Relazione

Daniele Cioffi, Marco Martino, Luca Komisarjevsky

Giugno 2024

Problema

Francesco 10 anni fa si è recato da un consulente finanziario della sua banca, il quale gli ha costruito un portafoglio di investimenti composto da 4 fondi comuni. Recentemente, guardando il rendimento del suo portafoglio, si è chiesto se fosse stato il caso di investire in un portafoglio di ETF. Si pone quindi le seguenti domande:[1]

Presentazione campione casuale

Statistica descrittiva riguardante il campione casuale da noi prelevato.

Quanto costano gli ETF?

Statistica descrittiva per i costi di gestione annuali del nostro campione casuale con rispettivo intervallo di confidenza per media con varianza incognita al 95%.

$$0.95 - I.C. : (0.1899306, 0.3002917) \tag{1}$$

Da cosa dipendono i costi di gestione (o TER)?

Statistica descrittiva per i possibili regressori dei costi di gestione annuali.

Formulazione modelli di regressione[2] dove Y = costi di gestione annuali (in %), x_1 = scarto rendimento medio annuale ETF-Indice (in %), x_2 = dimensione del fondo (in mld), x_3 = tipo di replica (sintetica, fisica totale, a campionamento), x_4 = indice di rischio (numero intero da 2-5):

1. $Y = \beta_0 + \beta_1 x_1 + \epsilon$, $\epsilon \sim N(0, \sigma^2)$
2. $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2 \cdot x_3 + \epsilon$, $\epsilon \sim N(0, \sigma^2)$ $\leftarrow$ modello migliore
3. $Y = \beta_0 + \beta_1 x_4 + \beta_2 x_1 + \epsilon$, $\epsilon \sim N(0, \sigma^2)$

Quanto costerebbe un ETF con fondo di dimensione 2mld, scarto di 0.05% rispetto all'indice e tipo di replica "fisica totale"?

Intervallo di confidenza al 95% per il valore atteso del modello 2 con $x_1 = 0.05\%$, $x_2 = 2$, $x_3 =$"fisica totale":

$$0.95 - I.C. : (0.2111689, 0.3240669) \tag{2}$$

Quanto rende un ETF in media su 10 anni?

Intervallo di confidenza al 95% per media con varianza incognita sul rendimento cumulativo (in %) degli ETF:

$$0.95 - I.C. : (87.2204, 165.1035) \tag{3}$$

Gli ETF raggiungono il loro obiettivo?

Anche se il nostro campione presenta dati accoppiati (rendimento medio annuale ETF - rendimento medio annuale indice) abbiamo svolto comunque un test di indipendenza:

- H_0 : rendimento annuale medio ETF e rendimento annuale medio dell'indice sono indipendenti
- H_1 : rendimento annuale medio ETF e rendimento annuale medio dell'indice **NON** sono indipendenti

Il p-value del test risulta: $2,499 \cdot 10^{-10}$ quindi concludiamo fortemente per H_1

Abbiamo poi effettuato un test per la differenza fra la media del rendimento annuale medio degli ETF e quello degli indici intendendo per replicazione dell'indice uno scarto maggiore del -0.30%

- $H_0 : \mu_1 - \mu_2 \leq -0.30\%$, $H_1 : \mu_1 - \mu_2 > -0.30\%$ $\rightarrow$ pvalue = 0.02338 (concludiamo fortemente per H_1)

[1] Tutti i dati usati sono stati reperiti dal sito justetf.com (per quanto riguarda gli ETF) e fondionline.it (per quanto riguarda i fondi attivi del portafoglio di Francesco). Tutti i dati sono raggruppati in un documento excel che verrà corredato alla seguente relazione

[2] Ogni modello di regressione, stima intervallare, test d'ipotesi è stato svolto su R. I risultati sono visibili nella presentazione

Per confermare ulteriormente che gli ETF raggiungono il loro obiettivo abbiamo impostato un intervallo di confidenza al 95% per la proporzione p: "probabilità che lo scarto tra rendimento annuale medio dell'ETF e quello dell'indice sia ≥ 0".

$$0.95 - I.C.: (0.2995055, 0.5987643) \tag{4}$$

Infine abbiamo sviluppato un modello che quasi completamente spiega la percentuale di variabilità del rendimento dell'indice tramire il rendimento dell'ETF (regressore x):

$$Y = \beta_0 + \beta_1 x + \epsilon,\ \epsilon \sim N(0, \sigma^2) \tag{5}$$

Mediamente un ETF fa meglio del portafoglio di Francesco?

