

Ch05: 多维随机向量及其数字特征

Multi-dimensional Random Vector and Distribution Function

November 6, 2024

多维随机向量

在很多实际问题中, 随机现象可能需要两种或两种以上的随机因素来描述, 仅仅用一个随机变量是不够的.

为考察某个地区儿童的身体素质, 可考虑他们的身高、体重、肺活量、视力等, 此时至少需要四个随机变量来进行描述.

这些随机变量之间可能存在某些关联, 分别对每个随机变量单独进行研究是不够的, 需要将其看作一个整体, 即多维随机向量.

定义 0.34 设 $X_1 = X_1(\omega)$, $X_2 = X_2(\omega)$, \dots , $X_n = X_n(\omega)$ 是定义在同一样本空间 Ω 上的 n 个随机变量, 其中 $\omega \in \Omega$, 由它们构成的向量 (X_1, X_2, \dots, X_n) 称为 n 维随机向量, 或称 n 维随机变量.

提纲：二维随机变量

- 二维随机变量的分布函数
- 二维随机变量的边缘分布函数
- 二维随机离散型变量
 - 联合分布列、边缘分布列
 - 联合分布函数、边缘分布函数
 - 常见的二维随机离散型变量：多项分布
- 二维随机连续型变量
 - 联合密度函数、边缘密度函数
 - 联合分布函数、边缘分布函数
 - 常见的二维随机连续型变量：均匀分布、正态分布

二维随机向量的分布函数

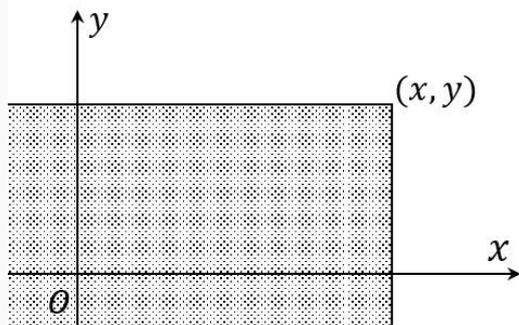
定义 0.35 设 (X, Y) 为二维随机向量, 对任意实数 x 和 y , 称

$$F(x, y) = P(X \leq x, Y \leq y)$$

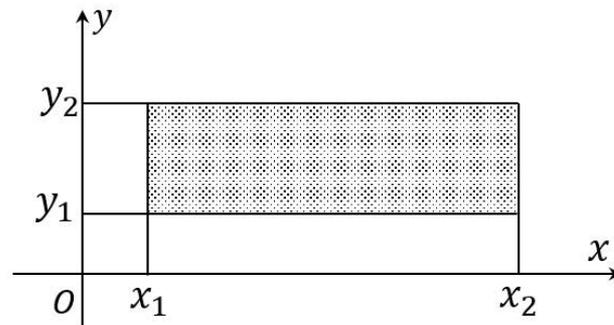
为二维随机向量 (X, Y) 的分布函数, 或称为随机变量 X 和 Y 的联合分布函数 (joint cumulative probability distribution function).

分布函数的几何解释: $F(x, y) = P(X \leq x, Y \leq y)$

若将 (X, Y) 看作平面上随机点的坐标, 则分布函数 $F(x, y)$ 在点 (x, y) 的值是随机向量 (X, Y) 落入以 (x, y) 为顶点的左下方无穷区域的概率, 如图 (a) 所示.



(a)



(b)

如图 (b) 所示, (X, Y) 落入矩形区域 $\{(x, y) : x \in (x_1, x_2], y \in (y_1, y_2]\}$ 的概率:

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

二维随机变量分布函数的性质

- **随坐标方向轴单调不减.** 分布函数 $F(x, y)$ 对每个变量单调不减
 - 固定 y , 当 $x_1 > x_2$ 时有 $F(x_1, y) \geq F(x_2, y)$
 - 固定 x , 当 $y_1 > y_2$ 时有 $F(x, y_1) \geq F(x, y_2)$
- **规范性.** 对任意实数 x 和 y , 分布函数 $F(x, y) \in [0, 1]$, 且

$$F(+\infty, +\infty) = 1, \quad F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0$$

