Non - Responsive

P.L. 86-36

EO 1.4.(c)
EO 1.4.(d)
P.L. 86-36

The Strength of the Bayes Score*

BY

~~Secret~~

The sigmage of the Bayes score is derived under the assumption that the score is normally distributed in right and wrong cases. Asymptotically there is a constant scoring rate per bit, and that rate is determined. Textlengths needed to attain certain sigmages for common attacks are calculated The authors verify the accuracy of these textlength calculations (given the validity of the underlying mathematical model).

1. INTRODUCTION

It is well known that the sigmage of the approximate Bayes score for a regularly stepping machine (number of standard deviations between right and wrong case means) is equal to $\sqrt{\alpha}\ T^{1/2}$ (α the expected value of the square of the putative bulges, T the textlength).

*Originally S12 Informal No. 283 of 8 September 1970, this paper won First Prize in the 1971 Crypto-Mathematics Institute Essay Contest.

EO 1.4.(c)
P.L. 86-36
EO 1.4.(d)

NSAL-S-199,795

~~GROUP 1~~
~~Excluded from automatic downgrading and declassification~~

[redacted]

(1) We prove for a lower bound model that the expected value of the factor in the right case has the form Cb^T. Our model is essentially an exact model, for although the score on T bits is weaker than the Bayes score on T bits, it is stronger than the Bayes score on $T/2$ bits.

[redacted]

(3) We are able to determine C [redacted] The value of C is important. Were it substantially larger or smaller than 1, [redacted] to the textlengths employed in the attack.

(4) Given the usual assumptions about scores being normally distributed, we have been able to calculate the sigmage of the Bayes score [redacted]. Before the asymptotic effect takes place, a better approximation to the sigmage is given by [redacted]

(5) We are able to do something else which [redacted] and that is to have some control over the accuracy of our estimates. Within our model there is an exact answer (given α and T) as to what the expected value of the factor in the right case is, which we temporarily will call E [redacted]

We are able to convert the statements (4) and (5) into quasi-practical COMSEC results at the end of section VI, where we list the textlengths needed to achieve certain sigmages for both Bayes and [redacted] with certain assumptions. These tables are presented with five reservations, one of which (the third) is analyzed in section VII. The significance of the other four is left as the subject for the further research. In section VII, where the statements we make in I, (5) are proved, it is also proved that the tables of section VI are quite

[redacted]

accurate (within the mathematical model that we have set up).

We summarize the asymptotic results [redacted] It is well-known that the Bayes score will always require less textlength than the [redacted] score. Our results indicate that the sigmage obtained from a textlength of T with a [redacted] score can be attained with a textlength of approximately [redacted] This result is especially important for primary attacks where α is small and the required textlength for [redacted] scoring might not be available. The Bayes score will never be cheaper to compute than the [redacted] score, for the work involved in calculating the former on a textlength of T is on the order of [redacted] rather than T. A measure of the "efficiency" of [redacted] over Bayes is thus given by U^2/T, which our results show to be the constant [redacted] this factor times the work needed to attain a certain [redacted] sigmage gives the work needed to attain a certain Bayes sigmage. [redacted]

We have not attempted to describe a specific COMSEC situation to which these results apply for the following reason, which, in fact, substantially limits the practical value of our findings. [redacted] The important question left unanswered by this paper, and one which, it is hoped, will be the subject of a future one, is the extent to which the Bayes scores [redacted] If, for example, it turns out that for the Bayes score one [redacted] Blenkin [redacted]

II. BAYES [redacted]

[redacted]

We think of $\{K_t\}$ as the observations from a sequence of independent Bernoulli random variables $\{X_t\}$. We are then asked to choose between two conflicting hypotheses, H_0 and H_1, where H_1 is the hypothesis that Prob$[X_t = 0] = 1/2$ for all $1 \le t \le T$. H_0 is a bit more complicated. We define a probability function P on the set S^T (all T-long sequences from S) so that if $Y = \{Y_t\} \in S^T$

