ENDER ERKAYA
56288

March 5, 2012

ELEC 530
Homework Set #3

1. Bias and Variance ...

$$x[n] = A + w[n] \quad , n = 0, 1, \ldots, N-1$$

$$\underline{w} \sim N(\underline{0}, \sigma^2 I) \quad \text{with } \sigma^2 = 1$$

$$\hat{A}_1 = \frac{1}{N} \sum_{n=0}^{N-1} x[n] = \bar{x} = \frac{1}{N} 1^T \underline{x}$$

$$\hat{A}_2 = \frac{1}{N+2} \left(2x[0] + \sum_{n=1}^{N-2} x[n] + 2x[N-1] \right)$$

$$E\{\hat{A}_1\} = \frac{1}{N} E\{1^T \underline{x}\} = \frac{1}{N} N . E\{x[n]\} = \underline{A}$$

Since, $\underline{x} \sim N(\underline{A}, \sigma^2 I)$

$$\mathrm{Var}\{\hat{A}_1\} = \frac{1}{N^2} N . \mathrm{Var}\{x[n]\} = \underline{\frac{\sigma^2}{N}}$$

Hence, $\underline{\hat{A}_1 \sim N\left(A, \frac{\sigma^2}{N}\right)}$ (due to affine comb. of gaussian rvs)

$$E\{\hat{A}_2\} = \frac{1}{N+2} \left(2\, E\{x[0]\} + (N-2)\, \overset{\text{all same}}{E\{x[n]\}} + 2 . E\{x[N-1]\} \right)$$

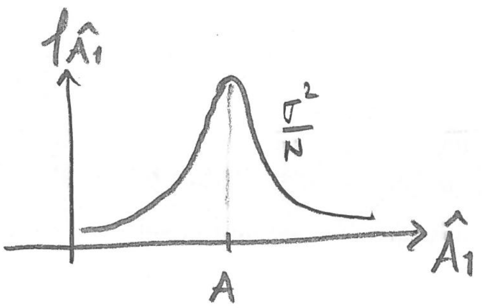
$$= \frac{1}{N+2} \left(2A + (N-2)A + 2A \right) = \frac{1}{N+2} (N+2) A = \underline{A}$$

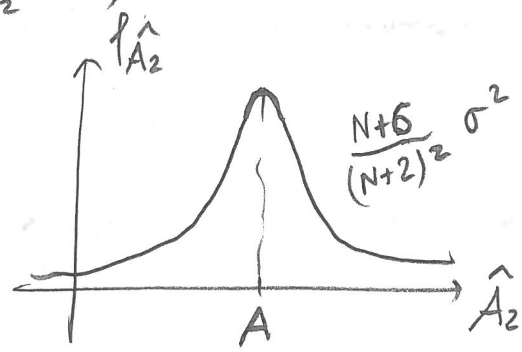
$$\mathrm{Var}\{\hat{A}_2\} = \frac{1}{(N+2)^2}\left(4\,\mathrm{Var}\{x[0]\} + (N-2)\,\mathrm{Var}\{x[n]\} + 4\,\mathrm{Var}\{x[N-1]\}\right)$$

$$= \frac{1}{(N+2)^2}(4+N-2+4)\,\sigma^2 = \frac{N+6}{(N+2)^2}\,\sigma^2$$

$$\hat{A}_2 = \frac{1}{N+2}\,\underline{a}^T\underline{x} \quad \text{with} \quad a = [2 \; 1 \; \ldots \; 1 \; 2]$$

Hence, $\hat{A}_2 \sim N\left(A, \frac{N+6}{(N+2)^2}\sigma^2\right)$





Both $\hat{A}_1, \hat{A}_2$ are unbiased estimators. To decide which one is better in MSE sense, we can compare their variances.

$$MSE(\hat{A}_1) = \mathrm{Var}(\hat{A}_1) + \cancel{\mathrm{Bias}(\hat{A}_1)^2} = \mathrm{Var}(\hat{A}_1)$$

$$MSE(\hat{A}_2) = \mathrm{Var}(\hat{A}_2) + \cancel{\mathrm{Bias}(\hat{A}_1)^2} = \mathrm{Var}(\hat{A}_2)$$

$$\frac{\sigma^2}{N} \overset{?}{>} \frac{(N+6)}{(N+2)^2}\sigma^2$$

$$\sigma^2 \overset{?}{>} \frac{N(N+6)}{(N+2)^2}\sigma^2 \iff N(N+6) \overset{?}{<} (N+2)^2$$

