Uses of Moments

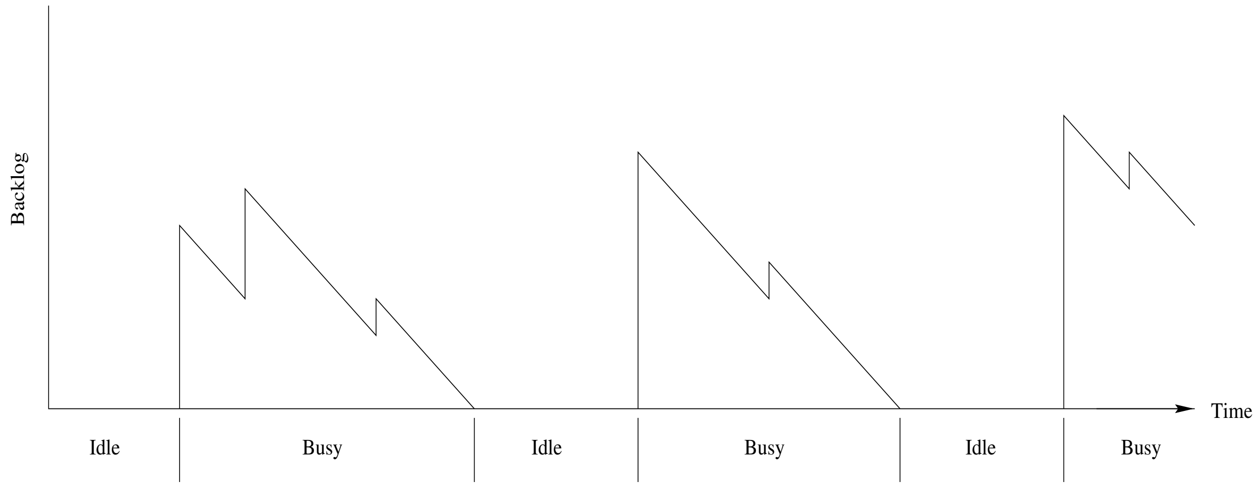
1. The first four moments are routinely used in statistical applications to describe the *location*, *spread*, *skewness*, and *heavy-tailedness* of a distribution (see *Kendall's Advanced Theory of Statistics – Stuart and Ord, 1994*).

2. The moments uniquely define the underlying probability distribution when its moment generating function exists.

3. The quality of any approximating probability distribution may be ascertained by comparing a large number of moments.

4. Higher order moments are often used to study tail behaviour and asymptotic properties of a distribution.

The Duration of a Busy Period



Let B denote the duration of a busy period with Laplace-Stieltjes transform (LST)

$$\Phi_B(s) = E(e^{-sB}), Re(s) \geq 0.$$

Now consider an M/G/1 queue with arrival rate λ and service time random variable S having LST $\phi(s)$ and moments $\mu_k = E(S^k)$.

- Kleinrock, 1975, p.212:

$$\Phi_B(s) = \phi(s + \lambda - \lambda\Phi_B(s)) \qquad (1)$$

- Related Item – The delay cycle transform (Takagi, 1991, pp.23-24):

$$\Phi_D(s) = \Phi_I(s + \lambda - \lambda\Phi_B(s))$$

Symbolic Derivation of Busy Period Moments

We will consider 3 methods for doing this:

1. Takács' explicit formula

2. Implicit differentiation and recursive evaluation

3. Algorithm based on *integer partitions*

Method 1

Takács' (1963) explicit formula:

$$E(B^r) = \frac{(-1)^{r-1}}{\lambda} \left\{ \frac{\partial^{r-1}}{\partial s^{r-1}} \Big(\frac{s}{s - \lambda + \lambda\phi(s)} \Big)^r \right\} \Big|_{s \to 0}$$

Method 2

* Implemented symbolically by Takagi and Sakamaki (1996)

