1 Choose-your-own-adventure (Binomial random variables).

How can we simulate a \(\text{Bin}(N,p)\) random variable on a computer? There is a direct method of flipping \(N\) coins with probability of success \(p\) and counting the number of successes, but this becomes expensive for large \(N\). What is a better strategy for large \(N\)? How well does it work? Try it on a computer or study it mathematically. (Hint: remember the central limit theorem.)

2 Choose-your-own-adventure (Poisson random variables).

How can we simulate a \(\text{Pois}(\lambda)\) random variable on a computer? There is a classical characterization of the Poisson distribution, in which we wait at a bus stop and count how many passengers arrive. The interarrival times are independent \(\text{exp}(\lambda)\) random variables and the number of passengers at time \(t=1\) is a \(\text{Pois}(\lambda)\) random variable. Use this characterization to derive a slick sampling algorithm. (This is how Matlab and Python do it for \(\lambda \leq 10\) or \(\lambda \leq 15\)).

3 Choose-your-own-adventure (Gaussian tails).

We can use rejection sampling to obtain samples from a conditional Gaussian distribution \(Z|Z \geq C\) where \(Z \sim \mathcal{N}(0,1)\) and \(C = 10\) or even larger.

(a) One idea is to use a \(\mathcal{N}(0,1)\) trial distribution. What is the resulting algorithm? What is the expected wait time to acceptance (try it on a computer or derive it mathematically)?

(b) A different idea is to use a \(\mathcal{N}(C,1)\) trial distribution. What is the resulting algorithm? What is the expected wait time to acceptance (try it on a computer or derive it mathematically)?

(Hint: for nice math results, use Mills’ ratio \(\frac{C}{C^2+1} \leq \frac{P\{Z>C\}}{\phi(C)} \leq \frac{1}{C}\) for \(C > 0\) and \(\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\)).

4 Coding (Digits of \(\pi\)).

In Matlab or Python, we can evaluate \(\pi\) with 16 digits of accuracy, but what if we want 100 digits?

(a) Write a Monte Carlo method to calculate \(\pi\). How many digits of accuracy can we get?

(b) As an alternative for calculating \(\pi\), use the formula \(\pi = 4 \arctan(1)\) and the Taylor series expansion of \(\arctan x\). Better yet, use the formula \(\pi = 4 \arctan(1/2) + 4 \arctan(1/3)\) (Euler, 1737) and the Taylor series expansion for \(\arctan x\). How many digits of accuracy can we get?

5 Coding (Repulsive eigenvalues).

Let \(G \in \mathbb{R}^{N \times N}\) be a square matrix with independent \(\mathcal{N}(0,1)\) entries. Then, \(X = \frac{1}{N} GG^*\) is a positive definite Wishart matrix.

(a) Simulate a Wishart matrix in high dimensions and show the histogram of eigenvalues converges to the Marchenko-Pastur distribution with density \(f(x) = (4/x - 1)^{1/2}/(2\pi)\) for \(0 < x < 4\).
(b) Write an exact Monte Carlo sampler for the Marchenko-Pastur distribution. (Hint: The square of a Unif(0, 1) random variable has density $f(x) = \frac{1}{2} x^{-1/2}$ for $0 < x < 1$. Use rescaling and rejection sampling.)

(c) Show that the Wishart matrix eigenvalues are more repulsive than independent Marchenko-Pastur random variables. How do you explain the repulsion, mathematically or intuitively?

6 Math (From one to many).

Prove that given just ONE Unif(0, 1) random variable, we can generate INFINITELY MANY independent Unif(0, 1) random variables. Question: can we make this work on a computer?

7 Math (Quantile problem).

For a random variable $X \in \mathbb{R}$, let $F(x) = \mathbb{P}\{X \leq x\}$ be the cumulative distribution function (cdf) and let $Q(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}$ be the quantile function.

(a) If $F$ is continuous and strictly increasing, prove that $Q$ is the inverse of $F$, i.e., $Q(F(x)) = x$.

(b) If $U$ is a Unif(0, 1) random variable, prove that $Q(U)$ is a random variable with cdf $F$.

8 Math (Importance beats rejection).

We want to simulate from a pdf $f$ and evaluate a statistic $\mu_h = \int f(x) h(x) \, dx$, but we can only simulate from a pdf $g$. First, we fix a number $M \geq \sup_x f(x)/g(x)$. Then, we draw independent samples $X_1, \ldots, X_N$ with pdf $g$, we calculate $w_i = f(X_i)/(g(X_i)M)$ for $1 \leq i \leq N$, and we draw independent random variables $U_1, \ldots, U_n \sim \text{Unif}(0, 1)$. Define the rejection sampling estimator

$$\hat{h}_{\text{rej}} = \frac{\mathbb{I}\{U_1 \leq w_1\} h(X_1) + \cdots + \mathbb{I}\{U_N \leq w_N\} h(X_N)}{\mathbb{I}\{U_1 \leq w_1\} + \cdots + \mathbb{I}\{U_N \leq w_N\}}.$$ 

and the importance sampling estimator

$$\hat{h}_{\text{imp}} = \frac{w_1 h(X_1) + \cdots + w_N h(X_N)}{w_1 + \cdots + w_N}.$$ 

As $N \to \infty$, prove the central limit theorems

$$\sqrt{N} (\hat{h}_{\text{rej}} - \mu_h) \xrightarrow{D} \mathcal{N}(0, V_{\text{rej}}^2), \quad \sqrt{N} (\hat{h}_{\text{imp}} - \mu_h) \xrightarrow{D} \mathcal{N}(0, V_{\text{imp}}^2),$$

where

$$V_{\text{rej}}^2 = M \int f(x) |h(x) - \mu_h|^2 \, dx \geq V_{\text{imp}}^2 = \int \frac{f(x)^2}{g(x)} |h(x) - \mu_h|^2.$$ 

Partial solution (avoiding the delta method). Set $h'(x) = h(x) - \mu_h$ and use the central limit theorem for averages of independent random variables to show

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbb{I}\{U_i \leq w_i\} h'(X_i) \xrightarrow{D} \mathcal{N} \left(0, \frac{1}{M} \int f(x) |h'(x)|^2 \, dx \right).$$
and
\[
\frac{1}{\sqrt{N}} \left( \sum_{i=1}^{N} \mathbb{1}\{U_i \leq w_i\} - \frac{1}{M} \right) = \mathcal{O}_p(1)
\]
as \(N \to \infty\). Next, recall that a quotient \(n/d\) has Taylor series expansion
\[
\frac{n}{d} = \frac{n}{d_0} - \frac{n_0}{d_0} \cdot \frac{d - d_0}{d_0} + \mathcal{O}(n - n_0)^2 + \mathcal{O}(d - d_0)^2
\]
Setting \(n_0 = 0, \ d_0 = 1/M\),
\[
n = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{U_i \leq w_i\} h'(X_i), \quad \text{and} \quad d = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{U_i \leq w_i\},
\]
we get
\[
\hat{h}_{\text{rej}} - \mu_h = \left( \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{U_i \leq w_i\} h'(X_i) \right) / \left( \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\{U_i \leq w_i\} \right)
\]
\[
= \frac{M}{N} \sum_{i=1}^{N} \mathbb{1}\{U_i \leq w_i\} h'(X_i) + \mathcal{O}_p \left( \frac{1}{N} \right)
\]
and by an application of Slutsky’s theorem
\[
\sqrt{N} (\hat{h}_{\text{rej}} - \mu_h) \xrightarrow{D} \mathcal{N} \left( 0, \frac{1}{M} \int f(x) |h'(x)|^2 \, dx \right).
\]