STATISTICAL TESTS AND MEASURES FOR THE PRESENCE AND INFLUENCE OF DIGIT PREFERENCE

Jay Beaman

Auctor Consulting, 676 Highland Ave., Ottawa, ON, K2A-2K4 (jaybman@igs.net) Canada and Adjunct Professor, Human Dimensions in Natural Resources Unit, Colorado State University, Fort Collins, Colorado

Michel Grenier

Senior Analyst, Parks Canada, Hull, Quebec, Canada (michel_grenier@pch.gc.ca-or-mgrenier@istar.ca)

Abstract: Digit preference which is really showing preference for certain numbers has often described as the heaping or rounding of responses to numbers ending in zero or five. Number preference, NP, has been a topic in the social science literature for some years. However, until recently concepts were not adequately rigorously specified to allow, for example, the estimation of number of responses for a given stated value that are a consequence of NP. This paper reviews developments, identifies current research directions and then builds on methods of estimating numbers exhibiting NP to provide measures and statistical tests. A chi-square test of significance of NP; a test for the significance of particular heaps and confidence intervals for the estimated proportion of a population exhibiting NP are presented. The interpretation of the tests and measure is illustrated. Of particular practical value is a measure of the potential for NP to bias the value of the mean and total of a stated response variable.

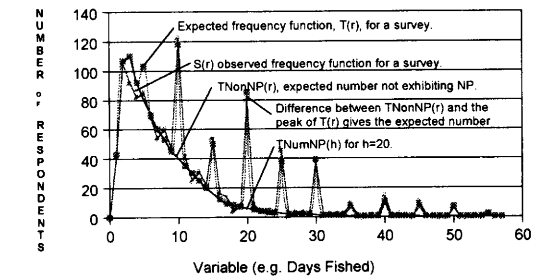
Note: In this article certain standard symbols are used. These are: $\approx$ or $\cong$ is approximately; $\equiv$ is equivalent; $\therefore$ is therefore; $\exists$ is there exists; $\forall$ is for all; $\subset$ is "is a subset of"; $\in$ is "is a member of"; $\ni$ is such that.

Keywords: digit preference, bias, statistical test, heaping, rounding, stated participation

Perspective

Digit preference, really number preference (NP), is a phenomenon that is recognized by the heaping on or rounding of responses to certain numbers such as numbers ending in zero or five. It has been a topic in the scientific literature for some years (Burton & Blair. 1991; Baker 1992; Hultsman, Hultsman, Black 1989; Huttenlocher, Hedges, & Bradburn 1990; Pickering, 1992; Rodgers, Brown, & Duncan1993; Rowe & Gribble 1993; Tarrant & Manfredo 1993; Wen, Kramer & Hoey 1993). Recent publications have operationalised the concept of exhibiting number/digit preference, NP. This has been done with adequate rigor to allow the estimation of the number of responses for a given stated value of a variable r for which NP was exhibited (Vaske, Beaman, Manfredo, Covey and Knox 1996; Beaman, Vaske, Donnelly and Manfredo 1997). A key idea is that heaping/rounding such as exhibited in response to questions about the frequency of doing something or the amount spent should be viewed as a generic phenomenon. The implication of this approach is that results can be derived that relate to a variety of practical problems of making estimates using stated responses.

Figure 1: Hypothetical Example of A Frequency Function Showing The Consequence of 0-5 Number Preference Without Random Variation (Series 1) and With Random Variation, S(r) (Series 2); and Respondents Not Exhibiting NP, NonNP(r)(Series 3)



This article presents results that are part of a larger research thrust that began in 1994 when the magnitude of heaping in Colorado fishing data was recognized (see Vaske et al. 1996). Figure 1 closely resembles an actual frequency function for the Colorado fishing data. In the figure the horizontal axis is used to indicate the stated number of days fished, r. One sees that for r=15, there are 50 responses. However, from the figure one also sees that there is a function marked as S(r) that is 50 for r=15 but that is relatively smooth for values that do not end in 0 or 5. In fact, the smooth function NonNP(r) is formed by ignoring heaps on numbers ending in zero or five. An important concept is elaborated on by Beaman et al. (1997). It is that the value of NonNP(r) for values of r at which heaps occur is an estimate of the number of people giving that responses who are not exhibiting NP. For example, one sees that for r=15 it is estimated that there are 10 people not exhibiting NP. Having estimated the number of people not exhibiting NP, one can conclude that 40 (=50-10) were respondents exhibiting NP.