[redacted]

To test which of the two hypotheses is true, a scoring function $\mathcal{S}$ (from S^T to the reals) is proposed and a threshold U is set so that H_0 is accepted if $\mathcal{S}(K) \geq U$ and H_1 is accepted if $\mathcal{S}(K) < U$. There are, of course, two possibilities for error: we may accept H_0 when H_1 is true (type II error) or accept H_1 when H_0 is true (type I error). The celebrated Neyman-Pearson Lemma (see, e.g., [2], p. 65) suggests that the "best" score is

$$\mathcal{S}_1(K) = \frac{\text{Prob}[K \mid H_0]}{\text{Prob}[K \mid H_1]}$$

in the following sense: if $\mathcal{S}_2$ is some other scoring function, and thresholds U_1 and U_2 are chosen for the respective scores so that the probabilities of type I error $= \alpha$, then the probability of a type II error using $\mathcal{S}_1$ is less than or equal to that using $\mathcal{S}_2$. What these thresholds are, and how small the probabilities of type II error then become, depend on knowledge of the distribution of the scoring function.

The scoring function $\mathcal{S}_1$ is commonly called the *Bayes factor*, since in order to obtain *a posteriori* odds in favor of H_0 from the *a priori* odds, one multiplies by $\mathcal{S}_1(\{K_i\})$ (in particular,

$$\frac{\text{Prob}[H_0 \mid K]}{\text{Prob}[H_1 \mid K]} = \mathcal{S}_1(K) \cdot \frac{\text{Prob}[H_0]}{\text{Prob}[H_1]}\Bigg) .$$

A unique "best" scoring function does not exist. In fact, it is easy to see that if f is monotonic increasing, the composite of f with $\mathcal{S}_1$ is as

"good" as $\mathcal{S}_1$ in the sense described above. In particular, a common scoring function, as good as the Bayes factor, is $\mathcal{S}_2(K) = \log_e(\mathcal{S}_1(K))$, commonly called the *Bayes score*. For the all-important example we described in our first paragraph, $\mathcal{S}_1(K) = 2^T \sum_{Y \in S^T} P(Y) \cdot \text{Prob}[K \mid Y]$,

[redacted] It would appear from this formula that $\mathcal{S}_1$ is difficult to evaluate; the computations can be arranged, however, so that their number is on the order of T^2 (see [3]).

[redacted] proposed a third score, an approximation to the Bayes score which we hereafter refer to as the [redacted] score. The [redacted] *score* $\mathcal{S}_3$ is defined as

[redacted]

where

[redacted]

$$= \binom{r}{t-r-1} 2^{-r}.$$

If many hypotheses H_0 are to be tested, the computation

$$\sum_{t=0}^{T} (-1)^{K_t} P_{t,r}$$

can be done as a one-time job for every $1 \leq r \leq T$, and the number of computations to evaluate the [redacted] score is on the order of T. In some applications this factor can be reduced to $\log_2 T$, [5] and in others the [redacted] might reduce the cost of calculating the [redacted] score, [redacted] could substantially reduce the cost of calculating the Bayes factor and score. [6]

A great deal is known about the [redacted] score. It can be shown that as T becomes large, the score [redacted]ally distributed, [7] with parameters

H_0 true, mean $= k\,\alpha\,T^{1/2}$
standard deviation [redacted]
H_1 true, mean $= 0$
standard deviation [redacted]

[redacted]

number [redacted] is also the number of standard deviations between the means, and from this it is easy to derive the probability of a type II error for a given type I error (because the scores are normally distributed). [redacted]

III. THE UNDILATED BAYES FACTOR

Our intention was to calculate mean and standard deviation of the Bayes factor, but this is complicated by the peculiar nature of the probability function P. It is possible that the techniques we employ in sections IV and V could be modified to apply to the exact Bayes factor, but the calculations would certainly be more cumbersome. For this reason we propose a fourth and fifth score, which we shall define shortly. We do not expect that these scores should ever be calculated for an actual key stream and hypothesis H_0, for the work would be comparable to the work in calculating the exact Bayes factor, and we know with certainty that the exact Bayes factor is the better score; however, the scores which we will introduce are closer in spirit to the Bayes score than is the [redacted] score, and we will be at least partially successful in calculating their means and variances.

