



**POLITECNICO**  
MILANO 1863

# PROGETTO STATISTICA

DANIELE CIOFFI, MARCO MARTINO, LUCA KOMISARJEVSKY

# COS'È UN ETF?



Un ETF è un fondo di investimento che, diversamente dai fondi comuni di investimento attivi gestiti da un team che cerca di battere il mercato, hanno una gestione passiva (tipicamente gestita da algoritmi)



L'obiettivo degli ETF è quello di replicare un indice di mercato a costi di gestione ridotti

# PRESENTAZIONE DEL PROBLEMA

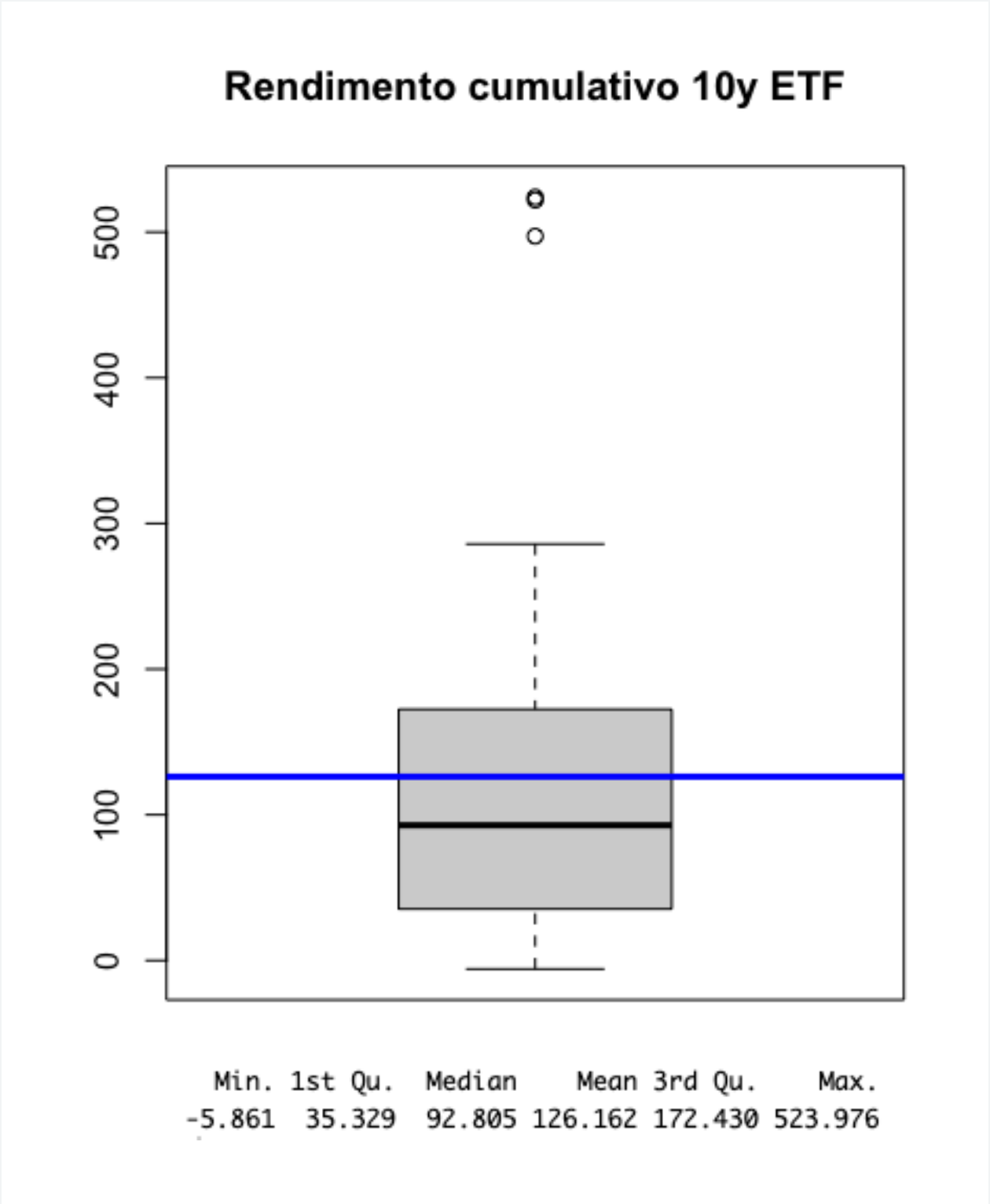
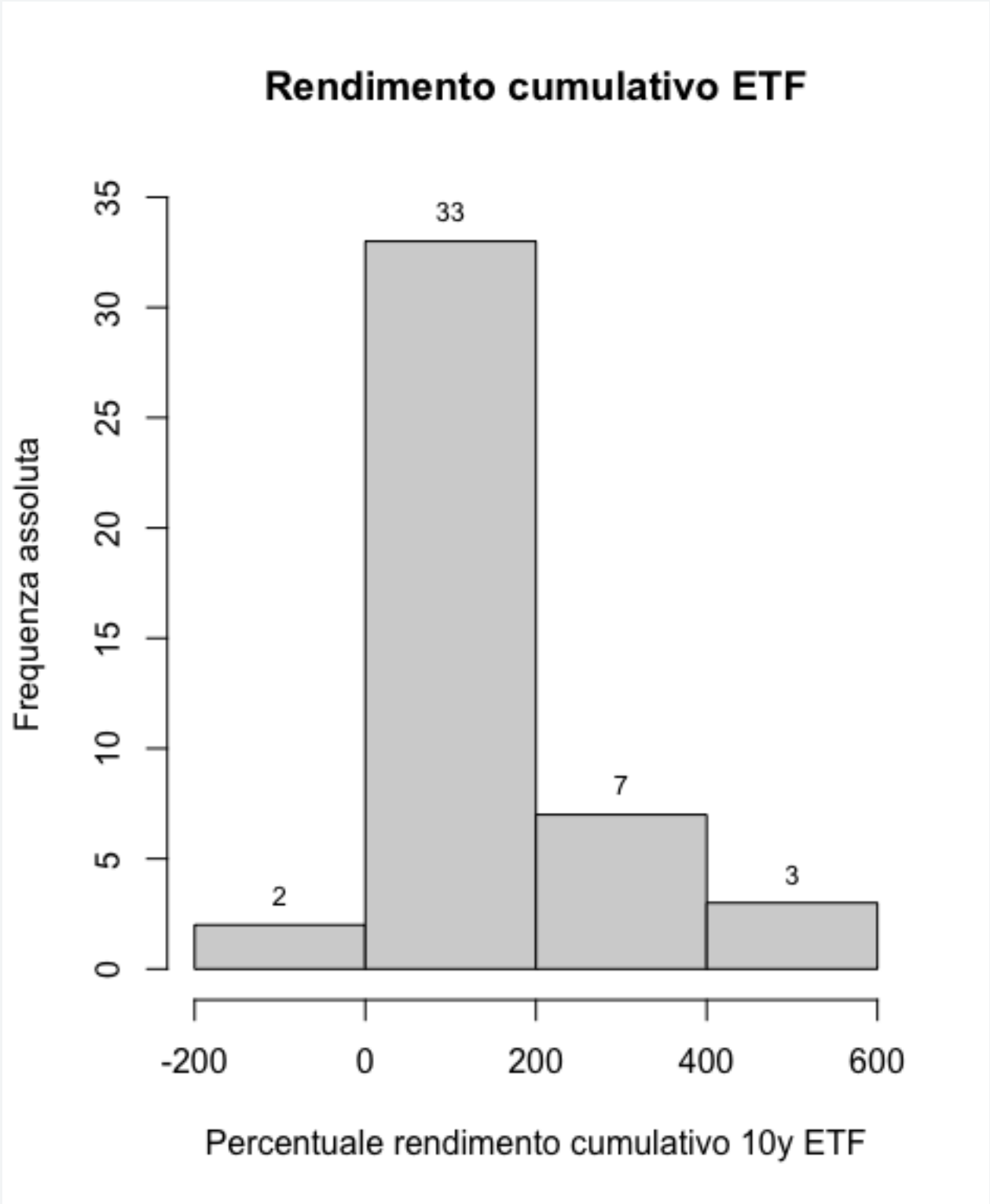
Francesco 10 anni fa si è recato da un consulente finanziario della sua banca, il quale gli ha costruito un portafoglio di investimenti composto da 4 fondi comuni. Recentemente, guardando il rendimento del suo portafoglio, si è chiesto se fosse stato il caso di investire in un portafoglio di ETF. Si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

- (1) Quanto costano gli ETF? Da cosa dipendono i costi di gestione? Quanto costerebbe un ETF di dimensione 2mld, scarto di 0.05% rispetto all'indice e tipo di replica «fisica totale»?
- (2) Quanto rende in media un ETF su 10 anni?
- (3) Gli ETF raggiungono il loro obiettivo? Mediamente un ETF fa meglio del portafoglio di Francesco?
- (4) Se dovesse comprare un ETF a caso, che scarto medio annuale potrebbe avere quest'ultimo su 10 anni?
- (5) Francesco decide di crearsi un portafoglio scegliendo casualmente 40 ETF e vuole sapere la probabilità che almeno 30 di questi abbiano replicato l'indice.
- (6) Come hanno reagito gli indici agli eventi più significativi degli ultimi 10 anni?

