

# Re: minimum Hamming distance among random bit vectors

Source: <http://www.derkeiler.com/Newsgroups/sci.crypt/2004-04/2640.html>

---

**From:** Michael Amling (*nospam\_at\_nospam.com*)

**Date:** 04/29/04

Date: Thu, 29 Apr 2004 12:16:28 GMT

Francois Grieu wrote:

- > Let  $p(b,n,d)$  be the probability that among a set of
- >  $n$  random vectors of  $b$  bits, there exists two distinct
- > vectors which differ by at most  $d$  bits out of  $b$ .
- >
- > We restrict to integers such that  $b > 0$ ,  $n > 0$ ,  $0 \leq d \leq b$
- >
- > Except for typo, we have:
- >
- > [1]  $p(b,1,d) = 0$
- >
- > [2]  $p(b,2,d) = \sum C(b,j)/2^b$  with  $C(i,j) = i! / j! / (i-j)!$
- >  $j=0..d$
- >
- > [3]  $p(b,2^b,d) = 1$

You assume  $d > 0$

- >
- > [4]  $p(b,n+1,d) > p(b,n,d)$  or  $p(b,n+1,d) = p(b,n,d) = 1$
- > (in other words:  $p$  grows with  $n$ , then becomes stationary with  $p=1$ )
- >
- > assuming  $n \leq 2^b$
- > [5]  $p(b,n,0) = 1 - \prod (1 - j/2^b)$
- >  $j=0..n-1$

How can there be a nonzero probability of "two distinct vectors which differ by at most" 0 bits?

- >
- > assuming  $n > 1$
- > [6]  $p(b,n,b) = 1$
- >
- > assuming  $n > 1$
- > [7]  $p(b,n,d+1) > p(b,n,d)$  or  $p(b,n,d+1) = p(b,n,d) = 1$
- > (in other words:  $p$  grows with  $d$ , then becomes stationary with  $p=1$ )

sci.crypt: Re: minimum Hamming distance among random bit vectors

- >
- >
- > *So far I fail to find a workable technique to exactly compute*
- >  *$p(n,b,d)$  in the general case. I wonder if in the domain*
- >  *$2 \leq n \leq 2^{(b/3)}$  a valid approximation could be:*
- > *[8]  $p(b,n,d) \sim 1 - (1 - p(b,2,d))^{(n-1)/2}$*
- >
- > *I also fail to characterise*
- >  *$N(b,d) = \min(n \text{ such that } p(b,n,d)=1)$*

Of course  $N(b,d) \leq (2^b)/v(b,d)+1$ , where  $v(b,d)$  is the sum for  $k=0$  through  $d$  of  $C(b,k)$ . The equality holds at least at  $N(7,1)=16+1$ , from the error correction code posted not too long ago.

- >  *$D(b,n) = \min(d \text{ such that } p(b,n,d)=1)$*
- > *(in other word: when it is impossible to find  $n$  vectors which*
- > *all differ by at least  $d$  bits)*
- >
- > *Any idea or reference ?*

—Mike Amling