A NOTE ON THE NON-NORMALITY OF DATA FROM MARK-RECAPTURE TRAPPING EXPERIMENTS

Richard F. Green
Department of Statistics
University of California
Riverside, California 92521

ABSTRACT

In studying animal movements the mark-release-recapture method is often used to measure dispersal rate or home range size. The recapture locations are often treated as if they are normally distributed. In this paper it is shown that if the movements of dispersing or of foraging animals are random, the recapture data will not have a normal distribution but rather will have a mixture of normals. This fact may help explain certain experimental observations.

A NOTE ON THE NON-NORMALITY OF DATA FROM MARK-RECAPTURE TRAPPING EXPERIMENTS

INTRODUCTION

When animal movements are studied by mark-recapture trapping, the data on retrapping is sometimes analyzed as if the locations where the animals are retrapped have a bivariate normal distribution (Dobzhansky and Wright, 1943, 1947; Calhoun and Casby, 1958; Jennrich and Turner, 1969). One mechanism suggested (Andrewartha and Birch, 1954) to produce such a bivariate normal distribution is random movement (Brownian motion) of the animals.

In fact, if animal movements are random, recapture locations should not show a bivariate normal distribution. If the recapture locations do show a bivariate normal distribution this is evidence that the animal movenents are not random.

The cases of animal dispersal and of home range will be discussed here.

Example 1. In their studies of dispersal distance Dobzhansky and Wright (1943, 1947) released a number of genetically marked drosophila and for several days they counted the number of flies recaptured in each of a number of traps arranged in the shape of a cross with the release point at the center of the cross.

According to Dobzhansky and Wright (1943, p. 320): "The simplest hypothesis with respect to the dispersion is that the distribution on any day is a radially symmetric bivariate normal one except as modified by local conditions." The observed distribution was not normal. There were too many flies trapped near the release point and far from the release point and too few trapped at intermediate distances. That is, the distribution was leptokurtic: the kurtosis (the ratio of the fourth central moment to the squared variance) was too great. For a normal distribution the kurtosis is 3 and for most experiments the kurtosis observed by Dobzhansky and Wright was four or more and was particularly high on the first day.

Example 2. In many studies on home range in small mammals the animals are captured, marked, released, and recaptured. The locations at which they are recaptured are used to determine the home range. Hayne (1949) suggested that the animals might not use their home range uniformly and suggested what was later called a "utilization distribution" by VanWinkle(1975). Calhoun and Casby (1958) suggested that a bivariate normal distribution be used to represent a home range. They considered a circular home range, while Jennrich and Turner (1969) suggested that the home range might have an elliptical shape.

In her work on peromyscus Myton (1974) demonstrated that mark-recapture data may show too many captures near the center of the home range or far from the center and too few at intermediate distances. She offered her data as evidence that the bivariate normal assumption may be wrong.

This paper shows that if animals forage at random this need not be reflected in a bivariate normal distribution of recapture locations. Further, it is pointed out that recapture locations may be more a measure of how the animals move than of where they spend their time.

THE MODEL

I assume that each animal moves at random, according to two dimensional Brownian motion, so that at time T after the animal starts moving the displacement (X(T), Y(T)) of the animal from the origin will have a bivariate normal distribution, with density

$$f(x,y) = \frac{1}{2\pi T}\exp(-(x^2+y^2)/2T).$$

This assumption of random movement is that mentioned by Andrewartha and Birch (1954, p. 94) in discussing the papers of Dobzhansky and Wright. The purpose of this model is to investigate the distribution of recapture locations if the animals move randomly.

Animals moving at random will at any fixed time be displaced from the origin by a distance that has a normal distribution. However, if animals are trapped they will stop moving and where they are trapped will be related to when they are trapped. If a large proportion of the animals are trapped during one trapping interval there will be a tendency for most of the animals trapped to be trapped early in the interval, while if a small proportion of the animals are trapped during an interval the capture times will tend to be spread uniformly over the interval. Animals that are trapped early will tend not to have moved as far as those that are trapped later.

The question examined here is: What will be the distribution of recapture locations when the fact that not all animals enter a trap at the same time is taken into account?

Since I assume that movements in the x- and y-directions are independent with the same variances I will concentrate on the x-direction. Besides the assumption of random movement I will make several other assumptions.

1. The trap grid is fine. That is, any value of X may be observed.

2. The trap "size" is small. That is, while a trap will be entered if it is encountered, it is easy to pass between traps.

3. The probability of being trapped in unit time (the interval at which traps are examined) is p and the probability of having been trapped by time T ($T \leq 1$) is $1 - \exp(-kT)$ for a constant k (trapping rate) such that $1 - \exp(-k) = p$.

If an animal is not trapped in unit time its displacement in the x-direction will have a standard normal distribution: $N(0,1)$. The question is, what is the distribution of X, the displacement in the x-direction of a trapped animal?

There are a number of different cases, corresponding to the different values of k, the trapping rate. High values of k, resulting in what has been called "trap interference", mean that many animals are trapped in each interval while low values of k mean that few animals are trapped. I will consider the limiting case $k = 0$ separately and will consider the cases $k > 0$ together.

