



NOTA TÉCNICA
N.º 02 | 2020

Media o mediana ¿cuál estimador usar? Caso de la expectativa de inflación a 12 meses

Andrea Collado Chaves

Fotografía de portada: "Presentes", conjunto escultórico en bronce, año 1983, del artista costarricense Fernando Calvo Sánchez. Colección del Banco Central de Costa Rica.



¿Media o mediana? ¿Cuál estimador usar? Caso de la expectativa de inflación a 12 meses

Andrea Collado Chaves ‡

Las ideas expresadas en este documento son de la autora y no representan las del Banco Central de Costa Rica.

Resumen

El objetivo de esta nota técnica es orientar a los usuarios de la Encuesta mensual de expectativas de inflación y tipo de cambio del Banco Central de Costa Rica, en el uso de la media y la mediana de las expectativas de inflación en aquellos meses en los cuales existen diferencias entre ambos estimadores.

Se verificó que las diferencias entre la media y la mediana son estadísticamente significativas. También, se comprobó que la media está sesgada por valores atípicamente altos que comienzan a aparecer con mayor frecuencia, cuando se amplía el tamaño de la muestra y se aumenta la proporción de empresarios en la Encuesta, y tienden a disminuir después de diciembre de 2014, cuando se reemplazó el 95% de los informantes.

A fin de eliminar el efecto de los valores atípicos y extremos, se calcularon dos series de medias: una sin atípicos y otra sin extremos. Los resultados señalan que las cuatro series: media, media sin atípicos, media sin extremos y mediana, muestran correlaciones bivariadas cercanas a 1, es decir, miden lo mismo.

Si los usuarios deben seleccionar una u otra medida para un mes específico, se recomienda usar la mediana siempre que la diferencia entre la media y mediana sea mayor a 0,1 puntos porcentuales (p.p.). Si se requiere de la serie temporal, se sugiere comparar los resultados utilizando la media y la mediana y evaluar cuál se ajusta mejor a los objetivos del análisis.

Palabras clave: Prueba de hipótesis, encuestas, nivel de precios, meta de inflación

Clasificación JEL: C12, C83, E31. D84

‡ Departamento de Investigación Económica, División Económica, BCCR, colladoca@bccr.fi.cr.

Le doy gracias a Evelyn Muñoz Salas, Carlos Segura Rodríguez y María Elena Ávila Gamboa por sus revisiones y comentarios y por el apoyo brindado durante el proceso de investigación.

Choosing an Estimator for Inflation Expectations: Mean vs Median

Andrea Collado Chaves ‡

The ideas expressed in this paper are those of the author and not necessarily represent the view of the Central Bank of Costa Rica.

Abstract

The main purpose of this technical note is to guide users of the Monthly Survey of Inflation Expectations and Exchange Rate, conducted by the Central Bank of Costa Rica, in selecting the mean and median of the inflation expectations when both estimators are different.

The time series of the mean and median since January 2006 to December 2019 were used to prove that, in average, the mean is statistically greater than the median. The positive outlier are in part responsible for this gap. These values appeared when the sample size was enlarged and the participation of entrepreneurs in the Survey was increased. They tend to decrease after December 2014 when 95% of the samples was replaced.

In order to eliminate the outliers, two series were calculated: one without outliers and one without extreme observations. The results indicate that the four series: mean, mean without outliers, mean without extremes and median have bivariate correlation close to 1, therefore they measure the same expectations.

When users must select only one estimator for a specific month, it is recommended to use the median if the mean is greater by 0.1 percentage points (p.p.) or more. If the time series is required, it is suggested to compare the results using the mean and median, and evaluate which one is better according to the research goals.

Key words: Hypothesis Testing, Survey Methods, Price Level, Inflation Target

JEL codes: C12, C83, E31, D84

‡ Department of Economic Research, Economic Division, BCCR, Email address colladoca@bccr.fi.cr.

I thank Evelyn Muñoz Salas, Carlos Segura Rodríguez and María Elena Ávila Gamboa for their reviews, comments and support during the research process.

Contenido

1.	Introducción	1
2.	Datos y métodos	1
2.1	Microdatos, informantes y cambios metodológicos de la Encuesta	2
2.2	Análisis estadísticos.....	5
2.2.1	Análisis univariado.....	5
2.2.2	Asociación, correlación y regresión lineal múltiple	7
3.	Resultados.....	7
3.1	Diferencias significativas entre la media y la mediana.....	12
3.2	Influencia de los valores atípicos y extremos	13
3.3	Variables que explican la diferencia entre la media y la mediana.....	15
4.	Conclusiones y recomendaciones	20
5.	Referencias.....	21
6.	Anexos	22

Lista de cuadros

Cuadro 1. Variables y sus fuentes	1
Cuadro 2. Cantidad de informantes que responden por primera y última vez a la pregunta sobre expectativas de inflación según meses específicos. Mayo, 2009 – Noviembre, 2017	3
Cuadro 3. Costa Rica: Estadísticas descriptivas de la meta de inflación, inflación del mes anterior, media y mediana de las expectativas de inflación y su diferencia, según período.	9
Cuadro 4. Costa Rica: Correlación de Spearman entre la meta de inflación, inflación del mes anterior, media y mediana de las expectativas de inflación, según período.	10
Cuadro 5. Resultado de la prueba de hipótesis T-Student para muestras pareadas Diferencia entre la media y la mediana	13
Cuadro 6. Resultado de la prueba de hipótesis T-Student para muestras pareadas Diferencia entre cada media y la mediana	15
Cuadro 7. Costa Rica: Estadística descriptiva de las variables incluidas en los modelos de regresión lineal múltiple	16
Cuadro 8. Resultados de las regresiones lineales múltiples.....	18
Cuadro 9. Resultados de la evaluación y pruebas de diagnóstico de los modelos.....	19
Cuadro 10. Costa Rica: Estadísticas descriptivas de la media y mediana de las expectativas de inflación, según período.....	22
Cuadro 11. Costa Rica: Matriz de correlación bivariada de Spearman entre la media y mediana de las expectativas de inflación, según período	23

Lista de gráficos

- Gráfico 1. Encuesta mensual de expectativas de inflación y tipo de cambio del Banco Central de Costa Rica: Proporción de empresarios y analistas dentro de la muestra Enero, 2006 - Diciembre, 2019 4
- Gráfico 2. Costa Rica: Series mensuales de la media y mediana de la expectativa de inflación, variación interanual del IPC del mes anterior y meta de inflación a 12 meses 8
- Gráfico 3. Costa Rica: Series mensuales de la media y la mediana de la expectativa de inflación a 12 meses y su rango de variación. Enero, 2006 - Diciembre, 2019 11
- Gráfico 4. Distribución de las diferencias entre las series de la media y la mediana de las expectativas de inflación 12
- Gráfico 5. Costa Rica: Series mensuales de la media y la mediana de las expectativas de inflación a 12 meses, rango intercuartil y rango. Enero, 2006 - Diciembre, 2019.. 14
- Gráfico 6. Distribución de las diferencias entre la media sin atípicos y la mediana y entre la media sin extremos y la mediana de las expectativas de inflación 24

1. Introducción

Se compara la media y la mediana mensual de las expectativas de inflación a 12 meses, las cuales se obtienen de la Encuesta de expectativas de inflación y tipo de cambio, con el fin de reflexionar sobre el uso de una u otra medida de posición central en aquellos meses que ambos estimadores difieren.

Por ejemplo, en abril de 2020, la media fue de 3,5% y la mediana fue de 3,0%; entonces, ¿cuál se debería usar? Antes de cualquier recomendación, hay que comprobar que la diferencia observada entre la media y la mediana (0,5 p.p.) no se debe a fluctuaciones aleatorias en las respuestas de los informantes.

Si el interés no está enfocado en un mes específico, sino en la serie temporal de la expectativa media o mediana, lo recomendable es analizar con mayor detalle las preguntas de investigación y los métodos estadísticos que serán implementados.

Esta nota técnica usa la teoría estadística frecuentista y la estadística paramétrica para guiar en la selección de uno u otro estimador, según los posibles usos que se den a la información. El documento se compone de cuatro secciones y las referencias bibliográficas. El título siguiente explica las variables y sus fuentes así como los análisis estadísticos aplicados. En la tercera parte, se presentan los resultados y finalmente las conclusiones y recomendaciones.

2. Datos y métodos

Los datos del Índice de Precios al Consumidor provienen del Instituto Nacional de Estadística y Censos de Costa Rica (Instituto Nacional de Estadística y Censos, 2020). La serie mensual de la meta de inflación es una reconstrucción a partir de las publicaciones del Programa Macroeconómico y sus revisiones. Se usan los micro datos de la Encuesta de expectativas de inflación y tipo de cambio (Banco Central de Costa Rica, 2020) para el cálculo de otras variables que se resumen en el Cuadro 1, junto con sus fuentes.

Cuadro 1. Variables y sus fuentes

Nombre	Variable	Fuente
Inflación	Variación interanual del IPC	INEC (2020)
Meta	Meta de inflación	BCCR (2019)
Media	Media mensual de expectativas de inflación	BCCR (2020)
Mediana	Mediana mensual de expectativas de inflación	BCCR (2020)
$\bar{x} - Me$	Diferencia entre la media y la mediana	BCCR (2020)
Empresario	Proporción de empresarios en la muestra	BCCR (2020)
Atípicos bajos	Proporción de atípicos bajos en la muestra	BCCR (2020)
Atípicos altos	Proporción de atípicos altos en la muestra	BCCR (2020)

Fuente: Elaboración propia

Asimismo, se usan, como variables dicotómicas, los quiebres estructurales de la inflación del país, fechados por Torres (2012) en julio de 2006 y mayo de 2009 y las fechas de cambios en el tamaño de la muestra o cambios en el diseño muestral que se obtuvieron a partir del análisis de los microdatos. Las series mensuales de la media y la mediana utilizadas inician en enero de 2006 y terminan en diciembre de 2019.

2.1 Microdatos, informantes y cambios metodológicos de la Encuesta

La medición de la expectativa se basa en una muestra de empresas y otros expertos que responden a la pregunta: ¿Cuál es su expectativa para la tasa de inflación, medida por el Índice de Precios al Consumidor, al término del período comprendido entre el 1 de mes de año y el 30 o 31 de mes de año (12 meses después)? Así por ejemplo, el cuestionario enviado en diciembre de 2019, preguntó por la expectativa del 1 de diciembre de 2019 al 30 de noviembre de 2020.

Hace algunos años, se hacían entrevistas cara a cara y telefónicas. También, se recibían los formularios por fax. De 2010 en adelante, el instrumento se completa en línea, el vínculo se envía durante la segunda semana de cada mes, luego de la publicación del Índice de Precios al Consumidor (IPC). Los participantes tienen aproximadamente tres semanas para llenar el cuestionario. Si un informante así lo desea, el formulario se completa mediante una entrevista telefónica o se reciben formularios por correo electrónico.

El estudio es longitudinal y la participación es voluntaria. De enero de 2006 a diciembre de 2019, habían participado 575 informantes. Al principio se seleccionó, por juicio experto, a 30 participantes, 57% eran analistas financieros, 20% académicos, 13% consultores y 10% empresarios (Arias Cubillo, Durán Viquez, & Kikut Valverde, 2006, pág. 7). A finales de enero de 2006, primer mes en que se envió el cuestionario, solo respondieron 27 personas: 52% analistas financieros, 33% consultores, 1 académico y 3 empresarios.

Durante los meses que han transcurrido desde que comenzó la encuesta, la cantidad promedio de respuestas ha sido de 60 y el 75% de los meses se han recibido 36 respuestas o más. En abril de 2009, respondieron 19 expertos, esta fue la muestra más pequeña que ha tenido la encuesta. De mayo de 2009 a mayo de 2012, se recibieron alrededor de 50 respuestas, en su mayoría de analistas financieros. En junio de 2012, la cantidad de respuestas aumentó a 80 y en noviembre de 2017 llegó a 135 (máximo) ¹.

¹ El cálculo de estimadores como la media o la desviación estándar requiere de al menos 30 respuestas para producir estimaciones consistentes. Además, esta cantidad de respuestas por mes favorece el enfoque paramétrico.

El porcentaje de no respuesta, durante el período en estudio, fue de 30%. Aunque 11% de los informantes solo responden una vez, la mayoría participa al menos por un año y 10% por 30 meses o más.

La fecha en que cada informante dio su primera y última respuesta se usó para determinar los cambios de muestra². El Cuadro 2 compara la cantidad de respuestas recibidas cada mes con la cantidad de informantes que entran por primera vez al estudio y los que dan su última respuesta en esa fecha (salen del estudio). Solo se muestran los meses que podrían estar asociados con variaciones en el diseño muestral.

Cuadro 2. Cantidad de informantes que responden por primera y última vez a la pregunta sobre expectativas de inflación según meses específicos. Mayo, 2009 – Noviembre, 2017

Mes	Muestra	Entran	Salen
May, 2009	25	5	1
Jun, 2009	36	9	0
Jul, 2009	40	2	0
May, 2012	52	3	26
Jun, 2012	80	56	9
Jul, 2012	81	17	11
Oct, 2014	73	1	11
Nov, 2014	71	0	63
Dic, 2014	70	66	2
Mar, 2017	60	13	9
Abr, 2017	59	12	5
Jun, 2017	74	12	6
Ago, 2017	100	35	2
Oct, 2017	109	6	21
Nov, 2017	135	40	5

Fuente: Elaboración propia con base en BCCR (2020)

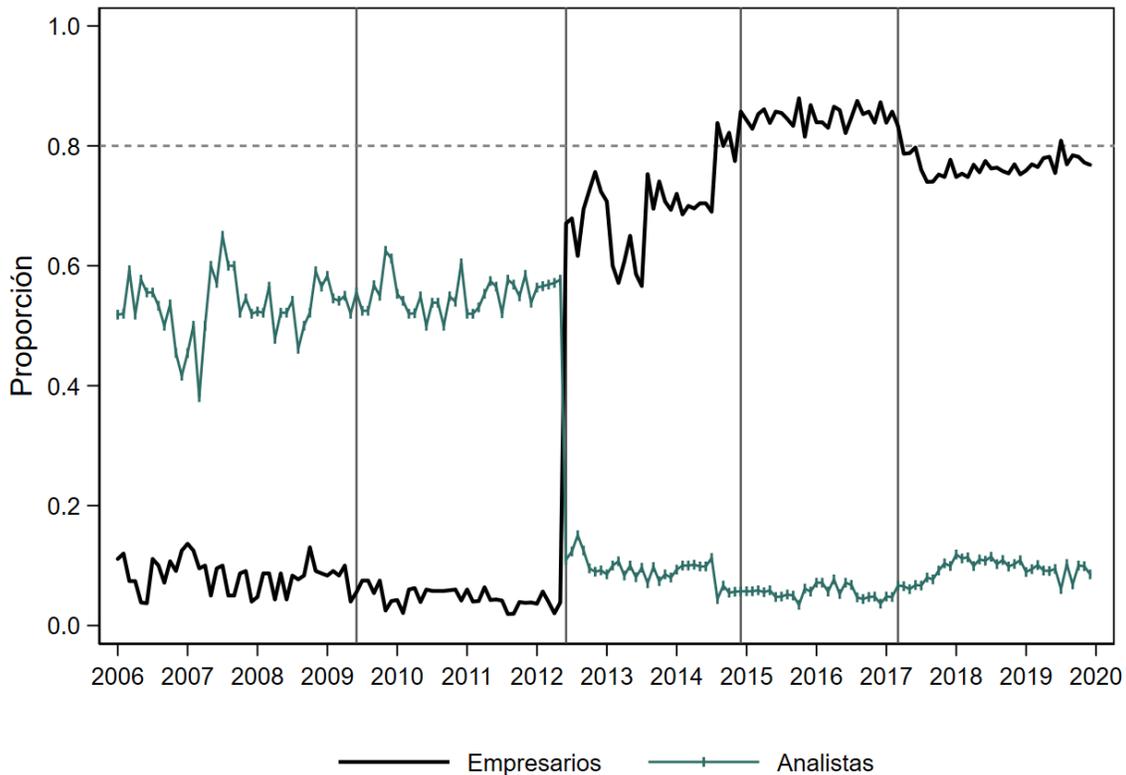
Gracias a esta reconstrucción, se hacen evidentes dos cambios metodológicos: uno en junio de 2012 y otro en diciembre de 2014. (Cuadro 2). Por otra parte, los cambios de diseño muestral en los años 2009 o 2017 son difíciles de visualizar con los patrones de entradas y salidas mostrados en el Cuadro 2.

² Los patrones que identifican la primera y la última respuesta, se basan en la expectativa de inflación. Este análisis no incluye las respuestas relacionadas con las expectativas de tipo de cambio.

En junio de 2009, participan 36 expertos, este tamaño de muestra es mayor al determinado inicialmente (Arias et. al, 2006), por eso se fechó un posible cambio de muestra en ese mes. Junio de 2012 registra el cambio más importante en la cantidad de respuestas por grupo de expertos. Se recibieron 80 respuestas, de las cuales 53 pertenecían al grupo de los empresarios. Un mes antes, la cantidad de respuestas fue de 52 y solo dos de los expertos eran empresarios.

El Gráfico 1 compara las proporciones de analistas financieros y empresas que respondieron cada mes. Las líneas verticales corresponden a junio 2009, junio 2012, diciembre 2014 y marzo 2017 y marcan posibles cambios metodológicos de la encuesta.

Gráfico 1. Encuesta mensual de expectativas de inflación y tipo de cambio del Banco Central de Costa Rica: Proporción de empresarios y analistas dentro de la muestra Enero, 2006 - Diciembre, 2019



Fuente: Elaboración propia con base en BCCR (2020)

Antes de junio de 2012, el grupo de analistas financieros representó entre el 40% y 65% de los informantes. Después de esa fecha, entre 60% y 80% de las respuestas provenían de las empresas. (Gráfico 1). De junio de 2012 en adelante, los diseños muestrales son probabilísticos y se usa, como marco muestral de empresas, un registro administrado por el BCCR.

En diciembre de 2014, fueron seleccionadas, como parte de la muestra, varias empresas que pertenecen a grupos empresariales, lo que modificó la relación entre empresario y empresa³.

