Prathamesh Prabhudesai
Student Id: 200111145
Unity Id: ppprabhu

Random Processes
ECE 514
Fall 2015
Project

Instructor: Dr. Hamid Krim
EBII #2099
e-mail: ahk@@ncsu.edu
http://www4.ncsu.edu/people/ahk

Traditionally, a fish population has been estimated by counting the number of fish in a net catch. That is WAY TOO EXPENSIVE AND INACCURATE. A modern and better approach is to use an ECHO SONAR. In echo sonar, sound pulses are transmitted, and the reflected waveform from a school of fish are modeled as a Gaussian random process.

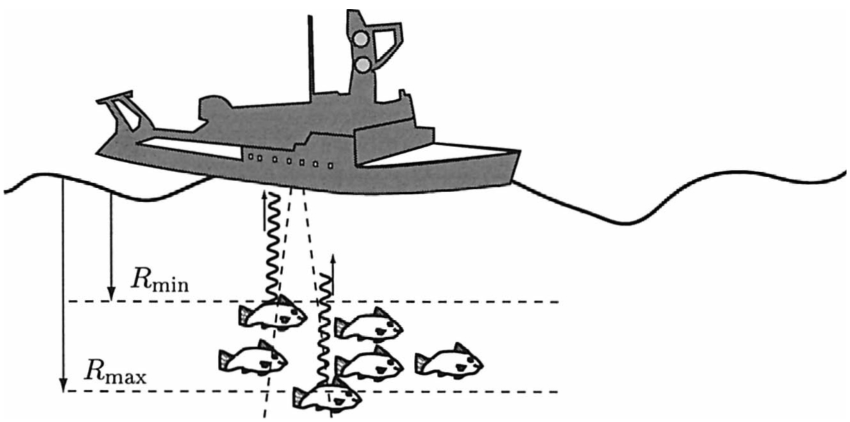


Figure 1: Sonar system

Simplifying for clarity, and referring to Figure 1, a sound pulse, as a sinusoidal signal, is transmitted from a ship. As this pulse encounters a school of fish, it will be reflected by each fish, making the received waveform the sum of all the individual pulses.

Let the time observation window be $T = \left[2\frac{R_{\min}}{c}, 2\frac{R_{\max}}{c} \right]$,

where R_{min} and R_{max} are the minimum and maximum ranges of interest respectively, and c is the speed of sound in water. Based on the received waveform, we wish to estimate the number of fish in the vertical direction in the desired range window.[1]

While strictly speaking, different angular regions require other pulses to be transmitted, here we assume that by the Central Limit Theorem, a large number of fish produce reflections, and hence the received waveform can be modeled as a Gaussian random process. As a result, the many reflected pulses will overlap in time, with two of the reflected pulses shown in Figure 1. Hence, each reflected pulse can be represented as

---

[1] Note that only the fish within the observation window denoted by the dashed horizontal lines,

$X_i(t) = A_i \cos(2\pi f_0(t-\tau_i)+\theta_i)$

where f_0 is the transmit frequency in $Hz(1khz)$ and $\tau_i = {}^{2R_i}/_{c}$ is the time delay of the pulse reflected from the i[th] fish. Since both A_i, and θ_j depend upon the fish's position, orientation, and motion, which are not known apriori, we assume that they are realizations of random variables (RV). Furthermore, since the ranges of the individual fish are unknown, we also do not know τ_i. The above model equation is rewritten as

$X_i(t) = A_i \cos(2\pi f_0 t + \Theta_i)$, where $\Theta_i = \theta_i - 2\pi f_0 \tau_i$

Hence the received waveform can be written as

$$X(t) = \sum_{i=1}^{N} X_i(t) = \sum_{i=1}^{N} A_i \cos(2\pi f_0 t + \Theta_i)$$

1. Rewrite the equation of the received waveform (over all fish), using $U_i = A_i \cos(\Theta_i)$ and $V_i = A_i \sin(\Theta_i)$.
   Assume that all fish are about the same size, and hence the echo amplitudes are about the same. Note that U and V are sums of independent RVs, i.e. different reflections of fish do not affect each other.

