

随机变量及其分布

Didnelpsun

目录

| | | |
|----------|----------------------|----------|
| 1 | 分布概念 | 1 |
| 1.1 | 分布函数存在性 | 1 |
| 1.2 | 分布函数计算 | 1 |
| 1.3 | 概率密度存在性 | 2 |
| 1.4 | 概率密度计算 | 2 |
| 2 | 一维随机变量 | 2 |
| 2.1 | 一维随机变量分布 | 2 |
| 2.1.1 | 二项分布 | 2 |
| 2.1.2 | 泊松分布 | 3 |
| 2.1.3 | 几何分布 | 3 |
| 2.1.4 | 均匀分布 | 4 |
| 2.1.5 | 指数分布 | 5 |
| 2.1.6 | 正态分布 | 5 |
| 2.2 | 一维随机变量函数分布 | 6 |
| 3 | 二维随机变量 | 6 |
| 3.1 | 二维离散型随机变量 | 6 |
| 3.2 | 二维连续型随机变量 | 6 |
| 3.2.1 | 联合概率 | 6 |
| 3.2.1.1 | 联合概率密度 | 6 |
| 3.2.1.2 | 联合概率函数 | 7 |
| 3.2.2 | 边缘概率 | 7 |
| 3.2.2.1 | 边缘概率函数 | 7 |

| | | |
|---------|----------------------|----|
| 3.2.2.2 | 边缘概率密度 | 8 |
| 3.2.3 | 二维均匀分布 | 8 |
| 3.2.4 | 二维正态分布 | 8 |
| 3.2.4.1 | 正态分布性质 | 8 |
| 3.2.4.2 | 标准正态化 | 9 |
| 3.3 | 二维随机变量函数分布 | 9 |
| 3.3.1 | 离散型 | 9 |
| 3.3.2 | 连续型 | 9 |
| 3.3.2.1 | 和的分布 | 9 |
| 3.3.2.2 | 差的分布 | 10 |
| 3.3.3 | 混合型 | 10 |

分布函数变量区域左闭右开，概率密度则不要求。

1 分布概念

1.1 分布函数存在性

分布函数的充要条件是：

1. $F(x)$ 单调不减。
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ 。
3. $\lim_{x \rightarrow +\infty} F(x) = 1$ 。
4. $F(x)$ 右连续。

这也是下面的计算需要考虑的性质。

例题：设随机变量 X 的分布函数 $F(x)$ ，判断 $F(ax)$ 、 $F(x^2 + 1)$ 、 $F(x^3 - 1)$ 、 $F(|X|)$ 是否可以做出分布函数。

解：逐个分析 $F(g(x))$ 中的 $g(x)$ 走向。

$$\lim_{x \rightarrow +\infty} F(ax) = a \lim_{x \rightarrow +\infty} F(x) = a, \text{ 不满足。}$$

由于 $x^2 + 1$ 在 $x \rightarrow -\infty$ 的趋向为 $+\infty$ 而不是 $-\infty$ ， $\lim_{x \rightarrow -\infty} F(x^2 + 1) = \lim_{x \rightarrow +\infty} F(x^2 + 1) = \lim_{x \rightarrow +\infty} F(x) = 1 \neq 0$ ，不满足。

$F(x^3 - 1)$ 满足分布函数的四个条件， $g(x) = x^3 - 1$ 在两边趋向与原函数一致。

$F(|x|)$ 的 $g(x) = |x|$ 在 $x \rightarrow -\infty$ 的趋向为 $+\infty$ 而不是 $-\infty$ ，不满足 $\lim_{x \rightarrow -\infty} F(x) = 0$ ，不满足。

1.2 分布函数计算

这里所说的分布函数都是不定式，需要利用分布函数性质来换算。

例题：随机变量 X 的概率密度函数 $f(x)$ 满足 $f(1+x) = f(1-x)$ ，且 $\int_1^2 f(x) dx = 0.3$ ，则求分布函数 $F(0)$ 的值。

