

## カテゴリカルデータを2値化する

	カレーはおいしいか？	2値化
回答者1	いいえ	0
回答者2	はい	1
回答者3	はい	1
回答者4	いいえ	0
回答者5	はい	1

$$\text{平均値 } \bar{x} = \frac{0+1+1+0+1}{5} = \frac{3}{5}$$

## 平方和

$$\begin{aligned} S_{xx} &= \left(0 - \frac{3}{5}\right)^2 + \left(1 - \frac{3}{5}\right)^2 + \left(1 - \frac{3}{5}\right)^2 + \left(0 - \frac{3}{5}\right)^2 + \left(1 - \frac{3}{5}\right)^2 \\ &= 0^2 + 1^2 + 1^2 + 0^2 + 1^2 - \frac{(0+1+1+0+1)^2}{5} \\ &= (0+1+1+0+1) \left(1 - \frac{0+1+1+0+1}{5}\right) \\ &= 3 \left(1 - \frac{3}{5}\right) \end{aligned}$$

## 分散

$$\frac{S_{xx}}{5} = \frac{3 \left(1 - \frac{3}{5}\right)}{5} = \frac{3}{5} \left(1 - \frac{3}{5}\right) = \bar{x}(1 - \bar{x})$$

## 標準偏差

$$\sqrt{\frac{S_{xx}}{5}} = \sqrt{\frac{3}{5} \left(1 - \frac{3}{5}\right)} = \sqrt{\bar{x}(1 - \bar{x})}$$