NPS-OR-96-003

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Katz Distributions, with Applications to Minefield Clearance

by

Alan Washburn

March 1996

Approved for public release; distribution is unlimited.

Prepared for: Office of Naval Research
Arlington, VA

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5000

Rear Admiral M. J. Evans
Superintendent

Richard Elster
Provost

This report was prepared for and funded by the Office of Naval Research (Code 322 TE).

This report was prepared by:

ALAN WASHBURN
Professor of Operations Research

Reviewed by:

FRANK PETHO
Chairman (Acting)
Department of Operations Research

Released by:

GORDON E. SCHACHER
Dean of Research (Acting)

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1996	Technical

4. TITLE AND SUBTITLE
Katz Distributions, with Applications to Minefield Clearance

5. FUNDING NUMBERS

6. AUTHOR(S)
Alan Washburn

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-OR-96-003

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research (Code 322 TE)
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

At the end of a mine clearance operation, some judgment must be made about the number of mines that remain uncleared. A Bayesian analysis will require a prior distribution for the number of mines. This report describes some desirable properties of Katz distributions for that purpose. The Katz class consists of binomial, Poisson, and negative binomial distributions.

14. SUBJECT TERMS
Bayes, minefield, clearance

15. NUMBER OF PAGES
27

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

Katz Distributions, With Applications to Minefield Clearance

Alan Washburn

Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943

1. Introduction

Consider a region S that contains M mines. Suppose that an action is taken to remove the mines from S, that Y mines are removed, and that a query is made about the number of mines $X \equiv M - Y$ that remain in S. If M is known, then, since Y is also known, so is X. In the common case where M is unknown, however, there is bound to be some uncertainty about the number of mines that remain after the clearance operation, and therefore about whether S has been sufficiently "sanitized". A quantitative analysis of the situation will require a probability distribution for X, which will in turn depend on the prior probability distribution for M. If the distribution of M is of a particular type, the Katz type that is shortly to be described, then the distribution of X will be of the same type, a simplifying feature. Katz distributions also have some other appealing properties, so there are reasons to begin a minefield clearance analysis by assuming a Katz distribution for M. The purpose of this paper is to summarize some of these properties.

The clearance action is assumed to be nonexhaustive, by which is meant that the operation cannot guarantee to remove every mine (if it were exhaustive, then of course X would be 0 regardless of M). Instead, the clearance action will be assumed to remove every mine independently with known probability p, with $0 \le p < 1$. the usual assumption in studying minefield clearance. In general p depends on the amount of clearance effort, sweep widths, and perhaps other parameters, but the nature of that dependence is not of concern here. Parameter p will be referred to as the "level" of clearance.

Although minefield clearance is the motivating application, of course the "mines" could actually be ore pockets, oil strikes, unexploded ordnance, piles of doggie-doo (domestic minefield), or whatever. The basic idea is that an imperfect attempt is made to find a scattered collection of objects.

Section 2 defines Katz distributions and summarizes what is already known about them. Other properties of Katz distributions are derived in Sections 3 and 4, particularly in Theorem 1 where it is shown that X still has a Katz distribution after Y is observed. Although Sections 2 - 4 will continue to refer to the objects being studied as "mines", the material in those sections is really just a collection of abstract but possibly useful facts about Katz distributions. The implications for models of minefield clearance are addressed in Section 5.

2. Katz Distributions (definition)

Katz (1965) describes a probability distribution $x_0, x_1, \ldots$ with the property that

$$x_{j+1}/x_j = \frac{\alpha + \beta j}{1+j}; \quad j \ge 0. \tag{1}$$

The distribution (1) will be referred to as a "Katz distribution with parameters α and β", provided α and β meet certain requirements that will be specified in the next paragraph. Given x_0, equation (1) determines $x_1, x_2, \ldots$. Since the sum $x_0 + x_1 + \ldots$ must be 1, x_0 is determined implicitly.

The parameter α must be nonnegative, since it is the ratio x_1/x_0, and β must be less than 1 to enforce convergence to 0 for large j. If $\beta < 0$, then (1) will eventually produce negative probabilities unless $-\alpha/\beta$ is an integer. To prevent this possibility, $-\alpha/\beta$ is required to be an integer when β is negative. The restrictions on parameters are thus that

$$\alpha \geq 0, \quad \beta < 1, \quad \text{and} \quad -\alpha/\beta \text{ is an integer when } \beta < 0. \tag{2}$$

Let the generating function be $g(z;\alpha,\beta) \equiv \sum_{j=0}^{\infty} x_j z^j$. Katz (1965) showed that

$$g(z;\alpha,\beta) = \left[(1-\beta)/(1-\beta z)\right]^{\alpha/\beta}, \tag{3}$$

with (3) being interpreted as $\exp(-\alpha z)$ (the limit as β approaches 0) if $\beta = 0$. It follows that the initial probability must be

$$x_0 = g(0;\alpha,\beta) = (1-\beta)^{\alpha/\beta}. \tag{4}$$

If M is a random variable with a Katz distribution with parameters α and β (hereafter abbreviated $M \sim K(\alpha, \beta)$), then

$$E(M) = \mu = \alpha/(1-\beta) \quad \text{and} \quad Var(M) = \sigma^2 = \alpha/(1-\beta)^2. \tag{5}$$

It is not hard to establish that a Katz distribution is

- if $\beta < 0$, a binomial distribution with $-\alpha/\beta$ trials and success probability $\beta/(\beta - 1)$,
- if $\beta = 0$, a Poisson distribution with mean α, or
- if $\beta > 0$, a negative binomial distribution. If α/β is an integer, this is the distribution of the number of failures until the α/β^{th} success when the failure probability is β. However, the "number of successes" α/β can actually be any positive real number in this case.