CALCULATIONS

Case 1. $k = 0$. This means that very few animals are recaptured in any interval and that those that are recaptured are trapped uniformly throughout the interval. If we write the recapture time as T then the probability density for the recapture distance X will be a uniform mixture of normals,

$$f(x) = \int_0^1 \frac{1}{\sqrt{2\pi T}} \exp(-x^2/2T)\,dT.$$

This density is symmetric about 0. Values of f(x) are given for various values of $x \geq 0$ in Table 1. The density f(x) is plotted in Figure 1 along with the normal density with the same mean and variance, namely $\mu = 0$, $\sigma^2 = 1/2$. It may be seen that f(x) is higher than the normal density near 0 and for large values of x but is lower for intermediate values. This is just what is observed in some of the work on dispersal and home range.

Dobzhansky and Wright suggest looking at kurtosis to see whether their data is normal. The normal distribution has a kurtosis of 3.

In Case 1 the fourth central moment is

$$3\int_0^1 T^2\,dT = \frac{3T^3}{3}\Big|_0^1 = 1$$

Table 1

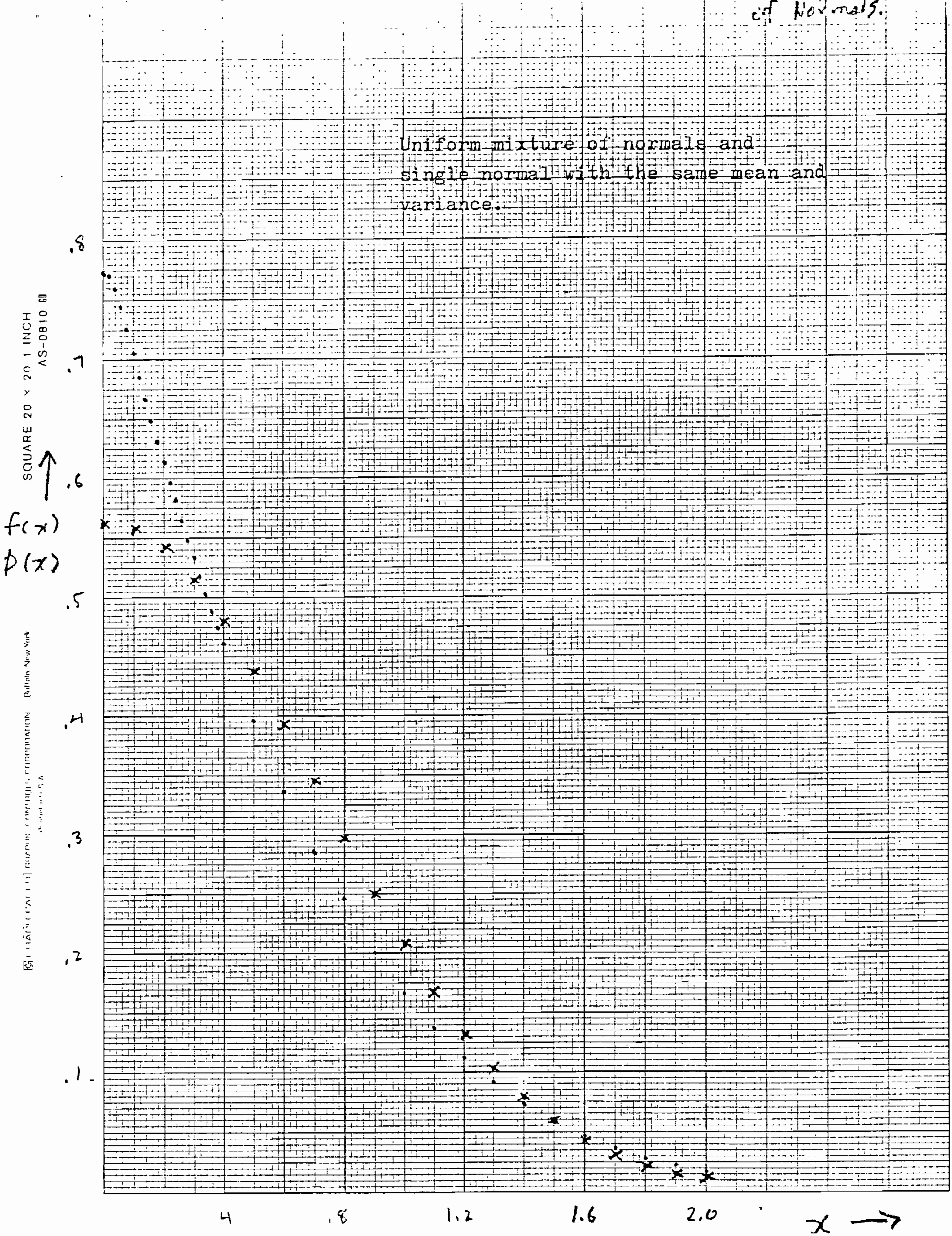
x	f(x)	$\phi(x)$
.0	.7738	.5642
.1	.7059	.5586
.2	.6135	.5421
.3	.5335	.5121
.4	.4609	.4808
.5	.3956	.4394
.6	.3373	.3936
.7	.2858	.3456
.8	.2404	.2975
.9	.2009	.2510
1.0	.1666	.2076
1.1	.1372	.1682
1.2	.1122	.1337
1.3	.0911	.1041
1.4	.0733	.0795
1.5	.0586	.0595
1.6	.0465	.0436
1.7	.0366	.0314
1.8	.0286	.0221
1.9	.0221	.0153
2.0	.0170	.0103
2.5	.0040	.0011
3.0	.0008	.0001

