Lect 1, MATH 5261

Returns

收益

Zongyi Liu

Fri, Sept 5, 2025

Contents

1 收益

1.1 收益的意義

- 投資的目標是獲取利潤.
- 投資的收益 (或在收益為負時的損失) 取決於資產價格的變化以及持有資產的數量.
- 作為投資者, 你關心的是相對於初始投資規模而言較高的收益.

- 報酬率衡量的正是這一點, 因為資產 (例如股票, 債券或股票與債券組合) 的報酬率, 是價格變動相對於初始價格的比例.

1.2 收益公式

- 設 P_t 為資產在時間 t 的價格. 假設期間未發放股息, 則淨報酬定義為:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1$$

- 簡單總報酬 (simple gross return) 定義為:

$$\frac{P_t}{P_{t-1}} = 1 + R_t$$

- 範例: 若 $P_{t-1} = 10$ 且 $P_t = 10.3$, 則:

$$1 + R_t = \frac{P_t}{P_{t-1}} = \frac{10.3}{10} = 1.03$$

得 $R_t = 0.03$ 或 $R_t = 3\%$.

- 最近連續 k 期 $(t-k, t-k+1, \ldots, t)$ 的總報酬定義為:

$$\begin{aligned} 1 + R_t(k) &= \frac{P_t}{P_{t-k}} \\ &= \frac{P_t}{P_{t-1}} \frac{P_{t-1}}{P_{t-2}} \cdots \frac{P_{t-k+1}}{P_{t-k}} \\ &= (1 + R_t)(1 + R_{t-1}) \ldots (1 + R_{t-k+1}) \end{aligned}$$

1.3 收益的性質

注意到, 收益都是:

- 無尺度 (scale-free), 意即它們不依賴於貨幣單位 (如美元, 分等).
- 但在時間上並非無單位 (not unit-less in time), 因為它們依賴於時間單位 t (例如小時, 天等).

1.4 案例

時間	t − 3	t − 2	t − 1	t
P	200	210	206	212
1 + R		1.05	0.981	1.03
1 + R(2)			1.03	1.01
1 + R(3)				1.06
1 + R				

$$1.05 = 210/200 \quad 0.981 = 206/210 \quad 1.03 = 212/206$$

$1 + \mathrm{R}(2)$

$$1.03 = 206/210 \quad 1.01 = 212/210$$

$1 + \mathrm{R}(3)$

$$1.06 = 212/200$$

- 對數價格 (log prices) 被定義為:

$$p_t = \log(P_t)$$

其中 $\log(x)$ 為 x 的自然對數.

- 連續複利報酬 (或稱對數報酬) 定義為總報酬的對數:

$$r_t = \log(1 + R_t) = \log\left(\frac{P_t}{P_{t-1}}\right) = p_t - p_{t-1}$$

- 注意到有:

$$P_t = P_{t-1}e^{r_t}$$

- 此外:

$$\begin{aligned} r_t(k) &= \log(1 + R_t(k)) \\ &= \log((1 + R_t)(1 + R_{t-1}) \ldots (1 + R_{t-k+1})) \\ &= \log(1 + R_t) + \log(1 + R_{t-1}) + \ldots + \log(1 + R_{t-k+1}) \\ &= r_t + r_{t-1} + \ldots + r_{t-k+1} \end{aligned}$$

1.5 對數收益

由於當 x 很小時, $\log(1+x) \approx x$, 因此對數報酬大約等於淨報酬. 在此情況下, $r_t = \log(1 + R_t) \approx R_t$.

