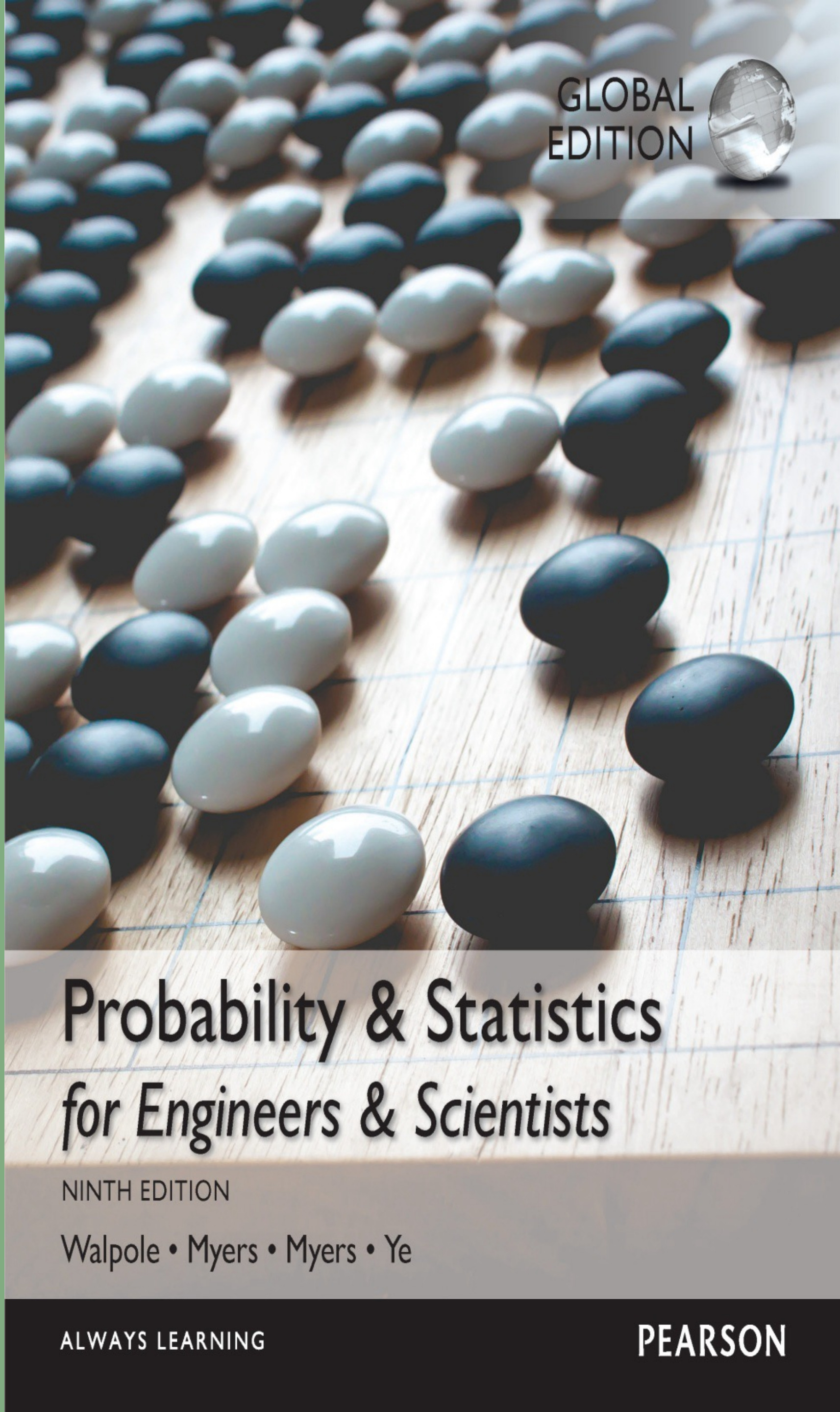
Section 6.5

Normal Approximation to the Binomial



Theorem 6.3

If X is a binomial random variable with mean $\mu = np$ and variance $\sigma^2 = npq$, then the limiting form of the distribution of

$$Z = \frac{X - np}{\sqrt{npq}},$$

as $n \to \infty$, is the standard normal distribution $n(z; 0, 1)$.

