

03043



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

**INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS
APLICADAS Y EN SISTEMAS**

**EL ENFOQUE ESTADÍSTICO EN EL ÁREA DE
ASEGURAMIENTO DE CALIDAD: UNA APLICACIÓN
CON MUESTREO SISTEMÁTICO.**

T E S I N A

**QUE PARA OBTENER EL DIPLOMA DE
ESPECIALIZACIÓN EN ESTADÍSTICA**

A P L I C A D A

P R E S E N T A:

ROBERTO RAMÍREZ HERNÁNDEZ



DIRECTOR: MTR. FRANCISCO SÁNCHEZ VILLARREAL

MEXICO, D.F.

2005

m343223



Universidad Nacional
Autónoma de México



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.

Dedico este trabajo de manera especial a:

Judith, mi esposa, que con tanto amor y apoyo ha contribuido más de lo que ella piensa. Gracias por ser mi impulso principal.

Mis padres Roberto y Catalina y mis hermanos, Arturo y Caro, que con sus palabras de aliento y sus inmerecidos elogios, me recuerdan lo indispensable que es la familia. Gracias.

Mi Dios, no es necesario decirlo. ÉL sabe por qué.

Mi querida UNAM, que me ha formado como un hombre de bien, me ha dado TODO, jamás podré retribuir suficientemente lo recibido. Gracias y espero ahora ser yo, el que contribuya a tu grandeza.

ÍNDICE

INTRODUCCIÓN.

iii

1. ANTECEDENTES GENERALES: UNA REVISIÓN RÁPIDA DEL PROCESO GENERAL DEL SURTIDO DE PEDIDOS.

1.1 Descripción General de las Actividades de la Compañía	1
1.2 Etapas en un Proceso de Venta Directa	1
1.3 El Proceso de Venta Directa en la Compañía	3
1.4 Primeros Procedimientos de Revisión de Pedidos	8
1.5 Problemática de los Métodos de Revisión de Pedidos	12

2. TÉCNICAS DE MUESTREO: ALGUNAS POSIBILIDADES PARA LA REVISIÓN DE PEDIDOS.

2.1 El Uso de Técnicas de Muestreo Probabilista ó Probabilístico	14
2.2 Control de Calidad y Bandas de Confianza	17
2.3 Algunos Elementos de Muestreo	18
2.4 Muestreo Aleatorio Simple Sin Reemplazo o Muestreo Irrestringido Aleatorio	19
2.5 Estimación de Proporciones	22
2.6 Estimador de la Proporción Poblacional Bajo Muestreo Irrestringido Aleatorio	24
2.7 Muestreo Aleatorio Estratificado	28
2.8 Muestreo por Conglomerados	38
2.9 Muestreo Sistemático	41
2.10 Otros Esquemas de Muestreo	48

3. ANÁLISIS DE CADA TÉCNICA DE MUESTREO: SELECCIÓN Y JUSTIFICACIÓN DEL USO DEL MUESTREO SISTEMÁTICO.

3.1 Un Análisis de las Características Prácticas de Cada Metodología	49
3.2 Aplicación del Muestreo Sistemático como Método de Selección	57
3.3 Beneficios Obtenidos	61
3.4 Diseño Sugerido de un Control Estadístico de Proceso	63

CONCLUSIONES.

66

APÉNDICE 1	69
APÉNDICE 2	70
APÉNDICE 3	71
APÉNDICE 4	72
BIBLIOGRAFÍA.	73

INTRODUCCIÓN.

En el presente trabajo se plantean algunos de los problemas más comunes a los que se enfrenta el personal que integra las áreas relacionadas con Aseguramiento de Calidad, profundizando, desde luego, en el punto de vista de la Estadística Aplicada y en su participación directa en esta área.

Específicamente se aborda un problema relacionado con la generación y surtido de pedidos de productos de higiene y belleza (donde dicho proceso de surtido se realiza a mano por un equipo de trabajadores). Aquí podemos observar cómo el factor de error humano al surtir dichos pedidos puede ser tan grande que puede provocar un proceso de surtido fuera de control. Es este el punto donde queda plenamente justificado el uso de técnicas de muestreo al azar en un afán de tener bajo control el proceso de surtido de tales pedidos.

En el presente trabajo se enfatiza también el beneficio desde el punto de vista del inherente ahorro de recursos y costos que se logra con el muestreo, teniendo así un beneficio global muy importante para el área de Aseguramiento de Calidad. Se llegó a la conclusión de que el beneficio más importante tal vez sea el que los usuarios de las técnicas de muestreo (los ingenieros de calidad) tengan plena conciencia de la importancia que tiene el muestreo, así como el uso correcto de las técnicas de muestreo, dejando de esta manera, la puerta abierta a usos futuros en el área para ésta y otras importantes técnicas estadísticas. En resumen, se fomenta con esto la cultura de la información estadística en áreas de aplicación como la de Aseguramiento de Calidad.

I

***ANTECEDENTES GENERALES: UNA
REVISIÓN RÁPIDA DEL PROCESO
GENERAL DEL SURTIDO DE PEDIDOS.***

1.1 DESCRIPCIÓN GENERAL DE LAS ACTIVIDADES DE LA COMPAÑÍA.

El giro de la compañía, consiste en la fabricación y distribución de cosméticos, artículos de perfumería, higiene y belleza, así como lencería, joyería y artículos para el hogar en general. La actividad comercial que practica es denominada “venta directa” debido a la mecánica de ésta y a la distribución de los productos. Específicamente parte de la promoción de los artículos mediante un folleto (catálogo) en el que aparecen fotografiados con su respectivo precio y las ofertas del momento.

1.2 ETAPAS EN UN PROCESO DE VENTA DIRECTA.

Cada compañía cuenta, desde luego, con procesos muy particulares en sus tareas, que pudieran ser diametralmente opuestos entre sí aún persiguiendo objetivos similares, pero podríamos hablar de cuatro etapas básicas en las que consta el proceso general llevado por una empresa de venta directa. Se describen a continuación:

a) Promoción de productos (Levantamiento de Pedidos).

Como primera etapa, la fuerza de ventas, es decir, el equipo humano encargado de la comercialización de los productos de la compañía, hace su labor en las áreas geográficas donde se tenga cobertura, en nuestro caso particular es en toda la República Mexicana.

b) Recepción y Validación de Pedidos.

Las vendedoras envían las órdenes de compra de los clientes (Pedidos) a la oficina central para su validación, es decir, la revisión que se hace para determinar los pedidos que preceden al surtido, lo cual depende de la disponibilidad de los productos en los almacenes, de la existencia de errores por parte de la vendedora al asentar el pedido, crédito disponible de la vendedora (pues los artículos los compra previamente la vendedora para ella cobrarles a sus clientes a la hora de entregar los pedidos), etc. Una vez validados los pedidos, se generan listas de surtido de productos que serán canalizadas al área de surtido de la compañía.

c) Surtido de Pedidos.

Se procede a abastecer las órdenes de compra con los productos solicitados. El proceso se basa en las listas de surtido previamente proporcionadas por el área que recibe las órdenes.

d) Distribución de Pedidos.

Los pedidos ya surtidos, se conducen a un área de acopio donde serán distribuidos, es decir, enviados a los lugares donde fueron originados, cerrándose así el ciclo de venta directa. Algunas compañías cuentan con canales de distribución propios (transportación terrestre, regularmente), aunque pueden ser distribuidos con proveedores externos (mensajería o camiones de reparto alquilados). Regularmente en esta etapa se da el proceso de cobro de los pedidos, es decir, los choferes repartidores pueden ser los encargados de la tarea, otras veces es la propia vendedora la que hace el depósito previamente por el importe completo de los pedidos logrados.

El siguiente diagrama describe con detalle el procedimiento:

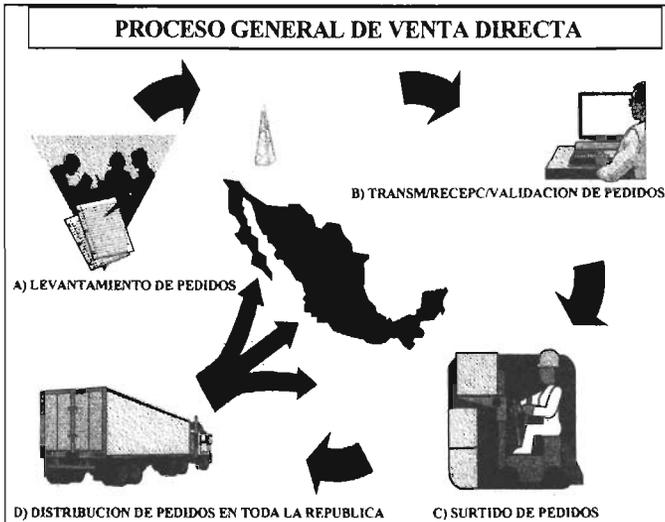


DIAGRAMA 1.1

1.3 EL PROCESO DE VENTA DIRECTA EN LA COMPAÑÍA

Como se mencionaba en un principio, las anteriores etapas tienen variaciones de una empresa a otra, en el caso particular de esta empresa, existen detalles que vale la pena describir. Por ello se expone a continuación, de manera general, el proceso de venta directa que lleva a cabo esta organización, de esta manera, se llega a una mejor comprensión de la problemática del surtido como uno de los objetivos de este trabajo, y que representa una de las etapas centrales en el proceso de venta directa.

1.3.1. Promoción de productos (Levantamiento de Pedidos).

La promoción de los artículos que produce y distribuye la compañía se hace a través de un folleto, como se había comentado antes, con la fotografía de los mismos, con este folleto la vendedora puede ejercer su labor de venta con facilidad pues el posible cliente puede conocer muchas de las características de los productos y por ello interesarse en algunos de los mismos, por ejemplo, en el precio ofrecido, las ofertas del momento, nuevos productos, etc.

Una vez que la vendedora ha logrado que la gente se interese realmente por al menos uno de los artículos ofrecidos, procede a registrar en un formato preimpreso todos los artículos pedidos por sus clientes, en este formato se encuentran las claves de todos los productos ofrecidos en el folleto, con cuadros para asentar las cantidades pedidas como se hace en una nota de remisión.

Este formato es el que se denomina “orden de compra del cliente” o “pedido” más comúnmente. La vendedora puede levantar cualquier cantidad de pedidos en una campaña, aunque lo normal es que genere como máximo un pedido por campaña (en él asienta los artículos pedidos por todos los clientes en la campaña). Se agrega como **anexo 1** un formato típico de registro de artículos pedidos por una vendedora,

es decir, un pedido. Consiste en varias hojas, aunque aquí se presenta sólo una para efectos de ilustración.

Existe un periodo determinado para la difusión y venta de los productos contenidos en el folleto, al que llamamos, como se dijo al principio, “campana”; sin embargo, es importante aclarar que la campana de venta difiere de la campana del surtido, esto significa que hay un desfase entre las dos, la campana de ventas (de difusión de los artículos) consta de un periodo de 20 días naturales, mientras que la campana de surtido consta de 14 días naturales, esta última comienza, por regla general, al doceavo día de iniciada la campana de venta, esto tiene lógica si pensamos en que debe haber un tiempo para hacer acopio y orden de pedidos, así como de transmisión de los mismos a la oficina central y por último su surtido y distribución, todo con cobertura nacional.

1.3.2. Recepción y Validación de Pedidos.

Al cierre de la campana de ventas, se realiza una junta por regiones geográficas ya determinadas por la empresa, donde las vendedoras de cada una de éstas entrega a una jefa designada del grupo (llamada gerente de distrito) los pedidos logrados para que sea ésta la que haga acopio de los mismos, una vez revisados con cuidado y validados, la gerente de cada distrito en la república, captura por computadora todos los pedidos de sus vendedoras, para mandar posteriormente (vía satélite) los archivos con los pedidos capturados de todo el distrito. Todo lo anterior se realiza en la mencionada junta, a la cual se le conoce como “cierre de valija”.

Al hacer la transmisión de los pedidos de cada distrito geográfico, el área de “Recepción de Pedidos” los reúne en la oficina central conforme van llegando, es decir, que esta área es la encargada de registrar todos los pedidos del país. Como segunda función, realizan una validación (independiente de la realizada por las gerentes de distrito) de los pedidos recibidos, que consiste en verificar la viabilidad de los productos, en el sentido de que exista suficiente surtido de cada producto en

los almacenes de la empresa. Se detectan en esta fase también, errores de captura de las gerentes y/o vendedoras. Por último, se determina si la vendedora cuenta con viabilidad crediticia, pues, como se comentaba en un principio, es la vendedora la que paga a la empresa los artículos pedidos, para que sea esta misma la que cobre a sus clientes a la hora de entregarles sus productos.

1.3.3. Surtido de Pedidos.

Una vez que el área de recepción de pedidos ha depurado y validado los mismos, genera (vía computadora) una serie de listas impresas, cada una de estas listas tiene indicados los artículos pedidos en cada hoja original. Estas listas son canalizadas al área de surtido para después servir como relación de artículos a surtir.

Se presenta como **anexo 2**, la copia de una lista de surtido típica, usada por el equipo de surtido en su trabajo cotidiano.

Al revisar el anexo, podemos apreciar en la parte superior de la hoja, los números de la zona y de la ruta, que son subdivisiones y “sub subdivisiones” del distrito geográfico, que se usan para control y programación del surtido. Vemos también la campaña (la de surtido), el número de pedido y la clave de la vendedora que lo emitió (en este caso FULL. 1028097).

En el cuerpo de la hoja es posible conocer a detalle los productos que se piden, podemos saber los productos que se requieren de cada estación de surtido (señalada con una letra) y de cada línea de surtido, la cantidad, la localización y la clave del producto en particular. Por ejemplo, en la hoja de surtido mostrada nos dice que de la estación de surtido “A” se debe poner en el pedido un artículo con clave “H”, el cual se encuentra localizado en el casillero AJ de la referida estación, de la misma manera, de la estación de surtido “B” es necesario poner en el pedido 3 artículos, que son los correspondientes a las claves “SCP” (uno) y dos de la clave “F6”, que se

encuentran ubicados en los casilleros AF y AJ respectivamente (Puede ser cualquier línea de surtido). Todo el proceso del surtido se describe con detalle a continuación.

El área de surtido emplea cuatro bandas mecánicas donde son desplazadas las cajas aún vacías (en esta etapa sólo contienen el instructivo o lista de surtido mencionada y en donde se hace referencia a los artículos que se deben de considerar en el pedido) y que contendrán todos y cada uno de los pedidos. A lo largo de las cuatro bandas o líneas de surtido está integrada una flotilla de trabajadores que harán la labor de surtido, se incorporan al proceso del surtido en promedio alrededor de 80 surtidores para las cuatro bandas mencionadas (20 por banda).

Cada pareja de surtidores, permanece en un punto fijo junto a la banda, a sus espaldas tiene una estación de surtido identificada con una letra, que consta de varios casilleros rotulados con claves llamadas localizaciones, que contienen un número de determinados productos como se había comentado anteriormente, con los cuales se surtirán los pedidos.

Es importante insistir en que ningún surtidor cuenta con todos los productos sino solamente algunos, por lo que al pasar la caja por su puesto o estación de surtido, deberá verificar (a través del instructivo o lista de surtido) si le toca surtir con alguno o algunos de los productos a su cargo, al terminar, debe dejar que la caja siga por la banda mecánica a otro punto de surtido de manera que el siguiente par de surtidores verificará igualmente si debe surtir con su grupo de productos o no en el mismo pedido. Como puede observarse, todos los surtidores de la banda tienen la obligación de verificar si les corresponde surtir o no. Dicho de otra manera, todos los pedidos que debe surtir la compañía pasan por muchas manos antes de ser enviados a su lugar de distribución.

