Appendix F. Correlation and Independence

Given a bivariate data sample (u_i, v_i) for $i = 1, 2, \ldots, n$, the traditional model associated with the statistical analysis of this sample is that the points (u_i, v_i) are the result of repeatedly sampling possible values of a pair of random variables (U, V). As in Section 4.4, the sample correlation coefficient r is then a statistic used to test the hypothesis that U and V are *linearly* related—that is, values of $|r|$ close to 1 are used to argue the existence of constants a, b, c such that $aU + bV + c \cong 0$.

F.1 Bivariate Random Variables

Definition F.1 The random variables U and V (discrete or continuous) are *independent* if and only if

$$\Pr\big((U \le u) \text{ and } (V \le v)\big) = \Pr(U \le u)\,\Pr(V \le v)$$

for all possible values u and v of U and V. If the random variables are not independent, they are said to be *dependent*.

Example F.1

The usual discrete-event simulation model of a single-server service node makes use of the *assumption* of independence. For example, for each job, the service time and arrival time are assumed to be independent. Also, generally the service time does not depend on the length of the queue, and so, unless the queue discipline is based upon a knowledge of a job's service times, the delay a job experiences in the queue is independent of the job's service time. The interarrival time and the delay in the queue, however, are *not* independent—that is, if the interarrival time is quite large, then any pre-existing queue has time to disappear, and, as a result, the delay in the queue experienced by a job with a large interarrival time is likely to be small. Similarly, a job's delay in the queue and wait in the service node are definitely not independent, particularly if the utilization is close to 1.

یعنی مقدار متغیر تصادفی - هیچ جوری (خطی یا غیرخطی) به دیگری بستگی ندارد

Independence is a very strong condition—it means that the value of one random variable does not depend in *any* (linear or nonlinear) way on the value of the other. How does independence relate to correlation? مهم

The answer is that independent random variables are uncorrelated. The converse is *not* necessarily true, however—uncorrelated random variables can be dependent. To understand why this is so, we begin by defining the covariance and correlation of two random variables. مهم

Covariance

Definition F.2 The *covariance* of two random variables U and V is

$$\gamma_{UV} = E\big[(U - \mu_U)(V - \mu_V)\big],$$

where

$$\mu_U = E[U] \qquad \text{and} \qquad \mu_V = E[V]$$

are the means of U and V, respectively.

The sample covariance (Definition 4.4.2) is an estimate of γ_{UV}, which is calculated from data pairs. Moreover, just as there is an alternative formula for the sample covariance, an equivalent expression for the covariance is

$$\gamma_{UV} = E[UV] - \mu_U \mu_V.$$

The derivation of this equation is left as an exercise.

Correlation Coefficient

Definition F.3 The *correlation coefficient* of two random variables U and V is

$$\rho_{UV} = \frac{\gamma_{UV}}{\sigma_U \sigma_V},$$

where

$$\sigma_U = \sqrt{E\big[(U - \mu_U)^2\big]} \qquad \text{and} \qquad \sigma_V = \sqrt{E\big[(V - \mu_V)^2\big]}$$

are the standard deviations of U and V respectively, which are assumed to be nonzero.

The sample correlation coefficient (Definition 4.4.2) is an estimate of ρ_{UV}. Therefore, by analogy, it is reasonable to expect the following:

- $|\rho_{UV}| \leq 1$;
- $|\rho_{UV}| = 1$ if and only if there exists constants (a, b, c) such that $aU + bV + c = 0$.

The proof that these two statements are true is left as an exercise.

Independence Implies Uncorrelated. One important consequence of independence (presented without proof) is that, if the random variables U and V are independent, then

$$E[UV] = E[U]E[V].$$

From this result, if U and V are independent, then $\gamma_{UV} = 0$. That is,

$$\begin{aligned}
\gamma_{UV} &= E[UV] - \mu_U \mu_V \\
&= E[U]E[V] - \mu_U \mu_V \\
&= \mu_U \mu_V - \mu_U \mu_V \\
&= 0,
\end{aligned}$$

as summarized by the following theorem. This theorem provides the justification for using a nonzero value of the sample correlation coefficient to argue the *non*independence of the two random variables from which the sample was drawn.

Theorem F.1 If the random variables U and V (discrete or continuous) are independent, then they are uncorrelated. That is

- if U and V are independent then U and V are uncorrelated.

