भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Generating Functions–Probability Generating Functions, Moment Generating Functions, Characteristic Functions

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

04/07 November 2024

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability Generating Function (PGF)

Probability Generating Function (PGF)

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

Let X be an integer-valued random variable w.r.t. $\mathscr{F}$.

Definition (Probability Generating Function)

The probability generating function (PGF) of the random variable X is defined as

$$G_X(z) = \mathbb{E}[z^X] = \sum_{k \in \mathbb{N}} z^k p_X(k), \qquad z \in \mathbb{R}.$$

Probability Generating Function (PGF)

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

Let X be an integer-valued random variable w.r.t. $\mathscr{F}$.

Definition (Probability Generating Function)

The probability generating function (PGF) of the random variable X is defined as

$$G_X(z) = \mathbb{E}[z^X] = \sum_{k \in \mathbb{N}} z^k p_X(k), \qquad z \in \mathbb{R}.$$

The region of convergence of a PGF is defined as the set

$$\mathrm{ROC} = \left\{ z \in \mathbb{R} : |G_X(z)| < +\infty \right\}.$$

Probability Generating Function (PGF)

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

Let X be an integer-valued random variable w.r.t. $\mathscr{F}$.

Definition (Probability Generating Function)

The probability generating function (PGF) of the random variable X is defined as

$$G_X(z) = \mathbb{E}[z^X] = \sum_{k \in \mathbb{N}} z^k \, p_X(k), \qquad z \in \mathbb{R}.$$

The region of convergence of a PGF is defined as the set

$$\mathrm{ROC} = \left\{ \mathrm{z} \in \mathbb{R} : |\mathrm{G_X(z)}| < +\infty \right\}.$$

$$\{z : |z| < 1\} \subseteq \mathrm{ROC}.$$

Examples

- If $X \sim \text{Poisson}(\lambda)$, then

$$G_X(z) =$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Examples

- If $X \sim \text{Poisson}(\lambda)$, then

$$G_X(z) = e^{\lambda(z-1)}, \qquad z \in \mathbb{R}.$$

- If $X \sim \text{Geometric}(p)$, then

$$G_X(z) =$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Examples

- If $X \sim \text{Poisson}(\lambda)$, then

$$G_X(z) = e^{\lambda(z-1)}, \qquad z \in \mathbb{R}.$$

- If $X \sim \text{Geometric}(p)$, then

$$G_X(z) = \frac{pz}{1-(1-p)z}, \qquad |z| < \frac{1}{1-p}.$$

Properties of PGF

- $G_X(1) = 1$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Properties of PGF

- $G_X(1) = 1$
- $\frac{\mathrm{d}}{\mathrm{d}z} G_X(z)\Big|_{z=1} = \mathbb{E}[X]$

Properties of PGF

- $G_X(1) = 1$
- $\left.\frac{\mathrm{d}}{\mathrm{d}z} G_X(z)\right|_{z=1} = \mathbb{E}[X]$
- More generally, for any $k \in \mathbb{N}$,

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k} G_X(z)\right|_{z=1} = \mathbb{E}[X(X-1)\cdots(X-k+1)].$$

Properties of PGF

- $G_X(1) = 1$
- $\frac{\mathrm{d}}{\mathrm{d}z} G_X(z)\Big|_{z=1} = \mathbb{E}[X]$
- More generally, for any $k \in \mathbb{N}$,

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k} G_X(z)\Big|_{z=1} = \mathbb{E}[X(X-1)\cdots(X-k+1)].$$

- $X \perp\!\!\!\perp Y \implies G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$. Furthermore,

$$\{\text{ROC of } \mathrm{G}_{\mathrm{X+Y}}\} = \{\text{ROC of } \mathrm{G}_{\mathrm{X}}\} \cap \{\text{ROC of } \mathrm{G}_{\mathrm{Y}}\}.$$

Properties of PGF

- $G_X(1) = 1$
- $\left.\frac{\mathrm{d}}{\mathrm{d}z} G_X(z)\right|_{z=1} = \mathbb{E}[X]$
- More generally, for any $k \in \mathbb{N}$,

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k} G_X(z)\right|_{z=1} = \mathbb{E}[X(X-1)\cdots(X-k+1)].$$

- $X \perp\!\!\!\perp Y \implies G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$. Furthermore,

$$\{\mathrm{ROC\ of\ G_{X+Y}}\} = \{\mathrm{ROC\ of\ G_X}\} \cap \{\mathrm{ROC\ of\ G_Y}\}.$$

