# midterm2

maryam norouzi

December 2023

## 1

### 1.1

To find the characteristic functions of $X_k$ and $Y_n$, we will proceed as follows:

For $X_k$, which can take the values $\pm\sqrt{k}$ each with probability $\frac{1}{2}$, the characteristic function $\phi_{X_k}(t)$ is defined by:

$$\phi_{X_k}(t) = E[e^{itX_k}] = \sum_{x \in X_k} e^{itx} P(X_k = x)$$

Since $X_k$ takes on two values with equal probability, the characteristic function is:

$$\phi_{X_k}(t) = \frac{1}{2}e^{it\sqrt{k}} + \frac{1}{2}e^{-it\sqrt{k}}$$

This can be simplified using Euler's formula $e^{ix} = \cos(x) + i\sin(x)$, recognizing that the sine terms will cancel out as they will be equal and opposite for $+\sqrt{k}$ and $-\sqrt{k}$:

$$\phi_{X_k}(t) = \frac{1}{2}(\cos(t\sqrt{k}) + i\sin(t\sqrt{k})) + \frac{1}{2}(\cos(t\sqrt{k}) - i\sin(t\sqrt{k}))$$

$$\phi_{X_k}(t) = \frac{1}{2} \cdot 2 \cdot \cos(t\sqrt{k})$$

$$\phi_{X_k}(t) = \cos(t\sqrt{k})$$

Now, for $Y_n$, which is the average of $n$ such $X_k$ variables, the characteristic function $\phi_{Y_n}(t)$ can be found by noting that the characteristic function of the sum of independent random variables is the product of their characteristic functions, and then we must adjust for the fact that $Y_n$ is an average, not a sum:

$$\Phi_{Y_n}(t) = \prod_{k=1}^{n} \Phi_{X_k}(t) = \prod_{k=1}^{n} \cos\left(t\sqrt{\frac{k}{n}}\right)$$

$$\phi_{Y_n}(t) = \left(\phi_{X_k}\left(\frac{t}{n}\right)\right)^n$$

Substituting the characteristic function of $X_k$ we found earlier:

$$\phi_{Y_n}(t) = \left(\cos\left(\frac{\sqrt{k}t}{n}\right)\right)^n$$

These are the characteristic functions for $X_k$ and $Y_n$ respectively.

## 1.2

b) The characteristic function of $X_k$ is given by

$$\phi_{X_k}(t) = E(e^{itX_k})$$

The characteristic function of $Y_n$ is then

$$\phi_{Y_n}(t) = \left(\cos\left(\frac{t}{\sqrt{n}k}\right)\right)^n$$

As $n \to \infty$, $\phi_{Y_n}(t)$ converges to

$$\phi_Y(t) = 1$$

By the Continuity Theorem, if $X_n \xrightarrow{d} X$, then

$$\lim_{n\to\infty} F_n(x) = F_X(x)$$

where $F_X(x)$ is the CDF of $X$. Hence,

$$\lim_{n\to\infty} F_n(x) = \delta(x)$$

Given the characteristic function for $Y_n$:

$$\Phi_{Y_n}(t) = \prod_{k=1}^{n} \cos\left(t\sqrt{\frac{k}{n}}\right)$$

we take the limit of the logarithm of $\Phi_{Y_n}(t)$ as $n$ approaches infinity:

$$\lim_{n\to\infty} \log\left(\Phi_{Y_n}(t)\right) = \lim_{n\to\infty} n \log\left(\cos\left(\frac{t}{\sqrt{n}}\right)\right)$$

Evaluating this limit, we find that it is $-\frac{t^2}{2}$, which corresponds to the logarithm of the characteristic function of a standard normal distribution:

$$\lim_{n\to\infty} \log\left(\Phi_{Y_n}(t)\right) = -\frac{t^2}{2}$$

Thus, the characteristic function for $Y_n$ converges to:

$$\lim_{n\to\infty} \Phi_{Y_n}(t) = e^{-\frac{t^2}{2}}$$

This implies that $Y_n$ converges in distribution to a standard normal distribution as $n$ goes to infinity. Therefore, the cumulative distribution function $F_n(x)$ converges to the CDF of a standard normal distribution:

$$\lim_{n\to\infty} F_n(x) = \Phi(x)$$

where $\Phi(x)$ is the CDF of a standard normal distribution.

