Latinoamericano de Demografía

CELADE



# EVALUACION Y CORRECCION DE DATOS DEMOGRAFICOS

## VI. Análisis de la población por edades

Juan Chackiel y Guillermo Macció



Serie B, N° 39
Octubre, 1978
Santiago de Chile

Este documento forma parte de un texto más amplio, constituido por 10 capítulos, cuyo título es Evaluación y Corrección de Datos Demográficos.

La publicación independiente de este capítulo y de otros que aparecerán posteriormente, se hace con la finalidad de utilizarlos como material docente.

Los autores agradecen, desde ya, los comentarios y sugerencias que les puedan hacer llegar, pues serán de gran utilidad para la versión definitiva.

# I N D I C E



## Indice de cuadros y gráficos

# VI. ANALISIS DE LA POBLACION POR EDADES

## 1. Importancia de la determinación correcta de la edad

La información de la población clasificada por edad es uno de los elementos más importantes en el análisis demográfico. En forma resumida se puede mencionar los siguientes aspectos que hacen de esta información una de las más valiosas, a tal punto que merece una consideración especial.

a) En primer lugar, la población por edades es importante en sí misma. La estructura por edades de una población expresa muchas cosas por sí sola; en ella se reflejan aspectos relevantes de la historia de las poblaciones y de su situación actual. Una alta mortalidad en el pasado -causada por una guerra o una fuerte epidemia- e incluso una variación brusca de la mortalidad, la fecundidad, las migraciones, dejan sus huellas marcadas, como fácilmente se percibe en la pirámide de población. Es en esta estructura donde se perciben con mayor facilidad la presencia de los llamados fenómenos hereditarios.

Los planificadores y estudiosos de casi todos los servicios a la comunidad, deben tomar en consideración de alguna manera a la población de acuerdo a su edad. Por ejemplo, el porcentaje de población joven muestra el potencial posible de asistencia escolar, la población entre 15 y 64 años es aproximadamente el potencial de mano de obra de una nación, etc.

b) Pero no sólo es útil la población por edades en sí misma, sino que resulta aún más interesante el estudio de otras variables cruzadas con la edad de las personas. Esto es así porque casi todos los aspectos relacionados con las poblaciones se comportan de distinta forma de acuerdo a la edad de las mismas. La edad es la variable más importante en el estudio de fenómenos como la población económicamente activa, la mortalidad, la fecundidad, la nupcialidad, la migración, etc.

c) Errores que a veces no se han detectado o que aparecen muy confusos, se manifiestan claramente al desagregar la información según la edad. Recuérdese que existe un comportamiento más o menos esperado de la población por edades, lo que unido a una serie de técnicas de evaluación, permite tener algunas ideas de la calidad de los censos.

d) La estructura por edades es también de gran valor para las estimaciones de las variables demográficas pues en ella se expresan niveles y variaciones. Existen muchas metodologías de estimación indirecta de fecundidad y mortalidad que se basan en parámetros que expresan la distribución relativa de la población según la edad. Un ejemplo muy conocido es la Relación Niños-Mujeres como indicador del nivel de la fecundidad.

Si los datos de población por edades son de mala calidad, los aspectos mencionados antes resultarán afectados, provocando conclusiones falsas; de ahí la necesidad de un concienzudo y profundo análisis de esa información.

Dentro del examen de la calidad de la población por edades, pueden distinguirse tres aspectos:

- La edad ignorada.
- La mala declaración de la edad.
- La omisión diferencial por edades.

## 2. El problema de la edad ignorada

Es común encontrar al pie de los cuadros que presentan la población por grupos de edades, un grupo denominado "edad ignorada" que se refiere a aquellas personas para las cuales no se cuenta con esa información. El problema que se plantea en esta sección es qué hacer con ese grupo; ¿qué edad asignarles? Para resolver esto se utilizan en general dos procedimientos:

### 2.1. Tratamiento manual

El más usado y que se recomienda en el Manual III de Naciones Unidas[1/] es el "prorrateo" por separado para cada sexo. Consiste en distribuir el grupo

1/ Naciones Unidas, Métodos para preparar Proyecciones de Población por Sexo y Edad, Manual III, ST/SOA/Serie A, N° 25. Nueva York, 1956, pág. 16.

de peroonas de edad ignorada en forma proporcional a la población que aparece en los distintos grupos de edades. El supuesto básico es que la ignorancia o falta de información de la edad no es diferencial según la edad.

Este procedimiento es útil cuando se trata de un número reducido de personas, pero cuando su número cobra importancia, prorratearlas puede ser riesgoso, porque probablemente se trata de personas con características muy especiales que pertenecen a edades particulares. En estos casos, quizás convenga conservar el total de población y proceder a una revisión completa de la distribución de todas las edades. En particular podrá recurrirse a fórmulas de ajuste o a modelos sustitutivos, pues la idea que subyace tiene que ver con la calidad que afecta a todo el censo, siendo una de sus manifestaciones el alto peso de la población ignorada.

2.2. Tratamiento mecánico

Consiste en programas de computación incorporados al procesamiento del mismo censo y que asignan una edad para cada caso individual cuando el dato no ha sido recogido. Toma en cuenta para ello diferentes criterios, como son: la unidad geográfica y las características de la persona en cuestión y va acotando las posibilidades hasta asignar un valor lógico. Por ejemplo, si se trata de una mujer que es activa y tiene hijos pequeños, se supone que su edad está entre 20 y 40 años y así se van limitando las posibilidades.

Si es el sexo el dato ignorado, la forma de asignación se ve más clara. Por ejemplo, si hay datos de fecundidad, se trata de una mujer.

En el caso que no haya indicio suficiente, se ubica a otra persona de la misma unidad geográfica con las mismas características conocidas y se le asigna su edad.

Si el número de datos ignorados es grande, este método también puede conducir a resultados muy equivocados. La debilidad está fundamentalmente en los criterios adoptados, pues si la selección de éstos fue mal hecha, puede conducir a errores sistemáticos. En este caso, lo recomendable es, aun realizando la asignación, conservar también la información en su forma original. En estas circunstancias para fines de control resulta muy aconsejable disponer de un listado independiente de las asignaciones hechas.

## 3. Mala declaración de la edad

La mala declaración de la edad es un fenómeno que ocurre, en mayor o menor medida, en todos los países del mundo. Esta situación se torna grave en los países en vías de desarrollo, poniendo en peligro la posibilidad de uso adecuado de la información, por lo que se hace necesario un detenido análisis de ese dato antes de proceder a utilizarlo.

Es oportuno advertir que los patrones de error en la declaración de la edad y el efecto que producen sobre los datos está estrechamente ligado con aspectos culturales propios de cada país o región.

Los tipos de error que más a menudo afectan la declaración de la edad, pueden clasificarse como sigue:

a) Preferencia de dígitos. Se refiere a que las personas suelen declarar su edad redondeando sistemáticamente el dígito final. Lo más común es que haya una atracción mayor por declarar edades terminadas en 0 y en menor medida en 5, lo que causa un aumento notable en las edades 10, 15, 20, 25, 30, etc. y un defecto en las demás. Los dígitos de preferencia dependen en gran medida de la cultura de los países y de la forma de investigar el dato en cuestión. En los países de América Latina es común que la preferencia sea por las edades terminadas en 0 y 5 y, en menor medida, por las edades que terminan en números pares. Este error se manifiesta también cuando la edad se investiga a través de la fecha de nacimiento, en cuyo caso la preferencia surge como consecuencia del redondeo del año de nacimiento.

b) Traslado de edades. Mientras que en el caso de la preferencia de dígitos existe un traslado desde edades próximas mayores y menores hacia el dígito de atracción, en el caso de "traslado de edades" se declara la edad sistemáticamente por debajo o sistemáticamente por arriba de la verdadera. Por ejemplo, con frecuencia ocurre que los ancianos tienen la tendencia a declarar una edad mayor a la real. También se constata que las mujeres a cierta edad están inclinadas a declararse más jóvenes de lo que son. Este tipo de error produce situaciones más graves que el anterior, a causa de que por un lado significa una modificación mayor en la estructura por edades, haciendo variar la edad media de la población, y además resulta más complicado detectar su existencia y cuantificar su importancia.

c) Preferencia o rechazo por una edad específica. Se trata del caso de una edad por la cual independientemente de la preferencia de dígitos, existe una atracción o un rechazo especial. Esto parecería observarse, por ejemplo, con la edad 60 de alta preferencia, en que a menudo se ubican muchas personas de edad avanzada, de las cuales se ignora su verdadera edad. Un caso bien conocido de rechazo es el de la edad 13; es fácil encontrar en algunos países atracciones o rechazos basados en creencias o valores locales.

¿Cuáles son las causas para que se declare la edad en forma errónea? Se podría mencionar fundamentalmente las siguientes:

a) Que se cometa un error al tratar de calcularla si no se la ha memorizado.

b) Que el informante desconozca la edad de la persona sobre la cual informa. Es muy común que en un censo la persona que informa no conozca en forma precisa algunas características de terceros.

c) Puede ocurrir que por determinada circunstancia, cierto sector de la población, ya sean hombres o mujeres, falseen deliberadamente la declaración de la edad. En América Latina existen abundantes ejemplos de preferencias, rechazos o traslado de edades derivados de creencias o valores muy acentuados en algunos sectores.

d) Que haya una errónea intelección del sentido de la pregunta y se dé una respuesta en el entendido de que se ha solicitado otra cosa, lo que constituye un caso más de la ya señalada falta de una adecuada comunicación entre el entrevistador y el entrevistado.

