Understanding Statistics and Queuing — Probability, Distributions and Statistical Models

Technical White Paper

microsystems

Contents

Foreword

Engineering has always been concerned with *predictability*, the expectation for development efforts to result in products which perform desired functions with acceptable reliability and availability on time and at acceptable cost. However, most branches of engineering experience strict physical and technical limitations that impact what can be economically achieved. As a result, strong emphasis is placed on estimation and calculation, such as the amount of materials involved, the strengths of structures, and whether multiple purposes can be served by any single construction.

The use of the word *architecture* in software design and hardware infrastructure is analogous to its use in civil engineering. Architecture is the science, and art, of bringing together all major elements of an intended construction or reconstruction so as to fulfill the primary purposes of the project. The goal of architecture is to simultaneously fulfill both primary and secondary project objectives while satisfying an engineering sense of correctness, appropriateness, and elegance. When designing a new IT hardware and software infrastructure, such as that needed by Web sites that utilize the capabilities of the Internet to promote interest to other businesses as well as customers, similar estimations and calculations must be done to ensure sufficient capacities are available without compromising economy.

There are many disciplines to engineering. Those involved in meeting performance, reliability, availability, security and other requirements — the techniques collected under the term *architecture* — are among the least understood. As a result, IT system architects often learn their craft on the job, from mentors, and through experience. Consequently, architects often fall prey to two tendencies: they tend to be effective only in developing systems seen before, and every problem is made to fit the architect's preconceptions.

Learning the art of architecture on the job has undesirable consequences. Basing all knowledge on mentoring and experience requires years observing and imitating more seasoned professionals. Eventually, such experience leads to good architectural judgement — but good judgement often comes from learning from mistakes. This slow and error-prone approach is unacceptable in a world where the need for new systems is great. Architecture must be taught to budding architects, and seasoned professionals need better tools with which to evaluate decisions without requiring systems to be built.

To be successful, architecture must be transformed into a rigorous process using quantitative mathematical tools. In an effort to help this process, Sun has developed a series of white papers on *Quantitative Architecture* — the discipline of developing architectures using rigorous techniques that enable architects to predict the non-functional properties of a system to ensure, from the very beginning, that the system meets its goals.

In order to develop a rigorous and quantitative approach to architecture, it is important to understand the disciplines that enable predictions to be made regarding CPU power, network bandwidth, response times, and so on. Toward that end, the white papers address the following topics:

- *Introduction to the Elements of Web Application Architecture*, introduces the sequence of considerations that go into what may be termed modern *dot-com* or *Web-based* architectures, and explains some of the terms and principles that apply.
- *What is Architecture?*, discusses the design processes that lead to an appropriate architecture that meets design and business goals.
- *Simplified Performance Estimation*, presents a simple and approximate performance estimation method for early capacity planning which enables architectural decisions to be made with confidence.
- *Understanding Growth*, discusses ways to make the growth problem easier to understand, and aims to facilitate the finding of good estimates of the effects of growth and alleviate some scalability concerns.

- *Understanding Statistics and Queuing — Probability, Distributions and Statistical Models*, delves deeply into the bases of these kinds of computations, making evident the assumptions in identifying the various kinds of scenarios in which it is possible to obtain useful, and interesting, results.
- *Understanding Capacity and Scalability*, presents both capacity and scalability in a basic mathematical model, and shows how common architectural decisions affect them.
- *Implementing Enterprise Security Policies*, discusses security policies and the potential for machine verification and other techniques.

This series is based on the Sun StartUp Essentials*SM* program, a class devoted to teaching entrepreneurs and their colleagues the steps needed to establish and operate business-to-client or business-to-business Web-enabled sites. Discussions with students revealed that quantifying many decisions enabled good architecture — and that the basis of the methodology is rarely fully understood. These papers aim to bridge the gap, making the underlying theory accessible and able to serve as a firm foundation, and demonstrating, by example, how these foundations support resulting hardware and software architectures.

Understanding Statistics and Queuing

Introduction

Statistics and queuing examples are found in nearly every aspect of daily life. Weather statistics are presented on TV news shows, and viewers are inundated with sports statistics and political polls — from approval ratings to voting predictions to exit polls and more. In addition, practical demonstrations of queuing are everywhere. People form lines in super markets, cars accumulate at tollbooths, people wait for the next available telephone agent, and many similar situations.

The connection between statistics and queuing comes with a general interest in:

- The arrival or departure rates of people, items or services
- The average queue length at various times
- The average queue wait time
- The longest expected time that one might have to wait in a queue

This last question causes consideration of *variation* — often it is not sufficient to consider just the average of some measure, but also how much variation exists in that measure. A supermarket manager may claim that the average waiting time in the checkout queue is just two minutes. Yet, if the variation on that two minutes time is one minute, the wait has an entirely different effect on a customer's day than if the variation were ten minutes, and that customer arrived at the peak waiting time.

Web site designers are concerned with similar matters — the average arrival rates of requests from users using client software, the variation in arrival rates, the time to wait for service of the request, and so on. While many of these considerations and the statistical computations surrounding them are discussed in *Capacity Planning on a Cocktail Napkin*, this paper delves more deeply into the bases of these kinds of computations, making evident the assumptions in identifying the various kinds of scenarios in which it is possible to obtain useful, and interesting, results.

Statistics is concerned with the use of numbers to describe population characteristics. *Population*, a term with a very special meaning in this context, is the complete set or collection of things — things that exist or could exist — in which there is interest and a wish to describe. There are two related concepts in the use of numbers to describe the *average* behavior of collections of activities: *probability* and *frequency distribution*.[1] While the details of the way in which mathematicians relate these two concepts is beyond the scope of this document, illustrations are given, making full appeal to intuition. Nevertheless, by reading the references listed at the end of this paper, more formal, and finer detail, can be investigated.

Probability

Probability is familiar in activities such a playing cards, throwing dice, or in simpler, but seemingly more contrived activities such as tossing a coin. However, coin tossing fosters thinking about some of the more subtle things related to probability.

Consider the concept of a fair coin in the following terms: toss a fair coin 10,000 times, making each toss a good toss in which the coin spins rapidly. The expected number of times the result is heads or is tails is equal — the definition of *fair*. In a real set of experiments, a coin, because of some wear or deformity, may not be fair. Whether a coin is or is not fair could be determined by throwing the coin multiple times (trying not to cause more wear or deformity), and recording the outcomes. This is called taking a *sample of the population*. It is not the whole population, since that is the set of all possible throws. What is needed is to deduce as much and as many of the characteristics of the population, knowing only the results obtained from the sample.

1. Some presentations on this subject differentiate between *frequency* and *relative frequency*, where the latter is the normalized version of the former (*i.e.*, is divided by the total number of occurrences). In this white paper, the term *frequency* always refers to the normalized frequency, and thus corresponds to a probability density.