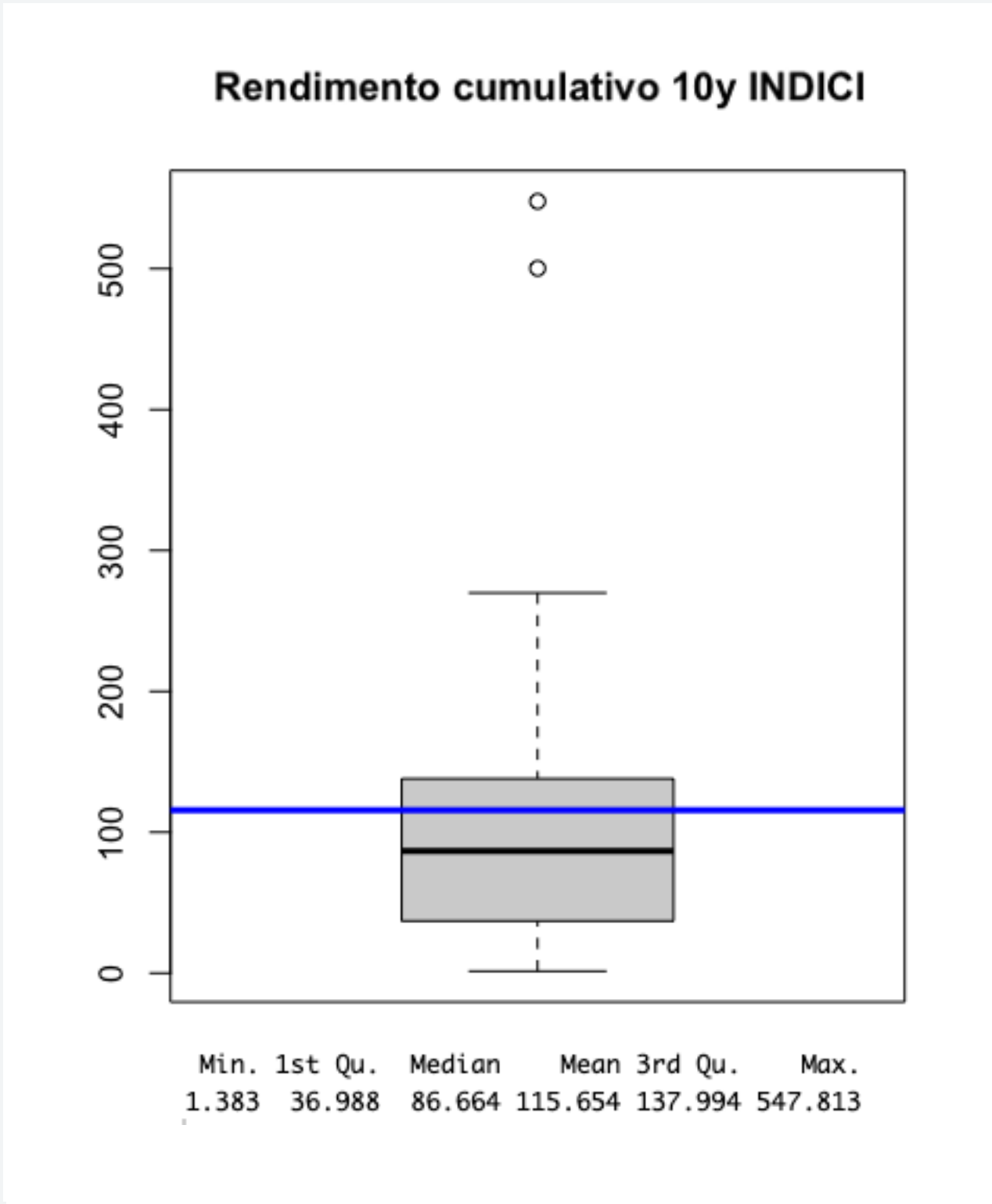
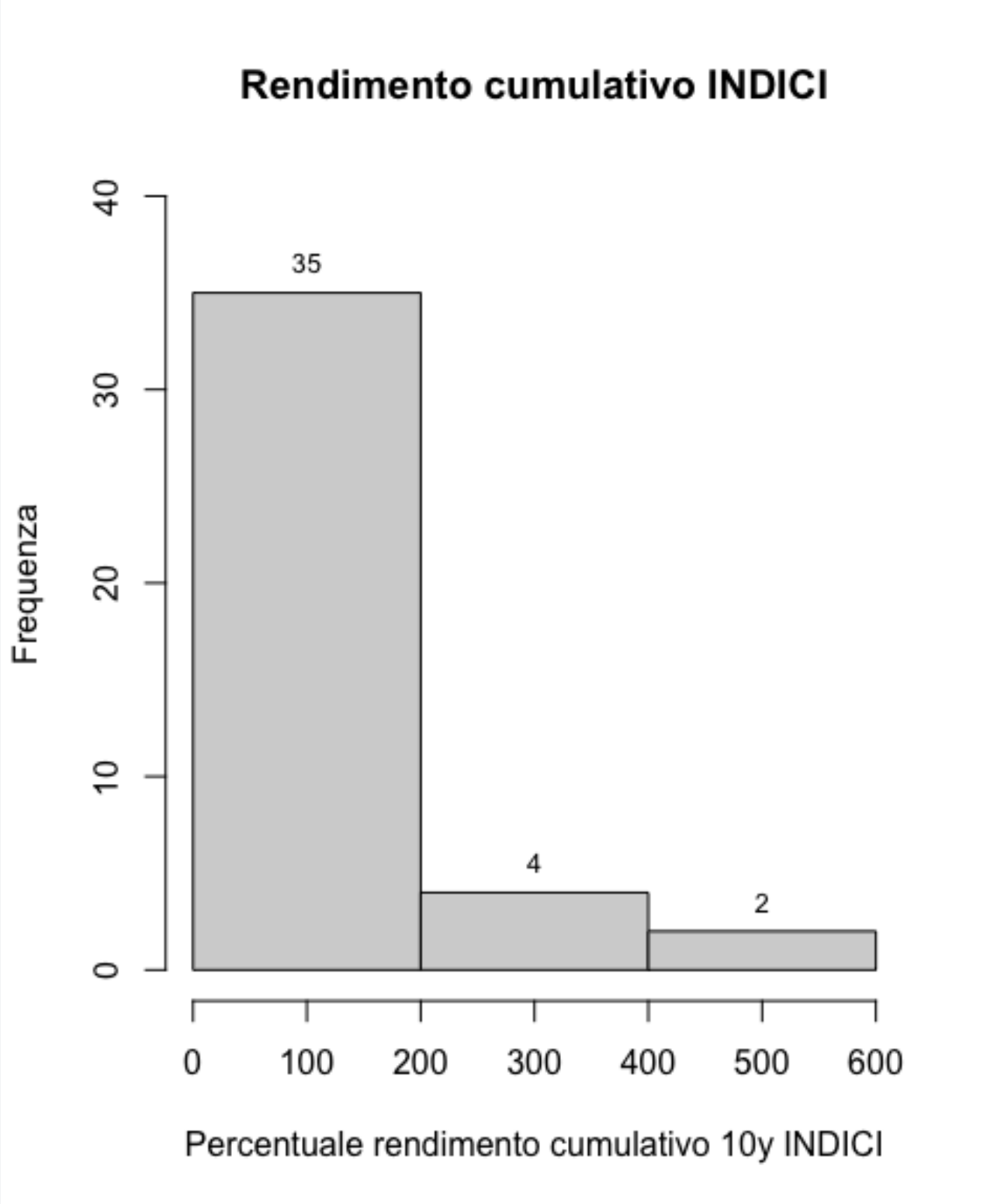
# IL PORTAFOGLIO DI FRANCESCO

FONDO	RENDIMENTO MEDIO ANNUO	RENDIMENTO CUMULATO	COSTI DI GESTIONE ANNUI	RENDIMENTO BENCHMARK MEDIO ANNUO	RENDIMENTO BENCHMARK CUMULATO	SCARTO ANNUALE MEDIO
8a+ Eiger R	3,95%	47,31%	2,00%	6,75%	92,17%	-2,80%
Eurizon AM Sicav High Yield Bond R	1,26%	13,34%	1,40%	2,63%	29,64%	-1,37%
8a+ SICAV - Eiger Class R	2,18%	24,07%	2,00%	6,75%	92,17%	-4,57%
8a+ Nextam Bilanciato R	2,18%	24,07%	1,65%	5,09%	64,29%	-2,91%
MEDIE	2,39%	27,20%	1,76%	5,31%	69,57%	-2,91%

# PRESENTAZIONE DEL CAMPIONE CASUALE

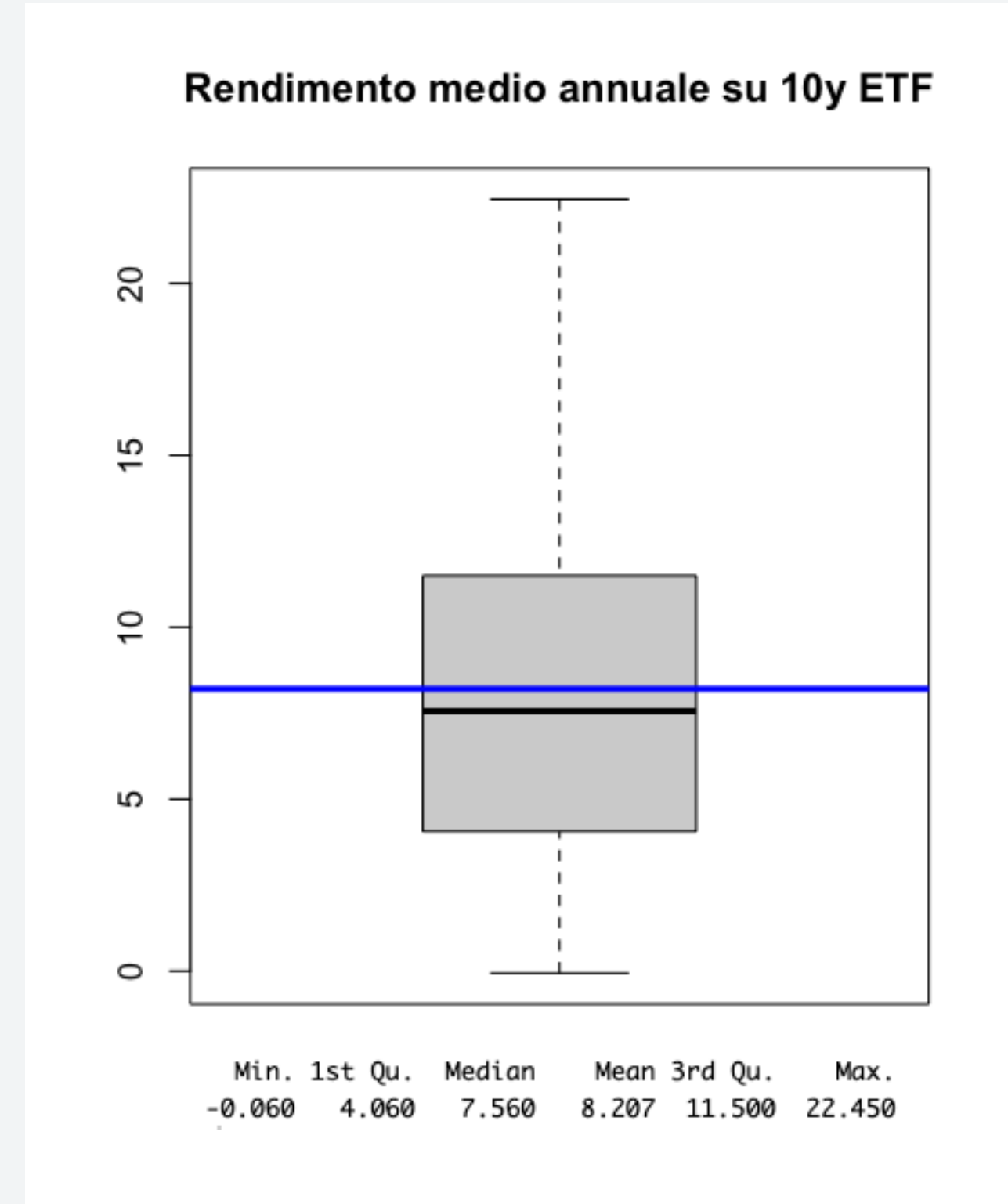
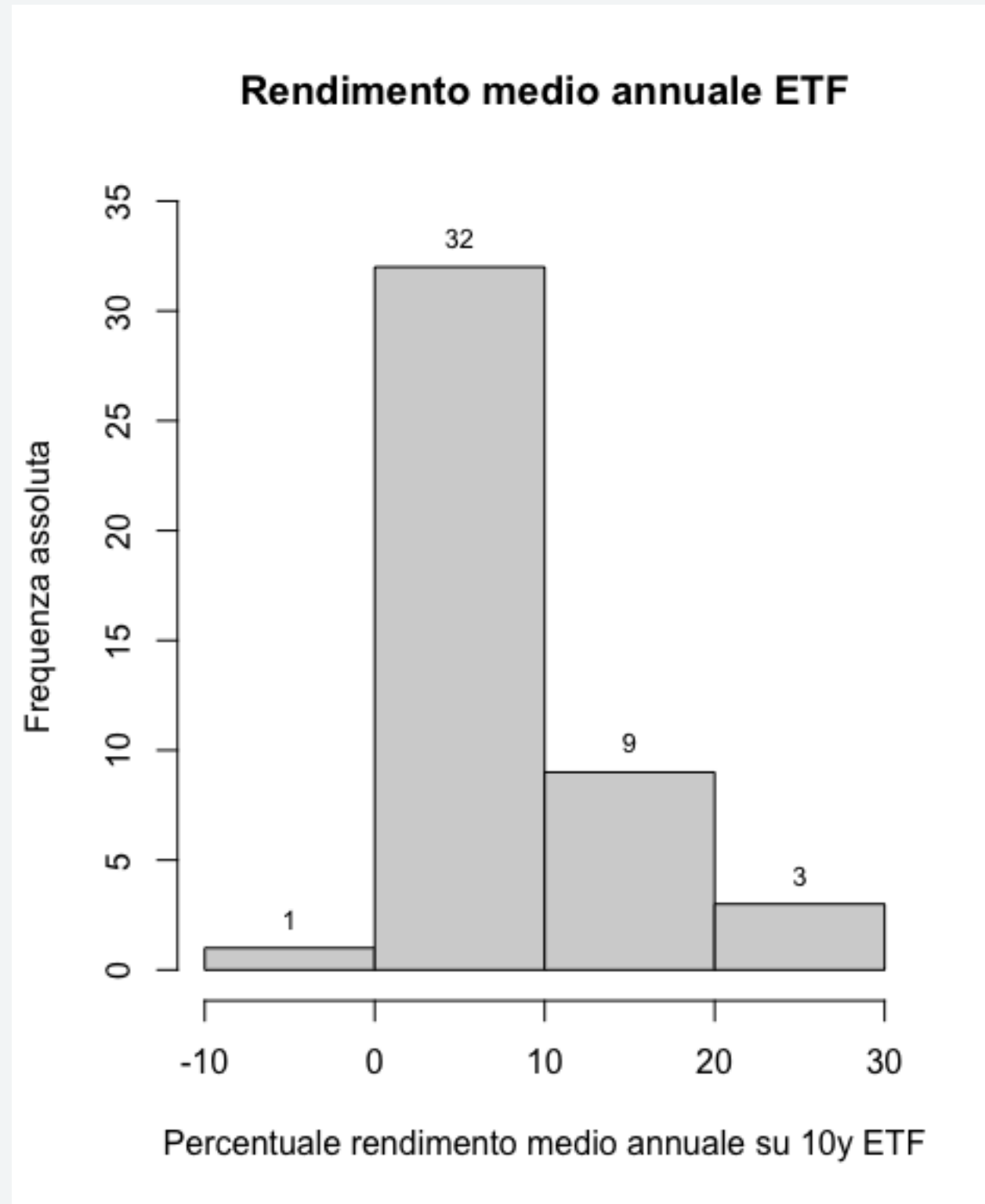