The Katz class includes no other distribution, so it can be thought of as the union of three familiar types.

Since the mean and variance are more familiar parameters than α and β, the solution of (5) for α and β in terms of μ and σ^2 may be useful:

$$\beta = 1 - \mu/\sigma^2 \quad \text{and} \quad \alpha = \mu^2/\sigma^2. \tag{6}$$

Clearly $\mu \geq 0$ and $\sigma^2 \geq 0$ in (6), but some nonnegative (μ, σ) pairs are impossible because of the restriction that $-\alpha/\beta$ must be an integer when β is negative. This restriction is <u>not</u> imposed by Katz (1965), who simply zeros all probabilities after (and including) the first that (1) would make negative. Unfortunately, this tactic falsifies equations (3) – (6). For example suppose $\alpha = 1$ and $\beta = -2$. Then (1) has $x_1/x_0 = 1$ and $x_2/x_1 = -1/2$, so Katz would take $x_0 = x_1 = 1/2$, $x_i = 0$ for $i \geq 2$. The mean of this distribution is $\mu = 1/2$, not 1/3 as would be obtained by (5). The fact that (3) – (6) are false when $\beta < 0$ and $-\alpha/\beta$ is not an integer is not recognized in Katz (1965), nor in subsequent restatements such as Johnson and Kotz (1969).

Since $\beta = 1 - \mu/\sigma^2$, all (μ, σ^2) pairs where $0 < \mu \leq \sigma^2$ are possible. This covers situations where there is great uncertainty about the number of mines present.

3. Useful Katz Properties

3.1 Sample-Observe-Subtract (SOS)

The main property that makes Katz distributions useful in minefield clearance is that the class is closed under the kind of SOS operations described in the introduction. Formally,

<u>Theorem 1</u>: Let M be the number of mines, suppose $M \sim K(\alpha, \beta)$, let Y be the number of mines removed when each mine is removed with probability p, independently of the others, and let $X = M - Y$ be the number of mines remaining (not removed). Then, conditional on the event $(Y = y)$ being given, X ~ $K(\alpha', \beta')$, where α' and β' are given by (10) with $q = 1 - p$.

<u>Proof</u>: Let $x_j = \Pr(M = j)$ and $x_j^* = \Pr(X = j|Y = y)$; $j = 0, \ldots$. Then

$$
\begin{aligned}
x_j^* \Pr(Y=y) &= \Pr(Y=y \ \underline{\text{and}}\ X=j) \\
&= \Pr(Y=y \ \underline{\text{and}}\ M=y+j) \qquad (7) \\
&= \Pr(Y=y \mid M=y+j)\Pr(M=y+j)
\end{aligned}
$$

But $\Pr(Y=y \mid M=y+j)$ is the binomial probability of y successes in $y+j$ trials, so, letting $q=1-p$,

$$x_j^* \Pr(Y=y) = \binom{y+j}{y} p^y q^j x_{y+j}; \ j=0,\ldots \qquad (8)$$

Taking the ratio of successive terms in (8), the factor $\Pr(Y=y)$ cancels and

$$x_{j+1}^* \big/ x_j^* = \left\{\frac{y+j+1}{j+1}\right\} q \left\{\frac{\alpha+\beta(y+j)}{y+j+1}\right\}. \qquad (9)$$

The first { } factor in (9) is a ratio of combinatorial coefficients, and the second is by assumption x_{y+j+1}/x_{y+j}. The two $(y+j+1)$ factors in (9) cancel, so (9) is again a linear function of j divided by $j+1$, as was to be shown. If α and β satisfy (2), it is easy to check that the same is true of the revised parameters α' and β', where

$$\alpha' = q(\alpha+\beta y) \qquad \beta' = q\beta. \qquad (10)$$

This concludes the proof. □

Glazebrook and Boys [1995] introduce a larger class of distributions that is still closed under the SOS operation. Binomial distributions are generalized to "light tailed" distributions, negative binomial distributions are generalized to "heavy tailed" distributions, and the Poisson distribution continues to play its central role. The Katz class can be regarded as a two-parameter subset with convenient analytic properties.

Theorem 1 resolves a certain minefield paradox. Suppose that a minefield is cleared to the .5 level, and that Y mines are removed in the process. One might argue that Y mines must remain, since only half have been removed. But how can it be that the number estimated to remain should increase with the number cleared, since clearance is by its nature subtractive? The paradox disappears when one realizes that clearance to a known level provides both evidence and removal. When $\beta > 0$, the evidence part dominates and the estimated number remaining does indeed increase with the number removed. When $\beta < 0$, the removal part dominates. In the Poisson case $\beta = 0$, the number removed does not affect the distribution of the number that remain.

Since clearance is a process carried out in time, it is likely that clearance times $T_1, \ldots, T_y$ will also be known when $(Y = y)$ is observed. If the magnitudes of these times influence the posterior distribution of M, then the clearance times, as well as the number of mines cleared, should be accounted for. However, there is no effect of this kind as long as the clearance level p is calculable, regardless of the initial distribution of the number of mines. The proof of this statement can be found in Appendix A.

