

Capítulo 6

Aplicación y resultados

En este capítulo se presenta la aplicación del *jackknife* y Repeticiones Balanceadas, para calcular la varianza de varios estimadores y los intervalos de confianza asociados a ellos. Las estadísticas de interés, como se mencionó en el Capítulo 5, se refieren a la Encuesta Nacional de Alimentación en el Medio Rural, 1996 (ENAL'96). Se comienza por la exposición de los estimadores utilizados y posteriormente se muestran e interpretan los resultados.

6.1 Estimadores

La fase aplicativa de esta tesis consiste en la obtención de la varianza de estimadores de razón y de la mediana de los indicadores PEDZ, PETZ y TEDZ. Los estimadores de razón en cuestión, se refieren a la proporción de niños en cada uno de los cinco grupos señalados en el Capítulo 5, por estado, región y a nivel nacional. Las estimaciones de varianzas (de los estimadores de razón) por estado y nacional, se realizaron a través del *jackknife* y por medio de las fórmulas conocidas que aproximan la estimación del error cuadrático medio (ECM), que se derivan de la linearización. Así también, en los estados donde los estratos consisten de únicamente dos localidades o unidades primarias de muestreo, se aplicó el método de Repeticiones Balanceadas (*Balanced Half-Sampling*) con la adecuación de Fay y $\lambda = 0.5$; también se calculó con otros valores de λ (0.3 y 0.7) y sin tal adecuación, pero todos los resultados fueron tan similares que se decidió mostrar solamente un caso ($\lambda = 0.5$). Los estimadores por regiones se obtuvieron sólo por *jackknife*.

Es importante aclarar que la estimación de la varianza de la mediana tiene un carácter exploratorio, ya que como se indicó en 3.2.4., el *jackknife* es inconsistente para la varianza de las estadísticas de orden (y en general las no suaves), bajo un m.a.s. y población infinita, pero prácticamente no se conoce su comportamiento para estas estadísticas, en un muestreo complejo. Es por ello que solamente se exhiben las estimaciones de la mediana y su varianza en los estados en los que resultó factible aplicar Repeticiones Balanceadas para lograr, de esta manera, cierto punto de comparación. (No se aplicó el método de Repeticiones Balanceadas por región pues no hubo ninguna donde todos los estados tuvieran dos localidades muestreadas en cada estrato). Vale la pena añadir que hay que tener presente que las verdaderas varianzas son desconocidas, por lo que simplemente es posible

apreciar cómo se asemejan o se alejan las varianzas de medianas *jackknife* y Repeticiones Balanceadas, contando también con la referencia de sus similitudes para los estimadores de razón y la media; siendo que ésta última se incluyó con el objetivo de ampliar el marco de referencia.

6.1.1 Estimadores de razón

La proporción de niños en determinada categoría se expresa como el total estimado de niños clasificados en ella, entre el total estimado de niños en la edad de interés que viven en comunidades rurales; o bien, como el promedio (estimado) por vivienda rural de niños menores de seis años en la categoría de interés, entre el promedio (estimado) por vivienda rural de niños de dicha edad. Como no se conoce el tamaño de dicha población, se requiere estimar igualmente el denominador, lo que conlleva a que el estimador de interés sea de razón. A continuación se da cierta nomenclatura¹ para posteriormente, dar las expresiones de los estimadores. Como se realizan estimaciones por estado, debe quedar claro que los subíndices de la nomenclatura que sigue, en realidad corren de esta manera: $h = 1, \dots, L^E$, donde L^E es el número de estratos en el estado E ; $i = 1, \dots, N_h^E$, donde N_h^E es el total de localidades en el estrato h del estado E ; $j = 1, \dots, M_{hi}^E$, donde M_{hi}^E es el total de viviendas en la localidad i -ésima del estrato h , del estado E . Sin embargo, para simplificar un poco la notación se consideró que se podía obviar el superíndice E en las expresiones de los estimadores que siguen en esta sección y la 6.1.2.

y_{hi}^c El número de niños menores de seis años en la familia j -ésima de la localidad i -ésima del estrato h , que son clasificados en la categoría c .

x_{hi} El número de niños menores de seis años en la familia j -ésima de la localidad i -ésima del estrato h .

xe_{hij} El número de niños menores de seis años **encuestados** en la familia j -ésima de la localidad i -ésima del estrato h

$VIVENC_{hi} = m_{hi}$ Viviendas encuestadas en la localidad i -ésima del estrato h .

$LOCENC_h = n_h$ Localidades encuestadas en el estrato h .

¹La razón por la que algunos conceptos tienen ser relación con los nombres de variables en la base de datos y también la relación con la nomenclatura propia usual en la Teoría del muestreo.

$VIVIEN_{hi} = M_{hi}$ Total de viviendas en la localidad i -ésima del estrato h .

$LOCXESTR_h = N_h$ Total de localidades rurales que reúnen los requisitos de inclusión en el estrato h . ($VIVIEN_{hi}$ y $LOCXESTR_h$ son conocidas.)

$VIVESTR_h, VIVEDO$ Total de viviendas en el estrato y en el estado respectivamente. (Valores conocidos)

W_h Peso de un estrato en un estado (Se usó $W_h = VIVESTR_h/VIVEDO_E$)

$F_{casa_{hij}} = \frac{x_{hij}}{xe_{hij}}$ Factor de expansión por la tercera etapa de muestreo implícita en el diseño

$FVIV_{hi} = \frac{VIVIEN_{hi}}{VIVENC_{hi}}$ Factor de expansión por vivienda (Segunda etapa de muestreo).

$FLOC_h = \frac{LOCXESTR_h}{LOCENC_h}$ Factor de expansión por localidad (Primera etapa de muestreo).

$POBRURAL_E$ Población rural en el Estado.

POB Población rural en la República Mexicana. ($POBRURAL_E$ y POB conocidos.)

$\Pi_E = \frac{POBRURAL_E}{POB}$ Peso del Estado E respecto a población rural en el Estado y Nacional.

Se decidió utilizar un estimador de razón combinado a nivel *estado*, y un estimador separado a nivel *Nacional* cuya formulación se basa en dos etapas de estratificación, pues para este efecto, los Estados son considerados estratos. Cochran (1977, pág. 167), explica que cuando la razón en cuestión no es constante de estrato a estrato, el estimador separado es más preciso pues su sesgo se desvanece. Por otra parte, también recomienda el uso del estimador *combinado* cuando la muestra es chica y del *separado* cuando la muestra en cada estrato es grande. Con base en estas recomendaciones, y bajo la sospecha de que los estimadores tuvieran mayores diferencias entre estados que dentro de los estados, se consideró adecuado el uniformar el método de estimación por estado y a nivel nacional como se indicó al inicio de este párrafo. Por otra parte, a nivel región el estimador que se aplicó fue el combinado, puesto que era de esperar que los estimadores de los estados

vecinos se comportaran de forma similar y además la muestra de cada estrato seguía siendo pequeña.

El estimador de razón combinado para la proporción de niños menores de seis años que viven en comunidades rurales y son clasificados en la categoría c , en un Estado E (o región)², de acuerdo al diseño muestral de la encuesta está dado por:

$$\begin{aligned}\hat{R}_E^c &= \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} FLO C_h FVIV_{hi} Fcasa_{hij} y_{hij}^c}{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} FLO C_h FVIV_{hi} x_{hij}} \quad (6.1) \\ &= \frac{\frac{1}{VIVEDO_E} \left\{ \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} FLO C_h FVIV_{hi} Fcasa_{hij} y_{hij}^c \right\}}{\frac{1}{VIVEDO_B} \left\{ \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} FLO C_h FVIV_{hi} x_{hij} \right\}} \\ &= \frac{\bar{y}}{\bar{x}}\end{aligned}$$

Hansen, Hurwitz y Madow (1953), dan la expresión anterior pero en Sukhatme y cols. (1984), se encuentra una forma del estimador en la que son explícitos los pesos de los estratos, W_h , de la siguiente forma³:

$$\hat{R}_E^c = \frac{\sum_{h=1}^L W_h \bar{y}_{est}}{\sum_{h=1}^L W_h \bar{x}_{est}} \quad (6.2)$$

Siguiendo la misma nomenclatura y suponiendo que el peso de los estratos de un estado está dado en función de las viviendas; es decir, $W_h = \frac{\text{Viviendas rurales en el estrato}}{\text{Viviendas rurales en el Estado}}$, se deriva que⁴:

$$\begin{aligned}\sum_{h=1}^L W_h \bar{y}_{est} &= \sum_{h=1}^L W_h \left[\frac{1}{LOCENC_h} \sum_{i=1}^{n_h} \frac{VIVIEN_{hi}}{VIVESTR_h} LOCXESTR_h \bar{y}_{hi} \right] \quad (6.3) \\ &= \frac{1}{VIVEDO_E} \left[\sum_{h=1}^L \frac{LOCXESTR_h}{LOCENC_h} \sum_{i=1}^{n_h} \frac{VIVIEN_{hi}}{VIVESTR_h} VIVESTR_h \bar{y}_{hi} \right] \\ &= \frac{1}{VIVEDO_E} \left[\sum_{h=1}^L FLO C_h \sum_{i=1}^{n_h} \frac{VIVIEN_{hi}}{VIVENC_{hi}} \sum_{j=1}^{m_{hi}} Fcasa_{hij} y_{hij}^c \right] \\ &= \frac{1}{VIVEDO_E} \left[\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} FLO C_h FVIV_{hi} Fcasa_{hij} y_{hij}^c \right]\end{aligned}$$

El denominador se puede desarrollar de igual forma, recordando que no se considera en éste la tercera etapa de selección (Por lo que no se incluye el factor $Fcasa_{hij}$) y se llega

²El estimador para una región tiene la misma expresión, pero en tal caso se consideran el total de estratos en todos los estados que conforman la región, $L = \sum L_E$.

³La definición de \bar{y}_{est} se encuentra implícita en el desarrollo de (6.3).

⁴Nótese que en el desarrollo de (6.3) se encuentran implícitas las definiciones de \bar{y}_{est} y \bar{y}_{hi} .

a que ambas expresiones son iguales si el peso de los estratos se da en términos de los elementos para la segunda etapa de muestreo, que en este caso fueron las viviendas. Se debe observar que (6.1) representa la razón de las medias de y y x , por vivienda. Por otro lado, (6.2) se podría obtener con pesos W_h que fuesen función de otra variable, como por ejemplo, población rural. Sin embargo, en tal caso los dos estimadores no coincidirían. Cabe resaltar que los estimadores que se expresan en términos de los factores de expansión son de construcción más lógica y operativamente son más sencillos de aplicar.

El estimador de razón separado para la proporción a nivel Nacional de niños menores de seis años que viven en comunidades rurales y son clasificados en la categoría c , está dado por:

$$R_{MEX}^c = \sum_{E=1}^{31} \Pi_E R_E^c \quad (6.4)$$

Se observa que de acuerdo a la nomenclatura antes expuesta, los pesos de los estados se basan en la población rural. Igualmente, se pudo haber tomado otra variable para este efecto. Por ejemplo, era factible dar los pesos por viviendas rurales, pero de existir diferencias en número de habitantes por vivienda en los estados, entonces resultaba más adecuado considerar la población rural directamente por lo cual se decidió usar este dato como ponderador. Cabría la duda de obtener este estimador separado que pondera los estados o considerar uno similar que se basara en regiones, en lugar de estados. Si se examinan los pesos en (6.4), se debe notar que a cada estrato que pertenece a un estado en particular, se le atribuye un peso a nivel nacional de Π_E/L_E , lo cual, en la mayoría de los casos, debe ser una buena aproximación a su peso verdadero. Si se considerara un estimador separado por regiones, como el que se mencionó anteriormente, en realidad se tendrían cinco grandes estratos (en la expresión del estimador) y se estaría igualando el peso de todos los estratos de una región. De tal suerte, se consideró más adecuado y menos riesgoso incluir la mayor cantidad posible de estratos a nivel nacional (en la expresión del estimador), sin llegar a los 372 que en verdad se dieron (lo que evitó una programación especial del *jackknife* a nivel nacional, además del programa que hace estimaciones por estado). En resumen, en la elección de este estimador, se involucró la filosofía de simplificar los métodos guardando la mayor coherencia posible.

