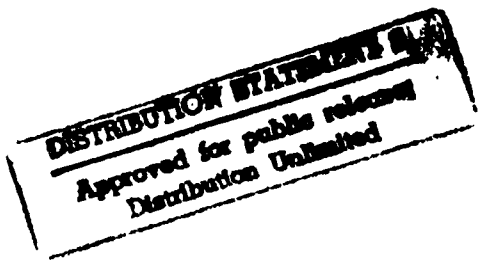
CHANGE *PP* PLOT AND CONTINUOUS SAMPLE QUANTILE FUNCTION

Technical Report # 189

October 1992



Emanuel Parzen

Texas A&M Research Foundation

Project No. 6547

Sponsored by the U. S. Army Research Office

Professor Emanuel Parzen, Principal Investigator

CHANGE *PP* PLOT AND CONTINUOUS SAMPLE QUANTILE FUNCTION

Emanuel Parzen*

Texas A&M University
College Station, Texas 77843

Key Words: *Functional inference; Goodness of fit; Comparison density; Renyi information; Quantile data analysis; Continuous sample quantiles; Maximum spacings parameter estimation; Distinct mid-value probability transform; Change PP Plot.*

ABSTRACT

Functional inference recommends data analysis of a sample of n observations by functional and graphical representations of its probability models using various functions on $0 < u < 1$, including the quantile function. This paper discusses: change PP plots and a continuous version of the sample quantile function which use the mid-distinct values probability integral transform; comparison density functions; comparison interpretation of probability integral transform; maximum spacings method of one sample parameter estimation.

1. My 15th Anniversary of Texas A&M and Functional Statistical Inference

As the Department of Statistics at Texas A&M University celebrates its 30th anniversary in 1992, each of us who are part of the department may want to celebrate our personal anniversaries marking the length and depth of our association. In 1992 I am completing 15 years of happy and deep association (since 1978) with Texas A&M. I would like to thank my colleagues for providing an enjoyable and stimulating environment, and proving that there is great life in College Station (as well as opportunities to travel to help maintain our department's national and international visibility). We can all

* Research supported by U. S. Army Research Office

take pride in 1992 in the fact that the Department of Statistics at Texas A&M has a reputation as one of the outstanding and very active statistics programs.

In 1978 my research emphasis expanded from time series analysis to functional statistical inference. I feel that time series analysis is important not only to provide analysis of time series data, but also to provide a background suitable for new approaches to mainstream statistical problems. What I call *functional inference* recommends data analysis of a sample of n observations by functional (and therefore graphical) representations of its statistical properties (probability models), using various functions on the unit interval $0 < u < 1$. An important function is the quantile function $Q(u)$, $0 < u < 1$, or inverse distribution function. Functional inference (introduced in Parzen (1979)) can be regarded as applying time series theoretical and function smoothing ideas to classical statistics.

What will be the benefits of functional statistical inference to applied statisticians? I believe that they will include (1) unification of statistical methods for discrete and continuous random variables, (2) change analysis, (3) information theory approaches to statistical inference (see Parzen (1989), (1991), (1991), (1992)).

Unification may have the most difficulty arousing interest from applied statisticians; its philosophy is that statistical problems should be solved in several ways (when I ask graduate students what are several ways to solve a problem in a textbook they usually tell me there is only one way!). Change analysis (on which my research interests have bloomed since December 1990) is an extension of changepoint analysis which has as initial goal to determine if a probability model fitted to a whole sample $Y_1, \ldots, Y_n$ fits all subsamples $Y_1, \ldots, Y_m$ for all $m < n$. Information theory is important to statistics because it provides measures of divergence between two probability distribu-

tions.

The research of Eubank, LaRiccia, and Hart (1992) can be considered to be fundamental research on functional statistical inference. The practice of statistics would be enhanced by applying their deep insights about the relations between goodness of fit tests and nonparametric regression.

2. One Sample Probability Model Fitting

A basic problem of statistics is fitting probability models to a sample $Y_1, \ldots, Y_n$ of a continuous random variable Y with true distribution function

$$F_Y(y) = \text{Prob}[Y \leq y],$$

probability density function $f_Y(y) = F'_Y(y)$, quantile function

$$Q_Y(u) = F_Y^{-1}(u).$$

The parametric approach to modeling a random sample assumes a parametric probability model $f_\theta(y)$ indexed by a vector parameter θ with k components θ_j.

Classical statistics assumes suitable regularity conditions on the parametric family of probability densities in order to assure desirable properties of parameter estimators formed by maximum likelihood estimators $\hat{\theta}$. These are defined to maximize the log likelihood

$$L(\theta) = \sum_{t=1}^{n} log f_\theta(Y_t),$$

and are usually computed as solutions of the estimating equations

$$\sum_{t=1}^{n} S_{\theta_j}(Y_t; \theta) = 0, j = 1, \ldots, k,$$

where $S_{\theta_j}(Y; \theta)$ is the partial derivative of $\log f_\theta(Y)$ with respect to the component θ_j of the k-dimensional vector parameter θ.

