

Programa de Muestreo Estadístico en Sanidad Animal (ProMESA)

León, Emilio A.; Duffy, Sergio J.

INTA- CICVyA- Inst. de Patobiología, Castelar, Prov. de Bs. As., ARGENTINA

Resumen

En respuesta a la necesidad creciente de contar con datos objetivos acerca del estatus sanitario de la ganadería, tanto a nivel de un establecimiento como de un área o un país, se ha desarrollado el programa ProMESA 1.3. El mismo está destinado a profesionales involucrados en el diseño, implementación y análisis de muestreos en poblaciones animales. Funciona mediante los programas MS Access y MS VisualBasic.

Permite el cálculo del número de unidades que deben ser incluidas en una muestra, para estimar una prevalencia o un promedio, a partir de diseño aleatorio simple, estratificado o en dos etapas. También calcula el tamaño de muestra para demostrar la ausencia de un evento, por diseño aleatorio simple o en dos etapas. Realiza la selección aleatoria de unidades a partir de una lista, de manera simple o con probabilidad proporcional al tamaño. Una vez implementado el muestreo, calcula los intervalos de confianza de los resultados, de acuerdo al diseño de muestreo seguido.

Hasta el presente ha sido utilizado por institutos de investigación y por servicios veterinarios oficiales. También ha demostrado ser de utilidad académica.

Abstract

The need of counting with objective data about animal health information on livestock at farm, local or regional level, is constantly growing. ProMESA 1.3 was developed in order to support this need. It is a computer program aimed to help professionals involved in designing, implementing and analyzing samplings in animal populations. It works using software MS Access and MS VisualBasic.

ProMESA 1.3 allows calculation of the number of units that should be included in a sample in order to estimate a mean or an average, from simple random, stratified or two stage design. Sample size needed to demonstrate the absence of an event is also calculated, for simple random or two stage design. Simple random selection or random selection with probability proportional to size from a sample frame may be implemented. Confidence intervals of results are estimated, according to sample design.

Up to present ProMESA 1.3 has been used by research institutes and official veterinary services, as well as an academic resource.

Palabras Clave

Muestreo estadístico, Sanidad animal

Introducción

El conocimiento del estatus sanitario de los animales y su evolución en el tiempo son fundamentales para la toma de decisiones por parte de los servicios veterinarios oficiales, organismos de investigación, profesionales privados y productores. Los estudios epidemiológicos que permiten obtener esta información frecuentemente están basados en datos colectados por muestreos. El objetivo de los muestreos, en el campo de la sanidad animal, generalmente está relacionado con estimar la frecuencia o demostrar la ausencia de una enfermedad o infección.

Los estudios epidemiológicos destinados a estimar la frecuencia de enfermedades, infecciones u otros eventos constituyen la base del conocimiento de la situación sanitaria de un establecimiento, área, país o región. La información que generan permite:

- Identificar las enfermedades existentes
- Estimar su frecuencia y distribución

- Determinar su importancia relativa
- Priorizar las actividades de control
- Planificar, implementar y evaluar los programas de control
- Mejorar la respuesta ante una emergencia sanitaria

Los estudios epidemiológicos para demostrar la ausencia de una enfermedad o infección son actualmente un requisito indispensable para buena parte de las operaciones de comercio internacional de animales y productos de origen animal. Los niveles de exigencia dependen tanto del país importador como del exportador. Las enfermedades que requieren este tipo de estudios actualmente en nuestro país son fiebre aftosa, encefalitis espongiforme bovina, peste porcina clásica e influenza aviar, entre otras.

La cantidad y calidad de información que puede obtenerse a partir de una muestra dependen de la representatividad y el tamaño de la muestra y de la variabilidad del evento de interés en la población.

La selección aleatoria de los individuos evita la existencia de sesgos, mientras que la variabilidad de los datos puede ser controlada parcialmente mediante el tamaño y el método de selección de la muestra. Estos procedimientos de selección constituyen el diseño del muestreo.

