

**MÁSTER UNIVERSITARIO INTERNACIONAL EN: "CATASTRO
MULTIPROPÓSITO AVANZADO" Y " CATASTRO Y VALORACIÓN"
TÍTULO PROPIO DE LA UNIVERSIDAD DE JAÉN
PARAJE LAS LAGUNILLAS S/N - JAÉN - 23071 - ESPAÑA**



Análisis Estadísticos de Series Pequeñas y Muy Pequeñas aplicado a la Valuación Inmobiliaria

Trabajo de Investigación para la obtención del
Título de Máster Universitario Internacional en
"Catastro Multipropósito Avanzado" y "Catastro y
Valoración"

Elaborado por:

Atilio Roberto Piol Puppio

[Mayo 2014]



FICHA DE TRABAJO

Alumno:	Ing. Atilio Roberto Piol Puppio
Título del Trabajo de Investigación:	Análisis Estadísticos de Series Pequeñas y Muy Pequeñas aplicado a la Valuación Inmobiliaria
Objetivo:	Demostrar que el uso de ciertas herramientas estadísticas no convencionales, permitirían al profesional tasador obtener una mejor precisión en el cálculo del "Estimador de la Tendencia Central", en series con muy pocos datos y una dispersión desconocida
Justificación:	En las grandes ciudades de Hispanoamérica, es posible seleccionar una suficiente cantidad de "Inmuebles Comparables" (muestra), necesarios para poder realizar una asignación valuatoria. No sucede lo mismo en la provincia, donde se formalizan muy pocas operaciones de Compra – Venta de inmuebles, haciendo casi imposible un estudio estadístico significativo para determinar el valor de un inmueble
Resultados Esperados:	Se clasificaron las serie en: Series Grandes, Series Pequeñas y Series Muy Pequeña; asignándole a las dos últimas metodologías estadísticas no convencionales diferentes a los modelos gaussianos convencionales basados en definir una "Función de Residuos" distinta a la propuesta por el método de los mínimos cuadrados
Actividades Realizadas:	Para las series grandes, se citó someramente la metodología estadística para datos agrupados en clases. Para las series pequeñas: Se adaptó a la Teoría Estadística



	<p>de Errores como "metodología alterna" para el cálculo del término central de una serie</p> <p>Para las series muy pequeñas, debido a que no es posible asumir en ellas una distribución normal y que por consiguiente no puede utilizarse el promedio aritmético como estimador del término central: Se adaptó la teoría de Estimación Robusta o No Paramétrica, tanto los tipo "MVT" (Estimador de Recorte Multivariante ó M-Estimadores); como los "MLT" (Estimador de Recorte de Máxima Verosimilitud) (Maximun Likelihood Test) ó (L-Estimadores). Diferenciándose que mientras los primeros eliminan o recortan los valores atípicos; los últimos, establecen "pesos" permitiendo aprovechar toda la data observada.</p>
Fecha y Firma:	Caracas, Mayo de 2014



DEDICATORIA Y AGRADECIMIENTO

A la memoria de mi hijo Ángel Camilo Piol Segura; quien a muy corta edad, fue llamado a la Presencia de nuestro Creador.

A mi querida esposa Iris del Rosario Segura Severino, compañera, colega, novia, camarada, amiga; con quien un día emprendí la más grande aventura de mi vida, que aún no termina.

Al Profesor Dr. Manuel Alcázar Molina, quien hizo posible que una gran cantidad de profesionales Iberoamericanos en valuación y catastro, pudiesen aspirar a un Título de Cuarto Nivel, no asequible en sus respectivos países. Quien tuvo la confianza en promoverme como profesor de estadística aplicada a la valuación inmobiliaria en la Universidad de Jaén, siendo uno de los honores más grandes que recibiese en mi vida y quien sin su ayuda no hubiese podido lograr esta Maestría.

A mis alumnos del Curso Profesional de Valuación Inmobiliaria de la Sociedad de Ingeniería de Valuación de Venezuela (SOITAVE). A quienes desde 1989, les he transferido todo mi conocimiento y experiencia.



INDICE GENERAL

ABSTRACT	7
1.0 LAS SERIES PEQUEÑAS.....	8
2.0 LA METODOLOGIA DE COMPARACION DIRECTA O APROXIMACION AL MERCADO.....	9
2.1 ¿Qué es el Método de Comparación Directa o Aproximación al Mercado?.....	9
2.2 Inmuebles Sustitutos: Semejantes o comparables	10
2.3 Definición del Método de Aproximación al Mercado.	10
3.0 METODOLOGIA PARA EL LEVANTAMIENTO DE LA DATA.....	11
3.1 Criterios de búsqueda y selección de referenciales o comparables	11
4.0 SERIES GRANDES, PEQUEÑAS Y MUY PEQUEÑAS	19
5.0 SERIES MAYORES DE VEINTE (20) DATOS:.....	21
6.0 SERIES ENTRE 10 Y 19 DATOS.....	25
7.0 SERIES MUY PEQUEÑAS	30



7.1 Definición	30
7.2 Justificación para el uso de Métodos "No Paramétricos" o "Robustos" en Series Muy Pequeñas:.....	31
7.3 Definición de Valores Atípicos:.....	31
7.4 Tipos de estimadores "No Paramétricos" o "Robustos":.....	32
7.5 El Estimador de Huber:.....	32
7.6 El Estimador- M de Huber:	33
7.7 El Estimador-L de Huber.	36
8.0 CONCLUSIONES	40
BIBLIOGRAFIA.....	43
APENDICE 1	46
EJEMPLO NUMERICO RESUELTO: La Teoría Estadística de Errores, como metodología para la determinación del Término Central de una "Serie Pequeña"	46
APENDICE 2	54
EJEMPLO NUMERICO RESUELTO: Uso del Estimador-M de Huber, como metodología para la determinación del Término Central de una "Serie muy Pequeña"	54
APENDICE 3	68
EJEMPLO NUMERICO RESUELTO: Uso del Estimador-L de Huber, como metodología para la determinación del Término Central de una "Serie muy Pequeña"	68



ABSTRACT

El presente Trabajo de Investigación, se basa en el estudio estadísticos de series pequeñas de inmuebles y de su aplicación a la Valuación Inmobiliaria.

En las grandes ciudades de Hispanoamérica, es posible seleccionar una suficiente cantidad de "Inmuebles Comparables", necesarios para poder determinar su "Término Central" mediante un análisis estadístico preciso de dichas series, conforme a los requerimientos contemplados en la Metodología. No sucede lo mismo en la provincia, donde se formalizan muy pocas operaciones de Compra – Venta de inmuebles, haciendo imposible un estudio estadístico significativo.

Se pretende demostrar que el uso de ciertas herramientas estadísticas no convencionales, permitirían al profesional tasador obtener una mejor precisión en el cálculo del "Estimador de la Tendencia Central", en series con muy pocos datos y una dispersión desconocida.

Se clasificará la data en: Series Pequeñas y Series Muy Pequeña; asignándoles metodologías estadísticas no convencionales, tales como "Teoría Estadística de Errores" y "Estimaciones Robustas o No Paramétricas"; fundamentadas estas últimas, en la definición de "Funciones de Residuos" distintos al propuesto en los modelos gaussianos convencionales.

Palabras Claves:

avalúo, mercado, "series pequeñas", "estimación robusta", "estimación no paramétrica", haber, residuos, "valores atípicos", outliers, pesos, errores, incertidumbre, "error accidental", M-Estimator, L-Estimator, Estimator-M, Estimator-L



1.0 LAS SERIES PEQUEÑAS

El presente Trabajo de Investigación, se basa en el estudio estadísticos de series pequeñas de inmuebles referenciales o comparables para su aplicación a la Valuación Inmobiliaria.

En las grandes ciudades de Hispanoamérica, es posible seleccionar una suficiente cantidad de "Inmuebles Comparables" (muestra), necesarios para poder determinar su "Término Central" mediante el análisis estadístico preciso de dichas series, conforme a los requerimientos contemplados en la Metodología Valuatoria de "Comparación Directa" o también conocida como "Aproximación al Mercado".

No sucede lo mismo en la provincia, donde se formalizan muy pocas operaciones de Compra – Venta de inmuebles, haciendo imposible un estudio estadístico significativo.

Se define como "Serie Pequeña", a una muestra de datos que se supone normalmente distribuidos, pero de la cual se desconoce su varianza o tendencia a desviarse del valor promedio.

También podría definirse como una colección menor o igual a Veinte (20) datos¹.

Hay que hacer notar que, en la mayoría de las pequeñas y medianas poblaciones del interior de Hispanoamérica y para cualquier asignación de investigación inmobiliaria; la tarea de seleccionar por lo menos Veinte (20) Inmuebles Comparables, resulta muy difícil o casi imposible.

¹ CROSBIE, J. (1993): "Interrupted time-series análisis with brief single-subject data". Publicación Periódica: Journal of Consulting and Clinical Psychology, 61, 966-974. Washington. EE.UU.



Se pretende demostrar en este Estudio, que el uso de ciertas herramientas estadísticas no convencionales, permitirían al profesional tasador obtener una mejor precisión en el cálculo del "Estimador de la Tendencia Central", en series con muy pocos datos y una dispersión desconocida.

2.0 LA METODOLOGIA DE COMPARACION DIRECTA O APROXIMACION AL MERCADO

Es importante entender la importancia que tiene el obtener una razonable serie de Inmuebles Referenciales o Comparables. Esa es la base indispensable para poder aplicar la "Metodología de Comparación Directa", también conocida como "Método de Aproximación al Mercado". Sin esa fuente de información, no hay posibilidad de determinar el valor de un inmueble por dicho método.

2.1 ¿Qué es el Método de Comparación Directa o Aproximación al Mercado?

El Método de Comparación Directa o de Aproximación Mercado consiste en la obtención del VALOR de un inmueble, mediante la definición de "Valor de Sustitución"; o sea, calcular el Valor de dicho bien raíz, aproximándolo su valor más probable a través del estudio estadístico de una serie de "Inmuebles Similares" o "Inmuebles Comparables", en su propio vecindario y en un lapso de tiempo razonable.

El término "Inmuebles Similares", es la clave de la aplicación del Método.

Son ejemplos de Inmuebles Similares:

- *Los apartamentos de igual área y distribución de un edificio.*
- *Las parcelas de terreno de similar superficie y topografía en una misma urbanización.*
- *Las casas pareadas de igual arquitectura de un conjunto residencial.*
- *Los galpones industriales de iguales características de un complejo manufacturero*

□

Como se podrá observar en el recuadro anterior, todos los inmuebles anteriormente listados son Inmuebles Similares entre sí; sin embargo, en la vida real es muy difícil o casi imposible obtener



una serie significativa de dicha clase de inmuebles al momento a que un Profesional Tasador se le contrate un avalúo.

Es fundamental, que el profesional tasador entienda la importancia de la obtención de referenciales apropiados, lo mas similares posibles al objeto de estudio, en una asignación.

2.2 Inmuebles Sustitutos: Semejantes o comparables

Los Inmuebles Sustitutos o Semejantes, existen gracias al principio de "competencia del mercado inmobiliario"; que exige que: Las propiedades similares tengan aproximadamente parecidos valores.

Partiendo de las siguientes premisas, Primera: No existen dos inmuebles "exactamente iguales" y Segunda: Es poco probable encontrar suficientes referenciales de "Inmuebles Similares" para un análisis estadístico significativo. El profesional tasador deberá recurrir a la búsqueda de "Inmuebles Sustitutos Similares", los cuales si bien podrían sustituir al inmueble objeto del avalúo diferirían de él en algunas características físicas o de ubicación o ambas; y el profesional tasador se obligaría entonces a "Corregirlos" o sea "HOMOLOGARLOS" a fin de poder forzarlos a través de expresiones lógico-matemáticas, a que se asemejen al inmueble a ser avaluado.

2.3 Definición del Método de Aproximación al Mercado.

El Método de Comparación Directa o Aproximación al Mercado (Marketing Approach, en idioma inglés); se basa en el "Principio de la Sustitución" y se define como: La Cantidad de Dinero estimado, en que un vendedor cedería un bien a un comprador, actuando ambos libremente y sin presiones externas, teniendo ambos un conocimiento del mercado existente en ese sector y para ese momento.



3.0 METODOLOGIA PARA EL LEVANTAMIENTO DE LA DATA

3.1 Criterios de búsqueda y selección de referenciales o comparables

Un error que cometen muchos tasadores profesionales, es no tener una planificación ni estrategia en la búsqueda de los inmuebles referenciales o comparables. La precisión del resultado de un avalúo es directamente proporcional a la calidad de los referenciales o comparables seleccionados.

Antes de proceder a la búsqueda de los referenciales o comparables, es indispensable que el profesional tasador haya realizado la inspección del inmueble objeto y su vecindario.

