第十九章 生成函数与卷积

19.1 动机

19.2 定义

定义 19.2.1 (生成函数) 已知序列 $\{a_n\}_{n=0}^{\infty}$. 它的生成函数被定义为

$$G_a(s) = \sum_{n=0}^{\infty} a_n s^n,$$

其中, s 是使这个和收敛的任意数.

在概率论中, a_n 可以看作:

① 非负随机变量等于n的概率.

② 随机变量的矩.

例. 求斐波那契数列 $(0, 1, 1, 2, \cdots)$ 的通项公式.

解:
$$G_a(s) = \sum_{n=0}^{\infty} a_n s^n$$
$$= 0 + s + \sum_{n=2}^{\infty} a_n s^n$$
$$= s + \sum_{n=2}^{\infty} (a_{n-1} + a_{n-2}) s^n$$

$$= s + s\sum_{n=2}^{\infty} a_{n-1} s^{n-1} + s^2 \sum_{n=2}^{\infty} a_{n-2} s^{n-2}$$

$$= s + s\sum_{m=1}^{\infty} a_m s^m + s^2 \sum_{m=1}^{\infty} a_m s^m$$

$$= s + s \cdot G_a(s) + s^2 G_a(s)$$

$$\Rightarrow \quad G_a(s) = \frac{s}{1-s-s^2}$$

$$= \frac{a}{1-As} + \frac{b}{1-Bs}$$

$$= \frac{1}{\sqrt{5}} \cdot \frac{1}{1-\frac{1+\sqrt{5}}{2}s} - \frac{1}{\sqrt{5}} \cdot \frac{1}{1-\frac{1-\sqrt{5}}{2}s}$$

$$= \frac{1}{\sqrt{5}} \left[\sum_{n=0}^{\infty} \left(\frac{1+\sqrt{5}}{2}s\right)^n - \sum_{n=0}^{\infty} \left(\frac{1-\sqrt{5}}{2}s\right)^n \right]$$

$$= \sum_{n=0}^{\infty} \left[\frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n \right] s^n$$

$$\Leftrightarrow \quad a_n = \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n$$

19.3 生成函数的唯一性和收敛性

根据序列$\{a_n\}_{n=0}^{\infty}$的具体形式，生成函数$G_a(s)$可能对全部s成立，部分s成立甚至仅对$s=0$成立。

$\Rightarrow$ 要保证生成函数收敛。

定理 19.3.1 (序列生成函数的唯一性) 设$\{a_n\}_{n=0}^{\infty}$和$\{b_n\}_{n=0}^{\infty}$是生成函数分别为$G_a(s)$和$G_b(s)$的两个数列。当$|s|<\delta$时，$G_a(s)$和$G_b(s)$均收敛。那么，这两个序列相等 (即对于所有的i，均有$a_i=b_i$)，当且仅当对于所有的$|s|<\delta$均有$G_a(s)=G_b(s)$。通过对生成函数求微分，我们可以重新得到序列：$a_n=\frac{1}{n!}\frac{\mathrm{d}^n G_a(s)}{\mathrm{d}s^n}$.

序列由生成函数唯一确定

与概率有关的生成函数通常是收敛的

19.4 卷积 I：离散型随机变量

定义 19.4.1 (序列的卷积) 已知两个序列$\{a_m\}_{m=0}^{\infty}$和$\{b_n\}_{n=0}^{\infty}$。它们的卷积被定义为下面这个新序列$\{c_k\}_{k=0}^{\infty}$

$$c_k = a_0b_k + a_1b_{k-1} + \cdots + a_{k-1}b_1 + a_kb_0 = \sum_{l=0}^{k} a_l b_{k-l}.$$

通常情况下，我们把它记作$c=a*b$.

定义来自于多项式乘法

若 $f(x) = \sum_{k=0}^{\infty} a_k \cdot x^k$, $g(x) = \sum_{k=0}^{\infty} b_k \cdot x^k$

那么 $h(x) = f(x) g(x) = \sum_{k=0}^{\infty} c_k x^k$

引理 19.4.2 设 $G_a(s)$ 是 $\{a_m\}_{m=0}^{\infty}$ 的生成函数, $G_b(s)$ 是 $\{b_n\}_{n=0}^{\infty}$ 的生成函数, 那么 $c = a * b$ 的生成函数是 $G_c(s) = G_a(s)G_b(s)$.