En el Capítulo 3 se mostró que el método de linealización produce el estimador de la varianza para un estimador de razón que usualmente se ve en libros de texto. Así pues el estimador de la varianza de (6.1) de acuerdo a esta aproximación, y según lo plantean Hansen, Hurwitz y Madow (1953)⁵, se encuentra de la siguiente forma:

$$\widehat{Var}(\hat{R}_E^c) = \frac{1}{\bar{X}^2} \sum_{h=1}^L \frac{LOCXESTR_h^2}{LOCENC_h} s_{ch}^2$$

⁵Para quien revise esta referencia se advierte que el desarrollo presentado por los autores citados se avoca al caso de la "varianza relativa" que se define como la varianza del coeficiente de variación al cuadrado. Pero de este estimador se deriva con facilidad el estimador de la varianza del estimador de razón. El uso de la varianza relativa radica en que es una estadística transportable ("portable"), según Kish (1977)

donde, $s_{ch}^2 = s_{chY}^2 + (\hat{R}_E^c)^2 s_{chX}^2 - 2\hat{R}_E^c S_{chYX}$

$$s_{chY}^2 = \frac{\sum_{i=1}^{n_h} \left(y'_{hi} - \frac{\sum_i^{n_h} y'_{hi}}{n_h} \right)^2}{n_h - 1}$$

$$y'_{hi} = FVIV_{hi} \left(\sum_{j=1}^{m_{hij}} Fcasa_{hij} y_{hij} \right)$$

$$x'_{hi} = FVIV_{hi} \left(\sum_{j=1}^{m_{hij}} xe_{hij} \right)$$

$$\hat{X} = \sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} FLOC_h FVIV_{hi} x_{hij}$$

$$s_{chYX} = \frac{\sum_{i=1}^{n_h} \left(y'_{hi} - \frac{\sum_i^{n_h} y'_{hi}}{n_h} \right) \left(x'_{hi} - \frac{\sum_i^{n_h} x'_{hi}}{n_h} \right)}{n_h - 1}$$

Por otra parte, el estimador de la varianza de (6.4) se da en función de las varianzas en cada estado, de la siguiente manera:

$$\widehat{Var}(R_{ME}^c) = \sum_{E=1}^{31} \Pi_E^2 \widehat{Var}(\hat{R}_E^c) \quad (6.5)$$

Se aprecia en la expresión anterior que una vez que se consiguen los estimadores de varianza por estado, la estimación a nivel nacional es fácil de obtener, pues los pesos son conocidos. En realidad no se aplicó un *jackknife* a nivel nacional para obtener las varianzas, sino que se obtuvo la ponderación de las varianzas de los estados, $\hat{V}_{E_j}(\hat{\Theta}_E)$, de esta manera:

$$\begin{aligned} \hat{V}_{N_j}(\hat{\Theta}) &= \sum_{E=1}^{31} \Pi_E^2 \hat{V}_{E_j}(\hat{\Theta}_E) \\ &= \sum_{E=1}^{31} \Pi_E^2 \left\{ \sum_{h=1}^{L_E} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\hat{\Theta}_{E(hi)} - \hat{\Theta}_E)^2 \right\} \\ &= \sum_{E=1}^{31} \left\{ \sum_{h=1}^{L_E} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\Pi \hat{\Theta}_{E(hi)} - \Pi \hat{\Theta}_E)^2 \right\} \end{aligned} \quad (6.6)$$

donde, $\hat{\Theta}_E$ es el estimador del estado; $\hat{\Theta}_E$ es el promedio de las iteraciones *jackknife* en el estado E lo cual representa otro estimador a nivel estado y $\hat{\Theta}_{E(hi)}$ es el estimador del estado E , al quitar la localidad i -ésima del estrato h . Ahora, interesa saber, cómo compara

esta expresión a la de un estimador *jackknife* basado en tantas iteraciones como localidades en la Nación. Para establecer tal comparación es conveniente revisar las expresiones de varios estimadores. En primer lugar, de acuerdo a la definición del estimador separado en (6.4), el estimador nacional sería:

$$\hat{\Theta}_N = \sum_{E=1}^{31} \Pi \hat{\Theta}_E. \quad (6.7)$$

Alguien podría sugerir otro estimador nacional que se basara en la ponderación de los estimadores *jackknife*, de la siguiente forma:

$$\hat{\Theta}_N = \sum_{E=1}^{31} \Pi \hat{\Theta}_E. \quad (6.8)$$

O bien, siendo muy puntuales, se podría pedir el estimador *jackknife* dado por el promedio de todas las iteraciones a nivel nacional, en cuyo caso, el estimador nacional sería :

$$\hat{\Theta}_{N_J} = \frac{\sum_{E=1}^3 1 \sum_{h=1}^{L_E} \sum_{i=1}^{n_h} (\sum_{e \neq E} \Pi_e \hat{\Theta}_e + \Pi_E \hat{\Theta}_{E(hi)})}{(\text{Localidades encuestadas})} \quad (6.9)$$

Ahora bien, si se realizara un *jackknife* para todo el País, la expresión de la varianza sería:

$$\sum_{E=1}^{31} \sum_{h=1}^{L_E} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\hat{\Theta}_{E(hi)} - A)^2 \quad (6.10)$$

donde A pudiera ser alguna de las expresiones en (6.9), (6.8) o (6.7).

Por otra parte, ocurre que el estimador nacional, al quitar la localidad i -ésima es:

$$\hat{\Theta}_{N_{E(hi)}} = \sum_{e \neq E} \Pi_e \hat{\Theta}_e + \Pi_E \hat{\Theta}_{E(hi)}. \quad (6.11)$$

Con estos elementos, se puede volver a revisar la expresión en (6.6). Sumando y restando el primer término de (6.11) dentro del paréntesis de (6.6), se tiene:

$$\begin{aligned} \hat{V}_{N_J} &= \sum_{E=1}^{31} \sum_{h=1}^{L_E} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} \left(\left(\Pi_E \hat{\Theta}_{E(hi)} + \sum_{e \neq E} \Pi_e \hat{\Theta}_e \right) \right. \\ &\quad \left. - \left(\sum_{e \neq E} \Pi_e \hat{\Theta}_e + \Pi_E \hat{\Theta}_E \right) \right)^2 \\ &= \sum_{E=1}^{31} \sum_{h=1}^{L_E} \frac{n_h - 1}{n_h} \sum_{i=1}^{n_h} (\hat{\Theta}_{E(hi)} - B)^2 \end{aligned} \quad (6.12)$$

Se observa que B no es igual por definición a (6.7), (6.8) ni a (6.9). Aunque parezca muy reiterativo, se recuerda que $\hat{\Theta}_E$, es el estimador del estado E , correspondiente al promedio

de las iteraciones *jackknife*, en las que se quita una localidad a la vez. Como este estimador puntual es muy parecido al global, es decir, a aquél en que no se elimina ninguna localidad (se constató en los resultados de la aplicación), se tiene que $B \approx A$, y por lo tanto, se considera (6.6) una buena aproximación de (6.10). Por otra parte, se debe notar que si en (6.6) se consideran las diferencias respecto a $\hat{\Theta}_E$ (el estimador global, que no es el promedio de iteraciones *jackknife*), entonces, $B = A$, y las dos formas de calcular la varianza son idénticas. Así pues, se tiene que el estimador que se aplicó a nivel nacional es muy similar al que hubiera requerido 854 iteraciones (pues hay 854 localidades) y en términos numéricos se pueden considerar iguales.

Sesgo relativo del estimador de razón

Otro aspecto del estimador de razón que merece discusión es su sesgo y estabilidad. Dada la definición del estimador de razón, resulta evidente que si el denominador se acerca a cero, los estimadores de varianza serán inestables. Kish (1987, pág. 133), recomienda verificar que el coeficiente de variación del denominador (que en esta aplicación es el promedio de niños menores de seis años por vivienda en el estado, o el total de niños de esta edad en el estado) sea menor a 0.2 o 0.1. Dicha recomendación se basa en el hecho de que el *sesgo relativo* del estimador de razón, definido por:

$$\frac{|Sesgo(\hat{R})|}{\sigma(\hat{R})}$$

disminuye en tanto menor sea el coeficiente de variación del denominador. Des Raj (1980, pág. 101) explica el origen de dicha relación, lo que se expone a continuación.

Sea $R = \frac{Y}{X}$, el valor poblacional y $\hat{R} = \frac{\hat{y}}{\hat{x}}$ su estimador; se considera,

$$Cov(\hat{R}, \bar{x}) = E(\hat{R}\bar{x}) - E(\hat{R})E(\bar{x}) = \bar{Y} - E(\hat{R})\bar{X}.$$

Con base en la expresión anterior para la covarianza, se llega a una expresión del sesgo, de esta forma:

$$\begin{aligned} \frac{Cov(\hat{R}, \bar{x})}{\bar{X}} &= R - E(\hat{R}) \quad , \\ Sesgo(\hat{R}) &= E(\hat{R}) - R = \frac{-Cov(\hat{R}, \bar{x})}{\bar{X}} \end{aligned} \quad (6.13)$$

Luego, ocurre que la covarianza de interés también se define como,

$$Cov(\hat{R}, \bar{x}) = -\rho(\hat{R}, \bar{x})\sigma(\hat{R})\sigma(\bar{x}),$$

donde ρ es el coeficiente de correlación lineal entre \hat{R} y \bar{x} . Por lo que se sigue que,

$$\begin{aligned} \text{Sesgo}(\hat{R}) &= -\rho(\hat{R}, \bar{x})\sigma(\hat{R})\frac{\sigma(\bar{x})}{\bar{X}} \text{ y, finalmente,} \\ \frac{|\text{Sesgo}(\hat{R})|}{\sigma(\hat{R})} &\leq \frac{\sigma(\bar{x})}{\bar{X}} = C_{\bar{x}} \end{aligned} \tag{6.14}$$

De (6.14), y recordando la definición de Error Cuadrático Medio (ECM), se deriva que:

$$ECM \leq \sigma^2(\hat{R}) (1 + C_{\bar{x}}^2) \tag{6.15}$$

Es decir, en la medida que aumenta $C_{\bar{x}}$, el coeficiente de variación del denominador del estimador de razón combinado, el error cuadrático medio del estimador crece.

Con el desarrollo anterior debe resultar evidente la importancia de la verificación de los coeficientes de variación del denominador de (6.1), a nivel estatal y regional, pues de hecho, se encontró que en algunos estados no es adecuado hacer inferencias sobre estos estimadores. Al respecto, cabe otra aclaración, ya que como las razones que se estiman en este trabajo son en sí proporciones, puede suscitarse en el lector inquietud porque en algunas categorías (a nivel estado o región estos valores son muy pequeños (menores a 0.2); situación ante la cual, de haberse obtenido los estimadores de proporción, se hubiera señalado la falta de estabilidad de su varianza. No interesa aquí entrar en detalles del estimador de proporción, pero sí puntualizar que son dos estimadores diferentes y que el trabajo que aquí se presenta, como es lógico, ha atendido las advertencias que le corresponden al estimador de razón.

6.1.2 Estimador de la mediana

La estimación de la mediana de los indicadores de estado nutricional provee información para comparar una población determinada con la población de referencia. Cuando se dan muestreos periódicos, un cambio en la mediana puede significar mejoría o deterioro de una población, según sea el caso. En un corto período de tiempo, tras la implementación de un programa de ayuda alimentaria, por ejemplo, es usual apreciar un cambio en la mediana del Peso para la Talla, aunque los otros indicadores no cambien (OMS(1983, págs. 20-32)). Si se quisiera llevar a cabo en futuras encuestas este tipo de inspección, sería indispensable un estimador de la varianza de la mediana. Sin embargo, en un muestreo complejo, la estimación de la varianza de la mediana no es tan evidente como la de la media, ya que la mediana es un cuantil, o estadística no suave.

Aunque se comentó al inicio del capítulo, se reitera que la estimación de la mediana se hizo en un plan exploratorio, pues se conoce poco sobre el comportamiento del *jackknife* al estimar su varianza en un muestreo complejo. A continuación se presenta el estimador de la mediana que se obtuvo en cada iteración *jackknife* o réplica de Repeticiones Balanceadas, basado en Kish (1972, pág. 572) y Kovar, Rao y Wu (1988, pág. 27).

De manera general, el estimador de la mediana está dado por \tilde{x} , que se define de la siguiente forma:

$$\tilde{x} = F_n^{-1}(0.5) = \min\{z : F_n(z) \geq 0.5\},$$

donde, F_n es la distribución acumulada muestral

La estimación de la mediana, como de cualquier otro cuantil, requiere del cálculo de la distribución muestral considerando el diseño, en las cercanías del valor estimado. Para evaluar la distribución muestral en un punto z , se requiere de una variable indicadora, $I_{hij}(z)$ ⁶, tal que:

$$I_{hij}(z) = \begin{cases} 1 & \text{si el valor del indicador en el elemento } hij \text{ es menor o igual a } z, \\ 0 & \text{en otro caso.} \end{cases}$$

Luego,

$$F_n(z) = \frac{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} FLOCh FVIV_{hi} Fcasa_{hij} I_{hij}(z)}{\sum_{h=1}^L \sum_{i=1}^{n_h} \sum_{j=1}^{n_{hi}} FLOCh FVIV_{hi} Fcasa_{hij}} \quad (6.16)$$

El primer valor para el cual se cumple que $F_n(z)$ es mayor o igual a 0.5, es el estimador de la mediana. Pero, en esta aplicación en particular, la precisión de los indicadores PEDZ, PETZ y TEDZ está dada en décimas, por lo que se interpoló, de manera similar a cuando se busca la mediana de datos agrupados (Ott, Mendenhall y Larson (1978), pág. 97), para evitar que todas las iteraciones resultaran en el mismo valor.

De modo que, si C es un valor tal que $F_n(C) < 0.5$, pero $F_n(C + 0.1) > 0.5$, entonces,

$$\tilde{x} = C + \frac{0.1}{F_n(C + 0.1) - F_n(C)} \{0.5 - F_n(C)\}. \quad (6.17)$$

Quiere decir que el estimador de la mediana tiene una precisión mayor a los datos de los indicadores, de acuerdo a la probabilidad acumulada hasta el valor C .

