

# 一、足球战术可视化

**任务：**根据上赛季若干场红蓝两队交战数据，分析双方战术优势与短板，并为红队不久将来与蓝队的交战提出建议。

**输入数据：**上赛季的4场联赛数据和1场杯赛数据，在每一时间点（间隔为0.2秒）下所有球员、足球的坐标数据，以及该时刻持球者的编号。

## 0. 引入必要的包

对于这一问题，我的想法较多，如绘制热力图，分析传接球次数等。首先引入必要的包，需要使用pip下载，如pandas、cv2、numpy、pyheatmap、scipy、matplotlib等，以便接下来的分析。

## 1. 读入数据，获取交战比分

由于数据特性，进球有效后足球会被放置在坐标(0, 0, 0)处，我们只需要统计球被放置在该坐标处的次数，以及持球人所在的队伍，即可得到交战比分。

```
3:2 ['红', '红', '红', '蓝', '蓝'] //联赛主场1
2:2 ['红', '蓝', '红', '蓝'] //联赛客场1
1:2 ['蓝', '蓝', '红'] //联赛客场2
0:1 ['蓝'] //联赛主场2
1:3 ['蓝', '蓝', '红', '蓝'] //杯赛第1回合
1:0 ['红'] //杯赛第2回合
```

四场联赛1胜1平2负，杯赛第一场1:3败北，第二场虽1:0点球绝杀获胜，但没能改变被淘汰的命运。

可以看出，红队先手进球时，结果往往是红胜或平局；而蓝队先手进球后，红队的士气会受到影响。因此**红队擅长踢顺风球，拿下先手进球权十分重要。**

接下来我们通过输出DataFrame来观察数据的结构。

	时间	红方 0	红方 1	红方 2	红方 3	红方 4	红方 5	红方 6	红方 7	红方 8	...	蓝方 4	蓝方 5	蓝方 6	蓝方 7	蓝方 8	蓝方 9
0	00'00"00	50 0	120 -320	120 320	225 0	300 -80	350 80	450 -300	450 300	500 -120	...	-250 -120	-300 0	-450 300	-450 -300	-500 120	-50 -12
1	00'00"12	31 0	103 -321	101 321	210 1	286 -84	335 79	434 -302	435 302	485 -118	...	-255 -106	-314 0	-441 288	-440 -287	-491 111	-49 -11
2	00'00"24	9 0	83 -321	80 321	193 2	271 -88	318 78	416 -305	417 305	468 -116	...	-260 -90	-330 0	-429 274	-427 -272	-479 101	-47 -99
3	00'00"36	-1 0	66 -321	61 322	178 2	257 -92	302 81	399 -307	401 307	453 -114	...	-265 -75	-343 0	-419 262	-415 -259	-468 93	-46 -91
4	00'00"48	-3 1	53 -311	42 321	163 3	244 -96	287 84	383 -307	385 309	437 -115	...	-270 -61	-355 -2	-408 250	-403 -247	-463 88	-46 -94
5	00'01"00	-21 1	44 -297	24 317	149 3	231 -97	272 85	367 -301	370 304	422 -114	...	-270 -66	-360 -13	-398 237	-388 -245	-470 83	-46 -10
6	00'01"12	-40 2	34 -283	6 313	136 2	218 -98	257 83	351 -298	356 297	407 -114	...	-266 -80	-355 -26	-391 223	-386 -258	-469 75	-45 -11
7	00'01"24	-58 3	25 -268	-10 318	123 0	205 -102	242 80	335 -300	343 289	393 -117	...	-264 -95	-349 -39	-389 211	-395 -262	-465 71	-45 -12
8	00'01"36	-76 4	17 -253	-29 322	110 -5	192 -106	227 77	319 -303	329 280	379 -120	...	-265 -110	-343 -52	-392 210	-388 -263	-471 66	-45 -13
9	00'01"48	-95 5	10 -238	-47 323	96 -6	180 -110	213 73	303 -307	315 272	366 -124	...	-266 -125	-339 -65	-391 206	-384 -272	-469 62	-46 -13
10	00'02"00	-113 8	2 -225	-66 325	82 -5	168 -115	198 69	287 -311	302 264	352 -128	...	-268 -140	-335 -78	-390 197	-391 -273	-466 57	-46 -14
11	00'02"12	-131 11	-2 -236	-85 325	67 -5	156 -120	183 68	272 -314	288 256	339 -133	...	-270 -155	-332 -92	-391 197	-384 -276	-470 53	-46 -14
12	00'02"24	-149 14	1 -252	-103 326	53 -4	145 -127	170 73	256 -318	274 248	325 -137	...	-271 -170	-330 -106	-391 193	-378 -282	-469 51	-46 -14
13	00'02"36	-167 18	1 -269	-122 327	39 -3	136 -136	156 79	241 -321	260 241	312 -141	...	-273 -185	-328 -120	-391 187	-381 -283	-467 47	-46 -15
14	00'02"48	-185 22	1 -286	-140 327	25 -2	124 -143	142 75	225 -324	246 234	299 -146	...	-274 -200	-326 -134	-388 189	-373 -290	-469 44	-46 -15
15	00'03"00	-203 26	1 -303	-159 328	10 0	111 -147	128 73	209 -326	231 227	287 -151	...	-274 -215	-325 -147	-374 195	-369 -296	-469 42	-46 -15
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

