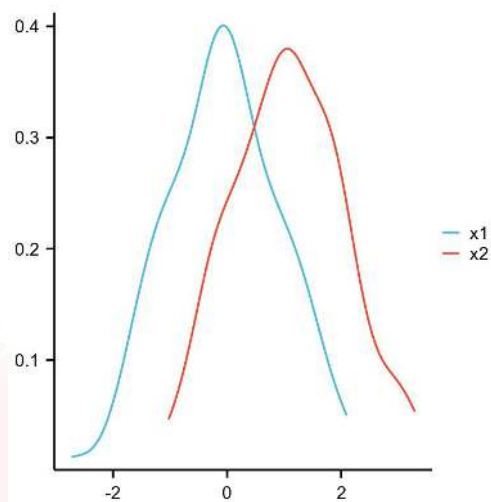


基础绘图 - 数据分布面积图



网址: <https://www.xiantao love>



更新时间: 2023.02.09

目录

基本概念	3
应用场景	3
分析过程	3
结果解读	4
数据格式	5
参数说明	6
样式	6
线	7
面积	8
标题文本	9
图注 (Legend)	10
坐标轴	10
风格	11
图片	11
结果说明	12
主要结果	12
方法学	13
如何引用	14
常见问题	15

基本概念

- 线图：通过线的趋势与高低，可以得到数据的分布趋势与数据的大小
- 面积图：通过线图包围的范围，显示不同数据随变量/时间/类别变化的趋势
线，围成的面积越大表明数据分布越广
- 数据分布图：数据分布图就是通过线+面积的方式来表示数据的分布情况

应用场景

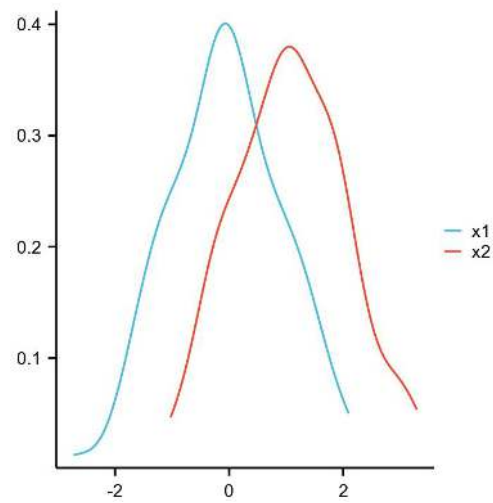
数据分布图主要用来展示数据分布情况。常应用于数据对比。

分析过程

上传数据 ➡ 数据清洗(处理) ➡ 数据分析 ➡ 可视化

- 数据格式：（具体数据格式要求可以看后面过程的“数据格式”部分）
 - 数据每一列都是数值类型，表示每一个分组
- 数据清洗：对数据进行相关处理
 - 非数值类型数据或者其它
- 数据分析：
 - 将清洗过后的数据进行相关分析（计算各分组数据的密度分布）
- 可视化：将分析得到的分布数据进行 ggplot2 包可视化

结果解读



- 横坐标表示分组数据的大小/值，对应每一列分组的数据
 - 样本数据中有 x1、x2 两个分组，横坐标就代表这两个分组的值
- 纵坐标表示分组数据分布密度的大小/值
- 不同颜色对应不同的分组，对应数据每一列分组变量，一个颜色表示一个分组/变量

数据格式

	A	B
1	x1	x2
2	1.19749354	0.40312409
3	-0.3691147	0.96056643
4	-0.3636738	0.43932862
5	1.03600859	2.70332469
6	0.2136652	-0.3732439
7	-0.2525561	0.73997775
8	0.75098549	1.68106813
9	-0.8446256	2.87044482
10	0.29363942	1.22942483
11	0.50584137	1.01204445
12	1.34686193	0.95719011
13	-2.7296267	-0.2510091
14	0.28173707	-0.3491964
15	-0.2683398	2.07035083
16	0.38746216	-0.7588097

