Basics of Probs. and Stats.

2. 随机变量与概率分布

梁雪吟

地球科学与资源学院
liangxueyin@email.cugb.edu.cn

April 2, 2023

Outline

2.1 基本概念

随机变量的含义

- 我们将随机试验的所有可能结果都转为“数值”(例如: 扔硬币正面 $=1$, 反面 $= 0$)
  这样, 我们就可以用随机变量 Y 来表示某次随机试验的可能结果及其发生概率(将随机事件数量化达到简化的目的)。
- 随机变量的两个组成要素:
  1. 随机变量的所有可能的取值
  2. 每一个可能取值对应的概率

累积分布函数 Cumulative Probability Distribution

- 如何描述随机变量取值在某一区间上的概率？
- $F(x) = P(X \le x)$

  随机变量 X 的 CDF 描述的是随机变量的取值落入“区间” $(-\infty, x)$ 的概率
- 特点：
  1. 单调不减
  2. 右连续
  3. $\lim_{x \to -\infty} F(x) = 0$

     $\lim_{x \to +\infty} F(x) = 1$

离散型随机变量的累积分布函数

Example

随机抽取一名中国居民，她/他的受教育程度可视作随机变量 Y 假设只存在 4 种可能的教育程度：文盲 $(Y=1)$、小学 $(Y=2)$、中学 $(Y=3)$、大学 $(Y=4)$ 而且已知这 4 个取值对应的概率： $P(Y=y_i)$

Y	Y=1	Y=2	Y=3	Y=4
$P(Y=y_i)$	10%	40%	40%	10%

那么，随机变量 Y 的累积分布函数就可以通过“累积加总”的方式计算出来：

离散型随机变量的累积分布函数

$$F(0) \equiv \mathrm{P}(Y \leq 0) = 0$$

$$F(1) \equiv \mathrm{P}(Y \leq 1) = 0.1$$

$$F(2) \equiv \mathrm{P}(Y \leq 2) = 0.1 + 0.4 = 0.5$$

$$F(3) \equiv \mathrm{P}(Y \leq 3) = 0.1 + 0.4 + 0.4 = 0.9$$

$$F(4) \equiv \mathrm{P}(Y \leq 4) = 0.1 + 0.4 + 0.4 + 0.1 = 1$$

$$F(5) \equiv \mathrm{P}(Y \leq 5) = 1$$

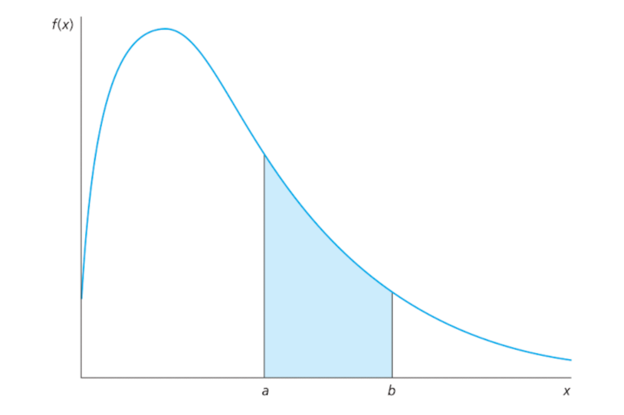
- $F(y) = F(Y \leq y_i)$ 描述的是随机变量 Y 的取值落入“区间” $(-\infty, y)$ 的概率

连续性随机变量的分布函数

- 连续型随机变量 X 的累积分布函数 (CDF)，是它的概率密度函数 (PDF) 的积分 (概率密度函数 f (x) 从负无穷开始积分，一直积到给定取值 x 为止)
- $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$
- $F(x)$ 描述的是随机变量的取值落入“区间” $(-\infty, x)$ 的概率

连续性随机变量的分布函数

FIGURE B.2 The probability that X lies between the points a and b.



概率密度函数 Probability Density Function

- 离散型随机变量的概率密度函数 $f_X(x) = P(X = x)$
- X 为某个中国公民的受教育程度，$f_X(x) = P(X = x)$ 可以描述此人受教育程度为文盲、小学、中学、大学的概率
- 连续型随机变量的概率密度函数 $f_X(x) = \dfrac{dF_X(x)}{dx}$
- 离散型随机变量 X 的概率密度函数，衡量的是随机变量 X 取值落在 x 某“邻域” $\left(x - \frac{\varepsilon}{2}, x + \frac{\varepsilon}{2}\right)$ 的概率大小
- 收入可以视作连续性随机变量:X 为某个中国公民的月收入，$f_X(x) = \dfrac{dF_X(x)}{dx}$ 衡量的是收入水平落在 x（如 2000 元附近）范围内的概率大小。

PDF 和 CDF 的应用

- 求随机变量 X 落入某个取值区间 $(a < x < b)$ 的概率
- 连续型：

$$\begin{aligned} F(a) &= P\{X < a\} = P\{X \leq a\} \\ F(b) &= P\{X < b\} = P\{X \leq b\} \\ F(b) - F(a) &= P\{a < X < b\} = P\{a \leq X \leq b\} \\ &= P\{a < X \leq b\} = P\{a \leq X < b\} \end{aligned} \tag{1}$$

- 连续型随机变量 X 取单点值的概率为零 $P(X = x) = 0$ ，但是这并不意味着 X 的取值就不可能恰好落在一个单点上面，而只是我们的测量工具测不出来.

PDF 和 CDF 的应用

- 连续型：

$$P\{a < X < b\} = F(b) - F(a) = \int_a^b f(x)\,dx \tag{2}$$

- 由积分与求导的关系，连续型随机变量的累积分布函数是概率密度函数的“原函数”。

例题

一汽车沿一街道行驶，需要经过三个设有红绿信号灯的路口，若设每个信号灯显示红绿两种信号的时间相等，且各个信号灯工作相互独立，以乘表示该汽车首次遇到红灯前已经通过的路口数，试求 X 的概率分布列

例题

一汽车沿一街道行驶，需要经过三个设有红绿信号灯的路口，若设每个信号灯显示红绿两种信号的时间相等，且各个信号灯工作相互独立，以乘表示该汽车首次遇到红灯前已经通过的路口数，试求 X 的概率分布列

解：

已知 X 的取值可以是 0，1，2，3，令 $A_i=$ 在第 i 个路口遇到红灯，则 A_1, A_2, A_3 相互独立，有:

$$\begin{aligned} P(X=0) &= P\left(A_1\right)=\frac{1}{2}, \\ P(X=1) &= P\left(\bar{A}_1 A_2\right)=P\left(\bar{A}_1\right) P\left(A_2\right)=\frac{1}{2} \times \frac{1}{2}=\frac{1}{4}, \\ P(X=2) &= P\left(\bar{A}_1 \bar{A}_2 A_3\right)=P\left(\bar{A}_1\right) P\left(\bar{A}_2\right) P\left(A_3\right)=\frac{1}{8} \\ P(X=3) &= P(\bar{A}_1 \bar{A}_2 \bar{A}_3)=\frac{1}{8} \end{aligned} \tag{3}$$

例题

设 $F_1(x)$ 和 $F_2(x)$ 为两个分布函数，其对应的概率密度为 $f_1(x)$ 和 $f_2(x)$，则必为概率密度函数的是：

A. $f_1(x)f_2(x)$

B. $2f_1(x)f_2(x)$

C. $f_1(x)F_2(x)$

D.$f_1(x)F_2(x)+f_2(x)F_1(x)$

例题

设 $F_1(x)$ 和 $F_2(x)$ 为两个分布函数，其对应的概率密度为 $f_1(x)$ 和 $f_2(x)$，则必为概率密度函数的是：

A. $f_1(x)f_2(x)$

B. $2f_1(x)f_2(x)$

C. $f_1(x)F_2(x)$

D.$f_1(x)F_2(x)+f_2(x)F_1(x)$

解

D（PDF 和 CDP 求导积分的关系）

2.2 常见离散型和连续性分布

离散型分布

- 二项分布：$X \sim B(n, p)$ （泊松定理）
- 超几何分布：$X \sim H(n, M, N)$
- 泊松分布：$X \sim P(\lambda)$
- 几何分布：$X \sim G(P)$

几何分布——例题

设某篮球选手命中率为 $\frac{3}{4}$，他连续投篮，直到第一次命中为止，求他投篮次数不超过 5 次就能命中的概率。

几何分布——例题

设某篮球选手命中率为 $\frac{3}{4}$，他连续投篮，直到第一次命中为止，求他投篮次数不超过 5 次就能命中的概率。

解

设首次命中所用的投篮次数为 X，则 X 服从几何分布，由已知可得：

$P(X=k)=(\frac{1}{4})^{k-1}\times\frac{3}{4}, k=1,2,\ldots$

所以有 $P(X\leq 5)=\sum_{k=1}^{5}P(X=k)=\frac{3}{4}\times\frac{1-(\frac{1}{4})^5}{1-\frac{1}{4}}=\frac{1023}{1024}$

泊松分布——例题

设某生产线生产产品的次品率为 0.001，若随机抽查 5000 件产品，求至少发现 2 件次品的概率？

泊松分布——例题

设某生产线生产产品的次品率为 0.001，若随机抽查 5000 件产品，求至少发现 2 件次品的概率？

解

生产线生产的产品的试验中可认为事件是相互独立的，于是本题就是一个独立重复的伯努利实验。设 X 表示 5000 件产品中的次品数，则 $X \sim B(5000, 0.001)$。我们发现直接利用二项分布计算比较繁琐，由于 n 很大，p 很小，所以可以用泊松分布近似，$np = \lambda = 5$：

$$P(X \geq 2) = 1 - P(X = 1) - P(X = 0) \approx 1 - e^{-5} - 5e^{-5} \approx 0.9596$$

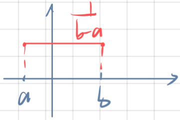
连续型分布

- **均匀分布**： $X \sim U[a, b]$
- **指数分布**: $X \sim e[\lambda]$
- **正态分布**: $X \sim N(\sigma, \mu^2)$

均匀分布

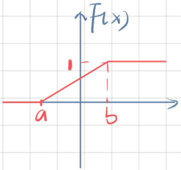
若随机变量的概率密度为:

$$f(x)=\begin{cases}\frac{1}{b-a} & x\in[a,b]\\ 0 & \text{otherwise}\end{cases}$$



则X服从均匀分布

$X\sim U[a,b]$



$$F(x)=\begin{cases}0 & x<a\\ \frac{x-a}{b-a} & x\in[a,b]\\ 1 & x>b\end{cases}$$

Figure: 均匀分布

均匀分布——例题

已知某地铁站每隔 5 分钟有一辆地铁到达，一位乘客去此地铁站乘车，等候时间超过 3 分钟的概率。

均匀分布——例题

已知某地铁站每隔 5 分钟有一辆地铁到达，一位乘客去此地铁站乘车，等候时间超过 3 分钟的概率。

解

令 X 为到站的时刻，可认为 $X \sim U[0,5]$

所以 $P(0 < X < 2) = 0.4$

均匀分布——例题

已知某地铁站每隔 5 分钟有一辆地铁到达，一位乘客每周要乘坐 10 次该地铁，求该乘客每周有 2 次以上等候时间超过 3 分钟的概率.

均匀分布——例题

已知某地铁站每隔 5 分钟有一辆地铁到达，一位乘客每周要乘坐 10 次该地铁，求该乘客每周有 2 次以上等候时间超过 3 分钟的概率。

解

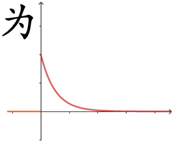
令 Y 为 10 次中等候超 3 分钟的次数，易知 $Y \sim b(n,p)$，其中 $n=10$，$p=0.4$

所以 $P(Y \geq 2) = 1 - 0.6^{10} - C_{10}^{1} 0.4^{1} 0.6^{9} = 4.6\%$

指数分布

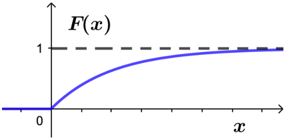
若随机变量X概率密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0,\\ 0, & x<0,\end{cases}$$



则称X服从参数为λ指数分布,记作

$$X\sim e(\lambda).$$



$$F(x)=\begin{cases}0, & x<0,\\ 1-e^{-\lambda x}, & x\geq 0.\end{cases}$$

Figure: 指数分布

常用来描述等待某事件发生的！等待时间 (或某东西的寿命)。**无记忆性**！

指数分布——例题

设变量 X 服从参数为 1 的指数分布，对 $a > 0$，有

$P(X \le a + 1 | X > a) =$

指数分布——例题

设变量 X 服从参数为 1 的指数分布，对 $a > 0$，有

$P(X \leq a+1|X > a) =$

解

由于 $X \sim e(\lambda)$，且指数分布具有无意义性：

$P(X \leq a+1|X > a) = 1 - P(X > a+1|X > a) = 1 - P(X > 1) = 1 - e^{-1}$

正态分布

其概率密度为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty,$$

其中 $\sigma>0,-\infty<\mu<+\infty$，$\varphi(x)$ 关于 $x=\mu$ 对称.

当 $\sigma=1,\mu=0$ 时，称 X 服从标准正态分布，记为 $N(0,1)$，其分布函数为

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt.$$

$\Phi(x)$ 有如下性质：$\Phi(-x)=1-\Phi(x)$，$\Phi(0)=\frac{1}{2}$，当 $a>0$ 时，$P\{|x|\leqslant a\}=2\Phi(a)-1$.

正态分布 $N(\mu,\sigma^2)$ 的标准化：设 $X\sim N(\mu,\sigma^2)$，则 $\frac{X-\mu}{\sigma}\sim N(0,1)$，故 X 的分布函数为

$F(x)=\Phi\left(\frac{X-\mu}{\sigma}\right)$，且

$$P\{a<X\leqslant b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

正态分布是概率统计中最常用的分布，有着广泛的应用.

正态分布是概率统计中，最常用最重要的分布，常用来描述量测误差或随机噪声.

例题——确定分布中的参数

设连续型随机变量 X 的分布函数为:

$$F(x) = \begin{cases} A + Be^{-\frac{x^2}{2}}, & x > 0, \\ 0, & x \leqslant 0. \end{cases} \tag{4}$$

求:(1)常数 A 和 B;

(2)X 的概率密度函数 $f(x)$;

(3)$P(-\sqrt{2} < X < 2)$

例题——确定分布中的参数

解 (1)

确定参数，用到的是分布函数的性质，由于 $F(X)$ 有右连续性，所以：

$F(0+) = \lim\limits_{x \to 0^+} F(X) = A + B = F(0) = 0$

$F(+\infty) = \lim\limits_{x \to +\infty} F(X) = A = 1$

所以 $A = 1$，$B = -1$

解 (2)

$$f(x) = F^{'}(x) = \begin{cases} xe^{-\frac{x^2}{2}}, & x > 0, \\ 0, & x \leqslant 0. \end{cases} \tag{5}$$

解 (3)

$P(-\sqrt{2} < X < 2) = F(2) - F(-\sqrt{2}) = 1 - e^{-2}$

2.2 二维随机变量及其分布

二维随机变量及其分布

- 有的事件需多个随机变量来描述，所以引入多维随机变量。
- 若 X, Y 为同一样本空间 Ω 上的两个随机变量，则由它们构成的有序组 (X, Y) 称为一个二维随机变量。
- e.g 扔两次骰子第一次和第二次的点数 (X, Y)。
- 仍是利用分布求概率（离散型、连续型）。

二维随机变量的分布函数

$$F(x,y)=P(X\leq x,Y\leq y)\quad \forall x,y\in\mathbb{R} \tag{6}$$

也称之为联合分布函数。仍然具有单调不减，右连续的特点。

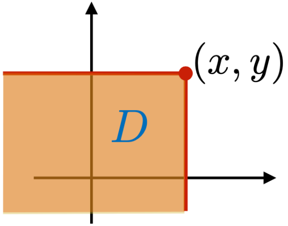
$$(1)\ 0\leq F(x,y)\leq 1;$$

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

$$\lim_{x\to+\infty}F(x,y)=F_Y(y)$$

$$\lim_{y\to+\infty}F(x,y)=F_X(x)$$

$$\lim_{x,y\to+\infty}F(x,y)=1$$



(X,Y)的边缘分布

书 P53 分布函数的性质（2）可用于求参数

二维离散型随机变量的分布

- 联合概率分布: $P(X=x_i, Y=y_i)=p_{ij}$
  其中 $p_{ij}\geq 0$, $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1$
- 边缘分布:
  1. 关于 X: $P(X=x_i)=\sum_{i=1}^{\infty}p_{ij}=p_i$
  2. 关于 Y: $P(Y=y_i)=\sum_{j=1}^{\infty}p_{ij}=p_j$
- 条件分布:
  1. 关于 X: $P(X=x_i|Y=y_i)=\dfrac{P(X=x_i, Y=y_i)}{P(Y=y_i)}=\dfrac{p_{ij}}{p_{\cdot j}}$
  2. 关于 Y: $P(Y=y_i|X=x_i)=\dfrac{P(X=x_i, Y=y_i)}{P(X=x_i)}=\dfrac{p_{ij}}{p_{i\cdot}}$
- 独立性: 任意 ij, 有 $p_{ij}=p_i\cdot p_j\Leftrightarrow X, Y$ 相互独立
- 举个例子

二维连续型随机变量的分布

- 分布函数：$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, du dv, \quad (x, y) \in \mathbf{R}^2$
- 边缘概率密度：
  1. 关于 X：$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy$
  2. 关于 Y：$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\, dx$
- 条件分布函数：$F_{X|Y}(x|y) = P(X \leq x | Y = y) = \int_{-\infty}^{x} \frac{f(u, y)}{f_Y(y)} du$
- 条件概率密度：$f_{X|Y}(x|y) = \dfrac{f(x, y)}{f_Y(y)}$
- 独立性：对于所有 $f(x, y)$ 连续的点 (x, y) 都有：

  $f(x, y) = f_X(x) f_Y(y) \Leftrightarrow X, Y$ 相互独立

常用的二维连续型随机变量分布

- 均匀分布：(X, Y) 服从区域上的 G 上的均匀分布，S 为该区域面积，则其概率密度为：

$$f(x, y) = \begin{cases} 1/S, & (x, y) \in G, \\ 0, & otherwise. \end{cases} \tag{7}$$

- 二维正态分布：$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，其中 $(-1 < \rho < 1)$ 当 $\rho = 0$ 时 $\Leftrightarrow$ X，Y 相互独立

例题——离散型

【例 2】 设袋中有 5 个球,其中 4 个是红球,1 个是白球,不放回抽样. 接连摸两次,令

$$X=\begin{cases}1, & \text{第一次摸到红球,}\\ 0, & \text{第一次摸到白球,}\end{cases}\quad Y=\begin{cases}1, & \text{第二次摸到红球,}\\ 0, & \text{第二次摸到白球,}\end{cases}$$

求:(1) (X,Y) 的联合分布;

(2) X,Y 的边缘分布;

(3) 在 $X=0$ 条件下 Y 的条件分布;

(4) $P\{X \geqslant Y\}$.

例题——连续型

【例 4】 设(X,Y)服从区域$D:0\leqslant y\leqslant 1-x^2$上的均匀分布，区域$B:y\geqslant x^2$，如图 3-1 所示，求：

(1) (X,Y)的联合概率密度；

(2) X,Y的边缘概率密度；

(3) 在$X=-\dfrac{1}{2}$的条件下Y的条件概率密度；

(4) $P\{(X,Y)\in B\}$.

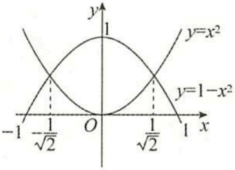


图 3-1

掌握分布函数、边缘分布函数、概率密度函数！

例题——独立性

设二维随机变量 (X, Y) 服从区域 $D\{(x, y)|a \leq x \leq b, c \leq y \leq d\}$ 上均匀分布，证明 X, Y 相互独立。

例题——独立性

设二维随机变量 (X, Y) 服从区域 $D\{(x, y)|a \le x \le b, c \le y \le d\}$ 上均匀分布，证明 X, Y 相互独立。

解

证 因为$(X,Y) \sim U(D)$,其中 $D = \{(x,y): a \leqslant x \leqslant b, c \leqslant y \leqslant d\}$,所以$(X,Y)$ 的联合密度函数为

$$p(x,y) = \begin{cases} \dfrac{1}{(b-a)(d-c)}, & a \leqslant x \leqslant b, c \leqslant y \leqslant d, \\ 0, & \text{其他}. \end{cases}$$

由此得,当 $a \leqslant x \leqslant b$ 时,

$$p_X(x) = \int_c^d p(x,y)\,\mathrm{d}y = \frac{1}{b-a};$$

当 $c \leqslant y \leqslant d$ 时,

$$p_Y(y) = \int_a^b p(x,y)\,\mathrm{d}x = \frac{1}{d-c}.$$

由此得 $p(x,y) = p_X(x) \cdot p_Y(y)$,即 X 与 Y 相互独立.

随机变量函数的概率分布

2.3 随机变量函数的概率分布

一维随机变量函数的概率分布

- 一维离散型随机变量函数的概率分布：$Y = g(X)$ 计算出不同 X 时 Y 的值，概率一一对应。（书 P70 例一）
- 一维连续型随机变量函数的概率分布，两个很严格的条件：
  1. $y = g(x)$ 为严格单调函数（但可以适当减弱书 P73 例 5、6）
  2. 其反函数 $x = h(y)$ 具有连续导数

  则 $Y = g(X)$ 才能为连续型随机变量，就有概率密度：

$$f_Y(y) = \begin{cases} f_X(h(y))|h'(y)|, & \alpha < y < \beta, \\ 0, & otherwise. \end{cases} \tag{8}$$

- 独立条件下分布的可加性：二项分布（p 要相同）、泊松分布、正态分布...

二维随机变量函数的概率分布

- 二维离散型随机变量函数的概率分布：$Z = g(X, Y)$。列表！一定要写清楚了，不要漏掉任何一种可能。如果对于不同的 (x_i, y_i)，$g(x, y)$ 有相同值，最后概率应该加起来。
- 二维连续型随机变量函数的概率分布：定义法 or 卷积

(1) 定义法：设 (X,Y) 的概率密度为 $f(x,y)$，求 $Z = g(X,Y)$ 的分布函数和概率密度.

① 先求 Z 的分布函数

$$F_Z(z) = P\{Z \leqslant z\} = P\{g(X,Y) \leqslant z\}$$
$$= P\{(X,Y) \in G\} = \iint\limits_G f(x,y)\mathrm{d}x\mathrm{d}y,$$

其中 G 由 $g(X,Y) \leqslant z$ 确定，$(X,Y) \in G$.

②Z 的概率密度为 $f_Z(z) = \dfrac{\mathrm{d}[F_Z(z)]}{\mathrm{d}z}$.

Figure: 定义法

二维随机变量函数的概率分布

- 卷积法：$Z = X + Y$，概率密度函数为：
  1. $f_Z(z) = \int_{-\infty}^{+\infty} f(z-y, y)\,dy$ （Y 型区域，先 Dy 再 Dx）
  2. $f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x)\,dx$ （X 型区域，先 Dx 再 Dy）
- 解题步骤：
  1. 换：换 X 或者 Y
  2. 确定 X、Y、Z 定义域
  3. 画图积分区域
  4. 使用卷积公式

二维随机变量函数的概率分布——连续型最值概率分布

④ 当 $Z=\max\{X,Y\}$，且 X 与 Y 相互独立时，有

$$F_Z(z)=P\{Z\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}$$
$$=P\{X\leqslant z\}P\{Y\leqslant z\}=F_X(z)F_Y(z),$$
$$f_Z(z)=\frac{\mathrm{d}[F_Z(z)]}{\mathrm{d}z}.$$

⑤ 当 $Z=\min\{X,Y\}$，且 X 与 Y 相互独立时，有

$$F_Z(z)=P\{Z\leqslant z\}=1-P\{Z>z\}=1-P\{X>z,Y>z\}$$
$$=1-P\{X>z\}P\{Y>z\}=1-(1-P\{X\leqslant z\})(1-P\{Y\leqslant z\})$$
$$=1-[1-F_X(z)][1-F_Y(z)].$$

Figure: 连续型变量最值概率分布

例题

【例 6】 设 X 与 Y 是相互独立的随机变量，其概率密度分别为

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad f_Y(y)=\begin{cases}\mu e^{-\mu y}, & y>0,\\ 0, & y\leqslant 0,\end{cases}$$

其中 $\lambda>0,\mu>0$，令 $Z=\begin{cases}1, & X\leqslant Y,\\ 0, & X>Y,\end{cases}$ 求 Z 的分布律和分布函数.

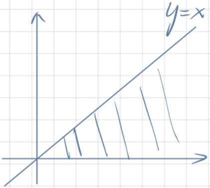
Figure: 综合题

解答

解：(1) $P(Z=1)=P(X\le Y)=\iint\limits_{X\le Y} f(x,y)\,dx\,dy$

由于 X 与 Y 相互独立，知 (X,Y) 的概率密度为：

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\lambda\mu e^{-\lambda x-\mu y}, & x>0, y>0\\ 0, & 其它\end{cases}$$



$$P(Z=1)=\iint\limits_{x\le y} f(x,y)\,dx\,dy=\lambda\int_0^{+\infty} e^{-\lambda x}dx\int_x^{+\infty} e^{-\mu y}d(\mu y)$$

$$=\lambda\int_0^{+\infty} e^{-(\lambda+\mu)x}dx=\frac{\lambda}{\lambda+\mu}$$

$\therefore P(Z=0)=1-P(Z=1)=\frac{\mu}{\lambda+\mu}$

(2) Z 的分布函数为

$$F_Z(z)=\begin{cases}0 & z<0\\ \frac{\mu}{\lambda+\mu} & 0\le z<1\\ 1 & z\ge 1\end{cases}$$

从而 Z 的分布律为：

Z	0	1
P	$\frac{\mu}{\lambda+\mu}$	$\frac{\lambda}{\lambda+\mu}$

例题

设 X，Y 概率密度为:

$$f(x, y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 2x \\ 0, & otherwise. \end{cases} \tag{9}$$

(1) 求 X，Y 的边缘概率密度 $f_X(x)$，$f_Y(y)$

(2) $Z = 2X - Y$ 的概率密度为 $f_Z(z)$

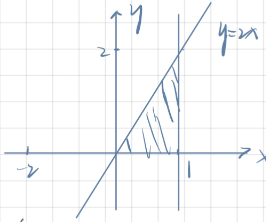
(3)$P(Y \leq \frac{1}{2} | X \leq \frac{1}{2}) =$

解答 1

解：(1) 当 $0<x<1$ 时，$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy=\int_0^{2x} dy=2x$

当 $x\le 0$ 或 $x\ge 1$ 时，$f_X(x)=0$

$$\therefore f_X(x)=\begin{cases}2x & 0<x<1\\ 0 & \text{其他}\end{cases}$$



当 $0<y<2$ 时，有 $f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\,dx=\int_{\frac{y}{2}}^{1} dx=1-\frac{y}{2}$

当 $y\le 0$，$y\ge 2$ 时，有 $f_Y(y)=0$

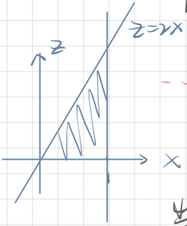
$$\therefore f_Y(y)=\begin{cases}1-\frac{y}{2} & 0<y<2\\ 0 & \text{其他}\end{cases}$$

解答 2

(2) 卷积公式: $f_Z(z)=\int_{-\infty}^{+\infty} f(x,2x-z)\,dx$ ---- ①换

$f(x,2x-z)=1$

$$\begin{cases}0<x<1\\ 0<2x-z<2x\end{cases} \Rightarrow 0<z<2x$$ ---- ②确定定义域



--- ③画图确定积分区域

$$f_Z(z)=\int_{\frac{z}{2}}^{1} 1\,dx = 1-\frac{z}{2}$$ ④卷积公式

当 $z\le 0$ 或 $z\ge 2$ 时, $f_Z(z)=0$

$$\therefore f_Z(z)=\begin{cases}1-\frac{z}{2} & 0<z<2\\ 0 & 其他\end{cases}$$

解答 3

- $P(Y \le \frac{1}{2} | X \le \frac{1}{2}) = \dfrac{P(X \le \frac{1}{2}, Y \le \frac{1}{2})}{P(X \le \frac{1}{2})} = \dfrac{\frac{3}{16}}{\frac{1}{4}} = \dfrac{3}{4}$
- **画图看面积即可**

Thank You!
April 2, 2023