# 小样本量下的马尔可夫链独立性检验与定阶

罗磊

2025年07月18日

## 一、主要内容

本文主要包括以下内容：

1. 案例构建和样本采集
2. 小样本量下的置换检验
3. 马尔可夫链定阶

## 二、案例构建和样本采集

假定有两个独立的骰子和，每个骰子有6个面。接下来，独立重复次掷骰子试验，每次试验持续较少的步。分别记录第次试验的第步骰子结果为和。最终获得样本数组和用于分析。

### 2.1 状态转移设计

在本案例中，使用状态转移矩阵控制骰子和各自的点数变化。状态转移矩阵每个元素表示从当前状态转移到下个状态的概率，因此的每一行之和必为1。由于骰子的点数为1至6，因此状态空间维数为6，也对应为的矩阵。

如果相邻两次投掷的点数独立，则状态转移矩阵为：

 (1)

如果相邻两次骰子点数不独立，比如如下的1阶马尔可夫过程状态转移：

 (2)

因此，可以通过设计状态转移的方式对每次试验的点数变化规律进行控制。需注意，由于骰子和相互独立，因此不论状态转移矩阵如何设置，和试验样本都应该无关。

### 2.2 采样结果

下图分别展示了某次独立和1阶Markov过程采样所得和的变化曲线。可见，由于Markov过程中相邻时刻状态存在依赖，和各自的变化更为平缓：

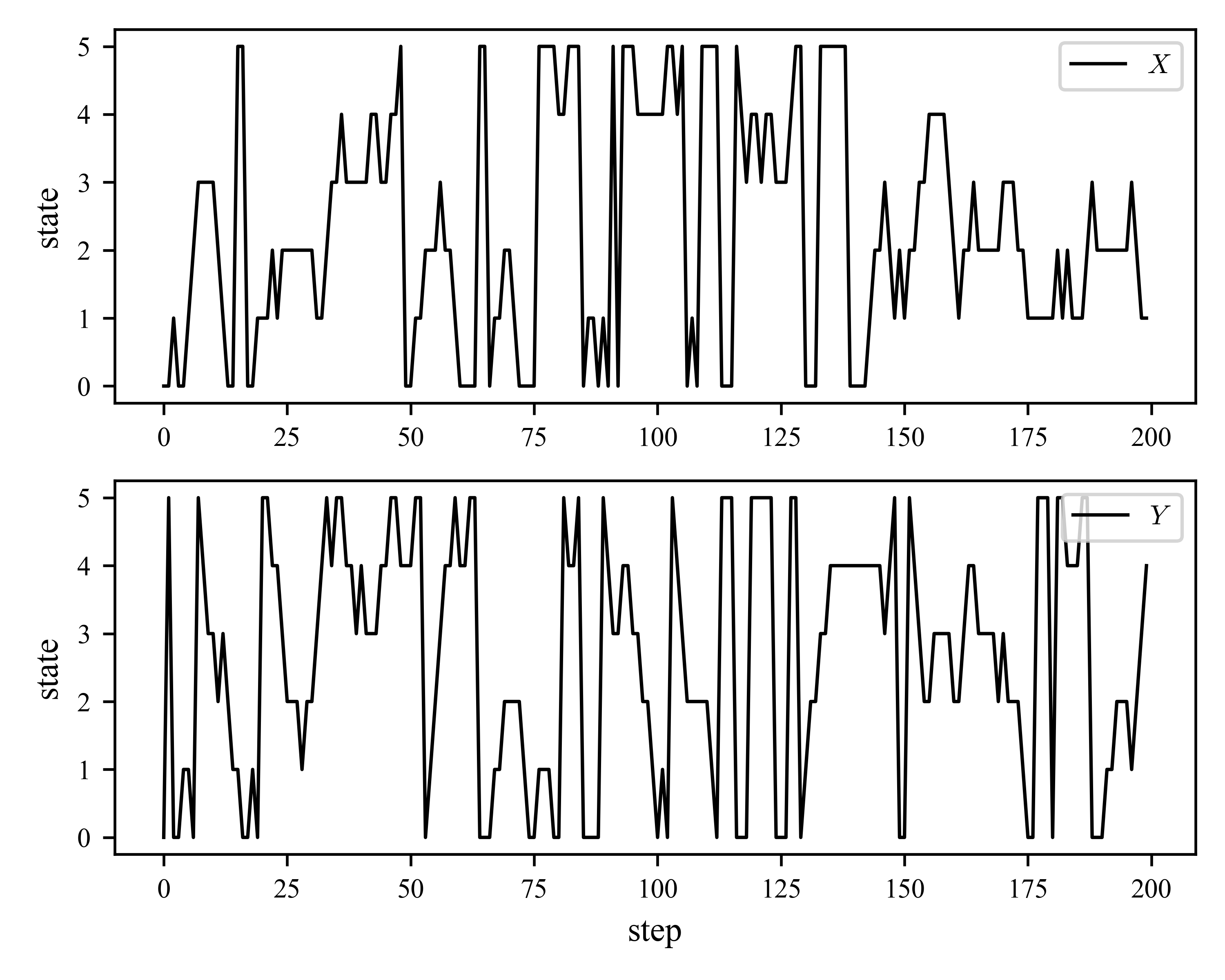
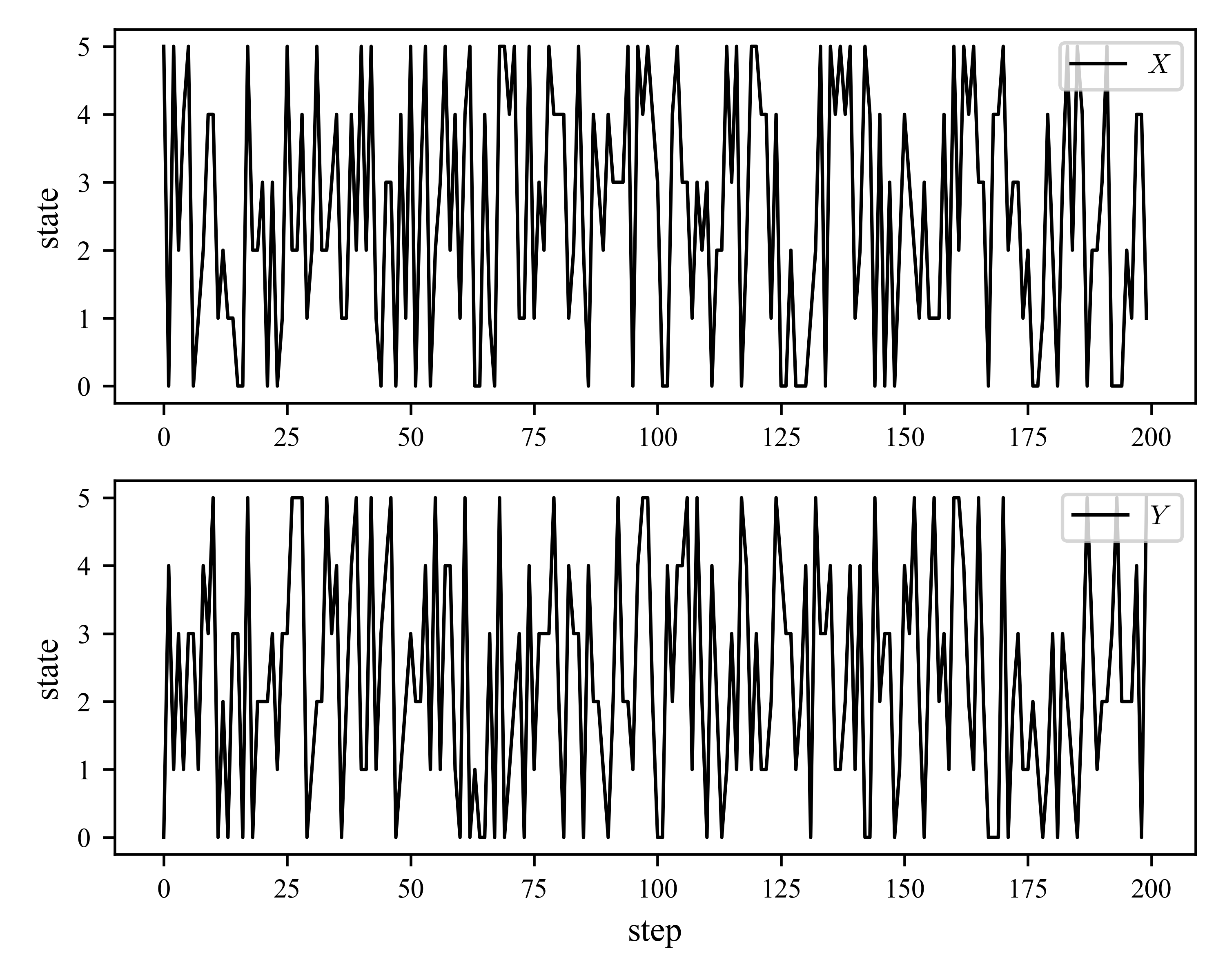