Imagine the results consisting of the information in Table 1.

Heads	Tails	Total Throws
2521	2479	5000

Table 1 Coin toss results

These numbers imply there is a bias in favor of heads. But this is a sample. Random variation may have resulted in these numbers being obtained. Another trial to obtain an additional sample could conceivably result in favor of tails. Ideally, a sample such as this would enable the determination of how well predictions can be made about the entire population. This difficult question is left to another paper on this topic.

Table 1 could be written a little differently, by dividing all the numbers by the total number of throws, as illustrated in Table 2. The coin toss data now has the form of a *frequency distribution*. The numbers under the *Heads* and *Tails* headings are now probabilities, or frequencies, and it can now be seen how the frequencies are distributed amongst the outcomes. The total frequency or probability must be 1.

Heads	Tails	Total Frequency
0.5042	0.4958	1.0000

Table 2 Frequency distribution of coin toss results

Laws of Probability

Before investigating these observations any further, consider the two major *laws of probability*.

Law 1

Assume the probability of an event is p_1, and the probability of some other event is p_2. If the events are independent — the occurrence or non-occurrence of one event has no effect on the occurrence or non-occurrence of the other event, and vice versa — the probability of both events happening is the product of the probabilities:

$$p_1 \times p_2$$

For example, when throwing a die, the probability of a six is 1/6. When throwing two dice, the probability of two sixes is 1/6 x 1/6, or 1/36. The cumulative effect of such computations is demonstrated by examples that follow below.

Law 2

Assume the probability of an event is p_1, and the probability of another event is p_2. If the events are independent — the occurrence or non-occurrence of one event has no effect on the occurrence or non-occurrence of the other event, and vice versa — the probability of either event happening is the sum of the probabilities, minus the product of the probabilities:

$$p_1 + p_2 - p_1 \times p_2$$

When throwing a die, the probability of a six is 1/6. When throwing two dice, the probability of at least one six is 1/6 + 1/6 – 1/36, or 11/36.

That this formula is correct is not as obvious as the first one. For the case of the dice, the correctness of the answer is easily verified by writing out all 36 of the possible results, and counting how many of the results have at least one six. All the possible results can be shown by writing the numbers 1 through 6 in two rows, where each of the vertically aligned pairs represent the outcomes on a visible face of each of the dice (Figure 1).

1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6

Figure 1 Possible dice toss results

TV Millionaires?

Few people succeed in becoming millionaires on a very popular TV program. Contestants need only answer 15 multiple-choice questions, with assistance from the audience, a friend, and a computer.

If, with assistance, and only having four choices, a contestant has a 99 percent chance of correctly answering each question, what is the probability of answering all 15 questions correctly?

Since each question is independent (see text), multiply the probabilities: $(0.99)^{15}$, or 0.86 (86%). Even with an almost certainty of answering each question, the chance correctly answering all 15 is much smaller. Yet, even 86% of contestants do not become millionaires. Experience shows less than 1 in 200 (0.005) succeed. So, the inverse calculation finds: $(0.005)^{1/15}$, or 0.702 (70%).

Game creators need only enough difficulty to keep the probability of each correct answer below 70% to ensure less than 1 in 200 succeed. As an exercise, check that were the questions slightly more difficult, about 63%, the success rate falls to under 1 in a 1000!

The creators take a different approach—earlier questions are easier, and later ones harder. Nevertheless, questions do not have to be very difficult to ensure that sponsors do not have to pay out too much money too often.

Clearly the number 6 occurs in the first row 6 times out of the 36 entries, and 6 times out of the 36 entries in the second row. Consequently, a reasonable first guess is that there are twelve columns containing at least one six. However, this accounting results in the final right-hand entry being counted twice (the fact that it has two sixes does not affect that it has a least one six). The reason that it has been counted twice is exactly because it has two sixes, and it is apparent that there are 1/36 outcomes with two sixes. As a result, obtaining the correct answer requires the number of entries with two sixes to be subtracted, ensuring that entry is counted only once. This is exactly what the formula for the second probability law says — the events that are counted twice have to be subtracted.

Data Path — An Example

The probability laws have an effect on the construction of Web site infrastructure. Assume there are two components, a black box firewall and a networking hub, connected as shown in Figure 2.

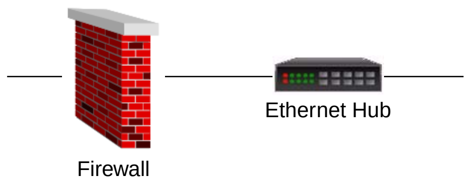

Figure 2 A firewall and network hub

Assume the reliability, or probability of not failing, of the firewall is 98 percent , and the reliability of the hub is 96.5 percent. Inquiring as to the reliability of the two, in terms of the signals or messages that must pass through both, and ignoring the possibility of failure of the connecting cables and their connectors, is a logical question. To obtain this, the probability of both units operating is needed. If it is assumed that failure events are independent, the probability of *both* of them operating is obtained by multiplying the probabilities of *each* of them operating, or 94.6%. Cables and connectors, once tested and proven to work, and not connected and disconnected as part of regular use, have reliabilities above 99.9 percent, or over *three nines*. [1]

1. Cables and connectors suffer are more likely to suffer from external effects, such as being cut by back-hoes, having things dropped on them, or being overly strained.

Even at just the three nines level, the four connections in the above arrangement only reduce the overall reliability by 0.4 percent to 94.2 percent. This is a general feature — as more components are put into the path of a communication or computation, the lower the reliability of the overall arrangement.

If a box fails due to external forces, such as a power spike, or earthquake, then it is likely that the second box would fail in a similar way if it were on the same power feed or in the same building. In this case, the assumption of independence of the failures would not be observed, requiring a different method of analysis of the effects.

Horizontal Scaling — An Example

Assume four computing systems are being used to run application servers in a distributed environment (Figure 3).

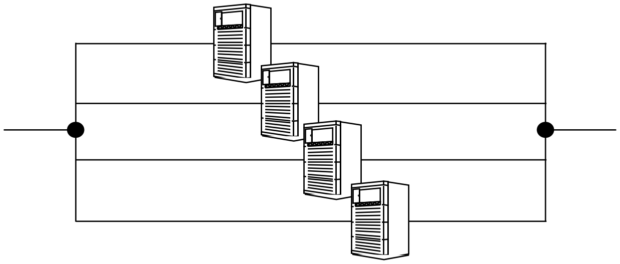

Figure 3 Horizontal server scaling

For the purposes of this discussion, assume the use of machines with low reliability, perhaps only 90 percent. What is the probability that at least *one* server is running at any given time? That is, what is the reliability of the system of four servers connected as shown?