f(x) = Uniform ($T = \sigma^2 \sim U(0,1)$) mixture of normals,

$\phi(x)$ = Normal ($\mu = 0, \sigma^2 = .5$), same mean and variance as f(x).

Figure 1

$\hat{f}(x)$: Uniform, $\sigma^2 \sim U(0,1)$ mixture of Normals.



while the variance is

$$\int_0^1 T dT = \frac{T^2}{2}\Big|_0^1 = 1/2.$$

Thus, the kurtosis is $1/(1/2)^2 = 4$.

Case 2. $k > 0$. In this case animals are more likely to be captured at the beginning than near the end of the interval. The probability density of the capture time will be truncated exponential

$$g(t) = k\exp(-kt)/(1-\exp(k)) \quad \text{for } t \leq 1$$
$$= 0 \text{ for } t > 1,$$

and the density of the capture distance, X, will be a truncated exponential mixture of normals,

$$f(x) = \frac{k}{1-\exp(-k)} \int_0^1 \frac{1}{\sqrt{2\pi T}} \exp(-kT - x^2/2T)dT.$$

For $k = 1$ values for $f(x)$ are given in Table 2 and $f(x)$ is plotted in Figure 2 along with the normal density with the same variance, namely, $\sigma^2 = .41802$.

For $k > 0$ the variance of X is given by $\frac{1-\exp(-k)-k\exp(-k)}{k(1-\exp(-k))}$ and the fourth central moment is given by

$$\frac{6(1-\exp(-k)-k\exp(-k)-k^2\exp(-k)/2)}{k^2(1-\exp(-k))}.$$

Table 3 gives the variance and the kurtosis of X for various values of k. It should be noticed that the kurtosis is larger for larger values of k.