解：已知 $f(1+x) = f(1-x)$ ，所以 $f(x)$ 关于 $x = 1$ 对称。所以 $\int_1^{+\infty} f(x) dx = \int_1^{-\infty} f(x) dx = \frac{1}{2}$ 。

$$\text{所以 } \int_1^2 f(x) dx = \int_0^1 f(x) dx = 0.3。$$

$$F(0) = \int_{-\infty}^0 f(x) dx = 1 - \int_0^{+\infty} f(x) dx = 1 - \int_1^{+\infty} f(x) dx - \int_0^1 f(x) dx = 0.2。$$

1.3 概率密度存在性

分布函数的充要条件是：

1. $f(x) \geq 0$ 。
2. $\int_{-\infty}^{+\infty} f(x) dx = 1$ 。

例题：随机变量 X 的概率密度函数 $f(x)$ ，判断 $f(2x)$ 、 $f(2-x)$ 是否可以作出概率密度函数。

解： $\int_{-\infty}^{+\infty} f(2x) dx = \frac{1}{2} \int_{-\infty}^{+\infty} f(2x) d(2x) = \frac{1}{2} \neq 1$ ，不满足。

$\int_{-\infty}^{+\infty} f(2-x) dx = - \int_{-\infty}^{+\infty} f(2-x) d(2-x)$ ，令 $2-x=t$ ，则 $x \rightarrow -\infty$ 时 $t \rightarrow +\infty$ ， $x \rightarrow +\infty$ 时 $t \rightarrow -\infty$ ， $= - \int_{+\infty}^{-\infty} f(t) dt = \int_{-\infty}^{+\infty} f(t) dt = 1$ 。

1.4 概率密度计算

例题：随机变量 X 的分布函数和概率密度函数为 $F(x)$ ， $f(x)$ ，求 $-X$ 的分布函数和概率密度。

解：已知 $F(x) = P\{X \leq x\}$ ，则 $F'(-x) = \{-X \leq x\}$ ，即 $F'(-x) = P\{X \geq -x\} = 1 - P\{X \leq -x\} = 1 - F(-x)$ 。

对 $F'(-x)$ 进行求导： $[1 - F(-x)]' = f(-x)$ 。

例题：随机变量 X 的概率密度为 $f(x)$ ，求随机变量 $2X+3$ 的概率密度函数。

解：设 $Y = 2X+3$ ，所以 Y 的分布函数 $F_Y(y)$ 为 $F_Y(x) = P\{Y \leq x\} = P\{2X+3 \leq x\} = P\left\{X \leq \frac{x-3}{2}\right\} = \int_{-\infty}^{\frac{x-3}{2}} f(x) dx$ 。

所以其概率密度为 $f_Y(x) = \left(\int_{-\infty}^{\frac{x-3}{2}} f(x) dx\right)' = \frac{1}{2} f\left(\frac{x-3}{2}\right)$ 。

2 一维随机变量

2.1 一维随机变量分布

2.1.1 二项分布

$P\{X = k\} = C_n^k p^k (1-p)^{n-k}$ ($k = 0, 1, \dots, n$, $0 < p < 1$), $X \sim B(n, p)$ 。

例题：已知随机变量 X 的概率密度为 $f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$ ， Y 表示对 X 进行 3 次独立重复试验中出现事件 $\left\{X \leq \frac{1}{2}\right\}$ ，求 $P\{Y = 2\}$ 。

解：已知对 X 进行独立重复试验，表示这个进行的是伯努利试验，从而 $Y \sim B(n, p)$ 。又是 3 次，所以 $Y \sim B(3, p)$ 。

只用求出这个 p 即 $\left\{X \leq \frac{1}{2}\right\}$ 的概率就可以了。又已知 $f(x)$ 。

$$\therefore p = \left\{X \leq \frac{1}{2}\right\} = \int_0^{\frac{1}{2}} 2x \, dx = \frac{1}{4}。 \therefore P\{Y = 2\} = B\left(3, \frac{1}{4}\right) = \frac{9}{64}。$$

