

# 1. 概率论

## 1.1 概率

符号约定（主要遵守《强化学习 第2版》）：

- 随机变量用大写字母表示，如  $X$
- 随机变量的具体取值用小写字母表示，如  $x$

1. 概率一定是针对某一事件的概率，条件也一定是某个具体事件

$$\circ \Pr\{R_{t+1} = r, S_{t+1} = s' | S_t = s, A_t = a\} \text{ 或者 } p(r, s' | s, a)$$

因此，里面一定是小写字母

## 1.2 期望

1. 期望是针对某一个随机变量的

$$E[R_t | S_t = s] = \sum_r r \cdot p(r | s)$$

因此，里面是一定是大写字母

2. （在《强化学习 第2版》中）期望的下标，表示对目标随机变量的一种约束，

$$\circ \text{如 } E_\pi[R_t | S_t = s] \text{ 中的 } \pi \text{ 表示按照策略 } \pi \text{ 下的期望}$$

### 1.2.2 多重条件期望

形如  $E_Y[E_X[X|Y]]$  或  $E_Y[E_X[X|Y]|Z = z_i]$  的

- 内层期望  $E_X[X|Y = y_i]$ ，是在  $y_i$  条件下关于  $X$  的期望

$$E_X[X|Y = y_i] = \sum_x x \cdot p(x|y_i)$$

最终是关于  $y_i$  的函数，简记为  $f(y_i) = E_X[X|Y = y_i]$

- 外层期望需要对每一个可能的  $y_i$  进行分析，因此可记作

$$E_Y[E_X[X|Y]|Z = z_i] = E_Y[f(Y)|Z = z_i]$$

$$\begin{aligned} E_Y[f(Y)|Z = z_i] &= \sum_y f(y) \cdot p(y|z_i) \\ &= \sum_y \left[ \sum_x x \cdot p(x|y) \right] \cdot p(y|z_i) \end{aligned}$$

#### 1.2.2.1 内外条件相等

如， $E_Y[E_X[X|Y]|Y = y_i]$

当条件  $Y = y_i$  确定，内侧的  $E_X[X|Y = y_i]$  中的  $y_i$  只有一个值可取，因此

$$\begin{aligned}
 E_Y[E_X[X|Y]|Y = y_i] &= \sum_y \left[ \sum_x x \cdot p(x|y) \right] \cdot p(y|y_i) \\
 &= \sum_x x \cdot p(x|y_i) \\
 &= E_X[X|Y = y_i]
 \end{aligned}$$

### 1.2.2.2 内侧条件信息多于外侧

如  $E_{Y,Z}[E_X[X|Y, Z]|Y = y_i]$

内侧的条件为当  $Y = y_i, Z = z_i$  时，而外侧的条件为当  $Y = y_i$  时。因此内侧条件的信息比外侧更多。

此时， $E_{Y,Z}[E_X[X|Y, Z]|Y = y_i] = E_X[X|Y = y_i]$

- 可简记为，**先细化再粗化** 等价于 **直接粗化**。

### 1.2.2.3 内侧条件信息少于外侧

如  $E_Y[E_X[X|Y]|Y = y_i, Z = z_i]$

此时， $E_Y[E_X[X|Y]|Y = y_i, Z = z_i] = E_X[X|Y = y_i]$

- 可简记为，**先粗化再细化** 等价于 **直接粗化**

### 1.2.2.4 Tower Property

塔性质 (Tower Property) 是对“内侧条件信息多于外侧”的进一步概述

(以下为结合deepseek问答和[proof wiki](#)整理得到)

对于随机变量  $X$  和子  $\sigma$ -代数  $\mathcal{G} \subseteq \mathcal{H}$ ，条件期望满足：  $E[E[X | \mathcal{H}] | \mathcal{G}] = E[X | \mathcal{G}]$

- 直观理解：先基于更多信息  $\mathcal{H}$  求期望，再基于较少信息  $\mathcal{G}$  求期望，等同于直接基于  $\mathcal{G}$  求期望。
- 特例：

当  $\mathcal{G}$  为平凡  $\sigma$ -代数时，Tower Property 退化为：  $E[E[X | \mathcal{H}]] = E[X]$

