Chapter 7
Sums of Random Variables and Long-Term Averages

ENCS6161 - Probability and Stochastic Processes

Concordia University
Sums of Random Variables

Let $X_1, \cdots, X_n$ be r.v.s and $S_n = X_1 + \cdots + X_n$, then

$$E[S_n] = E[X_1] + \cdots + E[X_n]$$

$$Var[S_n] = Var[X_1 + \cdots + X_n]$$

$$= E \left[ \sum_{i=1}^{n} (X_i - \mu_{X_i}) \sum_{j=1}^{n} (X_j - \mu_{X_j}) \right]$$

$$= \sum_{i=1}^{n} Var[X_i] + \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j)$$

If $Z = X + Y$ ($n = 2$),

$$Var[Z] = Var[X] + Var[Y] + 2Cov(X, Y)$$
Sums of Random Variables

Example: Sum of \( n \) i.i.d r.v.s with mean \( \mu \) and variance \( \sigma^2 \).

\[
E[S_n] = E[X_1] + \cdots + E[X_n] = n\mu
\]

\[
Var[S_n] = nVar[X_i] = n\sigma^2
\]

pdf of sums of independent random variables

\( X_1, \ldots, X_n \) indep r.v.s and \( S_n = X_1 + \cdots + X_n \), then

\[
\Phi_{S_n}(w) = E[e^{jwS_n}] = E[e^{jw(X_1+\cdots+X_n)}] = \Phi_{X_1}(w) \cdots \Phi_{X_n}(w)
\]

and

\[
f_{S_n}(s) = \mathcal{F}^{-1}\{\Phi_{X_1}(w) \cdots \Phi_{X_n}(w)\}
\]
Sums of Random Variables

Example: $X_1 \cdots X_n$ indep and $X_i \sim N(m_i, \sigma_i^2)$. What is the pdf of $S_n = X_1 + \cdots + X_n$?

For a Gaussian r.v.

$X \sim N(\mu, \sigma^2) \Rightarrow \Phi_X(w) = e^{j\mu w - \frac{w^2\sigma^2}{2}}$

(prove it by yourself)

So

$\Phi_{S_n}(w) = \prod_{i=1}^{n} e^{j\mu m_i - \frac{w^2\sigma_i^2}{2}} = e^{j\mu (m_1 + \cdots + m_n) - \frac{w^2(\sigma_1^2 + \cdots + \sigma_n^2)}{2}}$

$\therefore S_n \sim N(m_1 + \cdots + m_n, \sigma_1^2 + \cdots + \sigma_n^2)$

What if $X_1, \cdots, X_n$ are not indep?? (hint: use $Y = AX$)
Sums of Random Variables

- pdf of i.i.d r.v.s
  \[ \Phi_{S_n}(w) = (\Phi_X(w))^n \]

Example: Find the pdf of the sum of \( n \) i.i.d exponential r.v.s with parameter \( \lambda \).
\[ \Phi_X(w) = \frac{\lambda}{\lambda - jw} \] (see table 3.2 on page 101)

\[ \Rightarrow \Phi_{S_n}(w) = \left( \frac{\lambda}{\lambda - jw} \right)^n \]

\[ \Rightarrow f_{S_n}(s) = \frac{\lambda e^{-\lambda s} (\lambda s)^{n-1}}{(n-1)!}, \ s > 0 \]

This is the so called m-Erlang r.v.
Sums of Random Variables

When dealing with non-negative integer-valued r.v.s, we use the probability generating function:

\[ G_N(z) = E[z^N] = \sum_n z^n P_N(n) \]

\[ P_N(n) = \frac{1}{n!} \frac{d^n}{dz^n} G_N(z) \bigg|_{z=0} \]

For \( N = X_1 + \cdots + X_n \) where \( X_i \) are independent.

\[ G_N(z) = E[z^{X_1+\cdots+X_n}] \]
\[ = E[z^{X_1}] \cdots E[z^{X_n}] \]
\[ = G_{X_1}(z) \cdots G_{X_n}(z) \]
Sums of Random Variables

Example: Find the pdf of the sum of $n$ independent Bernoulli r.v.s with $p_0 = 1 - p = q$ and $p_1 = p$.

\[
G_X(z) = E[z^X] = q + pz
\]

\[\Rightarrow G_N(z) = (q + pz)^n\]

\[\Rightarrow P_N(k) = \binom{n}{k} p^k q^{n-k}, \ k = 0, 1 \cdots n\]

(See Table 3.1)

\[\Rightarrow N \sim \text{Binomial}(n, p)\]
The Sample Mean

Let $X$ be a r.v. with mean $\mu$ and variance $\sigma^2$. $X_1, \cdots, X_n$ denote $n$ independent, repeated measurement of $X$. That is, $X_i$’s are i.i.d r.v.s with the same pdf as $X$. The sample mean is defined as

$$M_n = \frac{S_n}{n} = \frac{X_1 + \cdots + X_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

