STAT 3600   Reference: Chapter 4 of Devore(8e)   Maghsoodloo

CONTINUOUS RANDOM VARIABLES

Let X be a continuous rv with range space R_x (a subset of real numbers). Since the Pr that the rv, X, takes on a single specific value in R_x is zero, then Prs over a continuum must be defined only over intervals within R_x. The function, f(x), is said to be a Pr density function (pdf) over R_x iff

(1) f(x) $\geq$ 0 for all x in R_x, (2) $\int_{R_x} f(x)\,dx$ = 1, (3) f(x) must be (piecewise) continuous, (4) f(x) = 0

for all x $\notin$ R_x, and (5) P(a $\leq$ X $\leq$ b) = P(a $\leq$ X < b) = P(a < X < b) = P(a < X $\leq$ b) = $\int_a^b f(x)\,dx$.

Note that all these last 4 Prs are equal because the Pr for a single point (or any set of measure 0) for any continuous rv is identically zero. Recall also that the subset of R_x on which f(x) > 0, is called the SUPS of f(x).

Example 25.   The time to failure, T measured in hours, of a computer that guides a 4-engine aircraft has the pdf $f(t) = ce^{-0.0004t}$, t $\geq$ 0 and f(t) = 0 for t < 0. (a) Determine the constant c such that f(t) is a Pr density function (pdf). The SUPS of f(t) = [0, ∞).

$$\int_{R_t} f(t)\,dt \;=\; \int_0^\infty c\,e^{-0.0004t}\,dt \;=\; c\left[\frac{e^{-0.0004t}}{-0.0004}\right]_0^\infty \;=\; c[-(0-1)/0.0004]$$

$$= c/0.0004 = 1 \qquad \rightarrow \qquad c = 0.0004.$$

Therefore, the only unique value of the constant c that makes the function $ce^{-0.0004t}$, t $\geq$ 0, a pdf (Pr density function) is c = 0.0004. Note that $f(t) = 0.0004e^{-0.0004t}$ is called the exponential density function at the rate of λ = 0.0004 failures (or Poisson events) per hour.

(b) Compute the Pr that the computer will survive a 7-hour flight. (TTF=Time-to-Failure)

$$P(\text{TTF} > 7 \text{ hours }) = \int_7^\infty 0.0004\,e^{-0.0004t}\,dt = e^{-0.0004\times7} = e^{-0.0028} = 0.997204.$$

The above Pr is called the reliability (or survival Pr) of the PC at t = 7 hours. This implies that the failure Pr of the PC during the 7-hour flight is P(T $\leq$ 7) = F_T(at 7 hours) = 1 $-$ 0.997204 = 0.002796,

while the Reliability(at 7 hours) = R(at 7 hours) = 0.997204 = 1– the cdf at 7 hours.

(c) Obtain the cdf (cumulative distribution function) of TTF, T, at time t.

$$F_T(t) = \Pr(T \le t) = \int_0^t \lambda e^{-\lambda x} dx = \left[\frac{e^{-\lambda x}}{-1}\right]_0^t = 1 - e^{-\lambda t} = 1 - e^{-0.0004t}$$

Note that once the cdf of a rv is obtained, then all Prs can be computed using the cdf. For example, we may use the cdf to re-compute the reliability of the PC in part (b) above.

$R(7 \text{ hours}) = \Pr(TTF > 7) = 1 - \Pr(T \le 7) = 1 - F_T(7) = 1 - (1 - e^{-\lambda t})$

$\qquad = e^{-\lambda t} = e^{-7\lambda} = e^{-0.0028} = 0.997204.$

(d) Use the cdf to compute the Pr that the PC will fail within the interval (10, 20 hours).

$\Pr(10 \le T \le 20 \text{ hours}) = \Pr(T \le 20) - \Pr(T \le 10) = F(20) - F(10) =$

$\qquad = (1 - e^{-20\lambda}) - (1 - e^{-10\lambda}) = e^{-10\lambda} - e^{-20\lambda} =$

$\qquad = R(10) - R(20) = e^{-0.004} - e^{-0.008} = 0.0039761.$

(e) Suppose it is known that on a certain flight the PC has already lasted 7 hours. What is the Pr that its lifetime will go beyond 20 hours?

$$P(T > 20 \mid T > 7) = \frac{P(T > 20 \cap T > 7)}{P(TTF > 7)} = \frac{P(T > 20)}{P(T > 7)} = \frac{e^{-20\lambda}}{e^{-7\lambda}} = e^{-13\lambda}$$

$= \Pr(T > 13 \text{ hours} \mid T > 0) = \Pr(T > 13 \text{ hours}) = R(13) = e^{-0.0052} = 0.9948145.$

The above developments show that the exponential density function is memory-less because the conditional Pr that the PC's lifetime exceeds 20 hours given that it has already lasted 7 hours is the same as the unconditional Pr that the PC lasts 13 hours from time zero. Thus, in general the exponential pdf has the following memory-less property

$$P(T > a+b \mid T > a) = P(T > b) = e^{-\lambda b}.$$

The discrete analogue of the exponential density is the Geometric distribution g(x; p). To my knowledge, these are the only two memory-less statistical distributions. The percentile (or quantile) function of the exponential at the rate λ is given by $x_p = \frac{1}{\lambda}\ln\left(\frac{1}{1-p}\right)$, $0 \le p \le 1$, i.e., the inverse

function of the exponential density is given by $F^{-1}(p) = \frac{1}{\lambda}\ln\left(\frac{1}{1-p}\right)$. For example,

the median, p = 0.50, is given by $x_{0.50} = F^{-1}(0.50) = \frac{1}{\lambda}\ln\left(\frac{1}{1-0.5}\right) = \frac{1}{\lambda}\ln(2)$ = 0.6931472/λ < as

expected because the exponential is positively skewed. The first quartile is given by

$F^{-1}(0.25) = \frac{1}{\lambda}\ln\left(\frac{1}{1-0.25}\right) = \frac{1}{\lambda}\ln(4/3)$ → IQR = $\frac{1}{\lambda}[\ln(\frac{1}{1-0.75}) - \ln(4/3)] = \ln(3)/\lambda$.

Exercise 29. The acidity of a chemical compound, measured on the
limited scale R_x = [1, 5%], has the pdf

$$f(x) = \begin{cases} a\,x + b, & 1 \le x \le 2.5 \\ c\,x + d, & 2.5 \le x \le 5.0 \\ 0, & \text{elsewhere} \end{cases}$$

The above pdf is called the asymmetrical triangular distribution. (a) Determine the constants a, b, c,
and d such that f(x) is a pdf given that 2.5a+b = 2.5c+d. (b) Obtain the cdf of X, F(x) and compute its
value at 0, 2, 4 and 7. (c) For a randomly selected item from the above distribution function, use the
cdf to compute the Pr that its acidity exceeds 4%. (d) Compute the $P(X \le 4 \,|\, X > 2)$, again using the
cdf.

THE EXPECTED VALUE OF A CONTINUOUS RV

Definition: Recall that if X is a discrete rv over the range R_x with pmf p(x), then the population
mean of the rv, X, is defined as its expected value and is given by the weighted average μ = E(X) =
$\sum_{R_X} xp(x)$, where the sum extends over the range space of X. However, if X is continuous with pdf f(x),
then

$$\mu = E(X) = \int_{R_x} xf(x)\,dx \qquad\qquad (10)$$

Since an integral is simply a limiting sum over a continuum and f(x)dx represents the Pr element at x
(actually over an interval about x of length dx), then the form of E(X) for a continuous rv is identical to
that of the discrete case.

Example 25 (continued). Compute the average lifetime (or mean time to failure = MTTF) of
the PC that guides the aircraft.

41

$$\text{MTTF} = \mu = E(X) = \int_0^\infty x\lambda e^{-\lambda x} dx \ , \quad \text{where } \lambda = 0.0004. \tag{11}$$

The above integration can be carried out in 2 distinct different ways: (i) thru integration by parts, (ii) or by partial differentiation.

(i) Using integration by parts : $\int_a^b u(x)\,dv = u(x)\,v(x)]_a^b - \int_a^b v(x)\,du$.

Letting u(x) = x, and v(x) = $\lambda e^{-\lambda x}$dx in equation (11) results in

$$E(X) = \left[-xe^{-\lambda x} \right]_0^\infty - \int_0^\infty -e^{-\lambda x} dx = 0 + \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^\infty = 1/\lambda = 1/0.0004 \ \rightarrow$$

$$\text{MTTF} = \mu = E(X) = 2500 \text{ hours.}$$

(ii) Using Partial Differentiation:

$$\mu = E(X) = \int_0^\infty x\,\lambda e^{-\lambda x} dx = -\lambda \frac{\partial}{\partial \lambda} \int_0^\infty e^{-\lambda x} dx = -\lambda \frac{\partial}{\partial \lambda}(\frac{1}{\lambda}) = -\lambda\,(-1/\lambda^2)$$

   = 1/λ = 2500, as before.

     Exercise 29(e).   Compute the population average of the rv acidity.
Explain why the population mean, μ, is not equal to 3.00 [= (1+ 5)/2]

THE Population VARIANCE OF A CONTINUOUS RV

     As before, the population variance is the long-run (or weighted) average of deviations from the population mean (μ) squared, i.e., the 2nd central moment of X.

$$V(X) = \sigma_X^2 = \sigma^2 = E[(X - \mu)^2] = \int_{R_x} (x-\mu)^2 f(x)\,dx = ... = \int_{R_X} x^2 f(x)\,dx - \mu^2.$$

     Exercise 29(f).   Fill in the steps above by expanding the quadratic under the integral, verifying again that V(X) = σ^2 = E(X^2) – μ^2. Then compute the E(X^2) for the exponential density and use it to show that the variance of the exponential is equal to 1/λ^2, and hence the CV$_x$ = 100%. Hint:

In computing $E(X^2)$ you may use the fact that $\int_0^\infty x^2 \lambda e^{-\lambda x} dx = \lambda \dfrac{\partial^2}{\partial \lambda^2} \int_0^\infty e^{-\lambda x} dx$.

THE MOMENTS OF A RANDOM VARIABLE

Definition: Let X be a rv with the range space R_x and let c be any known real constant. Then the kth moment of X, $k \geq 0$, about the constant c is defined as

$$M_k(c) = E[(X - c)^k]. \qquad\qquad (12)$$

In the field of statistics only 2 values of c are of interest: $c = 0$ and $c = \mu$. Moments about $c = 0$ are called origin (or raw) moments and are denoted by μ'_k, i.e., $\mu'_k = E(X^k)$, where the constant $c = 0$ has been inserted into Eq. (12). Moments about the population mean, μ, are called central moments and are denoted by μ_k, i.e., $\mu_k = E[(X - \mu)^k]$, where $c = \mu$ has been inserted into (12).

STATISTICAL INTERPRETATION OF MOMENTS

By definition of the kth origin moment, or origin-moment of order k, we have:

$$\mu'_k = \begin{cases} E(X^k) = \displaystyle\sum_{R_X} x^k p(x) , & \text{if } X \text{ is a discrete rv} \\[4ex] E(X^k) = \displaystyle\int_{R_X} x^k f(x)\, dx , & \text{if } X \text{ is continuous rv} \end{cases}$$

(1) Whether X is discrete or continuous, $\mu'_1 = E(X) = \mu$, i.e., the 1st origin moment is simply the population mean which measures central tendency, and μ'_1 is the center of gravity of the entire population. Note that $\mu'_0 \equiv 1$.

(2) Since the population variance, σ^2, is the weighted average of deviations from the mean squared over all elements of R_x, then $\mu_2 = E[(X - \mu)^2] = \sigma_X^2$. Therefore, the 2nd central moment, $\mu_2 = \sigma^2$, is a measure of dispersion (or variation, or spread) of the population. Further, the 2nd central moment can be expressed in terms of origin moments using the binomial expansion of $(X - \mu)^2$, as shown below.

43

$$\mu_2 = E[(X-\mu)^2] = E[(X^2 - 2\mu X + \mu^2)] = E(X^2) - 2\mu E(X) + \mu^2$$

$$= E(X^2) - \mu^2 = \mu_2' - (\mu_1')^2 = \sigma_X^2 = \sigma^2 . \tag{13}$$

Example 25 (continued). For the exponential density, $f(x) = \lambda e^{-\lambda x}$, $\mu_2' = \int_0^\infty x^2 \lambda e^{-\lambda x} dx =$

$2!/\lambda^2$ and $\mu_1' = \int_0^\infty x \lambda e^{-\lambda x} dx = 1/\lambda$ so that equation (13) yields $\sigma^2 = V(x) = \mu_2 = 1/\lambda^2$. (Note that the

exponential pdf, to my knowledge, is the only Pearsonian statistical pdf with $CV_x = 100\%$.)

(3) The 3rd central moment, μ_3, is a measure of skewness (bear in mind that $\mu_3 \equiv 0$ for all

symmetrical distributions). If X is continuous, then

$$\mu_3 = E[(X-\mu)^3] = \int_{R_X} (x^3 - 3\mu x^2 + 3x\mu^2 - \mu^3) f(x) dx = \mu_3' - 3\mu_2'\mu + 2\mu^3 \tag{14}$$

For the exponential pdf, we have shown that $\mu = \mu_1' = 1/\lambda$, $\mu_2' = 2!/\lambda^2$ and you may verify that $\mu_3' =$

$3!/\lambda^3 = 6/\lambda^3$. Hence, substitution into (14) yields

$$\mu_3 = \frac{6}{\lambda^3} - \frac{2\times 3}{\lambda^2}(\frac{1}{\lambda}) + \frac{2}{\lambda^3} = \frac{2}{\lambda^3} .$$

In order to show that $\mu_3' = 3!/\lambda^3$, it will be easier to use the partial differentiation and the fact that $\mu_3' =$

$$\int_{R_X} x^3 \lambda e^{-\lambda x} dx = -\lambda \frac{\partial^3}{\partial \lambda^3} \int_0^\infty e^{-\lambda x} dx .$$

If X is measured in hours, the units of μ_3 are expressed in terms of hours^3. To obtain a unit-

less measure of asymmetry (for comparative purposes), we standardize μ_3 to obtain the coefficient of

skewness given below:

$$\alpha_3 = \mu_3/\sigma^3.$$

Note that most authors refer to α_3 simply as skewness. Furthermore, some authors (such as Sir

Maurice G. Kendall) use the notation $\gamma_1 = \sqrt{\beta_1}$ for $\alpha_3 = \mu_3/\sigma^3$, i.e., $\gamma_1 = \sqrt{\beta_1} = \alpha_3 = \mu_3/\sigma^3$. For

example, the value of coefficient of skewness (or simply skewness) for the exponential pdf is $\alpha_3 =$

$\mu_3/\sigma^3 = (2/\lambda^3)/\lambda^{-3} = 2.00$, which is unit-less. It is not very common for the value of $\alpha_3 = \mu_3/\sigma^3$ to lie outside the interval [$-$ 2, 2].

Karl Pearson (1857-1936) proposed the measure Sk = $(\mu - MO)/\sigma$, or $(\mu - x_{0.50})/\sigma$ to evaluate the asymmetry (or skewness) of a statistical distribution. When a population distribution is symmetrical, for certain $\alpha_3 = \mu_3/\sigma^3 \equiv 0$; further, for a symmetrical distribution the values of μ, $x_{0.50}$, and the Modal point (MO) all three coincide. When $\alpha_3 = \mu_3/\sigma^3 > 0$, i.e., the population distribution is positively skewed, then it turns out that invariably its $\mu > x_{0.50} > MO$, if MO exists. However, when a population statistical distribution is negatively skewed (i.e., $\alpha_3 < 0$), then always $\mu < x_{0.50} < MO$. For the exponential pdf, Sk = $(\mu - x_{0.50})/\sigma = 0.30685282$, while $\alpha_3 = 2.00$. The above rules regarding μ, $x_{0.50}$, MO do not always apply to data sets because of sampling variation.

Exercise 29(g). Compute the value of skewness, $\alpha_3 = \mu_3 / \sigma^3$, for the density function of the rv acidity and also the value of Sk.

(4) The 4th central moment is a measure of Kurtosis (abundance of Prs at the tails of a distribution and sometimes peaked-ness in the middle) and for a continuous rv is given by

$$\mu_4 = \int_{R_X} (x-\mu)^4 f(x)\,dx = \int_{R_X} (x^4 - 4x^3\mu + 6x^2\mu^2 - 4x\mu^3 + \mu^4)f(x)\,dx$$

This last expression after simplifying reduces to

$$\mu_4 = \mu_4' - 4\mu_3'\mu + 6\mu_2'\mu^2 - 3\mu^4 \tag{15}$$

For the exponential density function $\mu_4' = \int_{R_X} x^4 \lambda e^{-\lambda x}\,dx = \lambda \frac{\partial^4}{\partial\lambda^4} \int_0^\infty e^{-\lambda x}\,dx = \lambda \frac{\partial^4}{\partial\lambda^4}(1/\lambda) =$

$4!/\lambda^4 = 24/\lambda^4$. Recall that for the exponential density $\mu_3' = 3!/\lambda^3$, $\mu_2' = 2!/\lambda^2$ and $\mu = 1/\lambda$. Substituting these into equation (15), for the exponential density, we obtain $\mu_4 = 9/\lambda^4$.

Again to obtain a unit-less measure of kurtosis we standardize μ_4 to obtain a (coefficient) of kurtosis defined as

$$\alpha_4 = \mu_4/\sigma^4.$$

For the exponential density function, the above measure of kurtosis is given by $\alpha_4 = 9.00$. Most

statistical distributions have their α_4 values within the interval [1.1, 9]. However, statistical packages normalize the value of $\alpha_4 = \mu_4/\sigma^4$ by 3 and use the terminology kurtosis = $(\mu_4/\sigma^4) - 3 = \beta_4$ so that the exponential kurtosis is equal to $\beta_4 = 6$. Statistical distributions whose x values far from μ (or the center) have high Pr of occurrence are called heavy-tailed and thus have β_4 values far above zero.

Exercise 29 (h). Compute the value of Kurtosis for the rv Acidity, which was assumed to possess a triangular pdf. ANSWER: All Triangular distributions in the universe have a Kurtosis $\beta_4 = \alpha_4 - 3 = \mu_4/\sigma^4 - 3 = -0.6000$. (i) For the rv Acidity, obtain the value of the REL-IQR = IQR/σ .

(5) Recall from chapter 1 that the IQR = Q3 – Q1 = $x_{0.75} - x_{0.25}$ is another measure of variability. To obtain the 25th percentile (or the 0.25 quantile) of the exponential density, we set its cdf equal to 0.25 and solve for the corresponding value of x : $1 - e^{-\lambda Q1} = 0.25 \rightarrow$ Q1 = 0.2876821/ λ, and similarly Q3 = 1.38629436/λ. Hence, the value of the IQR for the exponential density is IQR = 1.0986123/λ. Therefore, all exponential density functions in the universe have a relative IQR equal to IQR/σ = REL-IQR = 1.0986123 = ln(3), (far from the normal value of 1.3490 discussed below). Note that Sir Maurice G. Kendall (pp. 85-86), one of 3 premier statisticians of the 20th century, unusually refers to kurtosis, $\beta_4 = \alpha_4 - 3 = \mu_4/\sigma^4 - 3$, as γ_2. Two other statisticians who were also knighted are Sir. Ronald A. Fisher and Sir D. R. Cox (who was the father of Stochastic Processes).

THE pdfs OF SOME WELL-KNOWN CONTINUOUS RANDOM VARIABLES

1. The Uniform (or Rectangular) pdf Denoted U(a,b)

A continuous rv, X, is said to be uniformly distributed over the closed interval [a, b], written as X ~ U(a, b), iff its pdf (Pr density function) is given by

$$f(x) = \begin{cases} 1/(b-a), & a \le x \le b \\ 0, & \text{elsewhere} \end{cases}.$$

Exercise 30. (a) Prove that the cdf of a U(a, b) rv, X, is given by F(x) = (x–a)/ (b–a), $a \le x \le$ b. (b) Show that the 1st four moments of of a U(a, b) rv are μ = (a + b)/2, μ_2 = V(X) = $(b - a)^2$ /12, α_3 = 0, α_4 = 1.80, and kurtosis $\beta_4 = -1.20$; this kurtosis shows that the uniform distribution is totally flat relative to a Bell-Shape curve. (c) Show that the value of the relative IQR (REL-IQR = IQR/σ) for

all uniform distributions is equal to $\sqrt{3}$. (d) Obtain the percentile (or quantile function) of the rectangular distribution, i.e., obtain the inverse function.

