Performance Analysis (C336) - Answers to Tutorial Sheet 2

Lecturers: P. Harrison and J. Bradley

Prepared by: A. Trifunovic

February 14, 2005

1 The Poisson Distribution

(a) Given a single trial, let p be the probability that an event A is the outcome of that trial. Now assume that the trials are repeated N times, such that each trial is independent. Denote by A_N the total number of occurrences of event A in the N trials. Show that the probability of A occurring k times, where $0 < k \leq N$, is given by:

$$P(A_N = k) = \binom{N}{k} p^k (1-p)^{N-k}$$

Answer: Let the random variables $X_1, \ldots, X_N$ denote the N trials. There are

$$\frac{N!}{(N-k)!k!} = \binom{N}{k}$$

distinct assignments of exactly k occurences of A to N completed trials. Since each trial X_i is independent, we have that

$$P(X_1 = x_1, \ldots, X_N = x_N) = \prod_{i=1}^{N} P(X_i = x_i)$$

We also know that $P(X_i = A) = p$ and $P(X_i \neq A) = 1 - p$. The result then follows, since the $\binom{N}{k}$ distinct assignments are mutually exclusive. This is the *Binomial Distribution*. QED.

(b) Now suppose that the probability of the event A occuring in any given small time-step δt is $\lambda\delta t$, where λ is some constant. Consider N independent consecutive trials, each taking δt time, such that the time taken for the N trials is $t = N\delta t$. What is the probability that event A occurs k times during the N trials, where $0 < k \leq N$?

Answer: Let $A_{N\delta t}$ denote the total number of occurrences of A after the N consecutive time intervals of δt. We use the binomial distribution from 1(a) to get the expression for the probability of k occurrences of A.

$$P(A_{N\delta t} = k) = \binom{N}{k}(\lambda\delta t)^k(1-\lambda\delta t)^{N-k}$$

QED.

(c) Let A_t denote the number of times that event A occurs during time t. Using (b) or otherwise, show that the probability of the event A occuring k times during time t, for all $k > 0$, is:

$$P(A_t = k) = \frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

(Hint: consider the expression from (b), letting $N \to \infty$)

Answer: We let $N \to \infty$ in Eq. 1. Thus, intuitively we are now performing an 'infinite' number of trials, but the probability of each of the trials yielding A is becoming 'infinitesimal'. Treat the terms in Eq. 1 separately:

$$\begin{aligned}
\lim_{N\to\infty}\binom{N}{k}(\lambda\delta t)^k &= \lim_{N\to\infty}\frac{N!}{(N-k)!k!}\frac{(\lambda t)^k}{N^k} \\
&= \lim_{N\to\infty}\frac{N(N-1)\ldots(N-N+1)(\lambda t)^k}{N^k(N-k)\ldots(N-N+1)k!} \\
&= \lim_{N\to\infty}\frac{N(N-1)\ldots(N-k+1)(\lambda t)^k}{N^k k!} \\
&= \frac{(\lambda t)^k}{k!}
\end{aligned}$$

and in the second term, we substitute $\delta t = t/N$ to get:

$$\begin{aligned}
\lim_{N\to\infty}(1-\lambda\delta t)^{N-k} &= \lim_{N\to\infty}\left(1-\frac{\lambda t}{N}\right)^{N-k} \\
&= e^{-\lambda t}
\end{aligned}$$

Thus, combining the two terms yields the answer. QED.

2 Application to a simple server

The probability distribution in 1(c) is called the *Poisson Distribution*. In terms of applications to computer performance modelling, consider a server and requests being made to the server. Suppose that we think the probability of a request being made to the server in a given small time-step δt is proportional to the time step (with a given proportionality constant λ) and that each request to the server is independent of the other requests (this corresponds exactly to the "infinitesimal definition" given in the lectures). The Poisson distribution will then determine the probability of k requests to our server during a specified time t.

(a) What is the expected number of requests to the server during time t, under these assumptions?

Answer: We can just work it through:

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\
&= e^{-\lambda t} \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{k!} \\
&= (\lambda t) e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} \\
&= (\lambda t) e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \\
&= (\lambda t) e^{-\lambda t} e^{\lambda t} \\
&= \lambda t
\end{aligned}
$$

QED.

Alternatively, we can use the independence of each of the N trials from which we derived the Poisson distribution in $1(c)$. Once again, let the random variables $X_1, \ldots, X_N$ denote the N trials. We know that the X_i are independent. The mean number of occurences of the event A

in each trial is given by:

$$
\begin{aligned}
E(X_i) &= 1 \times \lambda\delta t + 0 \times (1 - \lambda\delta t) \\
&= \lambda\delta t \\
&= \frac{\lambda t}{N}
\end{aligned}
$$

Now summing over the N time-steps (trials), by the independence of each time-step (trial), we get that:

$$
\begin{aligned}
E(X) &= \lim_{N\to\infty} \Big(\sum_{i=1}^{N} E(X_i) \Big) \\
&= \lim_{N\to\infty} \Big(\sum_{i=1}^{N} \frac{\lambda t}{N} \Big) \\
&= \lambda t
\end{aligned}
$$

QED.

