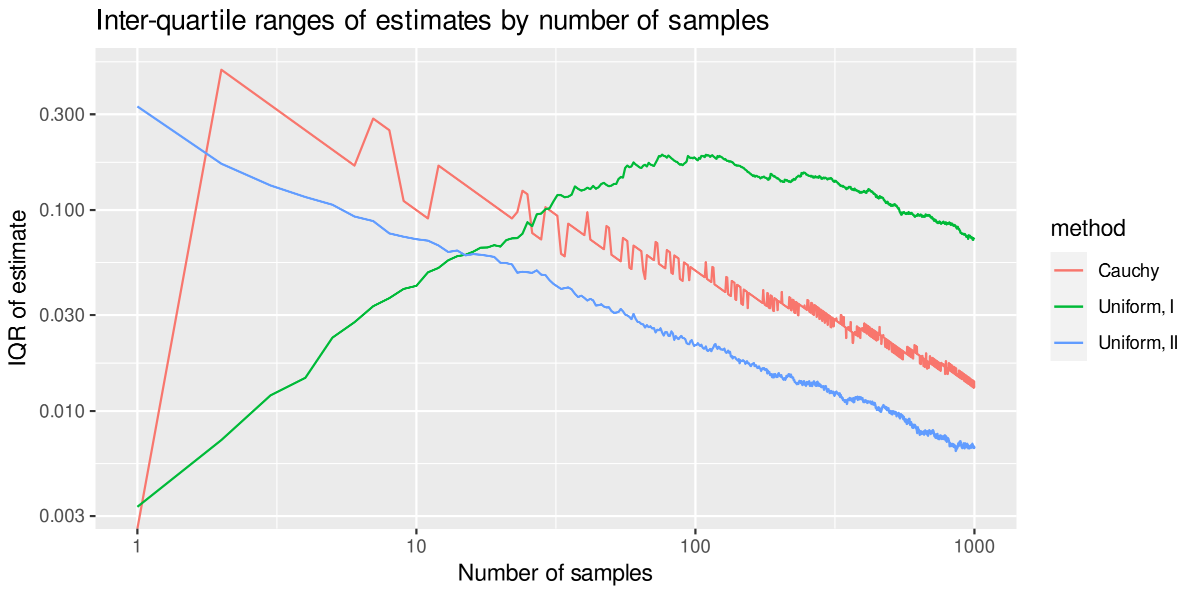
Next, note that for $U \sim \text{Uniform}(2, M)$ and $M \gg 2$, we have
$$I \;=\; \mathbb{E}\left[\frac{1}{\pi(1+U^2)}\cdot\frac{1}{f_U(u)}\right] + \int_M^\infty \frac{1}{\pi(1+x^2)}\,dx,$$
where the last term is very small; indeed,
$$\int_M^\infty \frac{1}{\pi(1+x^2)}\,dx < \int_M^\infty \frac{1}{\pi x^2}\,dx = \frac{1}{\pi M}.$$
We fix $M = 1000$, sample $U_1, \dots, U_n \overset{\text{iid}}{\sim} \text{Uniform}(2, M)$, and estimate
$$\hat{I}_{\text{Uniform,I}} \;=\; \frac{1}{n}\sum_{i=1}^n \frac{1}{\pi(1+U_i^2)}\cdot(M-2) \;\xrightarrow{p}\; \int_2^M \frac{1}{\pi(1+x^2)}\,dx \approx I.$$

In fact, we could also use the idea
$$I = \frac{1}{2} - \int_0^2 \frac{1}{\pi(1+x^2)}$$
and sample $V_1, \dots, V_n \overset{\text{iid}}{\sim} \text{Uniform}(0,2)$, then estimate
$$\hat{I}_{\text{Uniform,II}} \;=\; \frac{1}{2} - \frac{1}{n}\sum_{i=1}^n \frac{1}{\pi(1+V_i^2)}\cdot 2 \;\xrightarrow{p}\; I.$$

These methods have been illustrated in Figure 1. For $n = 10^7$ sample points, we obtain the estimates
$$\begin{aligned}\hat{I}_{\text{Cauchy}} &= 0.1475,\\ \hat{I}_{\text{Uniform,I}} &= 0.1479,\\ \hat{I}_{\text{Uniform,II}} &= 0.1476.\end{aligned}$$
The true value is $1/2 - \arctan(2)/\pi \approx 0.1476$.



(a) Trace plots of estimates of I, obtained via Monte-Carlo sampling from Student's t_{12}, Cauchy, and Gamma distributions, using n sample points. The colored line indicates the median of the estimates, and the ribbon encloses the first and third quartiles.



(b) Inter-quartile ranges of estimates of I against number of samples drawn.