19173 rows × 26 columns

可以看到球员位置的坐标是二维的，而足球的坐标是三维的。

下一步我们可以可视化球员与足球运动的热力图与站位的点击图。

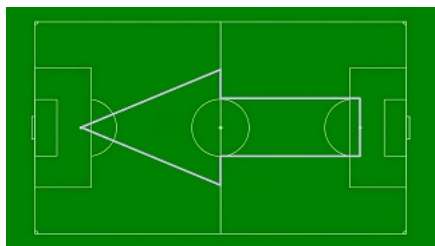
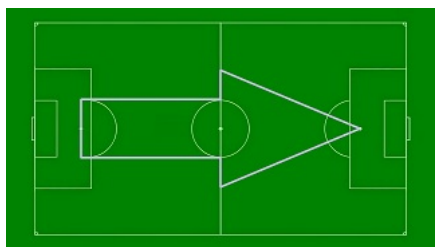
## 2. 判断球员位置

根据球员活动区域与行为，大致可以将球员分为以下10类：

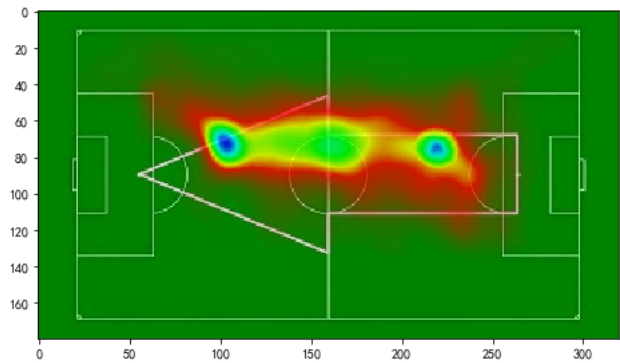
- 中锋
- （左/右）边锋
- 影锋
- 前腰
- （左/右）边前卫
- 中前卫
- 后腰
- （左/右）边后卫
- 中后卫
- 守门员

