Kolmogorov's Exponential Lower Bound and the Hartman-Wintner LIL

Stat 522 Handout
Rewrite of Pollard, **AUGTMP** pages 266-271
Wellner; 1/11/2004

Lemma (Kolmogorov's exponential lower bound). Suppose that $\{X_i\}_{i=1}^{\infty}$ are independent random variables with $E(X_i) = 0$, $\sigma_i^2 = Var(X_i) < \infty$, and let $V_n = \sigma_1^2 + \cdots + \sigma_n^2$ for $n \geq 1$. Suppose also that $|X_i| \leq \delta\sqrt{V_n}$ for some small $\delta > 0$. Then, for each $\theta > 1$ there exists an $x_0 > 0$ and a constant K both depending on θ only such that

$$P(S_n \geq x\sqrt{V_n}) \geq \exp\left(-\frac{1}{2}\theta x^2\right) \qquad \text{for } \ x_0 \leq x \leq K/\delta\,.$$

Proof. (Pollard, "AUGMTP", pages 266 - 268). We collect various restrictions on the constants involved at the beginning of the proof. The constant θ will determine a small $\epsilon > 0$, which in turn will determine another constant $\eta > 0$; as will be seen the constant η will be chosen somewhat smaller than $\epsilon^2/2$. Furthermore we will choose a small positive constant K depending on ϵ and η (and hence on θ so that

$$\psi(8K) \geq \max\{(1+\eta)^{-1}, (1+\epsilon)/2\} \tag{0.1}$$

and

$$\frac{\Delta(-2K)}{1+2K^2} > 1-\eta \tag{0.2}$$

where ψ is the function in Bennett's exponential bound and Δ is the function involved in the proof of Bennett's inequality. We will also need

$$\kappa \equiv (1+\eta) - \frac{\epsilon^2}{(1+\eta)(1+\epsilon)^2} < 1-\eta\,, \tag{0.3}$$

and

$$-\epsilon - (1+4\epsilon)(1+\epsilon) + \frac{1}{2}(1+\epsilon)^2(1-\eta) > -\theta/2\,. \tag{0.4}$$

The constant x_0 will need to be so large that

$$\frac{1}{2} \geq \exp(-\epsilon x_0^2)\,, \tag{0.5}$$

and

$$2 + 3(1+\epsilon)x\exp(\frac{1}{2}x^2(1+\epsilon)^2\kappa) \leq \frac{1}{2}\exp\left(\frac{1}{2}x^2(1+\epsilon)^2(1-\eta)\right), \qquad \text{for } \ x \geq x_0\,. \tag{0.6}$$

With this preparation we can begin the proof itself.

Without loss of generality, assume that $V_n = 1$ (equivalently, replace the X_i's by $X_i/\sqrt{V_n}$), and thus $\sigma_i^2 = E(X_i^2) \le \delta^2$ for each $i \le n$. By reasoning as in the proof of Bennett's inequality, for $t > 0$

$$\begin{aligned} E\exp(tS_n) &= \prod_{i=1}^n E\left(1 + tX_i + \frac{1}{2}t^2 X_i^2 \Delta(tX_i)\right) \\ &\ge \prod_{i=1}^n \left(1 + \frac{1}{2}t^2\sigma_i^2\Delta(-t\delta)\right) \qquad \text{since } \ \Delta \nearrow, \ X_i \ge -\delta \\ &\ge \exp\left(\sum_{i=1}^n \frac{(1/2)t^2\sigma_i^2\Delta(-t\delta)}{1 + (1/2)t^2\sigma_i^2\Delta(-t\delta)}\right) \end{aligned}$$

by using $\log(1+y) \ge y/(1+y)$. If $0 < t < 2K/\delta$, then $\Delta(-2K) \le \Delta(-t\delta) \le \Delta(0) = 1$. It follows that

$$E\exp(tS_n) \ge \exp\left(\sum_{i=1}^n \frac{(1/2)t^2\sigma_i^2\Delta(-2K)}{1+2K^2}\right) \ge \exp\left(\frac{1}{2}t^2(1-\eta)\right), \qquad \text{for } \ 0 < t\delta \le 2K\,.$$

Here the second inequality follows from (0.2) and the fact that $V_n = 1$.

