

RENATO IOVINE

METODOLOGIE STATISTICHE PER LA STIMA IMMOBILIARE

l'impiego nell'ambito della riforma del catasto
(legge n. 23/2014)
esempi applicativi

EXEO edizioni

GUIDE
OPERATIVE

ISBN formato pdf: 978-88-6907-150-8

professionisti
pubblica amministrazione

ESTIMO CATASTALE collana a cura di MASSIMO CURATOLO

MC09

RENATO IOVINE

Metodologie statistiche per la stima immobiliare

l'impiego nell'ambito della riforma del catasto
(legge n. 23/2014)
esempi applicativi

EXEO edizioni 

GUIDE OPERATIVE

ISBN formato pdf: 978-88-6907-150-8

professionisti

pubblica amministrazione

Il testo costituisce una sintetica, ma esaustiva guida operativa per l'approccio alle metodologie statistiche nel campo dell'estimo immobiliare. In particolare, in vista dell'imminente adozione nell'ambito dell'estimo catastale di metodologie statistico-matematiche volute dalla legge delega n. 23/2014, è viva una richiesta informativa nel campo delle procedure statistiche. L'ebook vuole estrapolare dall'estesissimo universo delle conoscenze statiche e matematiche, quelle nozioni essenziali ed utili, oltre che per comprendere il fenomeno, sufficienti anche per condurre alcune sperimentazioni ed applicazioni pratiche in proprio nell'ambito dell'estimo immobiliare e quello che sarà l'imminente estimo anche catastale. Le sperimentazioni sono illustrate con l'uso di un software nella disponibilità di quasi tutti i professionisti (foglio excel di Microsoft Office) che consente l'allestimento di un foglio di calcolo personalizzato, prodotto minimo ma già ampiamente sufficiente per lo scopo perseguito.

Copyright © 2015 Exeo S.r.l.. Tutti i diritti riservati. È consentita la stampa e l'utilizzo in più dispositivi ad esclusivo uso personale della persona fisica acquirente, o del destinatario del prodotto in caso di soggetto acquirente diverso da persona fisica, e comunque mai ad uso commerciale: ogni diversa utilizzazione e diffusione, con qualsiasi mezzo, con qualsiasi scopo e nei confronti di chiunque altro, è vietata senza il consenso scritto dell'editore. Quanto alla riproduzione dei contenuti, sono consentite esclusivamente citazioni di brevi brani in virgolettato a titolo di cronaca, studio, recensione, attività della pubblica amministrazione o professionale, accompagnate dal nome dell'autore, dell'editore, e dal titolo e anno della pubblicazione. Sarà perseguita nelle sedi opportune ogni violazione dei diritti d'autore e di editore. Alle violazioni si applicano le sanzioni previste dagli art. 171, 171-bis, 171-ter, 174-bis e 174-ter della legge 633/1941.

Renato Iovine: ingegnere, dottore di ricerca, membro della Commissione Catasto dell'Ordine degli ingegneri di Roma. Libero professionista. Fra, l'altro, collabora con studio di ingegneria specializzato nell'estimo immobiliare e catastale.

edizione: novembre 2015 | collana: ESTIMO CATASTALE, a cura di Massimo Curatolo | materia: estimo | tipologia: guida operativa | formato: digitale, pdf | codice prodotto: MCO9 - nic 9 | ISBN: 978-88-6907-150-8 | Editore: Exeo srl CF PI RI 03790770287 REA 337549 ROC 15200/2007 c.s.i.v. € 10.000,00, sede legale piazzetta Modin 12 35129 Padova - sede operativa: via Dante Alighieri 6 int. 1 35028 Piove di Sacco PD | Luogo di elaborazione presso la sede operativa. L'editore ringrazia per ogni segnalazione o suggerimento inviato a redazione@exeoedizioni.it.



professionisti

pubblica amministrazione

www.catastonline.it - www.exeoedizioni.it

CAPITOLO PRIMO

RICHIAMI DI STATISTICA

1.1 Definizione di statistica

Sono molteplici le definizioni di statistica formulate nel tempo dagli studiosi di questa scienza, per cui, per non fare torto ad alcuno si riporta, una per tutte, la definizione espressa dal "Vocabolario on line della Treccani).

«Scienza che ha per oggetto lo studio dei fenomeni collettivi suscettibili di misurazione e di descrizione quantitativa (spec. quando il numero degli individui interessato è talmente elevato da escludere la possibilità o la convenienza di seguire le vicende di ogni singolo individuo): basandosi sulla raccolta di un grande numero di dati inerenti ai fenomeni in esame, e partendo da ipotesi più o meno direttamente suggerite dall'esperienza o da analogie con altri fenomeni già noti, mediante l'applicazione di metodi matematici fondati sul calcolo delle probabilità, si perviene alla formulazione di leggi di media che governano tali fenomeni, dette leggi statistiche;.....»

La statistica, concepita inizialmente come attività descrittiva dei fenomeni studiati, ha via via ampliato i suoi scopi spostandosi anche verso procedimenti di inferenza statistica, finalizzati all'applicazione dei risultati, ottenuti su un campione statistico di dati effettivamente osservati, sull'intera popolazione.

Gli scopi sono chiaramente di elaborare grandi quantità di informazioni disponibili su certi fenomeni, altrimenti senza procedure statistiche, difficilmente gestibili e comprensibili. Lo studio statistico consente, in particolare, di esprimere in forma quantitativa (numerica) un'impressione anche di tipo qualitativo di un certo fenomeno statistico.

La statistica trova storicamente applicazione per le scienze che studiano un aspetto di una collettività di individui (demografia, economia, sociologia...), medicina, economia, ecc.

1.2 Rami di statistica

In relazione alla definizione in precedenza riportata degli studi statistici, la scienza si può suddividere in due sottoinsiemi:

- *Statistica descrittiva* – intendendo per questa l'insieme di procedure ed elaborazione atte ad analizzare e riassumere i dati che governano un certo fenomeno, per descriverli attraverso funzioni segnatamente mutuata dalla matematica. Lo studio può riguardare

l'intera popolazione o un campione significativo della stessa. Di norma deve sintetizzare tramite pochi valori significativi i contenuti di un vasto numero di rilevazioni.

- *Statistica inferenziale* – costituita dall'insieme delle procedure e analisi per derivare, con l'aiuto di modelli matematici, dai dati osservati su un campione ristretto affermazioni generali estendibili alla popolazione dalla quale il campione è stato estratto.

La statistica descrittiva univariata ha come obiettivo lo studio di fenomeni la cui distribuzione dipende da una sola variabile, o anche se governati da più variabili, ciascuna di essa è singolarmente considerata, all'interno della popolazione (analisi per colonna).

La statistica descrittiva bivariata o più in generale multivariata si occupa dello studio della distribuzione di dati descrittivi del fenomeno, considerando congiuntamente le due variabili (bivariata) ovvero le n variabili e come queste si influenzano reciprocamente (multivariata).