Differentiate both sides of (1) repeatedly so that the first few derivatives look like

$$\Phi_B^{(1)}(s) = \{1 - \lambda\Phi_B^{(1)}(s)\}\phi^{(1)}(s + \lambda - \lambda\Phi_B(s))$$

$$\begin{aligned}\Phi_B^{(2)}(s) = &\{1 - \lambda\Phi_B^{(1)}(s)\}^2\phi^{(2)}(s + \lambda - \lambda\Phi_B(s)) \\ &- \lambda\Phi_B^{(2)}(s)\phi^{(1)}(s + \lambda - \lambda\Phi_B(s))\end{aligned}$$

$$\begin{aligned}\Phi_B^{(3)}(s) = &\{1 - \lambda\Phi_B^{(1)}(s)\}^3\phi^{(3)}(s + \lambda - \lambda\Phi_B(s)) \\ &- 3\{1 - \lambda\Phi_B^{(1)}(s)\}\lambda\Phi_B^{(2)}(s)\phi^{(2)}(s + \lambda - \lambda\Phi_B(s)) \\ &\quad - \lambda\Phi_B^{(3)}(s)\phi^{(1)}(s + \lambda - \lambda\Phi_B(s)) \\ &\vdots\end{aligned}$$

Method 2 (continued)

When $r = 2$, for example, the solution for $\Phi_B^{(2)}(s)$ is given by

$$\Phi_B^{(2)}(s) = \frac{\{1 - \lambda\Phi_B^{(1)}(s)\}^2 \phi^{(2)}(s + \lambda - \lambda\Phi_B(s))}{1 + \lambda\phi^{(1)}(s + \lambda - \lambda\Phi_B(s))}.$$

By evaluating the above transform at $s = 0$ and making the substitution

$$\Phi_B^{(1)}(0) = -E(B) = -\frac{\mu_1}{1 - \lambda\mu_1},$$

one obtains

$$\Phi_B^{(2)}(0) = E(B^2) = \frac{\mu_2}{(1 - \lambda\,\mu_1)^3}.$$

Some Remarks

Methods 1 and 2 both lend themselves readily to symbolic computation, as:
(i) direct implementation is straightforward
(ii) pleasing in its simplicity
(iii) reflects the power of symbolic algebra

On the other hand, both methods are not terribly efficient and do not reflect the deeper elegant structure of the problem.

Intermediate-expression Swell (Kendall, 1998)

- the tendency for naive algorithms to generate huge expressions in a calculation where the final result is concise

Method 3

* Implemented symbolically by Drekic and Stafford (2002)

Consider differentiating a function of the form

$$f(\mathrm{s}) = (g \circ h)(\mathrm{s}) = g(h(\mathrm{s}))$$

Let:

- $s_i, s_j, \ldots$ denote the components of s
- $\partial_i, \partial_{ij}, \ldots$ denote the operators $\frac{\partial}{\partial s_i}, \frac{\partial^2}{\partial s_i \partial s_j}, \ldots$
- $h_i, h_{ij}, \ldots$ denote the quantities $\partial_i h(\mathrm{s}), \partial_{ij} h(\mathrm{s}), \ldots$
- $g_r = g^{(r)}(h(\mathrm{s}))$

Then:

$$\begin{aligned}
\partial_i f(\mathrm{s}) &= h_i g_1 \\
\partial_{ij} f(\mathrm{s}) &= h_{ij} g_1 + h_i h_j g_2 \\
\partial_{ijk} f(\mathrm{s}) &= h_{ijk} g_1 + h_{ij} h_k g_2 + h_{ik} h_j g_2 + h_{jk} h_i g_2 + h_i h_j h_k g_3
\end{aligned}$$

Proceeding in this fashion...

$$\partial_{i_1 i_2 \cdots i_r} f(\mathbf{s}) = \sum_{p \in \mathcal{P}_{V_r}} \left\{ \left[\prod_{b \in p} h_b \right] g_{|p|} \right\} \qquad (2)$$

where:

- $V_r = \{i_1, \ldots, i_r\}$
- $\mathcal{P}_{V_r}$ is the full partition of V_r
- $p = (b_1 | \cdots | b_k)$ is a particular partition of V_r into k blocks
- b is an arbitrary block of a partition
- $|p|$ denotes the number of blocks in the partition p

Example: Let $i_1 = i$, $i_2 = j$, $i_3 = k$. For the 3rd derivative:

$$V_3 = \{i, j, k\}$$

$$\mathcal{P}_{V_3} = \{(ijk), (ij|k), (ik|j), (jk|i), (i|j|k)\}$$

Consider the **scalar case** (i.e. $i = j = k$)

