A Statistical Theory of Cave Entrance Evolution

By RANE L. CURL

The question of the number of caves that may exist in a region or a country is one that is often raised by both laymen and scientists. In the past many estimates ranging from mere guesses to those based on empirical deductions have been proposed. To these estimates is now added a statistical approach which gives a more substantial foundation to the postulated number of caves. Rane L. Curl was introduced to cave exploration during the summer of 1952 in Charleston, W. Va. While at the Massachusetts Institute of Technology as a graduate student he was chairman of the Boston Chapter of the National Speleological Society. At present Dr. Curl is on the research staff of Shell Development Company in Emeryville, California.

INTRODUCTION

The natural creation and destruction of cave entrances are geological phenomena depending on the stratigraphic, meteorologic, hydrologic, mineralogic, etc., relations of a region in which caves occur. The existence of an entrance to a cave suggests from a statistical point of view that there is a *process* of entrance creation. The discovery of second entrances from within known caves suggests that there is a complimentary *process* of entrance destruction. It follows that there must be caves which have either not gained or have lost entrances and which are therefore unknown.

This conjecture is supported by the discovery of new caves in excavating and mining and by the appearance of a karst area having sinkholes *apparently* without traversable caves at their bases. The most common estimate is that there are ten times as many undiscovered as there are known caves. Folsom (1956) refers to approximately five thousand known caves in the United States and suggests upwards of fifty thousand unknown caves. The implication is often that these are mostly unreported caves with entrances, though this notion is not necessary and is indeed unlikely.

Supporting evidence of a logical nature is presented here for these otherwise empirical conclusions. Originally this study began after it had been noticed that since there are caves with one, two, and more entrances, there logically should be caves with no entrances. Such a notion found immediate support from the evidence mentioned above, but intriguing questions remained with regard to the number of such caves and their properties. It seemed that it should be possible to use data on the number of caves with one, two, or more entrances, and to *extrapolate* to find the number without entrances. This is, in essence, what has been done in this paper, although there is now a great difference between the first crude attempts and the present analysis based upon hypotheses concerning entrance evolution.

It was evident from the beginning that a mathematical relation would have to be "fitted" to the available information, in order to carry through an extrapolation. The main question was how this should be done. Although the earliest arbitrary functions considered gave an answer, they left the impression that the result was essentially meaningless. By what criterion was the extrapolation to be considered realistic? An example of the result of such an extrapolation, based upon an almost arbitrary mathematical relation is shown in Table 1. The mathematical relation in this case is called the *Poisson frequency distribution* (Cramer, 1955). It is a formula describing the manner in which items (caves) are sorted into different groups (groups of caves, each for a different number of entrances). The mathematical relation was fitted to the *observed frequencies* and then the expected frequency of caves with no entrances was calculated.

This use of an arbitrary though well known statistical distribution seems insufficient to provide *understanding* of the processes which may have occurred in nature to produce the observed distribution of caves by number of entrances. In addition, the characteristic parameter of the

TABLE 1.

Comparisons of Observed Frequencies of Caves of West Virginia with the Poisson Distribution Giving the "Best Fit" to Data

Number of entrances	0	1	2	3	4	5
Observed frequency	—	228	25	3	1	0
Expected frequency	893	226	29	2	0	0

Poisson formula was found to vary with the average length of the group of caves for which it was computed and hence it lacked the desirable feature of being a fundamental parameter of *all* the caves. Furthermore, other equations could have been fitted and there would have been little reason to prefer one in particular.

Following this first approach a simple mechanism of entrance evolution was assumed. The variable of cave length was not included and the resulting mathematical equation turned out to be the same Poisson formula which had been used first on intuitive grounds. Thus, Table I also represents the "fit" of a theoretical method based upon a simple mechanism. Two things were then observed which led ultimately to the theory in its present form. Firstly, it was noticed that the Poisson formula predicts fewer caves with multiple entrances than had been observed. Secondly, and more important, it was noticed that the multi-entrance caves are, in general, *longer caves* than those with one entrance. It thus became apparent that length was an important variable. As a result, the complete theory leads to a method of predicting the relative lengths of unknown caves as well as their number.

Statistical Theory of Entrances

We will not speak here of the peculiar histories and properties of individual caves. If we did, we would find that each probably could be argued as having some aspect which seems to place it out of the realm of applicability of a statistical theory. There would still remain, however, the group phenomena — properties of the whole ensemble of caves which cannot be attached to the individual members of the group. The *average number of entrances* is clearly such a property as are also the probability relations developed here. Although we will not be able to indicate what may happen to any given cave, we will speak of what is likely to occur to a group of caves with the passage of time.

Consider a large number of caves distributed over an area of relatively uniform geology, climate, and topography. A continuing *random* or *stochastic process* of formation or reopening of entrances and closure of entrances will be taking place. The agents of these processes are, in general, associated with surface erosion phenomena and, to a lesser extent, with the nature of the caves in question. The assumption will be made here that this process is taking place on a fixed population of caves of invariant lengths. The truth of this depends upon the relative time scales of external and internal changes. The assumption implies that the genesis of cave entrances is much more rapid than the genesis of internal cavern features. This is likely as the surface, where entrances form, is subject to a much harsher environment, involving more rapid chemical and physical weathering.

We start, then, with a population of caves of different lengths and with zero, one, or more entrances, at some instant of time. As time passes, the initial conditions of these caves with respect to numbers of entrances will change. In some given length of time, a few caves will lose entrances, a few will gain entrances and the rest will remain as they were. Let us consider each of these processes in turn, in order to deduce the logical form of a stochastic theory of entrance evolution.

In a large number of caves with one entrance, a certain fraction of the entrances will be closed in some interval of time. It is reasonable to say that in double the time interval about twice the number of caves will lose an entrance, as long as in either case the total number losing entrances is small compared to the total number of caves. Furthermore, if all of this group of caves had two entrances, it is apparent that the number of caves which would *lose an entrance* would be about twice that for single entrance caves, in the same interval of time. This presupposes that all present entrances are equally likely to close, an assumption which might not be quite true, but which will be taken as a first approximation. We therefore state the following hypothesis:

The probability (or fraction) of caves with n entrances losing one entrance in a time interval Δt is equal to $sn\Delta t$.

The proportionality constant s is a function of the influences of the environment.

As the length of a cave increases, it becomes ever more likely to possess those features which are conducive to entrance formation. While it may not be proper to say that the formation of an entrance is equally likely along every foot of a cave — we usually observe entrances at the terminations of caves — it is reasonable to assume that some related properties associated with entrances such as joint intersections, approaches to the surface, length relative to topographical scale, and others, are functions of the length. Therefore we might expect that on the average caves of twice the length would be twice as likely to gain an entrance in some time interval. This likelihood would also be proportional to the time interval, as in the case of losing entrances. We therefore state the following hypothesis:

The probability of caves of length l gaining an entrance in a time interval Δt is equal to $rl\Delta t$.

The proportionality constant r again depends on the environment.

It is possible that the existence of an initial entrance influences the creation of subsequent entrances, but if the new entrances are reasonably separated from the first, this effect should be negligible. In this analysis, such an effect will be assumed to be absent.

The remaining possible event of caves neither gaining nor losing an entrance follows from the previous two.

The probability of caves of length l and with n entrances neither gaining or losing an entrance in a time interval Δt is equal to $1\text{-}ns\Delta t\text{-}rl\Delta t$.

These relations define the mechanisms involved. It is now necessary to consider how they act in order to create the existing distribution of cave entrances among caves.

