CFA 一级百题预测

1. **ETHICAL AND PROFESSIONAL STANDARDS**
2. **QUANTITATIVE METHODS**
3. **ECONOMICS**
4. **FINANCIAL REPORTING AND ANALYSIS**
5. **CORPORATE FINANCE**
6. **EQUITY**
7. **FIXED INCOME**
8. **DERIVATIVES**
9. **ALTERNATIVE INVESTMENTS**
10. **PORTFOLIO MANAGEMENT**

1. Quantitative

1.1. Interest Rate

1.1.1. 重要知识点

1.1.1.1. Decompose required rate of return

- Interest rate = real risk free rate + expected inflation rate + risk premium
- Nominal risk free rate = real risk-free rate + expected inflation rate

1.1.2. 基础题

Q-1. Comparing to a company with a high credit rating and low default risk, another company with a low credit rating and high default risk is most likely to provide an interest rate of return with:

A. a lower level.

B. the same level.

C. a higher level.

Q-2. Now, the nominal risk-free rate decreases. Keep the credit risk, liquidity risk and maturity risk constant, if the inflation rate increases, the real risk-free rate will be:

A. Decrease.

B. No change.

C. Increase.

Q-3. The maturity premium can be best described as compensation to investors for the:

A. risk of loss relative to an investment's fair value if the investment needs to be converted to cash quickly.

B. increased sensitivit risk of the market value of debt to a change in market interest rates as maturity is extended.

C. possibility that the borrower will fail to make a promised payment at the contracted time and in the contracted amount.

Q-4. Which of the following risk premiums is most relevant in explaining the difference in yields between 15-year bonds issued by the US Treasury and 15-year bonds issued by a small private issuer?

A. Inflation.

B. Maturity.

C. Liquidity.

1.2. Calculation of HPY, EAR

1.2.1. 重要知识点

1.2.1.1. HPY 和 EAY/EAR 的计算及转化

- $HPR = \frac{FV - PV}{PV}$ 持有期 非年化
- Effective annual rate of return (EAY/EAR) 一年+复利+真实
  - $EAY = (1 + HPY)^{365/t} - 1$
  - $EAY = (1 + \frac{r}{m})^{m} - 1$
  - 性质：r↑，m↑，EAR↑；When m $\to \infty$, $EAR_{max} = e^{r} - 1$
- 单利时用 $\frac{360}{t}$，复利时用 $\frac{365}{t}$

1.2.2. 基础题

Q-5. The weekly closing prices of Mordice Corporation shares are as follows:

Date	Closing Price (€)
1 August	112
8 August	160
15 August	120

The continuously compounded return of Mordice Corporation shares stated rate for the period August 1 to August 15 is closest to:

A. 6.90%.

B. 7.14%.

C. 8.95%.

Q-6. An investment in 5,000 common shares of a company for one year earned a 7% return. The investor received a $1,250 dividend just prior to the sale of the shares at $48 per share. The price that the investor paid for each share one year earlier was closest to:

A. $44.05.

B. $45.09.

C. $45.90.

Q-7. A money manager has $1,000,000 to invest for one year. She has identified two alternative one-year certificates of deposit (CD) shown below:

EAR	Compounding frequency	Annual interest rate （r）
CD1	Quarterly m=4	4.10%
CD2	Continuously	4.00%

Which CD has the *higher* effective annual rate (EAR) and how much interest will it earn?

	Highest EAR	Interest earned
A.	CD1	$41,635
B.	CD2	$40,811
C.	CD2	$40,808

Q-8. The stated (quoted) annual interest rate on an automobile loan is r=10%. The effective annual rate (EAR) of the loan is 10.47%. The frequency of compounding per year for the loan is closest to:

A. quarterly.

B. monthly.

C. semi-annually.

1.3. Time Value of Money

1.3.1. 重要知识点

1.3.1.1. Annuities 年金：FV, PV, required interest, payment

- N = number of periods
- I/Y = interest rate per period
- PMT = amount of each periodic payment
- FV = future value
- PV = present value
- 考察方法：计算——N, I/Y, PMT, FV, PV 中任意给定四个，求另外一个

1.3.1.2. Ordinary annuity 后付年金

- The first cash flow occurs at the end of the first period (t=1)

1.3.1.3. Annuity due 先付年金 (BGN mode)

- The first cash flow occurs immediately (at t=0)

1.3.1.4. Perpetuity

- A perpetuity is a set of level never-ending sequential cash flows, with the first cash flow occurring one period from now.
- 计算： $PV = \frac{A}{r}$

1.3.2. 基础题

Q-9. As winning a lottery, Mikey has three options to get bonus.

Option 1: An ordinary annuity with 20 annual payments of $2,000.

Option 2: An annuity due with 20 annual payments of $2,000.

Option 3: A perpetuity with annual payments of $2,000.

Assuming the annual discount rate is 5 percent, which option is the last one for Mikey to choose?

A. Option 1.

B. Option 2.

C. Option 3.

Q-10. A saver deposits the following amounts in an account paying a stated annual rate of r=4%, compounded semiannually m=2

Year	End of Year Deposits ($)
1	4000
2	8000
3	7000
4	10000

At the end of Year 4, the value of the account is closest to:

A. $30,432

B. $30,447

C. $31,677

Q-11. For planning purposes, an individual wants to be able to spend €80,000 per year, at the end of each year, for an anticipated 20 years in retirement. In order to fund this retirement account, he will make annual deposits of €11,606.56 at the end of each of his working years. What is the minimum number of such deposits he will need to make to fund his desired retirement? Use 6% interest compounded annually for all calculations.

A. 29 payments

B. 30 payments

C. 31 payments

Q-12. A financial contract offers to pay €4,800 per month for 8 years with the first payment made immediately PMT 1 t=0 BGN. Assuming an annual discount rate of 6%, compounded monthly the present value of the contract is closest to:

A. € 84,484.

B. € 365,257.

C. € 367,083.

Q-13. A client invests €40,000 in a four-year certificate of deposit (CD) that annually pays interest of 7%. The annual CD interest payments are automatically reinvested in a separate savings account at a stated annual interest rate of 4% compounded monthly m=12. At maturity, the value of the combined asset is closest to:

A. € 51,200.00.

B. € 51,903.24.

C. € 58,831.19.

1.4. Types of Data & Data Visualization

1.4.1. 重要知识点

1.4.1.1. Types of Data

- **Numerical data (quantitative data)** are values that represent measured or counted quantities as a number. Has continuous data and discrete data.
- **Categorical data (qualitative data)** are values that describe a quality or characteristic of a group of observations and therefore can be used as labels to divide a dataset into groups to summarize and visualize. Has nominal data and Ordinal data.
- **One-dimensional array (simplest format):** suitable for representing a single variable. Has Time-series data and Cross-sectional data.
  - Time-series data is a sequence of returns collected at discrete and equally

spaced intervals of time (such as a historical series of monthly stock returns).

- Cross-sectional data are data on some characteristic of individuals, groups, geographical regions, or companies at a single point in time.

➢ **Two-dimensional rectangular array (data table) (popular forms)** for organizing data for processing by computers or for presenting data visually for consumption by humans. Have panel data.

1.4.1.2. Data Visualization

➢ **Relative frequency** of observations in an interval is the number of observations (the absolute frequency) in the interval divided by the total number of observations.

➢ **Cumulative relative frequency** cumulates (adds up) the relative frequencies as we move from the first interval to the last.

1.4.2. 基础题

Q-14. To perform meaningful mathematical analysis, an analyst must use data that are:

A. discrete.

B. numerical.

C. continuous.

Q-15. Consider the following contingency table from a political opinion poll:

	Supports Johnson	Supports Williams	Total
Supports Smith	42%	14%	56%
Supports Jones	10%	34%	44%
Total	52%	48%	100%

In this table, the value 34% represents:

A. a joint frequency.

B. a marginal frequency.

C. an absolute frequency.

Q-16. Frequency distributions summarize data in:

A. a tabular display.

B. overlapping intervals.

C. a relatively large number of intervals

Q-17. An analyst gathered the following information about the price-earning (P/E) ratios for the common stocks held in a foundation's portfolio:

Interval	P/E range	Absolute Frequency
I	7.00-15.00	12
II	15.00-23.00	24
III	23.00-31.00	11
IV	31.00-39.00	8

The relative frequency and the cumulative relative frequency, respectively, for interval III are closest to:

Relative frequency   Cumulative relative frequency

占比（%）

A. 20.00% 85.45%

B. 22.00% 36.00%

C. 20.00% 36.00%

Q-18. An analyst collects data relating to five commonly used measures of financial leverage and interest coverage for a randomly chosen sample of 500 firms. The data come from those firms' fiscal year 2020 annual reports. These data are best characterized as:

A. cross sectional data.

B. longitudinal data.

C. time-series data.

1.5. Measures of Central Tendency

1.5.1. 重要知识点

1.5.1.1. Median

- Odd numbers: the (n+1)/2 position
- Even numbers: the n/2 position

1.5.1.2. Mode

- Unimodal; bimodal & Trimodal.

1.5.1.3. Measures of mean

- The arithmetic mean: $\overline{X} = \frac{\sum_{i=1}^{N} X_i}{n}$
- The weighted mean: $\overline{X_W} = \sum_{i=1}^{n} w_i X_i = (w_1 X_1 + w_2 X_2 + \cdots + w_n X_n)$

- The geometric mean: $G=\sqrt[N]{X_1X_2X_3...X_N}=(\prod_{i=1}^{N}X_i)^{1/N}$
- The harmonic mean: $\overline{X_H}=\frac{n}{\sum_{i=1}^{n}(1/X_i)}$
- Harmonic mean≤ geometric mean ≤ arithmetic mean

1.1.1.1. Performance measurement with means

- The geometric mean of past annual return is the appropriate measure of past performance.
- The arithmetic mean is the statistically best estimator of the next year's returns.

1.5.2. 基础题

Q-19. XYZ Corp. Annual Stock Returns

2015	2016	2017	2018	2019	2020
23%	6%	-8%	12%	4%	12%

What is the median return for XYZ stock?

A. 8.2%.

B. 9.0%.

C. 12.0%.

Q-20. When analyzing investment returns, which of the following statements is correct?

A. The geometric mean measures an investment's compound rate of growth over multiple periods.

B. The geometric mean will exceed the arithmetic mean for a series with nonzero variance.

C. The arithmetic mean accurately estimates an investment's terminal value over multiple periods.

Q-21. The following information is available for a portfolio:

Asset Class / **Time**	Equities (60%)	Bonds (40%)
First year returns	15%	11%
Second year returns	11%	-5.6%
Third year returns	-13.86%	12%

The geometric mean return on the portfolio is closest to:

A. 3.9834%.

B. 3.5697%.

C. 4.5189%.

Q-22. A manager invests €5,000 annually in a security for 5 years at the prices shown in the following table.

Purchase Price of Security (€)	
Year 1	124.00
Year 2	152.00
Year 3	168.00
Year 4	180.00
Year 5	184.00

The average price paid for the security is closest to:

A. € 161.6000.

B. € 134.1841.

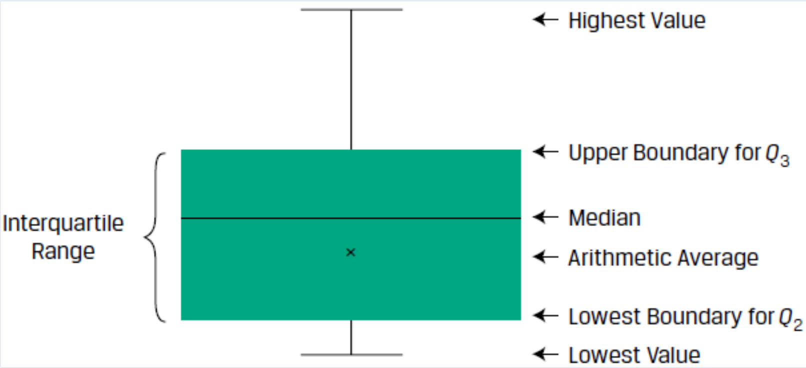
C. € 158.2971.

1.6. Describe, Calculate and Interpret Quartiles, Quintiles, Deciles and Percentiles

1.6.1. 重要知识点

1.6.1.1. Quantiles

- Quartile/Quintile/Deciles/Percentile
  - The third quintile: 60%, or there are three-fifths of the observations fall below that value.
- Calculation formula: $L_y = (n+1)y/100$,
  - Where L_y is the quantile position expressed in percentage.
- Box and whisker plot：



1.6.2. 基础题

Q-23. Which of the following statements is *most* accurate?

A. The first quintile generally exceeds the median.

B. The first quintile generally exceeds the first quartile.

C. The first quintile generally exceeds the first decile.

Q-24. The following exhibit shows the annual MSCI World Index total returns for a 10-year period:

Year 1	15.25%	Year 6	30.79%
Year 2	10.02%	Year 7	12.34%
Year 3	20.65%	Year 8	-5.02%
Year 4	9.57%	Year 9	16.54%
Year 5	-40.33%	Year 10	27.37%

The fourth quintile return for the MSCI World Index is closest to:

A. 20.65%.

B. 26.03%.

C. 27.37%.

1.7. Measure of Dispersion

1.7.1. 重要知识点

1.7.1.1. Absolute dispersion: 不需要和 benchmark 比较

- Range = highest value-lowest value
- $\text{MAD} = \frac{\sum_{i=1}^{n}|X_i - \bar{X}|}{n}$
- $\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$ (for population)

- $s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$ (for sample)

- **MAD 和 variance 掌握计算和比较**：

  - Variance 比 MAD 要好，因为 variance 是连续的，处处可导。MAD 计算的是绝对值，相对比较繁琐。但是 variance 和 MAD 都是表示风险的。注意 MAD ≤ σ.

1.7.1.2. Relative dispersion 相对离散程度，把均值（benchmark）

- **Coefficient of variation 变异系数 (CV)** measures the amount of risk (standard deviation) per unit of mean return.

$$CV = \frac{S_X}{\bar{X}} \times 100\%$$

1.7.1.3. The Sharpe ratio measures the reward, in terms of mean excess return, per unit of risk, as measured by standard deviation of return.

$$Sharpe\ Ratio = \frac{R_P - R_f}{\sigma_P}$$

- Sharpe ratio 只能用于 Rank, 没有实在含义；当小于零时，可能会得到错误的结论。

1.7.2. 基础题

Q-25. The least accurate statement about measures of dispersion for a distribution is that the:

A. range provides no information about the shape of the data distribution.

B. arithmetic average of the deviations around the mean will be equal to one.

C. mean absolute deviation will be either less than or equal to the standard deviation.

Q-26. In 2017, an analyst gathered the following annual return information about a portfolio since its inception on 1 January 2013:

Year	Portfolio return

2013	17.00%
2014	22.20%
2015	25.60%
2016	30.40%
2017	-19.00%

The portfolio's mean absolute deviation and variance of annual returns, respectively, for the five-year period are closest to:

	Mean absolute deviation	Population variance
A.	13.70%	3.12%
B.	13.70%	1.92%
C.	15.24%	1.54%

Q-27. If the risk-free rate is equal to zero and the mean is more than the standard deviation, compared with Sharpe ratio A, the coefficient of variation B is:

A. Same

B. Less

C. Greater

Q-28. An analyst gathered the following information:

Portfolio	Mean Return (%)	Standard Deviation of Returns (%)
1	9	18
2	10	20
3	12	32

If the risk-free rate of return is 5.0 percent, the portfolio that had the best risk-adjusted performance based on the Sharpe ratio is:

A. Portfolio 1.

B. Portfolio 2.

C. Portfolio 3.

Q-29. An analyst gathered the following information:

Portfolio	Mean Return (%)	Sharpe ratio (%)
1	10	34
2	10	37

If the risk-free rate of return is 5.0 percent, which portfolio's coefficient of variation is larger?