[redacted]

[redacted]

Our notation is awkward in that the second usage of m depends on the $\tilde{Y}$ currently being summed over, but the typist has already been overworked and the meaning of m here and later should always be clear from context. [redacted] This suggests that we define a new score $\mathcal{S}_4$ by

[redacted]

which we call the [redacted] *Bayes factor*, and

$$\mathcal{S}_5(K) = \log_e \mathcal{S}_4(K),$$

the [redacted] *Bayes score*. $\mathcal{S}_4$ is intuitively a weaker scoring function because it ignores the effect [redacted] it is formally weaker because of the Neyman-Pearson Lemma.

IV. EXPECTED VALUE OF THE BAYES FACTOR

From this point on the terms Bayes factor and Bayes score will refer to the [redacted] factor and score, and we set $N = T/2$. If H_1 is true, then for each $1 \le t \le N$, $\text{Prob}[K_{m,t} = 0] = 1/2$, so

[redacted]

and $E[\mathcal{S}_4\{K_t\}] = 1$.

Next suppose that H_0 is true and let [redacted] Let $\{Z_t\} \in S^N$. For $1 \le t \le N$ we say that the pair of sequences $\{\tilde{Y}_t\}$ and $\{Z_t\}$ match at t if [redacted] the sequences $\{\tilde{Y}_t\}$ and $\{Z_t\}$ [redacted] If there is a match at t, then Prob$[K_{m,t} = 0] = (1+\epsilon_t)/2$ and [redacted] If there is no match at t, then [redacted] if we are willing to assume that $E[\epsilon_t] = 0$, we get that in either case [redacted] Setting $b = 1 + E[\epsilon_t^2]$, it follows that $E[\prod_{t=1}^{N} (1 + (-1)^{K_{m,t}} Z_t^* \epsilon_t)] = b^k$, where k is the number of matches between $\{\tilde{Y}_t\}$ and $\{Z_t\}$. If we let $p(N, k)$ be the probability that a pair of randomly selected elements of S^N will have k matches, then

EO 1.4.(c)
EO 1.4.(d)

[redacted]

We devote the rest of this section to the calculation of the right hand expression, which we denote by $E(N,b)$.

Let $(U,V) \in S^N \times S^N$, with $U = (U_1, \ldots, U_n)$ and $V = (V_1, \ldots, V_n)$. Let $Q(N,k)$ = number of elements (U,V) with exactly k matches, so [redacted] N − 1 long sequences. Hence

$$L(N,k,l) = Q(N-1,k,l) \text{ for } 1 \le l \le N-1. \quad (1)$$

[redacted] Because $L(N,k,N) = 0$ for $k \ge 1$, this together with (1) yields

[redacted]

We next define a triangular array of integers, $M(l,i)$, for $0 \le l \le i$, as follows:

$$M(0,0) = 1$$
$$M(0,i) = 0$$
$$M(l,i) = \sum_{j=l-1}^{i-1} M(j,i-1). \quad (3)$$

[redacted]

Some values of $M(l,i)$ are

i/l	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	0	1									
2	0	1	1								
3	0	2	2	1							
4	0	5	5	3	1						
5	0	14	14	9	4	1					
6	0	42	42	28	14	5	1				
7	0	132	132	90	48	20	6	1			
8	0	429	429	297	165	75	27	7	1		
9	0	1430	1430	1001	572	275	110	35	8	1	
10	0	4862	4862	3432	2002	1001	429	154	44	9	1.

