

2) Sujeto



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

MUESTREO PARA LA IDENTIFICACION Y
CUANTIFICACION DE ENFERMEDADES EN
EL NOPAL DE TUNA BLANCA

T E S I S

Que para obtener el título de:

ACTUARIO

presenta:

ALFONSO GARCIA DURAN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

PROLOGO.

El campo de aplicación de la Estadística es bastante amplio, sin embargo la gente que no conoce dicho campo difícilmente hace un buen uso del análisis estadístico. Paradójicamente, la aplicación de la Estadística es mucho mayor en aquellas áreas como Biología, Química, Economía, Agricultura, Psicología, etc., en donde los egresados no tienen una buena formación de tipo estadístico, adecuado a su desarrollo profesional, que les permita aplicar la metodología estadística de una manera versátil a los problemas con los cuales se encuentran en el ejercicio diario de su profesión; ante esta situación es necesario promover el trabajo interdisciplinario que permita interactuar tanto a la gente que necesita el auxilio de la metodología estadística, como a la gente que maneje dicha metodología.

Sin embargo, esta necesidad no siempre se traduce en la formación de verdaderos grupos interdisciplinarios (2), lo cual hace que la metodología sea utilizada en forma inadecuada y ello trae como consecuencia que la actitud de la gente hacia la Estadística este distorsionada; así en muchos casos los "cálculos estadísticos" se emplean como sosten de investigaciones pobremente concebidas; es decir, estudios que en naturaleza son totalmente subjetivos son investidos de un falso matiz de objetividad.

El poder colaborar en una área diferente a la estadística permite trabajar de manera conjunta con personas de distintas disciplinas y con ello conocer la actitud que tienen estas gentes hacia la utilización de la Estadística, así mismo brinda la oportunidad de cambiar su actitud (1) de tal manera que les permita darse cuenta de cuales son los alcances y limitaciones de este tipo de disciplina científica.

El poder mostrar la utilidad de la Estadística a gente de diferente disciplina fue la causa principal que me impulsó a trabajar en una tesis de este tipo (2), sin embargo, para poder lograr de una manera eficiente que la gente en el área de Fruticultura tenga una aceptación real de la Estadística de manera rápida, se necesita de todo un grupo de traba

jo en Estadística que colabore en el area de Fruticultura (nacionalmente) para que de esta manera a través de ejemplos (elaborados de las mismas investigaciones) en las diferentes disciplinas de relevancia en la Fruticultura, mostrar que aspectos en las investigaciones realizadas se ven apoyados por la utilización de la metodología estadística.

-
- (1) El cambio de actitud se puede lograr mediante la presentación de los resultados obtenidos utilizando las técnicas estadísticas adecuadas para el manejo de información, así como de la metodología utilizada en la captación de dicha información.
 - (2) Un proyecto de investigación interdisciplinario.

INTRODUCCION

El presente trabajo se refiere a una investigación sobre las plagas y enfermedades que atacan a la planta de nopal de tuna blanca, el cual se realizó mediante una encuesta por muestreo en donde se utilizaron tanto las técnicas de muestreo, como las de computación y el análisis estadístico, las cuales se emplearon para llevar a cabo la planeación, conducción y análisis de dicha encuesta.

En el capítulo 1 se presenta todo lo referente a la investigación realizada; el objetivo de este capítulo es ubicar el trabajo en el contexto donde se desarrolló, así como el de mostrar los objetivos tanto de la investigación como de la tesis.

En el capítulo 2 se presenta la realización de la encuesta.- Esta parte es la más importante junto con el capítulo 3, ya que redacta todo el trabajo involucrado en la investigación y lo refiere a los anexos donde se presentan los elementos utilizados para la realización de la encuesta.

Los anexos son importantes, ya que se pretende que el trabajo sirva como parte metodológica en la realización de otros similares, de ahí la necesidad de incluirlos; además los anexos 3 y 4 aseguran la continuidad del trabajo.

En el capítulo 3 se presentan los resultados obtenidos, así como las conclusiones a las que se llegó.

Al final del trabajo se incluyen los Apéndices, el objetivo de esta parte del trabajo es proporcionar algunos elementos teóricos de algunos de los conceptos utilizados en la encuesta. Esta parte es necesaria ya que como se dijo anteriormente, se pretende que el trabajo, además de ser destinado a la Facultad de Ciencias, sirva como consulta en la COMISION NACIONAL DE FRUTICULTURA (CONAFRUT), de ahí también la justificación del porque a lo largo del trabajo se hace énfasis en algunos conceptos involucrados en él.

CONTENIDO.

PAGINA.

CAPITULO 1 : NATURALEZA DE LA INVESTIGACION.....	1
1.1 : Antecedentes	2
1.2 : Justificación	5
1.3 : Objetivos	7
a) De la Investigación	7
b) De la Tesis	9
CAPITULO 2 : LA ENCUESTA	11
2.1 : Definición del Marco	14
2.2 : Diseño de Muestreo	17
2.3 : Organización del trabajo de campo y de oficina	28
CAPITULO 3 : ANALISIS DE LOS DATOS, INTERPRETA- CION DE RESULTADOS Y CONCLUSIONES ..	32
ANEXOS : 1 : CUESTIONARIO UTILIZADO EN LA RECO- LECCION DE LA INFORMACION.....	51
2 : MARCO DE MUESTREO	62
3 : INSTRUCTIVO PARA LA MEDICION DIRECTA.	64
4 : MANUAL DE REFERENCIA DEL ENTREVISTA- DOR	70
5 : CODIFICACION UTILIZADA PARA PROCESAR LA INFORMACION	79
6 : DATOS ORIGINALES	86
APENDICES:	
A : MUESTREO POR CONGLOMERADOS	104
B : MUESTREO SISTEMATICO	108
C : ESTIMADORES DE RAZON	116
REFERENCIA BIBLIOGRAFICA	119

1.1 ANTECEDENTES.

El trabajo de tesis es una investigación que fue realizada en la Comisión Nacional de Fruticultura (CONAFRUT), por lo que es conveniente dar una visión muy general de las funciones de dicha institución ⁽¹⁾.

Una de las funciones de la CONAFRUT es: "Promover, fomentar, estimular y mejorar ⁽²⁾ la fruticultura del país".

Para realizar ésto, la institución cuenta con una infraestructura dentro de la cual, la Subdirección de Investigación y Docencia juega un papel importante, debido a que el objetivo de esta Subdirección es:

"Desarrollar integralmente la fruticultura a través de la investigación, el desarrollo experimental y la docencia".

Para lograr su objetivo, la Subdirección de Investigación y Docencia se divide en:

- a) Escuela Nacional de Fruticultura, donde se lleva a cabo estudios de postgrado (nivel especialización y nivel maestría), - así como cursos cortos a nivel medio de fruticultura general, floricultura general e industrialización de frutas a nivel - doméstico.
- b) División de Investigación y Desarrollo Experimental.

Dentro de esta última división, es donde se realizan las investigaciones necesarias para poder mejorar la fruticultura del país, - avocándose a ello los departamentos de:

(1) Las funciones a las cuales se hace mención, son solamente aquellas que están relacionadas con la investigación realizada.

(2) La investigación que se realizó tuvo como fin la mejora de la fruticultura.

- a) Microbiología y Fitopatología.
- b) Fisiología de Precosecha.
- c) Fisiología de Postcosecha.
- d) Análisis Especiales.
- e) Técnicas básicas agroindustriales.

Dentro de este contexto, surgió la necesidad de realizar una investigación que enfocara parte de sus esfuerzos a la planta de nopal de tuna blanca, debido a que en una de las zonas de mayor importancia - del país en el cultivo de esta planta (zona de San Martín de las Pirámides), se presentaba un problema cuyo efecto era disminuir tanto la producción de frutos como su calidad, dicho problema se atribuía a las diversas plagas y enfermedades que atacaban el nopal de tuna blanca.

Por lo anterior, la investigación general que se realiza sobre tuna blanca tiene como metas.

- a) Informar a los productores de tuna blanca de la zona cuales son las enfermedades que se presentan más frecuentemente.
- b) Saber cuales son las causas específicas que ocasionan que la planta se encuentre enferma.
- c) Poder controlar dichas enfermedades.

Por otro lado, los productores de tuna de la zona tienen en mente exportar el fruto y para ello se requiere tener fruta de primera calidad, la cual se deteriora con el transcurso del tiempo, inclusive - si es de primera calidad al momento del corte, por lo que se recurre a las denominadas técnicas de Fisiología de Postcosecha para la conservación del fruto.

Por todo lo anterior, la investigación general sobre esta - planta se realizo abordando algunos aspectos del problema, la cual dió origen a varios trabajos, incluyendo la presente tesis. Los aspectos -

estudiados son:

- 1) Análisis químicos de los componentes del fruto (ácido ascórbico y azúcares), y tiempo de almacenamiento del mismo.

Al considerar este aspecto surgieron dos trabajos:

- a) Uno donde se probaron diferentes tratamientos para ver cual prolongaba el tiempo de almacenamiento de la fruta, de tal manera que conservara sus valores nutritivos, y características organolépticas ya que sin -- aplicar tratamiento se conserva solamente 8 días como fruto de la. calidad, mientras que como resultado del estudio al aplicar uno de los tratamientos se puede -- prolongar la vida útil del fruto en almacenamiento -- hasta 25 días.
 - b) El otro es continuación del anterior y solamente se -- probaba un sólo tratamiento (el que resultó mejor en el trabajo anterior) en diferentes concentraciones.
- 2) Identificación de agentes patógenos⁽³⁾ en el fruto: considerando este aspecto surgió una tesis de maestría, la cual tenía como objetivo la identificación de agentes patógenos que producen enfermedades en el fruto, así como probar nuevos métodos de manejo del fruto y en base a ésto poder recomendar -- que manejo postcosecha debe utilizarse para prolongar la vida útil del fruto en el almacenamiento.

Los aspectos 1) y 2), se realizaron en la etapa de postcosecha del fruto, donde la fruta que se utilizó provenía para cada caso de una sola huerta.

(3) Microorganismo que causan algun tipo de alteración al desarrollo del fruto.

Un tercer aspecto abordado en la investigación es:

- 3) Identificación y cuantificación de enfermedades en la planta de nopal de tuna blanca.

A diferencia de los dos anteriores este fue estudiado en la etapa de precosecha del fruto.

La investigación se realizó considerando el apoyo de las técnicas de muestreo, de esta manera surgieron dos tesis que fueron desarrolladas en forma paralela, una de Biología y la presente.

1.2 JUSTIFICACION.

Como se mencionó anteriormente, en la zona de San Martín de las Pirámides, Edo. de México, se postulaba que:

"Diversas plagas y enfermedades atacaban el nopal de tuna blanca", ésto disminuía por un lado, la calidad de la fruta y por otro, la cantidad cosechada de la misma.

Debido a la naturaleza del problema, éste requería que la solución encontrada fuera válida para toda la zona⁽⁴⁾.

Para lograr conclusiones válidas para toda la zona, teníamos dos caminos a seguir:

- 1) Realizar un censo.

(4) Anteriormente, las investigaciones realizadas sobre el problema en estudio estaban limitadas, en el sentido de que las conclusiones que se pudieran obtener de dichos estudios sólo eran válidas para una pequeña parte de la zona, debido a que los estudios se realizaban con frutos de una huerta.

- 2) Hacer un muestreo, es decir, examinar sólo algunos de los objetivos de estudio (parte de la población) y extender nuestros hallazgos a toda la población en estudio, utilizando el proceso de inferencia Estadística.

El realizar un censo en la zona resultaba impráctico, debido a que la zona está compuesta por una gran cantidad de huertas⁽⁵⁾, así como por los pocos recursos humanos con que se contaba para abordar la investigación sobre la identificación y cuantificación de las enfermedades en la planta de nopal de tuna blanca (sólo 4 personas), de tal manera que se decidió utilizar los procedimientos de muestreo probabilístico, ya que como el problema se presentaba en la zona, entonces necesitamos conclusiones para ella.

Por otro lado, trabajos de este tipo (utilizando encuestas por muestreo) no se habían realizado en la CONAFRUT⁽⁶⁾, de tal manera que el presente serviría para sugerir una nueva línea de investigación.

Por otro lado es necesario mencionar que prácticamente es el primer trabajo en tuna blanca que se realiza en la etapa de precosecha del fruto en el Departamento de Microbiología y Fitopatología de la CONAFRUT; por esta razón, el trabajo es de gran interés, debido a que en trabajos posteriores se podrá detectar la relación que existe entre las enfermedades que presenta la planta en la etapa de precosecha, y las que presenta el fruto en la etapa de postcosecha; de tal manera que se puede saber que enfermedades (de las que se presentan en el nopal antes de que se corte el fruto) son las que afectan a la tuna en postcosecha.

Por otro lado, el trabajo de tesis presente servirá para:

- a) Mostrar a la gente la utilidad de la técnica de muestreo

(5) Además hay que considerar la gran cantidad de plantas que existen en cada huerta, lo que conlleva a costos altos.

(6) Ya que la línea seguida en las investigaciones es la relacionada con diseño de experimentos.

utilizada, mediante la presentación de los resultados obtenidos, así como de la metodología, comentando los alcances y limitaciones del estudio y el porque de estas limitaciones.

- b) Mediante el inciso anterior, mostrar a la gente que el tipo de análisis realizado sirve, y de esta manera poder ampliar el campo de aplicación real, de la estadística, adaptando principios teóricos y situaciones prácticas de la fruticultura en México.
- c) Cumpliendose los puntos a) y b), surge la necesidad de utilizar gente del área de estadística (con tendencia al muestreo y análisis de datos) que entre en apoyo para la planeación, conducción, procesamiento y análisis de información en la investigación frutícola, para que de esta manera la investigación básica al servicio de la fruticultura aplicada tenga mayor fundamento teórico.
- d) Generar material para la ESCUELA NACIONAL DE FRUTICULTURA, debido a que en las maestrias que imparte esta institución se da un curso de muestreo; por lo tanto, el trabajo servirá como apoyo de la clase y como material de consulta.

1.3 OBJETIVOS.

A) De la investigación.

Debido a la naturaleza del problema, éste dió origen a ciertas preguntas⁽⁷⁾ que era necesario contestar, y que al ser contestadas queda

(7) Preguntas como las siguientes:

- ¿ El daño esta muy avanzado en la planta ?
- ¿ La ubicación de la huerta influirá para que haya enfermedades ?, y otras de este tipo.

rían satisfechos los objetivos de la investigación; además por los conocimientos que se tenía acerca de la zona de estudio ⁽⁸⁾ y por lo que sabía sobre la naturaleza de las plantas de nopal, se plantearon los siguientes objetivos e hipótesis de trabajo:

Objetivos:

- 1) Principal: Identificación y cuantificación de las plagas y enfermedades más frecuentes en el nopal de tuna blanca.

- 2) Secundarios:
 - i) Conocer las causas que generan dichas enfermedades.
 - ii) Conocer la influencia que tienen las enfermedades sobre la producción.