3.2 Simple Sampling

Theorem 1 governs the case where the number of mines removed (Y) is observed. There are also circumstances where Y is not observed. One example is where M is the number of mines in region S, but only some fraction q of S (call it S') is of concern. If q is interpreted to be the probability that any given mine will be in S', then the number of mines X in S' is the number remaining after sampling M at the level q, but without observing the results of the sample. Theorem 2 states that X is still Katz.

Theorem 2: Let M be the number of mines, suppose $M \sim K(\alpha, \beta)$, and let X be the number of mines in the sample when each mine is included with probability q, independently of the others. Then $X \sim K(\alpha', \beta')$, where α' and β' are given by (14) with $p = 1 - q$.

Proof: Since X is binomial when M is given,

$$E\left(z^X\right)=\sum_{j=0}^{\infty} x_j \sum_{i=0}^{j}\binom{j}{i} q^i p^{j-i} z^i \tag{11}$$

$$=\sum_{j=0}^{\infty} x_j (qz+p)^j \tag{12}$$

$$=g(qz+p;\alpha,\beta). \tag{13}$$

Equation (12) is obtained from (11) by combining the factors q^i and z^i, and then employing the Binomial Theorem. Equation (13) is obtained from (12) by recalling the definition of the generating function g( ). After recalling (3) and rearranging (13), X can be shown to be Katz with parameters

$$\alpha'=\frac{\alpha q}{1-\beta p} \qquad \beta'=\frac{\beta q}{1-\beta p}. \tag{14}$$

If α and β satisfy (2), then so do α' and β'. □

Of course, the number of mines Y *removed* from M is also Katz, but with p and q reversed in (14).

3.3 Simple Initial Threat (SIT)

Uncertainty about the number of mines implies uncertainty about whether the minefield is safe for a transitor to cross. The simplest quantification is to define the parameter

$$t \equiv \text{probability that a given mine kills the transitor,} \tag{15}$$

and then assume that all mines act independently. For example, suppose that mines are distributed uniformly and independently in a minefield with width W, that each mine actuates with probability B if the transitor's straight line path takes it to within $A/2$ of the mine, and that the transitor will be killed with probability D, conditional on actuation. Then, as long as $W >> A$ and the transitor's path is near the center of the minefield (ignoring edge effects, in other words), the parameter t is ABD/W. However, t does not need to be calculated in that way – the calculation could involve actuation curves, navigation errors, and edge effects as in Odle (1977).

The transitor is assumed to encounter the mines one at a time. As long as the transitor survives, the probability that the next mine kills it is by assumption t, independently of any others. The probability that *all* M mines fail to kill the transitor is therefore $(1-t)^M$, and the probability that the first transitor to enter the minefield is killed is

$$\text{SIT} \equiv 1 - E((1-t)^M). \tag{16}$$

If $M \sim K(\alpha, \beta)$, equation (16) can be evaluated by recalling that $E(z^M) = g(z; \alpha, \beta)$ and substituting $1-t$ for z; that is, $\text{SIT} = 1 - g(1-t; \alpha, \beta)$. If clearance is carried out before the transitor enters the minefield, then of course α' and β' from (10) or (14) would be substituted for α and β.

3.4 Threat to Following Transitors

The second and following transitors are much harder to deal with analytically than the first. Odle (1977) gives formulas for several multi-transitor measures, but derivation is non-trivial even when the number of mines is known. An exception is the "catastrophic failure" probability c_n, the probability that none of n transitors are sunk. Let Q_n be the catastrophe probability for a single mine. Then c_n is simply Q_n^m for m independent mines. If $M \sim K(\alpha, \beta)$, then the (average) catastrophe probability is

$$c_n(\alpha,\beta) = E\left(Q_n^M\right) = g(Q_n;\alpha,\beta), \tag{17}$$

where g( ) is the generating function given by (3). Odle gives the formula when M is Poisson, a special case. As in the case of SIT, the important thing is that the generating function of M be known.

Q_n would be $(1 - t)^n$ if each transitor's track were chosen independently of the others, but multiple transitors are usually assumed to attempt to follow the same track. In that case the computation of Q_n can become a significant task in itself, particularly if navigation errors are involved, but the degree of difficulty has nothing to do with the distribution of the number of mines present.

There appear to be no simple, closed-form formulas other than (17) when multiple-transitors are present, even when the number of mines is known. There are practical methods for calculating the casualty distribution and other statistical measures (Odle, 1977), but the methods do not simplify when the number of mines has a Katz (or even a Poisson) distribution.

A simple upper bound on E_n, the expected number of casualties out of n transitors, can be obtained by observing that the number of casualties cannot exceed M, and therefore that E_n cannot exceed $E(M)$. If each mine causes a casualty with probability at most D whenever it detonates, then a better bound is

$$E_n \leq D\ E(M). \tag{18}$$

If $M \sim K(\alpha, \beta)$, then $E(M)$ is given by (5). Since E_n is necessarily a nondecreasing function of n, (10) is sharpest for large values of n. Of course, E_1 = SIT.

4. Sums and Partitions

Throughout this section there are n independent mine populations M_i, with $M_i \sim K(\alpha_i, \beta_i)$, $i = 1, \ldots, n$. It will be assumed that $\alpha_i > 0$, since otherwise $M_i \equiv 0$ and population i could be omitted. The total number of mines is $M \equiv M_1 + \ldots + M_n$. The clearance level for the i^{th} population is p_i, with $q_i \equiv 1 - p_i$. The number of type i mines cleared is Y_i, with $Y \equiv Y_1 + \ldots + Y_n$, and the number remaining is X_i, with $X \equiv X_1 + \ldots + X_n$. Of course $X_i + Y_i = M_i$ and $X + Y = M$. These mine populations might be different kinds of mines in one minefield, the numbers of mines in different minefields, or any other partition of M into n parts.

