Chapter 2
Normal Distribution

The normal distribution is one of the most important continuous probability distributions, and is widely used in statistics and other fields of sciences. In this chapter, we present some basic ideas, definitions, and properties of normal distribution, (for details, see, for example, Whittaker and Robinson (1967), Feller (1968, 1971), Patel et al. (1976), Patel and Read (1982), Johnson et al. (1994), Evans et al. (2000), Balakrishnan and Nevzorov (2003), and Kapadia et al. (2005), among others).

2.1 Normal Distribution

The normal distribution describes a family of continuous probability distributions, having the same general shape, and differing in their location (that is, the mean or average) and scale parameters (that is, the standard deviation). The graph of its probability density function is a symmetric and bell-shaped curve. The development of the general theories of the normal distributions began with the work of de Moivre (1733, 1738) in his studies of approximations to certain binomial distributions for large positive integer $n > 0$. Further developments continued with the contributions of Legendre (1805), Gauss (1809), Laplace (1812), Bessel (1818, 1838), Bravais (1846), Airy (1861), Galton (1875, 1889), Helmert (1876), Tchebyshev (1890), Edgeworth (1883, 1892, 1905), Pearson (1896), Markov (1899, 1900), Lyapunov (1901), Charlier (1905), and Fisher (1930, 1931), among others. For further discussions on the history of the normal distribution and its development, readers are referred to Pearson (1967), Patel and Read (1982), Johnson et al. (1994), and Stigler (1999), and references therein. Also, see Wiper et al. (2005), for recent developments. The normal distribution plays a vital role in many applied problems of biology, economics, engineering, financial risk management, genetics, hydrology, mechanics, medicine, number theory, statistics, physics, psychology, reliability, etc., and has been has been extensively studied, both from theoretical and applications point of view, by many researchers, since its inception.

M. Ahsanullah et al., *Normal and Student's t Distributions and Their Applications*,
Atlantis Studies in Probability and Statistics 4, DOI: 10.2991/978-94-6239-061-4_2,

2.1.1 Definition (Normal Distribution)

A continuous random variable X is said to have a normal distribution, with mean μ and variance σ^2, that is, $X \sim N(\mu,\ \sigma^2)$, if its pdf $f_X(x)$ and cdf $F_X(x) = P(X \leq x)$ are, respectively, given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty, \tag{2.1}$$

and

$$\begin{aligned} F_X(x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-(y-\mu)^2/2\sigma^2} dy \\ &= \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right], \quad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0, \end{aligned} \tag{2.2}$$

where $erf(.)$ denotes error function, and μ and σ are location and scale parameters, respectively.

2.1.2 Definition (Standard Normal Distribution)

A normal distribution with $\mu = 0$ and $\sigma = 1$, that is, $X \sim N(0, 1)$, is called the standard normal distribution. The pdf $f_X(x)$ and cdf $F_X(x)$ of $X \sim N(0, 1)$ are, respectively, given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \ -\infty < x < \infty, \tag{2.3}$$

and

$$\begin{aligned} F_X(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \quad -\infty < x < \infty, \\ &= \frac{1}{2}\left[1 + erf\left(\frac{x}{\sqrt{2}}\right)\right], \quad -\infty < x < \infty. \end{aligned} \tag{2.4}$$

Note that if $Z \sim N(0,\ 1)$ and $X = \mu + \sigma Z$, then $X \sim N(\mu, \sigma^2)$, and conversely if $X \sim N(\mu,\ \sigma^2)$ and $Z = (X - \mu)/\ \sigma$, then $Z \sim N(0,\ 1)$. Thus, the pdf of any general $X \sim N(\mu,\ \sigma^2)$ can easily be obtained from the pdf of $Z \sim N(0,\ 1)$, by using the simple location and scale transformation, that is, $X = \mu + \sigma\ Z$. To describe the shapes of the normal distribution, the plots of the pdf (2.1) and cdf (2.2),

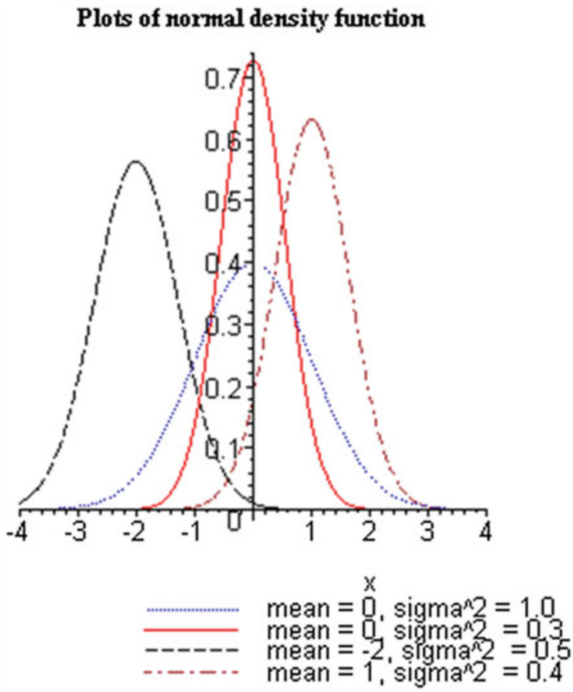

Fig. 2.1 Plots of the normal pdf, for different values of μ and σ^2

for different values of μ and σ^2, are provided in Figs. 2.1 and 2.2, respectively, by using Maple 10. The effects of the parameters, μ and σ^2, can easily be seen from these graphs. Similar plots can be drawn for other values of the parameters. It is clear from Fig. 2.1 that the graph of the pdf $f_X(x)$ of a normal random variable, $X \sim N(\mu,\ \sigma^2)$, is symmetric about mean, μ, that is $f_X(\mu + x) = f_X(\mu - x)$, $-\infty < x < \infty$.

2.1.3 Some Properties of the Normal Distribution

This section discusses the mode, moment generating function, cumulants, moments, mean, variance, coefficients of skewness and kurtosis, and entropy of the normal distribution, $N(\mu,\ \sigma^2)$. For detailed derivations of these, see, for example, Kendall and Stuart (1958), Lukacs (1972), Dudewicz and Mishra (1988), Johnson et al. (1994), Rohatgi and Saleh (2001), Balakrishnan and Nevzorov (2003), Kapadia et al. (2005), and Mukhopadhyay (2006), among others.

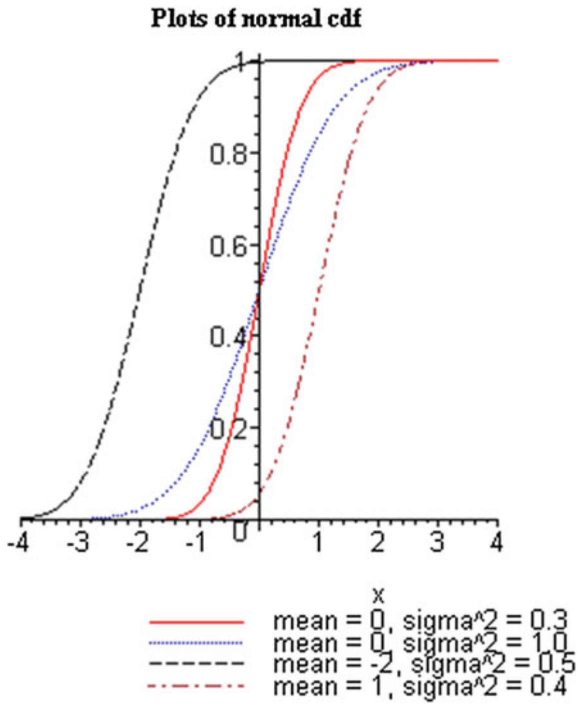

Fig. 2.2 Plots of the normal cdf for different values of μ and σ^2

2.1.3.1 Mode

The mode or modal value is that value of X for which the normal probability density function $f_X(x)$ defined by (2.1) is maximum. Now, differentiating with respect to x Eq. (2.1), we have

$$f'_X(x) = -\sqrt{\frac{2}{\pi}} \left[\frac{(x - \mu)\, e^{-(x-\mu)^2/2\sigma^2}}{\sigma^3} \right],$$

which, when equated to 0, easily gives the mode to be $x = \mu$, which is the mean, that is, the location parameter of the normal distribution. It can be easily seen that $f''_X(x) < 0$. Consequently, the maximum value of the normal probability density function $f_X(x)$ from (2.1) is easily obtained as $f_X(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$. Since $f'(x) = 0$ has one root, the normal probability density function (2.1) is unimodal.

2.1.3.2 Cumulants

The cumulants k_r of a random variable X are defined via the cumulant generating function

$$g(t) = \sum_{r=1}^{\infty} k_r \frac{t^r}{r!}, \text{ where } \mathrm{g(t)} = \ln\left(\mathrm{E(e^{tX})}\right).$$

For some integer $r > 0$, the rth cumulant of a normal random variable X having the pdf (2.1) is given by

$$\kappa_r = \begin{cases} \mu, & when\ r = 1; \\ \sigma^2, & when\ r = 2; \\ 0, & when\ r > 2 \end{cases}$$

2.1.3.3 Moment Generating Function

The moment generating function of a normal random variable X having the pdf (2.1) is given by (see, for example, Kendall and Stuart (1958), among others)

$$M_X(t) = E\left(e^{tX}\right) = e^{t\mu + \frac{1}{2} t^2 \sigma^2}.$$

2.1.3.4 Moments

For some integer $r > 0$, the rth moment about the mean of a normal random variable X having the pdf (2.1) is given by

$$E\left(X^r\right) = \mu_r = \begin{cases} \frac{\sigma^r (r!)}{2^{\frac{r}{2}} [(r/2)!]}, & \text{for r even;} \\ 0, & \text{for r odd} \end{cases} \tag{2.5}$$

We can write $\mu_r = \sigma^r (r!!)$, where $m!! = 1.3.5.\ldots(m-1)$ for m even.

2.1.3.5 Mean, Variance, and Coefficients of Skewness and Kurtosis

From (2.5), the mean, variance, and coefficients of skewness and kurtosis of a normal random variable $X \sim N(\mu,\ \sigma^2)$ having the pdf (2.1) are easily obtained as follows:

(i) **Mean:** $\alpha_1 = E(X) = \mu$;
(ii) **Variance:** $Var(X) = \sigma^2, \quad \sigma > 0$;
(iii) **Coefficient of Skewness:** $\gamma_1(X) = \frac{\mu_3}{\mu_2^{3/2}} = 0$;
(iv) **Coefficient of Kurtosis:** $\gamma_2(X) = \frac{\mu_4}{\mu_2^2} = 3$.

where μ_r has been defined in Eq. (2.5).

Since the coefficient of kurtosis, that is, $\gamma_2(X) = 3$, it follows that the normal distributions are mesokurtic distributions.

2.1.3.6 Median, Mean Deviation, and Coefficient of Variation of $X \sim N(\mu, \sigma^2)$

These are given by

(i) **Median:** μ

(ii) **Mean Deviation:** $\left(\frac{2\,\sigma^2}{\pi}\right)^{\frac{1}{2}}$

(iii) **Coefficient of Variation:** $\frac{\sigma}{\mu}$

2.1.3.7 Characteristic Function

The characteristic function of a normal random variable $X \sim N(\mu, \sigma^2)$ having the pdf (2.1) is given by (see, for example, Patel et al. (1976), among others)

$$\phi_X(t) = E\left(e^{itX}\right) = e^{it\mu - \frac{1}{2}t^2\sigma^2}, \; \mathrm{i} = \sqrt{-1}.$$

2.1.3.8 Entropy

For some $\sigma > 0$, entropy of a random variable X having the pdf (2.1) is easily given by

$$\begin{aligned} H_X\left[f_X(x)\right] &= E[-\ln(f_X(X)] \\ &= -\int_{-\infty}^{\infty} f_X(x)\ \ln\ [f_X(x)]\ dx, \\ &= \ln\left(\sqrt{2\pi e\sigma}\right) \end{aligned}$$

(see, for example, Lazo and Rathie (1978), Jones (1979), Kapur (1993), and Suhir (1997), among others). It can be easily seen that $\frac{d(H_X[f_X(x)])}{d\sigma} > 0$, and $\frac{d^2(H_X[f_X(x)])}{d\sigma^2} < 0$, $\forall\sigma > 0$, $\forall\mu$. It follows that that the entropy of a random variable X having the normal pdf (2.1) is a monotonic increasing concave function of $\sigma > 0$, $\forall\mu$. The possible shape of the entropy for different values of the parameter σ is provided below in Fig. 2.3, by using Maple 10. The effects of the parameter σ on entropy can easily be seen from the graph. Similar plots can be drawn for others values of the parameter σ.

