AM216 Stochastic Differential Equations

Lecture 03

List of topics in this lecture

- Characteristic function (CF) of a RV, relation with Fourier transform
- Properties of CF, CF of sum of two independent RVs, CF of a normal RV
- Sum of independent normal RVs is a normal RV.
- Monty Hall's game, incomplete description of a game
- The Wiener process

Review of probability theory (Continued)

Short notations:

RV = random variable

PDF = probability density function

CDF = cumulative distribution function

FT = Fourier transform

CF = characteristic function

We now develop tools to show that

$\boxed{\text{Sum of independent normal RVs is a normal RV}}$

Characteristic function (CF) of a random variable

Random variable: $X(\omega)$

PDF: $\rho_X(x)$

Characteristic function is defined as

$$\phi_X(\xi) \equiv E\Big(\exp(i\xi X)\Big) = \int_{-\infty}^{+\infty} \exp(i\xi x)\rho_X(x)dx$$

This is very similar to the Fourier transform (FT) of PDF $\rho_X(x)$

Fourier transform (FT): $f(x) \rightarrow \hat{f}(\xi)$

$$\hat{f}(\xi)=\int_{-\infty}^{+\infty}\exp(-i2\pi\xi x)f(x)dx$$

Inverse transform: $\hat{f}(\xi)\to f_2(x)$

$$f_2(x)=\int_{-\infty}^{+\infty}\exp(i2\pi\xi x)\hat{f}(\xi)d\xi$$

Theorem:

$$f_2(x)=f(x)$$

This theorem justifies the name "inverse transform".

Relation between characteristic function (CF) and Fourier transform (FT)

$$\phi_X(\xi)=\int_{-\infty}^{+\infty}\exp(i\xi x)\rho_X(x)dx$$

$$\hat{\rho}_X(\xi')=\int_{-\infty}^{+\infty}\exp(-i2\pi\xi' x)\rho_X(x)dx$$

==> $\phi_X(\xi)=\hat{\rho}_X(\xi')\Big|_{\xi'=\frac{-\xi}{2\pi}}$

Theorem (Properties of CF):

- $\phi_X(\xi)\Big|_{\xi=0}=1$

  Proof:

  $$\phi_X(\xi)\Big|_{\xi=0}=E\big(\exp(i\xi X)\big)\Big|_{\xi=0}=E(1)=1$$

- CF and the first moment

  $$\frac{d}{d\xi}\phi_X(\xi)\Big|_{\xi=0}=iE(X)$$

  Proof:

  $$\frac{d}{d\xi}\phi_X(\xi)=\frac{d}{d\xi}E\big(\exp(i\xi X)\big)=E\left(\frac{d}{d\xi}\exp(i\xi X)\right)=E\big(iX\exp(i\xi X)\big)$$

  ==> $\frac{d}{d\xi}\phi_X(\xi)\Big|_{\xi=0}=iE(X)$

- CF and the second moment

$$\left.\frac{d^2}{d\xi^2}\phi_X(\xi)\right|_{\xi=0} = -E(X^2)$$

- Expansion of CF around $\xi = 0$

$$\phi_X(\xi) = 1 + iE(X)\xi - \frac{E(X^2)}{2}\xi^2 + \cdots$$

- Mapping CF back to PDF.

  If $\phi_X(\xi) = \phi_Y(\xi)$, then $\rho_X(s) = \rho_Y(s)$.

  Proof: this property follows from the invertibility of FT:

  if $\hat{f}(\xi) = \hat{g}(\xi)$, then $f(x) = g(x)$.

- CF of the sum of two independent RVs.

  If random variables X and Y are independent, then we have

$$\phi_{(X+Y)}(\xi) = \phi_X(\xi)\cdot\phi_Y(\xi)$$

  Proof:

$$\phi_{(X+Y)}(\xi) = E\Big(\exp(i\xi(X+Y))\Big) = E\Big(\exp(i\xi X)\cdot\exp(i\xi Y)\Big)$$

(using the independence)

$$= E\Big(\exp(i\xi X)\Big)\cdot E\Big(\exp(i\xi Y)\Big)$$

$$= \phi_X(\xi)\cdot\phi_Y(\xi)$$

- CF of a shifted RV.