Hipótesis:

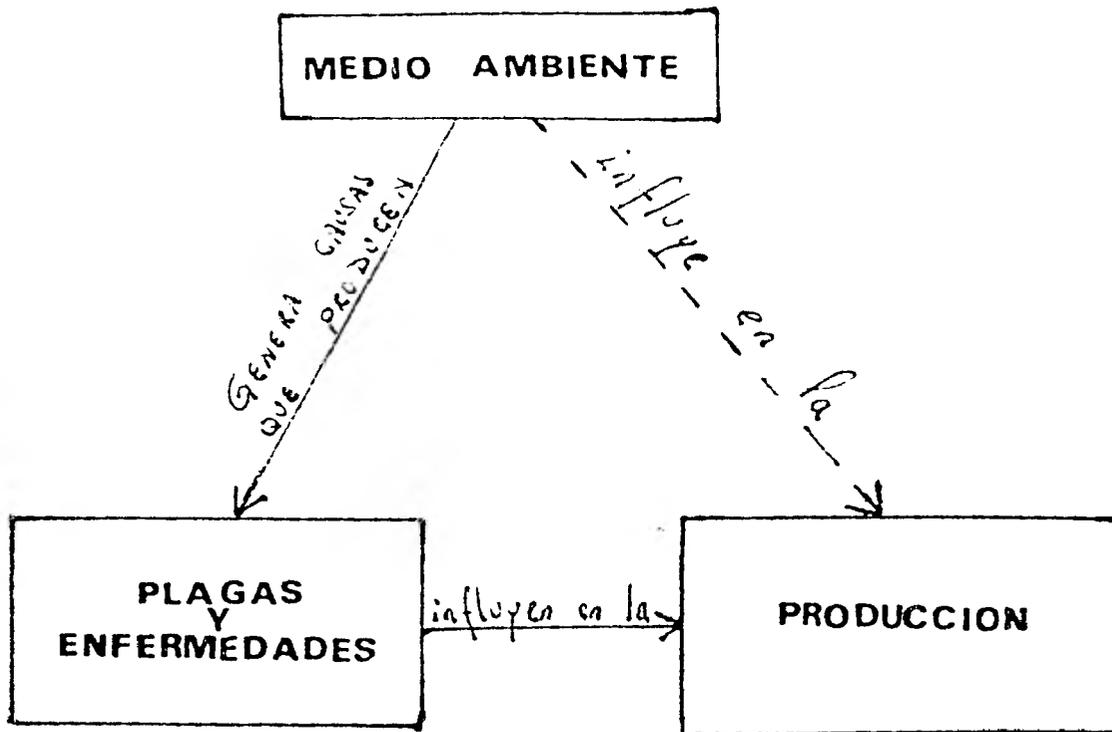
- 1) El estado fitosanitario de las huertas está en relación con la limpieza.

- 2) El grado de severidad de una plaga o enfermedad en una huerta está en relación directa con la producción de dicha huerta.

Como se puede observar, el objetivo principal y los objetivos secundarios están entrelazados entre sí y éstos a su vez con las hipótesis de trabajo; debido a que existe una estrecha relación entre las enfermedades, el medio ambiente que rodea a la planta y la producción de frutos que pueda tener dicha planta.

Gráficamente la relación de los objetivos y la hipótesis se puede visualizar de la siguiente manera.

(8) Ya que anteriormente se ha trabajado con tuna en esa misma zona.



En donde la flecha punteada indica que el medio ambiente influye de una manera directa para que haya una mayor o menor producción, y además dependiendo del medio ambiente que rodea a la planta ésta puede ser atacada por enfermedades, las cuales van a mermar su producción.

La hipótesis así como el primer objetivo secundario están involucrados en la flecha con la punta en el "Bloque de enfermedades"; mientras que el objetivo principal, el segundo objetivo secundario y la hipótesis 2 están comprendidos en la flecha que va del "bloque de enfermedades" al "bloque de producción"

B) De la tesis:

En la parte correspondiente a justificación se hace mención de algunas de las cosas para las cuales se pretende sirva la tesis, además en el prólogo se mencionan los motivos por los cuales se desarrolló la investigación, de tal manera que en esas partes ya han sido expuestos algunos de los objetivos de tesis, entre los cuales se encuentran los siguientes:

- 1) Mostrar la utilidad del muestreo a gente cuya formación profesional no es de "tipo estadístico".

- 2) Crear la necesidad de utilizar las técnicas de muestreo, así como las técnicas de computación en las investigaciones en Fruticultura.
- 3) Ampliar el campo de aplicación real de la Estadística (en este caso del muestreo). ○

CAPITULO 2
LA ENCUESTA

La ejecución de la encuesta para la identificación y cuantificación de enfermedades en las plantas de nopal de tuna blanca, fue realizada en los meses de mayo-septiembre de 1980 (tiempo de referencia⁽¹⁾), la cual fue financiada por la Subdirección de Investigación y Docencia de la CONAFRUT.

Para los fines de la encuesta la información fue colectada - por medio de un cuestionario, el cual fue dividido en dos rubros principales, dependiendo de la procedencia de la información obtenida; uno que corresponde a la entrevista y el otro que corresponde a la medición directa. (Véase Anexo 1). Se diseñó así porque se consideró que resultaba difícil saber donde vivía la gente que cuida las huertas (predios) y eso dificultaba la localización de los posibles entrevistados; para prevenirse de esta situación, se planeo que el cuestionario se dividiera en los dos rubros antes mencionados, de tal manera que los objetivos establecidos fueron alcanzados aún cuando no fuera posible realizar la entrevista, y que con la parte correspondiente a la medición directa pudieran cumplirse. La recolección de la información en los dos rubros fue hecha en fechas diferentes, realizando la medición directa en el período de mayo-julio de 1980, en tanto que la entrevista se llevó a cabo en el mes de agosto y principios de septiembre de 1980, ya que en la mayoría de los casos no era posible realizarlas al mismo tiempo, debido a las razones antes mencionadas.

Para diseñar el cuestionario se tomó en cuenta además de los objetivos de la encuesta, los conocimientos que sobre las enfermedades tenía la persona que fuera a realizar la parte correspondiente a medición directa, de tal manera que las enfermedades consideradas en el cuestionario estuvieran bien identificadas y las escalas de medida, de intensidad estuvieran definidas; para lo anterior; antes de realizar la encuesta se llevaron a cabo una serie de visitas a la zona de estudio para coleccionar material vegetativo y con esto determinar cuales eran las enfermedades más frecuentes y poder identificarlas, así como definir las escalas de medición, además en el recorrido de la zona la gente empieza a formular algunas hipótesis con las cuales se trabajaría posteriormente; para

(1) Periodo en que se recabó la información de la encuesta.

lelamente a ésto se fue construyendo el marco de muestreo el cual más tarde nos permitiría conocer la probabilidad de selección de cada elemento de la población de estudio (ver 2.1 Definición de Marco). Posteriormente, en la recolección de la información en el rubro correspondiente a la entrevista (parte I y II) se utilizó el método del cuestionario - (en donde se utiliza un conjunto de preguntas estandares); por otro lado, en el rubro correspondiente a medición directa (partes III, IV y V) se preparó una lista de renglones y se obtuvo información sobre ellos; la recolección de la información en la parte IV se obtuvo una vez que había sido obtenida la información de la parte III, la cual requería el recorrido dentro de la huerta y por lo tanto permitía tener mayor veracidad en la información levantada en la parte IV. En cuanto a la parte V, esta información era recabada paralelamente a la realización de la parte III y IV.

En relación al análisis de la información, dada la naturaleza del problema, la encuesta tenía como meta, aparte de la cuantificación de las enfermedades en la población en estudio (objetivo principal), dar una explicación tentativa acerca de las posibles causas que las originaron, así como ver que tanta influencia tienen estas en la producción (objetivos secundarios); de tal manera que la encuesta recolecta información que sirve para ambos propósitos, por tal motivo los análisis de la encuesta se consideran como descriptivos y analíticos, es decir, la encuesta es una combinación de ambos tipos (descriptivas y analíticas⁽²⁾).

Dadas las dos hipótesis planteadas anteriormente, los subgrupos que se pueden comparar para verificarlas o rechazarlas parcialmente son los siguientes:

-
- (2) Las encuestas por muestreo pueden ser clasificadas a grandes rasgos en dos grupos:
- a) Descriptivas: El objetivo es obtener información que describa a la población en estudio, en el caso presente, la cuantificación.
 - b) Analíticas: En este tipo de encuestas se hacen comparaciones - entre subgrupos diferentes de la población (poblaciones con distinto grado de generalidad), las cuales nos permiten verificar (o rechazar) hipótesis; en este caso es lo que se refiere a la explicación.

- a) Huertas limpias.
- b) Huertas no limpias.
- c) Huertas con poca incidencia (escala 0 y 1).
- d) Huertas con incidencia severa (escala 2 y 3).

Las comparaciones se realizan con subgrupos con igual grado de generalidad y se tendría por ejemplo: a) y b) para probar la hipótesis 1, y los subgrupos c) y d) para probar la hipótesis 2. Las huertas comparadas entre los dos últimos subgrupos deben ser aproximadamente de la misma edad entre otras cosas, para que la incidencia refleje o no la influencia real sobre la producción.

Los subgrupos a), b), c) y d) en la teoría de muestreo son llamados dominios de estudio⁽³⁾.

2.1 DEFINICION DEL MARCO.

Es necesario definir aquellos elementos sobre los cuales deseamos recolectar información, y usando métodos probabilísticos, seleccionar una muestra de ellos, la cual nos va a permitir llevar a cabo inferencias sobre todos los elementos de los cuales necesitábamos información, dichos elementos deben estar bien caracterizados y definidos en términos de: a) el contenido, b) las unidades, c) la extensión y d) el tiempo.

Los elementos definidos de esta manera forman una población a la cual se llama población objeto. Para seleccionar la muestra por métodos probabilísticos es necesario conocer la probabilidad de selección de cada elemento y para esto necesitamos establecer el marco de muestreo.

(3) Cuando la población es subdividida en un número de subpoblaciones, dentro de las cuales se realizan estimaciones diferentes, dichas subpoblaciones son conocidas como dominios de estudio.

Para establecer el marco de muestreo, se definió la población objeto como:

"Las plantas de nopal de tuna blanca mayores a 3 años en la zona de San Martín de las Pirámides en 1980".

Es conveniente mencionar que la población objeto no es la misma que la población que utilizamos comunmente en Estadística (población estadística); la primera se refiere a los objetos a quienes les queremos medir alguna característica de interés, mientras que la otra se refiere al conjunto de mediciones hechas sobre los objetos.

El marco fue construido por medio de fotografías aéreas, mapa topográfico, así como con la visita directa a las huertas de tuna blanca en la zona de San Martín de las Pirámides, el cual se llevó un tiempo de aproximadamente 5 meses para quedar terminado, de donde finalmente el marco quedó constituido como se presenta en el Anexo 2. Además, en las visitas realizadas en la zona también se tomaba información relacionada con el tipo de suelo y otros aspectos que posteriormente sirvieron para estratificar.

Debido fundamentalmente a la definición de población objeto y a la manera como fue construido el marco, este presenta el siguiente problema:

Contiene elementos extraños como son:

- a) Las huertas de nopal de tuna blanca que tienen una edad menor o igual de 3 años.
- b) Las huertas de maíz.

En cuanto a las huertas menores o iguales a 3 años estas no era posible detectarlas por medio de las fotografías aéreas, ni tampoco con la visita directa, ya que no siempre es posible saber de manera directa la edad de las plantas, incluso en algunos casos, solamente hasta que se realizó la entrevista se supo si las huertas eran menores e iguales a 3 años, y como la entrevista se hizo después de la medición -

directa, se realizó la medición en 13 huertas que no correspondían a nuestra población objeto, de tal manera que la proporción obtenida en la muestra fue de .288.

En cuanto a los elementos extraños mencionados en el inciso b), a pesar de haber realizado la visita directa no se tenía seguridad absoluta de que no existieran dichos elementos en nuestro marco.

Por otro lado, se decidió seleccionar las huertas de refacción para el caso de que la medición directa no se pudiera realizar por causas imprevistas y por si la huerta seleccionada correspondía a una de maíz.

Debido a que en la muestra no se presentó ningún elemento extraño de los mencionados en el inciso b), es posible concluir que es poco probable que dichos elementos se encuentren en nuestro marco, por otro lado, dado que no hubo ningún contratiempo no se utilizaron las huertas de refacción.

Por último mencionaremos que el tipo de marco utilizado para la investigación realizada fue un marco area ⁽⁴⁾, del cual se ha dado la siguiente definición.

"Un marco por areas consiste en un agregado de unidades identificables de terreno (segmentos), los cuales son generalmente áreas pequeñas y en donde cada unidad de observación está asociada con una y sólo una de tales áreas" (14).

(4) Existen dos tipos de marco principalmente.

- a) Marco Area: En donde las unidades de muestreo finales son áreas terrestres y las unidades informadoras sólo pueden identificarse por reglas geográficas que las asocien con áreas terrestre en la muestra.
- b) Marco Lista: Es aquel que no satisface las condiciones de marco área.

En nuestro caso, las unidades de muestreo fueron predios perfectamente identificables en el marco, de tal manera que al realizar la entrevista con el encargado de la huerta, éste pudiera ubicar su huerta en el mapa y así poder asegurar que en las huertas donde se realizó la medición directa fueran las mismas en donde se iba a obtener la información por medio de la entrevista.

2.2 DISEÑO DE MUESTREO.

Una vez construido el marco de muestreo (mapa) se estratificó considerando principalmente las diversas zonas geográficas, así como las delimitaciones naturales que tenían estas y como consecuencia se suponía que había diversas condiciones ambientales en ellas; conjuntamente a esta información se consideró también la relacionada con el tipo de suelo, la cual fue obtenida por medio de las visitas directas a la zona. El número de estratos (regiones) fue de cuatro.

En la planeación de la encuesta fue necesario decidir sobre que tamaño de muestra (n) era necesario utilizar para llevarla a cabo sin que se desperdiciaran recursos y además para que los resultados fueran útiles dentro de un margen de error preestablecido.

Es necesario mencionar que para determinar el tamaño de muestra satisfactoriamente se necesita conocer que exactitud se requiere, es decir, hay que establecer un margen de error (d) en los resultados, sin embargo, hay que hacer notar que no se puede garantizar una exactitud dentro del porcentaje deseado, excepto midiendo todas las unidades de la población. No importa que tan grande sea tomando n , hay posibilidad de obtener una muestra muy desafortunada que esté en error en más del porcentaje deseado, para esto necesitamos aceptar que hay un pequeño riesgo α , que hay que estar dispuestos a correr de que el error real sea mayor que d , es decir, queremos:

$$\Pr(|\bar{y} - \bar{Y}| \geq d) = \alpha$$

Donde:

- \bar{Y} es el promedio poblacional desconocido.
- \bar{y} es el estimador del promedio poblacional
- d es el margen de error elegido, y
- α es una probabilidad pequeña.

Se supone un muestreo simple aleatorio y además que \bar{y} se distribuye normalmente, entonces:

$$\sigma^2_{\bar{y}} = (1 - f) \frac{S^2}{n} = (1 - \frac{n}{N}) \frac{S^2}{n}$$

entonces

$$\sigma_{\bar{y}} = \sqrt{\frac{N-n}{N}} \frac{S}{\sqrt{n}}$$

Donde:

- S^2 = Varianza poblacional
- n = Tamaño de la muestra
- $f = \frac{n}{N}$ = Fracción de muestreo
- N = Tamaño de la población

Por lo tanto, la fórmula que conecta a n con el grado de precisión deseado es:

$$d = t \sqrt{\frac{N-n}{N}} \frac{S}{\sqrt{n}}$$

en donde, t es la abscisa en un punto que separa una área α en los extremos de la distribución de t de Student con $n - 1$ gl (para muestras grandes se puede tomar la distribución normal estandar). Resolviendo para n encontramos:

$$n = \frac{\left(\frac{ts}{d}\right)^2}{1 + \frac{1}{N} \left(\frac{ts}{d}\right)^2}$$

Si hacemos:

$$n_0 = \left(\frac{t_s}{d}\right)^2$$

entonces:

$$n = \frac{n_0}{1 + n_0/N}$$

Si se considera que $\frac{n_0}{N}$ es despreciable, entonces, n_0 es una aproximación satisfactoria de n .