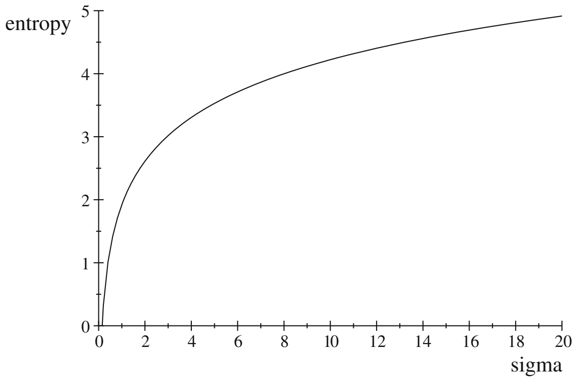

Fig. 2.3 Plot of entropy

2.1.4 Percentiles

This section computes the percentiles of the normal distribution, by using Maple 10. For any $p (0 < p < 1)$, the $(100p)th$ percentile (also called the quantile of order p) of $N(\mu,\ \sigma^2)$ with the pdf $f_X(x)$ is a number z_p such that the area under $f_X(x)$ to the left of z_p is p. That is, z_p is any root of the equation

$$\Phi(z_p) = \int_{-\infty}^{z_p} f_X(u)du = p.$$

Using the Maple program, the percentiles z_p of $N(\mu,\ \sigma^2)$ are computed for some selected values of p for the given values of μ and , which are provided in Table 2.1, when $\mu = 0$ and $\sigma = 1$. Table 2.1 gives the percentile values of z_p for $\mathrm{p} \geq 0.5$. For $\mathrm{p} < 0.5$, use $1 - \mathrm{Z}_{1-\mathrm{p}}$.

Table 2.1 Percentiles of $N(0,\ 1)$

p	z_p
0.5	0.0000000000
0.6	0.2533471031
0.7	0.5244005127
0.75	0.6744897502
0.8	0.8416212336
0.9	1.281551566
0.95	1.644853627
0.975	1.959963985
0.99	2.326347874
0.995	2.575829304
0.9975	2.807033768
0.999	3.090232306

Suppose $X_1, X_2, \ldots X_n$ are n independent N (0, 1) random variables and $M(n) = \max(X_1, X_2, \ldots X_n)$. It is known (see Ahsanullah and Kirmani (2008) p.15 and Ahsanullah and Nevzorov (2001) p.92) that

$P(M(n) \leq a_n + b_n x) \to e^{-e^{-x}}$, for all x as n $\to \infty$.

where $a_n = \beta_n - \frac{D_n}{2\beta_n}$, $D_n = \ln \ln n + \ln 4\pi$, $\beta_n = (2 \ln n)^{1/2}$, $b_n - (2 \ln n)^{-1/2}$.

2.2 Different Forms of Normal Distribution

This section presents different forms of normal distribution and some of their important properties, (for details, see, for example, Whittaker and Robinson (1967), Feller (1968, 1971), Patel et al. (1976), Patel and Read (1982), Johnson et al. (1994), Evans et al. (2000), Balakrishnan and Nevzorov (2003), and Kapadia et al. (2005), among others).

2.2.1 Generalized Normal Distribution

Following Nadarajah (2005a), a continuous random variable X is said to have a generalized normal distribution, with mean μ and variance $\frac{\sigma^2 \Gamma\left(\frac{3}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}$, where $s > 0$, that is, $X \sim N\left(\mu, \frac{\sigma^2 \Gamma\left(\frac{3}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}\right)$, if its pdf $f_X(x)$ and cdf $F_X(x) = P(X \leq x)$ are, respectively, given by

$$f_X(x) = \frac{s}{2\sigma\Gamma\left(\frac{1}{s}\right)} e^{-\left|\frac{x-\mu}{\sigma}\right|^s}, \tag{2.6}$$

and

$$F_X(x) = \begin{cases} \dfrac{\Gamma\left(\frac{1}{s}, \left(\frac{\mu - x}{\sigma}\right)^s\right)}{2\Gamma\left(\frac{1}{s}\right)}, & if\ x \leq \mu \\ 1 - \dfrac{\Gamma\left(\frac{1}{s}, \left(\frac{x - \mu}{\sigma}\right)^s\right)}{2\Gamma\left(\frac{1}{s}\right)}, & if\ x > \mu \end{cases} \tag{2.7}$$

where $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$, $s > 0$, and $\Gamma(a, x)$ denotes complementary incomplete gamma function defined by $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$. It is easy to see that the Eq. (2.6) reduces to the normal distribution for $s = 2$, and Laplace distribution for $s = 1$. Further, note that if has the pdf given by (2.6), then the pdf of the standardized random variable $Z = (X - \mu)/\sigma$ is given by

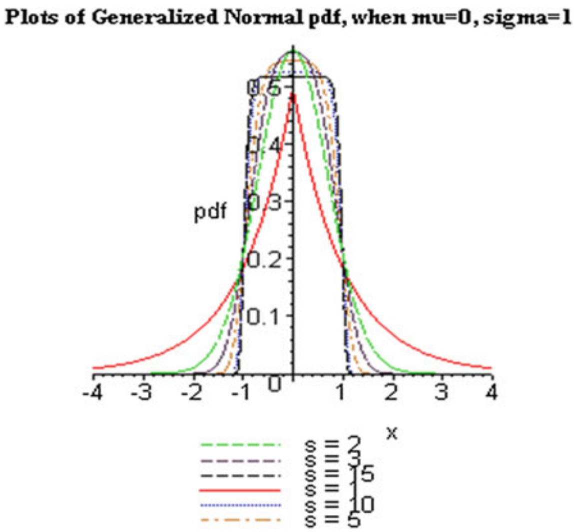

Fig. 2.4 Plots of the generalized normal pdf for different values of s

$$f_Z(z) = \frac{s}{2\Gamma\left(\frac{1}{s}\right)} e^{-|z|^s} \tag{2.8}$$

To describe the shapes of the generalized normal distribution, the plots of the pdf (2.6), for $\mu = 0, \sigma = 1$, and different values of s, are provided in Fig. 2.4 by using Maple 10. The effects of the parameters can easily be seen from these graphs. Similar plots can be drawn for others values of the parameters. It is clear from Fig. 2.4 that the graph of the pdf $f_X(x)$ of the generalized normal random variable is symmetric about mean, μ, that is

$$f_X(\mu + x) = f_X(\mu - x), \quad -\infty < x < \infty.$$

2.2.1.1 Some Properties of the Generalized Normal Distribution

This section discusses the mode, moments, mean, median, mean deviation, variance, and entropy of the generalized normal distribution. For detailed derivations of these, see Nadarajah (2005).

2.2.1.2 Mode

It is easy to see that the mode or modal value of x for which the generalized normal probability density function $f_X(x)$ defined by (2.6) is maximum, is given by $x = \mu$, and the maximum value of the generalized normal probability density function (2.6) is given by $f_X(\mu) = \frac{s}{2\sigma\Gamma\left(\frac{1}{s}\right)}$. Clearly, the generalized normal probability density function (2.6) is unimodal.

2.2.1.3 Moments

(i) For some integer $r > 0$, the rth moment of the generalized standard normal random variable Z having the pdf (2.8) is given by

$$E\left(Z^r\right) = \frac{1 + (-1)^r}{2\Gamma\left(\frac{1}{s}\right)}\Gamma\left(\frac{r + 1}{s}\right) \tag{2.9}$$

(i) For some integer $n > 0$, the nth moment and the nth central moment of the generalized normal random variable X having the pdf (2.6) are respectively given by the Eqs. (2.10) and (2.11) below:

$$E\left(X^n\right) = \frac{(\mu^n)\sum_{k=0}^{n}\binom{n}{k}\left(\frac{\sigma}{\mu}\right)^k\left[1 + (-1)^k\right]\Gamma\left(\frac{k+1}{s}\right)}{2\Gamma\left(\frac{1}{s}\right)} \tag{2.10}$$

and

$$E\left[(X - \mu)^n\right] = \frac{(\sigma^n)\left[1 + (-1)^n\right]\Gamma\left(\frac{n+1}{s}\right)}{2\Gamma\left(\frac{1}{s}\right)} \tag{2.11}$$

2.2.1.4 Mean, Variance, Coefficients of Skewness and Kurtosis, Median and Mean Deviation

From the expressions (2.10) and (2.11), the mean, variance, coefficients of skewness and kurtosis, median and mean deviation of the generalized normal random variable X having the pdf (2.6) are easily obtained as follows:

(i) **Mean**: $\alpha_1 = E(X) = \mu$;

(ii) **Variance**: $Var(X) = \beta_2 = \frac{\sigma^2\Gamma\left(\frac{3}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}, \quad \sigma > 0,\ s > 0$;

(iii) **Coefficient of Skewness**: $\gamma_1(X) = \frac{\beta_3}{\beta_2^{3/2}} = 0$;

(iv) **Coefficient of Kurtosis**: $\gamma_2(X) = \frac{\beta_4}{\beta_2^2} = \frac{\Gamma\left(\frac{1}{s}\right)\Gamma\left(\frac{5}{s}\right)}{\left[\Gamma\left(\frac{3}{s}\right)\right]^2}, \quad s > 0$;

(**v**) **Median** (X): μ;

(**vi**) **Mean Deviation**: $E\,|X - \mu| = \dfrac{\sigma\,\Gamma\left(\frac{2}{s}\right)}{\Gamma\left(\frac{1}{s}\right)}, \quad s > 0.$

2.2.1.5 Renyi and Shannon Entropies, and Song's Measure of the Shape of the Generalized Normal Distribution

These are easily obtained as follows, (for details, see, for example, Nadarajah (2005), among others).

(i) **Renyi Entropy:** Following Renyi (1961), for some reals $\gamma > 0$, $\gamma \neq 1$, the entropy of the generalized normal random variable X having the pdf (2.6) is given by

$$\begin{aligned}\Im_R(\gamma) &= \frac{1}{1-\gamma} \ln \int_{-\infty}^{+\infty} [f_X(X)]^{\gamma}\, dx \\ &= \frac{\ln(\gamma)}{s(\gamma-1)} - \ln\left[\frac{s}{2\sigma\Gamma\left(\frac{1}{s}\right)}\right], \quad \sigma > 0,\ s > 0,\ \gamma > 0,\ \gamma \neq 1.\end{aligned}$$

(ii) **Shannon Entropy:** Following Shannon (1948), the entropy of the generalized normal random variable X having the pdf (2.6) is given by

$$H_X[f_X(X)] = E[-\ln(f_X(X)] = -\int_{-\infty}^{\infty} f_X(x)\ \ln\ [f_X(x)]\ dx,$$

which is the particular case of Renyi entropy as obtained in (i) above for $\gamma \to 1$. Thus, in the limit when $\gamma \to 1$ and using L'Hospital's rule, Shannon entropy is easily obtained from the expression for Renyi entropy in (i) above as follows:

$$H_X[f_X(X)] = \frac{1}{s} - \ln\left[\frac{s}{2\sigma\Gamma\left(\frac{1}{s}\right)}\right], \sigma > 0,\ s > 0.$$

(iii) **Song's Measure of the Shape of a Distribution:** Following Song (2001), the gradient of the Renyi entropy is given by

$$\Im'_R(\gamma) = \frac{d}{d\gamma}\left[\Im_R(\gamma)\right] = \frac{1}{s}\left\{\frac{1}{\gamma(\gamma-1)} - \frac{\ln(\gamma)}{(\gamma-1)^2}\right\} \tag{2.12}$$

which is related to the log likelihood by

$$\Im_R'(1) = -\frac{1}{2} Var\left[\ln f(X)\right].$$

Thus, in the limit when $\gamma \rightarrow 1$ and using L'Hospital's rule, Song's measure of the shape of the distribution of the generalized normal random variable X having the pdf (2.6) is readily obtained from the Eq. (2.12) as follows:

$$-2\,\Im_R'(1) = \frac{1}{s},$$

which can be used in comparing the shapes of various densities and measuring heaviness of tails, similar to the measure of kurtosis.

