

---

# Proceso de simulación para evaluar el tamaño de muestra de la Encuesta de Acceso y Uso de los Servicios de Telecomunicaciones en Costa Rica

## Simulation Process to Evaluate Sample Size for the Survey on Access and Use of Telecommunication Services in Costa Rica

Vivian Aguilar<sup>a</sup>  
vivian.aguilaraguilar@ucr.ac.cr

Ricardo Alvarado-Barrantes<sup>b</sup>  
ricardo.alvarado@ucr.ac.cr

---

### Resumen

El Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT), realiza la Encuesta de acceso y uso de los servicios de telecomunicaciones en Costa Rica (EAU). Ante la situación fiscal que atraviesa el país es necesario revisar el tamaño de muestra y determinar un posible ajuste que, manteniendo la precisión en las estimaciones, reduzca los costos de la encuesta. Para lograrlo, se desarrolla un proceso de simulación en dos etapas: en la primera se determina el número de UPM y, en la segunda, la cantidad de entrevistas en cada UPM. Estos cálculos utilizan las variables más importantes de la encuesta, como lo son el índice y las escalas de uso de Internet. Para mantener la precisión y estabilidad en las estimaciones se establece que la diferencia entre el percentil 97,5 y 2,5 no sobrepase la cantidad de 0,5 en cada una de las regiones del país. Como resultado se concluye que, en las seis regiones del país, se requiere 141 UPM, por lo que se logra una reducción del 29%. Con ese número de UPM y con un tamaño de submuestra aproximado de 18 entrevistas por UPM, la muestra final pasa de 3 618 a 2 556 entrevistas.

**Palabras clave:** Proceso de simulación, Reducción del tamaño de muestra, Índice de uso de Internet, Escalas de uso de Internet, Precisión de estimaciones..

### Abstract

The Ministry of Science, Innovation, Technology, and Telecommunications (MICITT) conducts the "Survey on Access and Use of Telecommunication Services in Costa Rica (EAU)." Given the country's fiscal situation, it is necessary to review the sample size and determine a possible adjustment that, while maintaining precision in the estimates, reduces the survey's costs. To achieve this, a two-stage simulation process is developed: in the first stage, the number of Primary Sampling Units (UPM) is determined, and in the second stage, the number of interviews in each UPM is determined. These calculations use the most important variables of the survey, such as the Internet usage index and scales. To maintain precision and stability in the estimates, it is established that the difference between the 97.5th and 2.5th percentiles does not exceed 0.5 in each of the country's regions. As a result, it is concluded that, in the six regions of the country, 141 UPM are required, achieving a 29% reduction. With this number of UPM and with an approximate subsample size of 18 interviews per UPM, the final sample size decreases from 3,618 to 2,556 interviews.

**Keywords:** Simulation process, Sample size reduction, Internet usage index, Internet usage scales, Precision of estimates..

---

<sup>a</sup>Universidad de Costa Rica

<sup>b</sup>Universidad de Costa Rica

## 1. Introducción

Como parte de las acciones que ejecuta el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT), para generar información actualizada y periódica, sobre la evolución de las Tecnologías de Información y Comunicación (TIC) en las viviendas y población del país, se realiza la Encuesta de acceso y uso de los servicios de telecomunicaciones en Costa Rica (EAU).

En el país, la apertura del mercado de las telecomunicaciones, en el año 2008, provocó un dinamismo en el sector que beneficia el acceso de más personas a las TIC (SUTEL, 2017), por lo que, en una economía globalizada y en una sociedad cada vez más interconectada, son necesarios los indicadores y su medición.

La EAU brinda insumos que ayudan al sector a velar por el cumplimiento de la normativa vigente, la cual establece en la Ley General de Telecomunicaciones N 8642, la responsabilidad del Estado de garantizar a todos los usuarios el acceso universal y solidario a los servicios de telecomunicaciones, así como la protección de sus derechos en esta materia (Asamblea Legislativa, 2008). Además, esta ley contempla dentro de sus principios rectores el establecimiento de garantías y derechos a favor de los usuarios finales, para que puedan acceder y disfrutar de servicios de calidad a un precio asequible.

Por lo tanto, la EAU es de suma importancia, ya que brinda información del sector, lo que permite formular y actualizar periódicamente las políticas públicas. También, posibilita analizar la información tomando en cuenta diferentes sectores geográficos y establecer acciones diferenciadas para asegurar la inclusión de toda la población a la sociedad de la información y el conocimiento.

La encuesta se realizó anualmente desde el año 2012 hasta el 2018, y se retomó en mayo del 2023. Desde su primera versión hasta la fecha el contenido del cuestionario ha variado, buscando extraer indicadores actualizados en materia de telecomunicaciones, utilizando los estándares establecidos por la Organización para la Cooperación y el Desarrollo Económico (OCDE) y por la Unión Internacional de Telecomunicaciones (UIT). Además, debe tomarse en cuenta que, debido al proceso de incorporación del país a la OCDE, la periodicidad de la encuesta es clave para atender la solicitud de indicadores que demanda tanto la OCDE como la UIT. Sumado a esto, debe mencionarse que la situación fiscal que atraviesa el país implica la adopción de medidas de austeridad y contención del gasto público en aras de evitar el desperdicio de recursos. Bajo este contexto se consideró necesario revisar algunas etapas de la EAU, entre las que destaca la determinación del número de Unidades Primarias de Muestreo (UPM) a seleccionar en la muestra para realizar estimaciones con una precisión adecuada y el tamaño del submuestreo o número de entrevistas en cada UPM. Esta revisión es importante porque tiene un impacto directo en los recursos económicos que se destinan a la encuesta.

En la actualidad, no existe en el país un estudio similar que utilice procesos de simulación para evaluar tamaño de muestra de una encuesta, lo cual contribuye metodológicamente para que otras instituciones que realicen estudios cuantitativos puedan hacerlo y determinar si es posible una reducción de la misma en aras de reducir los costos y por ende el gasto público.

El propósito de la evaluación del tamaño de muestra de la EAU es determinar un tamaño óptimo, manteniendo la precisión en las estimaciones. La evaluación se realiza por medio de un proceso de simulación, alternativa metodológica que se está aprovechando en la estadística y que se aplica en diversas ramas de la ciencia.

El estudio tiene como objetivo calcular el tamaño de muestra o número de entrevistas que deben seleccionarse en la EAU para obtener estimaciones con una precisión determinada en cada una de las regiones de planificación del país. Los objetivos específicos de este estudio son:

- Determinar el número de UPM que deben seleccionarse en cada una de las regiones de planificación del país para lograr una precisión adecuada.
- Determinar la cantidad de entrevistas a submuestrear dentro de cada UPM para lograr una precisión adecuada.
- Desarrollar un proceso de simulación que permita balancear el número de UPM y de entrevistas a submuestrear dentro de cada UPM, para alcanzar una precisión determinada en las estimaciones.

## 2. Metodología

### 2.1. Fuente de datos y variables de estudio

Los datos utilizados en el estudio corresponden a la Encuesta de acceso y uso de los servicios de telecomunicaciones en Costa Rica 2017, realizado por el MICIIT (MICITT, 2019). La encuesta tiene cobertura nacional y se definen como dominios de estudio las seis regiones de planificación: Central, Chorotega, Pacífico Central, Brunca, Huetar Caribe y Huetar Norte. El diseño muestral es estratificado, bietápico, mediante selección con probabilidad proporcional al tamaño (PPT), utilizando como medida de tamaño el total de viviendas de cada UPM. La muestra planeada fue de 3 500 entrevistas y se asignaron 1 000 a la Región Central y 500 en cada una de las cinco regiones restantes.

La encuesta, entre otros, posee el módulo de usos de Internet que cuenta con una batería de 33 preguntas (Anexo A) y se consulta si ha utilizado o no Internet para realizar diferentes actividades. Con esta información se construyeron las escalas de uso de Internet y un índice de uso de Internet, que corresponden a las principales variables de estudio.

### 2.2. Técnicas estadísticas

Se desarrolló un proceso de simulación en dos etapas para evaluar el tamaño de muestra de la EAU: en la primera etapa se determinó el número de UPM a seleccionar y, en la segunda, el tamaño del submuestreo o número de entrevistas a realizar en cada UPM.

Desde un punto de vista estadístico, los procesos de simulación fueron usados en este estudio para modelar y analizar sistemas complejos mediante el uso de modelos matemáticos (Coss Bu, 2003). En este proceso pueden ajustarse parámetros (modificando el número de UPM o tamaño del submuestreo) lo que origina diferentes escenarios para analizar los cambios experimentados. Tal y como lo afirman Robert et al. (2010), sobre el uso de la simulación en el muestreo, ésta será aprovechada para estimar cantidades desconocidas (número de UPM o tamaño del submuestreo) o realizar nuevas inferencias (intervalos de confianza para determinar precisión).

Más concretamente, bajo esta perspectiva, la evaluación del tamaño de la muestra consiste, como lo afirma Kline (2023), en simular datos con un modelo, utilizando diferentes tamaños de muestra, ajustando el modelo y evaluando la precisión y estabilidad de los resultados. La evaluación consiste en analizar la precisión y estabilidad de los resultados después de aplicar los modelos de simulación (Lenth, 2001). Esto se logra calculando medidas de precisión, tales como los errores estándar o los intervalos de confianza. De acuerdo con Krejcie and Morgan (1970), para lograr lo anterior, se debe establecer el nivel de confianza con el que se desean los resultados y especificar un margen de error máximo para las estimaciones. Por lo general se utiliza un nivel de confianza del 95 %. Estos mismos autores señalan que el diseño de la investigación y los objetivos específicos deben tomarse en cuenta para determinar el tamaño de muestra óptimo.

### 2.3. Evaluación de la cantidad de UPM

Con el fin de evaluar el tamaño de muestra de la EAU, se establecen diferentes escenarios, a partir de variaciones en la cantidad de UPM, en cada región de planificación, revisando la precisión de las estimaciones y tomando en consideración el diseño muestral complejo de la EAU. Se toma en cuenta la teoría del muestreo para considerar el diseño muestral complejo de la encuesta, tanto en lo referente a la selección de la muestra como al cálculo de la precisión de las estimaciones.

La idea del análisis era establecer un criterio para seleccionar el número de UPM y, posteriormente, de submuestra en cada UPM, con el fin de mejorar la planificación de la encuesta, tomando en cuenta factores que pueden estar asociados tanto a la selección de la muestra como al trabajo de campo. En una primera etapa, se determinó la cantidad de UPM y, en la segunda, el tamaño del submuestreo.

El proceso para determinar la cantidad de UPM se realizó en fases, de la siguiente manera:

1. Para una región específica, se contabilizaron las UPM disponibles en el archivo de datos de la encuesta (Tabla 1).
2. Se extrajeron muestras aleatorias de UPM, sin exceder el máximo disponible de UPM de esa región.
3. Para desarrollar los escenarios, se inició con 5 UPM y se tomó en cuenta el máximo de UPM existente en cada región. Para cada región se desarrollaron 20 escenarios en total, generados de acuerdo con el número de UPM de esa región.
4. El procedimiento anterior se realizó 5 000 veces ( $K = 5\ 000$  iteraciones).
5. Para cada escenario y en cada iteración se calculó la media para cada una de las variables de estudio; en otras palabras, para una cantidad de 5 UPM se calcularon 5 000 medias de cada una de las variables, luego, para la siguiente cantidad de UPM, por ejemplo 8, se volvieron a calcular 5 000 medias de la variable y así sucesivamente.
6. Para cada uno de los 20 escenarios se calcularon los cuantiles 2,5 y 97,5, lo que determinó el intervalo de confianza, y la diferencia entre sus límites determinó el rango.
7. Como criterio para establecer la cantidad óptima de UPM se definió que el rango no debía sobrepasar el valor de 0,5 puntos. Se utilizó este criterio debido a que se deseaba una precisión suficientemente alta y utilizando este valor el error permitido en la estimación era bajo y se ajustaba con la teoría y las recomendaciones de los autores mencionados.
8. El procedimiento de simulación descrito se repitió en cada una de las seis regiones de planificación y para cada una de las seis variables de estudio.