No existe una expresión definida para el estimador de varianza de la mediana en un muestreo complejo. La estimación de ésta se realizó mediante el *jackknife* y Repeticiones Balanceadas, siguiendo las expresiones correspondientes que aparecen en los Capítulos 2 y 3. Más adelante, se comentará cómo comparan las estimaciones de ambos métodos, lo cual es de interés por la poca experiencia que se tiene con este estimador en un muestreo complejo.

6.1.3 Estimador de la media

Se estimó la media de los indicadores PEDZ, PETZ y TEDZ en los estados donde habían dos localidades (conglomerados) encuestadas por estrato, para comparar las estimaciones de *jackknife* y Repeticiones Balanceadas, en cuanto a su similitud en el caso de este estimador y en el caso de la mediana. La media se estima como el cociente de la suma de los datos ponderados por los factores de expansión entre la sumatoria de los factores de expansión:

⁶La nomenclatura que aquí se usa se refiere a un diseño estratificado bietápico, lo que obviamente cambiaría para otro diseño.

$$\hat{x} = \frac{\sum_{h=1}^L \sum_{i=1}^{\eta_h} \sum_{j=1}^{m_{hi}} FLOC_h FVIV_{hi} Fcasa_{hij} \sum_{k=1}^{\eta_{hij}} y_{hijk}}{\sum_{h=1}^L \sum_{i=1}^{\eta_h} \sum_{j=1}^{m_{hi}} \sum_{k=1}^{\eta_{hij}} FLOC_h FVIV_{hi} Fcasa_{hij}} \quad (6.18)$$

donde, η_{hij} es el número de niños encuestados en la vivienda j -ésima, de la localidad i -ésima en el estrato h ; y_{hijk} es el valor de PEDZ, PETZ o TEDZ, según sea al caso, para el niño k -ésimo de la vivienda j , en la localidad i , estrato h ; el resto de la notación corresponde a lo expuesto anteriormente en este capítulo. Como ya se dijo, se consideraron los estimadores de varianza *jackknife* y Repeticiones Balanceadas.

6.2 Selección de método

En el Capítulo 4 se resumieron los hallazgos que se han reportado sobre los métodos de remuestreo. En general, al *jackknife* se le han encontrado buenos atributos en el caso de estadísticas suaves, como es la fundamentación teórica para establecer intervalos de confianza y su consistencia. También resulta más sencilla su programación que la del *bootstrap*, además de que su estimador se ajusta fácilmente a situaciones de pos-estratificación, que aunque no se da en la ENAL'96. es interesante tener la noción práctica de su operación. Por otra parte, el estimador por fórmulas, que provienen de una linealización, es el más difundido en libros de muestreo, y el que proveen algunos paquetes. Por este motivo, se decidió aplicar el *jackknife* para el caso de los estimadores de razón de la proporción de niños por grupo nutricional, además del estimador de fórmula, y así verificar la similitud entre ellos.

El método de Repeticiones Balanceadas (*Balanced Half-Sampling*) tiene ventajas similares al *jackknife* pero se adscribe al diseño de dos conglomerados por estrato. La generalización a otros diseños es posible pero requiere de la colapsación de conglomerados. En el caso de estadísticas no-suaves no tiene el problema de inconsistencia que el *jackknife*. Se sabe además que su estimador es el más parecido al *bootstrap*, así como el estimador de linealización se parece al *jackknife*. De tal suerte, se decidió aplicar el estimador de Repeticiones Balanceadas en los ocho estados donde el diseño era adecuado para ello: Aguascalientes, Guanajuato, México, Nayarit, Oaxaca, Querétaro, Sonora y Tabasco.

La estimación de la mediana y su varianza, se realizó como estudio exploratorio. Se llevó a cabo en los estados arriba mencionados para contar así con el estimador de Repeticiones Balanceadas junto con el *jackknife*.

6.3 Resultados

6.3.1 Determinación de estados en los que no son válidas las inferencias

Se consideró conveniente empezar a hablar de resultados distinguiendo aquellos estados en los que se pueden hacer inferencias sobre los estimadores de razón combinados, de manera que el lector tenga conciencia de esto al ver las tablas. De hecho, ellas incluyen

los estimadores de todos los estados porque es necesario dar a conocer estos resultados. Se recordará, que en el capítulo anterior se dijo que el plan de muestreo se realizó considerando cinco grandes regiones como dominios de estudio para los tres indicadores de estado nutricional. De ahí que aunque los estados no fueron diseñados como dominios, teniendo las precauciones que aquí se señalan, se puede determinar si es válido hacer inferencias en algunos de ellos. Como se vio en (6.14), cuando el coeficiente de variación del denominador del estimador de razón combinado es grande su sesgo relativo se incrementa y por ende su error cuadrático medio. Kish (1987), recomienda verificar que este coeficiente sea menor a 0.1 ó 0.2. En las tablas E.1 o E.2 se incluye la información sobre este coeficiente en los estados. Se aprecia que en cinco estados el valor de este coeficiente excede 0.2, los cuales son: **Baja California Sur, Chihuahua, Nuevo León Quintana Roo y Tlaxcala**. Por lo tanto, se debe de tener presente que no es adecuado hacer inferencias en estos estados. Por otra parte, se puede decir que hay 11 estados de la República en los cuales no existe ningún problema para hacer inferencias sobre la proporción de niños en las categorías de nutrición porque este coeficiente tiene un valor menor a 0.1. Estos son: **Aguascalientes, Chiapas, Guerrero, Jalisco, México, Oaxaca, Puebla, San Luis Potosí Sinaloa, Tabasco y Veracruz**. En los 15 estados restantes el coeficiente tiene un valor menor a 0.2 y mayor a 0.1, por lo que según Kish (1987), podría considerarse un criterio más liberal que el de 0.1 y considerar también las inferencias en tales estados. (Estos últimos son: **Baja California, Campeche, Coahuila, Colima, Durango, Guanajuato, Hidalgo, Michoacán, Morelos, Nayarit, Querétaro, Sonora, Tamaulipas, Yucatán y Zacatecas**)

Sin embargo, en una encuesta se dañó otros problemas y los criterios, muchas veces, no se pueden circunscribir a un aspecto. El INNSZ reportó que detectó problemas de captura o confiabilidad de datos en los estados de Baja California Sur, Baja California, Colima y algo en Quintana Roo. Los problemas más serios fueron en Baja California Sur, pues de hecho, al hacer este trabajo se identificaron más problemas en los datos. Por tal motivo, en los estados mencionados se deben guardar reservas en cuanto a sus resultados; no sólo para los estimadores de razón sino para todos los que se fuesen a obtener.

6.3.2 Resultados sobre los estimadores de razón

La varianza de los estimadores de razón obtenida por fórmula se consiguió elaborando un programa en SPSS (versión 6.0), el cual se incluye en el Apéndice F. La primer autora se percató de que son muchos los detalles a cuidar para calcular esta varianza por fórmula, y si otra persona hubiera sido quien lo programara, hubiera sido más difícil la corrección de errores. Por otro lado, la varianza *jackknife* se obtuvo desarrollando un programa en Fortran32 (ver Apéndice G). Se realizaron los cálculos de acuerdo al ajuste de factores de expansión que propone Rao (1996), y de acuerdo al método de Wolter (1985), que considera explícitamente los factores de corrección por población finita y la obtención de Pseudo-valores. En la tabla E. 2 se aprecia que ambos estimadores *jackknife* proveen resultados muy cercanos. El único estado que presenta una diferencia importante entre los dos estimadores es Hidalgo. Sin embargo, el estimador de linealización, que sí considera factores de corrección por población finita fue muy parecido al de Rao, aún en Hidalgo.

los estimadores de todos los estados porque es necesario dar a conocer estos resultados. Se recordará, que en el capítulo anterior se dijo que el plan de muestreo se realizó considerando cinco grandes regiones como dominios de estudio para los tres indicadores de estado nutricional. De ahí que aunque los estados no fueron diseñados como dominios, teniendo las precauciones que aquí se señalan, se puede determinar si es válido hacer inferencias en algunos de ellos. Como se vio en (6.14), cuando el coeficiente de variación del denominador del estimador de razón combinado es grande su sesgo relativo se incrementa y por ende su error cuadrático medio. Kish (1987), recomienda verificar que este coeficiente sea menor a 0.1 ó 0.2. En las tablas E.1 o E.2 se incluye la información sobre este coeficiente en los estados. Se aprecia que en cinco estados el valor de este coeficiente excede 0.2, los cuales son: **Baja California Sur, Chihuahua, Nuevo León Quintana Roo y Tlaxcala**. Por lo tanto, se debe de tener presente que no es adecuado hacer inferencias en estos estados. Por otra parte, se puede decir que hay 11 estados de la República en los cuales no existe ningún problema para hacer inferencias sobre la proporción de niños en las categorías de nutrición porque este coeficiente tiene un valor menor a 0.1. Estos son: **Aguascalientes, Chiapas, Guerrero, Jalisco, México, Oaxaca, Puebla, San Luis Potosí Sinaloa, Tabasco y Veracruz**. En los 15 estados restantes el coeficiente tiene un valor menor a 0.2 y mayor a 0.1, por lo que según Kish (1987), podría considerarse un criterio más liberal que el de 0.1 y considerar también las inferencias en tales estados. (Estos últimos son: **Baja California, Campeche, Coahuila, Colima, Durango, Guanajuato, Hidalgo, Michoacán, Morelos, Nayarit, Querétaro, Sonora, Tamaulipas, Yucatán y Zacatecas**).

Sin embargo, en una encuesta se dan otros problemas y los criterios, muchas veces, no se pueden circunscribir a un aspecto. El INNSZ reportó que detectó problemas de captura o confiabilidad de datos en los estados de Baja California Sur, Baja California, Colima y algo en Quintana Roo. Los problemas más serios fueron en Baja California Sur, pues de hecho, al hacer este trabajo se identificaron más problemas en los datos. Por tal motivo, en los estados mencionados se deben guardar reservas en cuanto a sus resultados; no sólo para los estimadores de razón sino para todos los que se fuesen a obtener.

6.3.2 Resultados sobre los estimadores de razón

La varianza de los estimadores de razón obtenida por fórmula se consiguió elaborando un programa en SPSS (versión 6.0), el cual se incluye en el Apéndice F. La primer autora se percató de que son muchos los detalles a cuidar para calcular esta varianza por fórmula, y si otra persona hubiera sido quien lo programara, hubiera sido más difícil la corrección de errores. Por otro lado, la varianza *jackknife* se obtuvo desarrollando un programa en Fortran32 (ver Apéndice G). Se realizaron los cálculos de acuerdo al ajuste de factores de expansión que propone Rao (1996), y de acuerdo al método de Wolter (1985), que considera explícitamente los factores de corrección por población finita y la obtención de Pseudo-valores. En la tabla E. 2 se aprecia que ambos estimadores *jackknife* proveen resultados muy cercanos. El único estado que presenta una diferencia importante entre los dos estimadores es Hidalgo. Sin embargo, el estimador de linealización, que sí considera factores de corrección por población finita fue muy parecido al de Rao, aún en Hidalgo.

La tabla E. 1 muestra los estimadores *jackknife*, los obtenidos por linealización (o fórmula) y los estimadores de razón que consideran un muestreo aleatorio simple; de donde se deriva el efecto de diseño, o DEFF, que también se exhibe. En todos los estados y estimadores es evidente la similitud de la varianza *jackknife* y de la varianza obtenida por fórmula. Se aprecia que el DEFF varía entre estados y entre distintas variables para un mismo estado. Al respecto se ve que en los estados cuyo coeficiente de variación del denominador (del estimador de razón combinado) es grande, tiende haber algún DEFF menor a 0.5. Otros estados que presentan DEFF muy pequeños son Baja California, Colima, Durango y Sonora; siendo que los coeficientes de variación de Colima y Sonora son cercanos a 0.20 pero los de Baja California y Durango son de 0.1447 y 0.1239. (se debe recordar, que como se vio en (6.14), el sesgo relativo del estimador de razón combinado disminuye en la medida en que el coeficiente de variación del denominador es menor).

En los estados donde se muestrearon dos localidades en cada estrato, se obtuvo también la varianza de acuerdo a Repeticiones Balanceadas (*Balanced Half-Sampling*), ajustando un balance ortogonal completo y aplicando la adecuación de Fay con factor de 0.5. En el Apéndice G se muestra el programa en fortran32 que se desarrolló para este propósito. Cabe mencionar que se consideró también la varianza dada por el complemento de la matriz Hadamard utilizada (se invirtieron los 1's por -1, y viceversa), y al promedio de ambas varianzas. Las matrices Hadamard en las que se basó la estimación fueron las dadas por Wolter (1985), teniendo cuidado de no usar las columnas de unos, para poder lograr el balance ortogonal.

Más adelante, en este capítulo, se encuentra la tabla 6.1, en la que se pueden apreciar las estimaciones de razón *jackknife* y de Repeticiones Balanceadas (R.B.). En la tabla 6.4 se aprecia el cociente de ambos estimadores de varianza ($\text{Var}(\textit{jackknife}) / \text{Var}(\text{R.B.})$). En los estados donde el número de estratos es mayor, y por tanto así lo es el número de iteraciones o réplicas, ambos estimadores son muy cercanos. No ocurre esto en Nayarit, Querétaro y Sonora, estados con menos de 10 estratos. Si el teorema de Krewski y Rao que se expuso en 1.6 tuviera efecto en esta aplicación, se encontraría que este resultado es coherente, pues uno de los supuestos en que se basa la demostración de la consistencia de estos estimadores de varianza para estadísticas suaves, es que se tienen un número grande de estratos y pocas unidades primarias de muestreo en ellos. Sin embargo, en esta aplicación se tiene un m.a.s. sin reemplazo de unidades primarias, lo cual no cumple con otro de los requerimientos de este teorema (m.a.s. con reemplazo de las unidades en los estratos). Pero si las fracciones de muestreo, $\frac{n_h}{N_h}$, son pequeñas, digamos menores a 0.1, el muestreo con reemplazo tiende a ser semejante al de sin reemplazo.

Otro aspecto que se observa es que en Aguascalientes, a pesar de tener cinco estratos, los estimadores no resultaron tan diferentes (el *jackknife* es aproximadamente 1.5 veces el de R.B.), lo cual se debe a que su coeficiente de variación es menor a 0.1; mientras en los otros estados mencionados con pocos estratos, la varianza *jackknife* es en muchos casos, más de 10 veces la de R.B., siendo que sus coeficientes de variación son cercanos a 0.2. Es decir, pareciera ser que aún cuando el número de estratos fuera chico, si el coeficiente de variación del denominador del estimador de razón es pequeño, las varianzas obtenidas por ambas técnicas tenderán a parecerse. En relación a esto, se encontró un

comentario de J.N.K. Rao en un artículo de Kish (1977), diciendo que cuando más grande es dicho coeficiente de variación, tiende a haber diferencias significativas entre Repeticiones Balanceadas y linealización.

Por otra parte, en el Estado de México y Oaxaca los estimadores son prácticamente iguales y se encuentran dos factores que pueden influir en ello: ambos tienen muchos estratos y coeficientes de variación del denominador menor a 0.1. En los demás estados se observa la tendencia a que la varianza *jackknife* sea mayor, lo que repercute en intervalos más amplios. Este resultado es contrario a lo que encontraron Kovar, Rao y Wu (1988), y lo cual se platica en la Sección 4.3., ya que en su estudio, ellos vieron que las varianzas de Repeticiones Balanceadas eran más conservadoras. No obstante, en este caso se tienen situaciones diferentes, pues en el estudio citado se tenían 32 estratos y un muestreo estratificado simple y aquí ocurre que los estados donde se observa que la varianza *jackknife* es mayor, son los que tienen pocos estratos.

Es importante traer a colación otro resultado del estudio de Kovar, Rao y Wu (1988, pág. 35-37). Para el caso del estimador de razón, (en poblaciones simuladas de 32 estratos, con $n_h = 2$), se vio que a medida que el coeficiente de variación del denominador crece (> 0.1), R.B. y *bootstrap* tienden a sobreestimar la varianza. Llama la atención que en la aplicación que aquí se ha realizado, cuando el número de estratos es chico, parece que ocurre lo inverso, pues el estimador *jackknife* es mayor al de R.B. Se deriva de esta observación que en estudios futuros se pudieran considerar distintas combinaciones de coeficientes de variación y número de estratos para comparar los estimadores de métodos de remuestreo, considerando su sesgo y estabilidad.

En la tabla 6.13 se encuentran los resultados de estos estimadores por regiones. Como se mencionó en el Capítulo 5, estas regiones fueron consideradas como dominios de estudio, para los tres indicadores, al hacer el plan de muestreo. Como las categorías que nos interesan se basan en estos indicadores, es lógica la consecuencia de que los estimadores de razón de proporciones en tales grupos sean más precisos a nivel región. Lo primero que salta a la vista es que los coeficientes de variación son menores a 0.1 en las cinco regiones. Aunado a esto, los intervalos de confianza son más precisos que los obtenidos por estado (de manera general). Otro factor que beneficia una buena estimación a través del *jackknife* es que cada región se compone de una gran cantidad de estratos. No se pudo contar con la estimación de Repeticiones Balanceadas por región porque no se dieron las condiciones adecuadas de diseño. Tampoco se consideró de interés obtener la estimación de linealización (fórmulas) por regiones, porque ya a nivel estado se había observado una gran similitud; en realidad, a nivel región no se estableció comparación alguna, sino simplemente, se quisieron obtener las estimaciones puntuales e intervalos, para propósitos inferenciales en los dominios de estudio que se establecieron en el diseño.

6.3.3 Resultados del estimador de la mediana

Las estimaciones de la mediana y su varianza se obtuvieron a través de programas en fortran32, los que se pueden apreciar en el Apéndice G. Sólo se incluyen resultados de los estados en que habían dos localidades encuestadas por estrato, para poder comparar la

varianza *jackknife* con la de R.B. En la tabla 6.2 se exhiben estos resultados. Conviene señalar que debido a que la precisión de los datos es de décimas, y que se tuvo que caer en una interpolación para estimar la mediana, los resultados se presentan con precisión de milésimas y no con la misma de los estimadores de razón.

Como no se tiene referencia de los valores poblacionales, no es posible determinar si la varianza *jackknife* de la mediana en un muestreo complejo como el de esta aplicación, es adecuada o no. Sin embargo, como la varianza de la mediana de R.B. no tiene los problemas que la *jackknife* presenta en el marco de m.a.s. en población infinita, se pretendió comparar el desempeño de ambas varianzas con otros estimadores, como son los estimadores de razón ya mencionados y la media de los indicadores nutricionales (PEDZ, PETZ y TEDZ). La tabla 6.3 contiene los resultados de la estimación de la media a través de las dos técnicas. Si se observa nuevamente la tabla 6.4, se aprecia que las estimaciones de varianza de las medias comparan de forma similar al comportamiento de éstas en los estimadores de razón. Además se nota que la estimación *jackknife* de la varianza suele ser mayor y que en los estados con más estratos la similitud entre ambas estimaciones es muy grande.

Lo que ocurre en el caso de la mediana es que esa disyuntiva entre los estados con pocos o muchos estratos se hace más marcada. Donde el tamaño de muestra es mayor la mediana parece ser tan estable (en cuanto a que se asemejen los métodos) como los otros estimadores suaves. En la tabla 6.4 se observa que en los estados con mayor número de estratos, Oaxaca y México, el estimador de varianza *jackknife* tiende a ser un poco menor al de Repeticiones Balanceadas, lo que hace pensar que la relación encontrada por Kovar, Rao y Wu (1988) en su estudio⁷, se da básicamente, en las situaciones en que hay muchos estratos. Estos autores encontraron que el estimador *jackknife* era más preciso y el de Repeticiones Balanceadas, más conservador. Cabe señalar que el estimador *jackknife* de la mediana de PEDZ en Oaxaca es cerca de la mitad del estimador R.B., los intervalos de confianza son muy similares. Pero sobre los intervalos se habla a continuación.

6.3.4 Intervalos de confianza

En la tabla E.3 se exhiben los intervalos de confianza de dos colas, con $\alpha = 0.05$, de los estimadores de razón de la proporción de niños por grupo nutricional, para todos los estados de la República Mexicana. Los intervalos se basan en $(n - L)$ grados de libertad, es decir el número total de localidades encuestadas en el estado menos el número de estratos. Cuando el diseño es de dos conglomerados por estrato, los grados de libertad están dados por el número de estratos. Se aprecia que la tabla E.3 también incluye los intervalos que se obtendrían si se hiciera caso omiso a los grados de libertad que corresponden a un método de remuestreo y se basaran en la $N(0, 1)$, lo que en ocasiones puede repercutir en ciertas inferencias, ya que los intervalos basados en la $N(0, 1)$ son más cortos.

La tabla 6.5 de este capítulo presenta los intervalos de confianza de los estimadores de razón, de acuerdo a la técnica *jackknife* y la de R.B. Se observa que en los estados con pocos estratos, y por ende, pocos grados de libertad, se incurre en el problema de límites inferiores menores a cero, en particular en categorías con proporción muy baja. Casi invariablemente,

⁷Ver sección 4.3.

cuando se da esta situación, son los dos métodos los que incurren en ella; a excepción del intervalo para los niños *Mal para la edad y talla* en Sonora, en cuyo caso, el dado por el *jackknife* tiene límite inferior mayor a cero, lo que no ocurre con el de R.B. Es importante ver que en la medida en que el coeficiente de variación del denominador del estimador de razón es menor (menor a 0.1), los intervalos dados por ambos métodos son más parecidos. De tal suerte, los intervalos para Sonora, Querétaro y Nayarit dados por el *jackknife* son más amplios que los que se obtuvieron por R.B. y los de Aguascalientes, México, Oaxaca, y Tabasco (estados donde el coeficiente fue menor a 0.1) son prácticamente iguales. El coeficiente de variación en Guanajuato fue de 0.1064 y los intervalos de ambas técnicas también son bastante parecidos. Llama la atención el caso de Aguascalientes, pues aunque los grados de libertad son pocos, la coincidencia de los intervalos obtenidos por los dos métodos es también muy buena. Esto da a pensar que cuando el coeficiente de variación (del denominador del estimador de razón) es chico, la congruencia entre *jackknife* y R.B. se da aún con tamaño de muestra pequeño.

Se formularon los intervalos de confianza de dos colas, con un α de 0.05, para la mediana según la estimación *jackknife* y R.B., como se ve en la tabla 6.6. Conviene recordar que la precisión de los indicadores en la base de datos es de décimas, por lo que ahora resulta más evidente que los estados con mayor número de grados de libertad: Tabasco, Guanajuato, México y Oaxaca, son los que presentan clara congruencia entre ambos métodos.

6.3.5 Interpretación y discusión general

Para facilitar la interpretación se elaboraron las tablas 6.7 - 6.11, en las que se ven los estados ordenados de acuerdo al límite superior de los intervalos de confianza *jackknife*. En la tabla 6.7 se observa que los estados con coeficiente de variación del denominador mayor a 0.2 tienen una longitud de intervalo mayor a la de los estados con coeficiente menor a 0.1. (los cuales, en general, son más cortos de 0.1). Igualmente, en la tabla 6.8 se nota que el largo de intervalo de los estados con coeficiente grande es aproximadamente de 0.12 a 0.2, pero en aquellos donde el coeficiente es menor a 0.1 el largo de intervalo es menor (o alrededor) de 0.08. De manera similar se pueden inspeccionar las tablas 6.9-6.11 y percatarse de que en efecto, cuando el coeficiente de variación del denominador del estimador de razón es grande, los intervalos son más amplios, y muchas veces llegan a ser ridículos; tal como es el caso del intervalo para Baja California Sur en la tabla 6.9, que se da de 0 a 0.30 aproximadamente (se debe recordar además que el límite inferior se acotó en cero cuando resultaba ser menor). Si se consulta la tabla 6.12 se advierte que los estados con algunos intervalos demasiado amplios son los que contaron con menor tamaño de muestra. Son evidentes los problemas en Baja California Sur (donde, de hecho, se reportaron problemas de confiabilidad o extravío de datos), al estimar la proporción de niños *Mal para la edad*, *Mal para la talla* y *Mal para edad y talla*. Como estas categorías son las que representan los niños en peor estado nutricional, es claro que no se deben arrojar conclusiones sobre este estado, debido a la magnitud de sus intervalos. Se recomienda aumentar el tamaño de muestra en este estado en futuras encuestas (de querer hacer inferencias para el estado) y supervisar de manera más cerca la captura de sus datos.

Otros estados que presentan intervalos demasiado amplios para la categoría *Mal para la talla*, son Chihuahua, Nayarit, Tamaulipas, Nuevo León, Colima y Baja California. Todos ellos tuvieron un tamaño de muestra chico (ver tabla 6.12); como se dijo antes, el INNSZ reportó también problemas de confiabilidad de datos y de captura en Colima y Baja California (aunque aquí fue menor el problema que en Baja California Sur). En Chihuahua se presenta además la situación del cero muestral en la categoría *Mal para la edad y talla*.

Ya que la categoría *mal para la talla* se refiere a niños en mala situación en la actualidad, se recomienda hacer esfuerzos en los estados arriba mencionados por incrementar su tamaño de muestra en encuestas futuras y por cuidar la calidad de la información. El estado de Quintana Roo es un caso similar a los anteriores pues presenta algunos intervalos demasiado amplios, como ocurre con el de niños *Bajitos gorditos* y los que están *Mal para la edad*.

En general, si se observa la tabla E.3, se ve que los intervalos que se basan en la $N(0, 1)$ son más estrechos y el tipo de problema que se está planteando no es tan evidente como cuando se usan los intervalos calculados a partir de una distribución t con grados de libertad que varían según el número de estratos y localidades.

