

Stochastic Modeling of Patient Arrival Offset Times in Scheduled Visits

Kamran Eftakhari¹, John Fontanesi^{1,*}, Gregory Feld², Daniel Bouland³, Ajit B. Raisinghani²,
Kirk Knowlton²

¹School of Medicine, University of California, San Diego, Center for Management Science in Health, San Diego, CA 92093, USA

²School of Medicine, University of California, San Diego, Division of Cardiology, San Diego, CA 92103, USA

³School of Medicine, University of California, San Diego, Division of Hospital Medicine, San Diego, CA 92103, USA

Abstract A new model for patient offset times (i.e., patient deviation from scheduled appointment time) is developed. In previous studies, offset times was mostly assumed to be sampled from a normal distribution. Alexopoulos et al.[1] offered Johnson SU as the most suitable fit. A thorough analysis of patient offset times, obtained from workflow observations in a broad sampling of ambulatory care sites, revealed these assumptions are often not valid. Although Johnson SU is still largely acceptable, it is not the most stable fitted distribution of the observed data. Our study suggests that three distributions (Generalized Logistic, Johnson SU and Log-Logistic) are more suited to modeling patient offset times with Hosking[2] Generalized Logistic (GL) distribution the most stable in its estimated parameters. We will also consider uncertainty associated with computing parameters of a Generalized Logistic distribution fitted to observed data. This model is central in devising efficient scheduling strategies to reduce patient waiting time and improve patient throughput and satisfaction.

Keywords Stochastic Arrival Offsets

1. Introduction

Healthcare both more efficient and patient centered. In the ambulatory setting this has frequently been interpreted as having patients spend more of their appointment time in direct interaction with the provider and less time waiting. Excessive wait time has been associated with lower patient satisfaction and compliance with treatment[8-11], patients leaving without completing their appointments[14], missed opportunities to provide preventive services[12], disgruntled staff[13], and reduced revenue to cost ratios. Underscoring the importance of reducing unnecessary patient waiting is the Joint Commission standard LD.3.15 requiring major health-care organizations to reduce unnecessary patient waiting as part of their certification program[15].

Crucial to minimizing wait times is understanding and controlling the consequences of patient scheduling. Not surprisingly, there is considerable interest in optimizing scheduling. The typical scheduling convention is to give patients an appointment at a specific time. Patients, however, commonly deviate from their scheduled appointment. Given the randomness of ‘timeliness,’ queues develop and providers find themselves either rushed to service the queue or idled as they wait for a patient to be “roomed”[6,7]. Clinical

researchers have tried to address the queuing problem with ad hoc experiments using ‘open access’ or simulations – results are mixed[3,16].

Several authors have examined queue creation using discrete event simulation[1,4,5] while others have analyzed patient arrival patterns to identify the appropriate probability distribution for realization[1,15-19]. Most have recommended that offset times (i.e., deviation from scheduled appointment time) be sampled from normal distribution. Alexopoulos et al.[1] have suggested Johnson SU probability distribution. The consequences of using each of these different distribution families in modeling arrival patterns, developing discrete event simulations or in creating optimization schedules could be significant and warrants more study.

The normal distribution is a continuous probability distribution that often gives a good description of data that cluster symmetrically around the mean. The graph of the associated probability density function is bell-shaped, with a peak at the mean. In the Central Limit Theorem, the sum of a number of i.i.d (independent and identically distributed) random variables with finite means and variances approaches a normal distribution as the number of variables increases. The theorem will hold even if random variables are not i.i.d., although some constraints on the degree of dependence and the growth rate of moments still have to be imposed. Central Limit Theorem is an appropriate model for averaging observations.

We will show that normal distribution is seldom suitable

* Corresponding author:

jfontanesi@ucsd.edu (John Fontanesi)

Published online at <http://journal.sapub.org/ajor>

Copyright © 2012 Scientific & Academic Publishing. All Rights Reserved

for representing patient offset times. Johnson distribution SU[1], although providing an acceptable fit, is not the most stable fitted distribution.

Estimation of Distribution Function mainly contains three steps: choice of a model, finding the parameters, and analysis of error (e.g., checking that the model does not contradict observation). These steps are the core of a parametric estimation procedure to model the distribution. The choice of a family of distributions $F(X; \theta)$ to model $F(x)$ often depends on experience from studies of similar experiments or by analysis of data.

Before attempting to fit a probability distribution to a set of observed data, it is worth first considering the properties of the variable in question. The properties of the distribution or distributions chosen to be fitted to the data should match those of the variable being modeled. As an example, range of variable should match that of fitted distribution. Any interpretation of data requires subjective inputs, usually in the form of assumptions about the variable. The key assumption here is that observed data is randomly sampled from a probability distribution we are attempting to identify. It is assumed the observed data are both as reliable and representative as possible; anomalies in the data were checked and unreliable data points discarded. We also paid attention to possible biases that could be produced by method of data collection.

We are going to look at techniques to interpret observed data for a variable in order to derive a distribution that realistically models its true variability and our uncertainty about that true variability.

In this study, we will first find estimated parameters of statistical distributions that best fit patient offset times. Second, we will study the over-the-samples stability of the estimated parameters of the statistical distributions that best fit patient offset times and choose among such best fit distributions the

one that exhibits the largest degree of stability in its estimated parameters. That is to say, if the entire body of data is a set S of n elements, x_1, x_2, \dots, x_n , then a subsample S_i of size $n_i < n$ is a proper subset of the set S (that is, $S_i \subset S$). Moreover, for the purpose of random sampling, the elements of the sub-sample S_i are randomly chosen from the elements of the set S . If n_i is sufficiently smaller than n , then from S one can draw many sub-samples, S_1, S_2, \dots, S_n . Any suitable statistical distribution can be fitted to the data in these samples to obtain its estimated parameters. Obviously, there will be sampling variations in the estimated parameters. If the sample variations are within reasonable limits, estimated parameters are stable over the sub-samples.

We have considered numerous distributions such as Beta, Burr (4P), Cauchy, Chi-Squared (2P), Dagum (4P), Erlang (3P), Error, Error Function, Frechet (3P), Gamma (3P), Generalized Extreme Value, Generalized Gamma (4P), Generalized Logistic, Gumbel Min, Gumbel Max, Generalized Pareto, Hypersecant, Inv. Gaussian, Johnson-SU, Kumaraswamy, Laplace, Levy (2P), Log-Logistic(3P), Logistic, Normal, Pearson-5 (3P), Pearson-6 (4P), Pert, Rayleigh (2P),

Weibull, and Wakeby. It is worth noting that all these distributions, except the Normal distribution, are either asymmetric or non-mesokurtic or both. We expect the best fit distributions to be both skewed and non-mesokurtic. The goodness-of-fit of the distributions is measured by three statistics pertaining to Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Chi-squared (CS) tests. In addition, three *information* criteria; SIC (Schwarz information criterion)[20], AICC (Akaike information criterion)[21-22], and HQIC (Hannan-Quinn information criterion)[23], are also used.