P.L. 86-36

The importance of the $M(l,i)$ is that

[redacted] (4)

We prove this by induction on N. The result is trivial for $N = 0$. For $N = 1$ it becomes

[redacted]

Since the result clearly holds for $l = N$ and $l = 0$, we may assume $1 \le l \le N-1$ and argue that

[redacted]

Let us count the $Q(N,k)$ pairs of N sequences with k matches in terms of their initial bits. There are

[redacted]

Adding these together and applying (4), we get for $k \ge 1$

[redacted]

or translating into "p" notation

[redacted]

This last formula is also valid only for $k \geq 1$. If we give $p(0,k)$ its natural interpretation [redacted] we can trivially generalize the above formula to all k by writing

[redacted] (5)

[redacted] In fact, since

$$1 = \sum_k p(N,k)$$

[redacted]

it follows that [redacted]

We next find a closed form expression for $M(1,i)$, which we for notational simplicity refer to as $M(i)$. The $M(i)$ possess a convolution property

$$M(i) = \sum_{j=1}^{i-1} M(j)\, M(i-j). \qquad (6)$$

To prove (6), we observe from (3) that

$$M(l,i) = M(l-1,i-1) + M(l+1,i) \qquad 1 \leq l \leq i,\ i \geq 2.$$

Using this, we can prove by induction that

$$M(l,i+l) = \sum_{j=1}^{l} M(j+1,i+j) \qquad l \geq 1,\ i \geq 1. \qquad (7)$$

Now, by repeated use of this equation, we get

$$M(i) = M(i-1) + \sum_{j=2}^{i-1} M(j,i-1)$$

$$= M(1)\,M(i-1) + \sum_{j=1}^{i-2} M(j+1,j+1)\, M(j+1,i-1)$$

(since $M(j,j) = 1$)

EO 1.4.(c)
EO 1.4.(d)

P.L. 86-36

$$= \ldots$$

$$= M(1)\,M(i-1) + \ldots + M(l)\,M(i-l) + \sum_{j=1}^{i-l-1} M(j+1,j+l)\, M(j+1,i-l)$$

$$= M(1)\,M(i-1) + \ldots + M(l)\,M(i-l) + \sum_{j=1}^{i-l-1} M(j+1,j+l) \sum_{k=j}^{i-l-1} M(k,i-l-1) \quad \text{(by (3))}$$

$$= M(1)\,M(i-1) + \ldots + M(l)\,M(i-l) + \sum_{k=1}^{i-l-1} \left(\sum_{j=1}^{k} M(j+1,j+l) \right) M(k,i-l-1)$$

$$= M(1)\,M(i-1) + \ldots + M(l)\,M(i-l) + \sum_{k=1}^{i-l-1} M(k,l+k)\, M(k,i-l-1) \quad \text{(by (6))}$$

$$= M(1)\,M(i-1) + \ldots + M(l+1)\,M(i-l-1) + \sum_{k=1}^{i-(l+1)-1} M(k+1,k+l+1) \cdot M(k+1,i-(l+1))$$

$$= \ldots$$

$$= M(1)\,M(i-1) + \ldots + M(i-1)\,M(1).$$

Formula (6) having been verified, it follows that the power series

$$g(x) = 1 + \sum_{i=1}^{\infty} \frac{-M(i)}{2 \cdot 4^{i-1}}\, x^i,$$

which converges in some open neighborhood about the origin, has the property that $[g(x)]^2 = 1 - x$, so it follows from the generalized binomial theorem, since $g(x) = (1-x)^{1/2}$, that

$$\frac{M(i)}{4^{i-1}} = 2(-1)^{i-1} \binom{1/2}{i}. \qquad (8)$$

Recalling that the a_N appearing in (5) had the property

[redacted] (9)

EO 1.4.(c)
EO 1.4.(d)

we prove by induction that

[redacted] (10)

For $N = 1$, (10) is clear. Assume it holds for $N - 1$. Then

[redacted]

(by (9))

EO 1.4.(c)
EO 1.4.(d)

(by (9))

(by induction hypothesis)

(by (8)).