La tabla 6.13 resume los resultados más importantes para propósitos inferenciales, pues en ella se exhiben los resultados obtenidos mediante el *jackknife* para las cinco regiones del país que se planificaron como dominios de estudio. Examinando los intervalos de confianza al 95%, resulta notorio el hecho de que en los estados del Noroeste y Noreste la proporción de niños normales es mucho mayor que en el Centro y Sur, mientras que la proporción de Centro Occidente es similar a la que se da en el Noreste. Se observa la tendencia, cuanto más al Sur, a una mayor proporción de niños bajitos gorditos. Esto merece una adecuada interpretación de expertos en nutrición o quizás en antropología, a fin de determinar, qué tanto se puede atribuir a diferencias étnicas o de otra índole, o si hay algún problema de mala nutrición escondido tras estas cifras (no falta de nutrición, pero sí malos hábitos alimenticios). El mismo patrón de comportamiento se observa en la categoría *Mal para la edad*, es decir, los estados del Sur y Centro tienen mayor proporción que los del Norte y Centro Occidente. Por otra parte, se vuelve a notar similitud entre el Centro Occidente y Noreste. En este grupo se clasificaron niños que reflejan que tuvieron desnutrición en etapas anteriores de su vida. Por ello, hubiera resultado interesante analizar este grupo cuando menos, por edades; además de analizar las variables captadas sobre la alimentación del niño en sus primeras etapas (lactancia y ablactación). También puede revelar la necesidad de incluir en futuras encuestas datos sobre la alimentación de mujeres embarazadas.

Los resultados de la categoría *Mal para la talla* son preocupantes pues reflejan una situación similar en todo el país. Al respecto se debe recordar que se consiera que esta categoría incluye a los niños que están sufriendo desnutrición al momento de la encuesta. Por otra parte, la proporción de niños en peor situación, agrupados bajo el nombre *Mal para la edad y talla*, vuelve a ser mayor en los estados del Centro y Sur.

Al final de la tabla E.2, se aprecian los resultados nacionales, por categorías, basados en el estimador de razón separado. Aunque es importante contar con estas cifras globales, se considera que los resultados más importantes a nivel nacional, en términos interpretativos, se desprenden de mejor manera de las estadísticas por regiones o zonas.

Sobre la utilización de los métodos de remuestreo en esta encuesta, en particular el

jackknife, cabe mencionar que fue mucho más rápido y confiable que la estimación por fórmulas. Al decir confiable se refiere a que los errores en su programación fueron de mucho menor trascendencia que los que se dieron en la obtención de varianzas por fórmulas. Esto se traduce en que es más factible entrenar a un equipo de técnicos o programadores que colaboren en el análisis de encuestas, para evaluar varianzas por *jackknife*. Existe la ventaja de trabajar de la misma forma estimadores más complejos, pos-estratificación y no respuesta, sin importar lo complejo del diseño.

La aplicación de Repeticiones Balanceadas tiene la desventaja sobre el *jackknife* de necesitar las matrices Hadamard, las cuales hay que proporcionarlas, o bien establecer la rutina que las cree. Por otra parte, no tiene el problema teórico con las estadísticas no suaves de inconsistencia de la varianza. Pero, el ejemplo que aquí se pudo desarrollar parece apoyar la idea de que en un diseño complejo, donde se eliminan conglomerados y no unidades aisladas, la varianza *jackknife* es muy similar a la R.B. cuando el número de estratos es grande.

Al respecto, un punto importante de notar es que el que la varianza de R.B. sea consistente para la mediana significa que según el tamaño de muestra crece, ésta se acerca a la verdadera varianza. Por lo tanto, en los casos donde se tienen pocos estratos y pocas localidades encuestadas, importaría más conocer el sesgo para el estimador en cuestión, que la consistencia. De aquí se desprende que en el caso del estimador de razón, cuando el coeficiente de variación del denominador es chico, el sesgo relativo del estimador es menor, lo que hace esperar que cualquiera que sea el sesgo del método utilizado (*jackknife* o R.B.) éste sea también pequeño y los estimadores calculados por diferentes técnicas se parezcan. Tal parece ser lo que ocurrió en el caso particular de Aguascalientes. (Buena aproximación entre los métodos a pesar de que el tamaño de muestra chico se reflejó en algunos intervalos amplios, como se ve en las tablas 6.9 y 6.10).

6.4 Conclusiones

Este trabajo consistió en la revisión de técnicas de remuestreo y la aplicación de algunas de ellas en la ENAL'96, para evaluar la varianza e intervalos de estimadores de razón y la mediana. Resultó gratificante el desarrollar un tema sobre técnicas poco explotadas en México y descubrir en otras personas interés en ellas (en la medida en que se desarrolló el tema). A continuación se exponen las conclusiones que se derivan de este trabajo, considerando, en mayor medida, el aspecto estadístico. En segundo término, se resumen los resultados pertinentes a la nutrición, y de los cuales se hace la advertencia, de que requieren mayor interpretación de expertos en esa área.

1. Entre todos los métodos revisados, no se puede determinar que alguno de ellos sea *el mejor*. Sin embargo, se pueden observar las recomendaciones expuestas en 4.5 para decidir en un momento dado qué método es adecuado y factible de aplicar para una situación particular.
2. Los métodos alternativos de cálculo de varianza en muestreos complejos merecen

mayor difusión en México para así mejorar la calidad de los reportes de encuestas y proveer datos más cercanos a la realidad, evitando sobre todo, el cálculo de varianzas como si se tratara de muestreo aleatorio simple.

3. Para el caso de los estimadores de razón, en un muestreo complejo, el *jackknife* representa una forma sencilla y eficaz de obtener estimadores de varianza.
4. Se corrobora la gran similitud del estimador de varianza *jackknife* y el de linealización por series de Taylor. En la aplicación aquí realizada, el estimador por fórmula (derivado de la linealización) incluyó factores de corrección por población finita y el *jackknife* basado en Rao (ver varias referencias). Sin embargo, aún así las dos estimaciones fueron prácticamente iguales en todos los estados, incluyendo Hidalgo, donde el estimador *jackknife* basado en Wolter (1985), fue el único discordante.
5. Se observó que para el caso del estimador de razón, cuando el coeficiente del denominador es menor a 0.1, las estimaciones de varianza *jackknife* y de Repeticiones Balanceadas (*Balanced Half-Sampling*) son más cercanas entre sí.
6. Respecto al punto anterior, se pudo ver que cuando el número de estratos es grande y el coeficiente de variación del denominador chico, los dos métodos señalados en 5 coinciden en sus estimaciones. (No hubo un estado donde se tuvieran muchos estratos y coeficiente grande, pero esto se debe a que en la medida de que el tamaño de muestra crece, el coeficiente de variación decrece). Por otra parte, en Aguascalientes se dio la situación de pocos estratos (5) y coeficiente menor a 0.1 y se observó una buena congruencia entre los dos métodos. En los estados con pocos estratos y coeficiente de variación grande (mayor a 0.1), se observó que la estimación de varianza por *jackknife* era mucho mayor a la de R.B. Sin embargo, no se puede determinar cuál de los dos está más cercano al verdadero valor, pues es desconocido.
7. Los intervalos de confianza de estimadores de razón en los estados con coeficiente de variación del denominador grande son muy amplios y confirman la inadecuación de hacer inferencias en tales estados.
8. A nivel región se obtuvieron coeficientes de variación del denominador muy pequeños y por ende, intervalos precisos, lo que permite las inferencias sobre aspectos nutricionales a este nivel. Este resultado era de esperar pues la muestra se diseñó considerando las regiones como dominios de estudio para los indicadores (PEDZ, PETZ y TEDZ) y las categorías para las que se calculan las razones se basan en ellos.
9. El hecho de que en este ejemplo las varianzas e intervalos de la mediana obtenidos por Repeticiones Balanceadas (R.B.) y *jackknife* en los estados donde había alrededor de 30 estratos, fueran tan parecidos sustenta la necesidad de indagar en futuras investigaciones la posibilidad de que bajo ciertas condiciones, en un muestreo complejo, el estimador de la varianza *jackknife* (de la mediana) no sea inconsistente. (Esto se desprende de que el estimador de R.B. no es inconsistente y que ambas estimaciones resultaron similares).

10. Se encontró que la combinación de niveles de los tres indicadores de estado nutricional da lugar a la formación de categorías que permiten describir mejor la situación de los niños en determinada zona. En particular, en esta aplicación, se lograron identificar cinco grupos con adecuada representatividad. En un estudio futuro, se aconseja formar primero nueve grupos, como se menciona al final del capítulo 5. Es probable que los expertos en nutrición se interesen en explorar los mismos grupos, fijando límites distintos a los aquí utilizados. Pero si desean establecer comparaciones en el tiempo, se deben construir los grupos de la misma manera. También sería de interés, en un análisis estadístico posterior, hacer un análisis de conglomerados por niño, basado en los tres indicadores.
11. Se detecta que la proporción de niños menores de seis años, con estado nutricional *Normal o casi Normal* es mayor en los estados del Noroeste y Noreste, con intervalos de confianza al 95% de (71.12-74.77) y (59.32-64.44), en comparación con los estados del Centro y Sur, cuyos intervalos resultaron de (45.89 - 50.03) y (39.88 - 43.54) respectivamente. Pero la proporción en los estados de Centro Occidente es similar a la del Noreste, con un intervalo dado por (57.39-62.75).

En los estados del Centro y Sur, la proporción de niños *Bajitos Gorditos* es mayor que en las otras regiones, interpretación de lo cual debe ser del cuidado de un experto (en los estados del Centro y Sur las estimaciones y sus intervalos son menores pero cercanos a 30%, y en las otras zonas, es menor del 20%).

El grupo de los niños *Mal para la edad*, exhibe una proporción que crece cuanto más al Sur esté la región, ocurriendo la menor proporción en el Noroeste (4.17-6.37), y la mayor en el Sur (18.13-21.53); se vuelve a observar similitud en las zonas Noreste y Centro Occidente, con intervalos de confianza al 95% de (8.60-12.53) y (7.89-11.01), respectivamente.

La clasificación *Mal para la talla* exhibió una situación similar en todo el país, ya que en el Noroeste el intervalo obtenido fue de (5.53- 10.58) y en el Sur de (5.34-6.88).

Se detectó una mayor proporción de niños *Mal para la edad y talla* en los estados del Centro y Sur, con intervalos dados por (2.40-3.41) y (2.67-3.84), respectivamente. Se encontró que dicha proporción en el Centro Occidente es un poco menor (1.32-2.40); y nuevamente, en los estados del norte, se encuentra un mejor panorama con intervalos para estas proporciones de (0.23-1.18) en el Noroeste y (0.83-1.91) en el Noreste.

12. Finalmente, aunque este trabajo fue aplicativo y de revisión, surgen tres propuestas para posibles investigaciones o trabajos en un futuro, que se plantean de forma muy general:
- (a) Considerar el inciso 7 de la sección 4.4, que trata de otra posible conceptualización del *bootstrap* en un muestreo complejo y comparar sus estimaciones con las de otras versiones del *bootstrap*, lo cual debe ser bajo un estudio de simulación. Al respecto se señala que Sitter (1992), compara tres versiones del *bootstrap*.

- (b) Considerar una muestra tan grande como la que representa la ENAL'96 (u otra cualquiera) como si fuera una población; Luego, establecer diferentes diseños con distintos números de estratos y unidades primarias de muestreo. (Tomar la gran muestra como población serviría para poder calcular valores poblacionales). Comparar el sesgo y estabilidad de los estimadores de la varianza de la mediana, obtenidos por varios métodos de remuestreo, incluyendo el *jackknife*.
- (c) Analizar una muestra proporcionada por INEGI, para la que haya interés en la implementación de métodos alternativos para el cálculo de varianzas de estimaciones tanto a nivel nacional como a nivel de subclases. (Se expone aquí porque ya el INEGI ha manifestado esta necesidad).