The remainder of this article is organized as follows: We briefly review some theoretical concepts in Section 2 to keep this article self-contained. Section 3 introduces the genesis and main features of the Generalized Logistic model and notes the model is leptokurtic and with skewness and kurtosis governed only by one parameter. This section also provides some results concerning the Maximum Likelihood (ML) and Method of Moments (MOM) estimates of the Generalized Logistic parameters and its quantiles. In Section 4, a simulation study is carried out in order to appraise the over-the-sample performance of different candidate distribution functions that best fit patient offset times data. Section 5 explains fitting second-order distributions to observed data points and reports the results of application of Generalized Logistic distribution to the available sampled data and error estimation using Generalized Bootstrap method.

2. Theoretical Review

2.1. Von Neumann's test for independence

Von Neumann[24] proposed what is now known as the Von Neumann ratio:

$$v(X) = \frac{\sum_{i=2}^n (\hat{\epsilon}_i - \hat{\epsilon}_{i-1})^2}{\sum_{i=1}^n \hat{\epsilon}_i^2}, \text{ where } \hat{\epsilon}_i = X_i - \hat{\mu}, \hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i.$$

This can be approximated as:

$$v_n = \sqrt{\frac{n^2 - 1}{n - 2}} \left[1 - \frac{\sum_{i=2}^n (X_i - X_{i-1})^2}{2(n-1)S_n^2} \right], \text{ where } S_n^2 \text{ is sample}$$

variance of the data.

If data are i.i.d., Von Neumann ratio distribution is very close to normal distribution. One can reject the hypothesis of independence at level α when $|v_n| > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ -quantile of the standard distribution. The value α is the user specified type I error (type I error is rejecting the null hypothesis when in fact it is valid). The p-value of this test is approximately $2[1 - \Phi(v_n)]$, where $\Phi(\cdot)$ is the CDF of standard normal distribution. The p-value of a test is the probability that a test statistics larger than the current one would be obtained if the hypothesized distribu-

tion were correct.

2.2. Parameter Estimation

The distribution parameters that make a distribution best fit the data can be determined in several ways, among them the method of maximum likelihood, moment matching, quantile matching, and least square[25]. The choice of estimators is governed by principles and criteria that provide objective basis for comparing the alternatives, among them sufficiency, completeness and ancillary. Assume x_1, x_2, \dots, x_n are the observed values of n i.i.d. random variables X_1, X_2, \dots, X_n , each X_i having a density function $f_{X_i}(x_i; \theta)$ identical to $f_X(x; \theta)$. An estimator of θ is some function of the random variables and thus may be written as $\hat{\theta}_n = \theta(x_1, \dots, x_n)$, a notion that emphasizes this estimator is itself a random variable.

Various variable criteria have been proposed for an estimate to satisfy, among these are unbiased, consistent and with low variance of $\hat{\theta}$. It would also be desirable if $\hat{\theta}$ has, either exactly or approximately, a normal distribution since well-known properties of normal distribution can then be used.

2.2.1. Maximum Likelihood Estimator

The Maximum Likelihood Estimator (MLE) method is fundamental in finding estimates of parameters $\hat{\theta} = \theta(x_1, \dots, x_n)$ in a statistical distribution model and is the most widely used. The theory of MLE estimates has deep consequences for many fields in statistics (see[26]). Statistical properties of the MLE are also useful, as will be later discussed. The Maximum Likelihood Method considers n independent observations X_1, X_2, \dots, X_n and study the likelihood function $L(X; \theta)$ defined as joint probability density for the observed dataset. The maximum likelihood estimator (MLE) of a parametric distribution are the values of parameters that maximize $L(\theta)$. Consider a probability distribution type defined by a parameter vector θ . The likelihood function $L(\theta)$ of set of n data points (x_i) could be generated from the distribution with probability density function $f(x; \theta)$ as:

$$L(X; \theta) = \prod_i f(x_i; \theta).$$

The MLE $\hat{\theta}$ is then the value of θ that maximizes $L(X; \theta)$.

$$L_n(X; \hat{\theta}) = \sup_{\theta \in \Theta} [L_n(X; \theta)],$$

or equivalently

$$\hat{\theta}_n = \arg \max_{\theta \in \Theta} [L_n(X; \theta)].$$

In a majority of cases, whenever the density function is well behaved

$$\left. \frac{\partial L(X; \theta)}{\partial \theta} \right|_{\hat{\theta}} = 0.$$

For some distribution types, the MLE calculation is a relatively simple algebraic problem; for others the differential equation is extremely complicated and is solved nu-

merically. It is known that MLE has some asymptotic properties among them:

$$\begin{cases} E(\hat{\theta}_n) \rightarrow \theta \\ Var(\hat{\theta}_n) \rightarrow \frac{-1}{E\left(\frac{\partial^2}{\partial \theta^2} \log L\right)} \quad \text{for } n \rightarrow \infty. \end{cases}$$

One very important point is that MLE depends strongly on the parametric family chosen. Numerous studies examining the “robustness” of the MLE have identified how “wrong” a model can be when the incorrect distribution family is used. The best justification for the MLE is in its asymptotic properties, it turns out to be asymptotically optimal.

2.2.2. Moment Matching Estimator

The sample moments are functions of an i.i.d. sample X_1, X_2, \dots, X_n whose probabilistic structure is determined a priori by the statistical model chosen. The probability distribution moments are often the best way to handle the unknown parameters θ . This relationship is exemplified by the raw moments below:

$$\mu_k = E(X^k) = \int_{x \in \mathbb{R}_X} x^k f(x; \theta) dx = \mu_k(\theta), \quad k = 1, 2, \dots$$

Given a random sample x_1, x_2, \dots, x_n , the k -th sample moment is

$$M_k = \sum_{j=1}^n x_j^k / n. \quad \text{The moment estimator of population pa-}$$

rameters are obtained by matching the sample moments to the corresponding population moments and solving the resulting equations simultaneously.

2.2.3. The least square method

Let X_1, \dots, X_n be random variables, not necessarily identically distributed, and set $E_{\theta}(X_i) = g_i(x_i; \theta)$, $i = 1, \dots, n$. The

least-squares method is estimating values θ , say $\hat{\theta}_{LS}$ such that the sum of square of error minimizes the loss function:

$$Q(x; \theta) = \sum_{i=1}^n [x_i - g(x_i; \theta)]^2,$$

over $\theta \in \Theta$. That is:

$$\left. \frac{\partial Q(x; \theta)}{\partial \theta} \right|_{\hat{\theta}_{LS}} = 0.$$

Least square method is not necessarily asymptotically efficient and can be quite sensitive to heavy tails (i.e., outliers/error contamination)

2.3 Fitting a second-order parametric distribution to observed data

The main issue in estimating the parameters of distribution from data is that uncertainty distributions of estimated parameters are usually linked together in some way. Historically it is assumed that parameter uncertainty distributions are normally distributed; however, this is not true for most cases.