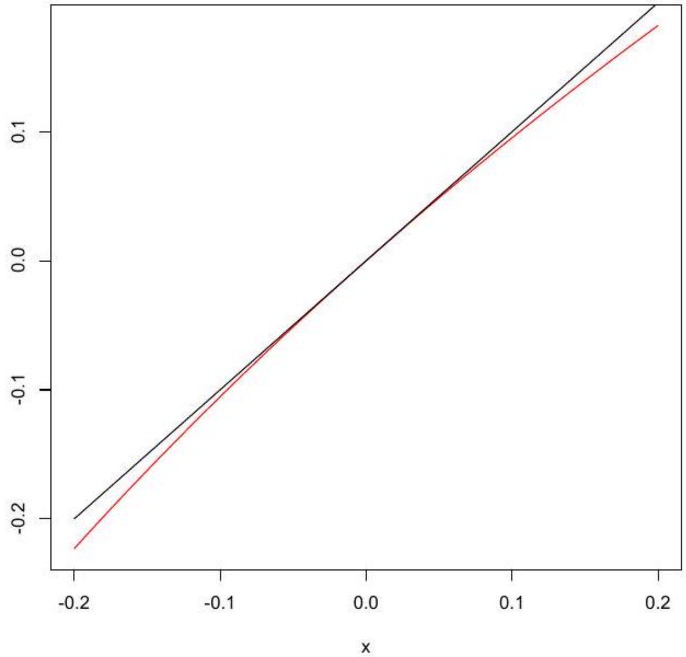
2 收益模型

2.1 正態模型

在時間 $t-1$, P_t 和 R_t 不僅未知, 而且我們也不知道它們的機率分布. 一個常見的模型假設報酬率是獨立, 同分布且服從常態分布的. 也就是說, 若 $R_1, R_2, \ldots, R_t$ 為同一資產的報酬率, 則:

- $R_1, R_2, \ldots, R_t$ 彼此獨立;

Figure 1: Figure - graphs of $f(x)=x$ (black) and $g(x)=\log(1+x)$ (red)



- 同分布;
- 且服從常態分布.

此模型存在兩個問題:

1. 該模型暗示可能出現無限損失, 但實際上損失通常是有限的 ($R_t \geq -1$), 因為你不可能損失超過原始投資金額.
2. 此外,

$$1+R_t(k)=\prod_{i=0}^{k-1}\left(1+R_{t-i}\right)$$

報酬率的乘積並不服從常態分布. 在某些條件下, 獨立常態分布隨機變數的「和」仍為常態分布, 但它們的「乘積」則不是.

2.2 對數-正態模型

2.2.1 性質

在此中我們假設:

$$r_t=\log\left(1+R_t\right)$$

是 IID 正態分佈, 即:

$$\log\left(1+R_t\right) \sim N\left(\mu, \sigma^2\right)$$

因此有:

1. $1+R_t=\exp(\text{常態隨機變數})\geq 0$

2. $R_t \geq -1$

- 注意到:

$$
\begin{aligned}
1+R_t(k) &= \prod_{i=0}^{k-1}\left(1+R_{t-i}\right) \\
&= \exp\left(r_t\right)\exp\left(r_{t-1}\right)\ldots\exp\left(r_{t-k+1}\right) \\
&= \exp\left(r_t+r_{t-1}+\ldots+r_{t-k+1}\right)
\end{aligned}
$$

因此:

$$
\log\left(1+R_t(k)\right)=\sum_{i=0}^{k-1} r_{t-i}
$$

又由於獨立常態變數的「和」仍為常態分布, 單期對數報酬若為常態, 則多期對數報酬亦為常態.

範例: 設簡單總報酬 $(1+R)$ 服從對數常態. Lognormal(0, 0.01), 即

$$
\log(1+R) \sim N\left(\mu=0,\ \sigma^2=0.01\right).
$$

求 $P(R \leq 0.05)$.

