

Bureau of Business Research  
Graduate School of Business Administration  
University of Michigan

August 1972

A Nonparametric Approach to  
the Construction of Prediction  
Intervals for Time Series Forecasts  
Working Paper No. 63

W. Allen Spivey  
William E. Wecker  
Graduate School of Business Administration  
University of Michigan

FOR DISCUSSION PURPOSES ONLY

None of this material is to be  
quoted or reproduced without  
the express permission of the  
Bureau of Business Research



## BACKGROUND OF THIS PAPER

This paper was given before the meetings of the American Statistical Association, August 16, 1972, in Montreal, Canada.

## ABSTRACT

The usual approach to the construction of confidence limits for time series forecasts begins with the assumption that the forecast errors are normally distributed. Williams and Goodman (JASA, December, 1971) have recently reported on a practical forecasting problem in which the data did not support an assumption of normality. They suggested an alternative procedure which requires that the underlying distribution of forecast errors be specified, although not necessarily as a normal distribution.

In this paper we develop a nonparametric approach to the construction of confidence limits which requires no such assumption concerning the distribution of forecast errors. This nonparametric procedure can be calculated from data more simply than either procedure above, and avoids the possibility of an erroneous distributional assumption and its attendant problems.



A Nonparametric Approach to the Construction  
of Prediction Intervals for Time Series Forecasts

1. Introduction. When forecasting a time series one is sometimes interested in specifying a random interval such that the actual future observation will lie in the interval with a given probability. This type of interval is sometimes called a prediction interval and is to be contrasted with a confidence interval, which, with a given probability, contains a fixed but unknown parameter.

Denote the time series to be forecast as

$$\{\dots X_{t-1}, X_t, X_{t+1}, \dots\},$$

the forecast of  $X_{t+1}$  made at time  $t$  as  $\hat{X}_{t+1}$  and the forecast error as  $e_{t+1} = X_{t+1} - \hat{X}_{t+1}$ . Then we seek a prediction interval  $(\hat{X}_{t+1} + L, \hat{X}_{t+1} + U)$  such that the probability is  $\alpha$  that the random variable  $X_{t+1}$  will appear in the interval, i.e., an interval such that

$$(1) \quad \Pr\{X_{t+1} \in (\hat{X}_{t+1} + L, \hat{X}_{t+1} + U)\} = \alpha,$$

$$L < 0 < U.$$

It is convenient to rewrite (1) so as to refer to the forecast error  $e_t$  and we have analogously for this random variable

$$(2) \quad \Pr\{e_{t+1} \in (L, U)\} = \alpha.$$

The standard approach to the construction of prediction intervals requires the assumption that the  $e_t$  are independent, identically distributed normal random variables [1], [2], [3]. Recently, Williams and Goodman described a practical forecasting study in which prediction intervals were desired, but in which also the normality assumption was found to be inappropriate [7]. As a result prediction intervals which were calculated on the basis of the normal theory failed to enclose the forecast error with the theoretically expected frequency. They developed an alternate procedure which requires one to determine a family of distributions having a member which gives a good fit to the empirical distribution of the forecast errors. For the data they considered the gamma family was chosen.

The Williams and Goodman procedure provides a general method for the construction of prediction intervals when the assumption of normality is not appropriate; however, a sophisticated data analysis capability is required to determine an appropriate alternate family of distributions. The purpose of this paper is to suggest another approach to the construction of prediction intervals -- distribution-free prediction intervals -- which does not require specification of an alternate family of distributions and so avoids the extensive data analysis necessary in the Williams and Goodman procedure.

2. Distribution-Free Prediction Intervals. We assume that the one step ahead forecast errors  $e_t$  are independent, identically distributed random variables with a fixed (continuous) density

function  $f(e)$ .<sup>1</sup> Following the suggestion of Williams and Goodman [7], we have a moving sample of fixed size  $n$ . That is, as the most recent forecast error is observed the error for the most remote time point in the sample is deleted so as to maintain a fixed sample size  $n$ . Let the sample observations be  $e_{t-n}, \dots, e_t$ , and let  $L$  and  $U$  the lower and upper end points, respectively, of the prediction interval be defined by the following functions:

$$(3) \quad L = g(e_{t-n}, \dots, e_t) = e_{(1)}$$

$$(4) \quad U = h(e_{t-n}, \dots, e_t) = e_{(n)},$$

where  $e_{(1)}$  and  $e_{(n)}$  are the first and  $n^{\text{th}}$  order statistics from a sample of size  $n$  from the distribution of errors  $f(e)$ .