To pursue fitting second-order distributions to observed data points, we need additional techniques for quantifying the uncertainty of distribution parameters. Among those techniques are bootstrap, Bayesian inference and some classical statistics methods. The parametric bootstrap technique is well suited, since one simply resamples from the MLE fitted distribution in the same size of the observed data. Data fitting using MLE again gives us random samples from the joint uncertainty distribution for the parameters.

2.3.1. Analysis of Estimation Error

The exact value of the estimation error is unknown-it is an uncertain value. The variability of the error can be studied using the following random variable,

$$\varepsilon = \theta - \hat{\theta},$$

the *estimation error*. For consistent estimators, ε tends to zero as n increases without bounds.

We can study the distribution of $\varepsilon, F_\varepsilon(e)$, which, for example, can be used to find intervals that, with high confidence, we can claim θ is in these intervals.

The ML estimators possess many good properties. For example, it can be shown (see[25] or, for a review,[26]) that ML method is a consistent estimator if $f(x; \theta)$ satisfy certain regularity conditions, and X_1, X_2, \dots, X_n be independent observation variables. The *consistent estimators* are defined as estimators that the error $\theta - \hat{\theta}$ tends to zero as the number of observations n goes to infinity[27].

It is shown in[25] that if $f(x; \theta)$ satisfies certain regularity conditions, ML estimators, behave asymptotically normal. Asymptotic normality means that for large n ,

$$P(\varepsilon \leq e) \approx \Phi(e / \hat{\sigma}_\varepsilon) \text{ where } \hat{\sigma}_\varepsilon = 1 / \sqrt{-\ddot{\ell}(\hat{\theta})}, \quad \ell(\hat{\theta}) = \ln L(\hat{\theta}),$$

and $\ddot{\ell}(\hat{\theta}) = \frac{\partial^2}{\partial \theta^2} \ell(\hat{\theta})$.

Estimation Error analysis can be performed numerically by parametric bootstrap method[28]. Bootstrap methods are most commonly used for complicated statistical problems, e.g. when the parameter θ is a large vector, or when an analytical approach is not possible.

For bootstrap methods, a computer program for Monte Carlo simulation is necessary. If the parameter $\hat{\theta}$, equivalently, the distribution $F_X(x; \hat{\theta})$ is known, such a program can simulate independent samples $x_i = (x_{i1}, \dots, x_{in})$, $i = 1, \dots, N$, where N is some large integer. All these samples have the same random properties as our initial sample x and from each sample estimated $\hat{\theta}_i = \hat{\theta}(x_i)$ are calculated

$$e_i^B = \hat{\theta} - \hat{\theta}_i^B, \quad i = 1, \dots, N.$$

The error distribution $F_\varepsilon^B(e)$ can be approximated by means of the empirical distribution of (e_1, e_2, \dots, e_N) , with increasing accuracy as N goes to infinity.

Let $F_\varepsilon^B(e)$ be the empirical distribution describing the variability of the sequence e_i^B . (Note that the empirical distribution depends both on the number n of observations in our original data set and the number N of bootstrap simulations). Usually N is much larger than n since it is only

limited by the computer time we wish to spend for the simulations. Finally, one can prove that, under suitable conditions, with $N > n$,

$$F_\varepsilon^B \rightarrow F_\varepsilon^B(e)$$

Using the last result, if n is large we have an approximation of the error distribution ε .

The bootstrap quantiles defined by

$$F_\varepsilon^B(e_{1-\alpha/2}^B) = \alpha / 2, \quad F_\varepsilon^B(e_{\alpha/2}^B) = 1 - \alpha / 2,$$

are close to the quantiles $e_{1-\alpha/2}, e_{\alpha/2}$. Thus an interval, which with (approximately) $1 - \alpha$ confidence, covers the unknown parameter θ is given by

$$\left[\hat{\theta} + e_{1-\alpha/2}^B, \hat{\theta} + e_{\alpha/2}^B \right]$$

2.4. Goodness-of-Fit Statistics

Many GOF (goodness-of-fit) statistics have been developed, but two are most commonly used. These are chi-square χ^2 and Kolmogorov-Smirnoff (K-S) statistics. The Anderson-Darling statistic is a modification of K-S statistics. The lower the value of these statistics, the closer the distribution fits the data. GOF statistics do not provide a true measure of the probability that the data actually come from the fitted distribution. Instead, they provide a probability that random data generated from the fitted distribution would have produced a GOF as low as that calculated for the observed data. Analysis of the χ^2 , K-S, and A-D statistics can provide confidence intervals proportional to the probability that fitted distribution could have produced the observed data.

Critical values are determined by the required confidence level α -they are the values of the goodness-of-fit statistics that have a probability of being exceeded that is equal to the specified confidence level. Critical values of K-S and A-D statistics have been found by Monte Carlo simulation[29]. K-S and A-D statistics are designed to test whether a distribution of known parameters could have produced the observed data. If the parameters of the fitted distribution have been estimated from the data, they will produce conservative results. One way to circumvent this problem is to use a portion of data for estimation and remaining data for GOF test.

2.4.1. The chi-square (χ^2) statistics

Chi-square (χ^2) statistics measures how well the expected frequency of the fitted distribution having a CDF $F(x)$ compares with the frequency of the n observed data points X_1, X_2, \dots, X_n . To conduct the most effective version of the test, we first divide the hypothesized distribution's support into k "equiprobable" non-overlapping intervals; we identify values a_0, a_1, \dots, a_K such that $a_i = F^{-1}(i/k)$, for $i = 1, 2, \dots, k$, where $F^{-1}(\cdot)$ is the inverse CDF. The respective intervals are $I_i = (a_{i-1}, a_i]$, for $i = 1, 2, \dots, k$. We then compare the number of observation that fall in each interval; O_1, O_2, \dots, O_k , to the corresponding expected number; E_1, E_2, \dots, E_k . The chi-square statistics is calculated:

$$\chi_{GOF}^2 = \sum_{i=1}^K \frac{(O_i - E_i)^2}{E_i},$$

where $E_i = n[F(a_i) - F(a_{i-1})] = n/k$, for $i = 1, \dots, k$.

Critical values for the χ^2 are found from the χ^2 distribution. The shape and range of the χ^2 distribution are defined by the degree of freedom d , where $d = k - s - 1$, s = number of parameters that are estimated. We reject the null hypothesis that $F(x)$ is the appropriate distribution, if $\chi^2_{GOF} > \chi^2_{1-\alpha, k-s-1}$, where $\chi^2_{1-\alpha, d}$ is the $(1-\alpha)$ -quantile of chi-square distribution with d degree of freedom.

