

Chapter 1

Quantile Functions

Abstract A probability distribution can be specified either in terms of the distribution function or by the quantile function. This chapter addresses the problem of describing the various characteristics of a distribution through its quantile function. We give a brief summary of the important milestones in the development of this area of research. The definition and properties of the quantile function with examples are presented. In Table 1.1, quantile functions of various life distributions, representing different data situations, are included. Descriptive measures of the distributions such as location, dispersion and skewness are traditionally expressed in terms of the moments. The limitations of such measures are pointed out and some alternative quantile-based measures are discussed. Order statistics play an important role in statistical analysis. Distributions of order statistics in quantile forms, their properties and role in reliability analysis form the next topic in the chapter. There are many problems associated with the use of conventional moments in modelling and analysis. Exploring these, and as an alternative, the definition, properties and application of L -moments in describing a distribution are presented. Finally, the role of certain graphical representations like the Q - Q plot, box-plot and leaf-plot are shown to be useful tools for a preliminary analysis of the data.

1.1 Introduction

As mentioned earlier, a probability distribution can be specified either in terms of the distribution function or by the quantile function. Although both convey the same information about the distribution, with different interpretations, the concepts and methodologies based on distribution functions are traditionally employed in most forms of statistical theory and practice. One reason for this is that quantile-based studies were carried out mostly when the traditional approach either is difficult or fails to provide desired results. Except in a few isolated areas, there have been no systematic parallel developments aimed at replacing distribution functions in modelling and analysis by quantile functions. However, the feeling that through

an appropriate choice of the domain of observations, a better understanding of a chance phenomenon can be achieved by the use of quantile functions, is fast gaining acceptance.

Historically, many facts about the potential of quantiles in data analysis were known even before the nineteenth century. It appears that the Belgian sociologist Quetelet [499] initiated the use of quantiles in statistical analysis in the form of the present day inter-quantile range. A formal representation of a distribution through a quantile function was introduced by Galton (1822–1911) [206] who also initiated the ordering of observations along with the concepts of median, quartiles and interquartile range. Subsequently, the research on quantiles was directed towards estimation problems with the aid of sample quantiles, their large sample behaviour and limiting distributions (Galton [207, 208]). A major development in portraying quantile functions to represent distributions is the work of Hastings et al. [264], who introduced a family of distributions by a quantile function. This was refined later by Tukey [568]. The symmetric distribution of Tukey [568] and his articulation of exploratory data analysis sparked considerable interest in quantile functional forms that continues till date. Various aspects of the Tukey family and generalizations thereto were studied by a number of authors including Hogben [273], Shapiro and Wilk [536], Filliben [197], Joiner and Rosenblatt [304], Ramberg and Schmeiser [504], Ramberg [501], Ramberg et al. [502], MacGillivray [407], Freimer et al. [203], Gilchrist [215] and Tarsitano [563]. We will discuss all these models in Chap. 3. Another turning point in the development of quantile functions is the seminal paper by Parzen [484], in which he emphasized the description of a distribution in terms of the quantile function and its role in data modelling. Parzen [485–487] exhibits a sequential development of the theory and application of quantile functions in different areas and also as a tool in unification of various approaches.

Quantile functions have several interesting properties that are not shared by distributions, which makes it more convenient for analysis. For example, the sum of two quantile functions is again a quantile function. There are explicit general distribution forms for the quantile function of order statistics. In Sect. 1.2, we mention these and some other properties. Moreover, random numbers from any distribution can be generated using appropriate quantile functions, a purpose for which lambda distributions were originally conceived. The moments in different forms such as raw, central, absolute and factorial have been used effectively in specifying the model, describing the basic characteristics of distributions, and in inferential procedures. Some of the methods of estimation like least squares, maximum likelihood and method of moments often provide estimators and/or their standard errors in terms of moments. Outliers have a significant effect on the estimates so derived. For example, in the case of samples from the normal distribution, all the above methods give sample mean as the estimate of the population mean, whose values change significantly in the presence of an outlying observation. Asymptotic efficiency of the sample moments is rather poor for heavy tailed distributions since the asymptotic variances are mainly in terms of higher order moments that tend to be large in this case. In reliability analysis, a single long-term survivor can have a marked effect

on mean life, especially in the case of heavy tailed models which are commonly encountered for lifetime data. In such cases, quantile-based estimates are generally found to be more precise and robust against outliers. Another advantage in choosing quantiles is that in life testing experiments, one need not wait until the failure of all the items on test, but just a portion of them for proposing useful estimates. Thus, there is a case for adopting quantile functions as models of lifetime and base their analysis with the aid of functions derived from them. Many other facets of the quantile approach will be more explicit in the sequel in the form of alternative methodology, new opportunities and unique cases where there are no corresponding results if one adopts the distribution function approach.

1.2 Definitions and Properties

In this section, we define the quantile function and discuss some of its general properties. The random variable considered here has the real line as its support, but the results are valid for lifetime random variables which take on only non-negative values.

Definition 1.1. Let X be a real valued continuous random variable with distribution function $F(x)$ which is continuous from the right. Then, the quantile function $Q(u)$ of X is defined as

$$Q(u) = F^{-1}(u) = \inf\{x : F(x) \geq u\}, \quad 0 \leq u \leq 1. \quad (1.1)$$

For every $-\infty < x < \infty$ and $0 < u < 1$, we have

$$F(x) \geq u \text{ if and only if } Q(u) \leq x.$$

Thus, if there exists an x such that $F(x) = u$, then $F(Q(u)) = u$ and $Q(u)$ is the smallest value of x satisfying $F(x) = u$. Further, if $F(x)$ is continuous and strictly increasing, $Q(u)$ is the unique value x such that $F(x) = u$, and so by solving the equation $F(x) = u$, we can find x in terms of u which is the quantile function of X . Most of the distributions we consider in this work are of this form and nature.

Definition 1.2. If $f(x)$ is the probability density function of X , then $f(Q(u))$ is called the density quantile function. The derivative of $Q(u)$, i.e.,

$$q(u) = Q'(u),$$

is known as the quantile density function of X . By differentiating $F(Q(u)) = u$, we find

$$q(u)f(Q(u)) = 1. \quad (1.2)$$

Some important properties of quantile functions required in the sequel are as follows.

1. From the definition of $Q(u)$ for a general distribution function, we see that
 - (a) $Q(u)$ is non-decreasing on $(0, 1)$ with $Q(F(x)) \leq x$ for all $-\infty < x < \infty$ for which $0 < F(x) < 1$;
 - (b) $F(Q(u)) \geq u$ for any $0 < u < 1$;
 - (c) $Q(u)$ is continuous from the left or $Q(u-) = Q(u)$;
 - (d) $Q(u+) = \inf\{x : F(x) > u\}$ so that $Q(u)$ has limits from above;
 - (e) Any jumps of $F(x)$ are flat points of $Q(u)$ and flat points of $F(x)$ are jumps of $Q(u)$.
2. If U is a uniform random variable over $[0, 1]$, then $X = Q(U)$ has its distribution function as $F(x)$. This follows from the fact that

$$P(Q(U) \leq x) = P(U \leq F(x)) = F(x).$$

This property enables us to conceive a given data set as arising from the uniform distribution transformed by the quantile function $Q(u)$.

3. If $T(x)$ is a non-decreasing function of x , then $T(Q(u))$ is a quantile function. Gilchrist [215] refers to this as the Q -transformation rule. On the other hand, if $T(x)$ is non-increasing, then $T(Q(1-u))$ is also a quantile function.

Example 1.1. Let X be a random variable with Pareto type II (also called Lomax) distribution with

$$F(x) = 1 - \alpha^c(x + \alpha)^{-c}, \quad x > 0; \alpha, c > 0.$$

Since $F(x)$ is strictly increasing, setting $F(x) = u$ and solving for x , we obtain

$$x = Q(u) = \alpha[(1-u)^{-\frac{1}{c}} - 1].$$

Taking $T(X) = X^\beta$, $\beta > 0$, we have a non-decreasing transformation which results in

$$T(Q(u)) = \alpha^\beta[(1-u)^{-\frac{1}{c}} - 1]^\beta.$$

When $T(Q(u)) = y$, we obtain, on solving for u ,

$$u = G(y) = 1 - \left(1 + \frac{y^{\frac{1}{\beta}}}{\alpha}\right)^{-c}$$

which is a Burr type XII distribution with $T(Q(u))$ being the corresponding quantile function.

