SOEE1475 Statistics and Data Analysis
Lecture 5: Spatial statistics

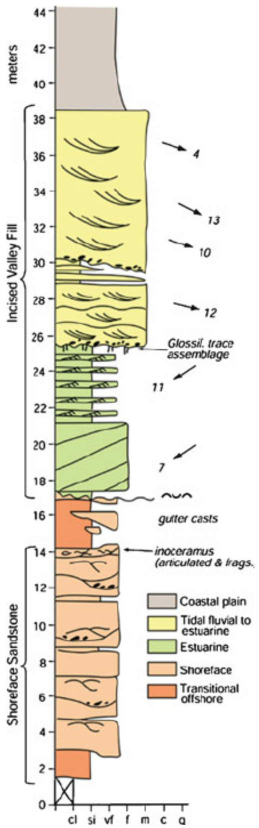
Graeme T. Lloyd

Today

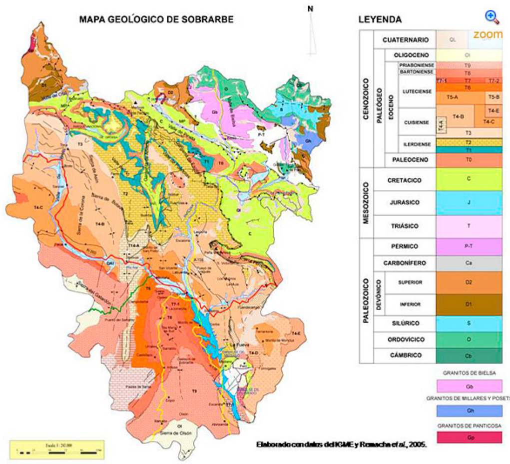
- Spatial patterns of point distributions
- Chi-square test
- Nearest neighbour distances