Since the χ^2 statistics sums of the square of all of the error $\{O_i - E_i\}$, it can be disproportionately sensitive to any large errors. However, it is very dependent on the number intervals. For better results, n usually needs to be sufficiently large and k sufficiently small that $E_i = n/k \geq 5$. It is recommended that the number of intervals to be chosen using Scott's [30] formula $k \approx 5/3\sqrt[3]{n}$.

2.4.2. Kolmogorov-Smirnoff (K-S) Statistic

Kolmogorov-Smirnoff (K-S) Statistics measures the vertical distance between CDF of the fitted distribution function and CDF of the observed data.

Assume data X_1, X_2, \dots, X_n arising from a continuous distribution having a CDF $F(x)$. Now let $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ denote the order statistics based on the sample X_1, X_2, \dots, X_n . The K-S statistics D_n is defined as

$$D_n = \sup_x [F_n(x) - F(x)],$$

where D_n is known as K-S distance, n is the number of observed data points, $F_n(x) = i/n$ for $i = 1, \dots, n$, where i is the commutative rank of the data point, and $F(x)$ is the distribution function of fitted distribution. It is well known Glivenko-Cantelli lemma [31] that, as the sample size n becomes large, the empirical CDF $F_n(x)$ converges uniformly to $F(x)$ for all x .

The K-S test quantifies both the maximum deviation of empirical CDF above or below the uniform line. The upper $F_U(i)$ and lower $F_L(i)$ empirical CDF are calculated as follows:

$$F_L(i) = \frac{i-1}{n}, \quad F_U(i) = \frac{i}{n}.$$

The K-S test rejects the hypothesized distribution when the test statistics

$$D_n = \max \left\{ \max_{1 \leq i \leq n} \left| \frac{i}{n} - F(X_{(i)}) \right|, \max_{1 \leq i \leq n} \left| F(X_{(i)}) - \frac{i-1}{n} \right| \right\},$$

is larger than a tabulated quantile based on the sample size and the type I error α .

The K-S statistic is generally more useful than χ^2 statistic in that the data are assessed at all data points which avoids the problem of determining the number of intervals into which the data must be split. However, its value is only determined by the one largest discrepancy and takes no account for lack of fit across the remainder of distribution.

The vertical distance between the observed distribution $F_n(x)$ and the fitted distribution $F(x)$ at any point has a

distribution with a mean of zero and a standard deviation σ_{K-S} given by binomial theory:

$$\sigma_{K-S} = \sqrt{\frac{F(x)[1-F(x)]}{n}}$$

This indicates that the position of D_n along the x axis is more likely to occur where σ_{K-S} is greatest, which generally is away from the low-probability tails. This insensitivity of K-S statistic to lack fit at the extremes of the distributions is corrected for in Anderson-Darling statistic.

2.4.3. Anderson-Darling (A-D) statistic

The A-D statistic is defined as:

$$A_n^2 = \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \Psi(x) f(x) dx,$$

where

$$\Psi(x) = \frac{n}{F(x)[1-F(x)]} = \frac{1}{\sigma_{K-S}^2}.$$

n is the number of observed data points, $F(x)$ is the CDF of fitted distribution, $f(x)$ is the density function of fitted distribution, $F_n(x) = i/n$, for $i = 1, \dots, k$ cumulative rank of the observed data point and k is the number of non-overlapping intervals.

The Anderson-Darling statistic is an improved version of Kolmogorov-Smirnoff statistic. $\Psi(x)$ compensates for the variance of the vertical deviation distance between sample distribution and fitted distribution (σ_{K-S}). $f(x)$ weights the distance by the probability that a value be generated at that x value. The vertical distances are integrated over all values of x to make maximum use of observed data (the K-S static only look at the maximum deviation distance).

The A-D statistic

$$A_n^2 = \int_{-\infty}^{\infty} \frac{[F_n(x) - F(x)]^2}{\sigma_{K-S}^2} f(x) dx,$$

is therefore generally a more useful measure of goodness of fit than the K-S, especially where it is important to place equal emphasis on fitting a distribution at the tails as well as at main body. Nonetheless, it still has the same problem as K-S statistic i.e., the fitted distribution should, in theory, not be estimated from the data.

2.4.4. A better Goodness-of-Fit Measure

For reasons explained above, the chi-square, Kolmogorov-Smirnoff and Anderson-Darling goodness-of-fit statistics are technically all inappropriate as a method of comparing fits of distributions to data. They are also limited to having precise observations and cannot incorporate censored, truncated or binned data. Realistically, most of the time we are fitting a continuous distribution to a set of precise data observations and, under these circumstances, Anderson-Darling proves adequate. However, for important work we should instead consider using statistical measure of fit called *information criteria*. Let n be the number of observations, s number of parameters to be estimated and L_{\max} be the maximized value of likelihood function.

SIC (Schwarz information criterion, aka Bayesian information criterion, BIC)[20]

$$SIC = \ln[n]s - 2\ln[L_{\max}]$$

AIC_C (Akaike information criterion)[21-22]

$$AIC_c = \left(\frac{2n}{n-s-1} \right) s - 2\ln[L_{\max}]$$

HQIC (Hannan-Quinn information criterion)[23]

$$HQIC = 2\ln[\ln[n]]s - 2\ln[L_{\max}]$$

The aim is to find the model with the *lowest* value of the selected information criterion. The $-2\ln[L_{\max}]$ term appearing in each formula is an estimate of the deviance model fit. The coefficients of S in the first part of each formula, shows the degree by which the number of model parameters is being penalized. For $n \geq 20$ the SIC[20] is the strictest in penalizing loss of degree of freedom by having more parameters in the fitted model. For $n \geq 40$ AIC_C ([21-22]) is the least strict of the three, and HQIC[23] is in between.

3. The Model

A theoretical representation of patient offset times distribution have traditionally relied on two-parameter cumulative density functions because these are relatively easy to estimate. However, the two-parameter models cannot deal with the existence of leptokurtic and skewness without the introduction of assumptions that limit their goodness-of-fit. Instead, Alexopoulos *et al.*[1] offered Johnson SU which treats skewness nicely but here we will show that it lacks the stability properties of Generalized Logistic distribution.

3.1. Genesis and Properties of the Hosking Generalized Logistic Distribution

The generalized logistic distributions are very useful classes of density functions as they possess a wide range of indices of skewness and kurtosis. Therefore, an important application of the generalized logistic (GL) distribution is its use in studying robustness of estimators and tests.

The GL distribution is a generalization of the two-parameter logistic distribution and is also a special case of the kappa distribution[32]. This generalization of the logistic distribution differs from other distributions defined in the literature. The cumulative distribution function and the probability density function of the GL distribution are defined respectively[2].

