APPENDIX B

Part I: JUSTIFICATION OF THE USE OF THE MEDIAN TEST

Assume that

$$X(t) = f(t) + U(t) \quad \text{for } t = t_{11}, t_{12}, \ldots, t_{1n}$$
$$Y(t) = g(t) + V(t) \quad \text{for } t = t_{21}, t_{22}, \ldots, t_{2k}$$

where

$$(U(t_{11}), \ldots, U(t_{1n}), V(t_{21}), \ldots, V(t_{2k}))$$

is a vector of independent random variables which are symmetrically distributed about zero (in particular, $E(U(t_{1i})) = E(V(t_{2j})) = 0, i = 1, 2, \ldots, n, j = 1, 2, \ldots, k$).

Let $a < t_{ij} < b, i = 1, 2, j = 1, 2, \ldots, n$ or k as the case may be. We wish to test the null hypothesis

$$H_0: f(t) = g(t) \text{ for } a \leq t \leq b$$

against the alternative

$$H_a: \text{either } f(t) < g(t) \text{ for } a < t < b, \text{ or } f(t) > g(t) \text{ for } a < t < b.$$

It is proposed to use the Median Test for this problem, and we will discuss the question whether this is justified. The test would be performed as follows: First, find a sample median m of the set of outcomes $x(t_{11}), \ldots, x(t_{1n}), y(t_{21}), \ldots, y(t_{2k})$ of $X(t_{11}), \ldots, X(t_{1n}), Y(t_{21}), \ldots, Y(t_{2k})$. Arrange a table as follows:

	Number of values $\leq$ m	Number of values $>$ m	Totals
'X – sample'	n_{1x}	n_{2x}	n
'Y – sample'	n_{1y}	n_{2y}	k
Totals	$n_{1.}$	$n_{2.}$	$n + k = n_{1.} + n_{2.}$

and test whether the ratios $\frac{n_{1x}}{n}$ and $\frac{n_{1y}}{k}$ differ significantly from the marginal ratio $\frac{n_{1.}}{n + k}$ by the usual chi-square or normal approximation (see e.g. Siegel, 1956).

It will now be shown that under certain conditions this is a reasonable test in the sense that it has the proper level of significance and that it is plausible that it will be sensitive with respect to the alternative hypothesis. Of course, in practice one cannot be sure that one of the conditions stated below is exactly true, but it seems rather unlikely that the null hypothesis will very often be rejected for the wrong reasons. Rather, we conjecture that the power of the test in general is not very large.

Assumption A. f and g are either both monotonically increasing functions or both

monotonically decreasing functions on [a, b]. We postulate that if there is a $t = t_0$ in (a, b) such that $f(t_0) = g(t_0)$, then $f \equiv g$ on [a, b].

Note: (a, b) denotes the open interval from a to b, i.e. the interval not containing the end points a and b themselves, and [a, b] denotes the closed interval from a to b, i.e. the interval containing a and b. The assumption excludes the possibility that the graphs of f and g intersect except perhaps at one of the end points a and b.

Assumption B. For any value of $t \in (a, b)$, U(t) and V(t) have the same probability distribution.

We will now discuss two different situations.

Case 1. $n = k$, and $t_{11} = t_{21}, t_{12} = t_{22}, \ldots, t_{1n} = t_{2n}$. That is, the X-es and Y's are to be measured at the same values of the independent variable t.

Let S_T denote the set $\{t_{11}, t_{12}, \ldots, t_{1n}\}$ and imagine the following chance experiment: an element T is selected at random from S_T, and we define

$$X^* = X(T), \quad Y^* = Y(T).$$

Proposition 1.1 If $f \equiv g$ on [a, b] and Assumption B holds, then X^* and Y^* have the same probability distribution, hence the same median.

Proof: For any number z, we have

$$\begin{aligned} P(X^* \leq z) &= \sum_{t \in S_T} P(X(t) \leq z \text{ and } T = t) \\ &= \sum_{t \in S_T} P(X(T) \leq z \mid T = t) P(T = t) \\ &= \sum_{t \in S_T} \frac{1}{n} P(f(t) + U(t) \leq z) \\ &= \frac{1}{n} \sum_{t \in S_T} P(U(t) \leq z - f(t)) \qquad (1) \end{aligned}$$

Similarly,

$$P(Y^* \leq z) = \frac{1}{n} \sum_{t \in S_T} P(V(t) \leq z - g(t)) \qquad (2)$$

Now if $f(t) = g(t)$ for all $t \in S_T$, and U(t) and V(t) have the same distribution, all individual terms in the sum (1) are equal to the corresponding terms in (2), therefore the sums are equal, and

$$P(X^* \leq z) = P(Y^* \leq z).$$

Since this is true for any z, X^* and Y^* have the same distribution function and therefore they have the same distribution, as stated.

Proposition 1.2 Suppose that Assemptions A and B hold and suppose further that for any t, the probability distribution of U(t) (hence also that of V(t)) is continuous, and that its probability density function has a unique maximum at zero. Then if X^* and Y^* have the same median, $f \equiv g$.