The second probability law can be used by considering this arrangement to be two pairs of machines each consisting of a pair. The reliability of a pair is

$$90\% + 90\% - 90\% \times 90\% = 180\% - 81\% = 99\%.$$

As a result, the reliability of the pair of pairs is:

$$99\% + 99\% - 99\% \times 99\% = 198\% - 98.01\% = 99.99\%.$$

Thus, the reliability of the combination is 99.99 percent, or *four nines*! It is important to remember this represents the probability that at least one server is working — it is not the probability that the power of all four servers is available.

It might appear that the use of a similar parallel computation arrangement allows any arbitrary reliability to be obtained. While it is true that systems can be made more reliable using this technique — the most powerful technique for building reliable systems — there are limits. For example, communications paths must be working, and connectors on both sides of the arrangement must be able to serve the last machine in the chain. Since machine failure order is unknown, all connections to all machines must be working. Considering the eighteen connections in this diagram, and making the same *three nines* assumption used previously, a reliability of approximately 98% is obtained for all these connections being functional all the time. Considering the connections to be in series with the four parallel machines, the combined reliability is 98% x 99.99%, or 97.9%. That is, the overall reliability is limited, in this case, by the wiring.

Clearly cabling and connectors tend to be more reliable than *three nines*. Indeed, they typically exhibit five or six *nines*, so the outlook is not quite so bleak. It does, however, need to be taken into consideration. Extra complexity is introduced when parallel systems are built, ultimately limiting improvements in reliability that can be gained by doubling up the machinery. Other sources of extra complexity include networking and software mechanisms needed to direct requests to the different machines, balance the load, and to direct output back via appropriate paths to get back to the original source of the request. Creating, installing, and maintaining this extra complexity is an ongoing cost.

In addition, such paralleling does not necessarily mitigate the risks that affect the systems in common. Such risks might include being in the same earthquake zone, being subject to building problems such as sharing power feeds, air-conditioning systems, common wiring channels, electromagnetic interference from storms, malfunctioning fire control systems, and so on. Other steps need to be taken to deal with such common source threats, such as separating the systems across geographical regions with differing earthquake risk, and using multiple power feeds from different sources. Nevertheless, the principle of putting system components in parallel to improve reliability remains one of the most powerful techniques available to increase reliability figures to the levels needed by such systems.

It is important to note that a frequent source of common threats to parallel systems is their control by copies of the same software systems. While installed software does not degrade over time in the same way as hardware, software does often contain bugs that can cause cumulative errors. As a result, the eventual software failure appears to be in a manner similar to a hardware failure. Having multiple copies of the same software on multiple (parallel) servers exposes server each to the same (common) set of risks. The good news is, of course, that when a bug is found and repaired on one, the same repair can be easily propagated to all systems.

Another surprising example of the sometimes unexpected effects of probabilistic interactions is demonstrated in the problem discussed in Appendix A.

Introducing Distributions

Information about dice can be viewed in a different way. When dice are used in games, it is the sum of the values on the dice that determine the course of the game. What are the probabilities of getting the various sums? Figure 4 illustrates the various combinations of results from throwing two dice.

1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6
2	3	4	5	6	7	3	4	5	6	7	8	4	5	6	7	8	9	5	6	7	8	9	10	6	7	8	9	10	11	7	8	9	10	11	12

Figure 4 Possible results for dice toss sums

Counting the outcomes, each sum shown in the top row of Figure 5 occurs as shown in the bottom row.

2	3	4	5	6	7	8	9	10	11	12
1	2	3	4	5	6	5	4	3	2	1

Figure 5 Possible frequency results for dice toss sums

This provides another example of a distribution. The information can be presented in terms of probabilities by dividing each of these values by 36 (the total number of outcomes). Consequently, when throwing two dice, there is one chance in 36, or a probability of 1/36, of getting a sum of 2, 6/36 = 1/6 of getting a sum of 7, and 1/18 of getting a sum of 11.

Another interesting question could arise. What is the probability of getting a sum of 9 or less? A naïve answer would be 3/4, because 9 is three-quarters of 12. However, if all the counts for 9 or less are added together, there are 30 outcomes with values less than or equal to 9. As a result, the answer is 30/36, or 5/6, or approximately 83 percent. It is said that the value 9 is at the *83rd percentile,* meaning that 83 percent of the possible outcomes are at, or less than, the value 9.

Describing Distributions

Certain numbers are used to help describe distributions, and three are used to give a sense of the *center* of the distribution: *mode*, *median*, and *mean*.

- *Mode* is the value that occurs most frequently in the distribution. In the dice distribution, the most frequent value is 7. Sometimes there are several values that occur more frequently than others, separated by values that are less. In this case, the distribution is described as being bi-, tri-, or even multi-modal.
- *Median* is the value corresponding to the 50^{th} percentile — the value which has as many occurrences above it as below it. In the dice distribution, the discreteness of the distribution results in no value that corresponds exactly to 50 percent. On the other hand, the value 7 has as many greater than it has less than it. As result, it is satisfactory to say that the value 7 is the median in most cases.
- *Mean* is the *balance* point of the distribution. Imagine that the distribution is represented by a shape whose height is proportional to the frequencies, and that the shape is balanced on a knife edge. The point at which it balances is the mean. In the dice distribution, the value 7 is also the mean.

  The mean is calculated by doing a calculation corresponding to the balancing described above — computing the *average moment*, the weight or frequency of each value multiplied by the lever arm, or value. The result is the *average moment* (which is why the mean is often called the average). If the observations for a distribution are designated as x_i, it is common to designate the mean by the notation $\bar{x}$. Using the dice distribution, the calculation to find the mean is as follows:

$$\frac{1}{36}\times 2+\frac{2}{36}\times 3+\frac{3}{36}\times 4+\frac{4}{36}\times 5+\frac{5}{36}\times 6+\frac{6}{36}\times 7+\frac{7}{36}\times 8+\frac{8}{36}\times 9+\frac{9}{36}\times 10+\frac{10}{36}\times 11+\frac{11}{36}\times 12$$
$$= (2+6+12+20+30+42+40+36+30+22+12)/36$$
$$= 7$$

The dice distribution is particularly uninteresting since the mode, median, and mean are all the same. Such a distribution is said to be *symmetric*, and many distributions are indeed so. It is, however, by no means always the case, as some examples later in this paper illustrate.

Other Measures of Interest

Other measures of interest are those that determine how close the rest of the distribution is to the center. One method is to measure and quote the *range* of the distribution.

- *Range* is the distance of certain points in the distribution from the mean. Typically the maximum and the minimum observations in the distribution are used.

  In the dice distribution, the range could be stated as 7 ± 5. The value 5 indicates how far the furthest value is from the mean, but does not give any sense of whether there are values anywhere in between. The distribution in Figure 6 has a mean of 7 and a range of 5, but is not in any useful way similar to the real dice distribution.

