Stat 512
Emerson, Fall 03
Homework #4
October 23, 2003

Supplemental written problems due Wednesday, October 29, 2003 at the beginning of class.

Consider a sample of independent, identically distributed X_i, $i = 1, \ldots, n$, and a sample of independent, identically distributed Y_i, $i = 1, \ldots, m$. We are interested in the null hypothesis that the X_i's have the same distribution as the Y_i's, and we are interested in finding the mean and variance of two statistics under that hypothesis.

1. Suppose we are interested in the probability that a randomly chosen X will be greater than a randomly chosen Y. We estimate this by
$$U = \sum_{i=1}^{n} \sum_{j=1}^{m} 1_{[X_i \geq Y_j]}.$$
Find the expectation and variance of U under the null hypothesis stated above.

2. Suppose we transform all of the random variables from the scale they were originally measured on to their ranks
$$R_i = rank(X_i) = \sum_{j=1}^{n} 1_{[X_j \leq X_i]} + \sum_{j=1}^{m} 1_{[Y_j \leq X_i]}$$
$$S_i = rank(Y_i) = \sum_{j=1}^{n} 1_{[X_j \leq Y_i]} + \sum_{j=1}^{m} 1_{[Y_j \leq Y_i]}$$
and define
$$R = \sum_{i=1}^{n} R_i.$$
Find the mean and variance of R. (Hint: Under the null hypothesis, R has the same distribution as the sum of m numbers randomly chosen without replacement from the integers $\{1, \ldots, m + n\}$.

3. Find the correlation between U and R.