4.1 Sums

If all of the mine populations have Katz distributions, does the total number of mines M also have a Katz distribution?

<u>Theorem 3</u>: If $\beta_i = \beta$ for all i, then $M \sim K(\alpha, \beta)$, where $\alpha = \alpha_1 + \ldots + \alpha_n$. Otherwise, M does not have a Katz distribution.

<u>Proof</u>: Since the M_i are all independent, the generating function of M is

$$g(z) = \prod_{i=1}^{n} \left[(1-\beta_i)/(1-\beta_i z)\right]^{\alpha_i/\beta_i}. \tag{19}$$

If $\beta_i = \beta$ for all i, then (19) reduces to $g(z) = \left[(1-\beta)/(1-\beta z)\right]^{\alpha/\beta}$, the generating function of a Katz random variable. Otherwise, (19) does not have the required form and M is therefore not Katz. □

<u>Corollary 1</u>: If $q_i\beta_i = \beta$ for some parameter β, $i = 1, \ldots, n$, and if Y_i is observed for $i = 1, \ldots, n$, then

$$X \sim K\left(\sum_{i=1}^{n}(\alpha_i q_i + \beta Y_i), \beta\right).$$

<u>Proof</u>: According to (10), $X_i \sim K(\alpha_i q_i + \beta Y_i, q_i\beta_i)$ when Y_i is given. Since $q_i\beta_i = \beta$, the conclusion that X has a Katz distribution then follows from Theorem 3. □

__Corollary 2__: Suppose $q_i = (1/\beta_i - 1)/(1/\beta - 1)$ for some β $i = 1, \ldots, n$. Then $X \sim K(\alpha_{\text{TOT}}, \beta)$, where

$$\alpha_{\text{TOT}} = \sum_{i=1}^{n} \beta(\alpha_i/\beta_i). \qquad (20).$$

If instead $p_i = (1/\beta_i - 1)/(1/\beta - 1)$ for $i = 1, \ldots, n$, then $Y \sim K(\alpha_{\text{TOT}}, \beta)$.

__Proof__: The condition on q_i enforces $\beta_i' = \beta$ and $\alpha_i' = \beta(\alpha_i/\beta_i)$ in (14), which applies when the number of mines cleared is not observed. The conclusion then follows from Theorem 3. If the condition on p_i holds, then the same logic applies to Y, the number of mines not removed. □

If X and Y in corollary 2 are __both__ to have Katz distributions, then it is necessary that p_i and q_i both be proportional to $(1/\beta_i - 1)$. This is not possible unless $\beta_i = 0$ for all i (in which case β is also 0 — this is the Poisson case), or if β_i and p_i are both independent of i. Either of these conditions also ensures that X will be Katz in the SOS condition where Y_i is observed for all i as in corollary 1.

If the only possibilities in clearance are that the clearance of each type is either observed or not observed, then the number of mines of each type remains Katz after clearance, and independence is preserved between populations. Other kinds of observations, however, can destroy this Katz structure. One example would be an observation that the total number of mines X cannot exceed some limit because of logistic considerations. Of more concern is the possibility that Y might be observed, without knowing all of its components.

To be precise, suppose that Y is known without knowing the identity of any of the mines that have been cleared. Are the residual numbers X_i still independent and Katz under this condition? The answer is yes if $Y = 0$, since the observation that $Y = 0$ is equivalent to the observation that $Y_i = 0$ for all i. The answer is also yes if $p = 1$, since in that case

$X_i = 0$ for all i. One might hope that the answer would still be yes even if $p < 1$ and $Y > 0$, provided $\beta_i = \beta$ for all i, since the latter condition is sufficient for Y to be Katz. Unfortunately, this is not true. Appendix B shows that conditional independence fails unless $\beta = 0$.

If $\beta_i = 0$ for all i, then the X_i are still Poisson and independent when Y is given because the information in Y is irrelevant; the distribution of X_i has nothing to do with the number of mines cleared. Thus, it is only in the Poisson case where the individual populations remain independent and Katz when Y alone is observed.

4.2 Partitions and MCK Distributions

If $M \sim K(\alpha, \beta)$, then a closed form expression for the probability mass function of M, valid if $\beta < 0$ or $\beta > 0$, is

$$P(M = m) = \frac{(-\alpha/\beta)_m}{m!}(1-\beta)^{\alpha/\beta}(-\beta)^m; \quad m \ge 0. \tag{21}$$

The notation $(x)_m$ is taken from Feller (1957) where $(x)_m$ is defined to be $x(x-1)\ldots(x-m+1)$ for $m \ge 1$, with $(x)_0 \equiv 1$ (m is a nonnegative integer, but x can be any real number). The limit as $\beta \to 0$ produces a Poisson distribution, so in that sense (21) is valid for all (α, β) satisfying (2). If $M_i \sim K(\alpha_i, \beta)$, and if $M_1, \ldots, M_n$ are all independent, then $M \sim K(\alpha, \beta)$ according to Theorem 3. Let $\underline{M} \equiv (M_1, \ldots, M_n)$, and $\underline{m} \equiv (m_1, \ldots, m_n)$. Then

$$P(\underline{M} = \underline{m} | M = m) = \frac{\prod_{i=1}^{n} P(M_i = m_i)}{P(M = m)}; \quad m_i \ge 0, m = m_1 + \ldots + m_n. \tag{22}$$

