

ETHNOLOGY

AN INTERNATIONAL JOURNAL OF
CULTURAL AND SOCIAL ANTHROPOLOGY



Spring 1993

Volume XXXII

Number 2

**DIMENSIONS OF SUBSISTENCE FORAGING
IN SOUTH INDIA** 109

Peter M. Gardner

**UNCERTAINTY, HUMILITY, AND ADAPTATION IN
THE TROPICAL FOREST: THE AGRICULTURAL
AUGURY OF THE KANTU'** 145

Michael R. Dove

**A RITUAL RESPONSE TO CLIMATIC PERTURBATIONS
IN THE HIGHLANDS OF PAPUA NEW GUINEA** 169

Paul Sillitoe

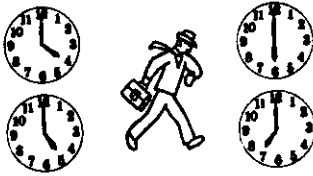
**FIRST-TIME TELEVIEWING IN AMAZÔNIA:
TELEVISION ACCULTURATION IN GURUPÁ, BRAZIL** 187

Richard Pace

SAMPLING IN TIME ALLOCATION RESEARCH 207

H. Russell Bernard
and Peter D. Killworth

SAMPLING IN TIME ALLOCATION RESEARCH



H. Russell Bernard
University of Florida

Peter D. Killworth
Oxford University

Sampling behavior by direct observation is a common technique in field studies of behavior (Lehner 1979; Martin and Bateson 1986). There are two ways to sample behavior: (1) focus on individuals and monitor them continuously for a period of time; (2) record the behavior of individuals at random times throughout the period of research. The former technique is called focal-animal sampling; the latter is known as spot sampling. Spot sampling is the basis for time allocation studies in anthropology, though there remains some controversy in primate studies about its merits compared to focal-animal sampling (Altmann 1974; Rhine and Flanigon 1978; Harcourt and Stewart 1984).

Spot sampling is based on a simple, appealing axiom: if you sample a representative number of moments in, say, a week or a year, and if you note what people are doing at those moments, then the percentage of times that people are seen doing things (eating, working, cooking, playing) is the percentage of time they spend doing those things (Erasmus 1955; Johnson 1975; Gross 1984).

The axiom holds, of course, only if the sample of moments is in fact representative, which raises the question: How big is a representative sample of moments? It seems to us that this question can sensibly be interpreted in three ways:

- (1) If some activity is fairly rare, how many times must an observer plan to make observations in order to be reasonably certain of even seeing the activity at all?
- (2) How many times must observations be made so that the frequency of an activity can be estimated within some desired error bounds?
- (3) How many times must observations be made so that the frequency of an activity can be known to be less than some desired amount, with reasonable accuracy?

The statistics to answer all three questions are well known, but do not seem to have been used in the current connection. Accordingly, we derive and tabulate the answers to these three questions and discuss them in the light of previous work.

ESTIMATING SAMPLE SIZE FOR SPOT OBSERVATION: AN EMPIRICAL SOLUTION

Michael Baksh (1990) had 4,182 observations of Machiguenga (Peru) men in 41 households of twelve activities and 9,673 observations of Embu (Kenya) women in 169 households on fifteen activities. Baksh randomly selected 90 per cent of the cases in his Machiguenga data and compared the percentages for each of the twelve activities (eating, child care, idle, etc.) against the percentages in 100 per cent of his data. The percentages were nearly identical. He repeated the exercise for random samples of 80 per cent of his data, 70 per cent of his data, and so on. All the percentages remained stable down to 20 per cent and changed dramatically only at 10 per cent of his data. In other words, he was able to recover the pattern of time allocation with a random sample of 137 observations.

Baksh then examined one activity in his Embu data: eating. Relatively little time is devoted to this activity so, Baksh reasoned, estimates of time spent eating should be sensitive to changes in sample size. He ran 25 trials of randomly selected observations drawn for thirteen sample sizes: 50 per cent, 10 per cent, 2 per cent, and increments of 0.1 per cent down to 1 per cent. The mean eating time for the 50 per cent sample was 4.4 per cent and the mean eating time for the 1 per cent sample was 5.0 per cent. The difference in the means was insignificant but, of course, the range increased as the sample size decreased.

