AI 中的数学

第九、十讲

方聪，概率统计部分参考章复熹和张原老师课件

2024 年秋季

1. 多元随机变量
2. 二元正态分布
3. 条件分布
4. 随机变量的独立性

- 例 1.3. 考察钢的硬度 X 与含碳量 Y, 含硫量 Z 之间的关系.
- 定义 1.1&1.1′. 设 $X_1, \cdots, X_n$ 是同一个概率空间 $(\Omega, \mathcal{F}, P)$ 上的随机变量, 则称
  $\xi = \vec{X} = (X_1, \cdots, X_n) : \Omega \to \mathbb{R}^n, \omega \mapsto (X_1(\omega), \cdots, X_n(\omega))$
  为 $(\Omega, \mathcal{F}, P)$ 上的 (n 维) 随机向量 / 变量.
- 定义 1.2. n 维随机向量的函数指新变量 $Y = f(X_1, \cdots, X_n)$, 其中 $f : \mathbb{R}^n \to \mathbb{R}, \quad (x_1, \cdots, x_n) \mapsto f(x_1, \cdots, x_n)$.
- 例 1.6. 三维空间中的一个随机点 (X, Y, Z) 与原点的距离为

$$f(X, Y, Z) = \sqrt{X^2 + Y^2 + Z^2}.$$

n 维随机向量：称 n 个随机变量 $X_1, \cdots, X_n$ 的整体 $\xi = (X_1, \cdots, X_n)$ 为 n 维随机向量（或者 n 维随机变量），一维随机向量简称随机变量。
n 维随机变量数学上的精确定义：设 $X_1 = X_1(\omega), \cdots, X_n = X_n(\omega)$ 都是概率空间 $(\Omega, \mathcal{F}, P)$ 上的随机变量，则称

$$\xi = \xi(\Omega) \triangleq (X_1(\omega), \cdots, X_n(\omega))$$

为概率空间 $(\Omega, \mathcal{F}, P)$ 上的 n 维随机向（变）量。

例如，用炮弹向远处目标攻击，炮弹的落点用平面坐标系中的坐标表示为 (X, Y)，是一个二维随机向量。

随机向量的函数：设 $X_1 = X_1(\omega), \cdots, X_n = X_n(\omega)$ 是 n 个随机变量，$f(x_1, \cdots, x_n)$ 是 n 元实值函数，则称随机变量 $Y \triangleq f(x_1, \cdots, x_n)$ 为随机变量 $X_1, \cdots, X_n$ 的函数（即随机向量 $(X_1, \cdots, X_n)$ 的函数）。

离散型情形

§3.2 二维随机向量的联合分布与边缘分布,§3.7 条件分布

- 定义 2.1&2.2. 若 $\xi = (X, Y)$ 取有限个或可列个"值"(二维向量), 则称 ξ 为离散型.
- ξ 是离散型当且仅当 X, Y 都是离散型.
- 定义 2.2. 设 X, Y 的可能值分别为 x_i, y_j, 则称

$$p_{ij} = P\left(X = x_i, Y = y_j\right), \quad i, j = 1, 2, \cdots$$

为 ξ 的联合分布 (列).
- 联合分布列满足: $p_{ij} \geqslant 0, i, j = 1, 2, \cdots$ (非负性);

$$\sum_{i,j} p_{ij} = 1 \text{ (规范性)}.$$

例 2.2&2.3：有大量粉笔，含白、黄、红三种颜色，比例分别为 p_1，p_2, p_3. 从中抽取 n 支. 求：恰好抽到 k_1 支白，k_2 支黄的概率.

- 设恰好抽到 X 支白，Y 支黄，即求 $(X, Y) = (k_1, k_2)$ 的概率.
- 可以理解为放回抽样，连续抽取 n 次.
- 所求事件包含了

$$C_n^{k_1} C_{n-k_1}^{k_2} = \frac{n!}{k_1! k_2! (n - k_1 - k_2)!}$$

个基本事件，其中，每一个的概率都为

$$p_1^{k_1} p_2^{k_2} p_3^{n-k_1-k_2}.$$

- 故，$\forall k_1, k_2 \geqslant 0, k_1 + k_2 \leqslant n$,

$$P(X = k_1, Y = k_2) = C_n^{k_1} C_{n-k_1}^{k_2} p_1^{k_1} p_2^{k_2} p_3^{n-k_1-k_2}.$$

- 称 $\xi = (X, Y)$ 服从三项分布.

例：有一大批量粉笔，其中 60% 是白的，25% 是黄的，15% 是红的，现从中随机的依次取出 6 支，问：其中恰有 3 支白色，1 支黄色，2 支红色的概率是多少？

例：有一大批量粉笔，其中 60% 是白的，25% 是黄的，15% 是红的，现从中随机的依次取出 6 支，问：其中恰有 3 支白色，1 支黄色，2 支红色的概率是多少？

解：令 X = "6 支中白粉笔的个数"，Y = "6 支中黄粉笔的个数"，则事件 "6 支中恰有 3 支白色，1 支黄色，2 支红色" 就是事件

$$\{X=3, Y=1\}, \text{即}\{(X,Y)=(3,1)\}.$$

由三项分布，概率可表示为

$$P((X,Y)=(3,1)) = \frac{6!}{3!1!2!}0.6^3 \times 0.25 \times 0.15^2.$$

用组合数方法同样可以得到上述结果。

一般的，对于满足 $k_1 \geqslant 0, k_2 \geqslant 0$ 及 $k_1+k_2 \leqslant 6$ 的 k_1, k_2，由三项分布有

$$P((X,Y)=(k_1,k_2)) = \frac{6!}{k_1!k_2!(6-k_1-k_2)!}0.6^{k_1}\times 0.25^{k_2}\times 0.15^{6-k_1-k_2}.$$

二维随机向量的边缘分布：对于二维随机向量 $\xi = (X, Y)$，分量 X 的概率分布称为 ξ 关于 X 的边缘分布，分量 Y 的概率分布称为 ξ 关于 Y 的边缘分布。

二维随机向量 $\xi = (X, Y)$ 的两个边缘分布均由 ξ 的概率分布完全确定。

例：从 $1,2,3,4$ 中任取一数记为 X，再从 $1,\cdots,X$ 中任取一数记为 Y，求 $(X\ Y)$ 的联合分布列及 $P(X=Y)$。
解：易知 X 的分布列为：

$$P(X=i)=\frac{1}{4},\quad i=1,2,3,4.$$

显然，$P(X=i,Y=j)=0, j>i, i=1,2,3,4$，当 $1\leqslant j\leqslant i\leqslant 4$ 时，由乘法公式得

$$P(X=i,Y=j)=P(X=i)P(Y=j|X=i)=\frac{1}{4}\times\frac{1}{i}=\frac{1}{4i}.$$

Y \ X	1	2	3	4
1	$\frac{1}{4}$	0	0	0
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

$$P(X=Y)=\sum_{i=1}^{4}P(X=Y=i)=\sum_{i=1}^{4}\frac{1}{4i}=\frac{25}{48}.$$

例：（对应郑书例 2.5）设随机变量 X 取值是 0 或 1，随机变量 Y 取值也是 0 或 1，且二维随机向量 (X,Y) 的概率分布是

$$P((X,Y)=(0,0))=\frac{1}{4}+\varepsilon,\quad P((X,Y)=(0,1))=\frac{1}{4}-\varepsilon,$$

$$P((X,Y)=(1,0))=\frac{1}{4}-\varepsilon,\quad P((X,Y)=(1,1))=\frac{1}{4}+\varepsilon,$$

其中 $0\leqslant\varepsilon\leqslant\frac{1}{4}$。

易知不同的 ε 对应不同的联合分布，但是

$$P(X=0)=P((X,Y)=(0,0))+P((X,Y)=(0,1))=\frac{1}{2},$$

$$P(X=1)=P((X,Y)=(1,0))+P((X,Y)=(1,1))=\frac{1}{2}.$$

同理，

$$P(Y=0)=P(Y=1)=\frac{1}{2},$$

有无穷多个不同的联合分布具有相同的边缘分布。

2. 连续型情形

- 定义 2.4. 设 $\xi = (X, Y)$. 若存在 $p(x, y)$ 使得

$$P(\xi \in D) = \iint_D p(x, y)dxdy,$$

  对任意开矩形 D 成立, 则称 ξ 为连续型随机向量, 称 $p(x, y)$ 为 ξ 的联合密度 (函数), 也记为 $p_{X,Y}(x, y)$.
- 联合密度满足:

$$p(x, y) \geqslant 0; \quad \iint_{\mathbb{R}^2} p(x, y)dxdy = 1.$$

- $\star$ 对更一般的集合 D 都成立, 例如, D 是单位圆盘.

例：（对应郑书例 2.6）设二维随机向量 $\xi=(X,Y)$ 的联合密度为

$$p(x,y)=\begin{cases} ce^{-(x+y)}, & x\geqslant 0\text{且}y\geqslant 0,\\ 0, & \text{其他},\end{cases}$$

其中 c 是一个常数，求：

（1）c 的值；　　（2）$P(0<X<1,0<Y<1)$.

解：（1）由归一性知

$$1=\int_0^{+\infty}\int_0^{+\infty}ce^{-(x+y)}dxdy=c\int_0^{+\infty}e^{-x}dx\cdot\int_0^{+\infty}e^{-y}dy$$

于是 $c=1$。

（2）取 $D=\{(x,y):0<x<1,0<y<1\}$，由定义知

$$\begin{aligned}P(0<X<1,0<Y<1)&=\int_0^1\int_0^1 e^{-(x+y)}dxdy\\&=\int_0^1 e^{-x}dx\cdot\int_0^1 e^{-y}dy=(1-e^{-1})^2.\end{aligned}$$

设 G 是平面上面积为 $a(0 < a < +\infty)$ 的区域，称二维随机向量 $\xi = (X, Y)$ 服从 G 上的均匀分布，若 $P((X, Y) \in G) = 1$，且 (X, Y) 取值属于 G 的任何部分 A（A 是 G 的子区域）的概率与 A 的面积成正比。容易推知二维随机向量 $\xi = (X, Y)$ 有联合密度为

$$p(x,y) = \begin{cases} \frac{1}{a}, & (x,y) \in G, \\ 0, & \text{其他}, \end{cases} \tag{1}$$

- 定理 2.1. 若 $\xi = (X, Y)$ 是连续型，则 X, Y 都是连续型，且

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x, y) dy, \quad p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(x, y) dx.$$

- 称 $p_X(\cdot)$ 与 $p_Y(\cdot)$ 为 ξ 的边缘密度.