Example 26.   At the end of a week, the closing price of a stock, X, is uniformly distributed over [$35.75, $43.75], i.e., X ~ U(35.75, 43.75). Then f(x) = 1/8 and the cdf of X is F(x) = (x – 35.75)/ 8. (a) Compute the Pr that the stock price will close below $40.00 at the end of next week.

$$P(X < \$40.00) = F(40) = (40 – 35.75)/8 = 0.53125.$$

(b) Compute the Pr that the stock price will close in the interval [39, $42.00].

$$P(39 < X < 42) = F_X (42) – F_X (39) = (42 – 39)/8 = 0.375.$$

It can easily be verified that the inverse (or percentile) function of the Uniform, U(a, b), distribution is given by F_X^{-1} = a+(b–a)x, $0 \le x \le 1$, i.e., the quantile function is x_p = a + (b–a)p.

2. THE NORMAL (GAUSSIAN OR LAPLACE) DISTRIBUTION N(μ, σ^2)

Definition: A continuous rv, X, is said to be normally distributed with mean μ and variance σ^2 iff its pdf (Pr density function) is given by

$$f(x; \mu, \sigma) = C\,e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})2} = C{\times}Exp[–(x – \mu)^2/(2\sigma^2)], \qquad –\infty < x < \infty$$

where C is a constant and must be evaluated in such a manner that the total area under the above pfd is identically equal to 1 (i.e., 100% Pr), i.e., we must require that

$$I = \int_{-\infty}^{\infty} C\,e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})2}\,dx = 1. \tag{16}$$

Recall from your calculus that the function e^{-x^2} has no closed-form antiderivative, i.e., there does not exist a closed-form function such that when differentiated will exactly yield e^{-x^2}. However, there does exist an infinite McLaurin series such that when differentiated will yield e^{-x^2}. Therefore, the integrand in equation (16) has no closed-form single-function antiderivative and the improper integral in Eq. (16) cannot be carried out directly.

In order to evaluate the improper integral I in (16), we first make the change of variable z = (x –

$\mu)/\sigma$ and $y = (x - \mu)/\sigma$ [note that these changes of variables are called standardization), and then we

compute the value of I^2 as follows: $I^2 = C^2\sigma^2 \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} e^{-z^2/2} \times e^{-y^2/2} dz\,dy =$

$$C^2\sigma^2 \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} e^{-(z^2+y^2)/2}\, dz\,dy = C^2\sigma^2 \int\limits_{0}^{2\pi} \int\limits_{0}^{\infty} e^{-r^2/2} r\,dr\,d\theta \qquad (17)$$

where we have made the transformation to polar coordinates by letting $z = r\cos(\theta)$ and $y = r\sin(\theta)$ into equation (17) which shows that $C^2 = 1/(2\pi\sigma^2)$. Therefore, the only unique value of C that makes the value of the improper integral in (16) equal to 100% is $C = 1/(\sigma\sqrt{2\pi})$. These developments show that a continuous rv, X, is $N(\mu, \sigma^2)$ [read as normally, or Gaussian, or Laplace-Gaussian distributed with mean μ and variance σ^2) only if its pdf is given by

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} = \frac{1}{\sigma\sqrt{2\pi}} \text{Exp}[-(x - \mu)^2/(2\sigma^2)], \quad -\infty < x < \infty \qquad (18)$$

The graph of the pdf in (18) is symmetrical about μ and is provided in Figure 2. The reader must be cognizant of the fact that the Gauss-Laplace density in Eq. (18) was first discovered by the French Mathematician Abraham De Moivre in 1738; A. Moivre arrived at the density in (18) as the limiting distribution of the binomial Pr mass as $n \to \infty$ so that $\mu = np$ and $\sigma^2 = npq$. However, unfortunately De Moivre did not study the properties of the density in (18); Karl Pearson referred to it as the "Normal" around 1894. Both Laplace and Carl Gauss studied the properties of the density in (18) around the turn of the 19th Century.

Exercise 31. Show that the modal point of any $N(\mu, \sigma^2)$ distribution, as shown in Figure 2, occurs at $x = \mu$. Secondly, show that its points of inflection occur at $x = \mu \pm \sigma$.

It can be verified that the 1st four moments of any Gaussian distribution are $\mu_1' = \mu$, $V(X) =$

$$\mu_2 = \sigma^2, \quad \alpha_3 \equiv 0, \text{ and } \alpha_4 = \mu_4 / \sigma^4 = E(X - \mu)^4/ \sigma^4 = E\{[(X - \mu)/\sigma]^4\} = E(Z^4) = \int\limits_{-\infty}^{\infty} \frac{z^4}{\sqrt{2\pi}} e^{-z^2/2} = 3.00,$$

and hence its kurtosis, β_4, is identically equal to zero. Thus, all normal distributions in the universe have a skewness of $\alpha_3 \equiv 0$ and a kurtosis $\beta_4 = \alpha_4 - 3.00 \equiv 0.00$; Note that because $\beta_4 = 0$, the normal

distribution is not heavy-tailed. Curves with zero kurtosis are called mesokurtic; those with $\beta_4 < 0$ are called platykurtic (flat relative to the normal such as the uniform) and those with $\beta_4 > 0$ are called leptokurtic (i.e., heavier-tailed relative to the normal distribution and perhaps more peaked).

It must be emphasized that the normal distribution has unlimited applications in all facets of

$f'' < 0$ (Concave)

$f'' = 0$ (Inflection Point)

$f'' > 0$ (Convex)

σ

0.05

μ

$X_{0.95}$

X
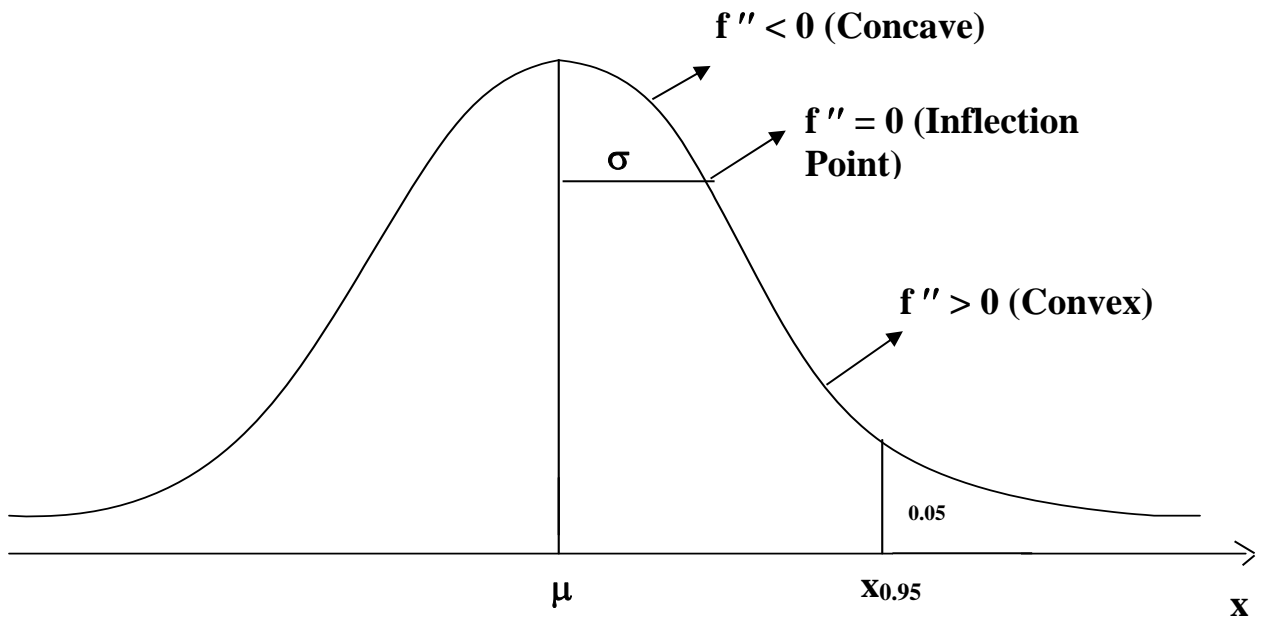
Figure 2

human life (for example, in all sciences, in all fields of engineering and manufacturing, in agriculture, medicine, pharmacy, etc.). In fact, the distribution of almost every dimension that Man manufactures can be modeled closely by a normal curve!   This implies that roughly 67% of all parent (or underlying) populations in the universe are assumed normally (or approximately normally) distributed with some mean μ and variance σ^2. Application examples will now follow.

Example 27.   The design tolerances (or consumers' specifications) on the length of steel pipes is 12 ± 0.125". The dimension or length of a pipe, X, is N(12, 0.0016 inches^2), as depicted in Figure 3. (a) Compute the Pr that a randomly selected pipe has length below 12.05". P(X < 12.05) =

$$P(X \leq 12.05) = \int_{-\infty}^{12.05} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \ , \text{ where } \mu = 12 \text{ and } \sigma = 0.04. \text{ To simplify the above}$$

integrand, let $(x - \mu)/\sigma = z$; then, $dz/dx = 1/\sigma$ or $dz = dx/\sigma$. Substituting the transformation (i.e.,

standardizing) $z = (x - \mu)/\sigma$ into the above integral results in:

$$P(X < 12.05) = P(X \leq 12.05) = \int_{-\infty}^{1.25} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{1.25} \phi(z)dz = \Phi(1.25).$$

The pdf $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ is called the standard (or unit) normal density function because its mean

is zero and its variance is equal to 1. Unfortunately, $\phi(z)$ has no closed-form simple antiderivative,

and therefore, its cdf is tabulated for different values of z in Table A.3, in Appendix A6 of Devore

(using numerical integration). The cdf of the standard normal density $N(0,1)$ is universally denoted by

$\Phi(z)$. For example, Table A.3, in Appendix A6 gives the Standard-Normal cdf only to 4 decimals;

while a spreadsheet can provide as many as 15 decimal for the cdf $\Phi(z)$. For example, $\Phi(1.5)$ =

0.93319, $\Phi(-1.5) = 0.06681$, $\Phi(1.959964) = 0.9750$, $\Phi(-1.96) \cong 0.025$, $\Phi(3) = 0.99865$, $\Phi(-3)$ =

0.00135 [Note that $\Phi(-z) = 1 - \Phi(z)$]. For part (a) of Example 27, we have: $P(X < 12.05) = F_X(12.05)$

$= P(Z \leq 1.25) = \Phi(1.25) = 0.89435$. The graph of $\phi(z)$ is shown in Figure 3.

