MONITORING TERRITORIAL RATING: A NONPARAMETRIC APPROACH

Bradford S. Gile

MONITORING TERRITORIAL RATING: A NONPARAMETRIC APPROACH

by Bradford S. Gile, FSA, MAAA

ABSTRACT

The primary concern in pricing is normally the overall adequacy of rates companywide, by state, and by territory. The primary concern of this paper, however, is the RELATIVE adequacy of rates by parts of a territory:

1. Is the rating plan for a given line or coverage in a particular territory equally "correct" in its various parts (counties, Zip Code groups, etc.)?
2. Is a particular part assigned to the right territory?

Because even whole territories often have experience of little or no credibility, traditional experience analysis is generally of little or no use. This paper circumvents the credibility problem by developing a nonparametric approach and statistical tests of the hypothesis that rates are "correct" throughout the territory under investigation. Actual applications of this process are shown.

I. THE BASIC PROBLEM

Rating territories are usually defined by the place of residence of the insured; the defining parameter is usually county or Zip Code, but other parameters are at least theoretically possible. The nature of the parameter is, for the purposes of this paper, immaterial. For this reason, we will use the parameter "county" throughout from now on, bearing in mind that we could just as easily use "Zip Code" or any other well defined parameter.

Territories, once defined, may remain unchanged for years without question. It might well seem reasonable to ask, after several years of experience, "Does our experience support our territorial definitions?" More commonly, however, the actuary will hear requests for changing the territory designation of one or more of its parts. Because experience by county is generally considered of little statistical value, the decision whether to make the change may be based solely on "judgment" (which may, unfortunately, be merely a synonym for political expediency). It is the purpose of this paper to develop a scientific approach from the experience by county in order to answer two basic questions:

QUESTION 1: Given no external information, does the experience indicate that one or more counties of a territory is improperly assigned?

This is a GLOBAL question about the territory.

QUESTION 2: Given external information about one or more counties in a particular rating territory, does the experience indicate that these counties do NOT belong in the territory to which they are assigned?

This second question is LOCAL; its focus is on one or more specified counties within a territory.

II. DEVELOPMENT OF THE STATISTICAL APPROACH

Consider the experience in a territory split up into M counties over an experience period of N years. For each year and county, suppose we have Earned Premium, losses on some uniformly consistent basis (e.g., estimated ultimate, calendar incurred, case incurred at a common age of loss), and exposure units (e.g.,policy-year, house-year, car-month,etc.). If we calculate the loss ratios by year and county we will, typically, get a matrix of loss ratios that fluctuate wildly due to lack of credibility in the individual cells. The result may appear to be meaningless. However, IF our pricing process is "correct", we would like to be able to assume the following KEY CRITERIA:

1. If in each year the loss ratios for each of the M counties are ranked by size , the M ranks are equally likely for any given county in any given year.
2. For any given county, rank in any given year is independent of the ranks that it has held in prior years.

It should be noted here that it makes no difference whether the ranking is done from low to high or high to low. Purely as a matter of personal preference, we will use the ordering 1 = lowest M =highest in this paper.

These key criteria will be satisfied if, for example, the territory is perfectly homogeneous and the M counties are all of equal size in exposure for each year. If the counties are NOT of equal size, however, we have a problem. Suppose that L(N) = losses for population size N, G = premium per unit and

$$LR(N) = \text{loss ratio} = \frac{L(N)}{N \cdot G}$$. Assuming perfect homogeneity,

the expected loss ratios will all be equal:

$$E(LR(N)) = \frac{E(L(N))}{N \cdot G} = \frac{N \cdot E(L(1))}{N \cdot G} = E(LR(1))$$

But for the loss ratio variances, we have

$$VAR(LR(N)) = \frac{VAR(L(N))}{(N \cdot G)^2} = \frac{N \cdot VAR(L(1))}{(N \cdot G)^2} = \frac{VAR(LR(1))}{N}$$

Thus, although all the expected values of county loss ratios are equal, the variances are not. Because the variance of the county loss ratio is inversely proportional to exposure, the smallest counties will have largest variance and may, therefore, be expected to have their rankings biased toward the upper and lower extremes. Because variation in exposure by county is a virtual certainty in real life, this problem must be dealt with.

The approach adopted in this paper is to substitute a set of linear transformations of the loss ratios for the loss ratios themselves. The transformed loss ratios will be called "adjusted loss ratios", and will be required to meet two conditions:

(A) The expected values of the adjusted loss ratios are equal to the expected values of the actual loss ratios.

(B) The adjusted ratios of the various counties will, in each experience year, have equal variances.

If we then rank the adjusted loss ratios, the bias due to unequal variances disappears.

Let LR(s,y) be the actual loss ratio for county s in year y, E(s,y) its exposure. We will rank ADJUSTED LOSS RATIOS ALR(s,y), $1 \le s \le M$, $1 \le y \le N$, in lieu of the actual LR(s,y):

$$(1)\ ALR(s,y) = Z(s,y)\cdot LR(s,y) + B(s,y)$$

If we write ELR(y) for the expected loss ratio in year y, (A) will require that (2) $B(s,y) = (1 - Z(s,y))\cdot ELR(y)$. Now since

$VAR(ALR(s,y)) = Z(s,y)^2\cdot VAR(LR(s,y))$ and the variance of the loss ratio is inversely proportional to exposure, we may write

$$(3)\ VAR(ALR(s,y)) = \frac{V(y)\cdot Z(s,y)^2}{E(s,y)},$$

where V(y) is the loss ratio variance in year y on one unit exposure. If we now take ratios in (3) for two counties s1 and s2 in year y, we get

$$(4)\ \left[\frac{Z(s1,y)}{Z(s2,y)}\right]^2 = \frac{E(s1,y)}{E(s2,y)}$$

as a necessary relationship for common variance amongst all counties in year y. Now the combination of (1) and (2) gives

$$(5)\quad ALR(s,y) = Z(s,y)\cdot LR(s,y) + (1 - Z(s,y))\cdot ELR(y)$$

which looks very much like a credibility formula. There are, of course, infinitely many ways in which the Z values may be defined to satisfy (4). This could well be a fertile area of further inquiry. The following definition of Z(s,y), however, is very appealing:

$$(6)\qquad Z(s,y) = \left[\frac{E(s,y)}{\max\limits_{1<=j<=M} \{ E(j,y) \}} \right]^{1/2}$$

This definition not only satisfies (4), but it also gives Z values between 0 and 1 which increase with exposure and equal 1 for the county of maximal exposure. Moreover, if all exposures ARE equal, the Z(s,y) = 1 and the adjusted loss ratios are equal to the actual loss ratios. The combination of (6) with (5) defines the adjusted loss ratios as credibility adjusted loss ratios such that the largest county is assigned full credibility and partial credibility is assigned to the other counties according to the traditional square root rule. Such adjusted loss ratios by county have the same expected values as the actual loss ratios and common variance, so that ranking of the adjusted loss ratios will not be biased due to unequal variances.

