

# ***p*-Values for Postmortem Intervals from Arthropod Succession Data**

Lynn Roy LAMOTTE and Jeffrey D. WELLS

Two approaches, one based on the likelihood-ratio statistic and the other based on unconditioning Fisher's exact test, are examined for obtaining a *p*-value in the comparison of the combination of arthropod species present on a mystery carcass to the observed frequency distribution of species combinations on carcasses exposed to the elements for a known time interval.

**Key Words:** Contingency tables; Fisher's exact test; Forensic entomology; Likelihood-ratio statistic; Multinomial outliers.

## **1. INTRODUCTION**

The time since death, or postmortem interval (PMI), often is of crucial importance in investigations of untimely deaths. Physical and biological changes that depend on time since death can provide useful clues. During the first day or two, the physical and histochemical consequences of death are usually the most reliable PMI indicators (Hennsge, Madea, Knight, Nokes, and Krompecher 1991), but they are less useful after that, when more accurate results may be obtained from ecological information concerning organisms on or in the vicinity of the body.

Two time-dependent processes make insects useful in estimating PMI. One is the growth of species whose larvae eat carrion, such as flies in the families Calliphoridae, Sarcophagidae, and Muscidae, and beetles in the families Silphidae and Dermestidae (Smith 1986). Sizes of larvae gathered from the carcass can be compared to experimental data on growth patterns in order to estimate the age of the larvae and hence to estimate a lower bound on PMI (adult females generally do not deposit eggs on a live body). Long used to estimate PMI (Catts and Goff 1992), only recently have statistical procedures been developed to assess the accuracy of estimates based on sizes of larvae (Wells and LaMotte 1995).

---

Lynn LaMotte is from the Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA 70803-5606. Jeffrey Wells is from the Department of Justice Sciences, University of Alabama at Birmingham, Birmingham, AL 35294-2060.

©2000 American Statistical Association and the International Biometric Society  
*Journal of Agricultural, Biological, and Environmental Statistics*, Volume 5, Number 1, Pages 58-73 (e)

The second time-dependent process is called succession. During the progressive decay of a carcass exposed to the elements, there is a fairly regular pattern of arthropod taxa present at a given time and the list of taxa changes over time. Presented with a carcass exposed for an unknown time, knowledge of this pattern of succession can be used to get some idea of the PMI (Schoenly and Reid 1987; Schoenly 1992). In contrast to larval growth, succession models can be used to estimate both the least and the greatest possible PMI.

In the remainder of this paper, we develop and investigate statistical procedures for obtaining  $p$ -values for PMI estimates based on succession data. This requires a probability model, developed in the next section, that incorporates variation inherent both in the experimental data that provide the basis for the estimate and in the pattern of species observed to be present on the mystery carcass. The third and fourth sections formulate extreme sets for two test statistics, and the fifth section shows some numerical comparisons.

Clearly, this setting is biologically complex, with many environmental factors that can affect succession patterns (Smith 1986). Accurate PMI estimation requires both careful reconstruction of crime scene conditions and the availability of matching reference data. Pigs have recently become the most common model for human decay used by forensic entomologists in North America. Reference data may come from experiments under conditions likely to be useful in the future (e.g., Anderson and VanLaerhoven 1996) or from experiments that try to duplicate conditions at a crime scene after the fact (e.g., Goff 1992).

In practice, it is not unusual to find several dozen species present on a carcass. To obtain reference data with sufficient sample sizes to provide useful detail on succession patterns of so many species over a range of potential crime scene conditions would endanger the world's pig population. Clearly, further work must be done to identify a practicable number of categories of species combinations that can be used for this purpose.

The succession process is complex. At the end, though, it requires a comparison between one observed combination of species and a frequency distribution of species combinations observed under controlled experimental conditions. Our objective is to provide basic statistical tools that can be used in this comparison.