1.3 L'econometria

Nell'ambito della statistica generale, agli inizi del ventesimo secolo, sono stati avviati i primi studi econometrici cioè sviluppati su dati economici oltre che per finalità descrittive, soprattutto per fini predittivi. Cioè per passare dall'analisi del fenomeno reale, all'individuazione della legge matematica (modello) che lo governa e che consente l'individuazione di un valore (visto che siamo nel campo dell'economia) per ogni elemento della popolazione in relazione alle sue caratteristiche specifiche. Come ogni studio statistico, anche quelli econometrici devono rispondere positivamente agli ordinari test statistici.

Questa branca della statistica, applicabile alle stime immobiliari, è quella che sarà oggetto di illustrazione nel presente testo.

Il più ricorrente strumento statistico utilizzato in econometria è sicuramente la regressione lineare (semplice o multipla), nonché la stima del modello lineare con il metodo dei minimi quadrati. Vi sono applicazioni più complesse che utilizzano altri metodi (della massima verosimiglianza e metodo dei momenti), che esulano dai fini della presente trattazione e sono perciò tralasciati.

È bene evidenziare comunque, con forza, i limiti della statistica nelle applicazioni pratiche ingegneristiche, quali la valutazione immobiliare, connessi alla quasi sempre impossibilità pratica di costituire campioni di studio sufficientemente numerosi (prima regola generale ed intrinseca degli studi statistici). I risultati potranno essere statisticamente molto significativi per la conoscenza e lo studio generale del problema, rispondenti per gran parte degli elementi

costituenti l'intera popolazione, ma insufficienti per trattare adeguatamente casi singolari, sempre presenti nel fenomeno osservato, che comunque debbono essere gestiti in queste applicazioni pratiche.

1.4 Alcune definizioni base

Preliminarmente pare opportuno, per comodità del lettore riepilogare alcune definizioni di termini statistici, propedeutici per la trattazione a seguire.

Unità statistica: minima unità di cui si raccolgono i dati (di norma definito anche come elemento della popolazione);

Popolazione: insieme di tutte le unità statistiche:

Caratteri: proprietà oggetto della rilevazione per ogni unità statistica. I caratteri possono essere di due tipi:

- Caratteri qualitativi (rappresentativi di determinati stati che non possono essere espressi con un numero) ed a loro volta possono essere suddivisi tra sconnessi: non c'è alcun ordinamento intrinseco tra le modalità (es: nazionalità di una persona, genere, professione, ecc.) ed ordinabili, quando è possibile ordinare le modalità del carattere in senso crescente o decrescente (es: in un edificio manutenzione ottima, buona, normale, mediocre).
- Caratteri quantitativi: (es: altezza o peso di un individuo, età di una costruzione, superficie di un immobile); questi possono poi essere suddivisi tra discreti se le modalità del carattere sono una serie di limitata (ad es numero ingressi, numero scale, numero stanze di una abitazione) ovvero continui quando le modalità del carattere sono misurate su una scala continua (es: superficie, età di un immobile)..

Variabile casuale: un carattere utilizzato come stimatore di un fenomeno che si presenta secondo modalità diverse nelle varie unità statistiche. In matematica una variabile casuale è detta anche variabile aleatoria o variabile stocastica, quando può assumere valori diversi in dipendenza di altrettanti eventi diversi in uno specifico fenomeno aleatorio.

Osservazione: Il/i valore/i assunto/i dalla/e variabile/i casuale/i in una data unità statistica.

Modalità: modo/valore con il quale si presenta ciascuna variabile causale nelle

varie unità

Campione statistico: è costituito da un sottoinsieme, rappresentativo dell'intera popolazione, ma con un numero di unità statistiche molto limitato, rispetto a quello della popolazione dalla quale è stato estratto.

Lo scopo del campionamento è di raccogliere dati che consentiranno di generalizzare, con un certo livello di probabilità, all'intera popolazione, le conclusioni ottenute dal campione.

La necessità del campionamento è dettata da ragioni pratiche, economiche ed organizzative:

- la popolazione è molto numerosa;
- le risorse economiche sono limitate;
- non tutta la popolazione è accertabile.

Di conseguenza, quando si effettua uno studio su un campione statistico dobbiamo essere consapevoli che si potranno ottenere risultati più o meno approssimati e mai il risultato vero (probabilità 100% coincidenza tra fenomeno osservato ed esito del sondaggio). I motivi sono vari motivi in relazione:

- ai criteri di scelta della popolazione
- al metodo di selezione del campione
- al periodo di osservazione
- alle tecniche di analisi
- alla precisione delle rilevazioni.

Sulla base delle indagini campionarie, si può pertanto accertare la cosiddetta validità interna ed esterna.

Per validità interna si intende che le conclusioni sono corrette per le unità che compongono il campione. La validità esterna rappresenta invece, il grado di generalizzabilità delle conclusioni tratte dallo studio ad ogni unità della popolazione.

1.5 Indici statistici

Nella statistica descrittiva si ricorre a vari indici sintetici o indicatori che possono essere classificati come di:

- Posizione
- Dispersione
- Forma

1.5.1 Indici di posizione

Sono indicatori che descrivono ed evidenziano alcune caratteristiche della distribuzione della popolazione.

I principali indici di posizione sono:

- Media
- Mediana
- Moda

1.5.1.1 La media

media aritmetica è la più semplice e conosciuta misura di posizione di una distribuzione statistica. Ricordiamo che essa è definita dalla somma dei valori di tutte le osservazioni diviso il numero delle osservazioni:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

supponiamo di avere la seguente distribuzione di dati:

1; 2; 10;20;100;150;153;160; 180;1000; 2000;

La media aritmetica è pari a 343,27

media aritmetica ponderata quando siamo di fronte ad una distribuzione di elementi di frequenze raggruppate in classi, la media ponderata è data dalla seguente formula:

$$\bar{X} = \frac{\sum_{i=1}^n f_i \cdot x_i}{\sum_{i=1}^n f_i}$$

dove:

\bar{X} = media della distribuzione in classi,

x_i = valore medio della i-esima classe di intervallo,

f_i = numero di osservazioni della classe i-esima classe,

n = numero di classi,

Σ = sommatoria per tutte le n classi.

Lo scopo perseguito nel calcolo di una media è di sostituire a più dati rilevati un solo numero che dia però una efficace rappresentazione del fenomeno atto ad esprimere l'ordine di grandezza o tendenza centrale dell'insieme dei dati relativi a un fenomeno.

A volte tale dato può non essere significativo se, come nell'esempio sviluppato per la media semplice, i dati sono molto differenti fra loro.

Esistono altri tipi di media (geometrica, media armonica) che non rilevano ai fini del nostro studio.

1.5.1.2 La mediana

Per mediana si intende il valore che occupa la posizione centrale in una serie ordinata di dati. Come la media, anche la mediana può essere calcolata solo se la serie è quantitativa o qualitativa ordinata.