即全期望公式：条件期望的期望等于原期望。

$\sigma$ -代数 (sigma-algebra) 是概率论和测度论中的一个核心概念，它用于描述“信息”或“可观测性”的结构。

## 1. 什么是 $\sigma$ -代数？

$\sigma$ -代数是一个集合的集合（即集合的族），它满足以下性质：

1. **包含全集**：全集  $\Omega$  属于  $\sigma$ -代数。
2. **对补集封闭**：如果集合  $A$  属于  $\sigma$ -代数，那么它的补集  $A^c$  也属于  $\sigma$ -代数。
3. **对可数并集封闭**：如果  $A_1, A_2, A_3, \dots$  属于  $\sigma$ -代数，那么它们的并集  $\bigcup_{i=1}^{\infty} A_i$  也属于  $\sigma$ -代数。

### 直观理解

- $\sigma$ -代数可以看作是一个“信息结构”，它告诉我们哪些事件是可观测的或可测量的。
- 例如，在概率空间中， $\sigma$ -代数描述了所有可能的事件（或集合），我们可以对这些事件分配概率。

## 2. $\sigma$ -代数和随机变量的关系

随机变量  $X$  会生成一个  $\sigma$ -代数，记作  $\sigma(X)$ 。这个  $\sigma$ -代数包含了所有形如  $X \leq a$  事件，其中  $a$  是实数。

换句话说， $\sigma(X)$  包含了所有可以通过  $XX$  观测到的事件。

**例子：**假设  $XX$  是一个随机变量，取值为 1 或 2。那么  $\sigma(X)$  包含以下集合：

- $\Omega = \{\text{所有可能的结果}\}$
- $X = 1$
- $X = 2$
- 空集  $\emptyset$

这些集合构成了  $\sigma(X)$ ，它描述了通过观测  $X$  所能获得的所有信息。

## 3. $\mathcal{G}$ 和 $\mathcal{H}$ 是什么？

$\mathcal{G}$  和  $\mathcal{H}$  是  $\sigma$ -代数，它们通常表示不同粒度的信息：

- $\mathcal{G}$  和  $\mathcal{H}$  可以是随机变量生成的  $\sigma$ -代数，例如  $\sigma(Y)$  和  $\sigma(Z)$ 。
- 它们也可以是更一般的  $\sigma$ -代数，描述不同层次的可观测性。

### 嵌套 $\sigma$ -代数

如果  $\mathcal{G} \subseteq \mathcal{H}$ ，这意味着  $\mathcal{H}$  包含的信息比  $\mathcal{G}$  更细。例如：

- $\mathcal{G} = \sigma(Y)$ ：仅包含  $Y$  的信息。
- $\mathcal{H} = \sigma(Y, Z)$ ：包含  $Y$  和  $Z$  的信息。

在这种情况下， $\mathcal{H}$  比  $\mathcal{G}$  提供了更多的细节。

## 4. Tower Property 中的 $\sigma$ -代数

在 Tower Property 中：

$$E[E[X \mid \mathcal{H}] \mid \mathcal{G}] = E[X \mid \mathcal{G}]$$

- $\mathcal{G}$  和  $\mathcal{H}$  是  $\sigma$ -代数，且  $\mathcal{G} \subseteq \mathcal{H}$
- $E[X \mid \mathcal{H}]$  是在更细的信息集下  $\mathcal{H}$  的条件期望
- $E[E[X \mid \mathcal{H}] \mid \mathcal{G}]$  是将这个条件期望进一步压缩到更粗的信息集  $\mathcal{G}$  下

## 1.2.3 指示函数的期望

### 1. 指示函数的定义

指示函数  $\mathbb{I}_{x=a}$  是一个二值函数，定义如下：

$$\mathbb{I}_{x=a} = \begin{cases} 1, & \text{如果 } x = a \\ 0, & \text{如果 } x \neq a \end{cases}$$

换句话说，当选择的动作  $x$  等于  $a$  时， $\mathbb{I}_{x=a} = 1$ ；否则， $\mathbb{I}_{x=a} = 0$ 。

### 2. 指示函数的期望

$$\mathbb{E}[\mathbb{I}_{x=a}] = \sum_x p(x) \cdot \mathbb{I}_{x=a} = p(a) \cdot 1 = p(a)$$