Es importante aclarar que las cuatro bandas mecánicas se deben considerar independientes desde el punto de vista de que un pedido en particular se moverá

solamente por una y sólo una banda en todo su proceso, por lo que las bandas surten de manera simultánea los pedidos (que son asignados a una banda en particular en función de la zona geográfica del país a donde vayan después).

Podemos conocer la manera que se ubican físicamente las cuatro bandas o líneas de surtido mencionadas, así como entender la forma de trabajo del área de surtido de pedidos a través del **anexo 3**.

1.3.4. Distribución de Pedidos.

Una vez que se ha realizado por completo el surtido del pedido, la misma banda mecánica conducirá las cajas a un área de sellado o empaque (véase diagrama de las líneas de surtido del anexo 3), cuando se hace esto, la banda transporta el pedido ya listo a donde es recogido por la gente de distribución (los choferes) y llevado a los camiones de distribución para su traslado al punto del país donde deben llegar, todo esto con sus órdenes de envío y plan de distribución previamente elaborado. Posteriormente los pedidos serán recibidos por la vendedora original, que se encargará de darle personalmente el pedido al cliente en cuestión.

De acuerdo a los cálculos hechos por la misma empresa, se estima que se generan diariamente 12,500 pedidos en promedio (12,500 cajas de pedido) que pasan por las cuatro bandas en conjunto, dado que las bandas trabajan al mismo ritmo y de modo simultáneo, se calcula que pasan cerca de 3,125 pedidos por cada banda mecánica a diario. Como podemos notar, la generación de pedidos es muy considerable.

Bajo este contexto, es claro que, en la mayoría de los procesos de la compañía, el factor humano tiene una intervención muy importante. Por otra parte, la cantidad de artículos y de pedidos en cada campaña son demasiado altos como para llevar a cabo las tareas sin un control adecuado, la situación en este punto era manejada bajo criterios entusiastas y con cierto sentido común pero con poca objetividad, es decir con la mejor de las intenciones pero sin ideas claras que nos permitieran conducir las operaciones con mayor eficacia. Lo

anterior será ilustrado con el problema de interés en este trabajo, es decir, con las tareas llevadas a cabo en el contexto del surtido de pedidos.

1.4 PRIMEROS PROCEDIMIENTOS DE REVISIÓN DE PEDIDOS.

Evidentemente (como se planteaba en el punto anterior), el hecho de que sean seres humanos los encargados de los procesos de surtido de pedidos, conllevan un factor de error (le llamaremos error de surtido) que no había sido cuantificado pero que se sabía que existía. Los errores de surtido pueden ser clasificados en cuatro tipos:

- a) Pedidos que contienen productos de menos (que implican las reclamaciones más frecuentes)

- b) Pedidos que contienen productos cambiados (con un número menor pero también considerable de reclamaciones)

- c) Pedidos que contienen productos de más (con una frecuencia de reclamaciones prácticamente nula por razones obvias, pero que tienen un costo para la compañía)

- d) Una combinación de las anteriores

Es evidente que lo anterior le implica a la compañía un deterioro en su imagen (con el costo implícito por ello) y en la de los productos. Estas razones tienen suficiente peso para intentar eliminar estos errores o por lo menos reducirlos a una proporción que la empresa considere “bajo control”.

Lo anterior, motivó a utilizar alguna técnica de muestreo para detectar el nivel de error en el surtido de los pedidos. Dentro del mismo equipo de ingenieros de calidad se propuso y se implantó un método que consistía en lo siguiente:

- Revisión del 30% (la gente de control de calidad determinó esa fracción de muestreo bajo criterios desconocidos) de los pedidos generados a diario, es decir alrededor 3,750 pedidos, para ello se contaba con un equipo de 23 personas, encargadas de hacer las revisiones.
- El criterio de determinación de pedidos erróneos consistía únicamente en decir “el pedido está bien surtido” o “mal surtido” si contiene los artículos correctos en cantidad y tipo.
- Selección de las cajas “*al azar*”, es decir se extraía de una sola vez todo un grupo de cajas (el 30%) bajo juicio de los encargados de las revisiones, en otras palabras, la selección era a través de una muestra a juicio (aunque los ingenieros de calidad consideraban e insistían en que era “al azar”).

A pesar de los posibles riesgos en que pudiesen incurrir con este procedimiento (dado que no es muestreo probabilístico), los ingenieros de calidad aseguraban que era eficiente porque lograban estimar un “nivel de error” (como ellos le llamaban) constante por lo que lo consideraban adecuado. Dicho nivel de error estaba en esa época en alrededor de 5% (es decir 5% de los pedidos diarios contenían alguno de los errores descritos anteriormente, esto es que, se concluía que se generaban alrededor de 625 pedidos en promedio en toda la población, con algún problema en el surtido).

Al paso del tiempo, la empresa tuvo que prescindir de los servicios de muchos de sus empleados, específicamente de casi todo el equipo de revisión de pedidos, que se redujo drásticamente de los 23 originales a sólo 8 para el mismo trabajo, con lo cual resultó inoperante el método debido a la carga de trabajo tan grande (el triple) para los 8 “sobrevivientes”.

Dada la problemática anterior, se pensó en el rediseño de los procedimientos de revisión. El adecuarlos a tan pocos recursos y garantizar una estimación confiable del nivel de error parecía difícil de lograrse. Se propuso y se implementó lo siguiente:

i) Revisión de una zona (geográfica) por cada línea de surtido, esto es, 4 zonas diarias, que contienen alrededor de 100 pedidos promedio, lo que daba un total de 400 pedidos aproximadamente, y que era la máxima carga a que se podía aspirar a absorber con tan poco personal de revisión. Esto nos da una fracción de muestreo de aprox. 3.2% a diario.

ii) La logística de revisión consistía en “bajar” las 100 cajas por línea prácticamente al mismo tiempo para ser revisadas en conjunto y tardarse varias horas en este proceso (justo como en el método anterior), evidentemente los surtidores se daban perfecta cuenta de todas las actividades del equipo de revisión, por ello es muy lógico sospechar que si se dedicaban a revisar determinadas cajas (el equipo de revisión) y a sabiendas de que ya no serían revisados el resto de los pedidos en el día, los surtidores no pondrían el mismo esmero en surtir correctamente (“¿Para qué, si ya no me van a revisar?”).

iii) El criterio de selección de las 4 zonas a ser revisadas indicaba las 4 zonas que más reclamaciones del campo se reportaban en la campaña anterior (los clientes y las vendedoras comisionistas), por concepto de errores en el surtido de sus pedidos. Esto evidentemente no es un indicador real de que esas 4 zonas fueran las de

mayores problemas, si acaso eran las que más reclamaciones reportaban en la campaña anterior.

iv) El hecho de revisar de una vez todo un grupo de cajas de esta manera, no contemplaba las posibles variaciones que hubiera en el ánimo y concentración que pudieran tener los surtidores y que podría derivar en diferentes niveles de eficiencia en el momento de surtir y en diversos intervalos de la jornada diaria. Esto último es importante considerarlo porque no es lo mismo, por ejemplo, para un surtidor, el realizar su labor a las 8:00 a.m. en que está fresco, descansado y más concentrado en su trabajo que a las 2:30 p.m. en que está terminando su jornada cotidiana, que ya se cansó, no tiene la misma concentración y por ello es más vulnerable a los errores de surtido. Visto de esta forma, podemos pensar en que el error de surtido pudiera tener un comportamiento cíclico que puede ser determinante en algún esquema de muestreo.

Bajo estos planteamientos, es de esperarse que se presentara mucha inconsistencia a la hora de estimar la proporción de pedidos con problemas de surtido. De hecho, al realizar las estimaciones, en un día se reportaba un nivel de error del 9% o 10% por ejemplo, mientras que para el día siguiente podía reportarse un nivel de error del 0.8% o de 1.2%. Lo cual, por supuesto, no resultaba creíble, además de que no podía ser usada esta cifra para tomar decisiones sobre acciones de mejora en la logística del surtido porque no se sabía “dónde estaban parados” realmente, y por ello no podía determinarse estrategia alguna.

1.5 PROBLEMÁTICA DE LOS MÉTODOS DE REVISIÓN DE PEDIDOS.

Ambos métodos de revisión implicaron una serie de deficiencias, amén de los resultados inservibles en el segundo y dudosamente creíbles en el primero, se describen a continuación para ambos métodos (salvo el inciso b aplicable sólo al primer método):

a) Dado que se trataba de una muestra a juicio, no era posible determinar una medida de variación (varianza) y por ello casi cualquier tamaño de muestra era dudosamente confiable.

b) Pese a estar reportando resultados aceptables, el tamaño de muestra empleado resultó muy costoso en términos de recursos usados, dado que para cumplir con este método, se requería de 23 personas y de una cantidad de horas considerable. (Todo el día laboral equivaldría a 8 horas, por lo tanto de 23 personas son $23 \times 8 = 184$ horas/hombre).

c) El hecho de extraer las cajas de pedido de una sola vez, como se comentó antes, muy bien podía provocar sesgos a la hora de estimar el nivel de error, pues al darse cuenta los surtidores de que ya no habría revisiones en el día bien podía tener “manos libres” en el surtido y con ello, a partir de ese momento, podría crecer considerablemente el error de surtido que, desde luego, sí sería detectado por un cliente o una vendedora comisionista.

d) El método no garantizaba (a pesar de los aparentes resultados) un verdadero control estadístico en el proceso debido a que, como no se tomaba en cuenta criterio formal alguno, no se establecían bandas de confianza y esto derivaba en que no se sabía a ciencia cierta en que momento el proceso estaría “fuera de control” y que por tanto se tendrían que tomar acciones emergentes para corregir el proceso, tal como se debe hacer en un control estadístico de procesos.

Como era de esperarse, la situación para la compañía comenzó a ser poco manejable ya que, por un lado, se conocía la necesidad de realizar muestreos en diferentes áreas de operación, muy particularmente en el área de surtido de pedidos, pero por otro no se tenía el conocimiento de cómo debían de llevarse a cabo ni con qué tamaños de muestra. En base a toda esta problemática, se pensó en replantear todo el proceso nuevamente, esta vez tomando en cuenta criterios estadísticos formales, que pudieran apoyar los procedimientos de revisión y obtener estimaciones confiables del nivel de error de surtido.

II

***TÉCNICAS DE MUESTREO: ALGUNAS
POSIBILIDADES PARA LA REVISIÓN DE
PEDIDOS.***

2.1 EL USO DE TÉCNICAS DE MUESTREO PROBABILISTA O PROBABILÍSTICO.

Un primer factor que puede advertirse desde el capítulo anterior es que no se cuenta con una muestra confiable, ya que el método para seleccionar los elementos (en este caso, los pedidos) no garantiza de forma alguna que el resultado sea veraz.

Es muy importante tener en cuenta que dentro de los esquemas disponibles para muestrear, aquellos que emplean muestreo probabilista, tienen varias ventajas. Estas técnicas tienen como característica el hecho de que la selección de los elementos que conformarán la muestra sea mediante un procedimiento de selección aleatoria, es decir, que involucre alguna forma en que se escojan los elementos al azar, por lo que cada unidad de la población tendrá un valor de probabilidad mayor a cero de ser escogido para integrar la muestra. Dado lo anterior, es posible también calcular un error de estimación, una varianza, o alguna otra medida de dispersión, o sea un valor con base en parámetros estadísticos que nos permita conocer la precisión de las inferencias realizadas por el experimentador a través de los estimadores que éste haya escogido.

Las técnicas de muestreo probabilista, por lo tanto, ofrecen algo que para los métodos de muestreo no probabilista es muy difícil de garantizar: precisión y confiabilidad conocidas, además de la posibilidad de tenerlas bajo control de aquel que muestrea, luego entonces, representa una metodología que en un amplio espectro de posibles circunstancias, resulta más eficiente a la hora de economizar recursos. William Cochran lo plantea de la siguiente manera: *"Para todo procedimiento que satisfaga estas condiciones, podemos calcular la distribución de frecuencia de las estimaciones que genera el proceso, si se aplica repetidamente a la misma población. Sabemos la frecuencia con que se elige cualquier muestra S_i , y sabemos cómo calcular la estimación a partir de los datos de S_i . Por lo tanto es claro que se puede desarrollar una teoría de muestreo para cada procedimiento de este*

tipo, aunque los detalles del desarrollo puedan ser intrincados. Un método de esta clase se conoce con el nombre de muestreo probabilista."¹

No obstante, los métodos de muestreo probabilista deben ser suficientemente conocidos si se quieren usar exitosamente, por eso es importante efectuar una revisión de lo que la teoría de muestreo probabilista nos puede ofrecer en el caso del surtido de pedidos.

Es importante decir que, el propósito de este capítulo no es plantear las técnicas de muestreo probabilista a fin de demostrar su validez estadística y matemática, ese objetivo ha sido cumplido cabalmente por muchos autores de textos sobre muestreo. La intención en este capítulo es simplemente plantear, a manera de una sencilla revisión, los esquemas teóricos de muestreo y conocer la manera en que mejor podemos ajustar la solución de nuestro caso de interés a estos esquemas. Además de señalar aquellos aspectos a cuidar, de manera que no se violen los supuestos teóricos mínimos que cada esquema exige.

Generalmente encontramos en los textos sobre técnicas de muestreo, que no se hace una distinción lo suficientemente clara respecto a la distinción entre los procedimientos de selección y los de estimación. No obstante es muy importante clarificar el concepto en cada caso. Los procedimientos de estimación se refieren a los métodos con que efectuamos el cálculo o estimación de un parámetro de nuestro interés, éste puede ser una media poblacional, un total o una proporción.

Los procedimientos de selección de la muestra consisten en los pasos y acciones que deben hacerse para escoger los elementos que conformarán la muestra de interés, de manera que podamos confiar en la representatividad de ésta. Evidentemente esos pasos y acciones van siempre orientados a usar métodos de selección de elementos al azar.

Conviene entonces, resumir lo anterior recordando los pasos que generalmente se siguen en el proceso de un diseño muestral probabilista:

¹ William G. Cochran, "Técnicas de Muestreo", p. 30

- a) **Diseño conceptual:** Se entiende globalmente el fenómeno que estamos conociendo y midiendo alguna característica particular, entonces determinamos los elementos de interés en el diseño: conocimiento de nuestra población objetivo, determinación del marco muestral, así como de las unidades de muestreo, variables de interés y definición de aspectos logísticos en la recolección de datos.
- b) **Procedimiento de selección de la muestra:** Una vez entendiendo el fenómeno, determinando el objetivo y qué vamos a medir, optaremos por un método probabilista de selección de los elementos de la muestra. Este método garantizará la representatividad de la muestra, en la inteligencia de que previamente fue bien aplicado.
- c) **Procedimiento de estimación:** Como parte del diseño, se deben considerar los métodos matemáticos que determinarán el cálculo de los parámetros que nos interesa conocer. Determinar cuáles serán esas expresiones (que llamaremos estimadores) incluye evidentemente, criterios estadísticos que nos garanticen una buena estimación. Los estimadores deberán cumplir por lo menos las propiedades de insesgamiento o de varianza mínima (en caso de estimadores sesgados)².

Para el caso del problema de las líneas de surtido, el primer punto a resolver consistía en resaltar y principalmente convencer de la necesidad de aleatorizar el proceso de revisión, es decir, cambiar el concepto original que los ingenieros de calidad de la organización tenían de lo que significa “azar”, esto es importante ya que existen muchos lugares en los que se considera que hacer muestreo al azar significa escoger o seleccionar una muestra bajo procedimientos que en realidad pertenecen al muestreo de juicio (o sea NO al azar o NO probabilista).

² Erwin Kreyszig, “Introducción a la Estadística Matemática”, pp. 175-181

Se desprenden de lo anterior muchas ideas que pueden resultar ineficientes, por ejemplo, es fundamental entender que decir por ejemplo: “escoge la caja de enfrente y la de atrás” no significa que sea aleatorio (pensemos que de esta manera, tal vez nunca se escogería una caja en el centro geométrico de un cúmulo gigante de cajas debido a lo impráctico que resultaría la revisión). En este proceso no se puede garantizar que todas las cajas tienen una probabilidad conocida de ser escogidas.

