9.3

a) $\theta = P(\sum_{i=1}^{5} iX_i \geq 21.6)$, $X_i \sim exp(1)$

so $E[\theta] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}(\sum i^{(k)} X_i^{(k)} \geq 21.6))$ $i^{(k)} = i \ \forall k$

to estimate θ we do the following

n times
1) $X = [\,]$
2) $U_i \sim unif(0,1)$ generation for $i = \{1,2,3,4,5\}$
3) if $\sum_{i=1}^{5} -i\log(U_i) \geq 21.6$, X.append(1)
   else X.append(0)
4) return $\frac{sum(X)}{len(X)}$

b) to estimate θ using antithetic variables we do the following

1) $X = [\,]$
n times
2) generate $U_i \sim unif(0,1)$ $\forall i$
3) set $X_i^{(1)} = -\log(U_i)$, $X_i^{(2)} = -\log(1-U_i)$
4) if $\sum_{i=1}^{5} iX_i^{(1)} \geq 21.6$ or $\sum_{i=1}^{5} iX_i^{(2)} \geq 21.6$, X.append(1)
   else X.append(0)
5) return $\frac{sum(X)}{len(X)}$

c) this is not efficient as variance is reduced by only a factor of ~1.006 after 10K runs of each estimator generating 1k samples.

this implements antithetic estimator

$$\overline{E[\theta]} = \frac{1}{2}\sum_{k=1}^{n} \mathbb{1}(\sum_{i=1}^{5} i^{(k)} X_i^{(1),k} \geq 21.6) + \mathbb{1}(\sum_{i=1}^{5} i^{(k)} X_i^{(2),k} \geq 21.6)$$

c.

Raw E: 0.038569, Antithetic E: 0.038298
Raw Var: 3.7e-05
Antithetic Var: 1.8e-05, Reduction over Raw: 2.0634
The use of antithetic variables is effective.

9.10

a. First note $E[I] = P(X \le a) = F(a)$, $F = P(X \le x)$

$$E[XI] = \int_0^a x f(x)\,dx$$

so $cov(X, I) = E[XI] - E[X]E[I]$

$$= \int_0^a x f(x)\,dx - F(a)\int_0^\infty x f(x)\,dx$$

and $var(X) = E[X^2] - E[X]^2 = \int x^2 f(x)\,dx - \left(\int x f(x)\,dx\right)^2$

$var(I) = E[I^2] - E[I]^2 = E[I](1 - E[I])$

$= F(a)(1 - F(a))$

Then b/c % reduction of var =

$$= \frac{var(I) + c^2 var(X) + cov(I, X)}{var(I)}$$

$$= \frac{var(I)(1 - corr^2(I, X))}{var(I)} = 1 - corr^2(I, X) = \frac{cov^2(I, X)}{var(I)\,var(X)}$$

so if $X \sim unif(0,1)$, $F(X) = X$, $f(x) = 1$,

$cov(X, I) = \int_0^a x\,dx - a\int_0^1 x\,dx = \frac{a^2 - a}{2}$

$var(X) = \int_0^1 x^2\,dx - \left(\int_0^1 x\,dx\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$

$var(I) = a(1-a)$

thus $corr^2(X, I) = 3a(1-a) \Rightarrow$ % reduction $= 1 - 3a(1-a)$

9.10. cont.

b. if $X \sim exp(1)$ then $F(X) = 1 - e^{-x}$, $f(x) = e^{-x}$

then $$cov(I, X) = \int_0^a x e^{-x} dx - (1 - e^{-a}) \int_0^\infty x e^{-x} dx$$

$$= -ae^{-a} - \int_0^\infty -e^{-x} dx - (1-e^{-a}) \left(\cancel{\infty} e^{-\infty} - \int_0^\infty -e^{-x} dx \right)$$

$$= -ae^{-a} - e^{-a} + 1 - 1 + e^{-a}$$

$$= -ae^{-a}$$

$$var(X) = \frac{1}{\lambda^2} = 1$$

$$var(I) = Fa(1 - F(a)) = 1 - e^{-a}(1 - (1 - e^{-a}))$$

$$= e^{-a}(1 - e^{-a})$$

and the % reduction in var $= 1 - corr^2(I, X)$

and % reduction $= \boxed{1 - \frac{a^2}{e^a - 1}} = 1 - \frac{(-ae^{-a})^2}{e^{-a}(1 - e^{-a})}$