利用热力图可视化其活动区域，可以大致判断球员的位置。

接下来准备背景图，用于叠加到热力图中，方便观察。

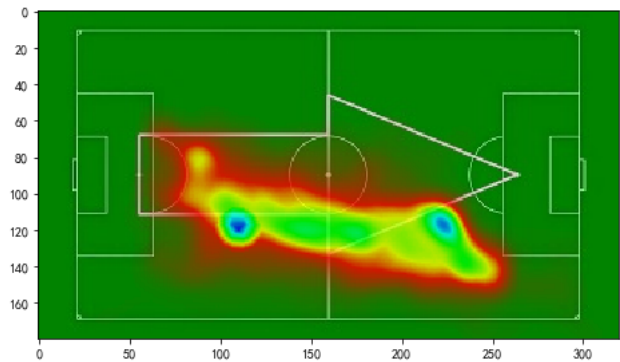


首先可视化红队11名球员的活动热力图（在此只列出一张图片，详见ipynb文件）。



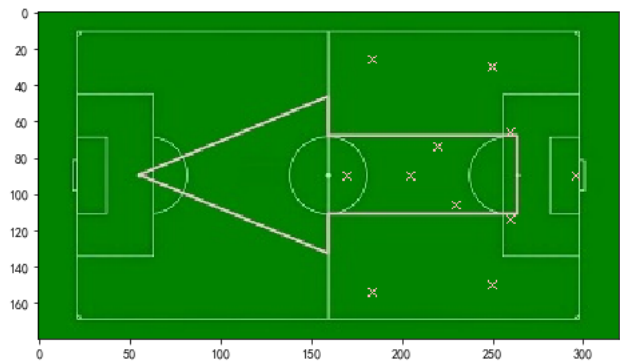
可以看到相对明显的4-3-3站位，0号为中锋，1、2号分别为左右边锋，3-5号为中场，6、7号为左右边后卫，8、9号为中后卫，10号为门将。

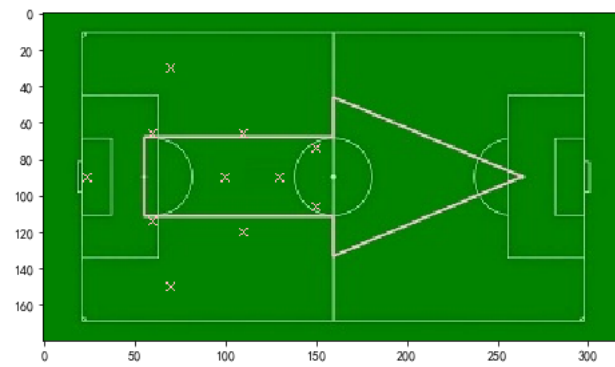
同样的方法可视化蓝队的球员活动区域。



根据热力图我们很难确定其站位。像是4-4-2的站位，但双中锋活动区域并没有深入对手防线；又像是4-2-3-1的站位，但影锋的回防未免太深了。可见蓝队的阵型布局更具迷惑性。

接下来我们绘制出比赛开始时所有球员的站位，来验证刚刚的猜想。



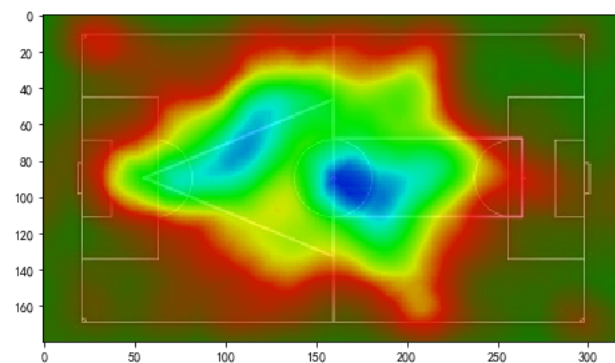
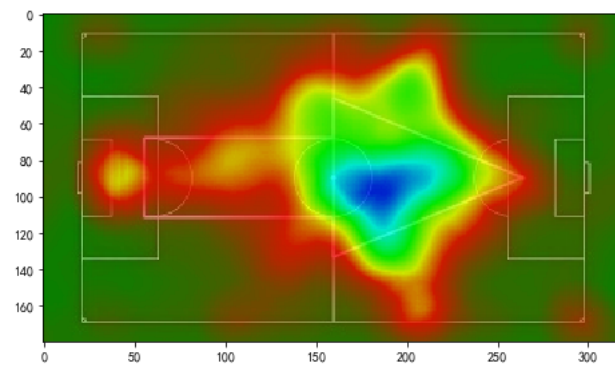
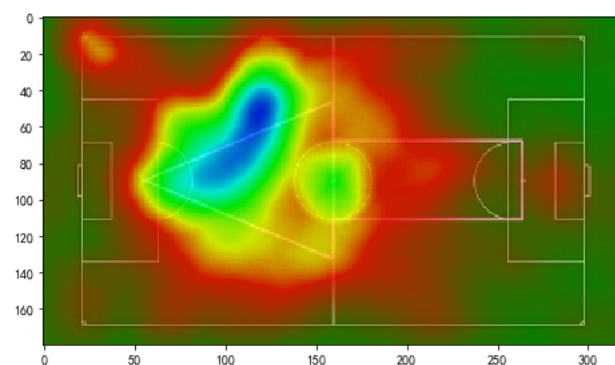