Per il calcolo della mediana bisogna ordinare i dati e rilevare quello che si trova a metà della distribuzione, in modo che il 50% dei valori della serie sia uguale o inferiore a esso e il restante 50% sia superiore, se i dati sono in numero dispari; diversamente è data dalla media aritmetica dei due dati centrali.

Supponiamo di avere la seguente distribuzione di dati (la stessa per la quale è stata calcolata la media aritmetica):

1; 2; 10; 20;100;150;153;160; 180;1000; 2000;

La mediana è pari a 150, mentre la media è 343,27. I valori sono sensibilmente differenti.

La media e la mediana presentano valori simili quando la distribuzione della variabile è simmetrica. Se però la distribuzione presenta forti asimmetrie le due misure possono divergere notevolmente.

CAPITOLO SECONDO

L'ANALISI DI REGRESSIONE

2.1 Definizione e concetti generali

L'analisi di regressione, nelle sue varie forme, è una delle tecniche statistiche maggiormente utilizzate per spiegare la relazione esistente tra una variabile Y detta variabile dipendente (oppure output) e una o più variabili indipendenti (dette anche covariate o regressori, predittori o variabili di input).

Le tecniche di regressione, sulla base della conoscenza di un fenomeno studiato su un campione ridotto della popolazione, possono consentire di costruire una funzione matematica atta ad esprimere la migliore descrizione del fenomeno stesso, applicabile all'intera popolazione.

In sostanza, tra le molteplici soluzioni al problema specifico che può svariare tra ogni campo applicativo della conoscenza umana (sociale, economico, fiscale, medico, ingegneristico, ecc.), la tecnica di regressione che si basa sulla teoria dei minimi quadrati (OLS) individua quella soluzione che statisticamente (sul campione) meglio lo approssima riducendo al minimo gli scarti tra valori reali osservati e quelli restituiti dalla funzione di regressione.

Si cercheranno di affrontare le principali tipologie di regressioni a partire da quella lineare semplice per passare alla lineare multipla e ad altre più complesse quali la logaritmica e semilogaritmica.

In merito alle analisi statistiche, preliminarmente è necessario precisare –come già peraltro un po' noto a tutti per conoscenza diffusa - che la statistica opera sui cosiddetti "grandi numeri" si occupa di fenomeni collettivi o di massa, influenzati da un determinato numero di variabili (inferenze), a volte elevatissimo come nelle stime immobiliari, per cui ogni analisi statistica, affinché abbia un fondamento, deve essere svolta su un numero di osservazioni di fenomeni ampiamente superiore a quello delle inferenze; deve quindi essere ben chiaro che le possibilità previsive dei modelli statistici migliorano all'aumentare del numero dei dati osservati.

Un piccolissimo ***campione statistico*** è formato indicativamente da 30 ÷ 60 osservazioni e per il suo studio è già necessario (considerato il numero di elementi ridotto) adottare una

distribuzione di Student.

La letteratura nel campo statistico evidenzia come criterio pratico, meno restrittivo, per un modello di regressione quello che preveda almeno 30 osservazioni ovvero un numero, se maggiore, pari a circa il numero delle inferenze (numero della variabili aumentato di uno) moltiplicato 5-6 per ogni inferenza.

Nel campo dell'estimo immobiliare, infine è da evidenziare, come meglio si vedrà nelle applicazioni concrete che seguiranno, che per un buon funzionamento dei suddetti modelli è necessario, o quanto meno utile, segmentare preventivamente il mercato immobiliare costituendo popolazioni già accomunate da similarità di caratteristiche incidenti sul valore. Quanto più gli immobili sono simili e più numeroso è il campione osservato, migliore è il valore di stima restituito dalla funzione valore (così come peraltro la stima svolta con procedimenti tradizionali). Di contro, con la segmentazione e l'allargamento del campione crescono gli oneri economici e i tempi per la costituzione dei vari campioni.

Per lo sviluppo dell'analisi di regressione multipla sono disponibili in commercio molteplici software specialistici, ma anche alcuni software di base (fogli di lavoro elettronici) comprendono funzioni statistiche di regressione.

Per agevolare la comprensione delle tematiche trattate saranno sviluppati alcuni esempi applicativi risolti con l'uso di un software molto diffuso nella disponibilità di tutti i professionisti che si occupano di problematiche connesse a contabilità, inventariazione o altri problemi tecnici ingegneristici. Si farà riferimento al foglio di lavoro Excel del pacchetto Office di Microsoft Windows.

Gli esempi, anche a partire dai più semplici, saranno principalmente indirizzati nel settore delle stime immobiliari.

2.2 Il modello lineare

2.2.1 La regressione lineare semplice con intercetta

Noti il valore (y) di una grandezza esprimibile secondo un numero, quindi dato un certo effetto misurabile (variabile spiegata) ed individuata la variabile x che lo ha generato (variabile esplicativa), si intende cercare la relazione funzionale - nel seguito definita **funzione** - che

lega i valori di x ai valori di y in modo da poter prevedere i valori della variabile y sulla scorta dei valori assunti dalla variabile x , in alcuni casi concreti.

La relazione funzionale è la seguente ed è rappresentata dall'equazione di una retta:

$$y = b + m x.$$

dove y è la variabile dipendente, x quella indipendente.

Nel caso delle stime immobiliari la y è costituita dal prezzo di compravendita degli immobili (quando è una grandezza osservata) e la x una caratteristica influente sul valore dell'immobile. La funzione individuata nel caso in esame si può definire **funzione valore** in quanto restituisce il valore (di stima) di ogni immobile una volta noto il valore della caratteristica x conosciuto nel campo di variabilità del valore x riferito all'intera popolazione, cui il campione di studio appartiene.

Oggetto della ricerca sono i parametri b ed m che legano la variabile esplicativa alla variabile spiegata e che rappresentano i coefficienti di regressione.

La determinazione di tali coefficienti conduce a specificare l'equazione della retta e quindi a posizionarla sul piano, dove sono pure riportati i dati osservati.

Ai fini estimativi b_1 rappresenta il prezzo marginale della caratteristica e b_0 il prezzo totale dell'immobile quando la generica caratteristica x è nulla.

Nella dinamica reale di un fenomeno scaturiscono una serie di effetti che non consentono alla relazione originaria di rappresentare un perfetto legame teorico matematico, pertanto si deve tener conto di una componente di indeterminazione "e" stocastica (casuale) nella formulazione del modello, che diventa il seguente:

$$y = b + m x + e$$

Nelle applicazioni la componente "e" viene tralasciata in quanto rappresenterà proprio lo scarto tra valore reale dell'immobile e valore stimato della funzione, che si cerca di minimizzare con una attenta costruzione del modello.

Nella figura seguente sono schematizzati i tre termini che entrano in gioco nella risoluzione dell'individuazione della retta di regressione che tra le infinite ad 1 possibili che meglio approssima il fenomeno.

