भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Joint MGF/CF, Probabilistic Inequalities–Markov's Inequality, Chebyshev's Inequality, Chernoff Bound, Jensen's Inequality

Karthik P. N.
Assistant Professor, Department of AI
Email: pnkarthik@ai.iith.ac.in

11 November 2024

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Joint MGF/CF

Joint MGF and Joint Characteristic Functions

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $X_1, \ldots, X_n$ be random variables w.r.t. $\mathscr{F}$.

Joint MGF and Joint Characteristic Function

1. The joint MGF of $X_1, \ldots, X_n$ is a function $M_{X_1,\ldots,X_n} : \mathbb{R}^n \to [0, +\infty]$, defined as

$$M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \mathbb{E}[e^{t_1 X_1 + \cdots + t_n X_n}] = \mathbb{E}[e^{\mathbf{t}^\top \mathbf{X}}],$$

where $\mathbf{t} = [t_1 \ \cdots \ t_n]^\top$ and $\mathbf{X} = [X_1 \ \cdots \ X_n]^\top$.

Joint MGF and Joint Characteristic Functions

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $X_1, \ldots, X_n$ be random variables w.r.t. $\mathscr{F}$.

Joint MGF and Joint Characteristic Function

1. The joint MGF of $X_1, \ldots, X_n$ is a function $M_{X_1,\ldots,X_n} : \mathbb{R}^n \to [0, +\infty]$, defined as

$$M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \mathbb{E}[e^{t_1X_1+\cdots+t_nX_n}] = \mathbb{E}[e^{\mathbf{t}^\top \mathbf{X}}],$$

   where $\mathbf{t} = [t_1 \ \cdots \ t_n]^\top$ and $\mathbf{X} = [X_1 \ \cdots \ X_n]^\top$.

2. The joint characteristic function of $X_1, \ldots, X_n$ is a function $C_{X_1,\ldots X_n} : \mathbb{R}^n \to \mathbb{C}$, defined as

$$C_{X_1,\ldots,X_n}(s_1, \ldots, s_n) = \mathbb{E}[j(s_1X_1 + \cdots + j_nX_n)] = \mathbb{E}[e^{j\mathbf{s}^\top \mathbf{X}}],$$

   where $\mathbf{s} = [s_1 \ \cdots \ s_n]^\top$.

Independence and Joint MGF/CF

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $X_1, \ldots, X_n$ be random variables w.r.t. $\mathscr{F}$.

Theorem (Independence and Joint MGF/CF)

1. Suppose that $M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) < +\infty$ for all $(t_1, \ldots, t_n) \in B(\mathbf{0}, \varepsilon)$ for some $\varepsilon > 0$, where $B(\mathbf{0}, \varepsilon)$ denotes a ball in $\mathbb{R}^n$ centered at the origin $\mathbf{0}$ and having radius ε. Then, the random variables $X_1, \ldots, X_n$ are independent if and only if

$$M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \prod_{i=1}^{n} M_{X_i}(t_i) \qquad \forall (t_1, \ldots, t_n) \in \mathbb{R}^n.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Independence and Joint MGF/CF

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $X_1, \ldots, X_n$ be random variables w.r.t. $\mathscr{F}$.

Theorem (Independence and Joint MGF/CF)

1. Suppose that $M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) < +\infty$ for all $(t_1, \ldots, t_n) \in B(\mathbf{0}, \varepsilon)$ for some $\varepsilon > 0$, where $B(\mathbf{0}, \varepsilon)$ denotes a ball in $\mathbb{R}^n$ centered at the origin $\mathbf{0}$ and having radius ε. Then, the random variables $X_1, \ldots, X_n$ are independent if and only if

$$M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \prod_{i=1}^{n} M_{X_i}(t_i) \qquad \forall (t_1, \ldots, t_n) \in \mathbb{R}^n.$$

2. The random variables $X_1, \ldots, X_n$ are independent if and only if

$$C_{X_1,\ldots,X_n}(s_1, \ldots, s_n) = \prod_{i=1}^{n} C_{X_i}(s_i) \qquad \forall (s_1, \ldots, s_n) \in \mathbb{R}^n.$$

Caution

Caution

To check that two random variables X and Y are independent, it DOES NOT suffice to check that

$$C_{X,Y}(s,s) = C_X(s)\, C_Y(s) \qquad \forall s \in \mathbb{R}.$$

Example:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{4}(1 + xy(x^2 - y^2)), & |x| < 1, |y| < 1, \\ 0, & \text{otherwise.} \end{cases}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probabilistic Inequalities

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Markov's Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a random variable w.r.t. $\mathscr{F}$.