Goodness of fit of the parametric model to the data is tested by forming the transformed data

$$W_t = F_{\theta^*}(Y_t)$$

and testing whether the discrete distribution function, denoted $F_W^{(n)}(w)$, $0 < w < 1$, of $W_1, \ldots, W_n$ is significantly different from the distribution function of U, a Uniform[0,1] random variable. Functional inference converts this question into a problem about the detection of signal in noise in data that is a process on $0 < w < 1$. Under the null hypothesis that the parametric model fits, the limit distribution of $n^{.5}(F_W^{(n)}(w) - w)$ is a Brownian Bridge $B(w)$, $0 < w < 1$, modified for the effect of parameter estimation (Shorack and Wellner (1986)).

Goal One of this paper is to define continuous versions of sample distribution functions that help provide tests of goodness of fit which provide non-parametric estimators of f_Y when a parametric model does not fit.

3. Comparison and information divergence of probability distributions

Goal Two of this paper is to raise statisticians' consciousness about the concept of *comparison density* function $d(u; F, G)$, $0 < u < 1$, of two distributions F and G (see Parzen (1992)).

For F and G continuous, we define the comparison distribution

$$D(u; F, G) = G(F^{-1}), 0 < u < 1,$$

and comparison density

$$d(u : F, G) = g(F^{-1}(u))/f(F^{-1}(u)), 0 < u < 1.$$

The graph of $D(u : F, G)$ is called a PP-plot because it is a plot of $(F(y), G(y))$ which compares the P values of an observation y under the two distributions.

For F and G discrete with respective probability mass functions p_F and p_G we first define the comparison density

$$d(u : F, G) = p_G(F^{-1}(u))/p_F(F^{-1}(u)), 0 < u < 1.$$

The comparison distribution is then defined as the integral

$$D(u : F, G) = \int_0^u d(t; F, G)dt, 0 < u < 1.$$

An equivalent way to describe $D(u; F, G)$ in the discrete case is $D(u; F, G) = G(F^{-1}(u))$ at F-exact u satisfying $F(F^{-1}(u)) = u$ and $D(u; F, G)$ is defined at other values of u by linear interpolation between its values at F-exact values of u. When a PP-plot $D(u; F, G)$ is determined by linear interpolation we call the values determining the plot the PP-plot values; thus the PP-plot values are $G(F^{-1}(u))$ for all values of u that are F-exact.

Goal Three of this paper is to recommend Change PP plots which plot $(F(y), G(y) - F(y))$ or equivalently $(u, D(u; F, G) - u)$.

Implicit in our definitions are assumptions that guarantee that $D(0; F, G) = 0$, $D(1; F, G) = 1$. Therefore $d(u; F, G)$ is a density, a non-negative function integrating to 1.

Goal Four of this paper is to remind statisticians that very useful measures of divergence of $D(u)$ from u are Renyi information measures (Renyi (1961)) of the divergence of $d(u)$ from 1. They provide "entropy detectors" to be used in addition to "non-linear detectors" and "linear detectors" which use norms of $D(u) - u$.

For a density $d(u)$, $0 < u < 1$, *Renyi information* of index λ, is defined for λ not equal to 0 or -1 by

$$IR_\lambda(d) = (2/\lambda(1 + \lambda)) \log \int_0^1 d(u)^{1+\lambda} du$$

For λ equal to 0 or -1, define

$$IR_0(d) = 2\int_0^1 (d(u)\log d(u))du$$

$$IR_{-1}(d) = -2\int_0^1 \log d(u))du$$

Hellinger information corresponds to $\lambda = -.5$;

$$IR_{-.5}(d) = -8\log\int_0^1 d(u)^{.5}du.$$

A very useful identity is

$$IR_\lambda(d(u;F,G)) = IR_{-(1+\lambda)}(d(u;G,F))$$

4. Comparison interpretation of probability integral transform

An important application of comparison concepts is to interpret explicit formulas for the true distribution and true quantile function of the probability integral transform $W = F_\theta(Y)$ assuming Y is continuous and the parametric model is continuous. One can show that

$$Q_W(u) = F_\theta(Q_Y(u)) = D(u;F_Y,F_\theta),$$

$$F_W(w) = F_Y(Q_\theta(w)) = D(w;F_\theta,F_Y)$$

Goal Five of this paper is to recommend that the divergence (comparison) between two distribution measures F_Y and F_θ be measured by the divergence from $D_0(u) = u$ of the comparison distribution functions $D(u;F_Y,F_\theta)$ or $D(u;F_\theta,F_Y)$.

To illustrate the different roles played by the two possible comparison distribution functions we note that (1) for estimation of the parameter one chooses $\hat{\theta}$ as the value of θ making the quantile function $D(u;F_Y,F_\theta)$ close to u, while (2) for goodness of fit, one tests if the distribution function $D(u;F_{\hat{\theta}},F_Y)$ is close to u.

