MATH 5301 Final Exam Review

1. Under certain asymptotic conditions, a researcher expects the residuals from a fitted model to approximately normal (see `residual-data.txt`).

   (a) Plot a histogram of the residuals. Do they appear to be approximately normal?

   (b) Build a qq-plot to assess normality of the residuals.

   (c) Test the null hypothesis that the residuals are normally distributed.

2. Customer service phone calls arrive at a call center at a rate of $\lambda = 56.3$ per minute. Under a Poisson process model, we would expect the waiting times between incoming calls to be exponentially distributed with rate parameter $\lambda = 56.3$. Test whether the waiting times W_i stored in `waiting-times.txt` have this distribution.

3. A research team is interested in testing whether data from two experimental processes, X_{i1} and X_{i2}, are identically distributed (see the file `experimental-data.txt`).

   (a) One problem is that the means $\bar{x}_1$ and $\bar{x}_2$ and standard deviations s_1 and s_2 from these samples are approximately equal. Despite this, the researchers have reason to believe that the shapes of the distributions may be different. Under these circumstances, which test would be better to use: Mann-Whitney-Wilcoxon or Kolmogorov-Smirnov?

   (b) Confirm that the means and standard deviations of the samples are approximately the same.

   (c) Plot a histogram of each data set to visually compare the distributions. Do they appear to be identically distributed?

   (d) Visually compare the samples with a qq-plot. Based on the qq-plot, do they appear to be identically distributed?

   (e) Test whether the samples are identically distributed using the Mann-Whitney-Wilcoxon test.

   (f) Test whether the samples are identically distributed using the Kolmogorov-Smirnov test.

   (g) Evaluate your answer to part (a) given the answers to parts (b) through (f).

4. An economist needs to fit models for two time series, X_t and Y_t, and would prefer to use an $MA(q)$ model with $q \leq 3$ or an auto-regressive model (see `time-series.txt`).

   (a) Plot the auto-correlation function for X_t. What model specification should be used?

   (b) Plot the auto-correlation function for Y_t. What model specification should be used?

5. Suppose Z_t is a white-noise process, such that $Z_t = \pm 1$, each with probability $\frac{1}{2}$, and consider the simple random walk

$$X_t = X_{t-1} + Z_t.$$

   Given that $X_{200} = 36$, compute a 95% prediction interval for X_{250}.

6. Let $Z_t \sim N(0, 5^2)$ be a white noise process, and consider the $MA(3)$ model

$$X_t = 0.6Z_t + 0.3Z_{t-1} + 0.1Z_{t-2}.$$

Given that $X_{200} = 0.87$, compute a 95% prediction interval for X_{250}.

7. A business analyst is modeling the time series data in `sales.txt` and is interested in periodic variation at three frequencies: weekly ($\omega = 2\pi/7$), monthly ($\omega = 2\pi/30$), and annually ($\omega = 2\pi/365$). Rank the contribution of these three frequencies to the variance from highest to lowest.