Assigned: 09 February 2017

Project #4 – Integrals and Intervals

EE 511: Spring 2017

Due: 16 Feb at 18:00 in EE511 project locker (near EEB B24). Late penalty: 15% per day.

1. Approximate the following integrals using a Monte Carlo simulation. Compare your estimates with the exact values (if known):
   a. $\int_{-2}^{2} e^{x+x^2}\, dx$.
   b. $\int_{-\infty}^{\infty} e^{-x^2}\, dx$.
   c. $\int_0^1 \int_0^1 e^{-(x+y)^2}\, dy\, dx$.

2. Define the random variable $X = Z_1^2 + Z_2^2 + Z_3^2 + Z_4^2$ where $Z_k \sim N(0,1)$. Then $X \sim \chi^2(4)$. Generate 10 samples from X by first sampling Z_i for $i = 1, 2, 3, 4$ and then computing X. Plot the empirical distribution $F_{10}^*(x)$ for your samples and overlay the theoretical distribution $F(x)$. Estimate a lower bound for $\|F_{10}^*(x) - F(x)\|_\infty$ by computing the maximum difference at each of your samples: $\max_{x_i} |F_{10}^*(x_i) - F(x_i)|$. Then find the 25th, 50th, and 90th percentiles using your empirical distribution and compare the value to the theoretical percentile values for $\chi^2(4)$. Repeat the above using 100 and 1000 samples from X.

3. A geyser is a hot spring characterized by an intermittent discharge of water and steam. Old Faithful is a famous cone geyser in Yellowstone National Park, Wyoming. It has a predictable geothermal discharge and since 2000 it has erupted every 44 to 125 minutes. Refer to the addendum data file that contains waiting times and the durations for 272 eruptions. Compute a 95% statistical confidence interval for the waiting time using data from only the first 15 eruptions. Compare this to a 95% bootstrap confidence interval using the same 15 data samples. Repeat these calculation using all the data samples. Comment on the relative width of the confidence intervals when using only 15 samples vs using all samples.