

2015考研数学概率论零基础入门讲义

主讲：张宇

张宇：新东方在线名师，博士，全国著名考研数学辅导专家，教育部“国家精品课程建设骨干教师”，全国畅销书《高等数学 18 讲》、《考研数学题源探析经典 1000 题》作者，高等教育出版社《全国硕士研究生入学统一考试数学考试参考书（大纲解析）》编者之一，2007 年斯洛文尼亚全球可持续发展大会受邀专家（发表 15 分钟主旨演讲）。首创“题源教学法”，对考研数学的知识结构和体系有全新的解读，对考研数学的命题与复习思路有极强的把握和预测能力，让学生轻松高效夺取高分。

欢迎使用新东方在线电子教材



目 录

第一讲	随机事件与概率.....	1
第二讲	一维随机变量及其概率分布.....	7
第三讲	随机变量的数字特征.....	12

【注】(1) 数二的考生不需要学习这部分内容。

(2) 老师没有完全按照讲义的顺序讲课，而是打乱了顺序，重新整合授课体系，但是老师所讲的内容多数是包含在讲义中的，讲义中没有的内容需要同学们自己做笔记。

第一讲 随机事件与概率

一、从古典概型讲起

1. 随机试验与随机事件

称一个试验为随机试验，如果满足：

- (1) 同条件下可重复
- (2) 所有试验结果明确可知且不止一个
- (3) 试验前不知哪个结果会发生

【注】①在一次试验中可能出现，也可能不出现的结果称为随机事件，简称为事件，并用大写字母 A, B, C 等表示，为讨论需要，将每次试验一定发生的事件称为必然事件，记为 Ω 。每次试验一定不发生的事件称为不可能事件，记为 ϕ 。

②随机试验每一最简单、最基本的结果称为基本事件或样本点，记为 ω_i 。

2. 古典概率

称随机试验(随机现象)的概率模型为古典概型，如果其基本事件空间(样本空间)满足：

- (1) 只有有限个基本事件(样本点)；
- (2) 每个基本事件(样本点)发生的可能性都一样。

【注】①等可能：对于可能结果： $\omega_1, \omega_2, \dots, \omega_n$ ，我们找不到任何理由认为其中某一结果 ω_i 更易发生，则只好（客观）认为所有结果在试验中发生的可能性一样。

②如果古典概型的基本事件总数为 n ，事件 A 包含 k 个基本事件，即有利于 A 的基本事件 k 个。则 A 的概率定义为

$$P(A) = \frac{k}{n} = \frac{\text{事件} A \text{ 所含基本事件的个数}}{\text{基本事件总数}}$$

由上式计算的概率称为 A 的古典概率。

3. 计数方法

(1) 穷举法: 样本点总数不大时

(2) 集合对方法:

①加法原理: 完成一件事, 有 n 类方法, 第一类方法中有 m_1 种方法, 第二类方法中有 m_2 种方法, ……第 n 类方法中有 m_n 种方法, 则完成此事共有 $m_1 + m_2 + \cdots + m_n$ 种办法.

②乘法原理: 完成一件事, 有 n 个步骤, 第一步中有 m_1 种方法, 第二步中有 m_2 种方法, ……第 n 步中有 m_n 种方法, 则完成此事共有 $m_1 \cdot m_2 \cdots m_n$ 种办法.

③排列: 从 n 个不同元素中取出 m ($m \leq n$) 个元素, 按照一定的顺序排成一列, 叫排列. 所有排列的个数叫做排列数, 记作 $P_n^m = n(n-1) \cdots (n-m+1) = \frac{n!}{(n-m)!}$. 当 $m = n$ 时,

$P_n^m = P_n^n = n!$, 称为全排列.

④组合: 从 n 个不同元素中取出 m ($m \leq n$) 个元素并成一组, 叫组合. 所有组合的个数叫做组合数, 记作 $C_n^m = \frac{P_n^m}{m!}$, 也有 $P_n^m = C_n^m \cdot m!$.