Figure 1: Estimation of I via Monte-Carlo methods.

Problem 2

We estimate

$$I = \mathbb{E}_{X\sim t_{12}}[h(X)], \qquad h(x) = \sqrt{\frac{|x|}{|1-x|}}.$$

First, we sample $X_1, \ldots, X_n \overset{\text{iid}}{\sim} t_{12}$, and estimate

$$\hat{I}_{t_{12}} = \frac{1}{n}\sum_{i=1}^{n} h(X_i) \xrightarrow{p} I.$$

Next, we sample $Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Cauchy}(0,1)$, and estimate

$$\hat{I}_{\text{Cauchy}} = \frac{1}{n}\sum_{i=1}^{n} h(Y_i)\cdot\frac{f_X(Y_i)}{f_Y(Y_i)} \xrightarrow{p} I,$$

where f_X is the density function of $X_1 \sim t_{12}$, and f_Y is the density function of $Y_1 \sim \text{Cauchy}(0,1)$. Finally, we sample $\tilde{Z}_1, \ldots, \tilde{Z}_n \overset{\text{iid}}{\sim} \text{Gamma}(1/2, 1)$, and iid $U_1, \ldots, U_n$ where each $P(U_i = 1) = P(U_i = -1) = 1/2$. Then, we set $Z_i = U_i\tilde{Z}_i + 1$; observe that $|Z_i - 1| \overset{\text{iid}}{\sim} \text{Gamma}(1/2, 1)$. Also, if $f_{\tilde{Z}}$ denotes the density function of $\tilde{Z}_1 \sim \text{Gamma}(1/2, 1)$, then

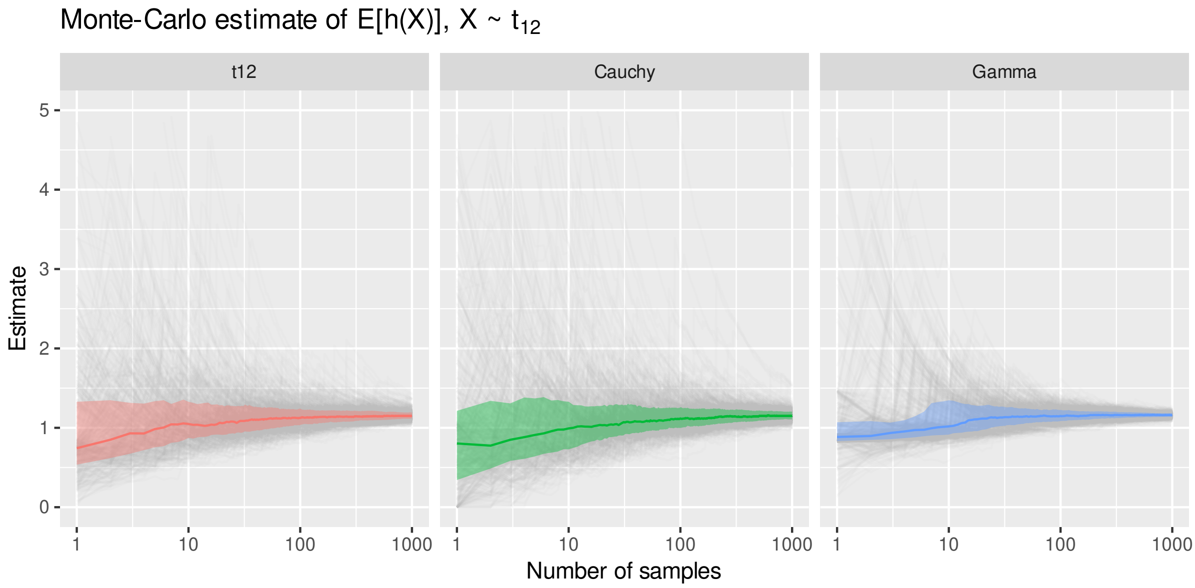
$$f_Z(z) = \frac{1}{2} f_{\tilde{Z}}(|z-1|).$$