At some instant of time, t, we would find that there exists in a population of caves a certain number, f_n, with n entrances; another number, f_{n-1} with n-1 entrances, and still another number of caves, f_{n+1} with $n+1$ entrances. We are interested in determining how many caves with n entrances there will be after some time interval Δt. Instead of discussing the actual number of caves corresponding to a certain num- of entrances, we may speak of the probability of caves having n, n-1, or $n+1$ entrances. If the interval of time Δt is so short that the event of a cave gaining or losing two or more entrances is extremely small compared with gaining or losing only one entrance, it follows that the probability of obtaining caves which are of length l and have n entrances at time $t + \Delta t$, $p(l,n,t+\Delta t)$, will depend upon:

(a) The probability of caves of length l with n-1 entrances at time t, $p(l,n\text{-}1,t)$, *and* the probability of such caves gaining an entrance, $rl\Delta t$,

(b) The probability of caves of length l with $n+1$ entrances at time t, $p(l,n+1,t)$, *and* the probability of these caves losing an entrance, $s(n+1)\Delta t$, and

(c) The probability of caves of length l with n entrances at time t, $p(l,n,t)$, *and* the probability of these caves neither gaining or losing an entrance, $1\text{-}rl\Delta t\text{-}sn\Delta t$.

Further mathematical development may be found in the appendix. It suffices for this discussion to state that an equation may be obtained which describes how the probability of having caves with n entrances varies with time, if we know the distribution of lengths. It may be solved numerically if we are given the lengths of all the caves and the initial distribution of entrances among the caves. It would be found that the probability of finding caves with n entrances would approach, in time, a value depending only on n, r, and s, and independent of time. This would be a mature population of caves in regard to entrance development and we would speak of the process as then being in *statistical equilibrium.*

When statistical equilibrium is attained, the probability distribution of entrances does not change with time. This does not mean that the opening and closing of entrances has ceased or slowed down, but rather that they appear and disappear at such relative rates that the probable number of caves with some number of entrances remains fixed. The actual number will vary about this value as a mean. For example, if we look at the cave population at different times, we may find that sometimes there are no caves with three entrances, and at other times there are one, two, or more caves with three entrances. The data we have on the number of caves with one, two, or more entrances is only one of a multitude of different possible arrangements. The time process occurs on an extremely large time scale for which reason we are not conscious of these fluctuations. It is possible to calculate the relative frequency with which other arrangements would occur and thereby obtain an estimate of the variation of f_n about the mean over a long time. These estimates are given in Tables 2 and 3 and are discussed in more detail in the section dealing with results and conclusions.

Since the probabilities do not change with time at equilibrium, the parameters which determine the distribution should also be independent of time. In the statistical theory, the only parameters which enter are r, the probability *per foot of cave per year* of an entrance forming; and s, the probability *per entrance per year* of an entrance closing. The only combination of these parameters which does not involve units of time, as required by the condition of statistical equilibrium, is the ratio of r to s. This ratio will be denoted by the Greek letter lambda (λ),

$$\lambda = r/s.$$

Since r and s are functions of the environment, λ is also. We may define as a *homogeneous region* a region encompassing a population of caves which have entrances developing by a process in which λ is everywhere constant. In this study, λ has been evaluated separately for caves in different length groups in the states of West Virginia and Pennsylvania. A comparison of the values obtained indicates to what extent the areas under consideration are subject to similar conditions with regard to cave entrance development.

While λ is defined in terms of the entrance genesis phenomenon, it would be expected that it is also related to the processes involved in the internal development of caves. If λ is constant similar internal cavern features might also be expected within the area.

The statistical theory may be used to correlate data. If we take data on the number of caves with one or more entrances and on their lengths, the theoretical equation may be "fitted" to obtain the value of λ which yields the closest agreement between observation and theory. In making this "fit", it is not necessary to try to make the theoretical form "pass through" all the data points. Only one constant is available for manipulation (λ), hence only one constant obtained from the observations may be used. This number from the data is the average number of entrances per cave for all caves in an area with one or more entrances and will be given the symbol alpha (α).

The quantity λ, determined from α and cave lengths, may be used to calculate the number of caves in the same area which have, at present, no entrance. In addition, the length distribution for caves with two or more entrances may be computed and compared with observations, to obtain a check on the theory. Finally, we may calculate the length distribution of unknown caves and obtain in this way the probable number of such caves longer than any given length.

Selection of Data

Because of the relative rarity of multi-entrance caves, it is necessary to study a very large population in order to obtain meaningful results from the statistical analysis. Only two sources were found to be satisfactory.

As regards the number of detailed reports of caves, the best reference was for the State of West Virginia (Davies, 1949). The second source, with fewer caves reported, was for Pennsylvania (Stone, 1953). The former describes about 400 caves of which the information on 257 was found suitable according to certain selection rules. Of the 272 caves described in Pennsylvania, 110 were used.

Even what we are to consider a cave depends on the size of the explorer, the length of enclosed passage, knowledge of geological relations, and the extent of exploration (which may make two caves into one with two entrances). For this study, the following selection rules were used.

Criteria for selecting caves

1. The cave is reported and described in either the West Virginia or Pennsylvania reference.
2. The cave occurs in limestone.
3. Sufficient data is given to ascertain the length of the cave and its number of entrances.
4. The cave's total length is greater than 50 feet for Pennsylvania and 100 feet for West Virginia.

All the caves in each state were considered to lie in homogeneous regions. While it may be argued that this is not strictly true, it was a necessary assumption, in order to have a sufficient number of caves upon which to conduct a statistical analysis. As a result, the possible variations in λ within each state are averaged in the final value obtained, according to the contribution by each county (or homogeneous region) to the total number of caves.

The study was restricted to caves occurring in limestone. Inclusion of the few caves in sandstone and other rocks would have had little effect on the results, but for the sake of being consistent with regard to geological environment they were excluded. Some reports were found of "numerous caves" in certain areas. These and similar indefinite references were ignored.

As shorter and shorter caves are considered, some question arises as to whether they are or are not caves. In reading collected descriptions, one gets the impression of decisions to omit caves which are "short" by some indefinite standard. It is reasonable that this be done, but in order to have an accurate probability distribution for lengths, some lower limit must be placed on the lengths of caves to be included. The calculated value of λ should not depend on this limit, but the calculated number of undiscovered caves will. In the case of West Virginia, the decision was made to include only those caves which are longer than 100 feet. In Pennsylvania, where there are fewer caves and the indication is that short caves have been retained in reports, the limit was placed at 50 feet. This "rule" accounts for most of the caves which were not included.

Criteria for determining length

1. Only horizontal distances are considered.
2. Total traversable length is summed, including different levels and parallel passages.

If a total length was unambiguously stated, it was used without question. In some caves, however, the final length came about by a process of interpretation, judgment, and addition, applied to each cave. The judgment used may have introduced errors, but, because even the best reported lengths are probably accurate within 10% and reported values like "300 yards" represent even greater probability of error, it was not thought of value to refine the process.

Criteria for counting entrances

1. The entrance is natural.
2. The entrance is large enough to permit entry by an adult human being.
3. The entrance exists now or is reported to have existed, but it is now closed as a result of human activity and the cave meets all other conditions for acceptability.

