2. Compare the TMR/simplex reliability with two-component and three-component redundant systems having standby redundancy. Graph the expressions on the same plot.
3. Repeat problem 2 for two and three component parallel redundant systems.
4. Show that the reliability expression (3.78) for k-out-of-n system reliability reduces to the expression

$$\sum_{i=k}^{n} \binom{n}{i} e^{-i\lambda t}(1 - e^{-\lambda t})^{n-i}.$$

5. Using equation (3.80) obtain an explicit expression for the reliability of a hybrid TMR system with one spare. Compare the reliability of this system with those of a TMR system and a simplex system by plotting. Use $\lambda = 1/10,000\ \text{h}^{-1}$ and $\mu = 1/100,000\ \text{h}^{-1}$.
6. Compare (by plotting) reliability expressions for the simplex system, the two-component parallel redundant system, and the two-component standby redundant system. Assume that the failure rate of an active component is constant at $1/10,000\ \text{h}^{-1}$, the failure rate of a spare is zero, and that the switching mechanism is fault-free.

3.9 FUNCTIONS OF NORMAL RANDOM VARIABLES

The normal distribution has great importance in mathematical statistics because of the central-limit theorem alluded to earlier. This distribution also plays an important role in communication and information theory. We will now study distributions derivable from the normal distribution. The use of most of these distributions will be deferred until Chapters 10 and 11.

THEOREM 3.6. Let $X_1, X_2, \ldots, X_n$ be mutually independent random variables such that $X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \ldots, n$. Then $S_n = \sum_{i=1}^{n} X_i$ is normally distributed, that is, $S_n \sim N(\mu, \sigma^2)$, where

$$\mu = \sum_{i=1}^{n} \mu_i \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{n} \sigma_i^2.$$

Owing to this theorem, we say that the normal distribution has the **reproductive** property. A proof of this theorem will be given in Chapter 4. The theorem can be further generalized as in problem 1 at the end of Section 4.4, so that if $X_1, X_2, \ldots, X_n$ are mutually independent random variables with $X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \ldots, n$ and $a_1, a_2, \ldots a_n$ are real constants, then $Y_n = \sum_{i=1}^{n} a_i X_i$ is normally distributed; that is, $Y_n \sim N(\mu, \sigma^2)$, where

$$\mu = \sum_{i=1}^{n} a_i\, \mu_i \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{n} a_i^2\, \sigma_i^2.$$

In particular, if we let $n = 2$, $a_1 = +1$, and $a_2 = -1$, then we conclude that the difference $Y = X_1 - X_2$ of two independent normal random variables $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ is normally distributed, that is, $Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Example 3.33

It has been empirically determined that the memory requirement of a program (called the **working-set size** of the program) is approximately normal. In a multiprogramming system, the number of programs sharing the main memory simultaneously (called the **degree of multiprogramming**) is found to be n. Now if X_i denotes the working-set size of the ith program with $X_i \sim N(\mu_i, \sigma_i^2)$, then it follows that the sum total memory demand, S_n, of the n programs is normally distributed with parameters $\mu = \sum_{i=1}^{n} \mu_i$ and $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$.

♯

Example 3.34

A sequence of independent, identically distributed random variables, $X_1, X_2, \ldots, X_n$, is known in mathematical statistics as a *random sample* of size n. In many problems of statistical sampling theory, it is reasonable to assume that the underlying distribution is the normal distribution. Thus let $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$. Then from Theorem 3.6, we obtain

$$S_n = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

One important function known as the sample mean is quite useful in problems of statistical inference. *Sample mean* $\overline{X}$ is given by

$$\overline{X} = \frac{S_n}{n} = \sum_{i=1}^{n} \frac{X_i}{n}. \tag{3.85}$$

To obtain the pdf of the sample mean $\overline{X}$, we use equation (3.55) to obtain

$$f_{\overline{X}} = n f_{S_n}(nx).$$

But since $S_n \sim N(n\mu, n\sigma^2)$, we have

$$\begin{aligned} f_{\overline{X}}(x) &= n\frac{1}{\sqrt{2\pi}(\sqrt{n}\sigma)} e^{-\frac{(nx-n\mu)^2}{2n\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}[\sigma(n)^{-1/2}]} e^{-\frac{(x-\mu)^2}{2(\sigma^2/n)}}, \quad -\infty < x < \infty. \end{aligned}$$

It follows that $\overline{X} \sim N(\mu, \sigma^2/n)$. Similarly, it can be shown that the random variable $(\overline{X} - \mu)\sqrt{n}/\sigma$ has the standard normal distribution, $N(0, 1)$.