All of the factors involving $(1 - \beta)$ and $(-\beta)$ raised to powers cancel in (22), leaving

$$P(\underline{M} = \underline{m} | M = m) = \frac{\prod_{i=1}^{n} \frac{(-\alpha_i/\beta)_{m_i}}{(m_i)!}}{\frac{(-\alpha/\beta)_m}{(m)!}}; \quad m_i \ge 0, m = m_1 + \ldots + m_n. \tag{23}$$

The distribution (23) will be referred to as a "multivariate conditional Katz distribution with parameters $\underline{\alpha}$, β, and m", or MCK for short. The MCK distribution is a multivariate hypergeometric distribution when $\beta < 0$, or a multinomial distribution in the limit as $\beta \to 0$ (Johnson and Kotz, 1969). When $\beta > 0$, the MCK distribution has been called a multivariate Polya-Eggenberger distribution (Johnson and Kotz, 1977) on account of its relationship to certain urn-sampling schemes, or simply the multivariate Polya distribution (Janardin and Patil, 1970). Thinking of M_i as the number of balls in an urn leads to a practical way of generating $\underline{M}$ in a Monte Carlo simulation, since only a single Katz sample of the total M is really required. This is the gist of Theorem 4.

Theorem 4: Let $M \sim K(\alpha, \beta)$, where $\alpha = \alpha_1 + \ldots + \alpha_n$, $\alpha_i \geq 0$ for $1 \leq i \leq n$. The pair (α_i, β) is assumed to satisfy (2) for $1 \leq i \leq n$. Consider the following procedure for placing M balls in n urns. For $k = 0, \ldots, M-1$, the $k + 1^{st}$ ball is placed in urn i with probability p_i, where

$$p_i = \frac{\alpha_i + \beta k_i}{\alpha + \beta k}, \tag{24}$$

and where k_i is the number of balls already in urn i. If M_i is the number of balls finally placed in urn i, then $M_i \sim K(\alpha_i, \beta)$, and all of the M_i are independent of each other, $i = 1, \ldots, n$. □

Proof: Let $\underline{M} \equiv (M_1, \ldots, M_n)$, and $\underline{m} \equiv (m_1, \ldots, m_n)$. It will be shown by induction that $P(\underline{M} = \underline{m} | M = m)$ is given by (23) for $m \geq 0$. Since (23) is equivalent to (22), the theorem follows upon removing the condition on M.

Let $Q(\underline{m})$ be $P(\underline{M} = \underline{m} | M = m_1 + \ldots + m_n)$, and note that $Q(\underline{0}) = 1$, a special case of (23) where $m_1 + \ldots + m_n = 0$. Suppose $Q(\underline{m})$ is given by (23) for all $\underline{m}$ such that $m_1 + \ldots + m_n = k$; let e_i be an n-vector all of whose components are zero except for component i, which is 1; and let $k_i = m_i - 1$, $i = 1, \ldots, n$ (if $k_i < 0$, then the corresponding term may be omitted from (25) below). Then, conditioning on the ball configuration after k balls have been placed,

$$Q(\underline{m}) = \sum_{i=1}^{n} Q(\underline{m} - e_i)\frac{\alpha + \beta k_i}{\alpha + \beta k} \tag{25}$$

where $\underline{m}$ is now any configuration such that $\sum_{i=1}^{n} m_i = k+1$. $Q(\underline{m} - e_i)$ on the right hand side of (25) is by assumption given by (23), and it is now only a matter of some algebra to conclude that $Q(m)$ on the left hand side is also given by (23). Since $\underline{m}$ is arbitrary except for its sum, this completes the inductive proof. □

Comment: When $\beta = 0$, (24) makes $p_i = \alpha_i/\alpha$ for every ball. The fact that a Poisson random variable produces independent Poisson parts when partitioned in this manner is well known (e.g. Ross (1993)). When $\beta \neq 0$, if each ball is placed in urn i with probability α_i/α, instead of according to (24), then by Theorem 2 $M_i \sim K(\alpha_i', \beta_i')$, where $\alpha_i' = \alpha_i/(1-\beta(1-\alpha_i/\alpha))$ and $\beta_i' = \beta(\alpha_i/\alpha)/(1-\beta(1-\alpha_i/\alpha))$. $E(M_i)$ is still $\alpha_i/(1-\beta)$, but it is not true that $M_i \sim K(\alpha_i, \beta)$, and furthermore $M_1, \ldots, M_n$ are not mutually independent. These latter properties require that the balls be allocated according to (24).

5. Tactical Decision Aids (TDA's) for Minefield Clearance

Barring the possibility of exhaustive search, any mine clearance campaign has got to cope with the problem of deciding when to stop. Stopping after a fixed time is of course an option, but it makes sense to let the stopping time depend on results achieved to date, particularly when there is as much initial uncertainty as is usually the case in mine clearance. As a minimum, therefore, any decision aid for mine clearance should be able to display the "status" of a clearance effort in terms that support the stopping decision, or, more generally, decisions about what should be done next. The natural status of a minefield is the risk that it poses to the traffic against which it was designed. In simplest terms this risk is measured by SIT. SIT depends strongly on the number of mines remaining, so it is hard to resist the conclusion that the number of mines initially present must be an input, even if the number is so vaguely known that the input must be a probability distribution. Without some input or assumption about the number of mines initially present, it is hard to imagine how a basically subtractive clearance activity could result in sufficient knowledge about the number of mines remaining to support a computation of SIT.

