

## A BAYESIAN MULTINOMIAL MODEL FOR ANALYZING CATEGORICAL HABITAT SELECTION DATA

DANA L. THOMAS<sup>1</sup>, Department of Mathematical Sciences, University of Alaska, Fairbanks AK 99701, USA  
COLLEEN IANUZZI, Department of Mathematical Sciences, University of Alaska, Fairbanks AK 99701, USA  
RON BARRY, Department of Mathematical Sciences, University of Alaska, Fairbanks AK 99701, USA

**Abstract:** Modeling the number of uses of discrete habitat types by animals with a multinomial distribution, we illustrate the use of Bayesian methods to estimate selection. An advantage of this approach in assessing selection is the construction of credibility intervals that do not rely on large sample normal theory. In addition, credibility intervals for ranked selection of habitats are easily obtained. Bayes factors and Bayesian posterior predictive values are used to test the hypothesis of selection. We compare our method to alternative methods with respect to general characteristics, such as the unit sum constraint, and for a real data set. Freely available WinBugs software is used to fit the model and test hypotheses.

**Key words:** Habitat selection; Bayesian analysis; Compositional Analysis;

---

### INTRODUCTION

A common resource selection study design is illustrated by Smith et al.'s (1982) study of gray partridges (*Perdix perdix*). They uniquely marked a sample of individual birds and classified their relocations over time into one of five habitat types. Thus, habitat use was recorded for each bird. A map of the study area was partitioned into the five habitat types and the proportional area of each type determined, i.e., a census of availability was conducted. Availability of the habitat categories was assumed to be the same for all animals. A summary of the information resulting from a portion of this study described by Alldredge and Ratti (1986) is given in Table 1. This data set has been used by others to illustrate new methods and compare hypothesis tests of selection (see for example, Dasgupta and Alldredge 1998 and 2002). We use this data set to illustrate a Bayesian analysis of categorical resource selection and to compare our results to those of Dasgupta and Alldredge (2002) and White and Garrott (1990).

### A BAYESIAN MODEL

For the  $i$ th animal, let the random vector  $\mathbf{X}_i$  be an  $h$  dimensional vector with elements  $X_{ij}$  = the number of uses of habitat  $j$  by animal  $i$  among  $n_i = \sum_{j=1}^h x_{ij}$  observations on animal  $i$ ; where  $i = 1$  to  $m$  and  $j = 1$  to  $h$ . If the  $n_i$  observations on animal  $i$  are independent, then  $\mathbf{X}_i$  is multinomially distributed with parameters  $n_i$  and  $\mathbf{p}_i$ , where  $\mathbf{p}_i$  is an  $h$  dimensional vector ( $p_{i1}, p_{i2}, p_{i3}, \dots, p_{ih}$ ) with elements  $p_{ij}$  = the probability that animal  $i$  uses habitat  $j$ . That is, the probability density of  $\mathbf{X}_i$  is

$$f(\mathbf{X}_i | n_i, \mathbf{p}_i) = \binom{n_i}{x_{i1}, x_{i2}, x_{i3}, \dots, x_{ih}} p_{i1}^{x_{i1}} p_{i2}^{x_{i2}} p_{i3}^{x_{i3}} \dots p_{ih}^{x_{ih}}$$

for  $i = 1$  to  $m$ . We note that the  $p_{ij}$  are subject to the unit sum constraint

$$\sum_{j=1}^h p_{ij} = 1.$$

<sup>1</sup>Email: ffdlt@uaf.edu

Table 1. The Gray Partridge data set.

	Small Grain	Row crop	hay	pasture	idle
Availability	.282	.417	.102	.135	.063

Animal	Use = number of relocations				
1	0	8	0	20	2
2	25	21	0	0	1
3	17	11	0	0	2
4	4	0	0	0	2
5	20	0	0	9	0
6	22	0	0	2	0
7	0	7	6	0	1
8*	10	26	2	8	0
9	21	0	4	0	3
10	44	1	0	0	5
Total	163	74	12	39	16

\* - note that animal 8 is the only animal for which use is not significantly different from availability

Similar to Manly et al (2002), Barmi and Pontius (2000) and Aebischer et al (1993) we use a selection function  $w$  to model the  $\mathbf{p}$  as

$$P_{ij} = \frac{w_{ij} a_j}{\sum_{j=1}^h w_{ij} a_j},$$

where  $a_j$  is the known proportion of the  $j$ th habitat available to all animals and  $w_{ij}$  is the unknown relative probability of selection of the  $j$ th habitat by animal  $i$  if all habitats were equally available (see, Manly et al 2002 for more detail).

