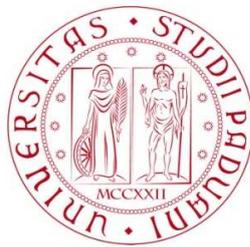


Università degli studi di Padova  
Dipartimento di Psicologia dello Sviluppo e della  
Socializzazione

Corso di Laurea Triennale in  
Scienze psicologiche dello sviluppo, della personalità e delle  
relazioni interpersonali



RELAZIONE FINALE

**SIMULARE NUMERI FUZZY CONTINUI PER DATI  
IMPRECISI: UN MODELLO CONDIZIONALE DI  
CAMPIONAMENTO**

Relatore:  
Prof. Antonio Calcagni

Laureanda: Yasmin El Morady  
Matricola: 1220671

Anno Accademico 2022/2023



# Indice

<b>Introduzione</b>	<b>1</b>
<b>1 Insiemi Fuzzy</b>	<b>3</b>
1.1 Teoria classica degli insiemi . . . . .	4
1.2 Teoria degli insiemi fuzzy . . . . .	6
1.3 Statistica fuzzy . . . . .	7
<b>2 Fuzzificazione</b>	<b>9</b>
2.1 Approcci nella fuzzificazione dei dati psicologici . . . . .	9
2.2 Numeri Fuzzy Beta . . . . .	10
<b>3 Modello condizionale di campionamento</b>	<b>13</b>
3.1 Modello . . . . .	13
3.2 Generazione di numeri fuzzy per l'analisi degli scenari . . . . .	14
3.3 Indici di riferimento . . . . .	15
3.4 Risultati e conclusioni . . . . .	16
<b>A Risultati grafici</b>	<b>18</b>
<b>Bibliografia</b>	

# Introduzione

Grazie alla sua flessibilità e capacità di gestire informazioni codificate secondo diversi gradi di possibilità, la teoria degli insiemi *fuzzy* è una risorsa preziosa per affrontare situazioni reali caratterizzate da incertezza. Nel primo capitolo sono state affrontate le basi teoriche necessarie a comprenderne i concetti e i principi fondamentali. Nel secondo capitolo se ne discutono le applicazioni affrontate nella ricerca statistica, le quali consentono di modellare matematicamente processi come quelli della valutazione soggettiva. Inoltre, viene presentata una famiglia di numeri fuzzy particolarmente flessibile, ossia i numeri *fuzzy beta*. Il terzo capitolo espone un modello teorico capace di modellare numeri fuzzy continui, unimodali e convessi insieme alla presentazione di uno studio che ha valutato un particolare meccanismo per la generazione stocastica di tali numeri.



# Capitolo 1

## Insiemi Fuzzy

La teoria degli insiemi *fuzzy* (in italiano *sfocati*) si è rivelata un'importante estensione della teoria classica degli insiemi. Introdotta da Lotfi Zadeh nel 1965, trova come ambiti di applicazione le discipline che richiedono l'elaborazione di informazioni ambigue, imprecise o soggettive. Il ricercatore propone la teoria degli insiemi fuzzy riscontrando nel pensiero una frequente difficoltà nell'associare gli elementi del mondo fisico e reale a precisi e universali criteri di appartenenza. Per esempio, il numero *10* rispetto alla classe "tutti i numeri reali molto più grandi di 1" presenta questo genere di ambiguità. La stessa questione può essere posta per diverse classi, come "quantità di acqua per formare un lago" o "esito positivo di un trattamento". Nonostante queste classi non vengano intese in senso strettamente matematico, è innegabile che esse costituiscano una parte fondamentale della comunicazione, del ragionamento astratto, del riconoscimento di schemi (Zadeh 1965) e della presa di decisioni. In tal senso, possono rappresentare un punto di incontro tra un'astratta esattezza matematica e una realtà percepita o descritta con vaghezza. Gli insiemi fuzzy sono costituiti da elementi la cui appartenenza è definita secondo un valore compreso tra  $[0, 1]$ . Gli estremi appartengono alla logica bivalente e rappresentano il completamente falso o il completamente vero di una proposizione, mentre, un modello che ammette valori compresi tra questi, permette la formalizzazione di logiche polivalenti capaci di valutare le proposizioni attraverso un grado di verità.

## 1.1 Teoria classica degli insiemi

Secondo la teoria classica degli insiemi, un insieme classico (chiamato anche *crisp*)  $A$  può essere definito come una collezione di elementi  $x$  appartenenti a un insieme universale  $X$ , tali per cui essi condividono una proprietà ben definita. Se un elemento  $x$  presenta tale proprietà, esso appartiene all'insieme  $A$ . Per rappresentare un insieme, è possibile elencare esplicitamente gli elementi appartenenti ad esso:

$$A = \{11, 13, 17, 19\}.$$

Oppure definire le proprietà che  $x$  deve soddisfare:

$$A = \{x \mid \mathcal{A}(x)\} \quad \text{dove}$$

$$\mathcal{A}(x) = 'x \text{ è un numero primo compreso tra } 10 \text{ e } 20'.$$

L'insieme può essere definito attraverso la funzione *indicatrice* o *caratteristica*  $\xi_A(x)$ , come:

$$\xi_A : X \rightarrow \{1, 0\}$$

$$\xi_A(x) = \begin{cases} 1 & \text{se } x \text{ è un numero primo compreso tra } 10 \text{ e } 20 \\ 0 & \text{altrimenti} \end{cases}$$

dove  $\xi_A(x) = 1$  indica l'appartenenza di  $x \in X$  all'insieme  $A$  e  $\xi_A(x) = 0$  indica la sua non appartenenza.

### Operazioni tra gli insiemi

*Inclusione.* Dato l'insieme universale  $X$  e gli insiemi  $A, B$ , si può dire che:

$$A \subseteq B \Leftrightarrow \forall x \in X : x \in A \Rightarrow x \in B. \quad (1.1)$$

$$A \subset B \Leftrightarrow A \subseteq B \wedge A \neq B. \quad (1.2)$$

Riformulando tale definizione secondo le funzioni caratteristiche  $\xi_A(x)$  e  $\xi_B(x)$  degli insiemi  $A$  e  $B$ , è possibile scrivere

$$A \subseteq B \Leftrightarrow \xi_A(x) \leq \xi_B(x) \quad \forall x \in X. \quad (1.3)$$

*Complemento.* L'insieme complementare  $A^c$  di  $A$  è l'insieme che contiene tutti gli elementi di  $X$  non membri di  $A$ .

$$A^c = \{x \mid x \in X \wedge x \notin A\} \quad (1.4)$$

La funzione caratteristica  $\xi_{A^c}(x)$  di  $A^c$  è definita

$$\xi_{A^c}(x) = \begin{cases} 1 & \text{se } \xi_A(x) = 0 \\ 0 & \text{se } \xi_A(x) = 1 \end{cases} \quad (1.5)$$

In cui è possibile distinguere la proprietà:

$$\xi_{A^c}(x) = 1 - \xi_A(x) \quad (1.6)$$

*Intersezione.* L'intersezione di due insiemi  $A, B$  appartenenti a  $X$ , è l'insieme  $A \cap B$  che contiene ogni elemento simultaneamente membro di  $A$  e  $B$ .

$$A \cap B = \{x \in X \mid x \in A \wedge x \in B\} \quad (1.7)$$

La funzione caratteristica e la relativa proprietà risultano

$$\xi_{A \cap B}(x) = \begin{cases} 1 & \text{se } \xi_A(x) = 1 \wedge \xi_B(x) = 1 \\ 0 & \text{altrimenti} \end{cases} \quad (1.8)$$

$$\xi_{A \cap B}(x) = \min\{\xi_A(x), \xi_B(x)\} \quad (1.9)$$

*Unione.* L'unione di  $A, B \subseteq X$  è l'insieme  $A \cup B$  che contiene sia tutti gli elementi di  $A$  che tutti gli elementi di  $B$ .

