

GOLDEN EDUCATION

CFA一级 每日知识点

• 高顿教育CFA学术中心

• 适用于2022年CFA一级考试

01 CFA-L1-每日知识点-数量

03 CFA-L1-每日知识点-财报

05 CFA-L1-每日知识点-权益

07 CFA-L1-每日知识点-衍生

09 CFA-L1-每日知识点-组合

02 CFA-L1-每日知识点-经济

04 CFA-L1-每日知识点-公发

06 CFA-L1-每日知识点-固收

08 CFA-L1-每日知识点-另类

10 CFA-L1-每日知识点-道德





01

CFA-L1

每日知识点-数量



要点解析

EAR

有效年利率(EAR)是指每单位货币在1年内获得的利息收益总额。

一般地, EAR的公式为:
$$EAR = \left(1 + \frac{r}{m}\right)^m - 1$$

连续复利情况下, EAR的公式为:
$$EAR = e^r - 1$$

随着复利频次m的增加, EAR越来越大。



要点解析

PV/ FV 转换

现值 (Present Value, PV) 与终值 (Future Value, FV) 是一组对应的概念。

现值与终值实际上是同一组现金流在不同时间点的价值, 两者可以相互转换, 其基本思想就是复利。

基本公式: $FV = PV \times (1 + r)^n$

离散复利: $FV = PV \times \left(1 + \frac{r}{m}\right)^{mn}$

连续复利: $FV = PV \times e^{rn}$



要点解析

Annuity 年金

年金广义上指等额且定期支付的一系列现金流。

根据每期收付时点与收付方式，年金分为以下三种：

Ordinary annuity 普通年金（后付年金）：是指在每期期末收到等额年金(END)

Annuity due 先付年金：是指在每期期初支付年金(BGN)

Perpetuity 永续年金：是指无限期支付年金，支付年金的时间点是每期期末。永续年金只有现值 $PV=A/r$ ，没有终值。



要点解析

数据类型 1

传统的数据分类方法

1. Numerical data (quantitative data) 数值数据（定量数据）

(1) Continuous data 连续型数据

(2) Discrete data 离散型数据

2. Categorical data (qualitative data) 分类数据（定性数据）

(1) Nominal data 名义数据

(2) Ordinal data 排序数据



要点解析

数据类型 2

实证研究常用的三种类型数据

1. Cross-sectional data 横截面数据

横截面数据是指在某一时点收集的不同对象的数据。它对应同一时点上不同对象所组成的一维数据集。

3. Panel data 面板数据

面板数据是截面数据与时间序列综合起来的一种数据资源。它可以用于分析各样本在时间序列上组成的数据的特征

2. Time-series data 时间序列数据

时间序列数据是指对同一对象在不同时间连续观察所取得的数据。它着眼于研究对象在时间顺序上的变化。



要点解析

数据类型 3

大数据工程使用的两种数据类型：

1. Structured data 结构化数据

高度组织和整齐格式化的数据。

它是可以放入表格和电子表格中的数据类型。

2. Unstructured data (alternative data) 非结构化数据（另类数据）

结构化数据之外的一切数据。

典型例子：

个人活动产生的数据：文字稿件、即时通讯消息、录音、视频；

企业生产活动产生的数据：订单流量、财务报表、市场指数、股票价格；

传感器(sensor)捕捉到的数据：地形、气温、交通状况、监控视频等



要点解析

平均数基本分类

中心趋势可以从均值这个指标进行度量，均值有以下四种：

1. Arithmetic Mean (算术平均)
2. Geometric Mean (几何平均)
3. Harmonic Mean (调和平均)
4. Weighted Mean (加权平均)

算术平均数、几何平均数与调和平均数之间的大小关系如下：

调和平均数 $H \leq$ 几何平均数 $G \leq$ 算术平均数 A



要点解析

分位数

分位数 (Quantile) 是指位于数据集中指定位置的数据。

常用分位数有Quartiles (四分位数) / Quintiles (五分位数) / Deciles (十分位数) / Percentiles (百分位数)