5. Sample quantile functions

Goal Six of this paper is emphasize to applied statisticians our opinion that the first step in data analysis should be to form the sample quantile function $Q^{(n)}(u)$, $0 < u < 1$, which is the inverse of the sample distribution function $F^{(n)}(y)$, $-\infty < y < \infty$. To compute it one determines u_j, v_j for $j = 1, \ldots, c$, where (1) the distinct values in the sample are denoted v_j, $j = 1, \ldots, c$, and (2) the cumulative relative frequencies are denoted

$$u_j = F^{(n)}(v_j) = \text{ fraction of sample } \leq v_j.$$

Note $u_c = 1$; define $u_0 = 0$. If all values in the sample are distinct, c=n and the distinct values are the order statistics $Y(1;n) < \ldots < Y(n;n)$.

The sample quantile function $Q^{(n)}$, the inverse of the sample distribution function $F^{(n)}$, can be calculated by

$$Q^{(n)}(u) = v_j, \; u_{j-1} < u \leq u_j,$$

or equivalently it is piecewise constant left continuous satisfying

$$Q^{(n)}(u_j) = v_j, \; j = 1, \ldots, c.$$

The sample median and quartiles are defined to be the values at $u = .5, .25, .75$ of $Q^{(n)}(u)$.

A nonparametric measure of location is the sample median. A nonparametric measure of scale is the quartile deviation

$$QD^{(n)} = 2IQR^{(n)},$$

defined as twice the interquartile range

$$IQR^{(n)} = Q^{(n)}(.75) - Q^{(n)}(.25).$$

An important characteristic of a distribution is its behavior at the tails or ends of the distribution; we like to joke that "in statistics the ends do justify

the means" (to adaptively efficiently estimate location parameters one must estimate tail shape parameters). Our experience is that tail behavior can be judged (in a quick and dirty "back of an envelope" way) from the values near 0 and 1 of the identification quantile function (Parzen (1983))

$$QI^{(n)}(u) = (Q^{(n)}(u) - Q^{(n)}(.5))/QD^{(n)}.$$

Intuitively, the identification quantile function is normalized to have at $u = .5$ value 0 and slope approximately 1.

Goal Seven of this paper is to remind statisticians that there is an extensive literature on the important question of whether one should use a nonparametrically smoothed sample quantile function $Q^{\wedge(n)}(u)$ rather than the raw sample quantile function $Q^{(n)}(u)$ at the initial stage of analysis. A comprehensive survey and exhaustive analysis of properties of smoothed sample quantile functions is given in the outstanding Ph.D. thesis of Cheng Cheng (1993). In this paper I discuss my proposal for a quick and dirty smoothing provided by a continuous version of the sample quantile function.

6. Continuous versions of sample quantile and distribution functions, and Change *PP* Plots

The sample distribution function $F^{(n)}$ of data is discrete. Goal Eight of this paper is to propose that to estimate a continuous distribution function we first form a continuous version $F^{c(n)}$ as follows. Define mid-values v_j^c, $j = 1, \ldots, c-1$, by

$$v_j^c = .5(v_j + v_{j+1}).$$

We do not propose a universal definition of v_0^c and v_c^c. Initially we define $v_0^c = v_1$, $v_c^c = v_c$.

Define $F^{c(n)}$ and $Q^{c(n)}$ to be piecewise linear between its values (for

$j = 0, \ldots, c$)

$$Q^{c(n)}(u_j) = v_j^c$$

$$F^{c(n)}(v_j^c) = u_j$$

Goal Nine of this paper is to note that our proposed continuous version may be regarded as related to another important modification of a discrete distribution (which we call the mid-distribution) that is being increasingly recognized as the way to express P-levels of significance tests (see Routledge (1992), Upton (1992)). The mid-distribution function of the sample is defined by

$$F^{mid(n)}(y) = F^{(n)}(y) - .5p^{(n)}(y)$$

where $p^{(n)}(y)$ is the fraction of the sample equal to y. One expects that approximately

$$F^{c(n)}(v_j) = F^{mid(n)}(v_j) = (u_j + u_{j-1})/2,$$

$$Q^{c(n)}((u_j + u_{j-1})/2) = v_j$$

Goal Ten of this paper is to propose that the problem of goodness of fit and parameter estimation of the parametric model F_{θ}. be treated as one of comparing with the Uniform[0,1] distribution $D_0(u) = u$ the continuous comparison distribution functions defined in terms of the PP-plot values u_j and

$$w_j(\theta) = F_{\theta}\left(Q^{c(n)}(u_j)\right) = F_{\theta}(v_j^c)$$

which we call the "*distinct mid-value probability transform*" since for $j = 1, \ldots, c-1$

$$w_j(\theta) = F_{\theta}((v_j + v_{j+1})/2).$$

Define the quantile-type PP-plot $D^c(u; F^{(n)}, F_{\theta})$ as piecewise linear connecting

$$(0,0), (u_1, w_1(\theta)), \ldots, (u_{c-1}, w_{c-1}(\theta)), (1,1).$$

Define the distribution-type PP plot $D^c(u; F_\theta, F^{(n)})$ as piecewise linear connecting

$$(0,0),(w_1(\theta),u_1),\ldots,(w_{c-1}(\theta),u_{c-1}),(1,1).$$

In practice we recommend plotting Change PP plots of $(u_j, w_j(\theta)-u_j)$ and comparison densities $d^c(u; F^{(n)}, F_\theta)$ and $d^c(u; F_\theta, F^{(n)})$.