  Let $Y = \mu + X$. The CFs of the two are related by

$$\phi_Y(\xi) = \exp(i\xi\mu)\phi_X(\xi)$$

  Proof:

$$\phi_Y(\xi) = E\Big(\exp(i\xi Y)\Big) = E\Big(\exp(i\xi(\mu+X))\Big) = \exp(i\xi\mu)E\Big(\exp(i\xi X)\Big) = \exp(i\xi\mu)\phi_X(\xi)$$

- CF of a scaled RV.

  Let $Y = \sigma X$. The CFs of the two are related by

$$\phi_Y(\xi) = \phi_X(\sigma\xi)$$

  Proof:

$$\phi_Y(\xi) = E\Big(\exp(i\xi Y)\Big) = E\Big(\exp(i\xi\sigma X)\Big) = E\Big(\exp\big(i(\sigma\xi)X\big)\Big) = \left.\phi_X(\xi')\right|_{\xi'=\sigma\xi}$$

<u>CF of a normal random variable</u>:

$X \sim N(\mu, \sigma^2)$

PDF: $\rho_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} \right)$

Characteristic function:

$$\phi_X(\xi) \equiv E\Big(\exp(i\xi X)\Big) = \int \exp(i\xi x)\rho_X(x)dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp\left(\frac{-(x-\mu)^2 + i2\sigma^2\xi x}{2\sigma^2} \right) dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp\left(\frac{-(x-\mu)^2 + i2\sigma^2\xi(x-\mu) - (i\sigma^2\xi)^2}{2\sigma^2} + i\mu\xi - \frac{\sigma^2\xi^2}{2} \right) dx$$

$$= \exp\left(i\mu\xi - \frac{\sigma^2\xi^2}{2} \right) \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp\left(\frac{-[(x-\mu) - i\sigma^2\xi]^2}{2\sigma^2} \right) dx$$

change of variables: $z = (x-\mu) - i\sigma^2\xi$

$$= \exp\left(i\mu\xi - \frac{\sigma^2\xi^2}{2} \right) \frac{1}{\sqrt{2\pi\sigma^2}} \int \exp\left(\frac{-z^2}{2\sigma^2} \right) dz$$

$$= \exp\left(i\mu\xi - \frac{\sigma^2\xi^2}{2} \right)$$

We obtain:

$$\boxed{X \sim N(\mu, \sigma^2) \quad \longrightarrow \quad \phi_X(\xi) = \exp\left(i\mu\xi - \frac{\sigma^2\xi^2}{2} \right)}$$

Since we can map the characteristic function (CF) back to the PDF, we can check if a RV has a normal distribution by inspecting its CF.

<u>Theorem</u> (CF of a normal RV):

$$X \sim N(\mu, \sigma^2) \quad \text{if and only if} \quad \phi_X(\xi) = \exp\left(i\mu\xi - \frac{\sigma^2\xi^2}{2} \right).$$

We apply the theorem above to the <u>sum of two independent normal RVs</u>.

<u>Theorem</u> (sum of two independent normal RVs)

Suppose X and Y are independent, and $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$.

Then we have

$(X + Y) \sim N(\mu_1+\mu_2, \sigma_1^2 + \sigma_2^2)$.

Proof:

$$\phi_{(X+Y)}(\xi) = \phi_X(\xi)\cdot\phi_Y(\xi) = \exp\left(i\mu_1\xi - \frac{\sigma_1^2\xi^2}{2}\right)\cdot\exp\left(i\mu_2\xi - \frac{\sigma_2^2\xi^2}{2}\right)$$

$$= \exp\left(i\left(\mu_1+\mu_2\right)\xi - \frac{\left(\sigma_1^2+\sigma_2^2\right)\xi^2}{2}\right)$$

which is the CF of $N(\mu_1+\mu_2, \sigma_1^2 + \sigma_2^2)$.