- **单侧连续.** 分布函数 $F(x, y)$ 关于每个变量**右连续**

$$F(x, y) = F(x + 0, y), \quad F(x, y) = F(x, y + 0)$$

- **矩形运算.** 对任意实数 $x_1 < x_2$ 和 $y_1 < y_2$ 有

$$F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$$

回顾：单变量分布函数的定义 – 连续性

一个函数 $F(x)$ 可以作为分布函数的充分必要条件为：满足以下三性质

- 单调不降
- $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$
- 右连续或左连续.
 - 右连续: $F_1(x + 0) = F_1(x)$ corresponding to $F_1(x) = P(X \leq x)$
 - 左连续: $F_2(x - 0) = F_2(x)$ corresponding to $F_2(x) = P(X < x)$

二者反映 X 的概率规律是相同的. 即

$$F_1(x) = F_2(x) + P(X = x)$$

当 X 为连续型随机变量时, $P(X = x) = 0$, 这两者无本质区别.

右连续或者左连续是分布函数的一个性质, 即通过定义可以证明其右连续或左连续性.

二维随机变量分布函数的性质

充要条件:

$F(x, y)$ 是某二维随机向量的分布函数 $\Leftrightarrow F(x, y)$ 满足前面四条性质

- 随坐标方向轴单调不减. 分布函数 $F(x, y)$ 对每个变量单调不减
- 规范性. 对任意实数 x 和 y , 分布函数 $F(x, y) \in [0, 1]$, 且

$$F(+\infty, +\infty) = 1, \quad F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0$$

- 单侧连续. 分布函数 $F(x, y)$ 关于每个变量右连续

$$F(x, y) = F(x + 0, y), \quad F(x, y) = F(x, y + 0)$$

- 矩形运算. 对任意实数 $x_1 < x_2$ 和 $y_1 < y_2$ 有

$$F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$$

二维随机变量分布函数的性质

充要条件:

$F(x, y)$ 是某二维随机向量的分布函数 $\Leftrightarrow F(x, y)$ 满足前面四条性质

值得说明的是, 当二元函数 $F(x, y)$ 仅仅满足前面的三条性质时, 不一定能成为某二维随机向量的分布函数. 例如

$$F(x, y) = \begin{cases} 1 & x + y \geq -1 \\ 0, & x + y < -1 \end{cases}$$

很容易验证 $F(x, y)$ 仅仅满足前面的三条性质, 但因为

$$F(1, 1) - F(1, -1) - F(-1, 1) + F(-1, -1) = -1$$

因此不满足第四条性质, 不构成一个分布函数.

边缘分布函数

根据联合分布函数 $F(x, y)$ 还可研究每个分量的统计特征, 将 X 和 Y 看做单独的随机变量, 根据联合分布函数 $F(x, y)$ 来研究随机变量 X 和 Y 的分布函数 $F_X(x)$ 和 $F_Y(y)$, 即 **边缘分布函数**.

定义 0.36 设二维随机变量 (X, Y) 的联合分布函数为 $F(x, y)$, 称

$$F_X(x) = P(X \leq x) = P(X \leq x, Y < +\infty) = F(x, +\infty) = \lim_{y \rightarrow +\infty} F(x, y)$$

为随机变量 X 的 **边缘分布函数**.

同理定义随机变量 Y 的边缘分布函数为

$$F_Y(y) = P(Y \leq y) = P(Y \leq y, x < +\infty) = F(+\infty, y) = \lim_{x \rightarrow +\infty} F(x, y)$$

遍历另一个随机变量的取值.

二维随机变量：例 0.94

例 0.94 设二维随机变量 (X, Y) 的联合分布函数为

$$F(x, y) = A \left(B + \arctan \frac{x}{2} \right) \left(C + \arctan \frac{y}{3} \right).$$

求随机变量 X 和 Y 的边缘分布函数, 以及概率 $P(Y > 3)$.

解答：例 0.94

题目：设二维随机变量 (X, Y) 的联合分布函数为

$$F(x, y) = A(B + \arctan \frac{x}{2})(C + \arctan \frac{y}{3})$$

求随机变量 X 和 Y 的边缘分布函数, 以及概率 $P(Y > 3)$.