2. Invoking the central limit theorem, why can you postulate a Gaussian PDF for U and V? Assume now U and V are independent so that $U,V \sim \mathfrak{N}(0, N\sigma^2)$

   As a result of this formulation, the received waveform can be modeled as a Rayleigh fading sinusoid model, where the envelope of the received waveform X(t), given by $A = \sqrt{U^2+V^2}$, has a Rayleigh distribution, written as

$$p_A(a) = \begin{cases} \dfrac{a}{\sigma^2 N}\exp\left(-\dfrac{1}{2}\dfrac{a^2}{N\sigma^2}\right) & a \geq 0 \\ 0 & a < 0 \end{cases}$$

   Hence, if we have previously measured the reflection characteristics of a single fish,then we will know σ^2.

3. Use the mean of a Rayleigh RV, $E[A] = \sqrt{\dfrac{\pi}{2}N\sigma^2}$ to derive the theoretical number of fish.
4. To estimate the number of fish empirically, we proceed to transmit P pulses and use the measured m^{th} envelope $\widehat{A}_m$ of each received waveform $X_m(t)$ for $m = 1, \cdots P$. How

would you go about computing the empirical $\widehat{N}$? [2]

5. Simulate $X_i(t) = A_i \cos(2\pi f_0 t + \Theta_i)$ with A_i Rayleigh distributed with $\sigma^2 = 1$, and $\Theta \sim U[0,2\pi]$,

   Hint:Replace $X_i[n] = X_i(n\Delta t) = A_i cos(2\pi f_o n\Delta t)$

   for $n = 0,1,\ldots,N\Delta t$, where $\Delta t = 1/N$ and N is large.

6. Taking P=50, plot the histogram of X(t). Is that consistent with above assertion? If yes, why? If not, why?
7. Now simulate 50 realizations of each of $U_i, V_i \sim N(0,1)$. Can you verify that A_i is indeed distributed as stipulated above?

---

[2] Recall that $A_m = \sqrt{U_m^2 + V_m^2}$ from $X_m(t)$

1] Rewrite the equation of the received waveform (over all fish), using $Ui = Ai\,cos(\Theta i\,)$ and $Vi = Ai\,sin(\Theta i\,)$.
Assume that all fish are about the same size, and hence the echo amplitudes are about the same. Note that U and V are sums of independent RVs, i.e. different reflections of fish do not affect each other.

$$X(t) = \sum_{i=1}^{N} X_i(t)$$

$$X(t) = \sum_{i=1}^{N} A_i * \cos\,(2\pi f_0 t + \Theta_i)$$

Expanding using Trigonometric identities

$$X(t) = \sum_{i=1}^{N} A_i * [\cos\,(2\pi f_0 t)\cos\,(\Theta_i) - \sin\,(2\pi f_0 t)\sin(\Theta_i)]$$

$$X(t) = \cos(2\pi f_0 t) * \sum_{i=1}^{N} A_i * \cos(\Theta_i) - \sin(2\pi f_0 t) * \sum_{i=1}^{N} A_i * \sin(\Theta_i)$$

Given: $U_i = A_i\,cos(\Theta_i\,)\ and\ V_i = A_i\,sin(\Theta_i\,)$

$$\boldsymbol{X(t) = \cos(2\pi f_0 t) * \sum_{i=1}^{N} U_i - \sin(2\pi f_0 t) * \sum_{i=1}^{N} V_i}$$

2] Invoking the central limit theorem, why can you postulate a Gaussian PDF for U and V?
Central Limit Theorem (CLT) states that, given n independent random variables xi, we form their sum,
$X = x1 + x2 + x3 + \cdots + xn$
With mean $\eta = \eta 1 + \eta 2 + \eta 3 + \cdots + \eta n$
and variance $\sigma^2 = \sigma 1^2 + \sigma 2^2 + \cdots + \sigma n^2$
So under certain general conditions, the distribution can be considered as Gaussian as n increases.
$F(x) = G(\frac{x-\eta}{\sigma})$ with mean η and variance σ^2.[1]