## 1.3 MP/MRP/MDP

《强化学习》一上来就是三者之中最复杂的MDP（当然，直接了解MDP毫无障碍）

MDP事实上是逐渐演化过来的：从马尔可夫过程（Markov Process）到马尔科夫奖励过程（Markov Reward Process）再到马尔可夫决策过程（Markov Decision Process）

（以下是基于Deepseek的介绍整理得到）

### 1.3.1 MP

**马尔可夫过程（Markov Process, MP）**，也称为**马尔可夫链（Markov Chain）**，描述了一个满足**马尔可夫性质**系统在不同状态之间随机转移的过程。

- **马尔可夫性质**: 即系统的下一状态只依赖于当前状态，而与过去的状态无关。

#### 1.3.1.1 MP的要素

马尔可夫过程由以下两个核心要素定义：

1. **状态空间（S）**：
  - 系统可能处于的所有状态的集合，可以是有限的或无限的。
  - 例如：天气模型中的状态空间可以是  $S=\{\text{晴天}, \text{雨天}, \text{阴天}\}$ 。
2. **状态转移概率矩阵（P）**：
  - 描述从一个状态转移到另一个状态的概率。
  - 转移概率  $P(s' | s)$  表示从状态  $s$  转移到状态  $s'$  的概率。
  - 转移概率矩阵满足：  $\sum_{s' \in S} P(s' | s) = 1$ ，即从任意状态出发，转移到所有可能状态的概率之和为 1

马尔可夫过程的描述

一个**状态序列**  $S_0, S_1, S_2, \dots, S_t, \dots$  其中：

- 每个状态  $S_t$  属于状态空间  $S$
- 状态转移遵循转移概率矩阵  $P$

#### 1.3.1.2 马尔可夫性质

MP的核心特性是**马尔可夫性质**，即：

$$P(S_{t+1} = s' | S_t = s, S_{t-1} = s_{t-1}, \dots, S_0 = s_0) = P(S_{t+1} = s' | S_t = s)$$

这意味着：

- 系统的下一状态  $S_{t+1}$  只依赖于当前状态  $S_t$ ，而与之之前的状态  $S_{t-1}, S_{t-2}, \dots, S_0$  无关
- 这种性质使得MP具有**无记忆性**

## 1.3.2 MRP

**马尔可夫奖励过程** (Markov Reward Process, MRP) 是**马尔可夫过程**的一个扩展，它在状态转移的基础上引入了奖励机制。

### 1.3.2.1 MRP的要素

1. MRP 由以下几个要素构成：

1. **状态空间 (S)**：系统可能处于的所有状态的集合。
2. **转移概率矩阵 (P)**： $P(s' | s)$  描述从一个状态转移到另一个状态的概率。
3. **奖励函数 (R)**：在状态转移时获得的即时奖励。
  - 可以是  $R(s)$ ，也可以是  $R(s, s')$ 。后者更通用一些
4. **折扣因子 ( $\gamma$ )**：用于衡量未来奖励的当前价值，取值范围为  $[0, 1]$ 。

2. MRP的描述

MRP 生成的序列是一个 **状态-奖励序列**，形式为  $(S_0, R_1, S_1, R_2, S_2, R_3, \dots)$

- 每个状态  $S_t$  属于状态空间  $S$
- 状态转移遵循转移概率矩阵  $P$
- 奖励遵循奖励函数  $R$

### 1.3.2.2 回报与价值

1. 奖励与回报

(奖励在《强化学习》中称为收益)