Equivalently, in terms of the contrapositive,

- if U and V are correlated (positive or negative) then U and V are dependent.

Example F.2
Consider an urn filled with black balls and white balls. Let p be the proportion of black balls, and suppose that two balls are drawn from the urn, in sequence, with replacement. Let the random variables U and V count the number of black balls (0 or 1) on the first draw and on the second draw respectively. Because the draws are *with* replacement, U and V are independent. Moreover, U and V are *Bernoulli*(p) random variables, and so $E[U] = E[V] = p$. The possible values of UV are 0, 1, with $\Pr(UV = 1) = p^2$. Therefore UV is a *Bernoulli*(p^2) random variable, and so $E[UV] = p^2$. From these calculations, $\gamma_{UV} = p^2 - p^2 = 0$, and so $\rho_{UV} = 0$: As is consistent with Theorem F.1, we see that U and V are uncorrelated.

$P(1) = p^2$ UV $(0,1)$

$P(0) = 1 - p^2$

$\gamma_{UV} = E[UV] - E[U]E[V]$

Example F.3

The converse of Theorem F.1 is not necessarily true. For example, let U be *Uniform*$(-1, 1)$ and define $V = U^2$. It is intuitive that U and V are *not* independent. They are, however, uncorrelated. The pdf of U is $f(u) = 1/2$ for $-1 < u < 1$, so

$$E[U] = \int_{-1}^{1} \frac{1}{2} u \, du = \cdots = 0$$

$$E[V] = E[U^2] = \int_{-1}^{1} \frac{1}{2} u^2 \, du = \cdots = \frac{1}{3}$$

$$E[UV] = E[U^3] = \int_{-1}^{1} \frac{1}{2} u^3 \, du = \cdots = 0.$$

From these calculations, $\gamma_{UV} = E[UV] - E[U]E[V] = 0 - 0 = 0$, and so $\rho_{UV} = 0$. Thus we see that U and V are uncorrelated but dependent random variables. The dependence of V on U despite ρ_{UV} being zero becomes clear when one views the scatterplot generated by, say, 100 (u_i, v_i) samples of (U, V), along with the associated regression line, as illustrated in Figure F.1.

کمیتی وابستگی U و V

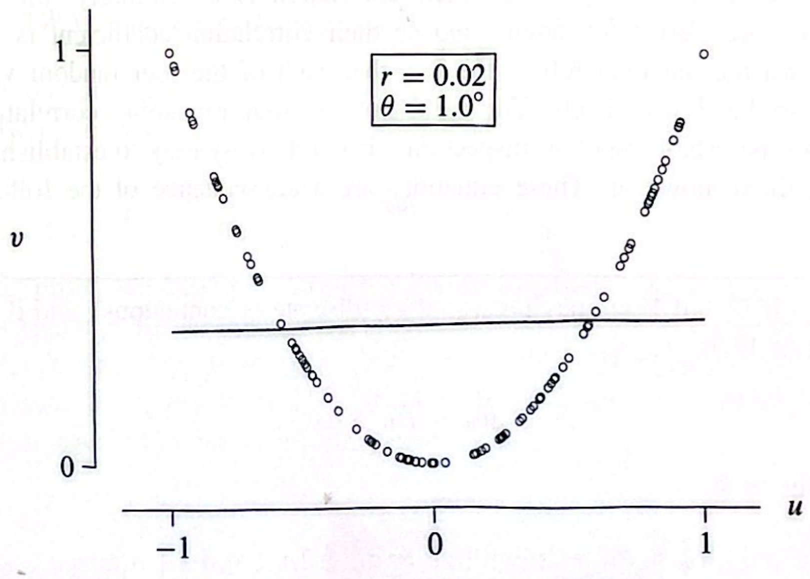

Figure F.1 Monte Carlo simulation of 100 (u_i, v_i) pairs.

چون وابستگی U و V غیرخطی است ... رگرسیون aV+bV+c ... منعکس نمی‌کند و این دو را خطی نمی‌توانند، و دیگری غیرخطی

The point here is that the correlation that exists between U and V is *nonlinear* and so is not reflected in the computation of the *linear* correlation coefficient.

