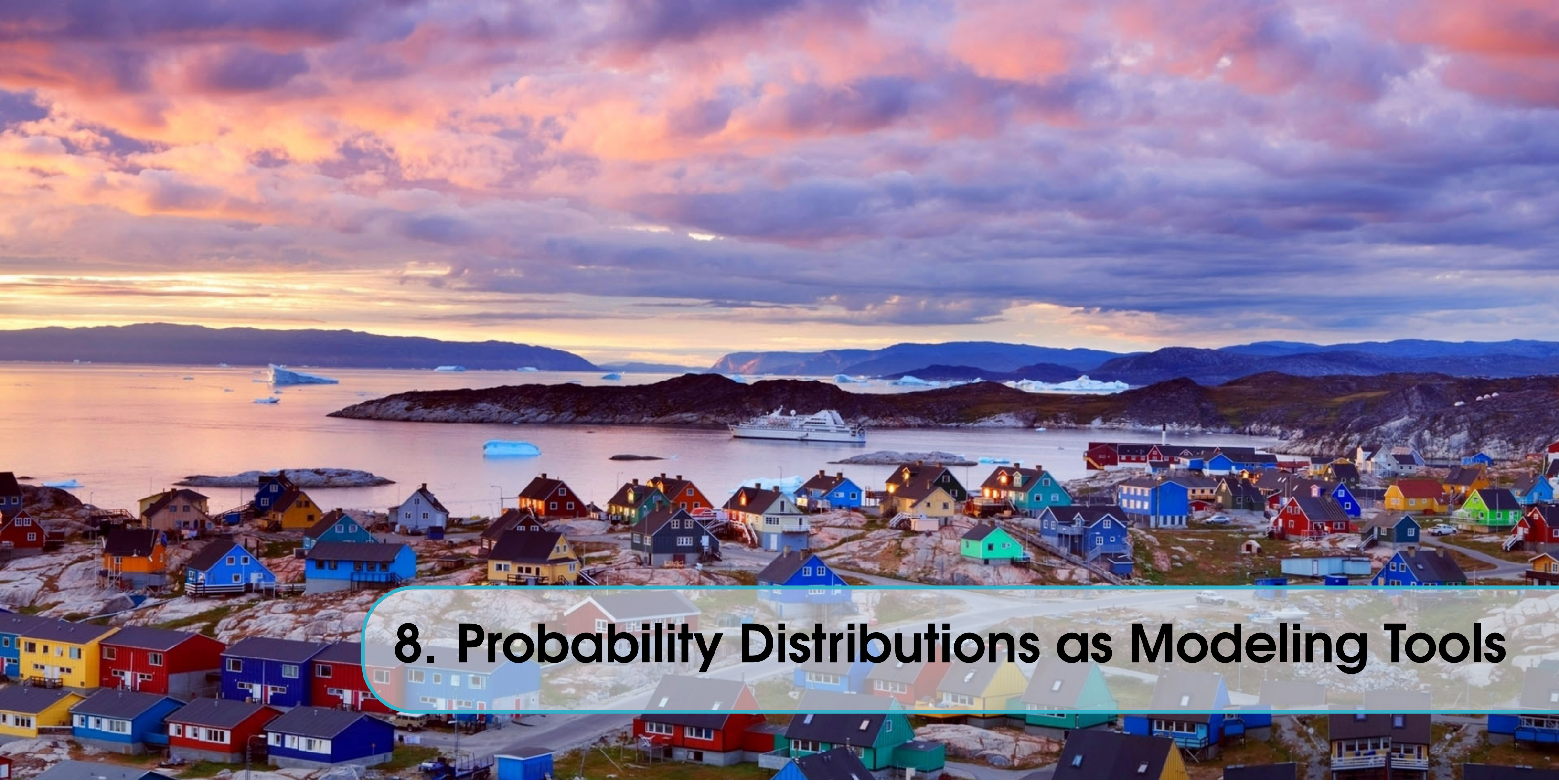

8. Probability Distributions as Modeling Tools

Recall that a probability distribution provides a characterization of a random variable – by enabling us to calculate the probability of a random variable assuming a particular value or range of values. Probability distributions are thus good "modeling" tools that we use to abstract away much of the details underlying the process responsible for instantiating the value of the random variable. For example, rather than modeling the various processes that determine whether an access to memory will hit or miss in the cache, we may simply abstract out such processes with a simple probability distribution (e.g., probability than an access would be served from level L_i of a cache hierarchy). We are now ready to learn about a number of probability distributions that we encounter frequently when modeling computing systems or when evaluating the performance of computing systems.

8.1 Important Probability Distributions

Let us now explore one by one some of the fundamental probability distribution that we will use throughout this course.

