



UNIVERSIDAD
NACIONAL
DE TUCUMÁN



FACULTAD DE
CIENCIAS ECONOMICAS
UNIVERSIDAD NACIONAL TUCUMAN

PROPUESTA DE DISEÑO MUESTRAL PARA ESTIMACIÓN DE RENDIMIENTOS EN LA PRODUCCIÓN DE SOJA

Autores: Gutiérrez Orozco, Sofía B.
Yalis, María Virginia

Director: López Cleip de Sosa, Aída

2013

Trabajo de Seminario: Contador Público Nacional

RESUMEN

La soja es un cultivo milenario, aunque en nuestro país su inclusión como tal en la agenda agrícola se remonta a pocas décadas atrás. Es así que durante los últimos años, en los cuales las extensiones sembradas y cosechadas se incrementaron notoriamente como consecuencia de un precio internacional muy atractivo por la rentabilidad que reportaba, se impuso como uno de los productos agropecuarios más importantes en la balanza comercial argentina.

Teniendo conocimiento de la importancia que tiene hoy la soja en el mundo y en la Argentina por ser uno de los países productores y uno de los principales exportadores, consideramos una problemática vinculada a ella - estimación de rendimiento y de volúmenes de producción- como tema de nuestro seminario y realizamos el aporte de una herramienta que permita agilizar y mejorar las decisiones vinculadas a su comercialización.

Observando el amplio abanico de posibilidades de temas a tratar respecto a este commodity, nos enfocamos en la parte estadística por haber interiorizado de la problemática planteada por productores a profesionales y a organizaciones asesoras en diversos aspectos referidos a la soja: los mismos no cuentan con un método objetivo que les permita realizar estimaciones significativas estadísticamente y tomar decisiones fundadas científicamente al asumir sus compromisos comerciales.

A continuación ofrecemos una propuesta al asunto planteado basándonos en los conocimientos adquiridos en la materia estadística, de la información relevada sobre el cultivo de la soja en los aspectos vinculados a la temática específica y de lo aprendido a largo de nuestra carrera, con la guía de docentes de la cátedra y de la directora del seminario. La misma brinda un diseño muestral que permita la estimación de rendimientos y producción de soja, en la etapa previa a la cosecha y con las limitaciones que oportunamente se señalan.

PRÓLOGO

El trabajo tiene por objetivo proponer una formalización, a partir de información de diversas fuentes, de un diseño muestral que permita la estimación de rendimientos y producción de soja en la etapa previa a la cosecha. La encuesta por muestreo basada en un diseño de este tipo permitirá conocer datos con una fundamentación científica que la toma de decisiones de empresas productoras, acopiadoras y organismos públicos o privados interesados en la temática requieren.

Decisiones sobre planificación de distintos aspectos asociados a la cosecha efectiva como necesidades financieras, logística de transporte de granos a lugares de acopio próximos y al puerto de embarque, organización de tiempos de recolección, relación con capacidad de almacenaje disponible, cantidad de horas máquina de cosecha, compras de granos a terceros por compromisos de exportación, reservas de transportes marítimos, entre otras, son decisiones que dependen de estadísticas fiables por cuanto **las estimaciones visuales de la producción, que muchas veces se emplean,** involucran el riesgo de la subjetividad del estimador porque no es posible conocer la precisión de las mismas o el grado de error que pueda afectarlas.

En el ámbito público, organismos de extensión y de producción de estadísticas macroeconómicas necesitan datos estadísticamente significativos para cumplir con sus objetivos y a nivel microeconómico la principal

preocupación del productor es si los rindes del cultivo habrán de justificar la inversión según el precio del grano al momento de la comercialización en el mercado con mercadería disponible y a término.

Conocer el dato con anticipación constituye una información valiosa que habrá de orientarlo en una decisión más racional en el sentido más económico del término: maximizar los beneficios o minimizar sus pérdidas.

CAPÍTULO I

Aspectos principales de la soja

SUMARIO: 1.1.- ¿Qué es la soja? 1.2.- Alimentos que derivan de la soja 2.- Importancia para las personas 2.1.- Beneficios y Propiedades 3.- Origen y Difusión 3.1.- Inicio y Desarrollo del Cultivo en Argentina 4.- Estadios Fenológicos Externos del Cultivo de Soja 5.- La Importancia de la Soja para Argentina 5.1.- Evolución 6.- La Reina de los Commodities Agrícolas 6.1.-Producción y Demanda Mundial 6.2.- Cotizaciones, precio y mercados 6.2.1.- M.A.G. y P. de la Nación-Valor de Mercado 6.2.2.- Consulta cotizaciones de precios Cámara Arbitral de Rosario 6.2.3.- Diferenciales de Soja 6.2.4.- Análisis: semana con subas para la Soja (semana del 15 al 19 de Abril del 2013)

1.- ¿Qué es la Soja?

La soja (o soya) es una leguminosa. En su forma es muy parecida a las semillas de las lentejas. No se trata de un cereal; es pariente cercana de los guisantes, habas o judías. La vaina guarda de 2 a 4 porotos (alubias), es oscura y vellosa. Para su cultivo requiere temperaturas cálidas y abundante agua. Es una planta dadora de nitrógeno al suelo, y de ella se aprovecha absolutamente todo. Fundamentalmente es en forma de ingredientes, producidos a partir del grano, como se utiliza la soja en la fabricación de alimentos industriales. Basta con estudiar la composición declarada de muchos alimen-

tos industriales para constatar que las diversas fracciones de la soja se utilizan de manera masiva en la confección de un gran número de alimentos, desde embutidos hasta pasteles.

Esencialmente contiene proteínas, lípidos, glúcidos y minerales. Los minerales que contiene el grano de soja son principalmente el calcio, el zinc y el hierro. La soja también es fuente de fibra soluble e insoluble, cuyos efectos fisiológicos son bien conocidos.

La calidad nutricional de las proteínas está determinada por su composición de aminoácidos esenciales y su digestibilidad. Muchos aminoácidos esenciales de la proteína vegetal concentrada en la soja, están en cantidades semejantes a las de las proteínas que contiene el huevo y tiene una excelente tolerancia gastrointestinal.

La soja contiene todos los aminoácidos esenciales necesarios para cubrir los requerimientos del ser humano para el crecimiento y el mantenimiento de las funciones orgánicas. Su patrón de aminoácidos es uno de los más completos dentro de las proteínas vegetales y es muy similar al de las proteínas animales de alta calidad, con excepción de los aminoácidos sulfurados como la metionina.¹

1.2.-Alimentos que derivan de la soja

- Aceite de soja: Cada vez se utiliza más, se obtiene de la semilla de una proteoleoginosa, la glycine máxima. El grano de soja se caracteriza por tener un elevado contenido lipídico (20%). El aceite de soja es fuente de ácidos grasos esenciales poliinsaturados, sobre todo de ácido linoleico y ácido alfa-linolénico. También tiene lecitinas.

La soja es proveedora de aceites con triglicéridos esterificados con ácidos grasos de alto grado de insaturación. No recomendable para frituras, pero sí en crudo para ensaladas. Pese haber sido perfectamente refinado, puede adquirir sabores extraños, (como a pescado y con aroma a barniz).

- Leche de soja: Su aspecto es similar al de la leche de vaca. Es nutritiva y refrescante. Sin colesterol, sin lactosa y caseína, contiene mucha lecitina. Proporciona más proteínas, menos grasas y menos carbohidratos que la leche de vaca y comparada con la leche materna contiene menos calcio.

-Salsa de soja: Se prepara con judías de soja y sal que se dejan fermentar y se mantienen durante unos años, al menos tres, en barricas de madera, como si del mejor vino se tratara. Resulta ser un condimento sano y sabroso que convierte en succulenta la más sosa de las materias primas. Un buen saborizante de variados tipos de comidas.

- Harina de soja: Se presenta desgrasada o integral, cualquiera de las dos son muy apreciadas en repostería porque apenas alteran el sabor del dulce. En una masa se puede prescindir de echar huevo, ya que contiene lecitina, que es un emulsionante natural.

-Como fuente de lecitina: Se utiliza en la industria alimentaria, en dulces y chocolate, como sustancia emulsionante; como estabilizante de la leche en polvo. Además se emplea en la producción de cintas magnéticas, barnices, colorantes e industria textil.

-Tofu: Es la concentración de la pulpa de la judía de soja. En su composición cuenta con 9 aminoácidos y el valor proteico es superior al de la carne, el huevo y la leche.

También tiene un aporte importante de minerales, en especial el hierro y el potasio. Su textura es muy similar a la del queso fresco.

2.-Importancia para las personas

El cultivo de soja, además de ser un factor muy valioso, ayuda al ser humano si se efectúa en el marco de un cultivo por rotación estacional, ya que fija el nitrógeno en los suelos, agotados tras haberse practicado otros cultivos intensivos. En cambio, el monocultivo de soja acarrea desequilibrios

ecológicos y económicos si se mantiene prolongadamente y en grandes extensiones.

2.1.-Beneficios y propiedades

- Hipoglucemiante: reduce la tasa de azúcares en sangre (tratamiento de diabetes).
- Fuente de proteínas en la alimentación.
- Previene los trastornos cardiovasculares; reduce el colesterol.
- Alivia los trastornos de la menopausia y menstruales por presentar:
 - Isoflavonoides: con acción hipocolesterolizante.
 - Fitoestrógenos: estrógenos de origen vegetal.
- Previene la osteoporosis por la reducción de estrógenos femeninos.
- De la soja se obtienen diversos derivados, como la bebida de soja o el tofu, excelentes alimentos para personas intolerantes a la lactosa o alérgicas a la proteína láctea.
- Por su composición lipídica, se obtienen derivados como la lecitina, utilizada como ingrediente por la industria agroalimentaria.¹

3.-Origen y difusión

De origen asiático, la soja cultivada es nativa del este asiático, probablemente originaria del norte y centro de China. Hacia el año 3000 AC los chinos ya consideraban a la soja como una de las cinco semillas sagradas. Su producción estuvo localizada en esa zona hasta después de la guerra chino-japonesa (1894-1895), época en que los japoneses comenzaron a importar tortas de aceite de soja para usarlas como fertilizantes. Es el alimento fuerte de los pueblos del oriente. En India se la promocionó a partir de 1935. Las primeras semillas plantadas en Europa provenían de China y su siembra se realizó en el Jardín des Plantes de París en 1740. Años más tarde (1765)

¹ Consulta Saludalia, Nutrición: La Soja, en Internet: <http://www.saludalia.com/nutrición/soja> (Abril 2013)

se introdujo en América (Georgia, EE.UU.) desde China, vía Londres. Sin embargo, no fue hasta la década del 40 donde se produce la gran expansión del cultivo en ese país, liderando la producción mundial de soja a partir de 1954 hasta la actualidad. En Japón se dice: "El que tiene soja, posee carne, leche y huevo". Ha servido de alimento a los paracaidistas alemanes, en forma de tabletas durante la segunda guerra mundial y su uso está generalizado en Estados Unidos y en el Brasil.

3.1.-Inicio y desarrollo del cultivo en Argentina

Las primeras plantaciones de soja en Argentina se hicieron en 1862, pero no encontraron eco en los productores agrícolas de aquellos años. En 1925, el Ministro de Agricultura Le Bretón, introdujo nuevas semillas de soja desde Europa y trató de difundir su cultivo, conocido en esa época entre los agrónomos del Ministerio como arveja peluda o soja hispida . Hacia 1956 en la Argentina no se conocían aún los aspectos básicos de la soja como cultivo. Los fracasos en la implantación hicieron que fuese considerada para esa época como cultivo "tabú". La primera vez que Argentina exportó soja fue el 5 de Julio de 1962, a través del buque "Alabama", que partió en esa fecha llevando en su interior 6.000 toneladas con destino a Hamburgo (Alemania). Su producción se incrementó notoriamente en los años 70 hasta alcanzar en la actualidad más de 6.000.000 de hectáreas cosechadas con una producción de más de 11.000.000 de toneladas, convirtiendo a la Argentina en el cuarto productor mundial de grano, el primer exportador mundial de aceite de soja y el segundo de harina de soja. No debe sorprender, entonces, que la soja represente en la actualidad el rubro de exportación de mayor incidencia en el Producto Bruto Agropecuario del país, y el mayor generador de divisas. Actualmente el cultivo de soja ocupa una amplia zona ecológica

que se extiende desde los 23° (en el extremo norte del país) a los 39° de latitud sur, concentrándose principalmente en la Región Pampeana.²

4.-Estadios fenológicos externos del cultivo de soja

En el año 1977 Fehr y Caviness crearon una escala que clasificó los distintos estadios de crecimiento que atraviesa el cultivo de la soja. A cada estado le han asignado un número. La numeración de estado vegetativo se determina mediante el recuento de los nudos existentes por encima del tallo principal que tienen o han tenido hojas totalmente desarrolladas. Se considera que una hoja ya está totalmente desarrollada cuando los bordes de los folíolos de la hoja ubicada inmediatamente superior no se tocan. Las etapas reproductivas se basan en la floración, en el crecimiento de las vainas y semillas y en la madurez de las plantas. La designación de cada fase se designa con una letra R seguida de un número y por una explicación resumida de cada estadio.³

Ve emergencia: cotiledones sobre la superficie del terreno

Vc Cotiledonar: hojas unifoliadas suficientemente desenrolladas de forma que sus bordes o laterales no se tocan

V1 Primer Nudo: hojas totalmente desarrolladas en el nudo de las hojas unifoliadas

V2 Segundo Nudo: dos nudos sobre el tallo principal con hojas totalmente desarrolladas comenzando por el nudo de hojas unifoliadas

Vn N° nudos: número de nudos sobre el tallo principal con hojas totalmente desarrolladas comenzando por el nudo de hojas unifoliadas.

R 1 Comienzo de Floración: una flor abierta en algún nudo del tallo principal

² Consulta Cultivo de soja, en Internet:

http://es.wikipedia.org/wiki/Cultivo_de_soja#Inicio_y_desarrollo_del_cultivo_en_Argentina (Abril 2013)

³ Ver Anexo, punto I.

R 2 Plena Floración: una flor abierta en uno de los nudos superiores del tallo principal con una hoja totalmente desarrollada.

R 3 Comienzo de Fructificación: vainas de 5 mm de largo en alguno de los cuatro nudos superiores del tallo principal con una hoja totalmente desarrollada

R 4 Plena Fructificación: vainas de 2 cm de largo en alguno de los cuatro nudos superiores del tallo principal con una hoja totalmente desarrollada

R 5 Comienzo de llenado de granos: granos de 3 mm de largo en una vaina en algunos de los cuatro nudos superiores del tallo principal con una hoja totalmente desarrollada.

R 6 Tamaño máximo de granos: las vainas tienen semillas verdes que llena completamente la cavidad del fruto, en algunos de los cuatro nudos superiores del tallo principal con hojas totalmente desarrolladas

R 7 Comienzo de Madurez: alguna vaina normal sobre el tallo principal ha alcanzado su color típico de madurez.

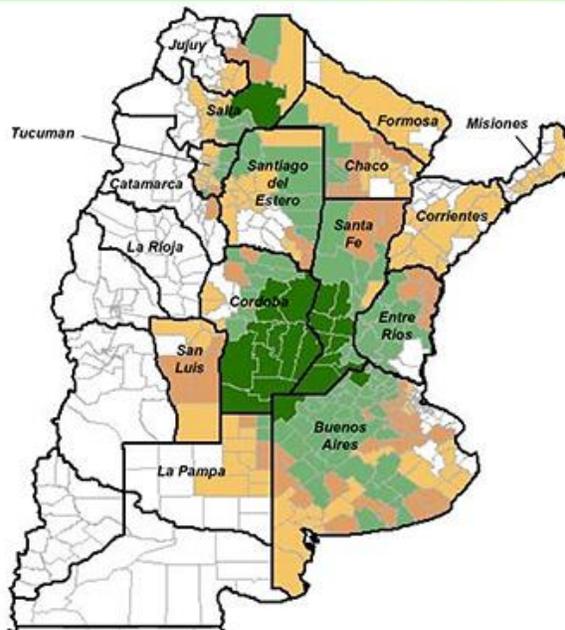
R 8 Madurez Comercial: 95 % de vainas con el color típico de madurez.⁴

5.-La importancia de la soja para Argentina

Es el cultivo que más ha crecido en los últimos 25 años y tiene una característica distintiva de los otros granos: se exporta en su casi totalidad (alrededor del 96% de la producción) ya sea como poroto, harina o pellets, aceite y biodiesel, siendo ésta una de las razones que explican que el productor siga sembrando cada vez más soja mientras disminuye en parte la siembra de otros cultivos, como trigo y el maíz, cuyos registros de exportación se han cerrado en repetidas ocasiones.

⁴ Consulta [Cultivo de Soja. Escala de los principales estados fenológicos](http://es.wikipedia.org/wiki/Cultivo_de_soja), en Internet: http://es.wikipedia.org/wiki/Cultivo_de_soja (Abril 2013)

Soja en Argentina



* % de Producción por Provincias	
Cordoba	29
Buenos Aires	27
Santa Fe	24
Entre Rios	7
Santiago del Estero	4
Salta	3
Chaco	3
Tucuman	2
La Pampa	1
Otras Provincias	~1

* Promedio 2005/06 a 2009/10
Fuente: SAGPyA



Imagen⁵

El centro del cultivo se encuentra cerca de Rosario; De todas maneras, a partir de ese centro, el cultivo se ha ido implantando en vastas regiones del noreste y noroeste argentino.

A raíz de que la producción sojera se centra en el área mencionada más arriba y que desde el año 1997 contamos con una vía navegable desde Puerto San Martín (km 457 del río Paraná) al mar, y contando toda la zona portuaria al norte y sur de la ciudad de Rosario con casi veinte puertos privados eficientes, se ha instalado una industria de crushing de soja muy poderosa.

⁵ Consulta [Mapas Productivos-Argentina](http://www.fyo.com/especiales/soja12-13/mapa.html), en Internet: [http:// www.fyo.com/especiales/soja12-13/mapa.html](http://www.fyo.com/especiales/soja12-13/mapa.html) (Abril 2013)

Lo que diferencia a la estructura de crushing de soja de nuestro país con respecto a los otros países, en especial en referencia a la exportación, es lo siguiente:

a) Las plantas argentinas están situadas en la zona de producción y a la vera del río Paraná. Ninguno de los otros países cuenta con esta ventaja.

b) Las plantas argentinas son las más modernas y se han instalado en dos oleadas: hacia la mitad de la década pasada y de ésta.

c) En conjunto totalizan alrededor de 125.000 toneladas diarias que hacen de la zona rosarina el principal polo de crushing de soja del mundo.⁶

5.1.-Evolución

La soja es el producto que mayor crecimiento ha tenido en los últimos años en el mercado argentino, ganando terreno frente a cultivos más tradicionales, como trigo, girasol o maíz. El área sembrada se incrementó desde las 11,6 millones de hectáreas en 2001/2002 a las 18,8 millones de 2011/2012, y la producción creció en aproximadamente 10 millones de toneladas en ese período, aunque los resultados fueron bajos en esta última campaña por efecto de la sequía.

Las razones van desde la buena adaptación a los suelos y la incorporación de tecnología, hasta los buenos precios del mercado internacional.

