

TRABAJO PRACTICO N° 1

Muestreo y Organización y Presentación de datos

I.-Muestreo

Objetivos

- Que el alumno comprenda la importancia de obtener buenos datos.
- Que el alumno comprenda la importancia del diseño de muestreo, la obtención de datos y la aplicación a los problemas presentados.

Ejercicio N° 1:

Clasifique las siguientes variables como cuantitativas (discretas o continuas), cualitativas (nominales u ordinales) y derivadas (tasas, razones o índices)

Tabla 1: Clasificación de variables

Variables	Clasificación
a) Kilómetros recorridos por un tractor con 10 litros de gas oil.	
b) Milímetros de precipitación caídos en el mes de enero en Paraná.	
c) Número de chauchas por planta de soja.	
d) Razas de ganado lechero en Argentina (Holando, Jersey, Pardo suizo).	
e) Calidad de grano de soja a cosecha. (1= Excelente, ..., 5 = Malo).	
f) Longitud de entrenudos de trigo pan (cm.).	
g) Número de plantas de avena por metro cuadrado.	
h) Concentración de cloruros disueltos en agua de pozo (ppm)	
i) Rendimiento en kg/ha de sorgo granífero.	
j) Presencia de gripe aviar en países de la Unión Europea	
k) Eficiencias de conversión en ganado lechero: <i>Leche producida / Kg. Alimento Balanceado (l/kgMS)</i>	
l) Porcentaje de Grasa Butirosa (%).	
ll) Número total de muertes fetales durante un año / número total de alumbramientos durante un año por 1000.	
m) Criaderos de semilla soja. (Pioneer, Monsanto, Nidera, La tijereta).	
n) Distancia entre surcos en el cultivo de soja.	
Ñ) Días Emergencia – Floración de híbridos de girasol.	
o) Temperatura media mensual durante el ciclo de cultivo de trigo.	
p) Intensidad de verticilosis (escala de 0 a 5), según la sintomatología en hoja de híbridos de girasol.	
q) Número de vacas en ordeño por tambo en la provincia de Entre Ríos.	
r) Kilogramos de materia seca / ha / año de un pastizal natural.	
s) Porcentaje de Proteína Bruta (%).	

Ejercicio N° 2:

Los datos que se presentan en la Tabla corresponden a un relevamiento realizado a los alumnos inscriptos en Estadística y D. Experimental FCA – UNER durante el año 2002. Supóngase que desea analizar algunas características del grupo.

- Identifique la población en estudio y unidad de muestreo.
- Identifique y Clasifique las variables en estudio.

Nota: ID: n° Individuo

Sexo: F = Femenino; M = Masculino

Origen: P = Paraná; I = Interior.

Procedimiento de Muestreo al Azar. Uso De Tabla de Números Aleatorios: El procedimiento correcto para extraer una muestra simple al azar, consiste en asignar un número de serie a cada elemento de la población (ID), escogiendo el número de serie al azar por medio de una tabla de números aleatorios, y usando como muestra aquellos elementos de la población cuyos números de series son escogidos.

Para ilustrar el uso de la Tabla de una población de $N = 72$ datos (Ej. N° 2) se extrae una muestra simple al azar $n = 25$.

- Se selecciona una de las páginas de la tabla de números aleatorios.
- Se deja caer un lápiz con la punta hacia abajo, para determinar el dígito correspondiente al punto de partida. La tabla puede ser leída vertical u horizontalmente.
- A partir de dicho dígito, sin saltar ninguno, se obtiene en forma horizontal sucesiva **n** números de dos cifras entre 01 y **N**. No se registra un número de individuo que aparezca por segunda vez (muestreo simple al azar sin reposición).
- Se toma de la lista de la población aquellos datos correspondientes a los individuos cuyos números de serie fueron seleccionados en el punto anterior. Se elabora la tabla.

Tabla 2: Relevamiento Inscriptos en Estadística Diseño Experimental.

ID	Sexo	Origen	Trabaja	Materias aprobadas	Materias regularizadas	Materias cursando	Peso (Kg.)	Altura (cm)	Año ingreso
----	------	--------	---------	-----------------------	---------------------------	----------------------	---------------	----------------	----------------

Ejercicio N° 3:

En cada uno de los siguientes enunciados indique la población en estudio, la muestra, las unidades de muestreo y la/s variables evaluadas.

a) Encuesta permanente de hogares (2002) en localidades de 100.000 y más habitantes, relevándose 890 hogares.

Áreas temáticas: Características Sociodemográficas, Ingreso de los Hogares, Mercado de Trabajo, Educación, Pobreza.

Población:

Muestra:

Unidad de Muestreo:

Variable:

b) Relevamiento censal nacional (2001) donde se censaron todas las personas que se encontraban presentes en el hogar a la hora cero del día del censo, residieran habitualmente o no en esa vivienda.

Cobertura geográfica disponible: total nacional, provincias, departamentos, localidades.

Áreas temáticas: Características Sociodemográficas, Necesidades Básicas Insatisfechas, Vivienda y Condiciones Sanitarias, Educación, Empleo y Capacidad de Subsistencia, Salud.

Población:

Muestra:

Unidad de Muestreo:

Variable:

c) El objetivo del presente estudio fue detectar animales portadores del virus del cólera porcino (VCP) luego de haber sido vacunados contra el cólera porcino (CP) en una granja de la provincia de Entre Ríos. Se colectaron muestras de suero de 166 lechones de 7 a 9 semanas de edad y que fueron vacunados contra el Cólera Porcino 15 días antes del muestreo. La detección de anticuerpos contra el VCP se realizó mediante la prueba de ELISA de competición. Se obtuvieron muestras adicionales de los animales que resultaron negativos o dudosos a la prueba de ELISA a los 30 días posteriores al primer muestreo para la detección del VCP mediante la prueba de inmunofluorescencia en linfocitos cultivados.

El 88% de los animales estudiados presentaron anticuerpos contra el VCP mientras que el 3% y el 9% de los animales presentaron resultados dudosos y negativos a anticuerpos, respectivamente.

Población:

Muestra:

Unidad de Muestreo:

Variable:

II. –Organización y presentación de datos.

Objetivos

- Que el alumno logre reducir los datos en forma compacta, comprensible y comunicable por la distribución de frecuencias.
- Que el alumno sea capaz de interpretar la información presentada en forma de cuadros y gráficos.
- Que el alumno sea capaz de interpretar y analizar la naturaleza de los datos.
- Que el alumno sea capaz de analizar, interpretar y concluir las salidas de los resultados de los paquetes estadísticos.

Ejercicio N° 4:

Los datos que se presentan en las tablas 3 y 4, corresponden al relevamiento de datos cuantitativos sobre la presencia de mujeres en la Universidad de Entre Ríos. El año seleccionado para el estudio fue 1995. La información fue proporcionada por diferentes fuentes de la UNER.

Tabla 3: Participación femenina en la matricula de grado en las Facultades – UNER

Facultades.	Total de Mujeres	Total de alumnos
Administración	775	1447
Agropecuarias	141	444
Alimentos	136	249
Bromatología	345	566
Económicas	961	1831
Educación	742	1116
Ingeniería	237	840
Salud	398	520
Trabajo Social	376	408
Total matricula	4111	7421

- Analice si el relevamiento realizado corresponde a una población o a una muestra. Justifique describiendo la población y/o la muestra o ambas.
- Identifique y Clasifique la variable.
- Calcule el porcentaje de mujeres en cada facultad.
- En que facultad se observa el mayor porcentaje de mujeres inscriptas.
- Sobre la matricula de grado de toda la Universidad que facultad tuvo el mayor número de inscriptos en 1995.
- Que porcentaje representan las mujeres en el total de la matricula de la Universidad?

Tabla 4: Participación femenina/masculina en el personal de las Facultades de Ciencias Agropecuarias.

Personal	Agropecuarias		Total	
		Total mujeres	Total varones	
Cuerpo docente	91	105		
Investigadores	55	55		
Extensionistas	12	19		
Personal superior	2	4		
Personal no docente	14	27		
Consejo Superior	1	4		

Total

Fuente: Ciencia, Docencia y Tecnología. N° 20 Año XI– Mayo 2000 UNER PID Relaciones de genero en la Universidad de Entre Ríos

- Identifique y clasifique la/s variables.
- Que porcentaje de mujeres forman parte del personal de la FCA?
- Que porcentaje de mujeres se encuentra en cada categoría de personal?
- Identifique en el grafico 1: variables, sus categorías y valores de porcentajes representados

Grafico 1: Porcentaje del Personal según Sexo y Categoría

Ejercicio N° 5:

Los datos que se presentan en la Tabla 5, corresponden al Relevamiento del Arbolado Urbano de la ciudad de Paraná, realizado por la Cátedra de Espacios Verdes FCA–UNER, en el año 1990. De las principales especies, por su abundancia o buena adaptabilidad ecológica, se tomaron las alturas a través de un muestreo al azar clasificándolas en baja (< 5 metros) medio (5 – 10 metros) y alta (> 10 metros)

Información adicional: el municipio de Paraná se dividió en 10 sectores para el relevamiento, el total de árboles relevados fue de 46217, número de especies halladas = 110.

