指标D值计算和筛选

**一、单个指标D值计算**

以指标“说明书页数”的D值计算为例。

1指标“说明书页数”的每个专利的数值都对应一个或几个专利（有数值相同的），将指标数值降序排列，列入第（2）列。

2确定“说明书页数”不同指标数值对应的专利中的“转让专利个数”和“未转让专利”的个数，分别列入第（3）、第（4）列。

3计算“转让专利累积频数”。第（3）列数依次累加，列入第（5）列。

和“未转让专利累积频数”。第（4）列数依次累加，列入第（6）列。

4计算“转让专利累积频率”。第（5）列的数分别除以第（5）列最后一行的数x，得出结果分别列入第（7）列。

同理，“未转让专利累积频率”。 第（6）列的数分别除以第（6）列最后一行的数y，得出结果分别列入第（8）列。

5计算累积频率差值。第（7）列的数对应减去第（8）列的数，取绝对值，列入第（9）列。

6计算第（9）列最大值，即为指标“说明书页数”的D值。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| （1）专利群编号 | （2）指标数值 | （3）转让专利个数 | （4）未转让专利个数 | （5）转让专利累积频数 | （6）未转让专利累积频数 | （7）转让专利累积频率 | （8）未转让专利累积频率 | （9）累计频率差值 | （10）D值 |
| 1 |  | a |  | a |  | a/x |  |  | =max第（9）列 |
| 2 |  | b |  | a+b |  | (a+b)/x |  |  |
| 3 |  | c |  | a+b+c |  | (a+b+c)/x |  |  |
| 4 |  |  |  | … |  |  |  |  |
| … |  |  |  | … |  | … |  |  |
| X |  |  |  | x | y | x/x=1 |  |  |

单个指标D值计算完毕。（需要计算全部指标的D值）

**二、通过相关性分析对指标进行筛选**

计算**同一准则层内**指标的相关系数。设：rqj-第q个指标和第j个指标的相关系数；xiq-第i个专利的第q个指标值；q-第q个指标平均值；xij-第i个专利第j个指标值；j-第j个指标的平均值。则相关系数rqj为：

在相关系数大于0.7的两个指标中，删除D值小的指标。

|  |  |
| --- | --- |
| 准则层 | 指标 |
| 技术方面的指标 | 说明书页数 |
| 附图个数 |
| 专利引证数 |
| 专利被引数 |
| … |
| 法律方面的指标 | 是否有优先权 |
| 专利的保护期 |
| … |

指标第一次筛选完成。。。。

**三、指标的第二次筛选。。。**

对第一次筛选后剩余的指标，进行指标的第二次筛选。（比如还剩20个指标）

1. 根据指标D值，对20个指标赋权。

某指标权重=该指标D值 除以 20个指标D值的和

1. 根据权重，计算每个专利的“得分”

某专利“得分”=100\*（）

1. 根据上式计算的专利“得分”，计算20个指标的D值。（跟第一次筛选步骤一样，只是把第一次数据中的“指标数值”换为“专利得分”）

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| （1）专利群编号 | （2）**专利得分** | （3）转让专利个数 | （4）未转让专利个数 | （5）转让专利累积频数 | （6）未转让专利累积频数 | （7）转让专利累积频率 | （8）未转让专利累积频率 | （9）累计频率差值 | （10）D值 |
| 1 |  | a |  | a |  | a/x |  |  | =max第（9）列 |
| 2 |  | b |  | a+b |  | (a+b)/x |  |  |
| 3 |  | c |  | a+b+c |  | (a+b+c)/x |  |  |
| 4 |  |  |  | … |  |  |  |  |
| … |  |  |  | … |  | … |  |  |
| X |  |  |  | x | y | x/x=1 |  |  |

Step1计算出20个指标的D值。确定出最大的D20max

依次去掉20个指标中的一个指标，计算剩余19个指标组合的D19max（比如去掉第一个指标后，剩余19个指标计算出来的D最大）

Step2比较去掉该指标前后D20max与D19max的大小，如果D20max小于D19max ，则删除去掉的指标。

重复step1和step2，继续删除指标。

删除指标终止。(依次去掉一指标后，剩余指标计算出的Dmax大于去掉指标前的Dmax）

第二次筛选完成，剩余指标即为最优指标体系。