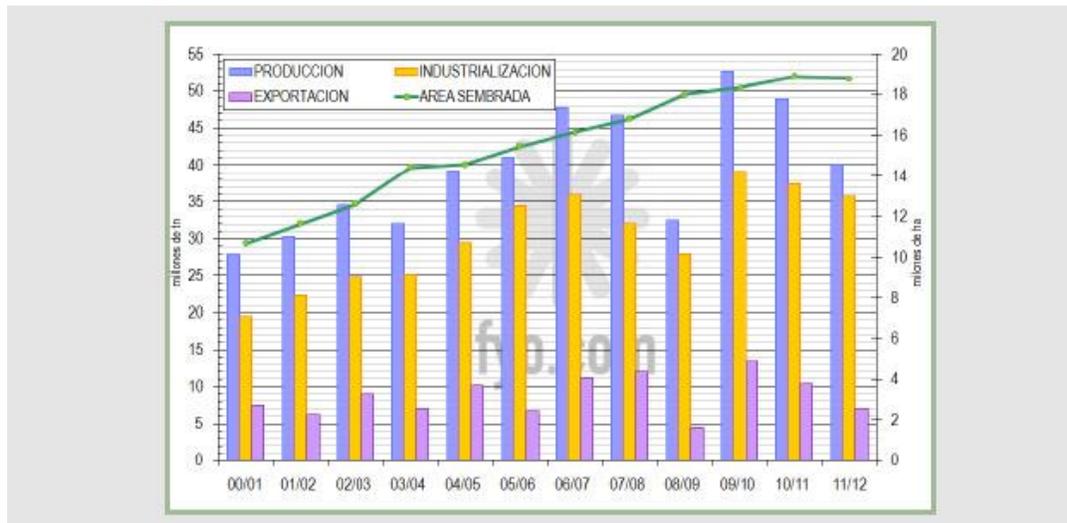
Por otro lado, desde 2008 es el único producto que refleja los precios del mercado externo. A diferencia del maíz y el trigo, que a partir de la intervención en los mercados perdieron dicha referencia –ya que hay momentos del año en que no hay exportación– la soja goza de una demanda constante durante todo el año, tanto por parte de los exportadores como de la industria.

⁶PONTÓN, Rogelio, La importancia de la soja para Argentina, en Bolsa de Comercio de Rosario, Capítulo 22, pág.1-3.

El gráfico muestra el incremento casi continuo del área de siembra, pero con resultados que no tienen el mismo desarrollo en el aspecto productivo, ya que existen campañas donde la producción fue menor que el ciclo precedente.

En las campañas 2003/2004, 2008/2009 y la actual 2011/2012, el aumento en el área de cobertura no se trasladó a la oferta por el clima adverso que acompañó el desarrollo de los cultivos.

Toda caída en la producción tiene primeramente un impacto en el volumen de exportaciones de poroto y luego en la industrialización, provocando esos años que el sector desarrolle su actividad con capacidad ociosa.⁷



⁷ Consulta La Soja en la Argentina-Evolución, en Internet: <http://www.fyo.com/especiales/soja12-13/mapa.html> (Abril 2013)

6.-La reina de los commodities agrícolas

La soja es el producto agrícola de mayor crecimiento en los últimos años: la oferta pasó de 185 millones de toneladas en el 2001/02 a 238 millones de toneladas en el 2011/12, mientras que la demanda de la oleaginosa creció de 184 millones de toneladas a 254 millones estimadas para 2011/12, es decir, un 28% en 10 años.

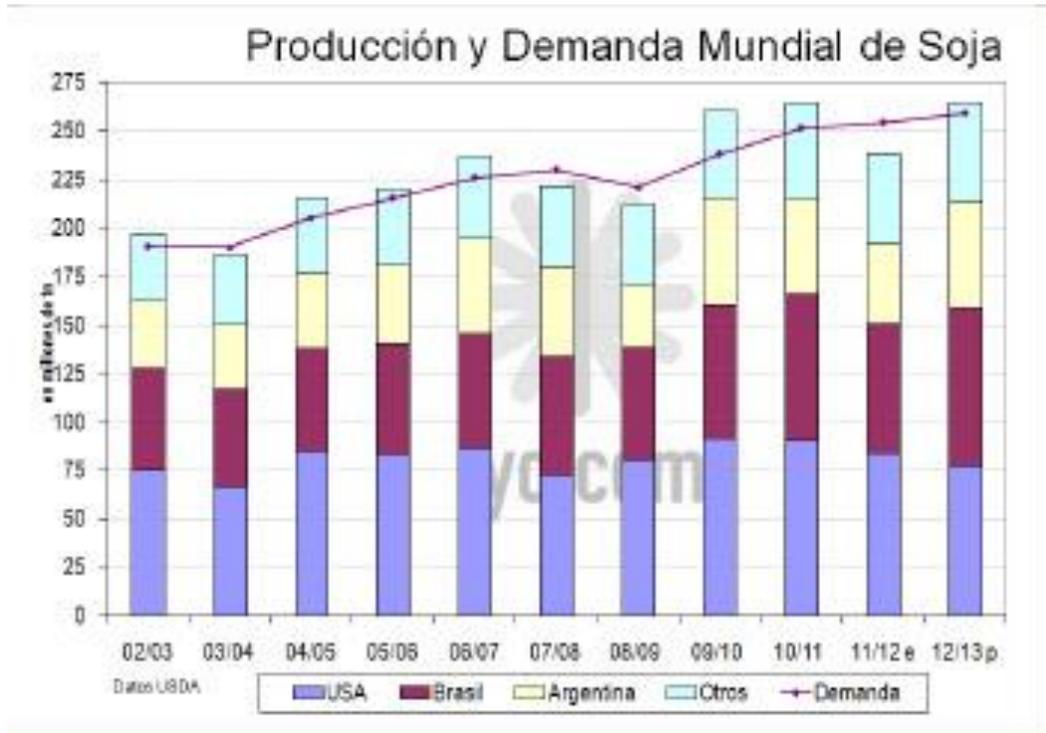
En el mismo período, China pasó de importar algo más de 10 millones de toneladas a 58 millones, creciendo un 83% en el período y participando con el 64% del comercio mundial.

Por el lado de la oferta, la producción creció notoriamente por la expansión en Sudamérica y concentró la oferta en tres grandes jugadores: Estados Unidos, Brasil y la Argentina.

Estados Unidos sigue liderando la producción mundial pero sin experimentar incrementos importantes en los últimos 10 años (con 78 millones de toneladas producidas en 2001/02 y 84 millones en 2011/12), mientras que la producción sudamericana muestra un sostenido crecimiento.

Brasil pasó de 43,5 millones de toneladas a 66,5 millones entre dichas campañas, mientras que Argentina tuvo una producción de 30 millones de toneladas en 2001/02 y de 40 millones en 2011/12 (habiendo alcanzado los 54 millones en la campaña 2009/10). Es decir, que mientras Estados Unidos incrementó su producción en 7,7%, Brasil lo hizo en un 52,8% y la Argentina en un 25% en los últimos 10 años.

6.1.-Producción y demanda mundial



En el gráfico donde se detalla la producción mundial por países participantes del mercado, se observa el crecimiento de aquellos provenientes del hemisferio sur.

Si bien Estados Unidos continúa predominando hasta 2011/2012, los países sudamericanos (Brasil, Argentina y Paraguay) sumados desde la campaña 2001/2002 superan la oferta estadounidense.

De esta forma, la caída en las cosechas de cualquiera de ellos repercute en la oferta mundial, que no puede abastecer la demanda, que disminuye por la suba de precios o absorbe parte de los stocks acumulados.

La participación de Sudamérica ha crecido en los últimos años pero sin dejar de ser el mercado de Estados Unidos (Chicago) el referente en la formación de los precios.

6.2.-Cotizaciones, precios y mercados

6.2.1.-M.A.G. y P. de la Nación - Valor de mercado

Resolución 42/07 - En pesos por tonelada año 2013⁸

MES	DIA	SOJA	ACEITE DE SOJA
Abril	3	1651	3482
	4	1582	3419
	5	1574	3370
	8	1556	3377
	9	1575	3445
	10	1598	3486
	11	1600	3508
	12	1598	3493
	15	1624	3487
	16	1597	3410
	17	1651	3488
	18	1680	3509
	19	1685	3484

⁸ Consulta Bolsa de Cereales, Cotizaciones, Valores de Mercado, en Internet:
<http://www.bolcereales.com.ar/>- Valores de Mercado (Abril 2013)

6.2.2. Consulta Cotización de Precio. Cámara arbitral de Rosario⁹

<u>Fecha de Operación</u>	<u>Producto</u>	<u>Precio Fijado</u>
03/04/2013	Soja	S/C
04/04/2013	Soja	\$ 1588.00
05/04/2013	Soja	\$ 1579.00
08/04/2013	Soja	\$ 1598.00
09/04/2013	Soja	\$ 1608.00
10/04/2013	Soja	\$ 1599.00
11/04/2013	Soja	\$ 1619.00
12/04/2013	Soja	\$ 1643.00
15/04/2013	Soja	\$ 1625.00
16/04/2013	Soja	\$ 1649.00
17/04/2013	Soja	\$ 1663.00
18/04/2013	Soja	\$ 1704.00

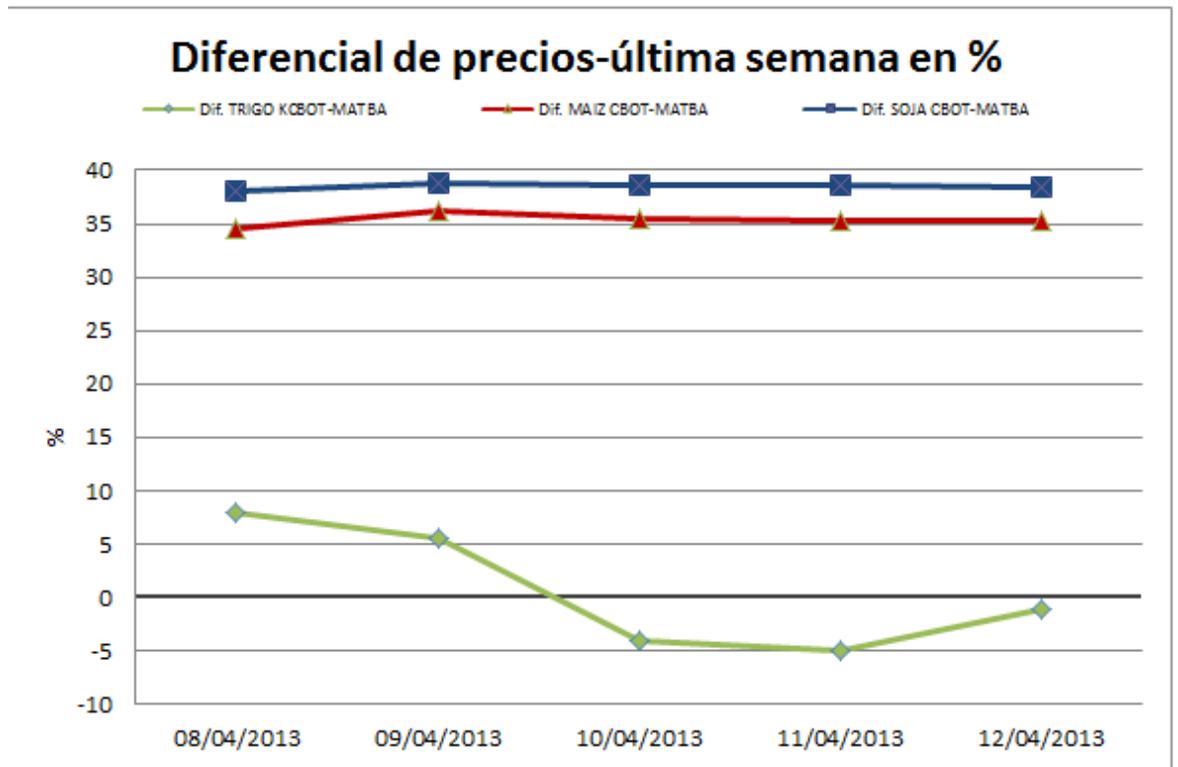
6.2.3.-Diferenciales soja

Fuente: MATBA, CBOT y KCBOT 1ra posición¹⁰

Dif. SOJA CBOT- MATBA	Soja Chicago	Soja Rosario	Dif %
08/04/2013	506,3	314,0	38,0
09/04/2013	512,8	314,0	38,8
10/04/2013	511,7	314,0	38,6
11/04/2013	515,1	316,5	38,6
12/04/2013	519,2	320,0	38,4

⁹ Consulta Bolsa de Comercio de Rosario, Mercado de Granos, Cotizaciones, en Internet: <http://www.bcr.com.ar/Pages/Publicaciones/infoboletinsemanal.aspx?IdArticulo=117>

¹⁰Consulta Bolsa de Cereales, Informe Semanal, Semana 15/04/2013, en Internet: <http://www.bolcereales.com.ar/ ESTUDIOS ECONOMICOS- Informe semanal, semana 15/04/2013>



6.2.4.-Análisis: Semana con subas para la soja (semana del 15/04/13 al 19/04/13)

El precio de la soja ha avanzado en la semana tanto en el mercado local como en el externo, impulsado por la necesidad de mercadería del consumo. El precio Pizarra de referencia para mercadería disponible que se dio a conocer el día de la fecha (y referido a las operaciones del jueves 18 de Abril del 2013) fue de \$ 1.704/tonelada, un 3,7% por encima del viernes anterior. Los precios disponibles escuchados en el recinto, por su parte, luego de avanzar a 1.700 pesos la tonelada el día jueves, cerraron la semana corrigiendo parte de la suba hasta \$1.670/tonelada.

Dicho precio aún resulta 50 pesos más alto del cierre del viernes anterior, y se manejaban rumores acerca de que en condiciones puntuales y para grandes lotes se podrían haber ofrecido hasta \$1.700/tonelada. En el plano interno, la decisión oficial conocida en la semana de llevar del 7% al

10% la participación del biodiesel en la mezcla de gasoil podría eventualmente actuar como factor de soporte adicional al mejorar las perspectivas de la demanda.

En Estados Unidos la semana cierra con una mejora del 0,3% luego de ajustar el viernes a u\$s 521/tonelada en el contrato con vencimiento en mayo, impulsados por la escasez de oferta tanto de poroto como de harina de soja. En efecto, en el mercado el ajustado balance de oferta y demanda global ante una demanda que no se serena se hace sentir.

Con respecto a la situación de la cosecha argentina, si bien la soja nueva ha comenzado a llegar al mercado, la necesidad de materia prima era imperiosa. Hasta tanto no se reanude el ritmo normal de venta al exterior de harina de soja (donde nuestro país abastece más del 50% del mercado internacional) la presión sobre los precios continuará actuando como sostén para este subproducto.

En lo que hace a la colocación del poroto, datos oficiales indican que el 26% de la producción prevista por GEA (Guía Estratégica para el Agro) ya se ha vendido, lo que representa 12,35 millones de toneladas. De ellas, 7,1 han sido colocadas bajo la modalidad "a fijar precio" mientras que las restantes 5,25 ya tienen precios en firme.

El grado de avance de la comercialización medido al 10 de abril muestra aún un retraso respecto a las campañas previas, aunque en nuestro recinto se ha registrado un fuerte impulso a los negocios en las últimas dos semanas (una estimación aproximada de 150.000 toneladas diarias la pasada y 300.000 toneladas durante la semana del 15/04 al 19/04 del 2013).

En las próximas semanas, el ritmo de avance de las ventas será un factor a seguir ya que estacionalmente comienza la época de mayor movimiento del grano, lo cual repercute a su vez en la actividad de una serie de industrias vinculadas al complejo oleaginoso.

CAPÍTULO II

Marco teórico

SUMARIO: 1.- Terminología a utilizar 1.1- Sesgo de Selección 1.2- Sesgo de Medición 1.3- El censo y La Encuesta 1.4- Muestreo por Juicio Personal 1.5- Muestreo Probabilístico 1.5.1.-Tipos de Muestras Probabilísticas 1.5.2.-Corrección y Precisión 1.6.- Métodos de Recolección de la Información 1.7.- Muestreo con Marco Imperfecto 2.- El Diseño de Muestras 2.1.- Estratificación 2.2.- Modo de Estratificar 2.3.- Número de Estratos 2.4.- Asignación de la Muestra a Estratos 2.5.- Muestreo Sistemático 3.- Encuestas de la Producción Agrícola 3.1.-Informe de la Cosecha 3.2.-Metodo de las Cosechas 3.2.1.- Selección de Campos 3.2.2.- Localización de los Cortes Muestrales 3.2.3.-Tamaño de los Cortes 3.2.4.- Forma del Terreno 3.3.- Experimento en los Campos de los Cultivadores 4.- Diseño Muestral para Encuestas de rendimientos 4.1.- Estratificación 4.2.- Tamaño de la muestra.

1.-Terminología a utilizar

- ✓ Unidad de observación: Es el objeto por el cual se realiza una medición. Esta es la unidad básica de medición, a veces llamado elemento.
- ✓ Población objetivo: En toda investigación existe la población-objetivo, aquella acerca de la cual se desea recolectar la información. La población objetivo puede ser, por ejemplo, la totalidad de viviendas de

una ciudad en un momento dado. Es la colección completa de observaciones que deseamos estudiar. La elección de la población-objetivo afectará profundamente a las estadísticas resultantes.

- ✓ Muestra: es un subconjunto de la población.
- ✓ Población muestreada: Es la población de donde se extrae la muestra.
- ✓ Población finita: Cuando el número de objetos o unidades contenidos en ella es limitado.
- ✓ Media de la población: : Es el cociente entre la suma de los valores de la variable y el total de ellos.
- ✓ Razón de la población: Es el resultado que se obtiene al dividir una cantidad con otra ($R= a/b$). En las ciencias, este término se utiliza con un significado muy particular, cuando el numerador y el denominador se refieren a cosas distintas.

razón de sexos= (No. de hombres)/(No. de mujeres)

- ✓ Proporción de la población: Una proporción es un tipo especial de razón, cuyo numerador está incluido en el denominador [$p= a/(a+b)$]. Por ejemplo, el resultado de la operación

$$\frac{\text{No. de hombres}}{(\text{No. de hombres}) + (\text{No. de mujeres})} \quad 11$$

1.1.-Sesgo de selección

Ocurre cuando alguna parte de la población objetivo no está en la población muestreada. Es una distorsión que se introduce debido a la forma en que se selecciona la muestra. Si el sesgo muestral no es tomado en cuenta, entonces algunas conclusiones propuestas pueden ser erróneas.¹²

¹¹ LOHR, Sharon L., Muestreo: Diseño y Análisis, International Thomson Editores (México, 2000), pág. 2 y 3, passim.

¹² Ibídem, págs. 4-8

1.2.-Sesgo de medición

Son los sesgos que se producen durante el proceso de la recolección de la información, ya sea por la obtención de información incompleta o errónea o por la modificación de la muestra (o parte de ella) durante la realización del trabajo.

Sesgos debidos al encuestado: la información que éste proporciona puede ser incorrecta debido a olvido, subjetividad, confusión, desconfianza, ignorancia, incomprensión o modificación de la respuesta por la propia encuesta o medición incorrecta de parámetros.