# 2

## 2.1

Given a sequence of iid random variables $\{X_i\}_{i=1}^{\infty}$ with all moments existing and finite, particularly with mean $m = E[X_i]$ and variance $\sigma^2 = \text{var}[X_i]$, and a sequence of positive constants $\{a_i\}_{i=1}^{\infty}$ where $a_1 = k$ and $a_n = a_{n-1}+k$ for $n > 1$, we seek to find the means and variances of the weighted sum $S_n = \sum_{i=1}^{n} a_i X_i$ and average $M_n = S_n / \sum_{i=1}^{n} a_i$.

The mean of $S_n$ is calculated as follows:

$$\begin{aligned}
E[S_n] &= E\left[\sum_{i=1}^{n} a_i X_i\right] \\
&= \sum_{i=1}^{n} a_i E[X_i] \\
&= m \sum_{i=1}^{n} a_i \\
&= m \sum_{i=1}^{n} (ik) \\
&= mk \sum_{i=1}^{n} i \\
&= mk \frac{n(n+1)}{2} \\
&= \frac{mk}{2} n(n+1).
\end{aligned}$$

The variance of $S_n$ is given by:

$$\begin{aligned}
\operatorname{var}(S_n) &= \operatorname{var}\left(\sum_{i=1}^{n} a_i X_i\right) \\
&= \sum_{i=1}^{n} a_i^2 \operatorname{var}(X_i) \\
&= \sigma^2 \sum_{i=1}^{n} a_i^2 \\
&= \sigma^2 \sum_{i=1}^{n} (ik)^2 \\
&= \sigma^2 k^2 \sum_{i=1}^{n} i^2 \\
&= \sigma^2 k^2 \frac{n(n+1)(2n+1)}{6} \\
&= \frac{\sigma^2 k^2}{6} n(n+1)(2n+1).
\end{aligned}$$

For $M_n$, the mean is the mean of $S_n$ divided by the sum of the weights $a_i$, and the variance is the variance of $S_n$ divided by the square of the sum of the weights $a_i$. Thus, we have:

$$\begin{aligned}
E[M_n] &= \frac{E[S_n]}{\sum_{i=1}^{n} a_i} \\
&= \frac{\frac{mk}{2} n(n+1)}{k \frac{n(n+1)}{2}} \\
&= m,
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{var}(M_n) &= \frac{\operatorname{var}(S_n)}{\left(\sum_{i=1}^{n} a_i\right)^2} \\
&= \frac{\frac{\sigma^2 k^2}{6} n(n+1)(2n+1)}{\left(k \frac{n(n+1)}{2}\right)^2} \\
&= \frac{6\sigma^2}{k^2 n^2 (n+1)^2} \cdot \frac{n(n+1)(2n+1)}{6} \\
&= \frac{\sigma^2 (2n+1)}{k^2 n(n+1)}.
\end{aligned}$$

The sum of the sequence $a_i$ where each term follows the pattern $a_n = a_{n-1} + k$ for $n > 1$ with $a_1 = k$ is an arithmetic sequence with common difference $k$. The $n$-th term of this sequence, $a_n$, is $nk$.

