Section 1.10. Important Inequalities

Note. We state and prove three inequalities in this section: Markov's Inequality, Chebyshev's Inequality, and Jensen's Inequality. In the first two inequalities, we give a proof for continuous random variables and leave the similar proofs for discrete random variables as exercises. We start with a preliminary result.

Theorem 1.10.1. Let X be a random variable and let$m \in \mathbb{N}$. Suppose $E[|X|^m]$ exists. If $k \in \mathbb{N}$ and $k \leq m$, then $E(X^k)$ exists.

Theorem 1.10.2. **Markov's Inequality.**
Let $u(X)$ be a nonnegative function of random variable X. If $E[u(X)]$ exists then for every positive constant c, $P(u(x) \geq c) \leq E[u(X)]/c$.

Note. If $E(X^2) < \infty$ then by Theorem 1.10.1, $\mu = E[X] < \infty$ so that μ exists and σ^2 exists.

Theorem 1.10.3. **Chebyshev's Inequality.**
Let X be a random variable where $E(X^2) < \infty$ (so that μ and σ^2 are define). Then for every $k > 0$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \text{ or } P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

Note 1.10.A. We could let $k\sigma = \varepsilon$ in Chebyshev's Inequality to get $P(|X - \mu| \geq \varepsilon) \leq \sigma^2/\varepsilon^2$ for all $\varepsilon > 0$.

Example 1.10.2. Let X be a discrete random variable where $P(X = -1) = P(X = 1) = 1/8$ and $P(X = 0) = 6/8 = 3/4$. Then $\mu = 0$ and

$$\sigma^2 = E[X^2] - \mu^2 = E[X^2] = \left(\frac{1}{8}\right)(-1)^2 + \left(\frac{1}{8}\right)(1)^2 + \left(\frac{3}{4}\right)(0)^2 = \frac{1}{4}.$$

With $k = 2$ in Chebyshev's Inequality we have

$$P(|X - \mu| \geq k\sigma) = P(|X| \geq 1) = 1/4 = 1/k^2.$$

In this case, we see that the upper bound on $P(|X - \mu| \geq k\sigma)$ given by Chebyshev's Inequality (namely, $1/k^2$) is attained for this specific example. This shows that Chebyshev's Inequality is best possible or "sharp."

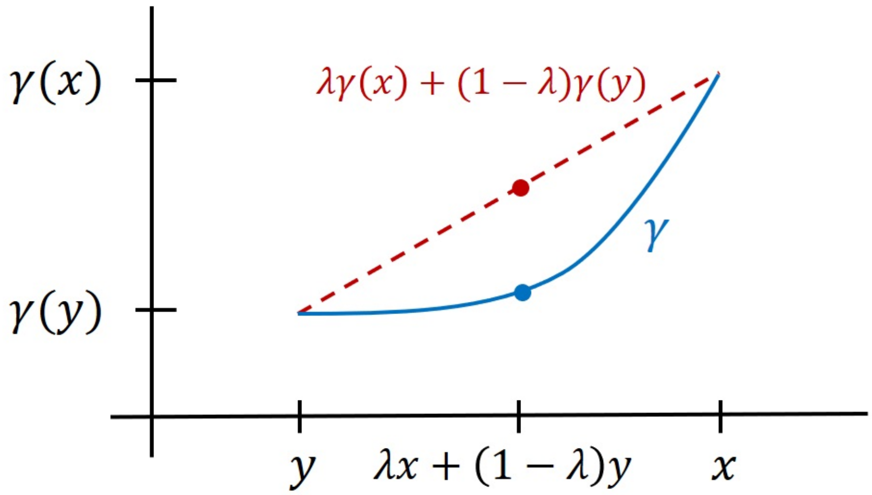
Note. Before stating and proving Jensen's Inequality we need a definition and a preliminary result.

