Lecture 3

多元随机变量与分布

ruanyl@buaa.edu.cn
September 2016, Beihang University

二元分布

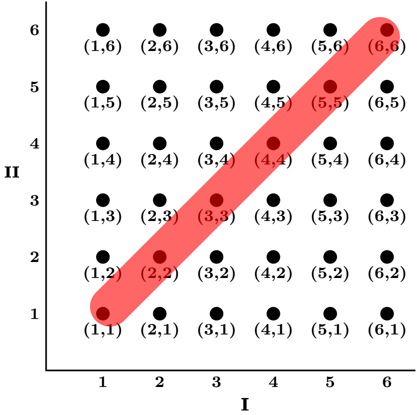
定义 假设 X, Y 是随机变量. 那么 (X, Y) 的**联合分布或者二元分布**定义为所有如下形式的概率的集合: $P((X, Y) \in C)$ 对任意 $C \subset \mathbb{R}^2$. 其中 $\{(X, Y) \in C\}$ 是 $\{\omega \in S : (X(\omega), Y(\omega)) \in C\}$ 的简写 . ω 通常表示一个样本.

二元离散分布

定义 如果联合分布 (X, Y) 的可能取值范围是有限点集或者无穷可数点集, 那么这个分布被称为**离散分布**.

定理 如果 X, Y 是离散随机变量, 那么 (X, Y) 是一个联合离散分布.

例题 投掷两个骰子, 分别用 X 和 Y 表示每个骰子出现的点数. 那么下图则是 (X, Y) 的联合分布的表示. 颜色区域表示事件 $\{X = Y\}$



与一元情形类似, 离散联合分布也有概率函数

定义 如果 (X, Y) 是一个离散联合分布, 那么下面的函数称为其联合概率函数

$$f(x, y) = P(X = x, Y = y).$$

定理 如果 (X, Y) 是一个离散联合分布, 那么其联合概率函数满足

$$\sum_{x,y} f(x,y) = 1$$

并且对任意 $C \subset \mathbb{R}^2$,

$$P((X,Y) \in C) = \sum_{(x,y)\in C} f(x,y) .$$

二元连续分布

定义 随机变量 X, Y 称为具有连续联合分布 , 如果存在定义在 $\mathbb{R}^2$ 上的非负函数 $f(x,y)$ 使得下面的等式成立,

$$P((X,Y) \in E) = \int\int_E f(x,y)\, dxdy$$

其中 E 是 $\mathbb{R}^2$ 上的任意有界, 无界，开或者闭的实数区域. **非负**函数 $f(x, y)$ 则称为该连续分布的**密度函数**.

需要注意的是在这个定义中我们并不假定密度函数 $f(x, y)$ 的连续性, 或是有界性.

定理 由概率公理可知, 密度函数应满足

$$\int\int_{R^2} f(x, y)\, dxdy = 1.$$

定理 假设随机变量 X, Y 具有连续联合分布. 那么

$$P\left((X, Y) \in C_0\right) = 0$$

其中 C_0 是 $\mathbb{R}^2$ 上的任何有限点集, 无穷点列或者函数曲线的图集.

例题 假设随机变量 X, Y 的联合分布具有密度函数

$$f(x,y) = \begin{cases} cx^2y, & x^2 \leqslant y \leqslant 1 \\ 0, & \text{其他} \end{cases}$$

试确定常数 c, 并求出 $P(X \geqslant Y)$.

- 密度函数在全空间积分应当为 1,

$$\begin{aligned} 1 &= \int_{-1}^{1} \left(\int_{x^2}^{1} cx^2 y dy \right) dx \\ &= c \int_{-1}^{1} \left(\frac{x^2}{2} - \frac{x^2 x^4}{2} \right) dx \\ &= c \left(\frac{1}{3} - \frac{1}{7} \right) = \frac{4c}{21} \end{aligned}$$

因而 $c = 21/4$.

$$P(X \geqslant Y) = \frac{21}{4} \int_{0}^{1} \left(\int_{x^2}^{x} x^2 y dy \right) dx$$

$$
\begin{aligned}
&= \frac{21}{4} \int_0^1 \left(\frac{x^2 x^2}{2} - \frac{x^2 x^4}{2} \right) dx \\
&= \frac{21}{8} \left(\frac{1}{5} - \frac{1}{7} \right) = \frac{3}{20}
\end{aligned}
$$