定义 19.4.3 (**概率生成函数**) 设 X 是一个取整数值的离散型随机变量. 设 $G_X(s)$ 是 $\{a_m\}_{m=-\infty}^{\infty}$ 的生成函数, 其中 $a_m = \text{Prob}(X = m)$. 那么 $G_X(s)$ 被称为概率生成函数. 如果仅当 X 取整数时概率不为零, 那么计算 $G_X(s)$ 的一个非常有用的方法就是

$$G_X(s) = \mathbb{E}[s^X] = \sum_{m=-\infty}^{\infty} s^m \text{Prob}(X = m).$$

更一般地, 如果 X 在一个至多可数的集合 $\{x_m\}$ 中取值的概率不为零, 那么

$$G_X(s) = \mathbb{E}[s^X] = \sum_{m} s^{x_m} \text{Prob}(X = x_m).$$

定理 19.4.4 设 $X_1, \cdots, X_n$ 是相互独立且都取非负整数值的离散型随机变量, 它们的概率生成函数分别是 $G_{X_1}(s), \cdots, G_{X_n}(s)$. 于是

$$G_{X_1+\cdots+X_n}(s) = G_{X_1}(s) \cdots G_{X_n}(s).$$

19.4.4 证明：

$n=2$ 的情况：

$$a_m = Prob(x_1 = m)$$

$$b_n = Prob(x_2 = n)$$

$$c = x_1 + x_2$$

$$G_c(s) = \sum_{l=0}^{k} Prob(x_1 = l)\, Prob(x_2 = k-l)$$

$$\Rightarrow c = a * b$$

有 $G_{x_1+x_2}(s) = G_{x_1}(s)\, G_{x_2}(s)$

$n=3,4,5,\ldots$ 时由分组证明法可解决

序列由生成函数唯一确定，而随机变量之和的生成函数是各生成函数的乘积。

如果恰好得到了这个乘积，就立马得到了和的概率密度函数。

例1. 求服从参数为5和7的泊松分布的随机变量 X_1、X_2 的和的pdf.

$$f_{X_{1,2}}(X=n) = \begin{cases} \dfrac{\lambda^n \cdot e^{-\lambda}}{n!} & n \in N \\ 0 & \text{其它} \end{cases}$$

$$G_X(s) = \sum_{n=0}^{\infty} Prob(X=n) s^n$$

$$= \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} s^n$$

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

$$= e^{-\lambda} \cdot e^{\lambda s}$$

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

$$G_{X_1+X_2}(s) = G_{X_1}(s)\, G_{X_2}(s)$$

$$= e^{5(s-1)} \cdot e^{7(s-1)}$$

$$= e^{12(s-1)}$$

$\Rightarrow$ $X_1 + X_2$ 服从参数为12的泊松分布

19.5 连续型随机变量

定义 19.5.1 (**概率生成函数**) 设 X 是一个连续型随机变量, 其概率密度函数为 f, 那么

$$G_X(s) = \int_{-\infty}^{\infty} s^x f(x)\mathrm{d}x$$

是 X 的概率生成函数.

定义 19.5.2 (**函数卷积**) 两个函数 f_1 和 f_2 的卷积, 即 $f_1 * f_2$ 被定义为

$$(f_1 * f_2)(x) = \int_{-\infty}^{\infty} f_1(t)f_2(x-t)\mathrm{d}t.$$

如果 f_i 都是概率密度函数, 那么上述积分收敛.