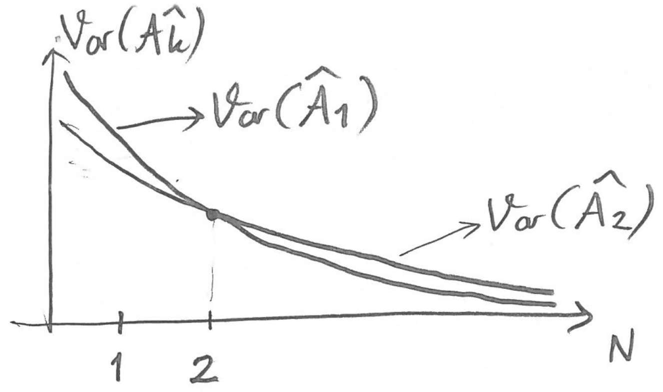
$$\cancel{N^2} + 6N \overset{?}{<} \cancel{N^2} + 4N + 4$$

$$2N \overset{?}{<} 4$$

$$N \overset{?}{<} 2$$

For $N < 2$

$$\mathrm{Var}\{\hat{A}_1\} > \mathrm{Var}\{\hat{A}_2\}$$



For $N<2$ or $N=1$, $\hat{A}_2$ is better than $\hat{A}_1$.

For $N=2$, $\hat{A}_2$ or $\hat{A}_1$ are same in MSE sense.

For $N>2$, $\hat{A}_1$ is better than $\hat{A}_2$.

It does not depend the value of A. It depends the value of N. $\hat{A}_1$ is asymptotically better than $\hat{A}_2$.

---

2. Estimating Variance

$\underline{x} \sim N(\underline{0}, \sigma^2 I)$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n]$$

$$E\{\hat{\sigma}^2\} = E\left\{\frac{1}{N}\sum_{n=0}^{N-1} x^2[n]\right\} = \frac{1}{N}\sum_{n=0}^{N-1} E\{x^2[n]\}$$

since iid

$$= \frac{1}{\cancel{N}} \cdot \cancel{N} \cdot E\{x^2[n]\}$$

$$= \left(Var\{x[n]\} + \underbrace{E\{x[n]\}^2}_{0}\right)$$

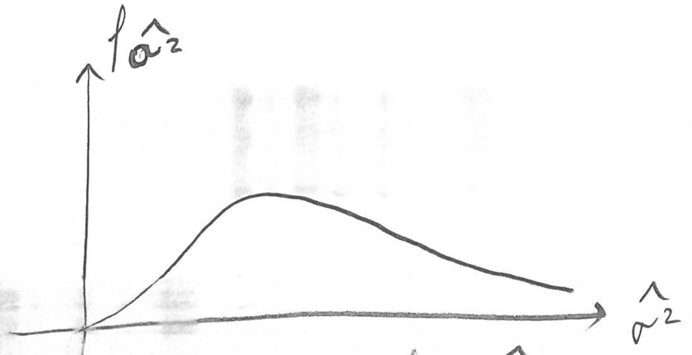
$$E\{\hat{\sigma}^2\} = \sigma^2$$ ✓ unbiased estimator

$$\mathrm{Var}\{\hat{\sigma}^2\} = \frac{1}{N^2}\,\mathrm{Var}\left\{\sum_{n=0}^{N-1} x^2[n]\right\} \underset{\text{iid } x^2[n]}{=} \frac{1}{N^2}\, N \cdot \mathrm{Var}\{x^2[n]\}$$

$$= \frac{1}{N^2} \cdot \left(E\{x^4[n]\} - E\{x^2[n]\}^2\right)$$

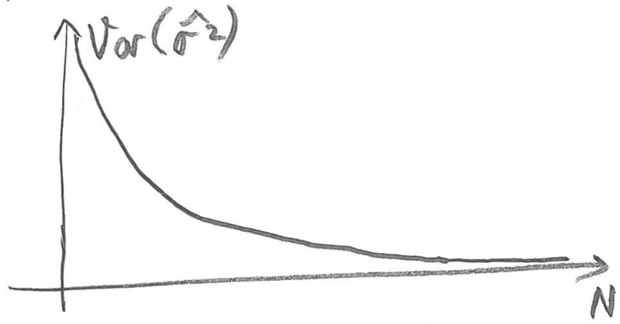
$$= \frac{1}{N^2}\left(3\sigma^4 - \sigma^4\right) = \frac{2\sigma^4}{N^2}$$