Cada una de estas causas que por lo general se dan en forma combinada, determinan un patrón particular de mala declaración lo que implica, por ejemplo, preferencia de dígitos y traslación de la edad. La segunda causa mencionada, el desconocimiento del informante, puede conducir a preferencia de dígitos por redondeo de la edad; la persona puede razonar que el tercero sobre el cual está informando tiene "aproximadamente 30 años" y declarar 30, aun cuando se trata de una estimación cuya proximidad al dato real se desconoce. Un ejemplo de error por falsedad deliberada, puede ser el rejuvenecimiento de las edades de las mujeres, o de los jóvenes que están en edad de entrar al servicio militar, lo que, como se ha dicho, lleva a una traslación de la edad.

Otro aspecto que repercute directamente sobre el tipo de error es la forma en que se investiga la edad. Esta información puede averiguarse directamente o indirectamente a través de la "fecha de nacimiento" o de "acontecimientos calendario".

La forma directa consiste en preguntar la edad de la persona en años cumplidos, es decir, la edad alcanzada en el último cumpleaños. Es probable que este tipo de preguntas conduzca a redondear la edad en los dígitos 0 y 5 y a cierto tipo de traslados.

Se cree que una solución para evitar estos errores es preguntar la fecha o año de nacimiento y, a partir de ese dato, deducir la edad al momento del censo. Sin embargo, es posible demostrar que esta forma de averiguar la edad genera sus propios errores. En primer lugar, puede ocurrir que la gente redondee el año de nacimiento en 0 y 5, lo que traería como consecuencia una atracción por ciertos dígitos muy particulares. Supóngase un censo en el año 1973 en el cual se investiga la edad a través del año de nacimiento; en ese caso si el año de nacimiento se redondea en 0 y 5, habrá preferencia en las edades terminadas en 3 y 8. Además, como muchas personas calculan su fecha de nacimiento a partir de la edad, se encontrará de todas maneras atracción por las edades terminadas en 0 y 5.

En algunos países de menor grado de desarrollo, debido a que es común que no se sepa la edad ni la fecha de nacimiento, se ha tratado de investigar este dato asociando la fecha de nacimiento con un listado de hechos importantes acaecidos en la región. Estos hitos ayudan a ubicar la fecha aproximada del nacimiento de la persona. Por ejemplo, si un acontecimiento importante considerado fuera el fin de la Segunda Guerra Mundial, y el informante señala que nació en el año que sucedió ese hecho, se supone que fue en 1945; si ya había nacido, por lo menos ya sabemos que nació antes de ese año. De esa forma, puede ir acotándose hasta dar con el año de nacimiento o un intervalo aceptable que permitirá calcular la edad.

Dos problemas difíciles de resolver son: por un lado, la confección de eventos de importancia para todos los años y, por otro, lograr que esos hechos tengan significación para todas las regiones y sectores de la población del país. Además, esta forma de averiguar la edad produce patrones de error

mucho más difíciles de detectar, aunque es de esperar que, de acuerdo al significado que tenga para la población cada evento histórico especificado, se den atracciones por algunas edades en particular.

Desde luego, existe siempre la posibilidad de combinar las tres formas comentadas para averiguar la edad.

Comentario aparte merece la omisión diferencial según la edad. No se trata de una edad mal declarada ni de la ignorancia de esa característica, sino de la no enumeración de la persona por muy distintas razones, pero que afecta la distribución relativa por edades de la población, en el caso que esa omisión sea diferencial según la edad.

Cuando una población presenta irregularidades en su estructura por edades, por ejemplo, poca población en edades centrales, a veces es difícil delimitar si se debe a mala declaración de la edad o a omisión diferencial. Ambos aspectos pueden ser tratados en forma separada, pero debido a lo anterior, en general se tratarán en conjunto.

## 4. Evaluación de la calidad de la declaración de la edad

Como ocurre en general cuando se trata de evaluar la calidad de cierta información, no es posible señalar en forma rígida cuáles serían los pasos que deberían darse. Esto depende de cómo viene presentada la información (por ejemplo, en edades individuales o en grupos quinquenales), de los elementos con que se cuenta para realizar la evaluación y de las irregularidades que se vayan detectando sobre la marcha. Por lo tanto, las etapas a cumplir en la evaluación, que se presentan en lo que sigue, pretenden ser sólo una guía, y de ninguna manera agotan todo lo que se puede y debe hacer para determinar la calidad de la declaración de la edad.

### 4.1. Examen minucioso de la información

La simple observación crítica de los datos brinda algunas ideas esenciales sobre la calidad de los mismos. Esto a la luz de lo que se sabe sobre el comportamiento esperado de la población por edades y del conocimiento que se tiene con referencia a ciertos aspectos sobresalientes del país en estudio cuya incidencia se refleja en la estructura de la población.

A título de ejemplo se analizarán los datos incluidos en el cuadro 1 y representados en el gráfico 1, de la población por sexo y edad de Guatemala tomada del Censo de población de 1973.

Del examen de la clasificación por edades simples, puede concluirse de inmediato que existe una preferencia por determinadas edades, sobre todo a partir de los 10 años. Como regla general, es mayor la atracción ejercida por las edades terminadas en 0, luego las terminadas en 5 y, en menor medida, las terminadas en 8 y en 2. Los rechazos mayores son por las edades terminadas en los dígitos 1 y 9, obviamente por ser las más próximas al 0 de mayor atracción y además recibir también la influencia absorbente de los dígitos 2 y 8, respectivamente. Esto bajo el supuesto que las atracciones afectan en mayor medida a los dígitos adyacentes.

La población reunida en grupos quinquenales de edades muestra un comportamiento más suave, lo que no significa que se elimine la preferencia de dígitos, sólo se disimula. La diferente atracción por el 0 y el 5 se mantiene, pero afecta ahora a los grupos en forma alternada; de esta forma el grupo 10-14 contiene el dígito 0; el 15-19, el dígito 5; el 20-24, nuevamente contiene el 0, y así sucesivamente.

Además de ese patrón de error en la declaración de la edad, puede detectarse otras situaciones irregulares que merecen ser investigadas con más detalle:

a) Los menores de 1 año tienen un valor desproporcionadamente grande con respecto a las edades que le siguen. Si no se debiera esto a un violento aumento de la fecundidad en el último año, tendría que vincularse a errores de empadronamiento o mala declaración en las edades iniciales de la vida.

b) Los grupos quinquenales de edades 30-34 y 35-39 son prácticamente iguales, sobre todo en el sexo femenino. Si no se tienen indicios de la existencia de hechos demográficos que afectan o han afectado a la población que ahora aparece en estos grupos de edades (brusco descenso en la fecundidad hace aproximadamente 30 años, migraciones o mortalidad diferenciales) se debería encontrar una explicación por el lado de la calidad de la información. Pudiera ser, y en parte se observa en el gráfico, que la edad 35 fuera de

Cuadro 1

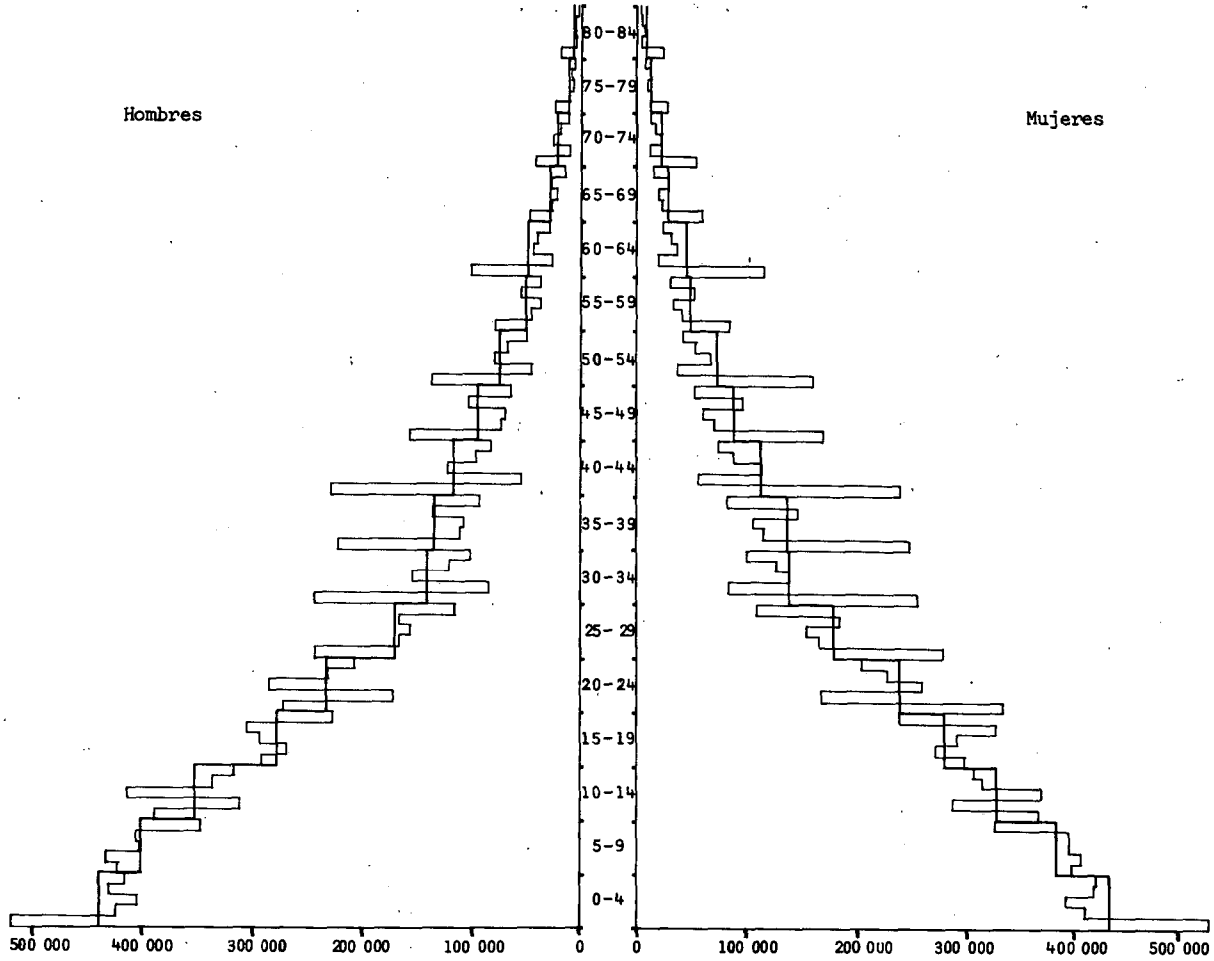
GUATEMALA: POBLACION TOTAL POR SEXO, EDADES SIMPLES Y GRUPOS DE EDADES, CENSO 1973