We now calculate the probability  $\alpha$  that the random variable  $e_{t+1}$  will appear in this interval,

$$(5) \quad \alpha = \Pr\{L < e_{t+1} < U\}.$$

If  $e_{t+1}$  is in the interval  $(L, U)$ , then it must with probability one be in one of the subintervals  $(e_{(i)}, e_{(i+1)})$ ,  $i = 1, \dots, n-1$ , and the probability that  $e_{t+1}$  will be in any one of these sub-

---

<sup>1</sup>We regard the assumption of independence as a reasonable one because if the errors were dependent, one could in principle utilize this in predicting future errors based on past errors and thereby improve the forecast. For example, if the observed serial correlation of the errors suggests that they can be modeled by  $e_t = .9 e_{t-1} + a_t$ , where  $a_t$  is serially uncorrelated with mean 0, then we can use this equation for forecasting future errors. The forecast accuracy would, of course, depend on the error term  $a_t$ .

intervals is the same as for any other subinterval. Since the endpoints of the subintervals are order statistics and since the order statistics can be selected in  $n!$  equally likely ways, there is a total of  $n!(n-1)$  equally likely ways for the random variable  $e_{t+1}$  to be in the prediction interval. Now there are  $(n+1)!$  equally likely ways for the order statistics  $e_{(1)}, \dots, e_{(n+1)}$  to be arranged and therefore

$$\alpha = \Pr\{L < e_{t+1} < U\} = \frac{(n-1)n!}{(n+1)!} = \frac{n-1}{n+1} .$$

More generally for  $L = e_{(r)}, U = e_{(s)}, r < s \leq n,$

$$(6) \quad \alpha = \Pr\{e_{(r)} < e_{t+1} < e_{(s)}\} = \frac{s-r}{n+1} .$$

3. Distribution-Free Tolerance Intervals. Yet another distribution-free interval of potential use in forecasting is the tolerance interval. Let  $L$  and  $U$  be symmetric functions of the sample  $e_{t-n}, \dots, e_t$  such that the random interval  $(L,U)$  will with probability  $\gamma$  cover at least a fixed, preassigned proportion  $p$  of the density  $f(e)$ , i.e., an interval such that

$$(7) \quad \Pr \left\{ \int_L^U f(e) de \geq p \right\} = \gamma .$$

Such an interval is called a  $100p\%$  tolerance interval at probability level  $\gamma$  [6, p. 334-336]. We observe that (7) can be also written

$$(8) \quad \Pr\{F(U) - F(L) \geq p\} = \gamma ,$$



where  $F(U) = \int_{-\infty}^U f(e) de.$

For the case in which L and U are given by  $L = e_{(r)}, U = e_{(s)},$   
 $r < s \leq n$  it is well known that  $F(e_{(s)}) - F(e_{(r)})$  has a beta  
distribution and is independent of  $F(e)$  [6, p. 238]. Therefore  
the tolerance intervals defined by the order statistics  $e_{(r)}$  and  
 $e_{(s)}$  will be distribution-free. For fixed preassigned  $p, r, s,$  and  
 $n$  the incomplete beta function can be used to evaluate  $\gamma$  for the  
tolerance interval (7),

$$(9) \quad \gamma = \int_p^1 n \binom{n-1}{s-r-1} u^{s-r-1} (1-u)^{n-s+r} du,$$

$$0 \leq u \leq 1 .$$

4. Relations Between Distribution-Free Prediction and Tolerance  
Intervals. For a fixed sample size  $n,$  the distribution-free  
tolerance interval  $(L,U)$  will have an expected coverage

$$E \left\{ \int_L^U f(e) de \right\} = E \{ F(e_{(s)}) - F(e_{(r)}) \} = \frac{s-r}{n+1} ,$$

the mean of the beta distribution parameterized as in (9).

Therefore the  $\alpha$  level of a prediction interval defined by  
 $(e_{(r)}, e_{(s)})$  is identically the expected coverage of the  
 $(e_{(r)}, e_{(s)})$  tolerance interval. Additionally the values  $p$  and  
 $\gamma$  of the tolerance interval satisfy the relation

$$(10) \quad \alpha \geq p\gamma ,$$

as can be easily shown.<sup>2</sup>

5. The Application of Distribution-Free Intervals in Forecasting.

In order to make distribution-free interval forecasts one begins by generating sequentially point forecasts for known values of the time series using only information which was available prior to the observation being forecast. In this way a time series of forecast errors can be constructed. In practice, one chooses two positive integers,  $r$  and  $s$ ,  $r < s$ , and values of  $\alpha$  from (6),  $p$  and  $\gamma$  from (7) to satisfy the needs of the particular forecasting problem under study. Relation (6) can then be used to determine the sample size  $n$  required for a chosen  $\alpha$  level distribution-free prediction interval  $(e_{(r)}, e_{(s)})$ . Tables of the incomplete beta function can be used to determine a sample size for choices of  $p$  and  $\gamma$ . Charts are also available [4], [5]