Control: El cuestionario debe ser correcto (sin posibilidad de respuestas subjetivas, vocabulario adecuado y preciso, corto). Es especialmente importante en los estudios longitudinales: se debe mantener el interés del encuestado para evitar que abandone. Otro aspecto importante será la utilización de aparatos de medida fiables, y evitar -en lo posible- las interpretaciones subjetivas.¹³

1.3.-El censo y la encuesta

Podría pensarse que el muestreo no es un método confiable y que sólo una enumeración completa de toda la población puede considerarse. Este argumento ha perdido ahora su validez por las razones siguientes:

- En primer lugar, no hay errores de muestreo en un censo, pero el error matemático es grande. Una encuesta por muestreo estaría sujeta a un error sistemático relativamente pequeño y a un error de muestreo menos importante aun. Por lo tanto el error total de una encuesta por muestreo es a menudo mucho menor que el solo error sistemático de un censo.
- En segundo lugar, el censo no provee medida alguna del margen de error al que estén sujetos los datos, aunque pueda existir un error consi-

¹³Consulta Los sesgos y su control de Jordi Casal y Enric Mateu, en Internet: http://www.epidemio.com/epidemio/img/datos/21_06_59_3LosSesgos41.pdf

derable. No ocurre así en una encuesta por muestreo bien organizada, que puede proveer los estimadores y los errores aleatorios a que estén sujetos.

Por ultimo tenemos la consideración primordial del costo involucrado en la recolección de datos y que los resultados pueden producirse con una velocidad considerablemente mayor cuando se emplea en método del muestreo.

El censo da una oportunidad para la recolección de datos por muestreo en renglones de carácter mas refinado. Los dos procedimientos pueden combinarse. El método de muestreo puede emplearse para producir estimaciones avanzadas a partir del censo.¹⁴

1.4.-Muestreo por juicio personal

Podría pensarse que la forma más rápida de selección de una muestra es el juicio personal. La dificultad de tal procedimiento es que en forma consciente o inconsciente el muestreador tenderá a cometer errores de juicio. Una segunda desventaja consiste en que el intervalo de variación observado en tal muestra no da una buena idea de la variabilidad en la población.

La razón básica de nuestra incapacidad para manejar las muestras a juicio personal es que no podemos calcular la probabilidad de que una unidad especificada se seleccione en la muestra. Por lo tanto, no podemos determinar la distribución de frecuencias de los estimadores producidos por este proceso. En ausencia de esta información, no puede determinarse de modo objetivo el error de muestreo.¹⁵

¹⁴ DES, Raj, Las estructuras de las encuestas por muestreo, 1° Edición en Español, Fondo de cultura económica, (México, 1979), pág. 21.

¹⁵ Ibidem, pág. 24.

1.5.-Muestreo probabilístico

En una muestra de probabilidad cada unidad de la población tiene una probabilidad de selección conocida; se emplea un método aleatorio para elegir las unidades específicas que se incluirán en la muestra.

Si un muestreo de probabilidad se realiza de manera adecuada, un investigador puede utilizar una muestra relativamente pequeña para llevar a cabo inferencias de una población arbitrariamente grande.

1.5.1.-Tipos de muestras de probabilidad:

✓ *Muestra aleatoria simple:* se obtiene una muestra aleatoria simple de tamaño "n" cuando cualquier subconjunto posible de "n" unidades en la población tiene la misma probabilidad de ser seleccionado para componer la muestra.

✓ *Muestra aleatoria estratificada:* la población se divide en subgrupos llamados estratos. Al llevar a cabo esta división, se extrae una muestra aleatoria simple de cada estrato la cual se elige de manera independiente.

Los estratos son, con frecuencia, subgrupos de interés para el investigador; los elementos del mismo estrato tienden, por lo regular, hacer mas similares que los elementos elegidos al azar de la población entera, de modo que, a menudo, la estratificación aumenta la precisión.

✓ *Muestra por conglomerado:* las unidades de observación que componen una población se reúnen en unidades de muestreo de mayor tamaño llamado conglomerado. Este tipo de muestreo con frecuencia disminuye la precisión.

1.5.2.-Corrección y precisión

Una estimación calculada a partir de la muestra se llama precisa si se acerca al valor esperado. La precisión se refiere a la aproximación al valor esperado, mientras que la corrección se refiere a la aproximación al valor verdadero. El valor esperado diferirá del valor verdadero cuando los errores presentes en los datos no tengan un promedio cero. Estos errores derivan de técnicas inadecuadas de medición, cuestionarios defectuosos, conceptos mal definidos, etc. Si los errores muestrales no son importantes, la precisión y la corrección no serán diferentes, y la muestra podrá arrojar luz sobre la corrección de los resultados.¹⁶

1.6.-Métodos de recolección de la información

Hay varios métodos para la recolección de la información: cuestionario enviado por correo, entrevista personal, observación directa, o pruebas de laboratorio. La decisión relativa al método que habrá de emplearse dependerá del carácter de la información buscada y de la población encuestada.

1.7.-Muestreo con marcos imperfectos

Para la selección de la muestra debe haber un marco -un mapa, una lista, u otra descripción aceptable del material- a partir del cual puedan construirse y seleccionarse las unidades muestrales. Diremos que un marco es perfecto si cada elemento del universo ocurre una y solo una vez en él por separado, no aparecen otros elementos, y la información es correcta. Siempre hay algunos defectos. Es posible que el marco no contenga algunas de las unidades de la población (falta de plenitud), o ciertas unidades pueden ocurrir mas de una vez (duplicación). Si parte de la información contenida en

¹⁶ Ibídem, pág. 25.

el marco no es correcta, el marco es incorrecto. Es obvio que los defectos del marco introducirán errores en los resultados.

- **Duplicación:** A veces dos unidades diferentes enumeradas en el marco pueden conducir a la misma unidad informadora en la población-objeto. Esto aumenta la probabilidad de selección de las unidades duplicadas cuando se selecciona una muestra.¹⁷

2.-El diseño de muestras¹⁸

2.1.-Estratificación

Esta técnica se emplea con frecuencia en el diseño de muestras, de acuerdo a la información derivada del marco. Las unidades se reparten en estratos colocando en el mismo estrato las unidades similares respecto de la característica que se desea medir, así los estratos serán internamente homogéneos, y la variabilidad entre una unidad y otra de un estrato será pequeña, lo que significa que el error muestral deberá ser pequeño. La estratificación genera una precisión mayor que las estimaciones hechas. A su vez las estimaciones de una precisión especificada pueden requerirse también para ciertas subdivisiones de ella.

2.2.-Modo de estratificación

Se forman los estratos cortando en un número adecuado de puntos de distribución de frecuencias de las unidades en cuanto a las características buscadas.

Pero no conviene utilizar estratificaciones distintas de la misma población para características diferentes, ya que eso nos llevaría a tomar una

¹⁷ Ibíd., págs. 174-176.

¹⁸ Ibíd., págs. 191-195, passim.

muestra de tamaño excesivo. En estos casos se requiere realizar estratificaciones múltiples.

2.3.-El número de estratos

Generalmente queremos hacer tantos como podemos, ya que un estrato heterogéneo siempre puede dividirse para generar mayor homogeneidad, pero sólo podrá crearse si tenemos información suficiente sobre varias características de cada unidad de la población.

El tamaño de la muestra fija un límite al número de estratos que puedan hacerse. Por lo menos una y preferentemente dos unidades deben seleccionarse de cada estrato para poder calcular el error muestral a partir de muestras probabilísticas. Aparte de aumentar precisión, el empleo de un gran número de estratos permite una estimación estable del error muestral.

2.4.-Asignación de la muestra a los estratos

Si se quiere seleccionar al azar las unidades dentro de los estratos, la mejor asignación de la muestra estará en proporción al tamaño del estrato y a la desviación estándar de la variable y del estrato. En la práctica resulta difícil el empleo exacto de este principio, no se conoce la desviación estándar de y dentro de los estratos; la asignación óptima de una característica diferirá de la asignación óptima de otra, pero no podemos emplear asignaciones diferentes, en consecuencia suele requerirse el empleo de una asignación en transacción.

Un procedimiento empleado comúnmente consiste en hacer los estratos de igual tamaño agregado. Significa que el valor total de la variable x se hace aproximadamente igual para cada estrato; este es un procedimiento conveniente y eficiente a la vez, si los estratos así formados son internamente homogéneo y difieren considerablemente entre sí respecto de x .

2.5.-Muestreo sistemático

La selección de cada k-ésima unidad de la lista, haciendo la primera selección al azar, constituye una muestra sistemática.

Se emplea ampliamente por su comodidad; resulta fácil verificar si la selección de la muestra cumple con las instrucciones. Hay ciertas situaciones en que tendera a dar una precisión mayor que en el muestreo al azar; esto ocurrirá cuando existe en la lista una tendencia respecto de la característica de interés y.

La selección sistemática centrada -la selección de la unidad central de las primeras k unidades y de cada k-ésima unidad en adelante- puede ser mejor todavía.

El muestreo sistemático funciona peor cuando hay periodicidad en los datos muestreados y el intervalo muestral se acomoda a ella; así la mayoría de las unidades de la muestra serán demasiado bajas o demasiado altas, lo que vuelve muy variable la estimación.

3.-Encuestas de la producción agrícola¹⁹

El problema de la estimación de la producción anual de cultivos alimenticios y de otros productos agrícolas tiene gran importancia para todos los países, sobre todo para países en desarrollo, que hacen grandes esfuerzos por alimentar su población y mejorar sus niveles de vida. Pero los cultivos que siembran millones de agricultores por todo el país plantean un problema formidable.

Un método de estimación de la producción total de tales cultivos consiste en estimar el área total bajo cultivo y la tasa de rendimiento; el pro-

¹⁹ Ibíd., pág. 250-258, passim.

ducto de los dos componentes es una estimación de la producción total del cultivo.

3.1.-Informes de cosecha

Las estadísticas de rendimientos agrícolas se basan en informes periódicos de informadores de cosechas, que pueden ser agricultores o funcionarios públicos.

Los informadores basan los informes en su propia evaluación y juicio de la condición de la cosecha. El método es muy subjetivo y está sujeto a sesgos desconocidos.

3.2.-El método de las cosechas

El método objetivo de la estimación de la tasa de rendimiento consiste en seleccionar una muestra de parcelas o campos al azar, y emplear en la observación el método de la medición física del rendimiento en el momento de la cosecha.

El método habitual consiste en elaborar una lista de unidades de primera etapa; se selecciona una muestra de unidades (provincias) y se elabora una lista de los campos donde se siembre el cultivo en cuestión para cada unidad de la muestra. Se toma una muestra de campos y se marca el terreno al azar en el campo seleccionado. Se cosecha el terreno y se pesa el producto una vez secado.

Las encuestas de diversos países han demostrado que éste es un método practicable, capaz de producir estimaciones de rendimientos libre de sesgo y con un alto grado de correlación. Pero es un método caro, que requiere el empleo de un gran número de encuestadores adiestrados.

3.2.1.-Selección de campos

Las encuestas de rendimientos agrícolas se realizan, de ordinario, cultivo por cultivo. La razón es que diversos cultivos maduran en épocas diferentes.

La muestra de un cultivo en particular sólo puede tomarse cuando la cosecha está madura; esto significa que solo hay un periodo muy breve, anterior a la cosecha efectiva, en que el campo está listo para el muestreo, por lo que hay una gran dificultad para asegurarse de que todos los campos sean visitados en el momento adecuado. La desventaja de este método es que los campos de la muestra se conocen por adelantado.

Existe otro método donde los campos de la muestra no son conocidos por adelantado. En este método se toma una muestra de agrupamientos de campos y se determinan los campos bajo el cultivo en cuestión que todavía no han sido cosechados y que están listos para el muestreo. Se toma una muestra de dos campos, y se determina el rendimiento por sub-muestreo de los campos. La dificultad aquí es que los cultivos listos para la cosecha en momentos diferentes están representados desigualmente en la muestra. Esto sólo puede evitarse si se ajusta el muestreo de modo de que su distribución en el tiempo corresponda a la distribución en el tiempo de la cosecha efectiva de los campos. En segundo lugar, no es constante el intervalo que media entre el momento en que el cultivo se considera listo para la cosecha y el momento efectivo de la cosecha. El sesgo involucrado sólo podrá descartarse si puede demostrarse que no hay correlación entre el rendimiento y este intervalo.

3.2.2.-Localización de los cortes muestrales

Para estimar el rendimiento de un campo seleccionado, deberá localizarse al azar en el campo un terreno o un corte muestral de forma y tamaño adecuados.

El corte muestral se localiza de ordinario tomando números al azar, digamos x e y . El encuestador camina x pasos a lo largo del campo e y pasos en dirección perpendicular. El punto alcanzado se considera como la esquina sudoeste del terreno, en caso de terrenos circulares se considera el centro del círculo. Así ciertas partes cercanas a los límites no tendrán oportunidad de selección, por lo que tenderá a colocarse un sesgo en los resultados si las áreas limítrofes difieren de las otras partes del campo en cuanto al rendimiento- sesgo del límite-.

3.2.3.-El tamaño de los cortes muestrales

Se ha considerado a fondo la cuestión del tamaño del terreno y de su forma.

Los terrenos pequeños tienen la ventaja de que pueden cosecharse rápidamente, el propio investigador puede trillar y pesar el producto. Pero hay un peligro obvio de sesgos debido a la localización o demarcación deficiente de los terrenos, si los encuestadores no están lo suficientemente calificados o los instrumentos empleados no son adecuados. El encuestador tiende a incluir plantas del límite del terreno, produciendo una sobre-estimación del rendimiento; por lo tanto resulta peligroso el empleo de terrenos muy pequeños en la estimación del rendimiento de cultivos en ciertas condiciones.

3.2.4.-La forma del terreno

Consideramos ahora, si la forma del terreno tiene alguna relación con la subestimación o sobre-estimación del rendimiento.

Al parecer los terrenos circulares están menos sujetos a sesgos que los terrenos triangulares, y se puede deber a que el sesgo deriva de la tendencia humana a incluir plantas del límite del corte que no le pertenecen en realidad; es así que la sobre-estimación es producida por perturbaciones en el límite. Ahora bien, para un área dada, el círculo tiene el perímetro más pequeño y por lo tanto la perturbación es mínima en este caso.

3.3.-Experimento en los campos de los cultivadores

El rendimiento obtenido en un corte muestral se obtiene de ordinario en condiciones controladas. Esto es diferente de las condiciones en que el agricultor cosechara sus campos. El resultado es que la estimación del rendimiento obtenido en la encuesta no será comparable con el rendimiento obtenido en la realidad por el agricultor. Por lo que los resultados de las encuestas deben corregirse. Para ello se pide a los agricultores que cosechen los campos empleando las herramientas y maquinarias usadas normalmente; esto se hace en una muestra de campos de los que se haya tomado los recortes. Una comparación entre las dos cifras nos provee el factor de corrección.

A veces se disponen de cifras de comercialización que permiten estimar los rendimientos comerciales en una sub-muestra de campos muestreados. Estas cifras pueden emplearse para hacer una comparación con los rendimientos muestrales correspondientes a fin de determinar el factor de corrección o para verificar su exactitud.

En otro enfoque se hace que los procedimientos de las encuestas se asemejen en la mayor medida posible a los procedimientos empleados en efecto por los agricultores. Se espera que la corrección involucrada sea menor en este caso, pero se requieren investigaciones para determinar el monto de la corrección, ya que las practicas de los agricultores no son uniformes en todo el país.

4.-Diseño muestral para encuestas de rendimientos

El diseño muestral de una encuesta de corte de cosechas solo puede decidirse en función a la información sobre el arreglo administrativo del país, equipo de campos disponibles, prácticas agrícolas locales, peculiaridades de los cultivos y de la información auxiliar.

4.1.-Estratificación

Las diversas sub-divisiones administrativas del país deben formar los estratos principales de la encuesta. Cada estrato deberá sub-dividirse en estratos que se esperan difieran considerablemente entre sí respecto de la tasa de rendimiento.

El muestreo dentro de un estrato se hace, de ordinario, en forma independiente de los demás. Se selecciona una muestra de unidades en el estrato con probabilidad proporcional al área cultivada de las unidades. Cuando se conoce el número total de unidades que se seleccionara en toda la muestra, el número asignado a un estrato puede basarse en la proporción que el área bajo cultivo en el estrato representa del área total bajo cultivo, aproximándonos a una asignación optima.²⁰

²⁰ Ibídem, pág. 259

4.2.-Tamaño de la muestra

El número de unidades que se seleccionara de un estrato principal estará representado por el nivel de precisión necesario para hacer estimaciones para el estrato.

Un aspecto práctico, que suele ser importante, es como conseguir el valor de σ^2 o alguna buena estimación del mismo, cuando no se conoce de encuestas o censos o estadísticas permanentes que traigan esta información. Esto suele ser frecuente en lugares donde hay pocas estadísticas y poca experiencia estadística.

La respuesta es que en esos casos debe tratar de conseguirse la mejor estimación o indicio sobre σ ; para ello, se puede tratar de hacer una muestra piloto, y junto a ello indagaciones personales de quien diseñará la muestra ante personas que están en el ambiente o conocen el tema a donde será dirigida la encuesta, aunque no tengan idea alguna sobre estadística. La tarea del encargado de la encuesta será la de buscar indicios opiniones, datos, etc. Evaluado todo ello, hacer un cálculo de la variabilidad de la variable que uno desea estudiar.

Al extraer una muestra, se pregunta por información referente a varias variables y en ese caso, se tiene que estimar y presentar intervalos de confianza para todas esas variables, y se pretende fijar cierto margen específico de precisión para cada una de ellas.

Sean las variables x , y , z , la primera de ellas puede tener una variabilidad diferente a las de la variable y y z y en ese caso, por ejemplo:

$$\sigma_x > \sigma_y > \sigma_z$$

¿Qué variabilidad se toma para fijar el tamaño de la muestra? Si se toma σ_x dado que es la variable que tiene mayor variabilidad, el tamaño de la

muestra será excesivo para las otras variables; si tomamos σ_z , el tamaño de la muestra será escaso para las otras variables que se desean estudiar. Si se calcula los tamaños posibles de la muestra, se obtendrá luego de aplicada la formula:

$n_x > n_y > n_z$ Para un determinado nivel de confianza.

Si los diferentes valores de los tamaños de muestras para las tres variables (n) son cercanos entre si, se tomará un valor promedio o cualquiera de ellos. Pero lo más frecuente, es que haya discrepancias importantes entre ellas, y en ese caso también entra en consideración el costo de la encuesta.

Un tamaño grande de muestra implica un costo alto y al mismo tiempo una precisión innecesaria en alguna de las variables. En ese caso, se suele eliminar aquellas variables que exijan un tamaño relativamente muy superior a las de las otras variables.

La consecuencia práctica de esto, es de extrema importancia ya que implica que al diseñar la muestra y cuestionario, habrá que estudiar críticamente las variables a incluir, y la consideración de sus variabilidades indicará:

1) qué variables incluir

2) qué variables excluir

y diseñar una encuesta paralela especial para aquellas variables con alta variabilidad, con miras a obtener una reducción en el tamaño de muestra requerido.²¹

²¹ MEDINA, Ramón Darío, Técnicas Muestrales, Publicación XIII, Cátedra de Estadística, Facultad de Ciencias Económicas, UNT, (San Miguel de Tucumán, Abril 2010), passim.