例 2.7. G 为由 $y=x^2$ 和 $y=x$ 所围成的有限区域. $\xi\sim U(G)$.
求: ξ 的联合密度与边缘密度.

- G 的面积: $a=\int_0^1 xdx-\int_0^1 x^2dx=\frac{1}{6}$.
- 联合密度: $p(x,y)=6,(x,y)\in G$.
- 边缘密度:

$$p_X(x)=\int_{-\infty}^{\infty}p(x,y)dy=\int_{x^2}^{x}6dy=6\left(x-x^2\right),\quad 0<x<1.$$

$$p_Y(y)=\int_{-\infty}^{\infty}p(x,y)dx=\int_{y}^{\sqrt{y}}6dx=6(\sqrt{y}-y),\quad 0<y<1.$$

- 注: X,Y 都取遍 $(0,1)$, 但 ξ 不能取遍 $(0,1)\times(0,1)$.

例：设二维随机向量 $\xi=(X,Y)$ 有联合密度

$$p_1(x,y)=\frac{1}{2\pi}\exp\{-\frac{1}{2}(x^2+y^2)\},$$

二维随机向量 $\eta=(U,V)$ 有联合密度

$$p_2(x,y)=\begin{cases}2p_1(x,y), & xy\geqslant 0,\\ 0, & \text{其他}.\end{cases}$$

则 X 与 U 有相同的分布密度，Y 与 V 有相同的分布密度。

一方面，当 $x \leqslant 0$ 时，

$$\int_{-\infty}^{+\infty} p_2(x,y)dy = \int_{-\infty}^{0} 2p_1(x,y)dy = \frac{1}{\pi}\int_{-\infty}^{0} e^{-(x^2+y^2)/2}dy$$

$$= \frac{1}{\pi}e^{-x^2/2}\int_{-\infty}^{0} e^{-y^2/2} = \frac{1}{\pi}e^{-x^2/2}\frac{1}{2}\int_{-\infty}^{+\infty} e^{-y^2/2} = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

类似的，当 $x > 0$ 时，

$$\int_{-\infty}^{+\infty} p_2(x,y)dy = \int_{0}^{+\infty} 2p_1(x,y)dy$$

$$= \frac{1}{\pi}\int_{0}^{+\infty} e^{-(x^2+y^2)/2}dy = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

即，对一切 x，$\int_{-\infty}^{+\infty} p_2(x,y)dy = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$。

同理，对一切 y，$\int_{-\infty}^{+\infty} p_2(x,y)dx = \frac{1}{\sqrt{2\pi}}e^{-y^2/2}$。

- 定义 2.6, 例 2.8& 例 7.5. 若 $\xi=(X,Y)$ 的联合密度 $p(x,y)$ 有如下表达式, 则称 ξ 服从二维 (元) 正态分布.

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{u^2+v^2-2\rho uv}{2\left(1-\rho^2\right)}\right\},$$

其中,

$$u=\frac{x-\mu_1}{\sigma_1},\quad v=\frac{y-\mu_2}{\sigma_2},$$

有 5 个参数: $\mu_1,\mu_2\in\mathbb{R}$,

$$\sigma_1,\sigma_2>0$$
$$\rho\in(-1,1)$$

一般二维随机向量及其联合分布函数：设 $\xi = (X, Y)$ 是二维随机向量，则称

$$F(x,y) = P(X \leqslant x, Y \leqslant y) \quad (x, y \in \mathbb{R})$$

为 ξ 的分布函数。也称为 (X, Y) 的联合分布函数。

分布函数 $F(x,y)$ 有以下性质：

（1） $0 \leqslant F(x,y) \leqslant 1$;

（2） $F(x,y)$ 是 x 的右连续增函数，也是 y 的右连续增函数;

（3） $\lim_{x\to-\infty} F(x,y) = 0, \lim_{y\to-\infty} F(x,y) = 0$;

（4） $\lim_{x\to+\infty} F(x,y) = P(Y \leqslant y), \lim_{y\to+\infty} F(x,y) = P(X \leqslant x)$;

（5） 对任何 $x_1 \leqslant x_2, y_1 \leqslant y_2$，有

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

关于性质（5），对一切 $x_1 \leqslant x_2, y_1 \leqslant y_2$，有

$$
\begin{aligned}
& P(x_1 < X \leqslant x_2, y_1 < Y \leqslant y_2) \\
& = P(x_1 < X \leqslant x_2, Y \leqslant y_2) - P(x_1 < X \leqslant x_2, Y \leqslant y_1) \\
& = P(X \leqslant x_2, Y \leqslant y_2) - P(X \leqslant x_1, Y \leqslant y_2) \\
& - [P(X \leqslant x_2, Y \leqslant y_1) - P(X \leqslant x_1, Y \leqslant y_1)] \\
& = F(x_2, y_2) - F(x_1, y_2) - [F(x_2, y_1) - F(x_1, y_1)]
\end{aligned}
$$