$$A \cup B = \{x \in X \mid x \in A \vee x \in B\} \quad (1.10)$$

Di conseguenza, la funzione caratteristica e la relativa proprietà sono definite come:

$$\xi_{A \cup B}(x) = \begin{cases} 1 & \text{se } \xi_A(x) = 1 \vee \xi_B(x) = 1 \\ 0 & \text{altrimenti} \end{cases} \quad (1.11)$$

$$\xi_{A \cup B}(x) = \max\{\xi_A(x), \xi_B(x)\} \quad (1.12)$$

## 1.2 Teoria degli insiemi fuzzy

Data una collezione di elementi  $X$ , un insieme *fuzzy*  $\tilde{A}$  in  $X$  è un insieme composto da coppie ordinate di elementi del tipo:

$$\tilde{A} = \{(x, \xi_{\tilde{A}}(x)) \mid x \in X\} \quad (1.13)$$

dove  $\xi_{\tilde{A}}(x)$  è detta *membership function* di  $x$  in  $\tilde{A}$  che associa  $X$  a uno spazio  $M$  compreso tra  $[0, 1]$ . Quando  $M$  contiene solo i valori 0 e 1,  $\tilde{A}$  è un insieme *crisp* e la funzione  $\xi_{\tilde{A}}$  si comporta come una funzione caratteristica. Quando il dominio  $X$  è continuo, l'insieme  $\tilde{A}$  può essere composto da un numero infinito di elementi. Quando il dominio  $X$  è discreto, è possibile esplicitare tutti gli elementi di  $\tilde{A}$ . Per esempio, definendo un insieme di "case grandi", considerando il numero di stanze come condizione di relativa appartenenza, si ha:

$$\tilde{A} = \{(3, 0.2), (4, 0.4), (5, 0.6), (6, 0.8), (7, 1), (8, 1)\}$$

Il primo elemento indica il numero di stanze, il secondo il grado di appartenenza alla categoria "casa grande". Solitamente gli elementi che hanno zero grado di appartenenza vengono omessi. Un insieme  $\tilde{A} =$  "numeri reali molto più grandi di 10" può essere definito come segue

$$\tilde{A} = \{(x, \xi_{\tilde{A}}(x)) \mid x \in X\}$$

dove

$$\xi_{\tilde{A}}(x) = \begin{cases} 0 & \text{se } x \leq 10 \\ 1 + (x - 10)^{-2} & \text{se } x > 10 \end{cases}$$

Il *supporto* di un insieme fuzzy  $\tilde{A}$ ,  $S(\tilde{A})$ , è l'insieme di tutte le  $x \in X$  tali che  $\xi_{\tilde{A}}(x) > 0$ . L'insieme contenente gli elementi dell'insieme fuzzy  $\tilde{A}$  a partire da un certo grado di appartenenza  $\alpha$  è chiamato  $\alpha$ -*level* o  $\alpha$ -*cut*, ovvero  $\tilde{A}_\alpha = \{x \in X : \xi_{\tilde{A}}(x) > \alpha\}$ , con  $\alpha \in (0, 1]$ . Il *nucleo* è l'insieme di tutti i punti massimi della funzione di appartenenza  $A_c = \{x \in X : \xi_{\tilde{A}}(x) = \max_{x \in X} \xi_{\tilde{A}}(x)\}$ . La *cardinalità*  $|\tilde{A}|$ , rispettivamente di un insieme discreto e

di un insieme continuo  $\tilde{A}$ , può essere definita come segue:

$$|\tilde{A}| = \sum_{x \in X} \xi_{\tilde{A}}(x) \quad (1.14)$$

$$|\tilde{A}| = \int_x \xi_{\tilde{A}}(x) dx \quad (1.15)$$

La funzione di appartenenza è la componente fondamentale degli insiemi fuzzy, pertanto le operazioni tra questi avvengono attraverso le rispettive funzioni. Le prime formalizzazioni permettono un'estensione delle operazioni elementari degli insiemi *crisp* derivati dalla logica bivalente o Booleana, come riportate in seguito.

$$\begin{array}{ll} \text{Complemento} & \xi_{\tilde{A}^c}(x) = 1 - \xi_{\tilde{A}}(x) \\ \text{Intersezione} & \xi_{\tilde{A} \cap \tilde{B}}(x) = \min\{\xi_{\tilde{A}}(x), \xi_{\tilde{B}}(x)\} \\ \text{Unione} & \xi_{\tilde{A} \cup \tilde{B}}(x) = \max\{\xi_{\tilde{A}}(x), \xi_{\tilde{B}}(x)\} \end{array}$$

Sebbene queste operazioni trovino diversi ambiti di applicazione, sono possibili ulteriori estensioni. Le operazioni tra questi insiemi possono essere incluse in sistemi costituiti da regole del tipo *IF-THEN* (Mamdani e Sugeno *Fuzzy Inference Systems*), quindi sistemi di inferenze logiche basate su proposizioni fuzzy che possono restituire un valore *crisp* finale.

Sino ad ora i valori della funzione di appartenenza sono stati considerati *crisp*, ma è possibile definire un insieme fuzzy di *tipo-n*, i cui valori della *membership function* sono definiti da  $n - 1$ ,  $n > 1$  insiemi fuzzy in  $[0, 1]$ . Tuttavia, se  $n > 2$ , essi risultano problematici e poco pratici, pertanto solitamente si utilizzano insiemi *fuzzy di tipo-2* (Zimmermann 2011, Hanss 2005).

## 1.3 Statistica fuzzy

L'incertezza legata al significato degli eventi differisce da quella riferita all'esito degli stessi, dunque un approccio fuzzy alla statistica e all'analisi dei dati permette di congiungere questi due aspetti. Si sono sviluppate principalmente due interpretazioni in ambito statistico, da una parte vi è un'interpretazione *ontica*, dall'altra una *epistemica*. Nel primo caso, il dato

raccolto viene inteso come rappresentazione di un oggetto esistente, ma essenzialmente impreciso. Generalmente, il termine "*fuzzy random set*" si riferisce a questo approccio. Nel secondo caso, i dati rappresentano una realizzazione imperfetta, non misurabile con esattezza, di un oggetto esistente e preciso. In questo senso, si utilizza il termine "*fuzzy random variable*". Entrambe le interpretazioni sono state applicate a diverse analisi statistiche (Couso et al. 2019, Renato Coppi 2006). Nel caso dell'inferenza, vi è la possibilità di costruire test di verifica d'ipotesi basati su questo tipo di dati, e, dal punto di vista epistemico, l'inferenza riguarda un parametro della distribuzione della variabile casuale sottostante. L'imprecisione nei dati viene così inclusa nella computazione del test statistico. Ulteriori metodi statistici per la stima dei parametri possono essere generalizzati ai dati sfocati. Ad esempio, è stata formalizzata l'estensione dell'algoritmo *Expectation-Maximization (EM)* basato sul metodo della massima verosimiglianza (Dencœux 2011).

# Capitolo 2

## Fuzzificazione

Al momento della raccolta di informazioni, spesso si ha una collezione di dati *crisp*, tuttavia, se è ragionevole ipotizzare che siano caratterizzati da qualche forma di ambiguità, questi si possono "fuzzificare": in altri termini si presume che ogni osservazione sia pervasa da un'incertezza descritta attraverso l'approccio fuzzy. La fuzzificazione è un passaggio cruciale nella teoria fuzzy in quanto consente di gestire l'incertezza e la sfumatura dei dati e di incorporarli nella modellazione e nelle analisi successive. Elaborare un modello generale che possa descrivere il processo di fuzzificazione per diverse tipologie di dati risulta una questione importante nel campo della statistica applicata e della psicometria.