- Let $Y = \sum_{i=1}^{N} X_i$, where $X_1, X_2, \ldots$ are i.i.d., positive integer-valued, and N is independent of $\{X_1, X_2, \ldots\}$. Then,

$$G_Y(z) = G_N(G_X(z)).$$

Moment Generating Function (MGF)

Moment Generating Function (MGF)

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

Definition (Moment Generating Function)

The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, +\infty]$ defined as

$$M_X(t) = \mathbb{E}[e^{tX}], \qquad t \in \mathbb{R}.$$

The region of convergence of MGF is defined as the set

$$\mathrm{ROC} = \left\{ t \in \mathbb{R} : M_X(t) < +\infty \right\}.$$

Examples

- If $X \sim \text{Exponential}(\mu)$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

$$M_X(t) = \begin{cases} \frac{\mu}{\mu - t}, & t < \mu, \\ +\infty, & t \geq \mu. \end{cases}$$

- If $X \sim \mathcal{N}(0, 1)$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

$$M_X(t) = \begin{cases} \frac{\mu}{\mu - t}, & t < \mu, \\ +\infty, & t \geq \mu. \end{cases}$$

- If $X \sim \mathcal{N}(0, 1)$, then

$$M_X(t) = e^{\frac{t^2}{2}}, \qquad t \in \mathbb{R}.$$

- If $f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, $x \in \mathbb{R}$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

$$M_X(t) = \begin{cases} \frac{\mu}{\mu - t}, & t < \mu, \\ +\infty, & t \geq \mu. \end{cases}$$

- If $X \sim \mathcal{N}(0, 1)$, then

$$M_X(t) = e^{\frac{t^2}{2}}, \qquad t \in \mathbb{R}.$$

- If $f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, $x \in \mathbb{R}$, then

$$M_X(t) = \begin{cases} 1, & t = 0, \\ +\infty, & t \neq 0. \end{cases}$$

MGF and Uniqueness of the Underlying Distribution

Theorem (MGF and Underlying Distribution)

1. Suppose there exists $\varepsilon > 0$ such that
$$M_X(t) < +\infty \qquad \forall t \in (-\varepsilon, \varepsilon).$$
Then, $M_X(t)$ determines the CDF of X uniquely.
2. If X and Y are random variables such that $M_X(t) = M_Y(t) < +\infty$ for all $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. Then, X and Y have the same CDF.

Properties of MGF

- $M_X(0) = 1$

- [Moment generating property]
  Suppose $M_X(t) < +\infty$ for all $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. Then,

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} M_X(t)\bigg|_{t=0} = \mathbb{E}[X^k] \qquad \forall k \in \mathbb{N}.$$

  In particular, for $k = 1$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} M_X(t)\bigg|_{t=0} = \mathbb{E}[X].$$

Properties of MGF

- If $Y = aX + b$, then
$$M_Y(t) = e^{bt}\, M_X(at).$$
As a corollary, it follows that if $Y = \sigma X + \mu$, where $X \sim \mathcal{N}(0, 1)$, then
$$M_Y(t) = e^{\mu t}\, M_X(\sigma t) = e^{\mu t}\, e^{t^2 \sigma^2 / 2}.$$

- If $X \perp\!\!\!\perp Y$, then
$$M_{X+Y}(t) = M_X(t)\, M_Y(t).$$
For example, if $X_1 \sim \mathrm{Exponential}(\mu_1)$ and $X_2 \sim \mathrm{Exponential}(\mu_2)$, then
$$M_{X_1+X_2}(t) = \begin{cases} \frac{\mu_1}{\mu_1 - t} \cdot \frac{\mu_2}{\mu_2 - t}, & t < \min\{\mu_1, \mu_2\}, \\ +\infty, & \text{otherwise.} \end{cases}$$

Properties of MGF

- Let $Y = \sum_{i=1}^{N} X_i$, where $X_1, X_2, \ldots$ are i.i.d., and N is positive integer-valued and independent of $\{X_1, X_2, \ldots\}$. Then,

$$M_Y(t) = G_N(M_X(t)) = M_N(\log M_X(t)),$$

where G_N is the PGF of N.