Then:

$$\begin{aligned}
\partial_i f(s) &= h_i g_1 \\
\\
\partial_{ii} f(s) &= h_{ii} g_1 + h_i h_i g_2 \\
&= h_{ii} g_1 + h_i^2 g_2 \\
\\
\partial_{iii} f(s) &= h_{iii} g_1 + h_{ii} h_i g_2 + h_{ii} h_i g_2 + h_{ii} h_i g_2 + h_i h_i h_i g_3 \\
&= h_{iii} g_1 + 3 h_{ii} h_i g_2 + h_i^3 g_3
\end{aligned}$$

In the case of the 3rd derivative:

$$V_3 \to \{i, i, i\}$$

$$\begin{aligned}
\mathcal{P}_{V_3} &\to \{(iii), (ii|i), (ii|i), (ii|i), (i|i|i)\} \\
&\to \{(iii), (ii|i), (i|i|i)\} \\
&\equiv \{\{3\}, \{2, 1\}, \{1, 1, 1\}\}
\end{aligned}$$

In other words, we're dealing with partitions of the integer 3, denoted by $\mathcal{P}_3$.

Therefore, equation (2) in the scalar case becomes

$$f^{(r)}(s) = \sum_{p \in \mathcal{P}_r} \left\{ c_p \left[\prod_{b \in p} h^{(b)}(s) \right] g^{(|p|)}(h(s)) \right\} \tag{3}$$

where c_p is the number of partitions in $\mathcal{P}_{V_r}$ that have block lengths given by the elements of p.

For the busy period LST (1), we set:

$$f \equiv \Phi_B$$

$$g \equiv \phi$$

$$h(s) \equiv s + \lambda - \lambda \Phi_B(s)$$

Then, equation (3) becomes

$$\Phi_B^{(r)}(s) = \sum_{p \in \mathcal{P}_r} \left\{ c_p \left[\prod_{b \in p} (\delta_{b,1} - \lambda \Phi_B^{(b)}(s)) \right] \phi^{(|p|)}(s + \lambda - \lambda \Phi_B(s)) \right\} \tag{4}$$

where $\delta_{i,j}$ denotes the Kronecker delta.

By evaluating (4) at $s = 0$, the following recursive procedure for calculating $E(B^r)$, $r \geq 2$, is obtained:

$$\frac{(-1)^r}{1-\lambda\mu_1} \sum_{p \in \{\mathcal{P}_r - \{r\}\}} \left\{ c_p \left[\prod_{b \in p} \left(\delta_{b,1} + (-1)^{b+1} \lambda E[B^b] \right) \right] (-1)^{|p|} \mu_{|p|} \right\}$$

It is this rule that we implement in *Mathematica* to compute moments of B. With this rule, note that:

- no derivative expression is ever generated in the computation of $E(B^r)$

- all that is needed are the moments of the service time distribution μ_k, the coefficients c_p, and the integer partitions $\mathcal{P}_r$.

Table 1: Comparison of Execution Times (in Seconds)

r	Method 1	Method 2	Method 3
2	0.031	0.016	0.001
4	0.109	0.094	0.047
6	1.27	0.609	0.203
8	74.6	2.72	0.718
10	4932	3.65	2.28
11	———	6.28	4.23
12	———	15.9	7.78
13	———	32.9	15.9
14	———	83.9	17.3
16	———	542	129
17	———	1437	346
18	———	———	934
19	———	———	2712
20	———	———	7177

References

Drekic, S., J.E. Stafford. 2002. Symbolic computation of moments in priority queues. *INFORMS Journal on Computing* **14** 261-277.

Kendall, W.S. 1998. Computer algebra. P. Armitage, T. Colton, eds. *Encyclopedia of Biostatistics*. John Wiley & Sons, New York.

Kleinrock, L. 1975. *Queueing Systems, Volume I: Theory*. John Wiley & Sons, New York.

Stuart, A., J.K. Ord. 1994. *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory*. John Wiley & Sons, New York.

Takács, L. 1963. Delay distributions for one line with Poisson input, general holding times, and various orders of service. *The Bell System Technical Journal* **42** 487-503.

Takagi, H. 1991. *Queueing Analysis: A Foundation of Performance Evaluation, Volume 1: Vacation and Priority Systems, Part 1*. North-Holland, Amsterdam, The Netherlands.

Takagi, H., K. Sakamaki. 1996. Moments for M/G/1 queues. *The Mathematica Journal* **6** 75-80.

Wolfram, S. 2003. *The Mathematica Book*, 5^{th} Edition. Wolfram Media, Inc.