



西安交通大学
XI'AN JIAOTONG UNIVERSITY

概率统计与随机过程

第二章：随机变量与概率分布

赵俊舟

junzhou.zhao@xjtu.edu.cn

2025 年 3 月 28 日

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

目录

- 1 一维随机变量
 - 随机变量与分布函数
 - 离散型随机变量
 - 连续型随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

目录

- 1 一维随机变量
 - 随机变量与分布函数
 - 离散型随机变量
 - 连续型随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

随机变量

- 很多随机试验的样本本身就是数值，例如掷一颗筛子观察点数、每天进入某超市的顾客数、手机的寿命等；
- 还有很多随机试验的样本不是数值，例如抛硬币观察正反面、人的血型等，这时需要**将样本映射为数值**以便于数学处理。

例 (检测一个产品的质量是否合格)

$$\Omega = \{\text{合格}, \text{不合格}\}$$

样本 ω	X
合格	1
不合格	0

- 如果用 p 表示产品的合格率，则 X 取不同值的概率为

X	0	1
P	$1 - p$	p

随机变量

例 (投掷一枚硬币)

$$\Omega = \{\text{正面}, \text{反面}\}$$

样本 ω	X
正面	1
反面	0

- 如果用 p 表示“出现正面”的概率, 则 X 取不同值的概率为

X	0	1
P	$1 - p$	p

例 (观察一次试验的结果)

$$\Omega = \{\text{成功}, \text{失败}\}$$

样本 ω	X
成功	1
失败	0

- 如果用 p 表示“试验成功”的概率, 则 X 取不同值的概率为

X	0	1
P	$1 - p$	p

随机变量

例 (检测三个产品中的合格情况)

用 X 表示“三个产品中的合格品数量”

样本 ω	X
$\omega_1 = (0, 0, 0)$	0
$\omega_2 = (1, 0, 0)$	1
$\omega_3 = (0, 1, 0)$	1
$\omega_4 = (0, 0, 1)$	1
$\omega_5 = (0, 1, 1)$	2
$\omega_6 = (1, 1, 0)$	2
$\omega_7 = (1, 0, 1)$	2
$\omega_8 = (1, 1, 1)$	3

- X 取各种值就是如下互不相容事件

$X = 0$ 表示事件: $\{\omega_1\}$

$X = 1$ 表示事件: $\{\omega_2, \omega_3, \omega_4\}$

$X = 2$ 表示事件: $\{\omega_5, \omega_6, \omega_7\}$

$X = 3$ 表示事件: $\{\omega_8\}$

- 若产品的合格率为 p , 则 X 取不同值的概率为

X	0	1	2	3
P	$(1-p)^3$	$3p(1-p)$	$3p^2(1-p)$	p^3

随机变量

定义 (随机变量)

随机变量是定义在样本空间 Ω 上的**实值函数** $X: \Omega \mapsto \mathbb{R}$, 常用大写字母 X, Y, Z 等表示, 其取值用小写字母 x, y, z 等表示。

- 如果一个随机变量仅取有限个或可列个值, 则称为**离散型随机变量**;
- 如果一个随机变量的可能取值充满数轴上的区间 (a, b) (其中 a 可以是 $-\infty$, b 可以是 $+\infty$), 则称为**连续型随机变量**。
- 随机变量是样本的函数, 允许不同样本对应不同的实数, 也允许对应同一实数;
- 函数的自变量 (样本) 可以是数, 也可以不是数, 但因变量一定是实数。

举例：随机变量

例 (射击评估)

记录某人进行射击训练的情况，以对其射击能力进行评估。假设给其 5 发子弹，命中记为 H，否则记为 N。问：如果关心命中数量，怎么定义随机变量？

- 样本空间： $\Omega = \{NNNNN, NNNNH, \dots, HHHHH\}$
- 随机变量 $X(\omega)$ 为样本点 ω 中 H 的数量。

例 (血糖值问题)

设空腹血糖高于 6.10mol/L 即为高血糖。随机抽一人检测血糖，研究事件“检测结果为高血糖”。

- 样本空间本身为实数集，可定义随机变量为 $X(\omega) = \omega$
- 事件“检测结果为高血糖”表示为： $A = \{\omega | X(\omega) > 6.10\}$

随机变量



- 引入随机变量，可以在更抽象和一般的层次上研究随机现象。
- 有些问题虽然具体背景不同，但数学本质完全一样。
- 例如前面的射击评估问题，可用于**类似的随机变量**的研究：
 - n 次试验中成功的次数；
 - n 件产品合格的件数；
 - n 个病人治愈的数量，等等
- 随机变量 $X(\omega)$ 就是随着试验结果 ω 的不同而变化的量。
- 随机变量 $X(\omega)$ 可以简写为 X 。

分布函数

- 若 $B \subset \mathbb{R}$ 是某些实数的集合, 则随机变量与随机事件的关系可以表示为 $\{\omega \in \Omega | X(\omega) \in B\}$, 简写为 $\{X \in B\}$ 。
- 记 X 表示掷一颗筛子出现的点数, $X \in \{1, 2, \dots, 6\}$ 。事件“点数不大于 3”可以表示为 $\{X \leq 3\}$
- 记 Y 表示一天内到达某超市的顾客数, $Y \in \{0, 1, 2, \dots\}$ 。事件“至少来 1000 位顾客”可以表示为 $\{Y \geq 1000\}$
- 记 T 表示某手机的使用寿命, $T \in [0, +\infty)$, 事件“使用寿命在 2 年到 3.5 年之间”可以表示为 $\{2 \leq T \leq 3.5\}$
- 注意到

$$\{X > x\} = \overline{\{X \leq x\}}$$

$$\{x_1 < X \leq x_2\} = \{X \leq x_2\} - \{X \leq x_1\}$$

因此, 我们只需研究形如 $\{X \leq x\}$ 的事件。

分布函数

定义 (分布函数)

设 X 是一个随机变量, 记

$$F(x) \triangleq P(X \leq x) \quad x \in \mathbb{R}$$

称 $F(x)$ 为随机变量 X 的**分布函数**。

- 与随机变量不同, 分布函数是定义在 \mathbb{R} 上的普通实函数。
- $X(\omega)$ 可以看成是样本 ω 在实数轴上的“坐标”, $F(x)$ 可以看成是坐标 $X(\omega)$ 落在区间 $(-\infty, x]$ 中的概率。
- 若已知随机变量 X 的分布函数, 就可以知道 X 落在任一区间 $(a, b]$ 的概率。
- 所以, **分布函数完整地描述了随机变量的统计规律**。

举例：分布函数

例

用随机变量 X 的分布函数 $F(x)$ 表示下列事件的概率：

(a) $\{X > a\}$

(b) $\{a < X \leq b\}$

$$P(X > a) = 1 - P(X \leq a) = 1 - F(a)$$

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

举例：分布函数

例 (收入分布)

据有关研究资料, 我国 2012 年家庭平均收入如下表:

x (千元)	1	2	4.5	9	15.9	25.8	34.3
收入不高于 x 的家庭比例	0.05	0.1	0.25	0.5	0.75	0.9	0.95

设 X 是随机抽取的一个家庭的平均收入, $F(x)$ 为其分布函数, 试用分布函数表示下列事件的概率:

(a) $\{X \leq 1\}$, (b) $\{X > 34.3\}$, (c) $\{2 < X \leq 25.8\}$.

(a) $F(1) = 0.05$

(b) $1 - F(34.3) = 1 - 0.95 = 0.05$

(c) $F(25.8) - F(2) = 0.9 - 0.1 = 0.8$

分布函数的基本性质

性质一：单调性

$F(x)$ 是一个单调非降函数。

当 $x_2 > x_1$ 时, $F(x_2) - F(x_1) = P(x_1 < X \leq x_2) \geq 0$

性质二：有界性

对任意实数 x , $0 \leq F(x) \leq 1$, 且

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$$

$$1 = P(-\infty < X < +\infty) = P\left(\bigcup_{i=-\infty}^{+\infty} \{i-1 < X \leq i\}\right) = \sum_{i=-\infty}^{+\infty} P(i-1 < X \leq i)$$

$$= \lim_{\substack{n \rightarrow +\infty \\ m \rightarrow -\infty}} \sum_{i=m}^n P(i-1 < X \leq i) = \lim_{n \rightarrow +\infty} F(n) - \lim_{m \rightarrow -\infty} F(m) \Rightarrow \begin{cases} \lim_{n \rightarrow +\infty} F(n) = 1 \\ \lim_{m \rightarrow -\infty} F(m) = 0 \end{cases}$$

分布函数的基本性质

性质三：右连续性

$F(x)$ 是右连续函数，即 $\lim_{\Delta x \rightarrow 0^+} F(x + \Delta x) = F(x)$ 或 $F(x + 0) = F(x)$ 。

因为 $F(x)$ 单调非降有界，则 $F(x)$ 在任意 x 的右极限都存在。任取一个以 x 为极限的递减数列 $x_1 > x_2 > \cdots > x_n > \cdots > x$ ，且事件列 $\{X \leq x_n\}$ 单调递减并收敛到可列交 $\bigcap_{n=1}^{\infty} \{X \leq x_n\} = \{X \leq x\}$ ，于是由概率的连续性得

$$\begin{aligned} F(x + 0) &= \lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} P(X \leq x_n) \\ &= P(\lim_{n \rightarrow \infty} \{X \leq x_n\}) = P(X \leq x) = F(x) \end{aligned}$$

💡 如果一个函数满足上述三条性质，那么这个函数一定是某个随机变量的分布函数。

举例：分布函数

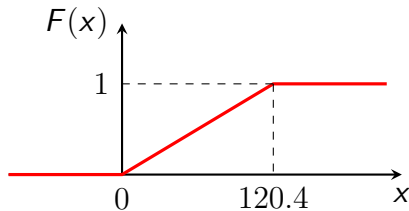
例 (污染问题)

已知空气中 PM2.5 一般在 $0.0 \sim 120.4 \mu\text{g}/\text{m}^3$ 之间。根据有关指数标准, PM2.5 含量在 100.5 以上为对人体有害。设 PM2.5 的值在任一小区间 $[a, b] \subseteq [0, 120.4]$ 中的概率与区间长 $b - a$ 成正比, 随机抽检空气质量, 求

(a) PM2.5 值 X 的分布函数并作图;

(b) 空气质量正常的概率。

- 当 $0 \leq x \leq 120.4$ 时, $F(x) = \frac{x}{120.4}$
- 当 $x < 0$ 时, $F(x) = 0$
- 当 $x > 120.4$ 时, $F(x) = 1$
- 空气质量正常的概率: $F(100.5) \approx 0.83$



目录

- 1 一维随机变量
 - 随机变量与分布函数
 - 离散型随机变量
 - 连续型随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

离散型随机变量

例

- ① 记 X 表示掷一颗筛子出现的点数, $X \in \{1, 2, \dots, 6\}$ 。事件“点数不大于 3”可以表示为 $\{X \leq 3\}$
 - ② 记 Y 表示一天内到达某超市的顾客数, $Y \in \{0, 1, 2, \dots\}$ 。事件“至少来 1000 位顾客”可以表示为 $\{Y \geq 1000\}$
 - ③ 记 T 表示某手机的使用寿命, $T \in [0, +\infty)$, 事件“使用寿命在 2 年到 3.5 年之间”可以表示为 $\{2 \leq T \leq 3.5\}$
- (1) 中的随机变量 X 取有限个值; (2) 中的随机变量 Y 取可数个值, 这些均为离散型随机变量;
 - (3) 中的随机变量 T 则连续取值, 不属离散型随机变量。

离散型随机变量

定义 (离散型随机变量)

如果一个随机变量仅取有限个或可列个值, 则称其为离散型随机变量。

定义 (分布律)

设 X 为离散型随机变量, 其所有可能的取值为 $\{x_1, x_2, \dots\}$, 记

$$p_i \triangleq P(X = x_i), \quad i = 1, 2, \dots$$

称为随机变量 X 的分布律, 也称概率函数或概率分布。

💡 由分布律可以求得分布函数:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_i$$

举例：离散型随机变量

例 (掷硬币问题)

将一枚硬币连抛三次，观察正、反面出现的情况，记 X 为正面出现的次数，求 X 的分布律。

X 的取值范围为 $\{0, 1, 2, 3\}$ ，样本空间为

$$\Omega = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$$

故 X 的分布律为

$$P(X = 0) = \frac{1}{8}, \quad P(X = 1) = \frac{3}{8}$$

$$P(X = 2) = \frac{3}{8}, \quad P(X = 3) = \frac{1}{8}$$

分布律的性质

性质 (分布律的性质)

离散型随机变量 X 的分布律有以下性质:

- **非负性:** $p_i \geq 0, i = 1, 2, \dots$

- **正则性:** $\sum_{i=1}^{\infty} p_i = 1$

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} P(X = x_i) = P\left(\bigcup_{i=1}^{\infty} \{X = x_i\}\right) = P(\Omega) = 1$$

💡 常用上式来验证所给出的随机变量的分布律是否正确。

分布律的表示方法

- 解析法：

$$p_i = P(X = x_i), \quad i = 1, 2, \dots$$

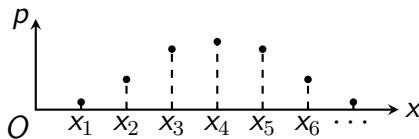
- 列表法：

X	x_1	x_2	\cdots	x_i	\cdots
P	p_1	p_2	\cdots	p_i	\cdots

- 矩阵法：

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{bmatrix}$$

- 图示法：



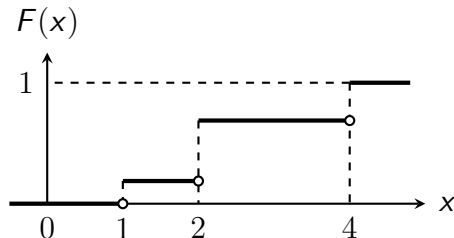
举例：离散型随机变量

例 (输血问题)

随机抽取一人观察血型， X 表示他可给别人输血的血型数。据有关资料，中国人 O, A, B, AB 血型的人分别占 41%, 28%, 24%, 7%。(1) 写出 X 的分布律；(2) 写出其分布函数，并作图。

X	1	2	4
P	0.07	0.52	0.41

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.07, & 1 \leq x < 2 \\ 0.59, & 2 \leq x < 4 \\ 1, & x \geq 4. \end{cases}$$