### 2.1 Approcci nella fuzzificazione dei dati psicologici

Nella misurazione in psicologia, le scale di valutazione occupano un ruolo fondamentale nell'analisi quantitativa di fenomeni psicologici e nelle rilevazioni di tratti di personalità, attitudini, motivazioni e opinioni. Esse sono strutturate come un insieme di item o indicatori a cui il soggetto risponde attraverso una valutazione numerica sul livello di accordo/disaccordo con la proposizione presentata, come nelle scale Likert. Sebbene queste rappresentino uno strumento pratico e agevole da somministrare e analizzare, sono soggette a diversi limiti quali distorsioni di tipo cognitivo, legate alla desiderabilità sociale, alle differenze individuali e alle differenze culturali. Inoltre,

non sono indipendenti da fattori contestuali. L'informazione raccolta risulta così compromessa, per cui estrarre la componente di verità ad essa associata diviene un compito cruciale. La teoria degli insiemi fuzzy come possibile soluzione è stata più volte utilizzata nell'ambito della ricerca psicometrica e statistica, non senza qualche criticità. E' possibile distinguere diversi metodi di fuzzificazione dei dati provenienti da queste misurazioni. Tra i principali vi sono le *fuzzy conversion scales* e le *fuzzy rating scales*. Il primo metodo consiste nel convertire ogni valore numerico proveniente da una scala categoriale ordinale in numero fuzzy, in genere utilizzando funzioni di appartenenza come quelle triangolari o trapezoidali (Vonglao 2017). Tali funzioni possono essere scelte a discrezione del ricercatore oppure sulla base di evidenze empiriche. Tuttavia, assumere una lista prefissata di potenziali numeri fuzzy per la descrizione della valutazione raccolta risulta controverso (Sáa et al. 2013). Il secondo metodo adotta, invece, procedure esplicite per raccogliere informazioni e definire la funzione di appartenenza associata alla singola valutazione da parte del soggetto (Lubiano et al. 2016). Ad esempio, in uno studio di Sáa et al. 2014, i partecipanti possono rispondere a ogni item indicando un intervallo su una scala grafica pseudo-continua. Successivamente, viene richiesto di definire il grado di incertezza relativo alla risposta precedente attraverso un secondo insieme. Infine, gli estremi degli intervalli vengono interpolati per descrivere la risposta finale attraverso un numero fuzzy. Anche in questo caso, è opinabile che il soggetto possa definire obiettivamente il proprio grado di incertezza e che questo venga utilizzato come insieme fuzzy in modo assoluto. Inoltre, sono state studiate procedure empiriche indirette, in cui l'insieme fuzzy è costruito a partire da misure implicite associate alla risposta crisp finale, come il lasso di tempo che intercorre tra la presentazione dell'item e la risposta del soggetto (Cao e Calcagnì 2022) oppure attraverso l'analisi delle traiettorie compiute da un computer durante il processo di scelta della risposta all'item (Calcagnì e Lombardi 2014).

## 2.2 Numeri Fuzzy Beta

I valori sfocati possono essere descritti con funzioni di appartenenza parametrizzate, che ne determinano la forma, la posizione e la dispersione. I

cosiddetti *beta fuzzy numbers* sono definiti da una funzione i cui parametri possono essere associati al processo di realizzazione di ogni dato sfocato. Essi nascono dall'estensione della distribuzione Beta di probabilità. La parametrizzazione di tale funzione permette di approssimare diverse tipologie di funzioni di appartenenza, come quelle trapezoidali, triangolari e gaussiane. Essa è definita:

$$\begin{aligned}\xi_{\tilde{A}}(x) &= \left(\frac{x - x_l}{m - x_l}\right)^a \left(\frac{x_u - x}{x_u - m}\right)^b \cdot \mathbb{1}_{x_l, x_u}(x) \\ m &= \frac{ax_u + bx_l}{a + b}\end{aligned}\quad (2.1)$$

dove  $x_l, x_u, a, b \in \mathbb{R}$ , e  $x_l, x_u$  sono gli estremi dell'insieme e  $m$  si riferisce alla moda dell'insieme fuzzy. I numeri fuzzy beta possono essere definiti secondo i parametri moda  $m \in \mathbb{R}$  e precisione  $s \in \mathbb{R}^+$ . Definendo l'intervallo  $x_l = 0$  e  $x_u = 1$ , l'equazione (2.1) può essere semplificata come segue:

$$\begin{aligned}\xi_{\tilde{A}}(x) &= \frac{1}{C} x^{a-1} (1-x)^{b-1} \\ a &= 1 + ms \\ b &= 1 + s(1 - m)\end{aligned}\quad (2.2)$$

dove  $C$  è una costante che assicura che l'insieme fuzzy soddisfi la condizione  $\max_{x \in A} \xi_{\tilde{A}}(x) = 1$ :

$$C = \left(\frac{a-1}{a+b-2}\right)^{a-1} \cdot \left(1 - \frac{a-1}{a+b-2}\right)^{b-1}\quad (2.3)$$

Le realizzazioni di un processo casuale, definito quindi  $Y \sim f_Y(y; \theta)$ , dove  $f_Y$  è la funzione di densità di probabilità caratterizzata da uno o più parametri  $\theta$ , sono legati al modello proposto in questo modo:  $m$  rappresenta il valore più plausibile,  $s$  è la precisione della realizzazione  $Y$ .

A partire dalle realizzazioni (vere e non osservabili) di  $Y \sim f_Y(y; \theta)$  e dai due vettori  $I \times 1$  moda ( $m$ ) e precisione ( $s$ ) la fuzzificazione del dato osservato  $\tilde{y}$  può seguire due modelli principali in termini di relazione tra quest'ultimo e i parametri proposti:

- **Fuzzificazione simmetrica:** il valore vero corrisponde al valore più plausibile  $m$ .

$$\begin{aligned}\tilde{\mathbf{y}} &= (\mathbf{m}, \mathbf{s}) = ((m_1, s_1), \dots, (m_i, s_i), \dots, (m_I, s_I)) \\ y &= m \\ s &\in \mathbb{R}^+ \sim F_s(s; \theta_s)\end{aligned}\tag{2.4}$$

Dove  $\tilde{y}$  indica il valore osservato e  $y$  il valore vero. È importante notare che, idealmente, se la precisione fosse molto alta, il grado di fuzziness tenderebbe a zero e il valore vero verrebbe osservato precisamente, ovvero  $Y = y = m$ .

- **Fuzzificazione asimmetrica:** la moda non corrisponde necessariamente al valore vero  $y$ .

$$\begin{aligned}\tilde{\mathbf{y}} &= (\mathbf{m}, \mathbf{s}) = ((m_1, s_1), \dots, (m_i, s_i), \dots, (m_I, s_I)) \\ m &\sim F_m(m; \theta_m) \\ s &\in \mathbb{R}^+ \sim F_s(s; \theta_s)\end{aligned}\tag{2.5}$$

In questo contesto, è possibile assumere che  $m \perp\!\!\!\perp s$ ,  $m \perp\!\!\!\perp y$  oppure che vi sia una relazione tra la moda  $m$  e i valori veri  $y$ .

# Capitolo 3

## Modello condizionale di campionamento

### 3.1 Modello

Sia  $\tilde{\mathbf{y}} = \{\tilde{y}_1, \dots, \tilde{y}_i, \dots, \tilde{y}_I\}$  un campione di  $I$  numeri fuzzy e  $\{Y_1, \dots, Y_i, \dots, Y_I\}$  una collezione di variabili casuali indipendenti, con distribuzione di probabilità  $f_Y(y; \theta) = \prod_{i=1}^I f_{Y_i}(y; \theta)$ , con  $\theta \in \Theta \subset \mathbb{R}^p$ . Facendo riferimento all'approccio epistemico dei numeri fuzzy, il meccanismo di generazione dei dati sfocati può essere formalizzato utilizzando uno schema condizionale di campionamento, in grado di legare i parametri dei numeri fuzzy beta come moda, precisione e informazioni riguardanti la funzione di appartenenza, all'esito casuale di  $f_Y(y; \theta)$ . Il modello è definito come segue:

$$Y_i \sim f_Y(y; \theta) \quad (3.1)$$

$$s_i \sim F_{s_i}(s; \theta_s) \quad (3.2)$$

$$m|s_i, y_i \sim \text{Beta}(m; y_i s_i, s_i - s_i y_i)_{[lb; ub]} \quad (3.3)$$

dove le eq. (3.2) e (3.3) descrivono il meccanismo di fuzzificazione e  $Y_i \sim f_Y(y; \theta)$  è il modello statistico parametrico che si vuole studiare. La moda, dunque, segue la distribuzione Beta condizionata agli esiti della precisione  $s_i$  e del valore non osservato  $y_i$  ed è definita nell'intervallo  $[lb, ub]$ , che identifica

il supporto empirico del valore vero non osservato. Si noti che la componente di precisione  $s_i$  può essere descritta utilizzando appropriate distribuzioni di probabilità (ad esempio distribuzione Lognormale, Gamma). Inoltre, l'assunzione  $Y_i \perp\!\!\!\perp s_i$  suggerisce che l'esito della precisione è indipendente dalla realizzazione vera. Lo studio del modello  $Y_i \sim f_Y(y; \theta)$  avviene mantenendo le eq. (3.2) e (3.3) fisse, mentre il primo può variare a seconda delle ipotesi di riferimento. Il presente capitolo ha lo scopo di illustrare le proprietà e le caratteristiche relative al modello proposto.

