A GENERAL MODEL OF RANDOM VARIATION

(Short Version)

Haim Shore

Department of Industrial Engineering and Management
Faculty of Engineering Sciences, Ben-Gurion University of the Negev
POB 653, Beer-Sheva 84105, Israel
shor@bgu.ac.il.

February, 2014

Comment: *This is a shortened version of Author's Accepted Manuscript of an article, by same title, accepted to be published in Communications in Statistics (Theory and Methods) [copyright Taylor & Francis], to be available online at: http://www.tandfonline.com/.*

ABSTRACT

A statistical distribution of a random variable is *uniquely* represented by its normal-based quantile function. For a symmetrical distribution it is S-shaped (for negative kurtosis) and inverted S-shaped (otherwise). As skewness departs from zero, the quantile function gradually transforms into a monotone convex function (positive skewness) or concave function (otherwise). Recently, a new general modeling platform has been introduced, Response Modeling Methodology, which delivers good representation to monotone convex relationships due to its unique "continuous monotone convexity (CMC)" property. In this shortened version of the original article, the CMC property is exploited to model the normal-based quantile function and explored using a test-set of 27 distributions.

1. Introduction

“All science is either physics or stamp collecting”. This assertion, ascribed to physicist Ernest Rutherford (the discoverer of the proton, in 1919) and quoted in Kaku (1994, p. 131), intended to convey a general sentiment that the drive to converge the five fundamental forces of nature into a unifying theory, nowadays a central theme of modern physics, represents science at its best. Furthermore, this is the right approach to the scientific investigation of nature. By contrast, at least until recently, most other scientific disciplines have engaged in taxonomy (“bug collecting” or “stamp collecting”). With “stamp collecting” the scientific inquiry is restricted to the discovery and classification of the "objects of enquiry" particular to that science; however, this never culminates, as in physics, in a unifying theory from which all these objects may be deductively derived as “special cases”.

Is statistics a science in a state of "stamp collecting"?

Observing the abundance of statistical distributions identified to-date an unavoidable conclusion is that statistics is indeed a science of "stamp collecting". This paper is an initial effort to unite all distributions under a unified "umbrella distribution" by presenting a new paradigm to model random variation. The new paradigm is based on three minimal assertions that hold true for all uni-modal non-mixture continuous distributions and for the requirements that a general model of random variation needs to fulfill:

Assertion [1]: Any distribution can be *uniquely* represented by its normal based quantile function (henceforth referred to as "Z-based quantile function", where Z is a standard normal variable); This function will henceforth be denoted $y=Q_Y(z;\theta)$, with Y being the response (the random variable (r.v.) modeled), θ is a vector of parameters (with values particular to the distribution being modeled) and $\{z,y\}$ are corresponding percentiles of Z and of Y;

Assertion [2]: Skewness and kurtosis jointly determine the *general* shape of $Q_Y(z;\theta)$ as follows:

- All symmetric r.v.s have S-shaped $Q_Y(z;\theta)$, for negative kurtosis, inverted S-shaped, for positive kurtosis, and linear, for Y normal;
- All non-symmetric r.v.s have $Q_Y(z;\theta)$ that, as skewness departs from zero, gradually transforms from S-shaped to monotone convex, for positively skewed distributions, or to monotone concave, for negatively skewed distributions;

Assertion [3]: A general model of $Q_Y(z;\theta)$ is required to deliver adequate representation to monotone convex relationships having any degree of convexity ("convex" will henceforth imply, unless otherwise stated, both convex and concave).

A crucial point for the feasibility of realizing the new paradigm is the existence of an effective general platform for modeling monotone convex relationships. In particular, Assertion [3] demands a model flexible enough to represent any degree of convexity. This can only be fulfilled if the model owns the "continuous monotone convexity" (CMC) property, discussed at some length in previous articles of this series. A proper platform to model the CMC property is Response Modeling Methodology (RMM; Shore, 2005, 2011, 2012a and references therein). With the CMC property, RMM transforms a distinct set of models of monotone convexity, which may be arranged in a hierarchical order according to their degree of convexity, into mere points on a continuous spectrum of monotone convexity. This renders RMM a unique modeling platform that is capable of delivering adequate representation to monotone convex relationships with any degree of monotone convexity (Assertion [3]). Furthermore, using RMM to model $Q_Y(z;\theta)$ renders differently shaped distributions into mere points on the continuous spectrum of monotone convexity, as the latter is spanned by RMM shape parameters.

