

This exam is open notes and open book. You have three hours. There are a total of 57 points. Good luck!

**Problem 1 (16 pts):** Suppose that we are testing a new product that will be marketed in six months. We start operating  $n$  prototypes of the product but will need to terminate the testing in four months (in order to deal with last-minute re-design issues, etc.). We model the test data as  $T_1 \wedge t, T_2 \wedge t, \dots, T_n \wedge t$ , where  $T_i$  is the lifetime of prototype  $i$  (in months),  $t = 4$  months,  $a \wedge b = \min\{a, b\}$ , and we assume, on the basis of past experience, that the  $T_i$ 's are iid and exponentially distributed with (unknown) parameter  $\lambda^*$ . Suppose  $n = 5$  and  $T_i \wedge t = 1, 3, 4, 1, 2$  for  $1 \leq i \leq 5$ .

1. Compute the maximum likelihood estimator  $\hat{\lambda}_5$ . **(4 Points)**

**Solution:** The likelihood of  $T_1 \wedge t, \dots, T_n \wedge t$  is

$$L_n(\lambda) = \prod_{i=1}^n (\lambda e^{-\lambda T_i} \mathbb{1}_{\{T_i < t\}} + e^{-\lambda t} \mathbb{1}_{\{T_i \geq t\}}) = \lambda^{N_n} e^{-\lambda \sum_{i=1}^n (T_i \wedge t)}$$

where

$$N_n = \sum_{j=1}^n \mathbb{1}_{\{T_j < t\}}.$$

So, the log-likelihood is

$$\mathcal{L}_n(\lambda) = \log L_n(\lambda) = N_n \log \lambda - \lambda \sum_{i=1}^n (T_i \wedge t).$$

It follows that the maximum likelihood estimator  $\hat{\lambda}_n$  satisfies

$$0 = \mathcal{L}'_n(\hat{\lambda}_n) = \frac{N_n}{\hat{\lambda}_n} - \sum_{i=1}^n (T_i \wedge t).$$

So,

$$\hat{\lambda}_n = \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i < t\}}}{\sum_{i=1}^n (T_i \wedge t)}.$$

For our data set,

$$\hat{\lambda}_5 = \frac{4}{11}$$

2. Show that

$$\hat{\lambda}_n \stackrel{\mathcal{D}}{\approx} \lambda^* + n^{-\frac{1}{2}} N(0, \sigma^2(\lambda^*))$$

for large  $n$  and compute  $\sigma^2(\lambda^*)$ . **(6 Points)**

**Solution:** Note that

$$\hat{\lambda}_n = \frac{\bar{Y}_n}{\bar{\beta}_n},$$

where  $Y_i = \mathbb{1}_{\{T_i < t\}}$  and  $\beta_i = (T_i \wedge t)$ . Observe that if the true underlying exponential parameter is  $\lambda^*$ , then

$$E Y = P \{T_i < t\} = 1 - e^{-\lambda^* t}$$

and

$$\begin{aligned} \mathbf{E}\beta &= \mathbf{E}(T_i \wedge t) = \int_0^t \lambda^* e^{-\lambda^* s} ds + t \mathbf{P}\{T_i > t\} = - \int_0^t sd(e^{-\lambda^* s}) + te^{-\lambda^* t} \\ &= - \left[ se^{-\lambda^* s} \right]_0^t + \int_0^t e^{-\lambda^* s} ds + te^{-\lambda^* t} = \int_0^t e^{-\lambda^* s} ds = \frac{1}{\lambda^*} (1 - e^{-\lambda^* t}). \end{aligned}$$