El cambio metodológico más reciente se hizo en marzo de 2017, cuando se calculó un tamaño de muestra de 197 expertos o empresas que se distribuyeron por afijación proporcional de la siguiente forma: 4% consultores, 7% analistas financieros, 8% académicos y 80% empresas. A partir de esa fecha, la media de las expectativas está ponderada por estas proporciones poblacionales. (BCCR, 2020).

2.2 Análisis estadísticos

El análisis estadístico usa métodos univariados y de asociación lineal.

El análisis univariado tiene un importante componente visual, también incluye la determinación de valores atípicos y extremos y la construcción de dos series: una de medias sin atípicos y otra de medias sin extremos.

Además, se prueban las hipótesis de normalidad de las distribuciones mensuales y según estos resultados se usan pruebas paramétricas o no paramétricas para determinar si la media y la mediana son estadísticamente diferentes de cero.

El análisis de asociación lineal incluye pruebas de hipótesis bivariadas y diversos modelos de regresión lineal múltiple.

2.2.1 Análisis univariado

Se comparó la media (\bar{x}), junto con su intervalo de confianza, con la mediana (Me) y el rango intercuartil. Las gráficas permiten determinar los períodos durante los cuales la media es muy diferente a la mediana⁴.

³ La distinción entre expertos y empresas se da con los cambios metodológicos hechos a la Encuesta a partir de diciembre de 2014. Hay varios expertos respondiendo por la misma empresa o un solo empresario responde por varias empresas del mismo grupo empresarial. Sin embargo, en las siguientes secciones, el término informante se usará para referirse a empresas y a expertos.

⁴ El análisis visual usa el intervalo de confianza al 95% para establecer el límite superior e inferior de la media: $\bar{x} - \tau_{(0,975; n-1)} \frac{s_x}{\sqrt{n}} < \bar{x} < \bar{x} + \tau_{(0,975; n-1)} \frac{s_x}{\sqrt{n}}$. Los percentiles 25 y el 75 se usan como límite inferior y superior de la mediana: $P_{25} < Me < P_{75}$.

Una de las explicaciones teóricas del sesgo de la media es el efecto de los valores influyentes. Estos valores se reconocen con dos métricas: 1- aquellos que se encuentran a más/menos dos o tres desviaciones estándar de la media en distribuciones normales y 2- los que se encuentran a una distancia mayor de 1,5 o tres del rango intercuartil (RI), se usan el percentil 25 como referencia para el límite inferior y el percentil 75 para el límite superior ⁵.

Se hicieron pruebas de normalidad a las respuestas de los informantes para determinar cuán apropiado era usar uno u otro criterio. Las pruebas demostraron que las distribuciones de las expectativas de inflación, en 53 de 168 meses, no se ajustan a una distribución normal, de modo que se utilizó el criterio del rango intercuartil para identificar los valores atípicos.

Se definió un *valor atípicamente bajo* como aquel cuya expectativa de inflación es menor que $P_{25} - 1,5RI$ y un *valor atípicamente alto* cuando es mayor que $P_{75} + 1,5RI$. Un *valor en extremo bajo* es aquel inferior a $P_{25} - 3RI$ y un *valor en extremo alto* correspondería a una expectativa superior a $P_{75} + 3RI$.

La identificación de estos valores permite controlar la proporción de atípicos y extremos en los modelos de regresión y construir una media sin valores atípicos y otra sin valores extremos para comparar los resultados. Estas series no forman parte del conjunto de estadísticos que se publican de forma regular en el sitio de Indicadores Económicos del BCCR.

Se calculó la diferencia entre la media y mediana mensual ($\bar{x} - Me$) de las expectativas de inflación y, tras comprobar que tiene distribución normal, se usó la prueba T-Student para verificar si las distribuciones son estadísticamente diferentes.

La hipótesis nula es que la diferencia de las medias poblacionales es cero, contra la alternativa de que es mayor. $H_0: \mu_{media} - \mu_{mediana} = 0$; $H_1: \mu_{media} - \mu_{mediana} > 0$. El estadístico de esta prueba se distribuye como T-Student, ecuación (1):

$$(1) \quad \tau_c = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

Donde:

$$\bar{d} = \frac{\sum_{t=1}^{n=168} d_t}{n}, \quad s_d = \sqrt{\frac{\sum d_t - \bar{d}}{n-1}}, \quad d_t = \bar{x}_t - Me_t = Media_t - Mediana_t,$$

t es cada uno de los meses de la serie temporal y $n = 168$ meses.

⁵ El intervalo o rango intercuartil es la diferencia entre los percentiles 75 y 25. ($RI = P_{75} - P_{25}$)

2.2.2 Asociación, correlación y regresión lineal múltiple

Se usó la misma prueba T-Student para comparar si las diferencias entre las medias y las medianas persistían en los períodos en los que se fraccionó la serie⁶. Asimismo, se usó la correlación de Spearman para medir la similitud entre las medias (completa, sin extremos y sin atípicos) y la mediana.

Después de comprobar que las diferencias entre la media y la mediana son significativas, se usó la herramienta de regresión lineal para comprender mejor cuáles factores podrían estar asociados con estas diferencias.

Las variables explicadas son tres: 1- diferencias entre la media calculada con todas las respuestas y la mediana, 2- media que elimina aquellos valores atípicos en la distribución menos la mediana de expectativa de inflación y 3- diferencia entre media libre de valores extremos y la mediana mensual.

Las variables explicativas son: el valor de la meta de inflación, la expectativa media o mediana del mes anterior, la variación interanual del IPC del mes anterior, la proporción mensual de empresarios, la proporción de valores atípicos y la proporción de valores extremos. Adicionalmente, se incluyeron interacciones con la variable categórica (períodos) y el resto de variables continuas.

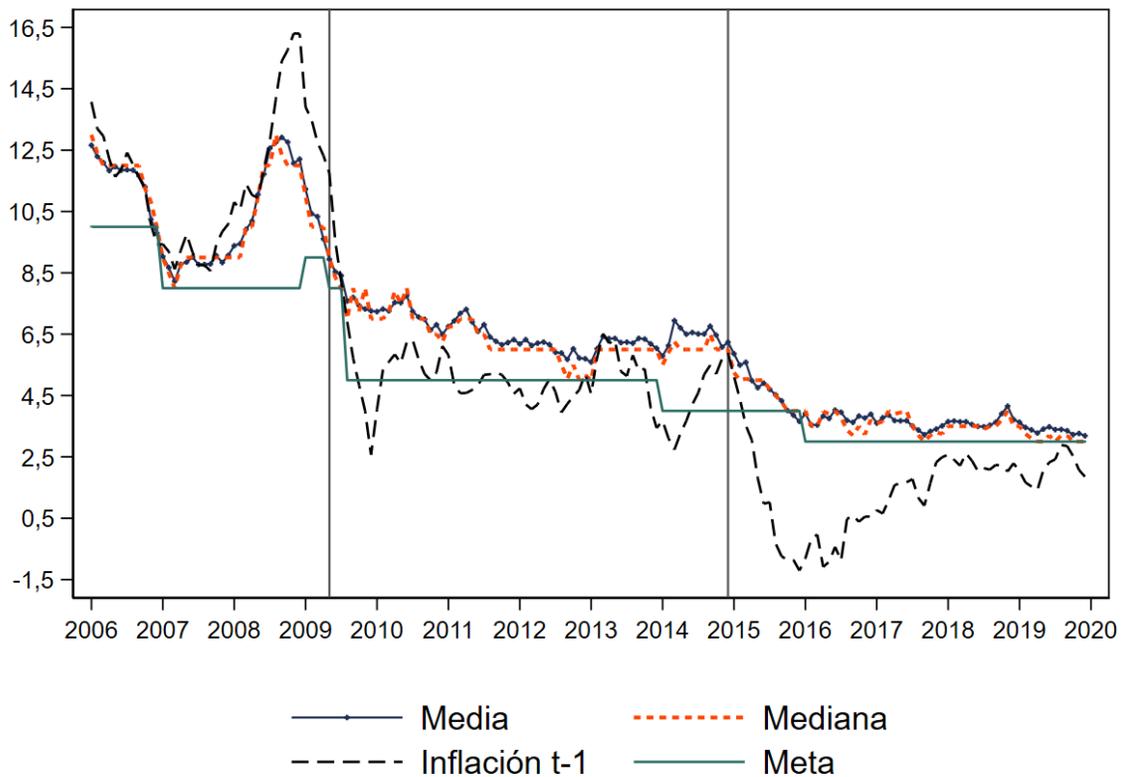
3. Resultados

El Gráfico 2 compara la evolución de la media y la mediana de las expectativas de inflación junto con la variación interanual del IPC con un mes de rezago (Inflación t-1) y la meta de inflación establecida por el Banco Central (Meta). Los ejes verticales corresponden a mayo de 2009 y a diciembre de 2014.

Cuando los informantes reciben la solicitud de cooperación por parte del Banco, el IPC publicado corresponde al mes anterior, por lo que al responder el cuestionario enviado en mayo, por citar un caso, los informantes podrían consultar la inflación de abril. También, pueden consultar la información de la meta de inflación a 12 meses plazo y por esa razón, el Gráfico 2 incluye estas dos series junto con la comparación de la media y la mediana.

⁶ La serie histórica se dividió en tres períodos: de enero de 2006 a abril de 2009, de mayo 2009 a noviembre de 2014 y de diciembre de 2014 en adelante. Estos periodos combinan los cambios de la muestra y los quiebres estructurales de la inflación.

