

AFCRL-64-737
SEPTEMBER 1964

AD 606 556

**Physical and Mathematical Sciences Research Papers
No. 48**

**Confidence Levels for the Sample Mean and
Standard Deviation of a Rayleigh Process**

LEO M. KEANE

**MICROWAVE PHYSICS LABORATORY PROJECT 4600
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES, OFFICE OF AEROSPACE RESEARCH, UNITED STATES AIR FORCE, L.G. HANSCOM FIELD, MASS.**

Abstract

The number of independent samples necessary to characterize the parameters of a Rayleigh-distributed process within arbitrary confidence limits is derived. Stationarity of the process is assumed, along with convergence of the Central Limit theorem with regard to the probability density function of the sample mean. Perfect measurement ability is also assumed. A graph of sample size for a range of confidence coefficients and error limits is presented.

BLANK PAGE

Confidence Levels for the Sample Mean and Standard Deviation of a Rayleigh Process

1. INTRODUCTION

The Rayleigh probability density function is encountered frequently in communications and radio astronomy detection problems. For example, it characterizes the voltage at the output of an envelope detector when the input to the detector is a narrowband gaussian noise. The narrowband gaussian noise could be the result of the band-limiting characteristic of an IF strip when a broadband gaussian process is under observation. It is common to estimate the true mean value and standard deviation of such a process from samples of the continuous data taken at selected discrete instants of time.

In practice it is necessary to decide the number (n) of independent samples that must be taken before, for instance, an estimate of the mean value calculated from these samples is within a specific percentage error of the true mean value of the process. This can never be done with absolute certainty but must have associated with it a certain confidence level denoting the probability with which the mean value calculated from the n samples will fall within the error limit specified.

The number of samples necessary for the specific case of a 99% confidence level and a $\pm 1\%$ error limit is calculated in Section 2. The result is extended to the general case of a confidence level of $c\%$ and an error of $\pm e\%$, and a plot of the extended function presented graphically.

(Received for publication 27 July 1964)

The true mean value completely characterizes a Rayleigh probability density function. The standard deviation is proportional to this value. Moreover, if the measured mean value is within certain error limits of the true mean value, it can be shown that a "sample" standard deviation calculated from the measured mean value is within identical error limits of the true standard deviation of the process.

2. ESTIMATION OF MEAN VALUE

The mean value of n samples of any one of the infinite number of sample functions is called the sample mean, defined by

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} . \quad (1)$$

The y_i , the values of the n samples, have probability density functions $p(y_i)$, which are given as Rayleigh density functions. The assumption of stationarity assures us that the $p(y_i)$ are equal. Since \bar{y} is the mean value of n random variables, it too is a random variable. The probability density function for the mean value of a large number n of the y_i will, by the Central Limit theorem, converge as $n \rightarrow \infty$ on a normal or gaussian density function characterized by

$$p(\bar{y}) = \frac{1}{\sqrt{2\pi \text{Var}(\bar{y})}} \exp \left\{ -\frac{[\bar{y} - E(\bar{y})]^2}{2 \text{Var}(\bar{y})} \right\} , \quad n \rightarrow \infty . \quad (2)$$

It can be readily shown that

$$E(\bar{y}) = E(y_i)$$

and

$$\text{Var}(\bar{y}) = \frac{\text{Var}(y_i)}{n} ,$$

or,

$$\text{SD}(\bar{y}) = \frac{\text{SD}(y_i)}{\sqrt{n}} .$$

The probability density function of \bar{y} can thus be approximated by

$$p(\bar{y}) = \frac{1}{\sqrt{2\pi \frac{\text{Var}(y_i)}{n}}} \exp \left\{ -\frac{[\bar{y} - E(y_i)]^2}{2 \frac{\text{Var}(y_i)}{n}} \right\} . \quad (3)$$

If we introduce

$$\mu = \frac{\bar{y} - E(y_i)}{\frac{SD(y_i)}{\sqrt{n}}}$$

a change of variable could show that

$$p(\mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2}\right) . \quad (4)$$

This is of course the standard tabulated normal probability density function. From the tables it can be seen that 99% of the area (or, the probability is .99 that μ) will fall between the values of ± 2.576 . Therefore,

$$\Pr \{-2.576 < \mu < 2.576\} = .99 ,$$

or,

$$\Pr \left\{ -2.576 < \frac{\bar{y} - E(y_i)}{\frac{SD(y_i)}{\sqrt{n}}} < 2.576 \right\} = .99 . \quad (5)$$

