

Econ 205

Continuous Random Variables and Probability
Distributions – 2

Course Outline

- **Probability**

- Probability and Inference
- The Discrete Case
- Calculating the Probability of an Event
- Conditional Probability and the Independence of Events
- Laws of Probability
- The Law of Total Probability and Bayes' Rule
- Introduction to Random Variables and their distributions

- **Discrete Random Variables and Their Probability Distributions**

- Discrete Random Variable
- The Probability Distribution for Discrete Random Variable
- The Expected Value of Random Variable
- Moments and Moment-Generating Functions
- The Bernoulli Probability Distribution
- The Binomial Probability Distribution
- The Geometric Probability Distribution
- The Negative Binomial Probability Distribution
- The Hypergeometric Probability Distribution
- The Poisson Probability Distribution
- Probability-Generating Functions

- **Continuous Random Variables and Their Probability Densities**

- Continuous Random Variable
- The Probability Distribution for Continuous Random Variable
- The Expected Value for Continuous Random Variable
- Moments and Moment-Generating Functions
- The Uniform Probability Distribution
- The Normal Probability Distribution
- Exponential Probability Distribution
- Chi-Square Probability Distribution
- Tchebysheff's Theorem

Course Outline (cont'd)

- **Multivariate Probability Distributions**

- Bivariate and Multivariate Probability Distributions
- Independent Random Variables
- Marginal and Conditional Probability Distributions
- The Expected Value of a Function of Random Variables
- Special Theorems
- The Covariance of Two Random Variables
- The Expected Value and Variance of Linear Functions of Random Variables
- The Multinomial Probability Distribution
- The Bivariate Normal Distribution
- Conditional Expectations

- **Functions of Random Variables**

- Finding the Probability Distribution of a Function of Random Variables
- The Method of Distribution Functions
- The Method of Transformations
- The Method of Moment-Generating Functions
- Multivariable Transformations Using Jacobians

Table of Contents of Devore & Berk

2	Probability	50
	Introduction	50
2.1	Sample Spaces and Events	51
2.2	Axioms, Interpretations, and Properties of Probability	56
2.3	Counting Techniques	66
2.4	Conditional Probability	74
2.5	Independence	84
3	Discrete Random Variables and Probability Distributions	96
	Introduction	96
3.1	Random Variables	97
3.2	Probability Distributions for Discrete Random Variables	101
3.3	Expected Values of Discrete Random Variables	112
3.4	Moments and Moment Generating Functions	121
3.5	The Binomial Probability Distribution	128
3.6	Hypergeometric and Negative Binomial Distributions	138
3.7	The Poisson Probability Distribution	146
4	Continuous Random Variables and Probability Distributions	158
	Introduction	158
4.1	Probability Density Functions and Cumulative Distribution Functions	159
4.2	Expected Values and Moment Generating Functions	171
4.3	The Normal Distribution	179
4.4	The Gamma Distribution and Its Relatives	194
4.5	Other Continuous Distributions	202
4.6	Probability Plots	210
4.7	Transformations of a Random Variable	220
5	Joint Probability Distributions	232
	Introduction	232
5.1	Jointly Distributed Random Variables	233
5.2	Expected Values, Covariance, and Correlation	245
5.3	Conditional Distributions	253
5.4	Transformations of Random Variables	265
5.5	Order Statistics	271

Content of this presentation

- **Continuous Random Variables and Probability Distributions**
 - The Gamma Distribution
 - The Exponential Distribution
 - The Chi-Squared Distribution
 - The Weibull Distribution
 - The Lognormal Distribution
 - The Beta Distribution
 - Chebyshev's/Tchebysheff's inequality/theorem
- **Examples**

The Gamma Distribution - 1

A family of pdf's that yields a wide variety of skewed distributional shapes is the gamma family.

Definition: For $\alpha > 0$, the **gamma function** $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

Properties:

The most important properties of the gamma function are the following:

1. For any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$ (via integration by parts)
2. For any positive integer, n , $\Gamma(n) = (n - 1)!$
3. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Definition:

A continuous random variable X is said to have a **gamma distribution** if the pdf of X is

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

where the parameters α and β satisfy $\alpha > 0$, $\beta > 0$. The **standard gamma distribution** has $\beta = 1$.

α is called the shape parameter. β is called the scale parameter.

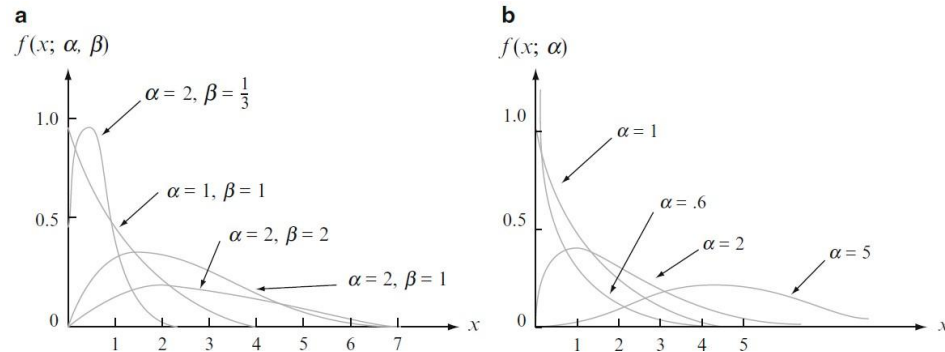


Figure 4.26 (a) Gamma density curves; (b) standard gamma density curves

Proposition:

The mean and variance of a random variable X having the gamma distribution $f(x; \alpha, \beta)$ are

$$E(X) = \mu = \alpha\beta \quad V(X) = \sigma^2 = \alpha\beta^2$$

The Gamma Distribution - 2

Proposition:

The moment generating function of a gamma random variable is

$$M_X(t) = \frac{1}{(1 - \beta t)^\alpha} \quad \text{for } t < \frac{1}{\beta}$$

Proof By definition, the mgf is

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-x/\beta} dx = \int_0^\infty \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-x(-t+1/\beta)} dx$$

One way to evaluate the integral is to express the integrand in terms of a gamma density. This means writing the exponent in the form $-x/b$ and having b take the place of β . We have $-x(-t + 1/\beta) = -x[(-\beta t + 1)/\beta] = -x/[\beta/(1 - \beta t)]$. Now multiplying and at the same time dividing the integrand by $1/(1 - \beta t)^\alpha$ gives

$$M_X(t) = \frac{1}{(1 - \beta t)^\alpha} \int_0^\infty \frac{x^{\alpha-1}}{\Gamma(\alpha)[\beta/(1 - \beta t)]^\alpha} e^{-x/[\beta/(1 - \beta t)]} dx$$

But now the integrand is a gamma pdf, so it integrates to 1. This establishes the result. ■

Definition:

When X is a standard gamma rv, the cdf of X , which is

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)} dy \quad x > 0 \quad (4.8)$$

is called the **incomplete gamma function** [sometimes the incomplete gamma function refers to Expression (4.8) without the denominator $\Gamma(\alpha)$ in the integrand].

Proposition:

Let X have a gamma distribution with parameters α and β . Then for any $x > 0$, the cdf of X is given by

$$P(X \leq x) = F(x; \alpha, \beta) = F\left(\frac{x}{\beta}; \alpha\right)$$

the incomplete gamma function evaluated at x/β .

Proof Calculate, with the help of the substitution $y = u/\beta$,

$$P(X \leq x) = \int_0^x \frac{u^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-u/\beta} du = \int_0^{x/\beta} \frac{y^{\alpha-1}}{\Gamma(\alpha)} e^{-y} dy = F\left(\frac{x}{\beta}; \alpha\right) \quad \blacksquare$$

For the second equality, note that

$$1) 0 < u < x \text{ iff } 0 < \beta y < x \text{ iff } 0 < y < \frac{x}{\beta} \quad \text{and } 2) du = \beta dy$$

The Exponential Distribution -1

Definition:

X is said to have an **exponential distribution** with parameter λ ($\lambda > 0$) if the pdf of X is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The exponential pdf is a special case of the general **gamma pdf** in which $\alpha = 1$ and β has been replaced by $1/\lambda$ [some authors use the form $(1/\beta)e^{-x/\beta}$].

Proposition:

The mean and the variance of the exponential distribution are:

$$\mu = \alpha\beta = \frac{1}{\lambda} \quad \sigma^2 = \alpha\beta^2 = \frac{1}{\lambda^2}$$

The cdf of the exponential distribution is:

$$F(x; \lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

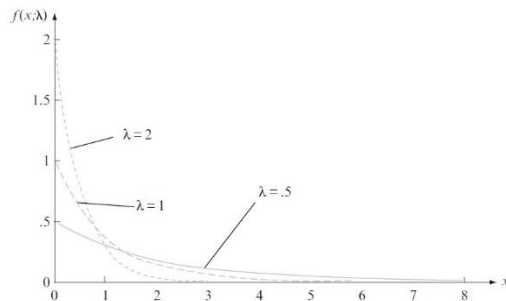


Figure 4.27 Exponential density curves

The Exponential Distribution -2

The exponential distribution is frequently used as a model for the distribution of times between the occurrence of successive events, such as customers arriving at a service facility or calls coming in to a switchboard. The reason for this is that the exponential distribution is closely related to the Poisson process discussed in Chapter 3.

Proposition:

Suppose that the number of events occurring in any time interval of length t has a Poisson distribution with parameter αt (where α , the rate of the event process, is the expected number of events occurring in 1 unit of time) and that numbers of occurrences in nonoverlapping intervals are independent of one another. Then the distribution of elapsed time between the occurrence of two successive events is exponential with parameter $\lambda = \alpha$.

Verification for the time X_1 until the first event occurs:

$$\begin{aligned} P(X_1 \leq t) &= 1 - P(X_1 > t) = 1 - P[\text{no events in } (0, t)] \\ &= 1 - \frac{e^{-\alpha t} \cdot (\alpha t)^0}{0!} = 1 - e^{-\alpha t} \end{aligned}$$

which is exactly the cdf of the exponential distribution.

the “memoryless” property of the exponential distribution:

Suppose component lifetime is exponentially distributed with parameter λ . After putting the component into service, we leave for a period of t_0 h and then return to find the component still working; what now is the probability that it lasts at least an additional t hours? In symbols, we wish $P(X \geq t + t_0 \mid X \geq t_0)$. By the definition of conditional probability,

$$P(X \geq t + t_0 \mid X \geq t_0) = \frac{P[(X \geq t + t_0) \cap (X \geq t_0)]}{P(X \geq t_0)}$$

But the event $X \geq t_0$ in the numerator is redundant, since both events can occur if and only if $X \geq t + t_0$. Therefore,

$$P(X \geq t + t_0 \mid X \geq t_0) = \frac{P(X \geq t + t_0)}{P(X \geq t_0)} = \frac{1 - F(t + t_0; \lambda)}{1 - F(t_0; \lambda)} = \frac{e^{-\lambda(t+t_0)}}{e^{-\lambda t_0}} = e^{-\lambda t}$$

This conditional probability is identical to the original probability $P(X \geq t)$ that the component lasted t hours. Thus *the distribution of additional lifetime is exactly the same as the original distribution of lifetime*, so at each point in time the component shows no effect of wear. In other words, the distribution of remaining lifetime is independent of current age.