Definition 1.10.1. A function φ defined on an interval (a, b), $-\infty \leq a < b \leq \infty$, is said to be a *convex function* if for all $x, y \in (a, b)$ and for all $0 < \gamma < 1$,

$$\varphi(\gamma x + (1 - \gamma)y) \leq \gamma\varphi(x) + (1 - \gamma)\varphi(y).$$

We say φ is *strictly convex* if this inequality is strict.

Note. As γ ranges over $(0,1)$, the values of $\gamma x + (1-\gamma)y$ ranges from y to x and the values of $\gamma\varphi(x) + (1-\gamma)\varphi(y)$ range "linearly" from $\varphi(y)$ to $\varphi(x)$. So the geometric interpretation is that the chord joining points $(x, \varphi(x)$ and $(y, \varphi(y))$ lies above the corresponding function values (as does a "concave up" function).



Theorem 1.10.4. If φ is differentiable on (a, b), then

(a) φ is convex if and only if $\varphi'(x) \leq \varphi'(y)$ for all $a < x < y < b$,

(b) φ is strictly convex if and only if $\varphi'(x) < \varphi'(y)$ for all $a < x < y < b$.

If φ is twice differentiable on (a, b) then

(a) φ is convex if and only if $\varphi''(x) \geq 0$ for all $a < x < b$,

(b) φ is strictly convex if and only if $\varphi''(x) > 0$ for all $a < x < b$.

Note. Given the figure above, Theorem 1.10.4 is not surprising. A proof can be found in my online notes for Real Analysis (MATH 5210/5220) on 6.6. Convex Functions (see Proposition 6.15).

Note. We now state Jensen's Inequality, which concerns convex functions applied to random variable. We give a proof for the special case in which φ is twice differentiable. For a proof assuming only convexity, see 6.6. Convex Functions (see Jensen's Inequality, though the proof is based on the interval $[0, 1]$; it also requires a slight background in Lebesgue integration).

Theorem 1.10.5. Jensen's Inequality. If φ is convex on an open interval I and X is a random variable whose support is contained in I and has finite expectation, then $\varphi(E[X]) \le E[\varphi(X)]$. If φ is strictly convex, then the inequality is strict unless X is a constant random variable.

Example 1.10.4/Definition. Let X be a discrete random variable with sample space $\{a_1, a_2, \ldots, a_n\}$, each a positive number. Suppose $P(X = a_i) = 1/n$ for $i = 1, 2, \ldots, n$. The *arithmetic mean* of X is

$$\text{AM} = E[X] = \frac{1}{n}\sum_{i=1}^{n} a_i.$$

Since $-\log x$ is a convex function, by Jensen's Inequality (in the discrete case)

$$-\log\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right) \le E[-\log X] = -\sum_{i=1}^{n} \log a_i = -\log(a_1 a_2 \cdots a_n)^{1/n},$$

or $\log\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right) \ge \log(a_1 a_2 \cdots a_n)^{1/n}$ or (since $\exp x$ is an increasing function)

$$(a_1 a_2 \cdots a_n)^{1/n} \le \frac{1}{n}\sum_{i=1}^{n} a_i.$$

The quantity $\text{GM} = (a_1 a_2 \cdots a_n)^{1/n}$ is the *geometric mean* of X. So we have

GM $\leq$ AM. Replacing a_i with $1/a_i$ (both positive) we then have

$$\left(\frac{1}{a_1}\frac{1}{a_2}\cdots\frac{1}{a_n}\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{a_i}$$

or

$$\frac{1}{\frac{1}{n}\sum_{i=1}^{n} a/a_i} \leq (a_1 a_2 \cdots a_n)^{1/n}.$$

The quantity $\mathrm{HM} = \dfrac{1}{\frac{1}{n}\sum_{i=1}^{n} 1/a_i}$ is the *harmonic mean* of X. So we have HM $\leq$ GM; combining with the above result we now conclude HM $\leq$ GM $\leq$ AM.

Revised: 1/1/2021