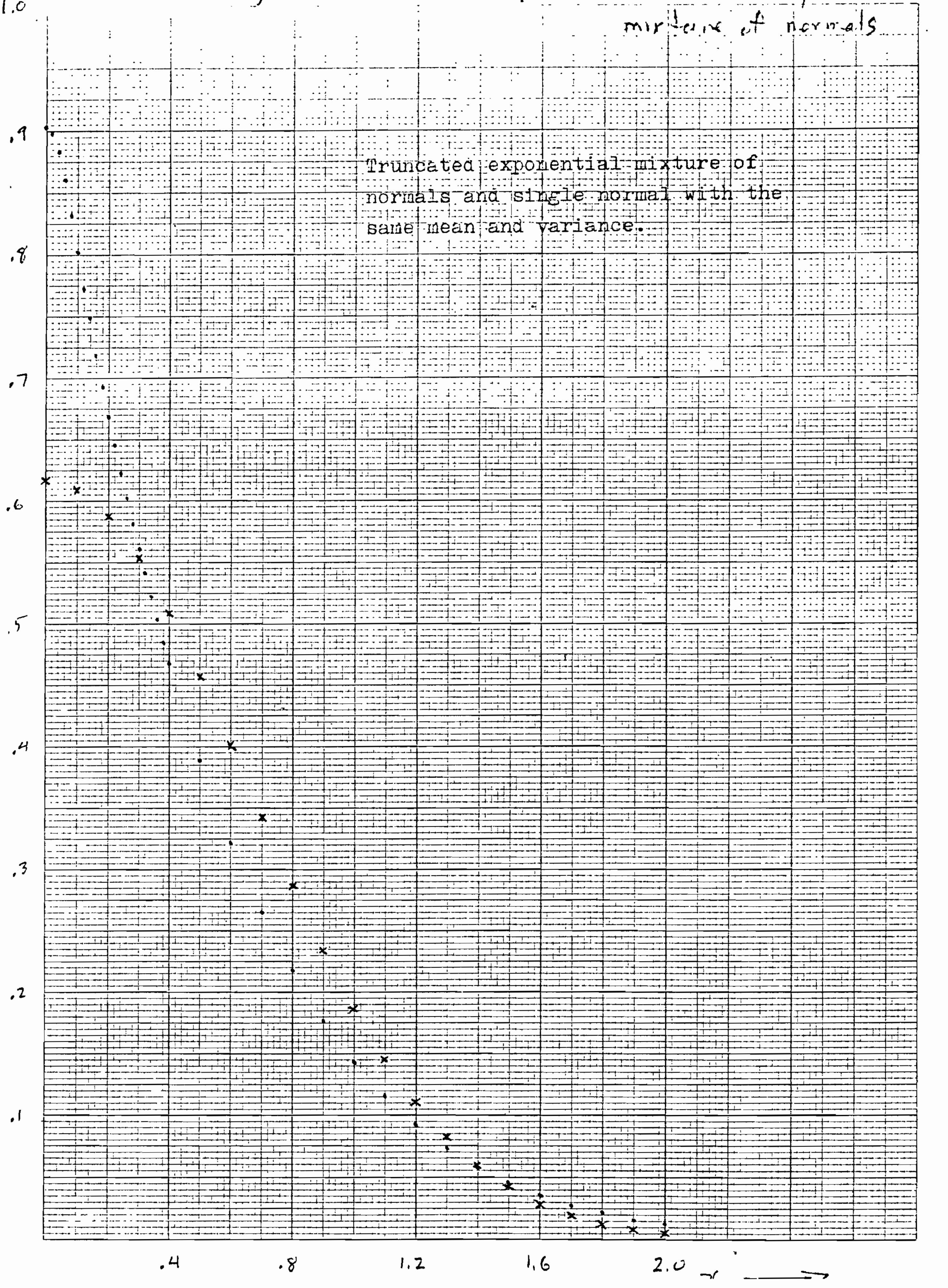
Table 2

x	f(x)	$\phi(x)$
.0	.9045	.6170
.1	.8021	.6097
.2	.6688	.5882
.3	.5609	.5541
.4	.4681	.5096
.5	.3891	.4576
.6	.3219	.4011
.7	.2650	.3434
.8	.2171	.2870
.9	.1769	.2342
1.0	.1434	.1866
1.1	.1155	.1451
1.2	.0926	.1102
1.3	.0737	.0817
1.4	.0583	.0592
1.5	.0458	.0418
1.6	.0358	.0289
1.7	.0277	.0195
1.8	.0214	.0128
1.9	.0163	.0082
2.0	.0124	.0052
2.5	.0028	.0003
3.0	.0005	.0000

f(x) = Truncated exponential mixture of normals with
$T \sim \exp(-kT)/(1 - \exp(-k))$,

$\phi(x)$ = Normal with $\mu = 0, \sigma^2 = .41802$ (same as for f(x)).

Figure 2



f(x) · Truncated exponential mixture of normals
Truncated exponential mixture of normals and single normal with the same mean and variance.
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
.4
.8
1.2
1.6
2.0

Table 3

k	Variance	Kurtosis
.0	.5	4.0
.1	.4917	4.034
.2	.4833	4.068
.3	.4750	4.103
.4	.4668	4.138
.5	.4585	4.174
.6	.4503	4.211
.7	.4421	4.248
.8	.4340	4.286
.9	.4260	4.324
1.0	.4180	4.362
1.5	.3794	4.557
2.0	.3435	4.754
2.5	.3106	4.946
3.0	.2809	5.127
3.5	.2546	5.293
4.0	.2313	5.438

Variance and kurtosis for truncated exponential mixtures of normal distributions where k is the exponential parameter in the mixing distribution of the normal variance.

DISCUSSION

The assumption that animal movements are random and can be approximated by Brownian motion is an appealing one. This assumption may be biologically reasonable in some cases and the resulting Brownian motion process is mathematically tractable. The more general case of diffusion is discussed by Pielou (1977, Chapt. 11) who mentions the examples of the spread of an invading species and the time of return of homing seabirds.

While animal movements are not likely ever to be completely random the assumption of randomness provides a convenient standard against which to compare the actual movements. The way the observed recapture locations differ from those expected under random movement will say something about how the animals actually do behave.

There is evidence that in the case of dispersal of drosophila (Dobzhansky and Wright, 1973, 1947) and peromyscus (Myton, 1974) that the kurtosis is higher than it would be if recapture locations were normally distributed. It is pointed out in this paper that if captures are made at random times within the trapping interval then the distribution of locations where animals are trapped should not show a normal distribution even if movements are random. Movements may not be random, however, and the nature of the deviation from randomness may be important.

In their work on drosophila Dobzhansky and Wright showed that dispersal was not uniform in all directions, that certain habitats (near old pines and oaks) were preferred, and that dispersal rate depended on the temperature. They suggested that the fact the distribution of recapture distances was leptokurtic was due to a mixture of flies, some of which flew a long way, other of which flew a short way. They showed that the kurtosis was highest on the first day and was lower later on. They suggested that this later decline might have been due to having many flies disperse beyond the furthest trap.

Part, but not all, of the excessive value of the kurtosis may be due to the fact that the flies enter the traps at different times and that those that enter the traps sooner will tend not to have travelled so far. This effect of mixing will be greatest for the first day, since for later days the relative differences in the amount of time the flies have been travelling will be smaller.

The variation in time until being trapped is certainly one factor which influences the distribution of recapture locations and leads to a leptokurtic distribution.