The Chi-Squared Distribution

The chi-squared distribution is important because it is the basis for a number of procedures in statistical inference. The reason for this is that chi-squared distributions are intimately related to normal distributions.

Definition:

Let v be a positive integer. Then a random variable X is said to have a **chi-squared distribution** with parameter v if the pdf of X is the gamma density with $\alpha = v/2$ and $\beta = 2$. The pdf of a chi-squared rv is thus

$$f(x; v) = \begin{cases} \frac{1}{2^{v/2}\Gamma(v/2)} x^{(v/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The parameter v is called the **number of degrees of freedom** (df) of X . The symbol χ^2 is often used in place of “chi-squared.”

The Weibull Distribution - 1

Definition:

A random variable X is said to have a **Weibull distribution** with parameters α and β ($\alpha > 0$, $\beta > 0$) if the pdf of X is

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4.11)$$

When $\alpha = 1$, the pdf reduces to the exponential distribution (with $\lambda = 1/\beta$), so the exponential distribution is a special case of both the gamma and Weibull distributions. However, there are gamma distributions that are not Weibull distributions and vice versa, so one family is not a subset of the other.

Proposition:

The mean and the variance of the Weibull distribution are:

$$\mu = \beta \Gamma\left(1 + \frac{1}{\alpha}\right) \quad \sigma^2 = \beta^2 \left\{ \Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \right]^2 \right\}$$

The cdf of a Weibull rv having parameters α and β is

$$F(x; \alpha, \beta) = \begin{cases} 0 & x < 0 \\ 1 - e^{-(x/\beta)^\alpha} & x \geq 0 \end{cases} \quad (4.12)$$

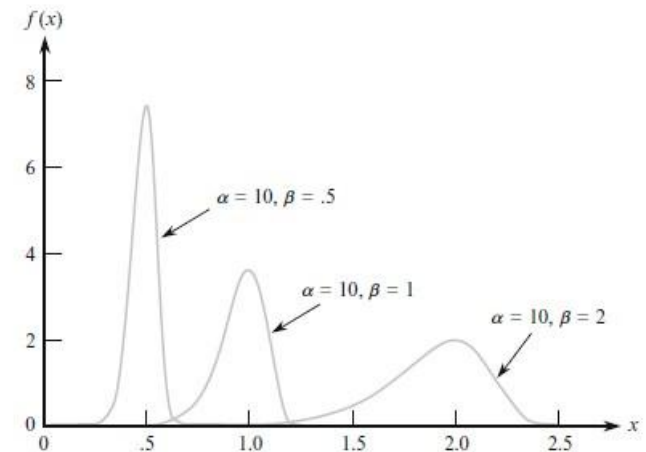
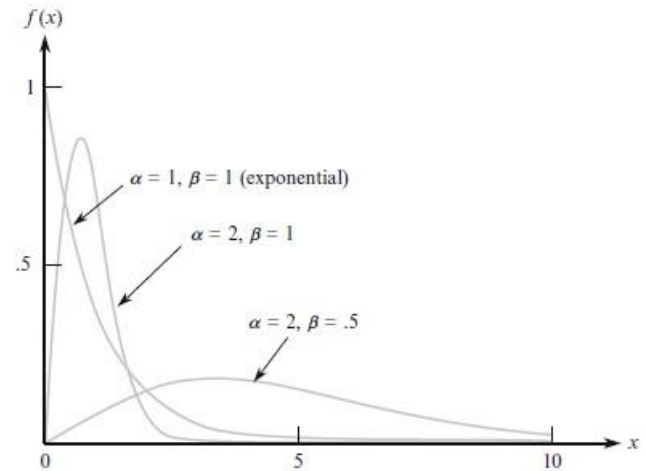


Figure 4.28 Weibull density curves

The Weibull Distribution - 2

Frequently, in practical situations, a Weibull model may be reasonable except that the smallest possible X value may be some value γ not assumed to be zero (this would also apply to a gamma model). The quantity γ can then be regarded as a third parameter of the distribution, which is what Weibull did in his original work. For, say, $\gamma = 3$, all curves in Figure 4.28 would be shifted 3 units to the right. This is equivalent to saying that $X - \gamma$ has the pdf (4.11), so that the cdf of X is obtained by replacing x in (4.12) by $x - \gamma$.