2.1.2 泊松分布

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, \dots, n, \lambda > 0), \quad X \sim P(\lambda)。$$

例题：设一本书的各页印刷错误的个数 X 服从泊松分布。已知只有一个和只有两个印刷错误的页数相同，则随机抽查的 4 页中无印刷错误的概率 p 为？

$$\text{解：} \because P\{X = 1\} = P\{X = 2\}, \therefore \frac{\lambda^1}{1!} e^{-\lambda} = \frac{\lambda^2}{2!} e^{-\lambda}, \quad \lambda = 2。$$

由于随机抽四页类似于伯努利试验是相互独立的，所以随机抽 4 页都无错误的概率为 $[P\{X = 0\}]^4 = e^{-8}$ 。

2.1.3 几何分布

$$P\{X = k\} = (1 - p)^{k-1} p \quad (k = 0, 1, \dots, n, 0 < p < 1), \quad X \sim G(p)。$$

例题：袋中有 8 个球，其中 3 个白球 5 个黑球，现在任意从中取出 4 个球，若四个球中有 2 个黑球和 2 个白球则试验停止，否则将其放回袋中重新抽取直到满足条件，用 X 表示试验次数，则求 $P\{X = k\}$ 。

解：由题目的停止，则说明这个题目的概率是服从几何分布的，最重要的就是求出单次满足事件概率 p 。

$$\text{根据组合和乘法原理，} p = \frac{C_3^2 C_5^2}{C_8^4} = \frac{3}{7}。$$

$$\text{则 } P\{X = k\} = \left(\frac{4}{7}\right)^{k-1} \cdot \frac{3}{7}。$$

例题：已知随机变量 X 的概率密度为 $f(x) = \begin{cases} 2^{-x} \ln 2, & x > 0 \\ 0, & \text{其他} \end{cases}$ ，对 X 进行独立重复观测，直到第 2 个大于 3 的观测值出现时停止，记 Y 为观测次数，求 Y 的概率分布。

解：由题目直到就停止，知道 $Y \sim G(p)$ 。

$$\text{又 } p = P\{X \geq 3\} = \int_3^{+\infty} 2^{-x} \ln 2 dx = \frac{1}{8}$$

这是对几何分布的变形, 首先进行 k 次试验, 第 k 次成功, 所以要乘 p , 而因为是第 2 个成功, 所以前面的 $k-1$ 次中有 $k-2$ 次失败和一次成功, 所以一共 $p^2(1-k)^{k-2}$ 。因为前面的成功的一次在 $k-1$ 中任意一个地方就可以了, 所以一共有 $k-1$ 中可能性, 要考虑到排列, 所以还要乘 $(k-1)$ 。

$$\therefore P\{Y = k\} = (k-1) \left(\frac{1}{8}\right)^2 \cdot \left(\frac{7}{8}\right)^{k-2}。$$

2.1.4 均匀分布

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases}, F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}, X \sim U(a, b)。$$

例题: 已知随机变量 $X \sim U(a, b)$ ($a > 0$) 且 $P\{0 < X < 3\} = \frac{1}{4}$, $P\{X > 4\} = \frac{1}{2}$, 求 X 的概率密度以及 $P\{1 < X < 5\}$ 。

$$\text{解: } \because P\{X > 4\} = \frac{1}{2}, 4 \text{ 在其区间中点上, } \frac{a+b}{2} = 4。$$

$$\because P\{0 < X < 3\} = \frac{1}{4}, 3 \text{ 若在 } a \text{ 左边则概率为 } 0, \text{ 所以必然在右边。}$$

$$\therefore P\{a < X < 3\} = \frac{1}{4}, P\{3 < X < 4\} = 1 - \frac{1}{4} - \frac{1}{2} = \frac{1}{4}, \frac{4-3}{b-a} = \frac{1}{4}。$$

$$\text{解得 } a = 2, b = 6, X \sim U(2, 6) = f(x) = \begin{cases} \frac{1}{4}, & 2 < x < 6 \\ 0, & \text{其他} \end{cases}。$$

$$P\{1 < X < 5\} = \frac{5-2}{6-2} = \frac{3}{4}。$$

例题: 已知随机变量 X 在区间 $[0, 1]$ 上服从均匀分布, 在 $X = x$ ($0 < x < 1$) 的条件下随机变量 Y 在区间 $[0, x]$ 上服从均匀分布。