Proof: Let μ be the common median of X^* and Y^*. This means that

$$P(X^* < \mu) \leq 0.5 \text{ and } P(X^* > \mu) \leq 0.5$$

From the computation in the proof of Proposition 1.1 it will be clear that

$$P(X^* = \mu) = \frac{1}{n} \sum_{t \in S_T} P(U(t) = \mu - f(t))$$

but as U(t) is continuously distributed, all probabilities $P(U(t) = \mu - f(t))$ are zero. Therefore $P(X^* = \mu) = 0$, and it follows that

$$P(X^* < \mu) = 0.5$$

or

$$\frac{1}{n} \sum_{t \in S_T} P(U(t) < \mu - f(t)) = 0.5 \tag{3}$$

Similarly, we may show that

$$\frac{1}{n} \sum_{t \in S_T} P(V(t) < \mu - g(t)) = 0.5 \tag{4}$$

By Assumption A, we only have three possibilities: either $f \equiv g$, or $f < g$, or $f > g$. But assuming that one of the last two possibilities holds leads into a contradiction; for instance, if $f < g$, we have $\mu - f(t) > \mu - g(t)$ for all t, therefore $P(U(t) \leq \mu - f(t)) > P(V(t) \leq \mu - g(t))$ for all t, hence the sum (3) is greater than the sum (4), in contradiction to the fact, already established, that they are equal.

Case 2. $(t_{11}, \ldots, t_{1n})$ and $(t_{21}, \ldots, t_{2k})$ are independent random samples from the same universe, with probability distribution P_T and distribution function F.

Imagine two types of chance experiment as follows: 1° a value t is sampled from this universe. Let T be the random variable which takes the value t in such an experiment, and define $X^* = X(T)$.

2° the same, but now define $Y^* = Y(T)$. Then $X(t_{11}), \ldots, X(t_{1n})$ all have the same distribution as X^*, and $Y(t_{21}), \ldots, Y(t_{2k})$ all have the same distribution as Y^*.

Proposition 2.1 If Assumption B holds and if $f \equiv g$, then X^* and Y^* have the same probability distribution, hence the same median.

Proof: For any number z, we have

$$P(X^* \leq z) = \int_a^b P(X(T) \leq z \mid T = t) dF(t)$$

$$= \int_a^b P(U(t) \leq z - f(t)) dF(t) \tag{5}$$

and also

$$P(Y^* \leq z) = \int_a^b P(V(t) \leq z - g(t)) dF(t) \tag{6}$$

By Assumption B, $P(V(t) \leq z - g(t)) = P(U(t) \leq z - g(t))$ for all t. Because $f \equiv g$, this is equal to $P(U(t) \leq z - f(t))$ and it follows that $P(X^* \leq z) = P(Y^* \leq z)$ for all z, hence X^* and Y^* have the same distribution function.

Proposition 2.2 If Assumptions A and B hold, and if for any t the probability distributions of U(t) and V(t) are continuous, their probability density functions having a unique maximum at zero, then if X* and Y* have the same median, then $f \equiv g$.

Proof: Let μ be the common median of X* and Y*. By an argument similar to that used in the proof of Proposition 1.2

$$\int_a^b P(U(t) \leq \mu - f(t))dF(t) = \int_a^b P(V(t) \leq \mu - g(t))dF(t) = 0.5$$

and again, if for example we have $f(t) < g(t)$ for all t, then

$$P(U(t) \leq \mu - f(t)) > P(V(t) \leq \mu - g(t)) \text{ for all } t$$

and it follows that the above equality can be true only if $f(t) = g(t)$ for all t.

We now replace Assumptions A and B with different assumptions; this leads us into a third case to be considered.

Case 3. The following assumptions apply:

Assumption C. The regression curves f and g are represented by straight lines:

$$f(t) = p + qt \qquad g(t) = r + st$$

where p, q, r, and s are constants. To simplify the algebra, we will assume that $p + qa > 0, r + sa > 0, q > 0, s > 0$.

Assumption D. The probability distributions of U(t) and V(t) do not depend on t, but need not be the same for U(t) and V(t). The distributions are continuous, or at least $P(U(t) = 0) = P(V(t) = 0)$ for all t,

Let

$$t_m = (a + b)/2, f_m = [f(a) + f(b)]/2, g_m = [g(a) + g(b)]/2$$

Assumption E. $\{t_{11}, \ldots, t_{1n}\}$ and $\{t_{21}, \ldots, t_{2k}\}$ are independent random samples from the same universe. The probability distribution concerned is continuous and the probability density function is symmetrical about t_m.

Imagine the chance experiments described for Case 2.

Lemma. If Assumptions C, D, and E hold, then f_m is the median of X* and g_m is the median for Y*.

Proof: We need only prove that f_m is the median of X*. We have

$$\begin{aligned} P(X^* \leq f_m) &= \int_a^b P(X(T) \leq f_m \mid T = t)dF(t) \\ &= \int_a^{t_m} P(X(T) \leq f_m \mid T = t)dF(t) + \\ &+ \int_{t_m}^b P(X(T) \leq f_m \mid T = t)dF(t) \\ &= \int_a^{t_m} P(U(t) \leq f_m - f(t))dF(t) \\ &+ \int_{t_m}^b P(U(t) \leq f_m - f(t))dF(t) \end{aligned}$$

Consider the first of these two integrals. We have assumed that $q > 0$, hence f is increasing; therefore $f_m - f(t) \geq 0$ for $a \leq t \leq t_m$, and we may write

$$\int_a^{t_m} [P(U(t) \leq 0) + P(0 < U(t) \leq f_m - f(t))]dF(t)$$

$$= \int_a^{t_m} P(U(t) \leq 0)dF(t) + \int_a^{t_m} P(0 < U(t) \leq f_m - f(t))dF(t)$$