从点击图中，可以看出红队确实是4-3-3，或者说是4-2-1-3站位的；而蓝队则是4-3-1-2站位。蓝队两前锋虽并不是完全的进攻球员，但活动区域较小且行为相似，因此都算作中锋；而前腰则更像影锋，活动区域较大；两个边后卫的行为也不尽相同，战术较红队变化更多。

### 3. 交战区域

#### 1. 进攻区域

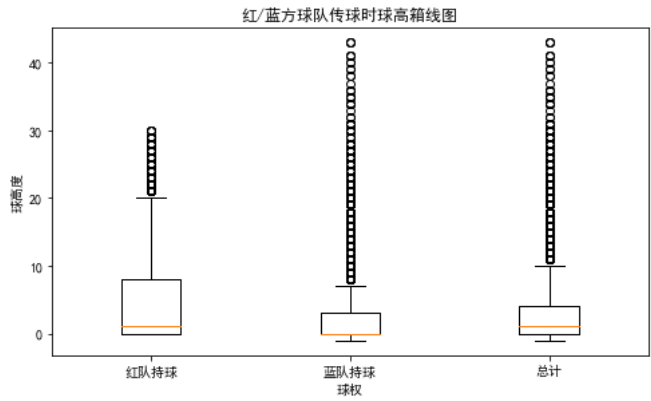
可视化足球在红方、蓝方分别持有球权时的活动区域的热力图，以分析其进攻区域。



看得出，红队的进攻区域更加宽广，在边路配合流畅；而蓝队擅长边路佯攻，而进攻核心在中路。整体形势来看，红队目的性更弱一些，进攻战术以传控为主；而蓝队目的性更强一些，进攻战术更偏重于防守反击。

## 2. 进攻方式

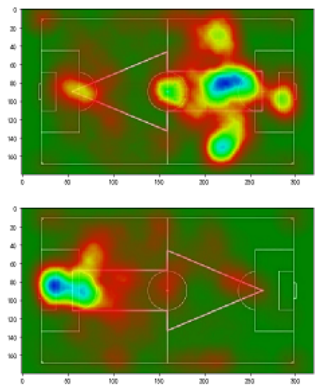
由于数据中还含有足球的高度信息，我们可以根据这一信息来判断球队的进攻方式，即进攻组织为长传还是短传。对此可以绘制箱线图来解决。