♯

If X is $N(0, 1)$, we know from Example 3.9 that $Y = X^2$ is gamma-distributed with $Y \sim \text{GAM}(\frac{1}{2}, \frac{1}{2})$, which is the chi-square distribution with

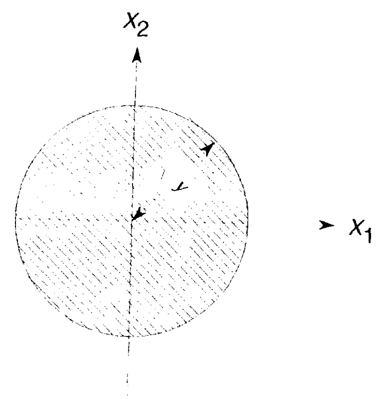

Figure 3.45. The area of integration for Example 3.35

one degree of freedom. Now consider X_1, X_2 that are independent standard normal random variables and $Y = X_1^2 + X_2^2$.

Example 3.35

If $X_1 \sim N(0,1)$, $X_2 \sim N(0,1)$ and X_1 and X_2 are independent, then $Y = X_1^2 + X_2^2$ is exponentially distributed so that $Y \sim \text{EXP}(\frac{1}{2})$.

To see this, we obtain the distribution function of Y:

$$\begin{aligned} F_Y(y) &= P(X_1^2 + X_2^2 \le y) \\ &= \iint_{x_1^2+x_2^2 \le y} f_{X_1X_2}(x_1, x_2)dx_1dx_2. \end{aligned}$$

Note that the surface of integration is a circular area about the origin with the radius $\sqrt{y}$ (see Figure 3.45). Using the fact that X_1 and X_2 are independent, and standard normal, we have

$$F_Y(y) = \iint_{x_1^2+x_2^2 \le y} \frac{1}{2\pi} e^{-\frac{x_1^2+x_2^2}{2}} dx_1dx_2.$$

Making a change of variables (to polar coordinates), $x_1 = r\cos\theta$, $x_2 = r\sin\theta$, so that $r^2 = x_1^2 + x_2^2$ and $\theta = \tan^{-1}(x_2/x_1)$, we have

$$\begin{aligned} F_Y(y) &= \int_{\theta=0}^{2\pi} \int_{r=0}^{\sqrt{y}} \frac{r}{2\pi} e^{-r^2/2} dr\, d\theta \\ &= \begin{cases} 1 - e^{-y/2}, & y > 0, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Therefore, Y is exponentially distributed with parameter $\frac{1}{2}$.

♯

This example is a special case of the following theorem.

THEOREM 3.7. If $X_1, X_2, \ldots, X_n$ is a sequence of mutually independent, standard normal random variables, then

$$Y = \sum_{i=1}^{n} X_i^2$$

has the gamma distribution, $\text{GAM}(\frac{1}{2}, n/2)$, or the chi-square distribution with n degrees of freedom, X_n^2.

This theorem follows from the reproductive property of the gamma distribution (see Theorem 3.8).

THEOREM 3.8. Let $X_1, X_2, \ldots, X_n$ be a sequence of mutually independent gamma random variables such that $X_i \sim \text{GAM}(\lambda, \alpha_i)$ for $i = 1, 2, \ldots, n$. Then $S_n = \sum_{i=1}^{n} X_i$ has the gamma distribution $\text{GAM}(\lambda, \alpha)$, where $\alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_n$.

This theorem will be proved in Chapter 4.

Since $X_n^2 \sim \text{GAM}(\frac{1}{2}, n/2)$, we have the following corollary.

