(4)

COMPUTER PROBLEMS IN GOODNESS-OF-FIT

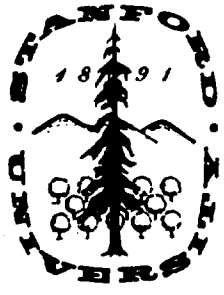
BY

MICHAEL A. STEPHENS

TECHNICAL REPORT NO. 420
AUGUST 10, 1989

PREPARED UNDER CONTRACT
N00014-89-J-1627 (NR-042-267)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA



COMPUTER PROBLEMS IN GOODNESS-OF-FIT

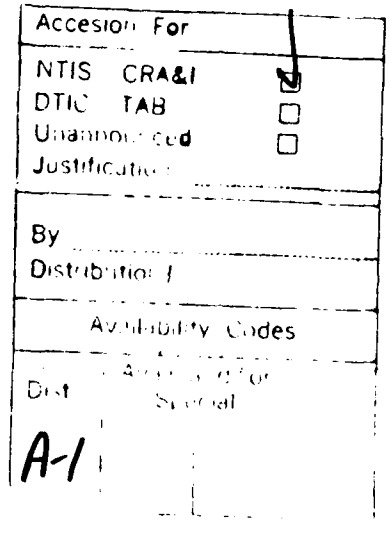
BY

MICHAEL A. STEPHENS

TECHNICAL REPORT NO. 420
AUGUST 10, 1989

Prepared Under Contract
N00014-89-J-1627 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director



DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

1. INTRODUCTION.

In this article, we consider some very elementary but important problems which arise from modern uses of the computer in statistics, particularly in connection with testing goodness-of-fit. These involve (a) estimating percentage points; (b) simulating a Gaussian process; and (c) approximating the inverse of a covariance matrix of order statistiscs.

2. ESTIMATING PERCENTAGE POINTS.

The first problem is that of estimating percentage points of an intractable distribution. For example, the distributions of many goodness-of-fit statistics are very difficult to find, particularly for a finite sample, and particularly if parameters are estimated. The standard procedure is to simulate the situation considered, and calculate the statistic, say S; then repeat this n times to obtain the Monte Carlo distribution of S. The p-th percentile, for example, would be estimated by the $[np]+1$ order statistic of the S-sample. Some years ago, Schafer (1974) suggested that it would be better to take, say, c samples of $m=n/c$ Monte Carlo values, and take the average of the c estimates of the p-th percentile as the overall estimate; thus if $S_{i(k)}$ denotes the k-th order statistic of the i-th subsample, the estimate would be given by $\Sigma_{i=1}^{c} S_{i(k)}/c$ where $k = [mp]+1$. Schafer's suggestion was investigated by two colleagues and myself (Juritz, Juritz, and Stephens, 1983); we showed that the bias in the estimate is less for the estimate from one full sample than it is for the estimate from the mean of c subsamples, except in somewhat contrived cases, while the confidence intervals for the estimates were approximately the same size. This work was subsequently confirmed by Dudewicz and van der Meulen

(1984), who illustrated their work on a statistic they had introduced for testing uniformity. Zelterman (1987) also discusses this problem.

Bias in estimates of percentage points.

Of course, for certain distributions, exact values of m_k, the expected values of $X_{(k)}$, will be known, and the bias in the Monte Carlo estimate can be estimated. Thus, for the normal distribution, for a sample of size 100, the 95% point will be estimated by $X_{[96]}$, which has expected value 1.6872, to compare with the true percentile value of 1.6449. In Juritz, Juritz, and Stephens (1983) we published a table showing the bias for the 95% and 97.5% points for the standard normal distribution, and for various sample sizes. Table 1 below is a similar table for the standard exponential distribution with mean 1. The values of r and s are those for which the interval $I = (X_{(r)}, X_{(s)})$ is a 95% confidence interval for the true percentile. These are found to high accuracy as follows. Let ξ_p be the percentile at level p, let $w = \{np(1-p)\}^{1/2}$, and let z_γ denote the (100γ) percentile of the standard normal distribution. Then the choice $r = -w\, z_\gamma + np + \frac{1}{2}$ and $s = w\, z_\gamma + np + \frac{1}{2}$, where $r = 1-\alpha/2$, gives a $(1-\alpha)$ 100% confidence interval for ξ_p.

