THE BOUNDED-COUNT METHOD FOR ANALYSIS OF LEK COUNTS

Douglas H. Johnson[1], Clait E. Braun[2], and Michael A. Schroeder[3]

[1]U.S. Geological Survey, Northern Prairie Wildlife Research Center, University of Minnesota, Saint Paul, Minnesota 55108
[2]Grouse Inc., 5572 North Ventana Vista Road, Tucson, Arizona 85750
[3]Washington Department of Fish and Wildlife, PO Box 1077, Bridgeport, Washington 98813

Abstract. Counts of displaying males at leks are the traditional method used to monitor populations of greater sage-grouse (*Centrocercus urophasianus*). Often, multiple daily counts are made during a single breeding season, and the maximum of those counts is used as an index to the population size. That maximum value clearly is a biased (low) estimate of the number of males because, as indicated by studies of radio-marked birds, some males in the population that use the lek may be either absent or present but not detected during the lek count. Here we consider an alternative, the Bounded-count (or jackknife) Estimator, which is always at least as large as the largest count. It is simple to compute, being twice the largest count minus the second-largest count, and allows confidence limits to be computed. We discuss the rationale for the estimator and apply it to simulated and actual data sets. When the recorded counts, or even a portion of them, are distributed uniformly between zero and the true population size, the Bounded-count Estimator performs well. In somewhat more realistic simulations with varying but low rates of occupancy and detection, the Bounded-count Estimator reduced the bias but did not result in values more closely correlated with the true population size.

Introduction

The greater sage-grouse (*Centrocercus urophasianus*) is a key species endemic to the sagebrush (*Artemisia* spp.) ecosystem of North America. Its range and population size have diminished in recent decades, and the species was proposed as a candidate for listing under the Endangered Species Act in the United States (U.S. Fish and Wildlife Service 2005). Populations of greater sage-grouse have been monitored by a variety of methods including brood surveys, harvest metrics, and age ratios in harvested samples, but the most common and widespread method involves counts of birds at strutting grounds (leks) in spring.

The methodology for conducting lek counts has become more standardized (Autenrieth et al. 1982) and consistent in recent years (Connelly et al. 2003). Details of the method are provided elsewhere (Connelly et al. 2003); for our purposes we note only that in any given year each lek is surveyed from zero to several times. Traditionally the maximum count of males from all the surveys in a year is used as a measure of the size of the population associated with that lek. For example, the Coalmont lek in Colorado was surveyed 5 times in 1989, yielding counts of 58, 18, 63, 67, and 81 males (Table 1). The Maximum Estimate resulting from those surveys then is 81 males. The objective of this paper is to consider an alternative to the Maximum as an estimator. Specifically, we examine the Bounded-count Estimator by exploring its properties analytically, with simulated data, and with actual counts of sage-grouse. Although we consider the estimator with regard to sage-grouse, its applicability may be more general.

Methods and Results

The count of males at a lek can depart from the true population size either because the occupancy rate (the probability that a male associated with a lek is present at the lek at the time of the survey) is less than 1 (resulting in an availability bias; Buckland et al. 2004), or the detection rate (the probability that a male present at the lek is recorded by the observer) is less than 1 (perception bias). Further, counts can vary dramatically within a single season (Table 1) due to variation in these rates. Influences on the occupancy rate include the date within a season, the time of the survey, weather conditions, variation in behavior of males, presence of predators near a lek site, and other disturbances (see Johnson and Rowland [2006] for summary). The detection rate depends upon the skill and

Table 1. Counts of male greater sage-grouse at known lek sites in North Park, Colorado, in 1989 (Braun, pers. obs.). Each lek was surveyed 1-5 times during the breeding season.

Lek site	Count 1	Count 2	Count 3	Count 4	Count 5
Alkali Lake	40	43	49	50	
Arapahoe	8	29	25	23	30
Aspen	11	10			
Bighorn	0	0	0		
Boettcher Junction	32	49	39	40	
Buteo	0	13	0	0	
Canuck	0	0			
Case Flats	0				
Cheyenne	53	59	11		
Coalmont	58	18	63	67	81
Deer Creek	13	12	11	6	
Delaney Butte	34	21	30	32	20
Denmark	12	45	32	21	
Eagle	0	0	0		
Fish Hatchery	14	26	36	35	
Hawk	4	6	0	0	
Hound	0	0			
Lost Creek # 1	23	21	20	18	
Migan	0	7	10	0	
Ortega	0	0	0		
Owl Creek	0	0			
Perdiz	8	0	8	0	5
Peregrine	0	0			
Prague	0	0			
Pronghorn	0	0	0		
Ptar	0	0	0		
Railroad	13	10	12	0	12
Ram	0	0	0		
Raven	33	24	31	33	
Ridge Road	38	32	6	3	
Riley	3	4	0	0	
Spring Creek # 1	46	39	43	48	61
Spring Creek # 2	0	0	0		
Thrasher	15	0	4		
Turkey	12	11	5	3	11
Ute	0	0	0		