(1) (X, Y) 的概率密度。

$$\text{解: } X \text{ 在区间 } [0, 1] \text{ 上服从均匀分布, 则 } X \sim f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}。$$

$$Y \text{ 在 } X = x \text{ 下均匀分布, 则 } f_{Y|X}(y|x) = \begin{cases} \frac{1}{x}, & 0 < y < x < 1 \\ 0, & \text{其他} \end{cases}。$$

(X, Y) 联合概率 = 条件概率 \times 边缘概率。

$$\text{即 } f(x, y) = f_{Y|X}(y|x)f_X(x) = \begin{cases} \frac{1}{x}, & 0 < y < x < 1 \\ 0, & \text{其他} \end{cases}。$$

(2) Y 的概率密度。

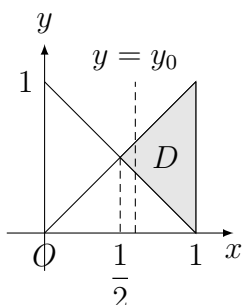
解: 首先求 Y 的边缘概率密度, 就需要积 X 。然后求 y 的区间, XY 的联

合区间是横坐标 $[0, 1]$ 到纵坐标 $[0, 1]$ 的下三角形, 则 $y \in [0, 1]$ 。

然后求 Y 就在联合概率密度所规定的区间中画一条 $y = y_0$ 的线, 从左先交到的是 $y = x$, 所以下限就是 y , 后交的是 $x = 1$, 所以上限为 1。最后将 y 的联合分布函数放在中间, 得到 $f_Y(y) = \begin{cases} \int_y^1 \frac{1}{x} dx = -\ln y, & 0 < y < x < 1 \\ 0, & \text{其他} \end{cases}$ 。

(3) 概率 $P\{X + Y > 1\}$ 。

解: 求 $P\{X + Y > 1\}$ 就是求一个区间的概率值, 即 $P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy$ 。



$$\begin{aligned} \text{所以 } P\{X + Y > 1\} &= \iint_D \frac{1}{x} d\sigma, \\ D &= x + y > 1 \cap 0 < y < x < 1. \\ \iint_D \frac{1}{x} d\sigma &= \int_{\frac{1}{2}}^1 dx \int_{1-x}^x \frac{1}{x} dy = 1 - \ln 2. \end{aligned}$$

2.1.5 指数分布

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{其他} \end{cases}, \quad F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad X \sim E(\lambda).$$

例题: 已知随机变量 $X \sim E(1)$, a 为常数且大于 0, 求 $P\{X \leq a+1 | X > a\}$ 。

$$\text{解: } P\{X \leq a+1 | X > a\} = \frac{P\{a < X \leq a+1\}}{P\{X > a\}} = \frac{\int_a^{a+1} e^{-x} dx}{\int_a^{+\infty} e^{-x} dx} = 1 - \frac{1}{e}.$$

也可以根据指数分布的无记忆性: $P\{X \leq a+1 | X > a\} = 1 - P\{X > a+1 | X > a\} = 1 - P\{X > 1\} = P\{X \leq 1\} = F(1) = 1 - \frac{1}{e}$ 。