For any t, $P(U(t) \leq 0) = \frac{1}{2}$; furthermore, $\int_a^{t_m} dF(t) = \frac{1}{2}$, and so we obtain for this first integral

$$\tfrac{1}{4} + \int_a^{t_m} P(0 < U(t) \leq f_m - f(t))F'(t)dt$$

Now we consider the second integral. As $f_m - f(t) \leq 0$ for $t_m < t \leq b$, we may write

$$P(U(t) \leq f_m - f(t)) = 1 - P(U(t) > f_m - f(t))$$
$$= 1 - [P(U(t) \geq 0) + P(0 > U(t) \geq f_m - f(t))]$$
$$= 1 - \tfrac{1}{2} - P(0 > U(t) \geq f_m - f(t))$$

and

$$\int_{t_m}^{b} P(U(t) \leq f_m - f(t))dF(t) =$$

$$= \tfrac{1}{2} - \tfrac{1}{4} - \int_{t_m}^{b} P(0 > U(t) \geq f_m - f(t))F'(t)dt$$

Combining these results, we have

$$P(X^* \leq f_m) = \tfrac{1}{2} + \int_a^{t_m} P(0 < U(t) \leq f_m - p - qt)F'(t)dt$$

$$- \int_{t_m}^{b} P(0 > U(t) \geq f_m - p - qt)F'(t)dt$$

In the first integral, we introduce the new variable $z = t_m - t$, and in the second integral we introduce the new variable $z = t - t_m$. Noting that

$$f_m = [f(a) + f(b)]/2 = (p + qa + p + qb)/2$$
$$= p + q(a + b)/2 = p + qt_m$$

we obtain

$$P(X^* \leq f_m) = \tfrac{1}{2} + \int_0^{t_m - a} P(0 < U(t_m - z) \leq qz)F'(t_m - z)dz$$

$$- \int_0^{b - t_m} P(0 > U(t_m + z) \geq -qz)F'(t_m + z)dz$$

As the distribution of U(t) does not depend on t and is symmetrical about zero, we

have $P(0 < U(.) \leq qz) = P(0 > U(.) \geq -qz)$. As the probability density function of t is symmetrical about t_m, we have $F'(t_m - z) = F'(t_m + z)$. It then follows that

$$P(X^* \leq f_m) = \tfrac{1}{2}$$

as stated.

As an immediate corollary of this lemma, we have

Proposition 3.1 If Assumptions C, D, and E hold, then X* and Y* have the same median if, and only if, $f \equiv g$.

We conclude this section with a brief discussion on the interpretation of the results obtained. With the median test, the 'X' values and the 'Y' values are pooled as though they did not depend on t. This was formalized in the above argument by imagining a sampling experiment resulting in random variables X* and Y* such that the 'X' values are values of X* and the 'Y'-values are values of Y*. Now if the two regression functions f and g are the same, there may be two reasons why X* and Y* could have different medians. In the first place, the t's on which X(t) and Y(t) depend might be differently distributed over the interval (a, b). Diagram 1 shows how this may lead to wrongly rejecting the null hypothesis. The black and the white points are scattered about the same regression line (so the null hypothesis should not be rejected), but there are more white points in the left half of the interval and more black ones in the right half. Because of this, there are more black points above the common median than there are white ones, and the difference would be significant according to the Median Test. In the second place, the distributions of the residuals U(.) and V(.) might differ in some peculiar fashion. An example is given by Diagram 2. Black and white points are scattered about the same regression line, hence the null hypothesis should not be rejected. But apparently the variance of the residuals of the black points decreases with t, whereas the variance of the residuals of the white points increases with t. The reader may verify that the Median Test would reject the null hypothesis of a common median.

We must therefore introduce certain assumptions about the distribution from

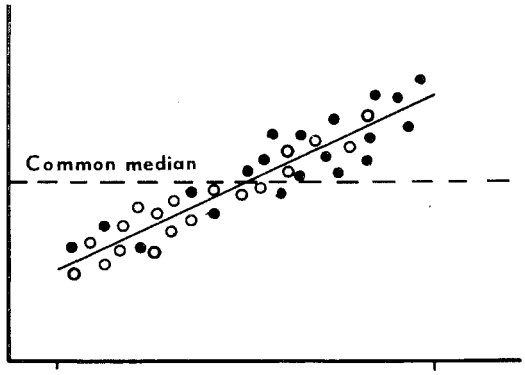


Diagram 1

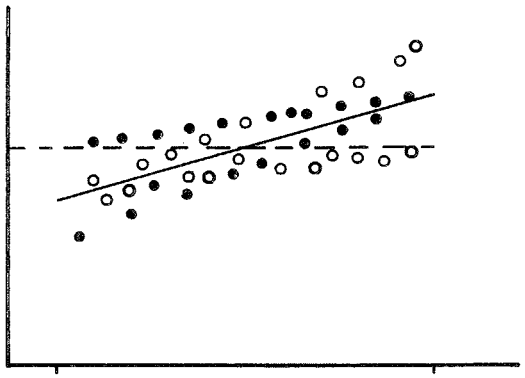


Diagram 2

which the t's are sampled as well as the distributions of the residuals. In the above discussion, some examples of such assumptions are given and we showed that under these assumptions, X* and Y* will have the same median if $f \equiv g$, and so if the Median Test leads to rejection of the null hypothesis, this implies that the hypothesis $f \equiv g$ has been rejected. The other side of the matter is: if f and g are not the same function, is it likely that the null hypothesis will be rejected, given sufficient data? This will not be so in cases where f and/or g are not monotonic, or if their graphs intersect each other. See for example Diagram 3, where the null hypothesis will not be rejected although f and g are different functions. However, in the present study such alternatives are not very interesting, and so it need not worry us too much if the power of the test with respect to them is low. On the other hand, simulation experiments which we performed gave the impression that one has to think up probability distributions depending in a very weird way on t in order to construct counterexamples where the application of the test would lead to wrongly rejecting the null hypothesis. Distributions as shown in Diagram 2 will not often be encountered in practice. Situations as shown in Diagram 1 may be avoided by inspection of the data and by choosing t-intervals which are not too wide.

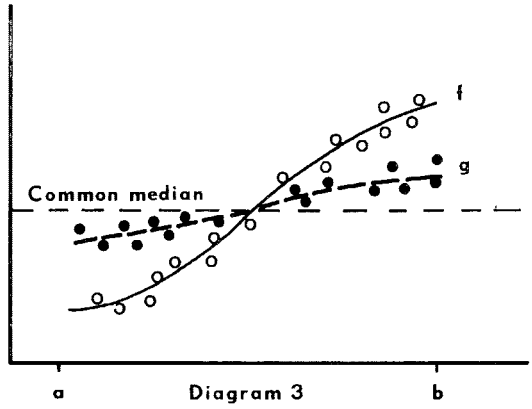


Diagram 3

Part II: THE MATHEMATICAL PROPERTIES OF DPS

0. We will discuss the mathematical properties of the quantity

$$DPS = 1 - \frac{\sum_{i=1}^{j} \ln (n_i + 1)}{j \ln (N + j)} \tag{1}$$

where

N = total number of individuals caught;

j = number of year-samples in which at least one individual was present;

n_i = number of individuals present in the i-th year-sample.

and

$$\sum_{i=1}^{j} n_i = N. \tag{2}$$

1. According to 5.3 items (1) and (2) it is required to show that

   (i) 'the higher the number of year-samples with low catches as compared to the number of year-samples with high catches, the higher DPS';

   (ii) conversely, 'the higher the number of year-samples with high catches as compared to the number of year-samples with low catches, the lower DPS',

in short: 'DPS is the lower the more evenly the numbers are distributed over the year-samples'.

In general, the numbers may be distributed over the year-samples in a great many ways, which cannot all be ordered on a simple scale from 'very unevenly' to 'very evenly'. However, we will show that DPS attains a unique minimum if all the n_i's are equal and a maximum if all but one of the n_i's are 1. It is easily seen from the definition that the graph of DPS as a function of the n_i should be a 'smooth' surface with certain symmetry properties (See Fig. on p. 181). Let us write

$$\theta_i = n_i/N, \tag{3}$$

then

$$\sum_{i=1}^{j} \theta_i = 1 \tag{4}$$

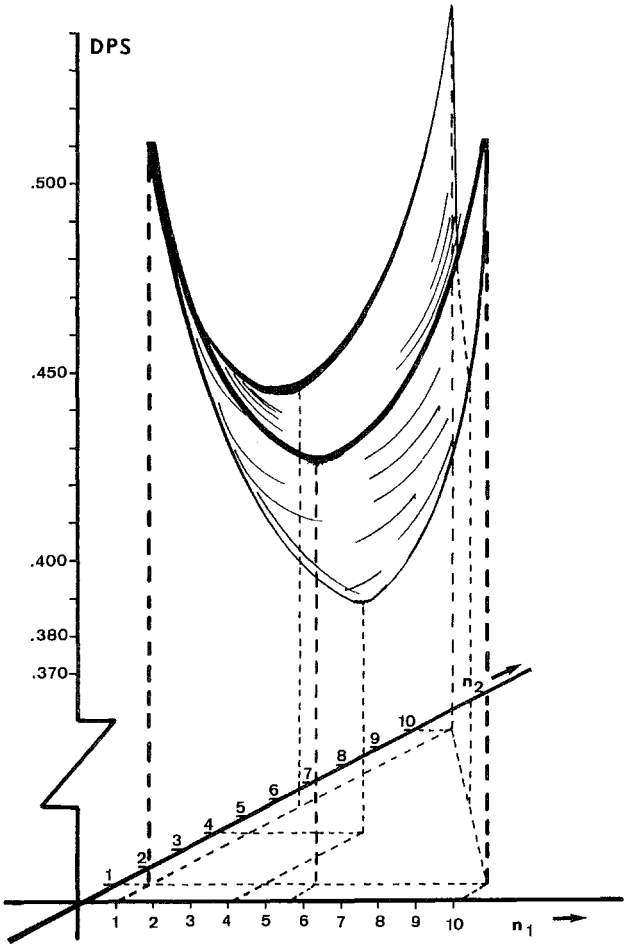
Let Θ denote the vector of the θ_i's, and let us consider the function

$$f(\Theta;N,j) = \sum_{i=1}^{j} \ln (\theta_i N + 1) \tag{5}$$

for j and N fixed.

Because of (4), only $j - 1$ of the θ_i's might be 'chosen freely'. To express this restriction we will write

$$\theta_j = 1 - \sum_{i=1}^{j-1} \theta_i \tag{6}$$



Graph of DPS as a function of n_1 and n_2 for j = 3, N = 12. Note that DPS is at its minimum when $n_1 = n_2 (= n_3) = 12/3 = 4$; given e.g. that $n_2 = 1$ DPS is at its minimum when $n_1 = n_3 = 11/2 = 5.5$ and so on; maxima are where one of the n_i is 10 and the others are 1.