La retta di regressione sarà quella che rende minima la sommatoria dei quadrati della

regressione, cioè i quadrati delle differenze tra valori stimati dalla funzione e valore medio tra quelli osservati.⁵

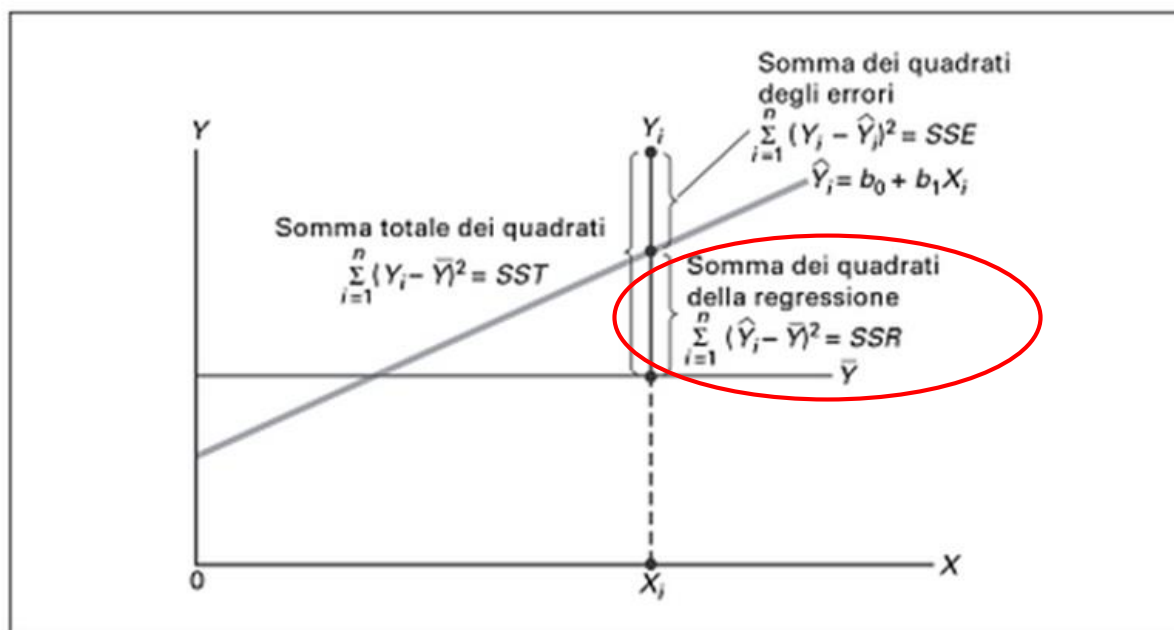


Figura 2.1

Si supponga di avere rilevato, a mero fine didattico, il prezzo di compravendita di un gruppo di 6 immobili appartenenti ad un quartiere di una città che può comprendere un numero notevolissimo di immobili simili tutti accomunati, per come detto, dalla sola ubicazione nel quartiere.

Nella tabella sottostante sono riportati, in terza colonna, i dati costituiti dal prezzo di compravendita (V); nella colonna centrale la superficie di ogni immobile; nella prima colonna è calcolato il valore unitario come rapporto tra prezzo dell'immobile e relativa superficie (V/S). Nei tre righi finali sono riportati dei dati statistici relativi al campione osservato costituiti da valori minimi, massimi e medi.

Si noti che la media aritmetica dei valori unitari è pari a €/mq 2.158,33 (sommatoria dei valori unitari, diviso 6, numero elementi).

Una media statisticamente più rappresentativa sarebbe la media ponderata alla superficie data dalla sommatoria dei prezzi di compravendita divisa la sommatoria delle relative superficie, o l'equivalente dato dal rapporto tra valore medio degli appartamenti e superficie media degli stessi. La media ponderata è pari ad €/mq 2.099,11.

Dati campione

⁵ Per approfondimenti vedi anche glossario

	V _{unit}	Superficie (S)	prezzo(V)
	2600	40	104.000
	2200	80	176.000
	2100	100	210.000
	2050	110	225.500
	2000	110	220.000
	2000	120	240.000
min	2.000	40	104.000
max	2.600	120	240.000
medio	2158,33	93	195.917

Figura 2.2

Nella figura n. 2.3 sono indicate nel piano cartesiano i prezzi delle unità immobiliari e relative superfici.

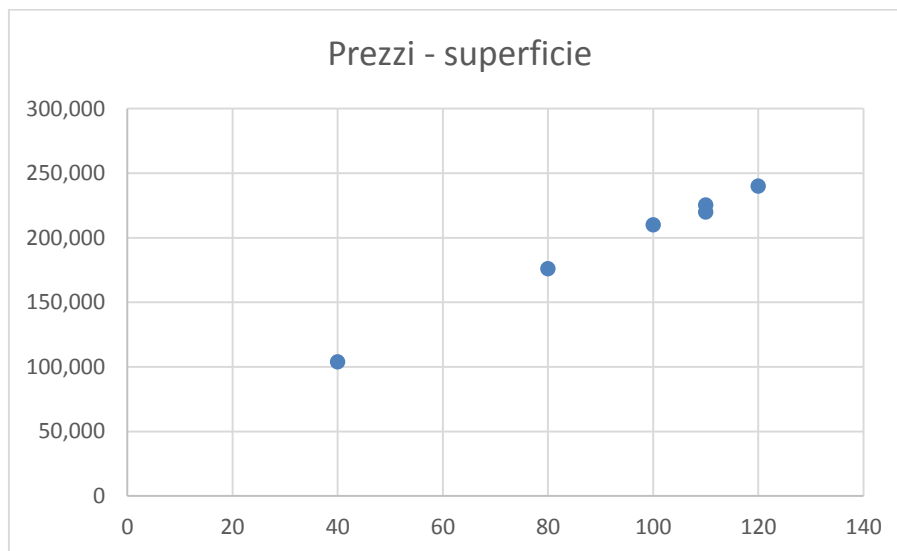


Figura 2.3

Si tratta di individuare la equazione della retta di regressione, disegnata nella figura n. 2.4.

CAPITOLO TERZO

LA STIMA IMMOBILIARE CON I MODELLI STATISTICO-MATEMATICI

3.1 Generalità

Il valore di un immobile, come è facile comprendere anche ai non esperti del settore immobiliare, dipende da molteplici fattori.

I fattori sono in un numero rilevante, ai fini della presente trattazione, si elencano le caratteristiche più importanti, che incidono sul valore dei fabbricati civili e pertanto debbono opportunamente essere tenuti presenti nella stima immobiliare.

Queste caratteristiche sono raggruppabili in due grandi categorie.

La prima relativa a fattori ambientali generali di zona detti caratteri estrinseci e la seconda a caratteri intrinseci.