A good analogy for the new paradigm of modeling random variation (by using RMM to model normal-based quantile functions) is modern perception of colors. Formerly treated as seperate entities, colors nowadays are perceived as mere points on the continuous spectrum of electromagnetic radiation. An RMM-based model for random variation may do likewise with respect to current statistical distributions.

The reader may find more details on the CMC property, as implemented with RMM, in the appendix (Appendix A in the original article).

Attempts to deliver general representation to distributions are not new in the published literature. These have culminated in what is generally recognized as "Distribution fitting". Various parameters-rich families of distributions have been suggested to attain that goal like the Pearson family, Johnson, generalized Lambda and others. A good recent source that reviews these endeavors is Karian and Dudewicz (2010). The current attempt to deliver a general model for random variation, based on RMM, departs from earlier efforts in three important respects:

1. The new paradigm is based on general shape-attributes common to all normal-based quantile functions (as delineated earlier in Assertion [2]). This is a departure from earlier efforts where the effectiveness of the families of distributions, used for distribution fitting, had only *empirically* been demonstrated or corroborated;

2. The new paradigm inserts a new element regarding how distributions are perceived. They are no longer separate entities but rather points arranged on a continuous spectrum of monotone convexity. This is a shift from current perception of statistical distributions in that the new paradigm delivers, via a model having the CMC property, a solid foundation to the practice of distribution fitting;

3. The new paradigm is well established theoretically, namely, the shape attributes related to in Assertion [2] above are shown to be preserved in the RMM-based model used for fitting. This implies that it does not need intensive *empirical* validation, as current models of distribution fitting require. Therefore, the generality of the new paradigm is better established.

In the next Section 2 we introduce intuitively the basis for the new paradigm. A sample of 27 distributions is explored to study properties that the general model is required to satisfy. Section 3 introduces RMM modeling of random variation in two modes:

A. The traditional RMM model, used in the past for general distribution fitting (Section 3.1);

B. RMM-based modeling according to the new paradigm (Section 3.2).

Some numerical examples for fitting current normal-based quantile functions via the RMM model, using minimization of the L2 norm as the objective function, are shown in Section 4.

2. Realization of the new paradigm for a set of 27 distributions

To appreciate whether a general model for random variation is feasible, based on Assertions [1] and [2] above, we select an arbitrary set of 27 differently-shaped distributions, and demonstrate some general properties shared by all. The selected distributions, with their parameters' values, are specified in Appendix B of the original article.

Plotting $Q_Y(z;\theta)$ for all distributions in the set we obtain Figure 1, where the quantile is given on the vertical axis in standardized form ("Std. Q"), namely:

$$x = \frac{y-\mu}{\sigma} ,$$

where x is the standardized y.

