

Homework02 - MATH 734

Boren(Wells) Guan

Date: February 28, 2024

Notation

Here I use $X \wedge Y$ for $\min(X, Y)$ and $X \vee Y$ for $\max(X, Y)$. r.v. for random variable.

Before Reading:

To make the proof more readable, I will miss or gap some natural or not important facts or notations during my writing. If you feel it hard to see, you can refer the appendix after the proof, where I will try to explain some simple conclusions (will be marked) more clearly. In case that you misunderstand the mark, I will add the mark just after those formulas between \$ and before those between \$\$.

And I have to claim that the appendix is of course a part of my assignment, so the reference of it is required. Enjoy your grading!

Ex.1

Let $(X_n)_{n \geq 0}$ be a stochastic process defined on a measurable space (X, \mathcal{F}) adapted to a filtration $(\mathcal{F}_n)_{n \geq 0}$. Let A be a measurable subset of the state space. For each $k \geq 1$, let $T_A^{(k)}$ denote the k^{th} time that the process X_n visits some state in A . That is

$$T_A^{(m)} = \begin{cases} \inf \{n > T_A^{(m-1)}, X_n \in A\} & T_A^{(m-1)} < \infty \\ \infty & T_A^{(m-1)} = \infty \end{cases}$$

Show that $T_A^{(k)}$ is a stopping time for all $k \geq 1$.

Sol.

We know

$$\{T_A^{(m)} \leq n\} = \bigcup_{k=0}^{\infty} \{\inf \{p > k, X_p \in A\} \leq n, T_A^{(m-1)} = k\} = \bigcup_{k=0}^{n-1} \left(\bigcup_{i=k+1}^n \{X_i \in A\} \cap \{T_A^{(m-1)} = k\} \right)$$

and hence $T_A^{(m-1)}$ is a stopping time will imply that $T_A^{(m)}$ is a stopping time.

Then notice

$$T_A^{(1)} = \inf \{n \geq 0, X_n \in A\}$$

and hence

$$\{T_A^{(1)} \leq n\} = \{\inf \{k \geq 0, X_k \in A\} \leq n\} = \bigcup_{k=0}^n \{X_k \in A\} \in \mathcal{F}_n$$

which means $T_A^{(1)}$ is a stopping time.

Therefore, $T_A^{(m)}$ is a stopping time for any $m \geq 1$ by the inducton. \square

Ex.2

Let $X = (X_n)_{n \geq 0}$ be a supermartingale w.r.t. a filtration \mathcal{F}_n and let $H = (H_n)_{n \geq 1}$ be any predictable sequence w.r.t. $(\mathcal{F}_n)_{n \geq 1}$. Suppose that H_n is bounded and nonnegative for $n \geq 1$. Show that $\int_0^n H dX$ is a supermartingale w.r.t. \mathcal{F}_n . Also show the similar results for submartingales and martingales.

Sol.

Notice that $E|\int_0^n H dX| \leq 2M_n \sum_{i=0}^n E|X_i| < \infty$ where $\sup\{|H_i|, 1 \leq i \leq n\} \leq M_n$,

$$E\left(\int_0^{n+1} H dX \middle| \mathcal{F}_n\right) = E\left(\sum_{k=1}^{n+1} H_k(X_k - X_{k-1}) \middle| \mathcal{F}_n\right) = H_{n+1}E(X_{n+1} - X_n | \mathcal{F}_n) + \int_0^n H dX \leq \int_0^n H dX$$

which means $\int_0^n H dX$ is a supermartingale.

For submartingales, we claim that $\int_0^n H dX$ is a submartingale w.r.t. \mathcal{F}_n since if X_n is a submartingale, then

$$E\left(\int_0^{n+1} H dX \middle| \mathcal{F}_n\right) = E\left(\sum_{k=1}^{n+1} H_k(X_k - X_{k-1}) \middle| \mathcal{F}_n\right) = H_{n+1}E(X_{n+1} - X_n | \mathcal{F}_n) + \int_0^n H dX \geq \int_0^n H dX$$

and naturally $\int_0^n H dX$ will be a martingale even if H_n may be nonnegative because

$$E(X_{n+1} - X_n | \mathcal{F}_n) = 0$$

for $n \geq 0$ and then we will have

$$E\left(\int_0^{n+1} H dX \middle| \mathcal{F}_n\right) = E\left(\sum_{k=1}^{n+1} H_k(X_k - X_{k-1}) \middle| \mathcal{F}_n\right) = H_{n+1}E(X_{n+1} - X_n | \mathcal{F}_n) + \int_0^n H dX = \int_0^n H dX$$

and hence $\int_0^n H dX$ is a martingale. \square

Ex.3

Find an instance of martingale converging in probability but not almost surely.

