ISSN 1853-3329

UNIVERSIDAD NACIONAL DE TUCUMAN
FACULTAD DE CIENCIAS ECONOMICAS
INSTITUTO DE INVESTIGACIONES ESTADISTICAS

Nº 110
SERIE CUADERNOS INIE
(Diciembre de 2010)

Algunos Comentarios sobre Muestreo de Poblaciones Bernoulli

Por

Viviana B. Lencina

Profesora Adjunta, Instituto de Investigaciones Estadísticas (INIE) Facultad de Ciencias Económicas, Universidad Nacional de Tucumán, Argentina.

Investigador Asistente del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

Editor responsable: Dr. Raúl Pedro Mentz
Instituto de Investigaciones Estadísticas (INIE)
Facultad de Ciencias Económicas, UNT
Avda. Independencia 1900
Casilla de correo nº 209
4000, San Miguel de Tucumán
Tucumán, Argentina
Tel. (0381) 410-7548
Fax: (0381) 436-4105
e-mail: inie@herrera.unt.edu.ar

Algunos Comentarios sobre Muestreo

de Poblaciones Bernoulli

Viviana B. Lencina

Instituto de Investigaciones Estadísticas, Facultad de Ciencias Económicas, UNT
Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET
Argentina
e-mail: lencina_viviana@yahoo.com.ar

Resumen

En este trabajo se aborda el problema de Inferencia en Poblaciones Bernoulli, y se muestra cómo, en el proceso de obtener los datos de la población mediante un muestreo aleatorio simple, diferentes fuentes de variabilidad producen parámetros en las variables aleatorias que serán observadas (muestra) que pueden no coincidir directamente con los parámetros en la población bajo estudio.

Palabras-Claves: Población Bernoulli, Errores de medida, Muestreo

1. Introducción

Al utilizar Metodología Estadística para resolver problemas aplicados es común el adoptar un **modelo probabilístico,** que se supone está regulando el proceso físico subyacente, y especificar las diversas preguntas de investigación en términos de los parámetros de ese modelo. Es fundamental interpretar correctamente los parámetros del modelo probabilístico para poder especificar operacionalmente los objetivos del problema y seleccionar la técnica estadística a ser utilizada. De esta manera, se podrá responder correctamente a las diversas preguntas de investigación planteadas.

Insistir en el cuidado que debe tenerse al realizar la asignación de significado práctico a los parámetros de un modelo supuesto no es un exceso ni algo superfluo. Particularmente, se conocen situaciones en la literatura estadística donde la asignación equivocada de significado a los parámetros ha conducido a controversias con relación a la forma de analizar algunas situaciones. Por ejemplo, en la segunda mitad del siglo pasado se ha prestado bastante atención al análisis de un experimento con dos factores (A y B), donde la respuesta se registra para cada nivel del factor A y para un subgrupo seleccionado aleatoriamente de niveles del factor B. Debido a que se proponen dos modelos diferentes para analizar la misma información, se generó una controversia al evaluar la hipótesis de "inexistencia de efecto principal debido al factor aleatorio en presencia de interacción". Bajo cada modelo, las pruebas de hipótesis propuestas en la literatura usan estadísticos de pruebas diferentes y pueden conducir a resultados o conclusiones diferentes. Se puede mostrar que esta discordancia en los resultados se debe a que las hipótesis puestas a prueba no son equivalentes; es decir, la forma en que se operacionalizó la hipótesis de interés, en ambos modelos, condujo a restricciones sobre los respectivos parámetros que no son equivalentes. Por lo tanto, las hipótesis nulas puestas a pruebas no coincidían (Hocking 1973, Lencina et al., 2005).

En este trabajo se aborda el problema de Inferencia en Poblaciones Bernoulli, siendo que por población Bernoulli se entiende una población cuyas observaciones son valores de una variable aleatoria Bernoulli. Este problema ocurre frecuentemente en la práctica estadística, por ejemplo cuando se quiere estimar la proporción de artículos defectuosos en una línea de producción, la prevalencia de alguna patología, o la proporción de personas a favor de alguna política gubernamental.

El objetivo principal de este trabajo es mostrar cómo, en el proceso de obtener los datos, diferentes fuentes de variabilidad pueden resultar en que los parámetros de las variables aleatorias que componen la muestra difieran de los parámetros en la población bajo estudio. Es pertinente tener en cuenta esta situación dado que generalmente los parámetros sobre los que se quiere hacer inferencia son los de la población.

En la Sección 2, se especifica lo que se entiende por Población Bernoulli; en la Sección 3 se muestra cómo las distribuciones de las variables observadas en el muestreo pueden tener diferente esperanza matemática que las de la población en presencia de error de respuesta interno y en la Sección 4 se muestra lo mismo bajo la presencia de error de respuesta externo (error de medición). Se finaliza este trabajo con una discusión general en la Sección 5.

2. Población Bernoulli

La **distribución Bernoulli** es una distribución de probabilidad que toma el valor 1 con probabilidad π y el valor 0 con probabilidad $1-\pi$. Un **proceso Bernoulli** es un proceso estocástico discreto compuesto por una secuencia finita o infinita de variables aleatorias independientes $X_1, X_2, ...X_n, ...$ tales que, para cada *i*, X_i tiene una distribución Bernoulli, y $P(X_i = 1) = \pi$. Las variables X_i (*i*=1,…) se denominan ensayos o ensayos de Bernoulli (en honor al matemático suizo James Bernoulli, 1654-1705).

Hay diversas variables aleatorias relacionadas con un proceso Bernoulli. Se pueden mencionar, por ejemplo, que (i) el número de éxitos en los primeros n ensayos tiene una *Distribución Binomial*, (ii) el número de ensayos necesarios para obtener r éxitos sigue una *Distribución Binomial Negativa*, y, como caso especial de la binomial negativa, (iii) el número de ensayos necesarios hasta conseguir el primer éxito sigue una *Distribución Geométrica.*

Usualmente se identifica a la distribución binomial como una distribución en el muestreo, porque describe el número de ensayos independientes, u observaciones, con un particular resultado (denominado éxito), cuando cada uno de los n ensayos tiene la misma probabilidad de éxito.

El caso más tipo donde se presenta esta distribución se puede enunciar en términos de un problema de urna. Específicamente, se dispone de una urna con pelotas de dos colores, se extrae con reposición n pelotas y se cuenta el número de pelotas con un color en particular. Este ejemplo corresponde a una población finita, cuyos elementos tienen asociados un atributo considerado como fijo para cada elemento y se observan los atributos de los elementos seleccionados en una muestra aleatoria simple con reposición de tamaño fijo n. Esta correspondencia se debe a que no hay nada de aleatorio en la población y la aleatoriedad en la respuesta que se observará, una vez realizado el muestreo, se debe exclusivamente a la aleatoriedad inducida por el muestreo.