Combining (5), (8) and (10) we get that, for $N \geq 0$, $k \geq 0$

[redacted] (11)

We now let $h(x,y)$ be the power series in two indeterminants defined by $h(x,y) = \sum p(N,k)\, x^N y^k$ (which converges in the region $|x| < |y|^{-1}$). It follows from (11) and the fact that $\sum_{i=0}^{\infty} \binom{1/2}{i} (-1)^i x^i = (1-x)^{1/2}$ that

$$g(x,y) = h(x,y) - 1/4\, xyh(x,y) + ((1-x)^{1/2} - 1)(h(x,y) - 1) + 1/4\, xh(x,y) = \sum_{N,k} g_{N,k}\, x^N y^k$$

has the property that

$$g_{N,k} = 0 \qquad k \neq 0$$

EO 1.4.(c)
EO 1.4.(d)

$$g_{N,0} = -2N \binom{1/2}{N} (-1)^N, \; N \neq 0.$$

But $-\sum_N 2N \binom{1/2}{N} (-1)^N x^N = -2x \sum N \binom{1/2}{N} (-1)^N x^{N-1}$

$= -2x \frac{d}{dx} [\sqrt{1-x}] = \frac{x}{\sqrt{1-x}}$ so we get that $h(x,y) - 1/4xyh(x,y)$

$+ (\sqrt{1-x} - 1)(h(x,y) - 1) + 1/4\, x\, h(x,y) - \frac{x}{\sqrt{1-x}}$ is a constant,

which is in fact 1, since the last four summands have a zero constant term. Solving for h, we get that

$$h(x,y) = \frac{1}{\sqrt{1-x}\left(\sqrt{1-x} - \frac{x(y-1)}{4}\right)}$$

$$= \frac{1 + \frac{x(y-1)}{4\sqrt{1-x}}}{1 - x - x^2\left(\frac{y-1}{4}\right)^2}.$$

[redacted]

$$h(x,y) - xh(x,y) - x^2\left(\frac{y-1}{4}\right)^2 h(x,y) = 1 + \frac{x(y-1)}{4\sqrt{1-x}},$$

[redacted]

Even more important, $h(x,b) = \sum_N E(N,b)\, x^N$.

Let

[redacted]

[redacted]

Since $(1-\theta x)^{-1} = \sum_N \theta^N x^N$, comparing coefficients yields

[redacted]

Let us look at the factor [redacted] in the third term. [redacted] we can rewrite our previous equation, getting

[redacted]

From (12) it is easy to see that [redacted] in the sense that the percentage error goes to 0 as N gets large. More precisely:

[redacted] (13)

(In section 7 we will prove the stronger result

[redacted] (14)

[redacted] able to estimate the accuracy of the [redacted] We here prove (13). It is obvious that [redacted] vergence of $\sum_{i=0}^{\infty} \binom{-1/2}{i} (-\theta_1^{-1})^i$ that $\lim_{N\to\infty} \sum_{i=N-1}^{\infty} \binom{-1/2}{i} (-\theta_1^{-1})^i = 0$. Finally, pick some θ_3 with $\frac{\theta_2}{\theta_1} < \theta_3 < \theta_2$. Then

[redacted] the series on the right hand side converges, and since $\theta_2 < \theta_1\theta_3$, $\lim_{N\to\infty}$ of the right hand side is 0.

EO 1.4.(c)
EO 1.4.(d)

V. VARIANCE OF THE BAYES FACTOR

Recalling the definition of $\mathfrak{S}_4(K)$, we see that $\mathfrak{S}_4^2(K)$ is equal to

[redacted]

We first take up the case where H_1 is true. Then if Z and Z' match at [redacted]

and otherwise the expected value is 1/4. It follows that the value of $E[\mathfrak{S}_4^2(K)]$ is given by

[redacted]

We have illustrated by this argument that the variance of the score when H_1 is true equals the mean of the score when H_0 is true. This is actually true in general for any score of the form Prob $[K|H_0]$ / Prob $[K|H_1]$ and can be proved by an elementary argument (see, e.g., [9]).

