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

LEO M. KEANE
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.
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 ± 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 ± \( e \% \), 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

\[ \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \]  \hspace{1cm} (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 \( n \to \infty \) on a normal or gaussian density function characterized by

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

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

If we introduce
\[ \mu = \frac{\bar{y} - E(y_i)}{\text{SD}(y_i) / \sqrt{n}} \]

A change of variable could show that

\[ p(\mu) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{\mu^2}{2} \right) . \]  

(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 \( \mu \)) will fall between the values of \( \pm 2.576 \). Therefore,

\[ \Pr \left\{ -2.576 < \mu < 2.576 \right\} = .99 , \]

or,

\[ \Pr \left\{ -2.376 < \frac{\bar{y} - E(y_i)}{\text{SD}(y_i) / \sqrt{n}} < 2.576 \right\} = .99 . \]  

(5)

This can be rearranged into the form:

\[ \Pr \left\{ \bar{y} - 2.576 \frac{\text{SD}(y_i)}{\sqrt{n}} < E(y_i) < \bar{y} + 2.576 \frac{\text{SD}(y_i)}{\sqrt{n}} \right\} = .99 . \]  

(6)

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

\[ \Pr \left\{ .99 E(y_i) < \bar{y} < 1.01 E(y_i) \right\} = .99 , \]

(7)

which can be rearranged into the form:

\[ \Pr \left\{ \bar{y} - .01 E(y_i) < E(y_i) < \bar{y} + .01 E(y_i) \right\} = .99 . \]  

(8)

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

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

or,

\[ n = \frac{(2.576)^2 \text{Var}(y_i)}{( .01 )^2 \left[ E(y_i) \right]^2} . \]  

(9)

For a process characterized by a Rayleigh density function,

\[ p(y_i) = \begin{cases} \frac{y_i}{\alpha} \exp \left( -\frac{y_i^2}{2\alpha^2} \right) , & 0 < y_i < \infty ; \\ 0 , & -\infty < y_i < 0 . \end{cases} \]  

(10)
Also,

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

\[ \text{Var}(y_i) = 2a^2 - \frac{a^2}{2} \; . \]

Therefore

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

(11)

Also,

\[ \text{SD}(y_i) = \sqrt{\frac{4}{\pi} - 1} \left[ E(y_i) \right] \; . \]  

(12)

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

\[ n = \frac{(2,576)^2 \left( \frac{4}{\pi} - 1 \right)}{(.01)^2} \approx 17,900 \; . \]  

(13)

Thus, for \( n \approx 17,900 \), the sample mean will fall within ±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 ±\( e\% \) error. Let ±\( m \) denote the values of the abscissa on the gaussian distribution between which the normalized variable \( \mu \) 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 ±\( e\% \), then Eq. (13) becomes

\[ n = \frac{(m^2) \left( \frac{4}{\pi} - 1 \right)}{\left( \frac{e}{100} \right)^2} \; . \]  

(14)

Equation (14) is plotted in Figure 1 for values of \( c \) between 1% and 50% and values of \( e \) from ±1% to approximately ±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 \( \text{SD}(y) \), defined in Eq. (15), will be within definable error limits of the
true process standard deviation.

Let

$$SSD(y) = \sqrt{\frac{\pi}{3}} - 1 (\bar{y}).$$

(15)

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

$$\Pr\left\{ y - \frac{E(y)}{100} < E(y) < y + \frac{E(y)}{100} \right\} = \frac{c}{100}.$$  

(16)

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

$$\Pr\left\{ \frac{1-c}{100} SD(y) < \sqrt{\frac{\pi}{3}} - 1 \bar{y} < \frac{1+c}{100} SD(y) \right\} = \frac{c}{100}.$$  

or,

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

(17)

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

$$\left( \sqrt{\frac{\pi}{3}} - 1 \approx 0.5227 \right)$$

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 17,000 independent samples of a process known to be Rayleigh-distributed, we can be 99% confident that the measured sample mean is within ±1% of the true mean. Similarly, 17,900 independent samples will assure that the standard deviation calculated from the sample mean is within ±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