Gráfico 2. Costa Rica: Series mensuales de la media y mediana de la expectativa de inflación, variación interanual del IPC del mes anterior y meta de inflación a 12 meses Enero, 2006 - Diciembre, 2019.



Fuente: Elaboración propia con base en INEC (2020) y BCCR (2019 y 2020)

Antes de mayo de 2009, la inflación observada fue mayor al 10% en la mayoría de los meses (29 de 40), las expectativas media y mediana eran prácticamente iguales y seguían las variaciones mensuales interanuales del IPC, en tanto que el valor de la meta de inflación siempre fue inferior a estas tres series.

De mayo de 2009 y hasta diciembre de 2014, la inflación no solo deja de ser de dos dígitos, sino que para 32 de 67 meses asumió valores en el rango de 2,5% a 5%, dichos valores no se observaban en el país desde finales de la década de los setenta. La meta disminuyó tres veces y la media superó la mediana durante la mayor parte del tiempo.

De diciembre de 2014 en adelante, la inflación observada presenta valores inferiores al 2% e incluso negativos, la meta vuelve a disminuir para permanecer en 3% pero nunca es inferior a la variación interanual del IPC. La expectativa mediana de inflación tiende a coincidir más con el valor de la meta y la media se mantiene por encima de ambas, en especial hacia el final de 2019. (Gráfico 2).

A lo largo del tiempo, la diferencia promedio entre la media y la mediana de las expectativas de inflación fue de 0,1 p.p. De enero de 2006 a abril de 2009, las medidas de posición de ambas series son muy similares y, en algunos casos, el estadístico de la distribución de las medias resulta inferior al estadístico de la distribución de las medianas. El promedio de las diferencias entre la media y la mediana es cero y ambas series se parecen más a la variación interanual del IPC que a la meta de inflación. (Cuadro 3).

Cuadro 3. Costa Rica: Estadísticas descriptivas de la meta de inflación, inflación del mes anterior, media y mediana de las expectativas de inflación y su diferencia, según período.

Período	Estadístico ¹	Meta	Inflación _{t-1}	Exp. Media	Exp. Mediana	$\bar{x} - Me$
Total (168 meses)	Media	5,2	5,3	6,6	6,4	0,1
	Desv. Estánd	2,2	4,2	2,7	2,8	0,3
	Mínimo	3,0	-1,2	3,2	3,0	-0,7
	Percentil ₂₅	3,0	2,2	3,8	4,0	0,0
	Mediana	5,0	4,7	6,3	6,0	0,2
	Percentil ₇₅	8,0	8,4	8,3	8,0	0,3
	Máximo	10,0	16,3	12,9	13,0	0,8
Ene, 2006 - Abr, 2009 (40 meses)	Media	8,7	11,6	10,6	10,6	0,0
	Desv. Estánd	0,9	2,2	1,5	1,5	0,3
	Mínimo	8,0	8,6	8,2	8,0	-0,6
	Percentil ₂₅	8,0	9,6	9,0	9,0	-0,2
	Mediana	8,0	11,5	10,4	10,4	0,0
	Percentil ₇₅	10,0	12,9	11,9	12,0	0,2
	Máximo	10,0	16,3	12,9	13,0	0,8
May, 2009 - Nov, 2014 (67 meses)	Media	5,0	5,1	6,7	6,4	0,2
	Desv. Estánd	0,8	1,4	0,7	0,8	0,3
	Mínimo	4,0	2,6	5,6	5,0	-0,7
	Percentil ₂₅	5,0	4,5	6,2	6,0	0,1
	Mediana	5,0	5,1	6,5	6,0	0,2
	Percentil ₇₅	5,0	5,7	7,2	7,0	0,4
	Máximo	8,0	11,8	8,9	9,0	0,7
Dic, 2014- Dic, 2019 (61 meses)	Media	3,2	1,5	3,9	3,7	0,1
	Desv. Estánd	0,4	1,5	0,7	0,7	0,2
	Mínimo	3,0	-1,2	3,2	3,0	-0,3
	Percentil ₂₅	3,0	0,6	3,5	3,3	0,0
	Mediana	3,0	1,7	3,7	3,5	0,1
	Percentil ₇₅	3,0	2,3	3,9	4,0	0,2
	Máximo	4,0	5,9	6,2	6,0	0,6

1/ Desv. Estánd= Desviación Estándar

Fuente: Elaboración propia con base en BCCR (2020)

Las brechas entre la media y la mediana son más amplias de mayo de 2009 a noviembre de 2014, el promedio de las diferencias en esos meses es de 0,2 p.p. Estas diferencias comienzan a disminuir después de diciembre de 2014.

Durante el último período, los estadísticos de la media y la mediana de las expectativas de inflación son más parecidos a los estadísticos de la meta de inflación que a los estadísticos de la variación interanual del IPC del mes anterior. (Cuadro 3).

El Cuadro 4 muestra los coeficientes de correlación de Spearman entre las series. La estadística bivariada demuestra que la media y mediana de las expectativas de inflación mantienen una asociación más fuerte con la meta de inflación que con la variación interanual de IPC del mes anterior y este grado de asociación varía con el tiempo.

Cuadro 4. Costa Rica: Correlación de Spearman entre la meta de inflación, inflación del mes anterior, media y mediana de las expectativas de inflación, según período.

Período	Variables ¹	Meta	Inflación _{t-1}	Exp. Media
Total (168 meses)	Meta	1		
	Inflación _{t-1}	0,9047 ***	1	
	Exp. Media	0,9301 ***	0,8946 ***	1
	Exp. Mediana	0,9432 ***	0,8984 ***	0,9917 ***
Ene, 2006 - Abr, 2009 (40 meses)	Meta	1		
	Inflación _{t-1}	0,2994 *	1	
	Exp. Media	0,4086 ***	0,8995 ***	1
	Exp. Mediana	0,5020 ***	0,8352 ***	0,9637 ***
May, 2009 - Nov, 2014 (67 meses)	Meta	1		
	Inflación _{t-1}	0,4043 ***	1	
	Exp. Media	0,2148 *	0,3728 ***	1
	Exp. Mediana	0,3627 ***	0,4212 ***	0,9424 ***
Dic, 2014- Dic, 2019 (61 meses)	Meta	1		
	Inflación _{t-1}	-0,0045	1	
	Exp. Media	0,6480 ***	-0,1574	1
	Exp. Mediana	0,6981 ***	-0,1568	0,9198 ***

1/ Exp.= Expectativa;

* $P(\tau_T \geq \tau_c) < 0,1$

** $P(\tau_T \geq \tau_c) < 0,05$

*** $P(\tau_T \geq \tau_c) < 0,01$

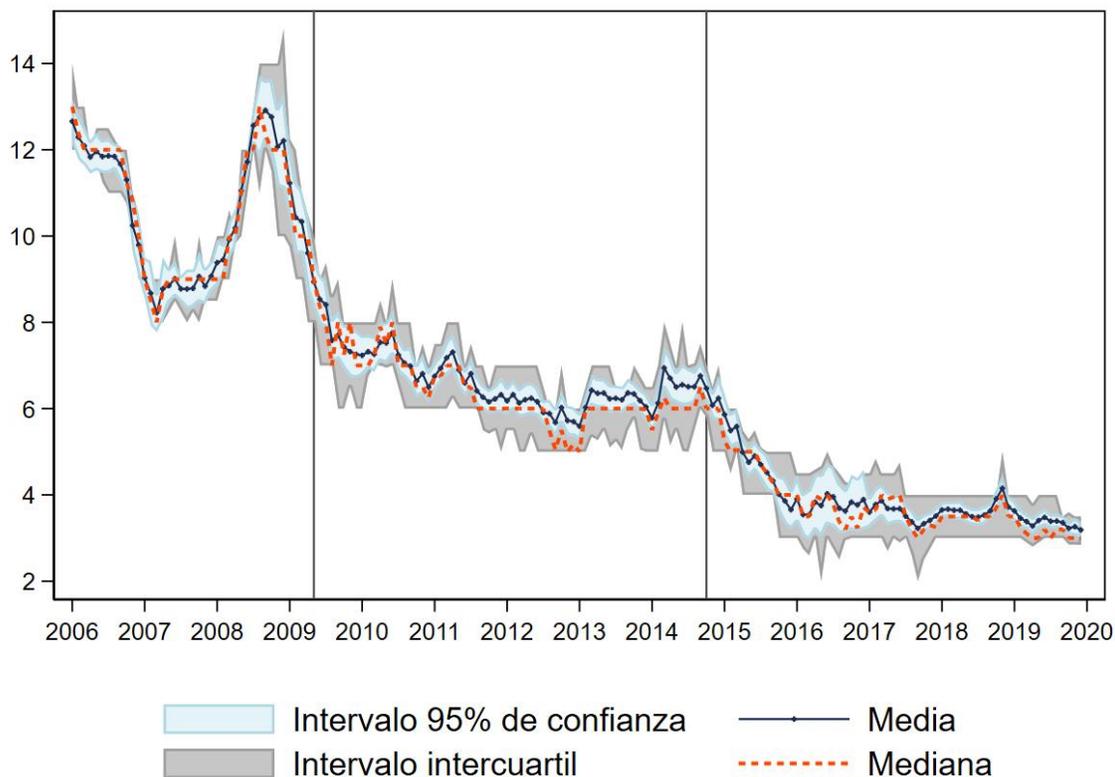
Fuente: Elaboración propia con base en BCCR (2020)

Antes de mayo de 2009, tanto la media, como la mediana tenían una asociación moderada y altamente significativa con la meta y una asociación alta y también muy significativa con la inflación. De mayo de 2009 a noviembre de 2014, la asociación lineal entre la media y la meta y entre la mediana y la meta disminuyen a sus niveles más bajos, pero conservan la significancia. Algo similar ocurre con la inflación. (Cuadro 4).