Baksh plotted the per cent of sample size against the ranges for the estimates of mean eating time in all the samples. By inspection, he found that the range of the estimates remained at 4 per cent or less in samples that were 1.6 per cent or greater. Considering this acceptable, Baksh concluded that a sample of "about 150 observations" was sufficient to estimate Embu women's time use (1.6 per cent of 9,673 is 157).

In the following section, we approach the sampling problem analytically. It turns out that Baksh's pioneering brute force estimates were excellent, but demonstrate an acceptance of what some may feel is too large an error bound. An analytic solution, however, offers several advantages, which we adumbrate later.

ESTIMATING SAMPLE SIZE FOR SPOT OBSERVATION: AN ANALYTIC SOLUTION

Suppose an activity occurs a fraction f of the time. We can derive three quantities of interest: (a) the number of observations needed in order to be certain to some degree (say, 95 per cent) of observing the activity at least once; (b) the number of observations necessary to be able to estimate the value of f to within some specified fractional accuracy; e.g., to estimate f to within, say, 5 per cent of its true value; and (c) the number of observations necessary to be able to bound f , i.e., to say that f is less than some value e with 95 per cent probability.

(a) To See Something Once

If n observations are made, the probability of not observing a particular activity on each occasion is $(1 - f)^n$. Thus, the probability that the activity is observed at least once is

$$q = 1 - (1 - f)^n \quad (1)$$

We require q to be some value—say, 95 per cent in the example above. The number of observations needed, then, is

$$n = \frac{\ln(1 - q)}{\ln(1 - f)} \quad (2)$$

where \ln is the natural logarithm. When f is small, the denominator is just f , so that the number of observations needed for rare activities is roughly inversely proportional to the frequency f .

The numerator, on the other hand, is not just unity (i.e., one expects to see one event with a frequency of 0.01 every 100 observations). Rather, it is $\ln(1 - q) = 3$ for 95 per cent probability. This indicates that one needs to make about three times as many observations than might have been expected, to be reasonably certain of observing the activity.

In this example, if an activity occurs 1 per cent of the time, one needs to make 299 rather than 100 observations to be 95 per cent certain of seeing the activity at least once. The last column in Table 1 gives figures for various values of f .

(b) To Estimate the Percentage f

Suppose we want to estimate f to within a fractional error r (say, 5 per cent). That is, if we observe a behavior 20 per cent of the time, we want to be able to say with 95 per cent confidence that it occurs between 19 per cent and 21 per cent of the time. Let g be the width of the normal distribution containing all but 5 per cent of its area; i.e., g is 1.96.

Now the expected value of occurrences of the activity following n observations will be $E = nf$, with a standard error $\sqrt{nf(1 - f)}$.

Thus rE (the required error) must be larger than $g\sqrt{nf(1 - f)}$ (the random error).

Therefore

$$E = nf > \left(\frac{g}{r}\right)^2 \cdot (1 - f) \quad (3)$$

or

Table 1

Number of Observations to Estimate f to Within Fractional Accuracy
(rounded up to an integer)

Frequency of Activity	Number of observations							
	f	0.05	0.10	0.15	0.20	0.30	0.40	0.50
0.01	152128	26277	9123	4055	1178	436	180	299
0.02	75295	13179	4515	2007	583	216	89	149
0.03	49685	8696	2979	1324	384	142	58	99
0.04	36879	6455	2211	983	285	105	43	74
0.05	29196	5110	1751	778	226	83	34	59
0.06	24074	4213	1443	641	186	69	28	49
0.07	20415	3573	1224	544	158	58	24	42
0.08	17671	3093	1059	471	136	50	20	36
0.09	15537	2719	931	414	120	44	18	32
0.10	13830	2420	829	368	107	39	16	29
0.15	8708	1524	522	232	67	24	10	19
0.20	6147	1075	368	163	47	17	7	14
0.25	4610	806	276	122	35	13	5	11
0.30	3585	627	215	95	27	10	4	9
0.40	2305	403	138	61	17	6	2	6
0.50	1537	268	92	40	11	4	2	5