CAPÍTULO III

Propuesta de diseño muestral para estimación de rendimientos en la producción de soja

SUMARIO: 1.- Introducción 2.- El procedimiento del muestreo
2.1.- Errores muestrales y no muestrales 2.2.- Medida del error muestral 3.- Planteamiento del problema 4.- Revisión de métodos de estimación empleados 4.1.- Estimaciones visuales y estimación por muestreo: críticas y ventajas 5.- Estimación de encuestas por muestreo 5.1.- Formulación de objetivo 5.2.- Población y unidades investigadas 5.3.- Principales variables investigadas por la encuesta 5.4.- Grado de precisión 5.5.- Período de referencia 5.6.- Presupuesto 6.- Marco muestral 6.1.- Diseño y selección de muestra 6.2.- Métodos de muestreo 6.3.- Tamaño de la muestra 6.4.- Diseño de instrumentos de recolección 6.5.- Organización del trabajo de campo 6.6.- Recursos humanos y capacitación 6.7.- Validación y crítica 7.- Grado de precisión de las estimaciones-Errores de muestreo 8.- Nota al diseño muestral propuesto 9.- Conclusión.

1.-Introducción

Esta es la época de las estadísticas. Se necesitan cifras para tomar decisiones correctas. El gobierno, las empresas y los profesionales buscan la base fáctica más amplia posible para la toma de decisiones. En ausencia de

datos sobre el tema, la decisión es como un salto en la oscuridad. Se necesitan estadísticas, en realidad estadísticas cada vez mejores.²²

2.-El procedimiento de muestreo

La encuesta por muestreo es reconocida como un instrumento organizado para conocer los hechos. Su importancia reside en que puede emplearse para resumir, como guía para la administración, hechos que de otro modo serían inaccesibles debido por ejemplo a la abundancia de las unidades involucradas. El “diseño de encuesta” incluye todos los preparativos necesarios para la iniciación de la misma, la conducción efectiva de las operaciones, el procesamiento subsecuente de los datos y la redacción del informe.²³

La idea básica es sencilla. Se necesita información acerca de una población de objetos: en nuestro caso el conjunto de todas las plantas de soja de un lote, finca o zona. Examinamos solo algunos objetos, una muestra o subconjunto de la población: algunas plantas por ejemplo y extendemos nuestros hallazgos a todo el lote, finca o zona, o sea para estimar la información en todo el dominio de interés. Se toma una muestra de plantas para inferir acerca de toda la zona plantada. No es esencial la enumeración de cada unidad del universo para obtener una cifra aceptable para el total. Una muestra diseñada con cuidado puede proveer la información necesaria y ello es más probable si se dispone de un marco para el muestreo.

Hay tres elementos en el proceso: selección de la muestra, recolección de datos y formulación de una inferencia acerca de la población (la plantación de interés) y no pueden considerarse en forma aislada, están entrelazados y cada uno de ellos tiene un efecto sobre los otros. El muestreo no es una selección caprichosa; incorpora reglas definidas para la selección de la

²² DES, Raj, Op. Cit., pág. 17.

²³ Ibíd., págs.. 9-13, passim.

muestra. Aun habiendo seguido tales reglas, el proceso de estimación no es independiente de ello pues se guía por la forma en que se seleccionó la muestra.²⁴

2.1.-Errores muestrales y no muestrales

La mayor parte de las encuestas informan de un margen de error. El margen de error es una expresión del error de muestreo, el cual resulta al considerar una muestra y no de examinar toda la población. En el caso de un dado lote, finca ó zona por ejemplo, si consideramos una muestra distinta (plantas de soja diferentes a las cosechadas en el muestreo, del mismo o diferente lote), es muy probable que obtengamos una producción promedio de vainas y granos por planta distinta de la efectivamente seleccionada. Este error de muestreo es posible reducirlo pero no eliminarlo, a menos que se cosechen todas las plantas del lote, finca o zona (población objetivo). Este camino no es práctico ni económico y entonces habrán de buscarse los mecanismos para reducir y controlar el error muestral a niveles aceptables a efectos de tomar decisiones afectadas con menos incertidumbre. El diseño y el tamaño muestral son aspectos que inciden en este tipo de error, por lo que es de relevancia tomar decisiones al respecto que traten de minimizarlo. En este sentido el análisis de datos previos y el trabajo en campo ayudarán a una solución aceptable del problema.

Las estimaciones realizadas a través de una encuesta por muestreo pueden estar afectadas por errores muestrales y no muestrales. “Por supuesto, los errores de una encuesta dependen considerablemente de las condiciones esenciales en que se conduzca. Si los recursos son escasos y no se dispone de personal calificado, los errores son grandes; éstos errores pueden reducirse si se dispone de dinero y de la clase de personal adecuado”.²⁵

²⁴ Ibíd., págs. 17-18.

²⁵ Ibíd., pág. 20.

Error muestral: es la diferencia entre el valor obtenido de la muestra (estimador) y el valor real de la población a la que representa (parámetro). Las características de éste tipo de error son:

❖ ***Puede medirse***

❖ ***Disminuye a medida que aumenta el tamaño de la muestra.***

A primera vista el muestreo parece algo riesgoso. La cifra obtenida de la muestra puede no ser exactamente igual al verdadero valor de la población. La razón es que la estimación se basa en una parte y no en el todo. Otra forma de decirlo es que la estimación está sujeta a errores de muestreo o a fluctuaciones del muestreo. Sin embargo éstos errores pueden controlarse y la moderna teoría del muestreo ayuda a diseñar encuestas en forma tal que los errores de muestreo se vuelvan pequeños.

Algunos autores sugieren el empleo de una muestra “representativa” como protección contra estos errores, término que no es de utilidad porque requiere conocer una población que es la que se quiere investigar.²⁶

Error no muestral: Incluye todos los errores que se pueden presentar en el transcurso de la investigación, salvo el error muestral.

Pueden provenir, entre otras razones de:

- Definición defectuosa del problema que hay que investigar.
- Mala definición del universo.
- Mala selección de la muestra.
- Mal diseño de la planilla de recolección de datos.
- No efectuar la prueba o test de la planilla.

²⁶ MEDINA, Ramón Darío, Op. Cit., pág. 18.

- Errores producidos por los encuestadores, sesgo, parcialidad, influencias, heterogénea valoración de las características observadas en el relevamiento por falta de capacitación.
- Errores de medición o conteo.
- Errores en el procesamiento de la información.
- Errores en el análisis de la información.²⁷

Características de Interés a Investigar en el Muestreo:

Dos características de la población que interesan de ordinario son:

- **El total de la población.** En el caso que nos ocupa será la **producción en toneladas de un lote**, finca o zona geográfica.
- **La media de la población:** rendimiento por ha. de lote, finca o zona.

2.2.-Medida del error muestral

Al analizar las estimaciones provenientes de una muestra se debe tener presente que las mismas están afectadas por el llamado **error muestral** y es importante conocer aproximadamente la magnitud de esos errores. Una medida de esos errores está dada por el **desvío estándar**.

Con el error de muestreo ó el desvío estándar es posible construir el intervalo numérico que tiene una cierta confianza, medida en términos de probabilidad de contener el valor verdadero que se desea estimar.

La teoría del muestreo indica que, bajo ciertas condiciones, y con una confianza aproximada del 95 %, el intervalo comprendido entre la estimación puntual de la producción promedio diaria (\bar{x}) menos 2 veces el des-

²⁷ LOHR, Sharon L., Op. Cit., passim.

vío estándar ($S_{\bar{x}}$) y la estimación puntual (\bar{x}) más 2 veces el desvío estándar ($S_{\bar{x}}$) contiene el valor verdadero que se desea estimar. Es por eso que éste intervalo se lo denomina “**intervalo de confianza del 95 %**”.

El valor del error muestral ó desvío estándar a usar para la construcción del intervalo de confianza correspondiente a la estimación de la producción y rendimiento de la planta ó unidad muestral considerada – se calcula aplicando la siguiente fórmula:

$$S_{\bar{x}} = \frac{s}{\sqrt{n}}, \text{ donde } s \text{ resulta de } s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n-1)}} \quad 28$$

En las expresiones anteriores:

X_i es el rendimiento de la planta ó unidad muestral muestreada en cada uno de las observaciones realizadas en el lote, finca o zona estudiada.

n es el tamaño de la muestra (número de plantas ó unidades muestrales definidas medidas para la estimación del rendimiento y de la producción).

\bar{x} es el rendimiento promedio calculado en la muestra de plantas o de unidades muestrales definidas para las estimaciones por mediciones (conteo de granos y peso) realizadas al efecto.

²⁸ Siempre y cuando la fracción de muestreo (n/N) sea pequeña, el factor $\sqrt{\frac{N-n}{N}}$ llamado

Corrección del error estándar debida a población finita toma valores cercanos a la unidad y el tamaño de la población como tal no tiene efecto directo en el error estándar de la media muestral ($S_{\bar{x}}$). En la práctica la corrección por finitud se puede ignorar siempre y cuando la fracción del muestreo (n/N) no exceda del 5%. El efecto de ignorar la corrección es sobre-estimar el error estándar de la estimación. En el caso bajo análisis se presenta la situación antes descrita y por ello las fórmulas presentadas no contienen la corrección por finitud.

Además del desvío estándar se puede calcular el **coeficiente de variación**, efectuando el cociente entre el desvío estándar y el correspondiente promedio de las estimaciones realizadas en el día.

$$\text{CVR} = \frac{\overline{sx}}{\overline{x}} \cdot 100$$

El coeficiente de variación nos brinda una idea de la precisión de la estimación. **Cuanto más pequeño es el coeficiente de variación, más precisa es la estimación.** Se recomienda que estimaciones con coeficientes de variación superiores al 10 % deben ser tratadas con cautela y es el usuario el que evaluará si un dato con cierto coeficiente de variación le es útil o no. Se considerará seriamente ésta advertencia en el caso de estimaciones correspondientes a muestras pequeñas, por ejemplo las correspondientes a unidades muestrales definidas en forma de conglomerados de unidades observacionales (plantas) con cvr superiores al 10 %, especialmente en caso de extensiones de terreno pequeñas.

3.-Planteamiento del problema

El diseño muestral de una encuesta de éste tipo requiere considerar información sobre el arreglo administrativo de la zona afectada al cultivo de interés, del equipo de campo disponible, las prácticas agrícolas, las particularidades del cultivo a muestrear y de la información adicional que se disponga.

La Información sobre las diferentes variables que afectan la producción de soja es importante porque hay características que están íntimamente vinculadas a la productividad de éste cultivo y de los lotes. Zona, región o subregión de implante, tipo de cultivar de soja (longitud de ciclo, porte de

planta, tolerancia al estrés hídrico y a las enfermedades de la sequía, estado fenológico), fecha de siembra, condiciones ambientales en el periodo de cultivo y especialmente durante el llenado de granos (cantidad y distribución de disponibilidad hídrica, radiación solar y temperatura en la etapa reproductiva, tipo de terreno, contenido de agua almacenada en suelo a la siembra), comportamiento al vuelco, grupo de madurez, arreglo espacial (densidad de plantas a siembra, espaciamiento entre surcos y altura de plantas), contenido de nitratos en el suelo a la siembra, micro-relieves del lote, sanidad ó control de problemas sanitarios (malezas, plagas y enfermedades), duración e intensidad de heladas antes del llenado de granos, número de nudos y largo de entrenudos de la planta, número de vainas y número de granos por vaina, la producción de biomasa vegetativa, índice de área foliar y altura, etc. son los principales factores asociados al rendimiento final del cultivo. Es útil entonces el conocimiento de los valores o categorías de las diferentes variables asociadas a los lotes de las diferentes fincas, a la hora de proceder a la estimación de la producción de cada uno de ellos.

El método de muestreo para la estimación del rendimiento de un cultivo se puede realizar mediante la recolección de frutos de plantas antes de la cosecha y/o mediante una estimación visual. Es muy importante, por tanto, la conducción de una serie de encuestas piloto para el desarrollo y la prueba de métodos. Esta tarea sólo puede ser cumplida con eficacia por un grupo de personas que pueda dedicarse a tiempo completo y utilice la experiencia acumulada.

4.- Revisión de métodos de estimación empleados

4.1.-Estimaciones visuales y estimaciones por muestreo: Críticas y ventajas

Las estimaciones visuales de la producción involucran el riesgo de la subjetividad del estimador, que se complica cuando mayor es el número de personas que realizan tal tarea. ¿la estimación de quién debe ser mas valorada?, ¿todas las estimaciones deben ser igualmente ponderadas?, ¿qué dice la experiencia acerca del acierto de cada una de las personas que estiman visualmente?, ¿cómo se realiza efectivamente la estimación visual, se visitan todos los lotes y se valoran las condiciones observadas de las plantas a simple vista, o se trabaja solo con las plantas seleccionadas?, ¿cómo se seleccionaron las plantas que se usaron en las estimaciones?, etc.

Las estimaciones visuales pueden ser muy incorrectas pero su ventaja radica en el bajo costo y el hecho de que el análisis de los datos resulta sencillo. Si los informes están sesgados en la misma dirección, la estimación de la producción total puede estar gravemente sesgada. Pero la estimación de la tendencia durante el periodo puede no ser tan mala. Sin embargo las estimaciones visuales pueden mejorar considerablemente si se realiza una encuesta objetiva para estimar la producción, lo que también dependerá de la extensión de los lotes, pues cuanto más pequeños sean también menores serán los sesgos o errores muestrales que se cometan. ***Los errores muestrales de las estimaciones pueden reducirse en las encuestas continuas reponiendo parcialmente la muestra cuando haya una alta correlación positiva entre ocasiones sucesivas;*** para evitar este sesgo en repeticiones de encuestas en períodos de tiempo breve (año a año), se exige necesariamente muestras nuevas con nuevos elementos poblacionales.

Un ejemplo de estimación no objetiva es el empleado por la EEAO (Estación Experimental Agroindustrial Obispo Colombes). En una encuesta

dirigida a productores se les pregunta “*En cuánto estima los rendimientos promedio, máximo y mínimo del cultivo de soja en la zona húmeda y en la zona seca de Cruz Alta Burreyacú) para la presente campaña y en que porcentajes se dieron los mismos.*” En este caso no existen evidencias que los diferentes productores hayan empleado un método objetivo para sus estimaciones y la opinión requerida se refiere a la superficie sembrada con soja en Cruz Alta (41360 has) o en Burreyacú (121820 has), según sea la ubicación de los cultivos del productor²⁹. En consecuencia se plantean los interrogantes presentados antes sobre el método de estimación empleado.

Un método de estimación objetivo de la producción requiere conocer el ***número de plantas a ser cosechadas y estimar la producción por planta***. El producto de los dos componentes dará una estimación de la producción total del cultivo: ¿hasta dónde producirá este método estimaciones fiables y precisas? La variabilidad presente en el rendimiento por planta, asociada a los numerosos factores que la causan, obliga a diseñar cuidadosamente un método que permita alcanzar el objetivo – estimaciones fiables y precisas - a un bajo costo y en tiempo que permitan la adopción de decisiones racionales desde el punto de vista económico.

En Argentina, como en otros países agrícolas, las estimaciones del rendimiento de las plantas, base para pronosticar la producción de diferentes cultivos, se realizan muchas veces mediante evaluaciones visuales de “expertos” a partir de sus propios juicios de la condición de la cosecha. Este método es muy subjetivo y por ende está sujeto a sesgos desconocidos. La experiencia de algunos lugares demuestra que ***los informadores tienden a inclinarse hacia lo normal***: el rendimiento se subestima cuando la estación es buena y se exagera cuando es mala.

²⁹ Ver Anexo, punto II.

En el caso específico que nos ocupa la comparación entre ambos métodos no puede realizarse porque el método alternativo utilizado – de cosecha completa de planta, aunque pueda usarse para comparar los rendimientos de los ejemplares seleccionados-, no responde a los requisitos que debe cumplir **un método objetivo basado en el azar.**

De todos modos se pueden comparar solo con un sentido descriptivo los resultados que arrojan uno y otro método, incluso los que resultan de los informes de cada “experto” que intervienen; el seguimiento de los resultados habrá de brindar indicios para “un mejoramiento” de la técnica empleada hasta ahora. Diferencias del 5 % ó más, entre las estimaciones subjetivas y las estimaciones por encuestas, pueden considerarse de importancia dado la necesidad de planificar la cosecha y comercialización en todos sus aspectos: mano de obra y equipamiento necesario para la recolección, transporte y almacenaje de la producción, necesidades de financiamiento y aprovechamiento de condiciones favorables del mercado disponible y a término del grano. En lugares que usan estimaciones visuales, cuando sea posible, se recomienda sustituir por un método objetivo.

Los productores de soja preocupados por el rendimiento de sus lotes recurren a sitios de la web para solicitar asesoramiento sobre un método de estimación que sea objetivo en la apreciación del fenómeno.

Pregunta realizada el: 08/03/2010

Nombre de quien consulta: María Pía

Consulta:

Hola Necesito saber cómo estimar el rinde de la soja, una vez me lo explicaron pero no recuerdo. Lo que tengo es la distancia entre líneas, nº de plantas por metro lineal y conté las chauchas pero no sé si saque bien las

muestras, agradecería que me expliquen cómo hacerlo. También tengo como peso promedio de las 1000 semillas 180 gramos, y no sé si esta acertado. Atte.-

Respuesta dada al: 22/03/2010

Estimada María Pía: Cuando se estima rendimiento en soja, es muy importante tomar varias muestras representativas del lote ó de sectores del lote (si el mismo tuviera zonas muy diferenciadas, por ejemplo lomas y bajos), de al menos un metro lineal de surco (con la densidad promedio del lote y con competencia de los surcos vecinos). Si los muestreos fueran representativos del lote y tuvieras un adecuado relevamiento de plantas y chauchas, te falta el número promedio de granos por chaucha y el peso promedio de los granos. Para determinar el número de granos promedio por chaucha, lo mejor es separar las chauchas por número de granos llenos (de 1, 2, 3 y 4 granos por chaucha), sumar granos llenos y dividirlos por el número de chauchas (generalmente los valores están entre 2,5 y 2,8 granos por chaucha y pueden ser mucho más bajos cuando hay ataques de chinches u otros problemas (sequía, enfermedades, otros). Con respecto al peso de los granos, si aún no estuviera terminado el llenado de granos, se deberá estimar, con el riesgo de cometer errores importantes; porque el peso de granos en soja varía generalmente entre 130 y 180 gramos los 1.000 granos (una estimación prudente sería de 150 gramos los 1.000 granos). Agradeciendo tu consulta y quedando a tu disposición, te saluda. Héctor.

La recomendación suele ser tomar varias muestras, por la variabilidad del número de granos por vainas, de vainas por planta y de plantas por ha; la unidad **muestral se sugiere sea un metro lineal de surco y el número de muestras, dependerá de la variabilidad del lote de producción;** si se identifican diferentes ambientes se sugiere tomar al menos 3 ó más muestras por ambiente (loma, media loma y bajo). Se cosechan las plantas

de cada unidad muestral (metro lineal) y se considera el peso de mil granos sobre la madurez ó corregidos por humedad, que se establece conveniente al momento efectivo de la cosecha (por ejemplo 15 %). En el asesoramiento se informa que el recorrido probable de éste peso está entre 130 gramos ó menos – si los granos fueran muy chicos y 170 ó más si los granos fueran muy grandes. ***El tamaño de semilla es una característica que presenta variaciones genéticas y ambientales en su expresión. De la tabla de características de los cultivares de los informes de la RECSO (Red de Evaluación de Cultivares de Soja) y dentro de cada grupo de madurez es posible tener una idea de la variación genética: se obtiene un valor promedio de muchos ensayos en los que participaron todos los cultivares. Por otra parte, la disponibilidad hídrica puede hacer variar el peso de la semilla de un mismo cultivar, entre 90 y 180 gramos las 1000 semillas, y se espera una mayor compensación del componente de rendimiento tamaño de semilla, mientras mayor es el tamaño de semilla del cultivar, bajo número de vainas promedio (35 vainas por planta), para la biomasa vegetativa que ha logrado formar, tendrá un peso de semillas alto y permitirá compensaciones para lograr un rendimiento alto.***

5.-Estimaciones en encuesta por muestreo

5.1.-Formulación de objetivos

Objetivo General: El objetivo principal de la investigación es proporcionar información sobre el rendimiento y **la producción de soja de lotes, fincas o zonas de interés para productores, acopiadores, otros agentes de la cadena de comercialización y a organismos estatales, ONGs ó privados vinculados a la actividad agrícola.** Para alcanzar el objetivo será necesario previamente estimar el número promedio de plantas por metro lineal y por hectárea a cosecha (etapa R8).