Figure 6 Hypothetical dice distribution with a mean of 7 and a range of 5

  A more useful way to measure the *centrality* of a distribution is to compute an average of the deviations from the mean. By definition, a simple average of the deviations is zero — they cancel from each side of the mean since the mean is the balance point. The sign of the deviation from the mean can be ignored by squaring the value, and then taking the mean of the squared values, resulting in a measure called the *variance*[1].

- *Variance* is the weighted average of the squares of the deviations of a distribution about its mean. Because it is in units of the *square* of the units of the values upon which the distribution is based, it is sometimes difficult to relate the variance to the visible characteristics of the distribution. Thus, a more useful value is the *standard deviation*, the square root of the variance. Being the square root, the standard deviation is in the same units as the distribution.

---

1. This procedure is not as arbitrary as it may seem. In terms of mechanics, the mean measures the first moment of the distribution (where it balances), and the variance measures the second moment (the resistance to rotation).

The variance computation for the dice distribution is as follows:

$$\frac{1}{36}\times(-5)^2+\frac{2}{36}\times(-4)^2+\frac{3}{36}\times(-3)^2+\frac{4}{36}\times(-2)^2+\frac{5}{36}\times(-1)^2+\frac{6}{36}\times 0^2+\frac{5}{36}\times 1^2+\frac{4}{36}\times 2^2+\frac{3}{36}\times 3^2+\frac{2}{36}\times 4^2+\frac{1}{36}\times 5^2$$

$$= \frac{(25+32+27+16+5+0+5+16+27+32+25)}{36}$$

$$= 5.83$$

Details on the above variance computation can be found in Appendix B later in this document.

- *Standard deviation* is the root mean square of the deviations of a distribution about its mean.

  Typically denoted by the Greek letter σ, the value of the standard deviation can be written as σ_x if a reminder is needed to relate it is the standard deviation for the observations designated by x. Thus, the variance is written as σ^2 or σ_x^2.

  For the dice distribution, the standard deviation is $\sqrt{5.83} = 2.42$. For comparison, the variance and the standard deviation of the second, hypothetical distribution are 25 and 5, respectively. (The deviation of every value in the hypothetical distribution is exactly 5, and, for this reason, the standard deviation is also 5. However, the deviations in the real dice distribution are themselves distributed, hence the standard deviation gives some measure of this distribution.)

  Computing the standard deviation helps measure the centrality of the distribution. Since the standard deviation of the real dice distribution is 2.42 and that of the hypothetical distribution is 5, the first distribution is more central than the second.

The Binomial Distribution

Reconsider the distributions generated by the tossing of coins.

For a coin toss, assume the probability of getting a head is p, and the probability of getting a tail is q. The first thing to note is that for any coin, $q + p = 1$ — the assumption is the result can only be a head or tail. By definition, a *fair* coin is one where $q = p = 1/2$.

Consider four successive tosses of a fair coin. Using the repeated application of the first probability law, the probability of four heads is $p^4 = 1/16$. The probability of achieving the sequence head, head, tail, tail is, by similar reasoning, also 1/16.

An interesting question arises when one considers four tosses of a fair coin — what is the probability of three heads and one tail, in any order? The possible results of four throws consisting of three heads and one tail, are:

heads	heads	heads	tail
heads	heads	tails	heads
heads	tails	heads	heads
tails	heads	heads	heads

The probability of each combination is 1/16. The four combinations yield a total probability for three heads and one tail in four tosses, in any order, of 4/16 = 1/4 or 25 percent.

Suppose a coin is not fair, with a lump of lead attached to one face that results in the probability of a head being $p = 2/3$, and the probability of a tail being $q = 1/3$. In this case, the probability of three heads and one tail, in that order, is $(2/3)^3 \times (1/3) = 8/81$. Since there are four combinations with three heads and one tail, the total probability for three heads and one tail in four tosses of this unfair coin, in any order, is 4 x 8/81 = 32/81 or nearly 40 percent — making heads more likely makes the probably of three heads in four tosses more likely.

In general, it is necessary to be able to count the number of combinations that have so many of one outcome and the remainder of the other. Let these numbers be x of the outcome with probability p, and $n - x$ be the outcome with probability q. Just as it is required that $q + p = 1$, equally, and equivalently, $(n - x) + x = n$, where n is the total number of discrete events in the outcome.

Considering n discrete events (such as the number of tosses), the number of ways to get x of the p-outcome is called the number of combinations of x out of n. This is written as:

$$\binom{n}{x}$$

See Appendix C for more information.

In the case of the three heads and one tail outcome discussed above:

$$\binom{4}{3} = \frac{4!}{3! \times 1!} = \frac{24}{6 \times 1} = 4$$

Putting this together with the computation of probability above, it can be seen that the frequency of occurrence of the outcomes with x_i of events with probably p (and thus $n - x_i$ events with probability q, where $q + p = 1$) is:

$$\binom{n}{x_i} p^{x_i} q^{n - x_i}$$

Distributions that result from populations with these frequencies are termed *binomial distributions*, since the above formula corresponds to the successive terms in the expansion of the binomial expression $(p + q)^n$.

To illustrate, the outcomes of 10 throws of an unfair coin. What is the frequency distribution of 0, 1, 2, … 9, and 10 heads, if the 10-throw experiment is repeated a large number of times? Figure 7 illustrates the results of the computations.

0	1	2	3	4	5	6	7	8	9	10
1.69E-05	3.39E-04	3.05E-03	0.016	0.057	0.137	0.228	0.260	0.195	0.087	0.017

Figure 7 Results of unfair coin toss

These results are easier to understand when viewed as a distribution (Figure 8).

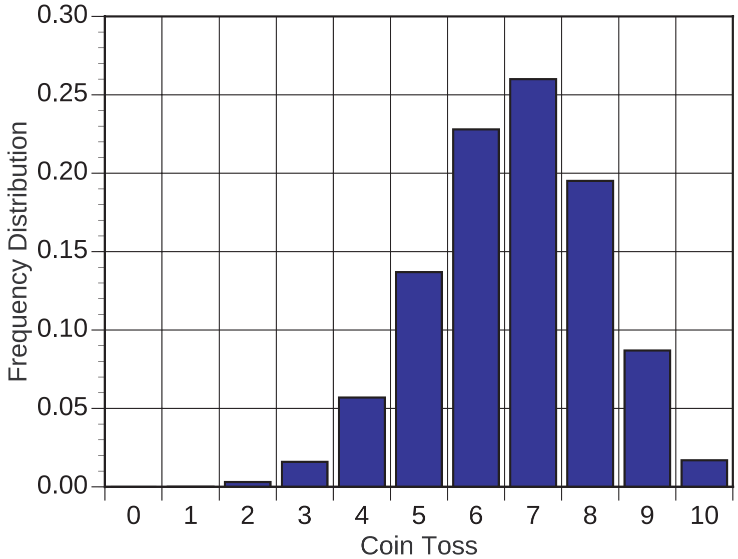

Figure 8 Results of unfair coin toss viewed as a distribution

The *mode* of this distribution is 7 — the outcome with 7 heads and 3 tails is seen most often, 26% of the time.