Tra la categoria dei fattori estrinseci rientrano:

- i caratteri del centro urbano in cui è posto il fabbricato, quali la dimensione demografica, i collegamenti pubblici generali a centri vicini, il clima, la presenza di servizi, la possibilità di offerte di lavoro, la criminalità, ecc.;
- i caratteri generali della zona del centro urbano in cui è sito il fabbricato, quali la posizione (zona centrale, semicentrale, periferica, suburbana, degradata, industriale, agricola, direzionale, ecc.), l'interesse artistico storico, turistico, la dotazione di verde, la salubrità dell'aria, la vicinanza al centro principale degli affari, il livello socio-economico della popolazione, il livello di urbanizzazione, i collegamenti con mezzi di trasporto pubblici, la disponibilità di parcheggi, la qualità edilizia media, l'epoca di costruzione dei fabbricati, ecc.;
- gli stessi caratteri del punto precedente nell'immediato intorno del fabbricato.

I fattori specifici ovvero intrinseci, perché esclusivi dell'immobile sono invece costituiti da:

- tipologia edilizia del fabbricato (edilizia intensiva, estensiva, abitazione in condominio o in villa);
- finiture edilizie del fabbricato (popolari, economiche, civili, di lusso) e la dotazione di impianti tecnologici;
- la vetustà;
- lo stato di conservazione;
- la consistenza del fabbricato;
- il livello di piano;
- gli affacci, la panoramicità, l'esposizione la luminosità.
- l'eventuale interesse artistico storico;

- altre caratteristiche specifiche;
- le servitù attive e passive ed eventuali altri vincoli di natura pubblica o privata.

Da quanto sopra si evince, in linea generale, come mal si presterebbe l'analisi di regressione multipla per la soluzione del problema della stima immobiliare proprio in relazione all'alto numero di variabili in gioco ed in particolare delle loro possibili reciproche interazioni.

L'alto numero di variabili in gioco e la correlata necessità di una rilevante rilevazione campionaria per impostare il modello ha frenato l'applicazione del modello di regressione anche per la complessità delle elaborazioni necessarie ed i connessi calcoli matematici.

Ad oggi, in relazione alle innovazioni tecnologiche dell'informatica che hanno reso disponibili, a basso costo, elaboratori elettronici indispensabili per trattare i complessi sistemi di equazione risultanti dai modelli, in passato possibili solo in pochi centri di ricerca, è venuto a mancare anche questo ultimo impedimento.

Per lo sviluppo delle analisi di regressione, anche per i modelli più complessi, è sufficiente un normale personal computer e programmi statistici, anche solo di base, per sviluppare le analisi.

Rimane tuttavia un altro grosso impedimento per lo sviluppo a tappeto delle analisi di regressione nel campo immobiliare, collegato alla scarsa disponibilità di numerosi dati di riferimento relativi ai prezzi di mercato e alle caratteristiche delle unità immobiliari simili a quella da stimare, per un corretto svolgimento del processo.

Tuttavia, con le opportune accortezze, l'analisi di regressione consente di svolgere previsioni del prezzo di un immobile assai fedeli alla realtà, in quanto il suo principio basilare consiste nell'interpolazione di una funzione teorica con i dati osservati, scelta in modo che meglio di ogni altra si approssimi ai dati stessi.

Inoltre alcuni test statistici sono in grado di fornire anche utili informazioni sull'attendibilità dei risultati.

In ogni caso, come già osservato nel paragrafo 2.2.3, una stima sviluppata con tale procedimento non può sostituire una stima immobiliare tradizionale, in quanto pur adottando ogni opportuno accorgimento (come illustrato nel seguito) non potrà mai, ordinariamente, raggiungere il livello di qualità di una stima puntuale, ovviamente sviluppata nella piena ortodossia della disciplina dell'estimo.

Per cui l'applicazione di questa metodologia estimativa è, in generale, relegata all'utilizzo nell'ambito di un sistema estimativo di massa, quale potrebbe essere quello necessario a chi deve, ad esempio, gestire l'imposizione immobiliare (sistema fiscale). Di fatto, per le stime di massa, quale quelle che occorre sviluppare per la formazione o la revisione del sistema estimativo di un catasto, tesse indifferentemente ad attribuire un valore immobiliare o reddito,

che potrebbero interessare, nel nostro Paese, decine di milioni di immobili, è impensabile di potere procedere con metodi di stima tradizionali.

3.2 Il campione statistico

In statistica un piccolissimo **campione** è formato indicativamente da 30 ÷ 60 osservazioni e per il suo studio è già necessaria adottare una distribuzione apposta (di Student).

Un criterio pratico, meno restrittivo, per un modello di regressione considera necessarie almeno 30 osservazioni aumentate del numero delle inferenze (numero della variabili aumentato di uno) ovvero 5-6 osservazioni per ogni inferenza.

È evidente come questi modelli possano fornire migliori risultati quanto più la popolazione statistica da esaminare sia omogenea. Omogeneità che nel campo immobiliare può essere conseguita segmentando¹⁵ preventivamente il mercato immobiliare per similarità di caratteristiche. Tuttavia ogni segmento richiederà un modello valutativo autonomo, con tutte le conseguenti operatività tra cui, molto onerosa, quella di campionamento, peraltro condizionata alle osservazioni di prezzi di mercato disponibili. Per cui tra le due esigenze di speditezza ed economicità del processo verso qualità dei risultati, occorre trovare un punto comune di equilibrio che soddisfi contemporaneamente le due condizioni.

Si richiamano alcuni concetti preliminari sul campionamento statistico.

Attraverso lo studio di un campione, si può soltanto stimare (cioè determinare con un certo margine di errore) il carattere della popolazione da cui il campione deriva e che pertanto non potrà mai essere determinato con esattezza. L'accuratezza della stima è direttamente correlata al numero di osservazioni che si compiono.

L'errore di campionamento è causa della differenza tra i risultati ottenuti nel campione e quelli reali della popolazione. Le metodologie statistiche ci forniscono indicazione per individuare i limiti probabili della entità dell'errore.

L'errore di campionamento può derivare da:

¹⁵ Un segmento di mercato è costituito da immobili accomunati da caratteri locazionali e tipologici omogenei. La prassi di segmentare il mercato immobiliare è comune a tutti gli operatori immobiliari, sia con professionalità tecnica o meno, per una questione di trasparenza e praticità di lavoro. Di fatto, ad esempio nell'ambito degli immobili urbani, tutte le borse, le quotazioni immobiliari fanno riferimento a suddivisioni per specifiche localizzazioni nell'ambito della città (zone centrali, semicentrali, periferiche, suburbane, ecc.), successivamente le stesse zone sono spesso suddivise in ulteriori segmenti per tipologia edilizia (abitazioni in edifici intensivi, in piccoli condomini, in ville o villini, ecc.), e, ancora, per taglio dimensionale (mini appartamenti, di media o di grande superficie), ecc.

- selezione viziata: se non rigorosamente individuato con casualità, si parla di campione distorto ed i dati ottenuti saranno difficilmente utilizzabili;
- variazione casuale: nonostante la scelta casuale, il campione può comunque essere costituito da elementi nei quali la misura da studiare assume un valore più alto o più basso rispetto al valore vero nella popolazione, senza una regola precisa; ciò può derivare da una incompletezza o non perfetta proporzionalità tra elementi del campione ed elementi della popolazione aventi medesimi caratteri, ad esempio perché un il campione è troppo ridotto.