| Edad | Hombres | Mujeres | Edad | Hombres | Mujeres | Edad | Hombres | Mujeres | Edad | Hombres | Mujeres |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Total | 2 589 264 | 2 570 957 | | | | | | | | | |
| 0- 4 | 439 571 | 430 806 | 25-29 | 168 921 | 178 339 | 50-54 | 75 332 | 71 672 | 75-79 | 11 911 | 11 845 |
| 0 | 108 876 | 106 434 | 25 | 48 160 | 56 171 | 50 | 27 638 | 31 506 | 75 | 4 852 | 5 508 |
| 1 | 83 897 | 81 109 | 26 | 32 703 | 32 952 | 51 | 8 776 | 7 448 | 76 | 2 057 | 1 906 |
| 2 | 79 061 | 77 432 | 27 | 30 836 | 30 641 | 52 | 15 553 | 13 170 | 77 | 1 699 | 1 408 |
| 3 | 83 263 | 82 664 | 28 | 34 232 | 36 706 | 53 | 12 995 | 10 461 | 78 | 2 165 | 2 038 |
| 4 | 84 474 | 83 167 | 29 | 22 990 | 21 869 | 54 | 10 370 | 9 087 | 79 | 1 138 | 985 |
| 5- 9 | 395 066 | 382 669 | 30-34 | 141 113 | 141 369 | 55-59 | 51 555 | 48 110 | 80-84 | 6 744 | 7 636 |
| 5 | 82 891 | 80 209 | 30 | 48 717 | 51 934 | 55 | 16 016 | 17 421 | 80 | 3 458 | 4 632 |
| 6 | 83 738 | 81 210 | 31 | 16 400 | 15 009 | 56 | 9 600 | 8 327 | 81 | 693 | 596 |
| 7 | 79 475 | 76 929 | 32 | 29 208 | 29 233 | 57 | 7 692 | 6 365 | 82 | 978 | 940 |
| 8 | 80 131 | 77 870 | 33 | 26 659 | 25 311 | 58 | 10 662 | 10 227 | 83 | 982 | 821 |
| 9 | 68 831 | 66 451 | 34 | 20 129 | 19 882 | 59 | 7 585 | 5 770 | 84 | 633 | 647 |
| 10-14 | 348 438 | 329 722 | 35-39 | 131 001 | 137 312 | 60-64 | 48 651 | 46 375 | 85-89 | 2 958 | 3 372 |
| 10 | 76 577 | 73 229 | 35 | 43 428 | 48 644 | 60 | 20 744 | 23 763 | 85 | 1 254 | 1 517 |
| 11 | 61 027 | 58 111 | 36 | 22 036 | 22 533 | 61 | 5 012 | 3 927 | 86 | 482 | 549 |
| 12 | 82 324 | 74 843 | 37 | 20 721 | 20 581 | 62 | 8 692 | 7 330 | 87 | 422 | 463 |
| 13 | 66 166 | 62 072 | 38 | 26 849 | 28 477 | 63 | 8 384 | 6 326 | 88 | 497 | 543 |
| 14 | 62 344 | 61 467 | 39 | 17 967 | 17 077 | 64 | 5 819 | 5 029 | 89 | 303 | 300 |
| 15-19 | 274 689 | 285 590 | 40-44 | 116 391 | 113 387 | 65-69 | 28 192 | 28 047 | 90-94 | 1 195 | 1 382 |
| 15 | 57 926 | 59 175 | 40 | 43 404 | 47 998 | 65 | 10 033 | 11 569 | 90 | 751 | 964 |
| 16 | 53 170 | 54 598 | 41 | 12 431 | 11 056 | 66 | 5 132 | 4 431 | 91 | 102 | 95 |
| 17 | 58 401 | 58 373 | 42 | 24 724 | 22 575 | 67 | 4 472 | 3 848 | 92 | 158 | 138 |
| 18 | 60 166 | 66 143 | 43 | 29 705 | 17 534 | 68 | 5 378 | 5 539 | 93 | 126 | 116 |
| 19 | 45 026 | 47 301 | 44 | 16 127 | 14 224 | 69 | 3 177 | 2 660 | 94 | 58 | 69 |
| 20-24 | 230 991 | 239 281 | 45-49 | 93 389 | 91 113 | 70-74 | 22 301 | 21 913 | 95-99 | 855 | 1 017 |
| 20 | 55 077 | 65 299 | 45 | 31 621 | 34 133 | 70 | 8 477 | 10 268 | 95 | 200 | 279 |
| 21 | 34 987 | 33 780 | 46 | 14 900 | 13 732 | 71 | 2 172 | 1 933 | 96 | 89 | 100 |
| 22 | 54 112 | 53 033 | 47 | 13 659 | 12 612 | 72 | 5 045 | 4 304 | 97 | 62 | 52 |
| 23 | 46 599 | 45 931 | 48 | 20 418 | 19 848 | 73 | 4 332 | 3 429 | 98 y + | 504 | 586 |
| 24 | 40 216 | 41 238 | 49 | 12 791 | 10 788 | 74 | 2 275 | 1 979 | | | |

Fuente: Dirección General de Estadística de Guatemala, VIII Censo de Población, 26 de marzo de 1973, Cifras Definitivas, Serie III, Tomo I, Guatemala, agosto de 1975.

Gráfico 1

GUATEMALA: POBLACION SEGUN SEXO Y EDAD EXPRESADA EN GRUPOS INDIVIDUALES Y QUINQUENALES, CENSO DE 1973




Fuente: Cuadro 1.

atracción especialmente mayor que las demás, produciéndose una sobreestimación en el grupo 35-39, en desmedro de 30-34. Otras posibles explicaciones estarían relacionadas con probables traslados de edades hacia 35-39 o por omisión diferencial entre estos grupos.

c) Con las edades 55-59 y 60-64 ocurre un fenómeno igual al anterior, pero esta vez con más énfasis en hombres que en mujeres. También la explicación puede estar en una atracción particular por la edad 60. En las personas ancianas es más difícil conocer la edad y a veces el redondeo se realiza de edades más alejadas. Por ejemplo, si una persona tiene 13 años, difícilmente se va a redondear en 20; en cambio, es más probable que a una persona de 53 años se le adjudique la edad 60. Es posible, por otra parte, que las personas ancianas exageren su edad[2/] y que ese fenómeno se haga más notable en Guatemala alrededor de los 55 años.

Aun cuando la percepción de errores evidentes varía de país a país, el ejercicio precedente sirve para identificar errores como los señalados y ayuda a orientar los pasos posteriores de la evaluación.

4.2. Comparación con una población estándar

Avanzando un poco más en el análisis, resulta aconsejable, bajo ciertas condiciones, realizar una comparación con alguna población estándar, ya sea de un país con similares condiciones demográficas y buena información o con un modelo teórico, como por ejemplo, una población estable. Este tipo de examen, al igual que el anterior, no tiene la intención de cuantificar errores, sino que, a través de las desviaciones de la población en estudio y el estándar elegido se trata de seguir estableciendo posibles errores y omisiones según la edad.