The probabilities from 0 to 6 account for approximately 44 percent of the total distribution, and those from 0 up to 7 total approximately 70 percent. Consequently, the *median* is a value greater than 6 and less than 7. Since this distribution is actually discrete — there are no values possible between the whole numbers — there is no exact value that can be assigned to the median. Observe that the 7-column represents 26 percent of the total probability, and that 0.62 of that is 16 percent. Combined with the 44 percent of the total probability represented by the total up to the 6-column, these results give a mathematical median of 6.62.

By direct computation, the *mean* value for this distribution is 6.67. While this distribution is discrete, the mean value can still be 6.67 as this is the point about which the distribution balances.

By direct computation, the variance of this distribution is 2.22, resulting in a standard deviation of 1.49. This indicates that the distribution is strongly centralized.

The fact that there is a way to mathematically describe the probabilities of the binomial distribution holds the promise for finding a formula for the mean and the variance (and hence the standard deviation). The mathematics needed to accomplish this is shown in Appendix D. Using such techniques, the following, and perhaps surprising, results are found:

$$\bar{x} = np$$
$$\sigma^2 = npq$$

From these formulæ, calculation shows that the mean is 10 x 2/3 = 6.67, and that the variance is 10 x 2/3 x 1/3 = 2.22, as formerly found by direct computation.

That the mean should be np is not surprising — it indicates that, on average, if an unfair coin is flipped, it has a chance of coming down on a given side with probability p. If this is repeated n times, np of those sides will be seen. It is by no means so obvious that the variance should be npq.

Statistical Models

A distribution such as the binomial distribution is called a *statistical model*. It is used by noting that certain processes or events in the real world fit the assumptions of the model. In so doing, it is possible to derive or infer additional characteristics of the real world situation by understanding the behavior of the model. Care must be taken to ensure that assumptions are met.

The assumptions that have been made for the binomial distributions include:

- The probability of the events that are the constituents of a single observation are independent, and that this probability, p, is constant.
- The events that are the constituents of a single observation either happen, or they do not happen.
- The number of events that are the constituents of a single observation is constant, and is the value n.

The values p and n are called the *parameters* of the binomial distribution, since when these are known, all the features of the distribution are known.

For example, the results of the statistical model can often be applied in reverse. Assume that CD-ROM discs are purchased in boxes of 500, and a shipment of 100 boxes is received. In use, manufacturing defects that render CD-ROMs unusable may be found. Assume the number of defective discs is recorded, with the following results:

Failures/Box	0	1	2	3	4	5	6	Total
Number of Boxes	45	36	13	4	1	1	0	100
Frequency	0.45	0.36	0.13	0.04	0.01	0.01	0	—

Table 3 Frequency distribution of CD-ROM defects

Note this experiment fulfills the requirements to be modeled as a binomial distribution.

By multiplying the failures per box by the frequencies and summing, a mean of 0.830 is obtained for this distribution, indicating there are, on average, 0.830 failures per box. What is the probability of failure for an individual disk? The mean is equal to np, where n is the number of CD-ROMs in a box. With 500 discs per box, an *estimate* of $p = 0.0166$, or approximately 1 in every 600 is obtained. The variance can also be estimated as npq, the value 0.828 or, equivalently, a standard deviation of 0.91.

Computing the binomial coefficients, based on the parameters $p = 0.0166$, and $n = 50$, the expected number of boxes with the given number of failures is predicted. It can be seen that the agreement between the observations and the predictions is reasonably good (Table 4).

Failures/Box	0	1	2	3	4	5	6	Total
Observed Boxes	45	36	13	4	1	1	0	100
Predicted Boxes	44	36	15	4	1	0	0	100

Table 4 CD-ROM defects, compared to binomial model

Another example illustrates how predicted and observed results may not match. Consider a stacked-platter disk drive with 1000 sectors per cylinder and 1000 cylinders. Assume the disk drive controller remaps failing tracks, those in which any sector on the track fails. Further assume the drive contains 40 tracks designated as targets for remapping, and that after two years all of these tracks are used. Table 5 describes the distribution of failures per cylinder found after testing. The mean is 0.147 and the predicted number of failures is shown, assuming a binomial distribution with $p = 0.147/1000$, and $n = 1000$.

Failures/Cylinder	0	1	2	3	4	5	6	7	8	Total
Observed Cylinders	919	45	19	10	4	1	1	1	1	100 0
Predicted Cylinders	863	127	9	0	0	0	0	0	0	100 0

Table 5 Disk drive failures, compared to binomial model

In this example, the predicted numbers do not match the observed numbers. It can be deduced that the assumptions underlying the binomial distribution are not operative here, most likely that the probability, p, of a cylinder failure is not the same for every cylinder. This is not unexpected, since the cylinders of a multiple platter disk occur at differing radii, with different recording densities, at different angles to the recording head, with differing linear speeds past the head sensors. In the previous CD case, the assumption of equal probabilities is more likely to hold as the production of each CD is independent, yet produced by an equivalent process.

The Poisson Distribution

Previous examples applied the binominal distribution to situations with a large number of items where the probabilities of the events of interest were small, such as a manufacturing process.

This situation is similar to those found when dealing with the behavior of systems. Consider the arrival of HTTP requests at a Web server that result from users clicking on links in browsers. Assume that 2400 URL requests, or *hits*, are observed over the course of an hour, for an average of 40 hits per minute. How would an estimate of how many of the minutes have no requests be found, how many would have 1 request, and 2 requests, and so on?

The first two assumptions that apply to the binomial distribution are likely to hold:

- The probability of the events that are the constituents of a single observation are independent, such as the arrivals in each minute, and that this probability, p, is constant (ignoring, for the moment, variations over the period of a day).
- Events that are the constituents of a single observation either happen, or they do not — an URL request is either made, or it is not.

However, in order to use the binomial statistical model, the total number of events must be known — those that occurred, and those that did not occur. Remember, n, the total number of events, is a parameter of the binomial distribution.

There is no way of knowing how may URL clicks users did *not* make on links in their browsers. But what can be safely assumed is that the number made is small compared to the number that could have been made. That is, the probability value, p, is small, and, by the same assumption, $q = 1 - p$ is very close to one.

In the situation discussed above, the average, or mean, is known — the number of hits per minute. In this case, the mean is 40 hits.

$$\bar{x} = np = 40$$

The variance, npq, of the binomial distribution can be written as $\bar{x}(1 - p)$. Since $1 - p$ is close to 1, this can be taken as being nearly equal to $\bar{x}$, in this case, also the value 40.

