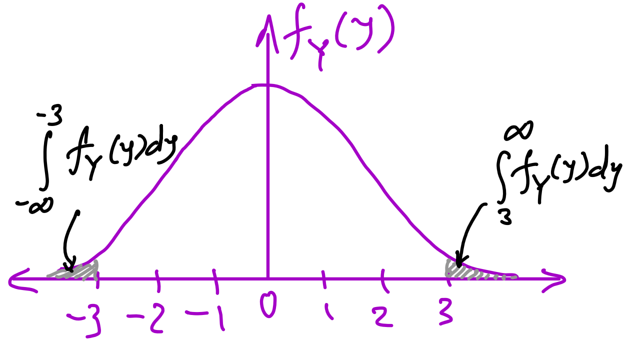
Lecture 24: Inequalities

Satyajit Thakor
IIT Mandi

Motivation

- Let's start with a familiar example: Find $P(|Y| \geq 3)$ if $Y \sim \underbrace{\mathcal{N}(0,1)}_{\text{standard normal distribution.}}$.



$$
\begin{aligned}
P(|Y| \geq 3) &= P(Y \leq -3) + P(Y \geq 3) \\
&= \int_{-\infty}^{-3} f_Y(y)\,dy + \int_{3}^{\infty} f_Y(y)\,dy \\
&= 0.0013 + 0.0013 \\
&= 0.0026
\end{aligned}
$$

(find using the table for CDF of $\mathcal{N}(0,1)$)

Inequalities

- Many times we like to solve such problems but we may only have limited information.
- How to solve such problems if we do not know the distribution but only know mean/variance?
- Markov and Chebyshev inequalities are useful to obtain upper bound on solutions, e.g., $P(|Y| \geq 3)$.

- Markov inequality: If X is an r.v. that takes only non-negative values, then for any value $a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Proof: $$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

Markov inequality

$$= \int_0^\infty x f_X(x)\,dx$$ ($\because$ X takes only non-neg. values

$$= \underbrace{\int_0^a x f_X(x)\,dx}_{\geq 0} + \int_a^\infty x f_X(x)\,dx$$

$$\geq \int_a^\infty x f_X(x)\,dx$$

$$\geq \int_a^\infty a f_X(x)\,dx$$ ($\because$ $x \geq a$)

$$= a \int_a^\infty f_X(x)\,dx = a P(X \geq a)$$

$$\Rightarrow P(X \geq a) \leq \frac{E(X)}{a}$$

Chebyshev inequality

- If X is an r.v. with mean μ and variance σ^2, then for any value $a > 0$

$$P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Proof: Note that $|X-\mu| \geq a \iff (X-\mu)^2 \geq a^2$

Hence, $P(|X-\mu| \geq a) = P\left((X-\mu)^2 \geq a^2\right)$

$$\leq \frac{E[(X-\mu)^2]}{a^2}$$

($\because$ considering $(X-\mu)^2$ as an r.v. and applying Markov ineq.

$$= \frac{\sigma^2}{a^2}$$

Example

- Find an upper bound on $P(|Y| \geq 3)$ if $Y \sim \mathcal{N}(0,1)$ using (1) Markov inequality (2) Chebyshev inequality.

(1) Using Markov inequality: $P(|Y| \geq 3) \leq \frac{E(|Y|)}{3}$

- Y is $\mathcal{N}(0,1)$. What is the mean of the r.v. $|Y|$?

- $E(|Y|) = \int_{-\infty}^{\infty} |y| \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy$ ($\because$ exp. of function of a random variable.)

$$= 2 \frac{1}{\sqrt{2\pi}} \int_0^{\infty} y e^{-y^2/2}\, dy = \sqrt{2/\pi} \underbrace{\int_0^{\infty} y e^{-y^2/2}\, dy}_{= 1 \text{ (Homework)}} = \sqrt{2/\pi}$$

- Hence, $P(|Y| \geq 3) \leq \frac{E(|Y|)}{3} = \frac{\sqrt{2/\pi}}{3} = 0.27.$

Example

(2) Using Chebyshev inequality:

$$P(|Y| \geq 3) \leq \frac{\sigma_Y^2}{3^2}$$

$$= \frac{1}{9}$$

$$= 0.11.$$

- Thus, the chebyshev inequality gives a better upper bound compared to the Markov inequality.
- To use chebyshev ineq. we need μ and σ^2. and to use markov ineq. we only need μ.

Example

- Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.
(a) What can be said about the probability that this week's production will exceed 74?

- Let Y be the no. of produced items in a week.

- Y is non-negative.
- Only mean is given.

$\Rightarrow$ We can use Markov ineq.

$$P(Y \geq 75) \leq \frac{E(Y)}{75}$$

$$= 50/75$$

$$= 2/3$$

Example

- (b) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be between 40 and 60?

$$P(40 < Y < 60) = P(|Y - 50| < 10)$$

$$= 1 - \underbrace{P(|Y - \overset{\mu_Y}{50}| \geq 10)}_{\leq \frac{\sigma_Y^2}{10^2} = \frac{25}{100} = 1/4}$$

by Chebyshev ineq.

$$\geq 1 - 1/4$$

$$= 3/4$$

- That is, probability that the production is between 40 and 60 is at least 0.75.

Example: Consider $X \sim \text{Bin}(n, p)$

Find upper bounds on $P(X \geq m)$

using Markov and Chebyshev inequalities.

(a) Using Markov inequality:

$$P(X \geq m) \leq \frac{E(X)}{m} = \frac{np}{m}$$

(b) Using Chebyshev inequality:

$$P(X \geq m) = P(X - np \geq m - np)$$

$$\leq P(|X - np| \geq m - np)$$

$$\leq \frac{\sigma_X^2}{(m-np)^2}$$

$$= P(\{X - np \geq m - np\} \cup \{X - np \leq -(m-np)\})$$

$$= \frac{np(1-p)}{(m-np)^2}$$

Alternatively:

(Using Markov inequality to obtain Chebyshev.)

$$P(X \geq m) = P(X - np \geq m - np)$$

$$= P((X - np)^2 \geq (m - np)^2)$$

$$\leq \frac{\sigma_X^2}{(m-np)^2} = \frac{np(1-p)}{(m-np)^2}$$