Math 114 Final

Lev Kruglyak

December 10, 2022

Problem 1. For a positive function f, consider two quantities

$$A := \int \left(\int f(x,y)^p \, dx \right)^{1/p} dy,$$
$$B := \left(\int \left(\int f(x,y) \, dy \right)^p dx \right)^{1/p},$$

for $1 \le p < \infty$. Assume all quantities are integrable and finite. Do we know that $A \ge B$ or $A \le B$ for all functions f? Prove your assertion.

We can use the following fact without proof, as given in the hint:

Claim. For functions h and g and p, q such that $1/p + 1/q = 1$, we have

$$\|h\|_p = \sup_{\|g\|_{L^q} \le 1} \int hg.$$

Now in our case, let's pick some function $g(x,y)$ with $g(x,y) \in L^q(x)$ with $\|g\|_q \le 1$ for all y. Then for any y we have:

$$\int f(x,y)g(x,y) \, dx \le \sup_{\|h\|_{L^q(x)} \le 1} \int f(x,y)h(x) \, dx = \|f\|_{L^p(x)} = \left(\int |f(x,y)|^p \, dx \right)^{1/p} = \left(\int f(x,y)^p \, dx \right)^{1/p}.$$

If we integrate with respect to y, we get an inequality

$$\iint f(x,y)g(x,y) \, dx \, dy \le \int \left(\int f(x,y)^p \, dx \right)^{1/p} dy = A.$$

But we can also write A as

$$A = \int \sup_{\|h\|_{L^q(x)} \le 1} \int f(x,y)h(x) \, dx \, dy \ge \sup_{\|h\|_{L^q(x)}} \iint f(x,y)h(x) \, dx \, dy.$$

All integrals above are assumed to be finite, so we can apply Fubini theorem to get:

$$\sup_{\|h\|_{L^q(x)}} \iint f(x,y)h(x) \, dx \, dy = \sup_{\|h\|_{L^q(x)}} \iint f(x,y) \, dy \, h(x) \, dx.$$

Let $t(x) = \int f(x,y) \, dy$. Notice that this is a positive function since f is positive. Applying the claim:

$$= \sup_{\|h\|_{L^q(x)}} \int t(x)h(x) \, dx = \|t\|_{L^p(x)} = \left(\int t(x)^p \, dx \right)^{1/p} = \left(\int \left(\int f(x,y) \, dy \right)^p dx \right)^{1/p} = B.$$

Putting these inequalities together, we get $A \ge B$.

Problem 2. Suppose $f_n \to f$ a.e. for al $x \in X = (0,1)$ and $\sup_n \|f_n\|_{L^2(X)} \leq M$ for some M fixed. Do we know that $\lim_{n\to\infty} \|f_n - f\|_{L^2(X)} = 0$? If, in addition, $\lim_{n\to\infty} \|f_n\|_{L^2(X)} = \|f\|_{L^2(X)} < \infty$, do we know that $\lim_{n\to\infty} \|f_n - f\|_{L^2(X)} = 0$?

Just with the first condition, it is not necessarily true. Consider the sequence of functions given by

$$f_n(x) = \sqrt{n}\chi_{(0,1/n)},$$

where χ_E is the characteristic function of E. Letting $f = 0$, it is clear that $f_n \to f$ a.e. and $\|f_n\|_{L^2(X)} = 1$ so $\sup_n \|f_n\|_{L^2(X)} = 1 \leq M$. However $\|f_n - f\|_{L^2(X)} = \|f_n\|_{L^2(X)} = 1$ so $\lim_{n\to\infty} \|f_n - f\|_{L^2(X)} \neq 0$.

If we add the second condition, the statement is true. Since $f_n \to f$ a.e, we have

$$4\|f\|_{L^2(X)}^2 = \left(\|f + f\|_{L^2(X)}\right)^2 = \int_X \lim_{n\to\infty} (f + f_n)^2 \, d\mu \leq \liminf_{n\to\infty} \int_X (f + f_n)^2 \, d\mu,$$

where the last inequality follows by Fatou's lemma. We can rewrite this integral by expanding $(f + f_n)^2 = 2f_n^2 + 2f^2 - (f - f_n)^2$:

$$\begin{aligned}
\liminf_{n\to\infty} \int_X (f + f_n)^2 \, d\mu &= 2\liminf_{n\to\infty} \int_X f_n^2 \, d\mu + 2\liminf_{n\to\infty} \int_X f^2 \, d\mu + \liminf_{n\to\infty} \left(- \int_X (f - f_n)^2 \, d\mu \right) \\
&= 2\liminf_{n\to\infty} \|f_n\|_{L^2(X)}^2 + 2\|f\|_{L^2(X)}^2 + \liminf_{n\to\infty} \left(- \int_X (f - f_n)^2 \, d\mu \right) \\
&= 4\|f\|_{L^2(X)}^2 - \limsup_{n\to\infty} \left(\int_X (f - f_n)^2 \, d\mu \right)
\end{aligned}$$