The Lognormal Distribution

Definition:

A nonnegative rv X is said to have a **lognormal distribution** if the rv $Y = \ln(X)$ has a normal distribution. The resulting pdf of a lognormal rv when $\ln(X)$ is normally distributed with parameters μ and σ is

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} e^{-[\ln(x) - \mu]^2 / (2\sigma^2)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Be careful here; the parameters μ and σ are not the mean and standard deviation of X but of $\ln(X)$. The mean and variance of X can be shown to be

Proposition:

$$E(X) = e^{\mu + \sigma^2/2} \quad V(X) = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$$

Because $\ln(X)$ has a normal distribution, the cdf of X can be expressed in terms of the cdf $\Phi(z)$ of a standard normal rv Z . For $x \geq 0$,

$$\begin{aligned} F(x; \mu, \sigma) &= P(X \leq x) = P[\ln(X) \leq \ln(x)] = P\left[\frac{\ln(X) - \mu}{\sigma} \leq \frac{\ln(x) - \mu}{\sigma}\right] \\ &= P\left[Z \leq \frac{\ln(x) - \mu}{\sigma}\right] = \Phi\left[\frac{\ln(x) - \mu}{\sigma}\right] \end{aligned}$$

a lognormal curve has a positive skew.

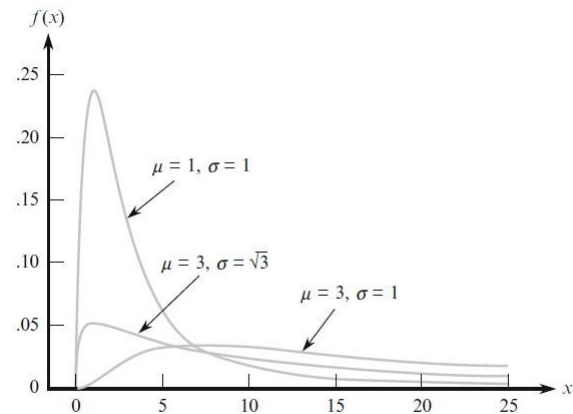


Figure 4.29 Lognormal density curves

The Beta Distribution

All families of continuous distributions discussed so far except for the uniform distribution have positive density over an infinite interval (although typically the density function decreases rapidly to zero beyond a few standard deviations from the mean). The beta distribution provides positive density only for X in an interval of finite length.

Definition:

A random variable X is said to have a **beta distribution** with parameters α, β (both positive), A , and B if the pdf of X is

$$f(x; \alpha, \beta, A, B) = \begin{cases} \frac{1}{B-A} \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\frac{x-A}{B-A}\right)^{\alpha-1} \left(\frac{B-x}{B-A}\right)^{\beta-1} & A \leq x \leq B \\ 0 & \text{otherwise} \end{cases}$$

The case $A = 0, B = 1$ gives the **standard beta distribution**.

Proposition:

The mean and the variance of the beta distribution

$$\mu = A + (B - A) \cdot \frac{\alpha}{\alpha + \beta} \quad \sigma^2 = \frac{(B - A)^2 \alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

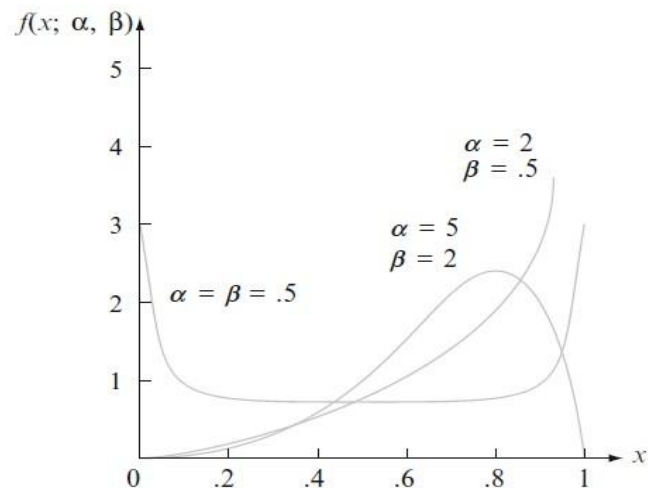


Figure 4.30 Standard beta density curves

Chebyshev's/Tchebysheff's inequality/theorem

One real value of Tchebysheff's theorem is that it enables us to find bounds for probabilities that ordinarily would have to be obtained by tedious mathematical manipulations (integration or summation). Further, we often can obtain means and variances of random variables without specifying the *distribution* of the variable. In situations like these, Tchebysheff's theorem still provides meaningful bounds for probabilities of interest.

Tchebysheff's Theorem Let Y be a random variable with finite mean μ and variance σ^2 . Then, for any $k > 0$,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad \text{or} \quad P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof: We will give the proof for a continuous random variable. The proof for the discrete case proceeds similarly. Let $f(y)$ denote the density function of Y . Then

$$\begin{aligned} V(Y) = \sigma^2 &= \int_{-\infty}^{\infty} (y - \mu)^2 f(y) dy \\ &= \int_{-\infty}^{\mu - k\sigma} (y - \mu)^2 f(y) dy + \int_{\mu - k\sigma}^{\mu + k\sigma} (y - \mu)^2 f(y) dy \\ &\quad + \int_{\mu + k\sigma}^{\infty} (y - \mu)^2 f(y) dy. \end{aligned}$$