Moreover, the variances of the adjusted loss ratios satisfy

$$(7)\quad VAR(ALR(s,y)) = VAR(ALR(m,y)) = \frac{V(y)}{\underset{1<=j<=M}{MAX} \{ E(j,y) \}}$$

where $E(m,y) = \underset{1<=j<=M}{MAX} \{ E(j,y) \}$.

Thus, the adjusted loss ratios also have, in a sense, minimum variance.

Now suppose that we tabulate the adjusted loss ratios and rank them by size, so that the county having the lowest loss ratio gets rank 1 and the highest is rank M. We may avoid the complication of ties by viewing the adjusted loss ratio as a continuous random variable. Then, given a county, each of the possible ranks from 1 to M is equally likely. Now do this same ranking process for each of N years:

(8) R(s,y) = Rank for county s, year y; 1<=s<=M, 1<=y<=N

Each of these values has, by itself, no statistical value. However, for each county s, consider the ranksum defined by

(9) RANKSUM(s) = R(s,1) + R(s,2) + $\cdots$ + R(s,N); 1 <= s <= M

which is simply the sum of the ranks for county s over the N year period. RANKSUM(s) is identically distributed in each of the counties. The possible values are the integers from N (when s has rank 1 in every year) to M*N (when s has rank M in every year). Except when M and N are large, the exact probabilities of each possible ranksum can be calculated by brute force on a Personal Computer in a reasonably short time. A BASIC program that will do this is shown as Appendix I. Because this distribution is symmetrical with respect to its mean, a Normal approximation may be useful in cases where M*N is unduly large .

The unconditional mean and variance of the ranksum for a given county are given by

$$(10) \qquad \text{MEAN} = \frac{M+1}{2} \cdot N$$

$$(11) \qquad \text{VARIANCE} = \frac{M^2 - 1}{12} \cdot N$$

because when N = 1,

$\text{MEAN} = p_1 = \frac{1}{M} \cdot \left[1 + 2 + \cdots + M \right] = \frac{M+1}{2}$ and the second moment is

$$p_2 = \frac{1}{M} \cdot \left[1^2 + 2^2 + \cdots + M^2 \right] = \frac{(2 \cdot M + 1) \cdot (M + 1)}{6}$$

and $\text{VARIANCE} = p_2 - \text{MEAN}^2 = \frac{M^2 - 1}{12}$.

Under the hypothesis that our pricing process is correct, we can determine confidence intervals for the N year rank sum for any county selected at random. We then select a confidence level of 100*p % so that

(12) Pr(a <= Ranksum <= b) = p

and we tabulate all actual ranksum values outside of that confidence interval. This should, of course, be a two tailed test, such as

$$(13)\quad Pr(Ranksum<a) = Pr(Ranksum>b) = \frac{1 - p}{2} .$$

Now suppose we had been told in advance to watch a specific county as one which should be in a lower cost territory. For brevity, let us call the county under investigation Q. Then the a priori probability that Q's rank sum will be outside the confidence interval should be 1 - p.

If, in fact, the ranksum IS outside the interval, then we have statistical evidence (but NOT proof) that all is not well with our pricing system within the territory. We might well be willing to consider such a result to be strong evidence to support moving county Q to a lower rate territory. If, on the other hand, the ranksum for Q is not an extreme value we can only conclude that the study did not give an indication that Q's experience was unusual relative to that of the other counties. Surely, if County Q turned out to have a high extreme ranksum, indicating unusually high cost, we would reject any notion that experience supports a move of County Q to a lower rate territory!

If, on the other hand, we had been told nothing in advance of our study, we would be unable to draw any conclusions about the rating of specific counties. We can , however, still evaluate the overall hypothesis that our rating structure is correct by looking at the number of extreme ranksum values. Using Monte Carlo simulation of ranking M counties over N years, one can get an excellent approximation to the density function for the NUMBER of extreme ranksum values to be encountered, ranging from zero to M. It is clear that the ranksum values for the various counties are NOT independent of one another, because

$$(14)\quad \mathrm{RANKSUM}(1) + \cdots + \mathrm{RANKSUM}(M) = N \cdot \left[\frac{M \cdot (M+1)}{2} \right]$$

We want to know the distribution of the number of extreme values to be encountered in a given year in order to get a confidence interval. Unfortunately, the probability distribution of the number of extreme values to be encountered in a given year is extremely complex. The ranksum process itself is, for a given county, equivalent to throwing an M sided die N times. The selection of extreme ranksum values is analogous to the selection of colored balls from an urn without replacement, but with the additional complication that the selected balls must meet an additional aggregate criterion (14). Fortunately, Monte Carlo simulation on a Personal Computer can give us a good approximation of the extreme value distribution. Such a program, written in BASIC is shown as Appendix II. Experimentation with Monte Carlo simulation shows that the Binomial Distribution

$$(15) \quad f(x) = \binom{M}{x} \cdot (1 - p)^{x} * p^{M - x} \quad ; \; x = 0, 1, 2, \ldots, M$$

where p is the probability that a given county will have a ranksum value that is NOT extreme, provides an excellent approximation to the number of extreme values distribution for the determination of confidence intervals. When M is large, the process is akin to distinguishing "extreme" balls from "non-extreme" balls among a large number of balls in an urn, so that complications of (a) non-replacement of "balls" selected and (b) the constraint that the sum of all ranksums is a constant become minor and the distribution of the number of extreme values will approach the binomial defined by (15). To demonstrate the usefulness of the binomial approximation, consider the case of 69 counties observed over a four year period so that M=69 and N=4. The four year ranksums will range from 4 to 276, inclusive. Brute force production of the rank sum distribution (Exhibit A) by computer tells us that the ranksum for a given county will range from 63 to 217, inclusive, approximately 95% of the time (exactly: 21,551,431 out of 22,667,121 possible combinations). Extreme values would thus be less than 63 or greater than 217. Exhibit B shows that Monte Carlo simulation of 1,000 four year periods resulted in generating one to seven extreme values 946 times out of 1,000. Use of the Binomial distribution with p = 21,551,431/22,667,121 = 0.950779 predicts 949.4 out of 1,000 periods will produce one to seven extreme values. This illustrates the power of the Binomial approximation in estimating confidence intervals for the number of extreme values. Thus, if we had eight or more actual extreme values, our hypothesis of "correct" pricing across counties would be considered suspect in general, without making any conclusion as to which counties were, in fact, problematical.

III. APPLICATION OF THE RANKSUM PROCESS

As has been noted, this ranksum procedure may be used to help answer the two basic questions:

QUESTION 1: Given no external information, does the experience indicate that one or more counties of a territory is improperly assigned?