2.2.2 *Half Normal Distribution*

Statistical methods dealing with the properties and applications of the half-normal distribution have been extensively used by many researchers in diverse areas of applications, particularly when the data are truncated from below (that is, left truncated,) or truncated from above (that is, right truncated), among them Dobzhansky and Wright (1947), Meeusen and van den Broeck (1977), Haberle (1991), Altman (1993), Buckland et al. (1993) , Chou and Liu (1998), Klugman et al. (1998), Bland and Altman (1999), Bland (2005), Goldar and Misra (2001), Lawless (2003), Pewsey (2002, 2004), Chen and Wang (2004) and Wiper et al. (2005), Babbit et al. (2006), Coffey et al. (2007), Barranco-Chamorro et al. (2007), and Cooray and Ananda (2008), are notable. A continuous random variable X is said to have a (general) half-normal distribution, with parameters μ (location) and σ (scale), that is, $X|\mu,\ \sigma \sim HN(\mu,\ \sigma)$, if its pdf $f_X(x)$ and cdf $F_X(x) = P(X \leq x)$ are, respectively, given by

$$f_X(x|\mu,\ \sigma) = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{2.13}$$

and

$$F_X(x) = erf\left(\frac{x-\mu}{\sqrt{2}\sigma}\right) \tag{2.14}$$

where $x \geq \mu,\ -\infty < \mu < \infty,\ \sigma > 0$, and $erf(.)$ denotes error function, (for details on half-normal distribution and its applications, see, for example, Altman (1993), Chou and Liu (1998), Bland and Altman (1999), McLaughlin (1999), Wiper et al. (2005), and references therein). Clearly, $X = \mu + \sigma|Z|$, where $Z \sim N(0,\ 1)$ has a standard normal distribution. On the other hand, the random variable $X = \mu - \sigma|Z|$ follows a negative (general) half- normal distribution. In particular, if $X \sim N\left(0,\ \sigma^2\right)$, then it is easy to see that the absolute value $|X|$ follows a half-normal distribution, with its pdf $f_{|X|}(x)$ given by

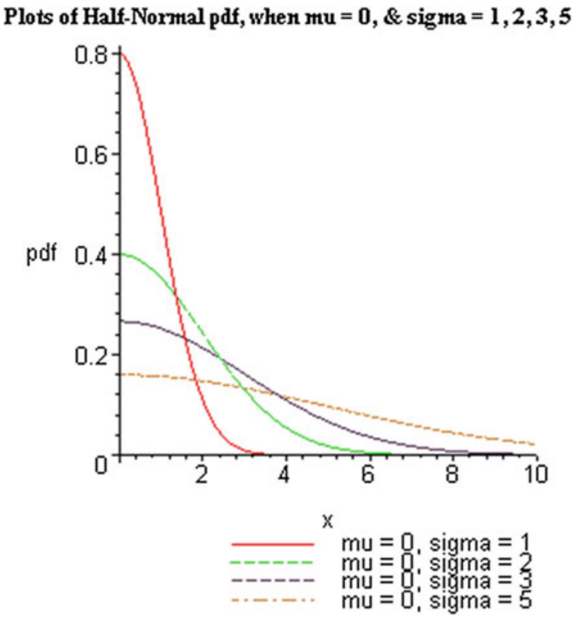

Fig. 2.5 Plots of the half-normal pdf

$$f_{|X|}(x) = \begin{cases} \frac{2}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases} \tag{2.15}$$

By taking $\sigma^2 = \frac{\pi}{2\theta^2}$ in the Eq. (2.15), more convenient expressions for the pdf and cdf of the half-normal distribution are obtained as follows

$$f_{|X|}(x) = \begin{cases} \frac{2\theta}{\pi} e^{-\left(\frac{x\theta}{\sqrt{\pi}}\right)^2} & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases} \tag{2.16}$$

and

$$F_{|X|}(x) = erf\left(\frac{\theta x}{\sqrt{\pi}}\right) \tag{2.17}$$

which are implemented in *Mathematica* software as HalfNormalDistribution[theta], see Weisstein (2007). To describe the shapes of the half-normal distribution, the plots of the pdf (2.13) for different values of the parameters μ and σ are provided in Fig. 2.5 by using Maple 10. The effects of the parameters can easily be seen from these graphs. Similar plots can be drawn for others values of the parameters.

2.2.3 *Some Properties of the Half-Normal Distribution*

This section discusses some important properties of the half-normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$.

2.2.3.1 Special Cases

The half-normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$ is a special case of the Amoroso, central chi, two parameter chi, generalized gamma, generalized Rayleigh, truncated normal, and folded normal distributions (for details, see, for example, Amoroso (1925), Patel and Read (1982), and Johnson et al. (1994), among others). It also arises as a limiting distribution of three parameter skew-normal class of distributions introduced by Azzalini (1985).

2.2.3.2 Characteristic Property

If $X \sim N(\mu, \sigma)$ is folded (to the right) about its mean, μ, then the resulting distribution is half-normal, $X|\mu, \sigma \sim HN(\mu, \sigma)$.

2.2.3.3 Mode

It is easy to see that the mode or modal value of x for which the half-normal probability density function $f_X(x)$ defined by (2.13) is maximum, is given at $x = \mu$, and the maximum value of the half-normal probability density function (2.13) is given by $f_X(\mu) = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}$. Clearly, the half-normal probability density function (2.13) is unimodal.

2.2.3.4 Moments

(i) kth **Moment of the Standardized Half-Normal Random Variable:** If the half-normal random variable X has the pdf given by the Eq. (2.13), then the standardized half-normal random variable $|Z| = \frac{X-\mu}{\sigma} \sim HN(0, 1)$ will have the pdf given by

$$f_{|Z|}(z) = \begin{cases} \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} & if\ z \geq 0 \\ 0 & if\ z < 0 \end{cases} \tag{2.18}$$

For some integer $k > 0$, and using the following integral formula (see Prudnikov et al. Vol. 1, 1986, Eq. 2.3.18.2, p. 346, or Gradshteyn and Ryzhik

1980, Eq. 3.381.4, p. 317)

$$\int_0^\infty t^{\alpha - 1} e^{-\rho t^\mu} dt = \frac{1}{\mu} \rho^{-\frac{\alpha}{\mu}} \Gamma\left(\frac{\alpha}{\mu}\right), \quad \text{where } \mu, \text{ Re } \alpha, \text{ Re } \rho > 0,$$

the kth moment of the standardized half-normal random variable Z having the pdf (2.18) is easily given by

$$E\left(Z^k\right) = \frac{1}{\sqrt{\pi}} 2^{\frac{k}{2}} \Gamma\left(\frac{k+1}{2}\right), \tag{2.19}$$

where $\Gamma(.)$ denotes gamma function.

(ii) **Moment of the Half-Normal Random Variable:** For some integer $n > 0$, the nth moment (about the origin) of the half-normal random variable X having the pdf (2.13) is easily obtained as

$$\begin{aligned} \mu_n' &= E\left(X^n\right) = E\left[(\mu + z\sigma)^n\right] = \sum_{k=0}^{n} \binom{n}{k} \mu^{n-k} \sigma^k E\left(Z^k\right) \\ &= \frac{1}{\sqrt{\pi}} \sum_{k=0}^{n} \binom{n}{k} 2^{\frac{k}{2}} \mu^{n-k} \sigma^k \Gamma\left(\frac{k+1}{2}\right) \end{aligned} \tag{2.20}$$

From the above Eq. (2.20), the first four moments of the half-normal random variable X are easily given by

$$\mu_1' = E[X] = \mu + \sigma\sqrt{\frac{2}{\pi}}, \tag{2.21}$$

$$\mu_2' = E\left[X^2\right] = \mu^2 + 2\sqrt{\frac{2}{\pi}}\mu\sigma + \sigma^2, \tag{2.22}$$

$$\mu_3' = E\left[X^3\right] = \mu^3 + 3\sqrt{\frac{2}{\pi}}\mu^2\sigma + 3\mu\sigma^2 + 2\sqrt{\frac{2}{\pi}}\sigma^3, \tag{2.23}$$

and

$$\mu_4' = E\left[X^4\right] = \mu^4 + 4\sqrt{\frac{2}{\pi}}\mu^3\sigma + 6\mu^2\sigma^2 + 8\sqrt{\frac{2}{\pi}}\mu\sigma^3 + 3\sigma^4. \tag{2.24}$$

(iii) **Central Moment of the Half-Normal Random Variable:** For some integer $n > 0$, the nth central moment (about the mean $\mu_1' = E(X)$) of the half-normal random variable X having the pdf (2.13) can be easily obtained using the formula

$$\mu_n = E\left[\left(X - \mu_1'\right)^n\right] = \sum_{k=0}^{n} \binom{n}{k} \left(-\mu_1'\right)^{n-k} E\left(X^k\right), \tag{2.25}$$

where $E\left(X^k\right) = \mu_k'$ denotes the kth moment, given by the Eq. (2.20), of the half-normal random variable X having the pdf (2.13).

Thus, from the above Eq. (2.25), the first three central moments of the half-normal random variable X are easily obtained as

$$\mu_2 = E\left[\left(X - \mu_1'\right)^2\right] = \mu_2' - \left(\mu_1'\right)^2 = \frac{\sigma^2(\pi - 2)}{\pi},$$

$$\mu_3 = \beta_3 = E\left[\left(X - \mu_1'\right)^3\right] \tag{2.26}$$

$$= \mu_3' - 3\mu_1'\mu_2' + 2\left(\mu_1'\right)^3 = \sqrt{\frac{2}{\pi}}\frac{\sigma^3(4 - \pi)}{\pi}, \tag{2.27}$$

and

$$\mu_4 = \beta_4 = E\left[\left(X - \mu_1'\right)^4\right] = \mu_4' - 4\mu_1'\mu_3' + 6\left(\mu_1'\right)^2\mu_2' - 3\left(\mu_1'\right)^4$$

$$= \frac{\sigma^4\left(3\pi^2 - 4\pi - 12\right)}{\pi^2}. \tag{2.28}$$

2.2.3.5 Mean, Variance, and Coefficients of Skewness and Kurtosis

These are easily obtained as follows:

(**i**) **Mean** : $\alpha_1 = E\,(X) = \mu + \sigma\sqrt{\frac{2}{\pi}}$;

(**ii**) **Variance** : $Var\,(X) = \mu_2 = \sigma^2\left(1 - \frac{2}{\pi}\right), \quad \sigma > 0$;

(**iii**) **Coefficient of Skewness** : $\gamma_1\,(X) = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\sqrt{2}(4 - \pi)}{\sqrt{(\pi - 2)^3}} \approx 0.9953$;

(**iv**) **Coefficient of Kurtosis** : $\gamma_2\,(X) = \frac{\mu_4}{\mu_2^2} = \frac{8(\pi - 3)}{(\pi - 2)^2} \approx 0.7614$;

2.2.3.6 Median (i.e., 50th Percentile or Second Quartile), and First and Third Quartiles

These are derived as follows. For any $p(0 < p < 1)$, the $(100p)th$ percentile (also called the quantile of order p) of the half-normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$, with the pdf $f_X(x)$ given by (2.13), is a number z_p such that the area under $f_X(x)$ to the left of z_p is p. That is, z_p is any root of the equation

$$F(z_p) = \int_{-\infty}^{z_p} f_X(t)dt = p. \tag{2.29}$$

For $p = 0.50$, we have the 50th percentile, that is, $z_{0.50}$, which is called the median (or the second quartile) of the half-normal distribution. For $p = 0.25$ and $p = 0.75$, we have the 25th and 75th percentiles respectively.

2.2.3.7 Derivation of Median(X)

Let m denote the median of the half-normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$, that is, let $m = z_{0.50}$. Then, from the Eq. (2.29), it follows that

$$0.50 = F(z_{0.50}) = \int_{-\infty}^{z_{0.50}} f_X(t)dt = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \int_{-\infty}^{z_{0.50}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt. \tag{2.30}$$

Substituting $\frac{t-\mu}{\sqrt{2}\,\sigma} = u$ in the Eq. (2.30), using the definition of error function, and solving for $z_{0.50}$, it is easy to see that

$$\begin{aligned} m = \text{Median}(X) = z_{0.50} &= \mu + \left(\sqrt{2}\right) erf^{-1}(0.50)\,\sigma \\ &= \mu + (\sqrt{2})(0.476936)\sigma \\ &\approx \mu + 0.6745\sigma, \quad \sigma > 0, \end{aligned}$$

where $erf^{-1}[0.50] = 0.476936$ has been obtained by using *Mathematica*. Note that the inverse error function is implemented in *Mathematica* as a *Built-in Symbol*, Inverse Erf[s], which gives the inverse error function obtained as the solution for z in $s = erf(z)$. Further, for details on Error and Inverse Error Functions, see, for example, Abramowitz and Stegun (1972, pp. 297–309), Gradshteyn and Ryzhik (1980), Prudnikov et al., Vol. 2 (1986), and Weisstein (2007), among others.

2.2.3.8 First and Third Quartiles

Let Q_1 and Q_3 denote the first and third quartiles of $X \sim HN(\mu, \sigma)$, that is, let $Q_1 = z_{0.25}$ and $Q_3 = z_{0.75}$. Then following the technique of the derivation of the Median(X) as in 2.2.3.7, one easily gets the Q_1 and Q_3 as follows.

