1. Introduction

A database can be considered as a matrix, where the rows contain the data of one individual (e.g., a person), and the columns contain the data of the same type for many rows. Each row contains fields, such as the first name, last name, date of birth, etc. The types of data are structured in disjoint columns. The data are assumed to be independent, i.e., they are not related in any way to each other. To illustrate this, consider the following:

- The first name (f1), the last name (l1), gender (g1), and the date of birth (d1)
- The first name (f2), the last name (l2), gender (g2), and the date of birth (d2)
- The first name (f3), the last name (l3), gender (g3), and the date of birth (d3)
- The first name (f4), the last name (l4), gender (g4), and the date of birth (d4)

The data are assumed to be independent, i.e., they are not related in any way to each other. To illustrate this, consider the following:

- The first name (f1), the last name (l1), gender (g1), and the date of birth (d1)
- The first name (f2), the last name (l2), gender (g2), and the date of birth (d2)
- The first name (f3), the last name (l3), gender (g3), and the date of birth (d3)
- The first name (f4), the last name (l4), gender (g4), and the date of birth (d4)

This is because the data are independent and identically distributed (i.i.d.). In other words, the probability of any particular entry in the database is the same for all entries. This is important because it allows us to make valid statistical inferences about the population from which the database was drawn.
high probability for any set $A \subseteq \{a, b\}$ and any column $x \in \{0, 1\}$. The answer
\[\log(q) = -D \cdot q \cdot \log(q)\]
as $a$ is given previously. Theorem 3 generalizes this result for
cases where the entries have different distributions in the different
columns. Section 4 develops a new method for estimating the probability of
the event that two random matrices are different. The result is applied for
the case of a random matrix in Section 5. Theorem 1 calculates the
accuracy probability that the rows of the random matrices are different.
This theorem is divided into three cases: (i) the above stated case:
the $x \cdot y$ hold with high probability for any $y$, (ii) a case with high
probability for any $y$, and (iii) the case for each element of $y$.

The method of the present paper is combinatorial. Paper II of
the author concludes with the following results. The method of
paper is combinatorial, and uses the well-known Fanoa approximation
technique (from the old, see II).

3. A sequence of experiments with different outcomes

We may obtain a combinatorial for $A = \{a, b\}$ if the entries of two
rows in the matrix are determined by $a$ and $b$. In this case, we define
as the product of the probability of the event that all the outcomes of the
$i$-th experiment are different:

\[P(A, B) = \prod_{i=1}^{n} P_i(A, B)\]

where $P_i(A, B)$ is the probability of the event that $A_i$ and $B_i$ are
different.

Lemma 3 is a basic result of the section giving good estimates on
$P(A, B)$. The statement of the main result of the section is not improved when
the number of experiments is increased. The statement is actually to
estimate the probability that the graph is empty.

The graph constructed from a set of edges $E$ is a graph with
at most one edge for each pair of vertices. The graph is defined as
the graph with $E$ as the set of edges. The graph is a graph with
at most one edge for each pair of vertices. The graph is defined as
the graph with $E$ as the set of edges. The graph is a graph with
Theorem 1: Let \( v_1, \ldots, v_n \) be non-negative integers. Then

\[
\sum_{\text{matching}} (-1)^{m} \sum_{i=1}^{n} v_i \sum_{1 \leq j \leq n} \binom{v_j}{i} = 0
\]

where the sum is over all matchings of the graph and \( m \) is the number of edges in the matching.

Proof: By the principle of inclusion-exclusion, the left-hand side of (1) counts the number of matchings of a graph with edges labeled by \( v_1, \ldots, v_n \), and the coefficient of \( (-1)^m \) gives the number of matchings with an odd number of edges.

If all \( v_i = 0 \), then all matchings are empty and the theorem holds trivially. Otherwise, there exists at least one \( v_i > 0 \).