Objetivos Secundarios:

- a) Proporcionar información sobre el rendimiento y producción de lotes, fincas o zonas **según variables asociadas que defina el productor, acopiador u organismo interesado en la información.**
- b) Proporcionar información sobre características de la soja a cosechar de los diferentes lotes de fincas y zonas (dominios de análisis).

5.2.-Población y unidades investigadas

El universo para la encuesta de rendimiento de soja está conformado por las plantas de ese cultivo implantadas en los diferentes lotes de cada una de las fincas y zonas que se definan como dominio de análisis al efecto de las estimaciones. Las unidades de recolección de los datos son **plantas en la etapa R8 – donde el 95 % de las vainas han alcanzado el estado de plenitud de madurez** - y en condiciones de ser cosechadas entre los cinco y diez días siguientes de clima seco para que la semilla tenga menos de 15% de humedad, estado requerido por el acopiador en la etapa de comercialización inmediata siguiente.

La encuesta se estructurará por lotes, fincas o zonas con idénticas características en cuanto a tipo de cultivar implantado, tipo de terreno, disponibilidad hídrica, fecha de siembra, condiciones ambientales durante el periodo de cultivo, grupo de madurez, arreglo espacial y control de problemas sanitarios. A éstos efectos será necesario realizar, según datos preexistentes suministrados por el productor y relevamientos satelitales, un agrupamiento de los diferentes lotes con similares condiciones en las características señaladas. Para los diferentes agrupamientos definidos se dispondrá de información sobre el número de semillas implantadas por metro lineal y por hectárea y para ellos se obtendrán estimaciones del número de plantas a cosechar por metro lineal y por hectárea, como etapa previa a la estimación

del rendimiento y la producción de soja por dominio de análisis (agrupamiento de zonas implantada con similares características).

Para reducir tiempo y costos operativos, de control, procesamiento y publicación de la información, como así también hacer factible la realización del trabajo de campo pueden definirse tamaños mínimos de lotes, zonas o agrupamientos de lotes que participan en la muestra con probabilidad igual 1; de modo que no todos los lotes, zonas o agrupamientos definidos participan en la muestra con probabilidad igual a 1.

5.3.-Principales variables investigadas por la encuesta

Las **variables de respuesta** consideradas en la encuesta correspondiente a los distintos dominios de análisis (lotes, finca, zona o agrupamiento de lotes) son:

- a) Número total de granos de soja por planta.**
- b) Peso en gramos de 1.000 granos de soja con el 15 % de humedad o el establecido por el acopiador.**
- c) Nº de plantas a cosechar por metro lineal.**
- d) Rendimiento en Kg. de soja por ha implantada.**
- e) Producción de soja en toneladas de lotes, fincas o zonas (dominio de análisis).**
- f) Proporción de la producción de soja con las características de interés definidas por el demandante de la información.**

Las **Variabes de clasificación** son: zona, finca, lote y cultivar implantado.

Las **Variables de asociación**: grupo de madurez, fecha de siembra, arreglo espacial, condiciones ambientales en el período de cultivo y control de problemas sanitarios.

5.4.-Grado de precisión

Se establece una precisión deseable (margen de error) de **más o menos 7,5 gramos** por 1.000 granos de soja con el 15 % de humedad o el establecido por el acopiador, para las distintas zonas y cultivares considerados. Las estimaciones de intervalos se harán a un nivel de confianza del 95% en los dominios lotes, fincas, zona y cultivares definidos de interés, cuando el tamaño muestral sea superior a 36 unidades terciarias y de observación relevadas.

5.5.-Periodo de referencia

En el diseño de la encuesta se han distinguido los siguientes períodos:

- De **encuesta**: período durante el cual se realiza el trabajo de campo.
- De **estimación**: se define a los efectos de realizar las estimaciones en base a un período común o estándar.
- De **referencia**: período acerca del cual las plantas informan sobre la ocurrencia de un suceso

Los periodos mencionados **no coinciden** en la encuesta. El **periodo de encuesta** está previsto que tendrá lugar en abril-mayo, meses de relevamiento de los datos, establecido a los fines de obtener la respuesta sobre las diferentes variables investigadas de las plantas seleccionadas para tal fin. Los **periodos de estimación y referencia** coinciden y corresponde al periodo de cosecha (año).

Se buscará reducir los errores muestrales al captar el mayor número de observaciones posibles. Este cometido se puede alcanzar trabajando con metros lineales como unidad de observación para estimar el número promedio de plantas por metro y realizando luego la correspondiente transformación para obtener el número de plantas a cosecha por metro cuadrado, según el tipo de arreglo espacial utilizado para el implante del cultivo en la zona de interés. También se buscará minimizar los costos operativos y los errores no muestrales como disponer de los datos en el tiempo que las decisiones a tomar lo requieran.

5.6.-Presupuesto

La implementación efectiva de la encuesta requiere determinar costos financieros que se dividen en

1. Costos fijos: corresponden a contratación de personal, vehículos, materiales y suministros;
2. Costos variables: corresponden a las necesidades específicas de cada zona ó dominio de análisis.

Personal y Honorarios

Los costos del trabajo operativo, teniendo en cuenta dos semanas de duración, implican las siguientes necesidades de personal, tareas a realizar y tiempos de ejecución tentativo previsto para cada una de ellas.

Cantidad	Denominación
1	Estadístico
12	Enumeradores
3	Supervisores
12	Recolectores
1	Codificador
1	Data Entry

CRONOGRAMA DE ACTIVIDADES

Actividades	Tiempo
1.- Listado de lotes	1 semana
2.- Confección de Planilla de Actualización	3 días
3.- Preparación de materiales	3 días
4.- Capacitación de Encuestadores y entrega de material	2 días
5.- Trabajo de Campo	1 semana
6.- Recepción, Supervisión y codificación de las encuestas	1 semana
7.- Ingreso de datos	2 días
8.- Obtención y depuración de la base de datos	1 día
9.- Confección de Cuadros Estadísticos y Análisis	2 días
10.- Redacción de Informe Final y Ejecutivo	3 días

6.-Marco muestral

El marco muestral es la expresión operacional y física de las unidades (lotes en producción) sujetos a muestreo. En este caso se utiliza la expresión marco de muestreo en un sentido amplio, dado que se toman los lis-

tados de elementos físicos que, a su vez, permiten obtener muestras de unidades elementales, sin el esfuerzo de haberlas listado a todas.

El marco original de la muestra es la enumeración o listado de los lotes de las fincas de cada zona de implante con su correspondiente cartografía.

La encuesta se realizará a través de una muestra probabilística de plantas. La muestra, en principio, representará a la población de ejemplares implantados en zonas que según datos a la fecha superan las 1000 hectáreas. La muestra se obtendrá a partir de un listado exhaustivo de los lotes en que se hallan ubicadas las unidades muestrales del universo, mediante una depuración previa que no considera lotes que no están implantados con el cultivo de soja.

6.1.-Diseño y selección de la muestra

Para estimar el rendimiento de un lote seleccionado deberán localizarse al azar, en el lote, las plantas a ser cosechadas. La selección se hace a partir del listado de los lotes, de su ubicación y utilizando una tabla de números al azar. Existe el riesgo de un *sesgo del límite*, si las plantas ubicadas en los límites del lote difieren en su rendimiento de otras ubicadas en otras áreas del mismo. Este sesgo se reduce cuando menor es la extensión del lote. A partir de la observación en campo, se sugiere eliminar de la muestra a plantas ubicadas sobre los bordes del lote, aunque puede no restringirse su inclusión.³⁰

La población objetivo fundamental es la población de plantas en etapa R8 – plenitud de madurez -. Toda planta del universo de referencia pertenece a una zona y a un solo lote ó sección del mismo. Al relevar un lote

³⁰ (Cfr.) CERISOLA, Juan A., GUTIERREZ, Jesús I., LOPEZ CLEIP DE SOSA, Aída, Propuesta de Diseño Muestral para Estimación de Rendimientos en la Producción de Limón, Cátedra de Estadística, Facultad de Ciencias Económicas, UNT, Diciembre 2001

o Sección, se seleccionará al azar el número de ejemplares del mismo que registren las características predominantes.

Relevar un lote ó Sección es seleccionar a todas las plantas que lo componen y allí se seleccionará al azar el número de ejemplares que reúnan las condiciones requeridas para formar el universo³¹.

6.2.-Métodos de muestreo

Teniendo en cuenta los objetivos de la Encuesta y las premisas fundamentales, se optó por una muestra Probabilística, estratificada, **por áreas y cultivar**, multietápica, con probabilidad de selección proporcional al tamaño de los lotes implantados con soja en cada una de ellas.

Probabilística significa que cada unidad del universo de estudio tiene una probabilidad de selección conocida y superior a cero. Este tipo de muestra permite establecer anticipadamente la precisión deseada en los resultados principales y calcular la precisión observada en todos los resultados obtenidos.

Estratificada: La división de las zonas cultivadas en lotes genera interés en la estimación de la producción de cada una de ellas y de los diversos lotes que las componen, dadas las características propias de los mismos. Las zonas de producción constituyen los estratos principales de la encuesta. El plan de muestreo dentro de un estrato principal debe escogerse de tal modo que la muestra pueda proveer una estimación para el estrato a un nivel de precisión prescripto. *Cada estrato principal se subdivide en sub-estratos que se espera difieran considerablemente entre sí respecto del rendimiento, según estudios del tipo mediante ensayos de la RECSO, o bien, según características de las plantaciones de los lotes altamente correlacionadas o asociadas al rendimiento de los mismos (por ej., tipos de suelo y arreglo es-*

³¹ Consultar en Tablas el esquema del número de ejemplares a muestrear por lote, zona y cultivar de los mismos y distancia entre surco de implante.

pacial) ó la disponibilidad hídrica en las etapas R2 a R7. El muestreo dentro de un estrato se hace en forma independiente de los demás.

Dentro de los diseños posibles se escogió un muestreo estratificado ya que este método asegura una mayor confiabilidad de la muestra, al disminuir la varianza de las estimaciones y mejorar la calidad de las mismas, con la única condición de que toda unidad de muestreo pertenezca a un estrato y sólo a uno. Para la estratificación y selección de la muestra se siguió el criterio de cultivar implantado y zona de implante del ejemplar. El número reducido de plantas que componen la muestra por cultivar y lote ó finca reducen la precisión de las estimaciones de producción a ese nivel de estratificación por lo que se recomienda realizar estimaciones para los estratos conformados con el conjunto de lotes ó fincas de la zona definida para la estimación para el cultivar especificado. Al aplicar este criterio de selección quedaran conformados estratos según zona de implante del ejemplar, sub-estratos según tipo de cultivar implantado y sub-estratos según zona y tipo de cultivar, que sería conveniente detallar en casos concretos.

Áreas: Cada zona (Norte, Centro ó Sur por ejemplo) se sugiere dividirla en áreas (lotes o secciones de lotes) según cultivar implantado.

Multietápica: La selección de la muestra se realiza en las siguientes etapas:

1. Selección de las unidades primarias de muestreo (Zona), dentro del universo muestral previamente ordenado, con probabilidad de selección igual a uno para cada una de las zonas especificadas: (por ejemplo **Norte, Centro y Sur**).
2. Selección de las unidades Secundarias de muestreo o áreas (lotes ó secciones de lotes), dentro del universo muestral previamente ordenado, con probabilidad de selección igual a uno de fincas, lotes o las secciones de

ellos con más de 1000 hectáreas de extensión y menor a uno si el lote o la sección del mismo posee menos de 1000 hectáreas de extensión.

3. Dentro de cada una de las secciones cartográficas previamente identificadas y seleccionadas (lote y zona) se escogen al azar las unidades terciarias de muestreo (**metro lineal**) de las que se obtendrá el **dato** sobre **número de plantas a cosecha en la etapa R8**, una de las variables de respuesta investigadas.

4. De cada una de las unidades terciarias de muestreo (**metro lineal**) seleccionada en la etapa 3 se escoge al azar la unidad observacional de muestreo (**planta**) de la que se obtendrá el **dato** sobre **número de vainas y de granos por planta a cosecha en la etapa R8**, otra de las variables de respuesta investigadas, a fin de estimar el rendimiento promedio por hectárea y la producción total del estrato o sub-estrato.

El proceso de selección por etapas sólo requiere una actualización cartográfica detallada de los lotes y secciones de éstos que componen cada finca en las zonas definidas de interés y para los diferentes cultivares implantados.

El siguiente cuadro resume las etapas en las cuales se dividió el plan de muestreo

Etapas de Selección	Definición de unidades	Probabilidades asignadas	Procedimientos de selección	Cantidad a muestrear
1. Primera etapa: selección de unidades primarias de muestreo (UPM)	Áreas Geográficas Implantadas con Soja	Proporcionales a las hectáreas implantadas con el cultivo de soja, según datos de la empresa	Cierta	3 Zonas: Norte, Centro y Sur, por ej. (compuesta de fincas y lotes definidas por la empresa)
2. Segunda etapa: selección de unidades secundarias de muestreo (USM)	Lotes ó Secciones de lotes (área pequeña dentro de cada zona)	Proporcionales a las hectáreas implantadas con el cultivo de soja, según datos de la empresa	Cierta (Si el área tiene 1000 has. ó más implantadas) Al azar (Si el área tiene menos de 1000 has. implantadas)	Nº de lotes ó secciones de lotes mayores o iguales a 1000 has. implantadas Nº de lotes ó secciones de lotes menores a 1000 has. implantadas
3. Tercera etapa: selección de unidades terciarias de muestreo (UTM)	Metros lineales implantados con Soja en etapa R8 (áreas definidas a cosechar)	Proporcionales a la cantidad de metros lineales de la UM (lote ó sección) a la que pertenece	Al azar	Al menos 36 metros lineales por cada área USM (lote ó Sección)
4. Cuarta etapa: selección de unidades observacionales de muestreo (UOM)	Plantas de Soja implantada en R8	Proporcionales a la cantidad de plantas en R8 por metro lineal de la UTM a la que pertenece	Al azar	1 planta en R8 por UTM (metro lineal) seleccionada en tercera etapa de muestreo.

Para cada una de las zonas definidas a efectos del muestreo (UPM) se realiza un listado completo de los lotes o secciones de lotes implantados con soja y en estado R8, a los que también se las identifica por otras características, como tipo de cultivar, arreglo espacial, condiciones ambientales en las diferentes etapas del cultivo, fecha de siembra, etc., que sea de interés a los fines de la toma de decisiones.

Debe advertirse que el número total de plantas a muestrear que se asignará a un estrato principal no debe ser menor que el número necesario para hacer una estimación razonablemente precisa allí. Dentro de cada sub-estrato se localizan al azar las plantas del mismo que integran la muestra y se determina su rendimiento. El diseño es de etapas múltiples: las unidades muestrales de primera y segunda etapa se seleccionan en la oficina, mientras que las de tercera etapa (metros lineales) y las unidades observacionales (plantas en R8) pueden hacerse en campo; si se disponen de mapas, herramientas para el uso de GPS y un conocimiento acabado de los lotes se pueden determinar previamente coordenadas de la ubicación de UTM y UOM.

La selección de las UTM y UOM (metros lineales y plantas en estado R8) al azar es una parte muy importante en todo el procedimiento. Es aquí donde tenderán a producirse sesgos que puedan invalidar toda la encuesta. Como una protección, deberán establecerse procedimientos rigurosos para que el trabajador de campo (cosechero) recoja los datos de esas unidades de muestreo ajustado a éste método.

6.3.-Tamaño de la muestra

Al planear una encuesta por muestreo hay que tomar una decisión respecto del tamaño de la muestra. Esta decisión es importante. Una muestra demasiado grande implica un uso no eficiente de recursos y una muy pequeña disminuye la utilidad de los resultados. La decisión no siempre es satisfactoria si no se dispone de la información suficiente para saber si el tamaño de la muestra seleccionada es el óptimo.

¿Cuántas plantas deberán integrar la muestra para estimar la producción de uno o varios lotes con características similares? El uso de un nú-

mero reducido en la estimación de la producción de los lotes es peligroso. El número de las plantas que se seleccionará de un estrato principal estará gobernado por el nivel de precisión necesario para hacer estimaciones para el estrato y puede ser necesario probar con varios tamaños para hacer comparaciones y decidir sobre el número mínimo de plantas a ser cosechadas en la muestra, atendiendo a los objetivos de tiempo, costo y precisión aceptable de la estimación. Es claro que ***el número de ejemplares que se tomarán de cada estrato deberá ser lo más pequeño posible.***

La elección del tamaño de la muestra involucra entre otros pasos la **especificación de un enunciado respecto de lo que se espera de la muestra**, ya sea en términos de límites de error, ó en términos de una decisión o acción que debe tomarse a partir de los resultados de la muestra. El enunciado deben realizarlo los usuarios de los resultados de la encuesta. **Si los datos se requieren para subdivisiones de la población (zonas, cultivar, zona y cultivar, fincas, etc.) es necesario calcular un tamaño muestral (ni) para cada subdivisión** y el tamaño muestral general (n) se obtiene por adición.

De todos modos el tamaño de la muestra por si solo no es garantía de precisión en las estimaciones. Las muestras grandes no representativas pueden comportarse tan mal como las muestras pequeñas no representativas. **El diseño de la encuesta** es mucho más importante que el tamaño absoluto de la muestra y deberá ponerse especial **cuidado en la interpretación de los resultados** *si se percibe un evidente sesgo de selección (diferencia entre la población muestreada y la población objetivo).*

Los dos aspectos – tamaño y diseño muestral – pueden continuar bajo análisis y la propuesta sobre ellos puede considerarse provisoria hasta la disponibilidad de información específica de la población bajo estudio. Finalmente, el tamaño muestral elegido (n) deberá sopesarse

que sea consistente con los recursos disponibles (tiempo, materiales y personal). En una primera etapa se pueden fijar tamaños y criterios operativos provisorios, para posteriormente realizar calibraciones a los mismos que permitan definir un diseño ajustado a los objetivos.

En este caso para determinar el tamaño de la muestra se ha tenido en cuenta que los estadísticos que se estimarán son promedios (número de plantas por metro lineal y rendimiento por ha.) y Totales (Producción en toneladas de lotes, fincas y zonas), que el proceso operativo de muestreo si se quiere lograr el objetivo principal planeado, si bien elimina el ejemplar para la producción en la temporada no constituye un costo relevante para la determinación del tamaño muestral y que considerar un número elevado de plantas puede poner en dudas las ventajas del procedimiento de estimación si el tiempo demandado por el proceso se extiende y se acerca demasiado al periodo de la efectiva cosecha.