Example 1.2. Assume X has Pareto type I distribution with

$$F(x) = 1 - \left(\frac{x}{\sigma}\right)^{\alpha}, \quad x > \sigma; \alpha > 0, \sigma > 0.$$

Then, working as in the previous example, we see that

$$Q(u) = \sigma(1 - u)^{-\frac{1}{\alpha}}.$$

Apply the transformation $T(X) = Y = X^{-1}$, which is non-increasing, we have

$$T(Q(1 - u)) = \sigma^{-1}u^{\frac{1}{\alpha}}$$

and equating this to y and solving, we get

$$G(y) = (y\sigma)^{\alpha}, \quad 0 \leq y \leq \frac{1}{\sigma}.$$

$G(y)$ is the distribution function of a power distribution with $T(Q(1 - u))$ being the corresponding quantile function.

4. If $Q(u)$ is the quantile function of X with continuous distribution function $F(x)$ and $T(u)$ is a non-decreasing function satisfying the boundary conditions $T(0) = 0$ and $T(1) = 1$, then $Q(T(u))$ is a quantile function of a random variable with the same support as X .

Example 1.3. Consider a non-negative random variable with continuous distribution function $F(x)$ and quantile function $Q(u)$. Taking $T(u) = u^{\frac{1}{\theta}}$, for $\theta > 0$, we have $T(0) = 0$ and $T(1) = 1$. Then,

$$Q_1(u) = Q(T(u)) = Q(u^{\frac{1}{\theta}}).$$

Further, if $y = Q_1(u)$, $u^{\frac{1}{\theta}} = y$ and so the distribution function corresponding to $Q_1(u)$ is

$$G(x) = F^{\theta}(x).$$

The random variable Y with distribution function $G(x)$ is called the proportional reversed hazards model of X . There is considerable literature on such models in reliability and survival analysis. If we take X to be exponential with

$$F(x) = 1 - e^{-\lambda x}, \quad x > 0; \lambda > 0,$$

so that

$$Q(u) = \lambda^{-1}(-\log(1 - u)),$$

then

$$Q_1(u) = \lambda^{-1}(-\log(1 - u^{\frac{1}{\theta}}))$$

provides

$$G(x) = (1 - e^{-\lambda x})^\theta,$$

the generalized or exponentiated exponential law (Gupta and Kundu [250]). In a similar manner, Mudholkar and Srivastava [429] take the baseline distribution as Weibull. For some recent results and survey of such models, we refer the readers to Gupta and Gupta [240]. In Chap. 3, we will come across several quantile functions that represent families of distributions containing some life distributions as special cases. They are highly flexible empirical models capable of approximating many continuous distributions. The above transformation on these models generates new proportional reversed hazards models of a general form. The analysis of lifetime data employing such models seems to be an open issue.

Remark 1.1. From the form of $G(x)$ above, it is clear that for positive integral values of θ , it is simply the distribution function of the maximum of a random sample of size θ from the exponential population with distribution function $F(x)$ above. Thus, $G(x)$ may be simply regarded as the distribution function of the maximum from a random sample of real size θ (instead of an integer). This viewpoint was discussed by Stigler [547] under the general idea of 'fractional order statistics'; see also Rohatgi and Saleh [509].

Remark 1.2. Just as $G(x)$ can be regarded as the distribution function of the maximum from a random sample of (real) size θ from the population with distribution function $F(x)$, we can consider $G^*(x) = 1 - (1 - F(x))^\theta$ as a generalized form corresponding to the minimum of a random sample of (real) size θ . The model $G^*(x)$ is, of course, the familiar proportional hazards model. It is important to mention here that these two models are precisely the ones introduced by Lehmann [382], as early as in 1953, as stochastically ordered alternatives for nonparametric tests of equality of distributions.

Remark 1.3. It is useful to bear in mind that for distributions closed under minima such as exponential and Weibull (i.e., the distributions for which the minima have the same form of the distribution but with different parameters), the distribution function $G(x)$ would provide a natural generalization while, for distributions closed under maxima such as power and inverse Weibull (i.e., the distributions for which the maxima have the same form of the distribution but with different parameters), the distribution function $G^*(x)$ would provide a natural generalization.

5. The sum of two quantile functions is again a quantile function. Likewise, two quantile density functions, when added, produce another quantile density function.

6. The product of two positive quantile functions is a quantile function. In this case, the condition of positivity cannot be relaxed, as in general, there may be negative quantile functions that affect the increasing nature of the product. Since we are dealing primarily with lifetimes, the required condition will be automatically satisfied.
7. If X has quantile function $Q(u)$, then $\frac{1}{X}$ has quantile function $1/Q(1-u)$.

Remark 1.4. Property 7 is illustrated in Example 1.2. Chapter 3 contains some examples wherein quantile functions are generated as sums and products of quantile functions of known distributions. It becomes evident from Properties 3–7 that they can be used to produce new distributions from the existing ones. Thus, in our approach, a few basic forms are sufficient to begin with since new forms can always be evolved from them that match our requirements and specifications. This is in direct contrast to the abundance of probability density functions built up, each to satisfy a particular data form in the distribution function approach. In data analysis, the crucial advantage is that if one quantile function is not an appropriate model, the features that produce lack of fit can be ascertained and rectification can be made to the original model itself. This avoids the question of choice of an altogether new model and the repetition of all inferential procedures for the new one as is done in most conventional analyses.

8. The concept of residual life is of special interest in reliability theory. It represents the lifetime remaining in a unit after it has attained age t . Thus, if X is the original lifetime with quantile function $Q(u)$, the associated residual life is the random variable $X_t = (X - t | X > t)$. Using the definition of conditional probability, the survival function of X_t is

$$\bar{F}_t(x) = P(X_t > x) = \frac{\bar{F}(x+t)}{\bar{F}(t)},$$

where $\bar{F}(x) = P(X > x) = 1 - F(x)$ is the survival function. Thus, we have

$$F_t(x) = \frac{F(x+t) - F(t)}{1 - F(t)}. \quad (1.3)$$

Let $F(t) = u_0$, $F(x+t) = v$ and $F_t(x) = u$. Then, with

$$x+t = Q(v), \quad x = Q_1(u), \text{ say,}$$

we have

$$Q_1(u) = Q(v) - Q(u_0)$$

and consequently from (1.3),

$$u(1 - u_0) = v - u_0$$

or

$$v = u_0 + (1 - u_0)u.$$

Thus, the quantile function of the residual life X_t becomes

$$Q_1(u) = Q(u_0 + (1 - u_0)u) - Q(u_0). \quad (1.4)$$

Equation (1.4) will be made use of later in defining mean residual quantile function in Chap. 2.

9. In some reliability and quality control situations, truncated forms of lifetime models arise naturally, and the truncation may be on the right or on the left or on both sides. Suppose $F(x)$ is the underlying distribution function and $Q(u)$ is the corresponding quantile function. Then, if the distribution is truncated on the right at $x = U$ (i.e., the observations beyond U cannot be observed), then the corresponding distribution function is

$$F_{RT}(x) = \frac{F(x)}{F(U)}, \quad 0 \leq x \leq U,$$

and its quantile function is

$$Q_{RT}(x) = Q(uQ^{-1}(U)).$$

Similarly, if the distribution is truncated on the left at $x = L$ (i.e., the observations below L cannot be observed), then the corresponding distribution function is

$$F_{LT}(x) = \frac{F(x) - F(L)}{1 - F(L)}, \quad x \geq L,$$

and its quantile function is

$$Q_{LT}(u) = Q(u + (1 - u)Q^{-1}(L)).$$

Finally, if the distribution is truncated on the left at $x = L$ and also on the right at $x = U$, then the corresponding distribution function is

$$F_{DT}(x) = \frac{F(x) - F(L)}{F(U) - F(L)}, \quad L \leq x \leq U,$$

and its quantile function is

$$Q_{DT}(u) = Q(uQ^{-1}(U) + (1 - u)Q^{-1}(L)).$$

Example 1.4. Suppose the underlying distribution is logistic with distribution function $F(x) = 1/(1 + e^{-x})$ on the whole real line \mathbb{R} . It is easily seen that the corresponding quantile function is $Q(u) = \log\left(\frac{u}{1-u}\right)$. Further, suppose we consider the distribution truncated on the left at 0, i.e., $L = 0$, for proposing a lifetime model. Then, from the expression above and the fact that $Q^{-1}(0) = \frac{1}{2}$, we arrive at the quantile function

$$Q_{LT}(u) = Q\left(u + (1-u)\frac{1}{2}\right) = \log\left(\frac{u + \frac{1}{2}(1-u)}{1 - u - \frac{1}{2}(1-u)}\right) = \log\left(\frac{1+u}{1-u}\right)$$

corresponding to the half-logistic distribution of Balakrishnan [47, 48]; see Table 1.1.