La generación de las muestras de UPM se realizó de manera aleatoria y los resultados se presentan mediante gráficos.

Tabla 1: Distribución del tamaño de muestra y cantidad de UPM según región de planificación

Región de planificación	Tamaño de muestra	Cantidad de UPM de la muestra	Promedio de entrevistas por UPM
Central	1006	60	17
Chorotega	492	28	18
Pacífico Central	520	27	19
Brunca	495	29	17
Huetar Caribe	508	27	19
Huetar Norte	598	32	19

## 2.4. Evaluación del tamaño de la submuestra

Una vez que se determinó la cantidad de UPM en cada región de planificación, se establecieron diferentes escenarios variando la submuestra, es decir, la cantidad de entrevistas en cada UPM. Esto se hizo de dos maneras: la primera consistió en tomar como base el número de UPM óptimo que se obtuvo de la etapa anterior para determinar, en esta etapa, el tamaño de submuestra. La segunda fue utilizar el número de UPM realizado en la EUA y calcular en esta etapa la submuestra por UPM.

Tomando en consideración lo anterior, el procedimiento seguido fue el siguiente:

1. Para una región específica, para cada UPM, se escogieron diferentes tamaños de submuestra, iniciando con una de 3 entrevistas, hasta alcanzar el máximo disponible en cada UPM. En promedio cada UPM tenía 18 entrevistas, por lo que el tamaño máximo de submuestra establecido fue de 17. En cada región se generaron 9 escenarios.
2. El procedimiento anterior se realizó 5 000 veces ( $K = 5\ 000$  iteraciones).
3. Para cada escenario y en cada iteración se calculó la media y desviación estándar de la variable de estudio; por ejemplo, para una cantidad de 3 entrevistas se sacaron 5.000 UPM y se calcularon 5 000 medias de la variable de interés, luego, para la siguiente cantidad de entrevistas y así sucesivamente.
4. Para cada uno de los escenarios o tamaños de submuestra, se calcularon los cuantiles 2,5 y 97,5 (para tener los intervalos de confianza) y la diferencia entre los límites fue el rango.
5. Como criterio para establecer el submuestreo óptimo se definió que el rango no sobrepasara el valor de 0,5 puntos, mismo que se utilizó en la etapa anterior.
6. El procedimiento se realizó en cada una de las regiones de planificación y para las variables de estudio que se consideraron pertinentes.

Igualmente, se revisó la precisión en las estimaciones considerando el diseño muestral de la EAU. Los resultados se presentan por medio de gráficos.

### 3. Resultados

La evaluación del tamaño de muestra se realizó en dos fases: la primera establece la cantidad de UPM que requiere la encuesta y la segunda define el tamaño de la submuestra, según se explicó en las secciones anteriores. El procedimiento se realizó para cada región de planificación con las estimaciones del índice de uso de Internet y, posteriormente, con las escalas.

La valoración en función de la variable del índice de uso de Internet es primordial, ya que es la variable que engloba a las escalas de uso de Internet, y aunque éstas también tienen un uso importante en los resultados, el índice sin duda será el principal criterio en caso de requerir algún recorte a nivel presupuestario.

#### 3.1. Número de UPM en la muestra

En esta primera fase se evalúa la cantidad de UPM en función de las estimaciones del índice de uso de Internet.

La Figura 1 resume la simulación del comportamiento de las medias del índice de uso de Internet para las diferentes cantidades de UPM. Entonces, para cada cantidad de UPM seleccionadas, se presentan las 5 000 medias del índice y esto se hace para cada una de las regiones de planificación.

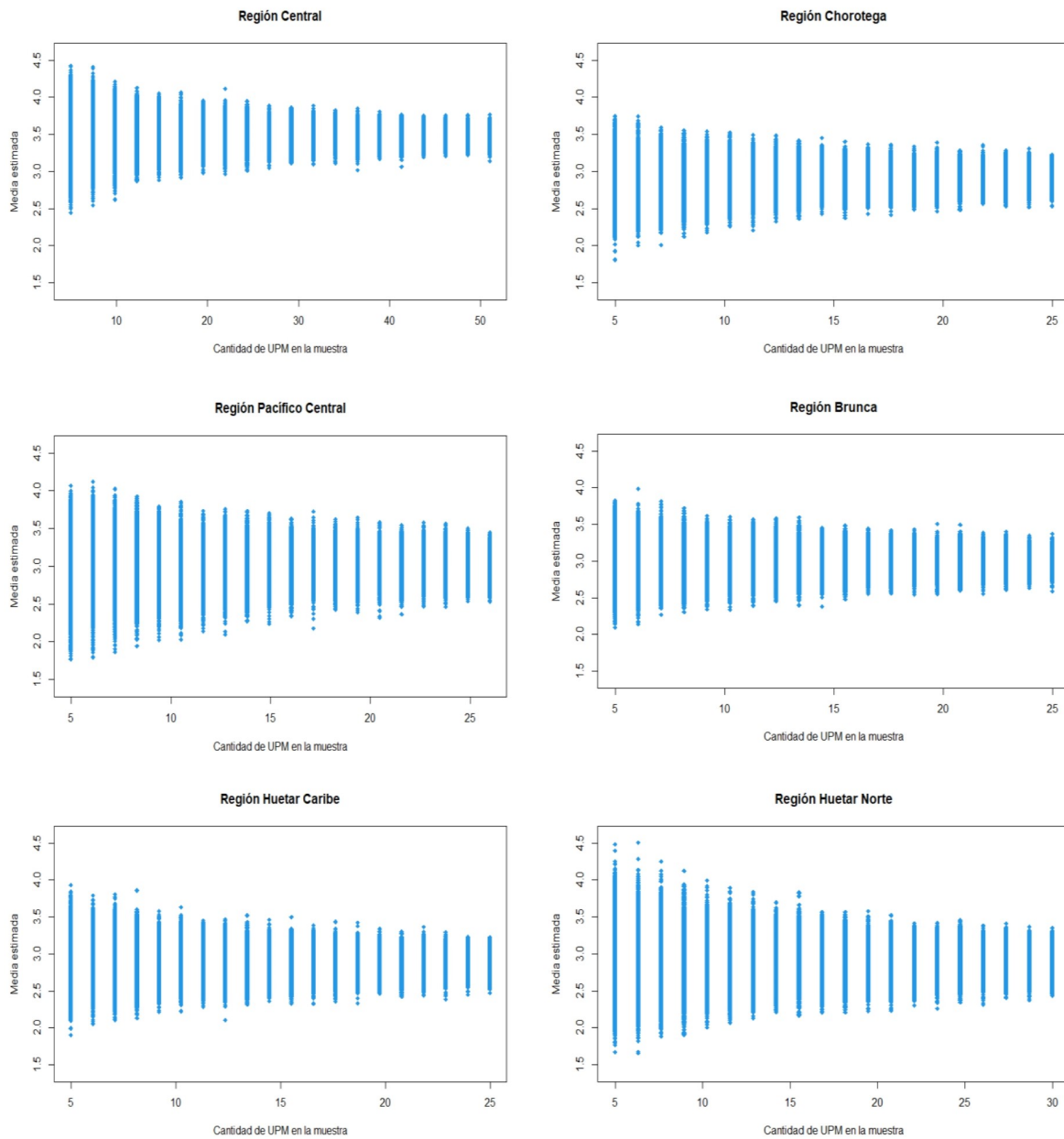


Figura 1: Comportamiento de las medias del índice de uso de Internet según número de UPM simuladas en cada región de planificación

Como es de esperar, a medida que aumenta el número simulado de UPM, la variabilidad de las medias simuladas tiende a decrecer. Las particularidades de cada región revelan varios detalles. En la Central, región que contiene 60 UPM, la variabilidad de las medias no decrece mucho más a partir de las 30 UPM, lo que evidencia una posible reducción de la muestra. En la Chorotega, donde hay 28 UPM, la variabilidad de las medias se estabiliza a partir de 20, por lo que también existen posibilidades de reducción. En la región Pacífico Central, la cual posee un total de 27 UPM, aún con 25 UPM la variabilidad de las medias no disminuye considerablemente, lo que tentativamente sugiere un aumento en la cantidad de UPM. En las regiones Brunca y Huetar Caribe, donde hay 29 y 27 UPM, respectivamente, a partir de 20 la variabilidad de las medias dejar de disminuir. La región Huetar Norte, la cual tiene un total de 32 UPM, no solo presenta

una alta variabilidad, sino que ésta disminuye poco conforme aumenta la cantidad de UPM, lo que hace suponer que requiere más cantidad de UPM.

Como complemento, en la Figura 2, se observa para cada número simulado de UPM la diferencia entre los cuantiles de las medias simuladas del índice de uso de Internet en cada una de las regiones. El criterio utilizado para establecer el número de UPM deseado en la muestra es que la diferencia entre los cuantiles 2,5 y 9,75 de las medias simuladas no sobrepase el valor de 0,5 puntos.

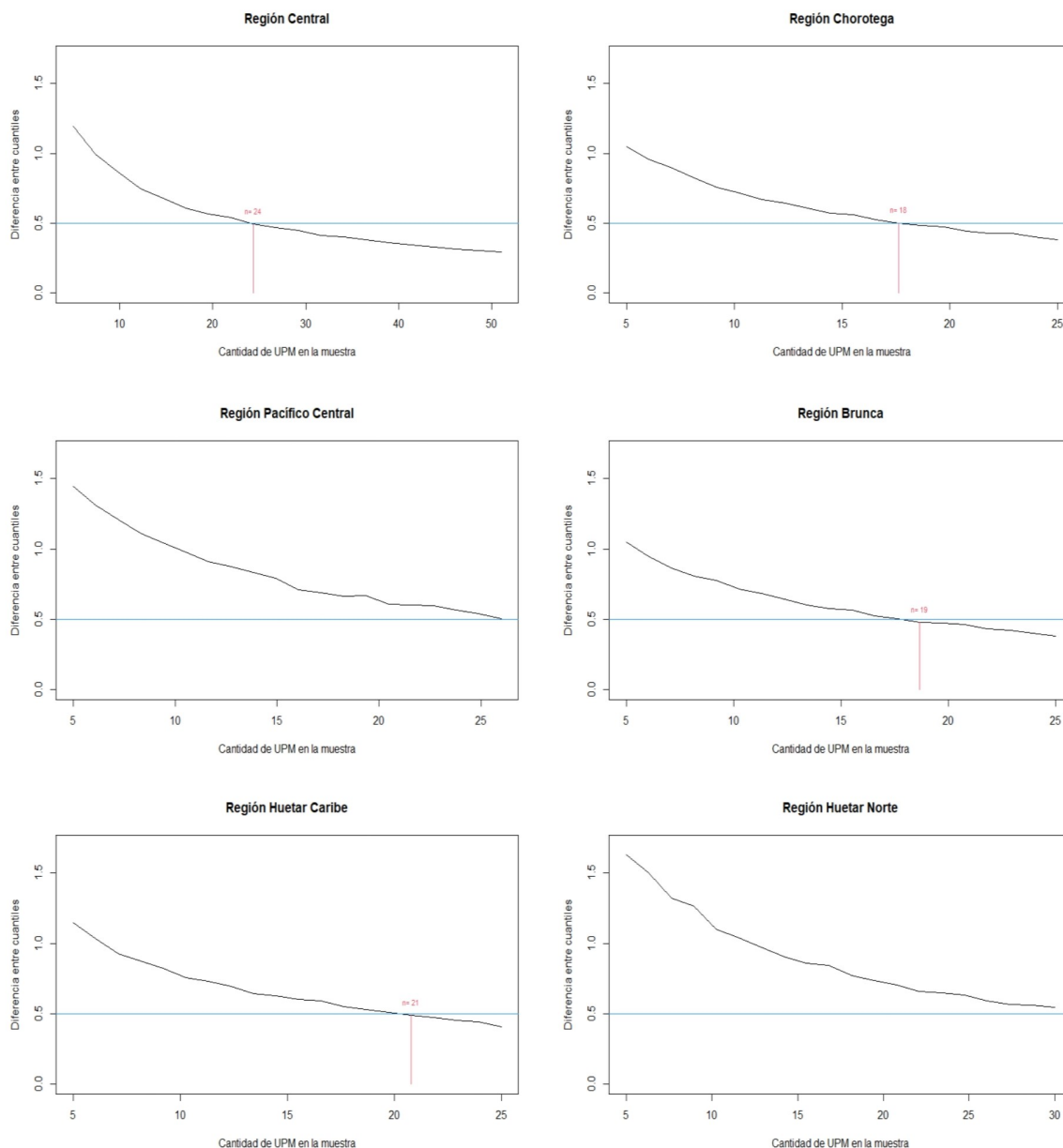


Figura 2: Diferencia entre los cuantiles de las medias del índice de uso de Internet según cantidad de UPM simuladas en cada región de planificación

