



---

---

## TRABAJO PRÁCTICO N° 1 POBLACIÓN Y MUESTRA. MUESTREO

**Ejercicio n° 1:** Para cada una de las siguientes situaciones de investigación describa:

- a) ¿Cuál es la población?
  - b) ¿Cuál es la muestra?
  - c) ¿Cuál es la unidad estadística, de análisis o de observación?
  - d) ¿Cuál es el método de muestreo más apropiado?
- 
1. En la encuesta permanente de hogares (EPH) se realiza un relevamiento de las condiciones estructurales de las viviendas para determinar las condiciones en el que habitan las familias.
  2. En una encuesta política para tratar de pronosticar el resultado de una elección en Capital Federal, se seleccionaron 1000 personas del padrón y se les preguntó la preferencia política.
  3. En un establecimiento apícola de 300 colmenas se realizó un estudio para conocer el contenido de nitrofurano en miel comercial, para lo cual se seleccionaron 10 colmenas al azar y se midió el contenido en cada una de ellas.
  4. Mientras un camión espera para descargar en un silo de granos la producción de trigo de un campo, se toma una muestra de semillas de trigo. Sobre la base de los resultados se clasifica a la semilla en excelente, muy buena, buena o mala, y se establece el precio de toda la carga del camión.
  5. El Ministerio de Trabajo de la Provincia vigila constantemente las cifras sobre empleo, desempleo, salario, movimiento laboral, etc. En base a una muestra de 100 trabajadores, el Ministerio presenta las siguientes cifras correspondientes a un mes en particular: "2 de cada 100 empleados cambiaron de trabajo el último mes; el sueldo mensual de los obreros metalúrgicos fue de \$7800; el número promedio de horas de trabajo en las empresas alimenticias fue 40,2".
  6. Se desea estudiar la relación existente entre el rendimiento de los alumnos de nivel secundario y la disponibilidad de recursos tecnológicos que posee en el hogar.
  7. Se desea estimar la edad promedio de los ingresantes a la U.Na.F. en 2013. (considere una situación en la cual se tiene la base de datos total de los ingresantes y otra en la que no se cuenta con tal información)

**Ejercicio n° 2:** Clasifique las siguientes variables

VARIABLE	CLASIFICACIÓN
Producción de leche (KG/GB) por vaca por período de lactancia	
Contenido de sodio en muestras de agua	
Concentración de colesterol en suero de humanos (mg/dl)	



Sabor del vino para diferentes cepas de uva	
Variedad de tomate	
Razas de bovinos	
Precio máximo de la tonelada de maíz en los últimos meses	
Color de las uvas	
Perímetros basales de troncos de quebracho blanco.	
Temperatura media anual	
Coficiente intelectual de alumnos universitarios.	
Altura de una montaña	
Duración de una prueba	
Número de semillas de alfalfa por metro de surco sembrado	
Temperaturas registradas en un laboratorio durante una semana	
Número de materias aprobadas con 4 por estudiantes de la FRN	
Número de vacas en ordeño por tambo	
Porcentaje de proteína bruta (%)	
Concentración de proteínas (baja, media, alta) en leche de cabra	
Estado (presñada o vacía) de una vaquillona (al tacto)	

**Ejercicio n° 4:** Una compañía cuenta para su servicio con 20 camiones similares numerados del 15 al 34 cuyos datos figuran en la siguiente tabla:

TONELAJE ALTO		TONELAJE MEDIO		TONELAJE BAJO	
Camión	Consumo (Km/lt)	Camión	Consumo (Km/lt)	Camión	Consumo (Km/lt)
15	2.6	18	5.6	25	8.1
16	3.0	19	5.2	26	7.5
17	2.5	20	5.0	27	7.6
		21	5.7	28	6.9
		22	4.9	29	7.3
		23	5.2	30	7.7
		24	5.5	31	7.4
				32	7.1
				33	7.5
				34	7.6

El gerente sospecha que hay un exceso en el consumo de combustible y para realizar una verificación técnica rápida decide tomar una muestra de tamaño  $n=6$ . Determine qué unidades formarán parte de la muestra utilizando:

- Muestreo aleatorio simple
- Muestreo estratificado proporcional
- Muestreo estratificado no proporcional
- Muestreo sistemático.



**Ejercicio nº 5:** Los siguientes datos corresponden a cincuenta niños que estuvieron bajo control y seguimiento para evaluar la incidencia del riesgo cardiovascular. Para ello se realizaron mediciones de distintos indicadores como PESO (en kg), EDAD (en años), TALLA (en centímetros) y COLESTEROL (en mg/dl).

ID	SEXO	EDAD	PESO	TALLA	COLEST
1	M	7,12	23,5	125,5	280
2	M	9,82	27,7	131	269
3	M	7,06	23,8	121,8	268
4	F	9,57	54,6	151	252
5	M	7,31	17,4	111,8	243
6	M	8,47	28,4	133,6	237
7	M	8,3	23,6	128,5	235
8	M	6,31	19,1	111,5	234
9	F	10,75	24,8	136,8	230
10	F	9,13	16,8	121,1	227
11	F	6,16	19,3	117,5	226
12	F	11,16	35,9	142,3	225
13	F	9,68	29,5	139,1	225
14	M	7,63	29,5	127,1	222
15	M	10,08	50,7	139,7	222
16	M	6,78	28,1	131,4	219
17	M	9,61	29,3	130,8	217
18	F	8	34,3	130	217
19	F	9,49	31,2	127,6	216
20	F	8,03	29,9	126,8	216
21	F	6,16	21,9	109	215
22	F	6,79	17,4	111,8	214
23	M	10,58	52,8	152,2	213
24	F	7,08	35,9	125	212
25	F	7,94	23,5	125,6	211

26	M	10,45	31,4	132,9	211
27	F	9,72	34,3	138,1	211
28	F	7,88	38,1	128,5	210
29	F	7,11	24	120,5	210
30	F	9,51	41,6	147,5	209
31	M	11,61	55,8	145,9	209
32	M	9,74	31,9	140,3	208
33	F	6,67	17,9	114,1	208
34	M	8,53	58,8	137,7	207
35	M	10,28	58,7	151,6	207
36	M	5,94	17,9	102,6	206
37	M	9,37	29,8	134,1	205
38	F	7,98	24,2	120,6	205
39	M	9,02	36,1	129,6	204
40	F	8,73	33	138,6	204
41	F	8,9	31,8	137,4	204
42	M	10,52	31,9	137,6	203
43	M	5,91	30,6	117,4	202
44	M	10,32	34,7	138,5	202
45	F	10,95	35,2	142,7	202
46	F	9,99	44,9	135,3	201
47	M	9,62	31,9	141,1	201
48	M	9,79	34,1	144,1	201
49	F	9,02	19,5	117,2	200
50	M	8,65	28,6	133,3	200

- a) Suponiendo que estos 50 niños forman una población, seleccione una muestra de tamaño 15 usando MAS y Muestreo sistemático.
- b) Estratifique la población según peso de la siguiente manera
- BAJO PESO: Menos de 23kg  
 PESO NORMAL: Más de 23 y menos de 50kg.  
 SOBREPESO: Más de 50kg.
- c) Seleccione una muestra de tamaño 15 usando el muestreo estratificado proporcional.



---

d) Estratifique la población según COLESTEROL de la siguiente manera

NIVEL DESEABLE: Menos de 200 mg/dl

NIVEL DE RIESGO: Entre 200 y 220mg /dl

NIVEL DE ALTO RIESGO: Más de 220 mg/dl.

e) Seleccione la muestra usando muestreo estratificado proporcional.



## TRABAJO PRÁCTICO N° 2 ORGANIZACIÓN Y PRESENTACIÓN DE DATOS

**Ejercicio n° 1:** Los siguientes datos se refieren al número de dientes por hoja en bulbos de ajo:

4	2	1	1	2	3	5	6	4	5
3	4	5	6	3	3	1	2	3	4
5	5	6	3	1	1	1	2	2	3
3	3	4	5	6	2	2	3	4	3

A partir de los mismos se pide:

- Identifique y clasifique la variable.
- Construya la tabla de distribución de frecuencias y complete las columnas de frecuencias acumuladas, relativas y relativas acumuladas.
- Grafique las frecuencias absolutas y las acumuladas.

Responda:

- ¿Cuántos bulbos de ajo poseen 3 dientes?
- ¿Cuántos tienen menos de 4 dientes?
- ¿Qué porcentaje presentan menos de 3 dientes?
- ¿Cuál es el número de dientes que presenta la mayor frecuencia de aparición?

**Ejercicio n° 2:** Los siguientes datos corresponden al número de hijos de mujeres mayores a 20 años que asistieron a un centro de salud.

0	2	3	3	4	0	2	1	3	4	1	3	3	0	1	3	4	3	0	1	1	2
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

- Presente los datos en una tabla de distribución de frecuencias.
- Complete las frecuencias acumuladas y las relativas.
- ¿Cuántas mujeres no tienen hijos?
- ¿Cuántas mujeres constituyen la muestra?
- ¿Qué porcentaje de mujeres tiene menos de 3 hijos?
- ¿Qué porcentaje de mujeres tiene 4 hijos?
- ¿Cuántas mujeres tienen al menos un hijo?
- ¿Cuántas tienen a lo sumo un hijo?
- Represente gráficamente las frecuencias absolutas y las acumuladas.

**Ejercicio n° 3:** Se tomaron 48 mediciones de peso (en mg) de una especie de polilla forestal en su tercer estadio larval. Los datos se muestran a continuación:

2	6	10	16	20	26	27	29	8	10	13	10
24	24	20	20	21	17	16	16	5	31	29	6
19	18	3	2	12	12	10	10	9	11	14	14
10	14	14	13	12	5	9	5	30	28	7	5



- a) Identifique y clasifique la variable.
- b) Construya una tabla de distribución de frecuencias con 6 intervalos cerrados a la izquierda de manera tal que el límite inferior del primero de ellos sea 2.
- c) Complete las columnas de frecuencias acumuladas, relativas y relativas acumuladas.
- d) Construya en histograma de frecuencias absolutas y el de frecuencias acumuladas.
- e) Represente en los gráficos el polígono de frecuencias y la ojiva.
- f) Responda las siguientes cuestiones:

- i. ¿Cómo calcularía el rango si solamente tiene como información la tabla anterior?
- ii. ¿Cómo calcularía la amplitud si solo dispone de la tabla?
- iii. ¿Cuántas larvas pesan entre 17 y 22 mg?
- iv. ¿Qué porcentaje pesa entre 2 y 7 mg?
- v. ¿Cuántas pesan menos de 27 mg?
- vi. ¿Qué porcentaje pesa menos de 12 mg.?
- vii. ¿Cuántos pesan 22 mg o más?

g) Complete:

- i. Sobre un total de \_\_\_\_\_ larvas solo \_\_\_\_\_ pesan menos de 7 mg.
- ii. \_\_\_\_\_ observaciones registraron un peso entre 22 y 27 mg, lo que representa un \_\_\_\_\_% del total.
- iii. Un \_\_\_\_\_ % del total de larvas pesaron 12 mg o más, mientras que 3 de ellas pesaron entre \_\_\_\_\_ y \_\_\_\_\_ mg.
- iv. El intervalo que presentó la mayor frecuencia fue el que va desde los \_\_\_\_\_ mg hasta los \_\_\_\_\_ mg, cuya frecuencia absoluta es \_\_\_\_\_ y representa el \_\_\_\_\_ del total.