Tabla 5: Especies relevadas según alturas.

Especie	Baja	Media	Alta	Total
Crespón	22	2	0	24
Jacaranda	0	32	26	58
Fresno	13	103	2	118
Lapacho	0	13	3	16
Ligustro	24	26	0	50
Palito dulce	0	16	0	16
Pino	1	9	0	10
Seibo	3	8	0	11
Tipa	0	10	26	36
Total				

- Analice si el relevamiento realizado corresponde a una población o a una muestra. Justifique describiendo la población y/o la muestra o ambas.
- Identifique y clasifique la/s variable/s.
- Calcule los porcentajes de altura para cada especie.
- Considerando la altura media, identifique las especies que poseen el mayor y menor porcentaje.
- Que altura de plantas posee el mayor porcentaje de individuos.

Ejercicio N° 6:

Se presentan los datos de Días transcurridos desde Emergencia – Floración, Altura de plantas (cm), Numero de plantas /ha, Porcentaje de Humedad en grano, Rendimiento (Kg/ha) y Porcentaje de aceite (Tabla n° 6) obtenidos en ensayos de Girasol *Paraíso – 30*, realizados en campos de productores del Sur de la provincia de Buenos Aires y la Pampa. Campaña 2001/02.

Tabla n° 6: Variables evaluadas en ensayo de Girasol Paraíso 30

Id	Emer – Flor	Altura	Plantas/ha	Humedad (%)	Rto (Kg/ha)	Aceite (%)
1	67	155	49524	10.3	4389	49.50
2	65	195	56190	11.5	4317	50.84
3	64	185	54286	11	4274	45.22
4	64	185	48571	10.6	4156	47.77
5	65	175	56190	10.7	4133	51.49
6	57	163	56190	9	4112	49.55
7	72	170	53333	10.8	4078	47.05
8	65	155	49524	8.6	3987	51.19
9	70	175	53333	8.8	3935	47.80
10	63	170	51429	11.3	3700	50.22
11	67	195	55238	9.8	3931	48.01
12	71	190	46667	9.6	3898	46.25
13	65	170	50476	11.4	3874	48.07
14	64	170	59048	10	3841	48.85
15	71	210	51429	10.2	3832	51.29
16	67	180	54286	10.8	3801	50.36
17	67	203	56190	10.2	3793	44.40
18	61	145	56190	10.1	3736	45.91
19	66	158	50476	11.4	3680	45.40
20	68	190	54286	11.1	3661	47.89
21	64	150	56190	10	3656	50.96
22	65	180	51429	10.9	3656	49.76
23	68	200	48571	8.9	3628	44.58
24	69	180	55238	10.2	3618	48.15
25	71	170	50476	8.8	3600	48.65
26	69	190	60000	10.4	3556	45.29
27	64	166	53333	9	3538	49.41
28	64	160	42857	9.3	3495	52.20
29	66	170	49524	9.5	3449	49.72
30	67	195	57143	9.7	3449	51.22
31	65	180	53333	9.6	3442	50.70
32	62	145	48571	10.4	3437	49.01
33	63	165	55238	10.6	3423	50.65
34	63	167	54286	8.9	3331	53.37
35	62	165	58095	9.9	3288	51.74
36	64	190	57143	10.9	3287	47.16
37	66	170	49524	9.5	3195	51.18
38	66	165	51429	10.5	3179	48.48
39	64	160	51429	9.7	3167	52.19

40	65	190	56190	10.2	3164	48.52
41	67	170	54286	10.4	3160	49.74
42	63	185	54286	11.3	3158	47.34
43	71	190	51429	10	3310	43.54
44	64	170	46667	9.6	3068	49.65
45	71	190	60000	9.1	3062	50.74
46	65	170	54286	10	3022	46.06
47	66	200	50476	13	2967	46.45
48	65	180	52381	9.8	3320	50.79
49	70	195	57143	9.8	3450	48.48
50	68	165	49524	9.8	2845	50.25
51	66	175	50476	9	2828	49.92
52	63	175	53333	9	3290	49.63
53	65	145	58095	9.4	2802	49.20
54	64	165	40000	10.1	2751	51.68

- Identifique y clasifique las variables de la tabla n° 6.
- Complete los datos de la variable ***Días emergencia-floración***, en tabla de distribución de frecuencias. (fa, fr, fr%, Fa, Far, FRA%)
 - Interprete en la clase 5: fa, fr, fr%, Fa.
 - Interprete el valor de Marca de Clase de la clase 3.

