$T = \cap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots)$ is the tail $\sigma$-field associated with the sequence $X_1, X_2, \ldots$

**Theorem:** If $\{X_n\}$ is a sequence of independent random variables, then $P(A) = 0$ or $1$ for $A \in T$.

**Proof:**

$\sigma(X_1, \ldots, X_k) \& \sigma(X_n, X_{n+1}, \ldots)$ are independent for $n > k$.

As $T \subset \sigma(X_n, X_{n+1}, \ldots)$ for all $n$, $\sigma(X_1, \ldots, X_k)$ and $T$ are independent for all $k$.

$\cup_{k=1}^{\infty} \sigma(X_1, \ldots, X_k)$ and $T$ are independent.

$\sigma(X_1, X_2, \ldots) = \sigma(\cup_{k=1}^{\infty} \sigma(X_1, \ldots, X_k))$ and $T$ are independent.

Thus $T$ is independent of itself.

---

**Kolmogorov’s Zero-One Law**

If $X_1, \ldots, X_n$ are independent with finite mean, then for $a > 0$,

$$1 - \frac{(a + 2c)^2}{\text{Var}(S_n)} \leq P \left( \max_{1 \leq k \leq n} |S_k - E(S_k)| \geq a \right) \leq a^{-2} \text{Var}(S_n).$$

The left side inequality holds if $P(|X_i| \leq c) = 1; S_k = X_1 + \cdots + X_k$.

**Proof:** Without loss of generality assume that $E(X_k) = 0$.

Define $A_k = \{|S_k| \geq a, |S_j| < a \text{ for } j < k\}$.

$$S_n^2 I_{A_k} = S_k^2 I_{A_k} + 2S_k I_{A_k}(S_n - S_k) + (S_n - S_k)^2 I_{A_k}$$

$$\geq S_k^2 I_{A_k} + 2S_k I_{A_k}(S_n - S_k)$$

$$\geq a^2 I_{A_k} + 2S_k I_{A_k}(S_n - S_k)$$

$$E(S_n^2 I_{A_k}) \geq a^2 E(I_{A_k}) = a^2 P(A_k), \quad \text{as}$$

$$E(S_k I_{A_k}(S_n - S_k)) = E(S_k I_{A_k})E(S_n - S_k) = 0$$

Note that $S_n - S_k$ and $S_k I_{A_k}$ are independent.

So

$$E(S_n^2) \geq a^2 \sum_{k=1}^{n} P(A_k) = a^2 P \left( \max_{1 \leq k \leq n} |S_k| \geq a \right).$$
Note that \(|X| \leq c\) implies \(E(|X|) \leq c\) and hence \(|X - E(X)| \leq 2c\).
So we continue to assume \(E(X_i) = 0, S_0 = 0\).
Put \(B_0 = \Omega, B_k = \{\max_{1 \leq j \leq k} |S_j| < a\}\).
Then \(A_k = B_{k-1} - B_k \subset \{|S_k| \geq a, |S_{k-1}| < a\}\). Note
\[
S_{k-1}I_{B_{k-1}} + X_kI_{B_{k-1}} = S_kI_{B_k} + S_kI_{A_k}.
\]
As \(S_{k-1}I_{B_{k-1}}\) and \(X_k\) are independent; and \(A_k \cap B_k = \emptyset\),
\[
E(S_{k-1}I_{B_{k-1}})^2 + \text{Var}(X_k)P(B_{k-1}) = E(S_kI_{B_k})^2 + E(S_kI_{A_k})^2.
\]
As \(|S_kI_{A_k}| \leq |S_{k-1}I_{A_k}| + |X_kI_{A_k}| \leq (a + 2c)I_{A_k}\), \(P(B_{k-1}) \geq P(B_n)\),
\[
E(S_{k-1}I_{B_{k-1}})^2 + \text{Var}(X_k)P(B_n) \leq E(S_kI_{B_k})^2 + (a + 2c)^2P(A_k).
\]
Summing over \(k = 1, \ldots n\), we get
\[
\text{Var}(S_n)P(B_n) \leq E(S_nI_{B_n})^2 + (a + 2c)^2 \sum_{k=1}^nP(A_k)
\]
\[
\leq a^2P(B_n) + (a + 2c)^2P(B_n^c) \leq (a + 2c)^2.
\]
This leads to the left side inequality.

