Last Updated: 10 Nov 2023

Date: 10 Nov 2023

Welcome Back

In this section, we will discuss:

- Inequality

Recap

Here are some important inequality recap!

1. Cauchy-Schwartz: $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$.
2. Monotonicity (of expectation): $E(Y_1) \leq E(Y_2)$ for $Y_1 \leq Y_2$.
3. Markov: for any $a>0, {P(|Y|a) }. $
4. Chebyshev: for any $Y \sim [\mu, \sigma^2(< \infty)]$ and $\epsilon > 0$, $P(|Y - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$.
5. Chernoff: for any Y with MGF $M(\cdot)$ and any $a > 0, t > 0$ which $M(t)$ exists, we have $P(Y \geq a) \leq e^{-at}M(t)$.
6. Concentration (Hoeffding): Let $Y_1, Y_2, \cdots, Y_n$ be bounded, independent and $\epsilon > 0$. If $|Y_j| \leq c$ for each j then $P(|\bar{Y_n} - E\bar{Y_n}| > \epsilon) \leq 2e^{-n\epsilon^2/(2c^2)}$. If $a_j \leq Y_j \leq b_j$, then $P(|\bar{Y_n} - E\bar{Y_n}| > \epsilon) \leq 2e^{-2n^2\epsilon^2/(\sum(b_j - a_j)^2)}$.
7. Convexity (Jensen): $Eg(Y) g(EY) $ for g convex ($g''(x) \geq 0$.)
8. Contraction: for any $r \geq 1$, $ | E(Y|X)| _r |Y|_r.$
9. Correlation inequality: for g, h increasing functions $ (g(Y),h(Y)) .$
10. Mills inequality: let $Z \sim \mathcal{N}(0, 1)$, then $ P(Z > t) $ for all $t > 0$.
11. Minkowski: for $1 \leq r < \infty$, $\|X + Y\|_r \leq \|X\|_r + \|Y\|_r$.
12. Monotonicity of norms: for any $1 \leq r \leq s < \infty$,$ { |Y|_r |Y|_s}.$
13. Conjugate Norms (Holder): for $\frac{1}{r} + \frac{1}{s} = 1$, $\|XY\|_1 \leq \|X\|_r \|Y\|_s$.
14. KL Divergence: $D(f, g) = E_f \log \frac{f(X)}{g(X)}$.

15. AM-GM-HM inequality: $AM \geq GM \geq HM$, where AM stands for Arithmetic Means, GM stands for Geometric Means, and HM stands for Harmonic Means.

Section Discussion Questions

✏️ Section Problem 1

Let X be a non-negative random variable with finite variance, and let $0 \leq \theta \leq 1$. (a) Prove that

$$\mathbb{P}(X > \theta E(X)) \geq (1-\theta)^2 \frac{E(X)^2}{E\left(X^2\right)}.$$

Hint: Write $E(X) = E\left(X \mathbf{1}_{X \leq \theta E(X)}\right) + E\left(X \mathbf{1}_{X > \theta E(X)}\right)$.

b. The above inequality can actually be improved. Show that

$$\mathbb{P}(X > \theta E(X)) \geq \frac{(1-\theta)^2 E(X)^2}{\operatorname{Var}(X) + (1-\theta)^2 E(X)^2}$$

and confirm that this inequality is strictly stronger lower bound than the one in part (a). Denoting $E(X) = \mu$ and $\operatorname{Var}(X) = \sigma^2$, conclude that

$$P(X > \mu - \theta\sigma) \geq \frac{\theta^2}{1+\theta^2}$$

for $0 \leq \mu - \theta\sigma \leq \mu$.

Solution

a. Following the hint, we write

$$E(X) = E\left(X \mathbf{1}_{X \leq \theta E(X)}\right) + E\left(X \mathbf{1}_{X > \theta E(X)}\right).$$

The first addend

$$E\left(X 1_{X \leq \theta E(X)}\right) \leq \theta E[X]$$

while the second addend

$$E\left(X 1_{X > \theta E(X)}\right) \leq E\left[X^2\right]^{1/2} \mathbb{P}(X > \theta E[X])^{1/2}$$

by the Cauchy-Schwarz inequality. The desired inequality then follows.