Johnson et al. (1996, pp.106) afirman que "la distribución binomial ocurre cuando una muestra de tamaño fijo n es tomada de una población *infinita*, donde cada elemento de la población, independientemente de los otros, tiene la misma probabilidad π de poseer un determinado atributo. Esta situación también acontece cuando una muestra de tamaño fijo n es tomada de una población *finita*, donde cada elemento de la población, independientemente de los otros, tiene la misma probabilidad π de poseer un determinado atributo y los elementos son seleccionados independientemente y secuencialmente con reposición".

Se observa que al asignar una probabilidad al evento de poseer un determinado atributo, implícitamente se infiere que el hecho de poseer el atributo es aleatorio y no fijo. A modo de ejemplo considérese que se tiene una caja o urna con N pelotas distinguibles, donde cada pelota posee luces intermitentes, que a veces están encendidas y a veces apagadas, independientemente la una de la otra, con probabilidad π o π_s, s=1,..N, dependiendo de si la probabilidad de estar encendida es común o no para todas las pelotas, respectivamente. En este ejemplo, la característica de interés (variable indicadora de si tiene la luz encendida o no) en cada elemento de la población es una variable aleatoria. Algunas de esas pelotas serán seleccionadas aleatoriamente y se registrará si están encendidas o apagadas en el momento en que se realicen las observaciones. Como en esta situación la población no tiene un atributo fijo sino aleatorio, se puede considerar que se está en presencia de una fuente de aleatoriedad interna en cada individuo de la población o error interno.

A las situaciones antes mencionadas también se puede añadir la situación en que una muestra de tamaño fijo n es tomada de una población *finita*, donde cada elemento de la población , independientemente de los otros, tiene la misma probabilidad π de poseer un determinado atributo y los elementos son seleccionados independiente y secuencialmente sin reposición (ver demostración Apéndice A).

Generalmente, es una de las más usadas en los cursos introductorios de Estadística y Probabilidades, junto con las distribuciones Poisson y Normal. Se usa, por ejemplo, para modelar:

(a) el número de caras obtenidas al arrojar "n" veces una moneda,

(b) el número de bolas rojas seleccionadas con reposición de una caja con bolas rojas y azules,

(c) el número de profesores a favor de alguna política universitaria en una muestra aleatoria de "n" profesores,

(d) el número de estudiantes en una muestra aleatoria de tamaño "n" que terminan una prueba en menos de una hora, ó

(e) el número de artículos defectuosos obtenidos en una muestra de "n" artículos seleccionados de una línea de producción.

Con lo mencionado hasta ahora, es sencillo justificar que las variables definidas en los ejemplos (a) y (b) siguen una distribución binomial, cuya probabilidad de éxito coincide con la probabilidad de éxito en la población. Específicamente en el caso (b), el número de bolas rojas en las "n" bolas seleccionadas con reposición de una caja con r bolas rojas y b azules tiene una distribución binomial con parámetros n y π, $b(n,\pi)$, donde π es la proporción de bolas rojas en la caja, $r/(r+b)$.

Por otra parte, los ejemplos (c) y (d) se diferencian de los anteriores debido a que en ellos se debe considerar la existencia de una fuente de aleatoriedad interna (error de respuesta interno); puesto que si el mismo individuo (profesor o estudiante) es observado en dos oportunidades diferentes la respuesta no es necesariamente la misma. En ambos ejemplos también puede ocurrir que la probabilidad de éxito dependa del individuo. En estos ejemplos, si bien es claro que las variables observadas en la muestra son Bernoulli, no es tan directo afirmar cuál es el parámetro π asociado a esta distribución, puesto que se está en presencia de dos fuentes de variabilidad, la que depende del individuo y la introducida por el muestreo. ¿Será que π es la proporción de estudiantes en la población que finalizan la prueba en menos de una hora? ¿Será la probabilidad que un estudiante cualquiera, pero fijo en la población, finalice la prueba en menos de una hora? ¿Será la probabilidad de que cualquier estudiante seleccionado en la muestra finalice la prueba en menos de una hora?

Por otro lado, para justificar que la variable definida en (e) sigue una distribución binomial, es necesario suponer que el proceso de producción es estable, es decir que la probabilidad de que un artículo fabricado sea defectuoso es constante a lo largo de toda la línea de producción, y además suponer que el hecho de que un artículo sea defectuoso no depende de si son o no defectuosos los otros artículos. Se debe destacar en este caso que si el sistema de control o monitoreo no es perfecto, el número de artículos observados como defectuosos puede no coincidir con el número de artículos defectuoso. El sistema de monitoreo, o de clasificación como defectuoso o no defectuoso, introduce una fuente de variabilidad extra en las variables que se observan (error de respuesta externa o error de medición). Específicamente, la probabilidad de que un artículo seleccionado sea calificado como defectuoso depende de la capacidad del sistema para detectar como defectuoso un artículo defectuoso y como no defectuoso un artículo no defectuoso, y de hecho puede diferir de la probabilidad de que el artículo sea realmente defectuoso. Esto debe ser tenido en cuenta para atribuir un correcto significado a los parámetros de las variables aleatorias con las cuales se trabaja.

Las preguntas de investigación, en general, están relacionadas con algún parámetro de la población (objeto de inferencia estadística). Consecuentemente, es importante saber cómo éstos se relacionan con los parámetros en la muestra (variables aleatorias que serán observadas una vez realizado el muestreo), y cómo los parámetros en la muestra están relacionados con las preguntas de investigación formuladas.

En la población, cada bola es roja o azul, cada estudiante puede terminar o no una prueba en menos de una hora, y cada artículo es defectuoso o no defectuoso y a su vez podrá ser detectado como defectuoso o no (ejemplos b, d y e respectivamente). Una vez seleccionada la unidad en la muestra, no hay nada de aleatorio en el ejemplo (b), puesto que la única fuente de aleatoriedad en ese caso es el muestreo; sin embargo, en los ejemplos (d) y (e) la respuesta en la unidad seleccionada todavía es aleatoria debido a la presencia de error de respuesta interno y externo, respectivamente.

3. Error de Respuesta Interno

Se considera la presencia de respuesta de error interno, cuando la característica de interés en cada elemento de la población es una variable aleatoria. En el caso considerado en este trabajo, la característica de interés es una variable Bernoulli, y es por eso que se habla de muestreo de poblaciones Bernoulli.