I metodi di campionamento più comunemente utilizzati sono:

- non probabilistico;
- per randomizzazione;
- sistematico;
- stratificato;
- a grappolo.

Il campionamento non probabilistico estratto in base a criteri di comodo o di praticità non fornisce dati affidabili in quanto non garantisce ad ogni elemento della popolazione la stessa possibilità di essere scelto.

Il campionamento casuale semplice (randomico) si esegue estraendo gli elementi della popolazione con metodo che garantisce la casualità, ottenuta con il sistema dell'estrazione di un numero a partire da un elenco in cui sono presenti tutti gli individui della popolazione. Un campionamento casuale, per essere rappresentativo, spesso richiede una numerosità molto alta, perché un campione piccolo potrebbe essere scarsamente rappresentativo.

Il campionamento sistematico prevede la scelta degli elementi da estrarre dalla popolazione ad intervalli regolari rispetto all'elenco generale. Si tratta ugualmente di un campione casuale come quello precedentemente descritto.

Il campionamento stratificato si effettua scomponendo la popolazione secondo una carattere in essa presente. La popolazione viene suddivisa in strati basati sul fattore che influenza il carattere da studiare e all'interno di ciascuno strato si sceglie un campione con tecnica casuale. Ogni strato avrà numerosità proporzionale alla presenza nella popolazione.

Il campionamento a grappolo prevede la formazione di gruppi (cluster) di unità: i cluster possono essere preformati naturalmente) oppure possono essere determinati. La metodologia facilita il reclutamento degli elementi, però l'errore di campionamento può essere più elevato rispetto ai precedenti metodi.

Nel campo delle stime immobiliari un campione statistico dovrebbe comprendere,

teoricamente, elementi che replicano per ogni caratteristica con le sue varie modalità di presentarsi a parità di altre caratteristiche e possibilmente in numero proporzionale quanto più possibile alla omologa proporzione esistente nella popolazione. Ovviamente tale composizione del campione che deve essere causale, implica fortissime complessità operative anche in relazione alla numerosità e natura dei dati disponibili dai quali formare il campione stesso.¹⁶ Occorrono quindi tecniche di campionamento adeguatamente studiate per l'applicazione pratica da sviluppare, ma che siano contestualmente rispettose dei principi della statistica.

Operativamente, per lo sviluppo dell'analisi di regressione multipla sono disponibili in commercio molteplici software specialistici.

Anche alcuni software di base (fogli di lavoro elettronici), come già si è fatto osservare, comprendono funzioni statistiche di regressione e che saranno utilizzati nella presente trattazione per gli esempi sviluppati.

3.3 I modelli utilizzabili

Come è stato evidenziato nel capitolo 2, i modelli statistico-matematici più idonei per le stime immobiliari sono rappresentati dalle tre seguenti tipologie:

Modello lineare

$$V = b_0 + b_1 x_1 + \dots + b_i x_i + \dots + b_n x_n + \varepsilon$$

Modello semilogaritmico esponenziale nelle variabili

$$V = b_0 b_1^{x_1} b_2^{x_2} \dots b_i^{x_i} \dots b_n^{x_n} + \varepsilon$$

Modello logaritmo esponenziale nei coefficienti

$$V = b_0 x_1^{b_1} x_2^{b_2} x_i^{b_i} \dots x_n^{b_n} + \varepsilon$$

¹⁶ Ad esempio per la riforma del catasto un limite insuperabile è costituito dai prezzi di mercato disponibili desunti da atti di compravendita. È evidente il limite intrinseco che in qualche modo deve essere superato; diversamente la funzione di stima non potrà mai avvicinarsi adeguatamente al valore di un immobile se le sue caratteristiche non sono possedute da qualche elemento campionario.

Dove:

- V = variabile dipendente (valore immobile)
- b_0 = fattore costante relativo alle caratteristiche locazionali riferite al segmento di mercato e al campione considerato, in € a m^2 di superficie
- b_i = prezzo marginale della caratteristica i -esima in € per unità di misura della caratteristica;
- x_i = variabile indipendente, caratteristica posizionale e tipologica generica.

Lo sviluppo della regressione porta alla determinare preliminarmente del valore dei suddetti coefficienti (b_i, b_0).

Una volta determinati questi coefficienti la funzione è nota ed è in grado di fornire il valore di tutti gli immobili (previo inserimento dei dati relativi alle caratteristiche X_i specifiche degli immobili stessi) appartenenti al segmento di mercato per il quale la funzione è costruita¹⁷, attraverso uno specifico campionamento di elementi, dei quali siano noti i parametri tecnici, oltre che economici (valore o reddito).

Il valore ricavato dalla funzione è, per come detto, effetto dell'indeterminazione e_j della quale può essere calcolato il valore più probabile, espresso ad esempio in termini percentuale di scostamento massimo e minimo rispetto al valore reale della grandezza osservata (nella fattispecie il valore immobiliare).

Questi modelli matematici mirano a spiegare le relazioni di tipo causa-effetto esistenti tra un insieme di variabili fisiche, tecniche, economiche, ecc., dette variabili esplicative o indipendenti, e la variabile spiegata (dipendente) costituita generalmente dal valore di mercato.

Affinché la funzione possa esprimere il valore immobiliare, occorre individuare correttamente il campo di variabilità di ciascuna variabile e la legge di associazione del valore di ciascuna variabile alla caratteristica specifica dell'immobile. E' indispensabile per una massima oggettivizzazione dei risultati del processo che i dati economici dei beni risultino da rilevazioni di mercato (prezzi dichiarati in atti di compravendita) e non siano frutto di valutazioni specifiche con metodologie tradizionali.¹⁸

¹⁷ I risultati della funzione potranno fornire una adeguata approssimazione del valore immobiliare reale soltanto se il valore dei parametri degli immobili per i quali si ricerca il valore interpolato della funzione ricade nel campo di variabilità del campione estimativo costituito dalle osservazioni di mercato sulle quali sono state eseguite le analisi di mercato.

¹⁸ Come noto fonti di informazioni dei valori immobiliari possono essere costituite da:

- indagini dirette presso pubblici Uffici (Ufficio delle entrate, ufficio del territorio, ecc.) per rilevare i prezzi dichiarati nei trasferimenti;
- agenzie immobiliari;
- annunci su quotidiani;
- bollettini immobiliari pubblici (Camere di commercio, ecc.) e privati;

Un'ulteriore problematica, preventivamente da superare, è costituita dal fatto che molto spesso, alcune caratteristiche qualitative dell'immobile sono espresse con termini descrittivi quali, mediocre, buono, ottimo, ecc. che devono essere trasformati in coefficienti numerici affinché il modello matematico possa elaborare i calcoli.

Ciò premesso, nel seguito si illustrano nelle linee generali i suddetti procedimenti, con degli esempi tipo, ancora in linea orientativa non essendo stato emanato il decreto legislativo attuativo della legge delega (n. 23/2014) per la riforma del catasto dei fabbricati.