8.1.1 Uniform Distribution

A continuous random variable x is said to follow a uniform distribution if the random variable has an equal probability of assuming any value in a range of values between a lower bound (say a) and an upper bound (say b). Thus, for a uniformly distributed random variable x, the density function $f(x)$ is given by:

$$f(x) = \begin{cases} c & \text{if } x \geq a \text{ and } x \leq b \\ 0 & \text{if } x < a \text{ or } x > b \end{cases} \tag{8.1}$$

In order to find the value of c in the above expression, recall that for $f(x)$ to be a valid density function, it must satisfy the condition that:

$$\int_{x=-\infty}^{x=+\infty} f(x) \cdot dx = 1 \tag{8.2}$$

From here we have:

$$\int_{x=-\infty}^{x=+\infty} f(x) \cdot dx = \int_{x=a}^{x=b} c \cdot dx = c \cdot b - c \cdot a = c(b-a) \tag{8.3}$$

And since we want $c(b-a) = 1$, it follows that $c = \frac{1}{b-a}$. It is also easy to show that for a uniformly distributed random variable x between a and b, the expected value (i.e. the mean) is given by:

$$\mu = \frac{b+a}{2} \tag{8.4}$$

and that the variance is given by

$$\sigma^2 = \frac{(b-a)^2}{12}. \tag{8.5}$$

8.1.2 Bernoulli Random Variable

Consider an experiment that has two outcomes (e.g. heads or tail, success or failure, packet lost or packet received, true or false, 0 or 1, etc.) Assume that one of these outcomes is deemed a success (e.g. packet received) and the other is deemed a failure. The outcome of such an experiment is clearly a random variable – call it x. Let $x = 0$ denote failure and $x = 1$ denote success. Let p denote the probability of failure. The random variable x is called a Bernoulli random variable, with the simple probability mass function (PMF):

$$f(x) = \begin{cases} p & \text{if } x = 0 \\ 1-p & \text{if } x = 1 \end{cases} \tag{8.6}$$

8.1.3 Geometric Distribution

Consider a "Bernoulli trial" experiment that has two outcomes (e.g. heads or tail, success or failure, packet lost or packet received, true or false, 0 or 1, etc.). Assume that we conduct these Bernoulli trials a certain number times and that each experiment is independent. Let n denote the number of times that a particular result (specifically "failure" of the Bernoulli trial) occurs in a sequence. We say that n is a discrete random variable which follows a geometric distribution. The probability mass function of n is given by:

$$f(n) = p^n \cdot (1-p) \tag{8.7}$$

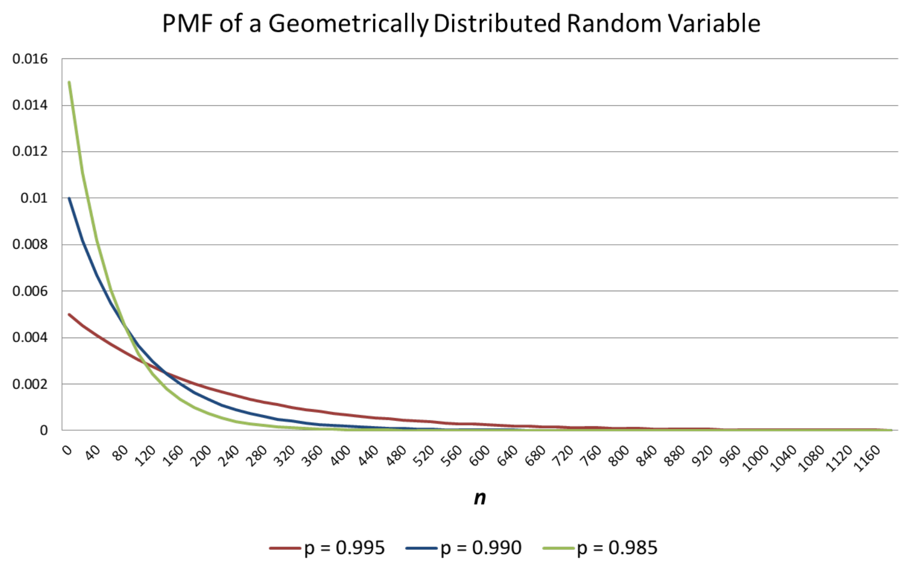

Figure 8.1: PMF of a geometrically distributed random variable n for various values of the parameter p.

where $n \in \mathbb{N}$, $n \geq 0$, and $0 \leq p \leq 1$.

Figure 8.1 illustrates the PMF of the geometrically distributed random variable n for various values of the parameter p, which is the probability that a single Bernoulli trial failed.