$/\!/\!/$

二元混合分布

定义 假设 X 是离散随机变量 , Y 是连续随机变量, 如果存在 $\mathbb{R}^2$ 上的非负函数 $f(x,y)$ 使得下面的积分对任何 $A, B \subset \mathbb{R}$ 都有意义并且成立

$$
P(X \in A, Y \in B) = \int_B \sum_{x \in A} f(x,y)\, dy.
$$

累积分布函数

定义 给定概率空间以及概率 P. 该空间上的随机变量 X, Y 的联合**分布函数**定义为

$$F(x,y) = P(X \leqslant x, Y \leqslant y), \forall x, y.$$

联合分布函数有如下性质

- 如果 X, Y 的分布函数分别记作 $F_X(x), F_Y(y)$, 那么成立

$$F_X(x) = P(X \leqslant x, Y < \infty) = \lim_{y\to\infty} F(x,y),$$

以及

$$F_Y(y) = P(X < \infty, Y \leqslant y) = \lim_{x\to\infty} F(x,y).$$

这里 $F_X(x)$ 和 $F_Y(y)$ 分别称为联合分布关于 X 和 Y 的**边际分布**.

// $F(\cdot, -\infty) = F(-\infty, \cdot) = 0, F(\infty, \infty) = 1.$

下面的定理给出了分布函数与密度函数的关系.

定理 假设 X, Y 的联合分布具有密度函数 $f(x, y)$. 那么 X, Y 的联合分布函数是连续的, 并且在其二阶导数存在时, 联合分布函数

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s, t)\, dsdt$$

满足以下关系

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial^2 F(x, y)}{\partial y \partial x}.$$

特别地, 如果 f 在 (x, y) 存在连续, 那么上面的求导关系式在该连续点成立.

定理 假设 X, Y 的联合分布具有密度函数 $f(x,y)$. 那么 X, Y 的联合分布函数关于 X 的分布具有密度函数

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

关于 Y 的分布具有密度函数

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx.$$

如果 X 或者 Y 或者两者同时是离散随机变量, 那么上面的相应的积分应改写为求和.

例题 本节开始投掷两个骰子的例题中我们给出了其联合分布, 试写出它的两个边际分布.

- 对 $x, y \in \{1,2,3,4,5,6\}$,

$$f_X(x) = \sum_y f(x,y) = \sum_y P(X=x, Y=y) = \frac{1}{6}$$

$$f_Y(y) = \sum_x f(x,y) = \sum_x P(X=x, Y=y) = \frac{1}{6}$$

这个例题中的联合分布做成一个**各个元素都非负**的矩阵, 其各列和都是相等的. 乘以适当的常数后可使得各列和变为 1, 这一列和为 1 的矩阵被称为**随机矩阵**.

例题 假设 X, Y 的联合分布由下面的矩阵给出

$$\frac{1}{3}\begin{pmatrix} 5/12 & 1/4 & 1/3 \\ 5/12 & 1/4 & 1/3 \\ 1/6 & 1/2 & 1/3 \end{pmatrix}$$

试写出它的两个边际分布.

例题 假设随机变量 X, Y 的联合分布具有密度函数

$$f(x,y)=\begin{cases}\frac{21}{4}x^2y, & x^2\leqslant y\leqslant 1\\ 0, & \text{其他}\end{cases}$$

试求出其边际密度函数

- 由前面的定理

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\,dy=\int_{x^2}^{1}\frac{21}{4}x^2ydy=\frac{21}{8}\left(x^2-x^6\right)$$

$$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\,dx=\int_{-y^{1/2}}^{y^{1/2}}\frac{21}{4}x^2ydx=\frac{7}{2}x^{\frac{5}{2}}$$

多元分布

许多二元分布的概念和相关性质都可以直接地推广到多元情形, 在此不做累述.