El diseño de los estudios para estimar la frecuencia o demostrar la ausencia de una enfermedad o infección se basa en la aplicación de métodos de selección probabilísticos, que permiten conocer la probabilidad de que cada individuo sea incluido en la muestra.

El diseño del muestreo debe tomar en consideración, además de las características de la población y la variabilidad del evento, otros aspectos tales como los costos y la operatividad de la toma de muestras.

Al estar estos estudios basados en muestreos siempre existe un margen de error sobre los resultados obtenidos. El investigador deberá fijar *a priori* el grado de precisión (error aceptable) y el nivel de confianza deseado.

Una vez elegido el método de selección de los individuos, identificado los parámetros necesarios y definido el nivel de confianza y el error aceptable, se puede calcular el tamaño de muestra requerido, mediante la aplicación de los cálculos pertinentes.

Finalmente, el procedimiento de análisis estadístico de los resultados depende de cuál haya sido el método de selección de unidades que se implementó.

Existen actualmente numerosos programas informáticos que pueden realizar estas operaciones. Algunos son paquetes estadísticos amplios que ofrecen la posibilidad de calcular el tamaño de muestra requerido para estimar la prevalencia o detectar la presencia de un evento, para determinados diseños. Otros son programas destinados exclusivamente al diseño de muestreos.

En el actual trabajo se presenta el programa ProMESA 1.3 (Programa de Muestreo Estadístico en Sanidad Animal), diseñado por investigadores del Área de Epidemiología, Instituto de Patobiología, Centro de Investigaciones en Ciencias Veterinarias y Agronómicas, Instituto Nacional de Tecnología Agropecuaria (INTA). El mismo fue concebido como herramienta de apoyo para el diseño de muestreos estadísticos, la selección de unidades y el análisis de los resultados. Su versión inicial fue finalizada en el año 2002. La versión actual es una ampliación de la primera.

Está principalmente destinado a veterinarios y otros profesionales que desarrollan actividades en el marco de los servicios veterinarios, de la investigación o del sector académico, si bien puede ser de utilidad para profesionales del ámbito privado.

Elementos del Trabajo y Metodología

Objetivos de ProMESA 1.3: ayudar en el diseño de muestreos estadísticos, en la selección aleatoria de unidades y en el análisis de los resultados.

Soporte informático: la programación de ProMESA 1.3 se realizó utilizando Microsoft® Visual Basic 6.3. Los formularios de ingreso de parámetros y de presentación de resultados fueron construidos mediante el programa Microsoft® Access 2002. Su utilización requiere la instalación de este último programa.

Procedimientos estadísticos: los diferentes cálculos estadísticos que se utilizan en el programa son convencionales. Han sido descriptos por [1], [2], [3], [4] y [5].

Diseño general: la primera pantalla que se visualiza al abrir el programa es el menú principal (Ilustración 1). A partir del mismo se puede acceder a cuatro módulos temáticos: a. prevalencia de un evento; b. promedio de una variable; c. detección de un evento; d. selección. A su vez, estos módulos presentan diferentes alternativas, de acuerdo al procedimiento que se desee ejecutar. En la Ilustración 2 se representa esquemáticamente el árbol de tareas que puede realizar el programa.

Cada módulo presenta diferentes alternativas hasta llegar a un formulario final. La estructura de los formularios finales es similar. Todos presentan:

- un sector de color naranja, en el cual se deben introducir los parámetros
- un sector amarillo, en el cual se presentarán los resultados
- un botón “Ayuda”, que despliega un texto explicativo de los parámetros a definir
- un botón “Teoría”, con explicaciones de los procedimientos matemáticos y estadísticos seguidos en el procedimiento correspondiente
- un botón “Informe”, que despliega un texto de resumen, mostrando los parámetros utilizados y los resultados obtenidos, que puede imprimirse
- un botón “Ejecutar”, que corre el programa



Ilustración 1: ProMESA 1.3: menú de inicio

Módulo “Prevalencia de un evento”: este módulo está diseñado para trabajar con variables cualitativas. Permite:

- Determinar el tamaño de muestra necesario para estimar una prevalencia: esto se lleva a cabo en función del método de selección de individuos que se haya elegido. El programa presenta tres alternativas: muestreo aleatorio simple, estratificado o en dos etapas.
- Calcular el intervalo de confianza de una prevalencia: el mismo es calculado de acuerdo al diseño de muestreo utilizado para estimar la prevalencia: muestreo aleatorio simple, estratificado o en dos etapas. Este último cálculo, debido a la complejidad matemática que requiere, presenta una dificultad operativa particular: necesita de una gran cantidad de datos. Para ello, en el formulario correspondiente se ha insertado una tabla en la cual se deben cargar los datos necesarios, ya sea escribiéndolos uno a uno o copiándolos desde otra fuente.
- Comparar dos prevalencias: esta alternativa permite por un lado determinar el tamaño de muestra necesario para comparar dos prevalencias, bajo el supuesto de que las muestras son independientes; por otro, calcular el intervalo de confianza de las diferencias, lo cual deja saber si las diferencias son estadísticamente significativas o no.
- Calcular la prevalencia real a partir de una prevalencia aparente: el resultado de un estudio en el cual se aplicó una prueba diagnóstica es la prevalencia aparente, o sea, la proporción de individuos positivos. Esto depende de la sensibilidad y la especificidad de la prueba utilizada. Conociendo ambos valores, ProMESA 1.3 permite estimar el valor de la prevalencia verdadera a partir de la prevalencia aparente.

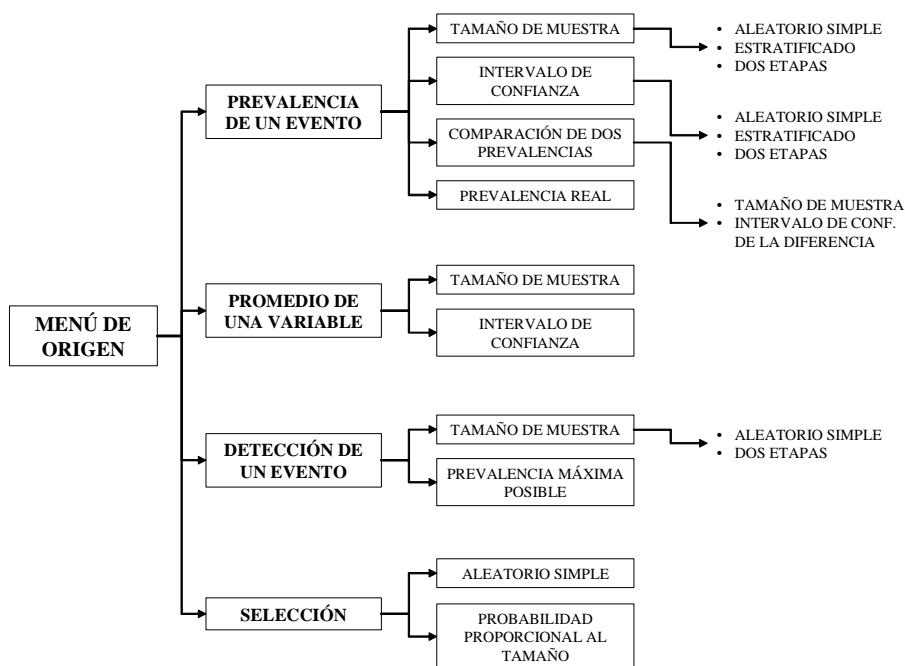


Ilustración 2: Esquema general de actividades que se pueden desarrollar con ProMESA 1.3

Módulo “Promedio de una variable”: orientado a trabajar con variables cuantitativas. Permite:

- Determinar la cantidad de muestras necesarias para estimar el promedio de una variable, solamente a partir de un muestreo con diseño aleatorio simple.
- Calcular el intervalo de confianza de un promedio estimado a partir de un muestreo aleatorio simple.

Módulo “Detección de un evento”: destinado a diseñar y/o analizar muestreos con la finalidad de demostrar la ausencia de un evento. Permite:

- Determinar la cantidad de muestras necesarias para detectar la presencia de un evento a partir de un muestreo aleatorio simple o de un muestreo en dos etapas. En el cálculo incluye el efecto de la sensibilidad y la especificidad de la prueba diagnóstica. Asimismo, permite incluir una o dos pruebas diagnósticas, combinadas en serie o en paralelo, y calcula los valores de sensibilidad y especificidad para las pruebas combinadas. En el diseño en dos etapas se puede optar entre fijar el nivel de confianza o el número de muestras que se desee tomar a nivel predio. Como resultado, además de indicar la cantidad de muestras necesarias para cumplir con los parámetros indicados, presenta el número de resultados falso positivo que se esperaría encontrar, en función de la especificidad declarada y del error tipo 2 que se esté dispuesto a aceptar.
- Calcular indicadores de calidad del muestreo: a. máxima prevalencia posible en una población en la que se realizó un muestreo y se obtuvo 100% de resultados negativos; b. nivel de confianza con el que se puede afirmar que la población no presenta una prevalencia superior a la indicada como mínima.

Módulo “Selección”: su función es seleccionar de manera aleatoria, a partir de listas, un número determinado de unidades a incluir en un muestreo. Permite:

- Seleccionar unidades a partir de una lista, por el método aleatorio simple
- Seleccionar unidades a partir de una lista, por el método aleatorio con probabilidad proporcional al tamaño

Ambos métodos de selección requieren que se ingresen o peguen los datos de la lista total de unidades, en el primer caso, y la lista de unidades con la cantidad de individuos en cada una en el segundo caso.

Resultados

ProMESA 1.3 presenta una serie de formularios que se abren consecutivamente, en función de las opciones que se vayan tomando. El último formulario de cada serie consecutiva de opciones es el operativo. En él se introducen los parámetros y después de ejecutar los cálculos se presentan los resultados.

Las operaciones matemáticas fueron programadas y se llevan a cabo mediante MS Visual Basic. Los códigos están abiertos y pueden ser visualizados por el usuario.

El tiempo requerido para las operaciones necesarias ínfimo. La respuesta es prácticamente inmediata.

A continuación se plantearán cuatro ejemplos, a fin de ilustrar el funcionamiento del programa.

Ejemplo 1: se desea estimar la prevalencia de un evento en una población bovina grande, como ser la de un área o una provincia. La distribución del evento es afectada por la edad de los animales, por lo tanto se decide implementar un muestreo estratificado. La población total consta de 1.5000.000 de terneros, 2.000.000 de novillitos/novillos/vaquillonas y 2.500.000 de vacas/toros. El evento de interés tiene una prevalencia esperada de aproximadamente 15% en terneros, 20% en bovinos de edades intermedias y 35% en animales adultos. Se desea una estimación que tenga un nivel de confianza del 95% y un error relativo aceptable del 20%.

En la Ilustración 3 se puede apreciar el formulario correspondiente al cálculo de tamaño de muestra para este planteo, con los parámetros ingresados y los resultados obtenidos. En “Ayuda” se puede obtener información acerca de los parámetros necesarios. En “Teoría” se describen los procedimientos matemáticos seguidos.

Como respuesta al ejemplo planteado se necesita tomar muestras de 283 animales, 52 del estrato 1 (terneros), 93 del estrato 2 (novillitos/novillos/vaquillonas) y 138 del estrato 3 (vacas/toros).

Ilustración 3: ProMESA 1.3. Formulario para cálculo de tamaño de muestra necesario para estimar una prevalencia, mediante un diseño estratificado.

Ejemplo 2: para descartar la presencia de animales infectados con el virus de la fiebre aftosa en una región se procederá a realizar un muestreo serológico, con diseño en dos etapas, es decir, selección aleatoria de establecimientos y luego de animales.

La región tiene 50.000 establecimientos, que cuentan con una población promedio de 300 bovinos. En el diseño del muestreo se tomarán en cuenta los supuestos sugeridos por la Organización Mundial de Sanidad Animal (OIE): si la infección está presente afecta al menos al 1% de los establecimientos, y en estos, al menos 10% de los animales presentan anticuerpos.