1.3 Quantile Functions of Life Distributions

As mentioned earlier, we concentrate here on distributions of non-negative random variables representing the lifetime of a component or unit. The distribution function of such random variables is such that $F(0-) = 0$. Often, it is more convenient to work with

$$\bar{F}(x) = 1 - F(x) = P(X > x),$$

which is the probability that the unit survives time (referred to as the age of the unit) x . It is also called the reliability or survival function since it expresses the probability that the unit is still reliable at age x .

In the previous section, some examples of quantile functions and a few methods of obtaining them were explained. We now present in Table 1.1 quantile functions of many distributions considered in the literature as lifetime models. The properties of these distributions are discussed in the references cited below each of them. Models like gamma, lognormal and inverse Gaussian do not find a place in the list as their quantile functions are not in a tractable form. However, in the next chapter, we will see quantile forms that provide good approximations to them.

1.4 Descriptive Quantile Measures

The advent of the Pearson family of distributions was a major turning point in data modelling using distribution functions. The fact that members of the family can be characterized by the first four moments gave an impetus to the extensive use of moments in describing the properties of distributions and their fitting to observed data. A familiar pattern of summary measures took the form of mean

Table 1.1 Quantile functions of some lifetime distributions

No.	Distribution	$\bar{F}(x)$	$Q(u)$
1	Exponential (Marshall and Olkin [412])	$\exp[-\lambda x]$ $x > 0; \lambda > 0$	$\lambda^{-1}(-\log(1-u))$
2	Weibull (Murthy et al. [434], Hahn and Shapiro [257])	$\exp[-(\frac{x}{\sigma})^\lambda]$ $x > 0; \lambda, \sigma > 0$	$\sigma(-\log(1-u))^{\frac{1}{\lambda}}$
3	Pareto II (Marshall and Olkin [412])	$\alpha^c(x+\alpha)^{-c}$ $x > 0; \alpha, c > 0$	$\alpha[(1-u))^{-\frac{1}{c}} - 1]$
4	Rescaled beta (Marshall and Olkin [412])	$(1 - \frac{x}{R})^c$ $0 \leq x \leq R; c, R > 0$	$R[1 - (1-u))^{\frac{1}{c}}]$
5	Half-logistic (Balakrishnan [47, 48], Balakrishnan and Wong [61])	$2 [1 + \exp(\frac{x}{\sigma})]^{-1}$ $x > 0; \sigma > 0$	$\sigma \log(\frac{1+u}{1-u})$
6	Power (Marshall and Olkin [412])	$1 - (\frac{x}{\alpha})^\beta$ $0 \leq x \leq \alpha; \alpha, \beta > 0$	$\alpha u^{\frac{1}{\beta}}$
7	Pareto I (Marshall and Olkin [412])	$(\frac{x}{\sigma})^{-\alpha}$ $x > \sigma > 0; \alpha, \sigma > 0$	$\sigma(1-u)^{-\frac{1}{\alpha}}$
8	Burr type XII (Zimmer et al. [604], Fry [204])	$(1+x^c)^{-k}$ $x > 0; c, k > 0$	$[(1-u)^{\frac{1}{k}} - 1]^{\frac{1}{c}}$
9	Gompertz (Lai and Xie [368])	$\exp[-\frac{B(C^x-1)}{\log C}]$ $x > 0; B, C > 0$	$\frac{1}{\log C} [1 - \frac{\log C \log(1-u)}{B}]$
10	Greenwich [225]	$(1 + \frac{x^2}{b^2})^{-\frac{a}{2}}$ $x \geq 0; a, b > 0$	$b[(1-u)^{\frac{2}{a}} - 1]^{\frac{1}{2}}$
11	Kus [364]	$\frac{1-e^{\lambda e^{-\beta x}}}{1-e^\lambda}$ $x > 0; \lambda, \beta > 0$	$-\frac{1}{\beta} \log[\lambda^{-1} \log\{1 - (1-u)(1 - e^{-\lambda})\}]$
12	Logistic exponential (Lan and Leemis [372])	$\frac{1+(e^{\lambda\theta}-1)^k}{1+(e^{\lambda(x+\theta)}-1)^k}$ $x \geq 0; \lambda > 0,$ $k > 0, \theta \geq 0$	$\frac{1}{\lambda} \log[1 + \{ \frac{(e^{\lambda\theta}-1)^k+u}{1-u} \}^{\frac{1}{k}}]$
13	Dimitrakopoulou et al. [178]	$\exp[1 - (1 + \lambda x^\beta)^\alpha]$ $x > 0; \alpha, \beta, \lambda > 0$	$\lambda^{-1} \{ [1 - \log(1-u)]^{\frac{1}{\alpha}} - 1 \}^{\frac{1}{\beta}}$
14	Log Weibull (Avinadav and Raz [41])	$\exp[-(\log(1 + \rho x))^k]$ $x > 0; \rho, k > 0$	$\rho^{-1} [\exp(-\log(1-u))^{\frac{1}{k}} - 1]$
15	Modified Weibull extension (Xie et al. [595])	$\exp[-\alpha \sigma (e^{\frac{x}{\sigma}})^\lambda - 1]$ $x > 0; \alpha, \sigma, \lambda > 0$	$\sigma [\log(1 + \frac{\log(1-u)}{\alpha \sigma})]^{\frac{1}{\lambda}}$
16	Exponential power (Paranjpe et al. [482])	$\exp[e^{-(\lambda x)^\alpha} - 1]$ $x > 0; \lambda, \alpha > 0$	$\frac{1}{\lambda} [-\log(1 + \log(1-u))]^{\frac{1}{\alpha}}$
17	Generalized Pareto (Lai and Xie [368])	$(1 + \frac{ax}{b})^{-\frac{a+1}{a}}$ $x > 0, b > 0, a > -1$	$\frac{b}{a} [(1-u)^{-\frac{a}{a+1}} - 1]$
18	Inverse Weibull (Erto [188])	$1 - \exp[-(\frac{\sigma}{x})^\lambda]$ $x > 0; \sigma, \lambda > 0$	$\sigma(-\log u)^{-\frac{1}{\lambda}}$

(continued)

Table 1.1 (continued)

No.	Distribution	$\bar{F}(x)$	$Q(u)$
19	Extended Weibull (Marshall and Olkin [411])	$\frac{\theta \exp[-(\frac{x}{\sigma})^\lambda]}{1-(1-\theta)\exp[-(\frac{x}{\sigma})^\lambda]}$ $x > 0; \theta, \lambda, \sigma > 0$	$\sigma [\log \frac{\theta+(1-\theta)(1-u)}{1-u}]^{\frac{1}{\lambda}}$
20	Generalized exponential (Gupta et al. [239])	$1 - [1 - \exp(-\frac{x}{\sigma})]^\theta$ $x > 0; \sigma, \theta > 0$	$\sigma [-\log(1 - u^{\frac{1}{\theta}})]$
21	Exponentiated Weibull (Mudholkar et al. [427])	$1 - [1 - \exp(-\frac{x}{\sigma})^\lambda]^\theta$ $x > 0; \sigma, \theta, \lambda > 0$	$\sigma [-\log(1 - u^{\frac{1}{\theta}})]^{\frac{1}{\lambda}}$
22	Generalized Weibull (Mudholkar and Kollia [426])	$[1 - \lambda (\frac{x}{\beta})^\alpha]^{\frac{1}{\lambda}}$ $x > 0$ for $\lambda \leq 0$ $0 < x < \frac{\beta}{\lambda^{\frac{1}{\alpha}}}, \lambda > 0$ $\alpha, \beta > 0$	$\beta [\frac{1-(1-u)^\lambda}{\lambda}]^{\frac{1}{\alpha}}, \lambda \neq 0$ $\beta [-\log(1 - u)]^{\frac{1}{\alpha}}, \lambda = 0$
23	Exponential geometric (Adamidis and Loukas [18])	$(1 - p)e^{-\lambda x}(1 - pe^{-\lambda x})^{-1}$ $x > 0, \lambda > 0, 0 < p < 1$	$\frac{1}{\lambda} \log(\frac{1-pu}{1-u})$
24	Log logistic (Gupta et al. [237])	$(1 + (\alpha x)^\beta)^{-1}$ $x > 0, \alpha, \beta > 0$	$\alpha^{-1} (\frac{u}{1-u})^{\frac{1}{\beta}}$
25	Generalized half-logistic (Balakrishnan and Sandhu [59], Balakrishnan and Aggarwala [49])	$\frac{2(1-kx)^{1/k}}{1+(1-kx)^{1/k}}$ $0 \leq x \leq \frac{1}{k}, k \geq 0$	$\frac{1}{k} \left\{ 1 - (\frac{1-u}{1+u})^k \right\}$