可以看出红队是稳定的高空球战术，而蓝队则更加重视地面球，除少数长传和解围外，极少传递高空球。

## 3. 防守区域

通过热力图可视化抢断、拦截传球的位置，来判断球队在哪些区域布有重兵防守。抢断和拦截传球的特点是球权交换，可以根据这一特点完成过滤。

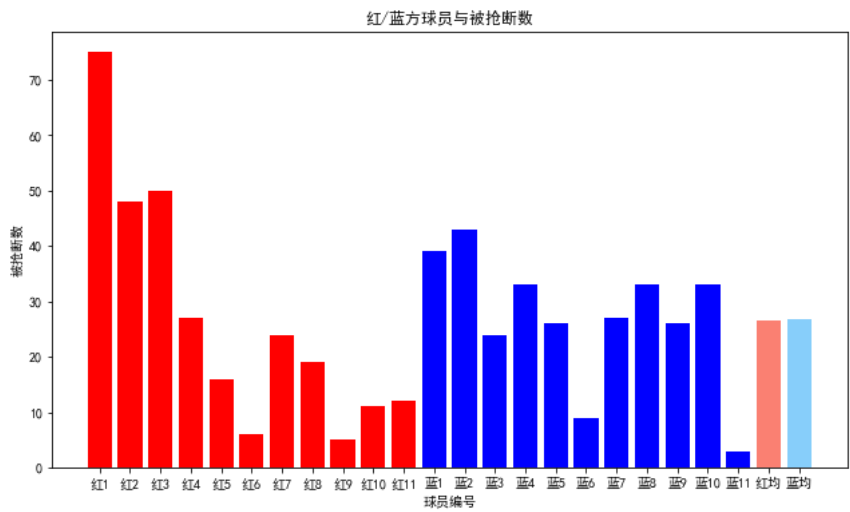


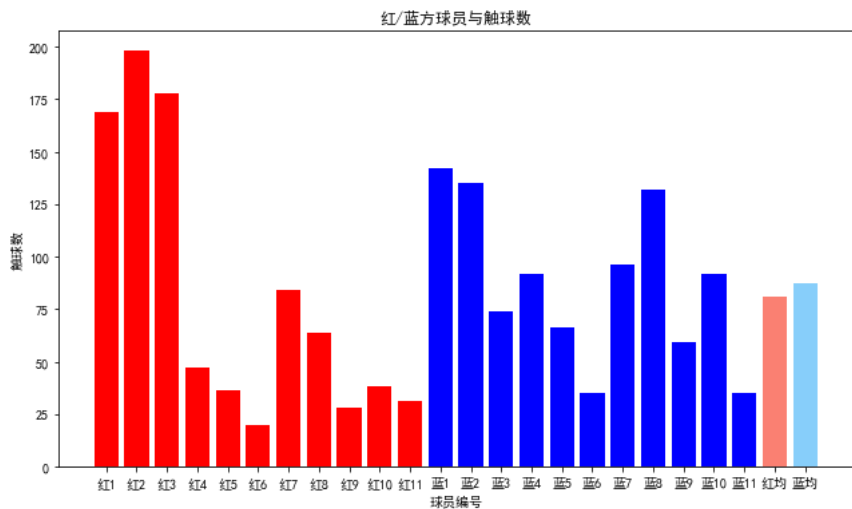
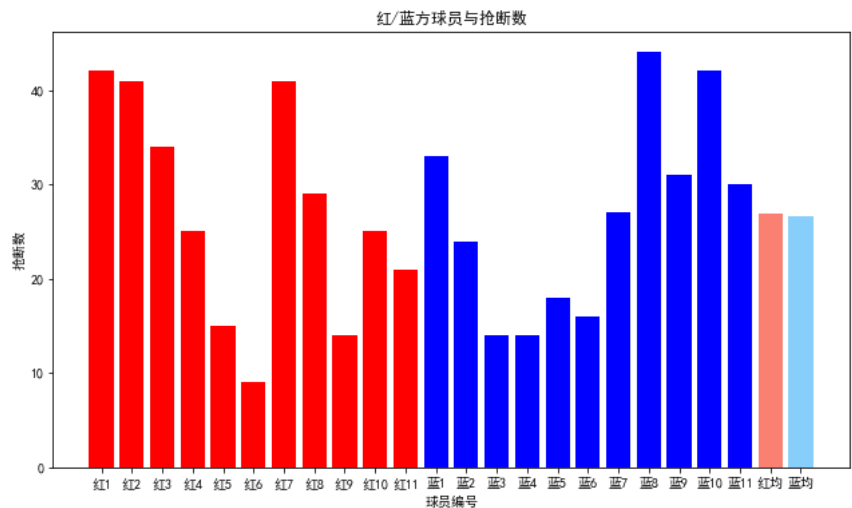
可以看出，两队重兵把守区域与进攻区域相似。红队在边路的抢断更多一些，蓝队则在中路的抢断更多一些。除此之外，我们发现，红队擅长前线施压夺球，而蓝队则更偏重于全攻全守。结合球员跑动热力图，我们证实了此前分析得到的结论——蓝队的战术可以概括为12字，**请君入瓮，瓮中捉鳖，快速反击。**