This can be rearranged into the form:

$$\Pr \left\{ \bar{y} - 2.576 \frac{SD(y_i)}{\sqrt{n}} < E(y_i) < \bar{y} + 2.576 \frac{SD(y_i)}{\sqrt{n}} \right\} = .99 . \quad (6)$$

If we want to assure ourselves with 99% confidence that the sample mean falls within $\pm 1\%$ of the true mean, we set

$$\Pr \{ .99 E(y_i) < \bar{y} < 1.01 E(y_i) \} = .99 , \quad (7)$$

which can be rearranged into the form:

$$\Pr \{ \bar{y} - .01 E(y_i) < E(y_i) < \bar{y} + .01 E(y_i) \} = .99 . \quad (8)$$

By equating the parts of Eq. (6) and (8) we find

$$.01 E(y_i) = 2.576 \frac{SD(y_i)}{\sqrt{n}} ,$$

or,

$$n = \frac{(2.576)^2 \text{Var}(y_i)}{(.01)^2 [E(y_i)]^2} . \quad (9)$$

For a process characterized by a Rayleigh density function,

$$p(y_i) = \begin{cases} \frac{y_i}{a^2} \exp\left(-\frac{y_i^2}{2a^2}\right) , & 0 < y_i < \infty ; \\ 0 , & -\infty < y_i < 0 . \end{cases} \quad (10)$$

Also,

$$E(y_i) = a \sqrt{\frac{\pi}{2}} ;$$

$$\text{Var}(y_i) = 2a^2 - \frac{a^2 \pi}{2} .$$

Therefore

$$\frac{\text{Var}(y_i)}{[E(y_i)]^2} = \frac{4}{\pi} - 1 . \quad (11)$$

Also,

$$\text{SD}(y_i) = \sqrt{\frac{4}{\pi} - 1} [E(y_i)] . \quad (12)$$

Combining Eqs. (9) and (11), we have

$$n = \frac{(2,576)^2 \left(\frac{4}{\pi} - 1\right)}{(.01)^2} \cong 17,900 . \quad (13)$$

Thus, for $n \cong 17,900$, the sample mean will fall within $\pm 1\%$ of the true mean with a 99% confidence level.

This result can of course be extended to the case of a $c\%$ confidence level and $\pm e\%$ error. Let $\pm m$ denote the values of the abscissa on the gaussian distribution between which the normalized variable μ falls in order to give $c\%$ confidence level or probability $c/100$. Then

$$\text{Pr} \{-m < \mu < m\} = \frac{c}{100} .$$

If the error is allowed to be $\pm e\%$, then Eq. (13) becomes

$$n = \frac{(m^2) \left(\frac{4}{\pi} - 1\right)}{\left(\frac{e}{100}\right)^2} . \quad (14)$$

Equation (14) is plotted in Figure 1 for values of c between 1% and 50% and values of e from $\pm 1\%$ to approximately $\pm 10\%$.

3. ESTIMATION OF STANDARD DEVIATIONS

Equation (12) defines a relationship between the true mean value and true standard deviation of a Rayleigh process. It is now shown that a "sample" standard deviation $\text{SSD}(y)$, defined in Eq. (15), will be within definable error limits of the

true process standard deviation.

Let

$$\text{SSD}(y) = \sqrt{\frac{4}{\pi} - 1} (\bar{y}) . \quad (15)$$

The extension of Eq. (8) to the general case is

$$\Pr \left\{ \bar{y} - \frac{e E(y_i)}{100} < E(y_i) < \bar{y} + \frac{e E(y_i)}{100} \right\} = \frac{c}{100} . \quad (16)$$

Substituting Eq. (12) into the left side of Eq. (16) and rearranging terms yields

$$\Pr \left\{ \frac{1-e}{100} \text{SD}(y_i) < \sqrt{\frac{4}{\pi} - 1} \bar{y} < \frac{1+e}{100} \text{SD}(y_i) \right\} = \frac{c}{100} .$$

or,

$$\Pr \left\{ \frac{1-e}{100} \text{SD}(y_i) < \text{SSD}(y) < \frac{1+e}{100} \text{SD}(y_i) \right\} = \frac{c}{100} . \quad (17)$$

In other words, if we multiply the measured sample mean by the factor

$$\left(\sqrt{\frac{4}{\pi} - 1} \approx .5227 \right)$$

we are assured that the "sample" standard deviation thus calculated is within the same error bounds of the true standard deviation as exist for the sample mean and the true mean. These error limits also exist under the same confidence conditions.