Distribución de frecuencias Variable Emer – Flor

Clase LI LS MC fa fr fr% FAA FRA FRA%

1 57 60 0,02 1 0,02

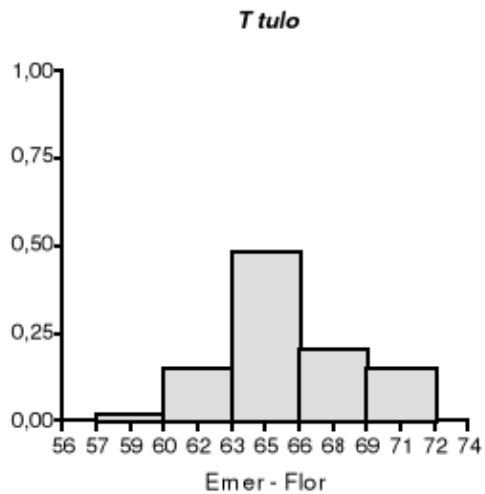
2 61,50 8 0,15 9 0,17

3 63 66 26

4 67,50 0,20 46 0,85

5 69 72 70,50 8 0,15 54 1,00

◊ Que frecuencias se han representado en el grafico siguiente? Interprete la forma de la distribución.



- Para la variable **Rendimiento en kg/ha**
 - ◆ Definir la amplitud de los intervalos de clase.
 - ◆ Completar la tabla.
 - ◆ Interpretar de la clase 4: fa, fr, fr%, FA.
 - ◆ Indique que tipo de grafico se presenta, y que frecuencias se han representado. Trace en el mismo el polígono e interprete la forma.

Distribución de frecuencias rendimiento en kg/ha

Clase LI LS MC fa fr fr% FA

1 2751 2985 5

2 2985 3219 9

3 3219 3453 12

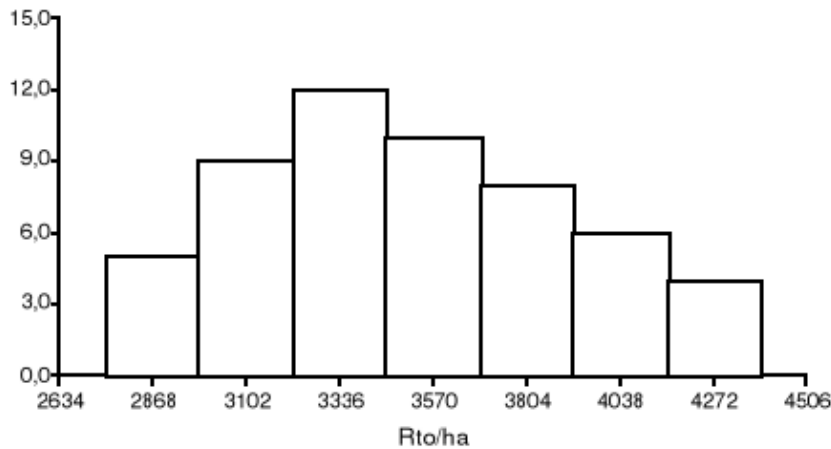
4 3453 3687 10

5 3687 3921 8

6 3921 4155 6

7 4155 4389 4

Grafico:



- Para la variable % **Humedad en grano**, se presenta el histograma de frecuencias.
 - ◆ Para la construcción de la tabla de frecuencias, indique: amplitud de intervalos, límites de intervalos de clase, marca de clase.
 - ◆ En que intervalo, se encuentran la mayor cantidad de datos.
 - ◆ Represente gráficamente el polígono de frecuencias.
 - ◆ Interprete la forma de la distribución.
- & Los datos de la variable **Altura de plantas**, se presentan en una tabla de distribución de frecuencias. (fa, fr, Fa, FAA, FRA)
 - ◆ Cual es el objetivo de una tabla de distribución de frecuencias. Que información puede extraer de la misma?.
 - ◆ Complete la tabla e indique y calcule que elementos necesita para su construcción.
 - ◆ Interprete del intervalo de mayor frecuencia absoluta: fa, fr, FAA, FRA.
 - ◆ Por debajo de que altura de plantas se encuentra el 0.85 de las plantas?.
 - ◆ Cuantas plantas tuvieron una altura entre 193.75 y 201.88 cm? A que medida nos estamos refiriendo?
 - ◆ Cual es el significado de MC5 = 181.56 cm?
 - ◆ Indique como se denomina el grafico que se presenta a continuación.

