Indian Institute of Technology Guwahati
Statistical Inference (MA682)
Problem Set 08

1. Let X be a random variable with probability density function

$$f(x) = \begin{cases} 2(1-x) & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find the transformation $Y = g(X)$ such that the random variable Y has the uniform distribution over $(-1, 1)$.

2. Ten points are chosen randomly and independently on the interval $(0, 1)$.

   (a) Find the probability that the point nearest 1 exceeds 0.90.

   (b) Find the number c such that the probability is 0.5 that the point nearest zero exceeds c.

3. Find the probability that the range of a random sample of size n from the population with probability density function

$$f(x) = \begin{cases} 2e^{-2x} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

does not exceed 4. Note that the range of a random sample is defined by $X_{(n)} - X_{(1)}$.

4. Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be two independent random sample from a population with continuous cumulative distribution function $F(\cdot)$. Let $S_m(\cdot)$ be the empirical distribution function based on the random sample $X_1, X_2, \ldots, X_m$.

   (a) Show that $S_m(Y_i)$, $i = 1, 2, \ldots, n$, is uniformly distributed over the set of points $\left\{0, \frac{1}{m}, \frac{2}{m}, \ldots, 1\right\}$.

   (b) Show that the distribution of $P_{(i)} = mS_m\left(Y_{(i)}\right)$ is given by

$$P\left(P_{(i)} = j\right) = \frac{\binom{m+n-i-j}{m-j}\binom{i+j-1}{j}}{\binom{m+n}{n}}, \quad j = 0, 1, \ldots, m.$$

   (c) Show that the distribution of $S_m(Y_{(j)}) - S_m(Y_{(k)})$, $k < j$, is the same as the distribution of $S_m(Y_{(j-k)})$.

5. The following are 30 time lapses in minutes between eruptions of Old Faithful geyser in Yellowstone National Park, recorded between the hours of 8 am and 10 pm on a certain day, and measured from the beginning of one eruption to the beginning of the next:

$$68, 63, 66, 63, 61, 44, 60, 62, 71, 62, 62, 55, 62, 67, 73$$
$$72, 55, 67, 68, 65, 60, 61, 71, 60, 68, 67, 72, 69, 65, 66$$

A researcher wants to use these data for inference purposes, but is concerned about whether it is reasonable to treat such data as a random sample. What do you think? Justify your answer.

6. The 20 observations below were chosen randomly from the continuous uniform distribution over $(0, 1)$, recorded to four significant figures, and rearranged in increasing order of magnitude. Test the null hypothesis that the square roots of these numbers also have the continuous uniform $(0, 1)$ distribution.

0.0123	0.1039	0.1954	0.2621	0.2802
0.3217	0.3645	0.3919	0.4240	0.4814
0.5139	0.5846	0.6275	0.6541	0.6889
0.7621	0.8320	0.8871	0.9249	0.9634

7. The Educational Testing Service reports that the 75th percentile for scores on the quantitative portion of the Graduate Record Examination (GRE) is 693 in a certain year. A random sample of 15 first-year graduate students majoring in statistics report their GRE quantitative scores as $690, 750, 680, 700, 660, 710, 720, 730,\ 650, 670, 740, 730, 660, 750,$ and 690 . Are the scores of students majoring in statistics consistent with the 75th percentile value for this year?