In what follows, we will need the partial derivatives of f with respect to the θ_i's. Substituting (6) into (5) we obtain

$$f(\Theta) = \sum_{i=1}^{j-1} \ln(\theta_i N + 1) + \ln\left(N + 1 - N \sum_{i=1}^{j-1} \theta_i\right)$$

Then for $k = 1, 2, \ldots, j-1$ we have

$$\frac{\delta f}{\delta \theta_k} = \frac{N}{\theta_k N + 1} - \frac{N}{\left(1 - \sum_{i=1}^{j-1} \theta_i\right) N + 1} \tag{7a}$$

$$\frac{\delta f}{\delta \theta_k} = \frac{N}{\theta_k N + 1} - \frac{N}{\theta_j N + 1} \tag{7}$$

where, to simplify the expression, we have resubstituted θ_j from (6).

Lemma 1. $f(\Theta; N, j)$ for given N and j attains a uniq..e maximum for $\theta_1 = \theta_2 = \ldots = \theta_{j-1} = \theta_j = j^{-1}$. The value of this maximum is $f_{max} = j[\ln(N + j) - \ln(j)]$.

Proof. A necessary condition for f to attain a maximum within the region defined by the inequalities

$$0 < \theta_i < 1, \; i = 1, 2, \ldots, j$$

on which f is differentiable is, that θ be the solution of the equations

$\frac{\delta f}{\delta \theta_i} = 0$, 1, 2, ..., j, i.e., from (7):

$$\frac{N}{\theta_i N + 1} - \frac{N}{\theta_j N + 1} = 0, \text{ or } \theta_i = \theta_j, \; i = 1, 2, \ldots, j-1.$$

The solution, obviously, is $\theta_1 = \theta_2 = \ldots = \theta_j = j^{-1}$.

Now consider the function

$$\delta(\Theta) = \begin{vmatrix} \frac{\delta^2 f}{\delta \theta_1^2} & \frac{\delta^2 f}{\delta \theta_1 \delta \theta_2} & \cdots & \frac{\delta^2 f}{\delta \theta_1 \delta \theta_{j-1}} \\ \frac{\delta^2 f}{\delta \theta_1 \delta \theta_2} & \frac{\delta^2 f}{\delta \theta_2^2} & \cdots & \frac{\delta^2 f}{\delta \theta_2 \delta \theta_{j-1}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\delta^2 f}{\delta \theta_1 \delta \theta_{j-1}} & \frac{\delta^2 f}{\delta \theta_2 \delta \theta_{j-1}} & \cdots & \frac{\delta^2 f}{\delta \theta_{j-1}^2} \end{vmatrix}$$

From (7a) and (6) we have

$$\frac{\delta^2 f}{\delta \theta_k^2} = -\frac{N^2}{(\theta_k N + 1)^2} - \frac{N^2}{(\theta_j N + 1)^2}, \; k = 1, 2, \ldots, j-1$$

$$\frac{\delta^2 f}{\delta \theta_k \delta \theta_m} = -\frac{N^2}{(\theta_j N + 1)^2}, \text{ for } m = 1, 2, \ldots, j-1; \; m \neq k$$

Let $\Theta_0 = \left(\frac{1}{j}, \frac{1}{j}, \ldots \frac{1}{j}\right)$. Then for $\Theta = \Theta_0$:

$$\frac{\delta^2 f}{\delta \theta_k^2} = -\frac{2N^2 j^2}{(N + j)^2} \text{ and } \frac{\delta^2 f}{\delta \theta_k \delta \theta_m} = -\frac{N^2 j^2}{(N + j)^2}$$

From this it may derived by evaluating the determinant that

$$\Delta = \delta(\Theta_0) = (-1)^{3j-5}\, j. \left(\frac{N}{N+j}\right)^{2(j-1)} \neq 0.$$

We now put $\Delta_0 = 1$ and let Δ_{j-1-k} be the determinant obtained by deleting the last k rows and columns of the determinant defining Δ. Then

$$\Delta_{j-1-k} = (-1)^{3(j-k)-5}\,(j-k) \left(\frac{N}{N+j}\right)^{2(j-k-1)}$$

These determinants (for various k) are positive or negative according as to whether $j-k$ is odd or even. The facts shown suffice to prove that f attains a local maximum at $\Theta = \Theta_0$, according to a theorem of Analysis (e.g. stated and proved as Theorem 7–9, p. 151–152 in Apostol (1957)). This maximum is unique since there is no other value of Θ where all partial derivatives of f vanish. The value of the maximum is found by substitution. End of Proof.

Lemma 2. $f(\Theta; N, j)$ for given N and j attains local minima at all points defined by

$$\theta_i = \frac{N-j+1}{N} \text{ for one i, } \theta_k = \frac{1}{N} \text{ for } k \neq i \qquad i = 1, 2, \ldots, j,$$

if Θ is restricted to the region defined by

$$\frac{1}{N} \leq \theta_i \leq 1,\ i = 1, 2, \ldots, j, \text{ and provided that } N > j.$$

The value of all these minima is $f_{min} = (j-1) \ln 2 + \ln (N - j + 2)$.

