

Recap: Multi-dimensional Gaussian density

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

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

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

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

Recap: Weak Law of large numbers

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

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

- ▶ By Chebyshev Inequality

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

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

Recap: Convergence in Probability

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

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

This is denoted as $X_n \xrightarrow{P} X_0$

- ▶ By the definition of limit, the above means

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

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

Example

- ▶ Let $\Omega = [0, 1]$ with the usual probability measure and let $X_n = I_{[0, 1/n]}$.
- ▶ $P[X_n = 1] = \frac{1}{n} = 1 - P[X_n = 0]$
- ▶ The probability of X_n taking value 1 is decreasing with n
- ▶ A good guess is that it converges to zero

$$P[|X_n - 0| > \epsilon] = P[X_n = 1] = \frac{1}{n}$$

which goes to zero as $n \rightarrow \infty$.

- ▶ Hence, $X_n \xrightarrow{P} 0$

Example

- ▶ Let X_1, X_2, \dots be a sequence of iid random variable which are uniform over $(0, 1)$.
- ▶ Let $M_n = \max(X_1, X_2, \dots, X_n)$
- ▶ Does M_n converge in probability?
- ▶ A reasonable guess for the limit is 1

$$P[|M_n - 1| \geq \epsilon] = P[M_n \leq 1 - \epsilon] = (1 - \epsilon)^n$$

- ▶ This implies $M_n \xrightarrow{P} 1$
- ▶ Suppose $Z_n = \min(X_1, X_2, \dots, X_n)$.
Then $Z_n \xrightarrow{P} 0$

Some properties of convergence in probability

- ▶ $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y \Rightarrow P[X = Y] = 1$
- ▶ $X_n \xrightarrow{P} X \Rightarrow P[|X_n - X_m| > \epsilon] \rightarrow 0$ as $n, m \rightarrow \infty$
- ▶ Suppose $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$ Then the following hold
 1. $aX_n \xrightarrow{P} aX$
 2. $X_n + Y_n \xrightarrow{P} X + Y$
 3. $X_n Y_n \xrightarrow{P} XY$
 4. $g(X_n) \xrightarrow{P} g(X)$ where g is a continuous function from \Re to \Re .

► $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y \Rightarrow P[X = Y] = 1$

► Proof: We have

$$|X - Y| = |X - X_n + X_n - Y| \leq |X - X_n| + |Y - X_n|.$$

$$\Rightarrow \left(\left[|X - X_n| \leq \frac{c}{2} \right] \cap \left[|Y - X_n| \leq \frac{c}{2} \right] \right) \subset [|X - Y| \leq c]$$

$$\Rightarrow [|X - Y| > c] \subset \left(\left[|X - X_n| > \frac{c}{2} \right] \cup \left[|Y - X_n| > \frac{c}{2} \right] \right)$$

$$\Rightarrow P[|X - Y| > c] \leq P\left[|X - X_n| > \frac{c}{2}\right] + P\left[|Y - X_n| > \frac{c}{2}\right], \forall n$$

► Hence, we can conclude $P[|X - Y| > c] = 0, \forall c > 0$ and hence $P[X = Y] = 1$.

► $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X + Y$

► We can prove this in the same manner

$$|X_n + Y_n - (X + Y)| = |(X_n - X) + (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$$

► All the other properties listed can be proved by following similar arguments

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

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

- ▶ Since random variables are also functions we can define convergence like this.
- ▶ We can demand $X_n(\omega) \rightarrow X_0(\omega), \quad \forall \omega$
- ▶ Such pointwise 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

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

or equivalently

$$P(\{\omega : X_n(\omega) \not\rightarrow X(\omega)\}) = 0$$

- ▶ Denoted as
 $X_n \xrightarrow{a.s.} X$ or $X_n \xrightarrow{w.p.1} X$ or $X_n \rightarrow X_0$ (*w.p.1*)
- ▶ We are saying that for 'almost all' ω , $X_n(\omega)$ converges to $X(\omega)$
- ▶ Given a sequence X_n , how does one establish almost sure convergence?

- ▶ X_n converges to X **almost surely** or **with probability one** if

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

- ▶ By notation, this is same as

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

- ▶ Next we will try to characterize the event
 $\{\omega : X_n(\omega) \rightarrow X(\omega)\}$
or the event
 $\{\omega : X_n(\omega) \nrightarrow X(\omega)\}$

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

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

This is equivalent to

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

- ▶ So, $x_n \not\rightarrow x_0$ means

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

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

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

This is equivalent to

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

- ▶ Hence, $X_n(\omega) \not\rightarrow X(\omega)$ is same as

$$\exists r \quad \forall N \quad \exists k \quad |X_{N+k}(\omega) - X(\omega)| > \frac{1}{r}$$

- ▶ Hence we can write this event as

$$\bigcup_{r=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{k=0}^{\infty} \left\{ \omega : |X_{N+k}(\omega) - X(\omega)| > \frac{1}{r} \right\}$$

- The event $\{\omega : X_n(\omega) \not\rightarrow X(\omega)\}$ can be expressed as

$$\bigcup_{r=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{k=0}^{\infty} \left[|X_{N+k} - X| > \frac{1}{r} \right]$$

- Hence X_n converges almost surely to X iff

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

- This is same as

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

- Same as

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

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

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

- ▶ We can also write it as

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

- ▶ We showed that this is equivalent to

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

- ▶ Same as

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

- ▶ 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 $X_n \xrightarrow{a.s.} X$ is same as

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

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

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

- ▶ X_n converges to X almost surely iff

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

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

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

- ▶ This also shows that

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

- ▶ 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 $n > 1/\omega$, $X_n(\omega) = 0$
- ▶ Hence, $\{\omega : X_n(\omega) \rightarrow 0\} = (0, 1]$

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

- ▶ Hence $X_n \xrightarrow{a.s} 0$

- ▶ X_n converges to X almost surely iff

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

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

- ▶ We normally do not specify X_n as functions over Ω
- ▶ We are only given the distributions
- ▶ How do we establish convergence almost surely?
- ▶ We will first look at limits of general sequences of events.

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

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

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

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

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

$$\begin{aligned} P(\cap_{N=1}^{\infty} \cup_{k=N}^{\infty} [|X_k - X| \geq \epsilon]) &= 0, \quad \forall \epsilon > 0 \\ \Leftrightarrow P(\limsup [|X_n - X| \geq \epsilon]) &= 0, \quad \forall \epsilon > 0 \end{aligned}$$

► We can show $\liminf A_n \subset \limsup A_n$

$$\begin{aligned}\omega \in \liminf A_n &\Rightarrow \omega \in \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \\ &\Rightarrow \exists m, \omega \in A_k, \forall k \geq m \\ &\Rightarrow \omega \in \bigcup_{j=n}^{\infty} A_j, \forall n \\ &\Rightarrow \omega \in \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j \\ &\Rightarrow \omega \in \limsup A_n\end{aligned}$$

- We can characterize $\liminf A_n$ as follows

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

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

- ▶ We can characterize $\limsup A_n$ as follows

$$\begin{aligned}\omega \in \limsup A_n &\Rightarrow \omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \\ &\Rightarrow \omega \in \bigcup_{k=n}^{\infty} A_k, \forall n \\ &\Rightarrow \omega \text{ belongs to infinitely many of } A_n\end{aligned}$$

Thus $\limsup A_n$ consists of points that are in infinitely many A_n

One refers to $\limsup A_n$ 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, \dots
- ▶ Recall

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad \liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

$$\bigcup_{k=n}^{\infty} A_k = A \cup B, \forall n \Rightarrow \limsup A_n = A \cup B$$

$$\bigcap_{k=n}^{\infty} A_k = A \cap B, \forall n \Rightarrow \liminf A_n = A \cap B$$

example

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

The sequence is

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

- ▶ 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

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

- ▶ 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: $\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$.
- ▶ Recall $A_n = [0, 1 + \frac{(-1)^n}{n})$
- ▶ First note that $[0, 1 - \frac{1}{n}) \subset \bigcap_{k=n}^{\infty} A_k \subset [0, 1 - \frac{1}{n+1})$
- ▶ Given any $\epsilon > 0$, $1 - \epsilon < 1 - \frac{1}{n}$ if $n > \frac{1}{\epsilon}$
- ▶ Hence, given any $\epsilon > 0$, $\exists n$ such that $1 - \epsilon \in \bigcap_{k=n}^{\infty} A_k$
- ▶ Hence $1 - \epsilon \in \liminf A_n$
- ▶ It is easy to see $1 \notin \bigcap_{k=n}^{\infty} A_k$ for any n .
- ▶ Hence $1 \notin \liminf A_n$
- ▶ This proves $\liminf A_n = [0, 1)$
- ▶ Since $\limsup A_n \neq \liminf A_n$, this sequence does not have a limit

- ▶ $X_n \xrightarrow{a.s.} X$ iff

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

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

$$P\left(\limsup A_n^\epsilon\right) = 0, \quad \forall \epsilon > 0$$

- ▶ The question now is: can we get probability of $\limsup A_n$
- ▶ 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, \dots

1. If $\sum_{i=1}^{\infty} P(A_i) < \infty$, then, $P(\limsup A_n) = 0$
2. If $\sum_{i=1}^{\infty} P(A_i) = \infty$ and A_i are independent, $P(\limsup A_n) = 1$

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 $B_N \subset B_{N+1}$.
- $\lim_{N \rightarrow \infty} B_N = \cup_{k=n}^{\infty} A_k$

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

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

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

► 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

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

► For the second part of the lemma:

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

- We can compute that limit as follows

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

because

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

- This finally gives us

$$P(\limsup A_n) = 1 - \lim_{n \rightarrow \infty} \prod_{k=n}^{\infty} (1 - P(A_k)) = 1$$

- ▶ 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

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

- ▶ By Borel-Cantelli lemma

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

- 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 $a_n \rightarrow 0$ 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