**熵的统计意义及其应用**

刘正浩2019270103005

电子科技大学英才实验学院

【摘要】简要阐述了熵的统计意义及熵在不同领域的应用

【关键词】熵 统计意义 信息熵

**1.背景**

熵（Entropy），最早在1865年由德国物理学家克劳修斯给出，他用熵来度量一个系统“内在的混乱程度”。简单地说，熵反映的是一个系统的混乱程度，一个系统越混乱，其熵就越大；越是有序，其熵就越小。在熵的概念被提出后，这个概念就不断地被渗透到各个学科中，用来解释各种现象。在本文中，将简要介绍信息熵这一概念。

**2.熵的统计意义**

**2.1熵是自发过程进行方向的标志**

我们考虑一个孤立系统，它可能有A、B、……等各种状态。热力学第一定律告诉我们，这系统的总能量是不变的。因此如果开始时它处于某一状态A，以后它只可能到达总能量状态A时的总能量相等的某些状态，如状态B、C、……等，而绝不可能到达总能量与A时的总能量不相等的那些状态，热力学第二定律则进一步告诉我们，一个孤立系统只能由熵小的状态向熵大的状态进行，因此，要使状态A和状态B的总能量相同，如果状态B的熵小于状态A的熵，就不可能自发地从状态A变到状态B，由此可见，热力学第一定律对自然界的变化发展过程作了能量方面的限制，即孤立系统只能到达总能量相等的那些状态；而热力学第二定律则对自然界的变化发展过程进一步作了进行方向方面的限制，即孤立系统只能从熵较小的状态向熵较大的状态变化发展，而决不能从熵大的状态向熵小的状态变化。

**2.2熵可以表示能量可被利用的程度**

假设有一个系统A，温度为，输出热量为，周围媒质的温度为，因而可以在系统A与媒质之间构成一个热机。

热机的效率为

因而产生的机械功为

热量中不可用的部分（即放到低温热源的热）为

现在假定先将热量Q传给温度为()的另一物体B，在B和媒质之间构成另一个热机，则热量Q的可用部分将减少为

而不可用部分则增加为

比较前后两次，热量Q的不可用部分增加了

同时，温度为和的这两个物体所构成的系统，其熵的增量为

因此，热量的不可用部分的增量为

可见，熵可以作为能量的不可用程度的一种量度。也就是说，在一切宏观过程中，能量的总和保持不变，但可利用的程度却随着熵的增加而降低。能量的可用程度降低也就意味着能量的品质变差了。

**3.信息熵简介**

在机器学习中，信息熵是一个很重要的概念。信息熵的公式如下：

它描述的是在结果出现之前对信息量的期望。因而，信息熵可以给决策提供一定的判断依据，比如决策树就利用熵来设置向下的分枝。

半个世纪前，时年32岁的香农发表了一篇论文：《通信的数学理论》（A mathematical theory of communication）。在这篇论文中，香农提出了用信息熵来定量衡量一个信息的大小。

我们先设随机事件发生的不确定性为发生概率的函数。这个函数具有以下性质：

（1）单调性：概率越大的事件，信息熵越小，反之亦然。

（2）非负性：的取值只能为0或者正数。

（3）可加性：多个随机事件同时发生的总不确定性的度量，可以表示为各个时间不确定性度量的和。例：

事件同时发生，其发生概率为，而应当满足：。

香农从数学上证明了满足上述三个性质的信息熵函数具有唯一的形式，也就是上文的信息熵公式。

信息熵不仅定量衡量了信息的大小，还为信息编码提供了理论上的最优值：使用的编码平均码长的理论下界就是信息熵。即信息熵为数据压缩的极限。至于对信息熵的详细推导与分析，作为通信理论的知识，在此不再赘述。

**参考文献**

【1】王可达,张之翔.熵的定义和物理意义[J].汕头大学学报(自然科学版), 1997(02):88-93**.**

【2】C. E. Shannon. 2001. A mathematical theory of communication. SIGMOBILE Mob. Comput. Commun. Rev. 5, 1 (January 2001), 3–55.