A. Portfolio 1

B. Portfolio 2

C. The same

1.8. Skewness and Kurtosis

1.8.1. 重要知识点

1.8.1.1. Skewness 掌握概念：

- 概念：A distribution that is not symmetrical is called skewed.
- 种类：
  - **Positively skewed**–A return distribution with positive skew has frequent small losses and a few extreme gains, (long right tail) (skewness > 0) (mean > median > mode);
  - **Negatively skewed**–A return distribution with negative skew has frequent small gains and a few extreme losses. (long left tail) (skewness < 0) (mean < median < mode)

1.8.1.2. Kurtosis 掌握概念：

- 概念：It deals with whether or not a distribution is more or less "peaked" than a normal distribution.
- 种类：leptokurtic, normal and platykurtic

	Leptokurtic	Normal distribution	Platykurtic
Sample kurtosis	>3	=3	<3
Excess kurtosis	>0	=0	<0

- 理解：

  A leptokurtic return distribution is more peaked and has fatter tails than the normal distribution.
- 可能在考试中会和 skew 合并考核综合知识

1.8.2. 基础题

Q-30. One year ago, an analyst expected his one year investment returns would present a normal distribution. However, the actual distribution of one year investment returns had an excess kurtosis. Based on the given information, which of following item would be mostly undervalued by the analyst one year ago?

A. The mean return of the one year investment returns.

B. The median return of the one year investment returns.

C. The probability that extreme returns occurs.

Q-31. Which of the following is most accurate regarding a distribution of returns that has a mean greater than its median?

A. It is positively skewed.

B. It is a symmetric distribution.

C. It has positive excess kurtosis.

Q-32. When analysing a distribtion, what is the power of the sample skewness and kurtosis respectively?

A. 4, 3.

B. 3, 4.

C. 2, 4.

Q-33. Equity return normal series are best described as, for the most part:

A. platykurtic (less peaked than a normal distribution).

B. leptokurtic (more peaked than a normal distribution).

C. mesokurtic (identical to the normal distribution in peakedness).

1.9. Empirical, Priori or Subjective Probability

1.9.1. 重要知识点

1.9.1.1. Empirical, priori, or subjective probability

- **Objective probability 客观概率**

  - **Empirical probability 经验概率（分析过去/历史，得到将来）**

    e.g. Historically, the Dow Jones Industrial Average has closed higher than the previous close two out of every three trading days. Therefore, the probability of the Dow going up tomorrow is two-thirds, or 66.7%.

  - **Priori probability 先验概率（分析过去/历史，得到过去）**

    e.g. Yesterday, 24 of the 30 DJIA stocks increased in value. Thus, if 1 of 30 stocks is selected at random, there is an 80%(24/30) probability that its value increased yesterday

- **Subjective probability 主观概率**
  - e.g. An investor judges that the probability that the Dow Jones Industrial Average will close higher tomorrow is 90%.

1.9.2. 基础题

Q-34. A fund manager would like to estimate the probability of a daily loss higher than 5% on the fund he manages. He decides to use a method that uses the relative frequency of occurrence based on historical data. The resulting probability is best described as a(n):

A. subjective probability.

B. a priori probability.

C. empirical probability.

1.10. Properties of Probability

1.10.1. 重要知识点

1.10.1.1. Properties of probability 掌握基本公式：

- "×" rule: P(AB)=P(B)×P(A|B)=P(A)×P(B|A);
- "+" rule: P(A or B) = P(A) + P(B) - P(AB)
- Total probability formula:
  - For unconditional probability of event A, where the set of events $\{W_1, W_2, \cdots, W_N\}$ is mutually exclusive and exhaustive.

$$P(A) = P(A|W_1)P(W_1) + P(A|W_2)P(W_2) + \ldots + P(A|W_N)P(W_N)$$

1.10.1.2. Mutually exclusive v.s. independent events

- For mutually exclusive events: P(AB) = 0, P(A or B) = P(A) + P(B)
  For independent events: P(A|B) = P(A), P(B|A) = P(B), P(AB) = P(A) × P(B)
- 注意：不独立未必互斥，互斥一定不独立。

1.10.1.3. Odds for an event

- **Odds for** an event: P(E)/(1-P(E))
- **Odds against** an event: (1-P(E))/P(E)

1.10.2. 基础题

Q-35. An analyst finds that the probability of stock A outperform the market is 65%. What is the odds against of the stock A underperform the market?

A. 0.5385

B. 0.4615

C. 1.8571

Q-36. The probability of price change is as follows:

Price change	0.9
Price increase	0.6

What is the probability that the two situations happen simultaneously?

A. 0.54

B. 0.6

C. 0.9

Q-37. If two events, A and B, are independent and the probability of A does not equal the probability of B [i.e., P(A) ≠ P(B)], then the probability of event A given that event B has occurred [i.e., P(A│B)] is best described as:

A. P(A).

B. P(B│A).

C. P(B).

Q-38. An analyst estimates that a share price has an 80% probability of increasing if economic growth exceeds 3%, a 40% probability of increasing if economic growth is positive but less than 3%, and a 10% probability of increasing if economic growth is negative. If economic growth has a 25% probability of exceeding 3% and a 25% probability of being negative, what is the probability that the share price increases?

A. 22.5%.

B. 42.5%.

C. 62.5%.

Q-39. Given the following information in the table below.

Portfolio 1 gains return in excess of 3%	0.3
Either portfolio 1 or portfolio 2 gains return in excess of 3%	0.8
Portfolio 1 gains return in excess of 3% when portfolio 2 does either	0.25

What is the probability that the return of portfolio 2 will excess 3%?

A. 0.52

B. 0.67

C. 0.75

1.11. Expected Value and Variance

1.11.1. 重要知识点

1.11.1.1. Expected value E(X)

- The expected value of a random variable is the probability-weighted average of the possible outcomes of the random variable.
- $E(X)=\sum P_i \times X_i$

1.11.1.2. Variance Var(X) or σ^2(X)

- The expected value (the probability-weighted average) of squared deviations from the random variable's expected value.
- $\sigma^2=\sum_{i=1}^{N} P_i(X_i-E(X))^2$

1.11.2. 基础题

Q-40. An analyst gathered the following information: the probability of economy prosperity is 80%, the probability of economy recession is 20%. For a company, when the economy is prosperity, there is 85% of probability that its EPS is $9.0 and 15% of probability that the EPS is $3.0. However, when the economy is recession, there is 10% of probability that the EPS is $9.0 and 90% of probability that the EPS is $3.0. What is the variance of this company's EPS, when the economy is prosperity?

A. 6.54

B. 4.59

C. 3.24

1.12. Correlation and Covariance

1.12.1. 重要知识点

1.12.1.1. Covariance:

- Covariance is a measure of the co-movement between random variables.
  - X 与 Y 同向变化，covariance >0.

- X 与 Y 反向变化，covariance < 0.
- $Covariance \in (-\infty, +\infty)$.

➢ Covariance ranges from negative infinity to positive infinity

- $Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$

➢ The covariance of a random variable with itself is its own variance.

- $Cov(X,X) = E[(X - E(X))(X - E(X))] = \sigma^2(X)$

1.12.1.2. Correlation:

➢ Correlation measures the co-movement (linear association) between two random variables.

➢ $$\rho_{XY} = \frac{\text{COV(X,Y)}}{\sqrt{\text{Var(X)Var(Y)}}} = \frac{\text{COV(X,Y)}}{\sigma_X \sigma_Y}$$

➢ Correlation is a number between −1 and +1.

➢ 理解：$\rho_{XY} \in [-1,1]$

- If $\rho_{x,y} = 0$, a correlation of 0 (uncorrelated variables) indicates an **absence** of any linear (straight-line) relationship between the variables.
- Increasingly positive correlation indicates an increasingly **strong** positive linear relationship (up to 1, which indicates a perfect linear relationship).
- Increasingly negative correlation indicates an increasingly **strong** negative (inverse) linear relationship (down to -1, which indicates a perfect inverse linear relationship).

1.12.1.3. Expected return, variance and standard deviation of a portfolio

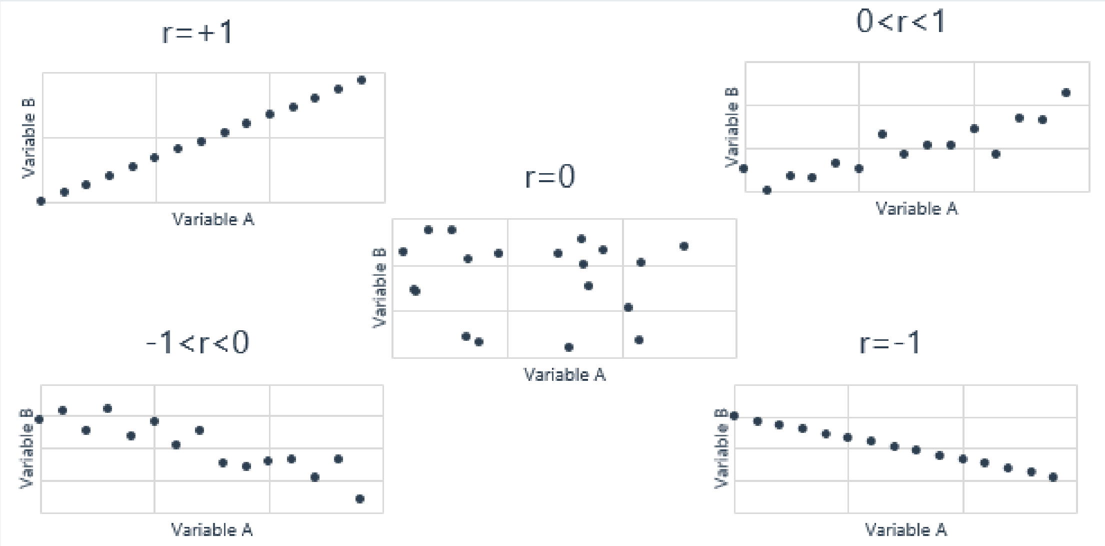
➢ $$E(r_p) = \sum_{i=1}^{n} w_i E(R_i)$$

➢ $$\sigma^2{}_p = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \operatorname{cov}(R_i, R_j)$$

1.12.1.4. Scatter plot & limitation of correlation

➢ A **scatter plots** is a graph that shows the relationship between the observations for two data series in two dimensions.

➢ **Scatter plots charts**



- Three limitations of correlation analysis.
  - **Nonlinear relationships**: Two variables can have a strong nonlinear relation and still have a very low correlation.
  - **Outliers**: Outliers are small numbers of observations at either extreme (small or large) of a sample.
  - **Spurious correlation**: Correlations can be spurious in the sense of misleadingly pointing towards associations between variables.

1.12.2. 基础题

Q-41. Having two high-risk investment products, an investment portfolio earns a risk-free rate of return. The value of correlation between these two high-risk investment products will most likely be:

A. -1

B. 0

C. +1

Q-42. The joint probability of returns, for securities A and B, are as follows:

Joint Probability Function of Security A and Security B Returns. (Entries are joint probabilities)

	Return on security B =32%	Return on security B=24%
Return on security A= 24%	0.70	0
Return on security A= 18%	0	0.30

The covariance of the returns between securities A and B is closest to:

A. 0.0005.

B. 0.0010.

C. 0.0032.

Solution: B.

Expected return on security A = 0.7 x 24% + 0.3 x 18% = 22.2%

Expected return on security B = 0.7 x 32% + 0.3 x 24% = 29.6%

Cov (R_A, R_B) = 0.7 x [(24% - 22.2%) x (32% - 29.6%)] + 0.3 x [(18% - 22.2%) x (24% - 29.6%)]

= 0.001008.

Q-43. The correlation between two variables is +0.3. The most appropriate way to interpret this value is to say:

A. a scatter plot of the two variables is likely to show a strong linear relationship.

B. when one variable is above its mean, the other variable tends to be above its mean as well.

C. a change in one of the variables usually causes the other variable to change in the same direction.

Q-44. An individual want to invest $300,000 in the following investment products:

Investment products	Expected Return	Weights	Standard deviation	Correlation
Stock	6%	80%	25%	0.2
Fund	8%	20%	30%	

What will be the rate of return and the standard deviation on the expected portfolio?

A. 6.4% and 4.84%.

B. 6.4% and 22%.

C. 6.2% and 4.84%.

Q-45. An analyst develops the following covariance matrix of returns:

	Hedge Fund	**Market Index**
Hedge fund	225	90
Market index	90	64

The correlation of returns between the hedge fund and the market index is closest to:

A. 0.005.

B. 0.75

C. 0.00625.

1.13. Bayes’ Formula

1.13.1. 重要知识点

1.13.1.1. Bayes’ formula 掌握计算：

➢ **Updated probability:** Given a set of prior probabilities for an event of interest, if you receive new information, the rule for updating your probability of the event is Updated probability of event given the new information = (probability of new information given event/ unconditional probability of new information) ×prior probability of event.

- $P(AB)=P(B|A)\times P(A)=P(A|B)\times P(B)$

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

➢ **Posterior probability（后验概率）**

- $P(A|B)=\frac{P(B|A)P(A)}{P(B)}=\frac{P(B|A)P(A)}{P(B|W_1)P(W_1)+P(B|W_2)P(W_2)}$
- $P(B)=P(B|W_1)\times P(W_1)+P(B|W_2)\times P(W_2)+\ldots+P(B|W_n)\times P(W_n)$

1.13.2. 基础题

Q-46. With Bayes' formula, it is possible to update the probability for an event given some new information. Which of the following most accurately represents Bayes' formula?