Using this, we estimate

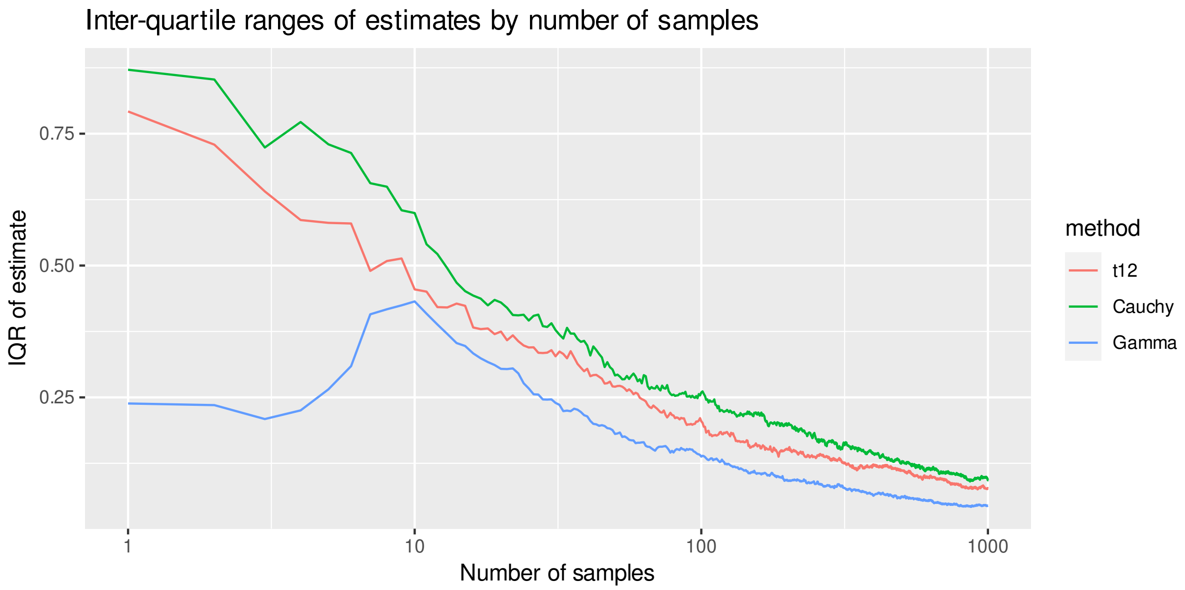
$$\hat{I}_{\text{Gamma}} = \frac{1}{n}\sum_{i=1}^{n} h(Z_i)\cdot\frac{f_X(Z_i)}{f_Z(Z_i)} \xrightarrow{p} I.$$

These methods have been illustrated in Figure 2. For $n = 10^7$ sample points, we obtain the estimates

$$\begin{aligned} \hat{I}_{t_{12}} &= 1.1603, \\ \hat{I}_{\text{Cauchy}} &= 1.1600, \\ \hat{I}_{\text{Gamma}} &= 1.1605. \end{aligned}$$



(a) Trace plots of estimates of I, obtained via Monte-Carlo sampling from a Cauchy and uniform distributions, using n sample points. The colored line indicates the median of the estimates, and the ribbon encloses the first and third quartiles.



(b) Inter-quartile ranges of estimates of I against number of samples drawn.

Figure 2: Estimation of I via Monte-Carlo methods.

Problem 3

We fix our covariates $\boldsymbol{X} = [x_{it}]_{1\le i\le I, 1\le t\le T}$, and have our responses $\boldsymbol{Y} = [y_{it}]$. Let $g(\cdot)$ denote the full conditional of the parameter $(\cdot)$. Then, we have the following.

$$
\begin{aligned}
g(y_{it}) &\propto f(y_{it} \mid \theta_{it} = \alpha_t + \beta_t x_{it} + \epsilon_{it}),\\
g(\epsilon_{it}) &\propto f(\epsilon_{it} \mid \sigma_t^2)\, f(y_{it} \mid \theta_{it}),\\
g(\sigma_t^2) &\propto f(\sigma_t^2) \prod_{i=1}^{I} f(\epsilon_{it} \mid \sigma_t^2),\\
g(\alpha_0) &\propto f(\alpha_0)\, f(\alpha_1 \mid \alpha_0, \tau_\alpha^2),\\
g(\alpha_t) &\propto f(\alpha_t \mid \alpha_{t-1}, \tau_\alpha^2)\, f(\alpha_{t+1} \mid \alpha_t, \tau_\alpha^2) \prod_{i=1}^{I} f(y_{it} \mid \theta_{it}),\\
g(\beta_0) &\propto f(\beta_0 \mid \tau_\beta^2)\, f(\beta_1 \mid \beta_0, \tau_\beta^2),\\
g(\beta_t) &\propto f(\beta_t \mid \beta_{t-1}, \tau_\beta^2)\, f(\beta_{t+1} \mid \beta_t, \tau_\beta^2) \prod_{i=1}^{I} f(y_{it} \mid \theta_{it}),\\
g(\tau_\alpha^2) &\propto f(\tau_\alpha^2) \prod_{t=1}^{T} f(\alpha_t \mid \alpha_{t-1}, \tau_\alpha^2),\\
g(\tau_\beta^2) &\propto f(\tau_\beta^2)\, f(\beta_0 \mid \tau_\beta^2) \prod_{t=1}^{T} f(\beta_t \mid \beta_{t-1}, \tau_\beta^2).
\end{aligned}
$$