Proof. It suffices to show that f attains a local minimum at

$$\theta_1 = \theta_2 = \ldots = \theta_{j-1} = \frac{1}{N},\ \theta_j = \frac{N-j+1}{N}$$

as all the other cases will follow in the same fashion. From (7) we have at this point

$$\frac{\delta f}{\delta \theta_k} = \frac{N}{2} - \frac{N}{N-j+2} = \frac{N(N-j)}{2(N-j+2)} > 0 \text{ for } k = 1, 2, \ldots, j-1.$$

This shows that f increases in all directions on a neighbourhood of this point inside the region considered, which proves our point. The value of f_{min} is found by substitution. End of Proof.

As an immediate corollary to these two lemmas we have the

Theorem 1. For N and j fixed, $N > j$, DPS attains a unique minimum for $n_1 = n_2 = \ldots = n_j = N/j$. The value of this minimum is $m = \ln(j)/\ln(N + j)$. Moreover DPS attains local maxima whenever $n_i = N - j + 1$ for one i, $n_k = 1$ for $k \neq i$. The value of these maxima is $M = 1 - [(j - 1) \ln 2 + \ln (N - j + 2)]/j \ln (N + j)$.

Proof. We have

$$\text{DPS} = 1 - \frac{f}{j \ln (N + j)}$$

therefore DPS attains a minimum where f attains a maximum and vice versa.

$$\text{From } m = 1 - \frac{f_{max}}{j \ln (N+j)},\ M = 1 - \frac{f_{min}}{j \ln (N+j)}$$

M and m are easily found by substitution. End of Proof.

2. According to 5.3 item (3), 'starting from an already high level of the total number of individuals caught (N), the addition of some year-samples with low catches should increase the value of DPS'.

Obviously it suffices to show that DPS increases if we add just one year-sample with a low catch. In order to prove anything, we have to specify 'an already high level' and a 'low catch'. This could perhaps be done in several ways; as an example, suppose we add a $j + 1$-st year with a catch n_{j+1} such that

$$\ln (n_{j+1} + 1) < \frac{1}{j} \sum_{i=1}^{j} \ln (n_i + 1)$$

i.e., one plus the new catch is lower than the geometric mean of all previous samples, each increased by 1. Let us write

$$A = 1 - \frac{\sum_{i=1}^{j} \ln (n_i + 1)}{j \ln (N + j)}, \quad B = 1 - \frac{\sum_{i=1}^{j} \ln (n_i + 1) + \ln (n_{j+1} + 1)}{(j + 1) \ln (N + n_{j+1} + j + 1)}$$

Then $B > A$ if, and only if,

$$\{j \ln (N + j)\} \sum_{i=1}^{j} \ln (n_i + 1) + \{j \ln (N + j)\} \ln (n_{j+1} + 1) <$$

$$\{(j + 1) \ln (N + n_{j+1} + j + 1)\} \sum_{i=1}^{j} \ln (n_i + 1) =$$

$$= j \ln (N + n_{j+1} + j + 1) \sum_{i=1}^{j} \ln (n_i + 1) +$$

$$+ \{\ln (N + n_{j+1} + j + 1)\} \sum_{i=1}^{j} \ln (n_i + 1) \qquad (8)$$

This will certainly be satisfied if

$$\{j \ln (N + j)\} \ln (n_{j+1} + 1) < \ln (N + n_{j+1} + j + 1) \sum_{i=1}^{j} \ln (n_i + 1)$$

and, since $N + j < N + n_{j+1} + j + 1$, this is certainly satisfied if

$$\ln (n_{j+1} + 1) < \frac{1}{j} \sum_{i=1}^{j} \ln (n_i + 1)$$

as stated. Note that this condition is sufficient, but not necessary; it should be possible to derive a less stringent condition on n_{j+1}, but this condition would look a bit more artificial.

3. It is furthermore stated in 5.3 item (3) that 'with high values of j, DPS should not necessarily be high, i.e. especially in very numerous and eurotopic species it should be possible that DPS (RPR) reaches a relatively low value – depending on the proportion of the year-samples that show relatively low catches'.

There is not much to prove about a statement such as this. As will be shown below, DPS in general will have a tendency to increase with j, but of course its actual value depends very much on the $\theta_i = n_i/N$. Let us look at the minimum value of DPS:

$$m = \ln(j)/\ln(N + j).$$

We may note that even with a 'large' j this quantity may be close to zero, because for any fixed j we may choose N so large that m is smaller than any preassigned number. On the other hand, the maximal value of DPS is

$$M = 1 - \frac{(j-1)\ln 2}{j \ln(N+j)} - \frac{\ln(N-j+2)}{j \ln(N+j)}$$

The value of this expression varies between $1 - \ln(2)/\ln(2j)$ (attained for $N = j$) and $1 - 1/j$ (limit for $N \to \infty$) and with j small, M will not be close to 1.

4. If the number of year-samples (j) is varied, the θ_i we introduced in (3) will vary also because by definition j is the number of year-samples with non-zero catches. Therefore the influence of j cannot be studied independently of the θ_i. However, we may study how the maximum and minimum values of DPS vary with N and j. Let us therefore first study the function

$$g(\Theta, N, j) = 1 - DPS = \frac{\sum_{i=1}^{j} \ln(\theta_i N + 1)}{j \ln(N + j)}$$

Its minimum for any given N and j will have the value

$$g_{min} = \frac{(j-1)\ln(2) + \ln(N-j+2)}{j \ln(N+j)} = h(j, N) \text{ (say).}$$

We have

$$h(j, N) = \frac{(j-1)\ln 2}{j \ln(N+j)} + \frac{\ln(N-j+2)}{j \ln(N+j)}$$

For N fixed, both terms are decreasing functions of j, so h(j, N) for fixed N is a decreasing function of j.