离散型随机变量的分布函数

- 阶梯函数，阶梯数为有限或可列个。
- 一般地，设 X 的分布律为 $P(X = x_i) = p_i, i = 1, 2, \dots$ ，则

$$F(x) = \begin{cases} 0, & x < x_1, \\ p_1, & x_1 \leq x < x_2 \\ p_1 + p_2, & x_2 \leq x < x_3 \\ \dots & \\ \sum_{j=1}^{i-1} p_j, & x_{i-1} \leq x < x_i \\ \dots & \end{cases}$$

- 分布函数又称为**累积分布函数** (Cumulative Distribution Function, CDF)。

举例：离散型随机变量

例 (路口红灯停车)

一汽车沿一街道行驶到终点，需要通过三个均设有红绿灯的路口，在每个路口遇到红灯的概率都为 $1/2$ ，且信号灯工作相互独立。以 X 表示该汽车首次停车时，在该街道上已经通过的路口数，求 X 的分布律及分布函数？

- 随机变量 $X \in \{0, 1, 2, 3\}$

- 则 X 的分布律为

X	0	1	2	3
P	$1/2$	$1/4$	$1/8$	$1/8$

- X 的分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

几种常用的离散型随机变量及其分布

- 单点分布
- 两点分布
- 二项分布
- 泊松分布
- 几何分布

单点分布、两点分布、伯努利分布

单点分布

若随机变量 X 的分布律为 $P(X = a) = 1$, 则称 X 服从**单点分布**。

两点分布

若随机变量 X 的分布律为 $P(X = a_0) = 1 - p, P(X = a_1) = p$, 或者统一表示为

$$P(X = a_i) = p^i(1 - p)^{1-i} \quad i = 0, 1$$

其中 $p \in [0, 1]$, 则称 X 服从**两点分布**。

0-1 分布 (伯努利分布, Bernoulli Distribution)

特别地, 当 $a_0 = 0, a_1 = 1$ 时, 称 X 服从**0-1 分布**, 也称为**伯努利分布**, 记为 $X \sim B(1, p)$ 。

举例：伯努利分布

例 (收入分布)

据有关研究资料, 我国 2012 年家庭平均收入如下表:

R (千元)	1	2	4.5	9	15.9	25.8	34.3
收入不高于 R 的家庭比例	0.05	0.1	0.25	0.5	0.75	0.9	0.95

若规定家庭平均收入 2 千元为贫困线。令

$$X = \begin{cases} 1, & R \leq 2 \\ 0, & R > 2 \end{cases}$$

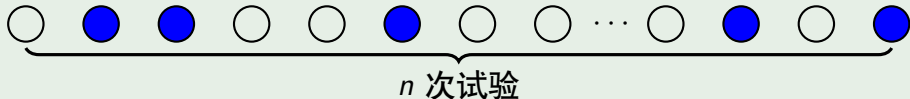
则 X 服从参数为 0.1 的伯努利分布, X 可以看成是某家庭是否是贫困家庭的**示性函数** (Indicator Function)。

伯努利试验

- 当随机试验的结果只有两个时，称为伯努利试验。
- 将伯努利试验独立地做 n 次，称为 n 重伯努利试验。
 - 两个要素：(1) 成功率 p 为常数；(2) 试验相互独立。

例 (n 重伯努利试验)

- 事件“ n 次试验中有某特定的 k 次试验成功”的概率为：
 $p^k(1-p)^{n-k}$ 。



- 事件“ n 次试验中有 k 次试验成功”的概率为
 $\binom{n}{k} p^k (1-p)^{n-k}$ ，因为一共有 $\binom{n}{k}$ 种情况，彼此互斥。

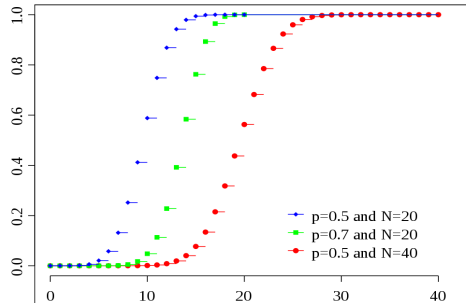
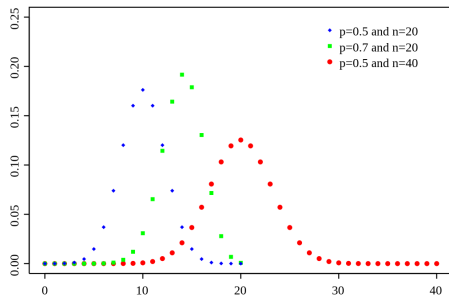
二项分布

二项分布 (Binomial Distribution)

若随机变量 X 的分布律为

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i} \quad i = 0, 1, \dots, n$$

其中 $p \in [0, 1]$, 则称 X 服从参数为 n, p 的**二项分布**, 记为 $X \sim B(n, p)$ 。



举例：二项分布

例 (电子元件)

一大批电子元件有 10% 已损坏，若从这批元件中随机选取 20 只来组成一个线路，问这个线路能正常工作的概率是多少？

- 因为元件数量很大，所以取 20 只元件可看作是有放回抽样。
- 令 X 表示“取出的 20 只元件中未损坏元件的数量”，则 $X \sim B(20, 0.9)$ 。
- 于是

$$P(\text{线路正常}) = P(X = 20) = \binom{20}{20} \times 0.9^{20} \times 0.1^0 \approx 0.1216$$

💡 二项分布理论上只适用于有放回抽样，但当总体规模很大时，也可近似用于无放回抽样。

举例：二项分布

例 (汽车故障)

有一繁忙的汽车站，每天有大量汽车通过，设每辆汽车在一天的某时段内出故障的概率为 0.001，在每天的该时段内有 1,000 辆汽车通过，问出事故的次数不小于 2 的概率是多少？

- 设 1000 辆车通过，出事故的次数为 X 。
- 由于每辆车是否发生事故相互独立，故 $X \sim B(1000, 0.001)$
- 所求概率为

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - 0.999^{1000} - \binom{1000}{1} \times 0.001 \times 0.999^{999} \approx 0.264 \end{aligned}$$

💡 应用二项分布时，往往计算比较繁琐，使用起来不太方便。

二项分布的泊松近似

定理 (泊松定理)

设随机变量 X_n 服从二项分布 $B(n, p_n)$, 其中 p_n 与 n 有关且满足 $\lim_{n \rightarrow \infty} np_n = \lambda > 0$, 则

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, \dots$$

- 实际应用中, 当 n 很大 (指 $n \geq 10$) 且 p 较小 (指 $p \leq 0.1$) 时, 有以下泊松近似公式

$$\binom{n}{k} p^k (1 - p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, \dots, n.$$

- 也称为二项分布的泊松近似。

二项分布的泊松近似

证明.

令 $np_n = \lambda_n$, 则 $p_n = \lambda_n/n$, 而

$$\begin{aligned} & \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\ &= \left[1\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right] \frac{\lambda_n^k}{k!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \end{aligned}$$

对任意固定的 k , 当 $n \rightarrow \infty$ 时, 因为

$$1\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right) \rightarrow 1, \lambda_n^k \rightarrow \lambda^k, \left(1 - \frac{\lambda_n}{n}\right)^{-k} \rightarrow 1$$

二项分布的泊松近似

及

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^n = \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda_n}{n}\right)^{\frac{n}{\lambda_n} \lambda_n} = e^{-\lambda}$$

所以

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

从而证明了泊松定理。



泊松分布

泊松分布 (Poisson Distribution)

若随机变量 X 的分布律为

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

其中 $\lambda > 0$ 为常数, 则称 X 服从参数为 λ 的**泊松分布**, 记为 $X \sim P(\lambda)$ 。

- 泊松分布由法国数学家泊松提出, 泊松分布是二项分布的极限分布。
- 故障、不幸事件、自然灾害等稀有事件在独立重复试验中出现的次数、固定时间段内到达的顾客数、显微镜下某区域中的细胞或微生物数量等, 往往服从泊松分布。

验证泊松分布为概率分布

根据指数函数 e^x 的泰勒展开, 有

$$\begin{aligned}\sum_{k=0}^{\infty} p_k &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \\&= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\&= e^{-\lambda} \left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots \right) \\&= e^{-\lambda} e^{\lambda} \\&= 1\end{aligned}$$

因此泊松分布确实为随机变量的分布律。

举例：泊松分布

例 (维修工人配备问题)

为保证设备正常工作，需配备适量的维修工人。现有同类设备 300 台，各台工作相互独立，故障率都是 0.01，一台设备的故障由一个人处理。问至少需配备多少工人，才能保证设备发生故障但不能及时维修的概率小于 0.005？

- 记发生故障的设备数为 X ，易知 $X \sim B(300, 0.01)$ 。
- 设需要配备 N 人，所要解决的问题是确定最小的 N ，使得 $P(X \leq N) \geq 0.995$ 。
- 由泊松定理得

$$P(X \leq N) \approx \sum_{k=0}^N \frac{3^k e^{-3}}{k!}$$

故希望上式 ≥ 0.995

举例：泊松分布

$$1 - F(x-1) = \sum_{r=x}^{+\infty} \frac{e^{-\lambda} \lambda^r}{r!}$$

附表 2 (续)

即希望

$$1 - \sum_{k=0}^N \frac{3^k e^{-3}}{k!} = \sum_{k=N+1}^{\infty} \frac{3^k e^{-3}}{k!} \leq 0.005$$

查泊松分布表得，满足上述不等式得最小值 N 为 8，故最少应配备 8 名工人。

x	$\lambda = 2.5$	$\lambda = 3.0$	$\lambda = 3.5$	$\lambda = 4.0$	$\lambda = 4.5$	$\lambda = 5.0$
0	1.000 000	1.000 000	1.000 000	1.000 000	1.000 000	1.000 000
1	0.917 915	0.950 213	0.969 803	0.981 684	0.988 891	0.993 262
2	0.712 703	0.800 852	0.864 112	0.908 422	0.938 901	0.959 572
3	0.456 187	0.576 810	0.679 153	0.761 897	0.826 422	0.875 348
4	0.242 424	0.352 768	0.463 367	0.566 530	0.657 704	0.734 974
5	0.108 822	0.184 737	0.274 555	0.371 163	0.467 896	0.559 507
6	0.042 021	0.083 918	0.142 386	0.214 870	0.297 070	0.384 039
7	0.014 187	0.033 509	0.065 288	0.110 674	0.168 949	0.237 817
8	0.004 247	0.011 905	0.026 739	0.051 134	0.086 586	0.133 372
9	0.001 140	0.003 803	0.009 874	0.021 363	0.040 257	0.068 094
10	0.000 277	0.001 102	0.003 315	0.008 132	0.017 093	0.031 828
11	0.000 062	0.000 292	0.001 019	0.002 840	0.006 669	0.013 695
12	0.000 013	0.000 071	0.000 289	0.000 915	0.002 404	0.005 453
13	0.000 002	0.000 016	0.000 076	0.000 274	0.000 805	0.002 019
14		0.000 003	0.000 019	0.000 076	0.000 252	0.000 698
15		0.000 001	0.000 004	0.000 020	0.000 074	0.000 226

举例：泊松分布

例 (疾病分布律)

设某地区患某种疾病的人数 $X \sim P(\lambda)$, λ 未知。若已知该地区患此病的概率为 0.001, 求 X 的分布律。

- 由 $1 - e^{-\lambda} = 0.001$ 得 $\lambda = \ln \frac{1000}{999} \approx 0.001$ 。
- 故分布律为

$$P(X = k) = 0.999 \frac{0.001^k}{k!}, \quad k = 0, 1, 2, \dots$$

几何分布

几何分布 (Geometric Distribution)

若随机变量 X 的分布律为

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

其中 $p \in (0, 1]$ 为常数, 则称 X 服从参数为 p 的几何分布, 记为 $X \sim \text{Geom}(p)$ 。

- 令 A_i 为彼此独立的伯努利试验, $P(A_i) = p$, 则试验首次成功需要的总的试验次数 $X \sim \text{Geom}(p)$:

$$P(X = k) = P(\overline{A_1}\overline{A_2}\cdots\overline{A_{k-1}}A_k) = (1 - p)^{k-1}p$$

- 有时会关心试验失败的总次数 Y , 随机变量 Y 同样被称为服从几何分布:

$$P(Y = k) = (1 - p)^k p \quad k = 0, 1, \dots$$

几何分布

- 若 $X \sim \text{Geom}(p)$, 则

$$P(X = k + n | X > k) = P(X = n), \quad k, n = 1, 2, \dots$$

称为**几何分布的无记忆性**。

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

$$P(X = k + n) = (1 - p)^{k+n-1} p$$

$$P(X > k) = \sum_{i=k+1}^{\infty} (1 - p)^{i-1} p = \frac{p(1 - p)^k}{1 - (1 - p)} = (1 - p)^k$$

$$P(X = k + n | X > k) = \frac{(1 - p)^{k+n-1} p}{(1 - p)^k} = (1 - p)^{n-1} p = P(X = n)$$

目录

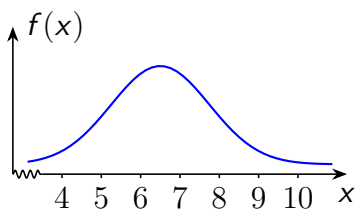
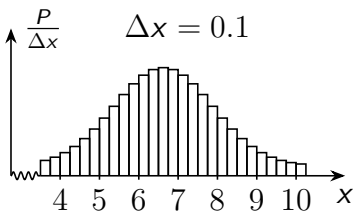
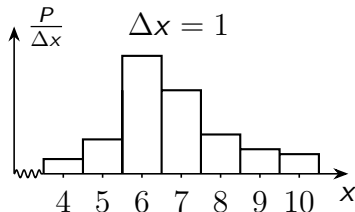
- 1 一维随机变量
 - 随机变量与分布函数
 - 离散型随机变量
 - 连续型随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

连续型随机变量

- 如果随机变量的可能取值充满某个实数区间，包含无穷不可列个实数，此时不能用分布律刻画这类**连续型随机变量**。

例 (新生儿体重)

随机变量 X 表示新生儿体重。统计大量新生儿体重，得到体重频率直方图，图中所有矩形面积之和为 1。当 $\Delta x \rightarrow 0$ ，体重频率图趋于光滑曲线 $f(x)$ 。此时 X 取值充满整个区间，成为**连续型随机变量**， $f(x)$ 成为**概率密度函数**。