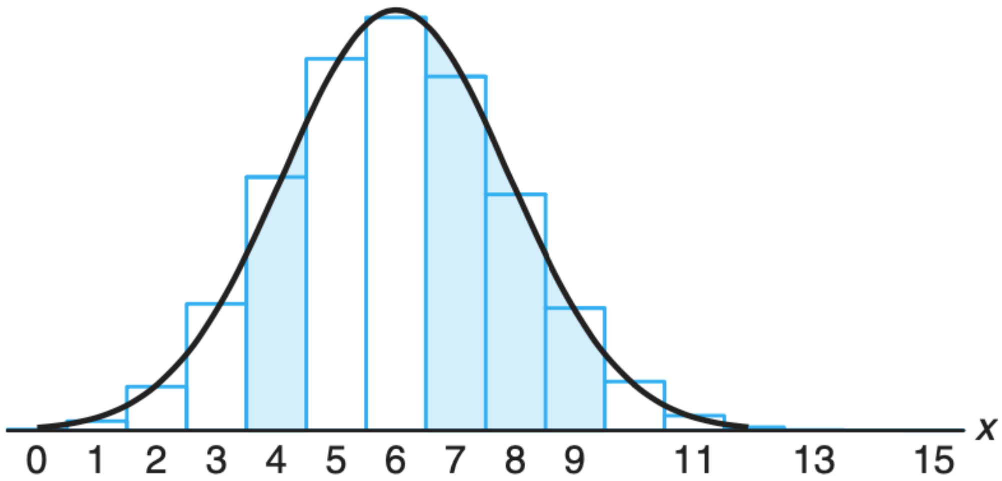
The normal distribution with $\mu = np$ and $\sigma^2 = np(1-p)$ provides a good approximation to the binomial distribution

a) when n is large and p is not extremely close to 0 or 1 (np and $n(1-p)$ are greater than or equal to 5); and

b) even when n is small and p is reasonably close to 1/2.

Normal approximation of *b*(*x*; 15,0.4)

$P(X = 4)$



Normal approximation to the Binomial

Normal Approximation to the Binomial Distribution

Let X be a binomial random variable with parameters n and p. For large n, X has approximately a normal distribution with $\mu = np$ and $\sigma^2 = npq = np(1-p)$ and

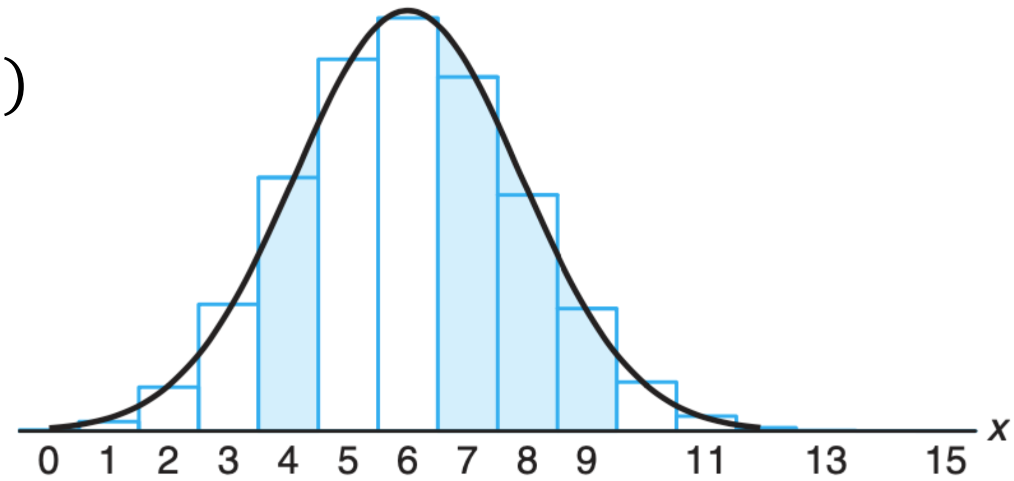
$$
\begin{aligned}
P(X \leq x) &= \sum_{k=0}^{x} b(k; n, p) \\
&\approx \text{area under normal curve to the left of } x + 0.5 \\
&= P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{npq}}\right),
\end{aligned}
$$

and the approximation will be good if np and $n(1-p)$ are greater than or equal to 5.

The correction +0.5 is called a continuity correction.