for location, variance for dispersion, and the Pearson’s coefficients $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ for skewness and $\beta_2 = \frac{\mu_4}{\mu_2^2}$ for kurtosis. While the mean and variance claimed universal acceptance, several limitations of β_1 and β_2 were subsequently exposed. Some of the concerns with regard to β_1 are: (1) it becomes arbitrarily large or even infinite making it difficult for comparison and interpretation as relatively small changes in parameters produce abrupt changes, (2) it does not reflect the sign of the difference (mean-median) which is a traditional basis for defining skewness, (3) there exist asymmetric distributions with $\beta_1 = 0$ and (4) instability of the sample estimate of β_1 while matching with the population value. Similarly, for a standardized variable X , the relationship

$$E(X^4) = 1 + V(X^2) \tag{1.5}$$

would mean that the interpretation of kurtosis depends on the concentration of the probabilities near $\mu \pm \sigma$ as well as in the tails of the distribution.

The specification of a distribution through its quantile function takes away the need to describe a distribution through its moments. Alternative measures in terms of quantiles that reduce the shortcomings of the moment-based ones can be thought of. A measure of location is the median defined by

$$M = Q(0.5). \quad (1.6)$$

Dispersion is measured by the interquartile range

$$\text{IQR} = Q_3 - Q_1, \quad (1.7)$$

where $Q_3 = Q(0.75)$ and $Q_1 = Q(0.25)$.

Skewness is measured by Galton's coefficient

$$S = \frac{Q_1 + Q_3 - 2M}{Q_3 - Q_1}. \quad (1.8)$$

Note that in the case of extreme positive skewness, $Q_1 \rightarrow M$ while in the case of extreme negative skewness $Q_3 \rightarrow M$ so that S lies between -1 and $+1$. When the distribution is symmetric, $M = \frac{Q_1 + Q_3}{2}$ and hence $S = 0$. Due to the relation in (1.5), kurtosis can be large when the probability mass is concentrated near the mean or in the tails. For this reason, Moors [421] proposed the measure

$$T = [Q(0.875) - Q(0.625) + Q(0.375) - Q(0.125)]/\text{IQR} \quad (1.9)$$

as a measure of kurtosis. As an index, T is justified on the grounds that the differences $Q(0.875) - Q(0.625)$ and $Q(0.375) - Q(0.125)$ become large (small) if relatively small (large) probability mass is concentrated around Q_3 and Q_1 corresponding to large (small) dispersion in the vicinity of $\mu \pm \sigma$.

Given the form of $Q(u)$, the calculations of all the coefficients are very simple, as we need to only substitute the appropriate fractions for u . On the other hand, calculation of moments given the distribution function involves integration, which occasionally may not even yield closed-form expressions.

Example 1.5. Let X follow the Weibull distribution with (see Table 1.1)

$$Q(u) = \sigma(-\log(1-u))^{\frac{1}{\lambda}}.$$

Then, we have

$$\begin{aligned} M &= Q\left(\frac{1}{2}\right) = \sigma(\log 2)^{\frac{1}{\lambda}}, \\ S &= \frac{(\log 4)^{\frac{1}{\lambda}} + (\log \frac{4}{3})^{\frac{1}{\lambda}} - 2(\log 2)^{\frac{1}{\lambda}}}{(\log 4)^{\frac{1}{\lambda}} - (\log \frac{4}{3})^{\frac{1}{\lambda}}}, \\ \text{IQR} &= \sigma \left[(\log 4)^{\frac{1}{\lambda}} - \left(\log \frac{4}{3} \right)^{\frac{1}{\lambda}} \right], \end{aligned}$$

and

$$T = \frac{(\log 8)^{\frac{1}{\lambda}} - (\log \frac{8}{3})^{\frac{1}{\lambda}} + (\log \frac{8}{5})^{\frac{1}{\lambda}} - (\log \frac{8}{7})^{\frac{1}{\lambda}}}{(\log 4)^{\frac{1}{\lambda}} - (\log \frac{4}{3})^{\frac{1}{\lambda}}}.$$

The effect of a change of origin and scale on $Q(u)$ and the above four measures are of interest in later studies. Let X and Y be two random variables such that $Y = aX + b$. Then,

$$F_Y(y) = P(Y \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-a}{b}\right).$$

If $Q_X(u)$ and $Q_Y(u)$ denote the quantile functions of X and Y , respectively,

$$F_X\left(\frac{y-a}{b}\right) = u \Rightarrow Q_X(u) = \frac{y-b}{a} = \frac{Q_Y(u) - b}{a}$$

or

$$Q_Y(u) = aQ_X(u) + b.$$

So, we simply have

$$M_Y = Q_Y(0.5) = aQ_X(0.5) + b = aM_X + b.$$

Similar calculations using (1.7), (1.8) and (1.9) yield

$$\text{IQR}_Y = a\text{IQR}_X, S_Y = S_X \text{ and } T_Y = T_X.$$

Other quantile-based measures have also been suggested for quantifying spread, skewness and kurtosis. One measure of spread, similar to mean deviation in the distribution function approach, is the median of absolute deviation from the median, viz.,

$$A = \text{Med}(|X - M|). \quad (1.10)$$

For further details and justifications for (1.10), we refer to Falk [194]. A second popular measure that has received wide attention in economics is Gini's mean difference defined as

$$\begin{aligned} \Delta &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x-y|f(x)f(y)dx dy \\ &= 2 \int_{-\infty}^{\infty} F(x)(1-F(x))dx, \end{aligned} \quad (1.11)$$

where $f(x)$ is the probability density function of X . Setting $F(x) = u$ in (1.11), we have

$$\Delta = 2 \int_0^1 u(1-u)q(u)du \quad (1.12)$$

$$= 2 \int_0^1 (2u-1)Q(u)du. \quad (1.13)$$

The expression in (1.13) follows from (1.12) by integration by parts. One may use (1.12) or (1.13) depending on whether $q(u)$ or $Q(u)$ is specified. Gini's mean difference will be further discussed in the context of reliability in Chap. 4.

