

## 点分布 (point distribution) とは

- 地点の集合
- より厳密には「地理空間において、長さや大きさを無視できるサイズの同一種類の施設または事象が、空間内に散らばったもの」



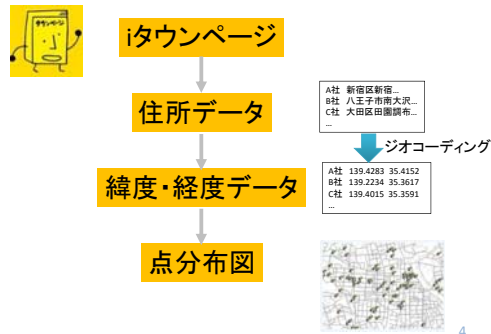
2

観光関連では、どんな点分布があるだろうか？



3

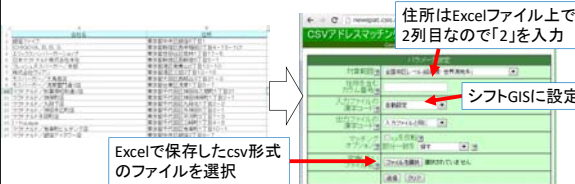
## 点分布データの取得の一例



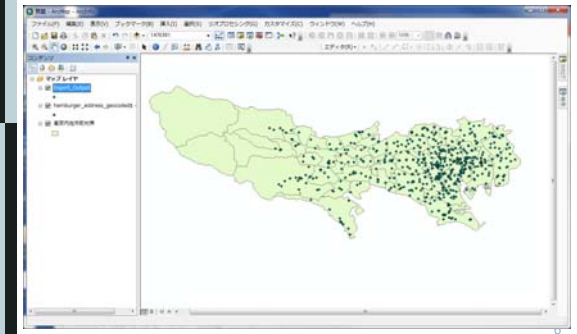
4

## ジオコーディングのやり方

1. Excel上で住所録をcsv形式で保存
2. 「東大CSVアドレスマッチングサービス」を利用  
→もとの住所録の右側に緯度・経度の列がついたテキストファイルが一瞬で帰ってくる
3. Excel上でxls形式で保存



## ArcGISでの点分布の読み込み



## 今日の授業の目的

地域内に分布する特定の施設や事象について、その空間的な分布特性を分析する手法について学ぶ



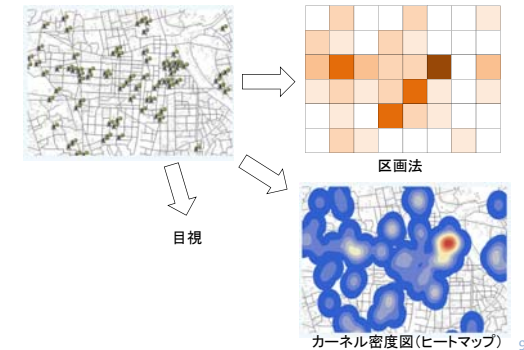
7

## 点分布分析のメニュー

- 地図上で眺める
- カーネル密度図(ヒートマップ)にする
- 各エリアにある点の個数を数える
- 密集/分散傾向を統計的に示す
- 2種類の点分布の場合:  
相互近接傾向/相互回避傾向を示す