Hay que hacer notar que se ha obtenido una fórmula para n , pero n depende de alguna propiedad de la población que va a ser muestreada. En este caso, la propiedad es la varianza de la población (S^2).

En la práctica para estimar la varianza de la población para la determinación del tamaño de muestra, procedemos de las siguientes formas;

- 1). A través de los resultados de una encuesta piloto.
- 2). Por conjetura sobre la estructura de la población, secundada por algunos resultados matemáticos.

Para obtener el tamaño de muestra en la investigación presentada en esta tesis, se supuso un margen de error del 5% y el riesgo que se quería correr era del 5% también, además, la estimación de la varianza de la población (S^2) se obtuvo por medio de información que se tenía de la feria de la tuna de 1979 en la zona de San Martín de las Pirámides, acerca de la "Enfermedad del Oro", "El Picudo" y "El Barrenador"; esta información provenía de 32 huertas en donde se hizo la medición en 15 plantas por cada huerta para detectar si existía alguna plaga o enfermedad. Hay que mencionar que el tamaño de muestra obtenida con esta información fue el correspondiente a las unidades de primera etapa (huertas). Como el diseño de muestreo es de dos etapas, es necesario calcular el tamaño de muestra para las unidades de segunda etapa (Plantas),-

de aquí se deduce que es necesario estimar la varianza poblacional entre plantas; para ésto se propuso el siguiente modelo:

$$Y_{ij} = \mu + H_i + \epsilon_{ij} \quad \text{-----} \quad (1)$$

$j = 1, 2, \dots, 15$ (plantas)

$i = 1, 2, \dots, 12$ (huertas).

Suposiciones:

$$H_i \sim N(0, \sigma_H^2)$$

$$\epsilon_{ij} \sim N(0, \sigma^2)$$

En el modelo (1):

Y_{ij} - Representa la intensidad de daño encontrado en la
 j - ésima planta de la i - ésima huerta.

μ - es el promedio (cantidad en %) de la enfermedad en
la zona de San Martín de las Pirámides.

$$H_i = \mu_i - \mu$$

(Diferencia entre el contenido promedio de la
 i -ésima huerta y el contenido promedio sobre toda
la población).

Al incluir este término se está tomando en cuenta el hecho -
de que el promedio de daño varía de una huerta a otra. Cada huerta de
la población tiene su valor de H_i , de suerte que podemos considerar H_i
como una variable aleatoria con una distribución sobre la población, de
aquí que resulta interesante estimar su varianza.

ϵ_{ij} - Se incluye porque la determinación está sujeta a error de medición, error de apunte, etc.

Se supone ϵ_{ij} independiente de H_i .

σ_H^2 - Se le denomina en este caso, componente de varianza por huerta.

Del modelo (1) se puede obtener que con p -plantas de cada una de las h -huertas, la media muestra $\bar{Y}..$ esta dada por:

$$\bar{Y}.. = \mu + \bar{H} + \bar{\epsilon}..$$

Donde:

\bar{H} - Es la media de los valores independientes de H_i (una por cada huerta), y

$\bar{\epsilon}..$ - Es la media de $h \times p$ ϵ_{ij} independientes, entonces la varianza de $\bar{Y}..$ (como estimación de μ) es:

$$V(\bar{Y}..) = \frac{\sigma_H^2}{h} + \frac{\sigma^2}{h \times p} = \frac{\sigma^2 + p \cdot \sigma_H^2}{h \times p} \quad (2)$$

La estimación del numerador de la ecuación anterior la proporciona el análisis de varianza del modelo (1); una vez hecha la estimación de σ_H^2 y de σ^2 , en el modelo (2) se puede ir variando h & p y determinar para que combinación de h & p , $V(\bar{Y}..)$ es mínima, Hay que hacer notar que h va a tener un valor mínimo y es el que se obtuvo cuando se determinó el tamaño de muestra para las unidades de primera etapa.

Finalmente, el tamaño de muestra utilizado en la investigación fue de:

45 huertas.

10 plantas en cada huerta, y en cada planta se seleccionaron 12 cladodios para medir intensidad de daño, de acuerdo a la escala previamente fijada.

Las 45 huertas fueron distribuidas dentro de los estratos de acuerdo al número de huertas que contenía cada región o estrato, es decir, se hizo una asignación proporcional a la cantidad de huertas en cada estrato, el número de huertas muestra fue de 17, 14, 12 y 2 en las regiones 1, 2, 3 y 4 respectivamente, más dos huertas de refacción para cada una de las regiones. El número de huertas totales en cada región era de 464, 376, 327 y 55 para las regiones 1, 2, 3 y 4 respectivamente.

Dada la naturaleza del problema y los objetivos de la investigación, el diseño de muestreo utilizado para la obtención de la muestra fue un diseño en dos etapas, en donde las unidades de primera etapa fueron las huertas y las unidades de segunda etapa fueron las plantas. Para los fines de la encuesta se entendía como huerta:

"Todo terreno que fuera trabajado por la misma persona utilizando los mismos implementos de trabajo, así como las mismas labores culturales y además que el terreno tuviera una delimitación del área trabajada".

Para estimar la intensidad promedio del % de predios afectados en la zona se utilizó el estimador que a continuación se describe:

Notación:

Valores Poblacionales.

H_r = Número de huertas totales en la región r -ésima ($r = 1, 2, 3, 4$), que vendrían siendo el número de unidades primarias.

Pr_w = Número total de plantas en la huerta w -ésima de la región r -ésima ($w = 1, 2, \dots, H_r$). Es decir, es el número de unidades secundarias en la unidad primaria w -ésima

y_{rwp} = Valor de la medición en la planta p -ésima de la huerta w -ésima en la región r -ésima ($p = 1, 2, \dots, 10$).

Y_{rw} = Total de la huerta w -ésima.

$$Y_{rw} = \sum_{p=1}^{Prw} y_{rwp}$$

\bar{Y}_{rw} = Promedio de la huerta w -ésima.

$$\bar{Y}_{rw} = \frac{1}{Prw} Y_{rw} = \frac{1}{Prw} \sum_{p=1}^{Prw} y_{rwp}$$

$Y_{r..}$ = Total de la población en la región r -ésima

$$\begin{aligned} Y_{r..} &= \sum_{w=1}^{Hr} Y_{rw} \\ &= \sum_{w=1}^{Hr} \sum_{p=1}^{Prw} y_{rwp} \end{aligned}$$

Valores muestrales y estimador:

h_r = Número de huertas en la muestra en la región r -ésima ($r=1, 2, 3, 4$).

= Número de unidades primarias en la muestra.

pr_w = Número de plantas en la muestra en la huerta w -ésima de la región r -ésima ($w = 1, 2, \dots, h_r$)

= Número de unidades secundarias muestreadas en la unidad primaria w -ésima ($pr_w = 10$)

\bar{y}_{rw} = Promedio de la muestra de plantas en la huerta w -ésima.

$$\bar{y}_{rw} = \frac{1}{p_{rw}} \sum_{p=1}^{p_{rw}} y_{rwp}$$

\hat{Y}_{rw} = Total estimado de la unidad primaria w-ésima.

$$\hat{Y}_{rw} = P_{rw} \bar{y}_{rw}$$

$\hat{Y}_{r..}$ = Total estimado para la población en la región r-ésima.

$$\hat{Y}_{r..} = H_r \frac{\sum_{w=1}^{h_r} \hat{Y}_{rw}}{h_r}$$

$$= \frac{H_r}{h_r} \sum_{w=1}^{h_r} P_{rw} \bar{y}_{rw}$$

Para estimar el promedio de toda la población

$$\bar{Y} = \frac{Y_{r..}}{H_r} = \frac{\sum_{w=1}^{h_r} P_{rw}}{H_r}$$

era necesario conocer los valores de P_{rw} para todas las unidades de la población objeto, como no fue posible obtenerlos el estimador propuesto toma la forma de un estimador de razón en donde es necesario estimar el numerador y el denominador.

El estimador propuesto es el siguiente

$$\hat{R} = \hat{Y} = \frac{\frac{H_r}{h_r} \sum_{w=1}^{h_r} \frac{P_{rw}}{p_{rw}} \sum_{p=1}^{p_{rw}} y_{rwp}}{\frac{H_r}{h_r} \sum_{w=1}^{h_r} \frac{P_{rw}}{p_{rw}} \sum_{p=1}^{p_{rw}} x_{rwp}} = \frac{\sum_{w=1}^{h_r} (P_{rw} \bar{y}_{rw} / p_{rw})}{\sum_{w=1}^{h_r} (P_{rw} \bar{x}_{rw} / p_{rw})}$$

$$\hat{Y} = \frac{\hat{Y}}{\hat{X}} \quad (1)$$

Donde:

$h_w = \frac{h_r}{H_r}$ es la probabilidad de seleccionar la unidad w -ésima en la muestra

$$\sum_{w=1}^{h_r} h_w = 1, \quad \sum_{w=1}^{h_r} \bar{x}_{rw} = 1$$

además \hat{Y} & \hat{X} estiman insesgadamente a Y & X respectivamente.

$$\text{donde: } X = \sum_{w=1}^{h_r} Pr_w$$

Una expresión más sencilla para la ecuación (1) es:

$$\hat{Y} = \frac{\sum_{w=1}^{h_r} Pr_w \bar{y}_{rw}}{\sum_{w=1}^{h_r} Pr_w} = \frac{\sum_{w=1}^{h_r} \frac{Pr_w}{pr_w} \sum_{\rho=1}^{pr_w} y_{rwp}}{\sum_{w=1}^{h_r} Pr_w} = \frac{\hat{Y}}{\hat{X}} \quad (2)$$

de donde el estimador de la varianza de este estimador es:

$$\hat{V}(\hat{Y}) = (1-h_r) \frac{1}{h_r} \frac{1}{\bar{Pr}^2} S_p^2 + \frac{1}{h_r \bar{Pr}^2} \sum_{w=1}^{h_r} Pr_w^2 (1-pr_w) \frac{S_s^2}{pr_w} \quad (A)$$

donde:

$$\bar{Pr} = \frac{\sum_{w=1}^{h_r} Pr_w}{h_r}$$

$$S_p = \frac{\sum_{w=1}^{h_r} Pr_w^2 (\bar{y}_{rw} - \hat{Y})^2}{h_r - 1}$$

$$S_s = \frac{\sum_{\rho=1}^{pr_w} (y_{rwp} - \bar{y}_{rw})^2}{pr_w - 1}$$

Ahora bien, si se considera que (1) es un estimador de razón en un sentido estricto, entonces una medida de la variabilidad para este estimador la proporciona el error cuadrático medio (E.C.M.), ya que en los estimadores sesgados se utiliza el E.C.M. para medir la variabilidad. El E.C.M. es una medida analoga a la varianza que es la que se utiliza en los estimadores insesgados. Cabe hacer mención que los estimadores de razón son sesgados sin embargo si la muestra es grande (≥ 30) el sesgo es pequeño (Ver apéndice C).

Supongamos que la expresión (1) nos denota un estimador de razón en sentido estricto, entonces, si las u.p.m. (unidades primarias de muestreo) son seleccionadas con probabilidades iguales sin reemplazo y al azar, entonces para estimar la media, el estimador resultante es:

$$\hat{R} = \frac{\sum_{w=1}^{hr} Prw \bar{y}_{rw}}{\sum Prw}$$

y como una estimación de la muestra del E.C.M. se tiene:

$$\frac{(1-hr)}{Hr} \frac{Hr^2}{hr Pr} \frac{\sum_{w=1}^{hr} Prw (\bar{y}_{rw} - \hat{R})^2}{hr - 1} + \frac{Hr}{hr Pr} \frac{\sum_{w=1}^{hr} Prw}{prw} \frac{\sum_{p=1}^{prw} (y_{rwp} - y_{rw})^2}{prw - 1} \quad (B)$$

Si en esta última expresión hacemos:

$$\hat{Pr} = Hr \bar{Pr} = Hr \frac{\sum_{w=1}^{hr} Prw}{hr}$$

Entonces la expresión (B) es igual a la expresión (A), es decir, en este caso $\hat{V}(\hat{Y}) = E.C.M.(\hat{Y})$, entonces el estimador propuesto es insesgado.

Ahora bien, en este estimador el valor de la medición en la unidad secundaria p -ésima de la unidad primaria w -ésima denotado anteriormente como y_{rwp} nos representa, la intensidad promedio en una planta y se obtiene por el método de Townsend y Heuberger (5). El método se resume en la fórmula siguiente :

$$y_{rwp} = \frac{\text{Suma } (n \times v)}{Z \times N} \times 100$$

donde:

- y_{rwp} = Intensidad de ataque o de daño
n = Número de cladodios en cada categoría de infestación
v = Valores numéricos de las categorías de infestación (valores escalares)
Z = Valor numérico de la categoría máxima ó número de categorías diferentes de cero
N = Número total de cladodios medidos en la planta.

El procedimiento general de muestreo en la zona de San Martín de las Pirámides fue el siguiente:

- a) Se tomó cada una de las 4 regiones geográficas definidas como estratos.
- b) En cada región se seleccionaron huertas con muestreo simple al azar sin reemplazo. El número de huertas seleccionadas en cada región fue proporcional al número de huertas por región.
- c) Dentro de cada huerta (donde era necesario estimar el número total de plantas) se seleccionaron 10 plantas con muestreo sistemático en dos dimensiones.
- d) Para cada planta (donde era necesario contar con el número total de cladodios) se seleccionaron 12 cladodios con muestreo simple al azar sin reemplazo, en donde se levanto la información.

El método de estimación utilizado en la investigación está dividido en dos fases.

- F.1: Predicción de un aspecto de interés en estudio de la planta de nopal de tuna blanca en una huerta.

F.2: Selección de la muestra de predios cuyo aspecto de interés se estima. Agregación de estas predicciones para la estimación del aspecto en cuestión a nivel regional y a nivel zona de San Martín de las Pirámides.

En la fase 1 se seleccionan con muestreo sistemático en dos dimensiones 10 plantas de nopal de tuna blanca y en cada planta se seleccionan 12 cladodios. (Vease anexo 3).

En la fase 2 se seleccionan las huertas en cada una de las regiones con muestreo aleatorio simple sin reemplazo.

En resumen, el diseño de muestreo utilizado fue Bietapico - - Estratificado, en donde las unidades de primera etapa (huertas) se seleccionan al azar sin reemplazo y las unidades de segunda etapa (plantas) se seleccionan con muestreo sistemático en dos dimensiones.

2.3. ORGANIZACION DEL TRABAJO DE CAMPO Y DE OFICINA

Como es bien sabido, la calidad de la información colectada por el personal asignado a las labores de campo determina en gran medida la calidad de los datos a tabularse de la encuesta; es entonces fundamental que las personas que realizan dichas labores posean características adecuadas para el buen desarrollo del trabajo, además es conveniente que estén enteradas de los objetivos de la encuesta, de esta manera se interesan en el trabajo que se les asigne y pondrán su máximo esfuerzo en él. Sin embargo, no solamente el esfuerzo de la gente que realiza las labores de campo es suficiente, sino que también es necesario preparar instructivos que les permita realizar el trabajo en forma adecuada. En nuestro caso se prepararon los siguientes instructivos:

- 1) Instructivo para la medición directa (vease anexo 3)
- 2) Manual de referencias del entrevistador (vease anexo 4)

En donde el instructivo 1 es creado para que la selección de las unidades de segunda etapa sean hechas en forma correcta, en tanto que el 2 es creado con dos propósitos básicos:

- a) Explicar con claridad el trabajo del encuestador a cada paso.
- b) Hacerle sentir al encuestador que debe ser capaz de crear una atmósfera cordial y hacer que el respondiente se sienta cómodo.