To calculate the mean (or expected value) of the geometrically-distributed random variable n, all we have to do is evaluate the summation:

$$\mu = \sum_{n=0}^{infty} n \cdot p^n \cdot (1-p) \tag{8.8}$$

It is easy to show that the summation above converges to the value

$$\mu = \frac{p}{1-p} \tag{8.9}$$

Consider now the random variable m describing the total number of trials until (and including) the first success. Since we already modeled the probability of a having n failures, then we can safely say that $m = n+1$. Hence, the probability of having to do a total of m trials is equal to the probability of having $n-1$ failures:

$$f(m) = f(n-1) = p^{n-1} \cdot (1-p) \tag{8.10}$$

From here it follows that the mean of m is:

$$\mu = \frac{1}{1-p} \tag{8.11}$$

8.1.4 Binomial Distribution

Consider a "Bernoulli trial" experiment that has two outcomes (e.g. heads or tails, success or failure, packet lost or packet received, true or false, 0 or 1, etc.). Assume that we conduct these Bernoulli trials n times and that each experiment is independent. Let x denote the number of times that a particular result occurs (and thus $n-x$ is the number of times the "other" result occurs). We say that x is a discrete random variable which follows a Binomial distribution. The PMF of x can be derived to be:

$$f(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \tag{8.12}$$

where $n \in \mathbb{N}$, $n \geq 0$, and $0 \leq p \leq 1$. And $\binom{n}{x}$ is defined as follows:

$$\binom{n}{x} = \frac{n!}{(n-x)! \cdot x!} \tag{8.13}$$

One can show that the mean of the Binomial distribution is $\mu = n \cdot p$ and that the variance is $\sigma^2 = n \cdot p \cdot (1-p)$.

Example: A web site consists of $n = 3$ servers. The probability that a server is "up" is 0.7 (i.e., it has an availability of 70%). What is the probability that i servers will be up? Obviously, the number of servers that are up is a random variable which is binomially distributed with $p = 0.7$ and $n = 3$. Substituting values into the above formula, we get the probability that 0, 1, 2, and 3 servers are up. This leads to the PMF shown in Figure 8.2.

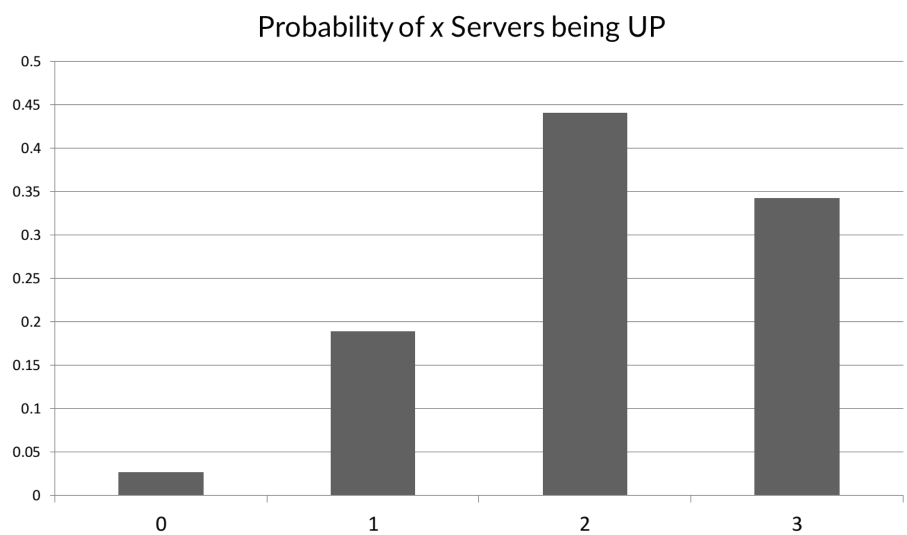

Figure 8.2: PMF modeling the probabilities of 0, 1, 2, or 3 servers being up in a set of 3 servers with a 70% availability for each.

Figure 8.3 shows the PMF of the Binomial distribution for $n = 10$ and for various values of the Bernoulli trial parameter p.