$\hat{\sigma}_2 \sim \frac{\sigma^2}{N}\chi^2_N \rightarrow$ chi-squared distribution with N degrees of freedom



example density function for $\hat{\sigma}^2$

As $N \to \infty$, $\mathrm{Var}\{\hat{\sigma}^2\} \to 0$



$*$ $\underset{N \to \infty}{\mathrm{l.i.m}}\ \hat{\sigma}^2 \rightarrow \sigma^2$

---

3. Bias-Variance Trade-off

$\underline{x} \sim N(A\underline{1}, \sigma^2 \underline{\underline{I}})$

We know $|A| \leq A_0$ apriori, ie $p_A(A) = u[A + A_0] - u[A - A_0]$

$\hat{A} = E\{A|\underline{x}\}$, conditional expectation optimum in MSE sense

$$p(A|\underline{x}) = \frac{p(\underline{x}|A)\,p(A)}{p(\underline{x})} = \frac{\left(\prod_{i=1}^{N} p(x_i|A)\right) \cdot p(A)}{p(\underline{x})}$$

due to conditional independence

$$p(\underline{x}, A) = \left(\prod_{i=1}^{N} p(x_i / A)\right) \cdot p_A(A)$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - A)^2\right\} \cdot p_A(A)$$

$$= (2\pi\sigma^2)^{N/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - A)^2\right\} \cdot \left(u[A+A_0] - u[A-A_0]\right)$$

$$p(\underline{x}) = \int p(\underline{x}, A)\, dA = \int_{-A_0}^{A_0} (2\pi\sigma^2)^{N/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - A)^2\right\} dA$$

$$E\{A \mid \underline{x}\} = \frac{\cancel{(2\pi\cdot\sigma^2)^{N/2}} \int_{-A_0}^{A_0} A \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - A)^2\right\} dA}{\cancel{(2\pi\cdot\sigma^2)^{N/2}} \int_{-A_0}^{A_0} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - A)^2\right\} dA}$$

I cannot handle the algebraic solution, But intuitively, I recommend the estimator.

$$\hat{A} = \sum_{i=0}^{N-1} I_{A_0}(x_i) \cdot x_i, \quad I_{A_0}(x) = u[x+A_0] - u[x-A_0]$$

I_{A_0} is indicator of A_0 threshold due to $p_A(A)$ prior density of A, $(I_{A_0}(x) = p_A(x))$

To reduce MSE, we use prior information about A, to get better estimator. This one has reduced variance hence MSE.

4. Unbiased Estimator

$\{x[0], x[1], \ldots x[N-1]\}$ each distributed $\sim U[0, \Theta]$

Intuitively,

$\hat{\Theta} = \max\{x[n]\}_{n=0}^{N-1}$ would be an efficient estimator

and $T(x) = \max\{x[n]\}_{n=0}^{N-1}$ is a sufficient statistic or minimal sufficient statistic in this case.

We can check if $\hat{\Theta} = T(\underline{x})$ is unbiased.

$$E\{\hat{\Theta}\} = \int p_{\hat{\Theta}}(\hat{\theta})\, \hat{\theta}\, d\theta$$

To find $p_{\hat{\theta}}(\hat{\theta})$, we can use cumulative distribution of $\hat{\theta}$

$$p_{\hat{\Theta}}(\hat{\Theta} \leq \theta) = p(T(x) \leq \theta) = \prod_{n=0}^{N-1} p(x[n] \leq \theta) = \left(\theta \cdot \frac{1}{\Theta}\right)^N$$

for $0 < \theta < \Theta$

$$p_{\hat{\theta}}(\hat{\theta} = \theta) = \frac{\partial}{\partial \theta}\left(\theta \cdot \frac{1}{\Theta}\right)^N = \Theta^{-N} \cdot N \cdot \theta^{N-1},$$

$$= \Theta^{-N} \cdot N \cdot \theta^{N-1} \cdot (u[\theta] - u[\theta - \Theta])$$

$$E_{\hat{\theta}}\{\theta\} = \int_0^{\Theta} \theta \cdot p_{\hat{\theta}}(\theta)\, d\theta$$

$$= \int_0^{\Theta} \theta \cdot \Theta^{-N} \cdot N \cdot \theta^{N-1} \cdot d\theta = \Theta^{-N} \cdot \int_0^{\Theta} N \cdot \theta^N \cdot d\theta$$

$$= \Theta^{-N} \cdot N \cdot \left.\frac{\theta^{N+1}}{N+1}\right|_0^{\Theta}$$

$$= \Theta^{-N}\left(\frac{N}{N+1}\right)\Theta^{N+1} = \frac{N}{N+1}\Theta \rightarrow \text{biased}$$

To make $T(x)$ unbiased we need to scale it.