Figure 3 graphs the
standard normal density
$\phi(z)$ versus z. Thus, the
ordinate represents $\phi(z)$,
while the abscissa is z.

Quality Specs:12 ± 0.125
Note that $\Phi(1.00) = 0.841345$

$f'' = 0$

1

$\Phi(1.25) = 0.89435$

0.10565

0

1.25

z
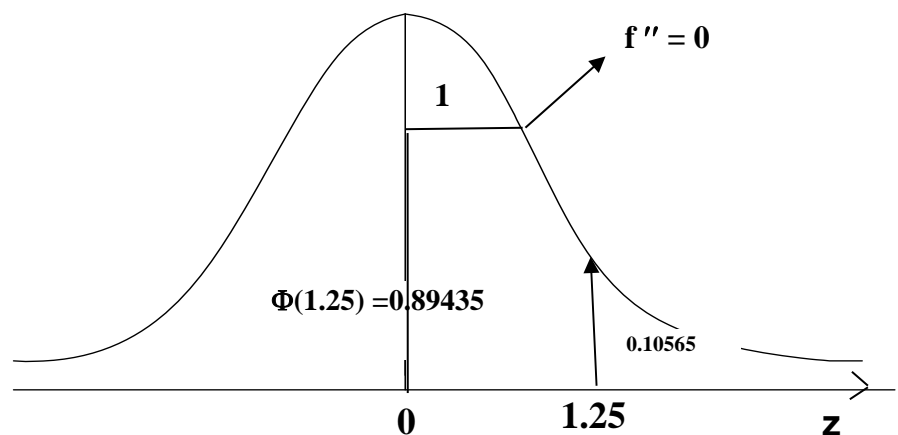
Figure 3. $\phi(z)$

(b) A pipe is selected at random. Compute the Pr that its length falls within 2σ of the mean

$\mu = 12$.

$$\Pr(|X - \mu| \leq 2\sigma) = P(11.92 \leq X \leq 12.08)$$

$$= \Pr(\frac{11.92 - 12}{0.04} \leq \frac{X - \mu}{\sigma} \leq \frac{12.08 - 12}{0.04}) = P(-2 \leq Z \leq 2) = \Phi(2) - \Phi(-2)$$

$= 0.97725 - 0.02275 = 0.95450$.

(c) Compute the Pr that the length of the pipe is within 3 sigma of μ.

$\Pr(|X - \mu| \leq 3\sigma) = \Pr(-3 \leq Z \leq 3) = \Phi(3) - \Phi(-3) = 0.99865 - 0.00135 = 0.99730$.

Finally, the $P(\mu - 4\sigma \leq X \leq \mu + 4\sigma) = \Phi(4) - \Phi(-4) = 0.99996833 - 0.000031671$

$= 0.99993666$. You may also use the Matlab's normcdf(z) to obtain $\Phi(4)$ and $\Phi(-4)$.

(d) Compute the fraction of the pipes that are not conforming to customer specifications LSL = 11.875" and USL = 12.125, i.e., compute the FNC (Fraction Nonconforming) of the process, denoted by p.

$p = FNC = P(X < 11.875) + P(X > 12.125) = P(Z \leq -3.125) + P(Z \geq 3.125)$

$= \Phi(-3.125) + [1 - \Phi(3.125)] = 2 \times \Phi(-3.125) = 2 \times 0.000889025 = 0.001778051$.

In fact, Chebyshev's Inequality states that for any frequency function, f(x), the **$\Pr(|X - \mu| > k\sigma) \leq$**
$1/k^2$ for all real constants $k \geq 1$.

Exercise 32. **(a)** Suppose that the process in Example 26 is not centered at $\mu = 12.00$. Compute the process FNC under the condition that $\mu = 12.03$". (ANS: p = 0.0088278). **(b)** Under the condition of part (a) compute the IQR and then the REL-IQR = IQR/σ and show that the value of REL-IQR, which is equal to 1.3489795004, is invariant for all De Moivre's distributions. **(c)** Under condition (a), obtain the value of σ such that the FNC is reduced back to 0.001778051.

THE PERCENTILES (or Inverse Function) of a NORMAL DISTRIBUTION

Definition: For the standard normal density, N(0, 1), the 95th percentile is denoted by $z_{0.05}$, and is defined such that the $P(Z \leq z_{0.05}) = 0.95$. Similarly, the 90th percentile of Z is defined as $P(Z \leq z_{0.10}) = 0.90$. This definition implies that for the standard normal density function (STNDF), $\Phi(z_{0.05}) =$

0.95 and $\Phi(z_{0.10}) = 0.90$. Because the normal cdf is not directly invertible, Table A.3 in Appendix A-6 of Devore(8e) must be used to invert the cdf in order to obtain the 95th percentile which is equal to $z_{0.05} = 1.6449$ while $z_{0.10} = 1.2816$. By symmetry of the STNDF, it follows that the 5th percentile of Z is equal to $z_{0.95} = -1.6449$ and the 10th percentile $z_{0.90} = -1.2816$, i.e., the $P(Z \leq -1.6449) = \Phi(-1.6449) = 0.05$ and the $P(Z \leq -1.2816) = \Phi(-1.2816) = 0.10$. It should be noted that $z_{0.05} = 1.6449$ is also called the 5 percentage point of the STNDF and $z_{0.10} = 1.2816$ is called the 10 percentage point of a STNDF. It turns out that in the field of Statistics, a percentage point, such as z_{α}, is almost always the distance from the origin beyond which the right tail Pr is exactly equal to α. Further, for the STNDF, it should be clear that $z_{\alpha} = -z_{1-\alpha}$ for all $0 < \alpha < 1$. Figure 3 shows that $z_{0.10565} = 1.250$ and hence $z_{0.89435} = -1.250$.

  If the rv, X, is normally distributed but either $\mu \neq 0$, or $\sigma \neq 1$ (or both), the 95th percentile of X is defined as $x_{0.95}$ such that the $P(X \leq x_{0.95}) = 0.95$, and similarly the 90th percentile of X is defined such that $P(X \leq x_{0.90}) = 0.90$. Furthermore, in general $x_{0.95} \neq - x_{0.05}$ or $x_p \neq -x_{1-p}$, while $z_{1-\alpha} = - z_{\alpha}$, where $Z \sim N(0, 1)$. The exact closed-form inverse function of the standard normal does not exist because $\Phi(z)$ is an integral. Matlab gives the inverse to 14 decimal accuracy using numerical analysis. However, a rough approximation to 2 decimals is given by $Z(p) \cong 4.91(p^{0.14} - q^{0.14})$, where $q = (1 - p)$ and $Z(p)$ represents the $100 \times p^{th}$ percentile (or the pth quantile) of $N(0, 1)$. For example, if $p = 0.10$, then $Z(0.10) \cong 4.91[0.10^{0.14} - (1 - 0.1)^{0.14}] = - 1.2811 \cong z_{0.10}$.

  Exercise 32(d).   For the $N(12.03, 0.0016)$ distribution, obtain the 5th, 10th, 20th, 50th, 80th, 90th, 95th, and 99th percentiles (or quantiles) of pipe length, X.

THE REPRODUCTIVE PROPERTY OF THE NORMAL DISTRIBUTION

  Suppose X_1, X_2, ..., X_n are n independently and normally (NID) distributed rvs with means μ_i and variances σ_i^2 . Further, c_1, c_2, ..., c_n are known or given constants. Then it has been proven in the theory of statistics that the linear combination (LC)

$$Y = c_1 X_1 + c_2 X_2 + ... + c_n X_n = \sum_{i=1}^{n} c_i X_i$$

is also normally distributed with mean $E(Y) = \sum_{i=1}^{n} c_i\mu_i$ and variance $V(Y) = \sum_{i=1}^{n} c_i^2\sigma_i^2$,

i.e., $Y \sim N(\sum_{i=1}^{n} c_i\mu_i , \sum_{i=1}^{n} c_i^2\sigma_i^2)$. Note that if X_i's are not independent, than $V(Y) = V(\sum_{i=1}^{n} c_iX_i)$

$= \sum_{i=1}^{n} c_i^2\sigma_i^2 + 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} c_ic_j\sigma_{ij}$, where $\sigma_{ij} =$ Covariance$(X_i, X_j) = E[(X_i-\mu_i)(X_j-\mu_j)]$, for all $i \neq j$.

Example 27 (continued, e).   Three steel pipes are selected at random from the N(12, 0.0016) population and placed end-to-end. Compute the Pr that the total length of the assembly exceeds 36.12". [n = 3, each c_i = 1, and each $X_i \sim$ N(12.00, 0.0016)].

  $Y = X_1 + X_2 + X_3 \rightarrow$ $E(Y) = 12 + 12 + 12 = 36.00$", & $V(Y) = 0.0016 + 0.0016 + 0.0016 =$

  $0.0048 \rightarrow$ Y $\sim$ N(36, 0.0048). $P(Y > 36.12) = P(\dfrac{Y-\mu_y}{\sigma_y} > \dfrac{36.12-36.00}{0.069282}) = P(Z > 1.732051) = 1 -$

$\Phi(1.732051) = 1 - 0.958368 = 0.0416323.$

Exercise 33.   Suppose 4 steel pipes are selected at random from the N(12, 0.0016) population and placed end-to-end. (a) Compute the Pr that the length of the assembly does not exceed 47.90". (b) Compute the Pr that the average length of the 4 pipes exceeds 12.04. ANS: 0.10565, (b) 0.02275.

THE CENTRAL LIMIT THEOREM (CLT)

If $X_1, X_2,..., X_n$ are independent, but not normal, rvs with $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$, then the sampling

distribution of $Y = X_1 + X_2 +...+ X_n$, as $n \rightarrow \infty$, approaches normality with $E(Y) = \sum_{i=1}^{n}\mu_i$ and $V(Y) =$

$\sum_{i=1}^{n}\sigma_i^2$. Note that the more skewed the frequency functions of the individual X_i's in the sum $\sum_{i=1}^{n}X_i$ are,

the more slowly (in terms of n) the distribution of Y approaches normality. For example, if the X_i's are uniformly distributed, then for n > 8, the distribution of the sum, Y, is approximately normal because α_3

= 0, while if X_i's are exponentially distributed (α_3 = 2.0 and thus highly skewed), then a value of n > 200 is needed.

