1. 势头与表现的关系

* 量化"势头"（Momentum）：

- "势头"通常用来描述球员在比赛中的连续胜局或得分趋势，这种趋势可能在一段时间内持续，然后突然改变。其强调的是比赛中的连胜局或连续得分趋势，通常关注比赛的整体动态。

* 量化"表现"：

"表现"通常指一方在某一局或一段时间内的出色表现，通常是指在得分、发球、接发球等方面的积极因素。量化"表现"的方式可以基于各种比赛统计数据。其强调的是在某一局或某一段时间内的具体技术和策略表现，关注的是短期内的局势。

## 任务二

1. "势头"定义:在网球比赛中,势头通常是"表现"与身体状态的结合。考虑到选手的身体状态不便估计,我们忽略每位选手的身体状态以及其对他们比赛的影响.因此我们将势头近似看成"表现分"。对于"表现分"的影响指标也同样适用于"势头".

5. 基于游程检验随机性检验

我们将p1与p2的势头序列导入到游程检验，从而判断势头序列是否具有随机性。通过python代码求得出p值<0.05,满足显著性检验,所以该序列是非随机序列.

为了验证比赛中的转折与选手的连胜是否具有随机性,我们提出一种游程检验的方法,将势头与转折点进行检验,来判段样本是否具有随机性.

其公式为$Z\_x=W(x)=(2n\_1\*n\_2+n)/n$

$σ^2\_x=2n\_1n\_2(2n\_1n\_2-n)/n^2(n-1)$

那我们可以得到统计量:

$$R=\frac{-Z\_x+r}{\sigma}\sim N(0,1)$$

r为游程总数,n1,n2分别表示价格上升和价格下降的个数.n为所取的样本数,即$n\_1+n\_2=n$.这样.当n很大时,x近似服从正态分布.假设$Z\_x$是正态分布条件下的期望值,$σ$为方差.

在给定显著性水平为$\alpha$时,样本为n的情况,可以统计出R的临界值,为$R\alpha/2$,如果样本中得到的r>$R\alpha/2$,就可以拒绝原假设,即序列不是随机的,否则接受原假设,即序列是随机的.

6. 多次比赛检验

通过对全部比赛进行游程检验,将所得的结果进行汇总如下表所示.

==见数据/工作簿2.xlsx==

通过对p1和p2的势头变化可知,所有比赛的势头对于球员而言都呈现出非随机性.其值都为1,这可能是因为一些因素所导致的,例如球员的实际水平,战术等等.

## 总结:

我们首先将势头序列进行差分处理,以此更好的描述势头的波动情况,（其次我们采用CUMSUM算法去寻找势头序列中的转折点.为第三问做铺垫.

最后运用游程检验将差分处理后的势头序列进行随机性检验,得出结论为比赛波动与成功不具备随机性.

问题三

2. 势头差分序列的计算:

为了更好的描述每个时间点势头的变化幅度,我们首先对每个球员的势头进行差分处理，以此来计算两点的势能差值.时间序列公式如下：

G(t),……

其中t=1,2....n

所以差分序列D(t)可以表示为:

D(t)=G(t)-G(t-1)

其中D(t)表示在t和t-1时间之间的势头变化量.

该差分序列的每个值都能够很好的反映每个球员表现的实时变化,正值表示势头的增加,负值表示势头的减少.

3. CUMSUM的使用:

累计和算法(CUMSUM)是通过累计差分序列中的正负变化来识别势头变化的算法。也就是说通过计算差分序列的累计总和来捕捉势头随时间的累计变化趋势,当CUMSUM曲线出现显著性变化时,表示势头可能经历了一个转折点.公式为:

$$F(t)=\sum^t\_{i=1}D(i)$$

其中,F(t)表示从比赛开始到t时间内的势头累计变化量.CUMSUM通过观察F(t)的变化趋势来识别潜在的转折点.

4. 转折点的识别与提取.

基于CUMSUM的结果,当累计和曲线的趋势发生显著性变化时,我们称为转折点.,这些点表示比赛中势头的趋势转变.

在这里放转折点的图片

图片显示1号与2号选手的势头随着比赛的进度表现出一定的波动,它的转折点表示势头发生显著变化的时刻.对应着球员表现的上升和下降.