The approximate confidence interval length is obtained from a formula in Juritz, Juritz and Stephens (1983). It is clear that the bias diminishes with sample size; since the mean of several estimates will give the same bias, it is better to use one larger sample for a point estimate.

Confidence intervals for percentage points.

Although in practice point estimates of percentage points are nearly always all that are given when tables are produced, it is useful to examine confidence intervals for the points, as a guide to their accuracy. In Table 2 the multi-sample and single-sample confidence intervals are compared for

the 95% point of a standard normal distribution, with true value 1.645. The intervals are obtained from 6 Monte Carlo tests, for each of two comparisons. In the left-hand half of Table 2 the comparison is between $c = 10$ runs each of size $m = 100$, (the multi-sample case) and one run of size 1000 (the single-sample case). In the right hand half of the Table, the multi-sample case has $c = 10$ and $m = 500$, to compare with the single-sample $n = 5000$.

Consider the left-hand part of Table 2. For each of the 6 tests, in the multi-sample case, the estimate is $\bar{X}_p$, the mean of the 10 values of X_p, where X_p is $X_{(\ell)}$ with $\ell = 96$. The standard deviation σ_p of $X_{(\ell)}$ is 0.210, found from the formula in David (1970, p. 65). This is to be compared with S, the standard deviation of the 10 values of $X_{(\ell)}$. For each test, the 95% confidence limits for ξ_p have been found, and the length I_m of the confidence interval recorded.

In the single-sample case, the estimate of $\xi_p = X_{(\ell)}$ with $\ell = 951$. For each test, the confidence limits have been found from $X_{(r)}$ and $X_{(s)}$, with r,s as given above, and the length I_s recorded.

From the top half of the table, it can be seen that the experimental values S cluster around the theoretical value $\sigma_p = 0.210$. More importantly, the confidence intervals from the top and bottom parts of the table are approximately the same, as was shown by Juritz, Juritz and Stephens (1983), but the bias in the one large run is much less than in the multi-sample case. These conclusions are supported by the right-hand half of the table, where, in the top part, $\ell = 476$ and in the bottom part, $X_p = X_{(\ell)}$ with $\ell = 4751$. Thus, the studies demonstrate conclusively the advantage of the single sample over the multi-sample method. Dudewicz and van der Meulen (1984) discuss confidence intervals further.

Estimates obtained from approximations using moments.

We now turn to another idea; would it perhaps be better to approximate percentiles by calculating the sample moments from the 1000 (say) values in the large Monte Carlo run, and fitting a suitable curve to the moments to approximate the distribution? Several curve-fitting families are useful for such a purpose, especially when, as for many statistics, the distribution required can be expected to be smooth. For many goodness-of-fit statistics, the values are usually required in the longer tail (usually the upper tail); for such purposes, Pearson curves using 4 moments to make the fit have been found to be very useful. It should be emphasized that this has been the case when theoretical (that is, exact) moments could be calculated. Here we propose to experiment with sample moments, based on large samples. When also, as often happens, the lower end point of the distribution is known (often it is zero), a 3-moment Pearson curve fit can be tried, or a 3-moment fit of the form $(c\chi^2_p)^k$. Since higher sample moments have notoriously high sampling variability, there is something to be said for approximations using only 3 moments.

We have recently explored the curve-fitting possibility, as opposed to direct estimation of the percentile from the Monte Carlo sample, by again taking samples from distributions for which the exact percentiles are known.

Example 1. The Weibull distribution. The first illustration is for a variable x which has a Weibull distribution with shape parameter 2, that is, x^2 has the standard exponential distribution. Thus, the exact value of $\xi_p = [-\log(1-p)]^{1/2}$. The steps in the curve-fitting technique are:

(a) Take a sample of size n (say 1000) from the Weibull distribution, and estimate, say ξ_p by $X_{(k)}$, where $k = [np]+1$.

(b) Also, calculate the first four sample moments M'_r, $r = 1,\ldots,4$, where

$M'_r = \Sigma X_i^r/n$.