En esta sección se considerarán las siguientes tres situaciones:

(i) Una población finita de distribuciones Bernoulli, todas ellas independientes con la misma probabilidad de éxito, π; es decir $X_1, X_2, ..., X_N$ $i.i.d$ $B(\pi)$.

(ii) Una población finita de k grupos de distribuciones Bernoulli, independientes con probabilidades de éxito π_j, es decir $X_{j1}, X_{j2}, ..., X_{jN_j}$ $i.i.d$ $B(\pi_j)$ para j=1,…,k.

(iii) Una población infinita de distribuciones Bernoulli, indexada por Ω, independientes con parámetros $\pi_\varpi, \varpi \in \Omega$.

Se supone que se selecciona una muestra aleatoria simple con reposición de tamaño "n" de elementos de la población y se denota con Y_i la respuesta observada para la i-ésima unidad seleccionada en la muestra. En esta sección se encontrará formalmente la distribución de $S = \sum_{i=1}^{n} Y_i$ para cada una de las situaciones antes mencionadas.

3.i. Población Finita de distribuciones Bernoulli con la misma probabilidad de éxito.

Se considera una población finita de distribuciones Bernoulli $\{X_1,\ X_2,...X_N\}$, todas independientes y con la misma probabilidad de éxito (π), para todos los $s \in U = \{1, 2,..., N\}$, de la cual se selecciona un muestreo aleatorio simple con reposición y se observa $Y_1, Y_2,..., Y_n$.

Las variables aleatorias $Y_1, Y_2,..., Y_n$ tienen dos fuentes de aleatoriedad, una proveniente del muestreo, que se denota con el subíndice ξ, y otra proveniente de la respuesta aleatoria (o población Bernoulli), denotada con el sub-índice I (*I*nterno). Específicamente, $Y_i = \sum_{s=1}^{N} U_{is} X_s$, donde U_{is} es una variable aleatoria que asume el valor 1 cuando la *i*-ésima unidad seleccionada en la muestra se corresponde con el elemento s de la población finita, y $E(U_{is}) = N^{-1}$, para todo *i* y *s*. Para calcular $P(Y_i = 1)$ se observa que,

$$\begin{aligned}
P(Y_i = 1) = E_{\xi \times I}(Y_i) &= E_{\xi}(E_{I/\xi}(Y_i)) \\
&= E_{\xi}\left(E_{I/\xi}\left(\sum_{s=1}^{N} U_{is} X_s \right)\right) \\
&= E_{\xi}\left(\sum_{s=1}^{N} U_{is} E_{I/\xi}\left(X_s\right)\right) \\
&= E_{\xi}\left(\pi \sum_{s=1}^{N} U_{is} \right) = \pi.
\end{aligned}$$

Se nota que las $Y_1, Y_2,..., Y_n$ son independientes puesto que el muestreo es con reposición. Por lo tanto, las variables aleatorias $Y_1, Y_2,..., Y_n$ constituyen una secuencia finita de variables aleatorias independientes Bernoulli (Proceso Bernoulli), donde π es la probabilidad de éxito, común para todas las unidades en la población, y consecuentemente $S = \sum_{i=1}^{n} X_i$ tiene distribución binomial con parámetros *n* y π.

Se observa que esta situación también ocurre cuando el muestreo es sin reposición (ver ejemplo presentado en Apéndice A).

3.ii. Población Finita de *k* grupos de distribuciones Bernoulli con probabilidades de éxito $\pi_1,...,\pi_k$

Se considera ahora una población finita compuesta por *k* grupos $\mho_1, \mho_2,..., \mho_k$ de distribuciones Bernoulli independientes, dentro y entre grupos, con probabilidades de éxito $\pi_1,...,\pi_k$ respectivamente. Se supone que cada grupo tiene $N_1, N_2,..., N_k$ elementos respectivamente ($N = \sum_{j=1}^{k} N_j$).

Bajo un muestreo simple con reposición, las variables aleatorias Y_i (i=1,…, n), asumen los valores 0 ó 1, y la probabilidad de éxito, $P(Y_i = 1)$, se calcula como sigue:

$$
\begin{aligned}
P(Y_i = 1) = E_{\xi \times I}(Y_i) &= E_{\xi}\left(\sum_{j=1}^{K} \sum_{s=1}^{N_j} \left[E_{I/\xi}(U_{is} X_s) \right] \right) \\
&= E_{\xi}\left(\left(\sum_{j=1}^{K} \sum_{s=1}^{N_j} \left[U_{is} E_{I/\xi}(X_s) \right] \right) \right) \\
&= E_{\xi}\left(\sum_{j=1}^{k} \left[\pi_j \sum_{s=1}^{N_j} U_{is} \right] \right) = \sum_{j=1}^{k} \left[\pi_j \frac{N_j}{N} \right].
\end{aligned}
$$

Se observa que las $Y_1, Y_2,..., Y_n$ son independientes debido a que el muestreo es con reposición, entonces ellas constituyen una secuencia finita de distribuciones Bernoulli independientes, donde $\pi = \sum_{j=1}^{k} \pi_j N_j / N$ no es necesariamente la probabilidad de éxito de algún elemento de la población, sino que es la media ponderada de las probabilidades de éxito en los diferentes estratos, ponderada por el tamaño de cada estrato. En otras palabras, se puede decir que π es la media de las probabilidades de éxito en la población, usando una medida uniforme discreta como medida de probabilidad en la población. Aquí también, se sigue que $S = \sum_{i=1}^{n} Y_i$ tiene distribución binomial con parámetros n y π.

3.iii. Población Infinita de Distribuciones Bernoulli

Para finalizar, se considera la situación donde la población es infinita, indexada por Ω, tal que para cada elemento de la población la característica de interés es una variable aleatoria con distribución Bernoulli con parámetros $\pi_{\varpi}, \varpi \in \Omega$, y que la función que asigna π_{ϖ} a cada $\varpi \in \Omega$ es una variable aleatoria cuya distribución de probabilidad viene dada por P_{Ω}.

Para todo i, Y_i asume los valores 0 ó 1, y $P(Y_i = 1)$ es dada por,

$$
\begin{aligned}
P(Y_i = 1) &= E_{\xi\times I}(Y_i) = E_\xi\left(E_{I/\xi}(Y_i)\right)\\
&= E_\xi\left(E_{I/\xi}(X_{\varpi(\xi)})\right)\\
&= \int_\Omega \pi_\varpi dP_\Omega = \pi,
\end{aligned}
$$

donde, dado el muestreo, $\varpi(\xi)$ denota la *i*-ésima unidad seleccionada de la población. Aquí, nuevamente, todas las variables $Y_1, Y_2, ..., Y_n$ tienen distribución Bernoulli con parámetro π, la media o esperanza matemática de las probabilidades de éxito en los elementos de la población.