A. $\mathrm{P}(Event\ A|Information\ B)=\frac{P(Information|Event)}{P(Information)}\mathrm{P(Event\ A)}$

B. $\mathrm{P}(Event|Information)=\frac{P(Information)}{P(Information|Event)}\mathrm{P(Event)}$

C. $\mathrm{P}(Event|Information)=\frac{P(Information|Event)}{P(Event)}\mathrm{P(Information)}$

Q-47. An analyst believes Davies Company has a 40% probability of earning more than $2 per share. She estimates that the probability that Davies Company's credit rating will be upgraded is 70% if its earnings per share are greater than $2 and 20% if its earnings per share are $2 or less. Given the information that Davies Company's credit rating has been upgraded, what is the updated probability that its earnings per share are greater than $2?

A. 50%.

B. 60%.

C. 70%.

Q-48. As from the record of CFA Institute and GARP, the pass-through rate of CFA level I exam is 30%, and the pass-through rate of FRM exam is 40%. And as an investigation led by the two institutes, among the people who has passed the FRM, the pass-through rate of CFA level 1 exams is 50%. So what is the pass-through rateof FRM members who has also passed the CFA level 1 exam before?

A. 48%

B. 60%

C. 67%

1.14. Principals of Counting

1.14.1. 重要知识点

1.14.1.1. Principals of counting

- Multiplication rule: $n_1 \times n_2 \times \ldots\ldots \times n_k$
- Factorial: n!
- Labeling (or Multinomial): $\dfrac{n!}{n_1! \times n_2! \times \ldots \times n_k!}$
- Combination: ${}_nC_r = \binom{n}{r} = \dfrac{n!}{(n-r)! \times r!}$
- Permutation: ${}_nP_r = \dfrac{n!}{(n-r)!}$

1.14.2. 基础题

Q-49. Given a portfolio of six stocks, how many unique covariance terms, excluding variances, are required to calculate the portfolio return variance?

A. 10.

B. 15.

C. 25.

Q-50. A firm decided to sell three stocks out of five in order. How many different options are possible?

A. 10.

B. 60.

C. 120.

1.15. Discrete and Continuous Random Variables

1.15.1. 重要知识点

1.15.1.1. Discrete and continuous random variables

- Discrete random variables: take on at most a **countable** number of possible outcomes **but do not necessarily to be limited**.
- Continuous random variables: cannot describe the possible outcomes of a continuous random variable Z with a list z_1, z_2, ... because the outcome $(z_1 + z_2)/2$, not in the list, would always be possible.
  - P(x) = 0 even though x can happen.
  - $P(x_1<X<x_2)$
- Probability function: p(x)=P(X=x)
  - For discrete random variables
  - $0 \leq p(x) \leq 1$
  - $\Sigma p(x)=1$
- Probability density function (p.d.f): f(x)
  - For continuous random variable commonly
- Cumulative probability function (c.p.f): F(x)
  - F(x)=P(X<=x)

1.15.2. 基础题

Q-51. Which of following is a discrete random variable?

A. Quoted prices of a stock.

B. A random variable from a normal distribution.

C. Probabilities of playing a dice.

Q-52. The value of the cumulative distribution function F(x), where x is a particular outcome, for a discrete uniform distribution:

A. sums to 1.

B. lies between 0 and 1.

C. decreases increases as x increases.

Q-53. For a contiuous random vaiable X, the probability that X exactly equals to 0 is:

A. 0.

B. 1/n.

C. 1/x.

1.16. Discrete Random Distribution

1.16.1. 重要知识点

1.16.1.1. Discrete uniform random variable would be a known, finite number of outcomes equally likely to happen. Every one of n outcomes has equal probability 1/n.

1.16.1.2. Bernoulli random variable: p(1) = P(X = 1) = p, p(0) = P(X = 0) = 1 - p

1.16.1.3. Binomial random variable X is defined as the number of success in n Bernoulli trials.

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

1.16.1.4. Expectations and variances

	Expectation	Variance
Bernoulli random variable (Y)	P	p(1-p)
Binomial random variable (X)	np	np(1-p)

1.16.2. 基础题

Q-54. For a binomial random variable with five trials, and a probability of success on each trial of 0.50, the distribution will be:

A. skewed.

B. uniform.

C. symmetric.

Q-55. Which of the following events can be represented as a Bernoulli trial ?

A. The flip of a coin.

B. The closing price of a stock.

C. The picking of a random integer between 1 and 10.

Q-56. The following table shows the discrete uniform probability distribution of gross profits from the purchase of an option.

Profit	Cumulative Distribution Function F(X)	Profit	Cumulative Distribution Function
$0	0.1	$2.5	0.6
$0.5	0.2	$3.0	0.7
$1.0	0.3	$3.5	0.8
$1.5	0.4	$4.0	0.9
$2.0	0.5	$4.5	1.0

The probability that X will take on a value of 1 to 3 is closest to:

A. 0.2.

B. 0.5.

C. 0.6.

Q-57. If the probability that a portfolio underperforms its benchmark in any quarter is 0.40, the probability that the portfolio outperforms its benchmark (Y=1) in two or fewer quarters over the course of a year is closest to:

A. 34.56%.

B. 49.92%.

C. 52.48%.

1.17. Continuous Uniform Distribution

1.17.1. 重要知识点

1.17.1.1. Definition

- All intervals of the same length on the Continuous Uniform Distribution's support are equally probable.

1.17.1.2. Properties

- $P(X < a \text{ or } X > b) = 0$
- For all $a \le x_1 < x_2 \le b$, $P(x_1 \le X \le x_2) = (x_2 - x_1)/(b - a)$

1.17.2. 基础题

Q-58. An energy analyst forecasts that the price per barrel of crude oil five years from now will range between USD$150 and USD$210. Assuming a continuous uniform distribution, the probability that the price will be less than USD$160 five years from now is closest to:

A. 5.8%.

B. 16.7%.

C. 43.4%.

1.18. Normal Distribution

1.18.1. 重要知识点

1.18.1.1. Properties

- $X \sim N(\mu, \sigma^2)$
- Symmetrical distribution: skewness=0, kurtosis=3
- A linear combination of two or more normal random variables is also normally distributed.

- As the values of x gets farther from the mean, the probability density get smaller and smaller but are always positive.

1.18.1.2. Confidence intervals

- 68% confidence interval is [μ - σ, μ + σ]
- 90% confidence interval is [μ – 1.65σ, μ + 1.65σ]
- 95% confidence interval is [μ – 1.96σ, μ + 1.96σ]
- 99% confidence interval is [μ – 2.58σ, μ + 2.58σ]

1.18.1.3. Standardization

- If $X \sim N(\mu,\sigma^2)$, then $Z=(X-\mu)/\sigma \sim N(0,1)$

1.18.1.4. Cumulative probabilities for a standard normal distribution

- F(-z)=1-F(z)
- P(Z>z) = 1 –F(z)

1.18.2. 基础题

Q-59. The total number of parameters that fully characterizes a normal distribution is:

A. 3.

B. 2.

C. 1.

Q-60. The return of a portfolio follows a normal distribution, with its mean return of 13% and its standard deviation of 5%. Given the following z-table, the probability that its return falls between 7% and 19% is *closest* to:

Cumulative Probabilities F(z) for a Standard Normal Distribution

P(Z ≤ z) = N(z) for z ≥ 0

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.20	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.60	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706

A. 83.84%.

B. 76.98%.

C. 93.32%.

Q-61. For a standard normal distribution, what is the probability that a random variable lies within 1 to 2 $P(1< x <2)$?

A. 13.5%

B. 27%

C. 15.5%

1.19. Safety First Ratio

1.19.1. 重要知识点

1.19.1.1. SFR 掌握计算及理解：

- $SFR = \left[E(R_p) - R_L\right] / \sigma_p$: the bigger, the better.
- Shortfall risk: R_L= threshold level return, minimum return required.
- Roy's safety-first criterion states that the optimal portfolio minimizes the probability that portfolio return, R_P, falls below the threshold level, R_L. In symbols, the investor's objective is to choose a portfolio that minimizes $P(R_P<R_L)$.

1.19.1.2. SFR 与 Sharpe ratio 的区别

- $SFR = \left[E(R_p) - R_L\right] / \sigma_p$
- $Sharpe\ ratio = \left[E(R_P) - R_F\right] / \sigma_P$
- Sharpe ratio will be a special case of SFR if $r_L = r_F$

1.19.2. 基础题

Q-62. On 1 January 2014, the value of an investor's portfolio is $90,000. The investor plans to donate $7,000 to charity organization and pay $3,000 to his insurance account on 31 December 2014, but meanwhile he does not want the year-end portfolio value to be below $90,000. If the expected return on the existing portfolio is 14% with a variance of 0.0225, the safety-first ratio that would be used to evaluate the portfolio based on Roy's criterion is closest to:

A. 0.193.

B. 0.465.

C. 0.415.

1.20. Lognormal Distribution

1.20.1. 重要知识点

1.20.1.1. Lognormal distribution

- Definition: If lnX is normal, then X is lognormal, which is used to describe the price of asset.
- Bounded from below by zero: the values of random variables that follow lognormal distribution are always be positive, so it is useful for modeling asset prices. $X \in (0, \infty)$
- Right skewed.
- Stock price follows lognormal distributionl, while the rate of return follows normal distribution.

1.20.2. 基础题

Q-63. In contrast to normal distributions, lognormal distributions:

A. are skewed to the left.

B. have outcomes that cannot be negative.

C. are more suitable for describing asset returns than asset prices.

1.21. The Chi-Square (χ^2) Distribution, Student's t-distribution & The F-Distribution

1.21.1. 重要知识点

1.21.1.1. The Chi-Square (χ^2) Distribution

- Definition: The distribution of the sum of the squares of k independent standard normally distributed random variables.

$$\mathrm{U} = \sum \boldsymbol{Z}_i^2 = \boldsymbol{Z}_1^2 + \boldsymbol{Z}_2^2 + \cdots + \boldsymbol{Z}_k^2 \sim \boldsymbol{\chi}_{(k)}^2$$

- Bounded from below by zero: Developed for testing hypotheses of positive parameters.
- As the degrees of freedom increase, the shape of the density function becomes more similar to a bell curve.

1.21.1.2. Student's t-distribution

- Definition: If Z is a standard normal variable and U is a chi-square variable with k degrees of freedom, then the random variable X follows a t-distribution with k degrees of freedom.

$$t = \frac{\boldsymbol{Z}}{\sqrt{U/k}}$$

- Symmetrical
- Degrees of freedom (df): n-1
- Less peaked than a normal distribution ("fatter tails")
- As the degrees of freedom increase, the Student's t-distribution approaches the standard normal distribution.

1.21.1.3. The F-Distribution

- Definition: If U1 and U2 are two independent Chi-Squared distributions with k1 and k2 degrees of freedom, respectively, then X:

$$\mathrm{F} = \frac{U_1/k_1}{U_2/k_2} \sim F(k_1, k_2)$$

- Asymmetrical distributions bounded from below by 0.
- As both the numerator (k1) and the denominator (k2) degrees of freedom increase, the shape of the density function will also become more bell curve–like.

1.21.2. 基础题

Q-64. Which of the following is least likely a property of Student's t-distribution?

A. As the degrees of freedom get larger, the variance approaches zero.

B. It is defined by a single parameter, the degrees of freedom, which is equal to n-1.

C. It has more probability in the tails fat tail and less at the peak than a standard normal distribution.

Q-65. Which of the following statements about the F-distribution and chi-square distribution is least accurate? Both distributions:

A. are typically asymmetrical.

B. are bounded from below by zero.

C. have means that are less than their standard deviations.

Q-66. Compared with normal distribution, which of the following statements about t-distribution is the most accurate?

A. It has no difference with normal distribution.

B. Its tails are fatter than the tails of normal distribution.

C. It has less moreprobability in the tails than the normal distribution.

Q-67. An analyst stated that as degrees of freedom increase, a t-distribution will become more peaked and the tails of the t-distribution will become less fat . Is the analyst's statement correct with respect to the t-distribution:

	Becoming more peaked?	Tails becoming less fat?
A.	No	Yes
B.	Yes	No
C.	Yes	Yes

1.22. Monte Carlo

1.22.1. 重要知识点

1.22.1.1. Lognormal distribution

- **Monte Carlo** simulation is to generate a large number of random samples from specified probability distribution(s) to represent the operation of risk in the system. It is used in planning, in financial risk management, and in valuing complex securities;
- Limitations:
  - The operating of Monte Carlo simulation is very complex and we must assume a parameter distribution in advance.
  - Monte Carlo simulation provides only statistical estimates, not exact results.

1.22.2. 基础题

Q-68. A Monte Carlo simulation can be used to:

A. directly provide precise valuations of call options.

B. simulate a process from historical records of returns.

C. test the sensitivity of a model to changes in assumptions.

1.23. Central Limit Theorem

1.23.1. 重要知识点

1.23.1.1. Central limit theorem

- Definition: The sampling distribution of the sample mean approaches a normal distribution as the sample size becomes large (>= 30);
- The mean of sample mean distribution=μ; The variance of sample mean distribution = σ^2/n.

1.23.1.2. Standard error of the sample mean

- Known population variance: $\sigma_{\bar{x}} = \sigma / \sqrt{n}$
- Unknown population variance: $s_{\bar{x}} = s / \sqrt{n}$

1.23.2. 基础题

Q-69. Which of the following items best describe the standard deviation of a sample statistic?

A. Sampling error

B. Standard error of the sample statistic

C. Standard deviation of population

Q-70. A population has a non-normal distribution with mean μ and variance σ^2**.** The sampling distribution of the sample mean computed from samples of large size from that population will have:

A. the same distribution as the population distribution.

B. its mean approximately equal to the population mean.

C. its variance approximately equal to the population variance/n.

Q-71. A research analyst makes two statements about repeated random sampling:

Statement 1: When repeatedly drawing large samples from datasets, the sample means are approximately normally distributed.

Statement 2: The underlying population from which samples are drawn must be normally distributed in order for the sample mean to be normally distributed.

Which of the following best describes the analyst's statements?

A. Statement 1 is false; Statement 2 is true.

B. Both statements are true.

C. Statement 1 is true; Statement 2 is false.

Q-72. The following sample of 10 items is selected from a normally distributed population. The population variance is unknown.

11	21	-7	3	-8	6	0	-7	4	22

The standard error of the sample mean is closest to:

A. 10.89.

B. 3.44.

C. 3.70.

1.24. Sampling and Estimation

1.24.1. 重要知识点

1.24.1.1. Concept of sampling and estimation

- **Methods of sampling:** Probability sampling and Non-probability sampling.
  - **Probability sampling:** sample random sampling, stratified random sampling , systematic sampling and cluster Sampling.
  - **Non-probability sampling:** convenience Sampling and Judgmental sampling.
  - **Definition of Sampling Distribution of a Statistic:** The sampling distribution of a statistic is the distribution of all the distinct possible values that the statistic can assume when computed from samples of the same size randomly drawn from the same population.
- Sample Statistic itself is a random variable, thus following specific distribution.
- **Sampling error:** sampling error of mean=sample mean-population mean.