## 4. 球员个人能力分析

除去分析球队战术外，我们还可以分析球员的个人能力。

首先我们可以分析每个球员被抢断、主动抢断和触球的次数，并绘制条形图。





可以看出，红方9号和红方8号虽然同为边后卫，在热力图中跑动区域也相似，但9号更少参与进攻与防守。而红方后腰7号则为红队防守核心人物。

由于红方战术要求球员更多地就地反抢，造成了前锋拥有较多的抢断数。

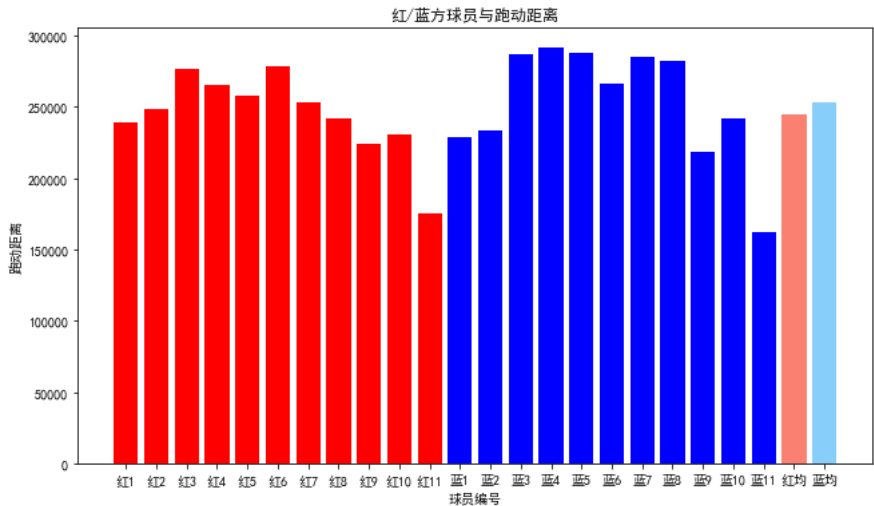
而红方三前锋更是进攻核心球员，从热力图中也可以看出，他们回撤到中圈接球，然后在中前场打出配合，进而创造出机会。而蓝方的配合则更加平均。

