



CFA 二级学习笔记

CFA Level 2 Learning Notes

作者：Ethan Wang

组织：纽约大学

时间：March 2, 2020

版本：0

鸣谢：特别感谢本笔记模板制作者：*Elegant \LaTeX Program*



Vini. Vidi. Vici. — Julius Caesar

目录

1	Introduction to Linear Regression	1
1.1	Linear Regression Introduction	1
1.2	ANOVA Table	1
1.3	Calculating R^2 and SEE	2
1.4	The F-Statistics	2
2	Multiple Regression	3
2.1	Indroduction	3
3	The Term Structure and Interest Rate Dynamics	4
3.1	Spot and Forward Rates	4
3.1.1	Spot Rates	4

第一章 Introduction to Linear Regression

1.1 Linear Regression Introduction

一元回归模型应该看起来长这样：

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i + \epsilon_i, i = 1, \dots, n \quad (1.1)$$

其中，

\hat{Y}_i (Predicted Value) 是对第 i 个因变量 (dependent variable) 的估计

\hat{b}_1 的置信区间是

$$\hat{Y} \pm (t_c \times s_f)$$

t_c 是 two-tailed t-value 检验值，自由度 (degree of freedom) 是 $n - 2$

$s_f = \text{SEE}^2 [1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{(n-1)s_x^2}]$ 是 standard error of the forecast，一般题目中会给

s_x^2 是 variance of the independent variable

\hat{X}_i 是对第 i 个自变量 (independent variable) 的估计

$\hat{b}_1 = \text{COV}_{XY} / \delta_X^2$ 是模型的坡度，slope coefficient.

\hat{b}_1 的置信区间是

$$\hat{b}_1 \pm (t_c \times s_{\hat{b}_1})$$

t_c 是 two-tailed t-value 检验值，自由度 (degree of freedom) 是 $n - 2$

所以检验 \hat{b}_1 用 $t_{b_1} = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}}$ ，并且拒绝 H_0 如果 $t > |t_{critical}|$

$s_{\hat{b}_1}$ 是 standard error of regression coefficient

$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$ 是模型的交点，intercept term.

1.2 ANOVA Table

首先，我们先看看 ANOVA Table 是什么样子的，再解释里面的各项是什么意思

表 1.1: ANOVA Table

Source of Variation	DoF (k)	Sum of Squares	Mean Sum of Squares
Regression (explained)	1	RSS	$\text{MSR} = \text{RSS}/k = \text{RSS}$
Error (unexplained)	$n - 2$	SSE	$\text{MSE} = \frac{\text{SSE}}{n-2}$
Total	$n - 1$	SST	

1. DoF (Degree of freedom) 自由度

回归模型的自由度是自变量的个数，也就是 k ；误差的自由度是 (观测样本的个数)

- (自变量个数) - 1，也就是 $n - k - 1$

2. RSS (Regression sum of squares) 回归平方和

回归模型能够解释的因变量变化量

$$RSS = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

3. SSE (Sum of squared errors) 残差平方和

回归模型不能够解释的因变量变化量

$$SSE = \sum_{i=1}^n (\hat{Y}_i - Y_i)^2$$

4. SST (Total Sum of squares) 总离差平方和

因变量的总变化量

$$SST = \sum_{i=1}^n (\bar{Y}_i - Y_i)^2$$

5. MSR (Mean regression sum of squares) 平均回归平方和

6. MSE (Mean squared error) 平均残差平方和

其中，SS 代表 sum of squares。MS 代表 Mean squared，也就是对应的 SS 除以对应的自由度。

1.3 Calculating R^2 and SEE

R^2 是 coefficient of determination，表明多少百分比的数据可以被回归模型解释。这个值越大越好。

$$R^2 = \frac{SST - SSE}{SST} = \frac{RSS}{SST}$$

SEE 是 standard deviation of the regression error terms，通过 MSE 开方得到。

$$SEE = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}}$$

1.4 The F-Statistics

F 值表明了该模型中的全部或部分自变量是否适合用来估计母体。In multiple regression, the F-statistic is used to test whether at least one independent variable in a set of independent variables explains a significant portion of the variation of the dependent variable.

$$F = \frac{MSR}{MSE} = \frac{RSS/k}{SSE/(n-k-1)}$$

F critical value 是通过分子和分母的自由度，还有 significance level 决定的。拒绝 H_0 的条件是 $F > F_c$ 。



第二章 Multiple Regression



2.1 Indroduction

多元回归一般的形式是

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i} + \hat{b}_k X_{ki}$$

第三章 The Term Structure and Interest Rate Dynamics

3.1 Spot and Forward Rates

首先，我们先定义一下 spot rate 和 forward rate 分别是什么。

Spot rate 是未来单笔收入的年化市场利率。也可以理解成为零息债券的收益率。

Forward rate 是对于某贷款事先定好的利率。

3.1.1 Spot Rates

我们定义，面值为 1 块钱的零息债券的现价为贴现因子 (discount factor)，记为 P_T 。
同时，这笔投资的产出 (yield to maturity) 记为 S_T

$$P_T = \frac{1}{(1 + S_T)^T}$$

3.1.2 Forward Rates