In spite of the above considerations, current (1996) mine clearance TDA's typically do not require the number of mines present to be an input. There are a variety of reasons for this, but the only important point is that the reader should understand that the necessity for a distribution for the number of mines to be an explicit input is arguable.

Even if one accepts the idea that the number of mines M must be thought of as a random variable, it does not necessarily follow that M should have a Katz distribution, since Bayes Theorem could just as well be applied to a general distribution. A general distribution would require storing 1000 numbers if the maximum conceivable number of mines were 999. A Katz distribution requires only 2, but performing a Bayesian update on a general distribution is trivial with a modern computer; in a different context, NODESTAR (Stone and Corwin, 1995) performs such updates with 10^6 states, rather than only 10^3. Using a general distribution would also have the advantage that any observation with a known conditional probability law could be the basis of a Bayesian update, which is not true in the Katz case. The idea of using general distributions does not become computationally unwieldy until multiple random variables must be described jointly. If there were for example 3 mine types, the number of each of which does not exceed 999, then there would be 10^9 joint possibilities. Today's computers cannot perform Bayesian updates on that scale. The Katz assumption provides no relief from this kind of explosion, since there is no useful theory for multiple Katz random variables unless they are all independent.

On the other hand, it is also true that very little is lost by restricting input distributions to be of the Katz type. The two Katz parameters are sufficient for quantifying the center and spread of a distribution, and it is hard to imagine knowing enough about the number of mines to need more detail. In fact, the Katz restriction may be operationally welcome, since the entire distribution is determined from only two estimated numbers. With these thoughts in mind, one TDA (MIXER) proposed by the author (Washburn, 1995) employs Katz distributions exclusively, requiring the user to quantify uncertainty by providing a mean and standard deviation for each mine type.

Katz distributions also have some computational advantages compared to the general case. At any point in a clearance operation where the clearance level and the number(s) of mines cleared are known, it is easy to compute revised Katz parameters (formula (10)) and then SIT (formula (16)) while the comparable operations in the general case would require extensive computation. The Katz divisibility properties described in Theorem 4 could also prove handy. If a region containing M mines must be divided into two parts, then Theorem 4 describes how the number of mines in the two parts can be independent, Katz, and still sum to M. The comparable operation in the general case may be difficult or impossible. Perhaps most important, the availability of an analytic expression for SIT opens up the possibility (as in MIXER) of posing the mathematical problem of minimizing SIT subject to constraints on the clearance effort, a computational problem that would be much more difficult in the general case.

Appendix B shows that one seemingly innocent observation (total mines cleared) can create analytical havoc when multiple Katz populations are present – the residual mine populations are not independent. Letting the populations have general distributions would not relieve this problem, since the general case includes the Katz case. Rigorously processing measurements of this sort will require Bayesian updates of joint distributions, regardless of the nature of the marginal distributions. The point, again, is that there is little to be gained by permitting marginal distributions to be general, rather than Katz.

The difficulty described in the paragraph above could be relieved by forcing all distributions to be Poisson, as noted on page 11. Poisson distributions also have other important analytic advantages, but unfortunately have only a simple parameter. For example a Poisson distribution with mean 100 necessarily has a standard deviation of (only) 10. Inclusion of distributions with $\beta > 0$ (negative binomial distributions) in the permitted class seems essential to model the large uncertainty about mine numbers that is to be expected.

In summary, the Katz class of distributions is large enough to support mine clearance TDA's, and offers several convenient analytic properties. The Poisson class would be even more convenient, but is not large enough. Permitting general distributions would lose the convenient Katz properties, and therefore should be done only if there is some use for the added flexibility.

In minefields where multiple mine types are present, all theorems proved above require an independence assumption. If observations that would destroy independence are contemplated, then, in spite of the implied computational burden, a Bayesian TDA will have to be based on general, multivariate distributions. In other words, Katz distributions appear to be the right answer only as long as observations preserve independence between populations.

APPENDIX A
Clearance times have no additional value

When the number of mines cleared (Y) is observed, it is likely that the clearance times $U_1, \ldots, U_Y$ will also be observed. These times turn out to have no additional value in making inferences about the initial number of mines M, whether or not M has a Katz distribution, and therefore no value for the residual number of mines $M - Y$. This result may seem counterintuitive. If one searches for 24 hours, finding 5 mines in the first hour and none thereafter, then intuition argues that there are probably no remaining mines, whereas there might be more mines if the clearance times were scattered over the whole clearance period. This intuition would be correct if the probability law $F(\,)$ governing the detection times were unknown, since there is information about $F(\,)$ in the clearance times. If $F(\,)$ is known, however (as it must be if the clearance level is calculable), then the corollary below states that the clearance times are useless.

Theorem: Let M be a nonnegative random variable, and let $T_1, \ldots, T_M$ be independent random variables with common distribution function $F(\,)$. Let t be any real number, let I_i indicate the event $(T_i \leq t)$, let $Y = I_1 + \ldots + I_M$, and let $\mathbf{U} = (U_1, \ldots, U_Y)$, where $U_1, \ldots, U_Y$ are the nondecreasing order statistics of those T_i for which $T_i \leq t$. If m and y are nonnegative integers for which $0 \leq y \leq m$, and if $\mathbf{u} = (u_1, \ldots, u_y)$ is a real vector, then either $\Pr(Y = y, M = m) = 0$, or

$$\Pr(\mathbf{U} = \mathbf{u} | Y = y, M = m) = \Pr(\mathbf{U} = \mathbf{u} | Y = y). \tag{A1}$$

Proof: Both sides of (A1) are well defined if $\Pr(Y = y, M = m) > 0$. Furthermore, both are 0 unless $u_i \leq t$ for $i = 1, \ldots, y$ and $u_i \leq u_{i+1}$ for $i = 1, \ldots, y - 1$, so suppose that those conditions hold. Define the event

$$E_{ym} \equiv (M = m) \bigcap_{i=1}^{y} (T_i = u_i) \bigcap_{i=y+1}^{m} (T_i > t). \tag{A2}$$

Then

$$\Pr\left(E_{ym}\right) = \Pr(M = m)\left[\prod_{i=1}^{y} dF(u_i)\right]\left[1 - F(t)\right]^{m-y}. \tag{A3}$$

The event $(\mathbf{U} = \mathbf{u}) \cap (M = m)$ includes E_{ym} and other mutually exclusive events that have the same probability, since the first y of the T_i are not necessarily the smallest. If all components of u are unequal, the number of these events is $y!\binom{m}{y}$, the number of permutations of m things taken y at a time. More generally, let there be K distinct components in u, with n_k being the number of times the k^{th} is repeated. Then there are $\binom{y}{\boldsymbol{n}}\binom{m}{y}$ such events, where $\binom{y}{\boldsymbol{n}}$ is the multinomial coefficient for y things taken n_1, $n_2, \ldots, n_K$ at a time. Every subset of $\{T_1, \ldots, T_m\}$ of size y can be assigned to the components of $\mathbf{u}$ in $\binom{y}{\boldsymbol{n}}$ different ways. Thus

$$\Pr(\mathbf{U} = \mathbf{u}, M = m) = \binom{y}{\boldsymbol{n}}\binom{m}{y}\Pr\left(E_{ym}\right). \tag{A4}$$

Since $\Pr(Y = y, M = m) = \Pr(M = m)\binom{m}{y}F(t)^y[1 - F(t)]^{m-y}$, it is a simple matter to take the ratio $\Pr(\mathbf{U} = \mathbf{u}, M = m)/\Pr(Y = y, M = m)$ to obtain

$$\Pr(\mathbf{U} = \mathbf{u} | Y = y, M = m) = \binom{y}{\boldsymbol{n}}\prod_{i=1}^{y}\left[dF(u_i)/F(t)\right]. \tag{A5}$$

But the right hand side of (A5) does not depend on m, so it must also be $\Pr(\mathbf{U} = \mathbf{u} | Y = y)$. Conditional on $(Y = y)$ being given, the order statistics $\mathbf{U}$ are distributed as if $\mathbf{U}$ were the order statistics of the truncated distribution $F(\ )/F(t)$, sampled y times. □

Corollary: Either $\Pr(\mathbf{U} = \mathbf{u}, Y = y) = 0$ or $\Pr(M = m | \mathbf{U} = \mathbf{u}, Y = y) = \Pr(M = m | Y = y)$.

Proof: Let $A = (\mathbf{U} = \mathbf{u})$, $B = (Y = y)$, $C = (M = m)$, and assume $\Pr(A \cap B) > 0$. It follows that $\Pr(C | A \cap B)$ and $\Pr(C | B)$ are well defined, and both are zero if

$\Pr(B \cap C) = 0$. If $\Pr(B \cap C) > 0$, then $\Pr(A | B \cap C) = \Pr(A | B) > 0$ by the theorem, and all that remains is to write the definition of $\Pr(C | A \cap B)$, cancel equal factors, and observe that the result is $\Pr(C | B)$. □

APPENDIX B
A counterexample to conditional independence

Suppose $M_i \sim K(\alpha, \beta)$, $i = 1, 2$, and let $M = X_1 + X_2$ and $\alpha = \alpha_1 + \alpha_2$. Each of the M mines is cleared with probability $p < 1$, independent of the others. Let Y_i be the number of mines of type i cleared, let $Y = Y_1 + Y_2$, let $X_i = M_i - Y_i$ be the number of type i mines remaining after clearance, and let $q = 1 - p$. It will be shown that X_1 and X_2 are not independent under the condition $Y = 1$, unless $\beta = 0$. It suffices to show the same thing about the events $(X_1 = 0)$ and $(X_2 = 0)$.

Let E be the event ($X_1 = 0$ and $X_2 = 0$ and $Y = 1$), and let F_1 be the event ($X_1 = 0$ and $Y = 1$). E is the same as the event that exactly one mine is present and that it is cleared, so, using (21) and the fact that $M \sim K(\alpha, \beta)$ by Theorem 2,

$$P(E) = pP(M = 1) = p\alpha(1-\beta)^{\alpha/\beta}. \tag{B1}$$

F_1 is the union of two mutually exclusive events

G = ($X_1 = 1$ and $Y_1 = 1$ and $Y_2 = 0$) and

H = ($X_1 = 0$ and $X_2 \geq 1$ and $Y_2 = 1$).