(c) Fit either a 4-moment or a 3-moment Pearson curve (knowing the lower end of the distribution is zero) and find the estimate Y_p (4-moment fit) or Z_p (3-moment fit) of ξ_p.

(d) Repeat steps (a), (b) and (c) 50 times, to obtain 50 estimates by each method.

(e) Calculate the average, the variance, and the mean square error (MSE) of the 50 estimates; from the average and the known value, we can estimate the bias.

The experiment can be repeated for different sample sizes; we used $n = 100$, 200, 500 and 1000. These sample sizes are small compared with those commonly used in Monte Carlo studies, but the trend of the results can easily be seen.

Table 3 gives a comparison of points all along the distribution, for both Pearson curve fits, for $n = 100$ and $n = 1000$. Type 1 refers to the 4-moment fit, and Type 2 to the 3-moment-and-lower-end-point fit.

Comments on Table 3.

(a) The Pearson curve fits often give somewhat greater bias to the estimate, but there is a smaller variance, so that the MSE of the Pearson curve estimate is better than for the Monte Carlo estimate.

(b) As expected, the 3-moment fit does better in the lower tail; but it is only very slightly worse in the upper tail; the marginal diference suggests that the 3-moment fit is to be preferred.

(c) There is a marked improvement in MSE as n gets larger in both methods, as one would expect; however, the relative sizes of MSE for Pearson curve fits compared to straight Monte Carlo estimation are still somewhat smaller as n increases.

Many goodness-of-fit statistics have asymptotic distributions which are sums of weighted chi-squares (see, e.g., Stephens, 1976, 1977, 1979), so our next example is a comparison for such a distribution. Statistic X has the distribution $X = .3z_1 + .2z_2 + .2z_3 + .1z_4 + .1z_5 + .1z_6$, where z are independent χ^2_1 variables. Table 4 gives a comparison of Monte Carlo and Pearson curve points as before, and again Pearson curves perform well in estimating points, measured by the MSE.

Implications for the bootstrap.

There are some interesting possible implications from this result. The bootstrap is now a very popular method for deducing properties of a statistic; the statistic is calculated many times over, by resampling from one sample. In its simplest form, the empirical distribution function (EDF) of the sample is used to estimate the population distribution and samples are then drawn from the estimate. It might, in some circumstances, be better to approximate the parent population by a smooth curve, such as a Pearson curve, fitted to the sample moments, and then to draw samples from the Pearson curve distribution when estimating properties of the relevant statistic by Monte Carlo methods.

3. SIMULATION OF A GAUSSIAN PROCESS

It is extremely useful, when finding percentage points for test statistics by Monte Carlo, to calculate the asymptotic points, instead of estimating them as described at the end of the previous section; then when the percentage points at level p are plotted against 1/n or $1/\sqrt{n}$, the curve is "anchored" at 1/n=0. If the curve can then be drawn with confidence, one can deduce percentage points for quite large samples without incurring the expense of large-sample Monte Carlo studies. Many

goodness-of-fit statistics are functionals of a process which is asymptotically a Gaussian process; such a process, based on the EDF, is often tied down at 0 and at 1, and is referred to as a Brownian bridge. Thus we wish to simulate a Brownian bridge, a Gaussian process $Z(t)$ with $Z(0) = 0$ and $Z(1) = 0$, and with known mean (often zero) and known covariance $\rho(s,t)$. For example, the Kolmogorov statistic D is the supremum of $Z(t)$, and the Cramer-von Mises statistic W^2 is $\int_0^1 Z^2(t)dt$.

Monte Carlo simulation of the process $Z(t)$ is difficult, and leads to further difficulties in finding the asymptotic distribution of D or W^2. Of necessity, on a computer, $Z(t)$ must be discretized. One way to construct a discrete approximation to $Z(t)$ is as follows:

(a) Choose values $t_1, t_2, \ldots, t_k$ equally spaced between 0,1.

(b) Generate u_i, a standard normal variate, at t_i, $i = 1,\ldots,k$. Let $u' =$ vector $(u_1, u_2, \ldots, u_k)$.