(b) Calculate the variance in the number of requests to the server, under these assumptions.

(Note: The variance of a random variable X is defined to be $Var(X) = E(X^2) - E(X)^2$, where $E(X)$ is the expectation of the variable X)

Answer: There are at least three ways to show this. First is to go via moment generating functions (see for example [1]). The moment generating function of a random variable X is defined as in Eq. 1:

$$
M_X(s) = E(e^{sX}) \tag{1}
$$

The moment generating function of the Poisson distributed random variable X is then given by:

$$
\begin{aligned}
M_X(s) &= E(e^{sX}) \\
&= \sum_{k=0}^{\infty} e^{sk} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\
&= \sum_{k=0}^{\infty} (e^s)^k \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\
&= e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t e^s)^k}{k!} \\
&= e^{(e^s - 1)\lambda t}
\end{aligned}
$$

By definition of the moment generating function, $E(X) = M_X'(0)$ and $E(X^2) = M_X''(0)$. We have:

$$M_X'(s) = (\lambda t)e^{(e^s-1)\lambda t}e^t$$

and

$$\begin{aligned} M_X''(s) &= \frac{d}{dt}(\lambda t e^{(e^s-1)\lambda t}e^t) \\ &= (\lambda t)^2 e^{(e^s-1)\lambda t} + (\lambda t)e^t \end{aligned}$$

by the chain and product and chain rule respectively. The variance is then given by:

$$\begin{aligned} Var(X) &= E(X^2) - E(X)^2 \\ &= M_X''(0) - (M_X'(0))^2 \\ &= \lambda^2 t^2 + \lambda t - \lambda^2 t^2 \\ &= \lambda t \end{aligned}$$

QED

The second way is to compute $E(X^2)$ directly.

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{\infty} k^2 \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\ &= e^{-\lambda t} \sum_{k=1}^{\infty} k^2 \frac{(\lambda t)^k}{k!} \\ &= \lambda t e^{-\lambda t} \sum_{k=1}^{\infty} k \frac{(\lambda t)^{k-1}}{(k-1)!} \\ &= \lambda t e^{-\lambda t} \sum_{k=1}^{\infty} \frac{d}{d(\lambda t)} \Big(\frac{(\lambda t)^k}{(k-1)!} \Big) \end{aligned}$$

Now, as the series expansion of the exponential function *converges uniformly* (standard result – see most textbooks on elementary analysis), we can interchange the limits of summation and differentiation:

$$\begin{aligned}
E(X^2) &= \lambda t e^{-\lambda t} \frac{d}{d(\lambda t)} \Big(\sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} \Big) \\
&= \lambda t e^{-\lambda t} \frac{d}{d(\lambda t)} \Big(\lambda t \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} \Big) \\
&= \lambda t e^{-\lambda t} \frac{d}{d(\lambda t)} (\lambda t e^{\lambda t}) \\
&= \lambda t e^{-\lambda t} (e^{\lambda t} + \lambda^2 t^2 e^{\lambda t}) \\
&= \lambda t + \lambda^2 t^2
\end{aligned}$$

and the result follows, given the value of expectation calculated in 2(b). QED

The third way is perhaps the neatest. We will use the independence of each of the N trials (from which we derived the Poisson distribution in $1(c)$) and first calculate the variance of each trial X_i.

$$\begin{aligned}
E(X_i^2) &= 1^2 \times \lambda \delta t + 0^2 \times (1 - \lambda \delta t) \\
&= \lambda \delta t \\
E(X_i)^2 &= (\lambda \delta t)^2 \\
Var(X_i) &= E(X_i^2) - E(X_i)^2 \\
&= \lambda \delta t - (\lambda \delta t)^2 \\
&= \lambda \delta t (1 - \lambda \delta t)
\end{aligned}$$

Now, the variance of the Poisson distribution is given by summing the variances of the N independent trials:

$$
\begin{aligned}
Var(X) &= \lim_{N\to\infty} \Big(\sum_{i=1}^{N} Var(X_i) \Big) \\
&= \lim_{N\to\infty} \Big(\sum_{i=1}^{N} \lambda\delta t(1-\lambda\delta t) \Big) \\
&= \lim_{N\to\infty} \Big(\sum_{i=1}^{N} \frac{\lambda t}{N}\Big(1-\frac{\lambda t}{N}\Big) \Big) \\
&= \lim_{N\to\infty} \Big(\lambda t \Big(1-\frac{\lambda t}{N}\Big) \Big) \\
&= \lambda t
\end{aligned}
$$

QED

References

[1] G. Grimmett and D. Stirzaker. *Probability and Random Processes.* Oxford University Press, Various Editions.