A number of people studying home ranges have considered a bivariate normal distribution for the observed trapping locations. Calhoun and Casby (1958) considered a radially symmetric (circular) bivariate normal distribution while Jennrich and Turner (1969) considered a general bivariate normal distribution. These and a number of other papers were reviewed by VanWinkle (1975). Myton's (1974) data suggest that observed recapture locations are not normally distributed, there being too many recaptures very near and very far from the center.

It is not clear what behavioral mechanism could produce a bivariate normal distribution in recapture data used in measuring home range. In dispersing animals it might be reasonable to assume random movement, in which case the locations of uncaptured animals after fixed time should have a bivariate normal distribution but recapture locations should not have a bivariate normal distribution.

Further, it is not clear what the relationship should be between the "utilization distribution" and the distribution of recapture locations for animals with home ranges. As defined by VanWinkle (1975) the utilization distribution is a measure of how much time an animal spends at each point in his range during the study period. Recapture data may be more likely to represent how much time an animal spends searching for food at a point, especially if captures are made in traps baited with food. If an animal spends a great deal of time feeding or resting at a particular site the large amount of time spent is not likely to be reflected by a correspondingly large number of recaptures at that site.

If an animal with a home range searches for food by leaving a nest or burrow and moving randomly for a fixed time, say until the end of a period suitable for feeding, and then returns directly to the nest or burrow, then the location of the animal at the time of return would have a bivariate normal distribution. However, if the animal could be trapped at various times within the foraging period, the recapture locations would not be bivariate normal but rather a mixture of such distributions. The resulting distribution would be leptokurtic, as were some of those observed by Myton (1974).

If an animal searches for food at random and returns to his nest after finding a given amount, say enough to feed young or to store, then neither the distribution of the location of the area traversed nor the point of return would be normally distributed.

It is possible that the recapture data does not accurately reflect the pattern of movement of animals. For example, animals might be more willing to enter traps encountered near their nest than ones encountered further away. Whether this is true might be determined by comparing recapture data with other measures of home range use.

Animal movements are not likely to be random and home range use is not likely to be accurately represented by a bivariate normal distribution. Assuming such a bivariate normal distribution may be useful for getting values for home range size to use in comparative studies, as the normality assumption may not be too far wrong for practical purposes, but deviations from normality may be important in themselves.

In studying home range the size of the home range is important but the pattern of land use is also important. If data suggest that the utilization distribution is not bivariate normal, this should not simply be regarded as an inconvenience because it calls into question home range area estimates based on normal theory. Attempts should be made to use the form of the utilization distribution to understand the animal's foraging behavior.

REFERENCES

Andrewartha, H. G., and Birch, L. C. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago.

Calhoun, J. B., and Casby, J. U. 1958. Calculation of home range and density of small mammals. Public Health Monograph No. 55. U. S. Dept. of Health, Education and Welfare.

Dobzhansky, T., and Wright, S. 1943. Genetics of natural populations. X. Dispersion rates in Drosophila pseudoobscura, Genetics 28: 304-40.

Dobzhansky, T., and Wright, S. 1947. Genetics of natural populations. XV. Rate of diffusion of a mutant gene through a population of Drosophila pseudoobscura, Genetics 32:303-24.

Hayne, D. W. 1949. Calculation of size of home range. J. Mammal. 30:1-18.

