◆ Section 1: The Time Value of Money

✓ Annuity, Bonds, Mortgages and Savings for college tuition

相關題型

1. An investor wants to have $140,000 available at the end of 12 years. This investor plans to make 12 equal year-end payments into an investment that is expected to earn an 8% annual rate of return. The required amount of each year-end payment is closet to:

a. $4,633
b. $6,831
c. $7,377
d. $10,802

N=12
I/Y=8 (P/Y=1)
PV=0
PMT=?
FV=140000
逐一輸入後，再按 CPT PMT 即可得，答案選 C

---

2. At an 8% rate of return, how much you have in your investment account on your 65th birthday so you can withdraw $30,000 on that birthday and on each of the next 19 birthdays?

a. $264,540
b. $288,120
c. $294,540
d. $318,108

N=20 (not 19, the first payment is due today and followed by 19 payments on each of your next 19 birthdays)
I/Y=8 (P/Y=1)
PV=?
PMT=30000
BGN Mode ON (2^{nd} + PMT =END, 2^{nd} + ENTER = BGN, 2^{nd} + CPT (=QUIT)
FV=0
逐一輸入或設定後，再按 CPT + PV 即可得，答案選 D

---

3. A bank lends you $3,000,000 to be paid down equally per month in the next seven years. If the bank charges 7% interest rate, the monthly installment payment will be closet to:

a. $66,667
b. $68,437
c. $45,015
d. $45,278

N=7*12 = 84
I/Y=7 (P/Y=12)
PV=3000000
PMT=?
FV=0
逐一輸入後，再按 CPT + PMT 即可得，答案選 D

---

4. Your father is preparing for your college tuition by creating an educating trust. You are 15 years old and you expect to enter a prestigious university in your state three years later. Your father expects your tuition along with living costs to be $80,000 per year for consecutive four years. The trustee assures you that that expected return to be 10% per year. How much does your father have to set aside to meet the goal?

a. $130,131
b. $209,578
c. $190,525
d. $155,622

N=4
I/Y=10 (P/Y=1)
PV=?
PMT=80000
BGN Mode ON (2^{nd} + PMT =END, 2^{nd} + ENTER = BGN, 2^{nd} + CPT (=QUIT)
FV=0
逐一輸入或設定後，再按 CPT + PV 求得 278948-

但是記得 278948-是你兒子進大學時你要準備的錢，但現在距離那個時間還有 3 年，所以現在你要準備的錢為
FV=278948-(將剛才算出來的 PV 值放在 FV 格中)
N=3
I/Y=10(P/Y=1)
PV=?
PMT=0
END Mode ON
逐一輸入或設定後，再按 CPT + PV 求得 209578，答案選 B

---

5. What is the value of a zero-coupon bond that pays $1,000 in five years if the market rate for this security is 7%?
a. $650.38
b. $708.92
c. $735.43
d. $751.37

先假設它的利息付出次數為一年兩次
N=5*2=10
I/Y=7 (P/Y=2)
PV=?
PMT=0(因為零息債券不付息)
FV=1000
逐一輸入後，再按 CPT + PV 即可得，答案選 B

---

6. Consider a 3 year coupon paying bond with par value of $1,000. The coupon rate is 5% paid annually and its yield to maturity is 6%. Compute the current value of the bond?
a. $973.3
b. $1,056.6
c. $935.4
d. $1,106.9

$PV = \frac{50}{1.06} + \frac{50}{1.06^2} + \frac{50}{1.06^3} + \frac{1000}{1.06^3}$ 利用 STO/RCL 的方法求出 PV
或是利用計算機第三行直接求得:
N=3
I/Y=6 (P/Y=1)
PV=?
PMT=50
FV=1000
逐一輸入後，再按 CPT + PV 即可得，答案選 A

---

7. If $5,000 is invested in a fund offering a rate of return of 12 percent per year, approximately how many years will it take for the investment to reach $10,000?
a. 4 years
b. 5 years
c. 6 years
d. 7 years

N= ?

I/Y=12 (P/Y=1)
PV=5000-
PMT=0
FV=10000
逐一輸入後，再按 CPT + N 即可得，N=6.11，選最接近的答案 C

✓ **PV of a series even cash flow and PV of Perpetuity**

$$PV = \frac{A}{r}$$

✓ **Uneven Cash Flow**

相關題型

1. An investor invests the following amounts at year end

Year 1	$ 1,000
Year 2	$ 4,500
Year 3	$ 2,000
Year 4	$ 4,000

If the invest earns a 10% interest rate, the accumulated amount at the end of Year 4 is closest to

a. $11,500
b. $12,976
c. $13,376
d. $14,274

1、 利用 2-1 的功能 CF 求之
2、 CF 按進去後第一個變數為 CF0，定義為當下的現金流量(在本題為 0)
3、 再按 4-1 下拉鍵後可得 C01，設 1000，F01 設 1(次數為 1 次)，依此類推，將 C02/F02，C03/F03，C04/F04 都設進去
4、 進入 3-2 的 NPV 功能，進入後，第一件事先輸入投資報酬率「I」是多少
5、 I 輸入後再按下拉鏈(4-1 鏈)，會看到 NPV，此時接 1-1 鏈(CPT)，即可得答案為 8863
6、 這一題題目要求 4 年後的終值是多少，所以:

N= 4
I/Y=10 (P/Y=1)
PV=8863-
PMT=0
FV=?
逐一輸入後，再按 CPT + FV 即可得 12976，故答案選 B

2. Assuming a discount rate of 10%, what is the present value of five $2,000 cash flows that occur at the end of year 1 to year 5, and an additional $10,000 cash flow at the end of year 5?

A. $17,779
B. $13,790
C. $16,209
D. $20,000

最後一期有兩個 cash flow 來源，分別是$2,000 及$10,000，其它期數的 cash flow 是 2000，算是 uneven cash flow 的問題。
進 CF 之功能(2-2)
CF0=0
C01=2000; F01=4
C02=12000; F02=1
進 NPV 之功能(3-2)
I=10
下拉至 NPV 後，接 CPT 求得，13790.7，答案選 B

但是，如果將最後一期的$10,000 看成是 FV，而其它各期數的$2,000 看成是 PMT，則這題又可用第三行解。這兩種方法解出來的值應該要一樣。

N=5
I/Y=10 (P/Y=1)
PV=?
PMT=2000
FV=10000
逐一輸入後，再按 CPT + PV 即可得，答案選 B

◆ Section 2: Statistical Concepts and Market Returns

✓ Sample, Population, and Parameter

✓ Central Tendency, Population mean and Sample mean

❒ Population Mean

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

❒ The Sample Mean

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

✓ Arithmetic vs. Geometric Mean Return

The *arithmetic mean return* is simply the sum of each period's asset return divided by the number of periods:

$$AR = \frac{\sum_{i=1}^{n} R_i}{n}$$

The *geometric mean Return* is used when calculating investment returns over multiple periods or to measure compound growth rates:

$$GR = \sqrt[n]{(1+R_1)\times(1+R_2)\times\Lambda\times(1+R_n)} - 1$$

In distinguishing between the two, the *geometric means is always less than or equal to the arithmetic mean*. The only time the arithmetic and geometric means are equal is when there is no variability in the observations (i.e., all observations are equal).