The second integral is always greater than or equal to zero, and $(y - \mu)^2 \geq k^2\sigma^2$ for all values of y between the limits of integration for the first and third integrals; that is, the regions of integration are in the tails of the density function and cover only values of y for which $(y - \mu)^2 \geq k^2\sigma^2$. Replace the second integral by zero and substitute $k^2\sigma^2$ for $(y - \mu)^2$ in the first and third integrals to obtain the inequality

$$V(Y) = \sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} k^2\sigma^2 f(y) dy + \int_{\mu + k\sigma}^{\infty} k^2\sigma^2 f(y) dy.$$

Then

$$\sigma^2 \geq k^2\sigma^2 \left[\int_{-\infty}^{\mu - k\sigma} f(y) dy + \int_{\mu + k\sigma}^{\infty} f(y) dy \right],$$

or

$$\sigma^2 \geq k^2\sigma^2 [P(Y \leq \mu - k\sigma) + P(Y \geq \mu + k\sigma)] = k^2\sigma^2 P(|Y - \mu| \geq k\sigma).$$

Dividing by $k^2\sigma^2$, we obtain

$$P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2},$$

or, equivalently,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

EXAMPLES

Example 1 : The Standard Gamma Distribution

Suppose the reaction time X of a randomly selected individual to a certain stimulus has a standard gamma distribution with $\alpha = 2$. Since

$$P(a \leq X \leq b) = F(b) - F(a)$$

when X is continuous,

$$P(3 \leq X \leq 5) = F(5; 2) - F(3; 2) = .960 - .801 = .159$$

The probability that the reaction time is more than 4 s is

$$P(X > 4) = 1 - P(X \leq 4) = 1 - F(4; 2) = 1 - .908 = .092 \quad \blacksquare$$

Table A.4 The Incomplete Gamma Function

$$F(x; \alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y} dy$$

$x \backslash \alpha$	1	2	3	4	5	6	7	8	9	10
1	.632	.264	.080	.019	.004	.001	.000	.000	.000	.000
2	.865	.594	.323	.143	.053	.017	.005	.001	.000	.000
3	.950	.801	.577	.353	.185	.084	.034	.012	.004	.001
4	.982	.908	.762	.567	.371	.215	.111	.051	.021	.008
5	.993	.960	.875	.735	.560	.384	.238	.133	.068	.032
6	.998	.983	.938	.849	.715	.554	.394	.256	.153	.084
7	.999	.993	.970	.918	.827	.699	.550	.401	.271	.170
8	1.000	.997	.986	.958	.900	.809	.687	.547	.407	.283
9		.999	.994	.979	.945	.884	.793	.676	.544	.413
10		1.000	.997	.990	.971	.933	.870	.780	.667	.542
11			.999	.995	.985	.962	.921	.857	.768	.659
12			1.000	.998	.992	.980	.954	.911	.845	.758
13				.999	.996	.989	.974	.946	.900	.834
14				1.000	.998	.994	.986	.968	.938	.891
15					.999	.997	.992	.982	.963	.930

Example 2 : The Gamma Distribution

Suppose the survival time X in weeks of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with $\alpha = 8$ and $\beta = 15$. (Data in *Survival Distributions: Reliability Applications in the Biomedical Services*, by A. J. Gross and V. Clark, suggests $\alpha \approx 8.5$ and $\beta \approx 13.3$.) The expected survival time is $E(X) = (8)(15) = 120$ weeks, whereas $V(X) = (8)(15)^2 = 1,800$ and $\sigma_X = \sqrt{1800} = 42.43$ weeks. The probability that a mouse survives between 60 and 120 weeks is

$$\begin{aligned} P(60 \leq X \leq 120) &= P(X \leq 120) - P(X \leq 60) \\ &= F(120/15; 8) - (60/5; 8) \\ &= F(8; 8) - F(4; 8) = .547 - .051 = .496 \end{aligned}$$

The probability that a mouse survives at least 30 weeks is

$$\begin{aligned} P(X \geq 30) &= 1 - P(X < 30) = 1 - P(X \leq 30) \\ &= 1 - F(30/15; 8) = .999 \end{aligned}$$

Table A.4 The Incomplete Gamma Function

$$F(x; \alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y} dy$$

$x \backslash \alpha$	1	2	3	4	5	6	7	8	9	10
1	.632	.264	.080	.019	.004	.001	.000	.000	.000	.000
2	.865	.594	.323	.143	.053	.017	.005	.001	.000	.000
3	.950	.801	.577	.353	.185	.084	.034	.012	.004	.001
4	.982	.908	.762	.567	.371	.215	.111	.051	.021	.008
5	.993	.960	.875	.735	.560	.384	.238	.133	.068	.032
6	.998	.983	.938	.849	.715	.554	.394	.256	.153	.084
7	.999	.993	.970	.918	.827	.699	.550	.401	.271	.170
8	1.000	.997	.986	.958	.900	.809	.687	.547	.407	.283
9		.999	.994	.979	.945	.884	.793	.676	.544	.413
10		1.000	.997	.990	.971	.933	.870	.780	.667	.542
11			.999	.995	.985	.962	.921	.857	.768	.659
12			1.000	.998	.992	.980	.954	.911	.845	.758
13				.999	.996	.989	.974	.946	.900	.834
14				1.000	.998	.994	.986	.968	.938	.891
15					.999	.997	.992	.982	.963	.930

Example 3: The Gamma Distribution

Prove the following using expected values:

If Y has a gamma distribution with parameters α and β , then

$$\mu = E(Y) = \alpha\beta \quad \text{and} \quad \sigma^2 = V(Y) = \alpha\beta^2.$$