Distribución de frecuencias de Altura (cm)

Clase LI LS MC fa fr FAA FRA

1 145.00 153.13 149.06 4 0.07 4 0.07

2 5 9

3 161.25 169.38 165.31 8 0.15 17 0.31

4 169.38 177.50 14 0.57

5 177.50 185.63 181.56 8 0.15 39

6 189.69 7 0.13 46 0.85

7 193.75 201.88 6 0.96

8 201.88 210.00 205.94 2 0.04 54 1.00

♦ &Para la variable **Número de plantas/ha**

- ◊ Determine la amplitud de los intervalos.
- ◊ Establezca para cada intervalo los límites inferior y superior.
- ◊ Complete las frecuencias faltantes.
- ◊ Cuando decimos que el 15 % del número de plantas /ha estuvo entre 47500 y 50000, a que medida nos estamos refiriendo?.
- ◊ Cual es el valor de la variable por debajo del cual se encuentra el 67 % del número de plantas /ha.?

Distribución de frecuencias

Clase LI LS MC fa fr FAA FRA

1 41250 1

2 43750 1

3 46250 2 0.04 4 0.07

4 48750 8 0.15 12 0.22

5 51250 12

6 53750 12 0.22 36 0.67

7 56250 13

8 58750 5 0.09 54 1.00

Ejercicio N° 7:

En la tabla n° 7 se presentan los datos de producción de un cultivar de frutilla.

- ♦ Identifique y Clasifique las variables.
- ♦ De las tablas de distribución de frecuencias presentadas, responda:
 - El valor del LI 1 corresponde al X min. de la serie?.
 - Cual es la amplitud de los intervalos para cada variable?
 - En cada una de las variables determine el intervalo y el valor de la mayor frecuencia porcentual.
 - Interprete para cada una de las variables las frec. absolutas y acumuladas de la clase n° 3.
 - Que grafico utiliza para representar: fa y Fa%.

Distribución de frecuencias Número de frutos

Clase LI LS MC fa fr FAA FRA

1 32.00 35.00 33.50 6 0.24 6 0.24

2 35.00 38.00 36.50 7 0.28 13 0.52

3 38.00 41.00 39.50 9 0.36 22 0.88

4 41.00 44.00 42.50 2 0.08 24 0.96

5 44.00 47.00 45.50 1 0.04 25 1.00

Distribución de frecuencias Rdto fruto comerciales (g/pl)

Clase LI LS MC fa fr FAA FRA

1 470.00 522.20 496.10 9 0.36 9 0.36

2 522.20 574.40 548.30 5 0.20 14 0.56

3 574.40 626.60 600.50 4 0.16 18 0.72

4 626.60 678.80 652.70 5 0.20 23 0.92

5 678.80 731.00 704.90 2 0.08 25 1.00

Distribución de frecuencias Rdto Total (g/pl)

Clase LI LS MC fa fr FAA FRA

1 550.00 604.00 577.00 9 0.36 9 0.36

2 604.00 658.00 631.00 5 0.20 14 0.56

3 658.00 712.00 685.00 3 0.12 17 0.68

4 712.00 766.00 739.00 7 0.28 24 0.96

5 766.00 820.00 793.00 1 0.04 25 1.00

Tabla N° 7: componentes de rendimiento de frutillas.

Id	Número de frutos comerciales por planta	Rendimiento de frutos comerciables (g/planta)	Rendimiento total (g/planta)
1	32	470	550
2	37	540	610
3	41	630	720
4	38	550.2	625.3
5	39	625.4	714.3
6	35	489	566
7	36	503	590
8	34	503	582
9	37	534	614
10	35	513	591
11	41	611.5	696.2
12	41	638.7	729.9

13	41	628.6	711.7
14	45	673.3	758.5
15	41	632.6	713.6
16	44	730.9	820
17	35	474.8	572.7
18	42	691.2	751.1
19	41	608.2	696.2
20	37	534	624.2
21	40	470.1	550
22	37	534	620.1
23	39	599.3	714.3
24	32	475.1	551.9
25	36	513.8	592

Ejercicio 8:

La tabla que se presenta a continuación corresponde a la distribución de frecuencias empíricas del número de plantas de *Carex spp* por cuadrado de superficie.