## 3.2 Generazione di numeri fuzzy per l'analisi degli scenari

La simulazione condotta per lo studio è stata implementata utilizzando il linguaggio di programmazione R. È necessario definire alcuni elementi chiave da cui i dati fuzzy hanno origine. La simulazione prende in considerazione tre elementi principali:

- L'origine dei dati di partenza: le distribuzioni da cui vengono generati i dati veri sono cinque, ovvero la distribuzione Normale, Beta, Esponenziale, Gamma e Weibull. I parametri associati a ciascuna distribuzione sono definiti in tre modi differenti, come riportato nella tabella 3.1.
- La componente di precisione: la precisione agisce casualmente e segue una distribuzione Gamma. Si possono distinguere tre casi, ovvero il caso di bassa precisione, di media precisione e alta precisione definiti dai parametri nella tabella 3.2.
- L'intervallo dei dati  $Y_i \sim f_{Y_i}(y; \theta)$  è stato considerato secondo due approcci: il primo consiste nell'utilizzo di un intervallo  $[lb, ub]$  opportunamente predefinito, il secondo si basa sui valori estremi ottenuti dal campione stesso.

Le possibili combinazioni del disegno di simulazione risultano essere  $15 \times 3 \times 2 = 90$ . Per ogni combinazione è stato generato un campione di numerosità  $n = 5000$ , producendo  $5000 \times 90 = 450000$  dati. Ognuno di essi è stato

**Tabella 3.1:** Distribuzioni utilizzate per  $F_{Y_i} \sim (y; \theta)$  e relativi parametri

Distribuzione	Parametri	<i>I</i>	<i>II</i>	<i>III</i>
<i>Normale</i>	$\mu$	0	-2.5	18.6
	$\sigma$	2	0.25	10
<i>Beta</i>	$\alpha$	0.5	0.2	0.8
	$\beta$	2.5	10.5	20
<i>Esponenziale</i>	$\lambda$	1/1.15	1/8.15	1/20.15
<i>Gamma</i>	$a$	1.85	6.04	51.33
	$b$	0.34	0.25	0.36
<i>Weibull</i>	$c$	2.15	5.15	1.15
	$d$	3.5	10.5	7.5

**Tabella 3.2:** Parametri per la precisione

Distribuzione	Parametri	Precisione		
		Bassa	Media	Alta
<i>Gamma</i>	$k$	12.48	72.83	128.55
	$t$	1.76	1.42	0.944

normalizzato attraverso il metodo *min-max scaling* (3.4), ottenendo un valore compreso nell'intervallo  $[0, 1]$  e non interferendo sulla forma della funzione e sulle caratteristiche della curva:

$$y'_i = \frac{y_i - \min(y)}{\max(y) - \min(y)}. \quad (3.4)$$

I valori della moda sono stati ottenuti secondo la funzione  $m_i^* \sim \text{Beta}(m_i; y'_i s_i, s_i - s_i y'_i)$  nell'intervallo  $[0, 1]$  a partire dai valori generati in precedenza.

### 3.3 Indici di riferimento

Per la valutazione dello studio sono stati utilizzati indici relativi agli insiemi fuzzy (Ross 2009, Calcagnì e Lombardi 2014) che riassumono in modo puntuale la funzione di appartenenza. Questi indici forniscono informazioni

significative sulla forma e sulle caratteristiche degli insiemi fuzzy considerati. Nel dettaglio, sono stati utilizzati i seguenti indici:

- *Centroide* (Ross 2009): rappresenta un valore di tendenza centrale dell'insieme fuzzy e permette di sintetizzare la distribuzione di appartenenza in modo puntuale. Esso è calcolato come il rapporto tra l'integrale del prodotto della funzione di appartenenza per il valore di variabile e l'integrale della funzione di appartenenza stessa.

$$CR(\tilde{\xi}_i) = \frac{\int \mu_{\tilde{\xi}(x)} x dx}{\int \mu_{\tilde{\xi}(x)} dx} \quad (3.5)$$

- *Cardinalità*: rappresenta la misura della quantità di informazione contenuta nell'insieme fuzzy. È ottenuta tramite l'integrazione della funzione di appartenenza sull'intervallo di definizione dell'insieme (1.15).
- *Indice di Kaufmann* (Kaufmann e Bonaert 1977): basato sulla cardinalità dell'insieme, rappresenta una misura della dispersione.

$$K(\tilde{\xi}) = \frac{2}{\int_x \xi_{\tilde{A}}(x) dx} \cdot \sum_x |\mu_{\tilde{\xi}}(x) - \delta(x)| \quad \text{dove} \quad (3.6)$$

$$\delta(x) = \begin{cases} 1 & \text{se } \mu_{\tilde{\xi}}(x) \geq 0.5 \\ 0 & \text{se } \mu_{\tilde{\xi}}(x) < 0.5 \end{cases}$$

Inoltre, è stata anche tenuta in conto la differenza tra la moda  $m_i^*$  e il valore vero  $Y_i \sim f_Y(y; \theta)_{[0,1]}$  generato da ogni distribuzione. Questa quantità permette di valutare la discrepanza tra il valore più frequente nell'insieme fuzzy e il valore reale generato dalla distribuzione di probabilità, con l'obiettivo di valutare quanto la moda si discosta dal valore vero generato da ogni distribuzione.

### 3.4 Risultati e conclusioni

I risultati della simulazione hanno fornito una panoramica approfondita sulle proprietà dei numeri fuzzy generati in molteplici scenari, come mostrato

dai grafici in Appendice A. In primo luogo, come era ragionevole aspettarsi, i diversi livelli di precisione hanno avuto un impatto notevole sulle proprietà dei numeri fuzzy beta. Una bassa precisione ha portato a una maggiore incertezza e imprecisione nei numeri fuzzy, mentre una precisione più alta ha ridotto la variabilità e ha prodotto numeri fuzzy più concentrati attorno alle loro mode. Questo è stato osservato in tutti i casi, ed è ben sintetizzato dagli indici di cardinalità e Kaufmann. Si evidenzia che questi indici mostrano una consistenza e una tendenza comune tra le diverse distribuzioni di partenza, ad eccezione di quella esponenziale, i cui indici assumono valori più bassi. Ciò suggerirebbe che il modello proposto è in grado di generare numeri fuzzy con caratteristiche simili e coerenti, indipendentemente dalla distribuzione di partenza  $Y_i \sim f_Y(y; \theta)$  del valore non noto. La media delle differenze tra il valore vero  $y$  e la moda  $m_i^*$  è sostanzialmente nulla in tutti i casi ma la dispersione cambia in relazione alla precisione: campioni con precisione più alta mostrano, in generale, una minore dispersione della differenza  $m_i^* - y$ .