COROLLARY 3.8. Let $Y_1, Y_2, \ldots, Y_n$ be mutually independent chi-square random variables such that $Y_i \sim X_{k_i}^2$. Then $Y_1 + Y_2 + \ldots + Y_n$ has the X_k^2 distribution, where

$$k = \sum_{i=1}^{n} k_i.$$

Example 3.36

Assume that $X_1, X_2, \ldots, X_n$ are mutually independent, identically distributed normal random variables such that $X_i \sim N(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$. It follows that $Z_i = (X_i - \mu)/\sigma$ is standard normal. Thus $Z_1, Z_2, \ldots, Z_n$ are independent standard normal random variables. Hence, using Theorem 3.7, we have

$$Y = \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2}, \tag{3.86}$$

which has the chi-square distribution with n degrees of freedom. Note that the random variable $\sum_{i=1}^{n}(X_i - \mu)^2/n$ may be used as an estimator of the parameter σ^2.

♯

Example 3.37

In the last example, we suggested that $\sum_{i=1}^{n}(X_i - \mu)^2/n$ may be used as an estimator of the parameter σ^2 assuming that $X_1, X_2, \ldots, X_n$ are independent observations from a normal distribution $N(\mu, \sigma^2)$. However, this expression assumes that the parameter μ of the distribution is already known. This is rarely the case in practice, and the sample mean $\overline{X} = \sum_{i=1}^{n} X_i/n$ is usually substituted in its place. Thus, the random variable

$$U = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} \tag{3.87}$$

is usually used as an estimator of the parameter σ^2 and is often denoted by S^2. (The reason for the value $n-1$ rather than n in the denominator will be seen in Chapter 10).

Rewriting, we have

$$S^2 = U = \frac{\sigma^2}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2, \tag{3.88}$$

where we note that n random variables $\{(X_i - \overline{X})/\sigma \mid 1 \le i \le n\}$ satisfy the relation

$$\sum_{i=1}^{n}\frac{X_i - \overline{X}}{\sigma} = 0 \tag{3.89}$$

(from the definition of the sample mean, $\overline{X}$). Thus they are linearly dependent. It can be shown that the random variable

$$W = \sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \tag{3.90}$$

can be transformed to a sum of squares of $(n-1)$ independent standard normal random variables, and hence $W = (n-1)U/\sigma^2 = (n-1)S^2/\sigma^2$ has a chi-square distribution with $n-1$ degrees of freedom (rather than n degrees of freedom).

♯

Just as the sums of chi-square random variables are of interest, so is the ratio of two chi-square random variables. First assume that X and Y are independent, positive-valued random variables and let Z be their quotient:

$$Z = \frac{Y}{X}. \tag{3.91}$$

Then the distribution function of Z is obtained using the formula

$$F_Z(z) = \iint_{A_z} f(x,y)\, dx\, dy,$$

where the set

$$A_z = \{(x,y) | y/x \le z\}$$

is shown in Figure 3.46. Therefore

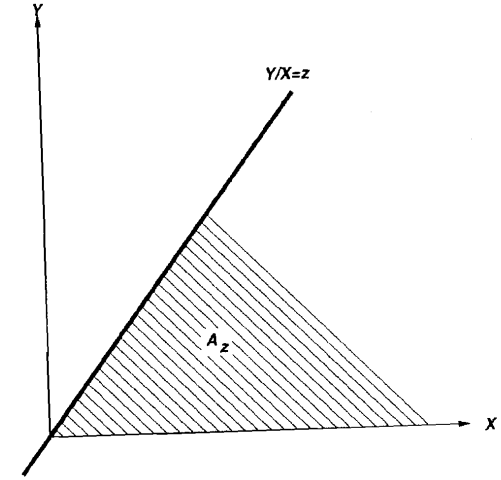

Figure 3.46. The area of integration for computing the CDF of $Y/X = Z$

$$\begin{aligned} F_Z(z) &= \int_0^\infty \left[\int_0^{xz} f(x,y)\, dy\right] dx \\ &= \int_0^\infty \left[\int_0^{z} x\, f(x,xv)\, dv\right] dx, \end{aligned} \tag{3.92}$$

after a change of variables to $y = xv$.