图 1. 独立（左）和Markov过程（右）采样结果展示

## 三、小样本量下的独立性检验

### 3.1 独立同分布序列样本

由于两个骰子相互独立，所以理论上任意试验中和都应相互独立，那么对样本和进行独立性检验所得结果都应为独立。此外，由于每次试验的步数很少（实际场景中很常见），我们希望在如此小的样本量下所得独立性检验结果依然可靠。

置换检验与Bootstrap检验等非参检验方法可用于小样本量下复杂分布数据之间的独立性检验。设对某-联合分布（注意不一定是来自马尔可夫链）进行独立同分布采样所得的样本分别为和，通过互信息计算其关联度为。接下来，我们希望通过置换检验判定仍然遵从零假设：不显著大于0即和无关，拒绝备择假设：显著大于0即和关联。

#### 3.1.1 排列检验

在零假设下，和的样本可视为来自同一总体，因此其标签（组别）可随机交换而不会导致统计量的显著改变。首先，将和数据合并为总样本集。接下来，进行轮置换，在每轮置换中，随机从抽取一半的样本作为置换组，剩下的另一半作为，计算关联系数。多次置换后，获得经验分布：

 (3)

进而计算值：

 (4)

以上便是对分别来自独立同分布采样的和进行小样本置换检验的计算步骤，式（1）所对应的掷骰子试验便符合该特点。对第1次掷骰子试验所得100步样本进行置换检验，结果如下图所示。设显著性水平为=0.01，单侧检验，此时值远大于，因此拒绝接受，认为该组和样本独立。