In case X_i's have the same mean and variance, then the distribution of Y approaches normality with E(Y) = $n\mu$ and V(Y) = $n\sigma^2$. As a general conservative rule (of thumb) a value of n > $25\alpha_3^2$ and n > $5\alpha_4$ (in most applications at least n > 50) is needed in order for the distribution of the sum Y =

$$\sum_{i=1}^{n} c_i X_i$$ to become approximately Gaussian, where c_i's are constants. Further if the underlying

distribution is symmetric ($\alpha_3 \equiv 0$), then the needed sample size can be relaxed to n_{Min} > $4\alpha_4$.

Example 28. A large freight elevator can transport a maximum of 10,000 pounds. A load of cargo containing 50 boxes of a certain product is transported on the elevator every hour. The weights of the boxes are independent (not normal but identically distributed) each with E(X_i) = 190lbs and V(X_i) = 900 lbs^2. Estimate (using the CLT) the Pr that the maximum limit is exceeded on a given hour and disrupts the transporting process? <u>Solution:</u>   Y = a Cargo Load = $\sum_{i=1}^{50} X_i$ → E(Y) = 50 ×190 =

9500 and V(Y) = 50×900 = 45,000.00. Figure 4 illustrates how to compute the Pr of process

disruption. P(Y > 10,000 lbs) = P(Z > $\frac{10000 - 9500}{212.132}$) = P(Z > 2.35702) = 1 − Φ(2.357023) = 1 −

0.990789 = 0.0092111.

Exercise 34. (a) For the Example 28 above, determine the maximum value of n such that the disruption Pr is reduced from 0.0092111 to 0.001. (b) If n = 50 boxes is kept fixed, how small should V(X_i) be so that the P(Y > 10,000 lbs) is reduced to 0.001?

The Normal Approximation to the Binomial

Recall that a binomial rv, X, is the sum of n independent Bernoulli rvs, i.e., X_{Bin} = $\sum_{i=1}^{n} X_i$, where

each X_i has a Bernoulli distribution with E(X_i) = p and V(X_i) = pq. As a result of the CLT, as n → ∞,
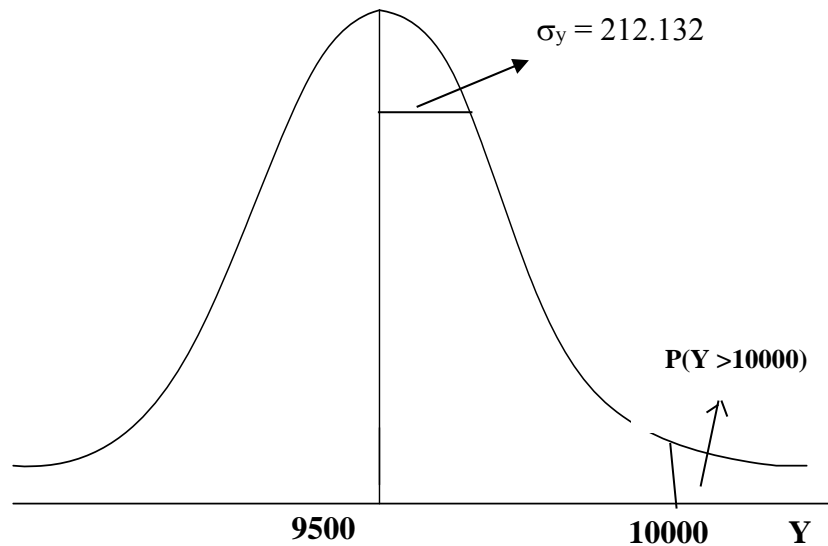
Figure 4. The approximate pdf of Y

the binomial pmf can be approximated by a normal pdf with E(X) = np and V(X) = npq. However, because a binomial rv has a discrete range space, a correction for continuity has to be applied as will be illustrated in the following example. The approximation is adequate when n is large enough such that np > 10 and 0.10 < p < 0.90.   If np < 10, I would recommend the Poisson approximation to a Binomial. The values of 10 ≤ np ≤ 15 form a gray region for which is not clear-cut as to which of the two distributions, Poisson or the normal, provide better approximation to the binomial. Recall that a general rule of thumb was for n to be large enough such that n > $25\alpha_3^2$ and simultaneously n > $5\alpha_4$, where for a single Bernoulli trial α_3^2 = (q − p)^2/ (p×q), and α_4 = (1 − 3 p + 3 p^2)/(p×q) = (1 − 3p q)/(p×q). It can be shown that for a Bin(n, p) distribution the value of skewness is α_3 = (q − p)/$\sqrt{npq}$

and the binomial kurtosis is given by α_4 − 3 = $\dfrac{1+3(n-2)pq}{npq}$ − 3 = $\dfrac{1-6pq}{npq}$.

Example 29. Consider a binomial distribution with n = 50 trials and p = 0.30 (so that μ = np = 15).   Note that we would like to have np > 15, but $25\alpha_3^2$ = 19.05 implies that the requirement on

55

skewness n = 50 > $25\alpha_3^2$ is easily satisfied. Then the range space of the discrete rv, X_D, is R_x = {0, 1, 2, 3, ..., 50}. The exact Pr of attaining exactly 12 successes in 50 trials is given by b(12; 50, 0.30) = $P(X_D = 12) = {}_{50}C_{12}(0.30)^{12}(0.70)^{38}$ = 0.08383.

To apply the normal distribution (with μ = 15 and σ^2 = npq =10.50) to approximate the $P(X_D = 12)$, we 1st have to select an interval on a continuous scale, X_C, to represent X_D = 12. Clearly this continuous interval on X_C has to be (11.5, 12.5), i.e., we will have a correction of 0.50 for continuity. In short, X_D = 12 $\cong$ (11.5, 12.5)$_C$. Figure 5 illustrates the normal approximation to the b(x; 50, 0.30) distribution. From Figure 5, $P(X_D = 12) \cong P(11.5 \le X_C \le 12.5) = P(\frac{11.5-15}{3.2404} \le Z \le \frac{12.5-15}{3.2404}) =$

$P(-1.0801 \le Z \le -0.77152) = \Phi(-0.77152) - \Phi(-1.0801) = 0.220200 - 0.140044 = 0.080157.$
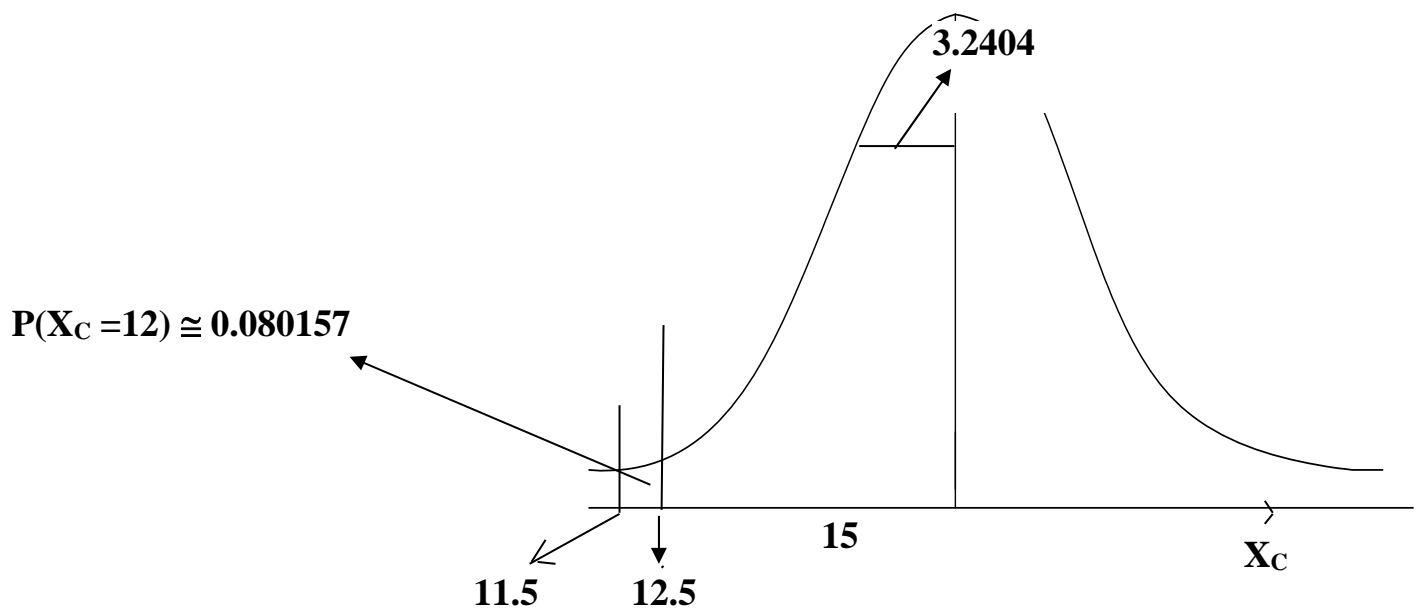
Figure 5. The Normal Approximation to Binomial

Since the exact Pr was 0.08383, the percent relative error in the above approximation is

$$\%\text{Relative Error} = (\frac{0.080157 - 0.08383}{0.08383}) \times 100 = -4.3815\%.$$

Next we compute the exact Pr that X exceeds 12, i.e., $P(X_D > 12) = 1 - B(12; 50, 0.30) =$

$1 - \sum_{x=0}^{12} {}_{50}C_x (0.30)^x (0.70)^{50-x}$ = 0.777134. The normal approximation to the above binomial Pr is

$P(X_C \geq 12.5) = P(Z \geq -0.77152) = 0.77980 \cong P(X_D > 12)$. The % error in this approximation is 0.343%.

