

**INSTITUTO DE ADMINISTRACION PÚBLICA DEL ESTADO DE CHIAPAS**

**MAESTRIA DE ADMINISTRACIÓN**

**Y POLÍTICAS PÚBLICAS**

ESTADISTICA ADMINISTRATIVA

**ACTIVIDAD 3:**

**EJERCICIOS DE PROBABILIDAD Y CONTROL DE LECTURA**

**CATEDRATICO:**

Dr. Enrique Antonio Paniagua Molina

**ALUMNO:**

Ma. del Carmen Espíndola Soto

Tapachula de Córdova y Ordoñez, Chiapas, 21 septiembre del 2015.

**Control de Lectura:**

* **TITULO DEL LIBRO:**

Metodología de la Investigación

* **AUTORES:**

Roberto Hernández Sampieri

Carlos Fernández-Collado

Pilar Baptista Lucio

* **TEMA QUE ABORDA:**

**Selección de la muestra**

* **IDEA PRINCIPAL:**

Obtención de la muestra, definición, características, tipos, validación.

* **IDEAS SECUNDARIAS:**

**Conceptos de muestra y tipos.**

* **CONCEPTOS POCO CLAROS:**
* **SINTESIS DE LOS CAPITULOS:**

Selección de la muestra:

Una vez que se ha definido cuál será la unidad de análisis, se procede a delimitar la población que va a ser estudiada y sobre la cual se pretende generalizar los resultados. Así, una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones. La Población o universo se conoce como el Conjunto de todos los casos que concuerdan con determinadas especificaciones. Toda investigación debe ser transparente, así como estar sujeta a crítica y réplica, este ejercicio no es posible si al examinar los resultados el lector no puede referirlos a la población utilizada en un estudio.

Se Definirán los siguientes conceptos:

Muestra es, en esencia, un subgrupo de la población. Digamos que es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que llamamos población.

Muestra Probabilística como el Subgrupo de la población en el que todos los elementos de ésta tienen la misma posibilidad de ser elegidos.

Muestra no probabilística o dirigida Subgrupo de la población en la que la elección de los elementos no depende de la probabilidad sino de las características de la investigación.

Elegir entre una muestra probabilística o una no probabilística depende de los objetivos del estudio, del esquema de investigación y de la contribución que se piensa hacer con ella. La elección entre la muestra probabilística y la no probabilística se determina con base en el planteamiento del problema, las hipótesis, el diseño de investigación y el alcance de sus contribuciones. Las muestras probabilísticas tienen muchas ventajas, quizá la principal sea que puede medirse el tamaño del error en nuestras predicciones. Se dice incluso que el principal objetivo en el diseño de una muestra probabilística es reducir al mínimo este error, al que se le llama error estándar[[1]](#footnote-1). Las muestras probabilísticas son esenciales en los diseños de investigación transaccionales, tanto descriptivos como correlaciónales causales (Encuestas de opinión o Surveys), donde se pretende hacer estimaciones de variables en la población. Estas variables se miden y se analizan con pruebas estadísticas en una muestra, donde se presupone que ésta es probabilística y todos los elementos de la población tienen una misma probabilidad de ser elegidos. Las unidades o elementos muéstrales tendrán valores muy parecidos a los de la población, de manera que las mediciones en el subconjunto nos darán estimados precisos del conjunto mayor. La precisión de dichos estimados depende del error en el muestreo, que es posible calcular.

Muestra probabilística estratificada: Es comparar sus resultados entre segmentos, grupos o nichos de la población.

Muestreo por racimos o clústers: Unidades de análisis se encuentran encapsuladas o encerradas en determinados lugares físicos o geográficos, a los que se denomina racimos,

Las unidades de análisis o los elementos muéstrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Se utilizan tres procedimientos de selección: Tómbola, Números Random o Cuadro de números aleatorios del Stats y Selección sistemática de elementos muéstrales. Listado y otros marcos muéstrales: El marco muestra constituye un marco de referencia que nos permita identificar físicamente los elementos de la población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muéstrales, Archivos, Mapas, Volúmenes y Periodos registrados son útiles para el análisis de estos elementos muéstrales.