例题 在一个临床试验中, 有 n 个病人接受治疗. 用 X_i 表示第 i 个病人在接收治疗后的状态,

$$X_i = \begin{cases} 1, & \text{恢复} \\ 0, & \text{其他} \end{cases}$$

假设每个病人能痊愈的概率是 p, 各个病人情况是相互独立的, 那么 $\mathbf{X} = (X_1, \ldots, X_n)$ 的联合密度应该是

$$f(\mathbf{x}) = p^{x_1+\ldots+x_n}(1-p)^{n-x_1-\ldots-x_n}$$

其中 $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

例题 排队系统是一个由客户以及服务组成的系统, 客户排队等待并按照某流程接受服务. 一个简单的情形是所有客户都排队

等待同一个终端提供的服务, (例如客户排队等待从同一台取款机取款). 假设由个客户. 第个客户接受服务所消耗的时间是. 为了描述这个系统中客户等待时间的分布情况, 我们通常假设 $\mathbf{X} = (X_1, ..., X_n)$ 的联合密度是

$$f(\mathbf{x}) = \begin{cases} \frac{c}{\left(2+\sum_{i=1}^{n} x_i\right)^{n+1}}, & x_i > 0, \forall i \\ 0, & \text{其他} \end{cases}$$

确定常数 c.

- 联合密度 $f(\mathbf{x})$ 应满足

$$I = \int_0^{\infty} \dots \int_0^{\infty} f(x_1, ..., x_n)\, dx_1 dx_2 ... dx_n = 1$$

将 $f(\mathbf{x})$ 的表达式代入并做累次积分, 首先对 x_n 积分得到

$$I = \int_0^{\infty} \dots \int_0^{\infty} \frac{c}{n\left(2+\sum_{i=1}^{n-1} x_i\right)^n} dx_1 dx_2 ... dx_{n-1} = 1$$

然后再依次对 $x_{n-1}, ..., x_1$ 积分得到

$$I = \frac{c}{n! \cdot 2} = 1$$

因此 $c = 2n!$

随机变量独立性

定义 随机变量 X, Y 称为是独立的, 如果:

$$P(X \in A, Y \in B) = P(X \in A) P(Y \in B)$$

对任意 $A, B \subset \mathbb{R}$. 严格地说 A, B 应该是由简单区间（开区间, 闭区间, 或半开闭区间）生成的“可测集”, 这部分内容已超出本课程要求, 在此不做深入讨论. 但我们有如下定理

定理 随机变量 X, Y 是独立的当且仅当对 $\forall x, y \in \mathbb{R}$

$$P(X \leqslant x, Y \leqslant y) = P(X \leqslant x) P(Y \leqslant y).$$

定理中 $\leqslant$ 替换为 $<$ 时仍然成立. 这一定理是告诉我们随机变量 X, Y 独立的充分必要条件是联合分布函数等于边际分布函数乘积. 从而借助于分布函数与密度函数的关系, 我们还有

定理 假设 X, Y 其具有密度函数. 随机变量 X, Y 是独立的当且仅当其联合密度函数 $f(x,y)$ 满足

$$f(x,y)=f_X(x)f_Y(y),\ a.e.\ x,y.$$

如果 $f(x,y), f_X(x), f_Y(y)$ 都是连续的, 那么独立性等价于

$$f(x,y)=f_X(x)f_Y(y),\ \forall\, x,y.$$

一元随机变量的函数

很多时候我们关心的不仅仅是随机变量本身, 同时也关心通过已知随机变量所描述的其他数量的分布情况, 例如

例题 假设 X 在整数集合 $C=\{1,2,...,9\}$ 上是均匀分布的. 想要知道 X 与集合 C 中点 5 距离的分布情况, 因而我们考虑 $Y=|X-5|$. 显然 Y 也是一个随机变量, 并且它是随机变量 X 的函数. Y 的分布可以由 X 的分布来确定. Y 的取值范围是 $\{0,1,...,4\}$. 由于 $P(X=i)=\frac{1}{9}, i=1,2,...,9$, 因此

$$P(Y=j)=\begin{cases} P(X\in\{5\})=\frac{1}{9}, & j=0 \\ P(X\in\{4,6\})=\frac{2}{9}, & j=1 \\ P(X\in\{3,7\})=\frac{2}{9}, & j=2 \\ P(X\in\{2,8\})=\frac{2}{9}, & j=3 \\ P(X\in\{1,9\})=\frac{2}{9}, & j=4 \end{cases}$$

一般地

定理 如果 X 是一个离散随机变量, Y 是随机变量 X 的函数, $Y=h(X)$, 那么 Y 的概率函数可以写作

$$g(y)=P(Y=y)=\sum_{x\in h^{-1}(y)} P(X=x)$$

其中 $h^{-1}(y)$ 表示 y 在 h 作用下的原象集 $\{x:h(x)=y\}$.

