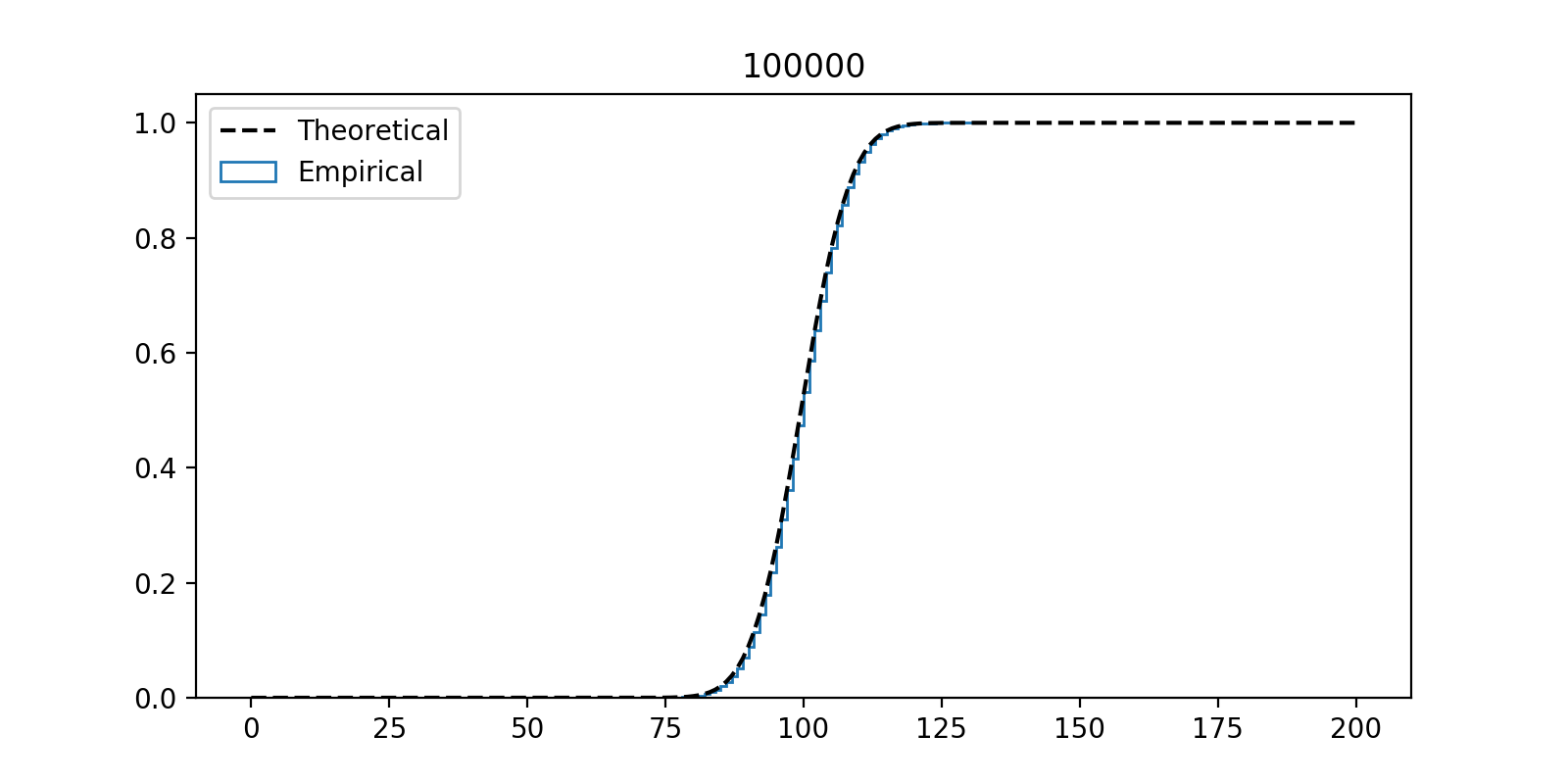
数学实验 统计量和MC算法 实验报告

计65 赖金霖 2016011377

一

1. 独立抛掷200次硬币，正面个数的近似分布p(x=i) 满足:

可以使用np.random.randint函数随机0和1，来模拟投掷硬币。随机进行100000次实验（投掷100000×200次），得到如下经验分布函数（黑色为近似分布函数）：