由 $P(x_1 < X \leqslant x_2, y_1 < Y \leqslant y_2) \geqslant 0$ 知性质（5）成立。

若二维随机向量 $\xi=(X,Y)$ 有联合密度 $p(x,y)$，则 ξ 的联合分布函数 $F(x,y)$ 与联合密度 $p(x,y)$ 有关系式

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)dudv. \tag{2}$$

例：设二维随机向量 (X, Y) 有密度函数

$$p(x,y)=\begin{cases} Ce^{-(2x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

求（1）常数 C 的值；（2）联合分布函数 $F(x,y)$；（3）概率 $P(X \leqslant Y)$。

例：设二维随机向量 (X, Y) 有密度函数

$$p(x,y)=\begin{cases} Ce^{-(2x+y)}, & x>0, y>0, \\ 0, & \text{其他}. \end{cases}$$

求（1）常数 C 的值；（2）联合分布函数 $F(x,y)$；（3）概率 $P(X \leqslant Y)$。
解：（1）由于

$$1 = C\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = C\int_0^{+\infty} e^{-2x}dx\int_0^{+\infty} e^{-y}dy = \frac{C}{2}$$

得 $C=2$。

(2) 利用公式

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}p(t,r)dtdr$$

$$=\begin{cases}\int_0^y\int_0^x 2e^{-2t-r}dtdr, & x>0,y>0,\\ 0, & \text{其他},\end{cases}$$

$$=\begin{cases}(1-e^{-2x})(1-e^{-y}), & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

(3) 设区域 $G=\{(x,y)|x\leqslant y\}$，则

$$P(X\leqslant Y)=P((X,Y)\in G)=\iint_G p(x,y)dydx$$

$$=\int_0^{+\infty}\int_x^{+\infty}2e^{-2x-y}dydx=\frac{2}{3}.$$

例：设 (X, Y) 的联合密度函数为

$$p(x,y)=\begin{cases} 1, & 0<x<1, |y|<x, \\ 0, & \text{其他}. \end{cases}$$

求边际密度函数。

解：根据定义

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy=\begin{cases} \int_{-x}^{x}1dy=2x & x\in(0,1), \\ 0, & x\notin(0,1). \end{cases}$$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\begin{cases} \int_{y}^{1}1dx=1-y & y\in[0,1), \\ \int_{-y}^{1}1dx=1+y, & y\in(-1,0), \\ 0, & y\notin(-1,1). \end{cases}$$

- 联合密度：$u=\frac{x-\mu_1}{\sigma_1}, v=\frac{y-\mu_2}{\sigma_2}$,

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{u^2+v^2-2\rho uv}{2\left(1-\rho^2\right)}\right\}.$$

- 边缘密度：$X\sim N\left(\mu_1,\sigma_1^2\right), Y\sim N\left(\mu_2,\sigma_2\right)^2$. 例如，

$$p_X(x)=\int_{-\infty}^{\infty}p(x,y)dy$$

$$=\int_{-\infty}^{\infty}\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}\exp\left\{-\frac{(v-2\rho u)^2+\left(1-\rho^2\right)u^2}{2\left(1-\rho^2\right)}\right\}dv$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}e^{-\frac{\left(1-\rho^2\right)u^2}{2\left(1-\rho^2\right)}}\int_{-\infty}^{\infty}\exp\left\{-\frac{(v-2\rho u)^2}{2\left(1-\rho^2\right)}\right\}dv$$

$$=\frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}}e^{-\frac{1}{2}u^2}\sqrt{2\pi\left(1-\rho^2\right)}=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}.$$

二维正态分布：若 $\xi=(X,Y)$ 的联合密度 $p(x,y)$ 有如下表达式，则称 ξ 服从二维（元）正态分布。

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{u^2+v^2-2\rho uv}{2\left(1-\rho^2\right)}\right\}, \tag{3}$$

其中，

$$u=\frac{x-\mu_1}{\sigma_1}, \quad v=\frac{y-\mu_2}{\sigma_2},$$

共有 5 个参数：$\mu_1,\mu_2\in\mathbb{R},\ \sigma_1,\sigma_2>0,\rho\in(-1,1)$

例：设二维随机向量 $\xi=(X,Y)$ 服从二维正态分布，试求出 X 的分布密度和 Y 的分布密度。

解：设 X 的分布密度为 $p_X(x)$，做变量代换 $v=\frac{y-\mu_2}{\sigma_2}$，得

$$
\begin{aligned}
p_X(x) &= \int_{-\infty}^{+\infty} p(x,y)dy \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right\} \\
&\cdot \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\frac{y-\mu_2}{\sigma_2}\frac{x-\mu_1}{\sigma_1}\right]\right\} dy \\
&= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right\} \cdot \\
&\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2\left(1-\rho^2\right)}[v^2 - 2\rho v\frac{x-\mu_1}{\sigma_1}]\right\} dv
\end{aligned}
$$