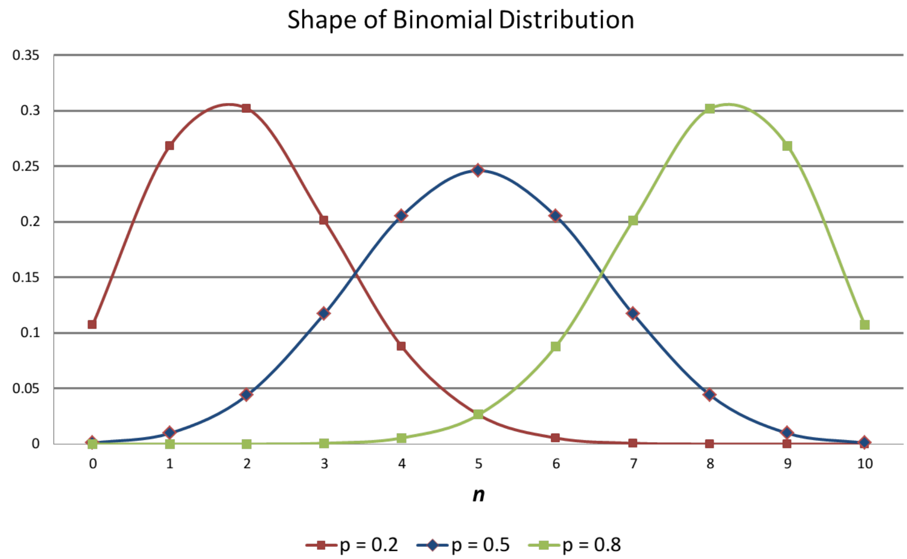

Figure 8.3: Examples of binomial distribution PMF functions for various values of p, when $n = 10$.

Notice that the PMF for the Binomial distribution is "symmetric" when p is 0.5. It is "skewed" to the right when $p > 0.5$ and skewed to the left when $p < 0.5$.

8.1.5 Poisson Distribution

Consider a binomial process with large n (number of events or experiments) and a small p (probability of "success"). Let n approach infinity and p approach 0, with the product $n \cdot p$ approaching a constant $\lambda > 0$. We call such a process a "Poisson" process and the resulting PMF a "Poisson Distribution." One can show that the Poisson distribution with parameter $\lambda, \lambda > 0$ is given by the following probability mass function:

$$f(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda} \tag{8.14}$$

where $x = 0, 1, 2, \ldots$

The Poisson distribution is useful to model random variables that capture the number of events that occur in a certain time interval. Under this interpretation, the parameter λ represents the average number of event in the considered time interval. The PMF of a Poisson distribution with $\lambda = 4$ is depicted in Figure 8.4

It is important to note that the *number* of observed events is inherently a discrete variable. Figure 8.5 shows the shape of the Poisson distribution as the value of λ increases (notice that the mean of the distribution is λ). The distribution exhibits a larger positive skew for smaller values of λ.

One can show that the Poisson distribution has a mean equal to the variance, equal to λ. Notice that we can obtain this result also by substituting $\lambda = n \cdot p$ and $p = 0$ in the formulae for the mean and variance of the binomial distribution.

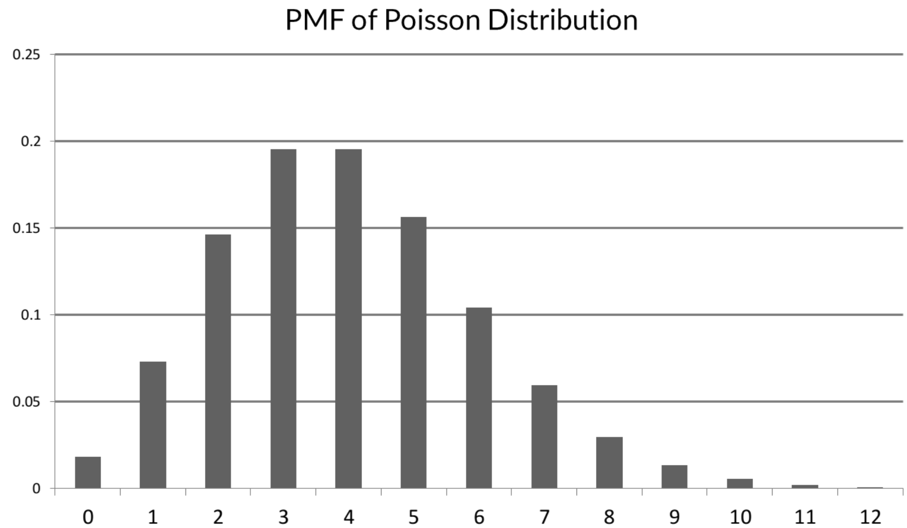

Figure 8.4: PMF of a Poisson Distribution with $\lambda = 4$.