So,

$$\frac{\mathbf{E}Y}{\mathbf{E}\beta} = \lambda^*.$$

Hence,

$$\hat{\lambda}_n - \lambda^* = \frac{\bar{Y}_n}{\bar{\beta}_n} - \frac{\mathbf{E}Y_1}{\mathbf{E}\beta_1} = \frac{\bar{Z}_n}{\bar{\beta}_n},$$

where  $Z_i = Y_i - \lambda^* \beta_i$ . Since

$$\bar{\beta}_n \rightarrow \mathbf{E}\beta_1 \quad \text{a.s.}$$

as  $n \rightarrow \infty$ , it follows that

$$\frac{\mathbf{E}\beta_1}{\bar{\beta}_n} \rightarrow 1 \quad \text{a.s.}$$

as  $n \rightarrow \infty$ . Also, note that  $\mathbf{E}Z_i = 0$ . So, the central limit theorem guarantees that

$$n^{\frac{1}{2}} \bar{Z}_n \Rightarrow \sigma_Z \mathbf{N}(0, 1)$$

as  $n \rightarrow \infty$ . Hence,

$$n^{\frac{1}{2}} (\hat{\lambda}_n - \lambda^*) = n^{\frac{1}{2}} \frac{\bar{Z}_n}{\bar{\beta}_n} = n^{\frac{1}{2}} \frac{\bar{Z}_n}{\mathbf{E}\beta_1} \cdot \frac{\mathbf{E}\beta_1}{\bar{\beta}_n} \Rightarrow \frac{\sigma_Z}{\mathbf{E}\beta_1} \mathbf{N}(0, 1) \cdot 1 = \sigma(\lambda^*) \mathbf{N}(0, 1)$$

where

$$\sigma^2(\lambda^*) = \frac{\mathbf{E}Z_1^2}{(\mathbf{E}\beta_1)^2} = \frac{\lambda^{*2} \mathbf{E}Z_1^2}{(1 - e^{-\lambda^* t})^2}$$

But

$$\begin{aligned} \mathbf{E}Z_1^2 &= \mathbf{E} \left[ (\mathbb{1}_{\{T_i < t\}} - \lambda^*(T_i \wedge t))^2 \right] = \mathbf{P}\{T_i < t\} - 2\lambda^* \mathbf{E}[\mathbb{1}_{\{T_i < t\}}(T_i \wedge t)] + \lambda^{*2} \mathbf{E}(T_i \wedge t)^2 \\ &= 1 - e^{-\lambda^* t} - 2\lambda^* \mathbf{E}[\mathbb{1}_{\{T_i < t\}} T_i] + \lambda^{*2} \mathbf{E}(T_i \wedge t)^2 \end{aligned}$$

Of course,

$$\mathbf{E}[\mathbb{1}_{\{T_i < t\}} T_i] = \int_0^t s \lambda^* e^{-\lambda^* s} ds = - \int_0^t sd(e^{-\lambda^* s}) = - \left[ se^{-\lambda^* s} \right]_0^t + \int_0^t e^{-\lambda^* s} ds = -te^{-\lambda^* t} + \frac{1 - e^{-\lambda^* t}}{\lambda^*}.$$

Also,

$$\begin{aligned} \mathbf{E}(T_i \wedge t)^2 &= \int_0^t s^2 \lambda^* e^{-\lambda^* s} ds + t^2 \mathbf{P}\{T_i > t\} = - \int_0^t s^2 d(e^{-\lambda^* s}) + t^2 e^{-\lambda^* t} \\ &= - \left[ s^2 e^{-\lambda^* s} \right]_0^t + 2 \int_0^t se^{-\lambda^* s} ds + t^2 e^{-\lambda^* t} = \frac{2}{\lambda^*} \mathbf{E}[\mathbb{1}_{\{T_i < t\}} T_i] = -\frac{2t}{\lambda^*} e^{-\lambda^* t} + \frac{2}{\lambda^{*2}} (1 - e^{-\lambda^* t}). \end{aligned}$$

So

$$\mathbf{E}Z_1^2 = 1 - e^{-\lambda^* t} + 2\lambda^* t e^{-\lambda^* t} - 2(1 - e^{-\lambda^* t}) - 2\lambda^* t e^{-\lambda^* t} + 2(1 - e^{-\lambda^* t}) = 1 - e^{-\lambda^* t}.$$