Hence,

$$\hat{\Theta}_u = \frac{N+1}{N} T(x) \quad \text{is unbiased.}$$

Now we have

$$E\{\hat{\Theta}_u\} = \frac{N+1}{N} E\{T(x)\} = \frac{N+1}{\cancel{N}} \frac{\cancel{N}}{N+1} \Theta = \Theta \checkmark$$

---

5. Simple CRLB Exercise

$$x[0] = A + w[0] \longrightarrow \sim p_w(w)$$

$$x[0] \sim p_x(x) = p_w(x-A)$$

(a)

$$CRLB = \frac{1}{\underbrace{-E\left\{\frac{\partial^2 \ln p(\underline{x};\theta)}{\partial \theta^2}\right\}}} \rightarrow I(\theta) : \text{Fisher information}$$

$$CRLB = \frac{1}{-E_x\left\{\frac{\partial^2 \ln p_w(x-A)}{\partial A^2}\right\}}$$

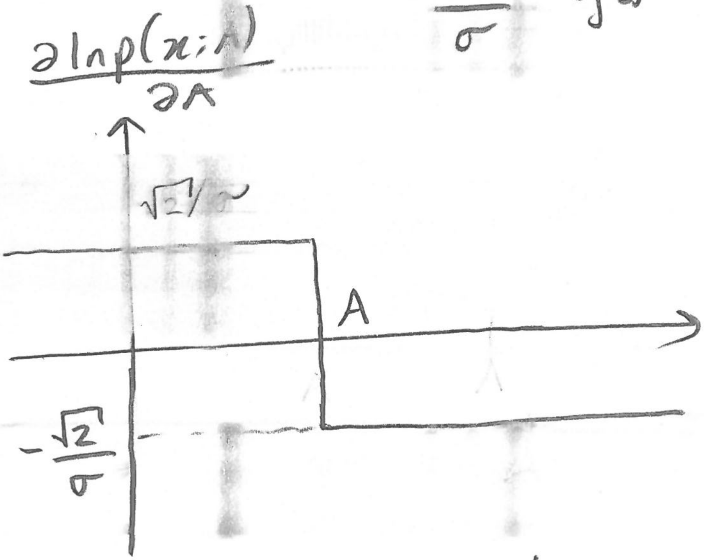
$$CRLB = \frac{1}{-\int \frac{\partial^2 \ln p_w(x-A)}{\partial A^2} p_w(x-A)\, dx}$$

(b) $\quad p_w(u) = \dfrac{1}{\sigma\sqrt{2}} \exp\left\{-\dfrac{\sqrt{2}}{\sigma}\cdot|u|\right\}$

$$I(\theta) = -E\left\{\frac{\partial^2 \ln p(x;\theta)}{\partial \theta^2}\right\} \quad : \text{Fisher information}$$

$$\ln p_x(x;A) = \ln p_w(x-A) = -\ln \sigma\sqrt{2} - \frac{\sqrt{2}}{\sigma}|x-A|$$

$$\frac{\partial \ln p(x;A)}{\partial A} = \begin{cases} \dfrac{\sqrt{2}}{\sigma} & \text{for } x>A \\[2ex] -\dfrac{\sqrt{2}}{\sigma} & \text{for } x<A \end{cases}$$



Since $\dfrac{\partial \ln p(x;A)}{A}$ is nondifferentiable (ie ∞ at 0) $\dfrac{\partial^2 \ln p(x;A)}{\partial A^2}$ does not apply

We can use the identity

$$* \quad E\left\{\left(\frac{\partial \ln p(x;\theta)}{\partial \theta}\right)^2\right\} = -E\left\{\frac{\partial^2 \ln p(x;\theta)}{\partial \theta^2}\right\}$$