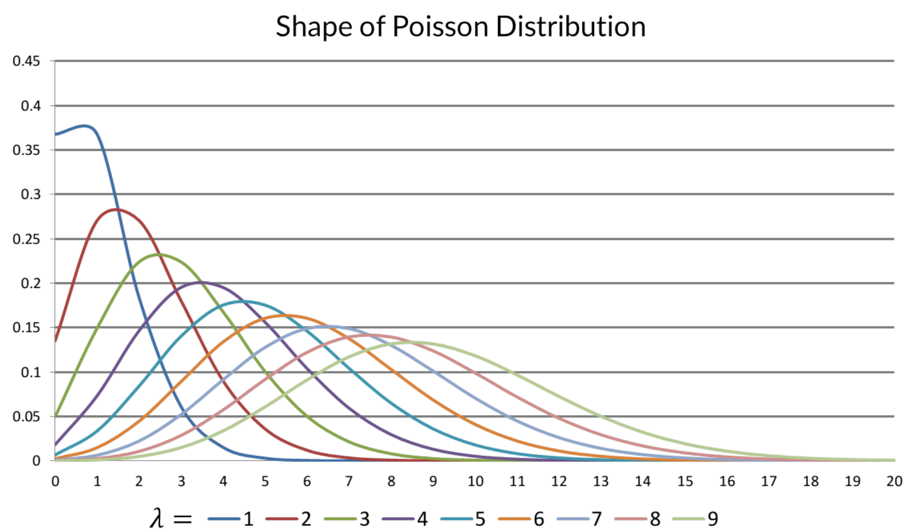

Figure 8.5: Shape of the Poisson distribution as value of λ increases – notice that the mean of the distribution is λ.

8.1.6 Exponential Distribution

The exponential distribution with parameter $\lambda\,(\lambda > 0)$ has a probability density function (PDF) given by:

$$f(x) = \lambda \cdot e^{-\lambda x} \tag{8.15}$$

One can show that the cumulative probability distribution function (CDF) for the exponential distribution is given by:

$$F(x) = 1 - e^{-\lambda x} \tag{8.16}$$

Figure 8.6 and 8.7 respectively depict the PDF and CDF functions of the Exponential distribution with parameter $\lambda = 1, 0.75, 0.5, 0.25$.

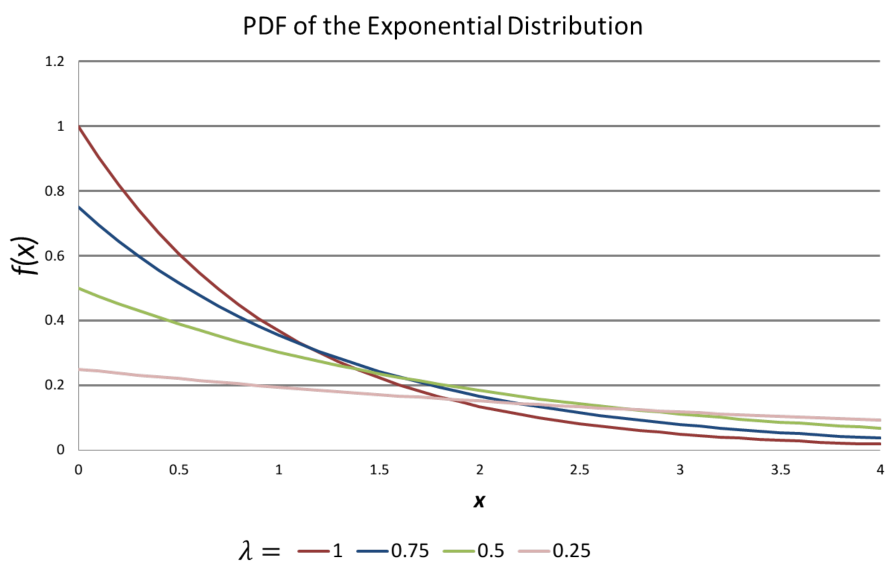

Figure 8.6: PDF of an Exponential Distribution with various values of λ.

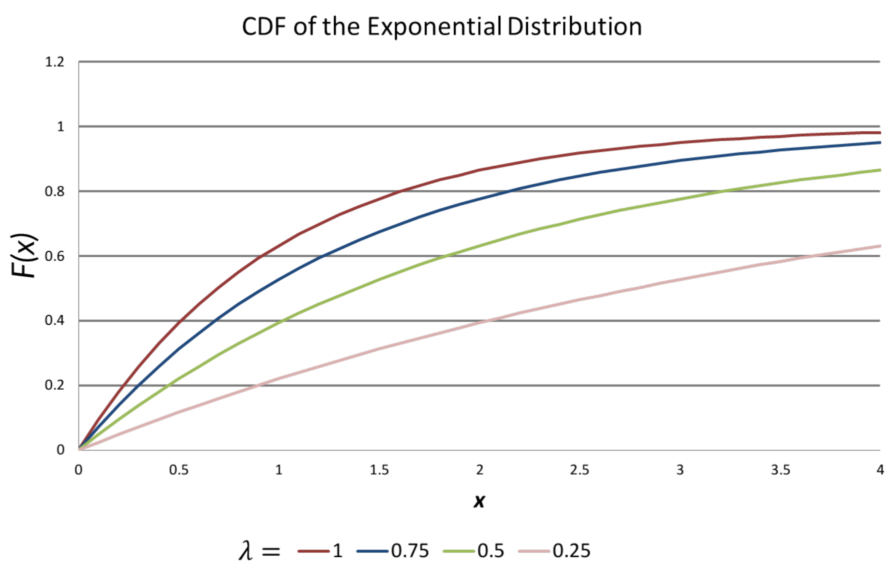

Figure 8.7: CDF of an Exponential Distribution with various values of λ.