Hence,

$$\sigma^2(\lambda^*) = \frac{\lambda^{*2}}{1 - e^{-\lambda^* t}}.$$

3. Use the result in part 2 to provide an approximate 95% confidence interval for  $\lambda^*$ . **(2 Points)**

**Solution:** A large-sample 95% confidence interval for  $\lambda^*$ , based on the above central limit theorem, is

$$\left[ \hat{\lambda}_n - z \frac{\sigma(\hat{\lambda}_n)}{\sqrt{n}}, \hat{\lambda}_n + z \frac{\sigma(\hat{\lambda}_n)}{\sqrt{n}} \right]$$

where  $P\{-z \leq N(0, 1) \leq z\} = .95$ , (i.e.  $z = 1.96$ ) and

$$\sigma(\hat{\lambda}_n) = \sqrt{\frac{\hat{\lambda}_n^2}{1 - e^{-\hat{\lambda}_n t}}}.$$

4. Discuss how you would apply the bootstrap algorithm to compute a confidence interval for  $\lambda^*$ . **(4 Points)**

**Solution:** An alternative to the previous solution would be

1. Compute

$$\hat{\lambda}_n = \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i < t\}}}{\sum_{i=1}^n (T_i \wedge t)}.$$

2. Sample  $n$  iid observations from an  $\text{Exp}(\hat{\lambda}_n)$  population; call the  $n$  observations  $T_1^*, \dots, T_n^*$ .

3. Compute

$$\hat{\lambda}_n^* = \frac{\sum_{i=1}^n \mathbb{1}_{\{T_i^* < t\}}}{\sum_{i=1}^n (T_i^* \wedge t)}.$$

4. Repeat steps 2. and 3.  $m$  independent times, thereby computing  $m$  iid copies  $\hat{\lambda}_{1n}^*, \dots, \hat{\lambda}_{mn}^*$  of the rv  $\hat{\lambda}_n^*$ .

5. Find a value  $\tilde{z}$  such that

$$\frac{1}{m} \sum_{i=1}^m \mathbb{1}_{\{-\tilde{z} \leq \hat{\lambda}_n - \hat{\lambda}_{in}^* \leq \tilde{z}\}} \approx 0.95.$$

6. The final bootstrap confidence interval is

$$[\hat{\lambda}_n - \tilde{z}, \hat{\lambda}_n + \tilde{z}].$$

**Problem 2 (10 pts):** Suppose that the position of a particle is described by a one-dimensional standard Brownian motion process  $B = (B(t) : t \geq 0)$  starting from  $B(0) = 0$ , with  $EB(t) = 0$  and  $\text{Var}(B(t)) = t$ . Because of measurement discretization, we observe

$$Z_n = \lfloor B(n) \rfloor,$$

where  $\lfloor x \rfloor$  is the greatest integer less than or equal to  $x$ .

1. Provide a formula for the conditional density of  $B(n)$ , conditional on  $Z_0, \dots, Z_n$ . **(7 Points)**

**Solution:** Let  $\phi(x)dx = P\{N(0, 1) \in dx\}$  so that  $\phi(\cdot)$  is the density of a  $N(0, 1)$  rv.

Conditional on  $Z(0) = 0, Z(1) = i_1, \dots, Z(n) = i_n$  (where  $i_j \in \mathbb{Z}$  for  $1 \leq j \leq n$ ),

$$\begin{aligned} & P \{B(1) \in dx_1, \dots, B(n) \in dx_n | Z(0) = 0, Z(1) = i_1, \dots, Z(n) = i_n\} \\ &= \begin{cases} \frac{\phi(x_1)\phi(x_2-x_1)\cdots\phi(x_n-x_{n-1})dx_1\cdots dx_n}{\int_{i_1}^{i_1+1} \int_{i_2}^{i_2+1} \cdots \int_{i_n}^{i_n+1} \phi(y_1)\phi(y_2-y_1)\cdots\phi(y_n-y_{n-1})dy_1\cdots dy_n} & i_j \leq x_i \leq i_j + 1, \quad 1 \leq j \leq n \\ 0 & \text{o.w.} \end{cases} \end{aligned}$$

It follows that

$$\begin{aligned} & P \{B(n) \in dx_n | Z(0) = 0, Z(1) = i_1, \dots, Z(n) = i_n\} \\ &= \frac{\int_{i_1}^{i_1+1} \int_{i_2}^{i_2+1} \cdots \int_{i_{n-1}}^{i_{n-1}+1} \phi(y_1)\phi(y_1-y_2)\cdots\phi(y_{n-1}-y_{n-2})\phi(x_n-y_{n-1})dy_1\cdots dy_{n-1}}{\int_{i_1}^{i_1+1} \cdots \int_{i_n}^{i_n+1} \phi(y_1)\phi(y_2-y_2)\cdots\phi(y_n-y_{n-1})dy_1\cdots dy_n}. \end{aligned}$$