Las labores de campo para esta encuesta las realizaron fundamentalmente dos personas.

Por una parte, la medición de la intensidad de las enfermedades la hizo un biologo (quien como ya se dijo, se entreno previamente en la identificación de enfermedades).

Por otro lado, un actuuario llevaba a cabo la supervisión del trabajo de campo, ya que verificaba que las instrucciones se llevaron a cabo tal como se habían planeado en el escritorio, de esta manera es posible darse cuenta de cuales son las dificultades que se presentan en el campo y así poder apreciar que tan real es lo que se propone en el escritorio y por lo tanto tener una participación tanto en la planeación como en la conducción de la encuesta

Las entrevistas las realizó una persona (biólogo) y se llevaron a cabo inmediatamente después de haber realizado la medición directa.

Una vez realizadas las entrevistas se procedió a revisar los cuestionarios, de donde se decidió que algunas preguntas no se incluyeran en el análisis, esto se hizo principalmente por dos razones.

- 1) Porque las respuestas eran muy ambiguas y esto ocasionaba que para una misma pregunta, las respuestas no fueran comparables.
- 2) Porque habia ausencia de respuesta.

Para el primer punto se presenta el caso de la pregunta número

10, en donde algunas de sus respuestas fueron las siguientes:

- "2 carretas"
- "1 bote"
- "1 carretilla"

de manera similar sucedia en las preguntas 11, 14 y 16.

En cuanto a la ausencia de respuesta, esta se debió principalmente porque la gente entrevistada no sabia la respuesta.

En conclusión, para el análisis no se incluyeron las preguntas 6,8,10,11,14 y 16.

Para facilitar el análisis necesitamos las respuestas en términos numericos, sin embargo no todas las preguntas del cuestionario aplicado pueden contestarse en esos términos debido al tipo de preguntas que se incluyen en él, para ellas se elaboró una lista de codigos de acuerdo a la frecuencia con que se repetia cada respuesta. Ejemplo de este tipo de preguntas es la número 9 entre otras (vease anexo 1).

El siguiente paso fue la transferencia de información a tarjetas de computo, en donde la codificación utilizada para los datos perforados se presenta en el anexo 5, mientras que los datos se encuentran en el anexo 6.

El proceso de computación fue realizado por personal del Instituto de Investigación en Matemáticas Aplicadas y Sistemas (I.I.M.A.S.) de la UNAM ya que la CONAFRUT no tenia los recursos para realizar este tipo de trabajo, por lo que se decidio pedir apoyo ⁽⁵⁾ de tipo computacional, para esto se les proporcionó los cuadros que se querian tener, así como la manera de llenarlos y las herramientas necesarias para ello, tal como las tarjetas perforadas, la codificación utilizada y el estimador propuesto.

(5)Aprovechando que en 1980 se firmo un convenio entre la U.N.A.M. Y CONAFRUT.

Finalmente se procedió a revisar los cuadros que se querían tener al final de la investigación con el objeto de verificar que los resultados no fueron extraños, dada la experiencia previamente adquirida en el campo, la cual permite tener una visión más real del problema.

Algunos de los cuadros obtenidos se presentan comentados en el próximo capítulo,

CAPITULO 3
ANALISIS DE DATOS, INTERPRETACION DE
RESULTADOS Y CONCLUSIONES

Los resultados obtenidos de la investigación y que cumplen con el objetivo principal se pueden resumir en los siguientes 10 cuadros.

Los primeros cinco se refieren a las plagas y enfermedades en donde se midió la intensidad del daño (cuadros con cuatro columnas).

Los restantes se refieren a los daños donde únicamente se midió presencia o ausencia (cuadros con 3 columnas).

Los resultados vienen dados por región (estrato) es decir, el estimador propuesto es por región, y para estimar la intensidad de daño en la zona se forma un promedio ponderado de las estimaciones de los estratos; en otras palabras, para estimar la intensidad de daño en la zona se uso la siguiente expresión:

$$\hat{R}_Z = \frac{\sum_{r=1}^4 \hat{R}_r H_r}{H.}$$

donde: \hat{R}_Z es la estimación de intensidad de daño en la zona

\hat{R}_r es la estimación de intensidad de daño en la región r-ésima.

H. es el total de huertas en la zona, es decir,

$$H. = \sum_1^4 H_r$$

Los cuadros contienen por cada celda, la estimación puntual y los intervalos de confianza al 95% y 99%. El intervalo que esta inmediatamente despues de la estimación puntual es el correspondiente al 95% de confiabilidad. Obviamente cuando el porcentaje de predios afectados es 1 ó 0 entonces no se puede dar la estimación por intervalo para el total de predios afectados (Y), ni para el porcentaje de predios afectados (p), ya que:

- Para obtener el intervalo de confianza para p, tenemos:

$$\hat{p} \pm t \hat{V}(\hat{p})$$

donde:

t es un valor obtenido de tablas correspondiente a la confiabilidad escogida.

$$\& \hat{V}(\hat{p}) = \frac{(N-n) p \cdot q}{N \cdot (n-1)}$$

- Para obtener el intervalo de confianza para Y, tenemos:

$$\hat{Y} \pm t \hat{V}(\hat{Y})$$

donde:

t es un valor obtenido de tablas correspondiente a la confiabilidad escogida.

$$\& \hat{V}(\hat{Y}) = \frac{N(N-n) p \cdot q}{(n-1)}$$

Para las plagas y enfermedades donde se midió la intensidad se tienen los siguientes resultados:

REGION No. 1

ENFERMEDAD	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS	INTENSIDAD PROMEDIO DEL % DE PREDIOS AFECTADOS
ORO	437 (385, 464) (368, 464)	0.9418 (0.8292, 1.00) (0.7936, 1.00)	0.0664 (0.0284, 0.1044) (0.0164, 0.1164)
BARRENADOR	464	1.00	0.0405 (0.0268, 0.0542) (0.0224, 0.0586)
PICUDO	437 (385, 464) (368, 464)	0.9418 (0.8292, 1.00) (0.7936, 1.00)	0.0115 (0.0049, 0.0180) (0.0029, 0.0200)

Se puede observar que en esta región, el barrenador es el que se presenta con mayor frecuencia, 464 huertas afectadas que es el 100% del total, sin embargo la intensidad es del .04 o sea que cae en la escala 1 de daño, lo cual significa que presenta únicamente huellas.

En lo que se refiere a la enfermedad del oro y al picudo, estos se presentan con la misma frecuencia, 437 predios afectados, lo cual es el 94% del total de predios en la región 1; sin embargo, de estos, el oro es el que presenta mayor intensidad, aproximadamente .07; por .01 del picudo ubicándose ambos en la escala número 1.

REGION No. 2

ENFERMEDAD	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS	INTENSIDAD PROMEDIO DEL % DE PREDIOS AFECTADOS
ORO	376	1.00	0.0675 (0.0379,0.0971) (0.0285,0.1064)
BARRENADOR	351 (301, 376) (285, 376)	0.9335 (0.8006,1.00) (0.7586,1.00)	0.0780 (0.0427,0.1133) (0.0315,0.1245)
PICUDO	351 (301, 376) (285, 376)	0.9335 (0.8006,1.00) (0.7586,1.00)	0.0471 (0.0240,0.0702) (0.0169,0.0775)

La enfermedad del oro se presenta en los 376 predios, con una intensidad de aproximadamente .07, es decir, la intensidad de daño cae en la escala 1; mientras que el barrenador y el picudo se presentan en 351 predios o sea en el 93% de el total de esta región, solamente que en el primero con una intensidad de aproximadamente .08 y en el otro con una intensidad de 0.5 aproximadamente, sin embargo los dos se encuentran en la escala 1.

REGION No. 3

ENFERMEDAD	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS	INTENSIDAD PROMEDIO DEL % DE PREDIOS AFECTADOS
ORO	268 (195, 327) (172, 327)	0.8196 (0.5966,1.00) (0.5260,1.00)	0.1281 (0.0123,0.2438) (0.0000,0.2805)
BARRENADOR	327	1.00	0.0864 (0.0547,0.1411) (0.0144,0.1584)
PICUDO	327	1.00	0.0882 (0.0348,0.1416) (0.0179,0.1567)

El Oro se presenta en 268 predios (aproximadamente 82% del total) con una intensidad de .13, lo que quiere decir que se encuentra en la escala 1. En lo que se refiere al barrenador y al picudo, ambos se presentan en 327 predios (el total de predios en la región) con una intensidad de aproximadamente .09 para ambos, es decir, con un daño cuantificado por la escala número 1.

REGION No. 4

ENFERMEDAD	TOTAL DE PREDIOS	PORCENTAJE DE PREDIOS	INTENSIDAD PROMEDIO DEL % DE PREDIOS AFECTADOS
ORO	55	1.00	0.0491 (0.0092,0.0890) (0.0000,0.1016)
BARRENADOR	28 (0, 55) (0, 55)	0.5091 (0.00,1.00) (0.00,1.00)	0.309 (0.000,0.0970) (0.000,0.1179)
PICUDO	55	1.00	0.0074 (0.0000,0.0152) (0.0000,0.0177)

Tanto el oro como el picudo se presentan en el total de los predios (55 en la región), pero mientras que el primero presenta una intensidad de .05 aproximadamente, el otro solamente de .007, sin embargo, ambos se encuentran en la escala 1.

El barrenador se presentó en 28 predios (aproximadamente el 51% del total) con una intensidad de .03, ubicandose tambien en la escala 1.

TODA LA ZONA

ENFERMEDAD	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS	INTENSIDAD PROMEDIO DEL % DE PREDIOS -- AFECTADOS
ORO	1136 (1045,1222) (1017,1222)	0.9296 (0.8554,1.00) (0.8319,1.00)	0.0825 (0.0156,0.1494) (0.0000,0.1705)
BARRENADOR	1170 (1098,1222) (1076,1222)	0.9575 (0.8990,1.00) (0.8804,1.00)	0.0638 (0.0256,0.1019) (0.0136,0.1140)
PICUDO	1170 (1098,1222) (1076,1222)	0.9575 (0.8990,1.00) (0.8804,1.00)	0.0427 (0.0119,0.0734) (0.0022,0.0832)

La enfermedad del oro se presenta en 1136 predios, en otras palabras aproximadamente el 93% de los predios de la zona estan afectados por oro donde la intensidad es de .08 aproximadamente. En términos más amplios podriamos afirmar (con probabilidad de .99 de acertar) que al menos hay 1017 predios afectados, lo cual corresponde al 83% del total de predios, con una intensidad de a lo más de .17; en términos de las escalas establecidas para medir la intensidad, se puede decir que la intensidad del daño en oro cae a lo más en la escala 1. La contribución de las regiones en esta enfermedad es como sigue: Tanto en la región 2 como en la 4, el 100% de los predios estan afectados, siguiendoles en importancia la región 1 con 94% y por ultimo la región 3 con el 82% de predios afectados.

El barrenador y el picudo se presentan ambos en 1170 predios (96% del total) con una intensidad de .06 para el primero y de .04 para el otro. En términos probabilísticos se puede decir con un 99% de confianza que al menos hay 5076 predios afectados, los cuales corresponden al 88% del total de predios, con una intensidad de a lo más de: .11 para el barrenador y de .08 para el picudo, cayendo ambos en la escala número 1.

Para los daños donde unicamente se midio presencia o ausencia se tienen los siguientes resultados:

REGION No. 1

DAÑOS	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS
PUDRICION DE EPIDERMIS	164 (57,271) (24,304)	0.3534 (0.1235,0.5833) (0.0507,0.6560)
PUDRICION CIRCULAR	0	0.0000
MANCHA CAFE	300 (193,407) (160,460)	0.6465 (0.4166,0.8764) (0.3438,0.9491)
COCHINILLA	0	0.0000
CARACOL	409 (337,464) (314,464)	0.8815 (0.7261,1.00) (0.6769,1.00)
RESEQUEDAD	109 (14, 204) (0, 233)	0.2349 (0.0310,0.4388) (0.0000,0.5033)
CARCOMA	437 (385,464) (368,464)	0.9418 (0.8292,1.00) (0.7936,1.00)
DAÑO MECANICO	464	1.00

El daño mecánico es el que se presenta con mayor frecuencia, ya que las 464 huertas de la región están afectadas, siguiéndole en importancia la carcoma con 437 (94% del total), el caracol con 409 (88%), la mancha café con 300 (65%) la pudrición de epidermis con 164 (35%) y por último la resequeidad con 109 (23%), la pudrición circular y la cochinilla no afectaron a ninguna huerta.

REGION No. 2

DAÑOS	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS
PUDRICION DE EPIDERMIS	175 (75, 275) (43, 307)	0.4654 (0.1993, 0.7314) (0.1152, 0.8156)
PUDRICION CIRCULAR	50 (0, 118) (0, 140)	0.1329 (0.0000, 0.3140) (0.0000, 0.3712)
MANCHA CAFE	276 (187,365) (159,376)	0.7340 (0.4983, 0.9697) (0.4238, 1.00)
COCHINILLA	100 (11, 189) (0, 217)	0.2659 (0.0302, 0.5016) (0.0000, 0.5761)
CARACOL	251 (156,345) (127,375)	0.6675 (0.4162, 0.9188) (0.3367, 0.9983)
RESEQUEDAD	226 (128,324) (97,355)	0.6010 (0.3398, 0.8622) (0.2572, 0.9448)
CARCOMA	351 (301,376) (285,376)	0.9335 (0.8006, 1.00) (0.7586, 1.00)
DAÑO MECANICO	251 (156,345) (127,375)	0.6675 (0.4162, 0.9188) (0.3367, 0.9983)

En esta región, los daños se presentaron de la siguiente manera: La carcoma con 351 huertas afectadas (93% del total) ocupa el primer lugar, siguiéndole la mancha café con 276 huertas afectadas (73%), el caracol y el daño mecánico ambos con 251 (67%), la resequeidad con 226 (60%), la pudrición de epidermis con 175 (46%), la cochinilla con 100 (27%) y por último la pudrición circular con 50 huertas afectadas (13%).

REGION No. 3

DAÑOS	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS
PUDRICION DE EPIDERMIS	119 (28, 210) (0, 239)	0.3639 (0.0848,0.6429) (0.0000,0.7312)
PUDRICION CIRCULAR	30 (0, 85) (0, 102)	0.0917 (0.0000,0.2591) (0.0000,0.3120)
MANCHA CAFE	238 (154,322) (127,327)	0.7278 (0.4696,0.9860) (0.3880,1.00)
COCHINILLA	89 (5, 173) (0, 200)	0.2722 (0.0140,0.5304) (0.0000,0.6120)
CARACOL	149 (54, 243) (25, 273)	0.4556 (0.1667,0.7445) (0.0754,0.8358)
RESEQUEDAD	267 (194,327) (170,327)	0.8165 (0.5920,1.00) (0.5210,1.00)
CARCOMA	327	1.00
DAÑO MECANICO	238 (154,322) (127,327)	0.7278 (0.4696,0.9860) (0.3880,1.00)

Aqui la carcoma con las 327 huertas afectadas es la que se presenta con mayor frecuencia, despues la resequedad con 267 (82%), la mancha cafe y el daño mecánico ambos con 238 (73%), el caracol con 149 (46%), la pudrición de epidermis con 119 (36%), la cochinilla con 89(27%) y por último la pudrición circular con unicamente 30 huertas afectadas (9% del total).

