# 1 Theorem for WLLN

We need finite second moment.

$$P(|\frac{\sum x_i}{n} - \frac{\sum \mu_i}{n}| > \epsilon) \leq \frac{\sum \sigma_{ij}}{\epsilon^2 n^2}.$$

## 1.1 Proof:

$$\begin{aligned} P(|\sum x_i - \sum \mu_i| > n\epsilon) &\leq \frac{E(\sum y_i)^2}{n^2\epsilon^2} \ (y_i := x_i - \ mu_i) \\ &= \frac{\sum \sigma_{ij}}{\epsilon^2 n^2} \end{aligned}$$

because $E(\sum y_i)^2 = E\sum_i \sum_j y_i y_j$.

## 1.2 Corollary

Suppose $\{x_i\}_{i=1}^{\infty}$ 1) uncorrelated r.v.s with 2) identical mean $\mu$ and 3) bounded variance. Then

$$\frac{\sum x_i}{n} \to \mu.$$

### 1.2.1 proof:

$$p(|\frac{\sum x_i}{n}| > \epsilon) \leq \frac{\sum \sigma_{ij}}{\epsilon^2 n^2} = \frac{\sum^n \sigma_i}{\epsilon^2 n^2} \leq \frac{c \cdot n}{\epsilon^2 n^2} \to 0.$$

## 1.3 Example

Let n balls drop into n boxes uniformly at random. Let $N_n$ be the number of empty boxes. Then

$$\frac{N_n}{n} \to_p \frac{1}{e}.$$

### 1.3.1 Proof:

We label boxes by 1,2,...,n. Define $I_i$ to be the indicator if box i is empty. Therefore

$$N_n = \sum I_i.$$

Note that

$EI_i = EI_1 = (1 - 1/n)^n.$
$EI_iI_j = P(both i, j boxes empty) = (1 - 2/n)^n.$
$cov(I_1, I_2) = EI_1I_2 - EI_1EI_2 = (1 - 2/n)^n - (1 - 1/n)^2 n \to e^{-2} - e^{-2} \to 0.$

So we can relax uncorrelation condition for wlln in this case.

$$
\begin{aligned}
&P(|\frac{N_n}{n} - \frac{EN_n}{n}| > \epsilon) \\
&\qquad = P(\frac{N_n}{n} - \frac{\sum EI_i}{n} > \epsilon) \qquad \leq \frac{\sum \sigma_{ij}}{\epsilon^2 n^2} \\
&\qquad = \frac{\sigma_1 1}{\epsilon^2 n} + \frac{n(n-1)\sigma_1 1}{\epsilon^2 n} \to 0.
\end{aligned}
$$

### 1.4 Example2

Suppose we have n different coupons. At each time, I get a coupon iid from the n coupons. I define $T_n$ = the first time that all the coupons are collected. Then $\frac{T_n}{nlogn} \to_p 1$.

#### 1.4.1 Proof:

Define $I_i$ to be the waiting time between the (i-1) and ith coupon.

$$T_n = \sum I_i,$$

each $I_i$ are independent but not identical, $I_i \sim Geo(\frac{n-i+1}{n})$.

$$E(I_i) = \frac{n}{n-i+1}, var(I_i) = \frac{(i-1)n}{(n-i+1)^2}$$

$$E|T_n| = \sum EI_i = n \sum \frac{1}{n-i+1} = n(1 + 1/2 + 1/3 + ... + n) \approx nlogn$$

for above proof is here https://math.stackexchange.com/questions/306371/simple-proof-of-showing-the-harmonic-number-h-n-theta-log-n

$$
\begin{aligned}
&P(\frac{T_n}{E|T_n|} - 1| > \epsilon) \\
&\qquad = P(|T_n - E|T_n|| > \epsilon E(T_n)) \\
&\qquad \leq \frac{var(T_n)}{\epsilon^2 (ET_n)^2} \\
&\qquad = \frac{\sum var(I_i)}{\epsilon^2 (ET_n)^2} \\
&\qquad \leq \frac{\sum \frac{(i-1)n}{(n-i+1)^2}}{\epsilon^2 (logn)^2 n^2} \\
&\qquad \frac{n^2(\sum \frac{1}{(n-i+1)^2})}{\epsilon^2 (logn)^2 n^2} \to 0.
\end{aligned}
$$

because $(\sum \frac{1}{(n-i+1)^2})$ converges.