De nada sirve una gran cantidad de datos, si los mismos no cumplen medianamente con los criterios de comparabilidad con el inmueble a valorar. Es preferible una serie pequeña de referenciales medianamente comprables al inmueble objeto, que una serie grande de referenciales disímiles al bien a tasar.

Para la búsqueda y selección de inmuebles referenciales o comparables, hay que recordar que el Método de Comparación Directa o Aproximación al Mercado, se basa principalmente en el "Principio de Competencia del Mercado Inmobiliario": "Inmuebles similares se venderán a precios similares". Por lo tanto el profesional tasador, deberá dirigir su búsqueda a los inmuebles lo mas similar posibles al bien a tasar.

Debido a que no existen dos avalúos iguales, tampoco existirán reglas universales para el sondeo y la selección de los referenciales o comparables. Cada búsqueda deberá ser tratada de forma

individual y en base a las características físicas, ubicación y uso de cada uno de los casos asignados al profesional tasador.

3.2 Primer Criterio: Selección de una poligonal de búsqueda o canevas

Se define como una "Poligonal de Búsqueda" o "Canevas", una superficie trazada sobre un mapa del vecindario donde se ubica el inmueble a tasar, de donde se seleccionarán los inmuebles referenciales o comparables que podrían ser incorporados para la determinación del valor del bien objeto a valorar.

Ejemplo de Aplicación

Se va a valorar el apartamento N° A-02-01 de la Torre "A" del edificio "Pórtico del Este", ubicado entre las Avenidas Las Acacias y Francisco Solano López del sector "Sabana Grande". Parroquia "El Recreo". Municipio Libertador. Caracas. Venezuela.



Area del Apartamento: 60,00 M2
Edad del Edificio: 5 años
Estado de Conservación: Normal

El primer criterio para la selección de los inmuebles referenciales o comparables, necesarios para la determinación del valor del apartamento, será trazar un "Poligonal de Búsqueda" en el vecindario donde se ubica el edificio "Pórtico del Este":



Nótese que la "Poligonal de Búsqueda" gravita alrededor del Edificio "Pórtico del Este", marcado con un círculo y una flecha.

Nótese que los linderos de la poligonal coinciden con las trazas de cuatro importantes avenidas del sector.

Por lo tanto, se seleccionarán únicamente apartamentos ubicados dentro de dicha "Poligonal de Búsqueda", como probables inmuebles referenciales o comparables, para la obtención del valor del apartamento objeto.

Los inmuebles ubicados fuera de la poligonal, no serán tomados en cuenta en el estudio valorativo.

Esta "Poligonal de Búsqueda", permitirá satisfacer las siguientes premisas:

- Inmuebles con similar ubicación o vecindario



- Inmuebles que compartan los mismos servicios públicos y privados
- Inmuebles con similar uso o zonificación

3.3 Segundo Criterio: Selección de inmuebles referenciales o comparables de similar área o superficie

Dentro del mismo Ejemplo N° 1, en la "Poligonal de Búsqueda" trazada, se ubicarán los inmuebles referenciales o comparables con un área igual o similar al Apartamento N° A-02-01 de la Torre "A" del edificio "Pórtico del Este" (60,00 M2).

Lo ideal sería seleccionar apartamentos de la misma área que el apartamento a valuar.

Sin embargo, utilizar un criterio tan drástico; traería la posibilidad de reducir la muestra u unos poquísimos referenciales o comparables o a ninguno.

Por lo tanto, aquí queda al solo criterio del profesional tasador, definir un rango de área aceptable para seleccionar o rechazar los probables inmuebles referenciales o comparables, que contribuirían a la formación del valor del apartamento objeto.

No se podría definir un criterio pre-establecido para definir los rangos de aceptación o rechazo. Cada caso es diferente y cada criterio será diferente.

Solo la experiencia del profesional tasador y un estudio de las áreas de los probables referenciales o comparables en comparación con el inmueble a valuar, podrían definir un criterio para aceptar o rechazar a los mismos. Ese criterio, será diferente en cada asignación confiada al evaluador.

Siguiendo el Ejemplo de Aplicación:

Área del Apartamento N° A-02-01: 60,00 M2

Rango de Selección: Apartamentos entre 40 y 80 M2 (incluidos)



Por lo tanto, ya existen Dos (2) criterios que cumplir, para seleccionar a los referenciales o comparables, que se utilizarán para determinar el valor del Apartamento N° A-02-01 del Edificio "Pórtico del Este":

- a) Todos los apartamentos ubicados dentro de la "Poligonal de Ubicación" trazada
- b) Todos los apartamentos ubicados dentro de la "Poligonal de Ubicación" trazada, que tengan un área comprendida entre los 40 y 60 M2

3.4 Tercer Criterio: Selección de inmuebles referenciales o comparables de similar edad

Dentro del mismo Ejemplo N° 1, en la "Poligonal de Búsqueda" trazada, se ubicarán los inmuebles referenciales o comparables con una edad igual o similar al Apartamento N° A-02-01 de la Torre "A" del edificio "Pórtico del Este" (5 años).

Al igual que en el punto anterior, aquí queda al solo criterio del profesional tasador, definir un rango de edad aceptable para seleccionar o rechazar los probables inmuebles referenciales o comparables.

Es de hacer notar en este punto, que el profesional tasador deberá tener mucho cuidado en poder identificar correctamente la edad de un inmueble.

Es costumbre definir la edad de un inmueble, como el tiempo transcurrido entre la fecha del avalúo y la fecha de registro del Documento de Condominio.

Sin embargo, esto será válido siempre y cuando la fecha de Protocolización del Documento de Condominio sea muy cercana a la fecha de finalización del edificio; si no fuera así, la edad del edificio estaría definiría por su "Acta de Finalización", "Constancia de Terminación", "Habitabilidad Sanitaria" u otro documento equivalente.

Siguiendo el Ejemplo de Aplicación:

Edad del Apartamento N° A-02-01: 5 años

Rango de Selección: Apartamentos entre 2 y 8 años de construcción



Por lo tanto, ya existen Tres (3) criterios que cumplir, para seleccionar a los referenciales o comparables, que se utilizarán para determinar el valor del Apartamento N° A-02-01 del Edificio "Pórtico del Este":

- 1) Todos los apartamentos ubicados dentro de la "Poligonal de Ubicación" trazada
- 2) Todos los apartamentos ubicados dentro de la "Poligonal de Ubicación" trazada, que tengan un área comprendida entre los 40 y 60 M2
- 3) Todos los apartamentos ubicados dentro de la "Poligonal de Ubicación" trazada, que tengan un área comprendida entre los 40 y 60 M2 y que tengan entre 2 y 8 años de construcción

Utilizando estos Tres (3) criterios de búsqueda precitados, es probable que se hayan filtrado el 65% de los probables referenciales o comparables, es decir de 100 operaciones de compra-venta seleccionadas, se han rechazado 65.

Esto se debe a que son "Criterios Anidados", equivalentes al procesamiento de Tres (3) operadores lógicos tipo ".AND." e implican altos niveles de rechazo. Por ende, mientras más criterios se analicen, la cantidad de referenciales o comparables resultantes del proceso será exponencialmente menor. El profesional tasador, deberá utilizar estas herramientas de búsqueda de referenciales o comparables, de una manera racional.

3.4 Filtraje de la información

En teoría, la aplicación de la técnica de "Criterios de Búsqueda" debería arrojar una muestra de referenciales o comparables suficientes y significativos para poder determinar el valor de un inmueble siguiendo los procedimientos sugeridos en la Metodología de "Comparación Directa" o "Aproximación al Mercado".

Sin embargo, eso no es siempre así. Existe en toda data, información que no es posible de filtrar aplicando los criterios de búsqueda antes citados. Información, que al no está referida a la ubicación, área, edad, estado de conservación, etc.; las técnicas de criterios anidados no la detectan.



Por lo tanto se deben prever "Filtros Adicionales". Estos filtros, serán aplicados a los referenciales o comparables resultantes de la aplicación de los criterios de búsqueda antes descritos.

Muchos "Filtros" podrían ser definidos por el profesional tasador, de acuerdo al criterio y tipo de avalúo. Aquí solo se nombrarán unos cuantos a fin de que sirvan de ejemplo.

a) Operaciones de Compra – Venta entre familiares se rechazan:

Este filtro tiene relación con la propia definición del Principio de Competencia del Mercado Inmobiliario: "Inmuebles similares se venderán a precios similares".

Si bien una operación de venta entre familiares es perfectamente legal. Se podría suponer que no es realmente un genuino contrato de compra - venta, sino mas bien, otro tipo de operación tal como una "herencia en vida" de un padre a un hijo o el enmascaramiento de los bienes de una persona frente a un juicio de embargo o cualquier otra razón de índole personal o familiar, que no tienen nada que ver con una real operación de compra – venta, donde no hay un traspaso formal de propiedad ni un intercambio real de dinero.

b) Operaciones de Compra – Venta entre empresas filiales o subsidiarias se rechazan:

Este filtro tiene relación con el anterior, pero no interviene personas naturales, sino jurídicas.

Si bien una operación de venta de un inmueble entre una empresa filial ("Holding") y una de sus empresas subsidiarias; o entre dos empresas subsidiarias de una misma filial es perfectamente legal.

Tampoco en este caso se puede suponer que se trate de una operación de compra – venta real. Mas bien, podría identificarse como una "Operación Accionaria", o una "Recapitalización", o cualquier operación financiera lícita, que no está relacionada de ninguna forma con una negociación de "compra – venta", en el sentido que antes hemos definido.

Ahora bien, lo realmente complicado de este tipo de referenciales o comparables; consiste en su identificación.



A veces, es relativamente sencillo deducir este tipo de operación, al analizar las razones sociales de cada una de las empresas. Por ejemplo: Banco Futuro, S.A. vende una oficina a Inversora Futuro, S.A.

Otras veces, es casi imposible identificar si las empresas presentes en una operación de compra – venta inmobiliaria; son filiales o subsidiarias. En estos casos, entra en juego la capacidad del profesional tasador de obtener “información privilegiada” o investigar sobre quiénes son los socios de dichas empresas.

c) Las operaciones de “ventas de derechos” sobre un inmueble se rechazan:

Las ventas de derechos o cuotas sobre bienes inmuebles, no implican el intercambio de ninguna propiedad parcial, más bien lo que se traspa en este tipo de operaciones son los “derechos de participación” sobre la totalidad de un inmueble.

Sin embargo, en las zonas agrícolas y ganaderas de Hispanoamérica, es muy común observar este tipo de operaciones, sustituyendo a los verdaderos contratos de compra – venta de fincas y parcelas.

3.5 Obtención de la data depurada.

Si el profesional tasador, es capaz de cumplir con los procedimientos arriba citados, de una forma racional y honesta; es probable que obtenga una excelente serie de referenciales o comparables, que permitirán como resultado un valor muy preciso en la determinación del “Término Central” de dicha serie, y por lo tanto una altísima exactitud en la obtención del valor del inmueble designado.



4.0 SERIES GRANDES, PEQUEÑAS Y MUY PEQUEÑAS

4.1 Definiciones:

Se define como "Serie Pequeña", a una muestra de datos que se supone normalmente distribuidos, pero de la cual se desconoce su varianza o tendencia a desviarse del valor promedio.

También podría definirse como una colección menor o igual a Veinte (20) datos.

Se define como "Serie Muy Pequeña", a una muestra de datos cuya varianza es desconocida, debido a que no es posible la selección "no aleatoria" de la misma.²

En este Trabajo de Investigación, se clasificarán las series en Tres (3) Grupos en función del número de datos, a cada grupo se le asignará una metodología específica para el cálculo del Término Central de cada una de dichas series:

² BARBOUR, A. (2001): "Simplified estimation of Widmark "r" values by the method of Forrest". Nota Técnica en la Publicación Periódica: Science & Justice 2001 Journal; 41(1): pp. 53–54. Amsterdam. Netherlands.



CANTIDAD DE DATOS DE LA SERIE	METODOLOGIA ESTADISTICA PROPUESTA PARA CALCULAR SU TERMINO CENTRAL
SERIES MAYORES A VEINTE (20 DATOS)	ANALISIS ESTADISTICO DE UNA (1) VARIABLE PARA DATOS AGRUPADOS EN CLASES ⁱ
SERIES ENTRE 10 Y 19 DATOS	TEORIA ESTADISTICA DE ERRORES
SERIES DE NUEVE (9) DATOS O MENOS	ANALISIS DE ESTIMADORES NO PARAMETRICOS



5.0 SERIES MAYORES DE VEINTE (20) DATOS:

5.1 Para el análisis de las Series de Veinte (20) datos o superiores, se utilizará la Metodología para el Análisis Estadísticos de Una (1) Variable, para Datos Agrupados en Clases.

