STATISTICA, anno LXIX, n. 1, 2009

THE UNIFORMLY MOST POWERFUL INVARIANT TEST FOR TWO MODELS OF DETECTION FUNCTION IN POINT TRANSECT SAMPLING

R. Borgoni, P. Quatto

1. INTRODUCTION

Numerous studies of wildlife populations require estimates of population abundance. Transect sampling methodology provides an effective approach for estimating population size ν or density $\delta = \nu/A$, where A is the area of the study region. A thorough review of this approach is given by Buckland et al. (2001, chap. 1).

The point transect design (Buckland, 1987) in particular assumes that:

- k points are randomly chosen within the study area;
- animals of interest are uniformly distributed with respect to distance in any direction from the points;
- at each of the selected points, an observer measures the distance from himself to any animal detected;
- animals at the observation points are detected with certainty;
- animals are detected at their initial location, prior to any movement;
- distances are measured without errors;
- detections are independent events.

Since the number of animals observed from each point is quite small in many contexts where this sample scheme is adopted (as for instance in ornithology), sampled distances are pooled together to increase the sample size.

Let $z_1,...,z_n$ be the sample of size n obtained by pooling together the distances measured at each of the k observation points. Let f be the probability density function (pdf) of the observed distances and let g be the detection function, that is to say $g(y)$ is the conditional probability of detecting an animal, given that it is at distance y from the observer.

From the above assumptions it turns out that the relation

$$f(z) = \frac{zg(z)}{\int_0^{+\infty} yg(y)dy} \tag{1}$$

holds for every distance z and the general form for estimating the population density δ is given by

$$\hat{\delta} = \frac{n}{2k\pi}\hat{f}'(0),$$

where $\hat{f}'(0)$ is an estimator of the derivative of f at 0 which satisfies the fundamental identity

$$f'(0) = \frac{1}{\int_0^{+\infty} y g(y) dy}.$$

The basic problem for estimating δ, or equivalently ν, is therefore to estimate $f'(0)$.

Assumptions and results regarding point transect sampling are extensively discussed in Buckland et al. (2001, chap. 2).

Borgoni et al. (2005) investigated the small-sample behaviour of different δ estimators, depending on the shape of the detection function. The authors considered two popular families of detection functions (Zhang, 2001; Eidous, 2005):
the half-normal family

$$g(y) = \exp\left(-\frac{y^2}{2\sigma^2}\right) \quad (\sigma > 0) \tag{2}$$

and the negative exponential family

$$g(y) = \exp\left(-\frac{y}{\sigma}\right) \quad (\sigma > 0). \tag{3}$$

The former satisfies the shape criterion

$$g'(0) = 0 \tag{4}$$

whereas the latter does not. This property, also known as the shoulder condition, ensures that animal detection is nearly certain at small distances from the observer (Buckland et al., 2001, pp. 42, 68-69; Buckland et al., 2004, p. 344). However, such a condition fails when detectability decreases sharply around the observation points because of low or inexistent visibility (e.g. in presence of fog or dense vegetation).

In a point transect framework, Borgoni et al. (2005) demonstrated several simulation results suggesting that the usual estimators of δ are extremely sensitive to departures from the shape criterion. A similar behaviour was found in the line transect context (Eidous, 2005). Therefore, testing the shape criterion is a preliminary step for any attempt to estimate wildlife population density via transect sampling (Zhang, 2003; Eidous, 2005). Although this problem has been previously addressed by Mack (1998) and Zhang (2001) in line transect sampling, no attempt has been made in the context of point transect so far.

The aim of this paper is to propose a procedure for testing the shoulder condition (4). As this condition is independent from the choice of the measure unit for the distance, the scale invariance seems to be a natural restriction for a statistical test. In particular we focus on a scale invariant test for discriminating between the two families (2) and (3). Because of (1) this turns out to be equivalent to testing that the distance pdf belongs to one of the two families

$$F_0 = \left\{ f(z) = \frac{z}{\sigma^2} \exp\left(-\frac{z^2}{2\sigma^2} \right) : \sigma > 0 \right\} \tag{5}$$

and

$$F_1 = \left\{ f(z) = \frac{z}{\sigma^2} \exp\left(-\frac{z}{\sigma} \right) : \sigma > 0 \right\}. \tag{6}$$

The proposed test is the uniformly most powerful (UMP) in the class of the scale invariant tests.