This *limiting* case of a binomial distribution, where the number of possible events is very large, and the probability amongst those large number of events is very, very small but the mean is known, is termed a *Poisson*[1] *distribution*. This mean value is designated λ, and is the (only) parameter of the distribution.

The following formula gives the probabilities of a Poisson distribution — the probability there will be exactly x_i events in a given interval, with parameter λ:

$$\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

1. This distribution is named after Siméon Denis Poisson, the French mathematician who investigated, and in 1837 published, many of the properties of this distribution.

Based on the discussion above, for the Poisson distribution:

$$\bar{x} = \lambda$$

$$\sigma^2 = \lambda$$

Based on the parameter of the example above, with $\lambda = 40$, the statistical model can be determined. Figure 9 illustrates the probabilities for values from 0 to 60. The mean of this distribution is 40, and appears close to being both the median and the mode value as well.

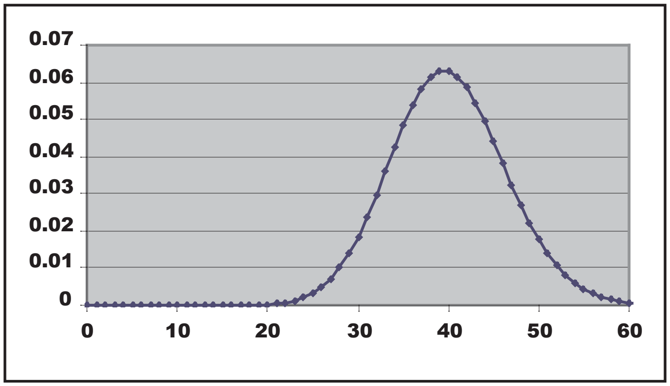

Figure 9 Probabilities for the example Poisson distribution

The graph and its underlying computations indicate that the probability of receiving exactly 40 hits per minute is 0.063, or 6.3 percent. While 40 is the mean value of this distribution, the probability of experiencing that exact value is small — a requirement for the Poisson statistical model.

Typically, the question of interest is not the probability of getting exactly the average number of events per interval, or even exactly the probability of getting some other number of events per interval. What is of more interest is looking at the graph and seeing that 99.9 percent of all hits occur at a rate of approximately 60 per minute or less. If 40 represents the mean over a period of time where the load averages remain constant, a Web site only needs to accommodate a peak load of 60 hits per minute to cope well. At this level, there is only 1 chance in 1000 (0.1 percent) that a URL request cannot be served immediately[1].

While missing only 0.1 percent of the requests is very good service, the purpose of the Web site determines the level of acceptability of these results. For some sites, such as those performing on-line stock trading, the chances of not serving a

URL request immediately must be less than 1 in 10000. In such an environment, the Web site would need an increase in handling capacity (estimated from the graph) to a service rate of 65 hits per minute.

Using the appropriate formula, the variance is also 40, resulting in a standard deviation for this distribution of 6.32. The typical rules of thumb are as follows:

- The *1 in 1000* level is the mean + 3σ, a value of 56.9 for this distribution
- The *1 in 10 000* level is the mean + 4σ, a value of 65.3 for this distribution

A look back at the CD failure scenario from the previous section may prove interesting. In that example, the average failure rate was 0.830. The numbers predicted by a Poisson model for the distribution with this same parameter, $\lambda = 0.830$, is shown in Table 6, together with the original data. Note that even in this case, the actual probability (which was found to be 0.0166) is small enough, and the number of instances (which was 500) are large enough, for there to be no practical difference between the binomial and the Poisson distribution.

Failures/Box	0	1	2	3	4	5	6	Total
Observed Boxes	45	36	13	4	1	1	0	100
Binomial Model	44	36	15	4	1	0	0	100
Poisson Model	44	36	15	4	1	0	0	100

Table 6 Binomial and Poisson model results for the CD failure example

Figure 10 illustrates the actual Poisson probabilities for the CD failure example. It is a general property of the Poisson distribution that if the parameter λ is less than 1, the graph is L-shaped, asymmetrical towards the left-hand side. If the parameter λ is greater than 1 (as in the example above), the graph is more rounded, and becomes more symmetric as the value of λ increases.

1. Another whte paper discusses the strategies that may be adopted if this maximal rate is exceeded. For the moment, assume that excessive requests are simply not answered.

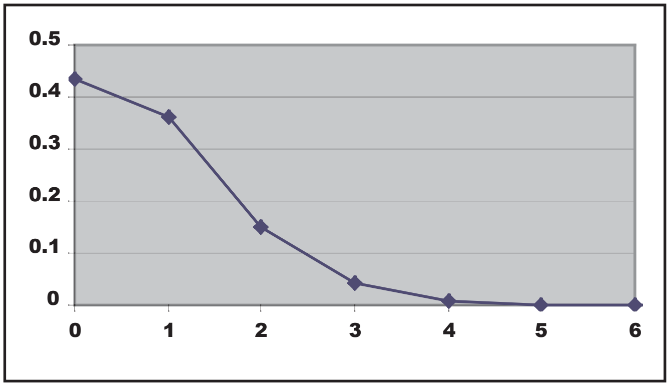

Figure 10 The Poisson probabilities for the CD failure example

Many similar arrival rate and occurrence per unit problems exist for which the Poisson statistical model is the appropriate way of describing and understanding the underlying process.

Summary

The white paper reviews the concepts of probability, and shows how this leads to the concept of a distribution. Distributions are the bases for the description of statistical models, which are ways of summarizing and making predictions about processes that are seen operating in the real world.

In another white paper on this topic, further statistical models are introduced, and the use of the models in more situations that occur in the architecting of a Web site, or similar e-business system, are investigated.

The Birthday Problem A

There is an age-old probability question known as *The Birthday Problem*. The question is deceptively simple — how many people must be in a room so that the probability of at least two of them having the same birthday is more than 50 percent?

Many conclude that the answer is 183, a value just greater than 50 percent of 365, the number of days in a year. While this answers the question of how many people are needed in a room so that the probability of one person having the same birthday as yourself (or some other designated person) is more than 50 percent, the original question really asked about at least *any* two people having the same birthday.

If a room contains 366 people, and if leap year considerations are ignored, a repeated birthday must exist in the room. That is, the probability of a repeated birthday is one. Because the question deals with probability, it should be thought about in that way. As with many such problems, it is actually easier to think about the probability of there *not* being two birthdays on the same day.