Remark: The normal distribution is closed to addition of independent RVs.

Conversely, we can use this property to "derive" the PDF of normal distribution.

Let $\Theta(\mu, \sigma^2)$ denote a hypothetical distribution with mean = μ and variance = σ^2.

Let $f(x; \mu, \sigma^2)$ be the PDF of distribution $\Theta(\mu, \sigma^2)$.

Theorem:

Suppose the distribution family Θ is closed to i) translation, ii) scalar multiplication and iii) addition of independent RVs. Then the PDF must have the expression

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

Proof: see Appendix A

Monty Hall's game:

Before we start the discussion of stochastic differential equations, let us look at another example to demonstrate the importance of the framework of repeated experiments.

A possible background:

Your group and Mike's group will do summer camping together. But each group has a different itinerary in mind. To decide on a joint itinerary, you and Mike play a game ONCE with Mike hosting.

Specifications of the simple game:

1) The host (Mike) puts a card in one of the 3 boxes without you looking

   (so he knows which box has the card but you don't know).

2) You select a box.
3) If the box you pick contains the card, you win

    (and you will have priority in the itinerary planning).

    Otherwise, you lose

    (and Mike will have priority in the itinerary planning).
4) At the end, all boxes are opened to verify that the host is not cheating.

The incident:

After you select box #1, Mike says "Let us play it the Monty Hall style".

- He opens box #2, which is empty, and
- he offers you the option of switching to box #3.

Question: Should you switch?

Answer: The behavior of the host (Mike) is incompletely specified.

There are many possible ways the game can be repeated.

Version 1: Monty Hall's game

This is Mathematicians' definition of Monty Hall's game. The real game show hosted by Monty Hall did not actually follow these rules.

- Upon your initial selection, the host must open a box that you did not pick.
- The host must open an empty box.
- The host must offer you the option of switching.

For this game, we have

Pr(winning | not switching) = 1/3

Pr(winning | switching) = 2/3

See Appendix B1 for derivation.

Version 2: (The greedy host)

The greedy host wants to lure you away from the correct box.

- Upon your initial selection, the host may or may not open a box that you did not pick. The greedy host will open a box if and only if you initial selection is correct.
- If the host opens a box, he must open an empty box.
- If the host opens a box, he must offer you the option of switching.

For this game, we have

Pr(winning | not switching if offered) = 1/3

Pr(winning | switching if offered) = 0

See Appendix B2 for derivation.

Caution: the condition in the first conditional probability is

"not switching if offered"

which includes two cases:

i) you are not offered the option of switching, and

ii) you are offered the option but you do not switch.

In particular, it does not say you are offered the option of switching.

The same description applies to the second probability.

Version 3: (The less greedy host)

The less greedy host still wants to lure you away from the correct box. But he wants to avoid this behavior being easily recognized in repeated games.

- Upon your initial selection, the host may or may not open a box that you did not pick. The less greedy host adds some randomness to the decision on whether or not to open a box.

  Pr(opening a box| your initial selection is incorrect) = p_1

  Pr(opening a box| your initial selection is correct) = p_2

- If the host opens a box, he must open an empty box.
- If the host opens a box, he must offer you the option of switching.

Version 2 is a special case of Version 3 with $p_1 = 0$ and $p_2 = 1$.

For this game, we have

Pr(winning | not switching if offered) = 1/3

Pr(winning | switching if offered) = $(1+2p_1-p_2)/3$

For example, for $p_1 = 0.25$ and $p_2 = 0.75$

Pr(winning | switching if offered) = 0.25

See Appendix B3 for derivation.

Key observation:

When you encounter an incompletely specified game only ONCE you have to make a model perceiving how the game is repeated. The model is subjective.

