NPS-OR-15-007 Rev.

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

DISTRIBUTION OF SOME MAXIMUM NORM SUMMARY SETS OF QUANTILE ESTIMATORS

by

Robert R. Read

July 2015

Approved for public release; distribution is unlimited

Prepared for: Unsponsored

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	3. DATES COVERED. From
31-07-2015	Technical Report	01-01-2009 to 31-07-2015

4. TITLE AND SUBTITLE
Distribution of Some Maximum Norm Summary Sets of Quantile Estimators

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Robert R. Read

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Operations Research Department
Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-OR-15-007 Rev.

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
None

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES
The views expressed in this report are those of the author(s) and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT

This report introduces an expanded class of quantile level sets $\{(j-\alpha)/(n+c),\ j = 1,\cdots, n\}$ to augment the popular ones, i.e., ($\alpha = ½$, $c = 0$) used in q-q probability plots and ($\alpha = 0$, $c = 1$) known as the Pyke alternative, for use in data analysis graphical studies of order statistics and for tests of distribution hypotheses. The expanded class can be useful in small sample studies in which their effects can be the greatest. The corresponding test statistics have the form $T_n = (\max |u_j - (j-\alpha)/(n+c)|,\ j = 1, \cdots, n)$ where the $\{u_j\}$ are the order statistics of a random sample of size n from a Uniform(0, 1) population. A sub family, called tail symmetric is described and shown to have greater efficacy than the other members of the family. The small sample distributions of these statistics are developed using Markov Chain methodology.

The computational aspects are illustrated, with n = 5 using a selected set of featured statistics. Some computational idiosyncrasies are attended to and some behavioral properties are illustrated graphically. A number of side issues are discussed.

15. SUBJECT TERMS. Probability distributions, quantile plots, goodness-of-fit testing, Markov Chains

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Robert R. Read
Unclassified	Unclassified	Unclassified	UU	80	**19b. TELEPHONE NUMBER** ***(include area code)*** 831-656-2382

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ronald A. Route
President

Douglas Hensler
Provost

The report entitled "Distribution of Some Maximum Norm Summary Sets of Quantile Estimators" was unsponsored and unfunded.

Further distribution of all or part of this report is authorized.

This report was prepared by:

Robert R. Read
Professor Emeritus of
Operations Research

Reviewed by:

Johannes O. Royset
Associate Chairman for Research
Department of Operations Research

Released by:

Patricia A. Jacobs
Chair
Department of Operations Research

Jeffrey D. Paduan
Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This report introduces an expanded class of quantile level sets $\{(j-\alpha)/(n+c), j = 1, \cdots, n\}$ to augment the popular ones, i.e., ($\alpha = ½$, $c = 0$) used in q-q probability plots and ($\alpha = 0$, $c = 1$) known as the Pyke alternative, for use in data analysis graphical studies of order statistics and for tests of distribution hypotheses. The expanded class can be useful in small sample studies in which their effects can be the greatest. The corresponding test statistics have the form $T_n = (\max |u_j - (j-\alpha)/(n+c)|, j = 1, \cdots, n)$ where the $\{u_j\}$ are the order statistics of a random sample of size n from a Uniform (0, 1) population. A sub family, called tail symmetric, is described and shown to have greater efficacy than the other members of the family. The small sample distributions of these statistics are developed using Markov Chain methodology.

The computational aspects are illustrated, with $n = 5$ using a selected set of featured statistics. Some computational idiosyncrasies are attended to and some behavioral properties are illustrated graphically. A number of side issues are discussed.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

This paper treats issues relating to the development and comparison of some goodness-of-fit statistics of the Kolmogorov type. The main idea is to define a class of quantile sets that relate to the order statistics of a sample of size n from a continuous distribution. These sets are linear modifications of the natural empirical quantiles used in the empirical cumulative distribution function (edf). Then the suggested use is made with the supremum of the magnitude of the deviations as statistics for several purposes, e.g., the testing hypotheses concerning the population distribution of the variables measured in the experiment, and the effective choice of a quantile set to be related to the estimated quantiles appearing in the popular q-q probability plots. The latter has the broader use in the screening of models if the candidate's inverse cdf's (distribution functions) are related to linear functions of the variate values.

Of particular importance are the small sample problems, as better methods are needed in these cases especially when the cost per observation is high. Methodology is developed for the computation of the small sample distributions.

A comparative study of the properties of the statistics can serve to understand their behavior and to screen them for various properties. Explicit attention is given to those that are used in practice or have been proposed for use in the literature.

The computational methods extend those presented in the Durbin (1968) paper. The methods presented there are conceptually simple but laborious and the Durbin paper is sharply curtailed. It is quite terse and many details are presumed or poorly stated and located within it (likely because of the typesetting costs of the day and of the pressure to reduce the length of the articles). Accordingly some of the present work expands upon the presentation of Durbin in order to facilitate its use and expansion by programmers, numerical analysts, students and other workers. It contains many supporting details, some generalizations, and provokes thought for continued work in a number of interesting directions.

In Section II appears the Markov chain methodology for computing the probability that the edf is encased within a corridor defined by two parallel straight lines, following the lead of Durbin and others. Durbin used the method to calculate the initial conditions for his extension of Massey's difference equation for finding the distribution of the Kolmogorov-Smirnov (KS) one sample statistic, D_n, and the Pyke alternative, and indicated extensions to other statistics. The

Durbin use of the Markov chain material is to provide the initial conditions for the difference equations.

Our present use is more direct for the distribution problem and requires an expanded presentation of the parallel lines development. By proper choice of the straight lines one can generate the distributions of statistics of the structure that are proposed.

Introduced in Section III is a family of maximum norm statistics and attention is drawn to their properties and to how their distributions can be obtained. The parameter set for the family includes those of popular statistics that are already in use. There is focus on an attractive subfamily, called tail symmetric, which is simpler to work with and has greater efficacy.

The relationship with parallel line corridors is not always complete in that there can be some distribution values for a particular statistic that cannot be obtained in this way. A discussion of this issue begins and methods for compensation of this and other issues are discussed. A set of five featured statistics are presented, members of the tail symmetric subfamily, and provide interesting examples.

Section III also contains an introduction to some basic properties in the setting of very small sample sizes; n = 1, 2. The distributions of our five featured statistics are developed and tabled. A comparison of their properties serves to set a tone for more general properties.

Section IV treats the case of general sample size n where, character, properties, and computational issues are discussed. Much of this is illustrated by means of actual computation using Maple 8. This software has symbolic computation capability. If our parameter c is a rational number, then the distributions can be presented as piecewise polynomials with rational coefficients, and hence, exact - no round off. The number of pieces grows with the sample size n.

The examples all use the sample size of n = 5. Plots of the cdf's of our featured statistics are included, as well as several plots of the Kolmogorov asymptotic distributions. When compared one is tempted to develop some approximate p-value methods connecting the two.

In Section V the flow of the report is disturbed with a diversion to Durbin's difference equation technique. It begins with a lengthy derivation that was omitted from the original Durbin paper, then his method is presented. It is limited in that non integral values of the parameter c are not treated. The material in Section IV provides initial equation methods for the non integral c case, provided that the statistic is one of our tail symmetric ones. The more general case is

similar, but more complicated. An example of his method for D_n and for the Pyke estimator is contained in Appendix C.

Section VI is entitled “Research Notes”. It contains some satellite material of interest to the analyst and a description of some ideas that have yet to be pursued. There is also some commentary about some issues that may be of interest to practitioners.

Section VII contains a summary. Some appendices are attached containing: the printout of the distributions of the five featured statistics for n = 5 (Appendix A); an example of the geometric method for n = 2 but not in the simpler tail symmetric case (Appendix B); a worksheet methodology for the difference equation approach (Appendix C).

THIS PAGE INTENTIONALLY LEFT BLANK

II. THE MARKOV CHAIN DEVELOPMENT

The foundation of the development centers on the probability that the edf is contained within two parallel straight lines in its sample space. This probability has form in terms of a Markov chain representation and this will be developed first.

This representation has two separate uses. The Durbin paper (1968) sharpens the Massey (1950) difference equation approach to the issue of computing the distribution function of the one-sample KS statistic and extends it to the Pyke (1959) modification of the KS statistic. The solution of the difference equations is accomplished through a generating function approach. There are initial conditions present that can be managed with the Markov representation, which is Durbin's intent for this representation.

The second and present work utilizes the Markov representation directly. In addition to extending the boundary condition methodology mentioned, it is adapted to the direct issue. This become practical because modern software is available that can perform the needed symbolic calculations. The immediate interest is for small sample sizes. Let us explain the foundation.

The data are the order statistics $x_1 \leq x_2 \leq \cdots \leq x_n$ of a random sample from a continuous distribution having distribution function F. The values x_0 and x_{n+1} are known values, possibly infinite in magnitude, and mark the endpoints of the positive sample space.

The empirical distribution function is defined

(2.1) $F_n(x) = j/n$, for $x_j \leq x < x_{j+1}$ and $j = 0, \cdots, n-1$

$= 0$ for $x < x_1 = 1$ for $x \geq x_n$

The edf is at the heart of the classical Kolmogorov one-sample statistic, D_n, which can be expressed in a number of useful ways. The following three should be kept in mind.

$$D_n = \sup\{|F_n(x) - F(x)|, -\infty < x < \infty\}$$

(2.2) $$= \sup\{|F(x_j) - j/n|, |F(x_j) - (j-1)/n|, \text{all } j = 1, \cdots, n\}$$

$$= \max\{0, max_{j=1,\cdots,n}(j/n - F(x_j)), max_{j=1,\cdots,n}(F(x_j) - (j-1)/n)\}$$

Setting.

Because of the properties of the probability integral transform, there is no loss of generality in assuming that F is the Uniform [0, 1] distribution and, for purposes of sharpening the clarity, use $u_1, u_2, \cdots, u_n$ to denote the associated order statistics; $u_0 = 0$ and $u_{n+1} = 1$.

Let S denote the sample path of $F_n(x)$ as x moves from zero to one. We consider the probability $p_n(a, b, c)$ that S lies entirely in the region R between the lines

(2.3) $y = a/n + mx$ and $y = – b/n + mx$; $m = 1+c/n$

or equivalently

(2.4) $ny = a + (n+c)x$ and $ny = –b + (n+c)x$ with $(a, b, a + c, b – c$ all $> 0)$.

Clearly the slope m is 1+c/n. The constraints provide broad structure and stability.

The Massey paper introduces a difference equation quantity which Durbin generalizes to

(2.5) $q_n(a, b, c) = p_n(a, b ,c)(n+c)^n /n!$

It will be convenient to work with q_n and indeed speak of it as if it were a probability, within our present work. Of course Equation (2.5) is used when converting to true probabilities.

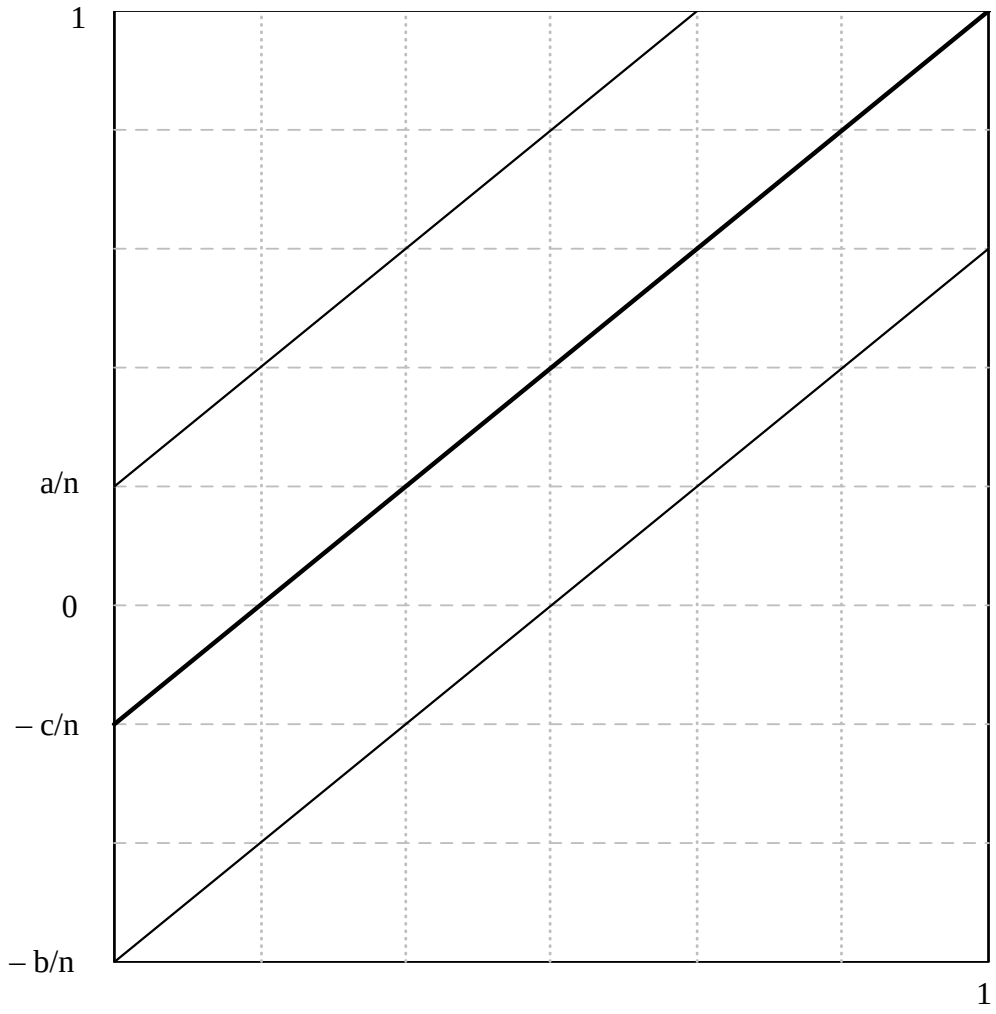
Figure 2.1 provides an image of the sample space included in a more extensive setting. Use is made of the relationship between the Poisson process and the order statistics. The horizontal grid work lines are spaced in increments of $\frac{1}{n}$ (called levels); and the jumps in the Poisson process should be viewed as having this size.

The vertical grid marks are determined in the following way. A base line of slope m is constructed through the point (1, 1). It intersects the j^{th} horizontal level at distance X_j from the y-axis and serves to define a collection of vertical cylinder sets as marked on the graph. Clearly $X_n = 1$ and $X_{–[c]}$ is the last value ≥ 0 in the sequence; the notation [c] refers to the greatest integer in the magnitude of c.

Two parallel lines, containing the baseline as a guide, form a corridor that may contain a sample edf. Notice the graph is presented as having all cylinder sets possessing the same width. Such will happen only if the value of c is a nonnegative integer. This will be a convenience in our early applications, but the immediate discussion is general and does not presume this. The development will proceed without utilizing this feature. In the general case the initial cylinder set can be narrower than the others.

Figure 2.1 Sample Space of the edf.

Levels - Cylinders - Corridors



The Poisson process has rate n + c for the interval [0, 1]. Each of the homogeneous subintervals has rate one. The initial interval, of rate c – [c] will be managed separately. It requires special treatment and will play an important role in the initial condition issue.

Let R be the corridor bounded by the parallel lines as indicated by the intercept parameters a and b; let S be the sample path of F_n and let S´ be the sample path of the Poisson process. The probability of the sample path S remains entirely within R as x moves from 0 to 1 is the same as the conditional probability that S´ remains inside R given that it passes through

(1, 1). Let us therefore consider the latter probability. For convenience, let us assume that b – c and a + c are not integers. Results for the integer cases will follow as limiting values.

The homogeneous cylinder sets intersect R in a way that forms a collection of congruent trapezoids, each of which is to be converted into a (pseudo, see Equations (2.5) and (2.12), the Markov matrix H). There are

(2.6) $p = [b - c] + [a + c] + 1$

states in this process; each representing the occupancy of the Poisson process in a level within R and in the cylinder set. Such a view introduces a localized coordinate system for H. The transitions from state j to state k are the number of Poisson events that occur within the confines of the cylinder.

Equation (2.6) may be viewed as follows. Let r_a, r_b, r_c be the residues of a, b, c, respectively. The number of full intervals from –c to a is $[a + c] = [a] + [c] + [r_a + r_c]$ and the last term can be either 0 or 1. The number of full intervals from –b to – c is $[b - c] + [r_b - r_c]$ and the last term is 0. The number of levels p containing these intervals must be one more than the total number of full intervals.