b. By the Cauchy-Schwarz inequality;

$$E(X - \theta E[X]) \leq E\left((X - \theta E[X]) 1_{X > \theta E(X)}\right) \leq E\left[(X - \theta E[X])^2\right]^{1/2} \mathbb{P}(X > \theta E[X])^{1/2}$$

which, after rearranging, implies that

$$\mathbb{P}(X > \theta E[X]) \geq \frac{(1-\theta)^2 E[X]^2}{E\left[(X - \theta E[X])^2\right]} = \frac{(1-\theta)^2 E[X]^2}{\operatorname{Var}(X) + (1-\theta)^2 E[X]^2}.$$

The lower bound in part (a) can be rewritten as

$$\frac{(1-\theta)^2 E[X]^2}{\mathrm{Var}(X)+E[X]^2}$$

which is strictly smaller than

$$\frac{(1-\theta)^2 E[X]^2}{\mathrm{Var}(X)+(1-\theta)^2 E[X]^2}$$

provided that $E[X]^2 > 0$ and $0 < \theta \leq 1$. The rest follows by considering the substitution $\theta = 1 - \tilde{\theta}\sigma/\mu$ for $0 \leq \mu - \tilde{\theta}\sigma \leq \mu$.

✏️ Section Problem 2 (9.8)

(Bound on sums of third absolute moments) Let $X_1, \ldots, X_n$ be r.v.s with finite fourth moments. By using Cauchy-Schwarz, show that the following inequality holds:

$$\sum_{j=1}^{n} E(|X_j|^3) \leq \sqrt{\left(\sum_{j=1}^{n} E(X_j^2)\right)\left(\sum_{j=1}^{n} E(X_j^4)\right)}.$$

Hint: consider X_J, where J is a random index supported on $\{1, \ldots, n\}$.

Solution

Let $Y = X_J$, where J is a random index supported on $\{1, \ldots, n\}$. By LOTP,

$$E\left(|Y|^\ell\right) = \frac{1}{n}\sum_{j=1}^{n} E\left(|X_j|^\ell\right)$$

for $\ell = 2, 3, 4$. By Cauchy-Schwarz,

$$E\left(|Y|^3\right) \leq \sqrt{E\left(|Y|^2\right) E\left(|Y|^4\right)}.$$

So from (1), we get the inequality

$$\frac{1}{n}\sum_{j=1}^{n} E\left(|X_j|^3\right) \leq \sqrt{\left(\frac{1}{n}\sum_{j=1}^{n} E\left(|X_j|^2\right)\right)\left(\frac{1}{n}\sum_{j=1}^{n} E\left(|X_j|^4\right)\right)}.$$

Removing the unnecessary absolute values on the left and multiplying both sides by n, we get the desired inequality.

✏️ Section Problem 3

Let $X_1, X_2, \cdots$ be independent with mean 0 and $\sigma_i^2 = \mathbb{E}\left(X_i^2\right) < \infty$ and define partial sums $S_k = X_1 + X_2 + \cdots + X_k$. Then

$$P\left(\max_{1 \leq k \leq n} |S_k| \geq \epsilon\right) \leq \frac{\mathbb{E}\left(S_n^2\right)}{\epsilon^2}.$$

Solution

Let $A_k = \{|S_k| \geq \epsilon, |S_i| < \epsilon \, \forall i < k\}$. This is the sets of events consisting of $X_1, \cdots, X_n$ such that $|S_k|$ is the first partial sum with absolute value greater than ϵ. We can see that A_k's are disjoint and the union

$$\cup_{i=1}^n A_i = \{\exists k \in \{1, \ldots, n\} \; s.t. \; |S_k| \geq \epsilon\} = \{\max_{1 \leq k \leq n} |S_k| \geq \epsilon\}.$$

We must have

$$\mathbb{E}\left(S_n^2\right) \geq \mathbb{E}\left(S_n^2 \mathbb{I}\left(\cup_{k=1}^n A_k\right)\right) = \sum_{k=1}^n \mathbb{E}\left(S_n^2 \mathbb{I}(A_k)\right),$$

because A_k's are disjoint and indicator function is always ≤ 1. For each k we must have $S_n^2 = S_k^2 + 2(S_n - S_k)S_k + (S_n - S_k)^2$ and therefore

$$E\left(S_n^2 \mathbb{I}(A_k)\right) = E\left(S_k^2 \mathbb{I}(A_k)\right) + E\left((S_n - S_k)^2 \mathbb{I}(A_k)\right) + 2E\left((S_n - S_k)S_k \mathbb{I}(A_k)\right).$$

