Lecture 20, Nov 24, 2023

Sampling Distributions in Practice

- Most math libraries have functions that generate uniformly distributed random real numbers in the range $(0, 1)$
  - e.g. `rand()` in MATLAB, `np.random.rand()` in Python
  - This interval will sometimes be closed or half-open, but practically we don't care
  - $f_u(u) = \begin{cases} 1 & u \in (0,1) \\ 0 & \text{otherwise} \end{cases}$
- Repeated calls to the RNG are independent
- The generator can usually be seeded, e.g. with `np.random.seed()`; this gives the same sequence of random numbers for the same seed
- How do we draw samples from arbitrary, non-uniform PDFs?

One Variable, Discrete

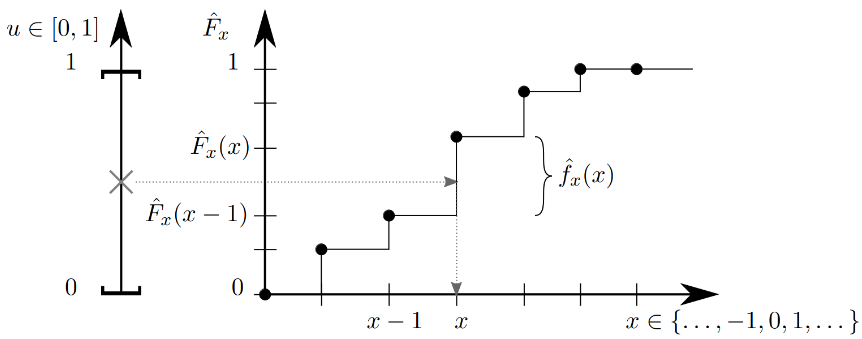

Figure 1: Algorithm to select one discrete random variable.

- Given a desired PDF $\hat{f}_x(x)$ for a DRV x, we want to come up with a procedure to generate x from u
- WLOG let $\mathcal{X} = \mathbb{Z}$ (note we can remap any DRV to be over the integers)
- The *cumulative distribution function* (CDF) of $\hat{f}_x$ is $\hat{F}_x(x) = \displaystyle\sum_{\bar{x}=-\infty}^{x} \hat{f}_x(\bar{x})$
  - Note $\hat{F}_x(-\infty) = 0$ and $\hat{F}_x(\infty) = 1$, and $\hat{F}_x$ is a non-decreasing function
  - We will make use of the fact that both $\hat{F}_x$ and the output of the RNG range from 0 to 1
- Let u be generated from $f_u(u)$; solve for x such that $\hat{F}_x(x-1) < u, \hat{F}_x(x) \geq u$; we claim that x will have PDF $\hat{f}_x(x)$
  - Intuition: we chop up the interval $[0, 1]$, so that each x gets a portion of the interval that is proportional to $\hat{f}_x(x)$
    - To see this note $\hat{F}_x(x) - \hat{F}_x(x-1) = \hat{f}_x(x)$
- Note that we can always solve for such an x given u, since $\hat{F}_x$ ranges from 0 to 1
  - There may be issues with $u = 0$ and $u = 1$, but this probability is technically 0 u is a CRV
  - In practice we can explicitly check for these cases and re-sample if we obtain them
- For a fixed x, to have $\hat{F}_x(x-1) < u \leq \hat{F}_x(x)$ we need $\hat{F}_x(x-1) < u \leq \hat{F}_x(x-1) + \hat{f}_x(x)$
  - Therefore $\displaystyle\int_{\hat{F}_x(x-1)}^{\hat{F}_x(x)} f_u(u)\,\mathrm{d}u = \int_{\hat{F}_x(x-1)}^{\hat{F}_x(x-1)+\hat{f}_x(x)} 1\,\mathrm{d}u = \hat{f}_x(x)$
  - So the probability that we get x is $\hat{f}_x(x)$

Multiple Variables, Discrete

- If we want $\hat{f}_{xy}(x, y)$

- If $\mathcal{X}$ and $\mathcal{Y}$ are finite, with N_x and N_y elements respectively, let $\mathcal{Z} = \{ 1, 2, \ldots, N_x N_y \}$, and define a one-to-one mapping between elements of $\mathcal{Z}$ and (x, y), and sample with the one-variable algorithm
- Otherwise, decompose $\hat{f}_{xy}(x, y) = \hat{f}_{x|y}(x|y)\hat{f}_y(y)$
  - Apply the one-variable algorithm to sample y first from $\hat{f}_y(y)$ (obtained by marginalizing the joint PDF)
  - Then apply the same algorithm again to get a value for x from $\hat{f}_{x|y}(x|y)$
- These algorithms both apply to any number of DRVs

One Variable, Continuous

- Let $\hat{F}_x(x) = \displaystyle\int_{-\infty}^{x} \hat{f}_x(\bar{x})\mathrm{d}\bar{x}$
- Let u be generated from $f_u(u)$; then have $x = \hat{F}_x^{-1}(u)$, and x will have PDF $f_x(x) = \hat{f}_x(x)$
  - x is any value that satisfies $u = \hat{F}_x(x)$; this will still work even if $\hat{f}_x$ is zero sometimes
- Assume that $\hat{F}_x$ is strictly increasing, then we can solve for a unique x given any u
  - For some arbitrary a, $F_x(a) = \Pr(x \leq a) = \Pr(\hat{F}_x^{-1}(u) \leq a)$
  - Applying $\hat{F}_x$ to both sides, this becomes $\Pr(u \leq \hat{F}_x(a))$
  - Since u is uniform, $\Pr(u \leq \hat{F}_x(a)) = \hat{F}_x(a)$
  - Therefore $F_x(a) = \hat{F}_x(a)$, so we must have $f_x(x) = \hat{f}_x(x)$

Multiple Variables, Continuous

- For multiple CRVs, again decompose $\hat{f}_{xy}(x, y) = \hat{f}_{x|y}(x|y)\hat{f}_y(y)$, and apply the one-variable algorithm to get values for y first from $\hat{f}_y(y)$, and then x from $\hat{f}_{x|y}(x|y)$