University of Cape Town - Department of Physics
Honours Computational Physics

TOPIC 4 - MONTE CARLO INTEGRATION : 1D IMPORTANCE SAMPLING AND SAMPLE MEAN METHODS

This worksheet accompanies the EJS simulation `MCIntegration_No2_1DImpSampAndSampleMean.jar`

The challenge: Our task is to evaluate the following integral[1]:

$$\int_a^b f(x)\ dx.$$

As introduced in a previous worksheet, the sample mean approximation estimates the integral as:

$$\int_a^b f(x)\ dx \approx \frac{(b-a)}{N} \sum_{i=1}^{N} f(x_i),$$

where the x_i are sampled uniformly from the interval $[a, b]$. The result is exact in cases where the integrand is a constant, but generally there is an uncertainty/error determined by the variance in the integrand, the size of the integration interval and the number of trials.

The importance sampling technique involves identifying an easily-sampled probability distribution function $p(x)$ that matches $f(x)$ in its main features. This approach estimates the integral as:

$$\int_a^b f(x)\ dx = \int_a^b p(x) \frac{f(x)}{p(x)}\ dx \approx \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)},$$

where the x_i are sampled according to $p(x)$. In order to sample according to $p(x)$, one can implement any of the techniques introduced earlier (e.g. the rejection, inverse-transform, combination analytic-rejection or Metropolis methods). The advantage of importance sampling is that the uncertainty in the integral estimate depends on the variance of $f(x)/p(x)$, which will be less than $(b-a)\,\mathrm{var}(f(x))$ if $p(x)$ is suitably chosen.

Questions:

1. Consider the integral,

$$\int_0^5 e^{-x^2}\ dx.$$

This integral cannot be performed analytically and since the integrand shows large variation over the integration region, the sample mean estimate is expected to have a large uncertainty.

   (a) Use the sample mean approach, as implemented in the associated EJS simulation, to evaluate this integral (perform 200 runs each with 100 trials).

      i. How is the final result that is quoted related to the histogram of run results? Overlay a suitable Gaussian on the histogram to confirm your answer.

[1] Monte Carlo techniques outperform traditional methods of integration only in higher dimensions. Our approach here, however, is to use the 1D case to introduce the concepts.

ii. How is the uncertainty quoted in the final result related to relevant quantities (e.g. N_{trials}, N_{runs}, integration interval size, variances etc.)?

iii. If you were able to perform just one run with N_{trials} evaluations, how would you calculate the uncertainty?

(b) Now investigate the use of the importance sampling approach, as implemented in the associated EJS simulation, to evaluate this integral (again performing 200 runs each with 100 trials).

i. Identify a suitable probability distribution $p(x)$ and a transformation function $x(y)$ that converts a uniformly sampled y in $[0, 1)$ into an x according to the chosen distribution (the simulation uses the inverse-transform method to sample according to $p(x)$).

ii. How is the final result that is quoted related to the histogram of run results? Overlay a suitable Gaussian on the histogram to confirm your answer.

iii. How is the uncertainty quoted in the final result related to relevant quantities (e.g. N_{trials}, N_{runs}, integration interval size, variances etc)?

iv. If you were able to perform just one run with N_{trials} evaluations, how would you calculate the uncertainty?

(c) Compare the results of the two integration techniques.

i. Are you convinced that the importance sampling technique outperforms the sample mean approach with the same number of trials and runs? Is a larger variance of $f(x)$ compared to that of $f(x)/p(x)$ the reason? Discuss.

ii. How many more trials would be required in each sample mean approach run in order for the sample mean approach to yield a more precise result than the importance sampling technique with $N_{\text{trials}} = 100$ and N_{runs}=200?

2. Use the associated EJS simulation to explore the effectiveness of the sample mean and importance sampling techniques in solving the integral:

$$\int_0^{10} x^2 e^{-x^2}\, dx.$$