解答:

- 由分布函数的性质 $F(+\infty, +\infty) = 1$, $F(-\infty, y) = 0$, $F(x, -\infty) = 0$, 计算常数 A, B, C 的值分别为 $A = \frac{1}{\pi^2}$, $B = \frac{\pi}{2}$, $C = \frac{\pi}{2}$.
- 由边缘分布函数的定义可得

$$F_X(x) = \lim_{y \rightarrow +\infty} \frac{1}{\pi^2} \left(\frac{\pi}{2} + \arctan \frac{x}{2} \right) \left(\frac{\pi}{2} + \arctan \frac{y}{3} \right) = \frac{1}{\pi} \left(\frac{\pi}{2} + \arctan \frac{x}{2} \right)$$

同理可得 $F_Y(y) = \frac{1}{\pi} \left(\frac{\pi}{2} + \arctan \frac{y}{3} \right)$.

- 由 Y 的边缘分布函数可得

$$P(Y > 3) = 1 - P(Y \leq 3) = 1 - F_Y(3) = \frac{1}{4}$$

二维离散型随机向量及其分布列

定义 0.37 若二维随机变量 (X, Y) 的取值是有限个或无限可列的, 称 (X, Y) 为 **二维离散型随机变量**. 设离散型随机向量 (X, Y) 的取值分别为 $(x_i, y_j), i, j = 1, 2, \dots$, 则称

$$p_{ij} = P(X = x_i, Y = y_j)$$

为 (X, Y) 的 **联合分布列**, 简称 **分布列**.

二维随机变量分布列具有以下性质:

非负性: $p_{ij} \geq 0$ 规范性: $\sum_{i,j} p_{ij} = 1$

二维离散型随机向量: 边缘分布列

根据二维随机变量 (X, Y) 的联合分布列 p_{ij} , 还可研究每个随机变量的统计特征, 将 X 和 Y 看做单独的随机变量.

- 随机变量 X 的边缘分布列

$$P(X = x_i) = \sum_{j=1}^{+\infty} P(X = x_i, Y = y_j) = \sum_{j=1}^{+\infty} p_{ij} = p_{i\cdot}$$

- 随机变量 Y 的边缘分布列

$$P(Y = y_i) = \sum_{i=1}^{+\infty} P(X = x_i, Y = y_j) = \sum_{i=1}^{+\infty} p_{ij} = p_{\cdot j}$$

- 遍历另一个随机变量的取值.

二维离散型随机向量: 边缘分布列

根据二维随机变量 (X, Y) 的联合分布列和边缘分布列可以通过下表表示:

$X \backslash Y$	y_1	y_2	\cdots	y_j	\cdots	$p_{i\cdot} = \sum_j p_{ij}$
x_1	p_{11}	p_{12}	\cdots	p_{1j}	\cdots	$p_{1\cdot}$
x_2	p_{21}	p_{22}	\cdots	p_{2j}	\cdots	$p_{2\cdot}$
\vdots	\vdots	\vdots		\vdots		\vdots
x_i	p_{i1}	p_{i2}	\cdots	p_{ij}	\cdots	$p_{i\cdot}$
\vdots	\vdots	\vdots		\vdots	\ddots	\vdots
$p_{\cdot j} = \sum_i p_{ij}$	$p_{\cdot 1}$	$p_{\cdot 2}$	\cdots	$p_{\cdot j}$	\cdots	1

二维离散型随机向量：联合分布函数、边缘分布函数

根据二维随机变量 (X, Y) 的联合分布列 p_{ij} , 可得:

- 联合分布函数

$$F(x, y) = \sum_{x_i \leq x, y_j \leq y} p_{ij}$$

- 边缘分布函数

$$F_X(x) = \sum_{x_i \leq x} p_{i.} = \sum_{x_i \leq x} \sum_{j=1}^{+\infty} p_{ij} \quad \text{和} \quad F_Y(y) = \sum_{y_j \leq y} p_{.j} = \sum_{y_j \leq y} \sum_{i=1}^{+\infty} p_{ij}$$

二维离散型随机变量：例 0.95

例 0.95 设某一个地区

- 15% 的家庭没有小孩
- 20% 的家庭有一个小孩
- 35% 的家庭有两个小孩
- 30% 的家庭有三个小孩

设每个家庭中小孩为男孩和女孩是等可能的且相互独立. 现从该地区随机任意选择一个家庭, 用随机变量 X, Y 分别表示该家庭中男孩和女孩的个数, 求概率 $P(X \geq 1)$, $P(Y \leq 2)$ 和 $P(X \leq Y)$.