An important feature of the exponential distribution is that both its mean and its standard deviation are equal to $1/\lambda$ One can easily show this by using the expressions for first and second moments presented earlier and substituting for the density function of the exponential distribution.

In computing systems, exponential distributions are used extensively to model random "times" (e.g. delay, inter-arrival time of events, etc.). The following are some examples of random variables that could be assumed to be exponentially distributed.

1. Time between subsequent hits to a web server;
2. Delay across a network;
3. Length of time it takes to transmit a file;
4. Time it takes a disk to write a block;
5. User thinking time for an interactive program.

Looking back at how we defined reliability in Section 4.4.1, now it makes more sense why $R(t) = e^{-t/MTBF}$. Indeed, we can define the random variable F that captures the time until the next failure when observing a system at a random point in time. F follows an exponential distribution with parameter 1/MTBF. So we can write $R(t) = P(F > t) = 1 - P(F \leq t) = 1 - F(t) = e^{-t/MTBF}$.

8.1.7 Memoryless Property

Another important characteristic of the exponential distribution is that it is "memoryless." In particular, if we use the exponential distribution to characterize "how long we have to wait for an event to happen," then the probability of waiting for an "additional time t" is independent of how long we have been waiting already.

To explain this, we use an analogy. Assume that the random variable x, which follows an exponential distribution, characterizes the length of time we have to wait for (say) the arrival of the next phone call. In this case, the probability that we have to wait "for another 5 minutes" is the same whether we have already been waiting for 1 minute, 10 minutes, 1 hour, or even a whole day! In other words, it doesn't matter what has happened in the past—and hence the term "memoryless."

In general, a distribution is considered to be memoryless if the following relationship holds:

$$1 - F(x+t) = (1 - F(t)) * (1 - F(x)) \tag{8.17}$$

The above relationship states that the probability that the random variable will be larger than $x+t$ is the same as the random variable being (independently) larger than t and larger than x. Another way of stating this property is as follows:

$$P(X > s+t | X > t) = P(X > s) \tag{8.18}$$

where $s > 0$ and $t > 0$.

Thinking of the random variable X as modeling inter-arrival time (or delay), the above relationship says that having already waited for a period t does not change the probability of waiting for an additional period s. In other words, passage of time does not make a sooner occurrence of the "next event" more likely.

8.1.8 The Exponential Distribution as an approximation of the Geometric Distribution

Consider an exponentially distributed random variable t with parameter λ and consider a very small time interval τ. The probability of an arrival (i.e. success) during an interval τ can be approximated by $\lambda \cdot \tau$. Now think about the complement of this probability (i.e. $1 - \lambda \cdot \tau$) as a Bernoulli trial parameter p. The number of Bernoulli trials necessary to perform until we finally get an arrival (or success) is nothing but a geometrically distributed random variable n. The length of time one has to wait for this arrival (or success) is therefore $n \cdot \tau$.

Consider the above scenario as the value of τ reaches 0. Clearly $(1 - \lambda \cdot \tau$, which is also the Bernoulli trial parameter, will get very close to 1. One can show that under such conditions, the geometric distribution reduces to an exponential distribution.

Thus, you can always think of the exponential distribution as the continuous counterpart of the geometric distribution where n (the number of Bernoulli trials necessary to get a success) becomes the time t necessary to observe an arrival (i.e. success).

Not surprisingly, both the exponential and geometric distributions are memoryless.

8.1.9 Relationship between the Poisson and Exponential Distributions

Consider a system in which events (e.g. requests for I/O, packets broadcast on an Ethernet, etc.) arrive in such a way that the number of event arrivals per unit time (i.e. rate of event arrivals) is a random variable that follows the Poisson distribution. Note that an event denotes an instantaneous atomic activity. Thus in any period, one can only observe the arrival of 0, 1, 2, ... events. Therefore, the number of event arrivals over a period of time is a discrete random variable (we cannot have "partial" events).

Let the average rate of event arrival be λ thus, the expected number of event arrivals over a period t is $\lambda \cdot t$ and the number of event arrivals follows a Poisson distribution with a probability mass function:

$$f(x) = \frac{(\lambda \cdot t)^x}{x!} \cdot e^{-\lambda \cdot t} \tag{8.19}$$

where $x = 0, 1, 2, \ldots$

Now consider the time separating the arrival of two events. We call this time the event interarrival time. One can show that the interarrival time of events that arrive according to a Poisson distribution is a random variable that follows the exponential distribution with parameter λ.