Normal approximation of $b(x; 15, 0.4)$

$P(7 \leq X \leq 9)$



Normal approximation to the Binomial

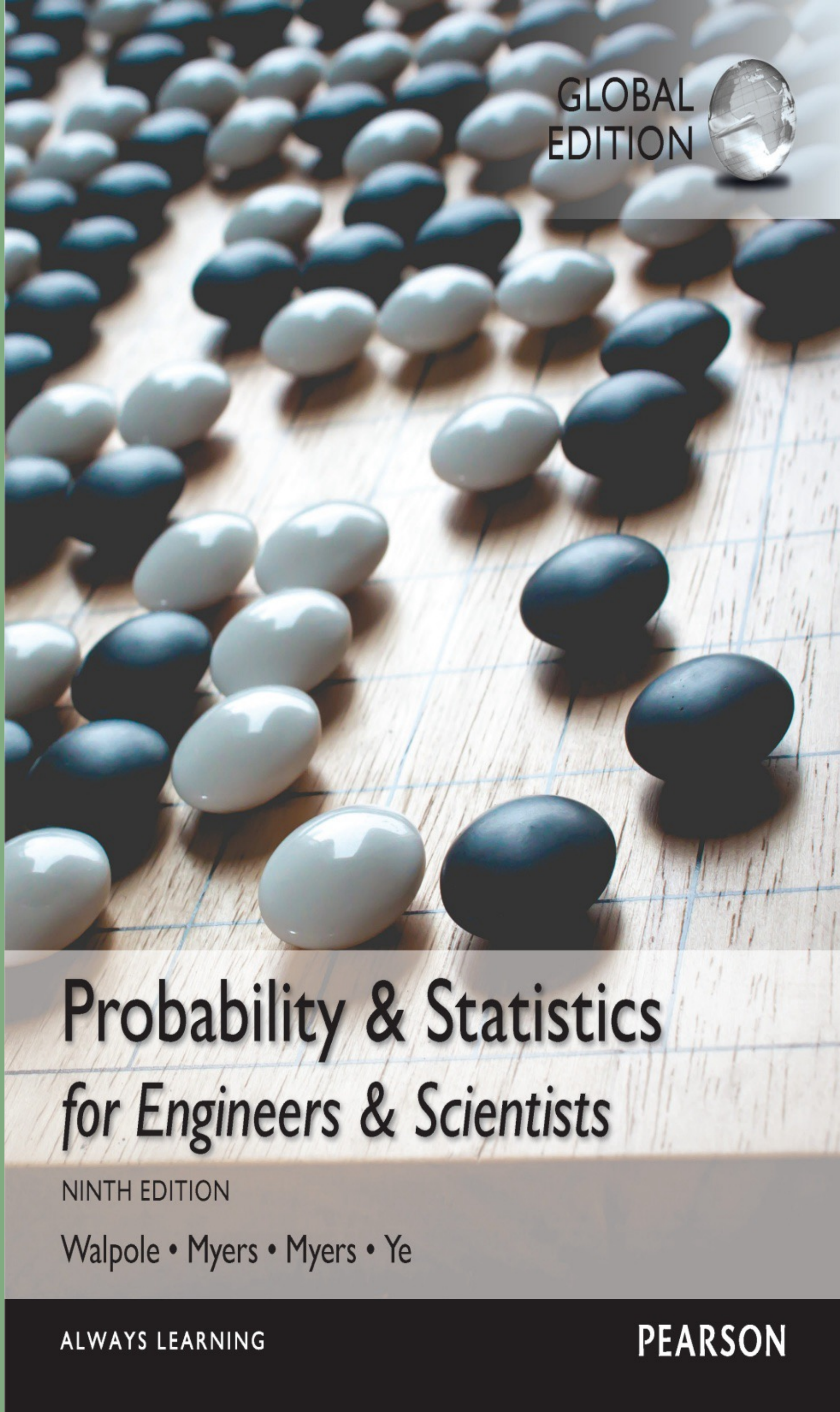
Example 1: The probability that a patient recovers from a rare blood disease is 0.4. If 100 people are known to have contracted this disease, what is the probability that fewer than 30 survive? **Use normal approximation**

Normal approximation to the Binomial

Example 2: A multiple-choice quiz has 200 questions, each with 4 possible answers of which only 1 is correct. What is the probability that sheer guesswork yields from 25 to 30 correct answers for the 80 of the 200 problems about which the student has no knowledge? **Use normal approximation**

Section 6.6

Gamma and Exponential Distributions



Gamma and Exponential Distributions

The exponential and gamma distributions play an important role in both queuing theory and reliability problems.