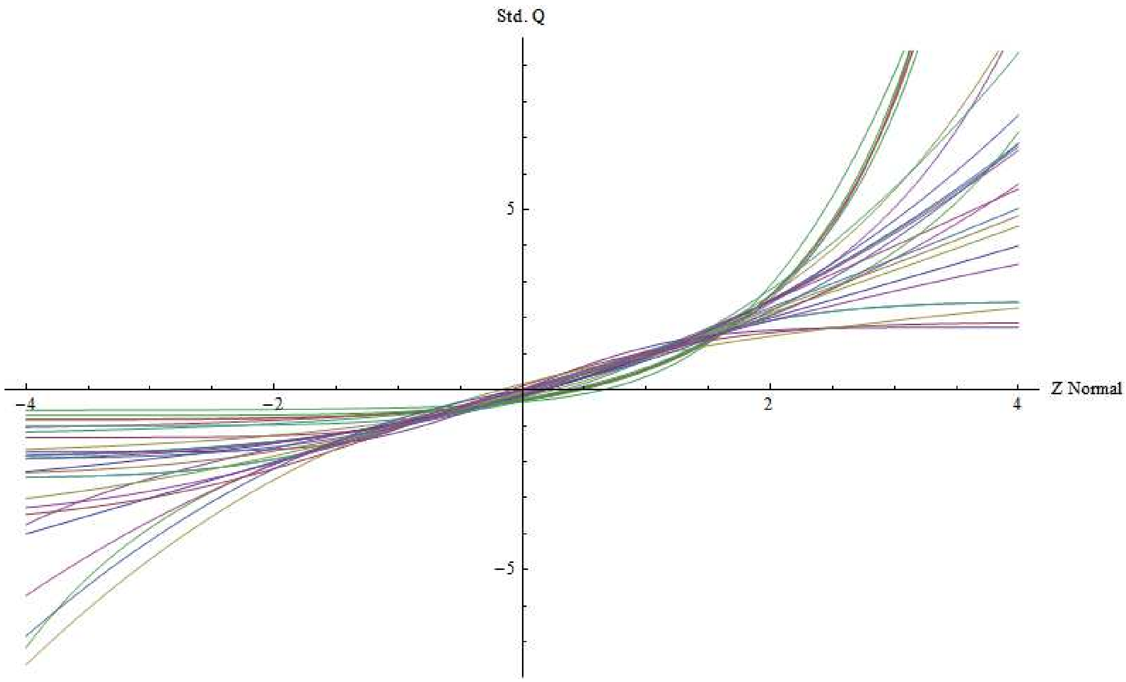

Figure 1. Quantile function for the standardized quantile ("Std. Q") as function of the standard normal quantile ("Z Normal") for 27 distributions (refer for details to Appendix B in original article).

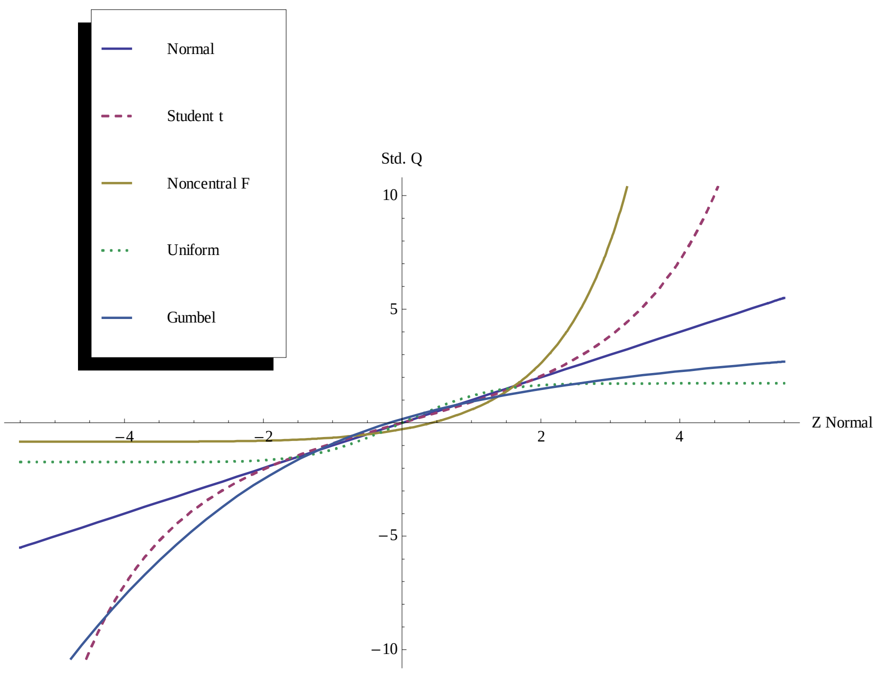

Figure 2. Quantile function for the standardized quantile ("Std. Q") as function of the standard normal quantile ("Z Normal") for: Normal(3,1/2) (Sk=0, Ku=0); Student-t(7) (Sk=0, Ku=2); Noncentral F-ratio(3,8,1) (Sk=6.38, Ku= Indeterminate); Uniform(2,6) (Sk=0, Ku=-1.2); Gumbel(4,2) (Sk=-1.14, Ku=2.4);

Note that all distributions seem to converge at around: $x=z$, with $x_U - x_L \cong 2$ (where x_L and x_U are the respective lower and upper quantiles for which x=z).

