ORDERING FOR BUEHLER BOUNDS

M.F. Guerrero[1] and H.T. David[2]

1. Introduction

Buehler upper bounds for a parametric function $H(\theta)$ of a discrete distribution F with finite sample space and parameter space Ω are defined, for a specified $\alpha \epsilon (0, 1)$, as:

$$b(x^{(k)}) \equiv \sup \{H(\phi): \ F(x^{(k)}; \theta) > \alpha \}$$

where $x^{(k)}$ is the k^{th} sample point in a given ordering or labeling of the points in $\mathcal{X}$. Guerrero and David (1985) established that the bounds $b(x^{(k)})$ are monotone nondecreasing and that among all similarly ordered bounds for $H(\theta)$, the Buehler bounds constitute a family of uniformly shortest ones.

In the computation of such bounds, one is faced with the initial task of ordering the points in $\mathcal{X}$. Guerrero and David (1985) illustrate how some orderings provide more reasonable bounds than others, indicating the need for care in the way one decides to order the sample space.

If the problem at hand is one of improving a set of upper confidence bounds, say $\{d(x)\}$, that are provided by some other confidence procedure, and magnitude is the criterion for improvement, then one would order according to the magnitude of the initial bounds; i.e., x_1 precedes x_2 whenever $d(x_1) < d(x_2)$. If this ordering provides upper Buehler bounds, then the Buehler bounds will be uniformly smaller than the initial bounds.

In some applications, a reasonable initial confidence procedure may be suggested by the problem at hand, and finding a suitable ordering would be tantamount to computing the bounds provided by this initial confidence procedure. Consider, for example, Buehler's recommended ordering for the problem of obtaining upper bounds for the function $H(p_1, p_2, \ldots, p_n)$, where the p_K's are the parameters of independent binomial variates X_K. The initial upper confidence bounds are obtained as follows: for the observation $X = (x_1, x_2, \ldots, x_n)$ let

---

1/ Assistant Professor in Statistics, Institute of Mathematical Sciences and Physics, College of Arts and Sciences, University of the Philippines at Los Banos, College, Laguna.

2/ Professor in Statistics, Department of Statistics and Statistical Laboratory, Iowa State University, Ames, Iowa.

$$d(x_1, x_2, \ldots, x_n) = \max \{ p_1 x p_2 x \ldots x p_n : 0 \le p_k \le p(x_k) \}$$

where $p(x_k)$ is the $(1-\alpha)^{1/n}$ upper confidence bound for p_k given by:

$$p(x_k) = \sup \{ p_k : \text{PROB } X_k \le x_k | p_k \} = 1 - (1-\alpha)^{1/n}$$

Note that these individual upper bounds are the Buehler bounds for the individual parameters p_k based on the "natural" ordering of the sample space $\{0, 1, \ldots, n_k\}$ of x_k. These upper bounds will be shown to be the smallest possible Buehler upper bounds for p_k for the given confidence level. The choice of $\{d(x)\}$ as initial upper confidence bounds for H is therefore intuitively appealing from the point of view of bounding a function which is monotone increasing in each argument by computing its value at the smallest-possible upper bounds for the individual arguments.

As a general rule, one would want the smaller upper bounds to correspond to sample points that are assigned large probability by those parameter points which make the function H small; similarly, one would want the larger upper bounds to correspond to sample points that are assigned large probability by the parameter points that make H large. This indicates that any ordering rule must take into account (i) the nature of the parametric function H and (ii) the nature of the likelihood function of the random vector in consideration. The likelihood function determines the nature of the distribution function of the ordered sample space, which in turn determines the nature of the monotone regions.

In the succeeding sections, we exhibit a condition under which an optimal class of orderings, or a single optimal ordering, may be identified in the case of monotone likelihood ratio families. We also provide a weaker sort of "sequentially optimizing" ordering for the general finite case. Finally, we point out that parametric functions which express the reliability of monotone systems share, with the product function, the property that a reasonable "initial" confidence procedure can be made to furnish an ordering.

2. Optimal Ordering Procedures

2.1 Optimality for Monotone Likelihood Ratio Families

In what follows, we assume that the random variable X is scalar and takes values in a finite sample space $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$ with:

$$x_1 < x_2 < \ldots < x_N.$$

We also assume that the parameter space is a subset of the real line. The next three definitions are essential to the development of the optimality criterion presented here.