The mean and variance of the sample mean are

$$E[M_n] = E\left[\frac{1}{n} \sum_{i=1}^{n} X_i\right] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \mu$$

$$Var[M_n] = E[(M_n - \mu)^2] = E \left[ \left( \frac{S_n - E(S_n)}{n} \right)^2 \right]$$

$$= \frac{1}{n^2} E[(S_n - E(S_n))^2] = \frac{1}{n^2} Var[S_n] = \frac{\sigma^2}{n}$$
The Laws of Large Numbers

- From Chebyshev inequality for any $\varepsilon > 0$

$$
P \{ |M_n - \mu| \geq \varepsilon \} \leq \frac{Var[M_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}
$$

So $P \{ |M_n - \mu| < \varepsilon \} > 1 - \frac{\sigma^2}{n\varepsilon^2}$

- The weak law of large numbers:

$$
\lim_{n \to \infty} P \{ |M_n - \mu| < \varepsilon \} = 1
$$

for any $\varepsilon > 0$.

- The strong law of large numbers:

$$
P \left\{ \lim_{n \to \infty} M_n = \mu \right\} = 1
$$

The proof is beyond the level of this course.
The Central Limit Theorem

Let $X_1, \cdots, X_n$ be i.i.d r.v.s with $\mu, \sigma^2$ and $S_n = X_1 + \cdots + X_n$. Let

$$Z_n = \frac{S_n - n\mu}{\sigma \sqrt{n}}$$

then as $n \to \infty$, the distribution of $Z_n$ tends to standard Gaussian.

$$\lim_{n \to \infty} P[Z_n \leq z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx = 1 - Q(z) = \Phi(z)$$
The Central Limit Theorem

Proof:

\[ \Phi_{Z_n}(w) = E[e^{jwZ_n}] = E\left[ e^{\frac{jw}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)} \right] \]

\[ = E\left[ \prod_{i=1}^{n} e^{\frac{jw(X_i - \mu)}{\sigma\sqrt{n}}} \right] = \prod_{i=1}^{n} E\left[ e^{\frac{jw(X_i - \mu)}{\sigma\sqrt{n}}} \right] (\because \text{indep}) \]

\[ = \left\{ E\left[ e^{\frac{jw(X_i - \mu)}{\sigma\sqrt{n}}} \right] \right\}^n (\because \text{i.i.d}) \]

(to be continued)
The Central Limit Theorem

Proof: (continues)

\[
E \left[ e^{\frac{jw(X_i - \mu)}{\sigma \sqrt{n}}} \right] = E \left[ 1 + \frac{jw}{\sigma \sqrt{n}}(X - \mu) + \frac{(jw)^2}{2!n\sigma^2}(X - \mu)^2 + R(w) \right]
\]

\[
= 1 + \frac{jw}{\sigma \sqrt{n}} E[X - \mu] - \frac{w^2}{2n\sigma^2} E[(X - \mu)^2] + E[R(w)]
\]

\[
= 1 - \frac{w^2}{2n} + E[R(w)]
\]

\(E[R(w)]\) becomes negligible compared to \(\frac{w^2}{2n}\) when \(n \to \infty\), therefore

\[
\lim_{n \to \infty} \Phi_{Z_n}(w) = \lim_{n \to \infty} \left(1 - \frac{w^2}{2n}\right)^n = e^{-\frac{w^2}{2}}
\]

So, when \(n \to \infty\), \(Z_n \sim N(0, 1)\)
Convergence of Sequence of R.V.s

- $X_1, \ldots, X_n$ are r.v.s, how to define the convergence of r.v.s? Recall: a r.v. is a function: $S \rightarrow R$. So $X_1(w), X_2(w), \ldots$ are functions.

- If $X_n(w) \rightarrow X(w)$ for all $w$, **sure convergence**

- If $P\{w | X_n(w) \rightarrow X(w)\} = 1$, **almost sure convergence**, $X_n \rightarrow X$ a.s. (or w.p. 1)

- If $E[(X_n(w) - X(w))^2] \rightarrow 0$ as $n \rightarrow \infty$, **mean square convergence**, $X_n \rightarrow X$ m.s.

- If $\forall \varepsilon > 0$, $P\{|X_n(w) - X(w)| > \varepsilon\} \rightarrow 0$, convergence in probability.
Convergence of Sequence of R.V.s

- a.s. convergence ⇒ convergence in probability
- m.s. convergence ⇒ convergence in probability
  But almost sure \( \neq \) mean square.

Convergence in distribution

\( X_n \) has cdf \( F_n(x) \) and \( X \) has cdf \( F(x) \). If \( F_n(x) \rightarrow F(x) \) for all \( x \) where \( F(x) \) is continuous. We call \( X_n \) converge to \( X \) in distribution.

Convergence in prob. ⇒ convergence in distribution.
Convergence of Sequence of R.V.s

Note:
weak LLN: convergence in prob.
\[ M_n \to \mu \text{ in prob.} \]
strong LLN: almost sure
\[ M_n \to \mu \text{ a.s.} \]
CLT: convergence in distribution
\[ Z_n \to Z \sim N(0, 1) \text{ in dist.} \]

In fact \( M_n \to \mu \text{ m.s. since } \)
\[ E[(M_n - \mu)^2] = Var[M_n] = \frac{\sigma^2}{n} \to 0 \text{, as } n \to \infty \]