Basic jargon

Defn: Stochastic process — family $\{X_i; i \in I\}$ of random variables indexed by a set I.

In practice the jargon is used only when the X_i are *not* independent.

If $I \subset$ Real Line, then we often call $\{X_i; i \in I\}$ a time series. Of course the usual situation is that i actually indexes a time point at which some measurement was made.

Two important special cases are I an interval in R, the real line, in which case we say X is a series in continuous time, and where $I \subset \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ in which case X is in discrete time.

Here is a list of some models used for time series:

- Stochastic Process Models (note the conflict of jargon).
  - Population models
    - Birth and Death Processes — which describe the size of a population in terms of random births and deaths.
    - Markov chain models — where the future depends on the present and not, in addition, on the past. Birth and Death processes are special cases.
    - Galton-Watson-Bienaymé process: a Markov chain model for the size of generations of a populations.

Model specifies: size of nth generation is sum of family sizes of each individual in $n-1$st generation and family sizes have an iid distribution.

* Branching processes: continuous time version of Galton-Watson-Bienaymé.

- Diffusion models
  * Brownian Motion
  * Random Walk
  * Stochastic Diff'l Equations: models like

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t$$

where B is a Brownian motion.

- Linear Time Series Models — linear filters applied to white noise.

This course: about discrete time version of these linear time series models.

Assume throughout that we have data

$$X_0, X_1, \ldots, X_{T-1}$$

where the X_t are real random variables.

A **model** is a family $\{P_\theta; \theta \in \Theta\}$ of possible joint distributions for $\{X_0, \ldots, X_{T-1}\}$.

Goal: guess the true value of θ. (Notice that it is an assumption that the distribution of the data is, in fact one of the possibilities.)

The question is this: is it possible to guess the true value of θ?

Will collecting more data (increasing T) make more accurate estimation of θ possible?

The answer is no, in general.

Example: Galton-Watson-Bienaymé process – even with infinitely many generations you don't get enough data to nail down parameters.

Example: Suppose $(X_0, \ldots, X_{T-1})'$ has multivariate normal distribution with mean vector $(\mu_0, \ldots, \mu_{T-1})'$ and $T \times T$ variance covariance matrix $\mathbf{\Sigma}$.

Big problem: T data points but $T + T(T-1)/2$ parameters to estimate; this is not possible.

To make progress: put restrictions on parameters μ and $\mathbf{\Sigma}$.

For instance you might assume one of the following:

1. Constant mean: $\mu_t \equiv \mu$.

2. Linear trend; $\mu_t = \alpha + \beta t$.

3. Linear trend and sinusoidal variation:
$$\mu_t = \alpha + \beta t + \gamma_1 \sin\left(\frac{2\pi t}{12}\right) + \gamma_2 \cos\left(\frac{2\pi t}{12}\right)$$

Can estimate parameters by regression but still have problem: can't get standard errors.

For instance, we might estimate μ in 1) above using $\bar{X}$.

In that case

$$\begin{aligned} \text{Var}(\bar{X}) &= T^{-2}\text{Var}(\sum X_t) \\ &= T^{-2}\mathbf{1}^t\mathbf{\Sigma}\mathbf{1} \\ &= T^{-2}\sum_{s,t}\mathbf{\Sigma}_{st} \end{aligned}$$

where $\mathbf{1}$ is a column vector of T 1s.

So: we must model $\mathbf{\Sigma}$ as well as μ.

The assumption we will make in this course is of stationarity:

$$\begin{aligned} \text{Cov}(X_t, X_s) &= \text{Cov}(X_{t+1}, X_{s+1}) \\ &= \text{Cov}(X_{t+2}, X_{s+2}) \cdots \end{aligned}$$