En la región Central, a partir 24 UPM, la diferencia entre los cuantiles presenta una amplitud menor a la establecida, por lo que la cantidad puede reducirse a menos de la mitad. En la Chorotega, el criterio usado

se cumple con 18 UPM, lo que implica una reducción de 10 UPM. En las regiones Pacífico Central y Huetar Norte, no se alcanza el criterio, pese a ello, es factible verificar que se encuentra muy cerca de alcanzarlo, por lo que puede mantenerse la cantidad de UPM actuales. En la región Brunca, el criterio se cumple para una cantidad de 19 UPM, resultado que evidencia una reducción considerable en el número actual de UPM en la muestra. En la región Huetar Caribe, la diferencia entre los cuantiles llega al límite propuesto para una cantidad de 21 UPM, por lo que la cantidad de UPM se reduce en 6 con respecto al número actual.

Después de realizar este análisis, no se encuentra evidencia para aumentar la cantidad de UPM actual en la muestra, ya que en los casos en que el criterio establecido no se cumple, es por muy poco, por lo que es factible mantener el número de UPM en estas regiones sin que se pierda precisión en el cálculo del índice de uso de Internet de una manera considerable en cada región.

La Tabla 2 resume los resultados del tamaño de muestra tomando en consideración las estimaciones del índice de uso de Internet. En la región Central, el número de UPM puede reducirse de 60 a 24 UPM (36 UPM menos), lo que representa una disminución en el tamaño de muestra de 603 entrevistas, si se mantiene el submuestreo actual de 17 entrevistas por UPM. En las regiones Chorotega, Brunca y Huetar Caribe el número de UPM se reduce respectivamente en 10, 10 y 6. La muestra total final pasa de 3 618 a 2 556, lo que se traduce en una reducción total de 1 062 entrevistas.

Tabla 2: Distribución del tamaño de muestra y cantidad de UPM al evaluar las estimaciones del índice de uso de Internet según región de planificación

Región de planificación	Cantidad de UPM de la muestra	Tamaño de muestra	Cantidad de UPM obtenidas en simulación	Tamaño de muestra propuesto
Central	60	1006	24	402
Chorotega	28	492	18	316
Pacífico Central 1/	27	520	27	520
Brunca	29	495	19	324
Huetar Caribe	27	508	21	395
Huetar Norte 1/	32	598	32	598

1/ Se mantiene la cantidad de UPM asumiendo un error un poco más grande.

Al evaluar el tamaño de muestra con la consideración de las estimaciones de las escalas de uso de Internet, se tienen diferentes escenarios (Tabla 3). En las estimaciones de la escala informativa no se alcanza el criterio para la reducción de la cantidad de UPM en ninguna de las regiones (se ilustra gráficamente en las Figuras 6 y 7 del Anexo B). Cabe destacar que este indicador está compuesto únicamente por dos variables, por lo que su comportamiento podría ser inestable. Además, de las escalas analizadas ésta es la que presenta mayor variabilidad pese a que su construcción se sustenta desde la teoría. Por lo tanto, el análisis se realiza con las otras escalas y no se toma en cuenta en la elección de la cantidad de UPM. De esta forma, se mantiene la cantidad actual.

Tabla 3: Distribución la cantidad de UPM al evaluar las estimaciones de las escalas de uso de Internet según región de planificación

Región de planificación	Cantidad de UPM	Cantidad de UPM en la simulación				
		E2	E1	E3	E4	E5/E6
Central	60	24	24	n.a.	32	29/39
Chorotega	28	18	14	n.a.	n.a.	24/21
Pacífico Central	27	19	n.a.	n.a.	n.a.	n.a./n.a.
Brunca	29	19	22	n.a.	n.a.	17/23
Huetar Caribe	27	18	16	21	n.a.	n.a./n.a.
Huetar Norte	32	29	n.a.	n.a.	n.a.	n.a./n.a.

Nota 1: Se indica n.a. (no aplica) cuando no alcanza el criterio para determinar un número menor de UPM.

Nota 2: Índice de uso de Internet (E1), Escala informativa (E2), Escala de interacción social (E3), Escala de entretenimiento (E4), Escala de productividad (E5), Escala de habilitación ciudadana (E6).

En la escala de interacción social, se cumple el criterio para todas las regiones (Figuras 8 y 9 del Anexo B), arrojando una reducción importante en la cantidad de UPM. Por otra parte, la escala de entretenimiento (Figuras 10 y 11) alcanza el criterio en la región Central y Huetar Caribe. Para las demás regiones, el criterio no se cumple, aunque en la Figura 11 se aprecia gráficamente que se encuentra muy cerca el límite establecido. En el caso de la escala de productividad (Figuras 12 y 13) y de habilitación ciudadana (Figuras 14 y 15), el criterio se cumple en la región Central, Chorotega y Brunca, y en Figuras 13 y 15 se aprecia que la diferencia entre los cuantiles se encuentra muy cerca del límite establecido.

Al incluir las estimaciones de las escalas de uso de Internet y analizar los resultados por región de planificación, se obtiene una cantidad de UPM de 39 para la región Central, ya que según el análisis es la cantidad de UPM de mayor tamaño requerida entre las variables analizadas. En la región Chorotega y Brunca, la cantidad de UPM queda en 28 y 29 respectivamente, tal y como se encuentra en la muestra original. Cabe mencionar que para ambas regiones la mayoría de las escalas cumplió el criterio, lo que evidencia una posible reducción en la cantidad de UPM, con excepción de la escala de entretenimiento. Pese a esto, en la Figura 11 se observa que para estas regiones el criterio estuvo muy próximo a cumplirse, pues casi roza la línea del límite que determina la precisión deseada, por lo que se sugiere mantener la cantidad de UPM.

En la región Huetar Caribe, también se mantiene la cantidad original de 27 UPM ya que, aunque las escalas de productividad y habilitación ciudadana no cumplen el criterio, la diferencia entre los cuantiles prácticamente se cumple con la cantidad de UPM de la muestra original, lo cual se evidencia en que la línea está muy próxima al límite establecido (Figuras 13 y 15).

Finalmente, en las regiones Pacífico Central y Huetar Norte, únicamente se alcanza el criterio para la escala de interacción social. Aunque en las Figuras 11, 13 y 15, se aprecia que el criterio se encuentra próximo a cumplirse en las escalas de las diferentes regiones, puede valorarse un pequeño ajuste en la cantidad de UPM en cada región. Por este motivo se considera que con 5 UPM adicionales en cada región el criterio se ajustaría para las demás escalas de uso de Internet.

En síntesis, en la Tabla 4 y con base en el análisis anterior se describen los resultados del número óptimo de UPM en la muestra. Obsérvese que al evaluar la cantidad de UPM, al utilizar únicamente la variable del índice de uso de Internet hay una disminución de 62 UPM, y al incluir las variables de las escalas de uso de Internet la disminución es de 11.

Tabla 4: Cantidad de UPM de la muestra original y cantidad de UPM óptimo

Región de planificación	Cantidad de UPM original	Cantidad de UPM óptimo 1/	Cantidad de UPM óptimo 2/
Central	60	24	39
Chorotega	28	18	28
Pacífico Central	27	27	32
Brunca	29	19	29
Huetar Caribe	27	21	27
Huetar Norte	32	32	37
<b>Total</b>	<b>203</b>	<b>141</b>	<b>192</b>

1/ Utilizando únicamente la variable del índice de uso de Internet.

2/ Incluyendo las variables de las escalas de uso de Internet.

### 3.2. Cantidad de entrevistas en cada UPM

El análisis de la submuestra por UPM se realizó evaluando únicamente las estimaciones del índice de uso de Internet.

El primer análisis considera la cantidad óptima de UPM obtenidas en el proceso de simulación anterior. En la Figura 3 se tiene para cada tamaño de submuestra, la diferencia entre los cuantiles de las medias simuladas del índice de uso de Internet en cada una de las regiones donde hubo una reducción en la cantidad de UPM (en la región Pacífico Central y Huetar Norte no hubo disminución en la cantidad de UPM). Obsérvese que el criterio está muy cerca de cumplirse en la cantidad máxima de submuestra, lo que sugiere mantener la

submuestra original (en promedio 18 entrevistas). Esto evidencia que al realizar una reducción en la cantidad de UPM para las regiones correspondientes hace que no sea viable una reducción de la submuestra.

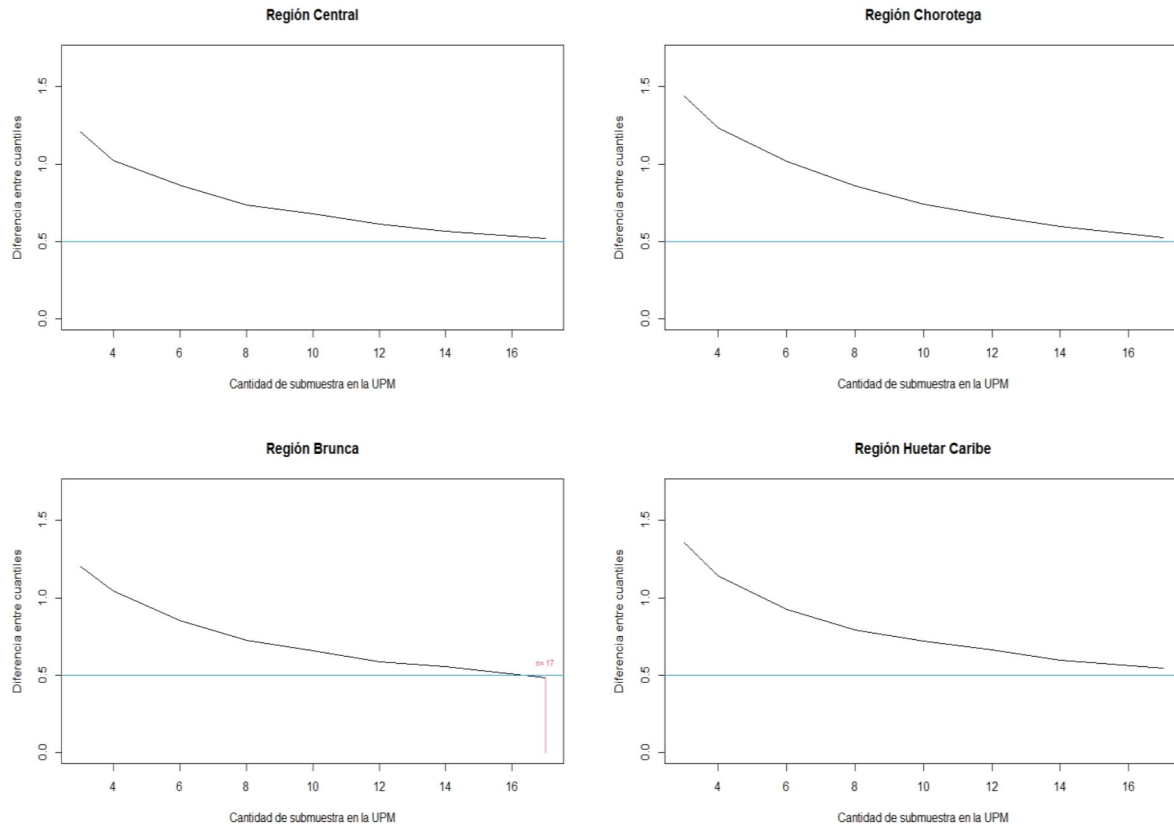


Figura 3: Diferencia entre los cuantiles de las medias del índice de uso de Internet utilizando la reducción en la cantidad de UPM simuladas según tamaño de submuestra

Para analizar otra alternativa, el análisis de la submuestra por UPM se realizó tomando en cuenta la cantidad original de UPM de la muestra. En la Figura 4, se tiene el comportamiento de las medias del índice para las cantidades de submuestra analizadas. Obsérvese que, en la región Central con un total de 60 UPM, la variabilidad de las medias se estabiliza a partir de la submuestra de tamaño 12, lo que hace suponer una disminución en la cantidad de entrevistas. En la región Chorotega y Pacífico Central, con el número original de UPM, la variabilidad de las medias tiende a mantenerse en los niveles esperados en la submuestra de tamaño 16, por lo tanto, se podría esperar una leve disminución en la cantidad de entrevistas. En la región Brunca se aprecia que la variabilidad de las medias no decrece mucho más a partir de 14 entrevistas, lo que evidencia una posible reducción en la submuestra por UPM. Para las regiones Huetar Caribe y Huetar Norte, las medias se estabilizan a partir de 16 entrevistas, por lo tanto, se esperaría que la submuestra disminuya levemente o se mantenga en la cantidad original.

