EM 509: Stochastic Processes

Class Notes

Transform Methods in Stochastic Theory

(Lecture 5)

Dr. R. Palamakumbura
Transform Methods in Stochastic Theory

• We saw in a previous lecture, to provide different statistical properties of the process and different amounts of information concerning the process, we need to calculate various types of characterizations.

• Some types of such characterizations are mean, variance, moments, the nth-order distributions, autocorrelation function, and spectral density etc.

• In stochastic theory we are dealing with an infinite family of random variables.

• Therefore to perform such operations in standard ways will sometimes be possible but, in many situations it will be very tedious or sometimes impossible.
For example:

- Suppose we have independent random variables $X_1, X_2, X_n, \ldots$ each has a Poisson distribution with parameter $\lambda_i$, $i = 1, 2, \ldots$.

- Suppose we need to find the distribution of the sum $X_1 + \cdots + X_n$. In this case we can use mathematical induction to show that the sum has Poisson distribution with parameter $\lambda_1 + \cdots + \lambda_n$.

- In the above we found a ‘natural’ way to manipulate the algebra so that we could recognize the answer.

- What would happen if we considered other sums of random variables? Will it be possible to come up with a mathematical tool as above?

- It is nice if we have a procedure that will work in general.
Transform Methods in Stochastic Theory

• Such an approach exists and it is called the theory of **generating functions or transform methods**.

• Some important transform methods are:

  1. Moment Generating Function

  2. Laplace Transform

  3. Characteristic Function

• Depending on your field of study, you can use the type that will be more effective than the others.

• Here we will study about these generating functions and their relationships.
Transform Methods in Stochastic Theory

- **Moment generating function (MGF)**

  - For $t \in \mathbb{R}$, the moment generating function of a random variable $X$ is defined as
    \[
    M_X(t) = E(e^{tX}) = \begin{cases} 
    \sum_{k=-\infty}^{\infty} e^{tx} P(X_k) & \text{discrete random variable} \\
    \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx & \text{continuous random variable}
    \end{cases}
    \]

    where $f_X, P$ are the probability density or mass function respectively.

  - The above is an infinite series or integral. Therefore we need to see whether it exists (finite).

  - For example it is clear from the definition that for the integral to exist, the right tail of the density has to go to zero faster than $e^{-x}$.

  - This is not the case for fat-tailed distributions.
Laplace Transform

Recall the Laplace transform of a function is defined as

\[ F(t) = \mathcal{L}(f(x)) = \int_0^\infty e^{-tx}f(x)\,dx. \]

Thus if we flip the sign on \( t \) in the definition of \( M_X(t) \), we have the two-sided Laplace transform of \( f_X \).

That is, the moment generating function of \( X \) at \( t \) is the two-sided Laplace transform of \( f_X \) at \(-t\).

Thus if the density function is zero for negative values, then the two-sided Laplace transform reduces to the more common (one-sided) Laplace transform.
Transform Methods in Stochastic Theory

- **Characteristic Function**

  - The characteristic function of a random variable is a variation on the moment generating function.

  - Rather than using the expected value of $tX$, it uses the expected value of $itX$.

  - This means the characteristic function of a random variable is the Fourier transform of its density/mass function.

  - Characteristic functions are easier to work with than moment generating functions.

  - Existence is not a problem for the characteristic function because the Fourier transform exists for any density/mass function.
Characteristic Function

- For $X$ a random variable and $t \in \mathbb{R}$, the characteristic function is defined as

$$
\psi_X(t) = E(e^{itX}) = \begin{cases} 
\sum_{k=-\infty}^{\infty} e^{itx_k}P(X_k) & \text{discrete random variable} \\
\int_{-\infty}^{\infty} e^{itx}f_X(x)dx & \text{continuous random variable}.
\end{cases}
$$

- Thus the characteristic function is the most general form of the transforms.

- Therefore we will study only about the characteristic function and all the results are appropriately applicable for other transforms.
Transform Methods in Stochastic Theory

- **Characteristic Function**

- Using Euler formula, \( E(e^{itx}) = E(\cos tx) + iE(\sin tx) \).

- The above gives the expectation of the complex random variable \( e^{itX} \) in terms of expectations of two real random variables.

- Since \( |e^{itX}| = 1 \), \( E(|e^{itx}|) = E(|e^{itx}|^2) = 1 \).

- This is a transformation that transforms probability density function or probability mass function to a complex function.
Transform Methods in Stochastic Theory

- **Characteristic Function**

  - Uniqueness: If two random variables $X_1$ and $X_2$ have the same characteristic functions, then they have the same distribution functions.

    \[ \psi_{X_1}(t) = \psi_{X_2}(t) \quad \forall t \in \mathbb{R}, \quad \text{then} \quad F_{X_1} = F_{X_2} \quad \forall x \in \mathbb{R}. \]

    This is written as $X_1 \overset{d}{=} X_2$.

