

# Técnicas de Muestreo I

Patricia Isabel Romero Mares

Departamento de Probabilidad y Estadística  
IIMAS UNAM

agosto 2015

## Algo de historia del Muestreo

- El muestreo es tan antiguo como el hombre.
- Se llega a conclusiones acerca de peces, árboles, lagos, comida, etc. con base en muestras no estructuradas relativamente pequeñas que tenemos disponibles en la vida diaria.
- Estadísticas sociales y económicas tienen una larga historia de estudios muestrales, en el sentido de cobertura parcial de la población relevante.

- Edmund Halley estima la mortalidad de la raza humana con base en las *curiosas tablas de nacimientos y funerales en la ciudad de Breslau, 1693*.
- En 1754 se hicieron estimaciones de la población de Inglaterra basadas en el número de casas en la lista de impuestos más una estimación burda del número de cabañas; el total de viviendas se multiplicó por 6 (personas por vivienda).
- Otras estimaciones se basaron en el número de bautismos, matrimonios y entierros.

- En 1800 Sir Frederick Morton estimó la población de Gran Bretaña en 9 millones de habitantes, usando datos muestrales del número promedio de habitantes por casa, así como el número de nacimientos. El primer censo en 1801 confirmó su estimación.
- Sir John Lawes usó el registro anual de producción de trigo en Rothamsted en 5 parcelas de 33 acres para estimar el cambio en la producción por acre de 1852 a 1879 para Inglaterra y Gales.
- Posteriormente se calcularon otras estimaciones como el número de caballos, ganado, ovejas y puercos, así como el área cultivada.
- Estudios por muestreo de cultivos; estadísticas económicas de precios, salarios, empleo; estudios sobre salud y opinión pública.

- En el siglo 19 con el nacimiento de sistemas nacionales de estadística y el desarrollo de estudios sociológicos, se crean dos movimientos, el primero de cobertura total (censo) y el segundo de casos de estudio (unas pocas familias, una sola ciudad, etc.).
- Kiaer en 1895 en una reunión del ISI (International Statistical Institute) propuso el **Método Representativo**. Explicó que una investigación parcial podría dar resultados confiables si las observaciones formaban una **pintura representativa** de todo el campo de estudio. Tuvo una feroz crítica por parte de los defensores de los conteos exhaustivos.
- Kiaer (1897) La muestra debe reflejar la población objetivo en características importantes.

- Arthur Lyon Bowley (1906) promovió activamente sus ideas sobre muestreo en general y aleatorización en particular.
- Realizó una prueba empírica del Teorema Central del Límite.
- En un artículo mencionó “... podemos obtener tan buenos resultados como queramos muestreando, y muy frecuentemente pequeñas muestras son suficientes; la única dificultad es el asegurar que cada persona o cosa tenga la misma oportunidad de inclusión en la investigación”.

- Para la reunión del ISI de 1925 el método representativo era aceptado. Surgieron dos formas de selección de la muestra:
  - selección aleatoria (probabilidades de inclusión iguales)
  - selección por intención (que refleje la población en ciertas variables)
- Neyman (1934) demuestra, entre otras cosas, que la distribución óptima en muestreo estratificado debe ser proporcional a la varianza del estrato, lo que lleva a probabilidades de inclusión desiguales.
- Neyman (1934) impulsó un rápido desarrollo de las técnicas de Muestreo.
- La I y II Guerras Mundiales ampliaron el alcance de las estadísticas oficiales pero al costo de usar muestras en lugar de medir a la población completa.



- Horvitz y Thompson (1952) formularon el muestreo con probabilidades desiguales (estimador  $\pi$ ).
- Godambe (1955) demuestra que no existe el estimador “óptimo” sino que depende del diseño de muestra.
- Mucho de la investigación en Muestreo se inició por estadísticos trabajando en la Oficina del Censo (Deming, Hansen, Hurwitz, Madow).
- Richard Royall (1970) da mucho impulso a la inferencia basada en modelo, que no ha reemplazado al enfoque de aleatorización o inferencia basada en diseño. Sigue el debate entre estos dos enfoques.

El Muestreo es diferente a las otras técnicas estadísticas.

- Población finita
- El principal objetivo es la estimación de ciertas características de la población bajo estudio
- $X_1, X_2, \dots, X_N$  fijas, desconocidas
- Lo aleatorio es introducido por el investigador a través del diseño muestral (probabilidad de selección de la muestra)
- En este curso usaremos la inferencia basada en diseño

# Bibliografía

- Bellhouse, D.R. (1988). A Brief History of Random Sampling Methods. P.R. Krishnaiah y C. R. Rao, eds. Handbook of Statistics, Vol 6 Elsevier Science Publishers B.V. pp 1-14.
- Godambe, V.P. (1955). A Unified Theory of Sampling from Finite Populations. *JRSS, Series B*, **17**, No. 2, pp 269-278.
- Horvitz, D.G. & Thompson, D.J. (1952). A Generalization of Sampling Without Replacement from a Finite Universe. *JASA*, **47**, No. 260, pp 663-685.
- Kiaer, A. (1897). The representative method of statistical surveys (1976 English translation of the original Norwegian). *Oslo: Central Bureau of Statistics of Norway*.

# Bibliografía

- Neyman, J. (1934). On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *JRSS*, **97**, No. 4, pp 558-625.
- Royal, R.M. (1970). On finite population sampling theory under certain linear regression models. *Biometrika*, **57**, pp 377-387.
- Stephan, F.F. (1948). History of the Uses of Modern Sampling Procedures. *JASA* **43**, No. 24, pp 12-39.