# PRESENTAZIONE DEL CAMPIONE CASUALE

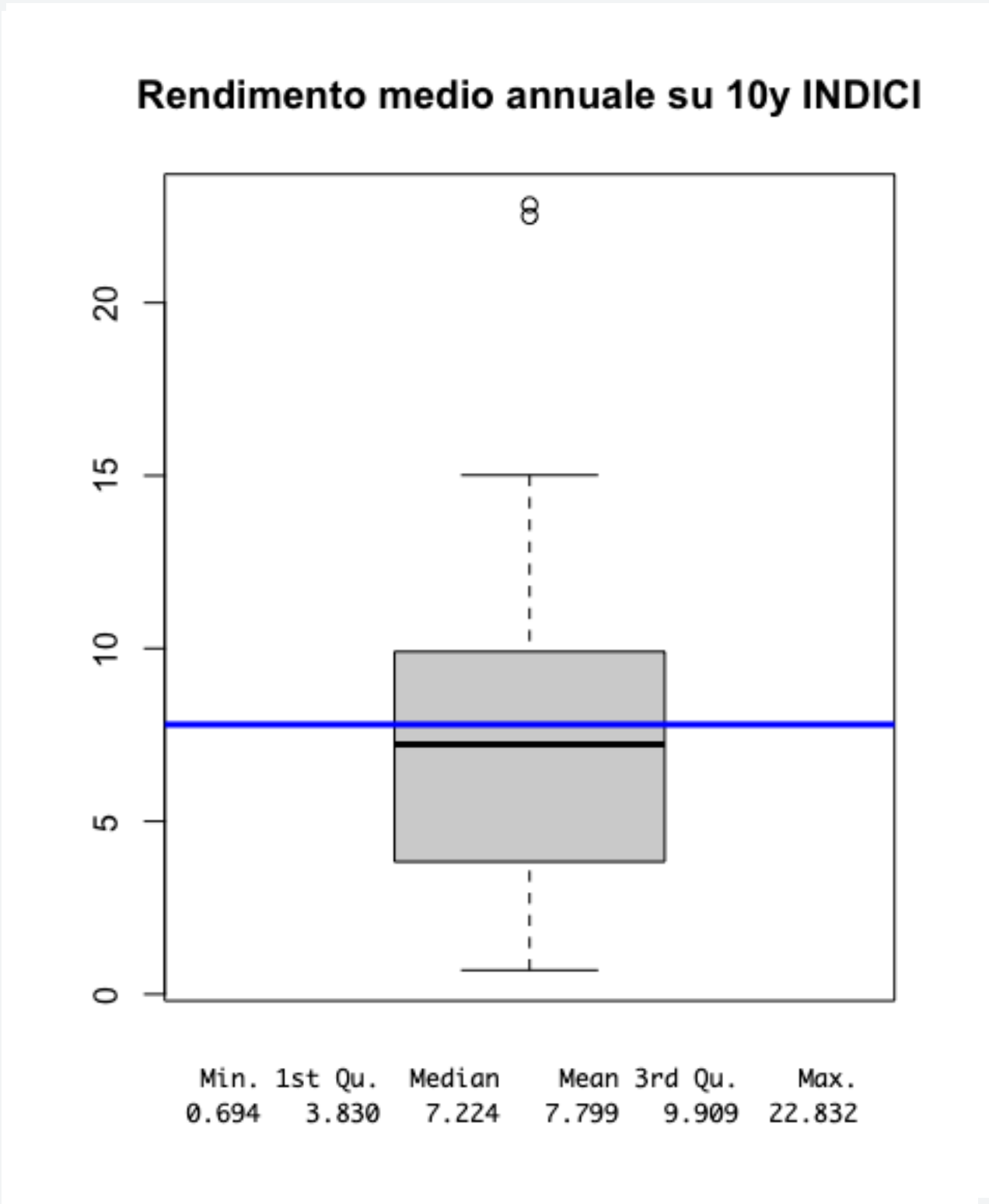
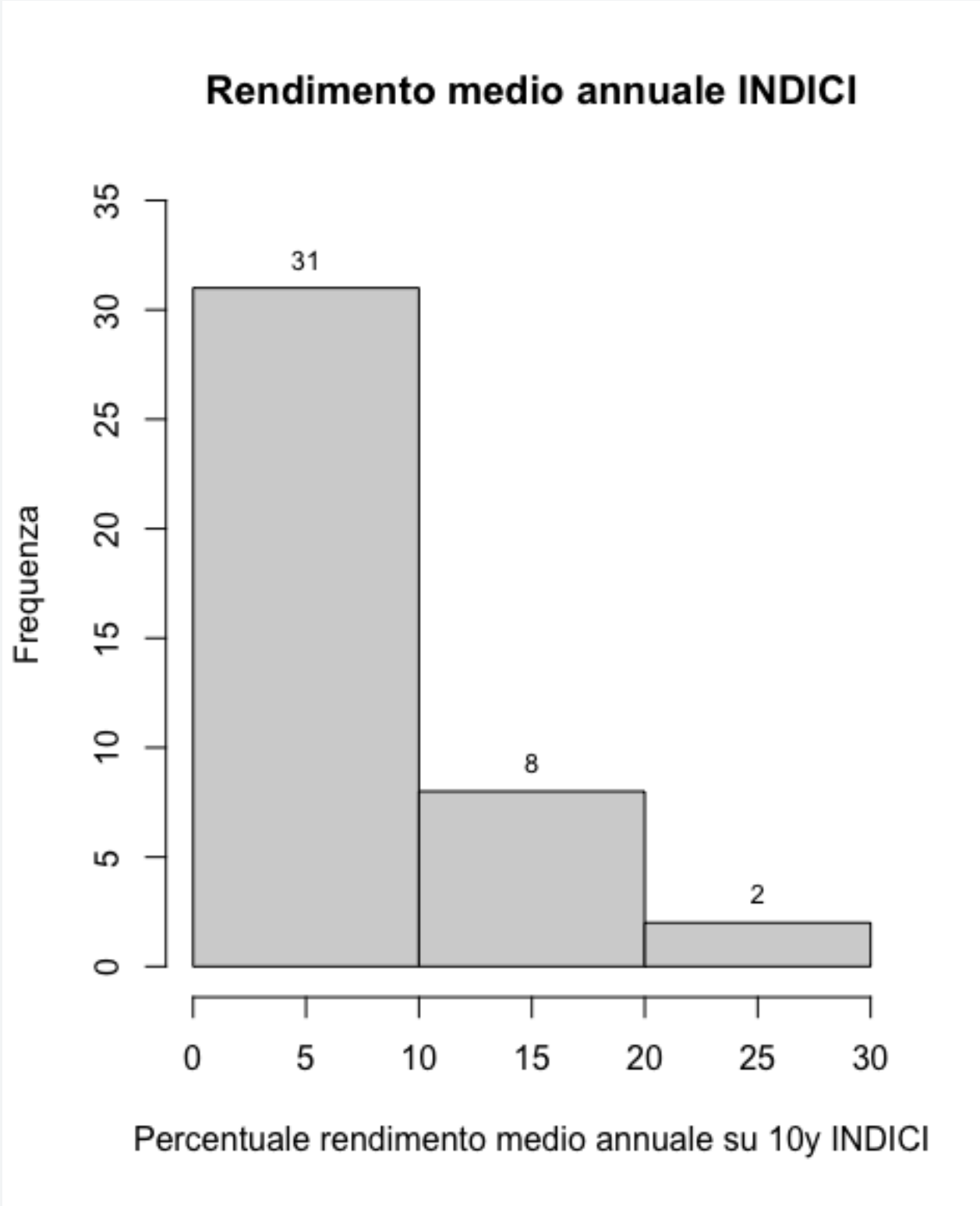