را ببینیم که و در این ترتیب ... رگرسیون را خطی ... نشان داده، ولی رگرسیون ... ببینیم ها ... نشان دهد

The primary significance of Theorem F.1 is that the sample correlation coefficient can be used as a test for *dependence*: If r is the sample correlation coefficient of the data (u_i, v_i)

و غیرخطی‌ها را نمی‌تواند بررسی نشان دهد

for $i = 1, 2, \ldots, n$ and if $|r|$ is significantly different from 0 then, with great confidence, we can conclude that the (U, V) pair of random variables that generated the data is *not* independent.*

Example F.4

In Example 4.4.2, a modified version of program ssq2 was used to simulate a FIFO $M/M/1$ service node and generate a steady-state sample of 100 interarrival times, delays, waits, and service times. Scatterplots and the associated correlation coefficients corresponding to four of the six possible pairings were illustrated. Intuition, coupled with the results of this experiment, suggests the following (symmetric) table of ρ_{UV} correlation coefficient values

	interarrival	delay	wait	service
interarrival	1			
delay	−	1		
wait	−	+	1	
service	0	0	+	1

where +, −, and 0 indicate positive, negative, and no correlation respectively. The 0's in this table are a consequence of Theorem F.1: Because the interarrival times and service times are independent, their correlation is 0. Similarly, the delay and service times are also independent, and so their correlation coefficient is 0 as well. The 1's down the diagonal reflect the fact that each of the four random variables is perfectly correlated with itself. The values of the four remaining correlation coefficients cannot be established "by inspection." It is relatively easy to establish equations for two of them, however. These equations are a consequence of the following two theorems.

Theorem F.2 If U and V are random variables (discrete or continuous), and if $W = U + V$, then the mean of W is

$$\mu_W = \mu_U + \mu_V$$

and the variance of W is

$$\sigma_W^2 = \sigma_U^2 + 2\gamma_{UV} + \sigma_V^2 = \sigma_U^2 + 2\rho_{UV}\sigma_U\sigma_V + \sigma_V^2.$$

The proof of this theorem is left as an exercise. Note that

$$\sigma_W^2 = \sigma_U^2 + \sigma_V^2$$

if and only if U and V are uncorrelated. The independence of U and V is a *sufficient* condition for this equation to be valid.

*There is no simple test for independence.

D | S

$M/M/1$ Correlations

Theorem F.3 Let the random variables W, D, and S represent the wait, delay, and service time experienced by a randomly selected job when a FIFO $M/M/1$ service node is in steady state. Define $\rho = \lambda/\nu$, where λ is the arrival rate and ν is the service rate. Then

- the steady-state correlation between W and S is $\rho_{WS} = 1 - \rho$, and
- the steady-state correlation between W and D is $\rho_{WD} = \sqrt{\rho(2-\rho)}$.

Proof. Let μ_W, μ_D, and μ_S and σ_W^2, σ_D^2, and σ_S^2 denote the mean and variance of W, D, and S, respectively. To prove this theorem, we begin with the covariance between W and S. Recall that $W = D + S$ and that D and S are independent. Therefore,

$$\mu_W = \mu_D + \mu_S \qquad \text{and} \qquad \sigma_W^2 = \sigma_D^2 + \sigma_S^2.$$

Moreover,

$$E[WS] = E[(D+S)S] = E[DS + S^2] = E[D]\,E[S] + E[S^2] = \mu_D\mu_S + E[S^2];$$

as a consequence,

$$\gamma_{WS} = E[WS] - \mu_W\mu_S = \mu_D\mu_S + E[S^2] - (\mu_D + \mu_S)\mu_S = E[S^2] - \mu_S^2 = \sigma_S^2.$$

Using an analogous argument, we find that the covariance between W and D is $\gamma_{WD} = \sigma_D^2$. Therefore, the two correlation coefficients of interest are

$\rho_{WS} = \frac{\gamma_{WS}}{\sigma_W\sigma_S} = \frac{\sigma_S^2}{\sigma_W\sigma_S}$

$$\rho_{WS} = \frac{\sigma_S}{\sigma_W} \qquad \text{and} \qquad \rho_{WD} = \frac{\sigma_D}{\sigma_W}.$$

To complete the proof, we need to relate the three standard deviations σ_W, σ_D, and σ_S to the steady-state utilization ρ. Recall (Section 8.5) that, for an $M/M/1$ service node, the service time is *Exponential*(μ_S). Similarly, if the queue discipline is FIFO, the steady-state wait in the node is *Exponential*(μ_W), where $\mu_W = \mu_S/(1-\rho)$. Because an *Exponential* random variable has its standard deviation equal to its mean, it follows that