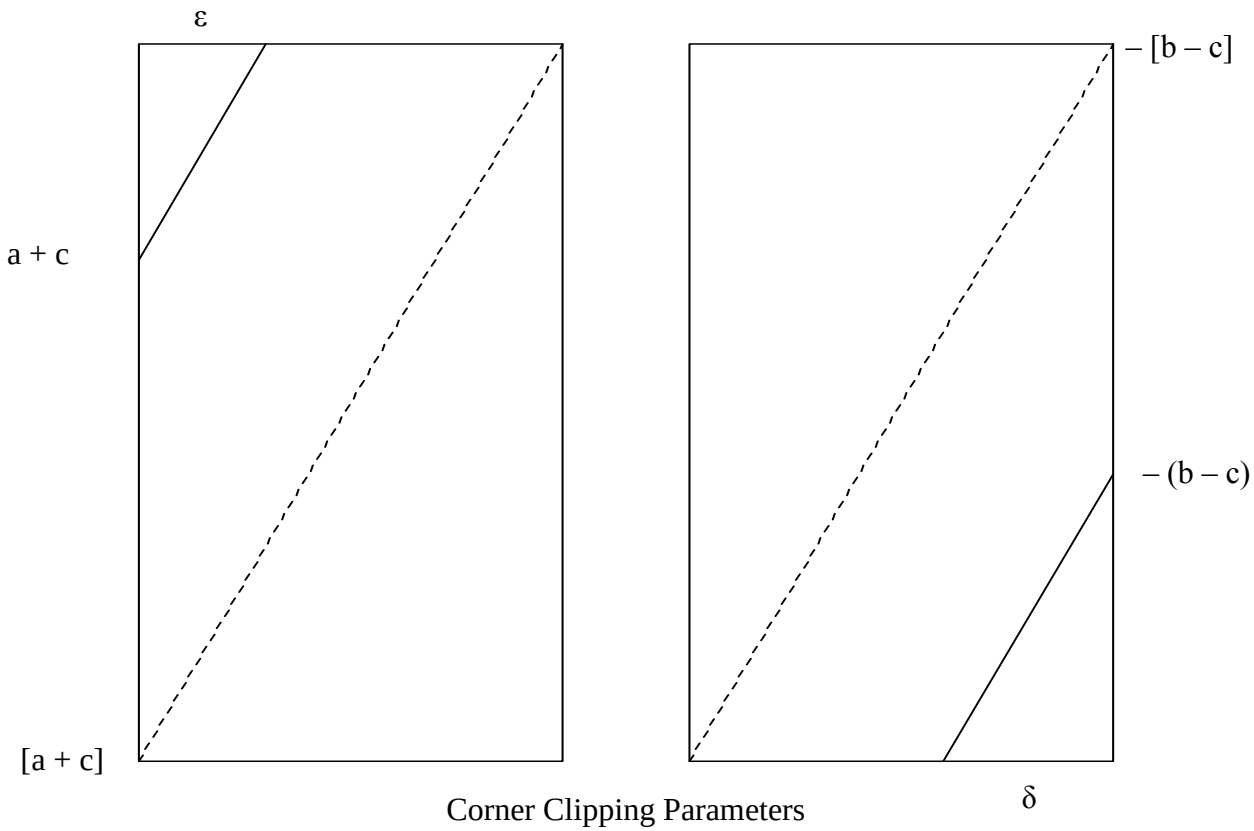
The possible points that the Poisson process can cross the line $x = X_j$ and remain inside R have y coordinates $[j - b + c+1]/n$, $j - b + c+2]/n$, ⋯, $[j + a+c]/n$. Denote these points by $A_{j, k}$, ⋯, $A_{j, p}$. Given that S´ passes through $A_{j, i}$ the probability that it passes through passes through $A_{j+1, k}$ is the probability of exactly $k - i + 1$ observations in the interval (X_j, X_{j+1}), that is, $e^{-1}/(k - i + 1)!$ $(i, k = 2, 3, \cdots, p–1; k \geq i + 1)$.

For k = 1, p we have to allow for the fact that b–c and a+c may not be integers. This leads us to the extremes of the trapezoid, as indicated in Figure 2.2. The right member of the figure shows the lower boundary of the corridor as it cuts the first cell and the level indexed by k = 1. The left member shows the upper boundary of the corridor as it cuts the most upper cell and the level indexed by k = p. The lines on the cell diagonals remind us that the cell, which is 1/n by 1/(n+c), shares the slope m. The proportions of the cell sides that are clipped away by the residue

(2.7) $1 - \delta = b - c - [b - c]$

(2.8) $1 - \varepsilon = a + c - [a + c]$

Figure 2.2 Extreme Cells for a Cylinder Set.



They are determined by the vertical measurements, but also apply to the horizontal distances because of the common slope.

Remarks:

i. A helpful fact connecting p to the line intercepts. As p is defined in (2.6), it follows that $p = b - c + a + c + \delta + \varepsilon - 1 = a + b + \delta + \varepsilon - 1$ and that

(2.9) $[a + b] = p$ if $\delta + \varepsilon \leq 1$ and $[a + b] = p - 1$ if $\delta + \varepsilon > 1$

In moving from $A_{j,1}$ to $A_{j,k}$ ($k = 1, \cdots, p-1$) the path of S′ will remain in R only if at least one observation occurs in the interval $(X_j, X_j + (1 - \delta)/(n+c))$. The probability of k observations in (X_j, X_j), at least one of which is in $(X_j, X_j + (1 - \delta)/(n+c))$, is $e^{-1}(1-\delta^k)/k!$. This is therefore the probability of S moving from $A_{j,1}$ to $A_{j+1,k}$ and remaining inside R. Similarly the probability of moving from $A_{j,i}$ to $A_{j+1,p}$ inside R is $e^{-1}(1-e^{p-i+1})/(p - i+1)!$,$(i = 2,\cdots,p)$. Similarly the probability from $A_{j,i}$ to $A_{j+1,p}$ inside R is $e^{-1}(1- \varepsilon^{p-i+1})/(p-i+1)!$, $i = 2,\cdots, p$. The probability of moving from $A_{j,1}$ to $A_{j+1,p}$ inside R is a bit more complicated. It involves moving from the first clipped cell to the last clipped cell. Refer to Figure 2.2. The result is

(2.10) $e^{-1}(1-\delta^p-\varepsilon^p+h)/p!$,

where h = 0 if $\delta + \varepsilon \leq 1$. That is, the possibility that all p events take place in the δ portion of the initial cell must be excluded and the possibility that all p events take place in the ε portion of the final cell must also be excluded. In other words there must be at least one event in the (1 – δ) portion of the initial level and at least one event arriving in the (1–ε) portion of the final level. Stated again, in one case if there are none in 1 – δ strip, then the smallest event in the cylinder cannot escape the first level. Further if all events are in 1 – ε strip then the largest one cannot remain below the upper boundary of R.

The second case specifies $\delta + \varepsilon > 1$ and then $h = (\delta + \varepsilon - 1)^p$ is added to the existing amount. In this contingency the cylinder strips marked by δ from below and ε from above are no longer disjoint. The adjustment for this case is a bit entangled. The problem is to ensure that the smallest Poisson count in the first cell is in the corridor and hence not in the interval marked with length δ, and that the largest Poisson count (in the top cell) is not in the interval marked as length ε. For the first case let A be the event that not all of the counts in the cell are in the portion marked δ; $P(A) \doteq 1-\delta^p$, ignoring the factor $(e^{-1}/p!)$. Similarly, let B be the event that not all counts are in the interval marked ε; $P(B) \doteq 1-\varepsilon^p$. But A and B are not mutually exclusive events; hence, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. But $A \cap B$ is the event that simultaneously not all of the counts are in the interval marked δ and in the one marked ε. This is the complement of the event that all counts are in the intersection of the marked intervals δ and ε; an event of probability $\doteq 1-(\delta + \varepsilon - 1)^p$. Combining all that has been said produces Equation (2.10).

Let $e^{-(j-c)}v_{j,i}$ be the probability that S′ passes through $A_{j,i}$ while remaining inside R and let $v_j' = [v_{j,1}, \cdots, v_{j,p}]$, j = c′, c′+1, ⋯, n, where c′ =– [c] for short. These are the states of S′ inside R. (The term in the exponent comes from the fact that (j+c)/(n+c) is the y coordinate of $A_{j,i}$.) The transition from v_j to v_{j+1} is then given by the relation

(2.11) $v_{j+1} = Hv_j$, j = c′, c′+1, ⋯, n–1 ,

and the transition matrix H is given by

(2.12) H =

$$\begin{bmatrix} 1-\delta & 1 & 0 & \cdots & \cdots & 0 & 0 \\ \frac{1-\delta^2}{2!} & 1 & 1 & 0 & \cdots & \cdot & 0 \\ \frac{1-\delta^3}{3!} & \frac{1}{2!} & 1 & 1 & 0 & \cdot & \cdot \\ \vdots & \frac{1}{3!} & & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \cdots & & & \\ \cdot & \cdot & & & & & 0 \\ \frac{1-\delta^{p-1}}{(p-1)!} & \frac{1}{(p-2)!} & & \cdots & & 1 & 1 \\ \frac{1-\delta^p-\varepsilon^p+h}{p!} & \frac{1-\varepsilon^{p-1}}{(p-1)!} & & \cdots & & \frac{1-\varepsilon^2}{2!} & 1-\varepsilon \end{bmatrix}$$

(2.13) h = 0 if $\delta + \varepsilon \leq 1$ and $h = (\delta + \varepsilon - 1)^p$ if $\delta + \varepsilon > 1$

The structure of H is taken from the above development. The ingredients of column 1 and row p are developed quite explicitly. The structure of h is in (2.10)ff. The remaining submatrix, of order p–1, is the lower triangular with ones on the main diagonal and on the first sub-diagonal. Each succeeding sub-diagonal contains constants, beginning with 1/ 2!, then 1/3!, and continuing until 1/(p–2)! in the lower left corner.

To obtain the probability of **S´** arriving at (1, 1) we require the element $v_{n,\,i}$ of v_n corresponding to the point (1, 1); this has i = [b–c] +1. Denote the p vector with one in the position[b–c] +1 and zeros elsewhere as ww. The required element is $ww' H^{[n+c]} v_{c'}$. The probability that S´ passes through (1, 1) after remaining in R is therefore

(2.14) $e^{-(n+c)}\ ww' H^{[n+c]} v_{c'}$.

Since the unconditional probability of S´ passing through (1, 1) is $e^{-(n+c)}(n+c)^n/n!$, the conditional probability $p_n(a, b, c)$ that S´ reaches (1, 1) after remaining in R, given that it reaches (1, 1) is

(2.15) $p_n(a, b, c) = ww' H^{[n+c]} v_{c'} tcof$; and $tcof = n!/(n+c)^n$

and from (2.5) we deduce the useful expression

(2.16) $q_n(a, b, c) = ww' H^{[n+c]} v_{c'}$.

If c is a nonnegative integer then $v_{c'}$ is a p vector with a one in position [b]+1 and zero elsewhere. This marks the entrance into the homogeneous Markov Chain. Otherwise, $v_{c'}$ is a boundary vector whose computation will be addressed in Section IV.

THIS PAGE INTENTIONALLY LEFT BLANK

III. STATISTICS AND RELATIONS WITH PARALLEL LINES

Our main development is the properties and distributions of some competitors to the Kolmogorov one-sample statistic, D_n. They have the structure of maximum norms applied to the difference between order statistics and the members of a family of quantile values, specifically linear functions of j/n for j = 1,⋯, n. Consider

(3.1) $T_n(\alpha, c) = \{\max | u_j - \frac{j-\alpha}{n+c} | \text{ for } j = 1,\cdots, n\}$.

The family of interest is $0 \leq c \leq 2$ and $-½ \leq \alpha \leq ½$ and focus will remain within this family, although the methodologies can be extended beyond these limitations.

The distributions of the random variables in the family (3.1) can (mostly) be expressed in terms of the $\{p_n(a, b, c)\}$; exceptions will be addressed later. The following result covers most, but not quite all, of the ground.

Proposition.

The following relationship is formally valid under the existing constraints, see Equation (2.4).

$$\{ -\frac{b-1+\alpha}{n+c} \leq \frac{j-\alpha}{n+c} - u_j \leq \frac{a-\alpha}{n+c}, \text{ for all } j = 1, 2, \cdots, n\}.$$

Proof. The event that the edf is always no higher than the upper line is

$\{F_n(x) \leq a/n + m\, x, \text{ for all } x \text{ in } (0,1)\}$, using m=(n+c)/n for the common slope of the parallel lines, and is the same as the event

$\{j/n \leq a/n + (n+c)x/n, \text{ for all } x \text{ in } [u_j, u_{j+1}) \text{ and for all } j = 0, 1, \cdots, n\}$;

that is, $F_n(x)$ is constant in each interval $[u_j, u_{j+1})$ and equal to j/n. From this, we can obtain

$\{j \leq a + (n+c)u_j, \text{ for all } j = 1, 2, \cdots, n\}$;

subtracting α from each side and adjusting produces

$\{\frac{j-\alpha}{n+c} \leq \frac{a-\alpha}{n+c} + u_j, \text{ for all } j = 1, 2, \cdots, n\}$.

In a similar way, using the lower member of the parallel lines, the event

$\{Fn(x) \geq -b/n + m\, x, \text{ for all } x \text{ in } (0, 1)\}$ is the same as the event

$\{(j-1)/n \geq -b/n + (n+c)x/n, \text{ for all } x \text{ in } [u_{j-1}, u_j), \text{ all } j= 1, \cdots, n+1\}$.

Proceeding in a manner similar to the previous leads to

$\{j - \alpha - 1 \geq -b - \alpha + (n + c)\, u_j, \text{ all } j\}$ or

$\{\frac{j-\alpha}{n+c} \geq \frac{-b+1-\alpha}{n+c} + u_j$, for all j$\}$. Then, combining the two results in a double inequality,

(3.2) $\{ -\frac{b-1+\alpha}{n+c} \leq \frac{j-\alpha}{n+c} - u_j \leq \frac{a-\alpha}{n+c}$, for all j = 1, 2, ⋯, n$\}$ qed.

Now the statistic T_n is the largest of the magnitudes of the center portion of the double inequality, and the variate value in its cdf is the common value of the bound on the right and the negative of the bound on the left, provided that b can be so chosen. It is a function of the parameters c and α. Formally, it appears the distribution functions can be obtained from

(3.3) $\Pr\{T_n < \frac{a-\alpha}{n+c}\} = p_n$ (a , a +1–2α, c) a , a+1–α, c,

but its use needs to be watched; the translation can produce surprising effects. Most certainly the quantile figures $\{q_j\}$ must be chosen so that $0 \leq (1-\alpha)/(n+c)$ and that $(n-\alpha)/(n+c) \leq 1$. The idea of using a translation parameter ($\alpha \neq 0$) is new and presents some further issues. For $0 < \alpha \leq ½$, the value a is required to be sufficiently large so that the upper boundary of the corridor is positive. Any edf in corridors having $a \leq \alpha$ will have probability zero. The entire cdf of T_n can be obtained from the Markov approach (2.13). On the other hand for $-½ \leq \alpha < 0$ there will be corridors that begin below the x–axis. This creates some (small) variate values for the cdf of T_n that cannot be obtained from Equation (2.15) and other methods need be found. The solution to this dilemma is a geometric one and will be presented in Section IV.

The Kolmogorov statistic D_n, Equation (2.2), is not a member of our family. But the calculation of its distribution is within our grasp. Others, e.g., Kemperman (1961) have shown that

(3.4) $\Pr\{D_n < \frac{a}{n}\} = p_n$ (a, a, 0),

which is more easily managed using one of our featured statistics (see below). But first we introduce our property of tail symmetry.

Symmetry. A member of the family always has a form of internal symmetry, based upon the separation of the successive quantiles used. That is, $\frac{j+1-\alpha}{n+c}$ is always an increment of value $\frac{1}{n+c}$ from its predecessor $\frac{j-\alpha}{n+c}$ for all j = 1, ⋯, n – 1. But such does not necessarily hold in the tails. In other words $\frac{1-\alpha}{n+c} - 0$ need not match $1 - \frac{n-\alpha}{n+c}$. But such a tail symmetry is desirable. The condition for it to occur is

(3.5) $\alpha = \frac{1-c}{2}$ or (re-stated) $c = 1 - 2\alpha$ (tail symmetry).

A comparison study will be made of five featured statistics, all of which are tail symmetric. In order of increasing c, these statistics are

$CL_n = T_n(½, 0)$

$TT_n = T_n(1/6, 2/3)$

(3.6) $C_n = T_n(0, 1)$

$FF_n = T_n(-1/8, 5/4)$

$CC_n = T_n(-½, 2)$.

The report focusses upon this set, but upon occasion some brief remarks will be made about the relaxation of the property. The two cases having negative α will be used to illustrate how one can compensate when the parallel lines cannot contain the distribution for certain of its variate values.

The translation value $\alpha = ½$ has a special property. Setting this value into (3.3) and applying the tail symmetry property, i.e., b = a + c, produces the relationship

(3.7) $\Pr\{T_n(½, c) + \frac{½}{n+c} \leq \frac{a}{n+c}\} = p_n(a, a, c)$

and using c = 0 as a special case it follows that $T_n(½, 0) + \frac{½}{n}$ is distributed the same as D_n which, as mentioned above, allows the distribution of D_n to be developed from our system. The relationship was noticed by Gibbons and Chakraborti (1992). In short, it says that the distribution of D_n can be computed from that of CL_n.

At this point it is convenient to look at the properties of our family from the view of small n, i.e., n = 1, 2. Direct methods will be used.

Very small values of n. The behavior of our statistics for the sample sizes of n = 1 and 2 are presented because of the insight that they provide.

n =1. The distribution of any family member is

(3.8) $\Pr\{T_1 < w\} = \Pr\{|u - \frac{1-\alpha}{1+c}| < w\}$,

and for those possessing the tail symmetry property this specializes to

(3.9) $\Pr\{|u - \frac{1-\alpha}{2(1-\alpha)}| < w\} = \Pr\{½ - w < u < ½ + w\}$,

and this is the uniform distribution on (0, ½). Variables from the family that are not tail symmetric cannot have this distribution. The Kolmogorov Statistic D_1 is distributed uniformly on (½, 1).

The family members that are not tail symmetric have piecewise uniform distributions. Example: Use α = 0 and c = ⅔.