- **奖励**：每次状态转移时，系统会获得一个即时奖励。
- **回报**：从当前状态开始，未来所有奖励的折扣总和。

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

2. 值函数  $V(s)$  表示从状态  $s$  开始的期望回报：

$$V(s) = \mathbb{E}[G_t | S_t = s]$$

3. 贝尔曼方程

$$V(s) = R(s) + \gamma \sum_{s' \in S} P(s' | s) V(s')$$

注意：MRP中不考虑动作，因此没有动作价值函数  $Q(s, a)$

## 1.3.3 MDP

**马尔可夫决策过程** (Markov Decision Process, MDP) 是强化学习的核心数学模型

- 用于描述一个智能体在环境中通过选择动作来影响状态转移和获得奖励的过程。

MDP 在马尔可夫奖励过程 (MRP) 的基础上引入了 **动作 (action)** 和 **策略 (policy)** 的概念

### 1.3.3.1 MDP 的要素

1. **状态空间 (S)**：系统可能处于的所有状态的集合。
2. **动作空间 (A)**：智能体在每个状态下可以选择的动作的集合。
3. **状态转移概率 (P)**： $P(s'|s, a)$  描述在状态  $s$  下选择动作  $a$  后，转移到状态  $s'$  的概率。
4. **奖励函数 (R)**：描述在状态  $s$  下选择动作  $a$  后，转移到状态  $s'$  时获得的即时奖励。
  - 可以表示为  $R(s, a)$  或  $R(s, a, s')$ 。
5. **折扣因子 ( $\gamma$ )**：用于衡量未来奖励的当前价值，取值范围为  $[0, 1]$ 。
6. **策略 ( $\pi$ )**： $\pi(a|s)$  策略定义了智能体在每个状态下选择动作的概率分布。

### 1.3.3.2 值函数

- **状态值函数  $v_\pi(s)$** ：  
在策略  $\pi$  下，从状态  $s$  开始的期望回报：

$$v_\pi(s) = \mathbb{E}[G_t | S_t = s]$$

- **动作值函数  $q_\pi(s, a)$** ：  
在策略  $\pi$  下，从状态  $s$  选择动作  $a$  后的期望回报：

$$q_\pi(s, a) = \mathbb{E}[G_t | S_t = s, A_t = a]$$

### 1.3.3.3 贝尔曼方程

(按照《强化学习》中，这里使用动态特性  $p(s', r|s, a)$ )

- **状态值函数的贝尔曼方程**：

$$v_\pi(s) = \sum_a \pi(a|s) \cdot \sum_{s', r} p(s', r|s, a) \cdot [r + \gamma v_\pi(s')]$$

- **动作值函数的贝尔曼方程**：

$$q_\pi(s, a) = \sum_{s', r} p(s', r|s, a) \cdot [r + \gamma \sum_{a'} \pi(a'|s) \cdot q_\pi(s', a')]$$

## 2. 数理统计

### 2.1 参数估计

(Deepseek)

#### 1. 参数估计的目的

参数估计的目的是通过样本统计量来推断总体参数。例如，通过某城市1000人的平均身高来估计全国人口的平均身高。

#### 2. 总体与样本

- **总体 (Population)**：研究对象的全体。例如，全国所有人的身高。
- **样本 (Sample)**：从总体中抽取的一部分个体。例如，某城市1000人的身高。

#### 3. 参数与统计量

- **参数 (Parameter)** : 描述总体特征的数值。例如, 全国人口的平均身高  $\mu$ 。
- **统计量 (Statistic)** : 描述样本特征的数值。例如, 某城市1000人的平均身高  $\bar{X}$ 。

#### 4. 参数估计的类型

- **点估计 (Point Estimation)** :
  - 提供一个单一的数值作为参数的估计。  
例如, 用样本均值  $\bar{X}$  作为总体均值  $\mu$  的估计。
- **区间估计 (Interval Estimation)** :
  - 提供一个区间作为参数的估计, 通常包括一个置信水平。  
例如, 全国人口的平均身高在165cm到175cm之间, 置信水平为95%。

#### 5. 常用的估计方法

(主要针对点估计)

1. **矩估计法 (Method of Moments)** :
  - 通过样本矩来估计总体矩。  
例如, 用样本均值估计总体均值。
2. **最大似然估计法 (Maximum Likelihood Estimation, MLE)** :
  - 选择使样本数据出现概率最大的参数值。  
例如, 通过最大化似然函数来估计参数。
3. **贝叶斯估计法 (Bayesian Estimation)** :
  - 结合先验分布和样本数据, 得到参数的后验分布。  
例如, 使用贝叶斯定理更新参数的估计。

#### 6. 评估估计量的标准

- **无偏性 (Unbiasedness)** : 估计量的期望值等于被估计参数的真实值。
- **有效性 (Efficiency)** : 估计量的方差越小越好。
- **一致性 (Consistency)** : 随着样本量的增加, 估计量收敛到被估计参数的真实值。
- **充分性 (Sufficiency)** : 估计量包含样本中关于参数的所有信息。

## 2.1.1 矩估计

核心思想: **设定总体矩等于样本矩**

假设我们有一个随机样本  $X_1, X_2, \dots, X_n$ , 并且我们想要估计总体均值  $\mu$  和总体方差  $\sigma^2$ 。

- 样本均值:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

- 样本方差:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

通过设定总体矩等于样本矩, 我们得到:

$$\mu = \bar{X}, \sigma^2 = S^2$$

正态分布和指数分布的矩方程是不同的? 高阶矩?