REGION No. 4

DAÑOS	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS
PUDRICION DE EPIDERMIS	0	0.0000
PUDRICION CIRCULAR	0	0.0000
MANCHA CAFE	28 (0, 55) (0, 55)	0.5091 (0.0000,1.00) (0.0000,1.00)
COCHINILLA	28 (0, 55) (0, 55)	0.5091 (0.0000,1.00) (0.0000,1.00)
CARACOL	55	1.000
RESEQUEDAD	55	1.00
CARCOMA	55	1.00
DAÑO MECANICO	55	1.00

En esta región se presentaron en las 55 huertas, el caracol - la resequedad, la carcoma y el daño mecánico; la mancha café y la cochinilla en 28 (51%) y los dos tipos de pudrición (circular y de epidermis) no se presentaron en ninguna huerta.

TODA LA ZONA

DAÑOS	TOTAL DE PREDIOS AFECTADOS	PORCENTAJE DE PREDIOS AFECTADOS
PUDRICION DE EPIDERMIS	458 (288,628) (235,681)	0.3747 (0.2343,0.5151) (0.1899,0.5595)
PUDRICION CIRCULAR	80 (0, 168) (0, 195)	0.0654 (0.0000,0.1371) (0.0000,0.1598)
MANCHA CAFE	842 (678,1006) (626,1058)	0.6890 (0.5548,0.8232) (0.5123,0.8657)
COCHINILLA	217 (82, 352) (39, 395)	0.1775 (0.0667,0.2883) (0.0316,0.3233)
CARACOL	864 (703,1025) (652,1076)	0.7070 (0.5750, 0.8390) (0.5333, 0.8807)
RESEQUEDAD	657 (480,834) (424,890)	0.5376 (0.3930,0.6822) (0.3473,0.7279)
CARCOMA	1170 (1098,1222) (1076,1222)	0.9574 (0.8988,1.00) (0.8803,1.00)
DAÑO MECANICO	1008 (873,1143) (831,1185)	0.8248 (0.7146,0.9350) (0.6797,0.9699)

En toda la zona el daño de mayor frecuencia es la carcoma con 1170 huertas afectadas (96%), despues con 1008 huertas afectadas (82%) el daño mecanico; el caracol con 864 (71%) ocupa el tercer lugar en importancia; la mancha cafe con 842 (69%), la resequedad con 657 (54%) y la pudrición de epidermis con 458 (37%) ocupan el cuarto, quinto y sexto lugares respectivamente, y por último la cochinilla con 217 (18%) y la pudrición circular con 80 (6%) ocupan los dos últimos lugares en importancia, de acuerdo a la frecuencia con que se presentan en la zona.

La contribución de cada región para todas las enfermedades se-comento anteriormente región por región.

Para probar las hipótesis planteadas en la investigación, se procedió de la siguiente manera:

La hipótesis 1 se planteó como a continuación se describe:

Hipótesis Estadística: La limpieza de las huertas no tiene relación con la incidencia de plagas y enfermedades.

Es decir,

$$\begin{aligned} \text{Vs } H_0 &: P_1 = P_2 \\ H_1 &: P_1 \neq P_2 \end{aligned}$$

donde P_1 = Porcentaje de predios afectados en las huertas limpias.

P_2 = Porcentaje de predios afectados en las huertas no limpias.

Para probar esta hipótesis, la información relevante se presenta en la siguiente tabla.

CANTIDAD Y PORCENTAJE DE PREDIOS AFECTADOS POR LAS DIVERSAS PLAGAS Y ENFERMEDADES.

Plagas y Enfermedades Fitosanitario	O R O	B A R R E N A D O R	P I C U D O	P. E P I D E R M I S	P. C I R C U L A R	M. C A F E	C O C H I N I L L A	C A R A C O L	R E S E Q U E D A D	C A R C O M A	D. M E C A N I C O
Huertas limpias	29 .90	31 .96	30 .94	9 .28	2 .06	22 .68	6 .18	16 .50	22 .68	30 .94	27 .85
Huertas no limpias	12 .92	12 .92	13 1.0	8 .62	1 .08	9 .69	1 .08	7 .54	9 .69	13 1.0	10 .77
	1	2	3	4	5	6	7	8	9	10	11

Sin embargo el análisis para la prueba de hipótesis se tiene

que hacer por separado para cada enfermedad, ya que en una huerta puede estar más de una enfermedad y por lo tanto las modalidades no son ajenas.

De tal manera que, por ejemplo, para el oro se tendría:

Estado Fitosanitario \ Plagas y enfermedades	O R O	NO O R O
Huertas limpias	29 .90	3 .10
Huertas no limpias	12 .92	1 .08

y cuadros similares para las demas plagas o enfermedades .

Para la prueba de hipotesis, se utilizó una prueba χ^2 , obteniéndose los siguientes resultados, con $\alpha = .05$

- Para el oro:

$$\chi^2_c = 0.032$$

$$\chi^2_t(.95) = 3.841$$

Como $\chi^2_c \neq \chi^2_t(.95)$ Entonces no rechazo Ho

- Para el barrenador:

$$\chi^2_c = 0.454 \quad \text{Entonces no rechazo Ho}$$

-Para el picudo:

$$\chi^2_c = 0.850 \quad \text{Entonces no rechazo Ho}$$

-Para la putrición de epidermis:

$$\chi^2_c = 4.39$$

$$\chi^2(.95) = 3.841 \quad ; \quad \chi^2(.99) = 6.63$$

con $\alpha = .05$ Rechazo Ho

$\alpha = .01$ No rechazo Ho

Haciendo la corrección por finitud

$$\chi^2_c = 2.27$$

con $\alpha = 0.5$ No rechazo Ho

$\alpha = 0.1$ No rechazo Ho

-Para la putrición circular:

$$\chi^2_c = 0.031$$

Entonces no rechazo Ho

-Para la Mancha cafe:

$$\chi^2_c = .000997$$

Entonces no rechazo Ho

-Para la cochinilla:

$$\chi^2_c = 0.860$$

Entonces no rechazo Ho

-Para el caracol:

$$\chi^2_c = .000997$$

Entonces no rechazo Ho

-Para la resequedad:

$$\chi^2_c = 0.05473$$

Entonces no rechazo Ho

- Para la carcoma:

$$\chi^2_c = 0.85029$$

Entonces no rechazo Ho

-Para el daño mecanico:

$$\chi^2_c = .35120$$

Entonces no rechazo Ho

Concluyendo, podemos decir que el porcentaje de predios afectados (de todas las plagas y enfermedades donde se realizo la medición) en las huertas limpias, es el mismo que el porcentaje de predios afectados en las huertas no limpias.

-La hipótesis 2, se planteo de la siguiente manera:

Hipotesis Estadística: La presencia de una enfermedad o plaga en la planta no disminuye la producción.

En otras palabras, se puede decir que:

La producción es la misma en las huertas sin incidencia que en las huertas con incidencia*.

Es decir

$$\begin{array}{l} H_0: \mu_1 = \mu_2 \\ \text{Vs} \\ H_1: \mu_1 \neq \mu_2 \end{array}$$

donde:

μ_1 = promedio de producción en las huertas sin incidencia
 μ_2 = promedio de producción en las huertas con incidencia.

*Únicamente se toma la producción de las 10 plantas muestra para cada huerta.

NOTA: Se comparo las huertas que no tenian oro contra las que si tenian oro, las que tenian barrenador contra las que no tenian y las que tenian picudo contra las que no tenian picudo. Ya que no hubo ningún caso en donde no se presentaran las 3 enfermedades por lo que se decidió hacer la comparación por separado para cada enfermedad.

Para probar la hipotesis, se utilizo la prueba t, obteniendose los siguientes resultados con $\alpha=.05$ y $\alpha=.10$.

- Para el oro

$$t_c = -.7865$$

$$t^{40} .975 = 2.021$$

$$t^{40} .95 = 1.684$$

Como $t_c \neq t$ tablas Entonces no rechazo H_0 en ninguno de los dos niveles de

-Para el barrenador

$$t_c = -1.7749$$

Entonces no rechazo H_0 al nivel del 5%

pero

Rechazo H_0 al nivel del 10%

-Para el picudo:

$$t_c = 1.5831$$

entonces

No rechazo H_0 al nivel del 5%

No rechazo H_0 al nivel del 10%

Concluyendo, el promedio de producción en las huertas con incidencia es el mismo que en las huertas sin incidencia con un nivel $\alpha = .05$ (para las 3 enfermedades).

Para el barrenador, con $\alpha = .10$ se concluye que el promedio de producción en las huertas con incidencia es diferente (mayor) al promedio de producción en las huertas sin incidencia.

ANEXO 1

CUESTIONARIO UTILIZADO EN LA RECOLECCION DE LA INFORMACION

CUESTIONARIO PARA LA ESTIMACION DE ALGUNAS ENFERMEDADES EN EL NOPAL EN -
LA ZONA DE SAN MARTIN DE LAS PIRAMIDES.

1980. (Mayo-Septiembre)

La información que nos proporcione servirá de base para la elaboración de investigaciones en la zona.

Se garantiza el caracter confidencial de la información proporcionada.

A: ENTREVISTA

PARTE I: DATOS GENERALES

1.- Nombre de la huerta: -----

Región (R₁) (R₂) (R₃) (R₄)

Poblado: (P₁) (P₂) (P₃) (P₄) (P₅)(P₆)

No. de huerta:

2.- Tipo de propiedad:

Ejidal	P.P. no asociados	Asoc. P.P.	Soc. Prod. Tuna
(1)	(2)	(3)	(4)

PARTE II: INFORMACION DE LAS PLANTAS DE TUNA BLANCA

3.- Has. de la huerta:

Totales:

De tuna blanca:

11.- ¿Cuándo fertilizó? ¿Antes o después de la cosecha?

A D
(1) 0 (2)
(3)

A D
(1) 0 (2)
(3)

A D
(1) 0 (2)
(3)

Especifique

Especifique

Especifique

12.- ¿Ha sido atacada la plantación por alguna enfermedad, plaga o siniestro?

Si
(1)
continúe

No
(2)
pase a 17

13.- ¿Cuál? _____

--	--

--	--

--	--

Continúe con la hilera de 14 y 15 respectivamente, después siga la columna de 16. Haga lo mismo para la 2a. y 3a. columna.

14.- ¿En cuántas has se presentó esto? ¿cuando?

HAS

MES

AÑO

			•		
			•		
			•		

1	9		
1	9		
1	9		

15.- ¿Que hizo para combatir esa enfermedad, plaga o siniestro?

usar insecticida
¿cual?

usar fungicida
¿cual?

otra medida
especifique

nada

16.- ¿Cuanto aplico por ha? (ponga unidades)

--	--	--	--	--	--	--	--	--	--

--	--	--	--	--	--	--	--	--	--

--	--	--	--	--	--	--	--	--	--

17.- Nombre del propietario: _____

Nombre del entrevistado: _____

Fecha de entrevista:

1	9	8	0			
AÑO		MEJ				DIA

18.- OBSERVACIONES GENERALES

- (1) _____
- (2) _____
- (3) _____
- (4) _____
- (5) _____

Planta	Cladodio	0	Enf. del oro				Mancha negra				Barrenador				Picudo				N				
			0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3					
9	1																						
	2																						
	3																						
	4																						
	5																						
	6																						
	7																						
	8																						
	9																						
	10																						
	11																						
	12																						
10	1																						
	2																						
	3																						
	4																						
	5																						
	6																						
	7																						
	8																						
	9																						
	10																						
	11																						
	12																						

20.- Otras enfermedades todavia no diagnosticadas. Si no hay, vaya a 22.

Anotar la Sintomatologia

E1 _____

E2 _____

sintomatologia

sobrenombre

21.- Ubicación de la enfermedad en el cladodio

E1
(1) (2) (3)

E2
(1) (2) (3)

ES BO PP

ES BO PP

PARTE V : INFORMACION DE PRODUCCION.

31.- En esta parte recabar la información siguiente:

Región: (R₁) (R₂) (R₃) (R₄)

N° de huerta:

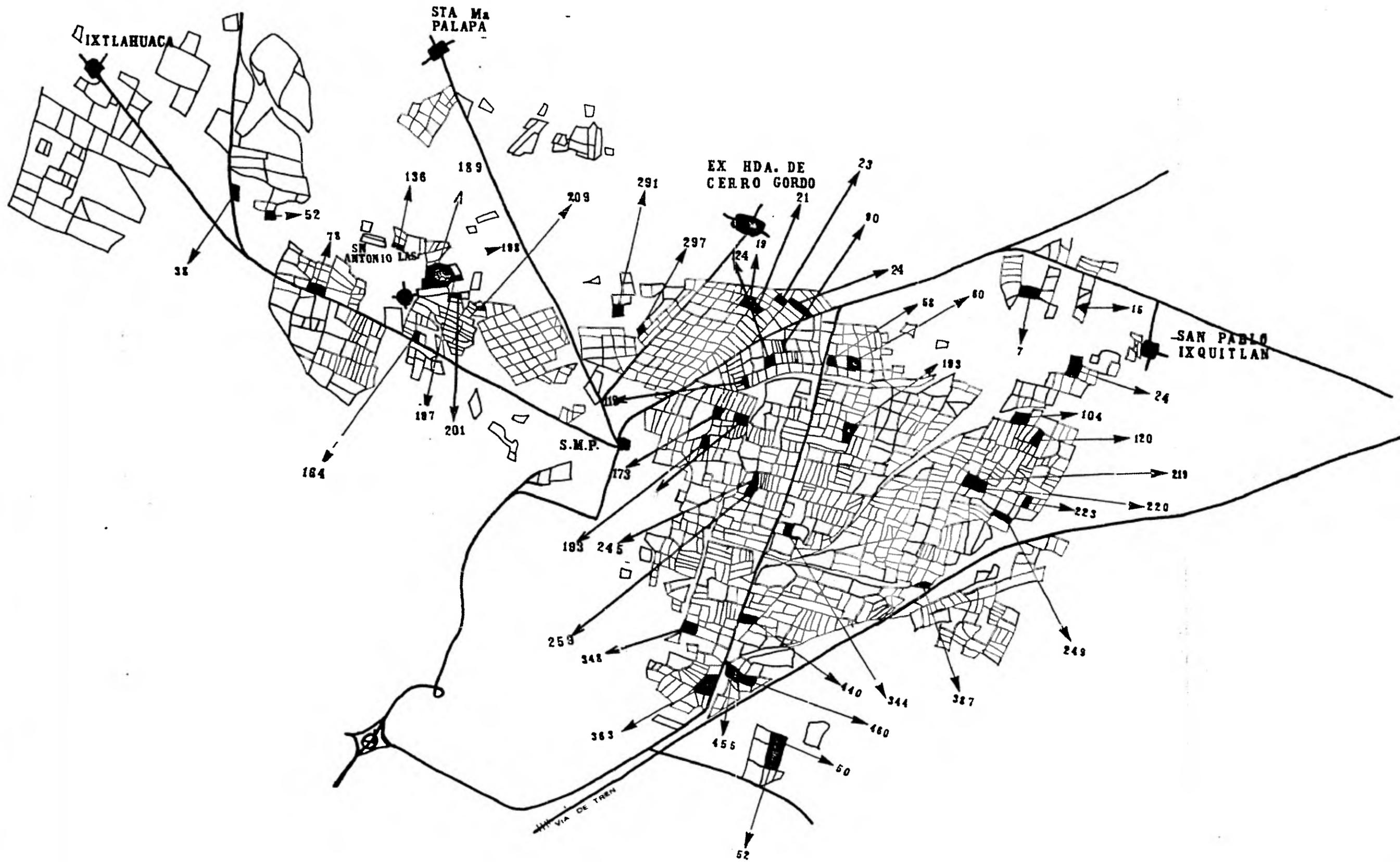
N° de plantas:

PLANTAS MUESTRA

CLADODIOS MUESTRA

N°	DE	CLADODIOS		N°	DE	TUNAS		N°	TUNAS	EIO
Y	CON	TUNA								
P1			C1							
			C2							
			C3							
			C4							
P2			C1							
			C2							
			C3							
			C4							
P3			C1							
			C2							
			C3							
			C4							
P4			C1							
			C2							
			C3							
			C4							
P5			C1							
			C2							
			C3							
			C4							
P6			C1							
			C2							
			C3							
			C4							
P7			C1							
			C2							
			C3							
			C4							
P8			C1							
			C2							
			C3							
			C4							
P9			C1							
			C2							
			C3							
			C4							
P10			C1							
			C2							
			C3							
			C4							

ANEXO 2
MARCO DE MUESTREO



ANEXO 3

INSTRUCTIVO PARA LA MEDICION DIRECTA

Recuerde que el cuestionario utilizado contiene dos partes principales, según la procedencia de la información:

- a) La parte correspondiente a la entrevista
- b) La parte correspondiente a la medición directa

Tanto en una como en otra, es necesario proveer de instructivos que permitan al personal de campo realizar adecuadamente su trabajo ya que la calidad de una encuesta depende de la preparación del personal de campo asignado para la recolección de la información.