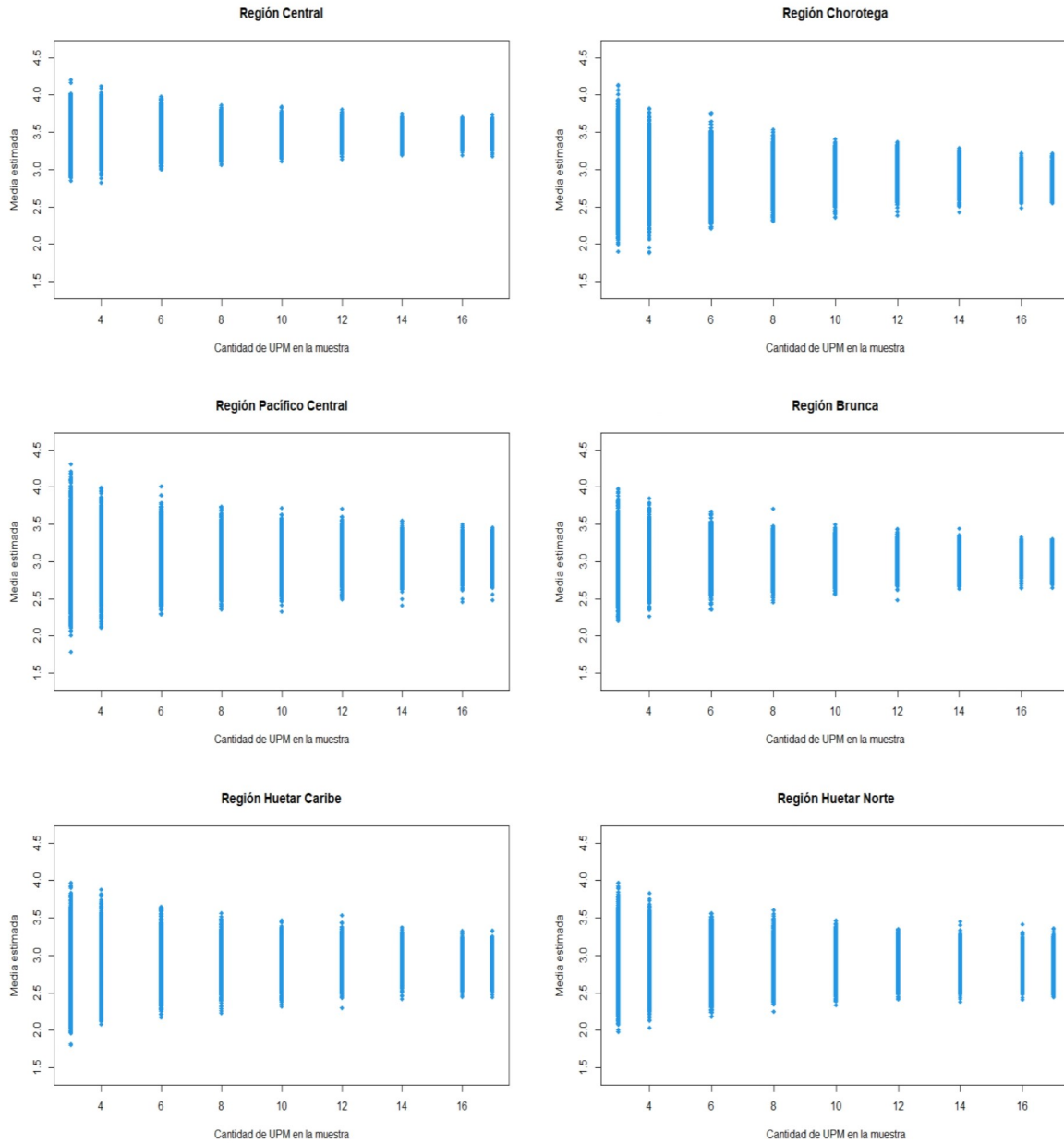


Figura 4: Comportamiento de las medias del índice de uso de Internet según cantidad de submuestra simulada en cada región de planificación

Como complemento a los resultados anteriores, en la Figura 5 se tiene para cada tamaño de submuestra, la diferencia entre los cuantiles de las medias simuladas del índice de uso de Internet. En la región Central, el análisis evidencia que la submuestra se puede disminuir a 8 entrevistas por UPM, reducción de poco más de 50% en la cantidad de entrevistas. En la región Chorotega, el criterio establecido se cumple para una submuestra de tamaño 14. Para la región Pacífico Central, la submuestra se establece en 17 entrevistas, lo que equivale a una pequeña disminución con respecto a las 19 entrevistas que en promedio se realizan en cada UPM de esta región. En la región Brunca, la reducción es importante pues la submuestra se establece en 10 entrevistas. Para la región Huetar Caribe, el criterio se cumple a partir de la submuestra de tamaño 17. Finalmente, en la región Huetar Norte, se aprecia que la línea prácticamente está pegando con el límite

de la precisión deseada, por lo que se considera un tamaño de submuestra aceptado de 17 entrevistas.

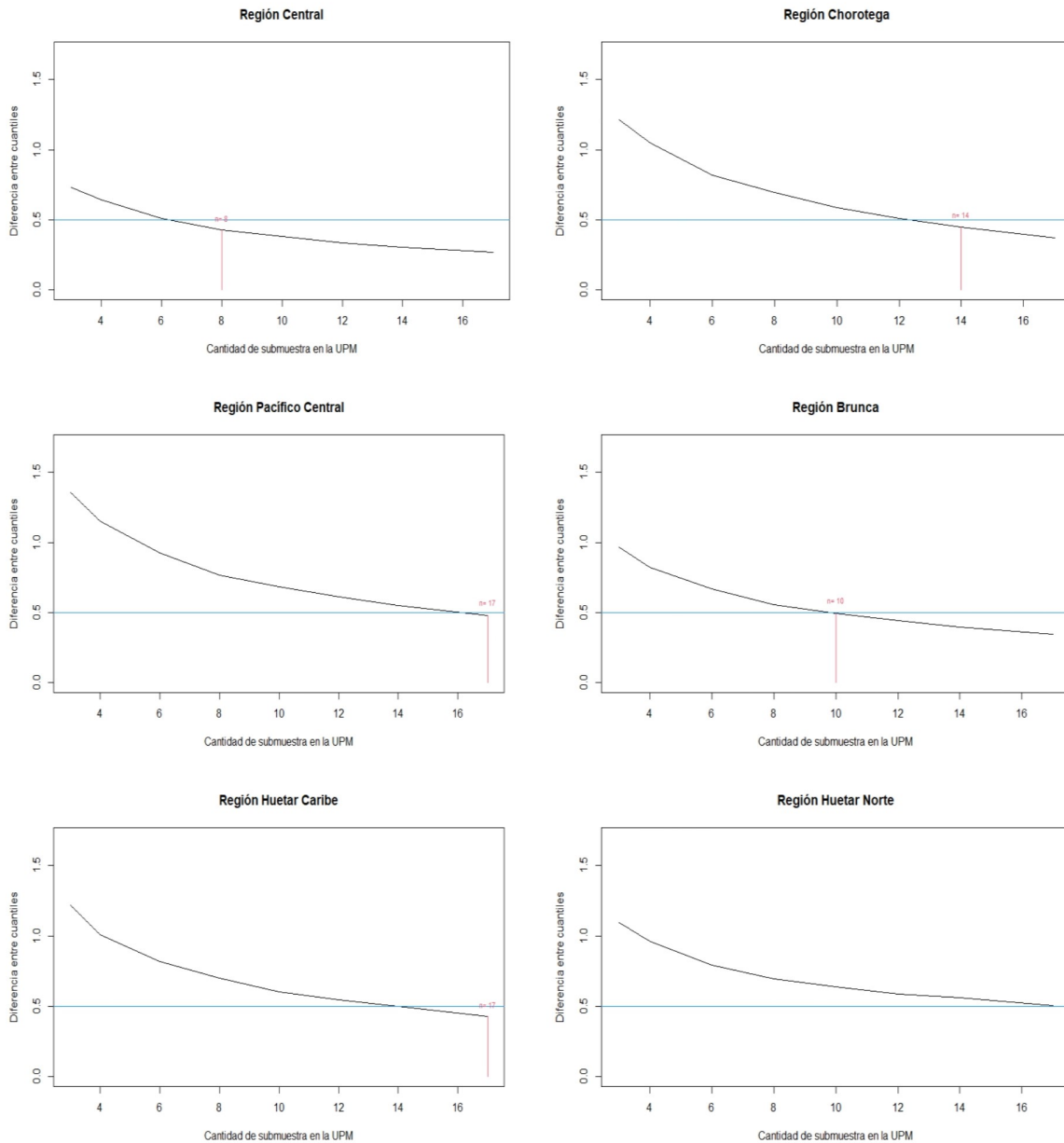


Figura 5: Diferencia entre los cuantiles de las medias del índice de uso de Internet utilizando la cantidad de UPM de la muestra original según tamaño de submuestra

En la Tabla 5, se comparan los resultados del tamaño de muestra al considerar la reducción del número de UPM del proceso anterior, esto es manteniendo la submuestra por UPM con el tamaño original. Estos resultados se comparan con los obtenidos al variar el tamaño de la submuestra y mantener la cantidad de UPM de la muestra original.

Tabla 5: Distribución de la muestra original, con cantidad de UPM y submuestra simuladas, según región de planificación, utilizando las estimaciones del índice de uso de Internet

Región de planificación	(1) Cantidad de UPM de la muestra	(2) Promedio submuestra por UPM	(3) Tamaño de muestra original	(4) Cantidad de UPM en simulación	(5) Tamaño de muestra con UPM simuladas	(6) Submuestra utilizando cantidad de UPM de la muestra	(7) Tamaño de muestra con submuestra simulada
Central	60	17	1006	24	402	8	480
Chorotega	28	18	492	18	316	14	392
Pacífico Central	27	19	520	27	520	17	459
Brunca	29	17	495	19	324	10	290
Huetar Caribe	27	19	508	21	395	17	459
Huetar Norte	32	19	598	32	598	17	544
<b>Total</b>	<b>203</b>		<b>3618</b>	<b>141</b>	<b>2556</b>		<b>2624</b>

La columna (1) presenta la cantidad de UPM de la muestra original y la columna (3) el tamaño de muestra original. La columna (2) se obtiene de dividir la columna (3) entre la columna (1). La columna (4) presenta la cantidad de UPM óptima que se obtiene del proceso de simulación, mientras que en la columna (5) se tiene el tamaño de muestra con esa cantidad de UPM y el submuestreo original (sale de la multiplicación de la columna (2) y (4)). Finalmente, en la columna (6) se tiene la submuestra simulada utilizando la cantidad de UPM de la muestra original y en la columna (7) el tamaño de muestra utilizando esa cantidad de submuestra (se obtiene de multiplicar la columna (1) y (6)).

