Problem 1. Problem 3.4 in Bertsekas and Gallager.

\[ \hat{\lambda}_i = \lambda_i \quad \text{for } i < k \]

\[ \hat{\lambda}_k = \frac{(1 - \rho_1 - \cdots - \rho_{k-1})}{\hat{x}_k} \]

For priorities \( i < k \) the arrival process is Poisson so the same calculation for the waiting time as before gives

\[ W_i = \frac{\sum_{j=1}^{k} \hat{\lambda}_j \hat{x}_j}{2(1 - \rho_1 - \cdots - \rho_{k-1})(1 - \rho_1 - \cdots - \rho_i)} \quad i < k \]

For priority \( k \) and above we have infinite average waiting time in queue.

3.40

(a) The algebraic verification using Eq. (3.79) listed below

\[ W_k = \frac{R(1 - \rho_1 - \cdots - \rho_{k-1})}{(1 - \rho_1 - \cdots - \rho_k)} \]

is straightforward. In particular by induction we show that

\[ \rho_1 W_1 + \cdots + \rho_k W_k = \frac{R(\rho_1 + \cdots + \rho_k)}{1 - \rho_1 - \cdots - \rho_k} \]

The induction step is carried out by verifying the identity

\[ \rho_1 W_1 + \cdots + \rho_k W_k + \rho_{k+1} W_{k+1} = \frac{R(\rho_1 + \cdots + \rho_k)}{1 - \rho_1 - \cdots - \rho_k} + \frac{\rho_{k+1} R}{(1 - \rho_1 - \cdots - \rho_k)(1 - \rho_1 - \cdots - \rho_{k+1})} \]

The alternate argument suggested in the hint is straightforward.

(b) Cost

\[ C = \sum_{i=1}^{k} c_k X_k^i \hat{\lambda}_k = \sum_{i=1}^{k} c_k \hat{x}_k W_k = \sum_{i=1}^{k} \left( \frac{c_k}{X_k} \right) \rho_k W_k \]

We know that \( W_1 \leq W_2 \leq \cdots \leq W_n \). Now exchange the priority of two neighboring classes \( i \) and \( j \), \( i < j \), and compare \( C \) with the new cost

\[ C' = \sum_{i=1}^{k} \left( \frac{c_k}{X_k} \right) \rho_k W_k' \]
Problem 1.
Problem 3.4 in Bertsekas and Gallager.

In $C'$ all the terms except $k = i$ and $j$ will be the same as in $C$ because the interchange does not affect the waiting time for other priority class customers. Therefore

$$C' - C = \frac{c_i}{X_j} \rho_j W_j' + \frac{c_i}{X_i} \rho_i W_i' - \frac{c_i}{X_j} \rho_j W_j - \frac{c_i}{X_i} \rho_i W_i.$$  

We know from part (a) that

$$\sum_{k=i}^n \rho_k W_k = \text{constant.}$$  

Since $W_k$ is unchanged for all $k$ except $k = i$ and $j (= i+1)$ we have

$$\rho_i W_i + \rho_j W_j = \rho_i W_i' + \rho_j W_j'.$$

Denote

$$B = \rho_i W_i - \rho_i W_i = \rho_j W_j - \rho_j W_j.$$

Clearly we have $B \geq 0$ since the average waiting time of customer class $i$ will be increase if class $i$ is given lower priority. Now let us assume that

$$\frac{c_i}{X_i} \leq \frac{c_j}{X_j}.$$  

Then

$$C' - C = \frac{c_i}{X_i} (\rho_i W_i' - \rho_i W_i) - \frac{c_i}{X_j} (\rho_j W_j' - \rho_j W_j) = B \left( \frac{c_i}{X_i} - \frac{c_j}{X_j} \right).$$

Therefore, only if $\frac{c_i}{X_i} < \frac{c_{i+1}}{X_{i+1}}$ can we reduce the cost by exchanging the priority order of $i$ and $i+1$. Thus, if $(1, 2, 3, ..., n)$ is an optimal order we must have

$$\frac{c_1}{X_1} \geq \frac{c_2}{X_2} \geq \frac{c_3}{X_3} \geq ... \geq \frac{c_n}{X_n}.$$
problem 2