$\Pr\{T_1(0, ⅔) \leq w\} = 2w$ for $0 < w < \frac{2}{5}$

$= \frac{2}{5} + w$ for $\frac{2}{5} < w < \frac{3}{5}$

n = 2. These distributions are most easily calculated from geometric considerations. The joint distribution of (u_1, u_2) is uniform over the upper triangular region of the unit square and the constant density is 2. One must deal with

$$\Pr\{\frac{1-\alpha}{2+c} - w < u_1 < \frac{1-\alpha}{2+c} + w \text{ and } \frac{2-\alpha}{2+c} - w < u_2 < \frac{2-\alpha}{2+c} + w,\ u_1 < u_2\},$$

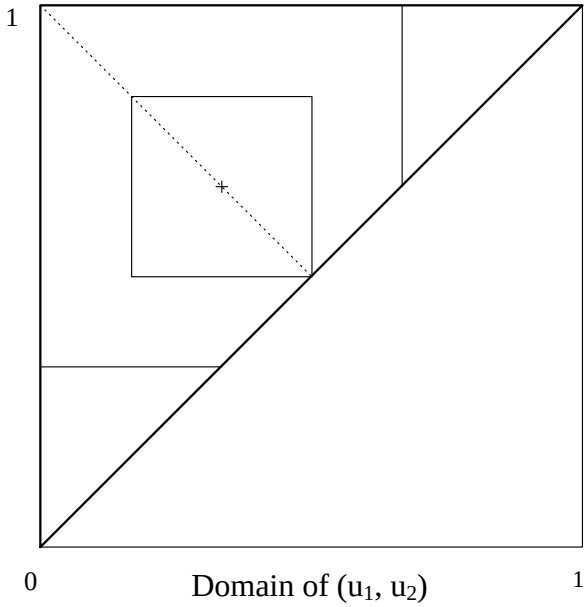
and the w = 0 point in the sample space of (u_1, u_2) is

$$\overline{u_1} = \frac{1-\alpha}{2+c} = \frac{1-\alpha}{3-2\alpha} \text{ (ts) and } \overline{u_2} = \frac{2-\alpha}{2+c} = \frac{2-\alpha}{3-2\alpha} = \text{(ts)},$$

where the (ts) tag on the second members mark specialization for tail symmetry.

Figure 3.1 shows the sample space and the geometric forms encountered for a tail symmetry case. In such cases the sum $\overline{u_1} + \overline{u_2} = 1$; the center point falls on the counter diagonal of the square. As the variate value w departs from 0 and grows, the contours of constant probability are squares. This continues until the tip of the square bumps into the edge of the positive sample; at (½, ½) for the ts case. The largest value for w in the square is $w_0 = ½ - \overline{u_1}$. (More generally this value is $w_0 = \frac{1}{2(n+c)}$). The cdf value is the area of a square times 2 (from the value of the density function), i.e., $8w^2$.

Figure 3.1 Location of Changes for the Contours of Constant Probability.



As w grows beyond w_0 the geometric shape over the positive sample space changes into a different shape and continues this shape until the size of w acquires $\overline{u_1}$. Then the shape changes again and w must grow to $\overline{u_2}$ and complete the process. This is an awkward method to execute directly. It is easier if we allow the original square to grow in size beyond w_0. As it does, we trim away the excess probability so acquired. The following algorithm makes the process explicit.

Step 1. Let $w_0 = \frac{1}{2(2+c)}$. Each side of the original set of squares has length 2w. It follows that

$$\Pr\{T_2 \leq w\} = 8w^2 \text{ for } 0 < w < w_0.$$

Step 2. Consider a square centered at (½, ½) having radius (a convenient terminology representing the distance from center to the square's edge) $w - w_0$. A triangular shaped half of this square is the part to be trimmed away. This continues until w acquires the value of $\overline{u_1}$. So

$$\Pr\{T_2 \leq w\} = 8w^2 - 4(w - w_0)^2 \text{ for } w_0 < w < \overline{u_1}.$$

Step 3. This process can be continued into the third stage. However the original square grows more slowly now because its left and upper sides have been stopped by the edge of the positive sample space. So

$$\Pr\{T_2 \leq w\} = 2(w + \overline{u_1})^2 - 4(w - w_0)^2 \text{ for } \overline{u_1} < w < \overline{u_2}.$$

Note. In the case of CL_2 we have $w_0 = \overline{u_1}$, and there are but two pieces to the distribution. The middle piece of the graph is absorbed.

Table 3.1 contains the cdf of our featured statistics for n = 2. They are all tail symmetric and were calculated using this algorithm.

Beneath the table are the cdf's of D_2 and T_2 (0, ⅔). The distribution of D_2 is deduced from that of CL_2 as shown in (3.7). The distribution of T_2 (0, ⅔) can be obtained by a more complicated version of the above algorithm and details appear in Appendix B. The difficulties stem from the fact that the zero point of a non-tail symmetric variable does not fall on the counter diagonal line of the sample space.

A natural competition between C_2 and CL_2 is present because of their popularity. It is seen from the table that, although the values are close, the former does not uniformly dominate the latter. The two distributions are

the same for $0 \leq x \leq \frac{1}{6}$

(3.10) $P\{CL_2 \leq x\}$ is larger for $0.17 \leq x \leq 0.27$ (approx.),$(\frac{1}{6}, \frac{5+\sqrt{22}}{36})$ exact

$P\{C_2 \leq x\}$ is larger for x beyond ~ 0.27.

For general n the range of C_n is $[0, \frac{n}{n+1}]$ and for CL_n it is $[0, 1- \frac{1}{2n}]$.

Table 3.1. Distributions.

0 $\frac{1}{8}$ $\frac{2}{13}$ $\frac{1}{6}$ $\frac{3}{16}$ $\frac{1}{4}$ $\frac{7}{26}$ $\frac{5}{16}$ $\frac{1}{3}$ $\frac{9}{26}$ $\frac{3}{8}$ $\frac{1}{2}$ $\frac{8}{13}$ $\frac{5}{8}$ $\frac{17}{26}$ $\frac{2}{3}$ $\frac{11}{16}$ $\frac{9}{13}$ $\frac{3}{4}$ 1				
CL_2	$8w^2$	$-2w^2+3w-\frac{1}{8}$	1	
TT_2	$8w^2$	$w^2+\frac{3}{2}w-\frac{9}{64}$	$-2w^2+\frac{11}{4}w+\frac{7}{128}$	1
C_2	$8w^2$	$4w^2+\frac{4}{3}w-\frac{1}{9}$	$-2w^2+\frac{8}{3}w+\frac{1}{9}$	1
FF_2	$8w^2$	$4w^2+\frac{16}{13}w-\frac{16}{169}$	$-2w^2+w+\frac{77}{169}$	1
CC_2	$8w^2$	$4w^2+w-\frac{1}{16}$	$-2w^2+\frac{5}{2}w+\frac{7}{32}$	1
D_2	0	$8w^2-4w+\frac{1}{2}$	$-2w^2+4w-1$	

$T_2(0, ⅔) = 8w^2$ $0 < w < 3/16$

$4w^2 + (3/2)w + 9/64$ $3/16 < w < ¼$

$(11/8)w - 9/64$ $¼ < w < 3/8$

$-2w^2 + (11/4)w + 21/64$ $3/8 < w < 5/8$

$-w^2 + (3/2)w + 7/16$ $5/8 < w < ¾$

This is an asymmetric form of TT_2. The shift expands the range but concentrates the probability to the left. The scale above identifies the partition marks for the distribution.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. GENERAL N AND FEATURED STATISTICS

This section discusses the properties and methods for general values of n. The cumulative distributions of our family of statistics are piecewise polynomials of order n or n–1. They can be calculated from (2.15) using software that has symbolic processing capabilities. Maple 8 is used in the present project. There are idiosyncrasies in the method, the nature of the statistics and in the software as well. The issues are nicely illustrated using the selection of featured statistics. The techniques will be covered by outlining how each is managed.

First, some generalities.

For all members of our family $\{T_n(\alpha, c)\}$ the small values of the distribution have common structure and can be managed geometrically, Kendall (1961) and Kendall & Moran (1963). Consider the set of event intersections

(4.1) $\{-w < \frac{j-\alpha}{n+c} < u_j < w + \frac{j-\alpha}{n+c}\}$

and $\{-w + \frac{j+1-\alpha}{n+c} < u_{j+1} < w + \frac{j+1-\alpha}{n+c}\}$; j =1,⋯, n–1.

For $0 \le w \le w_0 = \frac{1}{2(n+c)}$ they form a collection of disjoint events and hence a hypercube within the polygonal figure $\{0 < u_1 < u_2 < \cdots < u_n < 1\}$; the density is the constant n!. Its content is the n^{th} power of the length of a side. The density function is n! within the figure. It follows that

(4.2) $\Pr\{T_n \le w\} = n!(2w)^n$, for $0 < w < w_0$.

The value w_0 is the largest value that w can achieve in order to maintain the non-overlapping property of the successive intervals. This method is always available, and it will be a necessary alternative when α is negative.

The computation of probabilities using (2.16) requires the use of the transition matrix H. To this end, we must chose an order p for H by restricting attention to an interval of values of the intercept, a. The general formula is $p = [b - c]+[a + c\} = 1$. Restricting attention to the tail symmetric case means that b = a + c and hence

(4.3) $p = [a] + [a+ c] + 1$

The variate value of the distribution functions will always be in terms of $w = (a–\alpha)/(n+c)$. The various pieces of the cdf's are expressed in intervals of w which in

turn are intervals of the variable a associated with fixed values of the auxiliary parameter p; the order of H. A collection of values for p are required in order to obtain the entire distribution. Many of the intervals will be split because of the requirement of (2.14), repeated here:

(4.4) $h = 0$ if $\delta + \varepsilon < 1$; $h = (\delta + \varepsilon - 1)^p$ if $\delta + \varepsilon > 1$.

The matrix H must be raised to the power nn = n + [c], and such can be a major computational task. The term $H_{p,\,1}$ of H contributes some awkwardness because of the input quantity h. However this annoyance occurs only when using the smaller values of p. When p grows to possess a value circa nn, then $H_{p,1}$ does not enter into the elements of H^{nn} that are used in the probability calculation; splitting the interval becomes unnecessary.

An understanding of this phenomenon can be useful to the analyst. Referring to Equation (2.12), a matrix of type H has the schematic structure as the one on the left below.

(4.5)
$$H = \begin{matrix} x & x & 0 & 0 & 0 \\ x & x & x & 0 & 0 \\ x & x & x & x & 0 \\ x & x & x & x & x \\ \bullet & x & x & x & x \end{matrix} \qquad H^{nn} = \begin{matrix} \bullet & y & y & y & y \\ \bullet & \bullet & y & y & y \\ \bullet & \bullet & \bullet & y & y \\ \bullet & \bullet & \bullet & \bullet & y \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{matrix}$$

The matrix H, of order p, has an upper triangle containing entirely zeros, beginning with the (1, 3) element. In Equation (2.16) it is to be raised to the power nn = n + [c]. In our schematic the symbol • is used to mark any element that is a function of h, the contentious part of $H_{p,1}$. When p = nn, the H^{nn} power matrix has the character as shown in the diagram on the right: the awkward parameter h affects only members in the lower triangle. When p = nn +1 the diagonal members of H^{nn} change their symbol to y, a value not contaminated by h; no others change in the schematic. This effect continues on to the next lower diagonal when p = nn + 2, etc. On the other hand if p = nn – 1 the effect moves in the opposite direction; the first super-diagonal converts to the symbol • in place of a y, etc.

It is useful to identify conditions for which these members do not contribute to the computation at hand. Although nn is fixed, the value of p varies small to large for computing probabilities from the left; p grows until the full piecewise distribution is acquired. The location of the point at which the simplification begins is situation specific.

When c is an integer, the calculation requires the extraction of a single element of H^{nn}, specifically at the entrance column and the exit row. The given interval provides a value of p. The required p-vectors are $u_c = 1$ in the position [a+c]+1, zero elsewhere, and ww = 1 at position [a]+1, zero elsewhere. The location of the extracted value will be on the diagonal of H^{nn} if p = nn, and in the free region marked with y's if p is larger. Then the $H_{p,\,1}$ element becomes and remains uninvolved; it is of no further consideration in subsequent calculations when computing the distribution from left to right.

The cases involving integral c are the easiest to treat. (Durbin treats only these cases.) It follows from tail symmetry that $\delta = \varepsilon$. That is $1-\delta$ and $1-\varepsilon$ (of (2.7) and (2.8)) both reduce to = a – [a], a noticeable simplification. Further, then $\delta + \varepsilon = 2\,\delta \leq 1$ implies that h = 0; this in turn happens when $\delta = 1+[a] - a \leq ½$ or $a \geq [a] + ½$; and h has a positive value when a is smaller than this midpoint.

When c is not integral, the analysis is more complicated and will be addressed subsequently.

Specifics of the Featured Statistics

It is wise to have a visual image of the edf sample space (Figure 2.1) when developing the boundary conditions, especially when c is not integral. Such can be provided by the reader. The sample size value of n = 5 is used throughout. The software has its idiosyncrasies and the author developed his own peculiar way of dealing with them. Only a few intervals are selected in order to illustrate the techniques. When computing for successive intervals, the programmer is cautioned to use the Maple "unassign" function in order to avoid values from one interval to be passed on and contaminate a succeeding interval. Appendix A contains the full distribution for each statistic. This is done in the Maple 8 format. It is especially useful within the Maple graphics systems.

Our featured statistics calculation examples start with the three that have integral c. The other two are more complicated and follow afterwards.

A. FEATURE STATISTICS THAT HAVE INTEGRAL C

The easiest one is C_n also known as the Pyke alternative. It possesses a stronger form of symmetry because

(4.6) $\overline{u_j} = \frac{j}{n+1}$ for all j = 1, ⋯, n.

All separations, internal and external gaps, are the same.

1. $C_5 = T_5(0, 1)$; w = a/6 and a = 6w; b = a+1; again δ = ε.

p = [a]+[a+1]+1= 2[a] + 2. This case is the simplest; α = 0. Starting with the geometric produces $P\{w\} = 5!2^5w^5$ for $0 < w < \frac{1}{12}$ .This value of w_0 corresponds to a_0 = ½. Since p=2 for the entire interval 0 < a < 1, it is seen that the first half of the interval is managed using the geometric expression. It remains to work on the second half, which has h = 0.

The positive value of h appears when a is on the smaller side of the interval midpoint, and h = 0 on the larger side.

The factor tcof = $5!/6^5$; the power nn = 6. The Maple code for this H is

H_2 :=<< 1–δ,$\frac{(1-2\delta^2}{2}$ >|<1,1– ε >>; ww=<1|0>; uc = <0,1>; A:=ww. $H_2{}^6$;

AA:=A.uc; δ:=1– a; AAA:=expand(AA); a :=6w; AAAA:=tcof*expand(AAA). The result is AAAA= (i.e., P(w)) is

$-950w^5 + 320w^4 + \frac{200}{3}w^3 - \frac{20}{3}w^2 + \frac{5}{36}w$; and the interval endpoint values are

$P(\frac{1}{12}) = \frac{5}{324}$ and $P(\frac{1}{6}) = \frac{175}{648}$, for the interval ½ < a < 1.

Looking ahead, when [a] =2 then p = 6 and nn=6 and the entrance/exit point (ww, ucc) is (3,4), i.e., beyond the diagonal, and it is not necessary to split the interval 1 < a < 2. The issue of differing formula for $H_{4,1}$ is gone, and for all larger values of p.

2. $CL_5 = T_5(½, 0)$; w = (a – ½)/5 and a = 5w + ½; b = a+1, p = 2[a]+1.

Notice that w < 0 when a < ½. Such values must have probability zero. A graph would show but a single edf in the corridor whose upper boundary is a = ½, and such is an event of probability zero. The factor tcof = $5!/5^5$; the power nn = 5.

It is always easier to begin with the geometric calculation

$$P\{w\} = 5!\cdot 2^5w^5 \text{ for } 0 < w < \frac{1}{10}.$$

This value corresponds to a_0 = 1. So next let us consider the interval

1 < a < 2; p = 3; the Maple expression for H is

$$H_3 := <<1- \delta,\frac{1-\delta^2}{2!}, \frac{1}{3!}(1- 2\delta^3+ h)>|<1,1, \frac{1-\delta^2}{2!}>|<0,1,1- \delta>>$$

ww := <0|1|0>. That is ww =1 at position [a]+1 =2; this is the position of exit from the chain. Similarly, u_c = <0,1,0> as [a+c] +1=2 is the process starting position. Since c is

integral the pertinent value of $q_5(a, a, 0)$ is the (2,2) member of H^4. The code I prefer looks as follows.

A:= ww.H3^5; AA:=A.uc; δ:= 2 – a; a:=5w + ½; AAA:=expand(AA);

AAAA:= tcof*expand(AAA). Result AAAA = (i.e., P(w)) is

$-288w^4 + \frac{624}{5}w^3 - \frac{96}{25}w^2 - \frac{36}{125}w + \frac{6}{125}$. $P(\frac{1}{10}) = \frac{24}{625}$ and $p(\frac{1}{5}) = \frac{42}{125}$.

The first value is at the left endpoint and matches the right endpoint of the previous interval. The second value is at the right endpoint and serves to anticipate the correct starting value the next interval, i.e., the midpoint a = ½.

