

附录：直方图和柱状图

2023 年 9 月 8 日

5.5 直方图与柱状图

直方图

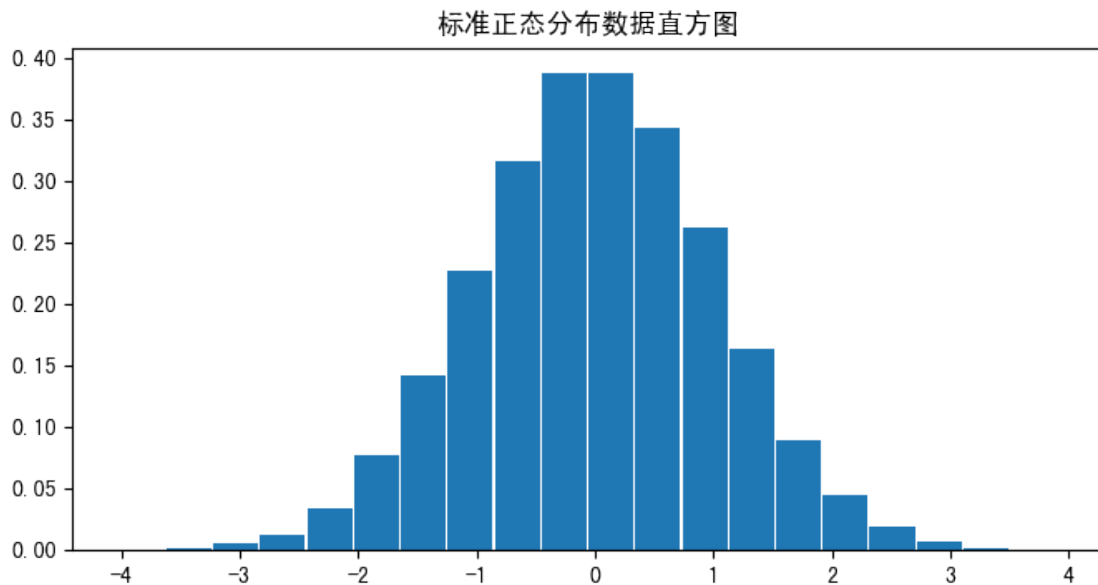
直方图（Histogram）又称频数直方图，由一系列宽度相等、长度不等的长方形来展示特征的频数情况。长方形的宽度表示组距（数据范围的间隔），长度表示在给定间隔内的频数（或频率）与组距的比值，以长方形的面积来表示频数（或频率）。由于分组数据具有连续性，直方图的长方形通常是连续排列的。直方图可以比较直观地展现特征内部数据，便于分析其分布情况。pyplot 中使用 hist 函数绘制直方图，其基本使用语法如下。

```
matplotlib.pyplot.hist(x, bins=None, range=None, density=None,
                        weights=None, cumulative=False, bottom=None,
                        histtype='bar', align='mid', orientation='vertical',
                        rwidth=None, log=False, color=None,
                        label=None, stacked=False, normed=None,
                        hold=None, data=None, **kwargs)
```

参数名称	说明
x	接收 array。表示 x 轴数据。无默认值
bins	接收 int 或 sequence。表示长方形条数。默认为“auto”
range	接收 tuple。筛选数据范围。默认为 None（最小到最大的取值范围）
normed	接收 bool。表示选择频率图还是频数图。默认为 True
rwidth	接收 0~1。表示长方形的宽度。默认为 None

```
[3]: import numpy as np
import matplotlib.pyplot as plt
plt.rcParams['font.sans-serif'] = 'SimHei'
plt.rcParams['axes.unicode_minus'] = False
```

```
plt.figure(figsize = (8, 4), dpi = 100)
mu = 0
sigma = 1
x = np.random.normal(mu, sigma, 10000) # 生成 10000 个服从标准正态分布数据
plt.hist(x, bins = 20, density = True, rwidth = 0.96) # 绘制直方图
plt.title('标准正态分布数据直方图')
plt.show()
```



柱状图

柱状图，也称条形图（Bar Chart）也是由一系列宽度相等、高度不等的长方形来展示特征的频数情况。但条形图主要展示分类数据，一个长方形代表特征的一个类别，长度代表该类别的频数，宽度没有数学意义。相较于面积，肉眼对于高度要敏感许多，故能很好显示数据间的差距。条形图不同类别之间是有空隙的。pyplot 中使用 bar 函数绘制条形图，其基本使用语法如下。

```
matplotlib.pyplot.bar(*args, **kwargs)
```

bar 函数在官方文档中只要求输入不定长参数，但其主要参数如下表所示。

参数名称	说明
x	接收 array。表示 x 轴的位置序列。无默认值

参数名称	说明
height	接收 <code>array</code> 。表示 x 轴所代表数据的数量（长方形长度）。无默认值
width	接收 0~1 之间的 <code>float</code> 。指定直方图宽度。默认为 0.8
color	接收特定 <code>str</code> 或者包含颜色字符串的 <code>array</code> 。表示直方图颜色。默认为 <code>None</code>