The sum of the first $n$ terms of an arithmetic sequence is given by the formula:

$$S_n = \frac{n}{2}(a_1 + a_n)$$

For our sequence, this becomes:

$$S_n = \frac{n}{2}(k + nk)$$
$$S_n = \frac{n}{2}(k(1 + n))$$
$$S_n = \frac{nk}{2}(1 + n)$$

Hence, the sum of $a_i$ from $i = 1$ to $n$ is:

$$\sum_{i=1}^{n} a_i = \frac{nk}{2}(1 + n)$$

## 2.2

To show that $M_n \xrightarrow{P} m$, we need to show that for any $\epsilon > 0$, the probability that the absolute difference between $M_n$ and $m$ is greater than $\epsilon$ goes to 0 as $n$ approaches infinity:

$$P(|M_n - m| > \epsilon) \to 0 \text{ as } n \to \infty.$$

From part a), we have that the expected value $E[M_n] = m$. To establish convergence in probability, we consider the variance of $M_n$, which is:

$$\text{Var}(M_n) = \frac{\sigma^2}{k^2} \cdot \frac{(2n+1)}{n(n+1)^2}.$$

As $n$ approaches infinity, the variance of $M_n$ approaches 0:

$$\lim_{n\to\infty} \text{Var}(M_n) = \lim_{n\to\infty} \frac{\sigma^2}{k^2} \cdot \frac{(2n+1)}{n(n+1)^2} = 0.$$

By Chebyshev's inequality, for any $\epsilon > 0$:

$$P(|M_n - m| > \epsilon) \leq \frac{\sigma_m}{\epsilon}.$$
$$\sigma_m = \sqrt{\text{Var}}$$

Since the variance of $M_n$ goes to 0 as $n$ approaches infinity, it follows that:

$$\lim_{n\to\infty} P(|M_n - m| > \epsilon) \leq \lim_{n\to\infty} \frac{\sqrt{\text{Var}(M_n)}}{\epsilon} = 0.$$

Hence, $M_n$ converges in probability to $m$.

$$(M_n \xrightarrow{P} m)$$

## 2.3

To show that the standardized version of $S_n$ converges in distribution to a standard normal random variable $Z$ with distribution $N(0, 1)$, we can invoke the Central Limit Theorem (CLT).

The CLT states that if $X_1, X_2, \ldots, X_n$ are independent and identically distributed (i.i.d.) random variables with mean $E[X_i] = \mu$ and variance $\mathrm{Var}(X_i) = \sigma^2 > 0$, then as $n$ approaches infinity, the standardized sum:

$$Z_n = \frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}(S_n)}}$$

converges in distribution to a standard normal random variable $Z$, that is:

$$Z_n \xrightarrow{d} Z$$

# Expected Value and Variance of $S_n$

Given the weighted sum $S_n = \sum_{i=1}^{n} a_i X_i$, where $X_i$ are i.i.d random variables with mean $m$ and variance $\sigma^2$, and $a_i$ are positive constants:

The expected value and the variance are calculated as:

$$E[S_n] = \sum_{i=1}^{n} a_i E[X_i] = m \sum_{i=1}^{n} a_i = mk \frac{n(n+1)}{2}$$

$$\mathrm{Var}[S_n] = \sum_{i=1}^{n} a_i^2 \mathrm{Var}[X_i] = \sigma^2 k^2 \frac{n(n+1)(2n+1)}{6}$$

# Limitations of CLT and the Need for Lyapunov's Condition

The Central Limit Theorem states that the sum (or average) of a large number of independent, identically distributed random variables with finite mean and variance will be approximately normally distributed, irrespective of the original distribution of the random variables.

The traditional CLT requires identically distributed random variables with finite variance. If the random variables are not identically distributed, or if the variances grow quickly (as with weighted sums), we use Lyapunov's version of the CLT.

# Lyapunov's Central Limit Theorem

Lyapunov's CLT is a generalization for sequences of independent, not necessarily identically distributed random variables, introducing a condition for their sum to be approximately normally distributed.

## Lyapunov's Condition

Lyapunov's condition, necessary for the CLT to apply, is for some $\delta > 0$:

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{n} E\left[|X_i - E[X_i]|^{2+\delta}\right] = 0$$

where $s_n^2$ is the sum of the variances of the random variables.