The section below titled "Derivation of Statistical Tests and Measures" formalizes the idea just introduced and proceeds to derive measures and statistical tests with some degree of rigor. In that section, the result just introduced for estimating numbers of respondents exhibiting NP is expressed by Equation 1. This result is important since Hultsman, Hultsman and Black (1989) recognized that simply taking the proportion of responses ending in zero or five as an estimate of the proportion of the population exhibiting NP is biased. This is because even if nobody exhibited NP about 20% of responses would still be expected to end in zero or five. Having identified the NonNP(r) function and noting that if one can estimate the number of people exhibiting NP for each heap, Beaman et al. (1997) deduced that the proportion of the population exhibiting NP can be estimated. It is just the sum of the estimates of numbers exhibiting NP for all heaps divided by the number of respondents (Equation 2). Their analysis shows that for a given variable r, unless some special conditions occur, this approach gives essentially unbiased estimates of the number of respondents exhibiting NP. They also identified the fact that if a NonNP(r) function can be defined, the same unbiased approach to estimating the proportion exhibiting NP can be applied regardless of whether heaping is on a set of numbers ending 0 and 5 or on aother set of preferred numbers.

Huttenlocher, Hedges and Bradburn (1990) were specific in identifying prototypes, groups of numbers on which heaping occurs. Prototypes include sets of numbers other than numbers ending in zero and five. The work of Huttenlocher et al. (1990) also plays an important role in ongoing research in that they propose a psychological theory of heaping. They propose that though there may be special cases where people intentionally give a response that is too large or too small, generally **heaping is a consequence of recall being imprecise.** In other words, the heaps that one sees in Figure 1 can be viewed as a consequence of respondents giving a "best estimate" of what they believe to be the true value with which they should respond. If asked about hours in western cultures responses will, for many respondents, be rounded to a prototype of 4, 8, 10 and 12. In other contexts, for some individuals, the mind as part of retrieving a response will "produce" responses heaped on 7; 14 or 15; 30; etc. reflecting weeks and months. Unfortunately, while based on a given question some respondents are disposed to respond by one prototype, others may use another. So, digit preference or more precisely number preference, NP, can be **generically** viewed as certain respondents minds using certain prototypes rather than giving exact responses (Huttenlocher et al. 1990).

The next section accepts the prototype model of NP. Based on accepting it, set theory, statistical theory and mathematical equations are used to deduce consequences. After deriving the proportion of the respondents exhibiting NP (Equations 1 and 2), attention turns to statistical tests for heaping. Firstly, the **null hypothesis** that there is no NP exhibited is examined. In this case, variation around NonNP(r) is random. The expected value of NonNP(r), TnonNP(r), gives the means of a multinomial distribution. So one has the information about the distribution of responses so that a chi-square distribution can be defined (Equation 4) to test whether any apparent exhibition of NP (any heaping) is statistically significant. Given that the chi-squared distribution is related to the sum of squared deviations from a normal distribution, a formula for the significance of a particular heap is derived (Equation 5). Equations 6 and 7 recast the consideration of deviations in term of the difference between an expected number of responses and the number actually observed by introducing the binomial distribution and a Poisson approximation for frequency functions with heaps. When one acknowledges that there are heaps, expected values and variances do not have the values that are expected under the null hypothesis. Equations 8, 9 and 10 formalize variance estimates, distributions and statistical tests when the null hypothesis obviously does not apply.

Though most of this article focuses on statistical tests and measures for stated response frequency functions, it also addresses the matter of the influence of heaps on the mean and totals. In applied research, one is not typically interested in what proportion of respondents may be exhibiting NP but rather in the influence their response heaping has on important parameters for managerial and operational decisions. In many cases the value of the stated responses to management or operations is in estimating such statistics as total participation, average catch or total kill. Knowing if such statistics are biased is the crucial matter to clarify in considering NP. The ability to estimate the numbers exhibiting NP is used to compute a ratio showing the potential of NP to influence the estimates of means and totals (see Equations 11 and 12).

