ECE 313 Notes

Lecture 4: Independence & Random Variables

1. Independence (独立性)

1.1 边际独立性（Marginal Independence）

定义:
假设有两个随机变量 A 和 B。如果 A 与 B 是独立的，那么 A 的值不会受到 B 的值的影响。这种独立性可以通过以下公式表示：

$$p(A \mid B) = p(A)$$

也就是说，给定 B 的情况下，A 的概率仍然是 A 的边际概率，不依赖于 B。
从条件概率的角度理解:
根据条件概率的定义，条件概率 $p(A \mid B)$ 表示在已知 B 发生的情况下，事件 A 发生的概率。公式为：

$$p(A \mid B) = \frac{p(A, B)}{p(B)}$$

如果 A 和 B 是独立的，那么：

$$p(A \mid B) = p(A)$$

因此，结合上述公式，可以得到：

$$p(A, B) = p(A) \cdot p(B)$$

这表示 A 和 B 的联合概率等于它们边际概率的乘积。如果两个变量独立，它们的联合分布可以简化为各自边际分布的乘积。
Example：
假设我们掷一枚公平的硬币两次，设 A 表示第一次掷出正面，B 表示第二次掷出正面。由于两次掷硬币是独立事件，有：

$$p(A, B) = p(A) \cdot p(B)$$

如果 $p(A) = 0.5$ 且 $p(B) = 0.5$，则：

$$p(A, B) = 0.5 \times 0.5 = 0.25$$

1.2 条件独立性（Conditional Independence）

定义:
假设有三个随机变量 A、B 和 C。如果在给定 B 的条件下，A 和 C 是条件独立的，意味着一旦我们知道了 B 的值，额外知道 C 的值不会给我们关于 A 任何新的信息。公式如下：

$$p(A \mid C, B) = p(A \mid B)$$

这意味着在已知 B 的情况下，A 和 C 之间没有直接联系
从条件概率的角度理解:
条件独立性是指在给定某个随机变量的情况下，另两个随机变量之间的独立性。即：

$$p(A, C \mid B) = p(A \mid B) \cdot p(C \mid B)$$

这表示在给定 B 的条件下，A 和 C 的联合分布可以分解为它们各自条件分布的乘积。
Example：
设：

- A：学生的期末考试成绩
- B：学生的学习时间
- C：学生的学前测验成绩
  如果在知道了学生的学习时间 B 后，学生的学前测验成绩 C 对期末考试成绩 A 没有额外的影响，那么有：

$$p(A \mid C, B) = p(A \mid B)$$

这表示一旦我们知道了学生的学习时间 B，C 不再影响我们对 A 的预测。

1.3 独立性与条件独立性的比较

- **独立性**（边际独立性）：两个变量在所有情况下都不相互影响。即使不知道其他信息，它们的发生是完全独立的。
- **条件独立性**：两个变量在给定某些特定变量值的情况下是独立的。但若没有给定这些条件，它们可能不是独立的。
  直观理解：
- 独立性 $\Rightarrow$ 在任何情况下都不会相互影响；
- 条件独立性 $\Rightarrow$ 在某些条件下不再相互影响。

2. Random Variables（随机变量）

2.1 随机变量 vs. 事件

- 当随机变量 X 取某个特定值（如 $X = 1$）时，这就构成了一个**事件**（event）。
- 对某个值的概率，例如 $P(X = 1)$，表示的是**事件 $X = 1$ 发生的概率**。
- 所以我们仍然可以像处理事件一样，处理随机变量的取值概率。

2.2 随机变量的定义

- **随机变量**是指：**将实验的每一个结果赋值为一个实数**的函数。
- 比如投硬币的实验中，我们可以定义：
  - 正面（H） $\Rightarrow X = 1$
  - 反面（T） $\Rightarrow X = 0$

随机变量是一个从样本空间 Ω 到实数集 $\mathbb{R}$ 的函数：

$$X : \Omega \to \mathbb{R}$$

Tips:

- 随机变量的取值**必须是数字**，不能是“正面”“反面”这类文字或符号。
- 所以我们在定义随机变量时，往往会用数字来表示不同的实验结果。
- 通过这个方式，我们可以将原本抽象的事件结果变成可以运算的数值，从而计算概率、期望、方差等统计量。

