Recap: Multi-dimensional Gaussian density

- $\mathbf{X} = (X_1, \cdots, X_n)^T$ are said to be jointly Gaussian if

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \, |\Sigma|^{\frac{1}{2}}} \, e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

- $E\mathbf{X} = \boldsymbol{\mu}$ and $\Sigma_X = \Sigma$.
- The moment generating function is given by

$$M_{\mathbf{X}}(\mathbf{s}) = e^{\mathbf{s}^T \boldsymbol{\mu} \, + \, \frac{1}{2} \mathbf{s}^T \Sigma \, \mathbf{s}}$$

- When X, Y are jointly Gaussian, the joint density is given by

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \, e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right)}$$

Recap

- If $X_1, \cdots, X_n$ are jointly Gaussian then they are independent if they are uncorrelated.
- When $X_1, \cdots, X_n$ be jointly Gaussian (with zero means), there is an orthogonal transform $\mathbf{Y} = A\mathbf{X}$ such that $Y_1, \cdots, Y_n$ are jointly Gaussian and independent.
- $X_1, \cdots, X_n$ are jointly Gaussian if and only if $\mathbf{t}^T\mathbf{X}$ is Gaussian for for all non-zero $\mathbf{t} \in \Re^n$.
- If $X_1, \cdots, X_n$ are jointly Gaussian and A is a $k \times n$ matrix of rank k, then, $\mathbf{Y} = A\mathbf{X}$ is jointly gaussian

Recap: Moment inequalities

- **Jensen's Inequality**: If g is convex and EX and $E[g(X)]$ exist
$$g(EX) \leq E[g(X)]$$
- **Holder Inequality**: For $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$
$$E|XY| \leq (E|X|^p)^{\frac{1}{p}} (E|Y|^q)^{\frac{1}{q}}$$
(assuming all expectations exist)
- For $p = q = 2$, the above is Cauchy-Schwartz inequality
- This implies $|\rho_{XY}| \leq 1$
- **Minkowski's Inequality**:
$$(E|X+Y|^r)^{\frac{1}{r}} \leq (E|X|^r)^{\frac{1}{r}} + (E|Y|^r)^{\frac{1}{r}}$$

Recap

- **Chernoff Bounds**

$$P[X > a] \le \frac{E[e^{sX}]}{e^{sa}} = \frac{M_X(s)}{e^{sa}}, \ \forall s > 0$$

- **Hoeffding Inequality** X_i iid, $X_i \in [a,\ b]$, $\forall i$ and $EX_i = \mu$

$$P\left[\left|\sum_{i=1}^{n} X_i - n\mu\right| \ge \epsilon\right] \le 2e^{-\frac{2\epsilon^2}{n(b-a)}},\ \epsilon > 0$$

Recap: Weak Law of large numbers

- X_i are iid, $EX_i = \mu$, $\mathsf{Var}(X_i) = \sigma^2$, $S_n = \sum_{i=1}^n X_i$

$$E\left[\frac{S_n}{n}\right] = \mu; \quad \text{and} \quad \mathsf{Var}\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$$

- By Chebyshev Inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right] \leq \frac{\mathsf{Var}(\frac{S_n}{n})}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}, \quad \forall \epsilon > 0$$

$$\Rightarrow \quad \lim_{n\to\infty} P\left[\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right] = 0, \quad \forall \epsilon > 0$$

Recap: Convergence in Probability

- A sequence of random variables, X_n, is said to **converge in probability** to a random variable X_0 is

$$\lim_{n \to \infty} P\left[|X_n - X_0| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

  This is denoted as $X_n \xrightarrow{P} X_0$
- By the definition of limit, the above means

$$\forall \delta > 0, \ \exists N < \infty, \ s.t. \ P\left[|X_n - X_0| > \epsilon\right] < \delta, \ \forall n > N$$

- We only need marginal distributions of individual X_n to decide whether a sequence converges to a constant in probability

- We mentioned point-wise convergence of a sequence of functions

$$g_n \rightarrow g_0 \quad \text{if} \quad g_n(x) \rightarrow g_0(x), \;\; \forall x$$

- Since random variables are also functions we can define convergence like this.
- We can demand $X_n(\omega) \rightarrow X_0(\omega), \; \forall \omega$
- Such pointise convergence is too restrictive.
- But we can demand that it should be satisfied for almost all ω

- A sequence of random variables, X_n, is said to converge **almost surely** or **with probability one** to X if

$$P\left(\{\omega \ : \ X_n(\omega) \rightarrow X(\omega)\}\right) = 1$$

or equivalently

$$P\left(\{\omega \ : \ X_n(\omega) \nrightarrow X(\omega)\}\right) = 0$$

- Denoted as $X_n \stackrel{a.s.}{\rightarrow} X$ or $X_n \stackrel{w.p.1}{\rightarrow} X$ or $X_n \rightarrow X_0 \ \ (w.p.1)$
- We are saying that for 'almost all' ω, $X_n(\omega)$ converges to $X(\omega)$
- We will first try and write the event $\{\omega \ : \ X_n(\omega) \nrightarrow X(\omega)\}$ in a proper form

- Recall convergence of real number sequences.
- A sequence of real numbers x_n is said to converge to x_0, $x_n \to x_0$, if

$$\forall \epsilon > 0, \ \exists N < \infty, \ s.t. \ |x_n - x_0| < \epsilon, \ \forall n \geq N$$

  This is equivalent to

$$\forall \epsilon > 0, \ \exists N < \infty, \ \forall k \geq 0 \ |x_{N+k} - x_0| < \epsilon$$

- So, $x_n \nrightarrow x_0$ means

$$\exists \epsilon \ \ \forall N \ \ \exists k \ \ |x_{N+k} - x_0| \geq \epsilon$$

- Note that given any ω, $X_n(\omega)$ is real number sequence.
- Hence $X_n(\omega) \to X(\omega)$ is same as

$$\forall \epsilon > 0 \;\; \exists N < \infty \;\; \forall k \geq 0 \;\; |X_{N+k}(\omega) - X(\omega)| < \epsilon$$

  This is equivalent to

$$\forall r > 0, r \text{ integer} \;\;\; \exists N < \infty \;\; \forall k \geq 0 \; |X_{N+k}(\omega) - X(\omega)| < \frac{1}{r}$$

- Hence, $X_n(\omega) \nrightarrow X(\omega)$ is same as

$$\exists r \;\; \forall N \;\; \exists k \, |X_{N+k}(\omega) - X(\omega)| \geq \frac{1}{r}$$

- Hence we can write this event as

$$\cup_{r=1}^{\infty} \cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \left\{ \omega \; : \; |X_{N+k}(\omega) - X(\omega)| \geq \frac{1}{r} \right\}$$

- The event $\{\omega \ : \ X_n(\omega) \nrightarrow X(\omega)\}$ can be expressed as

$$\cup_{r=1}^{\infty} \cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \left[|X_{N+k} - X| \geq \frac{1}{r} \right]$$

- Hence X_n converges almost surely to X iff

$$P\left(\cup_{r=1}^{\infty} \cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \left[|X_{N+k} - X| \geq \frac{1}{r} \right] \right) = 0$$

- This is same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \left[|X_{N+k} - X| \geq \frac{1}{r} \right] \right) = 0, \ \ \forall r > 0, \text{ integer}$$

- Same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \left[|X_{N+k} - X| \geq \epsilon \right] \right) = 0, \ \ \forall \epsilon > 0$$