All this must be performed twice. First for h:= $(2\delta-1)^3$ and a second time for h=0. The first case will produce for AAAA a polynomial in w valid for

1/10 < w < 1/5. The second AAAA quantity serves for 1/5 < w < 3/10. It is

$$160w^5 - 160w^4 + \frac{25}{4}w^3 + \frac{616}{25}w^2 - \frac{332}{125}w + \frac{6}{125}.$$

The next interval, 2 < a < 3, must also be split at its midpoint and the two sides be treated separately. This time p=5. When a 5 by 5 matrix is raised to the power 5, the $H_{5,1}$ element appears in all of the main diagonal elements (and in all of the lower subdiagonal elements). The solution extraction point from this matrix is the (3, 3) element, and is affected by the value used for h. But the simplification benefit begins in the next interval 3 < a < 4. This time p = 7 and of course nn remains at 5. The extraction point is (4,4), a diagonal point. The members of H_7^5 that are not visited by $H_{7,1}$ begin at the position. The Maple printout of the cdf is in Appendix A.

D_5. The edf. Because of the relationship (2.7) the distribution of D_5 is readily available from that above. One merely shifts all the interval values to the right using the quantity 1/2n = 1/10, then in the polynomials simply replace each w with $(w - \frac{1}{2n})$.

3. $CC_5 = T_5(-½, 2)$; $w = \frac{a}{4} + \frac{1}{8}$ and a = 7w –½; b = a+2;

p = [a]+[a+2]+1 = 2[a]+3. Again, h > 0 when a is on the smaller side of an interval midpoint, and h=0 on the larger side. The factor tcof is 5!/7^5; the power nn = 7.

The variate value w = 0 must initiate the probability mass, but it corresponds to a value of a = –½. A corridor having a negatively valued intercept for the line that marks its upper boundary cannot contain an edf, and cannot be used to compute probabilities in this range. Accordingly, the process must be started with

$P\{w\} = 5!2^5 w^5$ for $0 < w < 1/14$.

This value of w_0 corresponds to $a_0 = 0$.

The anticipated value for the front of the next interval is $P\{\frac{1}{14}\}$. So let us continue to the interval $0 < a < 1$. Again, $\delta = \varepsilon = 1 - a$; the midpoint is $a = ½$; $p = 3$; The factor $tcof = 5!/7^5$; the power $nn = 7$. The Maple code for H becomes

H_3 := <<1– δ,$\frac{1-\delta^2}{2!}$, $\frac{1}{3!}(1 - 2\delta^3 + h)$>|<1,1, $\frac{1-\delta^2}{2!}$>|<0,1,1– δ>>; ww:=<1|0|0>;

uc := <0,0,1>; h = $\frac{1}{3!}(2\delta - 1)^3$ for $0 < a < ½$; h = 0 for $½ < a < 1$. Start with $0 < a < ½$.

A:=ww.H_3^7; AA:= A.uc; AAA:= expand(AA); a:= 7w – ½:

AAAA:=tcof*expand(AAA). The resulting value is

eval(AAAA,{w = 1/14}) produces check for the beginning of the interval;

eval(AAAA,{w = 3/14}) produces the anticipation for the front of the next interval.

The pattern continues until full probability is achieved, i.e., a = 5. The benefit of large p occurs at [a] =2; then, nn=p = 7 and the $H_{3,1}$ element appears in H_p^7 only on the main diagonal and below. The value of ww is one at position 3 and uc = 1 at position 5, and it is no longer necessary to discriminate the midpoint of an interval.

Further Generalities.

Non integral values of c cause an expansion of the entrance vector. First let us add to the general remarks and treat the major issue. When c is an integer we are dealing with a homogenous Markov chain whose unit length time axis is portioned into n + c equal intervals. The total value of its Poisson rate parameter is taken to be n + c, hence each of the intervals has rate = 1; a useful convenient choice. If c is not an integer then the equal width intervals are n+[c] in number and each still has rate parameter equal to one. But there is an initial interval of length $x_0 = (c-[c])/(n+c)$ and its Poisson rate is $\gamma = c - [c]$. This adds another, but narrower, cylinder set to the front of our sample space. Further, it is related to the initial condition issue when utilizing the difference equation approach.

In any view the Poisson process enters this space at its zero point and can produce some counts in this narrowed cylinder set. It is possible that these counts can acquire any of the several levels in the corridor, prior to passing the signal on to the homogeneous part of the chain. The methodology from that point forward has already been described.

This narrowed cylinder set can be viewed as supporting a matrix similar to H, but the values it contains require adjusting. Further we only need one column of this matrix; the one that properly locates the entry position of the Poisson process. The entry point is no longer necessarily [a] +1, as it was before, and the values beyond the entry point must be calculated. The order p is the same as for the other cylinder sets. The boundary vector will be called u_{cc} (replacing $u_{c\prime}$) and a different way of determining the entry point is required.

The computation starts with specifying a and b, the intercepts (divided by n) of the upper and lower corridor boundaries. Next we compute p. The tail symmetric requirement specifies b = a + c and hence the computation of p is

(4.7) $p = [a] + [a + c] + 1$; further note $[a + c] = [a + \gamma] + [c]$. Let us introduce p*, which will be useful shortly,

(4.8) $p^* = p - [c] = [a] + [a + \gamma] + 1$.

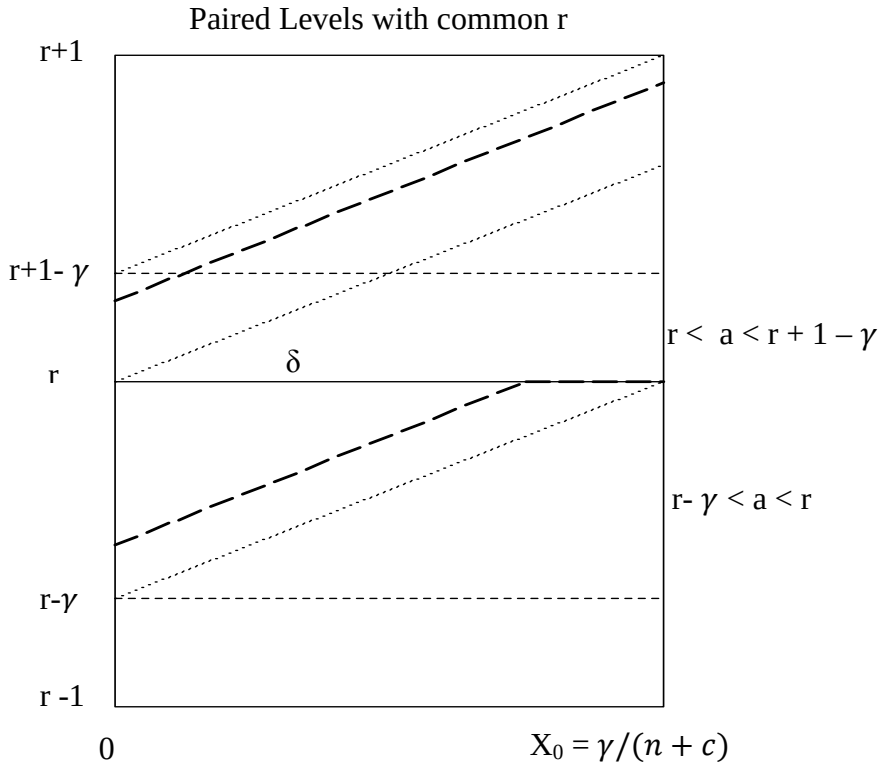
The smallest useful value of the intercept a leads to the initial value of p. It could be one (a very easy case) or it could be larger. But after that it increases in single steps. For example when a = [a] and p is determined, then the next value of p is acquired as soon as a increases to $[a] + 1 - \gamma$ because this causes $[a + \gamma]$ to gain a unit. Once that unit is acquired, then an additional increase in a by the amount γ will cause [a] to advance to [a] + 1, and hence p by another unit. These changes in p must occur in alternating sequence and each form has its own structure for the p^{th} member of u_{cc}.

A single column of an H matrix, other than the first, can serve as a model for the changes. It may begin with some zeros, before a value of one and subsequent values appear. When a one appears it should be viewed as γ^0 representing (i.e., the main input to) the probability of continuing at the present level without any fresh count(s). (This entire base interval is captured in our models without any clipping; the lower boundaries are distant, at – (a+c).) The value of one can be followed by the values $\frac{\gamma}{1}, \frac{\gamma^2}{2!}, \frac{\gamma^3}{3!}, \cdots$. The value of $u_{cc}(p)$ does not always follow the indicated progression. Let r + 1 be the number of non-zero members of the entrance vector. So u_{cc} begins with p – (r+1) zeros followed by r + 1 positive entries.

Let us discuss Figure 3.1; r must be at least two. Since $\gamma \neq 0$ there will be a change in p within each pair of consecutive intervals. The pair of basic intervals are

marked with the boundaries r – 1, r, r + 1. The in-between change points are marked with soft dashed lines in the graph.

Figure 4.1 An Adjacent Cell Pair.



	Algorithm			
	p – [c]	[a] + 1	p^* + [c] – r – 1	
Intervals	p^*	Exit point	No. of zeros	$u_{cc}(p)$
r – γ < a < r	2r	r	r – 1+ [c]	$(\gamma^r – \delta^r)/r!$
r < a < r + 1 – γ	2r + 1	r +1	2r + [c]	$\gamma^r/r!$

Consider an intercept value of the variable a within the lower interval, such as the one shown there with a heavy dashed line departing from the vertical axis. This line is a corridor upper boundary and it would intersect the right edge of the cylinder above the level r, but below the level r + 1 – γ. This boundary line clips a portion of the next level r, which cannot be reached by the Poisson process if it is to remain in the corridor.

The value of $u_{cc}(p)$ has a parallel to the last row of H_p, whose analysis is visualized in Figure 2.2 in the left side of the drawing. The comparable position contains

the symbol ε but the above figure contains a δ. This result is a product of the narrowed cylinder application and of the tail symmetry. The residues ε and δ relate to full cylinders of rate width unity and for the interval at hand: [a] = r – 1 and [a + γ] = r. Using Equation's (2.7, 2.8) and tail symmetry leads to δ = r – a and 1– ε = a + γ – r. In order to use ε in this context it must be shortened by 1 – γ for use in the cylinder of width γ. Accordingly,

(4.9) $\varepsilon - (1 - \gamma) = r - a = \delta$

and the use of δ is simpler in the calculation. Wherein the matrix H of (2.12) leads us to look for (1-ε^r)/r! in its pth row the modifications lead to (γ^r - δ^r)/r! as appears in the algorithm.

Turning to the second interval in the algorithm we see that the entry conforms to the natural progression indicated earlier. The reason may be seen in the Figure 4.1. The range for the heavy dashed line indicates that the boundary of the corridor does not intersect the next level at r + 1, yet the entire level r is captured.

In summary, the last entry in u_{cc} will be either with $\frac{\gamma^r}{r!}$ or $\frac{\gamma^r - \delta^r}{r!}$ (see Figure 4.1), according to whether the topmost level of the corridor is acquired in its entirety or whether part of it is clipped away by the corridor boundary.

To summarize the programmer's actions:

Probabilities for the smaller variate values are computed from the geometrically based formula. Convert w_0 to a_0 by inverting the relationship; determine γ from c. Determine r from the algorithm and the value of p. For the smaller values of p it is necessary to halve the established intervals in order to adjust for the two different expressions for h in $H_{p,1}$. Compare p with nn and prepare to economize on steps when p acquires an acceptable value circa nn. Advance r and continue until r = n, (a = n).

B. FEATURE STATISTICS THAT HAVE NON INTEGRAL C

1. $TT_5 = T_5(\frac{1}{6}, \frac{2}{3})$; $w = \frac{3}{17}a - \frac{1}{34}$ and $a = \frac{17}{3}w + \frac{1}{6}$; $b = a + \frac{2}{3}$; $gam = \frac{2}{3}$.

The parameter p becomes more intricate; $p = [a] + [a + \frac{2}{3}] + 1$.

Notice that w = 0 when $a = \frac{1}{6}$. So the interval $0 < a < \frac{1}{6}$ cannot have positive probability. Again it is simpler to start the process with $P\{w\} = 5!2^5w^5$ for $0 < w < w_0 = 3/34$. This

corresponds to 1/6 < a < 2/3. P(3/34) = $\frac{29160}{1419857}$ anticipates the beginning probability of the next interval, which begins at a = 2/3; p = [a] + [a+2/3] + 1 = 2 and 2/3 < a < 1. This interval contains the h > 0 input to H. tcof = 18/289; nn = 2.

The Maple code used is

H2 :=<<1– δ, (1– δ)(1–ε)>|<1, 1–ε>>; ww :=<1|0>; ucc := (1 , a – ⅓)´;

A:=ww.H2^2; AA:=A.ucc; AAA:= expand(AA); d := 1-a; ep:=4/3– a;

A:= (17/3)w +1/6; AAAA:= tcof*expand(AAA). The result is P(w) =

$-960w^5 + \frac{5760}{17}w^4 + \frac{21600}{289}w^3 - \frac{38880}{4913}w^2 + \frac{14580}{83521}w$

and P(3/34) = $\frac{29160}{1419857}$ and P(5/34) = $\frac{262200}{1419857}$; as check points

h is unimportant at 7/3 < a < 8/3 and ff p=6, nn=5

2. $FF_5 = T_5(-\frac{1}{8}, \frac{5}{4})$; w = $\frac{4}{25}$ a + $\frac{1}{50}$ and a = $\frac{25}{4}$w – $\frac{1}{8}$; b = a + $\frac{5}{4}$; gam = $\frac{1}{4}$.

The parameter p becomes more intricate; p = [a] + [a + $\frac{5}{4}$] + 1.

Notice that w = 0 then a = $-\frac{1}{8}$. So the interval $-\frac{1}{8}$ < a < 0 cannot contain an edf. The process must be started with P{w} = $5!2^5w^5$ for 0 < w < w_0 = 2/25. This corresponds to –1/8 < a < 3/8.

P(2/25) = $\frac{24576}{1953125}$ anticipates the beginning probability of the next interval, which begins at a = 3/8; p = [a] + [a+5/4] + 1 = 2 and 3/8 < a < 3/4. This interval contains the h = 0 input to H. tcof = $\frac{24576}{1953125}$; nn = 6; mx_0 = 1/20.

The Maple code used is

H2:= <<1-d,(1-d^2 – ep^2/2 >|< 1,1-ep>>;ww := <1|0>; ucc:=<0,1>;

A:=ww.H2^6; AA:= A.ucc; d:= 1– a; ep := 11/8 – a; AAA:= expand(AA)

a:=(25/4)w – (1/8); AAAA:= tcof*expand(AAA). Result is

$-960w^5 + \frac{1536}{5}w^4 + \frac{1536}{25}w^3 - \frac{18432}{3125}w^2 + \frac{9216}{78125}w$

The check points are P{$\frac{2}{25}$} = $\frac{24576}{1953125}$ and P{$\frac{7}{50}$} = $\frac{265398}{1953125}$.

Because of the large c the vector u_{cc} advances more rapidly. Consider the interval $\frac{7}{4}$ < a < $\frac{15}{8}$. The order p = 5 and nn = 6. The first super diagonal of H_5^6 is infected by the choice of h. The exit and entrance vectors are ww = 1 @ position 2 and ucc begins

nonzero values at position 3. Since this point is on the mentioned super diagonal, we must use the proper $H_{5,1}$ value and move on.

Approximate p-values.

Because of the Slutsky Theorem all of our statistics have this same asymptotic (Kolmogorov-Smirnov (K-S) distribution. It is given by

(4.8) Limit $P\{D_n \leq \frac{d}{\sqrt{n}}\} = 1 - 2\sum_1^\infty (-1)^{j-1}\exp(-2j^2d^2)$ as $n \to \infty$.

A matching Maple code is

$$KScdf := (t) \to 1 - 2\,sum\left((-1)^{j-1} \cdot \exp\left(-2j^2 \cdot t^2\right), j = 1..infinity\right) .$$

The work of Gibbons and Chakraborti tells us that the K-S large sample approximation becomes useful for the K-S procedure when samples are of size 25 to 30 and beyond. Our work is expected to be more useful when experiments are expensive and small samples must be relied upon for progressive decisions, often in an ad hoc manner. The piecewise feature generally means lots of pieces. There may be some alternative ways to use the approximation.

Figures 3.1 and 3.2 can help describe the suggestion. The first one contains six plots, five of which are the cdf's of our featured statistics, all with n = 5, and in reverse order of the parameter c. To their right is the asymptotic KS distribution for n = 5. The second figure contains plots of the asymptotic distribution for the marked values of n.

Figure 3.1 has a curious feature. Notice how they bunch up having common shape. The KS curve is of similar shape, but more distant. It may be useful to measure the distance of the KS approximation curve from a member of the collected pack of curves at, say, the 95% point, (such values have the greatest interest). Then without computing the selected distribution, use its test statistic value, translate it to the right using the recorded distance, followed by the percentile level from the approximation curve.