We know that $E\left((S_n - S_k)^2 \mathbb{I}(A_k)\right) \geq 0$ and since X_i's are iid and $\mathbb{I}(A_k)$ only depends on $X_1, \cdots, X_k$ we must have

$$E\left((S_n - S_k)S_k \mathbb{I}(A_k)\right) = E\left[S_n - S_k\right] E\left[S_k \mathbb{I}(A_k)\right] = 0.$$

More importantly

$$E(S_k^2 \mathbb{I}(A_k)) = \mathbb{P}(A_k) E\left[S_k^2 | A_k\right] \geq \mathbb{P}(A_k)\epsilon^2.$$

So

$$\mathbb{E}\left(S_n^2 \mathbb{I}(A_k)\right) \geq \mathbb{P}(A_k)\epsilon^2 \quad \forall k.$$

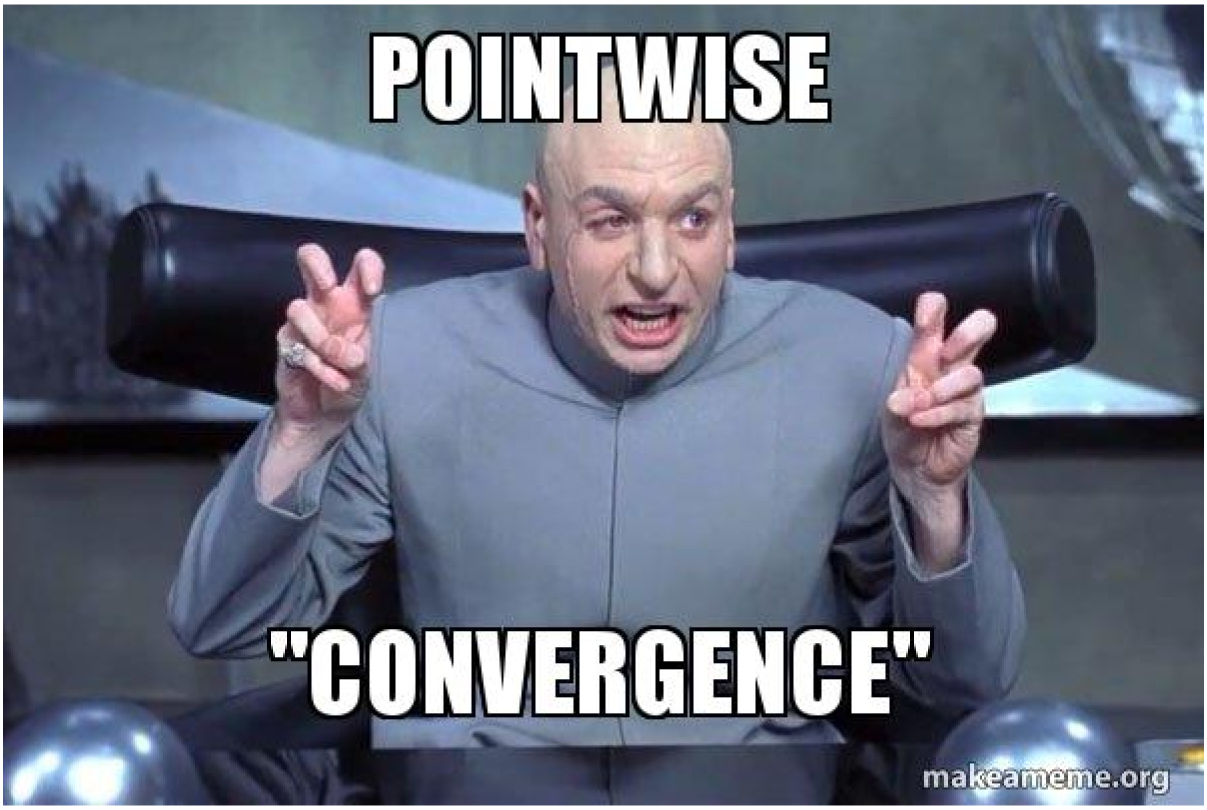
Finally

$$E(S_n^2) \geq \sum_{k=1}^n \mathbb{P}(A_k)\epsilon^2 = \epsilon^2 \mathbb{P}\left(\cup_{k=1}^n A_k\right) = \epsilon^2 \mathbb{P}\left(\max_{1 \leq k \leq n} |S_k| \geq \epsilon\right).$$

Next Week

Next week, we will discuss:

- Convergence



Feel free to upload the pencil problem you wish to be discussed next week here.

Note that a verified email address is needed in the GForm so we don't get scammy input! :)