例题: 随机变量 X 服从参数为 1 的指数分布, 求 $P\{3 > X > 2 | X > 1\}$ 。

已知 $F(X < x) = 1 - e^{-\lambda x}$, 则 $F(X > x) = e^{-\lambda x}$ 。且 $2 < X < 3 \cap 1 < X = 2 < X < 3$ 。

$$\begin{aligned} \text{则 } P\{3 > X > 2 | X > 1\} &= \frac{P\{3 > X > 2\}}{P\{X > 1\}} = \frac{P\{X > 2\} - P\{X > 3\}}{P\{X > 1\}} = \\ &= \frac{e^{-2} - e^{-3}}{e^{-1}} = e^{-1} - e^{-2}. \end{aligned}$$

2.1.6 正态分布

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty < x < +\infty, -\infty < \mu < +\infty, \sigma > 0), \quad X \sim N(\mu, \sigma^2).$$

例题：已知随机变量 $X \sim N(0,1)$ ，对给定的 α ($0 < \alpha < 1$)，数 μ_α 满足 $P\{X > \mu_\alpha\} = \alpha$ ，若 $P\{|X| < x\} = \alpha$ ，求 x 。

解： $P\{X > \mu_\alpha\} = \alpha$ 即表示 μ_α 为标准正态分布的上 α 分位点。

又 $P\{|X| < x\} = \alpha$ ，即 $-x < X < x$ 的面积为 α ，所以两边的面积各为 $\frac{1-\alpha}{2}$ ， $P\{X < x\} = P\{X > x\} = \frac{1-\alpha}{2}$ 。

\therefore 面积为 α 的下标为 α ， \therefore 面积为 $\frac{1-\alpha}{2}$ 的下标为 $\frac{1-\alpha}{2}$ ， $x = \mu_{\frac{1-\alpha}{2}}$ 。

2.2 一维随机变量函数分布

例题：随机变量 X 服从 $U(0,2)$ ，求随机变量 $Y = X^2$ 在 $(0,4)$ 内的概率分布密度 $f_Y(y)$ 。

解：求概率分布密度函数，可以求出其积分概率分布函数， $F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\} = P\{-\sqrt{y} \leq X \leq \sqrt{y}\}$ ，又 $X \sim U(0,2)$ ，所以 $f(x) = \frac{1}{2}$ 。

则概率分布函数就是概率密度的积分，此时已经将 Y 变为了关于 X 的积分， $= \int_{-\sqrt{y}}^{\sqrt{y}} f(x) dx = 2 \int_0^{\sqrt{y}} \frac{1}{2} dx = \frac{\sqrt{y}}{2}$ 。即 $F_Y(y) = \frac{\sqrt{y}}{2}$ 。

则 $f_Y(y) = F'_Y(y) = \frac{1}{4\sqrt{y}}$ 。

3 二维随机变量

使用定义法则直接用二重积分的分布函数来求，使用卷积公式则使用概率密度。

3.1 二维离散型随机变量

3.2 二维连续型随机变量

有两种主要方法，一个是分布函数法，一个是卷积公式法。如果给出联合概率密度使用分布函数法，如果给出边缘概率密度使用卷积公式法。

3.2.1 联合概率

3.2.1.1 联合概率密度

利用联合概率函数来计算，有两种方式，一种是以概率形式，一种是以积分形式。

例题：已知二维随机变量 (X_1, X_2) 的概率密度函数 $f_1(x_1, x_2)$ ，求 $Y_1 = 2X_1$ ， $Y_2 = \frac{1}{3}X_2$ 的联合概率密度 $f_2(y_1, y_2)$ 。

解：概率形式：设 $F_2(y_1, y_2) = P\{Y_1 \leq y_1, Y_2 \leq y_2\} = P\{2X_1 \leq y_1, \frac{1}{3}X_2 \leq y_2\} = P\{X_1 \leq \frac{y_1}{2}, X_2 \leq 3y_2\} = F_1(\frac{y_1}{2}, 3y_2)$ 。所以 $f_2(y_1, y_2) = \frac{3}{2}f_1(\frac{y_1}{2}, 3y_2)$ 。