Abbiamo svolto i seguenti test (concludiamo fortemente per H_1 in ogni test)[3]:

1. Test su rendimento medio annuale: $H_0: \mu \leq 2.39\%$, $H_1: \mu > 2.39\% \rightarrow$ p-value $= 8.526 \cdot 10^{-9}$
2. Test su rendimento medio cumulato: $H_0: \mu \leq 27.2\%$, $H_1: \mu > 27.2\% \rightarrow$ p-value $= 3.223 \cdot 10^{-6}$
3. Test per TER: $H_0: \mu \geq 1.76\%$, $H_1: \mu < 1.76\% \rightarrow$ p-value$=2.2 \cdot 10^{-16}$
4. Test su scarto annuale medio: $H_0: \mu \leq -2.91\%$, $H_1: \mu > -2.91\% \rightarrow$ p-value $= 3.223 \cdot 10^{-6}$

Notiamo quindi che in media un ETF fa meglio del portafoglio di Francesco

Se dovesse comprare un ETF a caso, che scarto medio annuale potrebbe avere quest'ultimo su 10 anni?

Abbiamo svolto un intervallo di predizione per una nuova osservazione dalla popolazione (scarto annuale medio degli ETF) dopo aver eseguito il test di shapiro-wilk che ha restituito un p-value del $8.279 \cdot 10^{-6}$:

$$0.95 - I.P.: (-1.195299, 1.192828) \tag{6}$$

A causa del risultato del test di SW questo intervallo non è per niente affidabile

Francesco decide di crearsi un portafoglio scegliendo casualmente 40 ETF e vuole sapere la probabilità che almeno 30 di questi abbiano replicato l'indice.

X = n° di ETF che replicano l'indice su 40 ETF pescate.

$X \sim Bin(40, p)$

Per replica dell'indice si intende uno scarto annuale medio $> -0.3\%$. Il nostro campione ha 34 successi su 45. A questo punto abbiamo stimato puntualmente la probabilità p con $\hat{p}$, ma pensiamo che la vera probabilità sia maggiore di $\frac{34}{45}$, abbiamo quindi svolto un intervallo di confidenza per la proporzione p:

$$0.95 - I.C.: (0.6013926, 0.8661495) \tag{7}$$

Considerando che, come ampiamente verificato precedentemente, gli ETF raggiungono il loro obiettivo, ci aspettiamo che questa probabilità sia intorno all'80%, verifichiamo con un opportuno test: $H_0: p = 0.8$, $H_1: p \neq 0.8$ che ha p-value $= 0.4561$ (45%). Possiamo quindi concludere debolmente che la proporzione p da noi cercata è 80%.

$X \sim Bin(40, 0.8)$ da cui $P(X \geq 30): 73.1777\%$

Come hanno reagito gli indici agli eventi più significativi degli ultimi 10 anni?

Probabilità annuale media di trovare un rendimento negativo = 0.27561

Attraverso dei test andiamo a vedere se i rendimenti degli indici negli anni 2018-2020-2022 (anni in cui si sono verificati rispettivamente dazi, covid e guerra) hanno subito significativi cambiamenti:

1. 2018 ($\hat{p} = 0.85$): $H_0: p = 0.27561$, $H_1: p \neq 0.27561 \rightarrow$ p-value $< 2.2 \cdot 10^{-16}$
2. 2020 ($\hat{p} = 0.24$): $H_0: p = 0.27561$, $H_1: p \neq 0.27561 \rightarrow$ p-value $= 0.6496$
3. 2022 ($\hat{p} = 0.92$): $H_0: p = 0.27561$, $H_1: p \neq 0.27561 \rightarrow$ p-value $< 2.2 \cdot 10^{-16}$

Concludiamo fortemente per H_1 negli anni 2018 e 2022, mentre concludiamo debolmente per H_0 nel 2020

[3] Tutti i test sono a varianza incognita