## 2.1.2 最大似然估计

### 1. 什么是似然函数？

**似然函数** 是用于描述在给定参数下，观察到当前样本数据的概率的函数。它的核心思想是：

- 假设总体分布的类型已知（例如正态分布、泊松分布等），但分布的参数（如均值  $\mu$ 、方差  $\sigma^2$  等）未知。
- 通过样本数据，找到最有可能生成这些数据的参数值。

换句话说，似然函数衡量的是：在某个参数值下，当前样本数据出现的可能性有多大。

### 2. 似然函数的数学定义

假设我们有一个随机样本  $X_1, X_2, \dots, X_n$ ，这些数据来自某个总体分布，其概率密度函数 (PDF) 为  $f(x | \theta)$ ，其中  $\theta$  是未知参数（可以是单个参数，也可以是向量参数，如  $\theta = (\mu, \sigma^2)$ ）。

- **单个数据点的似然**：对于单个数据点  $X_i$ ，其似然为  $f(X_i | \theta)$ 。
- **样本的联合似然**：由于样本是独立同分布的，整个样本的联合似然函数是单个数据点似然的乘积：

$$L(\theta) = \prod_{i=1}^n f(X_i | \theta)$$

这里， $L(\theta)$  就是似函数。

### 3. 为什么叫“似然”？

- **概率 (Probability)**：描述的是在已知参数  $\theta$  的情况下，观察到某一样本数据的可能性。
  - 概率是固定参数，变化数据。
- **似然 (Likelihood)**：描述的是在已知样本数据的情况下，某个参数值  $\theta$  的可能性。
  - 似然是固定数据，变化参数。

### 4. 最大似然 (Maximum Likelihood Estimation)

最大似然估计的目标是找到参数  $\theta$ ，使得似然函数  $L(\theta)$  最大

$$\hat{\theta} = \arg \max_{\theta} L(\theta)$$

### 5. 最大似然求解步骤

#### 1. 构建似然函数

假设我们有一个随机样本  $X_1, X_2, \dots, X_n$ ，这些数据来自某个总体分布，其概率密度函数 (PDF) 为  $f(x | \theta)$ 。似然函数为：取对数似然函数

$$L(\theta) = \prod_{i=1}^n f(X_i | \theta)$$

#### 2. 为了简化计算，通常对似然函数取对数，得到对数似然函数：

$$\ln L(\theta) = \sum_{i=1}^n \ln f(X_i | \theta)$$

#### 3. 对参数 $\theta$ 求导

对  $\ln L(\theta)$  关于参数  $\theta$  求导，并令导数等于零：

$$\frac{d}{d\theta} \ln L(\theta) = 0$$

#### 4. 解方程



解上述方程，找到使对数似然函数最大的参数值  $\hat{\theta}$ 。

## 2.1.3 贝叶斯估计

### 2.1.3.1 贝叶斯估计的核心思想

贝叶斯估计的核心思想是：**利用先验信息和样本数据，更新对参数的认知**。具体来说：

- **先验 (Prior)**：在观察到数据之前，对参数的初始认知。
- **后验 (Posterior)**：在观察到数据之后，对参数的更新认知。

贝叶斯估计通过 **贝叶斯定理** 将先验信息和样本数据结合起来，得到后验分布。

### 2.1.3.2 贝叶斯定理

1. 贝叶斯估计的核心思想是：**利用先验信息和样本数据，更新对参数的认知**。具体来说：

- **先验 (Prior)**：在观察到数据之前，对参数的初始认知。
- **后验 (Posterior)**：在观察到数据之后，对参数的更新认知。

贝叶斯估计通过 **贝叶斯定理** 将先验信息和样本数据结合起来，得到后验分布。

2. 贝叶斯定理是贝叶斯估计的基础，其数学形式为：

$$P(\theta | X) = \frac{P(X | \theta) \cdot P(\theta)}{P(X)}$$

其中：

- $\theta$ ：待估计的参数。
- $X$ ：观察到的样本数据。
- $P(\theta)$ ：**先验分布**，表示在观察到数据之前对参数的认知。
- $P(X | \theta)$ ：**似然函数**，表示在给定参数下观察到数据的概率。
- $P(\theta | X)$ ：**后验分布**，表示在观察到数据之后对参数的更新认知。
- $P(X)$ ：**边缘似然**，是一个归一化常数，确保后验分布的总概率为 1。

