

计算借出概率与传统统计方法的对比

- 图书馆买一本书，这本书被利用的情况如何？
- 传统上只能简单的统计**借阅的次数**。
- 很显然这样的统计是**过于简单的**。
- 因为简单的借阅次数统计**无法区分图书需求强度和馆藏实际情况的匹配程度**。
-

概率计算设计

- 仅需要单位周期（例如一年）之内图书“被借出的次数 C ”和“被借出的总天数 T_j ”，即可估算该图书在未来被借出的概率。
- 模拟计算可访问：<http://drdemo.mag4net.com/wei/>

计算借出概率的优势

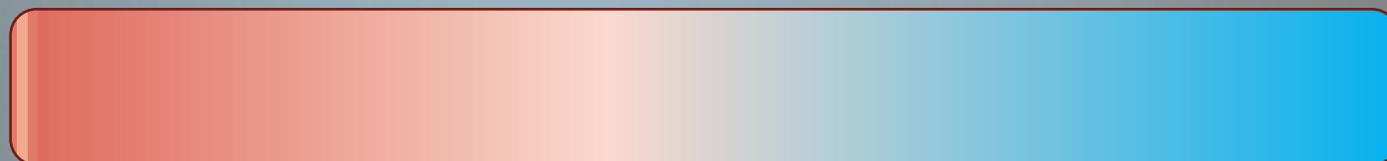
- 评价一本书的利用情况，必须要通过“想借”和“可借”这两个因素共同判定。
- 将流通的行为量化为一个可以预估的概率值。可以让图书馆掌握流通图书**真实利用和需求情况**。
- 通过比较概率值的大小就可以从读者**需求和资源**两个新的角度反应图书利用的特征数据。
- 通过对概率值的分析，可以为**调整馆藏量，副本量，剔旧，馆际之间图书调配**提供参考依据。

概率值分布的解释

高

中

低



需求大

需求模糊

需求低

可借到

不确定性强

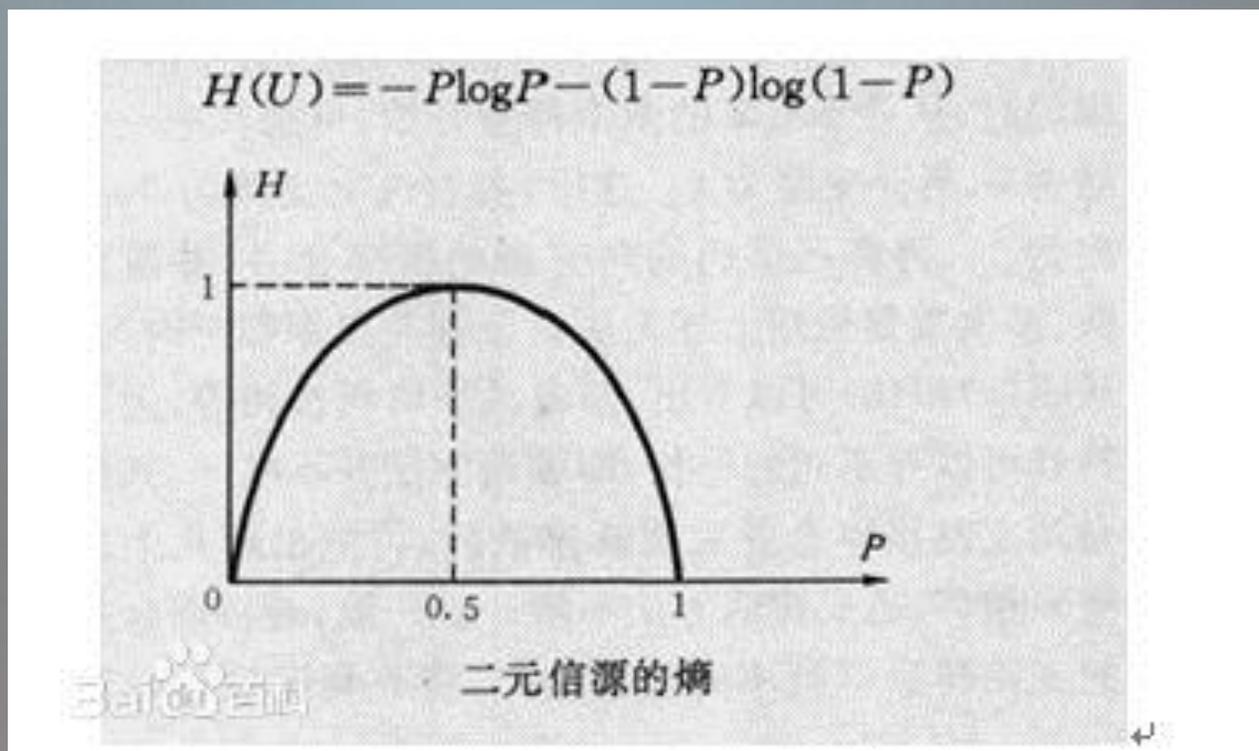
借不到

利用情况好

利用情况不明

可能副本不足

香农信息量公式计算整体流通馆藏的比特值



香农信息量公式计算整体流通馆藏的比特值

- 最简单的单符号信源仅取0和1两个元素，即二元信源，其概率为P和 $Q=1-P$ ，该信源的熵即为上图所示。
- 由图可见，离散信源的信息熵具有：
 - ①非负性：即收到一个信源符号所获得的信息量应为正值， $H(U) \geq 0$
 - ②对称性：即对称于 $P=0.5$
 - ③确定性： $H(1,0)=0$ ，即 $P=0$ 或 $P=1$ 已是确定状态，所得信息量为零
 - ④极值性：因 $H(U)$ 是 P 的上凸函数，且一阶导数在 $P=0.5$ 时等于0，所以当 $P=0.5$ 时， $H(U)$ 最大。

整体馆藏比特量计算

- 对于任意一本书均有借出概率 P_n ， n 属于集合（1.2.3.....）。
- 则 $H(P_n) = -P_n \log(P_n) - (1-P_n) \log(1-P_n)$ ；log以2为底
- 设整体馆藏比特值 H_B 。则有

$$H_B = \sum_{i=1}^n H(P_i)$$

H_B 的值越大，图书借出概率越介于50%左右。消除不确定性的所需比特值越大。

H_B 的值越小，图书的利用分布于概率两头。消除不确定性的所需比特值越小。

整体馆藏比特量的意义

- 通过比较**HB**的值就可以，量化衡量图书馆流通图书的利用效果。
- **HB**的值可以为图书馆馆藏的评估提供一个新的维度。

总结

- 智慧图书馆——**从传统的工作中挖掘新的智慧。**
- 智慧图书馆——**具备对需求的探索和研究。**
- 智慧图书馆——**从经验式的服务走向量化型的管理。**

祝智慧图书馆联盟更加繁荣！