De diciembre de 2014 en adelante, la media y la mediana se disocian de la inflación y aumenta el valor de los coeficientes de correlación de cada una de estas variables con la meta. (Cuadro 4).

Al graficar las series de la media y la mediana con sus rangos de variabilidad se logran identificar varios meses, durante los cuales la mediana es menor al límite inferior del intervalo de confianza de la media. (Gráfico 3). Esas diferencias comienzan a ser más frecuentes a partir de junio de 2012, cuando se incrementa el número de empresarios que participan en la Encuesta. Las diferencias desaparecen entre el 2015 y 2018 y reaparecen de forma persistente en 2019.

Gráfico 3. Costa Rica: Series mensuales de la media y la mediana de la expectativa de inflación a 12 meses y su rango de variación. Enero, 2006 - Diciembre, 2019



Fuente: Elaboración propia con base en BCCR (2020)

De las evidencias, hasta el momento presentadas, la más importante es que, durante el período de análisis, el promedio de las diferencias entre la media y la mediana de las expectativas de inflación es 0,1 p.p. y dicho promedio varía dependiendo de los cambios en el diseño de la encuesta o de los quiebres estructurales en la inflación observada.

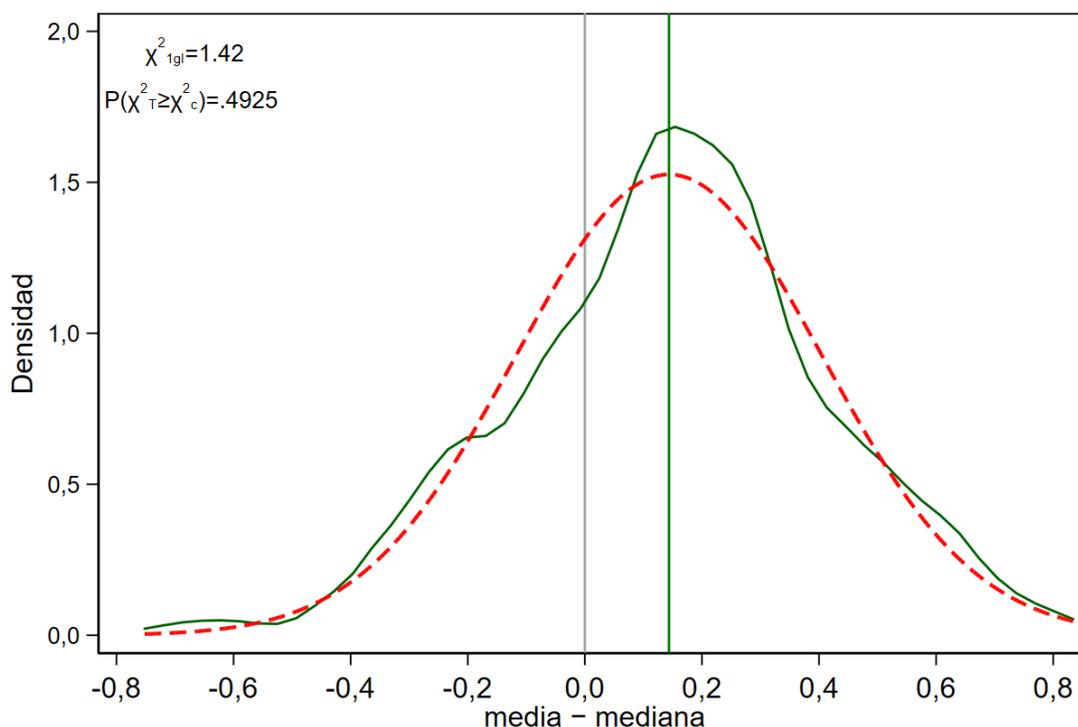
A pesar de estas diferencias, las dos series parecen medir lo mismo; su asociación lineal es 0,99 y, según el período, la media y la mediana de las expectativas se asocian más con la meta de inflación propuesta por el Banco Central que con el indicador de referencia dentro de la pregunta (la variación interanual del IPC del mes anterior).

3.1 Diferencias significativas entre la media y la mediana

El Gráfico 4 muestra la función de densidad de las diferencias entre la media y la mediana y la compara con una distribución normal (línea punteada). La distribución de las diferencias oscila entre un mínimo de -0,68 p.p. y un máximo de 0,76 p.p., es unimodal y está centrada en 0,14 p.p. con una desviación estándar de 0,26 p.p. El coeficiente de asimetría es -0,19 y el coeficiente de curtosis: 3,11.

La razón de asimetría a curtosis, como prueba de normalidad, no aporta suficiente evidencia estadística para rechazar la hipótesis nula de normalidad. El estadístico de esta prueba se distribuye como Chi-cuadrado con un grado de libertad ($\chi^2_{1gl} = 1,42$). La probabilidad de encontrar ese valor en la distribución teórica (χ^2_T) o uno más extremo $P(\chi^2_c \geq \chi^2_T)$ es casi de 0,5. (Gráfico 4).

Gráfico 4. Distribución de las diferencias entre las series de la media y la mediana de las expectativas de inflación
(n=168 meses: Enero, 2006 a diciembre, 2019)



Fuente: Elaboración propia con base en BCCR (2020)

La prueba de hipótesis para verificar si, en promedio, la diferencia entre la media y la mediana es cero, es la T-Student. La hipótesis nula es: $H_0: \mu_{media} - \mu_{mediana} \leq 0$ y, como se puede comprobar en el Cuadro 5, al 5% de significancia hay evidencia estadística para afirmar que el sesgo de la media es positivo.

Cuadro 5. Resultado de la prueba de hipótesis T-Student para muestras pareadas
Diferencia entre la media y la mediana
(n=168 meses: Enero, 2006 a diciembre, 2019)

Distribución	Promedio	Error estándar	Desviación estándar	Intervalo al 95% de confianza*	
Media	6,58	0,21	2,73	6,17	7,00
Mediana	6,44	0,21	2,78	6,01	6,86
Diferencia	0,14	0,02	0,26	0,10	0,18

$$\tau_{c(0,975;167gl)} = 7,127 \rightarrow P(\tau_T \geq \tau_c) = 0,000$$

Fuente: Elaboración propia con base en BCCR (2020)

3.2 Influencia de los valores atípicos y extremos

La distribución de la media mensual de las expectativas de inflación presentó 2% de valores atípicos bajos y el porcentaje de atípicos altos fue el doble. Los valores extremos bajos representaron 0,3% de la distribución y el porcentaje de valores en extremo altos resultó mayor al 1%. Una distribución normal, por azar, tendría a lo sumo 5% de valores atípicos: la mitad bajos y el otro 2,5% altos, de dichos valores, alrededor de 0,5% correspondería a valores extremos bajos y 0,5% a extremos altos. Así, la distribución de la media de las expectativas de inflación tiene más valores atípicos y extremos altos que una distribución normal⁷.

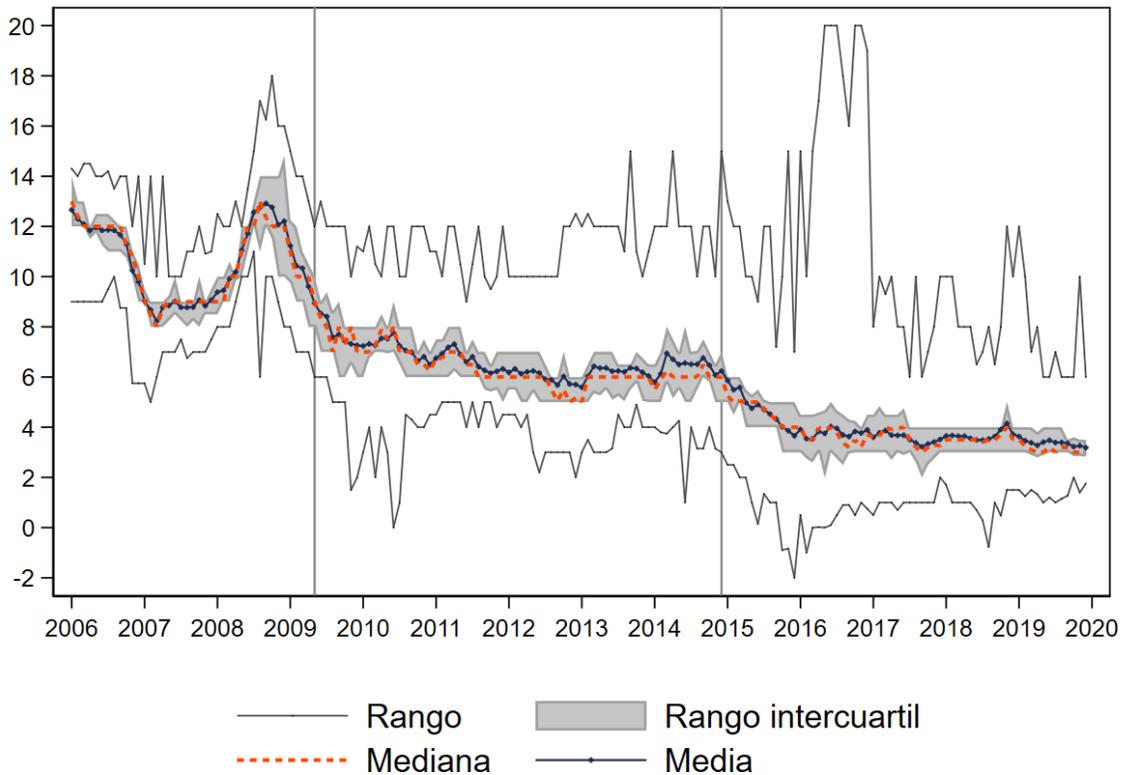
El Gráfico 5 muestra la media y mediana de las expectativas, el rango intercuartil y los valores mínimos y máximos que, usualmente, también son valores atípicos o extremos. La presencia de valores atípicos o extremos no implica sesgos en la media, pues para que el sesgo positivo o negativo se presente hay que estudiar la distribución mensual de las respuestas.