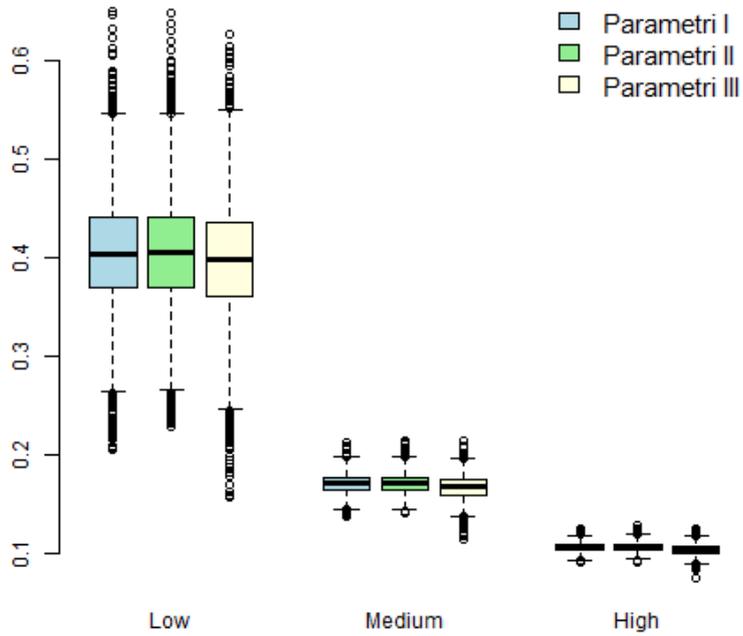
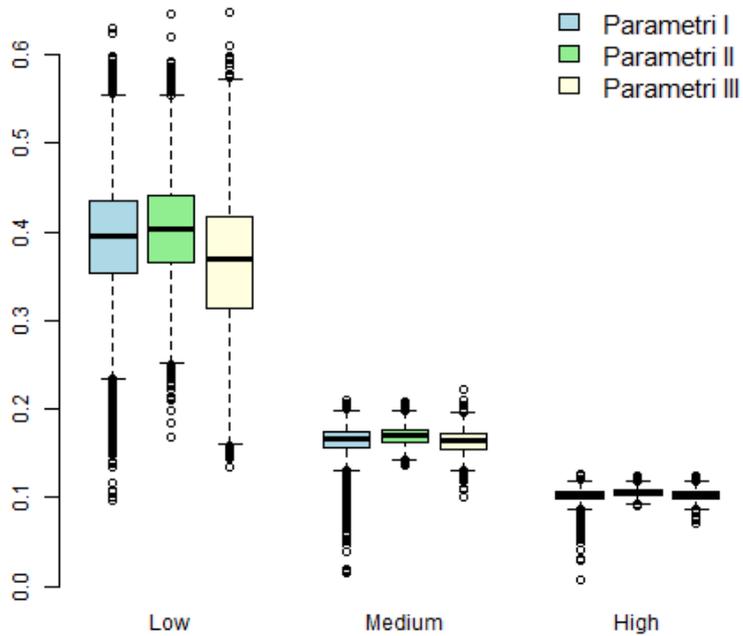
Appendice A

Risultati grafici

## CAMPIONI GENERATI DALLA DISTRIBUZIONE NORMALE

*Cardinalità*

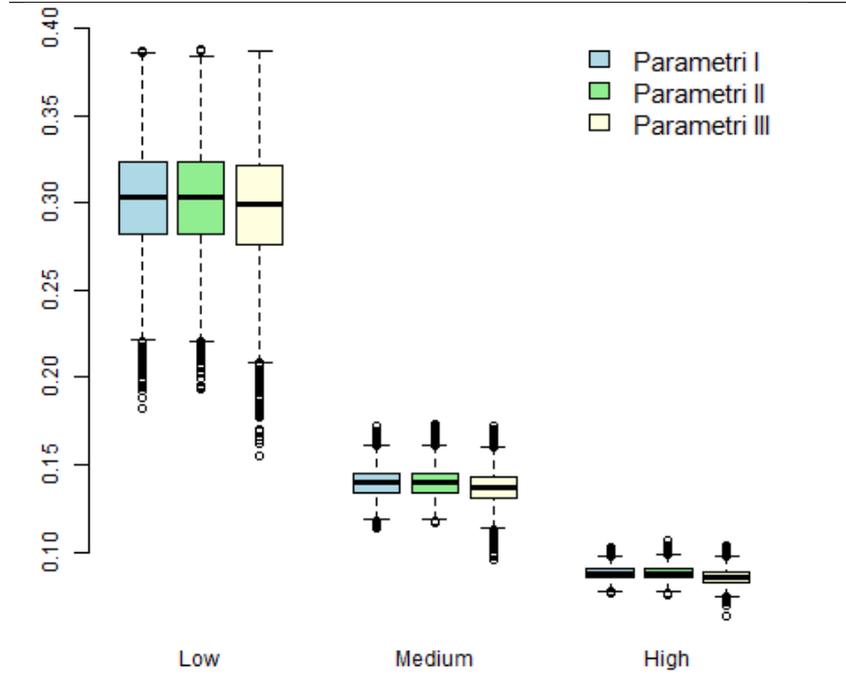
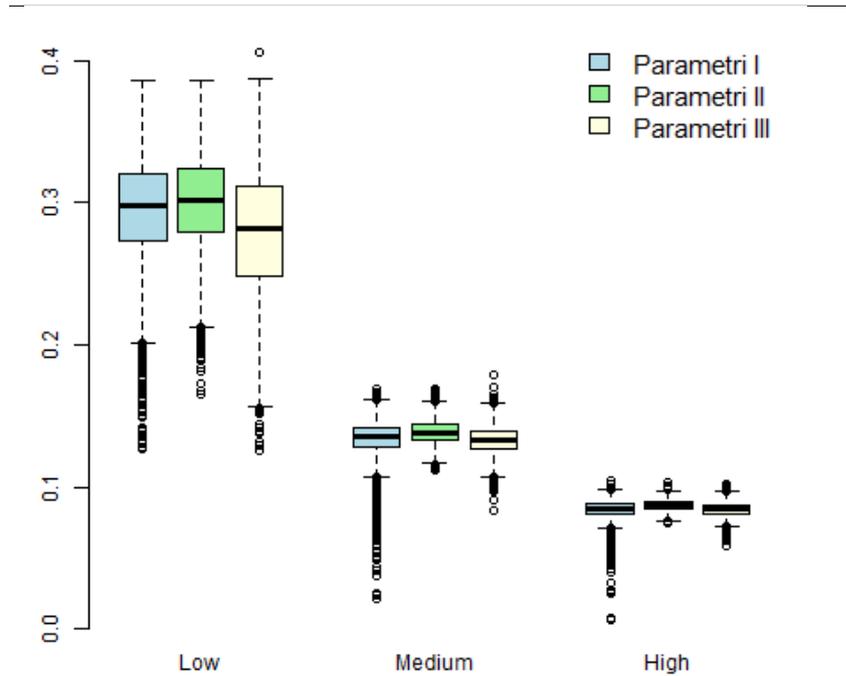
## Intervallo fisso

