# Statistical Inequalities in Probability Theory

Santosh Dhaladhuli - MS20BTECH11007

April 9, 2021

# Markov's Inequality

## Statement:

If X is a non-negative random variable and a $>$ 0, then the probability that X is at least a is at most the expectation of X divided by a.
Mathematically,

$$\Pr\left(X \geq a\right) \leq \frac{E[X]}{a} \tag{1}$$

## Importance

1. It is useful in providing bounds for the **Cumulative Distribution Function(CDF)** of a random variable.
2. It provides machinery to define the **Chebyshev's Inequality** ,which is an important underpinning in much of Probability theory and Asymptotic theory.

# Mathematical Proof of Markov's Inequality

## Proof:

Using the definition of E[X],let f(x) be the probability function,

$$E[X] = \int_{-\infty}^{\infty} xf(x)\,dx \tag{2}$$

$$= \int_{0}^{\infty} xf(x)\,dx \because \text{X is non-negative} \tag{3}$$

$$= \int_{0}^{a} xf(x)\,dx + \int_{a}^{\infty} xf(x)\,dx \geq \int_{a}^{\infty} xf(x)\,dx \tag{4}$$

$$\int_{a}^{\infty} xf(x)\,dx \geq \int_{a}^{\infty} af(x)\,dx \tag{5}$$

$$\int_{a}^{\infty} af(x)\,dx = a\Pr(X \geq a) \tag{6}$$

# Proof contd...

## Proof:

$$E[X] \geq \int_a^\infty af(x)\,dx = a\Pr(X \geq a) \tag{7}$$

$$E[X] \geq a\Pr(X \geq a) \tag{8}$$

$$\Pr(X \geq a) \leq \frac{E[X]}{a}, a > 0 \tag{9}$$

Hence, Markov's Inequality is proved.

# Chebyshev's Inequality

**Statement:**

Let X be a random variable with finite expected value $E[X]$ and finite non-zero variance $\sigma^2$. Then for any real number k $>$ 0,

$$\Pr\left(|X - E[X]| \geq k\sigma\right) \leq \frac{1}{k^2}, \forall k > 0 \tag{10}$$

## Importance:

1. **The Weak Law of Large Numbers(WLLN)**, one of the single most important theorems in probability theory, follows directly from application of Chebyshev's Inequality.

# Mathematical Proof of Chebyshev's Inequality

## Proof:

We can directly apply Markovs inequality:

$$\Pr\left(|X - E[X]| \geq k\sigma\right) = \Pr\left(|X - E[X]|^2 \geq k^2\sigma^2\right) \tag{11}$$

$$\Pr\left(|X - E[X]|^2 \geq k^2\sigma^2\right) \leq \frac{E[|X - E[X]|^2]}{k^2\sigma^2} \tag{12}$$

$$\leq \frac{\sigma^2}{k^2\sigma^2} \tag{13}$$

$$\leq \frac{1}{k^2} \tag{14}$$

$$\therefore \Pr\left(|X - E[X]| \geq k\sigma\right) \leq \frac{1}{k^2} \tag{15}$$

Hence, Chebyshev's Inequality has been proved

# Jensen's Inequality

## Pre - requisites:

In general, a twice differentiable function $\phi(X)$ is a convex function iff:

$$\frac{d^2\phi}{dX^2} \geq 0$$

$\phi(X)$ will be a concave function when,

$$\frac{d^2\phi}{dX^2} \leq 0$$

## Statement:

In the context of probability theory, it is generally stated in the following form: if X is a random variable and $\phi$ is a convex function, then

$$\phi(E(X)) \leq E(\phi(X))$$

Inequality gets reversed when $\phi$ is a strictly concave function

## Importance

1. There are situations where, for **mathematical convenience**,we may want to switch the order of Expected value and function. Jensen's Inequality provides a machinery to confidently make this switch,under appropriate conditions.
2. The Inequality pops up quite a bit in **Machine Learning** contexts. Most modern methods for **Deep Generative Learning** rely on Jensen's Inequality as an essential underpinning.

