Model Answers of Quiz II

1. (5 points) Let $y_1, \ldots, y_n\,(n \geq 2)$ be observed realization of a random sample. Suppose that we need to calculate $s^2 = \sum_{i=1}^{n} (y_i - \overline{y})^2$, where $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$. Show that the following algorithm can be used to compute s^2.

   $i = 1$;
   $\widehat{\mu}_1 = y_1$;
   $s_1 = 0$;
   for $i = 2$ **to** n **do**
   $\quad\delta_i = y_i - \widehat{\mu}_{i-1}$;
   $\quad\widehat{\mu}_i = \widehat{\mu}_{i-1} + \frac{\delta_i}{i}$;
   $\quad s_i = s_{i-1} + \frac{i-1}{i}\delta_i^2$;
   end
   return s_n

   Solution: We will prove it using mathematical induction. Note that for $n = 1$, $\widehat{\mu}_1 = y_1$ is the average of a sample of size 1 and $s_1 = 0 = (y_1 - \widehat{\mu}_1)^2$. Thus, the algorithm provides correct answer for $n = 1$. Now, assume that the algorithm provides correct answer for $n = k$. That means $\widehat{\mu}_k = \frac{1}{k}\sum_{i=1}^{k} y_i$ and $s_k = \sum_{i=1}^{k} (y_i - \widehat{\mu}_k)^2$. We need to prove that the algorithm provides correct answer for $n = k + 1$. Let us proceed as follows.

   $$\widehat{\mu}_{k+1} = \widehat{\mu}_k + \frac{\delta_{k+1}}{k+1} = \widehat{\mu}_k + \frac{1}{k+1}\left(y_{k+1} - \widehat{\mu}_k\right) = \frac{1}{k+1} y_{k+1} + \frac{k}{k+1}\widehat{\mu}_k = \frac{1}{k+1}\sum_{i=1}^{k+1} y_i.$$

   Also,

   $$\begin{aligned}
   s_{k+1} &= s_k + \frac{k}{k+1}\delta_{k+1}^2 \\
   &= \sum_{i=1}^{k}\left(y_i - \widehat{\mu}_k\right)^2 + \left(y_{k+1} - \widehat{\mu}_k\right)^2 - \left(y_{k+1} - \widehat{\mu}_k\right)^2 + \frac{k}{k+1}\left(y_{k+1} - \widehat{\mu}_k\right)^2 \\
   &= \sum_{i=1}^{k+1}\left(y_i - \widehat{\mu}_k\right)^2 - \frac{\delta_{k+1}^2}{k+1} \\
   &= \sum_{i=1}^{k+1}\left(y_i - \widehat{\mu}_{k+1} + \widehat{\mu}_{k+1} - \widehat{\mu}_k\right)^2 - \frac{\delta_{k+1}^2}{k+1} \\
   &= \sum_{i=1}^{k+1}\left(y_i - \widehat{\mu}_{k+1}\right)^2.
   \end{aligned}$$

   This completes the proof.

2. (5 points) Consider the importance sampling estimate of $\mu = E(X)$, where X has an exponential distribution with mean $\theta > 0$. Let the importance sampling probability density function

is of the form

$$q(x;\nu) = \begin{cases} \frac{1}{\nu}e^{-\frac{x}{\nu}} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\nu > 0$. Find the optimum choice of ν such that $\widehat{\mu}_{\text{imp}}$ has minimum variance.

Solution: Note that variance of $\widehat{\mu}_{\text{imp}}$ is minimum if $\int \frac{(fp)^2}{q}dx$ is minimum. Here $f(x) = x$, $p(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$ and $q(x) = \frac{1}{\nu}e^{-\frac{x}{\nu}}$ for $x > 0$. Thus,

$$\int_0^\infty \frac{(fp)^2}{q}dx = \frac{\nu}{\theta^2}\int_0^\infty x^2 \exp\left[-\left(\frac{2}{\theta} - \frac{1}{\nu}\right)\right] dx = \frac{2\theta\nu^4}{(2\nu-\theta)^3} \quad \text{if } \nu > \frac{\theta}{2}.$$

Therefore, we need to minimize $g(\nu) = \frac{\nu^4}{(2\nu-\theta)^3}$ with respect to ν over the interval $\nu > \frac{\theta}{2}$. Using standard technique, it can be shown that $g(\nu)$ attains its minimum at $\nu = 2\theta$. Hence, the optimum choice of ν is 2θ.

3. (5 points) Let $a \in [0,\, 0.99]$ be a real number. Consider the function $f : \mathbb{R} \to \{0, 10\}$ defined by

$$f(x) = \begin{cases} 10 & \text{if } a < x \le a + 0.01 \\ 0 & \text{otherwise.} \end{cases}$$

Consider the quantity of interest is $E(f(X))$, where $X \sim U(0,\, 1)$. Determine for what values of a antithetic sampling is helpful and harmful, respectively. You may do this by comparing the variances of $\widehat{\mu}$ under simple Monte Carlo and $\widehat{\mu}_{\text{anti}}$ under antithetic sampling.

Solution: For $U(0,\, 1)$ distribution, we know that the point of symmetry is $c = 0.5$, and hence, the opposite point of $x \in (0,\, 1)$ is $\tilde{x} = 1 - x$. Therefore,

$$E\left(f(X)\right) = E\left(f(\tilde{X})\right) = 10 \times 0.01 = 0.1.$$

We know that antithetic sampling is beneficial if $Cov\left(f\left(X\right),\, f(\tilde{X})\right) < 0$. Note that

$$f(\tilde{x}) = \begin{cases} 10 & \text{if } a < 1 - x < a + 0.01 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 10 & \text{if } 0.99 - a < x < 1 - a \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $E\left(f(X)f(\tilde{X})\right) = 0$ if the intersection of the intervals $[a,\, a + 0.01]$ and $[0.99 - a,\, 1 - a]$ is empty, and in this case, $Cov\left(f(X),\, f(\tilde{X})\right) = -0.01$. Now, the intervals $[a,\, a + 0.01]$ and $[0.99 - a,\, 1 - a]$ have non-empty intersection if

$$a \le 0.99 - a \le a + 0.01 \implies 0.490 \le a \le 0.495$$

or

$$a \le 1 - a \le a + 0.01 \implies 0.495 \le a \le 0.500.$$

For $0.490 \leq a \leq 0.495$,

$$\begin{aligned} E\left(f(X)f(\tilde{X})\right) &= 10^2 \times (a + 0.01 - 0.99 + a) \\ &= 100\,(2a - 0.98)\,. \end{aligned}$$

Therefore, antithetic sampling is beneficial if

$$Cov\left(f(X),\, f(\tilde{X})\right) = E\left(f(X)f(\tilde{X})\right) - 0.01 < 0 \implies a < 0.49005.$$

For $0.495 \leq a \leq 0.500$,

$$\begin{aligned} E\left(f(X)f(\tilde{X})\right) &= 10^2 \times (1 - a - a) \\ &= 100\,(1 - 2a)\,. \end{aligned}$$

Therefore, antithetic sampling is beneficial if

$$Cov\left(f(X),\, f(\tilde{X})\right) < 0 \implies a > 0.49995.$$

Hence, antithetic sampling is harmful if $0.49005 < a < 0.49995$ and helpful if $a \notin (0.49005,\, 0.49995)$.