QUESTION 2: Given external information about one or more counties in a particular rating territory, does the experience indicate that these counties do NOT belong in the territory to which they are assigned?

Question 1 is for routine periodic monitoring. Even if there are no requests to change territorial composition, we should still test whether our territorial composition is still reasonable. Question 2, however, is designed for queries about the appropriateness of a given county's territorial assignment, and should be asked IN CONJUNCTION with Question 1.

In Question 2, we focus on whether the particular county has an extreme value. In both cases, we start with the hypothesis that our rating system is perfect. If, as will generally be the case, the counties are of unequal size, we adjust the loss ratios by (5) and (6) for each county and year, rank the adjusted loss ratios and tabulate the ranksums for each of the counties. Using a predetermined criterion for extreme values, such as those ranksum values outside of a 95% confidence interval as defined by (3), tabulate the number of such extreme values and the identities of the counties generating such values.

In order to evaluate the overall "perfect system" hypothesis for question 1, we need only compare the actual number of observed extreme values with a confidence interval, such as 95%, for the number of extreme values one would expect under the hypothesis. Without external information, however, we can make no judgment as to which counties having extreme ranksum values are merely statistical fluctuations or are true abnormalities. The answer to that question is the subject of question 2, which requires information in advance of the analysis.

If the answer to QUESTION 1 is "yes" and the county under investigation has an extreme value, there is a strong case for the assertion that the particular county is misplaced in its rating territory. If the answer to QUESTION 1 is "yes" and the county under question does NOT have an extreme value, we are left with a need for further analysis. One approach would be to remove the experience of all counties in question and ask whether the answer to QUESTION 1 is still "yes" on the collection of all remaining counties. If it is not, there would seem to be evidence that one or more of the counties under study may be misplaced.

Now suppose that the answer to QUESTION 1 is "No". This does NOT mean that our rating process is, in fact, correct. It simply means that if it is not correct, the experience does not yet BY ITSELF expose the system's imperfections. If, in fact, we have advance external information about a county and that county does, indeed, generate an extreme value, there is then some evidence to support the assertion that the particular county is incorrectly placed and that the "perfect system" hypothesis may, on the basis of additional information, be faulty after all.

Finally, suppose that the answer to QUESTION 1 is "NO" and the counties in question do not have extreme values. In this case, the ranksum procedure fails to corroborate an assertion that the county is misplaced.

No matter what the results may be, the ultimate decision whether or not to modify the territory's composition will have to rest squarely on judgment. Unless the external information is compelling, however, it seems inappropriate to make a change unless the statistical evidence from the experience also supports such a change.

Although this paper focuses on territorial composition, it should be clear that other applications are possible. For example, one might test the hypothesis that a given state has been "correctly" rated by territory or even whether the various states themselves have been equitably treated in the rating process!

IV. THE REAL WORLD: ACTUAL APPLICATIONS

American Family Mutual Insurance Company has developed a fairly large block of health insurance business over the last 30 years; in 1991 we had $186.5 million premium written in the twelve states in which we operate. With the exception of our Medicare Supplement business, our Health rating territories are defined by county of residence. Medicare Supplement territories, on the other hand, are defined by Zip Code groupings.

The county definitions were originally set many years ago, and have been subject to periodic modification. The impetus behind such modifications has generally come from field requests. Frustrated by the absence of a rational and scientific method to apply for the evaluation of the merits of such requests, this ranksum approach was developed.

The first application is to the QUESTION 2 type problem: Is a given county improperly placed in its territory?

Over a period of two years, several requests from the field requested that a specific county in a 69 county territory be moved to a lower rated territory, with no evidence for such a move other than an unsupported assertion (which might not even be relevant!) that "our insureds in this county go to hospitals in nearby county X which is in a lower rated territory."

Whether the assertion is correct or not is really unimportant. What IS important is the empirical evidence to be found in the experience. For each of the years 1986 through 1989, earned premium, case incurred losses at age 21 months, and policy-years of exposure were tabulated by county for the 69 counties. The loss ratios were calculated, adjusted by exposure according to (5) and (6), ranked (1 = lowest, 69 = highest), and the four year ranksums tabulated.

In this case we have N = 4 and M = 69. There are 22,667,121 (69 to the fourth power) possible rank combinations. With the aid of a Personal Computer, an exact determination of the probability distribution for ranksums even in this case is not particularly tedious. Exhibit A shows the graph of this distribution and development of a 95% confidence interval for ranksum values ranging from 63 to 217, inclusive. Exhibit B then develops a 95% confidence interval of from 1 to 7, inclusive for extreme values, showing both simulation and Binomial approximation results.

We now compare the actual results with Exhibit A and Exhibit B. County number 27 is the one that we were asked to change. The 8 counties with extreme values are :

COUNTY	4 YEAR RANK SUM	1989 EXPOSURE
11	58	24
23	19	23
27	32	120
32	221	54
38	41	44
46	231	298
54	32	17
63	57	188

In this case, we have an unusual number (8) of extreme values for the territory AND the county named in advance (number 27) has one of the extreme (low) values. Moreover, county 27 is one of the larger counties in the territory.

It is interesting to note that if the correction for bias had not been made to the loss ratios before ranking, there would have been 13 extreme cases. Most of the above extrema, including county 27, would NOT have appeared among the extreme cases. Instead, the list of extreme cases was dominated by counties having trivially low exposures.

This suggests that, instead of applying this method to ALL counties in the territory, perhaps only those counties having some minimum 1989 exposure, such as 50 policy years, should be counted in the analysis. In this particular case, the number of counties would be reduced from 69 to 11. To augment credibility, we added the 1990 experience to give us five years on eleven counties. For those who like to follow actual cases from beginning to end, Appendix III shows the full detail in this shortened case.

Interestingly, if the 95% confidence standard for extreme values is maintained, County 27 is no longer extreme; in fact, county 45, which is the largest of all counties in the territory is the only extreme case at this level of confidence. If we had chosen a confidence standard of 90% rather than 95%, Counties 45, 27, 32, and 63 would have emerged as "extreme"; the occurrence of 4 extreme values at this level of confidence is highly unusual.

From these analyses, it should be reasonably clear that the questioned county, number 27, has had unusually good experience. County 27 was, in fact, moved to a lower cost territory. Because there was no external input on other counties, no other counties were moved to different territories.

The above was a "real life" answer to QUESTION 2. What about QUESTION 1? We will now look at a "real life" situation for this question.

