Lecture 13

ECE 278 Mathematics for MS Comp Exam

UCSD Jacobs | Electrical and Computer Engineering

- A random variable that describes an event may arise as the sum of **many independent** random variables $\underline{x}_i$ for repeated instances of some underlying constituent event.
- If the variances of the random variables $\underline{x}_i$ are finite and equal, then the probability density function $p_{\underline{x}}(x)$ for the normalized sum

$$\underline{x} = \frac{1}{\sqrt{N}} \sum_{i}^{N} (\underline{x}_i - \langle \underline{x} \rangle)$$

as N goes to infinity will usually tend towards a **gaussian probability density function**
  - with mean $\langle \underline{x} \rangle$
- Statement is true irrespective of the functional form of the probability density functions $p_{\underline{x}_i}(x_i)$ of the individual constituent events.
- When the density functions are all the same, the formal statement is called the ***central limit theorem***.
- The central limit theorem explains why the gaussian probability density function and its variants are ubiquitous in statistical analysis.

- As an example, consider the probability density function of the normalized sum of N independent and identically distributed (IID) complex random variables $\underline{A}_i e^{\mathrm{i}\underline{\phi}_i}$ added to a constant $Ae^{\mathrm{i}\theta}$,
- Random variables $\underline{A}_i$ and $\underline{\phi}_i$ are zero-mean, independent, and identically-distributed random variables
- The probability density function of $\underline{\phi}_i$ is uniform over $[0, 2\pi)$.
- The resulting normalized sum is a complex random variable written as

$$\underline{S} \quad = \quad \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\underline{A}_i e^{\mathrm{i}\underline{\phi}_i} - Ae^{\mathrm{i}\theta} \right)$$

UCSD Jacobs | Electrical and Computer Engineering

- This can be written as

$$\underline{S} = \underline{x} + \mathrm{i}\underline{y}$$

where

$$\underline{x} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\underline{A}_i \cos \underline{\phi}_i - A \cos \theta)$$

and

$$\underline{y} = \frac{1}{\sqrt{N}} \sum_{k=i}^{N} (\underline{A}_i \sin \underline{\phi}_i - A \sin \theta)$$

- In the limit as N goes to infinity, asserting the central limit theorem yields a joint probability density function $f_{\underline{x},\underline{y}}(x, y)$ for $\underline{S}$ that is a circularly-symmetric gaussian probability density function centered on the constant $Ae^{\mathrm{i}\theta}$.

- .

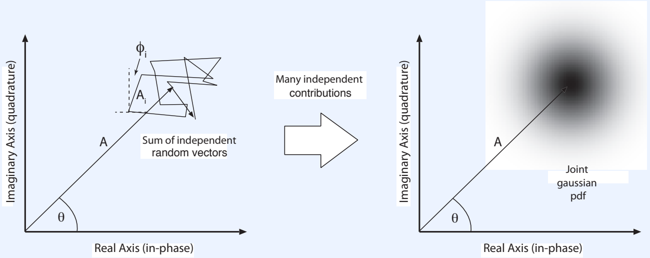

Figure: The limit of the sum of many independent random vectors superimposed on a constant signal is a circularly-symmetric gaussian probability density function centered on the constant $Ae^{i\theta}$.

The Central Limit Theorem-4

- Although the central limit theorem is quite powerful, the convergence to a gaussian distribution is not complete for any finite number of summand random variables.
- This means that calculations of **small probabilities of events** by using the central limit theorem to validate the use of a gaussian distribution may not be valid.

- Start with the expression

$$\underline{x} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\underline{x}_i - \langle \underline{x} \rangle)$$

and find the characteristic function

$$C_x(\omega) = \langle e^{i\omega \underline{x}} \rangle = \left\langle \exp\left[i\omega N^{-1/2} \sum_{i=1}^{N} (\underline{x}_i - \langle \underline{x} \rangle) \right] \right\rangle$$

- Now sum inside exponential can be converged into a product outside so that

$$C_x(\omega) = \langle e^{i\omega \underline{x}} \rangle = \prod_{i=1}^{N} \left\langle \exp\left[i\omega \underbrace{N^{-1/2}(\underline{x}_i - \langle \underline{x} \rangle}_{\text{new random variable X}} \right] \right\rangle$$
$$= C_X(N^{-1/2}\omega)^N$$

where $C_X(\omega)$ is the characteristic function of the random variable $\underline{x} - \langle \underline{x} \rangle$

- Now expand the characteristic function $C_X(\omega)$ in a power series to give

$$C_X(\omega) \approx 1 - \frac{1}{2}\sigma^2\omega^2 + \theta(\omega)|\omega|^3 + \ldots$$

where the remainder term is bounded so that $\theta(\omega) < M$ for some M

- The reason there is no linear term in the expansion is that term in the expansion would be the mean, but we are expanding with the mean so the highest order term is the quadratic term
- Expansion may or may not be possible, but can be checked

- Now find $C_X(N^{-1/2}\omega)$

$$C_X(N^{-1/2}\omega) \approx 1 - \frac{\sigma^2\omega^2}{2N} + \theta(N^{-1/2}\omega)N^{-3/2}|\omega|^3 + \dots$$

and raise to the Nth power

$$\left(C_X(N^{-1/2}\omega)\right)^N \approx \left(1 - \frac{\sigma^2\omega^2}{2N} + \theta(N^{-1/2}\omega)N^{-3/2}|\omega|^3 + \dots\right)^N$$

- As N goes to infinity, there is only one term that remains

$$\left(C_X(N^{-1/2}\omega)\right)^N \approx \left(1 - \frac{\sigma^2\omega^2}{2N}\right)^N$$

- Now use

$$\lim_{N\to\infty}\left(1+\frac{x}{N}\right)^N = e^x$$

on the first term so that

$$\left(C_X(N^{-1/2}\omega)\right)^N = e^{-\sigma^2\omega^2/2}$$

- The characteristic equation is approximated by a gaussian
- Therefore the pdf is a zero-mean gaussian with variance σ^2

- Approximate the probability using a gaussian so that

$$\Pr(y_1 < \underline{y} < y_2) \approx \operatorname{erfc}\left(\frac{y_1}{\sigma_y}\right) - \operatorname{erfc}\left(\frac{y_2}{\sigma_y}\right)$$

  where the random variable $\underline{y}$ has zero mean and variance σ_y^2
- The random variable $\underline{y}$ is the sum of iid random variables

$$\underline{y} = \sum_{i=1}^{N} \underline{x}_i$$

  so that

$$\Pr(s_1 < \sum_{i=1}^{N} \underline{x}_i < s_2) \approx \operatorname{erfc}\left(\frac{s_1 - N\langle x\rangle}{\sqrt{N}\sigma_x}\right) - \operatorname{erfc}\left(\frac{s_2 - N\langle x\rangle}{\sqrt{N}\sigma_x}\right) \tag{1}$$

  where σ_x^2 is the variance of the underlying distribution

- A batch of resistor is characterized by a uniform probability distribution between 950 Ω and 1050 Ω with a mean of 1000 Ω
- What is the probability that 100 of these resistors in series has a total resistance that is within 0.5% of 100k Ω?
- **Solution**
- To apply the CLT, we need the variance of the uniform distribution which is given by (see first lecture for probability)

$$\sigma_{\underline{x}}^2 \doteq \langle (\underline{x} - \langle \underline{x} \rangle)^2 \rangle = \frac{1}{b-a} \int_a^b (x - \langle \underline{x} \rangle)^2 = \frac{1}{12}(b-a)^2 = \frac{100^2}{12}$$

$$= \frac{1}{b-a} \int_a^b \left(x - \tfrac{1}{2}(b-a)\right)^2$$

$$= \frac{1}{12}(b-a)^2 = \frac{100^2}{12}$$

where $a =$950 and $b =$ 1050 so that $b - a = 100$.

UCSD Jacobs | Electrical and Computer Engineering

- The lower limit for (1) for 0.5% is

$$s_1 = 0.995N\langle \underline{x} \rangle$$

where $N = 100$ and $\langle \underline{x} \rangle = 1000$

- Similarly,

$$s_2 = 1.005N\langle \underline{x} \rangle$$

so that (1) in Ω is

$$\text{erfc}\left(\frac{s_1 - N\langle x \rangle}{\sqrt{N}\sigma_x}\right) - \text{erfc}\left(\frac{s_2 - N\langle x \rangle}{\sqrt{N}\sigma_x}\right)$$

$$= \text{erfc}\left(-\frac{0.005(100)}{10(0.1/\sqrt{12})}\right) - \text{erfc}\left(\frac{0.005(100)}{10(0.1/\sqrt{12})}\right)$$

$$= 1 - 2\text{erfc}\left(0.005(100)\sqrt{12}\right)$$

$$= 1 - \text{erfc}\left(1.732\right) \approx 0.971$$