There are many possible prior distributions that could be used for the selection proportions  $w_{ij}$ . For example, one choice would be the  $h$ -dimensional Dirchelet distribution because it is a conjugate prior for the multinomial distribution. However, like Atchison (1986), we found the Dirchelet lacks the flexibility needed to model complex settings such as differences in selection between genders or age groups, groups of animals selecting habitats together, or changes in selection over time. Following Atchison and Agresti (1990; Chapter 9), we prefer the use of a logit type model for flexibility and write

$$w_{ij} = e^{a_{ij}}.$$

This approach allows the flexibility needed for future development because now the  $a_{ij}$  may be modeled using multivariate distributions with a general covariance matrix.

To illustrate the utility of a logit model for a simple situation, we use independent log normal distributions for priors for the selection functions  $w_{ij}$ . An alternative perspective of these priors is that we use independent normal priors with mean 0 and variance 200 for the  $a_{ij}$  so the priors are vague (non-informative). For identifiability we impose the constraint  $a_{i1} = 0$  ( or  $w_{i1} = 1$ ) for each animal ( $i$ ). Thus, selection of habitat  $j$  is relative to habitat 1 for  $j = 2, 3, 4, \dots, h$ . Note that any other constraint, e.g.,

$$\sum_{j=1}^h a_{ij} = 0,$$

may be used and the choice of habitat 1 for comparison is arbitrary; any other habitat may be used for comparison.

The resulting posterior distribution of the selection proportions has the following form:

$$f_{\mathbf{w}_i | n_i, \mathbf{p}_i}(\mathbf{w}_i | n_i, \mathbf{p}_i) = \frac{f_{\mathbf{x}_i | n_i, \mathbf{p}_i}(\mathbf{x}_i, n_i, \mathbf{p}_i(\mathbf{w}_i)) f_{\mathbf{w}_i}(\mathbf{w}_i)}{\int f_{\mathbf{x}_i | n_i, \mathbf{p}_i}(\mathbf{x}_i, n_i, \mathbf{p}_i(\mathbf{w}_i)) f_{\mathbf{w}_i}(\mathbf{w}_i)}$$

$$\propto \binom{n_i}{x_{i1}, x_{i2}, \dots, x_{ih}} p_{i1}^{x_{i1}} p_{i2}^{x_{i2}} \dots p_{ih}^{x_{ih}} \cdot \prod_{j=2}^h \frac{1}{\mathbf{s} \cdot w_{ij} \cdot \sqrt{2\mathbf{p}}} \cdot e^{\frac{-1}{2\mathbf{s}^2} [\ln w_{ij}]^2}$$

where  $f_{\mathbf{x}}(\cdot)$  is a multinomial distribution with parameters  $n_i$  and  $\mathbf{p}_i$  and the prior,  $f_{\mathbf{w}}(\cdot)$ , is a log normal distribution with parameters 0 and 200 (the mean of  $\ln(w) = 0$  and variance of  $\ln(w) = 200$ ).

The Bayesian model proposed here is a log ratio model similar to that of Aebischer et al's (1993) compositional analysis approach where the ratio of proportion use of habitat  $j$  is compared to that of the first habitat via the following relationship:

$$\ln\left(\frac{p_{ij}}{p_{i1}}\right) = \ln\left(\frac{w_{ij}}{w_{i1}}\right) - \ln\left(\frac{a_j}{a_1}\right)$$

for  $j = 2, \dots, h$ . However, in our model we do not have to evaluate the log ratios of use or the log ratios of availability so zeros in any of these terms create no problems.

We use freely available WinBUGS software to fit this model (see <http://www.mrc-bsu.cam.ac.uk/bugs/>, Spiegelhalter et al 1999 and Gilks et al 1994). The name WinBUGS refers to the Windows version of Bayesian inference Using Gibbs Sampling (see, for example, Casella and George 1992). This software is commonly used for Bayesian analysis of statistical models. Fitted models yield estimated posterior distributions, Bayes estimates (means of these posterior distributions), and credibility intervals for each  $w_{ij}$ . The default credibility intervals, 2.5% and 97.5% quantiles of the posterior distribution, are shown here.