**Maximal Inequality**

What happens when we do not have information on mean or variance?

**Reflection Principle:**
If \(X_1, \ldots, X_n\) i.i.d., \(P(X_i = 1) = P(X_i = -1) = \frac{1}{2}\), then
\[
P\left(\max_{1 \leq k \leq n} S_k \geq a\right) = 2P\left(\max_{1 \leq k \leq n} S_k \geq a, S_n > a\right) + P(S_n = a) \leq 2P(S_n \geq a).
\]

**Theorem:** If \(X_1, \ldots, X_n\) are independent r.v., then for \(a > 0\),
\[
P\left(\max_{1 \leq k \leq n} |S_k| \geq 3a\right) \leq 3 \max_{1 \leq k \leq n} P(|S_k| \geq a).
\]

**Proof:** Let \(A_k = \{|S_k| \geq 3a, |S_j| < 3a\text{ for }j < k\}\).
\[
P\left(\max_{1 \leq k \leq n} |S_k| \geq 3a\right) \leq P(|S_n| \geq a) + \sum_{k=1}^{n-1} P(A_k \cap [|S_n| < a])
\]
\[
P(A_k \cap [|S_n| < a]) \leq P(A_k \cap [|S_n - S_k| > 2a])
\]
\[
\leq P(A_k)P(|S_n - S_k| > 2a) \quad \text{(independence)}
\]
\[
\leq P(A_k) \left(P(|S_n| > a) + P(|S_k| > a)\right)
\]
\[
\leq 2P(A_k) \max_{1 \leq j \leq n} P(|S_j| > a).
\]
The result follows as \(A_k\) are disjoint.
**Theorem 22.7**: Let \( \{X_n\} \) be a sequence of independent r.v. Then \( S_n \to S \) a.e. if and only if \( S_n \to_p S \).

**Proof**: If \( S_n \to_p S \), then by the maximal inequality,
\[
P\left( \max_{1 \leq j \leq k} |S_{n+j} - S_n| \geq 6a \right) \leq 3 \max_{1 \leq j \leq k} P(|S_{n+j} - S_n| \geq 2a)
\]
\[
\leq 6 \max_{0 \leq j \leq k} P(|S_{n+j} - S| \geq a),
\]
and hence
\[
P\left( \max_{j \geq 1} |S_{n+j} - S_n| \geq 6a \right) \leq 6 \max_{j \geq 0} P(|S_{n+j} - S| \geq a) \to 0.
\]

So \( S_n - S_m \to 0 \) (Cauchy) a.e.

**Theorem 22.6**: Let \( \{X_n\} \) be a sequence of independent r.v. with \( E(X_n) = 0 \). If \( \sum_n \text{Var}(X_n) < \infty \), then \( \sum_n X_n \) converges a.e.

**Proof**: Use Kolmogorov’s inequality for \( X_{n+1}, \ldots, X_{n+r} \), to conclude that \( X_{n+1} + \cdots + X_{n+r} \to 0 \) a.e.

(You may also use that \( S_n - S_m \to_p 0 \) implies \( S_n - S_m \to 0 \) a.e.)

**Kolmogorov’s three-series criterion**

**Theorem 22.8**: Let \( \{X_n\} \) be a sequence of independent r.v. and \( X_n^c = X_n I_{\{|X_n| \leq c\}} \). If \( \sum_n X_n \) converges a.e., then for all \( c > 0 \),

i) \( \sum_n P(|X_n| > c) < \infty \),

ii) \( \sum_n \text{Var}(X_n^c) < \infty \),

iii) \( \sum_n E(X_n^c) \) converges.

Conversely, if (i-iii) hold for some \( c > 0 \), then \( \sum_n X_n \) converges a.e.