Part of the problem of examining the distribution of stated responses is that some respondents may live near streams and fish hundreds of days a year. The statistical concern is that a very small heap can be concluded to be significant based on the fact that responses around it have a frequency of zero. Equation 13 addresses this matter. However, what this article does not do is examine the general estimation problem of redistributing heaped values back to the actual or "accurate" values. These are "accurate", not rounded, values. One may think of the mind as determining that, for example, determining by some information retrieval process determining that 27 is a best estimate. However, in a second step the mind, using a particular prototype, maps 27 into 30. This second step would be invoked when there was a certain level of uncertainty about a response such as 27. Still, "accurate" responses retrieved in a first step could be off because of recall error. On its own recall error should not cause heaping. Such matters are being pursued in other research where it is argued that not recalling all events or recalling some that should not be recalled simply does not result in heaping.

Specifically, promising research on unbiased estimation is proceeding based on a Huttenlocher, Hedges and Bradburn (1990) observation. This was that **typically a respondent is not trying to bias results.** A respondent's heaped response can be taken to give **what that individual considered to be** the most accurate report of their actual behavior that would be given in the particular context. If appropriately coached they might recall more accurately but that alters the context. Given that respondents are not attempting to introduce bias by exaggerating or understating what they do, algorithms to distribute heaps back to "where they came from" are being developed. Furthermore accepting that respondents are not trying to bias survey results allows for principles of **multiple imputation** (Rubin 1996) to be invoked. Multiple imputation has been used since the 70's to produce data sets in which accurate responses for individuals become part of a data set that can be processed to produce unbiased estimates of all statistics that may be biased by heaping. Papers reporting on such developments are being prepared for publication in 1997 and 1998.

In summary, a step toward dealing with the influence of NP on a variety of estimates is taken by providing a variety of measures and statistical tests. The following section derives these. These derivations may be of limited interest . However, it is suggested that until estimation methods under development are available (current references, manuscripts and programs can be requested from the authors), researchers should compute the "influence on the mean" measure (Equation 12). This suggestion is based on managerial or operational decisions depend on means computed from data in which NP is significant. Significance can be clear either because statistical tests show it or because it is obvious that the potential for bias from NP should not be ignored.

Derivation of Statistical Tests and Measures

Assume that there is a frequency function, S(r) (see Figure 1) that is defined on a set of response values, R. Also consider that there is a preferred set $P \subset R$. The observation by Beaman et al. (1997) facilitating new findings is that: All frequencies that are not for values in P are for members of R-P. Frequencies for members of R-P are ∴ responses of people who do not exhibit NP. Based on having many replications of a study, one can visualize the observed frequencies, S(r), from each survey as distributed around a function T(r) (see Figure 1). For $r \in R - P$, one has T(r) ≡ TNonNP(r) where TNonNP(r) gives the expected number of individuals not exhibiting NP. One does not observe TNonNP(r) but rather for $r \in R - P$ one observes NonNP(r) ≡ S(r) which is subject to statistical variation from survey to survey. So for a value, h $\in P$, at which there is a heap ENonNP(r),the estimated value of TNonNP(h), should be estimable using a "smooth" curve. Such a curve would joining together values of NonNP(h) for neighboring non-heaped values (e.g. a linear or quadratic function fit to surrounding values e.g. for $r \geq 5$ values of NonNP(r) $\ni r \in \{$ h-2, h-1, h+1, h+2 $\}$ see Beaman et al. 1997).