El gráfico 2 compara la población por edades del censo de Guatemala de 1973 con una población estable, tomada de las tablas de Coale y Demeny.[3/]

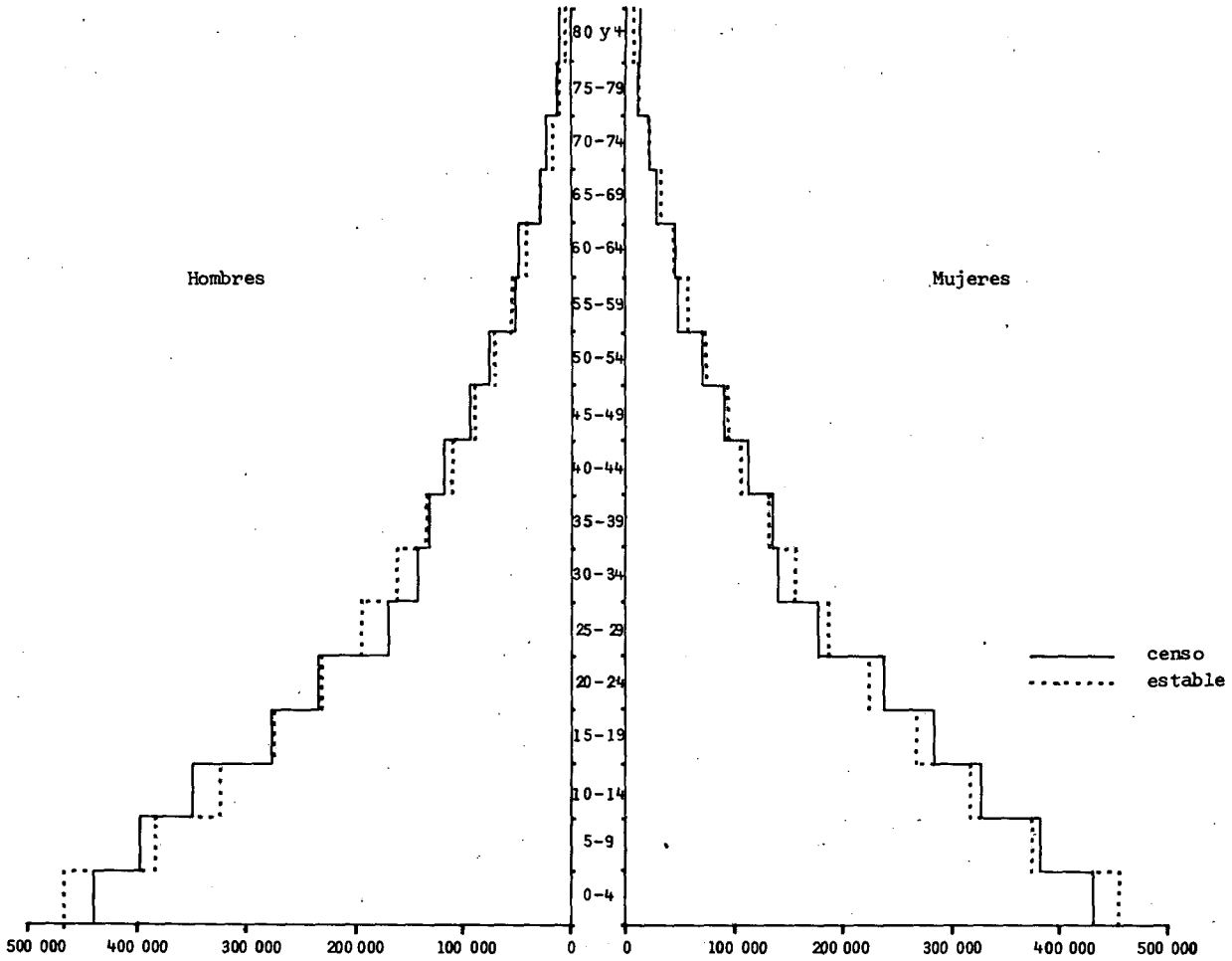
---

2/ Hobcraft, S., Seminario sobre Estimaciones Demográficas, San José, 1° al 9 de agosto de 1974, CELADE, Serie D, N° 94, Sesión IV.

3/ Coale, A. y Demeny, P., Regional Model Life Tables and Stable Populations, Princeton University Press, Nueva Jersey, 1966.

Gráfico 2

GUATEMALA 1973: COMPARACION DE LA POBLACION CENSAL SEGUN SEXO Y EDAD CON UNA POBLACION ESTABLE




Fuente: Cuadro 1.
Coale y Demeny; Regional Model Life Tables and Stable Populations. Princeton University Press. New Jersey 1966.

(Nivel 15 de mortalidad y Tasa Bruta de Reproducción de 3,0). De este gráfico puede sacarse las siguientes conclusiones tentativas:

a) Déficit en el grupo 0-4 años, lo que es un resultado esperado debido a lo que indica la experiencia en los Censos sobre omisión, particularmente importante en esas edades.

b) Déficit en edades centrales, fundamentalmente de 25 a 35 años y en el sexo masculino. Este hecho puede deberse a un posible saldo migratorio negativo o a omisión de cierta relevancia en este tramo de edades, aunque no se debería descartar un posible traslado de edades e incluso un efecto combinado de estos elementos.

c) Corroborando lo analizado con respecto a los grupos 55-59 y 60-64, se observa un déficit en el primero y un exceso en el segundo, asociado a la fuerte preferencia por la edad 60 o un posible traslado de edades de 55 años en adelante.

Conviene insistir en el carácter meramente exploratorio y limitado de estas comparaciones, pues es probable que la evolución no regular de la fecundidad y la mortalidad, lo mismo que los movimientos migratorios del país en consideración, sean la causa de que su estructura por edades se aparte de la distribución estándar elegida.

## 5. Cuantificación de la exactitud en la declaración de la edad

La sección está dedicada a presentar algunos métodos destinados a cuantificar la calidad de la declaración de la edad. Estos permitirán realizar comparaciones entre las declaraciones en los censos tanto a través del tiempo como del espacio. Para un país determinado, podrá investigarse si de un censo al siguiente hay una mejora en la declaración, si ésta es mejor en la zona urbana que en la rural, o para hombres que para mujeres, etc.

Como ilustración se mostrarán aquí los índices de Whipple, Myers, el de la Secretaría de las Naciones Unidas y el de usos de poblaciones estándar. Otros autores han elaborado índices un poco más complicados, pero de menor uso, como es el caso de Bacchi, Michalup, Das Gupta, Carrier, Ramachandran, etc.

5.1. El Indice de Whipple (o de concentración)[4/]

En su versión original está destinado a medir la preferencia por los dígitos 0 y 5 en conjunto, pero puede ser adaptado para establecer el nivel de atracción por cualquier dígito.

Así, si se deseara medir la preferencia por las edades terminadas en 0, el índice sería:

$$I = \frac{10(N_{30} + N_{40} + N_{50} + N_{60})}{\sum_{25}^{64} N_x} 100 = \frac{10 \sum_{x=3}^{x=6} N_{10x}}{\sum_{25}^{64} N_x} 100$$

donde $N_x$ es la población de edad $x$.

Bajo el supuesto de que sólo podría haber preferencias por el dígito 0, los resultados oscilarán entre un mínimo de 100, en el caso de información bien declarada y un máximo de 1 000 en el caso de que todos hayan declarado edades terminadas en 0, es decir, cuando la concentración es máxima. Si el índice resultara menor que 100, estaría indicando un rechazo por el 0.

Este mismo índice, tal como se dijo, puede aplicarse para medir la atracción de cualquier dígito o de dos a la vez. Para las edades terminadas en 0 y 5 sería:

$$I = \frac{5 \sum_{5}^{12} N_{5x}}{\sum_{23}^{62} N_x} 100$$

En general la versión más usada es ésta, que corresponde a la propuesta por Whipple.

---

4/ Naciones Unidas, Métodos para Evaluar la Calidad de los Datos Básicos Destinados a los Cálculos de la Población, Manual II, ST/SOA/Serie A, N° 23, Nueva York, 1955, pág. 45.

El supuesto básico del índice es que la población correcta varía en forma lineal dentro de los grupos de edades 23-27, 28-32, 33-37...58-62:

$$N_x = N_{23} + (x-23)a \qquad \text{para } x = 23 \text{ a } 27$$

$$N_x = N_{28} + (x-28)b \qquad \text{para } x = 28 \text{ a } 32$$

$$N_x = N_{33} + (x-33)c \qquad \text{para } x = 33 \text{ a } 37$$

.....

.....

.....

$$N_x = N_{53} + (x-53)g \qquad \text{para } x = 53 \text{ a } 57$$

$$N_x = N_{58} + (x-58)h \qquad \text{para } x = 58 \text{ a } 62$$

El índice sería:

$$\frac{5\left[(N_{23}+2a) + (N_{28}+2b) + \ldots + (N_{58}+2h)\right]}{(5N_{23}+10a) + (5N_{28}+10b) + \ldots + (5N_{58}+10h)} \; 100 = 100$$

es decir, que cuando la información es exacta, de acuerdo a la hipótesis, el índice vale 100, como se indicó anteriormente.