Proof:

$$E(Y) = \int_{-\infty}^{\infty} yf(y) dy = \int_0^{\infty} y \left(\frac{y^{\alpha-1} e^{-y/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \right) dy.$$

By definition, the gamma density function is such that

$$\int_0^{\infty} \frac{y^{\alpha-1} e^{-y/\beta}}{\beta^{\alpha} \Gamma(\alpha)} dy = 1.$$

Hence,

$$\int_0^{\infty} y^{\alpha-1} e^{-y/\beta} dy = \beta^{\alpha} \Gamma(\alpha),$$

and

$$\begin{aligned} E(Y) &= \int_0^{\infty} \frac{y^{\alpha} e^{-y/\beta}}{\beta^{\alpha} \Gamma(\alpha)} dy = \frac{1}{\beta^{\alpha} \Gamma(\alpha)} \int_0^{\infty} y^{\alpha} e^{-y/\beta} dy \\ &= \frac{1}{\beta^{\alpha} \Gamma(\alpha)} [\beta^{\alpha+1} \Gamma(\alpha+1)] = \frac{\beta^{\alpha} \Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta. \end{aligned}$$

$V(Y) = E[Y^2] - [E(Y)]^2$. Further,

$$\begin{aligned} E(Y^2) &= \int_0^{\infty} y^2 \left(\frac{y^{\alpha-1} e^{-y/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \right) dy = \frac{1}{\beta^{\alpha} \Gamma(\alpha)} \int_0^{\infty} y^{\alpha+1} e^{-y/\beta} dy \\ &= \frac{1}{\beta^{\alpha} \Gamma(\alpha)} [\beta^{\alpha+2} \Gamma(\alpha+2)] = \frac{\beta^2(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha(\alpha+1)\beta^2. \end{aligned}$$

Then $V(Y) = E[Y^2] - [E(Y)]^2$ where, from the earlier part of the derivation, $E(Y) = \alpha\beta$. Substituting $E[Y^2]$ and $E(Y)$ into the formula for $V(Y)$, we obtain

$$V(Y) = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha^2\beta^2 + \alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2.$$

Example 4: The Exponential Distribution

The response time X at an on-line computer terminal (the elapsed time between the end of a user's inquiry and the beginning of the system's response to that inquiry) has an exponential distribution with expected response time equal to 5 s. Then $E(X) = 1/\lambda = 5$, so $\lambda = .2$. The probability that the response time is at most 10 s is

$$P(X \leq 10) = F(10; .2) = 1 - e^{-(.2)(10)} = 1 - e^{-2} = 1 - .135 = .865$$

The probability that response time is between 5 and 10 s is

$$P(5 \leq X \leq 10) = F(10; .2) - F(5; .2) = (1 - e^{-2}) - (1 - e^{-1}) = .233 \quad \blacksquare$$

Recall the pdf and the cdf of the exponential distribution:

X is said to have an **exponential distribution** with parameter λ ($\lambda > 0$) if the pdf of X is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The cdf of the exponential distribution is:

$$F(x; \lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

Example 5: The Weibull Distribution

In recent years the Weibull distribution has been used to model engine emissions of various pollutants. Let X denote the amount of NO_x emission (g/gal) from a randomly selected four-stroke engine of a certain type, and suppose that X has a Weibull distribution with $\alpha = 2$ and $\beta = 10$.

$$P(X \leq 10) = F(10; 2, 10) = 1 - e^{-(10/10)^2} = 1 - e^{-1} = .632$$

Similarly, $P(X \leq 25) = .998$, so the distribution is almost entirely concentrated on values between 0 and 25. The value c , which separates the 5% of all engines having the largest amounts of NO_x emissions from the remaining 95%, satisfies

$$.95 = 1 - e^{-(c/10)^2}$$

Isolating the exponential term on one side, taking logarithms, and solving the resulting equation gives $c \approx 17.3$ as the 95th percentile of the emission distribution. ■

Recall

The cdf of a Weibull rv having parameters α and β is

$$F(x; \alpha, \beta) = \begin{cases} 0 & x < 0 \\ 1 - e^{-(x/\beta)^\alpha} & x \geq 0 \end{cases}$$

Example 6: The Lognormal Distribution

According to the article “Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines” (*Corrosion*, 2009: 332–342), the lognormal distribution has been reported as the best option for describing the distribution of maximum pit depth data from cast iron pipes in soil. The authors suggest that a lognormal distribution with $\mu = .353$ and $\sigma = .754$ is appropriate for maximum pit depth (mm) of buried pipelines. For this distribution, the mean value and variance of pit depth are

$$E(X) = e^{.353 + (.754)^2/2} = e^{.6383} = 1.893$$

$$V(X) = e^{2(.353) + (.754)^2} \cdot (e^{.754^2} - 1) = (3.57697)(.765645) = 2.7387$$

The probability that maximum pit depth is between 1 and 2 mm is

$$\begin{aligned} P(1 \leq X \leq 2) &= P(\ln(1) \leq \ln(X) \leq \ln(2)) \\ &= P(0 \leq \ln(X) \leq .693) \\ &= P\left(\frac{0 - .353}{.754} \leq Z \leq \frac{.693 - .353}{.754}\right) \\ &= \Phi(.45) - \Phi(-.47) = .354 \end{aligned}$$