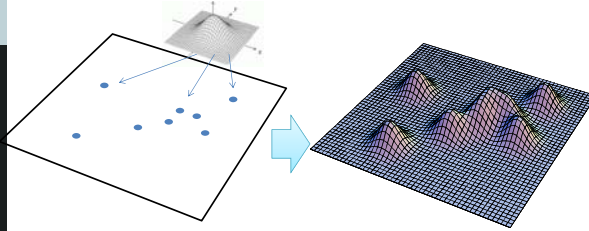
## 「点が多い箇所」を見つけるには？



9

## カーネル密度図の原理

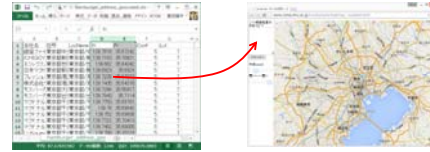
各点上に砂山(カーネル)を盛り、それらが重ねあつたできたものの高さを考える



10

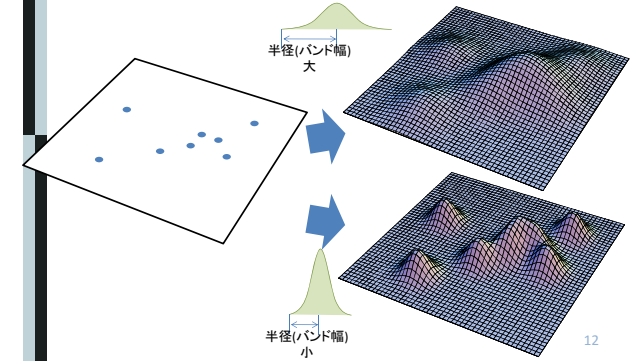
## カーネル密度図を体験してみよう

- zドライブの16観光地理情報学フォルダの中の
  - ①Webページショートカット「HeatmapVisualizer」と
  - ②Excelファイル「東京〇〇住所録.xlsx」をそれぞれ開く
- ②の「緯度・経度」を①のData欄にコピーして[Visualize!]
- 半径や濃さをいじりながら、いろいろ様子を見てみよう



11

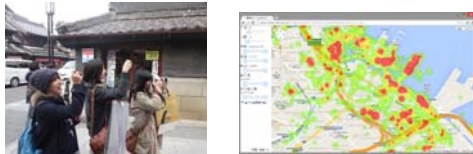
## 半径(バンド幅)の意味



12

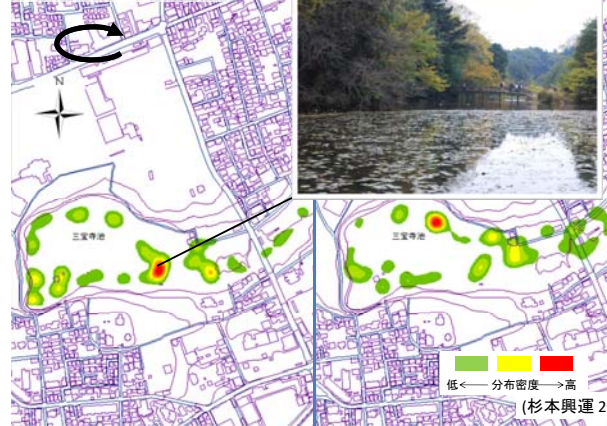
## 観光への応用: 写真撮影位置分析

- どのエリアが旅行者の関心を集めているかわかる
- 逆に何が見過ごされているかもわかる
- 観光動線や眺望ポイントの整備計画の役に立つ



13

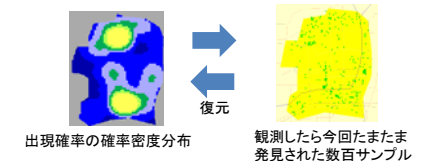
## 写真撮影位置分析の例



(杉本興運 20:)

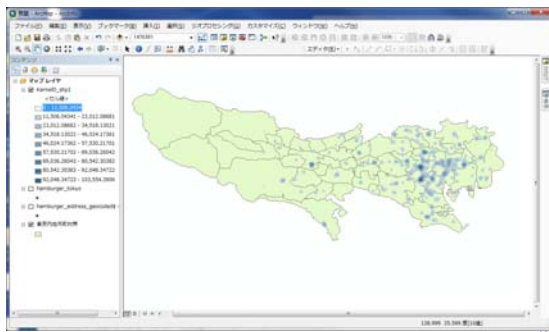
## なぜ「カーネル密度推定」というか

- 観測された点分布は、ある確率密度分布(出現ポテンシャル)に従って出現した「標本群」とみなせる
  - 例: ホタルの分布
- カーネル密度図をつくることは、観察された標本群から元の確率密度分布を「推定」することに相当



15

## ArcGISでのカーネル図



16

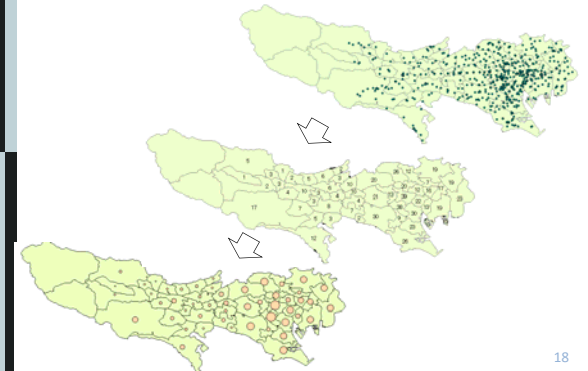
## 点分布分析のメニュー

- 地図上で眺める
- カーネル密度図(ヒートマップ)にする
- 各エリアにある点の個数を数える
- 密集/分散傾向を統計的に示す
- 2種類の点分布の場合:  
相互近接傾向/相互回避傾向を示す