Sol.

Let $X_0 = 0$ and $P(X_k = 1 | X_{k-1} = 0) = P(X_k = -1 | X_{k-1} = 0) = \frac{1}{2k}$, $P(X_k = 0 | X_{k-1} = 0) = 1 - \frac{1}{k}$ and $P(X_k = kX_{k-1} | X_{k-1} \neq 0) = \frac{1}{k}$, $P(X_k = 0 | X_{k-1} \neq 0) = 1 - \frac{1}{k}$, then we know $X_k \rightarrow 0$ in probability.

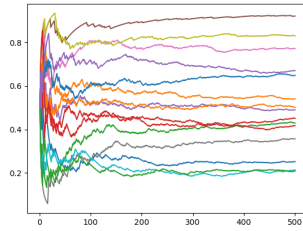
However, $P(X_k = 0, k \geq K)$ and it picks discrete values and hence X_k can not converge to 0 a.s. \square

Ex.4

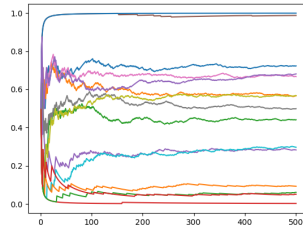
Use your favorite programming language (e.g., python, R, matlab, C++) and reproduce plots similar to the ones in Figure 5.3.1.

Sol.

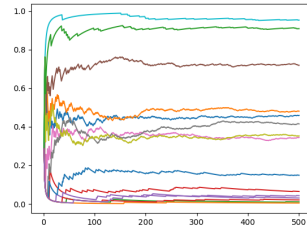
Here I tried with $(i, j, c) = (1, 1, 1), (1, 1, 3), (2, 5, 7)$ separately and get the images below:



(a) (1,1,1)



(b) (1,1,3)



(c) (2,5,7)

and with the code here:

```
import numpy
import matplotlib.pyplot as plt

M = 500
tag = range(1,M+1)
target = []

def simu(i, j, c, target, times = 0, stop = M):
    if times >= stop:
        return
    target.append(i/(i+j))
    t = numpy.random.uniform()
    if t <= i/(i+j):
        simu(i+c, j, c, target, times+1)
    else:
        simu(i, j+c, c, target, times+1)
    return

for i in range(15):
    target = []
    simu(1,1,1,target)
    plt.plot(tag,target)

plt.show()
```

□

Ex.5

Let X be a mean zero r.v. taking values from an interval $[-A, B]$. Fix a convex function $\phi : \mathbb{R} \rightarrow \mathbb{R}$. We will show that

$$E(\phi(X)) \leq \phi(-A) \frac{B}{A+B} + \phi(B) \frac{A}{A+B}$$

In words, over all possible distributions of X over $[-A, B]$, the most extreme distribution that maximizes $E[\phi(X)]$ is the one that puts point mass on $-A$ and B as in the right-hand side.

a. Let Y be a r.v. taking values from $[0, 1]$ and mean $p \in [0, 1]$. Suppose that for any convex function $\phi : \mathbb{R} \rightarrow \mathbb{R}$, we have

$$E(\phi(Y)) \leq (1-p)\phi(0) + p\phi(1)$$

then deduce the inequality above.

b. Here we will deduce the inequality in (a). Let Y be as before. Let $U \sim \text{Uniform}(0, 1)$ independent from Y . Argue that

$$\chi_{U \leq Y} | Y \sim \text{Bernoulli}(Y) \text{ and } \chi_{U \leq Y} \sim \text{Bernoulli}(p)$$

Then use Jensen's inequality to deduce

$$(1-p)\phi(0) + p\phi(1) = E(\phi(1(U \leq Y))) \geq E(\phi(Y))$$

c. Let $\phi(x) = e^{\theta x}$ for a fixed $\theta > 0$ and assume $A = B > 0$. Deduce that

$$E(e^{\theta X}) \leq \frac{E(e^{-\theta A}) + E(e^{\theta A})}{2} \leq e^{\theta^2 A^2 / 2}$$