C. since $1 - corr^2(I, X) = \frac{cov^2(I, X)}{var(I)\, var(I)}$ all we need to show is $cov(I, X) < 0$

$$cov(I, X) = E[XI] - E[X]E[I] = 1 - F(a) \int_0^a x f(x) dx - F(a) \int_0^\infty x f(x) dx$$

$$= -(1 - F(a)) \int_0^a F(x) dx + aF(a)(1 - F(a)) - F(a) \int_a^L x f(x) dx$$

$$= -(1 - F(a)) \int_0^a F(x) dx + aF(a)(1 - F(a)) - (1 - F(a)) \bar{X} F(a), \quad \bar{X} = E[X|F]$$

$$= -(1 - F(a)) \int_0^a F(x) dx + (\bar{X} - a) F(a)$$

then b/c $F(a) < 1$, $F(x) \geq 0$, $\bar{X} \geq a$ we must conclude

$$cov(I, X) < 0 \Rightarrow corr(I, X) = \frac{cov(I, X)}{\sqrt{var(I)\, var(X)}} < 0$$

9.18

$Y \sim N(1,1)$, $X \sim N(y,4)$, $y = Y$

so $X_{Y=y} \sim N(y,4)$, $\theta = P(X>1)$

a) raw estimator $E[\theta] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}(X>1)$ we do the following

1) $X = [\,]$
2) Gen $y = Y \sim N(1,1)$
3) Gen $X_i = X \sim N(y,4)$
4) if $X < 1$, X.append(1)
   else, X.append(0)

$\left.\right\}$ n times

5) return $\frac{\text{sum}(X)}{\text{len}(X)}$

b) First note $P(X>1 \mid Y=y_i) = P\left(\frac{X-y_i}{2} > \frac{1-y_i}{2} \mid y_i\right) = P\left(\frac{X-y_i}{2} > Z_{y_i} \mid y_i\right)$

$= P_{Y=y}(X>1 \mid y)$

so the conditional estimator is

$$E[\theta_c] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}\left(\frac{X-y_i}{2} > Z_{y_i}\right) =$$

and do the following to simulate it

1) $X = [\,]$
2) Gen $y = Y \sim N(1,1)$
3) find Z_y (z-score of y) $= \frac{y-1}{1}$
4) find $A_{Z_y} = P(X > y)$

$\left.\right\}$ n times

5) X.append(A_{Z_y})
6) average over X

9.18. cont.

c) using antithetic variables and conditioning we get an estimator of

$$E[\theta_{c,a}] = \frac{1}{2}\sum_{i=1}^{n} \mathbb{1}\left(\frac{X_k - y_k^{(1)}}{2} > Z_{y(1)}^{(k)}\right) + \mathbb{1}\left(\frac{X_k - y_k^{(2)}}{2} > Z_{y(2)}^{(k)}\right)$$

which we simulate by the following

1) $X = []$
2) Gen $y_1 \sim N(1, \frac{1}{2})$, $y_2 \sim N(0, \frac{\sqrt{3}}{2})$
3) set $A = y_1 + y_2$, $B = y_1 - y_2$
4) find Z_A, Z_B (respective z-scores)
5) find $A_{Z_A} = P(X > Z_A)$, $B_{Z_B} = P(X > Z_B)$
6) X.append($A_{Z_A} + B_{Z_B}$)
7) average over X
8) return $\frac{\text{avg}(X)}{2}$

n times

note that for $y_1 \sim N(1, \frac{1}{2})$, $y_2 \sim N(0, \frac{\sqrt{3}}{2})$ that $E[y_1 + y_2] = 1$, $\text{cov}(y_1 + y_2) < 0$; $E[y_1 - y_2] = 1$, $\text{cov}(y_1 - y_2) < 0$

d) we can use a control variate of Y so $E[Y] = 1$ and our new estimator is

$$E[\theta_{c,c}] = \sum_{i=1}^{n} A_{Z_{y_n}} - \gamma(y_n - 1) \quad \text{where } \gamma = \text{cov}(Z_y, y)$$

we find γ empirically and simulate the entire process by

1) $X = []$
2) Gen $y_n \sim N(1,1)$
3) find Z_y
4) find $A_{Z_{y_n}}$
6) X.append($A_{Z_{y_n}} - \gamma(y_n - 1)$)
7) return avg(X)

n times

1a) y_cov=[], Z_cov=[]
1b) y_cov.append($y \sim N(1,1)$)
1c) Z_cov.append(A_{Z_y})
1d) γ = cov(y_cov, Z_cov)

n times

e) f) g) h)
Raw E: 0.50137, Conditional E: 0.498847
Conditional + Antithetic Var E: 0.499851, Conditional + Control Var E: 0.500073