La mejor forma de confiar en la aleatoriedad de un proceso de selección de los elementos (las cajas de pedido) es emplear algún método conocido de generación de números aleatorios, por ejemplo, acudir a tablas publicadas de números aleatorios o mediante el uso de números aleatorios generados por una computadora (números pseudo-aleatorios), o bien, mediante un conteo que cumpla con lo establecido en un esquema de muestreo sistemático. En otro caso siempre se estará en la sospecha de que un ser humano sesgará los resultados de la selección, no permitiendo que la filosofía del muestreo probabilista beneficie al procedimiento.

Para entender cuál será el proceso de selección al azar más adecuado, es importante, como se planteó al principio de este capítulo, revisar ciertos conceptos importantes relacionados con el muestreo y las técnicas en sí mismas. Analizando y escogiendo posteriormente, en función a las características específicas del proceso de inspección de pedidos de la compañía.

2.2 CONTROL DE CALIDAD Y BANDAS DE CONFIANZA

A lo largo de este trabajo, se ha resaltado que uno de los aspectos de mayor relevancia para el área de Aseguramiento de Calidad, es justamente garantizar ésta mediante procedimientos que permitan a los interesados (los ingenieros de Calidad), mantener bajo su control, todos aquellos procesos que sean responsables del surtido de pedidos, de manera que también puedan corregir un eventual desfase que conlleve a errores graves en el surtido.

En palabras llanas, los ingenieros de Aseguramiento de Calidad requerían de diseñar e implantar un procedimiento de Control de Calidad basado en medir indicadores que dieran seguimiento a las condiciones generales del surtido y las variables que pudiesen desencadenar salidas a los parámetros normales en el proceso.

La discusión llevó a concluir que lo adecuado era implementar un control de tipo estadístico, esto es, emplear un dato que midiera el porcentaje de pedidos erróneos validado con procedimientos de estadística, además de sustentarlo con un procedimiento de muestreo probabilista.

El procedimiento consistiría en determinar un parámetro estadístico que indicara el nivel de error de surtido, establecer también un límite en el que se pudiera decidir si el proceso de surtido estaba bajo los parámetros de confiabilidad en el proceso o bien, si éste se había salido de “lo normal”. Esto quiere decir, implantar una banda o límite de confianza.

2.3 ALGUNOS ELEMENTOS DE MUESTREO.

Cuando se usa el muestreo, generalmente se tiene en mente conocer un dato específico, que centra las intenciones y motivos del estudio que se realiza. A ese dato lo llamamos parámetro³.

El parámetro de interés puede ser uno o más de los siguientes: un promedio, un total, o un porcentaje, por esta razón, cuando se habla de conocer un parámetro de interés, se habla también de estimar (calcular) una media, un total o una proporción. Es importante decir también, que al querer estimar un parámetro, existen también otros estimadores que a través del conocimiento de variables auxiliares, permiten también determinar nuevos parámetros de interés, quizá difíciles de obtener mediante los estimadores simples. Los métodos que usan variables auxiliares se denominan estimadores de razón, de regresión y

³ Ignacio Méndez Ramírez, “Conceptos muy elementales del muestreo con énfasis en la determinación práctica del tamaño de la muestra”, p. 2

de diferencia⁴. Sin embargo, no son de interés en este caso, como se comprobará más adelante, pues bastará con los estimadores simples.

2.4 MUESTREO ALEATORIO SIMPLE SIN REEMPLAZO O MUESTREO IRRESTRICTO ALEATORIO.

El esquema de muestreo irrestricto aleatorio (lo llamaremos MIA para mayor facilidad), que es quizá el más usado actualmente, sus ventajas saltan a la vista, especialmente su facilidad de comprensión y de aplicación.

Lo definimos de la siguiente manera:

Sea una población de N elementos o unidades, y extraemos una muestra de tamaño n , con las siguientes características:

Las unidades son escogidas sin reemplazo, esto es, selecciones independientes y secuenciales, de manera que la unidad escogida ya no será nuevamente seleccionada. Con lo que la probabilidad para cada selección será igual a:

$$\frac{1}{N-i} \text{ para } i = 0, 1, 2, 3, \dots, n-1$$

Las diferentes muestras que contengan exactamente las mismas unidades, independientemente de su orden dentro de las muestras en cuestión, se considera como muestras iguales, esto es, que su orden es poco relevante.

En otras palabras, dado que la selección de los elementos es sin reemplazo en el muestreo irrestricto aleatorio, la probabilidad de selección de una unidad en especial es $\frac{1}{N}$. Si tomamos en cuenta que con cada extracción se tienen eventos mutuamente excluyentes, la

⁴ William Mendenhall, "Elementos de Muestreo", p. 124

suma de las probabilidades de cada elemento, o dicho de otra manera, la probabilidad de observar el elemento en la 1ª, 2ª, 3ª o n-ésima selección es:

$$\frac{1}{N} + \frac{1}{N} + \frac{1}{N} + \dots + \frac{1}{N} \text{ n veces, es decir: } \sum_{i=1}^n \frac{1}{N} = \frac{n}{N}$$

En otros términos, en el muestreo irrestricto aleatorio extraemos de un universo de N elementos una muestra de tamaño n , debe ser posible identificar todos y cada uno de esos n elementos de manera que cualquiera pueda ser seleccionado para dicha muestra, esto quiere decir, que todos los elementos del universo, tienen la misma probabilidad de ser seleccionados.

Como se dijo con anterioridad, la selección al azar de un muestreo probabilista puede ser mediante el empleo de números aleatorios previamente generados de una computadora o consultados de una tabla para ese propósito.

Es importante ahora, identificar los elementos de un esquema irrestricto aleatorio en la notación típica:

N Tamaño de la población o Universo

n Tamaño de la muestra

y_i El valor de la variable estudiada en la i -ésimo elemento

Y Total de la población $Y = \sum_{i=1}^N y_i$

\bar{Y} Media de la población $\bar{Y} = \frac{\sum_{i=1}^N y_i}{N}$

f Fracción de muestreo $f = \frac{n}{N}$

\hat{Y} Estimador del Total

y Media de la muestra

$$y = \frac{\sum_{i=1}^n y_i}{n}$$

Visto lo anterior, los aspectos del problema en cuestión, son fácilmente identificables en la notación anteriormente definida, igualmente con el esquema del MIA. Tenemos entonces que los elementos de muestra son los pedidos a revisar, de allí podemos ver que el número total de pedidos generados diariamente representa el universo N , es decir $N=12,500$. El tamaño de muestra n , es decir, el número de pedidos que deberemos revisar a diario es precisamente uno de los datos que nos interesa conocer, por lo que, de momento no es conocido.

Lo que nos indica el MIA para este caso particular, sería que fuera posible que todos los pedidos generados diariamente puedan ser identificados mediante algún código, etiqueta o señalamiento, mismo que debería ser rápidamente identificable a fin de que la caja de pedido pueda ser extraída con la misma celeridad para su revisión. Se procedería entonces, una vez conocido el número de pedidos a revisar, es decir el tamaño de muestra n , a escoger aquellas cajas de pedido que nos indiquen los números aleatorios correspondientes para proceder a la inspección.

El resto corresponde al terreno de la estimación del parámetro de nuestro interés. Como se recordará, en el contexto del problema específico que nos ocupa, lo que interesa es estimar el porcentaje de pedidos que contienen algún error en el surtido, lo que enfoca la solución directamente a estimar una proporción, dicho lo anterior, la estimación de una media o de un total pierden interés en el presente trabajo.

2.5 ESTIMACIÓN DE PROPORCIONES.

Generalmente en la estimación de una proporción, se consideran solamente dos posibles resultados: 0 y 1, NO/SI, presencia/ausencia, o bien “pedido correcto”/“pedido con problemas de surtido”, o bien cualquier resultado de una variable dicotómica. Sin embargo, cabe decir que también es posible contemplar más de dos resultados posibles, como sucede frecuentemente en una encuesta, donde las posibles respuestas a una pregunta tienen varias opciones, como puede ser un “sí/no/no sé”, “Muy de acuerdo/De acuerdo/ni en acuerdo ni en desacuerdo/En desacuerdo/Muy en desacuerdo”, “PRI/PAN/PRD/Otro/No sé”, etc. Este tipo de proporciones, es decir las de más de dos clases conducen a un tratamiento más complejo y que no se aplica a nuestro caso de interés. No obstante, un tratamiento profundo a este respecto es abordado por autores como William G. Cochran⁵.

En un caso como el de este trabajo, tenemos dos posibilidades: “pedido correcto/pedido incorrecto” (0,1). Hablamos entonces de un ensayo Bernoulli, es decir, una prueba que solamente acepta dos posibles resultados: 0,1 o bien un “sí” o un “no”.

Podemos definir entonces como una proporción lo siguiente:

$$Y = \sum_1^N y_i = A$$
$$\bar{Y} = \frac{\sum_1^N y_i}{N} = \frac{A}{N} = P$$

En la misma forma para la muestra :

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{a}{n} = p$$

Donde:

N es el número total de elementos en el mismo universo.

α : Es la cantidad de "unos" en la muestra, es decir, la suma de los elementos que contienen la característica de interés, denotados por 1.

n : Es la cantidad de elementos muestreados.

y_i : Es el i -ésimo valor del elemento muestreado, de un total de n . Tal que $y_i = \begin{cases} 1 \\ 0 \end{cases}$

Es importante recordar que la proporción en la muestra $p=(\alpha/n)$ es una estimación insesgada de la población $P=A/N$.

A es la cantidad de "unos" en el universo.

Visto desde nuestro problema en particular, lo que haríamos simplemente será especificar la variable como "pedido incorrectamente surtido" $=1$, "pedido correctamente surtido" $=0$, esto es, definimos una variable y_i tal que: $y_i = \begin{cases} 1 & \text{Si el pedido fue incorrectamente surtido} \\ 0 & \text{Si el pedido fue correctamente surtido} \end{cases}$

Luego entonces A será el total verdadero de pedidos surtidos erróneamente en el día (puesto que son los únicos que suman). La proporción resultante de dividir entre A y el total de pedidos N será la verdadera proporción P que nos interesa estimar (puesto que no es posible conocerla). Algo similar sucedería con el procedimiento muestral, donde a es el número de pedidos detectados como mal surtidos en la inspección realizada ese día, que dividido entre el número de revisiones n , nos dará como resultado la proporción p de la muestra de pedidos mal surtidos.

Después de lo anterior, podemos establecer el estimador que emplearíamos bajo el MIA, así como su varianza.

⁵ William G. Cochran, "Técnicas de Muestreo", pp. 90-95.

2.6 ESTIMADOR DE LA PROPORCION POBLACIONAL BAJO MUESTREO IRRESTRICTO ALEATORIO.

Se pudo ver que el estimador adecuado para la proporción poblacional P , usando muestreo irrestricto aleatorio, es $p = \frac{a}{n}$

En cuanto a la varianza de p bajo el supuesto del MIA tenemos:

$$V(p) = \left(1 - \frac{n}{N}\right) \frac{S^2}{n} \text{ expresión derivada de } V(y)$$

Siendo $S^2 = \frac{NPQ}{N-1}$ y al sustituir esta expresión en $V(p)$, tenemos:

$$\begin{aligned} V(p) &= \left(1 - \frac{n}{N}\right) \frac{NPQ}{(N-1)n} \\ &= \frac{N-n}{N} \frac{NPQ}{(N-1)n} \\ &= \frac{N-n}{N-1} \left(\frac{PQ}{n}\right) \end{aligned}$$

Ahora, usaremos el estimador p para estimar la varianza muestral:

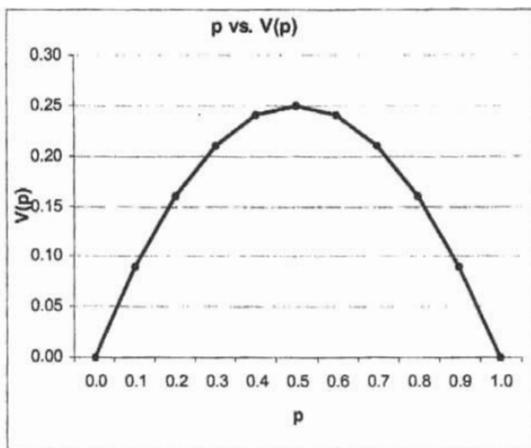
Sea:

$$s^2 = \frac{npq}{n-1} \quad \text{donde } q = 1 - p$$

Sustituyendo s^2 en $\hat{V}(p) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$ tenemos entonces :

$$\hat{V}(p) = \frac{N-n}{n-1} \frac{pq}{N}$$

Un aspecto muy relevante respecto a la varianza en una proporción, es el hecho de que la máxima varianza se da cuando $p=0.5$, la siguiente gráfica ilustra lo anterior:



Lo que significa, que la máxima varianza que puede existir para una proporción, es igual a 0.25 , es decir cuando $p = \frac{1}{2}$. Muchos autores abordan este hecho en sus textos⁶, donde se puede demostrar mediante técnicas de maximización con cálculo diferencial, que la máxima varianza se alcanza cuando $p = \frac{1}{2}$.

Este resultado es muy importante para la siguiente sección.

2.6.1 LA APROXIMACIÓN NORMAL PARA EL CASO DE p

El interés específico en nuestro caso, como hemos visto conforme se va desarrollando este trabajo, es conseguir alguna forma de mantener bajo control el proceso de surtido de pedidos. La intención entonces es, como se planteó en el punto 2.2, establecer un método de Control Estadístico de Proceso en el que, mediante un indicador pueda conocerse en todo momento el nivel de calidad del surtido y tomar decisiones oportunas. Lo anterior es posible a través de construir un intervalo de confianza para la proporción de pedidos mal surtidos que llamaremos p .

⁶ Ignacio Méndez Ramírez, "Conceptos muy elementales del muestreo con énfasis en la determinación práctica del tamaño de la muestra", p. 10

Construir un intervalo de confianza para una proporción p se facilita enormemente es el caso de que sea posible suponer normalidad en la distribución de p . Esto nos lleva a un tema discutido en diversos textos, la aproximación normal para una proporción p según algunos autores, es posible suponerla cuando tenemos un tamaño de muestra n ‘grande’ y una proporción p ‘relativamente pequeña’.

Para aclarar el concepto de n grande y p relativamente pequeña, Cochran y otros autores⁷ proporcionan una relación de ambos valores, con que es posible suponer normalidad para p :

p	np (núm. observado en la clase más pequeña)	Tamaño de muestra n
0.5	15	30
0.4	20	50
0.3	24	80
0.2	40	200
0.1	60	600
0.05	70	1400
$\sim 0^*$	80	∞

* Significa que p es extremadamente pequeña, de tal manera que np sigue la distribución Poisson.

Puede verse que, conforme disminuye la proporción p , podemos suponer normalidad si hacemos crecer el tamaño de muestra n y viceversa, si se emplea un tamaño menor, deberá estarse dispuesto a emplear una varianza mayor. Se estableció en el punto anterior que la máxima varianza que puede darse para una proporción p , es $p(1-p)=0.25$, con $p = \frac{1}{2}$ esto quiere decir que para la máxima varianza, el mínimo tamaño de muestra n necesario para suponer normalidad es $n=30$, dato conocido desde la estadística elemental.

Regresando al planteamiento original, si suponemos normalidad en la proporción p (cumpliendo la recomendación de la tabla anterior) entonces, nuevamente los mismos autores establecen que puede construirse un intervalo de confianza para la proporción p , con la siguiente expresión:

⁷ cfr. William G. Cochran, “Técnicas de Muestreo”, p. 88, Erwin Kreyszig, “Introducción a la Estadística Matemática”, pp. 137-138

$$p \pm \left[z_{(1-\alpha/2)} \sqrt{1-f} \sqrt{\frac{pq}{n-1}} + \frac{1}{2n} \right]$$

Donde $f = \frac{n}{N}$, $z_{(1-\alpha/2)}$ es un valor crítico a partir de la distribución normal que determina el nivel de confianza seleccionado, $\frac{1}{2n}$ es un factor de corrección por continuidad.

