

文章编号: 0427-7104(2011)01-

运用 Newton 微元法求解概率密度函数

朱慧敏

(复旦大学 数学科学学院, 上海 200433)

摘 要: 利用 Newton 微元法求连续型随机变量函数的概率密度, 可使其与求离散型随机变量函数的概率分布的计算方法相统一, 因而这种方法更直观简单, 尤其在求解复杂的随机变量函数的概率密度时, 此方法更胜一筹.

关键词: Newton 微元法; 概率密度函数; 连续型随机变量; 随机变量函数

中图分类号: O 1

文献标志码: A

在概率论中求随机变量函数的分布是一个很基本的问题, 即设 f 是已知函数, 若随机变量 ξ 和 η 满足 $\eta=f(\xi)$, 则称 η 是 ξ 的函数. 所谓求随机变量函数的分布, 就是已知随机变量 ξ 的分布, 要求 η 的分布.

若 ξ 是离散型随机变量, 其分布为 $P(\xi=x_k)=p_k, k=1, 2, \dots; \sum_{k=1}^{\infty} p_k=1$. 记 η 的可能取值为 $\{y_l, l=1, 2, \dots\}=\{f(x_k) \mid k=1, 2, \dots\}$, 则 η 的分布为

$$P(\eta=y_l)=\sum_{f(x_k)=y_l} p_k \quad l=1, 2, \dots .$$

若 ξ 和 η 都是连续型随机变量, 为了求 η 的密度函数, 常用的办法是先求出累积分布函数, 然后求导得到密度函数. 在本文中, 利用 Newton 微元法, 对连续型随机变量可以计算其在一点的微概率, 记 $\varphi_\xi(x)$ 为 ξ 的概率密度, 则 $P(\xi=x)=\varphi_\xi(x) |dx|$, 其中 $|dx|$ 是 x 轴上一点的长度, 由此, ξ 在一个区间上的概率:

$$P(\xi \in [a, b]) = \int_a^b \varphi_\xi(x) |dx| = \int_a^b \varphi_\xi(x) dx,$$

即可看成是它在每一点微概率的总和. 注意, 微概率的积分与积分上下限的顺序无关:

$$P(\xi \in [a, b]) = \int_b^a \varphi_\xi(x) |dx| = -\int_b^a \varphi_\xi(x) dx = \int_a^b \varphi_\xi(x) dx.$$

按照这样的想法, 我们就可以像处理离散型随机变量函数那样, 来求连续性随机变量函数 $\eta(\eta=f(\xi))$ 的概率密度了:

$$\varphi_\eta(y) |dy| = P(\eta=y) = P(f(\xi)=y).$$

1 求解一维连续型随机变量的函数的概率密度

当 ξ, η 均为连续型随机变量时, 若 ξ 的概率密度为 $\varphi_\xi(x)$, $f(x)$ 是一个连续函数, $\eta=f(\xi)$, 则为了求出 η 的概率密度 $\varphi_\eta(x)$, 通常的方法^[1]是:

1) 先求 $\eta=f(\xi)$ 的分布函数: $P(\eta \leq y) = F_\eta(y) = \int_{f(x) \leq y} \varphi_\xi(x) dx$;

2) 然后求导得出 η 的概率密度: $\varphi_\eta(y) = F'_\eta(y)$.

然而, 在某些情况下, $\{x | f(x) \leq y\}$ 是一个比较复杂的集合, 上述的方法就必须先在这复杂的集合上积分, 然后再求导以完成最终的计算. 这是一种比较繁琐的过程, 尤其是对于比较复杂的问题, 更是让人望而却步.

注意到积分与求导是互逆的运算, 利用 Newton 微元法可以避免这种“重复”计算, 如同离散型随机变量那样, 通过相应点的概率密度的相加直接得到所求的结果.