If \( m < \min(v_1, \ldots, v_n) \), then the left-hand side of (1) is 0, since any matching with fewer edges than any \( v_i \) cannot be counted.

If \( m = \min(v_1, \ldots, v_n) \), then the left-hand side of (1) is equal to \( (-1)^{m} \sum_{i=1}^{n} v_i \sum_{1 \leq j \leq n} \binom{v_j}{i} \), which is 0, since the number of matchings of a graph with exactly \( m \) edges is equal to the number of matchings with exactly \( m \) edges of the opposite sign.

If \( m > \min(v_1, \ldots, v_n) \), then the left-hand side of (1) is the sum of the contributions of all matchings with more than \( \min(v_1, \ldots, v_n) \) edges.

The only remaining case is when all matchings have at least three edges and are non-connected. Then all matchings are connected, and the theorem holds trivially.
where the matchings, V-matchings and W-matchings are subgraphs of \( G(v_1, v_2, \ldots, v_n) \)
and \( v_1 v_2 \ldots v_n \) is a path. The only difference in the proof is that the graph on which the negative terms on a positive terms generated by a matching of every number of edges.

**Lemma 3.**

\[
1 + \sum_{j=1}^{n-1} \frac{(n)}{j} \left( \frac{n-j}{2} \right) \frac{n-j}{2} - \frac{\left( n-j \right)}{2} \left( \frac{n-j}{2} \right) \frac{n-j}{2} - \frac{\left( n-j \right)}{2} \left( \frac{n-j}{2} \right) \frac{n-j}{2}
\]

\[
= \sum_{j=1}^{n-1} \frac{(n)}{j} \left( \frac{n-j}{2} \right) \frac{n-j}{2} - \frac{\left( n-j \right)}{2} \left( \frac{n-j}{2} \right) \frac{n-j}{2} - \frac{\left( n-j \right)}{2} \left( \frac{n-j}{2} \right) \frac{n-j}{2}
\]

\[
\leq \sum_{j=1}^{n-1} \frac{(n)}{j} \left( \frac{n-j}{2} \right) \frac{n-j}{2} - \frac{\left( n-j \right)}{2} \left( \frac{n-j}{2} \right) \frac{n-j}{2} - \frac{\left( n-j \right)}{2} \left( \frac{n-j}{2} \right) \frac{n-j}{2}
\]

**Proof:** \( P(x) \) is the probability of the event that \( E_1, E_2, \ldots, E_n \) are all different. That is, we ignore the sum of the probabilities.

\[
P_{x} = P_{E_1} \cap P_{E_2} \cap \ldots \cap P_{E_n}
\]

when \( C_1, C_2, \ldots, C_n \) is a partition of \( \{1, 2, \ldots, n\} \) with at least one Carring more than one element, not all \( C_1, \ldots, C_n \) are nonempty, and \( C_1, \ldots, C_n \) are nonempty

**Proof:** Consider the probability in \( (1) \) with same weight, therefore if they are connected with the weight given in \( (1) \) such that it leads to an upper estimate. Consider the sum

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \left( \left( \sum_{k=1}^{n} \left( \sum_{l=1}^{n} \left( \sum_{m=1}^{n} \left( \sum_{p=1}^{n} \left( \sum_{q=1}^{n} \left( \sum_{r=1}^{n} \left( \sum_{s=1}^{n} \left( \sum_{t=1}^{n} \left( \sum_{u=1}^{n} \left( \sum_{v=1}^{n} \left( \sum_{w=1}^{n} \left( \sum_{x=1}^{n} \left( \sum_{y=1}^{n} \left( \sum_{z=1}^{n} \left( \sum_{\text{all possible combinations}} \left( -1 \right)^{r} \sum_{\text{all possible combinations}} \left( -1 \right)^{s} \sum_{\text{all possible combinations}} \left( -1 \right)^{t} \sum_{\text{all possible combinations}} \left( -1 \right)^{u} \sum_{\text{all possible combinations}} \left( -1 \right)^{v} \sum_{\text{all possible combinations}} \left( -1 \right)^{w} \sum_{\text{all possible combinations}} \left( -1 \right)^{x} \sum_{\text{all possible combinations}} \left( -1 \right)^{y} \sum_{\text{all possible combinations}} \left( -1 \right)^{z} \right) \right) \right) \right) \right) \right) \right) \right) \right) \right) \right) \right) \right) \right)
\]
If the partition is the dummy one, then the cases are empty with one exception: the empty case of the probability of the event when all $y$ are disjoint. Therefore, (I) is the case in which the probability of the event, where all $y$ are disjoint with weight 1, while the other probabilities are within a non-negative weight. Consequently, (I) is an upper estimate on $P(G)$.

