Outline of Lecture

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
Where are we?

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
Bayesian Models

- Bayesian models differ from frequentist models only in that the parameters $\theta$ are random.
- This allows us to stack priors to create hierarchical models.
- The Gibbs sampler is a universal algorithm that allows us to efficiently sample from the posterior in hierarchical models.
Example: Rater Model
The Gibbs sampler provides samples from the posterior.

We can use these samples to estimate the posterior distribution (e.g., histogram) or the posterior mean.

JAGS will automatically simulate from the posterior.
Results for the Rater Model

I put priors on $\gamma_k$ and $\delta_\ell$ so that I could do inference on them.
Where are we?

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
What would a frequentist do?

- **Remember**: In the Bayesian framework, to perform inference on a parameter, you must put a prior on it. Otherwise, you must specify its value beforehand.

- **Example**: Suppose there are 3 sections of a class taught by different professors. Let $y_{ij}$ denote the final exam score of student $j$ in class $i$.
  - $y_{ij} | \theta_i \sim N(\theta_i, \sigma^2)$ (instructor effect)
  - $\theta_i \sim N(\mu, \tau^2)$

- Maybe we can instead try to estimate *hyperparameters* such as $\mu$ and $\tau^2$ from the data, then use $N(\hat{\mu}, \hat{\tau}^2)$ as the prior.

- What would our estimates of the parameters be then?
The idea of estimating hyperparameters from the data is called *empirical Bayes*.

It is ultimately a frequentist method because we don’t need to specify a prior on the hyperparameters we estimate!

We get hierarchical models without subjective priors! Is this too good to be true?

What are the challenges of doing this in practice?
Where are we?

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
Problem Setup

• El Niño is characterized by warmer sea surface temperatures (SST) in the equatorial Pacific Ocean.
• Therefore, to predict El Niño, one needs to forecast the SST several months in advance.
• We observe the average monthly SST at different locations in the Pacific:
  \[ z_t \overset{def}{=} (z_t(s_1), \ldots, z_t(s_m)) \]
• Want to forecast SSTs \( \tau \) months in advance: \( z_{T+\tau} \).
• This analysis is taken from Berliner, Wikle, and Cressie (2000).
First Model: Linear Dynamical Model

**State model:** \( y_{t+\tau} = \Phi y_t + \epsilon_t \)  \( \epsilon_t \sim N(0, \Sigma) \)

**Data model:** \( z_t = Ay_t + \delta_t \)  \( \delta_t \sim N(0, \sigma^2 I) \)

- \( A \) contains the first \( k \) principal components of the empirical covariance matrix over the spatial locations.
- \( y_t \) represents weights on those PCs.
- Unknown parameters are \( \Phi, \Sigma, \sigma^2 \). Need to put priors on all of these:

\[
\begin{align*}
\text{vec}(\Phi) & \sim N(\text{vec}(0.9I), 100I) \\
\Sigma^{-1} & \sim \text{Wishart}\left(\frac{1}{100(k-1)}I, k-1\right) \\
\sigma^2 & \sim \text{InverseGaussian}(0.1, 100)
\end{align*}
\]
First Model: Linear Dynamical Model

1997

2.5% and 97.5% quantiles of posterior

1998
Second Model: “Non-Linear” Model

State model: \( y_{t+\tau} = \Phi_t y_t + \epsilon_t \)
\( \epsilon_t \sim N(0, \Sigma) \)

Data model: \( z_t = A y_t + \delta_t \)
\( \delta_t \sim N(0, \sigma^2 I) \)

- Allow \( \Phi_t = \Phi(I_t, J_t) \) to vary with time.
- \( I_t \) classifies the current regime as “cool”, “normal”, or “warm”. Obtained by thresholding the Southern Oscillation Index (SOI):

\[
I_t = \begin{cases} 
0 & \text{if } SOI_t < \text{low threshold} \\
1 & \text{if } SOI_t \text{ in between} \\
2 & \text{if } SOI_t > \text{upper threshold}
\end{cases}
\]

- \( J_t \) is obtained by similarly thresholding a latent process \( W_t \):
\[
W_t | \beta, \tau^2 \sim N(x_t^T \beta, \tau^2)
\]
Second Model: “Non-Linear” Model

Case 1: Long-Lead Forecasting of Sea Surface Temperatures

2.5% and 97.5% quantiles of posterior
Where are we?