Example 1.6. The generalized Pareto distribution with (see Table 1.1)

$$Q(u) = \frac{b}{a} \left\{ (1-u)^{-\frac{a}{a+1}} - 1 \right\}$$

has its quantile density function as

$$q(u) = \frac{b}{a+1} (1-u)^{-\frac{a}{a+1}-1}.$$

Then, from (1.12), we obtain

$$\Delta = \frac{2b}{a+1} \int_0^1 u(1-u)^{-\frac{a}{a+1}} du = \frac{2b}{a+1} B\left(2, \frac{1}{a+1}\right),$$

where $B(m, n) = \int_0^1 t^{m-1}(1-t)^{n-1} dt$ is the complete beta function. Thus, we obtain the simplified expression

$$\Delta = \frac{2b(a+1)}{a+2}.$$

Hinkley [271] proposed a generalization of Galton's measure of skewness of the form

$$S(u) = \frac{Q(u) + Q(1-u) - 2Q(0.5)}{Q(u) - Q(1-u)}. \quad (1.14)$$

Obviously, (1.14) reduces to Galton's measure when $u = 0.75$. Since (1.14) is a function of u and u is arbitrary, an overall measure of skewness can be provided as

$$S_2 = \sup_{\frac{1}{2} \leq u \leq 1} S(u).$$

Groeneveld and Meeden [227] suggested that the numerator and denominator in (1.14) be integrated with respect to u to arrive at the measure

$$S_3 = \frac{\int_{\frac{1}{2}}^1 \{Q(u) + Q(1-u) - 2Q(0.5)\} du}{\int_{\frac{1}{2}}^1 \{Q(u) + Q(1-u)\} du}.$$

Now, in terms of expectations, we have

$$\begin{aligned} \int_{\frac{1}{2}}^1 Q(u) du &= \int_M^x xf(x) dx, \\ \int_{\frac{1}{2}}^1 Q(1-u) du &= \int_0^{\frac{1}{2}} Q(u) du = \int_0^M xf(x) dx, \\ \int_{\frac{1}{2}}^1 Q(0.5) du &= \frac{1}{2}M, \end{aligned}$$

and thus

$$S_3 = \frac{E(X) - M}{\int_M^\infty xf(x) dx - \int_0^M xf(x) dx} = \frac{\mu - M}{E(|X - M|)}. \quad (1.15)$$

The numerator of (1.15) is the traditional term (being the difference between the mean and the median) indicating skewness and the denominator is a measure of spread used for standardizing S_3 . Hence, (1.15) can be thought of as an index of skewness in the usual sense. If we replace the denominator by the standard deviation σ of X , the classical measure of skewness will result.

Example 1.7. Consider the half-logistic distribution with (see Table 1.1)

$$\begin{aligned} Q(u) &= \sigma \log \left(\frac{1+u}{1-u} \right), \\ \mu &= \int_0^1 Q(u) du = \sigma \log 4, \\ \int_{\frac{1}{2}}^1 Q(u) du &= \sigma \left(\log 16 - \frac{3}{2} \log 3 \right), \\ \int_0^{\frac{1}{2}} Q(u) du &= \sigma \left(\frac{3}{2} \log 3 - 2 \log 2 \right), \end{aligned}$$

and hence $S_3 = \log(\frac{4}{3}) / \log(\frac{64}{27})$.

Instead of using quantiles, one can also use percentiles to define skewness. Galton [206] in fact used the middle 50 % of observations, while Kelly's measure takes 90 % of observations to propose the measure

$$S_4 = \frac{Q(0.90) + Q(0.10) - 2M}{Q(0.90) - Q(0.10)}.$$

For further discussion of alternative measures of skewness and kurtosis, a review of the literature and comparative studies, we refer to Balanda and MacGillivray [63], Tajuddin [559], Joannes and Gill [299], Suleswki [552] and Kotz and Seier [355].

1.5 Order Statistics

In life testing experiments, a number of units, say n , are placed on test and the quantity of interest is their failure times which are assumed to follow a distribution $F(x)$. The failure times X_1, X_2, \dots, X_n of the n units constitute a random sample of size n from the population with distribution function $F(x)$, if X_1, X_2, \dots, X_n are independent and identically distributed as $F(x)$. Suppose the realization of X_i in an experiment is denoted by x_i . Then, the order statistics of the random sample (X_1, X_2, \dots, X_n) are the sample values placed in ascending order of magnitude denoted by $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$, so that $X_{1:n} = \min_{1 \leq i \leq n} X_i$ and $X_{n:n} = \max_{1 \leq i \leq n} X_i$. The sample median, denoted by m , is the value for which approximately 50% of the observations are less than m and 50% are more than m . Thus

$$m = \begin{cases} X_{\frac{n+1}{2}:n} & \text{if } n \text{ is odd} \\ \frac{1}{2}(X_{\frac{n}{2}:n} + X_{\frac{n}{2}+1:n}) & \text{if } n \text{ is even.} \end{cases} \quad (1.16)$$

Generalizing, we have the percentiles. The $100p$ -th percentile, denoted by x_p , in the sample corresponds to the value for which approximately np observations are smaller than this value and $n(1-p)$ observations are larger. In terms of order statistics we have

$$x_p = \begin{cases} X_{[np]:n} & \text{if } \frac{1}{2n} < p < 0.5 \\ X_{(n+1)-[n(1-p)]} & \text{if } 0.5 < p < 1 - \frac{1}{2n} \end{cases}, \quad (1.17)$$

where the symbol $[t]$ is defined as $[t] = r$ whenever $r - 0.5 \leq t < r + 0.5$, for all positive integers r . We note that in the above definition, if x_p is the i th smallest observation, then the i th largest observation is x_{1-p} . Obviously, the median m is the 50th percentile and the lower quartile q_1 and the upper quartile q_3 of the sample are, respectively, the 25th and 75th percentiles. The sample interquartile range is

$$iqr = q_3 - q_1. \quad (1.18)$$

All the sample descriptive measures are defined in terms of the sample median, quartiles and percentiles analogous to the population measures introduced in Sect. 1.4. Thus, iqr in (1.18) describes the spread, while

$$s = \frac{q_3 + q_1 - 2m}{q_3 - q_1} \quad (1.19)$$

and

$$t = \frac{e_7 - e_5 + e_3 - e_1}{iqr}, \quad (1.20)$$

where $e_i = \frac{i}{8}$, $i = 1, 3, 5, 7$, describes the skewness and kurtosis.

Parzen [484] introduced the empirical quantile function

$$\bar{Q}(u) = F_n^{-1}(u) = \inf(x : F_n(x) \geq u),$$

where $F_n(x)$ is the proportion of X_1, X_2, \dots, X_n that is at most x . In other words,

$$\bar{Q}(u) = X_{r:n} \quad \text{for } \frac{r-1}{n} < u < \frac{r}{n}, \quad r = 1, 2, \dots, n, \quad (1.21)$$

which is a step function with jump $\frac{1}{n}$. For $u = 0$, $\bar{Q}(u)$ is taken as $X_{1:n}$ or a natural minimum if one is available. In the case of lifetime variables, this becomes $\bar{Q}(0)$. When a smooth function is required for $\bar{Q}(u)$, Parzen [484] suggested the use of

$$\bar{Q}_1(u) = n \left(\frac{r}{n} - u \right) X_{r-1:n} + n \left(u - \frac{r-1}{n} \right) X_{r:n}$$

for $\frac{r-1}{n} \leq u \leq \frac{r}{n}$, $r = 1, 2, \dots, n$. The corresponding empirical quantile density function is

$$\bar{q}_1(u) = \frac{d}{du} \bar{Q}_1(u) = n(X_{r:n} - X_{r-1:n}), \quad \text{for } \frac{r-1}{n} < u < \frac{r}{n}.$$

In this set-up, we have $q_i = \bar{Q}(\frac{i}{4})$, $i = 1, 3$ and $e_i = \bar{Q}(\frac{i}{8})$, $i = 1, 3, 7, 8$.

It is well known that the distribution of the r th order statistic $X_{r:n}$ is given by Arnold et al. [37]

$$F_r(x) = P(X_{r:n} \leq x) = \sum_{k=r}^n \binom{n}{k} F^k(x) (1 - F(x))^{n-k}. \quad (1.22)$$