[1] Papoulis A, Pillai S. U., Probability, Random Variables, And Stochastic Processes, McGraw Hill Education (India) Private Ltd, Fourth Edition, New Delhi, 2014

Here in our case, each pulse is considered to be an independent pulse and identical. It is also mentioned that large number of fishes produce reflections. As we satisfy all the conditions for CLT, we can postulate a Gaussian pdf for U and V.

Assume now U and V are independent so that $U, V \sim \mathbb{N}(0, N\sigma^2)$.
As a result of this formulation, the received waveform can be modeled as a Rayleigh fading sinusoid model, where the envelope of the received waveform X(t), given by $A = \sqrt{(U^2 + V^2)}$, has a Rayleigh distribution, written as

$$p_A(a) = \begin{cases} \frac{a}{\sigma^2 N} \exp\left(-\frac{a^2}{2N\sigma^2}\right) & a \geq 0 \\ 0 & a < 0 \end{cases}$$

Hence, if we have previously measured the reflection characteristics of a single fish, then we will know σ^2.

$$Given: \ A = \sqrt{(U^2 + V^2)}$$

$$U = \sqrt{(A^2 - V^2)}$$

$$F_a(a) = \int_{-a}^{a} \int_{-\sqrt{(a^2-v^2)}}^{\sqrt{(a^2-v^2)}} f_{u,v}(u, v) \, du \, dv$$

Differentiating and removing the inner integral,

$$f_a(a) = \int_{-a}^{+a} \frac{a}{\sqrt{a^2 - v^2}} \{ f_{uv}(\sqrt{a^2 - v^2}, v) + f_{uv}(-\sqrt{a^2 - v^2}, v) \} \, dv$$

$$f_a(a) = \int_{-a}^{+a} \frac{a}{\sqrt{a^2 - v^2}} \left\{ \frac{1}{2N\sigma^2\pi} \exp\left(\frac{-a^2}{2N\sigma^2}\right) + \frac{1}{2N\sigma^2\pi} \exp\left(\frac{-a^2}{2N\sigma^2}\right) \right\} dv$$

$$f_a(a) = \int_{-a}^{+a} \frac{2a}{2N\sigma^2\pi} \left\{ \exp\left(\frac{-a^2}{2N\sigma^2}\right) \right\} \frac{1}{\sqrt{a^2 - v^2}} \, dv$$

$$f_a(a) = \frac{a}{N\sigma^2\pi} \left\{ \exp\left(\frac{-a^2}{2N\sigma^2}\right) \right\} \int_{-a}^{+a} \frac{1}{\sqrt{a^2 - v^2}} \, dv$$

$$f_a(a) = \frac{2a}{N\sigma^2\pi}\left\{\exp\left(\frac{-a^2}{2N\sigma^2}\right)\right\}\int_0^a \frac{1}{\sqrt{a^2-v^2}}dv$$

$Let\ v = a.cos\theta$
$dv = \ -a.sin\theta.d\theta$

$$\sqrt{a^2-v^2} \ = \ \sqrt{a^2-a^2\cos^2\theta} \ = \ \sqrt{a^2\sin^2\theta} \ = a.sin\theta$$

$$f_a(a) = \frac{2a}{N\sigma^2\pi}\left\{\exp\left(\frac{-a^2}{2N\sigma^2}\right)\right\}\int_{-\frac{\pi}{2}}^{0} \frac{1}{a.sin\theta}(-a).sin\theta\, d\theta$$

$$f_a(a) = \frac{2a}{N\sigma^2\pi}\left\{\exp\left(\frac{-a^2}{2N\sigma^2}\right)\right\}\int_0^{\frac{\pi}{2}} d\theta$$

$$f_a(a) = \frac{2a}{N\sigma^2\pi}\left\{\exp\left(\frac{-a^2}{2N\sigma^2}\right)\right\}.(\frac{\pi}{2})$$

$$f_a(a) = \frac{a}{N\sigma^2}\left\{\exp\left(\frac{-a^2}{2N\sigma^2}\right)\right\}$$