Sol.

a. Let $Y = \frac{X}{A+B} + \frac{A}{A+B}$ and we know Y is a r.v. taking values from $[0, 1]$ and mean $\frac{A}{A+B}$, so we know

$$E(f(Y)) \leq \frac{B}{A+B}f(0) + \frac{A}{A+B}f(1)$$

for any convex function $f : \mathbb{R} \rightarrow \mathbb{R}$. Let $f = \phi((A+B)x - A)$ and we will know

$$\begin{aligned} f((1-t)x + ty) &= \phi((A+B)[(1-t)x + ty] - A) = \phi((1-t)[(A+B)x - A] + t[(A+B)y - A]) \\ &\leq (1-t)\phi((A+B)x - A) + t\phi((A+B)y - A) = (1-t)f(x) + f(y) \end{aligned}$$

for any $x, y \in \mathbb{R}, t \in [0, 1]$ and hence f is a convex function, so we will have

$$E(\phi(X)) = E(f(Y)) \leq \frac{B}{A+B}f(0) + \frac{A}{A+B}f(1) = \phi(-A)\frac{B}{A+B} + \phi(B)\frac{A}{A+B}$$

b. Assume $\mu(A) = P(Y \in A)$ and then we know (U, Y) will have the distribution

$$\int_B du \mu(dy)$$

and hence $P(\chi_{U \leq Y} | Y \in B) = \int_{(\chi_{U \leq Y} \in B)_Y} du = Y\delta_1(B) + (1-Y)\delta_0(B)$ where

$$\delta_x(B) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

where E_Y is the Y -section, and which means $\chi_{U \leq Y} | Y \sim \text{Bernoulli}(Y)$. For $\chi_{U \leq Y}$, we know it can be only 1 or 0 and where

$$\begin{aligned} P(\chi_{U \leq Y} = 1) &= \int_{U \leq Y} du \mu(dy) = \int \int_{(U \leq Y)_Y} du \mu(dy) = \int Y \mu(dy) = p \\ P(\chi_{U \leq Y} = 0) &= \int_{(U \leq Y)^c} du \mu(dy) = \int \int_{(U \leq Y)_Y^c} du \mu(dy) = \int (1-Y) \mu(dy) = 1-p \end{aligned}$$

then we may know

$$(1-p)\phi(0) + p\phi(1) = E(\phi(1(U \leq Y))) = \int \int \phi(\chi(U \leq Y)) du \mu(dy) \geq \int \phi(Y) \mu(dy) = E(\phi(Y))$$

by the Fubini's theorem and the Jensen's inequality.

c. It suffices to show that

$$\frac{E(e^{-\theta A}) + E(e^{\theta A})}{2} = \frac{e^{-\theta A} + e^{\theta A}}{2} \leq e^{\theta^2 A^2 / 2}$$

and notice

$$e^{x^2/2} - (e^x + e^{-x})/2 = \sum_{k=0}^{\infty} \left(\frac{1}{2^k k!} - \frac{1}{(2k)!} \right) x^{2k} \geq 0$$

for any $x \in \mathbb{R}$ since $2^k \leq (2k)!/k!$ for any integer k . □

Ex.6

Let $T = T(n, p)$ denote the total number of triangles in $G(n, p)$.

a. For each three distinct nodes i, j, k in G , let $Y_{i,j,k} := \chi_{(ij,jk,ki \in E)}$, which is the indicator variable for the event that there is a triangle with node set i, j, k . Show that

$$Y_{i,j,k} \sim \text{Bernoulli}(p^3)$$

b. Show that we can write

$$T = \sum_{1 \leq i < j < k \leq n} \chi_{(ij,jk,ki \in E)}$$

Deduce that the expected number of triangles is

$$ET = C_n^3 p^3$$

c. Show that

$$\text{var}(T(n, p)) = C_n^3(p^3 - p^6) + 12C_n^4(p^5 - p^6) \sim \frac{n^4}{2}(p^5 - p^6)$$

Thus $\text{Std}(T(n, p)) = \Theta(n^2)$. If CLT holds for $T(n, p)$, then $T(n, p)$ should fluctuate around its mean by $\Theta(n^2)$. Can we conclude this by CLT?

d. Show that for each $t \geq 0$,

$$P\left(|T(n, p) - C_n^3 p^3| \geq t\right) \leq 2 \exp\left(-\frac{t^2}{n(n-1)(n-2)^2}\right)$$

Deduce that the above probability is $o(1)$ if $t \gg n^2$. Specifically, for any $\epsilon > 0$

$$P\left(|T(n, p) - C_n^3 p^3| \geq n^{2+\epsilon}\right) \leq 2 \exp(-n^{2\epsilon})$$

Thus, McDirmid's inequality almost confirms the upper tail of fluctuation of $T(n, p)$ predicted by CLT.