分位数的位置: $L_y = (n + 1)(y\%)$



要点解析

Downside Deviation and Coefficient of Variation

除了绝对离散程度之外比较不同随机变量之间的波动离散程度的两个指标：

1. 半方差 (Semi-deviation)：能更准确地描述数据小于所关注的target值的离散程度。

$$S_{Target} = \sqrt{\frac{\sum_{X_i \leq B}^n (X_i - B)^2}{n-1}}$$

2. 变异系数 (Coefficient of Variation, CV)：衡量单位均值收益率所承担的风险，因此，在对比不同资产时，变异系数越低越好。

$$CV = \frac{s}{\bar{X}}$$



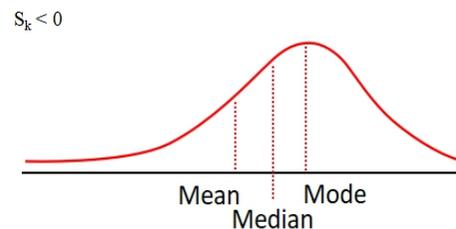
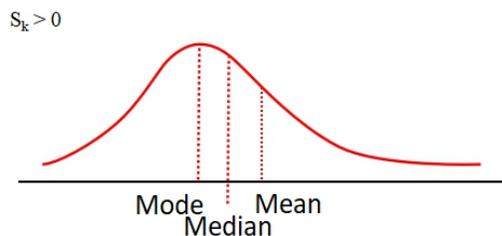
要点解析

Skewness & Kurtosis

偏度与峰度进一步描述投资收益率的分布。

1, 偏度 (Skewness) : 反映数据分布的对称性。

右偏时, 分布的众数 < 中位数 < 均值; 左偏时, 分布的均值 < 中位数 < 众数



2, 峰度 (Kurtosis) : 衡量尾部的厚度。实际运用中, 通常将正态分布的峰度作为基准。正态分布的峰度为3, 若峰度高于3, 则称为高峰 (Leptokurtic) ; 若峰度低于3, 则称为低峰 (Platykurtic)



要点解析

事件

随机变量的部分结果组成的集合称为随机事件，简称为**事件 (Events)**。

根据事件之间的关系，可以分为以下三类：

1. **互斥事件 (Mutually Exclusive Events)**：一组不可能同时发生的事件称为互斥事件。
2. **遍历事件 (Exhaustive Events)**：一组包含随机变量所有可能结果的事件称为遍历事件。
3. **独立事件 (Independent Events)**：如果一个事件的发生不会影响到另一个事件的发生，则称这两个事件独立。



要点解析

贝叶斯公式

事件A是研究问题中所关注的概率。如果得到了新的信息B，可以依据新的信息来更新对事件A概率的估计，即 $P(A|B)$ 。

我们将 $P(A)$ 称为**先验概率 (Prior Probability)**， $P(A|B)$ 称为**后验概率 (Updated Probability)**。

公式中分母 $P(B)$ 概率实际上是用全概率公式计算的。

$$P(A|B) = \frac{P(B|A)}{P(B)} \times P(A)$$



要点解析

Covariance & Correlation

协方差与相关系数衡量不同资产之间收益率的联动关系。

1. **协方差 (Covariance)**：用来度量不同资产之间的收益率联动性，如果两个资产之间收益率是正相关的，协方差为正数；如果资产之间收益率是负相关的，协方差为负数。

$$\text{Cov}(x,y) = E[(x - \bar{x})(y - \bar{y})] = \sum_{i=1}^N P_i (X_i - E(X))(Y_i - E(Y))$$

2. **相关系数 (Correlation)**：剔除了量纲的影响，直接用于比较两组资产之间联动性的高低。相关系数的取值范围在-1到+1之间。当相关系数为1时，称为完全正相关；当相关系数为-1时，称为完全负相关。相关系数绝对值越高，意味着资产之间的线性关系越强。如果相关系数为0时，意味着资产之间不存在线性关系。

$$\rho_{i,j} = \frac{\text{Cov}(R_i, R_j)}{\sigma_i \sigma_j}$$



要点解析

Combination & Permutation

计数的方法有两种，即排列与组合，两者均是有关“从n个元素中任取r个元素”取法的计算。不同之处在于：

排列 (Permutation) 区分取出元素的次序：
$$P_n^r = \frac{n!}{(n-r)!}$$

组合 (Combination) 不区分取出元素的次序：
$$C_n^r = \frac{n!}{(n-r)!r!}$$



要点解析

Bernoulli Distribution & Binomial Distribution

伯努利分布是指随机变量X只有两个可能取值结果1与0。当X=1时，代表“成功”，其对应概率为p；当X=0时，代表“失败”，对应概率为1-p。

二项分布（Binomial Distribution）就是度量进行n次伯努利试验，成功次数为X=x的概率，则P(X=x)的概率为：

$$P(x) = C_n^x p^x (1-p)^{n-x}$$



要点解析

标准正态分布

为了考察任一正态分布 $N(\mu, \sigma^2)$ 有关事件的概率，引入均值为0，方差为1的标准正态分布。

任一**正态分布 $N(\mu, \sigma^2)$** 都可以通过线性变换转换为**标准正态分布**：
$$z = \frac{x - \mu}{\sigma}$$

于是，可以通过标准正态分布表计算任一正态分布 $N(\mu, \sigma^2)$ 相关事件的概率。

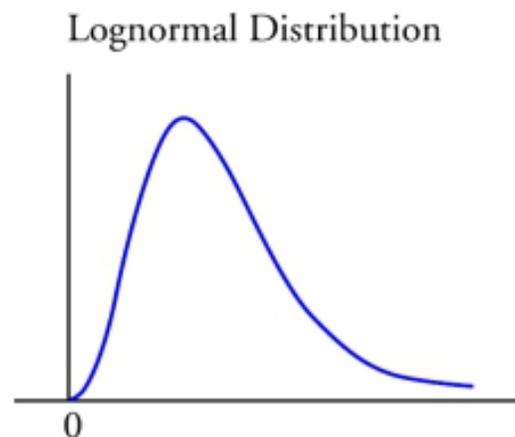


要点解析

对数正态分布

对数正态分布定义为：对于随机变量 X 来说，如果 $\ln(X)$ 服从正态分布，那么 X 服从对数正态分布。有关对数正态分布，需要掌握以下三个性质：

1. 对数正态分布的概率密度函数是**非负**的；
2. 对数正态分布的概率密度函数是**正偏**的；
3. 对数正态分布是用来衡量**资产价格**的。



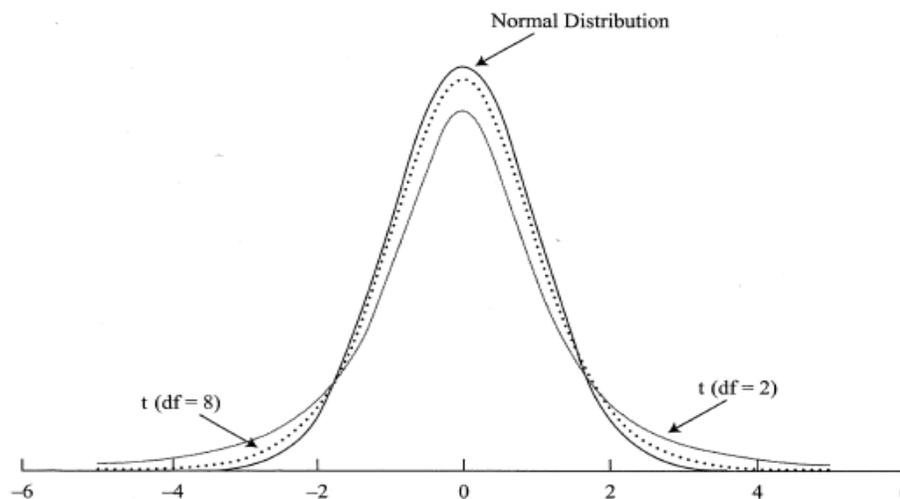
要点解析

Student's t-Distribution

t分布是统计学中非常重要的一类分布，开创了小样本统计推断的时代。

t分布有以下三个特征：

1. t分布的概率密度函数完全由一个参数刻画：自由度（degree of freedom, df）；
2. t分布的概率密度函数也是一条钟形曲线，左右对称，但尾部比正态分布更厚（即偏度=0，峰度>3）
3. 当t分布的自由度增加并趋近于无穷时，t分布的概率密度函数会无限趋近于标准正态分布的概率密度函数。