Let X be a positive continuous random variable that belongs to the family of Hosking GL distribution with three parameters. Its CDF and density distribution is:

$$F(X) = \left[1 + \left\{ 1 - \beta \left(\frac{x - x_0}{\alpha} \right) \right\}^{1/\beta} \right]^{-1}$$

$$f(x) = \frac{1}{\alpha} \left\{ 1 - \beta \left(\frac{x - x_0}{\alpha} \right) \right\}^{1/\beta - 1} \left[1 + \left\{ 1 - \beta \left(\frac{x - x_0}{\alpha} \right) \right\}^{1/\beta} \right]^{-2}$$

where x_0 is the location parameter, α is the scale parameter, and β is the shape parameter. The range of possible values for the GL distribution is given by

$$x_0 + \alpha / \beta \leq x < \infty \text{ for } \beta < 0$$

$$-\infty < x < x_0 + \alpha / \beta \text{ for } \beta > 0.$$

Note that as a special case, if $\beta = 0$ then the GL distribution is reduced to the two-parameter logistic distribution. Additional generalizations of the logistic distribution are discussed[33].

The mean, variance, and Fisher's coefficient of are[33]

$$\mu = x_0 + \frac{\alpha}{\beta}(1 - g_1), \quad \sigma^2 = \frac{\alpha^2}{\beta^2}(g_2 - g_1^2),$$

$$\gamma_1 = \frac{E[(X - E[X])^3]}{E[(X - E[X])^2]^{3/2}} = \frac{\beta}{|\beta|} \left(\frac{-g_3 + 3g_1g_2 - 2g_1^3}{(g_2 - g_1^2)^{3/2}} \right),$$

where $g_r = \Gamma(1+r\beta)\Gamma(1-r\beta)/\Gamma(1)^2$, $\Gamma(\cdot)$ is the gamma function, and g_r exists only if $|\beta| \leq 1/r$. Since γ_1 is location and scale invariant the skewness of the distribution depends only on parameter β . A random variable X with generalized logistic distribution has a variance depending on the parameters β and σ . The quantile estimator \hat{x}_T of the GL distribution can be obtained by substituting $F(x) = 1 - T$ and solving for x

$$\hat{x}_T = \hat{x}_0 + \frac{\hat{\alpha}}{\hat{\beta}} \left\{ 1 - (T - 1)^{-\hat{\beta}} \right\}$$

where \hat{x}_0 , $\hat{\alpha}$, and $\hat{\beta}$ are the parameter estimators, and T is the return period[33].

Method of moments (MOM)

The skewness coefficient γ_1 of the GL distribution is only a function of the shape parameter β . Then β can be approximated as follows[34]

$$\beta = \frac{2}{3\pi} \tan^{-1}(-0.5948\gamma_1).$$

A more precise estimate of the shape parameter can be obtained using a numerical approximation. The $\hat{\beta}$ that minimizes[33]

$$G(\hat{\beta}) = \frac{\hat{\beta}}{|\hat{\beta}|} \left(\frac{-\hat{g}_3 + 3\hat{g}_1\hat{g}_2 - 2\hat{g}_1^3}{(\hat{g}_2 - \hat{g}_1^2)^{3/2}} \right) - \hat{\gamma}_1,$$

is an approximation for the shape parameter. Once the shape parameter is known, \hat{x}_0 , and $\hat{\alpha}$ can be obtained as follows:

$$x_0 = \hat{\mu} - \frac{\hat{\alpha}}{\hat{\beta}} \left\{ 1 - \Gamma(1 + \hat{\beta})(1 - \hat{\beta}) \right\}$$

$$\hat{\alpha} = \frac{\hat{\sigma}\hat{\beta}}{\left\{ \Gamma(1 + 2\hat{\beta})\Gamma(1 - 2\hat{\beta}) - \Gamma^2(1 + \hat{\beta})\Gamma^2(1 - \hat{\beta}) \right\}^{1/2}}.$$

Method of maximum likelihood (ML)

Consider a sample of size n of independent positive random variables X_1, \dots, X_n . Let $\theta = (\beta, \lambda, \delta)$ the log-likelihood function of the GL distribution is given by[33]

$$\ell(x; x_0, \alpha, \beta) = -n \ln(\alpha) - (1 - \beta) \sum_{i=1}^n y_i - 2 \sum_{i=1}^n \ln(1 + e^{-y_i})$$

where $y_i = -\frac{1}{\beta} \ln \left\{ 1 - \beta \left(\frac{x_i - x_0}{\alpha} \right) \right\}$, n is the sample size, and \ln represents the natural logarithm. The MLEs $\hat{\theta}_n = (\hat{x}_0, \hat{\alpha}, \hat{\beta})$ are

obtained from the maximization of $\ell(\theta)$ as the solution of the following likelihood equations or score functions:

$$\frac{\partial \ell(\theta)}{\partial x_0} = \frac{Q}{\alpha} = 0, \quad \frac{\partial \ell(\theta)}{\partial \alpha} = \frac{P+Q}{\alpha\beta} = 0, \quad \frac{\partial \ell(\theta)}{\partial \beta} = \frac{R}{\beta} - \frac{(P+Q)}{\beta^2} = 0$$

where

$$P = n - 2 \sum_{i=1}^n \frac{e^{-y_i}}{1 + e^{-y_i}}, \quad Q = (\beta - 1) \sum_{i=1}^n e^{\beta y_i} + 2 \sum_{i=1}^n \frac{e^{(\beta-1)y_i}}{1 + e^{-y_i}},$$

$$R = n - \sum_{i=1}^n y_i + 2 \sum_{i=1}^n \frac{y_i e^{-y_i}}{1 + e^{-y_i}}$$

The system does not admit any explicit solution; therefore the ML estimates $\hat{\theta}_n = (\hat{x}_0, \hat{\alpha}, \hat{\beta})$ can be obtained only by means of numerical procedures.

4. Best Fit Parameter Analysis

4.1. Data collection

A total of 738 patient observations were obtained from a variety of ambulatory care clinics as part of ongoing workflow data collection efforts. These observations were collected using a workflow data acquisition tool described in Fontanesi et al.[8]. This tool, the Observational Checklist of Patient Encounters (OCPE), includes data fields to record individual patient scheduled appointment times and actual observed arrival times.

4.2. Data Analysis

Three distributions emerge as best fit based on GOF tests and *information criteria*: Generalized Logistic, Johnson SU and Log-Logistic. In the majority of cases, either Generalized Logistic or Log-Logistic does better than Johnson SU on the criterion of KS or AD test. However, on CS tests, Johnson SU emerges stronger than on KS or AD test. It may be noted that AD weights the fit more to the tails and CS weights the overall fit more. On information criteria test, again GL and Log-Logistic performed better than Johnson SU. This simply could be explained that information criteria penalizes for larger number of parameters (GL and Log-Logistic are 3 parameter distributions and Johnson is 4 parameter).