2 用牛顿微元法求解概率密度(一)

定理 1 设 ξ 是连续型随机变量, 其概率密度记为 $\varphi_\xi(x)$, $\eta = f(\xi)$. 设 $f(x)$ 满足: 整个实轴除去一个零集后, 可划分为若干个两两互不重叠的区间 (a_k, b_k) , 使得在每一个区间 (a_k, b_k) 上, $y = f(x)$ 严格单调并且有惟一具有连续导数的反函数 $x = g_k(y)$, 区间 (α_k, β_k) 为值域. 则 η 也是连续型随机变量, 其概率密度 $\varphi_\eta(y)$ 为

$$\varphi_\eta(y) = \begin{cases} \sum_{\{k | y \in (\alpha_k, \beta_k)\}} \varphi_\xi(x_k(y)) |g'_k(y)| & y \in \bigcup_{k=1}^{\infty} (\alpha_k, \beta_k), \\ 0 & y \notin \bigcup_{k=1}^{\infty} (\alpha_k, \beta_k). \end{cases} \quad (1)$$

证 利用 Newton 微元法求 $\varphi_\eta(y)$, 设 $y \in \bigcup_{k=1}^{\infty} (\alpha_k, \beta_k)$,

$$\varphi_\eta(y) |dy| = P(\eta=y) = P(f(\xi)=y). \quad (2)$$

由已知, 当 $y \in (\alpha_k, \beta_k)$ 时, $y = f(x)$ 在 $x \in (a_k, b_k)$ 上有对应的反函数 $x = g_k(y)$, 故

$$\begin{aligned} P(f(\xi)=y) &= \sum_{\{k | y \in (\alpha_k, \beta_k)\}} P(\xi = x_k(y)) = \\ &= \sum_{\{k | y \in (\alpha_k, \beta_k)\}} \varphi_\xi(g_k(y)) |d(g_k(y))| = \sum_{\{k | y \in (\alpha_k, \beta_k)\}} \varphi_\xi(g_k(y)) |g'_k(y)| |dy|, \end{aligned} \quad (3)$$

则由(2)和(3)式, 约去 $|dy|$ 得: $\varphi_\eta(y) = \sum_{\{k | y \in (\alpha_k, \beta_k)\}} \varphi_\xi(g_k(y)) |g'_k(y)|$.

当 $y \notin \bigcup_{k=1}^{\infty} (\alpha_k, \beta_k)$ 时, $P(f(\xi)=y) = 0$, 故 $\varphi_\eta(y) = 0$. (1) 式得证.

例 1 设连续型随机变量 $\xi \sim \text{Exp}(\lambda)$, 即 ξ 的概率密度为

$$\varphi_\xi(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0, \\ 0 & x \leq 0, \end{cases}$$

求随机变量 $\eta = \tan \xi$ 的概率密度 $\varphi_\eta(y)$.

解 如果直接利用定理 1 中的(1)式来计算 $\varphi_\eta(y)$, 需要找出定理 1 中的各个单调区间 (a_k, b_k) , 这既繁琐又不必要^[2]. 为此, 直接用 Newton 微元法计算 $\varphi_\eta(y)$, 就不需要考虑区间的变换关系, 而只要考虑点的变换关系, 既简便又清晰.

当 $y > 0$ 时, $\tan x = y$ 满足 $x > 0$ 的解为: $x = \arctan y + k\pi, k = 0, 1, 2, \dots$, 则

$$\begin{aligned} \varphi_\eta(y) \cdot |dy| &= P(\eta=y) = P(\tan \xi=y) = P(\xi = \arctan y + k\pi, k = 0, 1, 2, \dots) = \\ &= \sum_{k=0}^{\infty} P(\xi = \arctan y + k\pi) = \sum_{k=0}^{\infty} \lambda e^{-\lambda(\arctan y + k\pi)} |d \arctan y| = \\ &= \frac{1}{1+y^2} \lambda e^{-\lambda \arctan y} \sum_{k=0}^{\infty} e^{-k\pi\lambda} |dy|, \end{aligned}$$

所以,

$$\varphi_\eta(y) = \frac{\lambda e^{-\lambda \arctan y}}{1 - e^{-\pi\lambda}} \cdot \frac{1}{1+y^2}.$$