2.6.2 ESTIMACION DEL TAMAÑO DE MUESTRA PARA PROPORCIONES.

Uno de los aspectos más importantes es determinar el número de revisiones que se harán en cada jornada de inspección, es decir, determinar el tamaño de muestra n . La siguiente expresión nos permite determinar n , suponiendo un esquema de muestreo irrestricto aleatorio:

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

$$\text{Donde } n_0 = \frac{z_{(1-\alpha/2)}^2 pq}{d^2}$$

d es la precisión deseada y seleccionada previamente
 z es un valor crítico a partir de la distribución normal que determina el nivel de confianza seleccionado.

2.7 MUESTREO ALEATORIO ESTRATIFICADO.

El método de muestreo aleatorio estratificado (llamémoslo MAE) consiste en separar la población en dos o más grupos para obtener una muestra independiente en cada una de estas separaciones que llamaremos *estratos*. Posteriormente se harán las inferencias que sean de interés a través de los estimadores correspondientes (media, proporción, etc.) de manera que se obtendrá estimaciones para cada estrato y para la población en conjunto.

Separar en grupos más pequeños a la población puede parecer sencillo, sin embargo es importante considerar que para hacerlo debemos procurar hacer los grupos con elementos lo más homogéneos posibles, lo que quiere decir que éstos presentarían un comportamiento y características muy similares entre sí (respecto a las propiedades que se estén investigando), que reportará un beneficio muy importante al diseño.

Existen buenas razones para emplear el MAE, las ventajas que reporta se describen a continuación⁸:

- a) El construir estratos para la población tiende a reducir la variabilidad de los estimadores. Por ejemplo, si se quisiera levantar una encuesta por muestreo en tres ciudades, lo natural sería considerar como un estrato a cada ciudad, donde se esperaría que los elementos muestreados dentro de cada ciudad, presentaran características homogéneas (creencias religiosas, posiciones políticas, costumbres sociales, etc.) que permiten que la variación de los estimadores sea menor que si se muestreara las tres ciudades en un mismo marco.
- b) La disponibilidad de marcos puede ser más tratable. Por ejemplo, podemos considerar en un estrato, los elementos que tengan un marco de buena disponibilidad, como puede ser en una ciudad, donde se dispone de planos catastrales. Mientras que en otro estrato pueden incluirse conjuntos que no tengan la misma disponibilidad, como en un área rural, donde solamente sea posible usar fotografía aérea.
- c) Los costos inherentes al muestreo suelen ser menores al usar estratos ya que permite una mejor planeación de las actividades de campo (en el caso particular de una encuesta) y de otros quehaceres en el levantamiento y operación.

⁸ Ignacio Méndez Ramírez, "Conceptos muy elementales del muestreo con énfasis en la determinación práctica del tamaño de la muestra", pp. 11-13

Es posible usar diferentes formas de muestreo dentro de cada estrato, pero es importante decir que el caso más general es aplicar muestreo irrestricto aleatorio (MIA)⁹, para el caso de este trabajo igualmente es el que se tomará en cuenta.

Consideremos ahora la siguiente notación para el MAE, recordando que el parámetro de interés para este trabajo es una proporción:

N_h = Número de unidades en el estrato h - ésimo

$$h = 1, 2, \dots, L$$

L = Número de estratos

$$N = \sum_{h=1}^L N_h \quad \text{Total de unidades en la población}$$

n_i es el tamaño de muestra asignado al estrato i

\hat{p}_i es la proporción encontrada de la característica medida en el estrato i

es decir $\hat{p}_i = \frac{a_i}{n_i}$, para a_i que es la suma de elementos que poseen la característica medida en el estrato i .

$$\hat{q}_i = 1 - \hat{p}_i$$

Los elementos que se acaban de describir son, como en el caso del MIA, fácilmente identificables. El punto importante ahora sería precisamente definir el criterio a emplearse en la conformación de los estratos, es decir, cuántos y cuáles estratos se deben construir.

⁹ Ignacio Méndez Ramírez, "Conceptos muy elementales del muestreo con énfasis en la determinación práctica del tamaño de la muestra", p. 13

2.7.1 LA CONSTRUCCIÓN DE ESTRATOS EN EL MUESTREO ALEATORIO ESTRATIFICADO.

Existen métodos o reglas que establecen bases para construir estratos en el MAE, de manera que posean características estadísticamente deseables, por ejemplo la regla de Dalenius-Hodges¹⁰, sin embargo, es quizá más importante atender las recomendaciones generales que los textos de muestreo ofrecen al respecto¹¹, es decir, aspectos que debemos tomar en cuenta si es que deseamos tener estratos adecuados y funcionales, lo más importante para que los estratos tengan un buen desempeño, es formarlos en función de la homogeneidad observada entre los elementos que integren un estrato en particular, dicho en otras palabras, debemos pensar en variables con una alta correlación con los aspectos que se estén estudiando, por ejemplo, opiniones políticas similares, mismo nivel socioeconómico, ubicación geográfica, tamaños en la población, etc.

Justamente atendiendo a esas recomendaciones, se piensa que la mejor definición de estratos para el caso que ocupa al presente trabajo, sería considerar a cada línea de surtido como un estrato, de manera que tendríamos un número total de estratos $L=4$. La razón es que, los empleados encargados del surtido en cada línea, poseen un alto nivel de interacción con sus compañeros de la misma línea, tienen el mismo horario de trabajo, el conocimiento mutuo es extenso, fuerte sentido de pertenencia al grupo (a la línea de surtido en la que trabaja), etc. Con lo que se puede asumir cooperación y espíritu de equipo, pero al mismo tiempo una tendencia a cometer errores parecidos bajo circunstancias similares.

No obstante, el MAE para el caso de interés presenta algunas limitantes de orden logístico, que se abordarán en el capítulo 3.

¹⁰ William G. Cochran, "Técnicas de Muestreo", pp. 170-171

¹¹ cfr. William G. Cochran, "Técnicas de Muestreo", p. 169-177, Ignacio Méndez Ramírez, "Conceptos muy elementales del muestreo con énfasis en la determinación práctica del tamaño de la muestra", pp. 11-13

2.7.2 ESTIMADOR DE LA PROPORCION POBLACIONAL BAJO MUESTREO ALEATORIO ESTRATIFICADO.

Los estimadores a usarse, en caso de emplear un esquema de MAE, serán los siguientes:

Estimador de la proporción poblacional p_{st} :

$$\hat{p}_{st} = \frac{1}{N} \sum_{i=1}^L N_i \hat{p}_i$$

Varianza estimada de \hat{p}_{st} :

$$\begin{aligned} \hat{V}(\hat{p}_{st}) &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(\frac{N_i - n_i}{N_i} \right) \left(\frac{\hat{p}_i \hat{q}_i}{n_i - 1} \right) \\ &= \frac{1}{N^2} \sum_{i=1}^L N_i^2 \hat{V}(\hat{p}_i) \end{aligned}$$

Donde N_i es el universo en el estrato i

n_i es el tamaño de muestra asignado al estrato i

\hat{p}_i es la proporción encontrada de la característica medida en el estrato i

es decir $\hat{p}_i = \frac{a_i}{n_i}$, para a_i que es la suma de elementos que poseen la característica

medida en el estrato i .

$$\hat{q}_i = 1 - \hat{p}_i$$

Como en el caso del MIA, asumiendo normalidad es posible construir un intervalo de confianza para la proporción p_{st} :

$$\hat{p}_{st} \pm \left[z_{(1-\alpha/2)} \sqrt{\frac{1}{N^2} \sum_{i=1}^L N_i^2 \left(\frac{N_i - n_i}{N_i} \right) \left(\frac{\hat{p}_i \hat{q}_i}{n_i - 1} \right)} \right]$$

Es decir :

$$\hat{p}_{st} \pm \left[z_{(1-\alpha/2)} \sqrt{\hat{V}(\hat{p}_{st})} \right]$$

Donde $z_{(1-\alpha/2)}$ es un valor crítico a partir de la distribución normal que define el nivel de confianza seleccionado.

2.7.3 DETERMINACIÓN DEL TAMAÑO DE MUESTRA Y AFIJACIÓN DE LA MUESTRA.

Existe un punto muy importante al considerar el esquema del MAE y se refiere a la determinación del tamaño de muestra, sin embargo, el tema se hace especialmente importante cuando se plantea también, el tener que definir un criterio para distribuir ese tamaño de muestra en los diferentes estratos que se hayan definido previamente.

Por esa razón, es importante revisar nuevamente la teoría del muestreo. Aunque los autores no necesariamente coinciden, generalmente se manejan cuatro criterios de asignación o afijación de la muestra, que a continuación se describen muy brevemente.

2.7.3.1 AFIJACIÓN DE IGUAL NÚMERO EN CADA ESTRATO.

Esta forma, poco utilizada en la práctica, consiste simplemente en dividir el tamaño de muestra total al número de estratos definidos, es decir: $n_h = \frac{n}{L}$ donde n es el tamaño total de muestra, n_h es el estrato h-ésimo de un total de L estratos.

Quizá su única ventaja sea lo sencillo de su aplicación, no obstante, no posee una buena eficiencia en la realidad, menos aún si el costo de obtener información en cada estrato es diferente o si el tamaño del estrato h-ésimo N_h de cada estrato es muy variable.

La varianza estimada en este criterio de afijación es:

$$V(\hat{p}_{st}) = \frac{L}{N^2} \sum_{h=1}^L \frac{N_h^2 S_h^2}{n} - \frac{1}{N^2} \sum_{h=1}^L N_h S_h^2$$

$$\text{Donde } S_h^2 = P_h Q_h$$

La expresión anterior supone que se conoce un tamaño de muestra n , sin embargo, en caso de no conocerse, la siguiente expresión determina el tamaño de muestra total, bajo el esquema de MAE, usando afijación de igual número:

$$n = \frac{L \sum_{h=1}^L N_h^2 S_h^2}{N^2 V(\hat{p}_{st}) + \sum_{h=1}^L N_h S_h^2}$$

Donde $S_h^2 = P_h Q_h$

Donde la varianza de \hat{p}_{st} puede denominarse varianza deseada. Es decir, a libre arbitrio del diseñador.

2.7.3.2 AFIJACIÓN PROPORCIONAL AL TAMAÑO.

Este método basa su criterio en la medición previa del tamaño de cada estrato, para que en función de ese tamaño se asigne determinada cantidad de muestra.

Por lo tanto, tenemos la siguiente expresión:

$$n_h = \frac{N_h}{N} n$$

Esto es, se establece que el tamaño de muestra en el estrato h quedará en función del tamaño de la población contenida en el mismo estrato h , calculado éste a través de la razón de la población del estrato h entre la población total. Lo anterior toma en cuenta que conocemos el tamaño de muestra total n . La varianza correspondiente se define a continuación:

$$V(\hat{p}_{st}) = \frac{1}{N} \sum_{h=1}^L \frac{N_h S_h^2}{n} - \frac{1}{N^2} \sum_{h=1}^L N_h S_h^2$$

Donde $S_h^2 = P_h Q_h$ recordando que se está estimando una proporción para cada estrato h .

Ahora bien, para usar lo anterior es necesario conocer el tamaño de muestra total n , pero, como se explicó en el caso de la afijación de igual número en cada estrato, si no se conoce el tamaño n , la siguiente expresión es una buena estimación para dicho parámetro:

$$n = \frac{N \sum_{h=1}^L N_h S_h^2}{N^2 V(p_{st}) + \sum_{h=1}^L N_h S_h^2}$$

Donde $S_h^2 = P_h Q_h$ por estimar una proporción $V(p_{st})$ es la varianza deseada, es decir, fijada arbitrariamente.

2.7.3.3 AFIJACIÓN ÓPTIMA.

Para el caso de la afijación óptima y de la siguiente (la Afijación de Neyman) es necesario considerar un nuevo concepto: el costo por levantar información.

Este costo está asociado al gasto de recursos que se debe hacer para obtener la información al aplicar el esquema de muestreo escogido, por ejemplo, si se realiza una encuesta por muestreo, los recursos que se empleen al aplicar los cuestionarios (pago a encuestadores, transportación, viáticos, papelería, etc.) forman parte del costo.

Dichos costos representan una de las preocupaciones no solamente de la logística de recolección de datos por muestreo, sino que la misma teoría del muestreo ha abordado este problema en la afijación óptima y la de Neyman.

Desde el punto de vista matemático, la afijación óptima se convierte en un problema de optimización (de ahí se deriva el nombre), donde se pretende minimizar la varianza del estimador, sujeta (restricción) a una función de costos. La literatura generalmente se refiere a una función de costos lineal, que puede ser de la siguiente manera:

$C = C_0 + \sum_{h=1}^L C_h n_h$ donde C_0 es un costo fijo inicial y C_h son los costos asociados al estrato

h .

A través del empleo de técnicas de optimización como los Multiplicadores de Lagrange, se obtiene entonces, una expresión que minimiza la varianza sujeta a la función de costos mencionada. La asignación para cada estrato, resultante del proceso de optimización se plantea a continuación:

$$n_h = \frac{N_h S_h / \sqrt{C_h}}{\sum_{h=1}^L (N_h S_h / \sqrt{C_h})} n$$

Donde nuevamente $S_h = \sqrt{P_h Q_h}$

C_h es el costo asociado al estrato h

N_h es el tamaño del estrato h .

Donde la varianza del estimador de la proporción con afijación óptima será el siguiente:

$$V(p_{st}) = \frac{1}{N^2} \left(\sum_{h=1}^L \frac{N_h S_h \sqrt{C_h}}{n} \right) \left(\sum_{h=1}^L \frac{N_h S_h}{\sqrt{C_h}} \right) - \frac{1}{N^2} \sum_{h=1}^L N_h S_h^2$$

Donde $S_h = \sqrt{P_h Q_h}$

Como en los casos anteriores, se asume que n es conocido, sin embargo si no se conoce, es decir, que se pretende determinar, la siguiente expresión nos da n en el caso de afijación óptima:

$$n = \frac{\left(\sum_{h=1}^L N_h S_h \sqrt{C_h} \right) \left(\sum_{h=1}^L N_h S_h / \sqrt{C_h} \right)}{\left(V(p_{st}) N^2 + \sum_{h=1}^L N_h S_h^2 \right)}$$

Donde $S_h = \sqrt{P_h Q_h}$

C_h es el costo asociado al estrato h

$V(p_{st})$ es el nivel de varianza deseado, es decir determinado arbitrariamente.

Desde el punto de vista estadístico, asumiendo que realizamos un trabajo de asignación óptima adecuado, o como dice el propio Cochran, una “estratificación inteligente”, se obtiene generalmente una varianza más pequeña que en el caso de la afijación proporcional o del MIA, con lo que se tiene una ventaja inherente por el uso de la afijación óptima¹².

Sin embargo, de manera práctica es también importante considerar las ventajas de la afijación óptima. El hecho de tomar en cuenta los costos inherentes a la recolección de la información es crucial en muchos casos porque aunque se tenga un esquema de muestreo ideal, si los costos se elevaran demasiado podría ser catastrófico para el levantamiento. Por tal razón, la afijación óptima cobra relevancia, el control de los costos en cada estrato es siempre un tema a tomarse muy en cuenta, o por lo menos debe evaluarse si no causarían problemas.

2.7.3.4 AFIJACIÓN DE NEYMAN.

La afijación de Neyman representa un caso particular de la afijación óptima, ya que, para este caso se considera que los costos en todos y cada uno de los estratos son exactamente iguales, es decir que $C_h=C$ para toda h .