Figure 4.2 CC_5 FF_5 C_5 TT_5 CL_5 KS_5

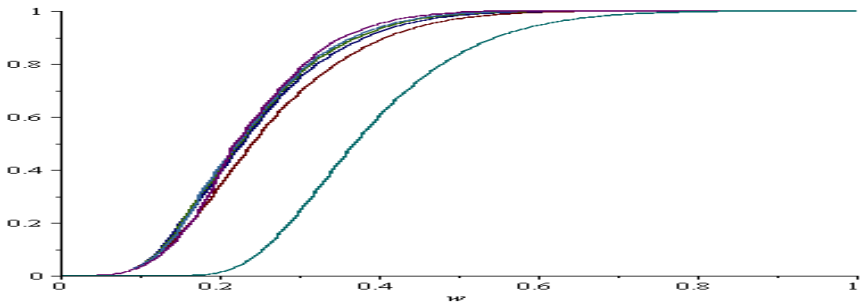
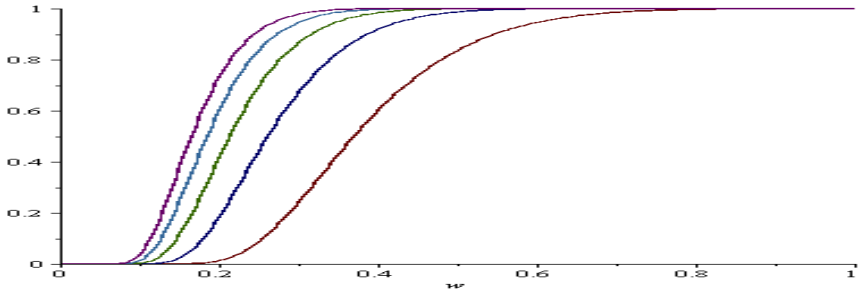


Figure 4.3 K-S Approximations



Sample size values are 25 20 15 10 5

This value might serve as an approximation to the percentile that we desire, and in that case a useful approximation. On the other hand it could produce a more distant value which indicates a weaker approximation (less representative of the selected curve).

The collection of K-S approximations in Figure 4.2 all translate left as n grows. Notice that the $KS_{n=10}$ curve may match up well with 95 percentile of the group in Figure 3.1. This observation triggers an alternative thought; use (say) the n=10 K-S percentiles to approximate n = 5 tail symmetric percentiles.

Ideas of this type should be more carefully explored. Section VI of this report contains a subsection marked “Use of p –values. It reflects upon broad, ad hoc use of approximate p-values and the selected need for sharper decision methods that can emerge at key times.

Table 4.1 Selected Quantiles.

Statistic	$q_{.25}$	$q_{.50}$	$q_{.60}$	$q_{.75}$	$q_{.80}$	$q_{.90}$	$q_{.95}$	$q_{.99}$	mean
KS_5	0.5500	0.6084	0.6326	0.6751	0.6926	0.7398	0.7793	0.8532	
CL_5	0.1776	0.2419	0.2697	0.3238	0.3470	0.4094	0.4633	0.5685	0.2583
TT_5	0.1653	0.2286	0.2549	0.3019	0.3229	0.3831	0.4327	0.5308	0.2427
C5	0.1631	0.2249	0.2511	0.2964	0.3155	0.3736	0.4217	0.4217	0.2386
FF_5	0.1659	0.2230	0.2491	0.2937	0.3120	0.3676	0.4149	0.5058	0.2368
CC_5	0.1753	0.2202	0.2465	0.2901	0.3071	0.3539	0.3994	0.4816	0.2361
D_5	0.2776	0.3419	0.3697	0.4238	0.4470	0.5094	0.5633	0.6685	0.3583

Table 4.1 provides a sharper comparison of the statistics studied for the marked percentiles.

Remark. A determination of the mean of KS_5 was attempted using the integral of the survivor function technique. Stabilization of the integral of the infinite series occurred at about 30,000 terms, producing a value of 0.3885. This value lacked credibility and was not placed in the table.

THIS PAGE INTENTIONALLY LEFT BLANK

V. DIFFERENCE EQUATION APPROACH

Durbin's intention for the Markov Chain development was to use the result for providing the initial conditions needed to solve his generalization of Massey's (1950) difference equation. He developed a generating function (pages 400-403 of his 1968 article approach in order to extract the difference equation. A portion of Section V is borrowed directly from his paper and serves to illustrate the method, its peculiarities, and limitations. A key step in this process however is a series formula used in the generating function that was presented undeveloped. Its derivation is produced here in a sub-section marked Determinant Development. The proof is intricate and is worthy of being recorded. Other parts of this section of Durbin are included for the sake of immediate reference.

The equations in this section are marked with a "D-" preceding the number. These designators are the same as in the Durbin paper; so is the notation used in the equations.

Durbin's generating function method can be applied to many maximum norm type statistics including the members of our family $\{T_n(0, c)\}$ and the original Kolmogorov statistic. The method of application is presumptuous as there is little direction for showing how to make choices. Further, little mention was presented to support the case of non-integral c; and the use of a translation parameter $\alpha \neq 0$ had not been raised. Our Section IV introduces the method for managing the non-integral c issue and is explicit for the tail symmetry set.

Generating Function for $q_n(a, b, c)$.

$p_n(a, b, c) = n!\, w' H^{[n+c]}\, u_{c'} / (n + c)^n$; $q_n(a, b, c) = w' H^{[n+c]}\, u_{c'}$.

The symbol w is used for the Markov exit vector and $u_{c'}$ is the entrance vector.

The quantity $q_n(a, b, c)$ is the coefficient of z^n in the generating function

$$f(z) = \sum_{r=0}^{\infty} w' H^r\, u_{c'}\, z^{r+c'} = z^{c'} w' [I - zH]^{-1} u_{c'}.$$

The series expansion of $[I - zH]^{-1}$ being valid for $|z\lambda| < 1$, where $\boldsymbol{\lambda}$ is the largest eigenvalue of H in modulus. Let Γ be the adjoint matrix of $I - zH$. Then

(D-3) $$f(z) = z^{c'}\, w' \Gamma\, u_{c'} / |I - z\, H|,$$

where $w' \Gamma\, u_{c'}$ and $|I - z\, H|$ are polynomials in z of orders p – 1 and p, at most, respectively. This is Equation (D-3) in Durbin.

Skipping to (D-7) in Durbin and using (D-4), which is justified in the marked determinant section, produces

(D-7) $|I - z\,H| = \sum_{j=0}^{[a+b]}(-1)^j((a+b-j)^j/j!)z^j = g(z, a+b)$ say.

Cross-multiplying in (D-3) and writing

(D-8) $f(z) = \sum_{r=c'}^{\infty} q_r(a, b, c)\, z^r$

leads to

(D-9) $g(z, a+b) \sum_{r=c'}^{\infty} q_r(a, b, c)\, z^r = z^{c'} w' \Gamma u_{c'}$.

Since the RHS of (D-9) is a polynomial of degree at most c'+ p – 1 = – [c] + [a+c] +[b–c], on equating the coefficients of z^r on both sides of (D-9) we have

(D-10) $\sum_{j=0}^{[a+b]}(-1)^j((a + b - j)^j/j!)\, q_{r-j}(a, b, c) = 0,$

$r = -[c] + [a+c] + [b-c] + 1, -[c] + [a+c] + [b-c] + 2, \cdots,$

where $q_s(a, b, c)\, z^r = 0$ for $s < 0$. This is the generalization of Massey's difference equation.

Determinant Development.

This separate section provides a proof of the formula

(D-4) $|H + yI| = \sum_{j=0}^{[a+b]}[(a+b-j)^j/j!]y^{p-j}$.

Let $D_r(\delta, \varepsilon) = |H + yI|$ and turn to the task of verifying the formula

(D-5) $D_r(\delta, \varepsilon) = \sum_{j=0}^{r'}[(r + 1 - \delta - \varepsilon - j)^j/j!]\, y^{r-j}$ r = 2, 3, ⋯.

Proof. By induction.

First. Show directly that (D-5) is true for r = 2.

$$D_2(\delta, \varepsilon) = \begin{vmatrix} 1-\delta+y & 1 \\ (1-\delta^2-\varepsilon^2+h)/2! & 1-\varepsilon+y \end{vmatrix}$$

$= (1-\delta+y)(1-\varepsilon+y) - (1-\delta^2-\varepsilon^2+h)/2!$

Case (i) $\delta + \varepsilon \leq 1$, which implies that h = 0 and $r' = r$. Writing the above as a polynomial in decreasing terms leads to

$y^2 + (2 - \delta - \varepsilon)\, y + (1-\delta-\varepsilon)^2/2!$ which conforms to (D-5) using r = 2.

Case (ii) $\delta + \varepsilon > 1$, which implies that $h = (\delta + \varepsilon - 1)^2$ and $r' = 1$. Then, (D-5) has two terms, the change leading to the constant term being omitted, hence zero. But the algebra of the direct computation also leads to the constant term being zero.

Second. Turn to the inductive step of the proof of (D-5). Assume that the form is valid for subscripts = r – 1, r –2, ⋯, 3, 2 and turn to the task that it is also valid for subscript r. The approach is to verify all of the coefficients of the powers of y. The first step in this process is the expand D_r (δ, ε) by its first column. The result is

(D-6) $D_r(\delta, \varepsilon) = (1- \delta +y)\, D_{r-1}(0, \varepsilon) - ((1- \delta^2)/2!)\, D_{r-2}(0, \varepsilon) + \cdots +$

$(-1^{r-i-1})\, ((1- \delta^{r-i})/(r-i!))\, D_i(0, \varepsilon) + \cdots +$

$(-1)^{r-3}((1- \delta^{r-2})/(r-2)!)\, D_2(0, \varepsilon) + \leq$

$(-1)^{r-2}((1-\delta^{r-1})/(r\text{-}1)!)\, (1 - \varepsilon + y) +$

$(-1)^{r-1}\, ((1- \delta^r - \varepsilon^r + h)/r!),\ r = 3, 4, \cdots$

(a) Starting with k = r it is seen that y^k appears only in the first term on the RHS of (D-6) and as y times the coefficient of y^{r-1}in $D_{r-1}(0, \varepsilon)$. This latter value is obtained from (D-5), which is valid when r is replaced by r–1, and we need only the term in the summation for j = 0. It follows that the sought after coefficient is unity.

Notice that the value of h plays a role only in the constant term of (D-6). We need not worry about the distinction between case (i) and case (ii) as long as we are inducing the coefficients of the positive powers of y.

(b) Turning to k = r–1, we must combine the coefficients of y^{r-2} and of y^{r-1} from the first term on the RHS of (D-6). This is because

Dr(δ, ε) = (1– δ + y)Dr-1(0, ε) + terms having powers of y lower that r-1.

These values are obtained from (D-5), i.e.,

$D_{r-1}(0, \varepsilon) = \sum_{j=0}^{r'-1} \frac{(r-1+1-\varepsilon-j)^j}{j!}\ y^{r-1-j}$.

It is seen that the coefficient of y^{r-2} is r–ε–1, and that of y^{r-1} is unity. It follows that the coefficient of y^{r-1} in (D-6) is (r – δ – ε). Since this value matches the one in (D-5), the inductive step is completed for the power r-1.

(c) For purposes of insight into the direction that we are taking, let us repeat this process for k = r–2. From (D-5) we see that the goal is to verify that

• $(r- 1- \delta- \varepsilon)^2/2!$ is the sought after coefficient of y^{r-2}.

This time we will need the first two terms on the RHS of (D-6). The first, $D_{r-1}(0, \varepsilon)$, is displayed in the previous paragraph. The second is

$D_{r-2}(0, \varepsilon) = \sum_{j=0}^{r'-2}[(r-2+1-\varepsilon-j)^j/j!]y^{r-2-j}$.

This time we must gather three terms: (1–δ) multiplies (r– 1– ε); 1 multiplies $(r-2-\varepsilon)^2/2!$; $-(1-\delta^2)/2!$ multiplies unity. Gather to obtain

$$\frac{1}{2}\{(r-2-\varepsilon)^2 + 2(1-\delta)(r-1-\varepsilon) - (1-\delta^2)\}.$$

The fact that this agrees with • can be shown by writing the bracketed terms as

$$[(r-1-\delta-\varepsilon) - (1-\delta)]^2 + 2(1-\delta)(r-1-\varepsilon) - (1-\delta^2),$$

followed by performing the square operation upon the indicated partition of the first term. Keep the goal term and the residue terms will add to zero.

(c) Let us turn to the general terms. According to (D-5), we need to verify that the coefficient of y^k for k = r-1, r-2, ⋯, 1 is given by

• goal $(k + 1 - \delta - \varepsilon)^{r-k}/(r-k)!$.

In order to induce the correctness of this form, we will extract the coefficient of y^k from each of the terms on the right hand side of (D-6). Recall that only the first r-k terms in (D-6) make a contribution.

The contribution of $D_{r-1}(0, \varepsilon)$ will be special; i.e., its first coefficient is

1 – δ + y. Accordingly, we need two terms from D_{r-1}. They are the coefficients of y^k *and* y^{k-1}. They are $(k+1-\varepsilon)^{r-k-1}/(r-k-1)!$ and $(k-\varepsilon)^{r-k}/(r-k)!$, respectively. The first one is to be multiplied by 1–δ and the other by unity. Accordingly, this term contributes two coefficients to y^k:

$(1-\delta)\,((k+1-\varepsilon)^{r-k-1}/(r-k-1)!$ and $(k-\varepsilon)^{r-k}/(r-k)!$.

All the other terms have a common pattern. The general case is, using (D-5),

$D_{r-i}(0, \varepsilon) = \sum_{j=0}^{r'-i}[(r-i+1-\varepsilon-j)^j/j!]\, y^{r-i-j}$.

We need the coefficient corresponding to k = r – i – j. It is

$(k+1-\varepsilon)^{r-k-i}$ / (r – k – i)! and is to be multiplied by

$(-1^{i-1})\,((1-\delta^i)/i!)$

because of its position in (D-6). Note that the contribution is zero when i=0.

Next, sum the coefficients. Let us treat case (i) first; h=0 and r' = r. Gather

$$(k-\varepsilon)^{r-k}/(r-k)! + \sum_{i=0}^{r-k}(-1)^{i-1}\binom{r-k}{i}\{(k+1-\varepsilon)^{r-k-i}$$

$$-\delta^i(k+1-\varepsilon)^{r-k-i}\}/(r-k)!.$$

The isolated first term is one of the two from D_{r-1} , while the other term from there melds with the general form. Continuing

$= \{(k-\varepsilon)^{r-k} - (k+1-\varepsilon-1)^{r-k}\}/(\mathrm{r}-\mathrm{k})!$

$+(k+1-\delta-\varepsilon)^{r-k}/(\mathrm{r}-\mathrm{k})!.$

The first two terms add out and the goal is confirmed.

It remains to verify (D-5) for k = 1 and 0 as well.

For k = 1, record the

goal $(2-\delta-\varepsilon)^{r-1}/(\mathrm{r}-1)!.$

There are three places in (D-6) that contribute to this linear term coefficient.

First, are r-2 terms from the established pattern of D_{r-1}, ⋯, D_2. These have the form, including their coefficients

$$(-1)^{i-1}\ ((1-\delta^i)/\mathrm{i}!)\ (2-\varepsilon)^{r-1-i}/(\mathrm{r}-1-\mathrm{i})!,$$

and to the sum of these we must add a contribution from the nonexistent D_1, whose role is filled by its position in (D-6) with the form

$$(-1)^{r-2}((1-\delta^{r-1})/(\mathrm{r}-1)!).$$

It is readily seen that this quantity matches the previous one when i is replaced with r – 1.

Further, there is a hidden linear term in the first term of (D-6), whose presence requires us to include the constant term of D_{r-1} (0, ε) (use (D-5) with

r ← r – 1), which is

$$(1-\varepsilon)^{r-1}/(\mathrm{r}-1)!.$$

Now we are positioned to gather and sum.

$$\frac{1}{(r-1)!}\{\sum_{i=0}^{r-1}(-1)^{i-1}\binom{r-1}{i}(1-\delta^i)(2-\varepsilon)^{r-1-i}\} + (1-\varepsilon)^{r-1}/(\mathrm{r}-1)!$$

$$= \frac{1}{(r-1)!}\{-\sum_{i=0}^{r-1}\binom{r-1}{i}(-1)^i\ (2-\varepsilon)^{r-1-i} +$$

$$\sum_{i=0}^{r-1}\binom{r-1}{i}(-\delta)^i(2-\varepsilon)^{r-1-i}\} + (1-\varepsilon)^{r-1}/(\mathrm{r}-1)!$$

$$= \frac{1}{(r-1)!}\{-(1-\varepsilon)^{r-1} + (2-\delta-\varepsilon)^{r-1} + (1-\varepsilon)^{r-1}\}$$

= goal.

For k = 0. Case (i) Recall r' = r and h = 0. Goal is $(1-\delta-\varepsilon)^r/\mathrm{r}!$.

We must gather all the constant terms from (D-6). For all, save the last two terms we have, coupled with their coefficients,

$(-1)^{i-1}((1-\delta^i)/i!)\ ((1-\varepsilon)^{r-i}/(r-i)!$ for i = 1, ⋯, r – 2.

The remaining two terms are

$(-1)^{r-2}((1-\delta^{r-1})/(r-1)!)(1-\varepsilon)$ and $(-1)^{r-1}(1-\delta^r-\ \varepsilon^r)/r!$

and the first of these has the pattern to qualify it for membership in the previous system at r = r – 1. So, gather and sum, letting i extend to r followed by an adjustment for the new value. The summation can be written

$$\frac{1}{r!}\{\sum_{i=1}^{r}(-1)^{i-1}\binom{r}{i}(1-\ \delta^i)\ (1-\varepsilon)^{r-i}+(-1)^r(1-\delta^r)\}+$$

$$(-1)^{r-1}(1-\delta^r-\ \varepsilon^r)/r!$$

$$= \frac{1}{r!}\{-(-\ \varepsilon)^r\ +(1-\delta-\varepsilon)^r+(-1)^r(1-\delta^r)+(-1)^{r-1}(1-\delta^r-\ \varepsilon^r)\}$$

= goal.

<u>For k = 0. Case (ii).</u> Recall r' = r – 1 and h = $(\delta+\varepsilon-1)^r$. Goal = zero.

The only change from the previous is in the contribution from the last term in (D-6). When that replacement is made in the above, it produces the negative value of the previous goal. Hence, the goal of zero is obtained. The formula (D-5) has been proven in its entirety.

The immediate exploitation of (D-5) utilizes the formula, when r = p,

(p + 1 – δ – ε – j) = (a + b – j) for all δ, ε in (0, 1).

Proof. Recall that a + b = [a] + [b] + δ + ε and that [a] + [b] = p or p – 1, depending upon whether δ + ε is ≤ 1 or > 1, respectively.

<u>Return to the Generating Function.</u>

It has been proved by induction that |H + yI| = $\sum_{j=0}^{[a+b]}\frac{(a+b-j^j}{j!}y^{p-j}$. Substituting y=$z^{-1}$allows us to define

g(z, a+b) = |I – zH | = $\sum_{j=0}^{[a+b]}(-1)^j\ \frac{(a+b-j)^j}{j!}z^{\ j} = z^{c\prime}w'\ \Gamma\ u_{c\prime}$.

Recall the generating function f(z) = $\sum_{r=c\prime}^{\infty} q_r$(a, b, c) z^r and cross-multiplying leads to

g(z, a+b) = $\sum_{j=0}^{[a+b]}(-1)^j \frac{(a+b-j)^j}{j!} z^j = z^{c\prime} w' \Gamma u_{c\prime}$.

The right-hand side is a polynomial of degree at most c' + p – 1, so equating coefficients allows us to write Durbin's equation

(D-10) $\sum_{j=0}^{[a+b]}(-1)^j \frac{(a+b-j)^j}{j!} q_{r-j}(a,b,c) = 0;\ r = -|c| + p - 1,\ -|c| + p, \cdots$ and

q_s(a, b, c) =0 for s < 0.

The Equation (D-3) can be simplified, since w′$\Gamma u_{c'}$ is then the (|b–c+1|,|b+1|)th element of Γ, which is the (|b–c+1|,|b+1|)th cofactor of I – zH. By deleting the |b+1|th row and the |b-c+1|th column of [I – zH], we see that this cofactor is $(-1)^c$ times $(-z)^c$ times the product of the two determinants of order [b – c] and p-1-[b] = [a], respectively, i.e., consider the below two determinants of submatrices of [I – zH. Both have the form of |I – zH|, with p = [b – c], ε = 0, and p = [a], δ = 0, respectively. Their product is therefore

g(z,b-c)g(z,a)

by (D-7). Substituting in (D-3) we have

(D-11) f(z) = g(z, a)g(z, b – c)/g(z, a+b).

Equating the coefficients of z^r on both sides of the equation

f(z)g(z, a+b) = g(z, a)g(z, b-c) for r = 0, 1, 2, ⋯, – c + [a] + [b]

is a way for getting the initial conditions needed for the application of (D-10)

I-zH with p = [b-c] and ε = 0

$$\begin{vmatrix} 1-z(1-\delta) & -z & 0 & \cdots & 0 \\ -z((1-\delta^2)/2! & 1-z & -z & \cdots & \cdots \\ \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ & & & \cdots & \cdot \\ -z(\frac{1-\delta^{[b-c]}}{[b-c]!}) & -z/[b-c]! & \cdots & \cdots & 1-z \end{vmatrix}$$

I-zH with p = [a] and δ = 0

$$\begin{vmatrix} 1-z & -z & 0 & \cdots & \cdots & 0 \\ -z/2! & 1-z & -z & \cdots & \cdots & 0 \\ \cdot & & & & & \\ \cdot & & & & \cdot\ \cdot & \\ & & & & \cdot & -z \\ -z(\frac{1-\varepsilon^{[a]}}{[a]!}) & & & \cdots & & 1-z(1-\varepsilon) \end{vmatrix}.$$

Multiplying the coefficient of z^n in (D-11) by $\frac{n!}{(n+c)^n}$ provides the exact form

(D-12) $p_n(a, b, c) = \frac{n!}{(n+c)^n} \sum_{r=0}^{[a]} \sum_{s=0}^{[b-c]} \frac{(-1)^{r+s}}{r!s!} \gamma_{n-r-s} \, (a - r)^r \, (b - c - s)^s$

for $n \geq [a] - [b - c]$, where γ_j is the coefficient of z^j in the expansion of $g(z, a + b)^{-1}$, i.e.,

(D-13) $$\gamma_j = (-1)^j \sum\nolimits' \{ \frac{(i_1 + \cdots + i_{m)!}}{(i_1! \cdots i_m!)} \prod_{h=1}^{m} (-1)^{i_h} \cdot \{ \frac{(a+b-h)^h}{h!} \}^{i_h},$$

where $\sum'$ indicates summation over all sets $(i_1, \cdots, i_m)$ of nonnegative integers satisfying $i_1 + 2i_2 + \cdots + mi_m = j$ and where $m = [a + b]$. This generalizes Kemperman's [5.40] result for $c = 0$.

Durbin also cites an avenue using Dempster (1959, 5′) as a way to the goal. The issue of tail symmetry (and translation to achieve it) appears simplified because $b = a + c$ achieves that goal. But the problem of c non-integral remains, although our description of the entrance vector $u_{c'}$ in Section IV may be used in (D-3).

VI. REASEARCH NOTES

Choice of Featured Statistics

$CL_n = T_n(½, 0)$. The quantile set used goes back to the days of normal probability paper and remains in use as the default quantiles for q-q plots. I chose the symbol because of the system's use in the work of Chernoff and Liebermann.

$C_n = T_n(0,1)$. Common designation for the Pyke modification. Pyke drew attention to the full symmetry property.

$TT_n = T_n(\frac{1}{6}, \frac{2}{3})$. TT stands for c = two-thirds, which was recommended in the early days of q-q plots, but I have lost the citation. It is useful to feature a statistic having c between zero and one.

$FF_n = T_n(-\frac{1}{8}, \frac{5}{4})$. FF stands for c = five-fourths, which played a role in the Read (1972) paper. It is useful to feature a statistic having c between one and two.

$CC_n = T_n(-\frac{1}{2}, 2)$. CC stands for c =2; a balancing endpoint for the set.

Record the Ranges and w_0 Values

statistic	CL_n	TT_n	C_n	FF_n	CC_n
range	$1-\frac{1}{2n}$	$\frac{6n-1}{2(3n+2)}$	$\frac{n}{n+1}$	$\frac{8n+1}{2(4n+5)}$	$\frac{2n+1}{2(n+2)}$
w_0	$\frac{1}{2n}$	$\frac{3}{2(3n+2)}$	$\frac{1}{2(n+1)}$	$\frac{2}{4n+5}$	$\frac{1}{2(n+2)}$

Note. Range is the largest distance from the center to an edge of the domain. If a statistic is not tail symmetric the range would be the larger of $|\frac{n-\alpha}{n+c}|$ and $|1-\frac{1-\alpha}{n+c}|$.

Same Distribution Does Not Mean the Same Random Variable.

The result that $CL_n + 1/2n$ has the same distribution of D_n serves as a useful example in which the two random variables are identically distributed are functions of the same random sample, yet they are not the same random variable. A simple example suffices to show this.

For n = 1, the two random variables (one shifted) are the same. But not for n =2. Let n = 2, u_1= ¼, u_2 = ¾. Then, CL_2 = max(|¼–¼|, |¾–¾| = 0 and $CL_2 + 1/2n$ = ¼. On the other hand, $D_2 = \max\{sup_{0 \le x < ¼} |x|; sup_{¼ \le x < ¾} |¼-x|; sup_{¾ \le x < 1} |¾-x|\}$= ½. The separation of ¼ and ½ is not zero.

Stochastic Smallness

Definition. The random variable X is stochastically smaller than the random variable Y if $P\{X \leq Y\} = 1$ and the two variables are not identically distributed. An alternative way of describing this is in terms of the distribution functions:

(7.1) $F_X(v) \geq F_Y(v)$ for all v with strict inequality for at least one value of v.

Should it emerge that a member of the family $\{T_n(\alpha, c)\}$ is stochastically smaller than the other members, then it should be considered for use in q-q probability plots, for confidence bands for continuous distributions, and would be expected have power function advantages in competition with the classical Kolmogorov-Smirnov one sample test.

A necessary condition for X to be stochastically smaller than Y is that the range of X is no more than the range of Y. The range of CC_n is the smallest of those in the family $\{T_n(\alpha, c\}$ for $(0 \leq c \leq 2 \; and \; -½ \leq \alpha \leq ½)$ and each n. Further, given c, the range of $T_n(\alpha, c)$ is smallest for the tail symmetric member.

Complete smallness is in doubt because the formula for small variate values uses $n!(2w)^n$, the same for all of our statistics. But at $w=w_0$ the cdf of a statistic must change and continue to rise at a slower rate. The statistic having the largest w_0 is CL_n, but it has the greatest range. A smallness comparison of C_2 and CL_2 has been made at the end of Section 2.

The graphs in Figure 4.2 appear to award stochastic smallness to CC_n , and it may be useful to treat it as such. It makes the case for replacing CL_n with CC_n for use in the qq-probability plots. The full symmetry of C_n creates a strong, general purpose case.

Statistic Too Small

The Mendel-Fisher controversy (look it up on Google) has drawn attention to the possibility that goodness-of-fit statistics can be too small; evidence of faking the data. The controversy makes interesting reading; issues of this type require study well beyond the selection of a method. It is curious that a poorly selected method might provide better evidence than a well selected one.

Duplicated Computations.

It has already been seen that the use of Equation (2.16) can lead to duplicated computations when the $H_{p,\,1}$ element is not used, i.e., when p is large compared to nn, case specific. Material in Section IV shows us how to plan for this situation.

There are other situations of this cut, but their full understanding has yet to emerge. Recall that the distribution of CL_2 has two pieces while its fellow statistics all have three. It may be seen from Figure 3.1 that if the center of the square is at (1/4, ¾) then when w increases to touching the square to the diagonal, it also touches the top and left boundaries as well. If this distribution is computed using Equation (2.16), then result for p =2 and p =3 produce the same quadratic expression.

This phenomenon has been observed elsewhere. The same expression appears in two adjacent partitions for differing values of p. The study of this point seems difficult and has yet to be attempted.

Use of N-Dimnsional Geometry

In Section III the distributions of the n=2 statistics were all produced with a geometry based algorithm. It seems that such could also be managed with n = 3, but such has yet to be attempted. The tail symmetric family should be the first to be studied in this way.

The extension to n-dimensional geometry may require the attention of an expert in that field. But we are encouraged by the success of the n-dimensional hypercube for small variate values. It could be helpful in understanding the expression duplication problem discussed earlier.

Commentary on p-Values.

It appears that many lay statisticians use the terms p-value and level of significance interchangeably, and the former terminology has evolved in a manner that undermines the original intent. I believe that this commentary can promote a better understanding of what has happened.

The originators introduced level of significance as a preassigned upper limit on the probability of a type one error in statistical decision making. Such provides a level of purity for the theoretical statistician in his search for methodologies and for making choices among competing methodologies. The issue of how to choose a level of

significance in an application of methodology is left to the users who, more often than not, find that the tenant value selection process is not very compatible with his statistical problem and the complex issues affecting his need to make choices.

As a side commentary, there have been professional societies who have debated the issue of specific value level selection to be used as policy for all workers in their trade. Although there are many who would regard this as ill-conceived, I can also imagine situations in which policy makers are concerned about legal issues. Such can create a desire to protect their members with acts that impose rigid procedures.

Let us continue our discussion with the convenient assumption that our test statistics are one dimensional. Such simplifies the explanations. The decision rule is to reject the null hypothesis whenever the test statistic is larger than a critical point c_0, which in turn corresponds to the level of significance by means of an inversion process applied to the null distribution. This is equivalent to the comparison of the test statistic itself with the value c_0. The next step of convenience is to feed the test statistic value through the null distribution and make the decision by comparing the resulting probability, i.e., p-value, with the level of significance. Use of this practice tempts the user to avoid the step of choosing a level of significance, look only at the p-value, and make an ad hoc decision after the fact; even after the data are reviewed.

I would wager that there is a great deal of this type of activity being practiced. There are a huge number of situations, fields of application, and academic disciplines that utilize statistical methods. Each has its own needs for statistical methodology and its own window of satellite concerns, especially concerns of the type “where do we go from here? Statistical methodology itself has seen broad layers of subjectivity incorporated. The use of ad hoc methods should be better understood.

The use of asymptotic distributions in place of exact null distributions presents a source of error that is usually ignored. Since the p-value is a random variable, the large sample approximation offers an additional source of error and adds to the tendency to indulge in “ad hoc-ry”. The use of the practice is acceptable because of its popularity. When this subjective situation leads to a “close call,” it seems that a discussion of the effect of error be made in the context of consequences.

The intermediate steps in a research project often involve the making of choices based upon multiple considerations, not merely the statistical interpretations. These settings are more inviting to the use of “adhoc-ry”. It is imagined that the resource commitments for sharp statistical analyses are delayed until the project is ready for finalization steps.

Small Sample Exploration of Properties of Continuous Distributions.

The popular q-q probability plots have long been valuable for discerning distribution characteristics, based upon small sample information. If the eyeball allows the presence of a straight line in the plot, then location and scale parameters are having a dominant effect. If smooth curvature in the tails are discernable then skewness is present and peakedness or lack of may show up in the intermediate portions. The key for this depends upon smoothness, rather than oscillatory behavior. The smaller the sample size, the more difficult it is.

A member of our family of statistics identifies a quantile model with it that might allow localized exploration of distributional features that may be of interest to the analyst. Such has yet to be examined, but we have the computing power needed to consider the idea, even using quantile models outside of our structured family.

Parametric models have been studied for appropriateness using computer “adhoc-ry”. For example, both the Weibull and Gamma models both have shape and scale parameters. A scale parameter falls in the category of linear and is easily spotted. On the other hand shape parameters lead to skewness in the distribution. The computer can be used to successively change the shape parameter values and hopefully arrive at a usable fit.

Data transformations can be tried directly as well: If ln(X) is normal, the X has a log normal distribution; If X has an exponential distribution the log of the survivor function is proportional to X.

When all else fails the analyst may wish to look at a simple smooth “interpolatory” distribution function. My suggestion is to form a graph of order statistics against a selected quantile model; then fit a cubic spline. (This choice will manage infinite end points nicely.) The inverse function is a cdf and it can be differentiated

producing a continuous density. It could be useful in providing a smooth look at what is going on.

One is tempted to go a step further and draw crude bounds about the estimated cdf curve. Suppose that we take plus and minus values of the 90 percentile (say) of the chosen T_n and mark these off of the estimated quantile verses p plot. When rotating the axes about the 45^0 line (inverting the functions), cropping all values the are either negative or exceed one, then one has the estimated cdf and smooth bounds that are in some way associated with 0.95 probability. These curves are not confidence bands, but some sort of fiduciary that places its faith in the smoothness of the inverted cubic spline functions.

Because of the Slutsky Theorem, all of our statistics have this same asymptotic distribution. The issue of stochastic smallness is important in small sample size problems. A main idea is that the set of n order statistics can serve as a set of n quantile estimates of F. At issue is: which set of quantile identifying probabilities are best served in this way? The statistics T_n contain linear modifications of the empirical probability levels, {j/n}, that can serve in this role. The inverse relationship is

(6.2) $\quad j/n = (1+c)\, F(x_j) + \alpha/n.$).

We seek statistics that serve the concept of stochastic smallness.

VII. SUMMARY

The report introduces a broad class of statistics of the Kolmogorov type, having structure $T_n(\alpha, c) = \max\{|x_j - \frac{j-\alpha}{n+c}| \text{ for } j = 1, \cdots, n\}$ and focusing on the parameter space $\{0 \leq c \leq 2, -½ \leq \alpha \leq ½\}$. The space is selected as both interesting and providing good competitors for the statistics presently in practice. The data $\{x_j\}$ are order statistics from a continuous population. The use of maximization of the magnitudes of the deviations should help analysts that prefer to reduce the appearance of strong oscillations in the models treated by the statistics.

The author has small sample sizes in mind, largely associated with expensive experiments. The distribution computations are lengthy, but recent advances in computerized symbolic computation have afforded considerable relief to this problem. For the intermediate work in a project, some suggestions are contained that utilize the Kolmogorov-Smirnov large sample approximations in differing ways that do not require large samples.

There are reasons to favor statistics that are stochastically small. They appear in the construction of confidence bounds, helping to render them narrower. Heuristically it seems that the popular use of q-q probability plot in the study of the characteristics of data sets would prefer to use theoretical quantile sets that are identified with stochastic smallness.

Our featured statistics associate stochastic smallness with the larger values of the parameter c. The relationship is not absolute but the graphs contained within are compelling. They generally support the notion from a visual point of view. When there is a commitment to a choice of c, then comparative stochastic smallness is acquired with the use of the tail symmetry version of the category.

In addition to the family of statistics introduced, the new material also provides ways to analyze the algorithmic adjustments needed to manage the cases having non integral values for the parameter c. Such is needed for both the Markov Chain method of distribution calculation and the earlier developed difference equation method. The latter of these two methods has the disadvantage of requiring the computation of results for all

previous values of n, whereas the Markov Chain method goes directly to the value of n in consideration.