## 2. A PROBABILITY MODEL FOR SUCCESSION DATA

In an experiment to obtain reference data,  $n$  carcasses are exposed and a list of species present on each carcass is compiled on each of several successive days. For example, consider hypothetical field data resulting from exposing  $n = 10$  carcasses and observing the presence or absence of two species for 7 days. Call the species A and B. There are four possible species combinations: A and not B (denote this by  $AB'$ ), A and B ( $AB$ ), B and not A ( $A'B$ ), and not A and not B ( $A'B'$ ). Table 1 shows the occurrence data that might result.

According to Table 1, combination  $AB'$  was observed on all 10 carcasses on day 1; on day 2,  $AB'$  occurred on 1 carcass,  $AB$  on 6 carcasses, and  $A'B'$  occurred on 3 of the 10 carcasses.

Suppose a mystery carcass is found and it has combination  $AB'$ , i.e., species A was found on it but species B was not. Clearly, the PMI for which this outcome is most likely

Table 1. Possible Occurrence Data

Species	Day						
	1	2	3	4	5	6	7
AB'	10	1	0	1	1	1	0
AB	0	6	10	5	0	0	1
A'B	0	0	0	4	8	3	1
A'B'	0	3	0	0	1	6	8

is 1 day, and it appears to be considerably less likely on other days. The statistical problem is to attach a  $p$ -value to each day. Consider, e.g., day 7. If 11 carcasses (corresponding to the 10 in the experiment and the mystery carcass) were exposed for 7 days, what is the probability that the 11th would differ from the first 10 as much as, or more than, the observed results (namely, frequencies 0, 1, 1, 8 for the four species combinations among the first 10 and 1, 0, 0, 0 for the 11th)? The answer to this question is the  $p$ -value that we seek.

It is apparent that the comparison between the training sample on day 7 and the mystery carcass can be represented as a contingency table, as shown in Table 2, where the objective is to test whether the distribution of the four species combinations (categories) is the same in the population (with unknown PMI) from which the mystery carcass came as it is in the population (with known PMI of 7 days) from which the training sample came.

The  $p$ -value for the chi-squared test of homogeneity in this table is 0.012 and the  $p$ -value for Fisher's exact test (FET) is 0.273. The chi-squared  $p$ -value is a large-sample approximation, and it can be shown that it is a poor approximation in this case. We shall argue below that the FET  $p$ -value is unnecessarily conservative. Our objective is to develop and evaluate other approaches that yield exact  $p$ -values for this setting.

A  $p$ -value comes from a probability model, which can be derived as a sampling model. Consider a population of subjects corresponding to the (conceptual) population of all carcasses exposed for 7 days. Each subject belongs to exactly one of  $c$  categories, with the categories corresponding to the species combinations. In this population, the relative frequency (proportion) of subjects in category  $i$  is  $\pi_i, i = 1, \dots, c$ . Let  $\pi = (\pi_1, \dots, \pi_c)$ . Denote each subject's response by  $\mathbf{Y}$ , a  $c$ -vector of zeros except for a one in the  $k$ th place, indicating that the subject is in category  $k$ . Thus,  $\mathbf{Y} = (Y_1, \dots, Y_c), Y_j \in \{0, 1\}, \sum_{j=1}^c Y_j = 1$ . Denote by  $\pi_* = (\pi_{*1}, \dots, \pi_{*c})$  the relative frequency distribution of cate-

Table 2. Contingency Table

Category	Training sample	Mystery	All
1	0	1	1
2	1	0	1
3	1	0	1
4	8	0	8
All	10	1	11

gories in the population from which the mystery carcass comes. The joint probability of realizing responses  $\mathbf{y}_i = (y_{i1}, \dots, y_{ic})$ ,  $i = 1, \dots, n$ , for  $n$  independently chosen subjects and  $\mathbf{y}_* = (y_{*1}, \dots, y_{*c})$  for the  $(n + 1)$ st subject is

$$\pi_1^{f_1} \pi_2^{f_2} \dots \pi_c^{f_c} \pi_{*\ell},$$

where  $f_j = \sum_{i=1}^n y_{ij}$  is the frequency of the first  $n$  subjects in category  $j$  and  $\ell$  is the category of the  $(n + 1)$ st subject (i.e.,  $y_{*\ell} = 1$ ). Let  $\mathbf{f} = (f_1, \dots, f_c)$  denote the vector of these frequencies. Now, given  $\pi$  and  $\pi_*$ , we can determine exactly the probability of any subset of the sample space. Clearly,  $(\mathbf{f}, \ell)$  is a sufficient statistic, and so we shall use the term outcome  $(\mathbf{f}, \ell)$  to mean the set of outcomes with frequency distribution  $\mathbf{f}$  and mystery specimen in category  $\ell$ . The probability distribution of  $\mathbf{f}$  is multinomial with parameters  $n$  and  $\pi$ .