Consider people entering the room one at a time. When the first person enters the room, the probability of there already being a person in the room *not* having the same birthday is one — it is a certainty as there is no one else in the room. When the second person enters, the probability that there is not already a person in the room with the same birthday is the probability that the first person does not have the same birthday as the person entering — 364/365 — that the first person's birthday is different. For the third person, the probability that the people already do not have the same birthday is 362/365. The probability that the first and the second and the third person do not have the same birthday is, these probabilities multiplied together. As the number of people increases, the probabilities continue to multiply. By the time there are, say, 6 people, the probability that no two people in the room have the same birthday is:

$$1 \times \frac{364}{365} \times \frac{363}{365} \times \frac{362}{365} \times \frac{361}{365} \times \frac{360}{365} = 0.9595...$$

Of course, the probability that two birthdays are the same can be obtained by subtracting the result from 1. In fact, the graph in Figure A-1 shows the value of the probability that two birthdays are the same as the number of people in the room increases from 1 to 60. This data illustrates that after a slow start, the probability that two people have the same birthday increases rapidly — by the time 60 people are present, the chances are indistinguishable from 1, a certainty.

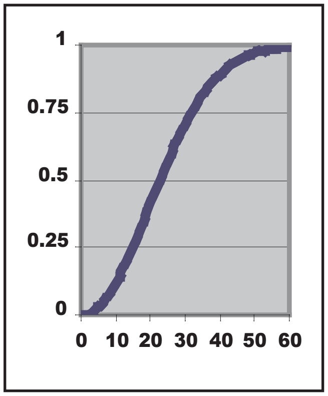

Figure A-1 The probability that two people in a room have the same birthday

The original problem asked how many people must be in the room for the probability to exceed 50 percent. The computations and graphical results show that at 22 people, the probability is just under 1/2, and at 23, it is just over. So, the answer is 23!

The Mathematics of Mean and Variance B

Designating the *variate* — the set of values whose frequency of occurrence that are being measured — by the symbol x, each different observed value can be designated as one of the set of values x_i, where i runs from 1 to k, and the corresponding relative frequency of each of the values x_i is f_i. Then:

$$\sum_{i=1}^{k} f_i = 1$$

The mean of the x_i values, written as $\bar{x}$, is defined by:

$$\bar{x} = \sum_{i=1}^{k} f_i x_i$$

The variance, equal to the standard deviation squared, where the standard deviation is written as σ_x, or more simply, when it is understood that it is the x-distribution that is being discussed, σ, is defined by:

$$\sigma^2 = \sum_{i=1}^{k} f_i (x_i - \bar{x})^2$$

This last formula can be simplified, through a sequence of steps. First, multiply the content of the parentheses, to obtain:

$$\sigma^2 = \sum_{i=1}^{k} f_i(x_i^2 - 2\bar{x}x_i + \bar{x}^2)$$

$$= \sum_{i=1}^{k} f_i x_i^2 - 2\bar{x}\sum_{i=1}^{k} f_i x_i + (\bar{x})^2 \sum_{i=1}^{k} f_i$$

Noting that $\sum_{i=1}^{k} f_i x_i = \bar{x}$, and that $\sum_{i=1}^{k} f_i = 1$, the equation above can be rewritten as:

$$\sigma^2 = \sum_{i=1}^{k} f_i x_i^2 - 2\bar{x}^2 + \bar{x}^2$$

$$= \sum_{i=1}^{k} f_i x_i^2 - \bar{x}^2$$

Using this formula, the variance of the dice distribution, σ^2_{dice}, may be recalculated as follows:

$$\frac{1}{36} \times 2^2 + \frac{2}{36} \times 3^2 + \frac{3}{36} \times 4^2 + \frac{4}{36} \times 5^2 + \frac{5}{36} \times 6^2 + \frac{6}{36} \times 7^2 + \frac{5}{36} \times 8^2 + \frac{4}{36} \times 9^2 + \frac{3}{36} \times 10^2 + \frac{2}{36} \times 11^2 + \frac{1}{36} \times 12^2 - 7^2$$

$$= \frac{(4 + 18 + 48 + 100 + 180 + 294 + 320 + 324 + 300 + 242 + 144)}{36} - 49$$

$$= 5.83$$

This calculation does indeed result in the same value.

Combinations and Combinatory Notation C

The number of ways that x things may be chosen from n is designated by the notation:

$$\binom{n}{x}$$

Mathematically, *choices* refer to groups of items where the order of the items in the group is of no interest. When order matters, these groups are called *arrangements*, or *permutations*.

The number of ways in which n things can be *arranged* is determined by considering the following. There are n selections from the available items for the first position in the arrangement. There are only $n - 1$ possible selections for the second position, and so on, until there is only one selection for the last position. This is an arrangement or permutation, and the computation is similar to the first law given for probabilities.

The number of different arrangements of five items is 5 x 4 x 3 x 2 x 1 = 120.

This value formed by multiplying an integer n by all the integers less than it down to 1, is designated mathematically by the notation $n!$, termed *n factorial*[1]).

If there is interest in only the *combinations* of just x of the n items, then first consider the arrangements of just x of the items. There are n choices for the first position in the arrangement, $n - 1$ for the second, and so on, stopping after x

1. By convention, 0! (zero factorial) is defined to have the value 1.

items are selected. The last number used in the multiplication is $n - x + 1$ — the numbers from $n - x$ down to 1 are not used. This is equivalent to $n!/(n-x)!$, in which all the numbers from $n - x$ to 1 are removed by division cancellation.

Secondly, there is no care for how the items are arranged in the combinations, and all arrangements of the x items chosen are equivalent. The number of arrangements must be reduced by dividing it by the number of arrangements of the items to be rearranged, which is $x!$. Thus,

$$\binom{n}{x} = \frac{\dfrac{n!}{(n-x)!}}{x!} = \frac{n!}{x!(n-x)!}$$

From this, it is noticed that:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \binom{n}{n-x}$$

The Mean and Variance of the Binomial Distribution D

As previously noted, the frequencies f_i of the values x_i of the binomial distribution, with parameters n and p, are the terms of the expansion of the binomial expression $(p + q)^n$, where $p + q = 1$. Keep in mind that, consequently, $(p + q)^n = 1$. This expansion can be represented as:

$$(p+q)^n = \sum_{i=1}^{n+1} \binom{n}{x_i} p^{x_i} q^{n-x_i}, \; x_i = i-1$$

Since the sum of the terms is the sum of the frequencies, the fact that this formula guarantees that the sum of the frequencies is one is reassuring.