We also have an upper bound: since

$$\exp(tx) = \int_{-\infty}^{x} te^{ty}dy = \int_{-\infty}^{\infty} te^{ty}1_{[y\le x]}\,dy$$

it follows by replacing x by S_n, taking expectations, and then using Fubini's theorem that

$$E\exp(tS_n) = \int_{-\infty}^{\infty} te^{ty}P(S_n \ge y)\,dy \le 1 + \int_0^{\infty} te^{ty}P(S_n \ge y)\,dy\,.$$

Combining these inequalities gives

$$\exp\left(\frac{1}{2}t^2(1-\eta)\right) \le 1 + \int_0^{\infty} te^{ty}P(S_n \ge y)\,dy\,.$$

Now the basic idea is to choose t so that the last integral is maximized in a small interval $J = [x, w]$. The interval J contributes the following term to the integral over $[0,\infty)$:

$$\int_x^w te^{ty}P(S_n \ge y)\,dy \le P(S_n \ge x)\int_x^w te^{ty}dy \le P(S_n \ge x)e^{tw}\,.$$

We will choose w so the contributions to the integral from $[0,\infty)\setminus J$ are small. Note that Bennett's inequality yields

$$P(S_n \ge y) \le \exp(-\frac{1}{2}y^2\psi(y\delta))\,,$$

and hence

$$\int te^{ty}P(S_n \ge y)dy \le \int te^{ty}\exp(-(1/2)y^2\psi(y\delta))\,dy\,.$$

Ignoring ψ, the integrand would be maximized by choosing $t = y$. and this suggests choosing $t = (1+\epsilon)x$, $w = (1+4\epsilon)x$. Note that $t \le 2x \le 2K/\delta$ as required. To handle large values of y we need to take account of the factor ψ. Now $y\psi(y) \nearrow$ in y, and since $x \le K/\delta$ we have

$$y\delta\psi(y\delta) \ge 8x\delta\psi(8x\delta) \ge 8x\delta\psi(8K) \qquad \text{for} \quad y \ge 8x\,,$$

so

$$\frac{1}{2}y^2\psi(y\delta) \ge \frac{1}{2}y8x\psi(8K) = \frac{yt4\psi(8K)}{1+\epsilon} \ge 2yt$$

by (0.1). Thus the contribution from the region $y \ge 8x$ is small if $x \le K/\delta$:

$$\begin{aligned}\int_{8x}^{\infty} te^{ty}P(S_n \ge y)dy &\le \int_{8x}^{\infty} t\exp(ty - 2ty)dy \\ &= \exp(-8tx) \le 1\,.\end{aligned}$$

For the interval $[0, 8x]$,

$$\psi(y\delta) \ge \psi(8x\delta) \ge \psi(8K) \ge \frac{1}{1+\eta}$$

sicne $x \le K/\delta$ and by (0.1) and since the integrand satisfies

$$\begin{aligned}te^{ty}P(S_n \ge y) &\le te^{ty}\exp\left(-\frac{y^2}{2(1+\eta)}\right) \\ &\le t\exp\left(\frac{1}{2}t^2(1+\eta) - \frac{|y-(1+\eta)t|^2}{2(1+\eta)}\right)\,.\end{aligned}$$

The exponent is maximized at $y = (1+\eta)t = (1+\eta)(1+\epsilon)x$ which lies in the interior of J with

$$\begin{aligned}&(1+\eta)t - x = (1+\eta)(1+\epsilon)x - x = (\epsilon + \eta + \epsilon\eta)x \ge \epsilon x\,, \\ &w - (1+\eta)t = (1+4\epsilon)x - (1+\eta)(1+\epsilon)x \ge (1+4\epsilon)x - (1+3\epsilon)x \ge \epsilon x\end{aligned}$$

since $(1+\eta)(1+\epsilon) \le (1+3\epsilon)$. (Draw a picture...!) Thus it follows that

$$\int_0^{8x} te^{ty}P(S_n \ge y)dy = \int_0^x + \int_x^w + \int_w^{8x} te^{ty}P(S_n \ge y)dy\,,$$

where

$$\begin{aligned}&\int_0^x + \int_w^{8x} te^{ty}P(S_n \ge y)dy \\ &\le \ t\exp\left(\frac{1}{2}t^2(1+\eta)\right)\sqrt{2\pi(1+\eta)}P(|\sqrt{1+\eta}Z| \ge \epsilon x) \\ &\le \ 3t\exp\left(\frac{1}{2}t^2(1+\eta) - \frac{\epsilon^2x^2}{2(1+\eta)}\right)\,.\end{aligned}$$