Next we do the case where H_0 is true. [redacted]

It follows from this and (a)-(c) that

[redacted]

where k is the number of matches among $\tilde{Y}$, Z and Z' (if the three sequences match at t this counts as three matches), and thus

[redacted]

where $q(N,k)$ is the probability that three N-long sequences chosen at random will have exactly k matches. The authors have been unable to calculate the density function q; the problem seems similar to that of calculating the function p of chapter IV, but involves three-dimensional arrays rather than the two-dimensional array $M(i,j)$.

Without doing any hard work we can get a lower bound on the variance. Let A be the random variable representing the number of matches between the first two sequences, B the number of matches between the first and third, D the number of matches between the second and third. A, B and D are identically distributed but patently not independent. [redacted]

The authors have been unable to come up with a useful upper bound. [redacted]

bound turns out to be useless for the considerations made in the following section; so the existence of an effective upper bound is still open.

Recalling that [redacted] let us calculate more explicitly the means and variances we have derived for small α. We have

[redacted]

EO 1.4.(c)
EO 1.4.(d)

Summarizing the results of sections 4 and 5 (the reader should check that the C and θ_i corresponding to b' are approximately the same as those corresponding to b), we get

	H_0 true		H_1 true
mean of $\mathfrak{S}_4$	[redacted]		1
variance of $\mathfrak{S}_4$	[redacted]	(upper bound)	[redacted],

since $E[\mathfrak{S}_4^2(K)]$ dominates $(E[\mathfrak{S}_4(K)])^2$.

VI. SIGMAGE OF THE BAYES SCORE

It has become part of the COMSEC folklore that log factors tend to be normally distributed, especially if scoring rates per bit are constant and low. For a further discussion of this principle, see [9]. Throughout the rest of this section, we assume that $\mathfrak{S}_5$ is normally distributed when either H_0 or H_1 is true. We seek the mean and standard deviation in the two cases.

Let μ' and σ' be the mean and standard deviation of S_4. It follows from the form of the moment generating function of a normal distribution (again, see [9]) that

$$\ln \mu' = \mu + \frac{\sigma^2}{2}$$

$$\ln(\sigma')^2 = 2\mu + 2\sigma^2 .$$

We have seen that H_1 is true, $\mu' = 1$ and [redacted] whereas when H_0 is true [redacted] Solving the simultaneous equations, we get:

H_1 true [redacted]

H_0 true [redacted]

For very large T this gives us the sigmage [redacted] referred to in section 1. The effect of $\ln 2$, being a positive number, is to keep the sigmage greater than [redacted]

Since there is a constant scoring rate per bit asymptotically, there is little advantage when T is large [redacted] We have calculated mean and variance for the score in the second case. We can approximate the first case by saying that the score we are deriving [redacted]

 [redacted]

[redacted]

To conclude this section we present tables which give for common values of α the textlength needed to attain certain sigmages using the formulas

[redacted]

The reader should understand the following reservations before using these tables:

(1) The textlength needed to attain a certain sigmage [redacted] Bayes score is greater than that needed for the true Bayes score, but not more than twice as much.

(2) The values we have listed in the Bayes column are only lower bounds for the textlength needed within the mathematical model we have set up, due to the independence assumption which we made at the end of section V.

(3) The values listed in the Bayes column are accurate only to the extent to which the error term discussed in the next section is small.

(4) We have assumed that the Bayes score is normally distributed.

[redacted]

[redacted]

TABLE 3

[redacted]

[redacted] respectively, are similar to tables I and III. In section VII it will be shown that all entries in our tables are accurate to $\pm 0.1\sigma$ (that is, the Bayes textlength which we purport to give rise to a sigmage of n does indeed give rise to a sigmage of $n \pm 0.1$). Note that were an entry for 1σ in the Bayes column computed, it would have been a negative textlength; this absurdity can be explained by [redacted] the property that [redacted]

VII. ACCURACY

We have seen in section IV that [redacted] We will prove in this section that

[redacted]

and in fact get an upper bound for $F(N)$.