It remains to define the subset of the sample space for which the probability is to be computed for the  $p$ -value. Heuristically, it is the set of outcomes as extreme as, or more extreme than, the observed outcome. The direction "more extreme than" is to be taken in relation to the condition  $\pi = \pi_*$ . We have not found a universal definition of this direction. (See Kempthorne and Folkes [1971] for a good discussion of this issue.) Most commonly, it is taken as the direction of decreasing probability so that outcomes are more extreme if they have less probability. Often the  $p$ -value is defined in terms of a test statistic, being the least level of significance at which the hypothesis would be rejected. In this case, the direction of more extreme outcomes is the direction of increasing (in some cases, decreasing) values of the test statistic. Because many such tests are equivalent to likelihood-ratio tests, this suggests that the likelihood ratio might be used to define the direction of more extreme outcomes.

Other schemes for ordering outcomes are possible, and there do not seem to be universally accepted criteria for determining which is best. For testing equality of two binomial proportions, as in  $2 \times 2$  tables, the uniformly most powerful unbiased test (see Lehmann 1959, p. 140) entails the less probable ordering, but its use in larger tables, as implemented by Mehta and Patel (1983), e.g., does not seem to have such compelling theoretical justification.

Given an ordering of outcomes, either in the direction of decreasing probabilities or in the direction of increasing likelihood ratio or in some other direction, the set of outcomes as extreme as, or more extreme than, the observed outcome is defined. Call this set of outcomes the extreme set of the observed outcome. Then the  $p$ -value of the observed outcome is the probability of its extreme set, computed with  $\pi = \pi_*$ . However, this probability generally depends on  $\pi$ , so different values of  $\pi$  give different  $p$ -values. One possible resolution of this problem (it is a problem of noninvariance) is to find the maximum of this probability over all relative frequency distributions  $\pi$ , which corresponds to Lehmann's (1959) definition of the size of the test as the maximum power over all distributions in the null hypothesis.

The statistical problem we face here, with a training sample to be compared to one subject, is conceptually the same as testing whether an observed response is an outlier in a regression model. There the question is the same: If the  $(n + 1)$ st unit were sampled at

random from the same population from which the first  $n$  units were sampled, what would be the probability of results as extreme as, or more extreme than, those actually observed? A solution to this problem, then, is equivalent to developing a test for outliers in multinomial samples.

There is much relevant work in the literature for  $2 \times 2$  contingency tables, corresponding here to observing the presence or absence of just one species, but not for the present situation, where one of the columns has only one observation. Suissa and Shuster (1985) give a good review of these efforts. The notion of defining the  $p$ -value as the maximum probability of an extreme set over nuisance parameters is widespread. Although FET, randomized in order to fix the size of the test, is uniformly most powerful among unbiased tests of  $H_0: \boldsymbol{\pi} = \boldsymbol{\pi}_*$ , dissatisfaction with conditional tests (e.g., FET is too conservative) and randomized tests (Suissa and Shuster [1985, p. 317] say, "in practice, decisions are not based on irrelevant events") is frequently mentioned in the literature. Other papers that are particularly relevant are McDonald, Davis, and Milliken (1977) and Storer and Kim (1990).

