ECE511: Analysis of Random Signals Spring 2014

Lecture 13 — 10, 20, 2014

Dr. Salim El Rouayheb *Scribe: Rawad Bitar*

1 Overview

In the last lecture, we talked about Chebyshev's and Markov's inequalities. Then, we proved a weak version of the Law of large numbers before closing by defining the Moment Generating Function of a RV.

In this lecture, we will introduce Chernoff bound, define the characteristic function of a RV, give some examples and conclude by proving the Central Limit Theorem with examples.

2 Moment Generating Functions (MGF)

Recall that the moment generating function $\mathcal{M}_X(t)$ is by definition equal to $E(e^{tX})$ for $t \in \mathbb{C}$.

Example: Let $X \sim \text{Poisson}(\lambda)$, $P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \ldots$.

1. Find $\mathcal{M}_x(t)$.

$$\begin{aligned}
\mathcal{M}_X(t) &= E(e^{t\lambda}) = \sum_{k=0}^{\infty} e^{tk} P(X=k) \\
\mathcal{M}_X(t) &= \sum_{k=0}^{\infty} e^{tk} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!} \\
\mathcal{M}_X(t) &= e^{-\lambda}e^{\lambda e^t}.
\end{aligned}$$

2. Find $E(X)$ from $\mathcal{M}_x(t)$.

$$E(X) = \frac{\partial \mathcal{M}_X(t)}{\partial t} \mid_{t=0} = \lambda e^t e^{\lambda(e^t-1)} \mid_{t=0} = \lambda.$$

3 Chernoff Bound

In this section, we introduce the Chernoff bound. Recall that to use Markov's inequality X must be positive.

Theorem 1. *(Chernoff's bound) For any RV X,*

$$P(X \geq a) \leq e^{-at}\mathcal{M}_X(t)\ \forall\, t > 0.$$

In particular,

$$P(X \geq a) \leq \min_t e^{-at}\mathcal{M}_X(t).$$

Proof. Apply Markov on $Y = e^{tX}$, but first recall that $P(X \geq a) = P(e^{tX} \geq e^{ta}) = P(Y \geq e^{ta})$, by Markov we get

$$P(Y \geq e^{ta}) \leq \frac{E(Y)}{e^{ta}} = e^{-ta}E(Y)$$
$$P(X \geq a) \leq e^{-ta}\mathcal{M}_X(t)$$

□

Example: Consider $X \sim N(\mu, \sigma)$ and try to bound $P(X \geq a)$ using Chernoff bound, this is an artificial example because we know the distribution of X.

From last lecture $\mathcal{M}_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$ hence

$$P(X) \leq \min_t e^{-at}e^{\mu t + \frac{\sigma^2 t^2}{2}} = \min_t e^{(\mu - a)t + \frac{\sigma^2 t^2}{2}}$$

Remark: You can check at home how the parameter t can affect the outer bound. For example pick $\mu = 0, \sigma = 1$ and change t; for $t = 0$ you will get the trivial bound $P \leq 1$ and for $t \to \infty$ you will get $P \leq \infty$. See how it varies.

$$\min_t e^{(\mu-a)t+\frac{\sigma^2 t^2}{2}} \Rightarrow \frac{\partial f(t)}{\partial t} = 0$$
$$\Rightarrow (\sigma^2 t + \mu - a)e = 0$$
$$\Rightarrow t^* = \frac{a - \mu}{\sigma^2}$$

Which gives us the following:

$$P(X \geq a) \leq e^{(\mu-a)t^* + \frac{\sigma^2 t^*}{2}}$$
$$P(X \geq a) \leq e^{\frac{-(a-\mu)(\mu-a)}{\sigma^2} + \frac{\sigma^2(a-\mu)^2}{2\sigma^4}}$$
$$P(X \geq a) \leq e^{\frac{-(a-\mu)^2}{2\sigma^2}}$$

We can compare this result with the reality where we know that $P(X \geq a) = \int_a^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$.

4 Characteristic Function

In this section, we define a characteristic function and give some examples. The characteristic function of a RV is similar to a Fourrier transform of a function without the '-'.

Definition: X is a RV,

$$\Phi_X(w) = E(e^{jwX}) = \int_{-\infty}^{+\infty} f_X(x) e^{jwx} dx, \tag{1}$$

is called the characteristic function of X where j is the complex number $j^2 = -1$.

