

Fourier transform applications: Probability and random processes

Dr. Edmund Lam

Department of Electrical and Electronic Engineering
The University of Hong Kong

ELEC8501: The Fourier Transform and Its Applications
(First Semester, 2023–24)

<https://www.eee.hku.hk/~elec8501>

Preamble

Although Fourier transform is a deterministic algebraic formula:

- ① as a mathematical tool, it helps to provide insight in probability calculations
- ② as a tool for analyzing linear systems, it supports the analysis when the signals are random (stochastic)

1 PMF and PDF (review)

2 Characteristic function

3 Linear systems

Random variables

Deterministic world:

- Numbers (arithmetic): $3 + 4 = 7$
- Variables (algebra): $x = 7$
- Functions of variables: $g(x) = x^2 \log_2 x$

Probabilistic world:

- Numbers: $\mathbb{P}(\text{head}) = \frac{1}{2}$
- Variables: $X = \text{sum of two dice rolls}$
- Functions of variables: $g(X) = X^2 \log_2 X$

We are interested in questions such as $\mathbb{P}(X = 7)$, the distribution of X , properties of such distribution (e.g., mean), $\mathbb{P}(g(X) = 192)$, the distribution of $g(X)$, etc.

Random variables

Random variable (r.v.) X :

- takes on numerical values
 - We may need to *assign* numerical values to each outcome
- convention is to use capital letter
- **discrete random variable** means the set of values it can take is either finite or countably infinite
- the probability of X being equal to each possible value in its range forms a **probability mass function (PMF)**

PMF = “ideal” histogram

(1)

- a function of X leads to another PMF

Notation

A r.v. X can take on different numerical values. If x is any real number within the range of this r.v., we write

$$p_X(x) = \mathbb{P}(X = x) \quad (2)$$

By definition, $\sum_x p_X(x) = 1$. Also, $p_X(x)$ must be nonnegative.

Notation

A r.v. X can take on different numerical values. If x is any real number within the range of this r.v., we write

$$p_X(x) = \mathbb{P}(X = x) \quad (2)$$

By definition, $\sum_x p_X(x) = 1$. Also, $p_X(x)$ must be nonnegative.

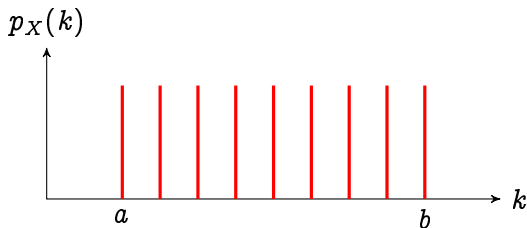
Example: Let X = number of heads in two tosses of a fair coin.

$$p_X(x) = \begin{cases} \frac{1}{4} & x = 0 \\ \frac{1}{2} & x = 1 \\ \frac{1}{4} & x = 2 \\ 0 & \text{otherwise.} \end{cases}$$

(1) Discrete uniform random variable

Discrete uniform random variable over a range $[a, b]$ has PMF

$$p_X(k) = \begin{cases} \frac{1}{b-a+1} & k = a, a+1, \dots, b \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$



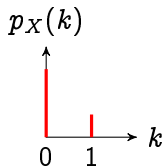
(2) Bernoulli random variable

Bernoulli random variable is useful for various binary state events: true/false, yes/no, 0/1, present/absent, etc. It has two possible values

$$X = \begin{cases} 1 & \text{head} \\ 0 & \text{tail} \end{cases}$$

Bernoulli PMF (coin toss with $\mathbb{P}(\text{head}) = p$) is

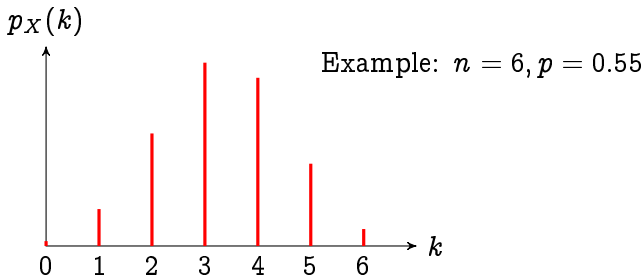
$$p_X(k) = \begin{cases} p & k = 1 \\ 1 - p & k = 0 \end{cases} \quad (4)$$



(3) Binomial random variable

Binomial random variable is the sum of n Bernoulli trials with probability p , with PMF

