概率论与数理统计笔记

Yu Yang

Wed Nov 6 17:53:20 2013

Contents

1 概率论

1.1 事件的概率

1.1.1 古典概率

全部试验结果已知并且等可能的情况下才可以应用古典概率, 古典概率一般是使用排列组合来求解

1.1.2 概率的公理化定义

所有的试验结果定义一个全集 Ω , 任何定义于该集合上的事件都是它的子集

1.1.3 事件

一个事件可以看作是实验或者观察的一个结果：

1. 必须建立在一个明确的试验的基础上
2. 这个试验的所有可能的结果已知 (每一个可能的结果叫一个基本事件或简单事件)
3. 事件明确的界定这些结果中的一部分。

如果规定事件 E 为掷骰子掷出偶数点, 那么它就符合上面的三点要求

1.1.4 基本事件 (简单事件)

如果只包含结果集中的一种结果的话, 那么这个时间可以称为基本事件, 比如掷骰子掷出 1 点就是一个基本时间

1.1.5 互斥事件

如果两个事件不可能同时发生，那么他们就是互斥事件，表现在集合上就是两个集合无交集互斥事件有一个发生的概率：P(A+B) = P(A) + P(B) (用集合很好理解)

1.1.6 对立事件

是互斥事件的特例，表现在集合上两个集合互为补集、

1.1.7 事件的和 (补集), 积 (交集) 与差

都可以从集合的角度理解, 分别是补集, 交集和差集, 记住这只是事件本身的运算, 不是事件的概率的运算

1. 和 若有两个事件 A 与 B C = {A, B 至少有一个发生}, 那么 C 就是 A, B 的和, 记作: C = A+B
2. 积 若有两个事件 A 与 B C = {A, B 至少有一个发生}, 那么 C 就是 A, B 的积, 记作: C = AB
3. 差 若有两个事件 A 与 B C = {A 发生但是 B 不发生}, 那么 C 就是 A 与 B 的差, 记作: C = A-B

1.1.8 加法定律

对于 n 个互斥时间 A1, A2, … An P(A1 + A2 + … + An) = P(A1) + P(A2) + … P(An)

1.1.9 条件概率 (后验概率)

有两个事件 A 与 B, 在确定 B 发生的条件下,A 发生的概率 (也就是 A 的条件概率, 也称为 A 的后验概率) P(A|B) = P(AB)/P(B)

可以从集合的角度理解, 如果已知 B 发生这个条件, 那么要 A 发生显然只能是 A 与 B 的交集部分 千万不要把条件概率与事件的积混同

1.1.10 独立事件与乘法定律

> 如果事件 B 发生对事件 A 发生的概率无影响，那么 A,B 就是独立事件。。P(A|B) = P(AB) / P(B) ==> 这个是条件概率的定义 P(AB) = P(A) P(B) ==> 乘法定律, 独立事件积的概率等于各独立事件概率的积乘法定律的推论：P(A B 负) / P(B 负) = P(A)

如果 A，B 是独立事件，那么 P(A|B)=P(AB)/P(B)=P(A), 也就是 P(AB)=P(A)P(B) 特别注意独立事件表现在集合上，两个集合是有交集的，独立事件的本质是：A 在 B 发生的前提下的条件概率等于 A 的本来概率，也就是 B 的发生对 A 的发生概率没有影响。在实际应用中并不经常使用 P(AB) = P(A) P(B) 来验证是否是独立事件, 而是根据常识来进行判断两个事件是否有关联

1.1.11 全概率公式与贝叶斯公式

> B1, B2,…Bn 是一个完备事件群, 也就是说它们的和事件是必然事件且两两互斥, 也就是和事件的概率为 1. 已知 B1 + B2 + … + Bn = Ω (样本空间) 那么 A = A(B1 + B2 + … + Bn) P(A)=P(AB1)+P(AB2)+…+P(ABn) P(A) = P(B1)P(A|B1) + P(B2)P(A|B2) + … +P(Bn)P(A|Bn) ==> 全概率公式

全概率公式可以这样 B1, B2, … Bn 都是导致 A 发生的途径, 先计算 A 在以上各途径发生的条件概率, 然后在加权平均 (也就是乘以 Bn 的概率)

> P(B|A)=P(AB)/P(A) 而 P(AB)=P(A|B)P(B) 所以 P(B|A)=P(A|B)*P(B)/P(A) ==> 贝叶斯公式

1.2 随机变量的概率

1.2.1 随机变量

和函数中自变量的含义类似, 随机变量可以取不同的值, 而随机变量能取的所有值也会组成一个全集 Ω, 与古典概率不同的是每一个值对应的概率可以是不同的, 而古典概率中全集 Ω 中每一个元素对应一个基本事件, 而每个基本事件的概率是均等的. 随机变量分为离散性随机变量和连续性随机变量

1.2.2 离散性随机变量

1. 概率函数 f(x) = x 这点的概率
2. 概率分布函数 F(x) = 所有小于等于 x 的变量的和
3. 离散性随机变量的几个分布
   (a) 二项分布: 放回抽样, 如果事件 A 在一次试验中发生的概率为 p, 把这个试验独立的重复 n 次, 以 X 记 A 在这 n 次试验中发生的次数, 那么有下式

   $$P(X=i) = b(i,n,p) = \mathrm{C}_n^i p^i (1-p)^{n-i} \tag{1}$$

   记号为 B(n, p)

   (b) 泊松分布 (Poisson distribution): 很重要的离散性分布, 一般出现在表示一定时间或者空间内出现的事件个数这种场合, 它可以看作是二项分布极限得到的, 当二项分布的 np 不太大时, 可以近似为泊松分布

   $$P(X=i) = \frac{e^{-\lambda}\lambda^i}{i!} \tag{2}$$

   很显然 P(x=1)+P(x=2)+P(x=3)+... = 1, 证明过程如下: 使用泰勒级数展开e^x, 可得:

   $$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \tag{3}$$

   两相比较是明显的倒数关系.
   下面来说明泊松分布与二项分布的关系, 以及泊松分布的推导过程: 对于二项分布, 如果它的试验次数 n 很大, 同时概率 p 又很小, 但是 λ=np 适中, 于是 p=λ/n , 根据二项分布:

   $$P(X=i) = \mathrm{C}_n^i \frac{\lambda}{n}^i (1-\frac{\lambda}{n})^{n-i} \tag{4}$$

因为

$$\lim_{n\to\infty}\frac{\mathrm{A}_n^i}{n^i}=1 \tag{5}$$

所以

$$\mathrm{C}_n^i\frac{\lambda}{n}^i=\frac{\mathrm{A}_n^i}{n^i}\frac{\lambda^i}{i!}=\frac{\lambda^i}{i!} \tag{6}$$

因为

$$\lim_{n\to\infty}(1+\frac{1}{n})^n=e \tag{7}$$

所以

$$\lim_{n\to\infty}(1-\frac{\lambda}{n})^n=\lim_{n\to\infty}(1+\frac{1}{(-\frac{n}{\lambda})})^{-\frac{n}{\lambda}(-\lambda)}=e^{-\lambda} \tag{8}$$

$$(1-\frac{\lambda}{n})^{n-i}=e^{-\lambda} \tag{9}$$

所以

$$P(X=i)=e^{-\lambda}\frac{\lambda^i}{i!} \tag{10}$$

If P(X=i) = $\mathrm{e}^{-\lambda}\frac{\lambda^i}{i!}$

(c) 超几何分布: 不放回抽样,

$$P(X=m)=\frac{\mathrm{C}_M^m\mathrm{C}_{N-M}^{n-m}}{\mathrm{C}_N^n} \tag{11}$$

(d) 负二项分布:

4. 连续性随机变量的几个分布

(a) 正态分布 [Normal distribution](高斯分布 [Gaussian distribution])

$$f(x)=(\sqrt{2\pi}\delta)^{-1}e^{-\frac{(x-\mu)^2}{2\delta^2}},\qquad -\infty<x<\infty \tag{12}$$