积分形式：已知 $\int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 = 1$ ，将 $x_1 = \frac{y_1}{2}$ ， $x_2 = 3y_2$ 代入 $\int_{-\infty}^{+\infty} d(\frac{y_1}{2}) \int_{-\infty}^{+\infty} f(\frac{y_1}{2}, 3y_2) d(3y_2) = 1$ ， $\frac{3}{2} \int_{-\infty}^{+\infty} dy_1 \int_{-\infty}^{+\infty} f(\frac{y_1}{2}, 3y_2) dy_2 = 1$ ，所以 $f_2(y_1, y_2) = \frac{3}{2}f_1(\frac{y_1}{2}, 3y_2)$ 。

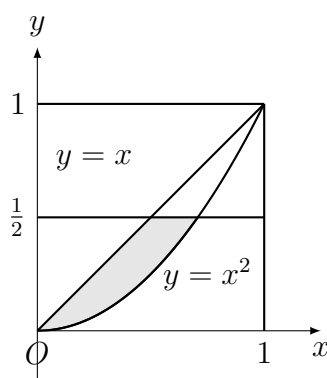
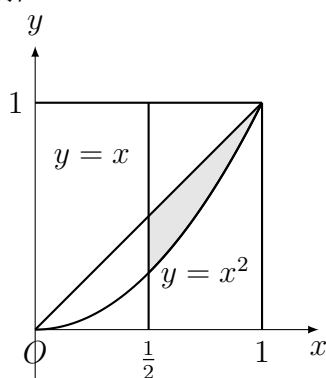
3.2.1.2 联合概率函数

已知联合概率密度，可以求概率函数，通过二重积分的方式，图像面积即是概率。

例题：已知概率密度为 $f(x, y) = \begin{cases} 6, & 0 < x^2 < y < x < 1 \\ 0, & \text{其他} \end{cases}$ ，

求 $P\left\{X > \frac{1}{2}\right\}$ ， $P\left\{Y < \frac{1}{2}\right\}$ 。

解：



$$P\left\{X > \frac{1}{2}\right\} = 6 \int_{\frac{1}{2}}^1 dx \int_{x^2}^x dy = 6 \int_{\frac{1}{2}}^1 (x - x^2) dx = 6 \left(\frac{1}{2}x^2 - \frac{1}{3}x^3 \right) \Big|_{\frac{1}{2}}^1 = \frac{1}{2}.$$

$$P\left\{Y < \frac{1}{2}\right\} = 6 \int_0^{\frac{1}{2}} dy \int_y^{\sqrt{y}} dx = 6 \int_0^{\frac{1}{2}} (\sqrt{y} - y) dy = \sqrt{2} - \frac{3}{4}.$$

3.2.2 边缘概率

3.2.2.1 边缘概率函数

往往是已知联合概率函数 $F(x, y)$ 求边缘概率函数 $F_X(x)$ 、 $F_Y(y)$ ，需要将联合概率函数中的 $x/y \rightarrow +\infty$ ，然后求这个函数的极限值。

例题：如果二维随机变量 (X, Y) 的分布函数为 $F(x, y) =$

$$\begin{cases} 1 - e^{-\lambda_1 x} - e^{-\lambda_2 y} + e^{-\lambda_1 x - \lambda_2 y - \lambda_{12} \max\{x, y\}}, & \lambda_1, \lambda_2, \lambda_{12} > 0, x > 0, y > 0 \\ 0, & \text{其他} \end{cases},$$

求 XY 各自的边缘分布函数。

解: $\lim_{x \rightarrow +\infty} \max\{x, y\} = x = +\infty$, $\lim_{y \rightarrow +\infty} \max\{x, y\} = y = +\infty$ 。

$F_X(x) = F(x, +\infty) = 1 - e^{-\lambda_1 x} - 0 + 0 = 1 - e^{-\lambda_1 x}$, $x > 0$, 当其他时 $= 0$ 。