(c) Create V, a $k \times k$ matrix with entries $V_{ij} = \rho(t_i, t_j)$, where $\rho(s,t)$ is the covariance function of the Gaussian process $Z(t)$. Suppose W is the square root matrix of V, that is, $W = V^{1/2}$. Since V is positive definite this is easily obtained. Suppose $V = P \Delta P'$ where P is orthogonal and Δ is diagonal, with elements on the main diagonal equal to $\lambda_1, \lambda_2, \ldots, \lambda_k$. Then $W = P \Delta^* P'$, where Δ^* is diagonal with elements $\sqrt{\lambda_i}$, $i = 1,\ldots,k$.

(d) Let z' be vector $(z_1, z_2, \ldots, z_k)$, given by $z = Wu$.

(e) Then let $\hat{Z}(t_i)$, the estimate of $Z(t)$, be z_i. The mean $E(z) = 0$, and the covariance $E(zz') = E(Wuu'W') = V$. Thus the covariance $E(\hat{Z}(t_i)\hat{Z}(t_j) = \rho(t_i, t_j)$ and the values $\hat{Z}(t_i)$, $i = 1,\ldots,k$ give a discrete k-variate approximation to the continuous $Z(t)$. Note also that there can be other matrices W such that $WW' = V$, so that various approximations are possible.

Suppose the number of points k is called the order of the approximation.

Even when an approximation to $Z(t)$ has been created, there are clearly further approximations involved in calculating D or W^2. The above procedure must be repeated n times, say, to find the distribution of the statistics. One might then suppose that the percentage points of the asymptotic distribution of, say D, will be found as the limit of the smoothed Monte Carlo values plotted against $1/k$ or $1/\sqrt{k}$, as k becomes larger. Unfortunately as k becomes larger, the manipulation of the $k \times k$ matrices V and $V^{1/2}$ becomes increasingly prone to numerical errors. Chandra, Singpurwalla and Stephens (1981), carried out this procedure to obtain points D for use in testing for a Weibull distribution with unknown parameters, and found that as k became larger so that $m = 1/k \to 0$, the plot of a typical percentile of D against m was not monotonic. In the end, it then becomes difficult if not impossible to extrapolate to get asymptotic percentage points.

Another method exists of constructing an approximation to the process $Z(t)$. Since $\rho(s,t)$ is positive semi-definite, one can proceed as follows.

(a) Solve the integral equation

$$f_i(s) = \lambda_i \int_0^1 \rho(s,t) f_i(t) dt$$

for eigenvalues λ_i and eigenfunctions $f_i(t)$.

(b) Let $u_1, u_2, \ldots, u_k$ be a set of k independent standard normal variables.

(c) Let $\tilde{Z}_k(t) = \sum_{i=1}^{k} f_i(t) u_i / \sqrt{\lambda_i}$. Then as $k \to \infty$, $\tilde{Z}_k(t)$ tends to $\tilde{Z}(t)$, say; $\tilde{Z}(t)$ is a Gaussian process with mean 0 and covariance $\tilde{\rho}(s,t) = \sum_{i=1}^{\infty} f_i(s) f_i(t) / \lambda_i$ and by well-known properties of integral

equations, this is equal to $\rho(s,t)$, the kernel of the equation in (a). Thus $\tilde{Z}_k(t)$ can be regarded as a k-th order approximation to $Z(t)$. Here the approximation arises because in real calculations the sum must be terminated at a finite u_k; then the $\tilde{Z}_k(t)$ can be calculated at any point in the interval (0,1). Once a realization has been made by the choice of u_i, the value of D or W^2 can be found accurately. This approximation appears, therefore, to have some advantage: the difficulties arise in calculating the $f_i(t)$ and the λ_i. It would be interesting to see this technique explored further: a good problem on which to test it would be that of finding the distribution of D or W^2 when parameters in the tested distribution are fully specified: then the distributions of D and W^2 are both exactly known. If the technique is successful, it could be used to find the distribution of D for cases where parameters are unknown. The accuracy can then be tested by finding the points for W^2, and comparing with the exact points, which are known for this statistic; these points are given, for many distributions, in Stephens (1986a). To find the distribution of D, the Kolmogorov statistic, in cases where parameters are unknown is a problem of considerable interest, since D is a well-known statistic for testing fit; although in fact it is often much less powerful than W^2 or the related Anderson-Darling statistic A^2 (see, for example, Stephens 1974, 1986a).