The mean, $\bar{x}$, is the weighted sum of the frequencies:

$$\bar{x} = \sum_{i=1}^{n+1} \binom{n}{x_i} p^{x_i} q^{n-x_i} x_i, \; x_i = i-1$$

Now,

$$\binom{n}{x_i} x_i = \frac{n!}{x_i!(n-x_i)!} x_i = n \frac{(n-1)!}{(x_i-1)!(n-x_i)!} = n \binom{n-1}{x_i-1}$$

and

$$p^{x_i} = p \bullet p^{x_i - 1}.$$

Thus, the above summation can be rewritten as:

$$\bar{x} = np\sum_{i=1}^{n+1}\binom{n-1}{x_i-1}p^{x_i-1}q^{n-x_i}, x_i = i$$

$$= np(p+q)^{n-1}$$

$$= np$$

Similarly,

$$\sigma^2 = \sum_{i=1}^{n+1}\binom{n}{x_i}p^{x_i}q^{n-x_i}(x_i-\bar{x})^2, x_i = i-1$$

$$= \sum_{i=1}^{n+1}\binom{n}{x_i}p^{x_i}q^{n-x_i}(x_i^2-2x_i\bar{x}+\bar{x}^2), x_i = i-1$$

$$= np + n(n-1)p^2 - 2(np)(np) + (np)^2$$

$$= np(1-p)$$

$$= npq$$

The Probabilities of the Poisson Distribution

E

Finding the *probability density function* of the Poisson distribution algebraically requires revisiting the binomial distribution. Recall that the probabilities for the binomial distribution are the terms of the binomial expansion of $(q + p)^n$. An easier way to examine this is to introduce an additional term into the binomial expression, $(q + pt)^n$. The binomial expansion then yields:

$$\binom{n}{0}q^n t^0 + \binom{n}{1}q^{n-1}pt + \binom{n}{2}q^{n-2}p^2t^2 + \binom{n}{3}q^{n-3}p^3t^3 + \ldots + \binom{n}{n}q^0p^nt^n$$

The term, t^i, acts as a marker of the value that is the probability of the occurrence of the value x_i in the distribution. The expression $(q + pt)^n$ is called a *probability generating function* (*pgf*), and in this case, is the pgf for the binomial distribution.

Write $\lambda = np$, from whence $p = \lambda/n$. Hence, the expression for q can be rewritten as:

$$q = (1 - p) = \left(1 - \frac{\lambda}{n}\right).$$

The binomial *pgf*, with p and q rewritten, is:

$$\left(\left(1 - \frac{\lambda}{n}\right) + \left(\frac{\lambda}{n}\right)t\right)^n = \left(1 + \frac{\lambda(t-1)}{n}\right)^n$$

The Poisson distribution is that obtained when λ remains constant, as n gets larger and larger.

Algebra texts define the exponential function as the limit as *n tends to infinity*:

$$e^k = \lim_{n \to \infty} \left(1 + \frac{k}{n}\right)^n$$

By writing $k = \lambda(t - 1)$, in the limit, as *n tends to infinity*, the binomial *pgf* formula above becomes the function:

$$e^{\lambda(t-1)}$$

This is the *pgf* for the Poisson distribution. Using the standard series expansion for the exponential function:

$$e^{\lambda(t-1)} = e^{-\lambda} \bullet e^{\lambda t} = e^{-\lambda}\left(1 + \frac{\lambda t}{1!} + \frac{\lambda^2 t^2}{2!} + \frac{\lambda^3 t^3}{3!} + \frac{\lambda^4 t^4}{4!} + \ldots + \frac{\lambda^i t^i}{i!} + \ldots\right)$$

This *pgf* reveals that the probability of exactly i events in a given interval, is, by picking the term with t^i in it, equal to:

$$\frac{e^{-\lambda}\lambda^i}{i!}$$

Similarly, the binomial variance, npq, is rewritten as $\lambda(1 - \lambda/n)$. The limit of this, as *n tends to infinity*, is λ — the same value as the mean. As a result:

$$\sigma^2 = \lambda.$$

References

Sun Microsystems, Inc. posts product information in the form of data sheets, specifications, and white papers on its Internet World Wide Web Home page at: *http://www.sun.com*.

Look for abstracts on these and other Sun technology white papers:

Capacity Planning on a Cocktail Napking, Sun Microsystems, 2001.

Implementing Enterprise Security Policies, Sun Microsystems, 2001.

Introduction to the Elements of Web Application Architecture, Sun Microsystems, 2001.

Understanding Capacity and Scalability, Sun Microsystems, 2001.

Understanding Growth, Sun Microsystems, 2001.

What is Architecture?, Sun Microsystems, 2001.

Other references of interest:

Teach Yourself Statistics, Richard Goodman, The English Universities Press Ltd., London, 1957.

The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling, Raj Jain, John Wiley and Sons, New York, 1991, ISBN 0-471-50336-3.

Mathematics for the Million, Lancelot Hogben, George Allen and Unwin Ltd., London (W. W. Norton, USA), 1936-1993. ISBN: 039331071X.

Web sites of interest:

http://www.sun.com

Sales Offices

Africa (North, West and Central): +9714-3366333
Argentina: +5411-4317-5600
Australia: +61-2-9844-5000
Austria: +43-1-60563-0
Belgium: +32-2-704-8000
Brazil: +55-11-5187-2100
Canada: +905-477-6745
Chile: +56-2-3724500
Colombia: +571-629-2323
Commonwealth of Independent States: +7-502-935-8411
Czech Republic: +420-2-3300-9311
Denmark: +45 4556 5000
Egypt +202-570-9442
Estonia: +372-6-308-900
Finland: +358-9-525-561
France: +33-01-30-67-50-00
Germany: +49-89-46008-0
Greece: +30-1-618-8111
Hungary: +36-1-202-4415
Iceland: +354-563-3010
India: +91-80-5599595
Ireland: +353-1-8055-666
Israel: +972-9-9710500
Italy: +39-039-60551
Japan: +81-3-5717-5000
Kazakhstan: +7-3272-466774
Korea: +822-2193-5114
Latvia: +371-750-3700
Lithuania: +370-729-8468
Luxembourg: +352-49 11 33 1
Malaysia: +603-264-9988
Mexico: +52-5-258-6100
The Netherlands: +00-31-33-45-15-000
New Zealand: +64-4-499-2344
Norway: +47 23 36 96 00
People's Republic of China:
    Beijing: +86-10-6803-5588
    Chengdu: +86-28-619-9333
    Guangzhou: +86-20-8755-5900
    Hong Kong: +852-2202-6688
    Shanghai: +86-21-6466-1228
Poland: +48-22-8747800
Portugal: +351-21-4134000
Russia: +7-502-935-8411
Singapore: +65-438-1888
Slovak Republic: +421-7-4342 94 85
South Africa: +2711-805-4305
Spain: +34-91-596-9900
Sweden: +46-8-631-10-00
Switzerland:
    German: 41-1-908-90-00
    French: 41-22-999-0444
Taiwan: +886-2-2514-0567
Thailand: +662-636-1555
Turkey: +90-212-335-22-00
United Arab Emirates: +9714-3366333
United Kingdom: +44-1-276-20444
United States: +1-800-555-9SUN OR +1-650-960-1300
Venezuela: +58-2-905-3800
Worldwide Headquarters:

    Sun Microsystems, Inc.
    901 San Antonio Road
    Palo Alto, CA 94303-4900 USA
    Phone: 650-960-1300 or 800-555-9SUN
    Internet: www.sun.com

LFC1.2 Printed in USA 6/99 FU010-0