This is discussed in section 2 where the asymptotic distributions of the test statistic under (5) and (6) are calculated.

In section 3 the critical values and the powers of the test are tabulated via Monte Carlo simulations for several typical α-levels and small sample sizes n.

In section 4 the proposed procedure is applied to a dataset coming from a large study conducted by the Rocky Mountain Bird Observatory, Colorado, in 2002.

Conclusions are provided in section 5.

2. THE UMP SCALE INVARIANT TEST

Given n independent observations $z_1,\ldots,z_n$ from Z with unknown pdf f, we consider the problem of testing

$$H_0 : f \in F_0 \text{ vs. } H_1 : f \in F_1, \tag{7}$$

where F_0 is the family of Rayleigh distributions with scale parameter σ, and F_1 is the family of Gamma distributions with shape parameter 2 and scale parameter σ, specified in (5) and (6) respectively. This problem is invariant under the group of scale transformations (Lehmann and Romano, 2005, p. 213)

$$G = \{ \gamma(z) = rz : r > 0 \}.$$

A maximal invariant under G (Lehmann and Romano, 2005, p. 215) is:

$$(z_1/z_n, z_2/z_n, ..., z_{n-1}/z_n) . \tag{8}$$

It can be proved (see Appendix 1) that the UMP test among all the invariant functions, i.e. the functions of this maximal invariant (Lehmann and Romano, 2005, p. 214), rejects the null hypothesis for large values of the likelihood ratio

$$\lambda = \frac{(2n-1)!}{2^{n-1}(n-1)!}\left[\frac{\sum_{i=1}^{n} z_i^2}{\left(\sum_{i=1}^{n} z_i\right)^2}\right]^n . \tag{9}$$

Given that random variable corresponding to λ is a monotonically increasing function of the statistic

$$Q_n = \frac{\frac{1}{n}\sum_{i=1}^{n} Z_i^2}{\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right)^2}, \tag{10}$$

the critical region of the UMP scale invariant test for the hypotheses (7) can be written as

$$Q_n \geq q_{n,\alpha} , \tag{11}$$

where α denotes the level of significance and $q_{n,\alpha}$ is the corresponding critical value so that

$$P(Q_n \geq q_{n,\alpha} \mid H_0) = \alpha .$$

Furthermore, the asymptotic normal distribution under H_0

$$\sqrt{n/(256/\pi^3 - 80/\pi^2)}(Q_n - 4/\pi) \xrightarrow{d} N(0,1) , \tag{12}$$

and under H_1

$$\sqrt{4n/3}(Q_n - 3/2) \xrightarrow{d} N(0,1) , \tag{13}$$

are derived from bivariate central limit theorem and delta method (see Appendix 2). For large n the approximate critical value and the power are given respectively by

$$q_{n,\alpha} \cong 1.27 + 0.39 z_{1-\alpha}/\sqrt{n}$$

and

$$1-\beta = P(Q_n \geq q_{n,\alpha} \mid H_1) \cong 1-\Phi(0.22 z_{1-\alpha} - 0.26\sqrt{n}),$$

where $z_{1-\alpha}$ is the $(1-\alpha)$th quantile and Φ is the cumulative distribution function of the standard normal distribution. Hence, the proposed test is also consistent (Lehmann, 2001, p. 158).

3. CRITICAL VALUES AND POWERS

As the distribution of the test statistic under (5) or (6) does not depend on the scale parameter, Monte Carlo simulations can be performed in order to obtain the empirical critical values and powers for small sample sizes.

The simulation design consists of randomly drawing n distances from the distribution (5) assuming, without loss of generality, $\sigma = 1$.

