STATISTICAL TESTING OF SURFACE COLLECTED AND EXCAVATED FAUNAL SAMPLES FROM THE PLATER-MARTIN SITE

Peter Hamalainen

ABSTRACT

Three faunal samples from the Plater-Martin site (BdHb-1) were compared using the Spearman's Rank Correlation Coefficient and the Student's t Test. Two of the samples were surface collected and the third was excavated. As far as the mammalian, avian and reptilian remains were concerned, the surface collected samples compared favourably with the excavated sample. A probable explanation for the poor representation of fish remains in the surface collected samples is the collection technique.

PURPOSE

Surface collected faunal samples are biased in favour of the larger, more easily noticeable specimens. The result is that a surface collected faunal sample is often suspected of being sufficiently unrepresentative to serve as a useful indicator of a site's faunal assemblage. The purpose of this paper is to test this hypothesis by statistical means in order to determine the degree of similarity or dissimilarity between surface collected and excavated faunal samples from the Plater-Martin site (BdHb-1).

METHODOLOGY

The Plater-Martin site is a historic Petun village located in Concession II, Lot 20 of Collingwood Township, Grey County, Ontario. The site was apparently occupied from circa 1639 to 1650 and is believed to have been the location of the Jesuit mission of St. Simon and St. Jude (Garrad 1980). Five midden loci have been found at the site, all of which have been surface collected and two, Middens 2 and 5, have been partially excavated.

Three faunal samples were selected for comparison. The first was an excavated sample from Midden 5. It was recovered from 3 five-foot squares dug in arbitrary 6-inch levels. All dirt was screened but no flotation was undertaken. The second sample consisted of surface collected remains from Midden 5, while the third was composed of all the surface collected faunal material from the site and includes the surface sample from Midden 5 (Hamalainen 1981; Wodinskv 1979).

Two statistical tests were used in the comparisons of the samples. These were the Spearman's Rank Correlation Coefficient and the Student's t Test.

The Spearman's Rank Correlation Coefficient is a common nonparametric test used in behavioural sciences which measures the degree of similarity or difference between two given samples. The variables are assigned ranks from the highest to the lowest. In this case, species were ranked by their elemental frequency, with the most frequent species in a sample assigned the rank of one, the second most frequent species the rank of two and so on. When

Hamalainen, Peter
1983 Statistical Testing of Surface Collected and Excavated Faunal Samples from the Plater-Martin Site. *Ontario Archaeology* 39: 57-64

two or more species had the same elemental frequency, the mean of all the ranks they would have otherwise occupied was taken and given to all. The results range from + 1, which represents perfect similarity, to - 1, which represents perfect dissimilarity.

An advantage of the Spearman's Rank Correlation Coefficient is that the results can be checked for validity. This has been done by using the Student's t Test. The level of significance chosen was .01, which is to say that if the results of the test were significant at this level, there is a 99% chance that the coefficient is statistically valid.

The frequencies of the species variables used in this paper, expressed as percentages, and their rankings are presented in Tables 1-3. An example of the calculations made for this paper is given in Table 4. For a full discussion of the mechanics of both the Spearman's Rank Correlation Coefficient and the Student's t Test see Siegel (1956:202-213).

The faunal samples consisted of only those vertebrate remains identified to taxa below the class level. Although the aim of faunal analysis is to identify each specimen as to the species, this is often not possible. The result is that specimens which in fact belong to the same species may be scattered over a number of taxa (order, family, genus and species). As one of the key elements in the Spearman's Rank Correlation Coefficient is the total number of ranks, the greater number of taxa for what is actually the same species produces greater sampling error. To partly compensate for this some identifications were combined together at the order, family or genus levels. Eastern Cottontail, Snowshoe Hare and Hare sp. were included under the heading Hare sp., Red Squirrel and Squirrel sp. under Squirrel sp., Dog, Wolf and *Canis* sp. under *Canis* sp., Red Fox and Fox sp. under Fox sp., *Anseriformes* sp., Goose sp. and Canada goose under Anseriforrmes sp., Northern Pike and Pike sp. under Pike sp., Longnose Sucker and Sucker sp. under Sucker sp. and Largemouth Bass and Bass sp. under Bass sp.

The faunal samples were divided into 3 groups when the statistical tests were applied to them. The first group includes all the classes in the calculations. The second group consisted of all the mammalian, avian and reptilian bone, while the third group was composed of the fish bone only. The reason for the separation of fish from the other classes is that fish remains were suspected of being the most susceptible to sampling error.