Theorem (Markov's Inequality)

Let X be a non-negative random variable with $\mathbb{E}[X] < +\infty$. Then,

$$\mathbb{P}(\{X > \alpha\}) \leq \frac{\mathbb{E}[X]}{\alpha}, \qquad \forall \alpha > 0.$$

Markov's Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a random variable w.r.t. $\mathscr{F}$.

Theorem (Markov's Inequality)

Let X be a non-negative random variable with $\mathbb{E}[X] < +\infty$. Then,

$$\mathbb{P}(\{X > \alpha\}) \leq \frac{\mathbb{E}[X]}{\alpha}, \qquad \forall \alpha > 0.$$

Remarks:

- Markov's inequality only applies to non-negative random variables

Markov's Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a random variable w.r.t. $\mathscr{F}$.

Theorem (Markov's Inequality)

Let X be a non-negative random variable with $\mathbb{E}[X] < +\infty$. Then,

$$\mathbb{P}(\{X > \alpha\}) \leq \frac{\mathbb{E}[X]}{\alpha}, \qquad \forall \alpha > 0.$$

Remarks:

- Markov's inequality only applies to non-negative random variables
- The inequality is useful only for $\alpha > \mathbb{E}[X]$

Example

The course AI5030 has a total registration of 88 students.
The class average in the midterm exam was evaluated to be 6.65 (out of 30).
What is the maximum number of students that could have scored more than 10 marks?

Chebyshev's Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a random variable w.r.t. $\mathscr{F}$.

Theorem (Chebyshev's Inequality)

Let X have mean $\mathbb{E}[X] = \mu \in \mathbb{R}$ and variance $\sigma^2 < +\infty$. Then,

$$\mathbb{P}(\{|X - \mu| > \alpha\}) \leq \frac{\sigma^2}{\alpha^2}, \qquad \forall \alpha > 0.$$

Chernoff Bound

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let X be a random variable w.r.t. $\mathscr{F}$.

Theorem (Chernoff Bound)

Suppose that $M_X(t) < +\infty$ for all $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. Then,

$$\mathbb{P}(\{X > \alpha\}) \leq \frac{M_X(t)}{e^{t\alpha}}, \qquad \forall \alpha > 0,\ t > 0,\ t \in \text{ROC of } M_X.$$

Noting that the left-hand side of the above inequality is independent of t, we can optimise the right-hand side over t to get

$$\mathbb{P}(\{X > \alpha\}) \leq \inf_{\substack{t>0,\\ t\in\text{ROC}}} \frac{M_X(t)}{e^{t\alpha}}, \qquad \forall \alpha > 0.$$

Jensen's Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Theorem (Jensen's Inequality)

Let $X : \Omega \to \mathbb{R}$ be a random variable w.r.t. $\mathscr{F}$.
Let $g : \mathbb{R} \to \mathbb{R}$ be a convex, differentiable function, i.e.,

$$g(y) \geq g(x) + g'(x)(y - x) \qquad \forall x, y \in \mathbb{R}.$$

If $\left|\mathbb{E}[X]\right| < +\infty$ and $\left|\mathbb{E}[g(X)]\right| < +\infty$, then

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]).$$

Jensen's Inequality

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Theorem (Jensen's Inequality)

Let $X : \Omega \to \mathbb{R}$ be a random variable w.r.t. $\mathscr{F}$.
Let $g : \mathbb{R} \to \mathbb{R}$ be a convex, differentiable function, i.e.,

$$g(y) \geq g(x) + g'(x)(y - x) \qquad \forall x, y \in \mathbb{R}.$$

If $\left|\mathbb{E}[X]\right| < +\infty$ and $\left|\mathbb{E}[g(X)]\right| < +\infty$, then

$$\mathbb{E}[g(X)] \geq g(\mathbb{E}[X]).$$

Corollary:

$$\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2, \qquad \mathbb{E}[|X|] \geq |\mathbb{E}[X]|, \qquad \mathbb{E}[\log X] \leq \log \mathbb{E}[X].$$