其中

$$\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2\left(1-\rho^{2}\right)}\left[v^{2}-2\rho v\frac{x-\mu_1}{\sigma_1}\right]\right\}dv$$

$$=\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2\left(1-\rho^{2}\right)}\left[\left(v-\rho\frac{x-\mu_1}{\sigma_1}\right)^{2}-\rho^{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^{2}\right]\right\}dv$$

$$=\exp\left\{\frac{\rho^{2}}{2(1-\rho^{2})}\left(\frac{x-\mu_1}{\sigma_1}\right)^{2}\right\}\cdot$$

$$\int_{-\infty}^{+\infty}\exp\left\{-\frac{1}{2\left(1-\rho^{2}\right)}\left(v-\rho\frac{x-\mu_1}{\sigma_1}\right)^{2}\right\}dv$$

$$=\exp\left\{\frac{\rho^{2}}{2(1-\rho^{2})}\left(\frac{x-\mu_1}{\sigma_1}\right)^{2}\right\}\sqrt{2\pi(1-\rho^{2})}.$$

于是

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}.$$

同理知

$$p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}.$$

这表明 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$.

例：假定 $(\xi_1, \xi_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，试求 (ξ_1, ξ_2) 落在

$$D = \left\{ (x, y) \middle| \left(\frac{x - \mu_1}{\sigma_1} \right)^2 - 2\rho \frac{(x - \mu_1)(y - \mu_2)}{\sigma_1 \sigma_2} + \left(\frac{y - \mu_2}{\sigma_2} \right)^2 \leqslant \lambda^2 \right\}$$

内的概率。

解：所求概率

$$\iint_D p(x,y)dxdy = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$

$$\iint_D \exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\} d$$

做变量代换 $u = \frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}$，$v = \sqrt{1-\rho^2}\frac{y-\mu_2}{\sigma_2}$，则

$$\left|\frac{\partial(u,v)}{\partial(x,y)}\right| = \begin{vmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_2} & \frac{\sqrt{1-\rho^2}}{\sigma_2} \end{vmatrix} = \frac{\sqrt{1-\rho^2}}{\sigma_1\sigma_2}, \quad |J| = \frac{\sigma_1\sigma_2}{\sqrt{1-\rho^2}}.$$

从而

$$\begin{aligned}\iint_D p(x,y)dxdy &= \frac{1}{2\pi(1-\rho^2)}\iint_{\{u^2+v^2\}\leqslant\lambda^2}\exp\left\{-\frac{u^2+v^2}{2(1-\rho^2)}\right\}dudv\\ &= \frac{1}{2\pi(1-\rho^2)}\int_0^{2\pi}\int_0^{\lambda}\exp\left\{-\frac{r^2}{2(1-\rho^2)}\right\}rdrd\theta\\ &= \int_0^{\frac{\lambda^2}{2(1-\rho^2)}} e^{-t}dt = 1-\exp\left\{-\frac{\lambda^2}{2(1-\rho^2)}\right\}.\end{aligned}$$

条件分布函数：设 X 和 Y 是两个随机变量，给定实数 y，如果 $P(Y=y)>0$，则称 x 的函数 $P(X\leqslant x|Y=y)$ 为在 $Y=y$ 的条件下 X 的条件分布函数，记作 $F_{X|Y}(x|y)$，显然，根据条件概率的定义，有

$$F_{X|Y}(x|y)=\frac{P(X\leqslant x,Y=y)}{P(Y=y)}.$$

离散型

设 (X,Y) 是二维离散型随机向量，其概率分布为

$$P(X=x_i, Y=y_j)=p_{ij}\quad (i=1,2,\cdots;j=1,2,\cdots),$$

这里 $P(Y=y_i)>0\quad (j\geqslant 1)$，则在 $Y=y_i$ 的条件下 X 的条件分布为

$$P(X=x_i|Y=y_j)=\frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{\sum_k p_{kj}}\quad (i=1,2,\cdots).$$

例：设随机变量 X 与 Y 相互独立，X 服从参数为 λ_1 的泊松分布，Y 服从参数为 λ_2 的泊松分布，试求在 $X+Y=n$ 条件下 X 的条件分布（n 为正整数）。

例：设随机变量 X 与 Y 相互独立，X 服从参数为 λ_1 的泊松分布，Y 服从参数为 λ_2 的泊松分布，试求在 $X+Y=n$ 条件下 X 的条件分布（n 为正整数）。

解：由于 $X+Y$ 服从参数为 $\lambda_1+\lambda_2$ 的泊松分布，故对 $k=0,1,\cdots,n$ 有

$$P(X=k|X+Y=n)=\frac{P(X=k,X+Y=n)}{P(X+Y=n)}$$

$$=\frac{P(X=k)P(Y=n-k)}{P(X+Y=n)}$$

$$=\frac{\lambda_1^k}{k!}e^{-\lambda_1}\frac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}\Big/\left[\frac{1}{n!}(\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)}\right]$$

$$=C_n^k\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}.$$

这表明，在 $X+Y=n$ 的条件下 X 的条件分布列为参数为 $n,\frac{\lambda_1}{\lambda_1+\lambda_2}$ 的二项分布。

例：设随机变量 X 与 Y 相互独立，都服从参数是 n，p 的二项分布，试求在 $X+Y=m(0\leqslant m\leqslant 2n)$ 条件下 X 的条件分布。