APPENDIX A. FEATURED STATISTIC DISTRIBUTIONS

$$CL5 := w \to piecewise\Big(0 \le w \text{ and } w < \frac{1}{10}, 3840\,w^5,\ \frac{1}{10} \le w \text{ and } w < \frac{1}{5},\ -288\,w^4 + \frac{624}{5}\,w^3 - \frac{96}{25}\,w^2 - \frac{36}{125}\,w + \frac{6}{625},\ \frac{1}{5} \le w \text{ and } w < \frac{3}{10},\ \frac{6}{125} - \frac{332}{125}\,w + \frac{616}{25}\,w^2 + \frac{24}{5}\,w^3 - 160\,w^4 + 160\,w^5,\ \frac{3}{10} \le w \text{ and } w < \frac{2}{5},\ \frac{343}{500}\,w - \frac{273}{1250} - \frac{318}{5}\,w^3 + \frac{542}{25}\,w^2 + 64\,w^4 - 20\,w^5,\ \frac{2}{5} \le w \text{ and } w < \frac{1}{2},\ \frac{2391}{500}\,w - \frac{3413}{6250} - \frac{62}{5}\,w^3 + \frac{6}{5}\,w^2 + 12\,w^5,\ \frac{1}{2} \le w \text{ and } w < \frac{7}{10},\ \frac{1383}{250}\,w - \frac{10527}{25000} - \frac{52}{5}\,w^3 - \frac{19}{5}\,w^2 + 18\,w^4 - 8\,w^5,\ \frac{7}{10} \le w \text{ and } w < \frac{9}{10},\ -\frac{9049}{50000} + \frac{6561}{1000}\,w - \frac{729}{50}\,w^2 + \frac{81}{5}\,w^3 - 9\,w^4 + 2\,w^5,\ \frac{9}{10} \le w \text{ and } w \le 1, 1\Big)$$

$$C5 := w \to piecewise\Big(0 \le w \text{ and } w < \frac{1}{12}, 3840\cdot w^5,\ \frac{1}{12} \le w \text{ and } w < \frac{1}{6}, -960\cdot w^5 + 320\cdot w^4 + \frac{200}{3}\cdot w^3 - \frac{20}{3}\cdot w^2 + \frac{5}{36}\cdot w,\ \frac{1}{6} \le w \text{ and } w < \frac{1}{4}, \frac{640}{27}\cdot w^2 - \frac{20}{3}\cdot w^3 - \frac{400}{3}\cdot w^4 + 160\cdot w^5 - \frac{275}{162}\cdot w + \frac{5}{648},\ \frac{1}{4} \le w \text{ and } w < \frac{1}{3},\ \frac{640}{27}\cdot w^2 - 40\cdot w^3 - \frac{425}{648}\cdot w - \frac{125}{1296},\ \frac{1}{3} \le w \text{ and } w < \frac{5}{12}, \frac{2855}{648}\cdot w - \frac{35}{3}\cdot w^3 + 12\cdot w^5 - \frac{1327}{3888} + \frac{5}{6}\cdot w^2,\ \frac{5}{12} \le w \text{ and } w < \frac{1}{2}, \frac{2855}{648}\cdot w - \frac{35}{3}\cdot w^3 + 12\cdot w^5 - \frac{1327}{3888} + \frac{5}{6}\cdot w^2,\ \frac{1}{2} \le w \text{ and } w < \frac{2}{3}, \frac{3125}{648}\cdot w - \frac{517}{3888} - \frac{25}{3}\cdot w^3 + \frac{50}{3}\cdot w^4 - 8\cdot w^5 - \frac{25}{6}\cdot w^2,\ \frac{2}{3} \le w \text{ and } w < \frac{5}{6}, \frac{763}{3888} + 2\cdot w^5 - \frac{25}{3}\cdot w^4 + \frac{125}{9}\cdot w^3 - \frac{625}{54}\cdot w^2 + \frac{3125}{648}\cdot w,\ \frac{5}{6} \le w \text{ and } w \le 1, 1\Big)$$

$$CC5 := w \to piecewise\Big(0 \le w \text{ and } w < \frac{1}{14}, 3840\,w^5, \frac{1}{14} \le w \text{ and } w < \frac{1}{7}, -960\,w^5 + \frac{1920}{7}\,w^4 + \frac{2400}{49}\,w^3 - \frac{1440}{343}\,w^2 + \frac{180}{2401}\,w,\ \frac{1}{7} < w < \frac{3}{14}, -\frac{1280}{7}\,w^4 + \frac{4160}{49}\,w^3 - \frac{320}{343}\,w^2 - \frac{940}{2401}\,w + \frac{160}{16807} + 320\,w^5,\ \frac{3}{14} < w \text{ and } w \le \frac{2}{7}, -\frac{190}{2401}\,w - \frac{1325}{16807} - \frac{1240}{49}\,w^3 + \frac{6320}{343}\,w^2 - \frac{80}{7}\,w^4,\ \frac{2}{7} < w \le \frac{5}{14}, \frac{2370}{2401}\,w - \frac{2349}{16807} - \frac{400}{7}\,w^4 + \frac{40}{49}\,w^3 + \frac{3760}{343}\,w^2 + 32\,w^5,\ \frac{5}{14} < w \le \frac{7}{14}, \frac{5765}{1372}\,w - \frac{7823}{33614} - \frac{550}{49}\,w^3 + \frac{30}{49}\,w^2 + 12\,w^5,\ \frac{7}{14} < w \le \frac{9}{14}, \frac{107861}{268912} + 2\,w^5 - \frac{55}{7}\,w^4 + \frac{605}{49}\,w^3 - \frac{6655}{686}\,w^2 + \frac{73205}{19208}\,w,\ \frac{9}{14} < w \le \frac{11}{14}, \frac{107861}{268912} + 2\,w^5 - \frac{55}{7}\,w^4 + \frac{605}{49}\,w^3 - \frac{6655}{686}\,w^2 + \frac{73205}{19208}\,w,\ \frac{11}{14} < w \le 1, 1\Big)$$

$$TT5 := w \to piecewise\Bigg(0 \le w \textbf{ and } w < \frac{3}{34}, 3840 \cdot w^5, \frac{3}{34} \le w \textbf{ and } w < \frac{5}{34}, \frac{21600}{289} w^3 - \frac{38880}{4913} w^2 + \frac{14580}{83521} w + \frac{5760}{17} w^4 - 960 w^5, \frac{5}{34} \le w \textbf{ and } w < \frac{3}{17}, -\frac{36180}{83521} w + \frac{12150}{1419857} + \frac{33840}{289} w^3 + \frac{2880}{4913} w^2 - \frac{5280}{17} w^4, \frac{3}{17} \le w \textbf{ and } w < \frac{7}{34}, -\frac{164700}{83521} w + 160 w^5 - \frac{1160}{289} w^3 - \frac{2400}{17} w^4 + \frac{119160}{4913} w^2 + \frac{24030}{1419857}, \frac{7}{34} \le w \textbf{ and } w < \frac{9}{34}, -\frac{164700}{83521} w + 160 w^5 - \frac{1160}{289} w^3 - \frac{2400}{17} w^4 + \frac{119160}{4913} w^2 + \frac{24030}{1419857}, \frac{9}{34} \le w \textbf{ and } w < \frac{11}{34}, -\frac{84510}{83521} w - \frac{8235}{83521} - \frac{13400}{289} w^3 + \frac{80}{17} w^4 + \frac{128880}{4913} w^2, \frac{11}{34} \le w \textbf{ and } w < \frac{6}{17}, \frac{675555}{334084} w - \frac{14950}{289} w^3 + \frac{960}{17} w^4 + \frac{73710}{4913} w^2 - 20 w^5 - \frac{623025}{2839714}, \frac{6}{17} \le w \textbf{ and } w < \frac{13}{34}, \frac{1504995}{334084} w - \frac{3430}{289} w^3 + \frac{270}{289} w^2 + 12 w^5 - \frac{1120689}{2839714}, \frac{13}{34} \le w \textbf{ and } w < \frac{1}{2}, \frac{1504995}{334084} w - \frac{3430}{289} w^3 + \frac{270}{289} w^2 + 12 w^5 - \frac{1120689}{2839714}, \frac{1}{2} \le w \textbf{ and } w < \frac{19}{34}, \frac{838475}{167042} w - \frac{1175}{289} w^2 - \frac{2580}{289} w^3 + \frac{290}{17} w^4 - 8 w^5 - \frac{2394731}{11358856}, \frac{19}{34} \le w \textbf{ and } w < \frac{23}{34}, \frac{838475}{167042} w - \frac{1175}{289} w^2 - \frac{2580}{289} w^3 + \frac{290}{17} w^4 - 8 w^5 - \frac{2394731}{11358856}, \frac{23}{34} \le w \textbf{ and } w < \frac{25}{34}, \frac{3536405}{668168} w + \frac{2206563}{22717712} + \frac{4205}{289} w^3 - \frac{121945}{9826} w^2 - \frac{145}{17} w^4 + 2 w^5, \frac{25}{34} \le w \textbf{ and } w < \frac{29}{34}, \frac{2206563}{22717712} + \frac{3536405}{668168} w - \frac{121945}{9826} w^2 + \frac{4205}{289} w^3 - \frac{145}{17} w^4 + 2 w^5, \frac{29}{34} \le w \textbf{ and } w \le 1, 1 \Bigg);$$

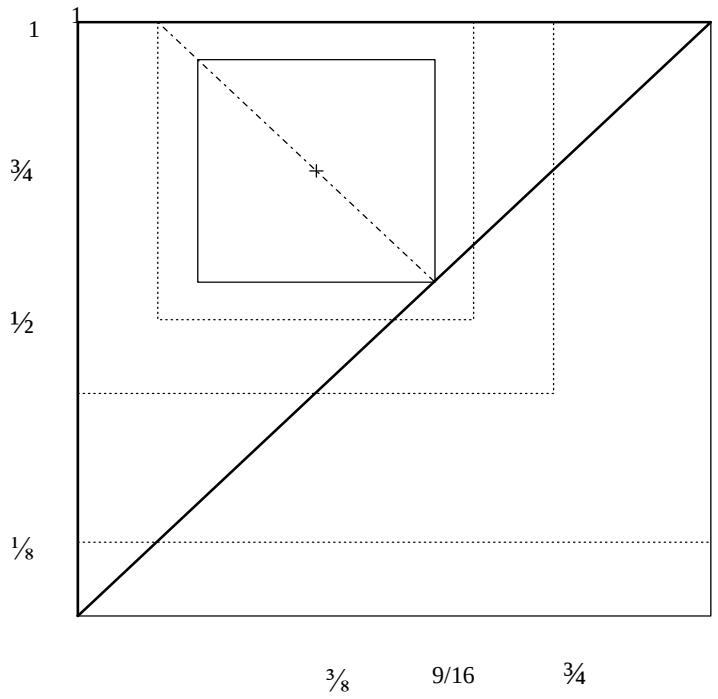
$$FF5 := w \to piecewise\Bigg(0 \le w \textbf{ and } w < \frac{2}{25}, 3840\,w^5, \frac{2}{25} \le w \textbf{ and } w < \frac{7}{50}, \frac{1536}{25}\,w^3 - \frac{18432}{3125}\,w^2 + \frac{9216}{78125}\,w + \frac{1536}{5}\,w^4 - 960\,w^5, \frac{7}{50} \le w \textbf{ and } w < \frac{4}{25}, \frac{1536}{25}\,w^3 - \frac{18432}{3125}\,w^2 + \frac{9216}{78125}\,w + \frac{1536}{5}\,w^4 - 960\,w^5, \frac{4}{25} \le w < \frac{9}{50}, -\frac{1024}{5}\,w^4 + \frac{13312}{125}\,w^3 - \frac{4096}{3125}\,w^2 - \frac{48128}{78125}\,w + \frac{32768}{1953125} + 320\,w^5, \frac{9}{50} \le w \textbf{ and } w < \frac{12}{50}, -\frac{118592}{78125}\,w + 160\,w^5 - \frac{1016}{125}\,w^3 - 128\,w^4 + \frac{72544}{3125}\,w^2 + \frac{1024}{390625}, \frac{12}{50} \le w \textbf{ and } w < \frac{3}{10}, -\frac{35648}{78125}\,w - \frac{181504}{1953125} - \frac{4472}{125}\,w^3 + \frac{69088}{3125}\,w^2 - \frac{16}{5}\,w^4, \frac{3}{10} \le w \textbf{ and } w < \frac{8}{25}, -\frac{35648}{78125}\,w - \frac{181504}{1953125} - \frac{16}{5}\,w^4 - \frac{4472}{125}\,w^3 + \frac{69088}{3125}\,w^2, \frac{8}{25} \le w \textbf{ and } w < \frac{17}{50}, \frac{95424}{78125}\,w - \frac{1956096}{9765625} - \frac{376}{125}\,w^3 + \frac{7264}{625}\,w^2 + 32\,w^5 - \frac{272}{5}\,w^4, \frac{17}{50} \le w \textbf{ and } w < \frac{23}{50}, \frac{1357411}{312500}\,w - \frac{3008056}{9765625} - \frac{1442}{125}\,w^3 + \frac{96}{125}\,w^2 + 12\,w^5, \frac{23}{50} \le w \textbf{ and } w < \frac{1}{2}, \frac{1357411}{312500}\,w - \frac{3008056}{9765625} - \frac{1442}{125}\,w^3 + \frac{96}{125}\,w^2 + 12\,w^5, \frac{1}{2} \le w \textbf{ and } w < \frac{31}{50}, \frac{733393}{156250}\,w - \frac{6486323}{78125000} - \frac{529}{125}\,w^2 + \frac{82}{5}\,w^4 - \frac{992}{125}\,w^3 - 8\,w^5, \frac{31}{50} \le w \textbf{ and } w < \frac{33}{50}, \frac{733393}{156250}\,w - \frac{6486323}{78125000} - \frac{529}{125}\,w^2 + \frac{82}{5}\,w^4 - \frac{992}{125}\,w^3 - 8\,w^5, \frac{33}{50} \le w \textbf{ and } w < \frac{39}{50}, \frac{40393799}{156250000} + 2\,w^5 - \frac{41}{5}\,w^4 + \frac{1681}{125}\,w^3 - \frac{68921}{6250}\,w^2 + \frac{2825761}{625000}\,w, \frac{39}{50} \le w \textbf{ and } w < \frac{41}{50}, \frac{40393799}{156250000} + 2\,w^5 - \frac{41}{5}\,w^4 + \frac{1681}{125}\,w^3 - \frac{68921}{6250}\,w^2 + \frac{2825761}{625000}\,w, \frac{41}{50} \le w \textbf{ and } w \le 1, 1\Bigg);$$

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. NON TAIL SYMMETRIC – GEOMETRIC

Use $T_2^* = T_2(0, ⅔)$ to illustrate; $\overline{u_1}$ = 3/8, $\overline{u_2}$ =3/4, w_0 = 3/16; center point is on the line $u_1 + u_2 = 9/8$; this line intersects $u_1 - u_2 = 0$ at (9/16, 9/16).

Figure B.2 Geometry of a Non Tail Symmetric Case.



Step 1. Twice the area of a square of radius (minimal distance from center to a side).

Step2. Continued expansion of the square beyond radius w_0 will be stopped at an edge; in this case, the top edge as the center point is above the symmetry guideline. The distance from the center point stops at $1-\overline{u_2}$ = ¼ and this is the limit of the second section. A square in this second section must be clipped by the excess or overlap induced into the lower right corner; it is a 45^o right triangle, whose edge is a distance w–3/16 from the intersection point. The probability excess of this triangle is $(2w\text{-}3/8)^2$; an amount to be subtracted from $8w^2$.

Step 3. The continued expansion beyond ¼ is rectangular and will be stopped again when it reaches the next boundary, the left in this example. But the growth in the area of the square is slowed because it cannot grow in the vertical direction. The formula for the area

becomes 2w(w + ¼), which must be doubled in order to convert to probability. The formula for clipping the excess probability caused by the triangle in the lower right remains the same, $(2w–3/8)^2$.

Step 4. Rectangular expansion can continue until it reaches the right edge, a distance of 5/8 from the center and a distance of 7/16 from the intersection point. It can expand in but two directions. The formula for area becomes (w+3/8)(w+1/4) and double that for probability. The corner clipping formula $(2w–3/8)^2$ is unchanged.

Step 5. The unanalyzed remaining part of the unit square is a rectangle at the bottom. The only part of it that concerns us is the triangle at the lower left. The additional growth of probability as w increases from 5/8 to ¾ is by means of trapezoidal regions starting with a baseline of 1/8 and a secondary line that diminishes to zero. The formula is

$(w – 5/8)(7/8 – w) + P\{5/8\}$.

$P\{T_2^* \leq w\} = 8w^2$	$0 < w < 3/16$	$P\{3/16\} = 9/32$
$= 4w^2 + \frac{3}{2}w - \frac{3}{16}$	$3/16 < w < ¼$	$P\{1/4\} = \frac{31}{64}$
$= \frac{5}{2}w - \frac{9}{64}$	$\frac{1}{4} < w < 3/8$	$P\{3/8\} = \frac{51}{64}$
$= -2w^2 + \frac{11}{4}w + \frac{3}{64}$	$3/8 < w < 5/8$	$P\{5/8\} = \frac{63}{64}$
$= -w^2 + \frac{3}{2}w + \frac{7}{16}$	$5/8 < w < ¾$	$P\{3/4\} = 1$

APPENDIX C. USE OF THE DIFFERENCE EQUATION

The technique for using the difference equations for purposes of computation and analysis is rather tersely described in the Durbin paper. In this appendix, we clarify the use of the equations and the boundary conditions developed for the Pyke statistics and the Chernoff-Lieberman statistics. The illustrated techniques can serve the numerical analyst wishing to write computer programs. Also the formulae produced might reveal some patterns that could be helpful in analytic work.

