THE CENTRAL LIMIT THEOREM FOR MEDIANS

RUSS WOODROOFE

1. Setting

The purpose of this note is to give a simple sketch of the following theorem:

Theorem 1. *Let $X_1, \ldots, X_n$ be iid random variables with pdf $f(x)$ and cdf $F(X)$. If $F(0) = \frac{1}{2}$, and f is continuous at 0, then $\sqrt{n}M \xrightarrow{D} N(0, \frac{1}{(2f(0))^2})$.*

We'll need two relatively heavy results:

Theorem 2. (Taylor approximation) *If $g(x)$ is differentiable at 0, then there is a function $h(x)$ with $\lim_{x\to 0} h(x) = 0$ so that*

$$g(x) = g(0) + g'(0) \cdot x + h(x) \cdot x.$$

Theorem 3. (Lebesgue dominated convergence) *Suppose that $f_n(x) \to f(x)$ as $n \to \infty$ (for every x), where f_n and f are all integrable. If every $f_n(x)$ satisfies $|f_n(x)| \leq g(x)$ for some $g(x)$ with $\int_{-\infty}^{\infty} g(x)\,dx < \infty$, then*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x)\,dx = \int_{-\infty}^{\infty} (\lim_{n\to\infty} f_n(x))\,dx = \int_{-\infty}^{\infty} f(x)\,dx.$$

Please refer to Wikipedia (or some other reference) for more background on these facts.

Recall that f is *asymptotically equivalent to* g if $\lim \frac{f}{g} = 1$. We write $f \sim g$ in this case. We'll also need:

Theorem 4. (Stirling's formula) *We have that $n! \sim \sqrt{2\pi n} \cdot (\frac{n}{e})^n$ as $n \to \infty$.*

Lemma 5. (Easy calculus fact) $\lim_{n\to\infty}(1 + \frac{x}{n})^n = e^x$.

2. Sketch of Proof

Proof. We now sketch the proof of Theorem 1. Some details are left as unassigned exercises.

We leave two gaps:

1) We'll work with odd samples $n = 2k + 1$. Even samples are similar, and an easy detail to fill in. This means that the pdf of the nth median is

$$f_M(x) = \frac{(2k+1)!}{(k!)^2} \cdot F(x)^k \cdot (1 - F(x))^k \cdot f(x).$$

2) We assume that $f_M(x)$ is dominated for all n by some $g(x)$ with finite integral, as in the Lebesgue dominated convergence theorem. Under these circumstances, it suffices to show that $f_{\sqrt{n}M} \to f_{N(0, \frac{1}{(2f(0)^2})}$.

Exercise 6. (Medium Hard) Find a dominating function for f_M for odd samples.

Exercise 7. (Easy) Explain the details of why Lebesgue dominated convergence gives convergence in distribution.

We now apply our above result: first, take a Taylor approximation of $F(x)$:

$$\begin{aligned} F(x) &= F(0) + F'(0)\cdot x + H(x)\cdot x \\ &= \frac{1}{2} + f(0)\cdot x + H(x)\cdot x, \end{aligned}$$

where $\lim_{x\to 0} H(x) = 0$. Then

$$\begin{aligned} f_M(t) &= \frac{(2k+1)!}{(k!)^2}\cdot\left(\frac{1}{2} + f(0)\cdot t + H(t)\cdot t\right)^k\cdot\left(\frac{1}{2} - f(0)\cdot t - H(t)\cdot t\right)^k\cdot f(t) \\ &= \frac{(2k+1)!}{(k!)^2}\cdot\left(\frac{1}{4} - f(0)^2\cdot t^2 - H(t)^2\cdot t^2 - 2H(t)\cdot t\right)^k\cdot f(t) \\ &= \frac{(2k+1)!}{(k!)^2}\cdot 2^{-2k}\cdot\left(1 - 4f(0)^2\cdot t^2 - 4H(t)^2\cdot t^2 - 8H(t)\cdot f(0)\cdot t\right)^k\cdot f(t). \end{aligned}$$

To simplify notation, write $4H(t)^2\cdot t^2 - 8H(t)\cdot f(0)\cdot t$ as $t\cdot G(t)$.