It follows that the pdf of Z is given by

$$\begin{aligned} f_Z(z) &= \int_0^\infty x\, f(x,xz)\, dx \\ &= \int_0^\infty x f_X(x) f_Y(xz)\, dx, \quad 0 < z < \infty \end{aligned} \tag{3.93}$$

(by independence of X and Y).

THEOREM 3.9. Let Y_1 and Y_2 be independent random variables with $X^2_{n_1}$ and $X^2_{n_2}$ distributions, respectively. Then

$$Z = \frac{Y_1/n_1}{Y_2/n_2}$$

has the F distribution, which is characterized by two parameters, (n_1, n_2), that is, $Z \sim F_{n_1,n_2}$. The pdf of Z is given by

$$f_Z(z) = \begin{cases} \dfrac{(n_1/n_2)\Gamma[(n_1+n_2)/2](n_1 z/n_2)^{(n_1/2)-1}}{\Gamma(n_1/2)\Gamma(n_2/2)[1+(n_1 z/n_2)]^{(n_1+n_2)/2}}, & z > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.94}$$

Proof: Recall that

$$f_{Y_1}(y_1) = \frac{y_1^{(n_1/2)-1} e^{-y_1/2}}{2^{n_1/2}\Gamma(n_1/2)}$$

and

$$f_{Y_2}(y_2) = \frac{y_2^{(n_2/2)-1} e^{-y_2/2}}{2^{n_2/2}\Gamma(n_2/2)}.$$

Let $Y = Y_1/n_1$ and $X = Y_2/n_2$. Using formula (3.55), it follows that

$$f_Y(y) = \frac{n_1(yn_1)^{n_1/2-1} e^{-(n_1 y)/2}}{2^{n_1/2}\Gamma(n_1/2)}$$

and

$$f_X(x) = \frac{n_2(xn_2)^{n_2/2-1} e^{-(n_2 x)/2}}{2^{n_2/2}\Gamma(n_2/2)}.$$

Now, applying equation (3.93), we get

$$\begin{aligned} f_Z(z) &= \int_0^\infty x \frac{n_1 n_2}{2^{(n_1+n_2)/2}\Gamma(n_1/2)\Gamma(n_2/2)} \\ &\quad \cdot (xzn_1)^{n_1/2-1}(xn_2)^{n_2/2-1} e^{-(n_1 xz + n_2 x)/2}\, dx \\ &= \frac{n_1 n_2 (n_1)^{n_1/2-1}(n_2)^{n_2/2-1} z^{n_1/2-1}}{2^{(n_1+n_2)/2}\Gamma(n_1/2)\Gamma(n_2/2)} \\ &\quad \cdot \int_0^\infty x^{n_1/2+n_2/2-1} e^{-x(n_1 z+n_2)/2}\, dx. \end{aligned} \tag{3.95}$$

Using equation (3.25), the last integral is evaluated as

$$\frac{\Gamma[(n_1+n_2)/2]}{[(n_1 z + n_2)/2]^{(n_1+n_2)/2}}.$$

Substituting this in (3.95), we get the required result as in (3.94).

Example 3.38

Suppose that $X_1, X_2, \ldots, X_m, X_{m+1}, \ldots, X_n$ are mutually independent normal random variables with the common distribution, $N(0, \sigma^2)$. Then by Theorem 3.7

$$Y = \sum_{i=1}^{m} \frac{X_i^2}{\sigma^2} \quad \text{and} \quad X = \sum_{i=m+1}^{n} \frac{X_i^2}{\sigma^2}$$

are chi-square distributed with m and $(n-m)$ degrees of freedom, respectively. Furthermore, X and Y are independent. It follows by Theorem 3.9 that

$$Z = \frac{\sum_{i=1}^{m} X_i^2/m}{\sum_{i=m+1}^{n} X_i^2/(n-m)} \tag{3.96}$$

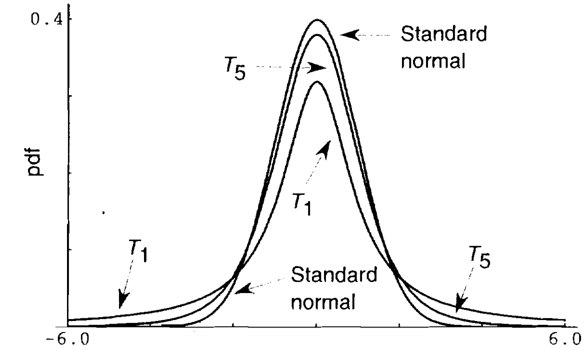

Figure 3.47. Student's t pdf and its comparison with standard normal pdf

has the $F_{m,n-m}$ distribution.