Two years ago, it became painfully clear to us that an entire territory, Territory A in State X, had a long term history of loss results that were unacceptably poor. This territory consists of 25 rural counties, so there was no clear reason why this particular territory had by far the worst experience in the Company. We decided to determine whether the cause might be due to an abnormal number of counties whose experience might identify them as the "bad apples". Appendix 4 shows the data and analysis of this territory by county and year for accident years 1986 - 1989. In this case, we have M=25 and N=4. The ranksum values of the adjusted loss ratios exhibit only two extreme values. This number of extreme values falls within a 90% confidence interval, so we do not conclude that we have an unusual number of extreme counties. Moreover, we are dealing here with a HIGH cost territory, so we are really interested in high extreme values rather than low ones. Interestingly enough, both of the observed extreme values are low rather than high. All of this suggests that, in essence, the territory experience is uniformly " bad". The answer to QUESTION 1 is, in this case, "No".

VI. CONCLUSIONS AND OBSERVATIONS

It should be emphasized that the process set forth in this paper does NOTHING to assess the adequacy or inadequacy of rates. That is a question of absolute magnitude. The process DOES attempt to assess RELATIVE adequacy of rates by county within territory. There are, no doubt, many questions that come to the reader which have not been addressed and should probably be researched further.

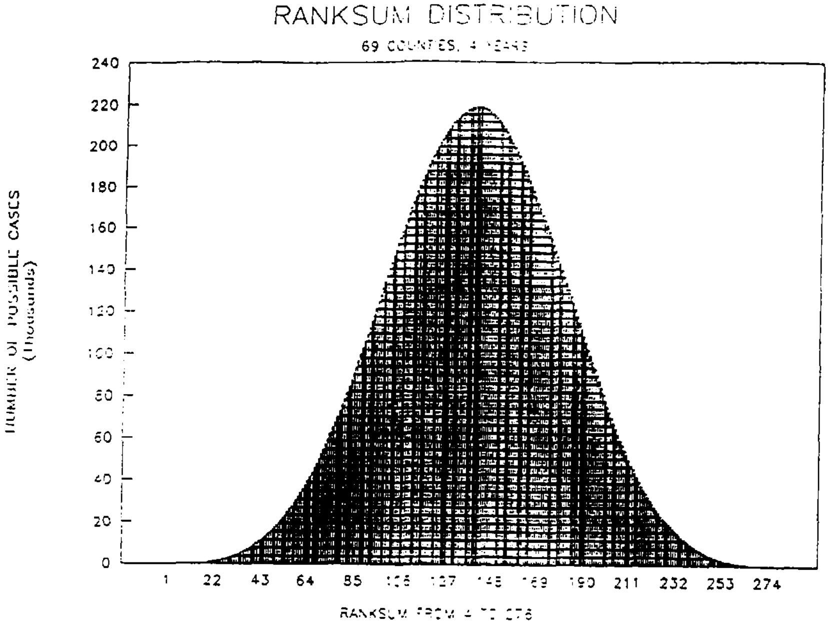
Examples that come to mind are:

1. The ranking process assumes, as part of the "correct pricing" hypothesis, that the territory is homogeneous in the sense that (1) the mean loss ratio is not changed by a population change and (2) the variance of the actual loss ratio is always inversely proportional to exposure. How much is lost with populations for which this does not hold?
2. Equation (4) defines the relationship between exposures and Z values in order that the M counties have a common variance. Although (6) turns out to be an extremely attractive choice, the possible choices are unlimited.
3. Nothing has been said about what data should be used, particularly losses. How does one deal with loss development on small populations? Are case incurred losses of equal maturity, for example, dependable as a proxy for "ultimate" losses for a long tailed coverage or line? The earned premiums for any county should, of course, be adjusted to the current territory if the county was in a different territory during part of the experience period.
4. To what extent should very small counties be removed from the analysis? What criteria should be employed?

Although the two historical examples given were in Health insurance, the methodology and principles should apply equally well to any personal line of insurance. Similarly, although the examples involved a county definition of territory, the way in which territory is defined is immaterial to the methodology.

Finally, although the problem to which this paper is addressed is territorial ratemaking, the nonparametric ranksum approach and analysis of extreme values of this paper (with particular emphasis on the use of a Personal Computer) should be applicable to an unlimited variety of actuarial questions involving comparative analysis.

EXHIBIT A

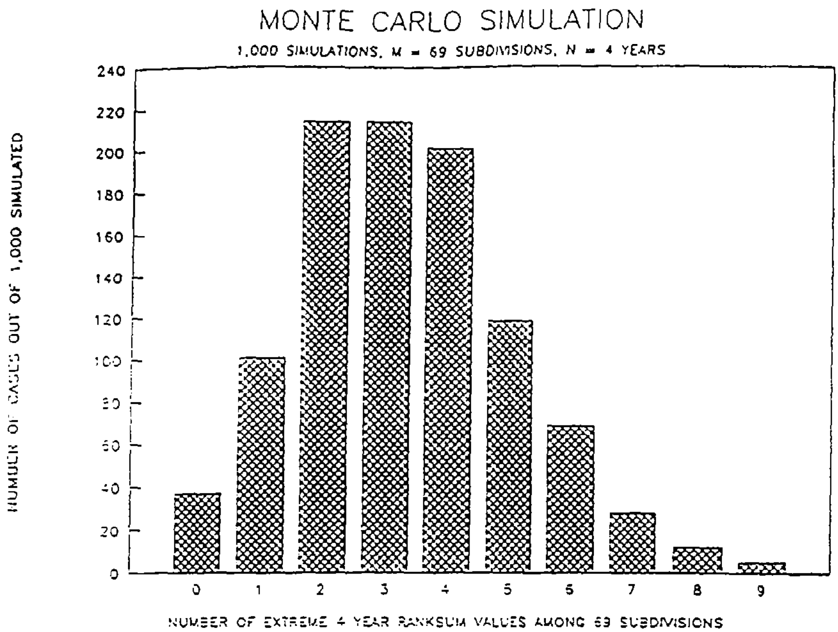


The above graph shows the exact probability density function for the ranksum values when M = 69 and N = 4. The possible ranksum values for a given subdivision range from 4 to 276, inclusive, as follows:

RANKSUM VALUES FROM	THROUGH	POSSIBLE COMBINATIONS	PERCENTAGE OF CASES IN RANGE
4	62	557,845	2.46 %
63	139	10,666,201	47.06
140	217	10,885,230	48.02
218	276	557,845	2.46
4	276	22,667,121	100.00 %

Thus, a two-tailed 95% confidence interval for the ranksum values is from 63 to 217, inclusive. Extreme values are (a) 4 to 62 and (b) 218 to 276.