Now, by hypothesis heaps only occur at h∈P. In Figure 1 this is for integers ending in 0 or 5. This means that observed values of S(r) for other r-values ($r \notin P \equiv r \in R - P$) are observed frequencies of not exhibiting NP, S(r) ≡ NonNP(r). This is the case as long as there is no reason for actual behavior to favor five, 10, etc. thus causing *real* peaks on such values (see Beaman et al. 1997). So, assume that there are no "real" heaps. Then, it is reasonable to conclude, as shown in the figure, that for integers $h \in P$ for which S(r)≠NonNP(r), an estimate of the value of TNonNP(h), ENonNP(r), can be obtained by extrapolating to it from non-heaped values above and below h. Furthermore, if ENonNP(h) is an estimate of the number of responses for h for which there is no NP the other responses at h are a consequence of exhibiting NP so:

Equation 1:

S(h)-ENonNP(h)=ENumNP(h)

The number of responses in which NP is exhibited, ENumNP(h), can be estimated. As indicated in Figure 1 this is an estimate of T(h)-TNonNP(h)=TNumNP(h).

The two recent articles identified above as clarifying generic NP concepts (Vaske et al. 1996 and Beaman et al. 1997) have suggested that the sum of ENumNP(h) values divided by *N*, the total number of responses on the stated response variable r, is an estimate of the proportion of a population(or of a subpopulation) exhibiting NP on the variable r. Here this measure is defined by:

Equation 2:

$$\text{PropXingNP}= \sum_{\forall h \in P} ENumDP(h) / \sum_{\forall r \in R} S(r) \cdot$$

The value of PropXingNP is .26 for the S(k) function shown. The notation $\text{PropXingNP}_{(var)}$ could be used to make it clear for which variable on a survey the measure was calculated. We do not further complicate description and notation by explicitly dealing with several variables and with nonresponse on different variables as producing subpopulations of different sizes form which estimates for a population would be made. Still, unless two variables had the same distribution for actual behavior $\text{PropXingNP}_{(var)}$ would be expected to have a different value for each. For example, if S(r) is for days fished and the other variable is total catch, c, then based on Vaske et al. (1996), PropXingNP_c would be predicted to be greater than .26. Based on the preceding, one can legitimately say that PropXingNP=.26 means that 26% of the respondents actually gave a response ending in 0 or 5 when their response should not have ended in 0 or 5; they exhibited NP. A proportion of .26 is a relative large percentage. Its large value is consistent with it being obvious by viewing Figure 1 that the heaps are too large and too regular in where they occur to reflect random variation. In Vaske et al. (1996 Figure 1) one sees that the only clear exceptions to 0-5 NP are heaping on 12 and 14. Though these might be thought to relate to avoiding 13, they also correspond to "once a month" and to "every day for two weeks." Clearly, one might ask for statistical proof of significance or more reasonably want a general significance criterion that would apply to smaller PropXingNP values such as .1 or .05 that might not be significant. What is needed is a test that allows one to reject the hypothesis that *NP is not being exhibited.*

To develop a test for the "not being exhibited" hypothesis, assume that for a given variable, r, and for *N* respondents no NP is being exhibited. Then for r the observed frequencies, S(r), are *all for people who are not exhibiting NP* thus they are just a set of observations, NonNP(r) that relate to the expected distribution, TNonNP(r). Based on the multinomial distribution TNonNP(r) gives expected frequencies, for a population of size *N* with an observation vector S(r) ≡ NonNP(r). For 0-5 NP the research cited in the introduction has defined estimates of NonNP(h), ENonNP(h), for numbers >0 ending in 0 or 5($h \in P$). For each $h \in P$, ENonNP(h) was estimated by the average of S(r) for 2 values above and two values below h:

Equation 3:

ENonNP(h)=mean(S(h-2),S(h-1),S(h+1),S(h+2))
≡Mean(NonNP(h-2), NonDp(h-1) ,NonDp(h+1), NonNP(h+2)).

It may appear "obvious" that the estimates defined by Equation 3 should be obtained by curve fitting. However, since one does not know what curve to use, results of curve fitting can produce poorer estimates than those obtained by using Equation 3. More importantly, because Equation 3 estimates each ENonNP(h) using different observations, the estimates for the different values of h are, for practical purposes, uncorrelated. This is since correlation only results from the constraint that ***the total number of respondents = the sum of all the S(r).*** Virtual independence of the estimates is important since given the definition of a chi-square with degrees of freedom, d, as the sum of squared deviates of d independent Normal(0,1) variates, based on and means, variances and asymptotic normality of the distribution of multinomial responses(e.g. see Kendal 1943, Ch. 12):

Equation 4:

$$\chi^2_d \cong \sum_{\forall h \in H} ((S(h)\text{-ENonNP}(h))^2/\text{ENonNP}(h))$$

where d=Size(H) and
ENonNP(h)>0 and $(S(h)\text{-ENonNP}(h))/\text{ENonNP}(h)^{1/2}$
≅ Normal(0,1).