相關題型

1. A portfolio had an initial market value of $100,000. At the end of next two years, the market value was $110,000 and $88,000 respectively. The portfolio's arithmetic and geometric mean returns, respectively, excluding any dividends, were:

a. -12.00% -6.00%
b. -6.00% -6.19%
c. -5.00% -6.19%
d. -5.00% -12.00%

arithmetic mean return = $\frac{\left(\frac{110000-100000}{100000}\right)+\left(\frac{88000-110000}{110000}\right)}{2} = \frac{10\%+(-20\%)}{2} = -5\%$

Geometric mean return=
$\sqrt{1.1*0.8}-1$ = -0.0619=-6.19%

✓ Compounding and Effective Annual Rate

■ Effective Annual Rate (EAR)

$$EAR = \left(1 + \frac{r}{m}\right)^{m} - 1$$

其中 r 等於 Stated Annual Rate

重點: 如何利用 TI 計算機求取 EAR? 請進入 2nd+2

相關題型

1. Which of the following options is the most desirable for a borrower taking out a loan?

a. pay 8% compounded daily
b. pay 8.2% compounded monthly
c. pay 8.3% compounded quarterly
d. pay 8.5% compounded annually

EAR(a)= $(1+(8\%/365))^{365}-1=8.33\%$
EAR(b)= $(1+(8.2\%/12))^{12}-1=8.52\%$
EAR(c)= $(1+(8.3\%/4))^{4}-1=8.56\%$
EAR(d)= $(1+(8.5\%/1))^{1}-1=8.50\%$
四個 EAR 比較下發現其實 A 選項最小, 所以答案選 A
這一題如果你用 2nd+2 的功能來計算 EAR, 可能速度會更快

---

2. A bank pays a stated annual interest rate of 6 percent. What is the effective annual rate with monthly compounding?

a. 6.10%
b. 6.14%
c. 6.17%
d. 6.20%

EAR= $(1+(6\%/12))^{12}-1=6.17\%$, 答案選 c

■ Continuously Compounded Rates of Return

$$FV_n = PV(1+\frac{r}{m})^{mt}$$

If m becomes infinitely, then

$FV = PVe^{rt}$ (其中 t 代表期數，單位是年，r 就是我們的主角，年連續複利率)

$$\frac{FV}{PV} = e^{rt}$$

上述二邊取自然對數(natural log，也就是以 e 為底的對數。Natural log 的符號通稱 LN，計算機的 1-6 有這個功能)，可得:

$$LN\frac{FV}{PV} = rt$$

Example:
假如你買台積電的成本為 50 元，一年後為 65 元，請問投資台積電的連續複利率是多少?

$LN\frac{65}{50} = rt$ (其中 t=1)

r=0.2624 or 26.24%

Example:

Suppose an investment with initial amount of $10,000 will earn 8 percent compounded continuously for two years. We can compute the future value using above formula as follows:

$$FV_n = PVe^{r\times N} = 10000e^{0.08\times 2} = 10000(1.173511) = 11{,}735.11$$

✓ **Variance (Covariance) and Standard Deviation**

■ The *variance* is defined as the mean of the squared deviations from the mean.

Population variance: $\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$

Sample variance: $s^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$

The **Standard Deviation** is simply calculated as the squared root of the variance. Since the standard deviation and the mean are expressed in the same units,

相關題型

1. An analyst develops the following frequency distribution for the monthly returns of aggressive equity funds over a forth-month period

Month 1 = 2%

Month 2 = -4%

Month 3 = 1%

Month 4 = 5%

Assume that this distribution can present either the population or a sample of aggressive equity fund. The population and sample standard deviation for this distribution of monthly return is closest to ：

	Population Standard deviation	Sample Standard deviation
a.	2.55%	2.94%
b.	2.94%	2.55%
c.	3.24%	3.74%
d.	3.74%	3.24%

要會使用 TI 計算機的統計功能:

$\boxed{2^{nd}}$ + $\boxed{7}$= DATA; $\boxed{2^{nd}}$ + $\boxed{8}$=STAT

由於樣本變異數的分母是 n-1，母體變異數的分母是 n。所以，相同資料之變異數，如果假設是來自於樣本，會大於假設其本身就是母體所求得的變異數

✓ **Chebyshev's inequiality:**

Let K be any positive constant greater than 1. The proportion of the observations within K standard deviations of the mean is at least

$1 - (1/K^2)$ for all K>1

Number of Standard deviation	Proportion falling within that SD
1.5	56%
2	75%
2.5	84%
3	89%
4	94%

✓ **Coefficient of Variation (這個數值愈小愈好)**

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

相關題型

1. A given sample of three securities has annual returns of 5 percent, 10 percent, and 15 percent. The coefficient of variation of the sample is closest to:

a. 0.3

b. 0.4

c. 0.5

d. 0.6

先求出資料的平均數，很簡單，用目測法即可求得。再求出其標準差，這時可用公式或計算機求得，標準差按得 4.0825%，除以平均數 10%，即可得 B 的答案最接近

✓ **Sharpe Ratio (這個數值愈大愈好)**

$$\text{Sharpe ratio}=\frac{\overline{r_P}-\overline{r_f}}{\sigma_P}$$

✓ **Roy's Safety First Ratio (這個數值也是愈大愈好)**

$$\text{SFRatio}=\frac{[E(R_P)-R_L]}{\sigma_P}$$

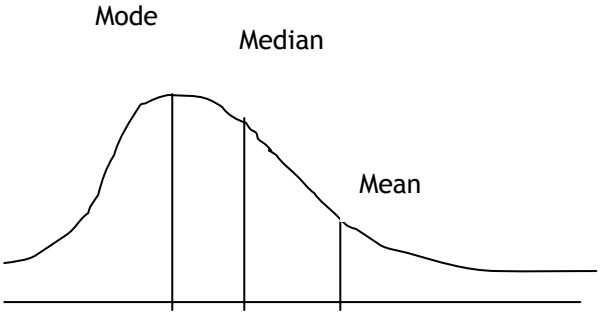
Example:

下列四個投資組合中，根據 Sharpe Ratio 及 Roy's Safety First Ratio，判斷那一個投資組合最好?

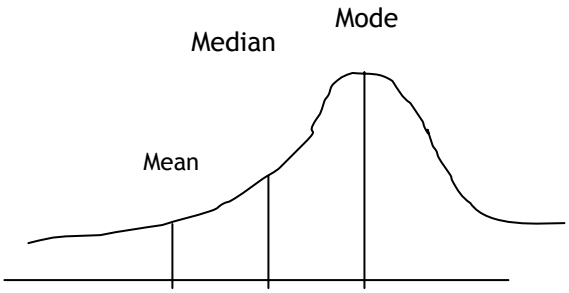
	R_p%	σ_p%	$R_{Threshold}$%	$R_{risk\ free}$%	**Sharpe Ratio**	**SFR**
投資組合 A	10	7	4	2	***1.143***	0.857
投資組合 B	11	8	4	2	1.125	0.875
投資組合 C	12	9	4	2	1.111	0.889
投資組合 D	13	10	4	2	1.100	***0.900***

✓ **Skewness**

Skewness represents the extent to which a distribution is not symmetrical. A *positively skewed* distribution has many outliers in the right tail, and the mean >median>mode. A *negatively skewed* distribution has many outliers in the left tail, and the mean<median<mode.



Distribution Skewed to the Right
(Positively Skewed)



Distribution Skewed to the Left
(Negatively Skewed)

相關題型

1. If the analyst concludes that the distribution of a large sample of return is positively skewed, which of the following relationship is correct？

a. Mean=median=mode
b. Mean>median>mode
c. Mean<median<mode
d. Mean>median<mode

答案選 B

---

2. You read GCC Trust company ten selective discretionary accounts' performance in 2003 as follows: 12%, 13%, 5%, 5%, 15%, 10%, 12%, 25%, 31%, 5%. Which of the following statements are true?

I. the distribution shows a positive skewness
II. the mean is 13.3% and the standard deviation is 8.65%
III. the median is 12%
IV. the mode is 5%

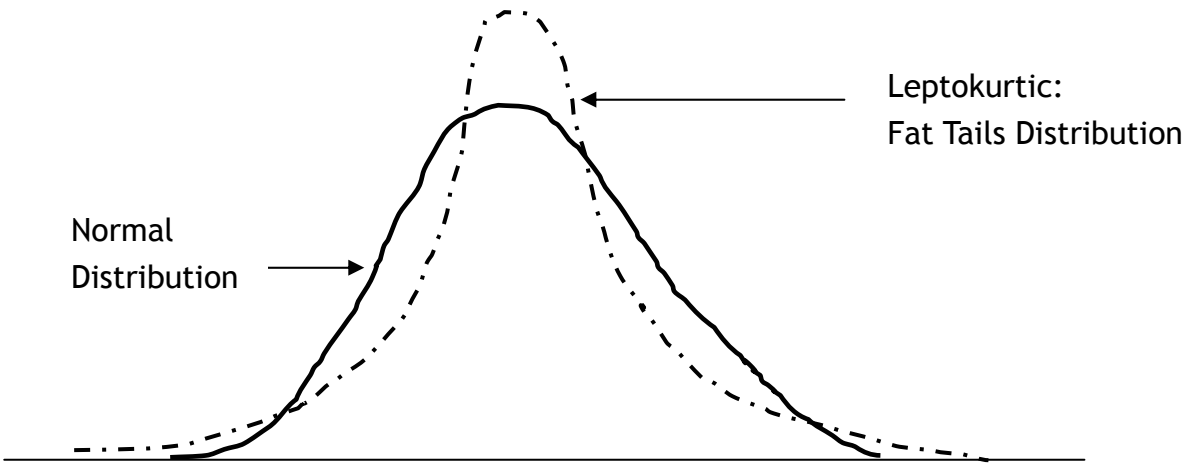
a. I and II only
b. I, III and IV only
c. II, III and IV only
d. I, II, III and IV

答案選 D
利用計算機求解即可

✓ **Kurtosis**

Kurtosis measures the peakedness of a distribution and affects the probability of extreme outcomes. Kurtosis is measured using deviations raised to the *fourth power*. A normal distribution has kurtosis equal to 3, so *excess kurtosis* is always measured relative to the number three. Positive values of excess kurtosis (kurtosis – 3) indicate a distribution that

is leptokurtic (fat tails, or more peaked), whereas negative values (kurtosis – 3) indicate a platykurtic distribution (thin tails, or less peaked).