زمان صف M/M/1 ها
دارای (نرخ $\bar{w}$ آن)

$$\rho_{WS} = \frac{\sigma_S}{\sigma_W} = \frac{\mu_S}{\mu_W} = \frac{\mu_W(1-\rho)}{\mu_W} = 1 - \rho.$$

Similarly,

$$\sigma_D^2 = \sigma_W^2 - \sigma_S^2 = \mu_W^2 - \mu_S^2 = \mu_W^2 - \mu_W^2(1-\rho)^2 = \mu_W^2\rho(2-\rho),$$

and, therefore,

$$\rho_{WD} = \frac{\sigma_D}{\sigma_W} = \frac{\mu_W\sqrt{\rho(2-\rho)}}{\mu_W} = \sqrt{\rho(2-\rho)},$$

which establishes the theorem. □

Example F.5
The following table summarizes the FIFO $M/M/1$ service node steady-state correlation coefficients ρ_{WS} and ρ_{WD} for selected values of ρ.

ρ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
ρ_{WS}	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100
ρ_{WD}	0.436	0.600	0.714	0.800	0.866	0.917	0.954	0.980	0.995

Both (W, S) and (W, D) have a positive correlation, as expected.

The analytic evaluation of the other two nonzero correlation coefficients, ρ_{RD} and ρ_{RW}, is significantly more difficult. In particular, the delay D_i and the interarrival time R_i of the i^{th} job are related to the delay D_{i-1} and service time S_{i-1} of the previous job (for a FIFO queue, per Section 1.2) by the nonlinear equation

$$D_i = \max\{0,\ D_{i-1} + S_{i-1} - R_i\} \qquad i = 1, 2, 3, \ldots.$$

Not only is this equation nonlinear; to evaluate γ_{RD}, one must know both the paired correlation between S_{i-1} and D_i and the serial correlation between D_{i-1} and D_i.

Example F.6
Fortunately, the two correlation coefficients ρ_{RD} and ρ_{RW} are relatively easy to *estimate* by simulation. Figure F.2 illustrates estimates of these two FIFO $M/M/1$ service

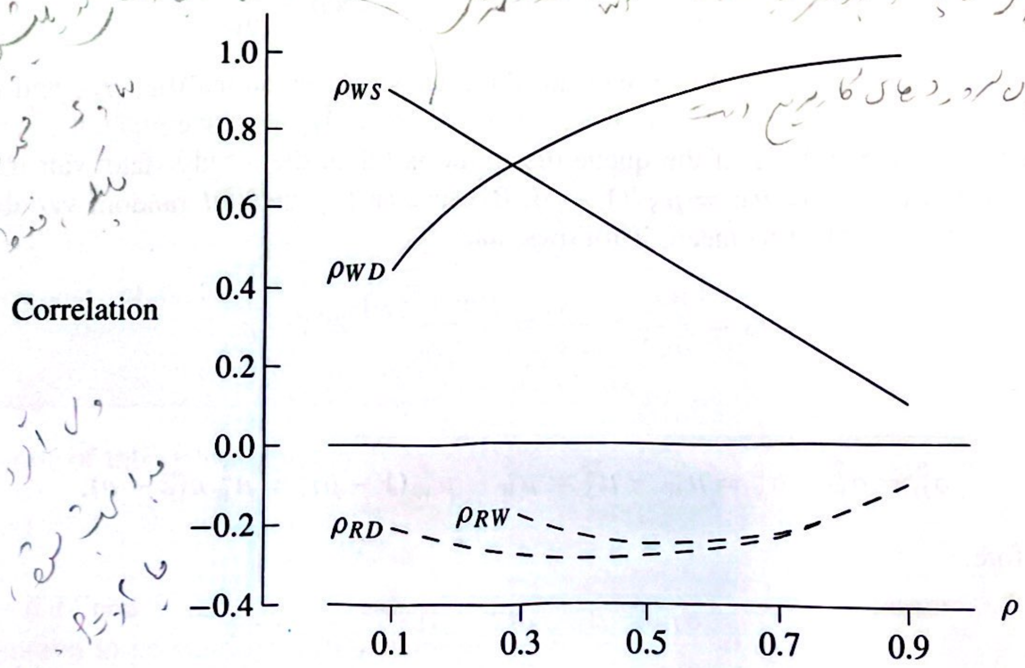

Figure F.2 Correlations between W, D, S, and R as a function of the traffic intensity $\rho = \lambda/\nu$ for an $M/M/1$ queue at steady state.

node steady-state correlation coefficients (as a dashed line). For comparison, the theoretical correlation coefficients from Theorem F.3 are also illustrated (as solid lines). Both (R, D) and (R, W) have a weak negative correlation, as expected.