Debido a que el muestreo es aleatorio simple sin reposición, las variables $Y_1, Y_2, ..., Y_n$ son independientes y la distribución de $S = \sum_{i=1}^{n} Y_i$ es binomial con parámetros n y π.

4. Error de respuesta externo.

En la Sección 3 se ha discutido la presencia de error interno, interpretándose por ello a que la característica de interés en cada individuo es en realidad una variable aleatoria. Una vez realizado el muestreo aleatorio simple con reposición, las variables observadas $Y_1, Y_2, ..., Y_n$ constituyen un proceso finito de Bernoulli con parámetro π, donde π es la media de las probabilidades de éxito en la población de donde se tomó la muestra.

Se considera en esta sección la presencia de error de respuesta externo, o de medición, en que, una vez seleccionada la unidad que ha de observarse, el proceso de medición no permite observar Y_i sin error, sino que se observa W_i. En el ejemplo sobre el número de artículos defectuosos en una muestra, Y_i indicaría si el *i*-ésimo articulo es o no defectuoso, y W_i indicaría si el artículo es detectado o no como defectuoso.

Si se considera que

$$P(W_i = 1 \mid Y_i = 1) = S \qquad (1)$$

y

$$P(W_i = 0 \mid Y_i = 0) = E \qquad (2)$$

son constantes que no dependen de i o π, se puede concluir que para cada una de las situaciones descritas en la Sección 3, la probabilidad de éxito de la variable realmente observada W_i es

$$\begin{aligned}\pi^* &= P(W_i = 1) \\ &= P(W_i = 1 | Y_i = 1)\ \pi + P(W_i = 1 | Y_i = 0)(1-\pi) \\ &= (1-E) + (S+E-1)\pi.\end{aligned} \qquad (3)$$

Por lo tanto, $W_1, W_2, ..., W_n$ constituyen una secuencia de distribuciones Bernoulli independientes con probabilidad de éxito π^*. En este punto es importante observar que dependiendo de los valores *E* y *S*, el parámetro π^* puede diferir bastante de π, y que la expresión (3) permite cuantificar el error cometido al interpretar π^* como π para diferentes niveles de *S* y *E*.

En Medicina es común el evaluar una prueba diagnóstica calculando *S* y *E*, que se denominan sensibilidad y especificidad respectivamente.

Existen situaciones donde *S* y *E* dependen de *i*, por ejemplo cuando la persona que realiza la medición, o el instrumento con que se mide no es el mismo para todos. De esta manera, la secuencia finita $W_1, W_2, ..., W_n$ no constituiría un proceso de Bernoulli, debido a que la probabilidad de éxito no sería constante.

5. Discusión

En este trabajo se aborda el problema de muestreo de una población donde la característica de interés es dicotómica, en presencia de error de respuesta interno (Población de variables aleatorias Bernoulli) o de error externo.

Se considera un muestreo aleatorio simple con reposición y se observa que en presencia de error interno las variables aleatorias en la muestra constituyen un proceso de Bernoulli con parámetro π, la media de las probabilidades de éxito de las variables aleatorias Bernoulli en la población. Sólo en caso de que las variables aleatorias en la población tengan todas las mismas probabilidades de éxito, π también será la probabilidad de éxito de cualquiera de ellas.

En presencia de error de medición externo, y bajo ciertos supuestos de estabilidad, también al realizar un muestreo aleatorio simple se observa un proceso de Bernoulli, pero el parámetro del proceso, es decir la probabilidad de éxito en cada ensayo Bernoulli, depende de la habilidad del proceso para detectar correctamente los éxitos y los fracasos.

En las situaciones abordadas en este trabajo, la proporción de éxitos en la población es en realidad una variable aleatoria, puesto que la característica de interés en cada elemento de la población es una variable aleatoria, y π es la media de esta variable proporción. Para ejemplificar lo que aquí se describe, supóngase, por ejemplo, que se quiere estimar el voto a favor de un candidato en una determinada elección; en esta situación el foco de interés es una realización de esa variable proporción, que sucederá el día de las elecciones y no la media π. Tener bien claro este concepto es de suma importancia, particularmente cuando se trata de hacer alguna afirmación acerca de la variabilidad o error cuadrático medio del correspondiente estimador.

Para finalizar se quiere enfatizar que la asignación correcta del significado a los parámetros tiene un papel importantísimo tanto en la educación como en la consultaría estadística, puesto que permite especificar tanto los objetivos operaciones en el problema que se quiera abordar como así también seleccionar la estrategia de análisis.

6. Agradecimientos

La autora agradece al Consejo de Investigaciones de la Universidad Nacional de Tucumán (CIUNT), y al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por el apoyo financiero brindado durante la elaboración de este trabajo.

7. References

Hocking, R.R. (1973). A Discussion of the Two-Way Mixed Model. *The American Statistician*, **27**, 148-152.

Johnson, N.L., Kotz, S. and Kemp, A.W. (1996). *Univariate Discrete Distributions*, 2^{nd}. New York: Wiley.

Lencina, V.B., Singer, J.M., and Stanek, E.J.I. (2005) Much ado about nothing: The mixed model controversy revisited. *International Statistical Review***, Vol 73**, Pages 9-20.

8. Apéndice A

Se considera la situación en que una muestra de tamaño fijo n es tomada de una población *finita*, donde:

(i) cada elemento de la población tiene la misma probabilidad π de poseer un determinado atributo (éxito),

(ii) el hecho de que un elemento de la población posea o no el atributo es independiente de lo que suceda con los otros elementos de la población, y

(iii) los elementos son seleccionados por un muestreo aleatorio simple sin reposición

Considerando que la población finita está compuesta por *N* elementos, sea X_s una variable aleatoria que asume el valor 1 cuando el elemento *s* presenta un éxito y 0 cuando no presenta éxito, para $s = 1,...,N$. Bajo el supuesto que la ocurrencia de éxito en cualquier subconjunto de elementos de la población no depende de la ocurrencia de éxito en los otros elementos se puede decir que las variables aleatorias $X_1, X_2, ..., X_N$ son independientes e idénticamente distribuidas como una Bernoulli con parámetro π ($X_1, X_2, ..., X_N \; i.i.d. B(\pi)$).

Si se realiza un muestreo aleatorio simple sin reposición de tamaño fijo *n*.