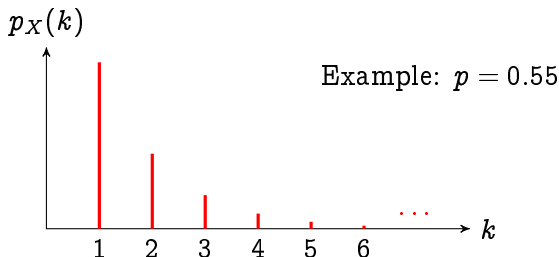
$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n \quad (5)$$



(4) Geometric random variable

Geometric random variable is the number of Bernoulli trials with probability p ($\mathbb{P}(\text{head}) = p$) before head turns up, with PMF

$$p_X(k) = (1 - p)^{k-1}p \quad k = 1, 2, \dots \quad (6)$$



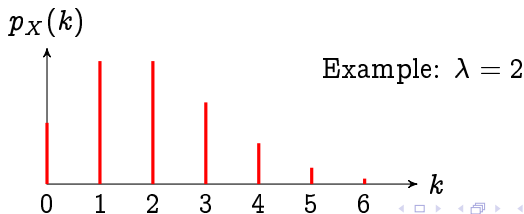
(5) Poisson random variable

Poisson random variable with parameter λ has PMF

$$p_X(k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad k = 0, 1, \dots \quad (7)$$

It is a good approximation of binomial r.v. with $\lambda = np$, n is very large, and p very small, i.e.

$$\frac{e^{-\lambda} \lambda^k}{k!} \approx \binom{n}{k} p^k (1-p)^{n-k} \quad k \ll n$$



Poisson random variable

Proof: For a binomial r.v. with parameters (n, p) , and let $\lambda = np$,

$$\begin{aligned} p_X(k) &= \frac{n!}{(n-k)! k!} p^k (1-p)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot \frac{(np)^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

We fix k but let $n \rightarrow \infty$, so for “small” $j = 1, \dots, k$,

$$\frac{n-k+j}{n} \rightarrow 1, \quad \left(1 - \frac{\lambda}{n}\right)^{-k} \rightarrow 1, \quad \left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda}$$

Therefore,

$$p_X(k) \rightarrow \frac{e^{-\lambda} \lambda^k}{k!}$$

Expectation

The **expectation** or **expected value** or **mean** of a r.v. X is

$$\mathbb{E}(X) = \sum_x x p_X(x) \quad (8)$$

It is a very useful number as a “representative” value of X , and is also viewed as the “center of gravity.”

Moment, variance, standard deviation

The n th moment of r.v. X is given by $\mathbb{E}(X^n)$

$$\text{first moment} = \mathbb{E}(X) = \text{mean} \quad (9)$$

$$\text{second moment} = \mathbb{E}(X^2) \quad \{\text{variance}\} \quad (10)$$

$$\text{third moment} = \mathbb{E}(X^3) \quad \{\text{skewness}\} \quad (11)$$

$$\text{fourth moment} = \mathbb{E}(X^4) \quad \{\text{kurtosis}\} \quad (12)$$

The variance is not exactly the second moment, but

$$\text{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (13)$$

Discrete random variable

It is useful to know the mean and variance of some common distributions. A summary:

	mean: $\mathbb{E}(X)$	variance: $\text{var}(X)$
uniform: $X \sim U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+1)}{12}$
Bernoulli: $X \sim \text{Ber}(p)$	p	$p(1-p)$
Binomial: $X \sim \text{Bin}(n, p)$	np	$np(1-p)$
Geometric: $X \sim \text{Geo}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson: $X \sim \text{Poisson}(\lambda)$	λ	λ

Continuous random variables

Continuous random variable X :

- has a nonnegative function

$$f_X(x) \geq 0 \quad (14)$$

for all x , called **probability density function (PDF)**

- Integrating (“continuous summing”) this PDF gives us the probability of an event:

$$\mathbb{P}(X \in B) = \int_B f_X(x) dx \quad (15)$$

for every subset B of the entire space

Expectation

The **expectation** or **expected value** or **mean** of a continuous r.v. X is

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx \quad (16)$$

Similar to the discrete case, it is a very useful number as a “representative” value of X , and is also viewed as the “center of gravity.”

