IEOR 6711: Stochastic Models I

Fall 2010, Professor Whitt

Topics for Discussion, Thursday, September 30

More on the Infinite-Server Queue and Staffing

We continued discussing the papers posted last Tuesday. We emphasized Theorem 1 of the 1993 "Physics" paper. It describes the $M_t/GI/\infty$ infinite-server queueing model, with a nonhomogeneous Poisson arrival process. See Prop. 2.3.2, Example 2.3(B), Example 2.3(C) and Section 2.4 of the Ross textbook for related material.

The following are the main topics discussed.

1. **the time-varying mean** $m(t)$

   We discussed three one-dimensional integral representations that can be obtained from the two-dimensional integral

   $$m(t) = \int \int_A \lambda(s) g(x) \, ds \, dx,$$

   where A is the upper triangular region in Figure 1 of the physics paper. We obtain these formulas by considering three different outer integrals: (i) we can integrate $g(x)$ from 0 to ∞, (ii) we can integrate $\lambda(s)$ from $-\infty$ to t, and (iii) we can integrate $\lambda(t-s)$ over s from 0 to ∞. We then define the appropriate inner integral in each case. In case (i) we get

   $$\int_0^\infty g(x) \int_{t-x}^t \lambda(u) \, du = E[\int_{t-S}^t \lambda(u) \, du].$$

   We get other formulas in the other cases, as in Theorem 1 of the physics paper and its proof.

2. **not a Poisson process**

   $Q(t)$ has a Poisson distribution for each t, but $\{Q(t) : t \geq 0\}$ is not a Poisson process, e.g., it does not have nondecreasing sample paths w.p.1.

3. **the covariance**

   Theorem 2 describes $Cov(Q(t), Q(t+u))$. The idea is to exploit the random measure representation and the picture, here Figure 3. We see that $Q(t) = X + Y$, while $Q(t+u) = Y + Z$, where X, Y and Z are independent. Hence

   $$Cov(Q(t), Q(t+u)) = Cov(Y, Y) = Var(Y) = E[Y],$$

   where $E[Y]$ has a simple integral formula, like the mean $m(t) = E[Q(t)]$.

4. **the departure process and the departure rate**

   Theorem 1 in the physics paper includes a description of the departure process and the departure rate. The random variables $Q(t)$ and $D(t)$ are independent, because they correspond to disjoint sets in Figure 1. Note that the departure rate has a formula closely related to the mean $m(t) \equiv E[Q(t)]$.

5. **ODE with M service**

   Theorem 6 and Corollary 4 show that the mean $m(t)$ satisfies an ODE when the service-time distribution is exponential. That reveals how the peaks of m and λ are related (when $E[S] = 1$). In particular, that explains why the curve for $m(t)$ crosses the curve for $\lambda(t)E[S]$ where the derivative $\dot{m}(t) = 0$, e.g., where $m(t)$ assumes its maximum.

6. **linear, quadratic and Taylor**

   We discussed simple formulas obtained when we approximate the arrival rate function by a linear function or a quadratic function. See (14) in the paper. One way to get that is to consider Taylor series approximations.

7. **relaxation time: approach to steady state**

   For a stationary model, it is important to understand how the system approaches steady state as time evolves, starting with various typical special initial conditions, such as starting empty. A very simple revealing formula exists for the stationary $M/GI/\infty$ model; formula (20).

8. **sinusoidal and other periodic arrival rates**

   The sine paper describes results for periodic arrival rate functions. The key fact is that m inherits structure from λ. Hence revealing formulas are available.

9. **starting in the infinite past**

   It is important to point out and emphasize that the simple formulas depend on starting in the infinite past. Starting at time 0 is covered as the special case in which $\lambda(t) = 0$ for $t < 0$.

10. **Erlangs: the concept of offered load**

    An Erlang is a dimensionless quantity indicating the average amount of load in a stochastic system. It is the mean number of busy servers in an associated infinite-server model. That notion is defined for a stationary model. We want to extend it to a nonstationary model. The first idea is to go beyond arrival rate and include the service requirements. The second idea is to adjust for nonstationarity.

11. **staffing: the 1996 paper**

    We briefly discussed the 1996 staffing paper. The infinite-server (IS) approximation, or offered-load approximation, we have been discussing is contrasted with the pointwise-stationary approximation (PSA) and the simple stationary approximation (SSA) there, for the case of a sinusoidal arrival-rate function. We noted that an explicit formula for $m(t)$ when $\lambda(t)$ is sinusoidal is given in the 1993 sine paper. The function $m(t)$ is also sinusoidal with the same frequency, but there is a time lag and space shift there too. An important concept and method is the modified offered load (MOL) approximation.

    There are other papers, which we did not discuss: There is a 1993 "networks" paper with Massey discussing the extension to networks of service facilities.

12. **staffing: the 2008 paper**

    We briefly discussed the 2008 paper with Feldman et al. discussing an iterative staffing algorithm (ISA), based on simulation, for stabilizing performance in even more general systems. Additional work is in progress with current doctoral student Yunan Liu.