7. Maximum Spacings method of one sample parameter estimation

Regular maximum likelihood estimators $\hat{\theta}$ are parameter values minimizing the negative of the average log likelihood

$$-L(\theta) = (1/n)\sum_{t=1}^{n} -\log f_\theta(Y(j;n))$$

$$= \sum_{j=1}^{c}(u_j - u_{j-1})(-\log f_\theta(Q^{(n)}(u_j))$$

A maximun spacings estimator, also denoted $\hat{\theta}$, minimizes

$$-2\sum_{j=1}^{c}(u_j - u_{j-1})\log(F_\theta(Q^{c(n)}(u_j)) - F_\theta(Q^{c(n)}(u_{j-1})))/(u_j - u_{j-1}))$$

Maximum spacings estimators have been discussed by Cheng and Iles (1987), Ranneby (1984), Cheng and Amin (1983), Titterington (1985); they could be called Maximum Grouped Likelihood estimators. Maximum spacings estimator can be shown to provide credible estimators in non-regular cases (where likelihood is unbounded and thus maximum likelihood does not provide a satisfactory estimator) and to provide efficient estimators in regular cases (they have the same properties as maximum likelihood estimators).

Goal Eleven of this paper is to note that (1) maximum spacings estimators can be represented in terms of comparison density functions whose neg-entropy is minimized to find parameter estimators:

$$2\int_0^1 (-\log d^c(u; F^{(n)}, F_\theta))du,$$

and (2) adapting Beran (1977) one can obtain robust parameter estimators by combining maximum spacings estimations with minimum information estimation criteria.

We propose for investigation minimum information estimators (more precisely, minimum Renyi information of index λ estimators), denoted $\hat{\theta}^{(\lambda)}$, defined to minimize

$$IR_\lambda(d^c(u; F^{(n)}, F_\theta)).$$

Minimum information estimators satisfy the estimating equations

$$\int_0^1 (d^c(u; F^{(n)}, F_\theta))^{1+\lambda} S_{\theta_j}(Q^{c(n)}(u), \theta) du = 0.$$

Regular maximum likelihood estimators correspond to $\lambda = -1$. Minimum information estimators of index λ are of interest because they provide ***robust estimators*** in the presence in the data of values not fitting the assumed parametric probability model (see Beran (1977)). To test if they should be computed in preference to regular maximum likelihood estimators one could test if the latter satisfy the estimating equations of the former. Research on these ideas is continuing.

8. Example of Change *PP* Analysis

The introductory statistics textbook by Friedman et al (1978) discusses a data set consisting of 100 measurements made at the National Bureau of Standards on the weight of NB 10. It is very interesting because it appears to follow a normal distribution with outliers. One can obtain this conclusion by an exploratory analysis (described below) or by robust estimation of parameters of a normal distribution using minimum information estimators.

Each measurement in the sample is the number of micrograms below ten grams. The sample standard deviaton is approximately 6 micrograms (the maximum likelihood estimator). But a normal distribution with parameters

equal to the sample mean and standard deviation does not pass tests of goodness of fit to the data. A trimmed sample (trimmed to omit the smallest and largest values) has standard deviation of approximately 4 microgram (the robust estimator) and is fit by a normal distribution.

To compare whole sample with a probability model, we first compute sample quantile function $Q^{(n)}$ of data normalized by subtracting sample mean and dividing by sample standard deviation. Figure 1 compares $Q^{(n)}$ with Φ^{-1}, the standard normal quantile function; we intuitively perceive that their slopes at $u = .5$ differ, indicating that the true scale parameter of the data is not well estimated by the sample standard deviation. Figure 2 compares to the standard normal the sample quantile function of the normalized trimmed sample; we perceive a fit.

Next we compute (what we have denoted by w_j) the mid distinct values of the normalized samples transformed by the standard normal distribution. We compare w_j to the sample cumulative frequencies u_j. A PP plot graphs the linear interpolation of (u_j, w_j). A Change PP plot graphs the linear interpolation of $(u_j, n^{.5}(w_j - u_j))$. Under the null hypothesis of goodness fit, the Change PP plot should be a sample path of a Brownian Bridge process, modified by the effect of parameter estimation. The asymptotic 95% significance level (found by simulation) is .97.

The Change PP plot of the whole sample in Figure 3 indicates lack of fit because of its maximum (which is 1.12) and its shape (which can be interpreted by an experienced analyst as a canonical shape indicating that the probability integral transformed data has a probability density whose graph looks like a bowl, implying outliers in the original data). This conclusion is reached with a minimum of computation; it would also be reached by a computer intensive density estimation analysis of the PP plot.