Bajo estas premisas y considerando las estimaciones previas de variabilidad en las variables de respuesta principales, el cálculo del tamaño de muestra necesario efectivamente recomendado (al menos 36 metros lineales implantados con plantas en estado R8)³² debe garantizar los requisitos de precisión prescriptos al nivel de confianza del 95 % y se encuentra en los límites presupuestarios posibles de afrontar, aún para al nivel de productores con áreas plantadas inferiores a las 1000 hectáreas, siempre que correspondan a características similares en otras variables relevantes en el resultado de las variables investigadas.

Un mayor tamaño puede estar relacionado con la necesidad de contar con datos sobre las variables investigadas en las distintas secciones (fincas ó lotes) que habrán de relevarse y que tengan un arreglo espacial con distancias entre surcos inferior a 52 cm.

³² Consultar Anexo, punto III.

Con éstas premisas y consideraciones referidas al tamaño y diseño muestral, la composición de la muestra correspondiente a cada estrato especificado se detalla a continuación.

Zona	d.e.s. 52 cm.	d.e.s. 35 cm.	d.e.s. 28 cm.	d.e.s. 21 cm.	Total Muestra
Zona Norte	n11	n12	n13	n14	n1
Zona Centro	n21	n22	n23	n24	n2
Zona Sur	n31	n32	n33	n34	n3
Total Muestra	n.1	n.2	n.3	n.4	$n = n1+n2+n3$

Zona Norte	d.e.s. 52 cm.	d.e.s. 35 cm.	d.e.s. 28 cm.	d.e.s. 21 cm.	Total Muestra
Lote N° 1	n11	n12	n13	n14	(n1)1
Lote N° 2	n21	n22	n23	n24	(n1)2
Lote N° m	nm1	nm2	nm3	nm4	(n1)m
Total Muestra	(n1).1	(n1).2	(n1).3	(n1).4	n1

Zona Centro	d.e.s. 52 cm.	d.e.s. 35 cm.	d.e.s. 28 cm.	d.e.s. 21 cm.	Total Muestra
Lote N° 1	n11	n12	n13	n14	(n2)1
Lote N° 2	n21	n22	n23	n24	(n2)2
Lote N° m	nm1	nm2	nm3	nm4	(n2)m
Total Muestra	(n2).1	(n2).2	(n2).3	(n2).4	n2

Zona Sur	d.e.s. 52 cm.	d.e.s. 35 cm.	d.e.s. 28 cm.	d.e.s. 21 cm.	Total Muestra
Lote N° 1	n11	n12	n13	n14	(n3)1
Lote N° 2	n21	n22	n23	n24	(n3)2
Lote N° m	nm1	nm2	nm3	nm4	(n3)m
Total Muestra	(n3).1	(n3).2	(n3).3	(n3).4	n3

Estos tamaños deben definirse para obtener estimaciones sobre número de plantas promedio por metro lineal, primero, y luego del rendimiento promedio por hectárea de soja en etapa R8 del cultivo para cada uno de los sub-estratos especificados en cada una de las zonas de interés. Si

razones de costos y tiempo lo requieren, pueden realizarse ajustes que reduzcan el número de unidades muestrales observadas en los sub-estratos – dominios de análisis - a efectos de obtener estimaciones de la producción en los mismos sin gran pérdida de eficiencia en las mismas, para lo cual se deberá cuidar la representación en la muestra de aquellos sub-estratos con tamaños menores especialmente.

Es importante recordar que el empleo de **personal especializado**, altamente formado, tiene un impacto significativo en la **reducción de errores no muestrales** que pueden atentar contra la validez de las estimaciones, más que los errores muestrales. A éste respecto es conveniente solicitar la colaboración de todo el personal afectado al objetivo en las diferentes tareas - antes, durante y posterior al momento del relevamiento - el apoyo necesario para eliminar o reducir el efecto de las potenciales fuentes de tal tipo de error.

6.4.-Diseño de instrumentos de recolección

La recolección de los datos sobre las variables investigadas por la encuesta resultante del conteo de número de plantas por metro lineal observado y la cosecha completa de la planta, con intervención de recolectores especialmente instruidos a los fines perseguidos por la misma, habrán de registrarse en formularios que se diseñarán a tal fin y que deberán contener los datos correspondientes a las variables de respuestas investigadas. Se sugiere realizar el diseño con la antelación que requiere la preparación del material para el proceso operativo de la encuesta. El sistema de recolección utilizado en la encuesta es el de barrido, que consiste en que cada recolector o grupo de recolectores trabaja en forma simultánea a otros grupos en los lotes o secciones asignados hasta relevar todas las unidades terciarias de

muestreo (metros lineales) y plantas comprendidas en ellos seleccionadas al azar (unidades observacionales de muestreo).

La siguiente es una planilla tipo que se propone para el registro de los datos que se relevan a partir de las unidades muestrales Terciaria (metro lineal) y observacional para el rendimiento del cultivo (planta).

Zona Norte. Unidad Secundaria de Muestreo: Lote N° i

Planilla de datos a relevar en UTM y UOM

UTM y UOM N°	Distancia entre Surcos (en cm.)	Ubicación de inicio metro lineal	Nº de plantas por metro lineal	Nº de vainas de planta observada	Nº de semillas en planta observada	Peso en gramos de semillas en planta observada
1						
2						
3						
m						
Total Lote i						

Zona Norte

Planilla Resumen de datos por Lote según Distancia entre Surco

Lote Nº	Distancia entre Surcos (en cm.)	Promedio de plantas por metro lineal	Promedio de vainas por planta	Promedio de semillas por planta	Peso Promedio de semillas por planta (en gramos)
1					
1					
1					
2					
2					
m					
m					
Total Zona					Promedio Semillas/m. cuadrado

Transformaciones de datos: Una vez depurada la información original, registrado el número de plantas por metro lineal, número de vainas y de semillas por unidad de observación (planta) correspondiente a los distintos dominios de análisis (empresa, zona, cultivar, zona y cultivar, distancia entre surco, zona y distancia entre surco) observadas en el muestreo, se procede a transformar los datos y generar las nuevas variables en planillas tabuladas especialmente diseñadas que permitan obtener la base de datos en términos de las variables de respuestas que interesan (rendimiento promedio por ha y producción por lote, producción por zona), CVR e intervalos de confianza correspondientes para las poblaciones objetivos de los dominios de análisis de interés.

Zona Norte.

Planilla Resumen de datos por Lote según Distancia entre Surco

Lote Nº	Distancia entre Surcos (en cm.)	Nº de surcos o líneas por hectárea	Promedio de vainas por hectárea	Promedio de plantas en R8 por hectárea	Número Promedio de semillas por planta	Peso Promedio de 1000 semillas con 15 % de humedad	Rendimiento promedio de soja por ha.
1							
1							
1							
2							
2							
M							
M							
Total Zona							

Zona Norte

Planilla Resumen de datos por Lote según Distancia entre Surco

Lote Nº	Distancia entre surcos (en cm.)	Promedio de plantas en R8 por hectárea implantada	Peso Promedio de 1000 semillas (15 % de humedad)	Rendimiento Promedio de soja (kg. por ha.)	Superficie del lote Nº de has.	Producción Estimada de soja (toneladas)
1						
1						
1						
2						
2						
M						
M						
Total Zona						

6.5.-Organización del trabajo de campo

Con el propósito de realizar una recolección y control con la oportunidad y calidad requerida, los grupos de trabajo se organizan en rutas. Una ruta la constituyen lotes y áreas seleccionadas dentro de cada zona, ordenadas lo más contiguamente posibles.

El rendimiento obtenido en un corte muestral se obtiene de ordinario en condiciones controladas. Esto es en diferentes condiciones de aquellas en que se cosecharán los lotes. El resultado será que la estimación del rendimiento obtenido en la encuesta no será comparable con el rendimiento obtenido en la realidad. Por ejemplo se perderá una parte de la cosecha por algún motivo y en consecuencia los resultados de la encuesta deben corregirse para tomar en cuenta estos factores. A tal fin pueden emplearse facto-

res de corrección según datos de pérdidas de cosecha en ésta etapa según información de experiencias realizadas por empresas proveedoras de las máquinas para tal servicio. Tal corrección puede hacerse a nivel de lotes, zona ó para el total del área bajo estudio.

Una comparación de los resultados de las estimaciones en las condiciones de la encuesta y las normales proveerá el factor de corrección a utilizar en futuras encuestas.

6.6.-Recursos humanos y capacitación

La calidad de la información obtenida depende de la buena disposición de los recolectores para respetar las instrucciones que reciban a los fines de encuesta y de su motivación. Esto último también será de importancia que se observe en supervisores, analistas y coordinadores para implementar la encuesta. Por ello es necesario contar con recursos humanos adecuadamente capacitados.

El personal afectado a la misma sería:

Estadístico o Consultor en el tema: Asesora la implementación efectiva de la encuesta.

Enumerador: realiza tareas antes, durante y después del trabajo de campo, tales como: corroborar que en la cartografía consten todos los datos correspondientes a la ubicación geográfica de las áreas de cultivo que conforman el marco muestral de la encuesta; planificar el recorrido del área; verificar la coincidencia entre la cartografía del área y las características del terreno; listar las áreas a relevar (lotes ó secciones de lotes) con la indicación de su extensión; ordenar las planillas ó listado de lotes ó secciones de lotes en función a la zona y fincas a que pertenecen; verificar el número de ejem-

plares de cada zona, lote ó sección a relevarse y entregar el material en condiciones.

Cartógrafo: suministra los planos necesarios para una correcta ubicación de lotes ó secciones de lotes seleccionadas y una mejor organización del relevamiento.

Personal de campo: se encarga coordinar las visitas a lotes y secciones de lotes seleccionadas, ubicar las unidades de muestreo terciarias (metros lineales) y observacionales (plantas) a ser relevadas por el muestreo que se encuentren en ellos siguiendo el sentido preestablecido. Cuentan el número de plantas en estado R8 en las UTM.

Recolectores: son las personas que realizan la recolección de las vainas de las unidades observacionales seleccionadas al azar; reúnen las vainas provenientes de todas las plantas cosechadas del lote para estimar el rendimiento promedio de la unidad transformada a tal efecto (metro cuadrado) y luego de la hectárea.

Supervisores de recolección: tienen a su cargo a un grupo de personal de campo y recolectores, se encargan de que se cumplan las normas establecidas y anotan los datos en el instrumento de registro de la encuesta (formulario).

Crítico-codificadores: son los encargados de hacer una verificación general de la cobertura del material, una revisión cuidadosa de los formularios desde el punto de vista de la cobertura individual y de la precisión y consistencia de la información recolectada, además de la constatación de las características propias de los lotes o secciones relevados.

Personal de grabación: son las personas encargadas de hacer la captura electrónica de la información recolectada y codificada.

En la capacitación se tiene en cuenta la dispersión de la encuesta y la cantidad de personal que participa en ella. Se recomienda elaborar un manual de procedimientos para cada uno de los puestos de trabajo de la encuesta, donde se detallan las tareas a cumplir y la forma de realizarlas.

6.7.-Validación y crítica

El objetivo de esta evaluación es documentar la revisión de aquellos aspectos más críticos del trabajo realizado, señalando los problemas y bondades de las soluciones aportadas de modo que esta revisión sirva como experiencia para acrecentar la calidad y las posibilidades de un mejor desarrollo de los futuros procesos de la encuesta.

Se pueden establecer posibles controles sobre los siguientes aspectos:

- Diseño de la muestra y Plan de muestreo
- Elaboración del instrumento de recolección de los datos (Formulario).
- Organización del trabajo de campo.
- Relevamiento de información durante el trabajo de campo.
- Procesamiento y Análisis de los datos
- Valores de las variables de respuestas principales al nivel de dominio de análisis para los que se disponen de datos suficientes para estimaciones significativas estadísticamente.
- Utilización combinada de métodos de estimación: por muestreo y visual.

- Respecto del Diseño, Plan de Muestreo y trabajo de campo, se tratará de organizar el relevamiento para disminuir los costos de recolección ligados a la movilidad de los encuestadores. Esta situación se debe a que la cantidad de unidades observacionales (plantas) seleccionadas por lotes o secciones de lotes a encuestar en muchos casos resulta un número bajo, obligando a realizar la mencionada concentración y prevenir los posibles errores que pudieran devenir de utilizar este procedimiento. Entre otras cosas, se decidió darle importancia, dentro de los dominios de análisis a las formas de organización de recolección, a la captura de datos y a la eficiencia muestral esperada, buscando de este modo realizar un uso eficiente de los recursos humanos y materiales que se afecten al operativo diario de recolección de la información.
- En lo referente a la construcción del formulario, los problemas a controlar se relacionan con las posibilidades de una buena aplicación del instrumento de recolección. Con este objeto se prestará especial atención a la definición de las variables transformadas asociadas a las principales variables de respuesta y el modo de registro del dato respectivo, según definiciones del Departamento técnico responsable de la producción de la información a informar para la toma de decisiones. Se recomienda, si ha de requerirse demasiados detalles, elaborar un instrumento diferente de aquel destinado a las variables de respuestas principales registradas en los relevamientos en campo y su procesamiento en forma independiente. La falta de una adecuada capacitación de los observadores para lograr la transformación más adecuada de las variables podría generar sesgos de medición importantes que invaliden los resultados. La planilla de datos a relevar al realizar más desagregaciones según variables asociadas las principales (tipos de cultivar, espaciamiento entre surco, disponibilidad hídrica, etc.) a efectos de estimar las diferentes rindes según ellas crean mayor oportunidad de vicios en los resultados (errores no muestrales) y es parte de problemas a definir con una

implementación provisoria de la encuesta. Todas éstas consideraciones se orientan a reducir los sesgos por errores de no muestreo que pueden colarse por hacer más complejo el proceso operativo.

- Para los controles a nivel de relevamiento de información es conveniente establecer que el primero de ellos sea realizado por los supervisores de recolección, quienes al recibir los formularios llenados por los recolectores tendrán que verificar la coherencia de la información a partir de las reglas generales y pautas de consistencias oportunamente señaladas en la capacitación y manuales de procedimientos elaborados a los fines operativos.
- Un segundo control lo realizarán los críticos-grabadores al cargar la información en los soportes informáticos. Posteriormente los analistas realizarán un doble control de la información: Controles previos al ingreso: controles de coherencia, controles sistemáticos de consistencia; Controles posteriores al ingreso: una vez ingresada la información, los analistas controlarán los errores o inconsistencias detectados en forma automática por el programa de ingreso.
- Con respecto a la utilización en forma conjunta de los métodos de estimación por muestreo y visual, se sugiere que el sistema de estimación visual por informadores experimentados puede complementar a un muestreo objetivo y solo cuando la corrección de los resultados de encuestas por muestreo haya sido probada por comparación con otros datos, el método del muestreo podrá sustituir totalmente al de informes visuales. Para que funcione el método de muestreo, los lotes y plantas se deberán seleccionar al azar y el momento del muestreo estar suficientemente cercano al de la cosecha total. El número de plantas que integren la muestra debe ser suficientemente grande para aumentar la eficiencia de la estimación. Hasta entonces, el método del muestreo doble puede emplearse para aumentar la precisión de la

estimación, realizando estimaciones visuales primero de una muestra grande de plantas, seguida de recolecciones o cosecha de una submuestra que permita calibrar las primeras en posteriores estimaciones. Se deberá cuidar que las dos estimaciones sean independientes, lo que exige que quienes intervengan en una de ellas no participen ó no dispongan de información sobre los resultados arrojados por la primera. En conclusión el procedimiento requiere de una prueba adecuada.

Es así que los controles fueron establecidos de la siguiente manera:

a) Controles realizados por el supervisor: durante la semana de la encuesta el supervisor realizará controles en campo, de rutina y especiales, observará muy especialmente la selección efectiva al azar de las unidades muestreadas.

b) Controles realizados por críticos – grabadores: los analistas realizarán un doble control de la información. Controles previos al ingreso: controles de coherencia, controles sistemáticos de consistencia. Controles posteriores al ingreso, una vez ingresada la información, los analistas controlarán los errores o inconsistencias detectados en forma automática por el programa de ingreso. Por ejemplo: al ingresar los datos de los diferentes lotes o secciones de lotes y para cada unidad muestreada (metro lineal y planta) se habrán de verificar que la sumatorias correspondan con al número de unidades muestrales observadas y previstas relevar.

c) Controles realizados por quienes procesan los datos: deberán verificar que las fórmulas que se apliquen para el cálculo de los promedios muestrales y totales estimados se ajusten a las definiciones y a los datos reales, base de los mismos. Iguales controles se realizarán en la aplicación de fórmulas para cálculo de los Coeficientes de Variabilidad Relativas (CVR) y de los límites de confianza, según instrucciones del equipo de estadísticos.

d) Controles realizados por el Coordinador sobre los resultados de las estimaciones. Las diferencias entre las estimaciones de producción informados resultantes del uso de muestras para la expansión a la población de los resultados y los volúmenes finales cosechados del área bajo análisis durante la campaña, pueden no ser adjudicables a las distintas muestras posibles de extraer – error muestral – sino a ***sesgos de selección o de medición. En este caso es relevante realizar las modificaciones en la implementación del proceso de manera de minimizarlos, a fin que los resultados finales permitan conocer con un mínimo error lo que buscamos mediante una estimación muestral. A la vez, el conocimiento de información relevante previa al muestreo proveniente de otras fuentes, si éstas existen,*** y la resultante del mismo, debidamente consideradas, ***permitirá mejorar el diseño que se presenta y obtener estimaciones más eficientes.***

Finalmente, en el análisis de los datos procesados deberán considerarse muy especialmente los CVR que sean superiores al 10%, indicativos de la importancia de los errores muestrales y evaluar permanentemente los sesgos de selección y de medición, fuentes de errores no muestrales que se intentarán minimizar.

7.-Grado de precisión de las estimaciones - Errores de muestreo

Las planillas diseñadas habrán de presentar los errores muestrales correspondientes a los diferentes dominios de análisis: Total, zonas, cultivos, y zona y cultivos, por ejemplo. Las mismas, al efecto del análisis de los datos habrán de mostrar la magnitud de los errores muestrales en forma de errores estándares relativos.

El error estándar relativo (EER o CVR) es un indicador de la precisión de los resultados obtenidos a partir de la muestra seleccionada. Cuanto menor es el EER, mayor es el nivel de precisión de la estimación correspondiente. En general se expresa en % y no tiene unidad de medida. Es habitual considerar que errores estándares relativos menores o próximos al 10% señalan resultados muestrales confiables, aunque en algunas circunstancias, dependiendo de las condiciones generales que se observan en los diferentes dominios de análisis, se pueden tolerar valores superiores.

El error estándar relativo (EER o CVR) de una estimación se usa para la construcción de un intervalo numérico al que se le asigna una cierta confianza - medida en términos de probabilidad - de contener el valor verdadero que se desea estimar (promedio de planta por metro lineal ó peso promedio de 1000 semillas en la población objetivo, por ejemplo). El cálculo de los límites del intervalo numérico depende en gran parte de los datos que se toman de las referidas variables en la muestra observada. En consecuencia las estimaciones se inician con los cálculos en una muestra preliminar o conjunto de muestras y del error estándar relativo (EER ó CVR) correspondiente a la misma, que servirán de base para fijar los límites de los intervalos³³.

Por ejemplo si se estima en un dominio (lote por ejemplo) que el número promedio de plantas en estado R8 por metro lineal es de 12 y el EER o CVR es igual al 5%, el verdadero valor estará comprendido entre 10,8 plantas y 13,2 plantas por metro lineal que surge de realizar el siguiente cálculo $12 \pm (2 \times 0,05 \times 12)$ asumiendo un nivel de confianza del 95%. Pero si se estima que la cantidad de plantas por metro lineal es de 15 el CVR es igual 10 %, se estima que el verdadero valor de cantidad de plantas a cosecha por metro lineal estará comprendido entre 12 y 18 plantas - suponiendo un nivel de confianza del 95%.