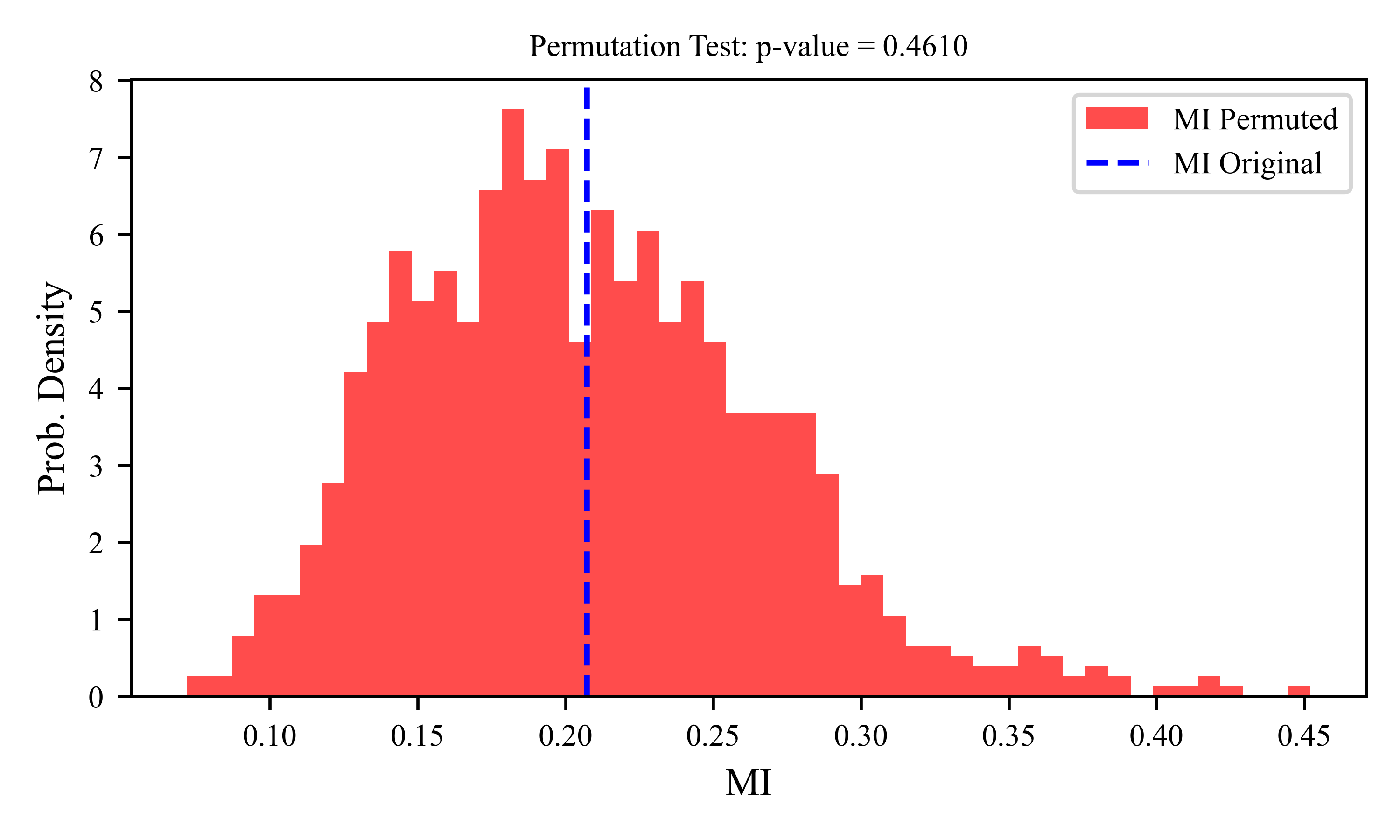


图 2. 式(1)状态转移过程对应第1次试验样本的排列检验结果

对所有=100次试验样本都进行上述置换检验，最终I类错误率为2%。对于这类样本量为=100的独立样本，置换检验有98%的概率正确接受，认为和无关；2%的概率错误拒绝而接受认为和关联。

#### 3.1.2 蒙特卡洛检验

与上述排列检验不同，蒙特卡洛置换检验的形式更适用于关联系数计算。对于一组和样本进行轮置换，在每轮置换时，随机打乱作为，然后计算关联系数。多次置换后，获得经验分布：

 (5)

进而计算值：

 (6)