  - There are several additional properties that follow immediately from the definition of the characteristic function.
Transform Methods in Stochastic Theory

- Characteristic Function

- Properties:

  1. Characteristic function $\psi_X(t)$ exists for any random variable.

  2. At $t = 0, \psi_X(0) = 1$ and $|\psi_X(t)| \leq 1$.

  3. Characteristic function $\psi_X(t)$ is uniformly continuous.

  4. Characteristic function of $a + bX$ for $a, b$ constants is

     $\psi_{a+bX} = e^{iat}\psi_X(bt)$. 

  5. Characteristic function of $-X$ is the complex conjugate $\bar{\psi}_X(t)$. 
Characteristic Function

Properties:

6. Characteristic function \( \psi_X(t) \) is real valued iff \( X \overset{d}{=} -X \).
   i.e the distribution is symmetric about zero.

7. For any complex numbers, \( z_l; \ l = 1,2,\ldots, n \) and for any real \( t_l; \ l = 1,2,\ldots, n \) we have

\[
\sum_{l=1}^{n} \sum_{k=1}^{n} z_l \overline{z_k} \psi_X(t_l - t_k) \geq 0.
\]

i.e. the characteristic function is positive semidefinite.
Transform Methods in Stochastic Theory

- **Characteristic Function**

Examples:

1. Standard Normal Distribution, $X \in N(0,1)$.

   \[ \psi_X(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{x^2/2} dx. \]

   Differentiating w.r.t the parameter $t$ and allowing to move differentiation inside the integral sign, we get

   \[ \psi_X(t)' = \frac{d\psi_X(t)}{dt} = \int_{-\infty}^{\infty} \frac{d}{dt} e^{itx} \frac{1}{\sqrt{2\pi}} e^{x^2/2} dx = \int_{-\infty}^{\infty} ixe^{itx} \frac{1}{\sqrt{2\pi}} e^{x^2/2} dx = \int_{-\infty}^{\infty} -ie^{itx} \frac{1}{\sqrt{2\pi}} (-xe^{x^2/2}) dx. \]

   By integration by parts we get

   \[ \psi_X(t)' = -ie^{itx} \frac{1}{\sqrt{2\pi}} e^{x^2/2} \bigg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left( -tt' e^{itx} \frac{1}{\sqrt{2\pi}} e^{x^2/2} \right) dx = 0 - t\psi_X(t). \]
Characteristic Function

This results in the first order linear ordinary differential equation

\[ \psi_X'(t) + t\psi_X(t) = 0. \]

Using the integrating factor we get \( \psi_X(t) = Ce^{-t^2/2}. \)

Since \( \psi_X(0) = 1 \) we have \( \psi_X(t) = e^{-t^2/2}. \)

Thus we have obtained \( X \in N(0,1) \Leftrightarrow \psi_X(t) = e^{-t^2/2}. \)

Note: Since this is real valued, therefore by the properties of the characteristic function we get \( -X \in N(0,1). \)
Transform Methods in Stochastic Theory

• Characteristic Function

2. Poisson Distribution, $X \in Po(\lambda), \lambda > 0$.

$$
\psi_X(t) = \sum_{k=0}^{\infty} e^{ikt} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{it} \lambda)^k}{k!} = e^{-\lambda} e^{e^{it} \lambda} = e^{(e^{it}-1)\lambda}.
$$

Thus we have obtained

$$
X \in Po(\lambda) \iff \psi_X(t) = e^{(e^{it}-1)\lambda}.
$$
Characteristic Functions and Moments of Random Variables

If the random variable $X$ has $E(|X|^K) < \infty$, then

$$\frac{d^k}{dt^k} \psi_X(t) \bigg|_{t=0} = \frac{d^k}{dt^k} \psi_X(0) = i^k E(X^K).$$

This can be proved changing the order of differentiation and expectation,

$$\frac{d}{dt} \psi_X(t) = E\left[ \frac{d}{dt} e^{itX} \right] = E[iX e^{itX}]$$

$$\frac{d}{dt} \psi_X(0) = iE[iX e^{itX}].$$
Transform Methods in Stochastic Theory

- **Characteristic Functions and Moments of Random Variables**

Example: Mean and Variance of the Poisson Distribution

\[ \psi_X(t) = e^{(e^{it} - 1)\lambda} \]

\[ \frac{d\psi_X(t)}{dt} = e^{(e^{it} - 1)\lambda}i\lambda e^{it} \]

\[ \frac{d^2\psi_X(t)}{dt^2} = e^{(e^{it} - 1)\lambda}i^2\lambda^2 e^{2it} + e^{(e^{it} - 1)\lambda}i\lambda e^{it}. \]

Thus

\[ E(X) = \frac{1}{i} \frac{d}{dt} \psi_X(0) = \lambda \]

\[ Var(X) = E(X^2) - (E(X))^2 = \frac{1}{i^2} \frac{d^2}{dt^2} \psi_X(0) - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \]
Transform Methods in Stochastic Theory

**Characteristic Functions of Sums of Independent Random Variables**

- Suppose $X_1, \ldots, X_n$ are independent random variables with respective characteristic functions $\psi_{X_k}(t), k = 1, 2, \ldots, n$.