Time between arrivals at service facilities and time to failure of components and electrical systems often are nicely modeled by the exponential distribution.

The relationship between the gamma and the exponential allows the gamma to be used in similar types of problems.

Definition 6.2

The **gamma function** is defined by

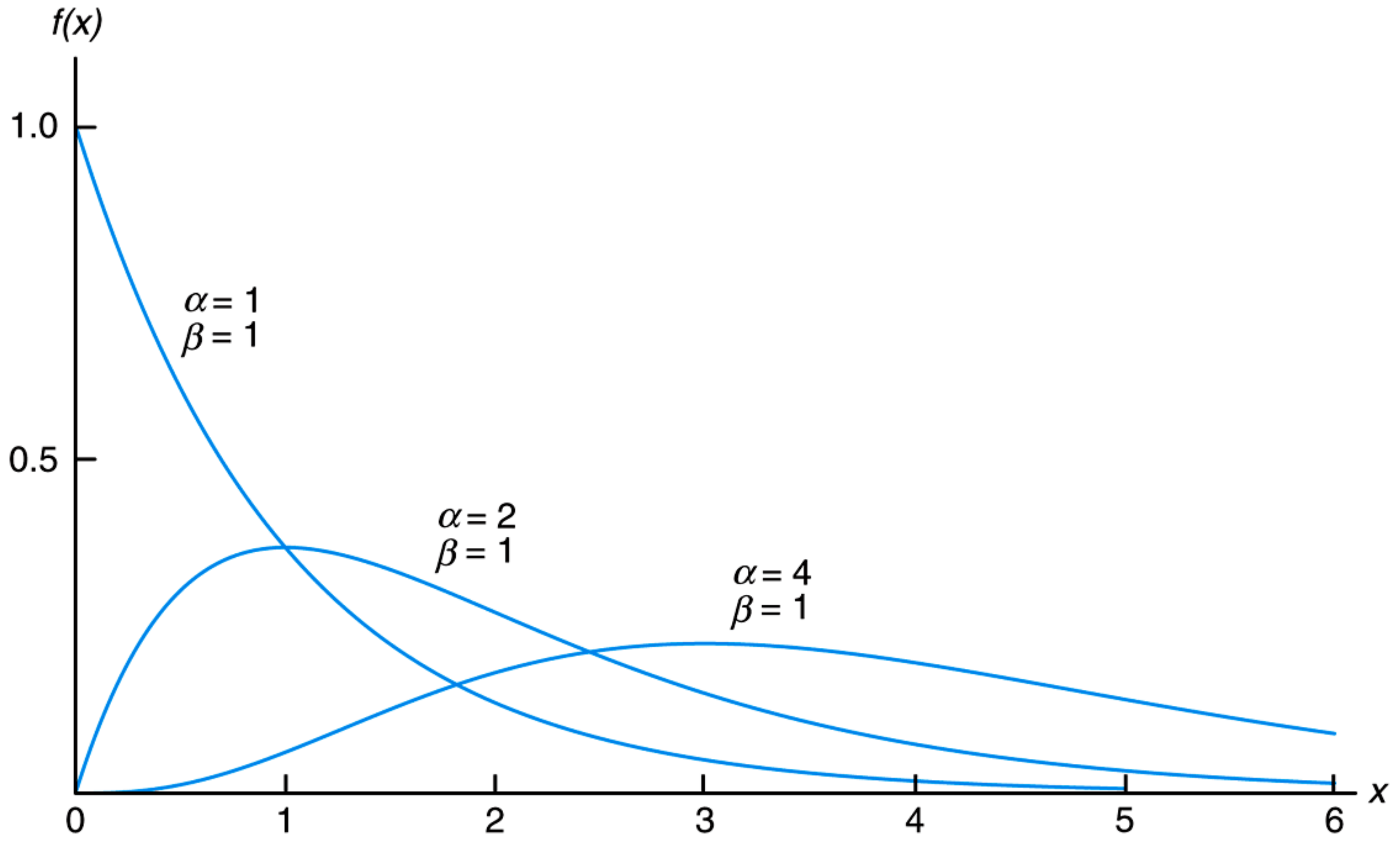
$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx, \qquad \text{for } \alpha > 0.$$

Gamma Distribution The continuous random variable X has a **gamma distribution**, with parameters α and β, if its density function is given by

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x > 0, \\ 0, & \text{elsewhere}, \end{cases}$$

where $\alpha > 0$ and $\beta > 0$.

