Biases in the Step-Point Method on Bunchgrass Ranges

DAVID STRAUSS AND DONALD L. NEAL

Abstract

During a study of grazing in the Great Basin, using the grazed plant method used for measuring utilization, we noticed a bias in the step point technique used to select plants. Subsequent mathematical calculations showed both the direction and order of magnitude of the bias: the step point method overestimates the *number* of large plants and underestimates the number of small plants; when used to estimate the *basal area* of bunchgrasses the method overestimates the area of small plants. The calculations were tested and verified on maps of real and artificial bunchgrass populations. When the plants are distributed at random, the biases can be removed. For estimation of numbers, one should select the plant whose center (rather than perimeter) is closest to the toe point to eliminate bias by size class; for estimations of basal area, the numbers in each size class should be scaled by the area of plants of that class.

The point method of sampling vegetation utilizes a sharp pointed pin to select a "dimensionless" plot. The method, which has many advantages and has been adapted to a wide variety of sampling objectives, has been in use since the 1920's (Tansley and Chipp 1926). Levy and Madden (1933) first described a frame holding 10 sliding vertical pins for estimating composition of pastures. Tinney et al. (1937) described the advantages of inclining the pins at 45 degrees. The use of a single pin for describing herbaceous vegetation was first reported by Tansley and Chipp (1926). Costello and Schwan (1946) used the single pin placed at the end of an observer's toe (step-point) when pacing a transect to sample condition and trend of rangeland. Roach (1950) described how to use the single pin in sampling to measure utilization by the ungrazed plant method. The use of the step-point method to determine cover and composition in the California annual type was described in detail by Evans and Love (1957). Since the method is as flexible as the man's ability to walk and count, the step-point has been modified for many purposes.

Goodall (1952) reported that the method could overestimate cover if the pins were not kept very sharp. Blunt pins begin to take on dimension and are more likely to "hit" a plant than sharp pins.

While using the step-point technique to select plants for the grazed plant method of utilization measurement (anonymous 1969), we observed what we thought were other biases in the technique. We were studying the effect of grazing on the Great Basin bunchgrass type. At each of two sites, one in Modoc County, Calif., and one in Lake County, Ore., we had 8 5-acre units in which cattle were grazed for 1 to 5-day periods, 5 to 6 times each season. At the end of each grazing period the grazed plant method was used to estimate utilization. Five 450-foot long transects were established in each unit and 20 step-points were taken on each transect. The plants were classified into several size classes; we were interested both in the numerical proportions of the size classes and their proportions of the total basal area.

If only direct step-point "hits" (i.e. when a plant is struck by the pin) were recorded, bias in the numerical proportions would probably be small. However, on low density ranges such as our western bunchgrass ranges one would be recording as many as 20 misses for every hit. The number of step-points required to get an acceptable sample for species composition or utilization would be very high. Therefore, if no plant is hit directly, the plant nearest to the toe point is recorded. This is the basis for the large errors we found; our samples contained a much higher ratio of large plants (which appeared to be preferred by the cattle) than our cover measurements taken by the area-list method (Pearse 1935) indicated. The result was a higher utilization figure than we thought correct. Conversely, when estimating basal area with the step-point method we found a substantial bias in favor of the small plants.

To check all this, we started recording each plant by size class as well as utilization state. After 3 seasons and an additional 30,000 plants were recorded by utilization state and size class, we became convinced that the biases were serious and should be studied further.

Methods

Our first step was to develop a mathematical model of a bunchgrass population and a formula for expected results with step-point sampling. For each size we calculated the number of plants expected, using a variety of plant sizes and densities of cover.

Next, we created maps of artifical populations, with randomly distributed plant centers, and simulated step-point sampling. We noted the number of plants in each size class in the sample. These numbers were compared with those expected on the basis of our theoretical calculations, to see whether the latter were reasonable. To investigate the bias, the numbers were also compared with those expected on the basis both of the true numerical proportions in the population, and the relative proportions of the total basal area.