(3) 用对立事件思想

4. 例题分析

【例 1】从 0 到 9 这十个数字中任取 3 个不同的数字, 求

- (1) 三个数中不含 0 和 5 的概率
- (2) 三个数中不含 0 或 5 的概率
- (3) 三个数中含 0, 但不含 5 的概率

【例 2】假设袋中有 5 个球, 3 白球 2 黑球, 求

- (1) 先后有放回取 2 球, 至少有一白球的概率;
- (2) 先后无放回取 2 球, 至少有一白球的概率;
- (3) 任取 2 球, 至少有一白球的概率.

【例 3】假设袋中有 100 个球, 40 个白球, 60 个黑球

- (1) 先后无放回取 20 个, 求取到 15 个白球 5 个黑球的概率;
- (2) 先后无放回取 20 个, 求第 20 次取到白球的概率;
- (3) 先后有放回取 20 个, 求取到 15 个白球 5 个黑球的概率;

(4) 先后有放回取 20 个, 求第 20 次取到白球的概率.

二、几何概型

1. 引例 天上掉馅饼

2. 几何概型的定义

如果(1)样本空间(基本事件空间) Ω 是一个可度量的几何区域;(2)每个样本点(基本事件)发生的可能性都一样, 即样本点落入 Ω 的某一可度量的子区域 A 的可能性大小与 A 的几何度量成正比, 而与 A 的位置及形状无关, 我们就称这样的随机试验的概率模型为几何概型,

在几何概型随机试验中, 如果 S_A 是样本空间 Ω 一个可度量的子区域, 则事件 A ="样本点落入区域 S_A "的概率定义为

$$P(A) = \frac{S_A \text{ 的几何测度}}{\Omega \text{ 的几何测度}}$$

由上式计算的概率称为 A 的几何概率

【评注】 基本事件有限、等可能的随机试验为古典概型; 基本事件无限、等可能的随机试验为几何概型.

3. 例题分析

【例 1】 君子有约, 上午 9:00-10:00 到新东方大厦门口见面, 先到者等 20 分钟即离开, 求甲、乙两人相遇的概率.

【例 2】 在区间 (0,1) 中随机取两个数, 则两数之和小于 $\frac{6}{5}$ 的概率为_____.

三、重要公式求概率

1. 重要公式总结

(1) 求逆公式 $P(\bar{A}) = 1 - P(A)$.

(2) 减法公式 $P(A - B) = P(A) - P(AB)$.

(3) 加法公式 $P(A \cup B) = P(A) + P(B) - P(AB)$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

【注】 ① 设 A_1, A_2, \dots, A_n 是两两互不相容的事件, 则 $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$

② 若 A_1, A_2, \dots, A_n 相互独立, 则 $P(\bigcup_{i=1}^n A_i) = 1 - \prod_{i=1}^n [1 - P(A_i)]$

(4) 条件概率公式 设 A, B 为任意两个事件, 若 $P(A) > 0$, 我们称在已知事件 A 发生的条件下, 事件 B 发生的概率为条件概率, 记为 $P(B | A)$, 并定义

$$P(B|A) = \frac{P(AB)}{P(A)} \quad (P(A) > 0).$$

【注】(1) 条件概率 $P(\cdot|A)$ 是概率，概率的一切性质和重要结果对条件概率都适用，例如：

$$P(\bar{B}|A) = 1 - P(B|A),$$

$$P(B - C|A) = 1 - P(B|A) - P(BC|A) > 0, \text{ 等等.}$$

(2) 条件概率就是附加一定的条件之下所计算的概率. 当说到“条件概率”时，总是指另外附加的条件，其形式可归结为“已知某事件发生了”.

(5) 乘法公式

如果 $P(A) > 0$ ，则 $P(AB) = P(A)P(B|A)$. 一般地，如果 $P(A_1 \dots A_{n-1}) > 0$ ，则

$$P(A_1 A_2 \dots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \dots P(A_n|A_1 \dots A_{n-1})$$

【注】 A_i 先于 A_{i+1} 发生时用此公式.

(6) 全概率公式 (全集分解思想)

如果 $\bigcup_{i=1}^n A_i = \Omega, A_i A_j = \emptyset (i \neq j), P(A_i) > 0$ ，则对任一事件 B ，有

$$B = \bigcup_{i=1}^n A_i B, P(B) = \sum_{i=1}^n P(A_i)P(B|A_i).$$

(7) 贝叶斯(Bayes)公式(逆概公式)

如果 $\bigcup_{i=1}^n A_i = \Omega, A_i A_j = \emptyset (i \neq j), P(A_i) > 0$ ，则对任一事件 B ，只要 $P(B) > 0$ ，有

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)} \quad (i=1, 2, \dots, n)$$

【注】① 要注意 $P(AB)$ 与 $P(B|A)$ 的区别：

$P(AB)$ 是在样本空间为 Ω 时， A 与 B 同时发生的可能性，而 $P(B|A)$ 则是表示在 A 已经发生的条件下， B 发生的可能性，此时样本空间已由 Ω 缩减为 A ，只要题目中有前提条件：“在 A 发生的条件下”或“已知 A 发生”等等，均要考虑条件概率.

② 全概率公式是用于计算某个“结果” B 发生的可能性大小. 如果一个结果 B 的发生总是与某些前提条件(或原因、因素或前一阶段结果) A_i 相联系，那么在计算 $P(B)$ 时，我们总是将 B 对 A_i 作分解： $B = \bigcup_i A_i B$ ，应用全概率公式计算 $P(B)$. 如果在 B 发生的条件下探求导致这

一结果的

各种“原因” A_i 发生的可能性大小 $P(A_i | B)$ ，则要应用 Bayes 公式.

2. 随机事件相互独立与独立试验序列概型

(1) 独立性定义

描述性定义(直观性定义) 设 A, B 为两个事件，如果其中任何一个事件发生的概率不受另外一个事件发生与否的影响，则称事件 A 与 B 相互独立. 设 A_1, A_2, \dots, A_n 是 n 个事件，如果其中任何一个或几个事件发生的概率都不受其余的某一个或某几个事件发生与否的影响，则称事件 A_1, A_2, \dots, A_n 相互独立.

数学定义 设 A, B 为事件，如果 $P(AB) = P(A)P(B)$ ，则称事件 A 与 B 相互独立，简称为 A 与 B 独立.

设 A_1, A_2, \dots, A_n 为 n 个事件，如果对其中任意有限个事件 $A_{i1}, A_{i2}, \dots, A_{ik} (k \geq 2)$ ，有 $P(A_{i1}A_{i2}\dots A_{ik}) = P(A_{i1})P(A_{i2})\dots P(A_{ik})$ ，则称 n 个事件 A_1, A_2, \dots, A_n 相互独立.

(2) 独立性的判定

1° 直观性判定：若试验独立其结果必相互独立. 例如：甲、乙各自试验结果相互独立；袋中有返回取球其结果相互独立等.

2° 充要条件.

$$\langle 1 \rangle A_1 \dots A_n \text{ 相互独立} \Leftrightarrow \text{任意 } k \geq 2; P\left(\bigcap_{j=1}^k A_{ij}\right) = \prod_{j=1}^k P(A_{ij}).$$

特别地 A, B 独立 $\Leftrightarrow P(AB) = P(A)P(B)$. 若 $0 < P(A) < 1$,

则 A, B 独立 $\Leftrightarrow P(B|A) = P(B|\bar{A}) = P(B)$.

$\langle 2 \rangle n$ 个事件相互独立的充要条件是，它们中任意一部分事件换成各自的对立事件所得到的 n 个事件相互独立.

3° 必要条件.

$\langle 1 \rangle n$ 个事件相互独立必两两独立，反之不然.

$\langle 2 \rangle n$ 个事件相互独立，则不含相同事件的事件组经某种运算后所得的事件是相互独立的. 例如， A, B, C, D 相互独立，则 AB 与 $C \cup D$ 相互独立， A 与 $BC - D$ 相互独立，等等.

4° 一定独立与一定不独立的判定.

概率为 1 或零的事件与任何事件都相互独立. 如果 $0 < P(A) < 1, 0 < P(B) < 1$ ， A 与 B 互

不相容或存在包含关系，则 A 与 B 不相互独立.

【评注】 在现实生活中，难于想像两两独立而不相互独立的情况，可以这样想：独立性毕竟是一个数学概念，是现实世界中通常理解的那种“独立性”的一种数学抽象，它难免会有些不尽人意的地方.

3.例题分析

【例 1】 假设有 10 份报名表，3 份女生报名表，7 份男生报名表。现从中每次任取一份，取后不放回，求下列事件的概率：

- (1) 第三次取到女生报名表的概率；
- (2) 第三次才取到女生报名表的概率；
- (3) 已知前两次没有取到女生报名表，第三次取到女生报名表的概率.

【例 2】 假设甲、乙两名射击手，甲命中的概率为 0.6，乙命中的概率为 0.5，求

- (1) 甲、乙中任选一人去射击；
- (2) 甲、乙各自独立去射击；

若目标命中，则是甲命中的概率为多少？

第二讲 一维随机变量及其概率分布

1. 基本概念与常用分布

1. 随机变量的定义

随机变量就是“其值随机而定”的变量. 设随机试验 E 的样本空间为 $\Omega = \{\omega\}$, 如果对每一个 $\omega \in \Omega$, 都有唯一的实数 $X(\omega)$ 与之对应, 并且对任意实数 x , $\{\omega: X(\omega) \leq x\}$ 是随机事件, 则称定义在 Ω 上的实单值函数 $X(\omega)$ 为随机变量. 简记为随机变量 X . 一般用大写字母 X, Y, Z, \dots 或希腊字母 ξ, η, ζ, \dots 来表示随机变量.

【评注】 随机事件是从静态的观点来研究随机现象, 而随机变量则是一种动态的观点, 如高等数学中常量与变量的区别与联系.

2. 随机变量的分布函数

(1) 定义 设 X 是随机变量, x 是任意实数, 称函数 $F(x) \triangleq P\{X \leq x\} (x \in \mathbb{R})$ 为随机变量 X 的分布函数, 或称 X 服从分布 $F(x)$, 记为 $X \sim F(x)$.

(2) 充分必要条件 函数 $F(x)$ 为某一随机变量 X 的分布函数的充要条件是

1° $F(x)$ 是 x 的单调不减函数, 即对任意 $x_1 < x_2$, 有 $F(x_1) \leq F(x_2)$;

2° $F(x)$ 是 x 的右连续函数, 即对任意 $x_0 \in \mathbb{R}$, 有 $\lim_{x \rightarrow x_0^+} F(x) \triangleq F(x_0 + 0) = F(x_0)$;

3° $F(-\infty) \triangleq \lim_{x \rightarrow -\infty} F(x) = 0, F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$.

【评注】 务必记住分布函数是事件的概率. 由此知 $0 \leq F(x) \leq 1$, 即 $F(x)$ 是有界函数.

3. 离散型随机变量及其概率分布

如果随机变量 X 只可能取有限个或可列个值 x_1, x_2, \dots , 则称 X 为离散型随机变量, 称 $p_i = P(X = x_i), i = 1, 2, \dots$, 为 X 的分布列、分布律或概率分布, 记为 $X \sim p_i$. 概率分布常

常用表格形式或矩阵形式表示, 即

X	x_1	x_2	\dots
p	p_1	p_2	\dots

或 $X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$

数列 $\{p_i; i = 1, 2, \dots\}$ 是离散型随机变量概率分布的充要条件是 $p_i \geq 0 (i = 1, 2, \dots)$ 且 $\sum_i p_i = 1$.

设离散型随机变量 X 的概率分布为 $p_i = P(X = x_i)$, 则 X 的分布函数

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

$$p_i = P(X = x_i) = P(X \leq x_i) - P(X < x_i) = F(x_i) - F(x_i - 0)$$

并且对实数轴上的任一集合 B 有

$$P(X \in B) = \sum_{x_i \in B} P(X = x_i)$$

4. 连续型随机变量及其概率密度

如果随机变量 X 的分布函数可以表示为

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

其中 $f(x)$ 是非负可积函数, 则称 X 为连续型随机变量, 称 $f(x)$ 为 X 的概率密度函数, 简称为概率密度或密度函数, 记为 $X \sim f(x)$.

$f(x)$ 为某一随机变量 X 概率密度的充要条件是, $f(x) \geq 0$ 且 $\int_{-\infty}^{+\infty} f(x) dx = 1$ (由此可知, 可以改变 $f(x)$ 有限个点的值, $f(x)$ 仍然是密度函数).

设 $X \sim f(x)$, 则 X 的分布函数 $F(x)$ 是 x 的连续函数; 在 $f(x)$ 的连续点 x_0 处有 $F'(x_0) = f(x_0)$; 如果 $F(x)$ 是连续函数, 除有限个点外, $F'(x)$ 存在且连续, 则 X 为连续型随机变量, 且 $f(x) = F'(x)$ (在 $F'(x)$ 不存在地方可以令 $f(x) = 0$ 或取其他值).

设 X 为连续型的, 即 $X \sim f(x)$, 则对任意实数 c 有 $P(X=c)=0$; 对实数轴上任一集合 B 有

$$P(X \in B) = \int_B f(x) dx.$$

特别是

$$\begin{aligned} P(a < X < b) &= P(a \leq X < b) = P(a < X \leq b) \\ &= P(a \leq X \leq b) = \int_a^b f(x) dx = F(b) - F(a). \end{aligned}$$

【评注】(1)“密度函数”这名词的来由可解释如下, 取定一个点 x , 则按分布函数的定义, 事件 $\{x < X \leq x+h\}$ 的概率 ($h > 0$ 为常数), 应为 $F(x+h) - F(x)$. 所以, 比值 $[F(x+h) - F(x)]/h$ 可以解释为在 x 点附近 h 这么长的区间 $(x, x+h)$ 内, 单位长所占有的概率. 令 $h \rightarrow 0$, 则这个比的极限, 即 $F'(x) = f(x)$, 也就是在 x 点处 (无穷小区段内) 单位长的概率, 或者说, 它反映了概率在 x 点处的“密集程度”. 你可以设想一条极细的无穷长的金属杆, 总质量为 1, 概率密度相当于杆上各点的质量密度. (2) $P(a < X < b) = \int_a^b f(x) dx$ 意味着 X 落入某一区间的概率等于该区间之上、密度函数之下曲边梯形的面积, 应用概率的这种几何意义, 常常有助于问题的分析与求解.

2. 常见的离散型、连续型分布

(1) 0—1 分布 $B(1, p)$

如果 X 的概率分布为 $X \sim \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$ 即 $P(X=1)=p, P(X=0)=1-p$, 则称 X 服从参数

为 p 的 0—1 分布, 记为 $X \sim B(1, p) (0 < p < 1)$.

(2) 二项分布 $B(n, p)$

如果 X 的概率分布为 $p_k = P(X=k) = C_n^k p^k (1-p)^{n-k}, k=0, 1, \dots, n, 0 < p < 1$, 则称 X 服从参数为 (n, p) 的二项分布, 记为 $X \sim B(n, p)$.

【评注】 ①如果 X 是 n 重贝努利试验中事件 A 发生的次数, 则 $X \sim B(n, p)$, 其中 $p = P(A)$. 这个结论在解题中我们要经常用到.

②泊松定理, 若 $X \sim B(n, p)$, 当 n 很大, p 很小, $\lambda = np$ 适中时, 二项分布可用泊松分布近似, 即 $C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}$.

(3) 泊松分布 $P(\lambda)$

如果 X 的概率分布为 $p_k = P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}, k=0, 1, \dots, \lambda > 0$, 则称 X 服从参数为 λ 的泊松分布, 记为 $X \sim P(\lambda)$.

(4) 几何分布 $G(p)$

如果 X 的概率分布为 $p_k = P(X=k) = q^{k-1} p, k=1, 2, \dots, 0 < p < 1, q=1-p$, 则称 X 服从参数为 p 的几何分布, 记为 $X \sim G(p)$.

(5) 超几何分布 $H(N, M, n)$

如果 X 的概率分布为 $p_k = P(X=k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k=0, 1, \dots, \min(M, n), M, N, n$ 为

正整数, 则称 X 服从参数为 N, M, n 的超几何分布, 记为 $X \sim H(N, M, n)$.

(6) 均匀分布 $U(a, b)$

如果 X 的概率密度或分布函数为

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

则称 X 在区间 (a, b) 上服从均匀分布, 记为 $X \sim U(a, b)$.

【评注】 区间 (a, b) , 可以是闭区间 $[a, b]$; 几何概型是均匀分布的实际背景. 用几何概率计算事件概率时已假设点在区域内服从均匀分布. 几何概率可以用均匀分布计算.

(7) 指数分布 $E(\lambda)$

如果 X 的概率密度或分布函数为

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

$F(x)$ 则称 X 服从参数为 λ 的指数分布, 记为 $X \sim E(\lambda)$.

(8) 正态分布 $N(\mu, \sigma^2)$

如果 X 的概率密度

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (-\infty < x < \infty)$$

其中 $-\infty < \mu < \infty$, $\sigma > 0$, 则称 X 服从参数为 (μ, σ^2) 的正态分布或称 X 为正态变量, 记为 $X \sim N(\mu, \sigma^2)$. 此时 $f(x)$ 图形关于直线 $x = \mu$ 对称, 即 $f(\mu - x) = f(\mu + x)$, 并在 $x = \mu$ 处有唯一最大值 $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$.

称 $\mu = 0$, $\sigma = 1$ 的正态分布 $N(0, 1)$ 为标准正态分布, 通常记标准正态分布密度函数为

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \text{ 分布函数为 } \Phi(x). \text{ 显然 } \varphi(x) \text{ 为偶函数, } \Phi(0) = \frac{1}{2}, \Phi(-x) = 1 - \Phi(x).$$

若 $X \sim N(0, 1)$, $\Phi(\mu_\alpha) = P(X \leq \mu_\alpha) = \alpha$, 则称 μ_α 为标准正态分布下 α 分位数. 如果 $X \sim N(\mu, \sigma^2)$, 则其分布函数

$$F(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right); F(\mu-x) + F(\mu+x) = 1$$

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

$$aX + b \sim N(a\mu + b, a^2\sigma^2) (a \neq 0).$$

注意: “分布”两字在离散型随机变量场合泛指概率分布或分布函数; 对连续型则泛指概率密度或分布函数; 一般地则是指分布函数. 二项分布, 泊松分布与正态分布是概率中最重要的三个分布.

3. 典型例题分析

1. 求随机变量的分布

【例 1】已知随机变量 X 的概率分布为

k	1	2	3
$P(X = k)$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

且 $P(X \geq 2) = \frac{3}{4}$, 求未知参数 θ 及 X 的分布函数 $F(x)$.

【例 2】已知随机变量 X 的分布函数为 $F(x)$, 概率密度函数为 $f(x)$. 当 $x \leq 0$ 时, $f(x)$ 连续且 $f(x) = F(x)$, 若 $F(0) = 1$, 求 $F(x)$, $f(x)$.

2. 求随机变量函数的分布

【例】设 $X \sim U[0, \pi]$, 求 $Y = \sin X$ 的概率密度 $f_Y(y)$.

【例】设随机变量 X 的概率密度为 $f_X(x) = \begin{cases} \frac{1}{2}, & -1 < x < 0, \\ \frac{1}{4}, & 0 \leq x < 2, \\ 0, & \text{其他} \end{cases}$, $Y = X^2$, 求 $f_Y(y)$.

第三讲 随机变量的数字特征

1 基本概念

1 随机变量的数学期望

1. 定义 设 X 是随机变量, Y 是 X 的函数: $Y=g(X)$,

(1) 如果 X 是离散型随机变量, 其分布律为 $p_i=P\{X=x_i\}(i=1, 2, \dots)$, 若级数

$\sum_{i=1}^{\infty} x_i P\{X=x_i\}$ 绝对收敛, 则称随机变量 X 的数学期望存在, 并将级数 $\sum_{i=1}^{\infty} x_i P\{X=x_i\}$ 和称

为随机变量 X 的数学期望, 记为 $E(X)$ 或 EX , 即

$$EX = \sum_{i=1}^{\infty} x_i P\{X=x_i\}$$

否则称 X 的数学期望不存在.

若级数 $\sum_{i=1}^{\infty} g(x_i) P\{X=x_i\}$ 绝对收敛, 则称 $Y=g(X)$ 的数学期望 $Eg(X)$ 存在, 且 $Eg(X)=$

$\sum_{i=1}^{\infty} g(x_i) P\{X=x_i\}$. 否则称 $g(X)$ 的数学期望不存在.

(2) 如果 X 是连续型随机变量, 其密度函数为 $f(x)$, 若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛, 则称 X 数学期望 EX 存在, 且 $EX \triangleq \int_{-\infty}^{+\infty} xf(x)dx$, 否则称 X 的数学期望不存在. 若积分 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛, 则称 $g(X)$ 的数学期望存在, 且 $Eg(X) = \int_{-\infty}^{+\infty} g(x)f(x)dx$. 否则称 $g(X)$ 的数学期望不存在.

【评注】(1) 数学期望又称为概率平均值, 常常简称为期望或均值. 数学期望是描述随机变量平均取值状况特征的指标, 它刻画随机变量一切可能值的集中位置.

(2) 在数学期望的定义中要求级数(或积分)绝对收敛, 否则说期望不存在. 这是因为 X 的期望存在与 X 的取值顺序无关, 任意改变 x_i 的次序不应改变 EX 的存在性, 这在数学上就要求级数(或积分)绝对收敛, 况且绝对收敛又有很多性质也便于数学上的处理.

2. 性质

(1) 对任意常数 a_i 和随机变量 $X_i(i=1, \dots, n)$ 有

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i EX_i$$

特别地 $Ec=c$, $E(aX+c)=aEX+c$, $E(X\pm Y)=EX\pm EY$.

(2) 设 X 与 Y 相互独立, 则

$$E(XY)=EX EY, E g_1(X)g_2(Y)=E g_1(X) E g_2(Y)$$

一般地, 设 X_1, X_2, \dots, X_n 相互独立, 则

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n EX_i, E\left(\prod_{i=1}^n g_i(X_i)\right) = \prod_{i=1}^n E g_i(X_i).$$

2 随机变量的方差(Variance)与标准差

1. 定义 设 X 是随机变量, 如果 $E(X-EX)^2$ 存在, 则称 $E(X-EX)^2$ 为 X 的方差, 记为 DX 或 $Var(X)$, 即 $DX=Var(X) \triangleq E(X-EX)^2=EX^2-(EX)^2$.

称 \sqrt{DX} 为 X 的标准差或均方差, 称随机变量 $X^* \triangleq \frac{X-EX}{\sqrt{DX}}$ 为 X 的标准化随机变量,

此时 $EX^*=0$, $DX^*=1$.

2. 性质

$$(1) DX \geq 0, EX^2 = DX + (EX)^2 \geq (EX)^2.$$

$$(2) Dc=0 (c \text{ 为常数}).$$

$$DX=0 \Leftrightarrow P\{X=EX\}=1.$$

$$(3) D(aX+b)=a^2 DX.$$

$$(4) D(X \pm Y) = DX + DY \pm 2E(X - EX)(Y - EY)$$

$$= DX + DY \pm 2\text{cov}(X, Y)$$

$$D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j E(X_i - EX_i)(X_j - EX_j)$$

$$= \sum_{i=1}^n a_i^2 DX_i + 2 \sum_{1 \leq i < j \leq n} a_i a_j E(X_i - EX_i)(X_j - EX_j)$$

$$= \sum_{i=1}^n a_i^2 DX_i + 2 \sum_{1 \leq i < j \leq n} a_i a_j \text{cov}(X_i, X_j)$$

(5) 如果 X 与 Y 独立, 则

$$D(aX+bY)=a^2 DX+b^2 DY,$$

$$D(XY)=DXDY+DX(EY)^2+DY(EX)^2 \geq DX DY$$

一般地, 如果 X_1, X_2, \dots, X_n 两两相互独立, $g_i(x)$ 为 x 连续函数, 则

$$D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 DX_i,$$

$$D\left(\sum_{i=1}^n g_i(X_i)\right) = \sum_{i=1}^n Dg_i(X_i),$$

(6) 对任意常数 c , 有 $DX = E(X - EX)^2 \leq E(X - c)^2$.

2 典型例题

【例 1】设试验成功的概率为 $\frac{3}{4}$, 失败的概率为 $\frac{1}{4}$, 独立重复试验直到成功两次为止, 求试验次数的数学期望.

【例 2】已知随机变量 X 与 Y 的相关系数为 0.9, 若 $Z = X - 0.4$, 则 Y 与 Z 的相关系数为_____.

【例 3】将一枚硬币重复掷 n 次, 以 X 和 Y 分别表示正面向上和反面向上的次数, 则

$$\rho_{XY} = \underline{\hspace{2cm}}.$$