相關題型

1. Which one is true regarding Kurtosis?
   - I. kurtosis is the statistical measure that tells when a distribution is more or less peaked than a normal distribution
   - II. a distribution that is more peaked than normal is call leptokurtic
   - III. leptokurtic means fatter tails and has excess kurtosis more than zero
   - IV. a normal distribution has excess kurtosis equal to 3

a. I only
b. I, II only
c. I, II, III only
d. I, IV only

答案要選 C，D 不能選，因為常態分配的 excess kurtosis 係數是 0，kurtosis 係數才是 3

◆ Section 3: Probability Concepts

■ **Three ways to obtain probabilities**

1. *Empirical probabilities*
2. *Priori probabilities*: based on logical analysis
3. *Subjective probabilities*: eg. Bayes' Theorem

✓ **Unconditional Probability**
✓ **Conditional Probability**
✓ **Joint Probability**

■ **Definition of Conditional Probability.** The conditional probability of B given that A

has occurred is equal to the joint probability of A and B divided by the probability of A (assumed not to equal 0)

$P(B|A) = P(AB)/P(A), \quad P(A) \neq 0$

- **Multiplication Rule for Probabilities**

$P(AB) = P(A) \times P(B/A), P(A) \neq 0$

- **Addition Rule for Probabilities.** Given events A and B, the probability that A or B occurs, or both occur, is equal to the probability that A occurs, plus the probability that B occurs, minus the probability that both A and B occur :

$P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$

- **Independent Events.**
  1. Two events A and B are independent if and only if P(AIB)=P(A) or, equivalently, P(BIA)=P(B)
  2. Multiplication Rule for Independent Events. When two events are independent, the joint probability of A and B equals the product of the individual probabilities of A and B: P(A∩B)=P(A)P(B)
  3. When A, B, C are independent events, then P(A∩B∩C)=P(A)P(B)P(C)

相關題型

1. At king investments, 80% of the portfolio mangers are male and 20% female. 95% of the female portfolio mangers have the CFA designation, while 80% of the males have the CFA designation. What is the probability that a portfolio manger would be female and not have a CFA designation?

a. 1%
b. 4%
c. 64%
d. 16%

Male/Female	Having CFA	Not Having CFA
0.80	0.8*0.80=0.64	0.8*0.20=0.16
0.20	0.2*0.95=0.19	0.2*0.05=***0.01***
1.00	0.83	0.17

The answer is 1%

2. The probability that portfolio A will yield positive annual returns is 0.30. The probability that portfolio B will yield positive annual returns is 0.60. The probability that

either portfolio A or B will yield a positive annual return is 0.70. What is the probability that both portfolios realize a positive return?

a. 0.40
b. 0.30
c. 0.20
d. 0.18

$P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$

0.7=0.3+0.6- $P(A \cap B)$

$P(A \cap B) = 0.2$

3. A dealer in a casino has rolled a five on a single die three times in a row. What is the probability of her rolling another five on the next roll assuming it is a fair die?

a. 0.500
b. 0.200
c. 0.001
d. 0.167

The answer is d

4. A sales representative has two leads. The first lead has a 40 percent chance of being successful, and the second lead has a 20 percent chance of being successful. The occurrence of success or failure with either lead is independent of the occurrence of success or failure with the other lead. The probability that the representative will be successful on exactly one lead is:

a. 0.08
b. 0.12
c. 0.32
d. 0.44

Lead 1 / Lead 2	0.20	0.80
0.40	0.4*0.2=0.08	0.4*0.8=**0.32**
0.60	0.6*0.2=**0.12**	0.6*0.8=0.48
1.00	0.20	0.80

0.12+0.32=0.44

The answer is d.

5. A fair die is rolled. Event A is the event that the number on the top face is a 1 or a 2. Event B is the event that the number on the top face is a 5 or a 6. Events A and B are:

a. Mutually exclusive and independent
b. Mutually exclusive, but not independent
c. Not mutually exclusive, but independent
d. Neither mutually exclusive nor independent

要檢驗是不是獨立事件的方法，可以用

P(A∩B)是否等於 P(A)P(B)

P(12∩34)=0；而 P(12)P(34)=1/3*1/3=1/9，既然兩者不相等，因此不是獨立事件

，但兩者卻是互斥事件，因此出現 1 或 2 時就不可能出現 3 或 4

The answer is b

6. Suppose that A and B are independent events. The probability of A is 0.6 and the probability of B is 0.8. What is the probability either A or B will occur?

a. 0.14
b. 0.48
c. 0.92
d. 0.90

A or not A / B or not B	0.80	0.20
0.60	0.6*0.8=0.48	0.6*0.2=**0.12**
0.40	0.4*0.8=**0.32**	0.4*0.2=0.08
1.00	0.80	0.20

把反黑的三個部分加起來，其實就是聯集的觀念。

注意，在英文的用法上，either A or B will occur 指的是 P(A∩B)，所以也包括 A 和 B 同時發生，所以左上方那一塊也要考慮進來

The answer is d

7. A consulting firm currently has two proposals outstanding. The probability that they will win proposal A is 0.40, the probability that they will win proposal B is 0.70, and the outcomes of proposals A and B are independent. The probability that they will win at least one of the two proposals is:

a. 0.18
b. 0.28
c. 0.54
d. 0.82

Proposal #1 / #2	0.70	0.30
0.40	0.4*0.7=0.28	0.4*0.3=0.12
0.60	0.6*0.7=0.42	0.6*0.3=0.18
1.00	0.70	0.30

把反黑的三個部分加起來，其實就是聯集的觀念。

The Answer is d

- ✓ **Total Probability Rule**

  1. $P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \ldots\ldots\ldots + P(B|A_n)P(A_n)$

  Where $A_1, A_2, \ldots\ldots, A_n$ are mutually exclusive and exhaustive scenarios or events

相關題型

1. Suppose you are modeling GNP, and you believe that the probability that GNP will expand if interest rates fall is 70%; if interest rates stay constant, you believe that there is a 29% chance of expanding GNP; if interest rates increase, you believe that there is a 1% chance of GNP expanding. You think that the likelihood of interest rates failing is 60%; of staying the same is 30%; of increasing is 10%. What is the unconditional probability of GNP expanding?

a. 50.0%
b. 50.8%
c. 50.9%
d. 49.8%

GNP expand			利率走勢
0.6*0.70=0.42			Fall =0.6
0.3*0.29=0.087			Stay =0.3
0.1*0.01=0.001			Increasing = 0.1
0.508(下方的邊際機率)			1.0

這一題是考 Total Probability Rule 的觀念。將各別邊際機率乘上各別的條件機率，再相加，會成為下方的邊際機率

The Answer is b

2. An analyst constructs the following probability table for the market and Company's M's stock:

State of the Economy	Probability of the Economic State	Stock Performance	Conditional Probability of Stock Performance
Good	0.4	Good Neutral Poor	0.5 0.25 0.25
Neutral	0.3	Good Neutral Poor	0.5 0.3 0.2
Poor	0.3	Good Neutral Poor	0.4 0.4 0.2

The total probability of good performance for Company M's stock is:

a. 0.10
b. 0.20
c. 0.40
d. 0.47

Stock Good	Stock Neutral	Stock Poor	
0.4*0.5=0.20			Econ Good=0.4
0.3*0.5=0.15			Econ Neutral=0.3
0.3*0.4=0.12			Econ Poor=0.3
0.47(下方的邊際機率)			

這一題是考 Total Probability Rule 的觀念。將各別邊際機率乘上各別的條件機率，再相加，會成為下方的邊際機率

The Answer is d

✓ **Bayes Theorem**

相關題型

1. You have developed a set of criteria for evaluating distressed credits. Firms that do not receive a passing score are classed as likely to go bankrupt within 12 months. You gather the following information when validating the criteria:

- Forty percent of the companies to which the test is administered will go bankrupt within 12 months
- Fifty-five percent of the companies to which the test is administered pass it
- The probability that a firm will pass the test (and be classed as a 12-month survivor), given that it will subsequently survive 12 months, is 0.85

Using Bayes' theory, calculate the probability that a firm is a survivor, given that it passes

the test?
a. 0.93
b. 0.80
c. 0.20
d. 0.85

Bankrupt	surviving	通過測驗/沒有過
	0.6*0.85(條件機率)=0.51	0.55
	0.09	0.45(題目有給)
0.40	0.60	1.00

根據上表分析，答案要我們求的是 P(surviving I 通過測驗) = 0.51/0.55

The answer is a.

2. A study found out that, on average, 10% of a pharmaceutical companies' drugs that are placed on the market sell more than $500 million their first year. If a drug sells more than $500 million on its first year on the market, its probability of selling more than $500 million on the second year goes up to 90%. If, on the other hand, the drug sells under $500 million during its first year on the market, its probability of selling more than $500 million the second year is only 30%. If a drug sold $750 million the second year after its launch, what is the probability that it sold more than $500 million the first year after its launch?

a. 0.500
b. 0.250
c. 0.125
d. 0.750

第二年營收>5 億	第一年營收<5 億	>5 億/<5 億
0.1***0.9**=0.09		0.1(題目有給)
0.9***0.3**=0.27		0.9
0.36	0.60	1.00

根據上表分析，答案要我們求的是 P(第一年營收>5 億 I 第二年營收>5 億) = 0.09/0.36

The answer is b.

- ✓ **Expected Value and Variance**

相關題型

- ✓ **Covariance and correlation**

■ **Definition of Covariance**

Given two random variables Ri and Rj, the covariance between Ri and Rj is

Cov (x_i, y_j) = $\sum p_i(x_i - \bar{x})(y_i - \bar{y})$ or $\frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$

相關題型

1. Which of the following are *true* regarding covariance?

I. covariance of returns on two assets is 0 when the returns are linearly unrelated

II. covariance will be negative if, when the return on one asset is below expected value, the other return will also be below its expected value

III. the covariance of returns of an asset with itself is the asset's return variance

IV. covariance is probability weighted average of the cross product of the deviation of each random variable from its expected value

a. I, II, and III only

b. I, III and IV only

c. I, II and III only

d. II, III and IV only

The answer is b

■ **Definition of Correlation**

The correlation between two random variables, Ri and Rj, is defined as

$$\rho(R_i, R_j) = \frac{Cov(R_i, R_j)}{\sigma(R_i)\sigma(R_j)}$$

✓ **Portfolio Expected Return and Variance**

✓ **Probability Counting Rules**

■ Combination formula (The Binomial Formula). The number of ways that we can choose r objects from a total of n objects, where the order in which the r objects is listed does not matter, is

$${}_nC_r = \frac{n!}{(n-r)! \times r!}$$

■ Permutation Formula. The number of ways that we can choose r objects from a total of n objects, where the order in which the r objects is listed does matter, is

$${}_nP_r = \frac{n!}{(n-r)!}$$

相關題型

1. How many three-digit numbers can you form using the digits 1, 3, 5, 7, 9 without repeating a digit within the number to be formed?

a. 20
b. 30
c. 60
d. 125

The Answer is c.

◆ Section 4: Common Probability Distributions

✓ Binomial Models

- You can compute the expected probability of *successes* in a given set of *n* trials using the following formula:

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

- **Mean and Variance of Binomial Random Variables**
  (二項式分配重覆實驗 n 次後，成功的期望值及變異數)

Mean	Variance
np	***np(1-p)***

相關題型

1. For a certain class of junk bonds, the probability of a default in a given year is 0.2. Whether one bond defaults is independent of whether another bond defaults. For a portfolio of five of these junk bonds, what is the probability that zero or one bond of the five defaults in the year ahead?

a. 0.0819
b. 0.4096
c. 0.7373
d. 0.5904

$C_0^5 0.2^0 0.8^5 + C_1^5 0.2^1 0.8^4$

The answer is c

2. An insurance agent has appointments with four prospective clients tomorrow. From past experience the agent knows that the probability of making a sale on any appointment is 1 out of 5. Using the rules of probability, what is the likelihood that the agent will sell a policy to 3 of the 4 prospective clients?

a. 0.0256
b. 0.0064
c. 0.4101
d. 0.4096

$C_3^4 0.2^3 0.8^1$

The answer is a.

3. A coin is tossed five times. What is the probability of obtaining exactly three heads?

a. 0.1250
b. 0.1875
c. 0.2500
d. 0.3125

$C_3^5 0.5^3 0.5^2$

The answer is d

4. Recent history has shown that the movements of more than 3% in the Dow Jone's Industrial Average occur in 2% of the trading days. What is the probability that within the next 30 trading days there will be one or more days in which the average will move more than 3%?

a. 0.4545
b. 0.2876
c. 0.1524
d. 0.0836

1- $C_0^{30} \times 0.02^0 \times 0.98^{30}$

The answer is a

✓ **Normal Distributions**

❒ The normal distribution is completely described by two parameters—its mean, μ, and variance, σ^2. We indicate this as follows

X ~ N(μ, σ^2)

❒ The normal distribution has a skewness of 0 (that is, it is symmetric). The normal distribution has a kurtosis (measure of peakedness) of 3; its excess kurtosis (kurtosis - 3) equals 0. As a consequence of symmetry, the mean, median, and the mode are all equal for a normal random variable

❒ A linear combination of two or more normal random variables is also normally distributed

- 約 68%的觀察值會落在母體平均值上下一個標準差的範圍內
- 約 95%的觀察值會落在母體平均值上下二個標準差的範圍內
- 約 99%的觀察值會落在母體平均值上下三個標準差的範圍內

- 90%的觀察值會落在母體平均數±1.645 個標準差內
- 95%的觀察值會落在母體平均數±1.96 個標準差內
- 98%的觀察值會落在母體平均數±2.33 個標準差內
- 99%的觀察值會落在母體平均數±2.58 個標準差內

✓ **Standard Normal Distribution**

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$

相關題型

1. A manager forecasts a bond portfolio return of 10% and estimates a standard deviation of annual returns of 4%. Assuming normality assumptions for returns, which of the following statements is **FALSE**

a. The 90 percent confidence interval will be 3.2% to 17.2%
b. The 95 percent confidence interval will be 2.16% to 17.84%
c. The 99 percent confidence interval will be -0.32% to 20.32%
d. A one standard deviation confidence interval for portfolio returns will run from 6% to 14%

A 90%覆蓋了從平均數算起正負 1.645 標準差的觀察值　$10\% \pm 1.645 \times 4\%$, 所以範圍為 3.42%到 16.58%

B. 95%覆蓋了從平均數算起正負 1.96 標準差的觀察值

C. 99%覆蓋了從平均數算起正負 2.58 標準差的觀察值

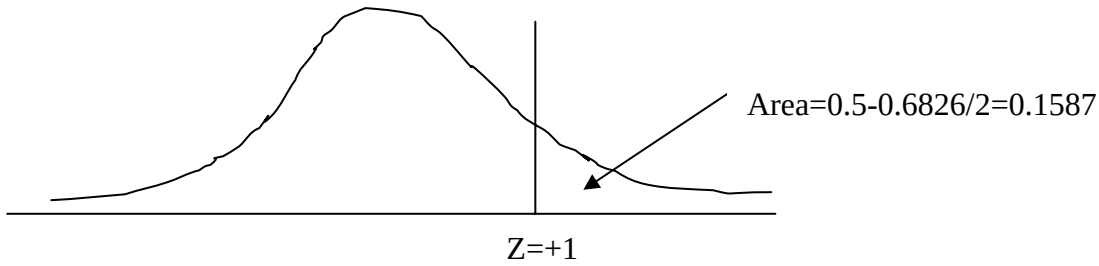
The answer is a

---

2. A food retailer has determined that the mean household income of her customers is $47,500 with a standard deviation of $12,500. She is trying to justify carrying a line of luxury food items that would appeal to households with incomes greater than $60,000. Based on her information and assuming that household incomes are normally distributed, what percentage of households in her customer base has incomes of $60,000 or more?

a. 34.13%
b. 5.00%
c. 2.50%
d. 15.87%

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{x - \mu}{\sigma} = \frac{60000 - 47500}{12500} = 1$$



The answer is d.