Note: There can be problematic cases with integrals: even if $f_X(x)$ is symmetric about 0, the expected value is in fact not defined, e.g. $f_X(x) = 1/\pi(1 + x^2)$

Expectation and variance

The formulas for continuous r.v. are similar to those for discrete:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx \quad (17)$$

$$\mathbb{E}(X^3) = \int_{-\infty}^{\infty} x^3 f_X(x) dx \quad (18)$$

$$\mathbb{E}(X^4) = \int_{-\infty}^{\infty} x^4 f_X(x) dx \quad (19)$$

$$\text{var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (20)$$

Discrete and continuous

We can make use of the delta function to unify both discrete and continuous random variables, and just say everything is continuous.

Example: A binomial PMF

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n \quad (21)$$

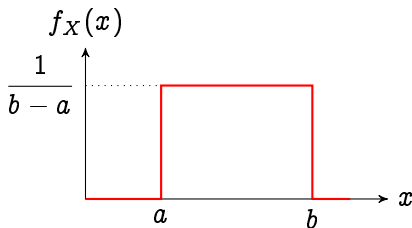
is equivalent to the PDF

$$f_X(x) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \delta(x - k) \quad (22)$$

(1) Uniform random variable

Continuous uniform random variable over a range $[a, b]$ has PDF

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases} \quad (23)$$



(1) Uniform random variable

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{2} x^2 \Big|_a^b = \frac{a+b}{2}$$

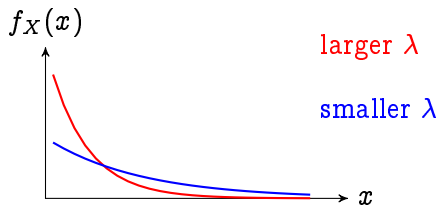
$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{1}{3} x^3 \Big|_a^b = \frac{a^2+ab+b^2}{3}$$

$$\text{var}(X) = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

(2) Exponential random variable

Exponential random variable with parameter $\lambda > 0$ has PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (24)$$



(2) Exponential random variable

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{E}(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = (-x e^{-\lambda x}) \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = 0 - \frac{e^{-\lambda x}}{\lambda} \Big|_0^{\infty} = \frac{1}{\lambda}$$

$$\mathbb{E}(X^2) = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = (-x^2 e^{-\lambda x}) \Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} dx = \frac{2}{\lambda} \mathbb{E}(X) = \frac{2}{\lambda^2}$$

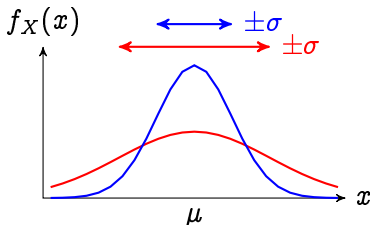
$$\text{var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

(3) Normal random variable

Gaussian or **normal** r.v. with parameters $(\mu, \sigma > 0)$ has PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad (25)$$

This is probably the most common distribution of all (with good scientific reason)!



(3) Normal random variable

Normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ has **mean** μ and **variance** σ^2 .

(3) Normal random variable

Normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ has **mean** μ and **variance** σ^2 .

Proof: For the mean, we will appeal to the fact that it is symmetric around μ and that the integral is defined. For variance, making use of a change of variable $y = (x - \mu)/\sigma$,

$$\begin{aligned}
 \text{var}(X) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-(x-\mu)^2/2\sigma^2} dx \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \left[\left(-ye^{-y^2/2} \right) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-y^2/2} dy \right] \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) = \sigma^2
 \end{aligned}$$

Continuous random variable

A summary:

	mean: $\mathbb{E}(X)$	variance: $\text{var}(X)$
uniform: $X \sim U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
exponential: $X \sim \text{Exp}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
normal: $X \sim \mathcal{N}(\mu, \sigma^2)$	μ	σ^2

- 1 PMF and PDF (review)
- 2 Characteristic function**
- 3 Linear systems

Definition

For a random variable X , define the **characteristic function** $\phi_X(s)$ (or simply $\phi(s)$) where

$$\phi_X(s) = \mathbb{E}(e^{jsX}) \quad (26)$$

For continuous random variable, $\phi(s) = \int_{-\infty}^{\infty} e^{jsx} f_X(x) dx$.