By its definition, j may vary between two extremes: $j = 1$ (all individuals were caught in just one sample) and $j = N$ (all samples contain just one individual).

We have

$$h(1, N) = 1 \text{ and } h(N, N) = \frac{\ln(2)}{\ln(2N)}$$

and so, as j increases from 1 to N, h decreases from 1 to $\ln(2)/\ln(2N)$.

If N is large, there is not much variation left because then $\ln(N)/\ln(2N)$ will differ very little from 1.

Let us now suppose that N and j increase at the same rate; that is, let us assume $N = cj$, c constant (the average number of individuals per year-sample, c, remains constant). For this case we write g_{min} as

$$h_c(j) = \frac{(j-1)\ln(2) + \ln[(c-1)j + 2)]}{j \ln[(c+1)j]}$$

This is again a decreasing function of j, with maximum value $h_c(1) = 1$ and

$$\lim_{j\to\infty} h_c(j) = 0.$$

5. Let us now consider the maximum value of g, viz.

$$g_{max} = H(j, N) \text{ (say)} = 1 - \frac{\ln(j)}{\ln(N+j)}$$

$$= \frac{\ln\left(\frac{N+j}{j}\right)}{\ln(N+j)} = \frac{\ln(1 + N/j)}{\ln(N+j)}$$

For N fixed, this is a decreasing fuction of j. Its maximum is $H(1, N) = 1$; its minimum $H(N, N) = \ln(2)/\ln(2N)$.

These findings turn out to be exactly the same as those we had for g_{min} in our Section 4. Again, let us consider the case $N = cj$ with c fixed; then

$$g_{min} = H_c(j) = \frac{\ln(c+1)}{\ln[(c+1)j]}$$

This is once again a decreasing function of j, with maximum $H_c(1) = 1$, and

$$\lim_{j\to\infty} H_c(j) = 0.$$

6. The findings of the above two sections may be interpreted and summed up in the following

Theorem 2.

(i) The minimum value DPS may have is 0; it is attained when $j = 1$, i.e. when all individuals are caught in one sample.

(ii) For any given N, the maximum value DPS may have is $\ln(N)/\ln(2N)$ which is attained when $j = N$, i.e., when all of the samples contain one and only one individual. Hence DPS always is less than 1, with the maximum approaching 1 as N and j increase indefinitely.

(iii) The maximum value DPS may take on for fixed N is an increasing function of j, increasing from zero ($j = 1$) to $\ln(N)/\ln(2N)$ ($j = N$). If the ratio of N to j is kept fixed ($N/j \equiv c$), the maximum value of DPS is also an increasing function of j (hence also of N) increasing from 0 tot 1, N and $j \to \infty$.

(iv) The same statements are true with respect to the minimum value of DPS, when j is decreasing from very large values to 1.

These results corroborate what we stated in our Section 3: that DPS in general will have a tendency to increase with j (see: 5.4.4).

7. Finally, we derive a result which gives an indication of the role of N. Suppose we keep $\theta_1, \theta_2, \ldots, \theta_j$ and j fixed and let N increase. For the time being, we replace N with a continuous variable x. That is, we consider the function

$$r(x) = 1 - \frac{\sum_{i=1}^{j} \ln(\theta_i x + 1)}{j \ln(x + j)}$$

We will show that r(x) is a decreasing function of x. Its derivative is

$$r'(x) = \frac{-j \ln(x + j) \sum_{i=1}^{j} \frac{\theta_i}{\theta_i x + 1} + \frac{j}{x + j} \sum_{i=1}^{j} \ln(\theta_i x + 1)}{j^2 [\ln(x + j)]^2}$$

$$= \frac{-\ln(x + j) \sum_{i=1}^{j} \frac{\theta_i}{\theta_i x + 1} + \frac{1}{x + j} \sum_{i=1}^{j} \ln(\theta_i x + 1)}{j[\ln(x + j)]^2}$$

We need only consider the numerator of this expression. By Lemma 1

$$-\ln(x + j) \Sigma \frac{\theta_i}{\theta_i x + 1} + \frac{1}{x + j} \Sigma \ln(\theta_i x + 1) \leq -\ln(x + j) \Sigma \frac{\theta_i}{\theta_i x + 1} +$$

$$\frac{j}{x + j} [\ln(x + j) - \ln(j)] = \left[\frac{j}{x + j} - \Sigma \frac{\theta_i}{\theta_i x + 1}\right] \ln(x + j) - \frac{j \ln(j)}{x + j}$$

We may write

$$\frac{j}{x + j} = \sum_{i=1}^{j} \frac{1}{x + j}, \text{ whence}$$

$$\frac{j}{x + j} - \Sigma \frac{\theta_i}{\theta_i x + 1} = \Sigma \frac{1 - \theta_i j}{(x + j)(\theta_i x + 1)}$$

Now let the smallest among the θ_i have the value δ; then $\theta_i \geq \delta$, $i = 1, 2, \ldots, j$ and

$$\Sigma \frac{1 - \theta_i j}{(x + j)(\theta_i x + 1)} < \Sigma \frac{1 - \theta_i j}{(x + j)(\delta x + j)} = \frac{1}{(x + j)(\delta x + j)} \Sigma (1 - \theta_i j)$$

But

$$\Sigma (1 - \theta_i j) = j - j \Sigma \theta_i = j - j = 0.$$

Therefore

$$\Sigma \left[\frac{j}{x + j} - \Sigma \frac{\theta_i}{\theta_i x + 1}\right] \ln(x + j) \leq 0.$$

Since moreover

$$-\frac{j \ln(j)}{x + j} < 0,$$

we have shown that

$$r'(x) < 0$$

which means that r is a decreasing function of x.

