Advanced Qualitative Data Analysis
STATISTICS 455

Noelle I. Samia, Ph.D.

Homework # 1:

- Let $\mathbf{X_i} = (X_{i1}, \cdots, X_{id})$ be a d-dimensional random vector, $i = 1, \cdots, K$.
  Let $(\mathbf{X}_1, \cdots, \mathbf{X}_K) \sim$ multinomial $(n, \pi_1, \cdots, \pi_K)$, where $\pi_i = (\pi_{i1}, \cdots, \pi_{id})$ is a d-dimensional vector, $i = 1, \cdots, K$.
  Show that $(X_{1+}, \cdots, X_{K+}) \sim$ multinomial $(n, \pi_{1+}, \cdots, \pi_{K+})$, where $X_{j+} = \sum_{i=1}^{d} X_{ji}$ and $\pi_{j+} = \sum_{i=1}^{d} \pi_{ji}$, for $j = 1, \cdots, K$.

- Let $(X_1, X_2, \cdots, X_6) \sim$ multinomial $(n, \pi_1, \cdots, \pi_6)$.
  Show that $(X_1 + X_3, X_2, X_4 + X_5) \sim$ multinomial $(n, \pi_1 + \pi_3, \pi_2, \pi_4 + \pi_5; \sum_{i=1}^{5} \pi_i \leq 1)$.

- The probability integral transform theorem shows that if X is continuous with cdf F_X, then $Y = F_X(X)$ is uniformly distributed on $(0, 1)$. In this problem, we investigate the relationship between discrete random variables and uniform random variables.
  Let X be a discrete random variable with cdf F_X and define the random variable Y as $Y = F_X(X)$. Let U be a uniform random variable on $(0, 1)$.
  Show that the cdf of Y satisfies $F_Y(y) \leq P(U \leq y) = y$, for all $0 < y < 1$ and $F_Y(y) < P(U \leq y) = y$, for some $0 < y < 1$.
  Note that in this case, Y is said to be stochastically greater than a Uniform $(0, 1)$ random variable.

- Problem # 1.7

- Problem # 1.8