Independence of Random Variables

November 11, 2024

随机变量的独立性

回顾: 对于随机事件的独立性有 $P(AB) = P(A)P(B)$. 而对于多维随机变量，各个分量的取值也可能是互不影响的, 这时称它们的相互独立的, 具体定义如下.

定义 0.44 设 X, Y 为二维随机变量, 对任意实数 x, y, 若事件 $X \leq x$ 和 $Y \leq y$ 相互独立, 即 $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$, 等价于

$$F(x, y) = F_X(x)F_Y(y)$$

则称随机变量 X 与 Y 相互独立.

Remarks:

- 随机变量 X 与 Y 相互独立等价于随机事件 $\{X \leq x\}$ 和 $\{Y \leq y\}$ 对任意实数 x, y 都相互独立
- 对于任意的常数 c 也与任意随机变量相互独立

离散型随机变量的独立性

定义 0.45 设二维离散型随机向量 (X,Y) 的分布列为

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

以及 X 和 Y 的边缘分布列为 $p_{i\cdot}$ 和 $p_{\cdot j}$.

则 X 与 Y 相互独立的充要条件是

$$p_{ij} = p_{i\cdot}\, p_{\cdot j} \ .$$

简而言之，分布列独立

- 某点处的概率等于所在“行”边缘分布列和“列”边缘分布列的乘积。

离散型随机变量的独立性：例 0.100

例 0.100 设离散型随机变量 X 和 Y 相互独立且它们的取值均为 $\{1, 2, 3\}$. 已知

- $P(Y = 1) = 1/3$
- $P(X = 1, Y = 1) = P(X = 2, Y = 1) = 1/8$
- $P(X = 1, Y = 3) = 1/16$

求 X 和 Y 的联合和边缘分布列.

解答:

- 由边缘分布列的定义有

  $$P(X = 3, Y = 1) = P(Y = 1) - P(X = 1, Y = 1) - P(X = 2, Y = 1) = 1/12$$

- 由离散型随机向量的独立性的定义，可得 $P(X = 1) = P(X = 2) =$

$3/8$ 和 $P(X=3)=1/4$, 同理计算其他概率后可得分布列为

$X \backslash Y$	1	2	3	$p_{i\cdot}$
1	$1/8$	$3/16$	$1/16$	$3/8$
2	$1/8$	$3/16$	$1/16$	$3/8$
3	$1/12$	$1/8$	$1/24$	$1/4$
$p_{\cdot j}$	$1/3$	$1/2$	$1/6$	

离散型随机变量的独立性：例 0.101

例 0.101 设离散型随机变量 X 和 Y 相互独立且它们的取值均为 $\{-1,0,1\}$. 已知

- $P(XY=0)=3/4$, $P(XY=1)=1/8$
- $P(X=0,Y=0)=1/4$
- $P(X=-1)=P(X=1)=1/4$

求 $P(X=-1,Y=1)$ 和 $P(Y=0)$ 的值.

离散型随机变量的独立性：例 0.101

题干: 设离散型随机变量 X 和 Y 相互独立且它们的取值均为 $\{-1, 0, 1\}$. 已知

- $P(XY = 0) = 3/4, P(XY = 1) = 1/8$
- $P(X = 0, Y = 0) = 1/4$
- $P(X = -1) = P(X = 1) = 1/4$

求 $P(X = -1, Y = 1)$ 和 $P(Y = 0)$ 的值.

解答:

- $P(X = -1) = 1/4, P(X = 0) = 1/2, P(X = 1) = 1/4$
- $P(Y = -1) = 1/3, P(Y = 0) = 1/2, P(Y = 1) = 1/6$

连续型随机变量的独立性

定义 0.46 设二维连续型随机向量 (X,Y) 的密度函数为 $f(x,y)$, 以及 X 和 Y 的边缘密度函数分别为 $f_X(x)$ 和 $f_Y(y)$, 则 X 与 Y 相互独立的充要条件是

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

Remarks:

- 注意: 在离散型的独立性定义中, 我们关心的是分布函数

$$F(x,y) = F_X(x)F_Y(y)$$

- 在连续型中, 积分独立

$$f(x,y) = f_X(x)f_Y(y) \quad \Leftrightarrow \quad F(x,y) = F_X(x)F_Y(y)$$

连续型随机变量独立性的性质

- 若随机变量 X 和 Y 相互独立, 则对任意集合 $A, B \subseteq \mathbb{R}$, 事件 $\{X \in A\}$ 和事件 $\{Y \in B\}$ 相互独立.
- 若随机变量 X 和 Y 相互独立, 则 $f(X)$ 与 $g(Y)$ 相互独立 (其中 $f(x)$ 和 $g(y)$ 是连续或分段连续函数).
- 若存在两个 $h(x)$ 和 $g(y)$ 函数, 使得 X 和 Y 的联合密度函数 $f(x, y)$ 对任意实数 x, y 均有

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

则随机变量 X 和 Y 相互独立. (注意: 定义域、存在性)

正态分布的独立性

定理 0.16 设二维随机向量 $(X,Y)\sim\mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则 X 和 Y 相互独立的充要条件是 $\rho=0$.

$$
\begin{aligned}
f(x,y) &= \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\left[\frac{(x-\mu_x)^2}{\sigma_x^2}+\frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right)\\
&= \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{(x-\mu_x)^2}{\sigma_x^2}\right)\cdot\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left(-\frac{(y-\mu_y)^2}{\sigma_y^2}\right)\\
&= f_X(x)\cdot f_Y(y)
\end{aligned}
$$

换言之，当 $\rho\neq 0$ 时，$f(x,y)\neq f_X(x)\cdot f_Y(y)$，进而 X 和 Y 不独立。

连续随机变量独立性：例 0.102

例 0.102 设二维随机向量 (X,Y) 的密度函数

$$f(x,y)=\begin{cases} cxe^{-y}, & 0<x<y<+\infty \\ 0, & \text{其它} \end{cases}$$

问随机变量 X 和 Y 是否相互独立?

解答：例 0.102

题目: 如上所述.

解答:

- 由密度函数的规范性求解常数 c 的值为 1.
- 利用分部积分可得, 当 $x > 0$ 时随机变量 X 的边缘密度为

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

  当 $y > 0$ 时随机变量 Y 的边缘密度为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_{y}^{0} xe^{-y}dx = \frac{1}{2}y^2e^{-y}$$

  因为 $f(x,y) \neq f_X(x)f_Y(y)$ 可得随机变量 X 和 Y 不独立.

连续随机变量独立性：例 0.103

例 0.103 设随机变量 X 和 Y 相互独立, 且 X 服从 $[-1, 1]$ 均匀分布, Y 服从参数为 $\lambda = 2$ 的指数分布, 求 $P(X + Y \leq 1)$.

解答：例 0.103

题目: 设随机变量 X 和 Y 相互独立, 且 X 服从 $[-1,1]$ 均匀分布, Y 服从参数为 $\lambda=2$ 的指数分布, 求 $P(X+Y\leq 1)$.

解答:

- 列出 X 和 Y 的边缘密度函数

$$f_X(x)=\begin{cases}\frac{1}{2}, & x\in[-1,1]\\ 0, & \text{其它}\end{cases} \qquad \text{和} \qquad f_Y(y)=\begin{cases}2e^{-2y}, & y\geq 0\\ 0, & \text{其它}\end{cases}$$

- 利用独立性求 X 和 Y 的联合密度函数

$$f(x,y)=\begin{cases}e^{-2y}, & -1\leq x\leq 1, y\geq 0\\ 0, & \text{其它}\end{cases}$$

- 利用联合密度函数有

$$P(X+Y\leq 1)=\int_{-1}^{1}\int_{0}^{1-x}e^{-2y}dy=\frac{3}{4}+\frac{1}{4}e^{-4}$$

Conditional Distribution

November 11, 2024

随机变量的条件分布

回顾: 随机事件的条件概率, 即在事件 B 发生的条件下事件 A 发生的条件概率

$$P(A \mid B) = \frac{P(AB)}{P(B)}\,.$$

推广到随机变量: 给定随机变量 Y 取值的条件下, 求随机变量 X 的概率分布, 即条件分布.

离散型随机变量的条件分布

定义 0.47 设二维离散型随机变量 (X,Y) 的分布列为

$$p_{ij} = P(X = x_i, Y = y_j) \quad (i, j \in \mathbb{N}^+) ,$$

若 Y 的边缘分布 $P(Y = y_j) = p_{\cdot j} > 0$, 则称

$$P(X = x_i \mid Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}$$

在 $Y = y_j$ 条件下随机变量 X 的条件分布列 (conditional probability distribution). 类似定义在 $X = x_i$ 条件下随机变量 Y 的条件分布列.

Remarks:

- 若出现条件概率 $P(X = x_i|Y = y_j)$, 一般都默认 $P(Y = y_j) > 0$.

离散型随机变量的条件分布列

条件分布列也可以通过下面的表格给出:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
$P(X=x_i \vert Y=y_j)$	$p_{1j}/p_{\cdot j}$	$p_{2j}/p_{\cdot j}$	$\cdots$	$p_{nj}/p_{\cdot j}$	$\cdots$

$$P(X = x_i \mid Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}$$

注意：该表格呈现的分布列是关于 X 取值展开的。

离散型随机变量条件分布列的性质

我们注意到条件概率是一种概率，而变量的条件分布列也是一种分布列，进而具有如下性质：

- 非负性: $P(X = x_i|Y = y_j) \geq 0$.
- 规范性: $\sum_{i=1}^{\infty} P(X = x_i|Y = y_j) = 1$.
- 若离散随机变量 X 和 Y 相互独立, 则有

$$P(X = x_i \mid Y = y_j) = P(X = x_i) \qquad P(Y = y_j \mid X = x_i) = P(Y = y_j)$$

离散型随机变量的条件分布：例 0.104

例 0.104 一选手随机进行射击训练, 击中目标的概率为 p, 射击进行到击中两次目标为止, 用 X 表示首次击中目标的射击次数, 用 Y 表示第二次射中目标的射击次数, 求 X 和 Y 的联合分布和条件分布.

解答：例 0.104

题目: 如上所述.

解答:

- $X=m$ 表示首次击中目标射击了 m 次, $Y=n$ 表示第二次击中目标射击了 n 次, X 和 Y 的联合分布列为

$$P(X=m,Y=n)=f(x,y)=\begin{cases} p^2(1-p)^{n-2} & 1\leq m<n<\infty \\ 0, & \text{其它} \end{cases}$$

- 由离散型随机向量的边缘分布列的定义可得随机变量 X 和 Y 的边缘分布列分别为

$$P\{X=m\}=\sum_{n=m+1}^{\infty}P\{X=m,Y=n\}=\sum_{n=m+1}^{\infty}p^2(1-p)^{n-2}=p(1-p)^{m-1}$$

$$P\{Y=n\}=\sum_{m=1}^{n-1}P\{X=m,Y=n\}=\sum_{m=1}^{n-1}p^2(1-p)^{n-2}=(n-1)p^2(1-p)^{n-2}(n\geq 2)$$

- 当 $n \geq 2$ 时, 随机变量 X 在 $Y=n$ 的条件下的分布列

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

当 $m \geq 1$ 时, 随机变量 Y 在 $X=m$ 的条件下的分布列

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

连续型随机变量的条件概率

可以注意到：对连续随机变量 (X,Y), 对任意 $x,y\in(-\infty,+\infty)$, 有

$$P(X=x)=0 \text{ 和 } P(Y=y)=0 \text{ 成立} \quad \text{(端点处的概率为 0)}$$

因此, 不能用离散随机变量的条件概率推导连续随机变量的条件分布.

连续型随机变量的条件概率

这里可以利用极限方式来考虑. 当 $P(y \le Y \le y+\epsilon) > 0$ 时, 利用积分中值定理来求解条件分布函数

$$
\begin{aligned}
F_{X|Y}(x \mid y) &= \lim_{\epsilon \to 0^+} P(X \le x \mid y \le Y \le y+\epsilon) \\
&= \lim_{\epsilon \to 0^+} \frac{P(X \le x, y \le Y \le y+\epsilon)}{P(y \le Y \le y+\epsilon)} \\
&= \lim_{\epsilon \to 0^+} \frac{\int_{-\infty}^{x} \int_{y}^{y+\epsilon} f(u,v)\,\mathrm{d}u\,\mathrm{d}v}{\int_{y}^{y+\epsilon} f_Y(u)\,\mathrm{d}v} \\
&= \lim_{\epsilon \to 0^+} \frac{\epsilon \int_{-\infty}^{x} f(u, y+\theta_1\epsilon)\,\mathrm{d}u}{\epsilon f_Y(y+\theta_2\epsilon)} \quad \text{积分中值定理}, \theta_1, \theta_2 \in (0,1) \\
&= \frac{\int_{-\infty}^{x} f(u,y)\,\mathrm{d}u}{f_Y(y)} = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)}\,\mathrm{d}u
\end{aligned}
$$

进一步得到条件密度函数 $f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)}$.

连续型随机变量的条件概率

定义 0.48 设连续随机变量 (X,Y) 的联合概率密度为 $f(x,y)$, 以及 Y 的边缘概率密度为 $f_Y(y)$, 对任意 $f_Y(y)>0$, 称

$$f_{X|Y}(x \mid y)=\frac{f(x,y)}{f_Y(y)}$$

在 $Y=y$ 条件下随机变量 X 的条件概率密度.

$$F_{X|Y}(x \mid y)=P[X \leq x \mid Y=y]=\int_{-\infty}^{x} f_{X|Y}(u \mid y)\,\mathrm{d}u$$

在 $Y=y$ 条件下 X 的条件分布函数.

Remarks: 射线/直线的比率

条件密度函数的性质

条件密度函数本质上是密度函数, 具有以下性质:

- 非负性: 对任意实数 x, y 有 $f_{X|Y}(x \mid y) \geq 0$
- 规范性: 对任意实数 $y : f_Y(y) > 0$ 有

$$\int_{-\infty}^{+\infty} f_{X|Y}(x \mid y) \, \mathrm{d}x = \int_{-\infty}^{+\infty} \frac{f(x, y)}{f_Y(y)} \, \mathrm{d}x = \frac{1}{f_Y(y)} \int_{-\infty}^{+\infty} f(x, y) \, \mathrm{d}x = 1$$

- 乘法公式: $f(x, y) = f_X(x) f_{Y|X}(y \mid x) = f_Y(y) f_{X|Y}(x \mid y)$
- 密度函数的贝叶斯公式. 设随机变量 X 和 Y 相互独立, 则有

$$f_{Y|X}(y \mid x) = f_Y(y) \qquad \text{和} \qquad f_{X|Y}(x \mid y) = f_X(x)$$

根据条件概率公式有

$$f(x \mid y) = \frac{f(x, y)}{f_Y(y)} = \frac{f_{Y|X}(y \mid x) f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y \mid x) f_X(x)}{\int_{-\infty}^{+\infty} f(x, y) \, \mathrm{d}x} = \frac{f_{Y|X}(y \mid x) f_X(x)}{\int_{-\infty}^{+\infty} f(y \mid x) f_X(x) \, \mathrm{d}x}$$

连续型随机变量的条件概率：例 0.105

例 0.105 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases}\dfrac{e^{-\frac{x}{y}}e^{-y}}{y}, & 0<x<+\infty, 0<y<+\infty\\ 0, & \text{其它}\end{cases}$$

求 $P(X>1 \mid Y=y)$.

解答：例 0.105

题目: 如上所述.

解答:

- 根据题意可得随机变量 Y 的边缘分布为

$$f_Y(y)=\int_0^{+\infty}\frac{e^{-\frac{x}{y}}e^{-y}}{y}\,\mathrm{d}x=e^{-y}\cdot -e^{-\frac{x}{y}}\ |_0^{+\infty}=e^{-y}\qquad (y>0)$$

- 进一步可得在 $Y=y$ 条件下随机变量 X 的条件概率密度为

$$f_{X|Y}(x\mid y)=\frac{e^{-\frac{x}{y}}}{y}$$

- 最后求解可得

$$P(X>1\mid Y=y)=\int_1^{+\infty}\frac{e^{-\frac{x}{y}}}{y}\,\mathrm{d}x=-e^{-\frac{x}{y}}\ |_1^{+\infty}=e^{-\frac{1}{y}}$$

连续型随机变量的条件概率：例 0.106

例 0.106 已知随机变量 $X \sim U(0,1)$, 当观察到 $X = x$ 的条件下, 随机变量 $Y \sim U(x,1)$, 求 Y 的概率密度.

解答：例 0.106

题目: 如上所述.

解答:

- 根据题意可得 $X=x$ 的条件下随机变量 Y 的条件概率密度为

$$f_{Y|X}(y \mid x)=\frac{1}{1-x} \qquad x>0$$

- 根据条件概率公式有

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

- 最后求解可得, 当 $y>0$ 时 Y 的边缘分布为

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x=\int_0^y \frac{1}{1-x}\,\mathrm{d}x=-\ln(1-y)$$

二维正态分布的条件分布

定理 0.17 设二维随机变量

$$(X, Y) \sim \mathcal{N}\left(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho\right) \ ,$$

则

- 在 $Y = y$ 条件下随机变量 X 的服从正态分布

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

- 在 $X = x$ 条件下随机变量 Y 的服从正态分布

$$\mathcal{N}\left(\mu_y + \rho(x - \mu_x)\frac{\sigma_y}{\sigma_x}, (1 - \rho^2)\sigma_y^2\right) \ .$$

该证明留作思考题.

二维正态分布的条件分布

提示:

$$\frac{x-\mu_x}{\sigma_x} = \rho\,\frac{y-\mu_y}{\sigma_y}$$

总结

- 独立性：要求分布函数独立 (定义法)
  - 离散型：分布列独立
  - 连续型：密度函数独立
    - 二维正态分布的独立性条件：$\rho = 0$
- 条件分布：一种特殊的分布
  - 离散型：相对于 $P(Y = y_j)$ 的关于 X 的分布列
  - 连续型：条件密度函数
    - 密度函数的 Bayes 公式
    - 二维正态分布的条件分布

如何构造二维 (连续型) 随机变量的联合分布函数?

为了求解二维连续型随机变量的概率, 需求解其联合分布函数, 进而需求解其联合密度函数 (本次课提供了两个可以构造多变量联合分布函数的方法)

- 根据实际问题或实际数据归纳为

$$f(x,y)$$

- 根据随机变量的独立性有

$$f(x,y) = f_X(x)f_Y(y)$$

- 根据条件分布的乘法公式

$$f(x,y) = f_X(x)f_{Y|X}(y \mid x)$$