As stated earlier, the normal approximation to the binomial should improve as n increases and $p \rightarrow 0.50$. To illustrate this fact, consider a binomial distribution with parameters n = 65 and p = 0.40. Then, μ = np = 26, σ = $(npq)^{1/2}$ = 3.9497, and $P(X_D = 22) = {}_{65}C_{22} (0.40)^{22}(0.60)^{43}$ = 0.0617. The normal approximation to the above b(22; 65, 0.40) is $P(21.5 \leq X_C \leq 22.5) = P(-1.13933 \leq Z \leq -0.88615)$ = 0.18777 − 0.12728 = 0.0605 with relative % error of 1.935%. Further, $P(X_D < 23)$ = B(22; 65, 0.40) = $\sum_{x=0}^{22} {}_{65}C_x (0.40)^x (0.60)^{65-x}$ = 0.188327. The normal approximation to this Binomial cdf corrected for continuity (cfc) is $P(X_c \leq 22.5) = P(Z \leq -0.88615) = \Phi(-0.88615)$ = 0.1877692 with a % relative error of − 0.2962%.

Exercise 35.   Work Exercises 32, 34, 37, 39, 42, 46, 48, and 52 on pages 162-164 of Devore(8e).

3. The Gamma (Pearson's Type III) pdf and its Related Distributions

A continuous rv, X, has a Gamma pdf at the rate λ > 0 iff its pdf is given by

$$f(x; \alpha, \lambda) = \begin{cases} \dfrac{\lambda}{\Gamma(\alpha)} (\lambda x)^{\alpha-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{19}$$

where X represents the time (measured from zero). When α = n is a positive integer, then X represents time to the occurrence of the nth Poisson event, and $\Gamma(n)$ = (n −1)! if n is a positive integer. More detailed definition of $\Gamma(n)$ will be forthcoming and these concepts will be more clear to you as you study these notes and go thru the examples. For the time being it is necessary to know that 0! = 1 (not zero). Furthermore, it is assumed that the rate of occurrence of the Poisson event does not change during the interval of length t that the model (19) is applicable to, i.e., $\lambda_1 = \lambda_2 = ... = \lambda_n = \lambda$. Note that Devore uses $1/\beta$ in (19) for λ in his Gamma pdf on page 168; λ is the-rate parameter, $1/\lambda = \beta$ is called the scale, while α is called the shape-parameter. When n is a positive

integer, the function in (19) is also referred to as the Erlang pdf. Further, when $\lambda =1$, the density function in Eq. (19) is called the Standard gamma distribution.

As an example to clarify the meanings of λ and n, consider an aircraft that is guided by 3 computers (i.e., n = 3), but the on-line computer and the two in standby mode have the same failure rates $\lambda_1 = \lambda_2 = \lambda_3 = \lambda = 0.0004$/hour, i.e., the failures rates are identical so that $\lambda = 0.0004$ for all 3 guiding computers. If λ_i's (i = 1, 2, ..., n) are not equal, then model (19) will not be applicable and the reliability of a 10-hour flight would have to be computed using other means taught in a reliability engineering course.

The First family Member of the Gamma pdf

Consider the specific gamma pdf for $\alpha = 1$, i.e., $f(x; 1, \lambda) = \lambda e^{-\lambda x}$, which is the exponential density function with rate parameter $\lambda > 0$, where $\alpha =1$ has been inserted into (19). To comprehend the concepts hereafter, you must recall that the cdf of the exponential density is given by $F(x) = 1 - e^{-\lambda x}$, $0 \le x < \infty$ so that the inverse function is given by $F^{-1}(p) = \frac{1}{\lambda} \ln[1/(1-p)]$, $0 \le p < 1$. The exponential pdf has widespread and enormous applications in reliability engineering (or life testing) and also in queuing theory. Majority of its applications occur because of its relation to the Poisson distribution as shown below.

Consider a Poisson event that occurs at a rate of λ per unit of time (such as the number of customers arriving per hour, or per minute, at a service center). Then the average number of occurrences during an interval of length t is λt. Let X(t) represent the number of Poisson events occurring during an interval of length t, i.e., the range space of X(t) is $R_{X(t)}$ = {0, 1, 2, 3, ...}. Then from the Poisson pmf

$$P[X(t) = x] = \frac{(\lambda t)^x}{x!} e^{-\lambda t} \quad \rightarrow \quad P[X(t) = 0] = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}.$$

Next, define a continuous rv, T, as follows:

T = the time between the occurrences of 2 successive Poisson events (or the inter-occurrence or intervening time of 2 successive Poisson events). Then the following two events are equivalent.

$$[X(t) = 0] = [T > t].$$

58

Thus,

$$P(T > t) = P[X(t) = 0] = e^{-\lambda t} = 1 - P(T \le t) = 1 - F(t) \rightarrow F(t) = 1 - e^{-\lambda t}$$

$$\rightarrow \quad f(t) = dF(t)/dt = \lambda e^{-\lambda t}$$

The above developments show that if the number of occurrences of an event is Poisson distributed, then the occurrence time, T, to the next Poisson event measured from the last occurrence is exponentially distributed.

Example 30. The number of downtimes of a computer network is Poisson distributed with an average of 0.20 downtime/week. (a) Compute the Pr of exactly 2 failures next week.

$$p(2; 0.20) = \frac{(0.20)^2}{2!} \times e^{-0.20} = 0.0164$$

(b) Compute the Pr that the time between 2 future successive downtimes exceeds one week.

$$P(T > 1 \text{ week}) = P[X(1 \text{ week}) = 0] = e^{-0.20} = 0.81873.$$

(c) Compute the Pr that exactly 2 failures occur in the next 3 weeks.

Y = number of failures/3weeks $\rightarrow \mu = \lambda t = (0.20) \times 3 = 0.60$

$$P(Y = 2 \text{ failures}) = \frac{(0.60)^2}{2!} e^{-0.60} = 0.09879.$$

(d) Compute the Pr that the time between 2 successive failures exceeds 3 weeks.

$$P(T > 3) = P[X(3 \text{ weeks}) = Y = 0] = e^{-0.60} = 0.548812.$$

Exercise 36. For the above computer network (with failure rate $\lambda = 0.20$ per week), compute the Pr that the time between the next 2 failures will be shorter than 2 weeks. (b) Compute the Pr that there will be exactly 4 failures in the next 6 weeks. At most 6 failures in the next 6 weeks. (c) Compute the Pr that the time (since the last failure) to the next downtime will be less than 4 weeks. ANS: (a) 0.32968, (b) 0.02602, 0.99975, (c) 0.550671.

The Gamma Function

Let α be a positive number. Then the Gamma of $\alpha > 0$, denoted $\Gamma(\alpha)$, is defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \qquad\qquad (20)$$

If the upper limit in the improper integral in Eq. (20) is not ∞, then the definite integral $\int_0^x y^{n-1} e^{-y} dy$

$(x \neq \infty)$ is called the Incomplete Gamma Function (IGF), whose values divided by $\Gamma(\alpha)$ are given in Table A.4 in Appendix A-8 of Devore for α = 1(1)10, and x = 1(1)15. However, the use of Table A.4 is unnecessary because MS Excel provides the values of the IGF/$\Gamma(n)$ as GAMMADIST(x, α = n, β = 1, true). Using Excel, GAMMADIST(4, α = 6, β = 1, true) = 0.214869603 so that the IGF(4; 6) = 25.78435236. The equivalent Matlab function for the IGF is gamcdf(x, n) = gamcdf(4, 6) =

$$0.214869613 = (\int_0^4 y^{n-1} e^{-y} dy)/\Gamma(6).$$

Integrating the infinite integral in (20) by parts $(u = x^{n-1}, dv = e^{-x} dx)$ yields:

$$\Gamma(\alpha) = \left[-x^{\alpha-1} e^{-x} \right]_0^{\infty} + (\alpha - 1)\int_0^{\infty} x^{(\alpha-1)-1} e^{-x} dx = 0 + (\alpha - 1)\,\Gamma(\alpha - 1).$$

$$\text{Or: } \Gamma(\alpha + 1) = \alpha\,\Gamma(\alpha) \tag{21}$$

Repeated integration of (20) by parts results in

$$\Gamma(\alpha) = (\alpha - 1)(\alpha - 2)(\alpha - 3)...(2)(1)\Gamma(1), \tag{22}$$

where (22) is valid only in the case when α is an exact positive integer; otherwise integration by parts

will not terminate. Since $\Gamma(1) = 0! = \int_0^{\infty} e^{-x} dx \equiv 1$, Eq. (22) shows that if n is a positive integer, then

$\Gamma(n) = (n - 1)!$ Further, it can be proven that $\Gamma(1/2) = \sqrt{\pi}$. In order to compute $\Gamma(\alpha)$ using MS Excel, use the Excel function gamma(α) for any α > 0. If α = 3/2, from Eq. (21), $\Gamma(\alpha + 1) = \alpha\,\Gamma(\alpha)$ so that

$\Gamma(3/2) = \Gamma(\frac{1}{2} + 1) = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{1}{2}\sqrt{\pi}$. Similarly, $\Gamma(5/2) = \Gamma(3/2 + 1) = \frac{3}{2}\Gamma(3/2) = 3\sqrt{\pi}/4$. You should

verify for yourself that $\Gamma(7/2) = 15\sqrt{\pi}/8$, etc. Note that if we divide both sides of Eq. (20) by $\Gamma(\alpha)$,

we obtain $1 \equiv \int_0^{\infty} [x^{\alpha-1} e^{-x}/\Gamma(\alpha)]dx$, and hence the integrand is a pdf over the interval [0, ∞). This

integrand $x^{\alpha-1} e^{-x}/\Gamma(\alpha)$ is called the Standard form of Gamma density, where α is the shape

parameter.

The Relationship Between the Gamma Function and the Gamma pdf

In the expression for the gamma function $\Gamma(\alpha)$, let $x = \lambda t \rightarrow dx = \lambda dt$. Substituting the transformation $x = \lambda t$ into Eq. (20) yields

$$\Gamma(\alpha) = \int_0^\infty (\lambda t)^{\alpha-1} e^{-\lambda t} \lambda \, dt \quad \rightarrow \quad 1 \equiv \int_0^\infty \frac{\lambda}{\Gamma(\alpha)} (\lambda t)^{\alpha-1} e^{-\lambda t} \, dt \ , \qquad (23)$$

which implies that the integrand of the 2nd integral must be a density function over the range $R_x = [0,$ $\infty)$. However, the integrand in (23), namely $\frac{\lambda}{\Gamma(\alpha)} (\lambda t)^{\alpha-1} e^{-\lambda t}$, is simply the Gamma pdf, $f(x; \alpha, \lambda)$, given in equation (19). Note that Devore uses the notation $1/\beta$ for λ; the rate parameter $\lambda = 1/\beta$ is scale because for a given α it determine the STDEV of T.

The Erlang pdf as the 2nd Family Member of the Gamma pdf (n is a positive integer and n > 1)

The application of gamma pdf, when n is a positive integer, occurs often in reliability engineering, where $T_n = X_1 + X_2 + ... + X_n$ represents the lifetime (or time to failure = TTF) of an n-unit standby system (i.e., units 2 thru n are cold spares). Unit 1 is energized at time 0 and its TTF is X_1. ASA unit 1 fails, unit 2 is energized (by means of a perfect switch) and its TTF measured from last failure is X_2; when unit 2 fails, unit 3 is put on-line with lifetime X_3, etc. When the value of n in the Gamma pdf (19) is a positive integer, then the pdf in (19) is called Erlang. The Erlang pdf has also widespread applications in queuing, where T_n represents the arrival time (measured from zero) of the nth customer to a queuing process.

It can be proven that if X_i's (i = 1, 2,..., n) are exponentially and independently distributed each with the same identical rate parameter λ, then the system lifetime, $T_n = \sum_{i=1}^{n} X_i$, is a rv having an Erlang pdf with parameters n (a positive integer) and rate λ as given in equation (19). As a result, $E(T_n) =$

$$\sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} (1/\lambda) = n/\lambda, \text{ and similarly } V(T_n) = n/\lambda^2 \text{ so that } \sigma_{T_n} = n^{1/2}/\lambda.$$ You should observe now

why $1/\lambda$ is called the scale parameter because it determines the standard deviation. It can be shown

that the 3rd and the 4th standardized moments of the Gamma pdf are given by $\alpha_3 = 2/\sqrt{n}$ and $\alpha_4 =$

$3 + 6/n$ so that the kurtosis of the Gamma density is equal to $\beta_4 = 6/n$. These last 2 moments show

that as $n \rightarrow \infty$, the distribution of T_n approaches normality because the 3rd and 4th standardized

moments of a normal distribution are $\alpha_3 = 0$ and $\alpha_4 = E(Z^4) = 3$. The parameter n is called the shape

parameter because it determines the skewness and kurtosis, implying that the shape of the Gamma

density changes as α (or n) change. Further, it can be shown that the model point occurs at MO =

$(n-1)/\lambda$, and thus we will have MO < $x_{0.50}$ < $E(T_n) = n/\lambda$ because $\alpha_3 = 2/\sqrt{n} > 0$ for all n > 0.

The reliability function for an n-identical-unit standby system is given by [T_n = time to the nth

failure, X(t) = No. of failures during t]

R(t) = P[(Time to the nth failure) > t] =P(T_n > t) = P[X(t) $\leq$ n – 1] =

$$= \sum_{x=0}^{n-1} \frac{(\lambda t)^x}{x!} e^{-\lambda t} = \int_{t}^{\infty} \frac{\lambda}{\Gamma(n)} (\lambda x)^{n-1} e^{-\lambda x} \, dx.$$

Therefore, the cdf of the Erlang rv, T_n, at time t is given by

$$F(t; n, \lambda) = P(T_n \leq t) = 1 - P(T_n > t) = 1 - \sum_{x=0}^{n-1} \frac{(\lambda t)^x}{x!} e^{-\lambda t} \qquad (24)$$

Example 31. Consider a 3-PC standby system that guides an aircraft each PC with a failure

rate of λ = 0.0004/hour. Compute the system reliability for a mission of 10-hour flight. [T_3 = Time to

the 3rd failure]. R(10) = P($T_3 \geq$ 10 hours) = P[No. of failures during 10 hours $\leq$ 2] = P[X(10) $\leq$ 2] =

$$\sum_{x=0}^{2} (0.004)^x e^{-0.004} / x! = 0.9^7893653 (= 0.9999999893653).$$ Note that X(10) denotes the number of

failures occurring during 10 hours, which has a Poisson pmf with mean $\mu = \lambda t$ = 0.004 failures per 10

hours. The system MTTF = $E(T_3)$ = n/λ = 3/0.0004 = 7500 hours, and $V(T_3) = \sigma^2 = n/\lambda^2$ = 18.75x10^6

hours^2 $\rightarrow \sigma$ = 4330.127 $\rightarrow$ CV(T_3) = 57.735%.

Exercise 37.   Consider a 4-unit standby system with failure rate $\lambda = 0.0002$/hour for each unit. Compute system reliability for a mission of length 200 hours. (b) Compute the CV(T_4). (c) Show that the modal point of the Gamma pdf occurs at MO = $(n - 1)/\lambda$. (d) I have not been able to obtain a general expression for the REL-IQR of the Gamma pdf for any n because the inverse function of the Gamma density (for n > 1) requires the inversion of an integral or a summation. Obtain the value the REL-IQR for n = 2. For n = 10, IQR/σ = 1.3243, at n = 20, IQR/σ = 1.3367, and at n = 40 the REL-IQR = 1.3428. It seems that as n → ∞, the Gamma REL-IQR approaches that of the normal IQR/σ = 1.3489795004. In fact, at n = 10^9, the Gamma REL-IQR = 1.348979500142027. Finally, it can be shown that $\Gamma(n) = \theta^n L\{t^{n-1}; \theta\}$, where L stands for the Laplace-transform with respect to θ.

The Chi-Square(d), χ^2, Distribution as the Last Family Member of the Gamma

Consider the general Gamma pdf in equation (19) specifically for the scale parameter $\lambda = 1/2$ and shape parameter n = $\nu/2$. This leads to a (central) χ^2 pdf with ν degrees of freedom (df). The pdf of χ^2_ν with ν df at a specific value w in its range space [0, ∞) is given below.

$$f(w; \nu) = \begin{cases} \dfrac{w^{(\nu/2)-1}}{2^{\nu/2}\,\Gamma(\nu/2)}\, e^{-w/2} & , \quad 0 \le w < \infty \\ 0, \quad \text{elsewhere} \end{cases} \tag{25}$$

The application of Chi-square distribution will be discussed in detail in chapters 6 and 7, where sampling distributions from a normal population are discussed. It will be shown that the sampling distribution of the statistic $(n - 1)S^2/\sigma^2$, from a normal parent population with variance σ^2, follows a χ^2 with df $\nu = n - 1$. The 1st 4 moments of χ^2_ν are $\mu'_1 = \nu$, $\sigma^2 = 2\nu$, $\alpha_3 = (8/\nu)^{1/2}$, and $\beta_4 = 12/\nu$.

4. THE WEIBULL DISTRIBUTION

Let T represent the TTF of any device (mechanical, electrical, or otherwise). Then its lifetime,

T, is said to have a Weibull pdf with minimum life $\delta > 0$ and characteristic life $\theta > \delta$ iff its pdf is given by

$$f(t; \delta, \theta, \beta) = \frac{\beta}{\theta - \delta} \left(\frac{t-\delta}{\theta-\delta}\right)^{\beta-1} e^{-\left(\frac{t-\delta}{\theta-\delta}\right)^{\beta}}, \quad t \geq \delta \tag{26}$$

where $\beta > 0$ is called the shape parameter (or the slope) of the Weibull. Note that the Weibull pdf in (26) at the slope $\beta = 1$ reduces to the exponential density with minimum life δ and failure rate $\lambda = 1/(\theta - \delta)$. To verify that the function in (26) is a pdf, consider an exponential pdf at a rate $\lambda = 1$ failure per unit of time, which is given by $f(x; \lambda = 1) = e^{-x}$, $x \geq 0$. Since e^{-x} is a pdf over the range [0,

∞), then clearly $\int_{0}^{\infty} e^{-x} dx = 1.$ $\tag{27}$

Now make the transformation $x = \left(\frac{t-\delta}{\theta-\delta}\right)^{\beta}$ in equation (27). Because the dx/dt $= \frac{\beta}{\theta-\delta}\left(\frac{t-\delta}{\theta-\delta}\right)^{\beta-1}$,

then the substitution of $x = \left(\frac{t-\delta}{\theta-\delta}\right)^{\beta}$ and dx $= \frac{\beta}{\theta-\delta}\left(\frac{t-\delta}{\theta-\delta}\right)^{\beta-1}$ dt into (27) yields

$$\int_{\delta}^{\infty} \frac{\beta}{\theta-\delta}\left(\frac{t-\delta}{\theta-\delta}\right)^{\beta-1} e^{-\left(\frac{t-\delta}{\theta-\delta}\right)^{\beta}} dt = 1 \tag{28}$$

In equation (28) the integrand has to be a Pr density function over the range [δ, ∞) because the value of the integral is identical to 1, but the integrand is simply the Weibull pdf given in equation (26). Note that Devore discusses the Weibull pdf on pp. 171-174 only for the case of minimum life $\delta = 0$ and uses somewhat uncommon notations α for the slope and β for the characteristic life. Further, δ is called the location parameter, θ is called the scale parameter, while $\beta > 0$ is called the shape parameter. It can be verified that the 1st four moments of the Weibull density in (26) are given by

$$E(T) = \delta + (\theta - \delta)\Gamma\left(\frac{1}{\beta} + 1\right) = \mu = \mu_1'$$

$$V(T) = \theta^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \left(\Gamma\left(1 + \frac{1}{\beta}\right)\right)^2\right] = \theta^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right] = \mu_2$$

Note that σ_T heavily depends on θ and hence θ is the scale-parameter.

$$-\infty < \alpha_3 = \frac{\Gamma(1+\frac{3}{\beta}) - 3\,\Gamma(1+\frac{2}{\beta})\,\Gamma(1+\frac{1}{\beta}) + 2\,\Gamma^3(1+\frac{1}{\beta})}{[\Gamma(1+\frac{2}{\beta}) - \Gamma^2(1+\frac{1}{\beta})]^{3/2}} = \mu_3/\sigma^3 < \infty$$

and

$$2.71050 < \alpha_4 = \frac{\Gamma(1+\frac{4}{\beta}) - 4\,\Gamma(1+\frac{3}{\beta})\,\Gamma(1+\frac{1}{\beta}) + 6\,\Gamma(1+\frac{2}{\beta})\Gamma^2(1+\frac{1}{\beta}) - 3\,\Gamma^4(1+\frac{1}{\beta})}{[\Gamma(1+\frac{2}{\beta}) - \Gamma^2(1+\frac{1}{\beta})]^2} = \mu_4/\sigma^4 < \infty$$

You may verify that the reliability function of a Weibull rv is given by $R(t) = 1$ for $t \le \delta$, $R(t) =$

$e^{-(\frac{t-\delta}{\theta-\delta})^\beta}$, for $\delta \le t < \infty$, and its cdf (i.e., the Pr that the device will fail by time t) is given by

$$F(t) = 1 - e^{-(\frac{t-\delta}{\theta-\delta})^\beta}, \qquad \delta \le t < \infty. \tag{29}$$

Example 32. The miles to failure of a car's water pump has a Weibull distribution with minimum life $\delta = 30{,}000$ miles, characteristic life $\theta = 65{,}000$ miles and slope $\beta = 4$. You buy such a brand new car. What is the reliability of its water pump at 50,000 miles, using equation (29).