Por esta razón, las expresiones para calcular las asignaciones por estrato, la varianza y el tamaño de muestra total se simplifican de manera significativa.

Para las asignaciones por estrato tenemos:

$$n_h = \frac{N_h S_h}{\sum_{h=1}^L N_h S_h} n$$

Donde $S_h = \sqrt{P_h Q_h}$
con n conocida o determinada
arbitrariamente.

¹² William G. Cochran, “Técnicas de Muestreo”, pp. 136-138

Para la varianza del estimador de proporciones tenemos:

$$V(\hat{p}_{st}) = \frac{1}{N^2} \frac{\left(\sum_{h=1}^L N_h S_h \right)^2}{n} - \frac{1}{N^2} \sum_{h=1}^L N_h S_h^2$$

Donde $S_h = \sqrt{P_h Q_h}$

Como en los casos anteriores, en caso de pretender determinar el tamaño de muestra total n , bajo afijación de Neyman, la expresión es la siguiente:

$$n = \frac{\left(\sum_{h=1}^L N_h S_h \right)^2}{N^2 V(\hat{p}_{st}) + \sum_{h=1}^L N_h S_h^2}$$

Donde $S_h = \sqrt{P_h Q_h}$

$V(\hat{p}_{st})$ es el nivel de varianza deseado, es decir determinado arbitrariamente.

La afijación de Neyman es sumamente útil en muchos casos, no solamente para aquellos donde el costo por estrato es el mismo sino también en donde se desconocen los costos, por lo que se asumen como iguales.

De hecho, para el ejemplo que justamente ocupa a este trabajo, en caso (como se planteó desde el principio del MAE) de que se optara por emplear el Muestreo Aleatorio Estratificado, al considerarse los costos por estrato, tomando en cuenta que un estrato sería una línea de surtido, se asume que el costo por estrato es exactamente el mismo en cada uno ya que el número de trabajadores y de cajas surtidas es el mismo en cada una de las cuatro líneas de surtido. Esto llevaría a usar la afijación de Neyman, con las ventajas ya descritas en la afijación óptima (que se extiende a la de Neyman por ser un caso particular de ésta) y con la simplicidad de este esquema.

2.8 MUESTREO POR CONGLOMERADOS.

El Muestreo por Conglomerados (lo llamaremos MC) es útil en muchas situaciones, específicamente cuando no se dispone de un marco muestral de unidades elementales de observación. Es útil también en otras limitaciones de tipo pragmático, pero muy particularmente en los aspectos de índole económica.

La utilidad del MC se hace evidente cuando se deben muestrear sitios de agrupamiento natural, como pueden ser las escuelas de una ciudad, hospitales, cárceles, etc.

Los conglomerados se definen como grupos que contienen las unidades a muestrearse, por lo que éstas permanecen concentradas en conglomerados, de allí el nombre, o como dice William Mendenhall: *“Una muestra por conglomerados es una muestra aleatoria en la cual cada unidad de muestreo es una colección, o conglomerado, de elementos”*.¹³

La idea del MC consiste en obtener primero una muestra de conglomerados, para que, en una segunda etapa efectuar otro muestreo o bien un censo, de los elementos contenidos en los conglomerados seleccionados previamente. Resumiendo, el MC se desarrolla en dos etapas:

- 1) Selección de conglomerados a través de una muestra
- 2) Muestreo o Censo de los elementos contenidos en los conglomerados que conforman la muestra.

No se especifica claramente en la literatura técnica si la primera muestra (es decir la de los conglomerados) debe ser construida mediante algún esquema en particular, por lo que para este trabajo, se optaría por emplear MIA en la selección de los conglomerados.

Debe considerarse ahora desde un punto de vista logístico este esquema para el caso de los pedidos de la compañía. Dado que la manera en que podría idearse la conformación de los conglomerados sería pensando, por ejemplo, en las zonas geográficas que la compañía definió, y que, como se recordará del capítulo I, la zonas se procesan (se surten) completas en la misma línea y en el mismo periodo de tiempo, es decir, no se continúa con otra zona

¹³ William Mendenhall, “Elementos de Muestreo”, p. 196

sino hasta terminar con la actual, por ello, los conglomerados vistos de manera natural serían precisamente dichas zonas.

El tamaño de conglomerado para este caso es de cien cajas de pedido en promedio. Según el interés de los ingenieros del área de Calidad, lo que procedería sería a hacer un censo dentro de cada conglomerado (no un muestreo de elementos dentro de cada conglomerado), por lo que, en cada conglomerado seleccionado se haría una revisión de los cien pedidos que aproximadamente tiene cada zona. De la misma manera que funciona el esquema que la misma área de Calidad implantó al quedarse solamente con ocho inspectores de pedidos. Conviene también recordar que, precisamente por la restricción del actual número de inspectores, el área de Aseguramiento de Calidad no podría revisar más allá de cuatro zonas (cuatro conglomerados, censando los 100 pedidos en cada uno) por jornada. Con lo que, el tamaño n de muestra sería igual a cuatro ($n=4$). A continuación se especifica la notación para este esquema de muestreo:

N = Número de conglomerados en la población

n = Número de conglomerados seleccionados en una muestra irrestricta aleatoria

m_i = Número de elementos en el conglomerado i , $i=1, \dots, N$

$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i$ = tamaño promedio del conglomerado en la muestra

$M = \sum_{i=1}^N m_i$ = número de elementos en la población

$\bar{M} = \frac{M}{N}$ = tamaño promedio del conglomerado en la población

a_i = total de elementos en el i -ésimo conglomerado que poseen la característica de interés

por lo que $a_i = \begin{cases} 0 & \text{Si el pedido es correctamente surtido} \\ 1 & \text{Si el pedido es incorrectamente surtido} \end{cases}$

El estimador de la proporción poblacional está dada por :

$$\hat{p} = \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n m_i}$$

De la misma manera, la varianza estimada de la proporción p será :

$$\hat{V}(\hat{p}) = \left(\frac{N-n}{Nn\bar{M}^2} \right) \frac{\sum_{i=1}^n (a_i - \hat{p}m_i)^2}{n-1}$$

Ahora bien, a efecto de simplificar los términos de los estimadores, es importante relacionar los datos que sí conocemos (en nuestro caso práctico) con dichas expresiones, esto es, en caso de que optáramos por emplear este esquema muestral, los estimadores de proporción y de varianza serían los siguientes:

$$M = 12500$$

$$n = 4$$

$$m_i = \bar{m} = \bar{M} = 100 \text{ (recordar que se efectuaría un censo dentro de cada conglomerado / zona).}$$

$$N = \frac{M}{\bar{M}} = \frac{12500}{100} = 125$$

$$\sum_{i=1}^4 m_i \approx 400 \Rightarrow \hat{p} = \frac{\sum_{i=1}^4 a_i}{400}$$

$$\hat{V}(\hat{p}) = \left(\frac{125 - 4}{(125)(4)(100^2)} \right) \left(\frac{\sum_{i=1}^4 (a_i - 100\hat{p})^2}{3} \right) = (2.42E - 05) \left(\frac{\sum_{i=1}^4 (a_i - 100\hat{p})^2}{3} \right)$$

Desde luego, como pasa con las anteriores técnicas de muestreo, si construimos un intervalo de confianza bajo la misma óptica ($\hat{p} \pm z \cdot \sqrt{\hat{V}(\hat{p})}$) se debe suponer que se tienen buenos estimadores empleando una muestra $n \geq 30$ como vimos anteriormente, sin embargo, este no es el caso puesto que $n=4$. Con lo que la factibilidad de emplear este esquema, es muy baja.

2.9 MUESTREO SISTEMÁTICO.

Este esquema de muestreo (que llamaremos en lo sucesivo MS), representa una de las mejores opciones en cuanto a control de procesos en serie, debido a su naturaleza que lo hace muy dinámico y fácil de aplicar, es importante, sin embargo, señalar algunas puntos relevantes para su correcta y adecuada aplicación.

Su metodología consiste en escoger una de cada k unidades (cada k -ésima), repartidas regularmente a lo largo de todo el universo, de tal manera que se hace un recorrido en el mismo de manera uniforme. Visto de otra forma, es análogo a conformar n estratos, escogiendo una unidad en cada uno de ellos, de manera que $k=N/n$. En este sentido, lo ideal es evidentemente que N/n fuera un número entero (N múltiplo de k), sin embargo no es lo más frecuente, pese a lo anterior, algunos autores¹⁴ sostienen que puede suponerse que probablemente son despreciables si $n>50$ las perturbaciones provocadas por el hecho de no obtener los “estratos” (cada k -ésimo visto como un estrato nuevo) del mismo tamaño.

2.9.1. PROCEDIMIENTO INSEGADO DE SELECCIÓN Y ARRANQUE ALEATORIO.

En principio suponemos que la muestra es un múltiplo de la población N . Esto es, $N=kn$. Llamaremos a k , constante de proporcionalidad. Si dicha proporcionalidad se cumple estrictamente, o sea que k es un entero, entonces podemos seleccionar un número aleatorio en el intervalo $[1,k]$ y escoger un elemento A de éste, con lo que, el siguiente elemento será $A+k$ y así, sucesivamente hasta el elemento $A+(n-1)k$.

Para aclarar lo anterior, pensemos por ejemplo en una población de $N=12$ elementos, esto es y_1, y_2, \dots, y_{12} . Si tomamos una muestra $n=4$, tenemos que $k=3$.

¹⁴ William G. Cochran, “Técnicas de Muestreo”, p. 258

Considerando un número A (que llamaremos 'de arranque') el cual esté dentro del intervalo $[1,3]$ tenemos 3 posibles muestras sistemáticas:

Muestra 1	Muestra 2	Muestra 3
y_1	y_2	y_3
y_4	y_5	y_6
y_7	y_8	y_9
y_{10}	y_{11}	y_{12}

Cada una de las muestras tiene una probabilidad $\frac{1}{k}$ de ser escogida.

Sin embargo, si ahora, en lugar de pensar en $k=3$ usamos 5 muestras, esto es que $N \neq nk$, de tal manera que tendríamos:

Muestra 1	Muestra 2	Muestra 3	Muestra 4	Muestra 5
y_1	y_2	y_3	y_4	y_5
y_6	y_7	y_8	y_9	y_{10}
y_{11}	y_{12}			

Puede verse ahora que no se obtiene el mismo número de elementos por cada una de las cinco muestras, con lo que no se cumple entonces el supuesto de proporcionalidad, es decir que k no sería entero y por lo tanto, la probabilidad de selección de los elementos no es igual (como sí sucede en el caso anterior). Lo anterior conlleva un problema mayor: que el estimador (el valor esperado) resultante no es insesgado.

Por lo anterior, se debe considerar en este caso, un procedimiento que conduzca a un estimador insesgado. Lo anterior se puede lograr generando números aleatorios como referencia o puntos de arranque, de manera que las probabilidades de selección se balancean, es decir, en función de su correspondiente número de elementos integrantes, será proporcionalmente mayor o menor la probabilidad de selección.

Para ilustrar lo anterior, supongamos que en el caso de $k=5$, de donde tomaremos un elemento en cada muestra, es decir, uno de cada cinco, si seleccionamos la j -ésima unidad por ejemplo $j=8$ tenemos que: $\frac{j}{k} = \frac{8}{5} = 1$ con residuo $r=3$ por lo tanto r puede tomar

residuos en 0, 1, 2, 3 o 4. Tomaremos ahora estos residuos como números aleatorios de arranque, es decir, si $r=1$ escogeremos y_1 , si $r=2$ la selección sería y_2 y así sucesivamente. Si $r=0$ entonces escogemos y_5 como punto de arranque.

Ahora bien, la probabilidad de seleccionar y_2 , y_7 , o y_{12} es $1/12$ para cada uno de los elementos. Por ello la probabilidad de seleccionar esta muestra es la suma de las probabilidades, esto es $3/12$. En pocas palabras, hay tres elementos, la probabilidad es $3/12$. Igualmente para la muestra 4 que contiene los elementos y_4 , y_9 la probabilidad es de $2/12$.

En resumen, el ‘arranque’ aleatorio que demos al seleccionar la muestra es esencial en el diseño de un MS.

En función de lo anterior, existe un método sugerido por Lahiri¹⁵ el cual basa la solución en algo similar, considerando el universo como si fuera un círculo o cola circular, de manera que el último elemento en el universo puede considerarse como antecesor del primer elemento, por lo que todos los elementos tienen la misma probabilidad de ser seleccionados.

2.9.2. TIPOS DE POBLACIÓN.

Una de las consideraciones más importantes a realizar cuando se emplea el MS es el tipo de población donde se aplicará. Esto último resulta de mucho peso a la hora de tomar en cuenta particularmente esta técnica, los tipos de poblaciones con que podemos encontrarnos son las siguientes¹⁶:

2.9.2.1. Población Aleatoria. Sus elementos están ordenados al azar, esto significa que no encontraremos tendencia alguna o correlación entre los elementos de la población, en este sentido, se puede afirmar que la aplicación del MS es equivalente a aplicar MIA.

¹⁵ William G. Cochran, “Técnicas de Muestreo”, p. 258

¹⁶ William Mendenhall, “Elementos de Muestreo”, pp. 173-174

- 2.9.2.2. Población con Tendencia Lineal:** Sus elementos están ordenados en magnitud de acuerdo con algún esquema o patrón, ascendente o bien descendente de modo continuo. Podemos ver que en este caso, la aplicación del MS queda en ventaja generalmente con respecto al MIA. Pensemos, por ejemplo, en que en cada segmento de k elementos seleccionamos uno por lo que reflejamos (del mismo modo que lo hace el MAE) diferentes segmentos en dicha población, con el MIA se corre el riesgo de no reflejar toda la población. Independientemente de lo anterior, existen métodos para mejorar la estimación de poblaciones lineales con MS, por ejemplo a través de usar muestras centralmente localizadas, (tomando como muestra el elemento central de cada segmento de k elementos).
- 2.9.2.3. Población Periódica:** Sus elementos tienen una variación cíclica, es decir con altibajos regulares, (pensemos por ejemplo en una curva sinusoidal simple) en este caso la efectividad del muestreo sistemático dependerá fundamentalmente del valor de k . Tomando un valor “demasiado grande” corremos el riesgo de captar sólo ciertas variaciones, interpretando así de manera parcial un parámetro estimado. Por ejemplo, si pensáramos en muestrear las ventas de la empresa sólo los días jueves de cada 2 semanas (cuando arranca formalmente una campaña o nuevo periodo) en que las ventas comienzan “lentas” es decir muy bajas en comparación con el resto de la campaña, corremos el riesgo de interpretar como “baja” toda la venta de la campaña cuando es evidente (para los que conocen la empresa) que no será así. Para estos casos, es recomendable aumentar el tamaño de muestra (reducir k), de forma tal que sea posible “cubrir”, es decir, detectar las variaciones de la población. Otra recomendación es emplear números aleatorios de arranque distintos para cada proceso de selección diferente, esto especialmente para procesos donde sea necesario efectuar muestreos de manera continua y normal.¹⁷

La conformación de los elementos en el universo es importante ya que en función de ésta, la varianza obtenida como resultado de un MS puede ser mayor o menor comparativamente con el MIA, incluso el MAE. Muchos de estos supuestos están ampliamente abordados en

¹⁷ cfr. William G. Cochran, “Técnicas de Muestreo”, pp. 271-273, William Mendenhall, “Elementos de Muestreo”, pp. 173-175

distintos textos de muestreo.¹⁸ En un caso como el de las revisiones de cajas de pedidos, se hace una serie de razonamientos que ubican las cajas de pedido como una *probable* población periódica, *sin tener suficientes elementos para asegurarlo*, por lo que para usar MS es aconsejable emplear varios números aleatorios de arranque¹⁹, de esta manera se asegura que el proceso no tenga varianza mayor a la obtenida con MIA. Todas estas consideraciones se amplían en el capítulo 3.