A key observation: **This is essentially a Fourier transform!** (except that the convention here is not to have the -2π in the exponential)

Computing moments

In exactly the same way that the Fourier transform can be used to find the moments of a function (see Lecture 4):

$$\begin{aligned}\int_{-\infty}^{\infty} g(t) dt &= G(0) \\ \int_{-\infty}^{\infty} t g(t) dt &= \frac{j}{2\pi} G'(0) \\ \int_{-\infty}^{\infty} t^2 g(t) dt &= -\frac{1}{4\pi^2} G''(0)\end{aligned}$$

we can also compute the moments of the PDF from the characteristic function:

$$\mathbb{E}(X) = -j \frac{d}{ds} \phi(s) \Big|_{s=0} \quad (27)$$

$$\mathbb{E}(X^2) = -\frac{d^2}{ds^2} \phi(s) \Big|_{s=0} \quad (28)$$

Example: Exponential random variable

Exponential random variable with parameter $\lambda > 0$ has PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

$$\phi(s) = \int_0^{\infty} \lambda e^{-\lambda x} e^{jsx} dx = \lambda \int_0^{\infty} e^{(js-\lambda)x} dx = \lambda \frac{e^{(js-\lambda)x}}{js-\lambda} \Big|_0^{\infty} = \frac{\lambda}{\lambda - js}$$

$$\mathbb{E}(X) = -j \frac{d}{ds} \frac{\lambda}{\lambda - js} \Big|_{s=0} = (-j)\lambda(-1)(\lambda - js)^{-2}(-j) \Big|_{s=0} = \frac{1}{\lambda}$$

$$\mathbb{E}(X^2) = -\frac{d^2}{ds^2} \lambda(\lambda - js)^{-1} \Big|_{s=0} = -j\lambda(-2)(\lambda - js)^{-3}(-j) \Big|_{s=0} = \frac{2}{\lambda^2}$$

$$\text{var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Example: Normal variable

For a Gaussian random variable, we can calculate the characteristic function (with $y = x - \mu$) as:

$$\begin{aligned}
 \phi(s) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{jsx} dx \\
 &= e^{j\mu s} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} e^{jys} dy \\
 &= e^{j\mu s} \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-j\sigma^2 s)^2}{2\sigma^2}} dy \right] e^{-\sigma^2 s^2/2} \\
 &= e^{j\mu s - \sigma^2 s^2/2}
 \end{aligned}$$

Example: Normal variable

Computing the derivatives:

$$\phi(s) = e^{j\mu s - \sigma^2 s^2/2}$$

$$\frac{d}{ds}\phi(s) = (j\mu - \sigma^2 s)e^{j\mu s - \sigma^2 s^2/2}$$

$$\frac{d^2}{ds^2}\phi(s) = (j\mu - \sigma^2 s)^2 e^{j\mu s - \sigma^2 s^2/2} + (-\sigma^2)e^{j\mu s - \sigma^2 s^2/2}$$

Therefore,

$$\mathbb{E}(X) = -j(j\mu - \sigma^2 s)e^{j\mu s - \sigma^2 s^2/2} \Big|_{s=0} = \mu$$

$$\mathbb{E}(X^2) = - \left[(j\mu - \sigma^2 s)^2 e^{j\mu s - \sigma^2 s^2/2} + (-\sigma^2)e^{j\mu s - \sigma^2 s^2/2} \right] \Big|_{s=0} = \mu^2 + \sigma^2$$

$$\text{var}(X) = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

Example: Binomial random variable

We can also deal with a discrete random variable in a similar way.
Let's try the binomial random variable:

$$\begin{aligned}
 f_X(x) &= \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \delta(x-k) \\
 \phi(s) &= \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} e^{jsk} \\
 &= \sum_{k=0}^n \binom{n}{k} (pe^{js})^k (1-p)^{n-k} \\
 &= [pe^{js} + (1-p)]^n
 \end{aligned}$$

Example: Binomial random variable

Computing the derivatives:

$$\phi(s) = [pe^{js} + (1-p)]^n$$

$$\phi'(s) = n [pe^{js} + (1-p)]^{n-1} \cdot jpe^{js}$$

$$\phi''(s) = n(n-1) [pe^{js} + (1-p)]^{n-2} j^2 p^2 e^{j2s} + n [pe^{js} + (1-p)]^{n-1} j^2 pe^{js}$$