F.2 Multivariate Random Variables

The high autocorrelation that typically exists in the time-sequenced stochastic data produced by a discrete-event simulation makes the statistical analysis of the data a challenge. Specifically, if we wish to make an *interval* estimate of some steady-state statistic—for example, the average wait in a service node—we must be prepared to deal with the impact of autocorrelation on our ability to make accurate estimates of the standard deviation. The following discussion introduces the notation and assumptions that are typically used in the analysis of discrete-event simulation output.

Let $X_1, X_2, \ldots, X_b$ be a *batch* (sequence) of random variables with common mean μ and common variance σ^2:

$$E[X_i] = \mu \qquad \text{and} \qquad E\big[(X_i - \mu)^2\big] = \sigma^2$$

for $i = 1, 2, \ldots, b$. These random variables are *not* assumed to be independent. Instead, we assume that the correlation between X_i and $X_{i'}$ is

$$\rho(i, i') = \frac{\gamma(i, i')}{\sqrt{\gamma(i, i)}\sqrt{\gamma(i', i')}} = \frac{\gamma(i, i')}{\sigma^2},$$

where the covariance is

$$\gamma(i, i') = E\big[(X_i - \mu)(X_{i'} - \mu)\big]$$

and $\gamma(i, i) = \sigma^2$ for $i = 1, 2, \ldots, b$. In addition, define the *batch mean* as

$$\overline{X} = \frac{1}{b}\sum_{i=1}^{b} X_i$$

and the *batch variance* as

$$S^2 = \frac{1}{b}\sum_{i=1}^{b}(X_i - \overline{X})^2.$$

Both $\overline{X}$ and S^2 are random variables. We will now compute the expected value of each, along with the variance of $\overline{X}$.

Expected Value of the Batch Mean. The expected value of $\overline{X}$ is particularly easy to calculate. Because the expectation operator $E[\,\cdot\,]$ is *linear*, it follows that

$$E[\overline{X}] = \frac{1}{b}\sum_{i=1}^{b} E[X_i] = \frac{1}{b}\sum_{i=1}^{b} \mu = \mu.$$

As was discussed in Section 8.1, this result confirms that the batch (sample) mean $\overline{X}$ is an *unbiased* estimator of μ.

Expected Value of the Batch Variance. To calculate the expected value of S^2, we begin with the observation that

$$\frac{1}{b}\sum_{i=1}^{b}(X_i-\mu)^2 = \frac{1}{b}\sum_{i=1}^{b}\left((X_i-\overline{X})+(\overline{X}-\mu)\right)^2$$

$$= \frac{1}{b}\sum_{i=1}^{b}(X_i-\overline{X})^2 + \frac{2}{b}(\overline{X}-\mu)\sum_{i=1}^{b}(X_i-\overline{X}) + \frac{1}{b}\sum_{i=1}^{b}(\overline{X}-\mu)^2$$

$$\vdots$$

$$= S^2 + (\overline{X}-\mu)^2.$$

Therefore,

$$E\left[S^2\right] + E\left[(\overline{X}-\mu)^2\right] = \frac{1}{b}\sum_{i=1}^{b}E\left[(X_i-\mu)^2\right] = \frac{1}{b}\sum_{i=1}^{b}\sigma^2 = \sigma^2,$$

where

$$E\left[(\overline{X}-\mu)^2\right] > 0$$

is the variance of $\overline{X}$. Because the variance of $\overline{X}$ is positive, we see that the batch variance S^2 is a *biased* estimator of the common variance σ^2; $E[S^2] < \sigma^2$, because the batch variance *underestimates* σ^2 by the amount $E[(\overline{X}-\mu)^2]$. As will now be shown, the extent of this expected underestimation (bias) is determined by the correlation and can be significant.