Se trabajará con un nivel de confianza del 95% y se aceptará una probabilidad de cometer error de tipo II (dar como positiva a una población realmente libre del evento) del 5%. Por razones operativas se decide tomar 15 muestras por establecimiento. Para los análisis de laboratorio se utilizarán dos pruebas diagnósticas, combinadas en serie. Los valores de sensibilidad y especificidad son, para la prueba 1: 94% y 96%; para la prueba 2: 98% y 99%.

En la Ilustración 4 se exhibe el formulario para el cálculo de tamaño de muestra que se ajusta al presente ejemplo. En el sector amarillo del formulario pueden apreciarse los resultados.

En primer lugar se ha calculado la sensibilidad y especificidad resultantes de combinar las dos pruebas diagnósticas en serie. La cantidad de muestras requeridas es de 33.120, lo cual surge de tomar 15 muestras a partir de 2.208 rodeos o establecimientos ganaderos. En este caso la cantidad de muestras a tomar por rodeo fue definida por el usuario, lo que produce un nivel de confianza de 13.3% a nivel rodeo. Para compensar un nivel tan bajo y obtener un nivel de confianza global de 95%, el programa incrementa el número de rodeos a analizar. Si en lugar de determinar 15 muestras por rodeo se hubiese determinado 50% de confianza a nivel rodeo, hubiese sido necesario tomar 66 muestras de 564 rodeos, o sea un total de 39.204 muestras.

TAMAÑO DE MUESTRA para detectar la presencia de un evento

Diseño aleatorio simple

Nivel de Confianza: 95%
 Prob. de Error tipo II: 0.05

POBLACION

Mínima prevalencia esperada de rodeos positivos: 0.01
 Cantidad de rodeos: 50000

Diseño en dos etapas

RODEO

Mínima prevalencia esperada de animales positivos: 0.01
 Cantidad promedio de animales por rodeo: 300
 # de muestras: 15

METODO DIAGNOSTICO

Una prueba diagnóstica: defínala en Prueba 1.
 Dos pruebas: defina ambas y la forma de combinarlas.

	Sensibilidad	Especificidad
Prueba 1:	0.94	0.96
Prueba 2:	0.98	0.99

Combinación: En serie

Método diagnóstico

Sensibilidad: **0.9212**
 Especificidad: **0.9996**

Tamaño de la muestra

Rodeos: **2208**
 Animales / rodeo: **15**
 Total de muestras: **33120**
 Sens. a nivel rodeo: **13.3%**

animales /rodeo	muestras /rodeo
50	13
100	14
250	14
500	15
750	15
1000	15
5000	15

FALSOS POSITIVOS
 (para Error tipo II: 0.05)

Analizando 15 muestras por rodeo con un método diagnóstico de especificidad 0.9996 podrían obtenerse hasta:

0 resultados falso positivo por rodeo.

Asimismo, de las 33120 muestras a analizar podrían obtenerse hasta:

18 resultados falso positivo

EJECUTAR

Ayuda

Teoría

Informe

Ilustración 4: ProMESA 1.3. Formulario para cálculo de tamaño de muestra necesario para detectar la presencia de un evento, mediante un diseño en dos etapas

Ejemplo 3: un muestreo en dos etapas para estimar la prevalencia de un evento ha sido realizado. Para calcular la prevalencia, el intervalo de confianza, el error estándar, el efecto del diseño y la tasa de homogeneidad, es necesario compilar los datos en función de la identificación del rodeo. Las variables de interés son: número de muestras tomadas y cantidad de resultados positivos por rodeo. Esos datos pueden ser copiados desde su planilla de origen y pegados en la tabla correspondiente (Ilustración 5). Además, debe definirse el nivel de confianza deseado para el cálculo del intervalo de confianza.

Al ejecutar el programa se obtendrán los valores de manera casi inmediata. Los procedimientos de cálculo seguidos están descritos en “Teoría”.

El valor de la tasa de homogeneidad es un insumo importante para el cálculo del tamaño de muestra de muestreos similares.