A manera de ejemplo, 2016, al menos un informante esperaba que la inflación alcanzara el 20% un año después, este valor se mencionó en mayo, junio, julio, octubre y noviembre. (Gráfico 5). En mayo de 2016, la media fue menor que la mediana porque 30 de 57 informantes reportaron una expectativa de inflación de 5% o menos, incluso hubo quienes esperaron una inflación de 0%. De esta manera el valor 20%, no sesga la media. En cambio,

⁷ Los *valores atípicos* son aquellos que se encuentran a $P_{25}/P_{75} \mp 1,5RI$ y los *valores extremos* se definieron como $P_{25}/P_{75} \mp 3RI$. La definición detallada se puede consultar en el punto 2.2.1, página 6.

para noviembre de ese mismo año, la expectativa de dos informantes volvió a ser 20% y la diferencia entre la media y la mediana fue 0,6 p.p. porque más del 95% de los informantes esperaba que la inflación estuviera entre 1% y 5%.

Gráfico 5. Costa Rica: Series mensuales de la media y la mediana de las expectativas de inflación a 12 meses, rango intercuartil y rango. Enero, 2006 - Diciembre, 2019



Fuente: Elaboración propia con base en BCCR (2020)

Se calculó una media libre de valores atípicos y otra libre de valores extremos para tratar de controlar su efecto. Las estadísticas descriptivas de cada distribución y las pruebas de asociación lineal muestran que la media sin valores atípicos se parece a la mediana y la media sin valores extremos es similar a la media calculada con todas las respuestas.

El Cuadro 10 permite apreciar cómo las distribuciones coinciden en las medidas de posición. El Cuadro 11 muestra los resultados de la asociación lineal entre las cuatro series. Ambos cuadros se pueden consultar los Anexos.

Las diferencias entre cada una de estas medias y la mediana se ajustan también a una distribución normal (Ver Gráfico 6 en los Anexos), por eso se aplica la misma prueba T-Student y se concluye que las diferencias entre cada media y la mediana persisten. (Cuadro 6).

Cuadro 6. Resultado de la prueba de hipótesis T-Student para muestras pareadas
Diferencia entre cada media y la mediana
(n=168 meses: Enero, 2006 a diciembre, 2019)

Variable	Promedio	Error estándar	$t_{c(0,975;167gl)}$	Valor de P ¹	Intervalo de Confianza	
$\bar{x} - Me$	0,1437	0,0202	7,1271	0,0000	0,1039	0,1836
$\bar{x}_{sin\ atípicos} - Me$	0,0511	0,0168	3,0445	0,0027	0,0180	0,0842
$\bar{x}_{sin\ extremos} - Me$	0,0919	0,0193	4,7636	0,0000	0,0538	0,1301

1/ Valor de P es la probabilidad de encontrar el estadístico calculado (τ_c) en la distribución teórica (T-Student) o uno más extremo. Se denota usualmente como $P(\tau_T \geq \tau_c)$ o $P > |\tau|$.

Fuente: Elaboración propia con base en BCCR (2020)

Las medias sin atípicos y sin extremos usan una mayor cantidad de respuestas que la mediana. La mediana a lo sumo usa 36 respuestas, en tanto que para calcular la media sin atípicos, como máximo, se eliminan 21 observaciones en un mes y en el caso de la media sin extremos, a lo sumo pierde nueve respuestas.

Debido a que la cantidad de valores atípicos y extremos depende de la distribución mensual, 27 meses no pierden ninguna respuestas, lo más frecuente es que se eliminen entre una y tres respuestas⁸.

3.3 Variables que explican la diferencia entre la media y la mediana

Esta sección comprueba que los valores atípicos no son el único factor que explica la diferencia positiva entre la media y la mediana. Se usan como variables explicadas las diferencias entre la media y la mediana y como variables explicativas: el valor de la meta, la variación interanual del IPC con un mes de rezago, la expectativa media o mediana del mes anterior, la proporción de empresarios dentro de la muestra y la proporción de valores atípicos. Además, se incluyó la variable categórica que divide la serie en tres períodos. La base para la comparación fue el período anterior a mayo 2009.

El Cuadro 7 muestra la estadística descriptiva de las variables incluidas en los modelos de regresión.

⁸ Aunque la Encuesta es longitudinal, para calcular las medias sin atípicos y sin extremos no se elimina del análisis a los informantes que las reportan. Así un informante que en un mes dio un valor atípicamente bajo, se excluye del cálculo de la media sin atípicos de ese mes, pero si al mes siguiente su respuesta no fue atípica, ese informante vuelve a la muestra.

Cuadro 7. Costa Rica: Estadística descriptiva de las variables incluidas en los modelos de regresión lineal múltiple
(n=168 meses: Enero, 2006 a diciembre, 2019)

VARIABLES ¹	Media	Desv. Estánd ²	Simetría	Curtosis	Mínimo	Mediana	Máximo
$\bar{x} - Me$	0,14	0,26	-0,19	3,11	-0,68	0,17	0,76
$\bar{x}_{sin\ at\ ípicos} - Me$	0,05	0,22	0,01	2,54	-0,44	0,06	0,56
$\bar{x}_{sin\ extremos} - Me$	0,09	0,25	-0,06	2,73	-0,53	0,11	0,76
Meta	5,22	2,21	0,86	2,51	3,00	5,00	10,00
Inflación	5,35	4,18	0,72	2,76	-1,20	4,70	16,30
Empresario ³	0,45	0,36	-0,11	1,10	0,02	0,66	0,88
Exp. Media (\bar{x})	6,58	2,73	0,67	2,56	3,19	6,25	12,91
Exp $\bar{x}_{sin\ at\ ípicos}$	6,49	2,79	0,70	2,56	3,03	6,16	13,02
Exp $\bar{x}_{sin\ extremos}$	6,53	2,76	0,67	2,53	3,15	6,23	12,91
Exp. Mediana (Me)	6,44	2,78	0,72	2,56	3,00	6,00	13,00
Atípicos bajos ³	0,02	0,03	2,95	14,48	0,00	0,00	0,22
Atípicos altos ³	0,04	0,03	1,65	7,26	0,00	0,03	0,22

1/ Me = mediana; \bar{x} =media; Exp.= Expectativa

2/ Desv. Estánd = Desviación estándar

3/ Proporciones mensuales

Fuente: Elaboración propia con base en BCCR (2020)

Varios elementos se pueden enfatizar con base en el cuadro anterior. El primero es que las estadísticas descriptivas de las diferencias y de las tres medias (media completa, media sin atípicos y media sin extremos) y la mediana, tienen valores muy similares. El segundo es que con excepción de las proporciones de valores atípicos, todas las variables tienden a ser simétricas. Además, todas las variables tienen un apuntamiento o curtosis superior al apuntamiento de la distribución normal. Finalmente, las varianzas de las diferencias son iguales, lo mismo que las varianzas de las tres medias y la mediana⁹.

El Cuadro 8 presenta el coeficiente, el error estándar (se) y el valor de P ($P > |\tau|$) de tres modelos, donde cada variable explicada (Y) es la diferencia entre la media y mediana, tal y como se especifica en la siguiente ecuación (2). Se llegó a esta selección de variables después de comparar varios modelos.

⁹ Las pruebas de hipótesis respectivas corroboran estas afirmaciones. Dichas pruebas puede ser solicitadas a la autora.

$$(2) \quad Y = \alpha + \beta(\text{meta}) + \gamma_1(\text{inflación}_{t-1}) + \delta_2(\text{período}_2) + \delta_3(\text{período}_3) + \gamma_2(\text{inflación}_{t-1}) * (\text{período}_2) + \gamma_3(\text{inflación}_{t-1}) * (\text{período}_3) + \kappa(\text{empresario}) + \lambda_1(\text{atípicos bajos}) + \lambda_2(\text{atípicos altos}) + \varepsilon$$

Los resultados del primer modelo indican que las diferencias entre la media (con todas las respuestas) y la mediana se amplían cuando aumenta la proporción de empresarios que participan en la Encuesta y cuando aumenta la proporción de valores atípicamente altos y disminuyen cuando la meta de inflación se reduce y conforme aumenta la proporción de valores atípicamente bajos. (Cuadro 8).

