

Assignment #7 Solutions

1. Telephone model.

(a) GSMP building blocks:

State space: $S = \left\{ (z_1, \dots, z_N) \in \{0, 1, \dots, K\}^N : \sum_{j=1}^N 1_{\{k\}}(z_j) = 0 \text{ or } 2 \text{ for } 1 \leq k \leq K \right\}$

Active events: for $s = (z_1, \dots, z_N) \in S$,

$e_i \in E(s)$ iff $z_i = 0$ ($1 \leq i \leq N$)

$e_{N+k} \in E(s)$ iff $\sum_{j=1}^N 1_{\{k\}}(z_j) = 2$ ($1 \leq k \leq K$)

State-transition probabilities: for $s = (z_1, \dots, z_N) \in S$, define

$m(s) = \frac{1}{2} \sum_{k=1}^K \sum_{j=1}^N 1_{\{k\}}(z_j) = \# \text{ of occupied links in state } s$

$a(s) = \min \left\{ k : \sum_{j=1}^N 1_{\{k\}}(z_j) = 0 \right\} = \text{ index of lowest-numbered avail. link in state } s$

Then the state-transition probabilities are defined as follows.

If $e^* = e_i$ ($1 \leq i \leq N$),

- $p(s'; s, e^*) = 1$

when $s = (z_1, \dots, z_N)$ with $z_i = 0$ and $m(s) = K$
and $s' = s$

- $p(s'; s, e^*) = \frac{1}{N-1} \sum_{\substack{1 \leq j \leq N \\ j \neq i}} 1_{\{1, 2, \dots, K\}}(z_j)$

when $s = (z_1, \dots, z_N)$ with $z_i = 0$ and $m(s) < K$

- $p(s'; s, e^*) = \frac{1}{N-1}$ and $s' = s$

when $s = (z_1, \dots, z_N)$ with $z_i = z_j = 0$ and $m(s) < K$

and $s' = (z_1, \dots, z_N) = s$, except that $z_i = z_j = a(s)$

If $e^* = e_{N+k}$ ($1 \leq k \leq K$),

- $p(s'; s, e^*) = 1$

when $s = (z_1, \dots, z_N)$ with $z_i = z_j = k$ for some distinct i, j

and $s' = s$ except that $z_i = z_j = 0$

Otherwise, $p(s'; s, e^*) = 0$.

Speeds: All speeds $r(s, e)$ are equal to 1

Clock-setting distributions:

if $e' = e_i$ ($1 \leq i \leq N$) then $F(x; s', e_j, s, e^*) = (1 - e^{-x/6})I(x \geq 0)$

if $e' = e_{N+k}$ ($1 \leq k \leq K$) then

$$F(x; s', e_j, s, e^*) = \begin{cases} 0 & \text{if } x < 0 \\ x/6 & \text{if } 0 \leq x \leq 6 \\ 1 & \text{if } x > 6 \end{cases}$$

Initial distribution:

$v(s) = 1$ iff $s = (0, 0, \dots, 0)$, and $F_0(x; e_i, s) = (1 - e^{-x/6})I(x \geq 0)$ for $1 \leq j \leq N$, $s_{0,j} = 1$ for $1 \leq i \leq N$.

(Events of the form e_{N+k} are not active at time 0.)

(b) See the class web page for a program that solves the problem. Our results were as follows.

Performance measure	Point Estimate	95% CI	CI Half-width
IQR for idle-links time	7.51	[6.59, 8.44]	0.924
Long-run frac. link 2 busy	0.20	[0.17, 0.22]	0.027
SD for idle-links time	5.54	[5.20, 5.89]	0.35

The successive times at which the GSMP enters state $s = (0, 0, 0, 0)$ form a sequence of regeneration points because (i) the state is trivially the same at each regeneration point and (ii) by the memoryless property, the clock reading for all active events (whether new or old) is $\exp(1/6)$, independent of past history.

To compute the approximate number of runs for $\pm 2\%$ accuracy of the third performance measure, note that the above confidence interval half-width is given by $h = z_{0.975} s_{500} / \sqrt{500}$. Thus

$$n \approx \left(\frac{z_{0.975} s_{500}}{0.02 \hat{\alpha}} \right)^2 = \left(\frac{h \sqrt{500}}{0.02 \hat{\alpha}} \right)^2 = \left(\frac{0.35 \sqrt{500}}{(0.02)(5.54)} \right)^2 \approx 4990$$

So we would need 4990 runs.