**Ejercicio n° 4:** Utilizando los mismos datos del ejercicio n° 3 construya ahora una tabla con 7 intervalos de clase de tal manera que el límite inferior del primero sea 0, realice lo indicado en los ítems c) y d) de dicho ejercicio y conteste las siguientes preguntas:

- a) ¿Cuántas larvas pesaron entre 10 y 15 mg?
- b) ¿Cuántas pesaron entre 20 y 25 mg?
- c) ¿Qué porcentaje pesó entre 0 y 5 mg?
- d) ¿Cuántas pesaron menos de 20 mg?
- e) ¿Qué porcentaje pesó menos de 10 mg?
- f) ¿Qué porcentaje pesó 25 mg o más?
- g) ¿Cuántos pesaron 10 mg o más?

**Ejercicio n° 5:** Se desea comparar el rendimiento alcanzado en un curso de capacitación por los operarios de los turnos mañana y tarde en una cierta fábrica. Para ello se dispone de los puntajes (de 0 a 100) obtenidos en la evaluación final.



Turno mañana						
63	55	68	83	71	56	65
54	47	55	75	81	69	73
68	69	73	34	76	69	61
59	54	66	46	43	64	72
79	82	92	69	57	73	75
72	49	48	67	71	65	64
36	59	89	59	66	49	62
55						

Turno tarde						
76	79	54	91	72	48	47
53	79	67	56	62	57	49
54	48	58	59	59	56	47
51	52	69	62	69	52	71
38	59	66	60	71	71	73
64	69	60	67	57	95	53
81	79	64	62	83	64	59
45						

- Elabore la tabla de distribución de frecuencias correspondiente a los resultados de cada turno. Utilice intervalos de longitud 10, comenzando desde cero.
- Dibuje los histogramas y los polígonos de frecuencias.
- Calcule el porcentaje de operarios del turno mañana que obtuvieron entre 50 y 80 puntos.
- Compare las dos distribuciones de frecuencias. Explique en qué se parecen, y en qué se diferencian.
- Formule una opinión sobre el rendimiento evidenciado en el curso por los dos turnos de trabajo.

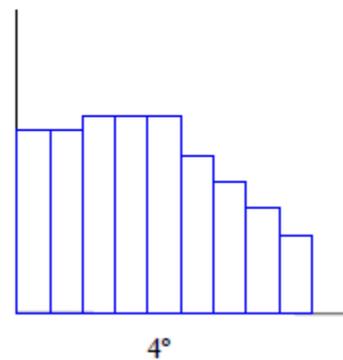
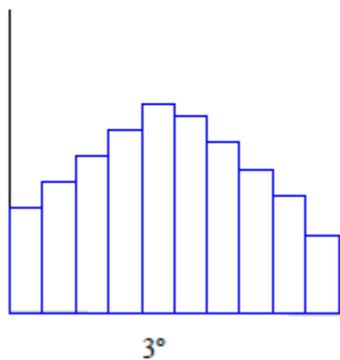
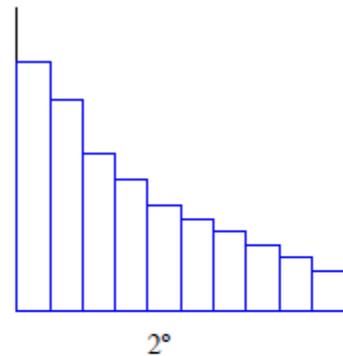
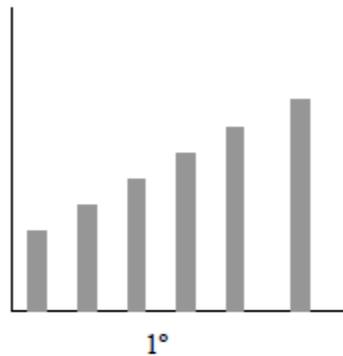
**Ejercicio nº 6 :** Coloque Verdadero (V) o Falso (F) según corresponda.

- La suma de todas las frecuencias absolutas es igual a la cantidad de observaciones analizadas.
- La frecuencia relativa porcentual acumulada está indicando el porcentaje de observaciones correspondiente a cierto intervalo.
- La frecuencia absoluta acumulada en el último intervalo es igual al total de observaciones.
- La suma de frecuencias relativas es igual a 100.
- La amplitud de un intervalo es igual a la diferencia entre su límite superior e inferior.
- Si sumamos todas las frecuencias relativas acumuladas el total es 100%.
- El punto medio de cada intervalo es la suma de su límite inferior y superior.
- Los valores de frecuencia acumulada deben ser siempre números enteros.


**Ejercicio nº 7:** A partir de la observación de los siguientes gráficos, ¿Cuál de ellos asocia con cada una de las siguientes descripciones?



- a) Distribución de la población argentina en 2012 según la edad (en años). El rango es de 0 a 90, el tamaño de la clase o amplitud del intervalo es 10.
- b) Distribución del número de plantas muertas con relación a la severidad de una enfermedad. La severidad se mide de acuerdo a una escala categórica de 0 a 5 en orden creciente de ataque.
- c) Distribución de altura de plantas (en cm) en un cultivo de trigo. Rango de 0 a 50, tamaño de la clase 5.
- d) Distribución de personas según la distancia (en km) que transitan desde su hogar al trabajo.





---

---

### TRABAJO PRÁCTICO N° 3 MEDIDAS DE RESUMEN (POSICIÓN, DISPERSIÓN Y FORMA)

**Ejercicio n° 1:** Calcule media, mediana, modo, varianza, desvío estándar y coeficiente para la siguiente serie que corresponde a la variable longitud recorrida por cosechadoras en metros.

12 28 11 5 24 21 14 16 29 25 19 20 16 10 16

**Ejercicio n° 2:** Los siguientes valores son los años de estudios de diez peones de una estancia.

2 3 5 8 8 7 2 3 4 6

a) Calcule e interprete las medidas de tendencia central y las medidas de dispersión.

**Ejercicio n° 3:** Las siguientes datos son las inasistencias de un grupo de alumnos a una clase:

4 6 0 1 2 3 4 5 6 5 1 2 3 4 5 3 3 2 3 4 3 2 5 6 6 5

- Clasifique la variable de estudio.
- Ordene los datos en una tabla de distribución de frecuencias.
- Calcule media, mediana y modo e interprete.
- Calcule e interprete los cuartiles.
- Calcule e interprete el decil 8.
- Calcule e interprete el percentil 15.
- Calcule el valor de la variable que deja por debajo al 45% de las observaciones.
- Calcule el valor de la variable que es superado por el 35% de los datos
- Calcule la varianza, el desvío estándar y CV.

**Ejercicio n° 4:** Considere la tabla construida en el ejercicio n° 1 del Trabajo Práctico n°2 y realice lo siguiente:

- Calcule las medidas de tendencia central e interprételas.
- Calcule e interprete los cuartiles.
- Calcule e interprete el decil 8.
- Calcule e interprete el percentil 15.
- Calcule el valor de la variable que deja por debajo al 45% de las observaciones.
- Calcule el valor de la variable que es superado por el 35% de los datos.
- Calcule la varianza, el desvío estándar y CV.
- Construya un gráfico de cajas para esta distribución.



**Ejercicio n° 5:** La siguiente tabla corresponde a los ingresos mensuales de 80 familias.

Clase (\$)	Cant. Familias
0 – 600	4
600 – 1200	32
1200 – 1800	21
1800 – 2400	14
2400 – 3000	8
3000 – 3600	1

- Calcule e interprete las medidas de tendencia central.
- Calcule e interprete el primer cuartil.
- Calcule e interprete el percentil 84.
- Calcule el valor que deja por debajo al 30% de los datos.
- Calcule el valor que deja por encima al 10% de los datos.
- ¿Qué ingreso posee como máximo el 15% de las familias?
- Calcule el modo y la mediana gráficamente.
- Calcule las medidas de dispersión.
- Construya el histograma y el gráfico de cajas para esta distribución.

**Ejercicio n° 6:** Considere la tabla construida en el ejercicio n° 3 del Trabajo Práctico n° 2 y realice lo siguiente.

- Calcule e interprete las medidas de tendencia central.
- Calcule e interprete el primer cuartil.
- Calcule e interprete el percentil 84.
- Calcule el valor que deja por debajo al 30% de los datos.
- Calcule el valor que deja por encima al 10% de los datos.
- ¿Qué ingreso posee como máximo el 15% de las familias?
- Calcule el modo y la mediana gráficamente.
- Calcule las medidas de dispersión.
- Construya el histograma y gráfico de cajas para esta distribución.

**Ejercicio n° 7:** La siguiente tabla muestra las frecuencias correspondientes a las razas bovinas disponibles en un establo para su comercialización.

RAZA	FRECUENCIA
Shorthorn	18
Hereford	10
Aberdeen Angus	10
Pardo	62

- ¿Cuál es el modo de la distribución?



**Ejercicio n° 8:** Se está estudiando dos tratamientos A y B para la ganancia de peso (en kg) en animales destinados a la producción de carne. Los registros de estos incrementos en ambos grupos son:

A	1	5	11	3	5	1	2
B	7	1	16	2	1	18	23

- ¿Qué grupo posee, en promedio, una mayor ganancia de peso?
- Calcule mediana, modo, varianza, desvío estándar y coeficiente de variación.
- ¿Qué distribución es más homogénea?
- Suponga que cada animal ha aumentado 3kg más, calcule nuevamente la media, varianza y coeficiente de variación aplicando propiedades.
- Suponga que luego de un tiempo cada animal ha duplicado su incremento de peso actual, calcule nuevamente la media, varianza y coeficiente de variación aplicando propiedades.