解:

$$
\begin{aligned}
P(R \leq 0.05) &= P(1+R \leq 1.05) \\
&= P(\log(1+R) \leq \log(1.05)) \\
&= P(N(0, 0.01) \leq \log(1.05)) \\
&= \Phi((\log(1.05)-0)/0.1)=0.68719
\end{aligned}
$$

第 k 期報酬的公式. 假設:

- $1+R_t(k)=\left(1+R_t\right)\left(1+R_{t-1}\right)\ldots\left(1+R_1\right)$
- $\log\left(1+R_i\right) \sim N\left(\mu, \sigma^2\right)$
- $\{R_i\}$ 是獨立的, 則:

$$
\log\left(1+R_t(k)\right) \sim N\left(k\mu, k\sigma^2\right)
$$

以及:

$$
P\left(R_t(k)<x\right)=\Phi\left(\frac{\log(1+x)-k\mu}{\sqrt{k}\sigma}\right)
$$

範例: 再次假設 $(1+R)$ 服從對數常態分布 $(0, 0.01)$. 求兩期簡單總報酬小於 1.05 的機率

答案: 兩期的總報酬服從對數常態分布 $(0, 0.02)$. 因此:

$$
\begin{aligned}
P(R \leq 0.05) &= P(1+R \leq 1.05) \\
&= P(\log(1+R) \leq \log(1.05)) \\
&= P(N(0, 0.01) \leq \log(1.05)) \\
&= \Phi((\log(1.05)-0)/\sqrt{0.02})=0.6349
\end{aligned}
$$

2.2.2 對數-正態密度

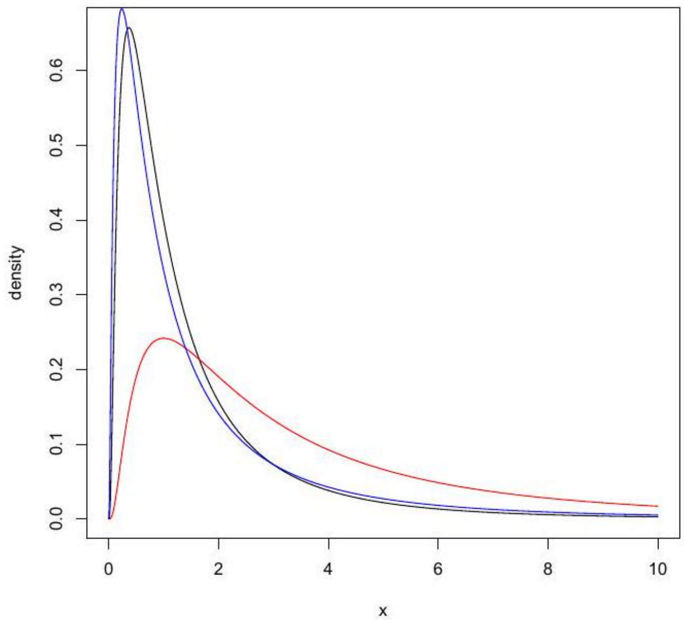


Figure 2: Figure - graphs of lognormal (0, 1) (black), lognormal (1, 1) (red) and lognormal ($0, 1.2$) (blue)

2.2.3 股利的調整

此處討論股利的調整 (Adjustment for Dividends), 許多股票會支付股利, 在計算報酬時必須將其納入考量. 若在時間 t 之前支付股利 D_t, 則在時間 t 的總報酬 (gross return) 定義為:

$$1 + R_t = \frac{P_t + D_t}{P_{t-1}}$$

淨收益為:

$$R_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}}$$

以及:

$$r_t = \log\left(1 + R_t\right) = \log\left(P_t + D_t\right) - \log\left(P_{t-1}\right)$$

多期的總報酬 (multiple-period gross returns) 為各單期報酬的乘積.

$$\begin{aligned} 1 + R_t(k) &= \left(\frac{P_t + D_t)}{P_{t-1}}\right)\left(\frac{P_{t-1} + D_{t-1}}{P_{t-2}}\right)\cdots\left(\frac{P_{t-k+1} + D_{t-k+1}}{P_{t-k}}\right) \\ &= \left(1 + R_t\right)\left(1 + R_{t-1}\right)\ldots\left(1 + R_{t-k+1}\right) \end{aligned}$$

類似地, k 期的對數報酬是各單期報酬之對數的總和

$$
\begin{aligned}
\log(1+R_t(k)) &= \log\left(\frac{P_t+D_t)}{P_{t-1}}\right)+\log\left(\frac{P_{t-1}+D_{t-1}}{P_{t-2}}\right)+\ldots \\
&+\log\left(\frac{P_{t-k+1}+D_{t-k+1}}{P_{t-k}}\right) \\
&= \log(1+R_t)+\log(1+R_{t-1})+\ldots+\log(1+R_{t-k+1})
\end{aligned}
$$

2.3 隨機漫步

2.3.1 引入

設 $Z_1, Z_2, \ldots, Z_n$ 為具有平均數 μ 與標準差 σ 的獨立同分布 (IID) 隨機變數. 令 Z_0 為任意起始點, 並定義:

$$
S_0 = Z_0, \quad \text{and} \quad S_t = Z_0 + Z_1 + \ldots + Z_t, t \geq 1.
$$

過程 $S_0, S_1, \ldots$ 稱為隨機漫步 (random walk), 而 $Z_1, Z_2, \ldots$ 稱為其步長 (steps). 在給定 S_0 的條件下, S_t 的期望與變異數分別為:

- $E(S_t \mid Z_0) = Z_0 + t\mu$
- $\operatorname{Var}(S_t \mid Z_0) = \sigma^2 t$
- $SD(S_t \mid Z_0) = \sigma\sqrt{t}$
- 參數 μ 稱為漂移 (drift), 它決定隨機漫步的大致方向.
- 參數 σ 為波動率 (volatility), 決定隨機漫步相對於條件平均值 $S_0 + \mu t$ 的波動程度.
- 由於在給定 S_0 的情況下, S_t 的標準差為 $\sigma\sqrt{t}$, 因此 $(S_0 + \mu t) \pm \sigma\sqrt{t}$ 表示平均值加減一個標準差. 對於正態隨機漫步 (即 Z_i 為正態分布) 而言, 這對應於條件平均值的 68% 信賴區間.
- 該區間的寬度與 $\sqrt{t}$ 成正比, 如下一張圖所示.