Hence it is shown that,

$$\boldsymbol{p_A(a)} = \begin{cases} \boldsymbol{\frac{a}{\sigma^2 N}\exp\left(-\frac{a^2}{2N\sigma^2}\right)} & \boldsymbol{a \geq 0} \\ \boldsymbol{0} & \boldsymbol{a < 0} \end{cases}$$

3] Use the mean of a Rayleigh RV, $E\{A\} = \sqrt{\frac{\pi N\sigma^2}{2}}$ to derive the theoretical number of fish.

$$E\{A\} = \int_0^\infty a.\frac{a}{N\sigma^2}\exp\,(\frac{-a^2}{2N\sigma^2})\,da$$

$$E\{A\} = \int_0^\infty \frac{a^2}{N\sigma^2}\exp\,(\frac{-a^2}{2N\sigma^2})\,da$$

$$Let\ \frac{a^2}{2N\sigma^2} = k$$

$$\frac{1}{2N\sigma^2}.2a.da = dk$$

$$a = \sqrt{2kN\sigma^2}$$

$$da = dk.\frac{2N\sigma^2}{2a} = dk.\frac{2N\sigma^2}{2\sqrt{2N\sigma^2}\sqrt{k}} = dk.\frac{\sqrt{2N\sigma^2}}{2\sqrt{k}}$$

$$E\{A\} = \int_0^\infty \frac{a^2}{N\sigma^2}\exp\left(\frac{-a^2}{2N\sigma^2}\right) da$$

$$E\{A\} = \int_0^\infty 2k.\mathrm{e}^{-\mathrm{k}}\sqrt{2N\sigma^2}\,k^{-\frac{1}{2}}.\frac{1}{2}dk$$

$$E\{A\} = \sqrt{2N\sigma^2}\int_0^\infty \sqrt{k}\,.\,e^{-k}dk$$

$As\ we\ know$

$$\Gamma(t) = \int_0^\infty x^{t-1}.e^{-x}dx \qquad \ldots Gamma\ Integral$$

$$Here, t = \frac{3}{2}$$

$$E\{A\} = \sqrt{2N\sigma^2}.\Gamma\left(\frac{3}{2}\right)$$

We know that, $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

$$\Gamma\left(\frac{1}{2} + 1\right) = \frac{(2 * 1)!}{4^1.1!}\sqrt{\pi} = \frac{\sqrt{\pi}}{2}$$

$$\boldsymbol{E\{A\} = \frac{\sqrt{2N\sigma^2}\sqrt{\pi}}{2} = \sqrt{\frac{\pi N\sigma^2}{2}}}$$

$$\boldsymbol{N = \frac{2E^2\{A\}}{\pi\sigma^2}}$$

4] To estimate the number of fish empirically, we proceed to transmit P pulses and use the measured m^{th} envelope $\hat{A}_m$ of each received waveform $X_m(t)$ for $m = 1 \ldots P$. How would you go about computing the empirical $\widehat{N}$?

We transmit P pulses and measure the m^{th} envelope $\hat{A}_m$ of each received waveform $X_m(t)$ for $m = 1 \ldots P$.
As we already know,

$$\boldsymbol{E\{A\} = \sqrt{\frac{\pi N \sigma^2}{2}}}$$

There for m^{th} envelope, we can write $\hat{A}_m = \sqrt{\boldsymbol{\frac{\pi N_m \sigma^2}{2}}}$

So,

$$\hat{A}_1 = \sqrt{\frac{\pi N_1 \sigma^2}{2}} \qquad \hat{A}_2 = \sqrt{\frac{\pi N_2 \sigma^2}{2}} \qquad \ldots \qquad \hat{A}_p = \sqrt{\frac{\pi N_p \sigma^2}{2}}$$