We refer to equation (12) in section IV. Since $|\theta_2| < 1$, it is clear that

[redacted]

We next estimate

[redacted]

By comparing

$$\binom{m}{n} = (-1)^n \prod_{k=1}^{n} \left(1 - \frac{m+1}{k}\right) \qquad (15)$$

with Weierstrass's infinite product expansion for the reciprocal of the gamma function

$$\frac{1}{\Gamma(x)} = x\, e^{\delta x} \prod_{k=1}^{\infty} \left(1 + \frac{x}{k}\right) e^{-\frac{x}{k}} \qquad (\delta = \text{Euler's constant})$$

and taking $x = -m-1$ it is easily seen that

$$\lim_{N\to\infty} \binom{m}{n} N^{m+1} (-1)^N = \frac{1}{\Gamma(-m)}.$$

In particular

$$\lim_{N\to\infty} (-1)^N \binom{-1/2}{N} N^{1/2} = \frac{1}{\Gamma(1/2)} = \frac{1}{\sqrt{\pi}} = \lim_{N\to\infty} \sqrt{N} \prod_{k=1}^{N} (1-1/2k). \qquad (16)$$

It is clear that $\left|\binom{-1/2}{i}\right|$ decreases as i becomes larger; so

[redacted]

The sequence $\sqrt{N} \prod_{k=1}^{N} \left(1 - \frac{1}{2k}\right)$ is increasing, since

$$\sqrt{N}\left(1 - \frac{1}{2N}\right) > \sqrt{N-1}.$$

Hence,

$$\sqrt{N+1} \prod_{k=1}^{N+1} \left(1 - \frac{1}{2k}\right) \le \frac{1}{\sqrt{\pi}} \qquad \text{by (16)},$$

and

[redacted]

Next we estimate

[redacted]

EO 1.4.(c)
EO 1.4.(d)

Although the series does not converge, by Taylor's theorem with a remainder we can write

[redacted]

where

$$R_N = \frac{\binom{-1/2}{N+1} (-\theta_2^{-1})^{N+1}}{(1 - y\,\theta^{-1})^{1/2+N+1}}$$

for some y with $0 < y < 1$. Since

[redacted]

we need only estimate $\theta_2^N\, R_N$. We have

$$|\theta_2^N \cdot R_N| = \text{[redacted]}$$

[redacted] (by 16)

sinc[redacted] This proves that

[redacted]

and thus that

$$F(N) = O\left(\frac{1}{\sqrt{N}}\right).$$

By using formula (12) and our previous estimates we can get a more precise upper bound for $F(N)$, namely:

[redacted]

EO 1.4.(c)
EO 1.4.(d)

[redacted]

For small α

[redacted]

[redacted]

EO 1.4.(c)
EO 1.4.(d)

P.L. 86-36

REFERENCES

[1] [redacted] "The Bayes Factor In Probabilistic Motion," *L91 Informal Note No. 92* (2 March 1966).

[2] E. L. Lehmann, *Testing Statistical Hypotheses* (New York, Wiley and Sons, 1959).

[3] [redacted] "Exact Computation of Bayes Score in Probabilistic Motion Attacks," *S12 Informal Note No. 255* (29 July 1969).

[4] [redacted]

[5] [redacted] *S12 Informal Note No. 276* (25 May 1970).

[6] [redacted] "Element of a Possible Analogue for Performing the Probabilistic Motion Attack," *L91 Informal Note No. 105* (27 June 1968).

[7] [redacted] "Two Scoring Functions and Their Distributions," *S12 Informal Note No. 212* (16 February 1968).

[8] [redacted]

[9] [redacted] "Multiple Hypothesis Testing and the Bayes Factor," *NSATJ*, XVI, 3 (Summer 1971) and *S12 Informal Note No. 291* (6 January 1971) (S-193,398).

EO 1.4.(c)
EO 1.4.(d)

EO 1.4.(c)

EO 1.4.(d)