The statistic (10) is then applied to each of the simulated samples and the procedure is repeated 5000 times.

The critical value $q_{n,\alpha}$ for a considered significance level α is obtained as $100\times(1-\alpha)$-th percentile of the Monte Carlo replicates. We obtained the power of the test in a similar manner. In this case, each sample is simulated according to the alternative distribution (6).

Monte Carlo approximations of the critical values $q_{n,\alpha}$ and powers are reported in Table 1 and in Table 2.

It can be noted that the power under (6) is reasonably good even in the case of a moderate sample and low α.

TABLE 1

Critical values $q_{n,\alpha}$ of the UMP scale invariant test

$q_{n,\alpha}$	$n = 30$	$n = 40$	$n = 50$	$n = 60$	$n = 100$
$\alpha = 0.01$	1.46	1.43	1.41	1.40	1.37
$\alpha = 0.05$	1.39	1.37	1.37	1.36	1.34
$\alpha = 0.10$	1.36	1.35	1.34	1.33	1.32

TABLE 2

Power $1-\beta$ of the UMP scale invariant test

$1-\beta$	$n = 30$	$n = 40$	$n = 50$	$n = 60$	$n = 100$
$\alpha = 0.01$	0.47	0.61	0.71	0.78	0.96
$\alpha = 0.05$	0.68	0.80	0.87	0.91	0.99
$\alpha = 0.10$	0.79	0.87	0.92	0.95	0.99

The test also performs well in terms of the power in the case of the data originating from a mixture of (5) and (6).

In particular, the case where the sample is drawn from a pdf

$$pz\ \exp(-z^2/2) + (1-p)\ z\ \exp(-z)$$

was considered, p being the average proportion of the observed distances simulated from a population distributed according to the null hypothesis.

Table 3 shows the power of the test of level $\alpha = 0.05$ for a range of mixture proportions p and some sample sizes.

It can be observed that the proposed procedure performs well even in the case of a sample of moderate size drawn from mixture model with a large p.

TABLE 3

Power $1-\beta$ of the UMP scale invariant test for different mixture proportions

$\alpha = 0.05$	$n = 40$	$n = 50$	$n = 100$
$p = 0.25$	0.80	0.87	0.99
$p = 0.50$	0.75	0.81	0.97
$p = 0.75$	0.54	0.60	0.83

4. A CASE STUDY

In this section, we apply the test procedure previously described to two datasets drawn from a large study conducted in 2002 by the Rocky Mountain Bird Observatory, Colorado (Panjabi, 2003, p. 100).

The first dataset is a point transect sample of 72 chipping sparrows (scientific name: *Spizella Passerina*) observed during the early morning of 28 May 2002. The 51 observation points were allocated in Pine Juniper shrub land in South Dakota. The distances collected range from between 8m to 183m (the first and third quartile were 20.75m and 75.25m, respectively) with an average distance of 53.18m and standard deviation equal to 37.59m.

The distribution of the observed distances is shown in Figure 1. The box plot suggests that one potential outlier is present in the data set at hand. This observation is therefore omitted in the subsequent analysis.

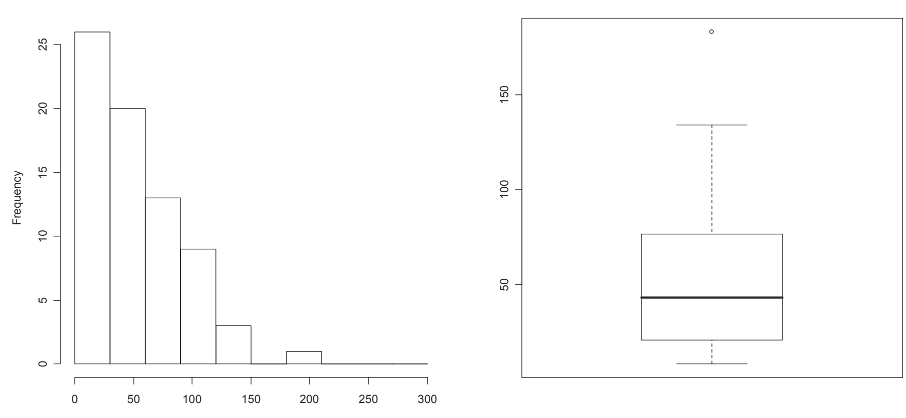

Figure 1 – Distribution of the observed distances for the chipping sparrow data in Pine Juniper shrub lands.