Definition 1: The probability mass function $f(x;\theta)$ of X is said to have monotone likelihood ratio in x if $x' < x''$ implies that the ratio of likelihood functions:

$$\frac{f(x'';\theta)}{f(x';\theta)}$$

is nondecreasing in θ.

Definition 2: Let $B = \{b(x_1), b(x_2), \ldots, b(x_N)\}$ and $B' = \{b'(x_1), b'(x_2), \ldots, b'(x_N)\}$ be two systems of $(1-\alpha)$ 100% confidence bounds for θ. Let

$$b_1 \leq b_2 \leq \cdots \leq b_N$$

and

$$b_1' \leq b_2' \leq \cdots \leq b_N'$$

be, respectively, the elements of B and B' ordered by magnitude. Then B is said to dominate B' if $b_i \leq b_i'$ for all i.

Naturally, one would prefer not to use a system of upper bounds that is dominated by another system.

Definition 3: A sample space ordering 0 corresponding to a system of upper confidence bounds $B = \{b(x_1), b(x_2), \ldots, b(x_N)\}$ is one that corresponds to the magnitudes $\{b_1, b_2, \ldots, b_N\}$.

For example, if $N = 4$ and

$$b(x_4) < b(x_1) = b(x_3) < b(x_2)$$

then $0 = (x_4, x_1, x_3, x_2)$ is a sample space ordering corresponding to B. Note that $0' = (x_4, x_3, x_1, x_2)$ is also a sample space ordering corresponding to B. This illustrates that if some of the upper bounds are equal, then we have an equivalence class of orderings corresponding to system B.

Consider the problem of constructing upper Buehler bounds for a real-valued parameter θ of a discrete distribution having monotone likelihood ratio. We now derive an optimal class of orderings for this problem.

For a specified pair (x_i, x_j) of points in the sample space, with $x_i < x_j$, define:

$$\theta_{ij} = \sup \{\theta: f(x_i;\theta) = f(x_j;\theta)\}$$

if $f(x_i;\theta) = f(x_j;\theta)$ for at least one θ, then θ_{ij} is the last point at which the curves $f(x_i, \theta)$ and $f(x_j; \theta)$ intersect. Since $f(x;\theta)$ has monotone likelihood ratio, it follows that:

$$
\begin{aligned}
&f(x_i;\theta) > f(x_j;\theta) \quad \text{if } \theta < \theta_{ij} \\
\text{and} \\
&f(x_i;\theta) < f(x_j;\theta) \quad \text{if } \theta > \theta_{ij}
\end{aligned}
\tag{2.1}
$$

Note that if $f(x_i;\theta) \neq f(x_j;\theta)$ for all θ, then it must be true that $f(x_j;\theta) > f(x_i;\theta)$ for all θ.

<u>Theorem</u>: Suppose $f(x;\theta)$ has monotone likelihood ratio in x. Let (x_i, x_j) be any pair of sample points for which $x_i < x_j$ and $f(x_i;\theta) = f(x_j;\theta)$ for some θ. Suppose that for a system B of $(1-\alpha)100\%$ upper confidence bounds for θ, an ordering O corresponding to B places x_j before x_i. Let B* be a Buehler system for θ. Also, let B** be a Buehler system of upper bounds for an ordering O' derived from O by interchanging the positions of x_i and x_j. If $f(x_i;\theta_{ij}) > \alpha$, then B** dominates B*.

<u>Proof</u>: Let $\{b_1^*, b_2^*, \ldots, b_N^*\}$ and $\{b_1^{**}, b_2^{**}, \ldots, b_N^{**}\}$ denote the upper Buehler bounds ordered by magnitude in B* and B**, respectively. Suppose x_j is in the k^{th} position and x_i is in the m^{th} position $(k < m)$ with respect to ordering O. Then,

$$b_l^* = b_l^{**} \quad \text{for } l < k \text{ and } l \geq m.$$

We now show that for $k \leq l < m$, $b_l^{**} \leq b_l^*$. Let x be a point in any one of these positions. Then the distribution function at x under O and O' are, respectively:

$$F(x;\theta) = G(\theta) + f(x_j;\theta)$$

and

$$F'(x;\theta) = G(\theta) + f(x_i;\theta)$$

where $G(\theta)$ is the sum of the likelihood functions of sample points that are in position 1 through $k-1$ and $k+1$ through l. Hence, for $k \leq l \leq m$

$$
\begin{aligned}
&b_l^* = \sup\{\theta : F(x;\theta) > \alpha\} \\
\text{and} \\
&b_l^{**} = \sup\{\theta : F'(x;\theta) > \alpha\}.
\end{aligned}
$$