What value c is such that only 1% of all specimens have a maximum pit depth exceeding c ? The desired value satisfies

$$.99 = P(X \leq c) = P\left(Z \leq \frac{\ln(c) - .353}{.754}\right)$$

The z critical value 2.33 captures an upper-tail area of .01 ($z_{.01} = 2.33$), and thus a cumulative area of .99. This implies that

$$\frac{\ln(c) - .353}{.754} = 2.33$$

from which $\ln(c) = 2.1098$ and $c = 8.247$. Thus 8.247 is the 99th percentile of the maximum pit depth distribution. ■

Recall the mean and the variance of the lognormal distribution::

$$E(X) = e^{\mu + \sigma^2/2} \quad V(X) = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$$

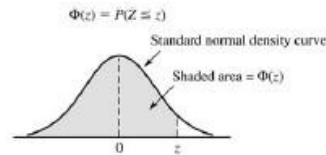
Also recall:

Because $\ln(X)$ has a normal distribution, the cdf of X can be expressed in terms of the cdf $\Phi(z)$ of a standard normal rv Z . For $x \geq 0$,

$$\begin{aligned} F(x; \mu, \sigma) &= P(X \leq x) = P[\ln(X) \leq \ln(x)] = P\left[\frac{\ln(X) - \mu}{\sigma} \leq \frac{\ln(x) - \mu}{\sigma}\right] \\ &= P\left[Z \leq \frac{\ln(x) - \mu}{\sigma}\right] = \Phi\left[\frac{\ln(x) - \mu}{\sigma}\right] \end{aligned}$$

Table A3 Devore & Berk

Table A.3 Standard Normal Curve Areas



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3482
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

(continued)

Table A.3 Standard Normal Curve Areas (cont.)

$$\Phi(z) = P(Z \leq z)$$

[illegible]

Example 7: The Beta Distribution

A gasoline wholesale distributor has bulk storage tanks that hold fixed supplies and are filled every Monday. Of interest to the wholesaler is the proportion of this supply that is sold during the week. Over many weeks of observation, the distributor found that this proportion could be modeled by a beta distribution with $\alpha = 4$ and $\beta = 2$. Find the probability that the wholesaler will sell at least 90% of her stock in a given week.

should be “standard” beta distribution

Solution If Y denotes the proportion sold during the week, then

$$f(y) = \begin{cases} \frac{\Gamma(4+2)}{\Gamma(4)\Gamma(2)} y^3(1-y), & 0 \leq y \leq 1, \\ 0, & \text{elsewhere,} \end{cases}$$

and

$$\begin{aligned} P(Y > .9) &= \int_{.9}^{\infty} f(y) dy = \int_{.9}^1 20(y^3 - y^4) dy \\ &= 20 \left\{ \frac{y^4}{4} \Big|_{.9}^1 - \frac{y^5}{5} \Big|_{.9}^1 \right\} = 20(.004) = .08. \end{aligned}$$

It is *not* very likely that 90% of the stock will be sold in a given week. ■

Example 8: The Beta Distribution

Project managers often use a method labeled PERT—for program evaluation and review technique—to coordinate the various activities making up a large project. (One successful application was in the construction of the *Apollo* spacecraft.) A standard assumption in PERT analysis is that the time necessary to complete any particular activity once it has been started has a beta distribution with A = the optimistic time (if everything goes well) and B = the pessimistic time (if everything goes badly). Suppose that in constructing a single-family house, the time X (in days) necessary for laying the foundation has a beta distribution with $A = 2$, $B = 5$, $\alpha = 2$, and $\beta = 3$. Then $\alpha/(\alpha + \beta) = .4$, so $E(X) = 2 + (3)(.4) = 3.2$.

For these values of α and β , the pdf of X is a simple polynomial function. The probability that it takes at most 3 days to lay the foundation is

$$\begin{aligned} P(X \leq 3) &= \int_2^3 \frac{1}{3} \cdot \frac{4!}{1! \cdot 2!} \left(\frac{x-2}{3}\right) \left(\frac{5-x}{3}\right)^2 dx \\ &= \frac{4}{27} \int_2^3 (x-2)(5-x)^2 dx = \frac{4}{27} \cdot \frac{11}{4} = \frac{11}{27} = .407 \end{aligned} \quad \blacksquare$$

Recall the pdf, the mean and the variance of the beta distribution:

A random variable X is said to have a **beta distribution** with parameters α , β (both positive), A , and B if the pdf of X is

$$f(x; \alpha, \beta, A, B) = \begin{cases} \frac{1}{B-A} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} \left(\frac{x-A}{B-A}\right)^{\alpha-1} \left(\frac{B-x}{B-A}\right)^{\beta-1} & A \leq x \leq B \\ 0 & \text{otherwise} \end{cases}$$

The case $A = 0$, $B = 1$ gives the **standard beta distribution**.

The mean and the variance of the beta distribution

$$\mu = A + (B - A) \cdot \frac{\alpha}{\alpha + \beta} \qquad \sigma^2 = \frac{(B - A)^2 \alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

Examples 9-10: Chebyshev's/Tchebysheff's inequality/theorem

Example 9:

Suppose that experience has shown that the length of time Y (in minutes) required to conduct a periodic maintenance check on a dictating machine follows a gamma distribution with $\alpha = 3.1$ and $\beta = 2$. A new maintenance worker takes 22.5 minutes to check the machine. Does this length of time to perform a maintenance check disagree with prior experience?