4. CONCLUSIONS

With 17,900 independent samples of a process known to be Rayleigh-distributed, we can be 99% confident that the measured-sample mean is within $\pm 1\%$ of the true mean. Similarly, 17,900 independent samples will assure that the standard deviation calculated from the sample mean is within $\pm 1\%$ of the true standard deviation. An extension of this conclusion to the general case is possible and can be obtained from Eq. (14).

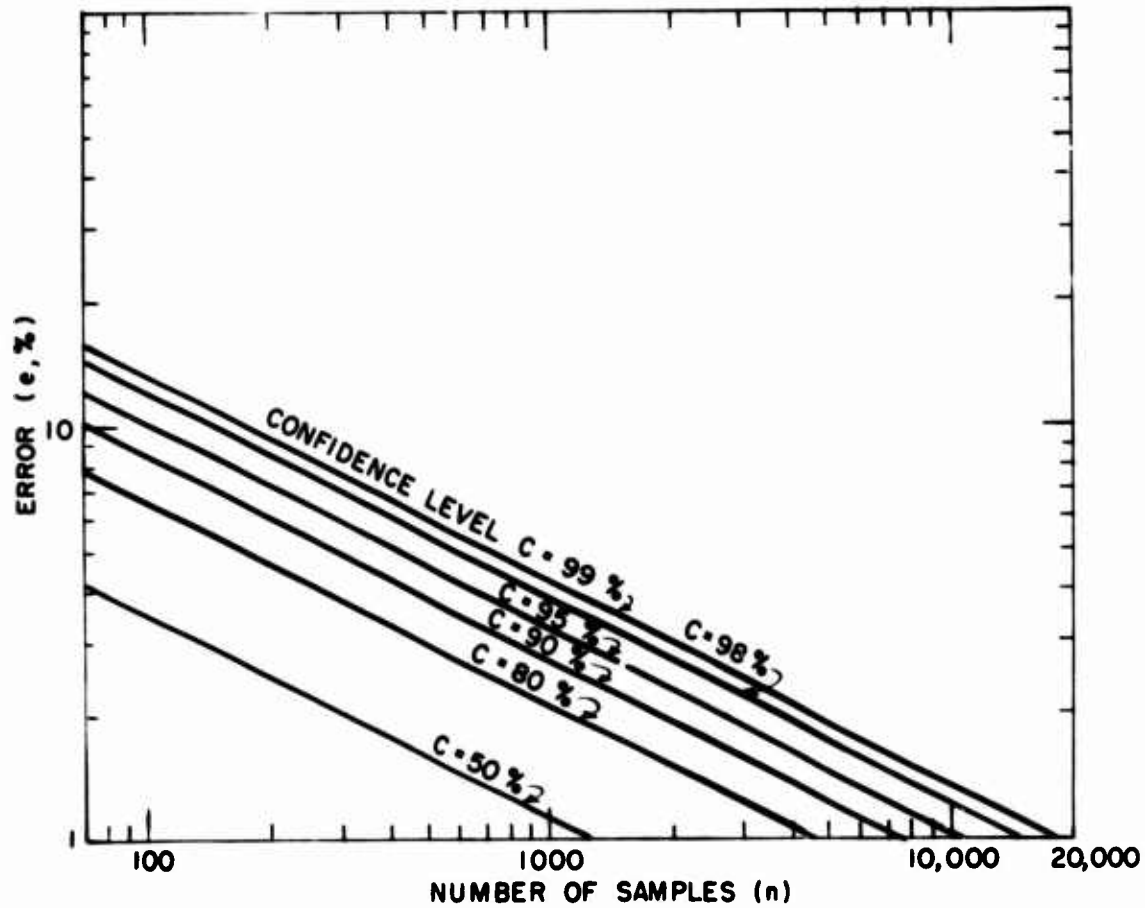


Figure 1. Error vs Number of Samples for Various Confidence Levels for the Sample Mean of a Rayleigh Process

Bibliography

Jerome C.R. Li, Introduction to Statistical Inference, Ann Arbor, Mich., 1957.
(Distributed by Edwards Brothers)

Acknowledgments

The author is indebted to Dr. S. Zahl (CRB) and Mr. J. Pierce (CRO) for many helpful comments during the course of the study; to Mr. R. Allen (CRF) for suggesting the problem; and to Mr. J. Pierce (CRO), Dr. E. Altshuler (CRD), and Mr. J. Short (CRD) for checking the report.