Figure 6.28 Gamma distributions



Theorem 6.4

The mean and variance of the gamma distribution are

$$\mu = \alpha\beta \text{ and } \sigma^2 = \alpha\beta^2.$$

Exponential Distribution

The special gamma distribution for which $\alpha = 1$ is called the exponential distribution.

Exponential Distribution The continuous random variable X has an **exponential distribution**, with parameter β, if its density function is given by

$$f(x;\beta) = \begin{cases} \frac{1}{\beta}e^{-x/\beta}, & x > 0, \\ 0, & \text{elsewhere}, \end{cases}$$

where $\beta > 0$.

The mean and variance of the exponential distribution are

$$\mu = \beta \text{ and } \sigma^2 = \beta^2.$$

Exponential Distribution

Exponential Distribution The continuous random variable X has an **exponential distribution**, with parameter β, if its density function is given by

$$f(x;\beta) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, & x > 0, \\ 0, & \text{elsewhere}, \end{cases}$$

where $\beta > 0$.

Example: Suppose that a system contains a certain type of component whose time, in years, to failure is given by T . The random variable T is modelled nicely by the exponential distribution with mean time to failure $\beta = 5$. If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years? Assume that the components fail independently.

Memoryless Property of the Exponential Distribution

We say that a nonnegative random variable X is *memoryless* if

$$P\{X > s + t \mid X > t\} = P\{X > s\} \quad \text{for all } s, t \geq 0$$

The exponential distribution is the unique distribution having this property.

Example: Let X denote the time until detecting a particle with a Geiger counter and assume that X has an exponential distribution with $E(X) = 1.4$ minutes. Suppose we turn on the Geiger counter and wait 3 minutes without detecting a particle. What is the probability that a particle is detected in the next 30 seconds?

Relationship to the Poisson process

Suppose that we have a Poisson process with parameter $\lambda > 0$, where λ is the average number of occurrences per unit time. Let $N(t)$ be the number of occurrences in $[0, t]$. We know that $N(t)$ is a Poisson random variable with parameter λt, i.e.

$$P\{N(t) = x\} = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, x = 0,1,2, \ldots$$

Let T_1 denote the time until the first occurrence, and for $n > 1$, let T_n denote the time elapsed between the $(n-1)^{th}$ and the n^{th} occurrences.

Proposition 1: $T_1, T_2, \ldots$ are independent exponential random variables with parameter $1/\lambda$.

Relationship to the Poisson process

Proposition 1 leads to an alternative equivalent definition of the Poisson process:

1. Start with a sequence of independent exponential random variables $T_1, T_2, \ldots,$ with common parameter $1/\lambda$, and let these represent the interoccurrence times.
2. Record an occurrence at times $T_1, T_1 + T_2, T_1 + T_2 + T_3$,. etc.

These occurrences form a Poisson process with parameter λ.

Relationship to the Poisson process

Let S_n be the time until the nth occurrence, $n \geq 1$. Then, $S_n = \sum_{i=1}^{n} T_i$.

Proposition 2: S_n has a gamma distribution with parameters $\alpha = n, \beta = 1/\lambda$.

When α is a positive integer n, the gamma distribution is also known as the Erlang distribution. Thus, S_n has an Erlang distribution with parameters $\alpha = n, \beta = 1/\lambda$.

We have also shown that if $T_1, T_2, \ldots, T_\alpha$ are independent exponential random variables with parameter β (α is a positive integer), then $T_1 + T_2 + \cdots + T_\alpha$ has the Erlang (gamma) distribution with parameters α and β.

Example 1

The distance to the first major crack and the distance between successive major cracks in a highway follow an exponential distribution with a mean of 10 km. Assume that these distances are independent of each other. So, the successive major cracks form a Poisson process with parameter 0.1 per km.

(a) What is the probability that there are no major cracks in a 20-km stretch of the highway?

(b) What is the probability that there are two major cracks in a 20-km stretch of the highway?

(c) What is the probability that the first major crack occurs between 24 and 30 km of the start of inspection?

(d) What is the probability that there are no major cracks in two separate 10-km stretches of the highway?

(e) Given that there are no cracks in the first 10 km inspected, what is the probability that there are no major cracks in the next 20 km inspected?