# PRESENTAZIONE DEL CAMPIONE CASUALE



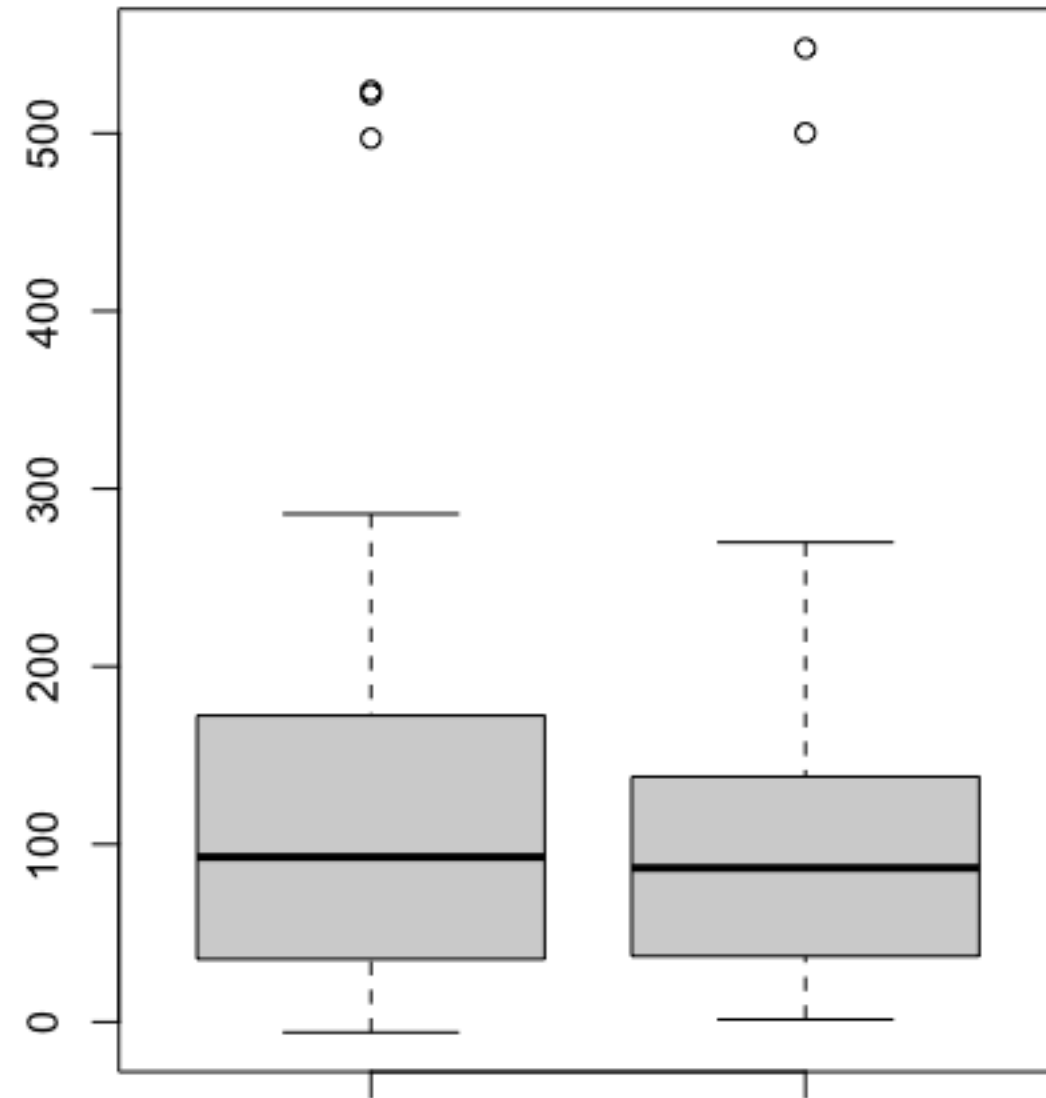
# PRESENTAZIONE DEL CAMPIONE CASUALE



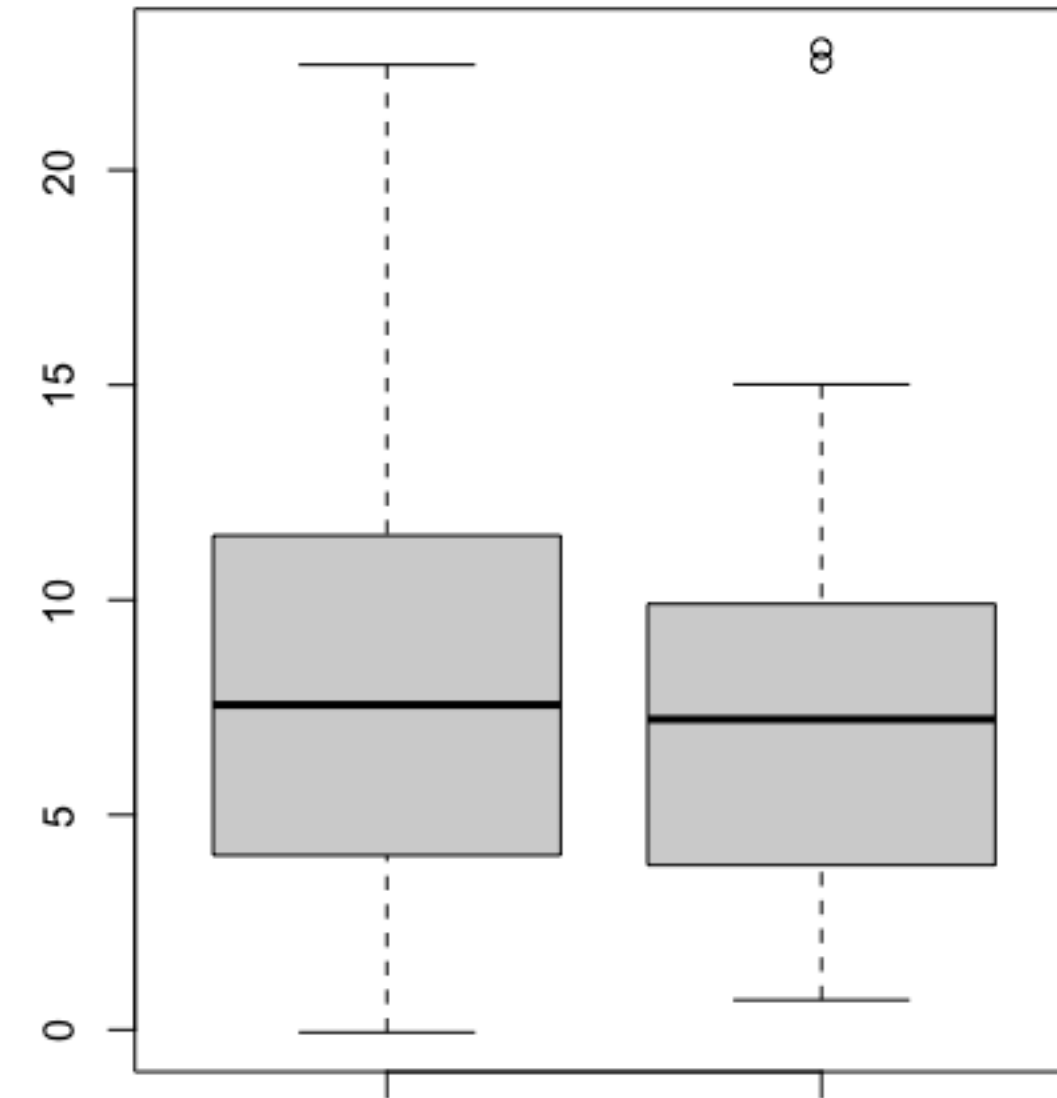


# PRESENTAZIONE DEL CAMPIONE CASUALE

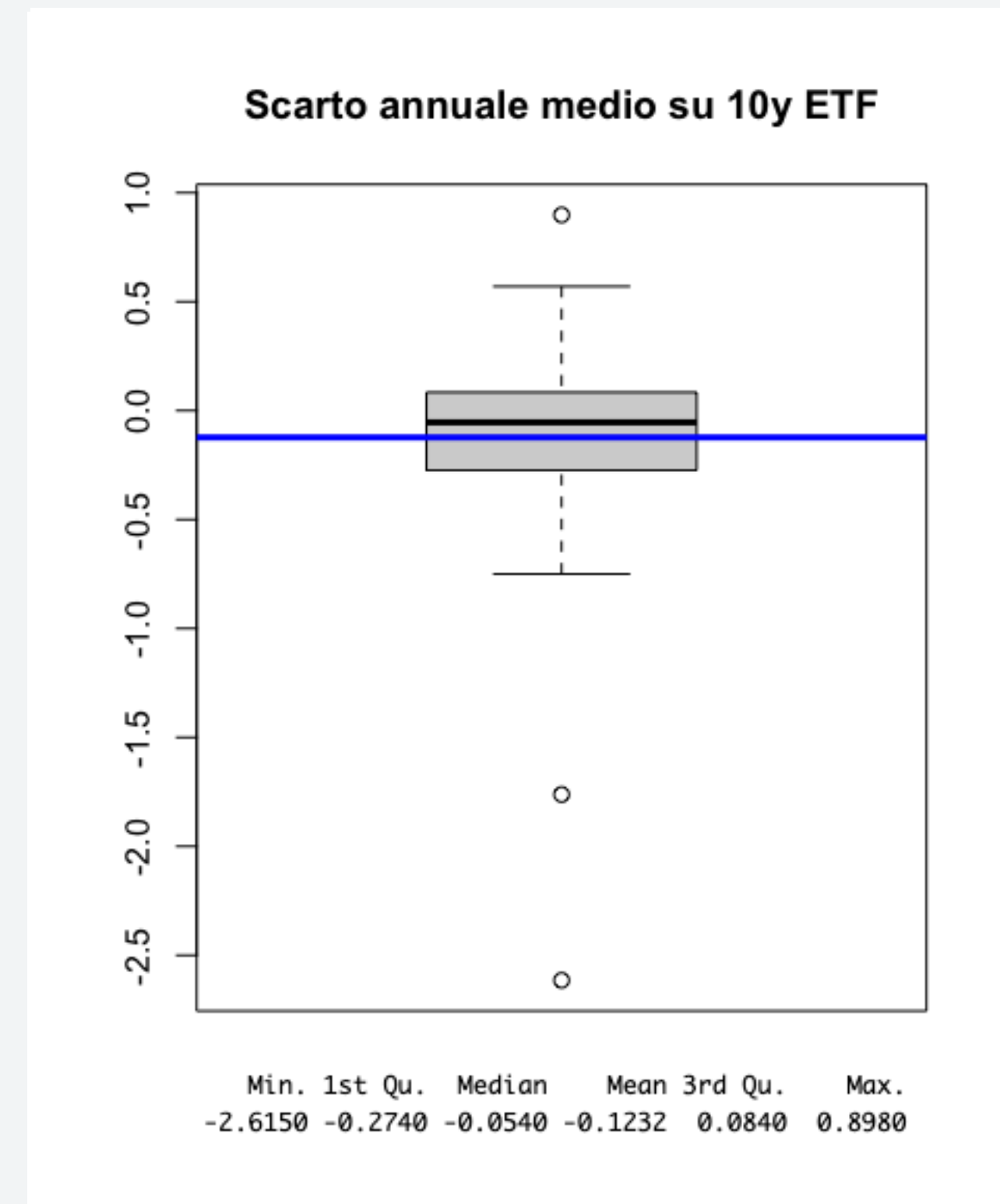
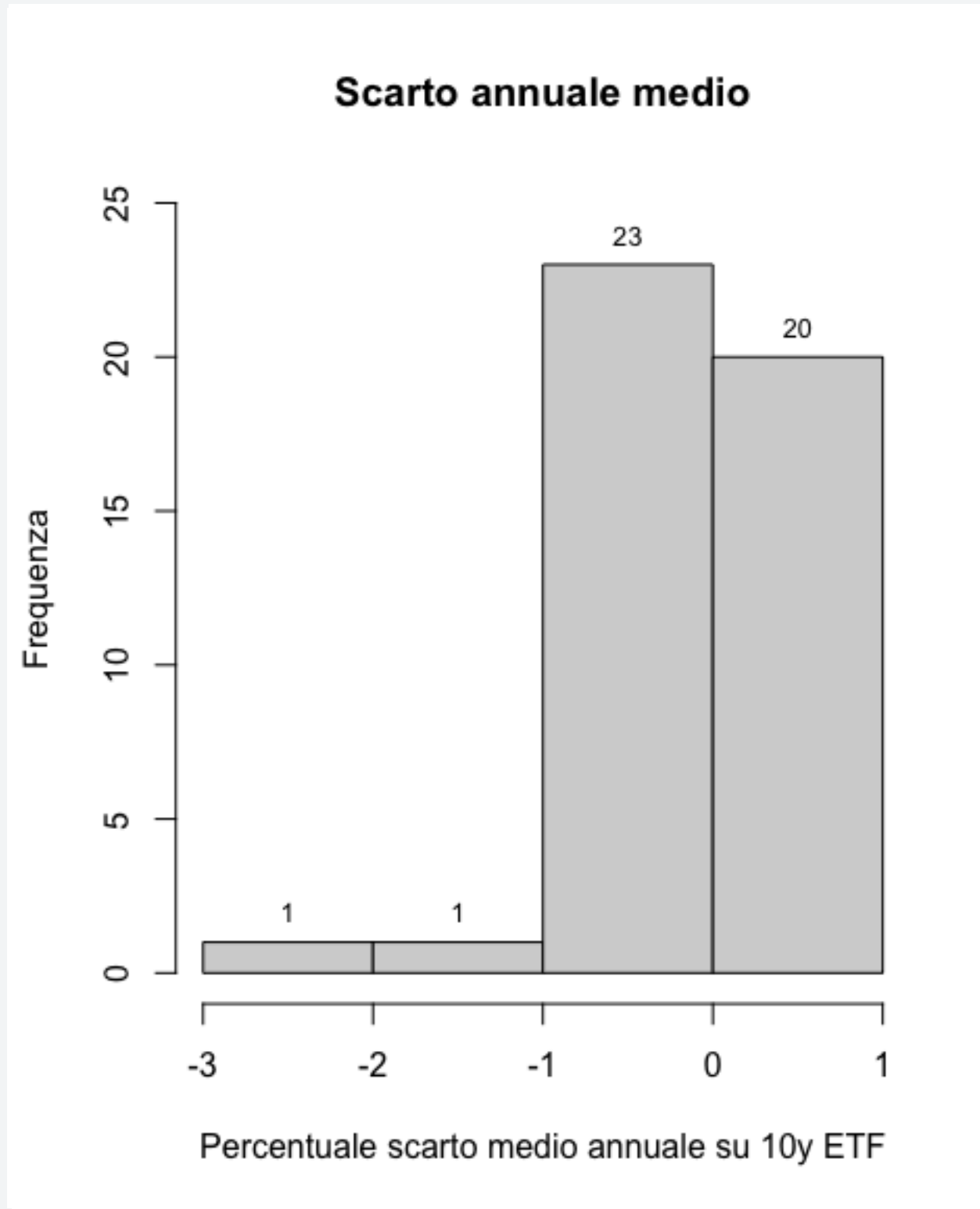
Rendimenti cumulativi a confronto