17

## 地区ごとの点の数を求める



18

## 点分布分析のメニュー

- 地図上で眺める
- カーネル密度図(ヒートマップ)にする
- 各エリアにある点の個数を数える
- 密集/分散傾向を統計的に示す
- 2種類の点分布の場合:  
相互近接傾向/相互回避傾向を示す



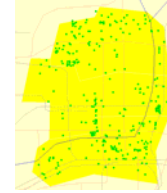
9

## 密集傾向か分散傾向か

- 密集傾向:互いに接近
- 分散傾向:互いに回避

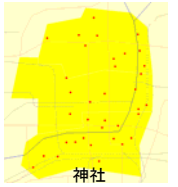
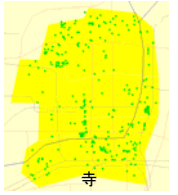


- 統計的に判定する方法
  - 最近隣距離法
  - K-関数法



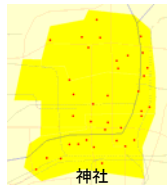
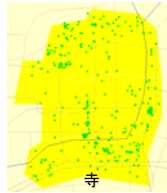
## 最近隣距離法

- 各点から見た最近隣点までの距離の平均値  $\bar{d}$  を調べる
  - 京都の寺:  $\bar{d}=54.9\text{m}$
  - 京都の神社:  $\bar{d}=340.8\text{m}$
- $\bar{d}$  が小さければ密集傾向
- $\bar{d}$  が大きければ分散傾向
- 有意性判定基準:
  - 面積  $a$  のエリアに  $n$  個の点をランダムに散らすと、 $\bar{d}$  は平均値  $\mu = \frac{1}{\sqrt{2}}$  標準偏差  $\sigma = \sqrt{\frac{a-\pi}{4\pi}}$  の正規分布に従う ( $\lambda = n/a$ )



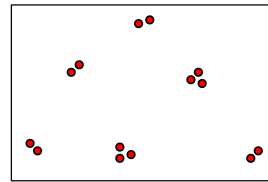
## 最近隣距離法

- 京都の寺の場合
  - 実測値:  $\bar{d}=54.9\text{m}$
  - ランダムと仮定したときの理論値:  
 $\mu=106.9$   $\sigma=2.6$
  - $p=2.8 \times 10^{-89} < 5\%$   
→ランダムでは明らかに発生し得ない  
→「密集傾向にある」と統計的に言える
- 京都の神社の場合:
  - 実測値:  $\bar{d}=340.8\text{m}$
  - ランダムと仮定したときの理論値:  
 $\mu=386.2$   $\sigma=34.2$
  - $p=9.2 \times 10^{-2} > 5\%$   
→ランダムなら気まぐれにおきる程度  
→「密集傾向にある」とは統計的に言えない



## 最近隣距離法の問題点

数個セットで離散分布している場合に「密集型」と誤判定されてしまう

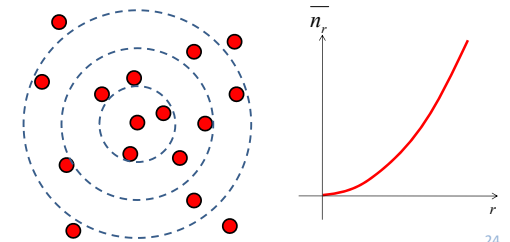


これを防げるのが、次に述べるK関数法

23

## K-関数法

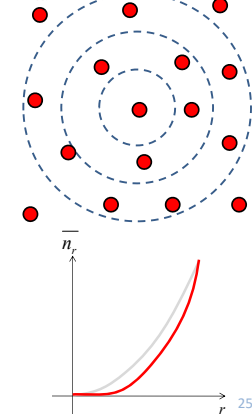
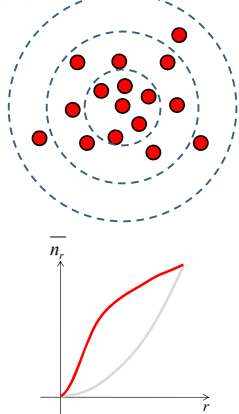
まず各点から見て半径  $r$  にある点の平均個数  $\bar{n}_r$  を数える