Algebraic form of the pdf of Generalized Logistic, Log-Logistic, and Johnson SU distributions are given as

i) Generalized Logistic Distribution

$$f(x) = \frac{1}{\alpha} \left\{ 1 - \beta \left(\frac{x - x_0}{\alpha} \right) \right\}^{1/\beta - 1} \left[1 + \left\{ 1 - \beta \left(\frac{x - x_0}{\alpha} \right) \right\}^{1/\beta} \right]^{-2},$$

Where x_0, α, β respectively are; continuous location parameter, scale parameter, and shape parameter.

ii) Log-Logistic (3P) Distribution

$$f(x) = \frac{\alpha}{\beta} \left(\frac{x - \gamma}{\beta} \right)^{\alpha - 1} \left[1 + \left(\frac{x - \gamma}{\beta} \right)^{\alpha} \right]^{-2},$$

where γ, β, α respectively are; continuous location parameter, scale parameter, and shape parameter.

iii) Johnson SU Distribution

$$f(x) = \frac{\delta}{\lambda \sqrt{2\pi} \sqrt{z^2 + 1}} \exp \left\{ -\frac{1}{2} \left[\gamma + \delta \ln(z + \sqrt{z^2 + 1}) \right]^2 \right\}$$

where

$$z = \frac{x - \xi}{\lambda},$$

and $\xi, \lambda, \delta, \gamma$ are respectively; continues location, scale ($\delta > 0$), and shape ($\delta > 0$) parameters.

The illustrative fits of Generalized Logistic, Johnson SU, and Log-Logistics distributions to sample data are presented in Fig.-1.

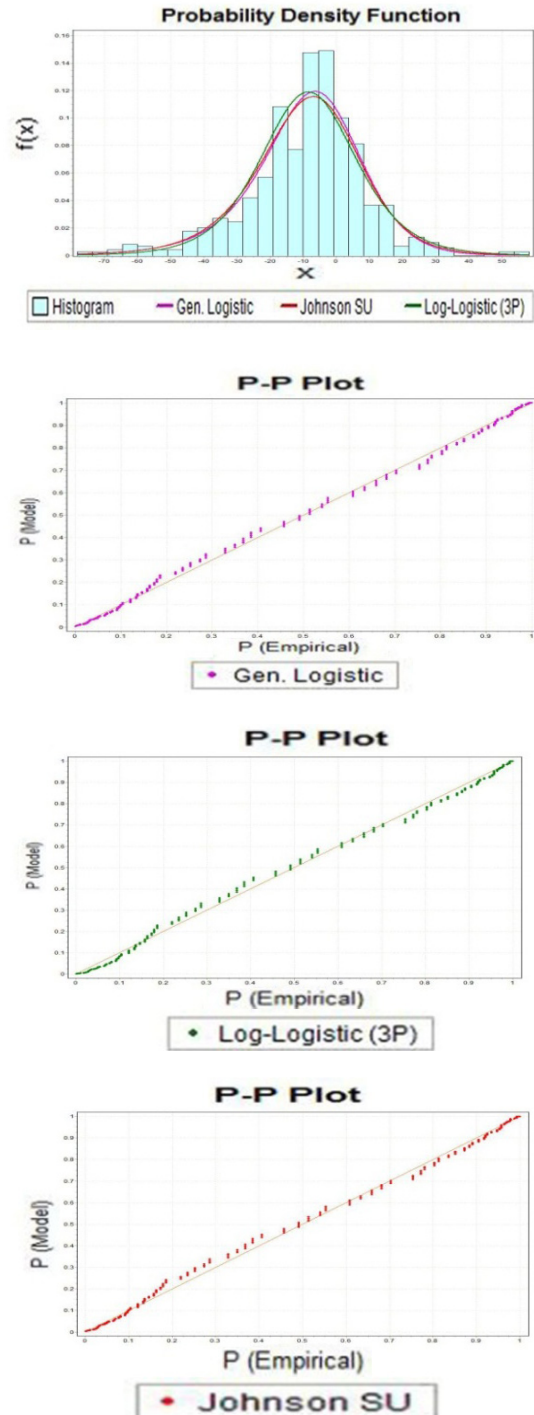


Figure 1. Generalized Logistic, Johnson SU, and Log-Logistics distributions fit to sample data

Table 2.1. Measures of Central Tendency and Dispersion of Parameters of Generalized Logistic Distributions over Sub-samples (Confidence level %95, Sample Size 100)

Parameter	Median	Mean	Mean LCL	Mean UCL	Std. Dev.	Mean LCL	Mean UCL
β	0.077	0.0781271	0.07383052	0.0819235	0.02068479	0.0182438	0.023712
α	9.5212	9.541834	9.46379622	9.60839026	0.3532221	0.31375933	0.40906781
x_0	-7.9558	-7.983801	-8.0881491	-7.8724036	0.5815828	0.51265048	0.67679522

Table 2.2. Measures of Central Tendency and Dispersion of Parameters of Johnson SU and Log-Logistic (3P) Distributions over Sub-samples (Confidence level %95, Sample Size 100)

Parameter	Johnson SU distribution				Lognormal(3P) Distribution		
	γ	δ	λ	ξ	α	β	γ
Mean	0.4963	2.0333	30.895	-0.744	265830162.78	2509587467.30	-2466933087.40
Median	0.3955	1.9943	30.389	-1.899	201130000.00	1919900000.00	-1919900000.00
Lower Limit	0.4325	1.9925	30.291	-1.717	219776380.18	2061683711.70	-2925285668.37
Upper Limit	0.5602	2.0741	31.498	0.229	311883945.38	2957491222.90	-2008580506.43
Sample Std Dev	0.3216	0.2058	3.042	4.904	232100540.02	2257332572.62	2309992264.72
Lower Limit	0.282377	0.18069	2.670869	4.305603	203785863	1981953451	2028189021
Upper Limit	0.373607	0.239067	3.533779	5.696666	269625454.1	2622287393	2683460854

4.3. Over-the-Samples Stability of the Estimated Parameters

Of the statistical distributions considered, the one that exhibits the largest degree of stability in its estimated parameters will be selected as the best fit best to patient offset times. For this analysis $N = 100$ subsets of data each of length $n = 600$ have been drawn randomly from a 738 points main sample data set. Estimated parameters $\hat{\theta}_i$, $i = 1, 2, \dots, N$ are tabulated in Tables 1.1, 1.2. Measures of central tendency and dispersion of the estimated parameters are presented in tables 2.1 and 2.2. The two measures of central tendency (median and mean) for all the parameters indicate their distributions are almost symmetrical and standard deviations are much smaller with respect to means[35].

It becomes readily apparent that the estimated parameters of the Generalized Logistic distribution exhibit better over-the-samples stability than the other two distributions.

5. Fitting Second-Order Distributions to Observed Data Points: the Generalized Bootstrap (GB) Analysis

In order to calculate uncertainty of distribution parameters[36] we will use the method of Generalized Bootstrap. The essence of the Generalized Bootstrap (GB) Method is to fit a distribution to the available data and then take samples from the fitted distribution (Bootstrap Method BM generally works with samples from the data). This method has been shown to perform better than the BM when the number of data points is not very large and do as well as the BM when the number of data points is large. Sun and Muller[37] have an excellent exposition with real-data examples.