Rendimenti medi annuali a confronto

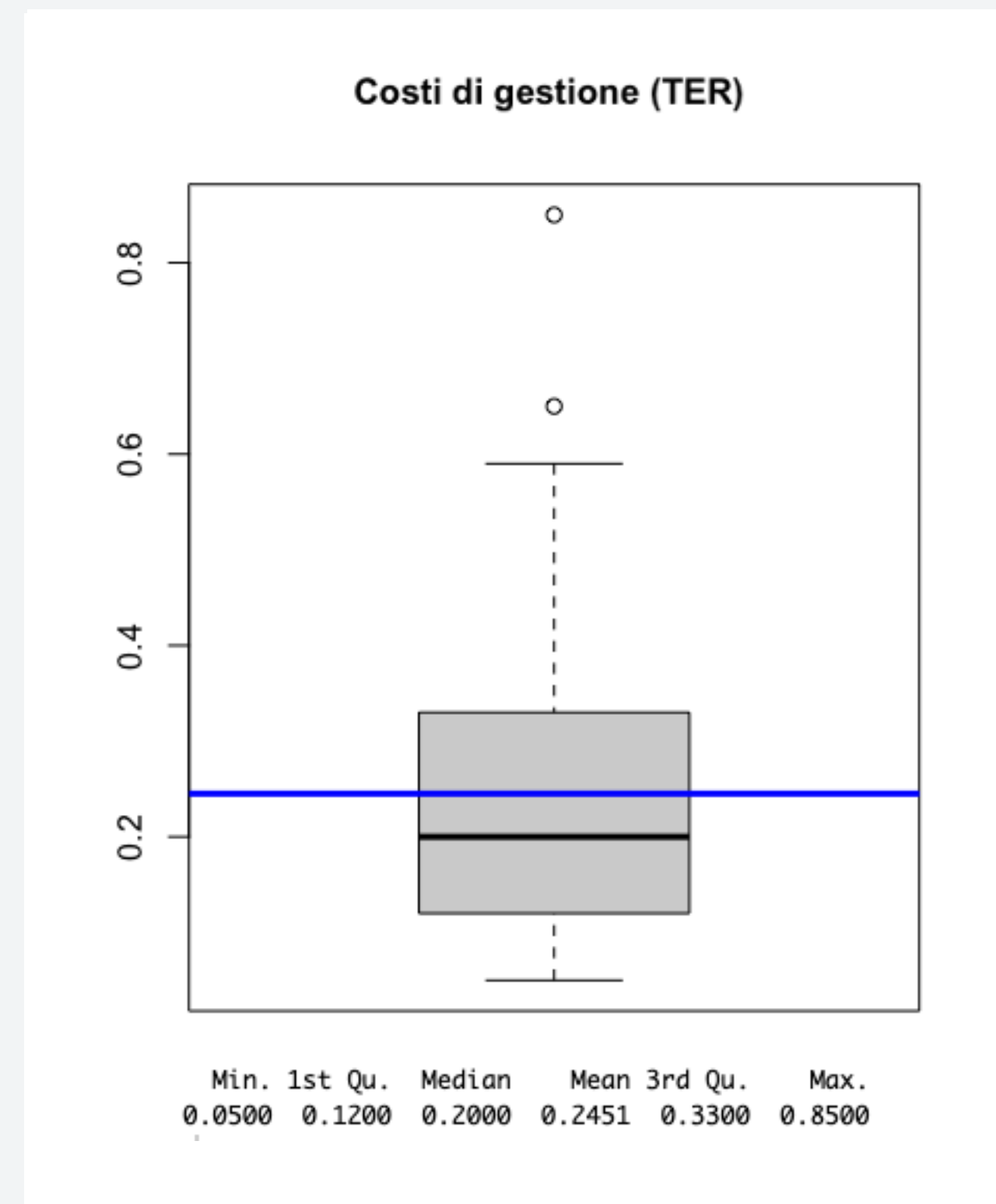
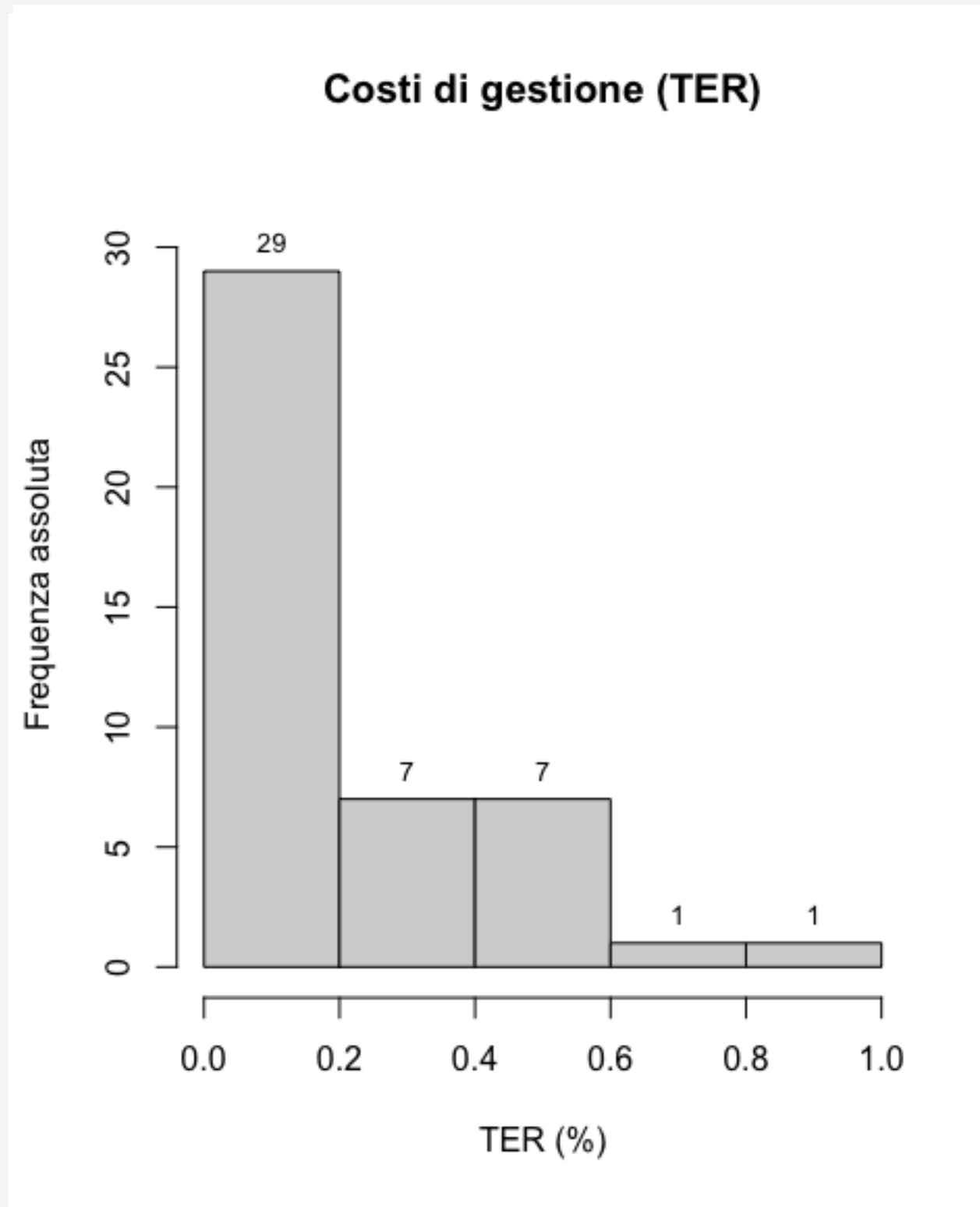


# PRESENTAZIONE DEL CAMPIONE CASUALE



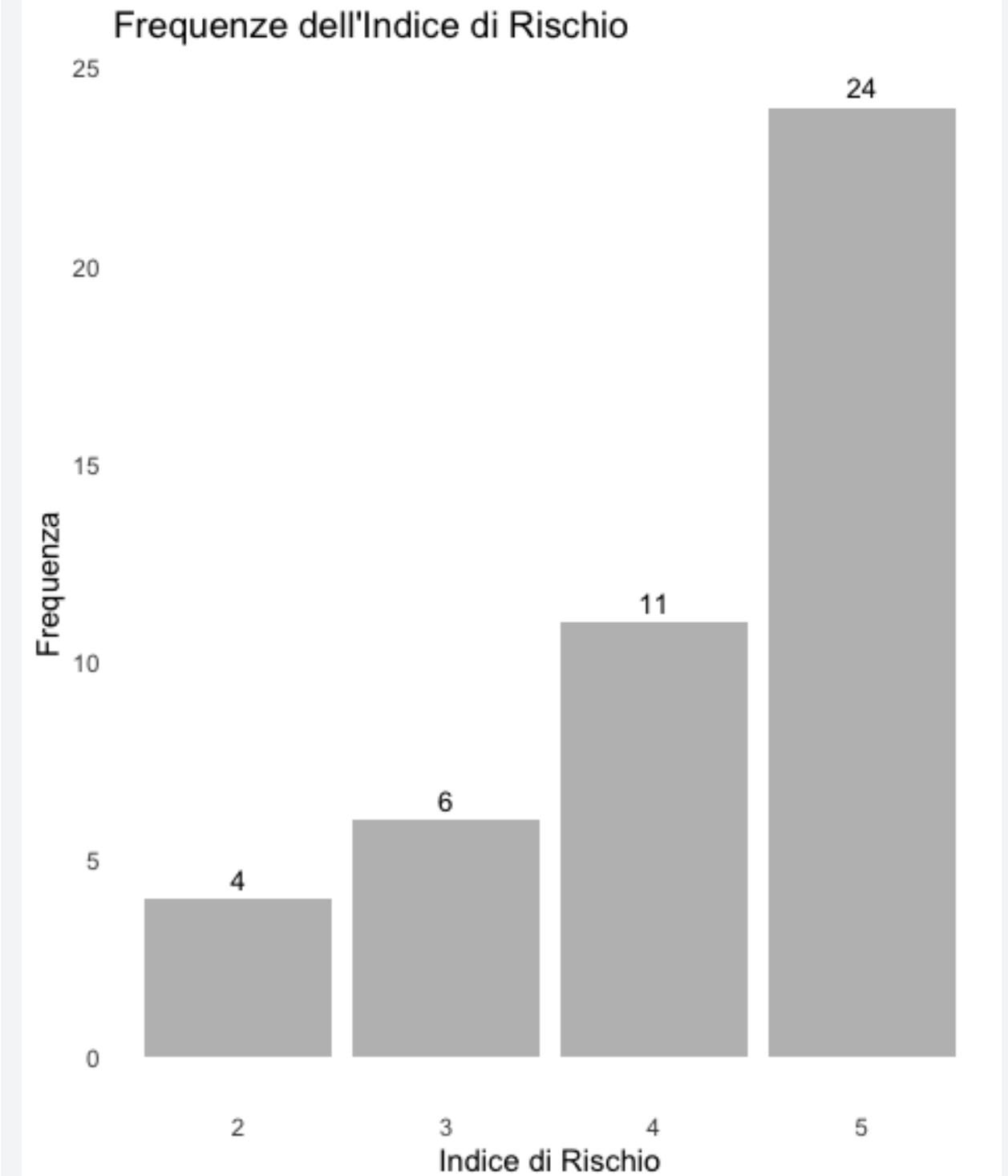
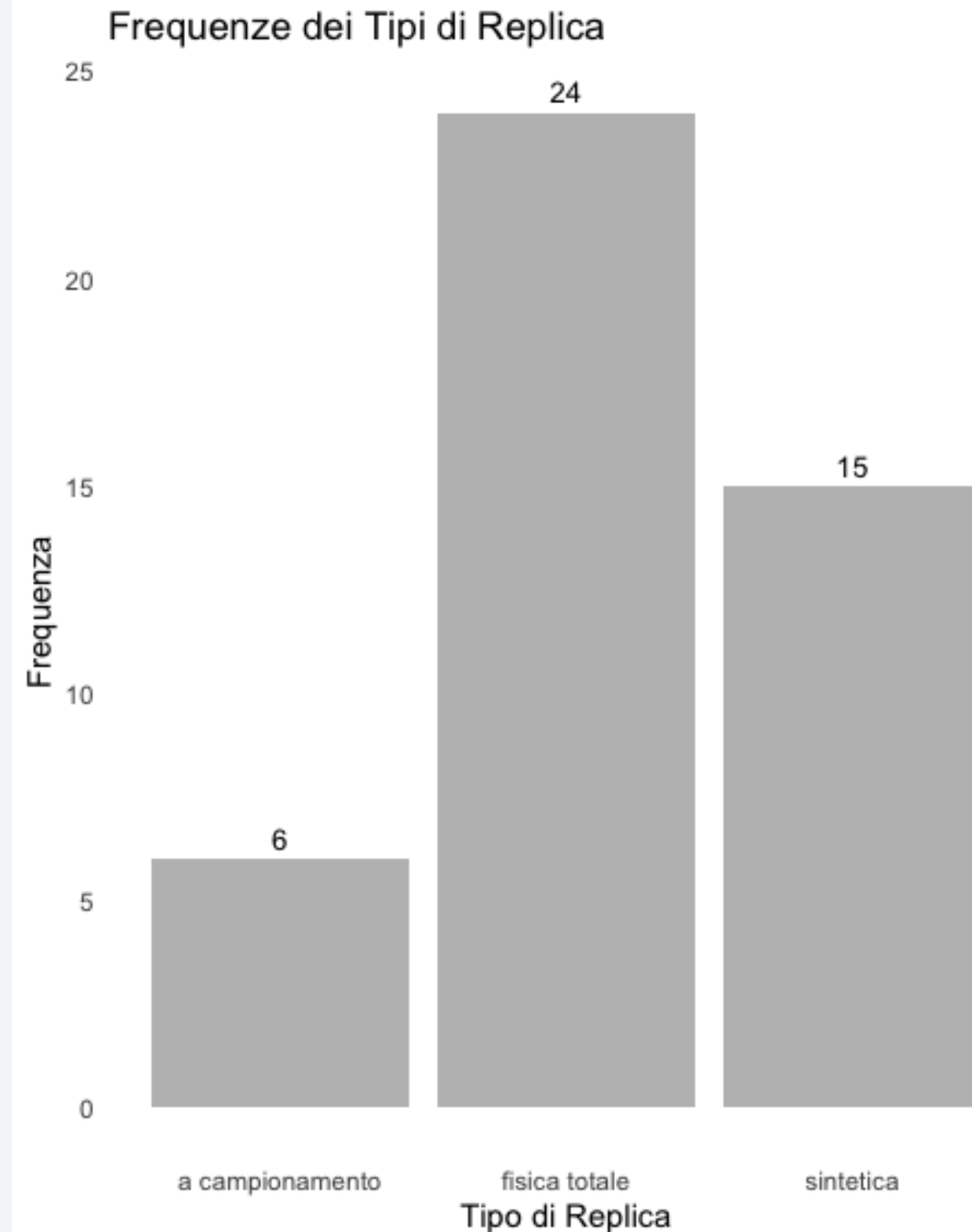
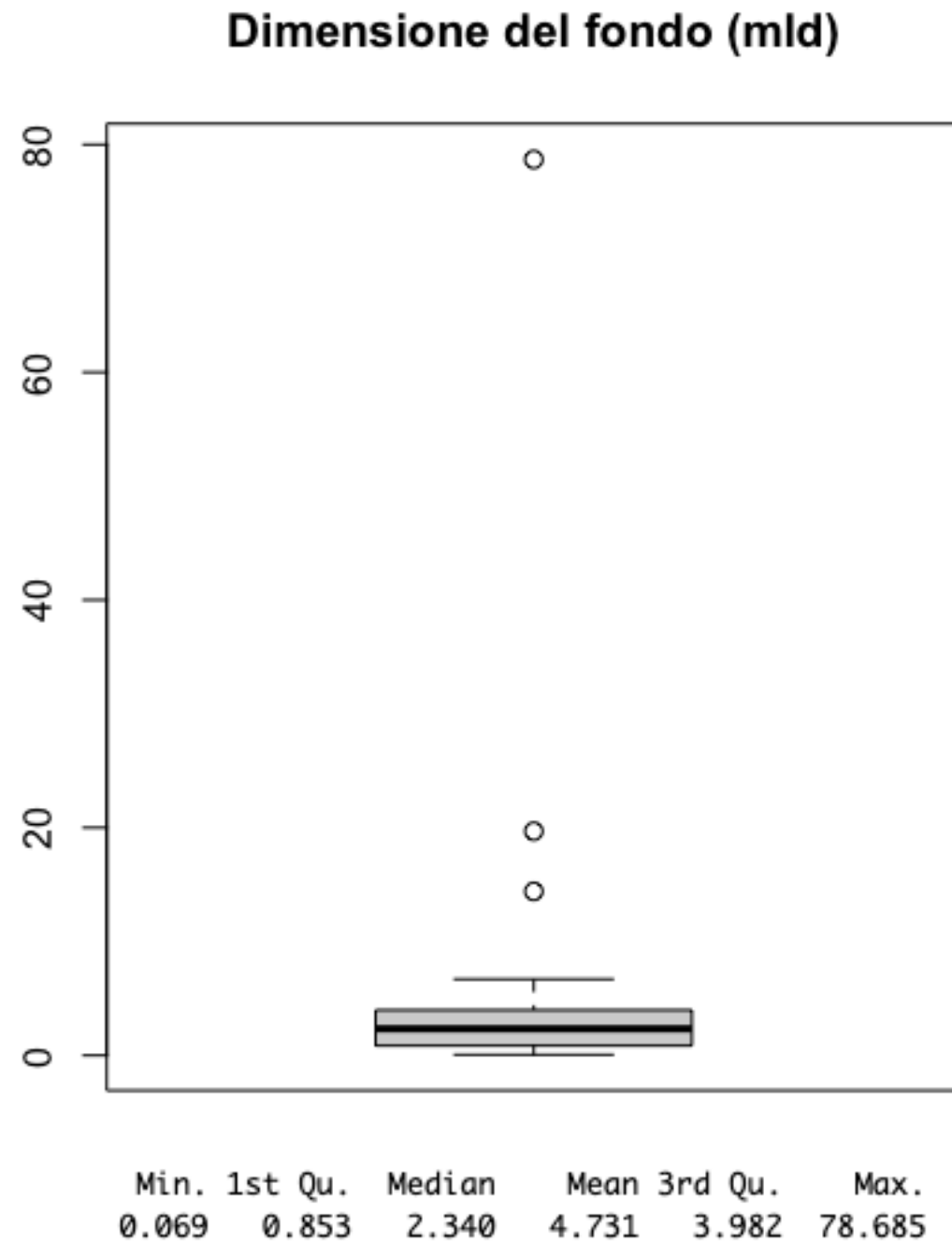
# 1. QUANTO COSTANO GLI ETF?