Also note that

$$\begin{aligned}\int_x^w te^{ty}P(S_n \ge y)dy &\le \int_x^w te^{ty}dyP(S_n \ge x) \\ &\le \exp(tw)P(S_n \ge x) \\ &= \exp(x^2(1+4\epsilon)(1+\epsilon))P(S_n \ge x)\,.\end{aligned}$$

Combining these pieces yields the following inequality:

$$\begin{aligned} 2 + 3(1+\epsilon)x \exp\left(\frac{1}{2}x^2(1+\epsilon)^2\left((1+\eta) - \frac{\epsilon^2}{(1+\eta)(1+\epsilon)^2}\right)\right) \\ + \exp(x^2(1+4\epsilon)(1+\epsilon))P(S_n \geq x) \\ \geq \exp\left(\frac{1}{2}x^2(1+\epsilon)^2(1-\eta)\right) \end{aligned}$$

for $0 < x \leq K/\delta$, or, equivalently with abbreviated notation,

$$A(x) + B(x)P(S_n \geq x) \geq C(x)$$

where A, B, and C depend on ϵ and η as well as x. If η is chosen so that (0.3) holds, then $A(x)/C(x) \to 0$ as $x \to \infty$, and we can find x_ϵ such that $A(x) \leq 2^{-1}C(x)$ for $x \geq x_\epsilon$. For x in this range we find that

$$B(x)P(S_n \geq x) \geq C(x) - A(x) \geq 2^{-1}C(x)\,,$$

and hence

$$\begin{aligned} P(S_n \geq x) &\geq \frac{C(x)}{2B(x)} \geq \exp(-\epsilon x^2)\frac{C(x)}{B(x)} \\ &= \exp\left(-\epsilon x^2 - x^2(1+4\epsilon)(1+\epsilon) + \frac{1}{2}x^2(1+\epsilon)^2(1-\eta)\right) \end{aligned}$$

if we choose x_ϵ so that (0.5) also holds. Finally we choose ϵ so small that (0.4) holds (which is possible since the left side of (0.4) converges to $-1/2$ as $\epsilon \searrow 0$). Now set $x_0 = x_\epsilon$. □

Question: What guarantees that the interval $[x_0, K/\delta]$ is not empty? For some history and commentary on this inequality, see Dudley (1989), page 379.

Now for the Hartman-Wintner version of the LIL:

Theorem. (Hartman and Wintner, 1941) Suppose that $X_1, X_2, \ldots$ are i.i.d. random variables with $E(X_i) = 0$ and $Var(X_i) = 1$. Let $S_n = X_1 + \cdots + X_n$. Then

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = 1 \qquad a.s. \tag{0.7}$$

and

$$\liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = -1 \qquad a.s.$$

The proof will proceed by a truncation argument with truncation levels chosen to mesh with our proof of the LIL for bounded random variables. Before proceeding with the proof we need the following lemma:

Lemma. The function

$$g(x) \equiv \sqrt{\frac{x}{2\log\log x}} = \frac{x}{L(x)}, \qquad x \geq e^e$$

is strictly increasing and

$$\int_{e^e}^{g^{-1}(t)} \frac{1}{L(x)} dx \le Ct, \qquad t \ge g(e^e)$$

where $C = 2e/(e-1)$ is an absolute constant.

Proof. Write

$$2g^2(x) = \frac{x}{\log\log x}\,;$$

taking logarithms gives

$$\log 2 + 2\log g(x) = \log x - \log\log\log x\,.$$

Differentiation across this identity gives

$$\begin{aligned} 2\frac{g'(x)}{g(x)} &= \frac{1}{x} - \frac{1}{\log\log x}\frac{1}{\log x}\frac{1}{x} \\ &= \frac{1}{x}\left(1 - \frac{1}{\log\log x}\frac{1}{\log x}\right) \\ &\ge \frac{1}{x}\left(1 - \frac{1}{e}\right), \qquad \text{for } \ x \ge e^e\,. \end{aligned}$$

This implies that

$$\frac{1}{L(x)} = \frac{g(x)}{x} \le \frac{2e}{e-1} g'(x) \equiv Cg'(x)\,,$$

and hence, for $t \ge g(e^e)$, that

$$\int_{e^e}^{g^{-1}(t)} \frac{1}{L(x)} dx \le \int_{e^e}^{g^{-1}(t)} Cg'(x)dx = Cg(x)\Big|_{g(e^e)}^{g^{-1}(t)} \le Ct\,.$$

□

Proof of upper Hartman-Wintner LIL: We first show that

$$\limsup_{n\to\infty} \frac{S_n}{L_n} \le \lambda \qquad \text{a.s.}$$

for a fixed $\lambda > 1$.

