**Advanced Signal Processing**

**Coursework 1**

Hao Ding

CID 00734091

Login hd1812

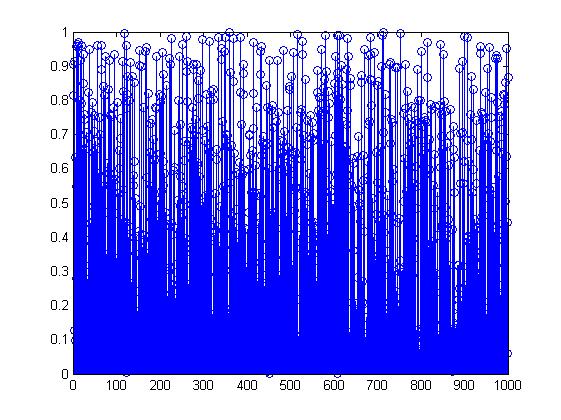
**CONTENTS**

**Coursework1: Random signals and stochastic processes**

* 1. **Statistical estimation**
  2. **Stochastic processes**
  3. **Estimation of probability distributions**

**1. Random signals and stochastic processes**

**1.1 Statistical estimation**

Using MATLAB command ‘rand’, 1000 samples of a uniform random variable X ∼ U (0, 1) are generated and plotted in Figure 1.

**Figure 1. 1000 samples of uniform distribution**

Although all samples have stochastic nature, they show uniformity since they are time-invariant and statistically stationary. Each sample is also referred as a realisation of the random process: X ∼ U (0, 1), ∀n.

**1.1.1** The theoretical mean, m = E{X}, of a uniform distribution between (0, 1) is 0.5. Using MATLAB function ‘mean’ to calculate sample mean, which performs as:

MATLAB code:

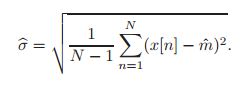


Result:

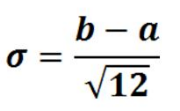
A=0.488

Sample mean calculated is has a deviation 2.4% from theoretical value and it is a good approximation of theoretical mean. However, different set of samples from the same distribution can have different sample average. Increasing sample number will lead to a more accurate estimator of theoretical average.

**1.1.2** Similar to 1.1.1, theoretical and sample standard deviation values are compared. MATLAB function ‘std’ computes sample standard deviation using the formula:



Whilst expected standard deviation can be derived from the expression:

For a uniform distribution, theoretical standard deviation is:

(b: higher boundary, a: lower boundary)

This can be calculated by taking integrals in pdf and in this example, result is approximately 0.2887.

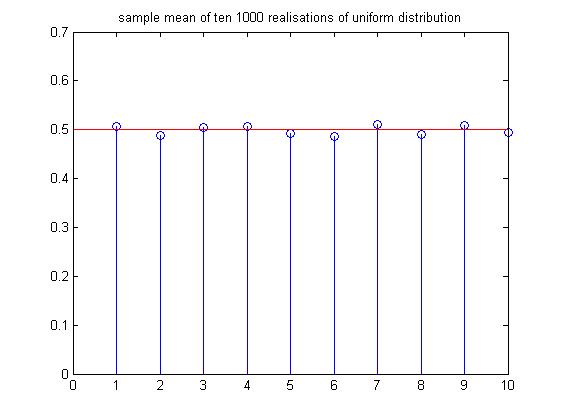
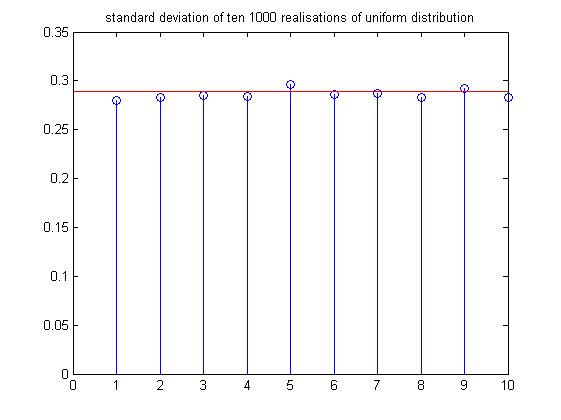
MATLAB code:



Result:

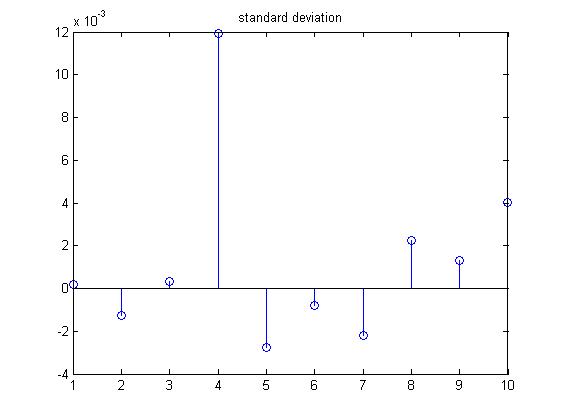
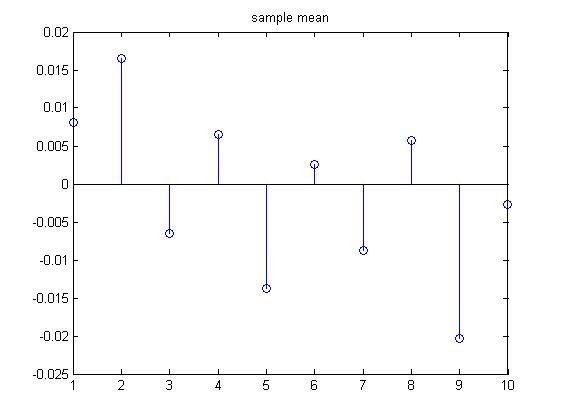
Std\_devs=0.2853

Sample value is 1.12% from the theoretical value. Therefore, deviation taken from 1000 realisation is a good estimator of theoretical deviation. Similar to sample mean, larger sample number results in a better estimation.

**1.1.3** Generate ten 1000-sample realisations and compare plot their bias of mean and standard deviation.

**Figure 2. Ten 1000-samples realisations. Sample mean and standard deviation**

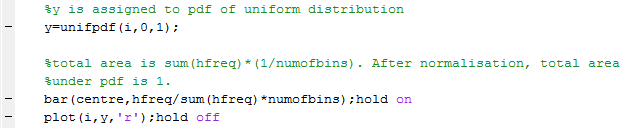


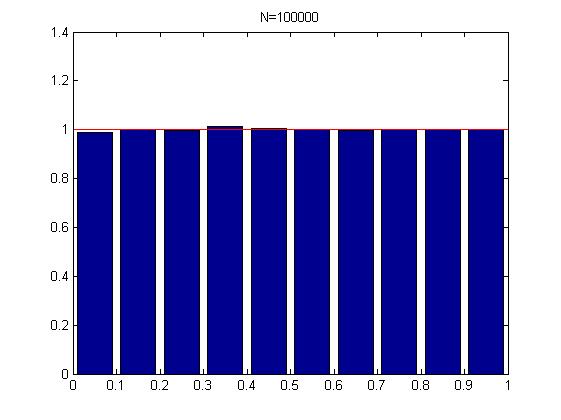
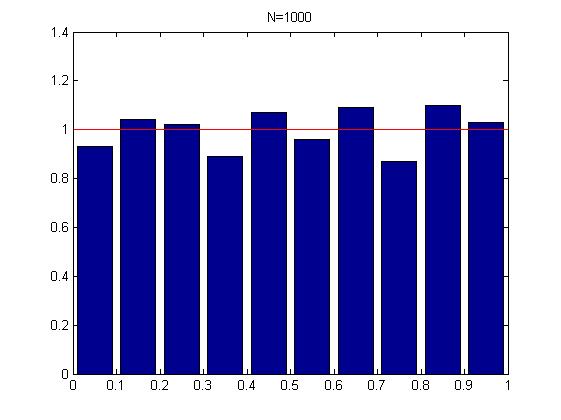
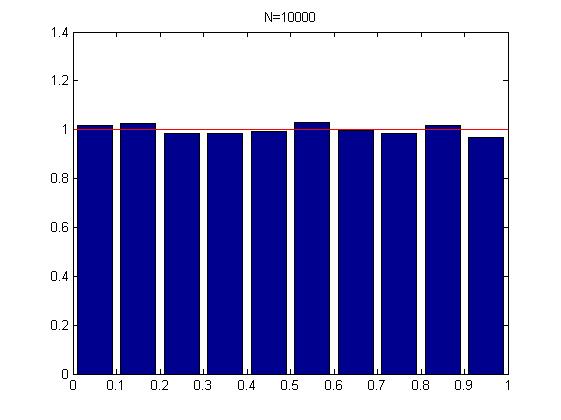
Bias is defined as Figure 2 illustrates that sample values are clustered closely around theoretical value. Therefore, bias is expected to be small, as shown below.