Obsérvese que después de conocer los resultados del proceso de simulación, la submuestra que utiliza la cantidad de UPM de la muestra original, disminuye considerablemente para la región Central, Chorotega y Brunca, donde en promedio la submuestra era de 17, 18 y 17, respectivamente, y pasa a 8, 14 y 10. Si se compara este resultado con el procedimiento alternativo de optar por un número óptimo de UPM y mantener el submuestreo por UPM originalmente establecido, se observa que la diferencia es de 68 entrevistas.

De ahí que, en términos de reducción de la muestra, se considera que resulta más factible optar por la simulación que provee una cantidad óptima de UPM (columna 4) y mantener el tamaño promedio de la submuestra por UPM (columna 2) ya que, para efectos logísticos y económicos, aunque la submuestra sea mayor, se concentra en una misma zona geográfica, lo que disminuye los gastos que implican el traslado, transporte y, por ende, hay una reducción de tiempo, posiblemente de personal humano y todo esto disminuye el costo final del trabajo de campo.

## 4. Conclusiones

La evaluación del tamaño de muestra de la EAU al utilizar un proceso de simulación brinda una herramienta útil para el planteamiento de la encuesta en años futuros. Como variables de interés para la EAU se estableció el índice de uso de internet y 5 escalas relacionadas con el uso de Internet. Para cada región de planificación y para las variables antes mencionadas se realizaron 228 procesos de simulación con 5 000 réplicas cada uno. Esto permitió calcular para cada muestra generada, la media de las variables, por lo que para valorar la precisión se calcularon los percentiles 2,5 y 97,5 de las 5 000 repeticiones y se estableció que la diferencia entre el máximo y el mínimo no debía sobrepasar la cantidad de 0,5.

Para analizar el número óptimo de UPM en la muestra se calculó el número requerido en cada región y se dejó sin cambios el número de entrevistas por UPM realizado en la encuesta. En este procedimiento el número de UPM en la muestra podía aumentar o disminuir. Los resultados evidenciaron la posibilidad de una reducción importante en la cantidad de UPM, al priorizar las estimaciones del índice de uso de Internet. Para la región Central puede lograrse una reducción del 60%, pasando de 60 a 24 UPM, lo que representa 603 entrevistas menos en el tamaño final de la muestra. Por otro lado, la cantidad de UPM utilizada en las regiones Chorotega, Brunca y Huetar Caribe puede reducirse considerablemente en 36%, 34% y 22%, respectivamente, lo que representa una reducción en las tres regiones de 462 entrevistas. En las

regiones Pacífico Central y Huetar Norte, no se alcanza el criterio establecido; no obstante, puede verificarse gráficamente que puede mantenerse la misma cantidad de UPM en la muestra sin perder precisión de manera considerable.

El análisis del tamaño de la submuestra se realizó utilizando únicamente las estimaciones del índice de uso de Internet y se analizaron dos posibilidades: la primera consistió en utilizar el número de UPM óptimo obtenido anteriormente para determinar el número de entrevistas requeridas por UPM y la segunda consideró mantener fijo el número de UPM utilizadas originalmente en la muestra y calcular el número de entrevistas requeridas por UPM. En el primer caso, al analizar el tamaño de la submuestra con la reducción en la cantidad de UPM se obtuvo que no fue posible disminuir la cantidad de entrevistas por UPM debido a que el número actualmente utilizado es el que alcanza la precisión deseada por región. En el segundo caso, es decir, al analizar la submuestra que utiliza la cantidad de UPM de la muestra original se obtiene que la muestra final no se reduce. Por lo anterior, termina siendo más factible, en términos logísticos y económicos, disminuir la cantidad de UPM y mantener el tamaño original de la submuestra por UPM. Esto significa, concretamente, un número de UPM y de entrevistas por UPM, como se muestra en la Tabla 6.

Tabla 6: Propuesta del nuevo tamaño de muestra para la EAU

Región de planificación	Número de UPM	Entrevistas por UPM	Tamaño de muestra
Central	24	17	402
Chorotega	18	18	316
Pacífico Central	27	19	520
Brunca	19	17	324
Huetar Caribe	21	19	395
Huetar Norte	32	19	598
<b>Total</b>	<b>141</b>		<b>2556</b>

Es importante tomar en cuenta que la planeación de la encuesta implica evaluar múltiples factores. Uno de los de los más importantes es el presupuesto y el tiempo disponible para la ejecución del trabajo de campo, lo que impactará de manera directa el tamaño de la muestra final. Se considera que los resultados de este estudio permitirán continuar profundizando en el análisis de los datos, incluso incluyendo otras variables que sean de interés para el MICITT. Esto permitirá planificar de una manera más eficiente el trabajo de campo, así como lograr una distribución óptima en el uso de los recursos disponibles.

## A. Encuesta

Tabla 7: Preguntas para construir índice y escalas de uso de Internet

---

Descargar imágenes, videos, música, jugar o descargar juegos
Descargar software o aplicaciones
Participar en redes sociales
Descargar aplicaciones
Llamar por teléfono por internet/VoIP
Enviar o recibir correo electrónico (uso privado/personal)
Buscar información sobre salud
Obtener información sobre bienes y servicios
Colgar texto, imágenes, fotos, videos, música en un sitio web para compartirlo
Leer o descargar periódicos o revistas en línea, libros electrónicos
Obtener información de sitios web gubernamentales o de servicios públicos
Escuchar la radio por Internet
Consultar Wikipedia, enciclopedias en línea
Ver televisión por la Internet
Enviar o recibir correo electrónico (trabajo)
Realizar trámites u operaciones bancarias
Utilizar espacio de almacenamiento en Internet para guardar documentos
Buscar empleo o enviar solicitudes de empleo
Descargar formularios oficiales de sitios web gubernamentales o de servicios públicos
Comprar o encargar bienes o servicios
Acceder a sitios de charla, blog, grupo de noticias o debates en línea
Edición en Internet de documentos, hojas de cálculo o presentaciones
Enviar formularios completos de sitios web gubernamentales o de servicios públicos
Utilizar servicios a viajes y el correspondiente alejamiento
Actualizar o añadir contenido a los blogs
Cursar estudios oficiales en línea
Vender bienes o servicios
Cargar contenido creado de forma independiente en sitios web compartidos
Participar en redes de profesionales
Interactuar con organizaciones gubernamentales generales
Publicar opiniones sobre cuestiones políticas o cívicas en los sitios web
Participar en consultas en línea o votar sobre asuntos cívicos o políticos
Gestionar una página web personal/propia

---

## B. Figuras

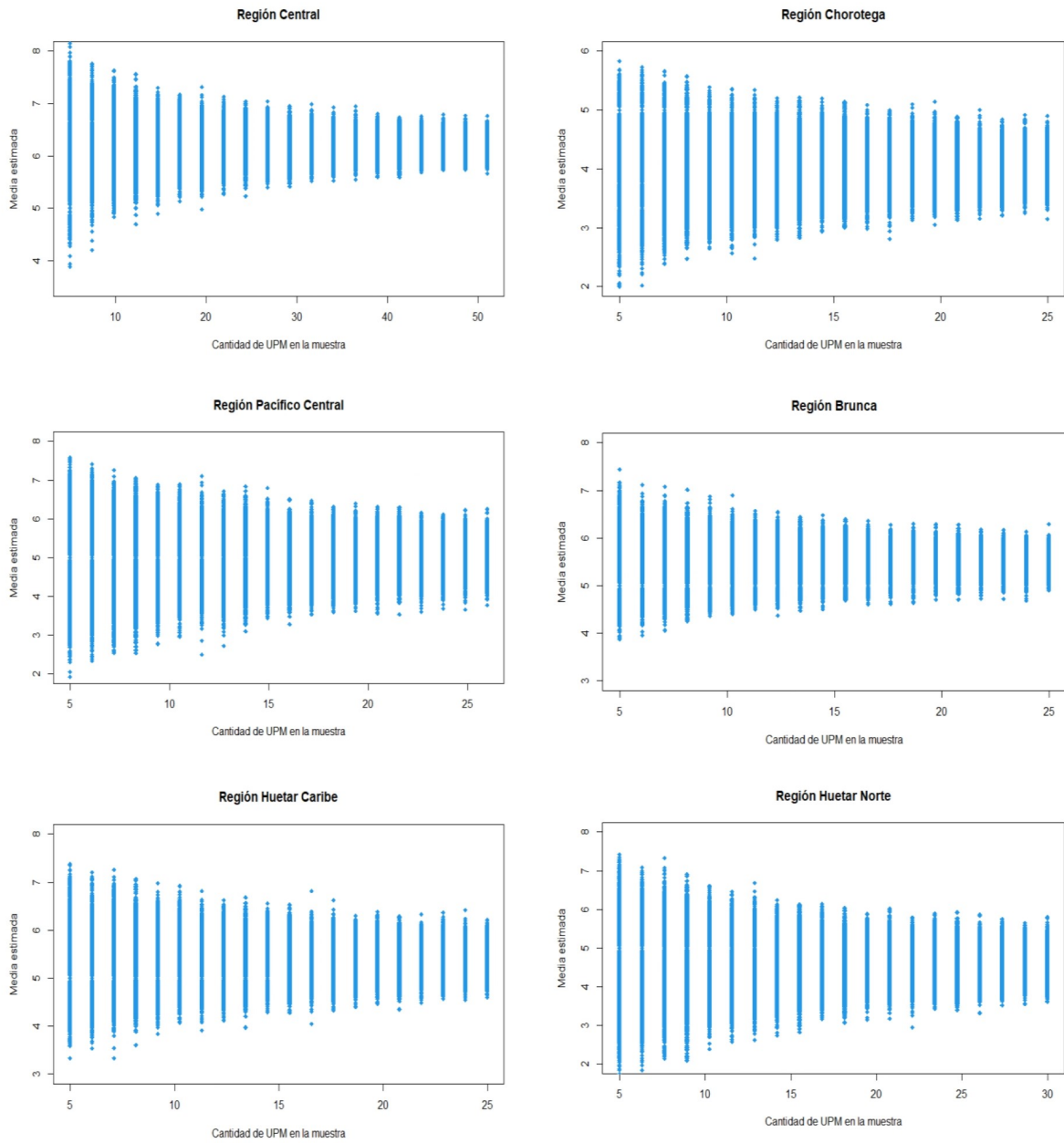


Figura 6: Comportamiento de las medias de la escala informativa según número de UPM simuladas en cada región de planificación