TABLA 6.1 : ESTIMADORES DE RAZON Y SUS VARIANZAS, OBTENIDOS POR REPETICIONES BALANCEADAS, JACKKNIFE Y FORMULA, EN LOS ESTADOS CON DOS LOCALIDADES MUESTREADAS POR ESTRATO

ESTADO	GRUPO NUTRICIONAL	REPETICIONES BALANCEADAS* (Half-Sampling)		JACKKNIFE		FORMULA o Linearización					
		ESTIMADOR DE RAZON**	VARIANZA	ESTIMADOR DE RAZON**	VARIANZA	ESTIMADOR DE RAZON	VARIANZA				
1 AGUASCALIENTES	L = 5 NORMALES y casi N, n=385 BAJITOS GORDITOS CV = 0.068466 MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.671295	0.000436	0.000436	0.000436	0.671309	0.000632	0.671347	0.000626		
		0.150036	0.000296	0.000296	0.000296	0.150033	0.000327	0.150159	0.000297		
		0.091842	0.000357	0.000357	0.000357	0.091889	0.000564	0.091869	0.000515		
		0.071799	0.000228	0.000228	0.000228	0.071746	0.000341	0.071514	0.000239		
		0.015027	0.000044	0.000044	0.000044	0.015022	0.000063	0.015111	0.000057		
		0.585936	0.000421	0.000415	0.000418	0.586051	0.000539	0.586011	0.000538		
		0.180433	0.000419	0.000413	0.000416	0.180556	0.000516	0.180556	0.000486		
		0.105298	0.000192	0.000196	0.000194	0.105145	0.000239	0.105159	0.000229		
		0.093071	0.000057	0.000061	0.000059	0.092979	0.000092	0.092989	0.000083		
		0.035263	0.000021	0.000022	0.000021	0.035269	0.000059	0.035284	0.000056		
15 MEXICO	L=27 NORMALES y casi N, n=1209 BAJITOS GORDITOS CV= 0.097556 MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.533425	0.000554	0.000564	0.000559	0.533731	0.000553	0.533757	0.000664		
		0.268736	0.000412	0.000415	0.000413	0.268386	0.000414	0.268416	0.00039		
		0.136376	0.000122	0.000127	0.000122	0.13611	0.000117	0.136056	0.000109		
		0.054562	0.000143	0.000147	0.000145	0.054748	0.000148	0.054723	0.000089		
		0.006902	0.000007	0.000007	0.000007	0.007024	0.000007	0.007048	0.000007		
		0.638343	0.000514	0.000514	0.000514	0.638913	0.00039	0.637705	0.0003404		
		0.180257	0.000138	0.000138	0.000138	0.17982	0.001537	0.180209	0.001439		
		0.074245	0.000256	0.000256	0.000256	0.074105	0.000732	0.074775	0.000687		
		0.095639	0.000040	0.000040	0.000040	0.095601	0.000926	0.095646	0.000572		
		0.011516	0.000064	0.000064	0.000064	0.011561	0.000066	0.011665	0.000061		
18 NAYARIT	L= 5 NORMALES y casi N, n=207 BAJITOS GORDITOS CV= 0.148956 MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.404475	0.000392	0.000429	0.000410	0.404345	0.000408	0.404333	0.000342		
		0.301328	0.000399	0.000407	0.000403	0.301366	0.000492	0.30132	0.000384		
		0.211758	0.000449	0.000451	0.000450	0.211899	0.00045	0.211909	0.000426		
		0.052913	0.000080	0.000085	0.000083	0.052892	0.000082	0.052922	0.000066		
		0.029526	0.000025	0.000025	0.000025	0.029498	0.000025	0.029514	0.000024		
		20 OAXACA	L= 30 NORMALES y casi N, 1728 BAJITOS GORDITOS CV= 0.090278 MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.671295	0.000436	0.000436	0.000436	0.671309	0.000632	0.671347	0.000626
				0.150036	0.000296	0.000296	0.000296	0.150033	0.000327	0.150159	0.000297
				0.091842	0.000357	0.000357	0.000357	0.091889	0.000564	0.091869	0.000515
				0.071799	0.000228	0.000228	0.000228	0.071746	0.000341	0.071514	0.000239
				0.015027	0.000044	0.000044	0.000044	0.015022	0.000063	0.015111	0.000057
0.585936	0.000421			0.000415	0.000418	0.586051	0.000539	0.586011	0.000538		
0.180433	0.000419			0.000413	0.000416	0.180556	0.000516	0.180556	0.000486		
0.105298	0.000192			0.000196	0.000194	0.105145	0.000239	0.105159	0.000229		
0.093071	0.000057			0.000061	0.000059	0.092979	0.000092	0.092989	0.000083		
0.035263	0.000021			0.000022	0.000021	0.035269	0.000059	0.035284	0.000056		

* Repeticiones balanceadas con balace ortogonal completo y la adecuación de Fay, factor =0.5.
 Var. Complemento es la varianza de las réplicas dadas por el complemento de la matriz Hadamard.
 ** Se exhibe la media de las iteraciones.

TABLA 6.1 : ESTIMADORES DE RAZON Y SUS VARIANZAS.
 OBTENIDOS POR REPETICIONES BALANCEADAS, JACKKNIFE Y FORMULA,
 EN LOS ESTADOS CON DOS LOCALIDADES MUESTREADAS POR ESTRATO

ESTADO	GRUPO NUTRICIONAL	REPETICIONES BALANCEADAS* (Half-Sampling)			JACKKNIFE		FORMULA o Linearización		
		ESTIMADOR DE RAZON**	VARIANZA	COMPLEMENTO VAR.	MEDIA DE VARIANZAS	ESTIMADOR DE RAZON**	VARIANZA	ESTIMADOR DE RAZON	VARIANZA
22 QUERETARO L=9 NORMALES y casi N, n=651 BAJITOS GORDITOS CV= 0.191388	MAL PARA LA TALLA MAL PARA LA TALLA	0.553715	0.000045	0.000044	0.000045	0.554111	0.000722	0.553556	0.000477
		0.228116	0.000061	0.000062	0.000061	0.227926	0.000308	0.228162	0.000028
		0.155317	0.000225	0.000217	0.000221	0.154936	0.000689	0.155423	0.000627
		0.048646	0.000080	0.000080	0.000080	0.04878	0.000137	0.048619	0.000121
26 SONORA L=6 NORMALES y casi N, n=243 BAJITOS GORDITOS CV= 0.187546	MAL PARA LA TALLA MAL PARA LA TALLA	0.014206	0.000003	0.000003	0.000003	0.014248	0.000031	0.01424	0.000028
		0.768257	0.000092	0.000101	0.000097	0.768715	0.000965	0.768588	0.000097
		0.133309	0.000193	0.000195	0.000194	0.133197	0.000376	0.13336	0.000359
		0.025261	0.000119	0.000119	0.000119	0.024944	0.000309	0.024864	0.000281
27 TABASCO L= 17 NORMALES y casi N, n=1367 BAJITOS GORDITOS CV= 0.61831	MAL PARA LA TALLA MAL PARA LA TALLA	0.068768	0.000032	0.000025	0.000029	0.068712	0.000202	0.068776	0.000382
		0.004405	0.000001	0.000001	0.000001	0.004431	0.000012	0.004412	0.000011
		0.58079	0.000254	0.000250	0.000252	0.58076	0.000287	0.581191	0.000454
		0.213225	0.000300	0.000295	0.000297	0.21323	0.000307	0.213235	0.000292
	MAL PARA LA TALLA MAL PARA LA TALLA	0.117903	0.000018	0.000018	0.000018	0.117906	0.000047	0.11777	0.000043
		0.07248	0.000044	0.000044	0.000044	0.072502	0.00007	0.072259	0.000067
		0.015603	0.000037	0.000037	0.000037	0.015603	0.000042	0.015544	0.000039

* Repeticiones balanceadas con balace ortogonal completo y la adecuación de Fay, factor =0.5.
 Var. Complemento es la varianza de las réplicas dadas por el complemento de la matriz Hadamard.
 ** Se exhibe la media de las iteraciones.

TABLA 6.2 : ESTIMADORES DE LA MEDIANA DE LOS TRES INDICADORES NUTRICIONALES, Y SUS VARIANZAS, OBTENIDOS MEDIANTE JACKKNIFE Y REPETICIONES BALANCEADAS EN LOS ESTADOS CON DOS LOCALIDADES ENCUESTADAS POR ESTRATO

ESTADO L = no. estratos NO. n= niños enc.	INDICADOR NUTRICIONAL	Jackknife		REPETICIONES BALANCEADAS* (Balanced Half-Sampling)			
		Mediana**	Varianza	Mediana**	Varianza	V(Complemento)	Media de Vars.
1 AGUASCALIENTES L=5 n=385	Peso para la Edad	-0.502	0.008601	-0.501	0.002145	0.002145	0.002145
	Peso para la Talla	-0.021	0.001757	-0.021	0.000291	0.000291	0.000291
	Talla para la Edad	-0.664	0.015055	-0.666	0.005046	0.005046	0.005046
11 GUANAJUATO L= 19 n=845	Peso para la Edad	-0.909	0.010048	-0.909	0.007264	0.008092	0.007678
	Peso para la Talla	-0.194	0.016098	-0.187	0.013345	0.013911	0.013628
	Talla para la Edad	-1.131	0.005424	-1.132	0.002545	0.002579	0.002562
15 MEXICO L=27 n=1209	Peso para la Edad	-0.846	0.005124	-0.849	0.004773	0.005237	0.005005
	Peso para la Talla	0.123	0.003039	0.122	0.002783	0.00278	0.0027815
	Talla para la Edad	-1.366	0.011277	-1.368	0.009914	0.01149	0.010702
18 NAYARIT L=5 n=207	Peso para la Edad	-0.553	0.019390	-0.549	0.00043	0.00043	0.00043
	Peso para la Talla	0.008	0.026707	-0.006	0.010468	0.010468	0.010468
	Talla para la Edad	-0.619	0.024550	-0.615	0.001293	0.001293	0.001293
20 OAXACA L=30 n=1728	Peso para la Edad	-1.212	0.001356	-1.207	0.003113	0.001903	0.002508
	Peso para la Talla	0.047	0.004945	0.047	0.00622	0.005919	0.0060695
	Talla para la Edad	-1.799	0.004957	-1.800	0.004846	0.005159	0.0050025
22 QUERETARO L=9 n=651	Peso para la Edad	-0.857	0.006873	-0.857	0.001797	0.001492	0.0016445
	Peso para la Talla	0.074	0.006114	0.073	0.000958	0.00089	0.000924
	Talla para la Edad	-1.287	0.004380	-1.285	0.000349	0.000351	0.00035
26 SONORA L=6 n=243	Peso para la Edad	0.044	0.037332	0.036	0.021089	0.018609	0.019849
	Peso para la Talla	0.320	0.048977	0.304	0.043526	0.043649	0.0435875
	Talla para la Edad	-0.043	0.097062	-0.04	0.0546	0.048985	0.0517925
27 TABASCO L= 17 n=1367	Peso para la Edad	-0.934	0.005094	-0.934	0.002532	0.002469	0.0025005
	Peso para la Talla	-0.110	0.006276	-0.109	0.004084	0.004528	0.004306
	Talla para la Edad	-1.230	0.001133	-1.229	0.001013	0.00094	0.0009765

* Repeticiones Balanceadas con balance ortogonal completo y la adecuación de Fay (factor =0.5).

V(complemento) = varianza de réplicas dadas por el complemento de la matriz Hadamard

**Se exhibe la media de las iteraciones o réplicas.

TABLA 6.3: ESTIMADORES DE LA MEDIA DE LOS TRES INDICADORES NUTRICIONALES, Y SUS VARIANZAS, OBTENIDOS MEDIANTE JACKKNIFE Y REPETICIONES BALANCEADAS EN LOS ESTADOS CON DOS LOCALIDADES ENCUESTADAS POR ESTRATO

ESTADO NO. (L = no. estratos)	INDICADOR NUTRICIONAL	Jackknife		REPETICIONES BALANCEADAS* (Balanced Half-Sampling)			
		Media**	Varianza	Media**	Varianza	V(Complemento)	Media de Vars.
1 AGUASCALIENTES L=5	Peso para la Edad	-0.422	0.006004	-0.433	0.003401	0.003401	0.003401
	Peso para la Talla	0.027	0.004985	0.035	0.004175	0.004175	0.004175
	Talla para la Edad	-0.653	0.026368	-0.664	0.013155	0.013155	0.013155
11 GUANAJUATO L= 19	Peso para la Edad	-0.795	0.008635	-0.800	0.006797	0.00694	0.006869
	Peso para la Talla	-0.112	0.015020	-0.113	0.014398	0.01433	0.014364
	Talla para la Edad	-1.058	0.006278	-1.062	0.004260	0.004138	0.004199
15 MEXICO L=27	Peso para la Edad	-0.732	0.002518	-0.739	0.002523	0.002521	0.002522
	Peso para la Talla	0.173	0.002614	0.175	0.002738	0.002731	0.002735
	Talla para la Edad	-1.283	0.009194	-1.292	0.009040	0.009195	0.009118
18 NAYARIT L=5	Peso para la Edad	-0.469	0.013076	-0.475	0.002095	0.002095	0.002095
	Peso para la Talla	0.005	0.021640	0.002	0.005935	0.005935	0.005935
	Talla para la Edad	-0.635	0.077107	-0.640	0.007608	0.007608	0.007608
20 OAXACA L=30	Peso para la Edad	-1.059	0.004829	-1.064	0.004857	0.004815	0.004836
	Peso para la Talla	0.099	0.007921	0.102	0.007760	0.008227	0.007994
	Talla para la Edad	-1.758	0.006317	-1.763	0.006058	0.006596	0.006327
22 QUERETARO L=9	Peso para la Edad	-0.793	0.004830	-0.801	0.001004	0.000966	0.000985
	Peso para la Talla	0.095	0.005243	0.103	0.000617	0.000607	0.000612
	Talla para la Edad	-1.280	0.008829	-1.290	0.001517	0.001508	0.001513
26 SONORA L=6	Peso para la Edad	0.218	0.038753	0.220	0.029445	0.029400	0.029423
	Peso para la Talla	0.381	0.068982	0.383	0.058103	0.057863	0.057983
	Talla para la Edad	0.059	0.029769	0.059	0.024003	0.023266	0.023635
27 TABASCO L= 17	Peso para la Edad	-0.793	0.003197	-0.801	0.001964	0.002021	0.001993
	Peso para la Talla	-0.051	0.002719	-0.059	0.001656	0.001593	0.001625
	Talla para la Edad	-1.118	0.003921	-1.126	0.003216	0.003283	0.003250

* Repeticiones Balanceadas con balance ortogonal completo y la adecuación de Fay (factor =0.5).