Caves in quarries were excluded. They are, indeed, to be classed with the zero entrance caves. Quite a few caves have been reported in quarries in Pennsylvania and relatively few in West Virginia. Possibly this may be because of greater quarrying activity in Pennsylvania. It is also not known where entrances may have existed prior to the quarrying operation. Unless reported otherwise, they were assumed not to have existed. Entrances produced by roadcuts and other construction work fall, of course, in the same category as caves in quarries.

A vague dividing line occurs for caves which were first entered by some excavating — perhaps such a trivial thing as dislodging a small stone. By the formal definition, the latter would be an *entranceless* cave. If necessary, the definition could be modified to allow some small excavating work such as that which can be done only by hand. However, since such a fact is seldom reported, some caves are probably incorrectly included. For the purposes of the present computations, it was assumed that reporting was accurate, and caves were accepted as reported.

After the data suitable for use in the computations had been selected, the calculations were performed. This involved the determination of the average number of entrances for known caves, separate numerical computations on the data of the 367 caves tabulated, and subsequent operations to obtain the desired numbers. The magnitude of this task would have made hand computation lengthy and subject to many errors. The computations were programmed for an electronic digital computer and run with the data on caves of West Virginia and Pennsylvania in groups having different lower lengths.

Results and Conclusions

In Tables 2 and 3 are shown some of the results of the calculations. Included are the observed and theoretical frequencies of caves with n entrances, an estimate of the standard errors of the predictions, the value of λ determined for different length groups and the comparable results obtained by use of the Poisson distribution. The values of expected frequencies are given to the nearest tenth, though they were calculated to more places.

Table 2. Results of Calculations for Caves of West Virginia
Constants and Frequencies of Caves with and without Entrances

Caves longer than: (feet)	Number (M)	Average number entrances (α)	λ	Number of entrances (n)	Observed frequency (f_n)	Expected frequency (m_n)	Poisson Distribution	
							λ*	Expected frequency (f_n*)
100	257	1.13	0.00032				0.25	
				0	(10)	2405 ± 395		893 ± 146
				1	228	228.7 ± 15		226 ± 15
				2	25	23.7 ± 4.9		28.5 ± 5.3
				3	3	3.6 ± 1.9		2.4 ± 1.6
				4	1	0.7 ± 0.9		0.1 ± 0.4
				5	0	0.2 ± 0.4		0.0 ± 0.1
500	124	1.24	0.00033				0.45	
				0	-	350 ± 59		218 ± 37
				1	99	99.9 ± 10		98.2 ± 9.9
				2	21	19.3 ± 4.4		22.1 ± 4.7
				3	3	3.7 ± 1.9		3.3 ± 1.8
				4	1	0.8 ± 0.9		0.4 ± 0.6
				5	0	0.2 ± 0.4		0.0 ± 0.2
1000	64	1.33	0.00031				0.60	
				0	-	104 ± 21		78 ± 15
				1	48	47.6 ± 6.9		46.8 ± 6.8
				2	12	12.8 ± 3.6		14.0 ± 3.7
				3	3	2.8 ± 1.7		2.8 ± 1.7
				4	1	0.6 ± 0.8		0.4 ± 0.6
				5	0	0.1 ± 0.3		0.0 ± 0.2

Table 3. Results of Calculations for Caves of Pennsylvania
Constants and Frequencies of Caves with and without Entrances

Caves longer than: (feet)	Number (M)	Average number entrances (α)	λ	Number of entrances (n)	Observed frequency (f_n)	Expected frequency (m_n)	Poisson Distribution	
							λ*	Expected frequency (f_n*)
50	110	1.12	0.00035				0.23	
				0	(55)	2109 ± 562		431 ± 115
				1	99	99.3 ± 10		98.0 ± 9.9
				2	9	8.8 ± 3.0		11.1 ± 3.3
				3	2	1.5 ± 1.2		0.8 ± 0.9
				4	0	0.0 ± 0.2		0.0 ± 0.2
				5	0	0.0 ± 0.0		0.0 ± 0.0
100	78	1.15	0.00034				0.31	
				0	(50)	701 ± 208		212 ± 63
				1	68	68.1 ± 8.2		66.4 ± 8.2
				2	8	8.2 ± 2.9		10.4 ± 3.2
				3	2	1.4 ± 1.2		1.1 ± 1.0
				4	0	0.3 ± 0.5		0.1 ± 0.3
				5	0	0.0 ± 0.2		0.0 ± 0.1
500	33	1.09	0.00013				0.18	
				0	-	238 ± 135		167 ± 95
				1	31	30.2 ± 5.5		30.1 ± 5.5
				2	1	2.6 ± 1.6		2.7 ± 1.6
				3	1	0.2 ± 0.4		0.2 ± 0.4
				4	0	0.0 ± 0.1		0.0 ± 0.1
				5	0	0.0 ± 0.0		0.0 ± 0.0
1000	21	1.14	0.00016				0.27	
				0	-	79 ± 44		67 ± 37
				1	19	18.3 ± 4.3		18.3 ± 4.3
				2	1	2.4 ± 1.6		2.5 ± 1.6
				3	1	0.3 ± 0.5		0.2 ± 0.5
				4	0	0.0 ± 0.1		0.0 ± 0.1
				5	0	0.0 ± 0.0		0.0 ± 0.0

Homogeneity of West Virginia and Pennsylvania

The data for caves of the two states were calculated separately for a number of lower limits on the length. This, in effect, separated the caves into a number of groups, representing different areas within each state. As the lower limit was raised, the number of included caves decreased.

If the assumption of homogeneity made in the analysis is not justified, it would show up in a variation of λ between the groups. It is an important result that such has not been found to be the case. In fact, the constancy of the parameter λ *between* the states as well as *within* the states is quite significant. We may conclude that the caves in these states are located in essentially the same geologic and climatic region; a result agreeing with simple observation.

The deviation of λ in the two upper length groups of Pennsylvania may be due to either inhomogeneity or to insufficient data. The numbers of caves represented in these groups have decreased to 31 and 21 respectively and results obtained from such limited data is subject to error.

Now λl is the ratio, for a cave of length l, of the probability of gaining an entrance to the probability of losing an entrance. With the value of λ given in Tables 2 and 3, the product λl would have the value $\lambda l = 1$ for a cave approximately 3000 feet long. Since most caves are shorter than this, it is a general result that caves are more likely to lose their entrances than to gain entrances. Of course, the calculated value of λ is a consequence of this fact. The destruction of a cave entrance might be looked upon as the "natural" event, in that almost any collapse, slip, silting, etc., tend to fill holes which might otherwise be traversable, and it is the rarer case when these actions not only disclose an aperture, but leave its entire length traversable to a human being.

Distribution of Cave Entrances

The agreement in Tables 2 and 3 between observed and calculated frequencies of caves with different numbers of entrances is excellent. One cannot say, though, that it is *appreciably* less excellent for the ordinary Poisson distribution. In both cases λ was calculated so as to have the best agreement, but the only value that was needed from the data on entrances was the average number of entrances for caves with one or more entrances. However, based solely on the Poisson distribution, an extrapolation to find the number of caves with no entrances is unconvincing. With the statistical theory to indicate a mechanism accounting for the existence of such caves, the extrapolation may be performed with some confidence. This fact and the agreement between theory and fact support the hypotheses in the theory.

Standard errors are given in Tables 2 and 3 for the expected frequencies. If we had an *a priori* value of λ for a population of caves, we would calculate the expected frequency distributions shown in the tables. If we inspected the population, we would find that there would be different frequencies than we had calculated. This has already been explained in terms of the fluctuations about the "probable" frequencies which would occur with time in a population in statistical equilibrium. The observed differences between the expected and observed frequencies are well within the standard errors so we are given by this test no reason to doubt the proposed mechanism.