**Ejercicio n° 9:** Coloque V o F según corresponda

a) La media no puede tomar valores negativos.	
b) Una distribución puede no tener modo.	
c) La media es la única medida que necesita todos los valores para calcularla.	
d) Ante la existencia de datos atípicos se recomienda el uso del modo.	
e) El modo no puede calcularse para variables cualitativas.	

**Ejercicio n° 10:** La siguiente muestra a 37 personas agrupadas según su edad en años.

Clase (pts.)	fi
0 – 5	7
5 – 10	8
10 – 15	9
15 – 20	6
20 – 25	4
25 – 30	3

- Represente el histograma y el polígono de frecuencias correspondiente a la distribución.
- Calcule el coeficiente de asimetría (sesgo) y de curtosis.

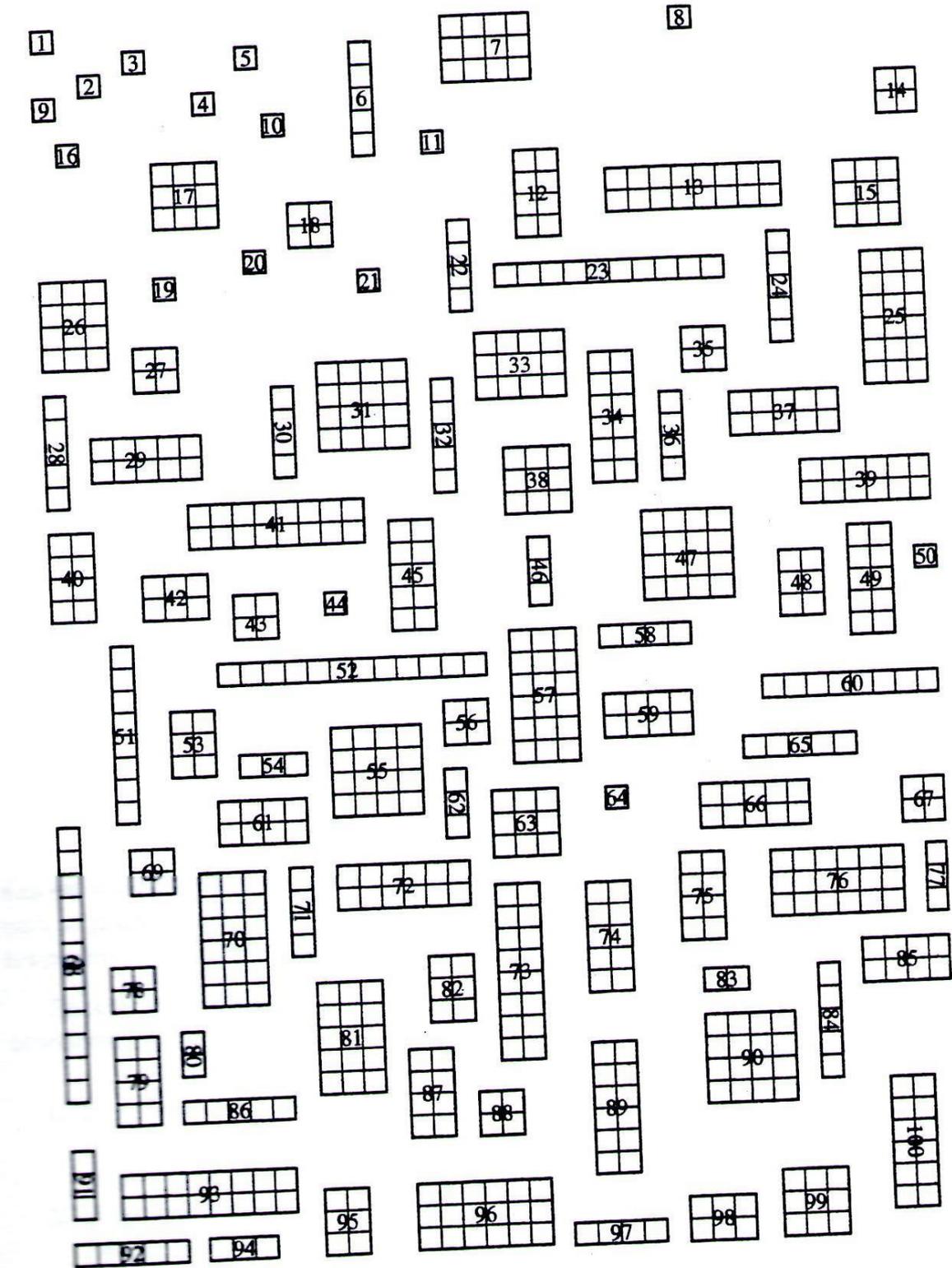
$$\text{Coef. de sesgo } SK = \frac{3.(\bar{X} - \tilde{X})}{D.S.} \quad \text{Coef. curtosis: } K = \frac{P_{90} - P_{10}}{1,9.(P_{75} - P_{25})}$$

**Ejercicio n° 11:** El objetivo del siguiente ejercicio consiste en elegir una muestra de cinco rectángulos y, a partir de ellos, estimar el área promedio de los 100 rectángulos de la ilustración.

- Sin estudiar con detalle la distribución de los rectángulos, elija rápidamente a cinco rectángulos que cree que representen a la población. Ésta es su muestra.
- Encuentre el área de cada rectángulo (expresada en celdas) en su muestra de cinco y calcule la media muestral, es decir, el área promedio de los cinco rectángulos.



- 
- c) Recopile las medias muestrales obtenidas por otros estudiantes de la clase. Genere un gráfico de medias.
  - d) Describa la forma, punto medio y dispersión de ese gráfico obtenida con muestras no probabilísticas.
  - e) Ahora elija aleatoriamente cinco rectángulos, es decir, generando números aleatorios con la calculadora, con una planilla de cálculo (ej. Excel) o la tabla de números aleatorios. Ésta es su muestra.
  - f) Repita los pasos b), c) y d) esta vez utilizando la muestra aleatoria.
  - g) Describa en qué se parecen y en qué se diferencian las dos distribuciones de medias muestrales.





**Ejercicio nº 12:** Coloque V o F según corresponda.

La varianza es la mejor medida de dispersión para realizar comparaciones entre distribuciones de distintas variables.	
A mayor coeficiente de variación, mayor homogeneidad.	
El desvío estándar utiliza la unidad de la variable.	
La varianza es siempre mayor o igual que cero.	
El C.V. es empleado para comparar distribuciones que poseen medias muy distintas entre sí.	
Una distribución perfectamente homogénea tiene C.V. = 100%	
Una de las medidas de dispersión menos utilizada es el rango	
En una distribución de frecuencias tiene asimetría negativa la relación entre moda y media es $\hat{X} > \bar{X}$	
El coeficiente de variación es una medida de orden.	



---

---

## TRABAJO PRÁCTICO N° 4 ANÁLISIS DE REGRESIÓN Y CORRELACIÓN

**Ejercicio n° 1:** Supongamos que un agrónomo posee la siguiente información referente a 6 años agrícolas:

$x_i$	$y_i$
15	6,5
17	7
20	9
18	8,5
21	9
20	8,2

Donde  $x$  = temperatura media en el mes de siembra del trigo e  $y$  = rendimiento del trigo.

- a) Grafique los datos utilizando un diagrama de dispersión.
- b) Estime la recta de regresión poblacional.
- c) Interprete  $a$  y  $b$ .

**Ejercicio n° 2:** Las materias primas que utiliza una fábrica en la producción de una fibra sintética son almacenadas en un lugar donde no se tiene control sobre la humedad. Para analizar si la humedad del almacenamiento influye sobre la humedad de la materia prima almacenada, se tomaron muestras de la materia prima durante 12 días. Cada día se midió la humedad de la muestra y del lugar de almacenamiento. Los resultados obtenidos fueron los siguientes:

Humedad de lugar ( $x_i$ )	Humedad de la muestra ( $y_i$ )
42	12
35	8
50	14
43	9
48	11
62	16
31	7
36	9
44	12
39	10
55	13
48	11

- a) Grafique los datos en un diagrama de dispersión.
- b) Estime los parámetros que caracterizan a la recta de regresión poblacional.



c) Interprete  $a$  y  $b$ .

**Ejercicio n° 3:** Se desea saber si las notas en matemática de una población de alumnos podrían explicarse a través de su rendimiento en Lengua. Para ello se tomó una muestra de 10 alumnos y se obtuvieron sus calificaciones en Matemática y en Lengua, las cuales figuran en la siguiente tabla.

Matemática	2	4	5	5	6	6	7	7	8	9
Lengua	2	2	5	6	5	7	5	8	7	10

- a) Confeccione el diagrama de dispersión.
- b) Estudie la tendencia de los datos y ajuste a través de la recta de los mínimos cuadrados.
- c) Interprete los parámetros del modelo construido.
- d) Determine mediante un coeficiente adecuado el grado y sentido de la relación lineal entre ambas variables.
- e) Estime el puntaje en lengua de un alumno que en matemática obtuvo un 3.
- f) Calcule el interprete el coeficiente de determinación.
- g) Calcule el error estándar e interprete.

**Ejercicio n° 4:** Un agricultor, que viene utilizando una determinada marca de fertilizante con el objeto de incrementar su producción, desea conocer cómo el mismo influye en el rendimiento del cereal que siembra. A tal fin aplicó distintas cantidades (en Kg.) del fertilizante en 8 parcelas y computó el rendimiento (en TN) luego de la cosecha:

Fertilizante (en Kg.)	Producción de trigo (en TN)
2	8
4	9
5	11
7	11
10	12
11	14
12	15
15	16

- a) Construya el diagrama de dispersión para los datos de la muestra e interprete.
- b) Identifique la variable respuesta (dependiente) y la variable explicativa (independiente), con el objeto de plantear un modelo de regresión adecuado a los datos.
- c) Estime en promedio el rendimiento del trigo cuando la cantidad de fertilizante aplicado sea 13 kg.
- d) Calcule  $r$  y  $r^2$  e interprete ambos valores.

**Ejercicio n° 5:** Para cada uno de los siguientes pares de variables, decida qué tipo de correlación lineal poseen.



- a) Demanda de un producto Vs. Precio del producto.
- b) Calificación obtenida Vs. Horas de estudio.
- c) Peso Vs. Talla.
- d) Cantidad de errores cometidos en una prueba Vs. Cantidad de veces que se rindió.
- e) Cantidad de calorías consumidas Vs. Nivel de colesterol en sangre.
- f) Tiempo que demora en vaciarse un estadio Vs. Cantidad de espectadores que asistieron
- g) Resistencia que presenta un conducto ante el paso de un fluido Vs. Diámetro del conducto
- h) Solubilidad Vs. Temperatura del agua.

