Time Series Analysis

Basics of Time Series Analysis

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

Basic Statistical Concepts

About This Lesson

Review of Basic Statistical Concepts

- *Moments of a Distribution:* Fully characterizing the distribution
- *Estimation Methods*: Method of Moments versus Maximum Likelihood Estimation
- *Basic Estimators*: Approach and Sampling Distribution
- *Multivariate Data*: Joint, Marginal, and Conditional Distribution
- *Statistical Inference*: Confidence Intervals and Hypothesis Testing

Georgia Tech

Moments of a Distribution

Moments of a random variable X with density $f(x)$:

l*-th moment:* $m_l' = \mathrm{E}[X^l] = \int_{-\infty}^{\infty} x^l f(x)\,dx$

l*-th central moment*: $m_l = \mathrm{E}[(X-\mu)^l] = \int_{-\infty}^{\infty} (x-\mu)^l f(x)\,dx$

<u>Examples of moments</u>:

- Expectation $\mathrm{E}[X]$
- Variance $\mathrm{E}[(X-\mu)^2]$
- Skewness $S(x) = \mathrm{E}\left[\frac{(X-\mu)^3}{\sigma^3}\right]$
- Kurtosis $K(x) = \mathrm{E}\left[\frac{(X-\mu)^4}{\sigma^4}\right]$

Statistical Estimation

Parametric Statistics: Observe $x_1, \dots, x_n$ (realizations) from a set of random variables $X_1, \dots, X_n \sim f(x; \theta)$ where $f(x; \theta)$ is a density function with parameter θ which is assumed <u>unknown</u>

- **Estimation**: evaluate the unknown parameter θ using the set of observations $x_1, \dots, x_n$ and using the distribution of the random variables $X_1, \dots, X_n$ from which we observe.
- **Approaches:**
  1. Method of Moments (MOM)
  2. Maximum Likelihood Estimation (MLE)

Examples of Classic Estimators

Given the data $\{X_1, \ldots, X_n\}$, estimate mean, variance, skewness and kurtosis with

Sample mean: $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Sample Variance: $\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu})^2$

Sample skewness: $\hat{S}(X_1, \ldots, X_n) = \frac{1}{(n-1)\hat{\sigma}^3}\sum_{i=1}^{n}(X_i - \hat{\mu})^3$

Sample kurtosis: $\widehat{K}(X_1, \ldots, X_n) = \frac{1}{(n-1)\hat{\sigma}^4}\sum_{i=1}^{n}(X_i - \hat{\mu})^4$

Estimators as Random Variables

Given the data $\{X_1, \ldots, X_n\}$, estimate mean, variance, skewness and kurtosis with

Sample mean: $\hat{\mu} = \bar{X}$

Sample Variance: $\hat{\sigma}^2 = S^2$

Sample skewness: $\hat{S}(X_1, \ldots, X_n)$

Sample kurtosis: $\widehat{K}(X_1, \ldots, X_n)$

Why are the estimators random?

- Estimates of statistical summaries are functions of the sample of realizations
- Each sample results in a different estimate hence called ‘sample’ estimates

Sampling Distributions of Estimators

Sampling Distributions: Under normality assumption

$$X_1, \dots, X_n \sim N(\mu, \sigma^2)$$

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

$$\frac{S^2(n-1)}{\sigma^2} \sim \chi^2_{n-1}$$

Properties of the Estimators

1. Unbiasedness
2. Consistency

Method of Moments Estimation

MOM Approach: Equate the distribution moments to the observed moments

$$\mathrm{E}[X^p] = \frac{1}{n}\sum_{i=1}^{n} X_i^p$$

Example: $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$(p = 1)$: $\mathrm{E}[X](= \mu) = \frac{1}{n}\sum_{i=1}^{n} X_i \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$(p = 2)$: $\mathrm{E}[X^2](= \sigma^2 + \mu^2) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$

Maximum Likelihood Estimation

MLE Approach: Maximize the likelihood function of θ given the data.

Joint distribution of $X_1, \ldots, X_n \sim f(x;\theta)$ under independence assumption

$$f(x_1, \ldots, x_n; \theta) = f(x_1; \theta) \ldots f(x_n; \theta)$$

Likelihood function

$$L(\theta : x_1, \ldots, x_n) = f(x_1, \ldots, x_n; \theta)$$

$$\text{MLE: } \hat{\theta} = \text{argmax}_{\theta \in \Theta} L(\theta : x_1, \ldots, x_n)$$

Multivariate Distribution

Joint, Marginal and Conditional distributions:

Joint distribution = *Conditional* × *Marginal*: $f(x,y) = f(x|y)f(y) = f(y|x)f(x)$

For three variables X, Y, Z:

$$f(x,y,z) = f(x|y,z)f(y,z) = f(x|y,z)f(y|z)f(z)$$

For n variables $X_1, \dots, X_n$:

$$f(x_1, \dots, x_n) = f(x_n|x_{n-1}, \dots, x_1)f(x_{n-1}|x_{n-2}, \dots, x_1) \dots f(x_2|x_1)f(x_1)$$

Example: if $X_i|X_{i-1}, \dots, X_1$ is normal with mean μ_i and variance σ_i^2 and if $f(x_1)$ is ignored

$$f(x_n, x_{n-1}, \dots, x_1) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[\frac{-(x_i - \mu_i)^2}{2\sigma_i^2}\right]$$

Georgia Tech

Statistical Inference

Hypothesis Testing

Parameter-based hypothesis testing: $H_0\colon \theta = \theta_0 \ \text{ vs } \ H_a\colon \theta \neq \theta_0$

Distribution-based hypothesis testing:

$$H_0\colon X_1, \ldots, X_n \sim N(\mu, \sigma^2) \ \text{ vs } \ H_a\colon \text{non}-\text{normal distribution}$$

P-value = a measure of the plausibility of H_0

Significance Level = the probability of type 1 error (reject H_0 when it is true)

Confidence Interval A $(1-\alpha)$ confidence interval for θ is $\left(\hat{\theta} - k, \hat{\theta} + k\right)$ with k s.t.

$$\Pr\left(\theta \in \left(\hat{\theta} - k, \hat{\theta} + k\right)\right) = 1 - \alpha$$

Summary