The Change PP plot of the trimmed sample in Figure 4 indicates fit (of

the trimmed sample by the normal with parameters equal to the maximum likelihood estimators from the trimmed sample) because of its maximum and its shape (which can be interpreted as a canonical shape whose derivative is a constant function, indicating the transformed data has a uniform distribution).

References

Beran, R. J. (1977). "Minimum Hellinger Distance Estimates for Parametric Models." *Annals of Statistics*, 5, 445–463.

Cheng, Cheng (1993). "Smoothing of Quantile Functions," Ph.D. Dissertation, Texas A&M University.

Cheng, R. C. H. and Amin, N. A. K. (1983). "Estimating Parameters in Continuous Univariate Distributions with a Shifted Origin." *J. R. Statist. Soc. B*, 45, pp. 394–403.

Cheng, R. C. H. and Iles, T. C. (1987). "Corrected Maximum Likelihood in Non-regular Problems," *J. R. Statist*, Soc. B, 49, pp. 95–101.

Eubank, R. L., Hart, J. D. and LaRiccia, V. N. (1992). "Testing Goodness-of-Fit Via Nonparametric Function Estimation Techniques," *Communications in Statistics.*

Freedman, D., Pisani, R. and Purves,R. (1978). *Statistics.* New York: W. W. Norton & Co., Inc.

Parzen, E. (1979). "Nonparametric Statistical Data Modeling", *Journal of the American Statistical Association*, (with discussion), 74. 105–131.

Parzen, E. (1983). "Informative Quantile Functions and Identification of Probability Distribution Types," *Proceedings of the 29th Army Conference on the Design of Experiments*, 97–107.

Parzen, E. (1989). "Multi-Sample Functional Statistical Data Analysis," *Staistical Data Analysis and Inference Conference in Honor of C. R. Rao*, ed. Y. Dodge, Amsterdam: Elsevier, 71–84.

Parzen, E. (1991). "Goodness of Fit Tests and Entropy,' *Journal of Combinatorics, Information, and System Science*, 16, 129–136.

Parzen, E. (1991). "Unification of Statistical Methods for Continuous and Discrete Data," *Proceedings Computer Science-Statistics INTERFACE '90*, (ed. C. Page and R. LePage), New York: Springer Verlag, 235–242.

Parzen, E. (1992). "Comparison Change Analysis." *Nonparametric Statistics and Related Topics* (ed. A. K. Saleh), Amsterdam: Elsevier, 3-15.

Parzen, E. (1993). "From Comparison Density to Two Sample Data Analysis," *The Frontiers of Statistial Modeling: An Informational Approach*, ed. H. Bozdogan, Amsterdam: Kluwers.

Ranneby, B. (1984). "The Maximum Spacing Method, An Estimation Method Related to the Maximum Likelihood Method," *Scand. J. Statist.*, 11, 93–112.

Renyi, A. (1961). "On measures of entropy and information." *Proc. 4th Berkeley Symp. Math. Statist. Probability, 1960*, 1, 547–561. Berkeley: University of California Press.

Routledge, R. D. (1992). "Resolving the Conflict Over Fisher's Exact Test." *The Canadian Journal of Statistics*, Vol. 20, pp. 201–209.

Shorack, Galen R. and Wellner, Jon A. (1986). *Empirical Processes With Applications to Statistics*. New York: John Wiley & Sons.

Titterington, D. M. (1985). "Comment on 'Estimating Parameters in Continuous Univariate Distributions' ". *J. R. Statist. Soc. B*, 47, No. 1, pp. 115–116.

Upton, Graham J. G. (1992). "Fisher's Exact Test." *J. R. Statist. Soc. A*, 155, pp. 395–402.