解答：例 0.95

- 联合分布列及边缘分布列

$X \backslash Y$	0	1	2	3	$p_{i\cdot}$
0	0.1500	0.1000	0.0875	0.0375	0.3750
1	0.1000	0.175	0.1125	0	0.3875
2	0.0875	0.1125	0	0	0.2000
3	0.0375	0	0	0	0.0375
$p_{\cdot j}$	0.3750	0.3875	0.2000	0.0375	1

- X, Y 所有可能取值为 $\{0, 1, 2, 3\}$, 由条件概率可知:

$$\begin{aligned} P(X = i, Y = j) &= P(\text{选择的家庭有 } i + j \text{ 个小孩, 其中 } i \text{ 个男孩和 } j \text{ 个女孩}) \\ &= P(\text{选择的家庭有 } i + j \text{ 个小孩}) \cdot P(i \text{ 个男孩和 } j \text{ 个女孩} \mid \text{选择的家庭有 } i + j \text{ 个小孩}) \\ &= \binom{i+j}{i} \frac{1}{2^{i+j}} \cdot P(\text{选择的家庭有 } i + j \text{ 个小孩}) \end{aligned}$$

- 可列出 X, Y 的联合分布列及边缘分布列, 最后可得:

$$P(X \geq 1) = 0.625, \quad P(Y \leq 2) = 0.9625, \quad P(X \leq Y) = 0.6625.$$

常用二维离散分布 – 多项分布

伯努利试验: 试验 E 有 2 种可能的结果 A 或 \bar{A} —— 二项分布

多可能试验: 试验 E 有 n 种可能的结果及概率, $p_1 + p_2 + \cdots + p_n = 1$

事件	A_1	A_2	\cdots	A_n
概率	p_1	p_2	\cdots	p_n
次数	X_1	X_2	\cdots	X_n

将试验 E 独立重复地进行 m 次, 有

- X_i 取值为 $\{0, 1, \dots, m\}$
- $X_1 + X_2 + \cdots + X_n = m$

则随机向量 (X_1, X_2, \dots, X_n) 服从 多项分布.

常用二维离散分布 – 多项分布

定义 **0.38** 若 n 维随机向量 (X_1, X_2, \dots, X_n) 的分布列为

$$P(X_1 = k_1, X_2 = k_2, \dots, X_n = k_n) = \binom{m}{k_1, k_2, \dots, k_n} p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n}$$

其中 k_1, k_2, \dots, k_n 是非负整数且满足

- $k_1 + k_2 + \cdots + k_n = m$
- $p_1 + p_2 + \cdots + p_n \leq 1$

则称随机向量 (X_1, X_2, \dots, X_n) 服从参数为 m, p_1, p_2, \dots, p_n 的**多项分布 (multinomial distribution)**, 记为 $(X_1, X_2, \dots, X_n) \sim M(m, p_1, p_2, \dots, p_n)$.

定义 **0.39** 设随机变量 X 的分布列为

$$P(X = k) = \binom{m}{k} p^k (1 - p)^{m-k}, \quad k \in \{0, 1, 2, \dots, m\}$$

称随机变量 X 服从为**参数为 m 和 p 的二项分布 (binomial distribution)**, 记为 $X \sim B(m, p)$.

多项分布的性质: (思考)

定理 0.14 若多维随机向量 $(X_1, X_2, \dots, X_n) \sim M(m, p_1, p_2, \dots, p_n)$, 则每个随机变量 $(X_i (i \in [n]))$ 的边缘分布为二项分布 $B(m, p_i)$.

一言以蔽之: [多项分布的边缘分布是二项分布] [如何证明?](#)

思路: 我们可以将第 i 类别定义为“成功”, 其他所有类别的结合定义为“失败”. 成功的概率就是第 i 类别的发生概率 p_i , 失败的概率就是其他所有类别的概率之和:

$$1 - p_i = p_1 + \dots + p_{i-1} + p_{i+1} + \dots + p_n$$

例 1: 在一个掷骰子的实验中, 每个骰子面可能被视为一个类别, 每个面被选中的概率可以不同.

多项分布的性质：例 1

例 1: 在一个掷骰子的实验中, 每个骰子面可能被视为一个类别, 每个面被选中的概率可以不同.

解答:

- 多可能试验: 6 种可能
- 独立重复进行 m 次试验
- 假设出现 $(1, 2, 3, 4, 5, 6)$ 取值出现的次数为 $k_1, k_2, k_3, k_4, k_5, k_6$, 则分布列为:

$$P(X_1 = k_1, \dots, X_6 = k_6) = \binom{m}{k_1, k_2, \dots, k_6} \left(\frac{1}{6}\right)^{k_1+k_2+k_3+k_4+k_5+k_6}$$

- 点数为 1 的边缘分布为

$$P(X_1 = k_1) = \binom{m}{k_1} \frac{1}{6} \left(1 - \frac{1}{6}\right)^{m-k_1}, \quad k \in \{0, 1, 2, \dots, m\}$$

多项分布的边缘分布是二项分布：例 0.96

例 0.96 假设我们进行了一项实验

事件	A	B	C
概率	0.3	0.4	0.3

总共进行了 5 次实验. 求结果 A 发生的次数的分布.

解答：例 0.96 — 咨询 ChatGPT-4

现在我们计算类别A发生1次、2次、3次、4次和5次的概率：

$$x_A = 1, x_B = 2, x_C = 2:$$

$$P(X_A = 1, X_B = 2, X_C = 2) = \frac{5!}{1!2!2!} \cdot (0.3)^1(0.4)^2(0.3)^2 \approx 0.3024$$

$$x_A = 2, x_B = 1, x_C = 2:$$

$$P(X_A = 2, X_B = 1, X_C = 2) = \frac{5!}{2!1!2!} \cdot (0.3)^2(0.4)^1(0.3)^2 \approx 0.216$$

$$x_A = 3, x_B = 1, x_C = 1:$$

$$P(X_A = 3, X_B = 1, X_C = 1) = \frac{5!}{3!1!1!} \cdot (0.3)^3(0.4)^1(0.3)^1 \approx 0.1512$$

$$x_A = 4, x_B = 0, x_C = 1:$$

$$P(X_A = 4, X_B = 0, X_C = 1) = \frac{5!}{4!0!1!} \cdot (0.3)^4(0.4)^0(0.3)^1 \approx 0.027$$

$$x_A = 5, x_B = 0, x_C = 0:$$

$$P(X_A = 5, X_B = 0, X_C = 0) = \frac{5!}{5!0!0!} \cdot (0.3)^5(0.4)^0(0.3)^0 \approx 0.00243$$

这些就是类别A发生1次、2次、3次、4次和5次的概率。

现在让我们来看看，如果我们把这些概率组合起来，得到了类别A发生次数的二项分布：

$$P(X_A = k) = \binom{5}{k} \cdot (0.3)^k \cdot (1 - 0.3)^{5-k}$$

其中，k可以是1、2、3、4或5。

Appendix: 二项分布和多项分布的关系

综上可知,

- 多项分布的边缘分布是二项分布 (已证)
- Label Distribution
 - 在机器学习中, 二项分布可对应于二分类任务的类别分布
 - 在机器学习中, 多项分布可对应于多分类任务的类别分布
- 第二次交互式课堂 (候选课题之一)

二维连续型随机变量

定义 0.40 设二维随机变量的分布函数为 $F(x, y)$, 如果存在二元非负可积函数 $f(x, y)$ 使得对任意实数 (x, y) 有

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) \, du \, dv$$

则称 (X, Y) 为二维连续型随机变量, 称 $f(x, y)$ 称为二维随机变量 (X, Y) 的概率密度, 或称为随机变量 X 和 Y 的联合密度函数.