De mayo de 2009 a noviembre de 2014, la diferencia estimada entre la media y la mediana es en promedio 0,5 p.p. mayor que la diferencia entre la media y la mediana de enero de 2006 a abril de 2009. Este es el único período durante el cual la diferencia es significativa. (Cuadro 8: Modelo 1).

La variación interanual del IPC solo contribuye a explicar las diferencias entre cada media y la mediana durante enero de 2006 a abril de 2009, luego su efecto desaparece.

Al usar la media sin valores atípicos a fin de calcular las diferencias, la proporción mensual de valores atípicos pierde significancia estadística, lo mismo ocurre con las variables dicotómicas de los períodos y con la participación de los empresarios. (Cuadro 8: Modelo 2).

En lo que respecta al caso de las diferencias entre la media libre de valores extremos y la mediana, los valores atípicamente altos siguen explicando el sesgo positivo de la media y desaparecen los otros efectos. (Cuadro 8: Modelo 3).

Cuadro 8. Resultados de las regresiones lineales múltiples
(n=168 meses: Enero, 2006 a diciembre, 2019)

Variable	Modelo 1			Modelo 2			Modelo 3		
		$\hat{Y} = \bar{x} - Me$		$\hat{Y} = \bar{x}_{\text{sin atípicos}} - Me$		$\hat{Y} = \bar{x}_{\text{sin extremos}} - Me$			
Meta	$\hat{\beta}$	-0,0738		-0,0674		-0,0742			
	$se_{\hat{\beta}}$	0,0326		0,0301		0,0317			
	$P > \tau $	0,0249		0,0267		0,0204			
Inflación_{t-1}	$\hat{\gamma}_1$	0,0612		0,0539		0,0581			
	$se_{\hat{\gamma}_1}$	0,0184		0,0174		0,0193			
	$P > \tau $	0,0011		0,0023		0,0030			
Períodos									
<u>Período 2:</u>	$\hat{\delta}_2$	0,4892		0,4122		0,4731			
May, 2009 a	$se_{\hat{\delta}_2}$	0,2506		0,2564		0,2566			
Nov, 2014	$P > \tau $	0,0527		0,1099		0,0671			
<u>Período 3:</u>	$\hat{\delta}_3$	0,1699		0,0682		0,0689			
Dic, 2014 a	$se_{\hat{\delta}_3}$	0,2360		0,2424		0,2495			
Dic, 2019	$P > \tau $	0,4726		0,7789		0,7829			
Interacción: período e inflación_{t-1}									
<u>Período 2</u>	$\hat{\gamma}_2$	-0,0485		-0,0441		-0,0476			
May, 2009 a	$se_{\hat{\gamma}_2}$	0,0267		0,0259		0,0267			
Nov, 2014	$P > \tau $	0,0712		0,0901		0,0771			
<u>Período 3</u>	$\hat{\gamma}_3$	-0,0184		0,0040		0,0027			
Dic, 2014 a	$se_{\hat{\gamma}_3}$	0,0249		0,0254		0,0276			
Dic, 2019	$P > \tau $	0,4613		0,8737		0,9227			
Empresario¹	$\hat{\kappa}$	0,1802		0,1132		0,1185			
	$se_{\hat{\kappa}}$	0,0790		0,0763		0,0749			
	$P > \tau $	0,0239		0,1399		0,1157			
Atípico bajo¹	$\hat{\lambda}_1$	-1,8380		1,0924		-1,3842			
	$se_{\hat{\lambda}_1}$	0,7457		0,6534		0,7884			
	$P > \tau $	0,0148		0,0966		0,0811			
Atípico alto¹	$\hat{\lambda}_2$	2,8013		-0,5208		1,5004			
	$se_{\hat{\lambda}_2}$	0,7125		0,6419		0,6643			
	$P > \tau $	0,0001		0,4184		0,0253			
Constante	$\hat{\alpha}$	-0,0961		-0,0375		-0,0325			
	$se_{\hat{\alpha}}$	0,3025		0,3075		0,3096			
	$P > \tau $	0,7512		0,9030		0,9164			

se= error estándar, por su abreviatura en inglés

1/ Proporciones mensuales

Fuente: Elaboración propia con base en BCCR (2020)

Los resultados de las pruebas de diagnóstico señalan que no hay suficiente evidencia estadística para rechazar la hipótesis de normalidad en la distribución de los residuos. Ninguna estimación cumple con el supuesto de homocedasticidad y, al 5% de significancia, todos los modelos tienen autocorrelación serial. (Cuadro 9). Si la significancia fuera 1%, la hipótesis nula de autocorrelación serial solo se rechazaría en el primer modelo.

Cuadro 9. Resultados de la evaluación y pruebas de diagnóstico de los modelos (n=168 meses: Enero, 2006 a diciembre, 2019)

Estadístico	Modelo 1	Modelo 2	Modelo 3
	$\hat{Y} = \bar{x} - Me$	$\bar{x}_{\text{sin atípicos}} - Me$	$\bar{x}_{\text{sin extremos}} - Me$
Estadístico F	13,9	5,5	12,3
Estadístico R ²	0,4	0,2	0,4
ECM ^{1/}	0,2	0,2	0,2
Verosimilitud ^{2/}	29,5	39,8	33,1
AIC ^{3/}	-39,1	-59,6	-46,2
BIC ^{4/}	-7,8	-28,4	-15,0
Heterocedasticidad ^{5/}	94,0	103,4	96,3
$P > \chi^2_{52}$	0,0003	0,0000	0,0002
ARCH ^{6/}	3,7	2,2	4,5
$P > \chi^2_1$	0,0535	0,1395	0,0342
Autocorrelación ^{7/}	8,1	6,4	4,1
$P > \chi^2_1$	0,0043	0,0117	0,0427
Normalidad ^{8/}	-0,2	-1,8	-0,5
$P > z $	0,59743	0,96353	0,67679

1/ Error cuadrático medio

2/ Logaritmo de la verosimilitud

3/ Criterio de información de Akaike. AIC por sus siglas en inglés

4/ Criterio de información Bayesiano de Schwarz. BIC por sus siglas en inglés

5/ Prueba de heterocedasticidad de Cameron y Trivedi (1990)

6/ Autocorrelación serial condicionada por heterocedasticidad. ARCH por sus siglas en inglés

7/ Prueba de alternativa de Durbin que no requiere exogeneidad estricta de los regresores.

8/ Test de Shapiro-Wilk

Fuente: Elaboración propia con base en BCCR (2020)

La estimación de los errores usó la corrección de Huber y White para controlar la heterocedasticidad. Si se quiere controlar la autocorrelación, se puede usar la diferencia entre la media y la mediana con uno, dos o tres rezagos como variables explicativas. La cantidad de rezagos depende de los valores atípicos y extremos. El modelo 1, requiere al menos 3 rezagos, el modelo 2 sin atípicos ajusta bien con un rezago y el modelo 3 sin extremos, requeriría de dos rezagos¹⁰.

¹⁰ Los resultados de los modelos con rezagos pueden ser solicitados a la autora.

De esta forma, la selección entre las series de medias o la mediana depende de los criterios que se prefieran para evaluar el modelo y de los resultados de las pruebas de diagnóstico posteriores a la estimación.

4. Conclusiones y recomendaciones

El promedio de las diferencias entre la media y la mediana de las expectativas de inflación fue mayor a 0,1 p.p, de enero de 2006 a diciembre de 2019, esta diferencia es estadísticamente significativa y varía según sean los cambios metodológicos de la Encuesta o de los quiebres estructurales en la inflación observada.

El sesgo positivo de la media es explicado, en parte, por la influencia de valores atípicamente altos que parecen responder al proceso de formación de las expectativas. Aunque este artículo no estudia a los individuos, ni conoce la información que utilizan para hacer sus pronósticos, sí encuentra que la brecha entre la media y la mediana disminuye cuando el Banco Central disminuye la meta de inflación de 12 meses. Asimismo, comprobó que, a inicios del periodo (enero 2006 a abril de 2009), la diferencia entre la media y la mediana aumentaba si la variación interanual del IPC aumentaba.

A partir del quiebre estructural de mayo del 2009 y con la participación de un mayor número de empresarios en la encuesta, la media comienza a distanciarse de la mediana. Durante ese período, la inflación bajó rápidamente y los informantes podrían haber pensado que la caída en la inflación era temporal y quizás por eso esperaban inflaciones altas. Los empresarios fueron más pesimistas que el resto de expertos y sus expectativas fueron extremas para las circunstancias observadas. Con el cambio de muestra de finales de 2014, aumenta la proporción de informantes cuya expectativa de inflación es un valor cercano o igual a la meta de inflación, por ende la diferencia entre la media y la mediana tiende a disminuir.

Si bien la media se ve afectada por valores atípicamente altos y la mediana no, estos podrían aportar información importante en tiempos de mucha incertidumbre. Por ejemplo, de marzo a junio de 2020, el 10% de los informantes espera una inflación superior al 5% para los próximos 12 meses, en tanto que la variación interanual del IPC continúa por debajo del 3%. Estos pronósticos al parecer extremos podrían ser una señal útil para las decisiones que se tomarán en los próximos meses.