Putting this together with the preivous inequality, we get

$$\begin{aligned}
4\|f\|_{L^2(X)}^2 &\leq 4\|f\|_{L^2(X)}^2 - \limsup_{n\to\infty} \left(\int_X (f - f_n)^2 \, d\mu \right) \\
&\implies \limsup_{n\to\infty} \left(\int_X (f - f_n)^2 \, d\mu \right) \leq 0
\end{aligned}$$

So $\limsup_{n\to\infty} \|f - f_n\|_{L^2(X)}^2 = 0$ and so we are done.

Problem 3. Suppose μ is a probability measure (on a set, say, $\mathbb{R}^n$) and $f \geq 0$ with $\int f \, d\mu = 1$. Prove that

$$S(f) := \lim_{\epsilon \downarrow 0} \int f \log(f + \varepsilon) \, d\mu$$

exists, $0 \leq S(f)$ (the limit $S(f)$ can be infinity). Prove that, for any bounded real function h,

$$\int hf \, d\mu - S(f) = \int (h - (\log f)) \, f \, d\mu \leq \log \left(\int e^h \, d\mu \right).$$

For any $\epsilon > 0$, consider the $(-\epsilon, \infty)$-convex function $x \log(x + \epsilon)$. Jensen's inequality gives:

$$\int f \log(f + \epsilon) \, d\mu \geq \int f d\mu \log \left(\int f d\mu + \epsilon \right) = \log(1 + \epsilon).$$

Since $\epsilon > 0$, $\log(1 + \epsilon) \geq 0$ as well. As $\epsilon \downarrow 0$, we know that $\int f \log(f + \epsilon) \, d\mu$ decreases. However since it is bounded below by $\log(1 + \epsilon)$, it must converge, hence $S(f)$ exists. Also since $\log(1 + \epsilon) > 0$, $S(f) \geq 0$.

Now suppose we have some bounded real function h on our space. As before, we pick some $\epsilon > 0$. Now consider the probability measure given by $P(E) = \int_E f \, d\mu$. This is a probability measure since $\int f \, d\mu = 1$ by assumption. It's a common result in statistics that for g any measurable function we have $\int g \, dP = \int fg \, d\mu$.

Rewriting, we get

$$\int (h - \log(f + \epsilon)) \, f \, d\mu = \int \left(\log(e^h) - \log(f + \epsilon) \right) f \, d\mu = \int f \log \left(\frac{e^h}{f + \epsilon} \right) \, d\mu = \int \log \left(\frac{e^h}{f + \epsilon} \right) \, dP.$$

Applying Jensen's on log for $x > 0$ gives

$$\int \log \left(\frac{e^h}{f + \epsilon} \right) \, dP \leq \log \left(\int \frac{e^h}{f + \epsilon} \, dP \right) = \log \left(\int \frac{fe^h}{f + \epsilon} \, d\mu \right) = \log \left(\int \left(\frac{f}{f + \epsilon} \right) e^h \, d\mu \right).$$

Since $f/(f + \epsilon) \leq 1$, this is less than $\log \int e^h \, d\mu$. So putting everything together, we have

$$\int (h - \log(f + \epsilon)) f \, d\mu \leq \log \left(\int e^h \, d\mu \right) \implies_{\epsilon \to 0} \int hf \, d\mu - S(f) \leq \log \left(\int e^h \, d\mu \right).$$

Problem 4. Let $X_1, X_2, \ldots, X_N$ be identically independent random variables with $\mathbb{E}X_j = 0$, $\mathbb{E}X_j^2 = 1$ and $\mathbb{E}|X_j|^3 \le M < \infty$. This problem gives a proof of CLT with an error bound. Let ϕ be a real function such that the first three derivatives are bounded, i.e., $\sum_{j=0}^{3} \|\phi^{(j)}\|_\infty \le M < \infty$. Let Y_j be i.i.d. normal distribution with mean zero and variance one (Y_j and X_i are independent). Prove that

$$\left| \mathbb{E}\left[\phi\Big(\frac{1}{\sqrt{N}} \sum_j X_j \Big) \right] - \mathbb{E}\left[\phi(\zeta)\right] \right| \le C_M N^{-1/2}$$

where ζ is a normal distribution with mean zero and variance one.

