Coupon collector's problems with statistical applications to rankings

Sigeo Aki · Katuomi Hirano

Abstract   Some new exact distributions on coupon collector's waiting time problems are given based on a generalized Pólya urn sampling. In particular, usual Pólya urn sampling generates an exchangeable random sequence. In this case, an alternative derivation of the distribution is also obtained from de Finetti's theorem. In coupon collector's waiting time problems with m kinds of coupons, the observed order of m kinds of coupons corresponds to a permutation of m letters uniquely. Using the property of coupon collector's problems, a statistical model on the permutation group of m letters is proposed for analyzing ranked data. In the model, as the parameters mean the proportion of the m kinds of coupons, the observed ranking can be intuitively understood. Some examples of statistical inference are also given.

Keywords   Generalized Pólya urn · Dirichlet distribution · Exchangeability · Likelihood ratio test · Permutation