PHYSICAL SCIENCES RESEARCH PAPERS

- No. 1. Central-Force Laws for an Elliptic Orbit, *Kurt Toman, March 1964 (REPRINT)*.
- No. 2. Structure of 10, 10-Dibromanthrone, *J. Silverman, N. F. Yannoni, February 1964 (REPRINT)*.
- No. 3. Ion Dissociation in the Drift Tube of a Time-of-Flight Mass Spectrometer: V. Analytic Solutions of the Flight-Time Shift Equation, *W. W. Hunt, Jr., M. J. Kennedy, February 1964*.
- No. 4. Asymptotic Form of the Electron Capture Cross Section in the Impulse Approximation, *R. A. Mapleton, March 1964 (REPRINT)*.
- No. 5. Intelligibility of Excerpts From Fluent Speech: Effects of Rate of Utterance and Duration of Excerpt, *J. M. Pickett, Irwin Pollack, March 1964 (REPRINT)*.
- No. 6. Back-Scatter by Dielectric Spheres With and Without Metal Caps, *David Atlas, Kenneth M. Glover, March 1964 (REPRINT)*.
- No. 7. An Adaptive Filter for the Design of Ionospheric Disturbance Detectors (U), *Richard D. Smallwood, 1/Lt, USAF, February 1964 (SECRET)*.
- No. 8. The Nonlinear Interaction of an Electromagnetic Wave With a Time-Dependent Plasma Medium, *Robert J. Papa, April 1964*.
- No. 9. Drastic Reduction of Warm-up Rate Within a Dewar System by Helium Desorption, *Peter D. Gianino, January 1964*.
- No. 10. The Antipodal Image of an Electromagnetic Source, *Kurt Toman, April 1964 (REPRINT)*.
- No. 11. Radiation Forces in Inhomogeneous Media, *E.J. Post, April 1964 (REPRINT)*.
- No. 12. Progressive Failure Prediction, *Walton B. Bishop, April 1964 (REPRINT)*.
- No. 13. Visual Data Transmission, *Ronald J. Massa, 1/Lt, USAF, April 1964*.
- No. 14. Rydberg Absorption Series of N_2 , *M. Ogawa and Y. Tanaka, May 1964 (REPRINT)*.
- No. 15. 600-A Band of Helium, *Y. Tanaka and K. Yoshino, May 1964 (REPRINT)*.
- No. 16. Charge Transfer Studies With a Time-of-Flight Mass Spectrometer: II. Kinetic Analysis, Including Attenuation of Both Neutrals and Ions by Scattering, *W. W. Hunt, Jr., May 1964*.
- No. 17. Photo-Induced Electron Transfer in Dye-Sulphydryl Protein Complex, *Eiji Fujimori, May 1964, (REPRINT)*.
- No. 18. Intelligibility of Excerpts From Fluent Speech: Auditory vs. Structural Context, *Irwin Pollack and J.M. Pickett, May 1964, (REPRINT)*.
- No. 19. A Study of Transverse Modes of Ruby Lasers Using Beat Frequency Detection and Fast Photography, *C. Martin Stickley, May 1964*.
- No. 20. Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech, *Philip Lieberman, June 1964 (REPRINT)*.
- No. 21. Infrared Absorption of Magnesium Stannide, *Herbert G. Lipson and Alfred Kahan, June 1964 (REPRINT)*.
- No. 22. On the Optimum Design of Multipath Signals, *Neil J. Bershad, 1 Lt USAF, June 1964*.
- No. 23. Area Properties of Television Pictures, *S. Nishikawa, R.J. Massa, J.C. Mott-Smith, June 1964*.
- No. 24. A Geometric Study of Coherence Properties of Partially Polarized Electromagnetic Radiation, *E.F. Bolinder, June 1964*.
- No. 25. The Preparation of High-Purity Boron via the Iodide, *A.F. Armington, G.F. Dillon, and R.F. Mitchell, June 1964 (REPRINT)*.
- No. 26. An Interpretation of the Far-Field Effects of a Rocket in the Ionosphere (U), *Thomas D. Conley and James E. Higgins, June 1964 (SECRET)*.
- No. 27. A Radon-Nikodym Theorem in Dimension Lattices, *S.S. Holland, Jr., June 1964 (REPRINT)*.
- No. 28. Plasma Produced Antenna Pattern Distortion, *Daniel J. Jacarunco, June 1964*.
- No. 29. Geometry and First-Order Error Statistics for Three- and Four-Station Hyperbolic Fixes on a Spherical Earth, *Eduard A. Lewis, June 1964*.
- No. 30. Ion Dissociation in the Drift Tube of a Time-of-Flight Mass Spectrometer: III Flight-Time Shift Equations for Spurious Fragment Peaks Arising From Charge Transfer and Dissociation Reactions Occurring Inside the Potential Barrier, *W.W. Hunt, Jr., June 1964*.
- No. 31. Dolph-Tchebyscheff Arrays of Many Elements and Arbitrary Uniform Spacing, *Charles J. Drane, Jr., June 1964*.

