Probability theory I

24/12/21

Fact: (Fundamental theorem of Calculus)

If $F : [a,b] \to \mathbb{R}$ is continuously differentiable, then

$$\int_a^b F'(x)\, dx = F(b) - F(a).$$

Note: The continuity of F' is not necessary. For example

Let

$$F(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \in (0,1] \\ 0 & \text{if } x = 0 \end{cases}$$

$\lim\limits_{x \to 0^+} \dfrac{F(x) - F(0)}{x - 0}$ exists.

H.w.

Then F is differentiable on $[0,1]$.

$$F'(x) = \begin{cases} 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \text{if } x \in (0,1] \\ 0 & \text{if } x = 0 \end{cases}$$

Here F' is not continuous at $x=0$. (HW). But F' is integrable. We shall have

$$\int_0^1 F'(x)\,dx = F(1) - F(0). \qquad \square$$

Definition (Probability generating function):

Let X be a discrete random variable taking values $0, 1, 2, \cdots$. The probability generating function (PGF) of X is a function $f : [-1, 1] \to \mathbb{R}$ defined by

$$f(s) = \sum_{n=0}^{\infty} P(X=n) s^n, \qquad s \in [-1, 1].$$

Note: The PGF of X can also be represented as $|f(s)|$

$$f(s) = E(s^X) = \sum_{n=0}^{\infty} s^n P(X=n) = \sum_{n=0}^{\infty} a_n s^n$$

<u>Theorem:</u> The PGF f of X is continuous on $[-1,1]$ and infinitely differentiable on $(-1,1)$. Furthermore,

$$\mathbb{P}(X=n) = \frac{1}{n!} D^n f(0) = \frac{1}{n!} \frac{d^n}{ds^n} (f(s)) \Big|_{s=0} \quad \text{for all } n \geq 0 \text{ with}$$

$D^0 f \equiv f$.

<u>Proof:</u> Denote $p_n := \mathbb{P}(X=n)$, for all $n \geq 0$.

Since,

$$\limsup_{n \to \infty} p_n^{1/n} \leq 1,$$

It follows that the radius of convergence of the power series $\sum_{n=0}^{\infty} p_n s^n$

is atleast 1. Hence it follows that f is infinitely differentiable in $(-1,1)$. Furthermore for a fixed $n \geq 0$,

$$D^n f(s) = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1)\, p_k\, s^{k-n}, \qquad |s| < 1.$$

putting $s=0$, we get $D^n f(0) = n!\, p_n$.

HW: Show that f is continuous on $s=+1$, $s=-1$.

$$f(1) = \lim_{s\to 1^-} f(s) = \lim_{s\to 1^-} \sum_{n=0}^{\infty} p_n s^n = \sum_{n=0}^{\infty} \lim_{s\to 1^-} (p_n s^n) = 1$$

if $\limsup a_n^{1/n} \leq R$ then $\sum a_n s^n$ is convergent on $\left(-\frac{1}{R}, \frac{1}{R}\right)$. (check)!

f is unif cont on (a,b)

& $f(a) = \lim_{s\to a} f(s)$

$f(b) = \lim_{s\to b} f(s)$

f unf. cont on $[a,b]$.

<u>Corollary:</u> If X and Y are random variables taking values $0, 1, 2, \cdots$ and having the same PGF, then

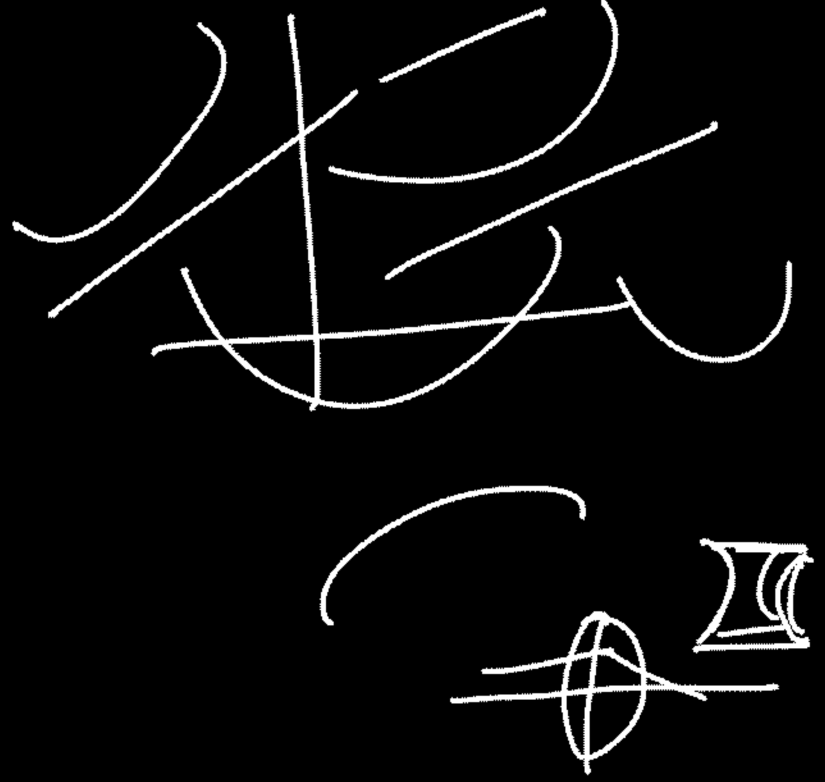
$$X \overset{d}{=} Y. \quad \left(\Rightarrow P(X=n) = P(Y=n) \quad \forall\ n \geq 0 \right).$$

<u>Pf:</u> $P(X=n) = D^n f(0) \cdot \frac{1}{n!} = P(Y=n)$ for all $n \geq 0$.

<u>Definition:</u> <u>(convex function)</u>:

A function f is convex on an interval $[a, b]$ if for any two points x_1 and x_2 in $[a, b]$ and any $0 < \lambda < 1$,

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2).$$



__Theorem:__ If f is a PGF of X, then $f(1)=1$ and f is convex on $[0,1]$. The convexity is strict if and only if $P(X\le 1) < 1$. (or $P(X>1) > 0$). [$\because P(X\le 1) + P(X>1) = 1$)

$1 = P(X\le 1) + P(X>1)$

$1 < 1 + P(X>1)$

$0 < P(X>1)$

__Proof:__ We have $f(s) = \sum_{n=0}^{\infty} p_n s^n$

So, $f'(s) = \sum_{n=1}^{\infty} n p_n s^{n-1}$

and $f''(s) = \sum_{n=2}^{\infty} n(n-1) p_n s^{n-2} \geq 0 \quad \forall\ s \geq 0.$

$f''(s) > 0$ if $\underbrace{p_0 + p_1}_{= \mathbb{P}(X \leq 1)} < 1$. $\square$

$$\sum_{n=2}^{\infty} n(n-1) p_n s^{n-2} > 0 \iff p_0 + p_1 < 1$$

$$\iff$$

Let us compute pgf's of some random variables.

Example: 1 Let $X \sim$ Binomial (n,p).

We know that if $X_1, X_2, \ldots, X_n$ are iid Bern(p). Then $X_1+X_2+\cdots+X_n \stackrel{d}{=} X$. (H.w.).

The pgf of X is

$$f_X(s) = E(s^X) = E\left(s^{X_1+X_2+\cdots+X_n}\right) = E\left(s^{X_1} s^{X_2} \cdots s^{X_n}\right)$$

$$= \left(E(s^{X_1})\right)^n \quad (\text{since } X_1, X_2, \ldots, X_n \text{ are iid}).$$

$$= \left(ps + (1-p)\right)^n.$$

② Let $X \sim \text{Poisson}(\lambda)$. Then

$$f_X(s) = E(s^X) = \sum_{n=0}^{\infty} s^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda s)^n}{n!}$$

$$= e^{-\lambda + \lambda s} = e^{\lambda(s-1)}$$

$$2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \text{cov}(X_i, X_j) \qquad \sum_{i=1}^{n} \text{var}(X_i)$$