♯

The last distribution we introduce here is Student's t distribution.

THEOREM 3.10. If V and W are independent random variables such that $V \sim N(0,1)$ and $W \sim X_n^2$, then the random variable

$$T = \frac{V}{\sqrt{W/n}} \tag{3.97}$$

has the t distribution with n degrees of freedom. The pdf of this random variable is given by

$$f_T(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma[n/2]}\left[1 + \frac{t^2}{n}\right]^{-(n+1)/2}, \quad -\infty < t < \infty. \tag{3.98}$$

For $n = 1$, this pdf reduces to

$$f_T(t) = \frac{1}{\pi(1+t^2)}, \tag{3.99}$$

which is known as the **Cauchy pdf**.

The pdf in (3.98) is plotted for various degrees of freedom in Figure 3.47. It may be noted that as n approaches infinity, the t distribution approaches the normal distribution.

Example 3.39

Assume that $X_1, X_2, \ldots, X_n$ are mutually independent identically distributed normal random variables such that $X_i \sim N(\mu, \sigma^2)$. Then from Example 3.34, it follows that

$$V = \frac{(\overline{X} - \mu)\sqrt{n}}{\sigma} \tag{3.100}$$

has the standard normal distribution. Also, from Example 3.37

$$\frac{(n-1)S^2}{\sigma^2} = W = \sum_{i=1}^{n} \left[\frac{X_i - \overline{X}}{\sigma}\right]^2 \tag{3.101}$$

has the X^2_{n-1} distribution. It follows that

$$\begin{aligned} T &= \frac{V}{\sqrt{\frac{W}{(n-1)}}} = \frac{(\overline{X} - \mu)\sqrt{n}/\sigma}{\left[S\frac{\sqrt{n-1}}{\sigma}\right]} \cdot \sqrt{n-1} \\ &= \frac{\overline{X} - \mu}{S/\sqrt{n}} \end{aligned} \tag{3.102}$$

has the t distribution with $(n-1)$ degrees of freedom.

♯

Problems

1. In communication theory, waveforms of the form

$$A(t) = x(t)\cos(\omega t) - y(t)\sin(\omega t)$$

appear quite frequently. At a fixed time instant, $t = t_1$, $X = X(t_1)$, and $Y = Y(t_1)$ are known to be independent Gaussian random variables, specifically, $N(0, \sigma^2)$. Show that the distribution function of the **envelope** $Z = \sqrt{X^2 + Y^2}$ is given by

$$F_Z(z) = \begin{cases} 1 - e^{-z^2/2\sigma^2}, & z > 0, \\ 0, & \text{otherwise.} \end{cases}$$

This distribution is called the **Rayleigh distribution**. Compute and plot its pdf.

2. The test effort during the software testing phase, such as human resources, the number of test cases, and CPU time, can be measured by the cumulative amount of testing effort during the time interval $(0, t]$. Yamada et al. [YAMA 1986] proposed a formula for $W(t)$: $dW(t)/dt = g(t)(1 - W(t))$ where $g(t)$ is the instantaneous consumption rate of the testing effort expenditures. $W(t)$ is defined as $W(t) = \int_0^t w(t)dt$ where $w(t)$ is the testing-effort consumption rate at time t. Find an explicit expression for $W(t)$ in terms of $g(t)$ and show that $W(t)$ is the Rayleigh distribution.

3. A calculator operates on two 1.5-V batteries (for a total of 3 V). The actual voltage of a battery is normally distributed with $\mu = 1.5$ and $\sigma^2 = 0.45$. The tolerances in the design of the calculator are such that it will not operate satisfactorily if the total voltage falls outside the range 2.70–3.30 V. What is the probability that the calculator will function correctly?