Discussions

Statistical Treatment of Values of First 2,000 Decimal Digits of $\pi$ Calculated on the ENIAC

The first 2,000 decimal digits of $\pi$ were calculated on the ENIAC by Mr. G. Reitwiesner and several members of the ENIAC Branch of the Ballistic Research Laboratories at Aberdeen, Maryland (MTAC, v. 4, p. 11-15). A statistical survey of this material has failed to disclose any significant deviations from randomness for $\pi$, but it has indicated quite serious ones for $e$.

Let $D_n^i$ be the number of digits $i$ (where $i = 0, 1, \cdots, 9$) among the first $n$ digits of $e$ or of $\pi$. The count begins with the first digit left of the decimal point. If these digits were equidistributed, independent random variables, then the expectation value of each $D_n^i$ (with $n$ fixed and $i = 0, 1, \cdots, 9$) would be $n/10$, and the $\chi^2$ would be

$$a_n = \chi_n^2 = \sum_{i=0}^{9} \frac{(D_n^i - \frac{n}{10})^2}{\frac{n}{10}}.$$

The system of the $D_n^i$s (where $i = 0, 1, \cdots, 9$) has 9 degrees of freedom. Therefore let

$$p = P^{(k)}(a) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{1}{2}k\right)}\int_{0}^{a} e^{-\frac{1}{2}x}\frac{1}{\sqrt{2\pi}}d\chi$$

be the cumulative distribution function of $a = \chi^2$ for $k$ degrees of freedom. Then

$$p_n = P^{(9)}(a_n)$$

is a quantity which would be equidistributed in the interval $[0, 1]$, if the underlying digits were equidistributed independent random variables.

Consider $n = 2000$. In this case, the $D_n^i$s for $e$ are

(1) 196, 190, 208, 202, 201, 197, 204, 198, 202, 202.

Hence $a_n = \chi_n^2 = 1.11$ and $p_n = .0008$. The $D_n^i$s for $\pi$ are

(2) 182, 212, 207, 189, 195, 205, 200, 197, 202, 211.

Hence $a_n = \chi_n^2 = 4.11$ and $p_n = .096$.

The $e$-value of $p_{2000}$ is thus very conspicuous; it has a significance level of about 1:1250. The $\pi$-value of $p_{2000}$ is hardly conspicuous; it has a significance level of about 1:10.

The relevant fact about the distribution (1) appears upon direct inspection. The values lie too close to their expectation value, 200. Indeed their absolute deviations from it are

$$4, 10, 8, 2, 1, 3, 4, 2, 2, 2,$$

and hence their mean-square deviation is $22.2 = 4.71^2$, whereas in the random case the expectation value is $180 = 13.4^2$.

In order to see how this peculiar phenomenon develops as $n$ increases to 2000, $a_n = \chi_n^2$ and $p_n$ of $e$ have been determined from $D_n^i$ for the following smaller values of $n$.
These numbers show that the abnormally low value of \( p_n \) which is so conspicuous at \( n = 2000 \) does not develop gradually, but makes its appearance quite suddenly around \( n = 1900 \). Up to that point, \( p_n \) oscillates considerably and has a decreasing trend, but at \( n = 2000 \) there is a sudden dip of quite extraordinary proportions.

Thus something number-theoretically significant may be occurring at about \( n = 2000 \). A calculation of more digits of \( e \) would therefore seem to be indicated. A conversion to a simpler base than 10, say 2, may also disclose some interesting facts.

We wish to thank Miss Home McAllister of the ENIAC Branch of the Ballistic Research Laboratories for sorting the digital material on which the above analyses are based, and Professor J. W. Tukey, of Princeton University, for discussions of the subject.