Define a truncation function $\tau(x, M)$ as follows:

$$\tau(x, M) \equiv -M1_{[x<-M]} + x1_{[|x|\le M]} + M1_{[X>M]}\,.$$

For $\epsilon > 0$ (which will be chosen in the end to depend on λ), define for $i \ge 17 > 1 + e^e$,

$$\begin{aligned} &Y_i = \tau(X_i, \epsilon g_i), \qquad g_i \equiv g(i) = \sqrt{\frac{i}{2\log\log i}}\,, \\ &Z_i = X_i - Y_i, \qquad \mu_i = E(Y_i), \qquad \xi_i = Y_i - \mu_i\,. \end{aligned}$$

Note that

$$x - \tau(x, M) = (x - M)1_{[x>M]} + (x + M)1_{[x<-M]}$$

and hence

$$|x - \tau(x, M)| \le |x| 1_{[|x|>M]} \,. \tag{0.8}$$

Also note that

$$|\mu_i| = |E(X_i - Z_i)| = |-E(Z_i)| \le E|Z_i|$$

since $E(X_i) = 0$. We now decompose S_n as

$$S_n = \sum_{i=1}^n \xi_i + \sum_{i=1}^n \mu_i + \sum_{i=1}^n Z_i \,.$$

The first term will be handled by our results for bounded random variables. To handle the second and third terms we bound $\sum_{i=17}^\infty |\mu_i|/L_i$ and then apply Kronecker's lemma:

$$\begin{aligned}
\sum_{i=17}^\infty \frac{|\mu_i|}{L_i} &\le \sum_{i=17}^\infty \frac{E|Z_i|}{L_i} \\
&\le \sum_{i=1}^{17} \frac{E\{|X_i| 1_{[|X_i|>\epsilon g(i)]}\}}{L_i} \qquad \text{using (0.8)} \\
&= \sum_{i=17}^\infty \frac{E\{|X_1| 1_{[g^{-1}(|X_1|/\epsilon)>i]}\}}{L_i} \qquad \text{since the } X_i\text{'s are i.i.d.} \\
&= E\left\{ |X_1| \sum_{i=17}^\infty \frac{\{1_{[g^{-1}(|X_1|/\epsilon)>i]}\}}{L_i} \right\} \\
&\le E\left\{ |X_1| \sum_{i=17}^\infty 1_{[g^{-1}(|X_1|/\epsilon)>i]} \int_{i-1}^i \frac{1}{L(x)} dx \right\} \\
&\le E\left\{ |X_1| \int \sum_{i=17}^\infty 1_{[g^{-1}(|X_1|/\epsilon)>i]} \frac{1_{[i-1<x\le i]}}{L(x)} dx \right\} \\
&\le E\left\{ |X_1| \int \frac{1_{[e^e<x\le g^{-1}(|X_1|/\epsilon)]}}{L(x)} dx \right\} \\
&\le E\left\{ |X_1| \frac{C|X_1|}{\epsilon} \right\} = \frac{C}{\epsilon} E|X_1|^2 < \infty
\end{aligned}$$

where we used the Lemma to get the first inequality in the last line. By Kronecker's lemma this yields

$$\frac{1}{L_n} \sum_{i=1}^n |\mu_i| = \frac{1}{L_n} \sum_{i=1}^n \frac{L_i |\mu_i|}{L_i} \to 0 \tag{0.9}$$

as $n \to \infty$. Similarly, finiteness of the expected value of $\sum_1^\infty |Z_i|/L_i$ implies

$$\sum_{i=1}^n Z_i / L_n \to_{a.s.} 0 \tag{0.10}$$

via Kronecker's lemma once again.