$P(G)$ is just $pP(X_1 = 1)\ P(Y_2 = 0)$, but evaluating $P(H)$ requires a summation:

$$P(H) = P(X_1 = 0)\sum_{j=1}^{\infty} P(X_2 = j)\left[jpq^{j-1}\right], \tag{B2}$$

with the factor in [ ] being a binomial probability. The sum can be evaluated by differentiating the generating function of X_2. Letting $\gamma = \left(\frac{1-\beta}{1-\beta q}\right)^{1/\beta}$, the result is that

$$P(H) = pP(X_1 = 0)\left(\frac{\alpha_2}{1-\beta q}\right)\gamma^{\alpha_2}. \tag{B3}$$

Since $P(Y_2 = 0) = \gamma^{\alpha_2}$ and $P(X_1 = 1) = \alpha_1 P(X_1 = 0)$, $P(F_1)$ can be obtained by summing $P(G) + P(H)$:

$$P(F_1) = pP(X_1 = 0)\gamma^{\alpha_2} f_1 = p(1-\beta)^{\alpha_1/\beta}\gamma^{\alpha_2} f_1. \tag{B4}$$

where $f_1 = \alpha_1 + \alpha_2/(1 - \beta q)$. Defining F_2 similarly to F_1, it follows by symmetry that

$$P(F_2) = p(1-\beta)^{\alpha_2/\beta}\gamma^{\alpha_1} f_2, \tag{B5}$$

where $f_2 = \alpha_2 + \alpha_1/(1 - \beta q)$.

If the events $(X_1 = 0)$ and $(X_2 = 0)$ are to be independent when the event $(Y = 1)$ is given, it should be true that $P(F_1)P(F_2) = P(E)P(Y = 1)$, so consider the ratio $R = P(F_1)P(F_2)/(P(E)P(Y = 1))$. Since $P(Y=1) = \frac{p\alpha}{1-\beta q}\gamma^{\alpha}$, all factors are known and R can be computed. Most factors cancel, leaving only

$$R = (1-\beta q)\, f_1 f_2/\alpha^2. \tag{B6}$$

It can be shown using simple algebra that $(1 - \beta q)f_1 f_2 \geq \alpha^2$ for all $\alpha_1,\ \alpha_2 \geq 0$, with equality possible only if $\beta = 0$. Thus $(X_1 = 0)$ and $(X_2 = 0)$ are not conditionally independent. A slightly stronger statement is possible: the probability of having 0 mines remaining, given $Y = 1$, is larger than the prediction based on conditional independence.

REFERENCES

Feller, W. 1957. *An Introduction to Probability Theory and its Applications* (2nd ed.), Wiley, p. 48.

Glazebrook, K.D and Boys, R.J. 1995. "A Class of Bayesian Models for Optimal Exploration," *Journal of the Royal Statistical Society (B)*, **57**, No.4, pp. 705-720.

Janardin, K. and Patil, G. 1970. "On the Multivariate Polya Distribution: A Model of Contagion for Data with Multiple Counts." Random Counts in Scientific Work II. Random Counts in Physical Science, Geo-Science, and Business (ed. G.P. Patil). The Pennsylvania State University Press, pp. 143-162.

Johnson, N. and Kotz, S. 1969. *Discrete Distributions*, Wiley, NY, p. 36-43.

Johnson, N. and Kotz, S. 1977. *Urn Models and Their Applications*, Wiley, NY, p. 194.

Katz, L. 1965. "Unified Treatment of a Broad Class of Discrete Probability Distributions," in *Classical and Contagious Discrete Distributions* (G. Patil, ed.), Statistical Publishing Society, Pergamon Press, Oxford.

Odle, J. 1977. *Minefield Analysis for Channelized Traffic*, NSWC/WOL TR 77-109, Naval Surface Weapons Center, White Oak, MD.

Ross, S. 1993. *Probability Models* (5th ed.) Academic Press, NY, pp. 216-217.

Stone, L. and Corwin, T. 1995. "NODESTAR: A Nonlinear, Discrete, Multiple-target, Correlator-tracker: Part 1," *U.S. Navy Journal of Underwater Acoustics*, **45**, No. 3, pp. 1-16.

Washburn, A. 1995. "MIXER: A TDA for Mixed Minefield Clearance," NPSOR-95-011PR, Naval Postgraduate School, Monterey, CA 93943.

DISTRIBUTION LIST

1. Research Office (Code 08) 1
Naval Postgraduate School
Monterey, CA 93943-5000

2. Dudley Knox Library (Code 52) 2
Naval Postgraduate School
Monterey, CA 93943-5002

3. Defense Technical Information Center 2
8725 John J. Kingman Rd., STE 0944
Ft. Belvoir, VA 22060-6218

4. Department of Operations Research 1
Editorial Assistant (Code OR/Bi)
Naval Postgraduate School
Monterey, CA 93943-5000

5. Prof. Alan R. Washburn (Code OR/Ws) 5
Naval Postgraduate School
Monterey, CA 93943-5000

6. COMINEWARCOM (N02R) 1
325 Fifth St. SE
Corpus Christi, TX 78419

7. Coastal Systems Station 1
6703 W. Highway 98
Panama City, FL 32407-7001

8. Center for Naval Analyses 1
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

9. Office of Naval Research (322TE) 1
800 North Quincy Street
Arlington, VA 22217-5660

10. Daniel H. Wagner, Assoc. 1
2 Eaton St., Suite 500
Hampton, VA 23669