当 $y < 0$ 时, $\tan x = y$ 满足 $x > 0$ 的解为: $x = \arctan y + k\pi, k = 1, 2, \dots$, 则

$$\begin{aligned} \varphi_{\eta}(y) \cdot |dy| &= P(\eta = y) = P(\tan \xi = y) = P(\xi = \arctan y + k\pi, k = 1, 2, \dots) = \\ &= \sum_{k=1}^{\infty} P(\xi = \arctan y + k\pi) = \sum_{k=1}^{\infty} \lambda e^{-\lambda(\arctan y + k\pi)} |d \arctan y| = \\ &= \frac{1}{1+y^2} \lambda e^{-\lambda \arctan y} \sum_{k=1}^{\infty} e^{-k\pi\lambda} |dy|, \end{aligned}$$

$$\text{所以, } \varphi_{\eta}(y) = \frac{\lambda e^{-\lambda(\pi + \arctan y)}}{1 - e^{-\pi\lambda}} \cdot \frac{1}{1+y^2}.$$

当 $y = 0$ 时, 由于 $\varphi_{\eta}(y)$ 在零集上的值不影响概率的计算, 故取 $\varphi_{\eta}(0) = \varphi_{\eta}(0+)$, 所以,

$$\varphi_{\eta}(y) = \begin{cases} \frac{\lambda e^{-\lambda \arctan y}}{1 - e^{-\pi\lambda}} \cdot \frac{1}{1+y^2} & y \geq 0 \\ \frac{\lambda e^{-\lambda(\pi + \arctan y)}}{1 - e^{-\pi\lambda}} \cdot \frac{1}{1+y^2} & y < 0 \end{cases} \quad k = 0, 1, 2, \dots.$$

3 用牛顿微元法求解联合概率密度(二)

概率密度反映的是连续型随机变量取 x 邻域内值的概率大小, 因此, 它是了解连续型随机变量的有力工具. 在实际问题中, 又常常需要计算多个连续型随机变量函数的概率密度.

计算二维连续型随机变量函数的概率密度时, 我们也可以计算其在一点的微概率:

$$P(\xi = x, \eta = y) = \varphi(x, y) dx dy,$$

其中 $\varphi(x, y)$ 为连续型随机变量 ξ 和 η 的联合概率密度, 则

$$P((\xi, \eta) \in D) = \iint_D \varphi(x, y) dx dy$$

为二维连续型随机变量在一个区域上的概率, 也可看成是它在每一点微概率的总和.

定理 2 设 ξ 和 η 为连续型随机变量, 其联合概率密度为 $\varphi(x, y)$. 随机变量函数 $\zeta = f(\xi, \eta), \tau = g(\xi, \eta)$. 设 $(u, v) = (f(x, y), g(x, y))$ 满足: (x, y) 平面除去一个零集后, 可分解成互不相交的开区域 D_k 的并, 在每个 D_k 上 $(u, v) = (f(x, y), g(x, y))$ 都有连续可微的反函数 $(x, y) = (s_k(u, v), t_k(u, v))$, 且其值域为 R_k . 则随机变量函数的联合概率密度 $\varphi_{(\zeta, \tau)}(u, v)$ 为

$$\varphi_{(\zeta, \tau)}(u, v) = \begin{cases} \sum_{\{k | (u, v) \in R_k\}} \varphi_{(\xi, \eta)}(s_k(u, v), t_k(u, v)) \left| \frac{D(s_k(u, v), t_k(u, v))}{D(u, v)} \right| & (u, v) \in \bigcup_{k=1}^{\infty} R_k, \\ 0 & \text{其他.} \end{cases}$$

证 利用 Newton 微元法求 $\varphi_{(\zeta, \tau)}(u, v)$, 设 $(u, v) \in \bigcup_{k=1}^{\infty} R_k$, 则

$$\varphi_{(\zeta, \tau)}(u, v) du dv = P(\zeta = u, \tau = v) = P(f(\xi, \eta) = u, g(\xi, \eta) = v).$$

由已知当 $(u, v) \in R_k$ 时, $(u, v) = (f(x, y), g(x, y))$ 在 $(x, y) \in D_k$ 上有对应的反函数 $(x, y) = (s_k(u, v), t_k(u, v))$, 故

$$P(f(\xi, \eta) = u, g(\xi, \eta) = v) = \sum_{\{k | (u, v) \in R_k\}} P(\xi = s_k(u, v), \eta = t_k(u, v)) =$$

$$\sum_{\{k | (u, v) \in R_k\}} \varphi_{(\xi, \eta)}(s_k(u, v), t_k(u, v)) \left| \frac{D(s_k(u, v), t_k(u, v))}{D(u, v)} \right| du dv,$$

$$\text{则 } \varphi_{(\zeta, \tau)}(u, v) = \begin{cases} \sum_{\{k | (u, v) \in R_k\}} \varphi_{(\xi, \eta)}(s_k(u, v), t_k(u, v)) \left| \frac{D(s_k(u, v), t_k(u, v))}{D(u, v)} \right| & (u, v) \in \bigcup_{k=1}^{\infty} R_k, \\ 0 & \text{其他.} \end{cases}$$

例2 设二维连续型随机变量 (ξ, η) 的联合概率密度为 $\varphi(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}$, $\zeta = \xi^2 + \eta^2$, 求 ζ 的概率密度.

解 用增加变量法, 设 $\begin{cases} \zeta = \xi^2 + \eta^2, \\ \tau = \xi. \end{cases}$

先用 Newton 微元法求 (ζ, τ) 的联合概率密度为 $\varphi_{(\zeta, \tau)}(u, v)$.

当 $u - v^2 > 0$ 时, $\begin{cases} u = x^2 + y^2 \\ v = x \end{cases}$ 有2个解 $\begin{cases} x = v \\ y = \sqrt{u - v^2} \end{cases}$ 和 $\begin{cases} x = v \\ y = -\sqrt{u - v^2} \end{cases}$, 所以

$$\varphi_{(\zeta, \tau)}(u, v) du dv = P(\zeta = u, \tau = v) = P(\xi^2 + \eta^2 = u, \xi = v) =$$

$$P(\xi = v, \eta = \sqrt{u - v^2}) + P(\xi = v, \eta = -\sqrt{u - v^2}) =$$

$$\varphi_{(\xi, \eta)}(v, \sqrt{u - v^2}) \left| \frac{D(v, \sqrt{u - v^2})}{D(u, v)} \right| du dv + \varphi_{(\xi, \eta)}(v, -\sqrt{u - v^2}) \left| \frac{D(v, -\sqrt{u - v^2})}{D(u, v)} \right| du dv.$$

约去 $du dv$ 得

$$\varphi_{(\zeta, \tau)}(u, v) = \frac{1}{2\pi} e^{-\frac{u}{2}} \cdot \frac{1}{\sqrt{u - v^2}}.$$

当 $u - v^2 < 0$ 时, $\varphi_{(\zeta, \tau)}(u, v) = 0$.

当 $u - v^2 = 0$ 时, $\varphi_{(\zeta, \tau)}(u, v)$ 的值可任取.

当 $u > 0$ 时, ζ 的边缘概率密度为

$$\varphi_{\zeta}(u) = \int_{-\infty}^{+\infty} \varphi_{(\zeta, \tau)}(u, v) dv = \int_{-\sqrt{u}}^{\sqrt{u}} \frac{1}{2\pi} e^{-\frac{u}{2}} \cdot \frac{1}{\sqrt{u - v^2}} dv =$$

$$\frac{1}{2\pi} e^{-\frac{u}{2}} \arcsin \frac{v}{\sqrt{u}} \Big|_{-\sqrt{u}}^{\sqrt{u}} = \frac{1}{2\pi} e^{-\frac{u}{2}} \cdot \pi =$$

$$\frac{1}{2} e^{-\frac{u}{2}};$$

当 $u \leq 0$ 时, $\varphi_{\zeta}(u) = 0$.

所以, ζ 的概率密度为

$$\varphi_{\zeta}(u) = \begin{cases} \frac{1}{2} e^{-\frac{u}{2}} & u \geq 0, \\ 0 & u < 0. \end{cases}$$

例3 设 $\xi_1, \xi_2, \dots, \xi_n$ 是 n 个独立同分布的连续型随机变量, 密度函数均为 $\varphi(x)$, 分布函数为 $F(x)$, $F'(x) = \varphi(x)$. 将 $\xi_1, \xi_2, \dots, \xi_n$ 按从小到大的顺序重排队得 n 个顺序随机变量: $\eta_1 \leq \eta_2 \leq \dots \leq \eta_n$. 求 1) $\eta_1, \eta_2, \dots, \eta_n$ 的联合概率密度 $g(y_1, y_2, \dots, y_n)$; 2) 第 k 个顺序随机变量 η_k 的边缘密度函数 $g_k(y)$.

解 1) 用 Newton 微元法求 $g(y_1, y_2, \dots, y_n)$.

为了方便起见记各顺序随机变量的变换函数为 $\eta_k = f_k(\xi_1, \xi_2, \dots, \xi_n)$, $k = 1, 2, \dots, n$.

①当 $y_1 < y_2 < \dots < y_n$ 时, $\begin{cases} y_1 = f_1(x_1, x_2, \dots, x_n) \\ y_2 = f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ y_n = f_n(x_1, x_2, \dots, x_n) \end{cases}$ 有 $n!$ 个解 (x_1, x_2, \dots, x_n) , 每个解都是

(y_1, y_2, \dots, y_n) 的一个全排列, 因此,

$$g(y_1, y_2, \dots, y_n) dy_1 dy_2 \cdots dy_n = P(\eta_1 = y_1, \eta_2 = y_2, \dots, \eta_n = y_n) =$$

$$P(f_1(\xi_1, \xi_2, \dots, \xi_n) = y_1, f_2(\xi_1, \xi_2, \dots, \xi_n) = y_2, \dots, f_n(\xi_1, \xi_2, \dots, \xi_n) = y_n) =$$

$$P(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n, (x_1, x_2, \dots, x_n) \text{为} (y_1, y_2, \dots, y_n) \text{的一个全排列}) =$$

$$\sum \varphi(x_1)\varphi(x_2)\cdots\varphi(x_n) \left| \frac{D(x_1, x_2, \dots, x_n)}{D(y_1, y_2, \dots, y_n)} \right| dy_1 dy_2 \cdots dy_n$$

(求和范围为 (y_1, y_2, \dots, y_n) 的全排列 (x_1, x_2, \dots, x_n)).

当 (x_1, x_2, \dots, x_n) 为 (y_1, y_2, \dots, y_n) 的一个全排列, 有 $\left| \frac{D(x_1, x_2, \dots, x_n)}{D(y_1, y_2, \dots, y_n)} \right| = 1$. 所以,

$$g(y_1, y_2, \dots, y_n) = n! \varphi(y_1) \varphi(y_2) \cdots \varphi(y_n).$$

② 当 $y_1 \leq y_2 \leq \cdots \leq y_n$ 时, 且不等式中至少有一个成立等式, 这些点组成 R^n 中的零集, 故密度函数 $g(y_1, y_2, \dots, y_n)$ 的值可任取.

③ 当 $y_1 \leq y_2 \leq \cdots \leq y_n$ 不成立时, $g(y_1, y_2, \dots, y_n) = 0$.

$$\text{所以, } g(y_1, y_2, \dots, y_n) = \begin{cases} n! \varphi(y_1) \varphi(y_2) \cdots \varphi(y_n) & y_1 \leq y_2 \leq \cdots \leq y_n, \\ 0 & \text{其他.} \end{cases}$$

$$\begin{aligned} 2) \quad g_k(y) &= \int \cdots \int_{R^{n-1}} g(y_1, y_2, \dots, y_n) dy_1 \cdots dy_{k-1} dy_{k+1} \cdots dy_n = \\ &= \int \cdots \int_{y_1 < \cdots < y_{k-1} < y_{k+1} < \cdots < y_n} \varphi(y_1) \cdots \varphi(y_{k-1}) \varphi(y) \varphi(y_{k+1}) \cdots \varphi(y_n) dy_1 \cdots dy_{k-1} dy_{k+1} \cdots dy_n = \\ &= n! \varphi(y) \left(\int_0^{y_k} \varphi(y_{k-1}) dy_{k-1} \int_0^{y_{k-1}} \varphi(y_{k-2}) dy_{k-2} \cdots \int_0^{y_2} \varphi(y_1) dy_1 \right) \cdot \\ &= \left(\int_{y_k}^{+\infty} \varphi(y_n) dy_n \int_{y_k}^{y_n} \varphi(y_{n-1}) dy_{n-1} \int_{y_k}^{y_{n-1}} \varphi(y_{n-2}) dy_{n-2} \cdots \int_{y_k}^{y_{k+2}} \varphi(y_{k+1}) dy_{k+1} \right). \end{aligned} \quad (4)$$