**Ejercicio n° 6:** Con el objetivo de determinar la correlación existente entre la circunferencia escrotal y algunos parámetros de calidad seminal en caprinos Criollos serranos de la provincia de Tucumán (Argentina), se desarrolló una investigación en la cual se midieron las siguientes variables:

Circunferencia escrotal (CE) en cm.

Volumen eyaculado (VE) en ml.

Concentración espermática (CEsp) en millones/ml

Espermatozoides totales (ET) millones por eyaculado.

Se trabajó con un grupo de 11 machos, de un año de edad al inicio de la experiencia. Las mediciones se realizaron cada 14 días, durante nueve meses y los resultados se consignan a continuación:

N°	CE	VE	CEsp	ET
1	24	0,59	6294	3606
2	40	0,58	6200	4200
3	41	0,57	6300	3950
4	41	0,62	6850	2930
5	29	0,68	6340	2700
6	23	0,68	5980	2850
7	21	0,57	5990	3200
8	19	0,49	5700	3780
9	18	0,48	6000	3600
10	19	0,63	6690	3990
11	23	0,65	6700	3980

a) Realice un análisis de correlación y estudie la importancia de considerar la circunferencia escrotal en la calidad reproductiva. Enuncie conclusiones.

**Ejercicio n° 7:** Los siguientes datos corresponden a la velocidad (y), la potencia (x1) y la cantidad de cilindros (x2) de 10 marcas de motos:



Velocidad (y)	Potencia (x <sub>1</sub> )	Cantidad de cilindros (x <sub>2</sub> )
160	60	2
156	26	4
193	80	6
191	67	2
200	70	2
190	70	2
194	70	2
170	36	1
132	27	1
111	17	1

La siguiente salida corresponde al análisis de regresión múltiple efectuado sobre esta base de datos. A partir de la misma se pide:

- Escriba el modelo de regresión correspondiente.
- Interprete sus parámetros.
- Enuncie las hipótesis que se plantearon y las conclusiones correspondientes.

### Análisis de regresión lineal

Variable	N	R <sup>2</sup>	R <sup>2</sup> Aj	ECMP	AIC	BIC
velocidad	10	0,80	0,75	678,72	87,05	88,26

### Coefficientes de regresión y estadísticos asociados

Coef	Est.	E.E.	LI(95%)	LS(95%)	T	p-valor	CpMallows
const	108,87	12,64	78,97	138,76	8,61	0,0001	
potencia	1,15	0,24	0,59	1,71	4,83	0,0019	22,54
cilindros	0,29	3,52	-8,03	8,61	0,08	0,9357	2,13

### Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	6478,49	2	3239,24	14,28	0,0034
potencia	5292,44	1	5292,44	23,34	0,0019
cilindros	1,59	1	1,59	0,01	0,9357
Error	1587,61	7	226,80		
Total	8066,10	9			

**Ejercicio n°8:** Los siguientes datos corresponden a la profundidad y los perfiles de suelo indicados como láminas de agua en el mismo. Se trata de un estudio



para tratar de modelar la disponibilidad de agua en función a la profundidad luego de una emergencia hídrica.

Profundidad (cm)	Agua (mm)
10	29,00
10	29,70
10	31,00
20	23,27
20	23,18
20	28,44
30	21,74
30	25,75
30	26,42
40	20,96
40	21,66
40	21,70
50	17,83
50	17,03
50	17,39
60	14,55
60	13,31
60	12,03

Analice la siguiente salida interpretando los resultados y efectuando conclusiones

### Análisis de regresión lineal

Variable	N	R <sup>2</sup>	R <sup>2</sup> Aj	ECMP	AIC	BIC
Agua (mm)	18	0,90	0,90	4,18	77,04	79,71

### Coefficientes de regresión y estadísticos asociados

Coef	Est.	E.E.	LI(95%)	LS(95%)	T	p-valor	CpMallows
const	32,83	0,99	30,72	34,93	33,08	<0,0001	
Profundidad (cm)	-0,31	0,03	-0,37	-0,26	-12,20	<0,0001	141,25

### Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	507,84	1	507,84	148,95	<0,0001
Profundidad (cm)	507,84	1	507,84	148,95	<0,0001
Error	54,55	16	3,41		
Total	562,39	17			

## CORRELACIÓN

**Ejercicio n° 9:** Los datos de la tabla se refieren al contenido de proteína



bruta (PB) y caseína (CA) en leche en una muestra de 23 tambos de la cuenca lechera del centro del país.

PB	2.74	3.19	2.96	2.91	3.23	3.04	3.08	3.23	3.11	3.11	3.1	3.25
CA	1.87	2.26	2.07	2.09	2.28	2.04	2.18	2.3	2.17	2.15	2.16	2.33
PB	2.95	3.08	3.14	3.22	3.14	3.15	3.2	2.95	3.19	3.12	2.99	
CA	2.04	2.16	2.16	2.22	2.22	2.24	2.22	2.07	2.25	2.23	2.16	

El coeficiente de correlación lineal muestral entre PB y CA es:  $r = 0.9327$ . ¿Es esta alta correlación estadísticamente significativa? Para contestar a esta pregunta se debe realizar un contraste de hipótesis.

**Ejercicio n° 10:** Los siguientes datos se refieren a la cantidad de turistas que ingresaron a la Argentina y la paridad del peso con el dólar en un período de 8 años.

Cantidad de turistas	Paridad
1599	1,16
1817	1,17
2070	1,20
1936	1,23
2160	1,23
2416	1,30
2694	1,37
3242	1,39

Estime el coeficiente de correlación lineal entre las variables del problema y pruebe su significación.

### TABLAS DE CONTINGENCIA

**Ejercicio n°11:** La siguiente tabla de contingencia muestra los resultados de una prueba que consistió en colocar un grupo de semillas al efecto de un fungicida y otro al control, luego se contabilizó la cantidad de semillas que germinaron y las que no.

Condición	No germinó	Germinó	Total
Control	245	1190	1435
Fungicida	123	1358	1481
Total			

Pruebe que la mayor germinación observada en las semillas tratadas es evidencia estadísticamente significativa de que el uso de fungicida mejora el poder germinativo.



**Ejercicio n° 12:** Evalúe la relación entre CC y éxito de la inseminación en 160 vacas según la siguiente tabla.

CC	Preñadas	No preñadas	Total
<b>2</b>	23	7	30
<b>3</b>	76	4	80
<b>4</b>	46	4	50
<b>Total</b>			



---

---

## TRABAJO PRÁCTICO N° 5 PROBABILIDAD Y VARIABLE ALEATORIA

**Ejercicio n° 1:** En cada una de las siguientes situaciones se pide:

- a) Determinar el espacio muestral.
- b) Realizar la asignación de probabilidad a cada evento elemental.
- c) Calcular  $P(A)$ .

1. Se extrae al azar una ficha de una caja con 20 fichas numeradas del 1 al 20 y se considera el siguiente evento:

*A: "El número extraído es par"*

2. Se extrae un paquete de un conjunto de cinco, de los cuales dos están defectuosos y se considera el evento:

*A: "El paquete extraído no es defectuoso"*

3. Se lanzan simultáneamente tres monedas equilibradas y se considera el evento:

*A: "Obtener exactamente dos caras"*

4. Se lanzan simultáneamente tres monedas equilibradas y se considera el evento:

*A: "Obtener al menos dos caras"*

5. Una mujer tiene dos hijos e interesa observar el sexo de los mismos.

*A: "Tiene un solo varón"*

**Ejercicio n° 2:** Se extrae una bolilla de una urna que contiene 6 bolillas rojas, 4 blancas y 5 azules. Determine la probabilidad de que la bolilla extraída:

- a) Sea roja
- b) Sea blanca
- c) Sea azul
- d) No sea roja
- e) Sea roja o blanca.
- f) Sea blanca o azul.
- g) Sea azul y roja.

**Ejercicio n° 3:** Considere la situación del Ejercicio n° 2 pero esta vez se extraen dos bolillas sin reposición. Determine la probabilidad de:

- a) Sacar dos rojas.
- b) Sacar roja la primera y blanca la segunda.



- c) Sacar una sola roja.
- d) Sacar al menos una azul.
- e) Sacar como máximo una blanca.
- f) Sacar ninguna blanca.

**Ejercicio nº 4:** Realice nuevamente los cálculos propuestos en el ejercicio nº 3 pero considerando esta vez el experimento con reposición.

**Ejercicio nº 5:** Veintisiete personas fueron clasificadas según sexo y contextura física resultando la siguiente tabla de contingencia:

	Delgada	Media	Robusta	TOTAL
Mujeres	1	5	3	
Varones	6	3	9	
TOTAL				

Si elegimos una persona al azar, calcule la probabilidad de que:

- a) Sea varón.
- b) Sea mujer y de contextura robusta.
- c) Sea de contextura robusta sabiendo que es varón.
- d) Sea varón sabiendo que es de contextura robusta.
- e) Sea de contextura delgada o sea mujer.
- f) Sea mujer siendo delgada.
- g) Sea de contextura delgada o robusta.
- h) Sea varón o mujer.
- i) ¿Los eventos Mujer y Robusta son independientes?. Justifique.

**Ejercicio nº 6:** La siguiente tabla muestra el conteo de vacas preñadas y no preñadas inseminadas artificialmente y clasificadas según su condición corporal.

CC	Preñada	Conteo
2	SI	23
3	SI	76
4	SI	46
2	NO	7
3	NO	4
4	NO	4

- a) Represente estos datos en una tabla de contingencia.
- b) Calcule la probabilidad de que una vaca elegida al azar esté preñada y tenga condición corporal igual a 3.
- c) Indique si los eventos preñada y CC=2 son independientes.

**Ejercicio nº 7:** Sean A y B dos eventos que verifican:  $P(A)=0,32$  ;  $P(B)=0,85$  y  $P(B/A)=0,8$ , calcule  $P(A/B)$ .



**Ejercicio nº 8:** Una urna contiene 10 bolillas numeradas de 1 a 10. Las bolillas 1 a 4 son rojas, y las 5 a 10 son azules. Se extrae una bolilla y consideramos los siguientes eventos:

- A: " la bolilla tiene el número 3"  
B: "la bolilla es roja"

Calcule  $P(A/B)$  y  $P(B/A)$

**Ejercicio nº 9:** Sea el experimento: "arrojar tres monedas equilibradas" y la variable aleatoria: "número de caras obtenidas".

