SSE 2300/ CEE 4990: Systems Design and Engineering

Probability Distributions

February 1, 2010

Bernoulli Experiment, Binomial and Geometric Distributions

The underlying assumptions are:

- An experiment consists of a sequence of n smaller experiments called *trials*. n is fixed for a given experiment.
- Each trial has the same two possible outcomes (dichotomous trials) - which we denote by success (S) and failure (F).
- The trials are independent - out come of any particular trial does not effect the impact of any other trial.
- Probability of success is constant from trial to trial and we denote it by p

An experiment that meets the above criteria is a Bernoulli Experiment and the random variable X associated with such an experiment is called a binomial random variable. It is defined as:

$$X = x = number\ of\ successes\ in\ n\ trials$$

The Binomial distribution function gives the probability of x successes in n trials and is given by $b(x; n, p)$ which can be read from the binomial tables:

$$b(x; n, p) = \begin{cases} {}^nC_x p^x (1-p)^{n-x}; x = 0, 1, 2, 3... \\ 0\ otherwise \end{cases}$$

The cumulative distribution function associated is:

$$F(x) = P(X \leq x) = \sum_{y=0}^{y=x} b(x; n, p)$$

Mean and variance of the Binomial distribution function is np and $np(1-p)$ respectively. The probability density function for inter-arrival times between failures or the time to failure after x successes is given by the Geometric distribution as follows:

$$P(X = x, n = x + 1) = (1-p)p^x = q(1-q)^x$$

where p is the probability of a successful trial and $q = (1-p)$. Mean: $1/p$ and variance: $(1-p)/p^2$.

Poisson Process, Poisson and Exponential Distributions

The Poisson arrival counting process is given by $\{N(t), t \geq 0\}$ where $N(t)$ denotes the total number of arrivals up to time t and $N(0) = 0$. In addition the following three assumptions will need to be satisfied.

- Probability that an arrival occurs between t and $(t+\Delta t)$ is given by $\lambda\Delta t+o(\Delta t)$ where Δt is an incremental element and the value of $o(\Delta t)$ compared to the value of Δt is negligible as Δt tends to 0:
$$\lim_{\Delta t\to 0}\frac{o(\Delta t)}{\Delta t}=0 \tag{1}$$
- Probability of more than 1 arrival between t and $t+\Delta t = o(\Delta t)$
- Number of arrivals in non-overlapping intervals are statistically independent

We wish to calculate $p_n(t)$ the probability of n arrivals in a time interval of length t, where n is an integer ≥ 0. We have:
$$p_n(t)=\frac{(\lambda t)^n e^{(-\lambda t)}}{n!}$$
Mean and variance of the Poisson distribution function is λ and λ respectively.

For a Poisson process the inter-arrival times are exponentially distributed and is given by the following distribution where λ is the arrival rate:
$$f(t)=\lambda e^{-\lambda t}, t>=0$$
Mean: $1/\lambda$ and variance: $1/\lambda^2$

Limiting behavior

The binomial distribution function tends to a Normal distribution when p is fixed and n tends to infinity. It tends to the Poisson distribution function when $p\to 0$, $n\to\infty$ and $np\to\lambda\Delta t$ remains constant and is very small.

The exponential function tends to a Normal distribution when $t\to\infty$ and λ remains constant and is very small.

Memoryless Property

Exponential and geometric distributions can be used to model distribution of component life times or inter-arrival times of failures in systems.They are the only distributions that exhibit a memoryless property. The memoryless property states that the distribution of additional lifetime (or the time to next failure) is exactly the same as the original distribution of lifetime (or the time to failure).

Suppose a component life time is exponentially distributed with parameter λ. Then we can say that if the component hasn't failed for a period of t_0 ($t_0>0$) hours then the probability of it not failing for at least another additional t hours is identical to it not failing for t hours. This is stated as:
$$P(X\geq t+t_0|X\geq t_0)=P(X\geq t)$$
The probability of a bus arriving in 40 minutes given that 30 minutes has passed is the same as the probability of the bus arriving in 10 minutes.