Stochastic differential equation

$$dX(t) = b(X(t),t)dt + \sqrt{a(X(t),t)}\,dW(t)$$

Or in a more concise form

$$dX = b(X,t)dt + \sqrt{a(X,t)}\,dW$$

Without the dW term, differential equation $dX = b(X, t)dt$ is easy to understand.

We need to introduce $W(t)$.

The Wiener process (Brownian motion)

Definition 1:

The Wiener process, denoted by $W(t)$ satisfies

1) $W(0) = 0$
2) For $t \ge 0$, $W(t) \sim N(0, t)$
3) For $t_4 \ge t_3 \ge t_2 \ge t_1 \ge 0$,

   increments $W(t_2)-W(t_1)$ and $W(t_4)-W(t_3)$ are independent.

Definition 2:

1) $W(0) = 0$
2) For $t_2 \ge t_1 \ge 0$, $W(t_2)-W(t_1) \sim N(0, t_2-t_1)$
3) For $t_4 \ge t_3 \ge t_2 \ge t_1 \ge 0$,

   increments $W(t_2)-W(t_1)$ and $W(t_4)-W(t_3)$ are independent.

Definition 2 appears to be stronger than Definition 1.

Question: Are these two definitions equivalent?

Answer: Yes.

Theorem:

Suppose X and Y are independent, and $X \sim N(\mu_1, \sigma_1^2)$ and $(X+Y) \sim N(\mu_2, \sigma_2^2)$.

Then we have

$$Y \sim N(\mu_2-\mu_1, \sigma_2^2- \sigma_1^2)$$

Proof: Homework problem.

Remark:

Suppose X and Y are independent.

A previous theorem: $X \sim N(\,)$ and $Y \sim N(\,)$ ==> $X+Y \sim N(\,)$.

<u>The current theorem</u>: $X \sim N(\ )$ and $X+Y \sim N(\ )$ ==> $Y \sim N(\ )$.

Using this theorem, we show that Definition 1 is as strong as Definition 2.

We start with Definition 1 and derive Definition 2.

Definition 1:

==> $W(t_1) \sim N(0, t_1)$ and

$W(t_2) \sim N(0, t_2)$

We write $W(t_2)$ as a sum

$$W(t_2) = W(t_1) + (W(t_2) - W(t_1))$$

For $t_2 \geq t_1 \geq 0$, $W(t_1)$ and $(W(t_2) - W(t_1))$ are independent.

Applying the theorem above, we conclude

$$W(t_2) - W(t_1) \sim N(0, t_2 - t_1)$$

which is Definition 2.

Appendix A

<u>Theorem</u>:

Suppose the distribution family Θ is closed to i) translation, ii) scalar multiplication and iii) addition of independent RVs. Then the PDF must have the expression

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2} \right)$$

<u>Proof:</u>

Consider independent RVs X and Y with $X \sim \Theta(0, 1)$ and $Y \sim \Theta(0, \varepsilon)$.

We first compare $Y \sim \Theta(0, \varepsilon)$ and $\sqrt{\varepsilon}\, X \sim \Theta(0, \varepsilon)$

==> $\phi_Y(\xi) = \phi_{\sqrt{\varepsilon} X}(\xi) = \phi_X(\sqrt{\varepsilon}\, \xi)$

Next we compare $X + Y \sim \Theta(0, 1+\varepsilon)$ and $\sqrt{1+\varepsilon}\, X \sim \Theta(0, 1+\varepsilon)$

==> $\phi_{\sqrt{1+\varepsilon} X}(\xi) = \phi_{(X+Y)}(\xi) = \phi_X(\xi)\phi_Y(\xi)$

==> $\phi_X(\sqrt{1+\varepsilon}\, \xi) = \phi_X(\xi)\phi_X(\sqrt{\varepsilon}\, \xi)$ (E01)

We expand the LHS and RHS of (E01) in terms of ε.