It remains only to treat $T_n \equiv \sum_1^n \xi_i$. We want to show that

$$\limsup_{n\to\infty} \frac{T_n}{L_n} \le \gamma \qquad \text{a.s.} \tag{0.11}$$

for any fixed $\gamma > 1$. Let $V_n = Var(T_n) = \sum_1^n Var(\xi_i)$. By the dominated convergence theorem $Var(\xi_i) \to 1$ as $i \to \infty$, and hence $V_n/n \to 1$ as $n \to \infty$. Let $\lambda = 1 + 2\delta$ for $\delta > 0$, and consider blocks defined by a subsequence n_k with $n_k/\rho^k \to 1$ as $k \to \infty$ for a $\rho > 1$ to be chosen. By the maximal inequality with $W_n \equiv n$ (possible since $Var(\tau(X, M)) \le Var(X)$ for each $M > 0$; exercise!), and with $\gamma = 1 + 2\delta$, $\lambda = 1 + \delta$, followed by Bennett's inequality,

$$\begin{aligned} &P(T_n \ge (\lambda+\delta)L_n \ \text{ for some } \ n_k \le n \le n_{k+1}) \\ &\le \ 2P(T_{n_{k+1}} \ge \lambda L_{n_k}), \qquad k \ \text{ large} \\ &\le \ 2\exp\left(-\frac{\lambda^2 2n_k \log\log n_k}{2V_{n_{k+1}}}\psi\left(\frac{2\epsilon g(n_{k+1})\lambda L_{n_k}}{V_{n_{k+1}}}\right)\right). \end{aligned}$$

Since $g(x) = x/L(x)$, the quantity inside the function ψ on the right side behaves as

$$\frac{2\epsilon n_{k+1}\lambda L_{n_k}}{L_{n_{k+1}} n_{k+1}} \sim 2\epsilon\frac{\lambda}{\sqrt{\rho}} < 2\epsilon\gamma$$

by choosing ρ close to 1. Thus the argument of the function ψ can be chosen arbitrarily close to zero, and hence the ψ factor can be made as close to 1 as desired by choosing ϵ small. The other term in the exponential behaves as $(\lambda^2/\rho)\log\log\rho^k$ for k large, and hence with appropriate choices of ρ and ϵ the bound decreases more rapidly than $k^{-\alpha}$ for some $\alpha > 1$, and this implies that (0.11) holds, as before via the Borel-Cantelli lemma. Combined with (0.9) and (0.10) it follows that

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} \le 1 \qquad \text{a.s.}$$

To show that (0.7) holds, we need to show that

$$\limsup_{n\to\infty} \frac{T_n}{\sqrt{2n\log\log n}} \ge 1 \qquad \text{a.s.}$$

where $T_n = \sum_1^n \xi_i$ as before.

For fixed $\gamma < 1$ we want to show that

$$\limsup_{n\to\infty} \frac{T_n}{L_n} \ge \gamma \qquad \text{a.s.}$$

In our lower bound proof of the LIL for normal random variables we chose $n_k = \lfloor a^k \rfloor$ with $a > 1$ large; here we take $n_k = k^k$. Set $T \equiv T_{n_k} - T_{n_{k-1}}$ and $V \equiv Var(T) = V_{n_k} - V_{n_{k-1}}$. Then $V/n_k \to 1$ as $k \to \infty$. The summands contributing to T are bounded in absolute value by $\delta\sqrt{V}$ where $\delta = 2\epsilon g(n_k)/\sqrt{V}$. We need to bound $P(T > \gamma L_{n_k})$ from below by a term of a divergent series. Fix a $\theta > 1$. Write x for $\gamma L_{n_k}/\sqrt{V}$. Kolmogorov's exponential lower bound then yields

$$\begin{aligned} P(T > \gamma L_{n_k}) \ &= \ P(T \ge x\sqrt{V}) \ge \exp\left(-\frac{1}{2}\theta x^2\right) \\ &= \ \exp\left(-\frac{\theta\gamma^2 2n_k\log\log n_k}{2V}\right) \end{aligned}$$

if $x_0 \leq x \leq K/\delta$; that is, if

$$x_0 \leq \gamma \sqrt{\frac{2n_k \log\log n_k}{n_k(1+o(1))}} \leq \frac{K}{2\epsilon} \sqrt{\frac{n_k(1+o(1))2\log\log n_k}{n_k}}\,.$$

By choosing ϵ sufficiently small, the range eventually contains the desired x value. The rest of the argument goes exactly as in the case of normal random variables. □