perseverance of the observer, distance of the observer from the lek, quality and use of optics, habitat features, weather conditions, and other factors.

Biologists have sought to reduce the error in the counts by 2 general types of procedures, design control and statistical control. Design control involves the standardization of methodology so that variables that influence the count are as fixed as feasible. For example, because counts vary by time of day, standardization restricts surveys to a brief period during early morning when lek attendance is thought to be highest. Similar restrictions apply to seasonal and weather effects. Standardization is important, but as Caughley and Goddard (1972: 136) noted, "Variability can be reduced by tight experimental design, but there usually remains a residual puddle resisting all efforts to drain it."

Recognizing that the maximum count (hereinafter termed the Maximum Estimator) does not accurately portray the number of male sage-grouse associated with a lek, some investigators have suggested remedies. Braun (unpublished annual lek count summaries, Colorado Division of Wildlife), Authenrieth et al. (1982) and Emmons and Braun (1984) specifically suggested one form of statistical control: divide the maximum count by 0.75, an estimate of the combined occupancy and detectability rates. That is equivalent to multiplying the Maximum Estimate by 1.33. For the Coalmont example, Braun's estimate would be 81/0.75 = 108 males. Earlier, Dalke et al. (1963) compared the highest of 3 counts 5 days apart, as was conducted operationally, to the highest of many counts conducted daily and indicated that the latter was 19 percent higher than the former; that would suggest multiplying the Maximum Estimate, based on the customary three counts, by at least 1.19. Jenni and Hartzler (1978) concluded that 3 counts during the period of peak attendance would yield an estimate within 10 percent of the seasonal maximum; essentially that might suggest that the Maximum Estimate should be multiplied by at least 1.11.

We consider here what is known as the Bounded-count or jackknife method (Robson and Whitlock 1964, Regier and Robson 1966, Overton 1969, Routledge 1982). Numerically, the estimator is simply twice the largest count minus the second-largest count. We use the following notation. Define the true but unknown population associated with a given lek to be N males. Suppose in a given year that k surveys are made of that lek, resulting in counts of $X_1, X_2, \ldots, X_k$ males. Define the combined occupancy-detection rate on the ith survey to be p_i, $i = 1, \ldots k$. Then $X_i = p_i N$ for all i. Define the ordered values of X_i to be $X_{(i)}$, where $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(k)}$. The fundamental assumption of the Bounded-count Estimator is that the X_i values are distributed uniformly in the interval (0, N). This is equivalent to assuming that the combined occupancy-detection rates p_i are distributed uniformly in the interval (0, 1).

We note the assumption that the Maximum Estimator ($X_{(k)}$) is accurate is equivalent to assuming that $\max\{p_i\}=1$ for some i; that is, that on one of the surveys, all birds associated with a lek are present and recorded. Braun's estimator essentially is based on the assumption that $\max\{p_i\} = 0.75$.

The Bounded-count Estimator can be motivated as follows. Imagine the line between zero and N. Within this line, k uniform random deviates are drawn. These divide the line into k+1 segments, whose expected values are identical. Thus, in expectation, the difference between N and $X_{(k)}$ is the same as the difference between $X_{(k)}$ and $X_{(k-1)}$. Equating these terms and solving for N yields the expression:

$$\hat{N} = X_{(k)} + (X_{(k)} - X_{(k-1)}) = 2X_{(k)} - X_{(k-1)}.$$

For the Coalmont example, $\hat{N}$ = 2Í81 – 67 = 95.

Approximate $(1 - \alpha)$ confidence limits for $\hat{N}$ were developed by Robson and Whitlock (1964): lower limit = $X_{(k)}$, and upper limit = $X_{(k)} + (X_{(k)} - X_{(k-1)})\,(1 - \alpha)/\alpha$. For the Coalmont example, 90% confidence limits (α = 0.10) are 81 and 81 + (81 – 67) Í (0.90/0.10) = 207. The breath of this confidence interval (81, 207) seems to be typical of Bounded-count Estimators.