FIGURE 1

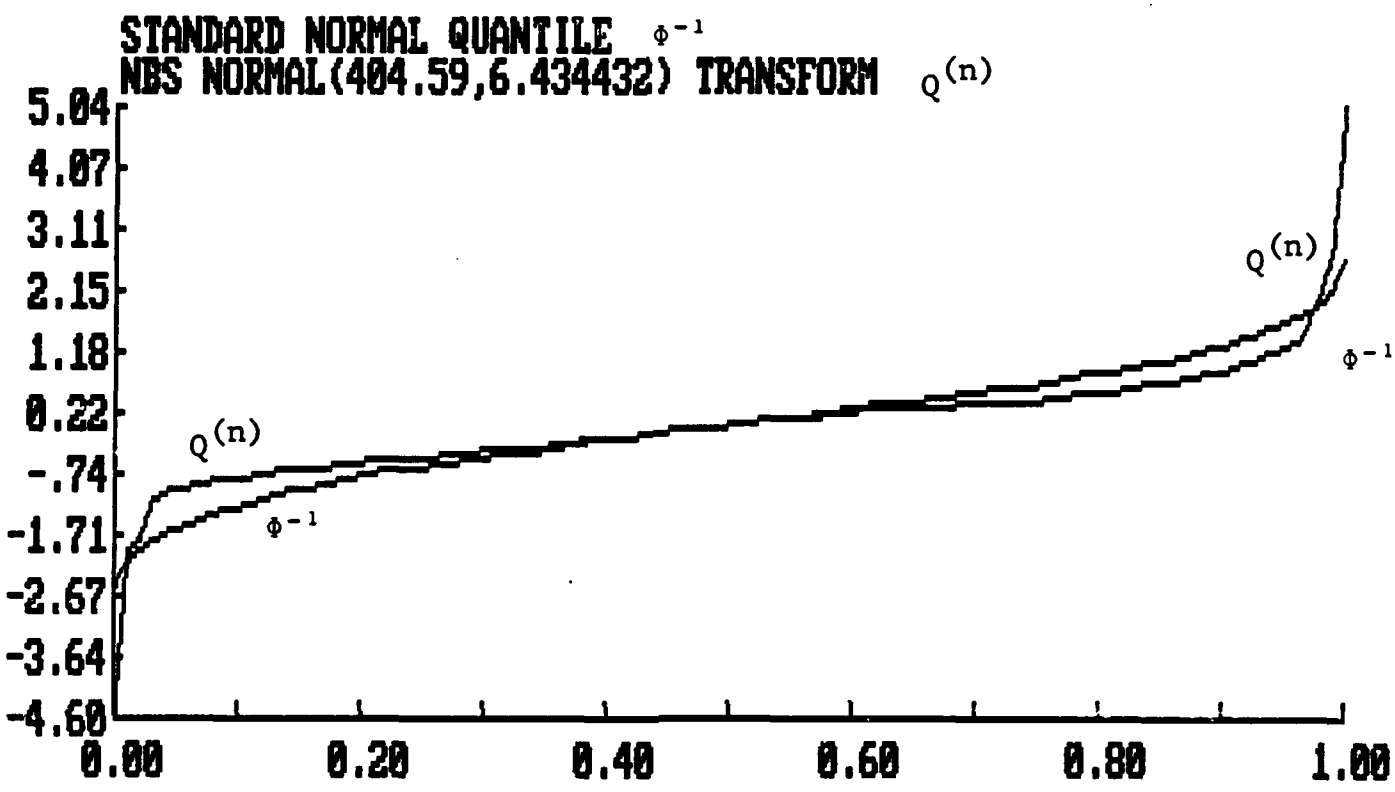

STANDARD NORMAL QUANTILE Φ^{-1}
NBS NORMAL(404.59,6.434432) TRANSFORM $Q^{(n)}$
5.04
4.07
3.11
2.15
1.18
0.22
-.74
-1.71
-2.67
-3.64
-4.60
0.00
0.20
0.40
0.60
0.80
1.00
$Q^{(n)}$
Φ^{-1}
$Q^{(n)}$
Φ^{-1}