数值上F(x)有如下的对比：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | 80 | 90 | 100 | 110 | 120 |
| Theoretical | 0.00284 | 0.08948 | 0.52817 | 0.93132 | 0.99818 |
| Empirical | 0.00274 | 0.08943 | 0.52927 | 0.93167 | 0.99826 |

可以看出，经验函数的估计十分准确，误差均不超过5%。

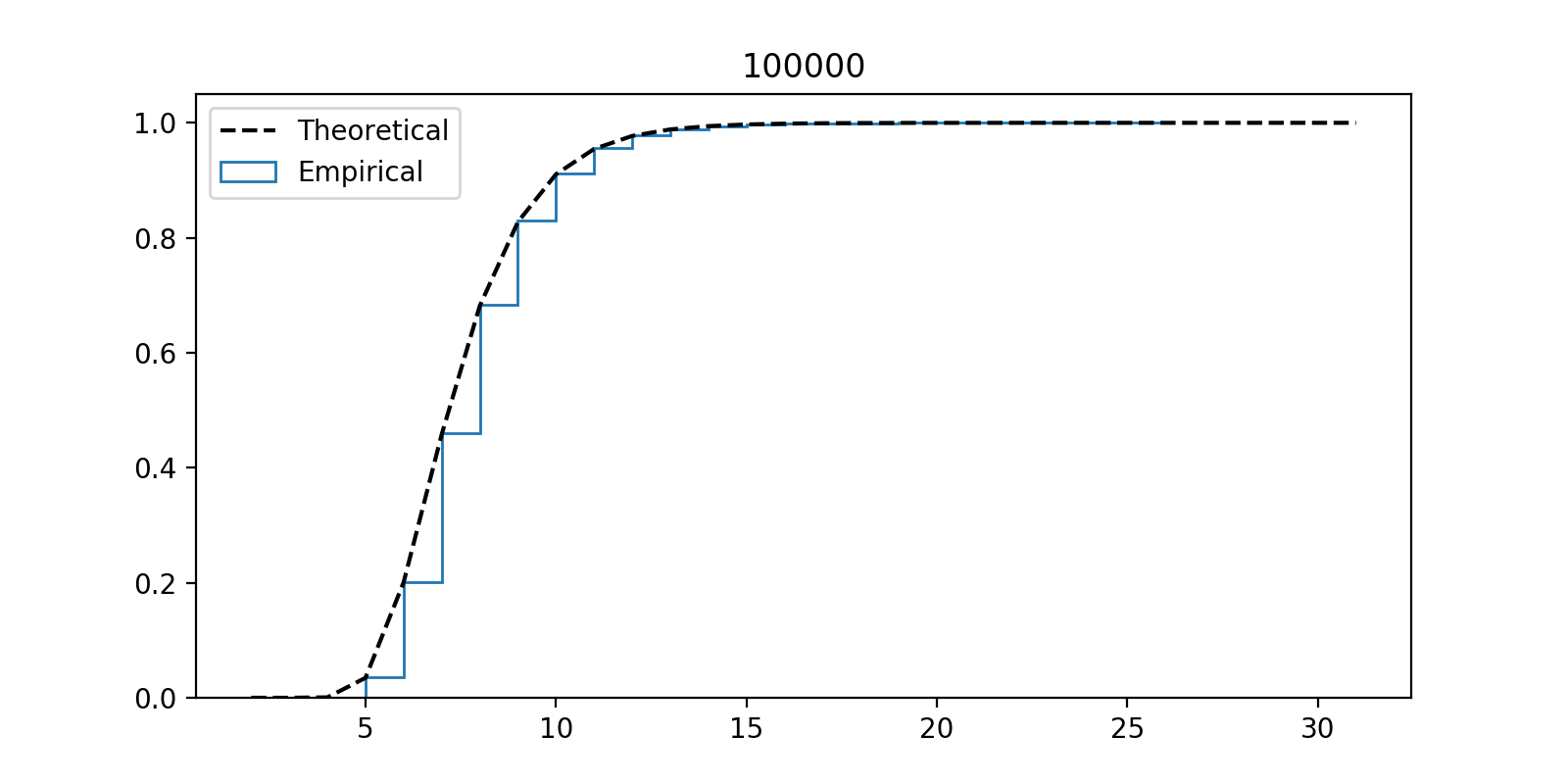
1. 独立抛掷硬币k次，设最长0串个数X的分布的cdf为Fk(x)，最长0或1串长度Y的分布的cdf为Gk(y)，其中x=0,1,…,200，y=1,2,…,200，那么我们有Gk(y)=Fk-1(y-1)。