For the Central Limit Theorem to apply, we need to check Lyapunov's condition. Specifically, we need to show that for some $\delta > 0$:

$$\lim_{n\to\infty} \frac{1}{(\mathrm{Var}(S_n))^{1+\delta/2}} \sum_{i=1}^{n} E[|a_i X_i|^{\delta}] = 0.$$

For $\delta = 1$, this becomes:

$$\lim_{n\to\infty} \frac{1}{(\sigma^2 \sum_{i=1}^{n} a_i^2)^{3/2}} \sum_{i=1}^{n} a_i^3 E[|X_i|^3] = 0.$$

The standardized sum $Z_n$ is then given by:

$$Z_n = \frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}(S_n)}} = \frac{\sum_{i=1}^{n} a_i X_i}{\sqrt{\sigma^2 \sum_{i=1}^{n} a_i^2}}.$$

If Lyapunov's condition is satisfied, the standardized sum $Z_n$ converges in distribution to $Z$ as $n \to \infty$, where $Z$ is a standard normal variable $N(0, 1)$.

We assume that $a_i$ are either constants or follow a pattern such that Lyapunov's condition is met. Under these assumptions, we can apply the Central Limit Theorem to conclude that $Z_n$ converges in distribution to a standard normal distribution.

## Proof Using Lyapunov's Condition

### Setting Up

Define $Z_i = \frac{a_i}{m}(X_i - m)$ and calculate $s_n^2$ as the sum of the variances of $Z_i$.

### Applying the Condition

Verify Lyapunov's condition by showing that for the chosen $\delta$, the sum of the $(2 + \delta)$-th central moments of $Z_i$ normalized by $s_n^{2+\delta}$ approaches zero as $n$ becomes large.

$$Z_i = \frac{a_i}{m} X_i - a_i$$

$$E[Z_i] = \frac{a_i}{m} E[X_i] - a_i = 0$$

$$E[Z_i^2] = \frac{a_i^2}{m^2} E[X_i^2] - a_i$$

As defined, all moments of $X_i$ are finite, therefore $E[Z_i^3] < \infty$.

The variance of $Z_i$ is given by:

$$\mathrm{Var}(Z_i) = \mathrm{Var}\left(\frac{a_i}{m} X_i - 1\right) = \frac{a_i^2}{m^2}\mathrm{Var}(X_i) = \frac{\sigma^2}{m^2} a_i^2$$

Then we have:

$$\sigma(n)^2 = \mathrm{Var}\left(\sum_{j=1}^{n} Z_j\right) = \sum_{j=1}^{n} \frac{\sigma^2}{m^2} a_j^2 =$$

$$\frac{k^2\sigma^2}{m^2}\left(\frac{n^4}{4} + \frac{n^3}{2} + \frac{n^2}{4}\right) + \frac{k^2\sigma^2}{m^2}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}\right) + \frac{k^2\sigma^2}{m^2}\left(\frac{n^2}{2} + \frac{n}{2}\right) =$$

$$\frac{\sigma^2}{m^2} k^2 \left(\frac{n(n+1)(2n+1)}{6}\right)$$

Given:

$$\sigma(n)^2 = \mathrm{Var}\left(\sum_{j=1}^{n} Z_j\right) = \sum_{j=1}^{n} \frac{\sigma^2}{m^2} a_j^2$$

$$= \frac{\sigma^2}{m^2} k^2 \left(\frac{n(n+1)(2n+1)}{6}\right)$$

We then examine the third moment:

$$\frac{1}{(\sigma(n))^3} \sum_{j=1}^{n} E\left[Z_j^3\right]$$

$$\leq \frac{1}{\left(\frac{\sigma^2}{m^2} k^2 \left(\frac{n(n+1)(2n+1)}{6}\right)\right)^{3/2}} \cdot \left(\frac{\sigma^2}{m^2} k^2 \cdot \left(\frac{n(n+1)(2n+1)}{6}\right)\right)^{3/2}$$