Pese a que el MS es, en apariencia, muy diferente del MIA, en realidad resulta generalmente equivalente al mismo. Será importante tomar en cuenta cuál es la situación de la población donde se aplicará.

Una consideración interesante es la comparación o equivalencia que se establece en algunos textos con respecto al MS y al MC, donde podemos apreciar al MS como un caso particular del MC, o como se describe en el texto de Cochran²⁰: *“El muestreo sistemático viene a ser la elección de una sola unidad de muestreo compleja, que constituye la muestra total. Una muestra sistemática es una muestra aleatoria simple de una unidad conglomerada, tomada en una población de k unidades conglomeradas”*.

Lo anterior es importante tomarlo en cuenta por lo siguiente: si pensamos que una muestra obtenida de MS es un caso particular de MC, de tamaño 1, eso implica la imposibilidad de estimar la varianza de la media para dicha muestra. Esta consideración ha sido abordada por distintos autores, por ejemplo, Cochran entre otros, establece varias expresiones alternativas de estimación de la varianza. En resumen, con dichas expresiones se aborda la estimación a través de considerar el caso del MIA además de tomar en cuenta la condición de la población estudiada, es decir, si existe correlación entre los elementos de la población, por ejemplo.

Específicamente en nuestro caso, se usaría un estimador de varianza de MIA, que iría acorde a la experiencia de los ingenieros de calidad, en el sentido de que se podría afirmar que la naturaleza de la población no es correlacionada, o en forma resumida, encontramos un razonamiento similar con Mendenhall²¹, el cual plantea: *“Notemos que la varianza*

¹⁸ cf. William G. Cochran, “Técnicas de Muestreo”, pp. 266-278, William Mendenhall, “Elementos de Muestreo”, pp. 173-174

¹⁹ William Mendenhall, “Elementos de Muestreo”, pp. 174-175

²⁰ William G. Cochran, “Técnicas de Muestreo”, pp. 259-260.

²¹ William Mendenhall, “Elementos de Muestreo”, p. 179

estimada de p_{xy} es idéntica a la varianza estimada de p usando muestreo aleatorio simple. Este resultado no implica que las varianzas poblacionales correspondientes sean iguales; sin embargo, si N es grande y si las observaciones dentro de una muestra sistemática no están correlacionadas (esto es $\rho=0$), las dos varianzas poblacionales serán iguales”.

2.9.3. ESTIMADORES DE PROPORCION POBLACIONAL Y DE VARIANZA.

Como sucede con los estimadores de varianza y de proporción poblacional del MIA, los del MS cumplen con las características de un buen estimador (insesgado para el caso del estimador de proporciones, consistencia, suficiencia, etc.) esto se puede comprobar fácilmente en diversos textos de estadística matemática²². A continuación se exponen los estimadores generalmente empleados en el MS:

Estimador de la proporción poblacional p :

$$\hat{p}_{xy} = \frac{\sum_{i=1}^n a_i}{n}$$

Donde $a_i = \begin{cases} 0 & \text{Si el pedido es correctamente surtido} \\ 1 & \text{Si el pedido es incorrectamente surtido} \end{cases}$

Varianza estimada de \hat{p}_{xy} :

$$\hat{V}(\hat{p}_{xy}) = \frac{\hat{p}_{xy}\hat{q}_{xy}}{n-1} \left(\frac{N-n}{N} \right)$$

donde $\hat{q}_{xy} = 1 - \hat{p}_{xy}$

Estableciendo un intervalo de confianza para \hat{p}_{xy} :

$$\hat{p}_{xy} \pm z \sqrt{\hat{V}(\hat{p}_{xy})}$$

Cabe decir que la expresión anterior es la misma que la usada en la sección 2.6.1 por lo que, es posible simplificar el intervalo de confianza como sigue :

$$\hat{p}_{xy} \pm \left[z \sqrt{1-f} \sqrt{\frac{\hat{p}_{xy}\hat{q}_{xy}}{n-1} + \frac{1}{2n}} \right] \text{ donde } 1/2n \text{ es un factor de corrección por continuidad}$$

$f = n/N$

donde z asume los valores $z = 1.96, 2.58, 3.29$ en función del nivel de confianza deseado.

²² Erwin Kreyszig, "Introducción a la Estadística Matemática", pp. 175-181

Donde resulta evidente que n es el tamaño de muestra, es decir, el número de pedidos a revisar bajo MS, asimismo N es el número total de pedidos generados en el día, esto es, $N=12500$.

Por último, es importante recordar que para el MS, el tamaño de muestra n se determina como se expresa a continuación:

$$n = \frac{N\hat{p}_{sy}\hat{q}_{sy}}{(N-1)V(\hat{p}_{sy}) + \hat{p}_{sy}\hat{q}_{sy}}$$

Donde :

$$\hat{q}_{sy} = 1 - \hat{p}_{sy}$$

$V(\hat{p}_{sy})$ puede ser fijada arbitrariamente

2.10 OTROS ESQUEMAS DE MUESTREO.

Existen otras alternativas en planes de muestreo, que desde luego pueden ser tomadas en cuenta, pero que resultan variantes de los esquemas mostrados anteriormente, o bien resultan combinaciones de los mismos. Un ejemplo resulta el denominado en etapas, el cual considera alguna técnica de muestreo de las mencionadas en este capítulo, seleccionado unidades primarias de muestreo, para que en una segunda etapa hacer selección de unidades secundarias de muestreo, es decir, dentro de las primarias, y así sucesivamente.

Este tipo de esquemas de muestreo se ocupan ampliamente en actividades como la Investigación de Mercados, donde al aplicar encuestas por muestreo, se seleccionan grandes zonas geográficas como estados, municipios, etc. para luego seleccionar colonias o zonas postales, y así sucesivamente hasta seleccionar casas determinadas.²³

Otras posibilidades consisten en la combinación de esquemas de muestreo, por ejemplo, se podría pensar en un diseño estratificado, para que en cada estrato se realizara muestreo sistemático, que de hecho su aplicación, podría ser factible para el diseño en las líneas de surtido, sin embargo, por efectos de logística y presupuesto, y como se verá en el siguiente capítulo, las técnicas de muestreo originales, específicamente el MS, resultan mucho más adecuados para el caso de interés.

²³ cfr. David Aaker, "Investigación de Mercados", pp. 303-305, Thomas C. Kinneer, "Investigación de Mercados".

III

***ANÁLISIS DE CADA TÉCNICA DE
MUESTREO: SELECCIÓN Y
JUSTIFICACIÓN DEL USO DEL
MUESTREO SISTEMÁTICO.***

3.1 UN ANÁLISIS DE LAS CARACTERÍSTICAS PRÁCTICAS DE CADA METODOLOGÍA

Cuando los ingenieros de calidad tomaron verdadero conocimiento de las fallas cometidas en sus diseños y principalmente, de la necesidad de comprender cabalmente lo que en verdad significa obtener una muestra al azar, fue posible también considerar los esquemas de muestreo anteriores, con las ventajas y desventajas en cada caso. Fue necesario hacer un cuidadoso análisis en todos ellos y así llegar, por descarte, al esquema que mejor se ajustaba a los procesos de surtido y revisión, cuidando al menos las condiciones mínimas que exige un diseño al azar, de manera que se garantizara un proceso confiable.

Lo anterior llevó a pensar en la implementación de un método basado en el control estadístico de procesos (CEP), que proporcionaría un beneficio muy importante ya que se tendría controlado el proceso del surtido, mismo que en un principio no era siquiera confiable en la estimación del error de surtido.

3.1.1. EL USO DEL MUESTREO ALEATORIO SIMPLE Y DEL MUESTREO ALEATORIO ESTRATIFICADO EN EL SISTEMA DE REVISIÓN.

Para usar muestreo aleatorio simple (MLA) o bien muestreo aleatorio estratificado (MAE), es necesario etiquetar todos los elementos del universo, es decir, rotular todas las cajas de la producción en el día para luego proceder a su selección al azar. Esto último debe de considerar la posibilidad de manejar fácilmente dichas “etiquetas” o “rótulos” a la hora ser seleccionadas y por lo tanto identificadas como elementos de la muestra.

Dados los procesos logísticos del surtido, no es posible tener gente que mantenga un control manual que identifique todas y cada una de las cajas que se usarán en el día. La única manera práctica de lograr lo anterior, se plantea a continuación:

Existe, dentro del proceso normal de ensamble de cajas de pedido, previo a la colocación en las bandas mecánicas para su posterior uso en el surtido, un sistema de generación y pegado de una etiqueta por caja, la cual contiene un número de identificación impreso por computadora (por lo tanto controlado por ésta), dicho número es único, por lo tanto existe, desde este punto de vista, cierta viabilidad para la selección al azar y para la identificación de cajas para su revisión. Pero existen los siguientes problemas que dificultan seriamente dicha identificación:

a) Los programas que controlan y generan las etiquetas mencionadas, tienen un código fuente inaccesible al área de Aseguramiento de Calidad, única y auténtica interesada en el nuevo proceso. El único departamento responsable, esto es, con acceso autorizado a estos programas es el área de Sistemas (igual de inaccesibles que sus programas).

b) Aún suponiendo que fuera posible tener acceso a los programas del área de Sistemas (ya convencidos los integrantes de esta área) es evidente que éstos deben ser modificados para insertar un algoritmo que genere números aleatorios y que imprima alguna marca o señal en la etiqueta de la caja seleccionada y que pueda entender posteriormente (y de manera exclusiva) el equipo de revisión (no los surtidores). Dicha tarea, de ser posible, tomaría (tiempo estimado por los señores del área de Sistemas) de tres a cuatro meses (que en tiempo de solución al problema, resulta excesivo para la empresa) ya que sería el personal de Sistemas y NO el área de Calidad la que haría las modificaciones, además habría que hacer requerimientos por escrito, análisis, estimación de tiempos y recursos, en fin *burocracia pura* en otras palabras.

c) Debido a que las cajas que tuvieran una marca en la etiqueta generada por la computadora (pensando en el hipotético caso de que hubiéramos salvado los obstáculos anteriores) serían las revisadas y dado que dicha etiqueta se tendría que

haber impreso y pegado en las cajas antes de ser surtidas (no existe forma de marcar las etiquetas después del surtido), ello hubiera implicado que el personal de surtido se hubiese dado cuenta perfectamente bien de qué cajas serían las auditadas y por tanto el error estimado de surtido estaría definitivamente sesgado (incluso llegaría prácticamente a cero) lo cual, por supuesto, hubiera hecho el procedimiento totalmente inoperante (Es evidente que dicho personal de surtido se hubiese “esmerado” y no cometería errores en tales cajas de manera deliberada). Este punto y el anterior descartan definitivamente las opciones de usar muestreo aleatorio simple y muestreo aleatorio estratificado. Por supuesto que lo anterior ocurrió en la realidad, pese a que se implantó un método (muestreo sistemático) con el que no era posible detectar cuáles cajas serían revisadas (gracias al conteo), una persona integrante del equipo de surtido, y a solicitud del equipo de revisión, contaba las cajas y las señalaba con un marcador de textos, y al identificar dicha marca, los integrantes de la revisión separaban las cajas marcadas para revisar, evitándose así el tener que contar las cajas, pero la marca se ponía al principio de la banda mecánica de surtido por lo que todos los surtidores veían la marca. Por esta razón fue necesario hacer presión de modo energético para que no hubiera marca alguna y se procediera a contar las cajas después de ser surtidas.

d) El muestreo aleatorio simple no garantiza hacer un “barrido” de toda la producción diaria en todos los casos en que se aplique, es obvio que gracias a que el proceso es aleatorio, podemos esperar que sí suceda de esa manera en todos los casos, pero NO necesariamente en el 100% de los casos y para siempre, esto resulta hasta cierto punto riesgoso para la estimación si se toma en cuenta que la población de pedidos es probablemente periódica, luego entonces, si por cuestiones del azar, se extraen y revisan todas las cajas antes de las 11:00 a.m. en ese día, por ejemplo, podría haber una estimación muy errónea, ya que al final del día se podrían dar la mayoría de los errores del surtido, y ya no se revisarían dichas cajas.

3.1.2. EL POSIBLE USO DEL MUESTREO POR CONGLOMERADOS (MC) EN EL SISTEMA DE REVISION.

En términos del problema, optar por esta técnica, equivaldría a seleccionar zonas geográficas “completas” las cuales contienen por regla general, alrededor de 100 pedidos, esto es, que un conglomerado consistiría en una zona geográfica completa, y el procedimiento de revisión es similar al usado en las revisiones por zona descritas anteriormente, sólo que en este caso la selección del conglomerado debe ser al azar (para garantizar un procedimiento probabilístico con varianza estimada, error de estimación, etc.) y aquí se presentan tres problemas.

i) No sería posible extraer los 100 pedidos de la zona seleccionada de una vez porque se provocarían conflictos como de por sí ya sucedían en el método anterior, con el área de Distribución (es decir al subir al camión que le llevará al destino final, el cual no puede regularmente esperar más de una hora de haberse surtido) además del tiempo de revisión de los 100 pedidos de la misma zona es mucho más tardado (4 horas aproximadamente).

ii) Extraer de la línea (la banda mecánica) los 100 pedidos de una sola vez provoca que el personal de surtido (como en el caso anterior) se dé cuenta de qué cajas serán revisadas y se confíe a cometer errores peores aún en los pedidos que obviamente no serán revisados, esto sucede principalmente porque no se podrían usar muchos conglomerados dado su tamaño (100 pedidos aprox.) quizá cinco pero mucho más probablemente cuatro, es decir revisar alrededor de 400 ó 500 cajas, y tomando en cuenta que todas las líneas deben ser revisadas diariamente, nos da como resultado sólo revisar un conglomerado por línea o a lo más dos (en una sola línea, en el resto sería uno nada más).

iii) Como en el caso anterior, este método no permite, por regla general, hacer un “barrido” de la producción a lo largo del día. Si por el efecto de la extracción al azar, resulta que se deben bajar todas las cajas de pedido de una zona que se surtió temprano, implicaría que la zona podría ser bajada y revisada, por ejemplo, a las 10:00 o 11:00 a.m., y ya no habría más extracciones en esa línea en el día (no habría suficientes recursos de tiempo, personal y dinero para la revisión) esto permitiría que se relajara la disciplina entre los surtidores, o bien, se podría propiciar que no se tenga conciencia sobre el posible efecto del cansancio periódico y de la distracción de los surtidores (ya en las últimas horas de trabajo), así que, si dicha zona se revisó temprano, puede crecer el error de surtido en la tarde con los efectos obvios en la proporción de pedidos erróneos (error de surtido).

iv) Un problema adicional usando muestreo por conglomerados, que no se prevé en otros esquemas, es que presenta limitaciones de tipo teórico, esto es que como se explicó en el capítulo II (en la sección referente al MC) que solamente podríamos pensar en un tamaño de muestra de conglomerados igual a cuatro (o cinco tal vez) es decir $n=4$, muestra que no garantiza que el estimador de la varianza sea insesgado, este hecho es vital a la hora de construir el intervalo de confianza necesario en el esquema teórico, con lo que con el hecho de que no se cumpla implica que no es factible usar el MC.

3.1.3. EL USO DEL MUESTREO SISTEMÁTICO (MS) EN EL SISTEMA DE REVISIÓN.

El método en este caso, consistiría en tomar y revisar una caja de cada k de ellas, es decir contar y dejar pasar k cajas para proceder a la revisión de la siguiente y así sucesivamente a lo largo del día. Esto nos lleva a realizar un análisis como sucedió en las técnicas anteriores,

en este esquema encontramos varias ventajas sobre los otros métodos de muestreo, a continuación se comentan:

1) El muestreo sistemático no es un método complicado, sino muy sencillo de ser entendido por el equipo de revisión, lo mismo de ser aplicado ya que sólo hay que contar una caja de cada k y proceder a la revisión.

