Applied Math and Statistics
Exam #2 **AMS-223.**

4. (Take-home, 50 points, **due on Tu March 7th**). The file **googletrendsUCSC.csv** (available online) contains a time series of the *interest over time* in the United States for the term "University of California, Santa Cruz" measured monthly as a Google index from February 2002 until January 2017.

   (a) (15 points) Perform a Bayesian spectral analysis of the first differences of these data based on a single component harmonic regression model. Discuss your results.

   (b) (35 points) Fit a polynomial trend model of order one to the original data, i.e., consider a model of the form:

$$
\begin{aligned}
y_t &= \theta_t + \nu_t, \quad \nu_t \sim N(0, v), \\
\theta_t &= \theta_{t-1} + w_t, \quad w_t \sim N(0, vW_t),
\end{aligned}
$$

   with v unknown and W_t specified by a discount factor $\delta \in (0, 1]$.

   i. (12 points) Find the optimal value of δ, $\hat{\delta}$, that maximizes the observed predictive density (as in page 58 of West & Harrison).

   ii. (15 points) For this optimal discount factor $\hat{\delta}$, provide summaries of the filtering distribution, the one-step ahead forecast distribution, and the smoothing distribution (i.e., the distributions of $(\theta_t|\mathcal{D}_t)$, $(y_t|\mathcal{D}_{t-1})$, and $(\theta_t|\mathcal{D}_T)$, with T the total number of observations and $t < T$. Discuss your results.

   iii. (8 points) Provide the 12-steps ahead forecast distributions for θ_t, and y_t given the observed data up to February 2017, i.e., summarize the distributions of $(\theta_{T+12}|\mathcal{D}_T)$ and $(y_{T+12}|\mathcal{D}_T)$. Do these distributions capture the structure of these data? What kind of DLM would you consider to improve this model?