$$= \frac{1}{\left(\frac{\sigma^2}{m^2} k^2\right)^{3/2}} \cdot \left(\frac{n(n+1)(2n+1)}{6}\right)^{3/2} \cdot \left(\frac{\sigma^2}{m^2} k^2\right)^{3/2} \cdot \left(\frac{n(n+1)(2n+1)}{6}\right)^{3/2}$$

$$= \frac{1}{\left(\frac{n(n+1)(2n+1)}{6}\right)^{3/2}} \cdot \left(\frac{\sigma^2}{m^2} k^2\right)^{3/2} \cdot \left(\frac{n(n+1)(2n+1)}{6}\right)^{3/2}$$

$$\sum_{j=1}^{n} a_j^3 \to k^3 \cdot \frac{n^2(n+1)^2}{4} \quad \text{because the nominator is dominated by the term} \quad n^4$$

As $n$ approaches infinity, the above expression tends towards zero, satisfying Lyapunov's condition. Hence: $\sum Z_j \sim N(0,1)$, then $S_n$ is also $\sim N(0,1)$.

So, if the sum of $Z_j$ converges in distribution to a normal distribution with mean 0 and variance 1, i.e., $\sum Z_j \sim N(0,1)$, then the sum $S_n$ also converges to a normal distribution:

$$\frac{1}{\sigma(n)} \cdot \sum_{j=1}^{n} Z_j \to N(0,1)$$

And thus, as $\{Z_j\}$ converges to $N(0,1)$, so does $S_n$.

## Concluding the Proof

Upon confirming Lyapunov's condition, it follows that:

$$\frac{S_n - E[S_n]}{\sqrt{\mathrm{Var}[S_n]}} \xrightarrow{D} N(0,1)$$

indicating that the standardized sum converges in distribution to a standard normal distribution. Hence, if $\sum Z_j \sim N(0,1)$, then $S_n$ is also normal as $n \to \infty$, confirming the convergence in distribution to $Z \sim N(0,1)$.

- The random variables $X_j$ are assumed to be independent with a mean of 0 ($E[X_j] = 0$) and a finite variance ($\mathrm{Var}(X_j) = \sigma^2$).
- They have a finite third absolute moment ($E[|X_j|^3] < \infty$), which allows for the use of Lyapunov's CLT.
- A standardized variable $Z_i$ is defined as a normalized form of the weighted random variables $a_i X_i$, with $E[Z_i]$ being 0 and the variance of $Z_i$ being the weighted variance of $X_i$.
- The variance of the sum $S_n$ is calculated, which involves the sum of the squares of the weights $a_j^2$.
- Lyapunov's condition is checked by considering the sum of $a_j^3 E[|X_j|^3]$ normalized by the cubed variance of the sum, which must approach 0 as $n$ goes to infinity for the CLT to apply.
- It is concluded that the normalized sum $\frac{1}{\sigma(n)} \sum_{j=1}^{n} Z_j$ converges in distribution to $N(0,1)$.
- If the sum of the $Z_j$ converges to $N(0,1)$, then the original sum $S_n$, normalized by its expectation and standard deviation, also converges to $N(0,1)$.

# 3

## 3.1

## Chernoff Bounds for Probability Inequality

Given $S_n$ as the sum of i.i.d random variables, we want to find the probability that the squared deviation of $S_n$ from its mean is greater than $n$:

$$P\left(\left(\frac{S_n - \mu_{S_n}}{\sigma_{S_n}}\right)^2 > n\right) = P(S_n > \mu_{S_n} + \sqrt{n}\sigma_{S_n}) + P(S_n < \mu_{S_n} - \sqrt{n}\sigma_{S_n})$$

