

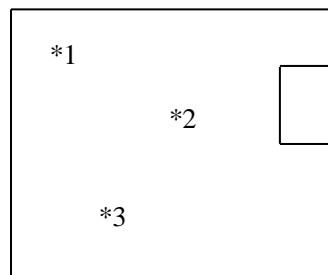
PRÁCTICA N° 1 – ESTUDIOS OBSERVACIONALES Y EXPERIMENTOS

- 1) En cada uno de los siguientes enunciados, indicar si se trata de un estudio observacional (retrospectivo o prospectivo) o experimental, la variable respuesta, unidad, factor o factores en estudio con sus respectivos niveles, tratamientos, y otras fuentes de variabilidad.

En el caso de tratarse de un experimento, desarrollar un posible diseño de campo.

- a) Un investigador desea estudiar los efectos de los medicamentos A y B sobre el recuento de linfocitos en ratones (expresados en miles por ml^3 de sangre) comparando A, B y el placebo P. Al diseñar el experimento se piensa que los ratones de una misma camada darán respuestas más homogéneas que ratones de camadas diferentes. Por tal motivo se utilizan en el experimento 7 camadas de 3 ratones cada una.
- b) El Instituto Nacional de Estadística y Censos realiza, con una periodicidad de cinco años, una encuesta de gastos e ingresos de los hogares, con el objetivo de obtener estadísticas que permitan conocer la situación socioeconómica de los hogares de un país, así como los patrones de consumo de la población. Por ello, selecciona mediante un diseño muestral complejo un conjunto de hogares de algunas localidades de Argentina para obtener la información necesaria. Las variables temáticas consideradas en la encuesta son el gasto y el ingreso de los hogares, y las variables de clasificación son las características demográficas, ocupacionales y educacionales de sus miembros, así como las características de la vivienda.
- c) Se pretende estudiar si el azufre produce una disminución del número de cierto tipo de parásitos de la patata. Para ello se divide el campo en 24 parcelas de igual tamaño. Interesa comparar 3 cantidades distintas de azufre en 2 estaciones del año (primavera y otoño). En cada estación del año y con cada cantidad de azufre se repite el experimento en 4 parcelas que son asignadas al azar entre las 24 parcelas preparadas para el experimento. En cada parcela se mide el número de parásitos de la patata.
- d) Un investigador realizó un estudio sobre las técnicas de dosificación del ácido ribonucleico. En una determinada etapa, el proceso comprende una extracción por una solución de ácido tricloroacético, donde se efectúa una medición del total de ácido ribonucleico extraído. El investigador estaba interesado en saber si el reemplazo del ácido tricloroacético por ácido perclórico asegura una extracción más completa y decidió ensayar el ácido perclórico a dos concentraciones diferentes, 5% y 10%. A tal fin se trataron 15 parte alícuotas de un mismo homogenato de páncreas (5 con cada tratamiento)
- e) El INDEC lleva a cabo la Encuesta Permanente de Hogares (EPH), la cuál es un programa nacional de producción permanente de indicadores sociales cuyo objetivo es conocer las características socioeconómicas de la población. En base a esta encuesta se proporcionan regularmente tasas oficiales de empleo, desocupación, subocupación y pobreza. Por ello, selecciona mediante un diseño muestral complejo un conjunto de personas de algunas localidades de Argentina para obtener la información necesaria. Las variables temáticas consideradas en la encuesta son la condición ocupacional de las personas, así como otras características demográficas, educativas. De igual forma se recaba información sobre las viviendas.
- f) Main y Tippett describen un experimento sobre el hilado de la tela de algodón. El propósito era investigar el efecto del tratamiento de apresto aplicado a la urdimbre. El criterio fue el número de roturas de ella durante el hilado. Cuatro urdimbres se hilaban simultáneamente en cuatro diferentes telares que fueron supervisados por un solo operario. En cada telar se aplicó cada tratamiento de apresto, de manera tal que cada apresto hubiera sido probado en los cuatro telares.