2. Provide a recursive (in  $n$ ) updating formula for the conditional density of the position  $B(n)$ , conditional on  $Z_0, Z_1, \dots, Z_n$ . **(3 Points)**

**Solution:** Let

$$\nu_n(x)dx = P \{B(n) \in dx | Z(0), \dots, Z(n)\}.$$

So

$$\nu_n(x) = \frac{\int_{Z(n-1)}^{Z(n-1)+1} \nu_{n-1}(y)\phi(x-y)\mathbb{1}_{\{Z(n)\leq x \leq Z(n)+1\}}dy}{\int_{Z(n-1)}^{Z(n-1)+1} \int_{Z(n)}^{Z(n)+1} \nu_{n-1}(y)\phi(z-y)dzdy}.$$

**Problem 3 (9 pts):** In many communication networking settings, the integrated Ornstein-Uhlenbeck process is used as a model of the aggregate traffic to arrive by time  $t$ . In particular, suppose that

$$Z(t) = t + \int_0^t X(s)ds,$$

where  $X = (X(t) : t \geq 0)$  is a stationary Gaussian process with  $E X(t) = 0$  and  $\text{Cov}(X(0), X(t)) = \sigma^2 e^{-t}$ .

1. Compute the covariance function of  $Z$ . [Hint: Note that  $Z$  is affine in  $X$ .] **(3 Points)**

**Solution:** Note that for  $s < t$

$$\begin{aligned} \text{Cov}(Z(s), Z(t)) &= \text{Cov}\left(s + \int_0^s X(r)dr, t + \int_0^t X(u)du\right) \\ &= \int_0^s \int_0^t \text{Cov}(X(r), X(u)) dudr \quad (*) \\ &= \int_0^s \int_{-r}^{t-r} \text{Cov}(X(0), X(v)) dvdr \\ &= \sigma^2 \int_0^s \left( \int_{-r}^0 e^v dv + \int_0^{t-r} e^{-v} dv \right) dr \\ &= \sigma^2 \int_0^s (1 - e^{-r} + 1 - e^{-(t-r)}) dr \\ &= \sigma^2 (2s - (1 - e^{-s}) - e^{-t}(e^s - 1)) \\ &\triangleq \sigma^2 \gamma(s, t) \end{aligned}$$

**Remark:** Equation (\*) follows directly or by recognizing that

$$\int_0^t X(s)ds = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} X\left(\frac{t}{n}\right) \frac{1}{n}$$

and noting that

$$\text{Cov}\left(\sum_{i=0}^{n-1} X\left(\frac{it}{n}\right) \frac{1}{n}, \sum_{j=0}^{n-1} X\left(\frac{jt}{n}\right) \frac{1}{n}\right) = \frac{1}{n^2} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \text{Cov}\left(X\left(\frac{it}{n}\right), X\left(\frac{jt}{n}\right)\right) \rightarrow \int_0^t \int_0^s \text{Cov}(X(r), X(u)) drdu$$

as  $n \rightarrow \infty$ .

2. Does  $Z$  have differentiable trajectories? Why or why not? **(2 Points)**

**Solution:** Note that  $X$  is continuous and  $Z$  is the integral of a continuous function. Hence,  $Z$  is (continuously) differentiable.

3. Compute the conditional distribution of  $Z(1/2)$ , given that  $Z(1) = z$ . **(2 Points)**

**Solution:** Note that for any  $s, t$ , we have

$$\mathbb{E}[Z(t)|Z(s)] = t + \frac{\gamma(s, t)}{\gamma(s, s)}(Z(s) - s)$$

so

$$\mathbb{E}[Z(1/2)|Z(1) = z] = \frac{1}{2} + \frac{\gamma(1/2, 1)}{\gamma(1, 1)}(z - 1).$$