联合密度函数的性质

- 非负性. $f(x, y) \geq 0$;
- 规范性. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$;
- 连续性. 若 $f(x, y)$ 在 (x, y) 连续, 则

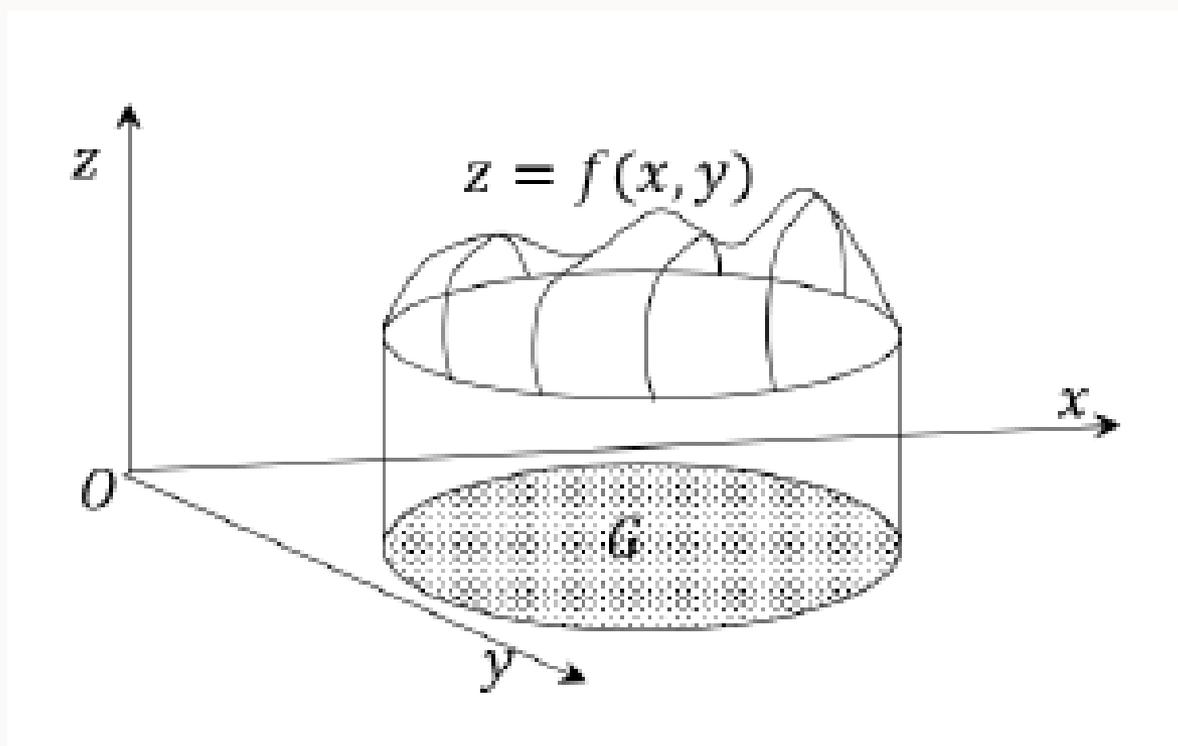
$$f(x, y) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(X \in (x, x + \Delta x], Y \in (y, y + \Delta y])}{\Delta x \Delta y} = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

- 有界区域可积. 若 G 为平面上的一个区域, 则点 (X, Y) 落入 G 的概率为

$$P((X, Y) \in G) = \iint_{(x, y) \in G} f(x, y) dx dy$$

联合密度函数的性质 — 有界区域可积的几何解释

$$P((X, Y) \in G) = \iint_{(x,y) \in G} f(x, y) dx dy$$



二维连续型随机变量: 边缘概率密度

设二维随机变量 (X, Y) 的概率密度为 $f(x, y)$, 由边缘分布函数的定义可知:

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(X \leq x, Y < +\infty) = F(x, +\infty) \\ &= \int_{-\infty}^x \int_{-\infty}^{+\infty} f(t, y) dt dy = \int_{-\infty}^x \left(\int_{-\infty}^{+\infty} f(t, y) dy \right) dt \end{aligned}$$

由此可得随机变量 X 的边缘概率密度为:

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

同理可得随机变量 Y 的边缘概率密度.