Lastly, we used vertical photographs of real bunchgrass populations to create maps at convenient scales. Again, transects were drawn, step-point samples simulated, and the numbers of each size class noted. As before, these were compared to values calculated theoretically and also to values based on the true proportions and basal areas.

Two assumptions were made in the creation of the model.

- (a) That there are k size classes, with all plants being circular
- and all plants in the ith class being of the same radius, r_i.
- (b) The centers of all plants are scattered independently and randomly (without clustering or segregation). Let ρ_i be the density (average number of plants per unit area) of class i plant centers.

We realize that bunchgrass plants are not all perfectly circular. We also recognize that it would require an infinite number of radii class to describe precisely a real population. Furthermore, for a variety of reasons, plants in a natural population are not located independently of one another. Our purpose here, however, is merely to illustrate the nature of the bias in the step-point method, and for this the assumptions are perhaps acceptable. (Obviously one would hardly expect the biases we describe to disappear in real range data simply because the plants are clustered, noncircular, etc.). We have not considered it worth while to investigate the bias when randomness fails, since the results depend critically on the nature of the departure from randomness. For example, if in

Authors are associate professor, Statistics Department, University of California, Riverside, 92521, and range scientist, Pacific Southwest Forest and Range Experiment Station, USDA, Forest Service, Fresno, Calif.

practice the small plants tend to be more clustered than the large plants, the step-point method will underestimate their proportions. On the other hand, if large plants are usually surrounded by a halo of small offsprings, the step-point method will tend to miss too many of the large plants, and thus overestimate the frequency of the small ones. Since both cases, and a whole spectrum of others, occur in practice it scarcely seems profitable to set up models for nonrandomness in this context.

By number, the size classes should appear in our sample in the ratio $\rho_i:\rho_2: ...: \rho_k$. By area of cover, the classes should appear in the ratios $\pi r_1^2 \rho_1: \pi r_2^2 \rho_2: ...: \pi r_k^2 \rho_k$. Our task is to calculate the expected proportions of the size classes obtained in step-point sampling. To do this, we need only compute the probability of getting the various size classes in a single step; for, as can easily be shown, the expected number of plants of class i in, say, 100 steps is simply 100 times the probability of getting a class i plant in a single step. Denote these probabilities by $p_i(i = 1, 2, ..., k)$.

We can get a plant of class i in 2 ways; either

(i) We hit on a plant of class i, or

(ii) We do not hit any plant, and the closest plant to our steppoint is of class i.

Event (i) has probability $\pi r_i^2 \rho_i$, since this is the proportion of land covered by the ith size class. Let s_i be the chance of event (ii). The s_i should sum to

$$k (1 - \sum_{i=1}^{k} \pi r_1^2 \rho_i)$$
 (1)

which is the chance that we do not step on any plant. Then p_i is given by

$$\mathbf{p_i} = \pi \mathbf{r_1^2} \boldsymbol{\rho_i} + \mathbf{s_i}.$$

Now s_i is the chance that for some x

- (a) we encounter no plant within a range x for our step-point hit, then
- (b) we encounter a class i plant in the range x to x + dx. See Fig.(1).

The probability of (b) is $2\pi\rho_i (x+r_i)dx$, since we require a plant center of class i to appear in an annulus of circumference $2\pi(x+r_i)$ and hence area $2\pi(x+r_i)dx$. The probability of (a) is the product from j=l to k of the chances that no plant of class j has its center in a circle with our toe as center and radius $x+r_j$. These probabilities are well known to be $\exp[-\pi\rho_j(x+r_j)^2]$; see for example, Pielou (1969, p. 112). Summing over all positive values x, and taking the limit as dx gets small, we get

$$s_{i} = \int_{\infty}^{\infty} \{ \prod [exp - \pi (x+r_{i})^{2}] \} 2\pi \rho_{i} (x+r_{i}) dx .$$

x=0 j=1

This can be written $s_i = 2\pi\rho_i \int_0^\infty (x+r_i) \exp(-(c_0x^2+2c_1x+c_2)dx)$ where, for n = 0, 1 or 2, we write $c_n = \sum_{i=1}^{k} \rho_i r_i^n$. After a fair amount of manipulation, this comes to