Hence $E\left\{\dfrac{2}{\sigma^2}\right\} = \dfrac{2}{\sigma^2}$ $\quad \left(p(x=A)=0 \ \& \ \dfrac{\partial \ln p(x;A)}{A} \text{ finite at } A\right)$

$$\text{CRLB} = \frac{1}{I(\theta)} = \boxed{\frac{\sigma^2}{2}}$$

6. Score Vector for Poisson Distribution

$\{x[0], \ldots, x[N-1]\}$ iid with Poisson dist.

$$p_{x[i]}(k) = \exp\{-g(\underline{\Theta})\} \cdot \frac{g(\underline{\Theta})^k}{k!}$$

(a) $s_{\Theta i} = \nabla_\Theta \ln(p_X(x;\underline{\theta}))$

$$= \frac{1}{p_X(x;\underline{\theta})} \cdot \nabla_\Theta p_X(x;\underline{\theta})$$

$$= \frac{1}{p_X(x;\Theta)} \cdot \nabla_\Theta \left(\exp\{-g(\underline{\Theta})\} \cdot \frac{g(\underline{\Theta})^k}{k!}\right)$$

$$= \frac{1}{p_X(x;\theta)} \left(\exp\{-g(\Theta)\} \cdot \frac{g(\Theta)^k}{k!} \cdot - \underline{\nabla_\Theta g(\Theta)} + \exp\{-g(\Theta)\} \cdot k \cdot \frac{g(\Theta)^{k-1}}{k!} \cdot \underline{\nabla_\Theta g(\Theta)}\right)$$

$$s_{\Theta i} = \frac{1}{p_X(x;\theta)} \cdot \exp\{-g(\Theta)\} \cdot \left(\frac{g(\Theta)^{k-1}}{(k-1)!} - \frac{g(\Theta)^k}{k!}\right) \cdot \underline{\nabla_\Theta g(\Theta)}$$

$$s_{\Theta i} = \frac{\cancel{k!}}{\cancel{\exp\{-g(\Theta)\}} \cdot g(\Theta)^k} \cdot \cancel{\exp\{-g(\Theta)\}} \cdot \left(\frac{g(\Theta)^{k-1}}{\cancel{(k-1)!}} - \frac{g(\Theta)^k}{\cancel{k!}}\right) \cdot \nabla_\Theta g(\Theta)$$

$$s_{\Theta i} = \left(\frac{k}{g(\Theta)} - 1\right) \cdot \nabla_\Theta g(\Theta) \quad \text{for an } p_{X[i]}(k)$$

$$s_\Theta = \nabla_\Theta \ln(p_x(\underline{x};\underline{\theta})) = \nabla_\Theta \ln\left(\prod_{i=0}^{N-1} p(x[i];\theta)\right) - \underset{\sim}{\underline{iid}}$$

$$= \nabla_\Theta \sum_{i=0}^{N-1} \ln(p(x[i];\theta))$$

$$= \sum_{i=0}^{N-1} \nabla_\Theta \ln(p(x[i];\theta))$$

$$= \sum_{i=0}^{N-1} s_{\Theta i}$$

$$= \left(\frac{1^T x}{g(\Theta)} - N\right) \cdot \nabla_\Theta g(\Theta)$$

(b) $$E\{s_\Theta\} = \left(\frac{N \cdot E\{x[i]\}}{g(\Theta)} - N\right) \cdot \nabla_\Theta g(\Theta)$$

$$= \left(\frac{N \cdot \cancel{g(\Theta)}}{\cancel{g(\Theta)}} - N\right) \cdot \nabla_\Theta g(\Theta) = 0$$

(c) $g(\Theta) = \Theta^2, \ \Theta \in \mathbb{R}$

$$\nabla g(\Theta) = 2\Theta$$

$$s_\Theta = \left(\frac{1^T x}{\Theta^2} - N\right) \cdot 2\Theta$$

$$Var_x\{s_\Theta\} = \frac{2^2}{\Theta^2} \cdot N \cdot Var\{x_i\} = \frac{4}{\cancel{\Theta^2}} \cdot N \cdot \cancel{\Theta^2} = 4N$$