3.23

Let
\[ p_m = P(\text{the 1st m servers are busy}) \]
as given by the Erlang B formula. Denote
\[ r_m = \text{Arrival rate to servers (m+1) and above} \]
\[ \lambda_m = \text{Arrival rate to server m} \]

We have
\[ r_m = p_m \lambda \]
\[ \lambda_m = r_{m-1} - r_m = (p_{m-1} - p_m)\lambda. \]

The fraction of time server m is busy is
\[ b_m = \lambda_m/\mu. \]

problem 3

3.35

Consider a gated all-at-once version of the limited service reservation system. Here there are m users, each with independent Poisson arrival rate \( \lambda/\mu \). Each user has a separate queue, and is allowed to make a reservation for at most one packet in each reservation interval. This packet is then transmitted in the subsequent data interval. The difference with the limited service system of Section 3.5.2 is that here users share reservation and data intervals.

Consider the ith packet arrival into the system and suppose that the user associated with packet i is user j. We have as in Section 3.5.2
\[ E(W_i) = E(R_i) + E(N_i)/\mu + (1 + E(Q_i) - E(m_i))E(V) \]

where \( W_i, R_i, N_i, \mu, E(V) \) are as in Section 3.5.2, \( Q_i \) is the number of packets in the queue of user j found by packet i upon arrival, and \( m_i \) is the number (0 or 1) of packets of user j that will start transmission between the time of arrival of packet i and the end of the frame in which packet i arrives. We have as in Section 3.5.2
\[ R = \lim_{i \to \infty} E(R_i) + E(N_i)/\mu + (1 + E(Q_i) - E(m_i))E(V) \]
\[ N = \lim_{i \to \infty} E(N_i) = \lambda W \]
\[ Q = \lim_{i \to \infty} E(Q_i) = \lambda W/m \]

so there remains to calculate \( \lim_{i \to \infty} E(m_i) \).

There are two possibilities regarding the time of arrival of packet i.
problem 3 cont.

a) Packet i arrives during a reservation interval. This event, call it A, has steady state probability \( (1 - \rho) \)

\[
P(A) = 1 - \rho.
\]

Since the ratio of average data interval length to average reservation interval length is \( \rho / (1 - \rho) \) we see that the average steady state length of a data interval is \( \rho E(V) / (1 - \rho) \). Therefore the average steady state number of packets per user in a data interval is \( \rho E(V) / ((1 - \rho)mE(X)) = \lambda E(V) / (1 - \rho)m \). This also equals the steady state value of \( E(m | A) \) in view of the system symmetry with respect to users

\[
\lim_{i \to \infty} E(m_i | A) = \frac{\lambda E(V)}{(1 - \rho)m}.
\]

b) Packet i arrives during a data interval. This event, call it B, has steady state probability \( \rho \)

\[
P(B) = \rho.
\]

Denote

\[
\alpha = \lim_{i \to \infty} E(m_i | B),
\]

\[
\alpha_k = \lim_{i \to \infty} E(m_i | B, \text{the data interval of arrival of packet } i \text{ contains } k \text{ packets}).
\]

Assuming \( k > 0 \) packets are contained in the data interval of arrival, there is equal probability \( 1/k \) of arrival during the transmission of any one of these packets. Therefore

\[
\alpha_k = \sum_{n=1}^{k} \frac{1}{k} \frac{k - n}{m} = \frac{k(k - 1)}{2km} = \frac{k - 1}{m}.
\]

Let \( P(k) \) be the unconditional steady-state probability that a nonempty data interval contains \( k \) packets, and \( E[k] \) and \( E[k^2] \) be the corresponding first two moments. Then we have using Bayes' rule

\[
\lim_{i \to \infty} P(\text{The data interval of arrival of packet } i \text{ contains } k \text{ packets}) = kP(k)/E(k).
\]