Cuando se dispone de número grande de datos, es útil el distribuirlos en "Clases o Categorías" y determinar el número de individuos pertenecientes a cada Clase, que es lo que se denomina la "Frecuencia de Clase".

Tabla de Distribución de Frecuencia, es una ordenamiento tabular de datos agrupados en Clases y sus correspondientes Frecuencias.

CLASE	LÍMITE INFERIOR L_i	LÍMITE SUPERIOR $L_{(i+1)}$	MARCA DE CLASE X_i	FRECUENCIA ABSOLUTA f_i	FRECUENCIA ACUMULADA F_i
1ra.	150,00	179,15	164,58	5	5
2da.	179,15	208,30	193,73	1	6
3ra.	208,30	237,45	222,88	3	9
4ta.	237,45	266,60	252,03	1	10
5ta.	266,60	295,76	281,18	1	11
			Σ	11	

5.2 Elementos que conforman una Tabla de Distribución de Frecuencias:



a) Límites Superior e Inferior de la Clase.

b) Marca de Clase (Xi): Se define como el punto medio del Intervalo de Clase y se obtiene sumando los Límites Inferior y Superior de la Clase y a este resultado dividirlo entre Dos (2). La "Marca de Clase" (Xi), también se conoce como "Punto Medio de la Clase" (o "Midpoint" en los textos anglosajones).

c) Frecuencia Absoluta (fi): Se define como Frecuencia Absoluta o Frecuencia Ordinaria de una Clase (fi), como el número de datos de la serie en estudio, que aparece ente el Límite Inferior y el Límite Superior de dicha Clase.

d) Frecuencia Acumulada (Fi): Se define como Frecuencia Acumulada de una Clase (Fi), como la sumatoria de las "Frecuencias Absolutas" (fi) de todas las "Clases" inferiores o igual a la "Clase" en estudio. En algunos textos también se conoce como "Frecuencia Acumulada Absoluta".

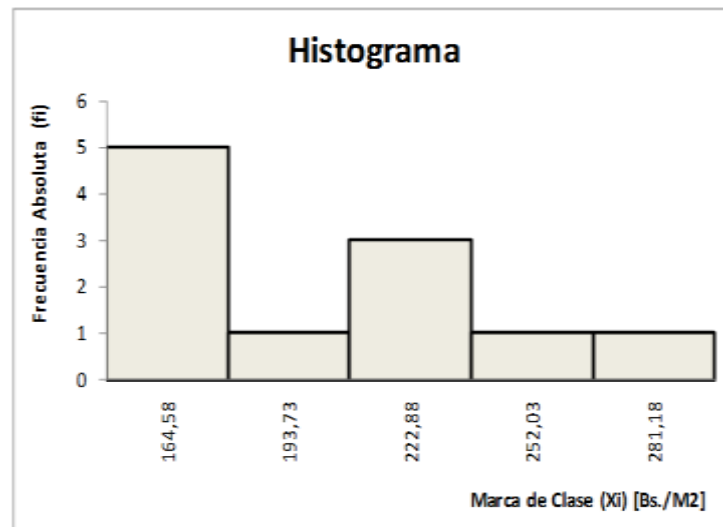
e) Representación Gráfica de la Tabla de Distribución de Frecuencia

e-1) El Histograma: Consiste en una serie de rectángulos o gráfico de barras, que tienen las siguientes características:

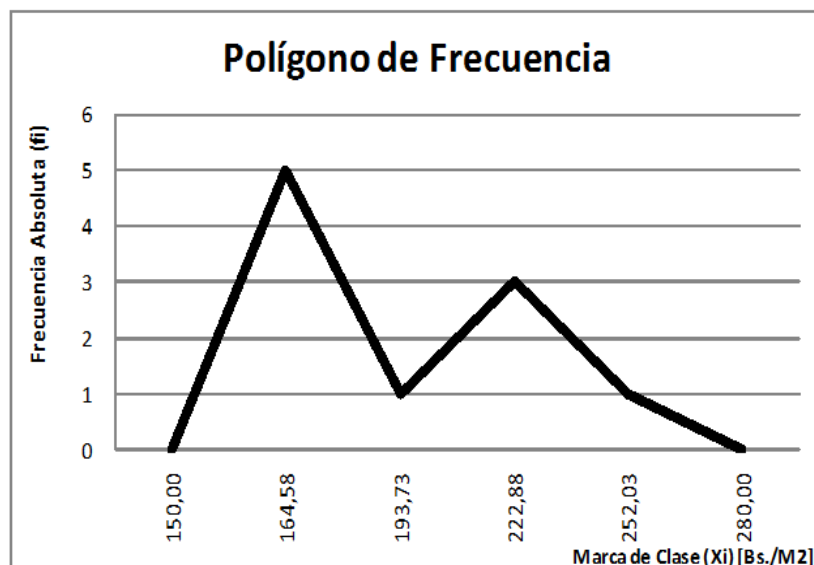
e-1-a) Su base sobre el eje horizontal (eje X) con centros en las "Marcas de Clase" (Xi) y anchos iguales al ancho de los intervalos de clase (C).

e-1-b) Sus superficies son proporcionales a las "Frecuencias Absolutas" de cada clase

e-1-c) Si los intervalos de clase tienen todos iguales tamaños, las alturas de los rectángulos serán proporcionales a las frecuencias de clase y se acostumbra en tal caso a tomar las alturas numéricamente iguales a las frecuencias absolutas (fi) de cada clase.



e-2) El Polígono de Frecuencias: El Polígono de Frecuencia, consiste en un "gráfico de línea", trazado sobre las "Marcas de Clase" (Xi). Se construye, uniendo los puntos medios de los techos de los rectángulos del "Histograma".





f) El punto inicial de un "Polígono de Frecuencias" es el "Límite Inferior de la Serie" y el punto final, el "Límite Superior de la Última Clase".

5.3 Generalidades:

Debido a que el objetivo de este estudio de investigación, está dirigido al análisis de "Series Pequeñas" y "Series Muy Pequeñas", no se profundizará más sobre el tema.

Sin embargo, se recomienda la siguiente bibliografía, donde podrá encontrar teoría y ejemplos resueltos, donde se pondrá a ahondar sobre la determinación del Término Central de una Serie, mediante la Metodología para el Análisis Estadísticos de Una (1) Variable, para Datos Agrupados en Clases:

a) ALCAZAR, M. (2012): *"Manual de Valoración Inmobiliaria"*. Capítulo 9 (Métodos Estadísticos y Económicos). Editorial Delta Publicaciones. Madrid. España.

Este libro de texto, podrá ser adquirido en línea, a través del Website: www.amazon.com

b) Acceso a Internet: www.rpiol.com

Enlace: *Estadística I: Análisis de Una Sola Variable. Determinación de la Tendencia Central.*

Correspondiente al material del Curso Profesional de Valuación Inmobiliaria, dictado por la Sociedad de Ingeniería de Tasación de Venezuela (SOITAVE). Caracas. Venezuela

c) RIVAS, Ernesto (1972): *Estadística General. Ediciones de la Biblioteca. UCV. Caracas. Venezuela*



6.0 SERIES ENTRE 10 Y 19 DATOS

6.1 Para el análisis de las Series entre Diez (10) y Diecinueve (19) Referenciales o Comparables, se utilizará la Metodología para el Análisis Estadísticos denominada "Teoría Estadística de Errores".

6.2 La Teoría Estadística de Errores, como metodología para la determinación del Término Central de una "Serie Pequeña":

La falta de exactitud o precisión en la medición de una serie de datos, se relaciona a los denominados "Errores Accidentales", que se analizarán en este Trabajo de Investigación.

Existen tres tipos de errores: "Los Errores Sistemáticos" (producidos por la mala calibración de un equipo de medición); los "Errores de apreciación" (producidos por erróneas lecturas de las escalas empleadas en cualquier instrumento de medición); los cuales no vienen al caso en este trabajo de investigación; y los "Errores Accidentales" (producidos por causas imponderables, que perturban aleatoriamente una serie de observaciones).

Si bien la finalidad de la "Teoría de Errores" está relacionada con la exactitud del valor numérico de las observaciones de los datos de una serie. No resulta posible determinar los errores absolutos ni relativos de una observación, pero si es realizable el conocer los límites superior e inferior del "Margen de Error".

6.3 Definición de "Errores Accidentales"



Es un hecho, que efectuadas observaciones a varios inmuebles referenciales o comparables, ubicados en la misma zona, con la misma edad, similar área y estado de conservación y mantenimiento; no se llegan a conseguir idénticos valores en dicha muestra. El Precio de Venta de los diferentes referenciales o comparables inspeccionados, estarán entre otras cosas siempre afectadas por perturbaciones que escapan de cualquier hipótesis.

Contrariamente a la hipótesis de que los Inmuebles Sustitutos, existen gracias al principio de "competencia del mercado inmobiliario"; que exige que: Las "propiedades similares" tengan aproximadamente "parecidos valores"; dicha "igualdad de condiciones", generalmente no se observa debido a la alteración en la apreciación humana del concepto de "Valor". Esto, principalmente, da origen a los denominados "Errores Accidentales" en una serie de inmuebles referenciales o comparables.

6.4 El concepto de "Incertidumbre", como estimador del "Error Accidental" de una Serie Pequeña es el siguiente:

Sea una Serie Pequeña " S " de la forma $S \{x_i\}_{i=1}^n$

Donde: " \bar{x} " es el "Promedio Aritmético Simple" para datos no agrupados en clases y su expresión matemática es la siguiente: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

El cálculo de la "Desviación Estándar de la Muestra" de dicha Serie " σ_{n-1} "; se determinará mediante la expresión matemática $\sigma_{n-1} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$



En una observación directa de varios inmuebles referenciales o comparables, se define como "Incertidumbre" o "Error" de una serie.. " $\delta_{\bar{x}}$ " Como "el mejor estimador de la dispersión de la data" y vendrá enunciada por la siguiente expresión: $\delta_{\bar{x}} = \frac{\sigma_{n-1}}{\sqrt{n}}$

La Incertidumbre o Error de una serie. También puede expresarse como: $\delta_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n*(n-1)}}$

Por tanto, en la observación directa de los valores de " n " inmuebles referenciales o comparables; el Término Central de la Serie será el siguiente: $T. C. S = \bar{x} \pm \delta_{\bar{x}}$

6.5 El rechazo o aceptación de observaciones en una serie mediana:

Para un mejor conocimiento de una medida de la tendencia central, es necesaria otra clase de estimador, que exprese el grado de representatividad de la primera, una de ellas se refieren al agrupamiento de los datos en torno a la tendencia central. Estas nuevas medidas se denominan "Estadísticos de Dispersión" o "Dispersores.

En el caso de la "Teoría Estadística de Errores", se han estudiado el dispersor "Desviación Estándar de la Muestra": σ_{n-1} y el dispersor "Incertidumbre de la Data": $\delta_{\bar{x}}$

Si \bar{x} se define como la mejor estimación posible del "Término Central de la Serie". El dispersor "Incertidumbre o Error de la Muestra". Por lo tanto, el estadígrafo $\delta_{\bar{x}}$; será el indicador para determinar si algunas de las observaciones de la serie son o no rechazables de acuerdo a los siguientes criterios:



CONDICION	RESULTADO
Si $ \bar{x} \pm x_i \leq 3\delta_{\bar{x}}$	Se acepta el Referencial o Comparable
Si $ \bar{x} \pm x_i > 3\delta_{\bar{x}}$	Se rechaza el Referencial o Comparable

Por lo tanto, si en una serie de Referenciales o Comparables, se observan que uno o más elementos cumplen con el criterio: $|x_i| > 3\delta_{\bar{x}}$; los mismos se considerarán como "Errores Accidentales" y deberán ser eliminados de la serie. Se procederá entonces, a recalcular el "Término Central de la Nueva Serie" $\bar{x}_{mod.}$ con el resto de los Referenciales o Comparables observados.