95 percent confidence interval:  
0.1899306 0.3002917



# 1. DA COSA DIPENDONO I COSTI DI GESTIONE?

## POSSIBILI REGRESSORI



# 1. DA COSA DIPENDONO I COSTI DI GESTIONE?

## MODELLO 1

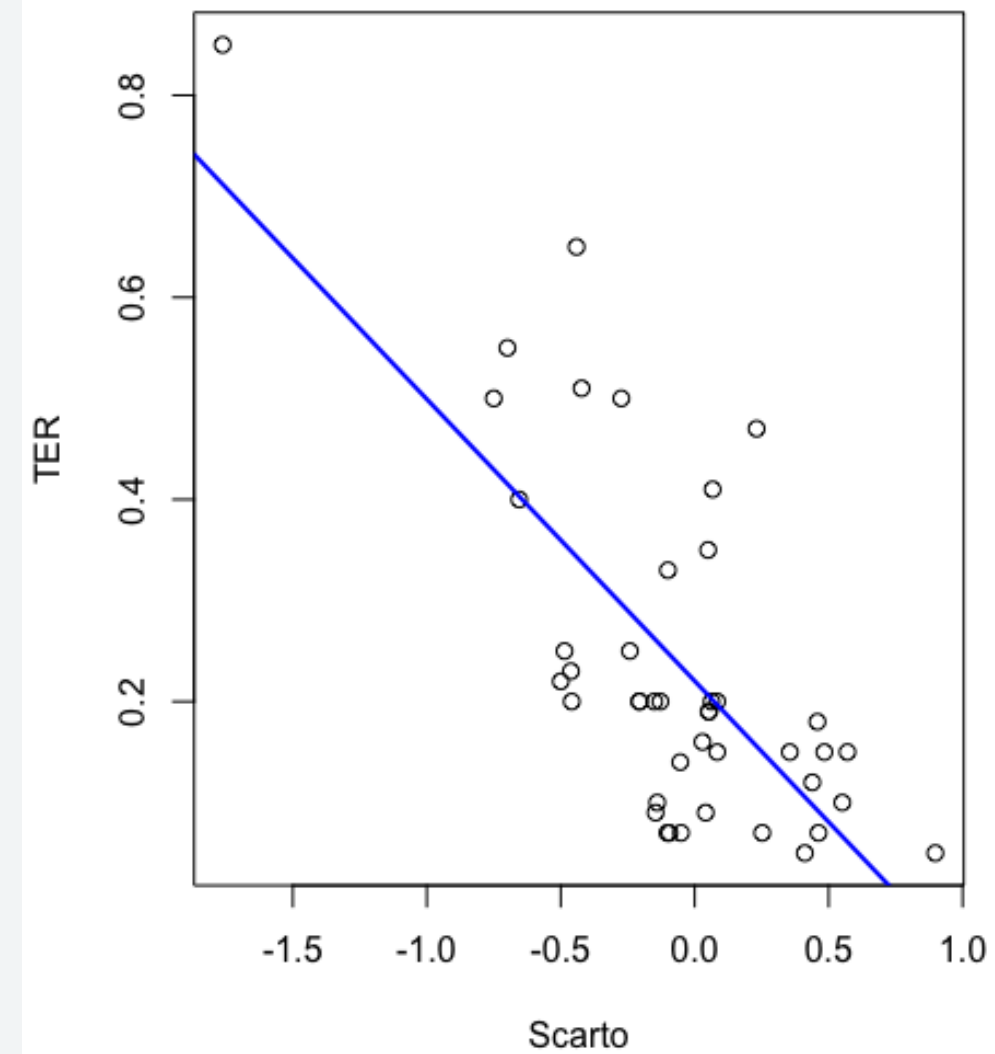
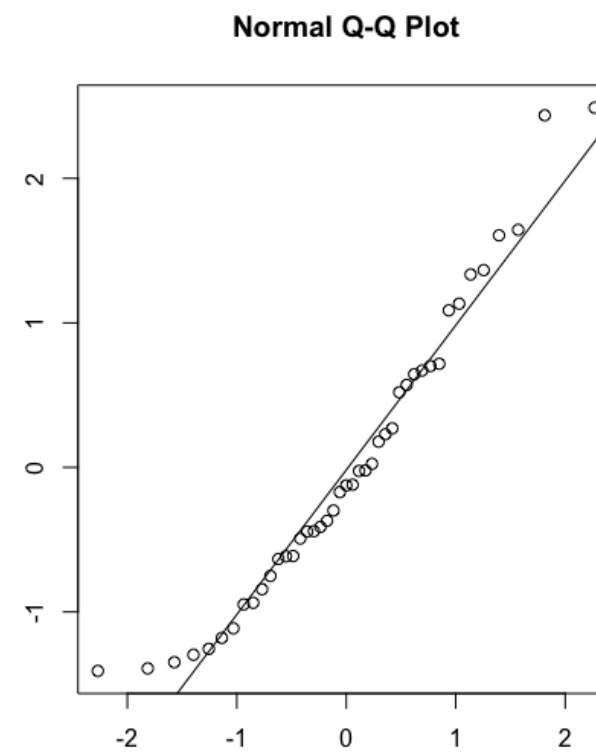
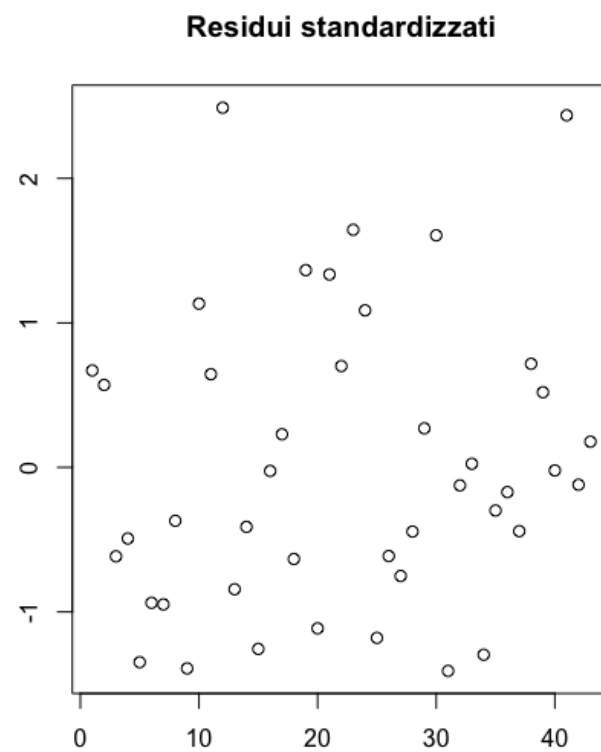
Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.22027	0.01980	11.124	5.87e-14 ***
Scarto	-0.27907	0.04345	-6.423	1.08e-07 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1284 on 41 degrees of freedom  
Multiple R-squared: 0.5016, Adjusted R-squared: 0.4894  
F-statistic: 41.26 on 1 and 41 DF, p-value: 1.082e-07



Shapiro-Wilk normality test

data: rstandard(ter\_vs\_scarto)  
W = 0.94922, p-value = 0.05585

# 1. DA COSA DIPENDONO I COSTI DI GESTIONE?

## MODELLO 2

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-0.53082	0.15199	-3.492	0.00132	**
Scarto	-0.26910	0.04126	-6.522	1.60e-07	***
Dimensione	0.22116	0.05111	4.327	0.00012	***
Tipo_di_replicafisica totale	0.85586	0.15681	5.458	4.00e-06	***
Tipo_di_replicasintetica	0.73783	0.14998	4.920	2.04e-05	***
Dimensione:Tipo_di_replicafisica totale	-0.24315	0.05191	-4.684	4.15e-05	***
Dimensione:Tipo_di_replicasintetica	-0.22220	0.05048	-4.402	9.62e-05	***

