

MUESTREO

ÍNDICE

3.0. Conceptos generales.....	4
3.1. Cálculo de tamaños de muestra	4
3.1.0. Introducción.....	4
3.1.1. Intervalos de confianza	6
3.1.1.1. Media.....	7
3.1.1.2. Proporción	9
3.1.1.3. Odds ratio	10
3.1.1.4. Riesgo relativo.....	12
3.1.1.5. Concordancia.....	13
3.1.1.6. Pruebas diagnósticas	14
3.1.2. Contrastes de hipótesis.....	17
3.1.2.1. Comparación de medias	18
3.1.2.2. Comparación de proporciones.....	21
3.1.2.3. Estudios de casos y controles.....	23
3.1.2.4. Estudios de cohorte	25
3.1.2.5. Estudios de equivalencia	26
3.1.2.6. Pruebas diagnósticas	29
3.1.2.7. Calidad de lotes.....	31
3.1.2.8. Supervivencia	33
3.1.2.9. Coeficiente de correlación	34
3.2. Selección muestral	35
3.2.0. Conceptos generales	35
3.2.1. Muestreo simple aleatorio (MSA).....	37
3.2.2. Muestreo sistemático (MS).....	39
3.2.3. Muestreo aleatorio estratificado (MAE).....	42
3.2.4. Muestreo por conglomerados monoetápico.....	44
3.2.5. Muestreo por conglomerados bietápico	46
3.2.6. Muestreo por conglomerados monoetápico estratificado.....	49
3.2.7. Muestreo por conglomerados bietápico estratificado.....	53
3.3. Asignación de sujetos a tratamientos	57
3.3.0. Conceptos generales	57
3.3.1. El proceso de asignación aleatoria.....	57
3.4. Estimación con muestras complejas	59
3.4.0. Introducción.....	59

3.4.1. Probabilidad de selección y ponderación.....	59
3.4.2. Diseños muestrales abarcados.....	60
3.4.3. Efecto de diseño.....	61
3.4.4. Tipo de variables	61
3.4.5. Trabajo con clases o subpoblaciones	61
3.4.6. Manejo del submódulo.....	62
Bibliografía	68
Anexo 1: Novedades del módulo de muestreo en la versión 4.....	70
Anexo 2: Fórmulas del módulo de muestreo	71

3.0. Conceptos generales

Virtualmente toda investigación epidemiológica exige la selección de una muestra de la población que se estudia. Necesariamente, entonces, habrá que tomar una decisión sobre:

- El tamaño que ha de tener la muestra.
- El procedimiento para seleccionar ese número de *unidades de análisis* (elementos o sujetos que componen la población y que podrían ser, en su momento, objeto de interrogatorio o medición).
- La realización de estimaciones puntuales con sus respectivos márgenes de incertidumbre (si el problema que se encara es “de estimación”; es decir, si es de naturaleza básicamente descriptivo-inferencial).

El presente módulo ofrece herramientas para las tres tareas.

El problema de determinar el tamaño de muestra necesario ha sido intensamente abordado desde el punto de vista informático; sin embargo, Epidat 4.0 incorpora un compendio bastante exhaustivo de soluciones para las diversas demandas potenciales de los epidemiólogos.

Análogamente, las alternativas disponibles para llevar adelante el proceso de selección son numerosas. Los recursos computacionales disponibles -especialmente, aquellos que aparecen en los grandes paquetes clásicos- para llevar adelante la selección de muestras son, sin embargo, bastante limitados. Epidat 4.0 aporta un instrumento apropiado para completar esta tarea, con un alto nivel de ductilidad para seleccionar los elementos de la población objeto de estudio que habrán de integrar la muestra, ocasionalmente realizando subselecciones de bases de datos previamente configuradas.

Como se ha dicho, Epidat 4.0 también incorpora procedimientos para realizar estimaciones de parámetros con acuerdo a los diseños muestrales.

Finalmente, cuando se realiza una investigación experimental, con frecuencia es necesario aplicar asignaciones aleatorias de los sujetos del estudio a dos o más grupos (decidir qué sujetos han de recibir cada uno de los tratamientos contemplados). El cuarto submódulo de este módulo de muestreo resuelve dicha tarea.

3.1. Cálculo de tamaños de muestra

3.1.0. Introducción

La determinación del tamaño de muestra es una tarea que se realiza durante la fase de diseño de un estudio para establecer el número de elementos o sujetos que deben seleccionarse. El abanico de criterios susceptibles de ser considerados para efectuar este cálculo es amplio y variado, y la elección concreta que se haga depende, entre otras cosas, de la información de que se disponga y de las características del procedimiento estadístico que se va a utilizar para analizar los datos.

Por ejemplo, en una encuesta transversal para estimar la prevalencia de diabetes en una población, se requerirá un tamaño de muestra que permita estimar dicha prevalencia con una precisión mínima. En cambio, en un ensayo clínico de un nuevo fármaco donde, pongamos por caso, se compara dicho medicamento con un placebo o con otro fármaco, el tamaño de muestra debe permitir declarar como *significativa* determinada diferencia entre

respectivos porcentajes de recuperación, en caso de que tal diferencia o una mayor sea empíricamente observada. En el primer caso, el objetivo es la estimación de un parámetro; en el segundo se pretende analizar si un tratamiento dado difiere de manera estadísticamente significativa de otro. En cada uno de estos dos contextos, estimación o realización de pruebas de hipótesis, se abren a su vez diversas posibilidades. Epidat 4.0 procura dar elementos que ayuden a resolver el problema para las situaciones que se presentan más frecuentemente en la práctica para ambos casos.

Debe enfatizarse, sin embargo, que los recursos disponibles para hacer esta determinación sólo tienen un carácter orientador. Los resultados formales como los que Epidat 4.0 produce, necesariamente han de ser combinados con otros elementos cruciales a la hora de decidir un tamaño muestral, tales como los recursos materiales disponibles y la “historicidad” del problema; una posible alternativa consiste en conjugar los resultados formales que ofrece Epidat 4.0 con el *sentido común* y tener en cuenta los tamaños usados en trabajos similares [1][2]. Aun cuando se elija aplicar recursos computacionales formales como los que aquí se ofrecen, estos elementos “laterales” a las fórmulas siempre habrán de desempeñar un papel en el proceso decisorio.

Cabe, en este punto, realizar las siguientes advertencias y recomendaciones generales:

- La determinación a priori de los valores que intervienen en las fórmulas del tamaño muestral es una decisión esencialmente subjetiva. En concreto, por ejemplo, no es fácil establecer cuál es el error máximo aceptable para la estimación de un parámetro. Esto puede conducir a que dos investigadores independientes, aplicando la misma fórmula, obtengan tamaños de muestra muy diferentes por el hecho de que tienen visiones distintas (defendibles ambas) del problema.
- En el cálculo de un tamaño de muestra tener que conocer en la fase de diseño el valor de un parámetro que solo se podrá estimar una vez extraída la muestra constituye una contradicción habitual. Para resolver este problema hay dos opciones:
 - Estimar el parámetro a partir de una muestra piloto.
 - Obtener un valor aceptable a partir de las referencias que se tengan de trabajos o experiencias similares ya realizadas.
- En la práctica, las encuestas están generalmente diseñadas para estimar un conjunto grande de parámetros, y no uno solo. Sin embargo, es necesario determinar un tamaño de muestra único. Esta dificultad suele solventarse adoptando uno de los parámetros como “el más importante”, lo que implica otra decisión subjetiva en el cálculo.
- El error muestral en que se incurre al estimar un parámetro depende del tipo de diseño establecido para seleccionar los elementos que integran la muestra. Por lo general, a tamaños de muestra iguales, el error es mayor si se utiliza un diseño complejo (por ejemplo, muestro estratificado bietápico) que bajo muestreo simple aleatorio (MSA). Sin embargo, las fórmulas que permiten obtener el tamaño de muestra en función de la precisión siempre asumen que se va a realizar un MSA, lo que conduce a tamaños de muestra más pequeños que los que realmente serían necesarios para garantizar el grado de precisión exigido si el diseño es otro. Una forma de “corregir” el tamaño de muestra en función del diseño muestral es multiplicar el tamaño obtenido con MSA por el denominado *efecto de diseño*. Naturalmente, no es fácil hallar una estimación adecuada de este valor, pero en la práctica suele asumirse algún valor entre 1,5 y 3. Un valor igual a 2, por ejemplo, significa que para obtener la misma precisión habrá que estudiar el doble de individuos que con el MSA.

Una abarcadora reflexión sobre el cálculo del tamaño de muestra puede verse en el capítulo 12 del libro “Diseño razonado de muestras y captación de datos para la investigación sanitaria” [3].

3.1.1. Intervalos de confianza

Como parte de una encuesta epidemiológica es habitual que se quiera estimar la prevalencia de una enfermedad en una población. En esas situaciones, el investigador se pregunta “¿cuántos sujetos necesito examinar para estimar esa prevalencia con un grado razonable de precisión?”. Los submódulos de Epidat 4.0 incluidos en el módulo de *Muestreo, Cálculo de tamaños de muestra*, bajo el epígrafe “Intervalos de confianza” basan el cálculo del tamaño de muestra necesario para un estudio en la amplitud máxima admisible para el intervalo de confianza del parámetro que se pretende estimar. Esos submódulos son:

- 3.1.1.1. Media
- 3.1.1.2. Proporción
- 3.1.1.3. Odds ratio
- 3.1.1.4. Riesgo relativo
- 3.1.1.5. Concordancia
- 3.1.1.6. Pruebas diagnósticas

Un intervalo de confianza (IC) es un recorrido de valores dentro del cual estaría el parámetro de interés con cierto grado de confianza, que habitualmente es del 95%, y debe presentarse siempre junto a la estimación puntual de un parámetro, puesto que permite cuantificar la magnitud del error asociado a la estimación, o error muestral. La amplitud del IC está directamente relacionada con ese error que, en el caso de una media o una proporción, por ejemplo, es exactamente la mitad de la amplitud del IC. A su vez, el error depende del tamaño de la muestra, por lo que el tamaño muestral mínimo estará en función del error máximo que se considere admisible. Si se procede a estimar, por ejemplo, una proporción, exigir que el error de esa estimación sea suficientemente pequeño como para considerar que la estimación es adecuada, equivale a demandar que el intervalo de confianza sea suficientemente estrecho.

En un contexto de estimación, por tanto, el tamaño muestral está determinado por varios factores:

- La *precisión de la estimación*, que se mide a través del error muestral. La magnitud del máximo margen de error admisible debe ser fijada por el investigador, pues es él quien mejor identifica las consecuencias prácticas de un error dado. Si, por ejemplo, lo que se quiere es conocer el salario medio de una población de mujeres en edad fértil y se declara que se admite un error de hasta $e=50€$, ello significa que se está pensando en términos como los siguientes: “Si el verdadero promedio fuera 500€ pero lo que se me informa es que asciende a 525€, entonces considero que conozco adecuadamente el dato; sin embargo, si lo que se me informa es que ese número es, por ejemplo, 418€ y luego me entero de la cifra verdadera (500€), considero que estaba trabajando con un dato equivocado”. En general se cumple que, al aumentar la precisión, es decir, al disminuir el grado de error admisible, aumenta el tamaño de muestra necesario, algo coherente con lo que indica el sentido común. Cuando el error se expresa como un porcentaje del valor que puede tener el parámetro, por ejemplo un 5% o un 10%, en ese caso se habla de precisión relativa.

- El *nivel de confianza*, que generalmente se fija en 95%. El efecto de este valor sobre el tamaño de muestra es directo; es decir, para una precisión dada, el tamaño de muestra aumenta al aumentar el nivel de confianza.
- Otros elementos directamente relacionados con el parámetro concreto que se pretende estimar y que se comentarán en cada caso.

Epidat 4.0 permite calcular tamaños de muestra de una forma flexible: genera una tabla de valores para la precisión y proporciona el tamaño de muestra correspondiente a cada uno de esos valores. La tabla se define partiendo de un valor *mínimo* fijado por el usuario, que va aumentando a intervalos regulares de *incremento* hasta llegar al valor *máximo* también declarados por el usuario. Por ejemplo, si se definen *mínimo*=3, *máximo*=15 e *incremento*=4, se obtiene una tabla con los valores 3, 7, 11 y 15 para la precisión y los correspondientes tamaños de muestra.

Un recurso incluido en Epidat 4.0 y que puede ser de utilidad en muchas situaciones es el de realizar el cálculo inverso; es decir, determinar la precisión asociada a la estimación de un parámetro que correspondería a diversos tamaños de muestra posibles. Igual que en la situación inversa, el cálculo es flexible; se puede definir una tabla de valores para los tamaños de muestra a partir de un *mínimo*, un *máximo* y un *incremento*, y se obtendrán las precisiones que corresponden a todos los casos.

3.1.1.1. Media

Los factores específicos que intervienen en el cálculo del tamaño de muestra (o precisión) para estimar una media son los que se mencionan y comentan a continuación:

- *Tamaño de la población:*

En la mayoría de las situaciones prácticas es difícil conocer el verdadero tamaño de la población objeto del muestreo. A menudo se dispone solo de una aproximación para tal valor, pero puede darse el caso de que se desconozca por completo. Ninguna de estas situaciones pone al investigador ante una restricción importante a la hora de calcular el tamaño de muestra, ya que el tamaño poblacional tiene escasa influencia sobre el tamaño de la muestra la mayor parte de las veces [2].

Supóngase, por ejemplo, que se planifican dos estudios de fecundidad: uno en una gran ciudad donde hay medio millón de mujeres en edad fértil, y otro en uno de sus municipios que tiene 55.000 mujeres entre 15 y 49 años. En ambos casos se quiere estimar, mediante MSA, el salario medio de dichas mujeres con un error no mayor de 50€ y admítase que, por estudios previos, se conoce que la desviación estándar de todos los salarios en la población es aproximadamente igual a 300€. Los tamaños de muestra que proporciona Epidat 4.0, para un nivel de confianza del 95%, son respectivamente, 139 mujeres en la ciudad y 138 en el municipio, números virtualmente iguales. Este hecho, válido siempre que la población sea mucho mayor que los posibles tamaños muestrales, pone de manifiesto que es errónea la recomendación, nada inusual, de fijar un porcentaje de la población como tamaño de muestra. En el ejemplo expuesto, seleccionar un 1% de la población conduciría a una muestra de 5.000 mujeres en la ciudad y de 550 en el municipio; puede comprobarse, haciendo el cálculo en Epidat, que el margen de error en el primer caso sería de unos 8,3€, valor que se triplica en el municipio, para el cual ascendería a 23,8€.

Por otra parte, si el tamaño poblacional se desconoce, pero puede asumirse que es mucho mayor que el de la muestra que se va a seleccionar, entonces se está en una situación

equivalente a la que se produciría ante el caso hipotético de una población infinita. En ese caso, en la pantalla de entrada de datos de Epidat 4.0 se puede dejar en blanco el campo correspondiente al tamaño de la población.

- *Desviación estándar de la variable de interés:*

Naturalmente, este número no se conoce con certeza virtualmente nunca. Como ya se había señalado, en tal caso, se emplea o bien el resultado surgido de una muestra piloto o bien un valor tomado de las referencias surgidas de trabajos o experiencias similares ya realizados.

Ejemplo

En un área sanitaria, la distribución del peso al nacer de niños cuyas madres cumplen su período de gestación de 40 semanas es aproximadamente normal con una media de $\mu=3.500$ gramos y una desviación estándar de $\sigma=430$ gramos. Un investigador planea llevar a cabo un estudio con una muestra simple aleatoria de los próximos partos, a lo largo del año entrante, en un hospital de maternidad. Supongamos que el número esperado de partos fuera $N=3.450$. Se quieren, como es usual, estimar diversos parámetros (algunos de ellos en subconjuntos tales como el de las madres que fumaron durante el embarazo, o el de las mayores de 38 años), pero a los efectos de fijar el tamaño muestral, consideremos que se quiere estimar el peso medio al nacer de los niños que llegan al término del embarazo asumiendo que la desviación estándar es la mencionada (430 gramos). Si el investigador desea que el error no supere los 50 gramos con una confianza del 95%, el tamaño de muestra requerido ascendería a 263 sujetos.

Resultados con Epidat 4.0:

Datos:	
Tamaño de la población:	3.450
Desviación est. esperada:	430,000
Nivel de confianza:	95,0%
Efecto de diseño:	1,0
Resultados:	
Precisión	Tamaño de la muestra
50,000	263

Nota: Adviértase que este resultado se obtiene dejando el valor por defecto para el "Efecto de diseño" (igual a 1,0) por tratarse de una muestra que se supone que será simple aleatoria.

3.1.1.2. Proporción

En la práctica, el caso más común, con diferencia, es el cálculo del tamaño de muestra (o precisión) para estimar una proporción. Los factores específicos que intervienen en él son:

- *Tamaño de la población:*

Sobre este dato, cabe hacer las mismas consideraciones que en el caso de la media, anteriormente comentadas.

- *Proporción esperada:*

Puesto que la demanda de proveer al programa de un valor anticipado para P es irremediablemente contradictoria (pues el tamaño que se busca es el de la futura muestra, que se estudiará precisamente con la finalidad de estimar P) es muy habitual encontrar en manuales y artículos [4][5] la recomendación de utilizar el valor 50 cuando no se conoce en absoluto el valor del porcentaje P . Para ello se esgrime el argumento de que dicho valor es el que produce el máximo tamaño de muestra, de modo que en tal caso no tendríamos el tamaño óptimo necesario sino una cota superior (nunca necesitaríamos uno mayor que el que se obtiene bajo ese supuesto). Cabe hacer algunas consideraciones con respecto a esta "receta", que ha sido objeto de discusión en la literatura [6][7][8][9]. En primer lugar, es cierto que la fórmula que permite calcular el tamaño de muestra para estimar una proporción P alcanza su máximo cuando $P=50$; pero esto solo es cierto si los demás elementos que intervienen en la fórmula, nivel de confianza y precisión del IC, están fijos. Sin embargo, el máximo error absoluto que se puede admitir para estimar una proporción P no puede establecerse de manera racional sin tener una idea acerca de la magnitud de P . Por ejemplo, un error absoluto del 3% puede ser adecuado si la proporción que se va a estimar está alrededor del 20%, pero resultaría del todo inapropiado si el rasgo cuya prevalencia se quiere estimar es muy poco frecuente, dígase un 2%, menor incluso que el error establecido.

Consecuentemente, tal receta resulta carente de sentido y no debe ser empleada. Para una fundamentación más detallada, acúdase a Silva [7]. Cabe señalar, sin embargo, que el investigador raramente se halla en esa situación de absoluto desconocimiento acerca del parámetro; basándose en su experiencia o el conocimiento científico sobre el tema, generalmente tiene de antemano una idea aproximada del recorrido de valores en el que razonablemente se podría ubicar. Repárese en el siguiente ejemplo tomado de la ayuda del módulo de "Análisis bayesiano" de Epidat 3.1:

"Es razonable suponer que de antemano se tengan ciertas ideas acerca de cuál puede ser ese valor desconocido; por ejemplo, podría considerarse altamente verosímil que la tasa de prevalencia de asma en la comunidad no esté muy distante de 9%. Desde luego, tal criterio inicial, nacido quizás de la experiencia propia y ajena, así como del encuadre teórico del problema, se produce en un marco de incertidumbre. Así, tal vez se piensa que 7% ó 13% son también valores posibles, aunque acaso menos verosímiles que 9%. Y también se pudiera estar convencidos de que 55% sería sumamente improbable, a la vez que se pudiera estar persuadidos de que cifras como 1% u 80% serían virtualmente imposibles. Tener de antemano puntos de vista como estos no solo es posible sino casi inevitable para alguien versado en la materia."

En una situación de este tipo, la determinación de la precisión adecuada no es una maniobra "ciega". Sin perder de vista la carga subjetiva que supone siempre tal decisión, por lo general es posible fijar un error que no sea disparatado teniendo en cuenta el recorrido de valores que se puede considerar razonable para P . En el ejemplo de la tasa de prevalencia, se puede suponer que la proporción que ha de estimarse se moverá probablemente entre un 5% y un 15%, y es razonable aspirar a que no se incurra en un

error mayor del 1%. Para el cálculo del tamaño de muestra, el valor de P que maximiza n es el más próximo al 50%, en este caso se tomaría P=15%.

Ejemplo (Adaptado de Lwanga y Lemeshow [5])

Un departamento local de salud quiere calcular la prevalencia de sujetos con reacciones positivas a la tuberculina entre los 450 menores de 5 años registrados (existentes) en su circunscripción. Se desea calcular cuántos de tales niños han de integrar una muestra simple aleatoria para que pueda estimarse dicha prevalencia con una precisión absoluta, expresada en porcentaje, entre el 3 y el 5% y un 95% de confianza si se cree, por estudios previos, que la tasa de prevalencia pudiera ser, aproximadamente, del 20%.

Epidat 4.0 arroja los tamaños de 272 niños, 208 y 159 para obtener precisiones de 3, 4 y 5% respectivamente.

Adviértase que se ha dejado el valor por defecto para el “Efecto de diseño” (igual a 1,0) por tratarse de una muestra simple aleatoria.

3.1.1.3. Odds ratio

Un estudio de casos y controles es un estudio observacional que comienza con la identificación de un grupo de casos (individuos con una enfermedad o condición) y un grupo de controles (individuos sin la enfermedad o condición). Para cada elemento del estudio se identifica si estuvo o no expuesto a cierto factor; entonces, se estima la prevalencia de la exposición en cada uno de los dos grupos y se comparan estas dos estimaciones.

Como señala Thompson [10], en los últimos años se ha producido un importante cambio de dirección en el análisis y presentación de los datos de estudios de casos y controles y otros tipos de investigaciones epidemiológicas. Cada vez se pone menos énfasis en las pruebas de significación estadística y los correspondientes valores p, y más en las estimaciones puntuales e intervalos de confianza de medidas cuantitativas del grado de asociación entre exposición y enfermedad [11]. Cuando el análisis previsto se va a orientar en tal dirección, el cálculo del tamaño de muestra necesario debe basarse en la estimación y no en la prueba de hipótesis.

La medida más frecuentemente utilizada para cuantificar la asociación a partir de datos procedentes de un estudio de casos y controles es la odds ratio (OR). Los factores específicos que intervienen en el cálculo del tamaño de muestra (o precisión) para estimar una OR en un estudio de este tipo son:

- *Proporción de casos expuestos o prevalencia de la exposición en los casos (P₁).*
- *Proporción de controles expuestos o prevalencia de la exposición en los controles (P₂).*
- *Odds ratio (OR).*

Estos tres valores están relacionados entre sí del modo siguiente:

$$OR = \frac{P_1(1-P_2)}{P_2(1-P_1)}$$

De modo que, especificando dos de estos parámetros, el otro queda automáticamente determinado. Epidat 4.0 permite que el usuario decida cuál pareja de valores desea introducir, y calcula automáticamente el tercero.

▪ *Número de controles por caso:*

La determinación del número de controles es una cuestión muy importante al planificar un estudio de casos y controles. Generalmente se utiliza igual número de controles que de casos. Se ha demostrado que hay muy poca ganancia en la precisión de un estudio cuando la razón (número de controles)/(número de casos) supera el valor 4 [12]. No obstante, Epidat 4.0 admite que se establezcan hasta 10 controles por caso.

Si se selecciona la opción de calcular tamaños de muestra, la precisión del intervalo de confianza o error se pide en términos relativos, es decir, la razón entre el error absoluto y la OR, y el resultado que produce Epidat 4.0 es el tamaño de cada uno de los grupos (casos y controles), por separado. Para fijar los errores relativos, se ha sugerido que el investigador fije un número que pudiera moverse entre 5% y 15%, ya que uno menor que 5% suele ser exageradamente exigente y uno mayor que 15% sería demasiado concesivo.

En el caso inverso, cuando se pide calcular precisión, el dato de entrada debe ser el tamaño total de la muestra.

Ejemplo

Se va a diseñar un estudio de casos y controles emparejado sobre tabaquismo y cáncer de laringe utilizando controles poblacionales. Suponiendo que la prevalencia del hábito tabáquico en la población es del 45% y que la OR esperada es aproximadamente 3, con 1.295 casos y otros tantos controles tomados aleatoriamente se estimará la OR con una precisión relativa del 15% y un nivel de confianza del 95%. Si se exigiera un error relativo de a lo sumo 8%, la muestra se elevaría a 4.919 pares.

Resultados con Epidat 4.0:

Datos:		
Proporción de casos expuestos:	71,053%	
Proporción de controles expuestos:	45,000%	
Odds ratio esperada:	3,000	
Número de controles por caso:	1	
Nivel de confianza:	95,0%	
Resultados:		
Precisión relativa (%)	Tamaño de la muestra	
	Casos	Controles
8,000	4.919	4.919
15,000	1.295	1.295

Nota: Adviértase que Epidat 4.0 calcula la proporción de casos expuestos (71,053%) empleando la OR y la proporción de controles expuestos que se han declarado.

3.1.1.4. Riesgo relativo

Un estudio de cohorte es un estudio observacional que típicamente comienza con la identificación de un grupo de personas libres de una enfermedad o que no han desembocado en determinada condición en estudio. En este grupo de personas se separan las que están sujetas a cierta exposición y las que no están en ese caso; luego se realiza un *seguimiento* en el tiempo de unos y otros con el fin de comparar la frecuencia de aparición de la enfermedad entre expuestos y no expuestos.

Un aspecto fundamental del análisis de los estudios de cohorte es su capacidad para producir medidas de incidencia de la enfermedad, es decir, del riesgo individual de desarrollar la enfermedad durante el período de seguimiento. La razón entre el riesgo en expuestos y en no expuestos, o riesgo relativo, expresa cuántas veces más probable es que ocurra la enfermedad en un grupo que en el otro, y a la vez constituye una medida que cuantifica la asociación entre la exposición y la enfermedad.

Los factores específicos que intervienen en el cálculo del tamaño de muestra (o precisión) para estimar un riesgo relativo son:

- *Riesgo en expuestos* (P_1).
- *Riesgo en no expuestos* (P_2).
- *Riesgo relativo esperado* (RR).

Estos tres valores están relacionados entre sí del modo siguiente:

$$RR = \frac{P_1}{P_2},$$

De modo que, especificando dos de estos parámetros, el otro queda determinado; consecuentemente, Epidat 4.0 permite que el usuario decida cuáles de los valores desea introducir y calcula automáticamente el tercero.

- *Razón entre el número de no expuestos y el de expuestos* (R):

Cuando se plantea el diseño de un estudio de cohortes, generalmente no se conoce de antemano la magnitud de la exposición en los sujetos de la población que se estudia, sino que la muestra se selecciona aleatoriamente de la población (o se asume que así ha sido) y posteriormente se determinan los individuos expuestos entre los que conforman la cohorte. En esa situación, la proporción de expuestos en la cohorte será aproximadamente igual a la prevalencia de la exposición en la población de la que procede la muestra. A partir de esta prevalencia (P) puede estimarse la razón entre el número de no expuestos y expuestos que conformarán la cohorte: $R=(1-P)/P$, valor que interviene en el cálculo del mínimo tamaño de muestra para cumplir los requisitos de precisión exigidos [5]. También puede darse la situación de que la exposición se conozca previamente a la selección de los sujetos y que se desee seleccionar, por ejemplo, el mismo número de expuestos que de no expuestos. En ese caso, la razón tomará el valor 1.

Si se selecciona la opción de calcular tamaños de muestra, la precisión del intervalo de confianza o error ha de comunicarse en términos relativos, es decir, hay que comunicar la razón máxima que puede admitirse entre el error absoluto y el riesgo relativo. En los resultados, Epidat 4.0 presenta los tamaños de expuestos y no expuestos por separado para el caso en que interesa realizar la selección de modo independiente en ambos grupos. En el caso contrario, cuando se pide calcular precisión, el dato de entrada debe ser el tamaño total de la muestra.

Ejemplo (Adaptado de Lwanga y Lemeshow [5])

Un epidemiólogo proyecta realizar un estudio sobre la posibilidad de que la aparición de cierta enfermedad pulmonar esté favorecida por la exposición a un contaminante atmosférico recién detectado cuyo efecto no ha sido examinado previamente. ¿Qué tamaño tendrá que tener la muestra de cada grupo (el de expuestos y el de no expuestos) si el epidemiólogo desea estimar el riesgo relativo con una precisión entre el 10% y el 20% del valor real (el cual, verosímilmente, pudiera ascender a 2,0) con un 95% de confianza? La enfermedad se manifiesta en el 20% de las personas no expuestas al contaminante atmosférico y los dos grupos serán de igual tamaño.