Letting $Z_i = X_i + X_{i+1} + \cdots + X_N$ and $W_i = Y_1 + \cdots + Y_i$, we can rewrite

$$\begin{aligned}
\mathbb{E}\left[\phi\Big(\frac{1}{\sqrt{N}} \sum_j X_j \Big) \right] - \mathbb{E}\left[\phi\Big(\frac{1}{\sqrt{N}} \sum_j Y_j \Big) \right] &= \mathbb{E}\left[\phi\left(\frac{Z_1}{\sqrt{N}} \right) \right] - \mathbb{E}\left[\phi\left(\frac{W_1 + Z_2}{\sqrt{N}} \right) \right] \\
&\quad + \mathbb{E}\left[\phi\left(\frac{W_1 + Z_2}{\sqrt{N}} \right) \right] - \mathbb{E}\left[\phi\left(\frac{W_2 + Z_3}{\sqrt{N}} \right) \right] \\
&\quad + \cdots + \mathbb{E}\left[\phi\left(\frac{W_{N-1} + Z_N}{\sqrt{N}} \right) \right] - \mathbb{E}\left[\phi\left(\frac{W_N}{\sqrt{N}} \right) \right].
\end{aligned}$$

Now by Taylor's theorem on $\phi(x)$, we have

$$\begin{aligned}
&E\phi\left(\frac{1}{\sqrt{N}}[X_1 + X_2 + \cdots + X_N] \right) - E\phi\left(\frac{1}{\sqrt{N}}[Y_1 + X_2 + \cdots + X_N] \right) \\
&= E\phi\left(\frac{X_1 + Z_2}{\sqrt{N}} \right) - E\phi\left(\frac{Y_1 + Z_2}{\sqrt{N}} \right) \\
&= E\left[\phi'\left(\frac{Z_2}{\sqrt{N}} \right) \frac{X_1 - Y_1}{\sqrt{N}} + \phi''\left(\frac{Z_2}{\sqrt{N}} \right) \frac{X_1^2 - Y_1^2}{2N} + \frac{X_1^3 \phi'''(\gamma_{X_1}) - Y_1^3 \phi'''(\gamma_{Y_1})}{6N^{3/2}} \right],
\end{aligned}$$

for some $\gamma_{X_1} \in [Z_2/\sqrt{N}, (X_1 + Z_2)/\sqrt{N}]$ and $\gamma_{Y_1} \in [Z_2/\sqrt{N}, (Y_1 + Z_2/\sqrt{N})]$. Recall that the first three derivatives of ϕ are bounded by M. Also Z_2 must be independent from X_1 and Y_1 and $\mathbb{E}[X_1] = \mathbb{E}[Y_1] = 0$. Thus we can write:

$$\mathbb{E}\left[\phi'\left(\frac{Z_2}{\sqrt{N}} \right) \frac{X_1 - Y_1}{\sqrt{N}} \right] = \frac{1}{\sqrt{N}} \cdot \mathbb{E}\left[\phi'\left(\frac{Z_2}{\sqrt{N}} \right) \right] \mathbb{E}\left[X_1 - Y_1\right] = 0.$$

In a similar fashion, since $\mathbb{E}[X_1^2] = \mathbb{E}[Y_1^2] = 1$ we can write:

$$\mathbb{E}\left[\phi''\left(\frac{Z_2}{\sqrt{N}} \right) \frac{X_1^2 - Y_1^2}{2N} \right] = \frac{1}{2N} \cdot \mathbb{E}\left[\phi''\left(\frac{Z_2}{\sqrt{N}} \right) \right] \mathbb{E}(X_1^2 - Y_1^2) = 0.$$

Laslty, since $\mathbb{E}[|X_1|^3] \le M$, we have

$$\left| \mathbb{E}\left[\phi\left(\frac{X_1 + Z_2}{\sqrt{N}} \right) \right] - \mathbb{E}\left[\phi\left(\frac{Y_1 + Z_2}{\sqrt{N}} \right) \right] \right| \le \mathbb{E}\left[\left| \frac{X_1^3 \phi'''(\gamma_{X_1}) - Y_1^3 \phi'''(\gamma_{Y_1})}{6N^{3/2}} \right| \right] \le \frac{M(M + \mathbb{E}[|X_1|^3])}{6} N^{-3/2},$$

where the first inequality is the triangle inequality.