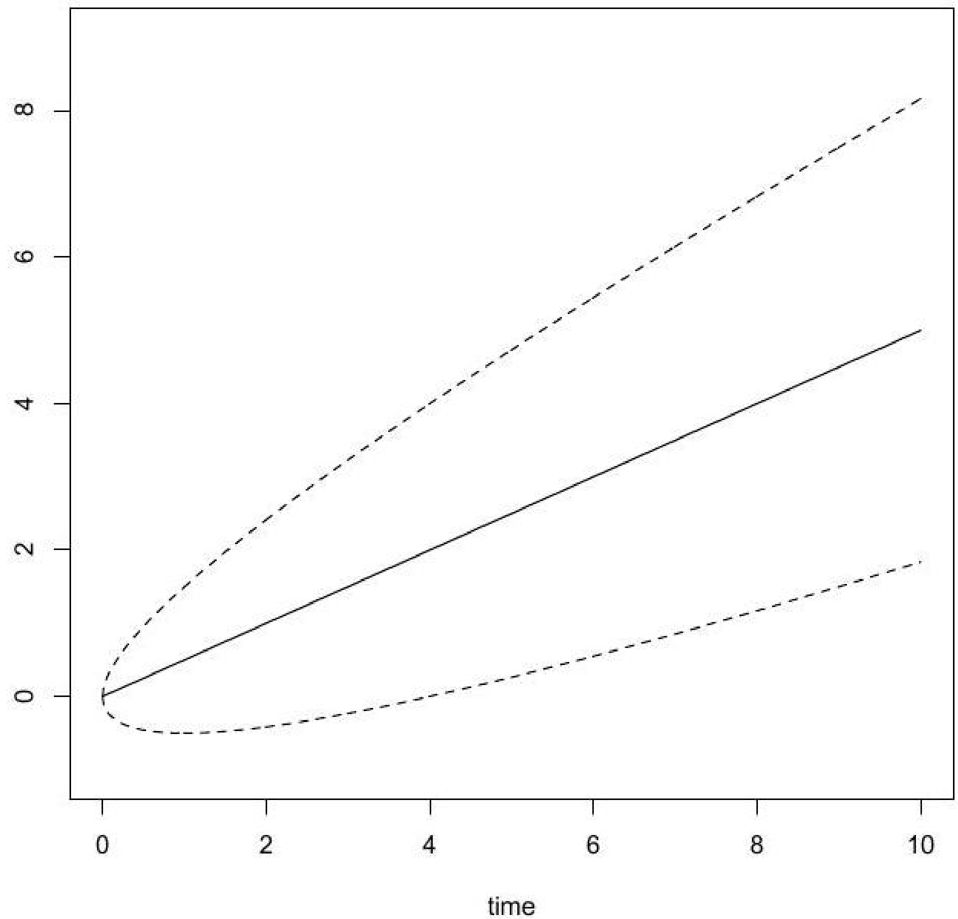


Figure 3: 對於具有 $S_0 = 0, \mu = 0.5, \sigma = 1$ 的隨機漫步, 其平均值與機率界限如下圖所示. 在任意時間點, 位於機率界限 (虛線曲線) 之間的機率為 68%. 此處的 Z_i 服從正態分布.

2.3.2 R 指令

生成這些圖片的 R 指令為:

```
T=100
x = 0 : T
y=c (0, cumsum (rnorm (T) ) )
plot (x, y, lty=1, type="I", xlab="", ylab="")
```

2.3.3 幾何隨機漫步

如下關係式:

$$\frac{P_t}{P_{t-k}} = 1 + R_t(k) = \exp\left(r_t + r_{t-1} + \ldots + r_{t-k+1}\right)$$

若取 $k = t$, 則有

$$P_t = P_0 \exp\left(r_t + r_{t-1} + \ldots + r_1\right)$$

若 $r_1, r_2, \ldots, r_t$ 為獨立同分布的 $N(\mu, \sigma^2)$, 即總報酬為對數常態且彼此獨立, 則有:

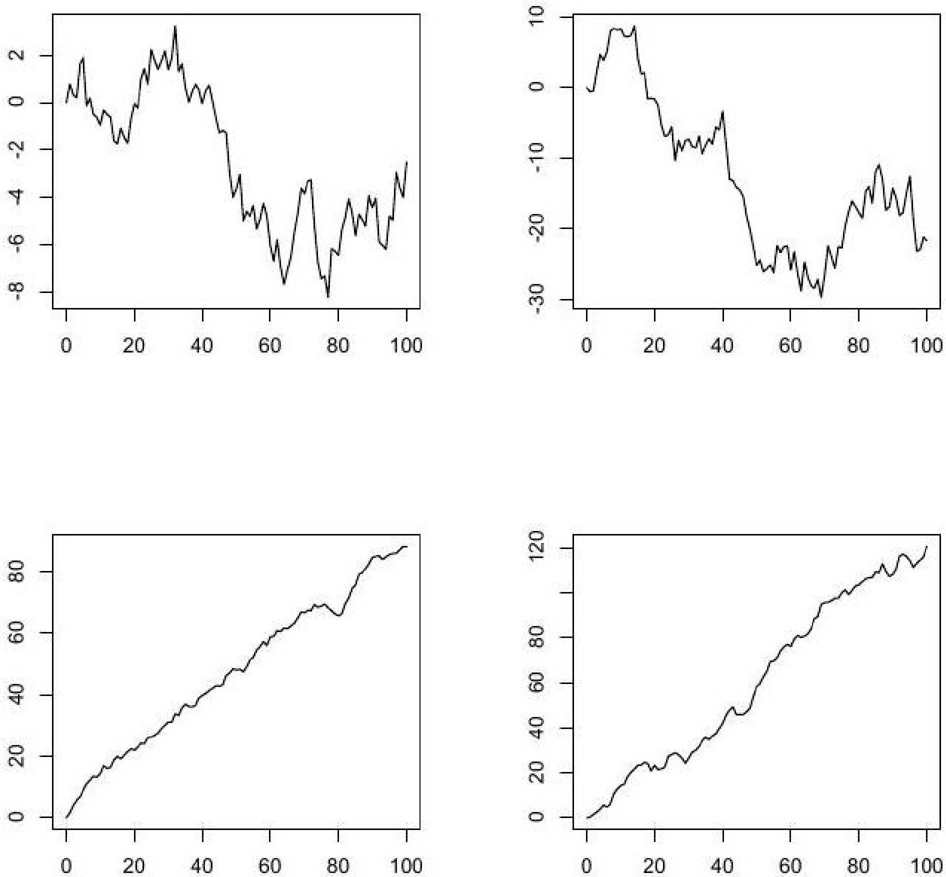


Figure 4: 具有參數 $(\mu, \sigma) = (0,1), (0,2), (1,1)$ 與 $(1,2)$ 的隨機漫步

- $\log(1 + R_t(k)) = r_1 + r_2 + \ldots + r_{t-k+1}$ 是一個隨機漫步.
- $\{P_t,\, t = 1, 2, \ldots, k\}$ 為隨機漫步的指數形式.
- 此類過程稱為幾何隨機漫步 (geometric random walk).
- $1 + R_t(k) = \exp(r_t + r_{t-1} + \ldots + r_1)$ 服從對數常態分布 (lognormal), 且具有偏態 (skewed).

飄移 μ 的作用:

- 幾何隨機漫步並不意味著無法獲利.
- 由於 μ 為正值, 該過程具有向上的漂移 (upward drift).
- 美國整體股市的對數報酬率平均約為 10%, 標準差約為 20%.

2.4 收益的分佈

2.4.1 三種常見檢測

對數報酬是否服從常態分布? 檢驗方法包括:

- 常態圖 (normal plot): 若資料點大致呈一直線, 則近似常態.

- 樣本偏態 (skewness) 與峰度 (kurtosis): 檢查其數值是否接近常態分布的理論值. 偏態與峰度衡量分配形狀, 且與平均數與變異數無關. 任何常態分布的偏態係數皆為 0, 峰度為 3.

設 $r_1, r_2, \ldots, r_T$ 為某資產的對數報酬, 令 $\hat{\mu}$ 為樣本平均值, $\hat{\sigma}$ 為樣本標準差. 則樣本偏態 (sample skewness) 定義為:

$$\hat{S} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \hat{\mu}}{\hat{\sigma}}\right)^3$$

樣本峰度則為:

$$\hat{K} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \hat{\mu}}{\hat{\sigma}}\right)^4$$

超額峰度 (excess kurtosis) 定義為 $\hat{K} - 3$. 它衡量與常態分布峰度 3 的偏離程度. 偏態與超額峰度的數值若接近 0, 則表示資料分布接近常態.

常態性檢定 (testing for normality): 令 $\hat{F}$ 為經驗累積分布函數 (empirical CDF).

$$\hat{F}(x) = \frac{1}{T}\sum_{t=1}^{T} I\left(r_t \le t\right)$$

$\hat{F}$ 是真實分布的估計量. 常態性檢定透過比較 $\hat{F}(x)$ 與 $\Phi((x-\hat{\mu})/\hat{\sigma})$ 來進行.

2.4.2 四種額外檢測

三種常用的常態性檢定方法, 皆比較 $\hat{F}(x)$ 與 $\Phi((x-\hat{\mu})/\hat{\sigma})$, 包括:

- 安德森–達令檢定 (Anderson–Darling test)
- 夏皮羅–威爾克檢定 (Shapiro–Wilks test)
- 柯莫哥洛夫–史密諾夫檢定 (Kolmogorov–Smirnov test)

我們將在未來進一步討論這些檢定.

柯爾莫哥洛夫–史密諾夫檢定 (Kolmogorov–Smirnov test) 是基於:

$$\sup_x |\hat{F}(x) - \Phi((x-\hat{\mu})/\hat{\sigma})|$$

這是 $\hat{F}(x)$ 和 $\Phi((x-\hat{\mu})/\hat{\sigma})$ 之間的最大距離.