例：设随机变量 X 与 Y 相互独立，都服从参数是 n，p 的二项分布，试求在 $X+Y=m(0\leqslant m\leqslant 2n)$ 条件下 X 的条件分布。

解：记 $l=\min\{n,m\}$，易知

$$
\begin{aligned}
P(X+Y=m) &= \sum_{i=0}^{l} P(X=i, Y=m-i) \\
&= \sum_{i=0}^{l} P(X=i)P(Y=m-i) \\
&= \sum_{i=0}^{l} C_n^i p^i (1-p)^{n-i} C_m^{m-i} p^{m-i} (1-p)^{n-m+i} \\
&= p^m (1-p)^{2n-m} \sum_{i=0}^{l} C_n^i C_m^{m-i} \\
&= C_{2n}^m p^m (1-p)^{2n-m}.
\end{aligned}
$$

于是，当 $k=0,1,\cdots,l$ 时，

$$P(X=k|X+Y=m)=\frac{P(X=k,X+Y=m)}{P(X+Y=m)}$$
$$=\frac{C_n^k p^k(1-p)^{n-k}C_m^{m-k}p^{m-k}(1-p)^{n-m+k}}{C_{2n}^m p^m(1-p)^{2n-m}}$$
$$=\frac{C_n^k C_n^{m-k}}{C_{2n}^m}.$$

当 $k>l$ 时，显然 $P(X=k|X+Y=m)=0$。

由此可见，在 $X+Y=m$ 条件下 X 的条件分布是超几何分布。

例：一射手进行射击，击中目标的概率 $p \in (0,1)$，射击至击中目标两次为止。若以 X 表示首次击中目标所进行的射击次数，以 Y 表示总共进行的射击次数。试求 X 和 Y 的联合分布列及条件分布列。

例：一射手进行射击，击中目标的概率 $p \in (0,1)$，射击至击中目标两次为止。若以 X 表示首次击中目标所进行的射击次数，以 Y 表示总共进行的射击次数。试求 X 和 Y 的联合分布列及条件分布列。

解：$Y=n$ 表示第 n 次击中目标且前 $n-1$ 次恰有一次击中目标，

$$P(X=m,Y=n)=p^2(1-p)^{n-2}, \quad n=2,3,\cdots,m=1,\cdots,n-1.$$

从而

$$\begin{aligned}
P(X=m) &= \sum_{n=m+1}^{\infty} P(X=m,Y=n) \\
&= \sum_{n=m+1}^{\infty} p^2(1-p)^{n-2} \\
&= p(1-p)^{m-1}, \quad m=1,2,\cdots
\end{aligned}$$

且

$$
\begin{aligned}
P(Y=n) &= \sum_{m=1}^{n-1} P(X=m, Y=n) \\
&= \sum_{m=1}^{n-1} p^2(1-p)^{n-2} \\
&= (n-1)p^2(1-p)^{n-2}, \quad n=2,3,\cdots
\end{aligned}
$$

于是当 $n=2,3,\cdots$ 时，

$$P(X=m|Y=n) = \frac{p^2(1-p)^{n-2}}{(n-1)p^2(1-p)^{n-2}} = \frac{1}{n-1}, \quad m=1,\cdots,n-1.$$

当 $m=1,2,\cdots$ 时，

$$P(Y=n|X=m) = \frac{p^2(1-p)^{n-2}}{p(1-p)^{m-1}} = p(1-p)^{n-m-1}, \quad n=m+1,m+2,\cdots$$

连续型

设二维随机向量 (X,Y) 有联合分布函数 $F(x,y)$，联合密度 $p(x,y)$，若 $p_Y(y)>0$，则在 $Y=y$ 条件下 X 的条件分布函数为

$$F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{p(u,y)}{p_Y(y)}du.$$

自然，在 $Y=y$ 条件下 X 的条件分布密度为

$$p_{X|Y}(x|y)=\frac{p(x,y)}{p_Y(y)}.$$

连续场合的全概率公式：由基本公式

$$p(x,y)=p_Y(y)p(x|y)=p_X(x)p(y|x),$$

连续场合的全概率公式为：

$$p_X(x)=\int_{-\infty}^{+\infty}p(x,y)dy=\int_{-\infty}^{+\infty}p_Y(y)p(x|y)dy,$$

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\int_{-\infty}^{+\infty}p_X(x)p(y|x)dx.$$

连续场合的贝叶斯公式：

$$p(y|x)=\frac{p_Y(y)p(x|y)}{\int_{-\infty}^{+\infty}p_Y(y)p(x|y)dy},$$

$$p(x|y)=\frac{p_X(x)p(y|x)}{\int_{-\infty}^{+\infty}p_X(x)p(y|x)dx}.$$

例：设二维随机向量 (X, Y) 满足二维正态分布，易知 Y 的分布密度为

$$p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}.$$

则在 $Y = y$ 条件下 X 的分布密度为

$$\begin{aligned} p_{X|Y}(x|y) &= \frac{p(x,y)}{p_Y(y)} \\ &= \frac{1}{\sqrt{2\pi(1-\rho^2)}\sigma_1} \exp\left\{-\frac{(x-m)^2}{2(1-\rho^2)\sigma_1^2}\right\}, \end{aligned}$$

其中 $m = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)$.