Combining the preceding equations we have

\[
\alpha = \sum_{k=1}^{m} \frac{kP(k)}{E[k]} \quad \alpha_k = \sum_{k=1}^{m} \frac{P(k)k(k - 1)}{2E[k]m} = \frac{E[k^2]}{2mE[k]} - \frac{1}{2m}.
\]
We have already shown as part of the analysis of case a) above that

\[ E[k] = \lambda E(V)/(1 - \rho) \]

so there remains to estimate \( E[k^2] \). We have

\[ E[k^2] = \sum_{k=1}^{m} k^2 P(k) \]

If we maximize the quantity above over the distribution \( P(k), k = 0, 1, ..., m \) subject to the constraints \( \sum_{k=m}^{m} kP(k) = E[k], \sum_{k=0}^{m} P(k) = 1, P(k) \geq 0 \) (a simple linear programming problem) we find that the maximum is obtained for \( P(m) = E[k]/m, P(0) = 1 - E[k]/m, \) and \( P(k) = 0, k = 1, 2, ..., m-1 \). Therefore

\[ E[k^2] \leq mE[k]. \]

Similarly if we minimize \( E[k^2] \) subject to the same constraints we find that the minimum is obtained for \( P(k') = k' - E[k], P(k) = 1 - (k' - E[k]) \) and \( P(k') = 0 \) for \( k \neq k' - 1, k' \) where \( k' \) is the integer for which \( k' - 1 \leq E[k] < k' \). Therefore

\[ E[k^2] \geq (k'-1)^2(k' - E[k]) + (k')^2[1 - (k' - E[k])] \]

After some calculation this relation can also be written

\[ E[k^2] \geq E[k] + (k' - 1)(2E[k] - k') \quad \text{for} \quad E[k] \in (k' - 1, k') \]

\[ k' = 1, 2, ..., m \]

Note that the lower bound above is a piecewise linear function of \( E[k] \), and equals \( (E[k])^2 \) at the breakpoints \( k' = 1, 2, ..., m \). Summarizing the bounds we have

\[ \frac{E[k] + (k' - 1)(2E[k] - k')}{2mE[k]} - \frac{1}{2m} \leq \alpha \leq \frac{1}{2} \cdot \frac{1}{2m}, \]

where \( k' \) is the positive integer for which

\[ k' - 1 \leq E[k] < k'. \]

Note that as \( E[k] \) approaches its maximum value \( m \) (i.e., the system is heavily loaded), the upper and lower bounds coincide. By combining the results for cases a) and b) above we have

\[ \lim_{i \to \infty} E(m_i) = P(A) \lim_{i \to \infty} E(m_i | A) + P(B) \lim_{i \to \infty} E(m_i | B) \]
Problem 3 cont.

\[ (1 - \rho) \frac{\lambda E(V)}{(1 - \rho) m} + \rho \alpha \]

or finally

\[ \lim_{i \to \infty} E(m_i) = \frac{\lambda E(V)}{m} + \rho \alpha \]

where \( \alpha \) satisfies the upper and lower bounds given earlier. By taking limit as \( i \to \infty \) in the equation

\[
E(W_i) = E(R_i) + E(N_i)/\mu + (1 + E(Q_i) - E(m_i))E(V)
\]

and using the expressions derived we finally obtain

\[
W = \frac{\lambda E(X^2)}{2(1 - \rho - \frac{\lambda E(V)}{m})} + \frac{(1 - \rho)E(V^2)}{2(1 - \rho - \frac{\lambda E(V)}{m})E(V)} + \frac{\left(1 - \rho \alpha - \frac{\lambda E(V)}{m}\right)E(V)}{1 - \rho - \frac{\lambda E(V)}{m}}
\]

where \( \alpha \) satisfies

\[
\frac{E(k) + (k' - 1)(2E(k) - k')}{2mE(k)} - \frac{1}{2m} \leq \alpha \leq \frac{1}{2} - \frac{1}{2m}
\]

\( E(k) \) is the average number of packets per data interval

\[ E(k) = \frac{\lambda E(V)}{1 - \rho} \]

and \( k' \) is the integer for which \( k' - 1 \leq E(k) < k' \). Note that the formula for the waiting time \( W \) becomes exact in the limit both as \( \rho \to 0 \) (light load), and as \( \rho \to 1 \) - \( \frac{\lambda E(V)}{m} \) (heavy load) in which case \( E(k) \to m \) and \( \alpha \to 1/2 - 1/2m \). When \( m = 1 \) the formula for \( W \) is also exact and coincides with the one derived for the corresponding single user one-at-a-time limited service system.
problem 4

Our objective is to derive the P-K formula for the mean waiting time in queue:

\[ W = \frac{\lambda X^3}{2(1 - \rho)}, \]

or equivalently, in terms of the number of customers in the system,

\[ N = \rho + \frac{\lambda^2 X^2}{2(1 - \rho)}, \]

which is easily verified by applying Little’s formula.