3. 边缘似然

$$P(X) = \sum_{\theta} P(X | \theta) \cdot P(\theta)$$

因此

$$\begin{aligned} \sum_{\theta} P(\theta | X) &= \sum_{\theta} \frac{P(X | \theta) \cdot P(\theta)}{P(X)} \\ &= \frac{1}{P(X)} \sum_{\theta} P(X | \theta) \cdot P(\theta) \\ &= 1 \end{aligned}$$

4. 在贝叶斯估计过程中，通常不会使用

$$P(\theta | X) = \frac{P(X | \theta) \cdot P(\theta)}{P(X)}$$

去准确计算出后验概率。而是使用正比关系，确定后验概率的分布形式，并从  $P(X | \theta) \cdot P(\theta)$  中获得估计值

$$P(\theta | X) \propto P(X | \theta) \cdot P(\theta)$$

两个式子只相差一个归一化系数 $P(X)$ 。而 $P(X)$ 不依赖于参数 $\theta$ ，因此在计算后验分布时， $P(X)$ 可以被视为一个常数忽略掉。

### 2.1.3.3 贝叶斯估计过程

#### 1. 过程

- 选择先验分布  $P(\theta)$
- 计算似然函数  $P(X | \theta)$
- 计算后验分布，并估计参数

#### 2. 示例

假设我们有一个正态分布模型  $X \sim \mathcal{N}(\mu, \sigma^2)$ ，其中  $\sigma^2$  已知， $\mu$  是未知参数。

##### 1. 先验分布

假设  $\mu$  的先验分布为：

$$\mu \sim \mathcal{N}(\mu_0, \sigma_0^2)$$

##### 2. 似然函数

给定数据  $X = \{X_1, X_2, \dots, X_n\}$ ，似然函数为：

$$P(X | \mu) = \prod_{i=1}^n \mathcal{N}(X_i | \mu, \sigma^2)$$

##### 3. 后验分布

根据贝叶斯定理：

$$P(\mu | X) \propto P(X | \mu)P(\mu)$$

将正态分布的概率密度函数代入：

$$P(\mu | X) \propto \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu)^2\right) \cdot \exp\left(-\frac{1}{2\sigma_0^2} (\mu - \mu_0)^2\right)$$

合并指数部分：

$$P(\mu | X) \propto \exp\left(-\frac{1}{2} \left( \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 + \frac{1}{\sigma_0^2} (\mu - \mu_0)^2 \right)\right)$$

通过完成平方，我们可以得到后验分布的形式：

$$\mu | X \sim \mathcal{N}(\mu_n, \sigma_n^2)$$

其中：

$$\mu_n = \frac{\sigma_0^2}{\sigma_0^2 + \frac{\sigma^2}{n}} \bar{X} + \frac{\frac{\sigma^2}{n}}{\sigma_0^2 + \frac{\sigma^2}{n}} \mu_0$$
$$\sigma_n^2 = \left( \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2} \right)^{-1}$$

## 2.1.3.4 贝叶斯预测

### 1. 后验预测分布的定义式

$$P(X_{new} | X) = \sum_{\theta} P(X_{new} | \theta) P(\theta | X)$$

其中：

- $P(X_{new} | \theta)$  是似然函数。
- $P(\theta | X)$  是参数的后验分布。
- 求和表示对所有可能的参数值  $\theta$  进行加权平均，相当于期望。

### 2. 独立性假设

(在大多数情况下)，当给定参数  $\theta$  时，新数据  $X_{new}$  的生成与已有数据  $X$  无关。因此，我们可以假设：

$$P(X_{new} | \theta, X) = P(X_{new} | \theta)$$

而当没有给定参数  $\theta$  时， $\theta$  的信息已经在历史的  $X$  中了，因此  $X_{new}$  和  $X$  是相关的：

$$P(X_{new} | X) \neq P(X_{new})$$

### 3. 定义式的简单推导

$$\begin{aligned} P(X_{new} | X) &= \sum_{\theta} P(X_{new}, \theta | X) \\ &= \sum_{\theta} P(\theta | X) P(X_{new} | X, \theta) \\ &= \sum_{\theta} P(\theta | X) P(X_{new} | \theta) \end{aligned}$$