EXHIBIT B



When M = 69 and N = 4, the two tailed 95% confidence interval for the ranksum of a given subdivision is from 63 to 217, inclusive. The program of Appendix II was run to simulate 1,000 four year experience periods, tabulate all ranksums and numbers of extreme values in order to approximate the distribution for number of extreme values. This result is compared with the BINOMIAL approximation

$$f(x) = \binom{M}{x} \cdot (1-p)^{x} \cdot p^{M-x}$$

$$\text{where} \quad p = \frac{21,551,431}{22,667,121} :$$

NUMBER OF EXTREME RANKSUMS	SIMULATION CASES OBSERVED	BINOMIAL APPROXIMATION PREDICTION
0	37	30.7
1	101	109.8
2	214	193.2
3	214	223.4
4	201	190.8
5	119	128.4
6	69	70.9
7	28	33.0
8	12	13.3
9	5	4.7
10+	0	2.0
ALL	1,000	1,000.0
1 TO 7	946	949.4

APPENDIX I: BASIC PROGRAM TO GENERATE EXACT DISTRIBUTION OF RANKSUM VALUES FOR ANY M AND N=3, 4 , 5, OR 6

(PAGE 1 OF 2)

```
MAIN:
CLS
INPUT "SUBDIVISIONS M";M
INPUT "YEARS N";N
DIM RS(M*N)
D=M^N
IF N=3 THEN GOSUB THREE
IF N=4 THEN GOSUB FOUR
IF N=5 THEN GOSUB FIVE
IF N=6 THEN GOSUB SIX
IF N>6 OR N<3 THEN GOTO MAIN
'REMARK: WE WILL CALCULATE THE TOTAL NUMBER OF WAYS OUT OF THE M^N
'        RANK COMBINATIONS THAT RANKSUM = J FOR EACH VALUE OF J
'        FROM M TO M*N. THESE VALUES WILL THEN BE WRITTEN TO A FILE
'        CALLED RESULTS.PRN.

OPEN "RESULTS.PRN" FOR OUTPUT AS 1
? #1,USING "RANKSUM DISTRIBUTION FOR ### SUBS AND ### YEARS";M;N
?  #1,"RANKSUM","NUMBER CASES"
FOR J=N TO M*N
? #1,J,RS(J)
NEXT J
RESET
?
? "FILE RESULTS.PRN  IS SET UP FOR FURTHER ANALYSIS."
END

'SUBROUTINES:
THREE:
FOR I1=1 TO M
FOR I2=1 TO M
FOR I3=1 TO M
S=I1+I2+I3
RS(S)=RS(S)+1
NEXT I3
NEXT I2
NEXT I1
RETURN

FOUR:
FOR I1=1 TO M
FOR I2=1 TO M
FOR I3=1 TO M
FOR I4=1 TO M
S=I1+I2+I3+I4
RS(S)=RS(S)+1
NEXT I4
NEXT I3
NEXT I2
NEXT I1
RETURN
```

APPENDIX I: BASIC PROGRAM TO GENERATE EXACT DISTRIBUTION OF RANKSUM VALUES FOR ANY M AND N=3, 4 , 5, OR 6
(PAGE 2 OF 2)

```
FIVE:
FOR I1=1 TO M
FOR I2=1 TO M
FOR I3=1 TO M
FOR I4=1 TO M
FOR I5=1 TO M
S=I1+I2+I3+I4+I5
RS(S)=RS(S)+1
NEXT I5
NEXT I4
NEXT I3
NEXT I2
NEXT I1
RETURN

SIX:
FOR I1=1 TO M
FOR I2=1 TO M
FOR I3=1 TO M
FOR I4=1 TO M
FOR I5=1 TO M
FOR I6=1 TO M
S=I1+I2+I3+I4+I5+I6
RS(S)=RS(S)+1
NEXT I6
NEXT I5
NEXT I4
NEXT I3
NEXT I2
NEXT I1
RETURN
```

APPENDIX II: BASIC PROGRAM FOR MONTE CARLO SIMULATION APPROXIMATION TO DISTRIBUTION OF NUMBER OF EXTREME VALUES

```
'EXTREME VALUE DISTRIBUTION GENERATOR
cls
INPUT "NUMBER SUBDIVISIONS";M
INPUT "NUMBER OF YEARS";N
INPUT "CONFIDENCE INTERVAL A,B";A,B
DIM NUMBER(M),R(M,N),RS(M)
OPEN "C:\TEMP\RESULTS.PRN" FOR OUTPUT AS 1
INPUT "TRIALS";T
T1=TIMER
RANDOMIZE TIMER
? #1,USING "#,###,### RANDOM TRIALS ON ### SUBS OVER ## YEARS";T;M;N
? #1,USING "EXTREME VALUES ARE LESS THAN ### OR GREATER THAN #,###";A;B
? #1,""
FOR TRIAL=1 TO T
FOR YEAR=1 TO N
X=RND
R(1,YEAR)=INT(M*X+1)
FOR S=2 TO M
     TEST1:
     R(S,YEAR)=INT(M*RND+1)
     FOR II=1 TO S-1
     IF R(S,YEAR)=R(II,YEAR) THEN
     'WE HAVE A DOUBLE COUNT
     GOTO TEST1
     END IF
     NEXT II
NEXT S
NEXT YEAR
W=0
FOR J=1 TO M
RS(J)=0
    FOR II=1 TO N
    RS(J)=RS(J)+R(J,II)
    NEXT II
IF RS(J)<A OR RS(J)>B THEN W=W+1
NEXT J
? #1,TRIAL,W
NEXT TRIAL
? "DONE!!!!!!"
? USING "RUN TIME ##,###.### SECONDS";TIMER-T1
? #1,USING "RUN TIME ##,###.### SECONDS";TIMER-T1
RESET
end
```

APPENDIX III: DETAILED DEVELOPMENT WHEN M = 11, N = 5
(PAGE 1 OF 5)

A. UNADJUSTED LOSS RATIOS
COUNTIES IN TERRITORY WITH 50 OR MORE POLICY YEARS EXPOSURE IN 1990

COUNTY	UNADJUSTED LOSS RATIOS BY YEAR AND COUNTY				
NUMBER	1990	1989	1988	1987	1986
7	58.4%	12.3%	43.0%	48.3%	77.5%
14	129.7%	34.0%	57.4%	49.3%	66.0%
27	53.6%	43.1%	34.1%	31.2%	17.3%
32	45.1%	98.8%	93.0%	70.4%	50.6%
35	140.5%	28.0%	50.6%	55.2%	156.7%
40	29.3%	92.2%	41.0%	90.4%	47.0%
46	67.2%	69.7%	67.2%	100.3%	59.1%
50	63.9%	20.0%	13.0%	63.4%	29.2%
52	101.1%	24.0%	38.9%	111.9%	35.4%
63	44.4%	51.4%	40.0%	49.3%	29.8%
67	37.4%	61.3%	52.2%	28.4%	30.6%

B. POLICY YEARS OF EXPOSURE
COUNTIES IN TERRITORY WITH 50 OR MORE POLICY YEARS EXPOSURE IN 1990

COUNTY	POLICY YEARS OF EXPOSURE BY COUNTY AND YEAR				
NUMBER	1990	1989	1988	1987	1986
7	79	82	62	49	30
14	156	146	134	127	122
27	151	120	97	74	47
32	52	54	44	46	45
35	79	63	51	50	51
40	82	89	89	88	101
46	297	298	273	286	273
50	53	50	46	45	51
52	71	65	64	64	71
63	198	188	189	208	188
67	125	121	125	121	115
	1,343	1,276	1,174	1,158	1,094