Sea $\mathbf{U}_i = (U_{i1}, U_{i2}, ..., U_{iN})$ un vector aleatorio cuyas componentes asumen el valor 1 en la posición correspondiente al elemento de la población extraído en la i-ésima selección/extracción y 0 en todas las otras posiciones, para $i = 1,...,n$.

La variable aleatoria que indica si presenta un éxito el elemento seleccionado en la *i*-ésima extracción se denota con Y_i y es tal que $Y_i = \sum_{s=1}^{N} U_{is} X_s$. Los valores posibles de Y_i son 0 ó 1, y asume el valor 1 con probabilidad $E(Y_i)$. Se puede observar que $E(Y_i) = \sum_{s=1}^{N} E(U_{is} X_s) = \sum_{s=1}^{N} E(U_{is}) E(X_s) = \sum_{s=1}^{N} \frac{1}{N} \pi = \pi$.

Los valores posibles de $(Y_1, ..., Y_n)$ son 2^n, donde las coordenadas de cada *n*-úpla son 0s y 1s. Si se tienen *N* elementos en la población y se extraen sin reposición *n* de ellos, considerando que el orden de selección interesa, se tiene $\binom{N}{n}$ posibles n-uplas de elementos de la población, todas ellas con la misma probabilidad de aparecer $1 \Big/ \binom{N}{n}$.

Dada cualquier *n*-úpla de elementos de la población $\mathbf{s}_{obs} = (s_1, ..., s_n)$, los valores observados de $(Y_1, ..., Y_n)$ todavía son los 2^n mencionados anteriormente, puesto que la respuesta o atributo en cada individuo de la población es aleatoria ($X_1, X_2, ..., X_N$ $i.i.d. B(\pi)$), y consecuentemente

$$f_{\mathbf{s}_{obs}}(y_1,...,y_n) = (1-\pi)^{n-\sum_{i=1}^{n} y_i} \pi^{\sum_{i=1}^{n} y_i} ,$$

donde $f_{\mathbf{s}_{obs}}$ denota la función de masa condicional de $(Y_1,...,Y_n)$ dado que se seleccionó la n-úpla $\mathbf{s}_{obs}$ de elementos.

Como todas las n-úplas de elementos de la población aparecen con la misma probabilidad la función de masa conjunta de $(Y_1,...,Y_n)$ será

$$\begin{aligned} f(y_1,...,y_n) &= (1-\pi)^{n-\sum_{i=1}^{n} y_i} \pi^{\sum_{i=1}^{n} y_i} \\ &= \prod_{i=1}^{n} (1-\pi)^{y_i} \pi^{y_i} , \end{aligned}$$

y por lo tanto, $(Y_1,...,Y_n)$ son independientes.

De esta manera queda probado que $(Y_1,...,Y_n)$ son i.i.d. $B(\pi)$ y consecuentemente $\sum_{i=1}^{n} Y_i \sim b(n,\pi)$.

Se observa que si en la población la probabilidad de éxito no fuera la misma en todos elementos, esta demostración ya no sería válida.

INSTITUTO DE INVESTIGACIONES ESTADISTICAS

Publicaciones

Serie "CUADERNOS"

1	(1973)	Raúl P. Mentz y Víctor J. Elías	*Modelos de Regresión para Estimaciones Estacionales.*
2	(1973)	Juan Carlos Abril	*Aplicaciones del Modelo de Regresión para el Caso Multiplicativo de las Series de Tiempo.*
3	(1973)	Jorge L. Cortigiani y Luis R. Acosta	*Algunos Problemas de Estimación Insesgada en la Distribución Hipergeométrica.*
4	(1973)	Juan Carlos Abril	*Estimaciones Estacionales en Series Monetarias Argentinas (1er. Informe).*
5	(1974)	Juan C. Abril	*Estimaciones Estacionales en Series Monetarias Argentinas (2do.informe).*
6	(1975)	Jorge L. Cortigiani y Luis R. Acosta	*Estimación Insesgada en la Distribución Hipergeométrica Negativa.*
7	(1975)	Ernesto Ramón Cerro	*Construcción de Tablas de Mortalidad para la Provincia de Tucumán y la República Argentina (años 1960, 1970).*
8	(1975)	Jorge L. Cortigiani y Luis R. Acosta	*Estimación Minimax en el Modelo de Respuestas Aleatorias.*
9	(1976)	Jorge L. Cortigiani y Luis R. Acosta	*Estimación Admisible en el Modelo de Respuestas Aleatorias.*
10	(1976)	Jorge L. Cortigiani y Santiago Di Lullo	*Estimación de la Probabilidad de Cero Fracasos en la Distribución Hipergeométrica.*
11	(1976)	Raúl Pedro Mentz	*Modelos Estocásticos para Series Cronológicas y el Modelo de Promedios Móviles.*
12	(1977)	Juan Carlos Abril	*Estudio Espectral de Algunos Métodos de Ajuste Estacional.*
13	(1978)	Raúl Pedro Mentz	*Introducción a los Métodos y Modelos para el Análisis Estacional Capítulo 7.*
14	(1980)	María I. Del Valle Mentz de Acosta	*Estimación por Máxima Verosimilitud en el Modelo Autorregresivo de Segundo Orden.*
15	(1980)	Ernesto Ramón Cerro	*Un Análisis del Crecimiento de la Población de la Provincia de Tucumán con Datos Provisorios del Censo de Población de 1980.*
16	(1980)	Raúl P. Mentz, María B. Ceballos, Juan C. Abril y Zulema Cardozo	*Modelos Estocásticos para Dinero y Precios de Argentina 1960-1973.*
17	(1981)	Raúl P. Mentz y María B. López	*Publicaciones de Estadísticos Latinoamericanos.*
18	(1982)	Aldo José Viollaz	*On the Reliability of the Chi-Square Test.*
19	(1982)	Blanca E. Manzur y María B. Ceballos	*El Desempleo en Tucumán.*
20	(1982)	Víctor Manuel Feijoo	*Beneficio y Rentabilidad de la Investigación Agropecuaria.*
21	(1982)	Lidia Rosa Elías	*Descripción y Modelos Estocásticos para un Conjunto de Indicadores Industriales Mensuales.*
22	(1982)	Juan Carlos Abril	*Aproximaciones a las Densidades de Estadísticos en Problemas de Series Cronológicas.*
23	(1982)	Raúl Pedro Mentz	*Estimación en los Modelos Autorregresivos y de Promedios Móviles.*

24	(1982)	Graciela Beatriz Mentz	*Estudio Algebraico de Cadenas de Marcov de dos Estados.*
25	(1982)	María B. Ceballos, María B. López y Patricia M. Fernandez	*Estructura de la Población en América Latina.*
26	(1982)	Aldo José Viollaz	*Interpretación de las Componentes Principales Muestrales.*
27	(1983)	Raúl P. Mentz y Nora M. Jarma	*Encuestas a Centros Estadísticos Latinoamericanos.*
28	(1983)	Raúl P. Mentz y Violeta Sonvico	*Enseñanza de Postgrado y Formación de Profesores e Investigadores en Estadística en la República Argentina.*
29	(1983)	María Beatriz Ceballos	*Estudio de la Fecundidad en América Latina.*
30	(1983)	Patricia M. Fernández y María B. López	*Análisis Estadístico de la Educación en la Argentina, 1900-1981.*
31	(1984)	Ernesto Ramón Cerro	*La Natalidad y la Mortalidad en Tucumán en el Período 1897-1980. La migración en Tucumán en el Período 1970-1980.*
32	(1984)	Graciela Beatriz Mentz	*Cadenas de Marcov Aplicadas al Estudio del Comportamiento de Alumnos Universitarios.*
33	(1984)	Rolf R. Mantel	*Economía Matemática, su Evolución Histórica y Estado Actual.*
34	(1984)	Juan M. Arranz y Lidia R. Elías	*Ciclo de Referencia para la Economía Argentina, 1960-1982.*
35	(1984)	Víctor Jorge Elías	*Economía Empírica y Teórica Económica.*
36	(1984)	Raúl Pedro Mentz	*Profesor Carlos Eugenio Dieulefait, 1901-1982.*
37	(1984)	María Beatriz Ceballos	*La Mortalidad en Algunos Países de América Latina.*
38	(1985)	Ricardo E. González y Elena Bru de Labanda	*Comparación de Estimadores del Parámetro del Modelo de Promedios Móviles de Primer Orden.*
39	(1985)	Raúl P. Mentz y Aldo J. Viollaz	*On Confidence Bands for Time Series Problems in the Time and Frequency Domains.*
40	(1985)	Elena Bru de Labanda y Raúl P. Mentz	*Estimación Máximo Verosímil Exacta Usando la Descomposición de Cholesky en Modelo de Promedios Móviles de Primer Orden.*
41	(1985)	Juan M. Arranz y Lidia R. Elías	*Medidas de los Ciclos Específicos y del Ciclo de Referencia.*
42	(1986)	Raúl P. Mentz y Oscar R. Galarza	*Información Estadística en los Medios de Comunicación Masiva.*
43	(1986)	Juan Carlos Abril	*La Densidad Aproximada de Algunas Formas Cuadráticas de Variables Aleatorias Estacionarias.*
44	(1986)	Aldo J. Viollaz y María R. Santillán	*Distribución del Cuadrado de la Máxima Correlación Canónica para Tamaños Muestrales Pequeños.*
45	(1986)	Juan Carlos Abril	*Edgeworth Approximation to the Density of Maximum Likelihood Estimators with Applications to Time Series Problems.*
46	(1986)	Aldo J. Viollaz y María R. Santillán	*Caso Especial de Correlaciones Canónicas Cuando Â=ß.*
47	(1987)	Lidia R. Elías y Carlos G. Rivas	*Indices Compuestos de Series Adelantadas Coincidentes y Atrasadas: Pronósticos.*
48	(1988)	Juan Carlos Abril	*Edgeworth Expansion for the Density of a Function of Statistics Whose Density Admits an Expansion.*
49	(1989)	Juan Carlos Abril	*Aproximaciones a las Densidades de Estimadores y Estadísticos de Tests.*

50	(1990)	Juan Carlos Abril	*Una Nota sobre Predicción Usando Información Externa al Modelo.*
51	(1990)	Graciela Beatriz Mentz	*Distribución Asintótica de los Estimadores de las Probabilidades de Transición de una Cadena de Marcov Usando Datos Agregados.*
52	(1990)	María Rosa Santillán	*Sobre la Aplicación del Método de Perturbación para Obtener Expansiones Asintóticas de la Distribución de Raíces Características de Matrices Aleatorias.*
53	(1990)	Raúl Pedro Mentz	*Sobre la Historia de la Estadística Oficial Argentina.*
54	(1990)	María Rosa Santillán	*Comentarios sobre la Distribución Asintótica de las Correlaciones Canónicas.*
55	(1991)	Raúl Pedro Mentz y Víctor J. Yohai	*Sobre la Historia de la Enseñanza de la Estadística en las Universidades Argentinas.*
56	(1991)	Pedro A. Morettin y Raúl P. Mentz	*Graduate Statistical Training in Argentina and Brazil.*
57	(1991)	Juan Carlos Abril y Koichi Maekawa	*Durbin-Watson Test in Non-Linear Regression Models.*
58	(1991)	María Beatriz Ceballos	*Población y Empleo del Noroeste Argentino por Departamento.*
59	(1993)	Juan Carlos Abril	*On the concept of approximate sufficiency.*
60	(1993)	R. P. Mentz, P. A. Morettin y C. M. C. Toloi	*On Residual Variance Estimation in Arma Models.*
61	(1993)	María Rosa Santillán	*Función Característica, Transformada de Fourier y Polinomios de Hermite.*
62	(1993)	María Rosa Santillán	*Expansiones Asintóticas de Edgeworth de la Distribución de Estadísticos.*
.63	(1993)	María Rosa Santillán	*Métodos de Perturbación.*
64	(1994)	María B. Ceballos y Christine A. Isgro	*Factores que Determinan La Fecundidad Un Estudio Para Argentina en 1980.*
65	(1994)	T. W. Anderson, R. P. Mentz, C. I. Martínez y N. M. Jarma	*Simulations of Iterative Procedures for Maximum Likelihood Estimation In MA(1) Models.*
66	(1995)	Marcela A. D'Urso Villar	*Comparaciones Empíricas entre los Métodos X-11.2 Y X-11 ARIMA (1988) de Ajuste Estacional.*
67	(1995)	María B. Ceballos y Christine A. Isgro	*Población Femenina Económicamente Activa en Tucumán*
68	(1995)	R.P. Mentz, P.A. Morettin y C. M. C. Toloi	*Residual Variance Estimation in Moving Average Models*
69	(1996)	María B. Ceballos y Christine A. Isgro	*Tasa de Actividad Femenina en el Gran San Miguel de Tucumán y Tafí Viejo 1990-1996.*
70	(1996)	María B. Ceballos y Christine A. Isgro	*La Tasa de Desocupación del Gran San Miguel de Tucumán y Tafí Viejo 1990-1995.*
71	(1997)	María B. Ceballos y Graciela B. Mentz	*Estacionalidad de La Mortalidad Infantil en Tucumán.*
72	(1997)	Juan C. Abril y María R. Santillán	*Approximation to the Finite Sample Distribution of a General Estimator of the Coefficient in A AR(1) Model.*
73	(1997)	J. M. Jorrat, N. M. Jarma y C. Hortt	*Estudio y Predicción del Ciclo Económico en Argentina.*
74	(1997)	J. M. Jorrat, N. M. Jarma y C. Hortt	*Sistema de Indicadores y Ciclos Económicos.*

75	(1997)	J. M. Jorrat, N. M. Jarma y C. Hortt	*Ciclos Económicos y Sistemas de Indicadores Coincidentes y Líderes.*
76	(1997)	M. B. Ceballos, M. I. Mentz, Ch. A. Isgro, y A. I. Pérez	*Tabla de Vida Activa de Argentina y de Provincias del NOA 1991.*
77	(1997)	A. I. Pérez, Ch. A. Isgro y M. B. Ceballos	*Estudio del Ingreso Femenino en el Gran San Miguel de Tucumán Y Tafí Viejo en el Período 1991-1995.*
78	(1997)	N. M. Jarma, M. D'Urso y R. P. Mentz	*Indicadores y Ciclos Económicos de la República Argentina, Período 1980-1997.*
79	(1998)	R. P. Mentz, P. A. Morettin y C. M. C. Toloi	*On Least Squares Estimation of the Residual Variance in The First Order Moving Average Model*
80	(1998)	N. R. De Martínez, E. Stigman, M. A. Martínez, M.A. D'Urso, N. M. Jarma y P. Fernandez	*Análisis de Factores de Riesgo del Cáncer de Mama.*
81	(1999)	M. B. Ceballos, M. A. D'Urso Villar y N. M. Jarma	*Población de 65 Años y Más en la República Argentina. Estudio de Algunos Indicadores Demográficos, Período 1947-1991.*
82	(1999)	María B. Ceballos y Nora M. Jarma	*Algunos Indicadores de la Población de 65 años y más en la República Argentina Desagregada por Sexo. Período 1947-1991.*
83	(1999)	Manuel L. Cordomí y Nora M. Jarma	*Non Linear Equations Systems in Economics*
84	(1999)	Adriana F. Panico de Bruguera y Carlos I. Martínez	*Trimestralización de Series Anuales. Análisis del Método de Harberger.*
85	(1999)	R. P. Mentz, N. M. Jarma y C. I. Martínez	*Bootstrap Estimation of Standard Errors in Seasonal Analysis.*
86	(2000)	Raúl Pedro Mentz	*Estacionalidad: Introducción a los Modelos y Métodos Estadísticos. Primera Parte.*
87	(2000)	Raúl Pedro Mentz	*Estacionalidad: Introducción a los Modelos y Métodos Estadísticos. Segunda Parte.*
88	(2000)	María Beatriz Ceballos	*La Migración Tucumana en el Período 1960-1980.*
89	(2001)	Nora M. Jarma y Raúl P. Mentz	*Smoothing Seasonal Time Series for Cyclical Analysis.*
90	(2001)	Raúl P. Mentz y Carlos I. Martínez	*Robustez de Métodos de Estimación en Series Cronológicas.*
91	(2001)	Raúl Pedro Mentz	*Estacionalidad: Introducción a los Modelos y Métodos Estadísticos . Tercera Parte.*
92	(2001)	Raúl Pedro Mentz	*Estacionalidad: Introducción a los Modelos y Métodos Estadísticos . Cuarta Parte.*
93	(2002)	María Beatriz Ceballos	*Cambios en la Estructura por edad de la Población y sus Consecuencias. Un Análisis de la Provincia de Tucumán, 1869-2001.*
94	(2002)	María B. Ceballos, Nora M. Jarma y Adriana I. Pérez	*Indicadores Demográficos Seleccionados de la Provincia de Tucumán, 1991-2001.*
95	(2003)	María B. Ceballos y Nora M. Jarma	*El Crecimiento de los grandes grupos de edad, en la República Argentina. Por Regiones. Período 1947-2001.*
96	(2004)	María B. Ceballos y Nora M. Jarma	*Perfil Sociodemográfico del Noroeste Argentino, 1991-2001.*

97	(2005)	Nora M. Jarma y María B. Ceballos	*Indicadores de privación en los hogares en base a datos censales. Fracciones censales del Gran San Miguel de Tucumán, 2001.*
98	(2006)	María Beatriz Ceballos	*Mortalidad Infantil según causas. Noroeste Argentino 1997-2002.*
99	(2007)	Nora M. Jarma y María I. Mentz	*Condiciones Sociodemográficas de la Población correspondiente a los Municipios de la Provincia de Tucumán, República Argentina. Año 2001.*
100	(2007)	María B. Ceballos, Nora M. Jarma y María I. Mentz	*Población Joven de la Provincia de Tucumán, Argentina. Condiciones Económicas y Sociales. Año 2001.*
101	(2007)	Nora M. Jarma y María B. Ceballos	*Indicadores de Privación en los Hogares en Base a Datos Censales. Departamentos de la Provincia de Tucumán, 2001.*
102	(2007)	Nora M. Jarma y María B. Ceballos	*Las Condiciones Sociodemográficas del Adulto Mayor en los Municipios de la Provincia de Tucumán, República Argentina. Año 2001.*
103	(2007)	María B. Ceballos y Patricia M. Fernández	*Mortalidad Femenina de 15 a 59 Años, Noroeste Argentino. 1997-2005.*
104	(2007)	María B. Ceballos y Patricia M. Fernández	*Condición Sociodemográfica de la Mujer en el Noroeste Argentino. 1980.2001.*
105	(2007)	Luciana I. Pérez Zamora y María R. Santillán	*Aplicación del Modelo Estructural de Espacio de Estado para la Proyección de la Demanda de Energía Eléctrica de la Provincia de Santa Fe, Argentina.*
106	2007	María R. Santillán y Luciana I. Pérez Zamora de Nahas	*Zonificación del Mercado Eléctrico de la Provincia de Santa Fe.*
107	2009	Nora M. Jarma, Christine A. Isgro y Adriana I. Pérez	*Discapacidad en la Provincia de Tucumán.*
108	2009	Viviana B. Lencina	*Regresión Logística con Muestras Complejas.*
109	2010	María L. Durbán Reguera	*Introducción a los Modelos Mixtos.*
110	2010	Viviana B. Lencina	*Algunos Comentarios sobre Muestreo de Poblaciones Bernoulli*

Serie "NOTAS"

