



Consejo Nacional de Evaluación
de la Política de Desarrollo Social

Nota Técnica 02/2006

**Pruebas de hipótesis estadísticas acerca de los
cambios en la incidencia de la pobreza
aplicando la metodología preliminar del CTMP
2000 – 2005**

Octubre de 2006

Av. Paseo de la Reforma #51, Piso 14, Col. Tabacalera
C. P. 06030, México, D. F.
www.coneval.gob.mx

Introducción

El pasado 29 de septiembre del presente año, el Instituto Nacional de Estadística, Geografía e Informática (INEGI) dio a conocer la nueva serie de bases de datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) correspondientes a los años 2000, 2002, 2004 y 2005. De acuerdo con el INEGI, la generación de esta nueva serie es producto de los ajustes a los ponderadores poblacionales de las encuestas, derivados de los resultados del Censo Nacional de Población 2005 y de las tareas de conciliación demográfica llevadas a cabo en conjunto con el Consejo Nacional de Población y con especialistas de El Colegio de México.

El Comité Directivo del CONEVAL aprobó realizar una nueva estimación de la pobreza para el periodo 2000-2005, utilizando la serie actualizada de la ENIGH y la propuesta metodológica del Comité Técnico para la Medición de la Pobreza (CTMP) de 2002, con el único propósito de permitir un análisis comparativo con estimaciones previas y de brindar información oportuna sobre la evolución de la pobreza en México, sin que esto constituya la metodología de medición oficial de la pobreza. El CONEVAL se encuentra actualmente desarrollando los trabajos para definir la metodología oficial para la medición de la pobreza en México.

Desde una perspectiva técnica, el CTMP recomendó que la medición de los cambios en la incidencia de la pobreza incluyera la aplicación de pruebas de hipótesis estadísticas. Esta recomendación obedece al hecho de que las estimaciones derivadas de la encuesta, como la ENIGH, están sujetas a errores y variaciones de muestreo, elementos que deben ser tomados en cuenta al comparar los cambios en el porcentaje de hogares o población pobre.

De acuerdo a lo anterior, en este documento se presentan los cálculos de la significancia estadística de los cambios en la proporciones de hogares y personas pobres entre los años 2000 y 2005, así como entre 2004 y 2005, según los resultados obtenidos por el CONEVAL al aplicar la metodología del CTMP. Las pruebas de hipótesis estadísticas fueron realizadas para la pobreza alimentaria, de capacidades y de patrimonio y a escala nacional, rural y urbana.

La evolución de la pobreza en hogares durante el quinquenio 2000-2005 se presenta en el cuadro 1, mientras que la información referente a la pobreza en personas durante el mismo periodo se detalla en el cuadro 2. Los cuadros 3 y 4 contienen la misma información sobre los cambios en la incidencia de pobreza, pero referidos al periodo 2004-2005.

Estructura de las pruebas de hipótesis

1. Planteamiento de hipótesis

Las pruebas de significancia requieren del planteamiento inicial de las hipótesis de trabajo que se desean contrastar. Las hipótesis a probar, de acuerdo a las estimaciones realizadas y tomando en cuenta los errores y variaciones de muestreo, son básicamente de dos tipos: hipótesis nula e hipótesis alternativa.

La hipótesis nula (H_0) es aquella que se formula en términos de igualdad; para el caso que nos ocupa esta hipótesis quedaría formulada para los dos periodos analizados 2000-2005 y 2004-2005 de la siguiente forma:

H_0 : "La proporción de la población (hogares) pobre(s) es la misma en los años 2000 (2004) y 2005."

La hipótesis alternativa (H_a) es la hipótesis de contraste, la cual se asume en caso de que se rechace la hipótesis nula. Para el caso que nos ocupa, esta hipótesis quedaría formulada como:

H_a : "Existe alguna diferencia en la proporción de la población (hogares) pobre(s) entre los años 2000 (2004) y 2005."

En términos formales el planteamiento de hipótesis quedaría como:

$$H_0: P_{2000} = P_{2005}$$
$$H_a: P_{2000} \neq P_{2005},$$

donde:

P_{2000} : es la proporción de personas pobres (o de hogares pobres) en la población del año 2000.

P_{2005} : es la proporción de personas pobres (o de hogares pobres) en la población del año 2005.

Para las pruebas entre los años 2004 y 2005 se substituye el subíndice 2000 por 2004.

2. Estadística de prueba

Dado que los parámetros P_{2000} y P_{2005} son estimados a partir de la muestra representada por la ENIGH y debido a que la selección de la muestra en los dos años se realiza de manera independiente, deberán hacerse las pruebas numéricas sobre la estadística:

$$Z = \frac{P_{2005} - P_{2000}}{\sqrt{s_{p_{2005}}^2 + s_{p_{2000}}^2}}$$

Los estadísticos p_{2000} y p_{2005} son los estimadores de los parámetros poblacionales respectivos. De acuerdo al teorema del límite central, cuando la hipótesis nula es verdadera, la estadística Z se distribuye asintóticamente conforme a una distribución normal estándar.

Las estadísticas muestrales $s_{p_{2000}}^2$ y $s_{p_{2005}}^2$ son los errores estándar al cuadrado de los estimadores de la incidencia, los cuales toman en cuenta el diseño de la muestra de las ENIGH 2000 y 2005, respectivamente.

El diseño de la muestra de la ENIGH es probabilístico, estratificado, polietápico y por conglomerados. El muestreo es probabilístico porque todas las unidades de muestreo tienen una probabilidad conocida y distinta de cero de ser seleccionadas. Es estratificado porque las unidades de muestreo con características similares de tipo geográficas y socioeconómicas se agrupan para formar estratos. Es polietápico porque la unidad última de selección (la vivienda) es seleccionada después de al menos dos etapas. Finalmente, es por conglomerados en virtud de que se conforman conjuntos de unidades muestrales de los cuales se obtiene la muestra.

La estimación de las estadísticas muestrales $s_{p_{2000}}^2$ y $s_{p_{2005}}^2$ se realiza mediante la aplicación del Método de Conglomerados Últimos. En la documentación metodológica de la ENIGH del año 2000 se señala que el término conglomerado último denota al conjunto de unidades seleccionadas como Unidades Primarias de Muestreo (UPM)¹ en cada uno de los estratos en que se divide a la población de estudio. El método está basado en la expansión en series de Taylor de la varianza de un cociente de variables aleatorias; las personas interesadas pueden consultar las fórmulas que se utilizan para generar los errores estándar en la documentación metodológica de la ENIGH o recurrir a algunos textos especializados en la materia.² En la práctica, las estimaciones fueron generadas utilizando la versión 9 del paquete Stata.

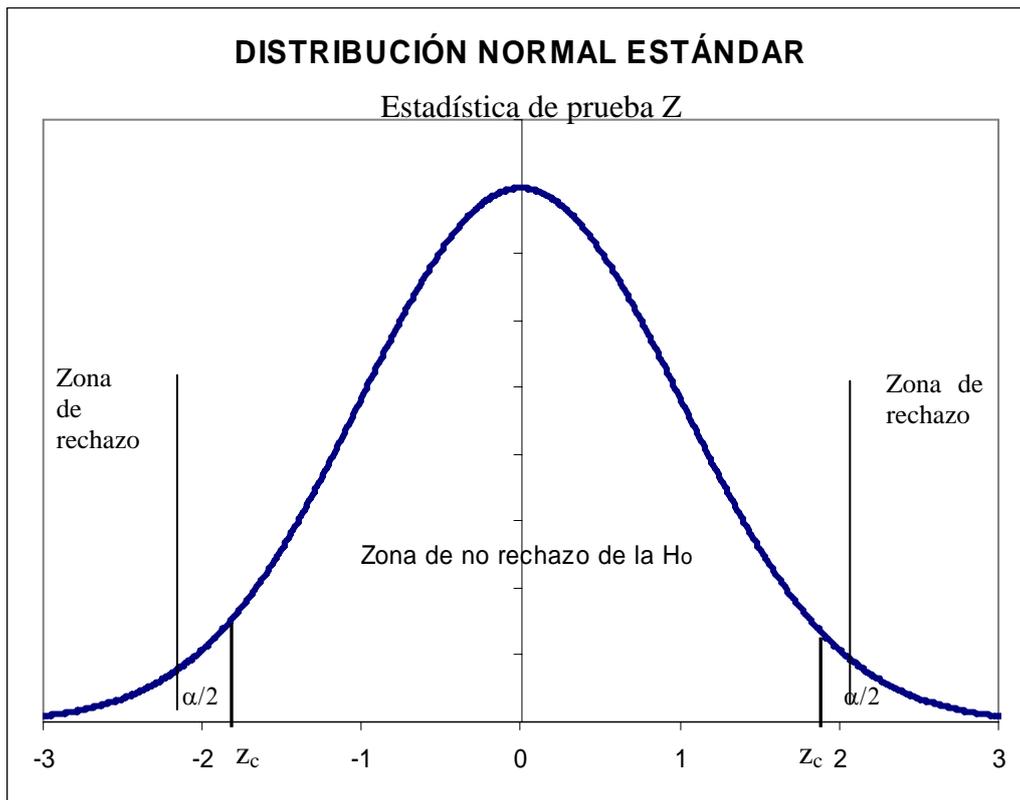
¹ Las unidades primarias de muestreo están constituidas por un Área Geostadística Básica (AGEB), una parte de una AGEB o por varias AGEB colindantes.