$$n > \left(\frac{g}{r}\right)^2 \cdot \left(\frac{(1-f)}{f}\right) \quad (4)$$

Put another way, the 95 per cent bounds of an estimate of f are

$$\pm g \sqrt{\frac{f(1-f)}{n}} \quad (5)$$

(This assumes that the normal distribution is relevant, and omits minor variations involving Student's t , Gaussian versus normal, and the like for simplicity.) Values for n are given in Table 1 for various values of f and r . Note that for small f , n is inversely proportional to f , so that the table can easily be extended to lower values

of f by inspection. For example, to estimate with a fractional accuracy of 50 per cent an activity that occurs .005 per cent of the time, simply multiply $180 \times 5 = 900$. Similarly, to compare two estimates for statistical difference, one simply checks the 95 per cent bounds of the two estimates for overlap; if overlap occurs, the two estimates do not differ significantly at the 5 per cent level.

(c) To Bound the Percentage f

We suppose that a frequency f has been observed within n observations, and wish to be able to say that f is less than some value e (say) with 95 per cent probability. Thus we must have

$$f + \left(g \sqrt{\frac{f(1-f)}{n}} \right) < e \tag{6}$$

as well as

$$e - \left(g \sqrt{\frac{e(1-e)}{n}} \right) > f \tag{7}$$

The second is usually the stronger of these restrictions, so that solving for n gives

$$n > e(1-e) \left(\frac{g}{(e-f)} \right)^2 \tag{8}$$

Values of n for a range of f (and $e > f$) are given in Table 2. When f is small, and e is quite large (0.4, say), values of n are tiny; only a few observations will convince the observer that f cannot be very large in this case.

DISCUSSION

Some of the numbers in Table 1 seem daunting. If an activity occurs just 1 per cent of the time, and you want to be able to say that it occurs no less than 0.0095 and no more than 0.0105 of the time, then 152,128 observations are required. Fortunately, 95 per cent accuracy (5 per cent error bounds) is generally unnecessary in time allocation research. If an activity occurs just 1 per cent of the time, and if you are comfortable saying that it occurs no less than 0.5 per cent and no more than 1.5 per cent of the time (error bounds of 50 per cent on 1 per cent), then just 180 observations are required. If you elect to make 299 observations in order to capture at least one instance of activities that occur with 1 per cent frequency, the error bounds are just 45 per cent.

Table 2

Number of observations necessary (to nearest integer) to ensure that an activity with frequency f will be estimated to occur not more than frequency e on 95 per cent of occasions. (Note that the low entries would give meaningless values in practice, where the last column of Table 1 should be used in preference.)

Frequency of activity	e															
	f	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20	0.25	0.30	0.40
0.01	-	753	279	164	114	87	69	58	49	43	25	17	13	10	6	4
0.02	-	-	1118	369	203	135	100	79	64	54	29	19	14	10	6	4
0.03	-	-	-	1475	456	241	156	113	87	71	34	21	15	11	7	4
0.04	-	-	-	-	1825	542	278	177	126	96	40	24	16	12	7	5
0.05	-	-	-	-	-	2167	625	314	197	138	49	27	18	13	8	5
0.06	-	-	-	-	-	-	2501	707	350	216	60	31	20	14	8	5
0.07	-	-	-	-	-	-	-	2827	787	384	77	36	22	15	8	5
0.08	-	-	-	-	-	-	-	-	3146	864	100	43	25	17	9	5
0.09	-	-	-	-	-	-	-	-	-	3457	136	51	28	18	10	6
0.10	-	-	-	-	-	-	-	-	-	-	196	61	32	20	10	6
0.15	-	-	-	-	-	-	-	-	-	-	-	246	72	36	15	8
0.20	-	-	-	-	-	-	-	-	-	-	-	-	288	81	23	11
0.25	-	-	-	-	-	-	-	-	-	-	-	-	-	323	41	15
0.30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	92	24
0.40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	96

Time allocation researchers usually code five to twenty activities. The activities are things like resting, eating, caring for children, manufacturing, gardening, and so on. These are really sets of activities. Packaging them ensures that each recorded activity occurs a significant fraction of the time in the data, but eliminates finer scale measurement and analysis of behavior. Many activities (weeding rather than gardening, suckling rather than child care, making baskets rather than manufacturing) are quite rare but are of interest to researchers.