El dispersor "Incertidumbre" $\delta_{\bar{x}} = \frac{\sigma_{n-1}}{\sqrt{n}}$, también llamado "Error Cuadrático Medio", permite asegurar que el "Valor Verdadero" del "Término Central de la Serie", tenga una probabilidad de estar comprendido en:

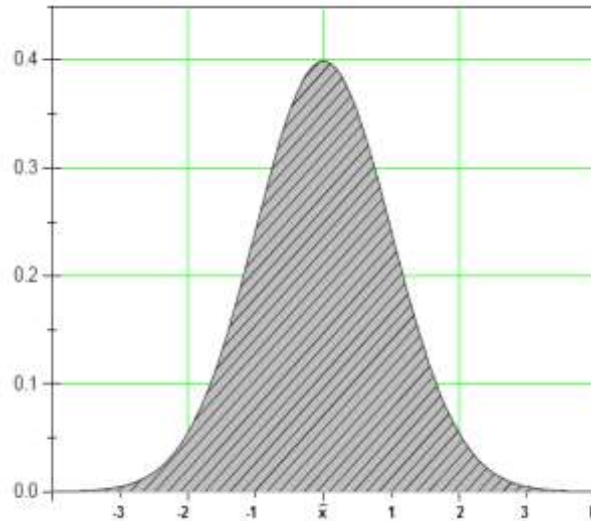
PROBABILIDAD	INTERVALO
68.27 %	Entre $\bar{x} - \delta$ y $\bar{x} + \delta$
94.45 %	Entre $\bar{x} - 2\delta$ y $\bar{x} + 2\delta$
99.70 %	Entre $\bar{x} - 3\delta$ y $\bar{x} + 3\delta$

Por tanto la expresión $|\bar{x} \pm k\delta|$ permitirá obtener el ancho del intervalo donde el "Término Central de la Serie" existe o sea la distribución de valores de una magnitud, medida en torno al "Valor Real".

6.6 El Diagrama de Dispersión:



Si se cumple que: $x_i \in |\bar{x} \pm k\delta|$ (para $k = 1, 2, 3, \dots$), se genera un "Diagrama de Dispersión" especialísimo, que se denominada "Distribución Gaussiana".



Toda "distribución Gaussiana", tiene la propiedad particular, de que el valor de una magnitud que se encuentre entre dos valores x_i y x_{i+1} , será igual al área bajo la curva entre dichos valores citados.

Esta distribución tiende asintóticamente tiende a cero en el infinito en ambas direcciones; por lo tanto, cualquier valor sería posible (aunque no probable).

Sin embargo, no por muy grande en que se establezca el ancho del intervalo de confianza, se tendrá una seguridad absoluta, a que el valor real de la magnitud estará dentro del mismo.

VER UN EJEMPLO NUMERICO RESUELTO DEL METODO EN EL APENDICE 1



7.0 SERIES MUY PEQUEÑAS

7.1 Definición

Se define como "Serie Muy Pequeña", a una muestra de datos cuya varianza es desconocida, debido a que no es posible una selección "no aleatoria" de la misma.³

Esta selección "No Aleatoria" de la muestra, se debe a que se estudia una cantidad muy limitada de datos, que en buena proporción nada tienen que ver con el estudio. La mayoría de la veces, esa muy pequeña muestra, resulta la única información disponible para estudiar un fenómeno.

En las grandes ciudades de Hispanoamérica, es posible seleccionar una suficiente cantidad de "Inmuebles Comparables" (muestras), necesarios para poder determinar su "Término Central" mediante el análisis estadístico preciso de dicha data, conforme a los requerimientos contemplados en la Metodología Valuadora de "Comparación Directa" o también conocida como "Aproximación al Mercado".

Sin embargo, no sucede lo mismo en las pequeñas ciudades y pueblos de la provincia, donde se formalizan muy pocas operaciones de Compra – Venta de inmuebles, haciendo imposible o casi imposible un estudio estadístico significativo con una pequeña muestra de inmuebles referenciales o comparables.

³ BARBOUR, A (2001): Technical Note: Simplified estimation of Widmark "r" values by the method of Forrest. Science & Justice Journal; 41(1): pp. 53–54. Amsterdam. Reinos Bajos



7.2 Justificación para el uso de Métodos "No Paramétricos" o "Robustos" en Series Muy Pequeñas:
La necesidad de la aplicación de análisis estadísticos a series muy pequeñas de datos, ha originado la búsqueda de soluciones no tradicionales.

Los Métodos "No Paramétricos" o "Robustos", ofrecen una alternativa a los modelos "paramétricos clásicos" o tradicionales, muy especialmente en el caso del análisis de datos sin ninguna distribución específica.

La "Estadística No Paramétrica" o "Robusta", es la rama de la estadística, que pretende estudiar el comportamiento de una serie, cuando exista la posibilidad de que su data esté contaminada por "outliers" (valores atípicos o errores groseros o errores espurios) que influyeran los resultados del análisis; a fin de poder garantizar una estimación más fiable del comportamiento del fenómeno en estudio

7.3 Definición de Valores Atípicos:

Se definen como "Valores Atípicos" o "Errores Groseros" o "Errores Espurios" o "Outliers", como aquellas observaciones; que aún desviándose marcadamente del centro de la data, constituyen parte de la serie estudiada.

Ahora bien; si bien es cierto que los "Valores Atípicos" constituyen parte de la serie estudiada; también es cierto que, la presencia de estos en dicha serie, alterarán enormemente el resultado del análisis estadístico.

En otras palabras: Por un lado, la inclusión de "Valores Atípicos" en un cálculo estadístico afectaría considerablemente los resultados; y por el otro, la eliminación de los mismos, sería equivalente a cerrar los ojos frente a una realidad incómoda.

Si bien es cierto que los Estimadores "Robustos" o "No Paramétricos"; son menos sensibles a los efectos de los "Valores Atípicos" en una serie de datos, que los "Estimadores Clásicos"; su gran ventaja reside en permitir emplear la totalidad de la data observada en un estudio estadístico.

Sin embargo, los Estimadores "Robustos" o "No Paramétricos", son especialmente útiles cuando la distribución de la data es desconocida, presumiéndose que dicha distribución ni es normal ni transformable a normal.



Además, el disponer de solamente unas pocas observaciones de un fenómeno; implica que los contrastes sobre normalidad serán de precaria validez. Precisamente aquí, los Estimadores "Robustos" o "No Paramétricos" se convierten en una herramienta esencial de análisis.

7.4 Tipos de estimadores "No Paramétricos" o "Robustos":

7.4.1. Estimadores tipo "MVT" (Estimador de Recorte Multivariante) (M-Estimadores): Se basan en procesos iterativos, fundamentados en la eliminación de valores atípicos.

7.4.2 Estimadores tipo "MLT" (Estimador de Recorte de Máxima Verosimilitud) (Maximun Likelihood Test) (L-Estimadores): Se basan en procesos, que en vez de eliminar valores atípicos, establecen "pesos" que permiten aprovechar toda la data observada.

7.4.3. Sobre los Estimadores MVT – MLT

Uno de los principales inconvenientes de los Estimadores "No Paramétricos" o "Robustos", es lo rápidamente que se alteran los estadísticos de control al aplicarlos de forma iterativa.

Principalmente, resultan alterados los estadísticos: "Promedio Aritmético" " \bar{x} ", la "Desviación Estándar de la Muestra" " σ_{n-1} " y la "Incertidumbre o Error de la muestra" $\delta_{\bar{x}}$.

Por lo tanto, se deberá tener cuidado en aplicar la metodología, a fin de minimizar este detrimento propio de la metodología.

7.5 El Estimador de Huber:

7.5.1 Definición: La teoría de la Estimación Robusta de la Tendencia Central de una serie, ideada por Peter J. Huber a principio de los años 60's; se basa en el cálculo del Término Central de una serie, que se presume normalmente distribuida, pero presenta uno o más valores atípicos o outliers (serie contaminada).

Peter J. Huber estudió y desarrolló los estimadores que llevan su nombre, para el análisis de dichas series, para lo menos en Dos (2) escenarios:

- a) En el caso de la eliminación de "Valores Atípico": Se conoce como: "Estimador MVT" o "Estimador de Recorte Multivariante" o también "Estimador-M" (M-Estimator, en inglés);



el cual se basa en un proceso iterativo, fundamentado en la eliminación o minimización de los valores atípicos de la serie.

b) En el caso del establecimiento de "pesos", que permitan utilizar toda la data observada sin eliminar los "Valores Atípicos": Se conoce como "Estimador MLT" o también "Estimador de Recorte de Máxima Verosimilitud" (Maximun Likelihood Estimator, en inglés) o "Estimador-L" (L-Estimator, en inglés).

7.6 El Estimador- M de Huber:

7.6.1 El Estimador-M de Huber: Hipótesis

La hipótesis para el desarrollo de este M-Estimador, consiste en definir una "Función de Residuos" $f_{(x-\bar{x})}$ distinta a la propuesta por el método de los mínimos cuadrados⁴

Por lo tanto, se elegirá una "Función de Residuos", de forma adecuada; que deberá cumplir que:

$$\sum f_{(x-\bar{x})} \rightarrow \text{sea mínima}$$

7.6.2 El Estimador-M de Huber: Procedimiento de Cálculo

La Norma Internacional ISO 13528, describe el procedimiento de cálculo de este "Estimador No Paramétrico" o "Estimador Robusto" del tipo "MVT"; creado por el Prof. Peter J. Huber en el año 1964, con el objetivo de obtener un estimador de la tendencia central en una pequeña muestra afectada por "valores extremos" u "outliers".

a) Sea una pequeña muestra de datos ordenada en orden creciente:

⁴ La "Teoría de los Mínimos Cuadrados", expresa que si un número es pequeño; al elevarlo al cuadrado será aún más pequeño. Por lo tanto, para que se cumpla la premisa citada en el texto, la "Función de Residuos" $f((x - \bar{x}))$ deberá ser diferente a la función $\sum (x - \bar{x})^2 = 0$



$$x_i \in \{x_1, x_2, x_3, \dots, x_n\}$$

b) Se denominará al estadígrafo \bar{x}^* , como "Promedio Robusto"

c) Se denominará al estadígrafo σ^* , como "Desviación Estándar Robusta"

d) Inicialmente, se le asigna al "Promedio Robusto", el valor de la Mediana para Datos No Agrupados en Clases":

$$\text{Valor Inicial de } \bar{x}^* = M_e = \text{Mediana}(x_1, x_2, x_3, \dots, x_n)$$

e) Se le asigna inicialmente a la "Desviación Estándar Robusta", el valor de la Desviación Mediana Basada en MAD (SMAD) para Datos No Agrupados en Clases":

$$\text{Valor Inicial de } \sigma^* = \text{SMAD}$$

$$\sigma^* = \text{SMAD} = 1,483 * \sum \frac{|x_i - M_e|}{n}$$

Donde:

$$\text{Desviación Mediana Absoluta} = \text{MAD} = D_{Me} = \sum \frac{|x_i - M_e|}{n}$$

$$\text{Desviación Mediana Estandarizada} = \text{SMAD} = 1,483 * \text{MAD}$$

f) Se calcula el Intervalo Máximo de Error, que tiene la forma: $\bar{x}^* \pm \varepsilon$ donde: $\varepsilon = 1,5 \sigma^*$

Quedando como: El Límite Superior del Intervalo Máximo de Error será $\bar{x}^* + \varepsilon$
El Límite Inferior del Intervalo Máximo de Error será $\bar{x}^* - \varepsilon$



g) Se sustituirán todos los datos x_i mayores que $\bar{x}^* + \varepsilon$, por el Límite Superior del Intervalo Máximo de Error $x_i = \bar{x}^* + \varepsilon$

$$\text{Sea } x_i \in \{x_1, x_2, x_3, \dots, x_n\}$$
$$\text{Si } x_i \geq \bar{x}^* + \varepsilon \quad \text{entonces } x_i = \bar{x}^* + \varepsilon$$

h) Se sustituirán todos los datos menores que $x_i = \bar{x}^* - \varepsilon$ por el Límite Inferior del Intervalo Máximo de Error $x_i = \bar{x}^* - \varepsilon$

$$\text{Sea } x_i \in \{x_1, x_2, x_3, \dots, x_n\}$$
$$\text{Si } x_i \leq \bar{x}^* - \varepsilon \quad \text{entonces } x_i = \bar{x}^* - \varepsilon$$

i) Teniendo ahora una nueva serie; se procederá a calcular con la nueva data los siguientes estadígrafos:

a) El Promedio Aritmético para Datos No Agrupados de la Nueva Serie:

$$\bar{x}^* = \frac{\sum_{i=1}^n x_i}{n} \quad ; \quad \text{el cual se designará como el nuevo "Promedio Robusto"}$$

b) La Desviación Estándar Robusta de la Nueva Serie: $\sigma^* = 1,134 * \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$
; la cual se designará como la nueva "Desviación Estándar Robusta"

j) Repetir o iterar los pasos, sustituyendo en cada iteración los datos originales, hasta lograr la convergencia de los valores de los estadígrafos: "media robusta" y "desviación estándar robusta" a valores estables.

k) La convergencia se asumirá, cuando se observe solo un pequeño cambio, entre los valores de los estadígrafos calculados en la última iteración realizada, con respecto a la anterior.