采用该方法对式（1）模组试验样本计算结果如下图所示，对应的I类错误率为3%。

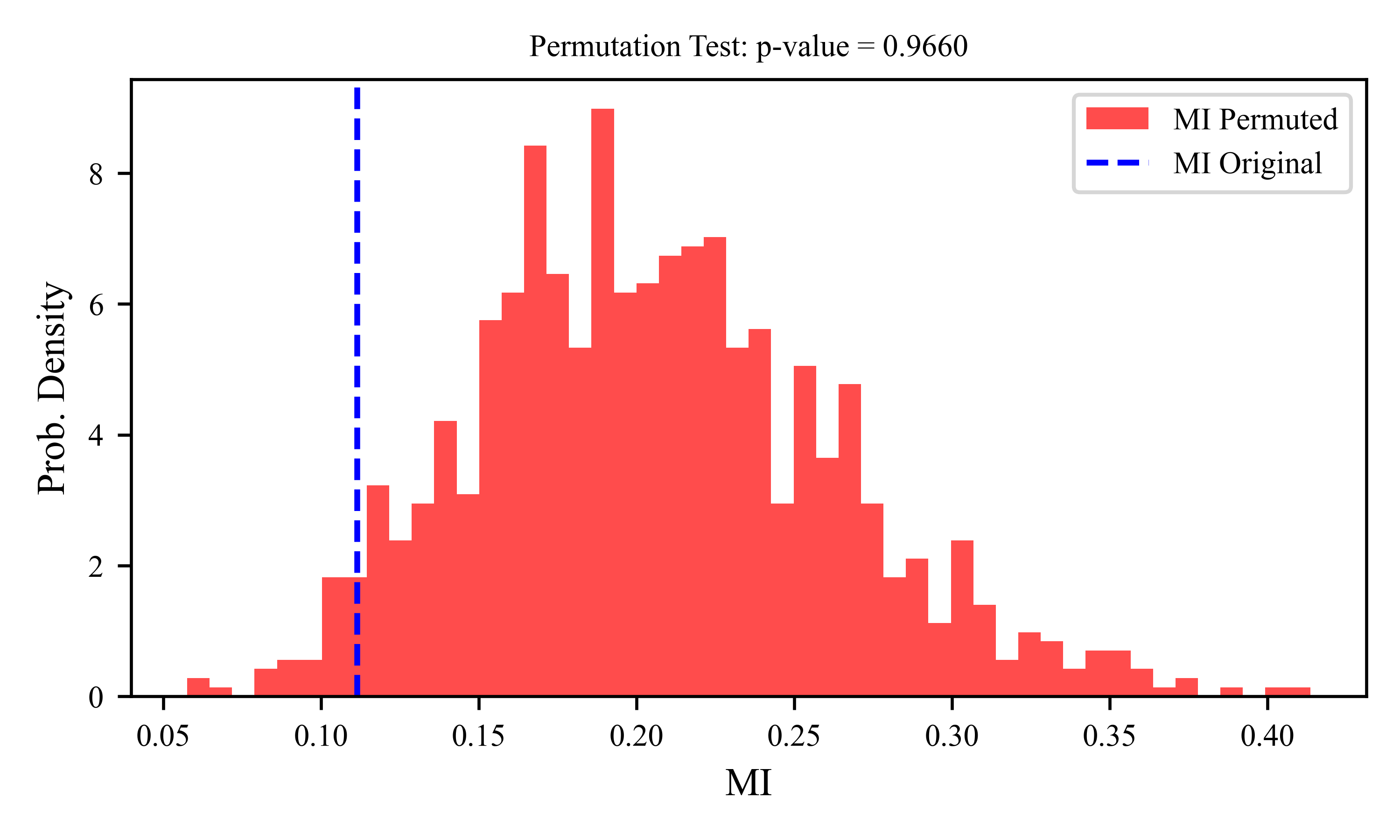


图 3. 式(1)状态转移过程对应第1次试验样本的蒙特卡洛检验结果

### 3.2 马尔可夫链——保序替代样本

式（2）所对应的过程即为1阶马尔可夫过程。对其样本直接按照3.1.2进行蒙特卡洛置换检验，所得I类错误概率暴涨到26%！这说明3.1.2中的方法将很多本应独立的样本误判为非独立。下图显示了式（2）所得的和互信息值实际分布（来自所有试验样本，蓝色）和置换样本的互信息值分布（红色）对比，可见基于3.1.2节方法所得置换样本的零假设互信息分布明显偏低，导致检验的I类错误率上升。这个结果表明，对于具有马尔可夫性的数据不能简单地采用3.1节中的随机置换检验，否则结果误差严重。