$F_Y(y) = F(+\infty, y) = 1 - 0 - e^{-\lambda_2 y} + 0 = 1 - e^{-\lambda_2 y}$, $y > 0$, 当其他时 $= 0$ 。

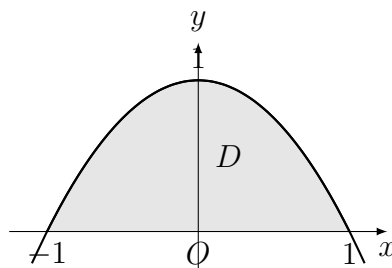
3.2.2.2 边缘概率密度

往往是已知联合概率密度 $f(x, y)$ 求边缘概率密度 $f_X(x)$ 、 $f_Y(y)$ ，需要将联合概率密度对另一个变量进行上下限无穷的一重积分，如果 xy 有上下限的定义域则需要画出图像取交集。

确定上下限时要注意，如果求 x 的边缘分布对 y 积分，表示 x 不动，求 y 的范围，求 y 的则反之。

例题：求 $f(x, y) = \begin{cases} \frac{5}{4}(x^2 + y), & 0 < y < 1 - x^2 \\ 0, & \text{其他} \end{cases}$ 的边缘概率密度。

如果求 x 的边缘概率密度，则 y 的取值范围为最底部的 0 到函数 $1 - x^2$ 。如果求 y 的边缘密度，则发现 D 为对称函数，所以可以拆为左右两个部分， x 的范围是 0 到函数 $\sqrt{1 - y}$ 。



$$f_X(x) = \frac{5}{4} \int_0^{1-x^2} x^2 + y \, dy = \frac{5}{4} \left(x^2 y + \frac{y^2}{2} \right) \Big|_0^{1-x^2} = -\frac{5}{8} x^4 + \frac{5}{8}.$$

$$f_Y(y) = 2 \frac{5}{4} \int_0^{\sqrt{1-y}} x^2 + y \, dx = \frac{5}{2} \left(\frac{x^3}{3} + yx \right) \Big|_0^{\sqrt{1-y}} = \frac{5}{6} (1 - y) \sqrt{1 - y} + \frac{5}{2} y \sqrt{1 - y}.$$

3.2.3 二维均匀分布

3.2.4 二维正态分布

概率密度为：

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right] \right\}$$

其中 $\mu_1, \mu_2 \in R$, $\sigma_1, \sigma_2 > 0$, $-1 < \rho < 1$ 。

3.2.4.1 正态分布性质

例题： $(X, Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; 0)$ ，分布函数为 $F(x, y)$ ，已知 $F(\mu_1, y) = \frac{1}{4}$ ，求 y 。

解：当 $\rho = 0$ 时， $F(X, Y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}$
 $= F_X(x)F_Y(y)$ ，即 XY 相互独立。

$X \sim N(\mu_1, \sigma_1)$ ， $Y \sim N(\mu_2, \sigma_2)$ ， $F_X(\mu_1) = P\{X \leq \mu_1\} = \frac{1}{2}$ 。

$F(\mu_1, y) = F_X(\mu_1)F_Y(y) = \frac{1}{2}F_Y(y) = \frac{1}{4}$ ，则 $F_Y(y) = \frac{1}{2}$ ，即根据性质 $y = \mu_2$ 。

3.2.4.2 标准正态化

$$F(x) = P\{X \leq x\} = P\left\{\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right\} = \Phi\left(\frac{x - \mu}{\sigma}\right)。$$

即将 XY 的相关系数消去。

例题：设随机变量 (X, Y) 的分布函数为 $\Phi(2x+1) \cdot \Phi(2y-1)$ ，其中 $\Phi(x)$ 为标准正态分布函数，求 (X, Y) 的分布函数。