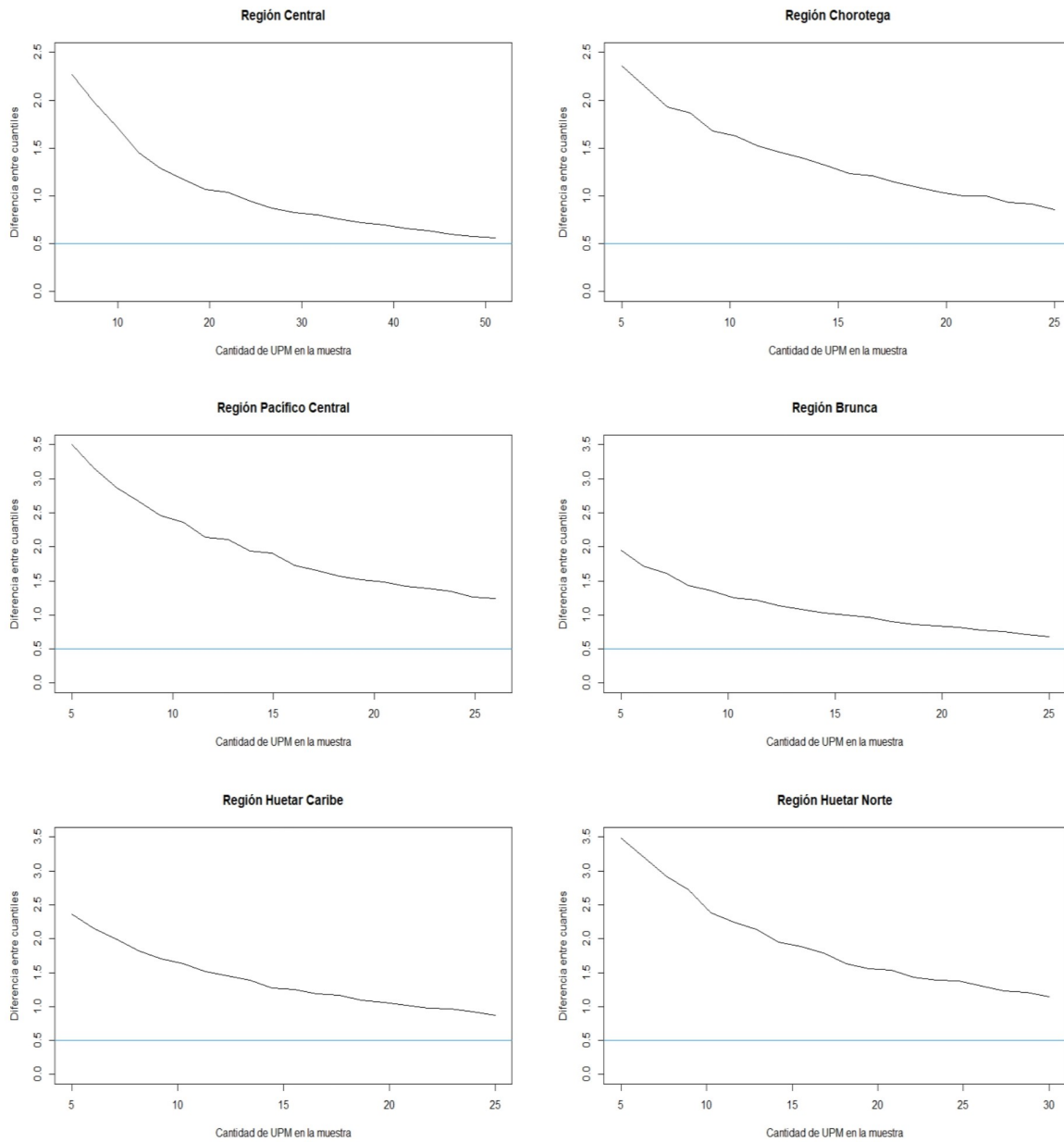


Figura 7: Diferencia entre los cuantiles de las medias de la escala informativa según cantidad de UPM simuladas en cada región de planificación

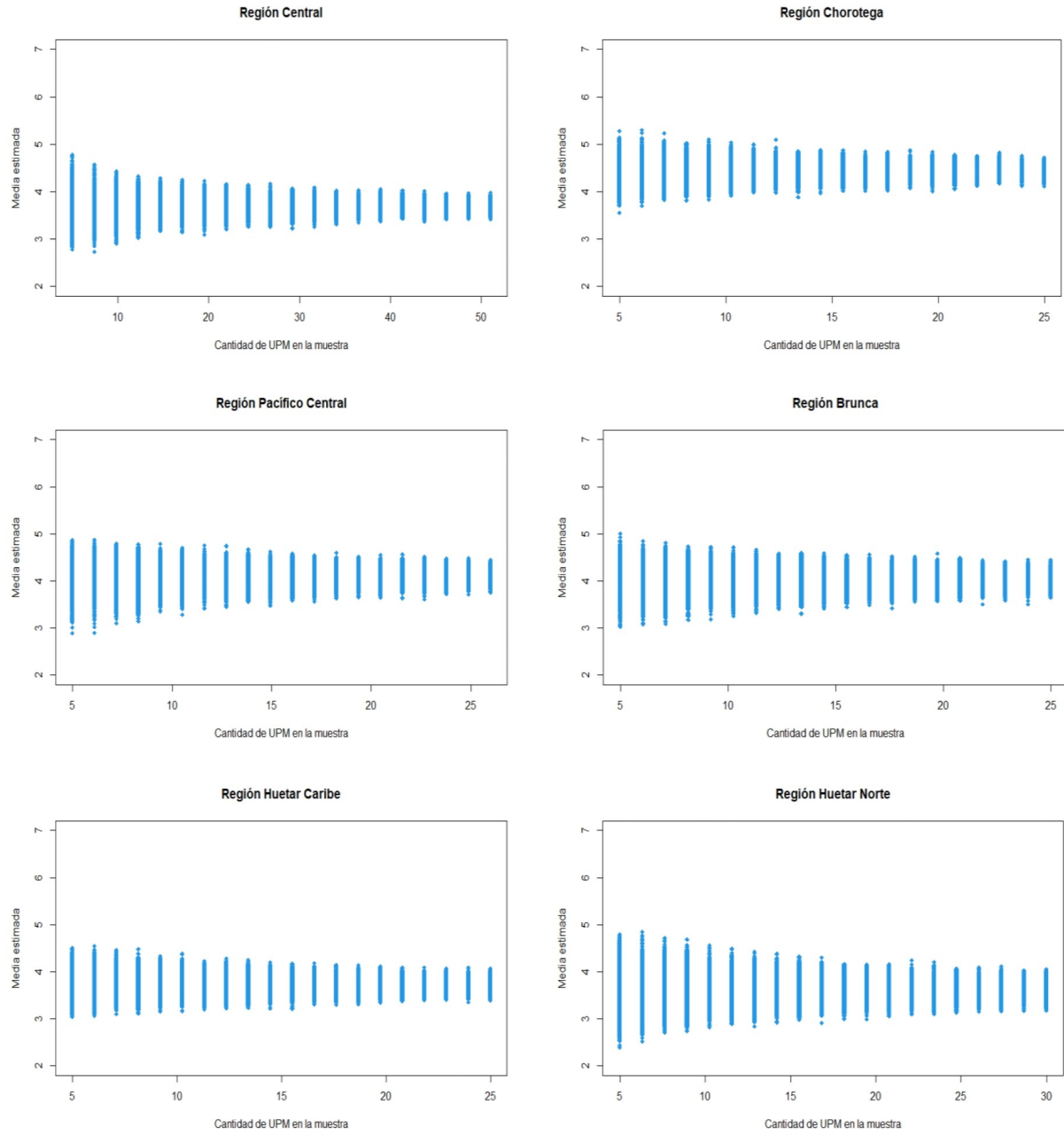


Figura 8: Comportamiento de las medias de la escala de interacción social según número de UPM simuladas en cada región de planificación

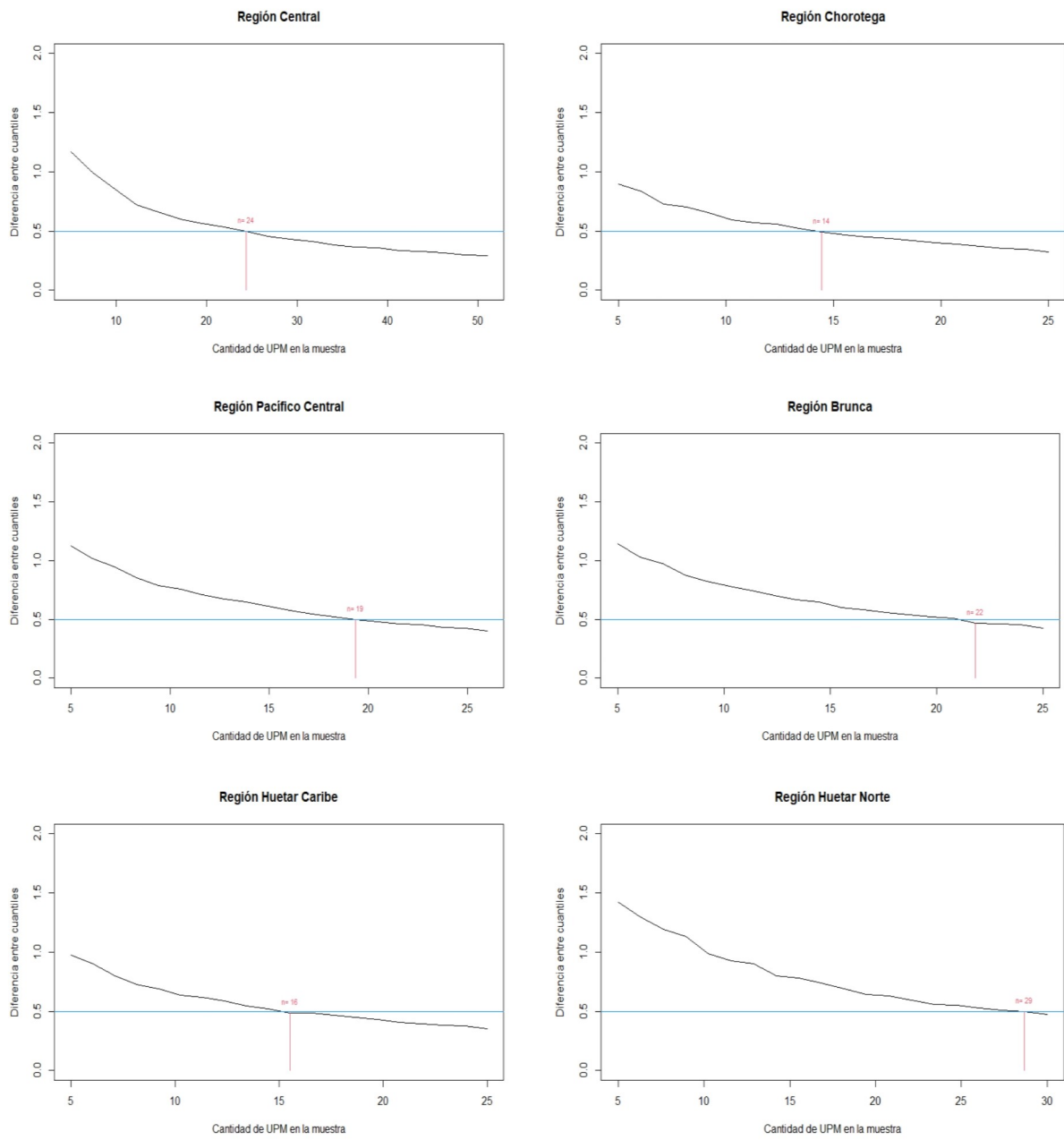


Figura 9: Diferencia entre los cuantiles de las medias de la escala de interacción social según cantidad de UPM simuladas en cada región de planificación