As defined in Equation 3, chi-square (Equation 4)must be computed for ENonNP(u)>0 while an estimated ENonNP(r) ≤ 0 can occur. This present problems. As one sees in Figures 1 and 2, for h=30 or greater, for some values of h, surrounding values are zero(as occurs in real data examined e.g. see Vaske et al. 1996). For h where adjacent values(re adjacent see Equation 3) are zero:

> **Equation 3 Alternative Conditions**: If the mean of surrounding values is 0 then:
> (a) if S(h)>0 then, based on the assumption that all the responses in the interval of interest have actually occurred at h by chance(testing the plausibility of this assumption is introduced below), an estimate of ENonNP(h)= S(h)/5;
> (b) if adjacent S(r)=0 and S(h)=0 then the potential heap should be ignored and d reduced by 1 since, effectively, an h has been removed from P.

In the case where ENonNP(r) is estimated by prorating its value to an interval, chi-square (Equation 4)is biased if there is actually a heap. The bias results since the allocation of S(r)/5 is based on there being no heap and thus mass is assigned to adjacent values that should not be assigned to them. If there is a heap of expected size TNumNP(h), then the mass allocated to each surrounding cell(should be) has an expected value of T(h)-TNumNP(h). Regardless, the effect on chi-square of the bias is conservative in that it results in a reduced probability of NP being found to be significant.

Equation 4 is important for more reasons than that it defines chi-square (Equation 4). Firstly, note that it is defined ∀h∈P for which ENonNP(h) is estimated. The definition does not depend on 0-5 NP being studied but applies to any set of preferred values P. It is only the way that ENonNP(h) is estimated(Equation 3) that is specific to 0-5 NP. If the variable being considered was average hours spent fishing and heaps were expected at 4, 8, 10 and 12 then an "alternative" Equation 3 could be formulated allowing a chi-square to be calculated.

The second matter to note about chi-square (Equation 4) is that it can be significant simply because NumNP(h) is very large for one or two values of h.. In fact, if NP is obviously significant, there may be good reason to focus on particular residuals, NumNP(h) rather than deriving a statistical test that just confirms the obvious. Based on the way that chi-square is defined in Equation 4 and derived from the multinomial(Kendal 1943 ch 12) each term in the sum can be treated as the square of an observation from a Normal(0,1). This means that confidence intervals can be based on the distribution:

Equation 5:
Distribution((S(r)-ENonNP(r))/(ENonNP(r))$^{1/2}$)
≅Normal(0,1)

Each S(r) being an observation from a multinomial the probability of a response, h, is:

Equation 6
Eprob(h| NP is not exhibited and h is not favoured in actual behavior)= p(h|~~NP~~)=ENonNP(h)/***N***

So, under the null hypothesis the difference between S(h) and ENonNP(h), NumNP(h), has a binomial distribution. Such a distribution is defined by a ***p*** and ***N***. Thus, the following defines alternative distributions to use for making statistical tests

Equation 7: Distribution(S(h)|~~NP~~) ≅
(a) Binomial(***N***, p(h|~~NP~~)); or
(b) Distribution((S(h)-***N****p(h|~~NP~~))/(***N****p(h|~~NP~~)(1-p(h|~~NP~~))$^{1/2}$)≈Normal(0,1); or
(c) Poisson(***N****p(h|~~NP~~) for ***N****p(h|~~NP~~)<<***N***.

Here, again, the methods apply to any heaping not just 0-5NP. The important matters are that:

1. there be an hypothesis about where the heaps occur(a set of preferred values, P, must be pre-specified); and
2. there be a way of computing ENonNP(h) that is appropriate for the set P.