Gibbs sampling as implemented in *WinBUGS*, has the ability to evaluate arbitrary functions of unknown model parameters at each iteration. We use this capability to rank the  $w_{ij}$  for each animal at each iteration. This yields a sample from the posterior distribution of ranks,  $P(w_{ij}$  is the  $k^{\text{th}}$  largest among the  $h$  habitats |  $\mathbf{x}_i$ ), from which credibility intervals for rank selection are determined. See the *WinBUGS* manual for additional information on posterior rank distributions.

In addition, we use *WinBUGS* for Bayesian model selection (hypothesis testing) by determining Bayesian posterior predictive values and the values of Bayes factors for competing models. Bayes factors are similar to likelihood ratio tests and  $2 \cdot \ln(\text{Bayes factor value})$  may be used to assess evidence against a null hypothesis (Kass and Raftery 1995). We determined Bayesian posterior predictive values following a process described by Congdon (2001). In each iteration we generated a new multinomial random variable,  $\mathbf{x}_i^{\text{rep}}$ , assuming no selection, that is,  $p_{ij} = a_j$  and new selection values,  $\mathbf{w}_i^{\text{rep}}$ , were estimated using the model described above. Posterior predictive values were determined by examining how often a statistic under the null hypothesis of no selection exceeded the value of the same statistic for fitted selection values ( $\mathbf{p}_i$  is a function of the fitted  $\mathbf{w}_i$ ). The statistic varied depending upon the hypothesis being examined. The hypothesis of no selection by an individual animal, say animal  $i$ , was assessed using a posterior predictive value based upon the usual chi-square statistic,

$$P\left(\sum_{j=1}^h \frac{(x_{ij}^{\text{rep}} - n_i a_j)^2}{n_i a_j} > \sum_{j=1}^h \frac{(x_{ij} - n_i p_{ij})^2}{n_i p_{ij}}\right).$$

Similarly, the hypothesis of no selection across animals was tested using the sum of the usual chi-square statistics,

$$P\left(\sum_i \sum_{j=1}^h \frac{(x_{ij}^{\text{rep}} - n_i a_j)^2}{n_i a_j} > \sum_i \sum_{j=1}^h \frac{(x_{ij} - n_i p_{ij})^2}{n_i p_{ij}}\right).$$

Finally, the hypothesis of no difference in selection of habitat  $j$  and habitat  $j'$  for animal  $i$  was tested using

$$P\left(|w_{ij}^{rep} - w_{ij'}^{rep}| > |w_{ij} - w_{ij'}|\right)$$

Alternative statistics could be used for any of these posterior predictive values. For example, the maximum of the chi-square statistic across animals could be used rather than the sum to test the hypothesis of no selection across animals.

To determine the sensitivity of our model to the choice of the normal prior, we also ran the model using uniform (-10, 10) priors on the  $a_{ij}$ .

## RESULTS

A two stage analysis is commonly performed when analyzing selection using data like that given in Table 1 (see White and Garrott 1990). First, an analysis of selection is conducted for each animal then an overall analysis of selection across animals is conducted. We follow this strategy using the Bayesian approach described here by first showing the analysis for partridge number 8 then showing overall results.

Table 2 provides a summary of the fit of our models with normal priors and uniform priors for animal number 8 and compares our results to those of the usual use divided by availability estimator of selection (selection ratio; see Manly et al 2002). It is clear from this table that our model provides a good fit and our estimates of selection are similar to those of the more usual estimator. In addition, this comparison shows little impact of choice of prior on the  $a_{ij}$ ; all further results are for normal priors only. Note that in Table 2 our Bayesian estimates have been re-scaled to sum to the total of selection ratios for comparison purposes. Our WinBUGS program yields  $w_{ij}^*$ , scaled values of  $w_{ij}$ , with values between 0 and 1 for each habitat using

$$w_{ij}^* = \frac{w_{ij}}{\sum_j w_{ij}}.$$

Table 3 provides a comparison of commonly used large sample normal theory confidence intervals ( $W \pm Z(1-a/2) \cdot se[W]$ ; see Manly et al 2002 equation 4.15) with Bayesian credibility intervals for each habitat for bird number 8. No effort has been made to control simultaneous error rates in this comparison; 95% confidence is used for each interval. Note that both types of intervals indicate no selection for small grain or pasture (1 is in both intervals). However, for hay and row crop habitats the Bayesian credibility interval suggests no selection while the large sample normal method suggests selection. Intervals for small grain and pasture were narrower for the large sample normal theory method while Bayesian intervals were narrower for hay and pasture habitats. We expect these differences are due to large sample versus small sample theory.

