

## 4.5互信息



## 4.5互信息

两个随机变量 $X, Y$ 的互信息定义为 $X, Y$ 的联合分布和各自边缘分布乘积的相对熵，用 $I(X, Y)$ 表示。



# 随机变量

- ★ 随机试验结果的实值单值函数
- ★ 可数量化
- ★ 主要对某些函数感兴趣



两个随机变量 $X, Y$ 的互信息定义为 $X, Y$ 的联合分布和各自边缘分布乘积的相对熵，用 $I(X, Y)$ 表示。

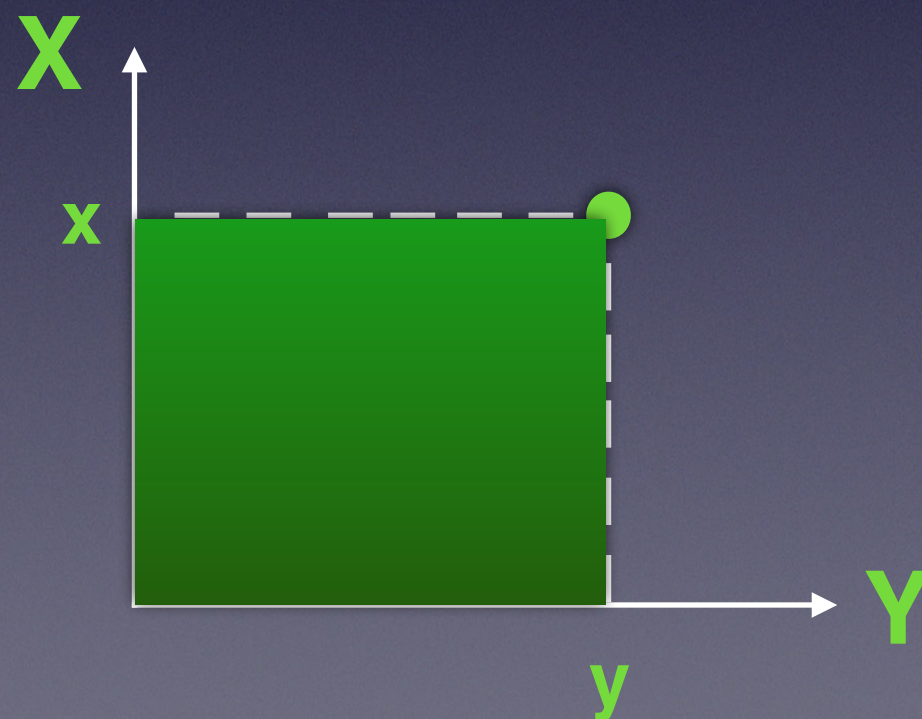


# 联合分布

设  $(X, Y)$  是二维随机变量

对于任意实数  $x, y$ :

$$F(x, y) = P(X \leq x, Y \leq y)$$





两个随机变量 $X, Y$ 的互信息定义为 $X, Y$ 的联合分布和各自边缘分布乘积的相对熵，用 $I(X, Y)$ 表示。



# 边缘分布

★ 只包含其中部分变量的概率分布

假设有一个和两个变量相关的概率分布： **$P(x,y)$**

其中一个特定变量的边缘分布：

$$P(x) = \sum_y P(x, y) = \sum_y P(x|y)P(y).$$



两个随机变量 $X, Y$ 的互信息定义为 $X, Y$ 的联合分布和各自边缘分布乘积的相对熵，用 $I(X, Y)$ 表示。



## 相对熵

$$\text{KL}(P \parallel Q) = \sum P(x) \log \frac{P(x)}{Q(x)}$$

$$\text{KL}(P \parallel Q) = \int P(x) \log \frac{P(x)}{Q(x)} dx$$



两个随机变量 $X, Y$ 的互信息定义为 $X, Y$ 的联合分布和各自边缘分布乘积的相对熵，用 $I(X, Y)$ 表示。



## 互信息—I(X,Y)

两个随机变量X,Y的联合分布为P(x,y)  
边缘分布为p(x),p(y)

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$



## 互信息—I(X,Y)

```
from sklearn import metrics
if __name__ == '__main__':
    A = [1, 1, 1, 2, 3, 3]
    B = [1, 2, 3, 1, 2, 3]

    result_NMI=metrics.normalized_
mutual_info_score(A, B)
    print("result_NMI:",result_NMI)
```