Automatic Monte Carlo Methods for Bayesian Inference

Noah Grudowski

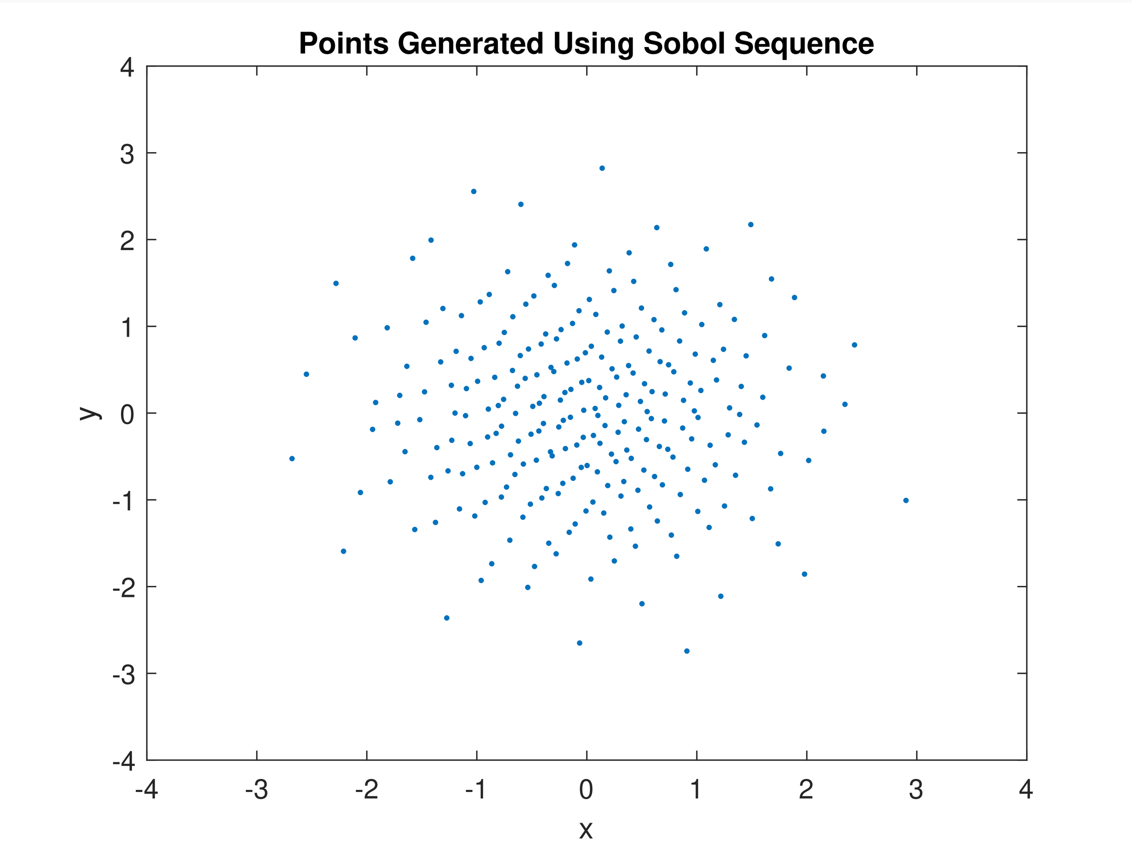
Applied Mathematics, Illinois Institute of Technology

Bayesian Statistics

- Model parameters are random
- Prior distribution, $\pi(b)$, reflects beliefs about the parameters
- Sampling yields a likelihood function, $L(b)$
- Posterior density: product of the prior and the likelihood
- Estimates involve integrals, e.g.
  $\hat{\beta}_j = \frac{\int_{\mathbb{R}^{d+1}} b_j L(b)\pi(b)db}{\int_{\mathbb{R}^{d+1}} L(b)\pi(b)db} =: \frac{\mu_j}{\mu},\ j = 0, \dots, d$

Quasi-Monte Carlo Cubature

- Helps approximate integrals:
  $\int_{[0,1]^{d+1}} f(x)dx \approx \frac{1}{n}\sum_{i=0}^{n-1} f(x_i)$
  where $\{x_i\}_{i=1}^{n}$ is a low-discrepancy sequence
- A low-discrepancy sequence is a correlated sequence whose empirical distribution matches the target distribution better than an IID sequence
- Plot of 256 Sobol sequence points that underwent the standard normal transformation:



Model Problem

- Bayesian inference is applied to logistic regression.
- $t_i \sim$ Ber $\left(\frac{\exp\left(\beta_0+\sum_{j=1}^{d}\beta_j s_{ij}\right)}{1+\exp\left(\beta_0+\sum_{j=1}^{d}\beta_j s_{ij}\right)}\right)$, for $i = 1, 2, \dots, M$
- The prior: standard normal with respect to $\hat{\beta}_j$
  $\pi(\beta) = \frac{\exp\left(-\frac{1}{2}\beta^T\beta\right)}{\sqrt{(2\pi)^{d+1}}}$
- Widely used Bayesian inference estimator:
  $\hat{\beta}_j = \frac{\int_{\mathbb{R}^{d+1}} b_j L(b)\pi(b)db}{\int_{\mathbb{R}^{d+1}} L(b)\pi(b)db} =: \frac{\mu_j}{\mu},\ for\ j = 0, 1, 2, \dots, d$
  where $\hat{\beta}$ are the unknown parameters