Therefore

$$\widehat{N}_1 = \frac{2}{\pi\sigma^2}\left(\hat{A}_1^2\right) \quad \widehat{N}_2 = \frac{2}{\pi\sigma^2}\left(\hat{A}_2^2\right) \quad \ldots. \quad \widehat{N}_p = \frac{2}{\pi\sigma^2}\left(\hat{A}_p^2\right)$$

$$\sum_{m=1}^{P} \widehat{N}_m = \frac{2}{\pi\sigma^2} \cdot \sum_{m=1}^{P} \hat{A}_m^2$$

$$\boldsymbol{\widehat{N} = \frac{1}{P}\sum_{m=1}^{P} \widehat{N}_m = \frac{2}{P\pi\sigma^2}\sum_{m=1}^{P} \hat{A}_m^2 = \frac{2}{P\pi\sigma^2}\sum_{m=1}^{P} (U_m^2 + V_m^2)}$$

5] Simulate $X_i(t) = A_i \cos(2\pi f_0 t + \Theta_i)$ with Ai Rayleigh distributed with $\sigma^2 = 1$ and $\Theta \sim U[0,2\pi]$.

Considering large number of fishes, N = 10000. The function $Xigenerator$ is used to generate every Xi.

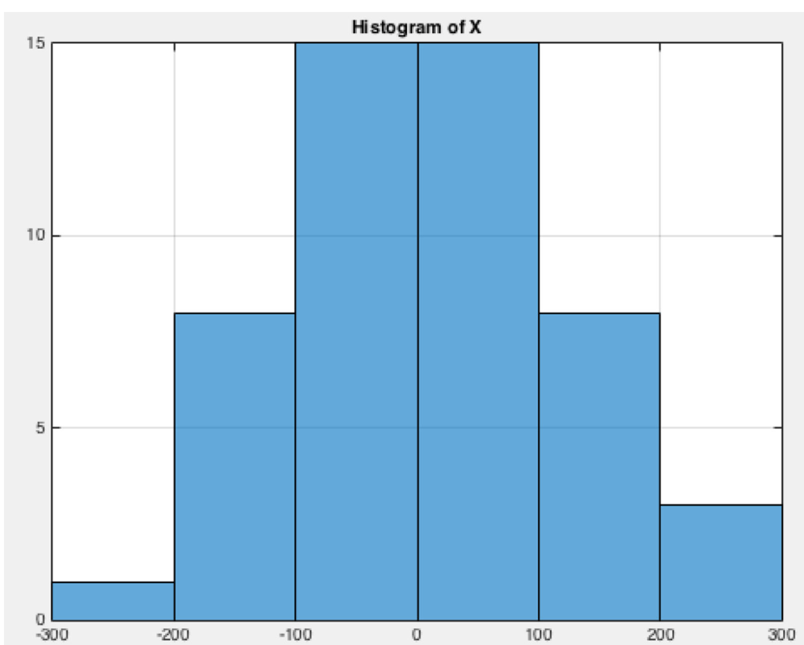
```
N = 10000; %Large Number of fishes
P = 500; %No of simulations

X = zeros(1,P);
for i = 1:P
    X(i) = Xigenerator(N);
end

   function [ X ] = Xigenerator(N)

   deltaT = 1/N;
   f0 = 1000; %Any frequency
   sigma = real(sqrt(1)); %Variance = 1

   %uniformly distributed theta U(0,2*pi)
   theta = random('uniform',0,2*pi,[N,1]);

   %Rayleigh Distributed amplitudes
   Ai = random('rayl',sigma,[N,1]);

   X = 0;
   for n = 0:1:N-1
       X = X + (Ai(n+1)*cos((2*pi*f0*n*deltaT) +
   theta(n+1)));
   end

   return;
   end
```

6] Taking P=50, plot the histogram of $X(t)$. Is that consistent with above assertion? If yes, why? If not, why?
So histogram of X is plotted with the inbuilt function, $histogram(\,)$. The histogram should be Gaussian or Normally Distributed.