FIGURE 2

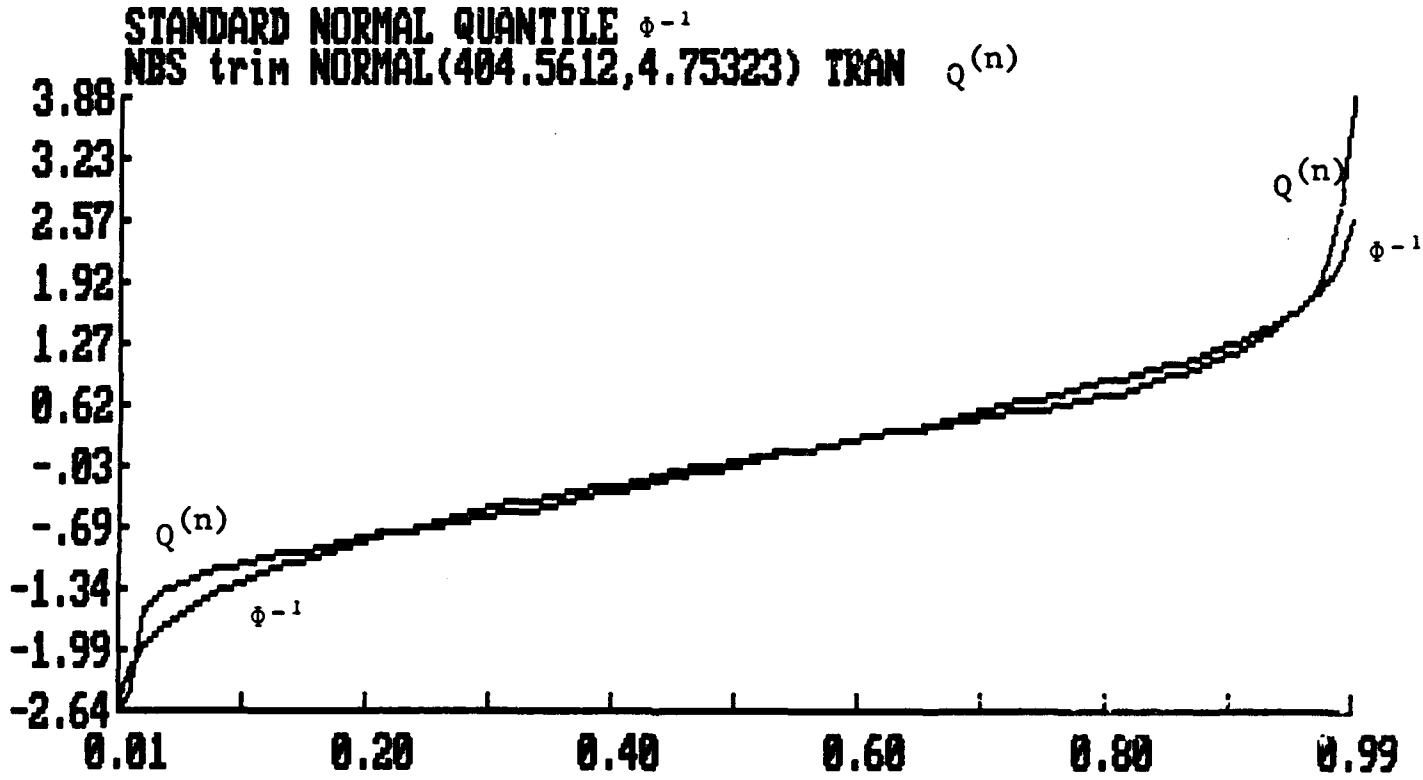

STANDARD NORMAL QUANTILE Φ^{-1}
NBS trim NORMAL(404.5612,4.75323) TRAN $Q^{(n)}$
3.88
3.23
2.57
1.92
1.27
0.62
-.03
-.69
-1.34
-1.99
-2.64
0.01
0.20
0.40
0.60
0.80
0.99
$Q^{(n)}$
Φ^{-1}
$Q^{(n)}$
Φ^{-1}