Choice of Density

- The ratio $\frac{\mu_j}{\mu}$ cannot be calculated analytically
- Thus, we rewrite μ and μ_j as such:
  $\mu = \int_{\Re^{d+1}} L(b)\pi(b)db = \int_{\Re^{d+1}} \frac{L(b)\pi(b)}{\rho(b)}\rho(b)db,$
  where $\rho(b)$ is some density. Thus, $\frac{\mu_j}{\mu}$ can be estimated by sampling from $\rho(b)$ and applying quasi-Monte Carlo cubature.
- We have the following choices for $\rho(b)$:
  1) $\rho = \pi$
  2) $\rho = \rho_{\text{MLE}} =$ Gaussian approximation to the likelihood
  3) $\rho \propto \pi \cdot \rho_{\text{MLE}}$

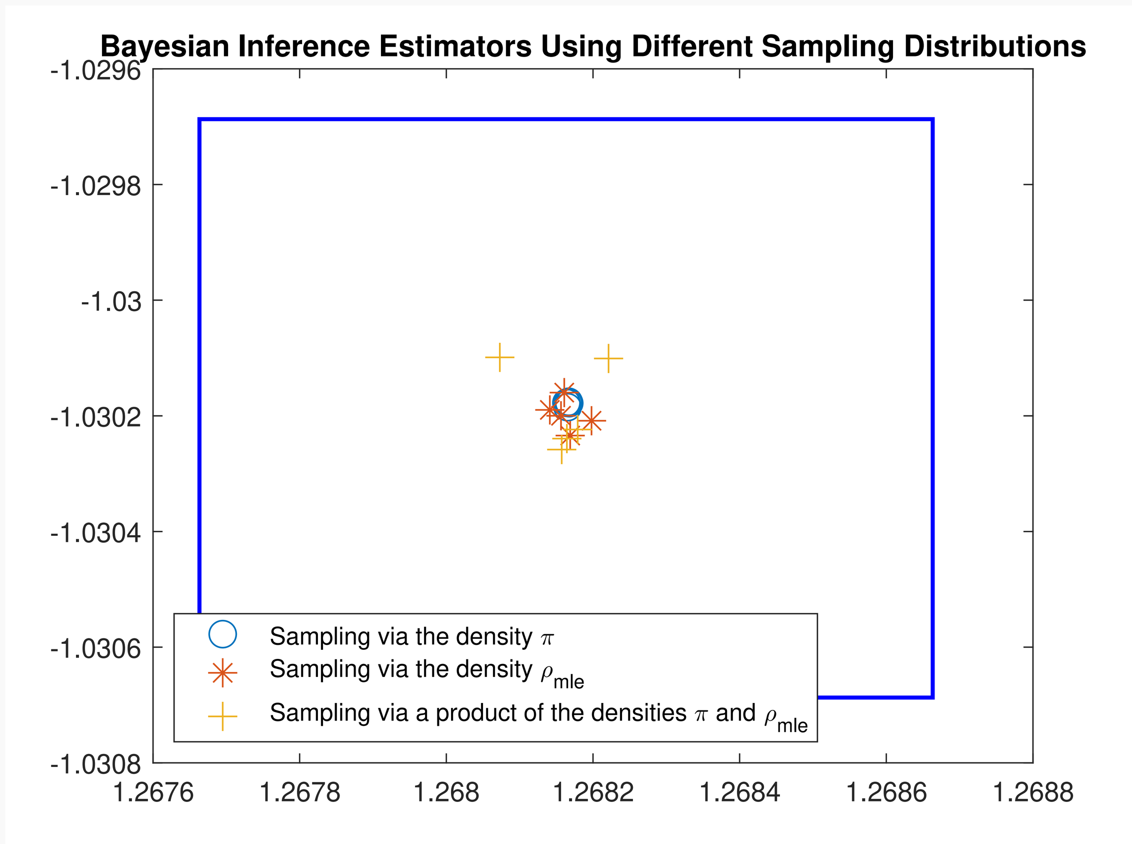
New MATLAB Function

- Developed function `BayesianFunction` that automatically estimates the unknown parameters $\hat{\beta}$ using GAIL (Choi Et al, 2015)
- User control over various inputs

Use Case: Uniform Random s-Data

```
% s-Data:100 uniform random values in [-2,6]
M = 100; %number of s-Data values
s = (6+2)*rand(1,M)-2; %s-Data uniform
dim=2; absTol=0.0005; n=5;
densityChoice = [true true true];
%Use all 3 densities
BayesianFunction(M, s, dim, absTol,
    densityChoice,n)
```