## 2.2 评价标准

还有一致性

### 2.2.1 无偏性 (Unbiasedness)

#### 1. 定义

设  $\hat{\theta}$  是参数  $\theta$  的估计量，如果满足  $E(\hat{\theta}) = \theta$ ，则  $\hat{\theta}$  是参数  $\theta$  的估计量。

例如：样本均值  $\bar{X}$  是总体均值  $\mu$  的无偏估计，因为  $E(\bar{X}) = \mu$

#### 2. 对无偏的理解：

- 无偏性意味着估计量在多次重复实验中，其平均值会接近真实参数值，没有系统性偏差。
- 无偏性并不保证单次估计的准确性，仅保证长期来看没有偏差。

#### 3. 对 $\theta$ 和 $E(\hat{\theta})$ 的理解

- $\theta$  是总体分布的参数，它是未知的，但为定值
- $\hat{\theta}$  是根据样本对  $\theta$  的估计值，具体取值受样本影响

而期望  $E(\hat{\theta})$  是估计量  $\hat{\theta}$  的理论平均值，表示在大量重复实验中  $\hat{\theta}$  的平均值。

计算  $E(\hat{\theta})$  时，我们是对  $\hat{\theta}$  的所有可能取值进行加权平均，权重是  $\hat{\theta}$  的概率分布。

$$E(\hat{\theta}) = \sum_{\hat{\theta}} \hat{\theta} \cdot p(\hat{\theta})$$

- 因此，当总体分布是确定的时候， $E(\hat{\theta})$  是一个固定的数值，表示  $\hat{\theta}$  的长期平均行为，与某次样本的取值是无关的。

这一点对无偏性的理解至关重要。

#### 4. 以矩估计总体均值为例

假设总计符合高斯分布  $X \sim N(\mu, \sigma^2)$ ，进行多次独立采样得到样本  $X_1, X_2, \dots, X_N$ ，很明显每个样本都符合原来的高斯分布，且样本与样本间是相互独立的。

- 此时，样本均值  $\bar{X} = \frac{1}{N} \sum_i X_i$
- 根据矩估计，我们认为样本均值等于总体均值（ $\bar{X}$  相当于  $\hat{\theta}$ ， $\mu$  相当于  $\theta$ ），即  $\bar{X} = \mu$

分析无偏性，

$$E[\bar{X}] = E\left[\frac{1}{N} \sum_i X_i\right] = \frac{1}{N} \sum_i E[X_i]$$

- 因为直接采样的样本必然符合总体分布，即  $X_i \sim N(\mu, \sigma^2)$ ，所以  $E[X_i] = \mu$   
也就是说，不管当前的样本  $X_i$  取了什么样的值， $E[X_i]$  都为  $\mu$   
由此，消除了计算  $E[\bar{X}]$  时，样本自身的影响
- 将  $E[X_i] = \mu$  带入，得到  $E[\bar{X}] = \mu$ ，即符合  $E(\hat{\theta}) = \theta$ ，因此矩估计总体均值是无偏的

## 2.2.2 有效性 (Efficiency)

### 1. 定义

对同一个  $\theta$  有两个估计量  $\hat{\theta}_1$  和  $\hat{\theta}_2$

若  $\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2)$ ，则称  $\hat{\theta}_1$  更有效

- 方差越小，估计值越稳定，波动越小，因此更有效

### 2. $\text{Var}(\hat{\theta})$ 也是一个和样本无关的定值

$$\text{Var}(\hat{\theta}_1) = E[(\hat{\theta} - E(\hat{\theta}))^2]$$

仍然可以用分析  $E(\hat{\theta})$  时的思路来考虑： $\text{Var}(\hat{\theta})$  对  $[\hat{\theta} - E(\hat{\theta})]^2$  所有可能的取值做加权平均

## 2.2.3 无偏和有效的结合

一种估计方法，可能无偏性和有效性，两者不能兼得。

在评价一种估计方法的好坏时，可以使用综合指标同时考虑两者，比如均方误差 (Mean Square Error, MSE)

$$\text{MSE} = E[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta}) + (E[\hat{\theta}] - \theta)^2$$

