MATH 565 Monte Carlo Methods

Fred J. Hickernell | Test 2 | Monday, November 3, 2025

Instructions:

i. *This test has* ***THREE*** *questions. Attempt them all. The maximum number of points is* ***100****.*
ii. *The time allowed is 75 minutes.*
iii. *This test is closed book, but you may use 4 double-sided letter-size sheets of notes.*
iv. *No calculators or other devices are allowed. Phones must be placed in your bags under your desks or face down on your desks. Hands mut be on top of your desks.*
v. *Keep at least* three *significant digits in your intermediate calculations and final answers.*
vi. *Show all your work to justify your answers. Answers without adequate justification will not receive credit.*
vii. *Off-site students may contact the instructor as directed by your syllabus.*

I understand these instructions and have not relied on any help for this exam beyond what is allowed, nor provided any help to anyone else beyond what is allowed.

______________________________ ____________________

Signature *Date*

1. (40 points) You want to approximate

$$\mu = \int_{[0,1]^2} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

for some $f : [0,1]^2 \to \mathbb{R}$ by IID Monte Carlo with control variates instead of the plain IID Monte Carlo estimator, which is $\hat{\mu} = \frac{1}{n}\sum_{i=0}^{n-1} f(\boldsymbol{X}_i)$ with $\boldsymbol{X} \sim \mathcal{U}[0,1]^2$.

a. (15 points) Does the following estimator

$$\hat{\mu}_{\mathrm{CV}} = \frac{1}{n}\sum_{i=0}^{n-1}[f(\boldsymbol{X}_i) - \beta_1 X_{i,1} - \beta_2 \sin(2\pi X_{i,2})], \qquad \boldsymbol{X}_i = (X_{i,1}, X_{i,2}) \overset{\mathrm{IID}}{\sim} \mathcal{U}[0,1],$$

utilizing the control variates X_1 and $\sin(2\pi X_2)$, satisfy $\mathbb{E}(\hat{\mu}_{\mathrm{CV}}) = \mu$ for all f and all $\boldsymbol{\beta} = (\beta_1, \beta_2)$? If not, adjust either or both of two control variates minimally to make $\mathbb{E}(\hat{\mu}_{\mathrm{CV}}) = \mu$.

Answer:

$$\begin{aligned}
\mathbb{E}(\hat{\mu}_{\mathrm{CV}}) &= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[f(\boldsymbol{X}_i) - \beta_1 X_{i,1} - \beta_2 \sin(2\pi X_{i,2})] \\
&= \int_{[0,1]^2}[f(\boldsymbol{x}) - \beta_1 x_1 - \beta_2 \sin(2\pi x_2)]\,\mathrm{d}\boldsymbol{x} \\
&= \int_{[0,1]^2} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \beta_1 \times \frac{1}{2} - \beta_2 \times 0
\end{aligned}$$

So the answer is, "No." We need to replace the control variate X_1 by $X_1 - 1/2$ so that $\mathbb{E}(X_1 - 1/2) = 0$. The correct estimator is

$$\hat{\mu}_{\text{CV}} = \frac{1}{n}\sum_{i=0}^{n-1}[f(\boldsymbol{X}_i) - \beta_1(X_{i,1} - 1/2) - \beta_2 \sin(2\pi X_{i,2})], \qquad \boldsymbol{X}_i = (X_{i,1}, X_{i,2}) \overset{\text{IID}}{\sim} \mathcal{U}[0,1],$$

b. (15 points) What is the formula for $\boldsymbol{\beta}_{\text{opt}}$, the optimal choice of $\boldsymbol{\beta}$ with properly chosen control variates according to part a.?

Answer: We choose $\boldsymbol{\beta}_{\text{opt}}$ by least squares regression to minimize $\text{var}\big(f(\boldsymbol{X}) - \beta_1(X_1 - 1/2) - \beta_2 \sin(2\pi X_2)\big)$. *Let*

$$\boldsymbol{y} = \begin{pmatrix} f(\boldsymbol{X}_0) - \hat{\mu} \\ \vdots \\ f(\boldsymbol{X}_{n-1}) - \hat{\mu} \end{pmatrix},$$

$$\mathsf{X} = \begin{pmatrix} X_{0,1} - \hat{\mu}_1 & \sin(2\pi X_{0,2}) - \hat{\mu}_2 \\ \vdots & \vdots \\ X_{n-1,1} - \hat{\mu}_1 & \sin(2\pi X_{n-1,2}) - \hat{\mu}_2 \end{pmatrix}, \qquad \hat{\mu}_1 = \frac{1}{n}\sum_{i=0}^{n-1} X_{i,1}, \quad \hat{\mu}_2 = \frac{1}{n}\sum_{i=0}^{n-1}\sin(2\pi X_{i,2})$$

Then the optimal $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta}_{\text{opt}} = (\mathsf{X}^\top \mathsf{X})^{-1}\mathsf{X}^\top \boldsymbol{y}.$$

c. (5 points) Is it possible that the variance of $\hat{\mu}_{\text{CV}}$ with properly chosen control variates according to part a. and for $\boldsymbol{\beta} = \boldsymbol{\beta}_{\text{opt}}$ is *larger* than the variance of the sample mean, $\hat{\mu}$?

Answer: No, since $\boldsymbol{\beta} = \boldsymbol{\beta}_{\text{opt}}$ must do no worse than $\boldsymbol{\beta} = \boldsymbol{0}$, which corresponds to $\mu_{\text{CV}} = \mu$.

d. (5 points) Is it possible that the variance of $\hat{\mu}_{\text{CV}}$ with properly chosen control variates according to part a.) is larger variance than the sample mean, $\hat{\mu}$ for some $\boldsymbol{\beta}$?