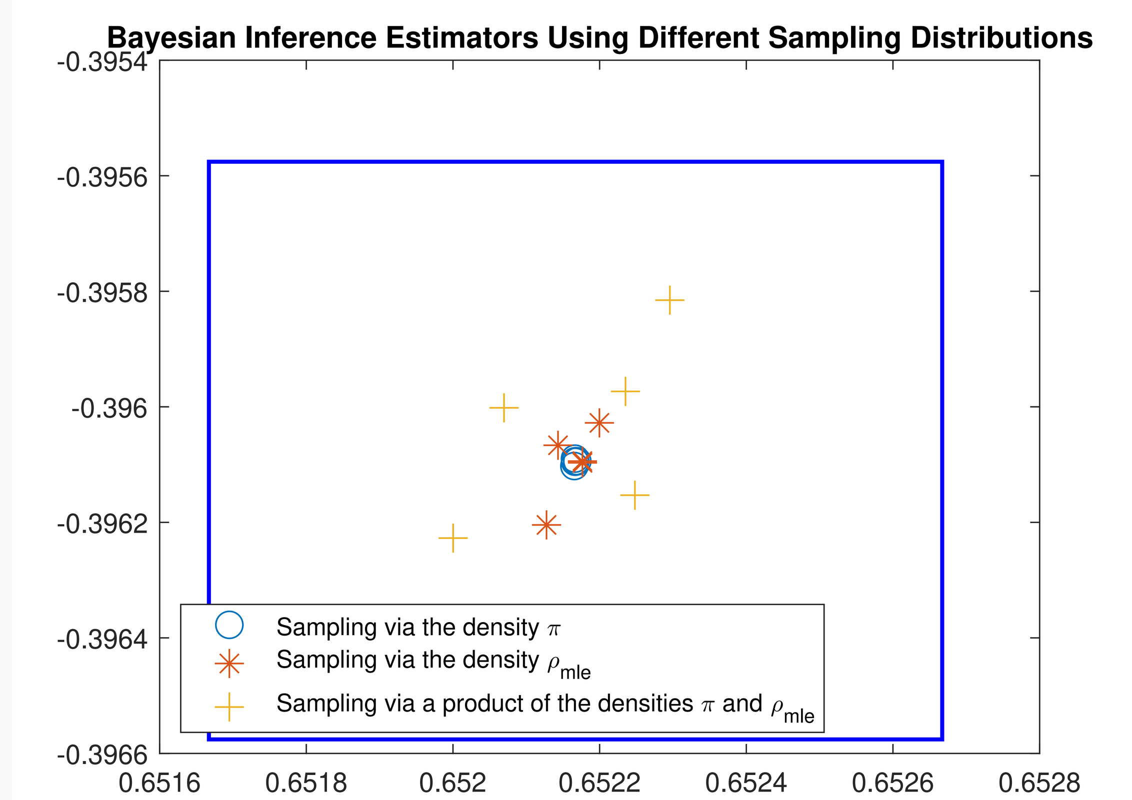
Density	Runtime (seconds)	Samples Taken
$\rho = \pi$	86.088	2621440
$\rho = \rho_{\text{MLE}}$	2.4954	81920
$\rho \propto \pi \cdot \rho_{\text{MLE}}$	1.3744	40960



Use Case: Normal Random s-Data

```
% s-Data: 100 normal random values
% with mean 2 and variance 4
M = 100; %number of s-Data values
u = 2; %mean
sd = 2; %standard deviation
s = sd.*randn(1,M)+u; %s-Data normal
dim=2; absTol=0.0005; n=5;
densityChoice = [true true true];
%Use all 3 densities
BayesianFunction(M, s, dim, absTol,
    densityChoice,n)
```

Density	Runtime (seconds)	Samples Taken
$\rho = \pi$	13.7314	327680
$\rho = \rho_{\text{MLE}}$	0.89835	20480
$\rho \propto \pi \cdot \rho_{\text{MLE}}$	1.4781	32768



Discussion

- Add some
- Concluding
- Remarks here

References

- Choi, S.C.T., Ding, Y., Hickernell, F.J., Jiang, L., Jimenez Rugama, Ll.A., Tong, X., Zhang,Y., Zhou, X. *GAIL: Guaranteed Automatic Integration Library (version 2.1)*, MATLAB software, 2015.
- Hickernell, F.J., Jimenez Rugama, Ll.A., Li, D. Adaptive Quasi-Monte Carlo Methods for Cubature.
- Brooks-Bartlett, Jonny. Probability Concepts Explained: Bayesian Inference for Parameter Estimation. Towards Data Science. Retrieved August 5, 2018.

A special appreciation and thank you to the IIT College of Science for providing the funding behind this research project.