Sampler for Different Ranking Models

Shuxiao Chen

Simulating farmers preference for different varieties requires sampling from different ranking models. Here I code three samplers for sampling from the Thurstone model (the method used in the preprint paper), the Plackett-Luce model and the Mallows model, repectively.

1 Sampler for the Thurstone model

Thurstone (1927) proposed a ranking process where the ranking π_j of t objects given by a judge j is determined by the relative ordering of t random utilities $\mathbf{y}_j = (y_{1j}, \ldots, y_{tj})^T$, drawing from some distribution F_j. The most basic case is:

$$\mathbf{y}_j \sim N(\boldsymbol{\mu}, \sigma^2 I)$$

where $\boldsymbol{\mu}$ is the vector of expected utilities and each y_{ij} is independent with the same variance σ^2. In the case of the preprint paper, the authors assume $\sigma = 1$.

So the R code is:

```
thurstoneSampler = function(mu, theta = c(0, 1)){
  #mu: vector of expected utility (like phenotypic traits)
  #theta = (mean, sd): the parameters passed to normal distributions
  nitem = length(mu)
  utility = mu + rnorm(nitem, theta[1], theta[2]) #add normal noise
  names(utility) = 1:nitem #assign labels to items
  #ranking is the items listed in the order
  #A succeeds B succeeds C, etc.
  ranking = as.numeric(names(sort(utility, decreasing = TRUE)))
  return(ranking)
}
```

2 Sampler for the Plackett-Luce model

According to Yellott (1977), when the utilities are drawn independently from a Gumbell distribution:

$$y_{ij} \sim \text{Gumbell}(\mu_i, \beta)$$

where μ_i is the location parameter and $\beta > 0$ is a arbitrary fixed scale parameter, then the ranking induced by the Thurstone model is identical to the Plackett-Luce model with parameterization:

$$\gamma_i = \exp\{\mu_i/\beta\}$$

This interpretation gives rise to the following sampler:

```
library(FAdist) #for sampling from Gumbel distribution
#this sampler is based on the equivalence
#of Thurstonian model and Luce's Choice
#when utilities are drawn independently
plSampler = function(ability, beta = 2){
  #ability is the vector of parameters in Plackett-Luce model
  #beta is the scale parameter of Gumbel distribution
  nitem = length(ability)
  #compute the location parameter of Gumbel distribution
  mu = beta * log(ability)
  #sample from Gumbell distribution
  utility = rgumbel(nitem, scale = beta, location = mu)
  names(utility) = 1:nitem #assign labels to items
  #get rank
  ranking = as.numeric(names(sort(utility, decreasing = TRUE)))

  return(ranking)
}
```

3 Sampler for the Mallows model

Given the reference ranking (the mode) σ and the dispersion parameter ϕ, we can sample from the Mallows model using Repeated Insersion Model (RIM) proposed by Doignon et al. (2004) The code is:

```
library(stats) #sampler from multinomial
```

```
mallowSampler = function(sigma, phi){
  #sigma is the reference ranking in the form (sig1, sig2,..., sigm )
  #phi is the dispersion parameter
  if(phi <= 0 | phi > 1){
    warning('phi should be in (0, 1]')
  }
  nitem = length(sigma)
  #initialize
  ranking = c(sigma[1])

  #repeated insersion
  for(i in 2:nitem){
    j = 1:i #the place to insert
    if(phi == 1){
      pvec = rep(1/i, i) #insersion probabilities
    } else{
      pvec = phi^(i-j) * (1-phi) / (1-phi^i) #insersion probabilities
    }
    ind = which(rmultinom(1, 1, pvec) == 1) - 1 #insert after which index
    ranking = append(ranking, sigma[i], ind)
  }

  return(ranking)

}
```