---

<sup>2</sup>Let  $f(U,L)$  be the joint density of  $U$  and  $L$  and  $\Gamma$  be the set of values of  $U$  and  $L$  such that

$$\int_L^U f(e_{t+1}) de_{t+1} \geq p;$$

then

$$\begin{aligned} \alpha &= \Pr(L < e_{t+1} < U) \\ &= \iint_{\substack{\text{all} \\ U,L}} \Pr(L < e_{t+1} < U | U,L) f(U,L) dUdL \\ &\geq \iint_{U,L \in \Gamma} \Pr(L < e_{t+1} < U | U,L) f(U,L) dUdL \\ &\geq p \iint_{U,L \in \Gamma} f(U,L) dUdL = p\gamma. \end{aligned}$$

which can be used to quickly determine  $n$  given  $r, s, p$  and  $\gamma$ .

When the intervals have been specified, the following statements can be made with respect to future forecasts based on these intervals. For continued one step ahead forecasting we can expect the interval forecast to enclose  $100\alpha\%$  of the actual forecast errors as in (2). We can also be  $100\gamma\%$  confident that at least  $100p\%$  of the future observed errors will be enclosed. Analogous statements can be made with respect to the observations on the time series of interest instead of the time series of forecast errors by making use of (1) instead of (2).

## 6. Empirical Results

A brief empirical investigation was conducted on six time series of forecast errors of 131 points each which were supplied by the Bell Laboratories. The series being forecast are the same series discussed in Williams and Goodman [7]; however, the forecasting equation used differs slightly from their equation (3.1) in that the quadratic term  $t^2$  is not present.

Five distribution-free prediction intervals were calculated each from a moving sample of size 19:

### Prediction Interval

- I  $0 \pm \text{Max}\{|e_{(19)}|, |e_{(1)}|\}$
- II  $(e_{(1)}, e_{(19)})$
- III  $0 \pm \text{2nd largest}\{|e_{(19)}|, |e_{(18)}|, |e_{(1)}|, |e_{(2)}|\}$
- IV  $0 \pm \text{4th largest}\{|e_{(19)}|, |e_{(18)}|, |e_{(17)}|, |e_{(16)}|, |e_{(1)}|, |e_{(2)}|, |e_{(3)}|, |e_{(4)}|\}$
- V  $(e_{(2)}, e_{(18)})$

The sample size was chosen using equation (6) so that the prediction intervals would enclose the forecast error with the following probabilities:

<u>Prediction Interval</u>	<u>Probability Forecast error enclosed</u>
I	.95
II	.90
III	.90
IV	.80
V	.80

The observed relative frequency of successful enclosures for the distribution-free prediction intervals was compared to the observed relative frequency of successful enclosures based on the normal theory. The latter were supplied by Williams and Goodman in a private communication. The results are presented in Table 1.

Table 1

Theoretically Expected Relative Frequency of Successful Enclosures	Observed Relative Frequency of Successful Enclosures	
	Based on Normal Assumption	Distribution-free
.95	.8948	.9479
.90	.8257	.8988 (interval II) .8929 (interval III)
.80	.7163	.8140 (interval IV) .7976 (interval V)

## REFERENCES

1. Box, G.E.P. and G.M. Jenkins, Time Series Analysis, San Francisco, California: Holden-Day, Inc., 1970.
2. Goldberger, A.S., Econometric Theory, New York: John Wiley and Sons, 1964.
3. Graybill, F.A., An Introduction to Linear Statistical Models, vol. I, New York: McGraw-Hill Book Co., 1961.
4. Murphy, R.B., "Nonparametric Tolerance Intervals," Annals of Mathematical Statistics, vol. 19, 1948, pp. 581-589.
5. Walsh, J.E., Handbook of Nonparametric Statistics, Princeton, N.J.: D. Van Nostrand Co., 1962.
6. Wilks, S.S., Mathematical Statistics, New York: John Wiley and Sons, 1962.
7. Williams, W.H. and M.L. Goodman, "Constructing Empirical Confidence Limits for Economic Forecasts," Journal of the American Statistical Assn., vol. 66, no. 336, December, 1971, pp. 752-754.