C. EXPOSURE ADJUSTED LOSS RATIOS
COUNTIES IN TERRITORY 6 WITH 50 OR MORE POLICY YEARS
EXPOSURE IN 1990

YEAR	EXPECTED LOSS RATIO	MAXIMUM EXPOSURE
1986	69.1%	273
1987	52.4%	286
1988	50.7%	273
1989	66.8%	298
1990	50.7%	297

APPENDIX III: DETAILED DEVELOPMENT WHEN M = 11, N = 5
(PAGE 2 OF 5)

D. VALUES OF Z(COUNTY,YEAR)

COUNTY NUMBER	1990	1989	1988	1987	1986
7	0.515745	0.524564	0.476557	0.413919	0.331497
14	0.724743	0.699952	0.700602	0.666375	0.668496
27	0.713034	0.634574	0.596080	0.508666	0.414923
32	0.418431	0.425685	0.401463	0.401048	0.405999
35	0.515745	0.459793	0.432219	0.418121	0.432219
40	0.525447	0.546496	0.570971	0.554700	0.608246
46	1.000000	1.000000	1.000000	1.000000	1.000000
50	0.422435	0.409616	0.410485	0.396664	0.432219
52	0.488935	0.467034	0.484182	0.473050	0.509974
63	0.816497	0.794275	0.832050	0.852803	0.829846
67	0.648749	0.637213	0.676665	0.650444	0.649034

E. ADJUSTED LOSS RATIOS

COUNTY NUMBER	1990	1989	1988	1987	1986
7	54.7%	38.2%	47.0%	50.7%	71.9%
14	108.0%	43.9%	55.4%	50.4%	67.0%
27	52.8%	51.8%	40.8%	41.6%	47.6%
32	48.3%	80.4%	67.7%	59.6%	61.6%
35	97.0%	48.9%	50.7%	53.6%	107.0%
40	39.5%	80.6%	45.1%	73.5%	55.7%
46	67.2%	69.7%	67.2%	100.3%	59.1%
50	56.3%	47.6%	35.2%	56.8%	51.8%
52	75.3%	46.8%	45.0%	80.5%	51.9%
63	45.5%	54.5%	41.8%	49.8%	36.5%
67	42.0%	63.3%	51.7%	36.8%	44.1%

F. RANKINGS OF ADJUSTED LOSS RATIOS
(1 = LOWEST, 11 = HIGHEST)

COUNTY NUMBER	1990	1989	1988	1987	1986	RANK SUM
7	6	1	6	5	10	28
14	11	2	9	4	9	35
27	5	6	2	2	3	18
32	4	10	11	8	8	41
35	10	5	7	6	11	39
40	1	11	5	9	6	32
46	8	9	10	11	7	45
50	7	4	1	7	4	23
52	9	3	4	10	5	31
63	3	7	3	3	1	17
67	2	8	8	1	2	21

APPENDIX III: DETAILED DEVELOPMENT WHEN M = 11, N = 5
(PAGE 3 OF 5)

G. EXACT RANKSUM PROBABILITY DISTRIBUTION: M = 11, N = 5
VALUES RANGE FROM 5 TO 55
TOTAL COMBINATIONS = 11^5 = 161,051

95% CONFIDENCE INTERVAL: FROM 16 TO 43. A=16, B=43
p = PROB(A<=RANKSUM<=B) = 0.954263

RANK SUM	POSSIBLE NUMBER CASES	PROBABILITY	CUMULATIVE
5	1	0.000006209	0.000006209
6	5	0.000031046	0.000037255
7	15	0.000093138	0.000130393
8	35	0.000217322	0.000347715
9	70	0.000434644	0.000782360
10	126	0.000782360	0.001564721
11	210	0.001303934	0.002868656
12	330	0.002049040	0.004917696
13	495	0.003073560	0.007991257
14	715	0.004439587	0.012430844
15	1001	0.006215422	0.018646267
16	1360	0.00844453	0.027090797
17	1795	0.011145537	0.038236335
18	2305	0.014312236	0.052548571
19	2885	0.017913580	0.070462151
20	3526	0.021893685	0.092355837
21	4215	0.026171833	0.118527671
22	4935	0.030642467	0.149170138
23	5665	0.035175193	0.184345331
24	6380	0.039614780	0.223960112
25	7051	0.043781162	0.267741274
26	7645	0.047469435	0.315210709
27	8135	0.050511949	0.365722659
28	8500	0.052778312	0.418500971
29	8725	0.054175385	0.472676357
30	8801	0.054647285	0.527323642
31	8725	0.054175385	0.581499028
32	8500	0.052778312	0.634277340
33	8135	0.050511949	0.684789290
34	7645	0.047469435	0.732258725
35	7051	0.043781162	0.776039888
36	6380	0.039614780	0.815654668
37	5665	0.035175193	0.850829861
38	4935	0.030642467	0.881472328
39	4215	0.026171833	0.907644162
40	3526	0.021893685	0.929537848
41	2885	0.017913580	0.947451428
42	2305	0.014312236	0.961763664
43	1795	0.011145537	0.972909202
44	1360	0.008444530	0.981353732
45	1001	0.006215422	0.987569155

G. EXACT RANKSUM PROBABILITY DISTRIBUTION: M = 11, N = 5
VALUES RANGE FROM 5 TO 55
TOTAL COMBINATIONS = 11^5 = 161,051

95% CONFIDENCE INTERVAL: FROM 16 TO 43. A=16, B=43
p = PROB(A<=RANKSUM<=B) = 0.954263

RANK SUM	POSSIBLE NUMBER CASES	PROBABILITY	CUMULATIVE
46	715	0.004439587	0.992008742
47	495	0.003073560	0.995082303
48	330	0.002049040	0.997131343
49	210	0.001303934	0.998435278
50	126	0.000782360	0.999217639
51	70	0.000434644	0.999652284
52	35	0.000217322	0.999869606
53	15	0.000093138	0.999962744
54	5	0.000031046	0.999993790
55	1	0.000006209	1.000000000
	161,051		

H. DISTRIBUTION OF EXTREME VALUE COUNTS
M = 11, N = 5
BINOMIAL p = 0.954263

NUMBER OF EXTREMA	MONTE CARLO OBSERVATIONS	BINOMIAL PREDICTION
0	613	597.5
1	298	315.0
2	80	75.5
3	9	10.9
4+	0	1.1
	1,000	1,000

I. RANKSUM TESTING FOR EXTREME VALUES
95% EXTREMA: UNDER 16 OR OVER 43

COUNTY	RANKSUM	EXTREME?
7	28	NO
14	35	NO
27	18	NO
32	41	NO
35	39	NO
40	32	NO
46	45	YES
50	23	NO
52	31	NO
63	17	NO
67	21	NO

APPENDIX III: DETAILED DEVELOPMENT WHEN M = 11, N = 5
(PAGE 5 OF 5)

J. REMARKS AND OBSERVATIONS

1. In this case, there is only ONE extreme value - not an unexpected result. County 27 just slightly misses, as do counties 32 and 63. Had a 90% confidence interval been the standard here, extreme values would be less than 19 or greater than 42 and counties 27, 32, and 63 would be added to the "EXTREME" category. At the 90% confidence level, 4 extreme values is highly unusual, occurring about 1% of the time.

