Stochastic Structural Dynamics

Lecture-25

Markov Vector Approach-5

Monte Carlo simulation approach-1

Dr C S Manohar
Department of Civil Engineering
Professor of Structural Engineering
Indian Institute of Science
Bangalore 560 012 India
manohar@civil.iisc.ernet.in

Summary : $p \equiv p\left(\tilde{x};t \mid \tilde{x}_0;t_0\right)$

$$dX(t) = f\left[t, X(t)\right]dt + G\left[t, X(t)\right]dB(t); t \geq 0; X(0) = X_0$$

$$\frac{\partial p}{\partial t} = -\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[f_j(x,t)\,p\right] + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left[\left(GDG^t\right)_{ij} p\right]$$

$$\frac{d}{dt}\left\langle h\left[X(t),t\right]\right\rangle = \left\langle \frac{\partial h}{\partial t}\right\rangle + \sum_{j=1}^{n}\left\langle f_j(X,t)\frac{\partial h}{\partial X_j}\right\rangle + \sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle \left(GDG^t\right)_{ij} \frac{\partial^2 h}{\partial X_i \partial X_j}\right\rangle$$

$$\left\langle X_1^m(t) X_1^m(t_0)\right\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^m v\left(x_1, x_2; t, t_0\right) dx_1 dx_2$$

$$v\left(x_1, x_2; t, t_0\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \eta_1^n p\left(x_1, x_2; t, t_0 \mid \eta_1, \eta_2; t, t_0\right) p\left(\eta_1, \eta_2; t, t_0\right) d\eta_1 d\eta_2$$

$$\frac{\partial v}{\partial t} = -\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[f_j(x,t)\,v\right] + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left[\left(GDG^t\right)_{ij} v\right]$$

PLUS : RELEVANT BCS & ICS

Summary (Continued) :

$$\frac{\partial p}{\partial t_0} = -\sum_{j=1}^{n} f_j\left[t_0,\tilde{x}_0\right]\frac{\partial p}{\partial x_{0j}} - \sum_{i=1}^{n}\sum_{j=1}^{n}\left[GDG^t\left(t_0,\tilde{x}_0\right)\right]_{ij}\frac{\partial^2 p}{\partial x_{0i}\partial x_{0j}}$$

$$R\left(t,\Gamma;t_0,\Omega\right) = \int_{\Omega} p\left(\tilde{x};t\,|\,\tilde{x}_0;t_0\right)d\tilde{x} = P\left[T > t - t_0\,|\,X\left(t_0\right) = \tilde{x}_0\right]$$

$$\frac{\partial R}{\partial t_0} = \left[-\sum_{j=1}^{n} f_j\left[t_0,\tilde{x}_0\right]\frac{\partial}{\partial x_{0j}} - \sum_{i=1}^{n}\sum_{j=1}^{n}\left[GDG^t\left(t_0,\tilde{x}_0\right)\right]_{ij}\frac{\partial^2}{\partial x_{0i}\partial x_{0j}}\right]R = \mathbf{L}R$$

$$\left(n+1\right)M_n + \mathbf{L}M_{n+1} = 0;\, n = 1,2,\ldots \qquad M_n = \langle T^n \rangle$$

PLUS : RELEVANT BCS & ICS

Illustration of deterministic averaging procedure

$$\ddot{u}+\omega_0^2 u=\varepsilon f\left(u,\dot{u}\right);t\geq 0;u\left(0\right)\&\dot{u}\left(0\right)\text{ specified.}$$

$$u\left(t\right)=a\left(t\right)\cos\left[\omega_0 t+\beta\left(t\right)\right]$$

$$\dot{u}\left(t\right)=-a\left(t\right)\omega_0\sin\left[\omega_0 t+\beta\left(t\right)\right]$$

$$\dot{a}\left(t\right)=-\frac{\varepsilon}{\omega_0}\sin\phi\left(t\right)f\left[a\left(t\right)\cos\phi\left(t\right),-a\left(t\right)\omega_0\sin\phi\left(t\right)\right]$$

$$\dot{\beta}\left(t\right)-\frac{\varepsilon}{\omega_0 a\left(t\right)}\cos\phi\left(t\right)f\left[a\left(t\right)\cos\phi\left(t\right),-a\left(t\right)\omega_0\sin\phi\left(t\right)\right]$$

$$\dot{a}\left(t\right)\approx-\frac{\varepsilon}{\omega_0}\frac{1}{2\pi}\int_0^{2\pi}\sin\phi f\left[a\cos\phi,-a\omega_0\sin\phi\right]d\phi$$

$$\dot{\beta}\left(t\right)\approx-\frac{\varepsilon}{\omega_0 a}\frac{1}{2\pi}\int_0^{2\pi}\cos\phi f\left[a\cos\phi,-a\omega_0\sin\phi\right]d\phi$$

Extension to randomly driven systems

$$\ddot{x}+\varepsilon^2 h(x,\dot{x})+\omega_0^2 x=\varepsilon z(t)$$

$$\langle z(t)\rangle=0;\langle z(t)z(t+\tau)\rangle=R_{zz}(\tau)\Leftrightarrow S(\omega)$$

$z(t)$ is taken to be broad banded.

•Characteristic time constant of excitation>> $\ll$
charateristic time constant of the system

•Time duration over which $R_{zz}(\tau)$ decays to 10% of $R_{zz}(0)$ >> $\ll$ Time duration over which the impulse response of the system decays by 90%.

•ε is small parameter.

$$\varepsilon=0\Rightarrow x(t)=a\cos(\omega_0 t+\phi)=a\cos\Phi$$ No excitation

$$\dot{x}(t)=-a\omega_0\sin(\omega_0 t+\phi)=-a\omega_0\sin\Phi$$

$$\varepsilon = 0 \Rightarrow$$

$$x(t) = a\cos(\omega_0 t + \phi) = a\cos\Phi; \Phi = \omega_0 t + \phi$$

$$\dot{x}(t) = -a\omega_0 \sin(\omega_0 t + \phi) = -a\omega_0 \sin\Phi$$

$$\varepsilon \neq 0 \Rightarrow$$

$$x(t) = a(t)\cos\left[\omega_0 t + \phi(t)\right] = a(t)\cos\Phi(t)$$

$$\dot{x}(t) = -a(t)\omega_0 \sin\left[\omega_0 t + \phi(t)\right] = -a(t)\omega_0 \sin\Phi(t)$$

$$\Rightarrow$$

$$\dot{a} = \frac{\varepsilon^2}{\omega_0} h\left[a\cos\Phi, -a\omega_0 \sin\Phi\right]\sin\Phi - \frac{\varepsilon z(t)}{\omega_0}\sin\Phi$$

$$\dot{\phi} = \frac{\varepsilon^2}{a\omega_0} h\left[a\cos\Phi, -a\omega_0 \sin\Phi\right]\cos\Phi - \frac{\varepsilon z(t)}{a\omega_0}\cos\Phi$$

- No approximations till now.

Averaging

- Two stage
  - ◦ Deterministic $\Rightarrow$ Replace "regular" oscillatory terms by their time averages
  - ◦ Stochastic $\Rightarrow$ Replace randomly fluctuating oscillatory terms by delta correlated processes

First stage follows the procedure used in deterministic averaging.
The second stage is based on the application of the Stratonovich-Khasminiski theorem.

Stratonovich - Khasminiski theorem

Consider the equation of motion

$$\dot{X} = \varepsilon^2 f\left[X,t\right] + \varepsilon g\left[X,t,Y(t)\right]; t \geq 0; X(0) \text{ specified.}$$

- ε=a small parameter
- $X(t) \sim n \times 1$ vector of response processes
- $Y(t) \sim m \times 1$ vector of random excitations

$\mathrm{E}\left[Y(t)\right] = 0; Y(t)$ is broad banded.

According to the Stratonovich-Khasminiski theorem the above equation can be approximated by a SDE

$$dX(t) = \varepsilon m(X)dt + \sigma(X)dB(t)$$

$$dX(t) = \varepsilon m(X)dt + \sigma(X)dB(t)$$

$$m = \mathrm{T}^{\mathrm{av}} E\{f\} + \int_{-\infty}^{0} E\left\{\left(\frac{\partial g}{\partial X}\right)_t \left(g^t\right)_{t+\tau}\right\} d\tau$$

$$\sigma\sigma^t = \mathrm{T}^{\mathrm{av}} \int_{-\infty}^{\infty} E\left\{(g)_t \left(g^t\right)_{t+\tau}\right\} d\tau$$

$$\mathrm{T}^{\mathrm{av}}\{\bullet\} = \lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{t_0+T} \{\bullet\} dt$$

Reference :

J B Roberts and P D Spanos, 1986, Stochastic averaging: an approximate method of solving nonlinear random vibration problems, Invited Review, International Journal of Nonlinear Mechanics, 21(2),111-134.

Exact

$$\dot{a} = \frac{\varepsilon^2}{\omega_0} h[a\cos\Phi, -a\omega_0 \sin\Phi]\sin\Phi - \frac{\varepsilon z(t)}{\omega_0}\sin\Phi$$

$$\dot{\phi} = \frac{\varepsilon^2}{a\omega_0} h[a\cos\Phi, -a\omega_0 \sin\Phi]\cos\Phi - \frac{\varepsilon z(t)}{a\omega_0}\cos\Phi$$

Averaging $\Rightarrow$

$$da(t) \approx -\frac{\varepsilon^2}{\omega_0} F(a)dt + \frac{\pi S(\omega_0)}{2a\omega_0^2}dt - \frac{[\pi S(\omega_0)]^{\frac{1}{2}}}{\omega_0} dB_1(t)$$

$$d\phi(t) \approx -\frac{\varepsilon^2}{a\omega_0} G(a)dt - \frac{[\pi S(\omega_0)]^{\frac{1}{2}}}{a\omega_0} dB_2(t)$$

$$F(a) = -\frac{1}{2\pi}\int_0^{2\pi} h[a\cos\Phi, -a\omega_0 \sin\Phi]\sin\Phi d\Phi$$

$$G(a) = -\frac{1}{2\pi}\int_0^{2\pi} h[a\cos\Phi, -a\omega_0 \sin\Phi]\cos\Phi d\Phi$$

$$da(t) \approx -\frac{\varepsilon^2}{\omega_0} F(a)\,dt + \frac{\pi S(\omega_0)}{2a\omega_0^2}\,dt - \frac{\left[\pi S(\omega_0)\right]^{\frac{1}{2}}}{\omega_0}\,dB_1(t)$$

$$d\phi(t) \approx -\frac{\varepsilon^2}{a\omega_0} G(a)\,dt - \frac{\left[\pi S(\omega_0)\right]^{\frac{1}{2}}}{a\omega_0}\,dB_2(t)$$

$\begin{Bmatrix} a(t) \\ \phi(t) \end{Bmatrix}$ is a Markov vector; **more interestingly** $\{a(t)\}$ is Markov.

- Forward equation : transient and steady state solutions
- One and two time moment equations
- Backward equation
- Reliability function
- GPV equations

FPK equation governing $p\left(a,\phi;t\,|\,a_0,\phi_0;t_0\right)$

$$\frac{\partial p}{\partial t}=-\frac{\partial}{\partial a}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F\left(a\right)+\frac{\pi S\left(\omega_0\right)}{2a\omega_0^2}\right\}p\right]-\frac{\partial}{\partial\phi}\left[\left\{-\frac{\varepsilon^2}{a\omega_0}G\left(a\right)\right\}p\right]+$$

$$+\frac{\pi S\left(\omega_0\right)}{2\omega_0^2}\left[\frac{\partial^2 p}{\partial a^2}+\frac{1}{a^2}\frac{\partial^2 p}{\partial\phi^2}\right]$$

Steady state

$$0=-\frac{\partial}{\partial a}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F\left(a\right)+\frac{\pi S\left(\omega_0\right)}{2a\omega_0^2}\right\}p\right]-\frac{\partial}{\partial\phi}\left[\left\{-\frac{\varepsilon^2}{a\omega_0}G\left(a\right)\right\}p\right]+$$

$$+\frac{\pi S\left(\omega_0\right)}{2\omega_0^2}\left[\frac{\partial^2 p}{\partial a^2}+\frac{1}{a^2}\frac{\partial^2 p}{\partial\phi^2}\right]$$

FPK equation governing $p\left(a;t\,|\,a_0;t_0\right)$

$$\frac{\partial p}{\partial t}=-\frac{\partial}{\partial a}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F(a)+\frac{\pi S(\omega_0)}{2a\omega_0^2}\right\}p\right]+\frac{\pi S(\omega_0)}{2\omega_0^2}\frac{\partial^2 p}{\partial a^2}$$

Steady state

$$-\frac{d}{da}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F(a)+\frac{\pi S(\omega_0)}{2a\omega_0^2}\right\}p\right]+\frac{\pi S(\omega_0)}{2\omega_0^2}\frac{d^2 p}{da^2}=0$$

$\Rightarrow$

$$\left\{\frac{\varepsilon^2}{\omega_0}F(a)+\frac{\pi S(\omega_0)}{2a\omega_0^2}\right\}p+\frac{\pi S(\omega_0)}{2\omega_0^2}\frac{dp}{da}=0$$

$$p(a)=Ca\exp\left\{-\frac{2\varepsilon\omega_0}{\pi S(\omega_0)}\int_0^a F(s)\,ds\right\};0<a<\infty$$

Remarks

$$0=-\frac{\partial}{\partial a}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F(a)+\frac{\pi S(\omega_0)}{2a\omega_0^2}\right\}p\right]-\frac{\partial}{\partial\phi}\left[\left\{-\frac{\varepsilon^2}{a\omega_0}G(a)\right\}p\right]+$$

$$+\frac{\pi S(\omega_0)}{2\omega_0^2}\left[\frac{\partial^2 p}{\partial a^2}+\frac{1}{a^2}\frac{\partial^2 p}{\partial\phi^2}\right]$$

$$0=-\frac{\partial}{\partial a}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F(a)+\frac{\pi S(\omega_0)}{2a\omega_0^2}\right\}p\right]+\frac{\pi S(\omega_0)}{2\omega_0^2}\frac{\partial^2 p}{\partial a^2}$$

$$\Rightarrow p(a,\phi)=\frac{p(a)}{2\pi};0<a<\infty;0<\varphi<2\pi$$

$$\Rightarrow p(u,\dot{u})=\frac{p(a)}{2\pi a}\bigg|_{a=\sqrt{x^2+\left(\frac{x}{\omega_0^2}\right)}};-\infty<u,\dot{u}<\infty$$

Remarks (continued)

•The transient solution of

$$\frac{\partial p}{\partial t}=-\frac{\partial}{\partial a}\left[\left\{-\frac{\varepsilon^2}{\omega_0}F(a)+\frac{\pi S(\omega_0)}{2a\omega_0^2}\right\}p\right]+\frac{\pi S(\omega_0)}{2\omega_0^2}\frac{\partial^2 p}{\partial a^2}$$

$p(a;t|a_0;t_0)$

$$p(a;0|a_0;0)=\delta(a-a_0)$$

can be obtained by using eigenfunction method.

•Similar approximate solutions for first passage times can also be obtained.

•The formulaiton can be generalized to deal with systems with random parametric excitation such as in

$$\ddot{u}+\varepsilon^2 h(u,\dot{u})+\omega_0^2\left[u+\varepsilon\varsigma(t)\right]=\varepsilon\xi(t);t\geq 0;u(0)\,\&\,\dot{u}(0)\text{ specified.}$$

$\varsigma(t)\,\&\,\xi(t)$: broad band, zero mean random excitations

Special case

$\ddot{u} + 2\eta\omega\dot{u} + \omega^2 u = z(t)$

$$\varepsilon^2 h(x, \dot{x}) = 2\eta\omega_0 \dot{u}$$

$$p(a) = \frac{a}{\sigma^2}\exp\left(-\frac{a^2}{2\sigma^2}\right); 0 < a < \infty$$

$$p(\phi) = \frac{1}{2\pi}; 0 < \phi < 2\pi$$

$$\sigma^2 = \frac{\pi S(\omega_0)}{2\eta\omega_0^3}$$

Note: compare this with results on envelope and peak distribution obtained earlier.

Remarks (continued)

•The method can also be generalized to deal with systems with nonlinear stiffness:

$$\ddot{u}+\varepsilon^2 h(u,\dot{u})+\omega_0^2\Lambda(u)\left[1+\varepsilon\varsigma(t)\right]=\varepsilon\xi(t)$$

$$t\geq 0; u(0)\,\&\,\dot{u}(0)\text{ specified}$$

The definition of the envelope here needs to be modified sutiably as

$$V(t)=\frac{\dot{x}^2}{2}+\omega_0^2\int_0^u \Lambda(s)\,ds$$

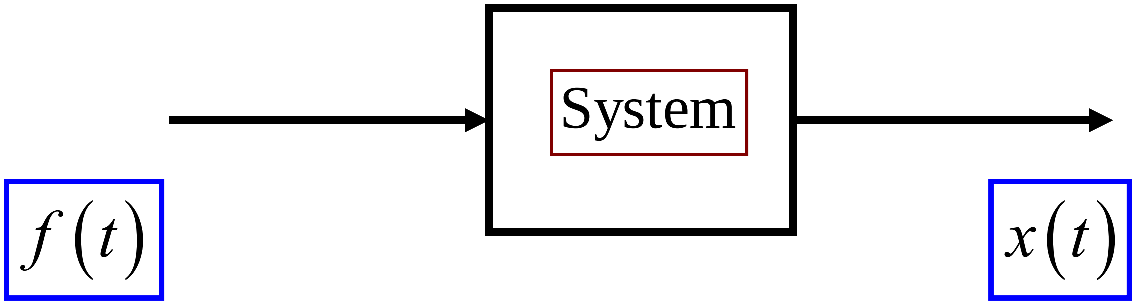
The method leads to a Markov approximation to the process $V(t)$.

Summary

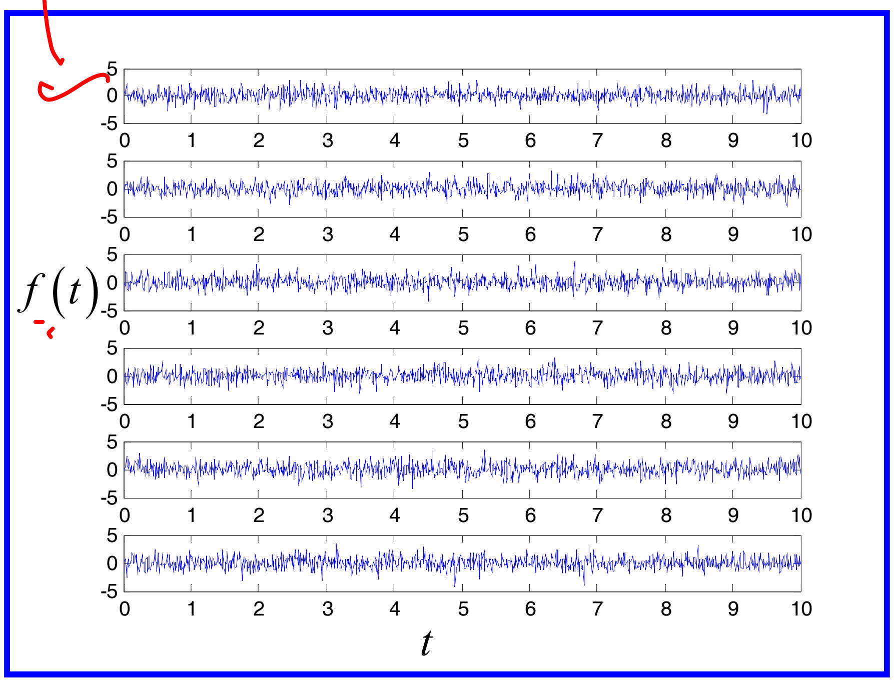
- Method of stochastic averaging enables us to study envelope and phase processes associated with weakly nonlinear system response to broad band excitations.
- The method also provides a framework to study first passage problems for the response envelope.
- The method is best suited to the study of sdof systems

Monte Carlo Simulation Methods in Stochastic Structural Dynamics

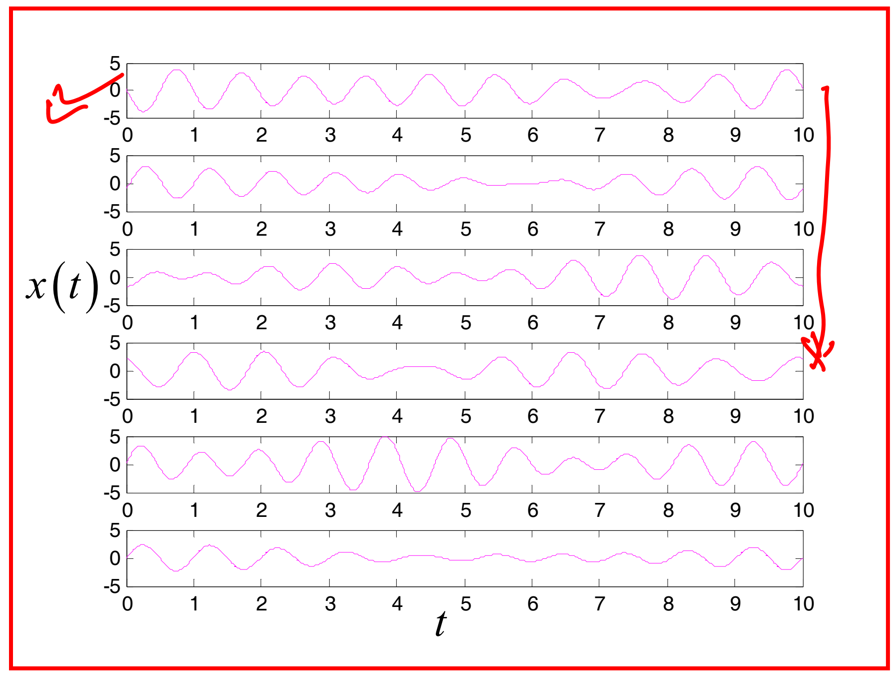
$$m\ddot{x}+c\dot{x}+g\left[x(t),\dot{x}(t)\right]=f(t);x(0)=x_0;\dot{x}(0)=\dot{x}_0$$



Ensemble of inputs



Ensemble of outputs



Generate ensemble of inputs obeying prescribed model for $f(t)$

System

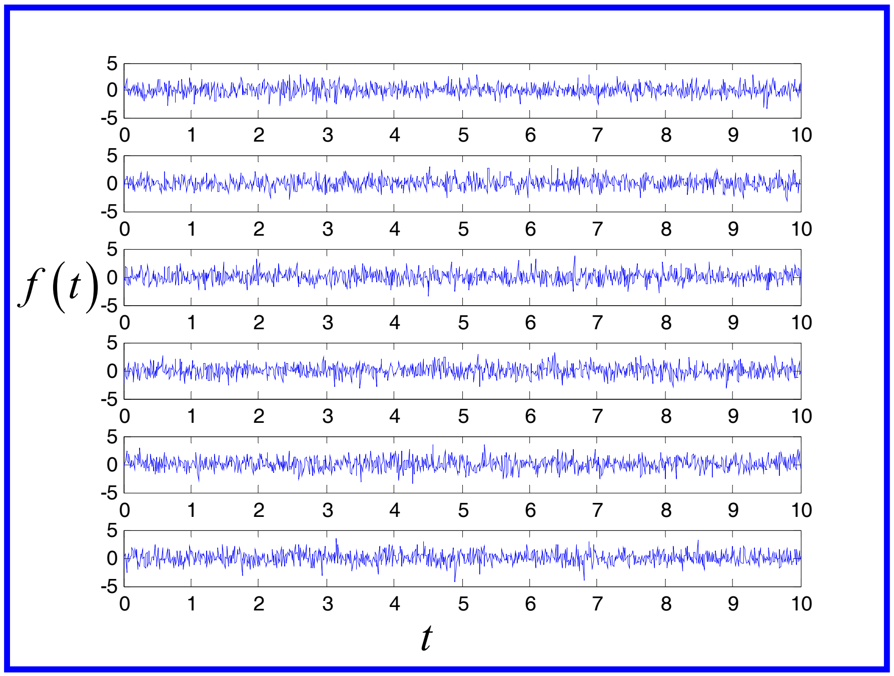
Process ensemble of outputs using statistical tools and arrive at probabilistic model for $x(t)$

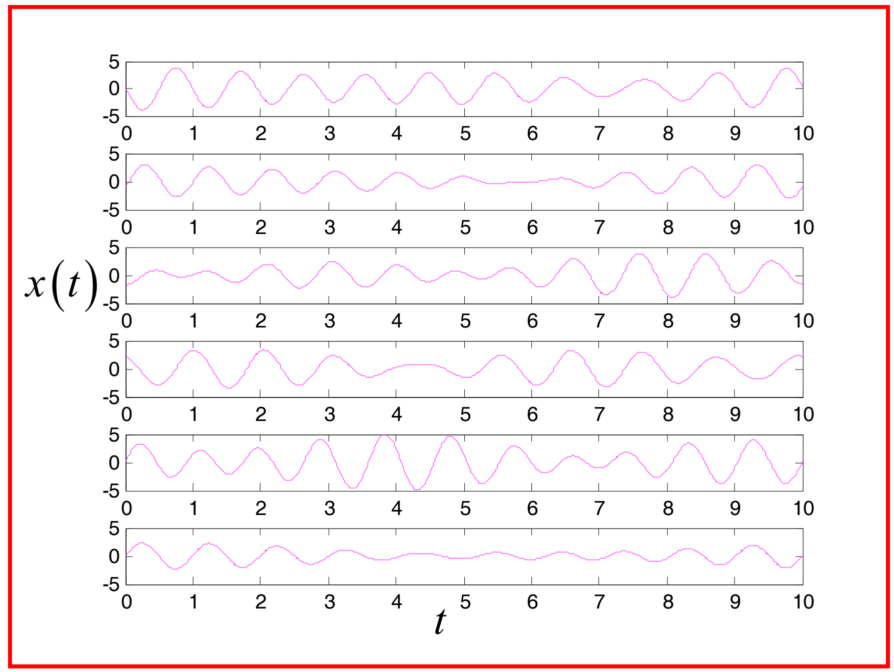
$f(t)$

$x(t)$

Ensemble of inputs

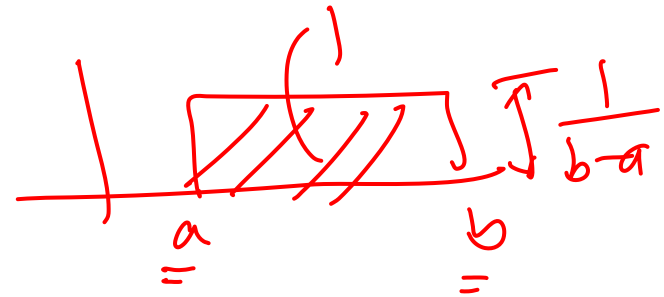
Ensemble of outputs





Another perspective

Consider the problem of evaluation of the definite integral $I = \int_a^b f(x)dx$.



This can be re-written as

$$I = (b-a)\int_a^b f(x)\left(\frac{1}{b-a}\right)dx = (b-a)\int_a^b f(x)\,p_X(x)\,dx$$

where $p_X(x) = \left(\frac{1}{b-a}\right); a < x < b$ is now interpreted as the pdf of a random variable that is uniformly distributed in a to b.

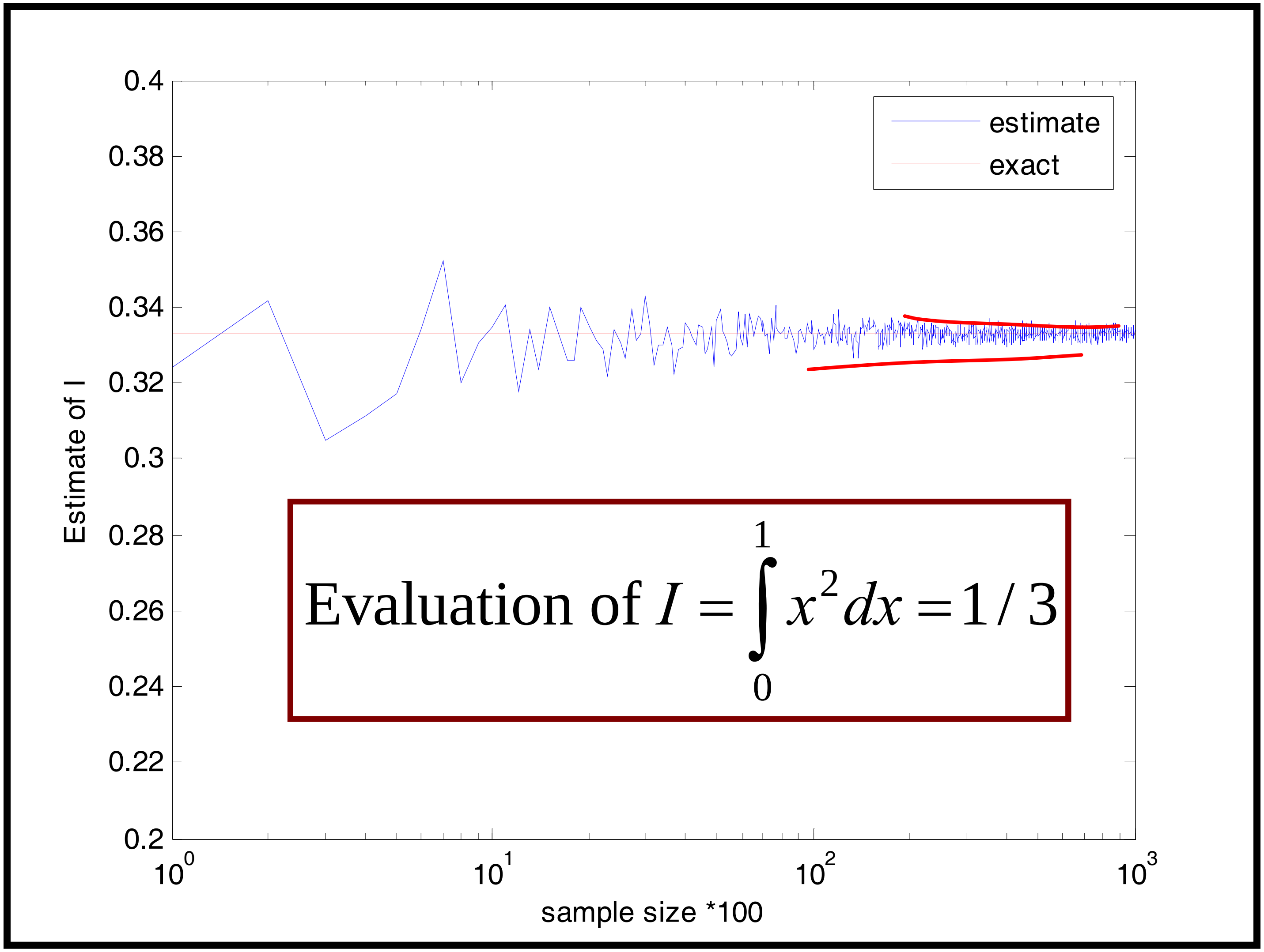
Following this, the integral I is now interpreted as an expectation

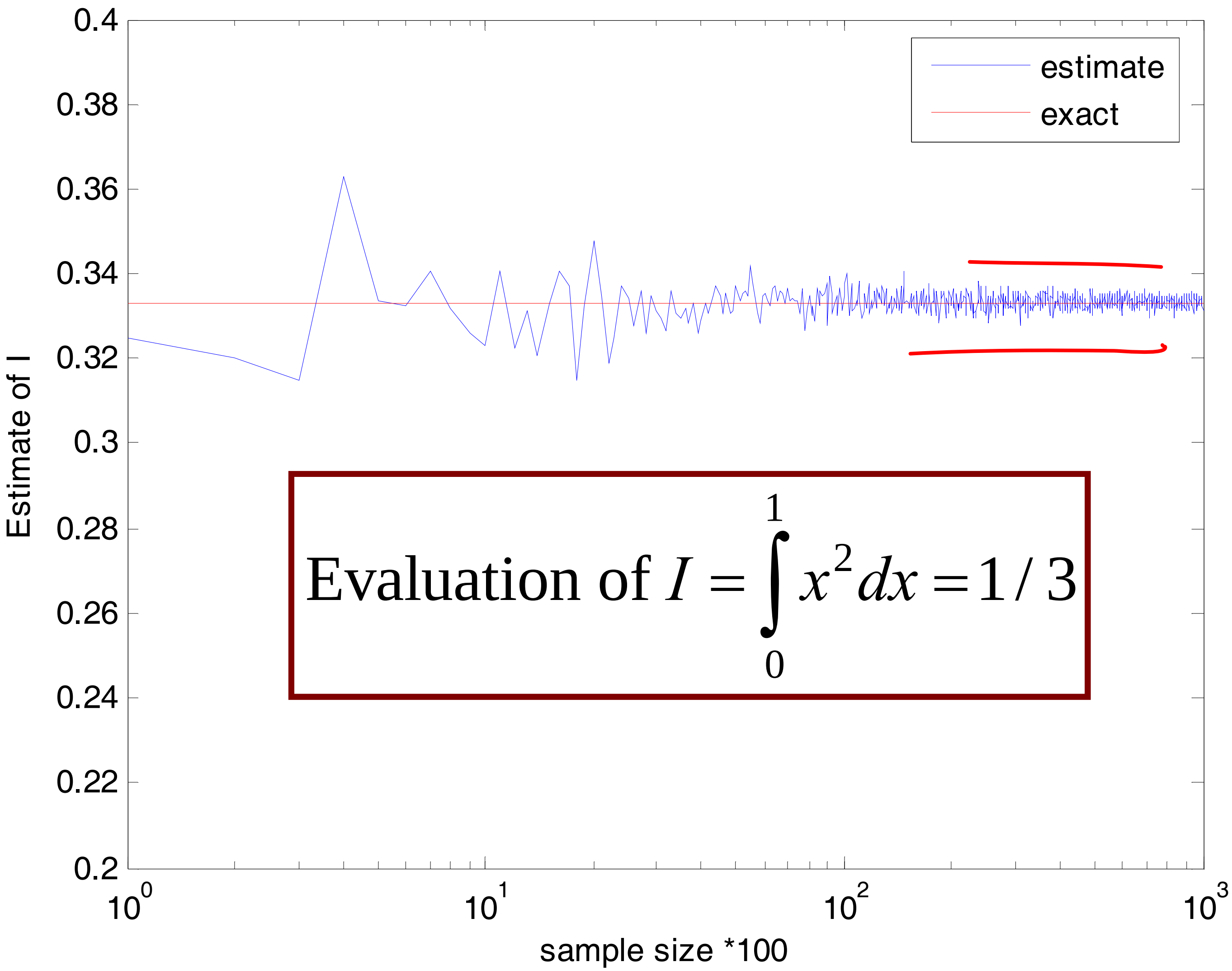
$$I = (b - a)\left\langle f\left(X\right)\right\rangle$$ wrt $p_x(x) \sim U(a,b)$

where the expectation is evaluated with respect to $p_X\left(x\right)$. Furthermore, I is now approximated by

$$\hat{I} = \frac{(b-a)}{N}\sum_{i=1}^{N} f(X_i)$$

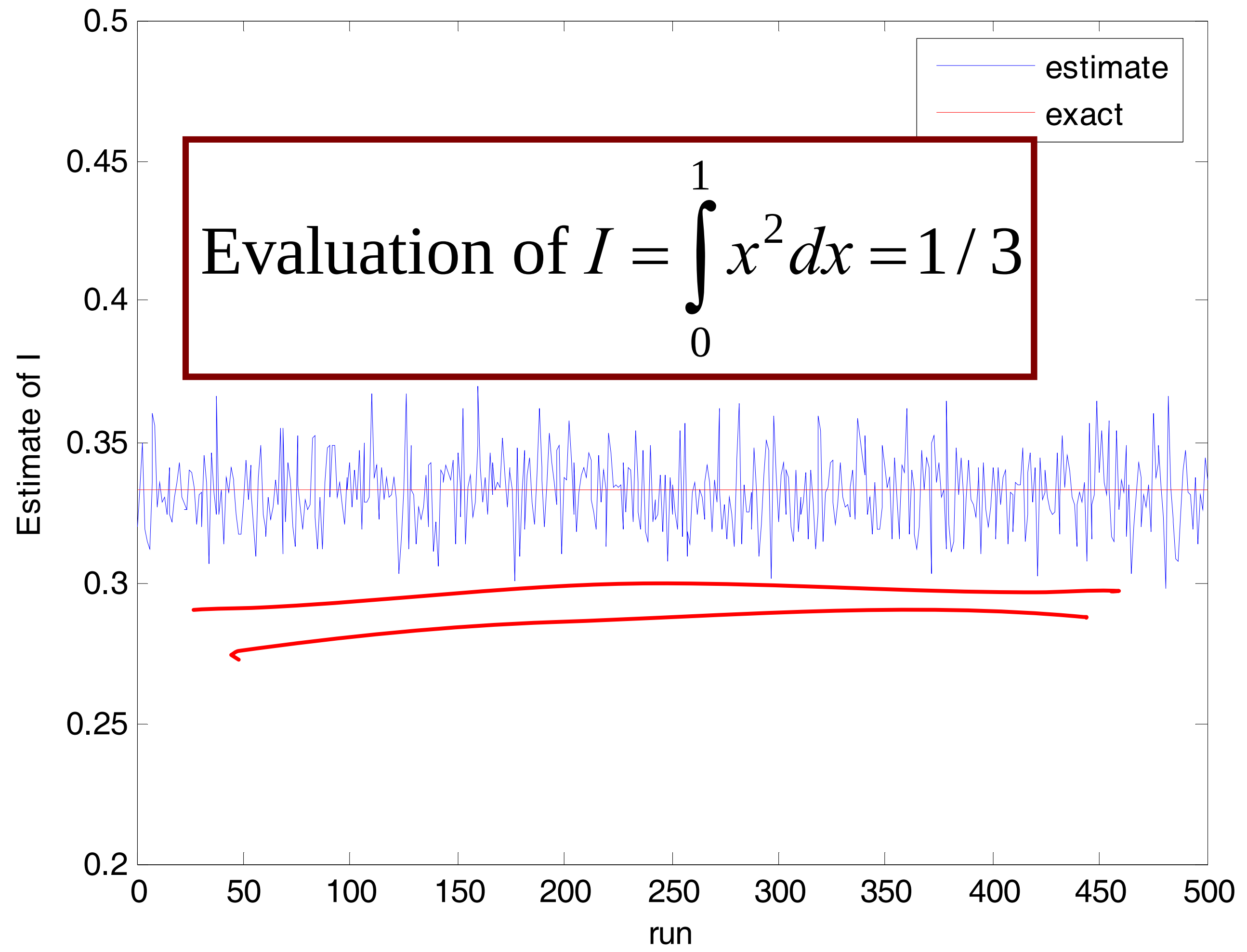
where X_i-s are uniformly distributed random numbers samples from $p_X\left(x\right)$.


0.4
0.38
0.36
0.34
0.32
0.3
0.28
0.26
0.24
0.22
0.2
Estimate of I
10^0
10^1
10^2
10^3
sample size *100
estimate
exact
Evaluation of $I = \int_0^1 x^2 dx = 1/3$



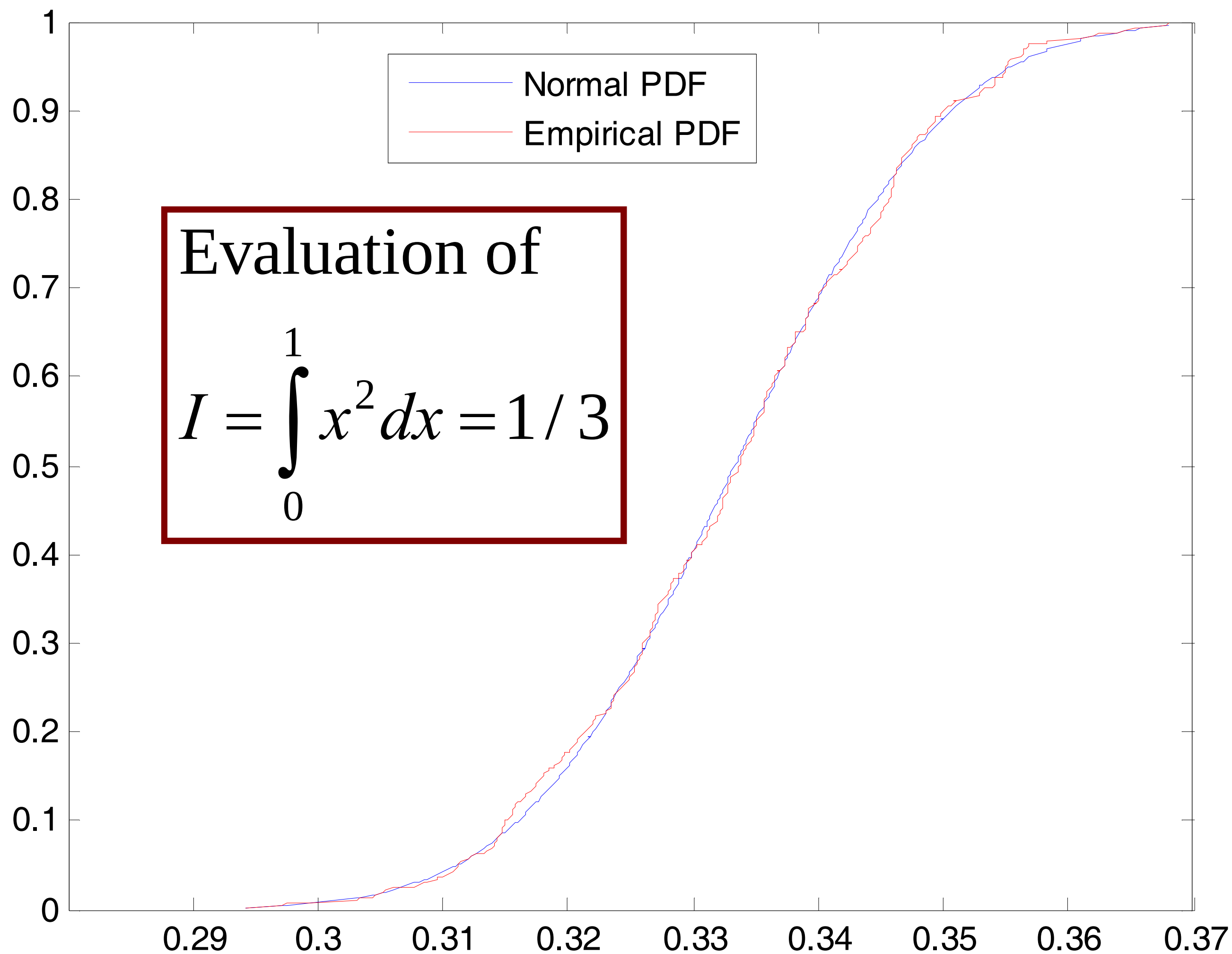
$$\text{Evaluation of } I = \int_0^1 x^2 dx = 1/3$$