(i) First Quartile: $Q_1 = \mu - 0.3186\sigma, \quad \sigma > 0;$

(ii) Third Quartile: $Q_3 = \mu + 1.150\sigma, \quad \sigma > 0.$

2.2.3.9 Mean Deviations

Following Stuart and Ord, Vol. 1, p. 52, (1994), the amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. These are known as the mean deviation about the mean and the mean deviation about the median, denoted as δ_1 and δ_2, respectively, and are defined as follows:

$$\textbf{(i)}\ \delta_1 = \int_{-\infty}^{+\infty} |x - E(X)| f(x) dx,$$

$$\textbf{(ii)}\ \delta_2 = \int_{-\infty}^{+\infty} |x - M(X)| f(x) dx.$$

Derivations of δ_1 and δ_2 for the Half-Normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$: To derive these, we first prove the following Lemma.

Lemma 2.2.1: Let $\delta = \frac{\omega - \mu}{\sigma}$. Then

$$\int_{\mu}^{\infty} \frac{1}{\sigma} |x - \omega| \sqrt{\frac{2}{\pi}} e^{-(1/2)(\frac{x-\mu}{\sigma})^2} dx$$
$$= \sigma \sqrt{\frac{2}{\pi}} \left(-1 - \delta \sqrt{\frac{\pi}{2}} + e^{-\frac{\delta^2}{2}} + \delta \sqrt{\frac{\pi}{2}} erf(\frac{\delta}{\sqrt{2}}) \right),$$

where $erf(z) = \int_0^z \frac{2}{\sqrt{\pi}} e^{-t^2} dt$ denotes the error function.

Proof: We have

$$\int_{\mu}^{\infty} \frac{1}{\sigma}|x-\omega|\sqrt{\frac{2}{\pi}}e^{-(1/2)(\frac{x-\mu}{\sigma})^2}dx$$
$$= \int_{\mu}^{\infty} \frac{|x-\mu-(\omega-\mu|}{\sigma}\sqrt{\frac{2}{\pi}}e^{-(1/2)(\frac{x-\mu}{\sigma})^2}dx$$
$$= \sigma\int_{0}^{\infty} |u-\delta|\sqrt{\frac{2}{\pi}}\, e^{-(1/2)u^2}du,$$

Substituting $\frac{x\ -\ \mu}{\sigma} = u,\ and\ \delta{=}\frac{\omega{-}\mu}{\sigma}$

$$= \sigma\int_{0}^{\delta} (\delta-u)\sqrt{\frac{2}{\pi}}e^{-(1/2)u^2}du + \sigma\int_{\delta}^{\infty} (u-\delta)\sqrt{\frac{2}{\pi}}e^{-(1/2)u^2}du$$
$$= \frac{\sigma}{\sqrt{\pi}}\left(\delta\sqrt{\pi}erf(\frac{\delta}{\sqrt{2}}) + \sqrt{2}e^{-\frac{\delta^2}{2}} - \sqrt{2}\right)$$
$$+ \frac{\sigma}{\sqrt{\pi}}\left(\delta\sqrt{\pi}erf(\frac{\delta}{\sqrt{2}}) + \sqrt{2}e^{-\frac{\delta^2}{2}} - \delta\sqrt{\pi}\right)$$
$$= \sigma\sqrt{\frac{2}{\pi}}\left(-1 - \delta\sqrt{\frac{\pi}{2}} + e^{-\frac{\delta^2}{2}} + \delta\sqrt{\frac{\pi}{2}}erf(\frac{\delta}{\sqrt{2}})\right).$$

This completes the proof of Lemma. □

Theorem 2.1: For $X|\mu,\ \sigma\ \sim\ HN(\mu,\ \sigma)$, the mean deviation, δ_1, about the mean, μ_1, is given by

$$\delta_1 = E|X-\mu_1| = \int_0^\infty |x-\mu_1|\, f(x)dx$$
$$= 2\sigma\sqrt{\frac{2}{\pi}}\left(-1 + e^{-\pi^{-1}} + erf(\pi^{-1/2})\right) \tag{2.31}$$

Proof: We have

$$\delta_1 = \int_0^\infty |x-\mu_1|\, f(x)dx$$

From Eq. (2.21), the mean of $X|\mu,\ \sigma\ \sim\ HN(\mu,\ \sigma)$ is given by

$$\mu_1 = E\ [X] = \mu + \sigma\sqrt{\frac{2}{\pi}}.$$

Taking $\omega = \mu_1$, we have

$$\delta = \frac{\omega - \mu}{\sigma} = \sqrt{\frac{2}{\pi}}.$$

Thus, taking $\omega = \mu_1$ and $\delta = \sqrt{\frac{2}{\pi}}$ in the above Lemma, and simplifying, we have

$$\delta_1 = 2\sigma\sqrt{\frac{2}{\pi}}\left(-1 + e^{-\pi^{-1}} + erf\left(\pi^{-1/2}\right)\right),$$

which completes the proof of Theorem 2.1. □

Theorem 2.2: For $X|\mu, \sigma \sim HN(\mu, \sigma)$, the mean deviation, δ_2, about the median, m, is given by

$$\begin{aligned}\delta_2 &= E|X - m| = \int_0^{\infty} |x - m| f(x)dx \\ &= \sigma\sqrt{\frac{2}{\pi}}\left(k\sqrt{\pi} - 1 + 2e^{-k^2} + 2k\sqrt{\pi}erf(k)\right),\end{aligned} \tag{2.32}$$

where $k = erf^{-1}(0.50)$.

Proof: We have

$$\delta_2 = \int_0^{\infty} |x - m| f(x)dx$$

As derived in Sect. 2.2.3.7 above, the median of $X|\mu, \sigma \sim HN(\mu, \sigma)$ is given by

$$m = \text{Median}(X) = \mu + \sqrt{2}erf^{-1}(0.50)\,\sigma = \mu + \sigma\sqrt{2}k,$$

where $k = erf^{-1}(0.50)$.
Taking $\omega = m$, we have

$$\delta = \frac{\omega - \mu}{\sigma} = \frac{m - \mu}{\sigma} = \frac{\mu + \sigma\sqrt{2}k - \mu}{\sigma} = \sqrt{2}k$$

Thus, taking $\omega = m$ and $\delta = \sqrt{2}k$ in the above Lemma, and simplifying, we have

$$\delta_2 = \sigma\sqrt{\frac{2}{\pi}}\left(k\sqrt{\pi} - 1 + 2e^{-k^2} + 2k\sqrt{\pi}erf(k)\right),$$

where $k = erf^{-1}(0.50)$. This completes the proof of Theorem 2.2. □

2.2.3.10 Renyi and Shannon Entropies, and Song's Measure of the Shape of the Half-Normal Distribution

These are derived as given below.

(i) **Renyi Entropy:** Following Renyi (1961), the entropy of the half- normal random variable X having the pdf (2.13) is given by

$$\begin{aligned}\Im_R(\gamma) &= \frac{1}{1-\gamma}\ln\int_0^{\infty}[f_X(X)]^{\gamma}\,dx,\\ &= \frac{\ln(\gamma)}{2(\gamma-1)} - \ln\left[\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}\right],\quad \sigma>0,\ \gamma>0,\ \gamma\neq 1.\end{aligned} \tag{2.33}$$

(ii) **Shannon Entropy:** Following Shannon (1948), the entropy of the half-normal random variable X having the pdf (2.13) is given by

$$H_X[f_X(X)] = E[-\ln(f_X(X)] = -\int_0^{\infty} f_X(x)\ \ln\ [f_X(x)]\ dx,$$

which is the particular case of Renyi entropy (2.31) for $\gamma \to 1$. Thus, in the limit when $\gamma \to 1$ and using L'Hospital's rule, Shannon entropy is easily obtained from the Eq. (2.33) as follows:

$$H_X[f_X(X)] = E[-\ln(f_X(X)] == \frac{1}{2} - \ln\left[\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}\right],\quad \sigma>0.$$

(iii) **Song's Measure of the Shape of a Distribution:** Following Song (2001), the gradient of the Renyi entropy is given by

$$\Im'_R(\gamma) = \frac{d}{d\gamma}\left[\Im_R(\gamma)\right] = \frac{1}{2}\left\{\frac{1}{\gamma(\gamma-1)} - \frac{\ln(\gamma)}{(\gamma-1)^2}\right\} \tag{2.34}$$

which is related to the log likelihood by

$$\Im'_R(1) = -\frac{1}{2}Var\,[\ln f(X)]\,.$$

Thus, in the limit when $\gamma \to 1$ and using L'Hospital's rule, Song's measure of the shape of the distribution of the half-normal random variable X having the pdf (2.13) is readily obtained from the Eq. (2.33) as follows:

$$\Im'_R(1) = -\frac{1}{8}\ (< 0),$$

the negative value of Song's measure indicating herein a "flat" or "platykurtic" distribution, which can be used in comparing the shapes of various densities and measuring heaviness of tails, similar to the measure of kurtosis.

2.2.3.11 Percentiles of the Half-Normal Distribution

This section computes the percentiles of the half-normal distribution, by using Maple 10. For any $p (0 < p < 1)$, the $(100p)th$ percentile (also called the quantile of order p) of the half-normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$, with the pdf $f_X(x)$ given by (2.13), is a number z_p such that the area under $f_X(x)$ to the left of z_p is p. That is, z_p is any root of the equation

$$F(z_p) = \int_{-\infty}^{z_p} f_X(t)dt = p. \tag{2.35}$$

Thus, from the Eq. (2.35), using the Maple program, the percentiles z_p of the half-normal distribution, $X|\mu, \sigma \sim HN(\mu, \sigma)$ can easily been obtained.

2.2.4 Folded Normal Distribution

An important class of probability distributions, known as the folded distributions, arises in many practical problems when only the magnitudes of deviations are recorded, and the signs of the deviations are ignored. The folded normal distribution is one such probability distribution which belongs to this class. It is related to the normal distribution in the sense that if Y is a normally distributed random variable with mean μ (location) and variance σ^2 (scale), that is, if $Y \sim N(\mu, \sigma^2)$, then the random variable $X = |Y|$ is said to have a folded normal distribution. The distribution is called folded because the probability mass (that is, area) to the left of the point $x = 0$ is folded over by taking the absolute value. As pointed out above, such a case may be encountered if only the magnitude of some random variable is recorded, without taking into consideration its sign (that is, its direction). Further, this distribution is used when the measurement system produces only positive measurements, from a normally distributed process. To fit a folded normal distribution, only the average and specified sigma (process, sample, or population) are needed. Many researchers have studied the statistical methods dealing with the properties and applications of the folded normal distribution, among them Daniel (1959), Leon et al. (1961), Elandt (1961), Nelson (1980), Patel and Read (1982), Sinha (1983), Johnson et al. (1994), Laughlin (http://www.causascientia.org/math_stat/Dists/Compendium.pdf,2001), and Kim (2006) are notable.

Definition: Let $Y \sim N(\mu, \sigma^2)$ be a normally distributed random variable with the mean μ and the variance σ^2. Let $X = |Y|$. Then X has a folded normal distribution with the pdf $f_X(x)$ and cdf $F_X(x) = P(X \leq x)$, respectively, given as follows.

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma}\left[e^{-\frac{(x-\mu)^2}{2\sigma^2}} + e^{-\frac{(-x-\mu)^2}{2\sigma^2}}\right], & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2.36}$$

Note that the μ and σ^2 are location and scale parameters for the parent normal distribution. However, they are the shape parameters for the folded normal distribution. Further, equivalently, if $x \geq 0$, using a hyperbolic cosine function, the pdf $f_X(x)$ of a folded normal distribution can be expressed as

$$f_X(x) = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\cosh\left(\frac{\mu x}{\sigma^2}\right)e^{-\frac{(x^2+\mu^2)}{2\sigma^2}}, \; x \geq 0.$$

and the cdf $\mathrm{F_X}(x)$ as

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\int_0^x \left(e^{-\frac{(y-\mu)^2}{2\sigma^2}} + e^{-\frac{(-y-\mu)^2}{2\sigma^2}}\right)dy,$$
$$x \geq 0, \; |\mu| < \infty, \; \sigma > 0. \tag{2.37}$$

Taking $z = \frac{y-\mu}{\sigma}$ in (2.37), the cdf $F_X(x)$ of a folded normal distribution can also be expressed as

$$F_X(x) = \frac{1}{\sqrt{2\pi}}\int_{-\mu/\sigma}^{(x-\mu)/\sigma}\left(e^{-\frac{1}{2}z^2} + e^{-\frac{1}{2}\left(z+\frac{2\mu}{\sigma}\right)^2}\right)dz,$$
$$z \geq 0, \; |\mu| < \infty, \; \sigma > 0, \tag{2.38}$$

where μ and σ^2 are the mean and the variance of the parent normal distribution. To describe the shapes of the folded normal distribution, the plots of the pdf (2.36) for different values of the parameters μ and σ are provided in Fig. 2.6 by using Maple 10. The effects of the parameters can easily be seen from these graphs. Similar plots can be drawn for others values of the parameters.

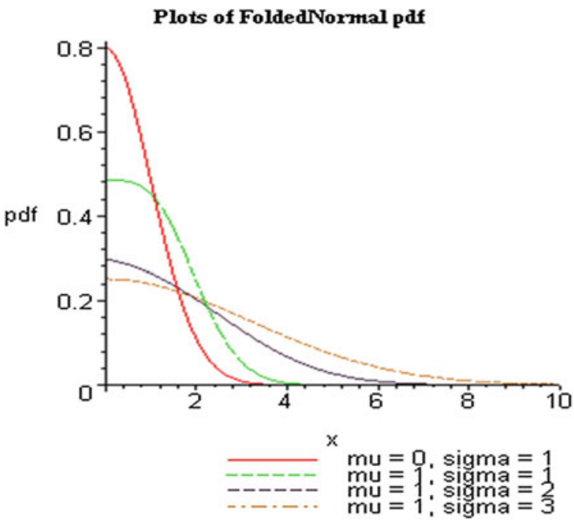

Fig. 2.6 Plots of the folded normal pdf

2.2.4.1 Some Properties of the Folded Normal Distribution

This section discusses some important properties of the folded normal distribution, $X \sim FN\left(\mu,\ \sigma^2\right)$.

2.2.4.2 Special Cases

The folded normal distribution is related to the following distributions (see, for example, Patel and Read 1982, and Johnson et al. 1994, among others).

(i) If $X \sim FN\left(\mu,\ \sigma^2\right)$, then (X/σ) has a non-central chi distribution with one degree of freedom and non-centrality parameter $\frac{\mu^2}{\sigma^2}$.

(ii) On the other hand, if a random variable U has a non-central chi distribution with one degree of freedom and non-centrality parameter $\frac{\mu^2}{\sigma^2}$, then the distribution of the random variable $\sigma\sqrt{U}$ is given by the pdf $f_X(x)$ in (2.36).

(iii) If $\mu = 0$, the folded normal distribution becomes a half-normal distribution with the pdf $f_X(x)$ as given in (2.15).

2.2.4.3 Characteristic Property

If $Z \sim N(\mu,\ \sigma)$, then $|Z| \sim FN\ (\mu,\ \sigma)$.

2.2.4.4 Mode

It is easy to see that the mode or modal value of x for which the folded normal probability density function $f_X(x)$ defined by (2.36) is maximum, is given by $x = \mu$, and the maximum value of the folded normal probability density function (2.35) is given by

$$f_X(\mu) = \frac{1}{\left(\sqrt{2\pi}\right)\sigma}\left[1 + e^{-\frac{2\mu^2}{\sigma^2}}\right]. \tag{2.39}$$

Clearly, the folded normal probability density function (2.36) is unimodal.

2.2.4.5 Moments

(i) rth **Moment of the Folded Normal Random Variable:** For some integer $r > 0$, a general formula for the rth moment, $\mu'_{f(r)}$, of the folded normal random variable $X \sim FN\left(\mu, \sigma^2\right)$ having the pdf (2.36) has been derived by Elandt (1961), which is presented here. Let $\theta = \frac{\mu}{\sigma}$. Define $\mathrm{I}_r(a) = \frac{1}{\sqrt{2\pi}}\int\limits_a^\infty y^r e^{-\frac{1}{2}y^2} dy, \quad r = 1, 2, \ldots,$ which is known as the "incomplete normal moment." In particular,

$$\mathrm{I}_0(a) = \frac{1}{\sqrt{2\pi}}\int\limits_a^\infty e^{-\frac{1}{2}y^2} dy = 1 - \Phi(a), \tag{2.40}$$

where $\Phi(a) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^a e^{-\frac{1}{2}y^2} dy$ is the CDF of the unit normal $N(0, 1)$. Clearly, for $r > 0$, $\mathrm{I}_r(a) = \left(\frac{1}{\sqrt{2\pi}}\right) a^{r-1} e^{-\frac{1}{2}a^2} + (r-1)\,\mathrm{I}_{r-2}(a)$. Thus, in view of these results, the rth moment, $\mu'_{f(r)}$, of the folded normal random variable X is easily expressed in terms of the I_r function as follows.

$$\begin{aligned}\mu'_{f(r)} &= E\left(X^r\right) = \int\limits_0^\infty x f_X(x)\, dx \\ &= \left(\sigma^r\right)\sum_{j=0}^{r}\binom{r}{j}\theta^{r-j}\left[\mathrm{I}_j(-\theta) + (-1)^{r-j}\,\mathrm{I}_j(\theta)\right].\end{aligned} \tag{2.41}$$

From the above Eq. (2.41) and noting, from the definition of the I_r function, that $\mathrm{I}_2(-\theta) - \mathrm{I}_2(\theta) = -\left[\left(\frac{2}{\sqrt{2\pi}}\right)\theta e^{-\frac{1}{2}\theta^2} + \{1 - 2\mathrm{I}_0(-\theta)\}\right]$, the first four moments of the folded normal random distribution are easily obtained as follows.

$$\begin{aligned}
\mu'_{f(1)} &= E\,[X] = \mu_f = \left(\frac{2}{\sqrt{2\pi}}\right)\sigma e^{-\frac{1}{2}\theta^2} - \mu\,[1 - 2\mathrm{I}_0\,(-\theta)] \\
&= \left(\frac{2}{\sqrt{2\pi}}\right)\sigma e^{-\frac{1}{2}\theta^2} - \mu\,[1 - 2\Phi\,(\theta)]\,, \\
\mu'_{f(2)} &= E\left[X^2\right] = \sigma_f^2 = \mu^2 + \sigma^2, \\
\mu'_{f(3)} &= E\left[X^3\right] = \left(\mu^2 + 2\sigma^2\right)\mu_f - \mu\sigma^2\,[1 - 2\Phi\,(\theta)]\,,
\end{aligned}$$

and

$$\mu'_{f(4)} = E\left[X^4\right] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4. \tag{2.42}$$

(ii) **Central Moments of the Folded Normal Random Variable:** For some integer $n > 0$, the nth central moment (about the mean $\mu'_{f(1)} = E\,(X)$) of the folded normal random variable X having the pdf (2.36) can be easily obtained using the formula

$$\mu_{f(n)} = E\left[\left(X - \mu'_{f(1)}\right)^n\right] = \sum_{r=0}^{n}\binom{n}{r}\left(-\mu'_{f(1)}\right)^{n-r} E\left(X^r\right), \tag{2.43}$$

where $E\ (X^r) = \mu'_{f(r)}$ denotes the rth moment, given by the Eq. (2.41), of the folded normal random variable X. Thus, from the above Eq. (2.43), the first four central moments of the folded normal random variable X are easily obtained as follows.

$$\begin{aligned}
\mu_{f(1)} &= 0, \\
\mu_{f(2)} &= \mu^2 + \sigma^2 - \mu_f^2, \\
\mu_{f(3)} &= \beta_3 = 2\left[\mu_f^3 - \mu^2\mu_f - \left(\frac{\sigma^3}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}\theta^2}\right],
\end{aligned}$$

and

$$\begin{aligned}
\mu_{f(4)} = \beta_4 = &\left(\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4\right) \\
&+ \left(\frac{8\sigma^3}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}\theta^2}\mu_f + 2\left(\mu^2 - 3\sigma^2\right)\mu_f^2 - 3\mu_f^4.
\end{aligned} \tag{2.44}$$

2.2.4.6 Mean, Variance, and Coefficients of Skewness and Kurtosis of the Folded Normal Random Variable

These are easily obtained as follows:

(i) **Mean:** $E\ (X) = \alpha_1 = \mu_f = \left(\frac{2}{\sqrt{2\pi}}\right)\sigma e^{-\frac{1}{2}\theta^2} - \mu\,[1 - 2\Phi\,(\theta)]$,

(ii) **Variance:** $Var\ (X) = \beta_2 = \mu_{f(2)} = \mu^2 + \sigma^2 - \mu_f^2, \quad \sigma > 0$,

(iii) **Coefficient of Skewness:** $\gamma_1\ (X) = \frac{\mu_3}{[\mu_2]^{\frac{3}{2}}}$,

(iv) **Coefficient of Kurtosis:** $\gamma_2\ (X) = \frac{\mu_{f(4)}}{[\mu_{f(2)}]^2}$,

where the symbols have their usual meanings as described above.

2.2.4.7 Percentiles of the Folded Normal Distribution

This section computes the percentiles of the folded normal distribution, by using Maple 10. For any $p(0 < p < 1)$, the $(100p)th$ percentile (also called the quantile of order p) of the folded normal distribution, $X \sim FN\left(\mu,\ \sigma^2\right)$, with the pdf $f_X(x)$ given by (2.36), is a number z_p such that the area under $f_X(x)$ to the left of z_p is p. That is, z_p is any root of the equation

$$F(z_p) = \int_{-\infty}^{z_p} f_X(t)dt = p. \tag{2.45}$$

Thus, from the Eq. (2.45), using the Maple program, the percentiles z_p of the folded normal distribution can be computed for some selected values of the parameters.

Note: For the tables of the folded normal cdf $F_X(x) = P(X \leq x)$ for different values of the parameters, for example, $\frac{\mu_f}{\sigma_f}$ = 1.3236, 1.4(0.1)3, and x = 0.1(0.1)7, the interested readers are referred to Leon et al. (1961).

Note: As noted by Elandt (1961), the family of the folded normal distributions, $N_f\left(\mu_f\ ,\ \sigma_f\right)$, is included between the half-normal, for which $\frac{\mu_f}{\sigma_f}$ = 1.3237, and the normal, for which $\frac{\mu_f}{\sigma_f}$ is infinite. Approximate normality is attained if , for which $\frac{\mu_f}{\sigma_f} > 3$.

2.2.5 *Truncated Distributions*

Following Rohatgi and Saleh (2001), and Lawless (2004), we first present an overview of the truncated distributions.

2.2.5.1 Overview of Truncated Distributions

Suppose we have a probability distribution defined for a continuous random variable X. If some set of values in the range of X are excluded, then the probability distribution for the random variable X is said to be truncated. We defined the truncated distributions as follows.

Definition: Let X be a continuous random variable on a probability space $(\Omega,\ S,\ P)$, and let $T \in$ B such that $0 < P\{X \in T\} < 1$, where B is a σ- field on the set of real numbers $\Re$. Then the conditional distribution $P\{X \le x \mid X \in T\}$, defined for any real x, is called the truncated distribution of X. Let $f_X(x)$ and $F_X(x)$ denote the probability density function (pdf) and the cumulative distribution function (cdf), respectively, of the parent random variable X. If the random variable with the truncated distribution function $P\{X \le x \mid X \in T\}$ be denoted by Y, then Y has support T. Then the cumulative distribution function (cdf), say, $G(y)$, and the probability density function (pdf), say, $g(y)$, for the random variable Y are, respectively, given by

$$G_Y(y) = P\{Y \le y \mid Y \in T\} = \frac{P\{Y \le y,\ Y \in T\}}{P\{Y \in T\}} = \frac{\int\limits_{(-\infty,\ y] \cap T} f_X(u)du}{\int\limits_T f_X(u)du}, \tag{2.46}$$

and

$$g_Y(y) = \begin{cases} \frac{f_X(y)}{\int\limits_T f_X(u)du}, & y \in T \\ 0, & y \notin T. \end{cases} \tag{2.47}$$

Clearly $g_Y(y)$ in (2.47) defines a pdf with support T, since $\int_T g_Y(y)dy = \frac{\int\limits_T f_X(y)dy}{\int\limits_T f_X(u)du} = 1$. Note that here T is not necessarily a bounded set of real numbers. In particular, if the values of Y below a specified value a are excluded from the distribution, then the remaining values of Y in the population have a distribution with the pdf given by $g_L(y;\ a) = \frac{f_X(y)}{1 - F_X(a)},\ a \le y < \infty$, and the distribution is said to be left truncated at a. Conversely, if the values of Y above a specified value a are excluded from the distribution, then the remaining values of Y in the population have a distribution with the pdf given by $g_R(y;\ a) = \frac{f_X(y)}{F_X(a)},\ 0 \le y \le a$, and the distribution is said to be right truncated at a. Further, if Y has a support $T = [a_1,\ a_2]$, where $-\infty < a_1 < a_2 < \infty$, then the conditional distribution of Y, given that $a_1 \le y \le a_2$, is called a doubly truncated distribution with the cdf, say, $G(y)$, and the pdf, say, $g(y)$, respectively, given by

$$G_Y(y) = \frac{F_X\{\max(\min(y,\ a_2),\ a_1)\} - F_X(a_1)}{F_X(a_2) - F_X(a_1)}, \tag{2.48}$$

and

$$g_Y(y) = \begin{cases} \frac{f_X(y)}{F_X(a_2) - F_X(a_1)}, & y \in [a_1,\ a_2] \\ 0, & y \notin [a_1,\ a_2]. \end{cases} \tag{2.49}$$

The truncated distribution for a continuous random variable is one of the important research topics both from the theoretical and applications point of view. It arises in many probabilistic modeling problems of biology, crystallography, economics, engineering, forecasting, genetics, hydrology, insurance, lifetime data analysis, management, medicine, order statistics, physics, production research, psychology, reliability, quality engineering, survival analysis, etc, when sampling is carried out from an incomplete population data. For details on the properties and estimation of parameters of truncated distributions, and their applications to the statistical analysis of truncated data, see, for example, Hald (1952), Chapman (1956), Hausman and Wise (1977), Thomopoulos (1980), Patel and Read (1982), Levy (1982), Sugiura and Gomi (1985), Schneider (1986), Kimber and Jeynes (1987), Kececioglu (1991), Cohen (1991), Andersen et al. (1993), Johnson et al. (1994), Klugman et al. (1998), Rohatgi and Saleh (2001), Balakrishnan and Nevzorov (2003), David and Nagaraja (2003), Lawless (2003), Jawitz (2004), Greene (2005), Nadarajah and Kotz (2006a), Maksay and Stoica (2006) and Nadarajah and Kotz (2007) and references therein.