Raw Var: 0.250001
Conditional Var: 0.083237, Reduction over Raw: 3.0035
Conditional + Antithetic Var: 0.041684, Reduction over Conditional: 5.9975
Conditional + Control Var: 0.003683, Reduction over Conditional: 67.8849

i) Exact value of $\theta = 0.5$

9.19

a) $U \sim unif(0,1)$, $C \sim poisson(\frac{15}{.5+U})$

$I = \begin{cases} 1 & \text{if } C \geq 20 \\ 0 & \text{otherwise} \end{cases}$ $\quad P = E[I] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(C_i \geq 20)$ is a raw estimator

so to simulate the estimator we do the following

1) $X = [\,]$
2) Gen $U \sim unif(0,1)$
3) Gen $C \sim poisson(\frac{15}{.5+U})$
4) if $C \geq 20$, X.append(1)
   else, X.append(0)
5) return avg(X)

b) A conditional estimator is as follows

$E[I|U] = \mathbb{P}(C \geq 20|U) = 1 - P(C \leq 19|U) = 1 - \sum_{k=0}^{19} \frac{e^{-\lambda}\lambda^k}{k!}$, $\lambda = \frac{15}{.5+U}$

then adding a control variable of $g(y) = \frac{15}{.5+U} = \frac{15}{E(U)+U}$

we get $= E[I|U] + \gamma(\frac{15}{.5+U} - E[\frac{15}{.5+U}])$

$E[I] = E[I|U] + \gamma(U-.5)$

$E[I] = 1 - \sum_{k=0}^{19} \frac{e^{-\lambda}\lambda^k}{k!} + \gamma(U-.5)$ where $\lambda = \frac{15}{.5+U}$, $\gamma = \frac{1}{\lambda}$

so to simulate we do the following

1) $X = [\,]$
2) Gen $U \sim unif(0,1)$
3) set $\lambda = \frac{15}{.5+U}$
4) $a = 1 - \sum_{k=0}^{19} \frac{e^{-\lambda}\lambda^k}{k!} + \gamma(U-.5) > 1$, X.append(a)

n times

5) return avg(X)

9.19. cont.

c) note that $1-\sum_{k=0}^{19} \frac{e^{-\lambda}\lambda^k}{k!}$, $\lambda = \frac{15}{.5+u}$ and

$$1-\sum_{k=0}^{19} \frac{e^{-\gamma}\gamma^k}{k!} , \gamma = \frac{15}{1.5-u}$$

have the same distribution and are negatively correlated.

Thus a conditional and antithetic estimator is as follows

$$E[I] = \frac{1}{2}\left[\sum_{i=1}^{2}\left(1-\sum_{k=0}^{19} \frac{e^{-\lambda_{(i)}}\lambda_{(i)}^k}{k!} + 1\right)\right] , \lambda_{(1)} = \frac{15}{.5+u} , \lambda_{(2)} = \frac{15}{1.5-u}$$

which we simulate by the following

1) $X = [\,]$

n times:

2) Gen $u \sim unif(0,1)$

3) set $\lambda_{(1)} = \frac{15}{.5+u}$, $\lambda_{(2)} = \frac{15}{1.5-u}$

4) set $\beta_{(1)} = 1-\sum_{k=0}^{19} \frac{e^{-\lambda_{(1)}}\lambda_{(1)}^k}{k!}$

$$\beta_{(2)} = 1-\sum_{k=0}^{19} \frac{e^{-\lambda_{(2)}}\lambda_{(2)}^k}{k!}$$

5) X.append $(\beta_{(1)} + \beta_{(2)})$

6) return $\frac{avg(X)}{2}$

d)

Raw E: 0.29208, Conditional + Control Var E: 0.297733, Conditional + Antithetic Var E: 0.290481
Raw Var: 0.206771
Conditional + Control Var: 0.096605, Reduction over Raw: 2.1404
Conditional + Antithetic Var: 0.016109, Reduction over Raw: 12.8357