### 3. THE LIKELIHOOD RATIO AND CHI-SQUARED $p$ -VALUE

Our objective is to formulate the likelihood ratio ordering of possible outcomes in order to define extreme sets. The likelihood function is

$$L(\boldsymbol{\pi}, \boldsymbol{\pi}_*; \mathbf{f}, \ell) = \pi_1^{f_1} \pi_2^{f_2} \cdots \pi_c^{f_c} \pi_{* \ell}^{\ell}, \quad 0 \leq \pi_j \leq 1, 0 \leq \pi_{*j} \leq 1, j = 1, \dots, c.$$

The likelihood ratio is the ratio of the maximum of  $L$  over  $\boldsymbol{\pi}$  and  $\boldsymbol{\pi}_*$  to the maximum of  $L$  over  $\boldsymbol{\pi} = \boldsymbol{\pi}_*$ , and it can be expressed as

$$\begin{aligned} R(f_\ell) &= n^{-n} (n+1)^{n+1} \quad \text{if } f_\ell = 0, \\ &= n^{-n} (n+1)^{n+1} \frac{f_\ell^{f_\ell}}{(f_\ell + 1)^{f_\ell + 1}} \quad \text{if } f_\ell > 0. \end{aligned}$$

The likelihood ratio can be used to order outcomes. An outcome is  $(\mathbf{y}_1, \dots, \mathbf{y}_n, \mathbf{y}_*)$ , a description that we have reduced to sets of outcomes specified by  $(\mathbf{f}, \ell)$ . Now, in terms of the likelihood ratio, sets of outcomes may be distinguished in terms of  $f_\ell$ . In this ordering, the extreme set for  $f_\ell$  consists of all outcomes  $(\mathbf{F}, j)$  such that  $R(F_j) \geq R(f_\ell)$ ; denote it by  $C_{LR}(f_\ell)$ . That  $R$  is a monotone decreasing function of the frequency  $f_\ell$  implies that  $R(F_j) \geq R(f_\ell)$  if and only if  $F_j \leq f_\ell$ . From these observations, it may be seen that

$$C_{LR}(x) = \bigcup_{j=1}^c \{(\mathbf{Y}_1, \dots, \mathbf{Y}_n, \mathbf{Y}_*) : Y_{*j} = 1 \text{ and } F_j \leq x\}.$$

The probability of  $C_{LR}(x)$  is

$$\begin{aligned} \Pr[C_{LR}(x)] &= \sum_{j=1}^c \Pr(Y_{*j} = 1) \Pr(F_j \leq x \mid Y_{*j} = 1) \\ &= \sum_{j=1}^c \pi_{*j} \text{CDF}_{\text{bin}}(x; n, \pi_j), \end{aligned}$$

where  $\text{CDF}_{\text{bin}}(x; n, \pi)$  is the binomial cumulative distribution function for  $n$  independent trials, each of which has a probability of success of  $\pi$ . This probability is a mixture of binomial probabilities for the observed frequency  $x$ . It is particularly simple because it is expressed in terms of the  $c$  marginal CDFs of the observed frequencies. Given  $\pi$ , it depends only on the frequency  $x$ , not on the category and not on the way the other frequencies are distributed.

The chi-squared statistic, mentioned earlier, has this same property. It can be shown that the chi-squared statistic is

$$\chi_{c-1}^2 = \frac{n+1}{n} \frac{n - f_\ell}{f_\ell + 1}.$$

Thus, the ordering of outcomes induced by the chi-squared statistic is the same as the ordering induced by the likelihood ratio. Consequently, extreme sets and  $p$ -values are identical for the likelihood ratio ordering and the chi-squared ordering. This equivalence is specific to this situation, in which there is only one observation from the second population. The approximation to the  $p$ -value, computed from the chi-squared distribution with  $c - 1$  d.f., appears to be very anticonservative, being consistently much smaller over a wide range of  $c$ ,  $n$ , and  $f_\ell$  than  $p$ -values computed as described in this section.