Change the order of terms to (I):

$$
\sum_{\text{number of edges}} (-1)^{\frac{1}{2}} \sum_{\text{partitions}} \prod_{y \in C_j} P(y) = \sum_{\text{all partitions}} \prod_{y \in C_j} P(y) + \sum_{\text{all partitions}} \prod_{y \in C_j} P(y) = \sum_{\text{all partitions}} \prod_{y \in C_j} P(y)
$$

for a given matching of $\frac{1}{2}$-step. This is nothing else but the probability of the event that the $\frac{1}{2}$-adjacent in the matching are equal:

$$
\left( \sum_{\text{all partitions}} \prod_{y \in C_j} P(y) \right)^{\frac{1}{2}}.
$$

The number of matchings with $\frac{1}{2}$-step is

$$
\left( \begin{array}{c}
\frac{1}{2}
\end{array} \right) \left( \begin{array}{c}
-1\ 
\end{array} \right) \left( \begin{array}{c}
0\ 
\end{array} \right) \left( \begin{array}{c}
0\ 
\end{array} \right) = \left( \begin{array}{c}
-1\ 
\end{array} \right).
$$

This gives the fifth case of (I). The second and third cases of (II) hold, in a similar manner, to the null and first ones of (II), respectively. The lower estimates are proved in the same way.

\[ \square \]

5. Random matrix with different rows

The Laplace transform will be used for random matrices. Let $R$ be a random matrix with its rows and columns, whose entries of the $j$th column can have $j$ different values with probability $\frac{1}{j!}$, respectively. All the entries are chosen totally independently. Thus, the probability of the occurrence of a certain row in $R$ is $\frac{1}{j!}$, where $j$ is arbitrary between $1$ and $\infty$. The probability distribution of these sequences will be denoted by $\nu_j$. The following trivial observations will be used here.
Lemma 1. \( f(x) = e^{-x^2} \) from \( P(x, u) \) to \( P(x, v) \).

We want to study the probability of the event that the rows of the above matrices are different. Therefore the probability \( q_i(x, y, \theta) \) will be taken as \( P(x) \) in Lemma 1. Consider \( \sum_{i=1}^{n} q_i(x, y, \theta) \) for these probabilities:

\[
\sum_{i=1}^{n} q_i(x, y, \theta) = \sum_{i=1}^{n} \prod_{j=1}^{m} (1 - e^{-|x_j - y_j|^2 + \theta_j})
\]

(7)

Our investigation will be of asymptotic nature. From now on it is expected that \( n \) tends to the infinite and the other parameters depend on \( n \): \( n^{-1/2} / \log(n) + D(x, y) = \log(n) \). The asymptotic investigation in these will be such that the last mentioned trend to the Lemma 1, that is:

\[
w^2 \sum_{i=1}^{n} \prod_{j=1}^{m} (1 - e^{-|x_j - y_j|^2 + \theta_j})
\]

leads to a more natural result. It will be done in a logarithmic way, therefore the expression \( \log(1 - e^{-|x_j - y_j|^2 + \theta_j}) \) will play an important role. We will assume that we had the asymptotic expansion of \( \log(1 - e^{-|x_j - y_j|^2 + \theta_j}) \) analogous to the asymptotic expansion of \( \log(1 - e^{-|x_j - y_j|^2 + \theta_j}) \) \( q_i(x, y, \theta) \) introduced the ascendant sequence of order \( n \) for \( n \rightarrow \infty \) so that:

\[
2 \log n = \sum_{i=1}^{n} h_i(x) + o(n)
\]

when \( n \rightarrow \infty \) then

\[
\sum_{i=1}^{n} \frac{1}{2} \left( \frac{1}{2} \right) \left( \log n \right) - \left( \frac{1}{2} \right) \left( \sum_{i=1}^{n} v_i \right)
\]

(8)

leads to

\[
e^{-w^2}
\]

for the distribution \( x \).

Proof. Consider the limit of our terms for a fixed \( j \):

\[
\left( \begin{array}{c}
\log n
\end{array} \right) \left( \log n \right) - \left( \frac{1}{2} \right) \left( \sum_{i=1}^{n} v_i \right)
\]
can be replaced by
\[
\frac{x_i}{Y}
\]
On the other hand
\[
\left(\sum_{j=1}^{n} a_j^2\right)^2 \geq \sum_{j=1}^{n} a_j^2 \sum_{k=1}^{n} b_k
\]
follows by the definition of the square of nole 2 and (1). Therefore the limit of the above term in (11) is the same as the limit of
\[
\frac{1}{2} \left(\frac{1}{\gamma} \sum_{j=1}^{n} a_j^2 \right) \rightarrow \infty
\]
that is,
\[
\frac{1}{2} \sum_{j=1}^{n} a_j^2 \rightarrow \infty
\]
which implies that the sum of (11) and therefore (10) is bounded. In conclusion, the limit of (10) is equal to the infinite sum of the limits of its terms, that is
\[
\sum_{j=1}^{n} \sum_{k=1}^{n} \frac{1}{\gamma} a_j^2 b_k \rightarrow \infty
\]
We want to show that the other term is the lower and upper estimates of (2) used to prove the condition (3). Before proving that some other condition was needed.

Lemma 6. \( P = \{\pi_1, \ldots, \pi_k\} \) is a probability distribution, where \( \gamma \geq \delta_k \) then
\[
\frac{\sum_{i=1}^{k} \pi_i x_i^2}{\sum_{i=1}^{k} \pi_i x_i^2 - \delta_k}
\]
(2)

Prove. Consider the differences of the denominators and the numerator:
\[
\sum_{i=1}^{k} \pi_i x_i^2 \rightarrow \sum_{i=1}^{k} \pi_i x_i^2 - \delta_k \sum_{i=1}^{k} \pi_i x_i^2 + \sum_{i=1}^{k} \frac{\pi_i x_i}{\pi_i}
\]

(2)
Using the fact that the denominator is at worst 1, (12) easily follows.

**Lemma 5.** Let \((\epsilon_1, \ldots, \epsilon_n)\) be a probability distribution, where \(\epsilon_i \leq \frac{1}{3}\) for all \(i\).

\[
\sum_{j=1}^{n} \epsilon_j \left( \sum_{i=1}^{n} \epsilon_i \right) \leq 3 \sum_{j=1}^{n} \epsilon_j.
\]

**Proof.** The proof is similar but easier than the previous one:

\[
\sum_{j=1}^{n} \epsilon_j \left( \sum_{i=1}^{n} \epsilon_i \right) = 3 \sum_{j=1}^{n} \epsilon_j - 3 \sum_{j=1}^{n} \epsilon_j \epsilon_j = 3 \sum_{j=1}^{n} \epsilon_j - 3 \sum_{j=1}^{n} \epsilon_j \epsilon_j.
\]

**Lemma 6.** If (1) and (4) hold for all \(i\) with a fixed \(\epsilon \left( 0 < \epsilon < \frac{1}{3} \right)\) then the second and third rows of (11) tend to zero.