- Halle su función de probabilidad y su función de distribución.
- Grafique ambas funciones.
- Calcule  $P(0 \leq X \leq 2)$ .
- Calcule  $P(1 \leq X \leq 2)$ .
- Calcule  $E(x)$

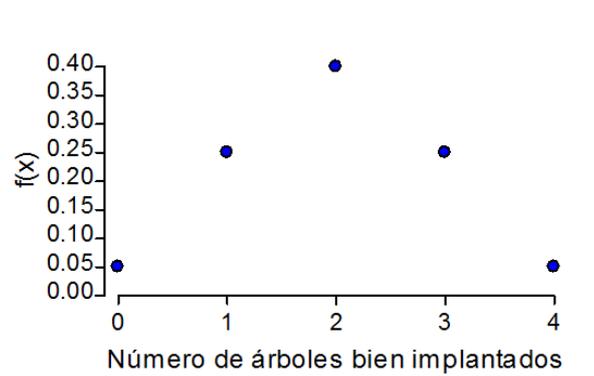
**Ejercicio nº 10:** Sea  $X$  una variable aleatoria que representa el número de animales vacunados la última semana contra una cierta enfermedad y conociendo que:

$X$	1	2	3	4	5	6	7
$p(X)$	0.05	0.08	0.17	0.35	0.22	0.10	0.03

se pide:

- Compruebe que se trata de una función de probabilidad.
- Grafique  $p(X)$ .
- Grafique  $F(X)$ .
- Calcule  $P(X \leq 2)$
- Calcule  $P(X > 5)$
- Calcule  $E(x)$

**Ejercicio nº 11:** La función de probabilidad de la variable aleatoria definida como el "número de árboles bien implantados", tiene la siguiente forma:





- a) Construya la función de distribución acumulada correspondiente.
- b) ¿Qué significa  $F(2)$ ?
- c) Calcular la media y la varianza poblacionales.

**Ejercicio n° 12:** Supongamos que se toma una muestra aleatoria con reposición de tamaño  $n=2$  a partir del conjunto  $\{1,2,3\}$ . Supongamos además que definimos la variable aleatoria  $Y=$  "suma de los dos números" y que estamos interesados en los siguientes eventos:

- A = "la suma es par"  
B = "la suma es impar"  
C = "la suma es 5"

Responda:

- d) ¿Cuál es la  $P(A)$ ?
- e) ¿Cuál es la  $P(B)$ ?
- f) ¿Cuál es la  $P(C)$ ?
- g) Los eventos A y B ¿son mutuamente excluyentes?
- h) ¿Cuál es la probabilidad de que ocurra A o B?
- i) ¿Cuál es la probabilidad de que ocurra B o C?
- j) Represente en una tabla a  $f(Y)$  y  $F(Y)$

**Ejercicio n° 13:** Se conoce que el cuantil 0.10 de la distribución de la variable  $X=$ "longitud de raíces" de plántulas de tomate al momento del trasplante es 3 cm, y se sabe que sólo las plántulas con raíces mayores de 3 cm tienen probabilidad de sobrevivir al trasplante:

- a) ¿Cuántas plántulas se deberían adquirir para lograr un lote de 2000 plántulas implantadas?

**Ejercicio n° 14:** Dada la siguiente función

$$f(x) = \begin{cases} 3x^2 & \text{si } 0 < x \leq 1 \\ 0 & \text{en el resto} \end{cases}$$

- a) Determine si se trata de una función de densidad
- b) Grafique
- c) Calcule  $P(x < 0,25)$
- d) Calcule  $P(x > 0,1)$



---

## TRABAJO PRÁCTICO N° 6 MODELOS DE DISTRIBUCIONES DE PROBABILIDAD

**Ejercicio n° 1:** Sea  $X$  una variable aleatoria discreta que tiene distribución binomial de parámetros  $n=5$  y  $p=0.25$ . ( $X \sim Bi(5; 0.25)$ ) Calcule las siguientes probabilidades:

- a)  $P(X = 2)$
- b)  $P(X < 4)$
- c)  $P(X > 3)$
- d)  $P(X \geq 3)$

**Ejercicio n° 2:** El 10% de las piezas producidas por una fábrica es defectuosa. Si de la producción total se toma un lote de 10 piezas seleccionadas al azar, responda las siguientes cuestiones:

- a) ¿Cuál es la variable de interés?
- b) Indique los eventos que se consideran, y cuál de ellos es "éxito" y cual es "fracaso".
- c) Indique cuales son los parámetros que definen esta distribución.
- d) Calcule la probabilidad de que todas las piezas estén sanas.
- e) Calcule la probabilidad de que haya más de dos defectuosas

**Ejercicio n° 3:** Se sabe que la probabilidad de encontrar una plaga cuarentenaria en una planta ornamental de importación es  $p=0,05$ . Si se examinan 25 plantas,

- a) ¿Cuál es la probabilidad de encontrar al menos una con la presencia de la plaga?
- b) ¿Cuál es la probabilidad de encontrar exactamente dos plantas de las 25 con la plaga?
- c) ¿Cuál es la probabilidad de detectar al menos una planta con la plaga si la probabilidad de éxito cambia a  $p=0,1$ ?

**Ejercicio n° 4:** Un técnico en semillas desea inspeccionar el funcionamiento de 20 cámaras de cría. Para esto toma dos cámaras al azar y registra la temperatura de las mismas. Si estas dos cámaras funcionan correctamente, el grupo de 20 será aceptado. ¿Cuáles son las probabilidades que tal grupo de 20 cámaras sea aceptado si contiene:

- a) 4 cámaras con registros de temperaturas no adecuadas?
- b) 8 cámaras con registros de temperatura no adecuadas?
- c) 12 cámaras con registros de temperaturas no adecuadas?

**Ejercicio n° 5:** Se extraen 4 bolillas con reemplazo de una urna que contiene 5 blancas y 3 negras.

- a) ¿Cuál es la probabilidad de que salgan al menos 2 blancas?
- b) ¿Cuál es la probabilidad de que no salga ninguna negra?
- c) ¿Cuál es la probabilidad de que salga a lo sumo dos blancas?

**Ejercicio nº 6:** Sea  $X$  una variable aleatoria con distribución de Poisson de parámetro  $\lambda = 3$  ( $X \sim Po(3)$ ). Calcule las siguientes probabilidades:

- a)  $P(X = 3)$                       b)  $P(X < 4)$                       c)  $P(X > 2)$                       d)  $P(X \geq 2)$

**Ejercicio nº 7:** A una central telefónica llegan 3 llamadas por minuto y los números de llamadas de minuto en minuto son independientes.

- a) ¿Cuál es la variable de interés?  
b) ¿Cuál es el parámetro que define a esta distribución?  
c) Calcule la probabilidad de que en un minuto no lleguen llamadas.  
d) Calcule la probabilidad de que en un minuto llegue a lo sumo una llamada.  
e) Calcule la probabilidad de que lleguen dos llamadas en dos minutos.  
f) Calcule la probabilidad de que en 30 segundos no lleguen llamadas.  
g) Calcule la probabilidad de recibir al menos dos llamadas en un minuto.

**Ejercicio nº 8:** Una máquina sufre, en promedio, tres desperfectos por mes que obligan a su detención para realizar reparaciones.

- a) Especifique cuál es la variable de estudio y cuál es el modelo probabilístico que permite representarla.  
b) Calcule la probabilidad de que durante un mes cualquiera no sufra ninguna detención.  
c) Calcule la probabilidad de que durante un mes cualquiera sufra al menos una detención.  
d) Calcule la probabilidad de que en dos meses sufra cinco detenciones.

**Ejercicio nº 9:** En una planta de embalaje de manzanas las frutas pasan por una cinta transportadora para su revisión. Las manzanas con manchas, picaduras y magulladuras, etc, son derivadas para la producción de jugos. En un establecimiento de embalaje se retiran de la cinta en promedio 15 manzanas por minutos. Cada empleado puede retirar un máximo de 20 manzanas por minuto.

- a) ¿Cuál es la probabilidad de que todas las manzanas retiradas por un empleado sean defectuosas?  
b) ¿Cuántos empleados se necesitan en una cinta si un lote de manzanas trae el doble de frutas defectuosas de lo normal y no se quiere que pase ninguna fruta defectuosa al embalaje definitivo en el 95% de las veces?

**Ejercicio nº 10:** Sea  $Z$  la variable aleatoria normal estandarizada de parámetros 0 y 1.

( $Z \sim N(0 ; 1)$ ) Calcule las siguientes probabilidades:

- a)  $P(-1 < z < 1.5)$                       e)  $P(z < -0,25)$   
b)  $P(-2 < z < -0.5)$                       f)  $P(0,28 < z < 1,8)$   
c)  $P(z < 1.05)$                       g)  $P(z < -1)$   
d)  $P(z > -0.25)$                       h)  $P(z > 1)$



i)  $P(z > 0,54)$

j)  $P(0 < z < \infty)$

**Ejercicio n° 11:** Halle el valor de  $Z_0$  que cumpla la condición pedida en cada caso.

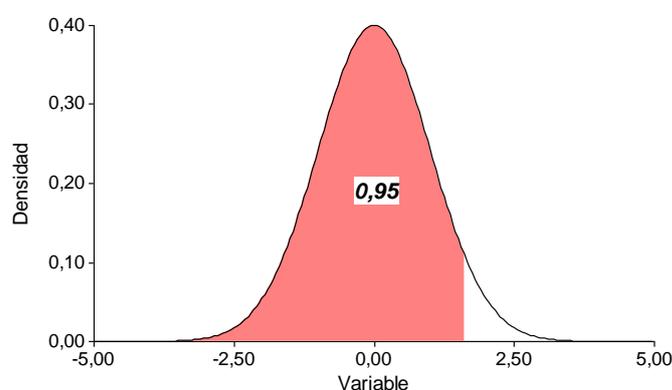
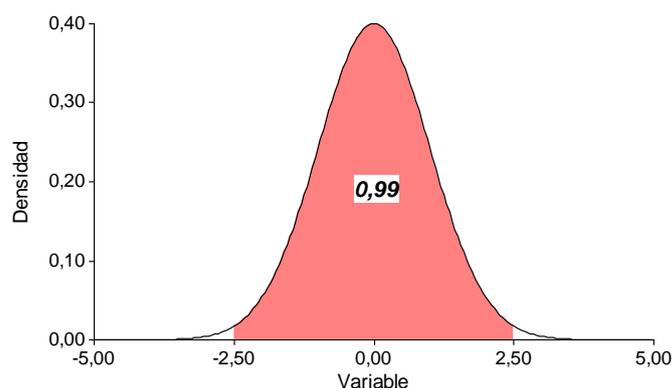
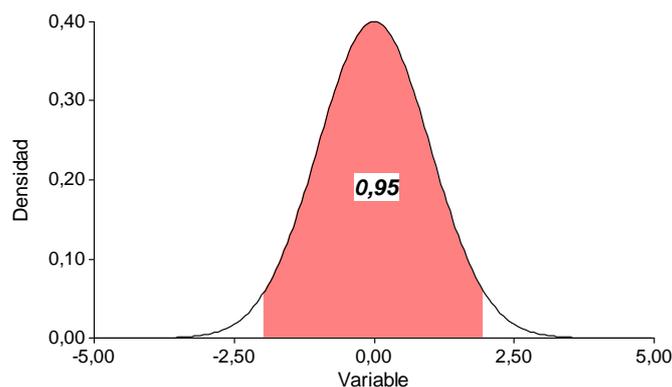
a)  $P(Z < Z_0) = 0.8023$