Para la parte correspondiente a la entrevista el instructivo elaborado se reproduce en el Anexo 4, para la parte correspondiente a la medición directa, el instructivo que se elaboró para tal fin se reproduce a continuación:

"El trabajo de selección de plantas y cladodios en una huerta se hará de la siguiente manera":

I.- Al llegar a una huerta, se procederá a contar la cantidad de plantas - que tiene de largo y las que tiene de ancho, con el fin de poder saber (de manera aproximada), el número total de plantas que existen en dicha huerta; además en caso necesario ⁽¹⁾ se deberá medir la huerta -- (cuantos metros de ancho y cuantos de largo). En ambos casos, es conveniente poner el mapa de la huerta al reverso de la primera hoja del cuestionario. Ya que se contabilizó el número de filas a lo largo y - ancho de la huerta se considerará:

- 1º Cuantas filas hay a lo ancho de una huerta y seleccionar de una manera sistemática 3 de ellas.
- 2º En dos de las filas seleccionadas se hará la selección de 3 plantas y en la fila restante se seleccionaran 4 plantas (en ambos casos, - con muestreo sistemático).

(1) Esto será cuando no sea satisfactorio en base al número de filas, saber el número de plantas.

Para saber en cual de las filas se toman cuatro plantas se realizará un sorteo entre las tres filas y a la fila seleccionada se le tomaran cuatro plantas como muestra, teniendo de esta manera un total de 10 plantas muestra.

De manera explicita, la selección sistemática de filas se hará de la siguiente manera:

En una huerta seleccionada, denotemos con N el número total de filas en el ancho de la huerta, y n las filas que vamos a seleccionar (en este caso tres), entonces se procede como sigue:

a) Dividir:

$$\frac{N}{n} = k$$

Si resulta que k es fraccionario, se toma k como el entero más próximo a la razón $\frac{N}{n}$.

En caso de que se tenga duda en cual entero elegir (por ejemplo, si $k=3.5$), se toma k como el entero mayor más proximo a la razón N/n (en el caso de que $k=3.5$, tomar $k=4$).

b) Obtener un número aleatorio entre 1 y k (a este número se le denomina el arranque aleatorio), y al número obtenido, agregar los intervalos k hasta seleccionar exactamente 3 filas (para esto hay que considerar la lista de filas como si fuera circular, de manera que la última fila sea seguida por la primera fila).

Como ilustración del método, supongamos que el número obtenido es j, entonces nuestra muestra constituida por 3 filas (en el ancho de la huerta), va ser la siguiente:

$$j, j + k, \text{ y } j + 2k$$

c) De las filas j , $j + k$, y $j + 2k$ seleccionar aleatoriamente una para decidir cual de ellas va a contener 4 plantas. Por otro lado, para la selección de plantas en las filas, se procede en cada fila de las 3 ya seleccionadas de una manera análoga a lo anteriormente mencionado, donde en este caso, N va a denotar el número total de plantas en una fila seleccionada, y n va a denotar las plantas que vamos a seleccionar en dicha fila (que pueden ser 3 ó 4).

II.- Después de haber seleccionado las plantas procedemos como sigue:

1º. Se seleccionan 12 cladodios por planta (para recabar la información de la parte III del cuestionario) de la manera siguiente:

Párese en el pasillo cuya dirección es el largo de la huerta y se leccione 6 cladodios. El primer cladodio seleccionado va a ser aquel que quede a la altura de la cabeza, los siguientes dos van a ser aquellos que queden a la altura de los hombros (teniendo los brazos extendidos de manera horizontal). Para seleccionar los otros 3 se seleccionará primero aquel que quede a la altura del pecho y los otros dos, se seleccionan de aquellos que queden a la altura de las caderas (suponiendo también que se tienen los brazos extendidos en forma horizontal (ver Fig. 1). Todo esto es de manera aproximada con el fin de poder esquematizar la situación.



Fig. No. 1: SELECCION DE CLADODIOS EN UNA PLANTA PARA EVALUAR ENFERMEDADES.

Para seleccionar los siguientes 6, trasladarse a la parte opuesta de la planta y proceder de la manera descrita anteriormente.

- 2º Para recabar la información de la parte V del cuestionario, se procede a contar cuantos cladodios hay en la planta y cuantos de estos tienen tuna blanca; después se seleccionan 4 de aquellos cladodios que tengan tuna blanca y se cuentan cuantas tunas tiene cada uno de ellos (haciendo la anotación correspondiente en la forma para información de producción).

La manera de seleccionar los cladodios con tuna es la siguiente:

Una vez considerada una planta seleccionada, pase en el pasillo cuya dirección es el largo de la huerta y seleccione un cladodio con tuna a la altura del pecho, el siguiente se seleccionará a la altura de la cintura; para seleccionar los otros dos solo hay que trasladarse al lado opuesto de la planta y proceder de la manera descrita anteriormente (vease Fig. No. 2).



Fig. No. 2: SELECCION DE CLADODIOS EN UNA PLANTA PARA EL CONTEO DE TUNAS,

El procedimiento de selección va a ser el mismo para todas las -- plantas muestra de la huerta, y se procederá de manera análoga en todas las huertas seleccionadas para cada una de las regiones.

III.- Por último, al término del trabajo en una huerta,

Anotar:

- De quien es la huerta
- Nombre de la huerta
- Donde vive el dueño: Congregación, Ranchería, etc., es de cir, hay que obtener una pista del dueño.
- Asociación a la que pertenece
- Referencias y ubicación de la huerta, con la finalidad de que el dueño a la hora de la entrevista identifique su -- huerta en el mapa.

Todo lo referente al punto III, ponerlo al reverso de la se gunda hoja del cuestionario.

ANEXO 4
MANUAL DE REFERENCIA DEL ENTREVISTADOR

INTRODUCCION:

Mediante este manual se da una guía al entrevistador para que maneje el cuestionario en forma adecuada; se cubren las posibles situaciones de ambigüedad que se pudieran presentar a lo largo de la ejecución de la entrevista.

El cuestionario consta de 5 partes:

- Parte I: Datos generales
- Parte II: Datos de las plantas de tuna blanca proporcionados por el productor.
- Parte III: Cuantificación de enfermedades y daños por insectos.
- Parte IV: Indicadores de enfermedades
- Parte V: Información de producción

Las partes I y II corresponden a información que se obtendrá por entrevista directa con el dueño de la huerta o bien el encargado de ella (de preferencia este último). En cuanto a las partes III, IV y V; estas se refieren a la medición directa que nosotros debemos realizar.

De la anterior se observa que:

- a) La calidad de la encuesta en las dos primeras partes (ENTREVISTA) dependerá de:
 - i) Que la entrevista se realice con personas que esten relacionados directamente con el cuidado de la huerta; -- así como también de que entrevistador informe suficientemente sobre el tema de investigación a dichas personas.
 - ii) Que al aplicar el cuestionario, el entrevistador sea capaz de crear una atmósfera cordial, y hacer que la persona entrevistada se sienta cómoda garantizandole el carácter confidencial de la información proporcionada).
 - iii) Que el entrevistador tenga tacto para desalentar la conversación superflua por un lado, y por otro debe saber -

provocar respuestas de quienes aparentemente no esten dispuestos a cooperar (argumentando que la enfermedad o daño, merma la producción y calidad de su producto).

- iv) El planteamiento de las preguntas. Estas deben realizarse del mismo modo como estan redactadas y de esa manera asegurar que los entrevistados contesten exactamente la misma pregunta, lo cual hace lo que los resultados sean comparables; también debe respetar el orden de las preguntas así como evitar inducir al entrevistado a constestar cosas de las que no este plenamente seguro.
- b) En tanto que las partes III, IV y V (MEDICION DIRECTA) dependeran de:
 - i) La concepción que se tenga acerca de la importancia del -- trabajo, teniendo claramente definida la necesidad de la - investigación, así como la forma en que se emplearán los - resultados.
 - ii) De los elementos de información profesional que se tengan para no tener ambigüedad en la identificación de los diferentes daños, labores culturales y otros, es decir que --- cuando se mida se sea objetivo, de tal manera que se des-- criba unicamente lo que se percibe y no lo que se infiere, en otras palabras, hay que evitar juicios dictados por el- deseo.

Es importante notar que, de no ser posible tener la información de las partes I y II, las otras servirán para cumplir los objetivos propuestos, además de que de estas últimas dependen la calidad de la investigación.

Dado que el cuestionario está diseñado para ser procesado por una máquina computadora, teniendo como base el orden de las preguntas consideradas con una secuencia lógica, hay que tener presente que no debe de quedar ninguna pregunta sin constestar, de acuerdo a las instrucciones de "Pase" o "Continue".

Es de vital importancia que los cuestionarios sean llenados correctamente, ya que evitará pérdidas de tiempo en la revisión y podremos tener la información al tiempo planeado.

A continuación se da a conocer la forma correcta de llenar el cuestionario y las posibles confusiones que pueda haber.

DATOS DE IDENTIFICACION:

Esta parte del cuestionario corresponde a las preguntas 1 y 2. La correspondiente a región y número de huerta podrá ser asignada desde el momento en que se seleccionen las huertas.

En cuanto al poblado, será necesario consultar la lista de claves en el momento de la entrevista ⁽¹⁾.

TIPO DE PREGUNTAS:

En general, tendremos tres tipos de preguntas dependiendo de la codificación; estas pueden ser codificación no variable, codificación variable y no codificadas.

Las preguntas que no llevan cuadros, son las denominadas no codificadas, por ejemplo, el nombre del entrevistado.

La de codificación variable son aquellas que al principio de la toma de información no tienen asignada todavía una identificación (un número), sino que está se tendrá después de tener la información en algunos cuestionarios, con los cuales, de acuerdo a la mayor frecuencia de las modalidades de cada una de las variables se les dará una identificación y -- entonces se codificará; ejemplo de estas variables son las preguntas: 9, - 13, 15, 20 y 23.

LINEAMIENTOS GENERALES DE USO DEL CUESTIONARIO:

- a) Los cuestionarios deberán llenarse únicamente con lápiz.
- b) Encierre en un círculo la respuesta dada por el entrevista-

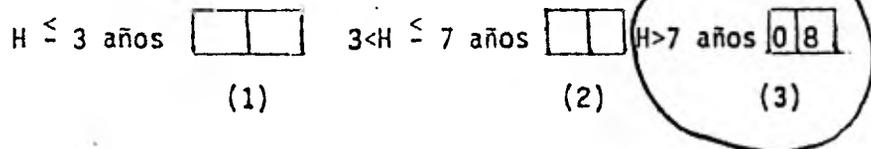
(1) La lista de claves se adjunta al final del manual de referen-
cia del entrevistador.

do cuando la pregunta sea cerrada, es decir, de opción múltiple.

Ejemplo:

4.- ¿Qué edad tienen sus plantas de tuna blanca?

$H \leq 3$ años (1) $3 < H \leq 7$ años (2) $H > 7$ años (3)



Supongamos que el respondiente mencionó que su huerta era de 8 hectáreas y que 6 de ellas tenían 8.5 años y las restantes 2 años, entonces para este caso específico debe registrarse aquella edad de la huerta - que tenga mayor extensión únicamente ⁽²⁾; debido a que la edad de mayor extensión es de 8.5 años, entonces se encerro el intervalo correspondiente.- Además, como puede observarse en los dos campos a la derecha se llenó con un 0 8, dado que se acordó que era conveniente registrar solo los años cumplidos. Por otro lado, este ejemplo también ilustra que las preguntas -- que tengan campos (cuadros) asignados cuando no haya punto, la información se cargara a la derecha, en el caso de que en los campos aparezca un punto, este debiera respetarse.

Nota: En caso de que queden campos en blanco, por la naturaleza de la información, estos llenarlos con ceros, como en el ejemplo anterior.

En la misma pregunta, si la huerta es menor o igual que 3 años entonces, continuar hasta la pregunta 6, luego pasar a 17, donde termina la información para esa huerta.

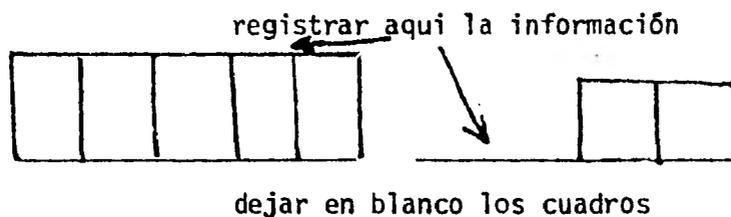
c) En la pregunta 7, insistir cuando la respuesta sea no porque quizá el respondiente no entienda que sea fertilizar.

(2) Dado que en la pregunta No. 5 se establecen las diferentes edades de alguna manera.

- d) En la pregunta 8, verificar que las hectáreas fertilizadas no sean mayores que las hectáreas de tuna blanca, en caso de que así sea, hay que insistir y decir que la información que queremos solo es de tuna blanca.
- e) En la pregunta 9, en caso de que sea más de un fertilizante el que se utilizó, proceder de la siguiente manera:
 - i) Que enuncie primero uno y obtener la información de las preguntas 10 y 11 y registrarlos en la misma columna en que se anota el fertilizante utilizado.
 - ii) Proceder de la misma manera para los otros dos fertilizantes.
- f) En las preguntas 9, 10, 13, 15, 16, 20 y 23 en el momento de llenar por primera vez la forma, registrar la información sobre la línea continua y los cuadros dejarlos en blanco.

Ejemplo:

Se tiene las siguientes formas:



- g) En la pregunta 13, proceder de la siguiente manera:
 - i) Anotar primero una y anotar a continuación la información de las preguntas 14 y 15 en el primer renglón y luego la primera columna de la 16.
 - ii) Proceder de la misma manera para las otras dos.
- h) En la pregunta 15, en la tercer columna (otra medida) se va a poner la información cuando la persona no sepa si es insecticida o fungicida y despues de identificarlo hacer la -

asignación correspondiente, es decir, llevarlo a fungicida, o bien dejarlo ahí.