In particular, $X_{n:n}$ and $X_{1:n}$ have their distributions as

$$F_n(x) = F^n(x) \quad (1.23)$$

and

$$F_1(x) = 1 - (1 - F(x))^n. \quad (1.24)$$

Recalling the definitions of the beta function

$$B(m, n) = \int_0^1 t^{m-1} (1-t)^{n-1} dt, \quad m, n > 0,$$

and the incomplete beta function ratio

$$I_x(m, n) = \frac{B_x(m, n)}{B(m, n)},$$

where

$$B_x(m, n) = \int_0^x t^{m-1} (1-t)^{n-1} dt,$$

we have the upper tail of the binomial distribution and the incomplete beta function ratio to be related as (Abramowitz and Stegun [15])

$$\sum_{k=r}^n \binom{n}{k} p^k (1-p)^{n-k} = I_p(r, n-r+1). \quad (1.25)$$

Comparing (1.22) and (1.25) we see that, if a sample of n observations from a distribution with quantile function $Q(u)$ is ordered, then the quantile function of the r th order statistic is given by

$$Q_r(u_r) = Q(I_{u_r}^{-1}(r, n-r+1)), \quad (1.26)$$

where

$$u_r = I_u(r, n-r+1) \quad (1.27)$$

and I^{-1} is the inverse of the beta function ratio I . Thus, the quantile function of the r th order statistic has an explicit distributional form, unlike the expression for distribution function in (1.22). However, the expression for $Q_r(u_r)$ is not explicit in terms of $Q(u)$. This is not a serious handicap as the $I_u(\cdot, \cdot)$ function is tabulated for various values of n and r (Pearson [489]) and also available in all statistical softwares for easy computation. The distributions of $X_{n:n}$ and $X_{1:n}$ have simple quantile forms

$$Q_n(u_n) = Q\left(u_n^{\frac{1}{n}}\right)$$

and

$$Q_1(u_1) = Q[1 - (1 - u_1)^{\frac{1}{n}}].$$

The probability density function of $X_{r:n}$ becomes

$$f_r(x) = \frac{n!}{(r-1)!(n-r)!} F^{r-1}(x)(1-F(x))^{n-r} f(x)$$

and so

$$\begin{aligned} \mu_{r:n} = E(X_{r:n}) &= \int x f_r(x) dx \\ &= \frac{n!}{(r-1)!(n-1)!} \int_0^1 u^{r-1} (1-u)^{n-r} Q(u) du. \end{aligned} \quad (1.28)$$

This mean value is referred to as the r th mean rankit of X . For reasons explained earlier with reference to the use of moments, often the median rankit

$$M_{r:n} = Q(I_{0.5}^{-1}(r, n-r+1)), \quad (1.29)$$

which is robust, is preferred over the mean rankit.

The importance and role of order statistics in the study of quantile function become clear from the discussions in this section. Added to this, there are several topics in reliability analysis in which order statistics appear quite naturally. One of them is system reliability. We consider a system consisting of n components whose lifetimes X_1, X_2, \dots, X_n are independent and identically distributed. The system is said to have a series structure if it functions only when all the components are functioning, and the lifetime of this system is the smallest among the X_i 's or $X_{1:n}$. In the parallel structure, on the other hand, the system functions if and only if at least one of the components work, so that the system life is $X_{n:n}$. These two structures are embedded in what is called a k -out-of- n system, which functions if and only if at least k of the components function. The lifetime of such a system is obviously $X_{n-k+1:n}$.

In life testing experiments, when n identical units are put on test to ascertain their lengths of life, there are schemes of sampling wherein the experimenter need not have to wait until all units fail. The experimenter may choose to observe only a prefixed number of failures of, say, $n-r$ units and terminate the experiment as soon as the $(n-r)$ th unit fails. Thus, the lifetimes of r units that are still working get censored. This sampling scheme is known as type II censoring. The data consists of realizations of $X_{1:n}, X_{2:n}, \dots, X_{n-r:n}$. Another sampling scheme is to prefix a time T^* and observe only those failures that occur up to time T^* . This scheme is known as type I censoring, and in this case the number of failures to be observed is random. One may refer to Balakrishnan and Cohen [51] and Cohen [154] for various methods of inference for type I and type II censored samples from a wide array of lifetime distributions. Yet another sampling scheme is to prefix the number of failures at $n-r$ and also a time T^* . If $(n-r)$ failures occur before time T^* , then the experiment is terminated; otherwise, observe all failures until time T^* . Thus, the time of truncation of the experiment is now $\min(T, X_{n-r:n})$. This is referred to as type I hybrid censoring; see Balakrishnan and Kundu [53] for an overview of various developments on this and many other forms of hybrid censoring

schemes. A third important application of order statistics is in the construction of tests regarding the nature of ageing of a device; see Lai and Xie [368]. For an encyclopedic treatment on the theory, methods and applications of order statistics, one may refer to Balakrishnan and Rao [56, 57].

1.6 Moments

The emphasis given to quantiles in describing the basic properties of a distribution does not in any way minimize the importance of moments in model specification and inferential problems. In this section, we look at various types of moments through quantile functions. The conventional moments

$$\mu'_r = E(X^r) = \int_0^\infty x^r f(x) dx$$

are readily expressible in terms of quantile functions, by the substitution $x = Q(u)$, as

$$\mu'_r = \int_0^1 \{Q(u)\}^r du. \quad (1.30)$$

In particular, as already seen, the mean is

$$\mu = \int_0^1 Q(u) du = \int_0^1 (1-u)q(u) du. \quad (1.31)$$

The central moments and other quantities based on it are obtained through the well-known relationships they have with the raw moments μ'_r in (1.30).

Some of the difficulties experienced while employing the moments in descriptive measures as well as in inferential problems have been mentioned in the previous sections. The L -moments to be considered next can provide a competing alternative to the conventional moments. Firstly, by definition, they are expected values of linear functions of order statistics. They have generally lower sampling variances and are also robust against outliers. Like the conventional moments, L -moments can be used as summary measures (statistics) of probability distributions (samples), to identify distributions and to fit models to data. The origin of L -moments can be traced back to the work on linear combination of order statistics in Sillito [537] and Greenwood et al. [226]. It was Hosking [276] who presented a unified theory on L -moments and made a systematic study of their properties and role in statistical analysis. See also Hosking [277, 279, 280] and Hosking and Wallis [282] for more elaborate details on this topic.

The r th L -moment is defined as

$$L_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r}), \quad r = 1, 2, 3, \dots \quad (1.32)$$

Using (1.28), we can write

$$L_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \frac{r!}{(r-k-1)!k!} \int_0^1 u^{r-k-1} (1-u)^k Q(u) du.$$

Expanding $(1-u)^k$ in powers of u using binomial theorem and combining powers of u , we get

$$L_r = \int_0^1 \sum_{k=0}^{r-1} (-1)^{r-1-k} \binom{r-1}{k} \binom{r-1+k}{k} u^k Q(u) du. \quad (1.33)$$

Jones [306] has given an alternative method of establishing the last relationship. In particular, we obtain:

$$L_1 = \int_0^1 Q(u) du = \mu, \quad (1.34)$$

$$L_2 = \int_0^1 (2u-1)Q(u) du, \quad (1.35)$$

$$L_3 = \int_0^1 (6u^2-6u+1)Q(u) du, \quad (1.36)$$

$$L_4 = \int_0^1 (20u^3-30u^2+12u-1)Q(u) du. \quad (1.37)$$

Sometimes, it is convenient (to avoid integration by parts while computing the integrals in (1.34)–(1.37)) to work with the equivalent formulas

$$L_1 = \int_0^1 (1-u)q(u) du, \quad (1.38)$$

$$L_2 = \int_0^1 (u-u^2)q(u) du, \quad (1.39)$$

$$L_3 = \int_0^1 (3u^2-2u^3-u)q(u) du, \quad (1.40)$$

$$L_4 = \int_0^1 (u-6u^2+10u^3-5u^4)q(u) du. \quad (1.41)$$

Example 1.8. For the exponential distribution with parameter λ , we have

$$Q(u) = -\lambda^{-1} \log(1-u) \quad \text{and} \quad q(u) = \frac{1}{\lambda(1-u)}.$$

Hence, using (1.38)–(1.41), we obtain

$$\begin{aligned} L_1 &= \int_0^1 \frac{1}{\lambda} du = \lambda^{-1}, \\ L_2 &= \int_0^1 u(1-u)q(u)du = \int_0^1 \frac{u}{\lambda} du = (2\lambda)^{-1}, \\ L_3 &= \int_0^1 u(1-u)(2u-1)q(u)du = (6\lambda)^{-1}, \\ L_4 &= \int_0^1 u(1-u)(1-5u+5u^2)q(u)du = (12\lambda)^{-1}. \end{aligned}$$

More examples are presented in Chap. 3 when properties of various distributions are studied.