In the above discussion we referred to the method of estimating asymptotic points of a distribution, as a parameter (k above) tends to infinity, by plotting points for finite k against $1/k$ or $1/\sqrt{k}$, and extrapolating a curve through these points to $1/k = 0$. How this extrapolation should be done is itself a problem. It often arises when asymptotic points (as sample size n tends to infinity) are required for making tables, say, and are to be obtained by extrapolating from Monte Carlo

results for finite samples. Suppose Monte Carlo experiments have given percentage points estimates $\hat{\xi}_{pn}$ for level p and sample size n. It is valuable to plot $\hat{\xi}_{pn}$ against 1/n, say, and then an extrapolation to $1/n = 0$ should give an estimate of the asymptotic point. However, how should this be done? It may even be known that the value ξ_{pn} can be expressed as $\xi_{pn} = a_o + a_1 m + a_2 m^2 + a_3 m^3, \ldots$, where $m=1/n$ or $1/\sqrt{n}$. The problem is then how best to estimate a_o from estimates $\hat{\xi}_{pn}$? We raise the question here because this appears to be an important practical problem in preparing tables of points, one which appears to need further examination. Knowledge of the values of $a_1, a_2, \ldots$ above would also be helpful to derive modified forms of test statistics, for example, of D or W^2, such as are used in some tables in Stephens (1986a). Such forms have the merit of drastically reducing the size of tables, and of making computerization of tables much easier.

4. APPROXIMATING THE INVERSE OF A COVARIANCE MATRIX OF ORDER STATISTICS

For some techniques of testing fit, based ultimately on the idea of probability plots, one needs V^{-1}, the inverse of V, where V is the covariance matrix of the order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ of a sample from a completely specified distribution. For a review of such tests see Stephens (1986b). The most notable example of the use of V^{-1} is with the Shapiro-Wilk test for normality, where the order statistics come from the normal distribution with mean 0 and variance 1.

When X has the uniform distribution with limts 0 and 1, the covariance matrix Q with entries $q_{ij} = \text{covariance}(X_{(i)}, X_{(j)})$ is given by $q_{ij} = (i/(n+1))(1-j/(n+1))/(n+2)$, $1 \le i \le j \le n$. The other entries, for $i > j$, are obtained from the symmetry of Q. More generally, if X has

continuous density $f(x)$, and if $m_i = E(X_{(i)})$, the covariance matrix V with entries v_{ij} = covariance $(X_{(i)},X_{(j)})$ has the approximation, for large n, $v_{ij} \approx q_{ij}/(f(m_i)f(m_j))$. (Blom, 1962). The inverse of cQ, where c is $(n+1)(n+2)$, is the matrix M given below. It then follows that V^{-1} can be approximated by $c\,DMD$, where D is the diagonal matrix, and M is the tri-diagonal matrix:

$$D = \begin{bmatrix} f(m_1) & & & \\ & f(m_2) & & \\ & & \ddots & \\ & & & f(m_n) \end{bmatrix} \qquad M = \begin{bmatrix} 2 & -1 & & & & & 0 \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & -1 & 2 & -1 & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ 0 & & & & & -1 & 2 \end{bmatrix}$$

M can also be written as $M=NN'$ where N is lower triangular. In this way, individual entries v^{ij} of V^{-1} can be approximated by easy formulas whose accuracy increases with n. However, for small n, this approximation is not accurate enough for many practical uses.

For the normal case, Davis and Stephens (1977) used various identities to give a good approximation to V, and this can be inverted to give V^{-1}. Can similar identities be used to give accurate approximations to V for other distributions, so that V can then be inverted to give V^{-1} accurately? Even if such identities were available, this technique seems a rather indirect way to approximate V^{-1}; can an accurate approximation for V^{-1} be found more directly? These are useful questions to answer not only in connection with tests of fit, but also for estimating parameters using linear combinations of order statistics.