二维连续型随机变量: 边缘概率密度

定义 **0.41** 设二维随机变量 (X, Y) 的联合密度函数为 $f(x, y)$, 随机变量 X 和 Y 的 **边缘密度函数** 分别为

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

其相对应的 **边缘分布函数** 分别为

$$F_X(x) = \int_{-\infty}^x \left(\int_{-\infty}^{+\infty} f(t, y) dy \right) dt$$
$$F_Y(y) = \int_{-\infty}^y \left(\int_{-\infty}^{+\infty} f(x, s) dx \right) ds$$

二维连续型随机变量：例 0.97

例 0.97 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} ce^{-(3x+4y)}, & x > 0, y > 0 \\ 0, & \text{其它} \end{cases}$$

求：

- 常数 c ;
- 联合分布函数 $F(x, y)$;
- X 和 Y 的边缘概率密度;
- $P(X + Y \leq 2)$ 的值.

解答：例 0.97

题目：如上所述.

解答：

- 由规范性可知 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$, 计算后得 $c = 12$.
- 由分布函数的定义可知, 当 $x > 0$ 和 $y > 0$ 时有

$$F(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = \int_0^x \int_0^y 12e^{-(3x+4y)} dx dy = (1 - e^{-3x})(1 - e^{-4y})$$

- 由边缘密度函数的定义可知,

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^{+\infty} 12e^{-(3x+4y)} dx dy = 3e^{-3x}$$

同理可得 $f_Y(y) = 4e^{-4y}$.

- 由联合密度函数的性质可知,

$$P(X + Y \leq 2) = 12 \int_0^2 f(x, y) dx \int_0^{2-x} e^{-(3x+4y)} dy = 1 - 4e^{-6} + 3e^{-8}$$

常用二维连续分布 — 均匀分布

定义 0.42 设 G 为平面上的一个有界区域, 其面积为 A_G , 若二维随机向量 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} \frac{1}{A_G}, & (x, y) \in G \\ 0, & (x, y) \notin G \end{cases}$$

则称 (X, Y) 服从区域 G 上的**二维均匀分布**.

二维均匀分布所描述的随机现象就是向区域 G 上随机投点, 该点落在区域 G 上的某一点是等可能的, 本质上就是 (平面) 几何概型的随机向量描述.

均匀分布：例 0.98

例 0.98 在坐标原点为中心、半径为 R 的圆内等可能随机投点. 用 (X, Y) 分别表示落点的横坐标和纵坐标.

求随机向量 (X, Y) 的联合密度函数, 边缘密度函数, 以及 (X, Y) 落入 $X^2 + Y^2 \leq r^2 (0 < r \leq R)$ 的概率.

解答：例 0.98

题目：如上所述.

解答：

- 由均匀分布的定义可知, 随机向量 (X, Y) 的联合密度函数为

$$f(x, y) = \begin{cases} \frac{1}{\pi R^2}, & x^2 + y^2 \leq R^2 \\ 0, & x^2 + y^2 > R^2 \end{cases}$$

- 由边缘密度函数的定义可知, 当 $x^2 \leq R^2$ 时有

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} \frac{1}{\pi R^2} dy = \frac{2}{\pi R^2} \sqrt{R^2 - x^2}$$

同理可得随机变量 Y 的边缘概率密度.

- 由联合密度函数的性质可知,

$$P(X^2 + Y^2 \leq r^2) = \iint_{x^2+y^2 \leq r^2} \frac{1}{\pi R^2} dx dy = \frac{r^2}{R^2}$$

常用二维连续分布 – 正态分布

定义 0.43 对任意实数 x, y , 若随机向量 (X, Y) 的密度函数 $f(x, y)$ 等于

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_x\sigma_y} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x^2\sigma_y^2}\right]\right)$$

其中常数 $\mu_x, \mu_y \in (-\infty, +\infty)$, $\sigma_x, \sigma_y \in (0, +\infty)$ 和 $\rho \in (-1, 1)$, 则称 (X, Y) 服从 **二维正态分布**, 记 $(X, Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$.