We can obtain a similar bound on the other $N-1$ terms in the telescoping sum for Z_3 and higher by an identical argument. Overall, by triangle inequality, we get

$$\left| \mathbb{E}\left[\phi\Big(\frac{1}{\sqrt{N}} \sum_j X_j \Big) \right] - \mathbb{E}\left[\phi\Big(\frac{1}{\sqrt{N}} \sum_j Y_j \Big) \right] \right| \le \sum_{i=1}^{N} \frac{M(M + \mathbb{E}[|X_i|^3|])}{6} N^{-3/2} \le N \times C_M N^{-3/2} = C_M N^{-1/2}$$

where C_M is the minimal $M(M + \mathbb{E}[|X_i|^3])/6$. Set $Z = W_N/\sqrt{N}$, so we have

$$\left| \mathbb{E}\left[\phi\Big(\frac{1}{\sqrt{N}} \sum_j X_j \Big) \right] - \mathbb{E}[\phi(Z)] \right| \le C_M N^{-1/2}$$

Note that Z is a normal distribution by the problem statement, and it's clear that it has mean zero and variance one.

Problem 5. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with $\mathbb{E}[X_j] = 0$ and $\mathbb{E}[X_j^2] = \sigma^2$. Let $S_n = X_1 + \cdots + X_n$. The weak law of large numbers states that for any $\delta > 0$,

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| > \delta\right) \leq \frac{\sigma^2}{n\delta^2}. \tag{1}$$

Suppose that, instead of $\mathbb{E}[X_j^2] = \sigma^2$, we only know that $(\mathbb{E}[X_j]^p)^{1/p} = M < \infty$ for some $1 < p < 2$. As in class, we let

$$\widehat{X}_j = X_j I(|X_j| \leq c),\ \widehat{Y}_j = X_j I(|X_j| > c),\ a_c = \mathbb{E}[\widehat{X}_j],\ \text{and } b_c = \mathbb{E}[\widehat{Y}_j].$$

Clearly, $X_j = \widehat{X}_j + \widehat{Y}_j$. Then we have

$$\begin{aligned}
\mathbb{E}\left|\sum_j (\widehat{X}_j + \widehat{Y}_j)\right| &\leq \mathbb{E}\left|\sum_j (\widehat{X}_j - a_c)\right| + E\left|\sum_j (\widehat{Y}_j - b_c)\right| \\
&\leq \left[\mathbb{E}\left(\sum_j (\widehat{X}_j - a_c)\right)^2\right]^{1/2} + 2n\mathbb{E}|\widehat{Y}_j| \\
&= \sqrt{n}\left[\mathbb{E}(\widehat{X}_1 - a_c)^2\right]^{1/2} + 2n\mathbb{E}|\widehat{Y}_1|.
\end{aligned}$$

Prove that

1. $\mathbb{E}[|\widehat{Y}_1|] \leq c^{1-p}M^p$,
2. $\mathbb{E}[\widehat{X}_1 - a_c]^2 \leq 2E[\widehat{X}_1^2] + 2a_c^2 \leq 4c^{2-p}M^p$,
3. $P\left(\left|\sum_j (\widehat{X}_j + \widehat{Y}_j)\right| \geq n\delta\right) \leq 4\delta^{-1}\inf_{c>0}[c^{1-p/2}M^{p/2}n^{-1/2} + c^{1-p}M^p]$.

To carry out the inf, prove by calculus that, for any $\alpha, \beta > 0$, there is a constant K such that

$$\inf_{x>0} Ax^\alpha + Bx^{-\beta} = KA^{1/(1+\gamma)}B^{\gamma/(1+\gamma)}, \quad \gamma = \frac{\alpha}{\beta}.$$

Let's begin by proving (1). Recall that Hölder's inequality gives us

$$\mathbb{E}[\widehat{Y}_1] \leq (\mathbb{E}[X_1]^p)^{1/p}(\mathbb{E}[I(|X_1| > c)])^{1-1/p} \leq M \cdot P(|X_1| > c)^{1-1/p}$$

since $\mathbb{E}[I(|X_1| > c)] = p(|X_1| > c)$ and $(\mathbb{E}[X_1]^p)^{1/p} = M$. Since $P(|X_1| > c) = P(|X_1|^p > c^p)$, Tchebychev's inequality gives us

$$P(|X_1| > c) \leq \frac{1}{c^p} \cdot \mathbb{E}[|X_1|^p] \leq \frac{M^p}{c^p} \implies \mathbb{E}[\widehat{Y}_1] \leq c^{1-p}M^p.$$