ACKNOWLEDGEMENTS

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada, and by the U.S. Office of Naval Research, and the author expresses thanks to both of these agencies. This paper was given at the First International Conference on Statistical Computing held in Cesme, Izmir, Turkey, in March-April 1987, and the author thanks the organisers for the invitation to attend.

REFERENCES

Blom, G., (1962). Nearly best linear estimates of location and scale parameters. Section 48 in Contribution to Order Statistics (Eds. A.E. Sarhan and B.G. Greenberg), New York: Wiley.

Chandra, M., Singpurwalla, N.D. and Stephens, M.A. (1983). Some problems in simulating the quantiles of the maximum and other functions of Gaussian processes. J. Statist. Comput. Simul. 18, 45-57.

David, H.A., (1970). Order Statistics, John Wiley, New York.

David, C.S. and Stephens, M.A., (1977). The covariance matrix of normal order statistics. Commun. Statist. Simula. Computa., B6, 135-149.

Dudewicz, E.J. and van der Meulen, E.C., (1984). On assessing the precision of simulation estimates of percentile points. Amer. Jour. of Math. Manage. Sc. 4, 335-343.

Juritz, J.M., Juritz, J.W.F., and Stephens, M.A., (1983), On the accuracy of simulated percentage points. J. Amer. Statist. Assn., 78, 441-444.

Schafer, R.E. (1974). On assessing the precision of simulations, J. Statist. Comput. Simul., 3, 67-69.

Stephens, M.A., (1974). EDF statistics for goodness-of-fit and some comparisons. J. Amer. Statist. Assn., 69, 730-737.

Stephens, M.A., (1976a). Asymptotic results for goodness-of-fit statistics with unknown parameters. Ann. Statistics. 4, 357-369.

Stephens, M.A. (1977). Goodness-of-fit for the extreme value distribution. Biometrika 64, 583-588.

Stephens, M.A. (1979). Tests of fit for the logistic distribution based on the empirical distribution function. Biometrika 66, 591-595.

Stephens, M.A., (1986a). Tests based on EDF statistics. Chapter 4 in Goodness-of-fit Techniques (eds. R.B. d'Agostino and M.A. Stephens). New York: Marcel Dekker.

Stephens, M.A., (1986b). Tests based on regression and correlation. Chapter 5 in Goodness-of-fit Techniques (eds. R.B. d'Agostino and M.A. Stephens). New York: Marcel Dekker.

Zelterman, D., (1987). Estimating percentage points by simulation. J. Statist. Comput. Simul. 27, 107-125.

Table 1

Expected values of order statistics from the exponential distribution, used in estimates and confidence intervals for ξ_p.

n	k	m_k	bias x 10^4	r	m_r	s	m_s	L_s: Exact Expected C.I. length	Expected C.I. Length (approx.)
					p = .95; True ξ_p = 2.9957				
100	96	3.1040	1083	92	2.4695	100	5.1874	2.718	1.802
400	381	3.0222	265	372	2.6428	390	3.6410	.998	.866
500	476	3.0169	212	466	2.6746	486	3.5413	.867	.772
1000	951	3.0063	106	937	2.7572	965	3.3387	.582	.543
2000	1901	3.0010	53	1882	2.8262	1920	3.2129	.387	.383
5000	4751	2.9978	21	4721	2.8843	4781	3.1259	.242	.242
10000	9501	2.9968	11	9458	2.9142	9544	3.0868	.173	.171
100000	95001	2.9958	1	94866	2.9692	95136	3.0232	.054	.054
					p = .975; True ξ_p = 3.6889				
100	98	3.6874	-15	95	2.9040	-	-	-	-
400	391	3.7410	521	385	3.2517	397	4.7366	1.485	1.221
500	488	3.6896	7	481	3.2451	495	4.5095	1.264	1.093
1000	976	3.7095	206	966	3.3673	986	4.2339	.867	.773
2000	1951	3.6992	103	1937	3.4501	1965	4.0316	.582	.547
5000	4876	3.6930	41	4854	3.5303	4898	3.8874	.357	.346
10000	9751	3.6909	20	9720	3.5738	9782	3.8236	.250	.248
100000	97501	3.6891	2	97404	3.6510	97598	3.7287	.078	.077

Table 2

Estimates for the 95% point of a standard normal distribution (true value = 1.645).

