Lecture Summary 1

Suppose we want to find

$$I = \int_0^1 f(x)\,dx$$

where $0 \le f \le 1$. As is well known this value is the area under the curve and I is defined (for as it is it is merely notation for something) using Lebesgue integration or Riemannian integration. We will provide a definition for Monte Carlo integration.

If we threw n random points uniformly in the unit square, $(0,1)^2$, the ratio of samples lying below the curve would be an estimate of the area; i.e. of I itself. That is,

$$I_n = \frac{\#\{\text{samples lying below the curve}\}}{n}.$$

And as n gets larger, so the accuracy gets better. We will see this property with the law of large numbers. More formally, if $\Omega\,(=1)$ is the area of the unit square and A the area under the curve,

$$\lim_{n\to\infty} I_n = \frac{A}{\Omega}.$$

Note the sequence I_n is a random sequence so the limit is a bit loose at the moment; we will rectify this later with the words "almost surely" or "in probability", once we have a better grip on what they mean.

To set the stage for the notation; the uniform sample is represented by $(X_i, Y_i)_{i=1}^n$ where (X_i) are independent and uniform on $(0,1)$, same for the (Y_i), and the (X_i) and (Y_i) are independent of each other. The joint density function for each (X,Y) is

$$f(x,y) = 1, \quad (x,y) \in (0,1)^2.$$

We write

$$I_n = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(Y_i < f(X_i)),$$

where $\mathbf{1}$ is the 0–1 indicator function given in class.

This clearly makes out the estimator is going for the area under the curve. However, we use

$$\widehat{I}_n = \mathrm{E}\left[I_n \mid X_1,\ldots,X_n\right] = \frac{1}{n}\sum_{i=1}^n f(X_i),$$

since both $\mathrm{E}\,I_n = \mathrm{E}\,\widehat{I}_n = I$ (both are unbiased) yet

$$\mathrm{Var}\,I_n > \mathrm{Var}\,\widehat{I}_n,$$

so $\widehat{I}_n$ is a better estimator of I than I_n is.