Epidat 4.0 arroja los tamaños de muestra 1.904, 800 y 425 en dependencia de que se admitan errores relativos de 10%, 15% y 20%, respectivamente:

Datos:			
Riesgo en expuestos:	40,000%		
Riesgo en no expuestos:	20,000%		
Riesgo relativo esperado:	2,000		
Razón no expuestos/expuestos:	1		
Nivel de confianza:	95,0%		
Resultados:			
Precisión relativa (%)	Tamaño de la muestra		
	Expuestos	No expuestos	Total
10,000	1.904	1.904	3.808
15,000	800	800	1.600
20,000	425	425	850

Puede observarse cómo disminuye el tamaño de muestra necesario a medida que aumenta el error permitido en la estimación. Por otra parte, el programa calculó el riesgo en expuestos, un 40%, a partir de los datos que se le proporcionaron: riesgo en no expuestos y riesgo relativo.

3.1.1.5. Concordancia

El coeficiente de concordancia kappa de Cohen permite cuantificar el grado de acuerdo entre dos o más observadores una vez que se ha eliminado el efecto del azar [13].

Los factores que intervienen en el cálculo del tamaño de muestra (o precisión) necesario para estimar adecuadamente el coeficiente kappa, cuando son dos los observadores y dos las categorías de clasificación (Positivo/Negativo), son:

- *Coefficiente kappa esperado (k):* debe ser un valor entre -1 y 1, recorrido de variación de kappa.
- *Proporción de sujetos clasificados como positivos por el primer observador.*
- *Proporción de sujetos clasificados como positivos por el segundo observador.*

- *Precisión absoluta (e)*: expresa la diferencia entre el valor del coeficiente kappa y el límite inferior (o el superior) de un intervalo de confianza. Téngase en cuenta que no puede introducirse un valor para el cual $k-e$ sea inferior a -1 o $k+e$ supere al 1 .

Si el valor especificado para el coeficiente kappa es incompatible con la proporción de sujetos que cada observador clasificará como positivos, el programa muestra el siguiente aviso: "El valor de kappa es incompatible con las proporciones marginales".

Los detalles sobre el procedimiento de cálculo pueden encontrarse en Cantor [14].

Ejemplo

Supóngase que las radiografías de tórax de los sujetos que están controlados en una unidad de neumología se reparten entre dos radiólogos. Se sabe que el radiólogo A clasifica el 80% como positivas y el radiólogo B clasifica el 76% como positivas, entendiéndose por "positiva" una radiografía sospechosa de lesión tuberculosa. Para estimar la concordancia entre los radiólogos, se piensa seleccionar una muestra de 100 radiografías que serán remitidas a los dos. En este ejemplo, en lugar de estimar el tamaño de muestra necesario, se trata de anticipar la precisión que tendrá el intervalo de confianza del 95% si el índice kappa estuviera alrededor de 0,4 para esa muestra de tamaño 100.

Epidat 4.0 informa que la precisión sería de solo 0,21. Para alcanzar una precisión de 0,15, puede corroborarse que haría falta una muestra de tamaño 201.

Resultados con Epidat 4.0:

<p>Datos:</p> <p>Kappa esperado: 0,400 Clasificaciones(+) observador 1: 80,000% Clasificaciones(+) observador 2: 76,000% Nivel de confianza: 95,0%</p> <p>Resultados:</p> <table border="1"> <thead> <tr> <th>Tamaño de la muestra</th> <th>Precisión</th> </tr> </thead> <tbody> <tr> <td>100</td> <td>0,213</td> </tr> </tbody> </table>	Tamaño de la muestra	Precisión	100	0,213	<p>Datos:</p> <p>Kappa esperado: 0,400 Clasificaciones(+) observador 1: 80,000% Clasificaciones(+) observador 2: 76,000% Nivel de confianza: 95,0%</p> <p>Resultados:</p> <table border="1"> <thead> <tr> <th>Precisión</th> <th>Tamaño de la muestra</th> </tr> </thead> <tbody> <tr> <td>0,150</td> <td>201</td> </tr> </tbody> </table>	Precisión	Tamaño de la muestra	0,150	201
Tamaño de la muestra	Precisión								
100	0,213								
Precisión	Tamaño de la muestra								
0,150	201								

3.1.1.6. Pruebas diagnósticas

La evaluación del desempeño de una prueba diagnóstica (PD) es un campo importante en la investigación médica. Este proceso comienza por la cuantificación (estimación, en rigor) de la magnitud de los errores que pueden cometerse, o su inverso, la magnitud de los aciertos que se cometen al intentar "adivinar" un diagnóstico a partir de los resultados que brinde dicho procedimiento.

La **sensibilidad** y la **especificidad**, introducidas por Yerushalmy [15], son los indicadores estadísticos tradicionales y básicos que evalúan el grado de eficacia inherente a una prueba diagnóstica. Estos índices miden la discriminación diagnóstica de una prueba con relación a un criterio de referencia (una prueba diagnóstica que se considera infalible, conocida como "Golden standard test" en inglés).

La sensibilidad indica la capacidad de la prueba para detectar a un sujeto enfermo, mientras que la especificidad indica su capacidad para clasificar como sanos a los que efectivamente lo son.

En un artículo publicado en 1998, Obuchowsky [16] revisa los métodos que permiten determinar el tamaño de muestra necesario para realizar un estudio de evaluación de una prueba diagnóstica. Entre los índices considerados en la literatura para abordar el cálculo del tamaño muestral se encuentran precisamente la sensibilidad y la especificidad.

Estas características pueden estimarse, respectivamente, como la proporción de verdaderos positivos y la proporción de verdaderos negativos si se tienen los datos dispuestos en una tabla 2x2 como la que sigue:

		Criterio de verdad		
		Enfermos	No enfermos	Total
Prueba diagnóstica	Positivos	a	b	a+b
	Negativos	c	d	c+d
	Total	a+c	b+d	a+b+c+d

Por tanto, el cálculo del tamaño de muestra para construir un IC, con una determinada precisión, para la sensibilidad, la especificidad, o para ambos índices de una prueba diagnóstica, se basa en el método para una proporción, al igual que para obtener la precisión en función de un determinado tamaño de muestra. Los factores que intervienen en el cálculo son:

- *Sensibilidad*, en porcentaje o *especificidad*, también en porcentaje.

Es necesario seleccionar al menos una de estas opciones, pero no son excluyentes. Si se seleccionan las dos, el tamaño de muestra se calculará por separado para cada uno de los dos parámetros y se adoptará como adecuado el máximo de los dos obtenidos, que es el que presenta Epidat 4.0.

- *Condición de enfermo*:

La selección de los pacientes que conformarán la muestra puede hacerse de varias maneras. La vía canónica consiste en seleccionar una muestra de N_1 sujetos enfermos y otra de N_2 no enfermos, caracterizados como tales por la prueba de referencia, y aplicar la nueva prueba a los $N = N_1 + N_2$ individuos, de suerte que pueda conformarse una clasificación cruzada tal como se expone en la anterior tabla de 2x2. Para poder aplicar este diseño es necesario conocer de antemano la condición de enfermo, antes de realizar la selección.

Una segunda opción para el diseño del estudio consiste en obtener una muestra simple aleatoria de N individuos y aplicar a cada uno de ellos tanto el criterio de verdad o prueba de referencia con que se cuenta, como la prueba diagnóstica que se evalúa. Este diseño se plantea cuando la verdadera condición de enfermedad de los sujetos no se conoce en el momento de la selección.

Consecuentemente, Epidat 4.0 solicita al usuario que indique si la condición de enfermo es *conocida* o *desconocida*, para tener en cuenta en el cálculo del tamaño de muestra o de la precisión el diseño que se va a aplicar. Según la opción escogida, conocida o desconocida, será necesario indicarle al programa o bien la *Razón entre los tamaños de muestra de no enfermos y enfermos*, o bien la *Prevalencia de la enfermedad*, respectivamente. Cuando se quiere calcular tamaño de muestra para sensibilidad y especificidad, y se conoce la condición de enfermo, no es necesario disponer de la *Razón entre los tamaños de muestra de no enfermos y enfermos* para realizar el cálculo; por ese motivo, el programa no solicita dicha información en este caso.

- *Razón entre los tamaños de muestra de no enfermos y enfermos (R):*

Debe especificarse en qué razón estarán los tamaños de los dos grupos de sujetos, no enfermos sobre enfermos: $R=N_2/N_1$. Frecuentemente, el investigador desea dos grupos de igual tamaño $N_1=N_2$, en cuyo caso R será igual a 1. Epidat 4.0 presenta los tamaños de enfermos y no enfermos por separado.

- *Prevalencia de la enfermedad:*

La proporción de enfermos en la muestra será aproximadamente igual a la prevalencia de la enfermedad en la población de la que proceden los sujetos. A partir de esta prevalencia (P), Epidat 4.0 estima la razón entre el número de no enfermos y enfermos que conformarán la muestra: $(1-P)/P$, valor que se utiliza en el cálculo del tamaño de muestra.

- *Precisión absoluta:*

Es el error absoluto asociado a la estimación mediante un intervalo de confianza. No es posible especificar valores diferentes para la sensibilidad y la especificidad, en el caso de que el tamaño de muestra esté basado en los dos parámetros. Se asume que los dos intervalos tendrán la misma precisión.

Si se selecciona la opción de calcular tamaños de muestra, Epidat 4.0 presenta el número necesario de enfermos y no enfermos por separado cuando se conoce la condición de enfermo, y el tamaño total cuando se desconoce. En el caso contrario, cuando se pide calcular precisión, el dato de entrada debe ser el tamaño total de la muestra en cualquiera de las dos situaciones.

Ejemplo

En un centro hospitalario se quiere evaluar una prueba diagnóstica de apendicitis que consiste en medir la temperatura sobre el punto de McBurney y sobre la posición simétrica del lado izquierdo; se diagnostica apendicitis aguda si la temperatura cutánea en el lado derecho es por lo menos 1°C superior. Mediante un estudio previo se han estimado la sensibilidad y la especificidad de la prueba en 83% y 92%, respectivamente.

Esta prueba se aplicará a un grupo de pacientes seleccionados para apendicectomía y con apendicitis confirmada por el examen histológico posterior, y a un grupo de controles, ingresados para un tratamiento ortopédico. El número de enfermos y no enfermos que el grupo investigador debe seleccionar para estimar la sensibilidad y la especificidad de la prueba con un nivel de confianza del 95% y una precisión absoluta del 3% asciende a 603 enfermos y 315 sanos.

Resultados con Epidat 4.0:

Datos:			
Sensibilidad esperada:		83,000%	
Especificidad esperada:		92,000%	
Razón no Enfermos/Enfermos:		1,00	
Nivel de confianza:		95,0%	
Resultados:			
Precisión (%)	Tamaño de la muestra		
	Enfermos	No Enfermos	Total
3,000	603	315	918

3.1.2. Contrastes de hipótesis

Como se comentó en el apartado 3.1.1.0, la determinación del tamaño de muestra en la fase de diseño de un estudio se basa en las características del procedimiento estadístico que se va a utilizar para analizar los datos. En los submódulos de Epidat 4.0 incluidos en el módulo de *Muestreo, Cálculo de tamaños de muestra*, bajo el epígrafe “*Contraste de hipótesis*”, el cálculo del tamaño de muestra necesario para un estudio se basa en la potencia (capacidad de detectar significación) deseada para efectuar una prueba de hipótesis. Esos submódulos, que incluyen las pruebas estadísticas aplicadas más frecuentemente en la investigación clínica y epidemiológica, son:

- 3.1.2.1. Comparación de medias
 - 3.1.2.1.1. Grupos independientes
 - 3.1.2.1.2. Grupos emparejados
- 3.1.2.2. Comparación de proporciones
 - 3.1.2.2.1. Grupos independientes
 - 3.1.2.2.2. Grupos emparejados
- 3.1.2.3. Estudios de casos y controles
 - 3.1.2.3.1. Grupos independientes
 - 3.1.2.3.2. Grupos emparejados
- 3.1.2.4. Estudios de cohorte
- 3.1.2.5. Estudios de equivalencia
 - 3.1.2.5.1. Equivalencia de medias
 - 3.1.2.5.2. Equivalencia de proporciones
- 3.1.2.6. Pruebas diagnósticas
 - 3.1.2.6.1. Grupos independientes
 - 3.1.2.6.2. Grupos emparejados
- 3.1.2.7. Calidad de lotes
- 3.1.2.8. Supervivencia
- 3.1.2.9. Coeficiente de correlación

Los contrastes o pruebas de hipótesis son recursos de inferencia estadística que, partiendo de la formulación de dos hipótesis complementarias sobre el posible valor de un parámetro (o de una expresión de varios parámetros), permiten pronunciarse acerca de la veracidad de una de ellas sobre la base de los datos de una muestra. En el planteamiento de un contraste se llama hipótesis nula (H_0) a aquella que afirma que el parámetro (o la expresión de varios parámetros) cumple determinada condición y se llama hipótesis alternativa (H_1) a la que expresa lo contrario, la negación de H_0 . El contraste entraña la comparación de estas dos hipótesis a la luz de la información muestral y la elección adecuada entre las dos.

Cuando se realiza una prueba de hipótesis, la decisión acerca de la veracidad de la hipótesis nula puede conducir a cometer dos tipos de errores:

- Error de tipo I: consistente en rechazar la hipótesis nula cuando es cierta. La probabilidad de incurrir en este error se denomina **nivel de significación** y se denota por α . El valor $1-\alpha$ expresa, por tanto, la probabilidad de no rechazar una hipótesis que es cierta y se conoce como el **nivel de confianza** de la prueba.
- Error de tipo II: consistente en aceptar la hipótesis nula cuando ésta es, en realidad, falsa. La probabilidad de cometer este error se denota por β ; y el valor $1-\beta$ que expresa la

probabilidad de rechazar la hipótesis nula cuando es falsa, se conoce como la **potencia** de la prueba.

Nota: Usualmente el investigador termina rechazando H_0 o “no rechazando” H_0 . Aunque este es el léxico en boga actualmente (debido a consideraciones que escapan al contenido de esta ayuda), cabe advertir que la definición original de Pearson y Neyman no aludía a “no rechazo” sino a “aceptación” de H_0 .

En el contexto de contraste de hipótesis, por tanto, el tamaño muestral está determinado por el nivel de confianza y la potencia de la prueba, además de otros elementos que dependen del diseño y de la prueba concreta que se vaya a aplicar.

Epidat 4.0 permite calcular tamaños de muestra de una forma flexible, generando una tabla de valores para la potencia y proporcionando el tamaño de muestra correspondiente a cada uno de esos valores. La tabla se define partiendo de un valor *mínimo* que aumenta a intervalos regulares de *incremento* hasta llegar al valor *máximo*. Por ejemplo, si se definen *mínimo*=80, *máximo*=90 e *incremento*=5, se obtiene una tabla con los valores 80, 85 y 90 para la potencia y los correspondientes tamaños de muestra.

Un recurso que incorpora Epidat 4.0 y que puede ser de utilidad en ciertas situaciones es la posibilidad de realizar el cálculo inverso: determinar la potencia asociada a una prueba de hipótesis en función del tamaño de muestra disponible. Igual que en la situación inversa, el cálculo es flexible, ya que permite definir una tabla de valores para el tamaño de muestra a partir de un *mínimo*, un *máximo* y un *incremento*, y proporciona la potencia que corresponde a cada uno.

Advertencias y recomendaciones

- La determinación a priori de los valores que intervienen en las fórmulas del tamaño muestral es una decisión esencialmente subjetiva.
- El cálculo del tamaño de muestra requerido en una investigación clínica o estudio epidemiológico debe basarse en la prueba estadística que se prevea utilizar en la fase de análisis. Sin embargo, es habitual en la mayoría de los estudios (especialmente los de tipo observacional) aplicar diferentes pruebas, además de la tomada como referencia para calcular el tamaño muestral.
- El empleo de pruebas o contraste de hipótesis ha sido seriamente cuestionado y no constituye la recomendación actual del Comité Internacional de Editores de Revistas Médicas (véanse [11][17][18][19]).

Una abarcadora reflexión sobre las implicaciones que tienen estas circunstancias en el cálculo del tamaño de muestra puede verse en el capítulo 12 del libro “Diseño razonado de muestras y captación de datos para la investigación sanitaria” de Silva [3].

3.1.2.1. Comparación de medias

3.1.2.1.1. Grupos independientes

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) con vistas a la comparación de medias de datos cuantitativos, cuando los dos grupos son independientes, asume que los datos siguen una distribución normal y se basa en la prueba t de Student. El estadístico de esta prueba, que sigue una distribución t de Student, tiene dos versiones en función de que las

varianzas de los dos grupos puedan asumirse iguales o no, de modo que también la fórmula para el cálculo del tamaño de muestra o de la potencia depende de esta circunstancia. Por esta razón, Epidat 4.0 solicita al usuario que, antes de introducir los datos, seleccione entre dos opciones excluyentes relativas a la varianza de los dos grupos:

- *Varianzas iguales:* en este caso, Epidat 4.0 ofrece dos alternativas para los datos de entrada necesarios. En la primera opción se piden la *Diferencia de medias a detectar* y la *Desviación estándar común a los dos grupos*. En la segunda opción, el valor que pide el programa es la *Diferencia estandarizada de medias*.
- *Varianzas distintas:* en este caso, es necesario introducir al programa la *Diferencia de medias a detectar* y las *Desviaciones estándar de cada grupo*.

Teniendo en cuenta estas alternativas, los factores que intervienen en el cálculo del tamaño de muestra o de la potencia para comparar las medias de dos grupos independientes son:

- *Diferencia de medias a detectar entre los dos grupos:*

Esta magnitud representa la mínima diferencia que se considera sustantiva (clínica, social o económicamente relevante) y que se quiere poder declarar, en caso de que exista, como estadísticamente significativa, con el nivel de confianza y la potencia prefijadas. Es lo que, en ensayos clínicos, se conoce como “tamaño del efecto”.

Un índice que cuantifica la magnitud de la diferencia entre los grupos, pero que es independiente de las unidades de medida, es la denominada *Diferencia estandarizada de medias* (δ), y que resulta de dividir la diferencia de medias, en valor absoluto, entre la desviación estándar común para los dos grupos. Esta alternativa ha sido utilizada para generar tablas de tamaños muestrales, en el caso de comparación de medias independientes, de uso general. Los valores tabulados de δ , siguiendo las recomendaciones de Cohen [20], oscilan entre 0,05 y 1,50. Según este autor, para un efecto “pequeño” se puede definir $\delta=0,2$ y si el efecto es “grande” se puede tomar $\delta=0,8$.

- *Desviación estándar:*

Si se asume que las varianzas de los dos grupos son iguales y se opta por la primera opción para introducir los datos, hay que comunicar al programa un valor para la desviación estándar común a los dos grupos. En el caso de varianzas distintas, habrá que introducir el valor de la desviación estándar que se supone tiene cada grupo. Pero aún en esta circunstancia, es muy habitual en la práctica asumir que los valores de las desviaciones estándar de los grupos son iguales y adelantar una estimación para ese valor común. Machin y cols [21] sugieren, cuando no se conoce la magnitud de la desviación estándar, una aproximación que consiste en dividir entre 4 el recorrido (diferencia entre el máximo y el mínimo) de las observaciones, magnitud que suele ser más fácil de estimar o conocer de antemano.

- *Razón entre los tamaños de muestra de los dos grupos:*

Habitualmente, el diseño de un estudio para la comparación de medias en dos grupos independientes contempla la selección de dos muestras de igual tamaño, pero puede que no sea así. En Epidat 4.0, es necesario especificar la razón entre los tamaños de muestra de los dos grupos ($R=n_2/n_1$).

- *Nivel de confianza de la prueba:*

Complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Si se selecciona la opción de calcular tamaños de muestra, el resultado que produce Epidat 4.0 es el tamaño de cada uno de los grupos en función del valor o valores de potencia especificada. En el caso contrario, cuando se pide calcular potencia, el dato de entrada debe ser el tamaño de muestra total, es decir, la suma de los dos grupos.

Ejemplo

En un ensayo clínico sobre el uso de una droga en embarazos gemelares, un tocólogo desea demostrar que hay un aumento estadísticamente significativo en la duración del embarazo al usar la droga frente a un placebo. El especialista estima que la desviación estándar de la duración de los embarazos en ambos grupos es de 1,75 semanas. ¿Cuántos embarazos debe observar como mínimo en cada grupo si considera que una semana es un aumento mínimo clínicamente importante en la duración del embarazo y quiere operar con una confianza del 95% y, una potencia del 80%?

Epidat 4.0, con varianzas iguales y mediante la opción 1, arroja que se necesitan 50 unidades de análisis en cada uno de los grupos. Lo mismo ocurre si se elige la opción 2, en este caso se ha de comunicar que la diferencia estandarizada de medias es $1/1,75=0,57$.

Nota: La razón entre los tamaños de muestra de los grupos es 1, ya que los dos tendrán el mismo tamaño.

Resultados con Epidat 4.0:

Datos:				Datos:			
Varianzas:		Iguales		Varianzas:		Iguales	
Opción:		Opción 1		Opción:		Opción 2	
Diferencia de medias a detectar:		1,000		Diferencia estandarizada de medias:		0,570	
Desviación estándar común:		1,750		Razón entre tamaños muestrales:		1,00	
Razón entre tamaños muestrales:		1,00		Nivel de confianza:		95,0%	
Nivel de confianza:		95,0%		Resultados:			
Resultados:				Resultados:			
Potencia (%)	Tamaño de la muestra			Potencia (%)	Tamaño de la muestra		
	Población 1	Población 2	Total		Población 1	Población 2	Total
80,0	50	50	100	80,0	50	50	100

3.1.2.1.2. Grupos emparejados

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) para la comparación de medias de datos cuantitativos, cuando los dos grupos de observaciones están emparejados, asume que los datos siguen una distribución normal y se basa en la prueba t de Student para una muestra aplicada a las diferencias entre los pares de observaciones. Teniendo en cuenta que Epidat 4.0 ofrece dos alternativas para los datos de entrada, los factores que intervienen en el cálculo son:

- *Diferencia de medias a detectar:*

Esta cantidad representa la magnitud mínima de la diferencia que se conceptúa como relevante y que se quiere poder calificar, en caso de que exista, como *estadísticamente significativa*, con el nivel de confianza y la potencia prefijadas. Es lo que, en ensayos clínicos, se conoce como “tamaño del efecto”.

Opción 1:

- *Desviación estándar de cada grupo:*

Es muy habitual en la práctica asumir que los dos valores son iguales y adelantar una estimación para ese valor común. Machin y cols [21] sugieren, cuando no se conoce la magnitud de la desviación estándar, una aproximación que consiste en dividir por 4 el recorrido (diferencia entre el máximo y el mínimo) de las observaciones, cantidad que suele ser más fácil de establecer.

- *Coefficiente de correlación de Pearson* entre los valores de la variable resultado en los dos grupos, que debe ser un valor entre -1 y 1.

Estos tres datos son necesarios para obtener una estimación de la *Desviación estándar de las diferencias* individuales entre los pares de observaciones.

- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Opción 2:

- *Desviación estándar de las diferencias* individuales entre los pares de observaciones.
- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Ejemplo

Un investigador quiere valorar la diferencia entre el valor de la presión sistólica en sangre en personas de mediana edad, antes y después de la práctica de ejercicio físico. Si el investigador aspira a poder declarar una diferencia de 2,6 mmHg o mayor como significativa, con un nivel de confianza del 95% y una potencia del 80% ¿cuántos pacientes deberá incluir en el estudio? Supóngase que se estima una desviación estándar de 11,0 mmHg y un coeficiente de correlación de 0,70 entre los valores de la presión sistólica antes y después del ejercicio.

El resultado según Epidat 4.0 es que harían falta 87 pares.

3.1.2.2. Comparación de proporciones

3.1.2.2.1. Grupos independientes

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) necesario para comparar dos proporciones, cuando los dos grupos son independientes, se basa en la prueba ji-cuadrado de Pearson, con o sin corrección de Yates. Los factores que intervienen en el cálculo son:

- *Proporción esperada en cada grupo.*
- *Razón entre tamaños muestrales (R):*
Debe especificarse en qué razón estarán los tamaños de los dos grupos de sujetos, grupo 2 sobre grupo 1: $R=n_2/n_1$. Frecuentemente, el investigador desea que los dos grupos sean de igual tamaño $n_1=n_2$, por lo que R será igual a 1.
- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Si se selecciona la opción de calcular tamaños de muestra, Epidat 4.0 presenta los tamaños de muestra para cada una de las poblaciones por separado. En el caso contrario, cuando se pide calcular potencia, el dato de entrada debe ser el tamaño total de la muestra.

Ejemplo

Supongamos que se realiza un ensayo clínico para evaluar si un remedio homeopático para la gripe se distingue de un tratamiento basado en agua. El criterio de respuesta es la curación transcurrida una semana desde la aparición de los primeros síntomas. Se trata por tanto de fijar un tamaño de muestra para cada uno de los grupos que se conformarán con asignación aleatoria. Se ha decidido tomar tamaños iguales en los dos grupos (es decir, $R=n_2/n_1$) y que no se aplicará la corrección de Yates. Se espera que el porcentaje de éxito empleando solo agua estaría alrededor de 35% y se considera que un aumento de 5 puntos porcentuales en la tasa para el tratamiento homeopático (que, en tal caso, sería del 40%) sería relevante y se decide asignar al ensayo una potencia de 80% (la cifra más habitual) y un nivel de confianza del 90%. En estas condiciones, el tamaño en cada uno de los grupos ascendería a 1.159.

Resultados con Epidat 4.0:

Datos:			
Proporción esperada en			
Población 1:		35,000%	
Población 2:		40,000%	
Razón entre tamaños muestrales:		1,00	
Nivel de confianza:		90,0%	
Resultados:			
Potencia (%)	Tamaño de la muestra*		
	Población 1	Población 2	Total
80,0	1.159	1.159	2.318
*Tamaños de muestra para aplicar el test χ^2 sin corrección por continuidad.			

3.1.2.2.2. Grupos emparejados

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) necesario para comparar dos proporciones, cuando los grupos están emparejados, se basa en la prueba de McNemar [22]. Los factores que intervienen en el cálculo son:

- *Proporción esperada en cada grupo.*
- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Ejemplo

Un investigador está diseñando un estudio para comparar la capacidad que tienen dos medios, A y B, para detectar la presencia de bacilos de la tuberculosis. Para ello es necesario seleccionar un número de muestras de esputo que se cultivarán en cada uno de los dos

medios. Por estudios previos, se piensa que aproximadamente el 60% de las muestras son positivas en el medio A y un 45% en el medio B. El tamaño de muestra necesario para comparar las dos proporciones con un nivel de confianza del 95% y una potencia del 85% sería de 201 pares de muestras de esputo.

Resultados con Epidat 4.0:

Datos:	
Proporción esperada en	
Población 1:	60,000%
Población 2:	45,000%
Nivel de confianza:	95,0%
Resultados:	
Potencia (%)	Número de pares
85,0	201

3.1.2.3. Estudios de casos y controles

3.1.2.3.1. Grupos independientes

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) necesario para estudiar si existe asociación entre la exposición y la enfermedad en un estudio de casos y controles, cuando los dos grupos son independientes, se basa en la prueba ji-cuadrado de Pearson, con o sin corrección de Yates. Los factores que intervienen en el cálculo son:

- *Proporción de casos expuestos o prevalencia de la exposición en los casos (P_1).*
- *Proporción de controles expuestos o prevalencia de la exposición en los controles (P_2).*
- *Odds ratio a detectar (OR).*

Los tres valores precedentes están relacionados entre sí del modo siguiente:

$$OR = \frac{P_1(1-P_2)}{P_2(1-P_1)},$$

por lo que, especificando dos de estos parámetros, el otro queda determinado. Epidat 4.0 permite que el usuario decida cuál pareja de valores desea introducir y calcula automáticamente el tercero.