We now examine $f_{\sqrt{n}M}(t)$, which we must show converges as $n\to\infty$ to a normal pdf:

$$\begin{aligned} f_{\sqrt{n}M}(t) &= \frac{1}{\sqrt{n}} f_M(\frac{t}{\sqrt{n}}) = \frac{1}{\sqrt{2k+1}}\cdot f_M(\frac{t}{\sqrt{2k+1}}) \\ &= \frac{1}{\sqrt{2k+1}}\frac{(2k+1)!}{(k!)^2}\cdot 2^{-2k}\cdot\left(1 - 4f(0)^2\cdot\frac{t^2}{2k+1} - G(\frac{t}{\sqrt{2k+1}})\cdot\frac{t}{\sqrt{2k+1}}\right)^k\cdot f(\frac{t}{\sqrt{2k+1}}). \end{aligned}$$

Then

$$\left(1 - 4f(0)^2\cdot\frac{t^2}{2k+1}\right)^k = \left(1 - \frac{4f(0)^2\cdot t^2\cdot\frac{k}{2k+1}}{k}\right)^k \to e^{-2f(0)^2\cdot t^2}.$$

Exercise 8. (Easy but tedious) Show that

$$\left(1 - 4f(0)^2\cdot\frac{t^2}{2k+1} - G(\frac{t}{\sqrt{2k+1}})\cdot\frac{t}{\sqrt{2k+1}}\right)^k \sim \left(1 - \frac{4f(0)^2\cdot t^2\cdot\frac{k}{2k+1}}{k}\right)^k.$$

(You'll need to look up a bound on the error term in the Taylor approximation).

Thus, we have

$$f_{\sqrt{n}M}(t) \sim \frac{1}{\sqrt{2k+1}}\frac{(2k+1)!}{(k!)^2}\cdot 2^{-2k}\cdot e^{-2f(0)^2\cdot t^2}\cdot f(\frac{t}{\sqrt{2k+1}}).$$

We further notice that $f(\frac{t}{\sqrt{2k+1}}) \to f(0)$, by continuity of f at 0, so

$$f_{\sqrt{n}M}(t) \sim \frac{1}{\sqrt{2k+1}}\frac{(2k+1)!}{(k!)^2}\cdot 2^{-2k}\cdot e^{-2f(0)^2\cdot t^2}\cdot f(0).$$

We now apply Stirling's formula:

$$f_{\sqrt{n}M}(t) \sim \frac{1}{\sqrt{2k+1}}\cdot\frac{\sqrt{2\pi(2k+1)}\cdot\left(\frac{2k+1}{e}\right)^{2k+1}}{2\pi k\cdot\left(\frac{k}{e}\right)^{2k}}\cdot 2^{-2k}\cdot e^{-2f(0)^2\cdot t^2}\cdot f(0).$$

We cancel the $\sqrt{2k+1}$ terms, and combine the $2k$th powers, to get

$$\begin{aligned} f_{\sqrt{n}M}(t) &\sim \frac{1}{\sqrt{2\pi}} \cdot \left(\frac{2k+1}{2k}\right)^{2k} \cdot \left(\frac{2k+1}{e \cdot k}\right) \cdot e^{-2f(0)^2 \cdot t^2} \cdot f(0) \\ &\sim \frac{1}{\sqrt{2\pi}} \cdot \left(1 + \frac{1}{2k}\right)^{2k} \cdot \left(\frac{2k+1}{e \cdot k}\right) \cdot e^{-2f(0)^2 \cdot t^2} \cdot f(0). \end{aligned}$$

We take limits of all remaining terms:

$$\begin{aligned} \lim_{n \to \infty} f_{\sqrt{n}M}(t) &= \frac{1}{\sqrt{2\pi}} \cdot e \cdot \frac{2}{e} \cdot e^{-2f(0)^2 \cdot t^2} \cdot f(0) \\ &= \frac{2f(0)}{\sqrt{2\pi}} e^{-2f(0)^2 \cdot t^2} = f_{N(0, \frac{1}{(2f(0)^2)})}(t). \quad \square \end{aligned}$$

Exercise 9. (Medium) Show that the $\sqrt{n}M$ pdf from an even sample is asymptotically equivalent with the odd sample pdf worked out above.