³³ (Cfr.) WALPOLE, Myers, Probabilidad y Estadística, 4° Edición, McGraw-Hill,(México, 1992).

El error estándar en el primer caso es de $\pm 1,2$ plantas y en el segundo caso es de $\pm 3,0$ y es lo que genera una mayor amplitud del intervalo numérico (**menor precisión en la estimación**) que deberá considerarse al evaluar la decisión a tomar.³⁴

En el caso que se estime en un dominio (lote por ejemplo) que el peso promedio de 1000 semillas en estado R8 es de 120 gramos y el EER o CVR es igual al 2,5%, el verdadero valor de la producción por planta estará comprendido entre 108 gramos y 132 gramos que surge de realizar el siguiente cálculo $120 \pm (2 \times 0,05 \times 120)$ asumiendo un nivel de confianza del 95%. Pero si se estima que tal peso es de 150 gramos y el CVR es igual 10 %, se estima que el verdadero valor del peso de 1000 semillas a cosecha con 15 % de humedad estará comprendido entre 120 y 180 gramos - suponiendo un nivel de confianza del 95%.

El error estándar en el primer caso es de ± 12 gramos y en el segundo caso es de ± 30 gramos, lo que genera una mayor amplitud del intervalo numérico (**menor precisión en la estimación**) que nuevamente deberá tenerse presente al evaluar la decisión a tomar.

Se establece una precisión deseable (margen de error) de **más o menos 1** planta por metro lineal en las estimaciones para los dominios de menor extensión (lotes) en las distintas zonas y cultivares considerados. En el caso de las estimaciones de peso para esos dominios se establece una precisión deseable (margen de error) de **más o menos 7,5** gramos para 1000 semillas. Las estimaciones de intervalos se harán a un nivel de confianza del 95% en los dominios lotes, fincas, zona y tramo de edad, zona, con muestras mayores a 36 unidades observacionales.

³⁴ D.A.S. Fraser, Fundamentos y Técnicas de la Inferencia Estadística, Editorial Limusa, (México, 1976)

El procedimiento descrito es conservador en el sentido que el estimador de la varianza tienda a sobreestimar la varianza verdadera y tenga en cuenta las características del diseño y del estimador empleado.

La selección del tamaño de muestra implica la disponibilidad equilibrada de recursos para ello y es posible que se requieran hacer cambios hasta alcanzar una solución aceptable. La estimación de los rendimientos y volúmenes de producción es un proceso continuo y a los resultados producidos por una muestra no siguen los de otras siempre. De este modo el tamaño muestral "efectivo" sugerido puede resultar diferente al resultante de la aplicación estricta de recomendaciones técnicas y éste procedimiento habrá de complementarse con datos históricos sobre pérdidas por el sistema de cosecha y diferencias en humedad al usado para las estimaciones, de modo que la información final para la toma de decisiones surgirá de la combinación de un método objetivo e información adicional, pero con menores errores totales del primero.

8.-Nota al Diseño Muestral Propuesto

La presente propuesta está basada en datos que surgen de consultas realizadas por productores de diversas regiones del país a profesionales y organizaciones especialistas en diversos aspectos del cultivo de soja. En una situación práctica concreta será el Departamento Técnico de la institución interesada en las estimaciones el encargado de implementar la encuesta, por lo que la definición de los dominios de análisis y la información necesaria para la adaptación de la propuesta será suministrada por el mismo. Se habrá de considerar entonces en el diseño una base de datos de la organización que se ha de procesar para la finalidad descrita. El tamaño muestral que se recomienda puede ajustarse a fines de reducir tiempos del procedimiento práctico y costo del mismo para producir los datos requeridos. El

tiempo y costo del operativo pueden resultar algo superiores si se lleva a cabo, en forma simultánea, relevamiento en campo de características vinculadas a otros objetivos que pueden resultar de interés para los decisores. Además, junto a la presentación de toda propuesta de diseño muestral los autores recomiendan prever como complemento indispensable para una adecuada ejecución de ella, una capacitación al Equipo Técnico que llevará a cabo el trabajo de campo con el fin de lograr mayor homogeneidad en la manera de realizar el mismo por los distintos grupos que lo conforman, buscando de ese modo reducir errores no muestrales en los resultados.

El metro cuadrado como unidad observacional para la estimación del rendimiento y la producción de soja: La estimación tanto del rendimiento como de la producción de soja, si se siguen las recomendaciones de profesionales y organizaciones que asesoran a productores está basada en considerar como unidad observacional el metro cuadrado implantado con soja, tanto para el número de plantas por hectárea como para el peso de las semillas a usar en la estimación del rendimiento. La recomendación sugiere tomar al menos 3 de esas unidades en lugares definidos como loma, media loma y bajo loma. La alternativa propuesta intenta evitar un error muestral importante por el número reducido de unidades a observar, al mismo tiempo que intenta incorporar a través del criterio sugerido una mayor diversidad de terrenos que permitan capturar una medida de variabilidad que produzca estimaciones más precisas. La sugerencia del metro cuadrado implica el uso de un muestreo por conglomerado, con lo que el número de observaciones requerido para un dado nivel de precisión será mayor que en un muestro al azar simple. El diseño propuesto si bien incorpora un mayor costo de recorrido del terreno para el relevamiento de los datos, permite relevar un número mayor de unidades observacionales (metros lineales y plantas), aumentando en modo sustantivo la precisión de las estimaciones.

9. CONCLUSIÓN

a) **El método de cosecha como método de estimación del rendimiento y la producción de soja.**

Resumiendo todo lo anterior en este acápite, el método de cosecha como método de estimación del rendimiento consiste en seleccionar una muestra de fincas o lotes al azar y emplear en la observación el método de la medición física del rendimiento en el momento de la cosecha. Las unidades muestrales seleccionadas (plantas) en efecto son cosechadas. Se elabora una lista de unidades de primera y segunda etapas (zonas y lotes) en el área que va a estudiarse. Se selecciona una muestra de lotes y se elabora un mapa de distribución de las plantas para cada lote de la muestra. Se toma una muestra al azar de las coordenadas de ubicación del inicio de los metros lineales y plantas a ser observados en el lote seleccionado, se cuenta el número de plantas por metro lineal y se cosechan las vainas en las plantas seleccionadas de los mismos. Posteriormente se calcularán los promedios y se hacen las transformaciones necesarias para obtener las variables de respuestas buscadas: rendimiento promedio por ha. y producción estimada del lote. ***La experiencia ha demostrado que éste es un método practicable, capaz de producir estimaciones de rendimientos libres de sesgos y con un alto grado de corrección.*** Pero es un método caro, en relación al de estimación visual, que requiere el empleo de un gran número de encuestadores adiestrados. ***El método es delicado en el sentido que pueden colarse en los resultados diversos sesgos si no se presta atención a los detalles.***

La selección de los lotes requiere que todos ellos se encuentren en la misma etapa de madurez del fruto previo a la cosecha, habrá que asegurarse que los lotes sean visitados en momentos adecuados y que se extraigan todos los frutos (vainas y semillas) a ser cosechadas en la temporada. ***Habrá que asegurarse que todos los lotes estén representados en la***

muestra. En lo posible habrá de ajustarse el muestreo de modo que su distribución en el tiempo corresponda con la distribución en el tiempo de la cosecha efectiva de los lotes. Por otra parte, no siempre es igual el tiempo que va desde el momento en que un lote se considera listo para la cosecha y el momento efectivo de la misma, por lo que habrá un sesgo por este motivo que deberá minimizarse.

b) Estimación de rendimientos y producción de soja en etapas anteriores a R8

Transcurridos las primeras etapas reproductivas del cultivo de soja, se plantea la necesidad de conocer estimaciones de la producción que permitan anticipar acciones que racionalicen las decisiones de los productores. El espacio temporal que medía entre el momento de la estimación y el de la cosecha efectiva puede ser acompañado de condiciones ambientales y sanitarias que generen un número de plantas y estado a cosecha que difieran demasiado que sumado al costo monetario, a pesar de la urgencia de disponer de la información, pueden constituir entonces límites a ejecutar una encuesta objetiva en esos estadíos previos a R8 que, de llevarse a cabo, deberá observar todos los recaudos ya señalados. Si bien ésta tarea no es recomendable realizarse por los motivos señalados, una estimación visual, con los peligros que ella implica, puede constituir una alternativa que se recomienda sea hecha por un número reducido de personas para evitar el problema de múltiples apreciaciones subjetivas, con el consiguiente ahorro de tiempo y dinero, aunque sujeta a sesgos desconocidos.

ANEXO

I.-Estadios Fenológicos Externos de la Soja

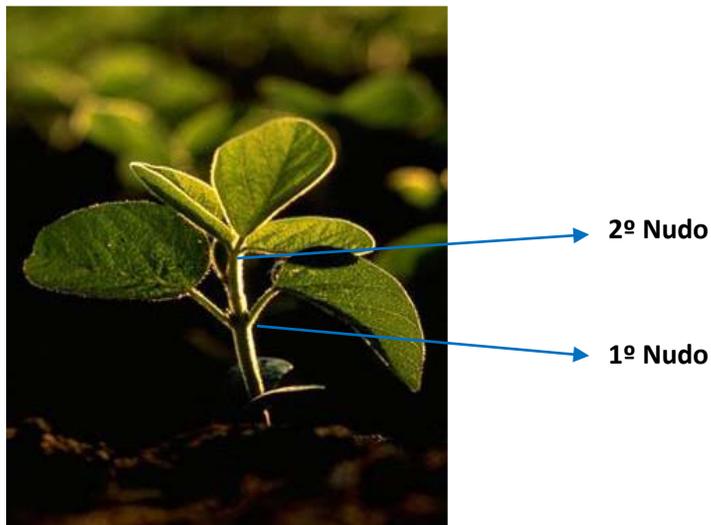
- **Etapas Vegetativa**

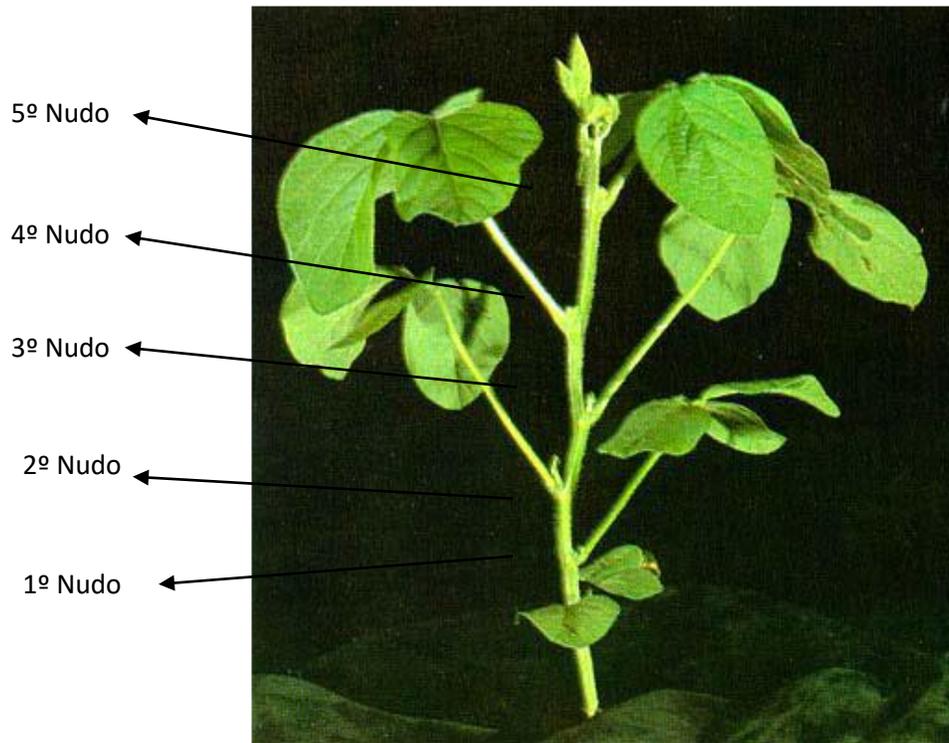
V1: (1er nudo), cuando el par de hojas opuestas unifoliadas están totalmente expandidas y en el nudo inmediato superior se observa a la 1er hoja trifoliada, en donde el borde de cada uno de sus folíolos no se tocan.

V2: (2do nudo), la 1er hoja trifoliada está totalmente expandida y en el nudo inmediato superior los bordes de los folíolos de la 2da hoja trifoliada no se están tocando.

V3: (3er nudo), La 2da hoja trifoliada está completamente desarrollada y la 3er hoja trifoliada presenta el borde de sus folíolos sin tocarse, así para cada uno de los nudos siguientes.

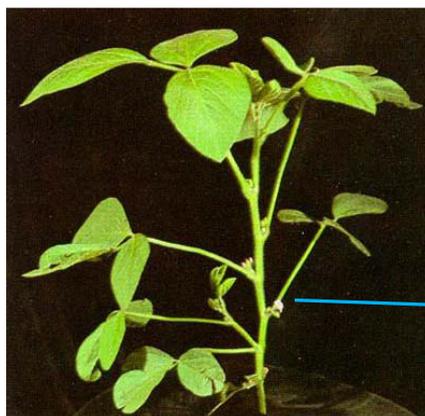
Vn: (n: número de nudos), la hoja trifoliada del nudo (n) está desarrollada y en el nudo inmediato superior el borde de cada uno de los folíolos no se tocan.





- **Etapa Reproductiva**

R1- Inicio de Floración: Presenta una flor abierta en cualquier nudo del tallo principal. Las flores miden entre 6 y 7 mm de largo y su color puede ser blanca o con distintos tonos de púrpura. La floración comienza en la parte media del tallo principal. La aparición de nuevas flores alcanza su máximo entre R2, 5 - R3 y culmina en R5.



Flor Abierta

El crecimiento vertical de las raíces aumenta marcadamente en esta etapa y el ritmo se mantiene hasta las etapas R4-R5

R2- Floración completa: Se observa una flor abierta en uno de los nudos superiores del tallo principal con hojas totalmente desplegadas. Esta etapa indica el comienzo de un período de acumulación diaria y constante de materia seca y nutriente que continuará hasta poco después de R6.



Detalles de la floración



Plantas en R2

R3- Inicio de formación de vainas: Una vaina en uno de los cuatro nudos superiores del tallo principal mide 5 mm. de largo y presenta hojas totalmente desplegadas. En este momento no es raro encontrar vainas formándose, flores marchitas, flores abiertas y pimpollos en la misma planta. La formación de vainas se inicia en los nudos inferiores.



Vainas en R3

R4- Vainas completamente desarrolladas: Una vaina de 2 cm en uno de los 4 nudos superiores del tallo principal con hojas totalmente desplegadas. Alguna de las vainas de los nudos inferiores del tallo principal han alcanzado su máximo tamaño, pero en general la mayoría lo logra en R5.



Vainas formándose y completamente desarrolladas

En esta etapa comienza el periodo crítico del cultivo, cualquier deficiencia en humedad de suelo, nutrientes, luz, defoliación por orugas, enfermedades foliares, ataque de chinches, etc., entre esta etapa y R6 repercutirá en el rendimiento.

El período entre R4,5 y R5,5 es el más crítico ya que ha finalizado la floración y las vainas y semillas más jóvenes son más propensas a abortar en condiciones de stress.

R5- Inicio de formación de semillas: Una vaina, ubicada en uno de los 4 nudos superiores del tallo principal contiene una semilla de 3 mm de largo.

Entre las etapas R5 -R6 ocurren eventos importantes:

- La planta logra la máxima altura, número de nudos y área foliar.
- Se registra incremento del ritmo de fijación de Nitrógeno llegando al máximo en este período, comenzando luego a caer abruptamente.
- Las semillas inician un período rápido de acumulación de materia seca y nutriente.

100% de pérdida de área foliar (granizo) entre R5 y R5,5 puede provocar disminución del rendimiento en general del 75 %.

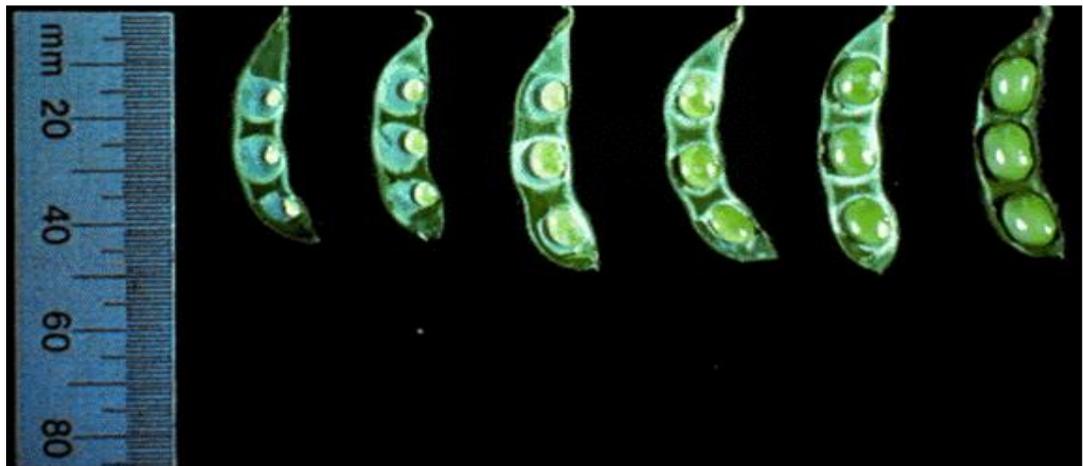


Vainas en R5

R6- Semilla completamente desarrollada: Una vaina en cualquiera de los cuatro nudos superiores del tallo principal contiene una semilla verde que llena la cavidad de dicha vaina, con hojas totalmente desplegadas.



Vainas en R6



R5  R6

Inicio de formación de semillas (R5), y semillas completamente desarrolladas (R6)

R7- Inicio de maduración: Una vaina normal en cualquier nudo del tallo principal ha alcanzado su color de madurez. La semilla alcanza la madurez fisiológica cuando ésta finaliza la acumulación de peso seco y generalmente junto con la vaina van perdiendo su coloración verde. La semilla en este momento contiene un 60 % de humedad.



Plantas en R7 (Madurez Fisiológica)

R8- Maduración completa: El 95 % de las vainas de la planta han alcanzado el color de madurez.

Se necesitan luego de R8 cinco a diez días de tiempo seco (baja humedad relativa ambiente), para que las semillas reduzcan su humedad a menos del 15%.³⁵



Cultivo en R8



Vainas y semillas en R8

³⁵Consulta Ing. Agr. Rubén E. Toledo, Fases del Desarrollo del Cultivo de la Soja, en Internet: http://agro.unc.edu.ar/~ceryol/documentos/soja/feno_soja.pdf (Abril 2013)



R6 → R7 → R8

Vainas en R6, R7 y R8

II.- Estimación no objetiva empleado por la EEAOC (Estación Experimental Agroindustrial Obispo Colombres) - Encuesta

Al realizar la investigación en el en el organismo que consideramos idóneo en el área de investigación agrícola - Estación Experimental Agroindustrial Obispo Colombres- logramos contactarnos con el departamento de Estadísticas y Economía del mismo, al cual dirigimos las siguientes preguntas:

1. ¿Qué hacen concretamente los encuestadores?
2. ¿Tienen un formulario con preguntas para cada productor entrevistado?
3. Si no existe tal formulario, ¿los encuestadores realizan mediciones de rendimiento?, ¿cuándo y cómo seleccionan los sitios para la medición?, ¿cuál es la instrucción de cómo hacerlo?, ¿qué tipo de datos deben informar los encuestadores a la EEAOC?, ¿cómo se obtiene la estimación del total de producción para la provincia a partir de los datos para los encuestadores?.