FIGURE 3

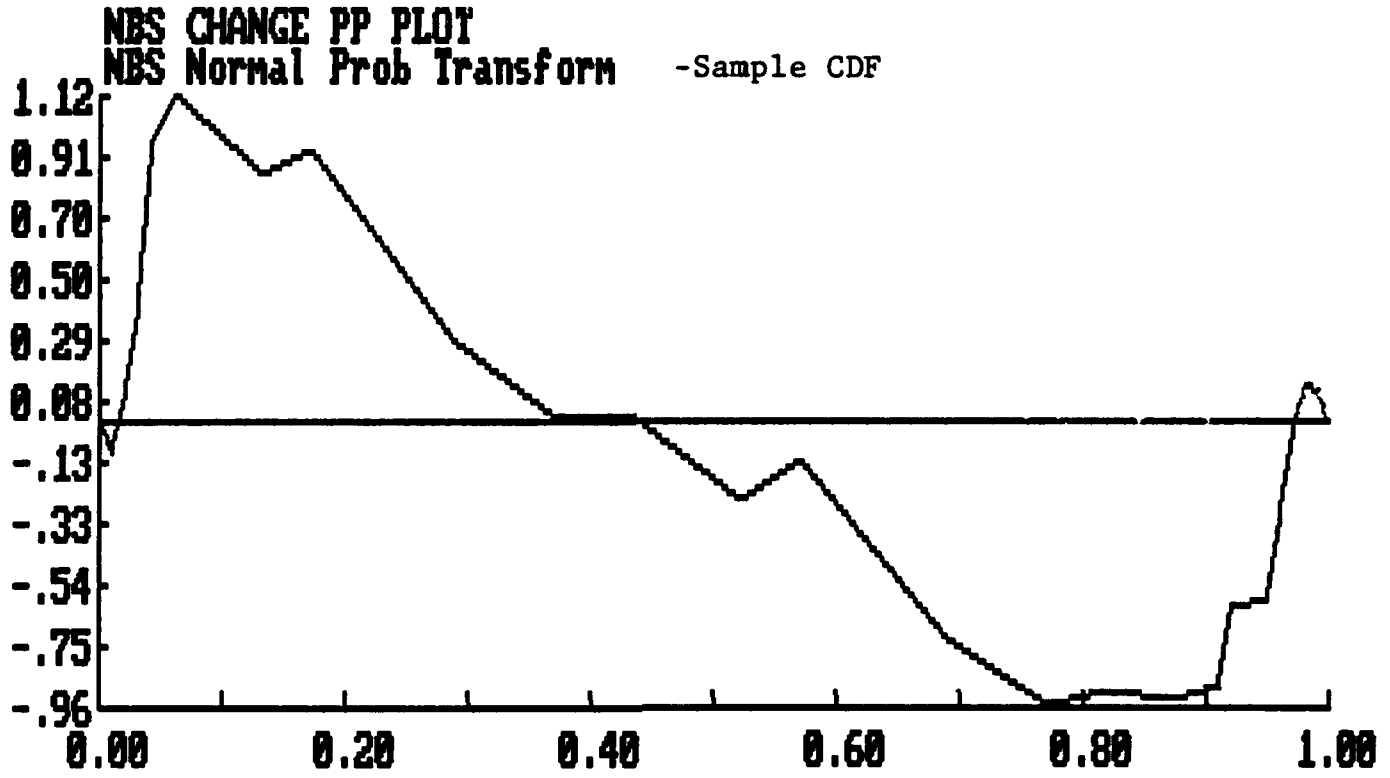
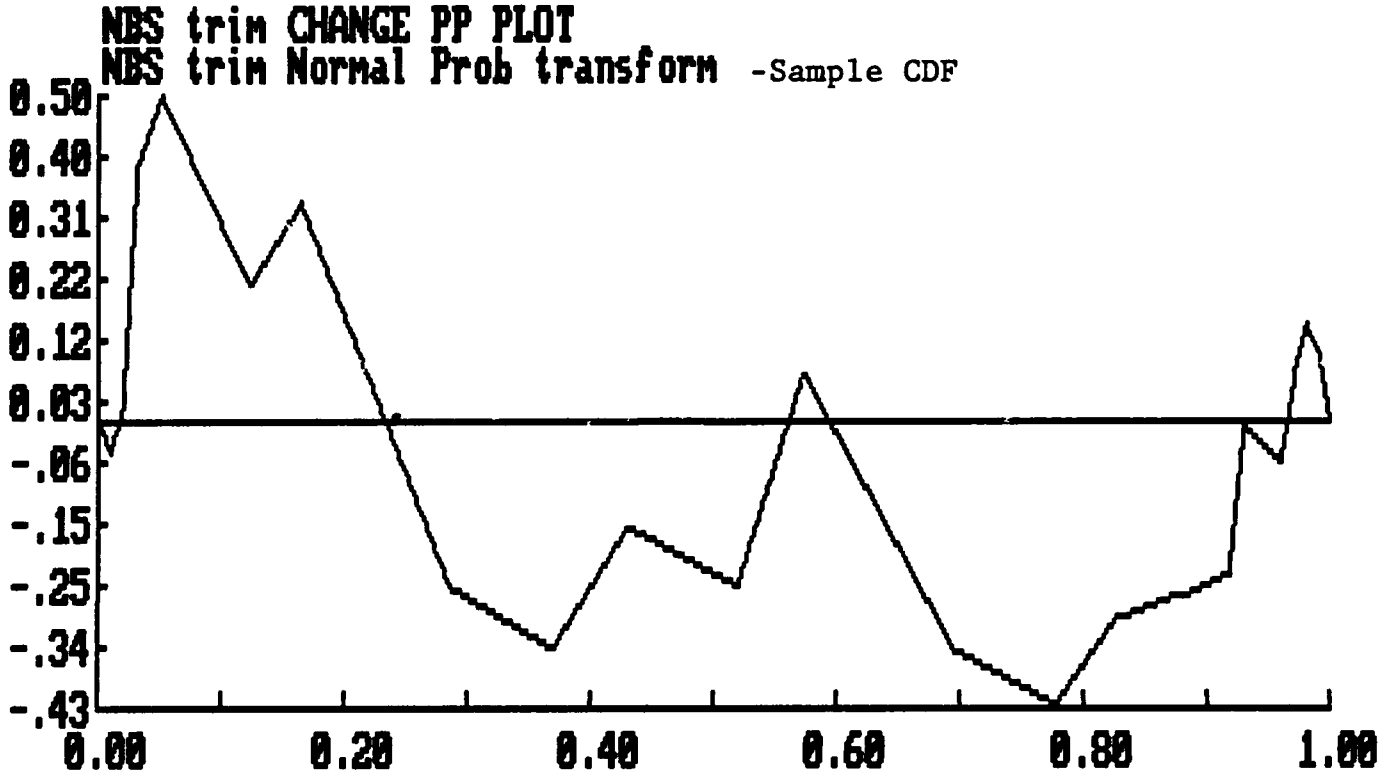

NBS CHANGE PP PLOT
NBS Normal Prob Transform
-Sample CDF
1.12
0.91
0.70
0.50
0.29
0.08
-.13
-.33
-.54
-.75
-.96
0.00
0.20
0.40
0.60
0.80
1.00

FIGURE 4


NBS trim CHANGE PP PLOT
NBS trim Normal Prob transform -Sample CDF
0.50
0.40
0.31
0.22
0.12
0.03
-.06
-.15
-.25
-.34
-.43
0.00
0.20
0.40
0.60
0.80
1.00

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Octoher 1992	

4. TITLE AND SUBTITLE
Change PP Plot and Continuous Sample Qauntile Function

5. FUNDING NUMBERS
DAAL03-90-G-0069

6. AUTHOR(S)
Emanuel Parzen

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

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

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Functional inference recommends data analysis of a sample of n observations by functional and graphical representations of its probability models using various functions on $0 < u < 1$, including the quantile function. This paper discusses: change PP plots and a continuous version of the sample quantile function which use the mid-distinct values probability integral transform; comparison density functions; comparison interpretation of probability integral transform; maximum spacings method of one sample parameter estimation.

14. SUBJECT TERMS
Functional inference; Goodness of fit; Comparison density; Renyi information; Quantile data analysis; Continuous sample quantiles; Maximum spacings parameter estimation; Distinct mid-value probability transform; Change PP plot

15. NUMBER OF PAGES
19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL