2023년 4월 6일 목요일 12:25

CM13-230406Thu

Announcements

1. Midterm exam will be taken from 19:00-22:00 on Thursday Apr. 13th at LG102.
2. HW#10 will be assigned.
3. Handouts, annotated handouts are available at http://cisl.postech.ac.kr/class/eece574/schedule.htm .
   Passwords are ...
4. Lecture videos are available at cisl.postech YouTube channel. Clickable time stamps, please.

Corrections; Clarifications; Common mistakes,misunderstandings, misconceptions, myth; Frequently Asked Questions; Conventional wisdom

1. Z~CN(0, sigma^2) vs. Z~CN(0, 2sigma^2)
2. Q for right tail, (generalized) Marcum's Q for outside the circle/hypersphere

$X \sim N\left(\begin{bmatrix} Re\{\mu\} \\ Im\{\mu\} \end{bmatrix}, \sigma^2 I_{2m}\right)$

(i) $X \sim N(\mu, \sigma)$ $\frac{r}{\sigma} - \frac{\mu}{\sigma}$ (ii) $\underline{X} \sim N(\underline{\mu}, \sigma^2 I_{2m})$ (iii) $\underline{Z} \sim CN(\underline{\mu}, 2\sigma^2 I_m)$ $\|\underline{\mu}\| = \sqrt{\|Re\{\underline{\mu}\}\|^2 + \|Im\{\underline{\mu}\}\|^2}$

$Pr(|X| \geq r) = Q\left(\frac{r-\mu}{\sigma}\right) + Q\left(\frac{r+\mu}{\sigma}\right) \overset{?}{=} Q_{\frac{1}{2}}\left(\frac{\mu}{\sigma}, \frac{r}{\sigma}\right)$ $\quad \frac{r}{\sigma} + \frac{\mu}{\sigma}$

$Pr(\|\underline{X}\| \geq r) = Q_m\left(\frac{\|\underline{\mu}\|}{\sigma}, \frac{r}{\sigma}\right)$ $\in \mathbb{R}^n = \mathbb{R}^{2m}$

$Pr(\|\underline{Z}\| \geq r) = ?$ $Q_m\left(\frac{\|\underline{\mu}\|}{\sigma}, \frac{r}{\sigma}\right)$

(iv) $\underline{Z} \sim CN(\underline{\mu}, \sigma^2 I_{\boxed{N}})$

$Pr(\|\underline{Z}\| \geq r) = ?$ $Q_N\left(\frac{\|\underline{\mu}\|}{\sigma/\sqrt{2}}, \frac{r}{\sigma/\sqrt{2}}\right)$

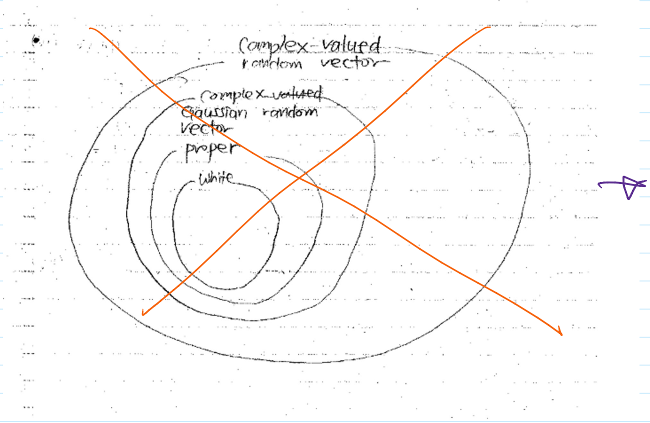
3. If Z~CN(mu, sigma^2 I_m), then any direction among 2m real directions is real Gaussian.

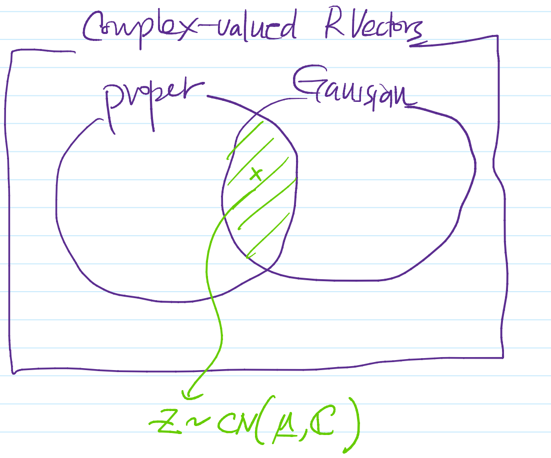
$Pr(Re\{\underline{Z}\} \geq r) = Q(\cdots)$, $Pr(\|\underline{Z}\| \geq r) =$

✓ $Pr(\|\underline{Z} - \underline{c}\| \geq \|\underline{Z} - \underline{\mu}\|) = Q(\cdots),$

$\underline{Z}$ ← complex, $\begin{bmatrix} Re\{\underline{Z}\} \\ Im\{\underline{Z}\} \end{bmatrix}$ real composite vector.

4. ...





Overview w/ Chapter numbers from the main textbook by Papoulis, (2)-(7) from Peebles

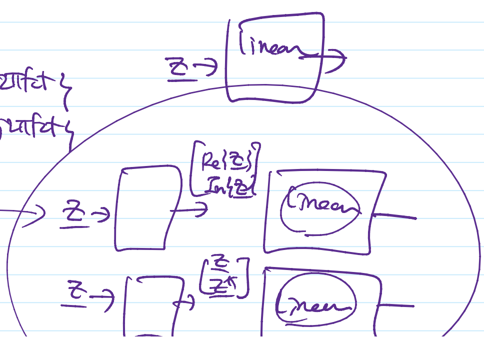
9. General Concepts
   1. RP concept
   (2) Stationarity and Independence
   (5) Gaussian RPs
   (6) Poisson RPs
   (7) Complex RPs
   (2) Time average and Ergodicity
   (3) Correlation functions
   (4) Measurement of correlation functions
7. Sequence of Random Variables
   9.1
   - positive semi-definiteness of a correlation function
   - alpha-dependent random process
   - correlation alpha-dependent random process
   - white noise in terms of correlation function
   - mean-square periodic random process
   - cyclostationary random process
10. Random Walks and Other Applications
11. Spectral Representation
12. Spectrum Estimation
13. Mean-Square Estimation

If WSS, then proper-complex

$x(t) = Re\{x_c(t)\, e^{j2\pi f_c t}\}$

$+) \; N(t) = Re\{N_c(t)\, e^{j2\pi f_c t}\}$

$y(t)$



Review

1. Why complex-valued signals and random processes in radio communications?
2. Why proper-complex random processes?
3. Why widely linear processing?

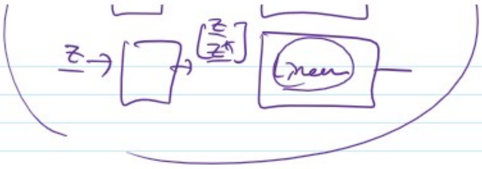
Preview

1. Inference vs. Statistical inference
2. Statistical signal processing and Random observation

EEEG45

Preview
1. Inference vs. Statistical inference
2. Statistical signal processing and Random observation
3. Detection and Estimation; Probability and Expectation
4. Observation of single sample path vs. Statistical inference
5. Time average and ergodicity

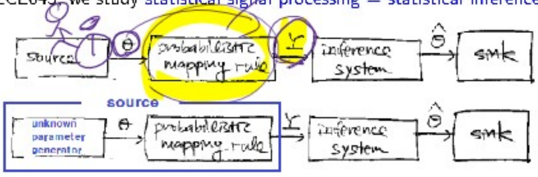
EECE645



6.2 (continued)

$f_{Y;\Theta}(y)$

- In EECE645, we study statistical signal processing = statistical inference.



$\theta \in \mathcal{X} = \{H, T\}$

$\theta \in \mathcal{X} = (0, 2\pi]$

$Y = \Theta + N$, $N \sim N(0, \sigma^2)$

$\theta = 1$ $Y_\theta \sim N(1, \sigma^2)$

$\theta = 0$ $Y_\theta \sim N(0, \sigma^2)$

$\{N(1, \sigma^2), N(0, \sigma^2)\}$

$f_\Theta(\theta)$

- A common signal and system model of statistical signal processing problems has
  1. a source that generates an unknown parameter $\underline{\theta} \in \mathcal{X}$ possibly with the prior probability distribution of Θ,
  2. a probabilistic mapping rule that has the parameter as the input and the observation as the output, $\{f_{Y;\underline{\theta}}(\underline{y}) : \underline{\theta} \in \mathcal{X}\}$ random
  3. an inference system to be designed has the observation as the input and generates an estimate as the output,
  4. a sink that admits the output of the system.
  - Often, the source and the mapping rule combined together is called a source.
  - cf) a communication problem

- In EECE645, we learn
  1. how to identify what class of problem I confront

     We encounter 12 classes:

     (a) what kind of observation $\underline{Y}$?
     - (scalar,) vector
       sequence, DT signal
       waveform, CT signal

     (b) unknown parameter to be inferred is in what set $\mathcal{X}$?
     - in a countable set ← detection problem
     - in an uncountable set ← point or interval estimation problems

     (c) unknown parameter has known *a priori* probabilistic distribution?
     - random parameter ← Bayesian problem
     - non-random parameter ← non-Bayesian or classical problem
  2. how to formulate an optimization problem
  3. how to solve the optimization problem
  4. how to analyze the performance of the
- The statistical inference problem is well defined if
  - the family of distribution functions $\{f_{\underline{Y};\underline{\theta}}(\underline{y}) : \underline{\theta} \in \mathcal{X}\}$ is completely given.

random observation

$\theta \in \mathcal{X}$ θ random ?

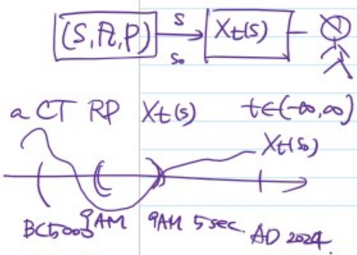
vector: D (O, X), E (O, X)

sequence DT signal: D, E

continuous time signal: D, E

$\frac{X_1 + \cdots + X_N}{N} \simeq E\{1_A\} = P(A)$

- Every statistical interference is about an ensemble average.
  - Even a probability is a kind of expectation.
  - Moreover, the expectation is meaningful when we can repeat the experiment.
- However, in many engineering problems involving a CT/DT signal X_t with uncertainty,
  - we observe only a single sample path and
  - we are not given statistical property of the process, but
  - we are asked do some inference.
  - Recall the discussion on the engineering implication of $t \in (-\infty, \infty)$, the ensemble, parallel universes, etc.



BC500 ... AD 2024.

– we are asked do some inference.

– Recall the discussion on the engineering implication of $t \in (-\infty, \infty)$, the ensemble, parallel universes, etc.

– Example: We are given an observation model $X(t) = R\cos(2\pi f_0 t + \Theta)$ with independent random variables R and Θ on $[0, 2\pi)$. But, we do not know the distributions of R and Θ.

$(f_0 t_0)$

event $\{X = x\}$

* Q1. Given a realization $\{X(t) = x(t), t \in (0, T]\}$ of the observation, predict this sample path at $t = 2T$.

$r_0\cos(2\pi f_0 t + \theta_0)$ an event

· A1. We can find the realizations r_0 and θ_0, and predict perfectly as $x(2T) = r_0 \cos(2\pi f_0 2T + \theta_0)$.

* Q2. Given a realization $\{X(t) = x(t), t \in (0, T]\}$ of the observation, predict this random process at $t = 2T$.

$X(2T)$

$x(t) = r_0\cos(2\pi f_0 t + \theta_0) \quad t \in (0, T]$

· A2. From the single sample path, we cannot obtain any information on distributions of R and Θ other than r_0 and θ_0 are in the sample spaces. So, we cannot make the statistical inference.

– Depending on the observation model, from a single sample path partially observed, either (i), or (ii) or (iii).

(i) we can obtain information on the distribution of the random process, ← entire sample paths,

(ii) we can obtain information on the conditional distribution of the random process, ← some sample paths,

(iii) we can obtain information only on the single sample path.

• We want to define a subset of random processes where an observation of a single sample path provides enough information on the statistical properties of the random process to perform some statistical inference.

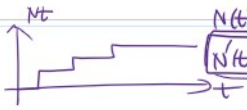
Boltzmann < 'ergon' work, 'hodos' path/way

– the subset of ergodic random processes

– As stationarity has many levels, ergodicity also has many levels.

– To extract some information, we use a time average instead of a statistical average. But we want the time average provide information on a statistical average.

Each level involves a different form of a time average linked with a statistical average.



$N(t) = \sum_{n=0}^{\infty} u(t - S_n)$

$\boxed{N'(t)} = \sum_{n=0}^{\infty} \delta(t - S_n)$

• Example. Given a Poisson process N_t with an unknown arrival rate λ, we observe a sample path $N_t(s)$ for $t \in (0, T]$. Let $\hat{\lambda}_1 \triangleq N_T/T$. Examine whether $\hat{\lambda} \to \lambda$ as $T \to \infty$.