V(complemento) = varianza de réplicas dadas por el complemento de la matriz Hadamard.

**Se exhibe la media de las iteraciones o réplicas.

TABLA 6.4 : COMPARACION DE LOS ESTIMADORES DE VARIANZA JACKKNIFE (1) Y REPETICIONES BALANCEADAS (2) PARA LA MEDIA, MEDIANA Y LOS ESTIMADORES DE RAZON DE LA PROPORCION POR GRUPO NUTRICIONAL, MEDIANTE EL COCIENTE [(1) / (2)]

ESTADO		ESTIMADOR DE VARIANZA JACKKNIFE / ESTIMADOR DE VARIANZA REPETICIONES BALANCEADAS									
		ESTIMADORES DE RAZON		PESO PARA LA EDAD		PESO PARA LA TALLA		TALLA PARA LA EDAD		TALLA PARA LA TALLA	
L= no. de estratos	n= niños encuestados	POR GRUPO NUTRICIONAL	(1)/(2)	MEDIA CV*	MEDIANA	MEDIA	MEDIANA	MEDIA	MEDIANA	MEDIA	MEDIANA
1	AGUASCALIENTES L= 5 n=385	NORMALES y casi N,	1.484	1.7654	3.0774	1.1940	6.0378	2.0044	2.9836		
		BAJITOS GORDITOS	1.1047	0.068466							
		MAL PARA LA EDAD	1.5783								
		MAL PARA LA TALLA	1.4984								
11	GUANAJUATO L= 19 n=845	MAL-EDAD Y TALLA	1.4255								
		NORMALES y casi N,	1.2881	1.2572	1.3087	1.0457	1.1812	1.4951	2.1171		
		BAJITOS GORDITOS	1.2406	0.106484							
		MAL PARA LA EDAD	1.2311								
15	MEXICO L=27 n=1209	MAL PARA LA TALLA	1.5467								
		MAL-EDAD Y TALLA	2.7593								
		NORMALES y casi N,	0.9889	0.9984	1.0238	0.9559	1.0926	1.0084	1.0537		
		BAJITOS GORDITOS	1.0016	0.097556							
18	NAYARIT L= 5 n=207	MAL PARA LA EDAD	0.9581								
		MAL PARA LA TALLA	1.0202								
		MAL-EDAD Y TALLA	0.9942								
		NORMALES y casi N,	7.5882	6.2415	45.0930	3.6462	2.5513	10.1350	18.9869		
20	OAXACA L= 30 1728	BAJITOS GORDITOS	11.1167	0.148956							
		MAL PARA LA EDAD	2.8538								
		MAL PARA LA TALLA	22.8983								
		MAL-EDAD Y TALLA	1.0360	0.9986	0.5407	0.9909	0.8147	0.9984	0.9909		
		NORMALES y casi N,	0.9946								
		BAJITOS GORDITOS	0.9974	0.090278							
		MAL PARA LA EDAD	0.9991								
		MAL PARA LA TALLA	0.9931								

* CV= coeficiente de variación del denominador del est. de razón combinado.

TABLA 6.4 : COMPARACION DE LOS ESTIMADORES DE VARIANZA JACKKNIFE (1) Y REPETICIONES BALANCEADAS (2) PARA LA MEDIA, MEDIANA Y LOS ESTIMADORES DE RAZON DE LA PROPORCION POR GRUPO NUTRICIONAL, MEDIANTE EL COCIENTE [(1) / (2)]

ESTADO	ESTIMADOR DE VARIANZA JACKKNIFE / ESTIMADOR DE VARIANZA REPETICIONES BALANCEADAS		PESO PARA LA EDAD		TALLA PARA LA TALLA		TALLA PARA LA EDAD	
	L= no. de estratos n= niños encuestados	ESTIMADORES DE RAZON	MEDIA	MEDIANA	MEDIA	MEDIANA	MEDIA	MEDIANA
		POR GRUPO NUTRICIONAL	(1) / (2)	CV*				
		MAL EDAD Y TALLA	0.9989					
22	QUERETARO	NORMALES y casi N, L=9 n=651	16.1150 5.0267 3.1186 1.7204 10.9736	4.9036 0.191388	4.1794	8.5670	6.6169	5.8374 12.5143
26	SONORA	NORMALES y casi N, L=6 n=243	9.9703 1.9365 2.6033 7.0529 23.7062	1.3171 0.187546	1.8808	1.1897	1.1236	1.2596 1.8741
27	TABASCO	NORMALES y casi N, L= 17 n=1367	1.1379 1.0328 2.6281 1.5903 1.1342	1.6045 0.061831	2.0372	1.6737	1.4575	1.2066 1.1603

* CV= coeficiente de variación del denominador del est. de razón combi

TABLA 6.5: INTERVALOS DE CONFIANZA AL 95% PARA LOS ESTIMADORES DE RAZON, SEGUN LAS VARIANZAS JACKKNIFE Y REPETICIONES BALANCEADAS EN LOS ESTADOS CON DOS LOCALIDADES ENCUESTADAS POR ESTRATO

ESTADO NO.	GRUPO NOMBRE	G. L. (I - L)	REPETICIONES BALANCEADAS		JACKKNIFE	
			LIM. INF.	LIM. SUP.	LIM. INF.	LIM. SUP.
1 AGUASCALIENTES cv* = 0.68466	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	5	0.617599	0.724991	0.606686	0.735932
			0.105809	0.194263	0.103549	0.196517
			0.043248	0.140436	0.030841	0.152937
			0.033020	0.110578	0.024277	0.119215
			0.000000	0.032116	0.000000	0.035425
11 GUANAJUATO cv= 0.106484	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	19	0.543122	0.628750	0.537459	0.634643
			0.137747	0.223119	0.133012	0.228100
			0.076135	0.134461	0.072788	0.137502
			0.076929	0.109213	0.072903	0.113055
			0.025585	0.044941	0.019192	0.051346
15 MEXICO cv= 0.097556	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	27	0.484905	0.581945	0.485480	0.581982
			0.227020	0.310452	0.226637	0.310135
			0.113702	0.159050	0.113916	0.158304
			0.029849	0.079275	0.029786	0.079710
			0.001457	0.012347	0.001595	0.012453
18 NAYARIT cv= 0.148966	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	5	0.580067	0.696619	0.478380	0.799446
			0.150031	0.210483	0.079042	0.280598
			0.033076	0.115414	0.004557	0.143653
			0.079292	0.111986	0.017378	0.173824
			0.000000	0.082994	0.000000	0.082444
20 OAXACA cv= 0.090278	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	30	0.363110	0.445840	0.363093	0.445597
			0.260327	0.342329	0.260419	0.342313
			0.168416	0.255100	0.168576	0.255222
			0.034355	0.071471	0.034398	0.071386
			0.019309	0.039743	0.019287	0.039709
22 QUERETARO cv= 0.191388	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	9	0.538573	0.568857	0.493327	0.614895
			0.210409	0.245823	0.188225	0.267627
			0.121693	0.188941	0.095557	0.214315
			0.028459	0.068833	0.022302	0.075258
			0.010404	0.018008	0.001653	0.026843
26 SONORA cv= 0.187546	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	6	0.744184	0.792330	0.692703	0.844727
			0.099213	0.167405	0.085750	0.180644
			0.000000	0.061919	0.000000	0.067957
			0.055673	0.081863	0.033935	0.103489
			0.002664	0.006146	0.000000	0.012907
27 TABASCO cv= 0.61831	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	17	0.547284	0.614296	0.545017	0.616503
			0.176850	0.249600	0.176263	0.250197
			0.108981	0.126825	0.103442	0.132370
			0.058483	0.086477	0.054850	0.090154
			0.002764	0.028442	0.001930	0.029276

* | = localidades encuestadas en el estado. L = estratos. cv = coef. de var. del denominador del est. combinado. Se encuentran sombreados los intervalos cuyo límite inferior resulto ser menor a cero, por lo que aparece "0" como dicho límite.

LA 6.6 : INTERVALOS DE CONFIANZA AL 95% PARA LA MEDIANA DE LOS TRES INDICADORES DE ESTADO NUTRICIONAL, DE ACUERDO A LAS ESTIMACIONES JACKKNIFE Y REPETICIONES BALANCEADAS EN LOS ESTADOS DONDE HABIAN DOS LOCALIDADES ENCUESTADAS POR ESTRATO

NO.	ESTADO	G.L. (n-L)	INDICADOR NUTRICIONAL	INTERVALOS DE CONFIANZA 95% PARA LA MEDIANA			
				REPETICIONES BALANCEADAS		JACKKNIFE	
				Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
1	AGUASCALIENTES	5	<i>Peso para la Edad</i>	-0.711	-0.293	-0.620	-0.382
			<i>Peso para la Talla</i>	-0.129	0.087	-0.065	0.023
			<i>Talla para la Edad</i>	-0.979	-0.349	-0.849	-0.483
11	GUANAJUATO	19	<i>Peso para la Edad</i>	-1.119	-0.699	-1.092	-0.726
			<i>Peso para la Talla</i>	-0.460	0.072	-0.431	0.057
			<i>Talla para la Edad</i>	-1.285	-0.977	-1.238	-1.026
15	MEXICO	27	<i>Peso para la Edad</i>	-0.993	-0.699	-0.994	-0.704
			<i>Peso para la Talla</i>	0.010	0.236	0.014	0.230
			<i>Talla para la Edad</i>	-1.584	1.148	-1.580	-1.156
18	NAYARIT	5	<i>Peso para la Edad</i>	-0.911	-0.195	-0.602	-0.496
			<i>Peso para la Talla</i>	-0.412	0.428	-0.269	0.257
			<i>Talla para la Edad</i>	-1.022	-0.216	-0.707	-0.523
20	OAXACA	30	<i>Peso para la Edad</i>	-1.287	-1.137	-1.309	-1.105
			<i>Peso para la Talla</i>	-0.097	0.191	-0.112	0.206
			<i>Talla para la Edad</i>	-1.943	-1.655	-1.944	-1.656
22	QUERETARO	9	<i>Peso para la Edad</i>	-1.045	-0.669	-0.949	-0.765
			<i>Peso para la Talla</i>	-0.103	0.251	0.004	0.142
			<i>Talla para la Edad</i>	-1.437	-1.137	-1.327	-1.243
26	SONORA	6	<i>Peso para la Edad</i>	-0.429	0.517	-0.309	0.381
			<i>Peso para la Talla</i>	-0.222	0.862	-0.207	0.815
			<i>Talla para la Edad</i>	-0.805	0.719	-0.597	0.517
27	TABASCO	17	<i>Peso para la Edad</i>	-1.085	-0.783	-1.040	-0.828
			<i>Peso para la Talla</i>	-0.277	0.057	-0.247	0.029
			<i>Talla para la Edad</i>	-1.301	-1.159	-1.295	-1.163

TABLA 6.7 : ORDENAMIENTO DE LOS ESTADOS DE ACUERDO A LOS INTERVALOS DE CONFIANZA AL 95%. BASADOS EN EL JACKKNIFE, PARA LA PROPORCION DE NIÑOS NORMALES Y CASI NORMALES

ESTADOS (En orden de Ilm. Sup.)	I. C. 95% JACKKNIFE	
	Normales y casi normales	
	Lim. Inferior	Lim. Superior
31 YUCATÁN	0.266767	0.391585
7 CHIAPAS	0.330698	0.403002
21 PUEBLA	0.362904	0.429442
12 GUERRERO	0.348648	0.434964
20 OAXACA	0.363093	0.445597
4 CAMPECHE	0.368336	0.506764
13 HIDALGO	0.400708	0.509930
23 QUINTANA ROO	0.345634	0.525576
30 VERACRUZ	0.492586	0.572658
15 MEXICO	0.485480	0.581982
6 COLIMA	0.408372	0.610302
22 QUERETARO	0.493327	0.614895
27 TABASCO	0.545017	0.616503
24 SAN LUIS POTOSI	0.500934	0.627500
11 GUANAJUATO	0.537459	0.634643
19 NUEVO LEON	0.557289	0.635110
17 MORELOS	0.474763	0.639563
29 TLAXCALA	0.416739	0.644435
16 MICHOACAN	0.525397	0.644645
32 ZACATECAS	0.591801	0.683083
28 TAMAULIPAS	0.588480	0.698480
3 B. C. SUR	0.661000	0.715264
10 DURANGO	0.621204	0.727040
1 AGUASCALIENTES	0.606686	0.735932
14 JALISCO	0.684341	0.750537
25 SINALOA	0.689042	0.754564
8 CHIHUAHUA	0.678075	0.772831
5 COAHUILA	0.588989	0.783847
18 NAYARIT	0.478380	0.799446
26 SONORA	0.692703	0.844727
2 BAJA CALIFORNIA	0.741883	0.906515

Nota: No se deben hacer inferencias en los estados sombreados, pues $cv > 0.2$

TABLA 6.8 : ORDENAMIENTO DE LOS ESTADOS DE ACUERDO A LOS INTERVALOS DE CONFIANZA AL 95%, BASADOS EN EL JACKKNIFE, PARA LA PROPORCION DE NIÑOS BAJITOS GORDITOS