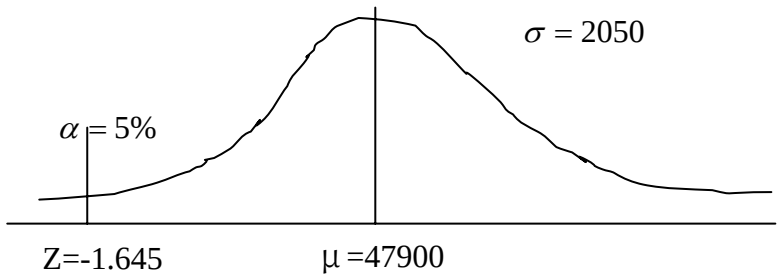
3. Suppose a tire manufacturer wants to set a mileage guarantee on its new XB 70 tire. Life test revealed that the mean mileage is 47,900 and the standard deviation of the normally distributed distribution of mileage is 2,050 miles. The manufacturer wants to set the guaranteed mileage so that no more than 5% of the tires will have to be replaced. What guaranteed mileage should the manufacturer announce?

a. 39,129
b. 44,528
c. 40,922
d. 49,621

在常態分配中，只是知道機率值，就可以求得對應的 z 值，而只是知到 z 值，就可以求出 x 值。

這一題問什麼呢?要先看的懂題目。輪胎商要保證一個哩數，讓客戶向他換新胎的機率只有 5%。如果輪胎磨損程度或耐用度是常態分配的話，則我就可以定一個公哩數(平均數)，讓 5% 的消費者會來跟我換胎。

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}}; -1.645 = \frac{x - 47900}{2050}; x = 44528$$



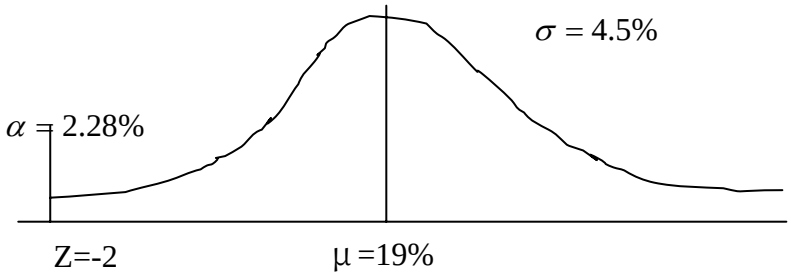
The answer is b.

4. A client will move his investment account unless you earn at least a 10% rate of return on the account. The rate of return for the portfolio you have chosen has a normal probability distribution with an expected return of 19% and a standard deviation of 4.5%. What is the probability you will keep this account?

a. 47.5%
b. 95.4%
c. 97.7%
d. 99.7%

$$z = \frac{10 - 19}{4.5} = -2$$

我們知道在標準常態分配中，z 正負 2 可以 cover 95.44%的機率，所以一邊尾的機率是 (1-0.9544)/2=0.0228



所以右邊的機率就是 1- 0.0228= 0.9772，投資經理人要達到 10%或以上的報酬率，有 97.72%的機率

5. If the annual returns of portfolio A follow a normal distribution with a mean-return of 17% and a standard deviation of 4%. What is the probability that the annual return of portfolio A will be between 16% and 20%?
(use the following selected areas under the normal distribution between 0 and z)

Positive Z value and area in the right hand side of Normal distribution

	0.03	**0.04**	**0.05**	**0.06**	**0.07**
0.20	0.0910	0.0948	0.0987	0.1026	0.1064
0.30	0.1293	0.1331	0.1368	0.1406	0.1443
0.70	0.2673	0.2704	0.2734	0.2764	0.2794
0.80	0.2967	0.2995	0.3023	0.3051	0.3078
0.90	0.3238	0.3264	0.3289	0.3315	0.3340

a. 0.3721
b. 0.3908
c. 0.4134
d. 0.4467

要先找出 16%和 17%之間，以及 17% 到 20%所對應的面積是多少。

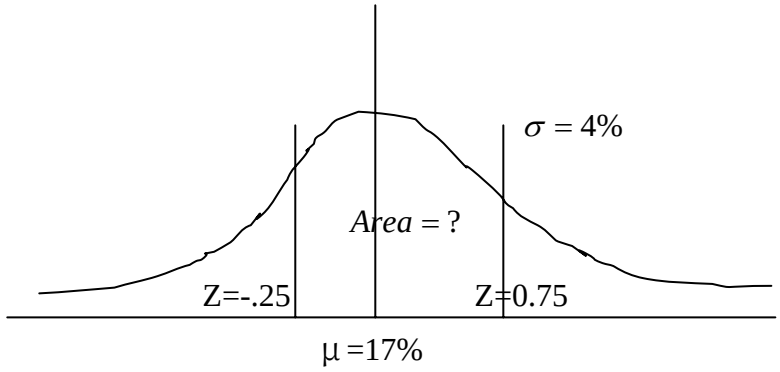
$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{16-17}{4} = -0.25$$

查表後所對應的面積是 0.0987 (z 值為-0.25 的面積等於 z 值為 0.25 的面積)

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{20-17}{4} = 0.75$$

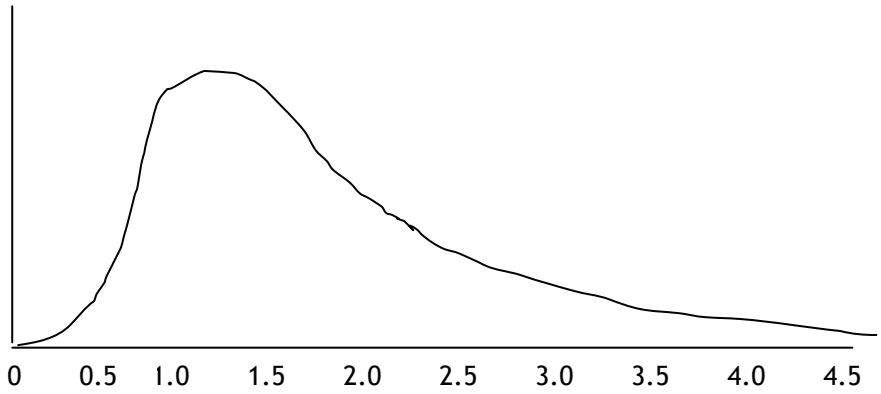
查表後所對應的面積是 0.2734 (直接查 z 值為 0.75 的面積)

加總後總面積是 0.3721



The answer is a.

✓ **Lognormal and normal distributions**



- Mean (μ_L) of a lognormal random variable = exp($\mu + 0.5\sigma^2$)
- Variance (σ_L^2) of a lognormal random variable =exp($2\mu + \sigma^2$) $\times \left[\exp(\sigma^2) - 1\right]$

✓ **Applications and limitations of Monte Carlo simulation and historical simulation**

- Monte Carlo simulation involves the use of a computer to find approximate solutions to complex problems
- Monte Carlo simulation involves identifying the risk factors associated with a problem and specifying probability distributions for them. Repeated random sampling from a probability distribution or distributions is used to simulate the risk factors
- Monte Carlo simulation is a complement to analytical methods
- Historical simulation is an established alternative to Monte Carlo simulation that involves repeated sampling from a historical data series
- Historical simulation can only reflect risks represented in the sample. Compared to Monte Carlo simulation, historical simulation does not lend itself to "what-if" analysis

◆ Section 5: Sampling and Estimation

✓ **Random Sampling**

✓ **The Central Limit Theorem**

定義: Given a population described by any probability distribution having mean μ and finite variance $\sigma^{2,}$ the sampling distribution of sample mean $\bar{x}$ computed from samples of size *n* from this population will be approximately normal with mean μ and variance σ^2/n when the sample size *n* is large.