Sol.

a. We know

$$P(Y_{i,j,k} = 1) = P(ij, jk, ki \in E) = P(ij \in E)P(jk \in E)P(ki \in E) = p^3$$

$$P(Y_{i,j,k} = 0) = 1 - P(Y_{i,j,k} = 1) = 1 - p^3$$

b. We know

$$T = \sum_{\{i,j,k\}, 1 \leq i,j,k \leq n, i \neq j \neq k} Y_{i,j,k} = \sum_{1 \leq i < j < k \leq n} Y_{i,j,k} = \sum_{1 \leq i < j < k \leq n} \chi_{(ij,jk,ki \in E)}$$

and hence

$$ET = E\left(\sum_{1 \leq i < j < k \leq n} \chi_{(ij,jk,ki \in E)}\right) = E\left(\sum_{1 \leq i < j < k \leq n} Y_{i,j,k}\right) = \sum_{1 \leq i < j < k \leq n} E(Y_{i,j,k}) = C_n^3 p^3$$

c. Denote $N = \{1, 2, \dots, n\}$ and notice

$$\begin{aligned} ET^2 &= E\left[\left(\sum_{1 \leq i < j < k \leq n} Y_{i,j,k}\right)^2\right] = E\left(\sum_{1 \leq i_1 < j_1 < k_1 \leq n} \sum_{1 \leq i_2 < j_2 < k_2 \leq n} Y_{i_1,j_1,k_1} Y_{i_2,j_2,k_2}\right) \\ &= \sum_{1 \leq i < j < k \leq n} EY_{i,j,k}^2 + \sum_{1 \leq i < j < k \leq n} \left(\sum_{p \in N - \{i,j,k\}} \sum_{\{m,n\} \subset \{i,j,k\}} EY_{i,j,k} Y_{m,n,p} \right) \\ &\quad + \sum_{1 \leq i < j < k \leq n} \left(\sum_{\{p,q\} \subset N - \{i,j,k\}} \sum_{m \in \{i,j,k\}} EY_{i,j,k} Y_{m,p,q} \right) \\ &\quad + \sum_{1 \leq i < j < k \leq n} \sum_{1 \leq p < q < r \leq n, \{i,j,k\} \cap \{p,q,r\} = \emptyset} EY_{i,j,k} Y_{p,q,r} \\ &= C_n^3 p^3 + 3(n-3)C_n^3 p^5 + 3C_n^3 C_{n-3}^2 p^6 + C_n^3 C_{n-3}^3 p^6 \end{aligned}$$

and hence

$$\begin{aligned} \text{var}(T(n, p)) &= ET^2 - (ET)^2 = C_n^3 p^3 + 3(n-3)C_n^3 p^5 + 3C_n^3 C_{n-3}^2 p^6 + C_n^3 C_{n-3}^3 p^6 - C_n^3 C_n^3 p^6 \\ &= C_n^3 p^3 + 12C_n^4 p^5 + C_n^3 (3C_{n-3}^2 + C_{n-3}^3 - C_n^3) p \\ &= C_n^3 (p^3 - p^6) + 12C_n^4 (p^5 - p^6) \end{aligned} \tag{6}$$

where you may check that $3C_{n-3}^2 + C_{n-3}^3 - C_n^3 = -3(n-3) - 1$. And we have

$$\lim_{n \rightarrow \infty} \frac{C_n^3 (p^3 - p^6) + 12C_n^4 (p^5 - p^6)}{\frac{n^4}{2} (p^5 - p^6)} = \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2)(n-3)}{n^4} = 1$$

and hence

$$\text{var}(T(n, p)) = C_n^3 (p^3 - p^6) + 12C_n^4 (p^5 - p^6) \sim \frac{n^4}{2} (p^5 - p^6)$$