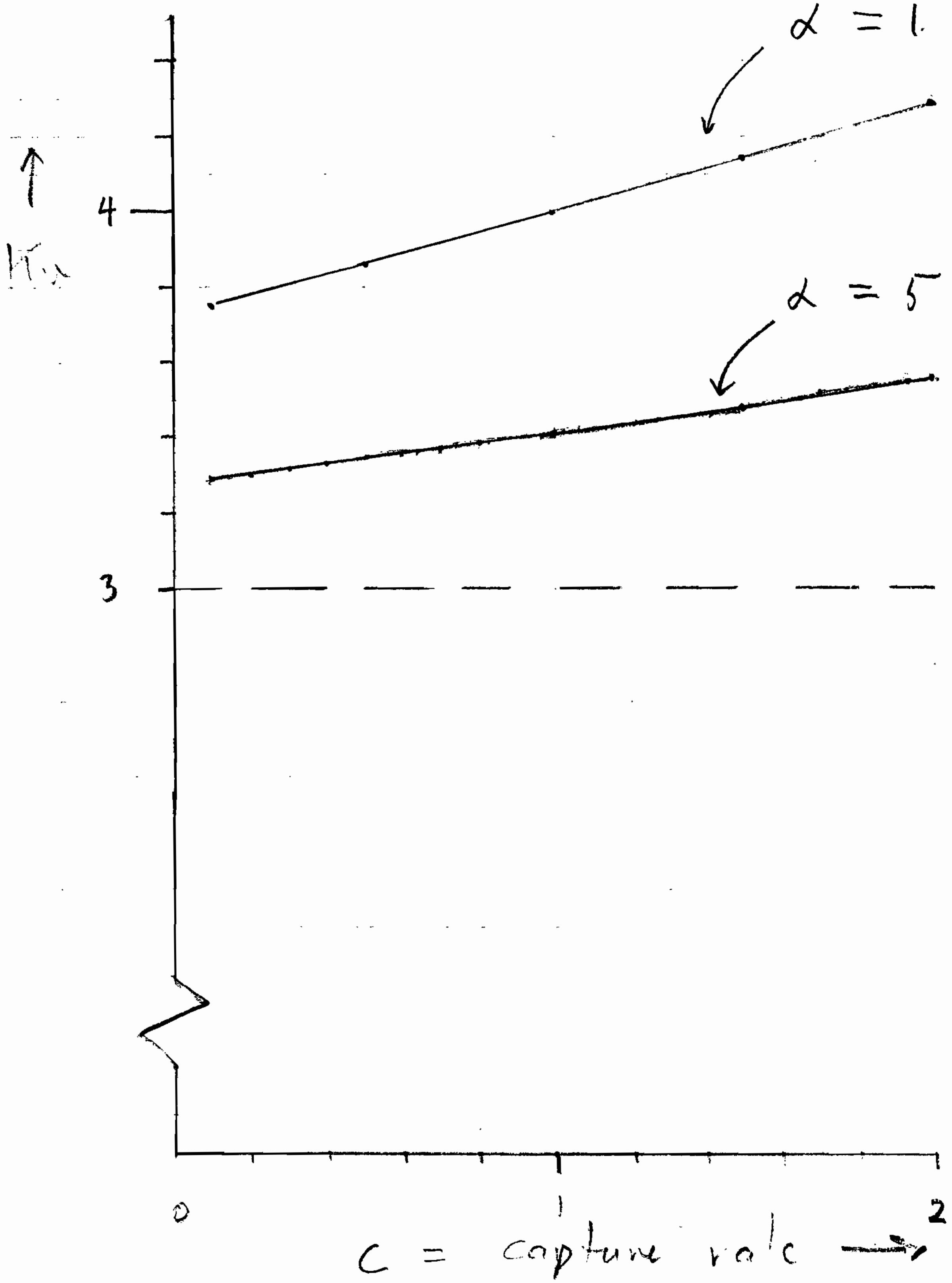
Jennrich, R. I., and Turner, F. B. 1969. Measurement of non-circular home range. J. Theoret. Biol. 22:227-37.

Myton, B. 1974. Utilization of space by Peromyscus leucopus and other small mammals. Ecology 55:277-90.

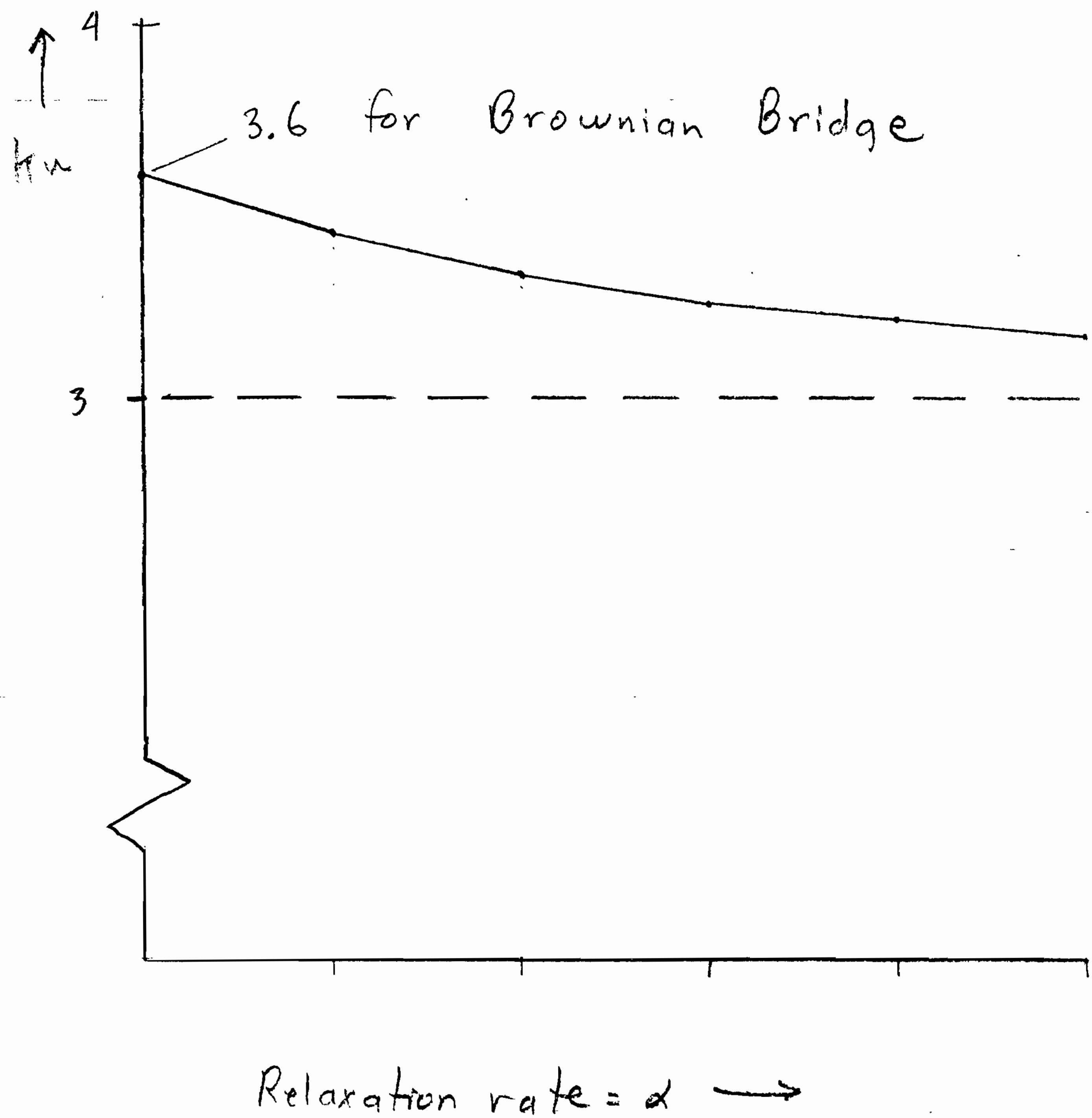
Pielou, E. C. 1977. Mathematical Ecology. John Wiley & Sons, New York.