The L -moments have the following properties that distinguish themselves from the usual moments:

1. The L -moments exist whenever $E(X)$ is finite, while additional restrictions may be required for the conventional moments to be finite for many distributions;
2. A distribution whose mean exists is characterized by $(L_r : r = 1, 2, \dots)$. This result can be compared with the moment problem discussed in probability theory. However, any set that contains all L -moments except one is not sufficient to characterize a distribution. For details, see Hosking [279, 280];
3. From (1.12), we see that $L_2 = \frac{1}{2}\Delta$, and so L_2 is a measure of spread. Thus, the first (being the mean) and second L -moments provide measures of location and spread. In a recent comparative study of the relative merits of the variance and the mean difference Δ , Yitzhaki [596] noted that the mean difference is more informative than the variance in deriving properties of distributions that depart from normality. He also compared the algebraic structure of variance and Δ and examined the relative superiority of the latter from the stochastic dominance, exchangeability and stratification viewpoints. For further comments on these aspects and some others in the reliability context, see Chap. 7;
4. Forming the ratios $\tau_r = \frac{L_r}{L_2}$, $r = 3, 4, \dots$, for any non-degenerate X with $\mu < \infty$, the result $|\tau_r| < 1$ holds. Hence, the quantities τ_r 's are dimensionless and bounded;
5. The skewness and kurtosis of a distribution can be ascertained through the moment ratios. The L -coefficient of skewness is

$$\tau_3 = \frac{L_3}{L_2} \tag{1.42}$$

and the L -coefficient of kurtosis is

$$\tau_4 = \frac{L_4}{L_2}. \tag{1.43}$$

These two measures satisfy the criteria presented for coefficients of skewness and kurtosis in terms of order relations. The range of τ_3 is $(-1, 1)$ while that of τ_4 is $\frac{1}{4}(5\tau_3^2 - 1) \leq \tau_4 < 1$. These results are proved in Hosking [279] and

Jones [306] using different approaches. It may be observed that both τ_3 and τ_4 are bounded and do not assume arbitrarily large values as β_1 (for example, in the case of $F(x) = 1 - x^{-3}$, $x > 1$);

6. The ratio

$$\tau_2 = \frac{L_2}{L_1} \quad (1.44)$$

is called L -coefficient of variation. Since X is non-negative in our case, $L_1 > 0$, $L_2 > 0$ and further

$$L_2 = \int_0^1 u(1-u)q(u)du < \int_0^1 (1-u)q(u)du = L_1$$

so that $0 < \tau_2 < 1$.

The above properties of L -moments have made them popular in diverse applications, especially in hydrology, civil engineering and meteorology. Several empirical studies (as the one by Sankarasubramonian and Sreenivasan [517]) comparing L -moments and the usual moments reveal that estimates based on the former are less sensitive to outliers. Just as matching the population and sample moments for the estimation of parameters, the same method (method of L -moments) can be applied with L -moments as well. Asymptotic approximations to sampling distributions are better achieved with L -moments. An added advantage is that standard errors of sample L -moments exist whenever the underlying distribution has a finite variance, whereas for the usual moments this may not be enough in many cases.

When dealing with the conventional moments, the (β_1, β_2) plot is used as a preliminary tool to discriminate between candidate distributions for the data. For example, if one wishes to choose a distribution from the Pearson family as a model, (β_1, β_2) provide exclusive classification of the members of this family. Distributions with no shape parameters are represented by points in the β_1 - β_2 plane, those with a single shape parameter have their (β_1, β_2) values lie on the line $2\beta_2 - 3\beta_1 - 6 = 0$, while two shape parameters in the distribution ensure that for them, (β_1, β_2) falls in a region between the lines $2\beta_2 - 3\beta_1 - 6 = 0$ and $\beta_2 - \beta_1 - 1 = 0$. These cases are, respectively, illustrated by the exponential distribution (which has $(\beta_1, \beta_2) = (4, 9)$ as a point), the gamma family and the beta family; see Johnson et al. [302] for details. In a similar manner, one can construct (τ_2, τ_3) -plots or (τ_3, τ_4) -plots for distribution functions or quantile functions to give a visual identification of which distribution can be expected to fit a given set of observations. Vogel and Fennessey [574] articulate the need for such diagrams and provide several examples on how to construct them. Some refinements of the L -moments are also studied in the name of trimmed L -moments (Elamir and Seheult [187], Hosking [281]) and LQ -moments (Mudholkar and Hutson [424]).

Example 1.9. The L -moments of the exponential distribution were calculated earlier in Example 1.8. Applying the formulas for τ_2 , τ_3 and τ_4 in (1.44), (1.42) and (1.43), we have

$$\tau_2 = \frac{1}{2}, \quad \tau_3 = \frac{1}{3}, \quad \tau_4 = \frac{1}{6}.$$

Thus, $(\tau_2, \tau_3) = (\frac{1}{2}, \frac{1}{3})$ and $(\tau_3, \tau_4) = (\frac{1}{3}, \frac{1}{6})$ are points in the τ_2 - τ_3 and τ_3 - τ_4 planes, respectively.

Example 1.10. The random variable X has generalized Pareto distribution with

$$Q(u) = \frac{b}{a} \{ (1-u)^{-\frac{a}{a+1}} - 1 \}, \quad a > -1, b > 0.$$

Then, straightforward calculations yield

$$\begin{aligned} L_1 &= b, & L_2 &= b(a+1)(a+2)^{-1}, \\ L_3 &= b(a+1)(2a+1)[(2a+3)(a+2)]^{-1}, \\ L_4 &= b(a+1)(6a^2+7a+2)[(a+2)(2a+3)(3a+4)]^{-1}, \end{aligned}$$

so that

$$\tau_2 = \frac{a+1}{a+2}, \quad \tau_3 = \frac{2a+1}{2a+3} \quad \text{and} \quad \tau_4 = \frac{6a^2+7a+2}{6a^2+17a+12}.$$

Then, eliminating a between τ_2 and τ_3 , we obtain

$$\tau_3 = \frac{3\tau_2 - 1}{\tau_2 + 1}.$$

Thus, the plot of (τ_2, τ_3) for all values of a and b lies on the curve $(\tau_2 + 1)(3 - \tau_3) = 4$. Note that the exponential plot is $(\frac{1}{2}, \frac{1}{3})$ which lies on the curve when $a \rightarrow 0$. The estimation and other related inferential problems are discussed in Chap. 7.

We now present probability weighted moments (PWM) which is a forerunner to the concept of L -moments. Introduced by Greenwood et al. [226], PWMs are of considerable interest when the distribution is expressed in quantile form. The PWMs are defined as

$$M_{p,r,s} = E[X^p F^r(X) \bar{F}^s(X)], \quad (1.45)$$

where p, r, s are non-negative real numbers and $E|X|^p < \infty$. Two special cases of (1.45) in general use are

$$\begin{aligned} \beta_{p,r} &= E(X^p F^r(X)) \\ &= \int x^p F^r(x) f(x) dx \\ &= \int_0^1 (Q(u))^p u^r du \end{aligned} \quad (1.46)$$

and

$$\begin{aligned}\alpha_{p,s} &= E(X^p \bar{F}^s(X)) \\ &= \int_0^1 (Q(u))^p (1-u)^s du.\end{aligned}\quad (1.47)$$

Like L -moments, PWMs are more robust to outliers in the data. They have less bias in estimation even for small samples and converge rapidly to asymptotic normality.

Example 1.11. The PWMs of the Pareto distribution with (see Table 1.1)

$$Q(u) = \sigma(1-u)^{-\frac{1}{\alpha}}, \quad \sigma, \alpha > 0,$$

are

$$\alpha_{p,s} = \sigma \int_0^1 (1-u)^{-\frac{p}{\alpha}+s} du = \frac{\sigma \alpha}{\alpha(s+1) - p}, \quad \alpha(s+1) > p.$$

Similarly, for the power distribution with (see Table 1.1)

$$Q(u) = \alpha u^{\frac{1}{\beta}}, \quad \alpha, \beta > 0,$$

we have

$$\beta_{p,r} = \alpha \int_0^1 u^{-\frac{p}{\beta}+r} du = \frac{\alpha \beta}{1 + \beta(r+1)}.$$

Further specializing (1.46) for $p = 1$, we see that the L -moments are linear combination of the PW moments. The relationships are

$$\begin{aligned}L_1 &= \beta_{1,0}, \\ L_2 &= 2\beta_{1,1} - \beta_{1,0}, \\ L_3 &= 6\beta_{1,2} - 6\beta_{1,1} + \beta_{1,0}, \\ L_4 &= 20\beta_{1,3} - 30\beta_{1,2} + 12\beta_{1,1} - \beta_{1,0}\end{aligned}$$

in the first four cases. Generally, we have the relationship

$$L_{r+1} = \sum_{k=0}^r \frac{(-1)^{r-k} (r+k)!}{(k!)^2 (r-k)!} \beta_{1,k}.$$

The conventional moments can also be deduced as $M_{p,0,0}$ or $\beta_{p,0}$ or $\alpha_{p,0}$. The role of PW moments in reliability analysis will be taken up in the subsequent chapters. In spite of its advantages, Chen and Balakrishnan [140] have pointed out

some infeasibility problems in estimation. While estimating the parameters of some distributions like the generalized forms of Pareto, the estimated distributions have an upper or lower bound and one or more of the data values lie outside this bound.