1	(1972)	Raúl P. Mentz	*Ejemplos para un Curso de Estadística (3ra. Parte).*
2	(1972)	Raúl P. Mentz	*Sobre la Enseñanza de Estadística.*
3	(1973)	Luis R. Acosta, Jorge L. Cortigiani y Santiago M. Di Lullo	*Estimabilidad en el Caso Discreto.*
4	(1873)	Instituto de Investigaciones Estadísticas	*Informe Anual 1972.*
5	(1973)	Raúl P. Mentz y Juan C. Abril	*Programas de Computación para el Análisis Estadístico de las Series Cronológicas (1ra. Parte).*
6	(1974)	Ernesto R. Cerro	*Tablas de Mortalidad para la Provincia De Tucumán, su Construcción para los Años 1897, 1947 Y 1960 Mediante Un Método Simple.*
7	(1974)	Instituto de Investigaciones Estadísticas	*Informe Anual 1973.*
8	(1974)	Raúl P. Mentz y Juan C. Abril	*Programas de Computación para el Análisis Estadístico de las Series Cronológicas (2da. Parte).*
9	(1974)	Raúl P. Mentz y José E. Santillán	*La Profesión de Estadístico (Traducción).*
10	(1975)	Instituto de Investigaciones Estadísticas	*Informe Anual 1974.*
11	(1976)	Instituto de Investigaciones Estadísticas	*Informe Anual 1975.*
12	(1976)	Víctor R. Valderrábano	*Estimaciones Estacionales en Series de Precios Argentinos.*
13	(1976)	María B. Ceballos	*Estadística de la Construcción en Tucumán.*
14	(1977)	Instituto de Investigaciones Estadísticas	*Informe Anual 1976.*
15	(1978)	Raúl P. Mentz	*Determinación del Orden de un Modelo Autorregresivo.*
16	(1978)	Instituto de Investigaciones Estadísticas	*Informe Anual 1977.*
17	(1978)	Raúl P. Mentz, Juan C. Abril, Zulema Cardozo y María B. Ceballos	*Gráficos de Series Cronológicas. Un Conjunto de Programas de Computación.*
18	(1978)	Santiago M. Di Lullo y Raúl P. Mentz	*Instituto de Investigaciones Estadísticas (INIE).*
19	(1979)	Instituto de Investigaciones Estadísticas	*Informe Anual 1978.*
20	(1979)	Juan C. Abril	*Programa de Regresión Múltiple.*
21	(1980)	Juan C. Abril	*Notas Sobre Análisis Estadístico Multivariado.*
22	(1980)	Instituto de Investigaciones Estadísticas	*Informe Anual 1979.*
23	(1980)	M. Lozano, Z. Cardozo, N. Jarma, C. Martínez y V. Feijoo	*Archivos Operativos de Series Estadísticas.*
24	(1981)	Instituto de Investigaciones Estadísticas	*Informe Anual 1980.*
25	(1981)	Instituto de Investigaciones Estadísticas	*Seminario de Actualización en Estadística. Abril 1981.*
26	(1982)	Instituto de Investigaciones Estadísticas	*Informe Anual 1981.*
27	(1982)	Lidia R. Elías	*Análisis Estadístico de los Ciclos Económicos Argentinos.*
28	(1982)	Carlos Martínez, Ricardo González y Nora Jarma de Cortes	*Computación para Estadística en el INIE.*
29	(1982)	Carlos Martínez y Nora Jarma de Cortes	*Gráficos Estadísticos. Un Conjunto de Programas de Computación.*
30	(1983)	Instituto de Investigaciones Estadísticas	*Informe Anual 1982.*
31	(1983)	Lidia Rosa Elías	*Programas del National Bureau of Economic Research.*

32	(1983)	Ricardo González y Carlos Martínez	*Archivos Bibliográficos del INIE.*
33	(1984)	Instituto de Investigaciones Estadísticas	*Informe Anual 1983.*
34	(1984)	Graciela B. Mentz	*Estudios Complementarios del Comportamiento de Alumnos Universitarios.*
35	(1985)	Instituto de Investigaciones Estadísticas	*Informe Anual 1984.*
36	(1996)	Instituto de Investigaciones Estadísticas	*Informe Anual 1985.*
37	(1997)	Instituto de Investigaciones Estadísticas	*Informe Anual 1986.*
38	(1988)	Instituto de Investigaciones Estadísticas	*Informe Anual 1987.*
39	(1989)	Instituto de Investigaciones Estadísticas	*Informe Anual 1988.*
40	(1990)	Instituto de Investigaciones Estadísticas	*Informe Anual 1989.*
41	(1991)	María Beatriz Ceballos	*Tabla de Vida Activa.*
42	(1991)	Instituto de Investigaciones Estadísticas	*Informe Anual 1990.*
43	(1992)	Instituto de Investigaciones Estadísticas	*Informe Anual 1991.*
44	(1993)	Instituto de Investigaciones Estadísticas	*Informe Anual 1992.*
45	(1993)	Juan Carlos Abril	*Nota Sobre Modelos de Espacio de Estado y Filtro de Kalman.*
46	(1993)	Instituto de Investigaciones Estadísticas	*Informe Anual 1993.*
47	(1995)	Instituto de Investigaciones Estadísticas	*Informe Anual 1994.*
48	(1995)	Lucía E. Ruiz	*Una Nota Sobre el Modelo de Media Constante.*
49	(1996)	Adriana I. Pérez	*Análisis Exploratorio de Datos.*
50	(1996)	Instituto de Investigaciones Estadísticas	*Informe Anual 1995.*
51	(1997)	Instituto de Investigaciones Estadísticas	*Informe Anual 1996.*
52	(1997)	Santiago Mario Di Lullo	*Los Primeros 50 años del I.N.I.E.*
53	(1998)	Instituto de Investigaciones Estadísticas	*Informe Anual 1997.*
54	(1999)	Instituto de Investigaciones Estadísticas	*Informe Anual 1998.*
55	(2000)	Instituto de Investigaciones Estadísticas	*Informe Anual 1999.*
56	(2000)	Raúl Pedro Mentz	*Enseñanza de Estadística al Nivel de Postgrado en la Facultad de Ciencias Económicas de la Universidad Nacional de Tucumán*
57	(2001)	Christine Adriane Isgro	*Análisis de los Resultados de los Operativos de Evaluación de Calidad de Finalización de la Escuela Secundaria para la Provincia de Tucumán, Período 1997-1999.*
58	(2001)	Christine Adriane Isgro y Adriana Inés Pérez	*Análisis de la Distribución del Ingreso de los Ocupados del Gran San Miguel de Tucumán (1995-2000).*