Therefore,

$$\mathbb{E}(X) = -j \cdot njp = np$$

$$\mathbb{E}(X^2) = - \left[n(n-1)(-p^2) + n(-1)(p) \right] = np(1-p + np)$$

$$\text{var}(X) = np - np^2 + n^2 p^2 - n^2 p^2 = np(1-p)$$

Characteristic functions

A summary of what we have calculated:

	characteristic function $\phi(s)$
exponential: $X \sim \text{Exp}(\lambda)$	$\phi(s) = \frac{\lambda}{\lambda - js}$
normal: $X \sim \mathcal{N}(\mu, \sigma^2)$	$\phi(s) = e^{j\mu s - \sigma^2 s^2 / 2}$
Binomial: $X \sim \text{Bin}(n, p)$	$\phi(s) = [pe^{js} + (1 - p)]^n$

Properties of characteristic function

Some properties of the characteristic function:

- $\phi(0) = 1$ (because PDF integrates to one)
- $\phi(s) = \text{hermitian}$ (remember from Lecture 1: Hermitian means the real part is even, the imaginary part is odd. Fourier transform of a real function is hermitian.) For example:

$$\left(\frac{\lambda}{\lambda - js} \right) \left(\frac{\lambda + js}{\lambda + js} \right) = \underbrace{\frac{\lambda^2}{\lambda^2 + s^2}}_{\text{even}} + j \underbrace{\frac{\lambda s}{\lambda^2 + s^2}}_{\text{odd}}$$

- $|\phi(s)| \leq 1$ because

$$\left| \int_{-\infty}^{\infty} e^{jsx} f_X(x) dx \right| \leq \int_{-\infty}^{\infty} |e^{jsx} f_X(x)| dx = 1$$

Use of characteristic function

Calculating the characteristic function seems like a pain. But it is very useful in dealing with multiple random variables.

For example, consider a sum of independent random variables:

$$Z = X + Y:$$

$$\text{Discrete:} \quad p_Z(z) = \sum_x p_X(x) p_Y(z - x) \quad (29)$$

$$\text{Continuous:} \quad f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx \quad (30)$$

The PMF (PDF) of Z is the **convolution** of the PMF (PDF) of X and Y . Consequently, a key result:

$$\phi_Z(s) = \phi_X(s) \phi_Y(s) \quad (31)$$

Example: Normal random variable

Example: Sum of Normal random variables.

Let $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$. Then,

$$\begin{aligned}\phi_Z(s) &= \phi_X(s)\phi_Y(s) = e^{j\mu_x s - \sigma_x^2 s^2/2} \cdot e^{j\mu_y s - \sigma_y^2 s^2/2} \\ &= e^{j(\mu_x + \mu_y)s - (\sigma_x^2 + \sigma_y^2)s^2/2}\end{aligned}$$

which implies that

$$Z \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

i.e., the resulting distribution is also Gaussian, and that the mean is the sum of the respective means, and the variance is the sum of the respective variances.

Example: Poisson random variable

Example: Sum of Poisson random variables.

Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ be two Poisson random variables. What is $Z = X + Y$?

(Remember that for a Poisson random variable, $p_X(k) = \frac{e^{-\lambda} \lambda^k}{k!}$ for $k = 0, 1, \dots$)

Example: Poisson random variable

We proceed as follows:

$$\begin{aligned}
 f_X(x) &= \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} \delta(x - k) \\
 \phi_X(s) &= \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} e^{jsk} \\
 &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{js} \lambda)^k}{k!} = e^{-\lambda} e^{e^{js} \lambda} = e^{\lambda(e^{js} - 1)}
 \end{aligned}$$

Therefore, $\phi_Z(s) = \phi_X(s)\phi_Y(s) = e^{\lambda(e^{js} - 1)} e^{\mu(e^{js} - 1)} = e^{(\lambda + \mu)(e^{js} - 1)}$, which means $Z \sim \text{Poisson}(\lambda + \mu)$.

Random number of independent random variables

The characteristic function can help us to compute more complicated probability questions, such as finding the distribution involving a random number of independent random variables.

Consider:

$$Y = X_1 + \dots + X_N$$

where

- N is a random variable that takes nonnegative integer values.
- $X_1, X_2, \dots$ are identically distributed random variables.
- $N, X_1, X_2, \dots$ are independent, meaning that any finite subcollection of these random variables are independent.
- $\mathbb{E}(X)$ and $\text{var}(X)$ are the common mean and variance of each X_i .