Example: Find the characteristic function of $X \sim \exp(\lambda)$. Recall that for $\lambda \geq 0$

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

Then

$$\begin{aligned} \Phi_X(w) &= \int_0^{\infty} \lambda e^{-\lambda x} e^{jwx} dx \\ &= \lambda \int_0^{\infty} e^{(jw-\lambda)x} dx \\ &= \frac{\lambda}{jw - \lambda} \left[e^{(jw-\lambda)x} \right]_0^{\infty} \end{aligned}$$

Since $\lambda \geq 0$ and jw is a unit quantity $\Rightarrow$ $(jw - \lambda) \leq 0$ therefore $\lim_{x \to \infty} e^{(jw-\lambda)x} = 0$. Which results in

$$\begin{aligned} \Phi_X(w) &= \frac{\lambda}{jw - \lambda}(0 - 1) \\ \Phi_X(w) &= \frac{\lambda}{\lambda - jw} \end{aligned}$$

Lemma 2. *if X, Y are two independent RV and $Z = X + Y$ then $\Phi_Z(w) = \Phi_X(w)\Phi_Y(w)$ and $\mathcal{M}_Z(t) = \mathcal{M}_X(t)\mathcal{M}_Y(t)$*

Remark: To find the distribution of $Z = X + Y$ it could be easier to find $\Phi_X(w)$, $\Phi_Y(w)$, multiply them and then invert the from "Fourrier" domain by integrating or by using tables of Fourrier inverse.

Example: Consider the example of problem 9 of homework 3:

Question: Let X_1 and X_2 be two independent RV such that $X_1 \sim N(\mu_1, \sigma_1)$ and $X_2 \sim N(\mu_2, \sigma_2)$ and let $X = aX_1 + bX_2$. Find the distribution of X.

Answer: Let $X_1' = aX_1$, $X_2' = bX_2$ it is clear that $X_1' \sim N(a\mu_1, a\sigma_1)$ and $X_2' \sim N(b\mu_2, b\sigma_2)$ and that $\Phi_X(w) = \Phi_{X_1'}(w)\Phi_{X_2'}(w)$.

$$\Phi_{X_1'}(w) = \ldots = e^{a\mu_1 jw - \frac{a^2\sigma_1^2 w^2}{2}}$$
$$\Phi_X(w) = e^{a\mu_1 jw - \frac{a^2\sigma_1^2 w^2}{2}} e^{b\mu_2 jw - \frac{b^2\sigma_2^2 w^2}{2}}$$
$$\Phi_X(w) = e^{j(a\mu_1 + b\mu_2)w - (a^2\sigma_1^2 + b^2\sigma_2^2)\frac{w^2}{2}}$$

Which implies that $X \sim N(a\mu_1 + b\mu_2, \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2})$.

Fact 3. *A linear combination of two independent Gaussian RV is a Gaussian RV.*

5 Central Limit Theorem

In this section we state the central limit theorem and give a rigourous proof.

Theorem 4. *Let $X_1, X_2, \ldots, X_n$ be n independent RVs with $\mu_{X_i} = 0$ and $V(X_i) = 1$ $\forall i$ then*

$$Z_n = \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \underset{n\to\infty}{\to} N(0,1)$$

In other words

$$\lim_{n\to\infty} P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx$$

This is for example a way to convert flipping a coin n times to a Gaussian RV (fig. 1,fig. 2).

$$X_i = \begin{cases} 0 & \text{if a tail is observed with } p = \frac{1}{2} \\ 1 & \text{if a head is observed with } 1 - p = \frac{1}{2} \end{cases}$$

And set $S_n = \frac{\sum_{i=0}^{n} X_i}{\sqrt{n}}$, notice that $S_n \in \{0, 1, \ldots, \frac{n}{\sqrt{n}}\}$ and according to CLT $S_n \underset{n\to\infty}{\sim} N(0,1)$.

CLT: says that no matter how far you are from the mean, the probability of $X = x$ being outside $|x - \mu| \leq \sqrt{n}$ decreases exponentially with n.

Remark: The RVs X_i have to be independent because if for example $X_i = X_1$ for $i \in \{2, 3, \ldots, n\}$ then

$$S_n = nX_i = \begin{cases} \sqrt{n} & \text{if } X_1 = 1 \\ 0 & \text{if } X_1 = 0 \end{cases}$$

which does not converge to a Gaussian distribution when $n \to \infty$.