Since the above was written (November 9, 1949), the ENIAC Branch of the Ballistic Research Laboratory very obligingly followed our suggestion and calculated the following 500 additional digits\(^1\) of \( e \). These should replace the last 10 digits of the value of \( e \) given in \( MTAC, v. 4, p. 15. \)

This makes it possible to extend the table of \( a_n = \chi_n^2 \) and \( p_n \) up to \( n = 2500 \)

\[
\begin{array}{ccc}
\hline
n & a_n = \chi_n^2 & p_n \\
\hline
500 & 6.72 & .33 \\
1000 & 4.82 & .15 \\
1100 & 5.93 & .25 \\
1200 & 4.03 & .093 \\
1300 & 3.83 & .080 \\
1400 & 4.74 & .145 \\
1500 & 3.69 & .070 \\
1600 & 2.47 & .019 \\
1700 & 3.22 & .046 \\
1800 & 2.85 & .031 \\
1900 & 2.22 & .013 \\
2000 & 1.11 & .0008 \\
\hline
\end{array}
\]
Thus the values of \( p_n \) for \( 2100 \leq n \leq 2500 \) are still significantly low but higher than the value of \( p_n \) at \( n = 2000 \).

Note that the general size and trend of \( p_n \), as well as its sudden deviation at \( n = 2000 \), indicate a non-random character in the digits of \( \pi \).

More detailed investigations are in progress and will be reported later.

Los Alamos Scientific Laboratory
N. C. Metropolis

Ballistic Research Laboratories
G. Reitwiesner

Institute for Advanced Study
J. von Neumann
Princeton, N. J.

1 Both \( \pi \) and \( 1/\pi \) were computed somewhat beyond 2500 D and the results checked by actual multiplication.

Notes on Numerical Analysis—2
Note on the Condition of Matrices

1. The object of this note is to establish the following theorem.

**Theorem.** Let \( A \) be a real \( n \times n \) non-singular matrix and \( A' \) be its transpose. Then \( AA' \) is more "ill-conditioned" than \( A \).

This theorem confirms an opinion expressed by Dr. L. Fox1 based on his practical experience. The term "condition of a matrix" has been used rather vaguely for a long time. The most common measure of the condition of a matrix has been the size of its determinant, ill-conditioned matrices being those with a "small" determinant. With this interpretation imposed, the theorem is clearly correct. More adequate measures of the condition of a matrix have been proposed recently by John von Neumann & H. H. Goldstine2 and by A. M. Turing.3 Their definitions concern all matrices, not just the ill-conditioned ones, characterized by very large condition numbers. The following two of these definitions will form a basis for the proof of the above-mentioned theorem:

The **P-condition number** is \( \frac{|\lambda_{\text{max}}|}{|\lambda_{\text{min}}|} \), where \( \lambda_{\text{max}} \) and \( \lambda_{\text{min}} \) are the characteristic roots of largest and smallest modulus.2

The **N-condition number** is \( N(A)N(A^{-1})/n \), where3

\[
N(A) = (\sum_{i,k} a_{ik}^2)^{1/2}
\]

2. Proof of the theorem in the P case:

Let \( \lambda_i \) be the characteristic roots of \( A \) and \( \mu_i \) those of \( AA' \) (which are in general distinct from the squares of the absolute values of \( \lambda_i \)). E. T. Browne4 has shown that

\[
\mu_{\text{min}} \leq \lambda_i \lambda_{i} \leq \mu_{\text{max}}.
\]

From this it follows that

\[
1 \leq \frac{|\lambda_{\text{max}}|}{|\lambda_{\text{min}}|} \leq \left( \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \right)^2 \leq \frac{\mu_{\text{max}}}{\mu_{\text{min}}},
\]

which implies the required result.

3. Proof of the theorem in the N case:

It is known that \( N(A) \) is the square root of the trace of \( AA' \) and therefore equal to \( (\sum \mu_i)^{1/2} \). The numbers \( \mu_i \) are all positive since \( AA' \) is symmetric and positive definite. Since the characteristic roots of \( A'A \) and \( AA' \) are