数据要求：

- 数据至少 1 列，每列至少 5 个观测(2 行)
 - 也就是说至少需要 1 各分组/变量进行分析
- 最多支持 10 列和 5000 行数据
 - 数据每一列都是数值类型
 - 其它：
 - ◆ 数据中不要上传特殊字符
 - ◆ 数据只能是数值类型的数据，不能上传非数值或者其它
 - ◆ 数据的列名不能重复
 - ◆

参数说明

(说明：标注了颜色的为常用参数。)

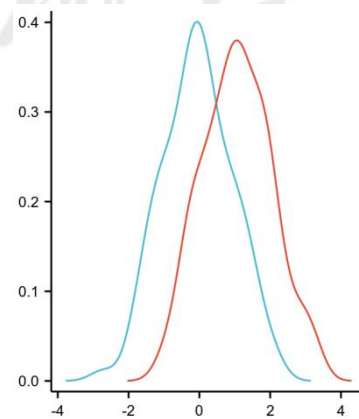
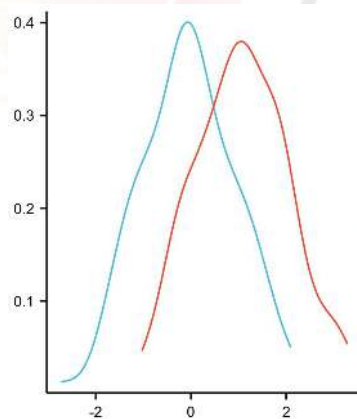
样式



➤ 数据范围截断：可以选择数据分布图的样式(数据范围截断)，默认为不截断，还可以选择截断，如下：左侧为截断，右侧为不截断

■ 截断表示每个密度都是在该组的范围内计算的，默认为截断

■ 不截断表示每个密度都是在整个数据范围上计算的



线

线

颜色

类型

实线

粗细

0.75pt

不透明度

1

- 颜色：可以修改数据分布图线条的颜色
- 类型：可以选择并修改数据分布图线条的类型，默认为实线，还可以选择虚线
- 粗细：可以选择并修改数据分布图线条的粗细，默认为 0.75pt
- 不透明度：可以选择并修改数据分布图线条的不透明度，1 表示完全不透明，0 表示完全透明

面积

面积

是否展示

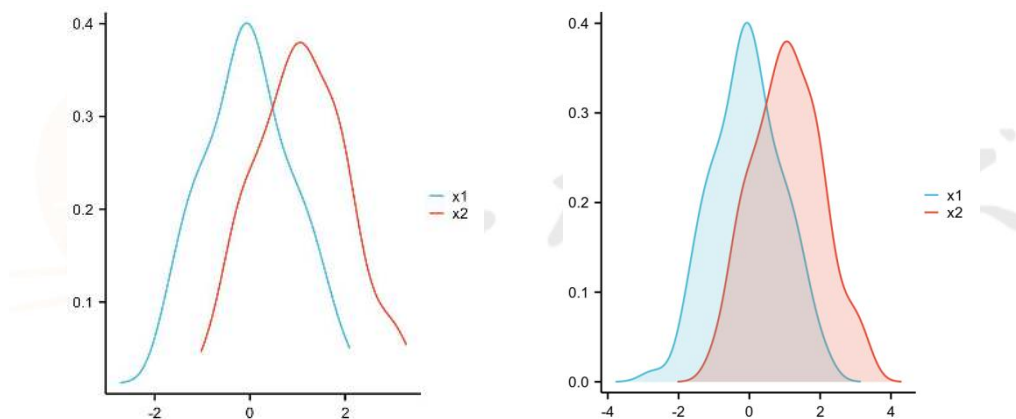
☐

填充色

不透明度

0.2

- 是否展示：可以选择是否对数据分布图进行面积部分展示，默认为不展示，还可以选择展示，如下：左侧为不展示的结果，右侧为展示



- 填充色：可以修改数据分布图面积部分的填充色
- 不透明度：可以修改数据分布图中面积部分的不透明度，默认为 0.2，1 表示完全不透明，0 表示完全透明

标题文本



标题	
大标题	大标题内容
x轴标题	x轴标题内容
y轴标题	y轴标题内容

- 大标题：大标题文本
- x 轴标题：x 轴标题文本
- y 轴标题：y 轴标题文本

补充：在要换行的中间插入\n。如果需要上标，可以用两个英文输入法下的大括号括住，比如 $\{2\}$ ；如果需要下标，可以用两个英文输入法下的中括号括住，比如 $[2]$