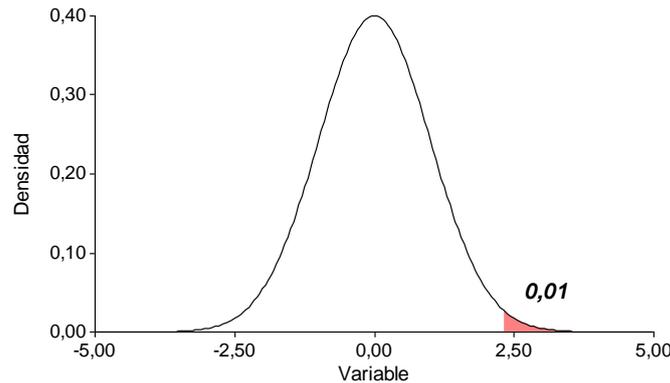
c)  $P(Z < Z_0) = 0.3263$

b)  $P(Z > Z_0) = 0.1002$

d)  $P(Z > Z_0) = 0.9177$

**Ejercicio n° 12:** Calcule en cada caso el o los valores de  $z$  que cumplan con la condición indicada en el gráfico sabiendo que  $z \sim N(0, 1)$ .





**Ejercicio nº 13:** Los salarios de un grupo de obreros siguen una distribución normal con media igual a \$380 y desviación estándar igual a \$70. Calcule cuántos obreros ganan:

- a) Menos de \$296.
- b) Más de \$320.
- c) Entre \$240 y \$420.
- d) Entre \$390 y 410.
- e) Entre \$350 y \$375.
- f) Exactamente \$320.

**Ejercicio nº 14:** El tiempo que funciona un televisor sin registrar fallas tiene distribución normal con media igual a 4 años y desvío estándar 1 año y medio.

- a) Calcule la probabilidad de que un televisor dure más de 5 años.
- b) Calcule la probabilidad de que un televisor dure entre 4 y 6 años.
- c) ¿Cuál tiene que ser el tiempo de garantía que debe ofrecer el fabricante si quiere que la probabilidad de que el televisor funcione hasta ese tiempo sea de 0.4?

**Ejercicio nº 15:** Por medio de un tamiz de malla de 8 mm de diámetro se zarandean 8000 granos de maíz. El *diámetro del grano* de maíz sigue una distribución normal con esperanza igual a 9 mm y una desviación estándar de 1.2 mm.

- a) ¿Qué proporción de granos serán retenidos por el tamiz?
- b) ¿Qué proporción de granos no retenidos, serán retenidos por un tamiz de diámetro de malla igual a 7.5 mm?
- c) ¿Qué proporción de granos pasará a través de los dos tamices?

**Ejercicio nº 16:** Si la variable espesor de un sedimento en un sustrato de suelo, se distribuye normalmente con media  $\mu = 15$  micrones y desviación estándar  $\sigma = 3$  micrones.

- a) ¿Cuál es el cuantil 0.75 de la distribución de la variable?
- b) ¿Cómo se interpreta este valor?





---

## TRABAJO PRÁCTICO N° 7 DISTRIBUCIÓN DE ESTADÍSTICOS MUESTRALES

**Ejercicio n° 1:** La variable aleatoria  $X$  tiene una distribución normal con media igual a 5 y varianza 9, esto se simboliza así  $X \sim N(5 ; 9)$ . Si se toma una muestra de tamaño 49 ( $n=49$ ) de observaciones correspondiente a esta variable, indique como se distribuye la media muestral, esto es  $\bar{X}$ .

**Ejercicio n° 2:** Supóngase que la variable "peso de 100 semillas" de una variedad de maíz sigue una distribución normal con la esperanza  $\mu=35$  g y varianza  $\sigma^2 = 30$  g<sup>2</sup>.

Si se elige al azar una muestra de 25 paquetes de 100 semillas, de los muchos que se llenan en el proceso de envasado:

- ¿Qué valor de peso promedio se esperaría?
- ¿Qué probabilidad hay de que el peso promedio de la muestra esté entre 36 y 38 gramos?
- ¿Cuál es la probabilidad si la muestra es de tamaño 29?
- ¿Cuál es la probabilidad si la muestra es de tamaño 49 y la varianza es 60 g<sup>2</sup>?

**Ejercicio n° 3:** Si la distribución de la variable aleatoria producción de leche de un establecimiento lácteo (en cientos de litros) se aproxima a una distribución normal con media  $\mu=70$  y desvío estándar  $\sigma=8$ .

- ¿Cuál es la probabilidad de que la media de una muestra de tamaño 5 exceda el valor 75?
- ¿Cuál es la probabilidad de que el peso promedio esté entre 68 y 75g?
- ¿Cuál es la producción promedio sólo superada por un 5% de las producciones promedio?

**Ejercicio n° 4:** Si se especifica que la esperanza de la variable "cantidad de kilómetros recorridos por litro" de un vehículo es  $\mu=12$  y tiene una desviación estándar  $\sigma=2$ .

- Calcule el error estándar de la media de una muestra de 10 recorridos.
- Calcule el error estándar de la media de una muestra de 20 recorridos. Compare con el obtenido en el punto anterior.
- ¿Cuál es la probabilidad de que la media de una muestra de 10 recorridos sea menor o igual que 11 km/lt, si el vehículo funciona de acuerdo a las especificaciones?

**Ejercicio n° 5:** El peso de bebés recién nacidos tiene una distribución aproximadamente normal con media igual a 3,5kg y desviación estándar 0,5kg. Si tomamos una muestra de niños de tamaño  $n$ , calcule la probabilidad de que el peso promedio de los  $n$  bebés esté entre 3 y 4,5kg si:



- a)  $n=25$
- b)  $n=36$
- c)  $n=121$

¿Qué conclusiones puede realizar?

**Ejercicio n° 6:** En un criadero de semillas se está probando una nueva variedad de maíz que saldrá a la venta si en una muestra de 50 parcelas experimentales el desvío estándar de su rendimiento no supera los 23 kg/ha.

- a) ¿Cuál es la probabilidad de que esto ocurra si la verdadera desviación estándar es 20?
- b) ¿Cuál es el valor por debajo del cual está el 99% de los valores posibles de desviaciones estándar muestrales basadas en muestras de tamaño 30, si la verdadera desviación estándar es 20?



## TRABAJO PRÁCTICO N° 8 ESTIMACIÓN DE PARÁMETROS

**Ejercicio n° 1:** Considere la variable rendimiento de maíz, cuya distribución es normal con media  $\mu$  y desviación estándar  $\sigma$ . Para estimar el rendimiento promedio del maíz bajo el efecto de un herbicida, se toma una muestra de tamaño 40 y se obtiene un promedio de 60 qq/ha. Se sabe por experiencias anteriores que la varianza poblacional  $\sigma^2$  es 25 (qq/ha)<sup>2</sup>.

- Construya los intervalos de confianza del 95% y 99% para  $\mu$ .
- ¿Cómo cambian los intervalos anteriores si el tamaño de la muestra fuese 100 y se obtiene el mismo promedio?
- ¿Cómo se modifica el intervalo del 95% calculado en a) si la desviación estándar fuese de 7 qq/ha?

**Ejercicio n° 2:** Una empresa dedicada a la comercialización de semillas desea estimar la altura promedio de un sorgo forrajero que ha desarrollado. Para ello se toma una muestra de 50 plantas y se calcula la media de la altura, la que resulta ser 130 cm. Se sabe por experiencias anteriores que la desviación estándar es 22cm que la altura se distribuye en forma normal.

- Construya los intervalos de confianza para  $\mu$  con un nivel de confianza del 95% y 99% respectivamente. Comparar ambos intervalos y concluir.
- Suponiendo que se desconoce la distribución de la variable altura, ¿qué estrategia de análisis utilizaría para realizar la estimación?

**Ejercicio n° 3:** Se desea establecer el contenido vitamínico de un alimento balanceado para pollos. Se toma una muestra de 49 bolsas y se encuentra que el contenido promedio de vitaminas por cada 100 g es  $\bar{X}=12$  mg. y que la desviación estándar  $S=2$  mg.

- Encontrar el intervalo de confianza del 95%, para el verdadero promedio del contenido de vitaminas usando la distribución normal.
- Encontrar el intervalo del punto a) pero utilizando la distribución T. Enuncie conclusiones.

**Ejercicio n° 4:** El espárrago es una planta perenne cuyo cultivo comercial puede tener una duración de 15 años y su implantación es costosa. Dada la extensión del sistema radicular, la profundidad del suelo es fundamental, considerándose indispensable contar con un promedio mínimo de 80 centímetros de sustrato permeable. Se realizan 14 determinaciones de la profundidad del sustrato permeable (en cm) en puntos tomados al azar en dos campos (A y B). Los valores registrados fueron los siguientes:

<b>A:</b>	<b>72</b>	<b>78</b>	<b>86</b>	<b>78</b>	<b>90</b>	<b>104</b>	<b>76</b>	<b>70</b>	<b>83</b>	<b>75</b>	<b>90</b>	<b>81</b>	<b>85</b>	<b>72</b>
<b>B:</b>	<b>78</b>	<b>82</b>	<b>68</b>	<b>68</b>	<b>74</b>	<b>81</b>	<b>85</b>	<b>73</b>	<b>75</b>	<b>89</b>	<b>100</b>	<b>91</b>	<b>82</b>	<b>75</b>



Los resultados del análisis estadístico fueron:

Variable	Parámetro	Estimación	E.E.	n	LI(95%)	LS(95%)
<b>A</b>	Media	81,43	2,45	14	76,13	86,73
<b>B</b>	Media	80,07	2,42	14	74,85	85,29

- A partir de los intervalos de confianza al 95% determinar si estos campos son aptos para el cultivo.
- ¿Hay diferencias en la profundidad del sustrato permeable entre ambos campos? *Ayuda : observar si los valores de LI y LS de ambos intervalos, se superponen*

**Ejercicio n° 5:** Un productor decide probar el funcionamiento de su máquina y para ello, luego de cosechar una parcela, cuenta en 10 unidades de 1 m<sup>2</sup> la cantidad de semillas que quedan en el suelo. Las normas técnicas indican que la desviación estándar del número de semillas caídas por m<sup>2</sup> no debería ser superior a 5. Los resultados, en semillas/m<sup>2</sup>, fueron:

77 73 82 82 79 81 78 76 76 75

- Construya un intervalo de confianza para  $\sigma^2$  con una confianza del 90%.
- Concluya sobre el funcionamiento de la máquina.