- A sequence X_n is said to converge **almost surely** or **with probability one** to X if

$$P\left(\{\omega \ : \ X_n(\omega) \rightarrow X(\omega)\}\right) = 1$$

- We can also write it as

$$P[X_n \rightarrow X] = 1$$

- We showed that this is equivalent to

$$P\left(\cap_{N=1}^{\infty} \cup_{k=0}^{\infty} \ [|X_{N+k} - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- Same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- let $A_k = [|X_k - X| \geq \epsilon]$
- Let $B_N = \cup_{k=N}^{\infty} A_k$.
- Then, $B_{N+1} \subset B_N$ and hence $B_N \downarrow$.
- Hence, $\lim B_N = \cap_{N=1}^{\infty} B_N$.
- We saw that $X_n \stackrel{a.s.}{\to} X$ is same as

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty}\ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

$$\Leftrightarrow \ \ P\left(\lim_{N\to\infty} \cup_{k=N}^{\infty}\ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

$$\Leftrightarrow \lim_{N\to\infty}\ P\left(\cup_{k=N}^{\infty}\ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- X_n converges to X almost surely iff

$$\lim_{n \to \infty} P\left(\cup_{k=n}^{\infty} [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- To show convergence with probability one using this one needs to know the joint distribution of $X_n, X_{n+1}, \cdots$
- Contrast this with $X_n \stackrel{P}{\to} X$ which is

$$\lim_{n \to \infty} P\left[|X_n - X| > \epsilon\right] = 0, \ \forall \epsilon > 0$$

- This also shows that

$$X_n \stackrel{a.s.}{\to} X \quad \Rightarrow \quad X_n \stackrel{P}{\to} X$$

- Almost sure convergence is a stronger mode of convergence

simple example: almost sure convergence

- Let $\Omega = [0,\ 1]$ with the usual probability measure and let $X_n = I_{[0,\ 1/n]}$.

$$X_n(\omega) = \begin{cases} 1 & \omega \leq \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

- Since $X_n \xrightarrow{P} 0$, zero is the only candidate for limit
- $X_n(\omega) = 1$ only when $n \leq 1/\omega$.
- Given any ω, for all $n > 1/\omega, \quad X_n(\omega) = 0$
- Hence, $\{\omega \ : \ X_n(\omega) \to 0\} = (0, 1]$

$$P[X_n \to X_0] = P(\{\omega \ : \ X_n(\omega) \to 0\}) = P((0, 1]) = 1$$

- Hence $X_n \xrightarrow{a.s} 0$

- X_n converges to X almost surely iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

$$\Leftrightarrow \ \ \lim_{n \to \infty} P\left(\cup_{k=n}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- We normally do not specify X_n as functions over Ω
- We are only given the distributions
- How do we establish convergence almost surely

- Let $A_1, A_2, \cdots$ be a sequence of events.
- How do we define limit of this sequence ?
- Define sequences

$$B_n = \cup_{k=n}^{\infty} A_k \qquad C_n = \cap_{k=n}^{\infty} A_k$$

- These are monotone: $B_n \downarrow$, $C_n \uparrow$. They have limits.
- Define

$$\begin{aligned} \lim\ \sup A_n &\triangleq \lim\ B_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \\ \lim\ \inf A_n &\triangleq \lim\ C_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \end{aligned}$$

- If $\lim\ \sup A_n = \lim\ \inf A_n$ then we define that as $\lim A_n$. Otherwise we say the sequence does not have a limit
- Note that $\lim\ \sup A_n$ and $\lim\ \inf A_n$ are events
- Note that $X_n \stackrel{a.s.}{\to} X$ iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

$$\Leftrightarrow \quad P\left(\lim\ \sup\ [|X_n - X| \geq \epsilon]\right) = 0, \quad \forall \epsilon > 0$$

- We can show $\lim\ \inf A_n \subset \lim\ \sup A_n$

$$\begin{aligned}
\omega \in \lim\ \inf A_n &\Rightarrow \omega \in \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \\
&\Rightarrow \exists m,\ \omega \in A_k,\ \forall k \geq m \\
&\Rightarrow \omega \in \cup_{j=n}^{\infty} A_j,\ \forall n \\
&\Rightarrow \omega \in \cap_{n=1}^{\infty} \cup_{j=n}^{\infty} A_j \\
&\Rightarrow \omega \in \lim\ \sup A_n
\end{aligned}$$

- We can characterize $\lim\ \inf A_n$ as follows

$$\begin{aligned} \omega \in \lim\ \inf A_n &\Rightarrow \omega \in \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k \\ &\Rightarrow \exists m,\ \omega \in A_k,\ \forall k \geq m \\ &\Rightarrow \omega \text{ belongs to all but finitely many of } A_n \end{aligned}$$

Thus, $\lim\ \inf A_n$ consists of all points that are there in all but finitely many A_n.

- We can characterize $\lim\ \sup A_n$ as follows

$$\begin{aligned} \omega \in \lim\ \sup A_n &\Rightarrow \omega \in \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \\ &\Rightarrow \omega \in \cup_{k=n}^{\infty} A_k, \ \forall n \\ &\Rightarrow \omega \text{ belongs to infinitely many of } A_n \end{aligned}$$

  Thus $\lim\ \sup A_n$ consists of points that are in infinitely many A_n
  One refers to $\lim\ \sup A_n$ also as 'A_n infinitely often' or 'A_n $i.o.$'
- What is the difference between
  Points that belong to all but finitely many A_n and
  Points that belong to infinitely many A_n
- There can be ω that are there in infinitely many of A_n and are also not there in infinitely many of the A_n

Example

- Consider the following sequence of sets: $A, B, A, B, \cdots$
- Recall

$$\lim\ \sup A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \quad \lim\ \inf A_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k$$

$$\cup_{k=n}^{\infty} A_k = A \cup B,\ \forall n \quad \Rightarrow \quad \lim\ \sup A_n = A \cup B$$

$$\cap_{k=n}^{\infty} A_k = A \cap B,\ \forall n \quad \Rightarrow \quad \lim\ \inf A_n = A \cap B$$

example

- Consider the sets $A_n = [0,\ 1 + \frac{(-1)^n}{n})$
  The sequence is

$$[0,\ 0),\ \left[0,\ 1+\frac{1}{2}\right),\ \left[0,\ 1-\frac{1}{3}\right),\ \left[0,\ 1+\frac{1}{4}\right)\cdots$$

- Guess: $\limsup A_n = [0,\ 1]$ and $\liminf A_n = [0,\ 1)$
- First note that $[0,\ 1+\frac{1}{n+1}) \subset \cup_{k=n}^{\infty} A_k \subset [0,\ 1+\frac{1}{n})$.
  Hence

$$x \in [0,\ 1]\ \Rightarrow x \in \cup_{k=n}^{\infty} A_k,\ \forall n\ \Rightarrow x \in \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\ \Rightarrow x \in \lim\ \sup A_n$$

- Given any $\epsilon > 0$, $1+\epsilon \notin [0,\ 1+\frac{1}{n})$ if $\epsilon > \frac{1}{n}$ or $n > \frac{1}{\epsilon}$.
- Hence, given any $\epsilon > 0$, $\exists n$ such that $1+\epsilon \notin \cup_{k=n}^{\infty} A_k$.
- This proves $\limsup A_n = [0,\ 1]$

- Now let us consider: $\lim\ \inf A_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k$.
- Recall $A_n = [0,\ 1 + \frac{(-1)^n}{n})$
- First note that $[0,\ 1 - \frac{1}{n}) \subset \cap_{k=n}^{\infty} A_k \subset [0,\ 1 - \frac{1}{n+1})$
- Given any $\epsilon > 0$, $1 - \epsilon < 1 - \frac{1}{n}$ if $n > \frac{1}{\epsilon}$
- Hence, given any $\epsilon > 0$, $\exists n$ such that $1 - \epsilon \in \cap_{k=n}^{\infty} A_k$
- Hence $1 - \epsilon \in \lim\ \inf A_n$
- It is easy to see $1 \notin \cap_{k=n}^{\infty} A_k$ for ay n.
- Hence $1 \notin \lim\ \inf A_n$
- This proves $\lim\ \inf A_n = [0,\ 1)$
- Since $\lim\ \sup A_n \neq \lim\ \inf A_n$, this sequence does not have a limit

- $X_n \stackrel{a.s.}{\to} X$ iff

$$P\left(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} \ [|X_k - X| \geq \epsilon]\right) = 0, \ \ \forall \epsilon > 0$$

- Let $A_n^{\epsilon} = [|X_n - X| \geq \epsilon]$
- Then $X_n \stackrel{a.s.}{\to} X$ iff

$$P\left(\lim\ \sup\ A_n^{\epsilon}\right) = 0, \ \forall \epsilon > 0$$

- The question now is: can we get probability of $\lim\ \sup A_n$
- We look at an important result that allows us to do this

Borel-Cantelli Lemma

- **Borel-Cantelli lemma**: Given sequence of events, $A_1, A_2, \cdots$
  1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\lim\ \sup A_n) = 0$
  2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and A_i are independent, $P(\lim\ \sup A_n) = 1$

  Proof:
- We will first show: $P(\cup_{i=n}^{\infty} A_i) \leq \sum_{i=n}^{\infty} P(A_i),\ \forall n$
- We have the result: $P(\cup_{i=n}^{N} A_i) \leq \sum_{i=n}^{N} P(A_i), n \leq N$
- For any n, let $B_N = \cup_{i=n}^{N} A_i$. Then $B_N \subset B_{N+1}$.
- $\lim_{N\to\infty} B_N = \cup_{k=n}^{\infty} A_k$

$$
\begin{aligned}
P(\cup_{i=n}^{\infty} A_i) = P(\lim_{N\to\infty} \cup_{i=n}^{N} A_i) &= \lim_{N\to\infty} P(\cup_{i=n}^{N} A_i) \\
&\leq \lim_{N\to\infty} \sum_{i=n}^{N} P(A_i) = \sum_{i=n}^{\infty} P(A_i)
\end{aligned}
$$

- If $\sum_{k=1}^{\infty} P(A_k) < \infty$, then, $\lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0$

$$
\begin{aligned}
0 \le P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&\le \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) \\
&= 0, \quad \text{if} \quad \sum_{k=1}^{\infty} P(A_k) < \infty
\end{aligned}
$$

- This completes proof of first part of Borel-Cantelli lemma

- Let $\sum_{k=1}^{\infty} P(A_k) = C < \infty$
- It means given any $\epsilon > 0$, $\exists n$

$$\left| \sum_{k=1}^{n} P(A_k) - C \right| < \epsilon \;\; \Rightarrow \;\; \left| \sum_{k=n}^{\infty} P(A_k) \right| < \epsilon$$

- This implies

$$\lim_{n \to \infty} \sum_{k=n}^{\infty} P(A_k) = 0$$

- For the second part of the lemma:

$$
\begin{aligned}
P\left(\lim\ \sup A_n\right) &= P\left(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= P\left(\lim_{n\to\infty} \cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} P\left(\cup_{k=n}^{\infty} A_k\right) \\
&= \lim_{n\to\infty} \left(1 - P\left(\cap_{k=n}^{\infty} A_k^c\right)\ \right) \\
&= \lim_{n\to\infty} \left(1 - \prod_{k=n}^{\infty} (1 - P(A_k))\ \right) \\
&\quad \text{because } A_k \text{ are independent} \\
&= 1\ -\ \lim_{n\to\infty} \prod_{k=n}^{\infty} (1 - P(A_k))
\end{aligned}
$$

- We can compute that limit as follows

$$\begin{aligned}\lim_{n\to\infty}\prod_{k=n}^{\infty}(1-P(A_k)) &\le \lim_{n\to\infty}\prod_{k=n}^{\infty}e^{-P(A_k)}, \quad \text{since } \ 1-x\le e^{-x}\\ &= \lim_{n\to\infty}\ e^{-\sum_{k=n}^{\infty}P(A_k)}\\ &= 0\end{aligned}$$

because
$\sum_{k=1}^{\infty}P(A_k)=\infty \quad \Rightarrow \quad \lim_{n\to\infty}\sum_{k=n}^{\infty}P(A_k)=\infty$

- This finally gives us

$$P\left(\lim\ \sup A_n\right)=1\ -\ \lim_{n\to\infty}\prod_{k=n}^{\infty}(1-P(A_k))=1$$

- Given a sequence X_n we want to know whether it converges to X
- Let $A_k^\epsilon = [|X_k - X| \geq \epsilon]$
- $X_n \xrightarrow{P} X$ if

$$\lim_{k\to\infty} P[|X_k - X| \geq \epsilon] = 0 \quad \text{same as} \quad \lim_{k\to\infty} P(A_k) = 0, \;\; \forall\epsilon > 0$$

- By Borel-Cantelli lemma

$$\sum_{k=1}^{\infty} P(A_k) < \infty \;\; \Rightarrow \;\; P(\lim\ \sup A_k) = 0 \;\; \Rightarrow \;\; X_k \xrightarrow{a.s.} X$$

- Consider a sequence X_n with

$$P[X_n = 0] = 1 - a_n; \quad P[X_n = c_n] = a_n$$

- We want to investigate convergence to 0.
- $A_n^\epsilon = [|X_n - 0| > \epsilon] = [X_n = c_n], \ \forall \epsilon > 0$
- Hence $P(A_n^\epsilon) = a_n, \ \forall \epsilon > 0$.
- If $a_n \to 0$ then $X_n \xrightarrow{P} 0$. (e.g., $a_n = \frac{1}{n}, \frac{1}{n^2}$)
- If $\sum a_n < \infty$, $X_n \xrightarrow{a.s.} 0$ (e.g., $a_n = \frac{1}{n^2}$)

- Consider a sequence X_n with

$$P[X_n = 0] = 1 - \frac{1}{n}; \quad P[X_n = 1] = \frac{1}{n}$$

- We can easily conclude $X_n \xrightarrow{P} 0$.
- But since, $\sum_n \frac{1}{n} = \infty$, Borel-Cantelli lemma is not really useful here
- We saw one example where such X_n converge almost surely.
- But, if X_n are independent, then by Borel-Cantelli lemma, they do not converge
- Convergence (to a constant) in probability depends only on distribution of individual X_n.
- Convergence almost surely depends on the joint distribution

Strong Law of Large Numbers

- Let X_n be iid, $EX_n = \mu$, $\text{Var}(X_n) = \sigma^2$, $S_n = \sum_{i=1}^{n} X_i$
- We saw weak law of large numbers:

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

- Strong law of large numbers says:

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu$$

- Let $A_n^\epsilon = \left[\left| \frac{S_n}{n} - \mu \right| > \epsilon \right]$
- As we saw, by Chebyshev inequality

$$P\left[\left| \frac{S_n}{n} - \mu \right| > \epsilon \right] \le \frac{\sigma^2}{n\epsilon^2}$$

- This shows $P(A_n^\epsilon) \to 0$ and thus we get weak law
- To prove strong law using Borel-Cantelli lemma, we need $\sum P(A_n^\epsilon) < \infty$
- Since $\sum_n \frac{\sigma^2}{n\epsilon^2} = \infty$, the Chebyshev bound is not useful
- We need a bound: $P[|\frac{S_n}{n} - \mu|] \le c_n$ such that $\sum_n c_n < \infty$.