As a corollary, suppose $X_1, X_2, \ldots \stackrel{\text{i.i.d.}}{\sim} \text{Exponential}(\mu)$ and $N \sim \text{Geometric}(p)$, then

Properties of MGF

- Let $Y = \sum_{i=1}^{N} X_i$, where $X_1, X_2, \ldots$ are i.i.d., and N is positive integer-valued and independent of $\{X_1, X_2, \ldots\}$. Then,

$$M_Y(t) = G_N(M_X(t)) = M_N(\log M_X(t)),$$

where G_N is the PGF of N.
As a corollary, suppose $X_1, X_2, \ldots \stackrel{\text{i.i.d.}}{\sim} \text{Exponential}(\mu)$ and $N \sim \text{Geometric}(p)$, then

$$M_Y(t) = \begin{cases} \frac{\mu p}{\mu p - t}, & t < \mu p, \\ +\infty, & t \geq \mu p. \end{cases}$$

Characteristic Functions

Characteristic Function

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

Definition (Characteristic Function

The characteristic function of the random variable X is a function $C_X : \mathbb{R} \to \mathbb{C}$, defined as

$$C_X(s) = \mathbb{E}[e^{jsX}] = \mathbb{E}[\cos sX] + j\,\mathbb{E}[\sin sX], \qquad s \in \mathbb{R}.$$

Characteristic Function

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

Definition (Characteristic Function

The characteristic function of the random variable X is a function $C_X : \mathbb{R} \to \mathbb{C}$, defined as

$$C_X(s) = \mathbb{E}[e^{jsX}] = \mathbb{E}[\cos sX] + j\,\mathbb{E}[\sin sX], \qquad s \in \mathbb{R}.$$

Remark:

$$\left| C_X(s) \right| \leq$$

Characteristic Function

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

Definition (Characteristic Function

The characteristic function of the random variable X is a function $C_X : \mathbb{R} \to \mathbb{C}$, defined as

$$C_X(s) = \mathbb{E}[e^{jsX}] = \mathbb{E}[\cos sX] + j\,\mathbb{E}[\sin sX], \qquad s \in \mathbb{R}.$$

Remark:

$$\left| C_X(s) \right| \leq 1 \qquad \forall s \in \mathbb{R}.$$

Examples

- If $X \sim \mathrm{Exponential}(\mu)$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

$$C_X(s) = \frac{\mu}{\mu - js}, \qquad s \in \mathbb{R}.$$

- If $f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, $x \in \mathbb{R}$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

$$C_X(s) = \frac{\mu}{\mu - js}, \qquad s \in \mathbb{R}.$$

- If $f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, $x \in \mathbb{R}$, then

Examples

- If $X \sim \text{Exponential}(\mu)$, then

$$C_X(s) = \frac{\mu}{\mu - js}, \qquad s \in \mathbb{R}.$$

- If $f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, $x \in \mathbb{R}$, then

$$C_X(s) = e^{-|s|}, \qquad s \in \mathbb{R}.$$

Properties of Characteristic Functions

- If $Y = aX + b$, then

$$C_Y(s) = e^{jbs}\, C_X(as), \qquad s \in \mathbb{R}.$$

Properties of Characteristic Functions

- If $Y = aX + b$, then
$$C_Y(s) = e^{jbs}\, C_X(as), \qquad s \in \mathbb{R}.$$

- If $X \perp\!\!\!\perp Y$, then
$$C_{X+Y}(s) = C_X(s)\, C_Y(s) \qquad \forall s \in \mathbb{R}.$$

  As a corollary, it follows that if X, Y are i.i.d. Cauchy, then $X + Y$ is also Cauchy (but with a different parameter).

Properties of Characteristic Functions

- If $Y = aX + b$, then
$$C_Y(s) = e^{jbs}\, C_X(as), \qquad s \in \mathbb{R}.$$

- If $X \perp\!\!\!\perp Y$, then
$$C_{X+Y}(s) = C_X(s)\, C_Y(s) \qquad \forall s \in \mathbb{R}.$$
As a corollary, it follows that if X, Y are i.i.d. Cauchy, then $X + Y$ is also Cauchy (but with a different parameter).

- If $M_X(t) < +\infty$ for all $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$, then
$$C_X(s) = M_X(js) \qquad \forall s \in \mathbb{R}.$$

Properties of Characteristic Functions

- If $C_X(s) = C_Y(s)$ for all $s \in \mathbb{R}$, then X and Y have the same CDF, i.e.,

$$F_X(x) = F_Y(x) \qquad \forall x \in \mathbb{R}.$$

Properties of Characteristic Functions

- If $C_X(s) = C_Y(s)$ for all $s \in \mathbb{R}$, then X and Y have the same CDF, i.e.,

$$F_X(x) = F_Y(x) \qquad \forall x \in \mathbb{R}.$$

- [Recovering moments from characteristic function]
  For $k \in \mathbb{N}$, if $\left| \frac{\mathrm{d}^k}{\mathrm{d}s^k} C_X(s) \Big|_{s=0} \right| < +\infty$, then

$$\mathbb{E}[X^k] = (-j)^k \left. \frac{\mathrm{d}^k}{\mathrm{d}s^k} C_X(s) \right|_{s=0}.$$

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

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

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

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

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

   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

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

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:

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