Bernoulli分布

目录

[Bernoulli分布 1](#_Toc85550445)

[1. 概要 1](#_Toc85550446)

[2. 概率分布 2](#_Toc85550447)

[2.1 概率质量函数 2](#_Toc85550448)

[2.2 概率分布函数 2](#_Toc85550449)

[2.3 生存函数，风险函数 2](#_Toc85550450)

[2.4 百分点函数 3](#_Toc85550451)

[2.5 示例代码 3](#_Toc85550452)

[3. 常用统计特征 3](#_Toc85550453)

[3.1 均值，Mean 3](#_Toc85550454)

[3.2 中位数，Median 3](#_Toc85550455)

[3.3 众数，Mode 3](#_Toc85550456)

[3.4 方差，Variance 4](#_Toc85550457)

[3.5 偏度，Skewness 4](#_Toc85550458)

[3.6 峰度，Kurtosis 4](#_Toc85550459)

**[Notations and Abbreviations]**

# 概要

设试验E只有两种可能的结果，这种试验称为伯努利试验。伯努利试验的结果为一个随机变量，它遵循伯努利分布。

伯努利分布也称为(0-1)分布，遵循伯努利分布的随机变量只有两种取值，分别用0和1表示（分别对应试验的两种结果）。典型的例子是扔硬币实验结果。将扔硬币的实验结果（正面向上或向下）看作是一个随机变量X，则X遵循伯努利分布，这是一种离散分布。记为：

其中通常p表示X取值为1的概率。

# 概率分布

## 概率质量函数

PMF: Probability Mass Function

有些书也写作分布律（如浙大版<<概率论与数理统计>>），与连续随机变量的PDF（Probability Density Function: 概率密度函数）相对应）。

伯努利分布的概率质量函数为：

当然，为了和条件概率区分开来，也有将写成的写法。一般情况下根据上下文也可以做出区分。

也可以进一步简记（未必简单，但是对于要做数学推导处理等就比较方便）为：

其中I(x)是所谓的Indicator函数。因此，以上式表示仅在x=0或x=1时成立，其它情况下则为0.

## 概率分布函数

伯努利分布的概率分布函数（CDF: Cumulative Distribution Function，累积分布函数。不管对于离散随机变量还是连续随机变量，CDF总是一个连续函数。只不过对于离散随机变量而言，CDF是一个分段函数：piecewise function）如下所示：

## 生存函数，风险函数

生存函数是概率分布函数的互补函数，定义为：。伯努利分布的生存函数为：

与生存函数相关联风险函数定义为：

注意，风险函数仅在S(x)的非零区间（In other words, only valid on S support）有定义。伯努利分布的风险函数为：

伯努利分布的累计风险函数为 (???待确认)

## 百分点函数

PPF: Percent Point Function. 也称百分位数(percentile)?

百分点函数PPF是CDF的反函数，它回答了“为了得到一定的累积分布概率，CDF相应的输入值是什么”的问题。伯努利分布的百分点函数为：

# 常用统计特征

## 均值，Mean

## 中位数，Median

## 众数，Mode

## 方差，Variance

## 偏度，Skewness

## 峰度，Kurtosis

[Reference]

[Revision history]