(c) D&D procedure. For $k = 3$ plans and $P^* = 0.95$, $d^* = 0.1$, and $n_0 = 40$, we have $h_1 = 2.872$ from Table 10.11 in the textbook. We obtained the following results.

Rate	stage-1 mean	stage-1 variance	N_i	$N_i - n_0$	stage-2 mean	weight	weighted mean
	3.93	0.16			3.93		3.93
0.05	3.93	0.16	132	92	3.93	0.309463	3.93
0.20	4.90	3.22	2656	2616	5.09	0.015397	5.09
0.60	4.22	10.24	8447	8407	4.35	0.00534	4.35

As can be seen, setting the call rate to \$0.20 yields the highest weighted mean revenue, and hence is identified as the best of the three plans.

2. Multiple control variates.

(a) For $m = 2$, we have

$$\begin{aligned}\text{Var}(X_C) &= \text{Var}(X) + a_1^2 \text{Var}(Y^{(1)}) + a_2^2 \text{Var}(Y^{(2)}) \\ &\quad - 2a_1 \text{Cov}(X, Y^{(1)}) - 2a_2 \text{Cov}(X, Y^{(2)}) + 2a_1 a_2 \text{Cov}(Y^{(1)}, Y^{(2)})\end{aligned}$$

Taking derivatives with respect to a_1 and a_2 , setting the derivatives equal to 0, and then solving

$$\begin{aligned}a_1 &= \frac{\text{Cov}(X, Y^{(1)}) - a_2 \text{Cov}(Y^{(1)}, Y^{(2)})}{\text{Var}(Y^{(1)})} \\ \text{yields} \quad a_2 &= \frac{\text{Cov}(X, Y^{(2)}) - a_1 \text{Cov}(Y^{(1)}, Y^{(2)})}{\text{Var}(Y^{(2)})}\end{aligned}$$

Similarly, for $m = 3$, we have

$$\begin{aligned}\text{Var}(X_C) &= \text{Var}(X) + a_1^2 \text{Var}(Y^{(1)}) + a_2^2 \text{Var}(Y^{(2)}) + a_3^2 \text{Var}(Y^{(3)}) \\ &\quad - 2a_1 \text{Cov}(X, Y^{(1)}) - 2a_2 \text{Cov}(X, Y^{(2)}) - 2a_3 \text{Cov}(X, Y^{(3)}) \\ &\quad + 2a_1 a_2 \text{Cov}(Y^{(1)}, Y^{(2)}) + 2a_1 a_3 \text{Cov}(Y^{(1)}, Y^{(3)}) + 2a_2 a_3 \text{Cov}(Y^{(2)}, Y^{(3)})\end{aligned}$$

and we can solve to get

$$\begin{aligned}a_1 &= \frac{\text{Cov}(X, Y^{(1)}) - a_2 \text{Cov}(Y^{(1)}, Y^{(2)}) - a_3 \text{Cov}(Y^{(1)}, Y^{(3)})}{\text{Var}(Y^{(1)})} \\ a_2 &= \frac{\text{Cov}(X, Y^{(2)}) - a_1 \text{Cov}(Y^{(1)}, Y^{(2)}) - a_3 \text{Cov}(Y^{(2)}, Y^{(3)})}{\text{Var}(Y^{(2)})} \\ a_3 &= \frac{\text{Cov}(X, Y^{(3)}) - a_1 \text{Cov}(Y^{(1)}, Y^{(3)}) - a_2 \text{Cov}(Y^{(2)}, Y^{(3)})}{\text{Var}(Y^{(3)})}\end{aligned}$$

(b) If the covariates are uncorrelated with each other, then $\text{Cov}(Y^{(i)}, Y^{(j)}) = 0$ for $i \neq j$, and we have

$$\text{Var}(X_C) = \text{Var}(X) + \sum_{i=1}^m a_i^2 \text{Var}(Y^{(i)}) - 2 \sum_{i=1}^m a_i \text{Cov}(X, Y^{(i)}), \text{ so that } a_i = \frac{\text{Cov}(X, Y^{(i)})}{\text{Var}(Y^{(i)})} \text{ for } 1 \leq i \leq m$$

(c) We can estimate the optimal weights by substituting sample variances and covariances into the foregoing formulas. If we know the variances and covariances we can just use the known values in the above formulas instead of estimating them from the data. This yields more estimates with lower variance.