Next, let's prove (2). Note that we have

$$\begin{aligned}
\mathbb{E}[\widehat{X}_1 - a_c]^2 &\leq \mathbb{E}[\widehat{X}_1 - a_c]^2 + \mathbb{E}[\widehat{X}_1 + a_c]^2 = 2\mathbb{E}[\widehat{X}_1^2] + 2a_c^2 \\
&\leq 4\mathbb{E}[\widehat{X}_1^2] = 4\mathbb{E}[|X_1|^p \cdot |X_1|^{2-p} I(|X_1|^{2-p} \leq c^{2-p})] \\
&\leq 4c^{2-p} \cdot \mathbb{E}|X_1|^p \\
&\leq 4c^{2-p}M^p
\end{aligned}$$

where the penultimate inequality follows from $|X_1|^{2-p}I(|X_1|^{2-p} \leq c^{2-p}) \leq c^{2-p}$.

Lastly, let's prove (3). Recall that we have

$$P\left(\left|\frac{S_n}{n}\right| \geq \delta\right) = P\left(\left|\sum_j \left(\widehat{X}_j + \widehat{Y}_j\right)\right| \geq n\delta\right) \leq \frac{1}{n\delta} \cdot \mathbb{E}\left[\left|\sum_j \left(\widehat{X}_j + \widehat{Y}_j\right)\right|\right]$$

where the last inequality follows from the Tchebychev inequality. Applying (1) and (2), we get

$$\begin{aligned}\frac{1}{n\delta}\cdot\mathbb{E}\left[\left|\sum_j\left(\widehat{X}_j+\widehat{Y}_j\right)\right|\right] \le \frac{1}{n\delta}\sum_j\sqrt{n}\left(\mathbb{E}[\widehat{X}_j-a_c]^2\right)^{1/2}+2n\mathbb{E}[|\widehat{Y}_j|] &= \frac{1}{\delta}\left(n^{-1/2}(4c^{2-p}M^p)^{1/2}+2c^{1-p}M^p\right)\\ &= \frac{2}{\delta}\left(c^{1-p/2}M^{p/2}n^{-1/2}+c^{1-p}M^p\right)\\ &\le \frac{4}{\delta}\inf_{c>0}\left(c^{1-p/2}M^{p/2}n^{-1/2}+c^{1-p}M^p\right).\end{aligned}$$

To prove the weak law of large numbers, we'll need one more lemma.

Claim. For any $\alpha,\beta>0$, there is a constant K such that

$$\int_{x>0}\left(Ax^\alpha+Bx^{-\beta}\right)=KA^{\frac{1}{1+\gamma}}B^{\frac{\gamma}{1+\gamma}},\quad \gamma=\frac{\alpha}{\beta}.$$

Proof. Let $f(x)=Ax+Bx^{-1/\gamma}$. The derivatives of f for $x>0$ are

$$f'(x)=A-\frac{B}{\gamma}x^{-\frac{1}{\gamma}-1},\quad f''(x)=\frac{B(1+\gamma)}{\gamma^2}x^{-1/\gamma-2}.$$

Since $f''(x)>0$ on $x>0$, by freshman calculus f has a minimum at $x=(\gamma A/B)^{-\gamma/(1+\gamma)}$. Plugging this into f, we get

$$\inf_{x>0}f(x)=\left(\gamma^{-\gamma/(1+\gamma)}+\gamma^{1/(1+\gamma)}\right)A^{1/(1+\gamma)}B^{\gamma/(1+\gamma)}=KA^{1/(1+\gamma)}B^{\gamma/(1+\gamma)}$$

where K is set to be the coefficient term involving γ. □

Back to the case we're trying to solve, set $x=c$, $A=M^{p/2}n^{-1/2}$, $B=M^p$, $\alpha=1-p/2$, and $\beta=p-1$. Applying the result to (3), we get

$$P\left(\left|\frac{S_n}{n}\right|>\delta\right)\le P\left(\left|\frac{S_n}{n}\right|\ge\delta\right)\le\frac{2KM}{n^{1-1/p}\delta}.$$

Importantly, this tends to zero as $n\to\infty$.