3.4 La scelta delle variabili

Le variabili (stimatori di regressione) da prendere in esame, perché più incidenti sul valore immobiliare, non sono sempre le stesse in ogni situazione operativa, ma variano da caso a caso a seconda del segmento di mercato oggetto di indagine (ad esempio tra abitazioni, negozi ed uffici) e da luogo a luogo in relazione alle caratteristiche specifiche del patrimonio immobiliare presente (ad esempio luoghi di interesse turistico e storico, zone rurali e zone urbane ordinarie).

Si premette in linea generale (cfr. paragrafo 1.5.5.5.8), che, per semplificazione operativa e per un migliore funzionamento del modello statistico, è opportuno ridurre quanto più possibile il numero delle variabili da utilizzare, verificando, però, contemporaneamente che la semplificazione non incida negativamente sul livello di qualità dei valori forniti dalla funzione valore.

Nei paragrafi a seguire si esamina un esempio di regressione multipla sviluppata su uno stesso campione con i tre modelli già analizzati ai paragrafi 2.2.4, 2.3 e 2.4 rispettivamente lineari ed esponenziali (il primo esponenziale nelle variabili, il secondo esponenziale nei coefficienti).

Gli esempi, sviluppati sullo stesso campione (per confrontare i risultati) illustrano una operatività specialistica per la individuazione della funzione di stima del valore immobiliare di un segmento costituito da immobili adibiti ad uso abitativo di un vasto quartiere di una città nel rispetto dei principi generali contenuti nella legge delega per la riforma del sistema estimativo del catasto dei fabbricati (salvo specifici accorgimenti che potrebbero essere introdotti dal decreto legislativo attuativo).

-
- notai;
 - stime istituti di credito;
 - stime tribunale;
 - altre indagini dirette.

3.5 Esempio di regressione lineare multipla con excel

Si sviluppa un esempio di regressione su un campione di 60 unità immobiliare di tipo abitativo ubicate in fabbricati condominiali del quale si è rilevato per ogni elemento il prezzo in una recente compravendita e le caratteristiche tecniche rappresentate dalle seguenti variabili, ritenute più influenti sul valore nel contesto urbano analizzato.¹⁹

X1 Superficie	X2 Intorno		X3 Tipologia		X4 Stato manut		X5 Piano		X6 Affaccio	
	Qualità	Valore	Qualità	Valore	Qualità	Valore	Qualità	Valore	Qualità	Valore
espressa in mq. di superficie commerciale	ricercato	1,20	signorile	1,10	ottimo	1,10	alto	1,10	pregiato	1,20
	ordinario	1,00	civile	1,00	normale	1,00	medio	1,00	ordinario	1,00
			economic	0,80	pessimo	0,80	basso	0,90		

Non è stata inserita la variabile presenza di ascensore, ipotizzando che l'ascensore sia sempre presente in quanto il quartiere analizzato è di epoca di costruzione successiva all'introduzione ordinaria, in ogni tipologia edilizia di costruzione, dell'impianto di elevazione. In caso di assenza di ascensore negli elementi campionari si è verificato che l'abitazione sia posta al piano terra o primo (per contemperare il possibile errore).²⁰ Per la valutazione di abitazioni a piani superiori, in assenza di ascensore si può applicare un coefficiente correttivo desunto dal mercato (rapporto ordinario tra valore unitario di un'abitazione con ascensore e quello di abitazione del tutto simile, ma priva di ascensore).²¹ Oltre a selezionare le variabili occorre individuare una scala di variabilità numerica associata alle caratteristiche di natura qualitativa. Nell'esempio, in pratica tutte tranne la superficie.²² Per la superficie si utilizzano i criteri

¹⁹ Si ricorda che le variabili da analizzare devono essere selezionate in rapporto alle caratteristiche della popolazione indagata. Per cui, a volte, anche se la variabile è in generale molto incidente sul valore, se in un determinato contesto è pressoché costante, non gioca alcun ruolo. Di contro potrebbero risultare determinanti altre caratteristiche, qui non prese in esame.

²⁰ È evidente che se nella popolazione esistessero fabbricati privi di ascensore occorre introdurre un coefficiente di aggiustamento (riduttivo) del valore della funzione per le abitazioni poste a piani diversi da terra e primo. Esistono tabelle parametriche che monitorano la differenza di valore al variare del piano di ubicazione in presenza ed in assenza dell'impianto di ascensore.

²¹ Diversamente può essere inserita una variabile dummy (presenza ed assenza di ascensore) ed allargare il campione statistico anche ad unità site in edifici privi di ascensore).

²² Occorre codificare per ogni variabile individuata caratteristiche che la descrivono, qualitativamente e numericamente. Ad esempio, nel caso in esame:

Intorno: caratteristica posizionale, rilevata a livello di fabbricato. Individua la qualità dell'ambito territoriale di prossimità del fabbricato in cui è ubicata l'unità immobiliare. La caratteristica presenta, di norma, le seguenti modalità: - *ricercato*: quando la qualità dell'ambito territoriale di prossimità del fabbricato presenta un livello significativamente superiore a quello riscontrato con prevalenza nell'ambito dell'ambito territoriale di ubicazione; - *ordinario*: quando la qualità dell'ambito territoriale di prossimità del fabbricato rispecchia il livello riscontrato. In alcuni casi potrà individuarsi anche un terzo livello qualitativo inferiore all'ordinario.

Tipologia edilizia: rilevata a livello di fabbricato, individua, in forma sintetica, una serie di caratteri originari del fabbricato in cui è ubicata l'unità immobiliare, come la distribuzione degli spazi, la qualità architettonica, la qualità dei materiali e delle finiture, la dotazione ed il pregio delle parti comuni. Sono state scelte le seguenti modalità: - *signorile*: quando il fabbricato presenta caratteri di ricercatezza progettuale e singolarità delle

commerciali per il calcolo previsti dagli operatori locali. Nell'esempio, coincidono con i criteri di cui all'allegato C al DPR n. 138/98.

Si evidenzia, che dal punto di vista del calcolo matematico non ha alcuna rilevanza sull'esito della regressione (Y stimata) il valore numerico associato al livello qualitativo della singola variabile. Ad esempio possono essere numeri ordinali (1,2,3,...) o coefficienti approssimativamente tarati, in base all'esperienza, sull'incidenza nel valore delle varie modalità secondo cui la caratteristica si può presentare; ciò che incide sull'output elaborato (y) è il numero dei livelli e la distanza che c'è tra gli stessi.²³

Nell'esempio che si sta sviluppando, per la variabile piano, l'interesse tra i livelli è costante. Per lo stato di manutentivo e la tipologia edilizia l'ampiezza dell'interesse non è costante.

Si fa osservare inoltre che le variabili "intorno" e "affaccio" potrebbero essere anche inserite come due variabili "dummy", di tipo binario (0 – 1).

Nella figura 3.1 tabella 1 parte 1^a e figura 3.2 tabella 1 parte seconda, di seguito riportate sono elencate le varie unità immobiliari costituenti il campione.

Ai margini della tabella sono state messe in evidenza righe e colonne del foglio Excel sul quale

soluzioni architettoniche; *civile*: quando il fabbricato presenta soluzioni progettuali ed architettoniche rispondenti alle ordinarie richieste del mercato immobiliare. Realizzato con componenti di rivestimento e di finitura di media qualità, dotato di spazi pertinenziali, comuni adeguati. - *economica*: quando il fabbricato si caratterizza per una estrema semplicità ed essenzialità delle soluzioni progettuali ed architettoniche. Il fabbricato presenta prospetti privi di qualsiasi articolazione compositiva, facciate principali prive di una regolare ed equilibrata distribuzione di balconi, utilizzo di materiali economici nella prevalenza dei rivestimenti esterni e nelle finiture, minimali spazi pertinenziali, comuni. Si presume che i caratteri dell'unità specifica siano conformi a quelli del fabbricato.

Stato di manutenzione:, rilevata a livello di fabbricato, individua, in forma sintetica, lo stato di conservazione complessivo, con riguardo alle strutture, alle facciate, agli infissi esterni e alle parti comuni in genere. Si presume che i caratteri dell'unità specifica siano conformi a quelli del fabbricato per i punti di interesse (pavimenti, infissi, servizi igienici ed impianti). La caratteristica è prevista con le seguenti modalità: - *ottimo*: quando il fabbricato presenta condizioni fisiche ottimali in tutte le sue parti comuni, ivi comprese quelle strutturali, di tamponamento e di finitura, tra cui le facciate, gli infissi, le coperture e le aree esterne, tali da non richiedere, nel breve periodo, alcun rilevante intervento di ripristino o rifacimento delle stesse. - *normale*: quando il fabbricato presenta una condizione fisica complessivamente accettabile, in quanto non richiede, nel breve periodo, interventi di ripristino o rifacimento della prevalenza delle sue parti comuni, ivi comprese quelle strutturali, di tamponamento e di finitura, tra cui le facciate, gli infissi, le coperture e le aree esterne. - *pessimo*: quando il fabbricato presenta un diffuso stato di degrado o di danneggiamento in tutte o nella prevalenza delle sue parti comuni, tale da richiedere, nel breve periodo, rilevanti interventi di ripristino o di rifacimento complessivo delle stesse.

Piano: caratteristica posizionale, rilevata a livello di singola unità immobiliare. Individua la qualità del livello di piano di ubicazione dell'unità immobiliare. La caratteristica presenta le seguenti modalità: - *alto*: quando l'unità immobiliare è posta ad un piano superiore ai 2/3 dei complessivi. *medio* quando l'unità immobiliare è posta ad un piano compreso tra 1/3 e 2/3 dei complessivi - *basso*: quando l'unità è posta al piano inferiore ad 1/3 dei piani complessivi dell'edificio.

Affaccio: individua la presenza, o meno, di affacci su vie, piazze e bellezze naturali di particolare pregio. La caratteristica presenta le i modalità: - *pregiato*: quando l'unità immobiliare presenta uno o più affacci su vie, piazze e bellezze di particolare pregio sotto il profilo storico, artistico o naturalistico. - *ordinario*: quando l'unità immobiliare non presenta affacci di particolare pregio.

²³ Usa una scala o l'altra di misurazione della variabilità di una caratteristica incide sul valore mutuo dei coefficienti di regressione, ma non sul valore della variabile dipendente.

è stata sviluppata l'esercitazione per i necessari riferimenti alle formule nel seguito illustrate.

Per ogni unità immobiliare è stato riportato il prezzo di compravendita (colonna R evidenziata con fondo verde) ed in corrispondenza di ogni colonna indicativa di una variabile è stato indicato il valore di competenza secondo le caratteristiche dell'immobile e la scala numerica di misurazione assunta per ciascuna variabile.

Nelle colonne a destra della R sono riportati gli esiti della regressione che si illustreranno più avanti.

In coda alla tabella (seconda parte), ai righe da 66 a 69 sono riportate alcune informazioni statistiche sul campione:

- valore minimo e massimo di ogni variabile (ad esempio si evince che la superficie delle unità immobiliari costituenti il campione varia da mq. 58 a mq. 189); la superficie media è di mq. 113.
- valore minimo e massimo del prezzo osservato (€ 156.000 - € 518.000);
- variazione percentuale tra valore minimo e massimo della superficie (325,86%) e del prezzo osservato (332,05%).

3.9 Verifiche finali sulla funzione di stima

Per pervenire al collaudo complessivo delle operazioni di definizione di una funzione di stima valida per l'intera popolazione da cui è stato estratto il campione, oltre alle verifiche già in precedenza esaminate ed elencate al paragrafo 1.5.5.5.7 è necessario verificare gli esiti della funzione su **un secondo campione, cosiddetto di controllo**, estratto casualmente dalla popolazione.

Dal confronto del valore di stima calcolato dalla funzione e del prezzo osservato sui vari elementi del campione di controllo si calcolano i relativi rapporti e si sviluppano i calcoli per pervenire al valore di:

- scarto medio;
- COV
- COD
- PRD

come illustrato negli esempi di paragrafi 3.5, 3.6 e 3.7 (per fini didattici direttamente sul campione principale piuttosto che su quello di verifica) per verificare se i valori degli indici rientrano nei parametri di tolleranza ordinariamente assunti o legalmente fissati.

Diversamente vanno adottati gli interventi per superare le criticità rilevate, sia per gli aspetti statistici che per quelli di coerenza tecnico economica delle stime.

Gli interventi sono molteplici e vanno dalla verifica dell'esatta rilevazione dei dati del campione di controllo, di aggiunta di ulteriori elementi al campione di controllo, di assunzione di un nuovo campione di controllo, fino a ripartire da zero dalla costituzione del campione principale, aggiungendo o eliminando una o più variabili.

Superate le criticità statistiche e di coerenza tecnico-economica, pervenendo ad un campione principale coerente con quello di controllo, normalmente si fondono i dati tra campione principale e quello di controllo, ottenendo una funzione di stima, che secondo i canoni statistici sarà un prodotto migliore, in quanto basata su un campione di studio più ampio.

Per gli aspetti pratici operativi estimativi, poiché la funzione di stima deve operare i suoi effetti sull'intera popolazione, quale potrebbe essere, ad esempio, il caso nell'ambito della riforma del catasto dei fabbricati con la rivalutazione di tutte le unità immobiliari presenti sul territorio nazionale, deve essere posta l'attenzione al problema degli outlier (elementi che hanno un'influenza sproporzionata sul modello stimato, fornendo così stime dei regressori non rappresentative della realtà).

Questi elementi della popolazione se (e sarebbe un bene) incappassero nel campione statistico di studio, per gli aspetti statistici, potrebbero essere eliminati solo e se il loro peso statistico dovesse essere modesto.