- g) Es sabido que entre los factores que influyen en la determinación del salario de ejecutivos de empresas radicadas en Brasil figuran el nivel educativo alcanzado y los años de experiencia. Es por ello, que se desea realizar una investigación para verificar si una situación similar se observa en Argentina. Se selecciona una muestra aleatoria de 100 ejecutivos a los cuales se les consulta sobre su salario, su nivel educativo y los años de experiencia.
- 2) La única manera satisfactoria de conocer la duración de una cubierta de automóvil es hacerla rodar en un coche de prueba bajo condiciones normales del camino. Se desea comparar 4 tratamientos, cada uno corresponde a gomas preparadas con diferentes procesos. Evalúa los siguientes planes:
- Utilizar 4 autos y en cada auto colocar las 4 cubiertas con el mismo tratamiento. Es decir, un tratamiento por auto.
 - Utilizar 4 autos y en cada uno de ellos asignar a la cubierta delantera derecha el tratamiento 1, a la delantera izquierda el tratamiento 2, a la trasera derecha el tratamiento 3 y a la trasera izquierda el tratamiento 4.
 - Utilizar 4 autos y en cada uno asignar aleatoriamente un tratamiento a una cubierta.
 - Utilizar 4 autos y en cada uno asignar un tratamiento a una cubierta de modo que todos los tratamientos estén una vez en cada posición (usando los 4 autos) y los 4 tratamientos estén presentes en cada auto.
- 3) Los tres principios básicos del diseño de experimentos son: replicación, aleatorización y control de los efectos de variables externas. Considere cada uno de los siguientes escenarios. ¿Se adhiere cada uno de ellos a todos los principios? En caso de respuesta afirmativa, describa de qué manera. En caso de respuesta negativa indique cómo cambiaría el diseño para incorporar el o los principios faltantes. Sea específico y conciso.
- Se realiza un experimento para comparar cómo los diferentes laboratorios miden la fuerza de rotura de una plancha de metal especial. Hay 6 laboratorios en el experimento. A cada laboratorio se le da una sola plancha (como la del gráfico) y se le pide que mida la fuerza en cada uno de tres puntos marcados. Las 6 planchas están manufacturadas bajo las mismas especificaciones. Se determina aleatoriamente qué plancha es asignada a qué laboratorio.



- Un químico desea comparar un nuevo método de ensayo más simple con un método estándar. Prepara una cierta solución y la divide en 40 tubos de ensayo. Aleatoriamente selecciona 20 de estos tubos y le pide a su ayudante técnico que los analice usando el método estándar. Ella por su parte analiza los restantes 20 usando el método nuevo. Al final los dos grupos de resultados de estos ensayos serán comparados.
- Un grupo de estudiantes universitarios cree que un cierto té de hierbas posee un remarkable poder para levantar el ánimo. Para probar lo que creían realizaron visitas semanales a un hogar de ancianos, visitando a los residentes y sirviéndoles de este té. El personal del hogar reportó que luego de varios meses muchos de los residentes estaban más alegres y saludables.

PRÁCTICA N° 2 – COMPARACIÓN DE DOS POBLACIONES

En cada uno de los siguientes enunciados realizar primero un análisis descriptivo de los datos, y luego el test correspondiente

- 1) Los siguientes resultados provienen de 10 determinaciones de porcentaje de cloro en cada uno de dos lotes de un polímero

LOTE 1: 58.6 – 58.5 – 59.6 – 58.9 – 58.4 – 59.0 – 59.3 – 57.9 – 58.6 – 58.3

LOTE 2: 56.7 – 55.5 – 55.7 – 56.6 – 57.6 – 56.7 – 56.1 – 56.2 – 57.2 – 55.2

¿Puede inferirse a partir de estos datos que el porcentaje medio de cloro del Lote 1 es superior al correspondiente del Lote 2?

- 2) Un fabricante de nafta hace pruebas para determinar el rendimiento relativo de los automóviles empleando dos aditivos diferentes. Los resultados (expresados en km. por litro de nafta) son:

ADITIVO 1: 12 – 12 – 15 – 13 – 11 – 10 – 14 – 12

ADITIVO 2: 12 – 14 – 22 – 19 – 25 – 11 – 18 – 17

¿Considera usted que en promedio rinde menos kilómetros la nafta con el Aditivo 1?

- 3) Se desea probar el comportamiento de un cultivar de soja cuando se somete a la semilla de alta calidad a un tratamiento químico, antes de la siembra y antes del almacenamiento. Para ello se tomaron 20 parcelas de terreno homogéneas y se les asignaron los tratamientos al azar. Los datos para rendimiento en kg/parcela son los siguientes:

Semillas almacenadas y luego tratadas (A): 10.2 – 13.3 – 13.2 – 10.0 – 15.1 – 13.3 – 11.1 – 10.0 – 12.0 – 14.0

Semillas tratadas y luego almacenadas (B): 9.5 – 10.3 – 11.3 – 10.0 – 11.5 – 9.5 – 10.5 – 8.9 – 8.7 – ND (dato perdido o no disponible)

Probar la hipótesis que no existen diferencias entre las medias de tratamientos.

- 4) Pese a controversias, los científicos creen que los cereales con alto contenido en fibra reducen la posibilidad de contraer muchos tipos de cáncer. Un investigador sugiere además que personas que desayunan cereales con alto contenido en fibra, consumen menos calorías en el almuerzo, mientras que sucede lo contrario con aquellas personas que no consumen en el desayuno ese tipo de cereales. Si esto fuera cierto, los fabricantes de cereales de alto contenido en fibra podrían sumar otra ventaja en el hecho de consumir sus productos.