To check this, we will use ***Anderson-Darling Test*** for Normality. If we get $p > 0.05$, the distribution is Gaussian.

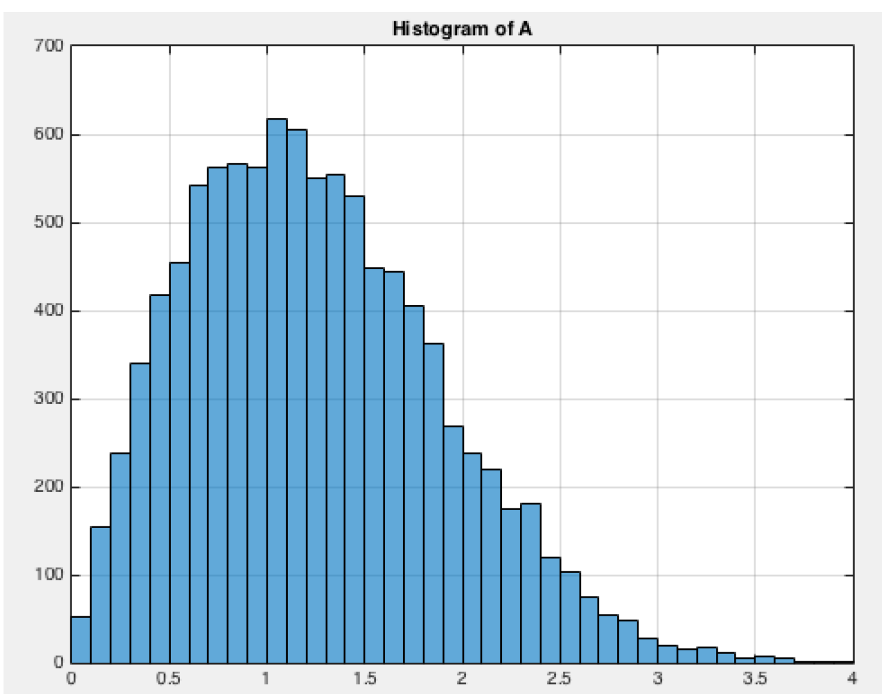
```
figure
histogram(X);
title('Hostogram of X');
grid on;
[h,p,adstat,cv] = adtest(X); % Anderson-Darling Test
if p>0.05
    disp(p);
    fprintf('\nWe failed to reject the null
hypothesis. So, the distribution of X is
normal.\n\n');
end
```

```
Output:
We failed to reject the null hypothesis. So, the distribution of X is normal. % p = 0.4076
```



7] Now simulate 50 realizations of each of $U_i, V_i \sim \mathbb{N}(0,1)$. Can you verify that A_i is indeed distributed as stipulated above?

```
% Generate U and V

U = random('normal',0,1,[N,1]);
V = random('normal',0,1,[N,1]);

% Calculation of A
A = zeros(1,N);
for i = 1:1:N
    A(i) = real(sqrt((U(i))^2 + (V(i))^2));
end
figure
histogram(A);
title('Histogram of A');
grid on;
test_cdf = makedist('Rayleigh','b',1);
[h,p] = kstest(A,'CDF',test_cdf);
if p>0.05
    fprintf('\n We failed to reject the null
hypothesis. A is Rayleigh Distributed.\n\n');
end
```



The original Ai was Rayleigh Distributed. So the Ai we are getting after the calculation should be Rayleigh distributed with $\sigma^2 = 1$.

So $test\_cdf$ is the cdf generated for Rayleigh Distribution with variance as 1. We will compare our histogram with this $test\_cdf$ using ***Kolmogorov-Smirnov Test***.

```
Output:
We failed to reject the null hypothesis. A is Rayleigh
Distributed. % p = 0.8660
```