24

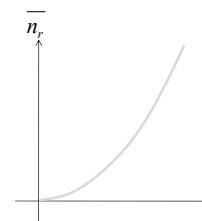
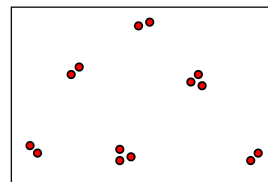
密集型

離散型



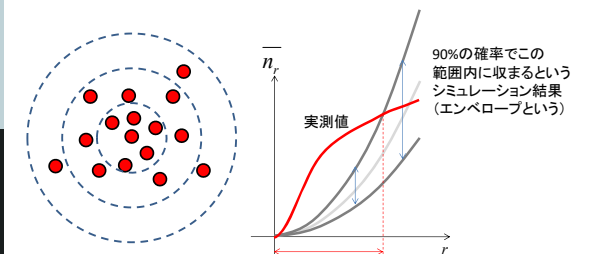
25

では、この場合は?



26

判定基準:シミュレーションを利用

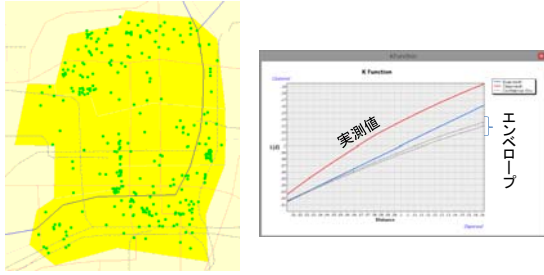


90%の確率でこの範囲内に収まるというシミュレーション結果(エンベロープという)

この範囲でエンベロープからはみ出ているので「ミクロスケールで点分布は密集型である」と判断

27

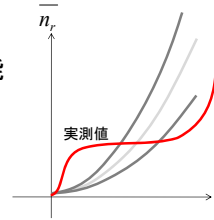
## 京都の寺の場合



28

## K-関数法のメリット

- 密集・分散傾向をスケールごとに判定可能
- 例: ホテル
- ミクروسケールでは密集
- マクروسケールでは分散



29

## 点分布分析のメニュー

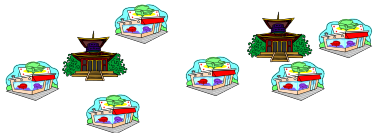
- 地図上で眺める
- カーネル密度図(ヒートマップ)にする
- 各エリアにある点の個数を数える
- 密集/分散傾向を統計的に示す
- 2種類の点分布の場合:  
相互近接傾向/相互回避傾向を示す



30

## 二種類の点分布の関係

- 近接傾向: AはBの近くに多い
  - 土産物屋はお寺の近くに多い
  - 温泉は火山の近くに多い
- 回避傾向: AはBの近くにが少ない
  - 病院は工場の近くに少ない



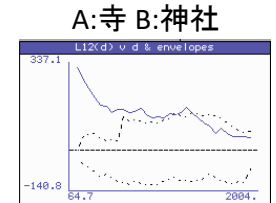
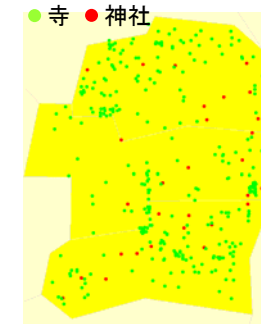
31

## 二種類の点分布の関係の分析手法

- **相互最近隣距離法**
  - 各Aから見て最も近いBまでの距離の平均値を調べる
  - そしてBが空間上にランダムに分布した場合のシミュレーション結果と比べる
- **相互K関数法**
  - 各Aから見て半径r mにあるBの個数の平均値を数える
  - そしてBが空間上にランダムに分布した場合のシミュレーション結果と比べる

32

## 寺・神社の場合



寺の近くに神社が多く立地する傾向にある

33

## 注意点!

- AはBの近くに多い ≠ BはAの近くに多い
- 駅の近くには銀行が多いが、銀行の近くには駅が多いわけではない
  - お寺の近くには土産物屋が多いが、土産屋の近くにはお寺が多いわけではない



34

## 点分布分析のメニュー

- 地図上で眺める
- カーネル密度図(ヒートマップ)にする
- 各エリアにある点の個数を数える
- 密集/分散傾向を統計的に示す
- 2種類の点分布の場合:  
相互近接傾向/相互回避傾向を示す



35

## 今日のキーワード

- 点分布分析
- ジオコーディング
- 緯度・経度データ
- カーネル密度図
- 最近隣距離法
- K関数法
- 相互最近隣距離法・相互K関数法