记

$$\begin{aligned} I_1 &= \int_{y_k}^{+\infty} \varphi(y_n) dy_n \int_{y_k}^{y_n} \varphi(y_{n-1}) dy_{n-1} \int_{y_k}^{y_{n-1}} \varphi(y_{n-2}) dy_{n-2} \cdots \int_{y_k}^{y_{k+2}} \varphi(y_{k+1}) dy_{k+1} = \\ &= \int_{y_k}^{+\infty} \varphi(y_n) dy_n \int_{y_k}^{y_n} \varphi(y_{n-1}) dy_{n-1} \int_{y_k}^{y_{n-1}} \varphi(y_{n-2}) dy_{n-2} \cdots \int_{y_k}^{y_{k+3}} (F(y_{k+2}) - F(y_k)) dF(y_{k+2}) = \\ &= \int_{y_k}^{+\infty} \varphi(y_n) dy_n \int_{y_k}^{y_n} \varphi(y_{n-1}) dy_{n-1} \int_{y_k}^{y_{n-1}} \varphi(y_{n-2}) dy_{n-2} \cdots \int_{y_k}^{y_{n-1}} \frac{1}{2} (F(y_{k+3}) - F(y_k))^2 dF(y_{k+3}) \cdots \\ &= \int_{y_k}^{+\infty} \frac{1}{(n - (k + 1))!} (F(y_n) - F(y_k))^{n - (k + 1)} dF(y_n) \\ &= \frac{1}{(n - k)!} (1 - F(y_k))^{n - k}, \end{aligned} \quad (5)$$

$$\begin{aligned} I_2 &= \int_0^{y_k} \varphi(y_{k-1}) dy_{k-1} \int_0^{y_{k-1}} \varphi(y_{k-2}) dy_{k-2} \cdots \int_0^{y_2} \varphi(y_1) dy_1 = \\ &= \int_0^{y_k} \varphi(y_{k-1}) dy_{k-1} \int_0^{y_{k-1}} \varphi(y_{k-2}) dy_{k-2} \cdots \int_0^{y_3} F(y_2) dF(y_2) = \\ &= \int_0^{y_k} \varphi(y_{k-1}) dy_{k-1} \int_0^{y_{k-1}} \varphi(y_{k-2}) dy_{k-2} \cdots \int_0^{y_4} \frac{1}{2} (F(y_3))^2 dF(y_3) \cdots = \\ &= \int_0^{y_k} \frac{1}{(k - 2)!} (F(y_{k-1}))^{k - 2} dF(y_{k-1}) \\ &= \frac{1}{(k - 1)!} (F(y_k))^{k - 1}. \end{aligned} \quad (6)$$

由(4) ~ (6) 式得

$$\begin{aligned} g_k(y) &= n! \varphi(y) \cdot \frac{1}{(k - 1)!} (F(y))^{k - 1} \cdot \frac{1}{(n - k)!} (1 - F(y))^{n - k} = \\ &= k C_n^k \varphi(y) (F(y))^{k - 1} (1 - F(y))^{n - k}. \end{aligned}$$

4 结 语

综上所述,运用牛顿微元法求解一维或多维连续型随机变量的函数的概率密度不失为一种简便有效的计算方法.

参考文献:

- [1] 费勒 W. 概率论及其应用(第二卷)[M]. 李志阡,郑元禄译. 北京: 科学出版社,1994: 52-61.
- [2] 张建华. 概率论与数理统计[M]. 北京: 高等教育出版社,2008: 60-63.

Solving Probability Density with Newton's Infinitesimal Method

ZHU Hui-min

(School of Mathematical Sciences, Fudan University, Shanghai 200433, China)

Abstract: The procedure to solve the probability density function for a continuous random variable with Newton's Infinitesimal method can be unified with that for a discrete random variable. Therefore it is more intuitive and simpler, particularly more practical for complicated random variables.

Keywords: Newton's infinitesimal method; probability density; continuous random variable; random variable function