函数卷积满足:

① 交换律

② 结合律

例. 对于服从参数为 λ 的指数分布的随机变量 X, 有

$$f_X(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & x \geq 0 \\ 0 & \text{其它} \end{cases}$$

$$G_X(s) = \int_0^{+\infty} s^x \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx$$

$$= \frac{1}{\lambda} \int_0^{+\infty} s^x e^{-\frac{x}{\lambda}} dx$$

$$= \frac{1}{\lambda} \int_0^{+\infty} e^{-\frac{x}{\lambda} + x \cdot \ln s} dx$$

$$= \frac{1}{\lambda} \cdot \frac{1}{\frac{1}{\lambda} - \ln s} \int_0^{+\infty} e^{-u} du$$

$$= \frac{1}{1 - \lambda \ln s}$$

$$\Rightarrow G_X(s) = \begin{cases} (1-\lambda \ln s)^{-1} & \ln s < \frac{1}{\lambda} \\ 0 & \text{其它} \end{cases}$$

例2. 设 X_1, X_2 分别是服从参数为 2, 5 的指数分布的随机变量. 探究 $X = X_1 + X_2$?

$$G_X(s) = G_{X_1}(s)\, G_{X_2}(s)$$

$$= (1-5\ln s)^{-1}(1-2\ln s)^{-1}$$

找不到λ使得上式变成服从指数分布的生成函数.

19.6 矩母函数的性质与定义.

定义 19.6.2 (矩母函数) 设 X 是一个概率密度函数为 f 的随机变量. X 的矩母函数记作 $M_X(t)$, 其定义为 $M_X(t) = \mathbb{E}[\mathrm{e}^{tX}]$. 具体地说, 如果 X 是离散的, 那么

$$M_X(t) = \sum_{m=-\infty}^{\infty} \mathrm{e}^{tx_m} f(x_m),$$

如果 X 是连续的, 那么

$$M_X(t) = \int_{-\infty}^{\infty} \mathrm{e}^{tx} f(x)\mathrm{d}x.$$

注意, $M_X(t) = G_X(\mathrm{e}^t)$, 或等价的 $G_X(s) = M_X(\log s)$.

定理 19.6.3 设 X 是一个随机变量, μ'_k 是它的矩.

(1) 我们有

$$M_X(t) = 1 + \mu'_1 t + \frac{\mu'_2 t^2}{2!} + \frac{\mu'_3 t^3}{3!} + \cdots;$$

特别是, $\mu'_k = \mathrm{d}^k M_X(t)/\mathrm{d}t^k \Big|_{t=0}$.

(2) 设 α 和 β 是常数. 那么

$$M_{\alpha X+\beta}(t) = \mathrm{e}^{\beta t} M_X(\alpha t).$$

有用的特殊情形是 $M_{X+\beta}(t) = \mathrm{e}^{\beta t} M_X(t)$ 和 $M_{\alpha X}(t) = M_X(\alpha t)$. 当证明中心极限定理时, $M_{(X+\beta)/\alpha}(t) = \mathrm{e}^{\beta t/\alpha} M_X(t/\alpha)$ 也很有用.

(3) 设 X_1 和 X_2 是两个相互独立的随机变量. 当 $|t| < \delta$ 时, 它们的矩母函数 $M_{X_1}(t)$ 和 $M_{X_2}(t)$ 均收敛. 于是

$$M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t).$$

更一般地, 如果 $X_1, \cdots, X_N$ 是相互独立的随机变量. 当 $|t| < \delta$ 时, 它们的矩母函数 $M_{X_i}(t)$ 均收敛. 那么

$$M_{X_1+\cdots+X_N}(t) = M_{X_1}(t) M_{X_2}(t) \cdots M_{X_N}(t).$$

如果所有随机变量都有同一个矩母函数 $M_X(t)$, 那么上式右端就变成了 $M_X(t)^N$.

证明:

以连续为例.

