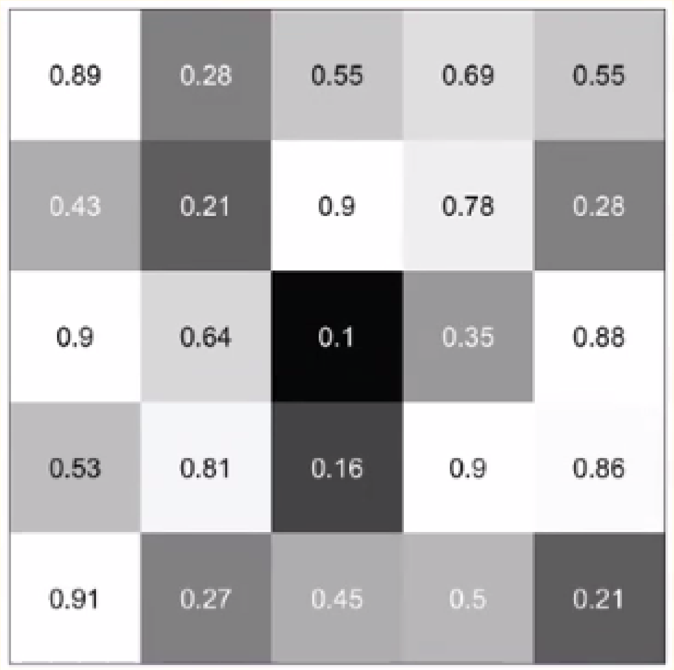
From Andrew Tomas' lecture

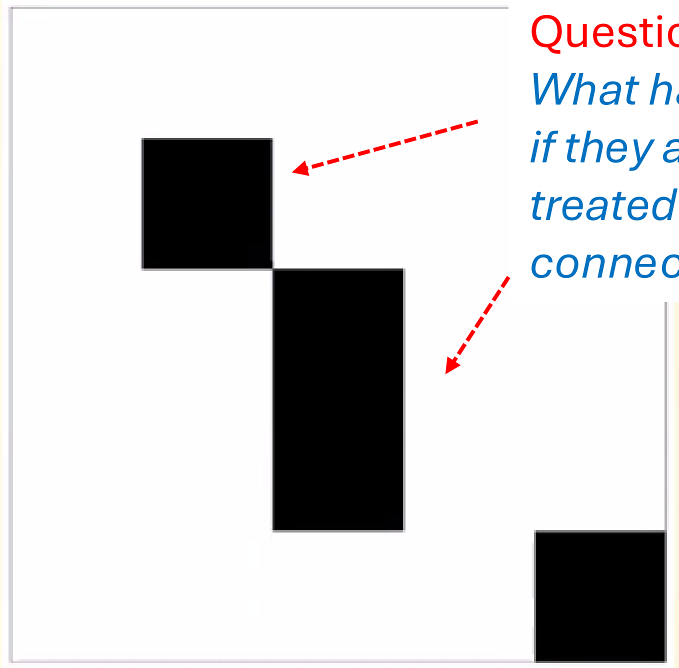
TDA: cubical homology

Grayscale image I

$\Rightarrow$ Binary images $\mathbf{1}\{I \leq t\}$



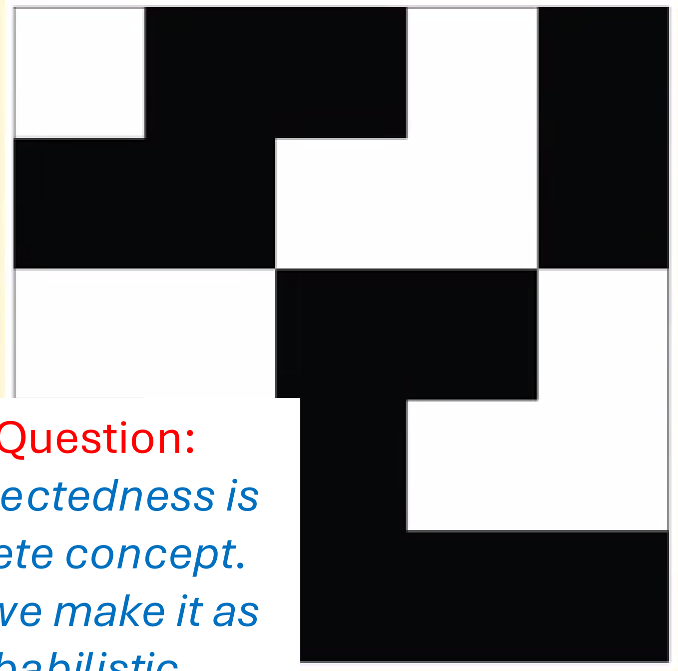
$t = 0.25;\ \beta_0 = 2,\ \beta_1 = 0$



Question:

What happen if they are treated as connected

$t = 0.55;\ \beta_0 = 1,\ \beta_1 = 0$



Next Question:

Connectedness is discrete concept. Can we make it as a probabilistic statement?

$t = 0.8;\ \beta_0 = 1,\ \beta_1 = 2$

$t = 0.85;\ \beta_0 = 1,\ \beta_1 = 1$

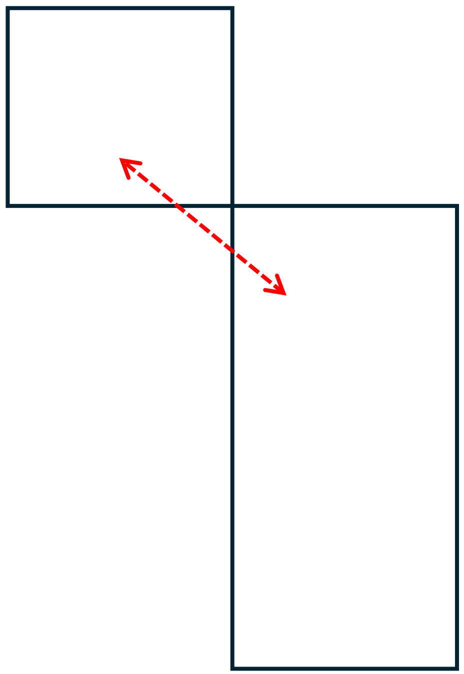
What should be reasonable approach?

Making discrete data into continuous quantity

Problem solved by
Paul Dirac (novel laureate in Quantum physics) through
Dirac delta function

$$\int_{\mathbb{R}} f(x)\,\delta(x - x_0)\,dx \;=\; f(x_0).$$

How statistics people addressed this problem?



Define indicator function I condition on the event that it is 1 if two objects are connected and 0 otherwise.

$$\begin{aligned}\mathbb{E}I &= 1 \cdot P(object\ connected) + 0 \cdot P(object\ not\ connected)\\ &= P(object\ connected)\end{aligned}$$

Then what next?

Logistic regression

$$P(I = 1 \mid x) = \underset{\uparrow}{\sigma}(\beta_0 + x^\top \beta) = \frac{1}{1 + \exp\left(-(\beta_0 + x^\top \beta)\right)}$$

Sigmoid function

Class do not cover logistic regression. Too textbook-like topic