At  $\pi = \pi_*$ , the probability of  $C_{\text{LR}}(x)$  is

$$PV_{\text{LR}}(\pi; x) = \sum_{j=1}^c \pi_j \text{CDF}_{\text{bin}}(x; n, \pi_j).$$

We would like to find the maximum of this probability. Note that

$$PV_{\text{LR}}(\pi; x) = \sum_{j=1}^c G(\pi_j; x),$$

the sum of  $c$  copies of the same function  $G$  evaluated at different arguments  $\pi_j$ . The function  $G$  is well behaved, with a unique maximum (say at  $p_*$ ). The maximum of  $PV_{\text{LR}}$  is approximated very well by  $(c - 1)G(p_*; x)$ . If the approximation is less than or equal to 0.10 (the approximation is always less than the true value), the error is at most 0.0014 for  $c = 2$  and  $n = 4$ , and it is not greater than 0.00011 for larger  $c$  and  $n$ . Thus, if the approximation is in a range that indicates a significant difference at the 10% level or less, then the approximation is very accurate. This makes the computation of the maximum  $p$ -value relatively simple, requiring only finding the maximum of  $G$ .

This simplification to a one-dimensional optimization problem was used in the calculations leading to Table 3; however, most results were double-checked with a multivariate optimization algorithm and a grid search, and in every case the results agreed with the one-dimensional algorithm.

For  $c = 2, 3, 4, 8$ , Table 3 shows the least sample size  $n_\alpha(x)$  such that  $PV_{\text{LR}}(\pi, x)$  is everywhere (with respect to  $\pi$ ) less than or equal to  $\alpha$ , with  $\alpha$  equal to 0.10, 0.05, and 0.01. This is a concise compilation of calculations in which the maximum  $p$ -value was computed for each  $c, n, x$  combination, with  $c = 2, 3, 4, 8$ ,  $n = 4(1)500$ , and  $x = 0(1)50$ . For example, suppose that, for  $c = 2$  and  $n = 50$ , the frequency  $x$  in the training sample



for the category in which the mystery sample falls is  $x = 5$ . With  $n_{0.10}(5) = 33$  and  $n_{0.05}(5) = 64$ , these results show a significant difference between the training sample and the mystery sample at the 10% level of significance but not at the 5% level.

Table 3 gives some indication of the need for sample sizes ranging upward from 20 or so as well as a judicious choice of species to follow. The fact that one researcher has exposed more than 200 experimental pig carcasses over 6 years (G. S. Anderson, personal communication) shows that sample sizes of 20 or so are certainly practicable. With sample size in that range (say, 25) and following two species or groups of species (so that  $c = 4$ ), it is possible to find a significant difference at the 10% level if the mystery carcass has a species combination that occurs in at most one of the training carcasses. However, tracking three species, so that  $c = 8$ , it is not possible for the mystery carcass to differ significantly from the training carcasses at the 10% level of significance with a sample size of 25.

#### 4. THE FISHER EXACT TEST *p*-VALUES

The FET *p*-value is computed from probabilities of outcomes  $(\mathbf{f}, \ell)$ , conditioned on  $(\mathbf{f} + \mathbf{y}_*)$ , that are *c*-category hypergeometric probabilities. If we adopt the convention of defining the conditional extreme set for the outcome  $(\mathbf{f}, \ell)$  as the set of tables for which this conditional probability is less than or equal to the conditional probability of the outcome  $(\mathbf{f}, \ell)$ , then it may be shown that the FET conditional *p*-value is

$$\frac{S(\mathbf{f}, \ell)}{n + 1},$$

where

$$S(\mathbf{f}, \ell) = 1 + \sum_{f_j \leq f_{\ell+1}} f_j.$$

Its least value is  $1/(n + 1)$ , which is attained only if the mystery carcass exhibits a species-combination category that is not present on any of the training carcasses (hence,  $f_{\ell} = 0$ ) and no other category occurs among the training carcasses with frequency as low as one (so that all  $f_j > 1, j \neq \ell$ ). For example, with  $n = 25$ , it is possible to attain a *p*-value as low as  $1/26 = 3.85\%$ .