By the way, we can not use CLT to $T(n, p)$ since $Y_{i,j,k}$ are not independent at all.

d. Let $Y_{i,j}$ be $\chi_{ij \in E}$ and T can be a function as $f(Y_{i,j}) : \mathbb{R}^{C_n^2} \rightarrow \mathbb{R}$, with $|f(x) - f(x')| \leq n-2$ if x, x' differs on only one coordinate. So we may use the McDirmid's inequality and we have

$$P(|T - C_n^3 p^3| \geq t) \leq 2 \exp\left(-\frac{t^2}{2C_n^2(n-2)^2}\right) = 2 \exp\left(-\frac{t^2}{n(n-1)(n-2)^2}\right)$$

then if $t = n^{2+\epsilon}$ for $\epsilon > 0$, we know

$$P(|T - C_n^3 p^3| \geq n^{2+\epsilon}) \leq 2 \exp\left(-n^{2\epsilon} \frac{n^3}{(n-1)(n-2)^2}\right) \leq 2 \exp(n^{-2\epsilon})$$

□

Durrett 4.3.1.

Give an example of a martingale X_n with $\sup_n |X_n| < \infty$ and $P(X_n = a \text{ i.o.}) = 1$ for $a = -1, 0, 1$. This example shows that it is not enough to have $\sup |X_{n+1} - X_n| < \infty$ in Theorem 4.3.1.

Sol.

Let ξ_n be i.i.d. and $P(\xi_n = 1) = P(\xi_n = -1) = P(\xi_n = 0) = 1/3$ and assume

$$X_{n+1} = \chi_{X_n=0} \xi_{n+1} + \chi_{X_n \neq 0} (1 + \xi_{n+1}) X_n$$

and $X_0 = \xi_0$. Assume $\mathcal{F}_n = \sigma(\xi_0, \xi_1, \dots, \xi_n)$ and we may use the induction to show that $X_n \in \mathcal{F}_n$. Notice

$$E(X_{n+1}|\mathcal{F}_n) = \chi_{X_n \neq 0} X_n = X_n$$

and we know X_n is a martingale w.r.t. to \mathcal{F}_n , also $\sup_n |X_n| < \infty$ but not uniformly bounded. Then we know

$$P(X_n = 0 \text{ i.o.}) = 1$$

since

$$P(X_n = 0 \text{ not i.o.}) = \sum_{k \geq 0} P\left(\bigcap_{m \geq k} \{X_m \neq 0\}\right) = 0$$

and hence $P(X_n = 1 \text{ i.o.}) = P(X_n = -1 \text{ i.o.}) = 1$. \square

Durrett 4.3.2.

(Assumes familiarity with finite state Markov chains.) Fine tune the example for the previous problem so that $P(X_n = 0) \rightarrow 1 - 2p$ and $P(X_n = -1), P(X_n = 1) \rightarrow p$, where p is your favorite number in $(0, 1/2)$, i.e., you are asked to do this for one value of p that you may choose. This example shows that a martingale can converge in distribution without converging a.s. (or in probability).

Sol.

Assume ξ_n independent with

$$P(\xi_k = k - i) = \frac{1}{4k} + \frac{1}{2k^2}, \quad P(\xi_k = -(k - i)) = \frac{1}{4k} + \frac{1}{2k^2}, \quad P(\xi_k = 0) = \frac{1}{2} - \frac{1}{k},$$

for $1 \leq i \leq k$ and let $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ and we know

$$X_{n+1} = \xi_{n+1} \chi_{X_n=0} + \frac{4n}{(n+2)(n-1)} X_n \chi_{X_n \neq 0}$$

and we know $X_n \in \mathcal{F}_n$ by the induction, then

$$E(X_{n+1}|\mathcal{F}_n) = \chi_{X_n \neq 0} X_n = X_n$$

Then

$$P(X_{n+1} = 0) = \left(\frac{1}{2} - \frac{1}{n+1}\right)P(X_n = 0) + \left(\frac{1}{2} - \frac{1}{n+1}\right)(1 - P(X_n = 0)) = \frac{1}{2} - \frac{1}{n+1} \rightarrow \frac{1}{2}$$

and

$$P(X_{n+1} = 1) = \left(\frac{1}{2} - \frac{1}{n}\right)\left(\frac{1}{2} - \frac{1}{n}\right) + \left(\frac{1}{2} + \frac{1}{n}\right)P(X_n = \frac{1}{2} + \frac{1}{n+1}) = \frac{1}{4} - \frac{1}{4(n+1)^2} \rightarrow \frac{1}{4}$$

and $P(X_{n+1} = -1)$ is similar. \square

Durrett Ex.4.3.3

Let X_n and Y_n be positive integrable and adapted to \mathcal{F}_n . Suppose $E(X_{n+1}|\mathcal{F}_n) \leq X_n + Y_n$, with $\sum Y_n < \infty$ a.s. Prove that X_n converges a.s. to a finite limit.