PHYSICAL SCIENCES RESEARCH PAPERS (Continued)

- No. 32. Measurement of Noise Figure of an X-Band Waveguide Mixer with Tunnel Diode, *Gustav H. Blaeser*, July 1964.
- No. 33. Transient Reflection and Transmission of a Plane Wave Normally Incident Upon a Semi-Infinite Anisotropic Plasma, *Carl T. Case, 1 Lt, USAF*, July 1964.
- No. 34. Low-Temperature Far-Infrared Spectra of Germanium and Silicon, *Peter J. Gielisse, James R. Aronson and Hugh G. McLinden*, June 1964.
- No. 35. Absorption Coefficients of Carbon Monoxide in the 1006-600-A Wavelength Region, *R.E. Huffman, J.C. Larrabee and Y. Tanaka*, July 1964 (REPRINT).
- No. 36. Asymptotic Form of the Electron Capture Cross Section in First Born and Distorted Wave Approximations, *R.A. Mapleton*, July 1964 (REPRINT).
- No. 37. A Computer Approach to Laser Design, *T.G. Purnhagen and J. Lubelfeld*, July 1964 (REPRINT).
- No. 38. Apparent Sky Temperatures at Millimeter-Wave Frequencies, *Karl V. Wulfsberg*, July 1964.
- No. 39. Observation of 2,1 Charge Transfer in a TOF Mass Spectrometer (Text of a paper presented at the Southwestern Meeting of the American Physical Society at Tucson, Arizona, on 28 February 1964), *W.W. Hunt, Jr., and K.E. McGee*, July 1964.
- No. 40. PMR Bi-Static Results During the Period 13 August to 14 December 1962, *T.D. Conley*, July 1964 (SECRET).
- No. 41. EM Pulses From 1962 USSR Nuclear Tests, Extracted From Sferics Records (U), *A. Ganio and J.L. Heckscher, Capt, USAF*, July 1964 (SECRET-RD).
- No. 42. Dislocation Structures in Single-Crystal Al_2O_3 , *D.L. Stephens and W.J. Alford*, August 1964 (REPRINT).
- No. 43. Anomalies in VLF Signals Observed During High-Altitude Nuclear Tests, 1962(U), *Alma Ganio*, August 1964 (SECRET-RD).
- No. 44. Molecular Structure of 2-(4-amino-5'-azamethenyl pyrimidyl)-3 pentene-4-ol, *M.F. Yannoni and Jerry Silverman*, August 1964 (REPRINT).
- No. 45. Output Power from GaAs Lasers at Room Temperature, *C.C. Gallagher, P.C. Tandy, B.S. Goldstein, and J.D. Welch*, August 1964 (REPRINT).
- No. 46. Weight Distribution of the Quadratic Residue (71,35) Code, *Vera Pless*, August 1964.
- No. 47. On the Convergence and Ultimate Reliability of Iterated Neural Nets, *R.H. Urbano*, September 1964 (REPRINT).
- No. 48. Confidence Levels for the Sample Mean and Standard Deviation of a Rayleigh Process, *Leo M. Keane*, September 1964.

BLANK PAGE