Using Chernoff bounds, this probability can be bounded by:

$$\leq M(t) \cdot e^{-t^2 n} \text{ for } t \geq 0, \text{ and } M(t) \cdot e^{t^2 n} \text{ for } t < 0$$

where $M_S(t) = (M_{X_i}(t))^n$ is the moment-generating function of $S_n$, given by:

$$M_{S_n}(t) = \left(\frac{1}{1-t^2}\right)^n$$

Then the Chernoff bounds can be expressed as:

$$\leq \left(\frac{1}{1-t^2}\right)^n \cdot e^{-t\sqrt{n}} \text{ for } t \geq 0, \text{ and } \left(\frac{1}{1-t^2}\right)^n \cdot e^{t\sqrt{n}} \text{ for } t < 0$$

Finally, we are interested in the limit as $n$ approaches infinity:

$$\lim_{n\to\infty} = 0$$

To apply the Chernoff bound, we first calculate the moment generating function (MGF) for a single random variable $X_i$. Given the symmetry and the exponential nature of the Laplace distribution, the MGF for $|t| < 1$ is given by

$$M_{X_i}(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \frac{1}{1-t^2}.$$

The Chernoff bound provides an exponentially decreasing bound on the tail probabilities of the sum of random variables. For our case, the Chernoff bound is

$$P(S_n \geq s) \leq \min_{t>0} \left\{e^{-ts}\left(M_{X_i}(t)\right)^n\right\}.$$

For $s = n$, we seek to minimize the above expression over $t$ in the range $(-1, 1)$. We performed a numerical minimization for a large value of $n = 1000$ and obtained an optimal $t \approx 0.414$ with a Chernoff bound approximately $7.16 \times 10^{-99}$.

This result indicates that the probability $P(S_n > n)$ is exceedingly small for large $n$, which is consistent with the Law of Large Numbers. As $n$ increases, the Chernoff bound decreases exponentially, supporting the claim that the probability in question approaches 0 as $n$ approaches infinity.

## 3.2

The problem is to find the limit:

$$\lim_{n\to\infty} \left(P\left(S_n^2 > n^2\right)\right)^{\frac{1}{n}}$$

This limit involves a probability raised to a power that diminishes as $n$ grows. The Large Deviation Principle (LDP) typically estimates probabilities of the form $P\left(\frac{S_n}{n} > x\right)$ as $n \to \infty$, not probabilities inside a power of $\frac{1}{n}$.

Assuming that $P\left(S_n^2 > n^2\right)$ can be approximated using LDP, we would have:

$$P\left(S_n^2 > n^2\right) \approx e^{-nI(2)}$$

where $I(x)$ is the rate function derived from the logarithmic moment-generating function (MGF) of the random variables $X_i$.

For a standard LDP problem, the rate function $I(x)$ for the random variable $X_i$ with PDF $f(x) = \frac{1}{2}e^{-|x|}$ is derived as follows:

1. Calculate the logarithmic moment-generating function:

$$\Lambda(\lambda) = \log E\left[e^{\lambda X_i}\right]$$

2. Find the Legendre-Fenchel transform, which is the rate function:

$$I(x) = \sup_{\lambda\in\mathbb{R}} \{\lambda x - \Lambda(\lambda)\}$$

Given the Laplace distribution, the MGF is:

$$M(\lambda) = \frac{1}{1-\lambda^2}$$

for $|\lambda| < 1$, and the log-MGF is:

$$\Lambda(\lambda) = \log\left(\frac{1}{1-\lambda^2}\right)$$

The supremum is found by setting the derivative of $\lambda x - \Lambda(\lambda)$ with respect to $\lambda$ to zero.

To calculate the rate function $I(x)$ for $x = n$ since we are dealing with $S_n^2 > n^2$, we solve for $\lambda$ in the equation:

$$\frac{d}{d\lambda}\Lambda(\lambda) = x$$

to find $\lambda^*$, and then find $I(n)$ using:

$$I(x) = \lambda^* x - \Lambda(\lambda^*)$$

The limit as $n \to \infty$ of the expression $\left(P\left(S_n^2 > n^2\right)\right)^{\frac{1}{n}}$ is calculated to be 0, which suggests that the probability $P(S_n^2 > n^2)$ decays to zero faster than any exponential decay of the form $e^{-cn}$ for $c > 0$.