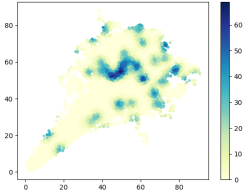
VE414 Presentation

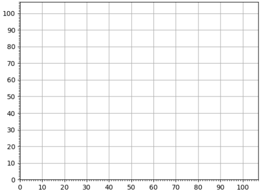
Group 9

August 5, 2019

Overview of Density Estimation

- Known: Many records of the number of Tayes
- Required: Estimation of Tayes in the Forest
- Solution:
  - Divide map into grid (107×107)
  - Calculate the density on the grid according to the records within 1 meter
    - For two records using *density estimation of two intersected circles*
    - For more than two records taking the mean of the estimation
  - Tayes sample according to the density



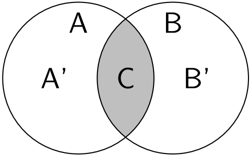


Density Estimation of two Intersected Circles

Suppose there are two detected circles which intersect with each other. We only know

$$\text{Tayes}[A = A' \cup C] = a, \quad \text{Tayes}[B = B' \cup C] = b$$

from the magic of Hermione.



If Tayes are (approx) uniformly distributed in a circle, and $S_A \approx S_B$, $a \approx b$,

$$\text{Density}[C] \approx \left(\frac{a}{S_A} + \frac{b}{S_B}\right)/2, \quad c = \text{Density}[C] \cdot S_C,$$

$$\text{Density}[A'] \approx \frac{a - c/2}{S_A - S_C}, \quad \text{Density}[B'] \approx \frac{b - c/2}{S_B - S_C}.$$

However, these conditions don't always hold, so we need to extend our model. Keep the assumption that Tayes are uniformly distributed in a circle, the number of Tayes x in area C should follow binomial distributions $B(a, p_a = S_C/S_A)$ and $B(b, p_b = S_C/S_B)$ in two circles.

$$P[X = x] = \binom{a}{x} p_a^x (1 - p_a)^{a-x} \binom{b}{x} p_b^x (1 - p_b)^{b-x},$$

$$P_x = P[X = x] \Bigg/ \sum_{i=0}^{\min\{a,b\}} P[X = i],$$

$$\text{Density}[A'] = \frac{1}{S'_A} \sum_{i=0}^{\min\{a,b\}} (a - i)P_i, \quad \text{Density}[B'] = \frac{1}{S'_B} \sum_{i=0}^{\min\{a,b\}} (b - i)P_i,$$

$$\text{Density}[C] = \frac{1}{S_C} \sum_{i=0}^{\min\{a,b\}} iP_i.$$

Guassian Mixture Distribution

- Assumption: The distribution of the fruits from a tree follows the *Bivariate Gaussian Distribution*.
- The distribution of Tayes on the ground $\sim$ Guassian Mixture Distribution

$$p(x) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k)$$

  where $\sum_{k=1}^{K} \pi_k = 1$. K is the number of trees, μ_k is the position of the tree, π_k and Σ_k determines some unique properties of a tree.
- Aim: Fit Guassian Mixture Model(GMM) with Tayes Samples

Fit GMM with Expectation-Maximization algorithm (EM)

- E-step: calculate the posterior probabilities for $k \in \{1, \cdots, K\}, n \in \{1, \cdots, N\}$

$$\gamma_{nk} = p(k|x_n) = \frac{\pi_k \mathcal{N}(x_n|\mu_n, \Sigma_n)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(x_n|\mu_j, \Sigma_j)}$$

- M-step: calculate the value for $\pi_k, \ \mu_k, \ \Sigma_k$

$$\mu'_k = \frac{\sum_{n=1}^{N} \gamma_{nk} x_n}{\sum_{n=1}^{N} \gamma_{nk}}$$

$$\Sigma'_k = \frac{\sum_{n=1}^{N} \gamma_{nk} (x_n - \mu'_k)(x_n - \mu'_k)^T}{\sum_{n=1}^{N} \gamma_{nk}}$$

$$\pi'_k = \frac{\sum_{n=1}^{N} \gamma_{nk}}{N}$$

- Loop until log likelihood convergence:

$$\ln p(X|\mu, \Sigma, \pi) = \sum_{n=1}^{N} ln\{\sum_{k=1}^{K} \pi_k \mathcal{N}(x_n|\mu_k, \Sigma_k)\}$$

- Pick a k = $\arg_k$ min BIC(Bayesian Information Creterion):

$$BIC = -2ln(L) + ln(N) * K$$

Generalize tree distribution to unobserved areas

- Assumptions
  - Assumption 1: The number of trees in each grid of a certain size follows Poisson distribution with parameter λ.
  - Assumption 2: In each grid, the trees are located uniformly.
- Estimate λ

$$L(\lambda) = \ln \prod f(k_i|\lambda) = -n\lambda - \sum \ln(k_i!) + (\sum k_i) \ln \lambda$$

$$\hat{\lambda}_{MLE} = \frac{1}{n} \sum k_i$$

- Random Generation
  1. For each unobserved grid, sample the number of trees in it according to the estimated Poisson distribution
  2. For grids with trees, randomly generate positions for these trees

Result

- Total estimated number of trees in the forest: $25 + 79 = 104$
- An example of possible tree positions