The process that results in arrivals that follow the Poisson distribution (or exponential interarrival times) is termed a *Poisson Process*. Poisson processes provide a very good approximation to the pattern of event arrivals in many "real-world" systems. Examples include the number of customers walking into a bank, the number of calls initiated at a phone switch, etc.

8.1.10 Independent Arrival

There is a good reason why Poisson Processes are important. To explain this point, we fall back to the analogy of phone calls. Basically, it suffices to note that if the arrival of one phone call is **independent** of the arrival of any other phone call, then knowing whether or not we have received a phone call in any past period of time should make absolutely no difference in the arrival of the other calls (hence the memoryless property of the exponential distribution).

The notion of "independent arrivals" is very important and is a common assumption when we do performance analysis. However, is it an accurate one? The answer is: *"most often it is not a correct assumption."* To understand this point, it suffices to note that in assuming that the arrival of (say) phone calls placed by different people should be fairly "independent," we neglect some factors (such as "people typically call in decent hours," "major events may trigger many phone calls," etc.).

Despite the "idealistic" (memoryless) nature of Poisson arrivals (exponential interarrivals), the use of these distributions in characterizing the number of arrivals per unit time (or the passage of time, e.g. interarrival times, delays, length of transmissions, etc.) is very common in computing systems, telephony, industrial engineering, business, etc. As we will see in the next few lectures, this is due to the nice mathematical properties of this distribution, which leads to closed-form solutions to many important performance evaluation problems.

8.1.11 Normal Distribution

A continuous random variable is said to follow the Normal Distribution if its probability density function is given by:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} e^{-0.5(\frac{x-\mu}{\sigma})^2} \tag{8.20}$$

A standard normal distribution has a mean of 0 and a variance of 1 and is denoted by $N(0,1)$. A random variable that follows the standard normal distribution is denoted by z. One can "standardize" a normally-distributed random variable (i.e. transform it into a random variable that follows the standard normal distribution) by applying the transformation:

$$z = \frac{x-\mu}{\sigma} \tag{8.21}$$

The cumulative distribution function $F(x)$ for the normal distribution does not have a "nice" closed-form that one can use readily in computing probabilities. Instead one must rely on "look-up tables" to find out the value of $F(x)$ for a given x. Since $F(x)$ will depend on the parameters of the normal distribution (namely its mean and its standard deviation), it is only practical to compute tables for the "standard" normal distribution. One must use the above transformation (from x to z) before using these tables. See the appendix below for values of $F(z)$.

8.1.12 The Normal Distribution as an approximation of the Binomial Distribution

Just as we established that the geometric distribution "at the limit" reduces to the exponential distribution, we can also show that a Binomial distribution "at the limit" reduces to the normal distribution. Thus, you can always think of the normal distribution as the continuous counterpart of the Binomial distribution where n (the number of Bernoulli trials) becomes large.

What depicted in Figure 8.8 is an example to illustrate how well a normal distribution could be fitted to a binomial distribution. Specifically, if $p = 0.3$, one can see that with a large enough n (e.g., $n = 20$, shown on the right), the normal distribution fits well the binomial distribution, whereas for small values of n (e.g., $n = 2$, shown on the left), it does not.

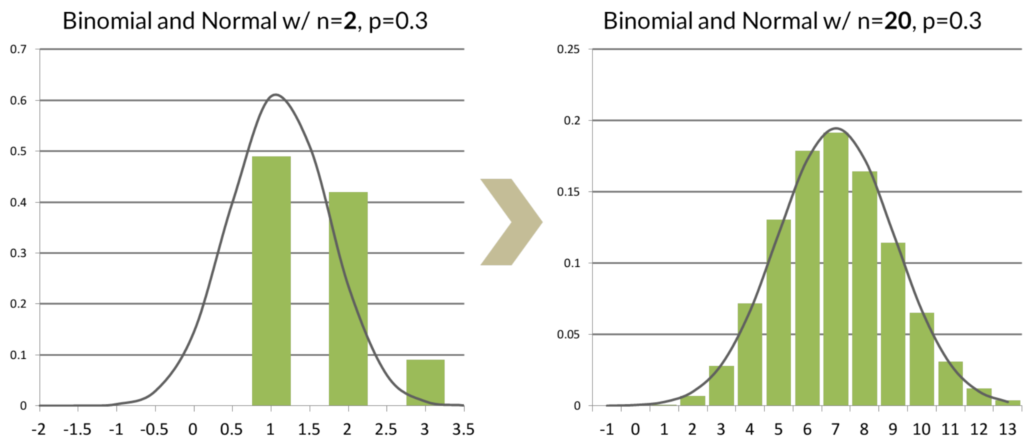

Figure 8.8: Illustration of how the normal distribution approximates the bionomial distribution as n gets large.