*Intervallo sample based*

## CAMPIONI GENERATI DALLA DISTRIBUZIONE NORMALE

*Indice di Kaufmann*

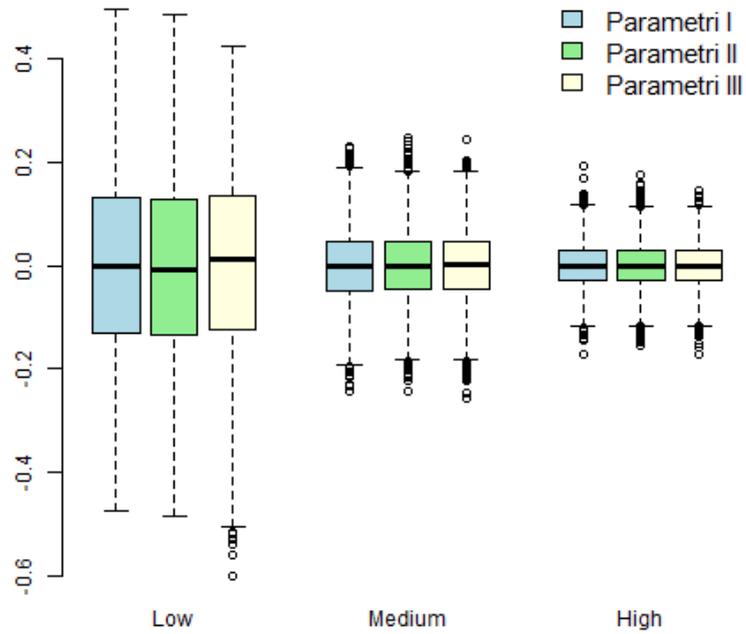
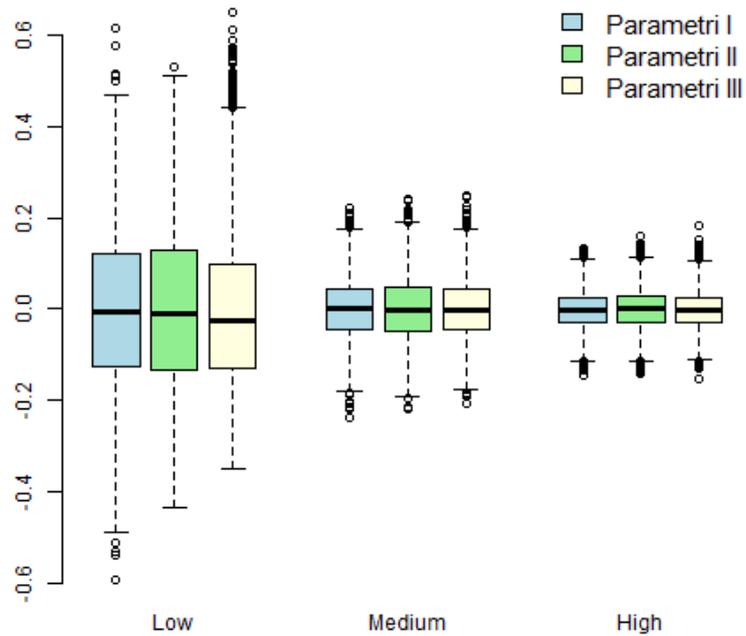
## Intervallo fisso

Intervallo *sample based*

## CAMPIONI GENERATI DALLA DISTRIBUZIONE NORMALE

 $Differenza (m_i^* - y_i)$ 

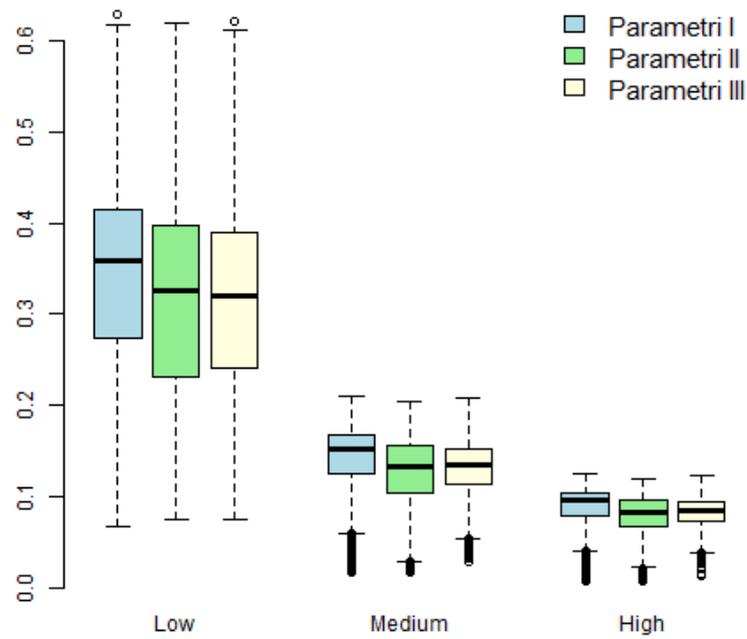
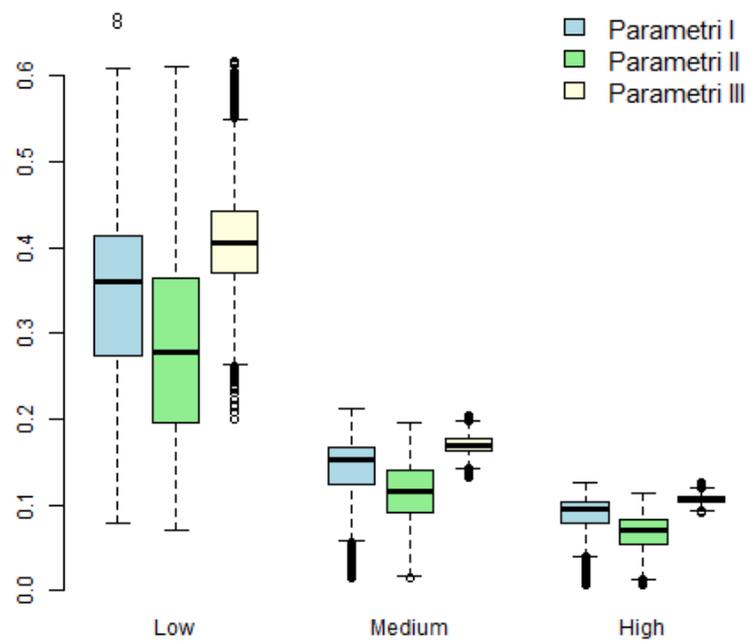
Intervallo fisso

Intervallo *sample based*

## CAMPIONI GENERATI DALLA DISTRIBUZIONE BETA

*Cardinalità*

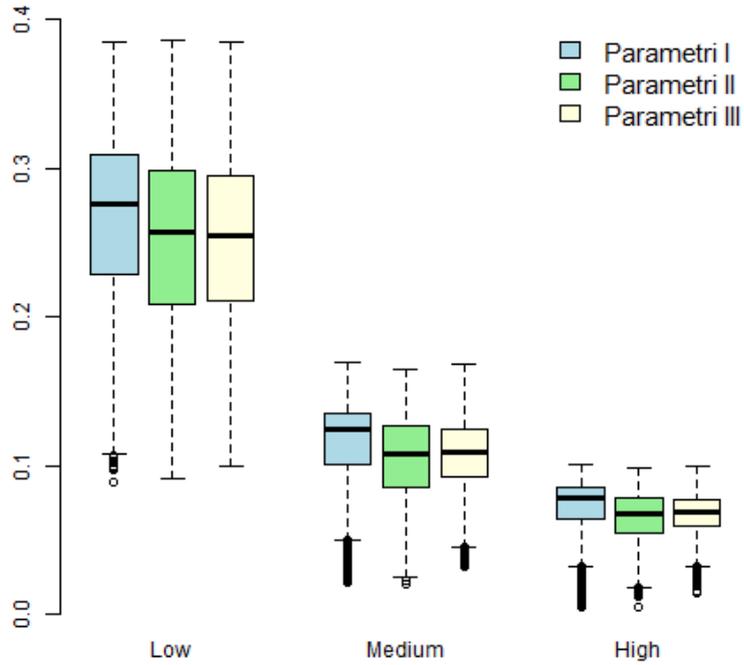
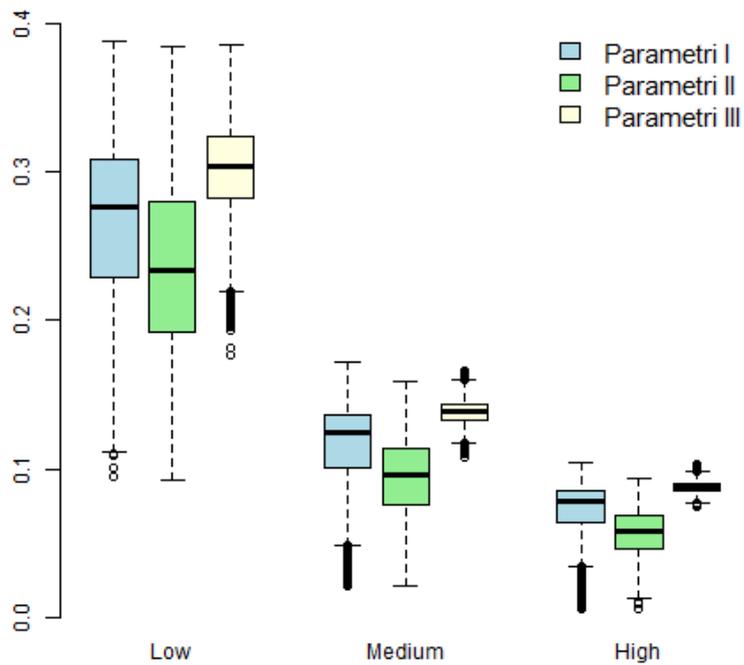
## Intervallo fisso