Standard errors are given also for the values of $m_{\bar{0}}$. If the theory is reasonably correct, the probability that m_0 differs from its predicted value in either direction by more than the given standard error is equal to 0.32. That it would differ by more than twice as much has a probability of 0.05.

The frequencies of "observed" caves with no entrances, given in Tables 2 and 3, refer to caves in quarries and roadcuts. It is interesting that of the eleven commercial caves reported in Pennsylvania, four were found during quarrying operations and one was found while grading for a highway. There seems little doubt that there are many more extensive caves which would be uncovered if such operations were carried on in other locations.

The ratio of unknown caves to known caves depends upon the length group considered. It is apparent from the theory that most very long caves would be likely to have an entrance. In the present cases, there are predicted to be about 20 times as many caves without entrances as with for caves over 50 feet long. For caves over 100 feet long the ratio is about 10, and for those over 1000 feet about 2. If we consider the 100 foot cave as the reasonable lower limit of our general

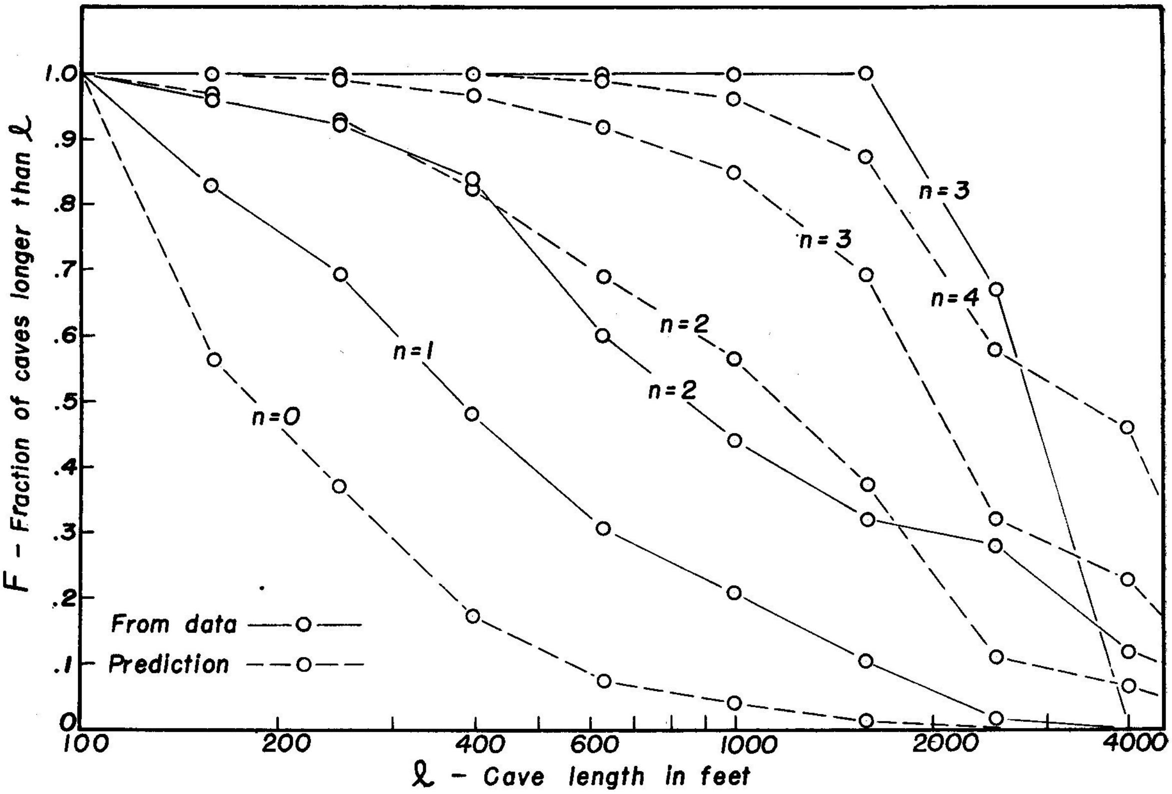

Figure 1
Distribution functions of length for caves of West Virginia over 100 feet long, according to number of entrances. Comparisons of observed and predicted distributions. The lines are shown only to connect related points.

concern about the existence or non-existence of sealed caves, the calculated ratio of 10 unknown caves to every known cave is comparable to the estimates which have been made by those familiar with caves and cave regions (Folsom, 1956).

Distribution of Cave Lengths

A mathematical result of the statistical theory is that it is possible to calculate, starting with the observed fraction of caves with one entrance whose lengths are greater than any given value, the same fraction for caves with zero, two or more entrances. This is a very useful result. Not only can we estimate the number of caves without an entrance, but we can also estimate their lengths. Figures 1 and 2 present the results of these calculations. In Figures 1 and 2 the fraction of caves longer than 100 feet, with lengths greater than some given length, is plotted versus cave length. This relation is called a ***distribution*** *function of lengths.* The logarithmic abscissa is used to contract the extent of the graph in the region of great length.

Because of the method of computation, the curves are not smooth. In order to simplify the calculations, only eleven values of length were used to find these distribution functions. Little would have been gained by a finer division.

The observed distribution functions for caves with one entrance in West Virginia and Pennsylvania are quite similar. The median length of caves over 100 feet long with one entrance is about 380 feet, and the average length about 680 feet. Using the single entrance distribution functions for caves of both states, the distribution functions were computed for caves with two entrances. In Figure 1 is shown the result for the caves of West Virginia. Quite good agreement is found between the observed and predicted length distributions for caves with two entrances. This result gives additional confirmation to the hypotheses in the statistical theory, since the prediction is based only on the data for caves with one entrance and is independent of λ, and hence of a. The length predictions, furthermore, did not require that a "fit" be made to data.