The truncated distributions of a normally distributed random variable, their properties and applications have been extensively studied by many researchers, among them Bliss (1935 for the probit model which is used to model the choice probability of a binary outcome), Hald (1952), Tobin (1958) for the probit model which is used to model censored data), Shah and Jaiswal (1966), Hausman and Wise (1977), Thomopoulos (1980), Patel and Read (1982), Levy (1982), Sugiura and Gomi (1985), Schneider (1986), Kimber and Jeynes (1987), Cohen (1959, 1991), Johnson et al. (1994), Barr and Sherrill (1999), Johnson (2001), David and Nagaraja (2003), Jawitz (2004), Nadarajah and Kotz (2007), and Olive (2007), are notable. In what follows, we present the pdf, moment generating function (mgf), mean, variance and other properties of the truncated normal distribution most of which is discussed in Patel and Read (1982), Johnson et al. (1994), Rohatgi and Saleh (2001), and Olive (2007).

Definition: Let $X \sim N\left(\mu,\ \sigma^2\right)$ be a normally distributed random variable with the mean μ and the variance σ^2. Let us consider a random variable Y which represents the truncated distribution of X over a support T = [a, b], where $-\infty < a < b < \infty$. Then the conditional distribution of Y, given that $a \le y \le b$, is called a doubly truncated normal distribution with the pdf, say, $g_Y(y)$, given by

$$g_Y(y) = \begin{cases} \dfrac{\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{\left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]}, & y \in [a,\ b] \\ 0, & y \notin [a,\ b] \end{cases}, \tag{2.50}$$

where $\phi(.)$ and $\Phi(.)$ are the pdf and cdf of the standard normal distribution, respectively. If $a = -\infty$, then the we have a (singly) truncated normal distribution from above, (that is, right truncated). On the other hand, if $b = \infty$, then the we have a (singly) truncated normal distribution from below, (that is, left truncated). The following are some examples of the truncated normal distributions.

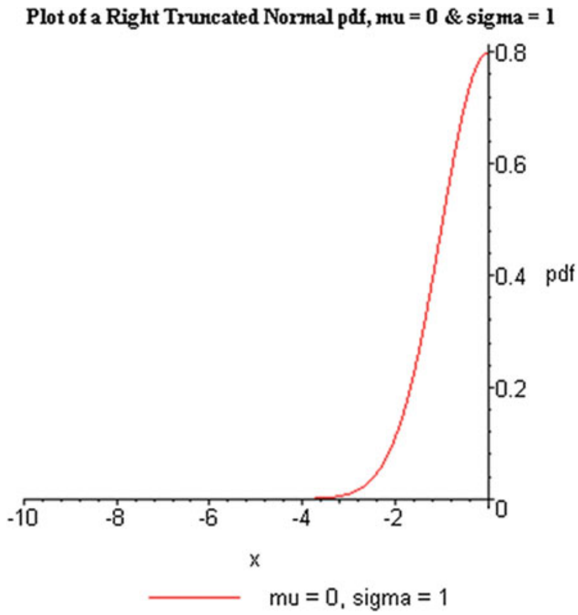

Fig. 2.7 Example of a right truncated normal distribution

(i) **Example of a Left Truncated Normal Distribution:** Taking $a = 0$, $\mathrm{b} = \infty$, and $\mu = 0$, the pdf $g_Y(y)$ in (2.50) reduces to that of the half normal distribution in (2.15), which is an example of the left truncated normal distribution.

(ii) **Example of a Right Truncated Normal Distribution:** Taking $a = -\infty, b = 0$, $\mu = 0$, and $\sigma = 1$ in (2.50), the pdf $g_Y(y)$ of the right truncated normal distribution is given by

$$g_Y(y) = \begin{cases} 2\phi(y), & -\infty < y \leq 0 \\ 0, & y > 0 \end{cases}, \tag{2.51}$$

where $\phi(.)$ is the pdf of the standard normal distribution. The shape of right truncated normal pdf $g_Y(y)$ in (2.51) is illustrated in the following Fig. (2.7).

2.2.5.2 MGF, Mean, and Variance of the Truncated Normal Distribution

These are given below.

(A) Moment Generating Function: The mgf of the doubly truncated normal distribution with the pdf $g_Y(y)$ in (2.50) is easily obtained as

$$M(t) = E\left(e^{tY}|Y \in [a, b]\right)$$
$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} \left\{ \frac{\left[\Phi\left(\frac{b-\mu}{\sigma} - \sigma t\right) - \Phi\left(\frac{a-\mu}{\sigma} - \sigma t\right)\right]}{\left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]} \right\} \tag{2.52}$$

(B) Mean, Second Moment and Variance: Using the expression for the mgf (2.52), these are easily given by

(i)
$$Mean = E(Y|Y \in [a, b]) = M'(t)\big|_{t=0}$$
$$= \mu + \sigma \left[\frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} \right] \tag{2.53}$$

Particular Cases:

(I) If $b \rightarrow \infty$ in (2.52), then we have

$$E(Y|Y > a) = \mu + \sigma h,$$

where $h = \frac{\phi\left(\frac{a-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)}$ is called the Hazard Function (or the Hazard Rate, or the Inverse Mill's Ratio) of the normal distribution.

(II) If $a \rightarrow -\infty$ in (2.53), then we have

$$E(Y|Y < b) = \mu - \sigma \left[\frac{\phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)} \right]. \tag{2.54}$$

(III) If $b \rightarrow \infty$ in (2.54), then Y is not truncated and we have

$$E(Y) = \mu \quad .$$
$$V(Y) = \sigma^2[1 + \alpha\phi]$$

(ii)
$$Second\ Moment = E\left(Y^2|Y \in [a, b]\right) = M''(t)\big|_{t=0}$$
$$= 2\mu \{E(Y|Y \in [a, b])\} - \mu^2$$
$$= \mu^2 + 2\mu\sigma \left[\frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} \right]$$

$$+\sigma^2\left[1+\frac{\left(\frac{a-\mu}{\sigma}\right)\phi\left(\frac{a-\mu}{\sigma}\right)-\left(\frac{b-\mu}{\sigma}\right)\phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}\right] \tag{2.55}$$

and

$$\begin{aligned}\textbf{(iii)}\quad \mathit{Variance} = \mathit{Var}\,(Y|Y\in[a,\ b]) &= \left\{E\left(Y^2|Y\in[a,\ b]\right)\right\}\\ &\quad-\{E\,(Y|Y\in[a,\ b])\}^2\\ &= \sigma^2\left\{1+\frac{\left(\frac{a-\mu}{\sigma}\right)\phi\left(\frac{a-\mu}{\sigma}\right)-\left(\frac{b-\mu}{\sigma}\right)\phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}\right.\\ &\quad\left.-\left[\frac{\phi\left(\frac{b-\mu}{\sigma}\right)-\phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}\right]^2\right\}\end{aligned} \tag{2.56}$$

Some Further Remarks on the Truncated Normal Distribution:

(i) Let $Y \sim TN\left(\mu,\ \sigma^2,\ a=\mu-k\sigma,\ b=\mu+k\sigma\right)$, for some real k, be the truncated version of a normal distribution with mean μ and variance σ^2. Then, from (2.53) and (2.56), it easily follows that $E\,(Y) = \mu$ and $Var\,(Y) = \sigma^2\left\{1-\frac{2k\phi(k)}{2k\Phi(k)-1}\right\}$, (see, for example, Olive, 2007).

(ii) The interested readers are also referred to Shah and Jaiswal (1966) for some nice discussion on the pdf $g_Y\,(y)$ of the truncated normal distribution and its moments, when the origin is shifted at a.

(iii) A table of the mean μ_t, standard deviation σ_t, and the ratio (mean deviation/σ_t) for selected values of $\Phi\left(\frac{a-\mu}{\sigma}\right)$ and $1-\Phi\left(\frac{b-\mu}{\sigma}\right)$ have been provided in Johnson and Kotz (1994).

2.2.6 Inverse Normal (Gaussian) Distribution (IGD)

The inverse Gaussian distribution (IGD) represents a class of distribution. The distribution was initially considered by Schrondinger (1915) and further studied by many authors, among them Tweedie (1957a, b) and Chhikara and Folks (1974) are notable. Several advantages and applications in different fields of IGD are given by Tweedie (1957), Johnson and Kotz (1994), Chhikara and Folks (1974, 1976,1977), and Folks and Chhikara (1978), among others. For the generalized inverse Gaussian distribution (GIG) and its statistical properties, the interested readers are referred to Good (1953), Sichel (1974, 1975), Barndorff-Nielsen (1977, 1978), Jorgensen

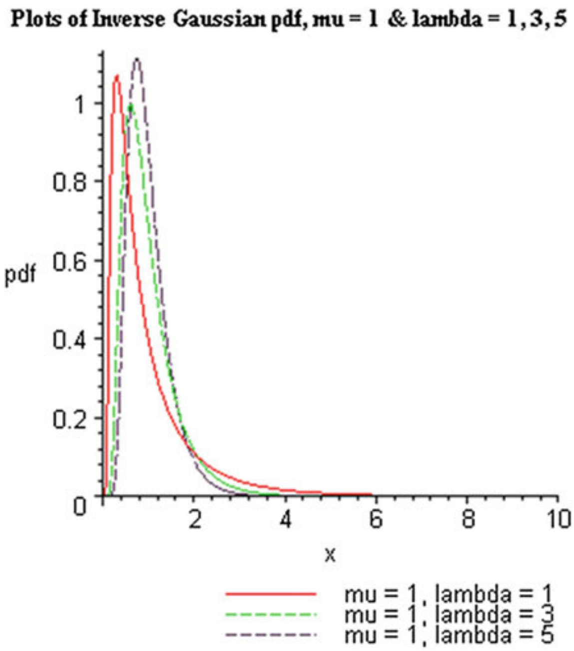

Fig. 2.8 Plots of the inverse Gaussian pdf

(1982), and Johnson and Kotz (1994), and references therein. In what follows, we present briefly the pdf, cdf, mean, variance and other properties of the inverse Gaussian distribution (IGD).

Definition: The pdf of the Inverse Gaussian distribution (IGD) with parameters μ and λ is given by

$$f(x,\ \mu,\ \lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left\{-\frac{\lambda}{2\mu^2 x}(x - \mu)^2\right\} \quad x > 0, \mu > 0, \lambda > 0 \tag{2.57}$$

where μ is location parameter and λ is a shape parameter. The mean and variance of this distribution are μ and μ^3/λ respectively. To describe the shapes of the inverse Gaussian distribution, the plots of the pdf (2.57), for $\mu = 1$ and $\lambda = 1, 3, 5$ are provided in Fig. 2.8 by using Maple 10. The effects of the parameters can easily be seen from these graphs. Similar plots can be drawn for others values of the parameters.

Properties of IGD:

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from the inverse Gaussian distribution (1.1). The maximum likelihood estimators (MLE's) for μ and λ are respectively given by

$$\hat{\mu} = \bar{x} = \sum_{i=1}^{n} x_i/n,\ \tilde{\lambda} = \frac{n}{V},\ \text{where } V = \sum_{i=1}^{n}\left(\frac{1}{x_i} - \frac{1}{\bar{x}}\right).$$

It is well known that

(i) the sample mean $\bar{x}$ is unbiased estimate of μ where as $\tilde{\lambda}$ is a biased estimate of λ.
(ii) $\bar{x}$ follows IGD with parameters μ and $n\,\lambda$, whereas $\lambda\,V$ is distributed as chi –square distribution with (n-1) degrees of freedom
(iii) $\bar{x}$ and V are stochastically independent and jointly sufficient for (μ,λ) if both are unknown.
(iv) the uniformly minimum variance unbiased estimator (UMVUE) of λ is $\hat{\lambda} = (n-3)/V$ and $Var\left(\hat{\lambda}\right) = 2\,\lambda^2/(n-5) = MSE\left(\hat{\lambda}\right)$.

2.2.7 Skew Normal Distributions

This section discusses the univariate skew normal distribution (SND) and some of its characteristics. The skew normal distribution represents a parametric class of probability distributions, reflecting varying degrees of skewness, which includes the standard normal distribution as a special case. The skewness parameter involved in this class of distributions makes it possible for probabilistic modeling of the data obtained from skewed population. The skew normal distributions are also useful in the study of the robustness and as priors in Bayesian analysis of the data. It appears from the statistical literatures that the skew normal class of densities and its applications first appeared indirectly and independently in the work of Birnbaum (1950), Roberts (1966), O'Hagan and Leonard (1976), and Aigner et al. (1977). The term skew normal distribution (SND) was introduced by Azzalini (1985, 1986), which give a systematic treatment of this distribution, developed independently from earlier work. For further studies, developments, and applications, see, for example, Henze (1986), Mukhopadhyay and Vidakovic (1995), Chiogna (1998), Pewsey (2000), Azzalini (2001), Gupta et al. (2002), Monti (2003), Nadarajah and Kotz (2003), Arnold and Lin (2004), Dalla Valle (2004), Genton (2004), Arellano-Valle et al. (2004), Buccianti (2005), Azzalini (2005, 2006), Arellano-Valle and Azzalini (2006), Bagui and Bagui (2006), Nadarajah and Kotz (2006), Shkedy et al. (2006), Pewsey (2006), Fernandes et al. (2007), Mateu-Figueras et al. (2007), Chakraborty and Hazarika (2011), Eling (2011), Azzalini and Regoli (2012), among others. For generalized skew normal distribution, the interested readers are referred to Gupta and Gupta (2004), Jamalizadeh, et al. (2008), and Kazemi et al. (2011), among others. Multivariate versions of SND have also been proposed, among them Azzalini and Dalla Valle (1996), Azzalini and Capitanio (1999), Arellano-Valle et al. (2002), Gupta and Chen (2004), and Vernic (2006) are notable. Following Azzalini (1985, 1986, 2006), the definition and some properties, including some graphs, of the univariate skew normal distribution (SND) are presented below.

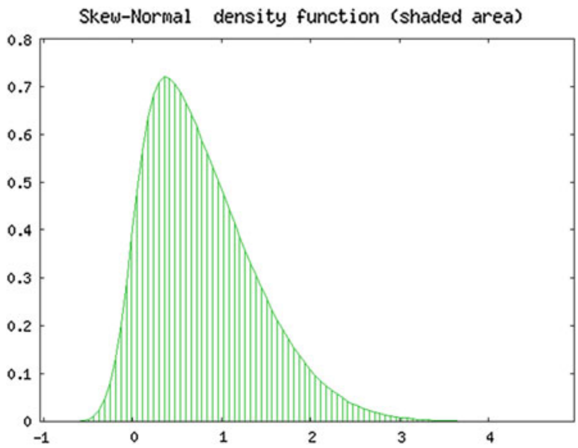

Fig. 2.9 Plot of the skew normal pdf: ($\mu = 0$, $\sigma = 1$, $\lambda = 5$)

Definition: For some real-valued parameter λ, a continuous random variable X_λ is said to have a skew normal distribution, denoted by $X_\lambda \sim SN\ (\lambda)$, if its probability density function is given by

$$f_X\,(x;\ \lambda) = 2\,\phi\,(x)\,\Phi\,(\lambda x)\,,\ -\infty < x < \infty,\,, \tag{2.58}$$

where $\phi\,(x) = \left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{1}{2}x^2}$ and $\Phi\,(\lambda x) = \int\limits_{-\infty}^{\lambda x} \phi\,(t)\,dt$ denote the probability density function and cumulative distribution function of the standard normal distribution respectively.

2.2.7.1 Shapes of the Skew Normal Distribution

The shape of the skew normal probability density function given by (2.58) depends on the values of the parameter λ. For some values of the parameters (μ, σ, λ), the shapes of the pdf (2.58) are provided in Figs. 2.9, 2.10 and 2.11. The effects of the parameter can easily be seen from these graphs. Similar plots can be drawn for others values of the parameters.

Remarks: The continuous random variable X_λ is said to have a skew normal distribution, denoted by $X_\lambda \sim SN\,(\lambda)$, because the family of distributions represented by it includes the standard $N\,(0,\ 1)$ distribution as a special case, but in general its members have a skewed density. This is also evident from the fact that $X_\lambda^2 \sim \chi^2$ for all values of the parameter λ. Also, it can be easily seen that the skew normal density function $f_X\,(x;\ \lambda)$ has the following characteristics:

1. when $\lambda = 0$, we obtain the standard normal density function $f_X\,(x; 0)$ with zero skewness;
2. as $|\lambda|$ increases, the skewness of the skew normal distribution also increases;

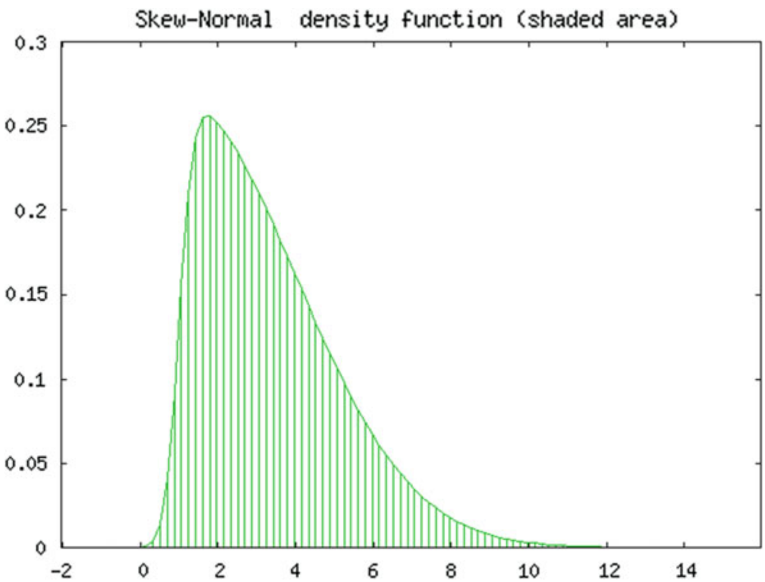

Fig. 2.10 Plot of the skew normal pdf: ($\mu = 1,\ \sigma = 3,\ \lambda = 10$)

3. when $|\lambda| \to \infty$, the skew normal density function $f_X(x;\ \lambda)$ converges to the half-normal (or folded normal) density function;
4. if the sign of λ changes, the skew normal density function $f_X(x;\ \lambda)$ is reflected on the opposite side of the vertical axis.

2.2.7.2 Some Properties of Skew Normal Distribution

This section discusses some important properties of the skew normal distribution, $X_\lambda \sim SN(\lambda)$.
Properties of $SN(\lambda)$:

(a) $SN(0) = N(0,\ 1)$.
(b) If $X_\lambda \sim SN(\lambda)$, then $-X_\lambda \sim SN(-\lambda)$.
(c) If $\lambda \to \pm\infty$, and $Z \sim N(0,\ 1)$, then $SN(\lambda) \to \pm|Z| \sim HN(0,\ 1)$, that is, $SN(\lambda)$ tends to the half-normal distribution.
(d) If $X_\lambda \sim SN(\lambda)$, then $X_\lambda^2 \sim \chi^2$.
(e) The MGF of X_λ is given by $M_\lambda(t) = 2e^{\frac{t^2}{2}}\Phi(\delta t)$, $t \in \Re$, where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.
(f) It is easy to see that $E(X_\lambda) = \delta\left(\sqrt{\frac{2}{\pi}}\right)$, and $Var(X_\lambda) = \frac{\pi - 2\delta^2}{\pi}$.
(g) The characteristic function of X_λ is given by $\psi_\lambda(t) = e^{\frac{-t^2}{2}}[1 + ih(\delta t)]$, $t \in \Re$, where $h(x) = \left(\sqrt{\frac{2}{\pi}}\right)\int_0^x e^{\frac{y^2}{2}}dy$ and $h(-x) = -h(x)$ for $x \geq 0$.

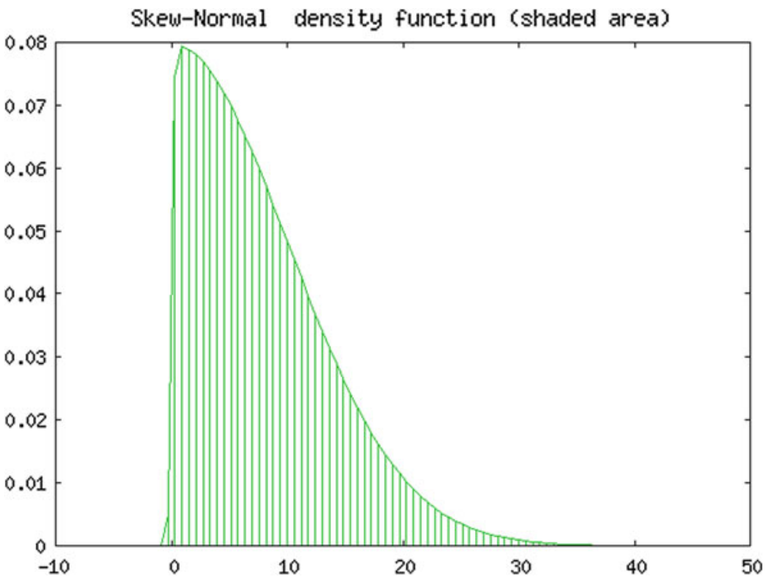

Fig. 2.11 Plots of the skew normal pdf: ($\mu = 0,\ \sigma = 10,\ \lambda = 50$)

(h) By introducing the following linear transformation $Y = \mu + \sigma X$, that is, $X = \frac{Y-\mu}{\sigma}$, where $\mu \geq 0,\ \sigma > 0$, we obtain a skew-normal distribution with parameters $(\mu,\ \sigma,\ \lambda)$, denoted by $Y \sim SN\left(\mu,\ \sigma^2,\ \lambda\right)$, if its probability density function is given by

$$f_Y(y;\ \mu,\ \sigma,\ \lambda) = 2\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left(\frac{\lambda(y-\mu)}{\sigma}\right),\ -\infty < y < \infty, \tag{2.59}$$

where $\phi(y)$ and $\Phi(\lambda y)$ denote the probability density function and cumulative distribution function of the normal distribution respectively, and $\mu \geq 0,\ \sigma > 0$ and $-\infty < \lambda < \infty$ are referred as the location, the scale and the shape parameters respectively. Some characteristic values of the random variable Y are as follows:

I. Mean: $E\,(Y) \;=\; \mu \;+\; \left(\sigma\delta\sqrt{\frac{2}{\pi}}\right)$

II. Variance: $Var\,(Y) \;=\; \dfrac{\sigma^2\left(\pi \;-\; 2\delta^2\right)}{\pi}$

III. Skewness: $\gamma_1 \;=\; \left(\dfrac{4 \;-\; \pi}{2}\right)\dfrac{[E\,(X_\lambda)]^3}{[Var\,(X_\lambda)]^{\frac{3}{2}}}$

IV. Kurtosis: $\gamma_2 \;=\; 2\;(\pi \;-\; 3)\,\dfrac{[E\,(X_\lambda)]^4}{[Var\,(X_\lambda)]^2}$

2.2.7.3 Some Characteristics Properties of Skew Normal Distribution

Following Gupta et al. (2004), some characterizations of the skew normal distribution (SND) are stated below.

(i) Let X_1 and X_2 be *i.i.d.* F, an unspecified distribution which admits moments of all order. Then $X_1^2 \;\sim\; \chi_1^2$, $X_2^2 \;\sim\; \chi_1^2$, and $\frac{1}{2}\,(X_1 \;+\; X_2)^2 \;\sim\; H_0\,(\lambda)$ if and only if $F \;=\; SN\,(\lambda)$ or $F \;=\; SN\,(-\lambda)$ where $H_0\,(\lambda)$ is the distribution of $\frac{1}{2}\,(X \;+\; Y)^2$ when X and Y are *i.i.d.* $SN\,(\lambda)$.

(ii) Let $H_0\,(\lambda)$ be the distribution of $(Y \;+\; a)^2$ where $Y \;\sim\; SN\,(\lambda)$ and $a \;\neq\; 0$ is a given constant. Let X be a random variable with a distribution that admits moments of all order. Then $X^2 \;\sim\; \chi_1^2$, $(X \;+\; a)^2 \;\sim\; H_0\,(\lambda)$ if and only if $X \;\sim\; SN\,(\lambda)$ for some λ.