The test statistic for a null hypothesis of an half-normal detection function equals 1.45. The null distribution of this statistic is tabulated according to the method described in the previous section for a sample size n = 71 and 10000 simulations obtaining a critical value at the 5% significance level equal to 1.35. The null hypothesis should therefore be rejected at the considered level. At this significance level, the power is about 95%. In fact, in this case, there seems to be a strong evidence of a detection function which is not half-normal as the Monte Carlo p-value is extremely small (0.0007).

The second dataset is a point transect sample of 82 chipping sparrows selected during the early morning of 26 May 2002 in a different environment. The 94 observation points were allocated in a burn area in South Dakota.

The distances collected ranged from between 10m to 200m (the first and third quartile were 42m and 100.2m, respectively) with an average distance of 77.2m and standard deviation equal to 44.84m.

The distribution of the observed distances is reported in Figure 2. The box plot suggests that two potential outliers are present in the data set at hand. These observations were therefore omitted in the subsequent analysis. The third largest observation in the original sample was nearly at the extreme of upper whisker of the box plot (183m). This value was identified as a further outlier by a second box plot constructed on the sample obtained by dropping the two largest observations from the original dataset. This value was also dropped from the subsequent analysis.

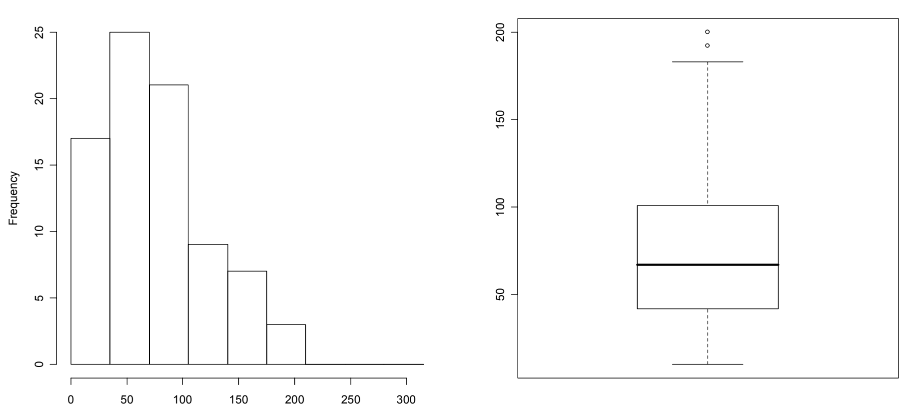

Figure 2 – Distribution of the observed distances for the chipping sparrow data in burn areas.

The test statistic for a null hypothesis of an half-normal detection function equals 1.291. In this case as well, the null distribution of this statistic is tabulated according to the method described in the previous section for a sample size n = 79 and 10000 simulations obtaining a 5% critical value equal to 1.345. Hence the null hypothesis should not be rejected at this considered significance level.

The power of the test was 96.2%. at this significance level. In this case, the Monte Carlo *p*-value results as being 0.304.

Finally, we can observe that, although the size of the two samples are not small, the Monte Carlo p-values still differ from the asymptotic approximation given by

$$\hat{\alpha} = 1 - \Phi\left(\frac{q - 1.273}{0.388}\sqrt{n}\right),$$

where Φ is the cumulative distribution function of the standard normal distribution and q the observed value of the test statistic in (10). The asymptotic p-values are equal to 0.0001 and 0.3437 for the first and second sample respectively.