Since $f(x_j;\theta_{ij}) > \alpha$, it follows that $F(x;\theta_{ij}) > \alpha$ and $F'(x;\theta_{ij}) > \alpha$. Hence, both b_l^* and b_l^{**} must be greater than or equal to θ_{ij}. Furthermore, (2.1) implies that:

$$F(x;\theta) \quad F'(x,\theta) \quad \text{for } \theta > \theta_{ij}$$

and it follows that $b_1^{**} \leq b_1^*$. Q.E.D.

The following corollaries are immediate consequences of the preceding theorem. They establish that Buehler upper bounds for the parameter of a discrete distribution that has monotone likelihood ratio obtained under certain classes of orderings will dominate those obtained under orderings outside these optional classes.

Corollary 1: Suppose $f(x;\theta)$ has monotone ratio in x. Let (x_i, x_j) be two sample points for which $x_i < x_j$ and $f(x_i;\theta) = f(x_j;\theta)$ for at least one θ. Suppose $f(x_i;\theta_{ij}) > \alpha$. Then, the orderings that order pairs (x_i, x_j) satisfying these properties in the natural way form a complete class.

Corollary 2: Suppose $f(x;\theta)$ has monotone likelihood ratio in x. If $f(x_i;\theta_{ij}) > \alpha$ for all pairs (x_i, x_j) for which $x_i < x_j$ and $f(x_i;\theta) = f(x_j;\theta)$ for at least one θ, then the set of natural orderings form a complete class.

Since the binomial density has monotone likelihood ratio, if α is such that $f(x_i;\ p_{ij}) >$ for all pairs $(x_i;\ x_j)$ then, Corollary 2 establishes that the quantities:

$$u(k) = \sup \{ p: \text{PROB} [X \leq k | p] = \alpha \}$$

for k = 0, 1,..., n are the smallest-possible $(1-\alpha)100\%$ upper Buehler bounds for the binomial parameter p. Furthermore,

$$l(k) = \inf \{(1-p): \text{PROB} [X \leq k | p] =$$

for k = 0, 1,..., n provides the largest-possible $(1-\alpha)100\%$ lower Buehler bounds for p.

In the next section, we propose a generally applicable ordering suggested by and specializing to the optimal natural ordering for exponential families, which guarantees a certain weaker sequential type of optimality.

2.2 A Sequentially Optimizing Ordering

A reasonable requirement to impose in obtaining upper Buehler bounds for a parametric function H is that, to the extent it is possible, the upper bounds $b(x^{(i)})$ should be as small as possible for all $x^{(i)}$. To ensure that this is true for $i = 1$, we would want to minimize the function:

$$p_1(x) = \sup \{H(\theta): f(x;\theta) > \alpha \}$$

where $f(x;\theta)$ is the likelihood function of the random vector X, over the sample space $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$. Suppose $x = y_1$ solves this minimization problem. Then to find the second smallest-possible upper bound for H given that y_1 is the first point in the ordering, we would minimize the function:

$$p_2(x) = \{\sup H(\theta): f(y_1; \theta) + f(x;\theta) > \alpha \}$$

over the set $\mathcal{X} - y_1$. Let y_2 denote a solution to this minimization problem. To find the ith smallest-possible Buehler upper bound, we proceed sequentially, as follows: given that $y_1, y_2, \ldots, y_{i-1}$ provide the first, second,..., (i-1)st smallest-possible Buehler upper bounds for H, the ith smallest-possible Buehler upper bounds is obtained by minimizing the function:

$$p_i(x) = \sup \{H(\theta); \sum_{j<i} f(y_j;\theta) + f(x; \theta) > \alpha \}$$

over the set $\mathcal{X} - \{y_1, y_2, \ldots, y_{i-1}\}$.

By construction, the ordering of the sample space provided by this sequential procedure has $x^{(i)} = y_i$ and provides a set of smallest-possible Buehler upper bounds for each i.