Bayesian intervals have several advantages. First, they are not dependent on large sample (normal) theory; They work with the sample sizes given. Second, negative lower bounds, like that for habitat hay in the normal theory intervals, are not possible. Third, Bayesian intervals for unused habitats, like idle habitat for animal 8, still result in intervals that include zero and based on the number of relocations of the bird, give some idea of likelihood of selection. Zero use for a given sample size does not mean that zero use is guaranteed for all sample sizes. Thus, Bayesian credibility intervals incorporate the variability of selection of the other habitats into the interval for selection of a zero use habitat. That is, this interval is a direct result of the unit sum constraint and the prior distribution on  $w_{ij}$ . Lastly, intervals for any function of the variables, e.g., estimating proportion use or selection when combining two habitats or the difference or ratio of selections are easy to implement in WinBUGS.

Figure 1 shows the posterior densities of selection for each habitat for animal 8 as estimated by WinBUGS. Quantiles of these estimated densities are used to form the credibility intervals in Table 2. The posterior densities for the selection functions for habitats 3 (hay) and 5 (idle) are clearly non-normal while the posterior densities for selection of habitats 1 (small grain), 2 (row crop) and 4 (pasture) appear approximately normal. The posterior density for habitat 5, idle, illustrates the case for a 0 use habitat.

Figure 2 shows the posterior rank distributions for selection functions  $w_{ij}$  and their associated credibility intervals generated by WinBUGS as part of the analysis for animal 8. Higher ranks relate to higher probability of selection in this presentation. Credibility intervals for rank use for each habitat for animal 8 are given in Table 4. Note that habitat 1, small grain, is ranked somewhere from 2<sup>nd</sup> to 4<sup>th</sup>, thus, it is not the most selected habitat by this animal nor the least selected by this animal. Habitat 5, idle, with rank 1 is the least selected habitat by bird 8.

Table 5 shows model selection results for each bird and across all birds using the two forms of Bayesian model selection statistics, Bayes factors and Bayesian posterior predictive values. Significance of Bayes factor results follow

Table 2. Estimates of selection for animal 8

	Small Grain	Row Crop	Hay	Pasture	Idle
Relocations	10	26	2	8	0
Observed % use	.217	.565	.043	.174	.000
% Availability	.282	.417	.102	.135	.063
Modeled % use					
normal priors	.217	.565	.044	.174	.001
uniform priors	.215	.566	.043	.174	.002
W = u/a (selection ratio)	.771	1.355	.425	1.288	.000
Scaled Bayesian w					
normal priors	.773	1.372	.416	1.269	.008
uniform priors	.762	1.367	.412	1.261	.036

Table 3. A comparison of 95% confidence intervals and 95% credibility intervals (normal priors) and interval widths for animal 8

Habitat	$W \pm Z(1-\alpha/2) \cdot se(W)$	width	Bayesian interval	width
Small Grain	(0.349, 1.195)	0.846	(0.381, 1.267)	0.886
Row Crop	(1.030, 1.718)	0.688	(0.882, 1.933)	1.051
Hay	(-0.159, 0.995)	1.154	(0.056, 1.038)	0.982
Pasture	(0.456, 2.078)	1.622	(0.652, 1.941)	1.289
Idle	-----	0.0	(0.000, 0.093)	0.093

Figure 1. Posterior densities for selection functions for habitat 1 (small grain), 2 (row crop), 3 (hay), 4 (pasture) and 5 (idle) for partridge number 8 using log normal priors. Note that the idle habitat had zero use by partridge number 8; normal priors.