Nº plantas/ cuadrado superficie	frecuencias observadas	fr	FA	FAR
0	185			
1	118			
2	97			
3	54			
4	32			
5	9			
6	5			
7	3			
8	1			

- ◆ Identifique y clasifique la variable en estudio.
- ◆ Explique por que se utilizo una tabla de distribución de frecuencias valor por valor en lugar de una tabla de frecuencias con intervalos de clase.
- ◆ Complete la tabla de distribución de frecuencias.
- ◆ Interprete para $X = 5$, f_a , f_r , FA y FAR .
- ◆ Indique en cuantos cuadrados se han encontrado 2 plantas de *Carex spp*. A que medida nos referimos.?
- ◆ Cual es el porcentaje de cuadrados en el que se han encontrado 6 ó mas plantas?
- ◆ Puede definir la forma de la distribución.?

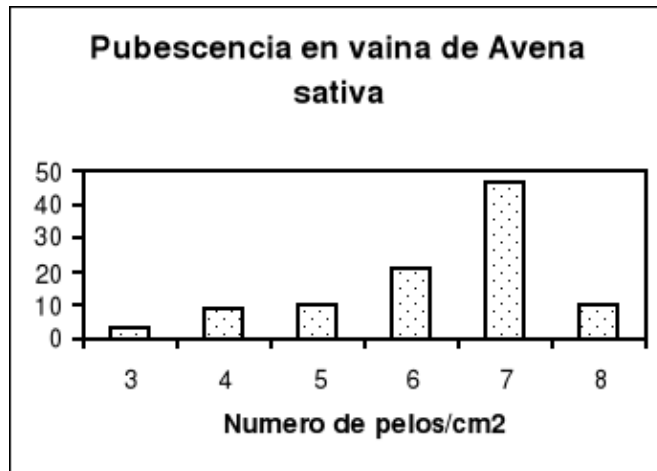
Ejercicio 9:

Los datos que se presentan a continuación corresponden a un estudio sobre la densidad de pubescencia: Número de pelos/cm², del tercio superior de la vaina de la primer hoja de un

cultivar de Avena sativa.

Num. de pelos/cm2	fa	fr	FA	FRA
3	3		3	
4	9			
5	10		22	
6	21			
7	47		90	
8	10		100	

- ◆ Identifique y clasifique la variable en estudio.
- ◆ Complete la tabla con las frecuencias faltantes.
- ◆ Cual es el numero de pelos mas frecuente?
- ◆ En el grafico que se presenta a continuación, indique: nombre del grafico, frecuencias que se han representado. Interprete la distribución de los datos.



Respuestas a los problemas seleccionados

Ejercicio 6:

Distribución de frecuencias de Altura (cm)

Clase LI LS MC fa fr FAA FRA

1 145.00 153.13 149.06 4 0.07 4 0.07

2 153.13 161.25 157.19 5 0.09 9 0.17

3 161.25 169.38 165.31 8 0.15 17 0.31

4 169.38 177.50 173.44 14 0.26 31 0.57

5 177.50 185.63 181.56 8 0.15 39 0.72

6 185.63 193.75 189.69 7 0.13 46 0.85

7 193.75 201.88 197.81 6 0.11 52 0.96

8 201.88 210.00 205.94 2 0.04 54 1.00

Distribución de frecuencias Número de plantas/ha

Clase LI LS MC fa fr FAA FRA

1 40000 42500 41250 1 0.02 1 0.02

2 42500 45000 43750 1 0.02 2 0.04

3 45000 47500 46250 2 0.04 4 0.07

4 47500 50000 48750 8 0.15 12 0.22

5 50000 52500 51250 12 0.22 24 0.44

6 52500 55000 53750 12 0.22 36 0.67

7 55000 57500 56250 13 0.24 49 0.91

8 57500 60000 58750 5 0.09 54 1.00

El símbolo &indica que las soluciones a estos problemas están en la sección ***Respuestas a los problemas seleccionados*** al final del Trabajo Práctico.