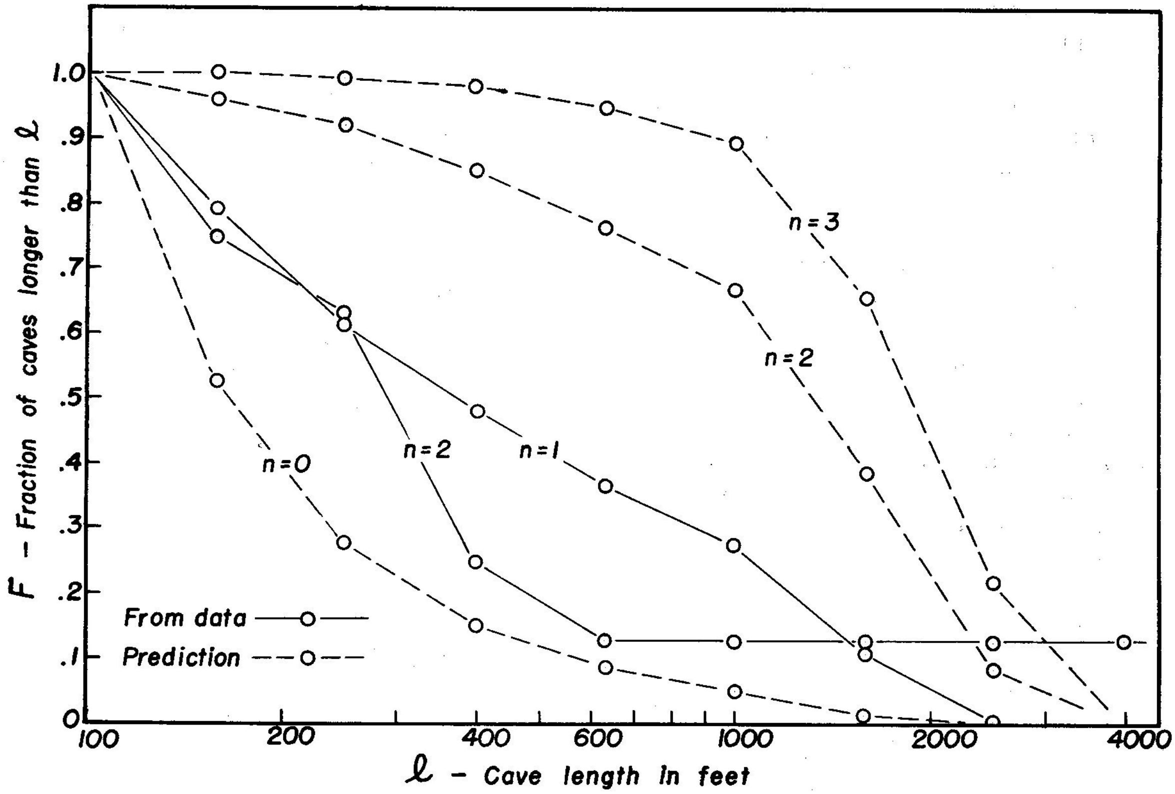

Figure 2
Distribution functions of lengths for caves of Pennsylvania over 100 feet long, according to number of entrances. Comparisons of observed and predicted distributions. The lines are shown only to connect related points.

The agreement between theory and observation for caves with three entrances in West Virginia is poorer. Only three caves in West Virginia are reported to have three natural entrances (Steeles Cave, Monroe County, 1700 feet; Mystic Cave, Pendleton County, 3700 feet; and Snedegars Cave, Pocahontas County, 3100 feet). It would be extremely fortuitous for their lengths to be just right to be correlated by the theoretical predictions. As it is there *is* agreement in the fact that the theory predicts that caves with three entrances should be long caves on the average, which is what is observed. The predicted average length of 2500 feet agrees well with the observed 2800 foot average length.

The predicted distribution function for caves with four entrances is included in Figure 1. Since there is only one such cave reported in West Virginia (Greenville Saltpeter Cave, Monroe County: over 13,000 feet), it is not possible to make a valid comparison except to note that the one four-entrance cave is also an exceptionally long cave.

The major source of error in extrapolating to n larger than one is the emphasis on the longer caves in forming these distribution functions. As n increases, the calculated function depends upon fewer and fewer caves and acquires the ills associated with working with limited data.

In West Virginia the longest cave with one entrance was reported to be Patton Cave, Monroe County, 5500 feet. Since the distribution function from which numerical calculations were made thus became zero at 5500 feet, so did all derived distribution functions. It would have been possible to fit some best equation to the observed distribution before performing the mathematical operations, but it was thought that this would distort the significance of the results. Instead it is concluded that the predictions at great length are probably in error.

The prediction of the length distribution function for caves with $n = 0$, depending most strongly on a large number of shorter caves having lengths known with greater accuracy, may be expected to be more accurate. We can conclude from an examination of Figure 1 that caves which at present have no entrances are short caves on the average. Their median length is approximately 175 feet and their average length about 290 feet. Since this plot gives the fraction of caves longer than some length, and we have an estimate from Table 2 of the number of caves without entrances, we can estimate the number of caves without entrances which are longer than some given length. For example, the probable number of unknown caves in West Virginia which are longer than 1000 feet is about 96.

Figure 2 presents the equivalent results for Pennsylvania. The conclusions are similar except that in Figure 2 there is no agreement between the predicted and observed distributions for $n = 2$. There appears to be an unusually large proportion of short caves with two entrances. Eight such caves were counted, and there was doubt whether some of them should have been included. The report on Dump Cave, Mifflin County; 110 feet, the shortest, leaves some question as to the history and condition of its second entrance. Even more questionable is the inclusion of Baker Cave No. 2, Franklin County; 150 feet. Excluding some of the doubtful cases from the two-entrance tabulation would improve the agreement between observation and prediction, but to do so would contribute no new knowledge — therefore the results are presented as originally obtained.

The Pennsylvania caves with three entrances are too few in number — Auchenbaugh Cave, Franklin County; 300 feet; and Bear Cave, Westmoreland County; 4000 feet — to allow comparison except to note that the average length of these caves is 2100 feet and the predicted average length is 2200 feet. This is fortuitous but agreeable.

The probable number of unknown caves over 1000 feet long in Pennsylvania is about 35.

Rate of Entrance Development

No information is available from this analysis to suggest the magnitude of the parameters r or s. These parameters depend on the time process and can be obtained only from records of the rate of entrance formation or closure. There may be some reason to believe that the present rate of these processes is different from that in the near past. There is the possibility of a very rapid rate having established statistical equilibrium with subsequent relatively static conditions. Both r and s could have varied together in the past, without the fact being evident because λ is a ratio of r to s.

In the state of West Virginia not one entrance was reported to have opened naturally in the period of human record. One cave of known length was reported to have closed (Mitchell Cave, Monroe County, 2000 feet), presumably by natural causes. This is inadequate data on which to base an estimate of r and s, but proceeding regardless, assuming a fifty year record, we find that the probability of an entrance closing in one year is about 0.00007. This is equivalent to a "life-time" of an entrance in the order of 14,500 years. This estimate is probably conservative with the present climatic conditions. However, under weather conditions such as occur immediately after glacial periods, the entrance life-time may have well been this value, or even less. Similarly, from the value of λ and r, we find that the probability of an entrance forming in one year per thousand feet of cave is of the order 0.00002 and the equivalent life-time of a 1000 foot cave without an entrance is of the order of 45,000 years. A hundred foot cave would have a life-time ten times as great. As implied at the beginning of this set of estimates, caution is required when stating values for r and s, based on the single available datum. However, the values estimated are not unreasonable and they suggest that entrance development is relatively more rapid than the development of the caves themselves; a hypothesis which was basic to the statistical theory.

Properties of Caves Without Entrances

It has been suggested here that undiscovered caves are, in general, caves without entrances. This would only be exactly true in an area which has been thoroughly explored. Although it is probable that the reporting of what we consider as caves with entrances is incomplete, it is not believed that further scouting and discovery will find anywhere near as many caves as, for example, the 2400 in West Virginia which are predicted to have no entrances. Furthermore, the discovery of new caves *with* entrances would only change the relative results. If further caves are found *without excavating* they should have much the same distribution of number of entrances, lengths, etc., as the presently known caves if there is no relation between length or number of entrances and the ease of discovery. This latter is true where the limitation on discovery comes about because of limitations on the areas searched rather than the ease of discovery. It is suggested that this is usually the case; new caves in most known cave areas commonly require some excavation. With the discovery of additional caves *not requiring excavation* the results of this paper would be scaled upward by the ratio of the total caves to the number now known. Any new caves discovered that require some excavating to enter would be examples of what this paper considers as *entranceless caves* and hence would not modify the numerical results.*

The caves without entrances have the length distribution shown in Figures 1 and 2. Other properties of such caves should, on the average, be the same as those of all the known caves, except for those aspects which *depend on having an entrance traversable by man.* An entrance of such a minimum size, or larger, will admit large animals and, depending upon the cave configuration, contribute some effect to the humidity, temperature, etc., of the cave interior. These in turn affect the environment for cave fauna and flora. Many of the caves which are considered *entranceless* by the criteria of this paper would still allow appreciable access to smaller animals and air and hence would be part of the same internal environment family as caves with entrances. As the cave entrance becomes smaller the internal environment will change, on the average, toward higher humidity, perhaps higher temperature, less illumination in the "entrance" passages, poorer air circulation, etc. These changes may, for example, change the pattern of flowstone deposition, make an interior room unsuitable for bat habitation, or lead to modifications (or exclusion) of life in the twilight zone of the cave. As the sealing of caves becomes more complete, these effects become more pronounced. A completely sealed cave with stagnant air may be expected to present a much altered environment for many cavern features from that found in presently accessible caves. The full range of extent of closure will be found among what are here called entranceless caves.