---

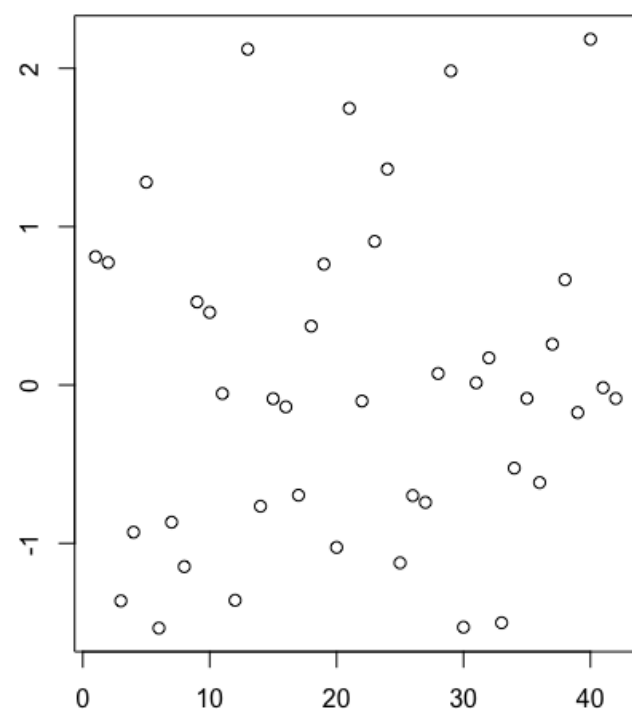
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1211 on 35 degrees of freedom

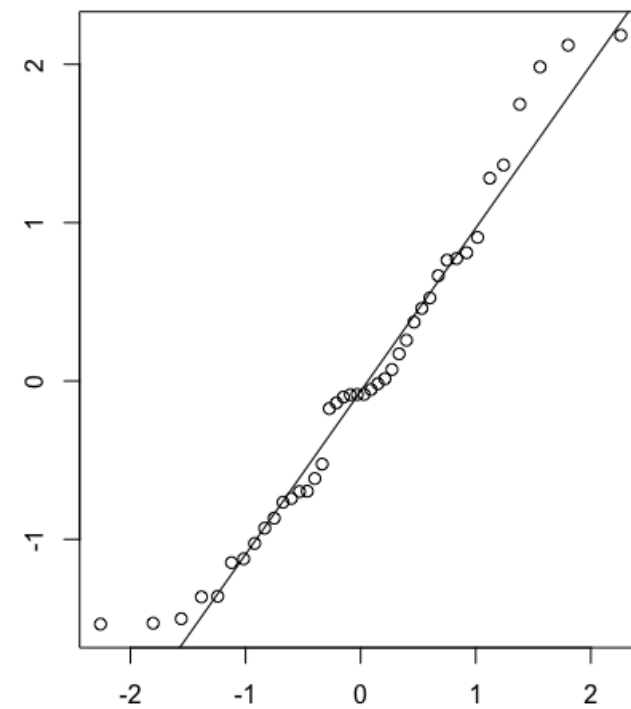
Multiple R-squared: 0.6148, Adjusted R-squared: 0.5488

F-statistic: 9.311 on 6 and 35 DF, p-value: 4.094e-06

Residui standardizzati



Normal Q-Q Plot



p-value = 0.1071

## MODELLO 3

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	0.01763	0.08189	0.215	0.8306	
Rischio	0.04578	0.01867	2.453	0.0188	*
Scarto	-0.28288	0.03931	-7.196	1.16e-08	***

---

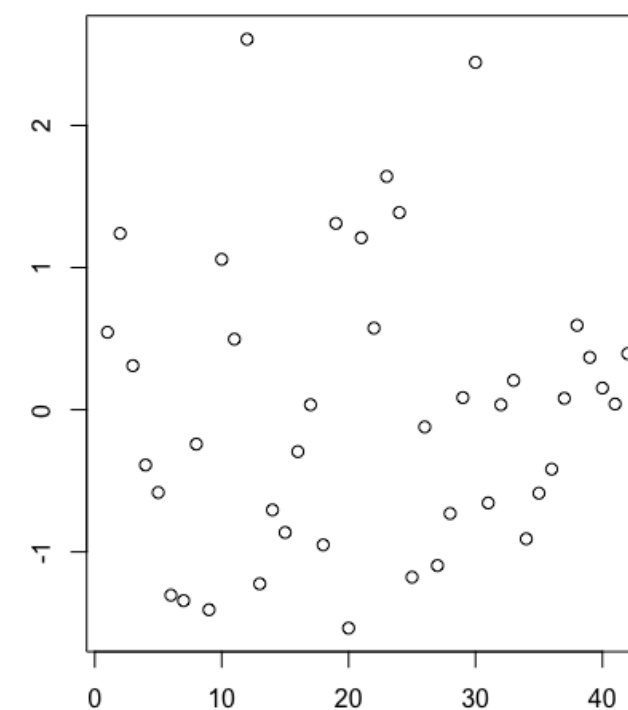
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1133 on 39 degrees of freedom

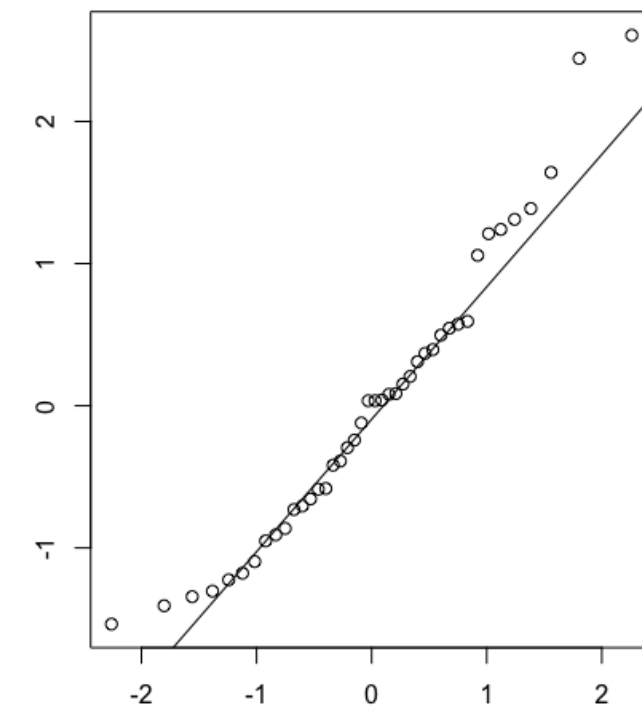
Multiple R-squared: 0.5767, Adjusted R-squared: 0.555

F-statistic: 26.57 on 2 and 39 DF, p-value: 5.234e-08

Residui standardizzati



Normal Q-Q Plot



p-value = 0.1077



# 1. QUANTO COSTEREBBE UN ETF CON FONDO DI DIMENSIONE 2MLD, SCARTO DI 0.05% RISPETTO ALL'INDICE E TIPO DI REPLICA “FISICA TOTALE”?

Mettendo a confronto i 3 modelli siamo arrivati alla conclusione che il modello 2 sia il migliore a causa di:

1. Ipotesi gaussiana verificata
2.  $R^2$  e  $R^2_{\text{corretto}}$  tra i più alti dei tre modelli
3. Significatività del modello
4. Significatività dei predittori
5. Outlier < 5%

Intervallo di confidenza al 95% per il valore atteso del modello 2 con dimensione = 2, scarto = 0.05% e tipo di replica = «fisica totale»

lwr	upr
0.2111689	0.3240669

## 2. QUANTO RENDE UN ETF IN MEDIA SU 10 ANNI?

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

95 percent confidence interval:
87.2204 165.1035



### 3. GLI ETF RAGGIUNGONO IL LORO OBIETTIVO?

Test di indipendenza per  $X_1$  = rendimento annuale medio ETF e  $X_2$  = rendimento medio annuale indice

	Var2_cat	
Var1_cat	[0.0069,0.118]	(0.118,0.228]
[-0.0006,0.112]	33	0
(0.112,0.225]	0	12

Pearson's Chi-squared test with Yates' continuity correction

```
data: contingency_table
X-squared = 40.032, df = 1, p-value = 2.499e-10
```

Test per differenza di media di popolazioni accoppiate

$\mu_1$  = media dei rendimenti annuali medi degli ETF

$\mu_2$  = media dei rendimenti annuali medi degli INDICI

$$H_0 : \mu_1 - \mu_2 \leq -0.30\%$$

$$H_1 : \mu_1 - \mu_2 > -0.30\%$$

```
data: record_vettori$RENDIMENTO.ANNUALE.MEDIO.ETF * 100 and record_vettori$RENDIMENTO.ANNUALE.MEDIO.INDICE * 100
t = 2.046, df = 44, p-value = 0.02338
alternative hypothesis: true mean difference is greater than -0.3
```

Intervallo di confidenza al 95% per la proporzione  $p$  = «probabilità che lo scarto tra il rendimento annuale medio dell'ETF e quello dell'indice sia  $\geq 0$ »

```
95 percent confidence interval:
 0.2995055 0.5987643
```

### 3. GLI ETF RAGGIUNGONO IL LORO OBIETTIVO?

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.10580	0.10285	-1.029	0.31
Rendimento_indice	1.00961	0.01028	98.172	<2e-16 ***

---

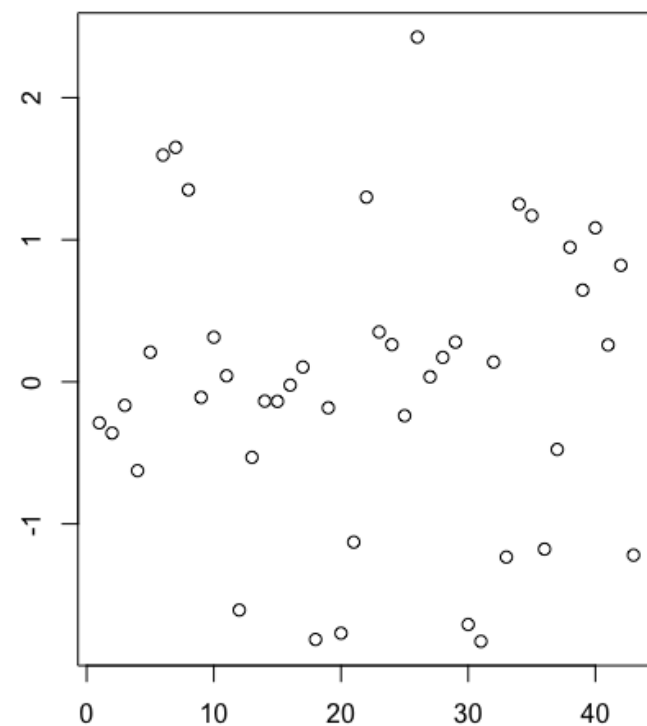
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3718 on 41 degrees of freedom