4. Compute the conditional distribution of  $Z(3/2)$ , given that  $Z(1) = z_1$  and  $Z(2) = z_2$ . **(2 Points)**

**Solution:** Define  $X_3 = [X(1), X(2), X(3/2)]^T$  and

$$\Sigma = \mathbb{E}[X_3 X_3^T] - \mathbb{E}[X_3] \mathbb{E}[X_3^T] = \begin{pmatrix} \gamma(1, 1) & \gamma(1, 2) & \gamma(1, 3/2) \\ \gamma(1, 2) & \gamma(2, 2) & \gamma(2, 3/2) \\ \gamma(1, 3/2) & \gamma(3/2, 2) & \gamma(3/2, 3/2) \end{pmatrix} \triangleq \begin{pmatrix} C & b^T \\ b & c \end{pmatrix}.$$

Now we have that

$$\mathbb{E}[Z(3/2)|Z(1) = x, Z(2) = y] = \frac{3}{2} + b^T C^{-1} \begin{pmatrix} x \\ y \end{pmatrix}.$$

**Problem 4 (12 pts):** Suppose Ms. Smith owns a car but lives close enough to her job that she always walks whenever the weather is nice. Whenever it rains and the car is where Ms. Smith currently is, she will drive. On any given trip from / to work, it rains with probability  $p$  and is nice with probability  $1 - p$ , independent of the weather on the preceding trips.

1. We first wish to compute the long-run fraction of time that Ms. Smith gets wet.

- (a) Model this situation via a suitable discrete-time Markov chain. **(3 Points)**

**Solution:** There are several different ways to approach this problem.

*Approach 1:* Model this system as a two-state Markov chain:

State 1: Both the vehicle and Ms. Smith are in the same. location

State 2: Ms. Smith is in a different place than her car.

The transitions are given by the following transition matrix (the first row corresponds to state 1; the second row to state 2)

$$P_2 = \begin{pmatrix} p & 1-p \\ 1 & 0 \end{pmatrix}.$$

*Approach 2:* We model the system as a three-state Markov chain. Each state encodes both if Ms. Smith and the car are in the same location, and whether or not Ms. Smith was rained on in her previous trip. The states are

State 1: Same location, not rained on

State 2: Same location, rained on

State 3: Different location, not rained on

State 4: Different location, rained on (note that this state is not compatible with the dynamics of the system and is ignored).

The transition matrix is given by

$$P_3 = \begin{pmatrix} p & 0 & 1-p \\ p & 0 & 1-p \\ 1-p & p & 0 \end{pmatrix}.$$

- (b) Compute the long-run fraction of time that Ms. Smith gets wet. **(3 Points)**

**Solution:** We consider both approaches.

*Approach 1:* We need to compute the limit, as  $n \rightarrow \infty$ , of

$$\frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_{\{X_i=2, \text{ it rains on the } (i+1)\text{'th trip}\}}.$$

This converges, as  $n \rightarrow \infty$ , to

$$p\pi(2).$$

where  $\pi(2)$  is the equilibrium probability of being in state 2. We obtain  $\pi(1)$  and  $\pi(2)$  by solving

$$\begin{aligned} p\pi(1) + \pi(2) &= \pi(1) \\ (1-p)\pi(1) &= \pi(2) \end{aligned}$$

subject to the constrain that,

$$\pi(1) + \pi(2) = 1.$$

So

$$\pi(1) = \frac{1}{2-p} \quad \text{and} \quad \pi(2) = \frac{1-p}{2-p}.$$

So, the fraction of trips on which Ms. Smith is rained on is

$$\frac{p(1-p)}{2-p}.$$

*Approach 2:* Solve for the left, stochastic eigenvector  $\pi$  which satisfies  $\pi P = \pi$ . Then  $\pi(2)$  is the desired probability.