概括一下红蓝两方的战术：

- **红方：分工明晰，配合默契。**
- **蓝方：全民皆兵，防守反击。**

## 除此之外做过的其它尝试

除上述分析外，我还尝试过利用其它信息进行分析，如跑动距离、控球率等。



红方控球率：48.43%

蓝方控球率：51.57%

至此对于两队的战术分析已经结束。

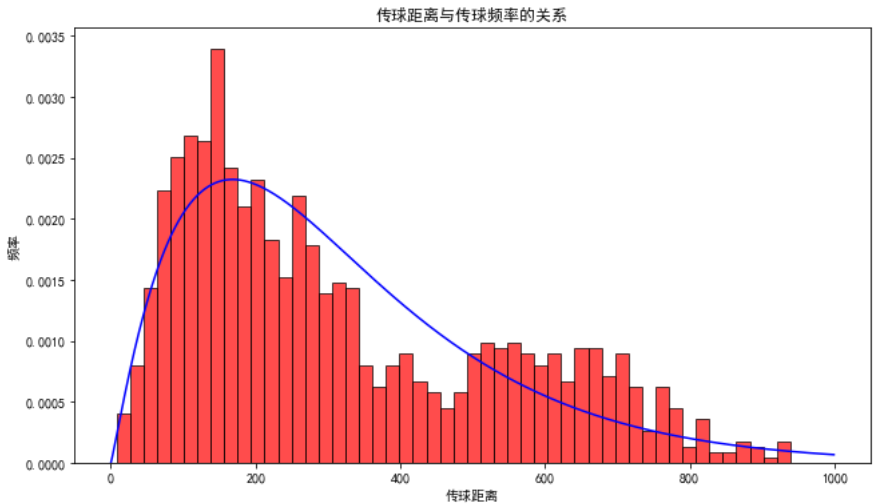
## 二、球员行为预测模型

有了这些数据，我们可以分析的绝对不止是以上信息。比如我们还可以建立统计模型，预测球员的行为，完成类似于“球赛模拟器”的软件。

### 1. 传球预测

假设球员的传球仅仅受持球球员的位置、目标球员的位置、球员间的距离、对方球员与传球路线的距离四个因素影响。

比如我们可以可视化所有发生传球行为时持球球员与目标球员间距离的分布情况。

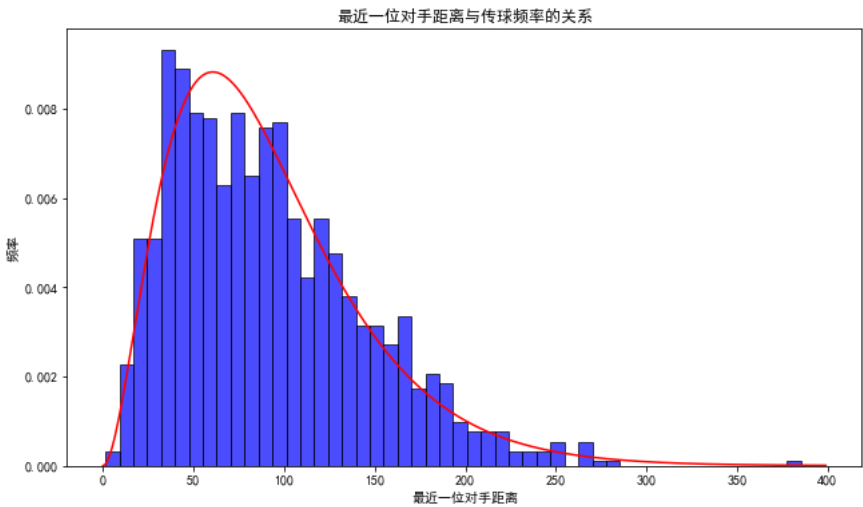


$\alpha=2.110576$ ,  $\beta=0.006605$

可以看出，上图平滑曲线是一个偏态分布——可以认为传球距离服从伽马分布，并且可以根据矩估计求出参数的估计值。

由于伽马分布参数的极大似然估计没有解析解，因此无法使用EM算法处理双峰问题，也无法分开预测短传与长传的参数。

接下来我们可视化传球发生时对方球员位置的分布情况。



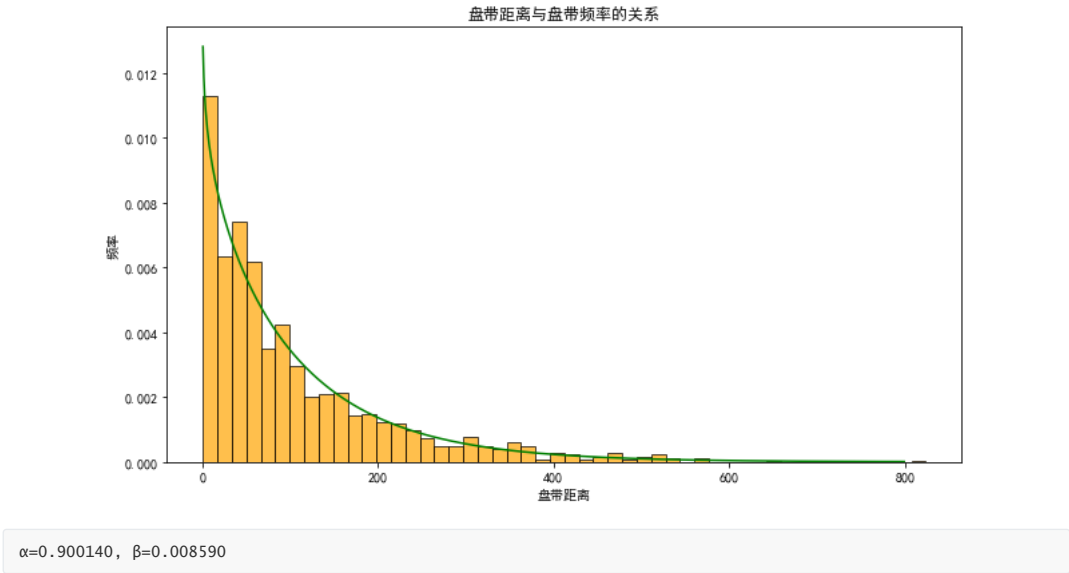
$\alpha=2.961396$ ,  $\beta=0.032346$

同样可以认为最近一位对手的距离近似服从伽马分布。

### 盘带预测

假设球员的盘带仅仅受持球球员的位置、对方球员位置和队友球员位置三个因素影响。

比如我们可以可视化所有发生盘带行为时盘带距离的分布情况。



有了这些分布，我们就可以构建指数回归模型，来实现球员行为分类预测。

经过上述分析，我们可以构建出球员行为预测模型。

### 三、球员行为预测模型

由于已知球员决策（后验分布）服从伽马分布，且根据直觉，其先验分布也服从伽马分布。由于伽马分布是指数分布的共轭先验分布，那么我们不妨设影响球员决策的战术因子服从相互独立的指数分布，且参数为上述决策因子。