- *Número de controles por caso:*

La determinación del número de controles es una cuestión muy importante al planificar un estudio de casos y controles. Generalmente se utiliza igual número de controles que de casos. Se ha demostrado que hay muy poca ganancia en la precisión de un estudio cuando la razón (número de controles)/(número de casos) supera el valor 4 [12]. No obstante, Epidat 4.0 admite que se establezcan hasta 10 controles por caso.

- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Si se selecciona la opción de calcular tamaños de muestra, el resultado que produce Epidat 4.0 es el tamaño de cada uno de los grupos, casos y controles, por separado. En el caso contrario, cuando se pide calcular potencia, el dato de entrada debe ser el tamaño total de la muestra.

Ejemplo

Se desea realizar un estudio de casos y controles, con 3 controles por caso, para analizar la influencia del consumo de alcohol en la aparición de cáncer de esófago. ¿Cuántos sujetos deberían estudiarse si se quiere poder declarar como significativa una OR de 2,5 o mayor, y se supone que la exposición entre los controles es del 60%? (Supongamos un nivel de confianza del 95% y una potencia del 80%, así como que la prueba ji-cuadrado se haría con corrección de Yates). Según arroja Epidat 4.0 harían falta 70 casos y 210 controles.

Resultados con Epidat 4.0:

Datos:			
Proporción de casos expuestos:	78,947%		
Proporción de controles expuestos:	60,000%		
Odds ratio a detectar:	2,500		
Número de controles por caso:	3		
Nivel de confianza:	95,0%		
Resultados:			
Potencia (%)	Tamaño de la muestra*		
	Casos	Controles	Total
80,0	70	210	280

*Tamaños de muestra para aplicar el test χ^2 con la corrección por continuidad de Yates (χ^2).

3.1.2.3.2. Grupos emparejados

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) para estudiar si existe asociación entre la exposición y la enfermedad en un estudio de casos y controles, cuando los dos grupos están emparejados, se basa en la prueba de McNemar [22]. Los factores que intervienen en el cálculo son:

- *Proporción de casos expuestos o prevalencia de la exposición en los casos (P_1).*
- *Proporción de controles expuestos o prevalencia de la exposición en los controles (P_2).*
- *Odds ratio a detectar (OR).*

Estos tres valores están relacionados entre sí del modo siguiente:

$$OR = \frac{P_1(1 - P_2)}{P_2(1 - P_1)},$$

por lo que, especificando dos de estos parámetros, el otro queda perfectamente determinado. Epidat 4.0 permite que el usuario decida cuál pareja de valores desea introducir y calcula automáticamente el tercero.

- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Ejemplo

Se diseña un estudio de casos y controles emparejados para valorar la influencia de tabaquismo en el cáncer de laringe utilizando controles poblacionales. Suponiendo que la prevalencia del hábito tabáquico en la población es del 45%, ¿cuántos pares de casos y controles se necesitarán para poder declarar que una OR de 2 o mayor es significativa con un nivel de confianza del 95% y una potencia del 85%? La respuesta que ofrece Epidat 4.0 es que harían falta 155 parejas.

3.1.2.4. Estudios de cohorte

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) para estudiar si existe asociación entre la exposición y la enfermedad en un estudio de cohortes se basa en la prueba ji-cuadrado de Pearson, con o sin corrección de Yates. Los factores que intervienen en el cálculo son:

- *Riesgo en expuestos (P_1).*
- *Riesgo en no expuestos (P_2).*
- *Razón entre el número de no expuestos y el de expuestos (R):*

Cuando se plantea el diseño de un estudio de cohortes, generalmente no se conoce de antemano el número de sujetos expuestos en la población que se estudia. Por lo general, se selecciona aleatoriamente una muestra de la población y posteriormente se determina cuáles individuos de los que conforman la cohorte están expuestos. En esa situación, la proporción de expuestos en la cohorte será aproximadamente igual a la prevalencia de la exposición en la población de la que procede la muestra. A partir de esta prevalencia (P) puede estimarse la razón entre el número de no expuestos y expuestos que conformarán la cohorte: $R=(1-P)/P$, valor que interviene en el cálculo del tamaño de muestra mínimo para que la prueba tenga la potencia exigida [21]. También puede darse la situación de que la exposición se conozca previamente a la selección de los sujetos y desee seleccionarse, por ejemplo, el mismo número de expuestos que de no expuestos. En ese caso, la razón tomará el valor 1.

- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Si se selecciona la opción de calcular tamaños de muestra, Epidat 4.0 presenta como resultado los tamaños de expuestos y no expuestos por separado, para el caso en que interesa realizar la selección de modo independiente en cada grupo. En el caso contrario, cuando se pide calcular potencia, el dato de entrada debe ser el tamaño total de la muestra.

Ejemplo

Un grupo de investigadores desea llevar a cabo un estudio de cohortes que exige determinar si existe asociación entre la infección crónica por *Chlamydia pneumoniae* y los eventos coronarios en varones adultos con un primer diagnóstico de angina inestable. Se sabe, por estudios previos, que la prevalencia de la exposición (infección) en este grupo (anginosos) está alrededor del 40%, que la proporción esperada de eventos coronarios es del 38% en el

grupo de expuestos y del 32% en el grupo de no expuestos ¿Cuál sería el tamaño de muestra necesario para contrastar la hipótesis nula –sin usar corrección de Yates- de que no existe asociación entre la infección y los eventos coronarios, con un nivel de confianza del 95% y una potencia del 80%?

Nota: Si la prevalencia de la exposición es del 40% ($P=0,4$), la razón de no expuestos con respecto a los expuestos será de $(1-0,4)/0,4=1,5$.

La muestra necesaria, según Epidat 4.0 sería de 823 sujetos expuestos y 1.235 no expuestos:

Datos:			
Riesgo en expuestos:	38,000%		
Riesgo en no expuestos:	32,000%		
Riesgo relativo a detectar:	1,188		
Razón no expuestos/expuestos:	1,50		
Nivel de confianza:	95,0%		
Resultados:			
Potencia (%)	Tamaño de la muestra*		
	Expuestos	No expuestos	Total
80,0	823	1.235	2.058

*Tamaños de muestra para aplicar el test χ^2 sin corrección por continuidad.

3.1.2.5. Estudios de equivalencia

Cuando se comparan dos tratamientos, no siempre se está interesado en encontrar diferencias estadísticamente significativas entre ellos. Ocasionalmente, el interés se centra en mostrar que ambos tratamientos son clínicamente equivalentes en eficacia. De forma que se podría estar interesado en corroborar que un nuevo tratamiento es tan bueno en términos de eficacia como uno estándar; esto es así especialmente cuando este nuevo tratamiento conlleve menos efectos secundarios, pueda ser más aceptado socialmente, sea más barato o tenga menor duración.

Estos estudios de equivalencia se basan en la comprobación de que uno de los tratamientos no es estadísticamente peor que el otro, lo que exige la aplicación de un contraste unilateral, basado en una de las dos colas de la distribución del estadístico del contraste. En este tipo de estudios el contraste que se plantea es el siguiente:

$$H_0 : \theta_1 = \theta_2 + \varepsilon$$

$$H_1 : \theta_1 < \theta_2 + \varepsilon$$

donde θ_1 y θ_2 representan o bien las medias o bien las proporciones de cada grupo de tratamiento, y ε representa la diferencia máxima, o bien de medias o bien de proporciones, respectivamente, por debajo de la cual se considera a todos los efectos prácticos que hay equivalencia.

Epidat 4.0 ofrece la posibilidad de calcular el tamaño de muestra (o potencia) para realizar estudios de equivalencia de medias o de proporciones.

3.1.2.5.1. Equivalencia de medias

Cuando se quiere realizar un estudio de equivalencia de medias, se está interesado en poder afirmar que dichas medias son iguales a todos los efectos prácticos, aunque realmente exista una pequeña diferencia entre ellas. Para ello se asume que los datos correspondientes a cada tratamiento siguen una distribución normal, con la misma desviación estándar.

Los factores que intervienen en el cálculo son:

- *Diferencia de medias esperada.*
- *Desviación estándar de cada grupo:*

Es muy habitual en la práctica asumir que los valores de las desviaciones estándar de los grupos son iguales y adelantar una estimación para ese valor común. Machin y cols [21] sugieren, cuando no se conoce la magnitud de la desviación estándar, una aproximación que consiste en dividir entre 4 el recorrido (diferencia entre el máximo y el mínimo) de las observaciones, magnitud que suele ser más fácil de estimar o conocer de antemano.
- *Diferencia de medias máxima que indica equivalencia (ϵ):*

Esta magnitud representa la máxima diferencia admisible entre los dos procedimientos o tratamientos por debajo de la cual puede considerarse que hay equivalencia.
- *Razón entre tamaños muestrales:*

Debe especificarse en qué razón estarán los tamaños de los dos grupos de sujetos, grupo 2 sobre grupo 1: $R=n_2/n_1$. Frecuentemente, el investigador desea dos grupos de igual tamaño $n_1=n_2$, por lo que R será igual a 1.
- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Ejemplo (Adaptado de Machin & Campbell [21])

Se sabe que los pacientes que toman determinado medicamento tienen una presión arterial diastólica (PAD) media de 96 mmHg mientras que la media es 94 mmHg para los que toman un medicamento alternativo. También se sabe que la desviación estándar de la PAD es aproximadamente 8 mmHg. Para confirmar que la verdadera diferencia entre los dos medicamentos es menor que 5 mmHg, lo que indicaría que son equivalentes, sería necesario seleccionar 88 individuos en cada grupo, asumiendo un nivel de confianza del 95% y una potencia del 80%.

Resultados con Epidat 4.0:

Datos:			
Diferencia de medias esperada:	2,000		
Desviación estándar esperada en las dos poblaciones:	8,000		
Diferencia de medias máxima que indica equivalencia:	5,000		
Razón entre tamaños muestrales:	1,00		
Nivel de confianza:	95,0%		
Resultados:			
Potencia (%)	Tamaño de la muestra		
	Población 1	Población 2	Total
80,0	88	88	176

3.1.2.5.2. Equivalencia de proporciones

Cuando se quiere realizar un estudio de equivalencia entre dos procedimientos usando proporciones, se está interesado en poder afirmar que las proporciones respectivas son iguales a todos los efectos prácticos, aunque realmente exista una pequeña diferencia entre ellas.

En este caso el límite inferior del intervalo de confianza (LI) para la diferencia de proporciones está determinado por la máxima diferencia posible entre los dos tratamientos, que se daría cuando la proporción en el grupo 1 fuera igual a 0 y la del grupo 2 igual a 1. De esta manera, LI toma el valor -1.

Los factores que intervienen en el cálculo son:

- *Proporción esperada en cada grupo.*
- *Diferencia de proporciones máxima que indica equivalencia (ϵ):*
Esta magnitud representa la máxima diferencia admisible por debajo de la cual puede considerarse que hay equivalencia.
- *Razón entre tamaños muestrales:*
Debe especificarse en qué razón estarán los tamaños de los dos grupos de sujetos, grupo 2 sobre grupo 1: $R=N_2/N_1$. Frecuentemente, el investigador desea dos grupos de igual tamaño $N_1=N_2$, por lo que R será igual a 1.
- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Ejemplo (Adaptado de Machin & Campbell [21])

Bradshaw et al diseñaron un ensayo clínico para contrastar si la quimioterapia combinada durante un corto período de tiempo es al menos tan buena como la terapia convencional en pacientes con meningitis criptocócica. Ellos reclutaron 39 pacientes para cada grupo y establecieron que una diferencia inferior al 20% en la tasa de respuesta entre los tratamientos

indica que son equivalentes. Asumiendo un nivel de confianza del 90%, una potencia del 80% y una tasa global de respuesta del 50%, ¿cuál sería un tamaño de muestra más realista si el ensayo tuviera que ser repetido?

Resultados con Epidat 4.0

Datos:			
Proporción esperada en			
Población 1:			50,000%
Población 2:			50,000%
Diferencia de proporciones máxima que indica equivalencia:			
			20,000%
Razón entre tamaños muestrales:			
			1,00
Nivel de confianza:			
			90,0%
Resultados:			
Potencia (%)	Tamaño de la muestra		
	Población 1	Población 2	Total
80,0	55	55	110

3.1.2.6. Pruebas diagnósticas

En su artículo de 1998, Obuchowsky [16] aborda los métodos que permiten determinar el tamaño de muestra necesario para realizar un estudio de evaluación de pruebas diagnósticas. El cálculo del tamaño de muestra (o potencia) para comparar dos pruebas diagnósticas a través de la sensibilidad, la especificidad, o ambos índices, se basa en la comparación de dos proporciones. En un diseño con grupos independientes, es decir, si las muestras de sujetos a los que se aplica cada una de las pruebas son independientes, se utiliza la prueba ji-cuadrado de Pearson, con o sin corrección de Yates. En el caso de observaciones emparejadas, como sucede si se aplican las dos pruebas diagnósticas a la misma muestra de pacientes, el cálculo se basa en la prueba de McNemar.

Nota: Se asume que las dos muestras, cada una correspondiente a una prueba diagnóstica, serán de igual tamaño, tanto si los grupos son independientes como emparejados. Además, cada una de ellas estará compuesta por sujetos enfermos y no enfermos, no necesariamente en la misma proporción; pero la razón entre el número de unos y otros sí será igual en las dos muestras.

En las dos opciones, grupos independientes o emparejados, los factores que intervienen en el cálculo del tamaño de muestra son:

- *Sensibilidad de cada prueba*, en porcentaje.
- *Especificidad de cada prueba*, en porcentaje.

Es necesario seleccionar al menos una de estas opciones, pero no son excluyentes. Si se seleccionan ambas, entonces el tamaño de muestra se calculará por separado para cada uno de los dos parámetros y se adoptará como adecuado el máximo de los dos obtenidos, que es el que presenta Epidat 4.0.

- *Condición de enfermo.*

La selección de los pacientes que conformarán cada una de las muestras puede hacerse de varias maneras. La vía canónica consiste en seleccionar una muestra de N_1 sujetos enfermos y otra de N_2 no enfermos diagnosticados por una prueba de referencia, y aplicar la nueva prueba a los $n=n_1+n_2$ individuos, de suerte que pueda conformarse una clasificación cruzada en una tabla de 2×2 . Para poder aplicar este diseño es necesario conocer de antemano la condición de enfermo, antes de realizar la selección.

Una segunda opción para el diseño del estudio consiste en obtener una muestra simple aleatoria de N individuos y aplicar a cada uno de ellos tanto el criterio de verdad o prueba de referencia con que se cuenta, como la prueba diagnóstica que se evalúa. Este diseño se plantea cuando la verdadera condición de enfermedad de los sujetos no se conoce en el momento de la selección.

En función de esto, Epidat 4.0 solicita al usuario que se indique si la condición de enfermo es *conocida* o *desconocida*, para tener en cuenta en el cálculo del tamaño de muestra (o de la precisión) el diseño que se va a aplicar. Según la opción escogida, conocida o desconocida, será necesario indicarle al programa o bien la *Razón entre los tamaños de muestra de no enfermos y enfermos*, o bien la *Prevalencia de la enfermedad*, respectivamente. Cuando se quiere calcular el tamaño de muestra para sensibilidad y especificidad, y se conoce la condición de enfermo, no es necesario disponer de la *Razón entre los tamaños de muestra de no enfermos y enfermos* para realizar el cálculo, por ese motivo el programa no solicita dicha información en este caso.

- *Razón entre los tamaños de muestra de no enfermos y enfermos (R):*

Debe especificarse en qué razón estarán los tamaños de los dos grupos de sujetos, no enfermos sobre enfermos: $R=n_2/n_1$. Frecuentemente, el investigador desea dos grupos de igual tamaño $n_1=n_2$, por lo que R será igual a 1. Con esta opción, Epidat 4.0 presentará los tamaños de enfermos y no enfermos por separado.

- *Prevalencia de la enfermedad:*

La proporción de enfermos en la muestra será aproximadamente igual a la prevalencia de la enfermedad en la población de la que proceden los sujetos. A partir de esta prevalencia (P), Epidat 4.0 estima la razón entre el número de no enfermos y enfermos que conformarán la muestra: $(1-P)/P$, valor que se utiliza en el cálculo del tamaño de muestra.

- *Potencia de la prueba:* el valor será el mismo para la sensibilidad y la especificidad, en el caso de que el tamaño de muestra esté basado en los dos parámetros.

- *Nivel de confianza de la prueba:* complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).

Si se selecciona la opción de calcular tamaños de muestra, Epidat 4.0 presenta los tamaños de muestra para enfermos y para no enfermos por separado cuando la condición de enfermo es conocida, y el tamaño total de la muestra cuando es desconocida. En el caso contrario, cuando se pide calcular potencia en función del tamaño, el dato de entrada debe ser el tamaño total de la muestra en cualquiera de las dos opciones.

Ejemplo

Un departamento de radiología se propone realizar un estudio para evaluar la precisión de una radiografía mamográfica respecto de la misma imagen digitalizada por monitor. La aspiración con que se realiza este estudio es la de demostrar que es posible sustituir las radiografías por las imágenes, que resultan mucho más baratas. Un radiólogo revisará las radiografías y las imágenes digitales de una única muestra de mujeres remitidas al hospital a las que se les determinará el verdadero diagnóstico mediante biopsia. Se espera que la sensibilidad de las pruebas sea de un 80% en el caso de las mamografías y de un 90% para las imágenes digitales. Se quiere calcular el tamaño de muestra necesario, sin usar corrección de Yates, para evaluar la hipótesis de igualdad entre las dos pruebas diagnósticas con una confianza del 95% y una potencia del 80%, teniendo en cuenta que se confirma el diagnóstico de cáncer de mama en un 10% de los casos revisados.

Usando la opción de muestras emparejadas, Epidat 4.0 arroja que, bajo tales condiciones, haría falta estudiar 1.990 radiografías.

Resultados con Epidat 4.0:

Datos:	
Sensibilidad esperada	
Prueba 1:	80,000%
Prueba 2:	90,000%
Prevalencia de la enfermedad:	10,000%
Nivel de confianza:	95,0%
Resultados:	
Potencia (%)	Tamaño de la muestra
80,000	1.990
*Tamaños de muestra para aplicar el test χ^2 sin corrección por continuidad.	

3.1.2.7. Calidad de lotes

El muestreo de aceptación de lotes, en inglés *Lot Quality Assurance Sampling* (LQAS), es un método de control de calidad desarrollado en la industria que puede resultar de gran utilidad para la monitorización en servicios de salud [23][24][25].

La evaluación periódica sobre el comportamiento de indicadores mediante métodos de muestreo tradicionales, que exigen estimaciones precisas, es una tarea que requiere de una importante cantidad de recursos humanos y materiales. Una opción es la de aplicar técnicas alternativas, tales como el plan de muestreo por lotes, el cual ofrece la posibilidad de utilizar muestras pequeñas con el fin de tomar decisiones sobre el nivel de cumplimiento de indicadores en un lote o población.

Tal como se define en la literatura sobre control de calidad en la industria, el muestreo para la aceptación de lotes de un determinado producto supone el examen de una porción de un lote con el propósito de aceptarlo o rechazarlo en su totalidad. Para la aplicación de esta técnica se requiere:

- Definir la característica que va a ser observada: variable dicotómica que permite identificar cada elemento del lote como “defectuoso” o no. La decisión de aceptar o rechazar el lote se basará en el porcentaje de elementos defectuosos en la muestra.
- Definir el lote o población, que debe ser lo más homogéneo posible respecto a la característica que se pretende evaluar.
- Definir los parámetros:
 - *Proporción de defectuosos a detectar*: es un valor que define mala calidad y por encima del cual se considera que un lote es inaceptable. La proporción complementaria (porcentaje de no defectuosos) se denomina nivel de calidad rechazable (NCR), y define un umbral inferior de cumplimiento de la calidad.
 - *Proporción esperada de defectuosos*: la proporción complementaria se llama nivel de calidad aceptable (NCA) porque es el nivel mínimo deseable de cumplimiento del indicador.
 - *Riesgo alfa (α)*: probabilidad de que un lote con proporción esperada de defectuosos inferior a la proporción de defectuosos a detectar sea valorado como defectuoso. $1-\alpha$ es el nivel de confianza.
 - *Riesgo beta (β)*: probabilidad de que un lote con nivel de cumplimiento inferior al NCR sea valorado como bueno. $1-\beta$ es la potencia.
- Definir el plan de muestreo que se desea emplear:
 - *n*: *Tamaño de la muestra*.
 - *d*: *Umbral de decisión*.

El par (n, d) permite tomar la decisión de aceptar o rechazar el lote con los niveles de calidad y riesgos prefijados. El procedimiento consiste en lo siguiente:

- En cada lote se extrae una muestra aleatoria simple de tamaño n.
- Se cuenta el número de elementos que no cumplen el requisito de calidad evaluado (defectuosos). Si este número es mayor que el umbral “d”, se rechaza el lote. En caso contrario se acepta el lote, es decir, se puede considerar que la proporción de defectuosos no supera el valor esperado.

Epidat 4.0 permite obtener los valores del par (n, d) para valores especificados de la proporción de defectuosos a detectar y la proporción esperada de defectuosos en un lote. La primera debe tomar un valor mayor que la esperada.

Ejemplo

En su departamento, la cobertura prevista para un programa de tamizaje de cáncer de mama es del 70%. Detectadas ciertas deficiencias en el programa, se desea encuestar a mujeres del grupo de edad seleccionado por el programa de suerte que se puedan detectar barrios y zonas donde la cobertura es menor o igual que 45%. Supongamos que se desea tener una confianza del 95% de que se reconocerán zonas donde se ha alcanzado la cobertura de tamizaje preestablecida.

¿Cuántas mujeres habrá que encuestar en cada barrio o zona y qué valor umbral habrá de utilizarse para contrastar la hipótesis de que la proporción de mujeres perdidas por el programa no supera el 55%?

Epidat 4.0 establece que el tamaño en el lote sería 24 y el umbral sería 9 unidades.

Resultados con Epidat 4.0:

Datos:		
Proporción de defectuosos		
esperada:		30,000%
a detectar:		55,000%
Nivel de confianza:		95,0%
Resultados:		
Potencia (%)	Tamaño de muestra	Valor umbral*
80,0	24	9
*Valor umbral de individuos con la característica.		

3.1.2.8. Supervivencia

En muchos ensayos clínicos se realiza un seguimiento de los pacientes asignados a cada tratamiento hasta que ocurre algún suceso crítico, y se utiliza el tiempo de seguimiento para comparar la eficacia de los tratamientos mediante técnicas de supervivencia.

En Epidat 4.0, el cálculo del tamaño de muestra para comparar curvas de supervivencia se basa en el test logrank [26], asumiendo un modelo de riesgos proporcionales para los tiempos de supervivencia.

Los factores que intervienen en el cálculo son:

- *Número de curvas que se comparan.*
- *Probabilidad de supervivencia para cada curva, en porcentaje.*
- *Proporción de pérdidas en el seguimiento de los pacientes.*
- *Razón entre los tamaños de los grupos, en el caso de dos curvas. Si se comparan tres o más curvas, se asume que los tamaños de los grupos son iguales.*
- *Nivel de confianza de la prueba: complemento del error de primer tipo que se estaría dispuesto a admitir (en porcentaje).*

Ejemplo

Se desea evaluar un nuevo tratamiento respecto de una terapia estándar y se sabe, por estudios previos, que las tasas de supervivencia a 5 años pudieran no estar lejos, respectivamente, del 35% y del 20%. ¿Cuántos pacientes deben seleccionarse para comparar los dos tratamientos con un nivel de confianza del 95% si se espera *perder* en el seguimiento a un 7% de los pacientes? Considérense diferentes potencias: 80, 85 y 90%.

El programa indica que tanto en uno como en el otro grupo, habría que tomar:

128 pacientes en cada tratamiento si se desea tener una potencia de 80%

146 pacientes en cada tratamiento si se desea tener una potencia de 85%

171 pacientes en cada tratamiento si se desea tener una potencia de 90%

3.1.2.9. Coeficiente de correlación

En Epidat 4.0, el cálculo del tamaño de muestra (o potencia) para evaluar si un coeficiente de correlación (ρ) es diferente de cero asume que los datos siguen una distribución normal y se basa en un procedimiento iterativo descrito en Machin y cols [21]. Además del nivel de confianza y la potencia, el único dato necesario es el valor que se anticipa para el coeficiente de correlación, que debe ser un valor entre -1 y 1.

Tanto si se selecciona la opción de calcular tamaños de muestra como la de calcular potencias, Epidat 4.0 presenta los resultados correspondientes a dos tipos de contrastes, el contraste unilateral y el bilateral.

Aunque el contraste más comúnmente usado en la práctica es el contraste bilateral (en este caso, con hipótesis nula $H_0: \rho=0$ e hipótesis alternativa $H_1: \rho \neq 0$), también se puede estar interesado en realizar un contraste unilateral que se basa en una de las colas de la distribución del estadístico de contraste, a diferencia del bilateral que tiene en cuenta las dos colas. En este caso, el contraste unilateral puede plantearse de dos formas:

$$\begin{array}{l} H_0 : \rho \leq 0 \quad H_0 : \rho \geq 0 \\ H_1 : \rho > 0 \quad \text{ó} \quad H_1 : \rho < 0 \end{array}$$

Por ejemplo, cuando se estudia la presión arterial sistólica y diastólica se sabe que éstas están asociadas positivamente, en este caso se podría estar más interesado en realizar un contraste de una cola (contraste unilateral con hipótesis nula $H_0: \rho \leq 0$ e hipótesis alternativa $H_1: \rho > 0$) que uno de dos colas (contraste bilateral).

Ejemplo

Supóngase que la correlación entre el volumen espiratorio forzado en un segundo y la capacidad vital forzada en individuos sanos es aproximadamente de 0,60. Supóngase, adicionalmente, que un grupo de pacientes con una enfermedad de pulmón está accesible en una clínica, e interesa contrastar si hay correlación entre ambas medidas en esos pacientes. Con un nivel de confianza del 95% y una potencia del 90%, el programa comunica que se necesitan 21 sujetos en la muestra cuando el contraste es unilateral y 25 cuando es bilateral.

Resultados con Epidat 4.0:

Datos:		
Coefficiente de correlación a detectar:	0,600	
Nivel de confianza:	95,0%	
Resultados:		
	Tamaño de la muestra	
Potencia (%)	Unilateral	Bilateral
90,0	21	25

3.2. Selección muestral

3.2.0. Conceptos generales

Muchas razones prácticas pueden mencionarse para justificar el uso del muestreo. Sin embargo, la fundamental es de naturaleza económica; por lo general, resulta literalmente imposible estudiar a todos los individuos que integran una población (hacer un censo). El imperativo de hallar recursos metodológicos que fuesen práctica y económicamente viables, pero a la vez científicamente válidos, es de muy larga data. Sin embargo, no fue hasta la década de los 30, con un célebre trabajo presentado por el polaco Jerzy Neyman ante la *Royal Statistical Society*, cuando se comenzó a desarrollar una teoría orgánica sobre las técnicas que permiten superar de manera científica esa dificultad [27].

Un paso de medular relevancia en el progreso histórico de consolidación de la teoría del muestreo fue dado con la introducción del azar en el procedimiento de selección muestral. El azar puede intervenir de muchos modos en el proceso, pero se considera que se ha conferido un carácter estadísticamente riguroso al método de selección sólo cuando se cumplen dos condiciones: que el azar opere de manera tal que el procedimiento otorgue a todos y cada uno de los elementos de la población cierta probabilidad *conocida* de antemano de integrar la muestra, y que tal probabilidad no sea nula para elemento alguno. En tal caso, se dice que se ha seguido un *método probabilístico* de muestreo. Los procedimientos probabilísticos satisfacen la exigencia intuitiva de eliminar -o mitigar al menos- la carga subjetiva que podría influir en la elección de los elementos que se van a examinar, interrogar o medir (y, por ende, en las conclusiones). Por otra parte, sólo cuando el método es probabilístico se podrá medir el grado de precisión con que se realizan las estimaciones, que es una de las tareas fundamentales del muestreo.

Si la probabilidad de selección es la misma para todos los sujetos de la población, entonces el procedimiento es *equiprobabilístico*. La propiedad más importante de los diseños muestrales que tienen esta última condición consiste en que, para hacer estimaciones puntuales, no es necesario emplear ponderaciones que corrijan los posibles sesgos introducidos por el hecho de que unos sujetos tengan más oportunidad de integrar la muestra que otros. Esto quiere decir que, por ejemplo, para estimar la media poblacional (o un porcentaje), si el diseño fue equiprobabilístico, simplemente se utiliza la media muestral (o el porcentaje muestral).