The FET conditional *p*-value is exact and it is invariant to  $\pi$ , conditional on the fixed marginal totals. Our next step is to find an unconditional probability based on FET. Following McDonald et al. (1977), extreme sets can be defined by the ordering of outcomes in terms of their FET conditional *p*-values, which is equivalent to ordering them in terms of  $S(\mathbf{f}, \ell)$ . In this ordering, the extreme set for  $(\mathbf{f}, \ell)$  is

$$C_{\text{FET}}(s) = \{(\mathbf{F}, j) : S(\mathbf{F}, j) \leq s\},$$

where  $s = S(\mathbf{f}, \ell)$ . For  $\pi = \pi_*$ , denote its probability by  $PV_{\text{FET}}(\pi; s)$ . Finding the maximum of  $PV_{\text{FET}}$  over  $\pi$  appears to involve burdensome computations, and we have been unable to find any clever shortcuts. In the computations shown in the next section, we performed these computations by listing all possible outcomes (with  $n = 10$  and  $c = 4$ ,

there are  $286 \times 4 = 1,144$ ), ordering them in terms of  $S(\mathbf{f}, \ell)$ , computing probabilities point by point, and maximizing this function with a multivariate nonlinear optimization algorithm. With larger sample size  $n$  and more categories  $c$ , performing the calculations this way becomes impractically burdensome, mainly because the number of outcomes becomes large rapidly with  $n$  and  $c$  (e.g., with  $c = 8$  and  $n = 20$ , there are 888,030 outcomes for  $(\mathbf{f}, \ell)$ ).

## 5. ILLUSTRATIONS AND COMPARISONS

The purpose of this section is to present some numerical results in order to give an impression of the relation between the exact  $p$ -values for the LR and FET extreme sets, on the one hand, and the chi-squared and conditional FET  $p$ -values on the other. Tables 4 and 5 illustrate that the conventional approximate  $p$ -values based on the chi-squared approximation and FET are poor. Table 4 compares exact LR  $p$ -values and their chi-squared probability approximation for  $c = 4$  and  $n = 10$  and 20. For moderately small exact  $p$ -values, the chi-squared probability approximation is much too small to be of any use.

Table 5 shows unconditional FET  $p$ -values and FET conditional  $p$ -values for  $c = 4$  and  $n = 10$  and 20. The FET conditional  $p$ -values are too conservative, being in some cases as much as twice the magnitude of the maximum attainable unconditional probability of the FET extreme set. Note also that the FET  $p$ -values in the small range, below about 0.15, are divided more finely than the LR  $p$ -values.

## 6. CONCLUSIONS AND CAVEATS

The comparison of the pattern of species present on a mystery carcass with experimental data on succession patterns can be represented as a contingency table. However, as we have illustrated here, the chi-squared approximation to the  $p$ -value is poor and the FET conditional probability is unnecessarily conservative. We developed the likelihood-ratio ordering of outcomes as a general approach to the problem in order to investigate a reasonable alternative to the unconditional FET  $p$ -value. The LR  $p$ -value, as it turns out, is simple to compute, it is an exact probability of a subset of the sample space, and its extreme set is the same as the extreme set for the chi-squared statistic. (It should be stressed that the equivalence with

Table 4. LR  $p$ -Values and  $\chi^2$  Approximations,  $c = 4$

$f\ell$	$n = 10$		$n = 20$	
	$p$ -Value	$\chi^2$ appr.	$p$ -Value	$\chi^2$ appr.
0	0.1051	0.0117	0.0538	0.0001
1	0.2479	0.1755	0.1241	0.0188
2	0.5256	0.4020	0.2045	0.0979
3	0.7759	0.5881	0.2925	0.2157
4	0.9219	0.7244	0.4148	0.3394

Table 5. FET  $p$ -Values and Conditional  $p$ -Values,  $c = 4$ 

$S(f, \ell)$	$n = 10$		$n = 20$	
	$p$ -Value	Conditional	$p$ -Value	Conditional
1	0.0431	0.0909	0.0223	0.0476
2	0.1091	0.1818	0.0567	0.0952
3	0.1698	0.2727	0.0873	0.1429
4	0.2349	0.3636	0.1215	0.1905
5	0.3770	0.4545	0.1570	0.2381
6	0.4241	0.5454	0.1943	0.2857
7	0.4881	0.6364	0.2301	0.3333
8			0.2692	0.3810
9			0.3086	0.4286
10			0.4119	0.4762

the chi-squared statistic does not generalize to tables in which both column totals are greater than one, even in  $2 \times 2$  tables.) Except for computations, the unconditional FET  $p$ -value shares these characteristics. Its computation is practicable for sample sizes  $n$  up to perhaps 50, but for  $n$  greater than about 20, it is cumbersome and time consuming. Even so, our limited experience suggests that the unconditional FET  $p$ -value is more sensitive than the LR  $p$ -value and so may be worth the extra computation. The fact that the LR  $p$ -value depends only on the training-sample frequency of the mystery carcass category is intuitively appealing. The dependence of the unconditional FET  $p$ -value on the training-sample frequencies in other categories is puzzling, but it may be this characteristic that makes it more sensitive.