我们经常用 $h^{-1}(C)$ 来表示使得 $h(x)$ 取值在 C 中的 x 的集合, $h^{-1}(C)=\{x:h(x)\in C\}$. 上面这个定理的一个特殊情形是 $h(x)$ 为严格单调函数, 此时 $h^{-1}(y)$, 我们有

定理 如果 X 是一个离散随机变量, 它的概率函数是 $f(x)$. Y 是随机变量 X 的函数, $Y=h(X)$, 那么 Y 的概率函数可以写作

$$g(y)=P\left(X=h^{-1}(y)\right)=f\left(h^{-1}(y)\right)$$

定理 如果 X 是一个以 $f(x)$ 为密度函数的连续型随机变量，$P(X \in \mathcal{S}) = 1$. Y 是随机变量 X 的函数 $Y = h(X)$, 并且 $h(x)$ 是 $\mathcal{S}$ 上的严格单调的可微分函数. 令 $\mathcal{T} = \phi(\mathcal{S})$, 那么 Y 的密度函数为

$$g(y) = f\left(h^{-1}(y)\right)\left|\frac{dh^{-1}(y)}{dy}\right| \cdot 1_{\mathcal{T}}(y)$$

这里 $1_{\mathcal{T}}(y)$ 表示集合 $\mathcal{T}$ 上的特征函数：$1_{\mathcal{T}}(y) = 1$ 如果 $y \in \mathcal{T}$, 否则为 0.

■ 我们证明 $h(x)$ 是严格单调减少的情形. 由于 $h(x)$ 严格单调，于是 $h^{-1}(y)$ 在 $\mathcal{T}$ 上存在, 并且也是严格单调减少的. 按定义 Y 的分布函数在 $y \in \mathcal{T}$ 时为

$$\begin{aligned} P(Y \leqslant y) &= P(h(X) \leqslant y) \\ &= P\left(X \geqslant h^{-1}(y)\right) \end{aligned}$$

$$
\begin{aligned}
&= 1 - P\left(X < h^{-1}(y)\right) \\
&= 1 - P\left(X \leqslant h^{-1}(y)\right) \\
&= 1 - f\left(h^{-1}(y)\right)
\end{aligned}
$$

这里我们用到了连续型随机变量分布函数的连续性. 等式两边对 y 求导数

$$
\begin{aligned}
g(y) &= -f\left(h^{-1}(y)\right) \cdot \frac{dh^{-1}(y)}{dy} \\
&= f\left(h^{-1}(y)\right) \left| \frac{dh^{-1}(y)}{dy} \right| .
\end{aligned}
$$

例题 已知微生物繁殖速度是指数增长的. 在 0 时刻将 v 个微生物放入一个水池中, 用 X 表示微生物增长率. 经过时间 t 小时之

后, 微生物的数量将变为 ve^{Xt}. 现假设 $v=10$, X 的分布函数为

$$f(x)=\begin{cases} 3(1-x)^2, & x\in(0,1) \\ 0, & \text{其他} \end{cases}$$

试确定 5 小时之后, 微生物数量的分布.

■ 令 $h(x)=10e^{5x}$, $\mathcal{S}=(0,1)$. 则 $h(x)$ 是一个严格递增函数, 它将 $\mathcal{S}$ 映射为 $\mathcal{T}=\left(10,10e^5\right)$, $h^{-1}(y)=\frac{1}{5}\ln\frac{y}{10}$, $h'(x)=50e^{5x}$. 因此

$$\frac{dh^{-1}(y)}{dy}=\frac{1}{h'\left(h^{-1}(y)\right)}=\frac{1}{5y}.$$

从而微生物的数量 $Y=h(X)$ 的分布为

$$g(y)=\begin{cases} \frac{3\left(1-\frac{1}{5}\ln\frac{y}{10}\right)^2}{5y}, & y\in\left(10,10e^5\right) \\ 0, & \text{其他} \end{cases}$$

多元随机变量的函数

与一元情形类似我们有

定理 如果 (X_1, X_2) 是一个离散联合分布，它的概率函数是 $f(x, y)$. (Y_1, Y_2) 是 (X_1, X_2) 的函数，

$$\phi : (X_1, X_2) \mapsto (Y_1 = \phi_1 (X_1, X_2), Y_2 = \phi_2 (X_1, X_2)).$$

那么 (Y_1, Y_2) 的联合概率函数可以写作

$$g(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) = \sum_{(x_1, x_2) \in \phi^{-1}(y_1, y_2)} P(X_1 = x_1, X_2 = x_2)$$

其中 $\phi^{-1}(y_1, y_2)$ 表示 (y_1, y_2) 在 $\phi = (\phi_1, \phi_2)$ 作用下的原象集 $\{(x_1, x_2) : \phi_1 (x_1, x_2) = y_1, \phi_2 (x_1, x_2) = y_2\}$.