Observing Figure 1, several properties emerge. First, one realizes that, as expected, all quantile functions increase with z. Second, we notice five types of curves, according to the skewness and kurtosis of the distribution:

Type 1: Line - this is typical to a normal distribution (Y is normally distributed);

Type 2: Monotone concave $Q_Y(z;\theta)$: this is typical to distributions with highly negative skewness;

Type 3: Monotone convex $Q_Y(z;\theta)$: this is typical to distributions with highly positive skewness;

Type 4: Symmetric S-shaped $Q_Y(z;\theta)$: this is typical to distributions with zero skewness and negative kurtosis; the standardized quantile function, $Q_x(z;\theta)$, where x is the standardized quantile, is symmetric with respect to the line $x=z$, namely: $Q_x(z;\theta)=-Q_x(-z;\theta)$;

Type 5: Inverted symmetric S-shaped $Q_Y(z;\theta)$: this is typical to distributions with zero skewness and positive kurtosis; $Q_x(z;\theta)$ is symmetric with respect to: $x=z$;

Examples for all five types of function are displayed in Figure 2.

One realizes that skewness and kurtosis are major factors in determining which of these shape patterns represents a given distribution. A suggested general model for random variation will be required to model all five of these shape patterns. It is also interesting to note that transition from monotone concavity (for negatively skewed distributions) to monotone convexity (for positively skewed distributions) passes through S-shaped or inverted S-shaped patterns, dependent on the kurtosis value (as expounded earlier). Figure 2 demonstrates this.

3. RMM Modeling of Random Variation

3.1 Traditional modeling

As shown in the appendix (also Appendix A in the original article), a traditional RMM model for $Q_Y(z;\theta)$, as used in the past for RMM-based distribution fitting, is:

$$\log(y-L)=\log(M_Y-L)+\left(\frac{a}{b}\right)[(1+cz)^b-1]+dz+\varepsilon, \quad (1)$$

or its approximation (assuming $cz<<1$):

$$\log(y-L)=\log(M_Y-L)+\left(\frac{a}{b}\right)(e^{bz}-1)+cz+\varepsilon\ . \quad (2)$$

Models (1) and (2) have median M_Y, parameters {*a,b,c,d*} for (1) and {*a,b,c*} for (2), and a location parameter *L* that may occasionally be needed (see discussion in the appendix). Note that for (2) to represent the Z-based quantile function of a proper distribution function we need to have: a>0, c>0 (Shauly *et al.*, 2014).

Examples of (1) and (2) fitted to existing distributions or to sample data are given in Shore (2005, 2007, 2010, 2011), Shore and A'wad (2010) and Shauly and Parmet (2011). Examples for using (1) and (2) to replace (approximate) published models of monotone convex relationships may be found in Shore (2012b), Shauly *et al.* (2014) and Benson-Karhi *et al.* (2014). While the current RMM model for random variation has *empirically* proven effective (as do other families currently used in distribution fitting), it lacks the required theoretical justification, based on the general shape-characteristics defined in Assertion [2]. In the next Section 3.2 we introduce the new RMM-based model that is compatible with the new paradigm, as delineated in the Introduction.

3.2 RMM-based general model for random variation

In this subsection we develop an updated RMM model that may serve as a general model for random variation, compliant with the five types of shapes detailed in Section 2. Furthermore, a distinction is made between inherently symmetric distributions and symmetric distributions that are not so (more details later). To develop the new model, we first introduce a requirement, which relates to symmetric distributions. Denote by $Q_{xs}(z;\theta)$ the standardized Z-based quantile function of a *symmetric* distribution. Since for the latter:

$$Q_{xs}(z;\theta) \;=\; -Q_{xs}(-z;\theta) \;, \tag{3}$$

and since a Taylor series expansion, around $z=0$, of $Q_{xs}(z;\theta)$ would fulfil this requirement if, and only if, it contains only non-even powers of z, any general model for $Q_{xs}(z;\theta)$ needs to have this property. Denote by $g(z;\beta)$ a function, with parameters vector β, that we wish to use in modeling $Q_x(z;\theta)$. Define the following function:

$$h(z;\beta) = \frac{g(z;\beta) + g(-z;\beta)}{2}. \tag{4}$$

Provided that $g(z;\beta)$ itself is non-symmetric, a Taylor series expansion of (4) in terms of z, around $z=0$, would contain all *even* powers of z in the Taylor series expansion of $g(z;\beta)$ and only these terms. Therefore, a model for the standardized quantile of a symmetric distribution, which contains only *non-even* powers of z in its Taylor series expansion, is:

$$\begin{aligned}\tilde{Q}_{xs}(z;\tilde{\beta}) &= \left(\frac{1}{\sigma_Y}\right)[g(z;\beta) - h(z;\beta)] = \\ &\left(\frac{1}{\sigma_Y}\right)\left[g(z;\beta) - \frac{g(z;\beta)+g(-z;\beta)}{2}\right] = \left(\frac{1}{\sigma_Y}\right)\left[\frac{g(z;\beta)-g(-z;\beta)}{2}\right],\end{aligned} \tag{5}$$

where $\tilde{\beta}$ is a vector of parameters. It is obvious that: $\tilde{Q}_{xs}(z;\tilde{\beta}) = -\tilde{Q}_{xs}(-z;\tilde{\beta})$, as required of a model for $Q_{xs}(z;\theta)$. Note that (5) applies only to distributions that are inherently symmetric and not necessarily to distributions that can be made symmetric by a proper choice of its parameters. For example, a Weibull distribution with parameters $\{\alpha,\beta\}$ has skewness that depends only on α. For α = 3.602349, Weibull has practically zero skewness (skewness of the order of E-7). However, its Taylor series expansion contains both even and non-even power terms, namely, the Taylor series constraint that results in (5) does not apply to it.

Conversely, it does apply to Student's t distribution, which is always symmetric (*Sk=0*). We denote the latter case an inherently symmetric distribution. A general model for random variation should therefore be able to deliver representation to non-symmetric distributions and to symmetric distributions that are either inherently symmetric (always: *Sk=0*) or non-inherently symmetric (*Sk≈0*). We now formulate the final general model for $Q_x(z;\theta)$, denote this model $\tilde{Q}_x(z;\tilde{\theta})$, as a weighted average of the standardized $g(z;\beta)$ for inherently symmetric distributions (represented by (5)) and for other distributions:

$$\tilde{Q}_x(z;\tilde{\theta}) = (w)\frac{g(z;\beta) - g(-z;\beta)}{(2\sigma_Y)} + (1-w)\frac{g(z;\beta) - \mu_Y}{\sigma_Y}\ . \qquad (6)$$

In (6), w is a weighting factor and $\{\mu_Y, \sigma_Y\}$ are the mean and standard deviation of the un-standardized Y. Later we will discuss how w is determined. However, note that $w=1$ revokes (5), the model for inherently symmetric distributions.

Various models may be used for $g(z;\beta)$. As addressed earlier, higher skewness means larger convexity of $Q_x(z;\theta)$. Thus, to be able to represent any degree of skewness $g(z;\beta)$ has to own the CMC property. This is offered by the RMM model. Therefore, we adopt for $g(z;\beta)$ the RMM models (1) or (2). However we replace therein the median, M_Y, by an additional parameter, m, so that (1) becomes:

$$\log[g(z;\beta)] = \log(y-L) = \log(m-L) + \left(\frac{a}{b}\right)[(1+cz)^b - 1] + dz + \varepsilon, \qquad (7)$$

or its approximation (assuming $cz<<1$):

$$\log[g(z;\beta)] = \log(y-L) = \log(m-L) + \left(\frac{a}{b}\right)(e^{bz} - 1) + cz + \varepsilon\ . \qquad (8)$$

To preserve the exact value of the standardized median, M_x, in (6) (using (7) or (8)), the following relationship needs to be maintained (assuming $L=0$ in $g(z;\beta)$):

$$w = 1 - \frac{M_Y - \mu_Y}{m - \mu_Y},\ m \neq \mu_Y\ . \qquad (9)$$

For non-symmetric distributions ($m=M_Y \neq \mu_Y$): $w = 0$, as expected;

For an inherently symmetric distribution ($M_Y=\mu_Y$): $w = 1$, irrespective of m. Introducing m as an additional parameter in the model (namely, assuming $m \neq M_Y$), w from (9) should be introduced into (6) to allow m participate in determining the value of w (while minimizing the L2 norm).