9.21

a) $\theta = P(S_n > c) = \sum_i P(S_n > c \mid m_n = X_i) P(m_n = X_i)$

$$= \sum_i P(S_n > c \mid m_n = X_i) \frac{1}{n}$$

$$= \frac{1}{n} \sum_i \frac{P(S_n > c \cap m_n = X_n)}{P(m_n = X_n)}$$

$$= \frac{1}{n} \sum_i \frac{P(S_n > c, m_n = X_n)}{\frac{1}{n}} = \sum_i P(S_n > c, m_n = X_n)$$

$$= n P(S_n > c, m_n = X_n)$$

b) $P(S_n > c, m_n = X_n \mid \{X_i\}_{i=1}^{n-1}) = P(S_n > c \mid \{X_i\}_{i=1}^{n}) P(m_n = X_n \mid \{X_i\}_{i=1}^{n-1})$

$$= P(S_n > c) P(S_{n-1} \leq c) P(m_n = X_n) P(m_{n-1} \neq X_{n-1})$$

$$= P(S_n > c)(1 - P(S_n > c)) P(m_n = X_n) P(m_{n-1} \neq X_{n-1})$$

$$= \frac{(\theta - \theta^2)(n-1)}{n^2}$$

9.24

a), b), c), d)

Raw E: 35.75642 Antithetic E: 34.727933

Control of S E: 35.81240339742783, Control of S and I E: 35.810366

Raw Var: 356.785666

Antithetic Var: 76.625111, Reduction over Raw: 4.6562

Control of S Var: 76.625111, Reduction over Raw: 3.5848

Control on S and I Var: 76.625111, Reduction over Raw: 4.3727

9.24

e) note that $E[T_i \mid N_i] = (N_i + 1)\mu$, μ is mean service time

so all we must know is $f(\cdot)$ for N_i. Observe the following

given $N_1 = 0$ we can say

$P(N_1 = 0) = 1$ then for N_2 we can say

$P(N_2 = 0) = \frac{1}{3}$, $P(N_2 = 1) = \frac{2}{3}$ and extrapolate this out to

for $k \geq 3$

$$P(N_k = 0) = \sum_{i=0}^{k-2} P(N_{k-1} = i)\left(\frac{1}{3}\right)^{i+1}$$

$$P(N_k = j) = \left(\frac{2}{3}\right)\sum_{i=j-1}^{k-2} P(N_{k-1} = i)\left(\frac{1}{3}\right)^{i+1-j} \quad \forall\ 1 \leq j \leq k-1$$

5.1.4

a) A raw estimator is as follows

$$E[Z] = E[P(X \in A = \{(x,y): x, y \geq a\})]$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X \in A)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(x \geq a \text{ and } y \geq a), \quad x, y \sim N(\bar{0}, C)$$

where $C = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$, $\bar{0} = (0,0)$

we can simulate through the following
1) $X = [\,]$
2) Gen $x, y \sim N(\bar{0}, C)$
3) if $x \geq a$ and $y \geq a$, X.append(1)
   else, X.append(0)
4) return avg(X)

from X we can easily find $E[X]$, $var(X)$ and the associated 95% CI

b) pretty clearly $(a,a)^T$ maximizes $N(0, C)$ over $A = \{(x,y): x, y \geq a\}$

So we just find $E[Z] = \frac{1}{n}\sum \mathbb{1}(x \geq a, y \geq a)$, $x, y \sim N((a,a)^T, C)$

by the following
1) $X = [\,]$
2) Gen $\bar{x}, y \sim N((a,a)^T, C)$
3) if $\bar{x} \geq a$, $y \geq a$, X.append(1)
   else X.append(0)
4) return avg(X)

c) we replace C with $\bar{C} = \delta C$ for some $\delta \in \mathbb{R}$ in the above methods

Crude MC simulation, a = 1
E: 0.06641, Var: 0.062000331903319036
95% CI: (0.06486670000476241, 0.06795329999523758)
Importance Sampling MC simulation, a = 1
Var: 0.01991542623748882
95% CI: (0.06422883527214322, 0.06597819105987184)

Crude MC simulation, a = 3
E: 0.00139, Var: 0.0013880817808178082
95% CI: (0.0011590803584807677, 0.0016209196415192322)
Importance Sampling MC simulation, a = 3
Var: 2.0454874191127908e-05
95% CI: (0.0013465624310992428, 0.0014026261302517906)