VanWinkle, W. 1975. Comparison of several probabalistic home-range models. J. Wildl. Manage. 39:118-23.

(4) Kurtosis for OU process, exponential capture.



(5) Kurtosis for OU Bridge, uniform capture.

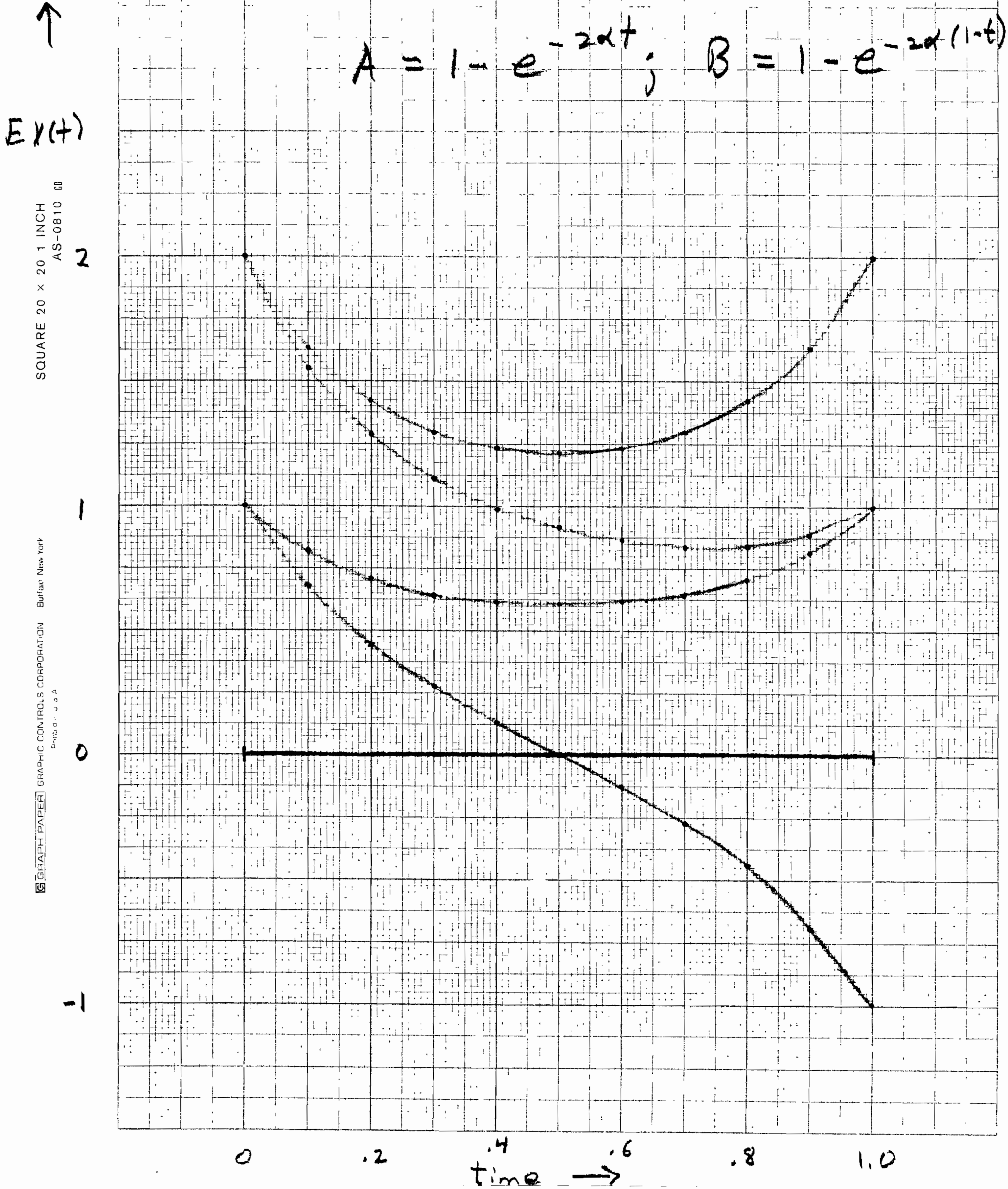


(6a) OU means : $E\,X(t) \mid X(0) = a,\ X(1) = b$

$$E\,X(t) = \frac{B e^{-\alpha t} a + A e^{-\alpha(1-t)} b}{A + B}$$