1.24.1.2. The desirable properties of an estimator

- **Unbiasedness:** the expected value of the estimator equals the population parameter.
- **Efficiency:** An unbiased estimator is efficient if no other unbiased estimator of the same parameter has a sampling distribution with smaller variance.
- **Consistency:** A consistent estimator is one for which the probability of estimates close to the value of the population parameter increases as sample size increases (the standard deviation of the parameter estimate decreases as the sample size increases).
  - As the sample size increases, the standard error of the sample mean falls.

1.24.1.3. Point estimation: the statistic, computed from sample information, which is used to estimate the population parameter.

1.24.1.4. Interval estimation:

- Level of significance (alpha)
- Degree of Confidence (1-alpha)
- Confidence Interval = [Point Estimate +/- (reliability factor) × Standard error]

1.24.1.5. Resampling: repeatedly draws samples from the original observed data sample for the statistical inference of population parameters.

- **Jackknife vs. Bootstrap**
  - **Jackknife produces similar results for every run.**
    - **Jackknife usually requires n repetitions. (n=sample size).**
  - **Bootstrap usually gives different results because bootstrap resamples are**

randomly drawn.

- ◆ **Bootstrap needs to determine how many repetitions are appropriate.**

1.24.1.6. Biases in sampling

- **Data snooping bias** comes from finding models by repeatedly searching through databases for patterns.
- **Sample selection bias** occurs when data availability leads to certain assets being excluded from the analysis.
  - **Survivorship bias** occurs if companies are excluded from the analysis because of having gone out of business or poor performance.
  - **Self-selection bias** reflects the ability of entities to decide whether or not they wish to report their attributes or results and be included in databases or samples.
  - **Implicit selection bias** may exist because of a threshold enabling self-selection.
  - **Backfill bias** occurs where past data, not reported or used before, is backfilled into an existing database.
- **Look-ahead bias** exists if the model uses data not available to market participants at the time when the market participants act in the model.
- **Time-period bias** is present if the time period used makes the results time-period specific or if the time period used includes a point of structural change.

1.24.2. 基础题

Q-73. Researchers found all data sampled from a population concentrating in tails of the sampling distribution. Which of the following sampling method is most likely used?

A. Stratified random sampling

B. Systematic sampling

C. Simple random sampling

Q-74. An important difference between two-stage cluster sampling and stratified random sampling is that compared to stratified random sampling, two-stage cluster sampling:

A. uses each member of all sub-group (strata).

B. takes random samples from all sub-groups (strata).

C. will not preserve differences in a characteristic across sub-groups.

Q-75. A sample of 64 observations has a mean of 8. The standard deviation of the sample is 15. Which of the following is the best estimate of the 95% confidence interval for this sample?

A. 4.325 to 11.675.

B. 4.906 to 11.094.

C. 3.031 to 12.969

Q-76. All else held constant, the width of a confidence interval for a population mean is most likely to be smaller if the sample size is:

A. larger and the degree of confidence is lower.

B. larger and the degree of confidence is higher.

C. smaller and the degree of confidence is lower.

Q-77. If an estimator is consistent, an increase in sample size will increase the: (2020 原版书)

A. accuracy of estimates.

B. efficiency of the estimator.

C. unbiasedness of the estimator.

Q-78. Which of the following techniques to improve the accuracy of confidence intervals on a statistic is most computationally demanding?

A. The jackknife.

B. Systematic resampling.

C. Bootstrapping.

Q-79. A report on long-term stock returns focused exclusively on all currently publicly traded firms in an industry is most likely susceptible to:

A. look-ahead bias.

B. survivorship bias.

C. intergenerational data mining.

1.25. Hypothesis Testing

1.25.1. 重要知识点

1.25.1.1. Steps of hypothesis testing

- Step 1: State null and alternative hypotheses
- Step 2: Identify the appropriate test statistic
- Step 3: Specify a level of significance

- Step 4: State a decision rule
- Step 5: Collect data and calculate the test statistic
- Step 6: Draw a conclusion

1.25.1.2. Hypothesis testing:

- $T\text{-}Statistic = \dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$; $T\text{-}Statistic = \dfrac{\overline{X}-\mu_0}{s/\sqrt{n}}$
- Test Statistic follows Normal, T, Chi Square or F distributions.
- Test Statistic has formula. Calculate it with the sample data.
- This is the general formula but only for Z and T distribution.

1.25.1.3. Relation between Confidence Intervals and Hypothesis Tests

- Confidence Interval = point estimate ± (critical value)×(standard error)
  - Center of Interval = point estimate (sample statistic)
  - Length of Interval = 2 × (critical value) × (standard error)

1.25.1.4. t-test 和 z-test 的不同应用：

Sampling from:	Statistic for small sample size (n<30)	Statistic for large sample size (n≥30)
Normal distribution with known variance	z-Statistic	z-Statistic
Normal distribution with unknown variance	t-Statistic	t-Statistic/z
Nonnormal distribution with known variance	not available	z-Statistic
Nonnormal distribution with unknown variance	not available	t-Statistic/z

1.25.1.5. Z 分布、T 分布、卡方分布、F 分布

Test type	Assumptions	H_0	Test-statistic	Critical value
Mean hypothesis testing	Normally distributed population, known population variance	$\mu=\mu_0$	$z=\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$	N(0,1)
	Normally distributed	$\mu=\mu_0$	$t=\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}$	t(n-1)

	population, unknown population variance			
	Independent populations, unknown population variances assumed equal	$\mu_1 - \mu_2 = 0$	$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{s_p^2/n_1 + s_p^2/n_2}}$ $Where\ s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$t(n_1 + n_2 - 2)$
	Independent populations, unknown population variances not assumed equal	$\mu_1 - \mu_2 = 0$	$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$	*t**
	Samples not independent, Paired comparisons test for example, two returns of stocks in the market, the return of gas and that of oil	$\mu_d = 0$	$t = \frac{\bar{d}}{s_{\bar{d}}}$	t(n-1)
		For example, the same population after and before an event, or comparison with the two different styles, which are affected by the same macro factors.		
Variance hypothesis testing	Normally distributed population	$\sigma^2 = \sigma_0^2$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2(n-1)$
	Two independent normally distributed population	$\sigma_1^2 = \sigma_2^2$	$F = \frac{s_1^2}{s_2^2}$	$F(n_1-1, n_2-1)$

1.25.1.6. Test Correlation

Test type	Assumptions	H_0	Test-statistic	Critical value

Correlation	Both of the variables are normally distributed	ρ = 0	$t = \frac{r-0}{\sqrt{\frac{1-r^2}{n-2}}}$	t(n-2)

1.25.2. 基础题

Q-80. Which of the following statements regarding a one-tailed hypothesis test is correct?

A. The rejection region increases in size as the level of significance alpha becomes smaller.

B. A one-tailed test more strongly reflects the beliefs of the researcher than a two-tailed test.

C. The absolute value of the rejection point is larger than that of a two-tailed test at the same level of significance.

Q-81. A sample has less than 30 data selecting from a normal distributed population with known variance. If an analyst wants to test the sample mean, which of the following distribution should be used?

A. t-student distribution

B. Z distribution

C. F distribution

Q-82. Investment analysts often use earnings per share (EPS) forecasts. One test of forecasting quality is the zero-mean test, which states that optimal forecasts should have a mean forecasting error of 0. (Forecasting error = Predicted value of variable − Actual value of variable.)

Performance in Forecasting Quarterly Earnings per Share		
Number of Forecasts	Mean Forecast Error (Predicted – Actual)	Standard Deviations of Forecast Errors
100	0.06	0.20

To test whether the mean forecasting error is 0, the t-statistic calculated is most likely:

A. 3.015.

B. 3.000.

C. 0.060.

Q-83. An analyst conducts a two-tailed test to determine if earnings estimates are significantly different from reported earnings. The sample size was over 90. The computed Z-statistic is 1.30 . Using a 5 percent significant level, which of the following statements is TRUE?

A. Both the null and the alternative are significant.

B. You cannot determine what to do with the information given.

C. Fail to reject the null hypothesis and conclude that the earnings estimates are not significantly different from reported earnings.

Q-84. Which of the following statements of null and alternative hypotheses requires a two-tailed test?

A. H_0: $\theta = \theta_0$ versus H_a: $\theta \neq \theta_0$

B. H_0: $\theta \leq \theta_0$ versus H_a: $\theta > \theta_0$

C. H_0: $\theta \geq \theta_0$ versus H_a: $\theta < \theta_0$

Q-85. Consider a two-tailed test of the hypothesis that the population mean is zero. The sample has 50 observations. The population is normally distributed with a known variance.

t Distribution			
Degree of freedom	p=0.10	p=0.05	p=0.025
49	1.299	1.677	2.010
50	1.299	1.676	2.009
z-Distribution	α=0.10	α=0.05	α=0.025
	1.645	1.960	2.330

At a 0.05 significance level, the rejection points are most likely at:

A. -2.010 and 2.010.

B. -2.009 and 2.009.

C. -1.960 and 1.960.

Q-86. A small-cap growth fund's monthly returns for the past 36 months have been consistently outperforming its benchmark. An analyst is determining whether the standard deviation of monthly returns is greater than 5%. Which of the following best describes the hypothesis to be tested?

A. $H_0: \sigma^2 \leq 0.25\%$

B. $H_a: \sigma^2 > 5\%$

C. $H_0: \sigma^2 \geq 0.25\%$

Q-87. Using a two-tailed test of the hypothesis that the population mean is zero, the calculated test statistic is 2.41. The sample has 24 observations. The population is normally distributed with an unknown variance.

Degrees of freedom	p = 0.10	p = 0.05	p = 0.025	p = 0.01	p = 0.005
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797

An analyst will most likely reject the null hypothesis at significance levels of:

A. 0.10, 0.05, and 0.01.

B. 0.10 and 0.05.

C. 0.10 only.

Q-88. The value of a test statistic is best determined as the difference between the sample statistic and the value of the population parameter under H_0 divided by the:

A. appropriate value from the t-distribution.

B. standard error of the sample statistic.

C. sample standard deviation.

Q-89. In setting the confidence interval for the population mean of a normal or approximately normal distribution and given that the sample size is small, Student's t-distribution is the preferred approach when the variance is:

A. large.

B. known.

C. unknown.

Q-90. Jill Batten is analyzing how the returns on the stock of Stellar Energy Corp. are related with the previous month's percent change in the US Consumer Price Index for Energy (CPIENG). Based on 248 observations, she has computed the sample correlation between the Stellar and CPIENG variables to be −0.1452. She also wants to determine whether the sample correlation is statistically significant. The critical value for the test statistic at the 0.05 level of significance is approximately 1.96. Batten should conclude that the statistical relationship between Stellar and CPIENG is:

A. significant, because the calculated test statistic has a lower absolute value than the critical value for the test statistic.

B. significant, because the calculated test statistic has a higher absolute value than the critical value for the test statistic.

C. not significant, because the calculated test statistic has a higher absolute value than the critical value for the test statistic.

1.26. P-Value

1.26.1. 重要知识点

1.26.1.1. P-value method

- The p-value is the smallest level of significance at which the null hypothesis can be reject
- $p-value \in [0,1]$
- p-value < α: reject H_0; p-value > α: do not reject H_0.
- P↓, easier to reject H_0

1.26.2. 基础题

Q-91. A two-tailed test of the null hypothesis that the mean of a distribution is equal to 4.00 has a p-value of 0.0567. Using a 5% level of significance (i.e., α = 0.05), the best conclusion is to:

A. fail to reject the null hypothesis.

B. increase the level of significance to 5.67%.

C. reject the null hypothesis.

Q-92. The null hypothesis of a two-tailed test is least likely to be rejected when the p-value of the test statistic:

A. exceeds a specified level of significance.

B. falls below half of a specified level of significance.

C. falls below a specified level of significance.

1.27. Type I and Type II Errors

1.27.1. 重要知识点

1.27.1.1. Type I and type II error

- **Type I error(拒真):** reject a true null hypothesis
  - Significance level (α): the probability of making a Type I error

- Significance level =P(Type I error)=P(H_0×|H_0√)

➢ **Type II error(取伪):** do not reject a false null hypothesis

- P(Type II error) = P(H_a×|H_a√)
- **Power of a test**: the probability of correctly rejecting the null hypothesis when it is false.
- Power of test = 1 − P(Type II error) = P(H_a√|H_a√)

1.27.2. 基础题

Q-93. All else equal, is specifying a smaller significance level in a hypothesis test likely to increase the probability of a:

	Type I error?	Type II error?
A.	No	No
B.	No	Yes
C.	Yes	No

Q-94. All else being equal, if the probability that fail to reject the null hypothesis when it's actually false increases, how about the width of confidence interval?

A. Increase

B. Decrease

C. No change

Q-95. When testing a hypothesis, the power of a test is best described as the:

A. probability of rejecting a true null hypothesis.

B. probability of correctly rejecting the null hypothesis.

C. same as the level of significance of the test.

1.28. Nonparametric tests & Tests of Independence

1.28.1. 重要知识点

1.28.1.1. Nonparametric tests

➢ A nonparametric test either **is not concerned with a parameter** or **makes minimal assumptions about the population** from which the sample comes.

➢ Nonparametric tests are used:

- when data do not meet distributional assumptions.
    - Example: hypothesis test of the mean value for a variable, but the distribution of the variable is not normal and the sample size is small so that neither the t-test nor the z-test are appropriate.
- when data are given in ranks.
- when the hypothesis we are addressing does not concern a parameter.

1.28.1.2. Tests of Independence

- Test whether there is a relationship between the size and investment type, we can perform a test of independence using a nonparametric test statistic that is chi- square distributed

$$\chi^2 = \sum_{i=1}^{m} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

- m = the number of cells in the table, which is the number of groups in the first class multiplied by the number of groups in the second class.
- O_{ij} = the number of observations in each cell of row i and column j.
- E_{ij} = the expected number of observations in each cell of row i and column j, assuming independence.
- degrees of freedom is (r − 1)(c − 1), where r is the number of rows and c is the number of columns.
- If test statistic χ^2 > critical value, there is sufficient evidence to conclude that size and investment type are related(not independent).

1.28.2. 基础题

Q-96. An analyst wants to test the samples selecting from a population are random or not, he should choice:

A. T-test.

B. χ^2-test.

C. nonparametric test.

Q-97. Given a large random sample, which of the following types of data are least appropriately analyzed with nonparametric tests?