The formula used depends on whether the population variance is known or unknown but

the equation is not substantially different.

known population variance: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

unknown population variance: $S_{\bar{x}} = \frac{s}{\sqrt{n}}$

❒ The desirable properties of an estimator are ***unbiasedness*** (the expected value of the estimate equals the population parameter), ***efficiency*** (the estimator has the smallest variance), and ***consistency*** (the estimator gets better as we use more data)

相關題型

1. Which of the following statements about the sampling and estimation is FALSE?

a. Sampling error is the difference between the observed value of a statistic and the quantity it is intended to estimate

b. A simple random sample is a sample obtained in such a way that each element of the population has an equal probability of being selected

c. The central limit theorem states that the sample mean for large sample size will be not distributed normally regardless of the distribution of the underlying population

d. 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

The answer is c

✓ **Confidence Intervals**

❒ Construction of Confidence Intervals. A (1-α)% confidence interval for a parameter has the following structure:

Point estimate ± Reliability Factor X Standard Error

Where

Point Estimate = a point estimate of the parameter (a value of a sample statistic)

Reliability factor = a number based on the assumed distribution of the point estimate and the degree of confidence (1-α)% for the confidence interval

Standard Error = the standard error of the sample statistic providing the point estimate

With a known variance, the formula for a confidence interval is

$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ or more commonly when we don't know the population's variance, we use $\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$

相關題型

1. An investor analyst takes a random sample of 100 aggressive equity fund and calculates the average beta as 1.7. The sample has a standard deviation of 0.4. Using 95% confidence interval and a z-statistic, which of the following is true：

a. 1.58 to 1.82 includes the mean of sample betas
b. 1.597 to 1.803 includes the mean of sample betas
c. 1.622 to 1.778 includes the mean of population betas
d. 1.634 to 1.766 includes the mean of population betas

$$\bar{x} \pm z_{0.025} \frac{s}{\sqrt{n}} = 1.7 \pm 1.96 \frac{0.4}{\sqrt{100}} = 1.7 \pm 1.96 \times 0.04$$

The confidence interval is between 1.6216 and 1.7784. We are 95% confident that the true beta will lie in this interval.

The answer is c

✓ **t-distributions**

1. When *compared to the normal distribution*, the *t*-distribution is flatter with more area under the tails (i.e., it has fatter tails).
2. *When working with confidence intervals,* due to the relatively fatter tails of the *t*-distribution, confidence intervals constructed using *t* reliability factors ($t_{\alpha/2}$) will be more conservative (wider) than those constructed using *z* reliability factors ($z_{\alpha/2}$).

Basis of Computing Reliability Factors

When sampling from a:	*Test Statistic*	
	Small Sample (n<30)	Large Sample (n≧30)
Normal distribution with *known* variance	***z-statistic***	***z-statistic***
Normal distribution with *unknown* variance	***t-statistic***	***t-statistic****
Nonnormal distribution with *known* variance	not available	***z-statistic***
Nonnormal distribution with *unknown* variance	not available	***t-statistic****

* Use of z also acceptable

(1- α)% confidence interval for the population mean μ is given by

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

相關題型

1. A money manager needs to estimate his clients' fund flow information for the next six months. He called a random sample of 10 clients. The manager then computes the change in cash flow for each client sampled as a percentage change in total funds placed with the manager. A positive percentage change indicates a net cash inflow to the clients account, and a negative percentage change indicates a net cash outflow from the client's account.

The money manager computes an average of 5.5 percent. The standard deviation of the observations in the sample is 10 percent. The manager assumes the population is normal. Which of the following is true for a 95 percent confidence interval for fund inflows and outflows over the next six months for the money manager's existing clients.

a. -1.65% to 12.65%
b. -0.69% to 11.69%
c. -17.12% to 28.12%
d. -14.10% to 25.10%

Selective t distribution

Df	P=0.10	P=0.05	P=0.025	P=0.01	P=0.005
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.106

Because the population is unknown and the sample size is small, the manager must use t-statistic to calculate the confidence interval. Based on the sample size of 10, df = n-1 = 9. For a 95 percent confidence interval, he needs to use the value of $t_{0.025}$ for df = 9. According to the appendix for t distribution table, this value is 2.262. Thus , a 95 percent confidence interval for the population mean is

$$\bar{x} \pm t_{0.025} \frac{s}{\sqrt{n}} = 5.5\% \pm 2.262 \frac{10\%}{\sqrt{10}} = 5.5\% \pm 2.262 \times 3.16\% = 5.5\% \pm 7.15\%$$

The confidence interval for the population mean spans –1.65 percent to +12.65 percent. The answer is a.

✓ **Sampling Selecting Bias**

Types of Bias	Description	Example
Data-snooping/data-mining Bias	Data mining is the practice of finding forecasting models by extensive searching through databases for patterns or trading rules.	"Fractured Four" VS "Foolish Four"
Sample Selection Bias / Survivorship Bias	Data availability leads to certain assets being excluded from the analysis	Low price to book stocks perform well in following years
Look-ahead Bias	Use information that was not available on the test date	In price to book ratio, book value per share is not known ahead
Time-period Bias	Based on a time period that may make the results time-period specific	Use an inadequate time periods to justify value stocks outperforming growth stocks

相關題型

1. Suppose you have found a statistically significant relationship that can be exploited for profit by means of trading stock. Your analysis indicates that if you buy a stock immediately after the close of a quarter in which the stock's EPS increases more than expected, you can earn above-market returns. Unfortunately, your trading rule suffers from:

a. time-period bias
b. sample selection bias
c. survivorship bias
d. look-ahead bias

The answer is d

2. Which of the following are likely to be subject to data mining?
   I. A high number of variables used by the researcher
   II. A trading rule that also works on out-of-sample data
   III. Research based on minor modifications to previously successful trading rules
   IV. Research based on price to book value ratios

a. I and II only
b. I and III only
c. I, II and III only
d. I, II, III and IV

The answer is b

◆ Section 6: Hypothesis Testing

✓ Steps in Hypothesis testing.

1. Stating the hypothesis

❒ 是雙尾檢定還是單尾檢定，要檢定的母體值是多少

2. Identifying the test statistic and its probability distribution

❒ 是常態分配還是 t 分配

相關題型

1. Given the following set of hypothesis:
   Ho: μ=100
   Ha: μ≠100

And the following random sample of n=9 values: *100, 105, 93, 100, 95, 88, 86, 93, 95*

The appropriate test statistic and its value are:

a. z=-0.83
b. t=0.83
c. t=-2.5
d. t=2.5

先利用 TI 計算機求出平均值和標準差，再將樣本之標準誤(standard error of the sample statistic)求出，樣本的標準誤=樣本觀察值的標準差除以觀察個數的開根號。

Test statistic =

$$\frac{Sample\ stastistic - Value\ of\ the\ population\ parameter\ under Ho}{S\tan dard\ error of\ the\ sample\ statistic} = \frac{95-100}{6/3} = -2.5$$

所以這一題的答案要選 c

3. Specifying the significance level

✓ Type I and Type II Errors

There are four possible outcomes when we test an null hypothesis:

1. We reject a false null hypothesis. This is a correct action

2. we reject a true null hypothesis. This is a **Type I error**
3. we do not reject a false null hypothesis. This is a **Type II error**
4. we do not reject a true null hypothesis. This is a correct action

- ***Type I error****:* the rejection of the null hypothesis when it is actually true.
- ***Type II error****:* the failure to reject the null hypothesis when it is actually false.

Type I and Type II Errors in Hypothesis Testing		
	True Condition	
Decision	*H_0 is true*	*H_0 is false*
Do not reject H_0	Correct decision	Incorrect decision ***Type II error***
Reject H_0	Incorrect decision ***Type I error*** Significance level, α, =P(Type I error)	Correct decision Power of the test =1—P(Type II error)

相關題型

1. If the significance level of a test is the probability of incorrectly rejecting the null hypothesis, then the "power of a test" is defined as which of the following?
a. the probability of a Type II error
b. the probability of a Type I error
c. the probability of incorrectly accepting the alternative hypothesis
d. the probability of correctly rejecting the null when it is false

power of the test 指的就是我們不去犯 type II Error 的機率，所以這題的答案應該選 d

2. Suppose we set the criterion for the rejection of the mull that is extremely lax, assuming us that the null will not be rejected. Then, which of the following is/are true?
I. the probability of a type I error is nearly zero
II. the probability of Type II error is nearly zero
III. the significance level of the test is nearly 1
a. II and III only
b. I only
c. I and III only
d. II only

我們所設的顯著水準愈低，我們就要用愈嚴格的證據來拒絕虛無假設，也就是我們需要計算出更極端的觀察值。因此，選項 I 是對的，選項 II 是錯的，選項 III 也是錯的。
The Answer is b.