2. It is also interesting that county 45 is the LARGEST county in the territory and has an extreme HIGH value even at the 95% confidence level.

APPENDIX IV: TESTING OVERALL TERRITORIAL CONSISTENCY
(PAGE 1 OF 7)

I. 21 MONTH CASE INCURRED LOSS RATIOS BY COUNTY AND ACCIDENT YEAR

County	1989	1988	1987	1986
1	109.5%	94.9%	65.0%	191.4%
2	108.0%	73.4%	94.7%	76.3%
3	211.0%	69.2%	63.6%	234.5%
4	81.4%	84.9%	338.7%	34.0%
5	39.8%	132.6%	101.0%	64.2%
6	81.2%	93.5%	61.8%	43.6%
7	45.8%	41.0%	215.5%	250.6%
8	114.7%	189.7%	97.4%	53.8%
9	66.4%	77.4%	73.7%	103.9%
10	115.0%	110.7%	148.7%	22.6%
11	65.7%	58.9%	75.5%	134.3%
12	62.4%	83.3%	71.5%	58.0%
13	52.6%	73.5%	77.2%	66.3%
14	63.0%	75.2%	130.9%	61.5%
15	23.0%	120.6%	17.9%	0.0%
16	108.5%	113.1%	47.2%	49.4%
17	110.7%	63.7%	107.7%	131.2%
18	107.2%	53.7%	67.2%	100.4%
19	137.7%	100.9%	51.5%	34.9%
20	146.3%	43.7%	87.8%	254.1%
21	63.5%	61.0%	53.9%	55.4%
22	88.9%	104.3%	59.1%	52.6%
23	95.4%	44.6%	82.4%	120.0%
24	95.1%	55.4%	60.8%	38.4%
25	136.9%	74.6%	196.1%	47.9%

II. POLICY YEAR EXPOSURES BY COUNTY AND ACCIDENT YEAR

County	1989	1988	1987	1986
1	181	187	152	141
2	205	196	211	212
3	134	123	100	105
4	15	22	24	27
5	86	78	79	87
6	121	138	131	133
7	130	151	143	152
8	108	122	143	162
9	92	90	91	94
10	69	57	54	54
11	195	196	208	215
12	267	249	251	246
13	151	160	133	121
14	111	107	98	103
15	27	27	15	18
16	86	86	79	75
17	110	109	101	106
18	91	100	107	120
19	60	71	69	65

APPENDIX IV: TESTING OVERALL TERRITORIAL CONSISTENCY
(PAGE 2 OF 7)

County	1989	1988	1987	1986
20	37	43	43	41
21	147	171	177	176
22	203	213	197	197
23	66	70	67	69
24	160	185	190	173
25	74	87	90	90
MAXIMUM	267	249	251	246

II. EXPECTED LOSS RATIOS AND Z VALUES

YEAR	EXPECTED LOSS RATIO
1986	92.8%
1987	90.2%
1988	82.2%
1989	91.4%

Z COEFFICIENTS FOR ADJUSTED LOSS RATIOS

County	1989	1988	1987	1986
1	0.823348	0.866605	0.778189	0.757080
2	0.876236	0.887214	0.916863	0.928326
3	0.708430	0.702834	0.631194	0.653322
4	0.237023	0.297243	0.309221	0.331295
5	0.567536	0.559690	0.561018	0.594692
6	0.673189	0.744457	0.722435	0.735289
7	0.697776	0.778733	0.754799	0.786057
8	0.635999	0.699971	0.754799	0.811503
9	0.587000	0.601204	0.602121	0.618154
10	0.508357	0.478451	0.463831	0.468521
11	0.854598	0.887214	0.910322	0.934871
12	1.000000	1.000000	1.000000	1.000000
13	0.752026	0.801605	0.727929	0.701334
14	0.644772	0.655529	0.624851	0.647070
15	0.317999	0.329293	0.244461	0.270501
16	0.567536	0.587692	0.561018	0.552158
17	0.641861	0.661628	0.634343	0.656425
18	0.583801	0.633724	0.652913	0.698430
19	0.474045	0.533986	0.524309	0.514031
20	0.372259	0.415561	0.413902	0.408248
21	0.741999	0.828702	0.839750	0.845841
22	0.871952	0.924890	0.885924	0.894882
23	0.497183	0.530212	0.516655	0.529611
24	0.774113	0.861958	0.870041	0.838601
25	0.526454	0.591099	0.598804	0.604858

APPENDIX IV: TESTING OVERALL TERRITORIAL CONSISTENCY
(PAGE 3 OF 7)

III. ADJUSTED LOSS RATIOS BY COUNTY AND ACCIDENT YEAR

County	1989	1988	1987	1986
1	106.6%	94.2%	68.8%	167.1%
2	106.2%	75.3%	93.7%	77.4%
3	176.5%	75.4%	70.4%	184.9%
4	90.1%	88.6%	161.5%	72.4%
5	62.7%	113.9%	92.7%	75.2%
6	85.0%	92.7%	67.4%	56.2%
7	60.0%	51.9%	182.8%	216.6%
8	106.7%	159.8%	93.6%	60.9%
9	77.3%	82.5%	77.1%	99.1%
10	104.1%	100.0%	113.0%	59.2%
11	69.6%	62.5%	76.1%	131.5%
12	62.4%	83.3%	71.5%	58.0%
13	62.5%	76.8%	78.5%	73.8%
14	73.6%	80.4%	112.6%	72.1%
15	70.6%	100.2%	66.4%	66.7%
16	101.7%	103.7%	62.5%	68.2%
17	104.3%	72.7%	98.3%	117.5%
18	101.2%	67.1%	72.4%	97.7%
19	114.1%	95.9%	66.1%	62.3%
20	112.7%	70.9%	84.5%	157.8%
21	71.0%	66.0%	58.4%	60.9%
22	89.4%	103.2%	61.7%	56.6%
23	94.1%	66.0%	82.3%	106.5%
24	94.5%	60.2%	63.6%	46.9%
25	116.0%	81.0%	150.4%	65.1%