要点解析

中心极限定理

中心极限定理 (Central Limit Theorem) 对样本均值的概率分布进行描述。我们将中心极限定理归纳为三个条件与三个结论：

条件1：抽样必须是简单随机抽样。

条件2：总体的均值与方差均有限，不为无穷大。

条件3：样本容量超过30。

结论1：样本统计量服从正态分布。

结论2：样本统计量的均值为 μ 。

结论3：样本统计量的方差为 σ^2/n 。样本统计量的标准差即标准误：



要点解析

Confidence Interval Estimate

区间估计指估计未知总体参数的区间范围。置信区间的一般公式为：

点估计量 \pm 置信因子(Reliability factor)*标准误

其中，置信因子的取值取决于总体分布与置信水平 $1-\alpha$ 。

判断采用z统计量还是t统计量：

- (1) 总体方差已知，z分布；
- (2) 总体方差未知，t分布；
- (3) 非正态小样本，不可估计；
- (4) 大样本 ($n \geq 30$)，z,t都可以。

关于标准正态分布的几个置信因子：90%置信区间-置信因子1.65；95%置信区间-置信因子1.96；99%置信区间-置信因子2.58。



要点解析

Hypothesis Testing

假设检验的步骤如下：

1. 建立需检验的假设： $=$ 、 \geq 、 \leq 出现在原假设中。
2. 选择合适的检验统计量，并确定其服从的概率分布。
3. 选择判断假设是否成立的显著性水平。

关键值是指判断是否拒绝原假设的临界值，拒绝域是使得原假设被拒绝的样本观测值所组成的区域。

4. 给出决策准则（Decision Rule），即拒绝域的形式。

p值指拒绝原假设的最小显著性水平。在给定显著性水平 α 的情况下，如果 $p \leq \alpha$ ，则拒绝原假设；如果 $p > \alpha$ 则无法拒绝原假设。

5. 收集数据，并计算检验统计量。
6. 做出判断。
7. 根据判断进行投资决策。



要点解析

Type I Error & Type II Error

在判断假设检验是否成立时有可能犯的误差可归为两类：

- 一是当原假设 H_0 真实成立时，却拒绝了原假设，称为第一类错误（Type I Error），也称为“拒真概率”；
- 二是当原假设 H_0 不成立时，却接受了原假设，称为第二类错误（Type II Error），也称为“受伪概率”。

Decision	True Situations	
	H_0 True	H_0 False
Do not reject H_0	Correct Decision	Type II Error (Probability= β)
Reject H_0	Type I Error (Probability= α)	Correct Decision (Power of test= $1-\beta$)



要点解析

相关系数的假设检验

相关系数的原假设与备择假设： $H_0: \rho=0$ vs $H_a: \rho \neq 0$

相关系数检验选取的统计量服从t分布，自由度为n-2，其公式为：
$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

其中

r为样本相关系数；

n为样本容量；

n-2为自由度。



要点解析

Spearman Rank Correlation

考生需知道相关性的另一种方式——斯皮尔曼秩相关系数，及其检验方法

Spearman秩相关系数公式:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

小样本时：对于秩相关系数的检验需要查秩相关系数表

大样本 (n>30)时：可以代入公式 $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ 进行t检验

当总体明显不服从正态分布，或者根本不知道总体服从什么分布时候用。属于非参数检验
(Nonparametric Test)



要点解析

Parametric Test & Nonparametric Test

知道参数检验和非参数检验的概念、形式

Parametric Test 参数检验

基于对总体分布已知或总体参数分布的假设下进行

Nonparametric Test 非参数检验

除了参数检验之外的检验，当：

- (1) 总体数据没有确定分布
- (2) 数据大小可以代表排序
- (3) 检验的对象无关于参数

非参数检验的例子：

1. 斯皮尔曼秩相关系数检验
2. 列联表行列相关性检验



要点解析

Simple Linear Regression

知道OLS估计方法的四个假设，以及OLS的含义。

Simple linear regression model:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i$$



Equation of regression line:

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$$

其中：

$$\hat{\beta}_1 = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}, \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

4 Assumptions:

线性关系，同方差，误差项 $\varepsilon_i \sim N(0, \sigma_e^2)$ 且 相互独立。

Ordinary least squares(OLS):
指普通最小二乘法，使估计出的截距项 \hat{b}_0 和系数 \hat{b}_1 让模型的SSE最小。

Sum of squared errors(SSE):
观测值和回归直线（估计值）的垂直距离(vertical distances)的平方和。



要点解析

Measures of Fitness

因变量总变化SST的拆解，R²的意义以及计算

Total sum of squares (SST) 总平方和

Measures the total variation in the dependent variable.