$\boxed{N(T)}/T = \frac{1}{T}\int_0^T N'(t)\,dt$

A.

The mean and variance of a Poisson random variable N_T is λT and λT.

Hence, $\mathbf{E}[\hat{\lambda}] = \lambda$ and $\mathrm{Var}[\hat{\lambda}] = \lambda/T$.

That is, $\mathbf{E}[(\hat{\lambda} - \lambda)^2] = \lambda/T \to 0$ as $T \to \infty$.

$\therefore \hat{\lambda}_1 \triangleq \frac{N(T)}{T}$.

Therefore, $\hat{\lambda} \to \lambda$ in the MS sense.

– Ergodicity is often understood a very restrictive form of stationarity.

* However, this example shows that a non-stationary random process can be ergodic.

– Apparently, the time average of a sample path is not performed. in this example.

* Of course, actually the derivative of a sample path is time averaged as

$$\hat{\lambda}(s) = \frac{1}{T}\int_0^T \frac{dN_t(s)}{dt}\,dt.$$

Q2. Any other way?

A2. X_n the interarrival process is i.i.d. exponential w/

$f_X(x) = \lambda \exp(-\lambda x)$

$E\{X\} = \frac{1}{\lambda}$..

Since $\frac{X_1 + X_2 + \cdots + X_N}{N} \to \frac{1}{\lambda}$. by LLN

$\therefore \hat{\lambda}_2 \triangleq \frac{N(T)}{X_1 + \cdots + X_{N(T)}}$

Q3 which estimate is better?

$\hat{\lambda}_1 = \frac{N(T)}{T}$ or $\hat{\lambda}_2 = \frac{N(T)}{X_1 + X_2 + \cdots + X_{N(T)}}$?

Time Average and Ergodicity

- Def. The time average or the mean value of a CT signal $x(t)$ is defined as the limit of an integral:

$$\mathbf{A}[x(t)] \triangleq \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)dt.$$

→ DC term extractor

  – Two-step understanding:
    1. We construct a function $\frac{1}{2T}\int_{-T}^{T} x(t)dt$ of T.
    2. We find the limit of this function when T goes to ∞. If it exists, it is called the time average of $x(t)$.

  – Lemma. The time average of a constant function $x(t) = c, \forall t$ is $\mathbf{A}[x(t)] = c$.

  – Def. The energy of a CT signal $x(t)$ is defined as

$y(t) = x(t-\tau)$

$$E_{xx} \triangleq \int_{-\infty}^{\infty} |x(t)|^2 dt.$$

(finite) Fourier-Plancherel $X(f)$ → $\int_{-\infty}^{\infty} |X(f)|^2 df$.

    * Def. If $0 < E_{xx} < \infty$, $x(t)$ is called an energy signal.
    * The energy is translation-invariant, i.e., $E_{xx} = \int_{-\infty}^{\infty} |x(t-\tau)|^2 dt, \forall \tau$.

  – Def. The power of a CT signal $x(t)$ is defined as

$$P_{xx} \triangleq \lim_{T\to\infty} \frac{1}{2T}\boxed{\int_{-T}^{T} |x(t)|^2 dt} = \mathbf{A}[|x(t)|^2].$$

    * Def. If $0 < P_{xx} < \infty$, $x(t)$ is called a power signal.

$x(t) = \infty, \forall t$

    * An energy signal has zero power. A power signal has infinite energy. There are signals with infinite energy and power.
    * The power is translation-invariant, i.e., $P_{xx} = \mathbf{A}[|x(t-\tau)|^2], \forall \tau$.

  – Def. The time-autocorrelation function of a CT signal $x(t)$ is defined as

$$\mathcal{R}_{xx}(\tau) \triangleq \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} x(t)x^*(t-\tau)dt = \mathbf{A}[x(t-\tau)x^*(t)]. = \mathbf{A}\left[\boxed{x(t)x^*(t-\tau)}\right]$$

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)x^*(t-\tau)\,dt$$

    * The time-autocorrelation function is translation-invariant, i.e., $\mathcal{R}_{xx}(\tau) = \mathbf{A}[x(t-\tau/2)x^*(t-\tau/2)] = \mathbf{A}[x(t)x^*(t-\tau)]$.

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t+\tfrac{\tau}{2})\, x^*(t-\tfrac{\tau}{2})\,dt$$