A continuación exponemos la **respuesta brindada por la EEAOC**:

"No tenemos encuestadores. Enviamos un cuestionario por mail a productores y/o asesores que sabemos en qué zona desarrollan su actividad y le preguntamos cuál es su rendimiento t/ha en finca y la estimación para el departamento en el que se encuentra. Y preguntas sobre la tecnología o los problemas de la campaña.

Hacemos un promedio por departamento y lo referenciamos a la superficie que se mide con imágenes satelitales.

Los departamentos se dividen en zona húmeda y seca.

Hacemos una estimación de rinde no una medición por eso no se hace chequeo a campo (no tenemos estructura para hacer esto)"

Cuestionario a productores y/o asesores utilizada por EEAOC

BURRUYACU - SOJA - CAMPAÑA 2008-2009

- 1 De la superficie sembrada con soja en Burruyacú (121.820 ha) qué *porcentaje* estima usted **no se cosechará** en ésta campaña?
Por favor escriba su estimación en el recuadro

%

- 2 En cuánto estima los rendimientos promedio, máximo y mínimo del cultivo de soja en la zona húmeda y en la zona seca de Burruyacú para la presente campaña y en que porcentajes se dieron los mismos ?
Por favor escriba su estimación en los recuadros según el área donde usted siembra

Zona húmeda 750-1000 mm de precipitación		
Rinde mínimo	t/ha	%
Rinde promedio	t/ha	%
Rinde máximo	t/ha	%
		100%

Zona Seca 500-750 mm de precipitación		
Rinde mínimo	t/ha	%
Rinde promedio	t/ha	%
Rinde máximo	t/ha	%
		100%

- 3 Realizó aplicaciones:

Control de roya: Cuantas y en qué zonas?

IGR:

Fertilizante:

- 4 Utilizó un esquema de herbicidas diferente para control de malezas? Qué productos?
-
-

- 5 A su juicio, cuáles fueron los principales factores que incidieron en los bajos rendimientos?
-
-

III.- Determinación del Tamaño de Muestra Recomendado

Al calcular el tamaño de la muestra a emplear para las estimaciones por lote se hizo uso de la fórmula que se recomienda para el caso de muestreo al azar simple:

$$n = (z \sigma / \alpha)^2 \text{ }^{37}$$

“Un aspecto práctico, que suele ser importante, es como conseguir el valor de s^2 o alguna buena estimación del mismo, cuando no se conoce de encuestas o censos o estadísticas permanentes que traigan esta información. Esto suele ser frecuente en lugares donde hay pocas estadísticas y poca experiencia estadística”.

En las dos consultas que se transcriben abajo se puede observar que el rango de la variable plantas por metro lineal puede “estimarse” en aproximadamente 10 plantas. A los fines de usar la relación empírica de magnitud que se menciona entre el rango y la desviación estándar de una variable que se distribuye como normal³⁸ se consideró un valor algo superior: 12 plantas. Según esa relación:

$$\sigma = R/4; \text{ en el caso que nos ocupa } \sigma = 12/4 = 3 \text{ plantas}$$

En el caso de la variable peso por 1000 granos de soja se ha considerado que los valores de la variable oscilan entre 90 y 180 gramos, por lo que el rango es de 90 gramos y la desviación estándar se “estima” en aproximadamente 22,5 gramos.

$$\sigma = 90/4 = 22,5 \text{ gramos}$$

A la vez se ha considerado que la precisión (α) requerida al nivel del 95 % de confianza para estar casi seguros que la muestra producirá estimaciones con una calidad estadística aceptable sea de más o menos una planta, en el caso de la variable número de plantas a cosecha por metro li-

³⁷ MEDINA, Ramón Darío, Loc. Cit., págs.. 44-45

³⁸ JULIANO, Víctor E. y ALCAIDE, Alejandro J., VARIABILIDAD ASIMETRÍA Y CURTOSIS, Cátedra de Estadística, Facultad de Ciencias Económicas, UNT, 2001.

neal y de más o menos 7,5 gramos en el caso del peso de 1000 granos de soja.³⁹

$$\alpha = 1 \text{ planta por metro lineal}$$

$$\alpha = 7,5 \text{ gramos por } 1000 \text{ granos de soja}$$

Si se trabaja al nivel de confianza del 95 %, el valor de z correspondiente en la distribución normal es 1,96; a los fines prácticos se considera el valor de $z = 2$

Con los datos anteriores y aplicando la fórmula arriba presentada para la determinación del tamaño muestral resulta⁴⁰:

$$n = (2 \cdot 3 / 1)^2 = 36 \text{ metros lineales con plantas en etapa R8}$$

$$n = (2 \cdot 22,5 / 7,5)^2 = 36 \text{ miles de granos de soja en etapa R8}$$

Finalmente y con el fin práctico de estimar la variable peso de 1000 granos de soja solo se ha de cosechar de cada metro lineal observado una de las plantas que se selecciona al azar. Los granos contenidos en las vainas de esas plantas habrán de proveer los 1000 granos que se han de pesar. A éste respecto puede usarse el método de muestreo sistemático para determinar la planta que habrá de suministrar la información de número de granos y peso por planta en cada metro lineal.

A continuación se transcriben algunas consultas a profesionales relacionadas con la actividad de donde se extrajeron datos que se usaron de referencia en la determinación del tamaño muestral.

³⁹ LIND-MARCHAL-MANSON, Estadística para Administración y Economía, 11ª Edición, Alfaomega Grupo Editorial, (México, 2005)

⁴⁰(Cfr.) D.A.S., Fraser, Op. Cit.

Pregunta realizada el: 30/11/2010

Nombre de quien consulta: Ariel

Consulta:

Estimado Héctor: quería hacerte las siguientes preguntas: 1- En la zona de Chivilcoy, Buenos Aires. Sembré la variedad DM 4670 a una distancia entre hilera de 52 cm., obteniendo una densidad de plantas emergidas de 20 plantas/metro ¿esta densidad está bien? ¿En caso de no ser la densidad adecuada que consecuencias tendré? 2- Para sembrar la misma variedad a la misma distancia entre hilera en siembra de segunda ¿Qué densidad se recomienda?

Respuesta dada al: 05/12/2010

Estimado Ariel: Van mis comentarios sobre tus consultas: 1-Como sabes, la soja es muy plástica a variaciones de la densidad de plantas. A una misma densidad de plantas por hectárea, esta plasticidad aumenta con el incremento de la uniformidad de la distribución de plantas y con la reducción del espaciamiento entre surcos. Es probable que para tu caso, entre 10 y 20 plantas por metro lineal a cosecha, no tengas una variación importante en rendimiento. Es decir que te encontrarías, casi en el límite superior de densidad recomendable. Entre los problemas pueden presentarse para esta densidad, estaría el vuelco y complicaciones sanitarias derivadas del vuelco, que son muy poco probables, para este cultivar y esta campaña. 2-Para la siembra de soja de segunda, puede funcionar bien el espaciamiento a 52 cm (con una densidad de 15 plantas por metro lineal a cosecha), en el caso que lo-gres un porte alto de planta. Si esto no ocurriera (recuerda que el espaciamiento entre surcos es más importante que la densidad de plantas), seguramente un menor espaciamiento entre surcos mejorará el rendimiento; por ejemplo a 35/38 cm y con una densidad de 10 plantas por metro lineal a co-

secha. Agradeciendo tu consulta y quedando a tu disposición, te saluda.
Héctor.

Pregunta realizada el: 08/12/2010

Nombre de quien consulta: Lautaro

Consulta:

Hola Héctor, la disponibilidad hídrica es buena, está sembrada los primeros de noviembre a 52cm, y la distribución es muy buena. Espero haber sido claro con los datos, muchas gracias.

Respuesta dada al: 09/12/2010

Consulta anterior: *Tengo campo en Arequito (Santa Fe) quisiera saber cuál es el stand ideal de plantas de soja 3900 sps. gracias..* **Respuesta anterior:** *Recuerda que en lo referente a arreglo espacial (espaciamiento entre surcos y densidad de plantas), es más importante el espaciamiento entre surcos, que depende de la altura más probable de plantas a lograr. En el caso de cultivares de GM III para tu zona, lo más recomendable son los espaciamiento de menos de 52 cm. y lo ideal es un espaciamiento entre 26 y 40 cm con sistema de distribución monograno. La densidad ideal (plantas por hectárea), generalmente se incrementa con la reducción del espaciamiento entre surcos y además varía con la campaña porque según la disponibilidad hídrica de cada campaña se modifica la altura de las plantas. Te agradeceré me indiques la calidad del ambiente, la fecha de siembra, el contenido de agua inicial a la siembra (estimativamente) y el espaciamiento entre surcos, para poder indicarte la densidad la densidad ideal. -----
----- Estimado Lautaro: Muchas gracias por el envío de la información adicional solicitada. Como sabes la soja presenta un amplio rango de densi-*

dades, en las que modifica mucho la arquitectura y el rendimiento de cada planta, pero con escasa variación en rendimiento por hectárea. Por tal motivo, estimo que para tu ambiente, fecha de siembra y un espaciamiento entre surcos de 52 cm, con un rango de densidades entre 10 y 20 plantas por metro lineal de surco a cosecha, SPS3900 presentará sus valores máximos de rendimiento y con una reducida variación. Agradeciendo tu consulta y quedando a tu disposición, te saluda. Héctor.

Pregunta realizada el: 08/03/2010

Nombre de quien consulta: María Pía

Consulta:

Hola Necesito saber cómo estimar el rinde de la soja, una vez me lo explicaron pero no recuerdo. Lo que tengo es la distancia entre líneas, nº de plantas por metro lineal y conté las chauchas pero no sé si saque bien las muestras, agradecería que me expliquen cómo. También tengo como peso promedio de las 1000 semillas 180 gramos, y no sé si esta acertado. Atte.-

Respuesta dada al: 22/03/2010

Estimada María Pía: Cuando se estima rendimiento en soja, es muy importante tomar varias muestras representativas del lote ó de sectores del lote (si el mismo tuviera zonas muy diferenciadas, por ejemplo lomas y bajos), de al menos un metro lineal de surco (con la densidad promedio del lote y con competencia de los surcos vecinos). Si los muestreos fueran representativos del lote y tuvieras un adecuado relevamiento de plantas y chauchas, te falta el número promedio de granos por chaucha y el peso promedio de los granos. Para determinar el número de granos promedio por chaucha, lo mejor es separar las chauchas por número de gra-

nos llenos (de 1, 2, 3 y 4 granos por chaucha), sumar granos llenos y dividirlos por el número de chauchas (generalmente los valores están entre 2,5 y 2,8 granos por chaucha y pueden ser mucho más bajos cuando hay ataques de chinches u otros problemas (sequía, enfermedades, otros). Con respecto al peso de los granos, si aún no estuviera terminado el llenado de granos, se deberá estimar, con el riesgo de cometer errores importantes; porque el peso de granos en soja varía generalmente entre 130 y 180 gramos los 1.000 granos (una estimación prudente sería de 150 gramos los 1.000 granos). Agradeciendo tu consulta y quedando a tu disposición, te saluda. Héctor.⁴¹

⁴¹ Consulta Planeta Soja, Manejo de Cultivares- Héctor Baigorri, en Internet: <http://planetasoja.com.ar/index.php?selenium=3&cons=1>

Índice Bibliográfico

A. General

CERISOLA, Juan A., GUTIERREZ, Jesús I., LOPEZ CLEIP DE SOSA, Aída, Propuesta de Diseño Muestral para Estimación de Rendimientos en la Producción de Limón, Cátedra de Estadística, Facultad de Ciencias Económicas, UNT, Diciembre 2001.

COCHRAN, William G., Técnicas de Muestreo, Compañía Editorial, México 1981.

DES, Raj, Las estructuras de las encuestas por muestreo, 1° Edición en Español, Fondo de cultura económica, (México, 1979).

JORDI, Casal y MATEU, Enric, Los sesgos y su control de, en Internet:
http://www.epidemiologia.com/epidemiologia/img/datos/21_06_59_3LosSesgos41.pdf .

LIND-MARCHAL-MANSON, Estadística para Administración y Economía, 11° Edición, Alfaomega Grupo Editorial, (México, 2005).

LOHR, Sharon L., Muestreo: Diseño y Análisis, International Thomson Editores (México, 2000).

PONTÓN, Rogelio, La importancia de la soja para Argentina, en Bolsa de Comercio de Rosario.

B. Especial

D.A.S. Fraser, Fundamentos y Técnicas de la Inferencia Estadística, Editorial Limusa, (México, 1976).

Ing. Daniel Rossana Pérez, Departamento Estadística y Economía, Estación Experimental Agroindustrial Obispo Colombes, Abril 2013.

JULIANO, Víctor E. y ALCAIDE, Alejandro J., VARIABILIDAD ASIMETRÍA Y CURTOSIS, Cátedra de Estadística, Facultad de Ciencias Económicas, UNT, 2001.

Planeta Soja, Manejo de Cultivares- Héctor Baigorri, en Internet:
<http://planetasoja.com.ar/index.php?selmenu=3&cons=1>

MEDINA, Ramón Darío, Técnicas Muestrales, Publicación XIII, Cátedra de Estadística, Facultad de Ciencias Económicas, UNT, (San Miguel de Tucumán, Abril 2010).

WALPOLE, Myers, Probabilidad y Estadística, 4° Edición, McGraw-Hill, (México, 1992).

C. Otras Publicaciones

Nutrición: La Soja, en Internet: <http://www.saludalia.com/nutrición/soja> (Abril 2013).

Consulta Cultivo de soja, en Internet:
http://es.wikipedia.org/wiki/Cultivo_de_soja#Inicio_y_desarrollo_del_cultivo_en_Argentina (Abril 2013).

Cultivo de Soja, Escala de los principales estados fenológicos, en Internet:
http://es.wikipedia.org/wiki/Cultivo_de_soja (Abril 2013).

Mapas Productivos-Argentina, en Internet: <http://www.fyo.com/especiales/soja12-13/mapa.html> (Abril 2013).

La Soja en la Argentina-Evolución, en Internet:
<http://www.fyo.com/especiales/soja12-13/mapa.html> (Abril 2013).

Bolsa de Cereales, Cotizaciones, Valores de Mercado, en Internet:
<http://www.bolcereales.com.ar/> - Valores de Mercado (Abril 2013).

Bolsa de Comercio de Rosario, Mercado de Granos, Cotizaciones, en Internet:
<http://www.bcr.com.ar/Pages/Publicaciones/infoboletinsemanal.aspx?IdArticulo=117>

Bolsa de Cereales, Informe Semanal, Semana 15/04/2013, en Internet:
<http://www.bolcereales.com.ar/> ESTUDIOS ECONOMICOS- Informe semanal, semana 15/04/2013

Ing. Agr. Rubén E. Toledo, Fases del Desarrollo del Cultivo de la Soja, en Internet: http://agro.unc.edu.ar/~ceryol/documentos/soja/feno_soja.pdf (Abril 2013).

Índice Analítico

Prólogo.....	Pág. 1
--------------	-----------

CAPITULO I

Aspectos Principales De La Soja

1.-. ¿Qué es la soja?	3
1.2.- Alimentos que derivan de la soja.....	4
2.- Importancia para las personas.....	5
2.1.- Beneficios y Propiedades	6
3.- Origen y Difusión.....	6
3.1.- Inicio y Desarrollo del Cultivo en Argentina.....	7
4.- Estadios Fenológicos Externos del Cultivo de Soja.....	8
5.- La Importancia de la Soja para Argentina	9
5.1.- Evolución.....	11
6.- La Reina de los Commodities Agrícolas	13
6.1.-Producción y Demanda Mundial	14
6.2.- Cotizaciones, precio y mercados.....	15
6.2.1.- M.A.G. y P. de la Nación-Valor de Mercado.....	15
6.2.2.- Consulta cotizaciones de precios Cámara Arbitral de Rosario.....	16
6.2.3.- Diferenciales de Soja.....	16
6.2.4.- Análisis: semana con subas para la Soja (semana del 15 al 19 de Abril del 2013).....	17

CAPITULO II

Marco Teórico

1.- Terminología a utilizar.....	19
1.1- Sesgo de Selección	20
1.2- Sesgo de Medición	21
1.3- El censo y La Encuesta	21

1.4- Muestreo por Juicio Personal.....	22
1.5- Muestreo Probabilístico.....	23
1.5.1.-Tipos de Muestras Probabilísticas	23
1.5.2.-Corrección y Precisión	24
1.6.- Métodos de Recolección de la Información.....	24
1.7.- Muestreo con Marco Imperfecto	24
2.- El Diseño de Muestras	25
2.1.- Estratificación.....	25
2.2.- Modo de Estratificar	25
2.3.- Número de Estratos.....	26
2.4.- Asignación de la Muestra a Estratos	26
2.5.- Muestreo Sistemático.....	27
3.- Encuestas de la Producción Agrícola	27
3.1.-Informe de la Cosecha	28
3.2.-Método de las Cosechas	28
3.2.1.- Selección de Campos.....	29
3.2.2.- Localización de los Cortes Muestrales	30
3.2.3.-Tamaño de los Cortes	30
3.2.4.- Forma del Terreno.....	31
3.3.- Experimento en los Campos de los Cultivadores.....	31
4.- Diseño Muestral para Encuestas de rendimientos	32
4.1.- Estratificación.....	32
4.2.- Tamaño de la muestra.....	33

CAPITULO III

Propuesta de Diseño Muestral para Estimación de Rendimientos en la Producción de Soja

1.- Introducción.....	35
2.- El procedimiento del muestreo.....	36
2.1.- Errores muestrales y no muestrales.....	37
2.2.-Medida del error muestral.....	39
3.- Planteamiento del problema.....	41
4.- Revisión de métodos de estimación empleados	43

4.1.- Estimaciones visuales y estimación por muestreo: críticas y ventajas.....	43
5.- Estimación de encuestas por muestreo	47
5.1.- Formulación de objetivo	47
5.2.- Población y unidades investigadas	48
5.3.- Principales variables investigadas por la encuesta	49
5.4.- Grado de precisión	50
5.5.- Periodo de referencia.....	50
5.6.- Presupuesto	51
6.- Marco muestral.....	52
6.1.-Diseño y selección de muestra	53
6.2.- Métodos de muestreo.....	54
6.3.- Tamaño de la muestra	58
6.4.- Diseño de instrumentos de recolección	63
6.5.- Organización del trabajo de campo	67
6.6.- Recursos humanos y capacitación.....	68
6.7.- Validación y crítica.....	70
7.- Grado de precisión de las estimaciones-Errores de muestreo.....	74
8.- Nota al diseño muestral propuesto	77
9.- Conclusión.....	79

ANEXO

I.- Estadios Fenológico Externos de la Soja.....	82
II.- Estimación no Objetiva Empleado por la EEAOC- ENCUESTA.....	90
III.- Determinación del Tamaño de Muestra Recomendado.....	92
Índice Bibliográfico.....	98
Índice Analítico.....	100