Let us start from the P-K transform equation:

\[ Q(z) = X^*(\lambda - \lambda z) \frac{(1 - \rho)(1 - z)}{X^*(\lambda - \lambda z) - z}. \]

Note that \( Q(z) \) here is the transform of the number of customers in the system at departure instants.

To obtain the mean number of customers, we need to calculate the first derivative of \( Q(z) \) evaluated at \( z = 1 \). We have as given in the lecture 4

\[
\left. \frac{dQ(z)}{dz} \right|_{z=1} = (1 - \rho) \frac{d\left(\frac{X^*(\lambda - \lambda z)(1 - z)}{X^*(\lambda - \lambda z) - z} \right)}{dz} \Bigg|_{z=1}.
\]

If we let \( y = (\lambda - \lambda z) \) for simplicity of notation, we have

\[
\left. \frac{dX^*(\lambda - \lambda z)}{dz} \right|_{z=1} = \left( \frac{dX^*(y)}{dy} \right) \left( \frac{dy}{dz} \right) = -\lambda X^*(y),
\]

and as \( z \to 1, y \to 0 \). Similarly, we calculate

\[
\left. \frac{d(X^*(y))}{dz} \right|_{z=1} = -\lambda X^*(y)
\]

for later use.

Now we can rewrite (1) as follows:

\[
\left. \frac{dQ(z)}{dz} \right|_{z=1} = (1 - \rho) \frac{1}{(X^*(y) - z)^2} \left\{ \left( -\lambda X^*(y)(1 - z) - X^*(y) \right)(X^*(y) - z) \right. \\
\left. - \left( -\lambda X^*(y) - 1 \right) X^*(y)(1 - z) \right\}_{y=0}^{y=0},
\]

and by cancelling terms we obtain

\[
\left. \frac{dQ(z)}{dz} \right|_{z=1} = (1 - \rho) \frac{X^*(y)(1 - z) + X^*(y)(1 - X^*(y))}{(X^*(y) - z)^2} \Bigg|_{y=0}^{y=0}
\]

Since \( X^*(y) \to 1 \) as \( y \to 0 \), both the numerator and the denominator are zero in the limit \( z \to 1 \).

By l’Hospital’s rule, we can evaluate this after differentiating by \( z \) both the numerator and the denominator. Therefore

\[
\left. \frac{dQ(z)}{dz} \right|_{z=1} = (1 - \rho) \frac{\lambda^2 X^*(y)(1 - z)}{2(1 - \lambda X^*(y) - 1)} \frac{\lambda X^*(y)(2 - X^*(y)) - X^*(y)(1 - z)}{2(\lambda X^*(y) - 1)} \Bigg|_{y=0}^{y=0}
\]

\[
= (1 - \rho) \frac{\lambda^2 X^*(y)(1 - z)}{2(\lambda X^*(y) - 1) \lambda X^*(y) - 1 + \lambda X^*(y) - 1} \Bigg|_{y=0}^{y=0}.
\]

Note that

\[ X^*(y)|_{y=0} = -\lambda, \quad X^*(y)|_{y=0} = \lambda. \]

and again by l’Hospital’s rule

\[
\frac{z(1 - z)}{X^*(y) - z} \Bigg|_{y=1}^{y=0} = \frac{1 - 2z}{-\lambda X^*(y) - 1} \Bigg|_{y=1}^{y=0} = \frac{-1}{\lambda - 1} = \frac{1}{1 - \rho}.
\]

Therefore, (2) becomes

\[ N = \frac{\lambda^2 X^2}{2(1 - \rho)} + \rho. \]