**1.1.4** Use MATLAB to generate pdf of uniform distribution.

**Figure 3. Bias of ten 1000-samples realisations. Sample mean and standard deviation**

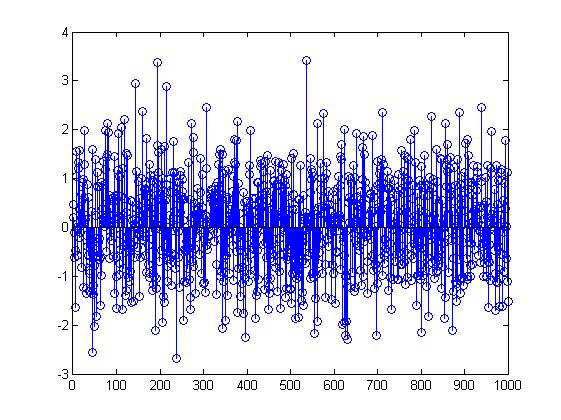
MATLAB code:



Figure 4 shows that the estimate converge as sample number increases. Regarding number of bins, if histogram is not normalised by number of bins, average height is inversely correlated to bin number. But larger bin number provides more details about the sample number in narrow intervals

**Figure 4. Pdf of uniform distribution for different N. (red line: theoretical value)**

**1.1.5** Repeat the process from 1.1.1 to 1.1.4 using ‘randn’ function and investigate the properties of random normal distribution.



**Figure 5. 1000 samples of Gaussian distribution**

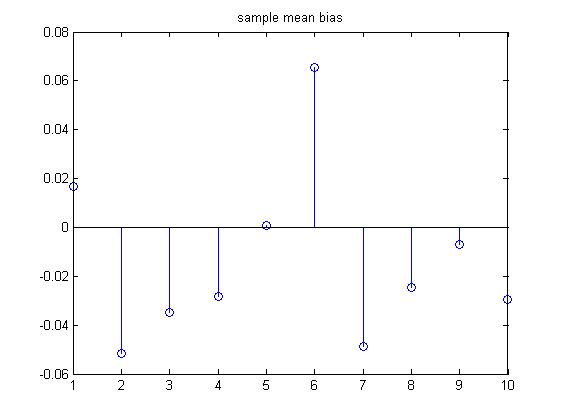
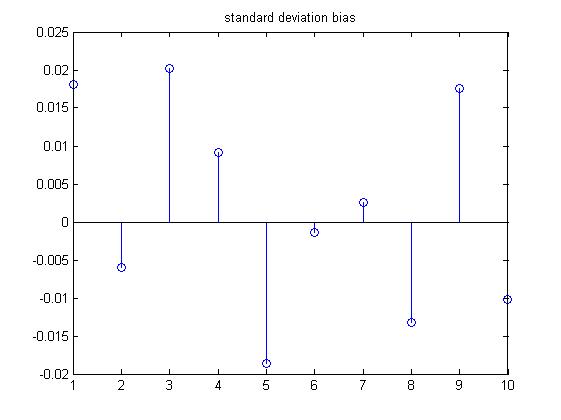
Theoretical mean: 0