En el cuadro 2 se muestran los índices de Whipple para los países de América Latina, Estados Unidos y Suecia en los censos de alrededor de 1950, 1960 y 1970. De acuerdo con una escala propuesta por Naciones Unidas[5/] que plantea una asociación entre la preferencia de dígitos y la calidad de los datos censales, los países de la región, en su gran mayoría, tendrían información que fluctuaría entre mala y muy mala, aunque mejorando en el transcurso

---

5/ Naciones Unidas, Métodos para Evaluar ... op.cit.
Se propone la siguiente escala:
Indices de 100 a 105 corresponden a "datos muy precisos"
Indices de 105 a 110 corresponden a "datos relativamente precisos"
Indices de 110 a 125 corresponden a "datos aproximados"
Indices de 125 a 175 corresponden a "datos malos"
Indices de 175 y más corresponden a "datos muy malos"

Cuadro 2

PAISES DE AMERICA LATINA Y ESTADOS UNIDOS Y SUECIA: INDICE DE WHIPPLE PARA LA POBLACION TOTAL

| Países | Censos alrededor de | | |
|---|---|---|---|
| | 1950 | 1960 | 1970 |
| Argentina | 98,8 | 104,8 | 103,9 |
| Bolivia | 211,1 | b/ | a/ |
| Brasil | 140,0 | 143,6 | 127,6 |
| Colombia | a/ | 143,5 | a/ |
| Costa Rica | 150,7 | 123,5 | 120,6 |
| Cuba | 112,6 | b/ | a/ |
| Chile | 128,0 | 129,8 | 123,0 |
| Ecuador | 197,1 | 175,3 | a/ |
| El Salvador | 198,3 | 181,6 | 161,7 |
| Guatemala | 196,8 | 158,3 | 163,9 |
| Haití | 219,6 | b/ | 170,6 |
| Honduras | a/ | 151,3 | 128,5 |
| México | 188,7 | 174,0 | 148,0 |
| Nicaragua | 195,0 | 180,7 | 185,2 |
| Panamá | 141,6 | 120,9 | 118,1 |
| Paraguay | a/ | 117,4 | 112,6 |
| Perú | b/ | 145,1 | 125,1 |
| República Dominicana | 184,9 | 182,4 | 181,0 |
| Uruguay | b/ | 105,8 | a/ |
| Venezuela | 166,0 | 120,3 | a/ |
| Estados Unidos | 107,2 | 100,9 | 101,6 |
| Suecia | 101,6 | 101,5 | 100,9 |

a/ No se disponía de la información por grupos individuales de edades.
b/ No existió censo.

del tiempo, con tendencia a datos aproximados. Para 1970 se tendría, sin incluir Estados Unidos y Suecia, la siguiente distribución:

| | |
|---|---|
| Datos muy precisos: | 1 país |
| Relativamente precisos: | 0 países |
| Datos aproximados: | 4 países |
| Datos malos: | 7 países |
| Datos muy malos: | 2 países |

Por supuesto los índices de los dos países desarrollados, incluidos a título de comparación, corresponden al grupo de los "muy precisos", no alcanzando en el último censo ni siquiera el valor 102.

En el cuadro mencionado se incluyen los índices correspondientes a la población de ambos sexos, pero también es posible realizar los cálculos para hombres y mujeres por separado.

5.2. El índice de Myers[6/]

El procedimiento propuesto por Myers es un avance respecto al anterior, ya que permite determinar la atracción o repulsión que tiene cada dígito.

Una forma podría ser considerar la suma de población de cada dígito en todas las decenas de edad y compararlos. Si la población tuviera una distribución rectangular en cada decena, se esperaría que para cada dígito, la suma mencionada representara el 10 por ciento de la suma total. Este procedimiento no es aplicable porque se sabe que en un grupo decenal, por ejemplo, el último valor (dígito 9), es siempre menor que el anterior (dígito 8) y así sucesivamente. La propuesta de Myers es un intento de usar parte de esta idea pero corrigiendo esa situación, de modo que al "sumar" las edades de cada dígito de las distintas decenas, represente, en ausencia de error, un 10 por ciento del total. El cuadro 3 incluye la rutina de cálculo del índice de Myers. La llamada "suma combinada" debería representar siempre bajo el supuesto de ausencia de error, un 10 por ciento del total en cada dígito;

6/ Myers, R.J., "Errors and Bias in the Reporting Ages in Census Data", en Handbook of Statistical Methods for Demographers, U.S. Department of Commerce, Washington D.C., 1960.

la preferencia o rechazo se mide en la última columna estableciendo el desvío del porcentaje observado con respecto a 10. La aplicación se hace de 10 años de edad en adelante y, en algunas oportunidades, de 20, porque en las primeras edades el patrón de preferencia no es muy regular. Obsérvese además que para este índice se necesita la población por edades simples.

Se demuestra analíticamente que el supuesto básico de este procedimiento consiste en aceptar un comportamiento lineal de la población con la variación de la edad.

$$N_x = N_0 + xd:$$

siendo $N_x$ la población de edad x.

| Dígito | Suma Combinada | Porcentaje | Desvío con respecto a 10 |
|---|---|---|---|
| 0 | $(8N_o+360d)\ 1 + (8N_o+440d)\ 9=80N_o+4320d$ | 10 | 0 |
| 1 | $(8N_o+368d)\ 2 + (8N_o+448d)\ 8=80N_o+4320d$ | 10 | 0 |
| 2 | $(8N_o+376d)\ 3 + (8N_o+456d)\ 7=80N_o+4320d$ | 10 | 0 |
| 3 | $(8N_o+384d)\ 4 + (8N_o+464d)\ 6=80N_o+4320d$ | 10 | 0 |
| 4 | $(8N_o+392d)\ 5 + (8N_o+472d)\ 5=80N_o+4320d$ | 10 | 0 |
| 5 | $(8N_o+400d)\ 6 + (8N_o+480d)\ 4=80N_o+4320d$ | 10 | 0 |
| 6 | $(8N_o+408d)\ 7 + (8N_o+488d)\ 3=80N_o+4320d$ | 10 | 0 |
| 7 | $(8N_o+416d)\ 8 + (8N_o+496d)\ 2=80N_o+4320d$ | 10 | 0 |
| 8 | $(8N_o+424d)\ 9 + (8N_o+504d)\ 1=80N_o+4320d$ | 10 | 0 |
| 9 | $(8N_o+432d)\ 10 + (8N_o+512d)\ 0=80N_o+4320d$ | 10 | 0 |

Una medida resumen de la preferencia de dígitos sería la suma en valores absolutos de la columna de desvío con respecto a 10. Teóricamente, el índice varía entre 0 y 180. En el caso de información correcta es obviamente 0 y en el caso de concentración absoluta en un dígito, por ejemplo, 0, el valor es 180.

Cuadro 3

CALCULO DEL INDICE DE MYERS PARA EL CENSO DE GUATEMALA DE 1973

| Dígito | Decenas de edades | | | | | | | Total A | Total B |
|---|---|---|---|---|---|---|---|---|---|
| | 10-19 | 20-29 | 30-39 | 40-49 | 50-59 | 60-69 | 70-79 | 10-69 | 20-79 |
| 0 | 151 080 | 121 0400 | 99 500 | 93 460 | 59 220 | 43 300 | 18 780 | 567 600 | 435 300 |
| 1 | 118 660 | 68 240 | 31 540 | 22 260 | 16 340 | 9 040 | 4 260 | 266 080 | 151 680 |
| 2 | 156 500 | 108 680 | 58 340 | 46 980 | 29 180 | 16 080 | 9 260 | 415 760 | 268 520 |
| 3 | 129 980 | 91 400 | 49 640 | 36 880 | 24 300 | 14 460 | 7 220 | 346 660 | 223 900 |
| 4 | 124 700 | 82 040 | 40 300 | 31 460 | 19 400 | 10 720 | 4 640 | 308 620 | 188 560 |
| 5 | 117 940 | 104 340 | 93 540 | 64 940 | 32 960 | 21 180 | 10 020 | 434 900 | 326 980 |
| 6 | 108 040 | 66 360 | 45 100 | 28 740 | 17 320 | 10 340 | 4 220 | 275 900 | 172 080 |
| 7 | 116 760 | 62 140 | 42 900 | 26 340 | 14 080 | 8 520 | 3 200 | 270 740 | 157 180 |
| 8 | 126 420 | 69 980 | 56 420 | 39 820 | 21 360 | 10 920 | 3 660 | 324 920 | 202 160 |
| 9 | 93 000 | 44 540 | 35 040 | 23 460 | 13 440 | 6 000 | 2 520 | 215 480 | 125 000 |
| Total | 1 243 080 | 818 760 | 552 320 | 414 340 | 247 600 | 150 560 | 67 780 | 3 426 660 | 2 251 360 |

| Dígito | Total A | Coeficiente A | Producto A | Total B | Coeficiente B | Producto B | Suma combinada Producto A+ Producto B | Porcentaje | Desvío respecto a 10 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 567 600 | 1 | 567 600 | 435 300 | 9 | 3 917 700 | 4 485 300 | 15,63 | 5,63 |
| 1 | 266 080 | 2 | 532 160 | 151 680 | 8 | 1 213 440 | 1 745 600 | 6,08 | -3,92 |
| 2 | 415 760 | 3 | 1 247 280 | 268 520 | 7 | 1 879 640 | 3 126 920 | 10,90 | 0,90 |
| 3 | 346 660 | 4 | 1 386 640 | 223 900 | 6 | 1 343 400 | 2 730 040 | 9,51 | -0,49 |
| 4 | 380 620 | 5 | 1 543 100 | 188 560 | 5 | 942 800 | 2 485 900 | 8,66 | -1,34 |
| 5 | 434 900 | 6 | 2 609 400 | 326 980 | 4 | 1 307 920 | 3 917 320 | 13,65 | 3,65 |
| 6 | 275 900 | 7 | 1 931 300 | 172 080 | 3 | 516 240 | 2 447 540 | 8,53 | -1,47 |
| 7 | 270 740 | 8 | 2 165 920 | 157 180 | 2 | 314 360 | 2 480 280 | 8,64 | -1,36 |
| 8 | 324 920 | 9 | 2 924 280 | 202 160 | 1 | 202 160 | 3 126 440 | 10,89 | 0,89 |
| 9 | 215 480 | 10 | 2 154 800 | 125 000 | 0 | 0 | 2 154 800 | 7,51 | -2,49 |
| Total | 3 426 660 | | | 2 251 360 | | | 28 700 140 | | /22,13/ |