3. An Application: Confidence Bounds for Reliability Functions of Monotone Systems

3.1 Background

Consider a system with n components where each of the components assumes two states: a functioning state or a failed state. Let p_i denote the probability that the ith component is in its functioning state. If the system has monotonic structure, then its reliability function, to be denoted by $H(p_1, p_2, \ldots, p_n)$, is nondecreasing in each p_i.

Suppose that for component k, n_k independent bernoulli trials are observed and x_k successes are recorded. If data for each component are independently obtained, then the observations $(x_1, x_2, \ldots, x_n)$ are values of the random vector $X = (X_1, X_2, \ldots, X_n)$ where the X_k's are independent binomial (n_k, p_k) variables.

For various kinds of systems, many researchers have proposed different ordering functions. A natural choice would be a point estimator for H(p). For the reliability function of a series system, for example, the maximum likelihood estimator $H(p) = \Pi_i (x_i/n_i)$ and certain modifications of it have been suggested. However, most of the results have been geared towards obtaining procedures which have the Buehler optimality property asymptotically but not necessarily for any fixed sample size. Easterling's (1972) modified maximum likelihood method and Madansky's (1965) linearization method provide an illustration of this approach. Epstein (1967) considered the problem of confidence sets for the product of two binomial parameters and considered the two ordering functions:

$$g_1(x) = x_1 x_2 / n_1 n_2$$

and

$$g_2(x) = (x_1 + 1)(x_2 + 1)/n_1 n_2$$

and concluded that the second was preferable to the first, since the partition of the sample space induced by the second is finer than that induced by the first.

In what follows, we present a confidence procedure for the reliability function of monotone systems. The procedure is based on binomial attribute data collected from life tests on the components of the system and we will establish that it is optimal in a certain class.

3.2 A Confidence Procedure for the Reliability Function of Monotone Systems

A reasonable set of easily computed $(1-\alpha)100\%$ upper confidence bounds for the reliability function H of a monotone system is provided by the construction:

$$d(x_1, x_2, \ldots, x_n) = \max \{H(p): \underline{p}_k \leq p_k \leq \bar{p}_k \} \quad (3.1)$$

where $[\underline{p}_k, \bar{p}_k]$ is a $(1-\alpha)^{1/n}$ two-sided confidence interval for p_k.

Since H is nondecreasing in each argument, its maximum over a cartesian product of intervals $[\underline{p}_k, \bar{p}_k]$ is attained at the upper bounds $\bar{p}_k$ of each argument. Hence, upper bounds for H would be small in magnitude if the upper bounds $\bar{p}_k$ are small. Consider the class of confidence procedures for H that utilize functions of the form defined in (3.1). Then, for a specified confidence level, those procedures which use one-sided upper intervals $[0, \bar{p}_k]$ for each parameter would provide upper bounds for H that are smaller than those obtained under procedures that use two-sided intervals. Now, suppose the Buehler upper bounds for p_k:

$$b(x_k) = \sup \{ p_k : \text{PROB} [Y_k \leq x_k | p_k] = 1 - (1-\alpha)^{1/n} \}$$

are employed in the construction defined by the function $d(\cdot)$. Since these individual Buehler upper bounds are optimally shortest, the upper bounds for H provided by d would then have the appealing property that they are smallest among all upper bounds for H obtained over cartesian product regions $[\underline{p}_k, \bar{p}_k]$. In the sense of Pavlov (1977 a and 1977b), these bounds are locally optimal and admissible. On the basis of these properties and the ease of computation, these upper bounds would be our recommendation for ordering the sample space for upper Buehler bound construction.

4. References

Easterling, R.G. 1972. Approximate Confidence Limits for System Reliability. Journal of the American Statistical Association. 67: 220-222.

Epstein, J.E. 1967. Upper Confidence Limits on the Product of Two Binomial Parameters. M.S. Thesis, Cornell University.

Guerrero, M.F. and H.T. David. 1985. Buehler Confidence Bounds. Philippine Statistician, 34(1): 86-105.

Madansky, A., 1965. Approximate Confidence Limits for the Reliability of Series and Parallel Systems. Technometrics. 7:495-503.

Pavlov. I.V. 1977a. Admissible Interval Estimate for Functions of Several Unknown Parameters I, Engineering Cybernetics. 15(3): 45-54.

Pavlov. I.V. 1977b. Admissible Interval Estimate for Functions of Several Unknown Parameters II. Engineering Cybernetics. 15(5): 52-61.