INTERVALO DE CONFIANZA de una prevalencia

Intervalo de confianza, según el método de selección de las muestras

ALEATORIO SIMPLE ESTRATIFICADO DOS ETAPAS

Introducir datos para análisis

id	Muestras	Eventos
01.056.0.00342/00	6	0
01.056.0.00438/00	3	0
01.056.0.00463/00	6	0
01.056.0.00730/00	6	0
01.056.0.01012/00	6	0
01.057.0.00035/00	6	0
01.057.0.00037/00	6	0
01.057.0.00494/00	6	0
01.057.0.00498/00	6	0
01.064.0.00059/00	5	0
01.064.0.00291/00	6	0
01.064.0.00636/01	6	1
01.065.0.00379/00	6	0
01.065.0.00681/00	6	0
01.065.0.00702/00	6	0
01.065.0.00775/00	4	0
01.065.0.00782/00	6	0

Nivel de confianza: 95% -

EJECUTAR

Ayuda

Teoría

Prevalencia: 0.0129

Int. conf., límite superior: 0.0193

Int. conf., límite inferior: 0.0066

Error estándar: 0.0032

Efecto del diseño: 1.5324

Tasa de homogeneidad: 0.1095

Total de establecimientos: 317

Total muestras: 1859

Total eventos: 24

Registro: 14 de 317

Ilustración 5: ProMESA 1.3. Formulario para el cálculo de la prevalencia, el intervalo de confianza, el error estándar y otras medidas, a partir de los datos de un muestreo en dos etapas

Ejemplo 4: se cuenta con una lista de 1390 establecimientos agropecuarios, de los cuales se conoce su identificación y el número de individuos que presentan. A partir de dicha lista se desea seleccionar de manera aleatoria y con probabilidad proporcional al tamaño, un total de 25 establecimientos, para realizar un muestreo destinado a estimar la prevalencia de un evento de interés.

La lista debe ser copiada desde su origen y pegada en la tabla incluida en el formulario que se exhibe en la Ilustración 6. En el mismo formulario debe indicarse el número de establecimientos que se desea seleccionar por este procedimiento.

Al ejecutarse el programa se abre un nuevo formulario, de presentación de resultados. En el mismo se presenta la lista de establecimientos seleccionados (Ilustración 7). Si se repitiese el procedimiento con los mismos datos, el resultado sería diferente, ya que en cada ejecución se corre un proceso aleatorio. Sin embargo, al ser mayor la probabilidad de selección de los establecimientos que tienen más animales, es posible, si en la lista existiesen unos pocos rodeos muy grandes en comparación con el resto, que esos siempre fuesen incluidos en la selección.

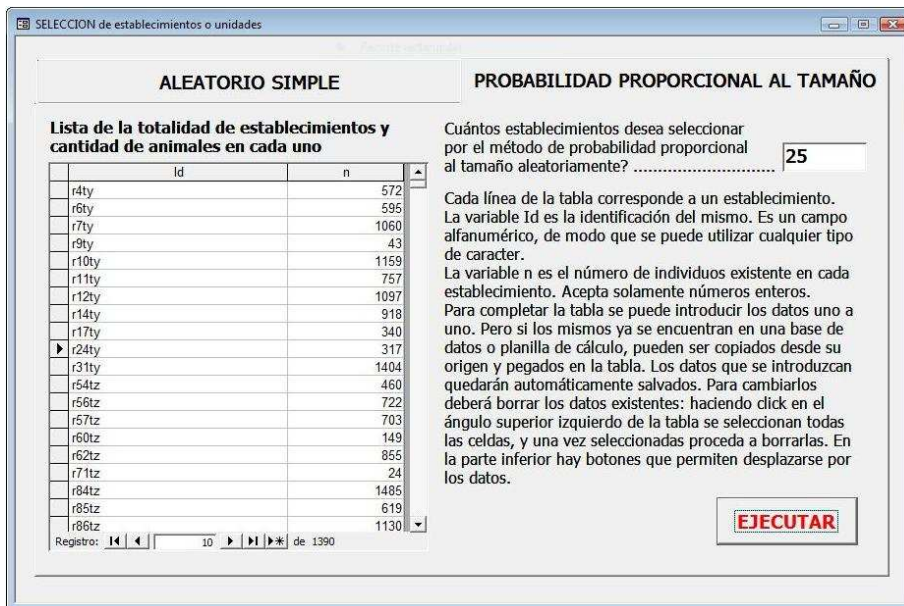


Ilustración 6: ProMESA 1.3. Formulario para seleccionar establecimientos o rodeos de manera aleatoria, con probabilidad proporcional al tamaño