The technique involves the selection of intervals of variate values and the computation of probabilities, starting with n =1 and then increasing sequentially in n prior to advancing to the next larger interval of variate values.

The Pyke Statistic

There are some important editorial omissions in the support of the Durbin difference equation [D-10] and subsequent derivates [D-19] and [D-23]. Let us expose them by using these equations to verify the small n distributions already established. Let us begin with restating the notation and equations

$P\{C_n \leq \frac{a}{n+1}\} = p_n(a, a+1, 1) = \frac{n!}{(n+1)^n} q_n(a)$,

where $q_n(a)$ is a convenient contraction. Durbin's equation [D-] may be written as

$0 = \sum_{j=0}^{[2a+1]} (-1)^j \frac{(2a+1-j)^j}{j!} q_{r-j}(a)$; r = 2[a] + 1, 2[a]+2, ⋯,

which we will call the main difference equation (MDE), and his boundary equations [21],

$q_r(a) = 0$, r = – 1, etc.

$= \frac{(r+1)^r}{r!}$ r = 0, ⋯, $[a]$

$= \frac{(r+1)^r}{r!} - 2(a+1)\sum_{j=0}^{[r-a]}((a+1+j)^{j-1}/\,j!)\cdot(r-a-j)^{r-j}/(r-j)!$

r = [a+1], ⋯, [2a].

The first interval of variate values is $0 \leq a < ½$, [a] = 0, [2a+1] = 1, and $\frac{n!}{(n+1)^n}$ is the multiplier that yields p_n when applied to q.

For n = 1we have $p_1 = q_1$ and a = x.

$q_1(a) = 2aq_0(a) = 2a$, $p_1(a) = a$, a = 2x, and $P\{C_1 \leq x\} = 2x$, for $0 \leq x < ½$; the Uniform [0, ½] distribution.

For n = 2 we have $p_2 = \frac{2}{9} q_2$ and a = 3x. Then

$q_2(a) = 2a\, q_1(a) = 4a^2$; $p_2(a) = \frac{8}{9} a^2$ and $P\{C_2 \le x\} = 8x^2$

for $0 \le x < 1/6$.

For n = 3 we have $p_3 = \frac{3}{2^5} q_3$ and a = 4x. Then

$q_3(a) = 2a\, q_2(a) = 8a^3$; $p_3(a) = \frac{3}{2^2} a^3$ and $P\{C_3 \le x\} = 3 \cdot 2^4\, x^3$ for $0 \le x < \frac{1}{8}$.

For n = 4 we have $p_4 = \frac{4!}{5^4} q_4$ and a = 5x. Then

$q_4(a) = 2a\, q_3(a) = 16a^4$; $p_4(a) = \frac{4! \cdot 2^4}{5^4} a^4$ and $P\{C_4 \le x\} = \frac{3}{5} 2^7\, x^4$

for $0 \le x < \frac{1}{10}$.

For n = 5 we have $p_5 = \frac{5!}{6^5} q_5$ and a = 6x. Then

$q_5(a) = 2a\, q_4(a) = 32a^5$; $p_5(a) = \frac{5! \cdot 2^5}{6^5} a^5$ and $P\{C_5 \le x\} = 5! \cdot 2^5\, x^5$

for $0 \le x < \frac{1}{12}$.

Second interval ½ ≤ a < 1, [a]=0, [2a+1] = 2.

For n = 1 we have $p_1(a) = ½\, q_1(a)$ and a = 2x. Then, $q_1(a) = 1$

but stop the development as $P\{C_1 \le x\} = 1$, all x ≥ ½.

For n = 2 we have $p_2 = \frac{2}{9} q_2$ and a = 3x. Then,

$q_2(a) = 2a\, q_1(a) - \frac{(2a-1)^2}{2!} q_0(a) = 4a^2 - (2a^2 - 2a + ½)$

$= 2a^2 + 2a - ½$ using the boundary rule for q_0

$p_2(a) = \frac{2}{9} q_2(a)$; $P\{C_2 \le x\} = 4x^2 + \frac{4}{3}x - \frac{1}{2}$ for $\frac{1}{6} \le x < \frac{1}{3}$

For n = 3 we have $p_3(a) = \frac{3}{2^5} q_3(a)$ and a = 4x. Then

$q_3(a) = 2aq_2(a) - \frac{(2a-1)^2}{2!} q_1(a) = 4a^3 + 2\, a^2 + a - ½$;

$p_3(a) = \frac{4 \cdot 3}{2^5} a^3 + \frac{3!}{2^5} a^2 + \frac{3}{2^5} a - \frac{3}{2^6}$;

$P\{C_3 \le x\} = 3 \cdot 2^2 x^3 + 3\, x^2 + \frac{3}{2^3} x - \frac{3}{2^6}$ for $\frac{1}{8} \le x < \frac{1}{4}$.

For n = 4 we have $p_4(a) = \frac{4!}{5^4} q_4(a)$ and a = 5x. Then

$q_4(a) = 2aq_3(a) - \frac{(2a-1)^2}{2!} q_2(a) = 2a(4a^3 + 2\, a^2 + a - ½)$

$$- \frac{(2a-1)^2}{2!}(2a^2 + 2a - ½);\ p_3(a) = \frac{4\cdot 4!}{5^4}a^4 + \frac{4\cdot 4!}{5^4}a^3 - \frac{3!\cdot 4!}{5^4}a^2 - \frac{4\cdot 4!}{5^4}a + \frac{3!}{5^4}$$

$$P\{C_4 \le x\} = 4\cdot 4!\, x^4 + \frac{4\cdot 4!}{5}x^3 + \frac{3!\cdot 4!}{5^2}x^2 - \frac{4\cdot 4!}{5^2}x + \frac{3!}{5^4} \quad \text{for } \frac{1}{10} \le x < \frac{1}{5}.$$

For n = 5 we have $p_5(a) = \frac{5!}{6^5}q_5(a)$ and a = 6x. Then

$$q_5(a) = 2aq_4(a) - \frac{(2a-1)^2}{2!}q_3(a) = 4a^2(4a^3 + 2a^2 + a - ½)$$

$$- (2a^2 - 2a + \frac{1}{4})(4a^3 + 2a^2 + a - ½);$$

$$p_5(a) = \frac{5!\cdot 2^3}{6^5}a^5 + \frac{12\cdot 5!}{6^5}a^4 + \frac{5\cdot 5!}{6^5}a^3 - \frac{5!}{2\cdot 6^5}a^2 - \frac{5\cdot 5!}{4\cdot 6^5}a - \frac{5!}{8\cdot 6^5};$$

$$P\{C_5 \le x\} = p_5(6x) \text{ for } \frac{1}{6} \le x < \frac{1}{4}.$$

Third interval $1 \le a < \frac{3}{2}$, [a]=1, [2a + 1] = 3.

For n = 1 we have $p_1 = ½\, q_1$ and a = 2x. But

stop because $P\{C_1 \le x\} = 1$ for all $x \ge ½$.

For n = 2 we have $p_2 = \frac{2}{9}q_2$ and a = 3x. But, since the MDE can be used only for n = 3 and higher, the case of n = 2 must be managed from the boundary equations,

$$q_2(a) = \frac{3^2}{2!} - 2(a+1)\frac{1}{a+1}\cdot\frac{(2-a)^2}{2!} = -a^2 + 4a + ½;\ p_2(a) = 1 - \frac{2}{9}(2-a)^2;$$

$$P\{C_2 \le x\} = -2x^2 + \frac{8}{3}x + \frac{1}{9} \text{ for } \frac{1}{3} \le x < ½$$

For n = 3 we have $p_3 = \frac{3!}{4^3}q_3$ and a = 4x. Then, using the MDE,

$$q_3(a) = 2a\, q_2(a) - \frac{(2a-1)^2}{2!}q_1(a) + 2(a-1)^2\, q_0(a)$$

$$= -2a^3 + 8a^2 + a + (-2a^2 + 2a - ½)2 + 2(a^2 - 2a + 1)$$

$$= -2a^3 + 6a^2 + a + 1;\ p_3(a) = -\frac{3}{2^4}a^3 + \frac{3^2}{2^3}a^2 + \frac{3}{8}a + \frac{3}{8};$$

$$P\{C_3 \le x\} = -3\cdot 4\ x^3 + 9\cdot 2\, x^2 + \frac{3}{2}x + \frac{3}{8} \text{ for } \frac{1}{4} \le x < \frac{3}{8}$$

For n = 4 we have $p_4 = \frac{4!}{5^4}q_4$ and a = 5x. Then, using MDE,

$$q_4(a) = 2aq_3(a) - \frac{(2a-1)^2}{2!}q_2(a) + 2(a-1)^2\, q_1(a) \text{ and } q_1 = 2;$$

$$p_4(a) = \frac{4!}{5^4}q_4(a);\ P\{C_4 \le x\} = p_4(5x) \text{ for } \frac{1}{5} \le x < \frac{3}{10}.$$

For n = 5 we have $p_5 = \frac{5!}{6^5}q_5$ and a = 6x. Then, using MDE,

$q_5(a) = 2aq_4(a) - \frac{(2a-1)^2}{2!} q_3(a) + 4(a-1)^2$; $p_5(a) = \frac{5!}{6^5} q_5(a)$;

$P\{C_5 \le x\} = p_5(6x)$ for $\frac{1}{6} \le x < \frac{1}{4}$.

The remaining intervals progress as follows:

$\frac{3}{2} \le a < 2$; $2 \le a < \frac{5}{2}$; etc.

Hopefully, the algorithmic pattern is established. One must be alert to the fact that the maximum value for C_n is $\frac{n}{n+1}$ and this occurs for a = n.

The Chernoff-Lieberman Statistics and D_n

Restating the notation and the new equations for MDE and boundary:

Since $CL_n \sim D_n - 1/2n$, it is convenient to deal with D_n

$P\{D_n \le \frac{a}{n}\} = p_n(a, a, 0) = \frac{n!}{n^n} q_n(a)$

and $q_n(a)$ is short for $q_n(a, a, 0)$. Durbin's MDE equation [D-23] is written as

$0 = \sum_{j=0}^{[2a]} (-1)^j \frac{(2a-j)^j}{j!} q_{r-j}(a)$; r = 2[a] + 1, 2[a] + 2, ⋯,

and the boundary equations [D-24]

$q_r(a) = 1$ r = 0

$= \frac{r^r}{r!}$ r = 1, ⋯, [a]

$= \frac{r^r}{r!} - 2a \sum_{j=0}^{[r-a]} \frac{(a+j)^{j-1}}{j!} \cdot \frac{(r-a-j)^{r-j}}{(r-j)!}$ r=[a+1], ⋯, 2[a]

It is not necessary to deal with the two differing uses of the notation $q_r(a)$ until comparisons between C_n and CL_n begin for a common variate value x. However, the translation by one-half is a nuisance that requires attention.

Technique

$P\{CL_n \le \frac{a_0}{n}\} = P\{D_n \le \frac{a}{n}\} = p_n(a, a, 0)$ and $a = a_0 + ½$

First interval $0 \le a_0 < ½$, $½ \le a < 1$, [a] = 0, [2a] = 1;

For n = 1: $p_1 = q_1$; $a_0 = x$. Then

$q_1(a) = (2a - 1)q_0(a)$; $p_1(a) = (2a - 1)$; $P\{CL_1 \le x\} = 2x$ for $0 \le x < ½$.

For n = 2: $p_2 = \frac{2!}{2^2} q_2$; $a_0 = 2x$. Then,

$q_2(a) = (2a - 1)q_1(a) = (2a - 1)^2 = 4a_0^2$; $p_2(a_0) = 2a_0^2$;

$P\{CL_2 \le x\} = 8x^2$ for $0 \le x < \frac{1}{4}$.

For n = 3: $p_3 = \frac{2}{9} q_3$; $a_0 = 3x$. Then

$q_3(a) = (2a - 1)q_2(a) = (2a - 1)^3 = 8a_0^3$; $p_3 = \frac{2}{9} 8a_0^3$;

$P\{CL_3 \le x\} = 3 \cdot 2^4 x^3$ for $0 \le x < \frac{1}{6}$.

For n = 4: $p_4 = \frac{3}{2^5} q_4$, $a_0 = 4x$. Then $q_4(a) = (2a - 1)q_3(a) = (2a - 1)^4 = 2^4 a_0^4$; $p_4 = \frac{3}{2} a_0^4$;

$P\{ CL_4 \le x\} = 3 \cdot 2^7 x^4$ for $0 \le x < \frac{1}{8}$.

For n = 5: $p_5 = \frac{4!}{5^4} q_5$, $a_0 = 5x$. Then $q_5(a) = (2a - 1)q_4(a) = 2^5 a_0^5$;

$p_5 = \frac{3 \cdot 2^8}{5^4} a_0^5$; $P\{ CL_5 \le x\} = 3 \cdot 2^8 \cdot 5\, x^5$ for $0 \le x < \frac{1}{10}$.

Second interval $½ \le a_0 < 1$, $1 \le a < \frac{3}{2}$, [a] = 1, [2a] +1 = 3, [a]+1 = 2

For n = 1: $p_1 = q_1$, $a_0 = x$. Then the boundary equations must be used

and stop as $P\{CL_1 \le x\} = 1$, all $x \ge 1$.

For n = 2: $p_2 = \frac{1}{2} q_2$, $a_0 = 2x$, so

$q_2(a) = 2 - 2a \sum_{j=0}^{0} \frac{1}{2a} \cdot \frac{(2-a)^2}{2} = 2 - ½ (\frac{3}{2} - a_0)^2 = 2 - \frac{9}{8} + \frac{3}{2} a_0 - ½ a_0^2$

$= - ½ a_0^2 + \frac{3}{2} a_0 + \frac{7}{8}$; $p_2 = -\frac{1}{4} a_0^2 + \frac{3}{4} a_0 + \frac{7}{16}$

$P\{ CL_2 \le x\} = - x^2 + \frac{3}{2} x + \frac{7}{16}$ for $½ \le x < \frac{3}{4}$

$- 2 x^2 + \frac{3}{2} x + \frac{1}{8}$

$p_2(a_0) = -\frac{1}{4} a_0^2 + \frac{3}{4} a_0 + \frac{7}{8}$; $P\{ CL_2 \le x\} = - 2 x^2 + 3 x + \frac{7}{8}$ for $\frac{1}{4} \le x < \frac{3}{4}$

For n = 2, the expressions have been developed already, but we develop them as $P\{CL_n + \frac{1}{2n} \le \frac{a}{n}\} = p_n(a, a, 0) = \frac{n!}{n^n} q_n(a)$, where $q_n(a)$ is a convenient contraction, and illustrate the use of the difference equation having this translation by one-half feature.

The first interval of variate values is $0 \le a < ½$, [a] = 0, [2a]= 0, and $\frac{n!}{n^n}$ is the multiplier that yields p_n when applied to q_n.

For n = 1, we have p1 = q1 and a = x.

q1(a) = 2aq0(a) = 2a, p1(a) = a, a = 2x, and

$P\{C1 \le x\} = 2x$, for $0 \le x < ½$; the Uniform [0, ½] distribution.

For n = 2, we have p2 = $\frac{2}{9}$ q2 and a = 3x. Then

$$q2(a) = 2a\ q1(a) = 4a^2;\ p2(a) = \frac{8}{9}\ a^2 \text{ and } P\{C2 \leq x\} = 8x^2.$$

LIST OF REFERENCES

Chernoff, H., & Lieberman, J. (1954). The use of normal probability paper. *J. of Amer.. Statist. Assoc*., **49**, 778-785.

Chernoff, H. and J. Lieberman. (1956). The use of generalized probability paper. *Ann. Math. Statist.*,**27**, 805-818.

Dempster, A. P. (1959). Generalized D_n^+ statistics. *Ann. Math. Statist.*, 30, 593.

Durbin, J. (1968). The probability that the sample distribution function lies between two parallel straight lines. *Ann. Math. Statist.*, **39**, 398.

Gibbons, J. D. & Chakraborti, S. (1992). *Nonparametric statistical inference.* New York, NY: Marcel Dekker.

Kemperman, J. H. B. (1961). *The passage problem for a stationary Markov Chain.* Chicago, Il., University of Chicago Press.

Kendall, M. (1961). *A course in the geometry of n dimensions.* London: Griffin & Co.

Kendall, D. and Moran, P. A. P. (1963). *Geometrical probability*. London: Griffin & Co.

Massey, F. J. (1950). A note on the estimation of a distribution by confidence limits. *Ann. Math. Statist.*,16, 116.

Pyke, R. (1959). The supremum and infimum of the Poisson process. *Ann. Math. Statist*., 30, 568.

Read, R. R. (1972). The asymptotic inadmissibility of the sample distribution function. *Ann. Math. Statist.*,43, 85.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, California

4. Richard Mastowski (Technical Editor).....1
   Graduate School of Operational and Information Sciences (GSOIS)
   Naval Postgraduate School
   Monterey, California

5. Sam Buttrey, OR/SB1.....1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA

6. Robert Dell, OR/Chair 2
   Operations Research Department
   Monterey, CA

7. Thomas Hamrick, OR/TH.....1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA

8. Patricia Jacobs.....1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA

9. Robert Koyak, OR/RK.....1
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA

10. David Olwell, SE/OL.....1
    Systems Engineering Department
    Naval Postgraduate School
    Monterey, CA

11. Robert Read, OR/RR 5
Operations Research Department
Naval Postgraduate School
Monterey, CA

12. Lynn Whitaker, OR/WH.....1
Operations Research Department
Naval Postgraduate School
Monterey, CA