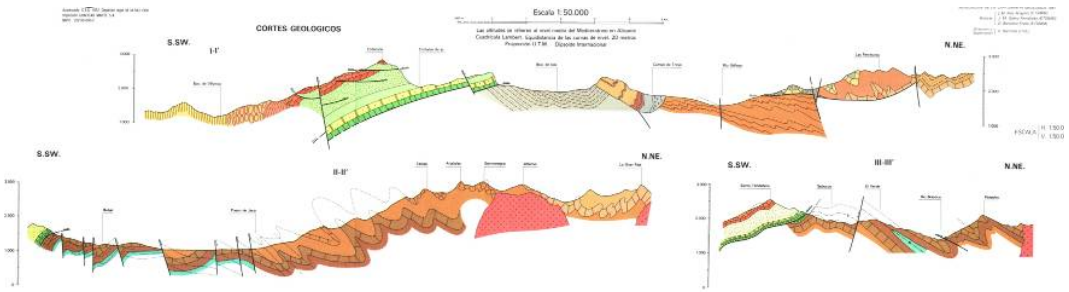
Spatial data in geology

1D

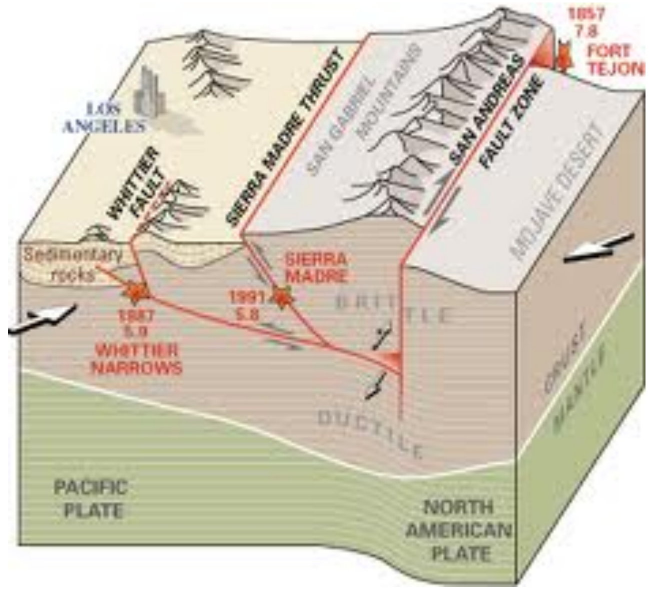


2D

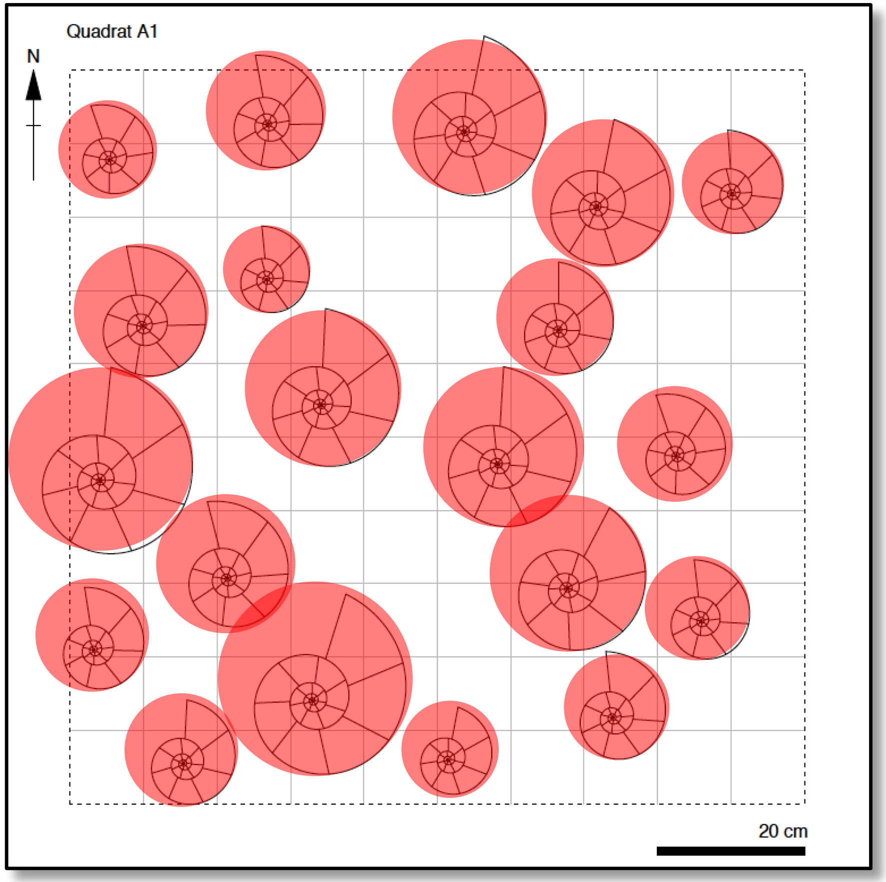


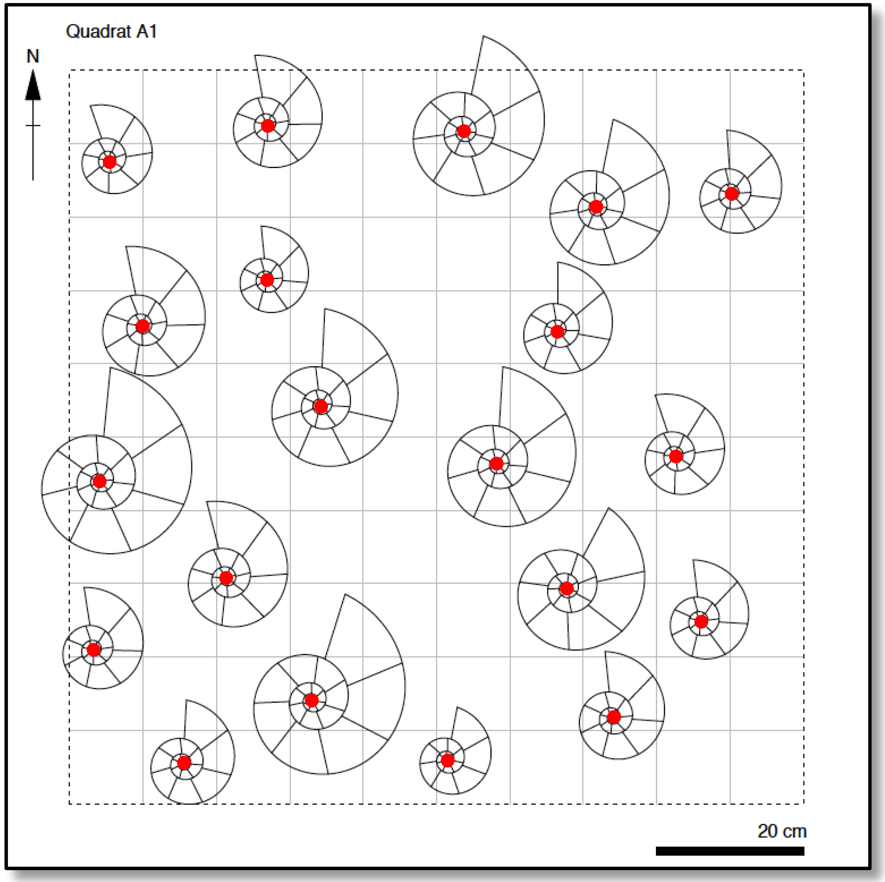


3D

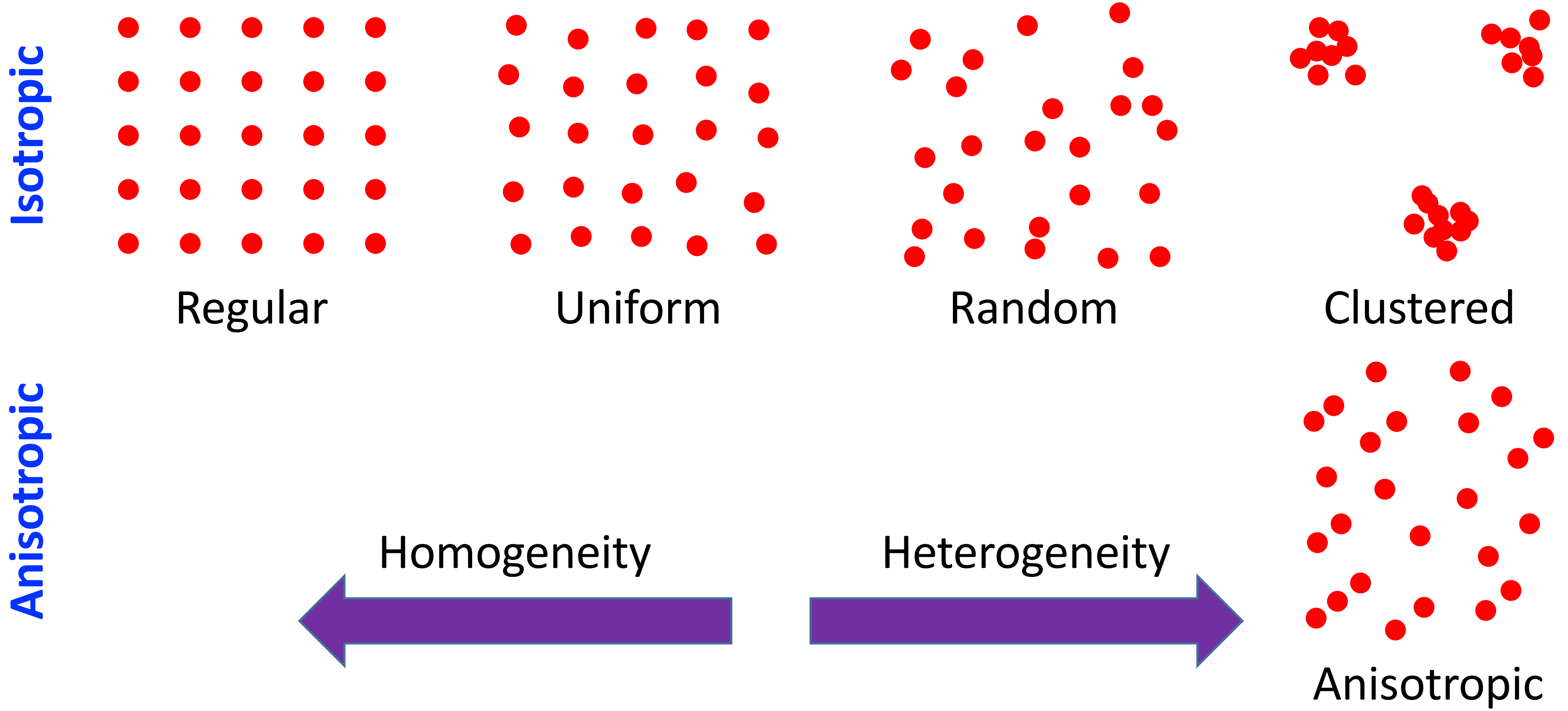


Point distributions

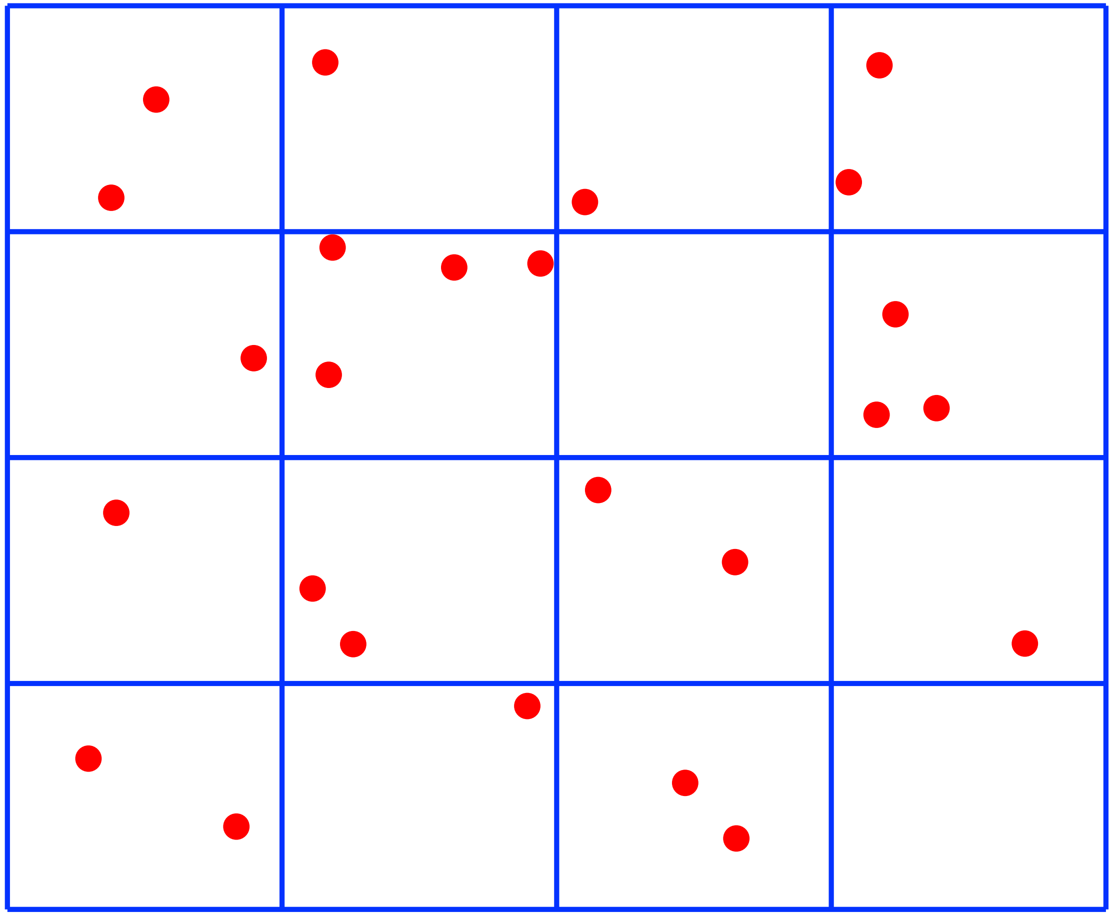




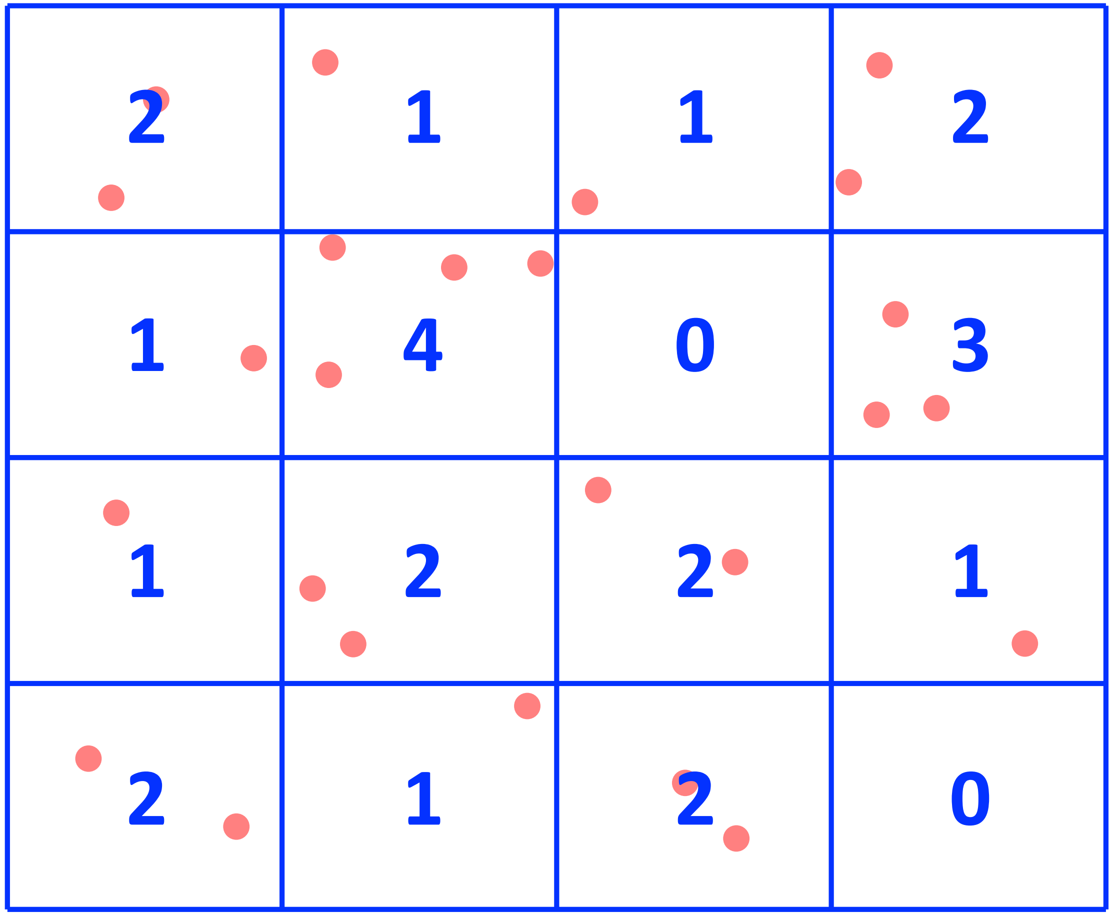
Point distributions



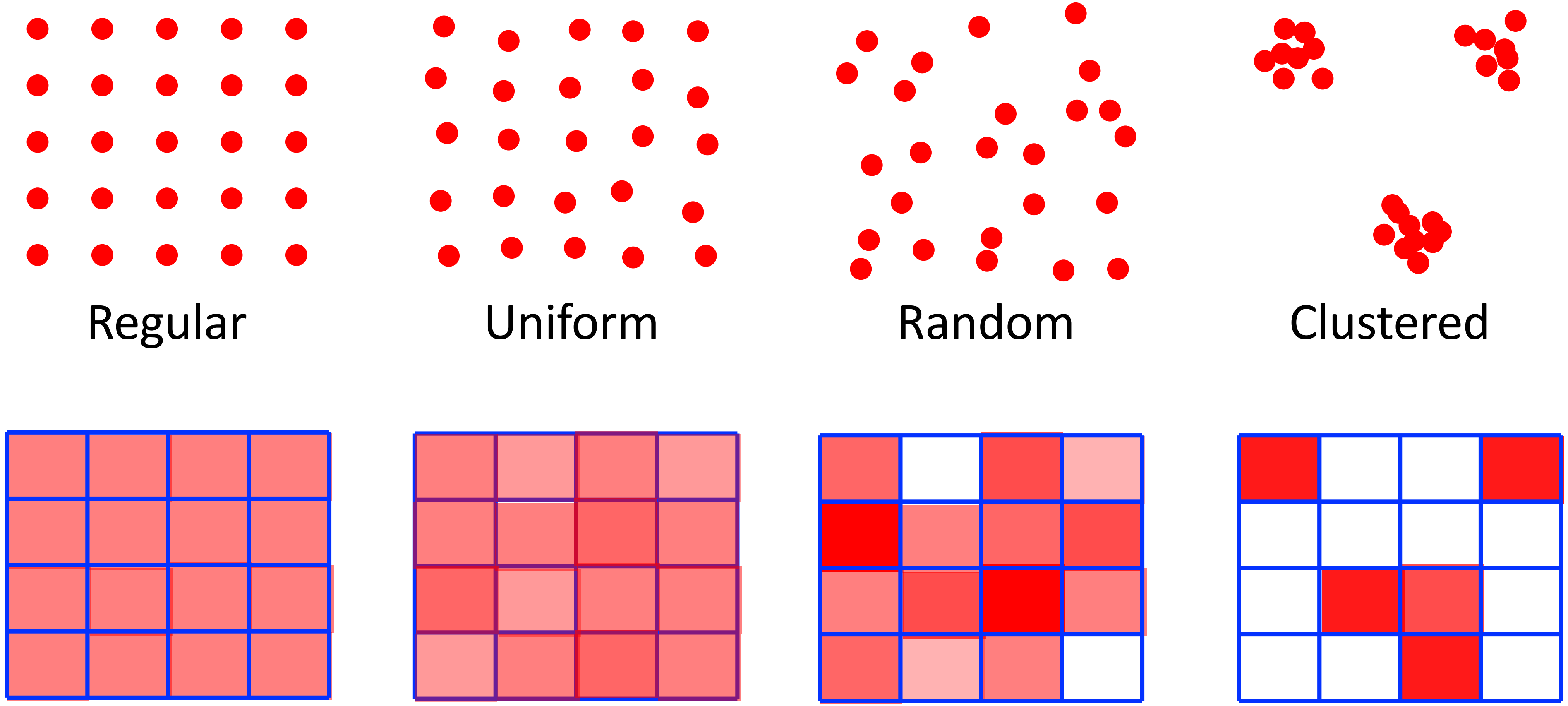
