ESERCIZIO SU TEST STATISTICO

Vogliamo verificare se il costo medio di un litro di latte è pari a 1 €, come dichiarato nel paniere per il calcolo dell'inflazione. Decidiamo una regione di accettazione di ±0.1 €. Andiamo quindi in 4 supermercati a controllare i prezzi di 10 marche differenti. Il valor medio dei 40 dati vale 1.12 €, con una deviazione standard campionaria pari a 0.23 € (che si può considerare la deviazione standard del processo, dato l'elevato numero di campioni).

- Calcolare il livello di significatività del test.
- Se il costo medio reale di un litro di latte fosse 1.15 €, quanto varrebbe la probabilità b di errore di tipo II?

SOLUZIONE

Ricordiamo che il **livello di significatività** di un test statistico è la probabilità di rifiutare l'ipotesi nulla, quando questa è vera. Si denota tipicamente con la lettera greca a, e corrisponde alla probabilità di commettere un errore di tipo I:

Livello di significatività a = P(errore di tipo I) = P(rifiutare H_0 quando H_0 è vera)

Il livello di significatività del test si ottiene calcolando la probabilità che la media campionaria esca dalla regione di accettazione (1±0.1€ nel nostro caso), supponendo vera l'ipotesi nulla, che in questo caso è $H_0 : m = 1$€. (Ricordiamo che per definizione l'ipotesi nulla è puntuale, poi nel test si decideranno le regioni di accettazione e di rifiuto).

La deviazione standard della media campionaria vale $s_{\overline{X}} = \frac{s_X}{\sqrt{n}} = \frac{0.23}{\sqrt{40}}$ € $\cong$ 0.036 €, avendo considerato $s_X = S_X$ come suggerito nel testo.

Per calcolare la probabilità standardizziamo la variabile casuale (gaussiana per il teorema del limite centrale) e ricorriamo quindi alla tabella dei valori della distribuzione cumulativa normale standard

$$a = P\left(\overline{X} < 0.9 \text{ o } \overline{X} > 1.1 \text{ con } m = 1\right) = P\left(z < \frac{0.9 - m}{s_{\overline{X}}}\right) + P\left(z > \frac{1.1 - m}{s_{\overline{X}}}\right) =$$