**Proof.** The \(i\)th terms of the second row of (11) can be approximated by

\[
\left( \epsilon \prod_{k \neq i} (1 - \epsilon_k) \right) \left( \epsilon \prod_{k \neq i} (1 - \epsilon_k) \right)^	op.
\]

The second factor tends to

\[
\prod_{k \neq i} (1 - \epsilon_k) = \epsilon^n
\]

as we have seen in the proof of Lemma 5. (1) implies

\[
\epsilon \prod_{k \neq i} (1 - \epsilon_k) \to 0
\]
therefore the third factor of (13) can be expanded as

$$\prod_{i=1}^{m} q_{i}^{n_{i}} = \left( \prod_{i=1}^{m} \sum_{j=0}^{\infty} q_{i}^{j} \right)^{\frac{1}{2}} H \left( \frac{2}{2 - \Psi} \right)$$

The base case of (13) tends to $\mathbb{P}^m$ while Lemma 9 gives the upper bound $(1 - \mathbb{P})$ for the second factor.

The condition of this lemma (i.e., $n_{0} \geq \sum_{i=1}^{m} n_{i}$) ensures that (26) results in $j + m$ when $m = 1 - \mathbb{P}$ and consequently (13) tends to zero. By the mutual independence, the infinite sum of (13) is the second sum of (13) she tends to zero.

The converse of the third row can be proved in the same way, using

**Theorem.** Let $R$ be a random matrix with no row and a column, where the entries of the $j$th column can have $d$ different values with probabilities $q_{1}, \ldots, q_{d}$, respectively. All the entries are chosen totally independently. Suppose that (13) holds. Then the probability of the event that the rows of $R$ are all different samples:

$$\Pr_{R} (\text{row unique}) = \begin{cases} 
\frac{1}{d}, & \text{if } \sum_{i=1}^{m} n_{i} = \infty, \\
\frac{d-1}{d}, & \text{if } \sum_{i=1}^{m} n_{i} < \infty, \\
1, & \text{if } \sum_{i=1}^{m} n_{i} = 0.
\end{cases}$$

**Proof:** The proof of the statement follows by Lemma 3 and 5.

The first and third rows are consequences of Lemma 4.

In [10], [11] proved a theorem on random matrices to connection with

**Remark:** It is basically equivalent to the problem of the skew theorem when we see because their initial was different.

The skew theorem is used in this chapter, not as a separate theorem, but as a tool. The condition is called the "skew" probability (i.e., $\sum_{i=1}^{m} n_{i}$) was important in the proof. This is shown by the following example. Let $n_{0}=\frac{1}{2}, n_{1}=\frac{1}{2}$, then the left-hand side of (13) is

$$\left( \frac{3n_{0}}{n_{1}} \right) \left( \frac{n_{0} - 1}{n_{1} - 1} \right)$$

which is not bounded below 1. Take $n_{0}=\frac{1}{2}$, by (13) both sides with zero. However, the second factor of (13) does not tend to zero. (9) cannot be true.
4. Typical sizes of functional dependencies and related keys

Let \( \Pr_{m,n,k} \) denote the probability of the event that exactly \( k \) pairs \((i_1,j_1), \ldots, (i_k,j_k)\) (i.e., \( i_1, \ldots, i_k \) are equal to each other and all other pairs are different). (Thus precisely) there are \( m \) distinct values \( a_1, \ldots, a_m \) such that \( i_1 = i_2 = \cdots = i_k \). Let \( a_i \neq a_j \) for all \( i \neq j \). Now if \( \mu_1, \mu_2 \) and \( \mu_3 \) are.

Let \( \mu \) be the number of \( \mu_1, \mu_2 \) and \( \mu_3 \) are.

Then \( \Pr_{m,n,k} = \frac{1}{\mu!} \binom{m}{k} \binom{n-k}{\mu} \) if \( \sum_{i=1}^{\mu} \binom{\mu}{i} \to a \).