---

3. What is the probability of making a Type II error if the null hypothesis is actually true?

a. 1-alpha
b. 0
c. 0.05
d. alpha

The Answer is b. 因為如果我們虛無假設是對的，則不可能犯下 type II error

4. **Stating the decision rule**
   - ❒ 如果 z 值或 t 值大於、小於，或大於小於臨界值(critical value)，我們就拒絕基本假設，反之，我們就不能拒絕

5. **Collecting the data and performing the calculations**
   - ❒ 計算 z 值或 t 值

6. **Making the statistical decision**
   - ❒ 將 5 所計算的 z 值或 t 值和 3 所定出的臨界 z 值或臨界 t 值利用 4 之基準做比較，以拒絕或接受基本假設

7. **Making the economic or investment decision**
   - ❒ 拒絕或接受基本假設所代表之經濟或投資意含

假設檢定綜合題型

Use the following data to answer the following four questions

Austin Roberts wants to test if the mean price of houses in the area is greater than $145,000 A random sample of 36 houses in the area has a mean price of $149,750. The population standard deviation is 24,000. Use the data for the next four questions, assuming a 1% level of significance

1. The alternative hypothesis will be that the population mean is
a. Less than $145,000
b. Not equal to $145,000
c. Greater than or equal to $145,000
d. Greater than $145,000

The Answer is d

2. The value of the calculated test statistic is
a. Z=0.67
b. Z=1.19
c. Z=4.00

d. Z=8.13

$$t_{calculated} = \frac{149750 - 145000}{\frac{24000}{6}} = \frac{4750}{4000} = +1.1875$$

3. The critical value of the z-statistic is

a. Z=±2.58

b. Z=+2.33

c. Z=-2.33

d. Z= ±2.33

因為是單尾檢定，所以 1%右尾所對應的 Z 值是+2.33

The answer is b

4. At 1% level of significance Robert should ：

a. Accept the null hypothesis

b. Fail to reject the null hypothesis

c. Reject the null hypothesis

d. Neither reject nor fail to reject the null hypothesis

因為+2.33 大於+1.185，所以我們沒有辦法拒絕虛無假設。

The answer is b

5. A pitching machine is calibrated to deliver a fastball at a speed of 98 miles per hour. Every day, a technician samples the speed of 25 fastballs in order to determine if the machine needs adjustment. Today, the sample showed a mean speed of 99 miles per hour. Assume that the population's variance is 3.0625 miles per hour and a significance level of 5%, what is the z-value in relation to the critical value? The:

a. critical value exceeds the z-value by 1.3

b. critical value exceeds the z-value by 0.7

c. z-value exceeds the critical value by 0.9

d. z-value exceeds the critical value by 1.5

$$t = \frac{99 - 98}{\frac{1.75}{\sqrt{25}}} = 2.8571$$

這題的題意是要我們做一個雙尾檢定。這題的顯著水準是 5% = 1 減掉信賴水準 (1-α)=1 - 0.95 = 0.05。在常態分配中，雙尾的顯著水準所對應的臨界值是+1.96 或-1.96。因此這一題

的觀察值比臨界值多出: 2.8571 - 1.96 = 0.8971，約等於 0.9,

The Answer is c

6. A portfolio manager's average return over the last five years is 10%, but a quantitative analyst states that in doing a hypothesis test it cannot be determined with 95% confidence that the return is actually different than zero because the t-statistic is only 1.5. How might the manger show his results are better than zero?

a. test a longer time period (have more observations)
b. Use an F test
c. Invest in stocks with lower standard deviation
d. Use a multiple regression model

這題是觀念問題，答案要選 a。觀察值不夠大的原因有很多，可能是樣本平均數和母體平均數的距離很小，或是樣本觀察數彼此之間的標準差太大，也可能是樣本中觀察個數值不夠多。增加觀察值，會讓分母變小，整個觀察值就會拉大，就變的比較容易超過臨界值，進而讓我們去拒絕虛無假設。

The Answer is a

✓ ***p-Values***

The *p-value* is the probability of obtaining a critical value that is the same as the computed test statistic, assuming the null hypothesis is true.

There are two decision rules for the *p*-value approach to hypothesis testing:

- Reject H_0 if the *p*-value is less than the significance level of the hypothesis test.
- Do not reject H_0 if the *p*-value is greater than the significance level.

相關題型

1. If the p-value for a statistic is less than the significance level, the decision rule is to:

a. reject the null hypothesis
b. accept the null hypothesis
c. reject the alternative hypothesis
d. conclude that the statistic does not differ significantly from the null hypothesis

The answer is a. p 值是假設虛無假設是正確的，則觀察值所應的機率。換句話說，如果觀察值是正確的，他發生的機率是 p 這麼多。如果 p 很小，我可以說:如果虛無假設是正確的，我不太可能觀察到這個數值。換句話說，我就比較能夠去拒絕虛無假設。

The answer is a

2. In a hypothesis test the computed p-value is 0.08. Our decision is to:

a. not reject Ho at the ten percent significance level and not reject Ho at the five percent significance level
b. not reject Ho at the ten percent significance level and reject Ho at the five percent significance level
c. reject Ho at the ten percent significance level and not reject Ho at the five percent significance level
d. reject Ho at the ten percent significance level and reject Ho at the five percent significance level

The Answer is c.

✓ **Other Tests**

When we want to test whether the observed difference between two means is because of chance, we must decide whether the samples are independent or dependent (related). If the samples are independent, we conduct tests concerning differences between means. If the samples are dependent, we often conduct tests of mean differences (paired comparison tests).

❒ **Tests Concerning Differences Between Means**

We often want to know whether a mean value differs between two groups.

1. H_0: $\mu_1 - \mu_2 = 0$ versus H_a: $\mu_1 - \mu_2 \neq 0$ (the alternative is that $\mu 1 \neq \mu 2$)
2. H_0 $\mu_1 - \mu_2 \leqq 0$ versus H_a: $\mu_1 - \mu_2 > 0$ (the alternative is that $\mu 1 > \mu 2$)
3. H_0: $\mu_1 - \mu_2 \geqq 0$ versus H_a: $\mu_1 - \mu_2 < 0$ (the alternative is that $\mu 1 < \mu 2$)

❒ Test Statistic for a Test of the Difference between Two Population Means (normally distributed populations, population variances unknown but assumed equal). When we can assume that two populations are normally distributed and that the unknown population variances are equal, a –test based on independent random samples is given by

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2})^{\frac{1}{2}}}$$

Where $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ is a pooled estimator of the common variance.

The number of degrees of freedom is $n_1 + n_2 - 2$

相關題型

1. A restaurant industry analyst thinks that after 9/11, families are eating out less often. He reviewed that results of an earlier study based on a survey of 21 families that showed that the mean meals eaten out per week 2.3 with a standard deviation of 0.9. He sampled

21 families after 9/11 and found that the mean meals eaten out per week were 1.8 with a standard deviation of 0.7. We assume that the population is normally distributed and that the population variances are equal. If we construct the null hypothesis that there is no difference in the mean number of meals eaten out per week, can we reject it at a 95 percent confidence level, and what is the statistical justification?

a. reject the null, the t-statistic is 0.0496 greater than the critical value
b. cannot reject the null, the t-statistic is 0.0114 less than the critical value
c. reject the null, the t-statistic is 0.0114 more than the critical value
d. cannot reject the null, the t-statistic is 0.0496 less than the critical value

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{20 \times 0.9^2 + 20 \times 0.7^2}{20 + 20 - 2} = 0.65$$

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{(\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2})^{\frac{1}{2}}} = \frac{2.3 - 1.8 - 0}{\sqrt{\frac{2 \times 0.65}{21}}} = 2.0096$$

查表得知臨界 t 值= $t_{0.025,40} = 2.021$

The answer is b

◆ Section 7: Correlation and Regression

✓ Covariance vs. the Correlation Coefficient

■ The sample correlation coefficient for two variables X and Y is

$$r = \frac{Cov(x_i, y_i)}{s_x s_y}$$

Caveats in interpretation of correlation coefficient

- ❒ Two variables can have a strong nonlinear relation and still have a very low correlation.
- ❒ Correlation may not be a reliable measure when outliers are present in one or both of the series.
- ❒ Correlation does not imply causation. *Spurious correlation* can make two series appear closely associated when no causal relation exists.