Note that in (9) we do not constrain: $0 \leq w \leq 1$, thus allowing m (and therefore also w, via (9)) be determined by the fitting procedure.

Numerical examples in the original article (Shore, 2014) demonstrate the effectiveness of the new general model for random variation. Further examples relating to all distributions detailed in Appendix B of the original article are given in "Supplementary Material" therein.

In the next section we demonstrate the new paradigm with two examples.

4. Numerical examples

We fit (6), with constraint (9) and RMM model (8), to two distributions (*Sk* is the skewness measure, *Ku* is kurtosis measure, with *Ku=0* for the normal distribution):

* The logistic distribution (inherently symmetric): $Q_Y = -log\left[\frac{P}{1-P}\right]$ with:

$$Sk = 0,\ Ku = 1.2,\ M_Y = 0;$$

* Gamma distribution (non symmetric), with parameters {2,3}:

$$Sk = 2^{1/2},\ Ku = 3,\ M_Y = 5.035,\ \mu_Y = 6;$$

Fitting is performed by minimizing the L2 norm, an integrated functional distance between the quantile function (6) (with RMM model given by (2)) and the origin distribution:

$$(L2)^2 = \int_0^1 [Q_x(z(p);\theta) - \tilde{Q}_x(z(p);\tilde{\theta})]^2 dp \quad , \tag{10}$$

In applying the fitting procedure, it is assumed that the response location parameter *L*=0 (cases when this assumption is inappropriate are discussed in the original article).

The logistic distribution

Minimizing (10) on the interval: *p*={0.001,0.999}, we obtain:

a = 0.9925; b = −0.5687; c = 0.3188; m = 1.217; w=1.

Figure 3 shows the error curve as function of the CDF (*"P*"). Note that plot of Figure 3 displays results in the range P={0.005, 0.995}.

We realize that the error in the standardized quantile does not exceed about ±0.0001.

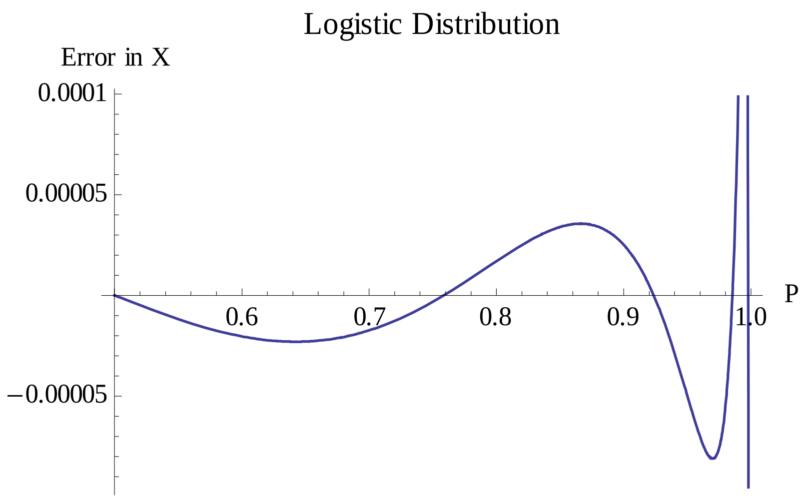

Figure 3. Error curve (as function of the CDF, denoted P) for fitting the general model to the logistic distribution.

Gamma distribution

Minimizing (10) we obtain:

a = 0.534368; *b* = –0.346663; *c* = 0.224263; *m* = 5.03577; *w* = –7E-4≈0.

Note that *m* is not the median (=5.03504) but slightly larger than the median however we can assume that *w=0*. Figure 4 shows the error curve as function of the CDF ("P"). We realize that the error in the standardized quantile does not exceed about ±0.0001.