Point distributions



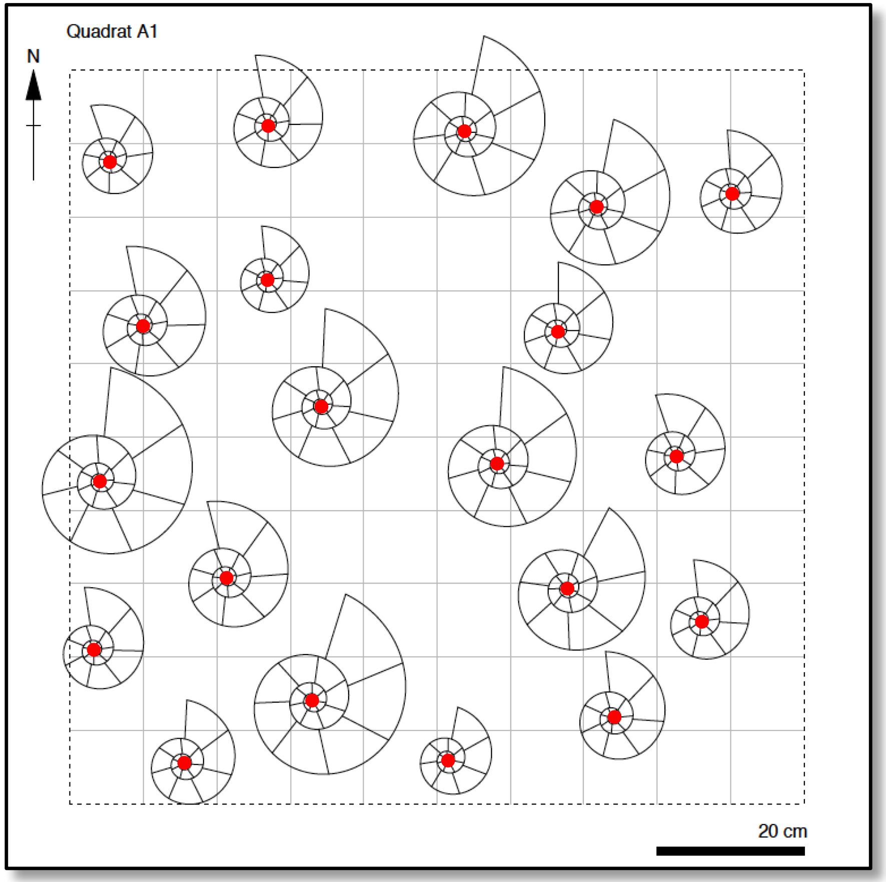
Point distributions



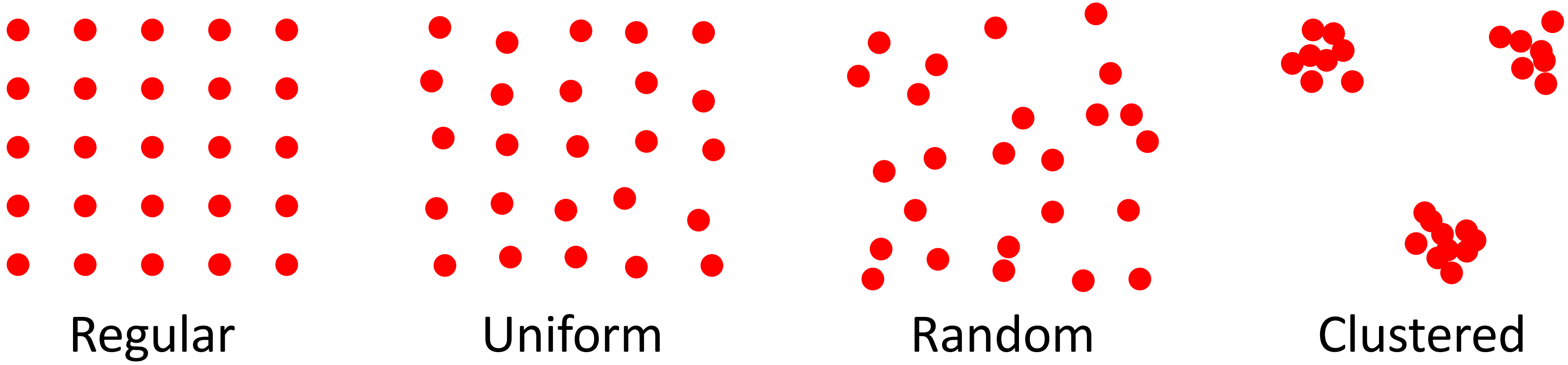
Point distributions



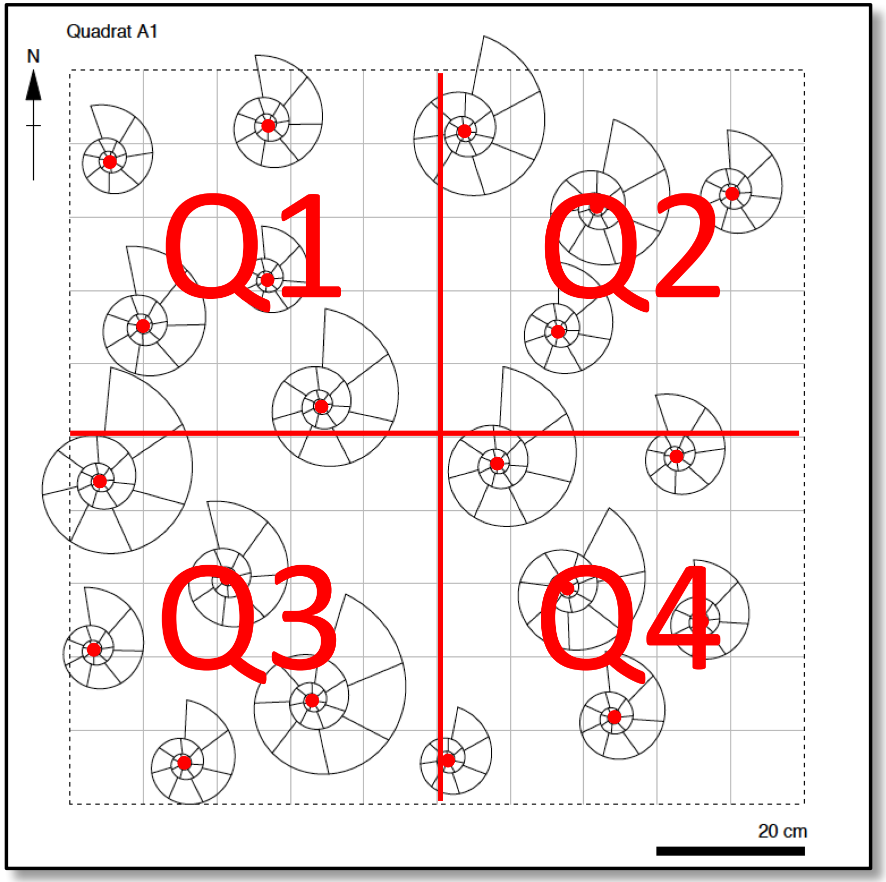
Exercise – part I



Exercise – part I



Exercise – part I



Q1 =

Q2 =

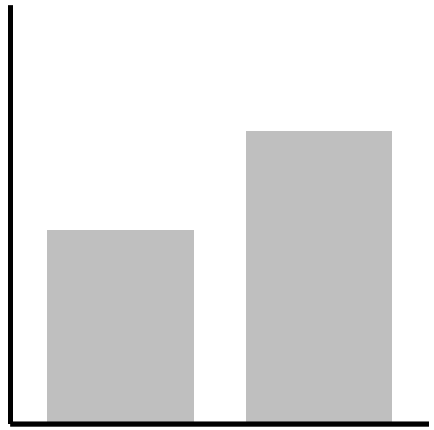
Q3 =

Q4 =

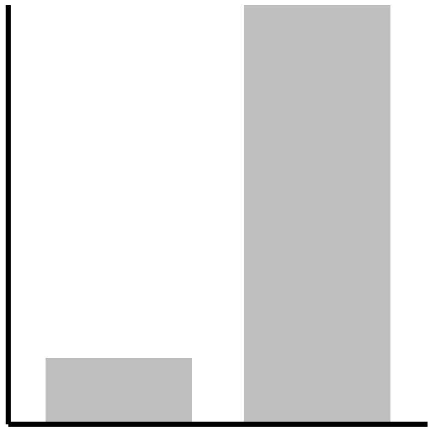
The Chi-squared test – testing for uniform