² Véase Lohr, Sharon L. (2000) *Muestreo: diseño y análisis*, editorial Thomson o el texto clásico de William Cochran (1977), *Técnicas de muestreo*, editorial CECSA.

3. Significancia estadística

El contraste de hipótesis se realiza considerando diferentes probabilidades de cometer errores al momento de rechazar la hipótesis nula, siendo ésta verdadera; a estos errores se les conoce como errores tipo I. En la literatura se refiere como α a la probabilidad de cometer este tipo de error, la cual se denomina nivel de significancia de la prueba.

Se considera pertinente plantear *a priori* algunas consideraciones analíticas sobre los posibles cambios de la pobreza en dos momentos diferentes. Estos cambios pueden ocurrir en cualquiera de dos sentidos y reflejarse ya sea en un aumento o en una disminución de los niveles de pobreza. Esta posibilidad queda especificada en la formulación de la hipótesis alternativa H_a . Cuando esto se asume, se dice que se realizarán "pruebas de dos colas"; por tal motivo, la zona de rechazo de la hipótesis nula se ubica en los dos extremos de la distribución, tal como se muestra en el siguiente dibujo.



Al estimar la probabilidad de cometer el error tipo I se calcula, bajo la distribución normal estándar, la probabilidad $\Pr [Z > |z_c|]$. Si esta probabilidad es menor de α entonces se rechaza la hipótesis nula, para lo cual se dice que el cambio es estadísticamente significativo; se rechaza la hipótesis nula a favor de la hipótesis alternativa.

El **no** rechazar la hipótesis nula implica la falta de evidencia estadística, dado el nivel de significancia seleccionado, para rechazar la hipótesis nula de igualdad. Esto puede deberse principalmente a la variación de los estimadores originados, entre otras cosas, por el diseño de muestreo. En tales casos, la evidencia empírica obtenida a partir de las muestras no permite *inferir* ninguna diferencia entre los parámetros poblacionales en los dos momentos en el tiempo.

Cabe señalar que el nivel de significancia α es relevante para el rechazo o no de la hipótesis nula. Mayores valores de α aumentarán la probabilidad de rechazo de la hipótesis nula, pero al mismo tiempo aumentará la probabilidad de cometer el error de rechazar la hipótesis nula cuando ésta es verdadera. El nivel de significancia α es determinado por el investigador, quien elige qué nivel de error está dispuesto a aceptar. El nivel de significancia estadística más comúnmente empleado es $\alpha = 0.05$ y es el que el CONEVAL ha utilizado en las pruebas de hipótesis que se reportan en este documento.