球队战术则可以看作是由上述分布产生的一个采样，因此每一个球员行为的得分都可看作是战术因子对球队战术的似然（的线性变换结果）。

#### Public 属性

ushort	keepForm	保持队形: [1, 10] 更多...
ushort	longPassMoti	长传积极性: [1, 10] 更多...
ushort	wingSideMoti	分边积极性: [1, 10] 更多...
ushort	advanceSpeed	推进速度: [1, 10] 更多...
ushort	defenseLine	防线高度: [1, 10] 更多...
ushort	supportDist	支持距离: [1, 10] 更多...
ushort	dribblingMoti	盘球积极性: [1, 10] 更多...
ushort	compactness	紧密度: [1, 10] 更多...
ushort	centerBlock	中路围堵: [1, 10] 更多...
ushort	backwardSpeed	回防速度: [1, 10] 更多...
ushort	pressureMoti	施压积极性: [1, 10] 更多...

球员每一次决策都是选择似然函数最大的一项对应的决策。球员每一帧调用函数计算似然并作出决策，即可完成预测。

Public 成员函数	
void Detect_hand_attack ()	【说明】检测到对手球员手球犯规时，分析球员犯规行为的严重程度。 更多...
void Detect_stand_attack ()	【说明】分析球员犯规时对手球员犯规行为，分析球员犯规行为的严重程度。 更多...
void Detect_defence ()	【说明】分析球员的防守行为，分析球员的防守行为。 更多...
void Detect_Action ()	【说明】分析球员的进攻行为，分析球员的进攻行为。 更多...
void Detect_FrontKick ()	【说明】分析球员的射门行为，分析球员的射门行为。 更多...
bool isGrounded ()	判断球员是否在地面上。 更多...
bool isInGoal ()	判断球员是否在球门内。 更多...
bool isFrontField ()	判断球员是否在球场的前场。 更多...
bool isMidField ()	判断球员是否在球场的中场。 更多...
bool isBackField ()	判断球员是否在球场的后场。 更多...
bool isWing ()	判断球员是否在球场的边路。 更多...

除去模型层面外，物理层面也十分重要。球员因为个人能力问题，可能会出现传球、射门有偏差的情况，身处尴尬位置也很难完成精确的操作，在代码中则体现为噪声；持球球员要追球，但这是一个三维追及问题，所以也需要解多元高次方程。



void Update ()	【说明】 每一帧分帧按照场上情况，按照逻辑更新球员行为。 更多...
void Start ()	第一帧调用，在Player与Team刚被创建时，按照各自位置初始化队重心跟新球员的位置和尺寸。 更多...
void turn(Vector3 ballDirect)	【说明】 转身并面向球 更多...
void dribble(Vector3 dribbleDirect)	【说明】 球点触时 更多...
void pass (Player target, bool smart, bool by45)	【说明】 球点传球 更多...
void shot (Vector3 goal)	【说明】 球点射门 更多...
void clear ()	【说明】 球点清理 球员传球人为阻止传球路线。 更多...
Vector3 detect_turn ()	【说明】 分析传球与射门 分析足球距离多久能够追上 更多...
void follow_ball ()	【说明】 球点跟球 球员按照预设方向传球。 更多...
void detect_dribble ()	【说明】 分析射门 分析球点触得分与进球方向。 更多...
void detect_ground_simple ()	【说明】 分析地面情况 分析球员位置与球得分与传球对象。 更多...

最终成果还是比较符合预期的。

代码文档：文档\文档.sh（或直接打开：文档\html\index.html）

最终成果：足球战术模拟器\Build\足球战术模拟器.exe