Se recomienda que cuando este sea el caso, poner el momento de hacer la entrevista un asterisco. En caso de que no lleve asterisco, --- quiere decir que se tiene la certeza que no es ni insecticida ni fungicida.

i) En la pregunta 12, considerar como siniestros cuando ocurren heladas, granizadas, incendios, temperaturas abajo de cero, unicamente durante este año. Es conveniente -- que al entrevistado se le comunique los siniestros que se consideran.

j) En cuanto a observaciones generales es conveniente encausar estas por tema, es decir, en cada hilera anotar un tópico específico.

- (1) _____
- (2) _____
- (3) _____
- (4) _____
- (5) _____

Digamos que (1) anotaciones para aclarar respuestas incompletas (3), confusas ó bien, fallas del cuestionario en (2) aspectos de fungicidas en (3) aspectos de limpieza etc.

Lo anterior tiene como finalidad ver si de acuerdo a la importancia y frecuencia de la información esta pudiera codificarse posteriormente y analizarla de manera formal.

Recomendaciones Finales.

1º Haga anotaciones en la parte observaciones, de aquellos datos que ayuden a aclarar las respuestas incompletas o confusas.

(3) En la pregunta 1 por ejemplo si no menciona alguno de los poblados, aclarar la cercanía a alguno de ellos.

- 2º Dé las gracias por las atenciones prestadas, lo más correcto posible.
- 3º Escriba con corrección y claridad, las palabras y los números.
- 4º Como los errores cometidos en los cuestionarios, son ocasionados principalmente por el poco dominio, al no estudiar, - principalmente, y a la poca importancia que le puedan dar a la entrevista, se les pide que estudien ⁽⁴⁾ bien para que - no cometan errores.
- 5º Lleve el material necesario para trabajar en el campo.
- 6º Cuando tenga confusión consulte la hoja de claves. Anexa.

(4) Además lea instructivo correspondiente a la parte de medición directa.

LISTA DE CLAVES

- Las claves asignadas a los poblados son los siguientes:

- P1 = Poblado # 1 = San Martín de las Pirámides
- P2 = Poblado # 2 = San Pablo Izquitlán
- P3 = Poblado # 3 = Cerro Gordo
- P4 = Poblado # 4 = San Antonio las Palmas
- P5 = Poblado # 5 = Ixtlahuaca
- P6 = Poblado # 6 = Sta. Ma. Palapa

- Para medir la intensidad de las enfermedades se propuso una escala del 0 - 3, donde:

- 0 - Significa que el cladodio no tiene daño
- 1 - Significa que se tiene un daño de un máximo del 25%
- 2 - Significa que el daño tiene un rango de variación del 25% - 50%.
- 3 - Significa que el daño que se presenta en el cladodio es más del 50%.

- Las abreviaturas utilizadas son:

- E.S. = Espina
- B.O. = Borde
- P.P. = Parte plana
- O. = Otros
- N = No tiene ninguna enfermedad, es decir, el cladodio examinado no tiene
 - Enfermedad del oro
 - Mancha negra
 - Barrenador
 - Picudo
 - Ninguna de las contenidas en el rubro "otros".

ANEXO 5

CODIFICACION UTILIZADA PARA PROCESAR LA INFORMACION

Codificación de las partes III y V de cuestionario (20 tarjetas por cuestionario, es decir, 2 tarjetas por planta).

Aclaración; la siguiente codificación solamente es para una planta; la codificación de las otras se hace de manera análoga a la primera.

Ia. Tarjeta

Parte V:

<u>Variable</u>	<u>Columnas</u>	<u>Valores</u>
Región	1-1	1-4
Blanco	2-3	-
No. de huerta	4-6	1-999
No. de planta	7-8	1-10
No. de tarjeta	9-10	1-20
Blanco	11-11	
No. total de cladodios	12-14	0-999
No. de cladodios con tuna	15-17	0-999
No. de tunas del cladodio 1	18-19	0-99
No. de tunas del cladodio 2	20-21	0-99
No. de tunas del cladodio 3	22-23	0-99
No. de tunas del cladodio 4	24-25	0-99

Parte III:

Enfermedades del cladodio No. 1	26-33	Vease nota 1
Enfermedades del cladodio No. 2	34-41	Vease nota 1
Enfermedades del cladodio No. 3	42-49	Vease nota 1
Enfermedades del cladodio No. 4	50-57	Vease nota 1
Enfermedades del cladodio No. 5	58-65	Vease nota 1
Enfermedades del cladodio no. 6	66-73	Vease nota 1

2a. Tarjeta

Parte V:

<u>Variable</u>	<u>Columnas</u>	<u>Valores</u>
Región	1-1	1-4
Blanco	2-3	-
No. de huerta	4-6	1-999
No. de planta	7-8	1-10
No. de tarjeta	9-10	1-20
Blanco	11-25	-

Parte III:

Enfermedades del cladodio No. 7	26-33	Vease nota 1
Enfermedades del cladodio No. 8	34-41	Vease nota 1
Enfermedades del cladodio No. 9	42-49	Vease nota 1
Enfermedades del cladodio No. 10	50-57	Vease nota 1
Enfermedades del cladodio No. 11	58-65	Vease nota 1
Enfermedades del cladodio No. 12	66-73	Vease nota 1

Nota 1: Los 8 campos que se le asignaron a las enfermedades, están distribuidos de la siguiente manera:

Los primeros 4 es para la codificación de OTROS y son:

- | | | |
|--|---|-------------------------|
| 1.- Corresponde a PUDRICION DE EPIDERMIS | } | ENFERMEDADES |
| 2.- Corresponde a PUDRICION CIRCULAR | | |
| 3.- Corresponde a MANCHA CAFE | | |
| 4.- Corresponde a COCHINILLA (PERA) | } | INSECTOS |
| 5.- Corresponde a CARACOL | | |
| 6.- Corresponde a RESEQUEDAD (TUZA) | } | MAMIFEROS
Y
OTROS |
| 7.- Corresponde a CARCOMA (DEFICIENCIA) | | |
| 8.- Corresponde a DAÑO MECANICO | | |

El campo No. 5 es para la codificación de la enfermedad del oro, y se codifica de acuerdo a la intensidad de daño que tenga.

El campo No. 6 es para la codificación del barrenador y se codifica de acuerdo a la intensidad de daño que tenga.

El campo No. 7 es para la codificación del picudo y se codifica de acuerdo a la intensidad de daño que tenga

El campo No. 8 es para la codificación de "No hay" y se codifica con el No. Cero.

Graficamente tenemos:

Enfermedades	Insectos	Mamíferos	Oro	Barrenador	Picudo	No hay	
1	2	3	4	5	6	7	8

El No. total de plantas de las huertas se codifica despues de la enfermedad del cladodio 12 de la planta No. 10, en las columnas 75-78.

Codificación de las partes I, II y IV del cuestionario (1 tarjeta por huerta)

Parte IV:

<u>Variable</u>	<u>Columnas</u>	<u>Valores</u>
Material de descho	1-2	1-4
Sintomas del material	3-8	Vease nota 2
Destino del material	9-9	1-3
Predio limpio	10-10	1-2
Situación del predio	11-11	1-4
Tipo de suelo	12-12	1-4
Ubicación del predio	13-14	1-4
Fecha	15-18	Mes y año

Nota 2: Los dos primeros campos se refieren a enfermedad, los segundos a insectos y mamíferos y los últimos a no se observe, y se codifican de -- acuerdo a las claves dadas en la nota 1, además de las siguientes:

- 9.- Corresponde a ENFERMEDAD DEL ORO
- 10.- Corresponde a BARRENADOR
- 11.- Corresponde a PICUDO
- 12.- Corresponde a "NO SE OBSERVA"

Partes I y II:

<u>Variable</u>	<u>Columnas</u>	<u>Valores</u>
Región	21-21	1-4
No. de huerta	22-24	0-999
Tipo de propiedad	25-25	1-4
Total de Hectáreas	26-29	0-9.99
Hectáreas de tuna blanca	30-33	0-9.99
Edad de plantas	34-35	0-99
Las sembro el mismo año	36-36	1-si, 2-no
Fertiliza	37-37	1-si, 2-no
Fertilizante (nombre)	38-38	Vease nota 3
Ha sido atacada	39-39	1-si, 2-no
Cual (problema 1)	40-41	Vease nota 4
Que hizo (combate 1)	42-42	Vease nota 3
Cual (problema 2)	43-44	Vease nota 4
Que hizo (combate 2)	45-45	Vease nota 3
Cual (problema 3)	46-47	Vease nota 4
Que hizo (combate 3)	48-48	Vease nota 3

Nota 3: La codificación correspondiente a Fertilizantes, Insecticidas y Fungicidas es la siguiente:

- Abono organico le corresponde el No. 1
- Abono químico le corresponde el No. 2
- Foley le corresponde el No. 4
- Folidoy le corresponde el No. 5
- Partil le corresponde el No. 6
- "Otra medida" le corresponde el No. 7
- "Nada" le corresponde el No. 8

Nota 4: La codificación correspondiente a Plagas, Enfermedades y Sinietros es:

- 1.- Corresponde a PUDRICION DE EPIDERMIS
- 2.- Corresponde a PUDRICION CIRCULAR
- 3.- Corresponde a MANCHA CAFE
- 4.- Corresponde a COCHINILLA (PERA)
- 5.- Corresponde a CARACOL
- 6.- Corresponde a RESEQUEDAD (TUZA)
- 7.- Corresponde a CARCOMA
- 8.- Corresponde a DAÑO MECANICO
- 9.- Corresponde a ENFERMEDAD DEL ORO
- 10.- Corresponde a BARRENADOR
- 11.- Corresponde a PICUDO
- 12.- Corresponde a "NO SE OBSERVA"
- 13.- Corresponde a GRANIZO

ANEXO 6
DATOS ORIGINALES

A P E N D I C E S

A. MUESTREO POR CONGLOMERADOS

Un procedimiento de muestreo, presupone dividir la población en un número finito de unidades distintas e identificables, llamadas - unidades de muestreo. Las unidades más pequeñas en que puede dividirse una población se llaman elementos de la población, y grupo de tales elementos son los llamados conglomerados. Cuando la unidad de muestreo es un conglomerado, el procedimiento de muestreo se designa como muestreo de conglomerados.

Para varios tipos de poblaciones, no siempre es factible obtener una lista de todos sus elementos ya que no existe ninguna lista utilizable (denominada marco) para ser enumerada, dentro de la cual seleccionar la muestra; incluso, suponiendo que existiera no sería económico basar la encuesta en una muestra simple aleatoria de tales unidades; por lo tanto, no resulta posible usar el elemento como unidad de muestreo. El método de conglomerados se usa de tales casos.

Es muy frecuente que los conglomerados esten definidos como - "areas" o partes bien delimitadas de terreno, de modo que todas las unidades últimas correspondientes al área sean las que constituyen, el conglomerado; de aquí que este generalizada la denominación del muestreo por -- áreas para designar estos procedimientos de muestreo.

El empleo de conglomerados o áreas como unidades de muestreo se justifica por razones de economía (costo, tiempo y recursos) y en ciertos casos por la disminución de sesgos al facilitarse la supervisión; además hay que hacer notar que la concentración de unidades disminuye la necesidad de desplazamiento.

El número de elementos de un conglomerado se llama tamaño de conglomerado. Los conglomerados son en la mayoría de las poblaciones de tamaño desigual. En general, mientras más pequeño sea el tamaño del conglomerado, más exacta será la estimación de la característica de la población para un número dado de elementos en la muestra por lo tanto, para un número determinado de elementos, es lógico esperar que se obtenga una ma-

yor precisión distribuyendolos en un gran número de conglomerados que tomando un pequeño número de conglomerados y muestreando un gran número de elementos de cada uno de ellos o enumerandolos completamente.

Cuando se compara una muestra de conglomerados con una muestra de elementos del mismo tamaño n , podemos esperar una varianza mayor pero un costo menor en la muestra de conglomerados. En general, la mayor distribución de una muestra de elementos en la población produce mayor precisión pero es más costosa.

El procedimiento de seleccionar primero conglomerados y luego escoger un número especificado de elementos de cada conglomerado seleccionado (utilizando también métodos probabilísticos) se conoce como submuestreo.

En el muestreo por conglomerados solo necesitamos disponer de la lista de los conglomerados, en el caso de que haya que afectar una segunda selección de unidades últimas o submuestreo, se necesita formar la lista de unidades componentes de cada conglomerado, pero ello resulta más económico que confeccionar la lista de todas las unidades en la población completa.

En el submuestreo pueden considerarse conglomerados de conglomerados o áreas de áreas, efectuandose una sucesión de submuestreo en varias etapas. Los conglomerados que resulten de la primera división se llaman unidades primarias de muestreo (U.P.M.) o unidades de primera etapa, los de la segunda, secundarias (U.S.M.) y así sucesivamente. A la muestra obtenida de la U.P.M. seleccionadas se les llama selecciones primarias (SP). Las selecciones primarias también se llaman conglomerados finales. El término conglomerado final se utiliza para denotar el total de unidades incluidos en la muestra, de una unidad primaria (Hansen, 1953, pág. 242).

En resumen:

Cuando la población contiene muchas unidades más o menos dispersas, puede haber dos razones fundamentales que impidan la toma de la mues-

tra directamente de la población.

- 1a. Que no se disponga de un marco para las unidades de la población y sea muy caro o imposible construirlo.
- 2a. El costo de muestreo se incrementa mucho por la dispersión de las unidades, siendo más deseable obtener la muestra de un modo menos disperso.

Cuando ocurra al menos una de las dos situaciones antes señaladas se recomienda el uso del muestreo por conglomerados.

La ventaja principal es que solo debe elaborarse un marco para las unidades primarias de muestreo y las de las etapas subsecuentes. El trabajo de campo resulta menos costoso porque la encuesta solo se realizara en las U.P.M. seleccionadas y por lo tanto se reduce el costo de los viajes.

En lo que se refiere a estimación de totales, medias razones y proporciones de la población a partir de un diseño determinado de submuestreo (o multietápico), el principio básico es el de formar las estimaciones empezando con las unidades de la última etapa y terminando con las de la primera.

En la encuesta presentada en esta tesis, las unidades que se requerían investigar directamente eran las plantas de nopal de tuna blanca por lo que resultaba imposible construir un marco para dichas unidades, tomando en cuenta estas consideraciones se necesitó construir un marco para seleccionar unidades más grandes o conglomerados (huertas), en vez de seleccionar directamente las plantas. Como puede notarse el tamaño era distinto para cada conglomerado.

Ahora bien, primero seleccionamos las huertas y en cada huerta seleccionada se obtuvo una muestra de plantas, a esto se le llama muestreo bietápico, puesto que ahora la muestra se selecciona en dos etapas, primero las huertas (llamadas unidades primarias de muestreo ó de primera eta-

pa) y despues las plantas (unidades de segunda etapa) dentro de las huertas. También se le llama submuestreo, puesto que se ha tomado una muestra adicional (de plantas) de la muestra original de huertas.

B. MUESTREO SISTEMATICO.

Cuando la población objeto se encuentra en el marco distribuida de manera aleatoria es más fácil obtener la muestra usando muestreo -- sistemático; por esta razón (simplicidad en la selección de la muestra) el método es usado extensivamente en la práctica.