What is the distribution of Y ?

Example

You visit a number of bookstores in search of a particular textbook on probability. Any given bookstore carries the book with probability p , independent of the others. In a typical bookstore, the amount of time you spend is exponentially distributed with parameter λ , and independent of the time you spend in other bookstores. You will keep on visiting bookstores until you find the book (because the lectures are too boring, you'd rather learn from a book). What are the mean, variance, and PDF of the total time spent in search of the book?

It is a sum of a *geometric number* of independent exponential random variables.

Random number of independent random variables

We aim to find $\phi_Y(s)$. We need to make use of this law of iterated expectations:

$$\mathbb{E}\left(e^{jsY}\right) = \mathbb{E}\left(\mathbb{E}\left(e^{jsY} \mid N = n\right)\right) \quad (32)$$

Random number of independent random variables

$$\begin{aligned}
 \mathbb{E}\left(e^{jsY} \mid N = n\right) &= \mathbb{E}\left(e^{js(X_1 + \dots + X_N)} \mid N = n\right) \\
 &= \mathbb{E}\left(e^{jsX_1} \dots e^{jsX_n} \mid N = n\right) \\
 &= \mathbb{E}\left(e^{jsX_1}\right) \dots \mathbb{E}\left(e^{jsX_n}\right) \\
 &= \left(\phi_X(s)\right)^n \\
 &= e^{n \log \phi_X(s)}
 \end{aligned}$$

Therefore,

$$\mathbb{E}\left(\mathbb{E}\left(e^{jsY} \mid N = n\right)\right) = \mathbb{E}\left(e^{N \log \phi_X(s)}\right)$$

Random number of independent random variables

We have:

$$\phi_Y(s) = \mathbb{E}\left(e^{N \log \phi_X(s)}\right)$$

$$\phi_N(s) = \mathbb{E}\left(e^{jsN}\right)$$

Therefore,

$$\phi_Y(s) = \phi_N(-j \log \phi_X(s)) \quad (33)$$

This equation allows us to relate the PDF of Y with those of X and N .

Random number of independent random variables

Back to our example:

X is exponential: $X \sim \text{Exp}(\lambda)$

N is geometric: $N \sim \text{Geo}(p)$

We first calculate the characteristic function of geometric random variable:

$$\begin{aligned}
 \phi_N(s) &= \sum_{n=1}^{\infty} (1-p)^{n-1} p \cdot e^{jsn} \\
 &= pe^{js} \sum_{n=1}^{\infty} \left[(1-p)e^{js} \right]^{n-1} \\
 &= \frac{pe^{js}}{1 - (1-p)e^{js}}
 \end{aligned}$$

Random number of independent random variables

Therefore,

$$\begin{aligned}
 \phi_Y(s) &= \frac{pe^{\log \phi_X(s)}}{1 - (1-p)e^{\log \phi_X(s)}} \\
 &= \frac{p^{\frac{\lambda}{\lambda-j s}}}{1 - (1-p)^{\frac{\lambda}{\lambda-j s}}} \\
 &= \frac{p\lambda}{p\lambda - js}
 \end{aligned}$$

Consequently, we can conclude that $Y \sim \text{Exp}(p\lambda)$.

Other uses

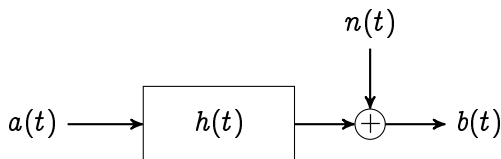
Another common use of the characteristic function is in proving the central limit theorem. But it requires more understanding of “convergence”, and therefore we will skip here. You can read, for example, the discussion in

<https://sas.uwaterloo.ca/~dlmcleis/s901/chapt6.pdf>

- 1 PMF and PDF (review)
- 2 Characteristic function
- 3 Linear systems**

Block diagram

We often deal with linear time-invariant (LTI) systems represented like this:



$$b(t) = h(t) * a(t) + n(t) \quad (34)$$

$$B(s) = H(s)A(s) + N(s) \quad (35)$$

Fourier transform inherently plays an important role.

Noise

We often call $n(t)$ “additive white Gaussian noise (AWGN). Why do we call it “white”?

$$\mathcal{F}\{n(t)\} \neq \text{flat}$$

For any *specific* realization, $n(t)$ does not have a flat spectrum. But *on average*, it is. (i.e., every frequency component is equally likely)

i.i.d. Gaussian

$n(t)$ is independent, identically distributed (i.i.d.) Gaussian means that each value is generated with the same Gaussian (mean and variance) without reference to each other.

Key fact: Its autocorrelation is a delta function!

$$n^*(t) \star n(t) = \sigma^2 \delta(t) \quad (36)$$

The factor σ^2 denotes the severity of the noise.

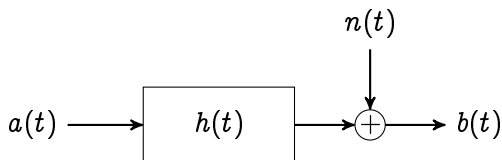
Power spectral density

The power spectral density (Φ) is the Fourier transform of the autocorrelation.

Key fact: The power spectral density of i.i.d. Gaussian noise is a constant!

$$\Phi_n(s) = \sigma^2 \quad (37)$$

LTI system with random input



If $a(t)$ is a random (stochastic) signal, knowing a particular $b(t)$ is not too helpful. Instead, we should work with their power spectral densities.

LTI system with random input

The autocorrelation function of a random process $X(t)$ is

$$R_X(t_1, t_2) = \mathbb{E}(X(t_1)X(t_2)) \quad (38)$$

When $X(t)$ is “wide-sense stationary”, we have

$$R_X(\tau) = \mathbb{E}(X(t)X(t + \tau)) \quad (39)$$

Let $Y(t) = X(t) * h(t)$. Then,

$$\begin{aligned} R_Y(\tau) &= \mathbb{E}(Y(t)Y(t + \tau)) \\ &= \mathbb{E}\left(\int_{-\infty}^{\infty} h(q)X(t - q)dq \int_{-\infty}^{\infty} h(r)X(t + \tau - r)dr\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(q)h(r)\mathbb{E}(X(t - q)X(t + \tau - r))dqdr \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(q)h(r)R_X(\tau + q - r)dqdr \end{aligned}$$

LTI system with random input

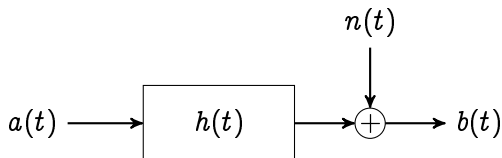
Now, let's go to the Fourier domain:

$$\begin{aligned}\Phi_y(s) &= \int_{-\infty}^{\infty} R_Y(\tau) e^{-j2\pi s\tau} d\tau \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(q)h(r)R_X(\tau + q - r)e^{-j2\pi s\tau} dqdrd\tau\end{aligned}$$

We let $u = \tau + q - r$, and therefore

$$\begin{aligned}\Phi_y(s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(q)h(r)R_X(u)e^{-j2\pi s(u-q+r)} dqdrdu \\ &= \int_{-\infty}^{\infty} h(q)e^{j2\pi sq} dq \int_{-\infty}^{\infty} h(r)e^{-j2\pi sr} dr \int_{-\infty}^{\infty} R_X(u)e^{-j2\pi su} du \\ &= H^*(s)H(s)\Phi_x(s) \\ &= |H(s)|^2\Phi_x(s)\end{aligned}$$

LTI system with random input

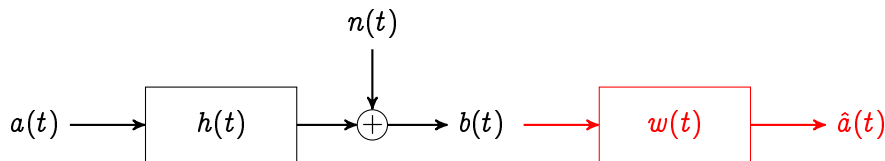


Using the results above:

$$\Phi_b(s) = |H(s)|^2 \Phi_a(s) + \Phi_n(s) \quad (40)$$

Wiener filter

Suppose we observe $b(t)$, and want to design a filter $w(t)$ such that the output is as close to $a(t)$ as possible:



What do we mean by $\hat{a}(t) \approx a(t)$?

Key result (Wiener filter):

$$W(s) = \frac{H^*(s)}{|H(s)|^2 + \frac{\Phi_n(s)}{\Phi_a(s)}} \quad (41)$$