Here: $\alpha = 1$

$$A = 1 - e^{-2\alpha t}; \quad B = 1 - e^{-2\alpha(1-t)}$$

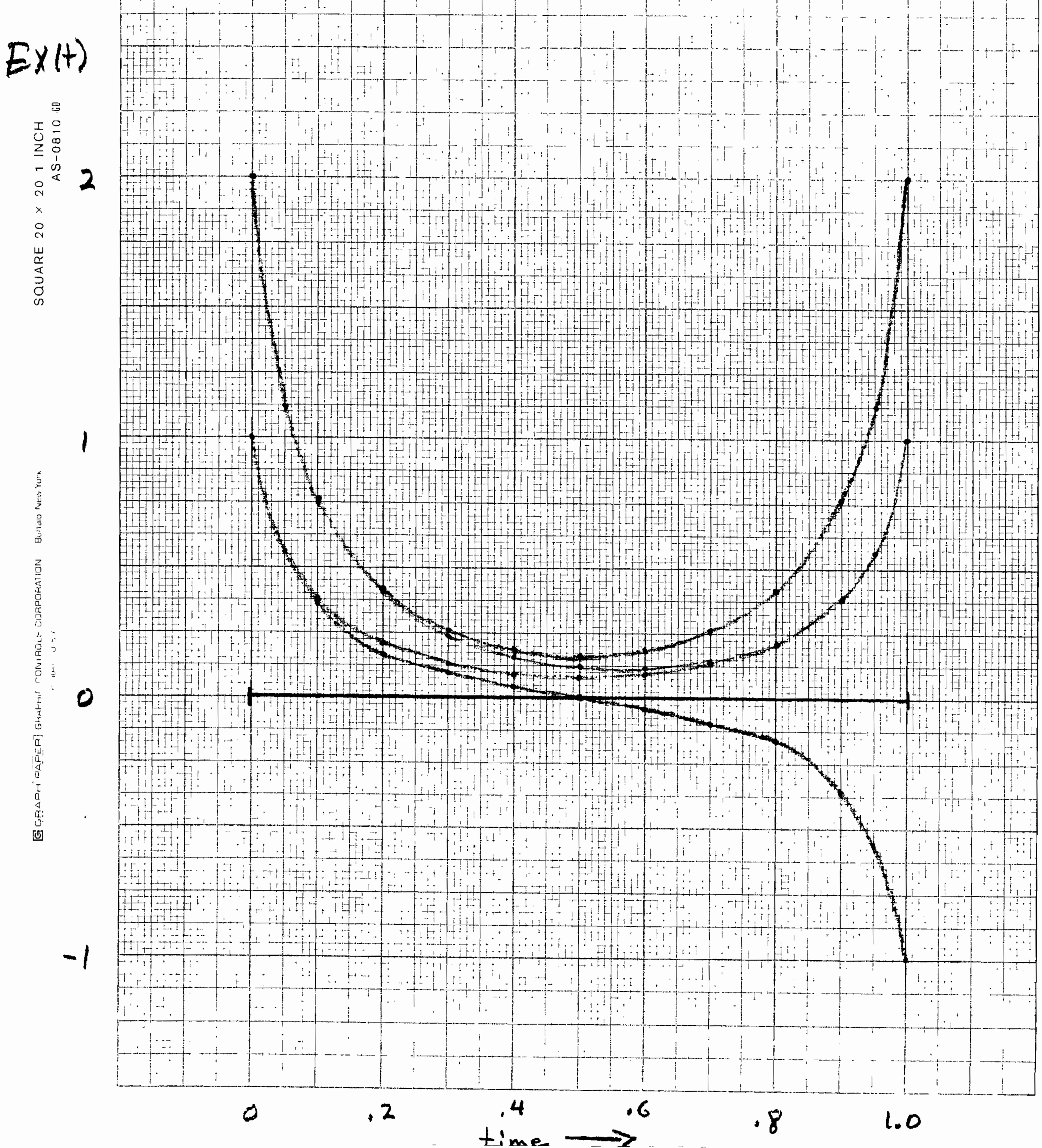


(6b) OU means: $E\,X(t) \mid X(0)=a,\ X(1)=b$

$$EX(t) = \frac{B e^{-\alpha t} a + A e^{-\alpha(1-t)} b}{A+B}$$

Here: $\alpha = 5$

$$A = 1 - e^{-2\alpha t}; \quad B = 1 - e^{-2\alpha(1-t)}$$



Variance at time = t for OU Bridge