例：设二维随机向量 (X, Y) 的联合密度为

$$p(x,y) = \begin{cases} \frac{1}{y}e^{-y} \cdot e^{-\frac{x}{y}}, & x > 0, y > 0, \\ 0, & \text{其他} \end{cases}$$

给定 $y > 0$，试求出条件概率 $P(X > 1|Y = y)$。

例：设二维随机向量 (X,Y) 的联合密度为

$$p(x,y)=\begin{cases}\frac{1}{y}e^{-y}\cdot e^{-\frac{x}{y}}, & x>0,y>0,\\ 0, & \text{其他}\end{cases}$$

给定 $y>0$，试求出条件概率 $P(X>1|Y=y)$。

解：在 $Y=y$ 条件下 X 的条件分布密度是

$$p_{X|Y}(x|y)=\frac{p(x,y)}{p_Y(y)}.$$

其中 $p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\int_0^{+\infty}\frac{1}{y}e^{-y}\cdot e^{-\frac{x}{y}}dx=e^{-x}$，于是

$$p_{X|Y}(x|y)=\begin{cases}\frac{1}{y}e^{-\frac{x}{y}}, & x>0,y>0\\ 0, & x\leqslant 0,y>0\end{cases}$$

因此

$$P(X>1|Y=y)=\int_1^{+\infty}\frac{1}{y}e^{-\frac{x}{y}}dx=e^{-\frac{1}{y}}.$$

例：设随机变量 X 在区间 $(0,1)$ 上随机取值，当观察到 $X=x(0<x<1)$ 时，随机变量 Y 在区间 $(x,1)$ 上随机取值，求 Y 的概率密度函数 $p_Y(y)$。

例：设随机变量 X 在区间 $(0,1)$ 上随机取值，当观察到 $X=x(0<x<1)$ 时，随机变量 Y 在区间 $(x,1)$ 上随机取值，求 Y 的概率密度函数 $p_Y(y)$。

解：X 服从区间 $(0,1)$ 上的均匀分布，对任意的 $x\in(0,1)$，在 $X=x$ 条件下，Y 的条件概率密度为

$$p(y|x)=\begin{cases}\frac{1}{1-x}, & y\in(x,1),\\ 0, & \text{其他}.\end{cases}$$

从而，

$$p(x,y)=p(y|x)p_X(x)=\begin{cases}\frac{1}{1-x}, & 0<x<y<1,\\ 0, & \text{其他}.\end{cases}$$

故

$$p_Y(y)=\int_{-\infty}^{+\infty}p(x,y)dx=\begin{cases}\int_0^y\frac{1}{1-x}dx=-\ln(1-y), & 0<y<1,\\ 0, & y\notin(0,1).\end{cases}$$

随机变量的独立性：设 X 和 Y 都是随机变量，如果对任何 $a < b, c < d$，事件 $\{a < X < b\}$ 和事件 $\{c < Y < d\}$ 相互独立，则称 X 与 Y 相互独立。

定理：设随机变量 X，Y 分别有分布密度 $p_X(x)$，$p_Y(y)$，则 X 与 Y 相互独立的充分必要条件是二元函数
$p(x, y) = p_X(x)p_Y(y)$ 是二维随机向量 (X, Y) 的联合密度。

证明：充分性：设 $p_X(x)p_Y(y)$ 是 (X,Y) 的联合密度，则对于任何 $a<b$，$c<d$ 有

$$P(a<X<b,c<Y<d)=\int_a^b\int_c^d p_X(x)p_Y(y)dxdy$$

$$=\int_a^b p_X(x)dx\cdot\int_c^d p_Y(y)dy=P(a<X<b)P(c<Y<d).$$

表明 X 与 Y 相互独立

必要性：设 X 与 Y 相互独立，则对任何 $a<b$，$c<d$ 有

$$P(a<X<b,c<Y<d)=P(a<X<b)P(c<Y<d)$$

$$=\int_a^b p_X(x)dx\cdot\int_c^d p_Y(y)dy=\int_a^b\int_c^d p_X(x)p_Y(y)dxdy$$

表明 $p_X(x)p_Y(Y)$ 是 (X,Y) 的联合密度。

设二维随机向量 (X, Y) 的联合密度 $p(x, y)$ 可以表示为

$$p(x, y) = f(x)g(y),$$

其中 $f(x) \geqslant 0$，$g(y) \geqslant 0$，且 $\int_{-\infty}^{+\infty} f(x)dx$ 收敛，则 X 与 Y 相互独立。

定理：设 $\xi = (X, Y)$ 是二维随机向量，X 的分布函数是 $F_X(x)$，Y 的分布函数是 $F_Y(y)$，则 X 和 Y 相互独立的充分必要条件是 ξ 的分布函数 $F(x, y)$ 等于 $F_X(x)$ 与 $F_Y(y)$ 之积，即

$$F(x, y) = F_X(x)F_Y(y). \tag{4}$$

定理：设 $\xi = (X, Y)$ 是二维随机向量，X 的分布函数是 $F_X(x)$，Y 的分布函数是 $F_Y(y)$，则 X 和 Y 相互独立的充分必要条件是 ξ 的分布函数 $F(x, y)$ 等于 $F_X(x)$ 与 $F_Y(y)$ 之积，即

$$F(x, y) = F_X(x)F_Y(y). \tag{4}$$

证明：必要性：设 X 与 Y 相互独立，则对任何 $n \geqslant 1$，事件 $\{-n < X \leqslant x\}$ 与事件 $\{-n < Y \leqslant x\}$ 相互独立，于是