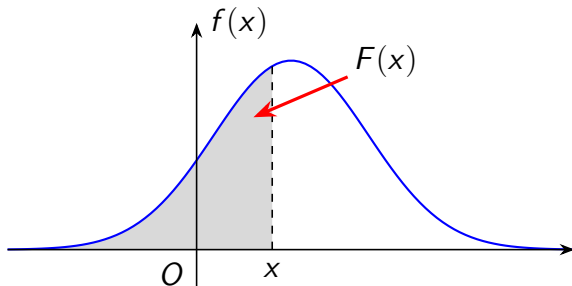
连续型随机变量

定义 (连续型随机变量及其概率密度)

设随机变量 X 的分布函数为 $F(x)$ ，如果实数轴上的一个非负可积函数 $f(x)$ 使得对任意实数 x 有

$$F(x) = \int_{-\infty}^x f(t) dt$$

则称 X 是**连续型随机变量**， $f(x)$ 是 X 的**概率密度函数**。



概率密度函数 $f(x)$ 的性质

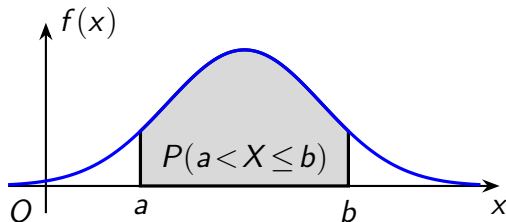
- 非负性: $f(x) \geq 0$

- 正则性: $\int_{-\infty}^{+\infty} f(t) dt = 1$

- 对任意实数 a, b 且 $a \leq b$, 有

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t) dt$$

- 若 $f(x)$ 在 x 点处连续, 则有 $F'(x) = f(x)$



连续型随机变量的性质

- 连续型随机变量的分布函数总是连续的，因为对任意点 x ，分布函数的增量总为

$$F(x + \Delta x) - F(x) = \int_x^{x+\Delta x} f(t) dt \rightarrow 0 \quad \text{as } \Delta x \rightarrow 0$$

- 对于连续型随机变量 X 有 $P(X = x) = 0, \forall x \in \mathbb{R}$ ，因为

$$P(X = x) = \int_x^x f(t) dt = 0$$

- 由于连续型随机变量 X 仅取一点的概率恒为 0，因此事件 $\{a \leq X \leq b\}$ 中去掉 a 或去掉 b 不影响其概率，即

$$P(a \leq x \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$

- 概率密度函数不唯一，改变 $f(x)$ 的“个别值”不影响积分值。

举例：连续型随机变量

例 (概率密度函数不唯一)

考虑以下两个概率密度函数

$$f_1(x) = \begin{cases} 1/a, & 0 \leq x \leq a \\ 0, & \text{其他} \end{cases} \quad f_2(x) = \begin{cases} 1/a, & 0 < x < a \\ 0, & \text{其他} \end{cases}$$

它们都是 $(0, a)$ 上均匀分布的概率密度函数。

- 可以发现

$$P(f_1(x) \neq f_2(x)) = P(X = 0) + P(X = a) = 0$$

- 可见这两个函数在概率意义上是无差别的，在此称 $f_1(x)$ 和 $f_2(x)$ “**几乎处处相等**”。
- 在概率论中可以剔除概率为 0 的事件后讨论两个函数相等及其他随机问题。

举例：连续型随机变量

例 (概率密度函数的性质)

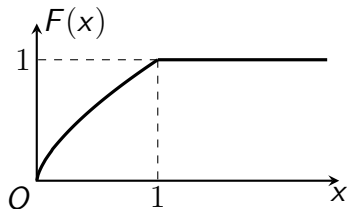
设 $f(x) = \begin{cases} kx^{-1/3}, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$ 为一概率密度函数, 求:

(a) 常数 k ; (b) 分布函数; (c) $P(1/4 < X \leq 2)$.

- $k = 2/3$
- 分布函数

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x^{2/3}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

- $P(1/4 < X \leq 2) = 1 - \frac{1}{2\sqrt[3]{2}}$



反之, 对分布函数逐段微分可得概率密度函数。

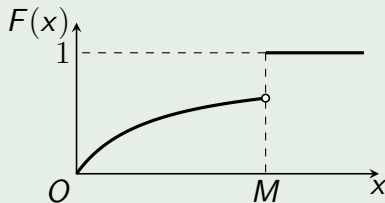
举例：连续型随机变量

- 并非所有随机变量非离散型即连续型，也有两者皆不是的。
- 例如下例的随机变量，分布函数既不是阶梯，又不连续，不属于本课程研究范围。

例 (电压问题)

前述电压问题，若电压表的量程到某 M 为止，设 X 为表测量值，则其分布函数应为

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{x}{1+x}, & 0 < x < M \\ 1, & x \geq M \end{cases}$$



X 既不能用概率密度函数，也不能用分布律刻画，但能用分布函数刻画。

几种常用的连续型随机变量概率分布

- 正态分布
- 均匀分布
- 指数分布

正态分布

- 正态分布是一种重要的分布：它所揭示的统计规律“两头小，中间大”在自然界与人类社会普遍存在：
 - 测量值与实际值的误差
 - 分子热运动时每个分子的运动速率
 - 大气中污染物的浓度

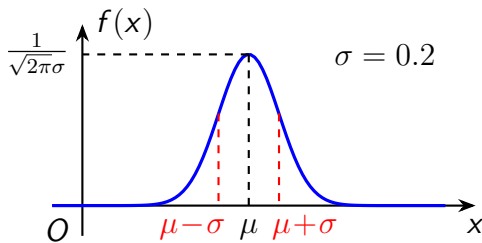
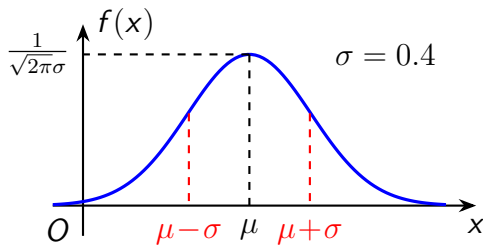
定义 (正态分布, Normal Distribution, Gaussian Distribution)

连续型随机变量 X 如果有如下形式的概率密度

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \mu \in \mathbb{R}, \sigma > 0$$

则称 X 服从参数为 μ, σ 的**正态分布**，也称作**高斯分布**，记为 $X \sim N(\mu, \sigma^2)$ 。

正态分布的性质



- 关于 μ 对称，在 $x = \mu$ 处达到极大值 $\frac{1}{\sqrt{2\pi\sigma}}$ ，越远离 μ ，概率密度值越小。
- $x = \mu \pm \sigma$ 是 $f(x)$ 的拐点。

$$\left. \frac{d^2}{dx^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right|_{x=\mu \pm \sigma} = \frac{1}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(\frac{(x-\mu)^2}{\sigma^2} - 1 \right)^2 \Big|_{x=\mu \pm \sigma} = 0$$

正态分布

- 它是否是一个概率密度？需要验证 $\int_{-\infty}^{+\infty} f(x) dx = 1$ 。
- 因为 μ 只对 $f(x)$ 左右平移，不影响积分大小，故令 $\mu = 0$ 。
- 令 $I = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} dx$ ，则

$$I^2 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} dx \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2\sigma^2}} dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2\sigma^2}} dx dy$$
 令 $x = \rho \cos \theta, y = \rho \sin \theta$ ，转化为极坐标，得

$$I^2 = \int_0^{2\pi} \int_0^{+\infty} e^{-\frac{\rho^2}{2\sigma^2}} \rho d\rho d\theta = 2\pi \int_0^{+\infty} e^{-\frac{\rho^2}{2\sigma^2}} d\frac{\rho^2}{2} = 2\pi\sigma^2$$
- 所以 $I = \sqrt{2\pi}\sigma$ ，从而 $\int_{-\infty}^{+\infty} f(x) dx = 1$

举例：正态分布

例（男性身高）

设成年男子身高 $X \sim N(1.7, 0.1^2)$ ，随机观察一男子，求其身高超过 1.8 的概率。

- 即要计算

$$P(X > 1.8) = \int_{1.8}^{+\infty} \frac{1}{\sqrt{2\pi} \times 0.1} e^{-\frac{(x-1.7)^2}{2 \times 0.1^2}} dx = ?$$

- 一般地，形如 $\int_{-\infty}^x e^{-t^2} dt$ 的积分得不到解析表达式，只能作数值积分，对每个 x 制表，查表计算。
- 但是，若 μ, σ 也变化怎么办？

正态分布的标准化

考虑变换 $Z = \frac{X-\mu}{\sigma}$, 称为 X 的**标准化**。

定理 (正态分布的标准化)

设 $X \sim N(\mu, \sigma^2)$, 令 $Z = \frac{X-\mu}{\sigma}$, 则 $Z \sim N(0, 1)$ 。

证明.

令 Z 的分布函数和概率密度函数分别为 $\Phi(x)$ 和 $\phi(x)$, 则

$$\Phi(x) = P(Z < x) = P(X < \sigma x + \mu) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\sigma x + \mu} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

令 $\frac{t-\mu}{\sigma} = u$, 则

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du \quad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

所以 $Z \sim N(0, 1)$ 。

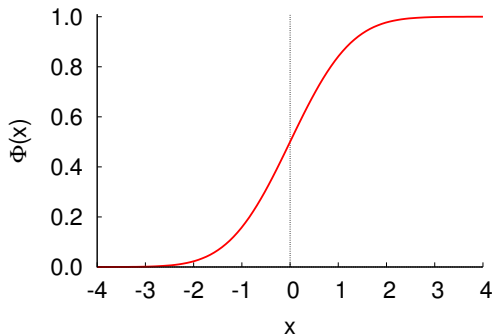
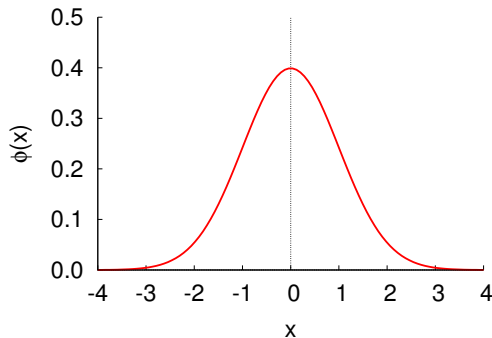


标准正态分布表

故只需对 $\Phi(x)$ 制表, 列出结果 $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.0251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

标准正态分布



- 称 $N(0, 1)$ 为**标准正态分布**，其概率密度函数 $\phi(x)$ 和分布函数 $\Phi(x)$ 形状如上图。
- 实际问题中，常有先给出 $\Phi(x_\alpha) = \alpha$ ，求反函数 $\Phi^{-1}(\alpha) = x_\alpha$ ，称 x_α 为 **α 分位数** (quantile)。

举例：正态分布

例（男性身高续，查表求解法）

令 $Z = \frac{X-1.7}{0.1}$ ，得 $X = 0.1Z + 1.7$ ，故

$$P(X > 1.8) = P(0.1Z + 1.7 > 1.8) = P(Z > 1) = 1 - \Phi(1)$$

查表得 $1 - \Phi(1) = 1 - 0.8413 = 0.1587$ 。

例（男性身高续，分位数）

城市公交车门高度设计中，受结构限制不能太高。但为了保证95%的男性都能顺畅登车，应保证车门的高度不低于多少？

- 设门的最低高度为 x 时可使 $P(X < x) \geq 0.95$
- 代入 $X = 0.1Z + 1.7$ 得
$$P(0.1Z + 1.7 < x) = P(Z < 10x - 17) = \Phi(10x - 17) \geq 0.95$$
- 查表得 $\Phi(1.65) = 0.9505$ ，故 $x = 1.865$

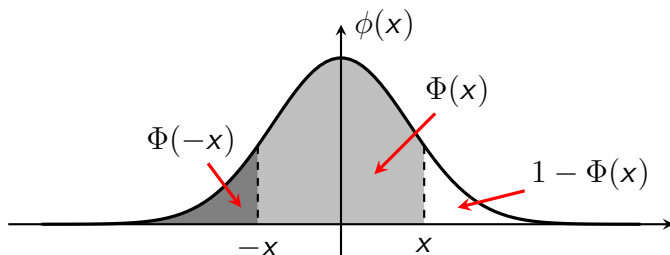
正态分布与 “ 3σ 法则”

- 标准正态分布表中，通常只列到 $x = 3.9$ ，为什么？

$$P(|X - \mu| \leq \sigma) = 0.6827 \quad P(|X - \mu| \leq 2\sigma) = 0.9545$$

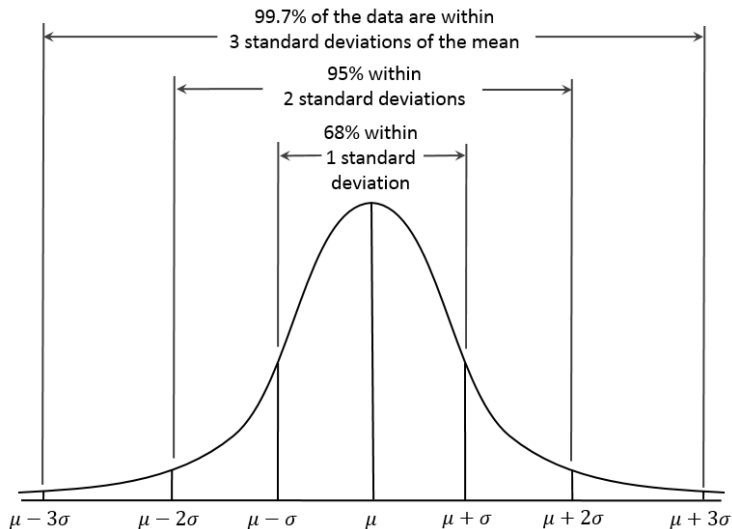
$$P(|X - \mu| \leq 3\sigma) = 0.9973$$

- 一个正态分布随机变量的取值几乎全部集中在区间 $(\mu - 3\sigma, \mu + 3\sigma)$ 内，超出这个范围的可能性不到 0.3%。
- 标准正态分布表中没有给出 $x < 0$ 的情况，为什么？



正态分布与“ 3σ 法则”

- 工程上采用的 3σ 法则：要求正态变量（例如产品精度指标等）的取值位于 $\mu \pm 3\sigma$ 之间。



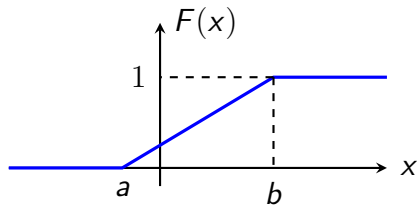
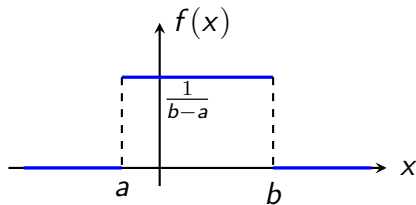
均匀分布

定义 (均匀分布, Uniform Distribution)

设随机变量 X 具有如下形式的概率密度函数

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{其他} \end{cases}$$

则称 X 服从区间 $[a, b]$ 上的**均匀分布**, 记为 $X \sim U(a, b)$ 。



在 $[a, b]$ 内任意小区间取值的概率只与小区间的长度有关, 与其位置无关。

指数分布

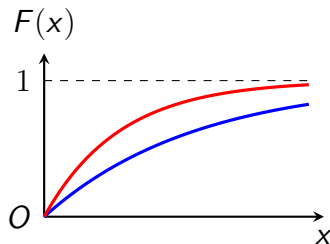
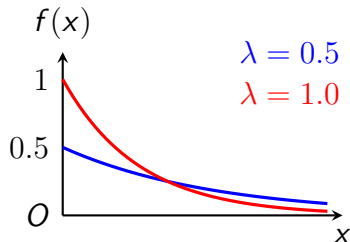
定义 (指数分布, Exponential Distribution)

若随机变量 X 具有如下形式的概率密度函数

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

称 X 服从参数为 λ 的**指数分布**, 记为 $X \sim \exp(\lambda)$ 。其分布函数为

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



指数分布

- 指数分布通常用来描述生命周期（生物、产品等）。
- 无记忆性**：设 $X \sim \exp(\lambda)$ ，则对于任意 $t, s > 0$ ，有
$$P(X > s + t | X > s) = P(X > t)$$

这是因为

$$P(X > s + t | X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t)$$

- 表示“永远年青”：用 X 表示某人寿命，在已知其寿命超过 s 的条件下，还能活 t 年的概率与在其出生时，问他还能活 t 年的概率是相同的。

目录

- ① 一维随机变量
- ② 二维随机变量
 - 二维随机变量与联合分布函数
 - 二维离散型随机变量
 - 二维连续型随机变量
- ③ 二维随机变量的条件分布
- ④ 随机变量的相互独立性
- ⑤ 随机变量函数的概率分布

目录

- 1 一维随机变量
- 2 二维随机变量
 - 二维随机变量与联合分布函数
 - 二维离散型随机变量
 - 二维连续型随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

二维随机变量

- 有些随机事件用一个随机变量无法描述, 需要两个或多个随机变量描述, 如:
 - 导弹的落点 $(X, Y) \rightarrow (\text{横坐标}, \text{纵坐标})$
 - 人的血压 $(X, Y) \rightarrow (\text{收缩压}, \text{舒张压})$
 - 一段时期某支股票行情 $(X, Y) \rightarrow (\text{平均市值}, \text{波动值})$
- 为什么不分别研究 X, Y , 而要整体研究 (X, Y) ?
- 分别研究不能体现 X 与 Y 之间的关系: 例如, (X, Y) 表示收缩压与舒张压, Y 必须满足 $Y \leq X$, 而且实际上两者悬殊不可能太大, 即 X, Y 不能分别自由取值。
- 即 X 与 Y 不独立。

n 维随机变量

定义 (n 维随机变量)

如果 $X_1(\omega), X_2(\omega), \dots, X_n(\omega)$ 是定义在同一样本空间 Ω 上的 n 个随机变量, 则称

$$X(\omega) = (X_1(\omega), X_2(\omega), \dots, X_n(\omega))$$

为 n 维随机变量或随机向量。

- 等价说法: 二维随机变量是 Ω 到 \mathbb{R}^2 的映射。
- 常见的一种情况是 $\Omega \subseteq \mathbb{R}^2$ 。例如 Ω 表示血压测量结果或导弹落点结果, 此时可定义 $X(x, y) \triangleq x, Y(x, y) \triangleq y$ 。

二维随机变量

例 (掷双筛子问题, 续)

掷两颗筛子, 则样本空间 $\Omega = \{(i, j) | i, j = 1, 2, \dots, 6\}$ 。定义 $X(i, j) = i + j$, $Y(i, j) = i \cdot j$, 则 (X, Y) 为 Ω 上的二维随机变量。

- 如何从整体上研究 (X, Y) 的分布?

- 即研究形如

$$\begin{aligned}\{\omega | (X(\omega), Y(\omega)) \in (a, b] \times (c, d]\} &\triangleq \{(X, Y) \in (a, b] \times (c, d]\} \\ &= \{a < X \leq b, c < Y \leq d\}\end{aligned}$$

的事件的概率。

- 为此, 须与一维情形类似, 研究二维随机变量的分布函数。

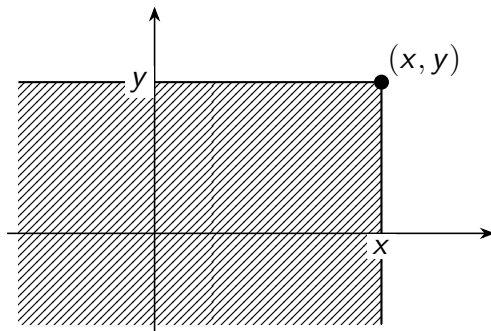
联合分布函数

定义 (联合分布函数, Joint Distribution Function)

设 (X, Y) 为一个二维随机变量, 对任意 $(x, y) \in \mathbb{R}^2$, 二元函数

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

称为二维随机变量 (X, Y) 的联合分布函数。



联合分布函数的性质

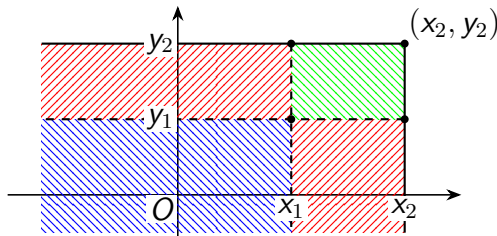
- 固定一个变元, $F(x, y)$ 是另一个变元的单调非减函数。

- $F(x, y)$ 对每个变元右连续, 即

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

- $\forall x, y$ 都有: $0 \leq F(x, y) \leq 1$, $F(+\infty, +\infty) = 1$,
 $F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0$ 。

- 对于任意两点 $(x_1, y_1), (x_2, y_2)$, 若 $x_1 \leq x_2, y_1 \leq y_2$, 则
 $F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \geq 0$



边缘分布函数

- 给定二维随机变量 (X, Y) 的联合分布函数 $F(x, y)$, 则它的两个分量 X, Y 的分布函数 $F_X(x), F_Y(y)$ 也随之确定。

$$F_X(x) \triangleq P(X \leq x) = P(X \leq x, Y \leq +\infty) = F(x, +\infty)$$

- 同理可得 $F_Y(y) \triangleq F(+\infty, y)$
- 称 $F_X(x)$ 和 $F_Y(y)$ 分别为二维随机变量 (X, Y) 关于 X, Y 的**边缘分布函数**。
- 联合分布函数唯一确定边缘分布函数, 但反过来不一定成立。

目录

- ① 一维随机变量
- ② 二维随机变量
 - 二维随机变量与联合分布函数
 - 二维离散型随机变量
 - 二维连续型随机变量
- ③ 二维随机变量的条件分布
- ④ 随机变量的相互独立性
- ⑤ 随机变量函数的概率分布

二维离散型随机变量及其联合分布律

定义 (二维离散型随机变量)

若二维随机变量 (X, Y) 的所有可能取值是有限对或者无穷可列对 $(x_i, y_i), i, j = 1, 2, \dots$, 则称为**二维离散型随机变量**, 并称

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

为二维离散型随机变量 (X, Y) 的**联合分布律**。

- 联合分布律常用表格表示, 满足

- 非负性: $p_{ij} \geq 0, \forall i, j$
- 正则性: $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$

- (X, Y) 的**联合分布函数**为

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

Y \ X	X			
	x_1	\cdots	x_j	\cdots
y_1	p_{11}	\cdots	p_{1j}	\cdots
\vdots	\vdots		\vdots	
y_j	p_{1j}	\cdots	p_{ij}	\cdots
\vdots	\vdots		\vdots	

举例：二维离散型随机变量

例

设二维离散型随机变量 (X, Y) 的联合分布律如下表，求 $P(X \geq 2, Y \geq 2)$ 及 $P(X = 1)$ 。

$\begin{matrix} Y \backslash X \\ \end{matrix}$	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

- $P(X \geq 2, Y \geq 2) = 0.1 + 0.2 + 0.2 = 0.5$
- $P(X = 1) = 0.1 + 0.3 + 0 = 0.4$

边缘分布律

- 设 (X, Y) 为二维离散型随机变量, 其联合分布律为

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

- 对于固定的 x_i , 考虑和式

$$\begin{aligned} p_{i\cdot} &\triangleq \sum_{j=1}^{\infty} p_{ij} = \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) = P\left(\bigcup_{j=1}^{\infty} \{X = x_i, Y = y_j\}\right) \\ &= P\left(\{X = x_i\} \cap \bigcup_{j=1}^{\infty} \{Y = y_j\}\right) = P(X = x_i) \end{aligned}$$

- 称 $\{p_{i\cdot}, i = 1, 2, \dots\}$ 为 X 的**边缘分布律**。

边缘分布律

同理称

$$p_{\cdot j} \triangleq \sum_{i=1}^{\infty} p_{ij} = P(Y = y_j), \quad j = 1, 2, \dots$$

为 Y 的**边缘分布律**，因为位于联合分布律表格的边缘。

$\begin{array}{c} X \\ Y \end{array}$	x_1	\cdots	x_i	\cdots	$p_{\cdot j}$
y_1	p_{11}	\cdots	p_{i1}	\cdots	$p_{\cdot 1}$
\vdots	\vdots		\vdots		\vdots
y_j	p_{1j}	\cdots	p_{ij}	\cdots	$p_{\cdot j}$
\vdots	\vdots		\vdots		\vdots
$p_{i\cdot}$	$p_{1\cdot}$	\cdots	$p_{i\cdot}$	\cdots	1

举例：二维离散型随机变量

例 (掷双骰子)

求 $X = |i - j|$, $Y = i + j$ 的联合分布律及边缘分布律。

$\begin{matrix} X \\ Y \end{matrix}$	0	1	2	3	4	5	$p_{\cdot j}$
2	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
3	0	$\frac{2}{36}$	0	0	0	0	$\frac{2}{36}$
4	$\frac{1}{36}$	0	$\frac{2}{36}$	0	0	0	$\frac{3}{36}$
5	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	$\frac{4}{36}$
6	$\frac{1}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{5}{36}$
7	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	$\frac{6}{36}$
8	$\frac{1}{36}$	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	$\frac{5}{36}$
9	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	$\frac{4}{36}$
10	$\frac{1}{36}$	0	$\frac{2}{36}$	0	0	0	$\frac{3}{36}$
11	0	$\frac{2}{36}$	0	0	0	0	$\frac{2}{36}$
12	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
$p_{i \cdot}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	1

目录

- ① 一维随机变量
- ② 二维随机变量
 - 二维随机变量与联合分布函数
 - 二维离散型随机变量
 - 二维连续型随机变量
- ③ 二维随机变量的条件分布
- ④ 随机变量的相互独立性
- ⑤ 随机变量函数的概率分布

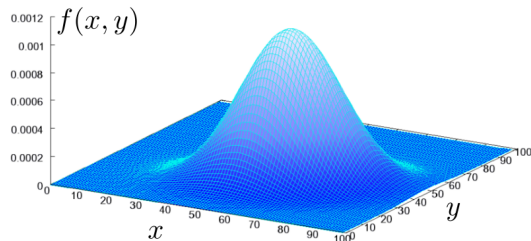
二维连续型随机变量

定义 (二维连续型随机变量)

如果存在非负函数 $f(x, y)$ 使得二维随机变量 (X, Y) 的分布函数 $F(x, y)$ 可以表示为

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

则称 (X, Y) 为**二维连续型随机变量**，称 $f(x, y)$ 为 (X, Y) 的**联合概率密度**，简称**概率密度**。



二维连续型随机变量的性质

- 非负性: $f(x, y) \geq 0, \forall x, y \in \mathbb{R}$
- 正则性:

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

- 在 $f(x, y)$ 的连续点处, 有

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$$

- 对于任何 \mathbb{R}^2 上的区域 G , 有

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

💡 对于二维连续型随机变量, 孤立点或曲线上的概率均为 0。

举例：二维连续型随机变量

例

设随机变量 (X, Y) 的分布函数为

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

求 (X, Y) 的联合概率密度。

显然, $F(x, y)$ 是连续的, 且除了直线 $x = 0$ 和 $y = 0$, $F(x, y)$ 在 \mathbb{R}^2 每点都可导, 故有

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y) = \begin{cases} \lambda^2 e^{-\lambda(x+y)}, & x > 0, y > 0, \\ 0, & \text{其他} \end{cases}$$

举例：二维连续型随机变量

💡 利用联合概率密度容易计算一些随机事件的概率，计算时要特别注意积分上下限。

例

设随机变量 (X, Y) 的联合概率密度为

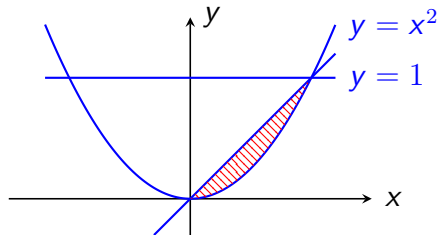
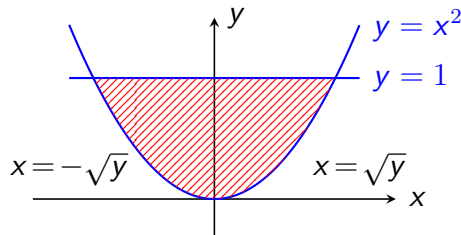
$$f(x, y) = \begin{cases} cx^2y, & x^2 \leq y \leq 1, \\ 0, & \text{其他} \end{cases}$$

(1) 确定常数 c ; (2) 计算 $P(X > Y)$ 。

举例：二维连续型随机变量

$$\begin{aligned}
 1 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = c \int_0^1 y dy \int_{-\sqrt{y}}^{\sqrt{y}} x^2 dx \\
 &= c \int_0^1 \frac{2}{3} y^{5/2} dy = \frac{4}{21} c \Rightarrow c = \frac{21}{4}
 \end{aligned}$$

$$P(X > Y) = \iint_{x>y} f(x, y) dx dy = \frac{21}{4} \int_0^1 y dy \int_y^{\sqrt{y}} x^2 dx = \frac{3}{20}$$



举例：二维连续型随机变量

💡 对于连续型随机变量，如果 (X, Y) 取值在某区域 G 中，即意味着 $f(x, y)$ 在 \bar{G} 上为 0。

例

设随机变量 (X, Y) 在 $[0, 2] \times [0, 2]$ 中取值，且分布函数在此区域的价值如下：

$$F(x, y) = \frac{1}{16}xy(x + y), \quad 0 \leq x \leq 2, 0 \leq y \leq 2$$

- (a) 完整写出 $F(x, y)$ 的定义；
- (b) 求其概率密度；
- (c) 求 $P(X + Y \leq 1)$ ；
- (d) 对于 $x, y \in [0, 2]$ ，分别求 $P(X \leq x)$ 和 $P(Y \leq y)$ 。

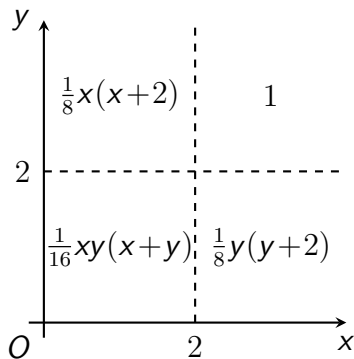
举例：二维连续型随机变量

(a) 由分布函数的连续性，得

$$F(x, y) = \begin{cases} 0, & x < 0 \text{ 或 } y < 0, \\ \frac{1}{8}x(x+2), & x \in [0, 2], y > 2, \\ \frac{1}{8}y(y+2), & y \in [0, 2], x > 2, \\ \frac{1}{16}xy(x+y), & x, y \in [0, 2], \\ 1, & x > 2, y > 2 \end{cases}$$

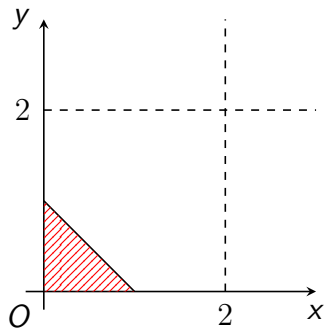
(b) 联合概率密度为

$$f(x, y) = \begin{cases} \frac{1}{8}(x+y), & x, y \in [0, 2], \\ 0, & \text{其他} \end{cases}$$



举例：二维连续型随机变量

$$\begin{aligned}
 P(X + Y \leq 1) &= \iint_{x+y \leq 1} f(x, y) \, dx \, dy \\
 &= \iint_{x+y \leq 1} \frac{1}{8}(x + y) \, dx \, dy \\
 &= \frac{1}{24}
 \end{aligned}$$



(d) 对于 $x \in [0, 2]$, 有

$$P(X \leq x) = F(x, +\infty) = F(x, 2) = \frac{1}{8}x(x + 2)$$

同理, 对于 $y \in [0, 2]$, 有 $P(Y \leq y) = \frac{1}{8}y(y + 2)$ 。

二维连续型随机变量的边缘概率密度

- 当给定二维连续型随机变量 (X, Y) 的联合概率密度 $f(x, y)$, 那么 X 和 Y 的概率密度 $f_X(x)$ 和 $f_Y(y)$ 也就随之确定。因为

$$F_X(x) = F(x, +\infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(u, y) dy du$$

- 对 x 求导, 得到 X 的**边缘概率密度**为

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

- 同理, 可得 Y 的**边缘概率密度**为

$$f_Y(y) \triangleq \int_{-\infty}^{+\infty} f(x, y) dx$$

举例：二维连续型随机变量的边缘概率密度

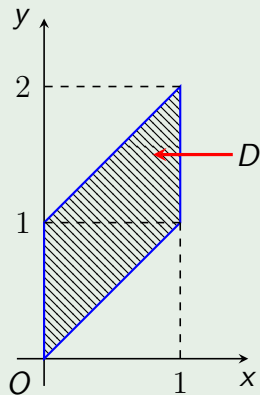
💡 联合概率密度往往是分片定义的，在计算边缘概率密度积分时，需要仔细确定积分限。

例

设 D 为 xy 平面上由 $x = 0$, $x = 1$, $x = y$, $x = y - 1$ 围成的区域，定义随机变量 X, Y 的联合概率密度如下

$$f(x, y) = \begin{cases} c, & (x, y) \in D, \\ 0, & \text{其他} \end{cases}$$

(a) 确定常数 c ; (b) 求 X, Y 的边缘概率密度 $f_X(x), f_Y(y)$ 。

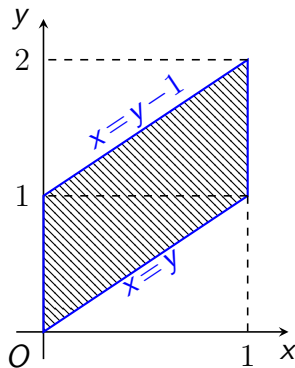


二维连续型随机变量的边缘概率密度

$$1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = \iint_D c dx dy = c$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_x^{x+1} dy = 1, x \in [0, 1]$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^y dx = y, & y \in [0, 1] \\ \int_{y-1}^1 dx = 2 - y, & y \in (1, 2] \\ 0, & \text{其他} \end{cases}$$



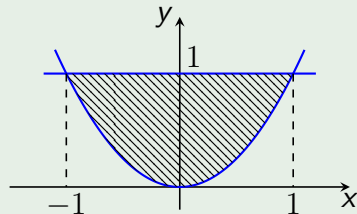
举例：二维连续型随机变量的边缘概率密度

例

设 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} \frac{21}{4}x^2y, & x^2 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

求 X, Y 的边缘概率密度 $f_X(x), f_Y(y)$ 。



$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_{x^2}^1 \frac{21}{4}x^2y dy = \frac{21}{8}x^2(1-x^4), \quad x \in [-1, 1]$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2y dx = \frac{7}{2}y^{\frac{5}{2}}, \quad y \in [0, 1]$$

二维正态分布

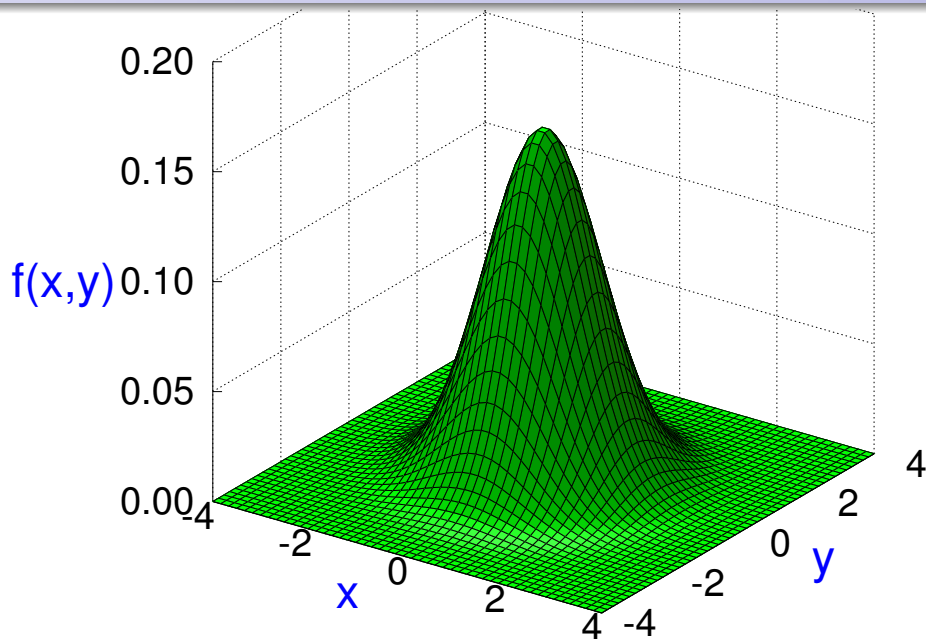
定义 (二维正态分布)

若二维随机变量 (X, Y) 的概率密度为

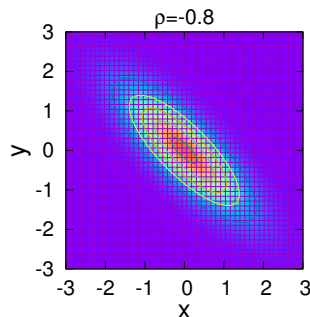
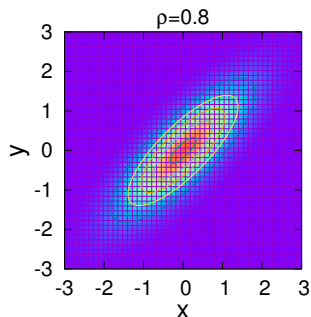
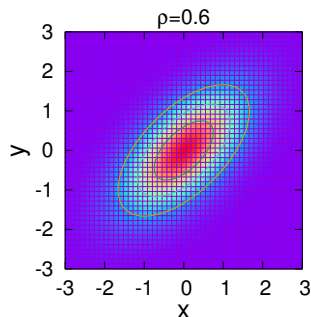
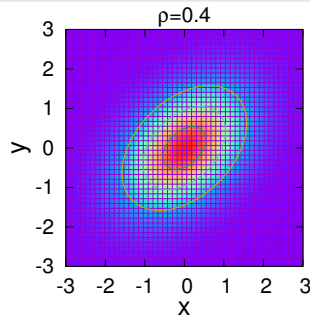
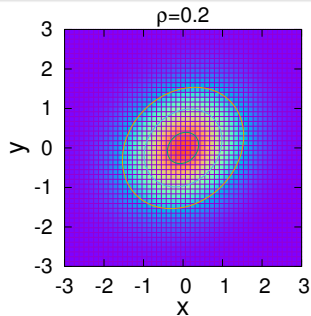
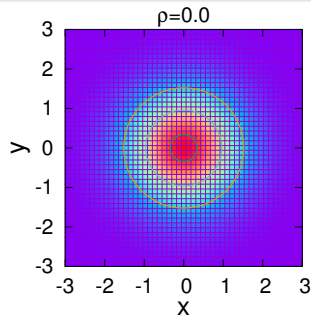
$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$

其中 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 为常数, 且 $\sigma_1 > 0, \sigma_2 > 0, |\rho| < 1$, 则称 (X, Y) 服从**二维正态分布** $N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$ 。

二维正态分布 $N(0, 1; 0, 1; 0)$



二维正态分布



二维正态分布的边缘概率分布

令 $u = \frac{x - \mu_1}{\sigma_1}$, $v = \frac{y - \mu_2}{\sigma_2}$, 则 X 的边缘概率密度为

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy \\
 &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \exp \left\{ \frac{-1}{2(1-\rho^2)} [u^2 - 2\rho uv + v^2] \right\} \sigma_2 dv \\
 &= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \exp \left\{ \frac{-1}{2(1-\rho^2)} [u^2(1-\rho^2) + (v-\rho u)^2] \right\} dv \\
 &= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \exp \left\{ \frac{-u^2}{2} - \frac{(v-\rho u)^2}{2(1-\rho^2)} \right\} dv \\
 &= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \exp \left(\frac{-u^2}{2} \right) \int_{-\infty}^{+\infty} \exp \left\{ -\frac{(v-\rho u)^2}{2(1-\rho^2)} \right\} dv
 \end{aligned}$$

二维正态分布的边缘概率分布

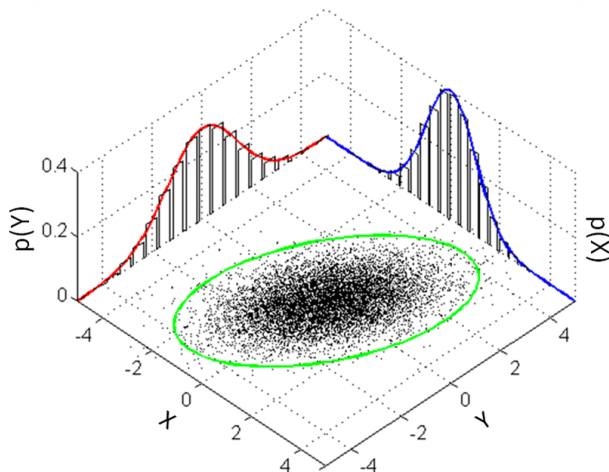
令 $t = \frac{v - \rho u}{\sqrt{1 - \rho^2}}$ 则继续上一页的推导, 有

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \exp\left(-\frac{u^2}{2}\right) \int_{-\infty}^{+\infty} \exp\left(-\frac{t^2}{2}\right) \sqrt{1-\rho^2} dt \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{u^2}{2}\right) \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \\ &= N(\mu_1, \sigma_1^2) \end{aligned}$$

同理可得 $f_Y(y) = N(\mu_2, \sigma_2^2)$ 。

二维正态分布的边缘概率分布

- 二维正态分布的边缘概率分布仍为正态分布。
- 注意两个边缘概率密度均与 ρ 无关。



二维均匀分布

定义 (二维均匀分布)

设 (X, Y) 为二维随机变量, 取值于平面上有界区域 $G \in \mathbb{R}^2$ 中, 面积为 A , 其概率密度为

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

称 (X, Y) 在区域 G 上服从**二维均匀分布**。

- 二维均匀分布描述了随机变量在取值空间的每一点概率密度都是相等的。

举例：二维均匀分布

例 (换乘问题)

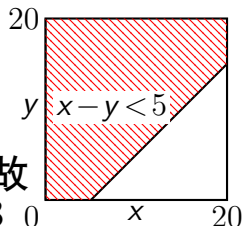
某人乘甲客车到某站换乘乙客车。已知两车到达该车站的时间均为等可能地取 8:00 到 8:20 之间, 又客车在此站停 5 分钟, 问该乘客能成功换乘的概率为多少?

- 设乘客到达车站的时间为 $X \in [0, 20]$, 所要转乘的客车到达车站时间为 $Y \in [0, 20]$ 。则 (X, Y) 服从 $[0, 20] \times [0, 20]$ 上的均匀分布, 概率密度为

$$f(x, y) = \begin{cases} \frac{1}{400}, & (x, y) \in D \\ 0, & \text{其他} \end{cases}$$

- 事件“能成功转乘”为 $G = \{(x, y) | x - y < 5\}$, 故

$$P\{\text{能成功转乘}\} = \iint_G \frac{1}{400} dx dy = \frac{23}{32}$$



目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
 - 离散型随机变量的条件分布
 - 连续型随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
 - 离散型随机变量的条件分布
 - 连续型随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

离散型随机变量的条件分布

定义 (条件分布律)

设 (X, Y) 为二维离散型随机变量, 其联合分布律以及关于 X 、关于 Y 的边缘概率分布分别为 $p_{ij}, p_{i\cdot}, p_{\cdot j}, i, j = 1, 2, \dots$

- 对于固定的 j , 若 $p_{\cdot j} = P(Y = y_j) > 0$, 则称

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

为在 $Y = y_j$ 条件下随机变量 X 的条件分布律。

- 同理, 对于固定的 i , 若 $p_{i\cdot} = P(X = x_i) > 0$, 则称

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

为在 $X = x_i$ 条件下随机变量 Y 的条件分布律。

离散型随机变量的条件分布

- 显然，条件分布律是个分布律，因为

$$P(Y = y_j | X = x_i) \geq 0$$

- 并且

$$\sum_j P(Y = y_j | X = x_i) = \frac{\sum_j p_{ij}}{p_{i\cdot}} = \frac{p_{i\cdot}}{p_{i\cdot}} = 1$$

条件分布律的直观意义

- 将某变量固定时，例如 $X = x_i$ ，将原来的样本空间缩为一维点集 $\{(x_i, y_1), (x_i, y_2), \dots\}$ 。
- 考虑其上的分布律，但因该点集概率之和不等于 1，而等于边缘分布律 $p_{i\cdot}$ ，故需要用它来规范化。

举例：离散型随机变量的条件分布

例

一袋中有 6 个大小形状相同的球，其中 2 个为红色，4 个为白色，每次从袋中任取一球，共取两次。定义随机变量 X, Y 如下

$$X = \begin{cases} 0, & \text{第一次取出红球} \\ 1, & \text{第一次取出白球} \end{cases} \quad Y = \begin{cases} 0, & \text{第二次取出红球} \\ 1, & \text{第二次取出白球} \end{cases}$$

考虑两种取球方式：有放回取球和无放回取球。求在 $X = 0$ 的条件下， Y 的条件分布律。

举例：离散型随机变量的条件分布

- 有放回取球：

$$P(Y = 0|X = 0) = \frac{P(X = 0, Y = 0)}{P(X = 0)} = \frac{1/9}{1/3} = \frac{1}{3}$$

$$P(Y = 1|X = 0) = \frac{P(X = 0, Y = 1)}{P(X = 0)} = \frac{2/9}{1/3} = \frac{2}{3}$$

- 无放回取球：

$$P(Y = 0|X = 0) = \frac{P(X = 0, Y = 0)}{P(X = 0)} = \frac{1/15}{1/3} = \frac{1}{5}$$

$$P(Y = 1|X = 0) = \frac{P(X = 0, Y = 1)}{P(X = 0)} = \frac{4/15}{1/3} = \frac{4}{5}$$

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
 - 离散型随机变量的条件分布
 - 连续型随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布

连续型随机变量的条件分布

- 对于连续型随机变量，无法计算条件概率 $P(X \leq x|Y = y)$ ，因为对于连续型随机变量 Y ，始终有 $P(Y = y) = 0$ 。
- 此时，可以将 $P(X \leq x|Y = y)$ 看作是当 $\delta \rightarrow 0$ 时 $P(X \leq x|y \leq Y \leq y + \delta)$ 的极限。
- 于是有

$$\begin{aligned}
 P(X \leq x|Y = y) &= \lim_{\delta \rightarrow 0} P(X \leq x|y \leq Y \leq y + \delta) \\
 &= \lim_{\delta \rightarrow 0} \frac{P(X \leq x, y \leq Y \leq y + \delta)}{P(y \leq Y \leq y + \delta)} = \lim_{\delta \rightarrow 0} \frac{\int_{-\infty}^x \int_y^{y+\delta} f(u, v) du dv}{\int_y^{y+\delta} f_Y(v) dv} \\
 &= \lim_{\delta \rightarrow 0} \frac{\int_{-\infty}^x \frac{1}{\delta} \int_y^{y+\delta} f(u, v) du dv}{\frac{1}{\delta} \int_y^{y+\delta} f_Y(v) dv}
 \end{aligned}$$

连续型随机变量的条件分布

$$P(X \leq x | Y = y) = \lim_{\delta \rightarrow 0} \frac{\int_{-\infty}^x \frac{1}{\delta} \int_y^{y+\delta} f(u, v) \, du \, dv}{\frac{1}{\delta} \int_y^{y+\delta} f_Y(v) \, dv}$$

当 $f(x, y), f_Y(y)$ 在 y 处连续时, 由积分中值定理可得

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_y^{y+\delta} f_Y(v) \, dv = f_Y(y)$$

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_y^{y+\delta} f(u, v) \, dv = f(u, y)$$

所以

$$P(X \leq x | Y = y) = \int_{-\infty}^x \frac{f(u, y)}{f_Y(y)} \, du \triangleq \int_{-\infty}^x f_{X|Y}(u|y) \, du$$

连续型随机变量的条件分布

定义 (连续型随机变量条件概率密度)

设 (X, Y) 为二维连续型随机变量, 其联合概率密度为 $f(x, y)$, (X, Y) 关于 X 、 Y 的边缘概率密度为 $f_X(x)$, $f_Y(y)$ 。

- 若对于固定的 y , $f_Y(y) > 0$, 则在给定 $Y = y$ 时 X 的条件分布定义为

$$F_{X|Y}(x|y) \triangleq P(X \leq x | Y = y) = \int_{-\infty}^x f_{X|Y}(u|y) du$$

其中

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

为在 $Y = y$ 条件下 X 的条件概率密度。

连续型随机变量的条件分布

- 若对于固定的 x , $f_X(x) > 0$, 则在给定 $X = x$ 时 Y 的条件分布定义为

$$F_{Y|X}(y|x) \triangleq P(Y \leq y | X = x) = \int_{-\infty}^y f_{Y|X}(v|x) dv$$

其中

$$f_{Y|X}(y|x) \triangleq \frac{f(x, y)}{f_X(x)}$$

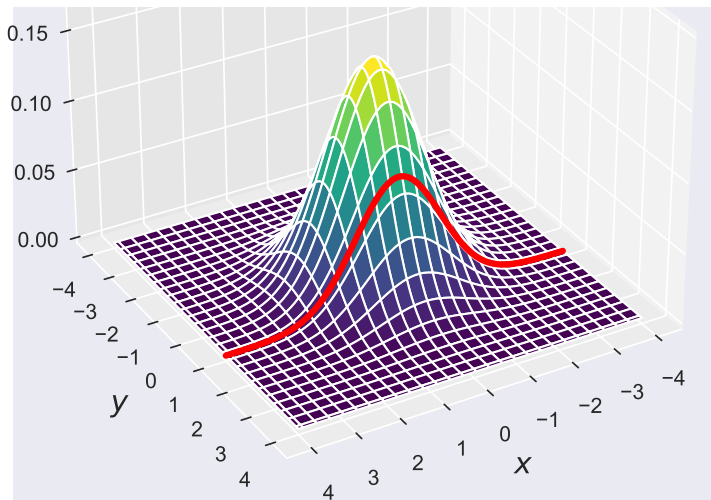
为在 $X = x$ 条件下 Y 的条件概率密度。

条件概率密度也是概率密度, 因为

- $f_{X|Y}(x|y) \geq 0$
- $$\int_{-\infty}^{+\infty} f_{X|Y}(x|y) dx = \int_{-\infty}^{+\infty} \frac{f(x, y)}{f_Y(y)} dx = \frac{f_Y(y)}{f_Y(y)} = 1$$

连续型随机变量的条件分布

- 条件分布的含义是将二维分布限制在 $Y = y$ 上, X 的分布。
- 其分布关系与二维分布一致, 但相差一个规范化因子 $f_Y(y)$ 。



举例：连续型随机变量的条件分布

例

设 (X, Y) 的联合概率密度为

$$f(x, y) = \begin{cases} \frac{21}{4}x^2y, & x^2 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

求 $P(Y \geq 1/4|X = 1/2)$ 及 $P(Y \geq 3/4|X = 1/2)$ 。

根据联合概率密度计算 X 边缘概率密度

$$f_X(x) = \int_{x^2}^1 f(x, y) dy = \int_{x^2}^1 \frac{21}{4}x^2y dy = \begin{cases} \frac{21}{8}x^2(1 - x^4), & |x| \leq 1 \\ 0, & \text{其他} \end{cases}$$

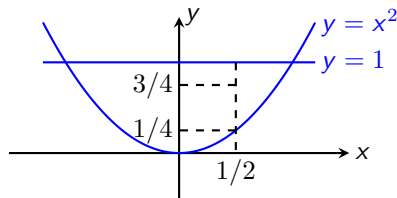
当 $|x| \leq 1$ 时，求条件概率密度

$$f_{Y|X}(y|x) = \begin{cases} \frac{21/4x^2y}{21/8x^2(1-x^4)} = \frac{2y}{1-x^4}, & x^2 \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$

举例：连续型随机变量的条件分布

于是

$$f_{Y|X}(y|\frac{1}{2}) = \begin{cases} \frac{32}{15}y, & \frac{1}{4} \leq y \leq 1 \\ 0, & \text{其他} \end{cases}$$



故

$$P(Y \geq \frac{1}{4} | X = \frac{1}{2}) = 1$$

$$P(Y \geq \frac{3}{4} | X = \frac{1}{2}) = \int_{3/4}^{+\infty} f_{Y|X}(y|\frac{1}{2}) dy = \int_{3/4}^1 \frac{32}{15}y dy = \frac{7}{15}$$

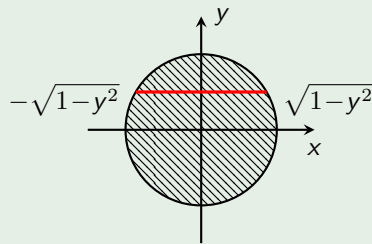
💡 在利用条件密度进行计算时，要特别注意定积分限。

举例：连续型随机变量的条件分布

例

设二维随机变量 (X, Y) 在单位圆上服从均匀分布，即有联合概率密度

$$f(x, y) = \begin{cases} 1/\pi, & x^2 + y^2 \leq 1 \\ 0, & \text{其他} \end{cases}$$



求条件概率密度 $f_{X|Y}(x|y)$ 。

当 $|y| < 1$ 时， Y 的边缘概率密度为

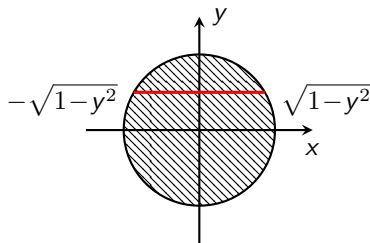
$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx = \frac{2}{\pi} \sqrt{1-y^2}$$

否则， $f_Y(y) = 0$

举例：连续型随机变量的条件分布

故当给定 $Y = y, |y| < 1$ 时, X 的条件概率密度为

$$f_{X|Y}(x|y) = \begin{cases} \frac{1/\pi}{2/\pi\sqrt{1-y^2}} = \frac{1}{2\sqrt{1-y^2}}, & |x| \leq \sqrt{1-y^2} \\ 0, & \text{其他} \end{cases}$$



表明：当已知 $Y = y, |y| < 1$ 时, $X \sim U(-\sqrt{1-y^2}, \sqrt{1-y^2})$ 。

举例：连续型随机变量的条件分布

例

已知数 X 在区间 $(0, 1)$ 上随机取值，当观察到 $X = x \in (0, 1)$ 时，数 Y 在区间 $(x, 1)$ 上随机取值，求 Y 的概率密度 $f_Y(y)$ 。

由题意知 X 具有概率密度

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

又知，当给定条件 $X = x \in (0, 1)$ 时， Y 的条件分布为区间 $(x, 1)$ 上的均匀分布，即

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

其中 $0 < x < 1$ 固定。

举例：连续型随机变量的条件分布

根据公式可得,

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

所以 Y 的边缘概率密度为

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_0^y \frac{1}{1-x} dx = -\ln(1-y), & 0 < y < 1 \\ 0, & \text{其他} \end{cases}$$

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性**
- 5 随机变量函数的概率分布

随机变量的相互独立性

- 随机事件 A, B 的独立性, 指 $P(AB) = P(A)P(B)$ 。
- 即若事件相互独立, 当且仅当乘积事件的概率等于事件概率的乘积。
- 考虑由随机变量确定的事件: $A = \{X \leq x\}, B = \{Y \leq y\}$ 。
- 这样事件 A, B 的相互独立性就可以扩展到随机变量 X, Y 的相互独立性。

随机变量的相互独立性

定义 (随机变量相互独立)

设二维随机变量 (X, Y) 的联合分布函数为 $F(x, y)$, X 与 Y 的边缘分布函数分别为 $F_X(x), F_Y(y)$, 如果对任意实数 x, y 恒有

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

即 $F(x, y) = F_X(x)F_Y(y)$, 则称**随机变量 X 与 Y 相互独立**。

定理 (离散型随机变量的相互独立性)

对于离散型随机变量 (X, Y) , X 与 Y 相互独立的充要条件为: 对 (X, Y) 的所有可能取值 (x_i, y_j) , 都有

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

即 $p_{ij} = p_{i \cdot} p_{\cdot j}$ 对一切 i, j 成立。

举例：随机变量的相互独立性

例

设二维随机变量 (X, Y) 的联合分布律如右表，若 X, Y 相互独立，试确定未知数 a, b, c 。

Y \ X	X		
	1	2	3
1	0.08	a	0.2
2	b	0.18	c

- 由 $p_{21} = p_{2.} \cdot p_{.1}$ 得 $(a + 0.18)(a + 0.28) = a$ ，故 $a = 0.12$ 或 0.42 。
- 当 $a = 0.12$ 时，由 $p_{11} = p_{1.} \cdot p_{.1}$ 得 $(b + 0.08) \times 0.4 = 0.08$ ，故 $b = 0.12$ 。由 $p_{31} = p_{3.} \cdot p_{.1}$ 得 $(c + 0.2) \times 0.4 = 0.2$ ，故 $c = 0.3$ 。
- 当 $a = 0.42$ 时，可计算出 $b = 6/175, c = 3/35$ 。

随机变量的相互独立性

定理 (连续型随机变量的相互独立性)

设连续型随机变量 (X, Y) 的联合概率密度与边缘概率密度分别为 $f(x, y)$, $f_X(x)$, $f_Y(y)$, 则 X 与 Y 相互独立的充要条件为: 对任意实数 x, y , 下式几乎处处成立

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

例

设二维随机变量 (X, Y) 的联合概率密度如下

$$f(x, y) = \begin{cases} e^{-(x+y)}, & x \geq 0, y \geq 0 \\ 0, & \text{其他} \end{cases}$$

问 X, Y 是否相互独立?

随机变量的相互独立性

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_0^{+\infty} e^{-(x+y)} dy, & x \geq 0 \\ 0, & x < 0 \end{cases} \\ &= \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \end{aligned}$$

同理

$$f_Y(y) = \begin{cases} e^{-y}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

故 $f(x, y) = f_X(x)f_Y(y)$, 即 X, Y 相互独立。

可见, 若 (X, Y) 为独立的连续型随机变量, 则其联合概率密度 $f(x, y)$ 可以表成分离形式 $f(x, y) = f_X(x)f_Y(y)$ 。

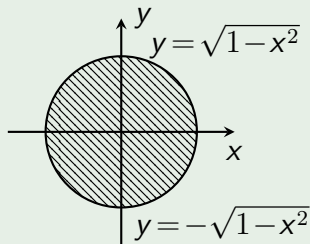
举例：随机变量的相互独立性

例

设 (X, Y) 的概率密度如下：

$$f(x, y) = \begin{cases} 1/\pi, & x^2 + y^2 \leq 1 \\ 0, & \text{其他} \end{cases}$$

求 $f_X(x), f_Y(y)$ ，并问二者是否独立？



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

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx = \frac{2\sqrt{1-y^2}}{\pi}, y \in [-1, 1]$$

故 $f(x, y) \neq f_X(x)f_Y(y)$ ，两者不独立。

随机变量相互独立性的推广

定义 (多个随机变量相互独立)

设 n 维随机变量 (X_1, \dots, X_n) 的联合分布函数为 $F(x_1, \dots, x_n)$, X_i 的边缘分布函数为 $F_{X_i}(x_i)$ 。若 $\forall (x_1, \dots, x_n) \in \mathbb{R}^n$, 有

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$$

则称 X_1, \dots, X_n 相互独立。

定义 (两组随机变量相互独立)

设 $(X_1, \dots, X_m) \sim F_1(x_1, \dots, x_m)$, $(Y_1, \dots, Y_n) \sim F_2(y_1, \dots, y_n)$, $(X_1, \dots, X_m; Y_1, \dots, Y_n) \sim F(x_1, \dots, x_m; y_1, \dots, y_n)$ 。若 $\forall (x_1, \dots, x_m) \in \mathbb{R}^m$, $\forall (y_1, \dots, y_n) \in \mathbb{R}^n$, 有

$$F(x_1, \dots, x_m; y_1, \dots, y_n) = F_1(x_1, \dots, x_m) F_2(y_1, \dots, y_n)$$

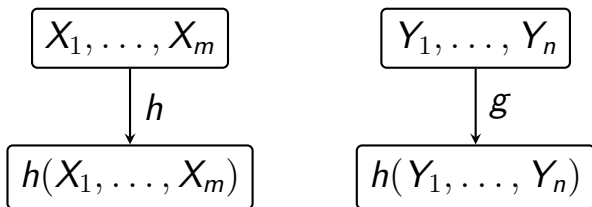
则称 (X_1, \dots, X_m) 与 (Y_1, \dots, Y_n) 相互独立。

随机变量相互独立性的推广

定理

设 (X_1, \dots, X_m) 与 (Y_1, \dots, Y_n) 相互独立, 若 h, g 是两个多元连续函数, 则 $h(X_1, \dots, X_m)$ 与 $g(Y_1, \dots, Y_n)$ 也相互独立。

两堆独立数据



处理后仍然独立

随机变量相互独立性的推广

定理

若随机变量 X_1, \dots, X_n 相互独立，把它们分为不相交的 k 个组，每个组中所有变量由一个连续函数复合而生成一个新的随机变量，则这 k 个新的随机变量仍相互独立。

例

设 X_1, X_2, X_3, X_4, X_5 相互独立，而 $Y_1 = \cos(X_1 + X_3)$, $Y_2 = \sin X_2$, $Y_3 = X_4 + 10$ ，由上述定理知 Y_1, Y_2, Y_3 相互独立。

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布
 - 一维随机变量函数的概率分布
 - 二维随机变量函数的概率分布

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布
 - 一维随机变量函数的概率分布
 - 二维随机变量函数的概率分布

随机变量函数的概率分布

- 有时我们关心的随机变量不能直接观测，而是某个可观测随机变量的函数。
- 一些随机变量的分布很难获得，但与其相关的另一些随机变量的分布却往往容易获得。

例 (电功率)

设某供电线路上电流值 X 为一随机变量，其分布函数为 $F_X(x)$ 。若线路上有一电阻 R ，试求 R 上的电功率 Y 的分布函数 $F_Y(y)$ 。

显然有 $Y = RX^2$ ，且 Y 取值为非负的，根据分布函数定义，有

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(RX^2 \leq y) = P\left(-\sqrt{\frac{y}{R}} \leq X \leq \sqrt{\frac{y}{R}}\right) \\ &= F_X\left(\sqrt{\frac{y}{R}}\right) - F_X\left(-\sqrt{\frac{y}{R}}\right) \end{aligned}$$

随机变量函数的概率分布

定义 (随机变量函数)

记随机变量 X 的一切可能取值集合为 D , 设 $g(x)$ 是定义在 D 上的连续函数或分段单调实函数, 若对于 X 的每一个可能值 $x \in D$, 随机变量 Y 相应地取 $y = g(x)$, 则称随机变量 Y 为随机变量 X 的函数, 记为 $Y = g(X)$ 。

- 类似地, 可以定义 n 维随机变量 (X_1, \dots, X_n) 的函数 $Y = g(X_1, \dots, X_n)$ 。
- 一般地, 我们需要研究问题: 设 X 为一随机变量, 分布已知。 $Y = g(X)$, 其中 g 为一确定的实函数, 求 Y 的分布。
- 分别讨论离散型随机变量与连续型随机变量两种情况。

一维离散型随机变量函数的概率分布

求一维离散型随机变量函数的分布律

设随机变量 X 的分布律为：

X	x_1	x_2	\cdots	x_i	\cdots
P	p_1	p_2	\cdots	p_i	\cdots

求 $Y = g(X)$ 的分布律。

- 计算 Y 的所有可能取值 $g(x_i)$ ，及其对应的概率 p_i
- 合并所有相同的 $g(x_i)$ ，并将对应的 p_i 相加

例

设离散型随机变量 X 的分布律为

$$P(X = n) = \frac{1}{2^n} \quad n = 1, 2, \dots$$

求随机变量 $Y = \cos(\frac{\pi}{2}X)$ 的分布律。

一维离散型随机变量函数的概率分布

$$\cos\left(\frac{n\pi}{2}\right) = \begin{cases} -1, & n = 2(2k-1) \\ 0, & n = 2k \\ 1, & n = 4k \end{cases} \quad k = 1, 2, \dots$$

所以 $Y = \cos(\frac{\pi}{2}X)$ 不同的可能取值为 $-1, 0, 1$ 。

由于 X 取 $2, 6, 10, \dots$ 时, 对应的 Y 取 -1 , 所以

$$P(Y = -1) = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^6 + \left(\frac{1}{2}\right)^{10} + \dots = \frac{1/4}{1 - 1/16} = \frac{4}{15}$$

同理可得

$$P(Y = 0) = \left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \dots = \frac{1/2}{1 - 1/4} = \frac{2}{3}$$

$$P(Y = 1) = \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^8 + \left(\frac{1}{2}\right)^{12} + \dots = \frac{1/16}{1 - 1/16} = \frac{1}{15}$$

举例：一维离散型随机变量函数的概率分布

例

设 X 的分布律为

X	0	1/4	1/2	3/4	1
P	0.1	0.2	0.4	0.2	0.1

令 $Y = \sin^2(\pi X)$, 求 Y 的分布律。

- 计算不同 X 值对应的 Y 值

X	0	1/4	1/2	3/4	1
Y	0	1/2	1	1/2	0
P	0.1	0.2	0.4	0.2	0.1

- 合并后得 Y 的分布律

Y	0	1/2	1
P	0.2	0.4	0.4

举例：一维离散型随机变量函数的概率分布

有时 X 是连续型的，但 Y 仍有可能是离散的。

例 (儿童智商)

设儿童智商 $X \sim N(100, 100)$ ，将儿童按智商分为 3 类，类标号 Y 规定如下：

$$Y = \begin{cases} 1, & X > 110 \\ 0, & 90 < X \leq 110 \\ -1, & X \leq 90 \end{cases}$$

求 Y 的分布律。

根据 X 的分布，查正态分布表可得：

Y	-1	0	1
P	0.16	0.68	0.16

一维连续型随机变量函数的概率分布

求一维连续型随机变量函数的概率密度

设 X 为连续型随机变量, 具有概率密度 $f_X(x)$, 又设 $Y = g(X)$ 亦为连续型随机变量, 求其概率密度 $f_Y(y)$ 。

- 从分布函数着手

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{x: g(x) \leq y} f_X(x) dx$$

其中积分区域 $\{x: g(x) \leq y\}$ 表示满足 $g(x) \leq y$ 的 x 的点集。

- 进一步化简的关键是寻求上述集合的 x 的显式表达。
- 对分布函数求导, 得到概率密度 $f_Y(y)$ 。

举例：一维连续型随机变量函数的概率分布

例 (电功率, 续)

设某供电线路上电流值 X 为一随机变量, 设 X 有概率密度 $f_X(x)$ 。又为方便计算, 令 $R = 1$, 求 $Y = X^2$ 的概率密度。

由前例已知

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

对分布函数求导得

$$f_Y(y) = \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y}))$$

或者直接对积分区间求导, 即对下式求导

$$F_Y(y) = P(X^2 \leq y) = \int_{\{x: x^2 \leq y\}} f_X(x) dx = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x) dx$$

举例：一维连续型随机变量函数的概率分布

例 (正态变量的平方)

设 $X \sim N(0, 1)$, 求 $Y = X^2$ 的概率密度。

- 当 $y \leq 0$ 时, $f_Y(y) = 0$
- 当 $y \geq 0$ 时,

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y})) \\ &= \frac{1}{\sqrt{y}}f_X(\sqrt{y}) \\ &= \frac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2} \end{aligned}$$

一维连续型随机变量函数的概率分布

定理

设连续型随机变量 X 的概率密度为 $f_X(x)$ 。若 $y = g(x)$ 为定义在实数域上的严格单调可导函数，则 $Y = g(X)$ 也是一个连续型随机变量，且其概率密度为

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|, & \alpha < y < \beta \\ 0, & \text{其他} \end{cases}$$

其中 $\alpha = \min_x g(x)$, $\beta = \max_x g(x)$ 。

一维连续型随机变量函数的概率分布

- 当 $g(x)$ 严格递增时, 则 $g^{-1}(y)$ 也严格递增, 此时
$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$
- 对上式求导, 得

$$f_Y(y) = f_X(g^{-1}(y)) \frac{dg^{-1}(y)}{dy} \quad (1)$$

- 当 $g(x)$ 严格递减时, 则 $g^{-1}(y)$ 也严格递减, 此时
$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$
- 对上式求导, 得

$$f_Y(y) = -f_X(g^{-1}(y)) \frac{dg^{-1}(y)}{dy} = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| \quad (2)$$

- 综合(1)和(2), 得

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|$$

举例：一维连续型随机变量函数的概率分布

例 (线性函数)

设 X 的概率密度为 $f_X(x)$, $Y = aX + b$, $a \neq 0$, 求 $f_Y(y)$ 。

当 Y 为 X 的线性函数时, 通常有

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

特别地, 若 $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, $a = \sigma$, $b = \mu$, 则 Y 的概率密度为

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

即 $Y \sim N(\mu, \sigma^2)$

- 正态分布的线性函数仍为正态分布。
- 任何正态分布都可表示为标准正态分布的线性函数。

举例：一维连续型随机变量函数的概率分布

例 (对数正态分布)

设 $X \sim N(\mu, \sigma^2)$, 令 $Y = e^X$, 求 Y 的概率密度 $f_Y(y)$ 。

因为 $y = e^x$ 满足定理中的条件, 当 $y > 0$ 时, $y = e^x$ 的反函数为 $x = \ln y$, 有

$$\begin{aligned} f_Y(y) &= \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} \left| \frac{1}{y} \right|, & y > 0 \\ 0, & y \leq 0 \end{cases} \\ &= \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} y^{-1} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}}, & y > 0 \\ 0, & y \leq 0 \end{cases} \end{aligned}$$

称具有上述概率密度的随机变量 Y 服从**对数正态分布**。

目录

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布
 - 一维随机变量函数的概率分布
 - 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

例

设二维随机变量 (X, Y) 的联合分布律如右表, 求 $Z = X + Y$; $Z = XY$; $Z = \max(X, Y)$; $Z = \min(X, Y)$ 的分布律。

$X \backslash Y$	0	1	2
-1	0.2	0.3	0.1
2	0.1	0.1	0.2

(X, Y)	$(-1, 0)$	$(-1, 1)$	$(-1, 2)$	$(2, 0)$	$(2, 1)$	$(2, 2)$
P	0.2	0.3	0.1	0.1	0.1	0.2
$X + Y$	-1	0	1	2	3	4
XY	0	-1	-2	0	2	4
$\max(X, Y)$	0	1	2	2	2	2
$\min(X, Y)$	-1	-1	-1	0	1	2

二维离散型随机变量函数的概率分布

故随机变量函数的分布律分别为：

$X + Y$	-1	0	1	2	3	4
P	0.2	0.3	0.1	0.1	0.1	0.2

XY	-2	-1	0	2	4
P	0.1	0.3	0.3	0.1	0.2

$\max(X, Y)$	0	1	2
P	0.2	0.3	0.5

$\min(X, Y)$	-1	0	1	2
P	0.6	0.1	0.1	0.2

举例：二维离散型随机变量函数的概率分布

例 (二项分布相加)

设随机变量 X 与 Y 相互独立, 且 $X \sim B(n_1, p)$, $Y \sim B(n_2, p)$, 求 $Z = X + Y$ 的分布律。

Z 的可能取值为 $0, 1, \dots, n_1 + n_2$ 。当 $0 \leq k \leq n_1 + n_2$ 时, 有

$$\begin{aligned} P(Z = k) &= P(X + Y = k) = \sum_{i=0}^k P(X = i, Y = k - i) \\ &= \sum_{i=0}^k \binom{n_1}{i} \binom{n_2}{k-i} p^k (1-p)^{n_1+n_2-k} \\ &= \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k} \end{aligned}$$