图注 (Legend)



图注配置面板，包含以下选项：

- 图注：下拉菜单
- 是否展示：开关按钮（当前开启）
- 图注位置：默认，下拉菜单

- 展示：可以选择是否展示图注操作
 - 选择展示：将会展示图注
- 图注位置：首先选择展示，则可以选择展示图注的位置

坐标轴



坐标轴配置面板，包含以下选项：

- 坐标轴：下拉菜单
- x轴标主旋转：0，下拉菜单
- y轴范围+刻度：逗号隔开

- x 轴标注旋转：可以选择 x 轴标注旋转的角度
- y 轴范围+刻度：可以控制 y 轴范围和刻度，可只提供 2 个值来控制范围。
形如 0.1, 0.2, 0.3（最小值和最大值不能超过可视化数据范围 20%，如果调整过大可能会无作用）

风格



- 边框：可以选择是否进行添加图形边框的操作
- 网格：可以选择是否进行添加图形网格的操作
- xy 颠倒：可以选择是否对图形进行 xy 颠倒的操作
- 文字大小：控制整体文字大小，默认为 7pt

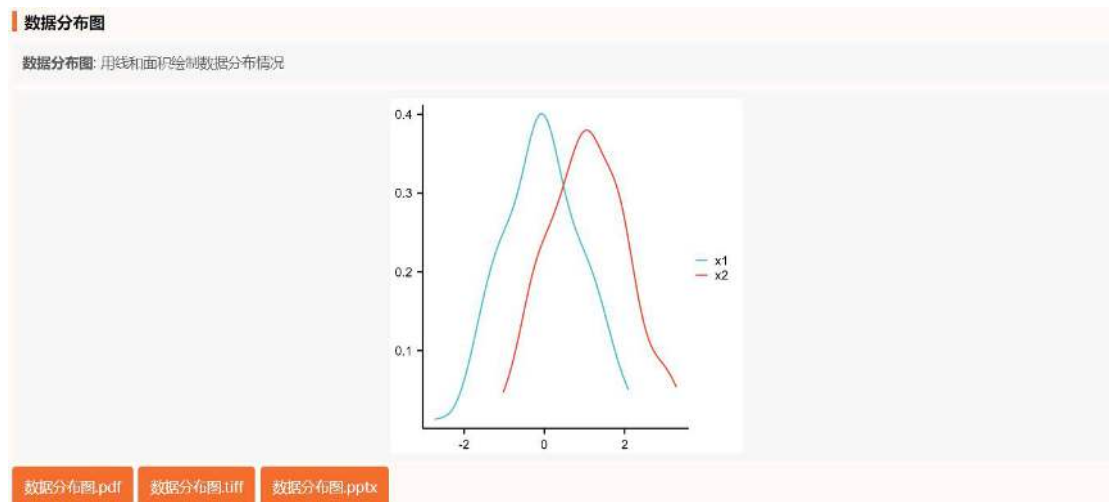
图片



- 宽度：图片横向长度，单位为 cm
- 高度：图片纵向长度，单位为 cm
- 字体：可以选择图片中文字的字体

结果说明

主要结果



方法学

统计分析和可视化均在 R 4.2.1 版本中进行

涉及的 R 包：ggplot2 包（用于可视化）

处理过程：

(1) 使用 ggplot2 包对数据进行数据分布图可视化。



如何引用

生信工具分析和可视化用的是 R 语言，可以直接写自己用 R 来进行分析和可视化即可，可以无需引用仙桃，如果想要引用仙桃，可以在致谢部分 (Acknowledge) 致谢仙桃学术 (www.xiantao love)。

方法学部分可以参考对应说明文本中的内容以及一些文献中的描述。



常见问题