The statistical theory unfortunately contributes nothing toward determining the location of any of the caves without entrances.

*The addenda to Davies (1949) in Bulletin 19 of the National Speleological Society (December 1957) reports 119 new caves in West Virginia. 21 meet the selection rules of this paper. The length distribution of the 20 of these caves with one entrance agrees very closely with the previous data up to 630 feet. For greater length the new caves exhibit a somewhat greater average length. This trend agrees with the correction required for the previous results to account for the additional length reported for some previously described caves, and especially for the introduction of a new very long cave (Culverson Creek Cave, 10,800 feet, one entrance) The length distribution functions should now become zero at 10,800 feet rather than 5,500 feet.

From a population of 21 new caves about 2 ± 1.4 should have 2 entrances. One is reported, the Carpenter's Pit-Swago Pit system, with a length of 8,850 feet.

A recent addition to Pennsylvania data is by Bernard L. Smeltzer and Ralph W. Stone in the Bulletin of the Pennsylvania Department of Internal Affairs, vol. 24, no. 5, April 1956. Seventeen new caves are described, nine being in quarries and cuts. Those over 100 feet long with natural entrances have lengths in rough agreement with Figure 2, $n = 1$ (112′, 600′, 880′ and 1620′). None have more than one entrance.

It is concluded that the newly reported caves in both states probably come from the same populations of "open" caves as the previously reported sets, and hence do not themselves represent any of the entranceless caves (except of course for the new quarry caves). Therefore the predictions of the number of zero entrance caves in each state should be scaled upward proportionately. These changes, + 215 and + 81 in West Virginia and Pennsylvania respectively for caves over 100 feet long, are less than the standard errors of the original estimates.

Conclusion

A more complex model would be justified only if some important phenomena are found to be omitted from the present analysis. Just as we might have said that the Poisson distribution was adequate if we did not realize that length was an important variable and that λ should be constant for all caves in an area, we may say that the present model is adequate because we do not yet perceive the relations which still require explanation. The completeness of the theory has been tested by its ability to correlate the frequencies of caves with different numbers of entrances, and by its ability to predict with reasonable accuracy the length distribution of cave with more than one entrance using only the lengths of caves with one entrance. The ultimate test is to *count* the number of caves without entrances and determine their lengths, but this procedure is unavailable to us.

References

Folsom, Franklin, (1956) Exploring American caves: New York, Crown Publishers, pp 3, 18, 187.
Davies, William E., (1949) Caverns of West Virginia: West Virginia Geological Survey, Vol. XIX.
Stone, Ralph W., (1953) Caves of Pennsylvania: National Speleological Society, Bulletin 15, December.
Cramer, Harold (1955) The elements of probability theory: New York, John Wiley.

APPENDIX

The stochastic process

Figure 3 is a schematic representation of the stochastic model of cave entrance genesis. The joint probability, $p(\ell,n,t+\Delta t)$, is the sum of the three mutually exclusive joint probabilities, $p(\ell,n-1,t,A)$, $p(\ell,n,t,C)$, and $p(\ell,n+1,t,B)$. In addition

$$(1) \qquad p(A/\ell,n-1,t) = r\ell\Delta t \ ,$$

$$(2) \qquad p(B/\ell,n+1,t) = s(n+1)\Delta t \ ,$$

and

$$(3) \qquad p(C/\ell,n,t) = 1-r\ell\Delta t-sn\Delta t$$

from the section Statistical Theory. Only terms first order in Δt have been retained.

The stochastic process may be written

$$(4) \qquad p(\ell,n,t+\Delta t) = p(\ell,n-1,t,A) + p(\ell,n,t,C) + p(\ell,n+1,t,B)$$

or, using (1), (2) and (3),

$$(5) \qquad p(\ell,n,t+\Delta t) = p(\ell,n-1,t)(r\ell\Delta t) + p(\ell,n,t)(1-r\ell\Delta t-sn\Delta t) + p(\ell,n+1,t)(s(n+1)\Delta t)$$

which, after rearranging and dividing through by $s\Delta t$, becomes

$$(6) \qquad \frac{p(\ell,n,t+\Delta t) - p(\ell,n,t)}{s\Delta t} = (n+1)p(\ell,n+1,t) - (n+\lambda\ell)p(\ell,n,t) + \lambda\ell p(\ell,n-1,t) \ ,$$

where $\lambda = r/s$.

In the limit $\Delta t \rightarrow 0$ the left hand term in (6) becomes the partial derivative of $p(\ell,n,t)$ with respect to t. We know from physical reasoning that the solution to (6) in the differential-difference form reaches an asymptotic value for large t which is independent of its initial value. The left hand term of (6) becomes zero when this asymptotic condition of statistical equilibrium is reached. Since we are interested here only in the case of statistical equilibrium, setting the right hand term of (6) equal to zero and rearranging, we obtain

$$(7) \qquad \left[(n+1)p(\ell,n+1) - \lambda\ell p(\ell,n)\right] - \left[np(\ell,n) - \lambda\ell p(\ell,n-1)\right] = 0$$

which is satisfied by the condition

$$(8) \qquad (n+1)p(\ell,n+1) - \lambda\ell p(\ell,n) = c_1 g(\ell) \ ,$$

where g is an undefined function of ℓ. The sum

$$(9) \qquad \sum_{n=0}^{\infty} p(n) = 1$$

by definition so p(n) must approach zero at least as fast as $1/n^2$ as $n \rightarrow \infty$. Hence the left hand terms of (8) must approach zero as $n \rightarrow \infty$, hence $c_1 = 0$. Equation (8) becomes

$$(10) \qquad (n+1)p(\ell,n+1) - \lambda\ell p(\ell,n) = 0$$

which has the solution

$$(11) \qquad p(\ell,n) = \frac{(\lambda\ell)^n}{n!} p(\ell,0)$$

Probability distribution of n

Equation (11) states that the joint probability $p(\ell,n)$ is a Poisson probability distribution in the variable $\lambda\ell$. The probability p(n) is one marginal distribution of (11) and may be obtained by integrating (11) over all ℓ.

$$(12) \qquad p(n) = \int_0^{\infty} \frac{(\lambda\ell)^n}{n!} p(\ell,0)d\ell$$

Now, $p(\ell,0)$ is not available since we can directly observe nothing about the case n = 0. However, from the available conditional probability $p(\ell/1)$, $p(\ell,0)$ may be obtained and hence (12) evaluated. From (11),

$$(13) \qquad p(\ell,0) = \frac{p(\ell,1)}{\lambda\ell}$$

and also

$$(14) \qquad p(\ell,1) = p(\ell/1)p(1)$$

so therefore, together with (12), we obtain the formula

$$(15) \qquad p(n) = p(1) \int_0^{\infty} \frac{(\lambda\ell)^{n-1}}{n!} p(\ell/1)d\ell$$