Suppose that we are interested in $\theta = \theta(F)$ some function of the distribution $F(\cdot)$, and $F(\cdot)$ is unknown. However, we have a random sample X_1, X_2, \dots, X_n , from $F(\cdot)$, and we want to estimate $\theta(F)$. The Generalized

Bootstrap (GB) approaches the problem as follows[38]:

Suppose that one would typically estimate θ by $\hat{\theta}(x_1, x_2, \dots, x_n)$. Then, instead, proceed as follows: First, estimate F by \hat{F} . Second, independently generate N random samples of n from \hat{F} , and estimate $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_N$ for each sample. Third, use the sample $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_N$ to estimate θ . For example, one may calculate[37]

$$\bar{\theta} = \left(\frac{1}{N}\right) \sum_{i=1}^N \hat{\theta}_i, \quad \bar{\sigma}^2 = \left(\frac{1}{N-1}\right) \sum_{i=1}^N (\hat{\theta}_i - \bar{\theta})^2.$$

which give the sample mean and sample variance, respectively, of the GB estimators. Then, assuming approximate normality, we have

$$Var(\bar{\theta}) = Var(\hat{\theta}_i) \approx \frac{\bar{\sigma}^2}{N},$$

and an approximate $100(1 - \alpha)\%$ confidence interval for θ based on standard method is

$$\bar{\theta} \pm (1 - \frac{\alpha}{2})\bar{\sigma}.$$

A widely used alternative to standard method is "percentile method," which uses the upper and lower $\alpha/2$ percentiles of the GB sample estimators as the confidence interval. Specifically, the percentile method proceeds as follows: Place the N estimates $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_N$ in increasing numerical order, obtaining

$$\hat{\theta}_{(1)} \leq \hat{\theta}_{(2)} \leq \dots \leq \hat{\theta}_{(N)}.$$

The percentile methods' (approximate) $100(1 - \alpha)\%$ confidence interval for θ is

$$(\hat{\theta}_{(a)}, \hat{\theta}_{(b)}), \text{ where } a = \left\lfloor N \frac{\alpha}{2} \right\rfloor, \quad b = \left\lceil N \left(1 - \frac{\alpha}{2}\right) \right\rceil.$$

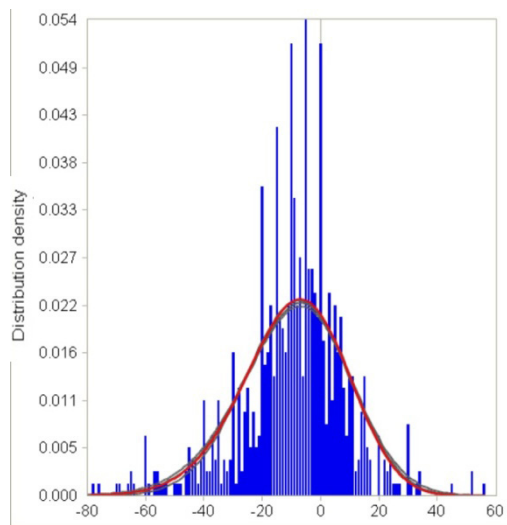
Sun and Muller-Schwarze[36] compared the performance of bootstrap method and generalized bootstrap and concluded that GB is more consistent in parameter estimation than BM. Asymptotic properties of GB have been shown by[28].

In this part of analysis we will independently generate $N = 517$ random samples of length $n = 738$ from a GL distribution

Table 3.2. Generalized Bootstrap Analysis of 517 sets of data of length 738

Parameter	Mean	Std. Dev.	99% Confidence Level				95% Confidence Level			
			Mean LCL	Mean UCL	Std. LCL	Std. UCL	Mean LCL	Mean UCL	Std. LCL	Std. UCL
β	0.082014893	0.020822709	0.07944258	0.084278897	0.01926404	0.02255096	0.080353383	0.083710216	0.01967632	0.022068573
α	9.566171262	0.327452621	9.53142313	9.607352605	0.3061453	0.34937013	9.537138208	9.592072363	0.310485248	0.347857048
x_0	-7.988701553	0.606562029	-8.0584773	-7.92046819	0.5577377	0.65340215	-8.04050821	-7.94285716	0.570660614	0.645357102

$F_{GL}(x_0 = -7.984179901, \alpha = 9.560168385, \beta = 0.0817128874)$, and estimated $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_{517}$ for each sample. Estimated values of parameters $\hat{\theta}_i$, $i = 1, 2, \dots, 517$ are shown in Table 3.1. The calculated values of mean, standard deviation and their lower and higher limits for different confidence intervals are presented in table 3.2 and 3.3.

**Figure 2.** Uncertainty in parameters of the Generalized Logistic distribution fitted to patient's offset times

6. Conclusions

Service industries, to include health care, have shown value is co-produced, perishable and time-sensitive.

In the health care clinic setting, surveys reveal the greater the amount of time a patient spends with a provider relative to total clinic time, the greater the potential degree of value co-production. To apply the science of health care quality to primary, practical application, one of our research groups' targets is the consequences of queues and crowding in the clinic setting and, concomitantly, in developing tools that may be utilized to create appointment schedules and allocate staff in order to provide a maximum number of patients the highest quality service.

Several interesting findings arose from our analysis of patient waiting times in the clinic venue. First, patient offset times in scheduled visits do not follow a normal distribution; indeed the distribution is not even symmetric. Second, three distributions, Generalized Logistic, Johnson SU and Log-Logistic, nicely fit the observed data points. Third, Generalized Logistic showed a high degree of stability in over-the-sample stability analysis. Due to these outcomes,

we recommend anyone engaged with modeling activities involving assumptions about patient arrival patterns be cautious when choosing the distribution of random variables.

To affect practical change, an important remaining task is to examine the real-life consequences of applying an ill-fitting distribution[1]. What is the implication, for example, in using a slightly incorrect patient offset time distribution in developing patient scheduling schemas? We predict failure to properly characterize patient's offset times is likely to result in unnecessary congestion and wait times as patient queue's emerge over the day[39, 40]. What are the broader implications to the quality of health care in using correct or ill-fitting distribution assumptions in scheduling patients? Thus, ongoing work includes optimal scheduling policies for patients and staff, improvement of medical practices to insure better levels of treatment, and the development of clinic policies in cases of unusual levels of patient demand. Provided that adequate and reliable data are obtained, one of the most obvious ways to study these problems is through the use of computer simulation techniques.