$$\mathbf{s}_{i} = \frac{\pi \rho_{i}}{c_{0}} \exp(-c_{2}) + \left\{ 2\pi \sqrt{\frac{\pi}{c_{0}}} \exp(-c_{2} - \frac{c_{1}^{2}}{c_{0}}) \right\} \rho_{i}(\mathbf{r}_{i} - \frac{c_{1}}{c_{0}}) \left[1 - \Phi(c_{1} \sqrt{\frac{2}{c_{0}}}) \right], (2)$$

where

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

As a partial check, we note that the s_i , sum to exp (-c₂), which is the probability that we do not step on a plant.



Fig. 1: Plant of ith class is closest to the toe point.

From equation (1), we now have an expression for the expected size class proportions P_i in toe point sampling. When the total area of cover becomes exceedingly small, we find that P_i tends to $p_i/(p_i+...+p_k)$, which is the numerical proportion of size class i in the sample. As the area increases, the P_i get closer to the areal proportions of the size classes. This reflects the fact that when the plant cover is very sparse and the nearest plant is, typically, far off, the areas of the plants are of much less importance than the frequencies. As distances decrease the areas of the plants play a larger part in determining which plant is closest to the toe point.

The bias in the toe point estimate of numerical proportion of size class i will here be taken as

$$(\mathbf{E}_i - \mathbf{T}_i)/\mathbf{T}_i$$

where T_i is the true proportion of the ith size class and E_i the proportion expected with the toe point method. The bias in estimating the areas of the size classes is similarly defined.

Size classes	Radius (r _i)	Density (ρ _i) (Number of plants per unit area)	Number of plants in 206 step points	Number expected (by proportion)	Number expected (by area)	Number expected by theory 2)	Number expected by perimeter
1	2	.000119	57	60.6	13.9	54.0	31.7
2	3	.0000849	32	43.1	22.3	40.8	33.9
3	4	.0000975	60	49.5	45.5	49.9	51.9
4	6.4	.000104	57	52.9	124.3	61.2	88.5
	χ^2 on 3 de	grees of freedom:	206	206 5.62 ⁽ⁿ⁾²	206 178.9**1	206 4.40 ⁽ⁿ⁾²	206 32.9**1

1**indicates significance at the 1% level; * indicates the 5% level.
2^(2k)indicates result not significant at 5%.

Table 1. Bunchgrass size and numbers from the California site.



Fig. 2(a),(b),(c): Bias in estimates of basal area (solid line) and proportion (broken line) for 3 size classes. Radii are in ratio 1:2:3.

2(a): Equal numbers of plants.

2(b): Frequencies in ratio 4(smallest plants):2:1.

2(c): Frequencies in ratio 1(smallest plants):2:4.

It has been suggested that the chance of hitting a plant is related to the length of perimeter. This is of little consequence since this perimeter is of no value in range management. However, the expected number of hits, based on perimeter, is presented in Table 1. This demonstrates that, in general, there will also be a bias in estimates of perimeter.

Results

Bias, Calculated Theoretically

Figure 2(a), (b) and (c) shows the bias in area (solid line) and in number (broken line) for the case of 3 size classes with radii in the ratio 1:2:3. In Figure 2(a) the numbers of plants in the 3 classes are always equal. In Figure 2(b) the plants occur with frequencies in ratio 4 (smallest plants):2:1 (largest). In Figure 2(c) the ratios are 1 (smallest):2:4.

Comments

(a) Bias in *area* estimates. As expected, the bias is greatest when the plants are far apart, since then the actual areas of the plants are irrelevant; only the numbers in the size classes are being estimated. Generally speaking, the bias decreases as the plant density increases, but still is typically large. For the smallest size class the bias is usually more than 100%.

(b) Bias in *numerical* estimates. There is, of course, no bias when the plants are very far apart, but the bias increases sharply as the density increases. Even at 5% total plant cover, the bias is often

of the order of 50%.

Examples from the Field

To illustrate the way the step-point method can lead to bias in practice, we simulated the method on several sets of real data; the maps are drawn from real populations at the California and Oregon sites. The first example is taken from a map showing bunchgrasses at the California site. The grasses ranged in size from 3/16'' radius up to 3/8''; the plants were not all circular, but this should not matter greatly since our concern is only to illustrate the bias inherent in the step-point method. For convenience, the plants were grouped into 4 sizes:

Group 1, consisting of the smallest plants. On the actual site there were 16 such plants, with a total area of .0015 (0.15% of cover).

Group 2, the next smallest set of plants. There were 25 of these, covering .0024 of the area.

Group 3, consisting of 33 plants with a total area of .0049. Group 4, the largest plants. There were 9 plants, covering .0134 of the area.

The average radii of the 4 groups were approximately in the ratio 2:3:4:6.4.

We conveniently took the 4 radii to be 2,3,4,6.4, as shown in the second column of Table 1. The densities ρ_i in column 3 are calculated from the fact that the ith group has area $\pi \rho_i r_i^2$. Column 4 shows the actual results obtained from 206 steps with the steppoint method.

Table 2. Bunchgrass size and numbers from the Oregon site.

Size classes	Radius	(\mathbf{r}_i) Density (ρ_i)	# Plants in 200 steps	# Expected proportion)	# Expected (by area)	# Expected by theory
1	6	.0000111	61	52.9	23.7	51.6
2	8 °	.0000214	83	102.0	81.3	99.2
3	13	.00000947	_56	45.1	95.0	49.2
			200	200	200	200
_	χ² on 2	degrees of freedom:		7.39*	74.6**1	5.29 ⁽ⁿ⁾

1**Indicates significance at the 1% level; *indicates the 5% level.

^{2(a)}Indicates result not significant at 5%.

Size classes	Radius (ri)	Density (p _i)	Observed # in 219 steps	# Expected (by proportion)	# Expected (by area)	# Expected by	theory
1	1	.00100	30	54.75	7.3	32.4	
2	2	.00100	57	54:75	29.2	51.7	
3	3	.00100	60	54.75	65.7	59.9	
4	4	.00100	72	54.75	116.8	75.0	
			219	219	219	219	9
				17.2**1	114.7**1	0.84 ⁽⁽ⁿ⁾²	

Table 3.	Bunchgrass	size and	numbers	from the	artificial	population
----------	------------	----------	---------	----------	------------	------------

¹⁺⁺Indicates sgnificance at the 1% level; ⁺indicates the 5% level.

^{2⁽ⁿ⁾Indicates results not significant at 5% level.}

If the step-point method were to be unbiased in estimating the proportions in the 4 size classes, then the expected numbers of the 4 size classes in 206 steps would be as in column 5, apart from sampling fluctuations; for these numbers are in the same proportions as the actual numbers in each class. The discrepancy between the observed frequencies and the true frequencies does not seem very serious here; indeed the chi-square statistic for comparing observed to expected is not significant ($\chi^2 = 5.62$ on 2 degrees of freedom). This result could have been anticipated because the plants are sparse in this example, the area of cover being only about 2%.

Column 6 shows what would be expected if the step-point method gave unbiased estimates of the *areas* of the 4 classes. Here, there is a large discrepancy, the total area of the smaller plants being overestimated at the expense of the larger plants.

As a check, we show in column 7 the expected numbers calculated on the basis of the theory of section 2, under the simplifying assumptions of circular plants whose centers are scattered at random over the plot. The difference between observed and expected numbers is clearly not very large, and indeed not statistically significant ($\chi = 4.40$, on 3 degrees of freedom). Notice that for each size class the theoretical expectation falls between the estimate by proportion and the estimate by area. Finally, column 8 indicates that, as expected, there is a bias in estimates of size class perimeters.

A similar pattern is revealed in the other set of data we analyzed. In this case, again bunchgrasses but from the Oregon site, we categorized the plants into 3 size classes: small, medium, and large. The results are displayed in Table 2 below; the format and the calculations are analogous to those of Table 1. Once more we find a serious bias in the step-point estimate of the areal proportions of the 3 size classes ($\chi^2 = 74.6$). Since the cover is again sparse, the bias in estimates of proportion is smaller (though still significant at 5%). This time the observed data does not fit the theoretical calculations quite so well, although the χ^2 of the 5.29 is not significant at 5%.