$$\phi_X(\sqrt{1+\varepsilon}\,\xi) = \phi_X(\xi + (\varepsilon/2)\xi) = \phi_X(\xi) + \phi_X'(\xi)\frac{\varepsilon}{2}\xi + \cdots$$

$$\phi_X(\sqrt{\varepsilon}\,\xi) = 1 - \frac{\varepsilon}{2}\xi^2 + \cdots$$

Substituting the expansions into (E01) yields

$$\phi_X(\xi) + \phi_X'(\xi)\frac{\varepsilon}{2}\xi + \cdots = \phi_X(\xi) - \phi_X(\xi)\frac{\varepsilon}{2}\xi^2 + \cdots$$

Equating the coefficients of corresponding ε terms on both sides, we get

$$\phi_X'(\xi) = -\phi_X(\xi)\xi$$

==> $$\frac{d}{d\xi}\ln\phi_X(\xi) = -\xi$$

This is an ODE for $\phi_X(\xi)$. Solving it with condition $\phi_X(0) = 1$ gives us

$$\phi_X(\xi) = \exp\left(\frac{-\xi^2}{2}\right)$$

==> $$\rho_X(x) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2}{2}\right)$$

With translation μ and scalar multiplication σ, we obtain

$$\phi_{\mu+\sigma X}(\xi) = \exp\left(i\xi\mu - \frac{\sigma^2\xi^2}{2}\right)$$

Therefore, for the general case of (μ, σ^2), we arrive at

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

Appendix B1

Suppose you never switch.

Pr(winning | not switching)

= Pr(your initial selection is correct) = 1/3

Suppose you always switch. Recall that the host must open an empty box.

You will switch to the correct box if and only if you initial section is incorrect.

Pr(winning | switching)

= Pr(your initial selection is incorrect) = 2/3

Appendix B2

Suppose you never switch if offered.

Pr(winning | not switching if offered)

= Pr(your initial selection is correct) = 1/3

Suppose you always switch if offered. Let

C = "your initial selection is correct"

O = "the host opens an empty box and offers you the option "

S = "switching if offered"

W = "winning"

The greedy host opens a box and offers you the option of switching if and only if your initial selection is correct.

O <==> C

We use the law of total probability.

Pr(winning | switching if offered) = Pr(W | S)

= Pr(W | C and S) Pr(C) + Pr(W| C^C and S) Pr(C^C)

= 0 × (1/3) + 0× (2/3) = 0

Appendix B3

Suppose you never switch if offered.

Pr(winning | not switching if offered)

= Pr(your initial selection is correct) = 1/3

Suppose you always switch if offered. Let

C = "your initial selection is correct"

O = "the host opens an empty box and offers you the option "

S = "switching if offered"

W = "winning"

The host decides whether or not to open a box with probabilities

$\Pr(O \mid C^C) = p_1$

$\Pr(O \mid C) = p_2$

We use the law of total probability.

Pr(winning | switching if offered) = Pr(W | S)

= Pr(W | C and O and S) Pr(C and O)

+ Pr(W | C and O^C and S) Pr(C and O^C)

+ Pr(W | C^C and O and S) Pr(C^C and O)

+ Pr(W | C^C and O^C and S) Pr(C^C and O^C)

We first calculate the various terms used in the law of total probability.

Pr(C and O) = Pr(O | C) Pr(C) = p_2*(1/3)

Pr(C and O^C) = Pr(O^C | C) Pr(C) = (1− p_2)*(1/3)

Pr(C^C and O) = Pr(O | C^C) Pr(C^C) = p_1*(2/3)

Pr(C^C and O^C) = Pr(O^C | C^C) Pr(C^C) = (1− p_1)*(2/3)

Pr(W | C and O and S) = 0

Pr(W | C and O^C and S) = 1

Pr(W | C^C and O and S) = 1

Pr(W | C^C and O^C and S) = 0

Substituting these terms into the law of total probability, we obtain

Pr(winning | switching if offered)

= 0 + (1− p_2)*(1/3) + p_1*(2/3) + 0

= (1+2p_1−p_2)/3