2) No requiere etiquetar cajas, pues basta contar mentalmente las mismas para seleccionarlas, esto nos lleva a que el equipo de surtido no tiene forma de saber qué caja será revisada. Lo cual evidentemente es una gran ventaja sobre cualquier otro esquema de muestreo.

3) La revisión se hace a la salida de las cajas ya surtidas, o sea cuando aún están en las bandas mecánicas pero aún no son subidas a los camiones de distribución.

4) Revisar una caja toma en promedio 2 minutos, por lo que puede ser bajada y revisada sin problemas, para luego reintegrarla al camión.

5) Con este método, bastaría con las 8 personas que integran actualmente el equipo de revisión ya que, dado que son cuatro bandas mecánicas (líneas), se ubicaría a dos personas en cada banda, una dedicada al conteo de los pedidos y otra revisando los pedidos que fueran seleccionando. Dado que revisar una sola caja no toma mucho tiempo (2 minutos) no se da la situación de que pase la siguiente caja a ser revisada cuando no se ha terminado de revisar la anterior (se verificó en un muestreo piloto). Con los otros métodos, es muy complicado ajustarse a un equipo de 8 personas o menos. Se presenta en el **anexo 4** un diagrama de las líneas de surtido y el lugar donde se estableció a los integrantes del equipo de revisión, según el esquema de muestreo sistemático propuesto.

6) Una de las mayores ventajas: dado que se revisa una caja de cada k , ello implica que se hace un “barrido” en toda la producción diaria, o sea, que se observa el error a todas horas del proceso, al iniciar el día, a medio día, en la tarde, al terminar el día, en fin, todo el tiempo, por lo tanto podemos afirmar que se cubre toda la población de cajas.

Existe un único y posible problema que debe ser considerado al emplear muestreo sistemático, y que se ha mencionado antes: si la población bajo estudio es periódica, entonces el muestreo sistemático puede dar como resultado una varianza del estimador mayor que la obtenida por el muestreo aleatorio simple, ello dependerá explícitamente del valor de k^1 , es decir, del tamaño de muestra (número de revisiones) que estemos dispuestos a emplear, a mayor tamaño de muestra, es evidente que se cubrirá con mayor eficiencia la población, reduciendo el riesgo de obtener una mala estimación por periodicidad en la muestra. Otro aspecto importante es que en cada proceso de revisión se usarán números aleatorios de arranque distintos cada día, con lo que se reduce la posibilidad de seleccionar observaciones (o sea cajas) con la misma posición relativa en una población periódica.² De la misma manera se puede, por ejemplo, tomar el mismo día con dos arranques aleatorios para la misma línea de surtido. Esto disminuye la probabilidad de sesgo en un día.

Pensando en la posibilidad de periodicidad, la información que proporcionaría el muestreo sistemático es menor que empleando muestreo aleatorio simple (por unidad de costo) considerando que el tipo de población sea éste. Al hacer una reflexión cuidadosa de lo anterior, se concluyó que sí existen razones para pensar que se tiene una población periódica:

El hecho de que sea en toda la jornada la misma gente la que está surtiendo los pedidos implica posiblemente que a ciertas horas del día dichos surtidores se cansan, ocasionando que la evidente falta de concentración les provoque descuidos al surtir. En otras palabras, a

¹ William G. Cochran, “Técnicas de Muestreo”, p. 271

ciertas horas del día (las mismas siempre), el nivel de cansancio y falta de concentración se elevaría lo suficiente para ocasionar errores de surtido.

Independientemente de que el método de selección de los elementos que conforman la muestra diariamente pudiera ser sistemático, se considera importante lo que sucede en cada línea de surtido, por eso se pensó en separar el proceso de revisión en cada banda mecánica, esto es, verlo como cuatro procesos de revisión independientes. Pese a lo anterior el método de selección sería sistemático para cada línea de surtido.

² William Mendenhall, "Elementos de Muestreo", pp. 174-175

3.2. APLICACIÓN DEL MUESTREO SISTEMÁTICO COMO MÉTODO DE SELECCIÓN.

Después del análisis anterior, es claro que el mejor método es el muestreo sistemático, específicamente su utilización como método de selección ofrece una gama de ventajas que los otros métodos no reportan, esta es una de las razones por las que el sistemático representa la mejor opción tratándose de procesos en serie, como es nuestro caso.

Es importante resaltar que desde el punto de vista estadístico matemático estricto pudiera ser mejor o no otra técnica de muestreo, sin embargo, los inconvenientes operativos de los otros métodos son los que impiden su aplicación y esto, a fin de cuentas es lo más importante para el presente trabajo, es decir, el trabajo no pretende demostrar qué técnica de muestreo es mejor desde el punto de vista matemático (eso es tarea de un texto de muestreo). Se insiste entonces que el objetivo del trabajo es aplicar el mejor método estadístico pero que además sea aplicable desde un enfoque práctico.

Como primer paso, se estableció la necesidad de implantar un muestreo piloto ya que no se contaba con una medición estadística confiable de medidas de dispersión en el proceso con la que se pudiera calcular un tamaño de muestra recomendado.

Otra intención con el muestreo piloto, era que los ingenieros se percataran de los beneficios de este esquema de muestreo, y observar el impacto del cambio de método en el grupo de surtidores.

Por cuestiones de presupuesto y para no alterar demasiado las formas de trabajo en el área de surtido. Se decidió respetar el número de revisiones llevadas a cabo hasta el momento (o sea un tamaño de muestra de 400) para efectuar el muestreo piloto, el cual se llevó a cabo de la siguiente manera:

- a) Dado que se pensó en las 400 revisiones originales y partiendo del hecho de que en cada línea de surtido se desplaza una cantidad similar de pedidos (3125 por línea) se planteó revisar cada 31 cajas, es decir: dejar pasar exactamente 30 cajas para revisar la número 31 ($k=3125/100=31.25\approx 31$ en cada línea o bien, $12500/400=31.25\approx 31$, por ser muestreo piloto, no se tomó en consideración la parte fraccionaria).
- b) Se instruyó a los encargados de las revisiones para efectuar su trabajo al terminar el surtido pero antes de que la caja llegara al camión de distribución, con la mayor discreción posible para no influir en el trabajo de los surtidores. Se recalcó también que este método era “sorpresa”, es decir ningún surtidor debía estar enterado, sólo el encargado general del surtido que controla las cuatro líneas de surtido y el gerente de operaciones, que evidentemente mantendrían discreción por razones obvias.
- c) Dos empleados de revisión estarían trabajando en cada línea, uno contando las cajas al pasar, y otro revisando las cajas seleccionadas y registrando los errores encontrados.
- d) El proceso de revisión se iniciaría al empezar el surtido (6:00 a.m.).

Los resultados del piloto fueron los siguientes:

Al efectuarse el muestreo piloto, se detectó un efecto negativo por la sorpresa del “operativo”, que descontroló a los surtidores, el error estimado en ese primer día fue de aproximadamente 7%, cifra muy elevada desde el punto de vista del personal de Calidad y el de Operaciones.

Al discutir el problema, se concluyó sobre el mencionado efecto negativo, por lo que se decidió efectuar durante varios días más el mismo operativo (5 días continuos) pero sin informar a los surtidores el propósito del método, con lo anterior se logró acostumbrarlos y el error se pudo considerar confiable. Los resultados coincidieron con lo reportado por los ingenieros de Calidad y los de Operaciones, es decir en 5% como promedio por la semana que duró el muestreo piloto.

En este momento, fue posible calcular tamaños de muestra confiables, se presentó a los ingenieros de calidad varias opciones en tamaños de muestra (número de revisiones diarias) jugando con un nivel de precisión razonable, para que tomaran la decisión de cuántas podían comenzar a realizar en función de aspectos logísticos y presupuestales.

A los ingenieros de Calidad se les dio la siguiente explicación: "si consideramos que se tienen 12,500 pedidos diarios como promedio, a lo cual consideramos el universo N , y ahora, si consideramos el nivel de error de surtido que los ingenieros habían estimado por mucho tiempo o sea, 5%, es posible calcular el tamaño de muestra, es decir, el número de revisiones recomendadas y cada cuántas cajas se debe extraer una para proceder a su inspección porque ya pudimos calcular una medida de dispersión (la varianza) bajo un esquema confiable. Por lo tanto, el número de revisiones dependerá ahora de el nivel de precisión que estemos dispuestos a considerar". Se les explicó también que dicho nivel de precisión se refería al tamaño del intervalo donde podíamos asegurar que estaba el error de surtido auténtico con un 95% de confianza, usando un determinado tamaño de muestra. Se le planteó que mientras más estrecho escogiéramos dicho intervalo o nivel de precisión, mejor nos aproximaríamos en nuestra estimación al nivel real de error de surtido, pero que ello implicaba que necesitábamos más información, es decir, más revisiones, de tal suerte que a mayor número de revisiones, mayor precisión e intervalo más estrecho y viceversa.

**ESTA TESIS NO SALI
DE LA BIBLIOTECA**

De esta manera, los ingenieros pudieron tener elementos para tomar una decisión sobre el tamaño de muestra en base a la precisión obtenida, costos del procedimiento y dificultades que pudieran surgir en el implante del proceso seleccionado.

Presentamos una tabla con tres opciones (entre varias posibles) que discutieron los ingenieros de Calidad antes de optar por alguna. Desde luego, los tamaños de muestra se calcularon en base a las fórmulas expuestas en el capítulo II, empleando un nivel de confianza de 95%. Se les explicó que en esta tabla, se debían manipular las celdas correspondientes al número total de pedidos surtidos a diario (N), el error de surtido obtenido en el muestreo piloto y el error de precisión, para que, se pueda obtener el número de revisiones y el intervalo de revisión (n y k respectivamente).

PROPUESTA PARA LA MEDICIÓN DE ERRORES EN LAS LINEAS DE SURTIDO EMPLEANDO MUESTREO SISTEMATICO

Pedidos surtidos (a diario)	12,500	12,500	12,500
proporc. est. (pedidos incompletos)	5.00%	5.00%	5.00%
Error de precisión obtenido	0.800%	1.147%	2.100%
Revisiones recomendadas:	2322	1250	401
Intervalo de revisión (k):	5	10	31

Una vez que se presentó esta información, los ingenieros de Calidad optaron por una muestra total de 1250 revisiones diarias en las cuatro líneas, es decir, revisar cada 10 cajas de pedido., Esto último al comprobar (desde el piloto) que la carga de trabajo para el equipo de revisión era bien asimilable.

Debido a la situación discutida sobre el posible problema que se pudiese presentar con respecto a la periodicidad de la población, se determinó tomar las dos medidas planteadas

anteriormente, es decir, emplear un número de revisiones (tamaño de muestra) suficientemente grande (al menos 1250 revisiones), y por otra parte, se sugirió la aplicación de números aleatorios de arranque, uno diferente para cada línea, de esta manera, independientemente del apoyo por el problema de periodicidad en la población, tuvo otra repercusión positiva a nivel práctico: al arrancar en una caja generalmente diferente en cada línea y cada día, los surtidores no podían darse cuenta qué cajas se revisarían (tratando de contar mentalmente) porque el patrón era distinto cada día. De la misma manera, los ingenieros de Calidad decidieron “jugar” con el número de revisiones cada día, es decir, un día revisaban cada 10, mientras que otro podría ser cada 8, aumentando o disminuyendo (con un mínimo de revisiones en total de 1,250 es decir cada 10 cajas). Se les proporcionó una hoja de cálculo con las fórmulas implementadas para que pudiesen calcular sin problemas el número de revisiones y el valor de k (cada cuantas cajas).

Se dio una situación que no estaba prevista, pero que representó un beneficio muy importante para la empresa: El nivel de error de pedidos bajó de manera considerable, de manera gradual, fue bajando hasta alcanzar un nivel promedio de 2%, manteniéndose estable desde entonces. De esta manera, el número de pedidos mal surtidos (en promedio) actualmente es de 250, de un total de 12,500 surtidos a diario. En la siguiente sección se explica la razón de ello.

3.3 BENEFICIOS OBTENIDOS

Se realizó un análisis minucioso de los beneficios obtenidos optando por este tipo de revisión y se pudo comprobar que representaba grandes ventajas el uso del muestreo sistemático en sus procesos normales:

- a) Se obtuvo un nivel de error confiable, en el sentido de que se tenía conocimiento sobre el nivel de error de surtido diario, reportado y validado estadísticamente.

- b) El nivel de error bajó considerablemente (3 puntos porcentuales) y se mantuvo estable en ese nivel, esto se debió principalmente a que el nuevo método representó un factor de presión positiva que motivó a los surtidores a poner más cuidado en su trabajo de manera permanente, pues el método llegó para quedarse.
- c) Fue posible manejar los pocos recursos humanos con que se contaba ahora (8 en vez de 23 inspectores). Con otro método igualmente confiable, esto no hubiese sido posible.
- d) El nuevo método motivó la generación de nuevas ideas de trabajo, especialmente la de implantar controles estadísticos en otras áreas (como en nuestro caso, que se expone en el punto 3.4) o bien se pensó en echar a andar un viejo proyecto (ya 'velado y enterrado') que consistía en el uso de un semáforo por línea de surtido, que indicara los errores por fracción de tiempo en el mismo día y así dar la alerta oportunamente y corregir tendencias desfavorables en el error de surtido.
- e) El número de revisiones con el nuevo método se redujo a una tercera parte, pues pasó de una tasa muestral (fracción de muestreo) de 30% a sólo 10%, que representa un ahorro muy considerable en tiempo y dinero para la organización. En tiempo por ejemplo, se redujo de 184 horas/hombre a 64 horas/hombre diarias invertidas (reducción de alrededor de 65%).

3.4 DISEÑO SUGERIDO DE UN CONTROL ESTADISTICO DE PROCESO.

Los resultados derivados del nuevo método de revisión superaron las expectativas originales de los ingenieros de Calidad, como se mencionó en el punto anterior, esto motivó a idear un control de proceso validado estadísticamente, el cual sirviera en la inspección de un modo rutinario, que permitiera detectar si el proceso del surtido está “fuera de control” o en línea.

Este punto no representó grandes dificultades, dado que, se contaba ya con el método de inspección correcto, es decir, que se disponía ya de un instrumento de medición confiable del error de surtido y con el cual (desde luego) era ya posible tomar decisiones acertadas en cuanto al surtido de pedidos.

El criterio estadístico para la construcción de un proceso en Aseguramiento de Calidad, tiene que ver necesariamente con la construcción de intervalos de confianza, en este caso, para una proporción, vimos en el capítulo II una expresión para un intervalo de confianza para una proporción, adecuado a estos propósitos, por esta razón, se sugirió el uso de esta expresión como límite de control superior (LCS), es evidente que no se requiere de un límite de control inferior (LCI) dado que el proceso es de una sola cola (sólo interesa que el error no “crezca” demasiado). Hecho lo anterior, será necesario la construcción de una gráfica de control para la proporción de error de surtido.³

Efectuando los cálculos, tenemos que $n=1,250$, $N=12,500$, $p=0.02$ y tomamos $z=2.58$ es decir con $\alpha=0.01$ que es (según algunos autores)⁴ un nivel de significancia usual. Por lo tanto, el límite superior (LCS) queda definido como:

³ Erwin Kreyszig, “Introducción a la Estadística Matemática”, p. 254

⁴ crf. Erwin Kreyszig, “Introducción a la Estadística Matemática”, p. 255, Acheson J. Duncan, “Estadística Industrial”

$$p \pm \left[z\sqrt{1-f}\sqrt{pq/(n-1)} + \frac{1}{2n} \right]$$

Donde $f = \frac{n}{N}$

$$\Rightarrow 0.02 + \left[2.58\sqrt{1-\frac{1250}{12500}}\sqrt{\frac{(0.02)(0.98)}{1249}} + \frac{1}{2500} \right]$$

$$= 0.02 + \left[2.58 * 0.0037581 + \frac{1}{2500} \right]$$

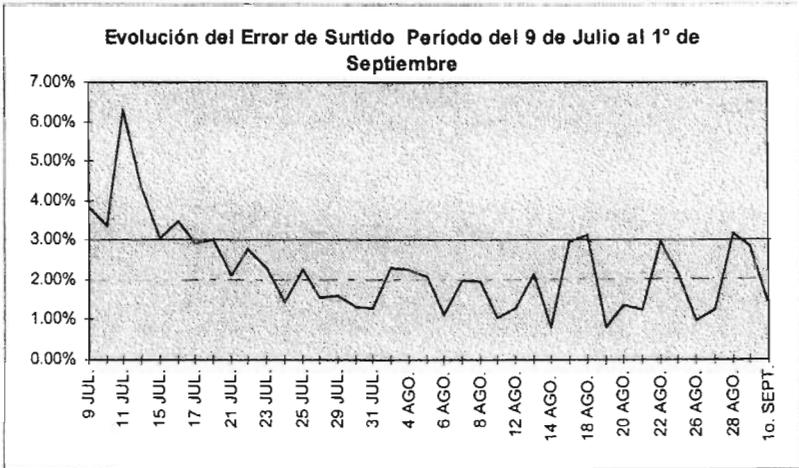
$$= 0.02 + 0.01$$

$$= 0.03 = 3\%$$

Entonces LCS=0.03 con un nivel de confianza de 99%.