解：由分布函数为 $\Phi(2x+1)\Phi(2y-1)$ 是 X 的分布函数和 Y 的分布函数的乘积，所以可知 XY 相互独立。

$$\text{所以根据标准化公式：}\Phi(2x+1) \cdot \Phi(2y-1) = \Phi\left(\frac{x + \frac{1}{2}}{\frac{1}{2}}\right) \Phi\left(\frac{y - \frac{1}{2}}{\frac{1}{2}}\right)。$$

$$\therefore (X, Y) \sim N\left(-\frac{1}{2}, \frac{1}{2}; \frac{1}{4}, \frac{1}{4}; 0\right)。$$

3.3 二维随机变量函数分布

3.3.1 离散型

3.3.2 连续型

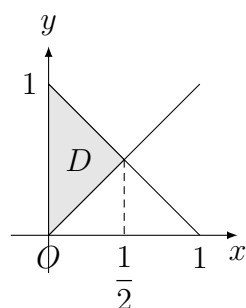
可以使用卷积公式法和分布函数法两种。

3.3.2.1 和的分布

例题：随机变量 (X, Y) 的概率密度函数 $f(x, y) = \begin{cases} e^{-y}, & 0 < x < y \\ 0, & \text{其他} \end{cases}$ ，

求 $P\{X + Y \leq 1\}$ 。

解：根据 $X + Y \leq 1$ 和 $0 < x < y$ 划分区域：



其中积分区域 D 如图所示，所

$$\text{以 } P\{X + Y \leq 1\} = \iint_D e^{-y} dx dy = \int_0^{\frac{1}{2}} dx \int_x^{1-x} e^{-y} dy = \int_0^{\frac{1}{2}} (e^{-x} - e^{x-1}) dx = 1 + e^{-1} - 2e^{-\frac{1}{2}}.$$

3.3.2.2 差的分布

例题：设 $X \sim N(\mu, \sigma_1^2)$, $Y \sim N(2\mu, \sigma_2^2)$, XY 相互独立, 已知 $P\{X - Y \geq 1\} = \frac{1}{2}$, 求 μ 。

解：若 $X \sim N(\mu, \sigma_1^2)$, $Y \sim N(2\mu, \sigma_2^2)$, 则 $X - Y \sim (-\mu, \sigma_1^2 - \sigma_2^2)$ 。则 $X - Y$ 的均值为 $-\mu$, 即其图像的对称轴为 $-\mu$ 。

又 $P\{X - Y \geq 1\} = \frac{1}{2}$, 则 $X - Y$ 在 1 这里均分, 则对称轴为 1, 即 $\mu = -1$ 。

3.3.3 混合型

使用全概率公式根据离散变量进行概率拆分。

例题：设随机变量 X_1 和 X_2 相互独立, 已知 $X_1 \sim B\left(1, \frac{3}{4}\right)$, X_2 的分布函数为 $F(x)$, 求 $Y = X_1 + X_2$ 的分布函数 $F_Y(y)$ 。

解：已知 $X_1 \sim B\left(1, \frac{3}{4}\right)$, X_2 的分布函数为 $F(x)$, 则 Y 为混合型。

则 $P\{X_1 = 0\} = C_1^0 \frac{3^0}{4} \frac{1^1}{4} = \frac{1}{4}$, $P\{X_1 = 1\} = C_1^1 \frac{3^1}{4} \frac{1^0}{4} = \frac{3}{4}$ 。

$F_Y(y) = P\{X_1 + X_2 \leq y\} = P\{X_1 + X_2 \leq y | X_1 = 0\}P\{X_1 = 0\} + P\{X_1 + X_2 \leq y | X_1 = 1\}P\{X_1 = 1\} = P\{X_2 \leq y | X_1 = 0\}P\{X_1 = 0\} + P\{1 + X_2 \leq y | X_1 = 1\}P\{X_1 = 1\}$, 由相互独立性 $= \frac{1}{4} \cdot P\{X_2 \leq y\} + \frac{3}{4} \cdot P\{X_2 \leq y - 1\}$ 。

根据分布函数定义, 则 $F_Y(y) = \frac{1}{4} \cdot F(y) + \frac{3}{4} \cdot F(y - 1)$ 。