En algunas ocasiones se cuenta con listas detalladas de la población que se quiere estudiar, y el proceso de selección se realiza entonces directamente haciendo uso de tales listas. Con más frecuencia, sin embargo, no se dispone de una lista completa de unidades de análisis. Cuando se da esa situación se suele emplear el llamado *muestreo en etapas* o *por conglomerados*. Este procedimiento, sin embargo, puede emplearse aun cuando se disponga de una lista con vistas a ahorrar recursos en la fase de terreno. Su estructura consiste en lo siguiente: antes de seleccionar la muestra, la población se divide en *unidades de primera etapa (UPE)*, cada una de las cuales contiene un conjunto de varias unidades de análisis. En cualquier caso, las unidades de muestreo de primera etapa deben cubrir por entero a la población sin intersectarse; o sea, todo miembro de la población o *unidad de análisis* pertenecerá a una y sólo una UPE.

Si se admiten en la muestra a todas las unidades de análisis que integran las UPE elegidas, se ha optado por un *muestreo por conglomerados monoetápico*. Si, luego de seleccionadas aquellas unidades de primera etapa que aportarán elementos a la muestra, dentro de cada una de ellas se eligen algunas unidades de análisis, entonces se está ante un *muestreo por conglomerados bietápico*. Epidat 4.0 contempla sólo esas dos modalidades en cuanto al número de etapas, pero en la práctica (especialmente para poblaciones muy grandes) se acude ocasionalmente a diseños con 3 o más etapas (por ejemplo, si se eligen 10 escuelas en una

ciudad, 2 grupos en cada escuela elegida y 10 alumnos en cada grupo, se habrán empleado tres etapas para seleccionar una muestra de 200 alumnos de la ciudad). Para tales situaciones Epidat 4.0 puede ser útil, pero hay que emplearlo más de una vez y creativamente en dependencia del diseño de que se trate.

Se llama *marco muestral* al conjunto de listas de todas las unidades de muestreo necesarias para llevar adelante el proceso. El conjunto de listas imprescindibles para la selección total (en el ejemplo: una lista de escuelas primero, después una lista para cada uno de los grupos que tiene cada escuela seleccionada, y finalmente una lista de niños dentro de cada uno de los grupos así elegidos) constituye el *marco muestral* del estudio.

Como ya se señaló, en la práctica se dan dos situaciones en relación al marco muestral:

- a) se dispone de todo el marco al comienzo del estudio
- b) el marco se va conformando en la medida que el muestreo avanza

Las diversas alternativas de selección que maneja Epidat 4.0 contemplan las dos posibilidades, siempre suponiendo que en el primero de estos casos el marco está recogido en una base de datos creada en un archivo de OpenOffice o de Excel.

Epidat 4.0 incluye los siete procedimientos de selección de muestras más importantes y brinda los recursos necesarios para aplicarlos a marcos muestrales reales. Los procedimientos incluidos son:

- 3.2.1. Muestreo simple aleatorio
- 3.2.2. Muestreo sistemático en fases
- 3.2.3. Muestreo aleatorio estratificado
- 3.2.4. Muestreo por conglomerados monoetápico
- 3.2.5. Muestreo por conglomerados bietápico
- 3.2.6. Muestreo por conglomerados monoetápico estratificado
- 3.2.7. Muestreo por conglomerados bietápico estratificado

Los tres primeros corresponden a una selección directa sobre una lista de unidades de análisis y los restantes cuatro son modalidades del muestreo en etapas. A continuación se explica cómo proceder para seleccionar muestras en cada caso.

Para ilustrar los procedimientos de selección incluidos en Epidat 4.0 se empleará un único libro Excel llamado INSTITUTO.xls, que contiene varias hojas. La primera de dichas hojas se denomina *Alumnado* y contiene 966 artículos o registros. Es importante comprender qué es lo que dicha base supuestamente representa. Se trata de un instituto de Enseñanza Media donde están registrados 966 alumnos por orden alfabético.

Los alumnos del Instituto se ubican en uno de 5 cursos: PRIMERO - SEGUNDO - TERCERO - CUARTO - QUINTO; dentro de cada curso, hay varias aulas (7, 4, 3, 3, 4 respectivamente). Para cada alumno está registrada la edad (entre 15 y 19 años). Los campos de la base contenida en la primera hoja son entonces los siguientes:

CAMPO	TIPO	VALORES POSIBLES
CODIGO	Numérica	Enteros de 1 a 966
NOMBRE	De texto	Texto
EDAD	Numérica	Enteros de 15 a 19
CURSO	Nominal	Primero - Segundo - Tercero - Cuarto - Quinto
AULA	Nominal	1.1, 1.2, ..., 7.3, 7.4

La distribución de los 966 alumnos registrados según aulas y cursos es la siguiente:

CURSO	AULA							Total
	1	2	3	4	5	6	7	
Primero	30	29	28	37	33	54	56	267
Segundo	62	43	64	48				217
Tercero	37	58	56					151
Cuarto	45	51	33					129
Quinto	48	47	86	21				202
Total	222	228	267	106	33	54	56	966

3.2.1. Muestreo simple aleatorio (MSA)

El muestreo simple aleatorio (MSA) es el procedimiento probabilístico de selección de muestras más sencillo y conocido. Es muy útil para obtener muestras de poblaciones pequeñas, pero no suele emplearse directamente para seleccionar muestras a partir de listas de unidades de análisis cuando las poblaciones son muy grandes (por ejemplo, de varios miles), debido tanto a la dificultad o imposibilidad de contar con tales listas como a razones económicas. En estas situaciones, el muestreo en etapas es mucho más recomendable.

El MSA se utiliza ampliamente, sin embargo, en los estudios experimentales, en lo que constituye el más importante punto de contacto de la teoría de muestreo en poblaciones finitas con la estadística inferencial clásica. Pero su mayor relevancia la debe al hecho de ser un procedimiento básico como componente de métodos más complejos (muestreo estratificado y en etapas).

Cuando se emplea de manera directa sobre un conjunto de unidades, el MSA otorga la misma probabilidad de ser elegidos a todos los subconjuntos posibles de ese conjunto que tengan un tamaño prefijado, de lo cual se deduce la propiedad de ser equiprobabilístico; es decir, de otorgar igual probabilidad de selección a todas las unidades de análisis que integran la población. Dicha probabilidad es el cociente entre el tamaño de la muestra y el tamaño de la población ($p=n/N$) y Epidat 4.0 presenta su valor en la ventana de resultados junto a los números seleccionados.

El muestreo simple aleatorio sólo exige dos datos para proceder: el tamaño poblacional y el de la muestra deseada.

Tamaño poblacional: Entero mayor que 1 que identifica el total de unidades de análisis pertenecientes al universo o población.

Tamaño de muestra: Entero mayor que cero y menor que el valor del tamaño poblacional. Indica el número de unidades de análisis que se tomarán de la población para integrar la muestra.

Cuando se cuenta con un marco muestral ya volcado en una base de datos, debe comunicarse el nombre y la ubicación del archivo que la contiene. En este caso, el programa identifica el *Tamaño poblacional* automáticamente y este campo se inhabilita, ya que el programa asume que dicho tamaño coincide con el número de artículos contenidos en el archivo.

En cualquiera de las dos situaciones, tanto si se introduce manualmente el tamaño de la población como si se cargan los datos a partir de un archivo, es posible guardar la muestra resultante de la selección en un archivo con formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods). En el primer caso, dicho archivo solo contendrá el listado de números correspondientes a las unidades de análisis seleccionadas; en el segundo caso, se guardará un archivo con la misma estructura del original y con todos los datos que éste contiene para cada una de las unidades de análisis que hayan resultado seleccionadas.

Ejemplo

Para ilustrar el empleo del MSA en Epidat 4.0, considérense las dos posibilidades previstas: el caso en que se cuenta con un marco muestral ya volcado en una base de datos, y aquel en que sólo se conoce el tamaño del futuro listado con que se va a operar (es decir, el tamaño poblacional).

Supóngase que se quiere obtener una MSA de 20 estudiantes de un instituto que tiene 966 alumnos registrados. Si no se contara con una base de datos donde se identifique a los 966 alumnos, simplemente se comunicarían esos dos números al programa y éste devuelve la identificación genérica de los 20 sujetos (entre 1 y 966) resultantes de la selección (es decir, simplemente el número que corresponde a cada cual), tal y como se observa a continuación:

Datos:

Tamaño de la población: 966
Tamaño de la muestra: 20

Número de los sujetos seleccionados:

838	396	530	488	547	727	134
740	771	701	869	922	750	950
49	273	873	82	697	625	

Probabilidad de selección: 2,0704%

Nota: Obviamente, si el usuario realiza esta operación, los 20 números que obtendrá no serán los que figuran en el cuadro precedente sino otros, ya que el programa hace una selección diferente en cada ocasión.

Indicando la hoja *Alumnado* dentro del archivo INSTITUTO.xls como fuente de datos para la selección de 20 alumnos, se obtiene un resultado similar al del caso anterior, pero ahora se tiene la posibilidad de guardar la muestra seleccionada en un archivo que contendrá para cada registro (cada estudiante) la misma información incluida en la hoja *Alumnado* de INSTITUTO.xls, pero sólo para los 20 sujetos elegidos, lo cual produciría una sub-base de datos tal y como se aprecia a continuación:

CODIGO	NOMBRE	EDAD	CURSO	AULA
743	MARIANO ARNAU	15	PRIMERO	1,4
863	RAFAEL JIMÉNEZ	16	SEGUNDO	2,3
933	SONIA MIRANDA	19	QUINTO	5,1
172	CELINA MACHIN	17	TERCERO	3,1
497	JESUS MENENDEZ	17	TERCERO	3,2
423	ISABEL AMBOS	15	PRIMERO	1,4
406	INDIRA RODRÍGUEZ	15	PRIMERO	1,2
927	SILVIA PLASENCIA	19	QUINTO	5,2
25	AGUSTIN JIMÉNEZ	17	TERCERO	3,3
759	MARLENE BAONZA	19	QUINTO	5,3
615	JOSÉMI MARTINEZ	19	QUINTO	5,2
422	ISABEL ALFONSO	17	TERCERO	3,2
914	ROSA RIERA	15	PRIMERO	1,2
346	HEBE RAMOS	19	QUINTO	5,1
851	PILAR PÉREZ	15	PRIMERO	1,3
71	ANAMARÍA POSADA	19	QUINTO	5,1
413	IRA AGUIRRE	15	PRIMERO	1,4
764	MARY CARBONELL	16	SEGUNDO	2,4
334	GRISEL NOGUERAS	19	QUINTO	5,1
136	BICHITO CASTAÑO	17	SEGUNDO	2,4

3.2.2. Muestreo sistemático (MS)

El muestreo sistemático (MS) es un procedimiento alternativo al muestreo simple aleatorio, que se usa para la selección de muestras equiprobabilísticas de una población organizada según cierto orden conocido. Una importante ventaja del MS sobre el MSA consiste en que no es estrictamente necesario contar con un marco *explícito* de todas las unidades de análisis de la población (podría aplicarse, por ejemplo, a los usuarios de un servicio de urgencias en la medida que van arribando a éste a lo largo de cierto período). Por conducto de este procedimiento, la muestra “recorre” toda la lista o población (por ejemplo, se extiende por todo el lapso durante el cual arriban pacientes al servicio).

En función del método empleado para la selección de la muestra, el MS se denomina muestreo sistemático *regular* o muestreo sistemático *en fases* según se fije, respectivamente, el tamaño de la muestra que se desea seleccionar o el porcentaje de población que debe quedar incluido en la muestra (equivalentemente, la probabilidad de selección). En ambos casos, al igual que ocurre en el MSA, se otorga igual probabilidad de integrar la muestra a todas las unidades de análisis de la población. En el caso del MS regular, dicha probabilidad es el cociente entre el tamaño de la muestra y el tamaño de la población ($p=n/N$) y Epidat 4.0 presenta su valor en la ventana de resultados junto a los números seleccionados.

El muestreo sistemático regular se basa en seleccionar n elementos de un listado de tamaño N a intervalos fijos de selección ($k=N/n$), comenzando por el arranque aleatorio r (valor

comprendido entre 1 y k). El muestreo sistemático en fases es una variante del muestreo sistemático convencional formalizada hace relativamente pocos años [28] y que consiste en seleccionar un porcentaje p de elementos de un listado de tamaño N . Los detalles de la forma en que opera cada uno de los métodos (especialmente, en lo que concierne a la definición de arranques para cada fase en el muestreo sistemático en fases) pueden hallarse en Silva [3].

El muestreo sistemático sólo exige dos datos para proceder: el tamaño poblacional, y el tamaño de la muestra o la probabilidad de selección (que también puede entenderse como el porcentaje de este total poblacional que se quiere obtener como tamaño de muestra), según se opte por el muestreo sistemático regular o por el muestreo sistemático en fases, respectivamente.

Tamaño poblacional: Entero mayor que uno que identifica el total de unidades de análisis pertenecientes al universo o población.

Tamaño de la muestra: Entero mayor que cero y menor que el tamaño poblacional que identifica el número de unidades de análisis que se tomarán de la población para integrar la muestra.

Probabilidad de selección: Es la probabilidad común que se quiere otorgar a cada elemento de la población y que a la vez representa la fracción del tamaño poblacional que quedará en la muestra. Para facilitar el trabajo con este campo, se define como un valor entero mayor que cero y menor que 100 (o sea, en rigor, el valor que ha de teclearse representa la probabilidad de selección multiplicada por 100); con esta información Epidat 4.0 determina los intervalos de selección (número que fija uno de cada cuántos han de tomarse) y arranques (valores a partir de los cuales se comienzan a aplicar dichos intervalos) que deben emplearse para aplicar el procedimiento.

Cuando se cuenta con un marco muestral ya volcado en una base de datos, debe comunicarse el nombre y la ubicación del archivo que la contiene. En este caso, el programa identifica el *Tamaño poblacional* automáticamente y este campo se inhabilita, ya que el programa asume que dicho tamaño coincide con el número de artículos contenidos en el archivo.

En cualquiera de las dos situaciones, tanto si se introduce manualmente el tamaño de la población como si se cargan los datos a partir de un archivo, es posible guardar la muestra resultante de la selección en un archivo con formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods). En el primer caso, dicho archivo solo contendrá el listado de números correspondientes a las unidades de análisis seleccionadas; en el segundo caso, se guardará un archivo con la misma estructura del original y con todos los datos que éste contiene para cada una de las unidades de análisis que hayan resultado seleccionadas.

Ejemplo

Para ilustrar el empleo del muestreo sistemático en Epidat 4.0, considérese las dos posibilidades previstas: el caso en que se cuenta con un marco muestral ya volcado en una base de datos y aquél en que sólo se conoce el tamaño del futuro listado con que se va a operar (es decir, el tamaño poblacional).

Supóngase que se quiere obtener una muestra sistemática en fases que contenga aproximadamente al 11% de los estudiantes de un centro universitario que tiene 966 alumnos registrados. Si no se contara con un listado en una base de datos, simplemente se comunicarían esos dos números al programa y éste devuelve alrededor de 106 números entre 1 y 966 resultantes de la selección.

Nota: Adviértase que en el muestreo sistemático la cifra de sujetos elegidos no siempre será exactamente igual para diferentes aplicaciones del programa aunque se introduzcan los mismos datos iniciales; puede producirse una ligera variación que es propia del método que se emplea, dependiente de la selección aleatoria que se realice; en este ejemplo concreto, dicho número puede ser 106 o 107, pues el 11% de 966 es un número no entero (106,26).

En el caso en que no usamos base de datos alguna, Epidat 4.0 produce un resultado como el siguiente:

Datos:

Tipo de muestreo: Sistemático en fases
 Tamaño de la población: 966
 Probabilidad de selección: 11%

Resultados:

Intervalos de selección(k1 y k2)		Arranques aleatorios(r1 y r2)	
11	48	5	3

Número de los sujetos seleccionados:

3	5	16	27	38	49	56
60	71	82	93	104	109	115
126	137	148	159	162	170	181
192	203	214	215	225	236	247
258	267	269	280	291	302	313
320	324	335	346	357	368	373
379	390	401	412	423	426	434
445	456	467	478	479	489	500
511	522	531	533	544	555	566
577	584	588	599	610	621	632
637	643	654	665	676	687	690
698	709	720	731	742	743	753
764	775	786	795	797	808	819
830	841	848	852	863	874	885
896	901	907	918	929	940	951
954	962					

Tamaño de muestra resultante: 107

Al aplicar Epidat 4.0 en esta situación, el programa ha seleccionado sistemáticamente 1 de cada 11 sujetos (empezando con el quinto sujeto como arranque) y, de los no seleccionados, añade uno de cada 48 (empezando con el arranque 3 en esa segunda fase).

Nota: Epidat 4.0 identifica los intervalos de selección (11 y 48 en este caso) y elige aleatoriamente los arranques (5 y 3 en este ejemplo). Cada vez que se emplee, elegirá arranques diferentes.

Cuando se cuenta con un marco muestral volcado en una base de datos, debe comunicarse el nombre y la ubicación de dicha base.

La hoja *Alumnado* dentro del archivo INSTITUTO.xls que se describe en el apartado 3.2.0 sirve a los efectos de ilustración del modo en que se usa una fuente de datos para la selección de 120 alumnos. Si se guarda la muestra en un archivo, éste contendrá las variables de la base original pero solo para la selección del 11% de los alumnos, es decir, sólo para los sujetos que resulten elegidos.

3.2.3. Muestreo aleatorio estratificado (MAE)

Este diseño muestral se utiliza cuando la población se puede dividir en diferentes grupos de elementos (estratos) cuya representación en la muestra quisiera asegurarse. La manera natural de lograrlo es construir listas separadas para cada uno de los estratos y proceder a seleccionar submuestras en cada uno de ellos. Con este diseño lo que se procura es obtener una muestra que tenga en sí una variabilidad similar a la de la población; lo ideal sería conseguir que los subconjuntos que se definen como estratos fuesen internamente homogéneos y diferentes entre sí.

Dentro de cada uno de los estratos, la muestra de unidades de análisis es seleccionada por un procedimiento probabilístico; Epidat 4.0 siempre usa muestreo simple aleatorio para ello.

Como en otros diseños, hay en principio dos situaciones posibles:

- a) Cuando se dispone de una base de datos que contiene el marco muestral detallado, es decir, cada registro corresponde a una unidad de análisis de la que se conoce a que estrato pertenece. En este primer caso no es necesario informar del número de unidades existentes en cada estrato, pues Epidat 4.0 lo calcula automáticamente una vez que uno comunica al programa cuál es el campo que identifica los estratos.
- b) Cuando la única información disponible es el número de estratos y el tamaño de cada uno, información que hay que proveer al programa; en este caso, hay dos alternativas:
 - Mediante la lectura de un archivo de OpenOffice o de Excel donde figure esta información. Dicha base ha de tener tantos registros como estratos haya, y una variable con la identificación o nombre de los estratos, otra columna correspondiente a los tamaños de éstos y, opcionalmente, una tercera donde figuren los tamaños muestrales deseados.
 - Entrando de manera enteramente manual los tamaños de los estratos.

Para ambas alternativas se ofrecen tres posibilidades para definir los tamaños de muestra de los estratos:

1. Ser idénticos para todos ellos,
2. Ser libremente determinados por el usuario,
3. Quedar establecidos mediante una asignación proporcional de un tamaño muestral general. En tal caso el propio programa se ocupa de dividir el tamaño muestral en tantos sumandos como estratos haya y de manera tal que la razón entre el tamaño muestral dentro de un estrato y el tamaño de todo el estrato sea la misma para todos los estratos. La muestra así elegida resulta ser equiprobabilística.

Además de presentar los resultados en pantalla, Epidat 4.0 permite guardar la muestra resultante del procesamiento en un archivo con formato de Excel (*.xls, *.xlsx) o de

OpenOffice (*.ods) cuya estructura depende de la variante elegida para realizar la entrada de datos:

- Entrada manual del número de estratos y sus tamaños: el archivo de resultados contendrá el listado de los números correspondientes a las unidades seleccionadas y el estrato al que pertenece cada una.
- Entrada automática de datos agregados: el archivo de resultados contendrá el listado de los números correspondientes a las unidades seleccionadas en cada estrato y toda la información relativa a los estratos que estaba contenida en el archivo original.
- Entrada automática de datos individuales: la información se salva en un archivo con la misma estructura del original y con todos los datos de cada una de las unidades de análisis seleccionadas.

En los tres casos el archivo de resultados incluye, además, la probabilidad de selección común a todas las unidades de un mismo estrato (cociente entre el tamaño de la muestra y el de la población) y la ponderación que les corresponde en la muestra, que es el inverso de la probabilidad de selección y se interpreta como el número de elementos de la población que representa cada unidad de la muestra.

Ejemplo

Supóngase que se quiere obtener una muestra de 120 estudiantes del Instituto cuyas características se han descrito al final del apartado 3.2.0. Dicho Instituto tiene 966 alumnos registrados, distribuidos en cinco cursos con los siguientes tamaños:

ESTRATO (curso)	TAMAÑO
PRIMERO	267
SEGUNDO	217
TERCERO	151
CUARTO	129
QUINTO	202

Si se cuenta con la información detallada (por ejemplo, la que figura en la hoja *Alumnado* dentro del libro INSTITUTO.xls), se marca *Abrir datos individuales* y se carga dicha hoja. Se señala entonces el campo que corresponde a los estratos (en este ejemplo, el campo CURSO) y se comunican los tamaños deseados para las muestras en los estratos según las alternativas (I, II y III) anteriormente enumeradas.

Si se desea que la muestra contemple distintos cursos como estratos pero no se cuenta con el marco detallado, se optará por la variante *Abrir datos agregados* y se emplea la hoja llamada MAE, la cual consta de solo cinco registros, uno por cada estrato, con un campo numérico llamado CURSO que contiene los 5 nombres (PRIMERO, SEGUNDO, etc), otro campo que recoge su tamaño y un tercer campo donde figuran los tamaños muestrales deseados para cada estrato. Al aplicar Epidat 4.0, el programa informa los sujetos seleccionados, así como las probabilidades de selección que tuvieron y las ponderaciones que le corresponden (diferentes para cada estrato si la asignación de tamaños no fue proporcional).

Nota: Epidat 4.0 realiza redondeos que pueden ocasionalmente producir que el tamaño muestral global no coincida con el previsto. Naturalmente, en la práctica se usarían los tamaños redondeados aunque la suma no sea igual al tamaño de muestra recabado (hecho que virtualmente nunca tiene importancia práctica alguna, ya que la diferencia diferirá a lo sumo en una unidad).

3.2.4. Muestreo por conglomerados monoetápico

Típicamente, se opta por este método de muestreo cuando no se pueda disponer de un listado de unidades de análisis y/o cuando se quieren eludir algunos problemas prácticos que supondría el empleo de muestreo simple aleatorio o el muestreo sistemático, tales como la dispersión geográfica de las unidades de análisis a lo largo del territorio en que se halla ubicada la población.

Una solución a este problema sería la selección de una muestra por conglomerados monoetápica: se trata de dividir la población en cierto número de partes o conglomerados, que se consideran como UPE, e incluir en la muestra a *todas* las unidades de análisis pertenecientes a los conglomerados que se elijan.

Epidat 4.0 selecciona los conglomerados mediante un MSA, lo que da lugar a una muestra equiprobabilística, y la probabilidad de selección común a todos los individuos es el cociente entre el número de UPE en la muestra y en la población (m/M). El programa incluye esta probabilidad en los resultados que muestra en pantalla.

Este procedimiento también puede aplicarse en dos posibles situaciones. Estas son:

- a) Cuando se cuenta con un archivo que contiene una base de datos donde cada uno de sus registros corresponde a una unidad de análisis (marco muestral detallado) de la que se conoce a qué UPE o conglomerado pertenece.
- b) Cuando sólo se tienen identificados los conglomerados y sus tamaños; en este caso se puede teclear la información (número de conglomerados y tamaños) manualmente, o leer esta información de un archivo, el cual ha de contener tantos registros como conglomerados, con una variable que identifique el conglomerado y otra que contenga su tamaño (número de unidades de que consta).

Nota: Cuando lo único disponible es el listado de conglomerados (por ejemplo, una lista donde figuran 45 escuelas con sus nombres y direcciones) pero no se conocen sus tamaños (número de alumnos por escuela), la selección de conglomerados (escuelas) debe realizarse usando alguno de los procedimientos directos de selección (muestreo simple aleatorio de las UPE, o muestreo sistemático o muestreo aleatorio estratificado) y en consecuencia, apelar a alguno de los módulos previamente explicados. El único aspecto que cabe advertir en este caso es que el usuario debe tomar un número de conglomerados aproximadamente igual a la razón entre el tamaño de muestra y el número medio de unidades de análisis por conglomerado. Es posible que este último dato no se conozca, pero casi siempre se tiene una idea aproximada del tamaño poblacional y, por ende, también del promedio por UPE.

En los dos casos de entrada automática (datos individuales o datos agregados) se ha de informar a Epidat 4.0 el archivo, en formato de Excel o de OpenOffice, en que se halla la información y cuál es la variable que identifica los conglomerados.

En el caso de que se lea un archivo de datos agregados, hay que indicar además la variable que identifica el tamaño de los conglomerados, es decir, el número de unidades de análisis que contiene cada uno de éstos. En caso de tener un marco muestral completo en la base de datos, esta información no es necesaria, pues Epidat 4.0 se ocupa de “contar” el número de sujetos en cada UPE.

Nota: Debe advertirse que, al examinar la base de datos, el sistema considera como pertenecientes a un mismo conglomerado a todos aquellos registros que tengan un mismo código en el campo que identifica al conglomerado. Esto quiere

decir, por ejemplo, que todos los registros para los cuales la variable correspondiente al conglomerado tenga el código 3, serán considerados como integrantes del tercer conglomerado de la población. Si la base de datos estuviera conformada de manera tal que se emplea un mismo número para identificar a conglomerados diferentes debido, por ejemplo, a que éstos pertenecen a zonas diferentes o a distintas unidades de muestreo de mayor jerarquía, el sistema “entenderá” indebidamente que se trata de registros de un mismo conglomerado. Consecuentemente, para evitar esta posible dificultad, es conveniente atribuir siempre un código diferente a cada conglomerado de la población. Por ejemplo, si se tuviera dos manzanas, una con 42 y otra con 23 viviendas (siendo éstas los conglomerados), en el campo correspondiente a estos últimos podrían usarse 65 números diferentes para identificarlos (por ejemplo, los números del 1 al 65) en lugar de emplear los números 1, 2, ..., 42 para la primera manzana y 1, 2, ..., 23 para la segunda.

Se comunicará además el *Tamaño de muestra*: Valor entero mayor que cero y menor que el tamaño poblacional (si los datos son agrupados, el tamaño poblacional no es el número de registros de la base, sino la suma de la variable que contiene los tamaños de los conglomerados). Indica el número de unidades de análisis que aproximadamente se tomarán de la población para integrar la muestra. El programa selecciona el número de conglomerados necesarios para alcanzar aproximadamente dicho tamaño. Alternativamente, puede solicitarse que el programa seleccione un número de conglomerados fijo en lugar de un número de unidades de análisis.

Además de presentar los resultados en pantalla, Epidat 4.0 permite guardar la muestra resultante del procesamiento en un archivo con formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods) cuya estructura depende de la variante elegida para realizar la entrada de datos:

- Entrada manual del número de conglomerados y sus tamaños: el archivo de resultados contendrá el listado de números correspondientes a los conglomerados seleccionados.
- Entrada automática de datos agregados: el archivo de resultados contendrá el listado de los conglomerados seleccionados y toda la información de cada uno de ellos contenida en el archivo original.
- Entrada automática de datos individuales: la información se salva en un archivo con la misma estructura del original y con todos los datos de cada una de las unidades de análisis pertenecientes a los conglomerados seleccionados.

Ejemplo

Supóngase que se quiere obtener una muestra de aproximadamente 120 estudiantes de un instituto de Enseñanza Media que tiene 966 alumnos registrados y que los alumnos están distribuidos en 21 aulas de tamaños variables y conocidos, tal y como se describió al final del apartado 3.2.0.