The development here has been in terms of an idealized setting in order to derive the LR and FET  $p$ -values. We regard this as a first step that provides useful tools. There are obvious difficulties in practice due to the many factors that can affect the comparability of the training sample and the mystery carcass. Training samples commonly are observed repeatedly over several days, so composition data for different days cannot be regarded as independent. This does not seem to affect the validity of any single  $p$ -value, but it may have some effect on the joint interpretation of several  $p$ -values.

In the development in this paper, we considered species combinations to be the same or to be different, ignoring the extent of the difference. As one reviewer pointed out, though, the amount of difference may be relevant, too. For example, AB would seem to be less different from AB' than it is from A'B'. It may be possible to use this kind of structure so that the dimensionality of the problem is reduced, resulting in much fewer than  $2^s$  possible species combinations with which to deal.

[Received July 1997. Accepted September 1999.]

## REFERENCES

- Anderson, G. S., and VanLaerhoven, S. L. (1996), "Initial Studies on Insect Succession in Southwestern British Columbia," *Journal of Forensic Sciences*, 41, 617-625.

- Catts, E. P., and Goff, M. L. (1992), "Forensic Entomology in Criminal Investigations," *Annual Review of Entomology*, 37, 253–272.
- Goff, L. (1992), "Problems in Estimation of Postmortem Interval Resulting from Wrapping of the Corpse: A Case Study from Hawaii," *Journal of Agricultural Entomology*, 9, 237–243.
- Hennsge, C., Madea, B., Knight, B., Nokes, L., and Krompecher, T. (1991), *The Estimation of the Time Since Death in the Early Postmortem Period*, London: Arnold.
- Kempthorne, O., and Folks, L. (1971), *Probability, Statistics, and Data Analysis*, Ames, IA: Iowa State University Press.
- Lehmann, E. (1959), *Tests of Statistical Hypotheses*, New York: John Wiley.
- McDonald, L. L., Davis, B. M., and Milliken, G. A. (1977), "A Nonrandomized Unconditional Test for Comparing Two Proportions in  $2 \times 2$  Contingency Tables," *Technometrics*, 19, 145–157.
- Mehta, C. R., and Patel, N. R. (1983), "A Network Algorithm for Performing Fisher's Exact Test in  $r \times c$  Contingency Tables," *Journal of the American Statistical Association*, 78, 427–434.
- Schoenly, K. (1992), "A Statistical Analysis of Successional Patterns in Carrion–Arthropod Assemblages: Implications for Forensic Entomology and Determination of the Postmortem Interval," *Journal of Forensic Sciences*, 37, 1489–1513.
- Schoenly, K., and Reed, W. (1987), "Dynamics of Heterotrophic Succession in Carrion Arthropod Assemblages: Discrete Series or a Continuum of Change?" *Oecologia*, 73, 192–202.
- Smith, K. G. V. (1986), *A Manual of Forensic Entomology*, Ithaca, NY: Cornell University Press.
- Storer, B. E., and Kim, C. (1990), "Exact Properties of Some Exact Test Statistics for Comparing Two Binomial Proportions," *Journal of the American Statistical Association*, 85, 146–155.
- Suissa, S., and Shuster, J. J. (1985), "Exact Unconditional Sample Sizes for the  $2 \times 2$  Binomial Trial," *Journal of the Royal Statistical Society, Series A*, 148, 317–327.
- Wells, J. D., and LaMotte, L. R. (1995), "Estimating Maggot Age from Weight Using Inverse Prediction," *Journal of Forensic Sciences*, 40, 585–590.