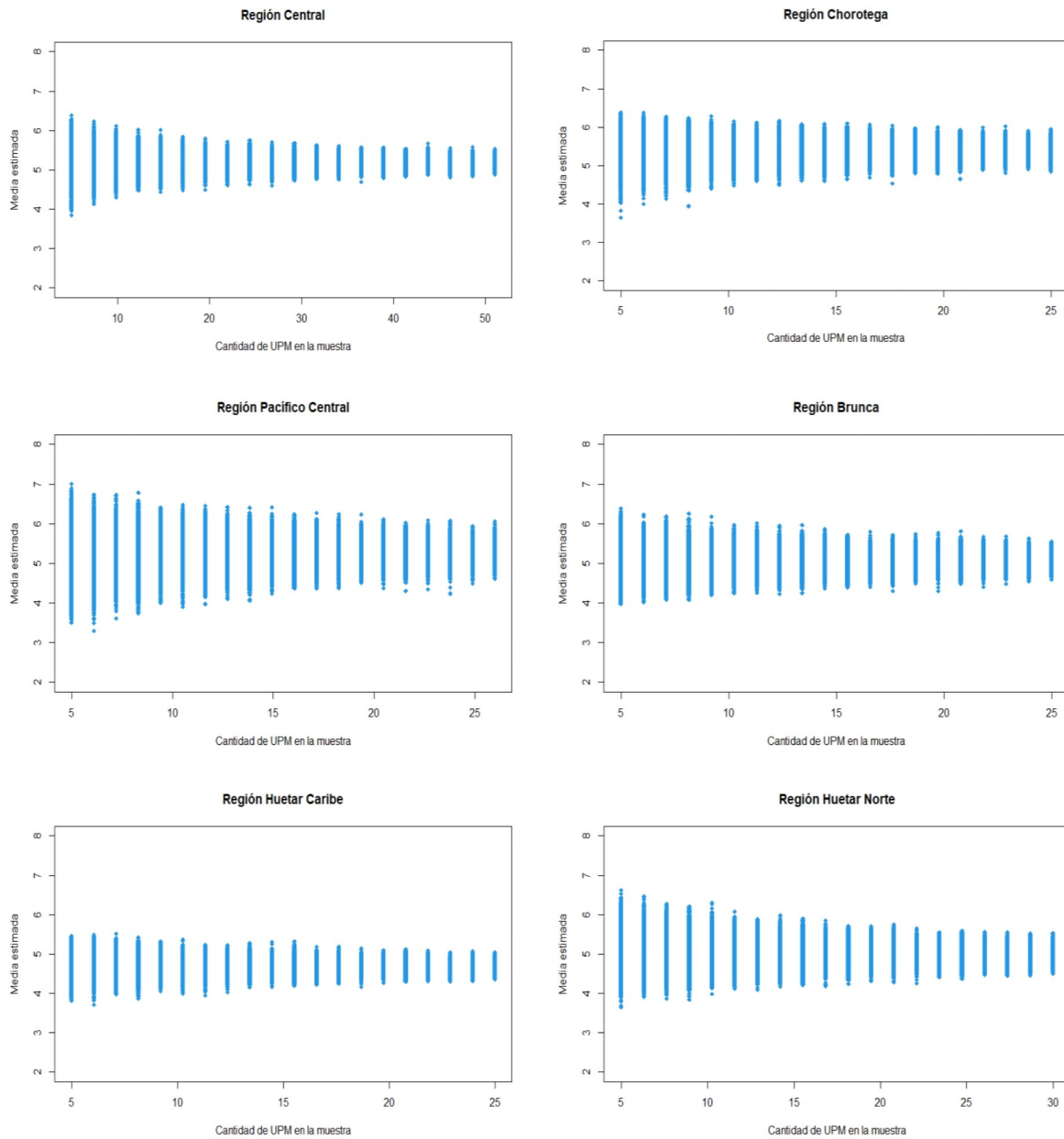


Figura 10: Comportamiento de las medias de la escala de entretenimiento según número de UPM simuladas en cada región de planificación

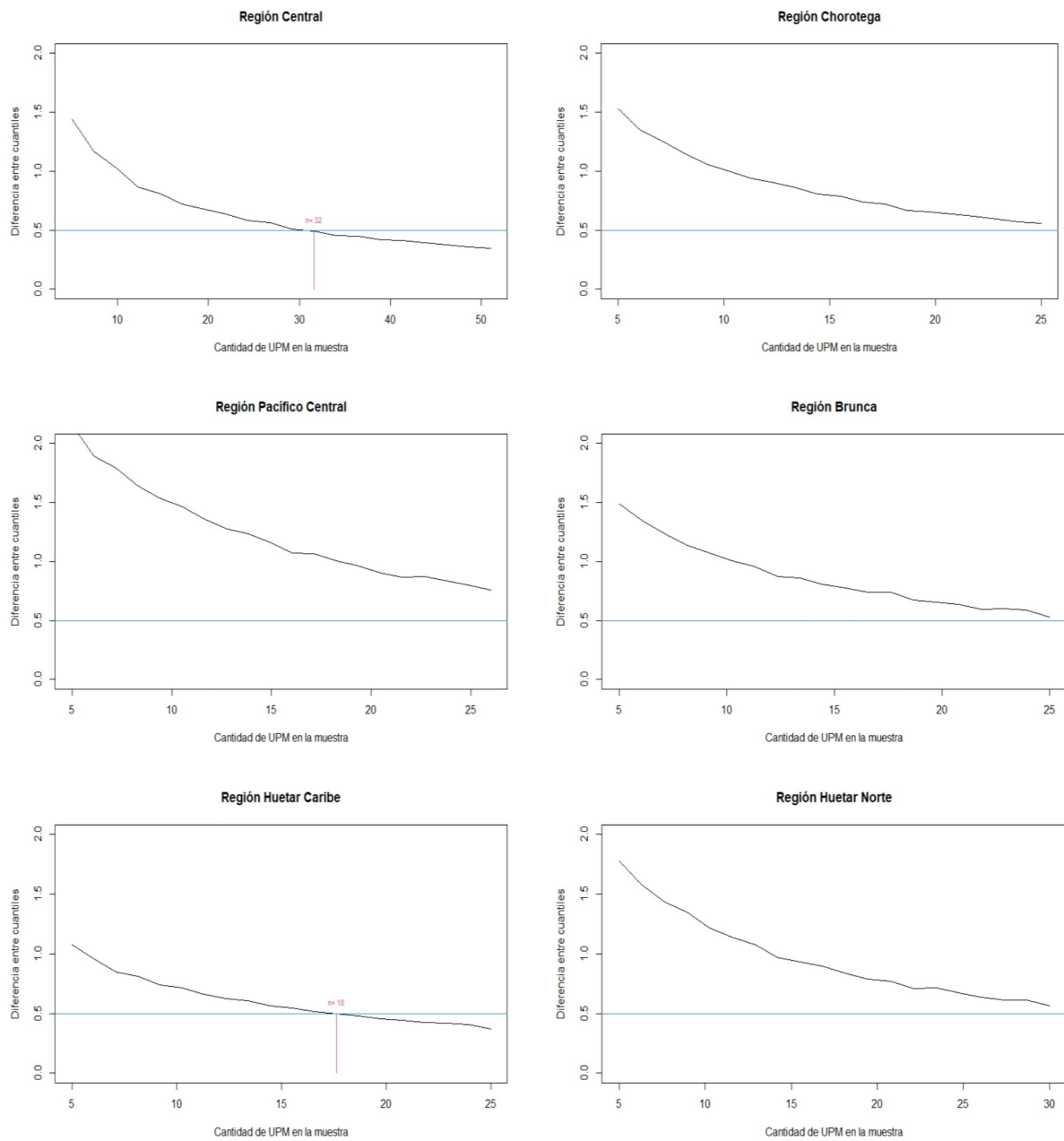


Figura 11: Diferencia entre los cuantiles de las medias de la escala de entretenimiento según cantidad de UPM simuladas en cada región de planificación

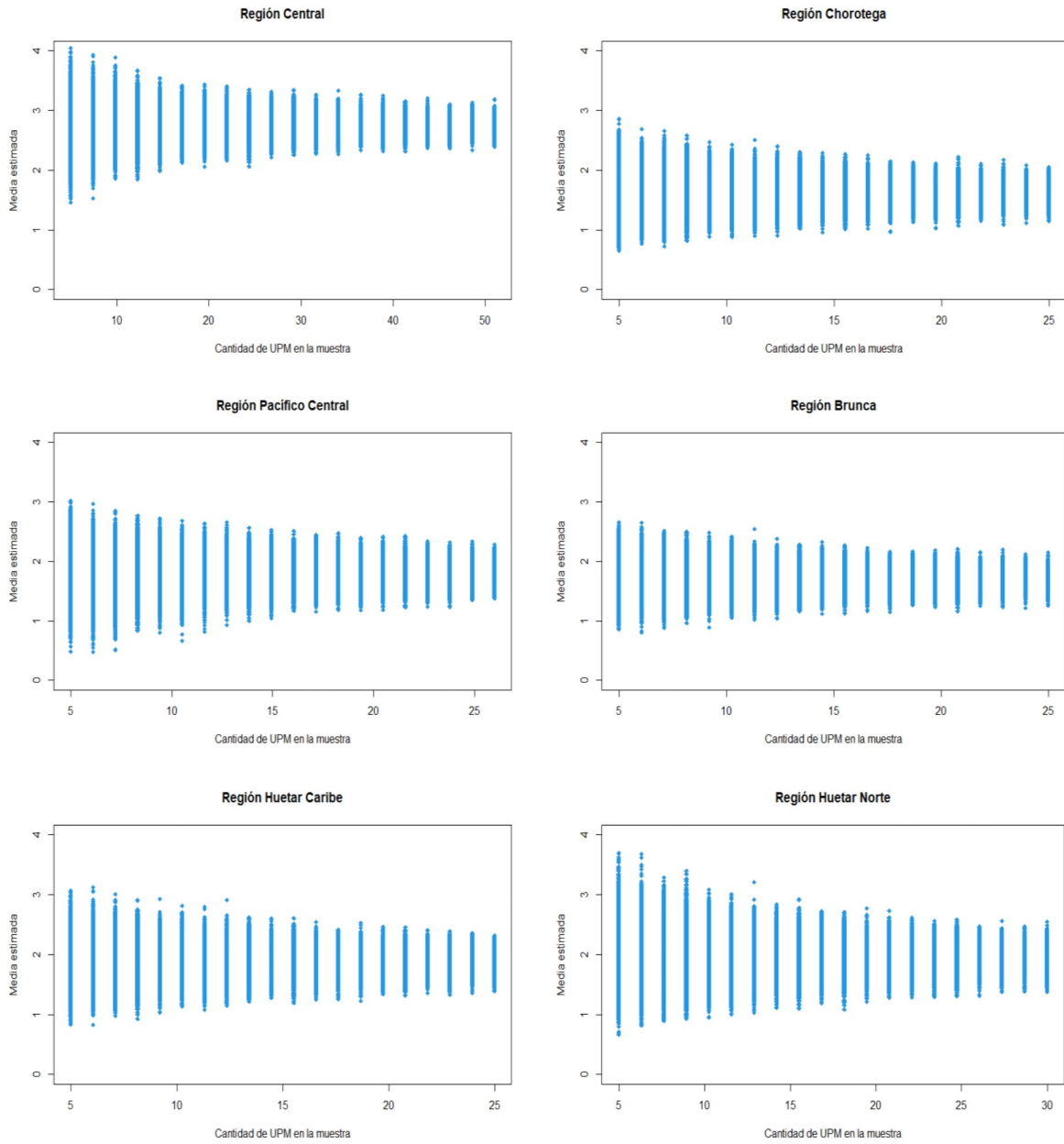


Figura 12: Comportamiento de las medias de la escala de productividad según número de UPM simuladas en cada región de planificación