ESTADOS (En orden de lIm. Sup.)	I. C. 95% JACKKNIFE Bajitos gorditos	
	Lim. Inferior	Lim. Superior
14 JALISCO	0.114164	0.162412
25 SINALOA	0.094869	0.164789
2 BAJA CALIFORNIA	0.000000	0.166034
5 COAHUILA	0.095327	0.176401
26 SONORA	0.085750	0.180644
1 AGUASCALIENTES	0.103549	0.196517
3 B. C. SUR	0.072454	0.203430
10 DURANGO	0.151000	0.214354
32 ZACATECAS	0.124984	0.215090
8 CHIHUAHUA	0.026930	0.216386
11 GUANAJUATO	0.133012	0.228100
27 TABASCO	0.176263	0.250197
12 GUERRERO	0.204928	0.255856
19 NUEVO LEON	0.077484	0.262934
30 VERACRUZ	0.210663	0.264853
24 SAN LUIS POTOSI	0.172292	0.265388
22 QUERETARO	0.188225	0.267627
28 TAMAULIPAS	0.102342	0.268808
16 MICHOACAN	0.210609	0.273275
18 NAYARIT	0.079042	0.280598
17 MORELOS	0.209153	0.291217
29 TLAXCALA	0.170702	0.292472
15 MEXICO	0.226637	0.310135
21 PUEBLA	0.246954	0.312868
6 COLIMA	0.118381	0.321205
13 HIDALGO	0.237650	0.338746
31 YUCATAN	0.278393	0.340177
20 OAXACA	0.260419	0.342313
4 CAMPECHE	0.273219	0.351855
7 CHIAPAS	0.327491	0.394679
23 QUINTANA ROO	0.265878	0.423776

Nota: No se deben hacer inferencias en los estados sombreados, pues $cv > 0.2$

TABLA 6.9 : ORDENAMIENTO DE LOS ESTADOS DE ACUERDO A LOS INTERVALOS DE CONFIANZA AL 95%. BASADOS EN EL JACKKNIFE, PARA LA PROPORCION DE NIÑOS MAL PARA LA EDAD

ESTADOS (En orden de ltm. Sup.)	I. C. 95% JACKKNIFE	
	Maj para la Edad	
	Lim. Inferior	Lim. Superior
2 BAJA CALIFORNIA	0.035222	0.050928
14 JALISCO	0.027441	0.052607
25 SINALOA	0.035038	0.067192
26 SONORA	0.000000	0.067957
10 DURANGO	0.027108	0.082944
28 TAMAULIPAS	0.006411	0.086109
6 COLIMA	0.020204	0.095282
16 MICHOACAN	0.060171	0.121581
8 CHIHUAHUA	0.025327	0.122331
27 TABASCO	0.103442	0.132370
19 NUEVO LEON	0.020966	0.132990
11 GUANAJUATO	0.072788	0.137502
18 NAYARIT	0.004557	0.143653
1 AGUASCALIENTES	0.030841	0.152937
32 ZACATECAS	0.046305	0.154401
15 MEXICO	0.113916	0.158304
24 SAN LUIS POTOSI	0.104231	0.166481
30 VERACRUZ	0.122421	0.168119
5 COAHUILA	0.021548	0.181416
13 HIDALGO	0.120942	0.192302
29 TLAXCALA	0.079599	0.192461
17 MORELOS	0.081343	0.207941
22 QUERETARO	0.095557	0.214315
21 PUEBLA	0.169743	0.215367
7 CHIAPAS	0.153943	0.216261
23 QUINTANA ROO	0.071619	0.227071
4 CAMPECHE	0.155180	0.242484
20 OAXACA	0.168576	0.255222
31 YUCATAN	0.151446	0.294018
12 GUERRERO	0.214627	0.300077
3 B. C. SUR	0.000000	0.302825

Nota: No se deben hacer inferencias en los estados sombreados, pues $cv > 0.2$

TABLA 6.10 : LISTADO DE LOS ESTADOS ORDENADOS DE ACUERDO A LOS INTERVALOS DE CONFIANZA AL 95%, BASADOS EN EL JACKKNIFE, PARA LA PROPORCION DE NIÑOS MAL PARA LA TALLA

RANGC	ESTADOS (En orden de lIm. Sup.)		I. C. 95% JACKKNIFE Mal Para la Talla	
	NO.	NOMBRE	Lim. Inferior	Lim. Superior
1	4	CAMPECHE	0.020643	0.040073
2	17	MORELOS	0.018885	0.059125
3	12	GUERRERO	0.046896	0.070244
4	20	OAXACA	0.034398	0.071386
5	30	VERACRUZ	0.048964	0.075140
6	22	QUERETARO	0.022302	0.075258
7	15	MEXICO	0.029786	0.079710
8	7	CHIAPAS	0.047300	0.080396
	24	SAN LUIS POTOSI	0.035182	0.084570
10	27	TABASCO	0.054850	0.090154
11	13	HIDALGO	0.032471	0.097841
12	5	COAHUILA	0.024828	0.098458
13	16	MICHOACAN	0.035112	0.103456
14	26	SONORA	0.033935	0.103489
15	25	SINALOA	0.056293	0.107391
16	21	PUEBLA	0.058572	0.110940
17	32	ZACATECAS	0.058535	0.111359
18	14	JALISCO	0.081039	0.111441
19	11	GUANAJUATO	0.072903	0.113055
20	3	B. C. SUR	0.000000	0.113605
21	23	QUINTANA ROO	0.000000	0.117440
22	29	TLAYCALA	0.025116	0.118794
23	1	AGUASCALIENTES	0.024277	0.119215
24	31	YUCATAN	0.054160	0.122356
25	10	DURANGO	0.037328	0.127754
26	8	CHIHUAHUA	0.007056	0.153326
27	18	NAYARIT	0.017378	0.173824
28	28	TAMAULIPAS	0.026843	0.210797
29	19	NUEVO LEON	0.018423	0.289205
30	2	BAJA CALIFORNIA	0.000000	0.295545
31	6	COLIMA	0.073983	0.301847

Nota: No se deben hacer inferencias en los estados sombreados, pues $cv > 0.2$

TABLA 6.11 : LISTADO DE LOS ESTADOS ORDENADOS DE ACUERDO A LOS INTERVALOS DE CONFIANZA AL 95%, BASADOS EN EL JACKKNIFE, PARA LA PROPORCION DE NIÑOS MAL PARA LA EDAD Y TALLA

RANGO	ESTADOS (En orden de lim. Sup.)		I. C. 95% JACKKNIFE Mal para Edad y Talla	
	NO	NOMBRE	Lim. Inferior	Lim. Superior
1	8	CHIHUAHUA		
2	2	BAJA CALIFORNIA	0.000000	0.005944
3	10	DURANGO	0.001312	0.009954
4	15	MEXICO	0.001595	0.012453
5	26	SONORA	0.000000	0.012907
6	14	JALISCO	0.002158	0.013860
7	28	TAMAULIPAS	0.000000	0.017206
8	32	ZACATECAS	0.000000	0.017776
9	17	MORELOS	0.000000	0.019511
10	16	MICHOACAN	0.005096	0.020658
11	22	QUERETARO	0.001653	0.026843
12	27	TABASCO	0.001930	0.029276
13	30	VERACRUZ	0.015016	0.029580
14	25	SINALOA	0.000465	0.030359
15	4	CAMPECHE	0.009112	0.032334
16	18	NAYARIT	0.000000	0.032444
17	19	NUEVO LEON	0.000000	0.033428
18	24	SAN LUIS POTOSI	0.009233	0.034189
19	23	QUINTANA ROO	0.000544	0.034450
20	1	AGUASCALIENTES	0.000000	0.035425
21	5	COAHUILA	0.000000	0.036539
22	7	CHIAPAS	0.009452	0.036780
23	20	OAXACA	0.019287	0.039709
24	6	COLIMA	0.010446	0.039978
25	29	TLAXCALA	0.011323	0.048397
26	11	GUANAJUATO	0.019192	0.051346
27	13	HIDALGO	0.012983	0.056427
28	21	PUEBLA	0.034755	0.058455
29	31	YUCATAN	0.028539	0.072559
30	12	GUERRERO	0.045497	0.078265
31	3	B. C. SUR	0.000000	0.146345

Notas: No se deben hacer inferencias en los estados sombreados pues $cv > 0.2$.
En el estado de Chihuahua se dio un cero muestral en esta categoría.

TABLA 6.12 : TAMAÑOS DE MUESTRA EN PRIMERA (LOCALIDADES) Y SEGUNDA ETAPA (NIÑOS) POR ESTADO

ESTADO	ESTRATOS	TAMAÑO DE MUESTRA	
		LOCALIDADES	NIÑOS
1 AGUASCALIENTES	5	10	385
2 BAJA CALIFORNIA	2	6	164
3 B. C. SUR	4	11	90
4 CAMPECHE	8	24	1024
5 COAHUILA	5	15	409
6 COLIMA	4	12	247
7 CHIAPAS	25	57	1897
8 CHIHUAHUA	4	9	250
10 DURANGO	8	21	581
11 GUANAJUATO	19	38	845
12 GUERRERO	20	54	1887
13 HIDALGO	15	32	997
14 JALISCO	12	35	886
15 MEXICO	27	54	1209
16 MICHOACAN	20	40	1435
17 MORELOS	11	25	669
18 NAYARIT	5	10	207
19 NUEVO LEON	3	9	206
20 OAXACA	30	60	1728
21 PUEBLA	25	59	2413
22 QUERETARO	9	18	651
23 QUINTANA ROO	3	9	209
24 SAN LUIS POTOSI	11	32	1192
25 SINALOA	14	33	742
26 SONORA	6	12	243
27 TABASCO	17	34	1367
28 TAMAULIPAS	5	12	332
29 TLAXCALA	4	9	380
30 VERACRUZ	36	77	2635
31 YUCATAN	6	17	674
32 ZACATECAS	9	20	746

TABLA 6.13 : RESULTADOS DE LOS ESTIMADORES DE RAZON
E INTERVALOS DE CONFIANZA AL 95%, OBTENIDOS MEDIANTE EL JACKKNIFE,
POR GRANDES REGIONES DE LA REPUBLICA MEXICANA

REGION <i>Estados en la region</i>	TAMANOS EN LA MUESTRA DE:			GRUPO NUTRICIONAL	ESTIMADOR JACKKNIFE		INTERVALO DE CONFIANZA AL 95% **	
	ESTRATOS	LOCS.	FAMILIAS*		PROP.	VARIANZA	LIM.	LIM.
							INFERIOR	SUPERIOR
1 NOROESTE <i>Baja California, B.C.Sur, Sonora, Chihuahua, Sinaloa, Durango</i> CV= 0.08472	38	92	1564	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.729457 0.130186 0.052695 0.08056 0.007102	0.00008247 0.0001308 0.00003029 0.00015899 0.00000557	0.711250 0.107257 0.041661 0.055280 0.002370	0.747664 0.153115 0.063729 0.105840 0.011834
2 NORESTE <i>Coahuila, Nuevo Leon, Tamaulipas, Zacatecas, San Luis Potosi, Aguascalientes.</i> CV= 0.05896	38	98	2316	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.618843 0.183781 0.105689 0.077972 0.013735	0.00016347 0.00011587 0.0000966 0.00004357 0.00000718	0.593268 0.162229 0.086029 0.064769 0.008375	0.844418 0.205293 0.125349 0.091175 0.019095
3 CENTRO OCCIDENTE <i>Nayarit, Jalisco, Colima, Guanajuato, Queretaro, Michoacan</i> CV= 0.06172	68	153	3042	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.600743 0.206682 0.094516 0.079439 0.018621	0.00018128 0.00009286 0.0000614 0.0000515 0.00000737	0.573973 0.187522 0.078936 0.065170 0.013223	0.827513 0.225842 0.110096 0.093708 0.024019
4 CENTRO <i>Mexico, Hidalgo, Tlaxcala, Puebla, Morelos, Veracruz.</i> CV= 0.03819	118	256	5881	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.479611 0.262861 0.160165 0.068318 0.029044	0.00010965 0.00006835 0.00003611 0.00002931 0.00000665	0.458906 0.246514 0.148283 0.057613 0.024003	0.500316 0.279208 0.172047 0.079023 0.034085
5 SUR <i>Guerrero, Oaxaca, Chiapas, Tabasco, Campeche, Yucatan, Quintana Roo.</i> CV= 0.03853	109	255	6185	NORMALES y casi N. BAJITOS GORDITOS MAL PARA LA EDAD MAL PARA LA TALLA MAL EDAD Y TALLA	0.417104 0.290987 0.198278 0.061132 0.0325	0.00008565 0.00007209 0.00007389 0.0000152 0.00000884	0.398813 0.274207 0.181289 0.053427 0.026624	0.435395 0.307767 0.215267 0.068837 0.038376

* Se indica el número de familias encuestadas, en las que hubo niños menores de 6 años.

** Los intervalos se basan en una "t" con grados de libertad dados por: #locs. encuestadas -estratos.