El muestreo sistemático consiste en dividir la población de N unidades en n subgrupos ordenados y tomar un elemento de cada subgrupo; - en este método únicamente la primera unidad es seleccionada con la ayuda de número aleatorios, el resto se selecciona automáticamente a partir de la primera. Para seleccionar la primera unidad tomamos aleatoriamente un número i, tal que $1 < i < k$ y a continuación de manera rígida o sistemática - (de aquí el nombre del procedimiento) se van tomando las unidades:

$$U_{i+k} , U_{i+2k} , \dots , U_{i+(n-1)k}$$

de tal manera que los elementos incluidos en la muestra son:

$$\{ U_i , U_{i+k} , U_{i+2k} , \dots , U_{i+(n-1)k} \}$$

Hay que hacer notar que estamos suponiendo que $N=nk$, donde n es el tamaño deseado de la muestra y k es un número entero tal que $k = \frac{N}{n}$. - A k se le llama intervalo de selección.

Evidentemente se cumple la siguiente expresión:

$$i+(n-1)k \leq nk = N$$

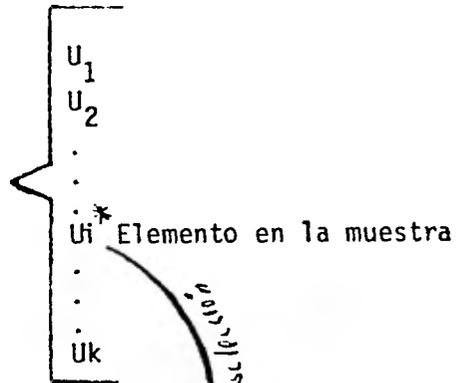
La conveniencia de este método esta en que la selección del - primer número de la muestra determina automáticamente toda la muestra.

Graficamente, el método se puede observar como sigue:

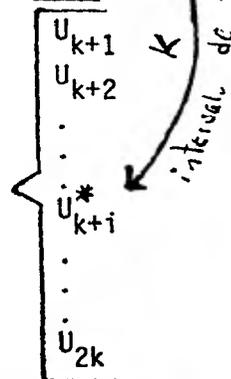
Subgrupos

Unidades

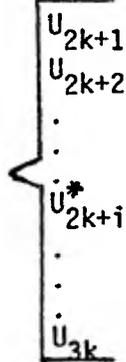
1



2

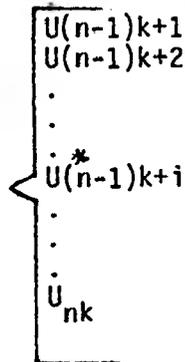


3



\vdots

n



Puesto que el primer número se selecciona al azar entre 1 y k, cada unidad recibe la misma probabilidad $\frac{1}{k}$ de selección.

Observación:

Una muestra sistemática de n unidades equivale a una muestra de un grupo o conglomerado de unidades, seleccionada de k conglomerados de n unidades cada uno (ver el siguiente diagrama), ya que cada unidad U_i en la población pertenece solo a un conglomerado. La probabilidad de seleccionar un conglomerado es $\frac{1}{k}$, y es por lo tanto la probabilidad con lo que se selecciona cualquier miembro del conglomerado en la muestra, es decir, la probabilidad de selección de un elemento poblacional cualquiera será igual o la de que resulte elegido el conglomerado que lo contiene, esto es, $\frac{1}{k} = \frac{n}{N}$ y es la misma que en el muestreo aleatorio simple, esto comprueba que el muestreo sistemático es un procedimiento de muestreo probabilístico.

Conglomerado	Composición del conglomerado
1	$U_1, U_{k+1}, \dots, U_{(n-1)k+1}$
.	
.	
.	
i	$U_i, U_{k+i}, \dots, U_{(n-1)k+i}$
.	
.	
.	
k	$U_k, U_{2k}, \dots, U_{nk}$

El muestreo sistemático es una herramienta delicada que puede ser igual, mejor o peor que el muestreo aleatorio simple.

IGUAL: En el caso que la población esta en orden aleatorio en lo que respecta a los valores de Y_i , el muestreo sistemático es equivalente al muestreo aleatorio simple y por lo tanto se usan las mismas expresiones para estimar la media \bar{y} o el total de la población cuando se utiliza un muestreo aleatorio simple.

MEJOR: Cuando la población esta ordenada en relación a los valores Y_i con tendencia a cambiar paulatinamente dichos valores (tendencia creciente o decreciente), el muestreo sistemático produce varianzas de los estimadores menores que el método del muestreo aleatorio simple.

PEOR: Si la población tiene un orden que se refleja en cambios periódicos de los valores Y_i , y que coinciden con el intervalo de selección k , el muestreo sistemático produce varianzas mayores.

Cuando se utiliza el muestreo sistemático en poblaciones que no tienen un orden aleatorio no existen expresiones válidas para estimadores y varianza de estimadores. Sin embargo, cuando el sistemático resulta mejor que el muestreo aleatorio simple se pueden usar las expresiones de este último como una aproximación y se sabe que en realidad las varianzas seran menores.

En el muestreo sistemático generalmente se supone una ordenación lineal, asimilable a una lista de los elementos de la población, sin embargo se han establecido métodos de muestreo sistemático sin prescindir de la bidimensionalidad del espacio que se considere.

Existen principalmente dos tipos de muestra sistemática bidimensional:

- a) La muestra alineada o "red cuadrada" (es la extensión más simple de la muestra unidimensional en el muestreo sistemático).

-Para la muestra desalineada

Seleccionamos de manera independiente, n enteros aleatorios i_1, i_2, \dots, i_n cada uno menor o igual a ℓ y m enteros aleatorios j_1, j_2, \dots, j_m cada uno menor o igual a k .

Entonces las unidades seleccionadas para la muestra tienen las siguientes coordenadas.

$$(i_1 + r\ell, j_{r+1}), (i_2 + r\ell, j_{r+1} + k), (i_3 + r\ell, j_{r+1} + 2k), \dots,$$

$$(i_n + r\ell, j_{r+1} + (n-1)k) \quad r = 0, 1, 2, \dots, (m-1)$$

Ejemplo:

Supongamos que tenemos las mismas condiciones del ejemplo anterior, es decir, $m\ell=9$, $nk=12$, $mn=9$, entonces $m=3$ y $n=3$

Supongamos que:

$$i_1 = 2; \quad i_1 \leq \ell$$

$$i_2 = 3; \quad i_2 \leq \ell$$

$$i_3 = 1; \quad i_3 \leq \ell$$

$$j_1 = 2; \quad j_1 \leq k$$

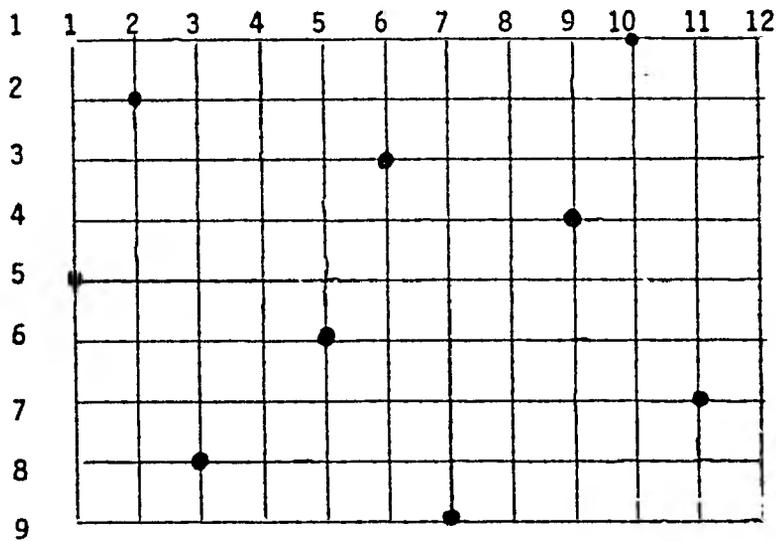
$$j_2 = 1; \quad j_2 \leq k$$

$$j_3 = 3; \quad j_3 \leq k$$

Entonces, las unidades seleccionadas en la muestra se presentan en el siguiente cuadro:

	i_1	i_2	i_3
$r = 0$	(i_1, j_1)	(i_2, j_1+k)	(i_3, j_1+2k)
$r = 1$	(i_1+l, j_2)	(i_2+l, j_2+k)	(i_3+l, j_2+2k)
$r = 2$	(i_1+2l, j_3)	(i_2+2l, j_3+k)	(i_3+2l, j_3+2k)

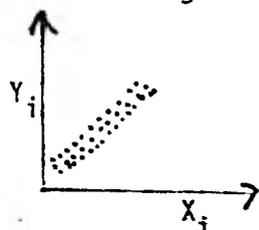
De donde, la muestra queda distribuida como se muestra en la figura:



C. ESTIMADORES DE RAZON

Generalmente se obtienen estimaciones basadas en medias aritméticas simples de los valores observados de la muestra, sin embargo, si se tiene conocimiento a priori que la razón de una variable (Y) entre otra (X) es bastante estable en toda la población (es decir, que tenga menor variabilidad que las mismas Y), se puede usar este conocimiento para cons-truir mejores estimadores.

Para usar el estimador de razón es fundamental considerar que Y_i es proporcional a X_i , es decir, que la relación entre esas dos varia-bles sea como se muestra en la figura:



Aproximadamente $Y_i = R X_i$

Si no hay proporcionalidad entre Y_i & X_i el estimador de razón no es adecuado, ya que es precisamente la proporcionalidad entre X_i & Y_i - la que se requiere para el buen uso de los estimadores de razón.

En la población $P = \{U_1, \dots, U_N\}$ se determinan dos mediciones - para cada unidad, y $(U_i) = Y_i$ & $X(U_i) = X_i$, además los valores de X_i se conocen. En este caso, los valores de X_i forman parte del marco. Se supone que existe el valor desconocido.

$$R = \frac{\sum_{i=1}^N Y_i}{\sum_{i=1}^N X_i} = \frac{Y}{X} = \frac{\bar{Y}}{\bar{X}}$$

En la práctica, X_i es a menudo el valor de Y_i en una ocasión anterior en que se hizo un censo completo.

Un estimador de Y , con base en una muestra donde se determinan valores $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$ es

$$\hat{Y}_R = X \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i} = X \frac{Y}{x} = X \frac{\bar{y}}{\bar{x}} = X\hat{R}$$

y es llamado el estimador de razón de Y , el total de población de las Y_i , donde y , x son los totales de Y_i y X_i en la muestra respectivamente.

Si X_i es el valor de Y_i en una ocasión previa, el método de la razón hace uso de la muestra para estimar el cambio relativo $\frac{Y}{X}$ que ha ocurrido desde ese tiempo.

El cambio relativo estimado y/x se multiplica por el total conocido de la población X en la ocasión anterior para obtener una estimación del total de la población actual. Si la razón Y_i/X_i es casi la misma en todas las unidades de muestreo, los valores de y/x varían poco de una muestra a la otra y el estimador de la razón es de gran precisión.

Si la cantidad a estimar es \bar{Y} , el valor medio de Y_i en la población, el estimador de razón es:

$$\hat{\bar{Y}}_R = \frac{\hat{Y}}{N} = \frac{y}{x} \bar{X}$$

Hay ocasiones en donde las X_i no se conocen y sin embargo se quiere estimar R , entonces se usa.

$$\hat{R} = \frac{\bar{y}}{\bar{x}} = \frac{y}{x}$$

Este último estimador es sesgado, es decir, $E(\hat{R}) \neq R$ entonces:

$$E(\hat{R}) = R + \text{Sesgo del estimador de razón}$$

$$E(\hat{R}) = R + B(\hat{R})$$

entonces:

$$B(\hat{R}) = E(\hat{R}) - R$$

donde:

$$B(\hat{R}) = - E(\bar{x})^{-1} \text{Cov}(\hat{R}, \bar{x}) \quad (1)$$

Hay que notar que si $\text{Cov}(\hat{R}, \bar{x})=0$, entonces $B(\hat{R})=0$ y \hat{R} es por lo tanto insesgado.

Para demostrar que \hat{R} es sesgado y que tiene un sesgo dado por la expresión (1), hay que demostrar que:

$$E(\hat{R}) = R - E(\bar{x})^{-1} \text{Cov}(\hat{R}, \bar{x})$$

Demostración:

$$\begin{aligned} R - E(\bar{x})^{-1} \text{Cov}(\hat{R}, \bar{x}) &= \frac{\bar{Y}}{\bar{X}} - \frac{1}{\bar{X}} \text{Cov}(\hat{R}, \bar{x}) \\ &= \frac{1}{\bar{X}} \left[E(\bar{Y}) - \text{Cov}\left(\frac{\bar{Y}}{\bar{X}}, \bar{x}\right) \right] \\ &= \frac{1}{\bar{X}} \left[E(\bar{Y}) - \left\{ E\left(\frac{\bar{Y} \cdot \bar{x}}{\bar{X}}\right) - E\left(\frac{\bar{Y}}{\bar{X}}\right) E(\bar{x}) \right\} \right] \\ &= \frac{1}{\bar{X}} \left[\cancel{E(\bar{Y})} - \cancel{E(\bar{Y})} + E(\hat{R}) E(\bar{x}) \right] \\ &= \frac{1}{\cancel{E(\bar{x})}} \left[E(\hat{R}) \cancel{E(\bar{x})} \right] \end{aligned}$$

$$R - E(\bar{x})^{-1} \text{Cov}(\hat{R}, \bar{x}) = E(\hat{R})$$

$$E(\hat{R}) = R - E(\bar{x})^{-1} \text{Cov}(\hat{R}, \bar{x})$$

REFERENCIA BIBLIOGRAFICA.

1. Azorin Poch, Francisco, Curso de muestreo y aplicaciones España, Aguilar, 1972, 375 p.
2. Carvajal Moreno, Raul. Interdisciplina: definición y aplicaciones Serie B: Investigación Vol. 6, No. 121, México, I.I.M.A.S. 1975.
3. Des, Raj. La estructura de las encuestas por muestreo. México, - Fondo de cultura económica, 1979, 475 p.
4. Des, Raj. Teoría del muestreo. México, Fondo de Cultura económica, 1980, 305 p.
5. Fr. W. Kremer y G. Unterstenhofer. Valoración de Resultados de ensayos fitosanitarios según el método de Townsend y Heuberger. Instituto Biológico de Farbenfabriken Bayer AG, Leverkusen.
6. G. Cochran, William. Técnicas de Muestreo. México, CECSA, 1971 507 p.
7. Kish, Leslie. Muestreo de Encuestas. México, Trillas, 1975, 739 p.
8. Méndez Ramírez, Ignacio. Conceptos muy elementales del muestreo - con énfasis en la determinación práctica del tamaño de muestra. - serie azul: Monografías. Vol. 3, No. 25, México, I.I.M.A.S. 1976. 55 p.
9. Méndez, et al. Diseño de muestra para estimación de cosecha de café en México. Serie Naranja: Investigaciones. Vol. 6, No. 91, México, - I.I.M.A.S. 1975. 19 p.
10. Méndez Ramírez, Ignacio. Modelos estadísticos lineales. interpretación y aplicaciones. México foccavi/Conacyt. 1976. 140 p.

11. Snedecor, & W.G. Métodos Estadísticos México CECSA, 1978, 703 p.
12. V. Sukhatme, P & V. Sukhatme B. Sampling Theory of Surveys with - applications. U.S.A. IOWA STATE UNIVERSITY PRESS, 1970, 452 p.
13. V. Sukhatme, P. Teoria de encuestas por muestreo con aplicaciones 2a. edición, México, Fondo de cultura económica: 1962. 495 p.
14. Villarreal y Dominguez, Ricardo German. El método de muestreo por areas en la investigación agrícola. Tesis profesional. Facultad - de Ciencias, U.N.A.M., México, 1970. 74 p.