1. Multi-Sample Results.

c = 10, m = 100 St. Dev. σ_p of $X_{(\ell)}$, (theory) = 0.210.

Test	Estimate $\bar{X}_p$	Bias x 10^3	S	C.I. Limits Lower	Upper	I_m - C.I. Length
1	1.794	149	.220	1.637	1.952	.352
2	1.792	147	.134	1.697	1.888	.192
3	1.676	31	.182	1.546	1.806	.260
4	1.668	23	.133	1.572	1.763	.190
5	1.685	40	.214	1.533	1.838	.305
6	1.587	-56	.246	1.413	1.765	.352

Average Bias x 10^3 = 56; Expected bias = 42.3

Single Sample Results

n = 1000

Test	Estimate X_p	Bias x 10^3	C.I. Limits Lower	Upper	I_s - C.I. Length
1	1.714	69	1.589	1.897	.308
2	1.721	76	1.627	1.822	.195
3	1.593	-52	1.490	1.699	.209
4	1.687	42	1.529	1.822	.293
5	1.612	-33	1.531	1.775	.244
6	1.554	-91	1.427	1.669	.242

Average Bias x 10^3 = 1.8: Expected bias = 4.1

2. Multi-sample Results

c = 10, m = 500, St. Dev. σ_p of $X_{(\ell)}$, (theory) = 0.094.

Test	Estimate $\bar{X}_p$	Bias x 10^3	S	C.I. Limits Lower	Upper	I_m - C.I. Length
	1.644	-1	.098	1.574	1.715	.141
	1.667	22	.113	1.586	1.748	.162
	1.640	-5	.113	1.559	1.722	.163
	1.688	43	.064	1.642	1.734	.092
	1.692	47	.105	1.617	1.767	.150
	1.645	0	.087	1.583	1.707	.124

Average Bias x 10^3 = 18: Expected bias = 8.2

Single Sample Results

n = 5000

Estimate X_p	Bias x 10^3	C.I. Limits Lower	Upper	I_s - C.I. Length
1.635	-10	1.581	1.698	.117
1.648	3	1.587	1.713	.126
1.638	-7	1.584	1.695	.111
1.673	28	1.626	1.746	.120
1.687	42	1.602	1.750	.148
1.639	-6	1.591	1.700	.109

Average Bias x 10^3 = 9: Expected bias = 0.8

Table 3

Comparison of estimates of percentage points obtained from (a) direct Monte Carlo estimates and (b) Pearson curve fits. The true distribution is Weibull with shape parameter 2. The results are based on 50 Monte Carlo runs, each with sample size n = 100 or n = 1000. The alpha-levels are measured from the lower tail.

Part 1. n = 100.

	Monte Carlo	Pearson Curve Type 1	Pearson Curve Type 2
Alpha level = 0.10		Exact Perc. Pt. = 0.324593	
A	0.33730	0.32783	0.32510
V	0.00363	0.00221	0.00218
B	0.01271	0.00323	0.00050
M	0.00379	0.00223	0.00218
Alpha level = 0.50		Exact Perc. Pt. = 0.832555	
A	0.82924	0.81905	0.82088
V	0.00411	0.00328	0.00307
B	-0.00331	-0.01350	0.01167
M	0.00412	0.00346	0.00320
Alpha level = 0.90		Exact Perc. Pt. = 1.517427	
A	1.52954	1.50708	1.50560
V	0.01196	0.00777	0.00675
B	0.01211	-0.01035	-0.01182
M	0.01211	0.00788	0.00689
Alpha level = 0.99		Exact Perc. Pt. = 2.145966	
A	2.21615	2.07182	2.07348
V	0.05308	0.02907	0.03337
B	0.07018	-0.07415	-0.07248
M	0.05801	0.03457	0.03862

Part 2. n = 1000.

	Monte Carlo	Pearson Curve Type 1	Pearson Curve Type 2
Alpha level = 0.10		Exact Perc. Pt. = 0.324593	
A	0.32454	0.32768	0.32726
V	0.00033	0.00025	0.00024
B	-0.00005	0.00309	0.00267
M	0.00033	0.00026	0.00025

A = Average; V = Variance; B = Bias; M = M.S.E.

	Monte Carlo	Pearson Curve Type 1	Pearson Curve Type 2
Alpha level = 0.050		Exact Perc. Pt. = 0.832555	
A	0.83302	0.83106	0.82760
V	0.00031	0.00028	0.00027
B	0.00046	-0.00149	-0.00495
M	0.00031	0.00028	0.00030
Alpha level = 0.90		Exact Perc. Pt. = 1.517427	
A	1.52723	1.52818	1.53567
V	0.00090	0.00066	0.00052
B	0.00980	0.01075	0.01824
M	0.00100	0.00078	0.00085
Alpha level = 0.99		Exact Perc. Pt. = 2.145966	
A	2.15864	2.15196	2.14043
V	0.00506	0.00241	0.00260
B	0.01268	0.00600	-0.00554
M	0.00523	0.00245	0.00263

Table 4

Comparison of estimates of percentage points (see Table 3). True distributions: sum of weighted chi-squares.

Part 1. n = 100.

	Monte Carlo	Pearson Curve Type 1	Pearson Curve Type 2
Alpha level = 0.10		Exact Perc. Pt. = 0.342	
A	0.34954	0.36148	0.32372
V	0.00170	0.00196	0.00154
B	0.00754	0.01948	-0.01828
M	0.00176	0.00234	0.00188
Alpha level = 0.50		Exact Perc. Pt. = 0.862	
A	0.86908	0.83069	0.86897
V	0.00382	0.00426	0.00298
B	0.00708	-0.03131	0.00697
M	0.00388	0.00524	0.00303

A = Average; V = Variance; B = Bias; M = M.S.E.

	Monte Carlo	Pearson Curve Type 1	Pearson Curve Type 2
Alpha level = 0.90		Exact Perc. Pt. = 1.831	
A	1.90994	1.90516	1.86493
V	0.04637	0.02837	0.02400
B	0.07894	0.07416	0.03393
M	0.05260	0.03387	0.02515
Alpha level = 0.99		Exact Perc. Pt. = 3.087	
A	3.59271	3.12139	3.11789
V	0.65287	0.19780	0.16248
B	0.50571	0.03439	0.03089
M	0.90861	0.19898	0.16343

Part 2. n = 1000.

Alpha level = 0.10		Exact Perc. Pt. = 0.342	
A	0.34527	0.34522	0.33629
V	0.00018	0.00016	0.00022
B	0.00327	0.00322	-0.00571
M	0.00019	0.00017	0.00025
Alpha Level = 0.50		Exact Perc. Pt. = 0.862	
A	0.86873	0.86074	0.87275
V	0.00057	0.00057	0.00044
B	0.00673	-0.00126	0.01075
M	0.00062	0.00057	0.00056
Alpha level = 0.90		Exact Perc. Pt. = 1.831	
A	1.85648	1.86264	1.84780
V	0.00325	0.00241	0.00217
B	0.02548	0.03164	0.01680
M	0.00390	0.00341	0.00245
Alpha level = 0.99		Exact Perc. Pt. = 3.087	
A	3.12492	3.08379	3.08268
V	0.02997	0.01909	0.01747
B	0.03792	-0.00321	-0.00432
M	0.03141	0.01910	0.01749

A = Average; V = Variance; B = Bias; M = M.S.E.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 420	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): Computer Problems In Goodness-Of-Fit		5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Michael A. Stephens		8. CONTRACT OR GRANT NUMBER(s): N00014-89-J-1627
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Statistics, Stanford University, Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Statistics & Probability Program Code 1111		12. REPORT DATE: August 10, 1989
		13. NUMBER OF PAGES: 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report): UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Goodness-of fit; Computer in statistics; Pearson curve fitting.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this article several problems are discussed involving the use of the computer in statistics. They have arisen in connection with the author's work in goodness-of-fit, but occur in many other applications. The subjects treated are: estimating percentage points (a) by Monte Carlo simulation and (b) by Pearson curve fitting; (c) estimation of asymptotic percentage points and (d) approximations to the covariance matrix of order statistics and its inverse.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)