3. Synchronization for common random numbers. Simply use streams 1 and 2 for generating inter-demand times and demand sizes for all five replications. As discussed in the text, model 1 and model 2 use the same number of random numbers in each replication, so that these two streams are always synchronized. We can also use only stream 3 for the delivery lags (DLs) in all replications, but we need a synchronization step after each replication. For example, suppose model 1 used k random

numbers for DLs in replication 1 and model 2 used $j (< k)$ random numbers. Then, when simulating model 1 in replication 2, simply continue to use stream 3 for generating DLs. When simulating model 2 in replication 2, first generate and discard $k - j$ random numbers from stream 3 before starting the simulation. In this way, the DL random numbers will be synchronized across the two models during replication 2, since they will both start out using the $(k+1)$ st random number in stream 3. We continue in this manner for the subsequent replications.

4. Process monitoring

(a) We have

$$\begin{aligned} L(h) &= \frac{P_{\theta+h}(X_1 = G, X_2 = B, X_3 = G)}{P_{\theta}(X_1 = G, X_2 = B, X_3 = G)} = \frac{P_{\theta+h}(X_1 = G)P_{\theta+h}(X_2 = B | X_1 = G)P_{\theta+h}(X_3 = G | X_2 = B)}{P_{\theta}(X_1 = G)P_{\theta}(X_2 = B | X_1 = G)P_{\theta}(X_3 = G | X_2 = B)} \\ &= \frac{(\theta+h)(1-\theta-h)(\theta+h)^2}{\theta(1-\theta)\theta^2} = \frac{(\theta+h)^3(1-\theta-h)}{\theta^3(1-\theta)} \end{aligned}$$

(b) Using the above formula,

$$\begin{aligned} L'(h) &= \frac{(\theta+h)^3(-1) + 3(\theta+h)^2(1-\theta-h)}{\theta^3(1-\theta)} \\ L'(0) &= \frac{3-4\theta}{\theta(1-\theta)} = \frac{3-2}{0.5(1-0.5)} = 4 \end{aligned}$$

Note: the formula for $L'(0)$ can also be obtained using the trick described in class:

$$\begin{aligned} L(h) &= r_1(h)r_2(h)r_3(h), \text{ where} \\ r_1(h) &= \frac{\theta+h}{\theta}, \quad r_2(h) = \frac{1-\theta-h}{1-\theta}, \quad r_3(h) = \frac{(\theta+h)^2}{\theta^2}, \text{ so} \\ r_1'(h) &= \frac{1}{\theta}, \quad r_2'(h) = \frac{-1}{1-\theta}, \quad r_3'(h) = \frac{2(\theta+h)}{\theta^2} \text{ and} \\ L'(0) &= r_1'(0) + r_2'(0) + r_3'(0) = \frac{1}{\theta} - \frac{1}{1-\theta} + \frac{2}{\theta} = \frac{3-4\theta}{\theta(1-\theta)} \end{aligned}$$

The cost function is given by $c(N, \theta) = 4N + 0.5\theta^2$, where N is the number bad batches produced. So

$$\begin{aligned} Z_1 &= \frac{\partial c(N, \theta)}{\partial \theta} + c(N, \theta)L'(0) = \theta + (4N + 0.5\theta^2)4 \\ &= 0.5 + [(4)(1) + 0.5(0.5)^2] \times 4 = 17 \end{aligned}$$

(c) The Robbins-Monro recursion is $\theta_{n+1} = \Pi\left(\theta_n - \frac{a}{n}Z_n\right)$, So

$$\theta_2 = \Pi\left(\theta_1 - aZ_1\right) = \Pi\left(0.5 - (0.01)(17)\right) = \Pi\left(0.33\right) = 0.33$$

5. Epidemiology.

(a) The arrival process to a specified quadrat is the superposition of the arrival process for the individual residents. Because residents move around in an i.i.d. manner, the individual arrival processes are independent and no arrival process dominates, so Khintchine's Theorem implies that the overall arrival process should be approximately a Poisson process. Therefore W , the number of arrivals in a two-hour period, should be Poisson: $P(W = m) = \frac{e^{-2\gamma}(2\gamma)^m}{m!}$.