En el cuadro 4 se incluyen los índices resumen para los países de América Latina y Estados Unidos y Suecia. Los países con muy buena información tendrían índices muy próximos a 0, por lo menos menores de 10. En América Latina se observan por lo general, valores superiores a 10 y en los últimos censos, inferiores a 30. Parecería que entre 20 y 30, se encontrarían los censos con alta preferencia de dígitos. Stockwell y Wicks[7/] utilizan como criterio de buena información el caso en que el índice es menor que 6 y de mala cuando el índice supera esta cifra. En un estudio realizado para la América Latina[8/] se propone la siguiente clasificación de los censos según la preferencia de dígitos:

| Nivel de atracción | Indice de Myers | Número de países de la América Latina | | |
|---|---|---|---|---|
| | | Censos de | | |
| | | 1950 | 1960 | 1970 |
| Bajo | 0,0 a 5,0 | 1 | 2 | 1 |
| Intermedio | 5,1 a 15,0 | 3 | 5 | 7 |
| Alto | 15,1 a 30,0 | 4 | 10 | 6 |
| Muy alto | 30,1 y más | 7 | 0 | 0 |

Obsérvese que en los últimos 20 años se ha producido una mejora sustancial en la declaración de la edad, pues de 11 países con "alta" y "muy alta" preferencia de dígitos en los censos de 1950, solamente hay 6 en "alta" preferencia y ninguno con "muy alta" en los censos de 1970.

El cuadro 4 permite ver la distinta incidencia de la preferencia de dígitos según el sexo. En general las mujeres parecen estar más afectadas por este tipo de error en la declaración de la edad, disminuyendo esta mayor preferencia a medida que los índices se hacen menores.

7/ Stockwell, E., y Wicks, H. "Age heaping in recent national censuses" en Revista Social Biology, Vol. 21, N° 2, 1974.

8/ Kamps, Jorge, La Declaración de la Edad en los Censos de Población de la América Latina, CELADE, Serie C, N° 1004, 1976.

Cuadro 4

INDICE DE MYERS PARA LOS CENSOS DE AMERICA LATINA Y ESTADOS UNIDOS Y SUECIA, ALREDEDOR DE LOS AÑOS 1950, 1960 Y 1970

| Países | Hombres | | | Mujeres | | |
|---|---|---|---|---|---|---|
| | 1950 | 1960 | 1970 | 1950 | 1960 | 1970 |
| Argentina | 0,9 | 1,8 | 1,7 | 1,9 | 2,5 | 1,9 |
| Bolivia | 35,88 | ____ | ____ | 47,4 | ____ | ____ |
| Brasil | 13,3 | 18,5 | 9,8 | 15,1 | 17,9 | 10,3 |
| Colombia | ____ | 14,6 | ____ | ____ | 18,7 | ____ |
| Costa Rica | 16,6 | 11,4 | 10,5 | 18,3 | 11,0 | 9,9 |
| Cuba | 9,7 | ____ | ____ | 9,0 | ____ | ____ |
| Chile | 12,9 | 11,0 | 9,3 | 15,8 | 14,6 | 11,3 |
| Ecuador | 31,4 | 26,8 | ____ | 39,5 | 33,1 | ____ |
| El Salvador | 32,0 | 26,1 | 20,4 | 34,1 | 28,8 | 20,9 |
| Guatemala | 25,7 | 15,3 | 19,9 | 40,2 | 24,6 | 24,4 |
| Haití | 45,8 | ____ | 26,2 | 43,0 | ____ | 24,7 |
| Honduras | ____ | 16,7 | 12,2 | ____ | 18,6 | 11,6 |
| México | 27,3 | 23,2 | 15,5 | 35,6 | 28,0 | 18,6 |
| Nicaragua | 33,8 | 29,5 | 28,3 | 35,9 | 29,7 | 27,9 |
| Panamá | 15,4 | 7,4 | 6,0 | 16,4 | 8,5 | 7,1 |
| Paraguay | ____ | 9,0 | 6,8 | ____ | 13,3 | 9,1 |
| Perú | ____ | 13,1 | 9,7 | ____ | 23,7 | 16,4 |
| República Dominicana | 22,2 | 22,8 | 25,2 | 32,5 | 29,0 | 25,9 |
| Uruguay | ____ | 4,0 | ____ | ____ | 5,2 | ____ |
| Venezuela | 21,6 | 6,0 | ____ | 28,4 | 9,2 | ____ |
| Estados Unidos | 2,9 | 1,2 | 1,1 | 4,0 | 1,1 | 1,2 |
| Suecia | 1,3 | 1,1 | 0,6 | 1,4 | 1,1 | 0,6 |

El cuadro 3 pone en evidencia que para el caso del censo de Guatemala de 1973, la preferencia mayor es por el 0, luego el 5 y en menor medida el 8 y el 2, siendo las edades de mayores rechazos las terminadas en 1 y 9. En este sentido, el estudio mencionado permite sacar las siguientes conclusiones:

i) Los dígitos más preferidos en orden decreciente son los siguientes: 0, 5, 8 y 2.

ii) Los dígitos más rechazados en orden decreciente son los siguientes: 1, 9, 7, 3, 4 y 6.

iii) Existen casos en que se dan preferencias marcadas por dígitos distintos del 0 y el 5, que parecen estar relacionados con el año particular en que se realizó el censo. Por ejemplo, en el Paraguay que realizó sus censos en 1962 y en 1972, existe una marcada preferencia por el dígito 2 (supera incluso al 8 y al 5). Esto se debe probablemente al hecho de que muchas personas en lugar de redondear su edad, redondean el año de nacimiento preguntado en el censo, a 0 (1940, 1950, 1960, etc.) y al estimar la edad aparecerá una preferencia sistemática por el 2.

5.3. El Indice de Naciones Unidas[9/]

El cálculo de este indicador de la calidad de la información sobre población según sexo y edad entraña tres etapas, cada una de las cuales brinda interesantes posibilidades de análisis.

Primera Etapa. Cálculo del índice de regularidad por sexo.

Se calculan las relaciones de masculinidad para la población observada y luego se realizan las diferencias sucesivas ($\Delta$ IM) entre un grupo de edades y el siguiente (véase el cuadro 5, columnas 4 y 5). Si la información es correcta, y bajo ciertas condiciones de comportamiento demográfico regular, se espera una evolución gradual de las diferencias y valores absolutos no demasiado grandes. En el caso de Guatemala 1973, llama la atención que para el

9/ Naciones Unidas, "Accuracy Tests for Census Age Distributions Tabulated in Five-year and Ten-year Groups" en Population Bulletin N° 2, Nueva York, 1952.

grupo 15-19, $\Delta$ IM = -10,9, lo que implica una violenta caída de la relación de masculinidad que se mantiene baja hasta la edad 35-39 y aumenta en forma pronunciada en 40-44, conduciendo a una diferencia de 7,0. Este defecto de hombres en edades centrales puede estar vinculado a una corriente emigratoria o a omisión diferencial por sexo.

Dentro de este tramo de edades escapa de la tendencia la diferencia del grupo 30-34 que adopta el valor 5,9 por tener índice de masculinidad más alto que los grupos adyacentes, lo que estaría de acuerdo con lo visto en la Sección 4, sobre la existencia de un déficit mayor del grupo 30-34 frente al 35-39 en el sexo femenino.

La $\Delta$ IM = -7,8 de la edad 65-69 se vincula al alto índice de masculinidad del grupo 60-64, fruto de los aspectos ya analizados de exceso de personas que declaran la edad 60 sobre todo en el caso de hombres.

Una medida resumen de este primer indicador está dada por el promedio de la suma en valores absolutos de las diferencias:

$\frac{\Sigma \mid \Delta IM_x \mid}{K-2}$, siendo K el número de intervalos de edad considerados. Para Guatemala 1973, el índice es 4,1, encontrándose en el mundo variaciones desde cerca del 1,0, en información de muy buena calidad, hasta valores cercanos al 15,0, en datos muy deficientes.

Segunda Etapa. Cálculo de cociente de edades.