Sol.

Let $S_n = X_n - \sum_{k=1}^{n-1} Y_k$ and we know

$$E(S_{n+1} | \mathcal{F}_n) = E(X_{n+1} | \mathcal{F}_n) - Y_n - (X_n - S_n) \leq X_n - (X_n - S_n) = S_n$$

which means S_n is a supermartingale. Assume $M > 0$ and

$$N = \inf \left\{ k, \sum_{i=1}^k Y_i > M \right\}$$

then we know

$$S_{n \wedge N} + M \geq X_{n \wedge N} \geq 0$$

and hence $S_{n \wedge N} + M$ will converge to a finite limit almost surely since $S_{n \wedge N}$ is also a supermartingale.

Now it is suffice to show that S_n converges to a finite limit a.s. since $X_n = S_n + \sum_{i=1}^{n-1} Y_i$ where $\sum_{i=1}^{n-1} Y_i$ converges to $\sum_{i=1}^{\infty} Y_i < \infty$ a.s. For all $\omega \in \{\omega, \sum_{i=1}^{\infty} Y_i(\omega) < M\}$, we know $N(M) = \infty$ and hence $S_{n \wedge N}(\omega) = S_n(\omega)$, then we know $S_n(\omega)$ converges to a finite limit a.s. on $\{\omega, \sum_{i=1}^{\infty} Y_i(\omega) < M\}$, and hence S_n will converge to a finite limit a.s. on $\bigcup_{M=0}^{\infty} \{\omega, \sum_{i=1}^{\infty} Y_i(\omega) < M\}$ and notice

$$\bigcup_{M=0}^{\infty} \{\omega, \sum_{i=1}^{\infty} Y_i(\omega) < M\} = \Omega$$

and hence S_n converges to a finite limit a.s. □

Durrett Ex.4.3.4.

Let $p_m \in [0, 1)$. Use the Borel-Cantelli lemmas to show that

$$\prod_{m=1}^{\infty} (1 - p_m) = 0 \iff \sum_{m=1}^{\infty} p_m = \infty$$

Sol.

Let B_n be independent events with $P(B_n) = p_n$ and $\mathcal{F}_n = \sigma(\{\emptyset, \Omega, B_1, \dots, B_n\})$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. By the Second Borel-Cantelli's Lemma, we know

$$1 - \prod_{m=1}^{\infty} (1 - p_m) \geq P(B_n \text{ i.o.}) = P\left(\sum_{n=1}^{\infty} P(B_n | \mathcal{F}_{n-1}) = \infty\right) = P\left(\sum_{n=1}^{\infty} p_n = \infty\right)$$

and

$$1 - \sum_{k \geq 1} \prod_{m=k}^{\infty} (1 - p_m) \leq P(B_n \text{ i.o.}) = P\left(\sum_{n=1}^{\infty} p_n = \infty\right)$$

since B_n is independent to \mathcal{F}_{n-1} and hence if $\sum_{m=1}^{\infty} p_m = \infty$, we know

$$\prod_{m=1}^{\infty} (1 - p_m) \leq 1 - P\left(\sum_{n=1}^{\infty} \chi_{B_n} = \infty\right) = 0$$

which means $\prod_{m=1}^{\infty} (1 - p_m) = 0$. If $\prod_{m=1}^{\infty} (1 - p_m) = 0$, then we know $\prod_{m=k}^{\infty} (1 - p_m) = 0$ for any integer

k and hence

$$1 = 1 - \sum_{k \geq 1} \prod_{m=k}^{\infty} (1 - p_m) \leq P\left(\sum_{n=1}^{\infty} p_n = \infty\right)$$

which means $\sum_{n=1}^{\infty} p_n = \infty$ since $P(\sum_{n=1}^{\infty} p_n = \infty)$ can be only 0 or 1.

□