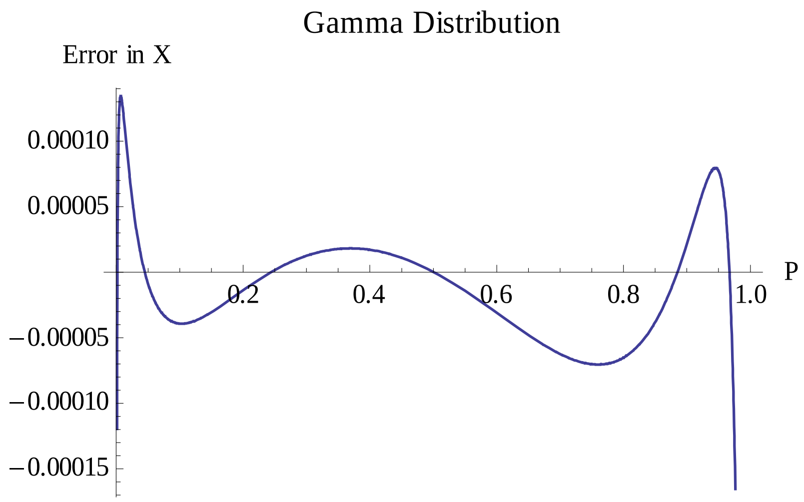

Figure 4. Error curve (as function of the CDF, denoted P) for fitting the general model to the gamma distribution.

5. Conclusions

Based on the new concept of continuous-monotone-convexity (CMC), modeled by RMM, a new paradigm is presented that perceives statistical distributions not as distinct discrete entities but rather as mere points on the CMC spectrum. The new paradigm regards deviations of quantiles of a particular statistical distribution (from respective quantiles of the fitted general model) as insignificant random deviations ("sampling errors"), not unlike the average of a particular sample is insignificant deviation from the population mean. For both inherently symmetric distributions and other distributions, the new model preserves same algebraic structure of the Taylor series expansion as the exact distribution. Observing the deviations from the fitted RMM-based model, one doubts that there is a sample-size large enough to differentiate between the approximating and approximated distributions (as have been shown, for example, in Shauly and Parmet, 2011, and repeatedly in an ongoing unpublished research).

Acknowledgment

TM Mathematica is a registered trademark of Wolfram Research.

REFERENCES

[1] Benson-Karhi D, Shore H, Malamud M (2014). Modeling fetal growth with Response Modeling Methodology (RMM). Submitted.

[2] Kaku M (1994). *Hyperspace- A Scientific Odyssey Through Parallel Universes, Time Warps and the Tenth Dimension.* Book. Oxford University Press Inc., NY.

[3] Karian ZA, Dudewicz EJ (2010). *Handbook of Fitting Statistical Distributions with R.* Taylor and Francis Group, LLC.

[4] Shauly M, Parmet Y (2011). Comparison of Pearson distribution system and response modeling methodology (RMM) as models for process capability analysis of skewed data. *Quality and Reliability Engineering International*. 27(5): 681–687.

[5] Shauly M, Parmet Y, Shore H (2014). Comparison of Pearson distribution system and Response Modeling Methodology (RMM) as models for distribution fitting. Under Review.

[6] Shore, H (2005). *Response Modeling Methodolog*y- *Empirical Modeling for Engineering and Science.* World Scientific Publishing Co. Ltd., Singapore.

[7] Shore H (2007). Comparison of Generalized Lambda Distribution (GLD) and Response Modeling Methodology (RMM) as general platforms for distribution fitting. *Communications in Statistics (Theory and Methods)*. 36(15): 2805-2819.

[8] Shore H (2010). Distribution fitting with the quantile function of Response Modeling Methodology. Chapter 13 in Karian, Z. A., Dudewicz, E. J. (Eds.). *Handbook of Fitting Statistical Distributions with R*. Taylor and Francis Group, LLC.

[9] Shore H (2011). Response Modeling Methodology - advanced review. *WIREs Computational Statistics*. 3(4): 357-372.

[10] Shore H (2012a). Estimating Response Modeling Methodology models - a focused article. *WIREs Computational Statistics.* 4: 323-333.

[11] Shore H (2012b). Modeling and monitoring ecological systems - a statistical process control approach. *Quality and Reliability Engineering International*. Accepted.

[12] Shore H (2014). A general model of random variation.*Communications in Statistics (Theory and Methods).* Forthcoming.