The arguments just made for p(h|~~NP~~) can actually apply for any S(r) where p(r)=S(r)/***N***). One can simply substitute p(r) for p(h|~~NP~~) into the expressions given in Equation 7 to obtain distributions. However, for p(h|NP) where the probability is estimated as p(h|NP)= ENumNP(h)/***N*** the same approach may seem appropriate but there is a complication. It arises because ENumNP(h)=S(r)-ENonNP(h). As indicated above, given 0-5 NP and the estimates specified in Equation 3, S(r) and ENonNP(h) are for practical purposes independent, so:

Equation 8:
Var(NumNP(h))=Var(.S(h)-ENonNP(h))
=Var(S(h))+Var(ENonNP(h)) and based on the Poisson
=S(h)+Var(ENonNP(h)) and based on Euation 3
=S(h)+Var((1/4)(S(h-2)+S(h-1)+S(h+1)+S(h+2))
≅ S(h)+(1/16)ENonNP(h)
and since for no heap E(S(h))=ENonNP(h) and otherwise
E(S(h))>ENonNP(h)
≤ S(h)+(1/16)S(h)=1.06S(h)

The preceding, in fact, provides proof that Equation 7 does not apply to NumNP. The variance derived is not

Np(h|NP)(1-p(h|NP)) which it would be if the distribution were binomial nor is E(NumNP(h))=Var(NumNP(h)) which must be the case for a Poission distribution. Of the options identified in Equation 7 one is left with approximating the distribution of NumNP(h) by a normal distribution or using weak distributional assumptions, such as expressed in Tchebycheff's inequality to define statistical tests and confidence intervals. Equation 9 shows these options:

Equation 9:

(a) the distribution of NumNP(h)

$\cong Normal(NumDP(h),(1.06S(h))^{1/2})$; or

(b) regardless of distribution using Tchebycheff's inequality(Kendal 1943, p. 203)
Probability(|NumNP(h)-TNumNP(h)|> α) $\leq 1/\alpha^2$
where α can be arbitrarily chosen e.g. as 10 to give a 1% level.

Another matter that might concern a person who is examining data in which NP occurs is whether for different data (e.g. for different segments) differences in estimates of the proportions of the populations exhibiting NP are significant. To develop a statistical test one must note that though by definition PropXingNP is an estimate of what appears to be the p of a binomial since it is based on NumNP(h) , as described above, its variance is defined by Equation 8 meaning that

Equation 10:

The distribution of PropXingNP is
(a) $\approx$ Normal(PropXingNP,((1.06/N)($\sum_{\forall h \in P} S(h)$)$^{1/2}$); or

(b) based on Tchebycheff's inequality and given "that a variance of a sampling distribution of proportions of an attribute" in samples of N
is $\leq 1/(4N)$(Kendal 1943, p. 203):
Probability(|PropXingNP-TPropXingNP|)>k)
$\leq 1/(4N*k^2)$
where k $\ni$ 0<k<1 and is e.g chosen for a 1% level.

The preceding provides tests for significance and confidence intervals, however a small part of a population may exhibit NP with high variability. Knowing NumNP(h) allows one to assess the degree to which NP is contributing to the mean or total of the response variable as follows:

Equation 11:

Sum(r|S(r))= $\sum_{\forall r \in R} rS(r)$

=Sum(rS(r) $\forall r \in$ R-P)+ Sum(hS(h) $\forall h \in$ P)
= Sum(rS(r) $\forall r \in$ R-P)+
Sum(h(ENonNP(h)+ENumNP(h)) $\forall h \in$ P)
=(Sum(rS(r) $\forall r \in$ R-P)+ Sum(hENonNP(h) $\forall h \in$ P))+
Sum(hENumNP(h) $\forall h \in$ P)
=Total related to not exhibiting NP+Total related to exhibiting NP

So, the contribution to the total and thus to the mean of r can be partitioned based on exhibiting and not exhibiting NP. This suggest that a measure of the potential for exhibiting NP to influence the mean of r be defined by:

Equation 12:

PropMeanNP= $\sum_{\forall h \in H} h * ENumDP(h) / \sum_{\forall r \in R} rS(r)$

Figure 2: Hypothetical Frequency Functions For Computing NP Measures And Tests

(for descriptions of the functions see the text)