The condition (9) suffices for the evaluation of p(1) from (15). The complete expression for p(n) is then

$$(16) \qquad p(n) = \frac{\int_0^{\infty} \frac{(\lambda\ell)^{n-1}}{n!} p(\ell/1)d\ell}{\int_0^{\infty} \frac{e^{\lambda\ell}}{\lambda\ell} p(\ell/1)d\ell}$$

The value of λ may be determined from data by obtaining the best fit of (16) to the empirical probability distribution. However, since data are only available for $n \neq 0$, a truncated probability distribution must be formed from (16) using the condition

$$(17) \qquad \sum_{n=1}^{\infty} p'(n) = 1$$

From (15) and (17) we obtain the truncated probability distribution, analogous to (16),

$$(18) \qquad p'(n) = \frac{\int_0^{\infty} \frac{(\lambda\ell)^{n-1}}{n!} p(\ell/1)d\ell}{\int_0^{\infty} \frac{e^{\lambda\ell}-1}{\lambda\ell} p(\ell/1)d\ell}$$

This is the relation which may be used to find λ from data by "fitting" to the observed frequencies f_n.

The technique used here to find the best estimate of λ for an observed cave population is known as the maximum likelihood method (4). In this method the derivative with respect to λ is taken of the logarithm of the likelihood function, and equated to zero, viz.:

$$(19) \qquad \frac{d}{d\lambda} \ln \prod_{n=1}^{\infty} p'(n)^{f_n} = 0$$

This derivation will not be continued here but the result is that the best estimate of λ is obtained by calculating from (18) the average number of entrances for caves with one or more entrances and setting this equal to the observed value. From (18) and the definition of an average value we find

$$(20) \qquad \bar{n} = \sum_{n=1}^{\infty} np'(n) = \frac{\int_0^{\infty} e^{\lambda\ell} p(\ell/1)d\ell}{\int_0^{\infty} \frac{e^{\lambda\ell}-1}{\lambda\ell} p(\ell/1)d\ell} \ ,$$

while the value from data is

$$(21) \qquad \alpha = \frac{1}{M} \sum_i n_i \quad \text{(i over all caves in M)} \ ,$$

where M is the total number of known caves.

Equating (20) and (21) we obtain the integral equation which determines λ from the observed value of α and the probability distribution of ℓ for n = 1.

$$(22) \qquad \alpha = \frac{\int_0^{\infty} e^{\lambda\ell} p(\ell/1)d\ell}{\int_0^{\infty} \frac{e^{\lambda\ell}-1}{\lambda\ell} p(\ell/1)d\ell}$$

The method used to solve this equation will be described in the subsequent section on Numerical Procedure.

When λ is obtained, equation (18) yields $p'(n)$ and the predicted values of the number of caves with n entrances, m_n, may be found from

$$m_n = Mp'(n) \qquad (23)$$

Frequency distribution of lengths

Inserting $p(\ell,0)$ from (13) into (11), and writing the joint probabilities as the product of the conditional and marginal probability distributions, we obtain

$$p(\ell/n)p(n) = \frac{(\lambda\ell)^{n-1}}{n!}\, p(\ell/1)p(1) \qquad (24)$$

Recognizing that equation (15) is equivalent to

$$p(n) = p(1)\, \frac{(\lambda)^{n-1}}{n!}\, \overline{\ell_1^{\,n-1}} \qquad (25)$$

the bar indicating the average value over all caves <u>with one entrance</u> of ℓ^{n-1} we may divide (24) by (25) to obtain

$$p(\ell/n) = \frac{\ell_1^{\,n-1}}{\overline{\ell_1^{\,n-1}}}\, p(\ell/1) \qquad (26)$$

or, from the definition of a <u>distribution function</u>,

$$dF(\ell/n) = \frac{\ell_1^{\,n-1}}{\overline{\ell_1^{\,n-1}}}\, dF(\ell/1) \qquad (27)$$

$F(\ell/n)$ is the <u>frequency distribution function of lengths</u> for caves with n entrances.

The average length of caves with n entrances may be obtained from (26) in the form

$$\overline{\ell}_n = \frac{\overline{\ell_1^{\,n}}}{\overline{\ell_1^{\,n-1}}} \qquad (28)$$

The averages are again over all caves with one entrance in the term on the right. Similar expressions arise for averages of higher powers of length.

Numerical procedure

Only basic aspects of the computational procedure will be described below. Details of the computer program and various arithmetical manipulations to obtain results in the desired forms are not presented.

The data used in this study consisted of length and entrance information on individual caves.

(a) Evaluation of α.

$$\alpha = \frac{1}{M}\sum_n nf_n = \frac{1}{M}\sum_i n_i \quad \text{(i over all caves in M)} \qquad (29)$$

(b) Evaluation of λ.

Equation (22) may be written, using the identity $p(\ell/1)d\ell = dF(\ell/1)$,

$$\alpha = \frac{\int_0^{\infty} e^{\lambda\ell}\, dF(\ell/1)}{\int_0^{\infty} \frac{e^{\lambda\ell}-1}{\lambda\ell}\, dF(\ell/1)} \qquad (30)$$

Now, $F(\ell/1)$ is not available as a continuous function, but rather as a series of steps as a function of ℓ. If plotted, each step would correspond to a single cave and would be of size $1/f_1$. At some lengths possessed by more than one cave, a number of such steps would add up to a larger step. Since the numerator and denominator of (30) represent the areas under curves of the arguments versus $F(\ell/1)$, equation (30) may be approximated, using the actual data, by the expression

$$\frac{\frac{1}{f_1}\sum_i e^{\lambda\ell_{i,1}}}{\frac{1}{f_1}\sum_i \frac{e^{\lambda\ell_{i,1}}-1}{\lambda\ell_{i,1}}} \quad \text{(i over all caves in } f_1\text{)}. \qquad (31)$$

Equation (31) becomes, after expanding the exponentials in their Taylor series (noting that the summations over each term, divided by f_1, are the average values),

$$\frac{1 + \lambda\overline{\ell_1} + \frac{1}{2!}\lambda^2\overline{\ell_1^{\,2}} + \frac{\lambda^3}{3!}\overline{\ell_1^{\,3}} + \ldots + \frac{\lambda^\nu}{\nu!}\overline{\ell_1^{\,\nu}} + \ldots}{1 + \frac{\lambda}{2!}\overline{\ell_1} + \frac{\lambda^2}{3!}\overline{\ell_1^{\,2}} + \frac{\lambda^3}{4!}\overline{\ell_1^{\,3}} + \ldots + \frac{\lambda^\nu}{(\nu+1)!}\overline{\ell_1^{\,\nu}} + \ldots} \qquad (32)$$

This may be rewritten as the polynomial expression in λ

$$0 = (1-\alpha) + (1-\tfrac{\alpha}{2})\lambda\overline{\ell_1} + (1-\tfrac{\alpha}{3})\frac{\overline{\ell_1^{\,2}}}{2}\lambda^2 + (1-\tfrac{\alpha}{4})\frac{\overline{\ell_1^{\,3}}}{6}\lambda^3 + \ldots + (1-\frac{\alpha}{\nu+1})\frac{\overline{\ell^{\nu}}\lambda^\nu}{\nu!} + \ldots \qquad (33)$$

In an actual numerical computation a choice must be made at this point as to how many terms of the infinite series will be used. This depends upon the rapidity of convergence of the series. In the present computation there is another consideration. The higher powers of ℓ depend essentially on the data for only one cave and are strongly influenced by small errors in the length. If the series has not converged before this happens, the resulting value of λ may be strongly in error. Therefore the series was terminated at the term which represented the data of essentially only one cave. In the machine computation, equation (33) was solved by means of Newton's method, keeping all terms up to the sixth power of λ. It was subsequently found that the last term became negligibly small with respect to the others.