A. Signed data

B. Ranked data

C. Numerical values

Q-98. A contingency table can be used to test:

A. a null hypothesis that rank correlations are equal to zero.

B. whether multiple characteristics of a population are independent.

C. the number of p-values from multiple tests that are less than adjusted critical values.

1.29. The basics of simple linear regression model

1.29.1. 重要知识点

1.29.1.1. The basics of simple linear regression model

- **Function:** $Y_i = b_0 + b_1 X_i + \varepsilon_i, i = 1, \dots, n$
  - **Interpretation of the parameters**
    - The dependent variable, Y is the variable whose variation about its mean is to be explained by the regression.
    - The independent variable, X is the variable used to explain the dependent variable in a regression.
    - Regression coefficients, b0 is intercept term of the regression, b1 is slope coefficient of the regression, regression coefficient.
    - The error term, εi is the portion of the dependent variable that is not explained by the independent variable(s) in the regression.
- **The assumptions of the linear regression**
  - A linear relationship exists between X and Y;
  - The independent variable, X, is not random, with the exception that X is random but also uncorrelated with the error term
  - The expected value of the error term is zero; (i.e., $E(\varepsilon_i) = 0$)
  - The variance of the error term is constant. If not, this refers to heteroskedasticity.
  - The error term is uncorrelated across observations; (i.e., $E(\varepsilon_i\varepsilon_j) = 0$ for all $i \neq j$)
  - The error term is normally distributed.

1.29.2. 基础题

Q-99. Which of the following is least likely a necessary assumption of simple linear regression analysis?

A. The residuals are normally distributed.

B. There is a constant variance of the error term.

C. The dependent variable is uncorrelated with the residuals.

Q-100. What is the most appropriate interpretation of a slope coefficient estimate of 10.0?

A. The predicted value of the dependent variable when the independent variable is zero is 10.0.

B. For every one unit change in the independent variable, the model predicts that the dependent variable will change by 10 units.

C. For every 1-unit change in the independent variable, the model predicts that the dependent variable will change by 0.1 units.

1.30. Estimate of Regression Coefficients & Hypothesis Test

1.30.1. 重要知识点

1.30.1.1. Point estimate:

- $\hat{b}_1 = b_1 \quad \hat{b}_0 = b_0$
- Calculation of $\hat{b}_1$ and $\hat{b}_0$
  - **Ordinary least squares (OLS):** Minimize the sum of squared vertical distances between the observations and the regression line (also called residuals or error terms).
    - $\hat{b}_1 = \frac{Cov(X,Y)}{Var(X)}$; $\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X}$
- Regression coefficient confidence interval
  - $\hat{b}_1 \pm t_c s_{\hat{b}_1}$ ← t_c 查表所得
    - If the confidence interval at the desired level of significance does not include zero, the null is rejected, and the coefficient is said to be statistically different from zero.

1.30.1.2. Hypothesis testing about the regression coefficient

- **Significance test for a regression coefficient**
  - H_0: b_1=The hypothesized value
  - $t = \frac{\hat{b}_1 - hypothesized\ value\ of\ b_1}{s_{\hat{b}_1}}$, df=n-2
  - Decision rule: reject H_0 if | t | > + t critical
  - Rejection of the null means that the slope coefficient is <u>significantly different from zero.</u>
- **P-value Method**
  - H_0: b_1= 0

- The **p-value** is the smallest level of significance at which the null hypothesis can be reject.
- p-value < α: reject H_0.
- reject H_0 means the coefficient is significantly different from zero.

➢ **Measure Fitness-ANOVA Table**

- **ANOVA table**

	df	SS	MSS
Regression	k	RSS	MSR=RSS/k
Error	n-2(n-k-1)	SSE	MSE=SSE/(n-2)
Total	n-1	SST	

- **Standard error of estimate**

$$SEE = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{MSE}$$

- **Coefficient of determination (R^2)**

$$R^2 = \frac{RSS}{SST} = 1 - \frac{SSE}{SST}$$

- Standard **error of estimate & coefficient of determination (R^2)**
  - ◆ The SEE is the standard deviation of the error terms in the regression.
  - ◆ The Coefficient of Determination (R2) is defined as the percentage of the total variation in the dependent variable explained by the independent variable.
  - ◆ Example: R^2 of 0.63 indicates that the variation of the independent variable explains 63% of the variation in the dependent variable.

➢ **Measure Fitness-F-test:** F test assesses the effectiveness of the model as a whole in explaining the dependent variable.

- Define hypothesis:
  - ◆ H_0: $b_1 = b_2 = b_3 = \ldots = b_k = 0$
  - ◆ H_a: at least one $b_j \neq 0$ (for j = 1, 2, …, k)
- **F-statistic**$= F = \frac{MSR}{MSE} = \frac{RSS/k}{SSE/(n-k-1)}$
- Critical value (查表)：F_α (k, n-k-1) "one-tailed" F-test; alpha=5%
- Decision rule
  - ◆ Reject $\mathbf{H_0}$: if F-statistic > F_α (k, n-k-1)

1.30.2. 基础题

Q-101. Bill Coldplay, CFA, is analyzing the performance of the Vigorous Growth Index Fund (VIGRX) over the past three years. The fund employs a passive management investment approach designed to track the performance of the MSCI US Prime Market Growth Index, a broadly diversified index of growth stocks of large U.S. companies. Coldplay estimates a regression using excess monthly returns on VIGRX (exVIGRX) as the dependent variable and excess monthly returns on the S&P 500 Index (exS&PSOO) as the independent variable. The data are expressed in decimal terms (e.g., 0.03, not 3%).

$$\mathbf{exVIGRX}_t = \boldsymbol{b}_0 + \boldsymbol{b}_1\left(\mathbf{exS\&P500_t}\right) + \boldsymbol{\varepsilon}_\mathbf{t}$$

Results from that analysis are presented in the following figures.

Estimated Coefficients		
Coefficients	Coefficient Estimate	Standard Error
$\boldsymbol{b}_0$	0.0023	0.0022
$\boldsymbol{b}_1$	1.1163	0.0624

Partial ANOVA Table	
Source of Variation	Sum of Squares
Regression (explained)	0.0228
Error (unexplained)	0.0024

Coldplay would like to test the following hypothesis: $\boldsymbol{H}_0: \boldsymbol{b}_1 \leq \mathbf{1}$ versus $\boldsymbol{H}_1: \boldsymbol{b}_1 > \mathbf{1}$ at the 1 % significance level. The calculated t-statistic and the appropriate conclusion are:

	Calculated t-statistic	Appropriate conclusion
A.	1.86	Reject $\boldsymbol{H}_0$
B.	1.86	Fail to reject $\boldsymbol{H}_0$
C.	2.44	Reject $\boldsymbol{H}_0$

Q-102. Bill Coldplay, CFA, is analyzing the performance of the Vigorous Growth Index Fund (VIGRX) over the past three years. The fund employs a passive management investment approach designed to track the performance of the MSCI US Prime Market Growth Index, a broadly diversified index of growth stocks of large U.S. companies. Coldplay estimates a regression using excess monthly returns on VIGRX (exVIGRX) as the dependent variable and excess monthly returns on the S&P 500 Index (exS&PSOO) as the independent variable. The data are expressed in decimal terms (e.g., 0.03, not 3%).

$$\mathbf{exVIGRX}_t = b_0 + b_1\left(\mathbf{exS\&P500_t}\right) + \varepsilon_t$$

Results from that analysis are presented in the following figures.

Estimated Coefficients		
Coefficients	Coefficient Estimate	Standard Error
b_0	0.0023	0.0022
b_1	1.1163	0.0624

Partial ANOVA Table	
Source of Variation	Sum of Squares
Regression (explained)	0.0228
Error (unexplained)	0.0024

The R^2 from the regression is closest to:

A. 0.095.

B. 0.295.

C. 0.905.

Q-103. Bill Coldplay, CFA, is analyzing the performance of the Vigorous Growth Index Fund (VIGRX) over the past three years. The fund employs a passive management investment approach designed to track the performance of the MSCI US Prime Market Growth Index, a broadly diversified index of growth stocks of large U.S. companies. Coldplay estimates a regression using excess monthly returns on VIGRX (exVIGRX) as the dependent variable and excess monthly returns on the S&P 500 Index (exS&PSOO) as the independent variable. The data are expressed in decimal terms (e.g., 0.03, not 3%).

$$\mathbf{exVIGRX}_t = b_0 + b_1\left(\mathbf{exS\&P500_t}\right) + \varepsilon_t$$

Results from that analysis are presented in the following figures.

Estimated Coefficients		
Coefficients	Coefficient Estimate	Standard Error
b_0	0.0023	0.0022
b_1	1.1163	0.0624

Partial ANOVA Table	
Source of Variation	Sum of Squares
Regression (explained)	0.0228

Error (unexplained)	0.0024

The standard error of estimate is closest to:

A. 0.008.

B. 0.014.

C. 0.049.

1.31. Estimate of Y

1.31.1. 重要知识点

1.31.1.1. Estimate of Y

- Predicted values are values of the dependent variable based on the estimated regression coefficients and a prediction about the value of the independent variable.
- Point estimate

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

- Confidence interval estimate

$$\hat{Y} \pm \left(t_c \times S_f\right)$$

  - t_c = the critical t-value with df=n−2
  - s_f = the standard error of the forecast

$$S_f = SEE \times \sqrt{1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{(n-1)S_X^2}} = SEE \times \sqrt{1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{\sum(X_i - \bar{X})^2}}$$

1.31.1.2. Forms of Simple Linear Regression

- Log-Lin Model

$$\ln Y = b_0 + \boldsymbol{b_1} X$$

- Lin-log model

$$\mathrm{Y} = \boldsymbol{b_0} + \boldsymbol{b_1} \ln X$$

- Log-log model

$$\ln \boldsymbol{Y} = \boldsymbol{b_0} + \boldsymbol{b_1} \ln \boldsymbol{X}$$

1.31.2. 基础题

Q-104. Elena Vasileva recently joined EnergyInvest as a junior portfolio analyst. Vasileva's supervisor asks her to evaluate a potential investment opportunity in Amtex, a multinational oil and gas corporation based in the United States. Vasileva's supervisor suggests using regression analysis to examine the relation between Amtex shares and returns on crude oil.

Vasileva runs a regression of Amtex share returns （Y）on crude oil returns （X）using the monthly data she collected. Selected regression output is presented in Exhibit 1. She uses a 1% level of significance in all her tests.

Vasileva expects the crude oil return next month, Month 37, to be −0.01. She computes the standard error of the forecast to be 0.0469.

Exhibit 1 Selected Data for Crude Oil Returns and Amtex Share Returns		
	Coefficient	Standard Error
Intercept	0.0095	0.0078
Oil return	0.2354	0.0760

Critical t-values for a 1% level of significance:

One-sided, left side: −2.441

One-sided, right side: +2.441

Two-sided: $\pm$2.728

Using information from Exhibit 1, the 99% prediction interval for Amtex share return for Month 37 is best described as:

A. $\hat{Y}_f \pm 0.0053$

B. $\hat{Y}_f \pm 0.0469$

C. $\hat{Y}_f \pm 0.1279$

1.32. 进阶题

Q-1. The table below shows three mutually exclusive $2,000,000 mortgage choices. Each of the three choices is compounded monthly.

Mortgage type	Quoted annual interest rate at initiation
32-year fixed rate	6.5%
24-year fixed rate	6.0%
32-year adjustable rate	4.5%

The adjustable-rate mortgage will reset its interest rate to 6.2% at the end of the year 4. After resetting the interest rate at the end of year 4, which mortgage will have the largest monthly payment?

A. 32-year fixed rate mortgage.

B. 24-year fixed-rate mortgage.

C. 32-year adjustable-rate mortgage.

Q-2. When rolling two six-sided dice and summing their outcomes, which of the following sums is most likely to occur?

A. Nine

B. Six

C. Five

Q-3. Independent samples drawn from normally distributed populations exhibit the following characteristics:

Sample	Size	Sample Mean	Sample Standard Deviation
A	28	210	50
B	21	195	65

Assuming that the variances of the underlying populations are equal, the pooled estimate of the common variance is 3,377.13. The t-test statistic appropriate to test the hypothesis that the two population means are equal is closest to:

A. 1.80.

B. 0.31.

C. 0.89.

Q-4. Two distributions have the same mean. One is negatively skew, the other is positive skew. Which one has the larger median?

A. Distribution with negative skew.

B. Distribution with positive skew.

C. The same.

Q-5.

Population	1	2
Sample size	$n_1 = 6$	$n_2 = 6$
Sample variance	$S_1^2 = 5$	$S_2^2 = 30$
The samples are drawn independently, and both populations are assumed to be normally distributed		

Using the above data, an analyst is trying to test the null hypothesis that the population variances are equal (H_0: $\sigma_1^2 = \sigma_2^2$) against the alternative hypothesis that the variances are not equal (H_a: $\sigma_1^2 \neq \sigma_2^2$) at the 5% level of significance. The table of the F-Distribution is provided below.

Table of the F-Distribution

Panel A: Critical values for right-hand tail areas equal to 0.05

df1 (read across) / **df2 (read down)**	1	2	3	4	5
1	161	200	216	225	230
2	18.5	19.0	19.2	19.2	19.3
3	10.1	9.55	9.28	9.12	9.01
4	7.71	6.94	9.59	6.39	6.26
5	6.61	5.79	5.41	5.19	5.05

Panel B: Critical values for right-hand tail areas equal to 0.025

df1 (read across) / **df2 (read down)**	1	2	3	4	5
1	648	799	864	900	922
2	38.51	39.00	39.17	39.25	39.30
3	17.44	16.04	15.44	15.10	14.88
4	12.22	10.65	9.98	9.60	9.36
5	10.01	8.43	7.76	7.39	7.15

Which of the following statements is most appropriate? The critical value is:

A. 9.36 and reject the null.

B. 9.60 and do not reject the null.

C. 7.15 and do not reject the null.

Q-6. Using the following sample results drawn as 25 paired observations from their underlying distributions, test whether the mean returns of the two portfolios differ from each other at the 1% level of statistical significance. Assume the underlying distributions of returns for each portfolio are normal and that their population variances are not known.

	Portfolio 1	Portfolio 2	Difference
Mean return	15.00	20.25	5.25
Standard deviation	15.50	15.75	6.25
t-statistic for 24 degrees of freedom and at the 1% level of statistical significance = 1.711			
Null hypothesis (H_0): Mean difference of returns = 0			

Based on the paired comparisons test of the two portfolios, the most appropriate conclusion is that H_0 should be:

A. accepted because the computed test statistic exceeds 1.711.

B. rejected because the computed test statistic exceeds 1.711.

C. accepted because the computed test statistic is less than 1.711.

Q-7. If the population distribution is unknown, the method that will lead to the *least* reliable estimation of a parameter is to:

A. use point estimates instead of confidence interval estimates.

B. use t-distribution instead of standard normal distribution to establish confidence intervals.

C. draw more samples.

Q-8. The table below reports the annual returns for two active portfolios in the same industry, namely, their returns are dependent with each other.