Es importante recordar que las muestras probabilísticas requieren dos procedimientos básicos: 1) la determinación del tamaño de la muestra y 2) la selección aleatoria de los elementos muéstrales.

El Teorema del límite central Señala que una muestra de más de cien casos será una muestra con una distribución normal en sus características, lo cual sirve para el propósito de hacer estadística inferencial.

En el enfoque cuantitativo las muestras probabilísticas son esenciales en diseños de investigación por encuestas, donde se pretenden generalizar los resultados a una población. La característica de este tipo de muestras es que todos los elementos de la población al inicio tienen la misma probabilidad de ser elegidos. Así, los elementos muéstrales tendrán valores muy aproximados a los valores de la población, ya que las mediciones del subconjunto serán estimaciones muy precisas del conjunto mayor. Tal precisión depende del error de muestreo, llamado también error estándar. Para una muestra probabilística necesitamos dos elementos básicos: Determinar el tamaño de la muestra y Seleccionar los elementos muéstrales en forma aleatoria.

Las muestras probabilísticas son: Simples, Estratificadas, Sistemáticas y por Racimos. La Estratificación aumenta la precisión de la muestra e implica el uso deliberado de submuestras para cada estrato o categoría que sea relevante en la población. Muestrear por racimos o conglomerados implica diferencias entre la unidad de análisis y la unidad muestra.

Las unidades de análisis o los elementos muéstrales se eligen siempre aleatoriamente para asegurarnos de que cada elemento tenga la misma probabilidad de ser elegido. Se utilizan tres procedimientos de selección:

Tómbola

Muy simple y no muy rápido, consiste en numerar todos los elementos muéstrales del uno al número n. Hacer fichas o papeles, uno por cada elemento, revolverlos en una caja, e ir sacando n número de fichas, según el tamaño de la muestra. Los números elegidos al azar conformarán la muestra.

Números random o números aleatorios

El uso de números random no significa la selección azarosa o fortuita, sino la utilización de una tabla de números que implica un mecanismo de probabilidad muy bien diseñado. Los números random de la Corporación Rand fueron generados con una especie de ruleta electrónica. Existe una tabla de un millón de dígitos, publicada por esta corporación

Selección sistemática de elementos muéstrales

Este procedimiento de selección es muy útil e implica elegir dentro de una población N un número n de elementos a partir de un intervalo K Este último (K) es un intervalo que se va a determinar por el tamaño de la población y el tamaño de la muestra. De manera que tenemos que K = N/n, en donde K = un intervalo de selección sistemática, N = la población y n = la muestra.

Listados y otros marcos muéstrales

Las muestras probabilísticas requieren la determinación del tamaño de la muestra y de un proceso de selección aleatoria que asegure que todos los elementos de la población tengan la misma probabilidad de ser elegidos. Todo esto lo hemos visto, aunque nos falta exponer sobre algo esencial que precede a la selección de una muestra: el marco muestra, este constituye un marco de referencia que nos permita identificar físicamente los elementos de la población, la posibilidad de enumerarlos y, por ende, de proceder a la selección de los elementos muéstrales (los casos de la muestra). Normalmente se trata de un listado existente o una lista que es necesario confeccionar ad hoc

Los listados existentes sobre una población son variados: guías telefónicas, listas de miembros de las asociaciones, directorios especializados, listas oficiales de escuelas de la zona, bases de datos de los alumnos de una universidad o de los clientes de una empresa, registros médicos, catastros, nóminas de una organización, etc. En todo caso hay que tener en cuenta lo completo de una lista, su exactitud, su veracidad, su calidad y su nivel de cobertura en relación con el problema a investigar y la población que va a medirse, ya que todos estos aspectos influyen en la selección de la muestra.

Archivos: Un gerente de reclutamiento y selección de una empresa quiere precisar si algunos datos que se dan en una solicitud de trabajo están correlacionados con el ausentismo del empleado. Es decir, si a partir de datos como edad, género, estado civil, nivel educativo y duración en otro trabajo, es factible predecir la conducta de ausentismo.