TABLE 1
TOTAL FAUNAL SAMPLE-PERCENTAGE AND RANKING

	Midden 5 Excavated		Midden 5 Surface		Total Surface	
	%	Rank	%	Rank	%	Rank
Human	0	40.5	0	36.5	1.27	8.5
Hare sp	0.37	25.5	0.59	15.5	0.56	18
Squirrel sp	0.37	25.5	0	36.5	0.14	31
Eastern Chipmunk	0	40.5	0.29	23	0.14	31
Woodchuck	8.05	4.	11.34	4	12.28	4
Beaver	17.97	3	17.31	3	20.90	2
Muskrat	0.18	33.5	0.59	15.5	0.28	22
Porcupine	0.37	25.5	0	36.5	0	41.5
Carnivore sp	0	40.5	0.59	15.5	1.83	6.5
Canis sp	22.28	1	20.89	2	16.66	3
Fox sp	0	40.5	0.59	15.5	0.84	12
Black Bear	19.18	2	23.88	1	28.10	1

Table 1 continued.

Raccoon	1.87	10.5	2.68	6	1.83	6.5
Mustelidae sp	0	40.5	0	36.5	0.28	22
Marten	0.74	17	0.29	23	0.28	22
Fisher	0.18	33.5	0.29	23	0.14	31
Striped Skunk	0	40.5	0	36.5	0.14	31
River Otter	0	40.5	0	36.5	0.28	22
Cervidae sp	2.80	7	1.49	9.5	0.84	12
White-tailed Deer	5.05	5	8.95	5	5.22	5
Moose	0.37	25.5	0.29	23	0.14	31
Elk	1.12	14.5	0.89	12	0.70	15.5
Loon sp	0.37	25.5	0.29	23	0.14	31
Anseriformes sp	2.05	9	1.79	7	1.27	8.5
Aythyinae sp	0	40.5	0	36.5	0.14	31
Hawk sp	0.18	33.5	0	36.5	0	41.5
Bald Eagle	0.37	25.5	0	36.5	0	41.5
Ruffed Grouse	0.37	25.5	0	36.5	0.14	31
Turkey	0.37	25.5	0.59	15.5	0.42	19
Sandhill Crane	1.49	13	0	36.5	0	41.5
Passenger Pigeon	1.87	10.5	0	36.5	0	41.5
Common Crow	0.37	25.5	0	36.5	0	41.5
Turtle sp	1.12	14.5	1.49	9.5	0.98	10
Snapping Turtle	0	40.5	0	36.5	0.14	31
Painted Turtle	0.37	25.5	1.49	9.5	0.70	15.5
Lake Sturgeon	0.56	18.5	0	36.5	0.84	12
Trout sp	0.56	18.5	0.29	23	0.14	31
Pike sp	0.93	16	0	36.5	0	41.5
Sucker sp	0.37	25.5	0.29	23	0.14	31
Perciformes sp	0	40.5	0.29	23	0.14	31
Bass sp	0.37	25.5	1.49	9.5	0.70	15.5
Percidae sp	0.18	33.5	0	36.5	0	41.5
Yellow Perch	2.99	6	0.29	23	0.14	31
Stizostedion sp	2.43	8	0.59	15.5	0.70	15.5
Freshwater Drum	1.68	12	0	36.5	0	41.5

Total number of ranks is 45

TABLE 2
MAMMALIAN, AVIAN AND REPTILIAN SAMPLE-PERCENTAGE AND RANKING

	Midden 5 Excavated		Midden 5 Surface		Total Surface	
	%	Rank	%	Rank	%	Rank
Human	0	30.5	0	28.5	1.31	8.5
Hare sp	0.42	18.5	0.62	14	0.58	15
Squirrel sp	0.42	18.5	0	28.5	0.14	26
Eastern Chipmunk	0	30.5	0.31	19	0.14	26
Woodchuck	8.96	4	11.73	4	12.65	4

Table 2, continued:

Beaver	20.00	3	17.90	3	21.51	2
Muskrat	0.21	25	0.62	14	0.29	19
Porcupine	0.42	18.5	0	28.5	0	33
Carnivore sp	0	30.5	0.62	14	1.89	6.5
Canis sp	24.79	1	21.61	2	17.15	3
Fox sp	0	30.5	0.62	14	0.87	11.5
Black *Bear*	21.25	2	24.69	1	28.92	1
Raccoon	2.08	8.5	2.78	6	1.89	6.5
Mustelidae sp	0	30.5	0	28.5	0.29	19
Marten	0.83	13	0.31	19	0.29	19
Fisher	0.21	25	0.31	19	0.14	26
Striped Skunk	0	30.5	0	28.5	0.14	26
River Otter	0	30.5	0	28.5	0.29	19
Cervidae sp	3.13	6	1.54	9	0.87	11.5
White-tailed *Deer*	5.63	5	9.26	5	5.38	5
Moose	0.42	18.5	0.31	19	0.14	26
Elk	1.25	11.5	0.93	11	0.73	13.5
Loon sp	0.42	18.5	0.31	19	0.14	26
Anseriformes sp	2.29	7	1.85	7	1.31	8.5
Aythyinae sp	0	30.5	0	28.5	0.14	26
Hawk sp	0.21	25	0	28.5	0.14	26
Bald Eagle	0.42	18.5	0	28.5	0	33
Ruffed Grouse	0.42	18.5	0	28.5	0.14	26
Turkey	0.42	18.5	0.62	14	0.44	16
Sandhill *Crane*	1.67	10	0	28.5	0	33
Passenger Pigeon	2.08	8.5	0	28.5	0.29	19
Common Crow	0.42	18.5	0	28.5	0	**33**
Turtle sp	1.25	11.5	**1.54**	9	1.02	10
Snapping Turtle	0	30.5	0	28.5	0.14	26
Painted Turtle	0.42	18.5	**1.54**	9	0.73	13.5

Total number of ranks is 35.

TABLE *3*

FISH BONE SAMPLE-PERCENTAGE AND RANKING

	Midden 5 Excavated		Midden 5 *Surface*		Total *Surface*	
	%	*Rank*	*%*	*Rank*	*%*	*Rank*
Lake Sturgeon	5.55	5.5	0	8.5	30.00	1
Trout sp	5.55	5.5	9.09	4.5	5.00	5.5
Pike sp	9.25	4	0	8.5	0	9
Sucker sp	3.70	7.5	9.09	4.5	5.00	5.5
Perciformes sp	0	10	9.09	4.5	5.00	5.5
***Bass** sp*	3.70	7.5	45.45	1	25.00	2.5
Percidae sp	1.85	9	0	8.5	0.00	9
Yellow *Perch*	29.62	1	9.09	4.5	5.00	5.5
Stizostedion sp	24.07	2	18.18	2	25.00	2.5
Freshwater Drum	16.66	3	0.	8.5	0.00	9

Total number of ranks *is* 10.

TABLE 4
AN EXAMPLE OF THE CALCULATION OF THE SPEARMAN'S RANK CORRELATION COEFFICIENT AND THE STUDENT'S t TEST-TOTAL GROUP: MIDDEN 5 EXCAVATED SAMPLE VS. MIDDEN 5 SURFACE SAMPLE

D	D^2
4	16
10	100
11	121
17.5	306.25
0	0
0	0
18	324
11	121
25	625
1	1
25	625
1	1
4.5	20.25
4	16
6	36
10.5	110.25
4	16
4	16
2.5	6.25
0	0
2.5	6.25
2.5	6.25
2.5	6.25
2	4
4	16
3	9
11	121
11	121
10	100
23.5	552.25
26	676
11	121
5	25
4	16
16	256
18	324
4.5	20.25
20.5	420.25
2.5	6.25
17.5	306.25
16	256
3	9
17	289
7.5	56.25
24.5	600.25
	6784.5

Spearman's Rank Correlation Coefficient:

$$R_S = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$= 1 - \frac{6(6784.5)}{45(45^2-1)}$$

$$= +0.56$$

Student's t Test:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

$$= \frac{.56\sqrt{45-2}}{\sqrt{1-56^2}}$$

$$= 4.432$$

$$< 0.01$$

OBSERVATIONS

Three observations were noted for the coefficients of all the groups. The first was the sign of the coefficients, that is, whether they were positive or negative. The second is the strength of the coefficients and the third is their statistical validity as indicated by the Student's t Test.