IV. RANKINGS AND RANKSUMS OF ADJUSTED LOSS RATIOS

County	1989	1988	1987	1986	RANKSUM
1	20	18	8	23	69
2	19	9	19	16	63
3	25	10	9	24	68
4	12	16	24	13	65
5	4	24	17	15	60
6	10	17	7	2	36
7	1	1	25	25	52
8	21	25	18	6	70
9	9	14	13	18	54
10	17	20	22	5	64
11	5	3	12	21	41
12	2	15	10	4	31
13	3	11	14	14	42
14	8	12	21	12	53
15	6	21	6	10	43
16	16	23	3	11	53
17	18	8	20	20	66
18	15	6	11	17	49
19	23	19	5	8	55

APPENDIX IV: TESTING OVERALL TERRITORIAL CONSISTENCY
(PAGE 4 OF 7)

IV. RANKINGS AND RANKSUMS OF ADJUSTED LOSS RATIOS

County	1989	1988	1987	1986	RANKSUM
20	22	7	16	22	67
21	7	5	1	7	20
22	11	22	2	3	38
23	13	4	15	19	51
24	14	2	4	1	21
25	24	13	23	9	69

V. DISTRIBUTION OF RANKSUMS WHEN M = 25, N = 4

RANKSUM VALUE	NUMBER COMBINATIONS	PROBABILITY	CUMULATIVE
4	1	0.000003	0.000003
5	4	0.000010	0.000013
6	10	0.000026	0.000038
7	20	0.000051	0.000090
8	35	0.000090	0.000179
9	56	0.000143	0.000323
10	84	0.000215	0.000538
11	120	0.000307	0.000845
12	165	0.000422	0.001267
13	220	0.000563	0.001830
14	286	0.000732	0.002563
15	364	0.000932	0.003494
16	455	0.001165	0.004659
17	560	0.001434	0.006093
18	680	0.001741	0.007834
19	816	0.002089	0.009923
20	969	0.002481	0.012403
21	1140	0.002918	0.015322
22	1330	0.003405	0.018726
23	1540	0.003942	0.022669
24	1771	0.004534	0.027203
25	2024	0.005181	0.032384
26	2300	0.005888	0.038272
27	2600	0.006656	0.044928
28	2925	0.007488	0.052416
29	3272	0.008376	0.060792
30	3638	0.009313	0.070106
31	4020	0.010291	0.080397
32	4415	0.011302	0.091699
33	4820	0.012339	0.104038
34	5232	0.013394	0.117432
35	5648	0.014459	0.131891
36	6065	0.015526	0.147418
37	6480	0.016589	0.164006
38	6890	0.017638	0.181645
39	7292	0.018668	0.200312
40	7683	0.019668	0.219981

V. *DISTRIBUTION OF RANKSUMS WHEN M = 25, N = 4*

RANKSUM VALUE	NUMBER COMBINATIONS	PROBABILITY	CUMULATIVE
41	8060	0.020634	0.240614
42	8420	0.021555	0.262170
43	8760	0.022426	0.284595
44	9077	0.023237	0.307832
45	9368	0.023982	0.331814
46	9630	0.024653	0.356467
47	9860	0.025242	0.381709
48	10055	0.025741	0.407450
49	10212	0.026143	0.433592
50	10328	0.026440	0.460032
51	10400	0.026624	0.486656
52	10425	0.026688	0.513344
53	10400	0.026624	0.539968
54	10328	0.026440	0.566408
55	10212	0.026143	0.592550
56	10055	0.025741	0.618291
57	9860	0.025242	0.643533
58	9630	0.024653	0.668186
59	9368	0.023982	0.692168
60	9077	0.023237	0.715405
61	8760	0.022426	0.737830
62	8420	0.021555	0.759386
63	8060	0.020634	0.780019
64	7683	0.019668	0.799688
65	7292	0.018668	0.818355
66	6890	0.017638	0.835994
67	6480	0.016589	0.852582
68	6065	0.015526	0.868109
69	5648	0.014459	0.882568
70	5232	0.013394	0.895962
71	4820	0.012339	0.908301
72	4415	0.011302	0.919603
73	4020	0.010291	0.929894
74	3638	0.009313	0.939208
75	3272	0.008376	0.947584
76	2925	0.007488	0.955072
77	2600	0.006656	0.961728
78	2300	0.005888	0.967616
79	2024	0.005181	0.972797
80	1771	0.004534	0.977331
81	1540	0.003942	0.981274
82	1330	0.003405	0.984678
83	1140	0.002918	0.987597
84	969	0.002481	0.990077
85	816	0.002089	0.992166
86	680	0.001741	0.993907
87	560	0.001434	0.995341

V. DISTRIBUTION OF RANKSUMS WHEN M = 25, N = 4

RANKSUM VALUE	NUMBER COMBINATIONS	PROBABILITY	CUMULATIVE
88	455	0.001165	0.996506
89	364	0.000932	0.997437
90	286	0.000732	0.998170
91	220	0.000563	0.998733
92	165	0.000422	0.999155
93	120	0.000307	0.999462
94	84	0.000215	0.999677
95	56	0.000143	0.999821
96	35	0.000090	0.999910
97	20	0.000051	0.999962
98	10	0.000026	0.999987
99	4	0.000010	0.999997
100	1	0.000003	1.000000
	390,625		

VI. DISTRIBUTION OF NUMBER OF EXTREME VALUES PER PERIOD

95% RANKSUM CONFIDENCE INTERVAL FROM 23 TO 79, INCLUSIVE

COMBINATIONS 23 TO 79	372,684
TOTAL COMBINATIONS	390,625
BINOMIAL p VALUE =	0.954071

x	BINOMIAL f(x)	PREDICTED CASES PER 1,000 TRIALS EXTREME = x	EXTREME <= x
0	0.308687	309	309
1	0.371504	372	680
2	0.214610	215	895
3	0.079207	79	974
4	0.020972	21	995
5	0.004240	4	999
6+	0.000780	1	

The binomial approximation predicts that 90% of the time, the number of extreme values is two or less, and 97.4% of the time it will be three or less.

Monte Carlo simulation of 1,000 four year periods gives the following results:

NUMBER OF EXTREME VALUES n	NUMBER OF OBSERVATIONS FOR WHICH EXTREME = n	EXTREME <= n
0	300	300
1	374	674
2	238	912
3	62	974
4	24	998
5	1	999
6	1	1,000
7+	0	
	1,000	

We thus confirm that an "unusual" number of extreme values is 3 or more at the 90% level, and 4 or more at the 97% level.

VII. OBSERVED EXTREME RANKSUM VALUES

RANKSUM IS LESS THAN 23 OR GREATER THEN 79:

COUNTY	RANKSUM
21	20
24	21

The number of extreme ranksum values is TWO, and both are at the LOW end of the range.