Forecasting and Time Series Econometrics

Derek Li

Contents

1 Introduction

1.1 Basic Definition

Definition 1.1. A ***time series*** is a sequence of numerical values ordered by time.

Definition 1.2. A ***trend*** is a slow, smooth, long-run evolution of a time series over time.

Definition 1.3. A ***cycle*** is a periodic fluctuation of a time series, which may be seasonal or nonseasonal.

1.2 Notation

Table 1.1: Notation

Description	Notation
Object to analyze - time series	$\{y_t\}$
Value at present time t - known value of the series	y_t
Future at time $t+h$ - random variable	Y_{t+h}
Value at future time $t+h$ - unknown value of the random variable	y_{t+h}
Collection of information - information set	$I_t = \{y_1, \cdots, y_t, x_1, \cdots, x_t\}$
Final objective - forecast h-step ahead	$f_{t,h}$
Uncertainty - forecast error	$e_{t,h} = y_{t+h} - f_{t,h}$

2 Statistics Review

2.1 Basic Definition

Definition 2.1. ***Population*** is the entire collection of elements about which information is desired.

Definition 2.2. ***Random process*** is the procedure involving a given population that can conceptually be repeated and leads to certain ***outcomes***. The outcome of a random process is a-priori uncertain.

Definition 2.3. ***Sample space*** is the set of all possible outcomes of the random process.

Definition 2.4. ***Random variable*** (r.v.) is the deterministic function from a sample space to the space of possible values of the variable ($\mathbb{R}$).

Definition 2.5. The ***cumulative distribution function*** (CDF) of r.v. X is

$$F_X(x) = P_X(X \leqslant x), \forall x \in \mathbb{R}.$$

Definition 2.6. The CDF of ***discrete r.v.*** is a step function, and its ***probability mass function*** (PMF) is

$$f_X(x_j) = P_X(X = x_j).$$

Definition 2.7. The CDF of ***continuous r.v.*** is a continuous function, and the Radon-Nikodym derivative of the CDF is the ***probability density function*** (PDF).

Note that $P_X(X = x) = 0$ if X is a continuous r.v.. The set $\{X = x\}$ is an example of a set of measure zero.

Definition 2.8. The ***expected value*** of a r.v. X is the weighted average of possible outcomes of X, where the weights are probabilities. In the discrete case,

$$\mathbb{E}[X] = \sum_x x f(x).$$

In the continuous case,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x.$$

Property 2.1.

$$\mathbb{E}\left[\sum_{i=1}^{K} a_i X_i\right] = \sum_{i=1}^{K} a_i \mathbb{E}[X_i],$$

where X is r.v., $a_i \in \mathbb{R}$.

Definition 2.9. The ***variance*** of a r.v. X is a measure of dispersion of X around its mean μ, denotes as $\mathrm{Var}[X]$ or σ_X^2, and defined as

$$\mathrm{Var}[X] = \mathbb{E}[(X - \mu)^2].$$

In the discrete case,

$$\mathrm{Var}[X] = \sum_x (x - \mu)^2 f(x).$$

In the continuous cases,

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) \mathrm{d}x.$$

Property 2.2. $\mathrm{Var}[aX+bY] = a^2\mathrm{Var}[X]+b^2\mathrm{Var}[Y]+2ab\mathrm{Cov}(X,Y)$, where X, Y are r.v.s., $a, b \in \mathbb{R}$.

Definition 2.10. The k^{th} ***moment*** (or ***non-central moment***) is

$$\kappa_k^* = \mathbb{E}[X^k].$$

Definition 2.11. The k^{th} ***central moment*** is

$$\kappa_k = \mathbb{E}[(X - \mathbb{E}[X])^k].$$

Note that the 0^{th} moment is 1, the 1^{st} non-central moment is mean, the 2^{nd} moment is variance, the 3^{rd} moment is skewness, and the 4^{th} moment is kurtosis. Moment of order $k > 2$ are typically called higher-order moment, and moment of order higher than a certain k may not exist for some distributions.

Definition 2.12. The ***standard deviation*** measures in the usual units:

$$\sigma_X = \sqrt{\mathrm{Var}[X]}.$$

Definition 2.13. ***Covariance*** measures the degree of joint variation of X and Y :

$$\mathrm{Cov}(X, Y) = \sigma_{XY} = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

Definition 2.14. ***Coefficient of correlation*** is the standardized covariance:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y},$$

where $-1 \leqslant \rho_{XY} \leqslant 1$.

Note that independence of X and Y implies $\rho_{XY} = 0$, but $\rho_{XY} = 0$ does not imply that $X \perp Y$.

Definition 2.15. The density of a r.v. Y conditional on the r.v. X taking on a specified value is called the ***conditional density*** of Y given X. In the discrete cases,

$$P(Y = y|X = x) = \frac{P(Y = y, X = x)}{P(X = x)}.$$

In the continuous cases,

$$f_{Y|X}(y|x) = \frac{f_{Y,X}(y, x)}{f_X(x)}.$$

2.2 Linear Regression Model

Assumption 1 (Linearity)**.** $Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + u$.

Assumption 2 (Zero Conditional Mean)**.** $\mathbb{E}[u|X_1, \cdots, X_k] = 0$.

Assumption 3 (Homoscedasticity)**.** $\mathrm{Var}[u|X_1, \cdots, X_k] = \sigma_u^2$.

Assumption 4 (No Serial Correlation)**.** $\mathrm{Cov}(u_t, u_{t-s}) = 0$ for $s = \pm 1, \pm 2, \cdots$.

Assumption 5 (No Perfect Collinearity)**.** There is no exact linear relationship among regressors.

Assumption 6 (Sample Variation in Regressors)**.** $\mathrm{Var}[X_j] > 0$ for $j = 1, \cdots, k$.

Theorem 2.1 (Gauss-Markov Theorem)**.** Under A1 to A6, the ordinary least squares estimators are the best linear unbiased estimators (BLUE) of the unknown population regression coefficients.

Note that for time series data, A4 is typically not satisfied since A1 is too restrictive for time series featuring complex dynamics. Hence, for time series data we typically use other models than linear regression with OLS.