In the situation in which the 2 largest counts have the same value ($X_{(k)} = X_{(k-1)}$), the upper and lower confidence limits above would be identical. To remedy this problem, Routledge (1982) suggested replacing the upper limit by $X_{(k)} + f(1 - \alpha)/\alpha$, where f is the number of tied high counts.

In 200 simple simulations of 6 Uniform (0, 100) random deviates, the Bounded-count Estimator did reduce the bias compared with the Maximum Estimator, on average from -14 to +1 (Table 2).

Examination of actual lek counts (e.g., Table 1) suggests that they are not uniformly distributed, a violation of the assumption of the Bounded-count method. Often there are numerous zeroes or other very low counts along with a variety of larger counts. The low counts likely reflect results of surveys done under imperfect conditions, or when occupancy was reduced due to poor conditions, a disturbance, or some other factor. We explored the effectiveness of the Bounded-count Estimator when several of the counts were zero. Specifically, we conducted simulations in which 3 of the counts were from a Uniform (0, 100) distribution and 3 of the counts were zeroes. The bias of the Maximum Estimator (-25) was, as expected, even greater than in the previous simulation (Table 3); the Bounded-count Estimator was virtually unbiased, although its variance was greater, also as expected.

We further examined the Maximum and Bounded-count estimators under conditions in which both the number of surveys per year and the occupancy-detectability rates varied. We simulated a 4-year monitoring program, in which the actual population increased by 5 males each year, from 20 to 35 males. Two to 7 surveys were conducted each year, and the fraction of males occupying leks and being counted was either low or high, under 2 scenarios. Low values were randomly drawn from a Beta distribution with a mean of 0.4 and a standard deviation of 0.28. High values were drawn from a Beta distribution with a mean of 0.8 and a standard deviation of 0.23. We performed 200 simulations under each scenario and calculated the average difference between the true population size and each estimate as a measure of bias. We found that, when occupancy-detectability was low, the Bounded-count Estimator indeed had less bias than the Maximum Estimator (-7 vs. -31%; Table 4). When occupancy-detectability was high, both estimators were fairly accurate.

We suspected that the Bounded-count Estimator might provide a better index to the population that could be useful in monitoring. We explored this idea with the simulation described above by examining the correlation coefficient over the four years of the survey between the true population size and both the Maximum and Bounded-count estimates. Large positive values of the correlation coefficient would indicate that an estimator provides a good index to the population size. Under conditions of low occupancy-detectability, however, neither estimator strongly correlated with the true population size (r = 0.68 and 0.53 for Maximum and Bounded-count estimators, respectively); conversely when occupancy-detectability was high, both estimators strongly correlated with true population size (r = 0.95 and 0.87). Perhaps surprisingly, the Maximum Estimator was a somewhat better index to the population, at least in terms of correlating with the true population, than was the Bounded-counts Estimator (Table 4).

Table 2. Maximum and Bounded-count estimates from 200 simulations of 6 random deviates from a Uniform (0, 100) distribution. The true value of the parameter being estimated is 100.

	Estimator	
	Maximum	Bounded-count
Mean	86	101
SD	12	20

Table 3. Maximum and Bounded-count estimates from 200 simulations of 3 random deviates from a Uniform (0, 100) distribution and 3 zero values. The true value of the parameter being estimated is 100.

	Estimator	
	Maximum	Bounded-count
Mean	75	99
SD	18	30

Discussion

Some properties of the Bounded-count Estimator are clear. It could be biased high if individuals on a lek can be counted twice. The Bounded-count Estimator will be biased low if some birds avoid a lek because certain other birds are present. Then $X_i \ll N$ for all i. In particular, if, say, no more than 70% of the males associated with the lek are likely to occur on the lek at any given time, then the Bounded-count Estimator will estimate, not N, but $0.70N$. These two problems can affect other estimators, of course, including the Maximum Estimator. More specific to the Bounded-count Estimator, it will have a large error if there is no more than one good count (close to N) and the others are bad (close to zero). Then $\hat{N}$ will be approximately 2N – 0 = 2N. The Bounded-count Estimator will perform poorly if p_i is nearly constant for all surveys (which seems unlikely; Walsh et al. 2004); it then will yield $\hat{N} \sim p^*N$, where p^* is the average value of the occupancy-detectability rates. In contrast to many statistical estimators, the error likely will be greater when there are more surveys.