$$2 \sum_{1 \le i < j \le n} \text{cov}(X_i; X_j)$$

③ Let $X \sim \text{Geometric}(p)$. Then,

$$f_X(s) = E(s^X) = \sum_{n=1}^{\infty} s^n q^{n-1} p = ps \sum_{n=1}^{\infty} (qs)^{n-1}$$

$$= ps \frac{1}{1-qs} = \frac{ps}{1-qs} \quad \text{where } q = 1-p.$$

Problem: (already done): Suppose X and Y are independent, $X \sim Poi(\lambda)$ and $Y \sim Poi(\mu)$. Find the distribution of $X+Y$.

Solution: $f_{X+Y}(s) = E\left(s^{X+Y}\right) = E(s^X)\,E(s^Y)$ ($\because$ X & Y are independent)

$$= e^{\lambda(s-1)}\, e^{\mu(s-1)}$$

$$= e^{(\lambda+\mu)(s-1)} = f_{Poi(\lambda+\mu)}(s).$$

Theorem: Let X be a discrete random variable with PGF $f_X(s)$ $(\equiv f(s))$.

Then,

1) $E(X) = f'(1)$

2) $E(\underbrace{X(X-1)\cdots\cdots(X-k+1)}_{k\text{th factorial moment}}) = f^{(k)}(1) = \frac{d^k}{ds^k} f(s) \Big|_{s=1}$

$E(X^k)$ is called the kth moment of X

Proof: 1) $f(s) = \sum_{n=0}^{\infty} s^n P(X=n)$ diff. w.r.t. s.

$f'(s) = \sum_{n=1}^{\infty} n s^{n-1} P(X=n)$

$$\therefore \; f'(1) = \sum_{n=1}^{\infty} n P(X=n) = E(X).$$

$$2) \quad f^{(k)}(s) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)\, s^{n-k} P(X=n)$$

$$\therefore \; f^{(k)}(1) = E\Big(X(X-1)\cdots(X-k+1)\Big) \qquad \square$$

Theorem: Suppose that $X_1, X_2, \dots, X_n$ are independent random variables and $Y = X_1 + X_2 + \cdots + X_n$. Then

$$f_Y(s) = \prod_{i=1}^{n} f_{X_i}(s) = f_{X_1}(s) f_{X_2}(s) \cdots f_{X_n}(s)$$

Proof: $f_Y(s) = E(s^Y) = E\left(s^{X_1+X_2+\cdots+X_n}\right) = E\left(s^{X_1} s^{X_2} \cdots s^{X_n}\right)$

$= E(s^{X_1})E(s^{X_2})\cdots E(s^{X_n})$ (by independence)

$= f_{X_1}(s) f_{X_2}(s) \cdots f_{X_n}(s)$.

Pointwise convergence: $\varepsilon > 0$, $\exists\, n \in \mathbb{N}$

$|f_n(x) - f(x)| < \varepsilon \quad \forall\, n \geq n_0$.

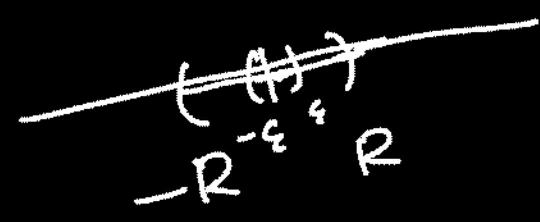
unif cv: $\varepsilon > 0$, $\exists\, n \in \mathbb{N}$

$|f_n(x) - f(x)| < \varepsilon \quad \forall\, n \geq n_0$ & $x \in D$. $\quad f_n : D \to \mathbb{R}$

Recall:

Fact: If two power serieses agree on an interval containing 0, however small, then all terms of the two series are equal.

$$\sum a_n x^n \to R_1$$
$$\text{"}$$
$$\sum b_n x^n \to R_2$$



Corollary: If two random variables have the same PGF in some interval containing 0, then the two random variables have the same distribution. So, every PGF corresponds to a unique distribution.