(d) $$s_\Theta = \frac{2}{\Theta} 1^T \underline{x} - N \cdot 2\Theta$$

$$\underset{\substack{\downarrow \\ g(\Theta)}}{x_i \sim Pois(\lambda_i)} \implies 1^T x \sim \underset{\substack{\downarrow \\ N \cdot g(\Theta)}}{Pois(1^T \lambda)}$$

$$p_{1^T \underline{x}}(1^T \underline{x}) = \exp\{-N \underbrace{g(\Theta)}_{\Theta^2}\} \cdot \frac{(N g(\Theta))^{1^T x}}{(1^T x)!}$$

$$p_{s_\Theta}(k) = \exp\{-N \cdot \Theta^2\} \cdot \frac{(N\Theta^2)^{\left(\frac{(k+N2\Theta)\cdot\Theta}{2}\right)}}{\left(\frac{(k+2N\Theta)\cdot\Theta}{2}\right)!}$$

$$p_{s_\Theta}(k) = \exp\{-N\Theta^2\} \cdot \cdot \frac{(N\Theta^2)^{\left(\frac{k\Theta+2N\Theta^2}{2}\right)}}{\left(\frac{k\Theta+2N\Theta^2}{2}\right)!}$$

(f) $I(\theta) = E\left\{\left(\frac{\partial}{\partial\theta} \ln p(x;\theta)\right)^2\right\}$

$$= E\{s_\theta^2\} = Var\{s_\theta\} = 4 \cdot N = \underline{40}$$

$$CRLB(\theta) = \frac{1}{I(\theta)} = \frac{1}{40} = 0.025$$

The variance of s_θ, $Var\{s_\theta\}$, is verified through simulations in Matlab part (e), attached.

(e) Matlab Part:

Score functions are calculated from the generated data. Calculated mean and variance are also given below as Figure 1. Analytical mean and variance in part(c,d) is verified by experimental results.

I(theta) = E{(s_theta)^2}= Var(s_theta) = 40

CRLB = 1/I(theta) = 1/40 = 0.025

```
>> Score_Vector_for_Poisson_Distribution
Analytical mean is calculated as 0 estimated mean through realizations is: 0.006
Analytical variance is calculated as 40 , estimated variance from realizations is: 39.992
```

Figure 1: Comparison of analytical vs experimental mean and variance

Figure 2 is the histogram estimator for density of s_theta, calculated from 1000 realizations.

Figure 3 is the comparison of analytical pdf of s_theta, calculated in part (d), and estimated pdf of s_theta, estimated by histogram estimator from the generated data.