El "cociente de edades" se define como el resultado de dividir el número de personas que figuran en un grupo de edades determinado entre la media aritmética de las cifras correspondientes a los dos grupos contiguos:

$$ {}_5CE_x = \frac{2 \cdot {}_5N_x}{{}_5N_{x-5} + {}_5N_{x+5}} $$

Donde: ${}_5CE_x$ = cociente de edades del grupo de edades $x$, $x+4$.
${}_5N_x$ = población con edades entre $x$ y $x+5$.

Es de esperar que ${}_5CE_x$ se desvíe muy poco de 100, salvo perturbaciones reales por cambios en la fecundidad, mortalidad o movimientos migratorios.

El supuesto básico consiste en asimilar la población real a una población en que el número de personas decrece con la edad en una progresión aritmética de primer grado:

$$_5N_x = {}_5N_o + xd$$

$$_5N_{x-5} = {}_5N_o + (x-5)d$$

$$_5N_{x+5} = {}_5N_o + (x+5)d$$

$$_5CE_x = \frac{2 \cdot {}_5N_x}{{}_5N_{x-5} + {}_5N_{x+5}} \; 100 = \frac{2({}_5N_o + xd)}{2({}_5N_o + xd)} \; 100 = 100$$

O sea que si la hipótesis se cumpliera, el índice debería ser exactamente 100. Una forma de medir la calidad de la información a través de los cocientes, es realizar el desvío con respecto a 100, el cual se espera que se aproxime a 0.

Para Guatemala 1973 (véase el cuadro 5), se calculó el índice y su desvío (CE-100) para hombres y mujeres (columnas 7 y 9). A los efectos del análisis se hace el supesto de que los desvíos positivos implican exceso de personas y los negativos déficit. Los desvíos mayores están en las edades avanzadas y como era de suponer, tienen un signo positivo en el grupo 60-64 y negativo en los contiguos. El grupo 35-39, del cual ya se tenían indicios de exceso de población, presenta un desvío de +10,0 en mujeres y +4,0 en hombres, lo que explica además el defecto del grupo 30-34. Este análisis puede proseguirse con las otras edades en forma similar.

Se ha propuesto como medida resumen de estos cocientes la suma en valores absolutos de los desvíos divididos por el número total de desvíos calculados: $\frac{\Sigma|DC|}{K-2}$. Si la información fuera buena, se esperaría un valor de 0. En el caso estudiado se encontró 8,2 para el sexo masculino y 7,4 para el femenino. Los valores más pequeños obtenidos son de alrededor de 1,5 y los mayores, en censos muy malos, cercanos a 30,0, aunque en muy pocas ocasiones.

Cuadro 5

CALCULO DEL INDICE DE NACIONES UNIDAS
PARA EL CENSO DE GUATEMALA DE 1973

| Grupos de edades | Población | | IM | ΔIM | CEH | DCH | CEM | DCM |
|---|---|---|---|---|---|---|---|---|
| | Hombres | Mujeres | | | | | | |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| 0- 4 | 437 180 | 432 880 | 100,99 | | | | | |
| 5- 9 | 399 620 | 385 040 | 103,79 | 2,79 | 101,34 | 1,34 | 101,02 | 1,02 |
| 10-14 | 351 460 | 329 460 | 106,68 | 2,89 | 104,18 | 4,18 | 98,04 | - 1,96 |
| 15-19 | 2751100 | 287 060 | 95,83 | -10,84 | 94,32 | - 5,68 | 100,90 | 0,90 |
| 20-24 | 231 860 | 239 540 | 96,79 | 0,96 | 104,63 | 4,63 | 102,74 | 2,74 |
| 25-29 | 168 100 | 179 260 | 93,77 | - 3,02 | 90,55 | - 9,45 | 94,49 | - 5,51 |
| 30-34 | 139 440 | 139 880 | 99,69 | 5,91 | 92,55 | - 7,45 | 87,69 | -12,31 |
| 35-39 | 133 220 | 139 780 | 95,31 | - 4,38 | 103,97 | 3,97 | 110,02 | 10,02 |
| 40-44 | 116 820 | 114 220 | 102,28 | 6,97 | 103,03 | 3,03 | 99,52 | - 0,48 |
| 45-49 | 93 540 | 89 760 | 104,21 | 1,93 | 97,33 | - 2,67 | 95,87 | - 4,13 |
| 50-54 | 75 400 | 73 040 | 103,23 | - 0,98 | 104,43 | 4,43 | 105,81 | 5,81 |
| 55-59 | 50 860 | 48 300 | 105,30 | 2,07 | 82,12 | -17,88 | 81,74 | -18,26 |
| 60-64 | 48 460 | 45 140 | 107,35 | 2,05 | 122,25 | 22,25 | 117,49 | 17,49 |
| 65-69 | 28 420 | 28 540 | 99,58 | - 7,78 | 80,56 | -19,44 | 84,94 | -15,06 |
| 70-74 | 22 100 | 22 060 | | | | | | |
| Total | | | | 52,58 | | 106,41 | | 95,68 |
| Promedio | | | | 4,04 | | 8,19 | | 7,36 |

INU = 3 x 4,04 + 8,19 + 7,36 = 27,68
=====

IM = Indice de masculinidad.

ΔIM = Diferencias sucesivas de IM.

CEH, CEM = Cocientes de edades de Hombres y Mujeres.

DCH, DCM = Desviación de CEH y CEM respecto a 100.

INU = Indice resumen de Naciones Unidas.

Tercera Etapa. Cálculo de un índice resumen (INU).

En las etapas precedentes, se tenían indicadores de la regularidad por sexo y edad. La Secretaría de las Naciones Unidas ha propuesto calcular un índice resumen de la siguiente manera:

$$INU = 3\,\frac{\Sigma|\Delta\ IM|}{K-2} + \frac{\Sigma|DC^{H}|}{K-2} + \frac{\Sigma|DC^{M}|}{K-2}$$

Se adoptó esta medida sintética por experiencia de un gran número de cálculos. Se pondera más el índice de regularidad por sexo argumentando que esta información es más confiable y refleja mejor los errores de los datos, ya que los índices de regularidad por edad se encuentran más afectados por perturbaciones reales, por cambios en la fecundidad o movimientos migratorios. La inclusión de los cocientes de edades, a pesar de lo dicho antes, ha sido justificado porque suele ocurrir que las diferencias de índices de masculinidad a veces ocultan errores porque la mala declaración tiene el mismo sentido en hombres y mujeres produciéndose compensaciones.

Una de las ventajas de este índice frente a los anteriores, es el hecho de que se trabaja con grupos quinquenales o decenales, ya que a veces la información no está dada por edades simples. Además, no sólo refleja errores por preferencia de dígitos, sino que todo tipo de irregularidades en los datos por sexo y edad. Esto que por un lado es una ventaja, por otra parte constituye una desventaja, pues la interpretación del INU se torna confusa y además sufre la influencia de perturbaciones reales que no se deben a defectos de los datos. La aplicación de este índice ha perdido importancia porque los censos más recientes han sido tabulados por edades detalladas permitiendo el cálculo de los índices de Whipple y Myers que son de más clara interpretación.

Las variaciones de este índice se interpretan de acuerdo a la experiencia, dado que no existen límites precisos. En ese sentido se ha propuesto la siguiente clasificación:[10/]

---

10/ Kamps, J.E., La Declaración de la Edad ... op. cit. pág. 7.

| | |
|---|---|
| Más de 40: | información deficiente |
| De 20 a 40: | información de calidad intermedia |
| Menos de 20: | información satisfactoria |

De acuerdo a esta clasificación y al cuadro 6 puede decirse que la gran mayoría de los países de América Latina tienen una información de calidad intermedia que ha mejorado, en general, a través de los tres últimos censos. De cinco países con mala información en 1950 sólo uno aparece en 1970. La mejoría sustancial se dio entre los censos de 1950 y 1960, siendo menos la de la última década.

La información suministrada en el cuadro correspondiente pone en evidencia, además, lo ya dicho sobre las debilidades del índice, en cuanto a que su valor puede verse afectado por perturbaciones reales, lo cual explicaría los índices superiores de Suecia y los Estados Unidos con respecto a algunos países de América Latina.

5.4. Estudio de errores a partir de una población estándar

a) Uso de ajuste de la población por edades

Se trata de estimar una población, supuestamente correcta, a través de algunos de los métodos de ajuste que se verán en la Sección VII, o similares, y compararlos con la población enumerada en el censo. Por ejemplo, si se quiere tener una idea de la preferencia por una edad $x$ determinada, el índice sería:

$$\frac{N_x}{N'_x} \; 100$$

siendo $N_x$ la población censada y $N'_x$ la estimada por ajuste. A su vez, para estudiar la preferencia de las edades terminadas en 0 y en 5 entre 23 y 62 años se operaría así:

$$\frac{N_{25}+N_{30}+N_{35}+N_{40}+N_{45}+N_{50}+N_{55}+N_{60}}{N'_{25}+N'_{30}+N'_{35}+N'_{40}+N'_{45}+N'_{50}+N'_{55}+N'_{60}} \; 100$$

Si la información fuera correcta, se esperaría un índice igual a 100.

