Distribution of Random Samples & Limit theorems

1 Distribution of i.i.d. Samples

Motivating example. Assume that the goal of a study is to demonstrate the digital transformation in the US population over the past 5 years. One feature of this technological revolution is for instance the time spent on smartphones. Let us denote $X$ that random variable for some US individual. If one is interested in $E[X]$ the mean time spent on smartphones by the US population, that quantity is exactly

$$E[X] = \frac{1}{N} \sum_{i=1}^{N} x_i$$

where $N \approx 320$ million, is the size of the US population and $x_i$ is the time spent by individual $i$ on his/her smartphone. However, in most statistical studies, it is rare to have access to the whole population. Therefore, we need some principled guidance to estimate $E[X]$ from a sample of the population of size $n$ with generally $n \ll N$.

Definition 1.1 (i.i.d. Sample). Let $X_1, \ldots, X_n$ be collection of $n \in \mathbb{N}^*$ random variables on probability space $(\Omega, \mathcal{A}, \mathbb{P})$. $(X_1, \ldots, X_n)$ is a random sample of size $n$ if and only if:

1. $X_1, \ldots, X_n$ are mutually independent
2. $X_1, \ldots, X_n$ are identically distributed, that is, each $X_i$ comes from the same distribution.

We say that $X_1, \ldots, X_n$ are independent and identically distributed (abbreviated i.i.d.).

A good statistical sample is such that the picked individuals are representative of the population. The latter is ensured by the mutual independence of the $n$ individuals. Upon existence, the population mean is denoted $\mu = E[X_1]$ and the variance of the population $\sigma^2 = \text{Var}(X_1)$.

Definition 1.2 (Sample mean). Let $(X_1, \ldots, X_n)$ be a random sample. Then the random variable

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

is called the sample mean.
We are interested in the distribution of $\mathbf{X}_n$. We need to define formally the distribution of more than 2 random variables. You should see that most of the following definitions/properties are natural extensions of the case $n = 2$.

**Definition 1.3** (Joint cumulative distribution function). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The joint (cumulative) distribution function (c.d.f.) of $X_1, \ldots, X_n$ is defined as follows:

$$F_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = \mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad (2)$$

for $(x_1, \ldots, x_n) \in \mathbb{R}^n$

**Definition 1.4** (Discrete rrvs). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ discrete rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The joint probability mass function (pmf) of $X_1, \ldots, X_n$, denoted is defined as follows:

$$p_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n), \quad (3)$$

for $(x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega)$.

**Property 1.1.** Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ discrete rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with joint pmf $p_{X_1, \ldots, X_n}$, then the following holds:

- $p_{X_1, \ldots, X_n}(x_1, \ldots, x_n) \geq 0$, for $(x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega)$
- $\sum_{x_1 \in X_1(\Omega)} \cdots \sum_{x_n \in X_n(\Omega)} p_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = 1$

**Definition 1.5** (Joint probability density function). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$. $X_1, \ldots, X_n$ are said to be jointly continuous if there exists a function $f_{X_1, \ldots, X_n}$ such that, for any subset $B \subset \mathbb{R}^n$:

$$\mathbb{P}((X_1, \ldots, X_n) \in B) = \int_{B} \cdots \int_{B} f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n \quad (4)$$

Then function $f_{X_1, \ldots, X_n}$ is called the joint probability density function of $X_1, \ldots, X_n$.

**Property 1.2.** Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with joint pdf $f_{X_1, \ldots, X_n}$, then the following holds:

- $f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) \geq 0$, for $(x_1, \ldots, x_n) \in \mathbb{R}^n$
- $\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n = 1$

**Definition 1.6** (Marginal distributions). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

- (Discrete case) If $X_1, \ldots, X_n$ have joint pmf $p_{X_1, \ldots, X_n}$. Then the marginal probability mass function of $X_i$ is obtained by summing over all $X_j$, $j \neq i$. For instance,

$$p_{X_i}(x_1) = \sum_{x_2 \in X_2(\Omega)} \cdots \sum_{x_n \in X_n(\Omega)} p_{X_1, \ldots, X_n}(x_1, \ldots, x_n), \quad \text{for } x_1 \in X_1(\Omega) \quad (5)$$
Continuous case) If $X_1, \ldots, X_n$ have joint pdf $f_{X_1 \ldots X_n}$. Then the marginal probability density function of $X_i$ is obtained by integrating over all $X_j \, j \neq i$. For instance,

$$f_{X_i}(x_i) = \int \cdots \int_{\mathbb{R}^{n-1}} f_{X_1 \ldots X_n}(x_1, \ldots, x_n) \, dx_2 \cdots dx_n, \quad \text{for } x_i \in \mathbb{R}$$