Sample mean: 0.0155

Theoretical deviation: 1

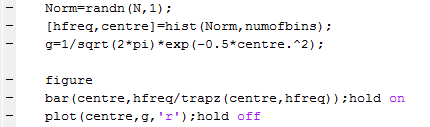
Sample deviation: 1.0002

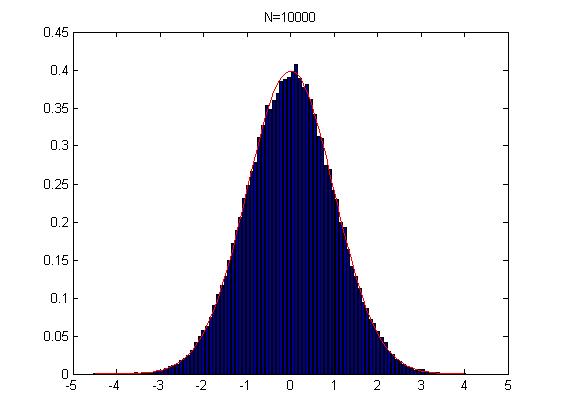
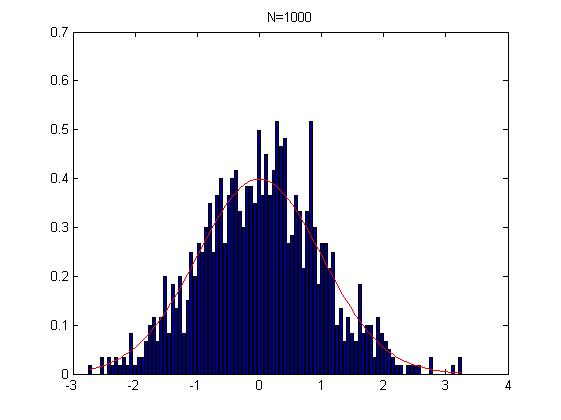
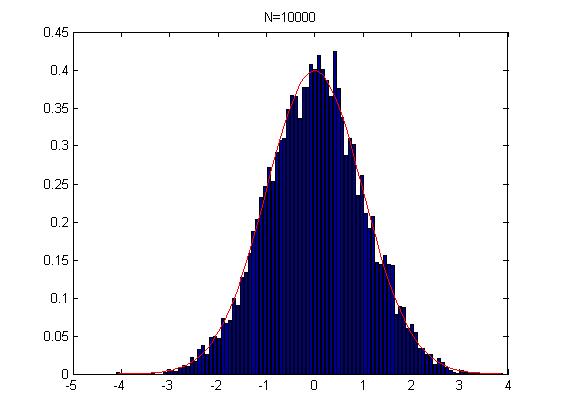
In this 1000-sample realisation, mean and deviation are good estimator and they are within 1% deviation of theoretical value. Larger sample number intended to reduce the estimation error.

Therefore, bias of mean and sample estimator is expected to be small. Figure 6 shows the bias of ten 1000-sample realisations and proves the accuracy of estimators.