这是因为长度为k的01串的最长0或1串和长度为k-1的01串的最长0串有对应关系：设有一长度为k的01串A[1~k]，B[i]=(A[i]等于A[i+1])，那么A中的0或1串对应B中的0串；反之，对长度为k-1的01串B[1~k-1]，也有对应的A串，使B的0串对应A的0或1串。

根据https://math.stackexchange.com/questions/59738/probability-for-the-length-of-the-longest-run-in-n-bernoulli-trials，Fk(y)满足



我们可以直接算出Gk(x)的分布函数。

当k=200时，随机进行100000次实验，最长0或1串的分布如下所示

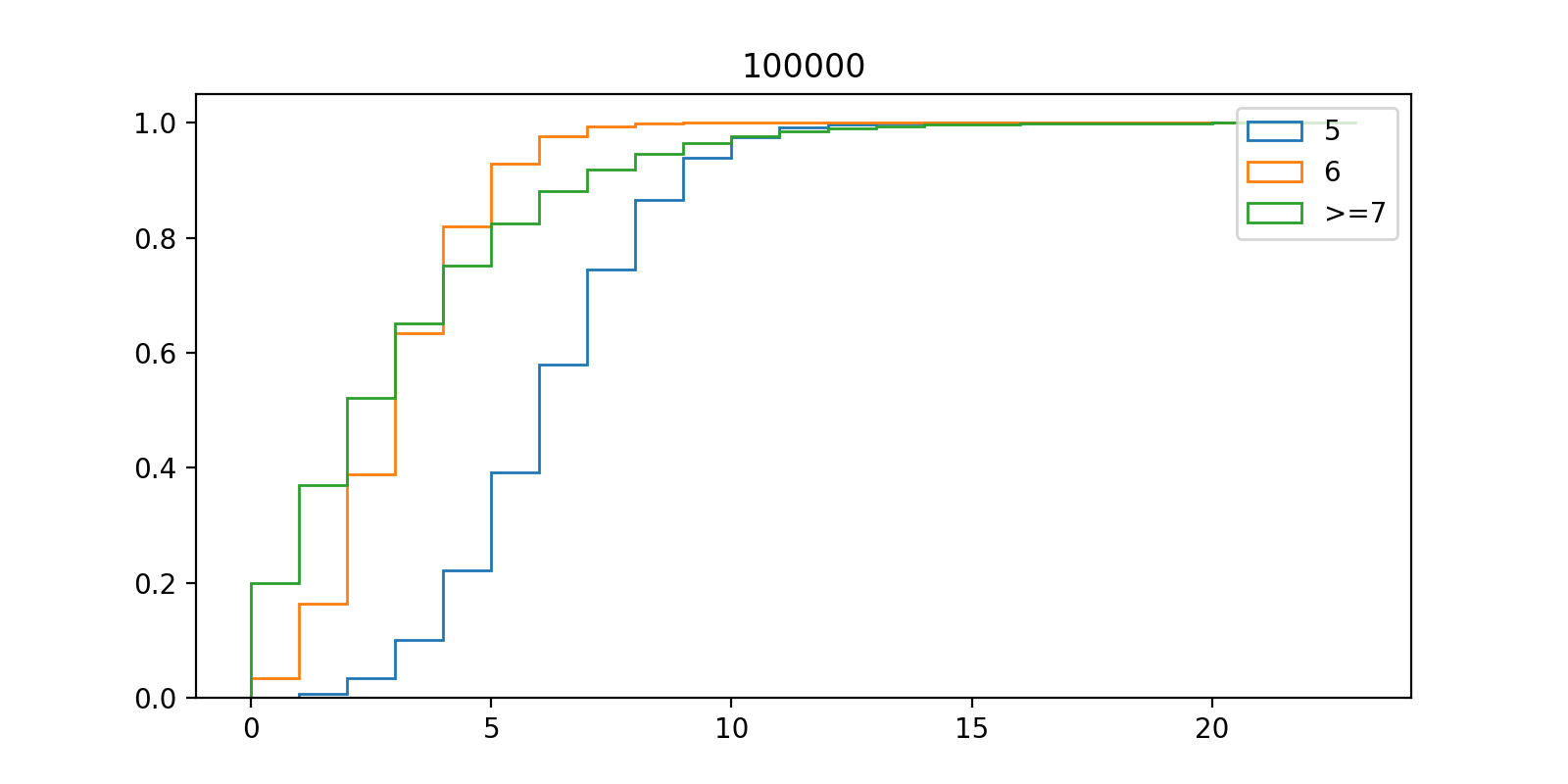
根据Dvoretzky–Kiefer–Wolfowitz 不等式，在样本数为n的情况下，经验cdf F\*n(x)和理论cdf F(x)的差距满足