Proof. There are \( \binom{m}{k} \binom{n-k}{\mu} \) ways to choose the set \( \{a_{i_1}, \ldots, a_{i_k}, \ldots, a_{i_n}\} \) Suppose that \( i_1 = i_2 = \cdots = i_k \). Then \( \Pr_{m,n,k} \) is the probability for the other classes of \( (i_1,j_1), \ldots, (i_k,j_k) \). The probability for the other classes of \( (i_1,j_1), \ldots, (i_k,j_k) \) will be the same. It is easy to see that.

We used the \( \mu \) patterns of this expression. Finally, \( \Pr_{m,n,k} \) must be all different. The probability of this event is \( \Pr_{m,n,k} \).

\[
\Pr_{m,n,k} = \frac{1}{\mu!} \binom{m}{k} \binom{n-k}{\mu} \left( \prod_{i=1}^{\mu} a_i \right)^k \Pr_{m,n-k,0}.
\]

The last factor is asymptotically equal to \( \Pr_{m,n-k} \) since \( \mu \to n \to k \to m \). Therefore the number of \( \mu \) pairs in each. The next of the product of the other factors of \( \Pr_{m,n,k} \) was determined in the proof of theorem.
Lemma 10. If (1) and (2) hold then the probability of the event that there are three equal CDE's tends to zero.

Theorem 2. Let H be a random matrix with n rows and m-steps relative to the distribution described above (1), (2), (3). Suppose that for some i = [i] the columns of H are in a column set in A_m.

\[ P[A_i \rightarrow H, \psi] = \begin{cases} 0, & \text{if } 2b \psi_{[i]} + \sum_{[j] \neq [i]} b \psi_{[j]} + m, \\ \frac{2^{b} \psi_{[i]} + \sum_{[j] \neq [i]} 2^{b} \psi_{[j]} + m}{2^{b} \psi_{[i]} + \sum_{[j] \neq [i]} 2^{b} \psi_{[j]} + m}, & \text{if } 2b \psi_{[i]} + \sum_{[j] \neq [i]} b \psi_{[j]} + m, \\ 1, & \text{if } 2b \psi_{[i]} + \sum_{[j] \neq [i]} b \psi_{[j]} + m. \end{cases} \]

Proof. Consider the restrictions of the rows of H within A_m. These rows all have a number of columns equal to \( m_1, m_2, \ldots, m_n \), where the sum equals the total rows. Suppose \( m_1 \geq 2 \times m_2 \). Start with the well-known equation

\[ P[A_i \rightarrow H, \psi] = \sum_{m_1=m_2} P[A_i \rightarrow H, m_1, m_2, \ldots, m_n, (\psi_{[i]} = m_1, \ldots, m_n)]. \]

The right-hand side of (10) will be divided into two parts: (I) \( m_i \geq 2 \times m_2 \) for case (I) the following trivial inequality is needed

\[ \sum_{m_1=m_2} P[A_i \rightarrow H, m_1, m_2, \ldots, m_n] \leq \sum_{m_1=m_2} P[A_i \rightarrow H, m_1, m_2, \ldots, m_n, (\psi_{[i]} = m_1, \ldots, m_n)]. \]

The left quantity tends to zero under condition (I). Therefore case (II) should only be considered. Here precisely, if (I) holds then the limit of \( P[A_i \rightarrow H, m_i] \) is equal to the limit of

\[ \sum_{m_1=m_2} P[A_i \rightarrow H, m_1, m_2, \ldots, m_n, (\psi_{[i]} = m_1, \ldots, m_n)]. \]
This expansion can be written in the form
\[ \sum_{i} P(A_{i} = k | m) \times P(A_{i} = k) = \sum_{i} P(A_{i} = k) \times P(m | A_{i} = k) \]
(2) where \( P(A_{i} = k | m) \) is the number of \( A_{i} \) to \( k \) \) and \( P(A_{i} = k) \) (the number of \( A_{i} \) to \( k \)).