(6)

**Definition 1.7** (Independence). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^\ast$ rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

• (Discrete case) If $X_1, \ldots, X_n$ have joint pmf $p_{X_1 \ldots X_n}$ with respective marginal pmfs $p_{X_1}, \ldots, p_{X_n}$. Then $X_1, \ldots, X_n$ are said to be independent if and only if:

$$p_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i), \quad \text{for all } (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$$

(7)

• (Continuous case) If $X_1, \ldots, X_n$ have joint pdf $f_{X_1 \ldots X_n}$ with respective marginal pdfs $f_{X_1}, \ldots, f_{X_n}$. Then $X_1, \ldots, X_n$ are said to be independent if and only if:

$$f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i), \quad \text{for all } (x_1, \ldots, x_n) \in \mathbb{R}^n$$

(8)

**Property 1.3** (Distribution of iid sample). Let $(X_1, \ldots, X_n)$ be a random sample of size $n \in \mathbb{N}^\ast$ on probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then

• (Discrete case)

$$p_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i), \quad \text{for all } (x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$$

(9)

• (Continuous case)

$$f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i), \quad \text{for all } (x_1, \ldots, x_n) \in \mathbb{R}^n$$

(10)

**Definition 1.8** (Expected Value). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^\ast$ rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and let $g : \mathbb{R}^n \to \mathbb{R}$.

• (Discrete case) If $X_1, \ldots, X_n$ have joint pmf $p_{X_1 \ldots X_n}$. Then, the mathematical expectation of $g(X_1 \ldots X_n)$, if it exists, is:

$$E[g(X_1 \ldots X_n)] = \sum_{x_1 \in X_1(\Omega)} \cdots \sum_{x_n \in X_n(\Omega)} g(x_1, \ldots, x_n)p_{X_1 \ldots X_n}(x_1, \ldots, x_n)$$

(11)
(Continuous case) If $X_1, \ldots, X_n$ have joint pdf $f_{X_1 \ldots X_n}$. Then, the mathematical expectation of $g(X_1 \ldots X_n)$, if it exists, is:

$$E[g(X_1 \ldots X_n)] = \int \cdots \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) f_{X_1 \ldots X_n}(x_1, \ldots, x_n)$$  \hspace{1cm} (12)

**Theorem 1.1.** Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ independent rrvs on probability space $(\Omega, \mathcal{A}, P)$ and let $g_1, \ldots, g_n$ be $n$ real-valued functions on $\mathbb{R}$. Then,

$$E[g_1(X_1) \ldots g_n(X_n)] = E[g_1(X_1)] \ldots E[g_n(X_n)]$$  \hspace{1cm} (13)

provided that the expectations exist.

**Theorem 1.2** (Variance of independent rrvs). Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ independent rrvs on probability space $(\Omega, \mathcal{A}, P)$. Then,

$$\text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i)$$  \hspace{1cm} (14)

provided that the variances exist.

**Property 1.4.** Let $(X_1, \ldots, X_n)$ be a random sample of size $n \in \mathbb{N}^*$ on probability space $(\Omega, \mathcal{A}, P)$ with mean $\mu = E[X_1]$ and variance $\sigma^2 = \text{Var}(X_1) < \infty$. Then

$$E[\overline{X}_n] = \mu$$  \hspace{1cm} (15)

and

$$\text{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$$  \hspace{1cm} (16)

## 2 Convergence of random variables

### 2.1 Convergence in probability

**Definition 2.1.** Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of rrvs on probability space $(\Omega, \mathcal{A}, P)$ and $X$ be a rrv on the same probability space. Sequence $(X_n)$ is said to converge in probability towards $X$ if, for all $\varepsilon > 0$:

$$\lim_{n \to \infty} \mathbb{P}( |X_n - X| > \varepsilon ) = 0$$  \hspace{1cm} (17)

Convergence in probability is denoted as follows:

$$X_n \overset{P}{\to} X$$

**Example 1.** Let $X$ be a discrete rrv with pmf $p_X$ defined by:

$$p_X(x) = \begin{cases} 1/3 & \text{if } x = 1 \\ 2/3 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$
and let \( X_n = (1 + \frac{1}{n})X \). Show that \( X_n \overset{P}{\to} X \).

**Answer.** We have:

\[
|X_n - X| = \left| X + \frac{X}{n} - X \right| = \frac{|X|}{n} = \frac{X}{n} \quad \text{since } X \text{ can only take nonnegative values.}
\]

Then, for any \( \varepsilon > 0 \), \( P(|X_n - X| > \varepsilon) = P(\frac{X}{n} > \varepsilon) \). Note that the event \( \{ \frac{X}{n} > \varepsilon \} \) can only occur when \( X = 1 \) and \( \varepsilon < \frac{1}{n} \) since \( \varepsilon > 0 \). Therefore, we get:

\[
P(|X_n - X| > \varepsilon) = \begin{cases} 
px(1) = 1/3 & n < 1/\varepsilon \\
0 & n > 1/\varepsilon
\end{cases}
\]

It now becomes obvious that \( P(|X_n - X| > \varepsilon) \) converges to 0, because it is identically equal to zero for all \( n > 1/\varepsilon \), which entails the desired result.

**Example 2.** For \( n \geq 1 \), let \( (X_n) \) be a sequence of random variables where \( X_n \) follows an exponential distribution with parameter \( n \). Show that \( (X_n) \) converges in probability to 0.

**Answer.** The probability density function of \( X_n \) is given by: \( f_{X_n}(x) = ne^{-nx}I_{(0,\infty)}(x) \). Let \( \varepsilon > 0 \) be an arbitrary constant, we have

\[
P(|X_n - 0| > \varepsilon) = P(X_n > \varepsilon) \quad \text{given that an exponential rrv can only take on nonnegative values}
\]

\[
= \int_{\varepsilon}^{\infty} ne^{-nx} \, dx = e^{-n\varepsilon} \overset{n\to\infty}{\longrightarrow} 0 \quad \text{since } \varepsilon > 0
\]

Hence, the result holds.

### 2.2 Almost sure convergence

**Definition 2.2.** Let \( (X_n)_{n\in\mathbb{N}} \) be a sequence of rrvs on probability space \((\Omega, \mathcal{A}, \mathbb{P})\) and \( X \) be a rrv on the same probability space. Sequence \( (X_n) \) is said to converge almost surely or almost everywhere or with probability 1 or strongly towards \( X \) if:

\[
P\left( \lim_{n\to\infty} X_n = X \right) = 1 \quad (18)
\]

Almost sure convergence is denoted as follows:

\[X_n \overset{a.s.}{\to} X\]
From Equation 18, we note that almost sure convergence is a slightly modified version of the concept of pointwise convergence of functions – recall that a random variable is formally a mapping from the sample space $\Omega$ to $\mathbb{R}$. That is,
\[
\forall \omega \in \Omega, \quad X_n(\omega) \overset{n \to \infty}{\longrightarrow} X(\omega)
\]
Requiring convergence for all $\omega \in \Omega$ is actually too stringent. To define almost sure convergence, we “relax” the above statement and allow that convergence might not be reached for some outcomes in $\Omega$. Rigorously, let $E$ be the following event:
\[
E = \{ \omega \in \Omega : X_n(\omega) \text{ does not converge to } X(\omega) \}
\]
and $F$ be an event with zero probability but $F$ is not the impossible event, i.e. $F \neq \emptyset$. Then we say that the sequence $(X_n)$ converges almost surely to $X$ if $E \subset F$. Almost sure convergence is a widely spread concept in the Probability and Statistics literature but proving almost sure convergence requires tools from measure theory, which is out of scope of this course.

2.3 Convergence in mean

Definition 2.3. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of rrvs on probability space $(\Omega, A, P)$ and $X$ be a rrv on the same probability space. Given a real number $r \geq 1$, sequence $(X_n)$ is said to converge in the $r$-th mean or in the $L^r$-norm towards $X$ if:
\[
\lim_{n \to \infty} E[|X_n - X|^r] = 0
\]
provided that for all $n$, $E[|X_n|^r]$ and $E[|X|^r]$ exist.

Convergence in the $r$-th mean is denoted as follows:
\[
X_n \overset{L^r}{\longrightarrow} X
\]

The most important cases of convergence in the $r$-th mean are:

- When Equation (19) holds for $r = 1$, we say that $(X_n)$ converges in mean to $X$

- When Equation (19) holds for $r = 2$, we say that $(X_n)$ converges in mean square to $X$

2.4 Convergence in distribution

Definition 2.4. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of rrvs on probability space $(\Omega, A, P)$. For any $n$, the distribution function of $X_n$ is denoted by $F_n$. Let $X$ be a rrv with distribution function $F_X$. Sequence $(X_n)$ is said to converge in distribution or converge weakly towards $X$ if:
\[
\lim_{n \to \infty} F_n(x) = F_X(x)
\]
for all $x \in \mathbb{R}$ at which $F_X$ is continuous.

Convergence in distribution is denoted as follows:

$$X_n \xrightarrow{D} X$$

The first fact to notice is that convergence in distribution, as the name suggests, only involves the distributions of the random variables. Thus, the random variables need not even be defined on the same probability space (that is, they need not be defined for the same random experiment), and indeed we don’t even need the random variables at all. This is in contrast to the other modes of convergence we have studied.

**Example 3.** Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of rrvs with cdf $F_n$

$$F_n(x) = \left( 1 - \left( 1 - \frac{1}{n} \right)^{nx} \right) \mathbb{1}_{(0, \infty)}(x)$$

What is the asymptotic distribution of $(X_n)$?

**Answer.** Note that for $x \in (-\infty, 0)$, we trivially have that $F_n(x) = 0 \xrightarrow{n \to \infty} 0$.

Now, let $x \in [0, \infty)$, a result in calculus gives:

$$\lim_{n \to \infty} \left( 1 - \frac{1}{n} \right)^{nx} = e^{-x}$$

Therefore,

$$F_n(x) = 1 - \left( 1 - \frac{1}{n} \right)^{nx} \xrightarrow{n \to \infty} 1 - e^{-x}.$$

We recognize the cumulative distribution function of an exponential distribution with parameter 1 – for those who are not convinced, you can differentiate the expression on the right-hand side. We conclude that the sequence $(X_n)$ converges to an exponential distribution with parameter 1.

**Theorem 2.1.** Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with respective mgfs $M_n$. Let $X$ be a rrv with mgf $M_X$. If the following holds:

$$\lim_{n \to \infty} M_n(x) = M_X(x) \quad (21)$$

for all $x \in \mathbb{R}$ where $M_n(x)$ and $M_X(x)$ exist, then sequence $(X_n)$ converges in distribution to $X$.

### 2.5 Implications between modes of convergence

The following summary gives the implications for the various modes of convergence; no other implications hold in general.
Proposition 2.1. 1. For $s > r \geq 1$, convergence in the $s$-th mean implies convergence in $r$-th mean.

2. Convergence in mean implies convergence in probability.

3. Almost sure convergence implies convergence in probability.


3 The Laws of Large Numbers

Property 3.1 (Markov’s Inequality). Let $X$ be a rrv that takes only on non-negative values. Then, for any $a > 0$, we have:

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

(22)

Property 3.2 (Bienaymé-Chebyshev’s Inequality). Let $X$ be a rrv that has expectation and variance. Then, for any $\alpha > 0$, we have:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \alpha) \leq \frac{\text{Var}(X)}{\alpha^2}$$

(23)

Theorem 3.1 (Weak law of large numbers). Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of i.i.d. rrvs, each having finite expectation. The weak law of large numbers (also called Khintchine’s law) states that the sample mean $\overline{X}_n$ converges in probability towards $\mathbb{E}[X_1]$, that is, for all $\epsilon > 0$:

$$\lim_{n \to \infty} \mathbb{P}(|\overline{X}_n - \mathbb{E}[X_1]| > \epsilon) = 0$$

(24)

Theorem 3.2 (Strong law of large numbers). Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of i.i.d. rrvs, each having finite expectation. The strong law of large numbers (SLLN) also called Kolmogorov’s strong law, states that the sample mean $\overline{X}_n$ converges almost surely towards $\mathbb{E}[X_1]$, that is:

$$\mathbb{P}\left(\lim_{n \to \infty} \overline{X}_n = \mathbb{E}[X_1]\right) = 1$$

(25)

Fundamental implication. Let $(X_n)$ be a sequence of independent Bernoulli random variables with parameter $p$, that is $X_n = 1$ when some event $E$ occurs with probability $p = \mathbb{P}(E)$ and $X_n = 0$ with probability $1 - p$ when $E$ does not occur. According to the Strong Law of Large Numbers,

$$\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{a.s.} \mathbb{E}[X_1] = p$$

In words, $\sum_{i=1}^{n} X_i$ is the number of times that $E$ occurs over $n$ trials. The SLLN thus states that the frequency of observing $E$ converges to $\mathbb{P}(E)$ as the size of the sample $n$ gets larger and larger. This justifies the frequentist school that sees the probability of an event as the theoretical frequency of observing that event.
4 The Central Limit Theorem

**Property 4.1.** Let $X_1, \ldots, X_n$ be $n \in \mathbb{N}^*$ independent rrvs on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with respective moment generating functions $M_1, \ldots, M_n$. Then the moment generating function of

$$S_n = \sum_{i=1}^{n} X_i$$

is:

$$M_{S_n}(x) = \prod_{i=1}^{n} M_i(x) \quad (26)$$

for all $x \in \mathbb{R}$ where $M_n(x)$ exist.