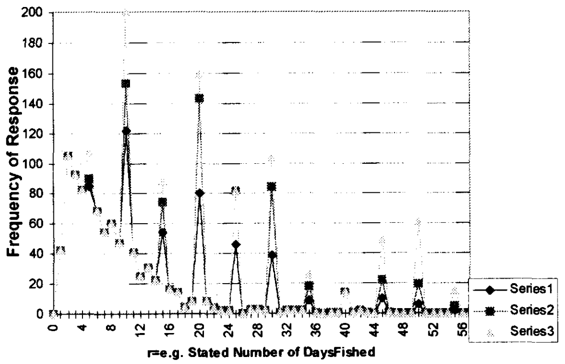
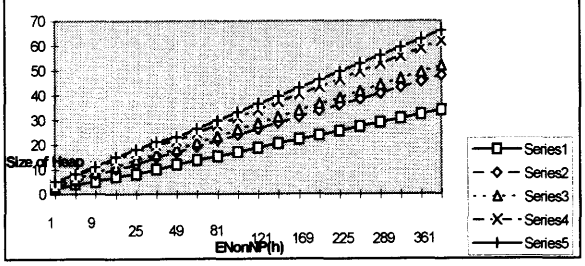


Figure 3: Deviations That Are Significant for a Given ENonNP(h) at the Levels Defined for the Different

(Series 1=.05, Series 2=.01, Series 3=.005, Series 4=.001, Series 5=.0005)



Results of Computations for Examples

Some measures and statistical tests have been calculated for the hypothetical S(r) data shown graphically in Figure 2. Series 1 in Figure 2 is Series 1 in Figure 1. The other two series in Figure 2 are variations on Series 1. Series 2 has about the same level of NP for small numbers as Series 1 but by r= 25 about twice as many people exhibit NP in Series 2. Series 3 has more people exhibiting NP than Series 2 except for r= 26. Because the three series share the same NonDp(r) function one knows that NumXingNP will increase in value from Series 1 to 2 to 3. Also chi-square (Equation 4) and PropMeanNP values are expected to satisfy the same relation. The PropXingNP values are .26, .38 and .45. Because the NP peaks are large and obvious chi-square is expected to be so large as to have an infinitesimal probability of occurring by chance. For chi-square with d=11 the .001 level for it is 28 while values calculated for it start at 3313 for Series 1 and are larger for Series 2 and 3 (11,962 and 19,301). One way to see this is by examining the probability that a heap like the one at 30 occurred by chance. For Series 1, based on a normal approximation, the difference between the heap frequency and the expected frequency is 7.3 times the expected standard deviation. Given that deviations of more than 4.1 standard deviations occur about 1 time in 10,000, the heap is clearly significant.

Given that there is significant NP being exhibited, if management is interested in a mean or total then one still has no idea whether there is potential for such values to be biased. The measures of the potential influence of NP on the mean (PropMeanNP, Equation 12) are .47, .63 and .71. There is obvious reason to be concerned that treating the heaps as valid observations could cause a significant bias in estimates of means and totals.

When chi-square (Equation 4) will obviously be significant, as is the case in the examples presented here, simply showing that one or more heaps is highly significant using distributions defined in Equation 7 is an easy alternative to computing chi-square (Equation 4) . Figure 3 was prepared based on Equation 7. It facilitates testing for the significance of heaps since one need only determine the value of S(r), estimate NonNP(h), to use the figure to determine the value of NumNP(h) that is significant at one of the levels given. The only issue then is if the estimate NumNP(h) is large enough to be significant. The .001 and .0005 levels are provided in Figure 3 so that if, for example, the significance of η heaps is being considered, one can use the Bonferonni principal that if η heaps, say 10, are significant at the λ (e.g. λ =.001). level, then all considered together are, at least, significant at a level $\lambda\times\eta$ (e.g. =10(.001)=.01). Ongoing research suggests that understanding the structuring of significant heaps is necessary to produce estimates of NP bias.

The Poission distributional approximation introduced allows statistical significance to be examined in another way. Assumption 3- Alternative Conditions (a) specified that an estimate of NonNP(h) be based on T(h)=S(h)/5 given that S(r)=0 $\forall r \in A = \{h-2, h-1, h+1, h+2\}$. One can examine the likelihood of heaps surrounded by zeros by considering:

Equation 13:

Prob(S(h) and S(r)=0 $\forall r \in A$) $\leq \Theta$

where it is assumed that there is no NP and no *real heaps; and* where Θ is the desired level of significance(e.g. .01).

In figure 2, one can see a potential heap of 2 at 55 surrounded by 4 zeros. Accepting the Poission distribution approximation as specified in Equation 7, based on T(r) $\cong$ S(h)/5, the probability of 0 for a given S(r) is $e^{-S(h)/5}$. There is only one way for all 4 adjacent observations to be 0 so the probability of this is $(e^{-S(h)/5})^4 = e^{-4*2/5}$=.2. For the heap of 6 at 50 the probability is $(e^{-S(h)/5})^4 = e^{-4*6/5}$=.008 so, conservatively, the heap is significant at the 1% level. Actually this shows that *any heap of 5 or greater,* ***considered on its own,*** *is significant at the 1% level.*

Conclusion

This research has immediate consequences and implications for future research. Its immediate practical importance is providing statistical tests for the significance of NP in data and a measure of the potential of NP to cause bias in the mean or total of a response variable (e.g. in average days fished or total catch). By taking a generic approach, attention is not on 0-5 NP or a preferred set of 4, 8, 10 and 12 (because a question is about hours). Continuing research is showing that as the "NP processes" are better understood criteria for rational decisions about ignoring, attenuating or correcting NP are being developed.

References

Beaman, J. G.; Vaske, J.; Donnelly, M. and Manfredo, M.. (1997). Individual versus Aggregate Measures of Digit Preference. *Human Dimensions of Wildlife*, 2(1):71-80.

Burton, S., & Blair, E. (1991). Task Conditions, Response Formulation Processes And Response Accuracy For Behavioral Frequency Questions In Surveys. *Public Opinion Quarterly*, 55:50-79.

Baker, M. (1992). Digit Preference In CPS Unemployment Data. *Economic Letters*, 39:117-121.

Hultsman, W. Z.; Hultsman, J. T. & Black, D. R. (1989). "Response Peaks" As A Component of Measurement Error: Assessment of Implications for Self-Reported Data in Leisure Research. *Journal of Leisure Research*, 21:310-315.

Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of Elapsed Time: Bounding and Rounding Processes in Estimation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16:196-213.

Pickering, R. M. (1992). Digit Preference in Estimated Gestational Age. *Statistics in Medicine*, 11:1225-1238.

Rodgers, W. L.; Brown, C. & Duncan, G. J. (1993). Errors In Survey Reports of Earnings, Hours Worked, And Hourly Wages. *Journal of The American Statistical Association*, 88:1208-1218.

Rowe, G. & Gribble, S. (1994). Income Statistics from Survey Data: Effects of Respondent Rounding. *Proceedings of The Business And Economic Statistics Section of The Annual Meetings of The American Statistical Association.*

Rubin, Donald B.(1996). Multiple Imputation after 18+ Years. *Journal of The American Statistical Association*, 91:434ff..

Tarrant, M. A. & Manfredo, M. J. (1993). DP, Recall Bias, And Non-Response Bias In Self Reports of Angling Participation. *Journal of Leisure Research*, 15:231-237.

Vaske, J. J.; Beaman, J.; Manfredo, M. J.; Covey, D. & Knox, R. (1996). Response Strategy, Recall Frame And Digit Preference In Self-Reports of Angling Participation. *Human Dimensions of Wildlife*, 1(4):54-68.

Wen, S. W.; Kramer, M. S. & Hoey, J. (1993). Terminal Digit Preference, Random Error and Bias in Routine Clinical Measurement of Blood Pressure. *Journal of Clinical Epidemiology*, 46:1187-1194.

Westat, Inc. (1989). *Investigation of Possible Recall/Reference Period Bias In National Surveys of Fishing, Hunting And Wildlife-Associated Recreation.* Washington, DC, U.S. Department of the Interior.