Con el objeto de probar esta suposición, se seleccionan en forma aleatoria 30 personas y se las consulta sobre sus hábitos alimenticios en el desayuno y el almuerzo. Cada persona fue identificada como consumidora o no de cereales de alto contenido en fibra, y fue medido el número de calorías consumidas durante el almuerzo.

Consumen: 660 – 605 – 529 – 596 – 564 – 615 – 560 – 635 – 623

No Consumen: 502 – 703 – 735 – 707 – 523 – 534 – 768 – 626 – 620 – 589 – 736 – 565 – 686 – 529 – 632 – 951 – 744 – 632 – 593 – 847

Qué conclusión se puede sacar a partir de los datos?

NOTA: mirar los resultados obtenidos con el supuesto de variancias iguales y distintas. Cuál debo utilizar? Por qué?

- 5) El encargado de planta de una compañía que fabrica equipos de oficina piensa que la productividad de los trabajadores es una función, entre otras cosas, del diseño del trabajo, que se refiere a la secuencias de los movimientos que deben realizar los trabajadores. Dos diseños van a ser considerados para la producción de un nuevo tipo de escritorio de computación ergonómico. Se realizó un estudio donde 10 trabajadores seleccionados al azar ensamblaron el escritorio utilizando el diseño A, y otros 10 lo

ensamblaron siguiendo el diseño B. Se consideró el tiempo de ensamblado en minutos, y el encargado desea testar si los tiempos de los dos diseños difieren.

Diseño A: 5.9 – 5.8 – 6.0 – 5.9 – 5.7 – 5.8 – 5.9 – 5.9 – 5.8 – 5.8

Diseño B: 5.6 – 5.8 – 5.7 – 5.9 – 5.8 – 5.6 – 5.7 – 5.9 – 5.8 – 5.7

Antes de realizar el test...

...cuál es la diferencia en segundos entre los promedios para ambos diseños?

...resulta una diferencia significativa?

Luego de realizar el test...

...resulta una diferencia significativa en términos estadísticos? A qué atribuye dicha conclusión?

- 6) Un grupo de investigadores encabezados por el biólogo Sergio Toledo de la Universidad Nacional de Córdoba, está probando un nuevo alimento para Tatú Carretas basado en una mezcla de invertebrados como hormigas, termitas, arañas, gusanos y larvas. El mismo es de costosa fabricación, pero si se demuestra que en el período de crecimiento comprendido entre el año y los dos años, el aumento promedio de peso de los mismos es 3 Kg mayor que el que se obtendría con el alimento tradicional, el Comité Internacional de Conservación del Tatú Carreta aportaría los fondos necesarios para la fabricación del mismo.

Por ello, se seleccionaron al azar 20 animales de iguales características, y en forma aleatoria se le asignaron los tratamientos al azar. Se les tomó el peso inicial y el final al cabo de un año, siendo los aumentos en los pesos los siguientes (en Kg)

Alimento tradicional: 3.2 – 2.1 – 4.1 – 1.5 – 3.0 – 2.7 – 1.5 – 2.0 – 3.4 – 2.7

Alimento nuevo: 7.5 – 5.1 – 6.2 – 7.2 – 5.7 – 8.1 – 10.0 – 6.3 – 7.1 – 8.0

¿Qué decisión se tomaría en base a los datos de las muestras?

NOTA: realizar este ejercicio utilizando el procedimiento *ttest* de SAS.

PRÁCTICA N° 3 – COMPARACIÓN DE MÁS DE DOS POBLACIONES

1. Se realizó un estudio de ingeniería de tránsito sobre los retrasos en las intersecciones con semáforos en las calles de la ciudad de Rosario. Se usaron tres tipos de semáforo: 1) programado, 2) semiactivado y 3) activado. Se utilizaron cinco intersecciones para cada tipo de semáforo. La medida de retraso utilizada fue el promedio de tiempo que cada vehículo permanece detenido en cada intersección (segundos/vehículo). Los datos son los siguientes:

Programado	Semiactivado	Activado
36.6	17.5	15.0
39.2	20.6	10.4
30.4	18.7	18.9
37.1	25.7	10.5
34.1	22.0	15.2

