**Model description**:

Say we have 3 activity centers (AC), 20 grid cells, and 10 time segments. Each time segment is associated with a single AC. Our goal is to identify the location of these AC’s based on the observed visitation frequency.

Subscripts:

l = grid cell

t = time segments

k = AC

The frequency data for time segment t is given by the vector and the total number of visits in time segment t is given by . Let denote the cluster membership of time segment t. We assume that:

where and is the distance between location l and activity center k.

In this model, we have to estimate and the location of the 3 activity centers.

Another way to express this same model is to integrate over the cluster membership , yielding:

This mixture model allows for a for more variability of the data to be accounted for through multiple , resulting in a better fit compared to a model with a multinomial likelihood but no latent clusters.

**Priors**:

We assume a uniform prior over the set of locations that were visited at least once for activity center k:

To determine the number of activity centers, we adopted a truncated stick-breaking prior on **:**

**FCD’s**:

* Location of activity center

where are a set of coordinates for one of the potential locations.

We will sample this using a categorical distribution

This implies that: