Joseph High
Hopkins ID: 9E1FDC

553.633 : Monte Carlo Methods
Homework # 1

(A.) Suppose that a random variable X has a symmetric triangular probability density function over the interval $[-1, 1]$ (i.e., with x the dummy variable for the density function, the density is $1-|x|$ for $x \in [-1,1]$ and 0 for $x \notin [-1,1]$). What is $Var(X)$ (the variance of X)?

Solution: Given the pdf, f, of the r.v. X

$$f(x) = \begin{cases} 1-|x| , & x \in [-1,1] \\ 0 , & x \notin [-1,1] \end{cases}.$$

Moreover, by definition of the variance: $Var[X] = E[X^2] - (E[X])^2$
Assuming X is a continuous random variable (since X has a triangular distribution on the interval $[-1,1]$), we have that

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{-1}^{1} x^2(1-|x|)\,dx = \int_{-1}^{1} x^2\,dx - \int_{-1}^{1} x^2|x|\,dx$$

$$= \int_{-1}^{1} x^2\,dx - \left[-\int_{-1}^{0} x^3\,dx + \int_{0}^{1} x^3\,dx \right] \quad \text{since } x^2|x| = \begin{cases} x^3 , & x \geq 0 \\ -x^3 , & x < 0 \end{cases}$$

$$= \frac{x^3}{3}\Big|_{-1}^{1} - \left[-\frac{x^4}{4}\Big|_{-1}^{0} + \frac{x^4}{4}\Big|_{0}^{1} \right] = \frac{2}{3} - \left[\frac{1}{4} + \frac{1}{4} \right] = \frac{1}{6}$$

$$\Rightarrow E[X^2] = 1/6$$

Now,

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-1}^{1} x(1-|x|)\,dx = \int_{-1}^{1} x\,dx - \int_{-1}^{1} x|x|\,dx$$

$\Longrightarrow$

continued...

$$= \int_{-1}^{1} x\,dx - \left[- \int_{-1}^{0} x^2\,dx + \int_{0}^{1} x^2\,dx \right] \quad \text{since } x|x| = \begin{cases} x^2, & x \geq 0 \\ -x^2, & x < 0 \end{cases}$$

$$= \frac{x^2}{2}\Big|_{-1}^{1} - \left[-\frac{x^3}{3}\Big|_{-1}^{0} + \frac{x^3}{3}\Big|_{0}^{1} \right] = 0 - \left[-\frac{1}{3} + \frac{1}{3} \right] = 0$$

$$\Rightarrow E[X] = 0$$

Then with $Var[X] = E[X^2] - (E[X])^2$

$$= \frac{1}{6} - (0)^2 = \frac{1}{6}$$

$$\Rightarrow \boxed{Var[X] = 1/6}$$

//

(B.) Exercise 1 in week 1 handout: Suppose a simulation output vector X has 3 components. Suppose that $\|\bar{X} - \mu\| = 2.276$ and $\bar{X} - E(X) = [1.0 \quad 1.9 \quad -0.1]^T$

(a) Using the information above and the standard Euclidean (distance) norm, what is a (strictly positive) lower bound to the validation/verification error $\|E(X) - \mu\|$?

Solution: Since the Euclidean norm satisfies the triangle inequality we have that

$$\|\bar{X} - \mu\| \leq \|\bar{X} - E(X)\| + \|E(X) - \mu\|$$

$$\Rightarrow \|E(X) - \mu\| \geq \|\bar{X} - \mu\| - \|\bar{X} - E(X)\|$$

where it is given that $\|\bar{X} - \mu\| = 2.276$

Since, $\bar{X} - E(X) = \begin{bmatrix} 1.0 \\ 1.9 \\ -0.1 \end{bmatrix} \Rightarrow \|\bar{X} - E(X)\| = \sqrt{(1.0)^2 + (1.9)^2 + (-0.1)^2}$