$R(50{,}000) = P(T > 50{,}000) = e^{-(\frac{50000-30000}{65000-30000})^4} = e^{-0.10662224} = 0.898865 \rightarrow$ Its failure Pr is equal to $F(50{,}000) = 0.101135$. Therefore, if 100,000 such cars are sold, the expected number of water pump failures by 50,000 miles is 10113.50.

It can easily be shown that the $100{\times}p$th percentile $(0 < p < 1)$ of the Weibull density function (i.e., the inverse or quantile function) is given by

$$x_p = \delta + (\theta - \delta)\left[\ln\left(\frac{1}{1-p}\right)\right]^{1/\beta} \tag{30}$$

and thus the IQR = $(\theta - \delta)[(\ln 4)^{1/\beta} - (\ln(4/3))^{1/\beta}]$.

The Weibull distribution can also be used to model the strength (such as tensile, or breaking strength) of a mechanical component, or engine emissions of various pollutants.
The Example 4.25 on page 173 of Devore's 8th edition illustrates one of these applications.

Exercise 38. Compute the first 4 moments of a Weibull rv with population parameters of the

Example 31 above and compare against those of the corresponding normal distribution. I have verified that the Weibull distribution is positively skewed for $\beta \leq 3.602349425718944$ and becomes negatively skewed if $\beta > 3.602349425718944$. I wrote a Matlab program to arrive at $\alpha_3 \cong 0$ at $\beta = 3.602349425718944$ and its corresponding kurtosis is equal to -0.283138917922285. Those of you who enjoy programming, should try to check on my answers! You may also verify that the Weibull Kurtosis (= $\alpha_4 - 3$) ranges widely from $1.378480203794640 \times 10^{11}$ at $\beta = 0.10$ to Kurtosis = $\alpha_4 - 3 = 2.125445834920493$ at $\beta = 100$. The values of Weibull Kurtosis at $\beta = 5.772781214230558$is and $\beta = 2.2520007003$ are equal to zero, while the corresponding skewness values are equal to -0.349229679644149 and 0.480129499304545, respectively. Further, the relationships between α_3 versus β and Kurtosis versus β are not at all linear. (b) Use a normal distribution to approximate the reliability of the water pump at 50,000 miles. (c) Compute the values of median life and the 10% life. Further, Show that all Weibull rvs in the universe have a reliability equal to 0.36788 at their characteristic life θ, i.e., $R(\theta) = e^{-1}$.

Devore covers two other statistical distributions, the Lognormal and Beta (Pearson's Type I), but we will not be using these two pdfs in chapters on statistical inference. Hence, I will not discuss them herein. The interested student should study pp. 174-177 of Devore or refer to my notes in the Reliability Engineering section of my website.

The Weibull Skewness can range within the interval $-\infty < \alpha_3 = \mu_3/\sigma^3 < \infty$ and is generally a decreasing function of the shape-parameter (or slope) $\beta > 0$. While, the Weibull kurtosis (W-kurtosis) ranges from -0.2895 to $+\infty$, i.e., $-0.2895 < $ W-kurtosis $< \infty$. The graphs of Weibull skewness and Kurtosis as a function of the shape parameter β, over the range $0.20 \leq \beta \leq 10$ & 100, respectively, are given below.

Exercise 38(d). Work exercises 66, 67, 68, and 70 on pages 170-1 of Devore. (e) Study pp. 178-187 of Devore on Probability Plots. (f) Show that the modal point of a Weibull rv is given by
$$MO = \delta + (\theta - \delta)[(\beta - 1)/\beta]^{1/\beta}$$

Further, its median is given by $x_{0.50} = \delta + (\theta - \delta) \times [0.69314718056]^{1/\beta}$.

The Weibull Kurtosis Versus the Shape Parameter Beta
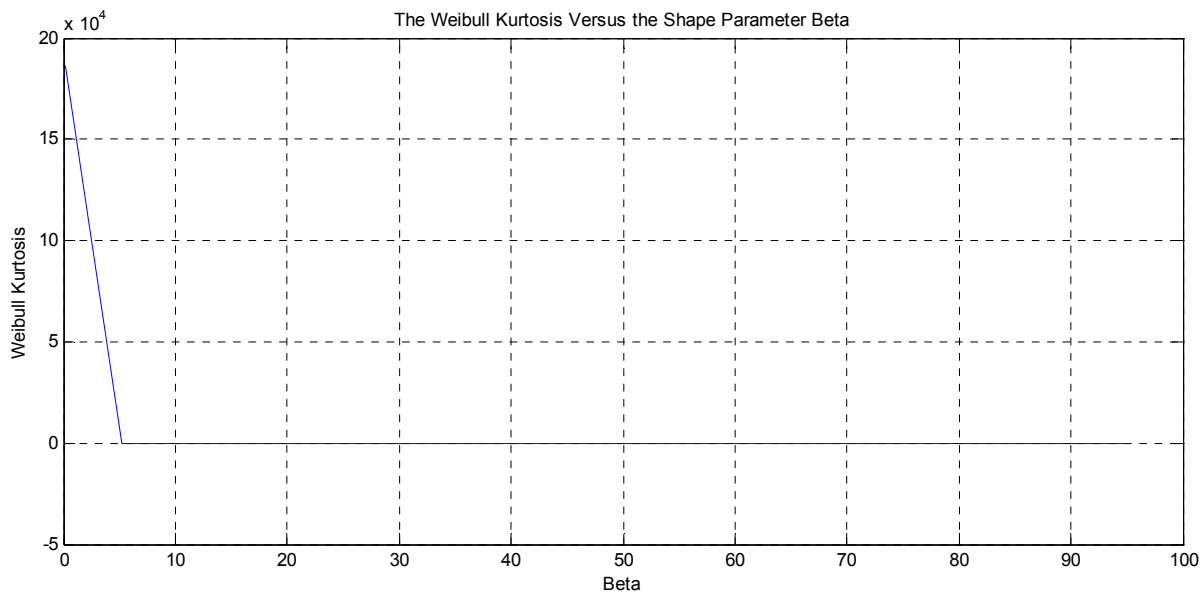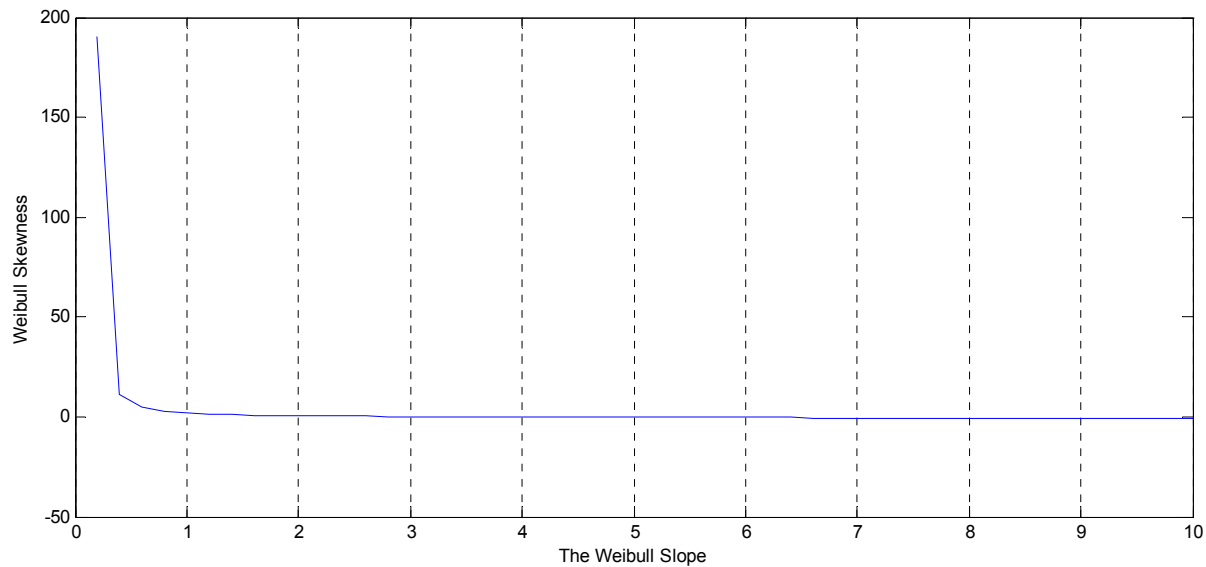
The Proof of the Chebyshev's Inequality which states the $P(|X-\mu| \geq k\sigma) \leq 1/k^2$.

Define the set $\psi = \{x: |X-\mu| \geq k\sigma)\}$; then

$$V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx \;=\; \int_{\psi} (x - \mu)^2 f(x)dx + \int_{\psi'} (x - \mu)^2 f(x)dx \;=\; \sigma^2.$$

Thus, $\sigma^2 \geq \int_{\psi} (x - \mu)^2 f(x)dx \geq \int_{\psi} (k\sigma)^2 f(x)dx = k^2\sigma^2 \int_{\psi} f(x)dx \;=\; k^2\sigma^2 \Pr(|x - \mu| \geq k\sigma) \rightarrow$

$$\Pr(|x - \mu| \geq k\sigma) \leq 1/k^2 \;.$$

My Chapter 4 notes are mostly edited by Mohammad-Ali Alamdar-Yazdi (mza0052), and also by Rong Huangfu (rzh0024).