## 2 Theorem SLLN

Let $\{X_n\}$ be a sequence of pairwise independent and identically distribution r.v.s with $E|X_1| < \infty$ then

$$\frac{\sum x_i}{n} \to Ex_1 = \mu \ a.s.$$

1. It suffices to show SLLN for $x_i \geq 0$
2. Let $Y_i = x_i 1_{x_i \leq i}$, we claim it truncated version.

### 2.1 proof 1:

Note that

$$\begin{aligned} &\sum^{\infty} P(x_i \neq y_i) \\ &= \sum P(x_i > i) \\ &= \sum P(x_1 > i) < \infty. \end{aligned}$$

Then By B-C,

$$P(X_i \neq Y_i i.o.) = 0.$$

Therefore, the event

$$\{w|x_i(w) = y_i(w) \ for \ all \ but \ finitely \ many i's.\}$$

has prob 1. hence for every $w \in A$,

$$\frac{\sum^{\infty} x_i(w)}{n} \to \mu \iff \frac{\sum y_i(w)}{n} \to \mu.$$

Thus, it is suffices to show

$$\sum y_i/n \to \mu \ a.s.$$

Next

$$\begin{aligned} E(y_i) - \mu &= E(y_i) - E(x_i) \\ &= -E(x_i 1_{x_i > i}) \\ &= -E(X_1 1_{x_1 > i}) \to_{MCT} 0. \end{aligned}$$

In the last equation, we use the fact that $EX_1 < \infty$ implies $X_1 < \infty$ a.s. and define $z_i := x_1 1_{x_1 > i}$, then $z_1 \geq z_2 \geq ....$

Therefore it suffices to show

$$Z_n := \frac{\sum(y_i - Ey_i)}{n} \to_{a.s.} 0.$$

Now we will show there is $\{k_n\}$ such that $z_{k_n} \to 0$ $a.s.$ since

$$P(|Z_{k_n}| > \epsilon) \leq \frac{\sum var(Y_i)}{\epsilon^2 k_n^2},$$

$$\begin{aligned}
\sum^{\infty} P(|Z_{k_n}| > \epsilon) &\leq \sum_{n=1}^{\infty} \frac{\sum_{i=1}^{k_n} var(Y_i)}{\epsilon^2 k_n^2} \\
&= \frac{1}{\epsilon^2} \sum_{i=1}^{\infty} \sum_{n:k_n \geq i} \frac{var(Y_i)}{k_n^2} && \text{(이해 잘 안감)} \\
&= \frac{1}{\epsilon^2} \sum_{i=1}^{\infty} var(y_i) (\sum_{n:k_n \geq i} \frac{1}{k_n^2}) && (1)
\end{aligned}$$

We know that, and (take any $\alpha > 1, k_n = \lfloor \alpha^n \rfloor$ 나중에 $k_n/(k_{n-1}) \to \alpha$.)

$$\sum_{n:k_n \geq i} \frac{1}{k_n^2} \leq \sum_{n:k_n \geq i} \frac{1}{i^2} \leq \frac{1}{i^2} c \tag{2}$$

By applying (2) into (1),

$$\sum^{\infty} P(|Z_{k_n}| > \epsilon) \leq \frac{c}{\epsilon^2} (\sum_{i=1}^{\infty} \frac{var(y_i)}{i^2})$$

Also,

$$\sum_{i=1}^{\infty} \frac{var(y_i)}{i^2} \leq \sum_{i=1}^{\infty} \frac{E(X_1^2 1_{x_1 \leq i})}{i^2} = E(x_1^2 \sum_{x_1 \leq i} \frac{1}{i^2}) \leq c' E(X_1^2 \frac{1}{X_1}) \leq c' E X_1 < \infty$$

Then by B-C again, and $P(Z_{k_n} > \epsilon) \leq P(|Z_{k_n}| > \epsilon)$,

$$P(Z_{k_n} > \epsilon \;\; i.o.\} = 0$$

therefore, $Z_{k_n} \to 0$ a.s.