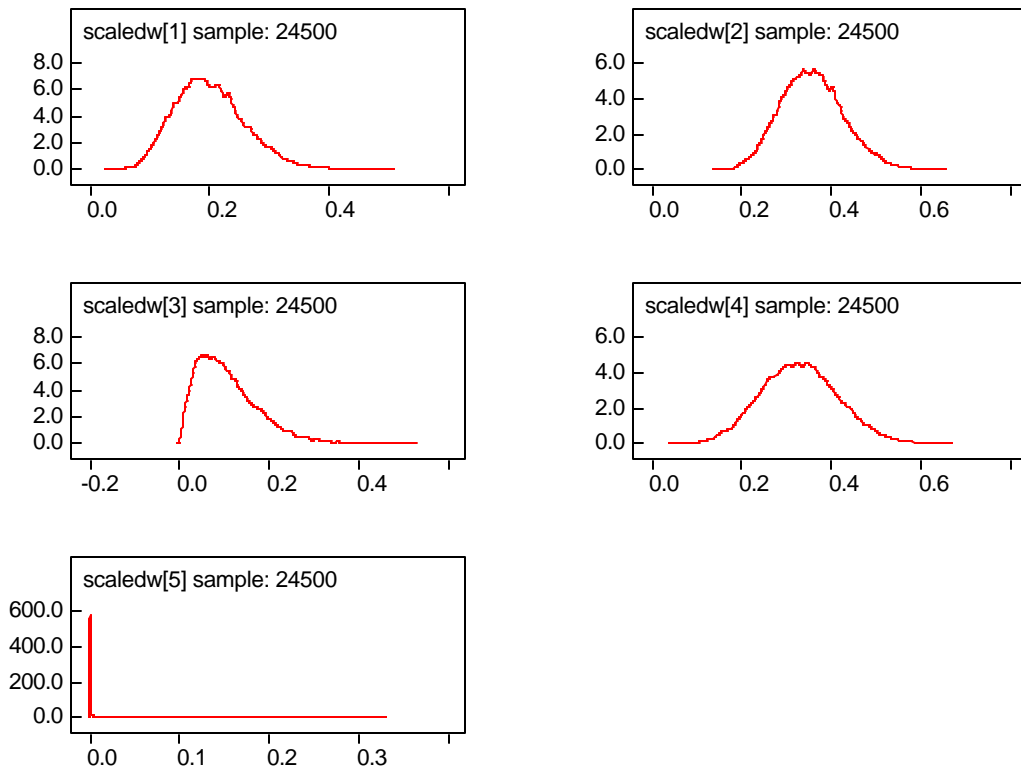


Figure 2. Posterior rank distributions for selection functions for partridge number 8. 1 = small grain, 2 = row crop, 3 = hay, 4 = pasture and 5 = idle; normal priors.