在实验中，若取z2=0.1，那么我们有

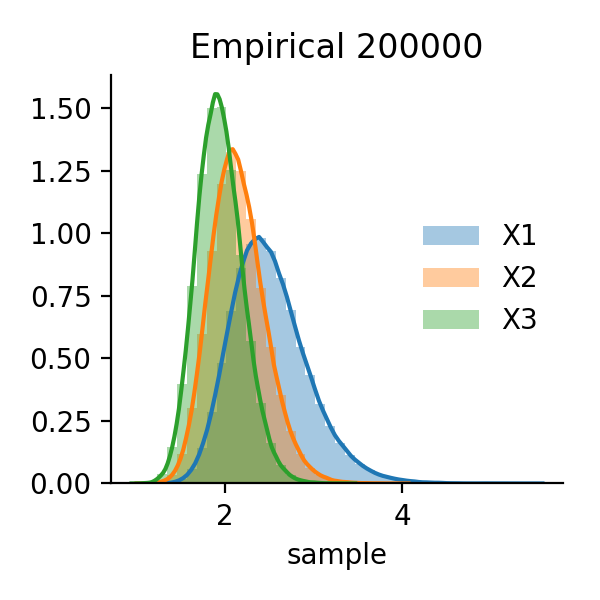
G200(x)的实验分布和理论分布数值如下

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | 6 | 7 | 8 | 9 | 10 |
| Theoretical | 0.20072 | 0.45813 | 0.68165 | 0.82760 | 0.91053 |
| Empirical | 0.20167 | 0.46006 | 0.68246 | 0.82683 | 0.91042 |

结果在理论上和时间上都特别准确。

1. 和1，2类似，进行100000次实验，长度为5、6和不小于7的0或1串的个数的经验分布如下所示（000000里有2个长度为5的0串）

可以明显看出，在连续200次抛掷硬币的实验里，有很大的概率出现连续5个或以上的连续相同串。

二、使用程序随机抽取100个标准正态分布并排序，X(1)，X(2)，X(3)是最大的三个数，可以看做一次实验。进行200000次实验，可以得到如下的的经验分布（pdf）

明显可以看出，X(1)、X(2)、X(3)三者，均值越来越小，方差越来越大。

有如下的数值（方差经过修正）：

|  |  |  |  |
| --- | --- | --- | --- |
| X | 1 | 2 | 3 |
| 均值 | 2.50869 | 2.14869 | 1.94711 |
| 方差 | 0.42952 | 0.30963 | 0.26186 |

三、

5. 炮弹落点X满足

带入数字，使用np.random.multivariate\_normal进行1000000次采样，在|X|≤100时视为炮弹落入圆形区域，有炮弹命中概率为0.698834。

6. 容易算出，球面体积为2/3π，而对锥面体积，使用积分方法有

所以理论上体积为π。

使用MC方法，则可以随机在[-1,1]×[-1,1]×[0,2]空间上采点，其在冰淇淋内的频率×空间体积8，即为冰淇淋的体积。进行1000000次采样，可以算出冰淇淋的体积为3.139952，与理论数值有一定差距。

四、参考资料

<https://en.wikipedia.org/wiki/Empirical_distribution_function>

<https://math.stackexchange.com/questions/59738/probability-for-the-length-of-the-longest-run-in-n-bernoulli-trials>

五、代码

代码路径：

https://github.com/lll6924/math\_exp/blob/master/exp7/statistics\_mc.py