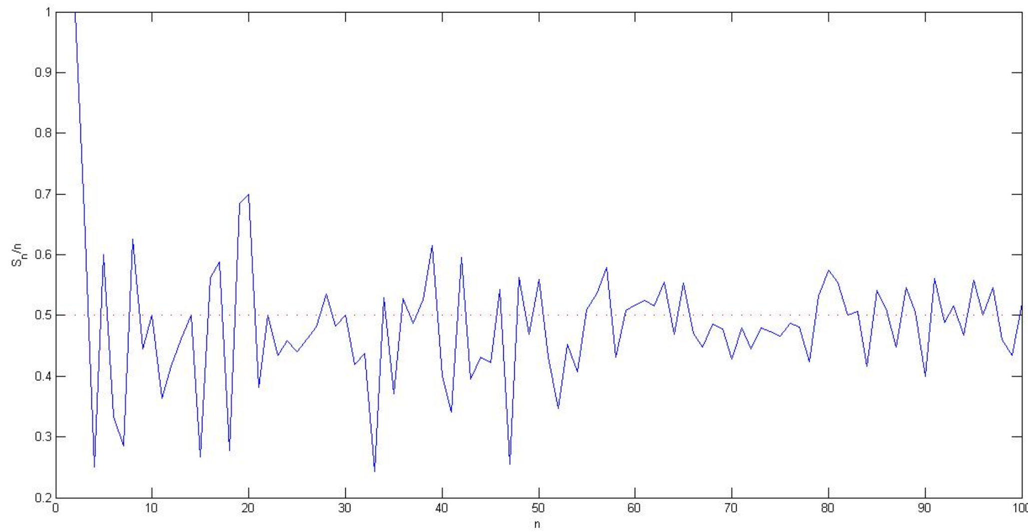

Figure 1: This is $\frac{S_n}{n}$ as a function of n, we can clearly see that when n grows $\frac{S_n}{n}$ goes to 0.5 for the equation of the example below for n goes to 100. Refer to section 6 for detailed code.

Proof. (of theorem 4)

$$\lim_{n\to\infty} \Phi_{Z_n}(w) = e^{-\frac{w^2}{2}} \Rightarrow f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

where this form of $\Phi_{Z_n}(w)$ is the characteristic function of a $N(0,1)$ RV.

$$Z_n = \underbrace{\frac{X_1}{\sqrt{n}}}_{W_1} + \underbrace{\frac{X_2}{\sqrt{n}}}_{W_2} + \cdots + \underbrace{\frac{X_n}{\sqrt{n}}}_{W_n}$$

$$\Phi_{Z_n}(w) = \Phi_{W_1}(w)\Phi_{W_2}(w)\Phi_{W_3}(w)\ldots\Phi_{W_n}(w) = [\Phi_{W_1}(w)]^n$$

$$\Phi_{W_1}(w) = E(e^{wjW_1}) = E(e^{\frac{jwX_1}{\sqrt{n}}}) = \Phi_{X_1}(\frac{w}{\sqrt{n}})$$

Taylor expansion: Using the Taylor expansion of $\Phi_{W_1}(w)$ around 0 we get,

$$\Phi_{W_1}(w) = \Phi_{W_1}(0) + \Phi'_{W_1}(0) + \frac{\Phi''_{W_1}(0)}{2!} + \ldots$$

.