Crude MC simulation, a = 10
E: 0, Var: 0
95% CI: (nan, nan)
Importance Sampling MC simulation, a = 10
Var: 1.0093322023063004e-32
95% CI: (1.1200873995965199e-17, 1.244624975505494e-17)

Importance Sampling MC simulation, a = 1, delta = 0.001
Var: 3.5908585150546344e-294
95% CI: (6.386310202812891e-149, 8.735307354325856e-149)
Importance Sampling MC simulation, a = 1, delta = 2
Var: 0.039487006633100515
95% CI: (0.09565626437097965, 0.09811952547760372)
Importance Sampling MC simulation, a = 1, delta = 10
Var: 0.08920734914819273
95% CI: (0.14933945547357524, 0.15304185822852984)

Importance Sampling MC simulation, a = 3, delta = 0.001
Var: 0.0
95% CI: (nan, nan)
Importance Sampling MC simulation, a = 3, delta = 2
Var: 0.0007208506592266336
95% CI: (0.009701072039568772, 0.010033889416399955)
Importance Sampling MC simulation, a = 3, delta = 10
Var: 0.02263870018307892
95% CI: (0.06889521052734648, 0.0707603404399142)

Importance Sampling MC simulation, a = 10, delta = 0.001
Var: 0.0
95% CI: (nan, nan)
Importance Sampling MC simulation, a = 10, delta = 2
Var: 6.302214284788805e-18
95% CI: (3.848106698983523e-10, 4.159299775524978e-10)
Importance Sampling MC simulation, a = 10, delta = 10
Var: 9.539689167495977e-06
95% CI: (0.0008941180978331379, 0.0009324050200036513)

5.4.1

A raw estimator is $E[\theta] = \frac{1}{k}\sum_{i=1}^{k} \mathbb{1}\left(\sum_{n=1}^{N} X_n D_n > C = 45\right)$ which we simulate by

1) $X = [\,]$

k times:

2) Gen $P \sim Beta(1,19)$, $X_n \sim N(3,1)$ for $n \in [1,100]$
3) Find $D_n \sim Bern(P)$ $\forall n = 1, \dots, 100$
4) Find $L = \sum_{n=1}^{N} D_n X_n$
5) if $L > 45$. X.append(1)
   else X.append(0)

6) return avg(X)

Then b/c $E\left[\sum_{n=1}^{N} X_n D_n \mid \sum_{n=1}^{N} D_n\right]$

$$= \sum_{n=1}^{N}\left(E\left[X_n D_n \mid \sum_{n=1}^{N} D_n\right]\right) = \sum_{n=1}^{N}\left(E\left[X_n \mid \sum_{n=1}^{N} D_n\right] E\left[D_n \mid \sum_{n=1}^{N} D_n\right]\right)$$

$$= \sum_{n=1}^{N}\left(E[X_n]\, E\left[D_n \mid \sum_{n=1}^{N} D_n\right]\right)$$

$$= \mu \sum_{n=1}^{N}\left(E\left[D_n \mid \sum_{n=1}^{N} D_n\right]\right) = \mu\, E\left[\sum_{n=1}^{N} D_n \mid \sum_{n=1}^{N} D_n\right]$$

$$= \mu \sum_{n=1}^{N} D_n$$

So $P(L > X) = 1 - \Phi\left(\dfrac{X - \mu \sum_{n=1}^{N} D_n}{\sqrt{\sum_{n=1}^{N} D_n}}\right)$ where $\Phi(\cdot)$ is the CDF of $N(0,1)$ at $(\cdot)$

So we simulate the conditional estimator of $E[\theta] = \frac{1}{k}\sum_{j=1}^{k} A_{Z_D}$ where $A_{Z_D} = 1 - \Phi\left(\dfrac{45 - \mu \sum_{n=1}^{N} D_n}{\sqrt{\sum_{n=1}^{N} D_n}}\right)$

by

1) $X = [\,]$

k times:

2) Gen $P \sim Beta(1,19)$, $D_n \sim Bern(P)$ $\forall n$
3) find $\dfrac{X - \mu \sum_i D_n}{\sqrt{\sum_i D_n}} = \beta$
4) find A_{Z_β}
6) X.append(A_{Z_β})

7) return avg(X)