Year	Portfolio A (%)	Portfolio B (%)
2013	11	9
2014	-10	4
2015	1	-3
2016	8	12
2017	21	23
2018	2	-4

If we want to test whether the two portfolios have the same mean return at a 5% significance level, the test statistics we shall use is *closest* to:

A. 1.96.

B. 1.66.

C. 0.45.

Q-9. In a head and shoulders pattern, if the neckline is at $23, the shoulders at $28, and the head at $33. The price target is closest to which of the following:

A. $13.

B. $19.

C. $40.

Q-10. An analyst has established the following prior probabilities regarding a company's next quarter's earnings per share (EPS) exceeding, equaling, or being below the consensus estimate.

	Prior Prababilities
EPS exceed consensus	23%
EPS equal consensus	56%
EPS are less than consensus	21%

Several days before releasing its earnings statement, the company announces a cut in its dividend. Given this new information, the analyst revises his opinion regarding the likelihood that the company will have EPS below the consensus estimate. He estimates the likelihood the company will cut the dividend, given that EPS exceeds/meets/falls below consensus, as reported below.

	Probabilities the Company Cuts Dividends, Conditional on EPS Exceeding/Equaling/Falling below consensus
P(Cut div/EPS exceed)	3%
P(Cut div/EPS equal)	11%
P(Cut div/EPS below)	86%

Using Bayes' formula, the updated (posterior) probability that the company's EPS are below the consensus is closest to:

A. 73%.

B. 84%.

C. 22%.

Q-11. Samples of size (n_1,n_2) are drawn respectively from two populations (X_1,X_2) with associated sample means and standard deviations of ($\bar{\mathbf{X}}_1$, $\bar{\mathbf{X}}_2$) and (S_1,S_2) and associated population means and standard deviations of(μ_1,μ_2) and(σ_1,σ_2) where ($\sigma_1 \neq \sigma_2$). In addition, $\bar{\mathbf{d}}$ is the sample mean of $\bar{\mathbf{X}}_1$-$\bar{\mathbf{X}}_2$ with a standard error of $\mathbf{S}_{\bar{\mathbf{d}}}$ and a population mean of $\boldsymbol{\mu}_d$ and $\mathbf{S}_\mathbf{p}^2$ is a pooled estimator of the common variance.

The most appropriate test statistic to determine the equality of the two population means assuming X_1 and X_2 are independent and normally distributed is:

A. $t = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}}$

B. $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2})^{0.5}}$

C. $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})^{0.5}}$

Q-12. Monte Carlo simulation is best described as:

A. a restrictive form of scenario analysis.

B. providing a distribution of possible solutions to complex functions.

C. an approach to backtest data.

Q-13. Which of the following most accurately describes how to standardize a random variable X?

A. Subtract the mean of X from X, and then divide that result by the standard deviation of X.

B. Subtract the mean of X from X, and then divide that result by the standard deviation of the standard normal distribution.

C. Divide X by the difference between the standard deviation of X and the standard deviation of the standard normal distribution.

Q-14. A descriptive measure of a population characteristic is best described as a:

A. parameter.

B. frequency distribution.

C. sample statistic.

Q-15. The discrepancy between a statistically significant result and an economically meaningful result is least likely the result of:

A. transaction costs.

B. sampling errors.

C. risk tolerance.

Solutions

1. Quantitative

1.1. 基础题

Q-1. Solution: C.

With a low credit rating and high default risk, the company would most likely give a higher required rate of return to investor.

Q-2. Solution: A.

The real risk free rate equals to the nominal risk-free rate minus the expected inflation rate. If the nominal risk-free rate decrease and expected inflation rate increase, the real risk-free rate should decrease.

Q-3. Solution:B.

Explain an interest rate as the sum of a real risk-free rate and premiums that compensate investors for bearing distinct types of risk.

B is correct. The maturity premium compensates investors for the increased sensitivity of the market value of debt to a change in market interest rates as maturity is extended, in general (holding all else equal).

Q-4. Solution: C.

C is correct. US Treasury bonds are highly liquid, whereas the bonds of small issuers trade infrequently and the interest rate includes a liquidity premium. This liquidity premium reflects the relatively high costs (including the impact on price) of selling a position.

Q-5. Solution: A.

A is correct. The continuously compounded return of an asset over a period is equal to the natural log of period's change. In this case: ln(120/112) = 6.90%.

Q-6. Solution: B.

Holding period return, HPR = $(P_1 – P_0 + D_1)/P_0$
where
P_0 = initial investment
P_1 = price received at the end of holding period
D_1 = dividend paid by the investment at the end of holding period
$P_1 = 48$, D_1= \$1,250/5,000 shares = \$0.25/shares
$(P_1 – P_0 +D_1)/P_0=(48 – P_0 +0.25)/P_0=7\%$ and solving for P_0 = \$45.09.

Q-7. Solution: A.

Quarterly: EAR = $(1+4.1\%/4)^4 - 1 = 4.1635\%$

Interest = $1,000,000 × 4.1635% = $41,635

Continuous: EAR = $e^{0.04} - 1 = 4.0811\%$

Interest = $1,000,000 x 4.0811% = $40,811

Therefore, the CD1 offers the highest effective annual rate with a $41,635 interest.

Q-8. Solution: B.

The effective annual rate is calculated as EAR = $(1 + \text{Periodic interest rate})^m - 1 = (1 + r/m)^m - 1$

For quarterly compounding, $(1 + 10\% / 4)^4 - 1 = 10.3813\%$

For monthly compounding, $(1 + 10\% / 12)^{12} - 1 = 10.471307\%$

For semi-annually compounding, $(1 + 10\% / 2)^2 - 1 = 10.25\%$

Therefore, the correct answer is monthly compounding.

Q-9. Solution: A.

All else being equal, due to the different payments, PV of option 1 will be the lowest, while PV of option 3 is the highest.

Calculation:

The present value for option 1 is $24,924. PMT=-2,000, N=20, I/Y=5, CPT: PV=24,924.

The present value for option 2 is $26,171. BGN mode,PMT=-2,000, N=20, I/Y=5, CPT: PV=26,171.

The present value for option 3 is $40,000, A=2,000, discount rate=5%.

PV=A/r=2,000/0.05=40,000

Option 1 (ordinary annuity) is the last option to choose.

Q-10. Solution: B.

To solve for the future value of unequal cash flows, compute the future value of each payment as of Year 4 at the semiannual rate of 2%, and then sum the individual future values, as follows:

Year	End of Year Deposits ($)	Factor	Future Value ($)
1	4,000	$(1.02)^6$	4,504.65
2	8,000	$(1.02)^4$	8,659.46
3	7,000	$(1.02)^2$	7,282.80
4	10,000	$(1.02)^0$	10,000.00
Sum =	-	-	30,446.91

Q-11. Solution: B.

Using a financial calculator, first calculate the needed funds at retirement:

N = 20, I/Y = 6, PMT =80,000, FV = 0; calculate PV =-917,593.6975.

Then use 917,593.6975 as the FV of the accumulation phase annuity:

I/Y = 6, PV = 0, PMT = -11,606.56, FV = 917,593.6975, CPT N=30.

Q-12. Solution: C.

Using a financial calculator: N = 8x12=96; the discount rate, I/Y = (6/12) = 0.5; PMT = 4,800; FV = 0; Mode = BGN; CPT PV = - 367,083.3329.

Q-13. Solution: B.

Annual payment=€40,000x7%=€2,800

1+EAR =$(1+\frac{0.04}{12})^{12}$=1.04074154

EAR=4.074154%

PV=0, PMT=2,800, I/Y=4.074154, N=4, CPT FV = -11,903.24.

Total Future value=€40,000 + €11,903.24 = €51,903.24

A is not correct since the reinvestment return is not considered.

€40,000 + €40,000 × 7% × 4 = €51,200.00.

C is not correct. The initial investment of 40,000 in the certificate of deposit (CD) does not need to count for compounding. Total Future value=€40,000 × (1+EAR)^4+€11,903.24 = €58,831.19

Q-14. Solution: B.

We can perform mathematical operations on numerical data but not on categorical data. Numerical data can be discrete or continuous.

Q-15. Solution: A.

The value 34% is the joint probability that a voter supports both Jones and Williams. Because it is stated as a percentage, this value is a relative frequency. The totals for each row and column are marginal frequencies. An absolute frequency is a number of occurrences, not a percentage of occurrences.

Q-16. Solution: A.

A is correct. A frequency distribution is a tabular display of data summarized into a relatively small number of intervals.

B is incorrect because intervals cannot overlap. Each observation is placed uniquely into one interval.

C is incorrect because a frequency distribution is summarized into a relatively small number of intervals.

Q-17. Solution: A.

The relative frequency is the number of observations in an interval divided by the total number of observations. For Interval III, relative frequency = 11/55 = 20%. The cumulative relative frequency is the sum of the relative frequencies of the relevant class and all the classes before it. For Interval III, the cumulative relative frequency = (12 + 24 + 11)/55 = 85.45%.

Q-18. Solution: A.

Data on some characteristics of companies at a single point in time are cross-sectional data.

Q-19. Solution: B.

To find the median, rank the returns in order and take the middle value: -8%, 4%, 6%, 12%, 12%, 23%. In this case, because there is an even number of observations, the median is the average of the two middle values, or (6%+ 12 %) / 2 = 9.0%.

Q-20. Solution: A.

A is correct. The geometric mean compounds the periodic returns of every period, giving the investor a more accurate measure of the terminal value of an investment.

Q-21. Solution: C.

The portfolio return is the weighted mean return and is calculated as follows:

$$\overline{X}_1 = \sum_{i=1}^{n} w_{e1}X_{e1} + w_{b1}X_{b1} = (15\% \times 60\%) + (11\% \times 40\%) = 13.4\%$$

$$\overline{X}_2 = \sum_{i=1}^{n} w_{e2}X_{e2} + w_{b2}X_{b2} = (11\% \times 60\%) + (-5.6\% \times 40\%) = 4.36\%$$

$$\overline{X}_3 = \sum_{i=1}^{n} w_{e3}X_{e3} + w_{b3}X_{b3} = (-13.86\% \times 60\%) + (12\% \times 40\%) = -3.52\%$$

$$Geometric\ mean = \sqrt[N]{X_1X_2X_3} = \sqrt[3]{(1+13.4\%)(1+4.36\%)(1-3.52\%)} - 1 = 0.045189$$

Q-22. Solution: C.

The harmonic mean is appropriate for determining the average price per unit. It is calculated by summing the reciprocals of the prices; then averaging that sum by dividing by the number of prices; and finally, taking the reciprocal of the average:

5/[(1/124.00) + (1/152.00) + (1/168.00) + (1/180.00)+ (1/184.00)] = €158.2971

Q-23. Solution: C.

The first quintile is the 20th percentile. The first decile is the 10th percentile, the first quartile is the 25th percentile, and the median is the 50th percentile. While it is possible that these various percentiles or some subsets of them be equal (for example the 10th percentile possibly could be equal to the 20th percentile), in general the order from smallest to largest would be: first decile, first quintile, first quartile, median.

Q-24. Solution: B.

1. Data arranged in ascending order:

−40.33%, −5.02%, 9.57%, 10.02%, 12.34%, 15.25%, 16.54%, 20.65%, 27.37%, and 30.79%

2. Fourth quintile $=\frac{1}{5}\times 4=80\%$

3. Ly 位置 = (n + 1) × 80% = 11 × 0.8 = 8.8.

4. The 8.8th position would be between the 8th and 9th entries, 20.65% and 27.37%, respectively. Using linear interpolation, $P_{80} = X_8 + (Ly – 8) \times (X_9 – X_8)$,

P_{80} = 20.65 + (8.8 – 8) × (27.37 – 20.65) = 26.03 (%)

Q-25. Solution: B.

The arithmetic sum of the deviations around the mean will always equal zero, not one.
A is incorrect. Range does not provide information about the shape of the distribution.
C is incorrect. The mean absolute deviation will always be less than or equal to the standard deviation. （18Mock AS Q30）未改编

Q-26. Solution: A.

Compute the mean portfolio return = (17.00% + 22.20%+ 25.60% + 30.40% - 19.00%)/5 = 15.24%；

MAD = (|17.00%-15.24%| + |22.20%-15.24%| + |25.60%-15.24%| + |30.40%-15.24%| + |-19.00% -15.24%|)/5 = 13.6960%

Variance = $[(17.00\%-15.24\%)^2 + (22.20\%-15.24\%)^2 + (25.60\%-15.24\%)^2 + (30.40\%-15.24\%)^2 + (-19.00\% -15.24\%)^2]/5$ ≈3.1221%

The population variance calculation is appropriate because the analyst is analyzing all the annual returns on the portfolio since its inception.

Q-27. Solution: B.

Sharpe ratio = [expected return (mean) – risk-free rate]/standard deviation = mean/standard deviation; CV = standard deviation/expected return. The mean is more than the standard deviation, so compared with Sharpe ratio, the coefficient of variation is less.

Q-28. Solution: B.

The Sharpe ratio is the mean excess return (mean return less risk-free rate of 5.0 percent) divided by the standard deviation of the portfolio. It is highest for portfolio 2 with a Sharpe ratio of (10-5)

/ 20= 0.2500. For portfolio 1, the Sharpe ratio is (9-5)/18 = 0.2222 and for portfolio 3 the Sharpe ratio is (12-5)/ 32 = 0.2188.

Q-29. Solution: A.

Sharpe ratio = $[E(R_p)- r_f]/\sigma$, based on the Sharpe ratio formula,

we can get the $\sigma = [E(R_p)- r_f]$/Sharpe ratio,

σ_1 = (10%-5%)/34% = 14.71%,

σ_2 = (10%-5%)/37% = 13.51%.

CV = $\sigma/\overline{X}$, CV_1 = 14.71%/10% = 1.471, CV_2 = 13.51%/10% = 1.351.

The portfolio 1's CV is larger.

Q-30. Solution: C.

Having an excess kurtosis, the actual distribution of one year investment returns is leptokurtic return distribution. It is more peaked and has fatter tails than the normal distribution, which means more extremely large deviations from the mean than a normal distribution and an undervalued probability than extreme returns occurs.

Q-31. Solution: A.

A distribution with a mean greater than its median is positively skewed, or skewed to the right. The skew pulls the mean. Kurtosis deals with the overall shape of a distribution, not its skewness.

Q-32. Solution: B.

Sample skewness is measured using deviations raised to the third power.

Sample kurtosis is measured using deviations raised to the fourth power.

Q-33. Solution: B.

Explain measures of sample skewness and kurtosis.

B is correct. Most equity return series have been found to be leptokurtic.

Q-34. Solution: C.

An empirical probability is a probability estimated from data as a relative frequency of occurrence.

A is incorrect. A subjective probability is a probability drawing on personal or subjective judgment.

B is incorrect. An a priori probability is a probability obtained based on logical analysis.

Q-35. Solution: C.

The probability of underperform = 1–65% = 35%

The odds against of underperform = 65%/35% =1.8571

Q-36. Solution: B.

Price change includes price increase.

Q-37. Solution: A.