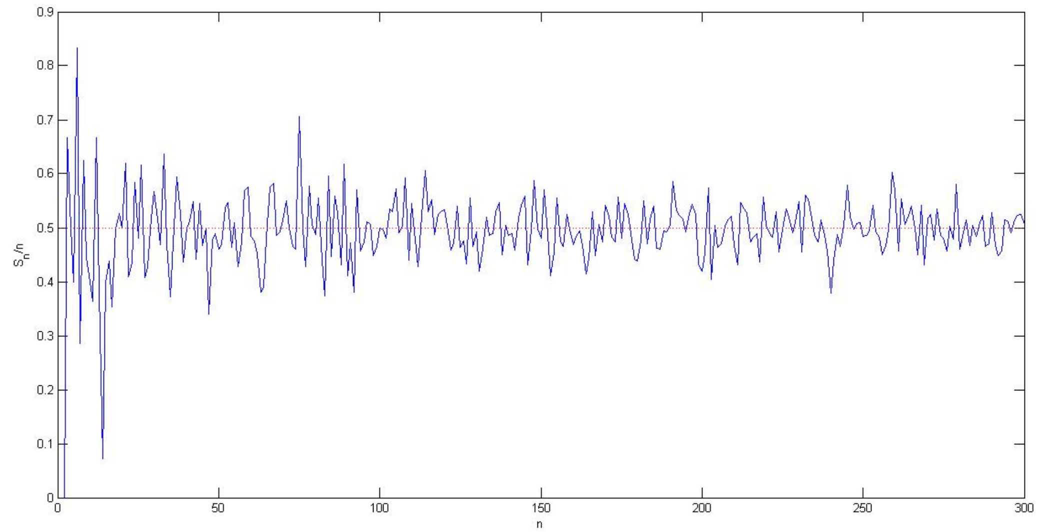

Figure 2: This is $\frac{S_n}{n}$ as a function of n, we can clearly see that when n grows $\frac{S_n}{n}$ goes to 0.5 for the equation of the example below for n goes to 300. Refer to section 6 for detailed code.

1. Find the value of $\Phi_{W_1}(0)$.

$$\begin{aligned}
\Phi_{W_1}(0) &= E(e^{\frac{jwX}{\sqrt{n}}})\,|_{w=0} \\
&= \int_{-\infty}^{+\infty} e^{\frac{j0X}{\sqrt{n}}} f_X(x)dx \\
&= \int_{-\infty}^{+\infty} f_X(x)dx \\
&= 1
\end{aligned}$$

2. Find the value of $\Phi'_{W_1}(0)$.

$$\begin{aligned}
\Phi'_{W_1}(0) &= \int_{-\infty}^{+\infty} \frac{jx}{\sqrt{n}} e^{\frac{j0X}{\sqrt{n}}} f_X(x)dx \\
&= \frac{j}{\sqrt{n}} \int_{-\infty}^{+\infty} x f_X(x)dx \\
&= \frac{j}{\sqrt{n}} E(X_1) \\
&= 0
\end{aligned}$$

3. Find the value of $\Phi''_{W_1}(0)$.

$$\begin{aligned}
\Phi''_{W_1}(0) &= \frac{d^2\Phi_{W_1}(w)}{dw^2} \\
&= \int_{-\infty}^{+\infty} (\frac{jx}{\sqrt{n}})^2 e^{\frac{j0X}{\sqrt{n}}} f_X(x)dx \\
&= \frac{-1}{n}\int_{-\infty}^{+\infty} x^2 f_X(x)dx \\
&= \frac{-1}{n}(\underbrace{V(X)}_{1} + \underbrace{E^2(X)}_{0}) \\
&= \frac{-1}{n}
\end{aligned}$$

Hence, using these results and Taylor's expansion, $\Phi_{W_1}(w) = 1 - \frac{w^2}{2n}$. Therefore $\Phi_{Z_n}(w) = [1 - \frac{w^2}{2n}]^n$. Recall that $\log{(1-\epsilon)} \simeq -\epsilon$, then

$$\begin{aligned}
\log \Phi_{Z_n} &= n \log{(1 - \frac{w^2}{2n})} \\
\log \Phi_{Z_n} &\simeq n(-\frac{w^2}{2n}) \\
\log \Phi_{Z_n} &\simeq -\frac{w^2}{2} \\
\Phi_{Z_n} &= e^{-\frac{w^2}{2}}
\end{aligned}$$

□

6 MATLAB Code generating the figures

In this section we give the MATLAB code used to generate fig. 1 and fig 2.

```
A=[];B=[];   % generate two empty vectors
for i=1:100  % in this loop i stands for the number of times the coin is flipped
    A=[A,binornd(i,0.5)/i]; % at each iteration generate a binomial random number with
end                         % parameters n=i, p=0.5 and divide it by n to have (S_n)/n
for n=1:300
B=[B,binornd(n,0.5)/n];     % same as previous but repeat it 300 times
end

x1=[1:i];x2=[1:n];          % x1 and x2 are used to represent n in each figure

figure(1)
plot(x1,A);
hold on
plot(x1,0.5,'r','linewidth',2);
xlabel('n');
```

```
ylabel('S_n/n');

figure(2)
plot(x2,B);
hold on
plot(x2,0.5,'r','linewidth',2);
xlabel('n');
ylabel('S_n/n');
```