Intervallo *sample based*

## CAMPIONI GENERATI DALLA DISTRIBUZIONE BETA

*Indice di Kaufmann*

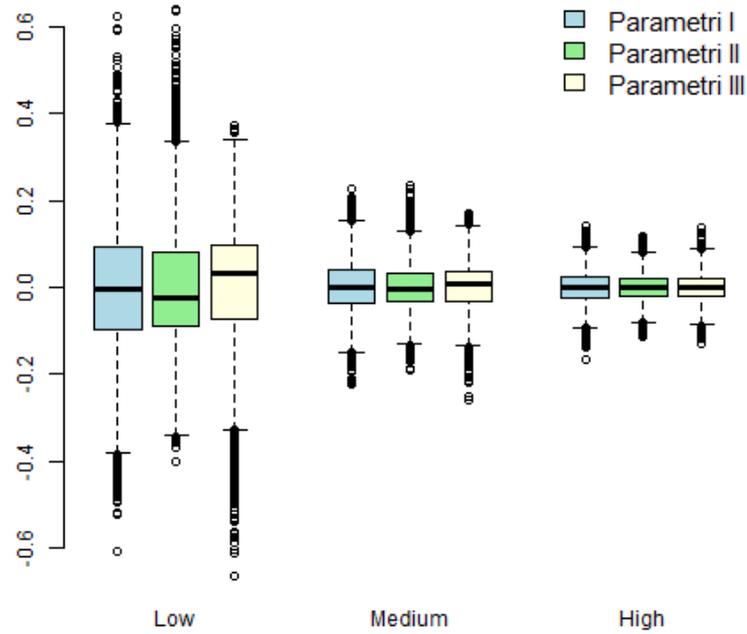
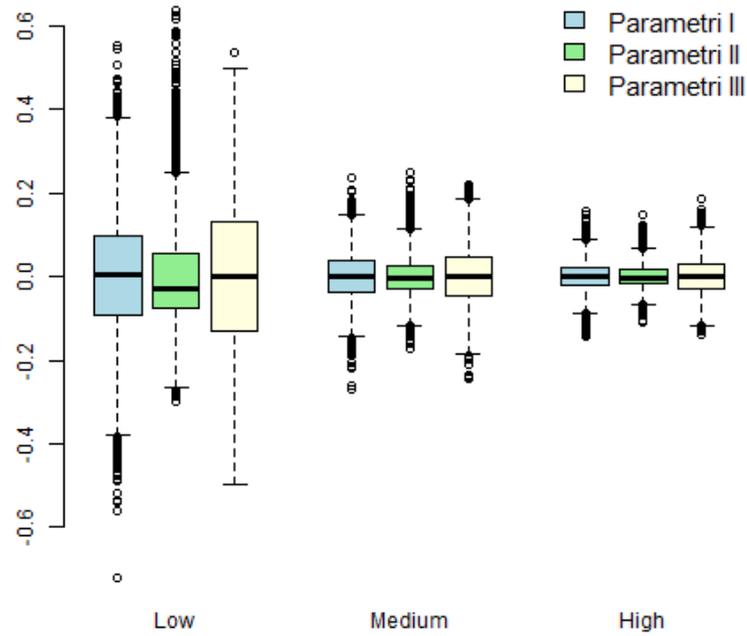
## Intervallo fisso

*Intervallo sample based*

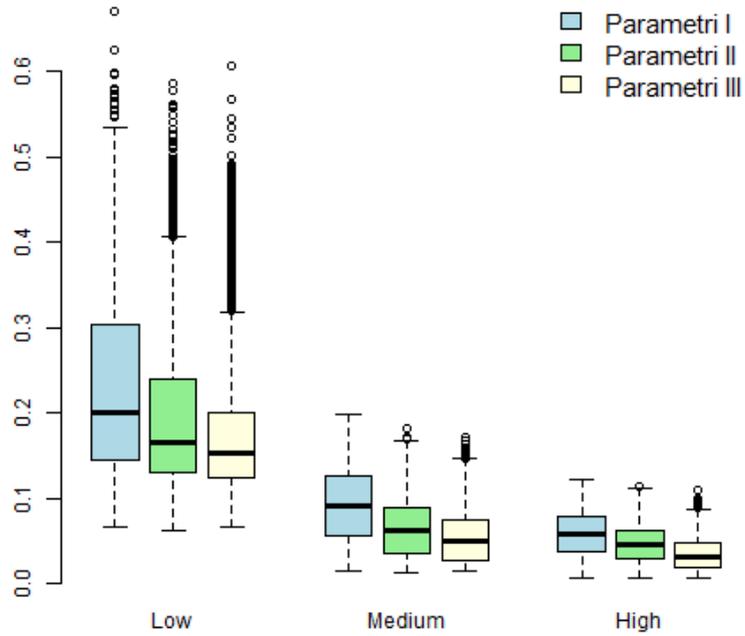
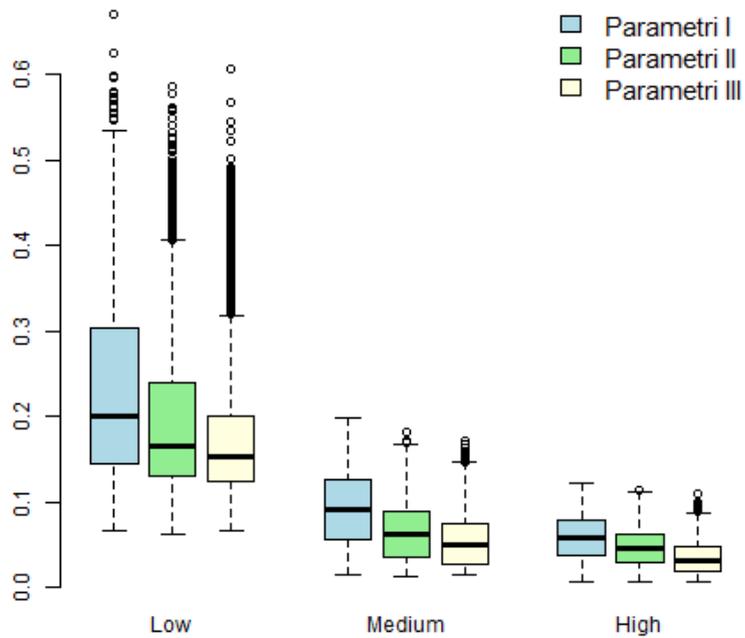
## CAMPIONI GENERATI DALLA DISTRIBUZIONE BETA