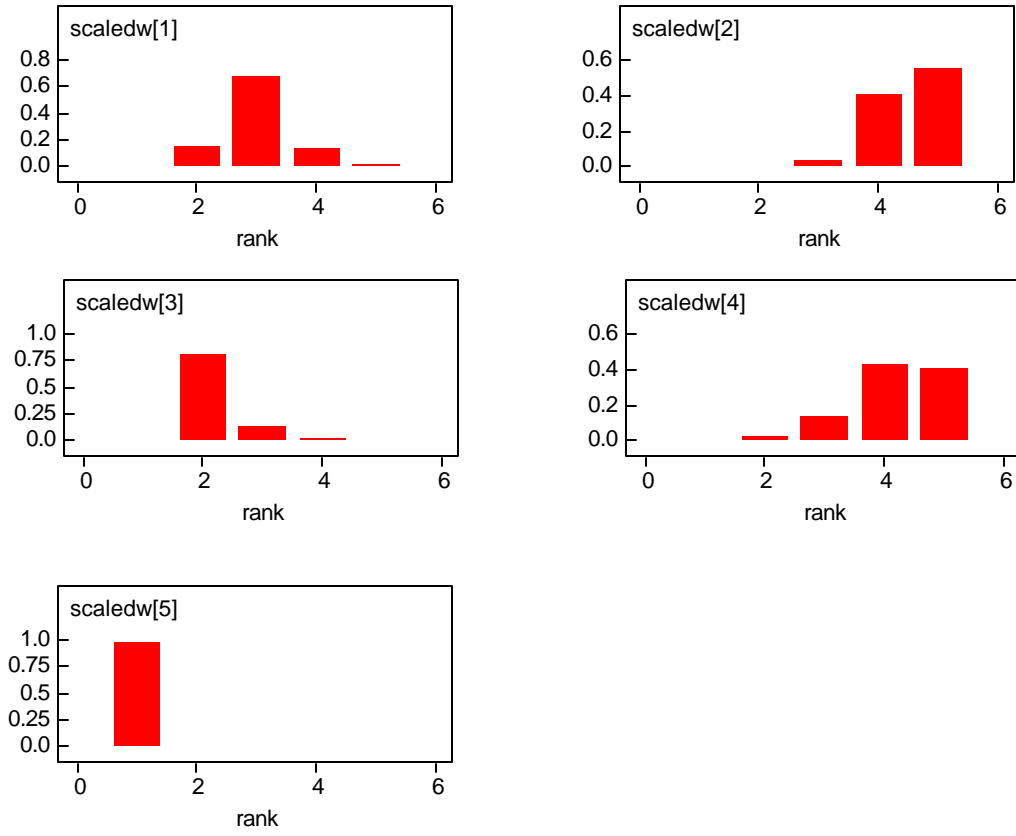


Table 4. Posterior rank distribution summary of selection function for partridge 8; normal priors.

	Rank Selection		
	2.5%	median	97.5%
Small Grain	2	3	4
Row Crop	3	5	5
Hay	2	2	4
Pasture	3	4	5
Idle	1	1	1

Table 5. Bayes factors and Bayesian posterior predictive values for assessing selection for each animal and across animals; normal priors.

Animal	Small Row					Bayes Factor	Bayesian posterior predictive values
	Grain	Crop	Hay	Pasture	Idle		
1	0	8	0	20	2	***	<.001
2	25	21	0	0	1	***	.001
3	17	11	0	0	2	**	.010
4	4	0	0	0	2	*	.022
5	20	0	0	9	0	***	<.001
6	22	0	0	2	0	***	<.001
7	0	7	6	0	1	*	.007
8	10	26	2	8	0	n.s.	.078
9	21	0	4	0	3	***	<.001
10	44	1	0	0	5	***	<.001
Overall (with and without animal 8)						***	<.001

\*\*\* = very strong evidence, \*\* = strong evidence, \* = evidence, n.s. = not significant

the guidelines given by Kass and Raftery (1995). Note that the Bayes factor for animal 8 agrees with the joint chi-square analysis by Dasgupta and Alldredge (2002) while Bayesian posterior predictive values do not (their analysis indicated non significance at the 10% level while we got a Bayesian posterior predictive value of .078). Thus, the results are close but different. We note that the Dasgupta and Alldredge (2002) approach is based upon large sample theory while the Bayesian posterior predictive value approach is not. Significance levels using uniform priors were within .002 of the levels given for normal priors.

The multiple comparison procedures of the White and Garrott (1990) and the Dasgupta and Alldredge (2002) technique are compared to a Bayesian posterior predictive value approach in Table 6. The Bayesian method generally identified fewer differences in selection between habitats than the other methods. We believe the differences between methods are largely caused by the fact that both the White and Garrott and Dasgupta and Alldredge methods rely on large sample normal theory while the Bayesian methods here do not.

## DISCUSSION AND CONCLUSIONS

The Bayesian model described here fits the example data well and has several advantages over existing methods. Credibility intervals and model selection methods are sample size specific and do not rely on large sample normal theory. In addition, credibility intervals can not yield negative lower bounds and zero use results in a meaningful interval.

We found it is challenging to determine Bayes factor values in fitting models for individual birds and across animals. Commonly only a lower bounds could be determined. In addition the determination of the Bayes factor must be made in a separate run from the model fitting in WinBUGS. This is necessary because both a null model and an alternative model are compared in one analysis. Because the parameters of the models may not have the same meaning under both the null and alternative models, the Bayes factor run cannot be used to estimate model parameters other than the Bayes factor. Thus, we consider these disadvantages of this model selection method and prefer Bayesian posterior predictive values for hypothesis testing here.

Our model can be extended to other common resource selection situations. First, replacing the common availability of resource  $j$  for all animals,  $a_j$ , with a value that is animal specific,  $a_{ij}$ , allows for different availability for each animal. Second, if availability is not known but is estimated through random point sampling, the vector of availabilities can be modeled as a multinomial random variable as part of a WinBUGS analysis. Third, dependencies among animals, e.g., animals in groups, can be modeled using a multivariate normal distribution for the  $a_{ij}$  rather than the independent univariate normal distributions we used. The structure of the covariance matrix in this case is given in Aitchison (1986). In addition, we can test for selection differences among gender or age groups by including additional parameters in our model such as in

$$w_{ij} = e^{a_{ij} + g_i}$$



Table 6. Multiple comparisons; S = small grain, R = row crop, H = Hay, P = pasture, and I = idle. W & G refers to the procedure by White and Garrott and D & A refers to the procedure by Dasgupta and Alldredge; normal priors.