5. CONCLUSIONS

In transect sampling the problem of testing the shoulder condition of a detection function is invariant under the group of scale transformations. Therefore, the scale invariance furnishes a natural restriction on the statistical procedure to be utilised.

The half-normal is perhaps the most widely used detection function family (possibly in conjunction with adjustment terms; Buckland et al., 2001) which satisfies the shoulder condition. In this paper we proposed a procedure for testing the half-normal family against the negative exponential family which violates the shape criterion. Hence the problem is reduced to testing between the Rayleigh family and a subclass of the Gamma family. For this we proposed the UMP scale invariant test for which the limiting normal distribution of the test statistic is provided. From this follows the consistency of the test. In the case of small samples we suggest a Monte Carlo approach for tabulating the critical values and related powers for a range of different sample sizes and significance levels. It turned out that the critical values and the power approximated via the Monte Carlo and the asymptotic distribution provide a very similar result for a sample size of 100 or more. For example, in the case of a sample size equal 100 the empirical and asymptotic critical values both were 1.34 at the 5% level; examining the power, only a slight difference was registered between the two approaches (0.99 and 0.97 respectively).

Finally, the test was applied to a point transect survey. As expected, the shoulder condition seems largely supported by data collected in an open space with good visibility whereas it is rejected in the case of data gathered in dense vegetation.

Department of Statistics RICCARDO BORGONI
University of Milano – Bicocca PIERO QUATTO

APPENDIX 1

The pdf of the sample $(z_1,...,z_n)$ can be written as

$$L(z_1,...,z_n)=\frac{1}{\sigma^n}\prod_{i=1}^{n} f\left(\frac{z_i}{\sigma}\right),$$

where

$$f(z)=\begin{cases} z\exp(-z^2/2) & : \text{ under } H_0 \\ z\exp(-z) & : \text{ under } H_1 \end{cases}.$$

Therefore, the maximal invariant (8) is expressed as

$$(r_1,...,r_{n-1})=(z_1/z_n\,, z_2/z_n\,,...,z_{n-1}/z_n)$$

and has pdf given by

$$\int_0^{+\infty} z_n^{n-1} L(z_n r_1,...,z_n r_{n-1}, z_n)\,dz_n = z_n^n \int_0^{+\infty} u^{n-1}\prod_{i=1}^{n} f(z_i u)du$$

$$=\begin{cases} 2^{n-1}(n-1)!\dfrac{\prod_{i=1}^{n} z_i}{\left(\sum_{i=1}^{n} z_i^2\right)^n} & : \text{ under } H_0 \\[2ex] (2n-1)!\dfrac{\prod_{i=1}^{n} z_i}{\left(\sum_{i=1}^{n} z_i\right)^{2n}} & : \text{ under } H_1 \end{cases}$$

from which the likelihood ratio (9) follows.

By the Neyman-Pearson Lemma, the most powerful test rejects the null hypothesis when (9) is too large. Given that its critical region does not depend on σ, the test is UMP among all invariant tests, as asserted.

APPENDIX 2

By bivariate central limit theorem (Lehmann, 2001, p. 291),

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i - \mu_1, \frac{1}{n}\sum_{i=1}^{n} Z_i^2 - \mu_2\right) \xrightarrow{d} N(\mathbf{0}, \mathbf{\Sigma}),$$

where

$$\mu_1 = E(Z), \quad \mu_2 = E(Z^2), \quad \mathbf{0} = (0,0)$$

and

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

with

$$\sigma_1^2 = Var(Z), \ \sigma_2^2 = Var(Z^2), \ \sigma_{12} = Cov(Z, Z^2).$$

Then, by delta method (Lehmann, 2001, p. 296),

$$\sqrt{n}(Q_n - \eta) \xrightarrow{d} N(0, \tau^2),$$

where

$$\eta = \frac{\mu_2}{\mu_1^2}$$

and

$$\tau^2 = \frac{4\mu_2^2}{\mu_1^6}\sigma_1^2 - \frac{4\mu_2}{\mu_1^5}\sigma_{12} + \frac{1}{\mu_1^4}\sigma_2^2 = \frac{\mu_4 - \mu_2^2}{\mu_1^4} + 4\mu_2 \frac{\mu_2^2 - \mu_1\mu_3}{\mu_1^6},$$