1.7 Diagrammatic Representations

In this section, we demonstrate a few graphical methods other than the conventional ones. The primary goal is fixed as the choice of model for the data represented by a quantile function. An important tool in this category is the Q - Q plot. The Q - Q plot is the plot of points $(Q(u_r), x_{r:n})$, $r = 1, 2, \dots, n$, where $u_r = \frac{r-0.5}{n}$.¹ For application purposes, we may replace $Q(u_r)$ by the fitted quantile function. One use of this plot is to ascertain whether the sample could have arisen from the target population $Q(u)$. In the ideal case, the graph should show a straight line that bisects the axes, since we are plotting the sample and population quantiles. However, since the sample is random and the fitted values of $Q(u)$ are used, the points lying approximately around the line is indicative of the model being adequate. The points in the Q - Q plot are always non-decreasing when viewed from left to right.

The Q - Q plot can also be used for comparing two competing models by plotting the r th quantile of one against the r th quantile of the other. When the two distributions are similar, the points on the graph should show approximately the straight line $y = x$. A general trend in the plot, like steeper (flatter) than $y = x$, will mean that the distribution plotted on the y -axis (x -axis) is more dispersed. On the other hand, S -shaped plots often suggest that one of the distributions exhibits more skewness or tail-heaviness. It should also be noted that the relationship in quantile plot can be linear when the constituent distributions are linearly related. This procedure is direct when the data sets from two distributions contain the same number of observations. Otherwise, it is necessary to use interpolated quantile estimates in the shorter set to equal the number in the larger sets. Often, Q - Q plots are found to be more powerful and informative than histogram comparisons.

Example 1.12. The times to failure of a set of 10 units are given as 16, 34, 53, 75, 93, 120, 150, 191, 240 and 390 h (Kececiloglu [322]). A Weibull distribution with quantile function

$$Q(u) = \sigma(-\log(1 - u))^{1/\lambda}$$

is proposed for the data. The parameters of the model were estimated by the method of maximum likelihood as $\hat{\sigma} = 146.2445$ and $\hat{\lambda} = 1.973$. The Q - Q plot pertaining to the model is presented in Fig. 1.1. From the figure, it is seen that the above model seems to be adequate.

¹There are different choices for these plotting points and recently Balakrishnan et al. [52] discussed the determination of optimal plotting points by the use of Pitman closeness criterion.

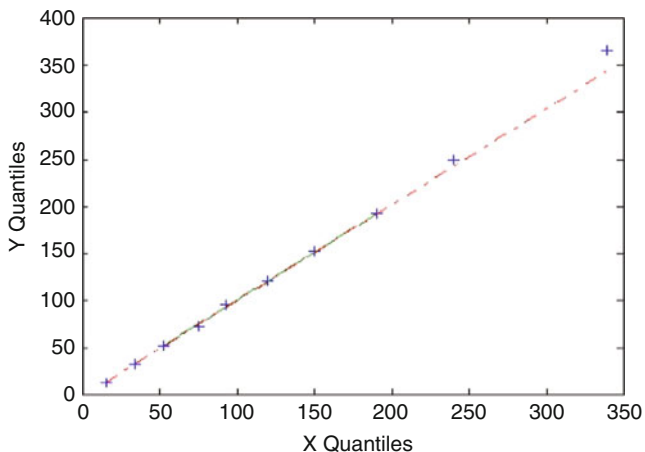


Fig. 1.1 *Q-Q* plot for Example 1.12

A second useful graphical representation is the box plot introduced by Tukey [569]. It depicts graph of the numerical data through a five-figure summary in the form of extremes, quartiles and the median. The steps required for constructing a box plot are (Parzen [484])

- (i) compute the median $m = \bar{Q}(0.50)$, the lower quartile $q_1 = \bar{Q}(0.25)$ and the upper quartile $q_3 = \bar{Q}(0.75)$;
- (ii) draw a vertical box of arbitrary width and length equal to $q_3 - q_1$;
- (iii) a solid line is marked within the box at a distance $m - q_1$ above the lower end of the box. Dashed lines are extended from the lower and upper ends of the box at distances equal to $x_{n:n} - q_3$ and $x_{1:n} - q_1$. This constitutes the *H*-plot, *H* standing for hinges or quartiles. Instead, one can use $\bar{Q}(0.125) = e_1$ and $\bar{Q}(0.875) = e_7$ resulting in *E*-box plots. Similarly, the quantiles $\bar{Q}(0.0625)$ and $\bar{Q}(0.9375)$ constitute the *D*-box plots;
- (iv) A quantile box plot consists of the graph of $\bar{Q}(u)$ on $[0, 1]$ along with the three boxes in (iii), superimposed on it.

Parzen [484] proposed the following information to be drawn from the plot. By drawing a perpendicular line to the median line at its midpoint and of length $\pm n^{-\frac{1}{2}} - (q_3 - q_1)$, a confidence interval for the median can be obtained. The graph $x = \bar{Q}(u)$ exhibiting sharp rises is likely to have a density with more than one mode. If such points lie inside the *H*-box, the presence of several distinct populations generating the data is to be suspected, while, if they are outside the *D*-box, presence of outliers is indicated. Horizontal segments in the graph may be the results of the discrete distributions. By calculating

$$\frac{\bar{Q}(\frac{1}{2}) - \frac{1}{2}[\bar{Q}(u) + \bar{Q}(1 - u)]}{\bar{Q}(1 - u) - \bar{Q}(u)}$$

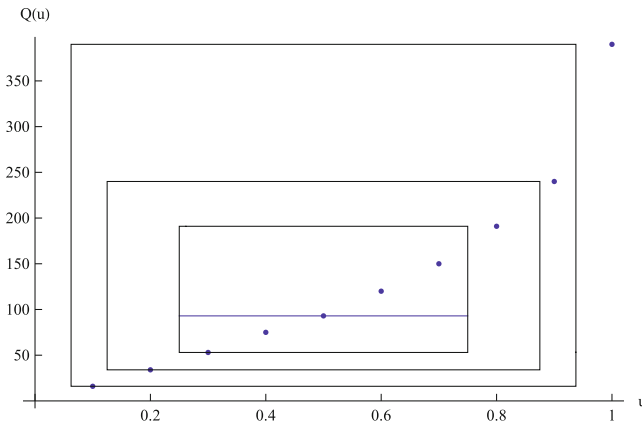


Fig. 1.2 Box plot for the data given in Example 1.12

for u values, one can get a feel for skewness with a value near zero suggesting symmetry. Parzen [484] also suggested some measures of tail classification.

Example 1.13. The box plot corresponding to the data in Example 1.12 is exhibited in Fig. 1.2. It may be noticed that the observation 390 is a likely outlier.

A stem-leaf plot can also be informative about some meaningful characteristics of the data. To obtain such a plot, we first arrange the observations in ascending order. The leaf is the last digit in a number. The stem contains all other digits (When the data consists of very large numbers, rounded values to a particular place, like hundred or thousand, are used a stem and leaves). In the leaf plot, there are two columns, first representing the stem, separated by a line from the second column representing the leaves. Each stem is listed only once and the leaves are entered in a row. The plot helps to understand the relative density of the observations as well as the shape. The mode is easily displayed along with the potential outliers. Finally, the descriptive statistics can be easily worked out from the diagram.

Example 1.14. We illustrate the stem-leaf plot for a small data set: 36, 57, 52, 44, 47, 51, 46, 63, 59, 68, 66, 68, 72, 73, 75, 81, 84, 106, 76, 88, 91, 41, 84, 68, 34, 38, 54.

3	4	6	8
4	1,	4,	6, 7
5	1	2 4	7 9
6	3	6 8	8 8
7	2	3 5	6
8	1,	4,	4, 8
9	1		
10	6		