[13] Shore H, A'wad F (2010). Statistical comparison of the goodness-of-fit delivered by five families of distributions used in distribution fitting. *Communication in Statistics (Theory and Methods)*, 39(10): 1707-1728.

Appendix: Brief Tutorial on RMM

RMM models a monotone convex relationship between the percentile of a response, *Y*, the linear predictor (LP, a linear combination of covariates, denoted herewith η) and the respective percentile of the standard normal variate, *Z*. The *origin* RMM model describes a modeled response, *Y*, in terms of the LP, two possibly correlated zero-mean normal errors, ε_1 and ε_2 (with correlation ρ and standard deviations $\sigma_{\varepsilon 1}$ and $\sigma_{\varepsilon 2}$, respectively), and a vector of parameters $\{\alpha, \lambda, \mu\}$:

$$\log(Y) = \mu + \left(\frac{\alpha}{\lambda}\right)[(\eta + \varepsilon_1)^{\lambda} - 1] + \varepsilon_2 . \qquad \text{(A.1)}$$

Note that ε_1 implies that there is uncertainty (either measurement imprecision or otherwise) in the covariates included in the LP. This is in addition to uncertainty associated with the response (ε_2). One may realize that various common scientific and engineering models can be derived from (A.1). For example, ignoring the errors and the scale parameter, μ, one obtains from (A.1) for λ=0:

$$\log(Y) = \lim_{\lambda \to 0}\{(\frac{\alpha}{\lambda})[\eta^{\lambda} - 1]\} = (\alpha)\log(\eta) = \log(\eta^{\alpha}). \quad \text{(A.2)}$$

From (A.2), a linear relationship between the response and the LP is obtained for $\alpha = 1$ and a power relationship for $\alpha \neq 1$. An exponential relationship is obtained from (A.1) for $\lambda = 1$; an exponential-power relationship for $\{\lambda \neq 1, \lambda \neq 0\}$ and so on. In fact, all models that appear on the "Ladder of monotone convex relationships", a core concept of RMM, may be derived from (A.1). Note that the power and exponential relationships, associated with the inverse Box-Cox transformation, are special cases of RMM.

The quantile function associated with (A.1) is (Shore, 2005, 2011):

$$\log(y) = \log(M_Y) + \frac{(a\eta^b)}{b}\{[1+(c/\eta)z]^b - 1\} + (d)z + \varepsilon, \quad \text{(A.3)}$$

where ε is the model's zero mean normal error with constant variance, σ^2, $\{a,b,c,d\}$ are parameters, and

$$\log(M_Y) = \mu + \left(\frac{a}{b}\right)[\eta^b - 1] = \log(m) + \left(\frac{a}{b}\right)[\eta^b - 1], \quad \text{(A.4)}$$

M_Y being the median of Y (the 50% percentile of Y, corresponding to z=0), and μ (or m) is an additional parameter. Note that the LP (with k covariates) is defined by

$$\eta = \beta_0 + \beta_1 X_1 + .. + \beta_k X_k \, . \quad \text{(A.5)}$$

This implies that in estimating the median (A.4), apart from the parameters of LP only two additional RMM parameters need estimating (assuming, without loss of generality: *a*=1). If the response data contain values that change sign or if the lowest response value is far from zero (for example, when data are left truncated), a location parameter, *L*, may be added to the response so that (A.3) and (A.4) become, respectively:

$$\log(y-L)=\log(M_Y-L)+\frac{(a\eta^b)}{b}\{[1+(c/\eta)z]^b-1\}+(d)z+\varepsilon, \qquad \text{(A.6)}$$

$$\log(M_Y-L)=\mu+\left(\frac{a}{b}\right)[(\eta)^b-1]. \qquad \text{(A.7)}$$

Assume now that LP is constant (no systematic variation) and, furthermore, that, without loss of generality: η=1. We obtain from (A.6) an RMM model for a *Z*-based quantile function:

$$\log(y-L)=\log(M_Y-L)+(\frac{a}{b})\{[1+cz]^b-1\}+(d)z+\varepsilon, \qquad \text{(A.8)}$$

Assuming: $cz<<1$ we obtain an approximation to (A.8), the RMM model (2).