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.
<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Authors</th>
<th>Date</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Central-Force Laws for an Elliptic Orbit</td>
<td>Kurt Toman</td>
<td>March 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>2</td>
<td>Structure of 10, 10-Dihomoanthracene</td>
<td>J. Silverman, N. F. Yannoni</td>
<td>February 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>4</td>
<td>Asymptotic Form of the Electron Capture Cross Section in the Impulse Approximation</td>
<td>R. A. Mapleton</td>
<td>March 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>5</td>
<td>Intelligibility of Excerpts From Fluent Speech: Effects of Rate of Utterance and Duration of Excerpt</td>
<td>J. M. Pickett, Irwin Pollack</td>
<td>March 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>6</td>
<td>Back-Scatter by Dielectric Spheres With and Without Metal Caps</td>
<td>David Atlas, Kenneth M. Glover</td>
<td>March 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>7</td>
<td>An Adaptive Filter for the Design of Ionospheric Disturbance Detectors(U)</td>
<td>Richard D. Smallwood, 1Lt USAF</td>
<td>February 1964</td>
<td>SECRET</td>
</tr>
<tr>
<td>8</td>
<td>The Nonlinear Interaction of an Electromagnetic Wave With a Time-Dependent Plasma Medium</td>
<td>Robert J. Papa</td>
<td>April 1964</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Drastic Reduction of Warm-up Rate Within a Dewar System by Helium Desorption</td>
<td>Peter D. Gianino</td>
<td>January 1964</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>The Antipodal Image of an Electromagnetic Source</td>
<td>Kurt Toman</td>
<td>April 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>11</td>
<td>Radiation Forces in Inhomogeneous Media</td>
<td>E. J. Post</td>
<td>April 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>12</td>
<td>Progressive Failure Prediction</td>
<td>Walton B. Bishop</td>
<td>April 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>13</td>
<td>Visual Data Transmission</td>
<td>Ronald J. Massa, 1 Lt, USAF</td>
<td>April 1964</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>An Intelligibility of Excerpts From Fluent Speech: Auditory vs. Structural Context, Irwin Pollack and J.M. Pickett</td>
<td>Irwin Pollack, J. M. Pickett</td>
<td>May 1964</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>A Study of Transverse Modes of Ruby Lasers Using Beat Frequency Detection and Fast Photography</td>
<td>C. Martin Stickley</td>
<td>May 1964</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Some Effects of Semantic and Grammatical Context on the Production and Perception of Speech</td>
<td>Philip Lieberman</td>
<td>June 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>17</td>
<td>Infrared Absorption of Magnesium Stanndide</td>
<td>Herbert G. Lipson, Alfred Kahn</td>
<td>June 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>18</td>
<td>On the Optimum Design of Multipath Signals</td>
<td>Neil J. Hershad, 1Lt USAF</td>
<td>June 1964</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>The Preparation of High-Purity Boron via the Iodide</td>
<td>A.F. Armington, G.F. Dillon, R.F. Mitchell</td>
<td>June 1964</td>
<td>REPRINT</td>
</tr>
<tr>
<td>21</td>
<td>A Radon-Nikodym Theorem in Dimension Lattices</td>
<td>S.S. Holland, Jr.</td>
<td>June 1964</td>
<td>SECRET</td>
</tr>
<tr>
<td>22</td>
<td>Plasma Produced Antenna Pattern Distortion</td>
<td>Daniel J. Jacobson</td>
<td>June 1964</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Geometry and First-Order Error Statistics for Three- and Four-Station Hyperbolic Fixes on a Spherical Earth</td>
<td>Edward A. Lewis</td>
<td>June 1964</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Dolph-Tchebycheff Arrays of Many Elements and Arbitrary Uniform Spacing</td>
<td>Charles J. Drake</td>
<td>June 1964</td>
<td></td>
</tr>
</tbody>
</table>
No. 32. Measurement of Noise Figure of an X-Band Waveguide Mixer with Tunnel Diode, Gustav H. Blaeser, July 1964.


No. 34. Low-Temperature Far-Infrared Spectra of Germanium and Silicon, Peter J. Gielisse, James R. Ammons and Hugh G. McIndoe, June 1964.


No. 36. Asymptotic Form of the Electron Capture Cross Section in First Born and Distorted Wave Approximations, R.A. Mynatt, July 1964 (REPRINT).


No. 40. PMR Hi-Static Results During the Period 13 August to 14 December 1962, T.D. Comley, July 1964 (SECRET).


No. 44. Molecular Structure of 24(1-amino-5-azamethylyprimimidyl)-3-pentene-1, N. Yannoni and James Silverman, August 1964 (REPRINT).

No. 45. Output Power from Coax Lasers at Room Temperature, C.G. Gallagher, P.C. Tandy, R.S. Goldstein, and J.H. Welsh, August 1964 (REPRINT).

No. 46. Weight Distribution of the Quadratic Residue (71,33) Code, Vera Pless, August 1964.