$$= \sqrt{4.62} \approx 2.149$$

$\Rightarrow$

$\Rightarrow \|E(x) - \mu\| \geq \|\bar{x} - \mu\| - \|\bar{x} - E(x)\|$

$$= 2.276 - \sqrt{4.62} \approx 2.276 - 2.149$$
$$= 0.127$$

$\Rightarrow \|E(x) - \mu\| \geq 0.127 > 0$

Thus, 0.127 is a lowerbound for $\|E(x) - \mu\|$. //

Moreover, since $\|\cdot\|$ is a norm, we have by definition that $\|E(x) - \mu\| \geq 0$,
and since $\|E(x) - \mu\| = 0$ iff $E(x) - \mu = 0$ iff $E(x) = \mu$
then we know that $\|E(x) - \mu\| > 0$ since it is not the case that $E(x) = \mu$ (strictly positive)

(b) In addition, suppose $\mu = [1 \;\; 0 \;\; 1]^T$ and $\bar{x} = [2.3 \;\; 1.8 \;\; 1.5]^T$.
What is $\|E(x) - \mu\|$? How does this compare with the lower bound in part (a)?

Solution: It is given that $\bar{x} - E(x) = \begin{bmatrix} 1.0 \\ 1.9 \\ -0.1 \end{bmatrix}$ and that

$$\bar{x} = \begin{bmatrix} 2.3 \\ 1.8 \\ 1.5 \end{bmatrix}$$

$$\Rightarrow E(x) = \bar{x} - \begin{bmatrix} 1.0 \\ 1.9 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.8 \\ 1.5 \end{bmatrix} - \begin{bmatrix} 1.0 \\ 1.9 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 1.3 \\ -0.1 \\ 1.6 \end{bmatrix}$$

We also know that $\mu = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$$\Rightarrow E(x) - \mu = \begin{bmatrix} 1.3 \\ -0.1 \\ 1.6 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.3 \\ -0.1 \\ 0.6 \end{bmatrix}$$

$$\Rightarrow \|E(x) - \mu\| = \sqrt{(0.3)^2 + (-0.1)^2 + (0.6)^2} = \sqrt{0.46}$$
$$\approx 0.678$$

This shows that 0.127 is indeed a lower bound for $\|E(x) - \mu\|$ $\Rightarrow$

$\|E(x)-\mu\| = \sqrt{0.46} \approx 0.678 > 0.127$

lower bound found in part (a)

This shows that 0.127 is indeed a lowerbound for $\|E(x)-\mu\|$.

The actual value compared to the LB: $\left|\frac{0.678-0.127}{0.678}\right| \times 100\%$

$= 81.3\%$ relative difference.

That is, the lower bound differs from the actual value by 81.3%.

In other words, 81.3% of the error cannot be explained by the model.

(c) Comment on whether the simulation appears to be a "good" model.

Answer: The lowerbound obtained in part (b) was obtained by the inequality

$$\|E(X)-\mu\| \geq \|\bar{X}-\mu\| - \|\bar{X}-E(X)\|$$

where $\|\bar{X}-\mu\|$ is the overall error of our model and $\|\bar{X}-E(x)\|$ can be interpreted as the systematic error. Then, the difference $\|\bar{X}-\mu\| - \|\bar{X}-E(x)\|$ is all of the error in the model excluding random error.

$\|E(X)-\mu\|$ is the validation/verification error, and is small when the model is valid or "good".

There is a 136.9% difference or 81.3% error model error, excluding all random error, which is quite large. Interpreted in this way, this might suggest that our model is not yet sufficient and that we should run additional simulations to decrease the validation error, $\|E(x)-\mu\|$.

Joseph High

If instead we compare the overall error $\|\bar{x} - \mu\|$ with the validation/verification error $\|E(x) - \mu\|$ where $\|\bar{x} - \mu\| = 2.276$ and $\|E(x) - \mu\| \approx 0.678$

$$\left|\frac{2.276 - 0.678}{2.276}\right| \times 100\% = 70.21\% \text{ difference}$$

Suggesting that 70.21% of the overall error in the model is not easily explained – which is a significant proportion. This suggests that the model is not "good" and that additional simulations should be run.