A is correct. Two events, A and B, are independent if and only if P(A|B) = P(A) or, equivalently, P(B|A) = P(B). The wording of the question precludes P(A) = P(B); therefore, P(B) and P(B|A) cannot be correct.

B is incorrect. Two events A and B are independent if and only if P(A|B) = P(A) or, equivalently, P(B|A) = P(B). As P(A) ≠ P(B), B cannot be correct.

C is incorrect. Two events A and B are independent if and only if P(A|B) = P(A) or, equivalently, P(B|A) = P(B). As P(A) ≠ P(B) and given that P(B) = P(B|A), C cannot be correct.

Q-38. Solution: B.

The three outcomes given for economic growth are mutually exclusive and exhaustive. The probability that economic growth is positive but less than 3% is 100% - 25% - 25% = 50%. Using the total probability rule, the probability that the share price increases is (80%)(25%) + (40%)(50%) + (10%)(25%) = 42.5%.

Q-39. Solution: B.

B is correct. Assuming A and B represent portfolio 1 gains return in excess of 3% and portfolio 2 gains return in excess of 3%, respectively.

P(A)=0.3, P(A or B)=0.8, P(A|B)=0.25.

P(A or B)=P(A)+P(B)-P(AB). P(AB)=P(A)+P(B)-P(A or B)

P(A|B)=P(AB)/P(B)= P(A)+P(B)-P(A or B)/P(B).

Calculate P(B)=0.667

Q-40. Solution: B.

When the economy prosperity:

E(EPS) = 85% × 9 + 15% × 3 = 8.1

Var(EPS) = 85% × $(9-8.1)^2$ + 15% × $(3 – 8.1)^2$ =4.59

Q-41. Solution: A.

For an extreme case in which $\rho_{XY} = -1$ (that is, the two asset returns move in opposite directions), the portfolio can be made risk free.

Q-42. Solution: B.

Expected return on security A = 0.7 x 24% + 0.3 x 18% = 22.2%

Expected return on security B = 0.7 x 32% + 0.3 x 24% = 29.6%

Cov (R_A, R_B) = 0.7 x [(24% - 22.2%) x (32% - 29.6%)] + 0.3 x [(18% - 22.2%) x (24% - 29.6%)]

= 0.001008.

Q-43. Solution: B.

Correlation of +0.3 indicates a positive linear relationship between the variables—one tends to be above its mean when the other is above its mean. The value 0.3 indicates that the linear relationship is not particularly strong. Correlation does not imply causation.

Q-44. Solution: B.

$$E(X) = \sum P_i \times X_i = 80\% \times 6\% + 20\% \times 8\% = 6.4\%$$

$$\sigma_p^2 = w_{stock}^2 \sigma_{stock}^2 + w_{fund}^2 \sigma_{fund}^2 + 2 \times w_{stock} \times w_{fund} \times \sigma_{stock} \times \sigma_{fund} \times \rho_{stock,fund}$$

$$= 80\%^2 \times 25\%^2 + 20\%^2 \times 30\%^2 + 2 \times 80\% \times 20\% \times 25\% \times 30\% \times 0.2 = 4.84\%$$

$$\sigma_p = \sqrt{4.84\%} = 22\%$$

The rate of return and the standard deviation of portfolio is 6.4% and 22%.

Q-45. Solution: B.

The correlation between two random variables R_i and R_j is defined as $\rho(R_i,R_j) = Cov(R_i,R_j)/[\sigma(R_i)\sigma(R_j)]$. Using the subscript i to represent hedge funds and the subscript j to represent the market index, the standard deviations are $\sigma(R_i) = 225^{1/2} = 15$ and $\sigma(R_j) = 64^{1/2} = 8$. Thus, $\rho(R_i,R_j) = Cov(R_i,R_j)/[\sigma(R_i)\sigma(R_j)] = 90/(15X8) = 0.75$.

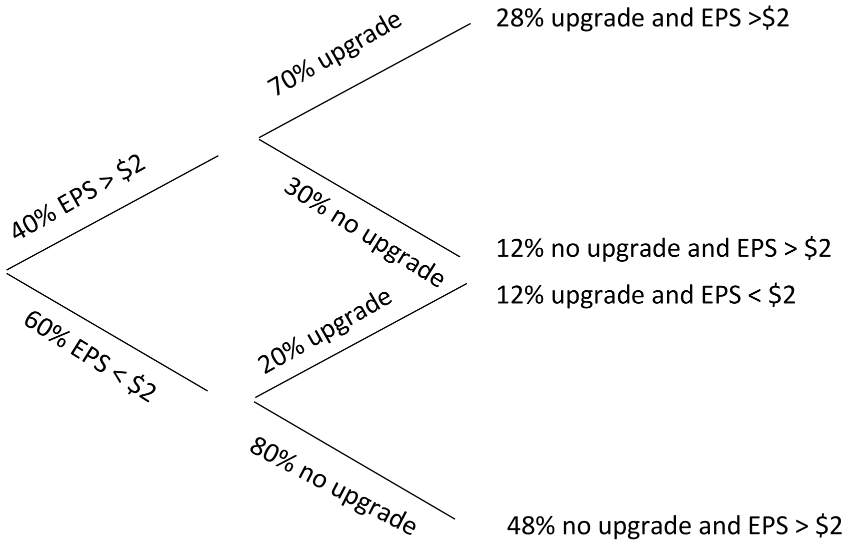
Q-46. Solution: A.

In probability natation, Bayes' formula can be writtrn concisely as:

$$\mathrm{P}(Event|Information) = \frac{P(Information|Event)}{P(Information)}\mathrm{P(Event)}$$

Q-47. Solution: C.

This is an application of Bayes'formula. As the tree diagram below shows, the updated probability that earnings per share are greater than $2 is 28%/(28% + 12%)= 70%



Q-48. Solution: C.

P(CFA I|FRM) =50%, P(FRM) = 40%

P(CFA I and FRM) = P(CFA I|FRM) × P(FRM) = 50% × 40% = 20%

P(FRM| CFA I)= P(CFA I and FRM)/ P(CFA I)=20%/30%=0.67

Q-49. Solution: B.

${}_6C_2$=15.

Q-50. Solution: B.

Because the order of the sales is important, the permutation formula is appropriate, which is 5!/(5-3)!=60.

Alternatively, using the BAII Plus financial calculator (5+2ND+-+3+=), to compute the anwser (60).

Q-51. Solution: A.

A is correct. Although quoted stock price is a discrete random variable with possible values NZ$0, NZ$0.01, NZ$0.02, we can also model stock price as a continuous random variable (as a lognormal random variable).

B is incorrect. A random variable from a normal distribution is a continuous random variable.

C is incorrect. Outcomes 1,2,3,4,5,6 of playing a dice belong to discrete random variable, not probabilities.

Q-52. Solution: B.

The value of the cumulative distribution function lies between 0 and 1 for any x: $0 \le F(x) \le 1$.

Q-53. Solution: A.

Since X is generated from a continuous distribution, the probability that X exactly equals to a certain nember is zero, even though the event can occur.

Q-54. Solution: C.

The binomial distribution is symmetric when the probability of success on a trial is 0.50.

Q-55. Solution: A.

A is correct. A trial, such as a coin flip, will produce one of two outcomes. Such a trial is a Bernoulli trial.

Q-56. Solution: B.

The problem deals with the discrete uniform distribution. This means that the five outcomes are all equally likely: P(X) = 1/10 = 0.1.

To find P(X=1)+P(X=1.5)+P(X=2)+ P(X=2.5)+P(X=3)=0.1+0.1+0.1+0.1+0.1=0.5

Q-57. Solution: C.

Y=1 outperform p=1-0.4=0.6; Y=0 underperform 1-p=0.4.

n=4 (quarter)

The probability that the performance is at or below the expectation is calculated by finding F(2) = p(2) + p(1)+ p(0) using the formula:

p(x)=P(X=x)=[n!/(n-x)!x!](p^x)[(1-p)^(n-x)]

Using this formula,

p(2)= [4!/(4-2)!2!](0.60^2)[(1-0.60)^(4-2)]=[24/4](0.36)(0.16)=0.3456

p(1)= [4!/(4-1)!1!](0.60^1)[(1-0.60)^(4-1)]=[24/6](0.60)(0.064)=0.1536

p(0)= [4!/(4-0)!0!](0.60^0)[(1-0.60)^(4-0)]=[24/24](1)(0.0256)=0.0256

Therefore,

F(2) =p(2) + p(1)+ p(0)= 0.3456 + 0.1536 + 0.0256 = 0.5248.

Q-58. Solution: B.

Because the price follows a continuous uniform distribution that ranges from \$150 to \$210, the probability that the price will be less than \$160 is P(X < 160) = (160 - 150)/(210 - 150) = 16.7%

Q-59. Solution: B.

B is correct. A normal distribution will have one mean and one variance.

Q-60. Solution: B.

First standardize the value of return for the given normal distribution:

$$P(7\% < X < 19\%) = P(\frac{7\% - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{19\% - \mu}{\sigma})$$
$$= P(\frac{7\% - 13\%}{5\%} < \frac{X - \mu}{\sigma} < \frac{19\% - 13\%}{5\%})$$
$$= P(-1.2 < z < 1.2)$$

Using the property of standard normal distribution,

P(-1.2 < z < 1.2) = 1 – 2 × P(z > 1.2)=1 – 2 × [1 – P(z ≤ 1.2)] = 1 – 2 × [1 – N(1.2)] = 2 × N(1.2) - 1

Given the z-table, N(1.2) = 0.8849, so P(-1.2 < z < 1.2) = 2 × N(1.2) – 1 = 76.98%

Q-61. Solution: A.

For a standard normal distribution, the probability that a random variable lies within 1 standard of the mean is about 68%.

$$P(\mu - \sigma \le X \le \mu + \sigma) = 68\%$$

The probability that a random variable lies within 1.96 standard of the mean is about 95%.

$$P(\mu - 2\sigma \le X \le \mu + 2\sigma) = 95\%$$

The probability that a random variable lies within 1 standard deviation to 2 standard deviation is about 13.5%.

$$P(\sigma \le X \le 2\sigma) = \frac{(95\% - 68\%)}{2} = 13.5\%$$

Q-62. Solution: A.

Roy's safety-first criterion states that the optimal portfolio minimizes the probability that the return of the portfolio falls below some minimum acceptable level. This minimum acceptable level is called the "threshold" level. Symbolically, Roy's safety-first criterion can be stated as: Maximize the SFR where SFR = $[E(R_p) - R_L] / \sigma_p$

Where: R_p = portfolio return; R_L = threshold level return

R_L=10,000/90,000=11.11%, SFR = (14%-11.11%)/($0.0225^{1/2}$) =2.89% / 0.15=19.27%

Q-63. Solution: B.

By definition, lognormal random variables cannot have negative values.

Q-64. Solution: A.

As the degrees of freedom get larger, the t-distribution approaches the normal distribution. As the degrees of freedom fall, the peak of the t-distribution flattens and its tails get fatter (more probability in the tails—that's why, all else the same, the critical t increases as the df decreases).

Q-65. Solution: C.

There is no consistent relationship between the mean and standard deviation of the chi-square distribution or F-distribution.

Q-66. Solution: B.

If the t-distribution and the normal distribution have a same significance level, the tails of the t-distribution are fatter than the tails of normal distribution. As the degrees of freedom increase, the t-distribution approaches the standard normal.

Q-67. Solution: C.

As degrees of freedom increase, the t-distribution will more closely resemble a normal distribution, becoming more peaked and having less fat tails.

Q-68. Solution: C.

A characteristic feature of Monte Carlo simulation is the generation of a large number of random samples from a specified probability distribution or distributions to represent the role of risk in the system.

Q-69. Solution: B.

The standard deviation of a sample statistic is known as the standard error of the statistic.

Q-70. Solution: B.

Given a population described by any probability distribution (normal or non-normal) with finite variance, the central limit theorem states that the sampling distribution of the sample mean will be approximately normal, with the mean approximately equal to the population mean, when the sample size is large.

Q-71. Solution: B.

When the population variance is unknown, the standard error of the sample mean is calculated as: $\text{Standard error} = \frac{s}{\sqrt{n}}$

Deviation from Mean	Squared Deviation
(11-4.5) = 6.5	42.25
(21-4.5) = 16.5	272.25
(–7 -4.5) = –11.5	132.25
(3-4.5) = –1.5	2.25
(–8 -4.5) = –12.5	156.25
(6-4.5) = 1.5	2.25
(0-4.5) = –4.5	20.25
(-7-4.5) = –11.5	132.25
(4-4.5) = –0.5	0.25
(22-4.5) = 17.5	306.25
Total	1066.5
Variance	1066.5/9 = 118.5
Standard deviation (s):	√118.5= 10.89

The standard error of the sample mean is: $10.89/10^{0.5}$=3.44

Q-72. Solution: B.

Stratified random sampling (known as proportional random sampling or quota random sampling) involves the division of a population into smaller sub-groups known as strata. The strata are formed based on members' shared attributes or characteristics such as income or educational attainment.

Sample members of systematic sampling are selected according to a random starting point but with a fixed, periodic interval. For example, to select a random group of 1,000 people from a population of 50,000, all the potential participants must be placed in a list and a starting point would be selected. If the selected starting point was 20, the 70^{th} person on the list would be chosen followed by the 120^{th}, and so on.

Random sampling is to take all the samples at random, and each sample has the same probability of being selected. For example, for a population (1,2,3…100), each data has equal probability of being drawn.

A 分层抽样：一个总体分分组，每个组抽抽样，例如，调查中国人的平均收入，此时如果只在中关村附近抽样很可能样本均是高收入，此时可以按照不同的地区抽样，这样的方法可以使得样本中的个体差异化较大，估计总体时更为准确。B 系统抽样：系统抽样又叫等距抽样。将总体分成均衡的几个部分，然后按照预先定出的规则，按照规则，从每一部分抽取。C 简单随机抽样：简

单随机抽样要求严格遵循概率原则，每个个体被抽中的概率相同，如果有 1-100 个数字，抽 10 个出来，那么抽去哪个可以有计算机产生随机数，按照这个来抽取

A 分层抽样的结果聚集在尾部的可能性较低；B 系统抽样比 C 简单随机抽样更方便，然而，如果在总体中选择时（可能是人为规定的抽取规则）存在一种潜在的模式，这也可能导致偏差(尽管这种情况发生的几率非常低)，本题是选择一个最优选，所以选 B。

Q-73. Solution: C.

With cluster sampling, the randomly selected subgroups may have different distributions of the relevant characteristic relative to the entire population. Cluster sampling uses only randomly selected subgroups, whereas stratified random sampling samples all subgroups to match the distribution of characteristics across the entire population.

Q-74. Solution: A.

The standard error of the sample mean, when the sample standard deviation is known, is:

$s_{\overline{X}}=\frac{s}{\sqrt{n}}$. In this case, $s_{\overline{X}}=\frac{15}{\sqrt{64}}=1.875$.