Descripción de resultados en las pruebas de significancia

Periodo 2000-2005

Dado el nivel de significancia seleccionado, los resultados permiten concluir que tanto a nivel nacional, como en las zonas urbanas y rurales, se ha presentado una disminución estadísticamente significativa entre 2000 y 2005 para cada uno de los tres tipos de pobreza: alimentaria, de capacidades y de patrimonio.

Periodo 2004-2005

Por su parte, las pruebas de hipótesis estadísticas aplicadas a los cambios en la pobreza experimentados entre 2004 y 2005 indican que no se cuenta con evidencia suficiente para concluir que en el país en su conjunto los cambios observados sean significativos; tampoco lo son las variaciones en las zonas rurales. Lo contrario sucede en las zonas urbanas, donde se observó una reducción estadísticamente significativa de la pobreza alimentaria, de capacidades y de patrimonio en hogares y de la pobreza de capacidades y de patrimonio en personas.

Cuadro 1
Cambio en la incidencia de la pobreza, México, 2000-2005
Hogares

Ámbito y tipo de pobreza	Incidencia de la pobreza ¹		Errores estándar ¹		Cambio en la incidencia	Error estándar de la diferencia	Estadística z	Nivel de significancia para la diferencia (dos colas)	Conclusión sobre la significancia de la diferencia ²
	2000	2005	2000	2005	P ₂₀₀₅ - P ₂₀₀₀				
Nacional									
Alimentaria	18.5	14.1	0.7858	0.4868	-4.466	0.924	-4.832	0.0000	Significativa
Capacidades	25.2	19.4	0.9218	0.5661	-5.790	1.082	-5.352	0.0000	Significativa
Patrimonio	45.7	39.6	1.1609	0.6162	-6.135	1.314	-4.668	0.0000	Significativa
Rural									
Alimentaria	34.1	26.1	1.7240	1.2699	-7.953	2.141	-3.714	0.0002	Significativa
Capacidades	41.3	32.9	1.9004	1.5249	-8.462	2.437	-3.473	0.0005	Significativa
Patrimonio	60.7	53.9	2.0538	1.5260	-6.817	2.559	-2.664	0.0077	Significativa
Urbana									
Alimentaria	9.7	7.7	0.7720	0.3666	-2.039	0.855	-2.386	0.0170	Significativa
Capacidades	16.1	12.4	1.0083	0.4386	-3.774	1.100	-3.432	0.0006	Significativa
Patrimonio	37.3	32.0	1.4344	0.5978	-5.214	1.554	-3.355	0.0008	Significativa

¹ La información está presentada en porcentaje.

² Significancia estadística del cambio en la incidencia de la pobreza en hogares ($\alpha = 0.05$).

Fuente: Estimaciones del Coneval con base en las ENIGH 2000 y 2005.

Cuadro 2
Cambio en la incidencia de la pobreza, México, 2000-2005
Personas

Ámbito y tipo de pobreza	Incidencia de la pobreza ¹		Errores estándar ¹		Cambio en la incidencia	Error estándar de la diferencia	Estadística z	Nivel de significancia para la diferencia (dos colas)	Conclusión sobre la significancia de la diferencia ²
	2000	2005	2000	2005	P ₂₀₀₅ - P ₂₀₀₀				
Nacional									
Alimentaria	24.1	18.2	1.0448	0.6450	-5.893	1.228	-4.800	0.0000	Significativa
Capacidades	31.8	24.7	1.1693	0.7011	-7.055	1.363	-5.174	0.0000	Significativa
Patrimonio	53.6	47.0	1.2490	0.6958	-6.561	1.430	-4.589	0.0000	Significativa
Rural									
Alimentaria	42.4	32.3	2.1017	1.5024	-10.078	2.583	-3.901	0.0001	Significativa
Capacidades	49.9	39.8	2.2144	1.6577	-10.116	2.766	-3.657	0.0003	Significativa
Patrimonio	69.2	61.8	1.9048	1.4327	-7.423	2.383	-3.115	0.0018	Significativa
Urbana									
Alimentaria	12.5	9.9	1.0085	0.4649	-2.550	1.110	-2.296	0.0216	Significativa
Capacidades	20.2	15.8	1.2751	0.5515	-4.379	1.389	-3.152	0.0016	Significativa
Patrimonio	43.7	38.3	1.6077	0.7248	-5.306	1.764	-3.009	0.0026	Significativa

¹ La información está presentada en porcentaje.