**VER UN EJEMPLO NUMERICO RESUELTO DEL CÁLCULO DEL ESTIMADOR-M DE HUBER EN EL
APENDICE 2**



7.7 El Estimador-L de Huber.

7.7.1 El Estimador-L de Huber: Hipótesis:

El Estimador-L de Huber, fue difundido por el Profesor PETER J. HUBER. En su libro "Robust Estimation"⁵ del año 1968. En él presenta en detalle la novedosa teoría de la estimación robusta de un parámetro de tendencia central en una distribución normal contaminada por valores atípicos u outliers, sin eliminarlos.

El Estimador-M de Huber, se basa en minimizar una función de residuos $\varphi(x_i - \bar{x})$, distinta a la propuesta por el método de los mínimos cuadrados $\sum(x_i - \bar{x})^2 \equiv 0$, de manera que se cumpla las siguientes condiciones:

$$\varphi(x_i - \bar{x}) = \begin{cases} (x_i - \bar{x})^2 & \text{si } |x_i - \bar{x}| \leq 2\sigma \\ 2\sigma * [2 * (x_i - \bar{x}) - 2\sigma] & \text{si } |x_i - \bar{x}| > 2\sigma \end{cases}$$

La iteración comienza suponiendo que todas las observaciones tienen el mismo peso. En la siguiente iteración, se calculan nuevos pesos para cada observación individual a partir de los residuos obtenidos en el ajuste anterior, y se repite el ajuste con los nuevos pesos. El proceso iterativo se continúa hasta que se logra la convergencia de $(\bar{x}_{i+1} - \bar{x}_i) \rightarrow 0$.

Para el cálculo de los Pesos, el Estimador-M de Huber utiliza las siguientes funciones:

$$\begin{cases} P = 1 & \text{si } |x_i - \bar{x}| \leq K * \delta_{\bar{x}} \\ P = \frac{K * \delta_{\bar{x}}}{|x_i - \bar{x}|} & \text{si } |x_i - \bar{x}| > K * \delta_{\bar{x}} \end{cases}$$

⁵ Huber, P. J. 1968: Robust Statistics, Wiley, New York.



NOTA 1: Ordinariamente, K adopta los valores de 2 ó 3.

NOTA 2: Obsérvese que la expresión $K * \delta_x$, representa el producto de la tolerancia de la medición (K) y la exactitud (δ_x) de la observación.

7.7.2 El Estimador-L de Huber: Procedimiento de Cálculo:

El procedimiento para el cálculo del Estimador-L de Huber, se puede resumir de la siguiente manera:

1) Sea una pequeña muestra de datos ordenada en orden creciente:

$$x_i \in \{x_1, x_2, x_3, \dots, x_n\}$$

2) Primera Iteración: Suponiendo que el Peso Inicial para cada uno de los datos es 1; se calcula el estadígrafo $\bar{x}_{inicial}$, como el "Promedio Simple para Datos No Agrupados" de la data original sin modificar.

3) Se procede al cálculo de la Desviación estándar de la Población σ_{n-1} y del Error o Incertidumbre de la muestra original $\delta_{\bar{x}}$, a través de las siguientes expresiones:



$$\delta_{\bar{x}} = \frac{\sigma_{n-1}}{\sqrt{n}}$$

$$\sigma_{n-1} = \sqrt{\sum \frac{(x_i - \bar{x})^2}{(n-1)}}$$

$$\delta_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n * (n-1)}}$$

4) Se Calcula los pesos de acuerdo a las expresiones:

$$\left\{ \begin{array}{l} P = 1 \text{ si } |x_i - \bar{x}| \leq K * \delta_{\bar{x}} \\ P = \frac{K * \delta_{\bar{x}}}{|x_i - \bar{x}|} \text{ si } |x_i - \bar{x}| > K * \delta_{\bar{x}} \end{array} \right.$$

5) Se le aplica a la data original, su ajuste respectivo mediante los pesos calculados en el paso anterior.

6) Se calcula el "Promedio Ponderado", de la serie ajustada, a través de la expresión:



$$\bar{x} = \frac{\sum(P_i * x_i)}{\sum P_i}$$

7) Se verifica la convergencia de los Promedios $(\bar{x}_i - \bar{x}_{inicial}) \rightarrow 0$. Si hay convergencia, \bar{x}_i será el Término Central de la Serie. Si no la hubiere, se procederá con una nueva iteración.

8) Nueva Iteración: Se calculan los nuevos Residuos $(x_i - \bar{x})$, desviación estándar de población σ_{n-1} y Error o Incertidumbre δ , en base al último promedio aritmético \bar{x}_i , tal como se hizo en el paso 3).

9) Se Recalculan los pesos de acuerdo a las expresiones arriba citadas.

10) Se le aplica a la data original, su nuevo ajuste respectivo mediante los pesos recalculados en el paso anterior.

11) Se vuelve calcular el nuevo "Promedio Ponderado", de la serie ajustada en esta iteración.

12) Se verifica la convergencia de los Promedios $(\bar{x}_{i+1} - \bar{x}_i) \rightarrow 0$. Si hay convergencia, \bar{x}_{i+1} será el Término Central de la Serie. Si no la hubiere, se procederá con una nueva iteración.

13) Repetir o iterar los pasos anteriores, hasta que la diferencia de los promedios sea mínima.

**VER UN EJEMPLO NUMERICO RESUELTO DEL CÁLCULO DEL ESTIMADOR-L DE HUBER EN EL
APENDICE 3**



8.0 CONCLUSIONES

8.1 Series Pequeñas

Mientras más pequeña sea una serie, mas se destacarán las discrepancias que aparecen en el levantamiento de la data. Cuando no es posible obtener un número de observaciones razonables, la precisión de los análisis estadísticos convencionales cae notablemente y es necesario la búsqueda de un "método alternativo".

Si bien la Teoría Estadística de Errores, define tres tipos de errores: "Los Errores Sistemáticos", los "Errores de Apreciación" y los "Errores Accidentales". Los dos primeros son inaplicables en el levantamiento de los referenciales o comparables; sin embargo el tercero, si podría asociarse al dinamismo del mercado inmobiliario, a la inflación desproporcionada, al cambio de la moneda y a otros factores fuera del ámbito de competencia del profesional tasador tales como: la insinceridad del contenido de los documentos de compra-venta registrados así como la promulgación de leyes y regulaciones arbitrarias; que son perturbaciones totalmente impredecibles. Validando así a la Teoría Estadística de Errores como "metodología alterna" para el cálculo del término central de una serie pequeña.

En una observación directa de varios inmuebles referenciales o comparables, al mejor estimador de la dispersión de la data, se le definirá como "Incertidumbre" o "Error de la serie" o "Error Cuadrático Medio", y vendrá a ser representado por la expresión:

$$\delta_{\bar{x}} = \frac{\sigma_{n-1}}{\sqrt{n}}$$



Este estadígrafo será definido como el criterio para determinar si algunas de las observaciones de la serie son o no rechazables (valores atípicos u outliers):

CONDICION	RESULTADO
Si $ \bar{x} \pm x_i \leq 3\delta_{\bar{x}}$	Se acepta el Referencial o Comparable
Si $ \bar{x} \pm x_i > 3\delta_{\bar{x}}$	Se rechaza el Referencial o Comparable

8.2 Series Muy Pequeñas

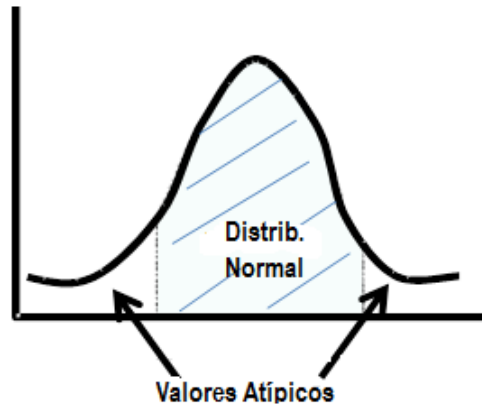
Cuando el número de observaciones es pequeño o muy pequeño, no es posible asumir una distribución normal para la población de donde se extrae la muestra; por consiguiente, el promedio aritmético como estimador del término central no debe utilizarse.

La estimación robusta o no paramétrica, nació de la necesidad de calcular el valor central de una serie a fin de que arroje un resultado preciso, aún cuando se sospeche que la data pueda estar "contaminada" por valores atípicos, que influyan en los resultados y conduzcan a una estimación errónea.

Los estimadores robustos o no paramétricos, evitan abusar del empleo de la media aritmética en toda serie, sea grande o pequeña, simétrica o no, homogénea o dispersa sin el menor análisis previo de la data.

También, evita el rechazo arbitrario de observaciones, alejadas del término central; sin verificar si realmente son o no valores atípicos.

Tradicionalmente el método utilizado para la detección de los valores atípicos, consiste en rechazar la data que no se ajuste a una distribución normal.



La problemática de esta técnica, reside en el hecho de que sólo se podrá reconocer los valores atípicos después del ajuste por mínimos cuadrados de la serie. Ocasionando que los residuos $|x_i - \bar{x}|$ sean encubiertos y distribuidos entre toda la data, haciendo muy difícil distinguir a un valor atípico de otro que no lo sea. Ni siquiera la presencia de residuos grandes, indicarían necesariamente una identificación correcta de los valores atípicos u outliers.

Los estimadores robustos o no paramétricos, reducen la influencia de los valores atípicos sobre el cálculo del término central, sin modificar en absoluto su información estadística, ya que no los elimina, sino más bien los pondera.

Sin embargo, el uso de estimadores robustos o no paramétrico, ha tenido muy poca difusión tanto académica como bibliográficamente. Por lo tanto, es imprescindible su enseñanza y aplicación, ya que es una herramienta indispensable en el estudio estadístico de series pequeñas y muy pequeñas, las cuales son muy frecuentes en los trabajos valuatorios.

En este trabajo de investigación, solo se desarrollaron el Estimador-M de Huber y el Estimador-L de Huber; por ser estos los mayormente estudiados y utilizados en el análisis exploratorio de datos (AED).

Ing. A. Roberto Piol Puppio / Mayo 2014



BIBLIOGRAFIA

ALCAZAR, M. (2012): «Manual de Valoración Inmobiliaria». Capítulo 9 (Métodos Estadísticos y Econométricos). Editorial: Delta Publicaciones. Madrid. España

ANSCOMBE, F. J. (1960): «Rejection of outliers», Publicación Periódica: Technometrics, vol. 2, 123-147. Acceso a Internet: www.tandfonline.com. Reino Unido.

BAIRD, D.C. (1962): «An Introduction to Measurement Theory and Experiment Design». Prentice-Hall, Inc. Primera edición. New Jersey. EE.UU.

BALAKRISHNAN, N. (2006); Permanents, Order Statistics, Outliers and Robustness. Tesis. McMaster University. Acceso a Internet: <http://library.mcmaster.ca/find/theses>. Ontario. Canada

BEVINGTON P., ROBINSON D.K. (1993): «Data reduction and error analysis for the physical sciences», 2nd edición; McGraw Hill, New York. EE.UU

BEERS, Y. (1962): «Theory of Error». Addison-Wesley Publishing Company, Inc. 2da. Edición. Londres. Reino Unido.

CAJAL, E.; PALMER, A. (2012): «When the mean fails, use an M-estimator». Publicación periódica: Anales de Psicología. Vol. 28, nº 1 (enero), 281-288. Servicio de Publicaciones de la Universidad de Murcia. Acceso a Internet: <http://revistas.um.es/analesps>. Murcia. España

DIAS, J.; PIMENTEL, M.; NORONHA, A.; PEÑA, T.; BROGUEIRA, P. (2001): «Introducción a la física». Editorial McGraw-Hill. 2001. 2da. Edición. México



CROSBIE, J. (1993): "Interrupted time-series análisis with brief single-subject data". Publicación Periódica: Journal of Consulting and Clinical Psychology, 61, 966-974. Acceso a Internet: <http://www.apa.org/pubs/journals/ccp/>. Washington. EE.UU.

DOMINGO, A. (2001): «Investigación sobre los Métodos de Estimación Robusta aplicados en los problemas fundamentales de la Fotogrametría». Publicación Periódica: Topografía y Cartografía del Colegio Oficial de Ingenieros Técnicos en Topografía (núm. 100, 103, 104, 105). España

HAMPEL, F. R. (1968): «Contributions to the theory of robust estimation», Ph. D. tesis, University of California, Berkeley. Acceso a Internet: <http://lib.berkeley.edu/find/types/dissertations.html>. EE.UU

HILL, R. W. (1977): «Robust regression when there are outliers in the carriers», Unpublished Ph. D. dissertation, Harvard University, Dept. of Statistics. Acceso a Internet: <http://www.math.harvard.edu/theses/index.html>. EE.UU

HUBER, P. J. (1964): «Robust estimation of a location parameter», Ann. Math. Statist. 35, 73-101. Wiley Online Library. Acceso a Internet: www.wiley.com. New York. EE.UU.