概念	含义
随机变量 X	是一个函数，把实验结果映射为实数
$X = x$	是一个事件，表示随机变量取值为 x
$P(X = x)$	表示事件 $X = x$ 的概率

2.3 Probability Mass Function (PMF) 概率质量函数

假设 X 是一个定义在**可数样本空间** S 上的**离散随机变量**，其中 $S \subseteq \mathbb{R}$，那么其**概率质量函数** $f_X(x)$ 定义如下：

$$f_X(x) = \begin{cases} \Pr(X = x), & x \in S \\ \\ 0, & x \in \mathbb{R} \setminus S \end{cases}$$

将所有可能结果的概率之和相加即可得到：

$$\sum_{all\ k} P(X = k) = 1$$

3. Binomial Distribution（二项分布）

3.1 二项试验（Binomial Trial）的特点

二项分布用于描述具有以下特征的随机过程：

1. 进行 **n 次相互独立的试验**（如抛硬币）。
2. 每次试验有相同的成功概率 **p**（例如，正面朝上）。
3. 我们关心的是：**在 n 次试验中，恰好有 k 次成功的概率是多少？**

3.2 二项分布的概率质量函数（PMF）

若随机变量 X 表示在 n 次试验中成功的次数，并且满足二项分布：

$$X \sim \mathrm{Bin}(n, p)$$

则其概率质量函数为：

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

其中：

- $\binom{n}{k}$：从 n 次试验中选择 k 次成功的组合数；
- p：每次试验成功的概率；
- $(1 - p)$：每次试验失败的概率；
- k：成功的次数；
- n：试验总次数。

Example: $X \sim \mathrm{Bin}(4, \frac{1}{4})$

设 X 表示某人一天内收养的猫的数量，满足：

$$X \sim \mathrm{Bin}(4, \frac{1}{4})$$

则我们可以计算以下概率：

- **0 只猫被收养的概率**：

$$P(X=0) = \binom{4}{0} \cdot \left(\frac{1}{4}\right)^0 \cdot \left(1-\frac{1}{4}\right)^4$$

- **1 只猫被收养的概率**：

$$P(X=1) = \binom{4}{1} \cdot \left(\frac{1}{4}\right)^1 \cdot \left(1-\frac{1}{4}\right)^3$$

- **2 只猫被收养的概率**：

$$P(X=2) = \binom{4}{2} \cdot \left(\frac{1}{4}\right)^2 \cdot \left(1-\frac{1}{4}\right)^2$$

- 二项分布适用于 **固定次数、每次成功概率相同且相互独立** 的实验。
- 只要知道 n 和 p，就能用公式计算出任意 k 次成功的概率。

Addition：概率质量函数（PMF） vs 概率密度函数（PDF）

但实际上根据《普林斯顿概率论读本》P231，概率质量函数可以看作概率密度函数在离散型随机变量中的体现

适用类型不同

项目	概率质量函数（PMF）	概率密度函数（PDF）
适用于	离散型随机变量	连续型随机变量
示例	掷骰子结果、硬币正反面	身高、温度、电压等

定义方式不同

概率质量函数 PMF

- 定义：给出随机变量 X 取某个具体值 x 的概率。

$$P(X=x)$$

- 特点：
  - $P(X=x) \geq 0$
  - 所有可能值的概率之和为 1：

$$\sum_x P(X=x) = 1$$

- Example：投掷一枚硬币：

$$P(X=0)=0.5, \quad P(X=1)=0.5$$

概率密度函数 PDF

- 定义：随机变量 X 的概率密度函数 $f(x)$ 表示在某一点附近的"相对可能性"，**不能直接表示概率**。

  要得到某一范围的概率，必须积分：

$$P(a \leq X \leq b)=\int_a^b f(x)\, dx$$

- 特点：
  - $f(x) \geq 0$
  - 总面积为 1：

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

- 注意：
  - $P(X=x)=0$，连续变量在某个点取值的概率是 0。
  - 只有区间概率才有意义。

类比维度	**PMF**	**PDF**
图像	条形图	曲线图
概率计算	直接查值 $P(X=x)$	需要积分 $P(a \leq X \leq b)$
在某点的概率	有意义 $P(X=x)$	总是为零 $P(X=x)=0$
例子	掷骰子、抽签	人的身高、气温、电阻值

---

3. Summary

- 独立性和条件独立性
- 随机变量
- 二项分布
- 概率质量函数和概率密度函数的比较