With modern, hand-held coding devices, field researchers can observe and code accurately up to 90 different activities (Whiten and Barton 1988; Hile 1991). Extrapolating from Table 1 shows that about 15,000 observations are sufficient to capture activities that occur just 0.1 per cent of the time (about one minute out of a sixteen-hour waking day) with 50 per cent error. The 50 per cent error means that if you find two occurrences of a behavior in 15,000 observations (about .0001), you can say the activity occurs no fewer than once out of 15,000 and no more than three times out of 15,000 on average.

For many research purposes in anthropology (and in animal behavior studies, for that matter) this may be more accuracy than is needed. If all one needs is to say that the activity does not occur more than, say, 0.5 per cent of the time (a 400 per cent error in estimating f !) only 239 observations are needed:

$$\left(\frac{1.96}{.004}\right)^2 \times 0.001 \times 0.999 = 239 \text{ only.}$$

The researcher must thus choose the accuracy required.

An analytic solution to the time sampling problem offers a solution to the problem of comparing time estimates by comparing the spreads using (5). Suppose your data show that men eat 4 per cent of the time while women eat 6 per cent of the time. If you have 180 observations, then the error bounds of the two estimates overlap considerably (0.01-0.07 for the men and 0.04-0.08 for the women). You need about 15,000 observations to tell whether 0.06 is really bigger than 0.04 comparing across groups. Similarly, for activities: if women are seen gardening 20 per cent of their time and caring for children 25 per cent of their time, then 1,066 observations are needed in order to tell if women really spend more time caring for children than gardening.

APPLYING THE ANALYTIC SOLUTION: FRATKIN'S LEWOKOSO DATA

Fratkin (1989) did two time allocation studies of a Lewokoso Lukumai settlement in northern Kenya. In the first study, he observed 39 households at random times between 5:30 a.m. and 8:00 p.m. for seven days in October, 1985. Fratkin asked the whereabouts of everyone in each household and combined the self-reports of household members with his own observations. We'll assume that the data contain no measurement error and focus here only on the possibility of sampling error.

Fratkin had 42 observations of married men and 79 observations of married women. His data show that married men spent 52.4 per cent of their time at rest, compared to 35 per cent for married women. According to our formula, the 95 per cent confidence limits on these two figures is 52.4 per cent ± 15.1 per cent for men (37.3-67.5 per cent) and 35 per cent ± 10.5 per cent for women (24.5-45.5 per cent). These data hint at a difference between the amount of time that Lewokoso Lukumai settlement men and women spend at rest, but we can't be sure at the 95 per cent confidence limits.

In his second study, Fratkin selected three poor households and four rich households. For one week in December, 1985, he and his assistants visited the seven households every 30 minutes between 5:30 a.m. and 8:00 a.m., and also between 5:50 p.m. and 8:00 p.m. From 8:00 a.m. to 4:00 p.m., they visited the households every two hours.

Fratkin wound up with 594 observations of men in poor households, 732 observations of men in rich households, 609 observations of women in poor households, and 1,208 observations of women in rich households. According to these data, men in poor households spend 40.8 per cent of their time at rest, compared to 58.7 per cent for men in rich households, while women in poor households spend 30.2 per cent of their time at rest, compared to 40.6 per cent for women in rich households.

From our formula (5) again, here are the four ranges:

Rich men	58.7%	$\pm 3.6\%$	(55.1%-62.3%)
Poor men	40.8%	$\pm 4.0\%$	(36.8%-44.8%)
Rich women	40.6%	$\pm 2.8\%$	(37.8%-43.4%)
Poor women	30.2%	$\pm 3.6\%$	(26.6%-33.8%)

Fratkin's data make the case clearly. When it comes to leisure, it's best to be rich and male in Lewokoso Lukumai settlement. However, to be able to assert this with confidence requires the level of observational effort (of order 600 observations) that Fratkin was able to bring to the problem. His original study, based on 42 observations of men and 79 of women did not permit such assertions.