For detailed derivations of the above and more results on other characterizations of the skew normal distribution (SND), see Gupta et al. (2004) and references therein. The interested readers are also referred to Arnold and Lin (2004), where the authors have shown that the skew-normal distributions and their limits are exactly the distributions of order statistics of bivariate normally distributed variables. Further, using generalized skew-normal distributions, the authors have characterized the distributions of random variables whose squares obey the chi-square distribution with one degree of freedom.

2.3 Goodness-of-Fit Test (Test For Normality)

The goodness of fit (or GOF) tests are applied to test the suitability of a random sample with a theoretical probability distribution function. In other words, in the GOF test analysis, we test the hypothesis if the random sample drawn from a population follows a specific discrete or continuous distribution. The general approach for this is to first determine a test statistic which is defined as a function of the data measuring the distance between the hypothesis and the data. Then, assuming the hypothesis is true, a probability value of obtaining data which have a larger value of the test statistic than the value observed, is determined, which is known as the p-value. Smaller p-values (for example, less than 0.01) indicate a poor fit of the distribution. Higher values of p (close to one) correspond to a good fit of the distribution. We consider the following parametric and non-parametric goodness-of-fit tests

2.3.1 χ^2(Chi-Squared) Test

The χ^2 test, due to Karl Pearson, may be applied to test the fit of any specified continuous distribution to the given randomly selected continuous data. In χ^2 analysis, the data is first grouped into, say, k number of classes of equal probability. Each class should contain at least 5 or more data points. The χ^2 test statistic is given by

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{2.60}$$

where O_i is the observed frequency in class i, $i = 1, \ldots, k$ and E_i is the expected frequency in class i, if the specified distribution were the correct one, and is given by

$$E_i = F(x_i) - F(x_{i-1}),$$

where $F(x)$ is the cumulative distribution function (CDF) of the probability distribution being tested, and x_i, x_{i-1} are the limits for the class i. The null and alternative hypotheses being tested are, respectively, given by:

H_0: The data follow the specified continuous distribution;
H_1: The data do not follow the specified continuous distribution.

The null hypothesis (H_0) is rejected at the chosen significance level, say, α, if the test statistic is greater than the critical value denoted by $\chi^2{}_{1-\alpha,\, k-1}$, with $k - 1$ degrees of freedom (df) and a significance level of α. If r parameters are estimated from the data, df are $k - r - 1$.

2.3.2 Kolmogorov-Smirnov (K-S) Test

This test may also be applied to test the goodness of fit between a hypothesized cumulative distribution function (CDF) $F(x)$ and an empirical CDF $F_n(x)$. Let $y_1 < y_2 < \ldots < y_n$ be the observed values of the order statistics of a random sample $x_1, x_2, \ldots, x_n$ of size n. When no two observations are equal, the empirical CDF $F_n(x)$ is given by, see Hogg and Tanis (2006),

$$F_n(x) = \begin{cases} 0, & x < y_1, \\ \frac{i}{n}, & y_i \le x < y_{i+1}, \quad i = 1, 2, \ldots, n-1, \\ 1, & y_n \le x. \end{cases} \tag{2.61}$$

Clearly,

$$F_n(x) = \frac{1}{n}\,[\text{Number of Observations} \le x].$$

Following Blischke and Murthy (2000), the Kolmogorov-Smirnov test statistic, D_n, is defined as the maximum distance between the hypothesized CDF $F(x)$ and the empirical CDF $F_n(x)$, and is given by

$$D_n = \max\left\{D_n^+, D_n^-\right\},$$

where

$$D_n^+ = \max_{i = 1, 2, \ldots, n}\left[\frac{i}{n} - F_n(y_i)\right]$$

and

$$D_n^- = \max_{i = 1, 2, \ldots, n}\left[F_n(y_i) - \frac{i-1}{n}\right].$$

For calculations of fractiles (percentiles) of the distribution of D_n, the interested readers are referred to Massey (1951). In Stephens (1974), one can find a close approximation of the fractiles of the distribution of D_n, based on a constant denoted by d_α which is a function of n only. The values of d_α can also be found in Table 11.2 on p. 400 of Blischke and Murthy (2000) for $\alpha = 0.15,\ 0.10,\ 0.05$, and 0.01. The critical value of D_n is calculated by the formula $d_\alpha / \left(\sqrt{n} + \frac{0.11}{\sqrt{n}} + 0.12\right)$. The null and alternative hypotheses being tested are, respectively, given by:

H_0: The data follow the specified continuous distribution;
H_1: The data do not follow the specified continuous distribution.

The null hypothesis (H_0) is rejected at the chosen significance level, say, α, if the Kolmogorov-Smirnov test statistic, D_n, is greater than the critical value calculated by the above formula.

2.3.3 Anderson-Darling (A-D) Test

The **Anderson-Darling test** is also based on the difference between the hypothesized CDF $F(x)$ and the empirical CDF $F_n(x)$. Let $y_1 < y_2 < \ldots < y_n$ be the observed values of the order statistics of a random sample $x_1, x_2, \ldots, x_n$ of size n. The A-D test statistic (A^2) is given by

$$A^2 = A_n^2 = \frac{-1}{n} \sum_{i=1}^{n} (2i - 1)\left\{\ln F_n(y_i) + \ln\left[1 - F_n(y_{n-i+1})\right]\right\} - n.$$

Fractiles of the distribution of A_n^2 for $\alpha = 0.15, 0.10, 0.05,$ and 0.01, denoted by a_α, are given in Table 11.2 on p. 400 of Blischke and Murthy (2000). The null and alternative hypotheses being tested are, respectively, given by:

H_0 : The data follow the specified continuous distribution
H_1: The data do not follow the specified continuous distribution.

The null hypothesis (H_0) is rejected if the A-D test statistic, A_n^2, is greater than the above tabulated constant a_α (also known as the critical value for A-D test analysis) at one of the chosen significance levels, $\alpha = 0.15, 0.10, 0.05,$ and 0.01. As pointed out in Blischke and Murthy (2000), "the critical value a_α does not depend on n, and has been found to be a very good approximation in samples as small as $n = 3$".

2.3.4 The Shapiro-Wilk Test for Normality

The **Shapiro-Wilk test** (also known as the W test) may be applied to test the goodness of fit between a hypothesized cumulative distribution function (CDF) $F(x)$ and an empirical CDF $F_n(x)$.

Let $y_1 < y_2 < \ldots < y_n$ be the observed values of the order statistics of a random sample $x_1, x_2, \ldots, x_n$ of size n with some unknown distribution function $F(x)$. Following Conover (1999), the Shapiro-Wilk test statistic, W, is defined as

$$W = \frac{\sum_{i=1}^{k} a_i (y_{n-i+1} - y_i)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2},$$

where $\overline{x}$ denotes the sample mean, and, for the observed sample size $n \leq 50$, the coefficients $a_i, i = 1, \ldots, k$, where k is approximately $\frac{n}{2}$, are available in Table A16, pp. 550–552, of Conover (1999). For the observed sample size $n > 50$, the interested readers are referred to D'Agostino (1971) and Shapiro and Francia (1972) (Fig. 2.12).

For the Shapiro-Wilk test, the null and alternative hypotheses being are, respectively, given by:

H_0: $F(x)$ is a normal distribution with unspecified mean and variance
H_1: $F(x)$ is non-normal.

The null hypothesis (H_0) is rejected at one of the chosen significance levels α if the Shapiro-Wilk test statistic, W, is less than the α quantile as given by Table A 17,

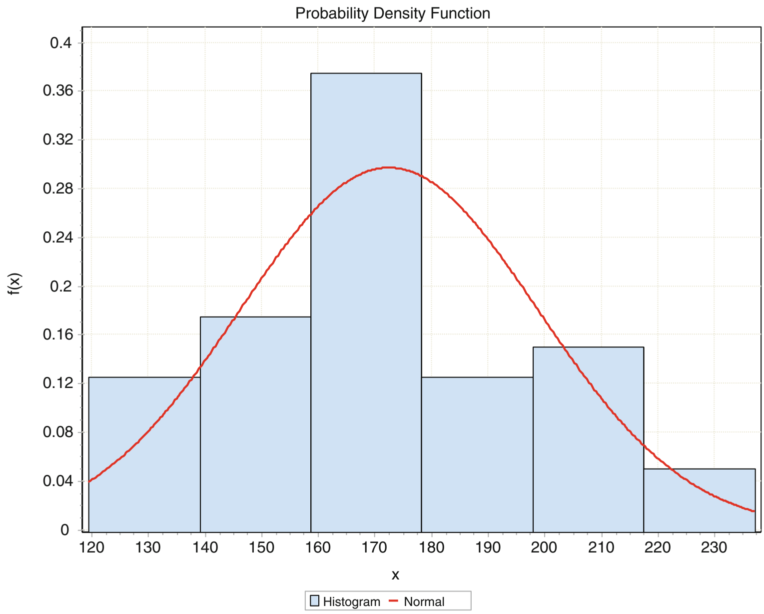

Fig. 2.12 Frequency histogram of the weights of 40 adult men

pp. 552–553, of Conover (1999). The p- value for the Shapiro-Wilk test may be calculated by following the procedure on p. 451 of Conover (1999).

Note: The Shapiro-Wilk test statistic may be calculated using the computer softwares such as R, Maple, Minitab, SAS, and StatXact, among others.

2.3.5 Applications

In order to examine the applications of the above tests of normality, we consider the following example of weights of a random sample of 40 adult men (Source: *Biostatistics for the Biological and Health Sciences*, Mario F Triola, Publisher: Pearson, 2005).

Example: We consider the weights of a random sample of 40 adult men as given below:

{169.1, 144.2, 179.3, 175.8, 152.6, 166.8, 135.0, 201.5, 175.2, 139.0, 156.3, 186.6, 191.1, 151.3, 209.4, 237.1, 176.7, 220.6, 166.1, 137.4, 164.2, 162.4, 151.8, 144.1, 204.6, 193.8, 172.9, 161.9, 174.8, 169.8, 213.3, 198.0, 173.3, 214.5, 137.1, 119.5, 189.1, 164.7, 170.1, 151.0}.

Table 2.2 Descriptive statistics

Statistic	**Value**	**Percentile**	**Value**
Sample size	40	Min	119.5
Range	117.6	5 %	135.11
Mean	172.55	10 %	137.56
Variance	693.12	25 % (Q1)	152.0
Standard deviation	26.327	50 % (Median)	169.95
Coefficient of variation	0.15258	75 % (Q3)	190.6
Standard error	4.1627	90 %	212.91
Skewness	0.37037	95 %	220.29
Excess Kurtosis	−0.16642	Max	237.1

Table 2.3 Normality for the Weights of 40 Adult Men

Test statistics	Value of the test statistics	P-value	Decision at 5 % level of significance
K-S test	0.112	0.652	Do not reject H_0
A-D test	0.306	0.552	Do not reject H_0
Chi-Squared test	2.712	0.844	Do not reject H_0
Shapiro-Wilk test	0.967	0.379	Do not reject H_0

Using the software EasyFit, the descriptive statistics are computed in the Table 2.2 below. The frequency histogram of the weights of 40 adult men is drawn in Fig. 2.12.

The goodness of fit (or GOF) tests, as discussed above, are applied to test the compatibility of our example of weights of the random sample of 40 adult men with our hypothesized theoretical probability distribution, that is, normal distribution, using various software such as EasyFit, Maple, and Minitab. The results are summarized in the Table 2.2 below. The chosen significance level is $\alpha = 0.05$. The null and alternative hypotheses being tested are, respectively, given by:

H_0: The data follow the normal distribution;
H_1: The data do not follow the normal distribution.

It is obvious from Table 2.3 is that the normal distribution seems to be an appropriate model for the weights of 40 adult men considered here. Since the sample size is large enough, all tests are valid for this example. In this section, we have discussed various tests of normality to test the suitability of a random sample with a theoretical probability distribution function. In particular, we have applied to test the applicability of normal distribution to a random sample of the weights of 40 adult men. It is hoped that this study may be helpful to apply these goodness of fit (or GOF) tests to other examples also.

2.4 Summary

The different forms of normal distributions and their various properties are discussed in this chapter. The entropy of a random variable having the normal distribution has been given. The expressions for the characteristic function of a normal distribution are provided. Some goodness of fit tests for testing the normality along with applications is given. By using Maple 10, various graphs have been plotted. As a motivation, different forms of normal distributions (folded and half normal etc.) and their properties have also been provided.

http://www.springer.com/978-94-6239-060-7

Normal and Student´s t Distributions and Their Applications
Ahsanullah, M.; Kibria, B.M.G.; Shakil, M.
2014, XII, 157 p. 36 illus., 27 illus. in color., Hardcover
ISBN: 978-94-6239-060-7
A product of Atlantis Press