2. Now, suppose that Ms. Smith can choose to drive if she has a car, even if it is nice out. Suppose that there are certain costs to driving or not driving under different circumstances (e.g. if you walk in the rain, you will have to take your clothes to the dry cleaner at a cost of  $\ell$ ; if you drive you must pay for the gas at a cost of  $g < \ell$ ). For  $\alpha > 0$ , let

$$C = \sum_{i=0}^{\infty} e^{-\alpha i} c(X_i, a_i).$$

- (a) What equation would you solve to minimize

$$E_x [C],$$

(where  $x$  denotes starting at home (with the car, dry))? **(3 Points)**

**Solution:** Let  $V(i)$  be the value function for the optimal control problem, so

$$V(i) = \min_{A_n: n \geq 0} E_i \left[ \sum_{j=0}^{\infty} e^{-\alpha j} c(X_j, A_j) \right].$$

The HJB equation is

$$\begin{aligned} V(1) &= \min[g + e^{-\alpha}V(1), e^{-\alpha}V(2)] \\ V(2) &= p\ell + e^{-\alpha}V(1). \end{aligned}$$

Hence,

$$V(1) = \min[g + e^{-\alpha}V(1), p\ell e^{-\alpha} + e^{-2\alpha}V(1)].$$

In other words,

$$V(1) = \min \left[ \frac{g}{1 - e^{-\alpha}}, \frac{p\ell e^{-\alpha}}{1 - e^{-2\alpha}} \right].$$

The optimal policy is to always drive if

$$\frac{g}{1 - e^{-\alpha}} < \frac{p\ell e^{-\alpha}}{1 - e^{-2\alpha}}.$$

- (b) Suppose we want to ensure that Ms. Smith gets wet only 20% of the time. How would you compute the optimal policy? **(3 Points)**

**Solution:** Here we consider the the case with  $\alpha = 0$  (so that we minimize the long-run average cost). Here,

$\pi(1, d)$  is the long-run fraction of time under the optimal policy that Ms. Smith should choose to drive, given that she is in state 1.

$\pi(1, w)$  is the long-run fraction of time under the optimal policy that Ms. Smith should choose to walk, given that she is in state 1.

We wish to

$$\min \pi(1, d)g + \pi(2)p\ell$$

subject to

$$\begin{aligned} \pi(1, d) + \pi(1, w) &= \pi(1, d) + \pi(2) && \text{(rate into 1 = rate out of 1)} \\ \pi(2) &= \pi(1, w) && \text{(rate into 2 = rate out of 2)} \\ p\pi(2) &\leq 0.2 \\ \pi(1, d) + \pi(1, w) + \pi(2) &= 1 \\ \pi &\geq 0. \end{aligned}$$

This linear program characterizes the optimal policy. In particular, the fraction of time that one should drive when the vehicle and Ms. Smith are in the same location is

$$\pi(d|1) = \frac{\pi(1, d)}{\pi(1, d) + \pi(1, w)}.$$

**Problem 5 (10 pts):**

1. Prove that when  $n$  is large

$$\pi \approx \frac{4}{n} \sum_{j=1}^n \mathbb{1}_{\{U_{j1}^2 + U_{j2}^2 \leq 1\}}.$$

where the  $U_{j1}$  and  $U_{j2}$  are all iid  $U(0, 1)$  rvs. **(2 Points)**

**Solution:** Note that for an arbitrary subset  $B \subseteq [0, 1]^2$ ,

$$P\{(U_{j1}, U_{j2}) \in B\} = |B|.$$

If  $B^* = \{(x, y) \in [0, 1]^2 : x^2 + y^2 \leq 1\}$ ,  $B^*$  is one quarter of a full disc centered on the origin having radius 1. So, the area of  $B^*$  is  $1/4$  the area of a circle having unit radius, i.e.

$$|B^*| = \frac{\pi}{4}.$$

So,

$$4P\{(U_{j1}, U_{j2}) \in B^*\} = \pi.$$

Hence, by the Law of Large Numbers,

$$\frac{4}{n} \sum_{j=1}^n \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}} \rightarrow \pi$$

as  $n \rightarrow \infty$ . So, when  $n$  is large,

$$\frac{4}{n} \sum_{j=1}^n \mathbb{1}_{\{U_{j1}^2 + U_{j2}^2 \leq 1\}} \approx \pi.$$

2. What control variate would you use to decrease the variance of a Monte Carlo simulation of  $\pi$ ? **(2 Points)**

**Solution:** Many control variates are possible. For example, consider

$$\mathbb{1}_{\{U_{j1} + U_{j2} \leq 1\}} = \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}}$$

where  $\tilde{B} = \{(x, y) \in [0, 1]^2 : x + y \leq 1\}$ . So

$$|\tilde{B}| = \frac{1}{2},$$

and

$$C_j = \mathbb{1}_{\{U_{j1} + U_{j2} \leq 1\}} - \frac{1}{2}$$

is a control variate (having zero expectation).

