... importance sampling

$$\int_0^1 \sin(e^x)\,dx = E\Big(\underbrace{\sin(e^x)}_{g(x)} \,\Big|\, X \sim R(0,1)\Big) = \theta$$

$$\hat{\theta}_{MC} = \frac{1}{n}\sum_{i=1}^{n} \sin(e^{x_i})$$

To evaluate $\theta = \int_{\mathcal{X}} g(x)\,dx$.

Consider a distn. f over $\mathcal{X}$, sampling from which is easier.

Then $$\theta = \int_{\mathcal{X}} g(x)\,dx = \int_{\mathcal{X}} \frac{g(x)}{f(x)} \cdot f(x)\,dx = E\left(\frac{g(x)}{f(x)} \,\Big|\, X \sim f\right)$$

The MC estimator of θ is therefore $\hat{\theta}_{IS} = \frac{1}{n}\sum_{i=1}^{n} \frac{g(x_i)}{f(x_i)}$, where $X_1, \ldots, X_n \overset{iid}{\sim} f$.

- f is called the importance sampling function

f is chosen such that

$$\left|\frac{g(x)}{f(x)}\right| \le c(x) \text{ with } E(X) < \infty \text{ is satisfied.}$$
(almost bounded)

Q. $X \sim N(0,1)$

$\theta = P(X > c)$, $c > 3$

$$\hat{\theta}_{MC} = \frac{1}{n}\sum_{i=1}^{n} I(x_i > c), \quad X_1, \ldots, X_n \overset{iid}{\sim} N(0,1)$$

$\hat{\theta}_{MC}$ gives inaccurate values as c exceeds the 3σ limit.

$$P(X > c) = \int_c^{\infty} \phi(x)\,dx = \int^{\infty} I(x > c)\,\phi(x)\,dx$$