# 第一章 文献翻译&个人理解

**15 极大极小下界**

在短暂的讨论信息论偏移之后，让我们回到k-armed随机赌博机上。在下面的内容中，我们规定范围和动作次数。本章有两个组成部分。首先是针对固定策略和不同的赌博机，精确计算典型赌博机模型中措施之间的相对熵。在第二部分中，我们证明了一个极大极小下界，将第13章给出的直觉论点形式化。

个人注：本章主要讨论决策理论问题中极大极小风险的下界。这种界限对于评估决策规则的质量很有用。

**15.1 赌博机之间的相对熵**

下面的结果将被反复使用，练习题中提供了一些概括。

**引理15.1：散度分解**

令为一个k-armed赌博机相关的奖励分布，令为另一个k-armed赌博机相关的奖励分布。固定一些政策，同时令和为由和（和）的轮相互联络所得到的正则赌博机模型（4.6节）的概率测度。

因此有：



证明：

假设对于所有都有。由此可见。定义，为任意可测集A定义的度量, 定理14.1表明，只要，则有：



回顾是上的计数测度，我们发现关于乘积测度的Radon-Nikodym导数在式（4.7）中给出：



除了被取代外，的密度相同，则有：



其中，我们使用了Radon-Nikodym衍生品的链规则和涉及策略的条款取消的事实。取双方的期望:



以及



其中，第二个等式在条件下，的分布是，回插到前式中得：



当（15.1）式的右边是无穷大时，由我们前面的计算不难看出，左边也是无穷大的。

我们注意到，无论动作集是否离散，发散分解都成立。在更一般的形式下，有关行动的总和必须用适当的非负措施的积分来代替，该积分概括了预期的臂拔出数量。详情见练习15.8。

个人注：相对熵（relative entropy）又称为KL散度（Kullback–Leibler divergence，简称KLD），信息散度（information divergence），信息增益（information gain）。KL散度是两个概率分布P和Q差别的非对称性的度量。 KL散度是用来 度量使用基于Q的编码来编码来自P的样本平均所需的额外的位元数。 典型情况下，P表示数据的真实分布，Q表示数据的理论分布，模型分布，或P的近似分布。简单来说，相对熵用来衡量两个取值为正的函数或概率分布之间的差异

**15.2 极大极小下界**

记是具有单位方差的高斯赌博机类，可以通过它们的平均向量进行参数化，令，为第个有奖励分布为的高斯赌博机。

**引理15.2：**

令且，因此对于任何政策，存在一个平均向量，使得：



从而，因此，当时，极大极小边界的极小边界当由上述等式右边主导的边缘下界：



证明的思想如图15.1所示。

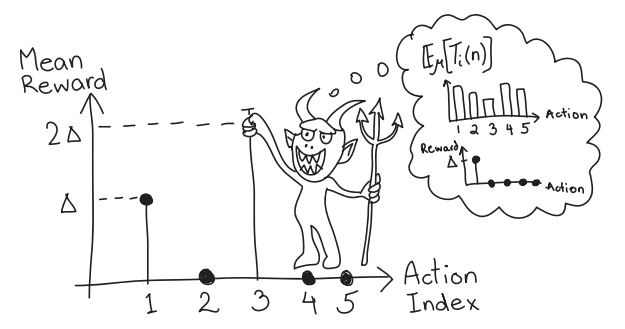


图15.1极大极小下界的概念。如果有一个政策和一个环境反对者选择另一个环境，使政策在至少一个环境中遭受巨大的偏差遗憾

证明：

固定策略，令是稍后选择的常数。根据第13章中的建议，我们从单位方差和平均向量的高斯赌博机开始。这种环境下和在正则空间上产生了分布和赌博机模型。为了简洁起见，我们将使用来代替。同时，将用表示。要选择第二个环境，令



从而，，并且。第二个高斯赌博机也是单位方差且均值为



其中。因此，除了指数和是第一个臂，而在中，臂是最佳的。我们缩写。将引理4.5进行一个简单的计算得到



然后，应用前一章中的Bretagnolle–Huber不等式（定理14.2），



它保持在上限。为此，我们使用引理15.1和获得



将此插入之前的显示，我们发现



通过选择，其中不等式遵循定理陈述中的假设。最后一步是下限和。

我们鼓励读者阅读练习15.2中概述的替代证明，其中采取了稍微不同的路径。

**15.3 注记**

1、我们使用高斯噪声模型，因为KL发散度在这种情况下很容易计算但我们实际使用的是，当平均值之间的差距很小。虽然并非所有情况都是这样，但通常情况下确实如此。为什么呢？令是上的一些参数分布族，并假设分布具有平均。假设密度是两次可微的，并且所有的都是积分和导数可以交换（几乎总是这样），我们可以用一个关于的泰勒展开式来表示



其中，第二行中引入的称为族的Fisher信息在。注意，如果是的的常用主要度量，则我们可以写



这是小学课文中通常给出的形式。这一切的结果是，因为实际上是的二次方，通过提供的缩放，因此，最糟糕的遗憾总是，提供所考虑的分配类别足够充分，也不太奇怪。

2、我们现在已经显示了一个下限，虽然许多上限是。这并不矛盾，因为对数界限取决于次优间隙的倒数，次优间隙可能非常大。

3、我们的下限仅为。在练习15.3中，我们要求您理解当时，有这样一个赌博机



4、用于证明定理15.2的方法可以看作是对统计中的Le Cam方法。回想一下，等式（15.2）对于任意和



为了解释Le Cam的方法，我们需要一点符号。令为结果空间，为上的一系列措施，，其中是一个度量空间。估计器是一个函数。Le-Cam方法用于证明期望误差的极大极小下界估计量



这个方法是用来选择去最大化，在任意的基础上



式中，。与赌博机下限相比，有两个区别：（1）我们处理顺序设置；（2）选择后，我们选择的方式取决于关于算法。这提供了一个非常需要的额外提升，如果没有它，该方法将是无法捕捉的特征如何反映在极大极小风险中（或遗憾，在我们的案例）

**15.4 文献综述**

我们知道的第一个关于下界的工作是Vogel [1960]对双臂伯努利赌博机的非常精确的极大极小分析。Bubeck等人[2013b]首次将Bretagnolle–Huber不等式（定理14.2）用于赌博机。正如注记中所述，利用这个不等式证明下界，在统计学中被称为Le Cam方法[Le Cam, 1973]。定理15.2的证明采用了与Gerchinovitz和Lattimore [2016]相同的思路，而练习题15.2中的另一种证明本质上是由Auer等人[1995]提出的，他们分析了奖励是伯努利的更困难的情况（见练习题15.4）。Yu [1997]描述了一些替代Le Cam方法的被动、统计设定。这些备选方案可以（而且经常可以）适应序列设置。

# 第二章 关键证明过程公式推导

**15.1 公式推导**

缩写，同时令。通过三角不等式得



令在上认为，在上认为



结果如下，因为