Solution The mean and variance for the length of maintenance check times (based on prior experience) are (from Theorem 4.8)

$$\mu = \alpha\beta = (3.1)(2) = 6.2 \quad \text{and} \quad \sigma^2 = \alpha\beta^2 = (3.1)(2^2) = 12.4.$$

It follows that $\sigma = \sqrt{12.4} = 3.52$. Notice that $y = 22.5$ minutes exceeds the mean $\mu = 6.2$ minutes by 16.3 minutes, or $k = 16.3/3.52 = 4.63$ standard deviations. Then from Tchebysheff's theorem,

$$P(|Y - 6.2| \geq 16.3) = P(|Y - \mu| \geq 4.63\sigma) \leq \frac{1}{(4.63)^2} = .0466.$$

This probability is based on the assumption that the distribution of maintenance times has not changed from prior experience. Then, observing that $P(Y \geq 22.5)$ is small, we must conclude either that our new maintenance worker has generated by chance a lengthy maintenance time that occurs with low probability or that the new worker is somewhat slower than preceding ones. Considering the low probability for $P(Y \geq 22.5)$, we favor the latter view. ■

Example 10:

The number of customers per day at a sales counter, Y , has been observed for a long period of time and found to have mean 20 and standard deviation 2. The probability distribution of Y is not known. What can be said about the probability that, tomorrow, Y will be greater than 16 but less than 24?

Solution We want to find $P(16 < Y < 24)$. From Theorem 4.8 we know that, for any $k \geq 0$, $P(|Y - \mu| < k\sigma) \geq 1 - 1/k^2$, or

$$P[(\mu - k\sigma) < Y < (\mu + k\sigma)] \geq 1 - \frac{1}{k^2}.$$

Because $\mu = 20$ and $\sigma = 2$, it follows that $\mu - k\sigma = 16$ and $\mu + k\sigma = 24$ if $k = 2$. Thus,

$$P(16 < Y < 24) = P(\mu - 2\sigma < Y < \mu + 2\sigma) \geq 1 - \frac{1}{(2)^2} = \frac{3}{4}.$$

In other words, tomorrow's customer total will be between 16 and 24 with a fairly high probability (at least 3/4).

Notice that if σ were 1, k would be 4, and

$$P(16 < Y < 24) = P(\mu - 4\sigma < Y < \mu + 4\sigma) \geq 1 - \frac{1}{(4)^2} = \frac{15}{16}.$$

Thus, the value of σ has considerable effect on probabilities associated with intervals. ■

Chebyshev's/Tchebysheff's inequality/theorem

One real value of Tchebysheff's theorem is that it enables us to find bounds for probabilities that ordinarily would have to be obtained by tedious mathematical manipulations (integration or summation). Further, we often can obtain means and variances of random variables without specifying the *distribution* of the variable. In situations like these, Tchebysheff's theorem still provides meaningful bounds for probabilities of interest.

Tchebysheff's Theorem Let Y be a random variable with finite mean μ and variance σ^2 . Then, for any $k > 0$,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad \text{or} \quad P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof: We will give the proof for a continuous random variable. The proof for the discrete case proceeds similarly. Let $f(y)$ denote the density function of Y . Then

$$\begin{aligned} V(Y) = \sigma^2 &= \int_{-\infty}^{\infty} (y - \mu)^2 f(y) dy \\ &= \int_{-\infty}^{\mu - k\sigma} (y - \mu)^2 f(y) dy + \int_{\mu - k\sigma}^{\mu + k\sigma} (y - \mu)^2 f(y) dy \\ &\quad + \int_{\mu + k\sigma}^{\infty} (y - \mu)^2 f(y) dy. \end{aligned}$$

The second integral is always greater than or equal to zero, and $(y - \mu)^2 \geq k^2\sigma^2$ for all values of y between the limits of integration for the first and third integrals; that is, the regions of integration are in the tails of the density function and cover only values of y for which $(y - \mu)^2 \geq k^2\sigma^2$. Replace the second integral by zero and substitute $k^2\sigma^2$ for $(y - \mu)^2$ in the first and third integrals to obtain the inequality

$$V(Y) = \sigma^2 \geq \int_{-\infty}^{\mu - k\sigma} k^2\sigma^2 f(y) dy + \int_{\mu + k\sigma}^{\infty} k^2\sigma^2 f(y) dy.$$

Then

$$\sigma^2 \geq k^2\sigma^2 \left[\int_{-\infty}^{\mu - k\sigma} f(y) dy + \int_{\mu + k\sigma}^{\infty} f(y) dy \right],$$

or

$$\sigma^2 \geq k^2\sigma^2 [P(Y \leq \mu - k\sigma) + P(Y \geq \mu + k\sigma)] = k^2\sigma^2 P(|Y - \mu| \geq k\sigma).$$

Dividing by $k^2\sigma^2$, we obtain

$$P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2},$$

or, equivalently,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

References

- Devore, J.L. and Berk, K.N. (2012) Modern Mathematical Statistics with Applications, 2nd Edition, New York: Springer.
- Wackerly, D.D., Mendenhall, W. and Scheaffer, R.L., (2008) Mathematical Statistics With Applications, 7th Edition, Pacific Grove: Duxbury.