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
Problem Setup

- The Eurasian Collared Dove (ECD) was first observed in North America in the 80s and are now spreading quickly throughout the continent.
- They pose a threat to native ecosystems, so we would like to forecast their spread.
- Observe $z_t(s_i)$: number of doves observed at location $s_i$.
- This analysis is taken from Hooten, Wikle, Dorazio, and Royle (2007).
Case 2: Modeling and Forecasting the Eurasian Dove Invasion

Dennis Sun
Stats 253 – Lecture 14
August 13, 2014

Maps showing the spread of the Eurasian Dove Invasion from 1986 to 2003 across the United States.
The Model

Data model: \( z_t(s_i) \mid y_t(s_i), \pi \sim \text{Bin}(y_t(s_i), \pi) \)
\[ y_t \mid \lambda_t \sim \text{Pois}(H \lambda_t) \]

State model: \( \lambda_t = B(\alpha) G(\lambda_{t-1}; \theta) \lambda_{t-1} \)

- \( \pi \) is the probability of observing an animal. Not estimable from this data alone, but the authors estimated it using data collected on a related species.
- \( G \) is a diagonal matrix that models growth, while \( B \) models dispersion.
- The authors go on to put priors on \( \alpha \) and \( \theta \).
Case 2: Modeling and Forecasting the Eurasian Dove Invasion

Results

Posterior means for years in sample

1998 1999 2000 2001 2002 2003
Results

Posterior means for future years

2003  2004  2005  2006  2007  2008

2009  2010  2011  2012  2013  2014

Where are we?

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
• The goal is to predict wind speeds and directions at different locations over the Mediterranean.

• \((x_t(s_i), y_t(s_i))\) is the vector indicating the wind speed in the \(x\)- and \(y\)-directions.

• We assume that \(x_t\) and \(y_t\) are noisy measurements of underlying states \(u_t\) and \(v_t\).
Data on February 1, 2005
A Physical State Model

- The state model is motivated by the Rayleigh friction equations:

\[
\frac{\partial u}{\partial t} - fv = -\frac{1}{\rho_0} \frac{\partial p}{\partial x} - \gamma u \quad \frac{\partial v}{\partial t} + fu = -\frac{1}{\rho_0} \frac{\partial p}{\partial y} - \gamma v.
\]

where \(u, v\) are the east-west and north-south components, \(p\) is the sea-level pressure, and the rest are (unknown) parameters.

- An approximate solution to these equations is given by

\[
\begin{align*}
    u &\approx -\frac{f}{\rho_0(f^2 + \gamma^2)} \frac{\partial p}{\partial y} - \frac{\gamma}{\rho_0(f^2 + \gamma^2)} \frac{\partial p}{\partial x} - 2\frac{\gamma}{\rho_0(f^2 + \gamma^2)} \frac{\partial u}{\partial t} - \frac{1}{\rho_0(f^2 + \gamma^2)} \frac{\partial^2 p}{\partial x \partial t} \\
v &\approx \frac{f}{\rho_0(f^2 + \gamma^2)} \frac{\partial p}{\partial x} - \frac{\gamma}{\rho_0(f^2 + \gamma^2)} \frac{\partial p}{\partial y} - 2\frac{\gamma}{\rho_0(f^2 + \gamma^2)} \frac{\partial v}{\partial t} - \frac{1}{\rho_0(f^2 + \gamma^2)} \frac{\partial^2 p}{\partial y \partial t}
\end{align*}
\]

- We can discretize this as follows:

\[
\begin{align*}
    u_t &= a_1 D_y p_t + a_2 D_x p_t + a_3 u_{t-1} + a_4 D_x p_{t-1} \\
v_t &= b_1 D_x p_t + b_2 D_y p_t + b_3 v_{t-1} + b_4 D_y p_{t-1}
\end{align*}
\]
Results for February 2, 2005
Where are we?

1. Recap of Bayesian Models
2. Empirical Bayes
3. Case 1: Long-Lead Forecasting of Sea Surface Temperatures
4. Case 2: Modeling and Forecasting the Eurasian Dove Invasion
5. Case 3: Mediterranean Surface Vector Winds
6. Wrapping Up the Course
A Few Last Thoughts

- The analysis of spatial and temporal data is really the analysis of correlated data.
- “Model the mean function, use spatial and temporal methods to model the residual.”
- There are two main ways to capture correlations: model the correlation directly (e.g., kriging) and via autoregressions.
- Many methods that work well for time series (e.g., Kalman filter) break down in space because the data are no longer ordered.

Thanks for a great quarter!