We now re-examine Baksh's (1990) estimates of 150 observations being necessary to permit accurate evaluations of frequencies. Baksh made these estimates, as noted above, by choosing a subsample size (1 per cent, say), and taking 25 subsamples of that size from the total 9,673 observations he had made. If we examine the 1.5 per cent subsamples, which led him to suggest the figure of around 150 observations, we find that while the mean was indeed 4.4 per cent, the maximum and minimum estimates were 7.1 per cent and 2.3 per cent respectively.

Now, knowing the mean was "correct" over 25 such subsamples is not too helpful, since the actual observer only takes one sample. Thus from Baksh's figures, the observer has a one in 25 chance of misestimating the 4.4 per cent as 7.1 per cent, and a one in 25 chance of misestimating it as 2.3 per cent. Baksh clearly felt this degree of error was acceptable in the at least 8 per cent (i.e., two out of 25) of the occasions on which it would occur.

How does this chance of misestimation to this degree match with the formulae above? Taking the worse of the two errors, 0.027 (= 0.071 - 0.044), we can ask how large the sample size would have to be for this to occur at most 8 per cent of the time. A spread in the data of ± 1.75 standard deviations covers 92 per cent of the normal distribution. Substituting this for g instead of 1.96, and using (5) for the spread, gives $n = 176$ for Baksh's data, very close to the 150 he estimated empirically.

CONCLUSION

Baksh's estimate of 150 observations as an all-purpose sample size for time allocation studies was correct, so long as:

- (1) the number of coded behavior categories is small (ten to fifteen);
- (2) the frequency of the behaviors is not too small;
- (3) the accuracy required for the estimate of frequency is not stringent; and
- (4) fine-grained distinctions in time estimates (across groups or across behaviors) are not required.

Where fine-grained distinctions in time estimates are required, the analytic solution presented in equations (2) and (4) can be used to calculate the number of observations needed for any given level of accuracy. The large numbers of observations required for reasonable accuracy are somewhat alarming, and should be taken into account in future experiment design. Only if the researcher is comfortable with wide error bars (e.g., if an upper bound on frequency is all that is required) can small numbers of observations be relevant.

BIBLIOGRAPHY

- Altmann, J. 1974. Observational Study of Behavior: Sampling Methods. *Behaviour* 49:227-67.
- Baksh, M. 1990. The Spot Observation Technique in Time Allocation Research (Continued). *Cultural Anthropology Methods Newsletter* 2(3):4-5.
- Erasmus, C. 1955. Work Patterns in a Mayo Village. *American Anthropologist* 57:322-33.
- Gross, D. R. 1984. Time Allocation: A Tool for the Study of Cultural Behavior. *Annual Review of Anthropology* 13:519-58.
- Fratkin, E. 1989. Household Variation and Gender Inequality in Ariaal Pastoral Production: Results of a Stratified Time-Allocation Study. *American Anthropologist* 91:430-39.
- Harcourt, A. H., and K. J. Stewart. 1984. Gorillas' Time Feeding: Aspects of Methodology, Body Size, Competition and Diet. *African Journal of Ecology* 22:207-15.
- Hile, M. G. 1991. Hand-held Behavioral Observations: The Observer. *Behavioral Assessment* 13:187-96.
- Johnson, A. 1975. Time Allocation in a Machiguenga Community. *Ethnology* 14:301-10.
- Lehner, P. N. 1979. *Handbook of Ethological Methods*. New York.
- Martin, P., and P. Bateson. 1986. *Measuring Behaviour: An Introductory Guide*. New York.
- Rhine, R. J., and M. Flanigon. 1978. An Empirical Comparison of One-zero, Focal Animal, and Instantaneous Methods of Sampling Spontaneous Primate Social Behavior. *Primates* 19:353-61.
- Whiten, A., and R. A. Barton. 1988. Demise of the Checklist: Using Off-the-shelf Miniature Hand-held Computers for Remote Fieldwork Applications. *Trends in Ecology and Evolution* 3:146-48.