Mapas

Los mapas son muy útiles como marco de referencia en muestras de racimos.

Volúmenes

En este caso supongamos que un estudioso del periodismo quiere hacer un análisis de contenido de los editoriales de los tres principales diarios de la ciudad durante los periodos del porfiriato en México, el gobierno sandinista en Nicaragua o el franquismo en España. El investigador va a la Hemeroteca Nacional y encuentra que los diarios son encuadernados por trimestre y año, lo cual le proporciona un marco de referencia ideal, a partir de donde seleccionará n volúmenes para su análisis.

Periodos registrados

En la investigación que se vincula con sesiones terapéuticas, entrevistas con pacientes o reos, etc., se dispone de los listados de archivos que por lo común guardan las instituciones.

Tamaño óptimo de una muestra

Las muestras probabilísticas requieren dos procedimientos básicos: 1) la determinación del tamaño de la muestra y 2) la selección aleatoria de los elementos muéstrales. El primer procedimiento fue descrito en su modalidad más simple en la sección sobre el tamaño de la muestra. Precisar adecuadamente el tamaño de la muestra puede tomarse muy complejo, esto depende del problema de investigación y la población a estudiar.

El tamaño de una muestra depende también del número de subgrupos que nos interesan en una población.

Repasemos que lo óptimo de una muestra depende de cuánto se aproxima su distribución a la distribución de las características de la población. Esta aproximación mejora al incrementarse el tamaño de la muestra. La "normalidad" de la distribución en muestras grandes no obedece a la normalidad de la distribución de una población. La distribución de diversas variables a veces es "normal" y en ocasiones está lejos de serlo. Sin embargo, la distribución de muestras de 100 o más elementos tiende a ser normal y esto sirve para el propósito de hacer estadística inferencial, sobre los valores de una población. A lo anterior se le llama teorema del límite central.

Distribución normal: Esta distribución en forma de campana se ogra generalmente con muestras de 100 o más unidades muéstrales, y es útil y necesaria cuando se hacen inferencias de tipo estadístico

Las muestras no probabilísticas, también llamadas muestras dirigidas, suponen un procedimiento de selección informal. Se utilizan en muchas investigaciones cuantitativas y cualitativas. No las revisaremos ahora, sino en el capítulo sobre muestras cualitativas. Por el momento comenta- remos su esencia y utilidad desde una perspectiva cuantitativa y ejemplificaremos la diferencia con las muestras probabilísticas.

La muestra dirigida selecciona sujetos "típicos" con la vaga esperanza de que sean casos representativos de una población determinada.

La única ventaja de una muestra no probabilística —desde la visión cuantitativa— es su utilidad para determinado diseño de estudio que requiere no tanto una representatividad de elementos de una población, sino una cuidadosa y controlada elección de sujetos con ciertas características especificadas previamente en el planteamiento del problema.

Muestreo al azar por marcado telefónico (Random Digit Dialing)

Ésta es una técnica que los investigadores utilizan para seleccionar muestras telefónicas. Involucra identificar áreas geográficas —para ser muestreadas al azar— y sus correspondientes códigos telefónicos e intercambios (los tres dígitos del número telefónico). Luego, los demás dígitos del número a marcar pueden ser generados al azar de acuerdo a los casos que requerimos para la muestra (n). Es posible identificar qué intercambios son usados de forma primaria para teléfonos residenciales y enfocar el muestreo en ese subgrupo.

Muestra multietapas o polietápica: Este concepto significa que para extraer la muestra hemos utilizado diversos procedimientos. Por ejemplo, en la investigación sobre la televisión y los niños de la ciudad de México, el procedimiento de selección implicó estratos y racimos: en una primera etapa se seleccionarían "escuelas" y en una segunda "niños".

Los estudios exploratorios regularmente emplean muestras dirigidas o no probabilísticas, aunque podrían usarse muestras probabilísticas. Las investigaciones experimentales, la mayoría de las veces utilizan muestras dirigidas, porque como se comentó, es difícil manejar grupos gran- des (debido a ello se ha insistido que, en los experimentos, la validez externa se consolida mediante la repetición o reproducción del estudio). Los estudios no experimentales descriptivos o correlaciónales-causales deben emplear muestras probabilísticas si quieren que sus resultados sean generalizados a una población.

Muy buen libro.

**Estadística Descriptiva**

**Ejercicio 1**

* **Si la señora López compra una de las casas anunciadas para su venta en un diario de TGZ, T es el evento de que la casa tiene tres o más baños, U es el evento de que tiene una chimenea, V es el evento de que cuesta más de $ 100 mil pesos y W es el evento de que es nueva.**
* **Describa (con palabras) cada uno de los siguientes eventos:**



T’= U, V, W (Casa con una chimenea, cuesta más de cien mil, nueva)

U’=T, V, W (Casa de 3 o más baños, cuesta más de cien mil, nueva)

V’=T, U, W (Casa de 3 o más baños, una chimenea, casa nueva)

W= casa nueva

T ∩ U= (Casa de 3 o más baños; tiene una chimenea)

T ∩ V= (Casa de tres o más baños; cuesta más de cien mil pesos)

U’ ∩ V= (Cuesta más de cien mil)

V U W= (Casa que cuesta más de cien mil; nueva)

V’ U W= (Casa con tres o más baños; chimenea; nueva)

T U U= (Casa con tres o más baños; chimenea)

**Ejercicio 2**

* **Un dado está arreglado de manera que cada número impar tiene el doble de probabilidad de ocurrir que un número par. Encuentra P(B), donde B es el evento que un número mayor que 3 ocurra en un solo tiro del dado.**
* **Espacio muestral:** S= ( 1,2,3,4,5,6)
* **Sub conjunto B:** B=(4,5,6)

**Probabilidad**

**Si x es la probabilidad que ocurra un número par, 2x sería la probabilidad que ocurra un número impar.**

**Entonces, encontramos que: 2 y + x + 2 y + x + 2 y+ x = 1**

**Esto se debe al postulado 2**

**La P(B) sería: 4 / 6**

**Ejercicio 3**

* **Calcula la muestra para una población desconocida con un 96% de confianza y 10% error. Para una prevalencia de .5 y .7**

**Para una prevalencia de 0.5**:**n= 105.47 ≈ 105.**

**Para una prevalencia de 0.7**:**n= 88.59 ≈ 89.**

**Ejercicio 4**

* **Calcula la muestra para una población de 350,000 familias, con un 99% de confianza y 5% error. Para una prevalencia de .5 y .7**

**Para una prevalencia de 0.5**……………**´n= 661.81 ≈ 662 familias.**

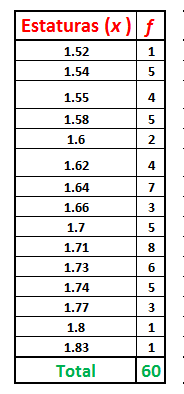
**Para una prevalencia de 0.7**……………**´n= 556.08 ≈ 556 familias.**

**Ejercicio 5**

* **De una población de 1,176 padres de familia de la ciudad de Tuxtla Gutiérrez, se pretende conocer la aceptación de los programas educativos mediante caricaturas. Se pretende obtener una muestra para saber el número de entrevistas y con ello obtener información estadísticamente confiable. Se asume un** [error standard](file:///F:\TAREA%20HOY\Error%20estándar.docx) **de 1.5% con un nivel de confiabilidad del 90%**

**Respuesta: Muestra 298 entrevistas.**

**Ejercicio 6**



* **Son los resultados de preguntarle la estatura a 60 trabajadores del departamento de limpia municipal de SCLC.**
* **Obtén la media aritmética (para datos agrupados)**
* **Obtén la desviación estándar y la varianza (para datos agrupados)**
* **Interpreta los resultados**

**Media: 1.553**

**Varianza: 0.6689358**

**Desviación Estandár: 0.81788496**

1. [↑](#footnote-ref-1)