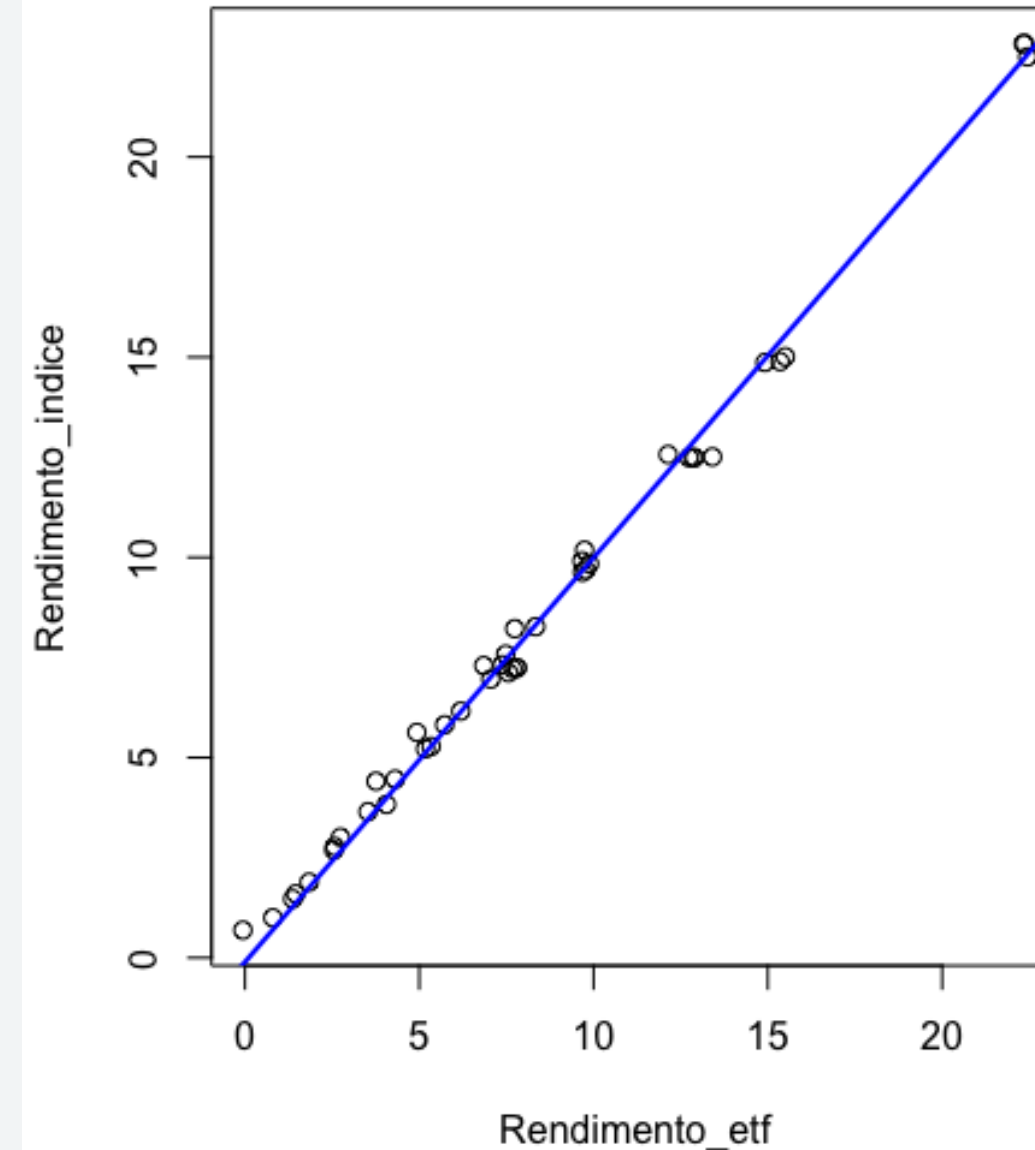
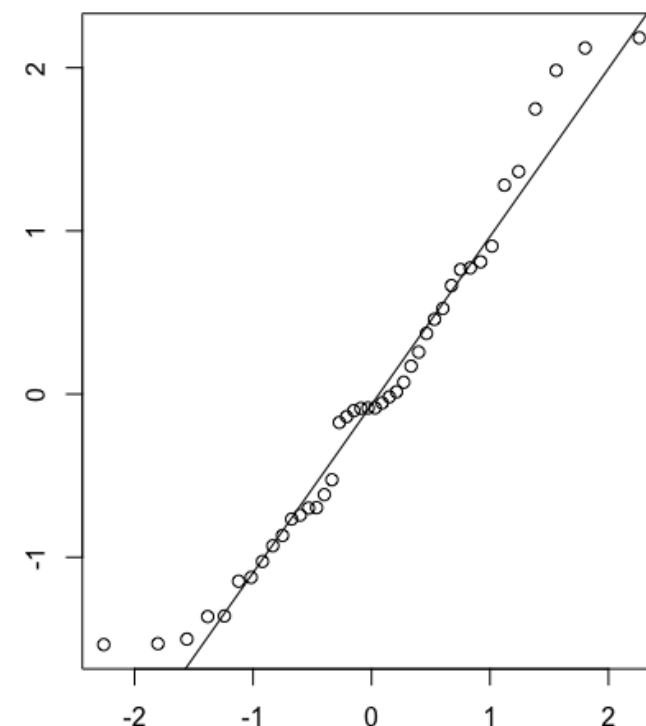
Multiple R-squared: 0.9958, Adjusted R-squared: 0.9957

F-statistic: 9638 on 1 and 41 DF, p-value: < 2.2e-16

Residui standardizzati



Normal Q-Q Plot



Shapiro-Wilk normality test

data: rstandard(reg2)  
W = 0.96872, p-value = 0.2854

# 3. MEDIAMENTE UN ETF FA MEGLIO DEL PORTAFOGLIO DI FRANCESCO?

Test per media con varianza incognita  
sul rendimento medio annuale degli ETF

$$H_0: \mu \leq 2.39\% \quad H_1: \mu > 2.39\%$$

```
data: record_etf$RENDIMENTO.MEDIO..ANNUALE. * 100  
t = 6.8816, df = 44, p-value = 8.526e-09  
alternative hypothesis: true mean is greater than 2.39
```

Test per media con varianza incognita  
sul TER degli ETF

$$H_0: \mu \geq 1.76\% \quad H_1: \mu < 1.76\%$$

```
data: record_etf$X * 100  
t = -55.329, df = 44, p-value < 2.2e-16  
alternative hypothesis: true mean is less than 1.76
```

Test per media con varianza incognita  
sul rendimento medio cumulato degli ETF

$$H_0: \mu \leq 27.2\% \quad H_1: \mu > 27.2\%$$

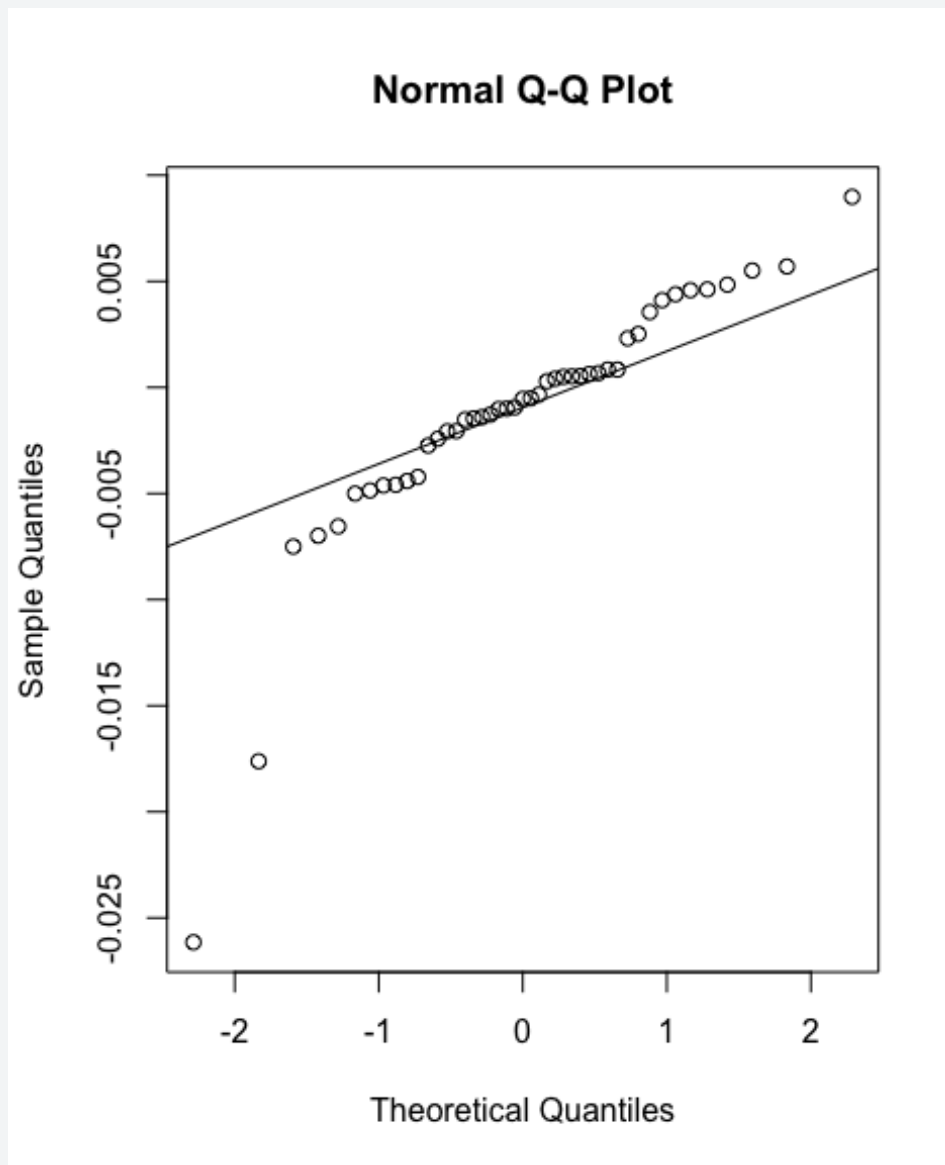
```
data: record_etf$RENDIMENTO.CUMULATIVO * 100  
t = 5.1216, df = 44, p-value = 3.223e-06  
alternative hypothesis: true mean is greater than 27.2
```

Test per media con varianza incognita  
sullo scarto medio annuale degli ETF

$$H_0: \mu \leq -2.91\% \quad H_1: \mu > -2.91\%$$

```
data: record_etf$SCARTO.ANNUALE.MEDIO * 100  
t = 31.924, df = 44, p-value < 2.2e-16  
alternative hypothesis: true mean is greater than -2.91
```

# 4. SE DOVESSE COMPRARE UN ETF A CASO, CHE SCARTO MEDIO ANNUALE POTREBBE AVERE QUEST'ULTIMO SU 10 ANNI?



Shapiro-Wilk normality test

```
data: record_etf$SCARTO.ANNUALE.MEDIO
W = 0.82379, p-value = 8.279e-06
```

Intervallo di predizione per una nuova osservazione dello scarto annuale medio:

Intervallo di predizione al 95%:( -1.195299 , 1.192828 )

Come possiamo notare dal grafico dei quantili e dal p-value di SW l'intervallo di predizione non è affidabile.

# 5. FRANCESCO DECIDE DI CREARSI UN PORTAFOGLIO SCEGLIENDO CASUALMENTE 40 ETF E VUOLE SAPERE LA PROBABILITA' CHE ALMENO 30 DI QUESTI ABBIANO REPLICATO L'INDICE

$X$  = numero di ETF che replicano l'indice su 40 ETF pescate

$X \sim \text{Bin}(40, p)$

Per replica dell'indice si intende uno scarto annuale medio  $> -0.3\%$

Successi: 34 (su 45)

Intervallo di confidenza al 95% per la proporzione  $p$ :

```
95 percent confidence interval:  
0.6013926 0.8661495
```

Test per proporzione  $p=0.8$ :

$H_0: p = 0.8$

$H_1: p \neq 0.8$

```
data: 34 out of 45, null probability 0.8  
X-squared = 0.55556, df = 1, p-value = 0.4561  
alternative hypothesis: true p is not equal to 0.8
```

$X \sim \text{Bin}(40, 0.8)$ :

$P(X \geq 30)$ : 73.17771 %



# 6. COME HANNO REAGITO GLI INDICI AGLI EVENTI PIU' SIGNIFICATIVI DEGLI ULTIMI 10 ANNI?

Probabilità annuale media di trovare un rendimento negativo = 0.27561

Tra il 2014 e il 2023, a nostro parere, ci sono stati tre eventi cardine:

1. Dazi (2018)
2. Covid (2020)
3. Guerra (2022)

Attraverso dei test andiamo a vedere se negli anni in cui si sono verificati questi eventi ci sono stati dei cambiamenti dei rendimenti degli indici che per noi possono essere imputabili agli eventi sopra citati:

- Neutro
- Negativamente

2018 (Probabilità  $\hat{p}$ = 0.85366)

$H_0: p = 0.27561$        $H_1: p \neq 0.27561$

```
data: 35 out of 41, null probability 0.27561
X-squared = 68.619, df = 1, p-value < 2.2e-16
alternative hypothesis: true p is not equal to 0.27561
```

2020 (Probabilità  $\hat{p}$ = 0.24390)

$H_0: p = 0.27561$        $H_1: p \neq 0.27561$

```
data: 10 out of 41, null probability 0.27561
X-squared = 0.20646, df = 1, p-value = 0.6496
alternative hypothesis: true p is not equal to 0.27561
```

2022 (Probabilità  $\hat{p}$ = 0.92683)

$H_0: p = 0.27561$        $H_1: p \neq 0.27561$

```
data: 38 out of 41, null probability 0.27561
X-squared = 87.091, df = 1, p-value < 2.2e-16
alternative hypothesis: true p is not equal to 0.27561
```

**GRAZIE PER  
L'ATTENZIONE**



**POLITECNICO**  
MILANO 1863