✓ The Regression Equation

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

$$\hat{Y}_i = b_0 + b_1 x_i$$

Linear regression computes a line that best fits the observations; it chooses values for the slope, b_1 and intercept b_0, that minimize the sum of the squared vertical distances between the observations and the regression line. That is, linear regression chooses the estimated or fitted parameters, $\hat{b}_0 \text{ and } \hat{b}_1$ in Equation to minimize

$$\sum_{i=1}^{n} (Y_i - \hat{b}_0 - \hat{b}_1 x_i)^2$$

✓ **SEE, Coefficient of Determination, and the Correlation Coefficient**

The Standard error of estimate (SEE) in a linear regression with one independent variable can be computed with the formula:

$$SEE = \left(\sum_{i=1}^{n} \frac{(Y_i - \hat{b}_0 - \hat{b}_1 x_i)^2}{n-2} \right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} \frac{(\hat{\varepsilon}_i)^2}{n-2} \right)^{1/2} \quad .$$

The standard error of estimate measures how well the regression model fits the data. If SEE is small, the model fits well.

The **coefficient of determination (R^2)** is formally defined as the percentage of the total variation in the dependent variable explained by the independent variable. In a linear regression with one independent variable, the simplest way to compute the coefficient of determination is to square the correlation of the dependent and independent variables.

$$R^2 = \frac{\text{total variation} - \text{unexplained variation}}{\text{total variation}} = \frac{\text{explained variation}}{\text{total variation}}$$

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

SST= $\sum (y_i - \bar{y})^2$; RSS= $\sum (\hat{y}_i - \bar{y})^2$; SSE= $\sum (y_i - \hat{y}_i)^2$

相關題型

1. Given the following information, determine what percentage of the variation of Y is NOT explained by the regression.

Y = 16 – 4X
σ_y = 12
σ_x = 15
cov_{yx} = 120

a. 66.7%
b. 33.3%
c. 44.4%
d. 55.6%

在單一變數迴歸中，要記住判定係數(R-Squared)=相關係數的平方值

相關係數的公式= $r = \frac{Cov(x_{i,}y_{i)}}{s_x s_y} = \frac{120}{12 \times 15} = 0.6667$ 所以 R-Square = 0.4444

1- 4444=0.5556，就是不能被迴歸線解釋的部分

The answer is d.

✓ **Confidence Intervals and hypothesis test for the slope and dependent value**

Use the following data to answer the following four questions:
To help gain a better understanding of the relationship between the return on the common stocks of small companies and the return on the S&P 500 index, you run a simple linear regression to quantify this relationship, using the monthly return on small stocks as the dependent variable and the monthly return on the S&P500 as the independent variable. The results of the regression are shown below:

	Coefficient	Standard error of Coefficient	t-Value
Intercept	1.71%	2.95	0.58
S&P 500	1.52	0.13	11.69

The t-statistic critical value at the 0.01 level is 2.58
Standard error of estimate=19.85%
Correlation coefficient=0.7740
N=100
F-Value=101.645 on 1/73 degrees of freedom

1. Use the regression statistics presented above and assume this historical relationship still holds in the future period. If the expected return on the S&P 500 over the next period were 3%, the expected return on small stocks over the next period would be:
a. 4.56%
b. 5.13%

c. 6.27%
d. 6.65%

$\hat{Y}_i = b_0 + b_1 x_i = 1.71\% + 1.52 \times 3\% = 6.27\%$

The answer is c

---

2. The percent of the variation in the return on the dependent variable (return on small stocks) explained by the return on the independent variable (return on the S&P 500) for the period under study was:

a. 10.07%
b. 19.85%
c. 59.91%
d. 77.40%

考 R-Squared (判定係數)
在單一變數迴歸中，判定係數會等於相關係數的平方。相關係數題目有給，是 0.774。所以判定係數等於 0.5991 或 59.91%
The answer is c.

---

3. The regression statistics presented above indicate that at the 0.01 level, the slope coefficient(1.52)

a. and the y-intercept (1.71) are both statistically significant
b. and the y-intercept (1.71) both lack statistical significance
c. is *not* statistically significant, but the y-intercept (1.71) is statistically significant
d. is statistically significant, but the y-intercept (1.71) is *not* statistically significant

---

這一題只是考觀念題，問你迴歸式之的截距和斜率之 t 觀察值和 t 臨界值的比較到底能不能拒絕截距和斜率等於 0 的虛無假設。截距(b_0)的 t 觀察值等於 0.58，斜率(b_1)的 t 觀察值等於 11.69。此題有給 t 臨界值是 2.58。因此斜率值在 1%的顯著水準下，可以拒絕是 0 的虛無假設，但截距值則不行。答案選 d

---

4. A analyst is estimating Microsoft Corp's stock beta using the past 60 trading day's closing prices against S&P 500's. His estimate is 1.3531 and he is wondering if the beta is significantly different from the market beta. The following is the table from his statistics software result:

Regression Statistics

Multiple R		0.5411	
R-squared		0.2928	
Standard error of estimate		0.0835	
Observations		60	
	Coefficient	Standard Error	*t*-statistics
Alpha	0.00267	0.0117	2.2819
Beta	1.3531	0.2761	

Calculate the t statistic and, using the significant level of 1%, conclude if Microsoft's Beta is different from the market beta?

	t-statistic	Microsoft's Beta
a.	4.0901	different from market beta
b.	2.2819	different from market beta
c.	1.3531	equal to market beta
d.	1.2789	equal to market beta

$t = \frac{1.3531 - 1}{0.2761} = 1.2789$ t 的臨界值牽涉到自由度和顯著水準。其中斜率的自由度是 n-2，顯著水準是 0.01/2=0.005 或 0.5%，查表後可以 t 臨界值接近 2.66。既然觀察值沒有大於正的臨界值或小於負的臨界值，我們不能拒絕微軟公司的系統性風險和市場風險(=1)有何差別

The answer is d.

✓ **ANOVA (Analysis of Variance)**

■ A typical ANOVA table looks like this:

ANOVA	Degrees of Freedom (df)	Sum of Squares (SS)	Mean Sum of Squares (MSS)	F
Regression	K (有多少組自變數)	**RSS**	MSR=SSR/K	MSR/MSE
Residual	n-k-1	**SSE**	MSE=SSE/n-k-1	
Total	n-1	SST		

In regression analysis, we use F-test to determine if all the slope coefficients in a linear regression are equal to 0. In a regression with one independent variable, this is a test of the null hypothesis Ho: $b_1=0$ against the alternative hypothesis Ha: $b_1 \neq 0$. The ANOVA table is the table to streamline the process in developing the F statistic and use the F distribution to measure how well the regression equation explains the variation in the dependent variable.

$$F = \frac{RSS/1}{SSE/(n-k-1)} = \frac{Mean\ \text{Re}gression\ Sum\ of\ \ Square}{Mean\ Sqaured\ \ Error}$$

相關題型

Use the following information for the following four questions:

Source of Variation	Degrees of Freedom	Variation	Variance
Regression	1	20,500	20,500
Error	498	2,500	5.02
Total	499	23,000	

1. What is the standard error of estimate of the regression on which the table is based?
   a. 5.02
   b. 2.24
   c. 1.43
   d. 2.05

The answer is b.

SEE 的全名是 Standard Error of Estimate，它在 ANOVA 表中沒有直接顯示，但其實它等於 MSE 的開根號取正的部分。在該題提供的 ANOVA 表中，你要找到 MSE 等於 5.02，開根號之後為 2.24

---

2. How many independent variables were used in the regression model?

a. 1
b. 2
c. 499
d. 498

這一題考觀念題，考這個迴歸的自變數有幾組。從 ANOVA 表中的斜率行中可以看到其自由度是 1，代表自變數的數目就是 1 組。在迴歸式有幾個參數(斜率)，就代表有幾組自變數。

The answer is a.

---

3. Determine the coefficient of determination, the coefficient of correlation, and the standard error of estimate of the regression on which the table is based?

a. 94.4%; -0.891; 5.02
b. 94.4%; +0.891; 2.24
c. 89.1%; 94.4%; 5.02
d. 89.1%; 94.4%; 2.24

這一題要你算三個數值，其中 SEE 剛才你已經算過了，是 2.24。在 ANOVA 中，判定係數 =RSS/SST=20500/23000=0.8913。在單變數迴歸式中，你知道相關係數等於判定係數的平方根，但這一題的平方根要取正還是取負，其實很難判定，這一題又沒有原始的 Data。但由於答案 b 所對應的 SEE 是正確的，所以我們選之。

The answer is b.

---

4. Calculate the F-statistic for the data given in the ANOVA table:

a. 4,084
b. 8.2
c. 498
d. 3,745

F=MSR/MSE=20500/5.02=4084

The answer is a