IZARRA, I; HERNANDEZ, J.N.; CARRERO, R: «Aplicación de la Estimación No Paramétrica para una variable en el tratamiento estadístico de la muestra de referenciales». Tesis. Sociedad de Ingeniería de Tasación de Venezuela (SOITAVE). Biblioteca. Caracas. Venezuela

MENDEZ, A. (2001): «Estimación robusta: una aplicación informática con fines didácticos». Publicación periódica: ESTADÍSTICA ESPAÑOLA, Vol. 43, Núm. 147, págs. 105 a 123. Universidad Politécnica de Madrid. España

MOLES, M; COLOM, A; MURGUI, S: «Valores Anómalos en Poblaciones Finitas». Tesis. Universidad de Valencia. Acceso a Internet: www.uv.es. Valencia. España

MUÑOZ-GARCIA, J; MORENO-REBOLLO, J. L.; PASCUAL-ACOSTA, A. (1990): «Outliers: A formal approach», Publicación Periódica: International Statistical Review, vol. 58, 215-226. Wiley Online Library. Acceso a Internet: www.wiley.com. New York. EE.UU.

PALMER, A. (1999). Análisis de datos. Etapa exploratoria. Editorial Pirámide. Madrid. España

PEÑA, D. (1991): «Estadística. Modelos y Métodos», vol. I, 2da. Ed. Madrid. España.



PIOL R.; «Métodos Estadísticos aplicados a la Valuación Inmobiliaria». Material de clase de la Sociedad de Ingeniería de Tasación de Venezuela (SOITAVE). Acceso a Internet: www.rpiol.com. Caracas. Venezuela

ROUSSEUW, P.; LEROY, A. M. (1987): «Robust regression and outlier detection», tesis, Vrije Universiteit Utrecht. Acceso a Internet: <http://stevenammerlaan.nl/vrije-universiteit/> Bruselas, Bélgica.

ZAMAR, R. (1994); «Estimación Robusta», Publicación periódica: Estadística Española; 36, 327–387. Acceso a Internet: <http://dialnet.unirioja.es>. La Rioja. España



APENDICE 1

EJEMPLO NUMERICO RESUELTO: La Teoría Estadística de Errores, como metodología para la determinación del Término Central de una "Serie Pequeña"

Sea la siguiente muestra, una serie de Inmuebles Referenciales o Comparables, correspondiente a apartamentos en la Urbanización Caribe. Parroquia Caraballeda del Estado Vargas (Venezuela), tomados en la Oficina de Registro Inmobiliario del Primer Circuito del Estado Vargas.

MASTER UNIVERSITARIO INTERNACIONAL EN: "CATASTRO MULTIPROPÓSITO AVANZADO" Y "CATASTRO Y VALORACIÓN"



ESTUDIO DEL MERCADO (REFERENCIALES)

APARTAMENTOS EN LA URBANIZACION CARIBE. PARROQUIA CARABALLEDA. ESTADO VARGAS

Identificac.	Urbanización	Dirección del Inmueble	Vendedor	Comprador	Precio de Venta Bs.	Bs. XMts2	Registro	Fecha	Año de Construcción	Area de Construcción
1	CARIBE	CALLE LOS APAMATES RESD.AVILAMAR PH. APTO.A	MARIA PEREZ	FABIOLA FRETTES	1.150.000,00	22.115,38	V1 TOMO 17 NRO.17	25-jun-13	1.994	55,20
2	CARIBE	AV. BOULEV. NAIGUATA RES. SOL Y MAR EDF. B P. 1 No. 9-B	ESPERANZA HIDALGO	ALOIS ROJAS	1.149.000,00	21.199,26	V1 TOMO 07 NRO.21	25-jun-13	1.994	54,20
3	CARIBE	CALLE LEONOR DE CACERES RES. CAROLINA MAR P. 3 No. 32	MONICA PEREZ	DARLING HERNANDEZ	1.151.600,00	20.436,56	V1 TOMO 11 NRO.13	25-jun-13	1.994	56,35
4	CARIBE	AV.LA PLAYA EDIF.REMANSO 2 P/5 NRO.52	WENDY ANGARITA	DAMARIS ALVAREZ	980.000,00	17.500,00	V1 TOMO 17 NRO.12	28-jun-13	1.994	59,00
5	CARIBE	AV.PRINCIPAL RESD.RUBI MAR 1 PH.NRO.5B	LUCIANO ROLO	ANABEL LOPEZ	1.300.000,00	21.666,67	V1 TOMO 5 NRO.16	20-jun-13	1.994	58,75
6	CARIBE	AV.LA PLAYA EDIF.PLAYA AMAR P0/5 NRO.5B	JOSE SOUKI	MIGUEL ALVAREZ	1.600.000,00	28.725,31	V1 TOMO 1 NRO.22	24-jun-13	1.995	55,70
7	CARIBE	AV.PRINCIPAL RESD.CARABALLEDA HUMBOLDT EDIF.F P/1 NºF11	ADRIANA MONTERO	YOLMAR ZURITA	1.000.000,00	17.241,38	V1 TOMO 18 NRO.35	25-jun-13	1.994	58,00
8	CARIBE	AV.PRINCIPAL RESD.CARABALLEDA CARIBE TORRE A P/4 NRO.41	MARIA MARTINEZ	BASILIO TUTUNJI	1.100.000,00	18.867,92	V1 TOMO 10 NRO.16	23-jun-13	1.994	58,30
9	CARIBE	AV.PRINCIPAL RESD.CARABALLEDA HUMBOLDT EDIF.I P/3 NRO.134	JOSE HERNANDEZ	ANA RENDEL	700.000,00	11.695,91	V1 TOMO 1 NRO.38	17-jun-13	1.994	59,85
10	CARIBE	AV.GUAIQUERI RESD.LA CARACOLA P/8 NRO.844	ELIGIA VACCARINO	SANDRA VALERO	1.300.000,00	22.413,79	V1 TOMO 3 NRO.27	29-jun-13	1.994	58,00
11	CARIBE	AV.LA PLAYA RESD.PQUE AZUL P/5 NRO.5D	GIOLY RANGEL	DANIELA BETANCOURT	1.100.000,00	18.803,42	V1 TOMO 5 NRO.20	26-jun-13	1.994	58,50
12	CARIBE	AV.LA PLAYA RESD.COSTAMAR 4 P/1 NRO.12	FATIMA COSME	FREDDY FERNANDES	1.100.000,00	19.047,62	V1 TOMO 5 NRO.20	25-jun-13	1.994	57,75
13	CARIBE	AV.SUR EDIF.LAGUNA BEACH P/12 Nº 14-G	PROMOTORA Y URIARICA	GLORIMAR VILLA FRANCA	1.100.000,00	18.487,39	V1 TOMO 5 NRO.20	25-jun-13	1.994	59,50
14	CARIBE	AV.PRINCIPAL RESD.COLI NA MAR P/8 NRO.82	JOSE ANGOLEMME	LUIS SISO	1.100.000,00	18.965,52	V1 TOMO 5 NRO.20	25-jun-13	1.994	58,00

Se pide calcular, para la fecha del avalúo; el valor del apartamento 2-B, ubicado en el piso 2 del Edificio "Punta Playa", ubicado en la Avenida Circunvalación de la Urbanización Caribe, Parroquia Caraballeda, Estado Vargas, Venezuela. El área del apartamento es de 57.65 M2, el edificio fue construido en el año 1994 y presenta un buen estado de conservación.

Edificio "Punta Playa": Fachada principal



Ubicación del Edificio "Punta Playa"





En primer lugar, se observa que se trata de una "Serie Mediana" (14 datos). Asimismo, al analizar la homogeneidad las variables; se observa que se trata de un "Análisis de Una Sola Variable" (Precio Unitario). Por lo tanto, se resolverá el ejemplo utilizando la "Teoría de Errores".

Cálculo del Promedio para Datos No Agrupados en clase:

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

$$\bar{x} = \frac{271.748,42}{14}$$

$$\bar{x} = 19.410,60 \text{ Bs./M}^2$$

Cálculo de la Desviación Estándar de la muestra:

$$s_{n-1} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$

Cálculo del valor de la "Incertidumbre" de la serie:

$$s_x = \frac{s_{n-1}}{\sqrt{n}}$$

Cálculo del valor de la "Incertidumbre" de la serie:

$$s_x = \frac{s_{n-1}}{\sqrt{n}}$$

Por lo tanto, el valor de la "Incertidumbre" de la serie:

$$s_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n \cdot (n-1)}}$$



Condición de Rechazo por ser considerado el Referencial o Comparable como un "Error Accidental":

CONDICION	RESULTADO
Si $ \bar{x} \pm x_i \leq 3\delta_{\bar{x}}$	Se acepta el Referencial o Comparable
Si $ \bar{x} \pm x_i > 3\delta_{\bar{x}}$	Se rechaza el Referencial o Comparable

Tabla de Cálculo:

Identificac.	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	Si $ \bar{x} \pm x_i > 3\delta_{\bar{x}}$	
1	22.115,38	2.276,21	5.181.152,98	ACEPTAR	
2	21.199,26	1.360,09	1.849.844,81	ACEPTAR	
3	20.436,56	597,39	358.874,81	ACEPTAR	
4	17.500,00	-2.339,17	5.471.716,29	ACEPTAR	
5	21.666,67	1.827,50	3.339.744,07	ACEPTAR	
6	28.575,25	8.736,08	76.319.093,77	RECHAZAR	Error Accidental !!!
7	17.241,38	-2.997,79	8.748.516,47	ACEPTAR	
8	18.867,92	-971,25	943.317,77	ACEPTAR	
9	12.428,25	-7.410,92	54.921.735,25	RECHAZAR	Error Accidental !!!
10	22.413,79	2.574,62	6.628.684,12	ACEPTAR	
11	18.803,42	-1.035,75	1.072.780,54	ACEPTAR	
12	19.047,62	-791,55	626.552,91	ACEPTAR	
13	18.487,39	-1.351,78	1.827.295,76	ACEPTAR	
14	18.965,52	-873,65	763.269,44	ACEPTAR	
$\sum_{i=1}^n (x_i - \bar{x})^2$			166.050.578,68		
$n * (n - 1)$			182		
$\delta_{\bar{x}}$			955,18	Bs/M2	
$3\delta_{\bar{x}}$			2.865,54	Bs/M2	

Error Accidental = Valor Atípico



Descarte de los Referenciales o Comparables y cálculo del nuevo Término Central de la Serie

Identificac.	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	Si $ \bar{x} \pm x_i > 3\delta_{\bar{x}}$
1	22.115,38	2.923,28	8.545.992,94	
2	21.199,26	2.007,16	4.028.691,27	
3	20.436,56	1.244,46	1.548.660,69	
4	17.500,00	-1.692,10	2.863.202,41	
5	21.666,67	2.474,57	6.123.460,19	
7	17.241,38	-1.950,72	3.805.311,21	
9	12.428,25	-6.763,85	45.749.666,82	
10	22.413,79	3.221,69	10.379.306,45	
11	18.803,42	-388,68	151.073,07	
12	19.047,62	-144,48	20.874,75	
13	18.487,39	-704,71	496.609,20	
14	18.965,52	-226,58	51.339,75	
\bar{x}		19.192,10		
$\sum_{i=1}^n (x_i - \bar{x})^2$		83.763.828,75		
$n = (n - 1)$		132		
$\delta_{\bar{x}}$		796,60		Bs/M2

Cálculo del valor de la "Incertidumbre" de la serie:

$$\delta_{\bar{x}} = \frac{\sigma_{n-1}}{\sqrt{n}} \quad \text{Bs/M2}$$

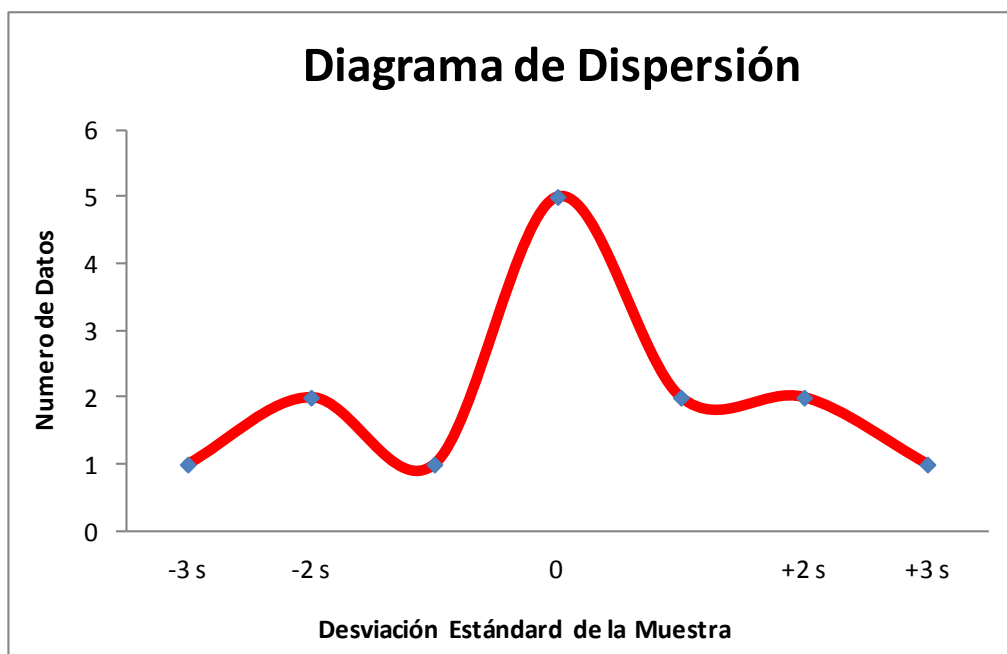
$$\delta_{\bar{x}} = 796,60 \quad \text{Bs/M2}$$



Cálculo del Intervalo de existencia del "Término Central de la Serie":

$$T.C.S = \bar{x} \pm \delta\bar{x}$$

$$T.C.S. = 19.192,10 \pm 796,6 \text{ Bs/M}^2$$





Cálculo del valor del apartamento a tasar::

AREA DEL APARTAMENTO:	57,65 M2
TERMINO CENTRAL DE LA SERIE:	19.192,10 Bs/M2
VALOR DEL APARTAMENTO:	1.106.424,57 Bs.



APENDICE 2

EJEMPLO NUMERICO RESUELTO: Uso del Estimador-M de Huber, como metodología para la determinación del Término Central de una "Serie muy Pequeña"

I) Planteamiento del Caso:

Se trata de determinar el Valor de un "Terreno de Secano" o "Terreno Inculto" en las cercanías de la población de Bruzual. Municipio Muñoz. Estado Apure. Venezuela.

La ciudad de Bruzual, está ubicado al Suroeste de Venezuela. Forma parte de los Llanos Venezolanos, tierra dedicada a la ganadería desde la época de la colonia:



Mapa del Estado Apure

Una vez efectuada la inspección del inmueble (lote de terreno inculto); se procedió en la Oficina de Registro Inmobiliario del Municipio Muñoz del Estado Apure, a seleccionar una muy pequeña serie de inmuebles referenciales o comparables al inmueble objeto del avalúo. Los únicos datos disponibles:

Lote de Terreno a valorar: Vía Puerto de Nutrias S/N Area (Has.): 156,25

REFERENCIALES O COMPARABLES
Lotes o Parcelas de Terreno Inculto en los alrededores de la ciudad de Bruzual

REF	INMUEBLE COMPARABLE	VALOR (Bs.F.)	AREA (Has.)	PRECIO UNITARIO (Bs.F./M2)
A	CASERIO LAS TIAMITAS. S/N	1.500	150,00	10,00
B	FINCA LA LOLA. SECTOR CHAPARRITO	1.378	125,25	11,00
C	CARRETERA NACIONAL.CAÑO SECO	1.953	177,55	11,00
D	CASERIO DOLORES.LOTE A	1.566	130,50	12,00
E	VIA CHORONCO ARRIBA.FINCA PAZ	12.000	120,00	100,00



Tal como se aprecia en los referenciales o comparables seleccionados (únicos disponibles); el Comparable "E" se destaca de los demás, al presentar un Precio Unitario de 100 Bs. /Ha

La primera intención del Profesional Tasador, sería considerarlo un "Error Grosero" y eliminarlo de la serie.

Sin embargo, la función principal de los Estimadores "No Paramétricos" o "Robustos"; es precisamente la determinación de un "Término Central" en una serie donde existen "Valores Extremos" o "Outliers".

Sin considerar que la escasez de "Referenciales y Comparables" del sector, no permitirían prescindir de ninguno de ellos, so pena de resultar en una serie muy "poco significativa" para la estimación de un "Término Central".

II) Procedimiento:

Inmuebles Referenciales o Comparables:

REF	AREA (Ha.)	VALOR Bs./Ha
A	150,00	10,00
B	125,25	11,00
C	177,55	11,00
D	130,50	12,00
E	120,00	100,00

PRIMERA ITERACION

a) Se calcula la Mediana de la Serie, como el "Promedio Robusto Inicial":

$$Me = \bar{x}' = 11,00 \text{ Bs./Ha.}$$



b) Se calcula la "MAD": "Desviación Mediana" o "Desviación Mediana", para Datos No Agrupados en Clases:

$$MAD = D_{Me} = \sum \frac{|x_i - Me|}{n}$$

REF	x_i	$ x_i - Me $	
A	10,00	1,00	
B	11,00	0,00	
C	11,00	0,00	
D	12,00	1,00	
E	100,00	89,00	
$\Sigma =$		91,00	
MAD =	$D_{Me} = \sum \frac{ x_i - Me }{n}$	18,20	Bs./Ha

c) Se calculará la Desviación Mediana Basada en MAD (SMAD) para "Datos No Agrupados en Clases"

$$SMAD = 1,483 * MAD$$

$$SMAD = 1,483 * \sum \frac{|x_i - Me|}{n}$$

$$SMAD = 26,99 \text{ Bs./Ha.}$$



d) Se asigna inicialmente como "Desviación Estándar Robusta" σ^* , al valor de la Desviación Mediana Basada en MAD (SMAD)

$$\sigma^* = \text{SMAD} = 26,99 \text{ Bs./Ha.}$$

e) Se calcula el "Intervalo Máximo de Error" $\bar{x}^* \pm \varepsilon$

$$\text{Donde: } \varepsilon = 1,5 \sigma^*$$

$$\varepsilon = 40,49 \text{ Bs./Ha.}$$

$$\begin{cases} \bar{x}^* - \varepsilon & = & -29,49 & \text{Bs./Ha.} \\ \bar{x}^* + \varepsilon & = & 51,49 & \text{Bs./Ha.} \end{cases}$$

SEGUNDA ITERACION

f) Se sustituirán todos los datos x_i mayores que el "Límite Superior del Intervalo Máximo de Error" por el valor del "Límite Superior del Intervalo Máximo de Error":

REF	x_i	Sustituído
A	10,00	10,00
B	11,00	11,00
C	11,00	11,00
D	12,00	12,00
E	100,00	51,49



g) Se calculará el "Promedio Robusto" de la nueva serie de datos:

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

$$\bar{x}^* = 19,10 \quad \text{Bs./Ha.}$$

h) Se calculará la Desviación Estándar para "Datos No Agrupados" de la nueva serie:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$

i) Se calculará la "Desviación Estándar Robusta" de la nueva serie:

$$\sigma^* = 1,134 = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$



REF	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	
A	10,00	-9,10	82,76	
B	11,00	-8,10	65,56	
C	11,00	-8,10	65,56	
D	12,00	-7,10	50,37	
E	51,49	32,39	1.048,98	
$\Sigma =$			1.313,23	
$(n-1)$			4	
$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$			18,12	Bs./Ha
$\sigma^* = 1,134 * \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$			20,55	Bs./Ha

j) Se calcula el nuevo "Intervalo Máximo de Error"

Donde: $\varepsilon = 1,5 \sigma^*$

$\varepsilon = 30,83 \text{ Bs./Ha.}$

$$\begin{cases} \bar{x}^* - \varepsilon & = & -11,73 & \text{Bs./Ha.} \\ \bar{x}^* + \varepsilon & = & 49,92 & \text{Bs./Ha.} \end{cases}$$



j) Se calcula el nuevo "Intervalo Máximo de Error"

Donde: $\varepsilon = 1,5 \sigma^*$

$\varepsilon = 30,83$ Bs./Ha.

$$\begin{cases} \bar{x}^* - \varepsilon & = & -11,73 & \text{Bs./Ha.} \\ \bar{x}^* + \varepsilon & = & 49,92 & \text{Bs./Ha.} \end{cases}$$

k) Resumen de las Iteraciones realizadas hasta el momento:

ITERACION	\bar{x}^*	σ^*
1	11,00	26,99
2	19,10	20,55



TERCERA ITERACION

l) Se sustituirán nuevamente todos los datos x_i mayores que el "Límite Superior del Intervalo Máximo de Error" por el valor del "Límite Superior del Intervalo Máximo de Error":

REF	x_i	Sustituído
A	10,00	10,00
B	11,00	11,00
C	11,00	11,00
D	12,00	12,00
E	51,49	49,92

m) Se calculará el nuevo "Promedio Robusto" de la nueva serie de datos:

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

$$\bar{x}^* = 18,78 \quad \text{Bs./Ha.}$$

n) Se calculará la nueva Desviación Estándar para "Datos No Agrupados" de la nueva serie:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$



o) Se calculará la nueva "Desviación Estándar Robusta" de la nueva serie:

$$\sigma^* = 1,134 * \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$

REF	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	
A	10,00	-8,78	77,17	
B	11,00	-7,78	60,60	
C	11,00	-7,78	60,60	
D	12,00	-6,78	46,03	
E	49,92	31,14	969,55	
$\Sigma =$			1.213,94	
$(n-1)$			4	
$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$			17,42	Bs./Ha
$\sigma^* = 1,134 * \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$			19,75	Bs./Ha

p) Se calcula el nuevo "Intervalo Máximo de Error"

Donde: $\varepsilon = 1,5 \sigma^*$

$\varepsilon = 29,63 \text{ Bs./Ha.}$

$$\begin{cases} \bar{x}^* - \varepsilon & = & -10,84 & \text{Bs./Ha.} \\ \bar{x}^* + \varepsilon & = & 48,41 & \text{Bs./Ha.} \end{cases}$$



q) Resumen de las Iteraciones realizadas

ITERACION	\bar{x}^*	σ^*
1	11,00	26,99
2	19,10	20,55
3	18,78	19,75

CUARTA ITERACION

r) Se sustituirán nuevamente todos los datos x_i mayores que el "Límite Superior del Intervalo Máximo de Error" por el valor del "Límite Superior del Intervalo Máximo de Error":

REF	x_i	Sustituído
A	10,00	10,00
B	11,00	11,00
C	11,00	11,00
D	12,00	12,00
E	49,92	48,41



s) Se calculará el nuevo "Promedio Robusto" de la nueva serie de datos:

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

$$\bar{x}^* = 18,48 \quad \text{Bs./Ha.}$$

t) Se calculará la nueva Desviación Estándar para "Datos No Agrupados" de la nueva serie:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$

v) Se calcula el nuevo "Intervalo Máximo de Error"

$$\text{Donde: } \varepsilon = 1,5 \sigma^*$$

$$\varepsilon = 28,47 \quad \text{Bs./Ha.}$$

$$\left\{ \begin{array}{l} \bar{x}^* - \varepsilon = -9,99 \quad \text{Bs./Ha.} \\ \bar{x}^* + \varepsilon = 46,95 \quad \text{Bs./Ha.} \end{array} \right.$$



u) Se calculará la nueva "Desviación Estándar Robusta" de la nueva serie:

$$\sigma^* = 1,134 * \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$

REF	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	
A	10,00	-8,48	71,94	
B	11,00	-7,48	55,98	
C	11,00	-7,48	55,98	
D	12,00	-6,48	42,01	
E	48,41	29,93	895,66	
$\Sigma =$			1.121,57	
$(n-1)$			4	
$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$			16,74	Bs./Ha
$\sigma^* = 1,134 * \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$			18,98	Bs./Ha

v) Se calcula el nuevo "Intervalo Máximo de Error"

Donde: $\varepsilon = 1,5 \sigma^*$

$\varepsilon = 28,47$ Bs./Ha.

$$\left\{ \begin{array}{l} \bar{x}^* - \varepsilon = -9,99 \text{ Bs./Ha.} \\ \bar{x}^* + \varepsilon = 46,95 \text{ Bs./Ha.} \end{array} \right.$$



u) Resumen de las Iteraciones realizadas

ITERACION	\bar{x}^*	σ^*
1	11,00	26,99
2	19,10	20,55
3	18,78	19,75
4	18,48	18,98

CONVERGENCIA

III Solución del problema planteado:

Calculo del Valor de la Parcela de Terreno

Identificación: Parcela de Terreno S/N

Ubicación: Km. 17 de la Carretera Nacional a Puerto de Nutrias. Municipio Santa Rosalía del Estado Portuguesa.

Area: 156,25 Has.

Valor Unitario: 18,48 Bs./Ha

Valor del Terreno: 2.887,79 Bs.