这表明 $Z \sim B(n_1 + n_2, p)$, 即二项分布具有可加性。

二维连续型随机变量函数的概率分布

求二维连续型随机变量函数的概率密度

设 (X, Y) 是二维连续型随机变量, $f(x, y)$ 是其概率密度, 又 $Z = g(X, Y)$ 是 X, Y 的连续函数, 且 Z 是连续型随机变量, 求 Z 的概率密度 $f_Z(z)$ 。

- 先求分布函数

$$F_Z(z) = P(Z \leq z) = P(g(X, Y) \leq z) = \iint_{g(x, y) \leq z} f(x, y) dx dy$$

- 再对分布函数求导, 得到概率密度 $f_Z(z)$ 。
- 计算的关键是确定上式的积分区域。

举例：二维连续型随机变量函数的概率分布

例 (瑞利分布)

设随机变量 X 与 Y 相互独立且 $X \sim N(0, \sigma^2)$, $Y \sim N(0, \sigma^2)$, 求 $Z = \sqrt{X^2 + Y^2}$ 的概率密度。

易知 (X, Y) 的联合概率密度为 $f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$ 。 $z > 0$ 时

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(\sqrt{X^2 + Y^2} \leq z) = P(X^2 + Y^2 \leq z^2) \\ &= \iint_{x^2+y^2 \leq z^2} \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} dx dy = \int_0^{2\pi} d\theta \int_0^z \frac{1}{2\pi\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}} \rho d\rho \\ &= \frac{1}{\sigma^2} \int_0^z e^{-\frac{\rho^2}{2\sigma^2}} \rho d\rho \end{aligned}$$

从而得到 $f_Z(z) = F'_Z(z) = \begin{cases} \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}, & z > 0 \\ 0, & \text{其他} \end{cases}$, 称为**瑞利分布**。

随机变量和的概率密度

求随机变量和的概率密度

设 X, Y 为连续型随机变量, 有联合概率密度 $f(x, y)$ 。设 $Z = X + Y$, 求 Z 的概率密度。

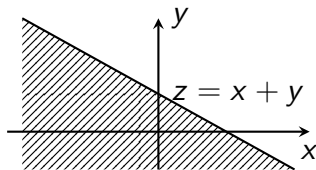
$$F_Z(z) = P(Z \leq z) = P(X + Y \leq z)$$

$$= \iint_{x+y \leq z} f(x, y) dx dy$$

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

令 $t = y + x$, 有

$$F_Z(z) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^z f(x, t-x) dt = \int_{-\infty}^z dt \int_{-\infty}^{+\infty} f(x, t-x) dx$$



随机变量和的概率密度

对 z 求导, 从而得到 Z 的概率密度为

$$f_Z(z) = F'_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx$$

利用 X 与 Y 的对称性, 也可以得到

$$f_Z(z) = \int_{-\infty}^{+\infty} f(z-y, y) dy$$

特别地, 当 X 与 Y 相互独立时:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$$

称为**卷积公式**。

举例：随机变量和的概率密度

例 (独立正态变量的和)

设 X, Y 均服从 $N(0, 1)$ 且相互独立, 求 $Z = X + Y$ 的概率密度。

由卷积公式, 有

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} dx \\ &= \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} e^{-(x-\frac{z}{2})^2} dx \stackrel{t=x-\frac{z}{2}}{=} \frac{1}{2\pi} e^{-\frac{z^2}{4}} \int_{-\infty}^{+\infty} e^{-t^2} dt \\ &= \frac{1}{2\pi} e^{-\frac{z^2}{4}} \sqrt{\pi} = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{z^2}{2(\sqrt{2})^2}} \end{aligned}$$

从而得到 $Z \sim N(0, 2)$

独立正态分布随机变量和的概率密度

- 若 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ 且相互独立, 则

$$Z = X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

- 若 $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \dots, n$ 且相互独立, 则

$$\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

- 利用前面关于正态线性函数的结论, 若 $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \dots, n$ 且相互独立, $a_i > 0$ 为常数, 则

$$\sum_{i=1}^n a_i X_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right)$$

举例：随机变量和的概率密度

例 (均匀分布与指数分布的和)

设随机变量 X, Y 相互独立, 且 $X \sim U(0, 2), Y \sim \exp(3)$, 求 $Z = X + Y$ 的概率密度。

X, Y 的概率密度分别为

$$f_X(x) = \begin{cases} \frac{1}{2}, & 0 < x < 2 \\ 0, & \text{其他} \end{cases} \quad f_Y(y) = \begin{cases} 3e^{-3y}, & y > 0 \\ 0, & \text{其他} \end{cases}$$

由卷积公式计算 Z 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_0^2 \frac{1}{2} f_Y(z-x) dx$$

$$\stackrel{t=z-x}{=} \frac{1}{2} \int_{z-2}^z f_Y(t) dt = \begin{cases} 0, & z < 0, \\ (1 - e^{-3z})/2, & 0 \leq z < 2, \\ (e^{-3(z-2)} - e^{-3z})/2, & z \geq 2 \end{cases}$$

举例：随机变量和的概率密度

例 (指数分布随机变量的和)

某系统的寿命由一个关键部件决定，该部件寿命 $X \sim \exp(\theta)$ ，另有一相同备用件，问系统的工作寿命服从什么分布？

设另一个备用件的寿命为 $Y \sim \exp(\theta)$ 且与 X 独立，故系统工作寿命为 $Z = X + Y$ 。由卷积公式计算 Z 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_0^z \frac{1}{\theta^2} e^{-\frac{x+z-x}{\theta}} dx = \frac{z}{\theta^2} e^{-\frac{z}{\theta}}$$

设 $X_i \sim \exp(\theta)$, $i = 1, \dots, n$ 且相互独立，则 $Z = \sum_{i=1}^n X_i$ 的概率密度为

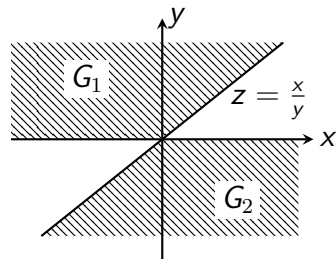
$$f_Z(z) = \begin{cases} \frac{z^{n-1}}{\theta^n (n-1)!} e^{-\frac{z}{\theta}}, & z > 0, \\ 0, & z \leq 0 \end{cases}$$

随机变量商的概率密度

求随机变量商的概率密度

设随机变量 X, Y 的联合概率密度为 $f(x, y)$, 求 $Z = X/Y$ 的概率密度。

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(X/Y \leq z) \\ &= \iint_{G_1} f(x, y) dx dy + \iint_{G_2} f(x, y) dx dy \\ \iint_{G_1} f(x, y) dx dy &= \int_0^{+\infty} dy \int_{-\infty}^{yz} f(x, y) dx \end{aligned}$$



令 $t = x/y$, 注意到 y 是固定的且 $y > 0$, 得

$$\int_{-\infty}^{yz} f(x, y) dx = \int_{-\infty}^z yf(yt, y) dt$$

随机变量商的概率密度

$$\begin{aligned}\iint_{G_1} f(x, y) dx dy &= \int_0^{+\infty} dy \int_{-\infty}^z yf(yt, y) dt \\ &= \int_{-\infty}^z \int_0^{+\infty} yf(yt, y) dy dt\end{aligned}$$

同理可得

$$\iint_{G_2} f(x, y) dx dy = - \int_{-\infty}^z \int_{-\infty}^0 yf(yt, y) dy dt$$

因此

$$\begin{aligned}F_Z(z) &= \int_{-\infty}^z \left[\int_0^{+\infty} yf(yt, y) dy - \int_{-\infty}^0 yf(yt, y) dy \right] dt \\ &= \int_{-\infty}^z \int_{-\infty}^{+\infty} |y| f(yt, y) dy dt\end{aligned}$$

随机变量商的概率密度

从而得到 $Z = X/Y$ 的概率密度

$$f_Z(z) = F'_Z(z) = \int_{-\infty}^{+\infty} |y| f(yz, y) dy$$

特别地, 当 X 与 Y 相互独立时, $Z = X/Y$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} |y| f_X(yz) f_Y(y) dy$$

举例：随机变量商的概率密度

例 (指数分布随机变量的商)

设 X, Y 独立同分布 $\exp(1)$, 求 $Z = X/Y$ 的概率分布。

由随机变量商的概率密度公式, 得

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} |y| f_X(yz) f_Y(y) dy \\ &= \int_0^{+\infty} e^{-(y+yz)} y dy = \frac{1}{1+z} \int_0^{+\infty} e^{-(y+yz)} dy \\ &= \frac{1}{(1+z)^2}, \quad z > 0 \end{aligned}$$

所以

$$f_Z(z) = \begin{cases} (1+z)^{-2}, & z > 0, \\ 0, & z \leq 0 \end{cases}$$

举例：随机变量商的概率密度

例

设 X, Y 服从 $N(0, 1)$ 且相互独立，求 $T = Y/\sqrt{X^2}$ 的概率密度。

由之前的结果知 $U = X^2$ 的概率密度为

$$f_U(u) = \begin{cases} \frac{1}{\sqrt{2\pi}} u^{-1/2} e^{-u/2}, & u > 0, \\ 0, & u \leq 0 \end{cases}$$

令 $V = \sqrt{U}$ 以及之前的定理，可得 V 的概率密度为

$$f_V(v) = \begin{cases} \frac{2}{\sqrt{2\pi}} e^{-v^2/2}, & v > 0, \\ 0, & v \leq 0 \end{cases}$$

于是得

$$f_T(t) = \int_{-\infty}^{+\infty} |v| f_Y(vt) f_V(v) dv = \frac{1}{\pi(1+t^2)}$$

随机变量取最大值与最小值的概率密度

- 有些随机变量函数不是可微函数，不能用前面做变换的方法，如 $\max(X, Y), \min(X, Y)$ 等。

随机变量取最大最小

设二维随机变量 (X, Y) 的联合分布函数及边缘分布函数分别为 $F(x, y), F_X(x)$ 和 $F_Y(y)$ 。设 $M = \max(X, Y), N = \min(X, Y)$ ，求 M, N 的分布函数。

因为对任意 $z \in \mathbb{R}$ 有

$$\{\max(X, Y) \leq z\} = \{X \leq z, Y \leq z\}$$

$$\{\min(X, Y) > z\} = \{X > z, Y > z\}$$

随机变量取最大值与最小值的概率密度

由分布函数定义，得到

$$\begin{aligned}F_M(z) &= P(M \leq z) = P(\max(X, Y) \leq z) \\&= P(X \leq z, Y \leq z) = F(z, z)\end{aligned}$$

$$\begin{aligned}F_N(z) &= P(N \leq z) = 1 - P(N > z) \\&= 1 - P(\min(X, Y) > z) = 1 - P(X > z, Y > z) \\&= 1 - (1 - F_X(z) - F_Y(z) + F(z, z)) \\&= F_X(z) + F_Y(z) - F(z, z)\end{aligned}$$

若 X 与 Y 相互独立，则有

$$\begin{aligned}F_M(z) &= F_X(z)F_Y(z) \\F_N(z) &= F_X(z) + F_Y(z) - F_X(z)F_Y(z) \\&= 1 - (1 - F_X(z))(1 - F_Y(z))\end{aligned}$$

随机变量取最大值与最小值的概率密度

- 设 X_1, \dots, X_n 是 n 个相互独立的随机变量, $F_{X_i}(x_i)$ 是 X_i 的分布函数。设 $M = \max(X_1, \dots, X_n)$, $N = \min(X_1, \dots, X_n)$, 则有

$$F_M(z) = F_{X_1}(z) \cdots F_{X_n}(z)$$

$$F_N(z) = 1 - (1 - F_{X_1}(z)) \cdots (1 - F_{X_n}(z))$$

- 特别地, 当 X_1, \dots, X_n 独立同分布且分布函数为 $F_X(x)$ 时, 有

$$F_M(z) = F_X(z)^n, \quad F_N(z) = 1 - (1 - F_X(z))^n$$

- 对上式求导, 得概率密度

$$f_M(z) = nF_X(z)^{n-1}f_X(z), \quad f_N(z) = n(1 - F_X(z))^{n-1}f_X(z)$$

举例：随机变量取最大值与最小值的概率密度

例

设 X_1, \dots, X_n 相互独立，都服从 $[a, b]$ 上的均匀分布，求 $M = \max(X_1, \dots, X_n)$ 与 $N = \min(X_1, \dots, X_n)$ 的概率密度。

M 的分布函数为

$$F_M(z) = F_X(z)^n = \begin{cases} 0, & z < a, \\ \left(\frac{z-a}{b-a}\right)^n, & a \leq z < b, \\ 1, & z \geq b \end{cases}$$

故 M 的概率密度为

$$f_M(z) = \begin{cases} \frac{n(z-a)^{n-1}}{(b-a)^n}, & a < z < b, \\ 0, & \text{其他} \end{cases}$$

举例：随机变量取最大值与最小值的概率密度

N 的分布函数为

$$F_N(z) = 1 - (1 - F_X(z))^n = \begin{cases} 0, & z < a, \\ 1 - \left(\frac{z-a}{b-a}\right)^n, & a \leq z < b, \\ 1, & z \geq b \end{cases}$$

故 N 的概率密度为

$$f_N(z) = \begin{cases} \frac{n(b-z)^{n-1}}{(b-a)^n}, & a < z < b, \\ 0, & \text{其他} \end{cases}$$

举例：随机变量取最大值与最小值的概率密度

例 (均匀分布)

设随机变量 $X_i \sim U(0, \theta), i = 1, \dots, n$ 且相互独立, 求随机变量 $M = \max(X_1, \dots, X_n)$ 的概率密度。

X_i 的分布函数和概率密度分别为

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{\theta}, & 0 \leq x \leq \theta, \\ 1, & x > \theta \end{cases} \quad f_X(x) = \begin{cases} \frac{1}{\theta}, & 0 \leq x \leq \theta, \\ 0, & \text{其他} \end{cases}$$

故 M 的概率密度为

$$f_M(x) = nF_X(x)^{n-1}f_X(x) = \begin{cases} \frac{nx^{n-1}}{\theta}, & 0 \leq x \leq \theta, \\ 0, & \text{其他} \end{cases}$$

小结

- 1 一维随机变量
- 2 二维随机变量
- 3 二维随机变量的条件分布
- 4 随机变量的相互独立性
- 5 随机变量函数的概率分布