**Ejercicio n° 6:** Para la construcción de silos una fábrica produce un tipo de metal el cual se desea que tenga una resistencia a la rotura más homogénea que un modelo anterior. Para evaluar esta homogeneidad toma 6 piezas de la producción y las somete a pruebas de resistencia. Las siguientes observaciones representan las cargas de rotura en miles de libras por pulgada cuadrada.

15.3	18.7	22.3	17.6	19.1	14.8
------	------	------	------	------	------

Conociendo que la variable cargas de rotura se distribuye normalmente, estime la variabilidad (varianza poblacional) con una confianza del 90%

**Ejercicio n° 7:** Un investigador desea estimar el rendimiento promedio de grasa que contiene la leche producida por vacas de cierta raza en un período de tiempo determinado.

Para ello extrae una muestra de 10 vacas obteniendo:

Media muestral : 36.4 kg.  $s^2 = 26.404 \text{ kg}^2$

Con un nivel de confianza de 95% estime el verdadero valor promedio de grasa en las vacas de esta raza.



**Ejercicio n° 8:** Un periodista especializado en política que ha hecho una encuesta para conocer la popularidad de un importante candidato presidencial. De los 380 encuestados, 325 estaba a favor del candidato. Estime por intervalos la verdadera proporción poblacional favorable al candidato en cuestión.

**Ejercicio n° 9:** En una encuesta, 1200 adolescentes expresaron su opinión con respecto a los principales problemas de la juventud en la actualidad: Los principales problemas planteados fueron:

Drogadicción	37%
Relación y comunicación con los padres	30%
Alcoholismo	17%
Desocupación	16%

Suponiendo que la muestra de adolescentes ha sido seleccionada al azar, estime por intervalos las verdaderas proporciones poblacionales para cada problema en particular. Trabaje con niveles de confianza del 95% y 99%. Comente los resultados.

**Ejercicio n° 10:** Ante los reiterados rechazos de productos que una fábrica de radiadores recibe por parte de sus clientes, el jefe de producción decide estimar el verdadero porcentaje de defectuosos que salen de la fábrica. En caso de que dicho porcentaje supere del 15% está dispuesto a realizar los ajustes que fueran necesarios en el proceso de producción. Con este fin toma una muestra aleatoria de 120 radiadores encontrando que 12 de ellos tenían algún defecto.¿ Debe el jefe de producción realizar algún ajuste? Use  $1 - \alpha = 0,90$

**Ejercicio n° 11:** En la campaña 2000/01, se cosecharon lotes de soja encontrándose altos porcentajes de granos verdes y arrugados. Los mismos correspondieron a zonas en los que el estrés hídrico y calórico fue intenso, acompañado generalmente por una muerte prematura de plantas provocadas por un hongo, *Macrophomina phaseolina* ("podredumbre carbonosa del tallo"). Los granos verdes son de casi la misma densidad y tamaño que los normales con los cuales las separaciones físicas no son posibles. Sobre un muestreo de 110 lotes se ha registrado que 33 presentan estas características.

a) Determine el intervalo de confianza del 95% para la proporción de lotes que presentan granos verdes.

**Ejercicio n° 12:** Se quiere diseñar el tamaño de una muestra para estimar  $\mu$  en una población normal con desviación estándar igual a 13.

a) ¿Cuál debería ser el tamaño mínimo de la muestra para asegurar una amplitud de 9 unidades para el intervalo de confianza al 95%?



---

b) ¿Qué sucede si la confianza cambia al 99%?

**Ejercicio n° 13:** En un estudio se supone una proporción poblacional  $p=0,35$ .  
¿De qué tamaño habrá que tomar una muestra para tener una confianza del 95% y un margen de error de 0,05?



## TRABAJO PRÁCTICO N° 9 PRUEBA DE HIPÓTESIS

**Ejercicio n° 1:** Se desea probar si una nueva variedad de soja lograda por un proceso de mejoramiento genético supera la base de 20 qq/ha. A tal efecto se toma una muestra aleatoria de 30 parcelas de soja obteniendo una media de rendimiento de 22qq/ha y desviación estándar igual a 40 qq/ha

- Plantee las hipótesis del problema
- Pruebe la hipótesis con un nivel de significación del 10%.
- Pruebe la hipótesis con un nivel de significación del 5%.
- Pruebe la hipótesis si la media muestral obtenida fue de 28qq/ha. ( $\alpha=0.05$ )

**Ejercicio n° 2:** Una variable aleatoria sigue una distribución  $N(\mu, 144)$  con  $\mu$  desconocido.

- ¿Se descartaría la hipótesis  $\mu=15$  en favor de la alternativa  $\mu \neq 15$ , si de una muestra aleatoria de  $n=64$  observaciones se obtiene una media igual a 20?
- Construya un intervalo de confianza del 95% para  $\mu$ .
- Considerando la misma hipótesis del punto a), ¿qué sucedería con un nivel de significación del 1%?
- Construya un intervalo de confianza del 99% para  $\mu$ .
- Pruebe  $H_0: \mu=15$  versus  $H_1: \mu > 15$  para  $\alpha=0.05$  y  $\alpha=0.01$ . Compare con los resultados obtenidos en los puntos a) y c).

**Ejercicio n° 3:** Un genetista afirma que el rendimiento de sus híbridos es menor al de sus progenitores, el cual es de 30 qq/ha ( $H_0: \mu=30$  versus  $H_1: \mu < 30$ ). Si la desviación estándar es de 2 qq/ha y trabaja con una muestra de 10 híbridos:

¿Cuál es la probabilidad de que concluya que el rendimiento de los híbridos es menor al de los progenitores, si el rinde promedio es verdaderamente de 29 qq/ha? Trabaje con  $\alpha=0.05$ .

**Ejercicio n° 4:** Un tipo de ratón de laboratorio muestra una ganancia media de peso de 65 gr. durante los primeros tres meses de vida. Doce ratones fueron alimentados con una nueva dieta desde su nacimiento hasta los primeros tres meses de vida, observándose las siguientes ganancias de peso (en gr):

65 62 64 68 65 64 60 62 69 67 62 71

a) ¿Hay razón para creer que la dieta produce una variación significativa en la cantidad de peso ganado? Trabaje con  $\alpha=0.05$  y considere el peso con distribución normal.



**Ejercicio nº 5:** A un chacarero se le ofrece un nutriente con el que obtendrá rendimientos medios superiores a 1850 kg / ha. Se sabe que en esa población los rendimientos tienen un desvío estándar de 210 kg/ha. El chacarero decide realizar una prueba sobre 9 parcelas, en las que pretende observar los rendimientos, promediarlos y de acuerdo al resultado optar o no por el nuevo producto. Acuerda correr un riesgo de rechazar el nutriente cuando en realidad éste cumple con las especificaciones promocionales, fijando para ello una probabilidad de error de 5%.

Si la experiencia arrojó los siguientes resultados para las 9 parcelas:

2150 1950 2170 1800 2050 2120 1820 1850 2100

- ¿Aconsejaría adquirir o no el nutriente?
- Pruebe ahora, un riesgo de error del 1%. Interprete.
- Calcule los puntos críticos y gráfíquelos.

**Ejercicio nº 6:** Se han cosechado 15 parcelas de híbridos de sorgo INTA Blanco, campaña 2003, con el objeto de determinar si el rendimiento obtenido de ese material es significativamente superior a los obtenidos por materiales tradicionales, 8165,0 kg/ha.

9273.8 9145.2 9045.9 1214.5 1112.7 9865.8 1654.4 9802.6  
9567.8 10002.5 9687.2 9956.7 10256.9 11023.4 9782.3

- Demuestre si el rendimiento promedio obtenido es superior al promedio tradicional.  $\alpha = 0.05$
- Calcule e interprete el rendimiento que define la zona de aceptación de la hipótesis nula.

**Ejercicio nº 7:** Un proceso de fabricación produce 12.3 unidades por hora. Esta producción tiene una varianza igual a 4. Se sugiere un nuevo proceso que es costoso de instalar, pero se piensa que puede incrementar la producción. Para decidir si se hace el cambio o no, se prueban 10 máquinas nuevas y se observa que éstas producen en promedio 13.3 unidades.

Calcular la probabilidad del error de tipo II en el contraste para  $H_0: \mu=12.3$  versus  $H_1: \mu>12.3$ , cuando la verdadera esperanza del nuevo proceso es  $\mu=14$ . Trabajar con  $\alpha=0.01$ .

**Ejercicio nº 8 :** Al contrastar la hipótesis  $H_0: \mu=50$  vs  $H_1: \mu\neq 50$ , en una distribución normal con  $\sigma=2$ :

- ¿Con qué frecuencia sería aceptada esta hipótesis si la media verdadera fuese 51? Trabajar con  $\alpha=0.10$  y  $n=10$ .
- Manteniendo una probabilidad de Error Tipo I previamente establecida, ¿cómo se podría disminuir la frecuencia de aceptación de una hipótesis nula falsa?



**Ejercicio n° 9:** Un tipo de ratón de laboratorio muestra una ganancia media de peso de 65 gr. durante los primeros tres meses de vida. Doce ratones fueron alimentados con una nueva dieta desde su nacimiento hasta los primeros tres meses de vida, observándose las siguientes ganancias de peso (en gr):

65	62	64	68	65	64	60	62	69	67	62	71
----	----	----	----	----	----	----	----	----	----	----	----

- a) ¿Hay razón para creer que la dieta produce una variación significativa en la cantidad de peso ganado? Trabajar con  $\alpha=0.05$ .
- b) Calcular para el contraste planteado, las potencias para diferentes valores de  $\mu_1$  variando en el intervalo [62 gr., 70 gr.] y dibujar la curva de potencia.