La hoja MC del archivo INSTITUTO.xls contiene la lista de las 21 aulas y los tamaños de ellas. Nótese que el campo que contiene los conglomerados no tiene que ser numérico (en este archivo, el campo se llama AULA, estos están identificados por los códigos que señalan de qué aula se trata dentro de cada curso). Se teclea el tamaño muestral deseado (120 en este caso) y Epidat 4.0 aplica un muestreo simple aleatorio para seleccionar 3 de los 21 grupos. Este número es el resultado de dividir el tamaño de la muestra (120) por el número medio de unidades de análisis de las UPE ($966/21=46$) y luego “redondear”: $120/46=2,6\approx 3$. Epidat 4.0

siempre tomará el entero más pequeño que supere o iguale a esa razón (en el ejemplo previo el menor número entero mayor que 2,6 es 3).

El resultado podría ser similar al siguiente:

Entrada automática (Datos agregados):

Archivo de trabajo: C:\Archivos de programa\Epidat 4.0\Ejemplos\Muestreo\INSTITUTO.xls
 Tabla: MC
 Variables:
 Conglomerados: AULA
 Tamaño de los conglomerados: ALUMNOS

Datos:

	Población	Muestra
Tamaño	966	120
Nº de conglomerados	21	3

Conglomerados seleccionados:

	Tamaño
1.3	28
2.2	43
4.3	33
TOTAL	104

Probabilidad de selección: 14,2857%

La muestra estará conformada por los sujetos que pertenezcan a estos 3 conglomerados, 104 en total (ya que los tamaños respectivos son 28, 43 y 33). Nótese que el tamaño muestral verdadero raramente coincidirá con el deseado (120 en este ejemplo). Este es un rasgo inherente al muestreo monoetápico. Nótese que tampoco en la variante de que se cuente con una hoja que contenga la información completa (listado de los 966 sujetos y, para cada cual, el código del grupo al que pertenece), el tamaño muestral esperado (deseado) coincide con el realmente obtenido. En un diseño como éste, el tamaño muestral es aleatorio, ya que depende de cuáles sean los conglomerados que a la postre resulten seleccionados.

3.2.5. Muestreo por conglomerados bietápico

Este procedimiento de selección muestral se utiliza cuando hay gran variabilidad entre los tamaños de los conglomerados, o cuando el tamaño poblacional es demasiado grande, casos en que no procede emplear muestreo por conglomerados monoetápico.

La población se divide inicialmente en conglomerados, que constituyen las unidades de primera etapa (UPE), parte de las cuales son seleccionadas para ser entonces objeto de subselección (cada una de ellas).

Este procedimiento también puede aplicarse según dos alternativas:

- Quando solo se dispone de listas de conglomerados y sus tamaños (y ocasionalmente, el tamaño muestral que se aplicaría en cada uno de los conglomerados en caso de que fuera seleccionado en la primera etapa). En esta situación es posible introducir los datos de forma manual o mediante un archivo que contenga la información necesaria, es decir,

cada registro de la base de datos debe corresponder a un conglomerado, uno de los campos de dicha base debe recoger el número de unidades de análisis de la UPE correspondiente (es decir, su tamaño) y, opcionalmente, el tamaño de muestra para dicha UPE.

- b) Cuando se cuenta con una base de datos en la que cada uno de sus registros corresponde a una unidad de análisis (marco muestral detallado). En este caso es posible guardar la muestra resultante del procesamiento en un archivo de datos con la misma estructura del original y con todos los datos de cada una de las unidades de análisis seleccionadas.

En los dos casos de entrada automática (datos individuales o datos agregados) se ha de indicar a Epidat 4.0 el archivo, en formato de Excel o de OpenOffice, en que se halla la información y cuál es la variable que identifica los conglomerados.

Cuando no se tiene base detallada, sino que sólo se cuenta con datos agrupados hay que indicar, además, la variable que contiene el tamaño de los conglomerados, es decir, el número de unidades de análisis que contiene cada uno. En caso de tener un marco muestral completo en la base de datos, esta información no es necesaria, pues Epidat 4.0 se ocupa de “contar” el número de sujetos en cada UPE.

Nota: Debe advertirse que, al examinar la base de datos, el sistema considera como pertenecientes a un mismo conglomerado a todos aquellos registros que tengan un mismo código en el campo que identifica al conglomerado. Esto quiere decir, por ejemplo, que todos los registros para los cuales la variable correspondiente al conglomerado tenga el código 3, serán considerados como integrantes del tercer conglomerado de la población. Si la base de datos estuviera conformada de manera tal que se emplea un mismo número para identificar a conglomerados diferentes debido, por ejemplo, a que éstos pertenecen a zonas diferentes o a distintas unidades de muestreo de mayor jerarquía, el sistema “entenderá” indebidamente que se trata de registros de un mismo conglomerado. Consecuentemente, para evitar esta posible dificultad, es necesario atribuir siempre un número diferente a cada conglomerado de la población. Por ejemplo, si se tuviera dos manzanas, una con 42 y otra con 23 viviendas (siendo éstas los conglomerados), en el campo correspondiente a estos últimos deben ponerse 65 números diferentes para identificarlos (por ejemplo, los números del 1 al 65) en lugar de emplear los números 1, 2, ..., 42 para la primera manzana y 1, 2, ..., 23 para la segunda.

Para seleccionar las UPE Epidat 4.0 ofrece dos posibilidades:

- Hacerlo mediante muestreo simple aleatorio.
- Elegirlas con probabilidad proporcional a sus tamaños.

Por su parte, los sujetos se seleccionan en cada UPE mediante un MSA, pero hay tres posibilidades para definir el tamaño de la muestra en cada UPE:

- Tomar un número igual de unidades de análisis en todos los conglomerados.
- Seleccionar una fracción fija del tamaño de los conglomerados, por ejemplo, el 10%.
- Definir el tamaño muestral de cada conglomerado.

Si se toma un número igual de unidades de análisis para todos los conglomerados, se deben introducir dos de los siguientes valores, ya que el tercero se calcula a partir de ellos:

Tamaño de muestra: Valor entero mayor que cero y no superior al producto del número de conglomerados por el tamaño del conglomerado más pequeño.

Sujetos a elegir por conglomerado: Valor entero mayor que cero y menor o igual que el tamaño del conglomerado (UPE) más pequeño con que cuente la población; identifica el número de unidades de análisis que han de seleccionarse dentro de cada uno de los conglomerados.

Número de conglomerados de la muestra: Valor entero mayor que cero y menor que el número de conglomerados (UPE) existentes.

Si se selecciona una fracción fija de unidades de análisis para todos los conglomerados, se deben proporcionar al programa los siguientes datos:

Porcentaje de muestra a seleccionar: Valor mayor que cero y menor que 100 que representa el porcentaje común que se quiere otorgar a cada conglomerado elegido.

Número de conglomerados de la muestra.

Si se define el tamaño muestral de cada conglomerado sólo es necesario indicar el *Número de conglomerados de la muestra*.

Además de presentar los resultados en pantalla, Epidat 4.0 permite guardar la muestra resultante del procesamiento en un archivo con formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods) cuya estructura depende de la variante elegida para realizar la entrada de datos:

- Entrada manual del número de conglomerados y sus tamaños: el archivo de resultados contendrá el listado de números correspondientes a los conglomerados e individuos seleccionados en las dos etapas del proceso.
- Entrada automática de datos agregados: el archivo de resultados contendrá el listado de los conglomerados e individuos seleccionados y toda la información de cada conglomerado contenida en el archivo original.
- Entrada automática de datos individuales: la información se salva en un archivo con la misma estructura del original y con todos los datos de cada una de las unidades de análisis seleccionadas.

En los tres casos el archivo de resultados incluye, además, la probabilidad de selección común a todas las unidades de un mismo conglomerado y la ponderación que les corresponde en la muestra, que es el inverso de la probabilidad de selección y se interpreta como el número de elementos de la población que representa cada unidad de la muestra.

Cuando los conglomerados se seleccionan con probabilidad proporcional a su tamaño y se define un tamaño de muestra igual para todos ellos se obtiene una muestra equiprobabilística.

Ejemplo

Supóngase que se quiere obtener una muestra de 120 estudiantes del instituto cuyas características se han descrito al final del apartado 3.2.0. Dicho instituto tiene 966 alumnos registrados, distribuidos en 21 aulas con tamaños que varían entre 21 y 86 alumnos.

La hoja MC del archivo INSTITUTO.xls contiene la lista de las 21 aulas y los tamaños de ellas. Nótese que el campo que contiene los conglomerados no tiene que ser numérico (en este archivo, el campo se llama AULA, estos están identificados por los códigos que señalan

de qué aula se trata dentro de cada curso). El campo ALUMNOS de la hoja mencionada contiene el número de sujetos existentes en cada uno de los 21 grupos.

Supongamos que se ha decidido seleccionar los 120 alumnos tomando el mismo número de sujetos de cada una de 10 aulas elegidas mediante MSA. En este caso, donde la muestra no es equiprobabilística, las probabilidades de selección (y sus inversos, las ponderaciones) varían de unos conglomerados a otros. Por ejemplo, para el aula 1.5, la probabilidad de cada uno de los 12 alumnos elegidos es 0,17316, resultante de multiplicar 10/21 (probabilidad de selección del conglomerado) por 12/33 (probabilidad de selección del alumno); consecuentemente, la ponderación para estos 12 individuos es 5,7750 (inverso de 0,17316):

Probabilidades de selección y ponderaciones:

Conglomerado	Probabilidad	Ponderación
2.3	8,9286	11,2000
3.2	9,8522	10,1500
4.1	12,6984	7,8750
2.4	11,9048	8,4000
3.3	10,2041	9,8000
1.5	17,3160	5,7750
5.1	11,9048	8,4000
4.2	11,2045	8,9250
5.3	6,6445	15,0500
1.2	19,7044	5,0750

Supóngase ahora que se cuenta con un archivo que contiene la información completa: listado de los 966 sujetos y, para cada cual, el dato del grupo al que pertenece. Cada registro puede contener o no, además, otra información sobre los individuos, que no será empleada en el acto de selección.

La hoja *Alumnado* del libro INSTITUTO.xls contiene los datos (nombre y edad) de cada uno de los 966 estudiantes y el campo AULA, donde se consigna el aula al que pertenece cada uno de ellos. Los registros están allí ordenados alfabéticamente (tal y como verosímilmente se obtendrían de un registro administrativo). Al correr el programa, se obtiene un resultado similar al del caso anterior.

3.2.6. Muestreo por conglomerados monoetápico estratificado

Típicamente, se opta por este método de muestreo cuando no se puede disponer de un listado de unidades de análisis y/o cuando se quieren eludir algunos problemas prácticos que supondría el empleo de muestreo simple aleatorio o el muestreo sistemático, tales como la dispersión geográfica de las unidades de análisis a lo largo del territorio en que se halla ubicada la población.

Una solución a este problema sería la selección de una muestra por conglomerados monoetápica estratificada: se trata de dividir la población en cierto número de partes o conglomerados, que se consideran como UPE, e incluir en la muestra a *todas* las unidades de análisis pertenecientes a los conglomerados que se elijan. Sin embargo, en esta variante, las

UPE tienen que estar inicialmente conformadas en estratos. En Epidat 4.0 la selección de los conglomerados en cada estrato se realiza mediante MSA.

Este procedimiento también puede aplicarse en dos posibles situaciones. Estas son:

- a) Cuando se cuenta con un archivo que contiene una base de datos donde cada uno de sus registros corresponde a una unidad de análisis (marco muestral detallado) de la que se conocen tanto el estrato como la UPE o conglomerado al que pertenece.
- b) Cuando sólo se tienen identificados los conglomerados dentro de cada estrato y sus tamaños; en este caso se puede teclear la información (número de estratos, conglomerados y tamaño de éstos) manualmente, o leer esta información de un archivo, el cual ha de contener tantos registros como conglomerados haya en total, con un campo que identifique el conglomerado, otro que señale a qué estrato pertenece y otro campo que contenga su tamaño (número de unidades de que consta).

Nota: Cuando lo único disponible es el listado de conglomerados (por ejemplo, una lista donde figuran 45 escuelas con sus nombres y direcciones) pero no se conocen sus tamaños (número de alumnos por escuela), la selección de conglomerados (escuelas) debe realizarse usando alguno de los procedimientos directos de selección (muestreo simple aleatorio de las UPE, o muestreo sistemático o muestreo estratificado) y en consecuencia, apelar a alguno de los módulos previamente explicados. El único aspecto que cabe advertir en este caso es que el usuario debe tomar un número de conglomerados aproximadamente igual a la razón entre el tamaño de muestra y el número medio de unidades de análisis por conglomerado. Es posible que este último dato no se conozca, pero casi siempre se tiene una idea aproximada del tamaño poblacional y, por ende, también del promedio por UPE.

En los dos casos de entrada automática (datos individuales o datos agregados) se ha de indicar a Epidat 4.0 el archivo, en formato de Excel o de OpenOffice, en que se halla la información y cuáles son las variables que identifican el estrato y el conglomerado.

Cuando no se tiene base detallada, sino que sólo se cuenta con datos agrupados hay que indicar, además, la variable que contiene el tamaño de los conglomerados, es decir, el número de unidades de análisis que pertenece a cada uno. En caso de tener un marco muestral completo en la base de datos, esta información no es necesaria, pues Epidat 4.0 se ocupa de “contar” el número de sujetos en cada UPE.

Nota: Debe advertirse que, al examinar la base de datos, el sistema considera como pertenecientes a un mismo conglomerado a todos aquellos registros que tengan un mismo código en el campo que identifica al conglomerado dentro de cada estrato. Esto quiere decir, por ejemplo, que todos los registros para los cuales la variable correspondiente al conglomerado tenga el código 3, serán considerados como integrantes del tercer conglomerado de la población en el estrato correspondiente. Si la base de datos estuviera conformada de manera tal que se emplea un mismo número para identificar a conglomerados diferentes dentro de un estrato debido, por ejemplo, a que éstos pertenecen a zonas diferentes o a distintas unidades de muestreo de mayor jerarquía, el sistema “entenderá” indebidamente que se trata de registros de un mismo conglomerado. Consecuentemente, para evitar esta posible dificultad, es conveniente atribuir siempre un código diferente a cada conglomerado de la población dentro de cada estrato. Por ejemplo, si se tuviera dos manzanas, una con 42 y otra con 23 viviendas (siendo éstas los conglomerados), en el campo correspondiente a estos últimos podrían usarse 65 números diferentes para

identificarlos (por ejemplo, los números del 1 al 65) en lugar de emplear los números 1, 2, ..., 42 para la primera manzana y 1, 2, ..., 23 para la segunda.

Al igual que en el muestreo por conglomerados monoetápico sin estratificar, en esta opción es posible definir el tamaño de muestra basándose en el número de individuos o en el número de conglomerados que han de seleccionarse. Para ambas alternativas se ofrecen tres posibilidades para definir los tamaños de muestra de los estratos:

1. Muestra igual para todos los estratos.
2. Reparto proporcional al tamaño de los estratos. En tal caso el propio programa se ocupa de dividir el tamaño muestral en tantos sumandos como estratos haya y de manera tal que la razón entre el tamaño muestral dentro de un estrato y el tamaño de todo el estrato sea la misma para todos los estratos. La muestra así elegida resulta ser equiprobabilística.
3. Definir libremente el tamaño muestral de cada estrato.

Además de presentar los resultados en pantalla, Epidat 4.0 permite guardar la muestra resultante del procesamiento en un archivo con formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods) cuya estructura depende de la variante elegida para realizar la entrada de datos:

- Entrada manual del número de estratos y conglomerados y los tamaños de éstos: el archivo de resultados contendrá el listado de números correspondientes a los conglomerados seleccionados y el estrato al que pertenecen.
- Entrada automática de datos agregados: el archivo de resultados contendrá el listado de los conglomerados seleccionados en cada estrato y toda la información de cada uno de ellos contenida en el archivo original.
- Entrada automática de datos individuales: la información se salva en un archivo con la misma estructura del original y con todos los datos de cada una de las unidades de análisis pertenecientes a los conglomerados seleccionados en cada estrato.

En los tres casos el archivo de resultados incluye, además, la probabilidad de selección común a todas las unidades de un mismo estrato y la ponderación que les corresponde en la muestra, que es el inverso de la probabilidad de selección y se interpreta como el número de elementos de la población que representa cada unidad de la muestra.

Ejemplo

Supóngase que se quiere obtener una muestra de aulas en cada curso de un instituto de Enseñanza Media que tiene 966 alumnos registrados, y que los alumnos están distribuidos en 21 aulas de tamaños variables y conocidos, tal y como se describió en el apartado 3.2.1.

La hoja *MCB* del archivo *INSTITUTO.xls* contiene la lista de las 21 aulas distribuidas en 5 cursos y los tamaños de cada una de las aulas. Nótese que el campo que contiene los conglomerados no tiene que ser numérico (en este archivo, el campo se llama *AULA*, estos están identificados por los códigos que señalan de qué aula se trata dentro de cada curso).

Si se quieren seleccionar dos aulas en cada curso y se piensa incluir en la muestra todos los alumnos de las aulas seleccionadas, hay que realizar un muestreo por conglomerados (aulas) monoetápico estratificado por cursos. En Epidat 4.0 hay que definir el tamaño de muestra en base a los conglomerados indicando que se desea una muestra igual en todos los estratos (de tamaño 2).

El resultado sería similar al siguiente:

Entrada automática (Datos agregados):

Archivo de trabajo: C:\Archivos de programa\Epidat 4.0\Ejemplos\Muestreo\INSTITUTO.xls

Tabla: MCB

Variables:

Estrato: CURSO

Conglomerado: AULA

Datos:

Tamaño de la muestra que se define: Conglomerados

Reparto de la muestra: Muestra igual para todos los estratos

Tamaño de la muestra en cada estrato: 2

	Tamaño		Nº de conglomerados	
	Población	Muestra	Población	Muestra
CUARTO	129	96	3	2
PRIMERO	267	59	7	2
QUINTO	202	134	4	2
SEGUNDO	217	91	4	2
TERCERO	151	93	3	2
Total:	966	473	21	10

Probabilidades de selección y ponderaciones:

Estrato	Probabilidad	Ponderación
CUARTO	66,6667	1,5000
PRIMERO	28,5714	3,5000
QUINTO	50,0000	2,0000
SEGUNDO	50,0000	2,0000
TERCERO	66,6667	1,5000

Número de los conglomerados seleccionados:

Estrato CUARTO:

Conglomerados	Tamaño
4.1	45
4.2	51
Total	96

Estrato PRIMERO:

Conglomerados	Tamaño
1.1	30
1.2	29
Total	59

3.2.7. Muestreo por conglomerados bietápico estratificado

Este procedimiento de selección muestral se utiliza cuando se quiere aplicar un muestreo bietápico pero habiendo separado antes las UPE según estratos. Se trata de un método cuyo uso está muy extendido en la práctica.

La población tiene que estar inicialmente conformada en estratos. En cada uno de ellos, se consideran conglomerados, que constituyen las unidades de primera etapa, de las cuales se selecciona cierto número en cada estrato. De las UPE elegidas se toman las unidades de análisis que integrarán la muestra; estas últimas son a su vez las unidades de segunda etapa.

Este procedimiento puede aplicarse en dos posibles situaciones:

- a) Cuando sólo se dispone de la lista de conglomerados y sus tamaños dentro de cada estrato; en este caso, es posible introducir los datos de forma manual o por medio de un archivo donde cada registro de la base de datos corresponda a un conglomerado. Uno de los campos de tal base debe recoger el número de unidades de análisis de la UPE correspondiente (es decir, su tamaño) y otro el estrato al que dicha UPE pertenece.
- b) Cuando se cuenta con una base de datos donde cada uno de sus registros corresponde a una unidad de análisis (marco muestral detallado). En este caso es posible guardar la muestra resultante del procesamiento en un archivo de datos con la misma estructura del original y con todos los datos de cada una de las unidades de análisis seleccionadas.

En los dos casos de entrada automática (datos individuales o datos agregados) se ha de indicar a Epidat 4.0 el archivo, en formato de Excel o de OpenOffice, en que se halla la información y cuáles son las variables que identifican los estratos y los conglomerados .

Cuando no se tiene base detallada, sino que sólo se cuenta con datos agrupados hay que indicar, además, la variable que contiene el tamaño de los conglomerados, es decir, el número de unidades de análisis que contiene cada uno. En caso de tener un marco muestral completo en la base de datos, esta información no es necesaria, pues Epidat 4.0 se ocupa de “contar” el número de sujetos en cada UPE.

Nota: Debe advertirse que, al examinar la base de datos, el sistema considera como pertenecientes a un mismo conglomerado a todos aquellos registros que tengan un mismo código en el campo que identifica al conglomerado dentro de cada estrato. Esto quiere decir, por ejemplo, que todos los registros para los cuales la variable correspondiente al conglomerado tenga el código 3, serán considerados como integrantes del tercer conglomerado de la población en el estrato correspondiente. Si la base de datos estuviera conformada de manera tal que se emplea un mismo número para identificar a conglomerados diferentes dentro de un estrato debido, por ejemplo, a que éstos pertenecen a zonas diferentes o a distintas unidades de muestreo de mayor jerarquía, el sistema “entenderá” indebidamente que se trata de registros de un mismo conglomerado. Consecuentemente, para evitar esta posible dificultad, es conveniente atribuir siempre un código diferente a cada conglomerado de la población dentro de cada estrato. Por ejemplo, si se tuviera dos manzanas, una con 42 y otra con 23 viviendas (siendo éstas los conglomerados), en el campo correspondiente a estos últimos podrían usarse 65 números diferentes para identificarlos (por ejemplo, los números del 1 al 65) en lugar de emplear los números 1, 2, ..., 42 para la primera manzana y 1, 2, ..., 23 para la segunda.

Para realizar un muestreo estratificado por conglomerados bietápico en Epidat 4.0 hay que optar por seleccionar una muestra equiprobabilística o una muestra “personalizada”.

Si se elige la opción de muestra equiprobabilística, Epidat 4.0 realiza el reparto de la muestra proporcional al tamaño de los estratos y la selección de conglomerados con probabilidad proporcional a su tamaño; además, en todos los conglomerados se selecciona el mismo número de unidades de análisis. Para definir el tamaño de muestra hay que proporcionar al programa dos de los siguientes valores, ya que el tercero se calcula a partir de ellos:

Tamaño de la muestra: Valor entero mayor que cero y no superior al producto del número de conglomerados por el tamaño del conglomerado más pequeño.

Número de conglomerados de la muestra: Valor entero mayor que cero y menor que el número de conglomerados (UPE) existentes.

Tamaño de muestra a seleccionar en cada conglomerado: Valor entero mayor que cero y menor o igual que el tamaño del conglomerado (UPE) más pequeño con que cuente la población; identifica el número de unidades de análisis que han de seleccionarse dentro de cada uno de los conglomerados.

Cuando se opta por seleccionar una muestra “personalizada”, Epidat 4.0 permite seleccionar las UPE con muestreo simple aleatorio o con probabilidad proporcional a sus tamaños dentro de cada estrato. Posteriormente, la selección de las unidades de análisis se realiza dentro de cada UPE elegida mediante muestreo simple aleatorio.

En cuanto al tamaño de muestra, hay que definir en primer lugar el número de unidades a seleccionar en cada conglomerado eligiendo entre 5 alternativas:

- Muestra igual para todos los conglomerados.
- Muestra igual para todos los conglomerados del mismo estrato.
- Fracción fija para todos los conglomerados.
- Fracción fija para todos los conglomerados del mismo estrato.
- Definir tamaño muestral de cada conglomerado.

A continuación debe definirse el número de conglomerados en la muestra y su reparto por estratos; Epidat 4.0 ofrece tres alternativas:

1. Muestra igual para todos los estratos.
2. Reparto proporcional al tamaño de los estratos. En tal caso el propio programa se ocupa de dividir el tamaño muestral en tantos sumandos como estratos haya y de manera tal que la razón entre el tamaño muestral dentro de un estrato y el tamaño de todo el estrato sea la misma para todos los estratos.
3. Definir libremente el tamaño muestral de cada estrato.

Además de presentar los resultados en pantalla, Epidat 4.0 permite guardar la muestra resultante del procesamiento en un archivo con formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods) cuya estructura depende de la variante elegida para realizar la entrada de datos:

- Entrada manual del número de estratos y conglomerados y los tamaños de éstos: el archivo de resultados contendrá el listado de números correspondientes a los conglomerados e individuos seleccionados en cada estrato en las dos etapas del proceso.
- Entrada automática de datos agregados: el archivo de resultados contendrá el listado de los conglomerados e individuos seleccionados en cada estrato y toda la información de cada conglomerado contenida en el archivo original.

- Entrada automática de datos individuales: la información se salva en un archivo con la misma estructura del original y con todos los datos de cada una de las unidades de análisis seleccionadas.

En los tres casos el archivo de resultados incluye, además, la probabilidad de selección común a todas las unidades de un mismo conglomerado y la ponderación que les corresponde en la muestra, que es el inverso de la probabilidad de selección y se interpreta como el número de elementos de la población que representa cada unidad de la muestra.

Ejemplo

Supóngase que se quiere obtener una muestra bietápica equiprobabilística de aproximadamente 120 estudiantes de un instituto que tiene 966 alumnos registrados. Tal y como se describió previamente (Sección 3.2.1) la distribución de los 966 alumnos registrados según aulas y cursos es la siguiente:

CURSO	AULA							Total
	1	2	3	4	5	6	7	
Primero	30	29	28	37	33	54	56	267
Segundo	62	43	64	48				217
Tercero	37	58	56					151
Cuarto	45	51	33					129
Quinto	48	47	86	21				202
Total	222	228	267	106	33	54	56	966

Si se quieren seleccionar 120 alumnos (tamaño de la muestra) de 6 aulas (número de conglomerados en la muestra), forzando a que sean 20 alumnos por aula, el resultado sería similar al siguiente:

Entrada automática (Datos agregados):

Archivo de trabajo: C:\Archivos de programa\Epidat 4.0\Ejemplos\Muestreo\INSTITUTO.xls

Tabla: MCB

Variables:

Estrato: CURSO

Conglomerado: AULA

Datos:

Tipo de muestreo: Equiprobabilístico

Tamaño de la muestra: 120

Número de conglomerados de la muestra: 6

Tamaño de muestra a seleccionar en cada conglomerado: 20

	Tamaño		Nº de conglomerados	
	Población	Muestra	Población	Muestra
Estrato				
CUARTO	129	20	3	1
PRIMERO	267	40	7	2
QUINTO	202	20	4	1
SEGUNDO	217	20	4	1
TERCERO	151	20	3	1
Total:	966	120	21	6

Conglomerados seleccionados para el estrato CUARTO:

Conglomerado	Tamaño	Muestra	Probabilidad	Ponderación
4.1	45	20	15,5039	6,4500
Total:	45	20		

Sujetos seleccionados para cada conglomerado:

Conglomerado 4.1:

37	20	22	35	39	18	17
28	15	4	26	11	31	38
19	9	6	10	12	29	

Conglomerados seleccionados para el estrato PRIMERO:

Conglomerado	Tamaño	Muestra	Probabilidad	Ponderación
1.1	30	20	14,9813	6,6750
1.5	33	20	14,9813	6,6750
Total:	63	40		

Sujetos seleccionados para cada conglomerado:

Conglomerado 1.1:

3.3. Asignación de sujetos a tratamientos

3.3.0. Conceptos generales

Como es bien conocido, la metodología de los ensayos clínicos controlados reposa sobre dos pilares fundamentales: el *enmascaramiento* de los tratamientos y la *asignación aleatoria* de los sujetos a las distintas alternativas terapéuticas en juego. El primero de estos recursos no es de índole estadística, sino estructural: concierne al hecho de que los pacientes involucrados no conozcan cuál de los tratamientos posibles es el que le ha correspondido [29]. Cuando esto se consigue, se dice que el estudio es "ciego".

Nota: Variantes más sofisticadas involucran otros niveles de enmascaramiento: si los que evalúan los resultados también ignoran el tratamiento que corresponde a cada paciente, el estudio es "doble ciego"; y si otro tanto ocurre, además, con los que administran el tratamiento (los propios terapeutas), se dice que es "triple ciego".