Let $T_n := \sum_{i=1}^{n} Y_i$, then for each m, we can find $k_n$, $k_{n-1}$ such that $k_{n-1} \leq m \leq k_{n+1}$. Then

$$\frac{T_{k_n}}{k_{n+1}} \leq \frac{T_m}{m} \leq \frac{T_{k_{n+1}}}{k_n}$$

Note that

$$\frac{T_{k_{n+1}}}{k_n} = \frac{T_{k_{n+1}}}{k_{n+1}} \frac{k_{n+1}}{k_n}$$

$$\frac{T_{k_n}}{k_{n+1}} = \frac{T_{k_n}}{k_n} \frac{k_n}{k_{n+1}}$$

We know $\frac{T_{k_{n+1}}}{k_{n+1}}$ converges. When $m \to \infty, n(m) \to \infty$, $\frac{T_{k_n}}{k_n} \to \mu$ and

$$\frac{k_n}{k_{n+1}} \to \alpha$$

Therefore

$$\frac{\mu}{\alpha} \leq liminf \frac{T_m}{m} \leq limsup \frac{T_m}{m} \leq \alpha\mu$$

take $\alpha \to 1$ then we have the result a.s.

# 3 The law of iterated logarithm

Let $X_1, X_2, ...$ be independent with mean 0 and variance 1. Then

$$P(limsup \frac{S_n}{\sqrt{2nloglogn}} = 1) = 1$$

In other words, for each $\epsilon > 0$,

$$P(S_n \geq (1+\epsilon)\sqrt{2nloglogn}\ i.o.) = 0,$$

$$P(S_n > (1-\epsilon)\sqrt{2nloglogn}\ i.o.) = 1.$$

Let $X_1, ..., X_n$ iid r.v.s with $EX_1 = 0, var(X_1) = 1$,

$$LLN\ \frac{\sum X_i}{n} \to EX_1 = 0, S_n = \sum X_i = o_p(n)$$

$$CLT\ \frac{\sum X_i}{\sqrt{n}} \to N(0,1),\ S_n = \sum X_i = O_p(\sqrt{n})$$

## 3.1 Theorem

Convergence in distribution implies bounded in probablity.

$$X_n \to_D X,\ then\ X_n = 0_p(1).$$

## 3.2 Lemma 1.

If $a_n \to \infty,\ and\ \frac{a_n}{\sqrt{logn}} \to 0$, then

$$\begin{aligned} P(\frac{\sum^n X_i}{\sqrt{n}} \geq a_n) &= exp(-1/2a_n^2(1+\xi_n)) \\ &= exp(-1/2a_n^2(1+0(1)),\ where\ \xi_n \to 0. \end{aligned}$$

## 3.3 Lemma 2. Maximal inequality

Define $M_n = \underset{i \leq n}{max} S_n$ be the maximum process.

For $\alpha \geq \sqrt{2}$, we have

$$P(\frac{M_n}{\sqrt{n}} \geq \alpha) \leq \alpha P(\frac{S_n}{\sqrt{n}} \geq \alpha - \sqrt{2}).$$

## 3.4 Proof for lemma2 maximal inequality

Using the previous result,

$$-S_{n_{k-1}} \leq 2\sqrt{2n_{k-1} loglogn_{k-1}}$$

with prob. 1 for a large enough k.