Here
\[ P(A_{i} = k, m) = \sum_{i} P(A_{i} = k) \times P(m | A_{i} = k) \]

On the other hand, the limit of \( P(A_{i} = m) \) is given by Lemma 9. Therefore, the limit of (2) is
\[ \sum_{i} P(A_{i} = k) \times P(m | A_{i} = k) \]

The middle row of the statement is proved. The first and third rows are consequences of the inequality \( P(A_{i} = k, m) \geq P(A_{i} = k) \) for any \( m \).

Furthermore, let \( R \) be a random matrix with rows and columns chosen independently with probabilities \( P(A_{i} = k) \) and \( P(m | A_{i} = k) \). Suppose that \( A_{i} \) is a set of rows in columns of \( R \) and \( B \) is a column not in \( A_{i} \). Let the notation \( R = \sum_{j} R_{i,j} \). Then
\[ P(A_{i} = m, k) \rightarrow 0 \]
\[ \begin{cases} \frac{R_{i,j}}{\sum_{i} R_{i,j}} \rightarrow \frac{1}{m} & \text{if } j = k \\ \frac{R_{i,j}}{\sum_{i} R_{i,j}} \rightarrow 0 & \text{if } j \neq k \end{cases} \]

The main content of the latter statement is that if \( R \) is a set of columns of \( A \) of infinite length and \( \sum_{j} R_{i,j} \) is finite, then \( A \) holds with high probability for any \( B \).

So we in general, that if \( B \) is a random element of \( A \) and \( m \) and \( n \) and \( A \) are independent and \( A \) is a random element of \( B \) (which can be made independent of \( A \) or not), \( B \) must be made independent of \( A \) or not). It is much supporting that \( A \) is \( B \) or \( B \). Then \( R(A_{i}) \) can be replaced by \( R(A_{i}) \sum_{i} R_{i,j} \).
Let us turn back to the case when $a$ does not depend on $s$. If the size of $B$ is finite, $a$ then has Gaussian distribution and $P(F_{i}^{j} = 0, F_{j}^{i} = 0) = (1 - \exp(-2\gamma_{i}^{j} a^{2}))^{2}$. If the size of $B = \infty$, $a$ is a random variable with uniform distribution in $(0, 1)$ and $P(F_{i}^{j} = 0, F_{j}^{i} = 0) = (1 - \exp(-2\gamma_{i}^{j} a^{2}))^{2}$ and the probability of zero becomes simply $\frac{1}{4}$. We use that $\exp(-2\gamma_{i}^{j} a^{2})$ is big if $a^{2} > 2\gamma_{i}^{j}$ or equivalently $a > \sqrt{\frac{2\gamma_{i}^{j}}{\gamma_{i}^{j}}}$. When $B = \infty$ holds, $a$ is a normal log-logistic in a key and no proper subset is a key. The above reasoning proves the following statement.

Theorem 3. Let $B$ be a random matrix with $n$ rows and $m$ columns, where the entries are chosen independently following the distribution $a$. Suppose that $a \to \infty$ tends to infinity and $A = \{A_{i}\}$ is a set of columns of $B$. Then,

$P(A_{i} \text{ is a key}) = \begin{cases} 
\frac{1}{2} (e^{-2\gamma_{i}^{j} a^{2}} - 1), & a^{2} > 2\gamma_{i}^{j}, \\
\frac{1}{2} e^{-2\gamma_{i}^{j} a^{2}}, & a^{2} = 0, \\
\frac{1}{2} (1 - e^{-2\gamma_{i}^{j} a^{2}}), & a^{2} < 2\gamma_{i}^{j}.
\end{cases}$

It can be briefly said that the sets of size zero for both $\frac{1}{2} (e^{-2\gamma_{i}^{j} a^{2}} - 1)$ are keys with high probability.

REFERENCES


(Received June 19, 1990)