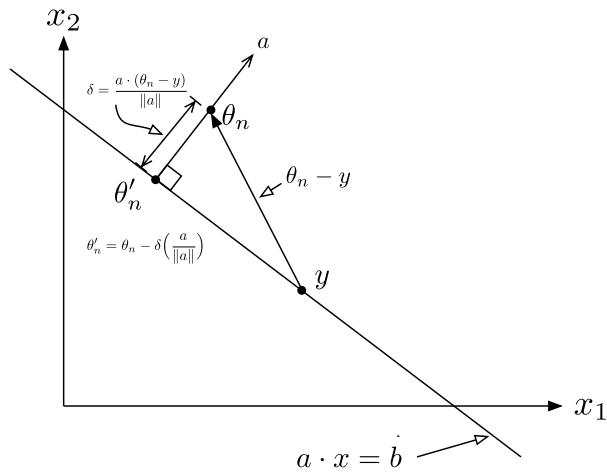
(b) For $B = 10$ bootstrap samples and an 80% confidence interval (so that $\delta = 0.2$), compute a pivot $\bar{X}_n^* - \bar{X}_n$ for each row, sort them, and then use elements $\lceil (\delta/2) \times 10 \rceil = 1$ and $\lceil (1 - \delta/2) \times 10 \rceil = 9$ as quantile estimates. As per the hint, when computing the pivots, you can delay subtracting \bar{X}_n until after you sort, since this does not change the sort order. So the sorted bootstrap means are: 2.0, 2.4, 2.8, 3.2, 3.6, 3.6, 4.6, 4.8, 5.4, 5.6. The 1st and 9th of these are 2.0 and 5.4. Now we can subtract $\bar{X}_n = 4.0$ to find that $\pi_{(l)}^* = 5.4 - 4.0 = 1.4$ and $\pi_{(u)}^* = 2.0 - 4.0 = -2.0$. So the 80% confidence interval is $[\bar{X}_n^* - \pi_{(l)}^*, \bar{X}_n^* - \pi_{(u)}^*] = [4.0 - 1.4, 4.0 + 2.0] = [2.6, 6.0]$.

(c) Our efficiency criterion says to compare the product of (expected cost per observation) times (variance per observation), and prefer the smaller value. For method A the product is $20\sigma_A^2$ and for Method B the product is $2\sigma_B^2$. Since $\sigma_B = 3\sigma_A$ (so that $\sigma_B^2 = 9\sigma_A^2$) the product for Method B is $18\sigma_A^2$, and so we prefer Method B, because it is more efficient.

6. (Extra credit) Maintaining feasibility in the R-M algorithm.

(a) We need to project the point θ_n onto the plane defined by $a \cdot x = b$. As per the hint, if y is an arbitrary point on the plane P then, considering the right triangle formed by θ_n , the projected point θ_n' , and y , we see that the perpendicular distance from θ_n to the plane is given by $\delta = \frac{a \cdot (\theta_n - y)}{\|a\|} = \frac{a \cdot \theta_n - b}{\|a\|}$. This is because $\frac{a}{\|a\|} \cdot (\theta_n - y) = 1 \times \|\theta_n - y\| \times \cos\phi = \|\theta_n - y\| \times \frac{\delta}{\|\theta_n - y\|}$. Thus, the actual projected point on the plane is computed as $\theta_n' = \theta_n - \delta \left(\frac{a}{\|a\|} \right) = \theta_n - \left(\frac{a \cdot \theta_n - b}{a \cdot a} \right) a$.

See the figure below; also, the website https://mathinsight.org/distance_point_plane has a nice interactive visualization of the geometry.



(b) If the projected point violates the lower-bound constraint, we can achieve feasibility by moving along the plane. (See the diagram below for the 2D case.) Let v be defined as in the hint; note that v is easy to calculate: e.g., set $v = (1, 1, \dots, 1, -(a_1 + a_2 + \dots + a_{d-1})/a_d)$. Let i^* denote the dimension corresponding to the largest violation of the lower bound constraint, i.e.,

$$c_{i^*} - \theta_{n,i^*} \geq c_i - \theta_{n,i} \text{ for } i = 1, \dots, d. \text{ Then set } \theta_n^+ = \theta_n' + r v, \text{ where } r = \frac{c_{i^*} - \theta_{n,i^*}}{v_{i^*}}. \text{ The figure below}$$

illustrates this for two dimensions: the example with θ_m shows the case where the projection remains feasible and no adjustment is needed, and the example with θ_n shows the case where the above adjustment is needed.