**Figure 6. Bias of ten 1000-samples realisations. Sample mean and standard deviation**

**PDF of Gaussian distribution**

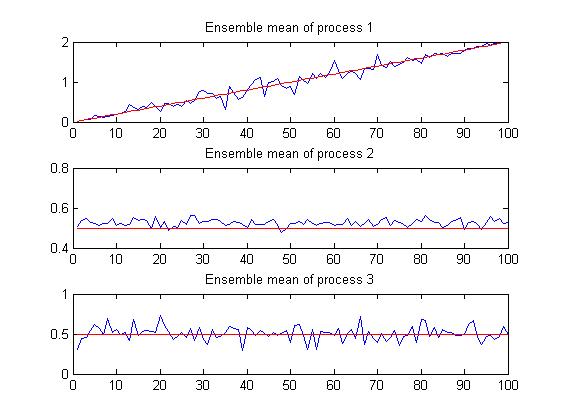


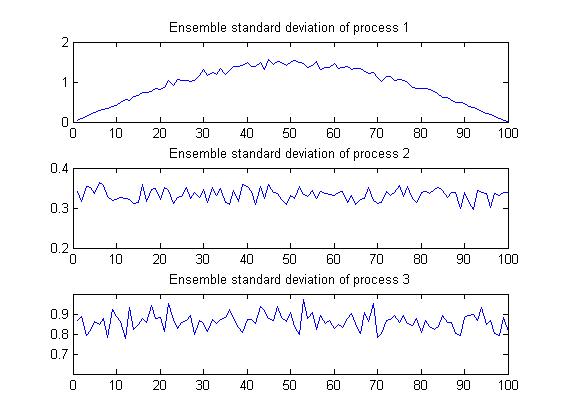


**Figure 7. PDF of Gaussian distribution when N=1000, 10000,100000. (Bin Number is 100)**

According to Figure 7, pdf converges as number of sample increases. Larger bin numbers provides a detailed view of small interval as well as a general trend.

**1.2 Stochastic process**

**1.2.1 Stationarity**



**Figure 8. Ensemble mean for three processes (red line: theoretical value, proved later)**

**Figure 9. Ensemble standard deviation for three processes**

Stationary processes can be categorised into Wide Sense Stationary (WSS) and Strict Sense Stationary (SSS). For WSS process, which has less restrictions, time average is constant and covariance only depends on time difference.

Figure 8 shows that process 1 is not stationary and process 2, 3 are stationary, since ensemble mean of first process varies with time.

**1.2.2 Ergodicity**

A stochastic process is ergodic if its properties, including mean and standard deviation, can be deduced from a sufficiently long sample.

Time average:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Process Number | Sample 1 | Sample 2 | Sample 3 | Sample 3 |
| 1 | 9.963 | 10.016 | 10.035 | 9.987 |
| 2 | 0.435 | 0.167 | 0.999 | 0.717 |
| 3 | 0.518 | 0.472 | 0.444 | 0.440 |

Standard deviation:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Process Number | Sample 1 | Sample 2 | Sample 3 | Sample 3 |
| 1 | 5.885 | 5.845 | 5.843 | 5.879 |
| 2 | 0.015 | 0.158 | 0.179 | 0.057 |
| 3 | 0.875 | 0.869 | 0.869 | 0.860 |

Process 1 is not ergodic because it is not a stationary process.

Process 2 is not ergodic since time average is not determinable.

Process 3 is ergodic.

**1.2.3 Mathematical Description of Sample Stochastic Processes**

**Process 1**

**Mathematical Expression:**

rand[n] ~ U ( -0.5 , 0.5 )

**Ensemble mean:**

= 0.02 \* n

**Ensemble variance:** Let a[n] =,

b[n] = 0.02n

Var(a[n]+b[n]) = Var(a[n])+Var(b[n]), a[n] and b[n] are independent at given n

Var(a[n]) = E(a2[n]) - E2(a[n])

= E (25\*rand2 [n]\*sin2 (nπ/N)) – 0

**=**  sin2 (nπ/N)

Var(b[n]) = 0

**Process 2**

**Mathematical Expression: (**a[n]-0.5)\*b[n] + c[n]

a[n] ~ U ( 0 , 1), b[n] ~ U ( 0 , 1), c[n] ~ U ( 0 , 1)

**Ensemble Mean:** E(a[n]-0.5)\*b[n] + c[n]) = 0.5

**Ensemble Variance:**

Var**((**a[n]-0.5)\*b[n] + c[n]) = Var ((a[n]-0.5)\*b[n]) + Var**(**c[n]**)**

= E2(b[n]) + Var (b[n]) E2(a[n]-0.5) + Var () Var ()+ Var**(**c[n]**)**

=

=

σ = 1 / 3

**Process 3**

**Mathematical Expression:** (a[n]-0.5)\*3 + 0.5

a[n] ~ U ( 0 , 1)

**Ensemble Mean:** E (a[n]-0.5)\*3 + 0.5) = 0.5

**Ensemble Variance:** Var(a[n]-0.5)\*3 + 0.5) = 9 \* Var (a[n]) = ¾

σ = = 0.866

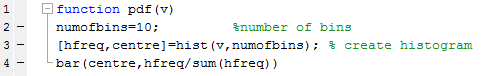
**Summary**

Results of ensemble mean and variance above match the curves plotted in Figure 8 and Figure 9.