Total group: All coefficients were positive. The strongest coefficient was between the 2 surface samples. The excavated Midden 5 sample was more strongly correlated with the Midden 5 surface sample than with the total surface sample (see Table 5). The values given by the Student's t Test indicated that all coefficients were significant at the .01 level.

TABLE 5
COEFFICIENTS FOR TOTAL FAUNAL SAMPLE

	Total Surface	Midden 5 Surface	Midden 5 Excavated
Total Surface	-	+0.80	+0.41
Midden Surface	+0.80	-	+0.56
Midden 5 Excavated	+0.41	+0.56	-

Mammalian, avian and reptilian group: Again, all coefficients were positive. The relationships of the strengths of the coefficients between the samples was the same as above, but with higher values (see Table 6). As above, all the values given by the Student's t Test were significant at the .01 level.

TABLE 6
COEFFICIENTS FOR THE
MAMMALIAN, AVIAN AND REPTILIAN SAMPLE

Total Surface	Total Surface	Midden 5 Surface	Midden 5 Excavated
Total Surface	-	+0.85	+0.52
Midden 5 Surface	+0.85	-	+0.69
Midden 5 Excavated	+0.52	+0.69	-

Fish group: Once again, all coefficients were positive. However, the coefficient between the 2 surface samples was considerably higher than those between the excavated sample and the surface samples (see Table 7). In all cases the Student's t Test produced values which were not significant at the .01 level.

TABLE 7
COEFFICIENTS FOR THE FISH SAMPLE

	Total Surface	Midden 5 Surface	Midden 5 Excavated
Total Surface	-	+0.62	+0.09
Midden 5	+0.62	-	+0.07
Midden 5 Excavated	+0.09	+0.07	-

CONCLUSIONS

In the total and the mammalian, avian and reptilian groups the coefficients were all strongly positive. In both groups the highest coefficients were between the *2* surface samples. This is as may be expected as the Midden 5 surface collected sample accounts for 47.31% of the total surface sample. The coefficients between the Midden 5 excavated and Midden 5 surface collected samples were also strongly positive. This too is as might be expected as both samples were drawn from the same raw data base. The strong coefficient also indicates that a carefully collected surface sample from Midden 5 serves as a reliable indicator of the mammalian, avian and reptilian samples in the midden. The weakest correlation was between the excavated Midden 5 sample and the total surface collected sample. What this most likely reflects is the differential deposition of faunal remains in the middens. The Student's t Test indicated that all the above findings are significant at the .01 level.

It is interesting to note that the coefficients in the mammalian, avian and reptilian group were higher than in the total group. This was due to the removal of the "interference" caused by the fish remains in the prior category. The fish group appears to be seriously affected by sampling error, as is indicated by the values derived from the Student's t Test which show that the coefficients for the fish group are not significant at the .01 level.

The most likely explanation for the sampling error affecting the fish group is the recovery method. Even though they were carefully collected, the surface samples do not truly reflect the findings made in the excavated sample. For example, the amount of fish bone in the excavated Midden 5 sample amounted to 10.07%, while the percentage of fish bone in the surface collected sample from the same midden was only 3.24%. It would seem, then, that faunal findings based on carefully collected surface samples are adequate for the interpretation of mammalian, avian and reptilian samples, but inadequate for the interpretation of fish samples.

At present the above conclusions apply only to the Plater-Martin site. Before it could be shown whether the findings made here be the exception or the rule for other surface collected and excavated faunal samples, more statistical testing of faunal assemblages from other sites should be carried out.

ACKNOWLEDGEMENTS

The writer would like to thank Dr. Conrad Heidenreich for his encouragement and guidance during the writing of this paper, Mr. Charles Garrad for making the faunal samples available for study and Dr. Howard Savage for granting access to the comparative material in the Faunal Archaeo-Osteology Collection at the University of Toronto.

REFERENCES CITED

Garrad, Charles
1980 Petun Pottery. In Proceedings of the 1979 Iroquois Pottery Conference. *Rochester Museum and Science Centre, Research Division, Research Report* No. 13.

Hamalainen, Peter
1981 Faunal Analysis of Twelve Petun Sites. Unpublished manuscript.

Siegel, Sidney
1956 *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill Book Company.

Wodinsky, Harold
1979 Faunal Analysis of the Plater-Martin Site (BdHb-1), Midden Five. Unpublished manuscript.

216 Woodmount Avenue
Toronto, Ontario M4C 3Z6