 $Differenza (m_i^* - y_i)$ 

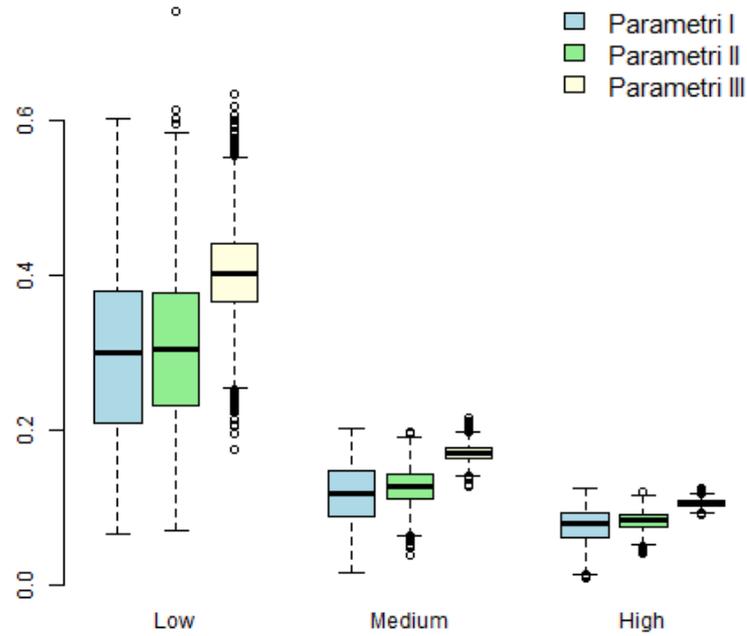
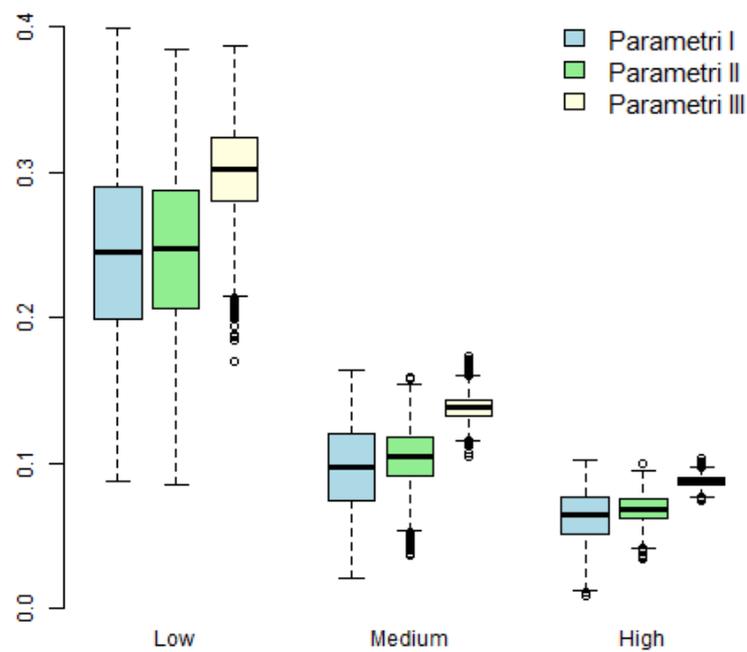
Intervallo fisso

Intervallo *sample based*

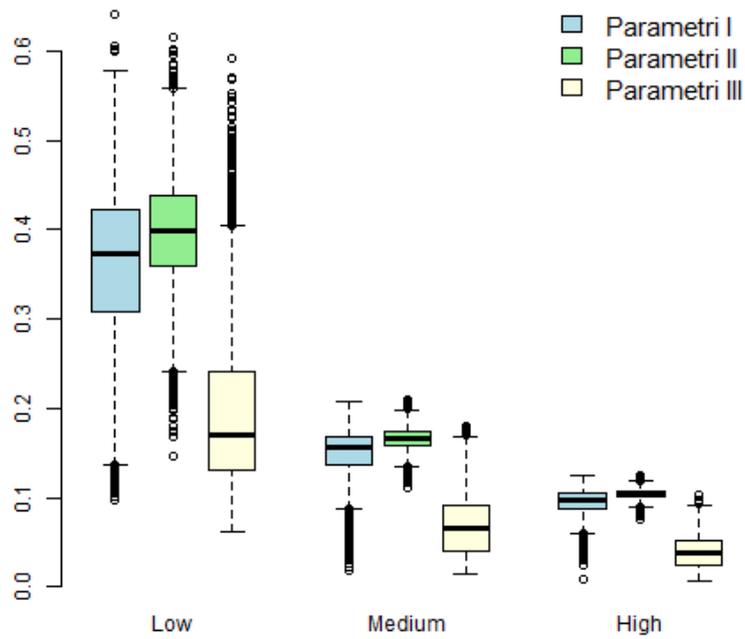
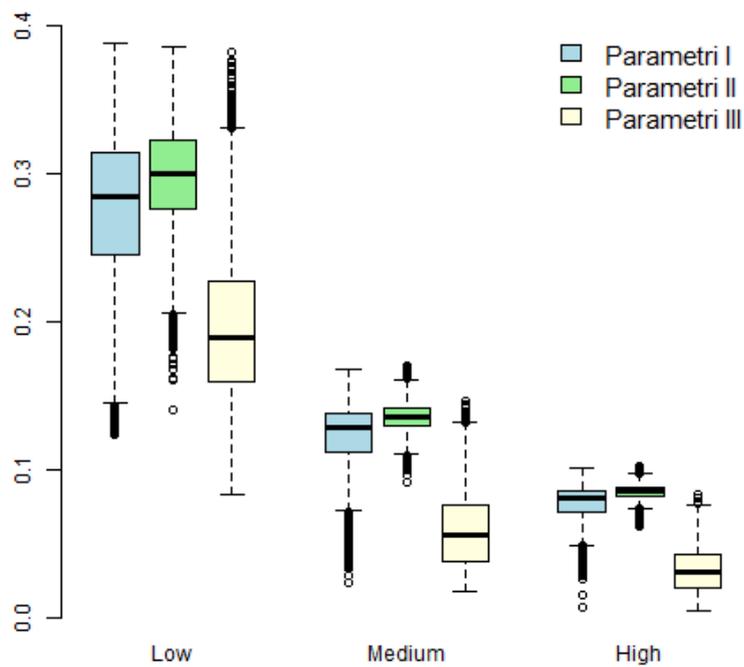
## CAMPIONI GENERATI DALLA DISTRIBUZIONE ESPONENZIALE

Intervallo *sample based**Cardinalità**Indice di Kaufmann*

## CAMPIONI GENERATI DALLA DISTRIBUZIONE GAMMA

Intervallo *sample based**Cardinalità**Indice di Kaufmann*

## CAMPIONI GENERATI DALLA DISTRIBUZIONE DI WEIBULL

Intervallo *sample based**Cardinalità**Indice di Kaufmann*

# Bibliografia

- Calcagnì, Antonio e L Lombardi (2014). «Dynamic Fuzzy Rating Tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales». In: *Applied soft computing* 24, pp. 948–961.
- Cao, Niccolò e Antonio Calcagnì (2022). «Jointly modeling rating responses and times with fuzzy numbers: An application to psychometric data». In: *Mathematics* 10.7, p. 1025.
- Couso, Ines et al. (2019). «Fuzzy sets in data analysis: From statistical foundations to machine learning». In: *IEEE Computational Intelligence Magazine* 14.1, pp. 31–44.
- Denceux, T. (2011). «Maximum likelihood estimation from fuzzy data using the EM algorithm. Fuzzy Sets Syst.» In.
- Hanss, Michael (2005). *Applied fuzzy arithmetic*. Springer.
- Kaufmann, Arnold e AP Bonaert (1977). «Introduction to the theory of fuzzy subsets-vol. 1: Fundamental theoretical elements». In: *IEEE transactions on systems, man, and cybernetics* 7.6, pp. 495–496.
- Lubiano, María Asunción et al. (set. 2016). «Descriptive analysis of responses to items in questionnaires. Why not using a fuzzy rating scale?» In: *Information Sciences* 160, 131–148. DOI: [10.1016/j.ins.2016.04.029](https://doi.org/10.1016/j.ins.2016.04.029).
- Renato Coppi Maria A. Gilb, Henk A.L. Kiersc (2006). «The fuzzy approach to statistical analysis». In.
- Ross, Timothy J (2009). *Fuzzy logic with engineering applications*. John Wiley & Sons.
- Sáa, Sara de la Rosa de et al. (2013). «Fuzzy rating vs. fuzzy conversion scales: an empirical comparison through the MSE». In: *Synergies of Soft*

- 
- Computing and Statistics for Intelligent Data Analysis*. Springer, pp. 135–143.
- Sáa, Sara de la Rosa de et al. (2014). «Fuzzy rating scale-based questionnaires and their statistical analysis». In: *IEEE Transactions on Fuzzy Systems* 23.1, pp. 111–126.
- Vonglao, Paotjai (2017). «Application of fuzzy logic to improve the Likert scale to measure latent variables». In: *Kasetsart Journal of Social Sciences* 38.3, pp. 337–344.
- Zadeh, L. A. (1965). «Fuzzy Sets». In.
- Zimmermann, Hans-Jürgen (2011). *Fuzzy set theory—and its applications*. Springer Science & Business Media.