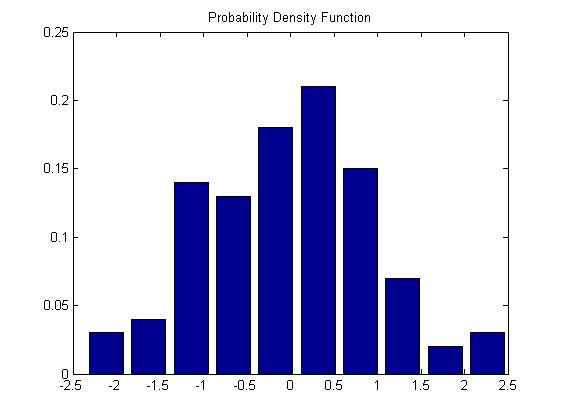
**1.3 Estimation of probability distributions**

**1.3.1 Probability Density Function design**

Function pdf takes v as input argument and plot a histogram. Bin number is chosen to 10 in this case since the sample number required to test is 100, 1000, 10000 respectively.

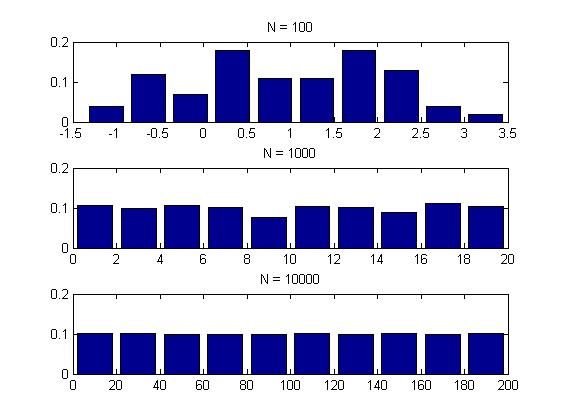
****

**Pdf of a Gaussian distribution with 100 samples is shown below:**

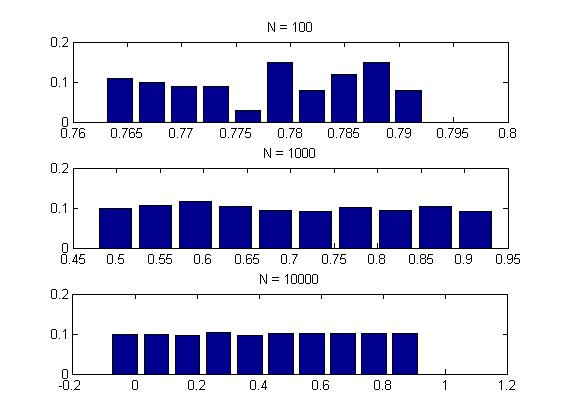


**1.3.2 Comparisons on pdf with different sample number**

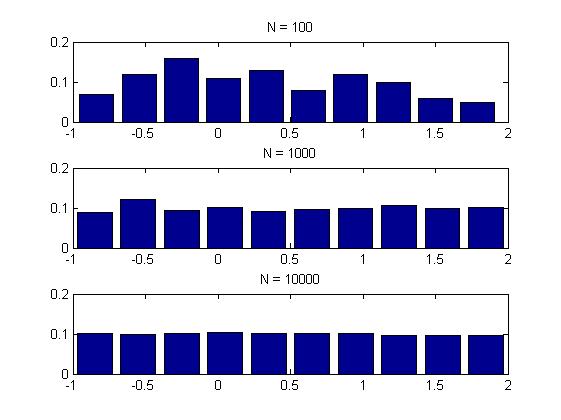
**Process 1:**

****

**Process 2:**

****

**Process 3:**

****

As N increases, variance converges

**1.3.3**