$$\begin{aligned} S_{n_k} \geq X_k + S_{n_{k-1}} &\geq X_k - 2\sqrt{2n_{k-1} loglogn_{k-1}} \\ &\geq (1-\theta^{-1})\sqrt{2n_{k-1} loglogn_{k-1}} - \frac{2}{\sqrt{\theta}} \frac{1}{\sqrt{2n_k loglogn_k}} \\ &\qquad\qquad\qquad\qquad \text{(definition of Xk)} \\ &= (1-\theta^{-1} - \frac{2}{\sqrt{\theta}})\frac{1}{\sqrt{2n_k loglogn_k}} \\ &\geq (1-\frac{3}{\sqrt{\theta}})\frac{1}{\sqrt{2n_k loglogn_k}} \\ &> \frac{1-\epsilon}{\sqrt{2n_k loglogn_k}}. \end{aligned}$$

Now

$$\{\frac{M_n}{\sqrt{n}} \geq \alpha\} \subset \{\frac{S_n}{\sqrt{n}} \geq \alpha\} \cup \{\frac{S_n}{\sqrt{n}} < \alpha - \sqrt{2},\ but\ \frac{S_i}{\sqrt{n}} > \alpha\ for\ some\ i\},$$

hence

$$P(\frac{M_n}{\sqrt{n}} \geq \alpha) \leq P(\frac{S_n}{\sqrt{n}} \geq \alpha) + p(\frac{S_n}{\sqrt{n}} < \alpha - \sqrt{2},\ but\ \frac{S_i}{\sqrt{n}} > \alpha\ for\ some\ i) \quad (1)$$

Let $A_i$ be the event that

$$\frac{M_i}{\sqrt{n}} \geq \alpha\ but\ \frac{M_j}{\sqrt{n}} < \alpha\ for\ every\ j \leq i.$$

$A_i$ are disjoint. Note that under $A_i$ and $\frac{S_n}{\sqrt{n}} < \alpha - \sqrt{2}$. We know $\frac{S_i}{\sqrt{n}} > \alpha$ and $\frac{S_n}{\sqrt{n}} > \alpha - \sqrt{2}$

$$\Rightarrow S_i - S_n > \sqrt{2n}.$$

Now,

$$
\begin{aligned}
P((1)) &= p(\frac{S_n}{\sqrt{n}} < \alpha - \sqrt{2}, \bigcup_{i=1}^{n} A_i) \\
&\leq \sum_{i=1}^{n} P(\frac{S_n}{\sqrt{n}} < \alpha - \sqrt{2} \bigcap A_i) && \text{(union bound)} \\
&\vdots \\
&\leq \sum_{i=1}^{n} P(\frac{S_i - S_n}{\sqrt{n}} > \sqrt{2} \bigcap A_i) && \text{(tricky)} \\
&= \leq \sum_{i=1}^{n} P(\frac{S_i - S_n}{\sqrt{n}} > \sqrt{2}) P(A_i) \\
&\leq \sum \frac{n-1}{2n} P(A_i) \\
&\leq \frac{1}{2} \sum^{n-1} P(A_i) \\
&= \frac{1}{2} P(\bigcup^{n-1} A_i) \\
&\leq \frac{1}{2} P(M_n \geq \alpha).
\end{aligned}
$$

$$
P(M_n \geq \alpha) \leq P(S_n \geq \alpha - \sqrt{2}) + \frac{1}{2} P(M_n \geq \alpha)
$$

Therefore

$$
P(M_n \geq \alpha) \leq P(S_n \geq \alpha - \sqrt{2}).
$$

□

## 3.5 Proof of LIL

. Fix $\epsilon > 0$ and choose $\theta$ such that $1 \leq \theta^2 < 1 + \epsilon$ and let $n_k = \theta^k$.