Debido a que los ingenieros de Calidad aceptaron la propuesta de usar este dato como LCS, se decidió su aplicación de manera inmediata, se hizo la sugerencia de observar el proceso y registrarlo durante algunos días, al principio se observaron niveles de error altos, debido a que había que esperar total adaptación por parte de los surtidores al nuevo método. Dos semanas después de iniciado, el error de surtido entró de manera constante en los niveles deseados, debido a ello, los ingenieros de Calidad buscaron la manera de mantener el nivel en un máximo de 3% como lo marcaba el LCS, a través de constante capacitación, concientización y supervisión, a partir de la tercera semana, se detectaron sólo tres desviaciones al nivel de error de surtido máximo, que al día siguiente de ocurridas, se tomó las medidas para volver a tener bajo control el nivel de error. Con lo que el proceso estadístico quedó definitivamente establecido y aceptado.

Se anexa una gráfica de cómo evolucionó el error de surtido en lo sucesivo:



CONCLUSIONES

Pudimos comprobar que esta aplicación del muestreo sistemático y del control estadístico de procesos contribuyó en varios frentes a que la organización alcanzara beneficios importantes, por un lado, se obtuvo un procedimiento confiable, es decir, serio y que le permitió, por lo tanto, tomar decisiones correctas y oportunas en una problemática que se hecho muy preocupante para la imagen de la empresa ante su clientela. Eso principalmente era lo que estaba en riesgo de deterioro. Por otro lado, se logró un ahorro de recursos muy considerable, sin restarle validez a sus procesos de revisión.

Cabe decir que resultó tan importante este proceso, que se instauró como prioritario en su proceso de certificación en la norma ISO 9001 (uno de los objetivos primordiales de la organización). Al cabo de un tiempo, el método quedó registrado formalmente como parte de las técnicas estadísticas de inspección y aseguramiento de Calidad. Hoy día es uno de los métodos más importantes y sigue formando parte de los manuales y procedimientos que se auditan regularmente en los procesos de ratificación del ISO 9001 para la organización.

Posiblemente en lo que más ganó la empresa, es que, por fin, se comprendió que cuando las técnicas estadísticas son aplicadas correctamente, es posible optimizar muchos procesos y diseñar correctamente otros. Fue posible adquirir gradualmente una cultura de información y tratamiento estadístico de datos, logrando con ello, obtener más información importante a un bajo costo.

Otra reflexión importante derivada del trabajo presentado, es que, el principal factor para que una técnica estadística pueda mostrar su utilidad en cualquier proceso de producción, de investigación o de cualquier índole, es definitivamente, que sea conocida realmente y no de manera marginal (“como me la enseñaron en la preparatoria o en la carrera”). La preparación a conciencia en estadística es esencial en este sentido, porque son las personas que realmente estudian este campo las que, con responsabilidad, pueden borrar esa imagen

que muchas veces se tiene acerca de los estadísticos (y de la estadística, lo que es más grave) en que se aplican técnicas "al vapor", saltándose preceptos importantes (recuérdese aquello de "muestreo al azar") que dan al traste con los resultados al aplicar un método y con el trabajo serio de otras personas.

Sólo existen dos caminos: la preparación con mayor conciencia o el recurrir a expertos estadísticos que nos hagan la tarea. En este sentido, es más factible pensar que lo primero se alcanza más difícilmente y sin embargo es más importante. La razón es que nadie recurrirá a un estadístico si no conoce a ciencia cierta qué es lo que realmente hace ... ¿Cómo detectar una necesidad si no se sabe qué puede hacer la estadística por nosotros?. Aquí radica una de las más importantes razones de la Especialización en Estadística Aplicada: que los profesionistas se "lleven" a su campo de acción todo este bagaje y comprendan dónde pueden aplicarlo y dónde sería necesario recurrir a un experto (con estudios a nivel de Maestría o Doctorado en Estadística).

A propósito de expertos, los estadísticos de carrera tienen en este sentido un campo de acción muy interesante y sin embargo, se encuentran con un gran reto: Una vez que se han preparado en estadística podrían advertir que para hacer aportaciones interesantes o simplemente resolver un problema es necesario conocer a fondo dicho problema y su entorno. Lo que los lleva al deber de informarse del campo donde se van a aplicar. Esto que puede parecer obvio, en realidad advierte seriamente que los conocimientos de los expertos en estadística podrían quedar aislados, sin poder ser apreciados y requeridos por otros profesionales. La estadística sin conocimientos relacionados a otro campo se convierte en algo abstracto y en un contrasentido para aquel que desea aplicarla ("La voy a aplicar pero ¿en qué?"). Es por ello, muy importante mantenerse en contacto constante con otros campos de conocimiento, entendiendo las necesidades de los mismos con sus principios elementales e imaginando dónde "entramos", es decir, cuál puede ser la participación del estadístico en ese problema en particular. Por otro lado, existen también cierta decepción en personas que esperan que un estadístico les resuelva sus problemas sin tomar en cuenta que

es vital para este último, la información sobre el fenómeno y que es necesario esperar cierto tiempo para hacer sugerencias válidas que contribuyan realmente a su resolución.

Todo esto engloba un problema importante: El verdadero conocimiento de la estadística como método de investigación y como abastecedor de soluciones depende de lo que difundamos de manera correcta, dentro y fuera de empresas, en recintos universitarios y en ámbitos de investigación. Seremos más eficaces como estadísticos si y sólo si somos capaces de inculcar una verdadera cultura estadística a nuestros clientes, a nuestras empresas y a nuestros alumnos.

APENDICE 2 * 5* ZONA: 5096 - RUTA: 038 CAMP. 22

PED. 118800

FULL. 1028097

"A" ESTACION "A"

~~1--AJ~~ H

"B" ESTACION "B"

~~1--AF~~ SCP
~~2--AJ~~ F6

"C" ESTACION "C"

~~4--AJ~~ 19A

"E" ESTACION "E"

~~1--BN~~ 17W
~~1--FB~~ 661

"G" ESTACION "G"

~~1--A~~ 92H
~~1--EK~~ 19M

"H" ESTACION "H"

~~1--AB~~ 54P
~~1--AF~~ 4B
~~1--FJ~~ CJK

"J" ESTACION "J"

~~1--A~~ 320
~~1--AB~~ 54Y

"K" ESTACION "K"

~~1--AD~~ M 24D
~~1--BU~~ 21C

"L" ESTACION "L"

~~1--B~~ H27

"P" ESTACION "P"

1--CB 37B

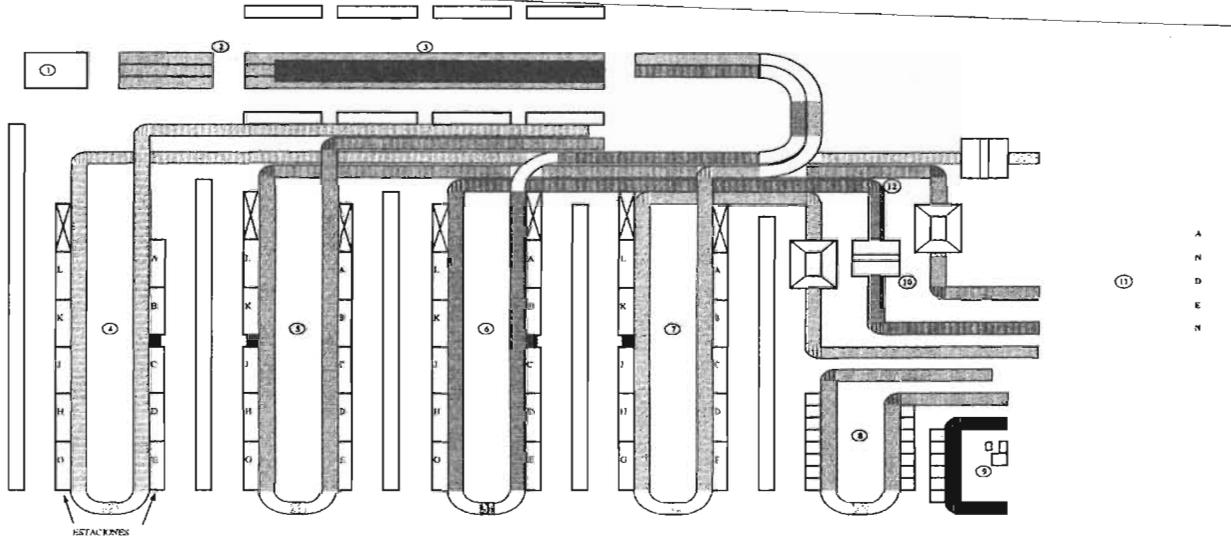
"X" ESTACION "X"

2--A P339
~~1--AB~~ P341

UNID.: 24 | 22

1E Z G M 1C
 6914

julio - 97



- ① ARMADORA DE CAJA
- ② ETIQUETADO
- ③ SET TABLE
- ④ LINEA 1
- ⑤ LINEA 2

- ⑥ LINEA 3
- ⑦ LINEA 4
- ⑧ LINEA DETALLES
- ⑨ LINEA PAPELERÍA
- ⑩ EMPAQUE

- ⑪ INTEGRACIÓN DE ZONAS
- ⑫ FACTURADO

✓ SET TABLE

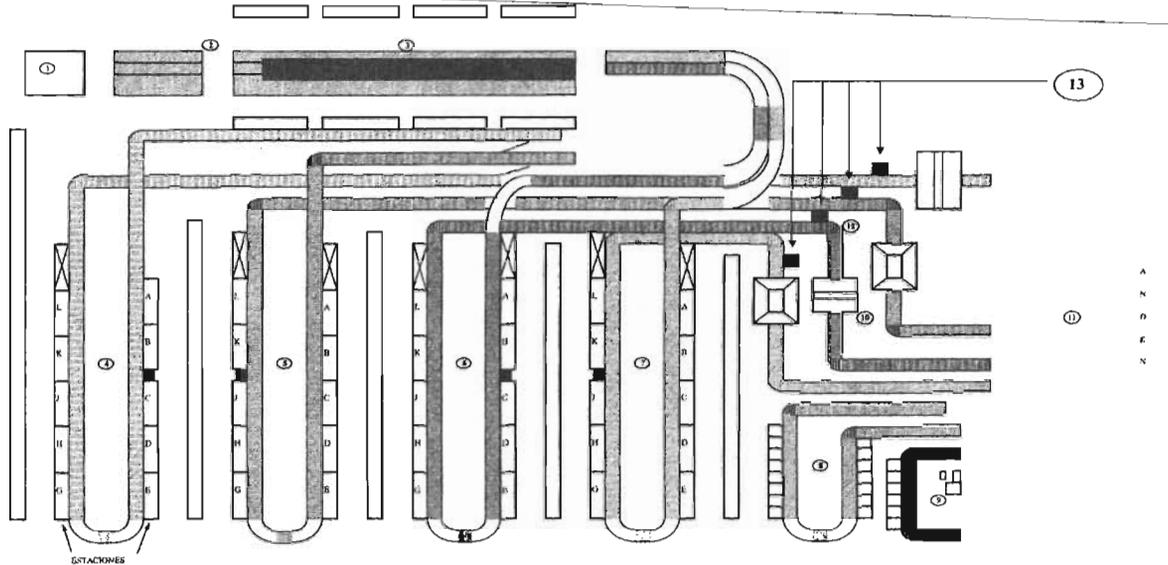
- ☐ PRODUCTO DE BAJO DESPLAZAMIENTO
- ☐ 4 ESTACIONES (110 SKU'S C/U) A - D
- ☐ 2 Y 4 NIVELES
- ☐ CAPACIDAD 440 SKU'S

✓ LÍNEAS DE SURTIDO

- ☐ PRODUCTO DE BAJO, MEDIANO Y ALTO PLAZAMIENTO
- ☐ 10 ESTACIONES (54 SKU'S C/U) A - L
- ☐ 3 NIVELES
- ☐ CAPACIDAD 540 SKU'S

✓ CAPACIDAD TOTAL 980 SKU'S

julio - 97



ALMACEN	CAMBIOS	OFICINAS	M. DE C.	P A S I L L O	P A S I L L O	DISTRIBUCIÓN
---------	---------	----------	----------	---------------------------------	---------------------------------	--------------

- ① ARMADORA DE CAJA
- ② ETIQUETADO
- ③ SET TABLE
- ④ LINEA 1
- ⑤ LINEA 2
- ⑥ LINEA 3
- ⑦ LINEA 4
- ⑧ LINEA DETALLES
- ⑨ LINEA PAPELERÍA
- ⑩ EMPAQUE

- ✓ SET TABLE
 - ☐ PRODUCTO DE BAJO DESPLAZAMIENTO
 - ☐ 4 ESTACIONES (110 SKU'S C/U) A - D
 - ☐ 2 Y 4 NIVELES
 - ☐ CAPACIDAD 440 SKU'S

- ⑪ INTEGRACIÓN DE ZONAS
- ⑫ FACTURADO
- ⑬ UBICACIÓN DEL EQUIPO DE REVISIÓN DE PEDIDOS NUEVO MÉTODO DE MUESTREO

- ✓ LÍNEAS DE SURTIDO
 - ☐ PRODUCTO DE BAJO, MEDIANO Y ALTO PLAZAMIENTO
 - ☐ 10 ESTACIONES (54 SKU'S C/U) A - L
 - ☐ 3 NIVELES
 - ☐ CAPACIDAD 540 SKU'S

- ✓ CAPACIDAD TOTAL 980 SKU'S

Bibliografía

- [1] **William G. Cochran, “Técnicas de Muestreo”, Editorial Limusa**

- [2] **Ignacio Méndez Ramírez, “Conceptos muy elementales del muestreo con énfasis en la determinación práctica del tamaño de la muestra”, Comunicaciones Técnicas, serie azul: monografías. Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM.**

- [3] **Erwin Kreyszig, “Introducción a la Estadística Matemática”, Editorial Limusa.**

- [4] **William Mendenhall y Richard Scheaffer, “Elementos de Muestreo”, Editorial Iberoamérica.**

- [5] **Acheson J. Duncan, “Estadística Industrial”, Editorial AlfaOmega.**

- [6] **David Aaker, “Investigación de Mercados”, Editorial McGraw-Hill.**

- [7] **Thomas C. Kinnear, “Investigación de Mercados”, Editorial McGraw-Hill.**

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: ROBERTO RAMIREZ
HERNANDEZ

FECHA: 11/ABR/2005

FIRMA: _____



**ESTA TESIS NO SALE
DE LA BIBLIOTECA**