Newton's method is an iterative procedure which converges very rapidly if an initial close value of λ is used. Since we hope that the calculated $p'(n)$ will fit the observed distribution, a good starting value should be that which makes the calculated curve fit the data <u>exactly</u> at $n = 1$ and $n = 2$. The ratio

$$\frac{p'(1)}{p'(2)} = \frac{1}{\frac{1}{2}\lambda\overline{\ell_1}} \qquad (34)$$

therefore

$$\lambda = \frac{2p'(2)}{\overline{\ell_1}p'(1)} \qquad (35)$$

which may be approximated by using the ratio of the observed frequencies f_1 and f_2.

$$\lambda \cong \frac{2f_2}{\overline{\ell_1}\, f_1} \qquad (36)$$

Using this initial value, Newton's method converged in about four iterations to 0.01%.

Using the notation of equation (31), equation (18) may be put into the form

$$p'(n) = \frac{\frac{1}{f_1}\sum_i \frac{(\lambda\ell_{i,1})^{n-1}}{n!}}{\frac{1}{f_1}\sum_i \frac{e^{\lambda\ell_{i,1}}-1}{\lambda\ell_{i,1}}}, \qquad (37)$$

or

$$p'(n) = \frac{\frac{1}{n!}\lambda^{n-1}\overline{\ell_1^{\,n-1}}}{1 + \frac{\lambda}{2!}\overline{\ell_1} + \frac{\lambda^2}{3!}\overline{\ell_1^{\,2}} + \ldots + \frac{\lambda^\nu}{\nu!}\overline{\ell^{\nu}} + \ldots}, \qquad (38)$$

from which $p'(n)$ was obtained numerically. With the formula (23) m_n was computed.

The distribution function for lengths was obtained from

$$(39)\qquad F(\ell/n) = \frac{1}{f_1 \ell_1^{\,n-1}} \sum_{i(\text{longest cave})}^{\text{cave of length } \ell} \ell_{i,1}^{\,n-1}$$

the summation being carried out with the caves arranged in order of decreasing length.

Poisson distribution

The Poisson distribution, mentioned in the INTRODUCTION and used for comparison purposes in RESULTS, is derived in detail in many books (4). The expression equivalent to (16) for the Poisson distribution is

$$(40)\qquad p^*(n) = \frac{e^{-\lambda^*} \lambda^{*n}}{n!}$$

and, equivalent to the truncated distribution of (18),

$$(41)\qquad p'^*(n) = \frac{\lambda^{*n}}{n!(e^{\lambda^*} - 1)}$$

Standard errors of predictions

The frequencies f_n would be found to be distributed according to the Poisson distribution

$$(42)\qquad p(f/n) = \frac{e^{-Mp'(n)} (Mp'(n))^{f_n}}{f_n!}$$

if an original population of M known caves and the associated ones with zero entrances could be observed over a very long time. The mean f_n is the expected frequency $Mp'(n)$. The standard deviation is

$$(43)\qquad \sigma_n = \sqrt{Mp'(n)} \cdot n \neq 0.$$

The standard error of the predicted values of m_0 may be shown to be given, to a first approximation, by

$$(44)\qquad \Delta m_0 \cong \frac{\Delta\alpha}{\alpha - 1} m_0 ,$$

in which $\Delta\alpha$ is the standard error of the sample estimate of α, and is equal to

$$(45)\qquad \Delta\alpha = \frac{\sigma}{M^{1/2}}$$

where σ, the standard deviation of $p'(n)$, may be approximated by using the Poisson distribution representation for $p'(n)$ and the sample value of α. For this case it is found that

$$(46)\qquad \sigma = \sqrt{(\alpha(1 + \lambda^* - \alpha))} ,$$

where λ^* is the transcendental function of α,

$$(47)\qquad \frac{\lambda^*}{1 - e^{-\lambda^*}} = \alpha$$

SYMBOLS

A	The event of a cave of length ℓ and n-1 entrances at time t gaining an entrance in the time interval Δt
B	The event of a cave of length ℓ and n+1 entrances at time t losing an entrance in the time interval Δt
C	The event of a cave of length ℓ and n entrances at time t neither losing or gaining an entrance in the time interval Δt
f	Observed frequency (number) of caves
F	Distribution function of cave lengths
ℓ	Cave length, feet
m	Predicted frequency (number) of caves
M	Total number of observed caves with one or more entrances
n	Number of entrances
p	Probability function
r	Probability per foot per year of a cave gaining an entrance
s	Probability per entrance per year of a cave losing an entrance
t	Time
α	Average number of entrances for caves with one or more entrances
Δ	Increment of ....
λ	Ratio of probability of a cave gaining an entrance per foot to probability of a cave losing an entrance per entrance; r/s

Subscripts

i	Summing index for individual caves
n	Number of entrances

Others

*	Poisson-distribution function
$\overline{(\ )}$	Average value
( )'	Truncated probability distributions

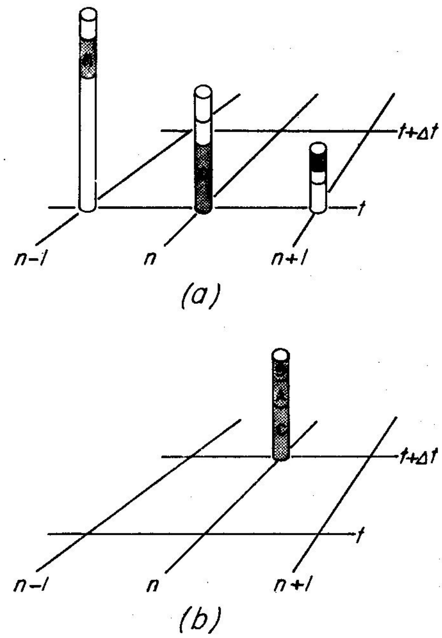

Figure 3. Schematic representation of the stochastic model of cave entrance genesis. The heights of the columns represent the number of caves with each number of entrances or, equivalently, the probabilities of caves having each number of entrances ($p(\ell,n,t)$, etc...). (a) The situation at time t. The shaded portions A, B, and C represent respectively the fraction of caves of length ℓ and n-1 entrances which will gain an entrance in the time interval Δt; the fraction of caves of length ℓ and n+1 entrances which will lose an entrance in the time interval Δt; and the fraction of caves of length ℓ and n entrances which will neither gain nor lose an entrance in the time interval Δt. The unshaded portions represent the fractions of the number of caves with each number of entrances which will not appear with n entrances at time $t+\Delta t$, i.e. caves with n+1 entrances gaining an entrance, etc... (b) The situation at time $t+\Delta t$. The number (or probability) of caves with n entrances at time $t+\Delta t$ is composed of the contributions from the states of n-1, n, and n+1 entrances at time t. Of course there also exist values at n-1 and n+1 but these have been omitted from the drawing.

A similar development may be used to find the probability of there being exactly f_n caves at time t with n entrances. The result of such an analysis is given in Equation (42).