Define

$$
\begin{aligned}
X_k &= \theta \sqrt{2loglogn_k} \\
&= \theta \sqrt{2loglog\theta^k} \\
&= \theta \sqrt{2(logk + log(log\theta))} \\
&= \theta \sqrt{2logk} \frac{\sqrt{2logk + 2loglog\theta}}{\sqrt{2logk}} \\
&= \theta \sqrt{2logk}(1 + 0(1)).
\end{aligned}
$$

$$
\begin{aligned}
p(\frac{M_{n_k}}{\sqrt{n_k}} \geq X_k) &\leq 2P(\frac{S_{n_k}}{\sqrt{n_k}} \geq X_k - \sqrt{2}) && \text{(lemma2)} \\
&= 2exp(-\frac{1}{2}(X_k - \sqrt{2})^2(1+0(1))) && \text{(lemma1)} \\
&= 2exp(-\frac{1}{2}(X_k - \sqrt{2})^2(1+0(1))) \\
&= 2exp(-\frac{1}{2}(\theta\sqrt{2logk})^2(1+0(1)) \\
&= 2exp(-\frac{1}{2}\theta^2 2logk(1+0(1)) \\
&= 2\frac{1}{k^{\theta^2(1+0(1))}}
\end{aligned}
$$

By Borel Cantelli-I, sum of the series is finite, hence we have that

$$p(\frac{M_{n_k}}{\sqrt{n_k}} \geq \theta\sqrt{2loglogn_k}\ i.o.) = 0$$

Suppose $S_n \geq (1+\epsilon)\sqrt{2nloglogn}\ i.o.$ i.e.,

$$P(S_n \geq (1+\epsilon)\sqrt{2nloglogn}\ i.o.) = 1.$$

then choose

$$n_{k-1} < n \leq n_k,$$

then

$$
\begin{aligned}
\frac{M_{n_k}}{X_k\sqrt{n_k}} &= \frac{M_{n_k}}{\theta\sqrt{2n_k loglogn_k}} \\
&\geq \frac{S_n\theta\sqrt{2nloglogn}}{\theta\sqrt{2nloglogn}\theta\sqrt{2n_k loglogn_k}} \\
&\geq \frac{S_n}{\theta\sqrt{2nloglogn}}\frac{\theta\sqrt{2n_{k-1}loglogn_{k-1}}}{\theta\sqrt{2n_k loglognk}} \\
&\geq \frac{S_n}{\theta\sqrt{2nloglogn}}\frac{1}{\theta^{1/2}}(1+0(1)) \\
&= \frac{S_n}{\theta^{1.5}\sqrt{2nloglogn}}(1+0(1)) \\
&\geq \frac{1+\epsilon}{\theta^{1.5}}(1+0(1)) && \text{(by assumption)} \\
&> 1 && \text{(contradiction)}
\end{aligned}
$$

Above implies that

$$\Rightarrow P(S_n \geq (1+\epsilon)\sqrt{2nloglogn}\ i.o.) = 0$$

For the other direction, fix any $\epsilon$and $\theta \in N^*$ such that $\frac{3}{\sqrt{\theta}} < \epsilon$ and $n_k = \theta^k$. Define

$$X_k = (1-\theta^{-1})\sqrt{2n_k loglogn_k}$$

$$\begin{aligned}
P(S_{n_k} - S_{n_{k-1}} \geq X_k) &= exp(-\frac{X_k^2}{2(n_k - n_{k-1})}(1+0(1))) \qquad \text{(lemma 1)} \\
& exp(-\frac{(1-\theta^{-1})^2 2n_k loglogn_k}{2n_k(1-\theta^{-1}})(1+0(1)) \\
&= exp(-(1-\theta^{-1})loglogn_k(1+0(1))) \qquad (1) \\
&= exp(-(1-\theta^{-1})logk(1+0(1))) \\
&= \frac{1}{k^{(1-\theta^{-1})}(1+0(1))}
\end{aligned}$$

Where in (1),

$$loglogn_k = logk(1 + \frac{loglog\theta}{logk}) = logk(1+0(1)).$$

Therefore

$$\sum P(S_{n_k} - S_{n_{k-1}} \geq X_k) = \infty,$$

by the Borel-Cantelli 2, we have that

$$P(S_{n_k} - S_{n_{k-1}} \geq X_k \ i.o.) = 1$$

$\square$