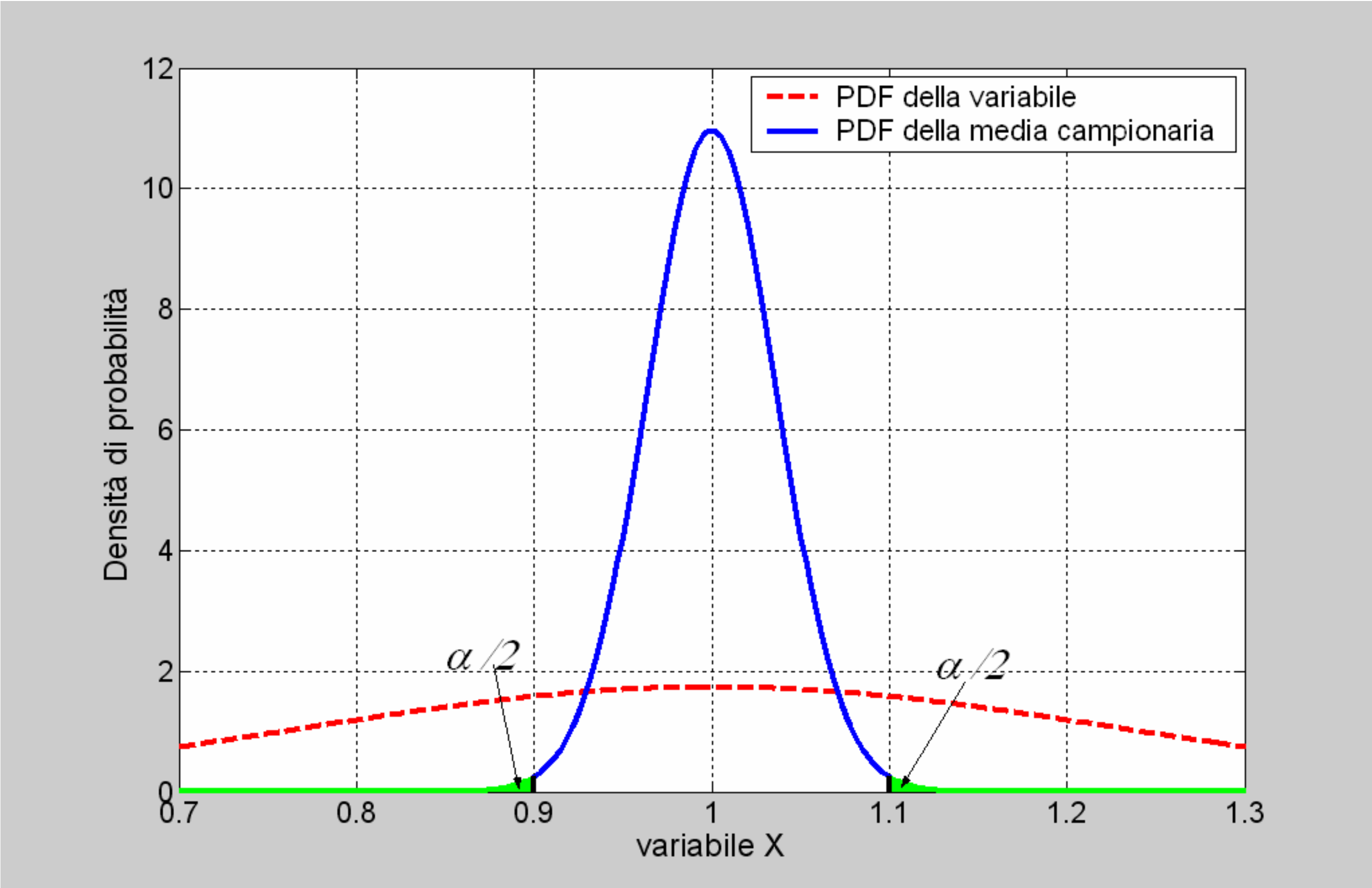
$$= P\left(z < \frac{0.9 - 1}{0.036}\right) + P\left(z > \frac{1.1 - 1}{0.036}\right) = P(z < -2.78) + P(z > 2.78) = 2 \cdot P(z < -2.78) \cong$$

$$\cong 2 \cdot 0{,}00272 = 0.00544$$

Il livello di significatività del test è pari dunque allo 0.54 %.

Riportiamo in figura il significato grafico di a.

In questo caso il test ha avuto come risultato il rifiuto dell'ipotesi nulla (perché la media campionaria misurata, pari a 1.12 €, cade al di fuori della regione di accettazione). Se l'ipotesi nulla fosse stata vera si sarebbe verificato un caso particolarmente sfortunato.



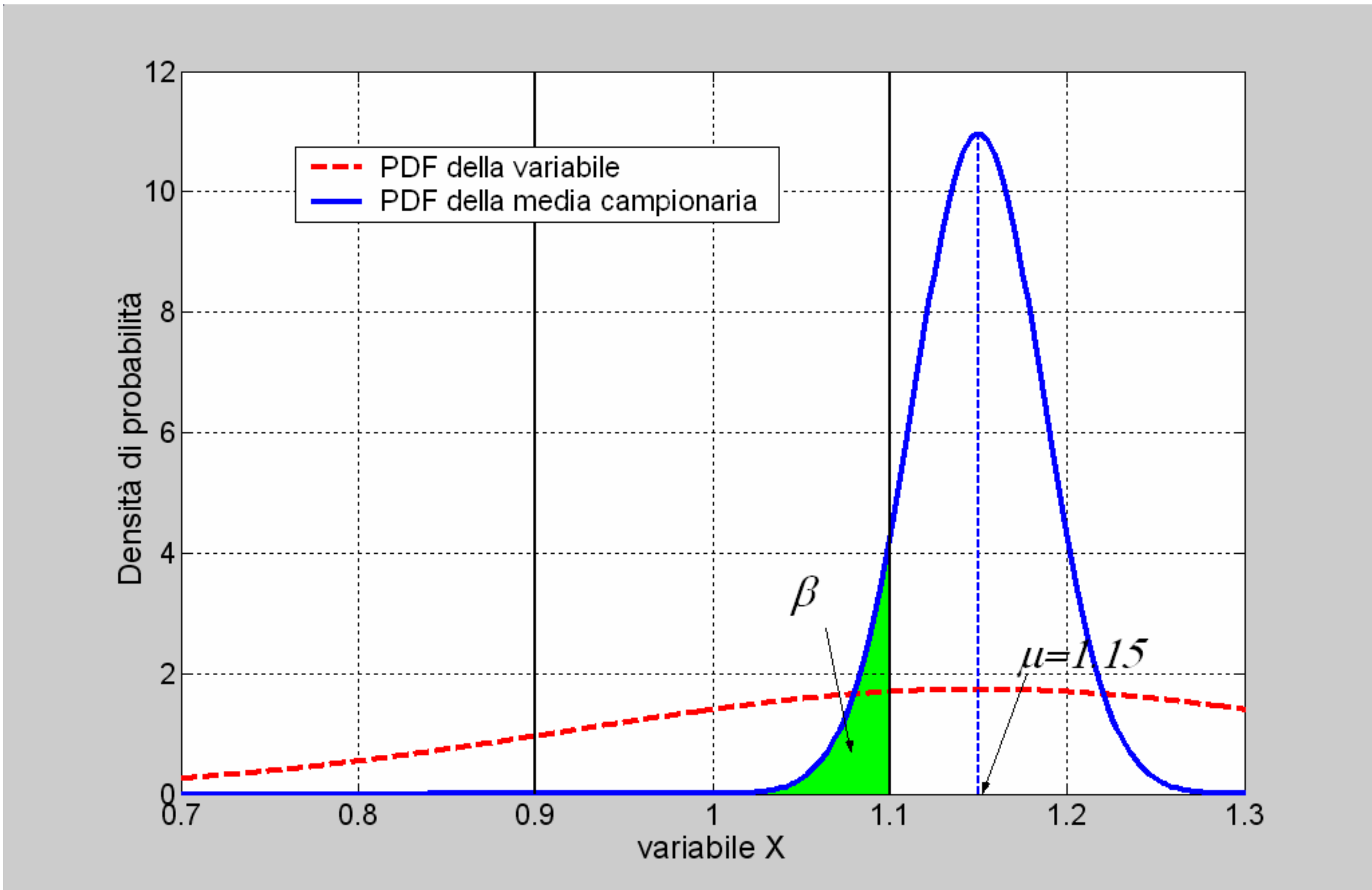
La probabilità *b* di errore di tipo II (di accettare H_0 quando è falsa) è la probabilità che la media campionaria cada all'interno della regione di accettazione, mentre il valor medio della popolazione ha un valore diverso da quello dell'ipotesi nulla. In questo caso $m = 1.15$ € e la deviazione standard del valor medio non è cambiata, $s_{\overline{X}} \cong 0.036$ €, per cui possiamo calcolare *b* per standardizzazione:

$$b = P\left(0.9 < \overline{X} < 1.1 \,\text{con}\, m = 1.15\right) = P\left(z < \frac{1.1 - m}{s_{\overline{X}}}\right) - P\left(z < \frac{0.9 - m}{s_{\overline{X}}}\right) =$$

$$= P\left(z < \frac{1.1 - 1.15}{0.036}\right) - P\left(z < \frac{0.9 - 1.15}{0.036}\right) == P(z < -1.39) - P(z < -6.94) \cong 0.08226 - 0 \cong 8.2\%$$

Riportiamo in figura il significato grafico di *b* per l'esercizio considerato.

La probabilità di errore di tipo II è abbastanza elevata (8%) in quanto le regioni di accettazione sono molto estese, difatti la probabilità di errore di tipo I è molto bassa.
Un modo per diminuire la probabilità di errore sia di tipo I che di tipo II è aumentare il numero dei campioni considerati (con conseguente restringimento della PDF della media campionaria).



PDF della variabile
PDF della media campionaria
Densità di probabilità
variabile X
β
μ=1.15
0
2
4
6
8
10
12
0.7
0.8
0.9
1
1.1
1.2
1.3