Advanced Econometrics

- **Reference**
  - 《高级计量经济学》 洪永淼著
  - Econometrics, Bruce E. Hansen
  - Introductory Econometrics, Jeffrey M. Wooldridge
  - 概率论与数理统计教程 茆师松等著
- Made By Jackson, any further question please contact me via Email.

Chapter 2 General Regression Analysis

A. Conditional Probability Distribution

1) Joint probability density function

我们通常用$f_{XY}(x,y)$作为随机变量X,Y的联合概率密度函数

2) Marginal PDF of X

在高维数据中，我们通常用$f_X(x)$表示随机变量X的边缘概率密度函数

3) Conditional PDF of Y given X

在条件概率密度函数中，我们一般使用以下关系式进行后续的结论证明

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} \tag{5}$$

4) Regression Function （Conditional Expectation）

在整个计量经济学的回归领域中，所有的回归函数都可以用$E(Y|X)$来进行表示，并具有以下关系式

$$E\left(Y|X\right)=E\left(Y|X=x\right)=\int_{-\infty}^{+\infty}y\cdot f_{Y|X}\left(y|x\right)dy \tag{6}$$

为了方便理解上式，我们可以通过回顾“期望的定义式”来进行加深理解：

$$\begin{cases} E\left(Y\right)=\int_{D}y\cdot f_{Y}\left(y\right)dy \\ E\left[g\left(Y\right)\right]=\int_{D}g\left(y\right)\cdot f_{Y}\left(y\right)dy \end{cases} \tag{7}$$

为了进一步理解条件期望，我这里设置了“自问自答”环节供自己复习时进行参考：

Q1: 为什么回归函数是对y求积分，而不是对x?

A1: 由于这部分涉及到了测度论类容，所以我主要从直观上进行理解。这里可以从条件概率密度函数的定义上进行理解：$f_{Y|X}(y|x)$是在$X=x$的条件下随机变量Y的函数，所以这里本质上还是求Y的期望，只是多了个条件而已。

Q2: 注意区分，条件期望$E(Y|X)$又是谁的函数?

A2: 从条件期望的积分式子中，我们可以明显看出积分后只剩下随机变量X，所以条件期望是X的函数，在Hansen的教材中，也把这个回归函数记作：$m(X)=E(Y|X)$

5) Conditional Variance

在初等概率论中，我们知道条件期望仍然继承了数学期望的性质，所以我们可以从方差的定义出发，导出条件方差的定义式

$$\begin{cases} Var\left(Y\right)=E[Y-E\left(X\right)]^{2}=\int_{D}\left[Y-E\left(X\right)\right]^{2}\cdot f\left(y\right)dy \\ Var\left(Y|X\right)=E[Y-E\left(Y|X\right)]^{2}=\int_{D}\left[Y-E\left(Y|X\right)\right]^{2}\cdot f_{Y|X}\left(y|x\right)dy \end{cases} \tag{8}$$

6) Conditional α-quantile

根据国内教材的习惯，本文中α分位数统一为上分位数。条件分位数与条件期望、方差类似，都是给定了$X=x$这个条件，然后在进行原本的运算，我们这里记$Q(x,\alpha)$为条件分位数，它的涵义为：在点$Q(x,\alpha)$右方的密度面积大小为α（这里实在不好描述），用数学公式表示为：$P\left[Y \geq Q\left(x,\alpha\right)\right]=\alpha$