Examples class 1
Computer Systems Modelling

November 13, 2000

Simulation Q4 Suppose that a simulation is constructed to estimate the response time, in milliseconds, of an interactive computer system. A number of repetitions are performed with the following results: 4.1, 3.6, 3.1, 4.5, 3.8, 2.9, 3.4, 3.3, 2.8, 4.5, 4.9, 5.3, 1.9, 3.7, 3.2, 4.1, 5.1.

Calculate the mean and variance of these results and hence derive bounds for a 95% confidence interval. You may wish to make use of the fact that $\mathbb{P}(Z > 1.96) > 0.025 > \mathbb{P}(Z > 1.97)$ where Z is a random variable with the unit normal distribution.

The question was written with the expectation that it would be answered using the central limit theorem and the supplied values from a unit normal distribution. However, the sample size of 17 is rather small for that (30 samples is a typical requirement), so the solution should instead use a t-distribution with 16 degrees of freedom. If you follow the same derivation through using the normal distribution then you will find the confidence interval is erroneously tight. [The use of the t-distribution in this way also requires us to assume that the response time itself has a normal distribution]

Let $\bar{x}$ be the sample mean, s^2 be the sample variance, $n = 17$ be the number of experiments and $x_1 \ldots x_n$ be the experimental results.

Recall the usual definitions for $\bar{x}$ and s^2:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Substituting the experimental results:

$$\bar{x} = \frac{64.2}{17} = 3.78$$

$$s^2 = \frac{13.03}{16} = 0.81$$

$100(1-\alpha)\%$ confidence bounds are given by

$$\bar{x} \pm \frac{t_{\alpha/2} s}{\sqrt{n}}$$

We want 95% bounds, so $\alpha = 0.05$ and $t_{0.025} = 2.12$ (the maximum value t such that $\mathbb{P}(T_{16} > t) > 0.025$). Hence $3.78 \pm (2.12 * \sqrt{0.81/17})$, i.e. 3.78 ± 0.46.

Simulation Q5 An experiment is performed to estimate the performance of a $M/M/1$ system with FIFO queueing. The response times of each of 1000 successive customers are recorded and the mean ($\bar{x}$) and variance (σ^2) of these numbers are calculated.

A naïve student believes that $100(1-\alpha)\%$ confidence bounds for the mean can be derived by the expression

$$\bar{x} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{1000}}$$

where $z_{\alpha/2}$ is defined from the unit normal distribution in the usual manner.

What mistake has this student made? Would their technique be valid if they were estimating the service time of this queue?

The expression for $100(1-\alpha)\%$ confidence bounds and the use of $z_{\alpha/2}$ are correct in themselves. The problem is that those results are not applicable here because they rely on assumptions not present in this system.

Recall that this expression for confidence bounds relies of the *central limit theorem*. That theorem shows that the mean of a set of independent random variates, drawn from an arbitrary distribution, has a limiting CDF approaching a normal random variable.

The issue here is that the samples have been obtained by measuring the response times of 1000 *successive* customers passing through the system. These measurements are not independent and so the central limit theorem should not be used. The lack of independence should be clear: if one customer in the queue experiences a large response time then the FIFO discipline means that subsequent customers are likely to experience a similarly poor response.

In contrast to *response times*, successive *service times* are independent and so it would be valid to generate confidence bounds in this manner.

1996 Paper 9 Q3 The Erlangian distribution E_r, with parameters (r, μ) is given by

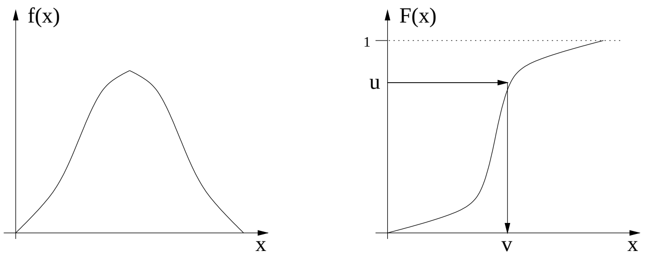
$$f_X(x) = \frac{\mu(\mu x)^{r-1}}{(r-1)!} \mathrm{e}^{-\mu x} \qquad x \geq 0$$

1. *Part (1) to be covered during the second examples class* [5 marks]
2. Describe how random variables from a given distribution function $f_X(x)$ can be sampled for use in a discrete event simulator. [3 marks]
3. Use your answer from the previous question to develop pseudo-code for a function in a discrete event simulator which when called returns a sampled value from the distribution function E_r. State any assumptions made and explain any arguments to the function. Comment on the efficiency of your code. [12 marks]

The general scheme for generating random variates from a particular distribution is to derive the inverse of the CDF. Let $f_X(x)$ and $F_X(x)$ be respectively the pdf and CDF of the required distribution.

Recall that

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(x)\, dx$$



Additionally, because the pdf is always positive, the CDF is a monotonically increasing function from 0 to 1. If v takes random values according to the desired distribution then $F_X(v)$ has uniform distribution between 0 and 1.

Conversely, if u takes uniformly distributed random values between 0 and 1 then $F_X^{-1}(u)$ has the required distribution.

The E_r distribution in this question can be considered to be an extension of the exponential distribution. It has two parameters, r and μ. The exponential distribution is achieved when $r = 1$. As shown here, the mean of the distribution is r/μ and the coefficient of variation is $1/\sqrt{r}$. Note that the parameters are sometimes expressed in different forms: the common alternatives give the mean directly proportional (rather than inversely) to the second parameter, or in fact give the mean equal to the second parameter (rather than depending on r).

In the case of the E_r it is generally not straightforward to integrate and invert the pdf. The preferred method is to realize that values from E_r can be calculated by summing r values from an exponential distribution with parameter μ.

The pdf for the exponential is $f_{exp}(x) = \mu e^{-\mu x}$, integrating gives the CDF $F_{exp}(x) = 1 - e^{-\mu x}$ and inverting this gives $F_{exp}^{-1}(x) = -\frac{1}{\mu}\ln(1-x)$.

So, if $u_1 \ldots u_r$ are uniformly distributed random numbers between 0 and 1, we can generate random E_r values by

$$-\frac{1}{\mu}\sum_{1}^{r}\ln(1-u_i)$$

In practise we can avoid r subtractions by using u_i in preference to $1-u_i$ (both are uniformly distributed between 0 and 1).

We may also wish to express the sum-of-logarithms as a logarithm-of-products:

$$-\frac{1}{\mu}\ln\prod_{1}^{r}u_i$$

As an extension, you could try to confirm the validity of the assertion that a value from E_r can be generated by summing r values from E_1. If $h_r(x)$ is the pdf for E_r then, if the assertion is true, we require

$$h_r(x) = \int_0^x h_{r-1}(y)h_1(x-y)\,dy$$

because a total of x from r E_1 variates can be obtained by a total of y from the first $r-1$ and the remaining $x-y$ from the last.

Operational analysis Q2 A printer and print queue were observed for 48 minutes during which time 16 jobs completed. The mean queue length was 4 jobs. The printer could print 8 pages per minute. The mean job size was 12 pages.

Assuming that the number of arrivals was equal to the number of completions, calculate (a) the throughput of the system in jobs-per-minute, (b) the mean residence time of a job, (c) the mean queuing time of a job and (d) the utilization of the printer.

Using the notation from the slides:

48 min	T	Observation period
16 jobs	C	Completion count
4 jobs	N	Mean queue length
$12/8 = 1.5$ min	S	Mean service requirement

(*a*) $X = C/T = 16/48 = 0.33\,\text{min}^{-1}$. (By definition of throughput)

(*b*) $R = N/X = 4 * 3 = 12\,\text{min}$. (Little's law)

(*c*) $R - S = 12 - 1.5 = 10.5\,\text{min}$. (Subtracting service requirement from residence time)

(*d*) $U = XS = 0.33 * 1.5 = 0.5$. (Utilization law)

Operational analysis Q4 Suppose that busses run according to a regular timetable with a fixed inter-arrival time of 10 minutes. What is the mean waiting time of a customer arriving at the bus stop?

Suppose instead that bus inter-arrival times are exponentially distributed, still with mean 10 minutes. What now is the mean waiting time of an arriving customer?

Derive a general relationship between the mean waiting time and the first and second moments of the inter-arrival distribution. *Hint: if the inter-arrival time pdf is $f(x)$ then note that the probability of encountering an interval of length τ is proportional to both $f(\tau)$ and to τ.*

Kleinrock contains rigorous derivations of these results in section 5.2.

In each case we assume that customer arrivals are independent of bus arrivals.

Firstly, suppose that the timetable is regular with a fixed 10 minute inter-arrival time. Evidently each inter-arrival period lasts 10 minutes and so, by the assumed independence, a customer arriving will have a mean waiting time of 5 minutes.

If the inter-bus arrival times are exponentially distributed then the mean waiting time is 10 minutes by the memoryless property. However, note that a similar argument shows that the mean time *since the last bus departed* is also 10 minutes. How can we reconcile that 20 minute interval with the mean of 10 minutes? The intuitive explanation is that an arriving customer is more likely to encounter a long inter-arrival period than a shorter one.

Suppose that the inter-bus arrival distribution is defined by a pdf $f(x)$. Let $h(x)$ be the pdf of the inter-arrival distribution *as seen by an arriving customer*. The probability of encountering an inter-arrival period of length x is proportional both to $f(x)$ and, by the assumed independence of customer arrivals, to x itself. Hence $h(x) \propto xf(x)$, or $h(x) = cxf(x)$ introducing c as the constant of proportionality.

Let $\mathbb{E}(X)$ and $\mathbb{E}(X^2)$ be the first and second moments of $f(x)$. Noting the $f(x)$ is only non-zero for positive values, recall that:

$$
\begin{aligned}
\mathbb{E}(X) &= \int_0^\infty xf(x)\,dx \\
\mathbb{E}(X^2) &= \int_0^\infty x^2 f(x)\,dx
\end{aligned}
$$

We know $h(x)$ integrates to one:

$$\int_0^\infty cxf(x)\,dx = c\int_0^\infty xf(x)\,dx = c\,\mathbb{E}(X) = 1$$

Hence $c = 1/\mathbb{E}(X)$.

We now derive the mean inter-arrival time encountered by an arriving customer:

$$\int_0^\infty xh(x)\,dx = c\int_0^\infty x^2 f(x)\,dx = \mathbb{E}(X^2)/\mathbb{E}(X)$$

A random customer arrival can occur anywhere between two busses with equal probability, so the mean waiting time is simply $\frac{1}{2}\mathbb{E}(X^2)/\mathbb{E}(X)$

We can use this expression to confirm our original results for deterministic and exponentially-distributed inter-arrival distributions. If we define σ^2 to be the variance of $f(x)$ then $\mathbb{E}(X^2) = \sigma^2 + (\mathbb{E}(X))^2$.

For the regular inter-arrival distribution: $\sigma^2 = 0$, $\mathbb{E}(X) = 10$, so the mean waiting time is $\frac{1}{2}10^2/10 = 5$ minutes.

For exponentially distributed inter-arrival times: $\sigma^2 = 100$, $\mathbb{E}(X) = 10$, so the mean waiting time is $\frac{1}{2}(100 + 10^2)/10 = 10$ minutes.