For large values of n, it is often handier to use the normal distribution as an approximation for the probabilities returned by the binomial distribution. In particular, given a large enough n and a value for p, then one can approximate a binomial $f(x)$ by a normal distribution with mean $\mu = n \cdot p$ and a standard deviation of $\sigma = \sqrt{n \cdot p \cdot (1-p)}$.

Example: A total of 1,000 packets were transmitted over a noisy wireless link. Assuming that the loss probability is 0.2, we can compute the probability that less than 100 packets will be lost by looking up the probability that a normally-distributed random variable with mean $1,000 \cdot 0.2 = 200$ and standard deviation of $\sqrt{1,000 \cdot 0.2 \cdot 0.8} = 12.65$ will have a value less than 100.

8.2 Central Limit Theory

Let $x_1, x_2, x_3, \ldots, x_n$ be a set of random variables (continuous or discrete) with identical probability distributions, each with a finite mean μ and a finite standard deviation σ. Let the value $X_n = x_1 + x_2 + \ldots + x_n$ be the sum of these random variables. The central limit theorem states that as the value of n approaches infinity, the distribution of X_n approaches a Normal distribution with mean $n \cdot \mu$and variance $n \cdot \sigma^2$. Another way of stating the Central Limit Theorem is that the random expression S_n follows a *standard normal distribution* as n grows large (typically $n > 30$ is considered large). S_n is defined as follows:

$$S_n = \frac{X_n - n \cdot \mu}{\sigma \sqrt{n}} = \frac{(\sum_{i=1}^{n} x_i) - n \cdot \mu}{\sigma \sqrt{n}} \tag{8.22}$$

The Central Limit Theorem is interesting because it allows us to characterize the sum of random variables **even when we do not know the original distributions** that govern these random variables – we only need to know the mean and standard deviation.

To illustrate how one can apply the central limit theory, we consider the following example. Assume that the number of bytes transferred per unit time (say per second) from a computer on an Ethernet bus is a random variable, the distribution of which is unknown. However, we know (possibly through measurement) that the mean of that distribution is 100 KB and that the standard deviation is 30 KB. If we assume that there are 16 machines connected to the Ethernet, we can use the Central Limit Theory to characterize the distribution that governs the "total" traffic on the Ethernet per unit time. In particular, the total traffic could be approximated by a normal distribution with mean $16 \cdot 100\ KB = 1.6\ MB$ and a standard deviation of $\sqrt{16} \cdot 30\ KB = 120\ KB$.

As we will see in future lectures, the Central Limit Theorem provides the basis for establishing "confidence" in making estimates about a distribution from a limited number of observations.

Example: We have a set of n files, each of which has an integer number of blocks ranging from 1 to k. Assuming that the probability that a file will have a specific number of blocks (between 1 and k) is uniform, what is the PMF for the total number of blocks in the n files?

Assuming we regard the individual files as independent, there are obviously k^n possible sets of file sizes. If n and k are small, then the exact answer for the PMF can be easily computed combinatorially for each possible total number of blocks. However, for large values of n and k, such as $n = k = 256$, we are better off applying the central limit theorem and dealing with proportions rather than absolute numbers.

Taking $n = k = 256$ as an example, we have 256 independent random variables, each of which has a uniform distribution with a mean of $256/2 = 128$ and a standard deviation of $256/(2 \cdot \sqrt{3})$. Our "outcome population" is the sum of these individual populations, so it has a mean of $256 \cdot (256/2) = 32768$ and a standard deviation given by $\sigma = \sqrt{256} \cdot 256/(2 \cdot \sqrt{3})$. Therefore, we have $\sigma = 1182.4$. By the central limit theorem, it follows that the density of sequences for a given sum x is (see Equation 8.20):

$$f(x) = \frac{1}{1182.4 \cdot \sqrt{2\pi}} \cdot e^{-0.5(\frac{x-32768}{1182.4})^2} \tag{8.23}$$

8.3 Appendix: Values of the Standard Normal Distribution for $z > 0$

Table 8.2 lists the value of the probability distribution function for the standard normal distribution. Recall that the distribution function is the probability that the random variable is less than a specified value. Thus, the table below gives the value of the "area under the curve" for the standard normal density function illustrated below – namely:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-0.5z^2} \tag{8.24}$$

Thus the table below gives the value of:

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-0.5y^2} \cdot dy \tag{8.25}$$

Notice that for $z < 0$ one can use the symmetry of the normal distribution to calculate $F(z)$. Specifically, $F(z) = 1 - F(-z)$.

Table 8.1: Normal cumulative distribution function for $\mu = 0$ and $\sigma = 1$

NORMAL CUMULATIVE DISTRIBUTION FUNCTION										
x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 8.2: Normal cumulative distribution function for $\mu = 0$ and $\sigma^2 = 1$