Sin embargo, lo que se relaciona directamente con el muestreo es la tarea de *asignar* los sujetos a los grupos experimentales. Los cánones universalmente aceptados exigen que tal distribución se realice *al azar*. De modo que el objetivo de este submódulo de Epidat 4.0 es facilitar la asignación de sujetos experimentales a grupos, en especial de los pacientes que tomarán parte en un ensayo clínico con varios tratamientos.

En principio, puede plantearse lo más natural: tomar un tamaño inicial y repartirlo por igual entre el número de grupos de tratamientos existentes. Sin embargo, tal estrategia puede no ser acorde con las circunstancias o restricciones reales. Ocasionalmente no se puede delimitar de antemano con exactitud cuántos pacientes quedarán en el estudio. Eso ocurre, por ejemplo, si en lugar de precisar el número de individuos que quedarán incluidos, lo que se fija es un lapso durante el cual se irán incorporando pacientes al ensayo. Puesto que es conveniente que cuando se haga el análisis, los grupos tengan tamaños similares entre sí, una manera de ayudar a conseguirlo consiste en asignar los pacientes del modo siguiente: el sujeto se asigna al grupo con probabilidad variable, más concretamente: menor cuanto mayor sea el número de sujetos ya asignados a dicho grupo hasta ese momento. Esto quiere decir que se aplica un algoritmo tal que cada vez que se va a asignar un nuevo sujeto a algún grupo, la probabilidad de que ese sujeto quede asignado a dicho grupo sea inversamente proporcional al número de individuos incluidos en él hasta ese momento. De ese modo, la asignación va prosperando "equilibradamente".

Epidat 4.0 contiene la posibilidad de dar solución al problema que se acaba de describir.

3.3.1. El proceso de asignación aleatoria

A continuación se enumeran y explican los datos que han de introducirse:

Número de grupos (tratamientos): Valor entero mayor que uno y menor o igual que 5. Identifica cuántos grupos hay que formar.

Número total de sujetos: Valor entero igual o mayor que el número de tratamientos considerados. En caso de que se opte por formar grupos de idéntico tamaño, este número será igual al producto de dicho tamaño por el número de tratamientos. Consecuentemente, bajo esta opción el número de sujetos tiene que ser un múltiplo del número de tratamientos.

Tipo de grupos a crear: Debe seleccionarse de qué forma se van a crear los grupos de sujetos; existen dos posibilidades: grupos de igual tamaño o grupos equilibrados. La activación de una de estas opciones inhabilita la otra.

Ejemplo

Supóngase que se planifica un ensayo clínico para evaluar la eficacia de la homeopatía en la curación de cierto tipo de conjuntivitis, y que se compararan tres tratamientos: un colirio convencional (A), cierto remedio homeopático (B) y agua destilada (C). Supóngase que se ha decidido aplicar cada uno de ellos a 20 pacientes. Asígnese que es entonces necesario distribuir aleatoriamente a cada uno de los próximos 60 portadores de esa forma de conjuntivitis que acudan a la consulta a uno de los tres tratamientos, y hacerlo de modo que queden 20 en cada cual.

Aplicar una asignación aleatoria a los tres grupos en las condiciones mencionadas no es otra cosa que lo siguiente: primero, sacar una muestra simple aleatoria de tamaño 20 de dicho listado y asignarle el tratamiento A a sus integrantes; luego obtener otra muestra simple aleatoria de los 40 que restan y asignar el tratamiento B a éstos, y finalmente ubicar en el C a los 20 pacientes que no fueron elegidos hasta entonces.

El procedimiento es algo engorroso, aun contando con un programa informático que realice una selección simple aleatoria dentro de un listado. En efecto, después de seleccionar el primer grupo de 20 habría que reenumerar del 1 al 40 a los sujetos no elegidos, y aplicar entonces por segunda vez el programa. Por eso es conveniente contar con programas *ad hoc* para realizar la asignación, aunque lo que hacen en esencia tales programas es aplicar sucesivamente el MSA.

En Epidat 4.0 se puede escoger en este caso la alternativa de formar 3 grupos de tamaño 20 (Grupos iguales):

Datos:						
Tipo de grupos a crear:	Grupos de igual tamaño					
Número de grupos:	3					
Número total de sujetos:	60					
Número de los sujetos seleccionados:						
Grupo 1						
1	2	4	6	15	18	22
25	29	30	32	34	37	44
45	46	48	55	58	60	
Grupo 2						
3	5	7	9	12	13	16
19	20	26	28	36	40	41
42	50	51	53	54	57	
Grupo 3						
8	10	11	14	17	21	23
24	27	31	33	35	38	39
43	47	49	52	56	59	

Si se hubiera optado por el recurso de formar grupos equilibrados, los subconjuntos que define Epidat no necesariamente tendrán 20 sujetos exactamente como ocurre en el que se acaba de ver, sino que la distribución será sólo aproximadamente equitativa.

3.4. Estimación con muestras complejas

3.4.0. Introducción

El manejo de una muestra simple aleatoria (MSA) se explica en textos, cursos básicos de estadística, programas informáticos, etc. Sin embargo, el tratamiento estadístico de los datos de un estudio donde se empleó un diseño muestral complejo debe contemplar tanto el ocasional carácter no equiprobabilístico de la muestra como la estructura del diseño muestral. Este problema exhibe cierta complejidad y, consecuentemente, su tratamiento exige contar con programas especializados. Aunque hay programas disponibles, éstos son por lo general bastante complicados de manejar o se hallan en el contexto de grandes paquetes estadísticos. Epidat 4.0 resuelve el problema en un entorno amigable y tiene el atractivo de que opera con archivos de OpenOffice o de Excel.

3.4.1. Probabilidad de selección y ponderación

El contexto de un diseño muestral complejo es, en general, el siguiente: la población inicial se divide en L estratos; dentro de cada estrato se eligen al menos dos conglomerados o unidades de primera etapa (UPE). Dentro de cada una de las UPE así elegidas en la población se selecciona cierto número de unidades de análisis (las que serán medidas o interrogadas) usando una o más etapas de selección. Por otra parte, además de la estratificación y las etapas del muestreo, la selección en cada una estas etapas puede llevarse a cabo de distintos modos: por ejemplo, mediante MSA o con probabilidades proporcionales al tamaño de las unidades de muestreo. Obviamente, en la práctica se utilizan distintas combinaciones de los elementos que intervienen en el diseño muestral: puede ocurrir que la selección se realice en varias etapas pero sin estratificación, o que se aplique un muestreo estratificado de conglomerados en una única etapa.

En este proceso, a cada individuo de la muestra le corresponde cierta probabilidad de selección, que depende, naturalmente, de la forma en que se haya llevado adelante la selección de la UPE a la que pertenece en su estrato y la que se empleó para el propio sujeto dentro de dicha UPE. En cualquier caso, la probabilidad de selección de un individuo es el producto de las probabilidades de selección en las distintas etapas que conducen a su inclusión en la muestra. Una vez que se conoce dicha probabilidad, la ponderación se calcula, generalmente, como su inverso, y se interpreta como el número de individuos de la población que representa uno de la muestra. Para procesar muestras complejas en Epidat 4.0 es necesario que los pesos estén definidos de este modo, debido a los estimadores que se utilizan.

Puesto que las variantes en este proceso son muy diversas, resulta imposible exponerlas exhaustivamente. No obstante, consideremos un ejemplo a modo de ilustración.

Ejemplo

Supongamos que la población de ciudadanos en una pequeña ciudad está dividida en 3 estratos (Zona Norte, Zona Centro y Zona Sur). En cada zona se han seleccionado al azar 4

manzanas (UPE) y en cada manzana seleccionada se ha obtenido una MSA de 20 sujetos. El tamaño de muestra global será: $3 \cdot 4 \cdot 20 = 240$ individuos. La probabilidad de selección de una manzana perteneciente al estrato h será:

$$\frac{4}{M_h}$$

donde M_h es el número de manzanas existentes en la zona o estrato h. En la segunda etapa, la probabilidad de selección de un individuo que reside en la manzana j del estrato h será:

$$\frac{20}{N_{hj}}$$

donde N_{hj} es el número de individuos de dicha manzana. El producto de estos dos factores da como resultado la probabilidad de selección de cada individuo. Por ejemplo, la probabilidad del octavo individuo elegido en la segunda manzana del tercer estrato es:

$$\frac{4}{M_3} \frac{20}{N_{32}}$$

La ponderación que corresponde a ese individuo es el inverso de la probabilidad de selección.

3.4.2. Diseños muestrales abarcados

El diseño que se ha expuesto suele denominarse con el nombre de muestreo estratificado polietápico (MEP). Cabe señalar que se trata de un diseño sumamente general, que abarca diversos casos particulares. Entre ellos, se destacan los siguientes casos especiales:

- A. Muestreo simple aleatorio o sistemático:** Caso del MEP en que se tiene un solo estrato, en que cada conglomerado contiene un único sujeto y en que la ponderación es constante para todos los elementos.
- B. Muestreo aleatorio estratificado con asignación proporcional:** Caso del MEP en que se trabaja con estratos pero donde cada conglomerado contiene un único sujeto y a todos los individuos corresponde la misma ponderación.
- C. Muestreo aleatorio estratificado con asignación no proporcional:** Caso del MEP en que cada conglomerado contiene un único sujeto y a los individuos de diferentes estratos pueden corresponderles ponderaciones diferentes (el número de sujetos existentes en el estrato al que pertenece dividido por el número de los que fueron seleccionados en dicho estrato).
- D. Muestreo por conglomerados:** Caso del MEP en que hay un solo estrato. Las ponderaciones pueden ser o no iguales, dependiendo de cómo se hayan subseleccionado los sujetos dentro de las UPE o conglomerados. Por ejemplo, si el muestreo fue monoetápico o si fue un muestreo bietápico con probabilidades proporcionales al tamaño, donde más tarde se selecciona igual número de sujetos en cada UPE seleccionada, entonces el diseño es equiprobabilístico (iguales ponderaciones para todos los sujetos); si fuera bietápico con las UPE elegidas al azar y un número fijo de sujetos elegidos por muestreo simple aleatorio dentro de cada UPE seleccionada, entonces los sujetos de distintos conglomerados tendrán ponderaciones diferentes.

3.4.3. Efecto de diseño

Típicamente, la realización de un diseño muestral complejo entraña una pérdida de precisión respecto de una muestra simple aleatoria. La manera usual de medir a posteriori el grado en que ese diseño complejo es más ineficiente que el MSA (el cual sería el caso en que "no hay diseño" pues solo el azar es responsable de lo que pueda obtenerse como muestra) es a través del llamado *efecto de diseño* (*design effect* en inglés), que suele denotarse como *d_{eff}* y se define del modo siguiente:

$$d_{eff} = \sqrt{\frac{\text{var}(\hat{R})}{\text{var}_{MSA}(\hat{R})}}$$

donde $\text{var}_{MSA}(\hat{R})$ es la varianza si se hubiera usado MSA, y \hat{R} es una estimación cualquiera (por ejemplo, la de una media o una proporción, que son los casos más comunes).

Lo más frecuente es que la varianza del diseño complejo sea mayor, a veces mucho mayor, que la que correspondería si se manejara la muestra como si fuera una MSA, de modo que *d_{eff}* suele ser mayor que la unidad.

3.4.4. Tipo de variables

Epidat 4.0 puede trabajar con variables cuantitativas, tal como la tensión arterial o la edad, en cuyo caso, \hat{R} es una estimación de la media poblacional. Pero también admite variables categóricas o politómicas. Por ejemplo, dicha variable pudiera registrar la RELIGIÓN de un sujeto y estar conformada por 4 categorías: CATÓLICO, PROTESTANTE, ATEO, OTRO; o pudiera identificar el GRADO en que se halla un escolar de Primaria, y estar conformada por las 6 categorías: 1º, 2º, 3º, 4º, 5º y 6º. En el caso de tratarse de una variable categórica, Epidat 4.0 computará tantas estimaciones puntuales y respectivos intervalos para los porcentajes como categorías diferentes contenga la variable en cuestión.

Un caso particular de variable categórica que tiene especial relevancia es aquella con solo dos categorías, ya que es la situación en que se quiere estimar un porcentaje.

3.4.5. Trabajo con clases o subpoblaciones

Epidat 4.0 estima las medias y/o proporciones de las variables que se indiquen utilizando para ello toda la muestra; es decir, considera que se quieren realizar estimaciones globales. Pero ocasionalmente el usuario puede desear hacer tales estimaciones para una "clase" o subpoblación (por ejemplo, solo para las mujeres, o solo para los mayores de 60 años que además hayan declarado estar de acuerdo con cierta ley). Epidat 4.0 contempla esta posibilidad a través del recurso de establecer "filtros", de modo que el análisis se reduzca a los integrantes de la muestra que cumplan las condiciones establecidas por dicho filtro. Dicho de otro modo, se pueden fijar filtros que definen, de hecho, un subconjunto de la muestra y las estimaciones corresponderán a la respectiva subpoblación.

Por otra parte, Epidat 4.0 ofrece también la posibilidad de segmentar los resultados en función de las categorías de una variable cualitativa, de modo que calcula las estimaciones solicitadas en cada subpoblación de individuos definida por dichas categorías.

3.4.6. Manejo del submódulo

En este submódulo de Epidat 4.0 solo se puede realizar la entrada de datos de forma automática, no es posible introducirlos manualmente. Para ello hay que importar los datos individuales a partir de un archivo en formato de Excel (*.xls, *.xlsx) o de OpenOffice (*.ods) por medio del asistente de datos. Es aconsejable que el archivo que se esté procesando no esté abierto por otra aplicación mientras Epidat 4.0 esté operativo. Aunque Epidat 4.0 no introduce ningún cambio en el archivo de trabajo, no se ha establecido ninguna protección explícita para el uso múltiple de un archivo.

Las tablas de datos con que Epidat 4.0 opera deben tener una estructura específica. En este caso, cada columna corresponde a una variable y cada fila a un individuo o unidad de análisis de la muestra; en la primera fila se ubica el nombre de las variables y a partir de la segunda fila han de figurar los valores de dichas variables. Además, cada columna debe contener solo el nombre y los valores de la variable.

Una vez declarado el libro de OpenOffice o de Excel con que se habrá de trabajar, el usuario debe indicar el nombre de las variables donde se consignan el ESTRATO y el CONGLOMERADO en que se hallan los sujetos, así como la variable que contiene la PONDERACIÓN. Todas estas variables son opcionales, es decir, solo tienen que indicarse las necesarias. Consecuentemente, si el estudio fuera equiprobabilístico, no se señalará ningún campo para la ponderación. Si el diseño no usó estratos (es decir, operó con un único estrato) no se indicará ningún campo para los estratos. Si la muestra procede de un diseño que no usó conglomerados, entonces no se señalará ningún campo para los conglomerados. Además de las variables que definen el diseño de la muestra, hay que seleccionar las variables cuantitativas para las que se desea estimar medias y/o las variables cualitativas para las que estimar proporciones. Por último, es posible indicar una variable categórica para segmentar los resultados, como ya se comentó.

Si para alguna o algunas variables involucradas en el proceso se desconocen algunos datos, simplemente las celdas correspondientes deben dejarse vacías. Epidat 4.0 operará solo con los datos existentes en esas filas.

El programa puede acceder a todas las variables de una hoja que, consecutivamente y a partir de la primera columna, posean información en la primera fila. Si la primera fila de la primera columna de una hoja está vacía, entonces Epidat 4.0 asume que esa hoja no dispone de información alguna.

Ha de advertirse que los nombres de las variables no tienen que ser diferentes, pero el usuario ha de tener la precaución de ponerle diferentes nombres si desea evitar ocasionales confusiones.

Por último, téngase en cuenta que, si el diseño declarado involucra conglomerados, la muestra que se está procesando (resultado o no de un filtrado previo), debe contener representantes de al menos dos conglomerados diferentes en cada estrato. De no ser así, el programa no podrá prosperar y se detendrá.

Ejemplo

Imaginemos que se ha realizado una encuesta a una muestra de 18 sujetos. Para cada cual se han registrado las siguientes variables:

- **MEDIDA:** variable nominal que clasifica al sujeto según grado de acuerdo con cierta medida; las tres respuestas posibles a dicha variable son: "A favor", "En contra", "Indeciso";

- **DIENTES:** registra el número de dientes naturales que tiene el sujeto;
- **ESTU:** Establece si el sujeto ha realizado o no estudios superiores y puede ser positiva (valor 1) o negativamente contestada (valor 0).

Supóngase que la muestra de 18 individuos procede de una muestra estratificada bietápica. La población para comenzar fue dividida en tres estratos. Éstos tenían 5, 16 y 20 UPE cada uno. Se tomaron al azar 2, 3 y 3 UPE, respectivamente. Los tamaños de dichas 8 UPE resultaron ser, por ese orden, 20, 30, 40, 67, 130, 12, 8 y 12. En cada una de estas 8 UPE se tomaron 2 sujetos mediante MSA, salvo en la segunda donde se eligieron 4 sujetos.

Es fácil convencerse de que el resumen de la estructura probabilística correspondiente a esta situación es la que sigue:

SUJETO	ESTRATO	UPE	PROBABILIDAD			PONDERACIÓN
			UPE	SUJETO	TOTAL	
1	1	1	0,40	0,10	0,04	25,00
2	1	1	0,40	0,10	0,04	25,00
3	1	2	0,40	0,13	0,05	18,75
4	1	2	0,40	0,13	0,05	18,75
5	1	2	0,40	0,13	0,05	18,75
6	1	2	0,40	0,13	0,05	18,75
7	2	1	0,19	0,05	0,01	106,67
8	2	1	0,19	0,05	0,01	106,67
9	2	2	0,19	0,03	0,01	178,67
10	2	2	0,19	0,03	0,01	178,67
11	2	3	0,19	0,02	0,00	346,67
12	2	3	0,19	0,02	0,00	346,67
13	3	1	0,15	0,17	0,03	40,00
14	3	1	0,15	0,17	0,03	40,00
15	3	2	0,15	0,25	0,04	26,67
16	3	2	0,15	0,25	0,04	26,67
17	3	3	0,15	0,17	0,03	40,00
18	3	3	0,15	0,17	0,03	40,00

El libro INSTITUTO.xls (hoja *Dientes*) contiene el resultado de una hipotética encuesta realizada a estos 18 sujetos donde, además de las variables (MEDIDA, DIENTES, ESTU) figuran el estrato (ESTR), el conglomerado (CONG) y la ponderación (W) que corresponde a cada uno de ellos.

Los resultados y los datos necesarios que han de figurar en la hoja de trabajo si el estudio fue un MEP, aparecen en la siguiente tabla:

NOMBRE	NUMERO	ESTR	CONG	W	MEDIDA	DIENTES	ESTU
JULIETTA AZANZA	1	1	1	25,00	A favor	15	1
EDITH BENAVIDEZ	2	1	1	25,00	Indeciso	14	0
IÑAKI MANRESA	3	1	2	18,87	En contra	7	0
TOMAS MOLINER	4	1	2	18,87	Indeciso	8	1
JORGE RODRÍGUEZ	5	1	2	18,87	A favor	8	0
MILAGROS DEYBIS	6	1	2	18,87	En contra	4	0
ARSENIO CAÑAS	7	2	1	111,11	En contra	32	1
REGLA TABOADA	8	2	1	111,11	En contra	30	1
IVAN RODRÍGUEZ	9	2	2	166,67	En contra	27	0
DANIEL JIMÉNEZ	10	2	2	166,67	En contra		1
WILFREDO CABRERA	11	2	3	333,33	A favor	20	1
ANTONIO RODRÍGUEZ	12	2	3	333,33	A favor	19	0
JAIME ARRIZABALAGA	13	3	1	40,00	Indeciso	3	0
PATRICIA GRANDES	14	3	1	40,00	Indeciso	0	1
AURORA HORTA	15	3	2	26,31	Indeciso	7	0
GRACIELA PÉREZ	16	3	2	26,31	A favor	8	0
ROBERTO TORRALBA	17	3	3	40,00	En contra	12	0
JESÚS NAZUNTALO	18	3	3	40,00	A favor	13	1

Notas:

- Los dos primeros campos (NOMBRE y NÚMERO) no son necesarios para llevar adelante el procesamiento, pero Epidat 4.0 no exige que en la hoja de trabajo solo estén presentes los campos necesarios para el procesamiento; consecuentemente, pueden figurar en ella algunas variables que no vayan a ser empleadas.
- Obsérvese que el número de dientes para el sujeto 10 (Daniel Jiménez) no figura en la base (quizás porque este número no pudo determinarse en esta unidad), de modo que la variable DIENTES solo está presente para 17 individuos.

Las estimaciones para las 3 variables del estudio en este caso son las siguientes (ha de indicarse al programa para cuales se quieren estimar medias (DIENTES) y para cuáles proporciones (MEDIDA, ESTU):

Entrada automática:

Archivo de trabajo: C:\Archivos de programa\Epidat 4.0\Ejemplos\Muestreo\INSTITUTO.xls

Tabla: Dientes

Variables:

Estratos: ESTR
 Conglomerados: CONG
 Ponderacion: W
 Estimar medias: DIENTES
 Estimar proporciones: MEDIDA, ESTU

Datos:

Nivel de confianza: 95,0%

Filtro: No

Resultados:**Medias:**

Variable	N	Media	Media ponderada	E.E.	IC (95,0%)		Deff
DIENTES	17	13,353	19,473	1,769	16,006	22,939	0,771

Proporciones para la variable MEDIDA:

Categoría	N	%	% ponderado	EE (%)	IC (95,0%)		Deff
A favor	18	33,333	49,787	30,604	-10,195	109,769	2,677
En contra	18	38,889	40,588	28,306	-14,890	96,066	2,394
Indeciso	18	27,778	9,625	5,160	-0,489	19,739	0,475

Proporciones para la variable ESTU:

Categoría	N	%	% ponderado	EE (%)	IC (95,0%)		Deff
0	18	55,556	45,775	8,175	29,752	61,797	0,678
1	18	44,444	54,225	8,175	38,203	70,248	0,678

*EE: Error estándar

*Deff: Efecto de diseño

Nota: Adviértase que el programa siempre informa el valor del promedio (o porcentaje) no ponderado de los datos, además de la estimación correctamente ponderada. Ambas coinciden, naturalmente, para los diseños equiprobabilísticos; pero típicamente difieren en otro caso. La estimación no ponderada carece por lo general de interés intrínseco; su inclusión se debe exclusivamente a que ocasionalmente el usuario desea tener una idea del efecto de la ponderación.

Muestreo por conglomerados monoetápico

Si, por ejemplo, esta misma muestra hubiera sido el resultado de haber empleado un muestreo por conglomerados monoetápico, entonces se dejan en blanco los campos correspondientes a los estratos (ya que no hubo estratificación) y a las ponderaciones (ya que el muestreo sería equiprobabilístico o autoponderado); solo se marca que el de los conglomerados es CONG, y el resultado hubiera sido:

Entrada automática:

Archivo de trabajo: C:\Archivos de programa\Epidat 4.0\Ejemplos\Muestreo\INSTITUTO.xls

Tabla: Dientes

Variables:

Conglomerados: CONG
 Estimar medias: DIENTES
 Estimar proporciones: MEDIDA, ESTU

Datos:

Nivel de confianza: 95,0%
 Filtro: No

Resultados:**Medias:**

Variable	N	Media	Media ponderada	E.E.	IC (95,0%)		Deff
DIENTES	17	13,353	13,353	2,166	9,108	17,598	0,944

Proporciones para la variable MEDIDA:

Categoría	N	%	% ponderado	EE (%)	IC (95,0%)		Deff
A favor	18	33,333	33,333	13,981	5,931	60,736	1,223
En contra	18	38,889	38,889	7,484	24,220	53,558	0,633
Indeciso	18	27,778	27,778	11,906	4,443	51,113	1,096

Proporciones para la variable ESTU:

Categoría	N	%	% ponderado	EE (%)	IC (95,0%)		Deff
0	18	55,556	55,556	14,022	28,073	83,038	1,163
1	18	44,444	44,444	14,022	16,962	71,927	1,163

*EE: Error estándar

*Deff: Efecto de diseño

Muestreo simple aleatorio segmentando datos

Finalmente, supóngase que la muestra hubiera resultado de la aplicación de un muestreo simple aleatorio, pero que se quieren las estimaciones solo para los sujetos que tienen estudios superiores; entonces se indica que se segmente el análisis según los niveles de la variable ESTU y se dejan todos los demás campos en blanco (es decir, no se marca nada en PONDERACIONES, ni en ESTRATOS ni en CONGLOMERADOS). El resultado para DIENTES y MEDIDA en este caso sería:

Entrada automática:

Archivo de trabajo: C:\Archivos de programa\Epidat 4.0\Ejemplos\Muestreo\INSTITUTO.xls

Tabla: Dientes

Variables:

Estimar medias: DIENTES
 Estimar proporciones: MEDIDA
 Segmentar resultados : ESTU

Datos:

Nivel de confianza: 95,0%
 Filtro: No

Resultados para ESTU=0:

Medias:

Variable	N	Media	Media ponderada	E.E.	IC (95,0%)		Deff
DIENTES	10	10,900	10,900	2,340	6,313	15,487	1,000

Proporciones para la variable MEDIDA:

Categoría	N	%	% ponderado	EE (%)	IC (95,0%)		Deff
A favor	10	30,000	30,000	15,275	0,061	59,939	1,000
En contra	10	40,000	40,000	16,330	7,994	72,006	1,000
Indeciso	10	30,000	30,000	15,275	0,061	59,939	1,000

Resultados para ESTU=1:

Medias:

Variable	N	Media	Media ponderada	E.E.	IC (95,0%)		Deff
DIENTES	7	16,857	16,857	4,345	8,341	25,373	1,000

Proporciones para la variable MEDIDA:

Categoría	N	%	% ponderado	EE (%)	IC (95,0%)		Deff
A favor	8	37,500	37,500	18,298	1,636	73,364	1,000
En contra	8	37,500	37,500	18,298	1,636	73,364	1,000
Indeciso	8	25,000	25,000	16,366	-7,077	57,077	1,000

*EE: Error estándar

*Deff: Efecto de diseño

Bibliografía

- 1 Rothman JK. Modern epidemiology. Boston: Little, Brown and Col; 1982.
- 2 Silva LC. Cultura estadística e investigación científica en el campo de la salud: una mirada crítica. Madrid: Díaz de Santos; 1997.
- 3 Silva LC. Diseño razonado de muestras y captación de datos para la investigación sanitaria. Madrid: Díaz de Santos; 2000.
- 4 Marrugat J, Vila J, Pavesi J, Sanz F. Estimación del tamaño de muestra en la investigación clínica y epidemiológica. Med Clin (Barc). 1998;111:267-76.
- 5 Lwanga SK, Lemeshow S. Determinación del tamaño de las muestras en los estudios sanitarios: manual práctico. Ginebra: OMS; 1991.
- 6 Suárez P, Alonso JC. Sobre el supuesto de máxima indeterminación, el tamaño muestral y otras consideraciones sobre muestreo. Gacet Sanit. 1999;13(3):243-6.
- 7 Silva LC. Nueva visita al supuesto de máxima indeterminación y al empleo de errores absolutos y relativos. Gacet Sanit. 2000;14(3):254-7.
- 8 Marrugat J, Vila J, Pavesi J. Supuesto de máxima indeterminación: ¿error absoluto o error relativo en el cálculo del tamaño de la muestra?. Gacet Sanit. 1999;13(6):491-3.
- 9 García C, Almenara J. Determinación del tamaño de muestra en variables cualitativas en las que se desconoce el valor del parámetro. Med Clin (Barc). 1999;112:797-8.
- 10 Thompson WD. Análisis estadístico de los estudios de casos y controles. Bol Oficina Sanit Panam. 1996;121(1):41-61.
- 11 International Committee of Medical Journal Editors [página en Internet]. Uniform requirements for manuscripts submitted to biomedical journals [actualizado Oct 2008; citado 11 Jul 2009]. Disponible en: <http://www.icmje.org>
- 12 Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection on controls in case-controls studies. III. Design options. Am J Epidemiol. 1992;135 (9):1042-9.
- 13 Fleiss JL. Statistical methods for rates and proportions. New York: John Wiley & Sons; 1981.
- 14 Cantor AB. Sample-size calculations for Cohen's kappa. Psychological Methods. 1996;1(2):150-3.
- 15 Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. Pub Health Rep. 1947;62:1432-49.
- 16 Obuchowsky NA. Sample size calculations in studies of test accuracy. Statistical Methods in Medical Research. 1998;7:371-92.