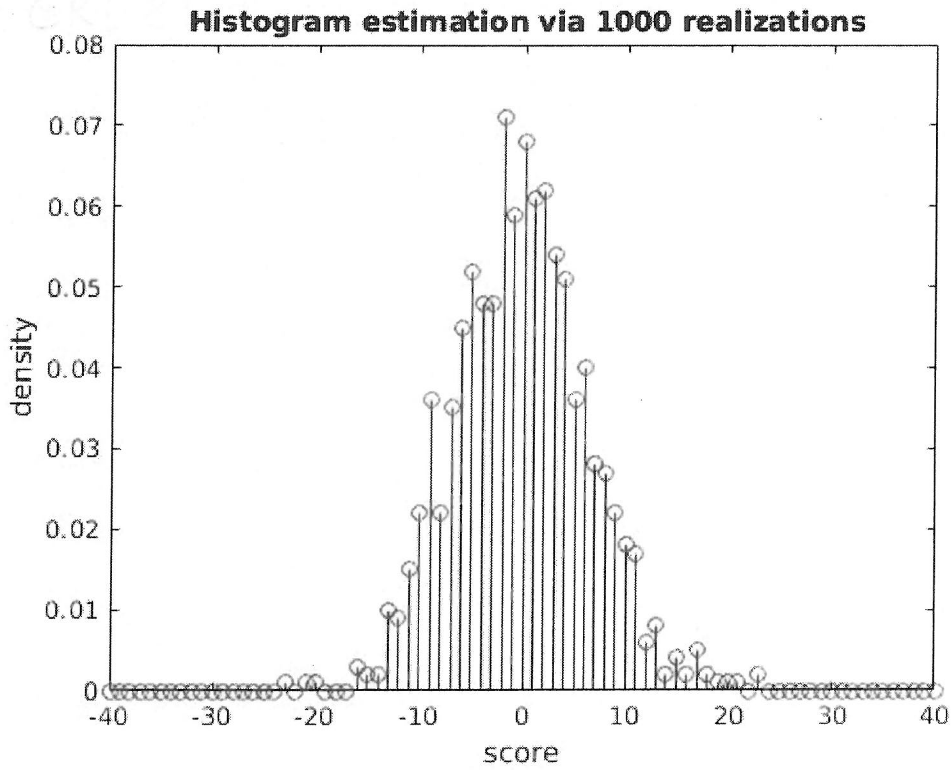


Figure 2: Histogram Estimator

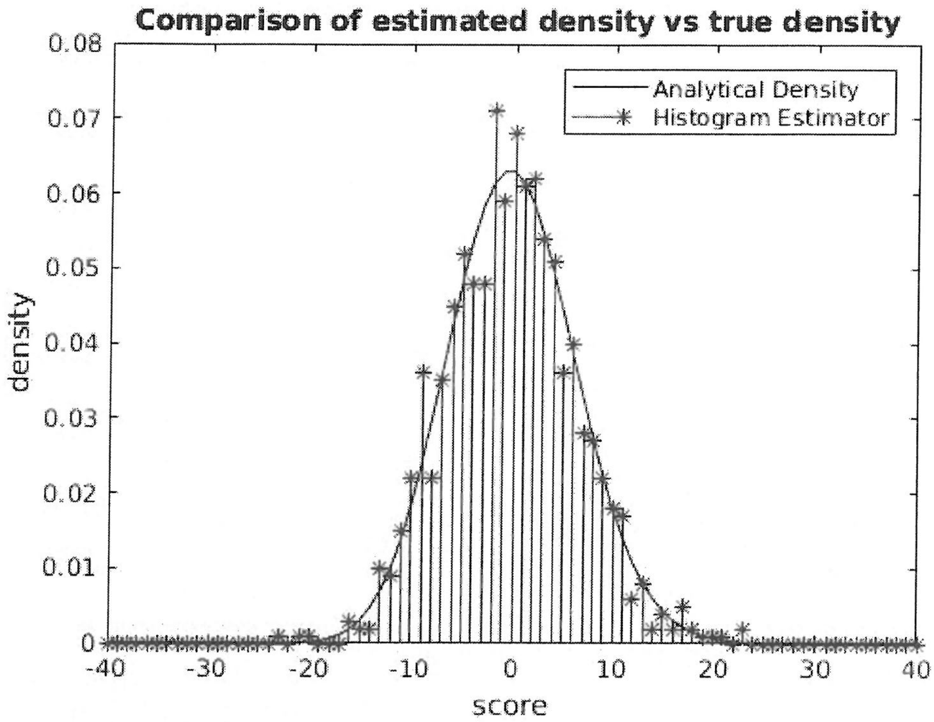


Figure 3 : Comparison of Analytical pdf and Histogram estimator

Matlab Code:

```
% ELEC-530 Detection and Estimation Theory
% Homework 3
% Question 6 - Part (e)
% written by Ender Erkaya
% 3/16/2018
%
% Score Vector for Poisson Distribution

% Global constants and defaults
theta = 2;
N = 10;
gtheta=theta^2;
Num_of_Realizations = 1000;

% Generate poisson random variables with parameter \lambda=gtheta
xdata = poissrnd(gtheta,N,Num_of_Realizations);

% Score vector generation
% for g_theta=theta^2
score = (2/theta) * sum(xdata) - 2 * N * theta;

% Density Estimation for score
bins = -40:40 ;
s_hist = hist(score,bins); % histogram
s_hist = s_hist/sum(s_hist); % for normalization
```

```
mean_score = mean(score); % mean estimation for score
% analytically calculated as 0.
display(['Analytical mean is calculated as ', num2str(0) , ' estimated mean through realizations is: ',
num2str(mean_score)]);

var_score = var(score); % variance estimation for score
% analytically calculated as 4 * N = 40
display(['Analytical variance is calculated as ', num2str(4*N),' , estimated variance from realizations
is: ', num2str(var_score)]);

% Comparison of Densities
s_theta = bins ;
% analytical density
fs_theta = exp(-N * (theta^2)) * ((N * theta^2).^((s_theta * theta + 2 * N * theta^2)/2)) ./
factorial((s_theta * theta + 2 * N * theta^2)/2);
% to check if it is a proper probability density
% max(cumsum(fs_theta)) - 1 < 1e-2 %true if it sums up 1

% Plotting and comparison
figure
stem(s_theta,s_hist,'r');
title('Histogram estimation via 1000 realizations');
xlabel('score');
ylabel('density');

figure
plot(s_theta,fs_theta,'b');
hold on
stem(s_theta,s_hist,'*r');
xlabel('score');
ylabel('density');
title('Comparison of estimated density vs true density');
legend('Analytical Density','Histogram Estimator');
```