A CENSORED NONPARAMETRIC TEST OF Ho: $G = 1-(1-F)^K$

Ronald Regal, SUNY Albany

To indicate why one might be interested in the null hypothesis Ho: $G = 1-(1-F)^K$, consider the data in Table 1 adapted from Restle and Davis (1962) and Davis and Restle (1963). First, 251 subjects were separated into M=163 single (individual) units and N=22 groups units with K=4 subjects in each group. Each of the M+N=185 units was given the same problem to solve. Table 1 gives approximate results for Restle and Davis' gold problem. For details on the relation of these data to the charts in Restle and Davis (1962), see Regal (1975). Table 1 gives the times in seconds of the m=71 individuals and n=17 groups who solved the problem before the time limit Of 678 seconds. The 17 group times are identified by (G). The times of the M-m=92 and N-n=5 groups who had not solved by the time limit are considered to have been right censored by the time limit.

1. Solution Times

62 (G)	241	370	520
92	252	372	528
95	256	375	538
100(G)	259(G)	378	547
101	261	381	548(G)
111	280	383	558
130	290(G)	386	597
131(G)	295	388	611
135	300	390	615
140(G)	310	392(G)	618
141	313	395	620(G)
162	317	399	637
165	320	399(G)	639(G)
168	340	409	649
170	343	418	649(G)
181	346	440(G)	666
181(G)	348	452	667(G)
191	350	459	669
201	352	469	672
210(G)	355	481	675
221	358	489	677
229	361	509	677(G)

(G) = group time
92 individuals and 5 groups did not solve by the time limit of 678 seconds.

To introduce some notation, let F be the distribution function of the time required for a randomly chosen individual to solve working alone, and let G be the distribution function for the time needed by K=4 subjects working as a group. One could test Ho: F=G by a number of nonparametric methods which allow for censoring and ties such as we have in Table 1. A summary of some such methods is given by Gehan (1976).

However, even if one knew, say, that a group of size K=4 is expected to perform better than a single individual, one would still not know whether a problem should be solved sooner by four subjects working together or four subjects working separately. The true test of the effectiveness of the grouping comes in comparing the group solution time to the best (minimum) time of K independently working individuals. The null hypothesis that one is equally likely to receive a solution before any time t from a group of size K or from K independently working individuals is

$$H_0: \; G(t) = 1 - (1 - F(t))^K.$$

Lorge and Solomon (1955) proposed such a model for group problem solving when one only observes the numbers of solving individuals and groups, m and n. Fienberg and Larntz (1971) gave methods for testing the Lorge-Solomon model given such data. The problem here is to develop nonparametric methods of analyzing and testing the Lorge-Solomon type model, H_0: $G = 1 - (1-F)^K$, with timed data containing right censoring and ties.

As first step, define

S_j = # of individual (single unit) solutions among the first j combined solutions.

For the data of Table 1 for example

$S_1 = 0$ $\quad S_{16}$ = unknown } tie with time of 181 seconds
$S_2 = 1$ $\quad S_{17} = 12$
$S_3 = 3$ $\quad S_{88} = 71$
$S_{15} = 11$ $\quad S_{185} = 163$

Results from Koul and Staudte (1972) can be used to give approximations to the distribution of S_j under H_0: $G = 1-(1-F)^K$. Define

$$V_j = j/(M + N)$$

and let V_j^* be the unique value in [0,1] such that

$$V_j = \lambda V_j^* + (1 - \lambda)[1 - (1 - V_j^*)^K]$$

where

$$\lambda = M/(M + N).$$

Then

$$E(S_j) \simeq M V_j^*$$

and for $i \le j$

$$\mathrm{Cov}(S_i,S_j) \approx \left(\frac{MN}{M+N}\right) \frac{(1-V_j^*)^K}{\lambda+(1-\lambda)K(1-V_j^*)^{K-1}} \cdot$$

$$\frac{(1-\lambda)K^2V_i^*(1-V_i^*)^{K-1}+\lambda[1-(1-V_i^*)^K]}{\lambda+(1-\lambda)K(1-V_i^*)^{K-1}}$$

Suitably standardized and extended, the S_j process converges weakly to a normal stochastic process. For details and justifications see Regal (1975). As an example of the approximations consider S_{88} under the conditions of Table 1. Using $V_{88} = 88/185 = 0.4757$, the above results suggest approximations of 68.486 and 2.166 for the mean and variance of S_{88} compared to exact values of 68.498 and 2.174 found through recursive methods (Regal, 1975). Since $S_{88} = 71$ for Table 1, there are more than the expected number of individuals or too few groups compared to expectations under H_0. Hence at the 88th checkpoint the groups are doing worse than expected under the Lorge-Solomon model.

Similar comparisons of S_j to the expected value of S_j under Ho can be made at those values of j for which S_j is known. A graphical presentation of the deviations from the Lorge-Solomon model is provided by the plot of $S_j-E(S_j)$ as a function of j. Figure A gives such a plot for the data of Table 1. A possible interpretation of Figure A is that at the beginning the groups did nearly as well as independently working individuals, but as time went by, the grouping started impeding solution. Since the $\mathrm{Var}(S_j)$ is smaller for small j and large j than for intermediate j, one might wonder how much of the apparent peaking in Figure A can be explained by increased variability. Figure B shows a plot of the standardized variable $(S_j-E(S_j))/(\mathrm{Var}(S_j))^{\frac{1}{2}}$. Figure B lends itself to the same sort of interpretations as Figure A in this case.

Although Figures A and B suggest that the group performance falls short of the performance of an equal number of independently working individuals, we still need an overall test of the Lorge-Solomon model, Ho: $G = 1-(1-F)^K$. One possibility is the statistic

$$\frac{\Sigma(S_j-E(S_j))}{(\mathrm{Var}(\Sigma S_j))^{\frac{1}{2}}} .$$

In the case of no censoring or ties this can be shown to be equivalent to the Wilcoxon rank sum or Mann-Whitney statistic, and results from Lehman (1953) can be used to give the exact mean and variance. See Regal (1975) for details, including a comparison of the normal approximation and the exact distribution.

With ties, including ties due to censoring, one possibility is to make inferences conditional on the observed pattern of ties and assign midvalues. For example in Table 1 there is a tie between an individual and a group for places number 16 and 17, and S_{16} is unknown. Giving S_{16} a value of $(S_{15} + S_{17})$ can be shown to be equivalent to using midranks in the Wilcoxon rank sum test. Graphically, the statistic $\Sigma(S_j-ES_j)$ with midvalues attempts to integrate Figure A extended out to $S_{185}-E(S_{185})\equiv 0$. Using $(S_{15} + S_{17})/2$ for S_{16}, the statistic ΣS_j involves S_{15} and S_{17} each multiplied by 1.50 in this case. The variance of ΣS_j would be figured accordingly, using the approximation given above for $\mathrm{Cov}(S_i,S_j)$. Since the inference is conditional on the pattern of ties in the data, ties between individuals or ties between groups would be treated similarly for variance calculations. For the data of Table 1 the resulting standardized score is 2.31 which corresponds to a 2-sided normal significance level of 0.021. Hence the Lorge-Solomon model would be rejected at the 5% level but not at the 1% level.

In summary, methods have been displayed for graphical presentation of deviations from the Lorge-Solomon type model for timed data and for testing the significance of these deviations. These methods allow for ties, including ties due to censoring by a single common time limit. More complicated forms of censoring can be handled along the line of Mantel (1966).

REFERENCES

Davis, J. H. and Restle, F. (1963). The analysis of problems and prediction of group problem solving. J. Abnorm. Soc. Psychol., 66, 103-116.

Fienberg, S. E. and Larntz, F. K. (1971). Some models for individual-group comparisons and group behavior. Psychometrika, 36, 349-368.

Gehan, E. A. (1976). Analyzing survival data by distribution-free methods. Presented at the Thirty-second Annual Princeton Conference on Applied Statistics.

Koul, H. L. and Staudte, R. G. (1972). Weak convergence of weighted empirical cumulatives based on ranks. Ann. Math. Statist., 43, 832-841.

Lehmann, E. L. (1953). The power of rank tests. Ann. Math. Statist., 24, 23-43.

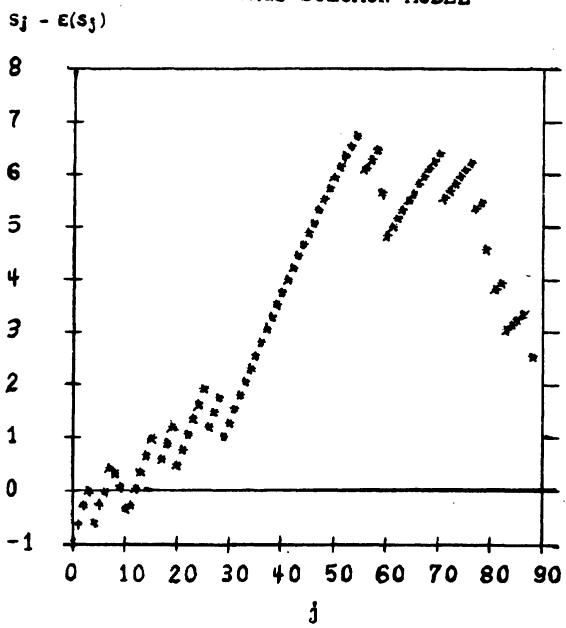
Lorge, I. and Solomon, H. (1955). Two models of group behavior in the solution of eureka-type problems. Psychometrika, 20, 139-148.

Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemotherapy Reports, 50, 163-170.

Regal, R. (1975). Some parametric and

nonparametric methods for individual and group performance. Unpublished Ph.D. thesis, University of Minnesota.
Restle, F. and Davis, J. H. (1962). Success and speed of problem solving by individuals and groups. Psychol. Rev., 69, 520-526.

A. DEVIATIONS $S_j-E(S_j)$ FROM LORGE-SOLOMON MODEL



B. STANDARDIZED DEVIATIONS FROM LORGE-SOLOMON MODEL