Monte Carlo Integration

Timothy Hanson

Department of Statistics, University of South Carolina

Stat 740: Statistical Computing

LLN & CLT

Recall that the strong law of large numbers implies, under general conditions, for

$$X_1, \ldots, X_m \overset{iid}{\sim} f(\cdot),$$

that

$$\bar{g}_m = \frac{1}{m}\sum_{i=1}^{m} g(X_i) \overset{a.s.}{\to} \underbrace{\int_{-\infty}^{\infty} g(x)f(x)dx}_{E_f[g(X)]} .$$

Multivariate versions work the same.

The central limit theorem with a plug-in variance helps estimate/bound error

$$\bar{g}_m \overset{\bullet}{\sim} N(E_f[g(X)], v_m),$$

where $v_m = \frac{1}{m^2}\sum_{i=1}^{m}[g(x_i) - \bar{g}_m]^2$.

Example: $E(X^4)$ for $X \sim N(1, 2^2)$.

Quantiles

Let $X_1, \ldots, X_m \overset{iid}{\sim} f(\cdot)$ with cdf $F(\cdot)$.

The sample quantile from a Monte Carlo sample also estimates its population counterpart. Let $F_m(x) = \frac{1}{m} \sum_{i=1}^{m} I\{X_i \leq x\}$ be the empirical cdf.

The pth sample quantile q_m satisfies $F_m^{-}[q_m] = p$, although R uses a weighted version in the `quantile` function; see `https://en.wikipedia.org/wiki/Quantile` for several alternative different definitions. Simplest is $q_m = X_{[(mp)]}$. Can also replace $[\cdot]$ with $\lceil \cdot \rceil$ or $\lfloor \cdot \rfloor$.

Asymptotically,

$$q_m \overset{\bullet}{\sim} N\left(q, \frac{p(1-p)}{mf(q)^2}\right), \text{ where } F(q) = p.$$

Sketch of proof

Since $mF_m(q) \sim \text{binom}(m, F(q))$, we have $F_m(q) \overset{\bullet}{\sim} N(F(q), \frac{1}{m}F(q)[1 - F(q)])$. Consider the transformation $g(p) = F^{-1}(p)$ and use the delta method:

$$F^{-1}\{F_m(q)\} \overset{\bullet}{\sim} N\left(q, \frac{F(q)[1 - F(q)]}{mf\{F^{-1}[F(q)]\}^2}\right).$$

noting that $F^{-1}\{F(q)\} = q$ and $\frac{d}{dp}F^{-1}(p) = 1/f\{F^{-1}[F(q)]\}$. Now let q_m solve $p = F_m(q_m)$ and note $|q_m - F^{-1}\{F_m(q_m)\}| \overset{a.s.}{\to} 0$, giving

$$q_m \overset{\bullet}{\sim} N\left(q, \frac{p(1 - p)}{mf(q)^2}\right).$$

Quantiles of $g(X)$

Take $y_i = g(x_i)$. The sample quantile from $y_1, \ldots, y_m$ estimates the corresponding population quantile of the r.v. $Y = g(X)$.

The Monte Carlo solution is especially easy compared to an exact solution, which requires change-of-variables and an integral equation!

Example: $90th$ percentile of X^4 from $X \sim N(1, 2^2)$.

Probability intervals

Often interested in (a, b) s.t. $P(a \leq X \leq b) = p$. An *equal-tailed* interval uses the $2.5th$ and $97.5th$ percentiles of a Monte Carlo sample.

A highest density interval finds the (a, b) such that $P(a \leq X \leq b) = p$ and the interval length $b - a$ is the smallest.

Highest density *regions* R for multivariate $\mathbf{x}$ satisfy $P(\mathbf{x} \in R) = p$ and $f(\mathbf{x}_1) \geq f(\mathbf{x}_2)$ for $\mathbf{x}_1 \in R$ and $\mathbf{x}_2 \notin R$. They usually are not "nice" shaped.

Examples for X^4 from $X \sim N(1, 2^2)$.

Comments

- Monte Carlo is a very easy way to estimate means and quantiles of complex functions.
- *Need to be able to sample $f(\cdot)$ to use it directly!*
- If cannot sample from $f(\cdot)$ can use Markov chain Monte Carlo (coming up).
- Monte Carlo is a terrible way to solve some problems, e.g. estimating tail probabilities.
- Another alternative is importance sampling and sampling importance resampling; can be used when cannot sample $f(\cdot)$ but can sample proposal $g(\cdot)$, and also useful in situations like estimating tail probabilities.
- Rao-Blackwellization can also improve the Monte Carlo estimation of integrals (more later).

Bayesian inference

Historically, Bayesian inference relied on normal approximations, Laplace approximations, and importance sampling.

The MCMC era was ushered in by Gelfand and Smith (1990). They realized that earlier Markov chain methods (Metropolis et al., 1953; Hastings, 1970; Geman and Geman, 1984; Tanner and Wong, 1987) had much broader applicability than previously thought. Note: although Metropolis was given the honor of the algorithm's name, M. Rosenbluth actually invented it.

We will use MCMC to obtain inference for posterior densities $\pi(\boldsymbol{\theta}|\mathbf{x})$, both hand-coded in R, and also using JAGS.