with

$$\mu_3 = E(Z^3), \ \mu_4 = E(Z^4).$$

Therefore, under H_0

$$\eta = 4/\pi, \ \tau^2 = 256/\pi^3 - 80/\pi^2,$$

whereas under H_1

$$\eta = 3/2, \ \tau^2 = 3/4.$$

Hence (12) and (13) follow.

ACKNOWLEDGMENT

The authors thank David Hanni and the Rocky Mountain Bird Observatory for having kindly provided the data used in the paper. The authors also thank Kendall Katze for editing the paper.

REFERENCES

R. BORGONI, M. CAMELETTI, P. QUATTO (2005), *Comparing estimators of animal abundance: a simulation study*, Atti del Convegno della Società Italiana di Statistica "Statistica e Ambiente", Università degli Studi di Messina, pp. 181-184.

S. T. BUCKLAND (1987), *On the variable circular plot method of estimating animal density*, "Biometrics", 43, 2, pp. 363-384.

S. T. BUCKLAND, D. R. ANDERSON, K. P. BURNHAM, J. L. LAAKE, D. L. BORCHERS, L. THOMAS (2001), *Introduction to Distance Sampling*, Oxford University Press, Oxford.

S. T. BUCKLAND, D. R. ANDERSON, K. P. BURNHAM, J. L. LAAKE, D. L. BORCHERS, L. THOMAS (2004), *Advanced Distance Sampling*, Oxford University Press, Oxford.

O. M. EIDOUS (2005), *On Improving Kernel Estimators Using Line Transect Sampling*, "Communications in Statistics – Theory and Methods", 34, pp. 931-941.

Y. P. MACK (1998), *Testing for the shoulder condition in transect sampling*, "Communications in Statistics – Theory and Methods", 27, pp. 423-432.

E. L. LEHMANN (2001), *Elements of Large-Sample Theory*, Springer, New York.

E. L. LEHMANN, J. P. ROMANO (2005), *Testing Statistical Hypotheses*, Springer, New York.

A. PANJABI (2003), *Monitoring the birds of the Black Observatory*, Fort Collins, Colorado.

S. ZHANG (2001), *Generalized likelihood ratio test for the shoulder condition in line transect sampling*, "Communications in Statistics – Theory and Methods", 30, pp. 2343-2354.

S. ZHANG (2003), A *Note on Testing the Shoulder Condition in Line Transect Sampling*, "Proceedings of the 2003 Hawaii International Conference on Statistics and related topics".

SUMMARY

The uniformly most powerful invariant test for two models of detection function in point transect sampling

Estimating population abundance is of primary interest in wildlife population studies. Point transect sampling is a well established methodology for this purpose. The usual approach for estimating the density or the size of the population of interest is to assume a particular model for the detection function (the conditional probability of detecting an animal given that it is at a certain distance from the observer). Two popular models for this function are the half-normal model and the negative exponential model. However, it appears that the estimates are extremely sensitive to the shape of the detection function, particularly to the so-called shoulder condition, which ensures that an animal is nearly certain to be detected if it is at a small distance from the observer. The half-normal model satisfies this condition whereas the negative exponential does not. Testing whether such a hypothesis is consistent with the data at hand should be a primary concern. Given that the problem of testing the shoulder condition of a detection function is invariant under the group of scale transformations, in this paper we propose the uniformly most powerful

test in the class of the scale invariant tests for the half-normal against the negative exponential model. The asymptotic distribution of the test statistic is calculated by utilising both the two models while the critical values and the power are tabulated via Monte Carlo simulations for small samples. Finally, the procedure is applied to two datasets of chipping sparrows collected at the Rocky Mountain Bird Observatory, Colorado.