Cuadro 6

PAISES DE AMERICA LATINA Y ESTADOS UNIDOS Y SUECIA: INDICE DE NACIONES UNIDAS

| Países | Censo alrededor de | | |
|---|---|---|---|
| | 1950 | 1960 | 1970 |
| Argentina | 17,7 | 12,26 | 10,29 |
| Bolivia | 46,2 | ——— | 32,21 |
| Brasil | 23,4 | 19,32 | 11,41 |
| Colombia | 42,9 | ——— | 30,57 |
| Costa Rica | 28,6 | 24,52 | 20,66 |
| Cuba | 34,1 | ——— | 22,35 |
| Chile | 31,0 | 15,80 | 13,74 |
| Ecuador | 39,3 | 34,66 | 26,38 |
| El Salvador | 34,6 | 33,89 | 25,72 |
| Guatemala | 42,3 | 27,77 | 27,68 |
| Haití | 53,2 | ——— | 37,59 |
| Honduras | 20,2 | 24,42 | 27,07 |
| México | 35,3 | 25,66 | 18,65 |
| Nicaragua | 39,0 | 38,50 | 31,13 |
| Panamá | 30,7 | 17,02 | 14,62 |
| Paraguay | 23,2 | 26,47 | 26,71 |
| Perú | ——— | 24,73 | 23,32 |
| República Dominicana | 66,0 | 50,96 | 41,16 |
| Uruguay | ——— | 16,30 | 16,97 |
| Venezuela | 37,1 | 19,23 | 19,51 |
| Estados Unidos | 11,0 | 12,59 | 13,67 |
| Suecia | 11,9 | 13,26 | 14,42 |

Los supuestos de estos índices dependen de los supuestos que implica el ajuste adoptado. Constituye una seria limitación el hecho de que cada ajuste conlleva resultados distintos, lo que implica un alto grado de arbitrariedad.

b) Uso de una población modelo estándar

En el Manual IV de las Naciones Unidas[11/] se presenta un método que permite determinar los errores en la declaración de la edad. El método consiste en estudiar los desvíos que se dan entre la población censada con respecto a una población estable adecuada, a través de las siguientes relaciones:

$$\frac{{}_5c_x}{{}_5c_x^s} \quad y \quad C(x) - C^s(x)$$

donde:

${}_5c_x$ es la proporción de personas por grupos quinquenales de edades de la población censada.

${}_5c_x^s$ la proporción de personas por grupos quinquenales de edades de la población estable elegida.

$C(x)$ distribución por edad acumulada de la población censal.

$C^s(x)$ distribución por edad acumulada de la población estable elegida.

El procedimiento para elegir la población estable se presenta en la sección en que se trata de ajustes mediante modelos.

Si la población censal correspondiera exactamente a la estable, se esperarían los siguientes resultados:

$$\frac{{}_5c_x}{{}_5c_x^s} = 1 \quad y \quad C(x) - C^s(x) = 0$$

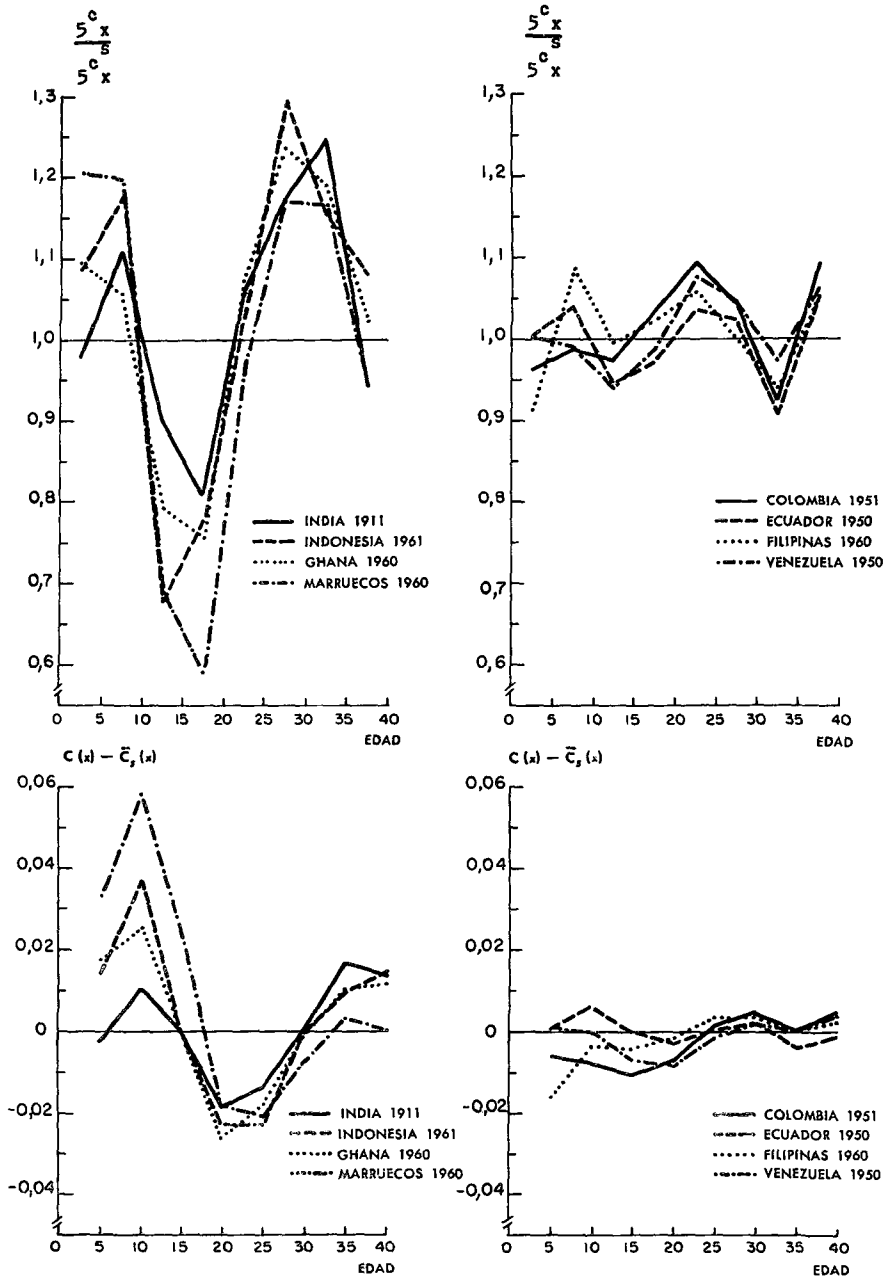
11/ Naciones Unidas, Métodos para Establecer Mediciones Demográficas Fundamentales a Partir de Datos Incompletos, Manual IV, ST/SOA/Serie A/42, Nueva York, 1968, pág. 18.

Cuando $\frac{{}_5C_x}{{}_5C_x^s}$ es mayor que 1 se interpreta que los errores tienden a exagerar ese grupo. En el caso de ser menor que 1 se estaría frente a la situación contraria. Si $C(x) - C^s(x)$ es mayor que 0, habría un traslado hacia edades más jóvenes, por debajo de x; en cambio, un valor negativo implica un traslado hacia edades mayores por encima de la edad x. Se propone usar esta metodología con la intención de descubrir patrones generales de error. Con ese fin se realizó la comparación basándose en más de 150 censos o encuestas para cada sexo, en los que se cumple aproximadamente la condición de estabilidad o cuasi-estabilidad. Cabe, sin embargo, alertar sobre la posibilidad de que parte de las desviaciones se deben al apartamiento de la población investigada de los supuestos de estabilidad.

En el gráfico 3 se presentan algunos casos de comportamiento de esta función del error que permiten sacar algunas conclusiones generales. Se distinguen claramente dos modalidades de error: una perteneciente a los países africanos y asiáticos y otra a los países latinoamericanos. En el caso de los primeros, los desvíos $\frac{{}_5C_x}{{}_5C_x^s}$ con respecto a 1 son mucho más pronunciados, estando en ciertos tramos de edades sistemáticamente por abajo o sistemáticamente por arriba. En cambio, para los países de América Latina, existe un comportamiento del tipo "diente de sierra". Esto implica que, en el primer grupo de países, a la izquierda del gráfico, se observe una declaración sistemática y altamente sesgada, perturbando también la ojiva de la población por edades. En los países latinoamericanos aun cuando se notan errores importantes, éstos afectan poco a la población acumulada por no ser sistemáticos. Este hecho constituye un argumento importante para usar, en lo posible, las ojivas de población como indicadores de la estructura por edades de las poblaciones de la región.

Gráfico 3

Comparaciones de la distribución de la población femenina por edad, por grupos de edad de cinco años -$_5c_x$- y su ojiva -C(x)-, como se registró en diversos censos, con los valores correspondientes en poblaciones estables medianas entre los definidos por un $\mathring{e}_o$ de 40 años y por la concordancia con C(5), C(10), ..., C(40), en las poblaciones del censo. Las comparaciones ilustran dos modalidades típicas del fenómeno de las declaraciones falsas de la edad: la modalidad "africana-asiática meridional" (diagramas de la izquierda) y la modalidad "latinoamericana" (diagramas de la derecha)




Fuente: Tomado de: Naciones Unidas, Métodos para establecer mediciones ... op. cit.