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

令 $n \to \infty$，即知4式成立。

充分性：设4式成立，对任何 $a < b$，$c < d$，有

$$\begin{aligned} P(a < X \leqslant b, c < Y \leqslant d) &= F(b, d) - F(a, d) - F(b, c) + F(a, c) \\ &= F_X(b)F_Y(d) - F_X(a)F_Y(d) - F_X(b)F_Y(c) + \\ &= (F_X(b) - F_X(a))(F_Y(d) - F_Y(c)) \\ &= P(a < X \leqslant b)P(c < Y \leqslant d). \end{aligned}$$

定理：若随机变量 X 和 Y 相互独立，且方差 $D(X)$ 和 $D(Y)$ 存在，则 $D(X+Y)=D(X)+D(Y)$。
证明：根据方差定义：

$$
\begin{aligned}
& D(X+Y)=E(X+Y)^2-(E(X+Y))^2 \\
=&(E(X^2)+E(Y^2)-2E(XY))-((E(X))^2+(E(Y))^2+2E(X)E(Y)).
\end{aligned}
$$

由独立的性质知 $E(XY)=E(X)E(Y)$，则

$$
D(X+Y)=E(X^2)-(E(X))^2+E(Y^2)-(E(Y))^2=D(X)+D(Y).
$$

例：设二维随机向量 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 的二维正态分布，则 X 与 Y 相互独立的充分必要条件是 $\rho = 0$。

例：设二维随机向量 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 的二维正态分布，则 X 与 Y 相互独立的充分必要条件是 $\rho = 0$。

证明：已求出 X 和 Y 的分布密度：

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}.$$

$$p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}.$$

于是

$$p_X(x)p_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right\}.$$

结合联合密度 $p(x,y)$（式3），知当 $\rho = 0$ 时，

$$p(x,y) = p_X(x)p_Y(y).$$

故 X 与 Y 相互独立。反之，成立.

例：设 (X, Y) 联合密度为

$$p(x,y) = \begin{cases} 8xy, & 0 \leqslant x \leqslant y \leqslant 1, \\ 0, & \text{其他}. \end{cases}$$

问 X 与 Y 是否独立？

例：设 (X, Y) 联合密度为

$$p(x,y) = \begin{cases} 8xy, & 0 \leqslant x \leqslant y \leqslant 1, \\ 0, & \text{其他}. \end{cases}$$

问 X 与 Y 是否独立？

解：易得

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)dy = \begin{cases} \int_x^1 8xydy = 4x(1-x^2), & x \in [0,1], \\ 0, & \text{其他}. \end{cases}$$

$$p_Y(y) = \int_{-\infty}^{+\infty} p(x,y)dx = \begin{cases} \int_0^y 8xydx = 4y^3, & y \in [0,1], \\ 0, & \text{其他}. \end{cases}$$

从而，

$$p(x,y) \neq p_X(x)p_Y(y) = \begin{cases} 16x(1-x^2)y^3, & x,y \in [0,1], \\ 0, & \text{其他}. \end{cases}$$

故 X，Y 不独立。

例：假定一天内进入邮局的人数为服从参数 λ 的泊松分布的随机变量，如果每个进入邮局的人为男性的概率为 p，为女性的概率为 $1-p$，证明进入邮局的男人数和女人数是相互独立的泊松随机变量，且参数分别为 λp 和 $\lambda(1-p)$。

例：假定一天内进入邮局的人数为服从参数 λ 的泊松分布的随机变量，如果每个进入邮局的人为男性的概率为 p，为女性的概率为 $1-p$，证明进入邮局的男人数和女人数是相互独立的泊松随机变量，且参数分别为 λp 和 $\lambda(1-p)$。

解：设 X 和 Y 分别是进入邮局的男人数和女人数，则对任意的自然数 i 和 j，

$$P(X=i,Y=j)=P(X=i,Y=j|X+Y=i+j)P(X+Y=i+j).$$

注意到

$$P(X+Y=i+j)=e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}.$$

且在给定 $i+j$ 人进入邮局的条件下，恰有 i 个男人和 j 个女人的概率是 $C_{i+j}^{i}p^{i}(1-p)^{j}$，从而

$$P(X=i,Y=j)=C_{i+j}^{i}p^{i}(1-p)^{j}e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}$$

故 $i \in \mathbb{N}$

$$P(X=i)=e^{-\lambda p}\frac{(\lambda p)^i}{i!}\sum_{j=0}^{+\infty}e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}=e^{-\lambda p}\frac{(\lambda p)^i}{i!}.$$

且 $j \in \mathbb{N}$

$$P(X=j)=e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}\sum_{i=0}^{+\infty}e^{-\lambda p}\frac{(\lambda p)^i}{i!}=e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}.$$