The Bounded-count Estimator should perform well if values of X_i are uniformly distributed between zero and

Table 4. Average bias (and percent of true value) and correlation coefficient with true population size for Maximum and Bounded-count estimators, under conditions of low and high rates of occupancy-detectability.

Statistic	Low occupancy-detectability		High occupancy-detectability	
	Maximum	Bounded-count	Maximum	Bounded-count
Bias	-8.4 (31%)	-1.9 (7%)	-1.0 (4%)	+2.0 (-8%)
Correlation	0.68	0.53	0.95	0.87

N. More useful in real-world situations is the fact that it seems to do well even if only some of the X_i values are so distributed.

In a real-world application in which sage-grouse were radio-marked and their locations could be determined accurately, Walsh et al. (2004) found that the Bounded-count method consistently underestimated the true population size, indicating that it does not totally correct for the bias. Because the Bounded-count Estimate is always at least as large as the usual estimate (maximum count), it clearly is less biased than that estimator. Walsh et al. (2004) also found that the specified confidence intervals for Bounded-count Estimates did not attain their nominal coverage levels.

While the Bounded-count Estimator is far from perfect, it does tend to reduce the bias associated with the Maximum Estimator. It is worthy of further investigation with both simulated data but especially with actual field studies to examine the extent of the improvement. Further, additional studies might shed light on the actual distribution of p_i values. If they follow some general distribution other than the Uniform, perhaps an estimator could be developed to capitalize on that fact.

Acknowledgments

We are grateful to R. R. Cox, Jr., K. P. Reese, M. M. Rowland, M. H. Sherfy, R. K. Steinhorst, and G. C. White for useful comments on previous versions of this paper.

Literature Cited

Autenrieth, R., W. Molini, and C. Braun. 1982. Sage grouse management practices. Western States Sage Grouse Committee Technical Bulletin **1**. Twin Falls, Idaho, USA.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. Advanced distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, UK.

Caughley, G., and J. Goddard. 1972. Improving the estimates from inaccurate censuses. Journal of Wildlife Management **36**:135-140.

Connelly, J. W., K. P. Reese, and M. A. Schroeder. 2003. Monitoring greater sage-grouse habitats and populations. Station Bulletin **80**. College of Natural Resources Experiment Station, University of Idaho, Moscow, Idaho, USA.

Dalke, P. D., D. B. Pyrah, D. C. Stanton, J. E. Crawford, and E. F. Schlatterer. 1963. Ecology, productivity, and management of sage grouse in Idaho. Journal of Wildlife Management **27**:811-840.

Emmons, S. R., and C. E. Braun. 1984. Lek attendance of male sage grouse. Journal of Wildlife Management **48**:1023-1028.

Jenni, D. A., and J. E. Hartzler. 1978. Attendance at a sage grouse lek: implications for spring censuses. Journal of Wildlife Management **42**:46-52.

Johnson, D. H., and M. M. Rowland. 2007. The utility of lek counts for monitoring greater sage-grouse. Pages 15-23 *in* K. P. Reese and R. T. Bowyer (editors). Monitoring populations of sage-grouse. College of Natural Resources Experiment Station Bulletin 88, College of Natural Resources, University of Idaho, Moscow, Idaho.

Overton, W. S. 1969. Estimating the numbers of animals in wildlife populations. Pages 403-456 *in* R. H. Giles, editor. Wildlife Management Techniques, third edition, revised. The Wildlife Society, Washington, DC, USA.

Regier, H. A., and D. S. Robson. 1966. Estimating population numbers and mortality rates. Pages 31-66 *in* S. D. Gerking, editor. The biological basis of freshwater fish production. Blackwell Scientific Publications, Oxford, UK.

Robson, D. S., and J. H. Whitlock. 1964. Estimation of a truncation point. Biometrika **51**:33-39.

Routledge, R. D. 1982. The method of bounded counts: when does it work? Journal of Wildlife Management **46**:757-761.

US Fish and Wildlife Service. 2005. Endangered and threatened wildlife and plants; 12-month finding for petitions to list the Greater Sage-Grouse as threatened or endangered; proposed rule. Federal Register **70**:2244-2282

Walsh, D. P., G. C. White, T. E. Remington, and D. C. Bowden. 2004. Evaluation of the lek-count index for Greater Sage-Grouse. Wildlife Society Bulletin **32**:56-68.