**Proof**: If (ii) and (iii) hold, then by Theorem 22.6,
\( \sum_n X_n^c \) converges a.e.

Now by (i) and Borel-Cantelli Lemma, \( \sum_n X_n \) converges a.e.
**Lemma:** If $|Y_n| \leq c$ and $\sum_n Y_n$ converges a.e., then $\sum_{n=1}^{\infty} \text{Var}(Y_n) < \infty$.

We use this lemma to prove the converse of Kolmogorov’s three-series criterion.

First assume $|X_n| \leq c$ and $\sum_n X_n$ converges a.e.

By the Lemma, $\sum_n \text{Var}(X_n) < \infty$.

Thus by Theorem 22.6, $\sum_n (X_n - E(X_n))$ converges a.e.

So $\sum_n E(X_n)$ converges.

In the general case, a.e. convergence of $\sum_n X_n$ implies $X_n \to 0$ a.e.

So (i) holds by the second Borel-Cantelli Lemma.

The rest follows.

**Proof of the Lemma.** We use *Symmetrization*:

Let $Y'_i$, $Y''_i$; $i, j \geq 1$, be independent r.v. defined on the same probability space $(\Omega^0, \mathcal{F}^0, P^0)$ such that $Y_n$, $Y'_n$, $Y''_n$ have the same distribution.

The symmetric r.v. $Y^s_n = Y'_n - Y''_n$ satisfy $|Y^s_n| \leq 2c$, $\text{Var}(Y^s_n) = 2\text{Var}(Y_n)$, $E(Y^s_n) = 0$.

Let $S_n = \sum_{i=1}^{n} Y_i$, $S'_n = \sum_{i=1}^{n} Y'_i$, $S''_n = \sum_{i=1}^{n} Y''_i$, and $S^s_n = \sum_{i=1}^{n} Y^s_i$.

As $\sum_n Y_n$ converges to $S$ a.e., $P(\cap_{m=1}^{\infty} \cup_{n\geq m} (|S_n - S| \geq \frac{a}{4})) = 0$, & hence

$$P^0(\max_{n \geq m} |S^s_n - S^s_m| \geq a) \leq P^0(\max_{n \geq m} |S'_n - S'_m| \geq \frac{a}{2}) + P^0(\max_{n \geq m} |S''_n - S''_m| \geq \frac{a}{2})$$

$$\leq 2P(\max_{n \geq m} |S_n - S_m| \geq \frac{a}{2})$$

$$\leq 4P(\max_{n \geq m} |S_n - S| \geq \frac{a}{4}) \to 0 \text{ as } n \to \infty.$$  

So by Kolmogorov’s inequality,

$$1 - \frac{(a + 4c)^2}{\text{Var}(S^s_n - S^s_m)} \leq P^0(\max_{n \geq m} |S^s_n - S^s_m| \geq a) \leq \frac{1}{2}$$

for all large $m \leq n$, and hence, $\text{Var}(X^s_m) + \cdots + \text{Var}(X^s_n) \leq 2(a + 4c)^2$.

So $\sum_n \text{Var}(X_n) < \infty.$
Let \( \{X_n\} \) be a sequence of i.i.d. r.v. with finite mean. Put \( S_n = X_1 + \cdots + X_n \). Suppose that \( \tau \) is a stopping time: \( \tau \) is a positive integer valued r.v., and \([\tau = n] \in \sigma(X_1, \ldots, X_n)\). Then \( E(\tau) < \infty \) implies \( E(S_\tau) = E(X_1)E(\tau) \).

Without loss of generality assume that \( X_i \) are non-negative. As \([\tau \geq k]^c = [\tau \leq k - 1]\) and \( X_k \) are independent,

\[
E(S_\tau) = E\left( \sum_k X_k I_{[\tau \geq k]} \right) = \sum_k P([\tau \geq k]) E(X_k) = \sum_k P([\tau \geq k]) E(X_1) = E(X_1)E(\tau).
\]