**Ejercicio n° 10:** Un grupo inversor desea montar una cadena de restaurantes y para asegurar que el precio no será un factor fundamental en la decisión de por lo menos el 70% de los potenciales clientes, encarga a una empresa de encuestas investigar sobre el tema. La empresa decide tomar una muestra aleatoria de 400 personas y luego de sistematizar la información se observó que el 38% de los encuestados consideraba al precio como factor preponderante en la elección.

- a) Formule las hipótesis que se plantearon
- b) Realice la prueba y concluya al 5% de significación

**Ejercicio n° 11:** Un asesor de empresas considera que para invertir en una acción determinada esta debe tener bajo riesgo. Para que sea considerada de bajo riesgo, la variabilidad de su cotización diaria no debe tener una varianza igual a 5. Estudiando durante 30 días la cotización de esta acción, se obtuvo una varianza de 6- Puede considerarse que invertir en esta acción es una inversión de bajo riesgo? Concluya con un nivel de significación de 5%.

## COMPARACIÓN DE POBLACIONES

**Ejercicio n° 12 :** Se busca comparar dos formulaciones de un mismo insecticida sobre el porcentaje de mortalidad de un chinche verde evaluada como el número de insectos muertos en un total de 100 iniciales. El ensayo se realizó tomando 20 lotes de 100 insectos cada uno y asignando al azar 10 lotes para la formulación A y el resto para la formulación B. Los valores obtenidos fueron los siguientes

Formulación A	85	86	92	87	92	90	95	90	92	91
Formulación B	87	86	84	80	89	85	92	89	86	90

¿Existen diferencias estadísticamente significativas entre formulaciones considerando la mortalidad promedio de los insectos? Trabaje con  $\alpha=0,10$



**Ejercicio 13:** Dos lotes de pollos de la misma raza y edad fueron alimentados durante 30 días con dos tipos diferentes de alimento balanceado. Los aumentos de peso, en ramos, fueron:

Balanceado A: 329 363 298 243 391 333 369 432 440 397  
 409 350  
 Balanceado B: 353 405 372 345 377 409 428 421 357 372  
 409 367

- a) Pruebe si existen diferencias estadísticamente significativas entre los aumentos de peso promedio de los dos lotes. Trabaje con un nivel de significación de 5%.  
 b) Estimar la diferencia entre las medias de los tratamientos, con una confianza del 95%.  
 ¿Recomendaría algún balanceado?

**Ejercicio n° 14:** Para probar la eficacia de un tratamiento de poda en un bosque de Raulí, un investigador decide comparar el incremento del diámetro de los fustes de los árboles podados, con el incremento en árboles sin poda. Para ello se localizan 20 lotes de los cuales a 10 se los poda y al resto no. Al cabo de 3 años se obtienen los incrementos promedio para cada lote siendo los resultados los siguientes (en cm):

Con poda: 0.290 0.305 0.280 0.320 0.350 0.297 0.300 0.298  
 0.315 0.324  
 Sin poda: 0.300 0.303 0.270 0.300 0.320 0.310 0.280 0.302  
 0.298 0.301

- a) ¿Cuál es el efecto de la poda? Trabaje con un nivel de significación del 5%.

**Ejercicio n° 15:** Un investigador supone que el estrés que se produce en vacas fistuladas puede disminuir los niveles de fósforo en sangre. Para probar su hipótesis selecciona 8 vacas y a cada una de ellas le extrae una muestra de sangre antes de la fistulación y otra muestra después. Los resultados son:

Vaca	1	2	3	4	5	6	7	8
Antes de fistulación	8,69	7,13	7,79	7,93	7,59	7,86	9,06	9,59
Después de la fistulación	7,24	7,10	7,80	7,95	7,50	7,79	9,00	9,48

- a) ¿Qué conclusión se puede extraer acerca de la fistulación? Utilizar  $\alpha = 0.01$ .



## TRABAJO PRÁCTICO N° 10 DISEÑO Y ANÁLISIS DE EXPERIMENTOS

### A UN CRITERIO DE CLASIFICACIÓN

**Ejercicio n° 1:** Se desea conocer el efecto de las cepas de inoculantes de Rhizobium, fijadoras de nitrógeno atmosférico, sobre el contenido de nitrógeno de plantas de trébol rojo. Para ello se dispone de 30 macetas de trébol rojo en un invernadero. Se asignan al azar 5 macetas para cada una de las cepas y se procede a inocularlas. Los resultados son los siguientes (en mg. de nitrógeno):

Cepa I	Cepa II	Cepa III	Cepa IV	Cepa V	Cepa VI
19.4	17.7	09.1	18.6	11.6	16.9
27.0	24.3	11.9	18.8	11.8	17.3
32.1	24.8	15.8	20.5	14.2	19.1
32.6	25.2	17.0	20.7	14.3	19.4
33.0	27.9	19.4	21.0	14.4	20.8

- ¿Cuales son las unidades experimentales?, ¿Cuántas repeticiones hay?
- Plantee  $H_0$  y  $H_1$
- Realice el Análisis de la Varianza ( $\alpha = 0.05$ ) y concluir sobre si las distintas cepas producen el mismo nivel de fijación de nitrógeno o no.
- Si corresponde, realizar una prueba "a posteriori" e indique que cepa o cepas recomendaría.

**Ejercicio n° 2:** En un estudio sobre el efecto de la adición de azúcares sobre diámetro de secciones de poroto criados en un medio de cultivo, se obtuvieron los siguientes datos:

Control	75	67	70	75	65	71	67	67	76	68
Glucosa	57	58	60	59	62	60	60	57	59	61
Fructosa	58	61	56	58	57	56	61	60	57	58
Gluc. + Fruc.	58	59	58	61	57	56	58	57	57	59
Sacarosa	62	66	65	63	64	62	65	65	62	67

- ¿Qué se puede decir sobre el efecto de los distintos medios de cultivo? Concluya trabajando con un nivel de significación de 0.05.
- Realice los cálculos utilizando software estadístico y haga representaciones gráficas apropiadas.



**Ejercicio n° 3:** Se desea estudiar el efecto de la carga animal sobre la producción de materia seca en una pastura implantada. Para ello se divide un lote en 28 potreros y se asignan aleatoriamente 7 potreros a cada una de las 4 cargas animales en estudio (2 nov./ha., 4 nov./ha, 6 nov./ha. y 8 nov./ha.). Los resultados fueron los siguientes expresados en toneladas de materia seca por hectárea.

								Media
carga 2	2.6	1.9	3.1	2.8	2.2	2.0	2.7	2.47
carga 4	3.3	3.6	3.0	3.5	3.2	3.9	3.4	3.41
carga 6	3.1	2.0	2.5	3.1	2.3	3.0	2.2	2.60
carga 8	2.5	2.3	2.8	1.8	2.7	2.6	2.0	2.39

- Plantee un modelo lineal que permita recomendar alguna carga en especial.
- ¿Qué supuestos se requieren para el análisis de este ensayo?
- Realice el análisis y concluya. Trabaje con un nivel de significación de 0.05.

**Ejercicio n°4:** Una empresa agrícola necesita establecer si le conviene fertilizar sus cultivos de soja y si es así, seleccionar el mejor fertilizante. Para este propósito se realizó un ensayo en un lote de 5 has., dividido en parcelas de 1/4 ha. cada una, asignando los tratamientos en forma aleatoria. Los rendimientos obtenidos (qq/ha) fueron:

Control (sin fertilizar)	Fert. A	Fert. B	Fert. C
23	30	28	27
20	32	36	25
22	29	31	24
20	35	32	28
21	33	34	26

- Haga una representación gráfica comparativa de los rendimientos
- ¿Se recomendaría la fertilización?
- De ser así, ¿cuál de los fertilizantes se recomendaría?

**Ejercicio n° 5:** Un equipo de investigación en bioquímica desea averiguar la efectividad de tres medios de cultivos diferentes en cuanto al retardo del crecimiento de bacterias en recipientes de leche. Los análisis bacteriológicos correspondientes se efectúan en un laboratorio y solamente se pueden realizar 3 ensayos por día.



A causa de que se piensa que los días podrían representar una fuente potencial de variabilidad, el equipo decide usar un diseño en bloques aleatorizados tomando

a cada día como un bloque.

La variable de respuesta fue la cantidad de bacterias en una unidad de medición especificada.

Los análisis se efectúan durante 4 días obteniéndose los siguientes resultados:

	<b>Días (bloques)</b>			
<b>Solución</b>	1	2	3	4
<b>1</b>	13	22	18	29
<b>2</b>	16	24	17	44
<b>3</b>	5	4	1	22

Analice convenientemente los datos y extraiga conclusiones

**Ejercicio n° 6:** Con la finalidad de establecer si la inversión media realizada por la industria en los últimos 12 meses es igual en cuatro ramas, una consultora decide tomar muestras aleatorias de 7 empresas de igual tamaño correspondiente a cada una de ellas: metalúrgica, telecomunicaciones, electrónica y química.

A cada una de ellas se les preguntó el monto de la inversión en el período considerado y los resultados obtenidos fueron los siguientes:

<b>Metalúrgica</b>	<b>Telecomunicaciones</b>	<b>Electrónica</b>	<b>Química</b>
<b>5,4</b>	9,3	4,8	2,7
<b>1,4</b>	3,1	3,3	3,7
<b>3,6</b>	4	1,5	2,3
<b>4,9</b>	8,2	6,7	1,8
<b>2,7</b>	4,1	2,7	5,7
<b>5,2</b>	11,4	7,8	2,3
<b>3,4</b>	5,7	3,5	3,9

- ¿Cuál sería la hipótesis a plantear en esta situación?
- ¿Puede considerarse que las 4 ramas industriales han invertido, en promedio, los mismos montos?
- Comente los resultados obtenidos.

## **A VARIOS CRITERIOS DE CLASIFICACIÓN**

**Ejercicio 7:** El siguiente conjunto de datos corresponde a proteína bruta en leche obtenida con dos suplementos (A y B) en dos dosis (1 y 2). Cada observación corresponde al contenido de proteína bruta en leche de una muestra compuesta obtenida por tambo.



---

---

Tambo	Control	A1	A2	B1	B2
I	3.19	3.03	3.06	3.22	3.33
II	3.16	3.07	3.08	3.28	3.20
III	3.25	3.23	3.24	3.45	3.45
IV	3.48	3.30	3.33	3.44	3.39
V	3.25	3.25	3.24	3.35	3.54
VI	3.10	3.05	2.93	3.28	3.35

---

---

- Calcular la estadística descriptiva básica.
- Identifique el modelo lineal correspondiente.
- Calcule la tabla de ANAVA y si corresponde, utilice una prueba de comparaciones múltiples.
- ¿Qué suplementación se recomendaría si el objetivo es maximizar la concentración de proteína bruta en leche?