3. Discuss how you would choose the control coefficient  $\alpha$ . **(1 Point)**

**Solution:** We use the estimator

$$\frac{4}{n} \sum_{j=1}^n \left[ \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}} - \alpha \left( \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}} - \frac{1}{2} \right) \right].$$

The optimal choice of  $\alpha$  is

$$\alpha^* = \frac{\text{Cov} \left( \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}}, \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}} \right)}{\text{Var} \left( \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}} \right)} = \frac{\text{Cov} \left( \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}}, \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}} \right)}{\frac{1}{2} \frac{1}{2}}.$$

But

$$\text{Cov} \left( \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}}, \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}} \right) = \mathbb{E} \left[ \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^* \cap \tilde{B}\}} \right] - \frac{\pi}{4} \frac{1}{2} = \frac{1}{2} - \frac{\pi}{8}$$

So,

$$\alpha^* = \frac{4 - \pi}{8}.$$

Since we are using the Monte Carlo method to compute  $\pi$ , we estimate  $\alpha^*$  via

$$\hat{\alpha}_n = \frac{4 - \hat{\pi}_n}{8},$$

where  $\hat{\pi}_n$  is the crude Monte Carlo estimator given by

$$\hat{\pi}_n = \frac{4}{n} \sum_{j=1}^n \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}}.$$

In other words, our control variate estimate for  $\pi$  is

$$\frac{4}{n} \sum_{j=1}^n \left[ \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}} - \hat{\alpha}_n \left( \mathbb{1}_{\{(U_{j1}, U_{j2}) \in \tilde{B}\}} - \frac{1}{2} \right) \right].$$

Note: There are many other reasonable estimators for  $\alpha^*$ .

4. Compute  $\mathbb{E}[Y|U_1]$  where  $Y = \mathbb{1}_{\{U_1^2 + U_2^2 \leq 1\}}$  where  $U_1$  and  $U_2$  are independent  $U(0, 1)$  rvs. **(2 Points)**

**Solution:**

$$\mathbb{P} \{U_1^2 + U_2^2 \leq 1 | U_1 = u\} = \mathbb{P} \{U_2^2 \leq 1 - u^2\} = \mathbb{P} \{U_2 \leq \sqrt{1 - u^2}\} = \sqrt{1 - u^2}.$$

So,

$$\mathbb{E}[Y|U_1] = \sqrt{1 - U_1^2}.$$

5. Show that for any rv  $Z$  such that  $\text{Var}(Z) < \infty$ , and any other rv  $W$ , that

$$\text{Var}(\mathbb{E}[Z|W]) \leq \text{Var}(Z).$$

**(2 Points)**

**Solution:** Note that the rv  $Z - E[Z|W]$  is orthogonal in  $L^2$  to  $E[Z|W]$ . So if  $E Z = 0$ ,

$$\begin{aligned} \text{Var}(Z) = E Z^2 &= E [(Z - E[Z|W]) + E[Z|W]]^2 = E [(Z - E[Z|W])^2] + E [(E[Z|W])^2] \\ &= \text{Var}(Z - E[Z|W]) + \text{Var}(E[Z|W]). \end{aligned}$$

Hence,

$$\text{Var}(E[Z|W]) \leq \text{Var}(Z).$$

If  $E Z \neq 0$ , pass to  $\tilde{Z} = Z - E Z$  and apply the above to  $\tilde{Z}$ .

6. Provide an alternative estimator for  $\pi$ , with lower variance than  $Y$ , based on parts 4 and 5 above. (1 Point)

**Solution:** Use

$$\frac{4}{n} \sum_{j=1}^n \sqrt{1 - U_j^2}$$

in place of

$$\frac{4}{n} \sum_{j=1}^n \mathbb{1}_{\{(U_{j1}, U_{j2}) \in B^*\}}.$$

(Note that we only need to generate half the uniforms in this method.)