$$\chi^2 = \sum_{j=1}^{k} \left[\frac{\left(O_j - E_j \right)^2}{E_j} \right]$$

Low χ^2



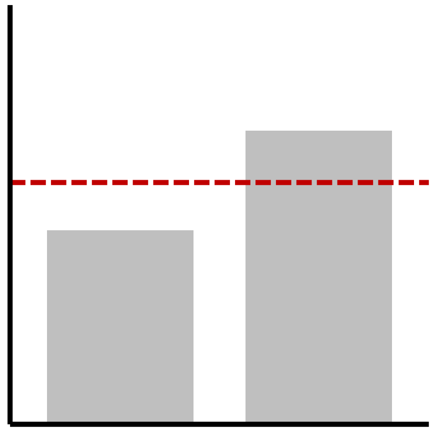
High χ^2



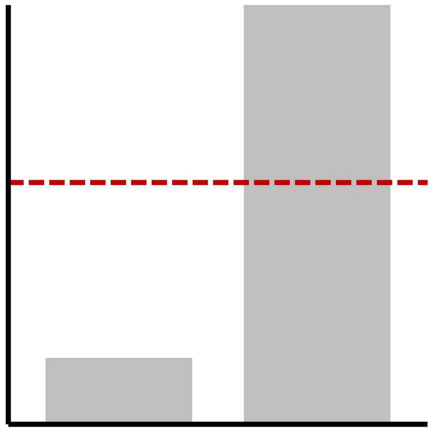
The Chi-squared test – testing for uniform

$$\chi^2 = \sum_{j=1}^{k} \left[\frac{(O_j - E_j)^2}{E_j} \right]$$

Low χ^2



High χ^2

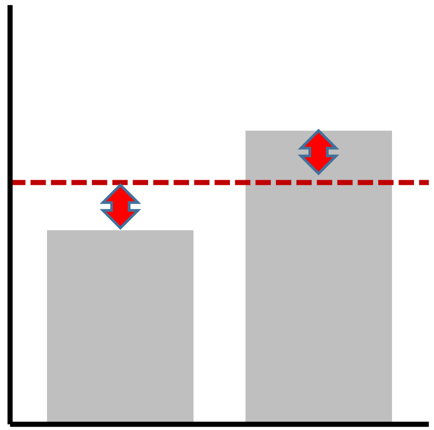


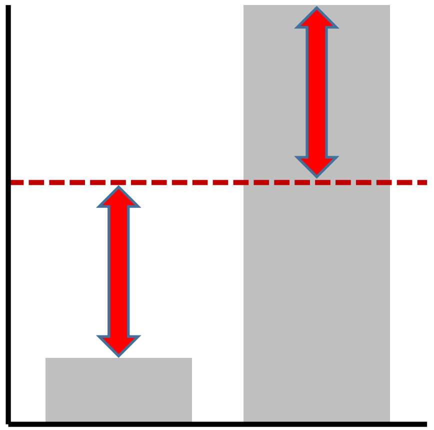
The Chi-squared test – testing for uniform

$$\chi^2 = \sum_{j=1}^{k} \left[\frac{\left(O_j - E_j\right)^2}{E_j} \right]$$

Low χ^2

High χ^2





Exercise – part II

$$\chi^2 = \sum_{j=1}^{k} \left[\frac{(O_j - E_j)^2}{E_j} \right]$$

$$E_j = \frac{N_{ammonites}}{N_{quadrants}}$$

$$\chi^2 = \frac{(Q1-5)^2}{5} + \frac{(Q2-5)^2}{5} + \frac{(Q3-5)^2}{5} + \frac{(Q4-5)^2}{5}$$

The Chi-squared test – testing for uniform

$$0 \leq \chi^2 \leq \infty$$

Degrees of freedom

Sample standard deviation

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Pearson correlation coefficient

$$r = \frac{\sum_{i=1}^{n}(Z_{xi} \times Z_{yi})}{n-1}$$

t-distribution parameter

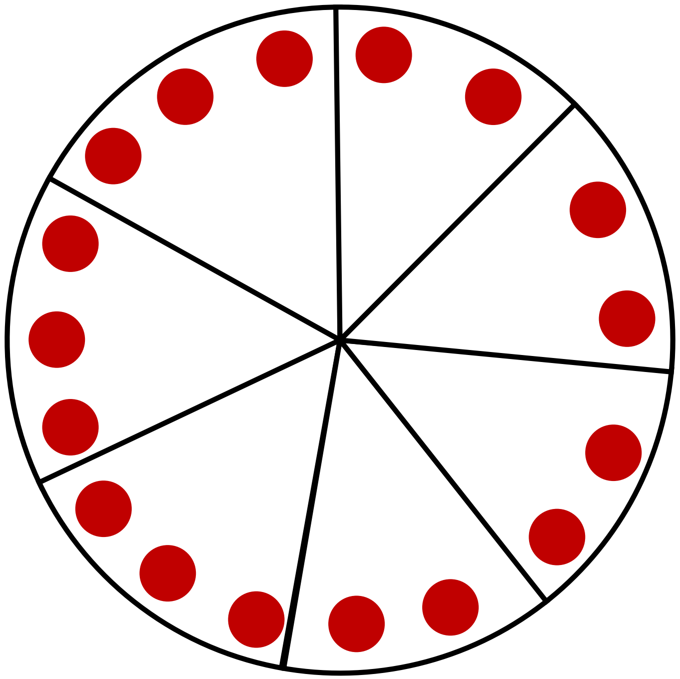
$$v = n - 1$$

F-distribution parameters

$$d_1 = N_{replicates} - 1$$

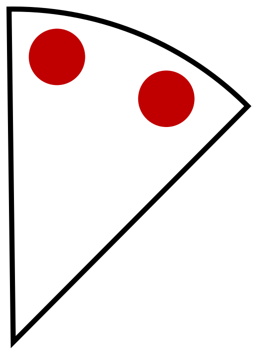
$$d_2 = N_{treatments} - 1$$

Degrees of freedom



Degrees of freedom

Person 1: Choice
Person 2: Choice
Person 3: Choice
Person 4: Choice
Person 5: Choice
Person 6: Choice
Person 7: No choice



Degrees of freedom

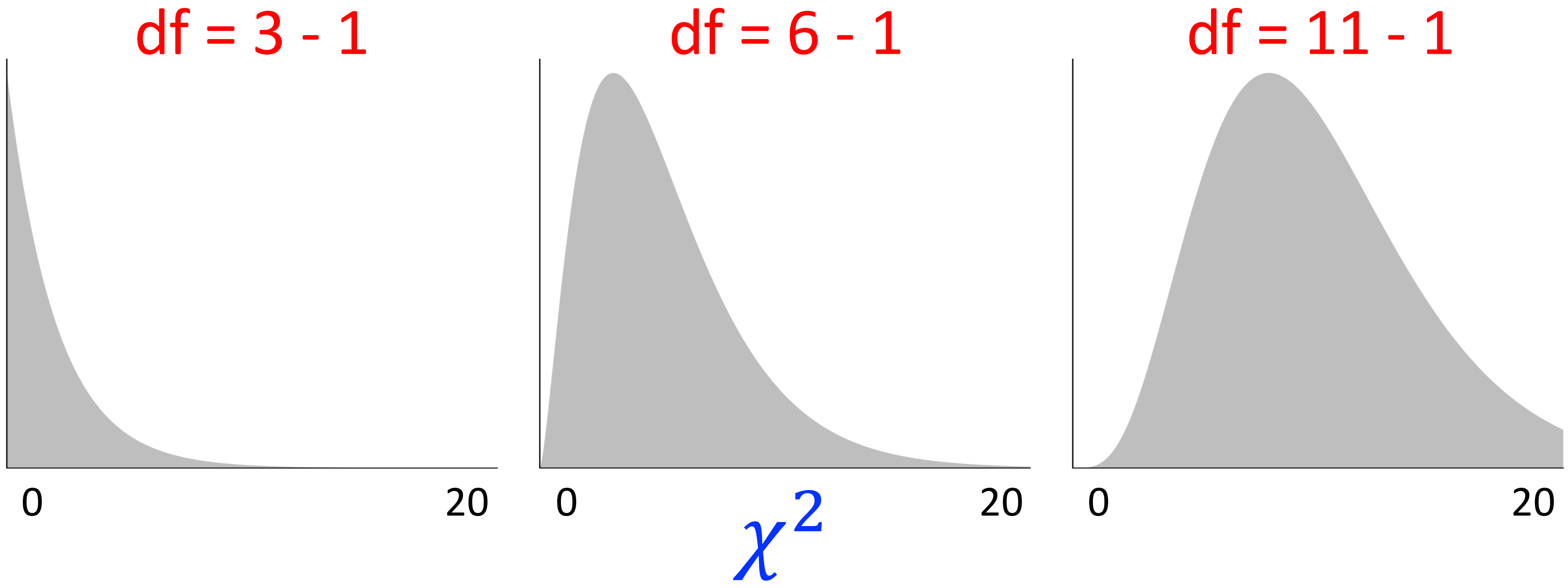
$$x_1 = 5 \qquad\qquad x_1 - \bar{x} = -5$$

$$x_2 = 10 \qquad \bar{x} = 10 \qquad x_2 - \bar{x} = 0$$

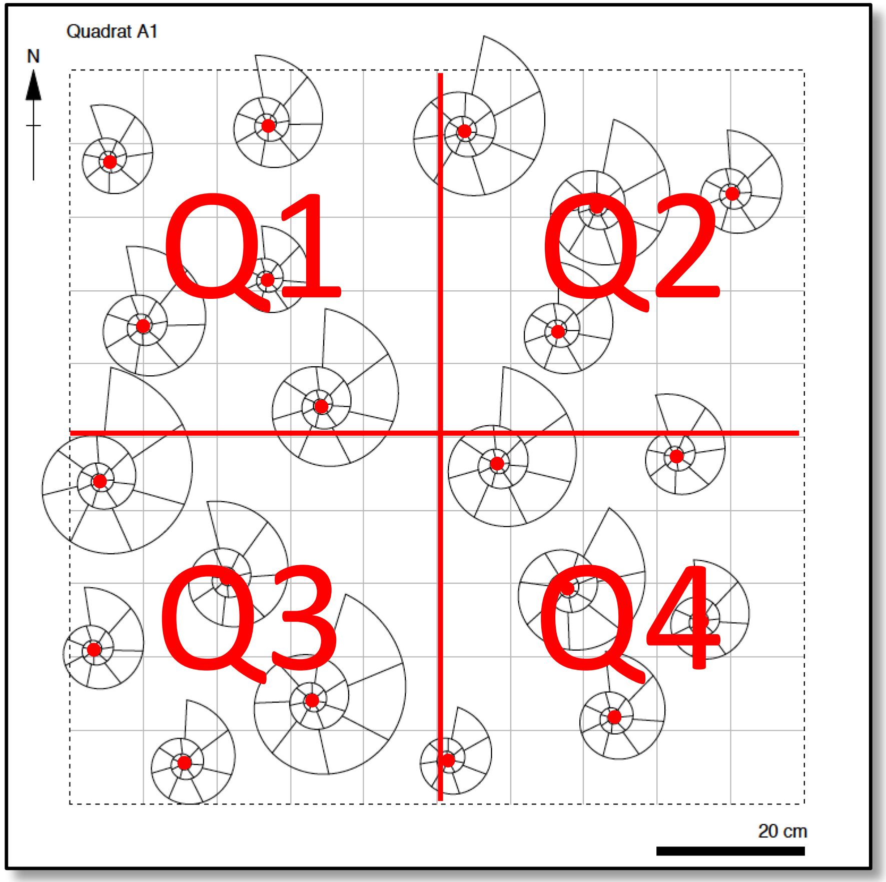
$$x_3 = 15 \qquad\qquad x_3 - \bar{x} = ?$$