500 runs with 500 samples



Evaluation of $I = \int_0^1 x^2 dx = 1/3$

Estimate of PDF



Evaluation of

$$I = \int_0^1 x^2 dx = 1/3$$

Ingredients of MCS

- Methods for generating samples of excitations and system parameters compatible with the prescribed probabilistic models
- Test statistically if the generated samples indeed obey the prescribed probabilistic laws.
- A computational model for the system dynamics which accepts samples of inputs and system paramters produced above and generates an ensemble of response quantities.
- Statistical processing of ensmble of response time histories and inferences on system behavior

We will begin with a review of elements of statistical methods

Statistics

(a) Data (used in plural) (birth, death, marraige).

(b) Science of statistics (used in singular).

(c) Statistic: a random variable; statistics: a set of random variables.

(It is in this sense that we use the word in the present course).

.

Average: a single number that describes data.

A material is described by its density, viscosity, stiffness, strength etc.

In the same sense there exist different measures to describe data-
e.g., arithmetic mean, geometric mean, mode, median, percentile, range, minimum, maximum, variance, standard deviation, skewness, kurtosis, histogram, cumulative frequency distribution, correlation, etc.

Population

Campus with 5000 persons.

Height $X_1 \quad X_2 \quad \cdots \quad X_{5000}$

Weight $Y_1 \quad Y_2 \quad \cdots \quad Y_{5000}$

Income $I_1 \quad I_2 \quad \cdots \quad I_{5000}$

specs? $Y \quad N \quad \cdots \quad Y$

gender M $F \quad \cdots \quad F$

In statistics each of this is a **population**.

That is, population of heights, population of weights, etc.

Population (Universe)

is a collection of all possible observations on a particular characteristic with respect to the problem on hand.

-starting point in statistics

-analogous to sample space in probability.

Any collection of measurements capable of being described by a random variable constitutes a population.

Sample

In practice it is not possible to study the entire population.

Sample is a part of the population which we want to study and draw conclusions about property of population.

-it is not enough to say that sample is a subset of population; the subset needs to be representative.

Sampling: Procedure of drawing samples.

Sampling design: development of sampling procedures to meet a requirement.

Random sample

Let X be a random variable with pdf $p_X(x)$.

Let $\{X_i\}_{i=1}^{n}$ be a set of iid random variables with common pdf $p_X(x)$.

The set of random variables $\{X_i\}_{i=1}^{n}$ is called a random sample of size n of X.

Consider the real valued function $S(X_1, X_{2,} \cdots, X_n)$.

This function is called a statistic. It is a random variable.

Let the pdf $p_X(x)$ be of the form $p_X(x;\theta)$ where

θ = unknown parameter.

The joint pdf of $\{X_i\}_{i=1}^n$ is of the form

$$p_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n) = \prod_{i=1}^n p_X(x_i;\theta)$$

$\{x_i\}_{i=1}^n$ = values of observed data taken from the random sample.

An estimator of θ is a statistic $S(X_1, X_2, \cdots, X_n)$ denoted as $\Theta = S(X_1, X_2, \cdots, X_n)$

For a particular set of observations $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$, the value of the estimator $S(x_1, x_2, \cdots, x_n)$ is called as an estimate of θ and is denoted by $\hat{\theta}$.

Estimator: a random variable

Estimate: the realization of the estimator.

Consider a population of four numbers x=[1 2 3 4]^t .
Population mean=2.5.

N=1

Samples	Mean
1	1
2	2
3	3
4	4
Mean	2.5
Std dev	1.29

N=2

Samples	Mean
1,2	1.5
1,3	2.0
1,4	2.5
2,3	2.5
2,4	3.0
3,4	3.5
Mean	2.5
Std dev	0.7071

N=3

Samples	Mean
1,2,3	2.0
1,2,4	7/3
2,3,4	3.0
1,3,4	8/3
Mean	2.5
Std dev	0.4303

N=4

Sample	Mean
1,2,3,4	2.5
Mean	2.5
Std dev	0.0

0

Sampling PDF

Estimator: $T = \frac{1}{n}\sum_{i=1}^{n} X_i$

The PDF of T is known as the sampling distribution of T.

A realization of T is known as an estimate.

The estimator is said to be unbiased if $\langle T \rangle =$ population mean.

The estimator is said to be consistent if $\lim_{n\to\infty} \mathrm{Var}(T) \to 0$.

Estimation : Finding a realization of T as an approximation to a population parameter.

Estimation of mean

Let X be a random variable with PDF $P_X(x)$, pdf $p_X(x)$, mean μ, and standard deviation σ.

Let $\{X_i\}_{i=1}^{n}$ be an iid sequence with common pdf $p_X(x)$.

That is, $X_i \perp X_j \forall i \neq j \in [1,n]$,

$\langle X_i \rangle = \mu,\ Var[X_i] = \sigma^2, p_{X_i}(x) = p_X(x) \forall i \in [1,n].$

Let $\Theta = \sum_{i=1}^{n} a_i X_i$ be an estimator of μ.