**Corollary 4.1** (Mgf of iid sample). Let $(X_1, \ldots, X_n)$ be a random sample of size $n \in \mathbb{N}^*$ on probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with moment generating function $M = M_{X_1}$. Then the moment generating function of

$$S_n = \sum_{i=1}^{n} X_i$$

is:

$$M_{S_n}(x) = (M(x))^n \quad (27)$$

for all $x \in \mathbb{R}$ where $M(x)$ exists.

**Theorem 4.2** (Central Limit Theorem). Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of i.i.d. rrvs, each having expectation $E[X_1] = \mu$ and finite variance $\text{Var}(X_1) = \sigma^2 < \infty$. The Central Limit Theorem states that the sequence of variables $(Z_n)_{n \in \mathbb{N}^*}$ defined by:

$$Z_n = \frac{X_n - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

converges in distribution towards $Z$ following a standard normal distribution $\mathcal{N}(0,1)$, that is:

$$\lim_{n \to \infty} F_{Z_n}(x) = \Phi(x), \quad \text{for all } x \in \mathbb{R} \quad (28)$$

Equivalently, the CLT can be rewritten as:

$$\overline{X}_n \xrightarrow{D} \mathcal{N} \left( \mu, \frac{\sigma^2}{n} \right)$$

**Applications of the CLT.** With the Strong Law of Large Numbers, the CLT is the other most important result in Probability and Statistics. In words, the CLT states that the distribution of the sum (or mean) of any iid random variables converges to a normal distribution provided that the population distribution has finite variance. As a consequence, you can use the normal distribution to approximate probabilities as long as the sample size $n$ is large enough.

How large is “large enough”? The answer depends on two factors.
• Requirements for accuracy. The more closely the sampling distribution needs to resemble a normal distribution, the more sample points will be required.

• The shape of the underlying population. The more closely the original population resembles a normal distribution, the fewer sample points will be required.

Empirical evidence shows that a sample size of 30 is large enough when the population distribution is roughly bell-shaped. Some statisticians may recommend a sample size of at least 40 though. But if the original population is distinctly not normal, the sample size should be even larger.

Example 4. Let \((X_1, \ldots, X_{15})\) be a random sample with probability density function:

\[
f(x) = \frac{3}{2} x^2 I_{(-1,1)}(x)
\]

What is the approximate probability that the sample mean \(\overline{X}_{15}\) falls between \(-2/5\) and \(1/5\)?

**Answer.** The CLT states that \(\overline{X}_{15}\) follows approximately a normal distribution \(\mathcal{N}(\mathbb{E}[X_1], \text{Var}(X_1)/15)\). Let us compute \(\mathbb{E}[X_1]\) and \(\text{Var}(X_1)\):

\[
\mathbb{E}[X_1] = \int_{-1}^{1} \frac{3}{2} x^2 dx = 0
\]

\[
\text{Var}(X_1) = \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2
\]

\[
= \int_{-1}^{1} x^2 \cdot \frac{3}{2} x^2 dx - 0^2
\]

\[
= \frac{3}{5}
\]

Therefore, \(\text{Var}(X_1)/15 = 3/75 = 1/25\)

\[
\overline{X}_{15} \overset{D}{\sim} \mathcal{N} \left( 0, \frac{1}{25} \right)
\]

Hence,

\[
P \left( \frac{-2}{5} \leq \overline{X}_{15} \leq \frac{1}{5} \right) = P \left( \frac{-2/5 - 0}{\sqrt{1/25}} \leq \frac{\overline{X}_{15} - 0}{\sqrt{1/25}} \leq \frac{1/5 - 0}{\sqrt{1/25}} \right)
\]

\[
\approx P (-2 \leq Z \leq 1) \quad (Z \text{ is a standard normal random variable})
\]

\[
\approx \Phi(1) - \Phi(-2) \quad (\Phi \text{ is the cdf of } \mathcal{N}(0, 1))
\]

\[
\approx \Phi(1) + \Phi(2) - 1
\]

\[
\approx 0.8413 + 0.9772 - 1
\]

\[
\approx 0.8185
\]