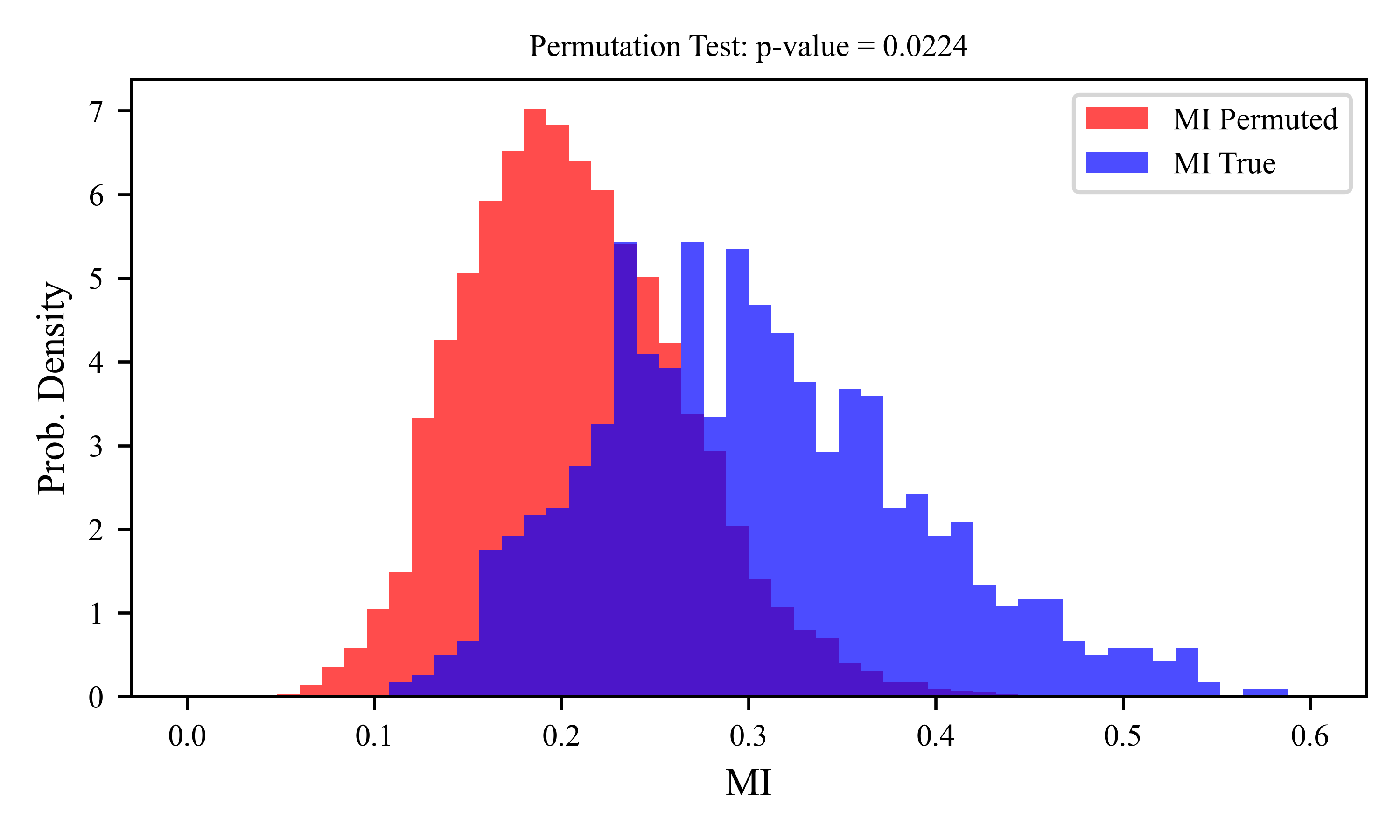


图 4. 无关样本关联值真实分布和置换分布对比

在常见的置换检验中，和样本来自对潜在联合分布的独立采样（i.i.d.），各变量样本内部不具有时序性或马尔可夫性，此时可以通过多次随机打乱样本来生成替代数据，进而进行置换检验来判断和是否独立。但是在时间序列和马尔可夫链的情形下，和的样本序列具有时序性和马尔可夫性，直接打乱样本会破坏其时序结构和依赖关系，丢失本底关联信息。

需要针对马尔可夫链数据构建替代样本（surrogates），对应的保序替代样本（order-preserving surrogates）构造步骤如下：

1. 设定的马尔可夫链阶数，根据原始数据估计状态转移矩阵；
2. 利用估计的转移概率矩阵，从的初始状态出发，按照马尔可夫过程生成新的状态序列，作为替代数据，然后计算互信息作为零假设样本；
3. 多次重复步骤2，获得替代样本互信息的经验分布，计算值并判断独立性。

下图显示了关联值真实分布和替代样本所得分布的对比，可见二者吻合良好。在100次掷骰子试验中，I类错误率为0。

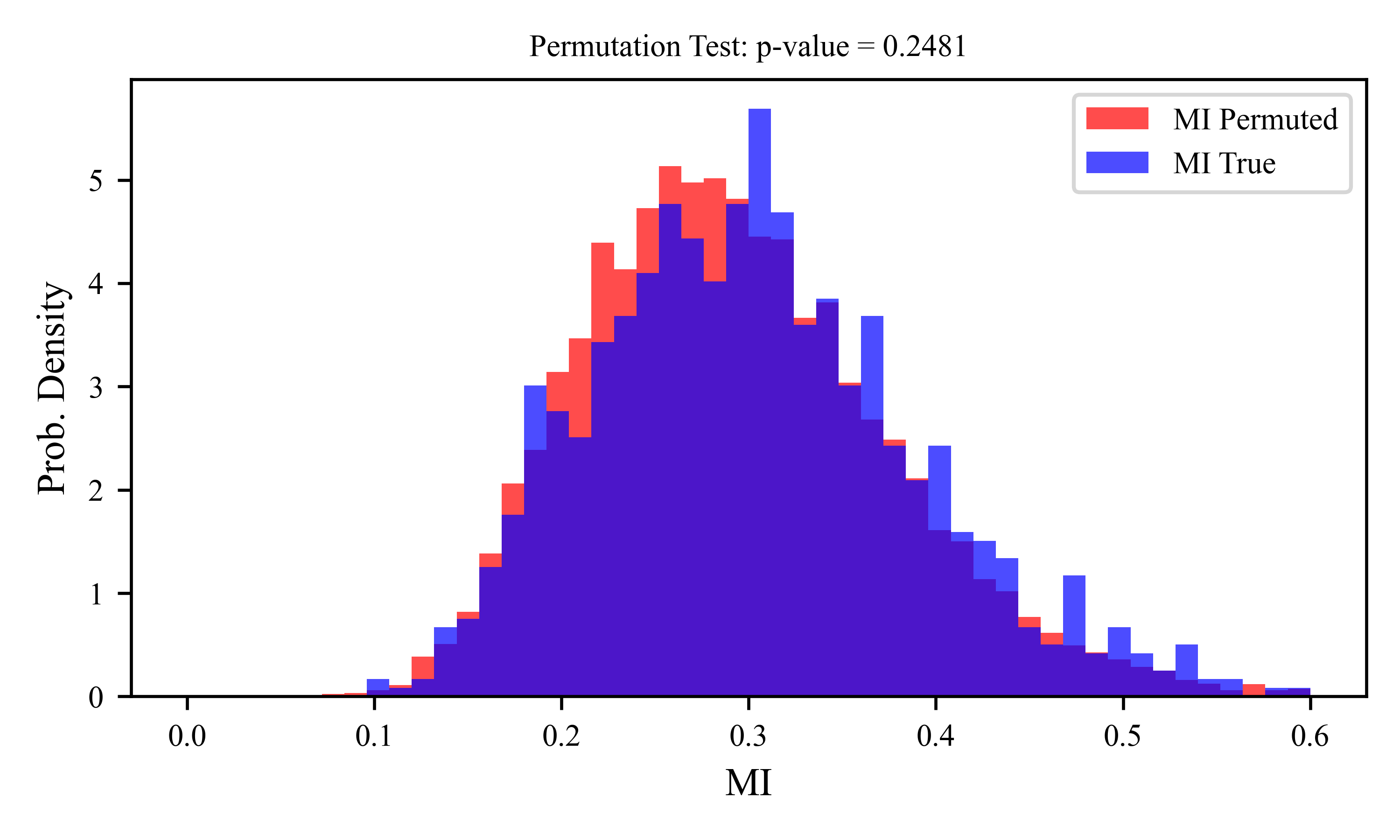


图 5. 无关样本关联值真实分布和替代样本所得分布对比

## 四、马尔可夫链定阶

上述检验的前提是准确确定马尔可夫链的阶数。这里通过破坏高阶依赖生成的替代数据，检验**阶假设是否被违反。步骤如下：

1. 计算观测序列的阶统计量（如块熵**）；
2. 生成阶保序替代数据，计算零假设统计量分布；
3. 对比阶统计量和阶统计量分布差异，判定是否接受阶的假设。

块熵是信息论中用于衡量序列不确定性的指标，描述长度为的连续符号块（block）的熵，反映了系统在特定时间尺度下的复杂度，计算公式为：

 (7)

如果序列完全随机，则；否则，。如果在时趋于稳定，说明阶马尔可夫性成立。

## 参考文献

1. T. Schreiber, A. Schmitz: Improved Surrogate Data for Nonlinearity Tests. Physical Review Letters, 1996.
2. T. Schreiber, A. Schmitz: Surrogate time series. Physica D, 2000.
3. S.D. Pethel, D.W. Hahs: Exact Test of Independence Using Mutual Information. Entropy, 2014.