Animal	W & G	D&A	Bayesian
1	<u>S R H I</u> P	<u>S R H I</u> P	<u>S R H I</u> P
2	P <u>H I R</u> S	P <u>H I R</u> S	P <u>H I R</u> S
3	P <u>H R I</u> S	P <u>H R I</u> S	P <u>H R I</u> S
4	R P <u>H I</u> S	R P <u>H I</u> S	R P H S I
5	R <u>H I</u> P S	R H I <u>P</u> S	<u>R H I</u> <u>P</u> S
6	R <u>H I</u> P S	R <u>H I</u> P S	R <u>H I</u> P S
7	<u>S P I</u> R H	<u>S P I</u> R H	<u>P S I</u> R H
8	<u>S I H</u> P R		
9	R P <u>H I</u> S	R P <u>H I</u> S	R P <u>H I</u> S
10	R P <u>H I</u> S	R P <u>H I</u> S	<u>P H R I</u> S * * *

\* - Due to the small number of relocations in the asterisked habitats, differences in selection were not detected for pasture, hay and small grain even though other ordered differences were significant.

## REFERENCES

- AEBISCHER, N. J., ROBERTSON, P. A., AND KENWARD, R. E. (1993a), Compositional analysis of Habitat Use From Animal Radio-Tracking Data, *Ecology* **74**: 1313-1325.
- AGRESTI, A. (1990), *Categorical Data Analysis*, John Wiley & Sons, New York.
- AITCHISON, J. (1986), *The Statistical Analysis of Compositional Data*, Chapman and Hall, New York.
- ALLDREDGE, J. R. AND RATTI, J. T. (1986), Comparison of Some Statistical Techniques for Analysis of Resource Selection, *Journal of Wildlife Management*, **50**, 157-165.
- ALLDREDGE, J. R., THOMAS, D. L., AND McDONALD, L. L. (1998). Survey and comparison of methods for study of resource selection, *Journal of Agricultural, Biological, and Environmental Statistics* **3**: 237 - 253.
- BARMI, H.E. AND PONTIUS, J.S. (2000). Testing Ordered Hypotheses in Animal Resource Selection Studies, *Journal of Agricultural, Biological, and Environmental Statistics* **5**: 88 - 101.
- CASELLA, G. AND GEORGE, E. I. (1992). Explaining the Gibbs Sampler, *The American Statistician*, **46**, 167 - 174.
- CONGDON, P. (2001). *Bayesian Statistical Modeling*, John Wiley & Sons, LTD, New York.
- DASGUPTA, N., AND ALLDREDGE, J. R. (1998), A Multivariate  $\chi^2$  analysis of resource selection data, *Journal of*

- Agricultural, Biological, and Environmental Statistics* **3**: 323-334.
- DASGUPTA, N., AND ALLDREDGE, J. R. (2002), A Single Step Method for Identifying Individual Resources, *Journal of Agricultural, Biological, and Environmental Statistics* **7**: 208 - 221.
- GILKS, W. R., THOMAS, A., AND SPIEGELHALTER, D. J. (1994), A Language and Program for Complex Bayesian Modelling, *The Statistician*, **43**, 169 - 178.
- KASS, R. E., AND RAFTERY, A. E. (1995), Bayes Factors, *Journal of the American Statistical Association*, **90**, 773 - 795.
- MANLY, B. F. J., McDONALD, L. L., THOMAS, D.L., McDONALD, T.L., AND ERICKSON, W.P. (2002). *Resource Selection by Animals; Statistical Design and Analysis for Field Studies: Second Edition*. Chapman and Hall, London.
- SPIEGELHALTER, D. J., THOMAS, A., AND BEST, N. G. (1999). *WinBUGS Version 1.2 User Manual*. MRC Biostatistics Unit.
- SMITH, L.E., HUPP, J. W., AND RATTI, J.T. (1982) Habitat use and home range of Grey Partridge in Eastern South Dakota.. *Journal of Wildlife Management*, **46**, 580-587.
- WHITE, G.C., AND GARROTT, R. A. (1990), *Analysis of Wildlife Radio-Tracking Data*, San Diego: Academic.