Degrees of freedom

$$0 \le \chi^2 \le \infty$$



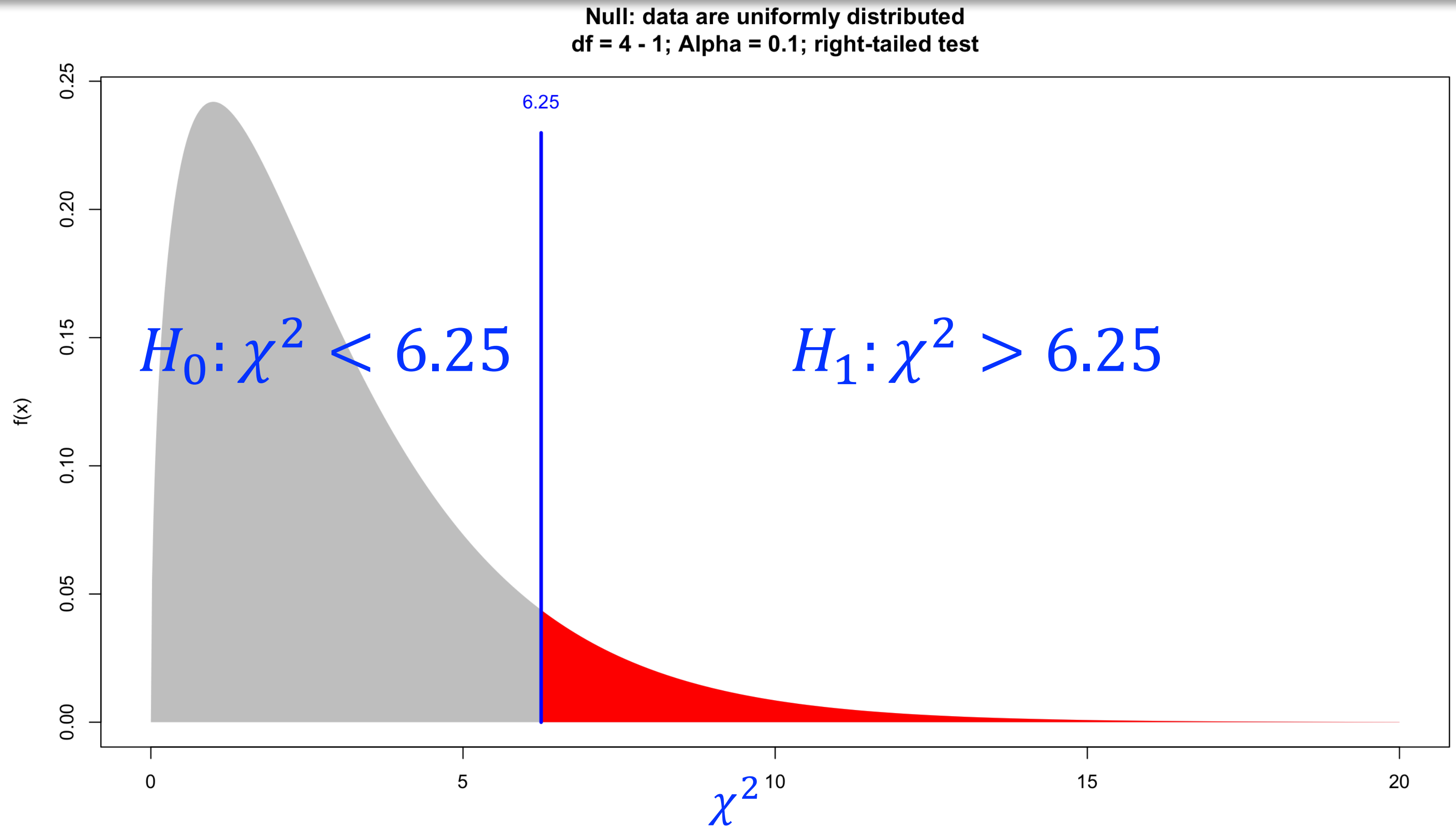
Exercise – part III



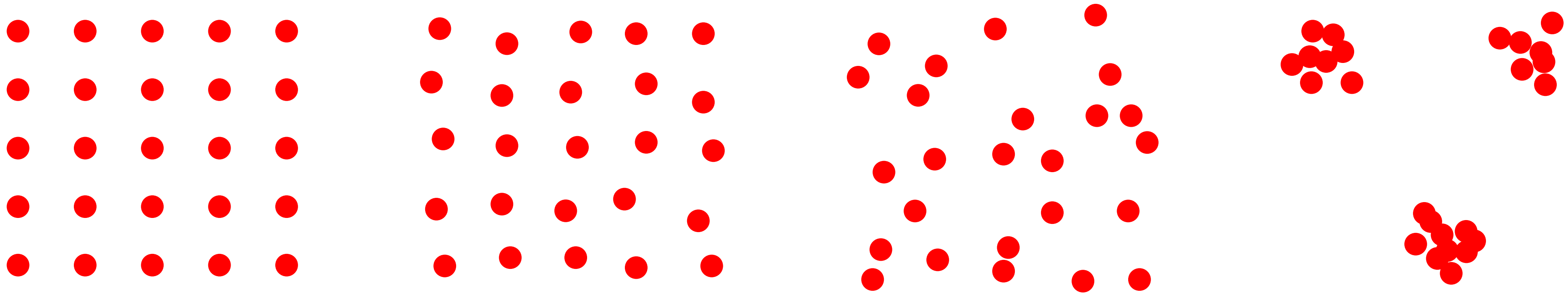
$$\chi^2 = \frac{(Q1-5)^2}{5} + \frac{(Q2-5)^2}{5} + \frac{(Q3-5)^2}{5} + \frac{(Q4-5)^2}{5}$$

df = ?

Exercise – part III



Exercise – part III



Regular Uniform Random Clustered

Poisson distribution – testing for random

Discrete probability distribution

Has single parameter (λ), the mean

$$P(X = x) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Expresses probability of given number of events

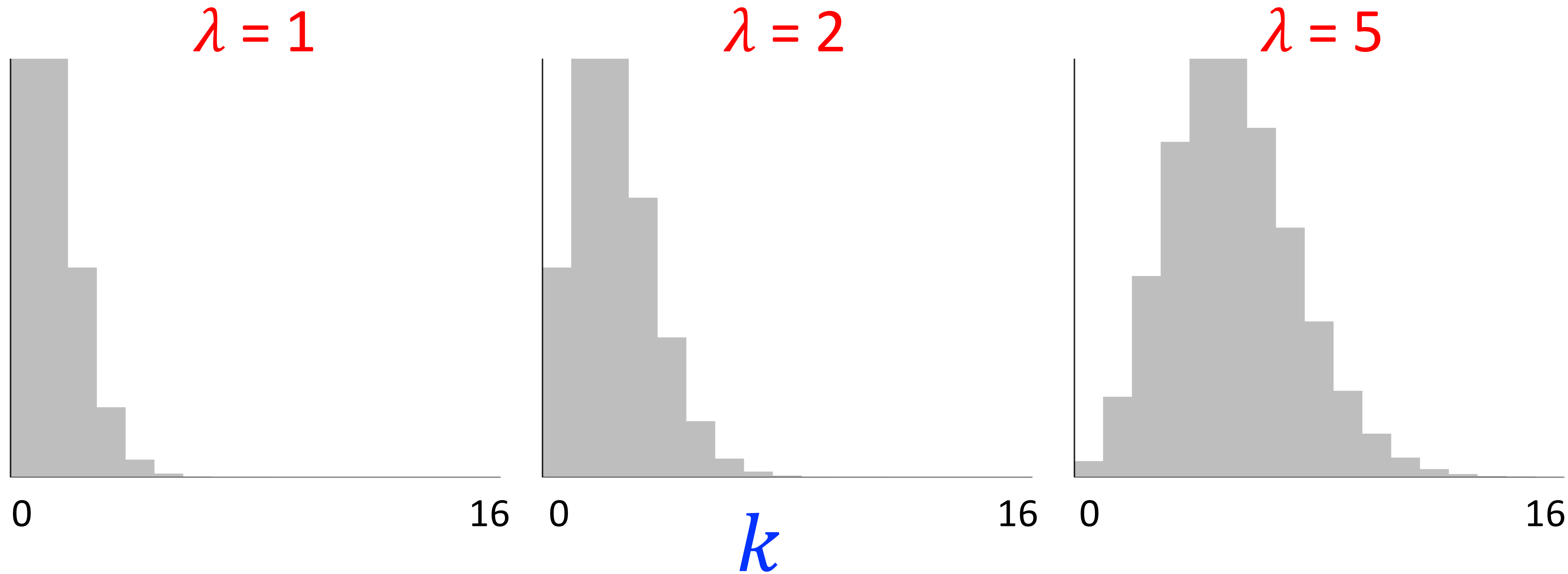
Assumes randomness; events can occur in space or time