- 17 Silva LC. Los laberintos de la investigación biomédica. En defensa de la racionalidad para la ciencia del siglo XXI. Madrid: Díaz de Santos; 2009.
- 18 Orlitzky M [página en Internet]. Institutionalized dualism: Statistical significance testing as myth and ceremony (8 Jun 2009). Disponible en: <http://ssrn.com/abstract=1415926>
- 19 Nickerson RS. Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*. 2000;5:241-301.
- 20 Cohen J. *Statistical power analysis for the behavioral sciences*. 2ª ed. New Jersey: Lawrence Erlbaum; 1988.
- 21 Machin D, Campbell MJ, Fayers PM, Pinol APY. *Sample size tables for clinical studies*. 2ª ed. Blackwell Science Ltd; 1997.
- 22 Connor RJ. Sample size for testing differences in proportions for the paired-sample design. *Biometrics*. 1987;43:207-11.
- 23 Saturno PJ. La distribución binomial y el muestreo para la aceptación de lotes (LQAS) como métodos de monitorización en servicios de salud. *Rev Calidad Asistencial*. 2000;15:99-107.
- 24 López-Picazo JJ. Aceptación de muestras por lotes: aplicabilidad en la evaluación de cartera de servicios de Atención Primaria. *Aten Primaria*. 2001;27(8):547-53.
- 25 Corbella A, Fernández-Fernández O, Pérez-Orcuna X, Grima P. Aplicación de la técnica de planes de muestreo por lotes en la evaluación de indicadores: asma en pediatría. *Rev Calidad Asistencial*. 2000;15:79-87.
- 26 Ahnn S, Anderson SJ. Sample size determination for comparing more than two survival distributions. *Stat Med*. 1995;14:2273-82.
- 27 Neyman J. On the two different aspects of the representative method: The method of stratified sampling and the method of positive selection. *Journal of the Royal Statistical Society*. 1934;25:558-606.
- 28 Fariñas H, Silva LC. Selección equiprobabilística mediante muestreo sistemático para una probabilidad prefijada. *Revista Investigación Operacional*. 1984;5:51-77 (Cuba).
- 29 Hulley SB, Cummings SR. *Designing clinical research: An epidemiological approach*. Baltimore: Williams and Wilkins; 1988.

Anexo 1: Novedades del módulo de muestreo en la versión 4

- ⇓ Cálculo de tamaños de muestra: se utiliza una nueva fórmula para el caso de comparación de medias independientes.
- ⇓ Cálculo de tamaños de muestra: en comparación de medias emparejadas se da la opción de introducir la desviación estándar de las diferencias en lugar de las desviaciones de cada muestra con el coeficiente de correlación.
- ⇓ Cálculo de tamaños de muestra: al calcular precisión para un riesgo relativo se pide el tamaño de muestra total en lugar del número de expuestos.
- ⇓ Cálculo de tamaños de muestra: al calcular la potencia en estudios de cohortes, se pide el tamaño de muestra total en lugar del número de expuestos. Además, se da la opción de introducir como dato de entrada el riesgo relativo.
- ⇓ Cálculo de tamaños de muestra: se añadieron tres nuevas opciones:
 - ⇓ Comparación de proporciones independientes.
 - ⇓ Estudios de equivalencia de medias.
 - ⇓ Estudios de equivalencia de proporciones.
- ⇓ Selección de muestras: hay más flexibilidad a la hora de definir los elementos que intervienen en el diseño muestral, de modo que se permiten diseños no equiprobabilísticos.
- ⇓ Selección de muestras: en muestreo sistemático se añadió la posibilidad de utilizar muestreo sistemático regular.
- ⇓ Selección de muestras: se incluyó una nueva opción: muestreo por conglomerados monoetápico estratificado.
- ⇓ Se incluyó un nuevo submódulo para estimación con muestras complejas.

Anexo 2: Fórmulas del módulo de muestreo

Esquema del módulo

1. Cálculo de tamaños de muestra
 - 1.1. Intervalos de confianza
 - 1.1.1. Media
 - 1.1.2. Proporción
 - 1.1.3. Odds ratio
 - 1.1.4. Riesgo relativo
 - 1.1.5. Concordancia
 - 1.1.6. Pruebas diagnósticas
 - 1.2. Contrastes de hipótesis
 - 1.2.1. Comparación de medias
 - 1.2.1.1. Grupos independientes
 - 1.2.1.2. Grupos emparejados
 - 1.2.2. Comparación de proporciones
 - 1.2.2.1. Grupos independientes
 - 1.2.2.2. Grupos emparejados
 - 1.2.3. Estudios de casos y controles
 - 1.2.3.1. Grupos independientes
 - 1.2.3.2. Grupos emparejados
 - 1.2.4. Estudios de cohorte
 - 1.2.5. Estudios de equivalencia
 - 1.2.5.1. Equivalencia de medias
 - 1.2.5.2. Equivalencia de proporciones
 - 1.2.6. Pruebas diagnósticas
 - 1.2.6.1. Grupos independientes
 - 1.2.6.2. Grupos emparejados
 - 1.2.7. Calidad de lotes
 - 1.2.8. Supervivencia
 - 1.2.9. Coeficiente de correlación

Esquema del módulo (continuación)

2. Selección muestral
 - 2.1. Muestreo simple aleatorio
 - 2.2. Muestreo sistemático en fases
 - 2.3. Muestreo aleatorio estratificado
 - 2.4. Muestreo por conglomerados monoetápico
 - 2.5. Muestreo por conglomerados bietápico
 - 2.6. Muestreo por conglomerados monoetápico estratificado
 - 2.7. Muestreo por conglomerados bietápico estratificado
3. Asignación de sujetos a tratamientos
4. Estimación con muestras complejas

1.- CÁLCULO DE TAMAÑOS DE MUESTRA

1.1.- INTERVALOS DE CONFIANZA

Este submódulo permite calcular el tamaño de muestra mínimo para estimar un parámetro θ mediante un intervalo de confianza de nivel de confianza $(1-\alpha)\%$ con una cierta precisión (absoluta o relativa). También calcula la precisión del intervalo de confianza para un tamaño dado n .

En todas las opciones:

- z_α es el percentil de la distribución normal estándar, $N(0,1)$, que deja a la izquierda una cola de probabilidad α ,
- $1-\alpha$ es el nivel de confianza.

1.1.1.- **MEDIA** [Machin (1997, p. 133-135)]

Tamaño de muestra:

$$n = \left(z_{1-\alpha/2} \frac{\sigma}{e} \right)^2, \text{ si la población es infinita,}$$

$$n_F = \frac{Nn}{N+n}, \text{ si la población es finita,}$$

Los tamaños resultantes se multiplican por el efecto de diseño ($deff$).

Donde:

- σ es la desviación estándar esperada en la población,
- e es la precisión absoluta de un intervalo de confianza para la media,
- N es el tamaño de la población,
- $deff$ es el efecto de diseño.

Precisión absoluta:

$$e = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \text{ si la población es infinita,}$$

$$e_F = e\sqrt{1-f}, \text{ si la población es finita,}$$

Los valores resultantes se multiplican por el efecto de diseño (deff).

Donde:

- n es el tamaño de la muestra,
- $f = \frac{n}{N}$.

1.1.2.- PROPORCIÓN [Machin (1997, p. 132-133)]

Tamaño de muestra:

$$n = \left(\frac{z_{1-\alpha/2}}{e} \right)^2 P(1-P), \text{ si la población es infinita,}$$

$$n_F = \frac{Nn}{N+n}, \text{ si la población es finita,}$$

Los tamaños resultantes se multiplican por el efecto de diseño (deff).

Donde:

- P es la proporción esperada en la población,
- e es la precisión absoluta de un intervalo de confianza para la proporción,
- deff es el efecto de diseño,
- N es el tamaño de la población.

Precisión absoluta:

$$e = z_{1-\alpha/2} \sqrt{\frac{P(1-P)}{n}}, \text{ si la población es infinita,}$$

$$e_F = e\sqrt{1-f}, \text{ si la población es finita,}$$

Los valores resultantes se multiplican por el efecto de diseño (deff) y por 100.

Donde:

- n es el tamaño de la muestra,
- $f = \frac{n}{N}$.

1.1.3.- ODDS RATIO [Machin (1997, p. 134)]

Tamaños de muestra:

$$n_1 = \frac{z_{1-\alpha/2}^2 \left[\frac{1}{P_1(1-P_1)} + \frac{1}{\phi P_0(1-P_0)} \right]}{[\ln(1-\varepsilon)]^2} \text{ (casos)}$$

$$n_0 = \phi n_1 \text{ (controles)}$$

Donde:

- P_1 es la proporción de casos expuestos,
- P_0 es la proporción de controles expuestos,
- OR es la odds ratio,
- P_1 , P_0 y OR se relacionan del modo siguiente:

$$P_1 = \frac{OR P_0}{(1-P_0) + OR P_0}, P_0 = \frac{P_1}{OR(1-P_1) + P_1}, OR = \frac{P_1(1-P_0)}{P_0(1-P_1)},$$

- ε es la precisión relativa de un intervalo de confianza para la odds ratio,
- ϕ es el número de controles por caso.

Precisión relativa:

$$\varepsilon = \left[1 - \exp\left(-z_{1-\alpha/2} \sqrt{\text{Var}(\ln OR)}\right) \right] \times 100$$

Donde:

- $\text{Var}(\ln OR) = \frac{1}{n_1} \left[\frac{1}{P_1(1-P_1)} + \frac{1}{\phi P_0(1-P_0)} \right]$,
- n_1 es el tamaño de la muestra de casos.

1.1.4.- RIESGO RELATIVO [Lwanga (1991, p. 52)]

Tamaños de muestra:

$$n_1 = \frac{z_{1-\alpha/2}^2 \left[\frac{1-P_1}{P_1} + \frac{1-P_0}{\phi P_0} \right]}{[\ln(1-\varepsilon)]^2} \text{ (expuestos)}$$

$$n_0 = \phi n_1 \text{ (no expuestos)}$$

Donde:

- P_1 es el riesgo en expuestos,
- P_0 es el riesgo en no expuestos,
- RR es el riesgo relativo,
- P_1, P_0 y RR se relacionan del modo siguiente:
 - $P_1 = P_0 RR, P_0 = \frac{P_1}{RR}, RR = \frac{P_1}{P_0},$
- ε es la precisión relativa de un intervalo de confianza para el riesgo relativo,
- ϕ es la razón entre los tamaños de no expuestos y expuestos.

Precisión relativa:

$$\varepsilon = \left[1 - \exp\left(z_{1-\alpha/2} \sqrt{\text{Var}(\ln RR)} \right) \right] \times 100$$

Donde:

- $\text{Var}(\ln RR) = \frac{1}{n_1} \left[\frac{1-P_1}{P_1} + \frac{1-P_0}{\phi P_0} \right],$
- $n_1 = \frac{n}{1+\phi}$ es el tamaño de la muestra de expuestos,
- n es el tamaño total de muestra.

1.1.5.- CONCORDANCIA [Cantor (1996)]

Tamaño de muestra:

$$n = \left(\frac{z_{1-\alpha/2}}{e} \right)^2 Q$$

Donde:

- e es la precisión absoluta de un intervalo de confianza para el coeficiente kappa,
- $Q = \frac{A+B-C}{(1-P_e)^2}$,
- $A = P_{11}[1-(P_{1.}+P_{.1})(1-k)]^2 + P_{22}[1-(P_{2.}+P_{.2})(1-k)]^2$,
- $B = (1-k)^2 [P_{12}(P_{1.}+P_{2.})^2 + P_{21}(P_{2.}+P_{1.})^2]$,
- $C = [k - P_e(1-k)]^2$,
- k es la concordancia esperada (coeficiente kappa),
- $P_{1.}=P_1$, es la proporción de clasificaciones positivas por el primer observador,
- $P_{.1}=P_2$, es la proporción de clasificaciones positivas por el segundo observador,
- $P_{2.}=1-P_1$ y $P_{.2}=1-P_2$,
- $P_e = P_{1.}P_{.1} + P_{2.}P_{.2}$,
- $P_0 = k(1-P_e) + P_e$,
- $P_{22} = \frac{P_0 - P_{1.} + P_{2.}}{2}$,
- $P_{11} = P_0 - P_{22}$, $P_{12} = P_{1.} - P_{11}$ y $P_{21} = P_{.1} - P_{11}$.

Precisión absoluta:

$$e = z_{1-\alpha/2} \sqrt{\frac{Q}{n}}$$

Donde:

- n es el tamaño de muestra.

1.1.6.- SENSIBILIDAD Y ESPECIFICIDAD [Obuchowski (1998)]Sensibilidad

Tamaños de muestra:

n_E y $n_{NE} = \phi n_E$, si se conoce a priori la condición de enfermo

$n = \frac{n_E}{P}$, si no se conoce a priori la condición de enfermo

$$n_E = \left(\frac{z_{1-\alpha/2}}{e} \right)^2 \theta_S (1 - \theta_S) \text{ (Fórmula para una proporción, población infinita)}$$

Donde:

- θ_S es la sensibilidad esperada,
- n_E es el número de enfermos,
- n_{NE} es el número de no enfermos,
- n es el tamaño total de muestra,
- ϕ es la razón entre los tamaños de no enfermos y enfermos,
- P es la prevalencia de enfermedad en la población,
- e es la precisión absoluta de un intervalo de confianza para θ_S .

Precisión absoluta:

$$e = z_{1-\alpha/2} \sqrt{\frac{\theta_S (1 - \theta_S)}{n_E}} \text{ (Fórmula para una proporción, población infinita)}$$

Donde:

- $n_E = \frac{n}{1 + \phi}$, si se conoce a priori la condición de enfermo,
- $n_E = nP$, si no se conoce a priori la condición de enfermo,
- n es el tamaño total de muestra.

Especificidad:

Tamaños de muestra:

$$n_{NE} \text{ y } n_E = \frac{n_{NE}}{\phi}, \text{ si se conoce a priori la condición de enfermo}$$

$$n = \frac{n_{NE}}{1-P}, \text{ si no se conoce a priori la condición de enfermo}$$

$$n_{NE} = \left(\frac{z_{1-\alpha/2}}{e} \right)^2 \theta_E (1 - \theta_E) \text{ (Fórmula para una proporción, población infinita)}$$

Donde:

- θ_E es la especificidad esperada,
- e es la precisión absoluta de un intervalo de confianza para θ_E .

Precisión:

$$e = z_{1-\alpha/2} \sqrt{\frac{\theta_E (1 - \theta_E)}{n_{NE}}} \text{ (Fórmula para una proporción, población infinita)}$$

Donde:

- $n_{NE} = \frac{n\phi}{1+\phi}$, si se conoce a priori la condición de enfermo,
- $n_{NE} = n(1-P)$, si no se conoce a priori la condición de enfermo,
- n es el tamaño total de muestra.

Sensibilidad y especificidad

Tamaños de muestra:

$$n_E \text{ y } n_{NE}, \text{ si se conoce a priori la condición de enfermo}$$

$$n = \max \left\{ \frac{n_E}{P}, \frac{n_{NE}}{1-P} \right\}, \text{ si no se conoce a priori la condición de enfermo}$$

n_E se calcula con la fórmula para la sensibilidad

n_{NE} se calcula con la fórmula para la especificidad

La precisión se calcula por separado para sensibilidad y especificidad.

1.2.- CONTRASTES DE HIPÓTESIS

Este submódulo permite calcular el tamaño de muestra mínimo para aplicar una prueba de hipótesis bilateral (salvo que se indique unilateral) con nivel de confianza $(1-\alpha)\%$ y potencia $(1-\beta)\%$. También calcula la potencia de la prueba para un tamaño de muestra dado n .

En todas las opciones:

- z_α es el percentil de la distribución normal estándar, $N(0,1)$, que deja a la izquierda una cola de probabilidad α ,
- $1-\alpha$ es el nivel de confianza,
- $1-\beta$ es la potencia,
- Φ es la función de distribución normal estándar.

1.2.1.1.- COMPARACIÓN DE MEDIAS, GRUPOS INDEPENDIENTES [Schouten (1999)]

Tamaños de muestra:

$$n_1 = \left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{d} \right)^2 \frac{(\tau + \phi) \sigma_1^2}{\phi} + \frac{(\tau^2 + \phi^3) z_{1-\alpha/2}^2}{2\phi(\tau + \phi)^2}, \text{ si las varianzas son distintas}$$

$$n_1 = \left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{\Delta} \right)^2 \frac{1 + \phi}{\phi} + \frac{z_{1-\alpha/2}^2}{2(1 + \phi)}, \text{ si las varianzas son iguales}$$

$$n_2 = \phi n_1$$

Donde:

- σ_i^2 es la varianza esperada en la población i , $i=1,2$,
- $\tau = \frac{\sigma_2^2}{\sigma_1^2}$
- d es la diferencia esperada entre las dos medias,
- ϕ es la razón entre los dos tamaños muestrales,
- $\Delta = \frac{d}{\sigma}$ es la diferencia estandarizada de medias,
- σ es la varianza común esperada en las dos poblaciones.

Potencia:

$$1 - \beta = \Phi(Z_{1-\beta})$$

$$z_{1-\beta} = |d| \sqrt{\left(n_1 - \frac{(\tau^2 + \phi^3) z_{1-\alpha/2}^2}{2\phi(\tau + \phi)^2} \right) \frac{\phi}{(\tau + \phi) \sigma_1^2} - z_{1-\alpha/2}}, \text{ si las varianzas son distintas}$$

$$z_{1-\beta} = |\Delta| \sqrt{\left(n_1 - \frac{z_{1-\alpha/2}^2}{2(1+\phi)} \right) \frac{\phi}{1+\phi} - z_{1-\alpha/2}}, \text{ si las varianzas son iguales}$$

Donde:

- $n_1 = \frac{n}{1+\phi}$, donde n es el tamaño total de la muestra.

1.2.1.2.- COMPARACIÓN DE MEDIAS, GRUPOS EMPAREJADOS [Machin (1997, p. 73)]

Número de pares:

$$n = \left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{\Delta} \right)^2 + \frac{z_{1-\alpha/2}^2}{2}$$

Donde:

- $\Delta = \frac{d}{\sigma_d}$ es la diferencia estandarizada de medias,
- d es la diferencia esperada entre las dos medias,
- $\sigma_d = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$ es la desviación estándar de las diferencias,
- σ_i es la desviación estándar esperada en la población i, i=1,2,
- ρ es el coeficiente de correlación entre ambas poblaciones.

Potencia:

$$1 - \beta = \Phi(Z_{1-\beta})$$

$$z_{1-\beta} = \left[|\Delta| \sqrt{n - \frac{z_{1-\alpha/2}^2}{2}} \right] - z_{1-\alpha/2}$$

Donde:

- n es el número de pares en la muestra.

1.2.2.1.- COMPARACIÓN DE PROPORCIONES, GRUPOS INDEPENDIENTES

[Machin (1997, p. 19-20)]

Tamaños de muestra para aplicar el test χ^2 :

$$n_1 = \frac{\left(z_{1-\alpha/2} \sqrt{(1+\phi)\bar{P}(1-\bar{P})} + z_{1-\beta} \sqrt{\phi P_1(1-P_1) + P_2(1-P_2)} \right)^2}{\phi(P_1 - P_2)^2}; n_2 = \phi n_1$$

Tamaños de muestra para aplicar el test χ^2 con la corrección por continuidad de Yates (χ_c^2) o el test exacto de Fisher:

$$m_1 = \frac{n_1}{4} \left[1 + \sqrt{1 + \frac{2(1+\phi)}{\phi n_1 |P_1 - P_2|}} \right]^2; m_2 = \phi m_1$$

Donde:

- P_i es la proporción esperada en la población i , $i=1, 2$,
- ϕ es la razón entre los dos tamaños muestrales,
- $\bar{P} = \frac{P_1 + \phi P_2}{1 + \phi}$.

Potencia del test χ^2 :

$$1 - \beta = \Phi(Z_{1-\beta})$$

$$z_{1-\beta} = \frac{|P_1 - P_2| \sqrt{\phi n_1} - z_{1-\alpha/2} \sqrt{(1+\phi)\bar{P}(1-\bar{P})}}{\sqrt{\phi P_1(1-P_1) + P_2(1-P_2)}}$$

Potencia del test χ^2 con la corrección por continuidad de Yates (χ_c^2) o del test exacto de Fisher:

$$1 - \beta = \Phi(Z_{1-\beta})$$

$$Z_{1-\beta} = \frac{|P_1 - P_2| \sqrt{\phi \left(m_1 - \lambda + \frac{\lambda^2}{4m_1} \right)} - z_{1-\alpha/2} \sqrt{(1+\phi)\bar{P}(1-\bar{P})}}{\sqrt{\phi P_1(1-P_1) + P_2(1-P_2)}}$$

Donde:

- $\lambda = \frac{1 + \phi}{\phi |P_1 - P_2|}$,
- $n_1 = \frac{n}{1 + \phi}$ y $m_1 = \frac{m}{1 + \phi}$, donde n ó m es el tamaño total de la muestra.

1.2.2.2.- COMPARACIÓN DE PROPORCIONES, GRUPOS EMPAREJADOS [Connor (1987)]

Número de pares para aplicar el test de McNemar:

$$n = \frac{\left(z_{1-\alpha/2} \sqrt{P_d} + z_{1-\beta} \sqrt{P_d - (P_1 - P_2)^2} \right)^2}{(P_1 - P_2)^2}$$

Donde:

- P_i es la proporción esperada en el grupo i, i=1, 2,
- $P_d = P_1(1 - P_2) + P_2(1 - P_1)$ es la proporción de pares discordantes.

Potencia:

$$1 - \beta = \Phi(Z_{1-\beta})$$

$$Z_{1-\beta} = \frac{|P_1 - P_2| \sqrt{n} - z_{1-\alpha/2} \sqrt{P_d}}{\sqrt{P_d - (P_1 - P_2)^2}}$$

Donde:

- n es el número de pares en la muestra.

1.2.3.1.- ESTUDIOS DE CASOS Y CONTROLES, GRUPOS INDEPENDIENTES [Machin (1997, p. 19-20)]

Se aplican las fórmulas del epígrafe 1.2.2.1, donde:

- n_1 ó m_1 es el número de casos en la muestra,
- n_2 ó m_2 es el número de controles en la muestra,
- ϕ es el nº de controles por caso,
- P_1 es la proporción de casos expuestos,
- P_2 es la proporción de controles expuestos,
- P_1 y P_2 se relacionan con OR del modo siguiente:

$$\text{▪ } P_1 = \frac{OR P_2}{(1 - P_2) + OR P_2}, P_2 = \frac{P_1}{OR(1 - P_1) + P_1}.$$

1.2.3.2.- ESTUDIOS DE CASOS Y CONTROLES, GRUPOS EMPAREJADOS [Connor (1987)]

Se aplican las fórmulas del epígrafe 1.2.2.2, donde:

- P_1 es la proporción de casos expuestos,
- P_2 es la proporción de controles expuestos,
- P_1 y P_2 se relacionan con OR del modo siguiente:

$$\text{▪ } P_1 = \frac{OR P_2}{(1 - P_2) + OR P_2}, P_2 = \frac{P_1}{OR(1 - P_1) + P_1}.$$

1.2.4.- ESTUDIOS DE COHORTE [Machin (1997, p. 19-20)]

Se aplican las fórmulas del epígrafe 1.2.2.1, donde:

- n_1 ó m_1 es el tamaño de la muestra de expuestos,
- n_2 ó m_2 es el tamaño de la muestra de no expuestos,
- P_1 es el riesgo en expuestos,
- P_2 es el riesgo en no expuestos,
- P_1 y P_2 se relacionan con RR del modo siguiente: $P_1 = P_2 RR$, $P_2 = \frac{P_1}{RR}$,
- ϕ es la razón entre el tamaño muestral de no expuestos y el de expuestos.

1.2.5.1.- EQUIVALENCIA DE MEDIAS [Machin (1997, p. 102)]

Tamaños de muestra para un contraste unilateral con varianzas iguales:

$$n_1 = \frac{\sigma^2(1+\phi)(z_{1-\alpha} + z_{1-\beta})^2}{\phi(|d| - \varepsilon)^2}$$

$$n_2 = \phi n_1$$

Donde:

- d es la diferencia esperada entre las dos medias,
- σ es la varianza común esperada en las dos poblaciones,
- ε es la diferencia máxima que indica equivalencia entre las medias,
- ϕ es la razón entre los dos tamaños muestrales.

Potencia:

$$1 - \beta = \Phi(z_{1-\beta})$$

$$z_{1-\beta} = \left(\frac{|d - \varepsilon|}{\sigma} \sqrt{\frac{\phi n_1}{1 + \phi}} \right) - z_{1-\alpha}$$

Donde:

- $n_1 = \frac{n}{1 + \phi}$,
- n es el tamaño total de la muestra.

1.2.5.2.- EQUIVALENCIA DE PROPORCIONES [Machin (1997, p. 101-102)]

Tamaños de muestra para un contraste unilateral:

$$n_1 = \frac{\left(z_{1-\alpha} \sqrt{\phi \bar{P}_1(1 - \bar{P}_1) + \bar{P}_2(1 - \bar{P}_2)} + z_{1-\beta} \sqrt{\phi P_1(1 - P_1) + P_2(1 - P_2)} \right)^2}{\phi(|P_1 - P_2| - \varepsilon)^2}$$

$$n_2 = \phi n_1$$

Donde:

- P_i es la proporción esperada en la población i , $i=1, 2$,
- ε es la diferencia máxima que indica equivalencia entre las proporciones,
- ϕ es la razón entre los dos tamaños muestrales,
- $\bar{P}_1 = 2u \cos(w) - \frac{b}{3a}$ y $\bar{P}_2 = \bar{P}_1 - \varepsilon$,
- $w = \frac{1}{3} \left[\pi + \arccos\left(\frac{v}{u^3}\right) \right]$,
- $v = \left(\frac{b}{3a}\right)^3 - \frac{bc}{6a^2} + \frac{d}{2a}$,
- $u = \text{signo}(v) \sqrt{\left(\frac{b}{3a}\right)^2 - \frac{c}{3a}}$,
- $a = \frac{1+\phi}{\phi}$,
- $b = -\frac{1}{\phi} [1 + \phi + \phi P_1 + P_2 + \varepsilon(1 + 2\phi)]$,
- $c = \frac{1}{\phi} [\phi \varepsilon^2 + \varepsilon(2\phi P_1 + 1 + \phi) + \phi P_1 + P_2]$,
- $d = -P_1 \varepsilon(1 + \varepsilon)$.

Potencia:

$$1 - \beta = \Phi(z_{1-\beta})$$

$$z_{1-\beta} = \frac{||P_1 - P_2| - \varepsilon| \sqrt{\phi n_1} - z_{1-\alpha} \sqrt{\phi \bar{P}_1(1 - \bar{P}_1) + \bar{P}_2(1 - \bar{P}_2)}}{\sqrt{\phi P_1(1 - P_1) + P_2(1 - P_2)}}$$

Donde:

- $n_1 = \frac{n}{1 + \phi}$,
- n es el tamaño total de la muestra.

1.2.6.1.- PRUEBAS DIAGNÓSTICAS, GRUPOS INDEPENDIENTES [Machin (1997, p. 19-20)]

Esta opción calcula el tamaño de muestra necesario para comparar la sensibilidad y/o la especificidad de dos pruebas diagnósticas en grupos independientes de igual tamaño mediante el test χ^2 .