Los hallazgos también señalan que la medias: completa, sin atípicos, sin extremos y la mediana, tienen correlaciones bivariadas superiores a 99% de modo que las cuatro series miden lo mismo. Los resultados de las regresiones múltiples, donde estas variables se utilizan para explicar la variabilidad de las diferencias, llegan a la misma conclusión.

Cuando las series de media y mediana se usan como variables explicadas, se comprueba que los coeficientes tienen los mismos signos y en algunos casos hasta el mismo valor. También, los errores estándar son similares.

Cada medida de posición tiene sus propias ventajas y debilidades. Las medias completa, sin atípicos y sin extremos, son mayores que la mediana pero usan más respuestas de los informantes. La mediana a lo sumo resume la opinión de 36 informantes y una o dos respuestas serían suficientes para calcularla. La media sin valores atípicos es muy similar a la mediana y como máximo se eliminan 21 respuestas en un mes.

La media sin atípicos y la media sin extremos fueron también de utilidad en la estimación de los modelos, pero deben ser estimadas con los microdatos. Entonces, elegir una u otra serie para su inclusión como variable dependiente o independiente en un modelo de regresión debería responder a los objetivos específicos de cada estudio. Otras investigaciones también han probado con medidas alternativas de las expectativas de inflación. Trabajos futuros deberían aprovechar la riqueza de los datos tipo panel y evaluar los errores de pronóstico centrados en la media o en la mediana o usar un enfoque bayesiano.

Las recomendaciones pueden resumirse en los siguientes puntos:

1. Si se desea usar una estimación para un mes específico, se puede usar la mediana siempre que la diferencia (media-mediana) exceda 0,1 p.p.
2. En informes técnicos se recomienda valorar en cuáles ocasiones conviene comentar la media y la mediana, la desviación estándar, los deciles 1 y 9 o limitarse al uso de solo uno de estos estimadores. A su vez, se recomienda evaluar el beneficio que puede aportar a los usuarios de la información, la publicación de los cuartiles 1 y 3 y los porcentajes de atípicos y extremos.
3. Si se debe usar alguna de las series temporales de expectativas de inflación, se aconseja modelar la media y la mediana y probar cuál brinda un mejor ajuste.

5. Referencias

Arias Cubillo, E., Durán Víquez, R., & Kikut Valverde, A. (2006). *Diseño de la Encuesta Mensual de Expectativas de Inflación*. Nota Técnica, DIE-04-2006-NT, Banco Central de Costa Rica, Departamento de Investigaciones Económicas, División Económica, San José.

Banco Central de Costa Rica. (Enero de 2020). Encuesta de Expectativas de Inflación y Tipo de cambio. *Área de Encuestas Económicas, Departamento de Gestión de la Información Económica*(División Económica, Banco Central de Costa Rica). San José, San José, Costa Rica.

Instituto Nacional de Estadística y Censos. (Enero de 2020). *Indicadores Económico, Índice de Precios, IPC*. Obtenido de Índice de precios al consumidor (IPC). Niveles y Variaciones:

<https://gee.bccr.fi.cr/indicadoreseconomicos/Cuadros/frmVerCatCuadro.aspx?idioma=1&CodCuadro=%202732>

Stata Corp LP. (6 de junio de 2017). Stata/SE (version 15.00 for Window, 64 bits). College Station, Texas, USA.

Torres, C. (2012). *Costa Rica: Determinación de cambios estructurales en el nivel de la tasa de inflación: período 1997-2011*. Banco Central de Costa Rica, Departamento de Investigación Económica, División Económica. San José, Costa Rica: Documento de Trabajo 02-2012.

6. Anexos

Cuadro 10. Costa Rica: Estadísticas descriptivas de la media y mediana de las expectativas de inflación, según período

Período	Estadístico	Media	Mediana	Media sin atípicos	Media sin atípicos
Total (168 meses)	Media	6,6	6,4	6,5	6,5
	Desv. Estánd	2,7	2,8	2,8	2,8
	Mínimo	3,2	3,0	3,0	3,2
	Percentil ₂₅	3,8	4,0	3,7	3,7
	Mediana	6,3	6,0	6,2	6,2
	Percentil ₇₅	8,3	8,0	8,3	8,3
	Máximo	12,9	13,0	13,0	12,9
Ene, 2006 - Abr, 2009 (40 meses)	Media	10,6	10,6	10,6	10,6
	Desv. Estánd	1,5	1,5	1,5	1,5
	Mínimo	8,2	8,0	8,3	8,2
	Percentil ₂₅	9,0	9,0	9,1	9,1
	Mediana	10,4	10,4	10,5	10,4
	Percentil ₇₅	11,9	12,0	12,0	12,0
May, 2009 - Nov, 2014 (67 meses)	Media	6,7	6,4	6,5	6,6
	Desv. Estánd	0,7	0,8	0,8	0,7
	Mínimo	5,6	5,0	5,3	5,5
	Percentil ₂₅	6,2	6,0	6,0	6,2
	Mediana	6,5	6,0	6,3	6,4
	Percentil ₇₅	7,2	7,0	7,1	7,2
Dic, 2014- Dic, 2019 (61 meses)	Media	3,9	3,7	3,7	3,8
	Desv. Estánd	0,7	0,7	0,6	0,6
	Mínimo	3,2	3,0	3,0	3,2
	Percentil ₂₅	3,5	3,3	3,4	3,4
	Mediana	3,7	3,5	3,5	3,5
	Percentil ₇₅	3,9	4,0	3,7	3,7
	Máximo	6,2	6,0	5,9	6,0

Fuente: Elaboración propia con base en BCCR (2020)

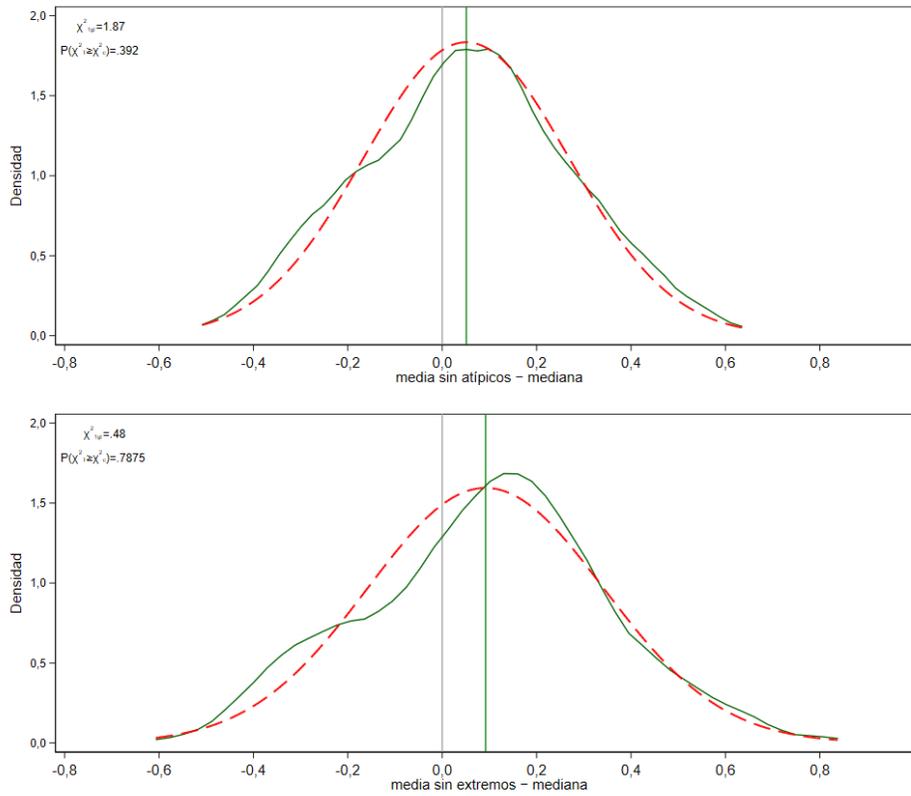
Cuadro 11. Costa Rica: Matriz de correlación bivariada de Spearman entre la media y mediana de las expectativas de inflación, según período

Período	Variables	Media	Mediana	Media sin atípicos
Total (168 meses)	Media	1		
	Mediana	0,9917	1	
	Medias sin atípicos	0,9941	0,9923	1
	Media sin extremos	0,9948	0,9920	0,9967
Ene, 2006 - Abr, 2009 (40 meses)	Media	1		
	Mediana	0,9637	1	
	Medias sin atípicos	0,9942	0,9733	1
	Media sin extremos	0,9970	0,9667	0,9949
May, 2009 - Nov, 2014 (67 meses)	Media	1		
	Mediana	0,9424	1	
	Medias sin atípicos	0,9798	0,9429	1
	Media sin extremos	0,9892	0,9419	0,9888
Dic, 2014- Dic, 2019 (61 meses)	Media	1		
	Mediana	0,9198	1	
	Medias sin atípicos	0,9110	0,9308	1
	Media sin extremos	0,9137	0,9285	0,9478

Todos los coeficientes de correlación tienen un significancia inferior al 1%.

Fuente: Elaboración propia con base en BCCR (2020)

Gráfico 6. Distribución de las diferencias entre la media sin atípicos y la mediana y entre la media sin extremos y la mediana de las expectativas de inflación (n=168 meses: Enero, 2006 a diciembre, 2019)



Fuente: Elaboración propia con base en BCCR (2020)