Good null for random distribution of points in space

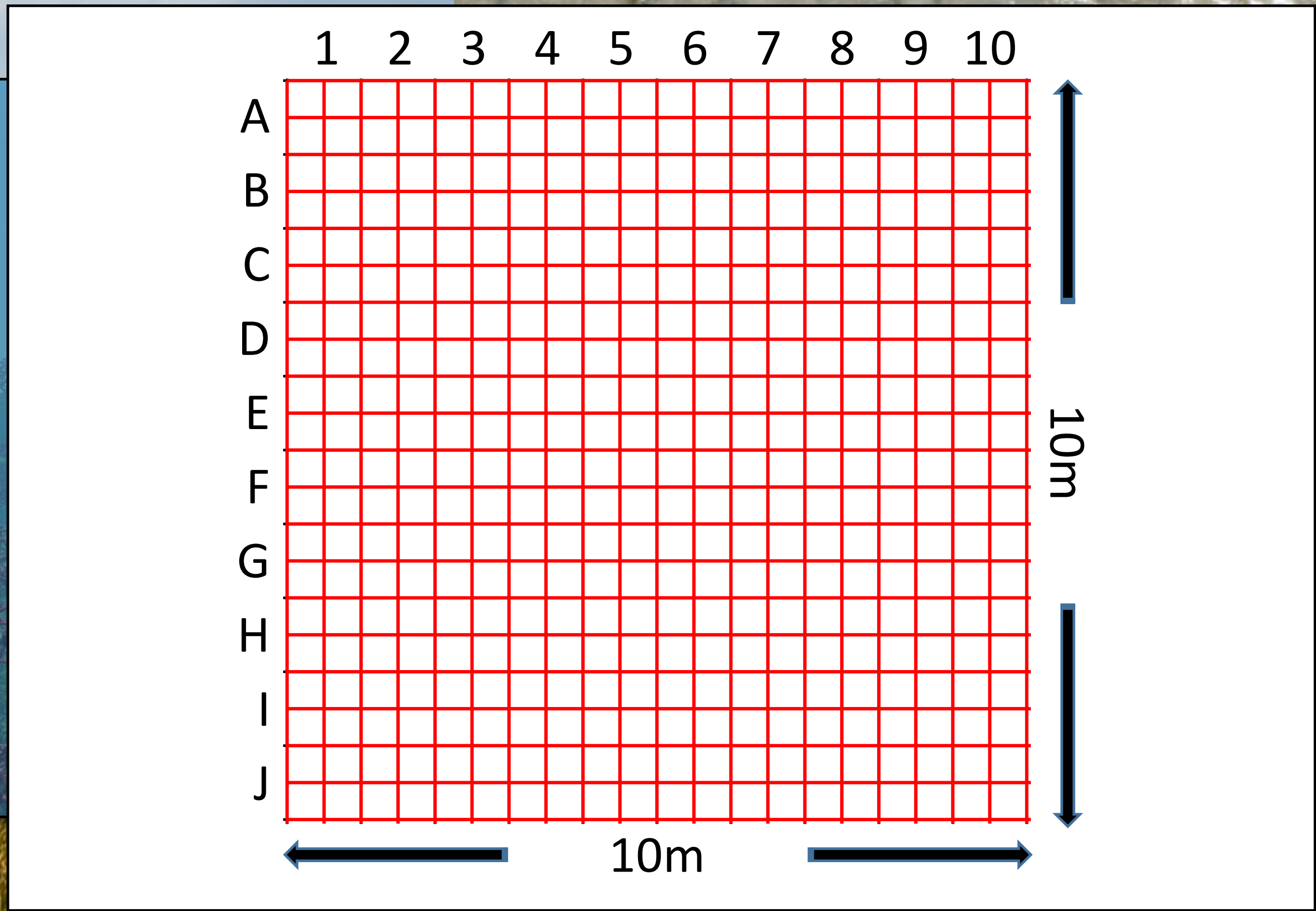
I.e., in ammonite example expected N quadrants with k ammonites

Poisson distribution – testing for random

$$0 < \lambda \leq \infty$$



Poisson distribution – testing for random

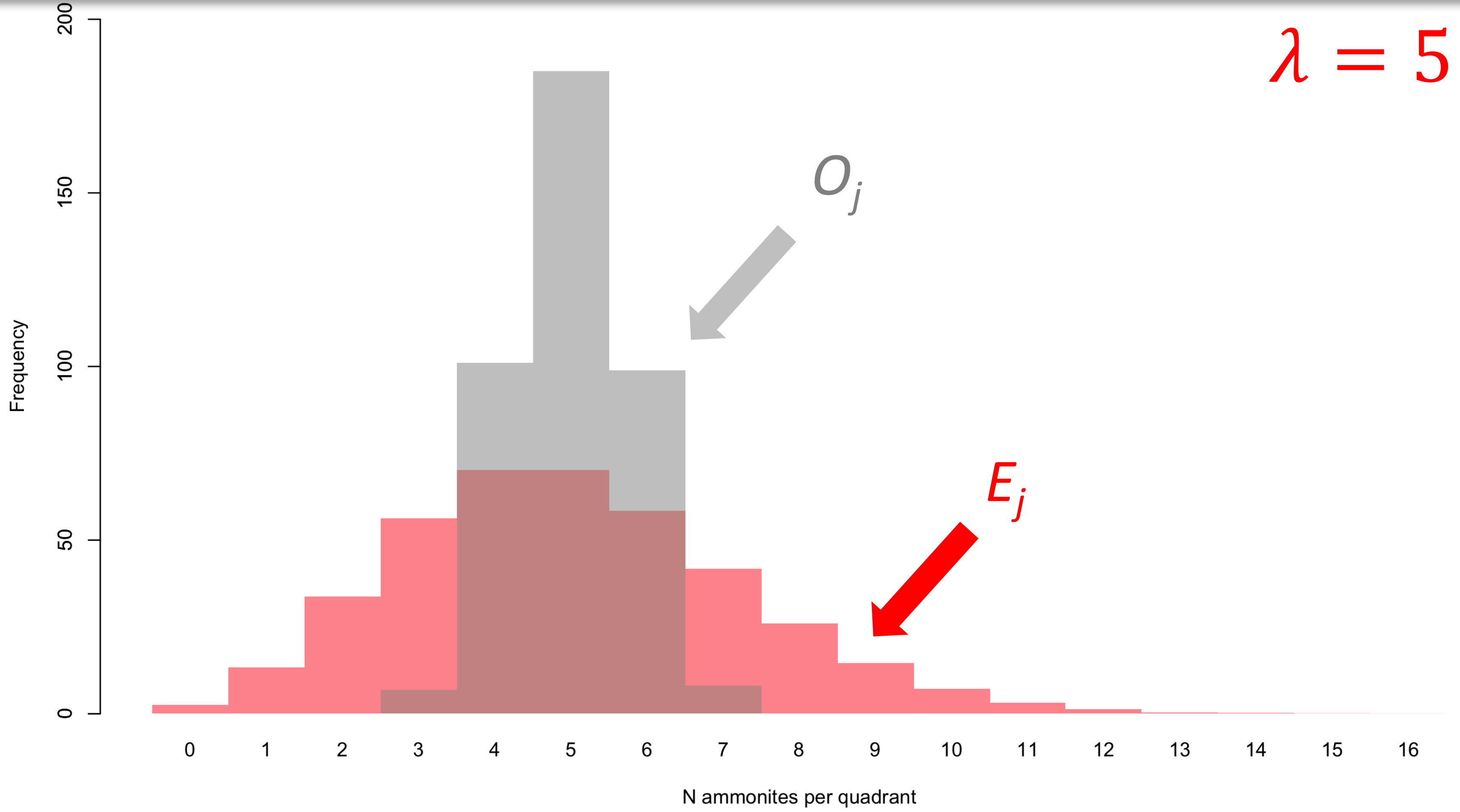


Poisson distribution – testing for random

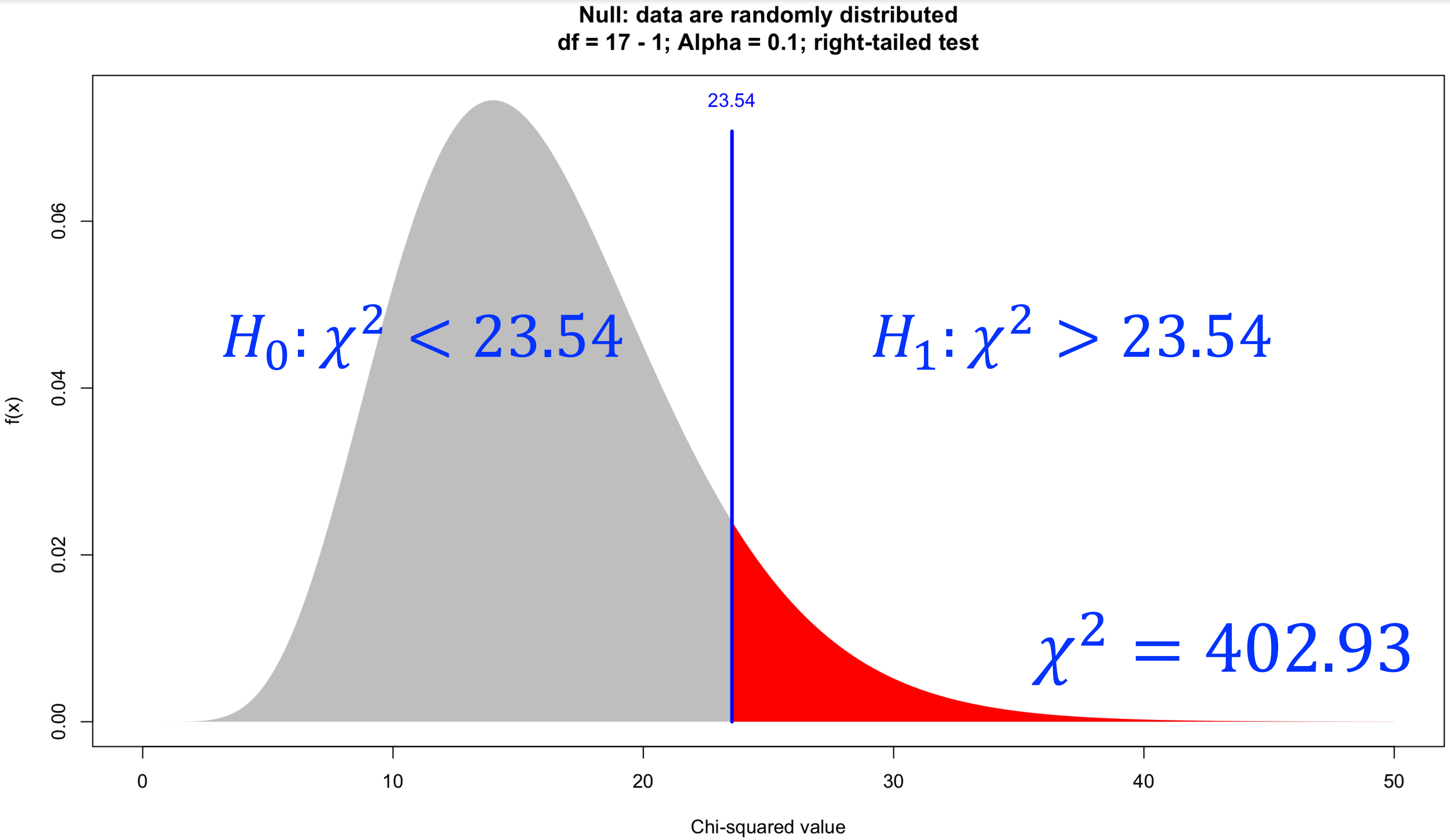
$$\lambda = \frac{N_{ammonites}}{N_{quadrants}}$$