Finally, Table 3 shows the results on an artificial population, consisting of 4 size classes. These are circles with radii in the ratio 1:2:3:4 in equal frequencies, scattered at random over the "plot." Although the results of the sampling are fitted well by the theoretical expectations from equation (1), there is serious bias in the estimates both of the proportions and of the areas of the size classes. The bias in proportions is not surprising, as the density of cover is about 10%, which is rather high.

Conclusions and Recommendations

The step-point method has proved to be a valuable technique on annual grass ranges. Its simplicity has caused it to be employed on other range types, including the bunchgrass ranges of the Great Basin. Among its uses are the estimation of cover and composition, and systematic selection of plants for utilization estimates. Our calculations have shown the method to have serious biases, which vary with plant size, in the estimation of cover and density. The biases are clearly illustrated in the sampling maps of real bunchgrass populations. Admittedly, the bias in basal area would decrease for very dense stands, or for cases when plant sizes are roughly the same. However, bunchgrass ranges in the Great Basin seldom exceed 10% basal cover and there is usually a variety of plant sizes in the population. Therefore, the biases in area estimates are likely always to be serious and use of the method should be considered carefully. Moreover, we have seen that with as little as 5% cover the bias in estimates of size class frequency can be serious.

What can be done about the bias? For estimates of the numerical proportions of the size classes, a simple device is to list the plant whose *center*, rather than perimeter, is closest to the toe point. It is clear that if the plants are really circular and scattered at random, there will then be no bias associated with the unequal sizes of the plants. When basal areas, rather than numerical proportions, are being estimated it is possible to categorize into a few convenient size classes, estimate the proportions of the size classes as above (using the centers of the plants rather than the perimeters), and finally multiply each proportion by the squared radius of the size class. This takes account of the differing basal areas of plants in various classes. Once again, if the plants were really circular and scattered at random, and if the plants within each class were really all of the same size, this device would remove the bias.

As we have noted in section 2, one cannot predict in general what will be the effect on the bias of departures from randomness. It is easy to see how, even with the above suggestions—using plant centers, and multiplying by area if basal area estimates are required—departure from the assumptions can lead to systematic bias. But it seems that the suggestions can be recommended: they require minimal extra effort over the standard step-point method, and are superior regarding bias at least when the grasses are scattered approximately at random. There seems no reason to suppose the reverse to be true in other cases.

Bibliography

- Anonymous. 1969. Range analysis handbook. U.S. Forest Service, Region 5.
- Costello, David F., and H.E. Schwan. 1946. Conditions and trend on ponderosa pine ranges in Colorado. USDA, Forest Service, Rocky Mountain Forest and Range Exp. Sta. (Processed).
- Evans, Raymond A., and R. Morton Love. 1957. The step-point method of sampling—a practical tool in range research. J. Range Manage. 10:208-212.
- Goodall, D.W. 1952. Some consideration in the use of point quadrats for analysis of vegetation. Aust. J. Sci. B. 5:1-41.
- Levy, E.B., and E.A. Madden. 1933. The point method of pasture analysis. N.Z. J. Agr. 46:267-279.
- Pearse, Kenneth. 1935. An area-list method of measuring range plant populations. Ecology 16:573-79.
- Pielou, E.C. 1969. An introduction to mathematical ecology. Wiley-Interscience.
- Roach, Mack E. 1950. Estimating perennial grass utilization on semidesert cattle ranges by percentage of ungrazed plants. J. Range Manage. 3:182-5.
- Cockayne, Leonard. 1926. In: Tansley, A.G. and T.F. Chipp. (Ed.) Aims and Methods in the Study of Vegetation in New Zealand, London.
- Tinney, F.W., O.S. Asmodt, and H.L. Ahlgren. 1937. Preliminary report of study of methods used in botanical analysis of pasture sevards. J. Amer. Soc. Agron. 29:835-840.