定理 如果 (X_1, X_2) 是具有密度函数 $f(x_1, x_2)$ 的连续型分布. $P((X_1, X_2) \in \mathcal{S}) = 1$. (Y_1, Y_2) 是 (X_1, X_2) 的可微分函数,

$$\phi : (X_1, X_2) \mapsto (Y_1 = \phi_1(X_1, X_2), Y_2 = \phi_2(X_1, X_2)).$$

并且 (Y_1, Y_2) 是 (X_1, X_2) 一一对应, 即 (X_1, X_2) 也可以表示为 (Y_1, Y_2) 的函数

$$\psi : (Y_1, Y_2) \mapsto (X_1 = \psi_1(Y_1, Y_2), X_2 = \psi_2(Y_1, Y_2)).$$

令 $\mathcal{T} = \phi(\mathcal{S})$. 那么 (Y_1, Y_2) 的联合概率函数是（注意绝对值）

$$g(y_1, y_2) = f(\psi_1(y_1, y_2), \psi_2(y_1, y_2)) \left| \frac{\partial(\psi_1, \psi_2)}{\partial(y_1, y_2)} \right| \cdot 1_{\mathcal{T}}(y_1, y_2)$$

其中

$$\frac{\partial(\psi_1, \psi_2)}{\partial(y_1, y_2)} = \det \begin{pmatrix} \frac{\partial \psi_1}{\partial y_1} & \frac{\partial \psi_1}{\partial y_2} \\ \frac{\partial \psi_2}{\partial y_1} & \frac{\partial \psi_2}{\partial y_2} \end{pmatrix}$$

为 Jacobian 行列式.

上面的定理可以直接地推广到多个变量的情形, 在此不做详述.

例题 如果 (X_1, X_2) 是具有密度函数 $f(x_1, x_2)$ 的连续型分布. $Y = a_1X_1 + a_2X_2 + b$, $(a_1 \neq 0)$. 那么 Y 具有连续型分布, 其密度函数为

$$g(y) = \int_{-\infty}^{\infty} f\left(\frac{y - a_2x_2 - b}{a_1}, x_2\right) \frac{1}{|a_1|} dx_2.$$

- 假设 $a_1 > 0$. （$a_1 < 0$ 证明类似）. 令

$$A_y = \{(x_1, x_2) : a_1x_1 + a_2x_2 + b \leqslant y\}$$

那么 Y 的分布函数为

$$G(y) = \int\int_{A_y} f(x_1, x_2)\, dx_1 dx_2$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{(y-a_2x_2-b)/a_1} f(x_1, x_2)\, dx_1 dx_2$$

对内层积分进行变量代换, 令 $z = a_1x_1 + a_2x_2 + b$, 那么 $dz = a_1 dx_1$,

$$\begin{aligned} G(y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{y} f\left(\frac{z - a_2x_2 - b}{a_1}, x_2\right) \frac{1}{a_1} dz dx_2 \\ &= \int_{-\infty}^{y} \left(\int_{-\infty}^{\infty} f\left(\frac{y - a_2x_2 - b}{a_1}, x_2\right) \frac{1}{a_1} dx_2 \right) dz \end{aligned}$$

由此可知

$$\int_{-\infty}^{\infty} f\left(\frac{y - a_2x_2 - b}{a_1}, x_2\right) \frac{1}{a_1} dx_2$$

正是所求的密度函数.

该结论的一个特殊情形是 $Y = X_1 + X_2$, 此时我们有

定理 如果 (X_1, X_2) 是具有密度函数 $f(x_1, x_2)$ 的连续型分布. 那么 $Y = X_1 + X_2$ 具有连续型分布, 其密度函数为

$$g(y) = \int_{-\infty}^{\infty} f(y - x_2, x_2)\, dx_2 = \int_{-\infty}^{\infty} f(x_1, y - x_1)\, dx_1.$$

随机变量 Y 的分布称为 X_1 与 X_2 **分布的卷积.**