APENDICE 3

EJEMPLO NUMERICO RESUELTO: Uso del Estimador-L de Huber, como metodología para la determinación del Término Central de una "Serie muy Pequeña"

Planteamiento del Caso:

Se trata de determinar el Valor de una "Parcela de Terreno Urbana" con un área de 195.75 M2, identificada con el Nro. 23-1 en la población de El Chaparro. Municipio McGregor. Estado Anzoátegui. Venezuela.



Ubicación de la población de El Chaparro

La ciudad de El Chaparro, está ubicado al Oriente de Venezuela. Forma parte de los Llanos Venezolanos, tierra dedicada a la ganadería y agricultura desde la época de la colonia:



Ubicación de la parcela urbana Nro. 23-1 en la población de El Chaparro



Una vez efectuada la inspección del inmueble (parcela de terreno urbano Nro. 23-1); se procedió en la Oficina de Registro Inmobiliario del Municipio Aragua de Barcelona (jurisdicción donde se registran los inmuebles del Municipio Gral. Sir Arthur McGregor) del Estado Anzoátegui, donde se seleccionó una muy pequeña serie de inmuebles referenciales o comparables al inmueble objeto del avalúo. Los únicos datos disponibles:

Identificac.	Ciudad	Dirección del Inmueble	Vendedor	Comprador	Precio de Venta Bs.F.	P.U.	Area m2
A	EL CHAPARRO	CALLE JOSEFITA CALATRAVA PC. No. 37	RAUL OLIVEIRA	NATALIA DIAZ	9.538	50,00	190,75
B	EL CHAPARRO	CALLE MCGREGOR PC. S/N	VISTA AL VALLE.	NELSON MOSCHELLA	10.361	52,00	199,25
C	EL CHAPARRO	CALLE JOSEFITA CALATRAVA PC. No. 23	WIDAD TORBAY	ARTURO JIMENEZ	11.487	59,00	194,70
D	EL CHAPARRO	CALLE JOSEFITA CALATRAVA PC. No. 25	YALITZA GAZZANEO	INVERS. ARDAGNA FORTE C.A.	191.093	950,00	201,15
E	EL CHAPARRO	CALLE LIBERTAD PC No. 43	CARLOS DA SILVA	ELIANA TEIXEIRA	193.602	985,00	196,55

Calculo del Promedio Inicial: 419,20

SERIE ORIGINAL	\bar{x} =	419,20	Bs./M2
PRIMERA ITERACION	\bar{x} =		Bs./M3
SEGUNDA ITERACION	\bar{x} =		Bs./M4
TERCERA ITERACION	\bar{x} =		Bs./M5
CUARTA ITERACION	\bar{x} =		Bs./M6
QUINTA ITERACION	\bar{x} =		Bs./M7



PRIMERA ITERACION:

Cálculo de los Residuos, de la Desviación Estándar y la Incertidumbre o Error:

Donde:

$$\delta_{\bar{x}} = \frac{\sigma_{n-1}}{\sqrt{n}}$$

$$\sigma_{n-1} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n-1)}}$$

$$\delta_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n * (n - 1)}}$$



REF	x_i	\bar{x}	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
A	50	419,20	369,20	136.308,64
B	52	419,20	367,20	134.835,84
C	59	419,20	360,20	129.744,04
D	950	419,20	530,80	281.748,64
E	985	419,20	565,80	320.129,64
Σ				1.002.766,80
$n*(n-1)$				20
$\delta_{\bar{x}}$				223,92

Donde:

$$\delta_{\bar{x}} = 223,92 \text{ Bs./M}^2$$

Cálculo de los Pesos:

$$K = 2$$

Entonces: $K * \delta_{\bar{x}} = 447,83 \text{ Bs./M}^2$

Se calcularán los pesos de acuerdo a los siguientes modelos:

$$\left\{ \begin{array}{l} p = 1 \text{ si } |x_i - \bar{x}| \leq K * \delta_{\bar{x}} \\ p = \frac{K * \delta_{\bar{x}}}{|x_i - \bar{x}|} \text{ si } |x_i - \bar{x}| > K * \delta_{\bar{x}} \end{array} \right.$$



REF	x_i	\bar{x}	$ x_i - \bar{x} $	Estatus	Pi	$P_i * x_i$
A	50	419,20	369,20	$\leq K * \delta_{\bar{x}}$	1,00	50,00
B	52	419,20	367,20	$\leq K * \delta_{\bar{x}}$	1,00	52,00
C	59	419,20	360,20	$\leq K * \delta_{\bar{x}}$	1,00	59,00
D	950	419,20	530,80	$> K * \delta_{\bar{x}}$	0,84	801,51
E	985	419,20	565,80	$> K * \delta_{\bar{x}}$	0,79	779,63
Σ					4,64	1742,14

Se calcula el nuevo promedio: $\bar{x} = \frac{\sum(P_i * x_i)}{\sum P_i} = 375,85 \text{ Bs./M}^2$

SERIE ORIGINAL	$\bar{x} =$	419,20	Bs./M²
PRIMERA ITERACION	$\bar{x} =$	375,85	Bs./M³
SEGUNDA ITERACION	$\bar{x} =$		Bs./M⁴
TERCERA ITERACION	$\bar{x} =$		Bs./M⁵
CUARTA ITERACION	$\bar{x} =$		Bs./M⁶
QUINTA ITERACION	$\bar{x} =$		Bs./M⁷



SEGUNDA ITERACION:

Cálculo de los Nuevos Residuos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
A	50	375,85	325,85	106.178,13
B	52	375,85	323,85	104.878,73
C	59	375,85	316,85	100.393,83
D	950	375,85	574,15	329.648,38
E	985	375,85	609,15	371.063,89
Σ				1.012.162,97
$n*(n-1)$				20
$\hat{\sigma}_{\bar{x}}$				224,96

Los Nuevos Pesos se calcularán de acuerdo a los siguientes modelos:

$$\left\{ \begin{array}{l} P = 1 \text{ si } |x_i - \bar{x}| \leq K * \hat{\sigma}_{\bar{x}} \\ P = \frac{K * \hat{\sigma}_{\bar{x}}}{|x_i - \bar{x}|} \text{ si } |x_i - \bar{x}| > K * \hat{\sigma}_{\bar{x}} \end{array} \right.$$

Donde: $K * \hat{\sigma}_{\bar{x}} = 449,93$



Cálculo de los Nuevos Pesos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	Estatus	Pi	$P_i * x_i$
A	50	375,85	325,85	$\leq K * \delta_{\bar{x}}$	1,00	50,00
B	52	375,85	323,85	$\leq K * \delta_{\bar{x}}$	1,00	52,00
C	59	375,85	316,85	$\leq K * \delta_{\bar{x}}$	1,00	59,00
D	950	375,85	574,15	$> K * \delta_{\bar{x}}$	0,78	744,45
E	985	375,85	609,15	$> K * \delta_{\bar{x}}$	0,74	727,53
Σ					4,52	1632,99

Se calcula el nuevo promedio: $\bar{x} = \frac{\sum(P_i * x_i)}{\sum P_i} = 361,10$

SERIE ORIGINAL	$\bar{x} =$	419,20	Bs./M2
PRIMERA ITERACION	$\bar{x} =$	375,85	Bs./M3
SEGUNDA ITERACION	$\bar{x} =$	361,10	Bs./M4
TERCERA ITERACION	$\bar{x} =$		Bs./M5
CUARTA ITERACION	$\bar{x} =$		Bs./M6
QUINTA ITERACION	$\bar{x} =$		Bs./M7



TERCERA ITERACION:

Cálculo de los Nuevos Residuos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
A	50	361,10	311,10	96.783,64
B	52	361,10	309,10	95.543,24
C	59	361,10	302,10	91.264,83
D	950	361,10	588,90	346.802,39
E	985	361,10	623,90	389.250,34
Σ				1.019.644,45
$n*(n-1)$				20
$\delta_{\bar{x}}$				225,79

Donde: $K * \delta_{\bar{x}} = 451,58$

Cálculo de los Nuevos Pesos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	Estatus	Pi	$P_i * x_i$
A	50	361,10	311,10	$\leq K * \delta_{\bar{x}}$	1,00	50,00
B	52	361,10	309,10	$\leq K * \delta_{\bar{x}}$	1,00	52,00
C	59	361,10	302,10	$\leq K * \delta_{\bar{x}}$	1,00	59,00
D	950	361,10	588,90	$> K * \delta_{\bar{x}}$	0,77	728,49
E	985	361,10	623,90	$> K * \delta_{\bar{x}}$	0,72	712,95
Σ					4,49	1602,44



Se calcula el nuevo promedio: $\bar{x} = \frac{\sum(P_i * x_i)}{\sum P_i} = 356,84$

SERIE ORIGINAL	\bar{x} =	419,20	Bs./M2
PRIMERA ITERACION	\bar{x} =	375,85	Bs./M3
SEGUNDA ITERACION	\bar{x} =	361,10	Bs./M4
TERCERA ITERACION	\bar{x} =	356,84	Bs./M5
CUARTA ITERACION	\bar{x} =		Bs./M6
QUINTA ITERACION	\bar{x} =		Bs./M7

CUARTA ITERACION:

Cálculo de los Nuevos Residuos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
A	50	356,84	306,84	94.150,91
B	52	356,84	304,84	92.927,55
C	59	356,84	297,84	88.708,78
D	950	356,84	593,16	351.838,55
E	985	356,84	628,16	394.584,74
Σ				1.022.210,52
$n*(n-1)$				20
$\delta_{\bar{x}}$				226,08

Donde: $K * \delta_{\bar{x}} = 452,15$



Cálculo de los Nuevos Pesos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	Estatus	Pi	$P_i * x_i$
A	50	356,84	306,84	$\leq K * \delta_x$	1,00	50,00
B	52	356,84	304,84	$\leq K * \delta_x$	1,00	52,00
C	59	356,84	297,84	$\leq K * \delta_x$	1,00	59,00
D	950	356,84	593,16	$> K * \delta_x$	0,76	724,16
E	985	356,84	628,16	$> K * \delta_x$	0,72	709,01
Σ					4,48	1594,17

Se calcula el nuevo promedio:
$$\bar{x} = \frac{\sum(P_i * x_i)}{\sum P_i} = 355,68$$

SERIE ORIGINAL	$\bar{x} =$	419,20	Bs./M2
PRIMERA ITERACION	$\bar{x} =$	375,85	Bs./M3
SEGUNDA ITERACION	$\bar{x} =$	361,10	Bs./M4
TERCERA ITERACION	$\bar{x} =$	356,84	Bs./M5
CUARTA ITERACION	$\bar{x} =$	355,68	Bs./M6
QUINTA ITERACION	$\bar{x} =$		Bs./M7



QUINTA ITERACION:

Cálculo de los Nuevos Residuos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
A	50	355,68	305,68	93.438,22
B	52	355,68	303,68	92.219,51
C	59	355,68	296,68	88.017,04
D	950	355,68	594,32	353.220,24
E	985	355,68	629,32	396.047,87
Σ				1.022.942,88
$n*(n-1)$				20
$\delta_{\bar{x}}$				226,16

Donde: $K * \delta_{\bar{x}} = 452,31$

Cálculo de los Nuevos Pesos:

REF	x_i	\bar{x}	$ x_i - \bar{x} $	Estatus	Pi	$P_i * x_i$
A	50	355,68	305,68	$\leq K * \delta_{\bar{x}}$	1,00	50,00
B	52	355,68	303,68	$\leq K * \delta_{\bar{x}}$	1,00	52,00
C	59	355,68	296,68	$\leq K * \delta_{\bar{x}}$	1,00	59,00
D	950	355,68	594,32	$> K * \delta_{\bar{x}}$	0,76	723,01
E	985	355,68	629,32	$> K * \delta_{\bar{x}}$	0,72	707,95
Σ					4,48	1591,96

Se calcula el nuevo promedio: $\bar{x} = \frac{\Sigma(P_i * x_i)}{\Sigma P_i} = 355,36$



SERIE ORIGINAL	\bar{x} =	419,20	Bs./M2
PRIMERA ITERACION	\bar{x} =	375,85	Bs./M3
SEGUNDA ITERACION	\bar{x} =	361,10	Bs./M4
TERCERA ITERACION	\bar{x} =	356,84	Bs./M5
CUARTA ITERACION	\bar{x} =	355,68	Bs./M6
QUINTA ITERACION	\bar{x} =	355,36	Bs./M7

CONVERGENCIA

Cálculo del Valor del inmueble:

Parcela de Terreno N° 23-1 de la Calle Independencia de la población de El Chaparro. Municipio McGregor. Estado Anzoátegui. Venezuela

AREA DE LA PARCELA N° 23-1:	195,75	M2
VALOR UNITARIO:	355,36	Bs./M2
VALOR DE LA PARCELA N° 23-1:	69.562,50	Bs.