(1)

$$M_X(t) = E(e^{tX})$$

$$= \int \sum_{i=0}^{\infty} \frac{(tx)^i}{i!} f_X(x)\,dx$$

$$= \sum_{i=0}^{\infty} \frac{t^i}{i!} \int x^i f(x)\,dx$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \cdots$$

(2)

$$M_{\alpha X+\beta}(t) = E(e^{t(\alpha X+\beta)})$$

$$= E(e^{\alpha tX} \cdot e^{\beta t})$$

$$= e^{\beta t} E(e^{\alpha tX})$$

$$= e^{\beta t} M_X(\alpha t)$$

(3) $$M_{X_1+X_2}(t) = E(e^{t(X_1+X_2)})$$

$$= E(e^{tX_1} \cdot e^{tX_2})$$

X_1、X_2 为独立的随机变量，

e^{tX_1}、e^{tX_2} 也相互独立.

$$\Rightarrow M_{X_1+X_2}(t) = E(e^{tX_1}) \cdot E(e^{tX_2})$$

$$= M_{X_1}(t) \cdot M_{X_2}(t)$$

定理 19.6.5 (离散型随机变量的矩母函数的唯一性) 设 X 和 Y 是两个取非负整数值 (即仅当在 $\{0,1,2,\cdots\}$ 中取值时, 概率才不为 0) 的离散型随机变量, 它们的矩母函数 $M_X(t)$ 和 $M_Y(t)$ 在 $|t|<\delta$ 时收敛. 那么, X 和 Y 是同分布的, 当且仅当存在一个 $r>0$, 使得当 $|t|<r$ 时 $M_X(t)=M_Y(t)$.

离散型随机变量是由矩母函数唯一确定的.

证:

由引理 19.3.1, 两序列相等, 当且仅当两生成函数相等.

往证: $M_X(t) = M_Y(t)$

$$\Longleftrightarrow G_X(t) = G_Y(t)$$

$$E(e^{tX}) = E(e^{tY})$$

令 $e^t = s$, 有

$$E(s^X) = E(s^Y)$$

$$\int_{-\infty}^{+\infty} s^x f(x)\, dx = \int_{-\infty}^{+\infty} s^y f(y)\, dy$$

$\Longleftrightarrow$ 两生成函数相等.

得证.

例. 设随机变量 X_1、X_2 服从参数为 λ_1、λ_2 的泊松分布. 求证 X_1+X_2 服从参数为 $\lambda_1+\lambda_2$ 的泊松分布.

$$\mathrm{Prob}(X=n)=\begin{cases} \lambda^n \cdot \frac{e^{-\lambda}}{n!} & n \in N \\ 0 & \text{其它} \end{cases}$$

矩母函数 $M_{X_{1,2}} = E(e^{tn})$

$$= \sum_{n=0}^{\infty} e^{tn} f(n)$$

$$= e^{-\lambda} \sum_{n=0}^{\infty} e^{tn} \cdot \lambda^n \cdot \frac{1}{n!}$$

$$= e^{-\lambda} \cdot e^{\lambda e^t}$$

$$= e^{(e^t-1)\cdot\lambda}$$

$$M_{X_1+X_2} = M_{X_1} \cdot M_{X_2}$$

$$= e^{(e^t-1)\lambda_1} \cdot e^{(e^t-1)\lambda_2}$$

$$= e^{(e^t-1)(\lambda_1+\lambda_2)}$$

$\Rightarrow$ X_1+X_2 是服从参数为 $\lambda_1+\lambda_2$ 的泊松分布随机变量

存在具有相同矩的不同概率分布. 换句话说, 知道所有的矩并不总是可以唯一确定概率分布.

19.7 矩母函数的应用

如果可以求出一个随机变量的矩母函数，则可以通过求导得到任一阶的矩

例. 求泊松分布的均值 μ. 方差 σ^2

法一：微分恒等式法.

法二：

$$f_X(X=n) = \begin{cases} \lambda^n \cdot \dfrac{e^{-\lambda}}{n!} & n \in N \\ 0 & 其它 \end{cases}$$

$$M_X(t) = \sum_{n=0}^{\infty} e^{tx} \lambda^n \cdot \frac{e^{-\lambda}}{n!} = e^{(e^t-1)\lambda}$$

$$\frac{d}{dt} M_X(t) \Big|_{t=0} = e^{(e^t-1)\lambda} \cdot \lambda \cdot e^t \Big|_{t=0}$$

$$= \lambda$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0}$$

$$= \cdots \Big|_{t=0}$$

$$= \lambda^2 + \lambda$$

$$\Rightarrow \quad \sigma^2 = E(X^2) - E^2(X)$$

$$= \lambda.$$