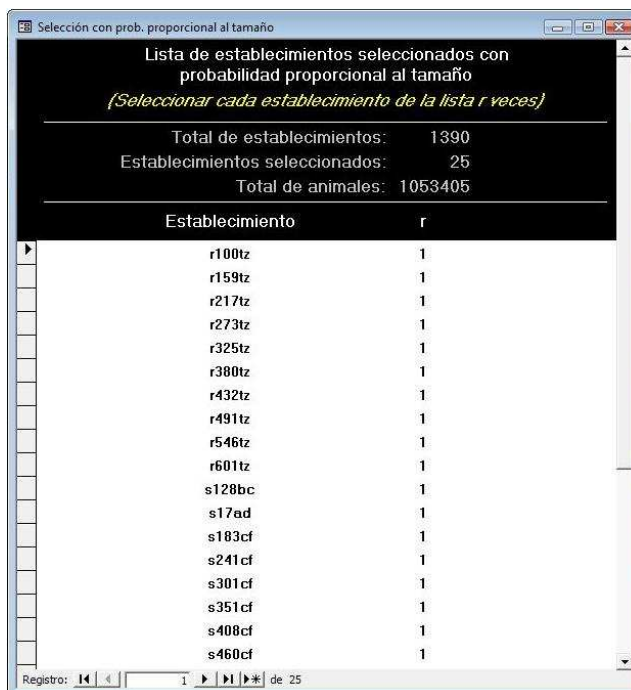


Ilustración 7: ProMESA 1.3. Formulario con presentación de resultados de una selección aleatoria de establecimientos o rodeos, con probabilidad proporcional al tamaño

Discusión

ProMESA 1.3 es una herramienta de apoyo a los profesionales involucrados en el diseño, implementación y análisis de muestreos en poblaciones animales.

Hasta el presente los usuarios más frecuentes han sido instituciones de investigación y los servicios veterinarios oficiales. También ha demostrado ser de utilidad con fines académicos.

La presente versión de ProMESA ha sido verificada en numerosas oportunidades, no habiéndose detectado errores en su funcionamiento. Sin embargo, la detección de errores en el futuro no puede ser descartada y la comunicación de los mismos a los autores será agradecida.

La próxima etapa será transferir ProMESA 1.3 a un programa utilitario estable e independiente de MS Access.

Referencias

1. Levy, P.S.; Lemeshow, S. *Sampling of Populations: Methods and Applications*. New York. Wiley - Interscience Publication. ISBN 0-471-50822-5. 1991.
2. Scheaffer, R.L.; Mendenhall, W.; Ott, L. *Elementary Survey Sampling*, 3ra Edición. Boston. PWS-KENT Publishing Company. ISBN 0-87150-943-1. 1986.
3. Cochran, WG. *Sampling Techniques*, 3ra Edición. New York. John Wiley & Sons. ISBN 0-471-16240-X. 1977.
4. Bennett, S.; Woods, T.; Liyanage, W.M.; Smith, D.L. A simplified general method for cluster-sample surveys of health in developing countries. *World health statistics quarterly*, 44, 98-106. 1991.
5. Otte, M.J.; Gumm, I.D. Intra-cluster correlation coefficients of 20 infections calculated from the results of cluster-sample surveys. *Preventive Veterinary Medicine*, 31, 147-150. 1997.

Datos de Contacto:

Emilio A. León, Méd. Vet., PhD
Instituto de Patobiología, CICVyA-INTA
CC 25. CP 1712. Castelar. República Argentina
E-mail: eleno@cnia.inta.gov.ar