- 第一项评估有效性
- 第二项评估无偏性（若无偏，则为零）

简单推导

$$\begin{aligned}
\text{MSE} &= E[(\hat{\theta} - \theta)^2] \\
&= E[(\hat{\theta} - E(\hat{\theta})) - (\theta - E(\hat{\theta}))]^2] \\
&= E[(\hat{\theta} - E(\hat{\theta}))^2 - 2[\hat{\theta} - E(\hat{\theta})][\theta - E(\hat{\theta})] + [\theta - E(\hat{\theta})]^2] \\
&= E[(\hat{\theta} - E(\hat{\theta}))^2] - 2E[(\hat{\theta} - E(\hat{\theta}))(\theta - E(\hat{\theta}))] + E[(\theta - E(\hat{\theta}))^2]
\end{aligned}$$

- 第一项

$$\text{第一项} = \text{Var}(\hat{\theta})$$

- 第三项

因为  $\theta - E(\hat{\theta})$  是常数，详见“无偏性”中的讨论

$$\text{第三项} = [\theta - E(\hat{\theta})]^2$$

- 第二项

$\theta - E(\hat{\theta})$  和  $E(\hat{\theta})$  是常数

$$\begin{aligned}
\text{第二项} &= -2[\theta - E(\hat{\theta})] \cdot E[\hat{\theta} - E(\hat{\theta})] \\
&= -2[\theta - E(\hat{\theta})] \cdot [E[\hat{\theta}] - E(\hat{\theta})] \\
&= 0
\end{aligned}$$

## 2.3 采样

### 2.3.1 重要性采样

#### 1. 基本概念

- 假设我们有一个目标分布  $p(x)$ ，但我们很难直接从  $p(x)$  中采样。
- 我们可以选择一个**提议分布 (proposal distribution)**  $q(x)$ ，这个分布更容易采样。
- 然后，通过从  $q(x)$  中抽取样本，并调整这些样本的权重，来估计  $p(x)$  下的期望值。

#### 2. 公式推导

假设我们要估计  $f(x)$  在  $p(x)$  下的期望值：  $E_p[f(x)] = \sum f(x)p(x)$

使用重要性采样，我们可以改写为：

$$E_p[f(x)] = \sum \frac{f(x)p(x)}{q(x)} q(x) = E_q\left[\frac{f(x)p(x)}{q(x)}\right]$$

其中，  $\frac{p(x)}{q(x)}$  称为**重要性权重**。

#### 2. 步骤

- 选择提议分布  $q(x)$
- 按  $q(x)$  抽取样本
- 对每个样本计算重要性权重  $w_i = \frac{p(x)}{q(x)}$
- 估计期望值：

$$E_p[f(x)] \approx \frac{1}{N} \sum_{i=1}^N w_i \cdot f(x_i)$$

约等于是因为采样个数有限，只有当 $N$ 为 $\infty$ 时才变成等于

### 3. 普通重要性采样，加权重要性采样

#### ○ 普通重要性采样

$$E_p[f(x)] \approx \frac{1}{N} \sum_{i=1}^N w_i \cdot f(x_i)$$

- 无偏性：普通重要性采样的估计值是无偏的，即期望值等于真实期望值。
- 高方差：由于重要性权重可能非常大，估计值的方差可能会很高，导致估计不稳定。

#### ○ 加权重要性采样

$$E_p[f(x)] \approx \frac{\sum_{i=1}^N w_i \cdot f(x_i)}{\sum_{i=1}^N w_i}$$

- 有偏性：加权重要性采样的估计值是有偏的，尤其是在样本数量较小时。但随着样本数量的增加，偏差会逐渐减小。
- 低方差：通过对权重进行归一化，加权重要性采样的方差通常比普通重要性采样低，因此估计更稳定。

#### ○ 场景对比

特性	普通重要性采样	加权重要性采样
无偏性	无偏	有偏（但随着样本数量增加，偏差减小）
方差	高	低
适用场景	样本数量较大时	样本数量较小时

4. 注意：在重要性采样中，被估计的参数是  $E_p[f(x)]$ ，估计量可记为  $\hat{E}_p[f(x)]$