Now,

$$
\begin{aligned}
f(y_{it} \mid \theta_{it}) &\propto \frac{e^{\theta_{it} y_{it}}}{\Gamma(y_{it}+1)} e^{-e^{\theta_{it}}} \sim \mathcal{P}(e^{\theta_{it}}),\\
f(\epsilon_{it} \mid \sigma_t^2) &\propto (\sigma_t^2)^{-1/2} e^{-\epsilon_{it}^2/2\sigma_t^2} \sim \mathcal{N}(0, \sigma_t^2),\\
f(\sigma_t^2) &\propto (\sigma_t^2)^{-2} e^{-1/\sigma_t^2} \sim \mathcal{IG}(1,1),\\
f(\alpha_0) &\propto e^{-\alpha_0^2/200} \sim \mathcal{N}(0, 100),\\
f(\alpha_t \mid \alpha_{t-1}, \tau_\alpha^2) &\propto (\tau_\alpha^2)^{-1/2} e^{-(\alpha_t - \alpha_{t-1})^2/2\tau_\alpha^2} \sim \mathcal{N}(\alpha_{t-1}, \tau_\alpha^2),\\
f(\beta_0 \mid \tau_\beta^2) &\propto (\tau_\beta^2)^{-1/2} e^{-\beta_0^2/2\tau_\beta^2} \sim \mathcal{N}(0, \tau_\beta^2),\\
f(\beta_t \mid \beta_{t-1}, \tau_\beta^2) &\propto (\tau_\beta^2)^{-1/2} e^{-(\beta_t - \beta_{t-1})^2/2\tau_\beta^2} \sim \mathcal{N}(\beta_{t-1}, \tau_\beta^2),\\
f(\tau_\alpha^2) &\propto (\tau_\alpha^2)^{-2} e^{-1/\tau_\alpha^2} \sim \mathcal{IG}(1,1),\\
f(\tau_\beta^2) &\propto (\tau_\beta^2)^{-2} e^{-1/\tau_\beta^2} \sim \mathcal{IG}(1,1).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&g(y_{it}) \sim \mathcal{P}(e^{\theta_{it} y_{it}}),\\
&g(\sigma_t^2) \propto (\sigma_t^2)^{-(2+I/2)} e^{-\left(2+\sum_{i=1}^{I} \epsilon_{it}^2\right)/2\sigma_t^2} \sim \mathcal{IG}\left(1 + I/2,\, 1 + \frac{1}{2}\sum_{i=1}^{I} \epsilon_{it}^2\right),\\
&g(\alpha_0) \propto e^{-\alpha_0^2/200 - (\alpha_1 - \alpha_0)^2/2\tau_\alpha^2} \sim \mathcal{N}\left(\frac{100\alpha_1}{100+\tau_\alpha^2}, \frac{100\tau_\alpha^2}{100+\tau_\alpha^2}\right),\\
&g(\beta_0) \propto e^{-\beta_0^2/2\tau_\beta^2 - (\beta_1 - \beta_0)^2/2\tau_\beta^2} \sim \mathcal{N}(\beta_1/2,\, \tau_\beta^2/2),\\
&g(\tau_\alpha^2) \propto (\tau_\alpha^2)^{-2} e^{-1/\tau_\alpha^2} (\tau_\alpha^2)^{-T/2} e^{-\sum_{t=1}^{T} (\alpha_t - \alpha_{t-1}^2)/2\tau_\alpha^2} \sim \mathcal{IG}\left(1 + T/2,\, 1 + \frac{1}{2}\sum_{t=1}^{T} (\alpha_t - \alpha_{t-1})^2\right),\\
&g(\tau_\beta^2) \propto (\tau_\beta^2)^{-2-(T+1)/2} e^{-1/\tau_\beta^2 - \beta_0^2/2\tau_\beta^2 - \sum_{t=1}^{T} (\beta_t - \beta_{t-1})^2/2\tau^2\beta} \sim \mathcal{IG}\left(1 + (T+1)/2,\, 1 + \frac{\beta_0^2}{2} + \frac{1}{2}\sum_{t=1}^{T} (\beta_t - \beta_{t-1})^2\right).
\end{aligned}
$$

The remaining full conditionals do not resolve into standard distributions; they will be sampled via the Metropolis-Hastings algorithm.

For instance, to sample from $g(\alpha_t)$, we may use a proposal distribution of the form $q(\cdot \mid x) = \mathcal{N}(x, \tau_\alpha^2)$.