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

Regression sum of squares (RSS) 回归/解释平方和

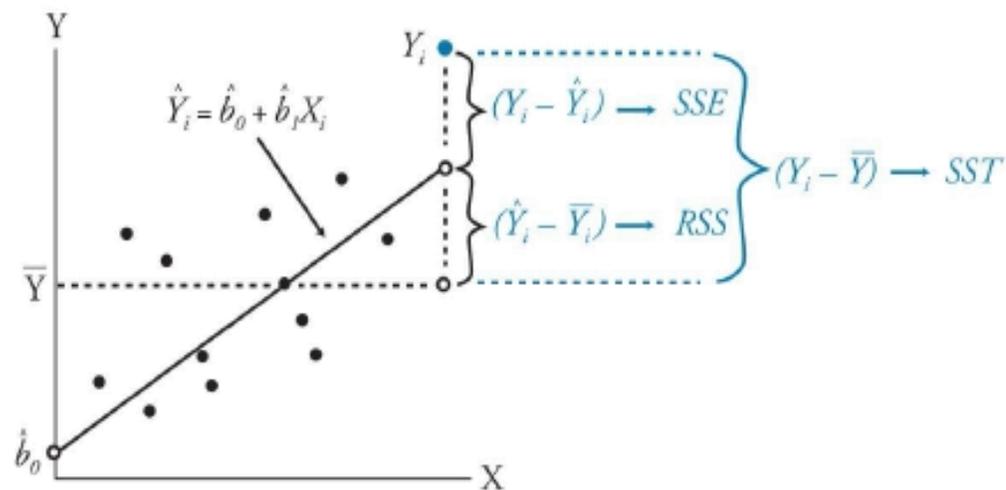
Measures the variation in the dependent variable that is explained by the independent variable.

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

➤ **Sum of squared errors/residuals (SSE) 误差平方和**

- Measures the unexplained variation in the
- dependent variable.

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



$$R^2 = \frac{\text{Explained variation}}{\text{Total variation}} = \frac{RSS}{SST} = \frac{SST - SSE}{SST}$$

The **greater** the R², the **better** the fit.

For **simple** linear regression: **R²=r²=ρ_{x,y}²**



要点解析

ANOVA

利用方差分析判断自变量变动对因变量变动的解释程度。

ANOVA 方差分析	df 自由度	Sum of Squares (SS) 平方和	Mean Sum of Squares (MS) 均方和
Regression (explained)	1	RSS	MSR=RSS/1
Error (unexplained)	n-2	SSE	MSE=SSE/(n-2)
Total	n-1	SST	-

$$F = \frac{MSR}{MSE} = \frac{RSS/1}{SSE/n-2} = t^2, \quad df_{\text{numerator}} = 1; \quad df_{\text{denominator}} = n-2$$

方差分析 第二个作用就是构造F检验统计量。F-test用于检验模型总体有效性。

在简单回归中，模型总体有效等价于 $b_1\text{-hat}$ 不等于零。



要点解析

Predicted & Inference

使用回归方程进行预测和推断。对因变量的预测只需将自变量值代入估计方程即可。

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 X_p$$

\hat{Y} : predicted value of the dependent variable;

X_p : forecasted value of the independent variable.

预测值 \hat{Y} 类似“点估计”，与真实值Y之间一般会存在误差。因此，Y的置信区间为： $\hat{Y} \pm t_{critical} S_f$

s_f : standard error of the prediction.

$$s_f^2 = s^2 \left[1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{(n-1)s_x^2} \right]$$

s^2 : the squared standard error of estimate (=SEE²);

s_x^2 : variance of the independent variable (=Var(X)).



要点解析

Other forms & Limitations

了解简单回归模型的其他形式（非线性模型）和意义，以及回归模型的局限性。

对数-线性模型：

$$\ln Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

线性-对数模型：

$$Y_i = \beta_0 + \beta_1 \ln X_i + \varepsilon_i$$

对数-对数模型：

$$\ln Y_i = \beta_0 + \beta_1 \ln X_i + \varepsilon_i$$

Limitations:

1. Regression relations can change over time (parameter instability). 参数不稳定
2. To investment contexts, public knowledge of regression relationships may negate their future usefulness. 回归结果被公开后就会失效
3. If the regression assumptions are violated, hypothesis tests and predictions based on linear regression will not be valid. 假设条件难以满足