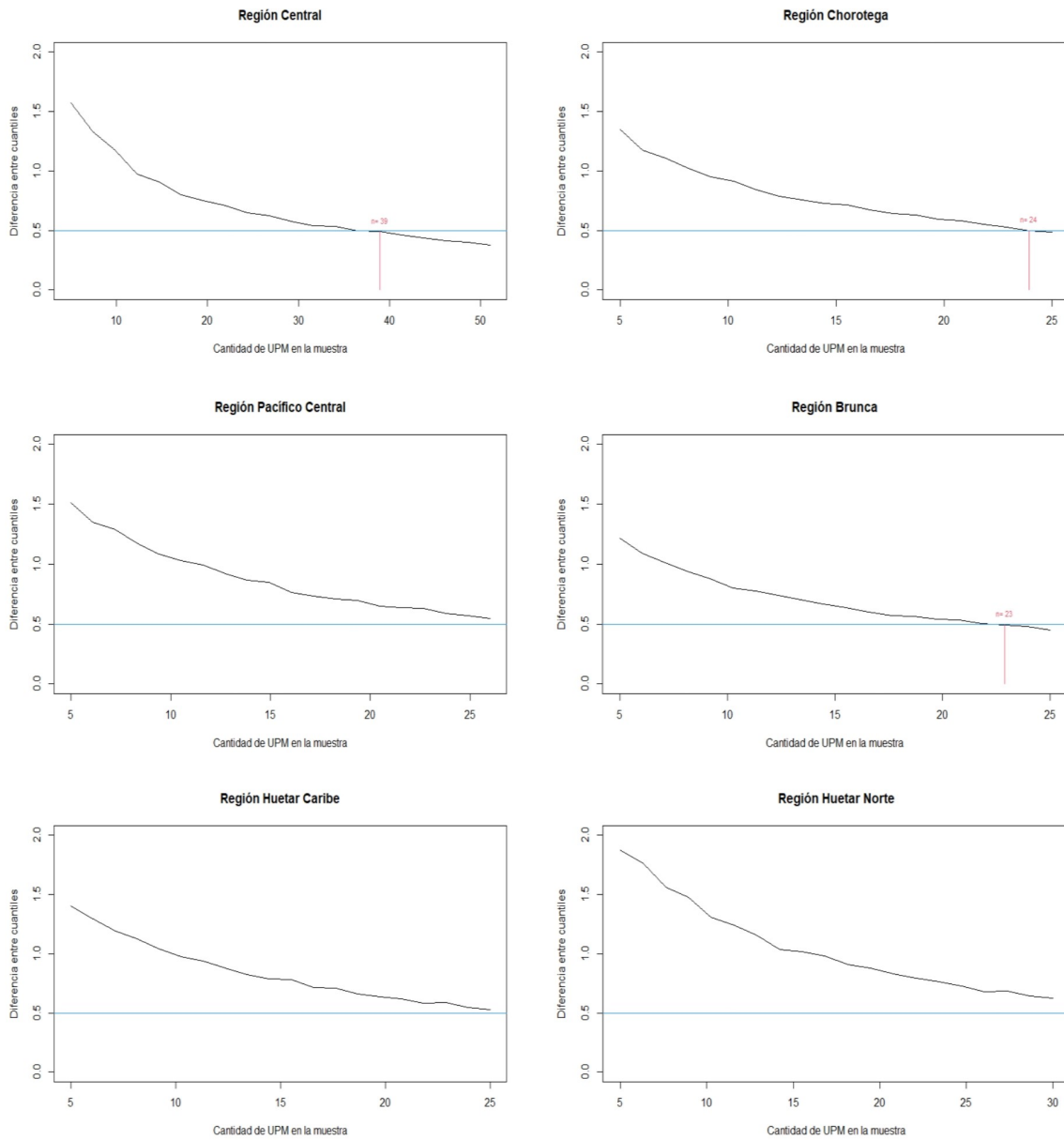


Figura 13: Diferencia entre los cuantiles de las medias de la escala de productividad según cantidad de UPM simuladas en cada región de planificación

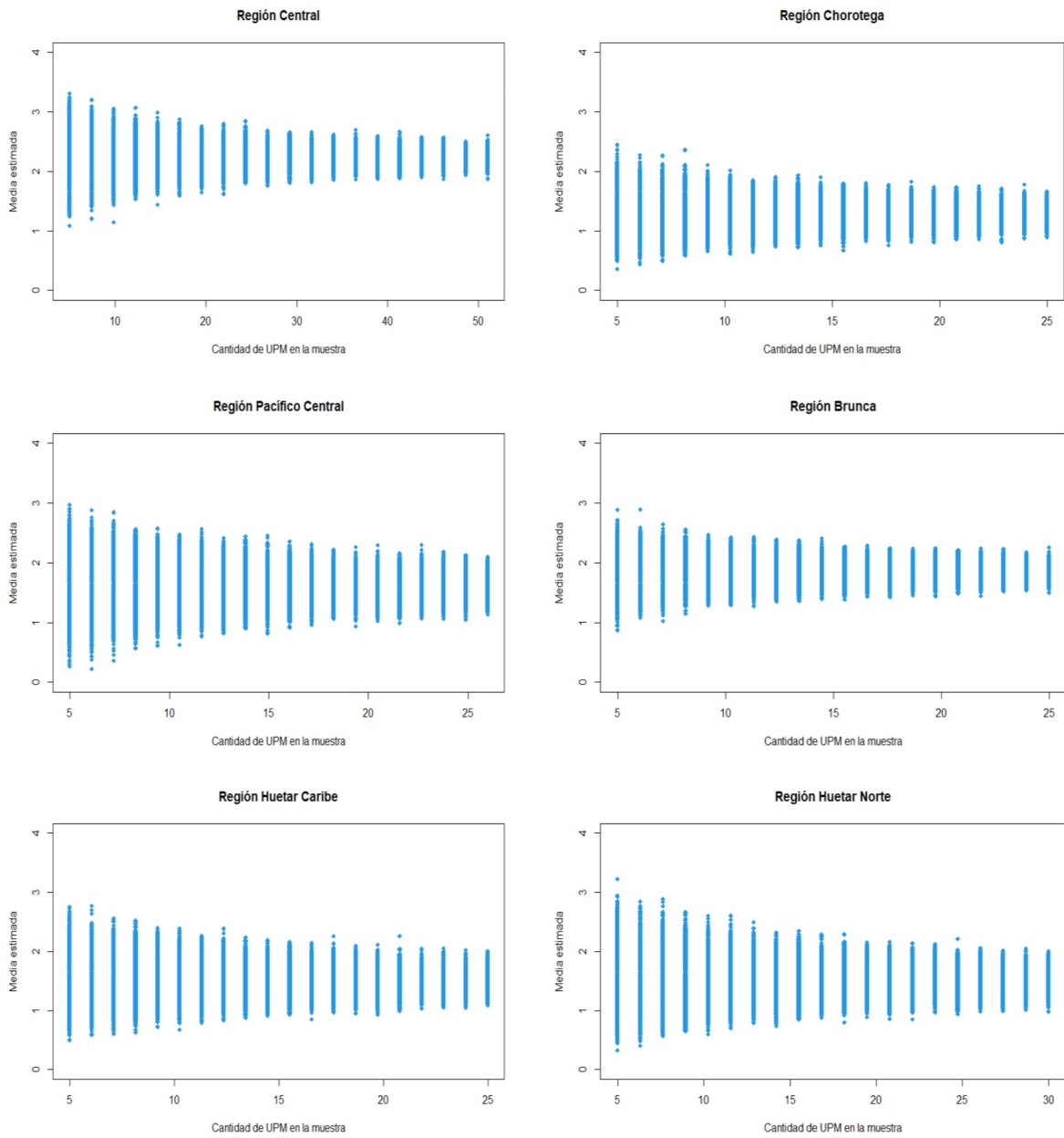


Figura 14: Comportamiento de las medias de la escala de habilitación ciudadana según número de UPM simuladas en cada región de planificación

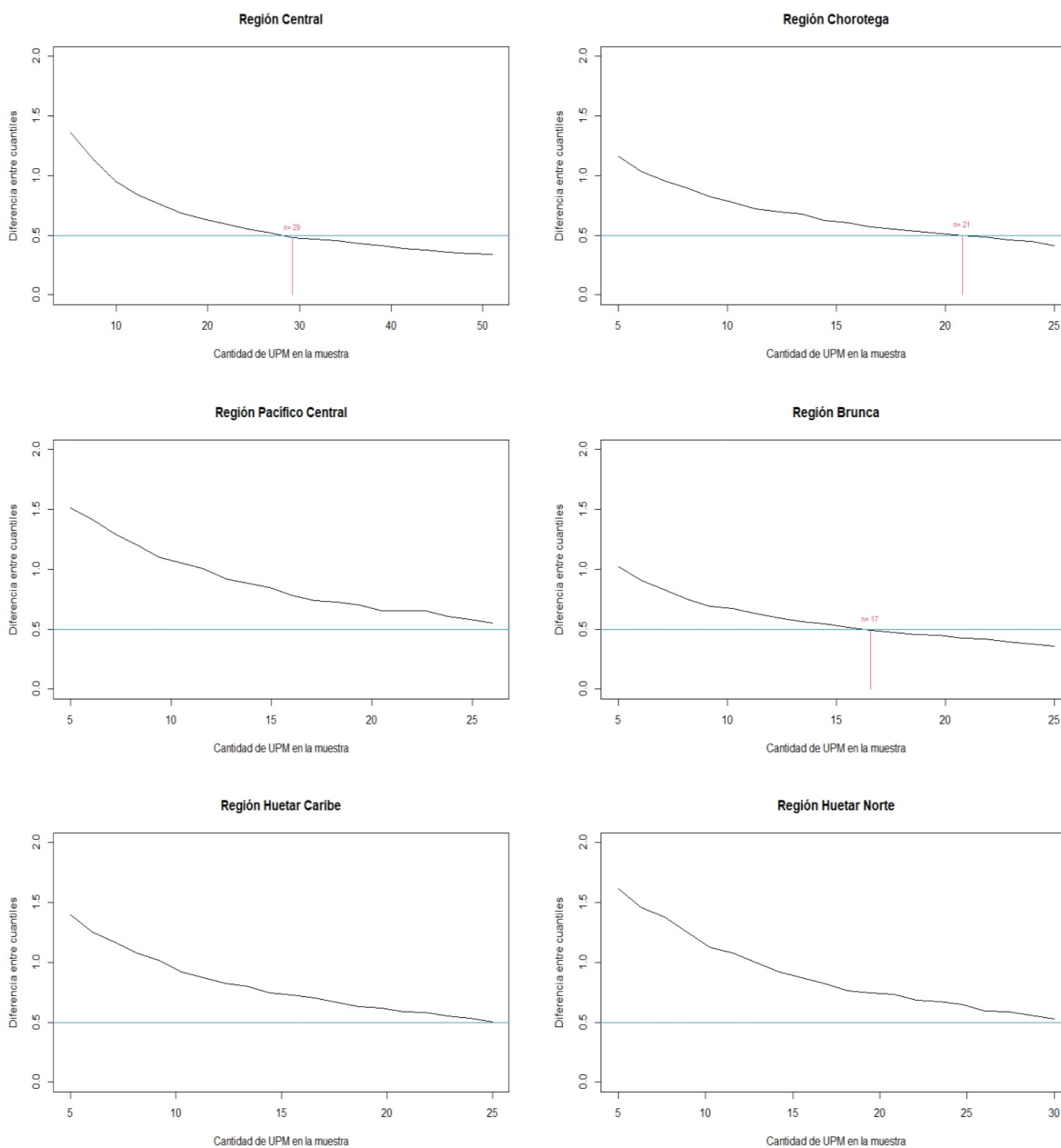


Figura 15: Diferencia entre los cuantiles de las medias de la escala de habilitación ciudadana según cantidad de UPM simuladas en cada región de planificación

Recibido: Diciembre de 2024  
 Aceptado: Julio de 2025

## Referencias

Asamblea Legislativa. Ley n 8642. ley general de telecomunicaciones, 2008.  
 Raúl Coss Bu. *Simulación: un enfoque práctico*. Limusa, 2003.  
 Rex B Kline. *Principles and practice of structural equation modeling*. Guilford publications, 2023.

- Robert V Krejcie and Daryle W Morgan. Determining sample size for research activities. *Educational and psychological measurement*, 30(3):607–610, 1970.
- Russell V Lenth. Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3):187–193, 2001.
- MICITT. Acceso y uso de los servicios de telecomunicaciones en costa rica 2017. Technical report, Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones, San José, Costa Rica, 2019.
- Christian P Robert, George Casella, and George Casella. *Introducing Monte Carlo methods with R*, volume 18. Springer, 2010.
- SUTEL. Estadísticas del sector de telecomunicaciones. informe 2017. Technical report, Superintendencia de Telecomunicaciones, 2017.