Writing

$$r(x) = \frac{j \ln (x + j) - \Sigma \ln (\theta_i x + 1)}{j \ln (x + j)}$$

and substituting

$$\ln (x + j) = \ln(x) + \ln (1 + j/x)$$
$$\ln (\theta_i x + j) = \ln(x) + \ln (\theta_i + 1/x)$$

we obtain

$$r(x) = \frac{j \ln (x + j/x) - \Sigma \ln (\theta_i + 1/x)}{j \ln(x) + \ln (1 + j/x)}$$

As $x \to \infty$, the denominator increases beyond all bounds, whereas the numerator approaches a finite limit. Therefore we see that

$$\lim_{x \to \infty} r(x) = 0.$$

In conclusion, we have

Theorem 3. For a fixed number of year-samples (j) and a given distribution sequence $\{\theta_i, i = 1, 2, \ldots, j\}$, DPS is the lower the higher the total number N of individuals caught. In the limit ($N \to \infty$) DPS becomes zero.

This result makes good sense from the point of view of the theory given in the paper. If two species would have been caught in a same number of year-samples and would have the same proportional distribution over these samples, the species with the higher total number of individuals must have the poorer power of dispersal etc., for otherwise it would have been caught in more samples.

References

Apostol, T. M., 1957: Mathematical Analysis. A Modern Approach to Advanced Calculus. Reading, Mass: Addison-Wesley, xii + 553 pp.

Siegel, S., 1956: Nonparametric Statistics for the Behavioral Sciences, New York: Mac Graw Hill.

Appendix B by:
J. REDDINGIUS
Zoological Laboratory
University of Groningen
Kerklaan 30
Haren
Groningen, Netherlands

GLOSSARY

RPR: Realization of Population Replacement, i.e. the degree to which (within a given area) the rate of extinction of populations is currently compensated by the rate of (re)founding.

DPS: frequency Distribution of observed Population Sizes. For comparative purposes the cumulated distributions are used (Fig. 3). From chapter 6 onwards DPS is characterized by the expression $1 - \frac{\Sigma \ln (n_i + 1)}{j. \ln (N + j)}$ (cf. Appendix B.II), in which:

n_i: number of individuals in year-sample i, i.e. caught during one year in one standard set of pitfalls,

j: number of year-samples (out of 175) in which the pertinent species is represented by at least one specimen, and

N: total number of specimens of one species caught in all (175) year-samples together.

f: maximal number of potential flyers, i.e. the maximal number of individuals (from N) that could have been capable of flying during some period of the life-cycle; cf. Table 3.

w: number of individuals actually caught in window-traps, i.e. caught during the act of flying.

100w/f: rough estimate of the powers of dispersal (by flight).

A-species: species with a low dispersal power, i.e. no or only very few individuals (less than 70 out of N ind., which at the same time is less than 1% of N ind.) are expected sometimes to be capable of flying (f), although up till now no individuals are caught in window-traps, cf. 4.2.2.

B-species: species with a high dispersal power, i.e. at least one individual is caught in a window-trap, and at the same time is $f \geq 70$ and – with only one exception : 3.7% (species 133, Table 3) – between 17 and 100% of N ind.; cf. 4.2.1.

C-species: species of which no individuals are caught in window-traps, but which cannot unambigiously be considered A-species: the dispersal power is uncertain; cf. 4.2.3.

D-species: species mainly occupying woodlike habitats (cf. Table 4).

E-species: species mainly inhabiting areas of blown sand that are partly fixed by vegetation (cf. Table 4).

F-species: Species mainly living in heathlike areas and/or in old peat moor (cf. Table 4).

G-species: species mainly occupying temporary habitats, viz. places in natural localities that are readily inundated or the structure of which is significantly changing in the course of a few – less than 10 – years (in most cases after human intervention): cf. Table 4 and 3.7.

H-species: eurytopic species, i.e. species that inhabit – without much preference – localities belonging to more than one habitat group (cf. Table 1), among which some kind of temporary habitat is generally represented (cf. Table 4 and 4.3).

Permanent habitats: Localities of the types D, E and F; cf. Table 1 and 3.7.

Temporary habitats: Localities of the G-type; cf. Table 1 and 3.7.

Hypothesis 3.2: 'Overflow' hypothesis of dispersal, or more generally 'escape' hypothesis.

Hypothesis 3.4: 'Founding' hypothesis of dispersal.

% macr. ind.: the percentage of the individuals that is macropterous in wing di(poly)-morphic species; cf. Table 3.

Z: deviation from the mean of the standardized normal distribution measured in standard deviation units.

U-test: the nonparametric Mann-Whitney U-test or Wilcoxon two-sample test for the comparison of the values of two independent samples (cf. 5.2.1).

Med.-test: Median test as used according to a proposal by REDDINGIUS (cf. 5.2.3 and Appendix B.I) for the comparison of two bivariate samples.

τ: coefficient of rank correlation (cf. note to 4.4.1).