$$\lambda = \frac{2000}{4 \times 100} = 5$$

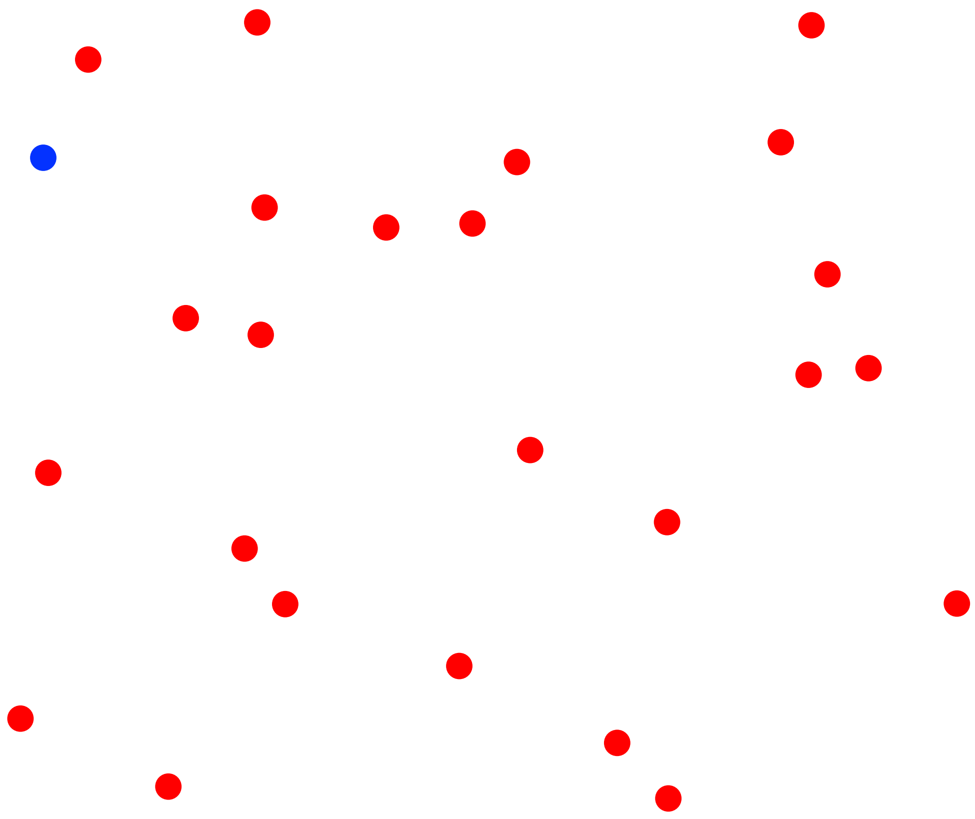
Poisson distribution – testing for random



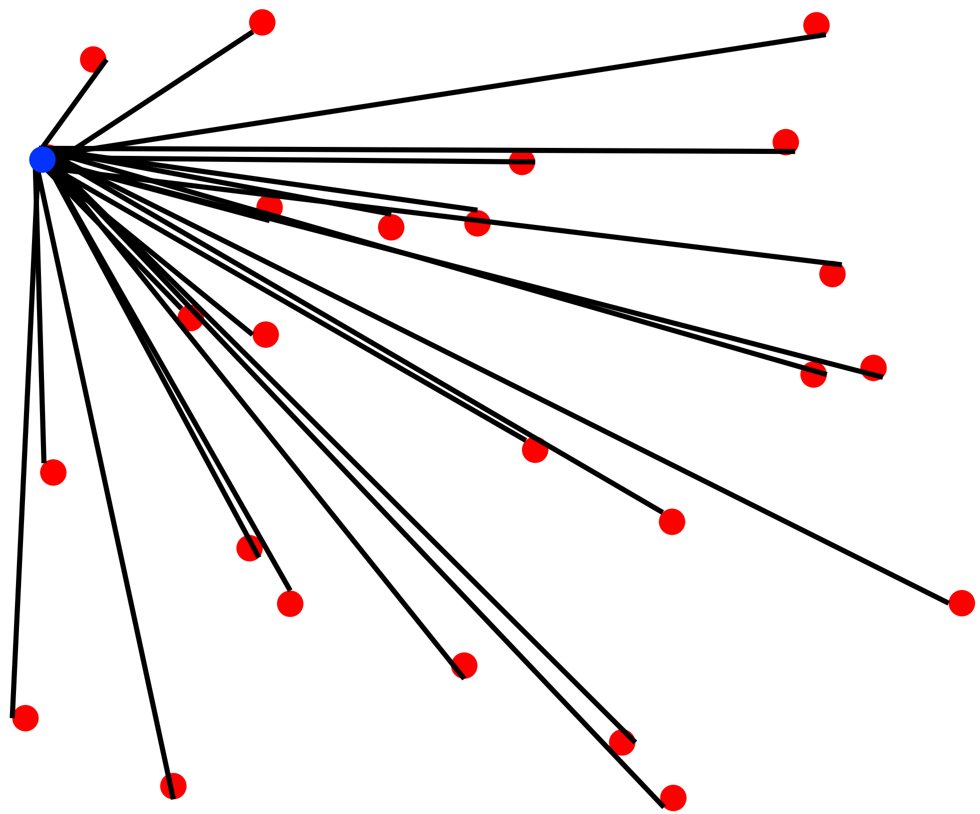
Poisson distribution – testing for random



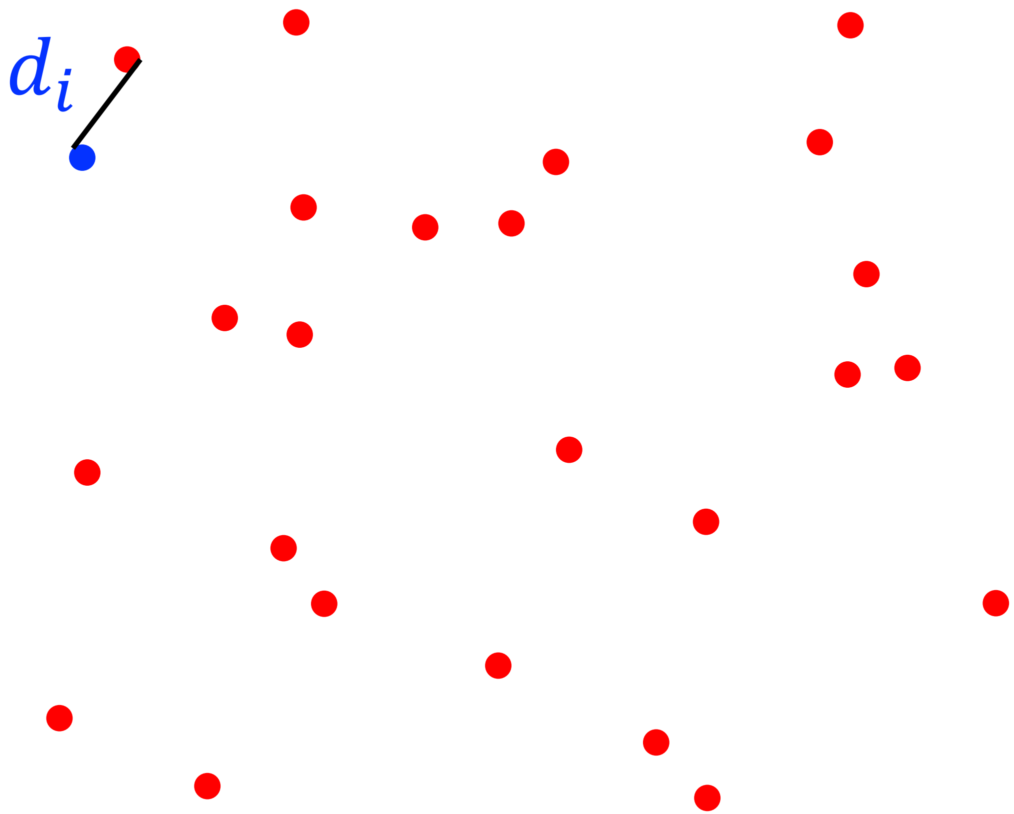
Other tests - nearest neighbour



Other tests - nearest neighbour



Other tests - nearest neighbour



Other tests - nearest neighbour

Mean nearest neighbour distance

Captures average distance to nearest point

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$$

Expected mean nearest neighbour distance

Uses total area (A) and number of objects (n)

$$\bar{\delta} = \frac{1}{2}\sqrt{A/n}$$

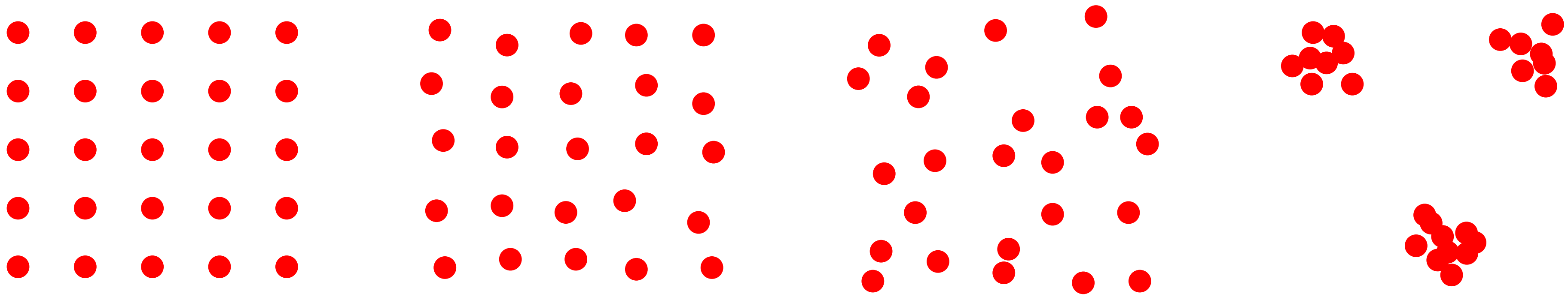
Nearest neighbour statistic

Varies from 0 to 2.15

$$R = \frac{\bar{d}}{\bar{\delta}}$$

Other tests - nearest neighbour

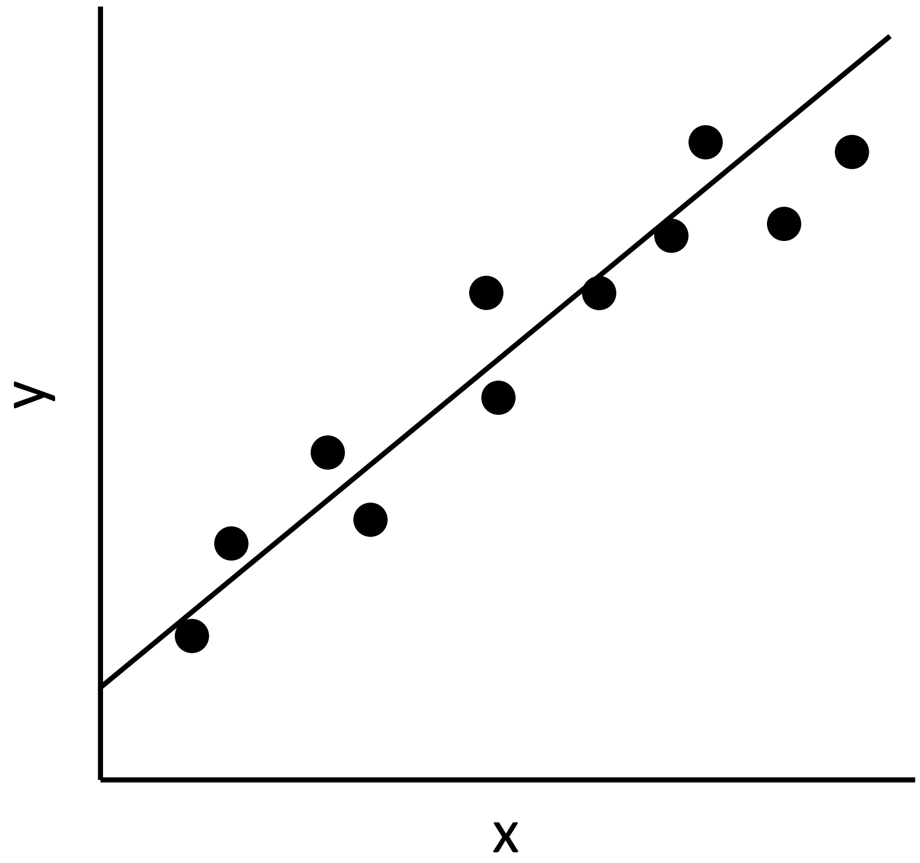
Nearest neighbour statistic

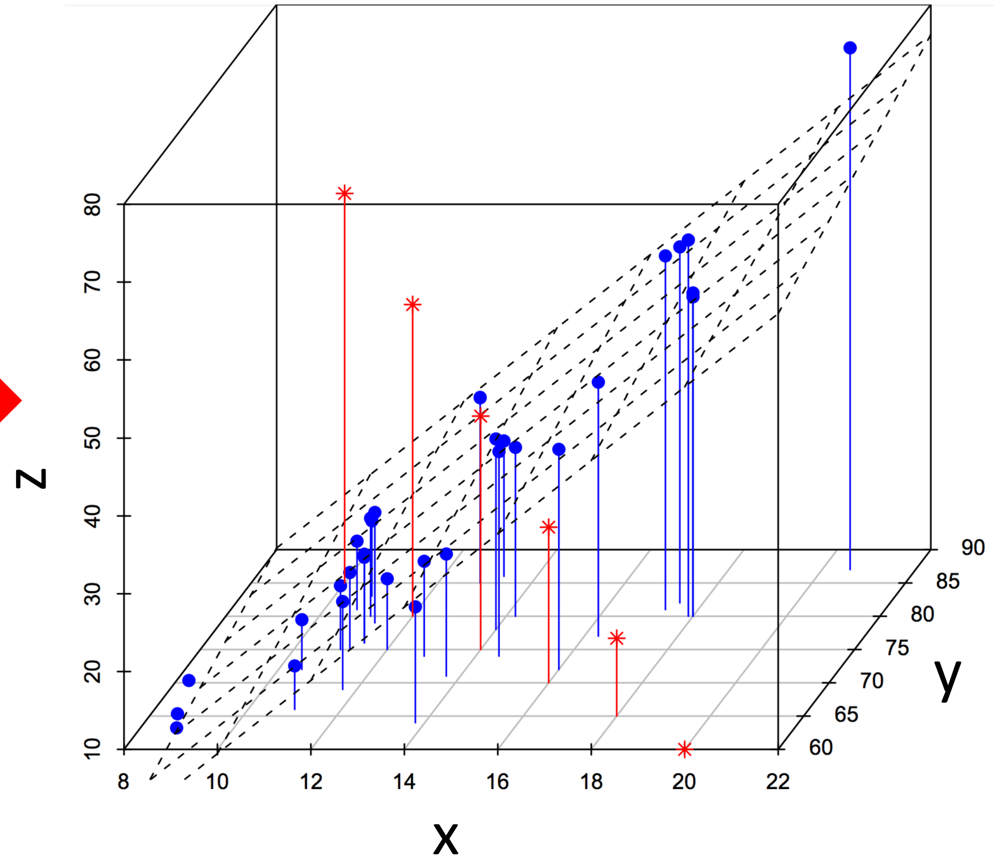


Regular | Uniform | Random | Clustered

$R = 2.15$ | $2.15 > R > 1$ | $R \approx 1$ | $R < 1$

3D regression





3D regression

Planes $$z = ax + by + c$$

3D equivalent of straight line

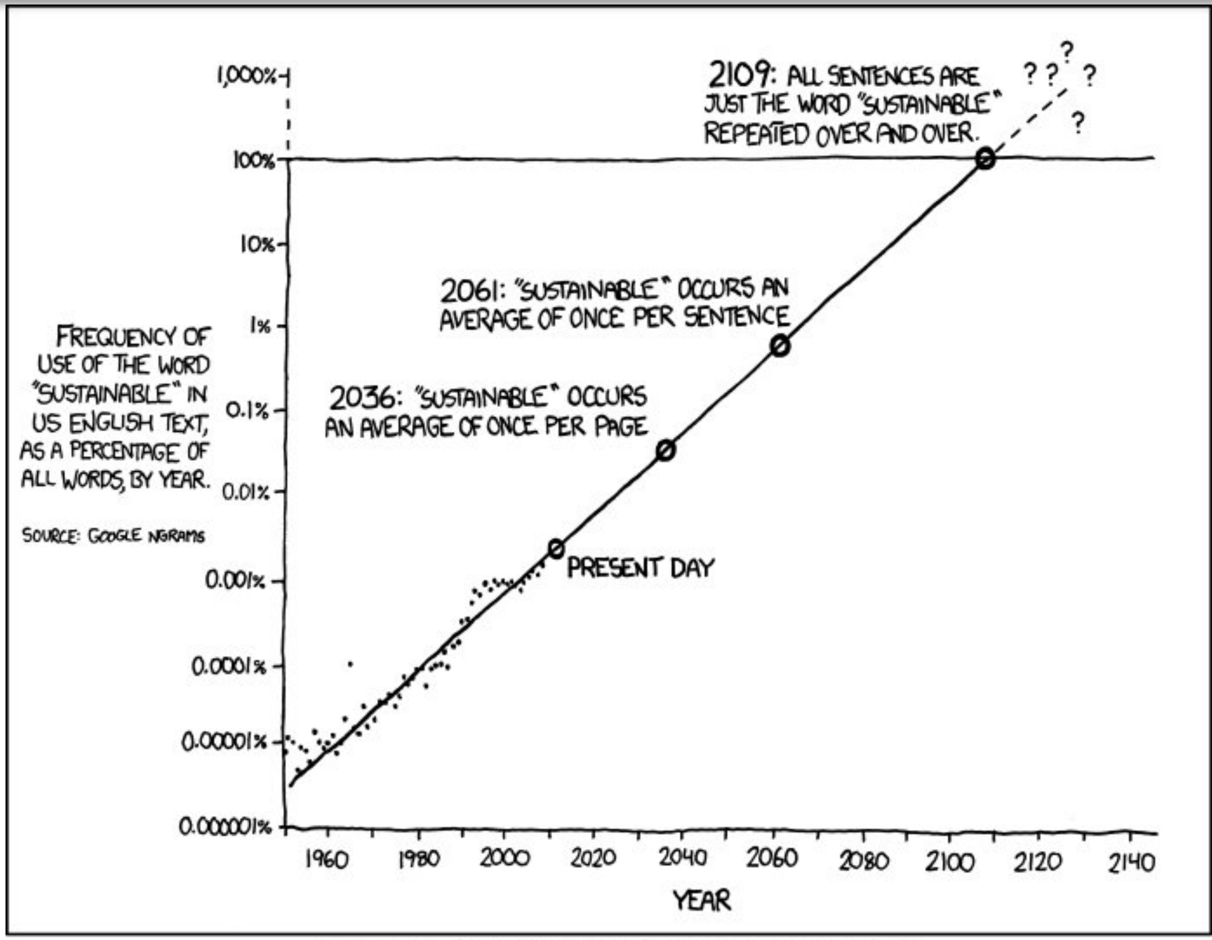
Domes/bowls $$z = ax + by + cx^2 + dy^2 + exy + f$$

3D equivalent of (simple) curvilinear

More complex $$z = ax + by + cx^2 + dy^2 + ex^3 \ldots$$

Capture more complex landscapes

The dangers of extrapolation



THE WORD "SUSTAINABLE" IS UNSUSTAINABLE.

Source: xkcd