# Mathematical Proof of Jensen's Inequality

## Proof:

Let $E[X] = \mu$, and let $L_\mu(X) = a + bX$ be the tangent line to the strictly convex function $\phi$ at $\mu$.
We have that $\phi(\mu) = L_\mu(\mu)$, and we know by convexity $\phi(X) \geq L_\mu(X) \forall X$.
Thus we have:

$$\phi(X) \geq L_\mu(X) \tag{16}$$

$$\implies E[\phi(X)] \geq E[L_\mu(X)] \tag{17}$$

$$\implies E[\phi(X)] \geq a + bE[X] = a + b\mu = L_\mu(\mu) = \phi(\mu) = \phi(E[X]) \tag{18}$$

$$\therefore \phi(E[X]) \leq E[\phi(X)] \tag{19}$$

Hence, Jensen's Inequality is proved.

# Example Question 1

### GATE ST 2021 Q.1 st. section

Let X be a non-constant positive Random Variable such that $E(X) = 9$. Then which of the following statements is True?

1. $E(\frac{1}{X+1}) > 0.1$ and $\Pr(X \geq 10) \leq 0.9$
2. $E(\frac{1}{X+1}) < 0.1$ and $\Pr(X \geq 10) \leq 0.9$
3. $E(\frac{1}{X+1}) > 0.1$ and $\Pr(X \geq 10) > 0.9$
4. $E(\frac{1}{X+1}) < 0.1$ and $\Pr(X \geq 10) > 0.9$

# Solution

Given, for X $> 0$ ,$E(X) = 9$, $E(\frac{1}{X+1})$ can be estimated by Jensens's Inequality.

$$\text{So for } \phi(X) = \frac{1}{X+1}, \tag{20}$$

$$\frac{d\phi}{dX} = -\frac{1}{(X+1)^2} \tag{21}$$

$$\frac{d^2\phi}{dX^2} = \frac{2}{(X+1)^3} \implies \frac{d^2\phi}{dX^2} \geq 0, (\because X > 0) \tag{22}$$

So, $\phi(X) = \frac{1}{X+1}$ is a convex function.

# Solution Contd...

$$E[(X+1)^{-1}] \geq \frac{1}{E[X]+1} \quad (23)$$

$$\implies E[(X+1)^{-1}] \geq \frac{1}{9+1} \quad (24)$$

$$\implies E[(X+1)^{-1}] \geq 0.1 \quad (25)$$

# Solution contd...

$\Pr(X \geq 10)$ can be estimated by Markov's Inequality.
for a = 10,using (9)

$$\Pr(X \geq 10) \leq \frac{E(X)}{10} \tag{26}$$

$$\implies \Pr(X \geq 10) \leq \frac{9}{10} \tag{27}$$

$$\therefore \Pr(X \geq 10) \leq 0.9 \tag{28}$$

So by equations (25) and (28),

$$\boxed{\textbf{Option 1 is the Correct Answer}}$$

# Example Question 2

**gov/stats/2015/statistics-I(1), Q.1(d)**

Let X be a Random Variable with $E[X] = 3$, $E[X^2] = 13$. Use Chebyshev's Inequality to obtain $\Pr(-2 < X < 8)$

# Solution

Computing the Variance($\sigma^2$),

$$\sigma^2 = E[X^2] - E[X]^2 \quad (29)$$

$$\implies \sigma^2 = 13 - 9 = 4 \quad (30)$$

$$\sigma = 2 \quad (31)$$

using (31),

$$\Pr(-2 < X < 8) = 1 - \Pr(|X - 3| > 5) \quad (32)$$

$$\Pr(|X - 3| > 5) = \Pr(|X - E[X]| > k\sigma) \quad (33)$$

$$k\sigma = 5 \quad (34)$$

$$\implies 2k = 5 \quad (35)$$

$$\therefore k = \frac{5}{2} \quad (36)$$

# Solution contd...

Using (10) , (33) and (36) in (32),

$$\Pr(-2 < X < 8) \geq 1 - (0.4)^2 \tag{37}$$

$$\implies \Pr(-2 < X < 8) \geq \frac{21}{25} \tag{38}$$