$\square$

Joseph High
Hopkins ID. 9E1FDC

Exercise 1.2 (Chapter 1, Rubinstein and Kroese)
Prove the product rule (1.4) for the case of three events.

Proof: (1.4) gives the product rule for the general finite case of n events.

We want to show that for any sequence of 3 events A_1, A_2, A_3, we have that

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1 A_2 A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 A_2).$$

For the events A_1 and A_2, we know that $\mathbb{P}(A_2 | A_1) = \dfrac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)}$

(*) $\Rightarrow \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 | A_1)$ (by def of conditional probability).

Moreover, we have by definition of the conditional probability of A_3 given the event $A_1 \cap A_2$:

$$\mathbb{P}(A_3 | A_1 \cap A_2) = \frac{\mathbb{P}(A_3 \cap (A_1 \cap A_2))}{\mathbb{P}(A_1 \cap A_2)}$$

$$= \frac{\mathbb{P}(A_1 \cap A_2 \cap A_3)}{\mathbb{P}(A_1) \cdot \mathbb{P}(A_2 | A_1)} \leftarrow \text{(from above (*))}$$

$$\Rightarrow \mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 | A_1) \cdot \mathbb{P}(A_3 | A_1 \cap A_2)$$

i.e. $\mathbb{P}(A_1 A_2 A_3) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 | A_1) \cdot \mathbb{P}(A_3 | A_1 A_2)$ // $\square$

Exercise 1.4 : Consider the random experiment where we toss a biased coin until heads comes up. Suppose that the probability of heads on any one toss is p. Let X be the number of tosses required. Show that $X \sim G(p)$. (Assume independent tosses)

Proof : [Reference : Example 1.1 in course textbook]

Suppose we must toss the biased coin n times before achieving a heads, with $n \in \mathbb{N}$.
Then, we are interested in the probability that the first head that appears is on the n^{th} toss i.e. $\mathbb{P}\{X = n\}$.

Now, let H_i be the event that the first heads is attained on the i^{th} trial, $i \in \{1, \ldots, n\}$.
Then, by the given information $\mathbb{P}(H_i) = p$

Clearly, H_i^c is the event of the coin landing on tails on the i^{th} trial. Letting $H_i^c = T_i \quad \forall i \in \{1, \ldots, n\}$

$$\Rightarrow \mathbb{P}(T_i) = \mathbb{P}(H_i^c) = 1 - \mathbb{P}(H_i) = 1 - p$$

Moreover, the event of attaining the first heads on the n^{th} toss of the biased coin is the event

$$\left(\bigcap_{i=1}^{n-1} H_i^c \right) \cap H_n = \left(\bigcap_{i=1}^{n-1} T_i \right) \cap H_n$$

That is,

$$\mathbb{P}\{X = n\} = \mathbb{P}\left[\left(\bigcap_{i=1}^{n-1} T_i \right) \cap H_n \right] = \mathbb{P}(T_1 \cap T_2 \cap \cdots \cap T_{n-1} \cap H_n)$$

since the tosses are independent $= \left(\prod_{i=1}^{n-1} \mathbb{P}(T_i) \right) \cdot \mathbb{P}(H_n) =$

$\Longrightarrow$

$$= \mathbb{P}(T_1)\,\mathbb{P}(T_2)\cdots\mathbb{P}(T_{n-1})\cdot\mathbb{P}(H_n)$$

$$= \underbrace{(1-p)\cdot(1-p)\cdots(1-p)}_{n-1 \text{ times}}\cdot p = (1-p)^{n-1}p$$

$$\Longrightarrow \mathbb{P}\{X=n\} = (1-p)^{n-1}p$$

Therefore, the probability density function for X is

$$f(x) = \begin{cases} (1-p)^{x-1}p & ; \ x = 1, 2, \ldots \\ 0 & ; \ \text{otherwise} \end{cases}$$

which is the pdf of a geometric random variable.

$$\Longrightarrow X \sim G(p) \qquad // \quad \square$$