Answer: Yes, for $\boldsymbol{\beta}$ far from $\boldsymbol{\beta}_{\text{opt}}$.

You need to know how to take the expectation of an expression of a random vector.

2. (40 points) Consider a single queue with customers arriving with according to an exponential random variable with mean t_{arr} minutes and with a service time that is a uniform random variable with minimum 1 minute and maximum 5 minutes.

a. (5 points) How small can t_{arr} be so that the queue does grow infinitely long as time goes to infinity?

Answer: Since the mean service time is 3 minutes, then we need $t_{\text{arr}} > 3$.

b. (10 points) For the rest of this problem, assume that $t_{\mathrm{arr}} = 4$. What is the probability that two customers arrive within 1 minute of each other?

Answer: The cumulative distribution for the arrival time is defined as $F(t) = 1 - \exp(t/4)$, *so the probabilty of a customer arriving within one minute of the previous one is* $F(1) = 1 - \exp(1/4)$.

c. (20 points) Given the following times between the arrivals of successive customers and the times to serve customers in the queue, show the clock times and events until the last customer arrives. Assume that you start with an empty queue.

time between customer arrivals	2	1	6	5
service times of customers	2	4	2	3

Answer:

Step	*time, t*	*Event*	*In system*	*Time to next arrival*	*Remaining service time*
0	*0.0*	*Start*	*0*	*2.0*	*0.0*
1	*2.0*	*Arrival*	*1*	*1.0*	*2.0*
2	*3.0*	*Arrival*	*2*	*6.0*	*1.0*
3	*4.0*	*Departure*	*1*	*5.0*	*4.0*
4	*8.0*	*Departure*	*0*	*1.0*	*0.0*
5	*9.0*	*Arrival*	*1*	*5.0*	*2.0*
6	*11.0*	*Departure*	*0*	*3.0*	*0.0*
7	*14.0*	*Arrival*	*1*	—	*3.0*

d. (5 points) How many other customers are in the queue when the last customer arrives?

Answer: None. All others have been served.

3. (20 points) Consider the expected squared discrepancy—defined by kernel K—between a single random point, $\boldsymbol{X}_0$ and a point cloud, $\{\boldsymbol{Z}_i\}_{i=0}^{n-1}$, where $\boldsymbol{X}_0, \boldsymbol{Z}_0, \boldsymbol{Z}_1, \ldots$ are IID from a distribution with probability density ϱ. What is the formula for this expected squared discrepancy in terms of K, ϱ, and n?

Answer: The formula for the squared discrepancy is

$$D^2\big(\{\boldsymbol{X}_0\}, \{\boldsymbol{Z}_i\}_{i=0}^{n-1}; K\big) = K(\boldsymbol{X}_0, \boldsymbol{X}_0) - \frac{2}{n}\sum_{i=0}^{n-1} K(\boldsymbol{X}_0, \boldsymbol{Z}_i) + \frac{1}{n^2}\sum_{i,j=0}^{n-1} K(\boldsymbol{Z}_i, \boldsymbol{Z}_j)$$

which means that the mean squared discrepancy is

$$\mathbb{E}\big[D^2\big(\{\boldsymbol{X}_0\}, \{\boldsymbol{Z}_i\}_{i=0}^{n-1}; K\big)\big] = \mathbb{E}[K(\boldsymbol{X}_0, \boldsymbol{X}_0)] - \frac{2}{n}\sum_{i=0}^{n-1} \mathbb{E}[K(\boldsymbol{X}_0, \boldsymbol{Z}_i)] \\ + \frac{1}{n^2}\sum_{i,j=0}^{n-1} \mathbb{E}[K(\boldsymbol{Z}_i, \boldsymbol{Z}_j)]$$

The tricky part is $\mathbb{E}[K(\boldsymbol{Z}_i, \boldsymbol{Z}_j)]$, *where the answer depends on whether* $i = j$ *or not. Note that* $\mathbb{E}[K(\boldsymbol{X}, \boldsymbol{Z})] = \int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{z})\varrho(\boldsymbol{x})\varrho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{z}$ *if* $\boldsymbol{X}$ *and* $\boldsymbol{Z}$ *are independent and* $\mathbb{E}[K(\boldsymbol{X}, \boldsymbol{X})] = \int_{\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{x})\varrho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. *Then*

$$
\begin{aligned}
&\mathbb{E}\big[D^2\big(\{\boldsymbol{X}_0\}, \{\boldsymbol{Z}_i\}_{i=0}^{n-1}; K\big)\big] \\
&\qquad = \int_{\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{x})\varrho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{2}{n}\sum_{i=0}^{n-1}\int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{z})\varrho(\boldsymbol{x})\varrho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{z} \\
&\qquad + \frac{1}{n^2}\left\{\sum_{i=0}^{n-1}\int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{x})\varrho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \sum_{\substack{i,j=0\\ i\neq j}}^{n-1}\int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{z}, \boldsymbol{z}')\varrho(\boldsymbol{x})\varrho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{z}'\right\} \\
&\qquad = \left(1 + \frac{1}{n}\right)\int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{x})\varrho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \left(-2 + \frac{n^2 - n}{n^2}\right)\int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{z})\varrho(\boldsymbol{x})\varrho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{z} \\
&\qquad = \left(1 + \frac{1}{n}\right)\left[\int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{x})\varrho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\mathcal{X}\times\mathcal{X}} K(\boldsymbol{x}, \boldsymbol{z})\varrho(\boldsymbol{x})\varrho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{z}\right]
\end{aligned}
$$