The estimator is said to be unbiased if $\langle \Theta \rangle = \mu$.

$$\langle \Theta \rangle = \left\langle \sum_{i=1}^{n} a_i X_i \right\rangle = \mu \sum_{i=1}^{n} a_i. \quad = \mu$$

$$\Rightarrow \sum_{i=1}^{n} a_i = 1 \Rightarrow \Theta \text{ is an unbiased estimator of } \mu.$$

$$\sum_{i=1}^{n} a_i = 1 \Rightarrow \Theta \text{ is an unbiased estimator of } \mu.$$

$\Rightarrow$ The above unbiased estimator is not unique.

$$\mathrm{Var}(\Theta) = Var\left\{\sum_{i=1}^{n} a_i X_i\right\} = \left\langle \left\{\sum_{i=1}^{n} a_i (X_i - \mu)\right\}^2 \right\rangle = \sigma^2 \sum_{i=1}^{n} a_i^2$$

To get an unbiased estimator with minimum variance, we minimize

$$\mathrm{Var}(\Theta) = \sigma^2 \sum_{i=1}^{n} a_i^2 \text{ subject to the constraint } \sum_{i=1}^{n} a_i = 1.$$

Lagrangian

$$L = \sigma^2 \sum_{i=1}^{n} a_i^2 + \lambda \left\{\sum_{i=1}^{n} a_i - 1\right\}$$

Necessary conditions for optima

$$\frac{\partial L}{\partial a_k} = 0 \Rightarrow 2\sigma^2 a_k + \lambda = 0; k = 1, 2, \cdots, n. \Rightarrow a_k = -\frac{\lambda}{2\sigma^2}.$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow \sum_{i=1}^{n} a_i = 1 \Rightarrow \sum_{i=1}^{n}\left(-\frac{\lambda}{2\sigma^2}\right) = 1$$

$$\lambda = -\frac{2\sigma^2}{n} \Rightarrow a_k = \left(-\frac{1}{2\sigma^2}\right)\left(-\frac{2\sigma^2}{n}\right) \Rightarrow a_k = \frac{1}{n}.$$

The optimal $\text{Var}(\Theta) = \sigma^2 \sum_{i=1}^{n} \frac{1}{n^2} = \frac{\sigma^2}{n}.$

Summary:

$\Theta = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of μ with minimum variance and the lowest variance is $\frac{\sigma^2}{n}$.

Maximum likelihood estimation

Let X be a random variable with pdf $p_X(x;\theta)$.

Here θ is a vector of parameters of the distribution.

For the moment assume that θ is known.

Let $\{X_i\}_{i=1}^{n}$ be an iid sequence of random variables with the common pdf given by $p_X(x;\theta)$.

Consider now the function

$$p_{X_1 X_2 \cdots X_n}\left(x_1, x_2, \cdots, x_n; \theta\right) =$$

$$p_{X_1}\left(x_1; \theta\right) p_{X_2}\left(x_2; \theta\right) \cdots p_{X_n}\left(x_n; \theta\right) = \prod_{i=1}^{n} p_{X_i}\left(x_i; \theta\right)$$

For example, if X is exponentially distributed,

$$p_X\left(x; \lambda\right) = \lambda \exp\left(-\lambda x\right); x \geq 0.$$

$$\Rightarrow p_{X_1 X_2 \cdots X_n}\left(x_1, x_2, \cdots, x_n; \theta\right) =$$

$$\lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right); x_i \geq 0 \forall i \in \left[1, n\right].$$

Clearly,

$$\prod_{i=1}^{n} p_{X_i}\left(x_i;\theta\right)dx_i =$$

$$P\left(x_1 < X_1 \le x_1 + dx_1 \cap x_2 < X_2 \le x_2 + dx_2 \cap \cdots \cap x_n < X_n \le x_n + dx_n\right).$$

Maximum likelihood estimation (continued)

Let us now consider the case when θ is unknown and let us observe a sample $\{x_i\}_{i=1}^{n}$.

We interpret $\prod_{i=1}^{n} p_{X_i}(x_i;\theta)dx_i = L(\theta \mid x_1, x_2, \cdots, x_n)$

as the likelihood of making the observation $\{x_i\}_{i=1}^{n}$.

It is a function of observed samples $\{x_i\}_{i=1}^{n}$ and the unknown parameter vector θ.

Definition

The maximum likelihood estimator of θ is the value of θ for which $L(\theta \mid x_1, x_2, \cdots, x_n)$ is the maximum.

Example 1

Let $X \sim \lambda \exp(-\lambda x); x \geq 0.$

$$L(\lambda \mid t_1, t_2, \cdots, t_n) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} t_i\right)$$

$$\Rightarrow \ln L(\lambda \mid t_1, t_2, \cdots, t_n) = n \ln \lambda - \lambda \sum_{i=1}^{n} t_i$$

Let $\hat{\lambda}$ maximize this function.

$$\frac{\partial}{\partial \lambda} \ln L(\lambda \mid t_1, t_2, \cdots, t_n) = 0$$

$$\Rightarrow \frac{n}{\hat{\lambda}} - \sum_{i=1}^{n} t_i \Rightarrow \frac{1}{\hat{\lambda}} = \frac{1}{n} \sum_{i=1}^{n} t_i.$$

Recall $\langle X \rangle = \lambda$ and the above estimator is consistent with the unbiased estimator with minimum variance derived earlier.

Example 2

Let $X \sim N(\mu, \sigma)$.

$$L\left(\mu,\sigma \mid t_1,t_2,\cdots,t_n\right)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma}}\exp\left[-\frac{1}{2}\left(\frac{t_i-\mu}{\sigma}\right)^2\right]$$

$$L\left(\mu,\sigma \mid t_1,t_2,\cdots,t_n\right)=\left[\frac{1}{\sqrt{2\pi\sigma}}\right]^n\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i-\mu}{\sigma}\right)^2\right].$$

$$\ln L\left(\mu,\sigma \mid t_1,t_2,\cdots,t_n\right)=-n\ln\sqrt{2\pi}-n\ln\sigma-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i-\mu}{\sigma}\right)^2$$

$$\ln L\left(\mu,\sigma \mid t_1,t_2,\cdots,t_n\right) = -n\ln\sqrt{2\pi} - n\ln\sigma - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i-\mu}{\sigma}\right)^2$$

Let $\hat{\mu}$ & $\hat{\sigma}$ maximize the above function.

$$\Rightarrow \frac{\partial}{\partial\mu}\ln L\left(\mu,\sigma \mid t_1,t_2,\cdots,t_n\right) = 0$$

$$\& \frac{\partial}{\partial\sigma}\ln L\left(\mu,\sigma \mid t_1,t_2,\cdots,t_n\right) = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} t_i$$

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(t_i - \hat{\mu}\right)^2}$$

Sampling distribution for the estimator of mean

Consider the estimator $\Theta = \frac{1}{n}\sum_{i=1}^{n} X_i$

Θ is an unbiased estimator of μ with variance $\frac{\sigma^2}{n}$.

Let us consider the case in which σ^2 is known.

If X is Gaussian, it would mean that $\{X\}_{i=1}^{n}$ is an iid sequence of Gaussian random variables and consequently Θ would also be Gaussian distributed.

If X is not Gaussian, by virtue of central limit theorem, for large n, we may still consider Θ to be Gaussian.

It may be infered that $\Theta \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ or, $\frac{\Theta - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$.