Remark:

$$\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x^2\sigma_y^2}\right]$$

NOT refer to

$$\left(\frac{x-\mu_x}{\sigma_x} - \frac{y-\mu_y}{\sigma_y}\right)^2$$

refer to

$$\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{\left(y-\mu_y - \rho\frac{x-\mu_x}{\sigma_x}\right)^2}{2\left(\sqrt{1-\rho^2}\sigma_y\right)^2}$$

二维随机正态分布的性质

定理 0.15 设二维随机变量 $(X, Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, 则随机变量 X 和 Y 的边缘分布为

$$X \sim \mathcal{N}(\mu_x, \sigma_x^2), \quad Y \sim \mathcal{N}(\mu_y, \sigma_y^2).$$

Remark:

- 该证明可以参考《概率论与数理统计 (人工智能或计算机专业用书)》
- ρ 为结合曲率, 确保边缘分布 $F_X(x)$ 中关于 $Y \in (-\infty, +\infty)$ 的积分为 1.

$$\frac{(x - \mu_x)^2}{\sigma_x^2} + \frac{\left(y - \mu_y - \rho \frac{x - \mu_x}{\sigma_x}\right)^2}{2 \left(\sqrt{1 - \rho^2} \sigma_y\right)^2}$$

正态分布：例 0.99

例 0.99 假设有两个随机变量 X 和 Y , 它们的联合分布是一个二维正态分布, 均值向量为 $(1, 2)$. 求

- $P(X > 2, Y < 3)$.
- 边缘分布 $P(X > 2)$ 和 $P(Y < 3)$

解答：例 0.99— 咨询 ChatGPT-4

答案：

1. $P(X > 2, Y < 3)$

我们要求 $P(X > 2, Y < 3)$ ，可以通过对概率密度函数在给定区域上的积分来计算：

$$P(X > 2, Y < 3) = \iint_R f(x, y) dx dy$$

这里 R 是满足 $x > 2$ 且 $y < 3$ 条件的区域。

1. 边缘分布 $P(X > 2)$ 和 $P(Y < 3)$

边缘分布是指对所有可能的 y 值上积分以得到 X 的概率分布，或者在所有可能的 x 值上积分以得到 Y 的概率分布。

$$P(X > 2) = \int_2^\infty \int_{-\infty}^\infty f(x, y) dy dx$$

$$P(Y < 3) = \int_{-\infty}^3 \int_{-\infty}^\infty f(x, y) dx dy$$

同样，这些计算可能会比较繁琐，可能需要使用数值积分方法或者计算机软件来进行估算。

请给出上述积分的数值解

很抱歉，由于我目前无法进行数值计算或使用实时数据，我不能直接提供数值解。然而，我可以指导你如何使用常见的数学软件（如 Python 中的 SciPy 库）来计算这些积分。

Appendix: 联合概率分布和边缘分布的关系

综上所述,

- 已知二维随机变量的联合分布可以唯一确定二维随机变量的边缘分布 (根据定义)
- 反之, 不成立 (why?)
- Copula 函数和 Sklar 定理

Appendix: Copula 函数和 Sklar 定理

考虑：二维随机变量的联合分布、边缘分布关系

Copula（耦合或称关联结构），为处理统计中随机变量相关性问题的方法，由一组随机变量的边缘分布来确定它们的联合分布。

- 通过边缘分布和关联结构可以确定一个联合分布
 - 通过分别将每个边缘分布都转换为平均分布的转换组成。这样，一个关联结构（dependence structure）就可以表达为一个基于上述所得平均分布之上的联合分布，而关联结构（copula）即是边缘均匀随机变量之上的一个联合分布。
- Sklar 定理
 - 必然存在一个 Copula 函数，使得

$$F = \text{Copula}(F_1, F_2, \dots, F_n)$$

Appendix: 正态分布的 Copula 函数和 Sklar 定理

已知

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \quad \text{和} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}},$$

思考：该 Copula 函数应该如何构造？