Variance of the Batch Mean. To calculate the expected value of $(\overline{X}-\mu)^2$, recognize that

$$\overline{X}-\mu = \frac{1}{b}\sum_{i=1}^{b}(X_i-\mu),$$

and so

$$(\overline{X}-\mu)^2 = \frac{1}{b^2}\sum_{i=1}^{b}(X_i-\mu)\sum_{i'=1}^{b}(X_{i'}-\mu) = \frac{1}{b^2}\sum_{i=1}^{b}\sum_{i'=1}^{b}(X_i-\mu)(X_{i'}-\mu).$$

Therefore, from the definition of $\gamma(i,i')$ and $\rho(i,i')$,

$$E[(\overline{X}-\mu)^2] = \frac{1}{b^2}\sum_{i=1}^{b}\sum_{i'=1}^{b}\gamma(i,i') = \frac{\sigma^2}{b^2}\sum_{i=1}^{b}\sum_{i'=1}^{b}\rho(i,i').$$

The following theorem summarizes the previous discussion. The details of the proof are left as an exercise.

Theorem F.4 Let $X_1, X_2, \ldots, X_b$ be a batch (sequence) of random variables with common mean μ and common variance σ^2. Let $\rho(i, i')$ be the correlation between X_i and $X_{i'}$ for $i = 1, 2, \ldots, b$ and $i' = 1, 2, \ldots, b$. Then the expected value of the batch mean is

$$E[\overline{X}] = \mu \qquad \text{(the batch mean is an } \textit{unbiased} \text{ estimator of } \mu\text{)},$$

the expected value of the batch variance is

$$E[S^2] = \beta\sigma^2 \qquad \text{(the batch variance is a } \textit{biased} \text{ estimator of } \sigma^2\text{)},$$

and the variance of the batch mean is

$$E\left[(\overline{X} - \mu)^2\right] = (1 - \beta)\,\sigma^2,$$

where the *bias* in the batch variance is

$$\beta = 1 - \frac{E\left[(\overline{X} - \mu)\right]^2}{\sigma^2} = 1 - \frac{1}{b^2}\sum_{i=1}^{b}\sum_{i'=1}^{b}\rho(i, i').$$

The bias in the batch variance is determined by the extent to which β is different from 1. This is illustrated by two important examples.

Example F.7
If $\rho(i, i') = 0$ for $i \neq i'$ (this is the case if the X_i's are independent, and thus uncorrelated), then, because $\rho(i, i) = 1$ for $i = 1, 2, \ldots, b$,

$$\beta = 1 - \frac{1}{b^2}\sum_{i=1}^{b}\sum_{i'=1}^{b}\rho(i, i') = 1 - \frac{1}{b^2}\sum_{i=1}^{b}\rho(i, i) = 1 - \frac{1}{b} = \frac{b-1}{b}.$$

Therefore, in this case, the expected value of the batch variance is

$$E[S^2] = \left(\frac{b-1}{b}\right)\sigma^2.$$

The result in Example F.7 is the basis for the interval-estimation equation established in Section 8.1: To estimate the variance of an *independent* sample (batch) of size b, the (slight) bias in the batch variance can be removed by using

$$\frac{1}{\beta}S^2 = \frac{1}{b-1}\sum_{i=1}^{b}(X_i - \overline{X})^2 \qquad \text{in place of} \qquad S^2 = \frac{1}{b}\sum_{i=1}^{b}(X_i - \overline{X})^2.$$

Unless b is small, this bias correction is largely irrelevant. As the next example illustrates, however, the bias correction is typically *not* irrelevant for correlated observations—which are the usual case in time-sequenced data generated by a discrete-event simulation.

Example F.8

If $\rho(i, i') = \rho_{|i-i'|}$ for $i \neq i'$, then the correlation matrix has the form

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{b-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{b-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{b-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{b-1} & \rho_{b-2} & \rho_{b-3} & \cdots & 1 \end{bmatrix}.$$