- Then the characteristic function of their sum $S_n = \sum_{k=1}^{n} X_k$ is given by $\psi_{S_n}(t) = \psi_{X_1}(t) \cdots \psi_{X_n}(t)$.

- Thus if $X_1, \ldots, X_n$ iid random variables with characteristic function $\psi_X(t)$ then

$$\psi_{S_n}(t) = (\psi_X(t))^n.$$
Central Limit Theorem

- Suppose \( X_1, X_2, \ldots \) is an infinite sequence of iid random variables with

\[
E(X_k) = \mu, \quad \text{Var}(X_k) = \sigma^2; \quad k = 1, \ldots.
\]

- Standardize each random variable by subtracting the common mean and then dividing the difference by the common standard deviation,

\[
Y_k = \frac{X_k - \mu}{\sigma}.
\]

- Then \( Y_k \)'s are iid with \( E(Y_k) = 0, \text{Var}(Y_k) = 1. \)

- Now define \( W_n \) by adding the first \( n \) of the \( Y_k \)'s and scale the sum by the factor \( \frac{1}{\sqrt{n}} \) so that,

\[
W_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k = \sum_{k=1}^{n} \frac{Y_k}{\sqrt{n}}.
\]
Central Limit Theorem

Note: A useful result:
If \( E(|X|^n) < \infty \), for some \( n \) then
\[
\psi_X(t) = 1 + \sum_{k=1}^{n} E(X^k) \frac{(it)^k}{k!} + o(|t|^n).
\]

Next we will compute the characteristic function of \( W_n \).

\[
\psi_{W_n}(t) = (\psi_{Y/\sqrt{n}}(t))^n, \quad \vdash Y_k's \ iid
\]
\[
= (\psi_Y(t/\sqrt{n}))^n, \quad \vdash \psi_{Y/\sqrt{n}}(t) = \psi_Y(t/\sqrt{n})
\]

Now expanding \( \psi_Y(t/\sqrt{n}) \) as given in the above note along
with \( E(Y_k) = 0, Var(Y_k) = 1 \) we get,
\[
\psi_Y(t/\sqrt{n}) = 1 - \frac{t^2}{2n} + o(t^2/n).
\]
Central Limit Theorem

• Thus the characteristic function of $W_n$ is given by

$$\psi_{W_n}(t) = (1 - \frac{t^2}{2n} + o(t^2/n))^n.$$  

• Now we will see what happens when $n \to \infty$. Taking the limit,

$$\lim_{n \to \infty} \psi_{W_n}(t) = e^{-t^2/2}.$$  

• Thus we observe that the characteristic function of $W_n$ converges for all $t$ to the characteristic function of $N(0,1)$. Thus $W_n \in N(0,1)$.

• This is central limit theorem.
Exercises

1. Prove the property, if the characteristic function of $X$ is $\psi_X(t)$ then the characteristic function of $a + bX$, $a, b$ constants is $\psi_{a+bX} = e^{iat}\psi_X(bt)$.

2. In the notes we obtained the characteristic function of a random variable $Z \in N(0,1)$ . Using this and the property given in question 1 above, show that the characteristic function of $X \in N(\mu, \sigma^2)$ is $\psi_X(t) = e^{i\mu t-\sigma^2 t^2/2}$.

3. If $X$ has Bernoulli distribution with probability of success $p$ then show that the characteristic function of $X$ is given by $\psi_X(t) = (1 - p) + e^{itp}$.
Exercises

4. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables and each $X_k \in N(\mu_k, \sigma_k^2)$; $k = 1, 2, \ldots, n$. Then show that for any constants $a_1, a_2, \ldots, a_n$ by using characteristic functions

$$S_n = \sum_{k=1}^{n} a_k X_k \in N \left( \sum_{k=1}^{n} a_k \mu_k, \sum_{k=1}^{n} a_k^2 \sigma_k^2 \right).$$

Deduce that if $X_1, X_2, \ldots, X_n$ are iid and $N(\mu, \sigma^2)$ then

$$\bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k \in N \left( \mu, \frac{\sigma^2}{n} \right).$$