The reliability factor for a 95% confidence interval with unknown population variance and sample size greater than 30 is $Z_{0.025}$=1.96.

The confidence interval estimate is:

$\overline{X}\pm Z_{0.025}\left(s/\sqrt{n}\right)$.

With sample standard error of 1.875, the estimated confidence interval is:

$$8\pm1.96\times1.875=8\pm3.675$$

Q-75. Solution: A.

A is correct. As the degree of confidence is increased, the confidence interval becomes wider. A larger sample size decreases the width of a confidence interval.

B is incorrect. As the degree of confidence is increased, the confidence interval becomes wider. A larger sample size decreases the width of a confidence interval.

C is incorrect. As the degree of confidence is increased, the confidence interval becomes wider. A larger sample size decreases the width of a confidence interval.

Q-76. Solution: A.

A consistent estimator is one for which the probability of estimates close to the value of the population parameter increases as sample size increases. More specifically, a consistent estimator's sampling distribution becomes concentrated on the value of the parameter it is intended to estimate as the sample size approaches infinity.

Q-77. Solution: C.

Bootstrapping, repeatedly drawing samples of equal size from a large data set, is more computationally demanding than the jackknife. We have not defined "systematic resampling" as a specific technique.

Q-78. Solution: B.

B is correct. A report that uses a current list of stocks does not account for firms that failed, merged, or otherwise disappeared from the public equity market in previous years. As a consequence, the report is biased. This type of bias is known as survivorship bias.

Q-79. Solution: B.

B is correct. One-tailed tests in which the alternative is "greater than" or "less than" represent the beliefs of the researcher more firmly than a "not equal to" alternative hypothesis.

Q-80. Solution: B.

For testing a sample mean with a small sample size and known population variance, Z distribution should be used.

Q-81. Solution: B.

The t-test is based on $t = \frac{X - \mu_0}{s / \sqrt{n}}$.

For this test, we have $t = \frac{0.06 - 0}{0.20 / \sqrt{100}} = 3$

Q-82. Solution: C.

The sample size was over 90, which was more than 30, so z-test is appropriate. Using a 5% significant level, the critical value of a two-tailed test is 1.96. The z-statistic is 1.30, which is less than 1.96, so the analyst fail to reject null hypothesis.

Q-83. Solution: A.

Define a hypothesis, describe the steps of hypothesis testing, describe and interpret the choice of the null and alternative hypotheses, and distinguish between one-tailed and two-tailed tests of hypotheses.

A is correct. When the null and alternative hypotheses are of the form: H_0: θ = θ_0 versus H_a: θ ≠ θ_0, the correct approach is to use a two-tailed test.

Q-84. Solution: C.

The appropriate test statistic is a z-statistic because the sample comes from a normal distributed population with known variance. A z-test does not use degrees of freedom. This test is two-sided at the 0.05 significance level, and the rejection point conditions are z > 1.960 and z < –1.960.

Q-85. Solution: A.

A is correct. This is a one-tailed hypothesis testing with a "greater than" alternative hypothesis. A squared standard deviation is being used to obtain a test of variance.

The hypotheses are $H_0: \sigma^2 \leq 0.25\%$ versus $H_a: \sigma^2 > 0.25\%$

B and C are incorrect as explained in choice A.

Q-86. Solution: B.

This is a two-tailed hypothesis testing because it concerns whether the population mean is zero.

$H_0: \mu = 0$ versus $H_a: \mu \neq 0$

With degrees of freedom (df) = n – 1 = 24 – 1 = 23, the rejection points are as follows:

Significance level	Rejection points for t-test
0.10	t < –1.714 and t > 1.714
0.05	t < –2.069 and t > 2.069
0.01	t < –2.807 and t > 2.807

Because the calculated test statistic is 2.41, the null hypothesis is thus rejected at the 0.05 and 0.10 levels of significance but not at 0.01.

Q-87. Solution: B.

A test statistic is determined by the following formula:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Q-88. Solution: C.

Describe the properties of Student's t-distribution and calculate and interpret its degrees of freedom.

C is correct. When the sample size is small, the Student's t-distribution is preferred if the variance is unknown.

Q-89. Solution: B.

The calculated test statistic is:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{-0.1452\sqrt{248-2}}{\sqrt{1-(-0.1452)^2}} = -2.30177$$

Because the absolute value of t = −2.30177 is greater than 1.96, the correlation coefficient is statistically significant.

Q-90. Solution: A.

The p-value is the smallest level of significance at which the null hypothesis can be rejected. The smaller the p-value, the stronger the evidence against the null hypothesis. Because the p-value (0.0567) exceeds the stated level of significance (0.05), the null hypothesis cannot be rejected. Therefore, A is correct and C is incorrect.

B is incorrect. A 5% confidence level does not allow the significance level to be increased beyond 5%.

Q-91. Solution: A.

If the p-value is less than the specified level of significance, the null hypothesis is rejected. Definetion of p-value: the smallest level of significance at which the null hypothesis can be reject.

Q-92. Solution: B.

Decrease the significance level can increase type II error and decrease type I error.

Q-93. Solution: A.

When the probability of Type II error increase, the probability of Type I error will decrease, which means that the significance level (α) will decrease. So the width of confidence interval will increase.

Q-94. Solution: B.

The power of a test is the probability of correctly rejecting the null hypothesis-that is, the probability of rejecting the null when it is false.

Q-95. Solution: C.

Test the selected samples are random or not, nonparametric test should be used.

Q-96. Solution: C.

C is correct. Nonparametric tests are primarily concerned with ranks, signs, or groups, and they are used when numerical parameters are not known or do not meet assumptions about distributions. Even if the underlying distribution is unknown, parametric tests can be used on numerical data if the sample is large.
A is incorrect because nonparametric tests can be used on grouped or counted data.
B is incorrect because nonparametric tests can be used on ranked data.

Q-97. Solution: B.

A contingency table is used to determine whether two characteristics of a group are independent.

Q-98. Solution: C.

The model does not assume that the dependent variable is uncorrelated with the residuals. It does assume that the independent variable is uncorrelated with the residuals.

Q-99. Solution: B.

The slope coefficient is best interpreted as the predicted change in the dependent variable for a 1-unit change in the independent variable. If the slope coefficient estimate is 10.0 and the independent variable changes by 1 unit, the dependent variable is expected to change by 10 units. The intercept term is best interpreted as the value of the dependent variable when the independent variable is equal to zero.

Q-100. Solution: B.

Note that this is a one-tailed test. The critical one-tailed 1 % t-statistic with 34 degrees of freedom is approximately 2.44. The calculated t-statistic for the slope coefficient is (1.1163 - 1) / 0.0624 = 1.86. Therefore, the slope coefficient is not statistically different from one at the 1 % significance level and Coldplay should fail to reject the null hypothesis.

Q-101. Solution: C.

C is correct. SST is equal to the sum of RSS and SSE: 0.0228 + 0.0024 = 0.0252. R^2 = RSS /SST= 0.0228 / 0.0252 = 0.905

Q-102. Solution: A.

A is correct. Because n = 36, and the degrees of freedom for the sum of squared errors (SSE) is n - 2 in simple linear regression, the degrees of freedom for SSE is 34, and the mean squared error is SSE / 34. The standard error of estimate (SEE) is equal to the square root of the mean squared error:

$$\text{SEE} = \sqrt{\frac{0.0024}{34}} = 0.008$$

Q-103. Solution:C.

C is correct.

Confidence interval estimate: $\hat{\boldsymbol{Y}} \pm \left(\boldsymbol{t_c} \times \boldsymbol{S_f}\right)$

In the 1% significance level of two tailed test, the value of t statistic is 2.728. So $\boldsymbol{t_c} = 2.728$. the standard error of the forecast($\boldsymbol{S_f}$) is 0.0469.

$$\boldsymbol{t_c} \times \boldsymbol{S_f} = 2.728 \times 0.0469 = 0.1279$$

1.2. 进阶题

Q-1. Solution: B.

After year 4, the 24-year fixed-rate mortgage has the largest payment.

The loan payments are summarized in the table below.

Mortgage type	Initial Payment ($)	Payment after adjustment($)
32-year fixed	12,389.92	12,389.92
24-year fixed	13,119.56	13,119.56
32-year adjustable	9,836.93	11,785.90

Payment on the 32-year fixed is calculated as:

N = 12×32= 384, I/Y = 6.5/12, PV = -2,000,000, FV = 0; CPT PMT = 12,389.92

Payment on the 24-year fixed is calculated as:

N = 12×24 = 288, I/Y = 6/12, PV = -2,000,000, FV = 0; CPT PMT = 13,119.56

Payment on the 32-year adjustable is calculated as:

Initial payment

N = 12×32 = 384, I/Y = 4.5/12, PV = -2,000,000; FV = 0; CPT PMT = 9,836.93

Balance at end of year 4:

N = 12×28 = 336, I/Y = 4.5/12, FV = 0, PMT = 9,836.93; CPT PV =-1,877,349.82

Payment after the end of year 4:

N = 336, I/Y = 6.2/12, PV = -1,877,349.82; FV = 0; CPT PMT =11,785.90

Q-2. Solution: B.

This scenario provides an example of a discrete random variable. The paired outcomes for the dice are indicated in the following table. The outcome of the dice summing to six is the most likely to occur of the three choices because it can occur in five different ways, whereas the summation to five and nine can occur in only four different ways.

Summed Outcome Paired Outcomes (Die 1, Die 2) Possible Combinations

5 (1, 4), (2, 3), (3, 2), and (4, 1) 4

6 (1, 5), (2, 4), (3, 3), (4, 2), and (5, 1) 5

9 (3, 6), (4, 5), (5, 4), and (6, 3) 4

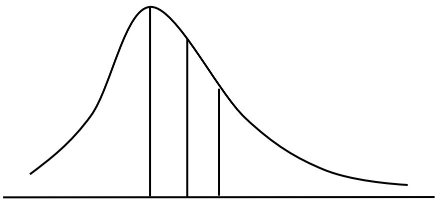
Q-3. Solution: C.

The t-statistic for the given information (normally distributed populations, population variances assumed equal) is calculated as:

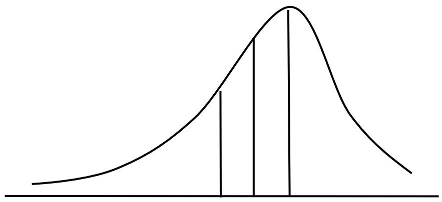
$$t = \frac{(210-195)-0}{(\frac{3377.13}{28}+\frac{3377.13}{21})^{0.5}} = 0.89$$

Q-4. Solution: A.

As shown in the following figure, the median is smaller than the mean for the positive skew. In contrast, the median is larger than the mean for the negative skew.



Mode < Median < Mean
Positive (right) skew



Mean < Median < Mode
Negative (left) skew

Therefore, if the two means equal, the median of the negative skew is larger than that of positive skew.

Q-5. Solution: C.

Identify the appropriate test statistic and interpret the results for a hypothesis test concerning 1) the variance of a normally distributed population, and 2) the equality of the variances of two normally distributed populations based on two independent random samples.

C is correct. The test statistic is the ratio of the variances, with the larger variance in the numerator. Here, the test statistic is 30 ÷ 5 = 6. The degrees of freedom are 5 by 5. Because it is a two-tailed test, the correct critical value at α = 5% is 7.15. And because the test statistic is less than the critical value, we cannot reject the null hypothesis.

Q-6. Solution: B.

The test statistic is : $\frac{d-\bar{\mu}_{d0}}{s_d/\sqrt{n}}$ where d is the mean difference, $\bar{\mu}_{d_0}$ is the hypothesized difference in the means, s_d is the sample standard deviation of differences, and n is the sample size. In this case, the test statistic equals : $(5.25-0)\big/(6.25/\sqrt{25}) = 4.20$. Because 4.20 > 1.711, the null hypothesis that the mean difference is zero is rejected.

Q-7. Solution: A.

Point estimates are less reliable than confidence interval estimates.

Using the *t*-distribution rather than the normal distribution is a more conservative approach to construct confidence intervals, and thus increase the reliability of the confidence interval.

Increasing the sample size can also increase the reliability of the confidence interval.

Q-8. **Solution: C.**

First, calculate the return difference each year:

Year	Portfolio A (%)	Portfolio B (%)	Differences (%)
2013	11	9	-2
2014	-10	4	14
2015	1	-3	-4
2016	8	12	4
2017	21	23	2
2018	2	-4	-6

And calculate the mean difference of returns using a financial calculator: $\bar{d} = \frac{1}{n}\sum d_i = 1.33\%$

Then, calculate the sample standard deviation and the standard error of the mean difference using a financial calculator:

$$S_d = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n-1}} = 7.23\%$$

$$S_{\bar{d}} = \frac{S_d}{\sqrt{n}} = \frac{7.23\%}{\sqrt{6}} = 2.95\%$$

Finally, calculate the *t*-statistic: $t = \frac{\bar{d} - 0}{S_{\bar{d}}} = 0.45$

Q-9. **Solution: A.**

Head and shoulders pattern: Price target = neckline – (head – neckline) = 23 – (33 – 23) = 13.

Q-10. **Solution: A.**

Bayes' formula: P(A|B)= [P(B|A)P(A)]/P(B)

Updated probability of event given the new information:

where

Updated probability of event given the new information: P(EPS below|Cut div);

Probability of the new information given event: P(Cut div|EPS below) = 86%;

Unconditional probability of the new information: P(Cut div) = P(Cut div|EPS exceed)P(EPS exceed)+P(Cut div|EPS equal)P(EPS equal)+P(Cut div|EPS below)P(EPS below)=23%*3%+56%*11%+21%*86%=0.69%+6.16%+18.06%=24.91%;

Prior probability of event: P(EPS below) = 21%.

Therefore, the probability of EPS falling below the consensus is updated as:

P(EPS below|Cut div) = [P(Cut div|EPS below)/P(Cut div)] × P(EPS below)

= (0.86/0.2491) × 0.21≈73%

Q-11. Solution: C.

The most appropriate test statistic for the difference between two population means (unequal and unknown population variances) is $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})^{0.5}}$.

Q-12. Solution: B.

Monte Carlo simulation provides a distribution of possible solutions to complex functions. The central tendency and the variance of the distribution of solutions give important clues to decision makers regarding expected results and risk.

Q-13. Solution: A.

There are two steps in standardizing a random variable X: Subtract the mean of X from X, and then divide that result by the standard deviation of X. This is represented by the following formula: $Z = (X - \mu)/\sigma$.

Q-14. Solution: A.

Any descriptive measure of a population characteristic is called a parameter.

Q-15. Solution: B.

Sampling errors will result in statistical error. A statistically significant result might not be economically meaningful after an analyst accounts for the risk, transaction costs, and applicable taxes.