Sensibilidad

Número de enfermos para aplicar el test χ^2 :

$$n_E = \frac{\left(z_{1-\alpha/2} \sqrt{2\bar{\theta}(1-\bar{\theta})} + z_{1-\beta} \sqrt{\theta_1(1-\theta_1) + \theta_2(1-\theta_2)} \right)^2}{(\theta_1 - \theta_2)^2}$$

Número de enfermos para aplicar el test χ^2 con la corrección por continuidad de Yates o el test exacto de Fisher:

$$m_E = \frac{n_E}{4} \left[1 + \sqrt{1 + \frac{4}{n_E |\theta_1 - \theta_2|}} \right]^2$$

Si se conoce a priori la condición de enfermo se calculan los tamaños de enfermos y no enfermos:

$$n_E \text{ y } n_{NE} = \phi n_E$$

$$m_E \text{ y } m_{NE} = \phi m_E$$

Si no se conoce a priori la condición de enfermo, se calcula el tamaño total de muestra:

$$n = \frac{n_E}{P} \text{ ó } m = \frac{m_E}{P}$$

Donde:

- θ_i es la sensibilidad esperada para la prueba i , $i=1,2$,
- $\bar{\theta} = \frac{\theta_1 + \theta_2}{2}$,
- ϕ es la razón entre el tamaño muestral de no enfermos y el de enfermos,
- P es la prevalencia de enfermedad en la población.

Potencia:

$$1 - \beta = \Phi(Z_{1-\beta})$$

- para el test χ^2 :

$$z_{1-\beta} = \frac{|\theta_1 - \theta_2| \sqrt{n_E} - z_{1-\alpha/2} \sqrt{2\bar{\theta}(1-\bar{\theta})}}{\sqrt{\theta_1(1-\theta_1) + \theta_2(1-\theta_2)}}$$

- para el test χ^2 con la corrección por continuidad de Yates o el test exacto de Fisher:

$$z_{1-\beta} = \frac{|\theta_1 - \theta_2| \sqrt{\left(m_E - \lambda + \frac{\lambda^2}{4m_E}\right) - z_{1-\alpha/2} \sqrt{2\bar{\theta}(1-\bar{\theta})}}}{\sqrt{\theta_1(1-\theta_1) + \theta_2(1-\theta_2)}}$$

Donde:

- $n_E = \frac{n}{1+\phi}$ ó $m_E = \frac{m}{1+\phi}$ si se conoce a priori la condición de enfermo,
- $n_E = nP$ ó $m_E = mP$ si no se conoce a priori la condición de enfermo,
- n es el tamaño total de la muestra de cada grupo (los dos iguales),
- $\lambda = \frac{2}{|\theta_1 - \theta_2|}$.

Especificidad

Las fórmulas para calcular el tamaño de muestra son las mismas que para la sensibilidad, con la diferencia de que se obtienen los no enfermos n_{NE} y m_{NE} , respectivamente, para el test χ^2 con o sin corrección por continuidad. También coinciden las fórmulas para la potencia, sustituyendo n_E y m_E por n_{NE} y m_{NE} , respectivamente.

Si se conoce a priori la condición de enfermo se calculan los tamaños de no enfermos y enfermos:

$$n_{NE} \text{ y } n_E = \frac{n_{NE}}{\phi}$$

$$m_{NE} \text{ y } m_E = \frac{m_{NE}}{\phi}$$

Si no se conoce a priori la condición de enfermo, se calcula el tamaño total de muestra:

$$n = \frac{n_{NE}}{1-P} \text{ ó } m = \frac{m_{NE}}{1-P}$$

Para la potencia, se calcula primero el tamaño de muestra de no enfermos a partir del tamaño total de muestra:

$$n_{NE} = \frac{\phi n}{1+\phi} \text{ ó } m_{NE} = \frac{\phi m}{1+\phi} \text{ si se conoce a priori la condición de enfermo}$$

$$n_{NE} = n(1-P) \text{ ó } m_{NE} = m(1-P) \text{ si no se conoce a priori la condición de enfermo}$$

Sensibilidad y especificidad

Si se conoce a priori la condición de enfermo se calcula el tamaño de enfermos para sensibilidad y el de no enfermos para especificidad.

Si no se conoce a priori la condición de enfermo, se calcula el tamaño total de muestra:

$$n = \text{máx} \left\{ \frac{n_E}{P}, \frac{n_{NE}}{1-P} \right\}$$

La potencia se calcula por separado para sensibilidad y especificidad.

1.2.6.2.- PRUEBAS DIAGNÓSTICAS, GRUPOS EMPAREJADOS [Connor (1987)]Sensibilidad

Se aplican las fórmulas del epígrafe 2.2.2 para obtener el número de enfermos, n_E , donde:

- P_1 es la sensibilidad esperada para la prueba 1,
- P_2 es la sensibilidad esperada para la prueba 2.

Si se conoce a priori la condición de enfermo se calculan los tamaños de enfermos y no enfermos:

$$n_E \text{ y } n_{NE} = \phi n_E$$

Si no se conoce a priori la condición de enfermo, se calcula el tamaño total de muestra:

$$n = \frac{n_E}{P}$$

Para la potencia se calcula primero el tamaño de muestra de enfermos a partir del tamaño total de muestra:

$$n_E = \frac{n}{1 + \phi}, \text{ si se conoce a priori la condición de enfermo}$$

$$n_E = nP, \text{ si no se conoce a priori la condición de enfermo}$$

Donde:

- ϕ es la razón entre el tamaño muestral de no enfermos y el de enfermos,
- P es la prevalencia de enfermedad en la población,

Especificidad

Se aplican las fórmulas del epígrafe 2.2.2 para obtener el número de no enfermos, n_{NE} , donde:

- P_1 es la especificidad esperada para la prueba 1,
- P_2 es la especificidad esperada para la prueba 2.

Si se conoce a priori la condición de enfermo se calculan los tamaños de no enfermos y enfermos:

$$n_{NE} \text{ y } n_E = \frac{n_{NE}}{\phi}$$

Si no se conoce a priori la condición de enfermo, se calcula el tamaño total de muestra:

$$n = \frac{n_{NE}}{1-P}$$

Para la potencia, se calcula primero el tamaño de muestra de no enfermos a partir del tamaño total de muestra:

$$n_{NE} = \frac{\phi n}{1+\phi}, \text{ si se conoce a priori la condición de enfermo}$$

$$n_{NE} = n(1-P), \text{ si no se conoce a priori la condición de enfermo}$$

Donde:

- ϕ es la razón entre el tamaño muestral de no enfermos y el de enfermos,
- P es la prevalencia de enfermedad en la población,

Sensibilidad y especificidad

Si se conoce a priori la condición de enfermo se calcula el tamaño de enfermos para sensibilidad y el de no enfermos para especificidad.

Si no se conoce a priori la condición de enfermo, se calcula el tamaño total de muestra:

$$n = \text{máx} \left\{ \frac{n_E}{P}, \frac{n_{NE}}{1-P} \right\}$$

La potencia se calcula por separado para sensibilidad y especificidad.

1.2.7.- CALIDAD DE LOTES [Lwanga (1991, p. 63, 69)]

Tamaño de muestra:

$$n = \frac{\left(z_{1-\alpha}\sqrt{P_0(1-P_0)} + z_{1-\beta}\sqrt{P_a(1-P_a)}\right)^2}{(P_0 - P_a)^2}$$

Valor umbral de sujetos con la característica:

$$d = nP_0 - z_{1-\alpha}\sqrt{nP_0(1-P_0)}$$

Donde:

- P_0 es la proporción umbral de defectuosos,
- P_a es la proporción de defectuosos esperada en la población.

1.2.8.- SUPERVIVENCIA

Comparación de dos curvas de supervivencia [Schoenfeld (1983)]:

Esta opción calcula el tamaño de muestra necesario para comparar dos curvas de supervivencia mediante el test log-rank, asumiendo un modelo de riesgos proporcionales para los tiempos de supervivencia.

Número de eventos necesarios:

$$E = \left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 \frac{1}{\phi} \left(\frac{1+\phi}{\ln \Delta}\right)^2$$

Tamaño total de muestra (corregido por las pérdidas de seguimiento):

$$n_c = \frac{n}{1-P_c}, \text{ con } n = \frac{E}{P_e}$$

Tamaño de muestra de cada grupo:

$$n_1 = \frac{n_c}{1+\phi} \text{ y } n_2 = \frac{\phi n_c}{1+\phi}$$

Donde:

- ϕ es la razón entre los dos tamaños muestrales (n_2/n_1),
- $\Delta = \frac{\ln P_2}{\ln P_1}$ es el hazard ratio,

- P_i es la probabilidad de supervivencia en el grupo i , $i=1,2$,
- P_c es la proporción de pérdidas de seguimiento,
- $P_e = \frac{(1-P_1) + \phi(1-P_2)}{1+\phi}$ es la probabilidad de ocurrencia de un evento durante el período de estudio.

Comparación de tres o más curvas de supervivencia [Ahnn (1995)]:

Esta opción calcula el tamaño de muestra necesario para comparar k curvas de supervivencia mediante el test log-rank, asumiendo un modelo de riesgos proporcionales para los tiempos de supervivencia y que todos los grupos son de igual tamaño.

Número de eventos necesarios:

$$E = \frac{\delta k^2}{(k-1) \sum_{j=2}^k (\ln \Delta_j)^2 - 2 \sum_{j=2}^k \sum_{q>j} \ln \Delta_j \ln \Delta_q}$$

Tamaño total de muestra (corregido por las pérdidas de seguimiento):

$$n_c = \frac{n}{1-P_c}, \text{ con } n = \frac{E}{P_e}$$

Tamaño de muestra de cada grupo:

$$n_j = \frac{n_c}{k}, \forall j=1,\dots,k$$

Donde:

- k es el número de grupos que se comparan,
- $\Delta_j = \frac{\ln P_j}{\ln P_1}$, $j=2,\dots,k$,
- P_j es la probabilidad de supervivencia en el grupo j , $j=1,2,\dots,k$,
- δ es el parámetro de no centralización de la distribución del estadístico bajo la hipótesis alternativa: ji-cuadrado no central con $k-1$ grados de libertad. Depende del nivel de confianza y de la potencia del test [Narula y Desu (1981)],
- P_c es la proporción de pérdidas de seguimiento,

- $P_e = \frac{1}{k} \sum_{j=1}^k (1 - P_j)$ es la probabilidad de ocurrencia de un evento durante el período de estudio.

1.2.9.- COEFICIENTE DE CORRELACIÓN [Machin (1997, p. 168-9)]

El tamaño de muestra se calcula siguiendo un procedimiento iterativo:

Paso 1:

$$U_1 = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right)$$

$$n_1 = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{U_1^2} + 3$$

Paso i (i>1):

$$U_i = U_1 + \frac{\rho}{2(n_{i-1} - 1)}$$

$$n_i = \frac{(z_{1-\alpha} + z_{1-\beta})^2}{U_i^2} + 3$$

Si $|n_i - n_{i-1}| < 1$ se para el proceso y se toma $n=n_i$

Potencia:

$$1 - \beta = \Phi(z_{1-\beta})$$

$$z_{1-\beta} = (U\sqrt{n-3}) - z_{1-\alpha}$$

Donde:

- ρ es el coeficiente de correlación a detectar,
- $U = \frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right) + \frac{\rho}{2(n-1)}$.

2.- SELECCIÓN DE MUESTRAS

2.1.- MUESTREO SIMPLE ALEATORIO

Algoritmo de selección sin reemplazamiento:

- Se asigna un número de 1 a N a cada elemento de la población,
- Paso 1: se selecciona un número aleatorio entre 1 y N,
- Paso k+1: se selecciona un número aleatorio entre los N-k números no seleccionados en los pasos anteriores, k=1, ..., n-1.

Probabilidad de selección (%):

$$f = \frac{n}{N} \times 100$$

Donde:

- N es el tamaño de la población,
- n es el tamaño de la muestra.

2.2.- MUESTREO SISTEMÁTICO [Silva (2000, p. 126-131)]

Muestreo sistemático regular:

- Se asigna un número de 1 a N a cada elemento de la población,
- Se calcula el intervalo de selección $k = \text{Parte entera} \left(\frac{N}{n} \right)$,
- Se selecciona el arranque aleatorio r entre 1 y k,
- Los números seleccionados son los de la secuencia $\{r+jk, j=0, 1, \dots, s-1\}$, donde s es el menor entero tal que $r+sk > N$.

Probabilidad de selección (%):

$$f = \frac{n}{N} \times 100, \text{ siendo } n \text{ el tamaño de muestra resultante.}$$

Muestreo sistemático en fases:

- Se asigna un número de 1 a N a cada elemento de la población,
- Se calcula h:

$$h = \begin{cases} \frac{f}{100} & \text{si } f < 51 \\ 1 - \frac{f}{100} & \text{si } f > 50 \end{cases}$$

- En la Tabla 1 se buscan los valores de k_1 y k_2 que corresponden a h,
- Se selecciona un número aleatorio entre 1 y k_1 (r_1) y, si $k_2 \neq 0$, un número aleatorio entre 1 y $|k_2|$ (r_2),
- Se calcula la secuencia de números: $S_1 = \{r_1 + jk_1, j=0, 1, \dots, s-1\}$, donde s es el menor entero tal que $r_1 + sk_1 > N$,
- Si $k_2 = 0$, se acaba el proceso, y los números seleccionados son los de S_1 ,
- Si $k_2 > 0$, a la secuencia S_1 se adicionan algunos valores tomados de entre los no seleccionados. Para ello, se consideran los N-s números resultantes de eliminar la secuencia S_1 de los números 1 a N. En esa nueva lista se seleccionan los números S_2 que ocupan las posiciones $r_2 + jk_2$, con $j=0, 1, \dots, t-1$, donde t es el menor entero tal que $r_2 + tk_2 > N-s$. Los números seleccionados son la unión de las dos secuencias S_1 y S_2 .
- Si $k_2 < 0$, de la secuencia S_1 se eliminan los números que ocupan las posiciones $r_2 + j|k_2|$, con $j=0, 1, \dots, g-1$, donde g es el menor entero tal que $r_2 + g|k_2| > s$. Los números seleccionados son los que quedan en S_1 luego de eliminar esos valores.
- **Nota:** Si $f > 50$, los números seleccionados corresponden al complementario de los números resultantes del proceso anterior respecto del conjunto de los N primeros números naturales.

Donde:

- N es el tamaño de la población,
- n es el tamaño de la muestra,
- f es la probabilidad de selección o fracción de muestreo.

Tabla 1.- Intervalos de selección (k_1 , k_2) para aplicar el muestreo sistemático en fases con probabilidad de selección h .

h	k_1	k_2	h	k_1	k_2
.01	50	-2	.26	6	9
.02	50	0	.27	4	38
.03	33	0	.28	4	25
.04	25	0	.29	4	19
.05	20	0	.30	3	-10
.06	16	-25	.31	3	-14
.07	14	-50	.32	3	-25
.08	12	-25	.33	3	0
.09	10	-10	.34	3	60
.10	10	0	.35	3	40
.11	11	48	.36	3	25
.12	8	-25	.37	3	18
.13	7	-11	.38	3	14
.14	7	-50	.39	3	12
.15	6	-10	.40	2	-5
.16	6	-25	.41	3	9
.17	7	32	.42	2	-6
.18	5	-10	.43	2	-7
.19	5	-20	.44	2	-8
.20	5	0	.45	2	-10
.21	9	9	.46	2	-12
.22	5	40	.47	2	-17
.23	5	27	.48	2	-25
.24	4	-25	.49	2	-50
.25	4	0	.50	2	0

2.3.- MUESTREO ALEATORIO ESTRATIFICADO [Silva (2000, p. 126-131)]

Tamaño total de la población:

$$N = \sum_{h=1}^L N_h$$

Algoritmo de selección:

- En el estrato h ($h=1, \dots, L$) se selecciona una muestra simple aleatoria de tamaño n_h .

Probabilidad de selección (%) en el estrato h :

$$f_h = \frac{n_h}{N_h} \times 100, h=1, \dots, L$$

Donde:

- L es el número de estratos,
- N_h es el tamaño poblacional del estrato h , $h=1, \dots, L$,
- n_h es el tamaño de muestra del estrato h , $h=1, \dots, L$. Si el reparto es proporcional:

$$n_h = n \frac{N_h}{N}, h=1, \dots, L, \text{ siendo } n \text{ el tamaño total de la muestra.}$$

2.4.- MUESTREO POR CONGLOMERADOS MONOETÁPICO [Silva (2000, p. 157-159)]

Tamaño total de la población:

$$N = \sum_{k=1}^M N_k$$

Algoritmo de selección:

- Se selecciona una muestra simple aleatoria de m conglomerados.

Probabilidad de selección (%):

$$f = \frac{m}{M} \times 100$$

Donde:

- N_k es el tamaño poblacional del conglomerado k , $k=1, \dots, M$,
- M es el número de conglomerados en la población,
- n es el tamaño de muestra previsto,
- m es el número de conglomerados a seleccionar: $m = n \frac{M}{N}$

2.5.- MUESTREO POR CONGLOMERADOS BIETÁPICO [Silva (2000, 177-180)]

Tamaño total de la población:

$$N = \sum_{k=1}^M N_k$$

Algoritmo de selección en la primera etapa. Hay dos opciones:

1. Se seleccionan m conglomerados con probabilidades proporcionales a sus tamaños N_k , $k=1, \dots, M$:
 - Se calculan los tamaños acumulados de los conglomerados:
 - $C_k = \sum_{i=1}^k M_i$, $k=1, \dots, M$
 - Se calcula el intervalo de selección $I = \text{Parte entera} \left(\frac{N}{m} \right)$,
 - Se selecciona el arranque aleatorio r entre 1 e I ,
 - Se calculan los números $z_j = r + jI$, $j=0, 1, \dots, m-1$,
 - Para cada z_j se selecciona el primer conglomerado k que cumpla $C_k \geq z_j$.
2. Se selecciona una muestra simple aleatoria de m conglomerados.

Algoritmo de selección en la segunda etapa:

- En cada conglomerado de la muestra, se selecciona una muestra simple aleatoria de tamaño u_k .

Probabilidad de selección (%) en el conglomerado k:

1. $f_k = \frac{m}{N} u_k \times 100$, $k=1, \dots, M$, si los conglomerados se seleccionan con probabilidades proporcionales al tamaño,
2. $f_k = \frac{m}{M} \frac{u_k}{N_k} \times 100$, $k=1, \dots, M$, si se selecciona una muestra simple aleatoria de conglomerados.

Donde:

- N_k es el tamaño poblacional del conglomerado k, $k=1, \dots, M$,
- M es el número de conglomerados en la población,
- m es el número de conglomerados a seleccionar,
- u_k es el tamaño de muestra a seleccionar en el conglomerado k, $k=1, \dots, M$.

2.6.- MUESTREO POR CONGLOMERADOS MONOETÁPICO ESTRATIFICADO

Tamaño de la población de cada estrato:

$$N_h = \sum_{k=1}^{M_h} N_{hk}$$

Tamaño total de la población:

$$N = \sum_{h=1}^L N_h$$

Algoritmo de selección:

- En el estrato h ($h=1, \dots, L$) se selecciona una muestra simple aleatoria de m_h conglomerados.

Probabilidad de selección (%) en el estrato h:

$$f_h = \frac{m_h}{M_h} \times 100, h=1, \dots, L$$

Donde:

- L es el número de estratos,
- M_h es el número de conglomerados en el estrato h, $h=1, \dots, L$,
- N_{hk} es el tamaño poblacional del conglomerado k del estrato h, $k=1, \dots, M_h$, $h=1, \dots, L$,
- m_h es el número de conglomerados a seleccionar en el estrato h, $h=1, \dots, L$.
 - Si el reparto es proporcional a partir del número total de conglomerados en la muestra, m:

$$m_h = m \frac{N_h}{N}, h=1, \dots, L$$

- Si m_h se calcula a partir del tamaño de muestra del estrato h, n_h :

$$m_h = n_h \frac{M_h}{N_h}, h=1, \dots, L$$

2.7.- MUESTREO POR CONGLOMERADOS BIETÁPICO ESTRATIFICADO

[Silva (2000, p. 192-194)]

Tamaño de la población de cada estrato:

$$N_h = \sum_{k=1}^{M_h} N_{hk}$$

Tamaño total de la población:

$$N = \sum_{h=1}^L N_h$$

Algoritmo de selección en el muestreo equiprobabilístico:

1. Primera etapa: en el estrato h ($h=1, \dots, L$) se seleccionan m_h conglomerados con probabilidades proporcionales a sus tamaños N_{hk} , $k=1, \dots, M_h$ (véase sección 2.5). El reparto por estratos del número total de conglomerados en la muestra, m, es proporcional:

$$m_h = m \frac{N_h}{N}, h=1, \dots, L$$

2. Segunda etapa: en cada conglomerado de la muestra del estrato h ($h=1, \dots, L$), se selecciona una muestra simple aleatoria de tamaño u .

Algoritmo general de selección:

1. Primera etapa. Hay dos opciones:

1.1. En el estrato h ($h=1, \dots, L$) se seleccionan m_h conglomerados con probabilidades proporcionales a sus tamaños N_{hk} , $k=1, \dots, M_h$ (véase sección 2.5).

1.2. En el estrato h ($h=1, \dots, L$) se selecciona una muestra simple aleatoria de m_h conglomerados.

2. Segunda etapa: en cada conglomerado de la muestra del estrato h ($h=1, \dots, L$), se selecciona una muestra simple aleatoria de tamaño u_{hk} .

Probabilidad de selección (%) en el conglomerado k del estrato h :

$$1. f_k = \frac{m_h}{N_h} u_{hk} \times 100, k=1, \dots, M_h, h=1, \dots, L, \text{ si los conglomerados se seleccionan}$$

con probabilidades proporcionales al tamaño,

$$2. f_k = \frac{m_h}{M_h} \frac{u_{hk}}{N_{hk}} \times 100, k=1, \dots, M_h, h=1, \dots, L, \text{ si se selecciona una muestra simple}$$

aleatoria de conglomerados,

Donde:

- L es el número de estratos,
- M_h es el número de conglomerados en el estrato h , $h=1, \dots, L$,
- N_{hk} es el tamaño poblacional del conglomerado k del estrato h , $k=1, \dots, M_h$, $h=1, \dots, L$,
- m es el número total de conglomerados a seleccionar,
- m_h es el número de conglomerados a seleccionar en el estrato h , $h=1, \dots, L$,
- u es el tamaño de muestra a seleccionar en cada conglomerado, igual para todos,
- u_{hk} es el tamaño de muestra a seleccionar en el conglomerado k del estrato h , $k=1, \dots, M_h$, $h=1, \dots, L$.

3.- ASIGNACIÓN DE SUJETOS A TRATAMIENTOS [Silva (2000, p. 248-249)]

Pasos para formar grupos de igual tamaño:

- Se calcula $m = \frac{n}{k}$
- Paso 1: se selecciona una muestra simple aleatoria de tamaño m entre los números de 1 a n . Los sujetos de esta muestra se asignan al grupo 1.
- Paso j ($j=2, \dots, k-1$): se selecciona una muestra simple aleatoria de tamaño m entre los números de 1 a n que no han sido asignados a los grupos 1, 2, .. $j-1$. Los sujetos de esta muestra se asignan al grupo j .
- Paso k : tras $k-1$ selecciones, quedan m números que se asignan al grupo k .

Pasos para formar grupos equilibrados:

- Sujetos del 1 al k : se asigna aleatoriamente un sujeto a uno cada de los k grupos.
- Sujeto $s+1$: se supone que se han asignado s_i sujetos al grupo i ($i=1, \dots, k$) y

$$s = \sum_{i=1}^k s_i :$$

- Se calculan: $w_i = \frac{1}{s_i}$, $w = \sum_{i=1}^k w_i$, $w_i^* = \frac{w_i}{w}$,
- Se selecciona un número aleatorio r entre 0 y 1,
- Si $r < w_1^*$, el sujeto $s+1$ se asigna al grupo 1,
- Si $\sum_{i=1}^{j-1} w_i^* < r \leq \sum_{i=1}^j w_i^*$, el sujeto $s+1$ se asigna al grupo j ($j=2, \dots, k$).

Donde:

- k es el número de tratamientos,
- n es el número de sujetos.

4- ESTIMACIÓN CON MUESTRAS COMPLEJAS [Silva (2000, p. 195-196)]

Estimador sin ponderar:

$$\bar{x} = \frac{1}{n} \sum_{h=1}^L \sum_{i=1}^{m_h} \sum_{j=1}^{n_{hi}} x_{hij}$$

Estimador ponderado (estimador de razón):

$$r = \frac{x}{y}$$

Error estándar:

$$EE = \sqrt{\text{Var}(r)}$$

$$\text{Var}(r) = \frac{1}{y^2} \left(\sum_{h=1}^L U_h + r^2 \sum_{h=1}^L V_h - 2r \sum_{h=1}^L W_h \right)$$

Intervalo de confianza con nivel de confianza $(1-\alpha)\%$:

$$\left(r - z_{1-\alpha/2} EE, r + z_{1-\alpha/2} EE \right)$$

Efecto de diseño:

$$deff = \sqrt{\frac{n \text{Var}(r)}{s^2}}$$

Donde:

- x_{hij} es el valor de la variable X en el sujeto j del conglomerado i del estrato h,
- w_{hij} es la ponderación del sujeto j del conglomerado i del estrato h,
- L es el número de estratos,
- m_h es el número de conglomerados del estrato h, $h=1, \dots, L$,
- n_{hi} es el número de sujetos del conglomerado i, $i=1, \dots, m_h$, del estrato h, $h=1, \dots, L$,

- n es el número total de observaciones: $n = \sum_{h=1}^L \sum_{i=1}^{m_h} n_{hi}$,

- $x = \sum_{h=1}^L x_h$, $x_h = \sum_{i=1}^{m_h} x_{hi}$, $x_{hi} = \sum_{j=1}^{n_{hi}} w_{hij} x_{hij}$,

- $y = \sum_{h=1}^L y_h$, $y_h = \sum_{i=1}^{m_h} y_{hi}$, $y_{hi} = \sum_{j=1}^{n_{hi}} w_{hij} y_{hij}$,

- $$U_h = \frac{1}{m_h - 1} \left(m_h \sum_{i=1}^{m_h} x_{hi}^2 - x_h^2 \right),$$
- $$V_h = \frac{1}{m_h - 1} \left(m_h \sum_{i=1}^{m_h} y_{hi}^2 - y_h^2 \right),$$
- $$W_h = \frac{1}{m_h - 1} \left(m_h \sum_{i=1}^{m_h} x_{hi} y_{hi} - x_h y_h \right),$$
- $$s^2 = \frac{1}{n - 1} \sum_{h=1}^L \sum_{i=1}^{m_h} \sum_{j=1}^{n_{hi}} (x_{hij} - \bar{x})^2,$$
- $z_{1-\alpha/2}$ es el percentil de la distribución normal estándar, $N(0,1)$, que deja a la izquierda una cola de probabilidad $1 - \frac{\alpha}{2}$,
- $1-\alpha$ es el nivel de confianza.

Bibliografía

- Ahnn S, Anderson SJ. Sample size determination for comparing more than two survival distributions. *Stat Med.* 1995;14:2273-82.
- Cantor AB. Sample size calculations for Cohen's kappa. *Psychological Methods.* 1996;1(2):150-3.
- Connor RJ. Sample size for testing differences in proportions for the paired-sample design. *Biometrics.* 1987;43:207-11.
- Freedman LS. Tables of the number of patients required in clinical trials using the logrank test. *Stat Med.* 1982;1:121-9.
- Haynam GE, Govindarajulu Z, Leone FC. Tables of the cumulative non-central chi-square distribution. En: *Selected tables in mathematical statistics, vol.1.* American Mathematical Society; 1970. pp. 1-78.
- Lwanga SK, Lemeshow S. Determinación del tamaño de las muestras en los estudios sanitarios. Manual práctico. Ginebra: OMS. 1991.
- Machin D, Campbell MJ, Fayers PM, Pinol APY. Sample size tables for clinical studies. 2ª ed. Blackwell Science Ltd. 1997.
- Narula SC, Desu MM. Algorithm AS 170: Computation of probability and non-centrality parameter of a non-central chi-squared distribution. *Journal of the Royal Statistical Society.* 1981;30(3):349-52.
- Obuchowski NA. Sample size calculations in studies of test accuracy. *Statistical Methods in Medical Research.* 1998;7:371-92.
- Rosner B. *Fundamentals of biostatistics.* 5ª ed. Belmont, CA: Duxbury Press. 2000.
- Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. *Biometrics.* 1983;39:499-503.
- Schouten HJA. Sample size formula with a continuous outcome for unequal group sizes and unequal variances. *Stat Med.* 1999;18:87-91.
- Silva LC. *Diseño razonado de muestras y captación de datos para la investigación sanitaria.* Madrid: Díaz de Santos. 2000.