REFERENCES

- [1] Alexopoulos C., Goldsman D., Fontanesi J., Kopald D., Wilson J. Modeling patient arrivals in community clinics. The International Journal of Management 2008; 36:33-43.
- [2] Hosking JRM. The theory of probability weighted moments. IBM Watson Research Center, New York; 1986.
- [3] Rust G, Ye J, Baltrus P, Daniels E, Adesunloye B, Fryer GE. Practical barriers to timely primary care access: impact on adult use of emergency department service Arch Intern Med. 2008 Aug 11;168(15):1705-10.
- [4] Christl HL. Some methods of operations research applied to patient scheduling problems. Medical Progress Through Technology 1973; 2: 19-27.
- [5] Stiglic G, Kokol P. Intelligent patient and nurse scheduling in ambulatory health care centers. Conf Proc IEEE Eng Med Biol Soc. 2005;5:5475-8.
- [6] Kalai E, Kamien MI, Rubinovitch M. Optimal service speeds in a competitive environment. Management Science 1992; 38: 1154-1163.
- [7] Kapustiak J, Ling H. Evaluation of patient waiting times at an academic ophthalmology clinic. Journal of Medical Practice Management 2000; 15: 228-33.
- [8] Fontanesi JM, DeGuire M, Chiang J, Holcomb K, Sawyer MH. Application of workflow analysis tools in outpatient

- primary care settings. Joint Commission Journal on Quality Improvement 2000; 26:654–60.
- [9] Waghorn A, McKee M. Surgical outpatient clinics: Are we allowing enough time? International Journal for Quality in Health Care 1999; 11:215–9.
 - [10] Dexter F. Design of appointment systems for preanesthesia evaluation clinics to minimize patient waiting times: A review of computer simulation and patient survey studies. Anesthesia & Analgesia 1999; 89:925–31.
 - [11] Saunders CE, Makens PK, Leblanc LJ. Modeling emergency department operations using advanced computer simulation systems. Annals of Emergency Medicine 1989; 18: 134–40.
 - [12] Fontanesi J, De Guire M, Holcomb K, Kopald D, Sawyer MH. Can the doctor still see me: What happens when patients arrive late? Journal of Medical Practice Management 2003; 18(5): 239–43.
 - [13] Reeves C. How many staff members do you need? Family Practice Management 2002; 9(8):45–9.
 - [14] Kyriacou DN, Ricketts V, Dyne PL, McCollough MD, Talan DA. A 5-year time study analysis of emergency department patient care efficiency. Annals of Emergency Medicine 1999; 34:326–335.
 - [15] Joint Commission Perspectives on Patient Safety, February 2005, Volume 5, Issue 2 Copyright 2005 Joint Commission on Accreditation of Healthcare Organizations
 - [16] Callahan NM, Redmon WK. Effects of problem-based scheduling on patient waiting and staff utilization of time in a pediatric clinic. Journal of Applied Behavioral Analysis 1987;20:193–9.
 - [17] Clague JE, Reed PG, Barlow J, Rada R, Clarke M, Edwards RH. Improving outpatient clinic efficiency using computer simulation. International Journal for Health Care Quality Assurance 1997; 10:197–201.
 - [18] Dada M, Babad Y. In: 1991 Joint meeting of The Operations Research Society of America/The Institute of Management Science (ORSA/TIMS). Anaheim, CA: Institute for Operations Research and Management Science (INFORMS); 1991.
 - [19] Hashimoto F, Bell S. Improving outpatient clinic staffing and scheduling with computer simulation. Journal of General Internal Medicine 1996; 11:182–4.
 - [20] Jennings M, Audit of a new appointments system in a hospital outpatient clinic (published erratum appears in BMJ 1991 Feb 23; 302(6774):455) (see comments). British Medical Journal (Clin Res Ed) 1991; 302:148–9.
 - [21] Schwarz G, Estimating the Dimension of a Model. *The Annals of Statistics* 6, No.2 (Mar., 1978), 461–464.
 - [22] Akaike H, A new look at the statistical model identification. *IEEE Transactions on Automatic Control* AC 1974;19: 716–723.
 - [23] Akaike H, Canonical correlation analysis of time series and the use of an information criterion, in *System Identification: Advances and Case Studies*, ed. by Mehra, R. K. and Lainotis, D. G. New York, NY: Academic Press 1976; 52–107.
 - [24] Hannan E. J, and Quinn B. G, The determination of the order of an autoregression. *Journal of the Royal Statistical Society*. 1971, B 41, 190–195.
 - [25] von Neumann J, Distribution of the Ratio of the Mean Square Successive Difference to the Variance, *Annals of Mathematical Statistics*. 1941; 12 367.
 - [26] Lehmann E.L., Casella G. *Theory of Point Estimation*. Springer-Verlag, New York, 1998.
 - [27] Pawitan Y, *In All Likelihood: Statistical Modeling and Inference Using Likelihood*. Oxford University Press, Oxford, 2001.
 - [28] Ryden J, Rychlik I, Probability and Risk Analysis. Springer-Verlag, Berlin, 2006.
 - [29] Lin, Y. (1997). Asymptotics of bootstrapping mean on some smoothed empirical distribution, *Statistics & Decisions*. 1997; 15: 301–306.
 - [30] Stephens, M. A. EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc.* 1974; 69(347) 733.
 - [31] Stephens, M. A, Goodness of fit for the extreme value distribution. *Biometrika*. 1977; 64(3): 583–588.
 - [32] Chandra M, Singpurwalla N. D, and Stephens M. A, Kolmogorov statistics for tests of fit for the Extreme value and Weibull distribution. *J. Am. Stat. Assoc.* 1981; 76(375): 729–731.
 - [33] Johnson N.L, Kotz S, Balakrishnan N, Continuous Univariate Distributions Volume 2. Wiley, New York, 1995.
 - [34] Shin H, Kim T, Kim S., Heo JH, Estimation of asymptotic variances of quantiles for the generalized logistic distribution. *Stoch Environ Res Risk Assess.* 2010; 24: 183–197.
 - [35] Rao AR, Hamed KH, Flood frequency analysis. CRC Press, Florida, 2000.
 - [36] Mishra Sk, Empirical Probability Distribution of Journal Impact Factor and Over-the-Samples Stability in its Estimated Parameters. Dept. of Economics, North-Eastern Hill University.
 - [37] Vose D, Risk analysis, a quantitative guide. Wiley, New York, 2008.
 - [38] Sun L, Müller-Schwarze D, Statistical resampling methods in biology: a case study of beaver dispersal patterns. *American Journal of Mathematical and Management Sciences*. 1996; 16: 463–502.
 - [39] Karian Z. A, Dudewicz E.J, Fitting statistical distributions: the Generalized Lambda Distribution and Generalized Bootstrap methods. CRC Press, New York, 2000.
 - [40] Podgorelec V, Kokol P. Genetic algorithm based system for patient scheduling in highly constrained situations. *Journal of Medical Systems* 1997;21:417–27.
 - [41] Reilly T, Marathe V, Fries B. A delay-scheduling model for patients using a walk-in clinic. *Journal of Medical Systems* 1978; 2:303–13.