This assumption is known as *weak stationarity*. (See Alexopoulos and Seila, 1998.) In this case,

$$\begin{aligned} \beta &= 1 - \frac{1}{b^2}\sum_{i=1}^{b}\sum_{i'=1}^{b}\rho(i, i') \\ &= 1 - \frac{1}{b^2}\big(b + 2(b-1)\rho_1 + 2(b-2)\rho_2 + \cdots + 2\rho_{b-1}\big) \\ &= \frac{b-1}{b} - \frac{2}{b}\sum_{j=1}^{b-1}\left(1 - \frac{j}{b}\right)\rho_j, \end{aligned}$$

where ρ_j is the autocorrelation between X's separated by a lag j. Therefore, in this case, the expected value of the batch variance is

$$E[S^2] = \beta\sigma^2 \qquad \text{where} \qquad \beta = \frac{b-1}{b} - \frac{2}{b}\sum_{j=1}^{b-1}\left(1 - \frac{j}{b}\right)\rho_j.$$

We can remove the bias introduced by the autocorrelation by using S^2/β to estimate σ^2. To do so, however, we must know $\rho_1, \rho_2, \ldots, \rho_{b-1}$. In practice, the best we can hope for is to *estimate* these autocorrelations by using Definitions 4.4.5 and 4.4.6—that is, we can estimate β as

$$\hat{\beta} = \frac{b-1}{b} - \frac{2}{b}\sum_{j=1}^{b-1}\left(1 - \frac{j}{b}\right) r_j,$$

where program acs is used to compute r_j for $j = 1, 2, \ldots, k$ and it is assumed that $\rho_j = 0$ for $j > k$, so that, if $b - 1 > k$, then $r_j = 0$ for $j = k, k+1, \ldots, b-1$.

Would the assumptions for Example F.8 hold for the wait times for jobs in a congested $M/M/1$ service node in steady state? First, consider the individual wait times. Does each job's wait time have the same mean and variance, i.e., does $E[X_i] = \mu$ and $V[X_i] = \sigma^2$ for some large index i? Intuitively, the answer is "yes"—all wait times should have identical distributions at steady state. Second, consider the correlation structure. Does the correlation between the wait times for job i and job i' depend only on $|i - i'|$ for the service node at steady state? As a specific instance, is the correlation between the wait times X_{946} and X_{949} (three jobs apart) the same as the correlation between X_{971} and X_{974} (also three jobs apart)? As in the first instance, the answer is probably "yes"—it is reasonable so assume that correlation depends only upon the difference in the indices of the jobs.

Example F.9
For example, the autocorrelation estimates from Example 4.4.3 were used to compute $\hat{\beta}$ for various batch sizes b. The "cut-off" autocorrelation lag was, somewhat arbitrarily, taken to be $k = 100$. The results were

b	16	32	64	128	256	512	1024
$\hat{\beta}$	0.17	0.29	0.43	0.60	0.77	0.88	0.94

and we see, as expected, $\hat{\beta} \to 1$ as $b \to \infty$—that is, for *large* batch sizes, the bias in the variance estimate vanishes. Note, however, that, for this example, positive autocorrelation causes small batches to yield a variance estimate that is (on average) too small by a *significant* factor.

F.3 Exercises

F.1 Prove that the two equations

$$\gamma_{UV} = E[(U - \mu_U)(V - \mu_V)] \quad \text{and} \quad \gamma_{UV} = E[UV] - \mu_U\mu_V$$

are equivalent.

F.2 Given that definition of the correlation coefficient of two random variables U and V is

$$\rho_{UV} = \frac{\gamma_{UV}}{\sigma_U \sigma_V},$$

prove that (*a*) $|\rho_{UV}| \le 1$; (*b*) $|\rho_{UV}| = 1$ if and only if there exist constants (a, b, c) such that $aU + bV + c = 0$. *Hint*: Minimize $E\left[(\alpha(U - \mu_U) + \beta(V - \mu_V))^2\right] \ge 0$ subject to the constraint $\alpha^2 + \beta^2 = 1$.

F.3 Relative to Example F.2, suppose the draw is *without* replacement, from an urn with n black and n white balls. (*a*) What is ρ_{UV}? (*b*) Note that $\rho_{UV} \to 0$ as $n \to \infty$. Why? (*c*) Verify the correctness of your ρ_{UV} equation via a Monte Carlo simulation for the case $n = 5$.

F.4 Suppose the relation in Example F.3 is $V = U^3$ (rather than $V = U^2$). (*a*) What is ρ_{UV}? (*b*) Comment.

F.5 Prove Theorem F.2.

F.6 How would Examples F.4 and F.6 change if the service discipline is "shortest job first"? Conjecture first, then simulate.