² Significancia estadística del cambio en la incidencia de la pobreza en personas ($\alpha = 0.05$).

Fuente: Estimaciones del Coneval con base en las ENIGH 2000 y 2005.

Cuadro 3
Cambio en la incidencia de la pobreza, México, 2004-2005
Hogares

Ámbito y tipo de pobreza	Incidencia de la pobreza ¹		Errores estándar ¹		Cambio en la incidencia	Error estándar de la diferencia	Estadística z	Nivel de significancia para la diferencia (dos colas)	Conclusión sobre la significancia de la diferencia ²
	2004	2005	2004	2005	P ₂₀₀₅ - P ₂₀₀₄				
Nacional									
Alimentaria	13.8	14.1	0.5397	0.4868	0.226	0.727	0.311	0.7558	No significativa
Capacidades	19.9	19.4	0.6706	0.5661	-0.462	0.878	-0.527	0.5982	No significativa
Patrimonio	39.7	39.6	0.8615	0.6162	-0.142	1.059	-0.134	0.8935	No significativa
Rural									
Alimentaria	22.9	26.1	1.4947	1.2699	3.261	1.961	1.663	0.0964	No significativa
Capacidades	29.9	32.9	1.9139	1.5249	2.938	2.447	1.200	0.2300	No significativa
Patrimonio	49.3	53.9	2.3255	1.5260	4.577	2.781	1.646	0.0998	No significativa
Urbana									
Alimentaria	8.7	7.7	0.3788	0.3666	-1.041	0.527	-1.976	0.0482	Significativa
Capacidades	14.2	12.4	0.4220	0.4386	-1.886	0.609	-3.099	0.0019	Significativa
Patrimonio	34.3	32.0	0.5672	0.5978	-2.281	0.824	-2.767	0.0056	Significativa

¹ La información está presentada en porcentaje.

² Significancia estadística del cambio en la incidencia de la pobreza en hogares ($\alpha = 0.05$).

Fuente: Estimaciones del Coneval con base en las ENIGH 2004 y 2005.

Cuadro 4
Cambio en la incidencia de la pobreza, México, 2004-2005
Personas

Ámbito y tipo de pobreza	Incidencia de la pobreza ¹		Errores estándar ¹		Cambio en la incidencia	Error estándar de la diferencia	Estadística z	Nivel de significancia para la diferencia (dos colas)	Conclusión sobre la significancia de la diferencia ²
	2004	2005	2004	2005	P ₂₀₀₅ - P ₂₀₀₄				
Nacional									
Alimentaria	17.4	18.2	0.7660	0.6450	0.842	1.001	0.841	0.4003	No significativa
Capacidades	24.7	24.7	0.9014	0.7011	0.001	1.142	0.001	0.9993	No significativa
Patrimonio	47.2	47.0	0.9085	0.6958	-0.169	1.144	-0.148	0.8825	No significativa
Rural									
Alimentaria	28.0	32.3	1.9970	1.5024	4.254	2.499	1.702	0.0887	No significativa
Capacidades	36.2	39.8	2.4209	1.6577	3.617	2.934	1.233	0.2177	No significativa
Patrimonio	57.4	61.8	2.3035	1.4327	4.429	2.713	1.633	0.1026	No significativa
Urbana									
Alimentaria	11.0	9.9	0.5017	0.4649	-1.070	0.684	-1.564	0.1177	No significativa
Capacidades	17.8	15.8	0.5497	0.5515	-2.023	0.779	-2.598	0.0094	Significativa
Patrimonio	41.1	38.3	0.6584	0.7248	-2.786	0.979	-2.845	0.0044	Significativa

¹ La información está presentada en porcentaje.

² Significancia estadística del cambio en la incidencia de la pobreza en personas ($\alpha = 0.05$).

Fuente: Estimaciones del Coneval con base en las ENIGH 2004 y 2005.