- a) Indique:
- i) Objetivo del estudio
 - ii) Factor en Estudio
 - iii) Unidad experimental
 - iv) Variable respuesta
- b) Realice un estudio descriptivo de los datos.
- c) Escriba el modelo lineal estadístico para este estudio y explicar sus componentes
- d) Establezca las suposiciones necesarias para un análisis de variancia de los datos
- e) Realice el análisis de variancia
- f) Pruebe la hipótesis de que no hay diferencia entre las medias de retraso para los tipos de semáforo a un nivel de significación de 0.05
- g) Plantee las ecuaciones normales mínimo-cuadráticas para este problema y resolverlas para μ y τ_i , siendo τ_i el "efecto semáforo" y teniendo en cuenta la restricción $\sum_{i=1}^3 \tau_i = 0$
- h) Calcule el intervalo de confianza del 95% estimado para $\tau_2 - \tau_3$
2. Los siguientes datos provienen de una operación particular utilizada por cuatro máquinas en un establecimiento industrial. Cada número representa la cantidad de unidades producidas por la máquina en un día:

Máquina A	Máquina B	Máquina C	Máquina D
293	308	323	333
298	353	343	363
280	323	350	368
288	358	365	345
260	343	340	330

- a) Indique:
- i) Objetivo del estudio
 - ii) Factor en Estudio
 - iii) Unidad experimental
 - iv) Variable respuesta
- b) Realice un estudio descriptivo de los datos.
- c) Plantee un modelo que crea adecuado para estos datos
- d) Realice el test de la hipótesis que las cuatro máquinas producen la misma cantidad de unidades. Interprete el resultado en términos del problema.
- e) Plantee matricialmente los siguientes modelos:
- i) $y_{ij} = \mu + \tau_i + e_{ij}$
 - ii) $y_{ij} = \mu_i + e_{ij}$
- f) Obtener para cada uno de los modelos en d) estimaciones de los parámetros. Para el modelo i) dar una solución obtenida a partir de una inversa generalizada y las soluciones obtenidas a partir de las 3 reparametrizaciones vistas en teoría. La solución para el modelo ii) coincide con una de las reparametrizaciones del modelo i).
- g) Obtener predicciones para las unidades (1,4) y (3,4) para cada una de las soluciones obtenidas en e). Qué conclusiones puede obtener?

3. Un fabricante de televisores está interesado en el efecto que tienen sobre los cinescopios de televisores a color, cuatro diferentes tipos de recubrimiento. Se obtuvieron los siguientes datos de conductividad

Recubrimiento			
1	2	3	4
143	152	134	129
141	149	136	127
150	137	132	132
146	143	127	129

- a) Indique:
- Objetivo del estudio
 - Factor en Estudio
 - Unidad experimental
 - Variable respuesta
- b) Realice un estudio descriptivo de los datos.
- c) Plantee un modelo que crea adecuado para estos datos
- d) Existe diferencia en la conductividad producida por los recubrimientos a un nivel de significación de 0.05?
- e) Estime la media general y los efectos de los tratamientos.
- f) Calcule una estimación por intervalo de 95% de confiabilidad para la media del recubrimiento tipo 4. Calcule una estimación por intervalo con 99% de confiabilidad para la diferencia media entre los recubrimientos 1 y 4.
4. Un fabricante sospecha que los grupos de materia prima suministrados por su proveedor difieren significativamente en contenido de calcio. Por lo general hay un gran número de grupos en el depósito. Cinco de ellos son elegidos al azar para el experimento. Un químico hace 5 determinaciones en cada grupo y obtiene los siguientes resultados:

Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
23.46	23.59	23.51	23.28	23.29
23.48	23.46	23.64	23.40	23.46
23.56	23.42	23.46	23.37	23.37
23.39	23.49	23.52	23.46	23.32
23.40	23.50	23.49	23.39	23.38

- a) Indique:
- Objetivo del estudio
 - Factor en Estudio
 - Unidad experimental
 - Variable respuesta
- b) Realice un estudio descriptivo de los datos.
- c) Plantee un modelo que crea adecuado para estos datos
- d) Existe variación significativa en contenido de calcio de grupo a grupo?
- e) Estime los componentes de variancia
- f) Calcule el intervalo de confianza del 95% para $\sigma_{\tau}^2 / (\sigma^2 + \sigma_{\tau}^2)$
5. Se investigan si un catalizador puede afectar la concentración de un componente en una mezcla líquida. Para ello se seleccionan al azar cuatro catalizadores para realizar la prueba. Se obtienen las siguientes concentraciones:

Catalizador			
1	2	3	4
58.1	56.3	50.1	52.9
57.2	54.9	54.2	49.9
58.4	57.0	55.4	50.0
55.8	55.3		51.7
54.9			

- a) Indique:
- Objetivo del estudio
 - Factor en Estudio

- iii) Unidad experimental
 - iv) Variable respuesta
 - b) Realice un estudio descriptivo de los datos.
 - c) Plantee un modelo que crea adecuado para estos datos
 - d) Existe variación significativa en contenido de calcio de grupo a grupo?
 - e) Estime los componentes de variancia
 - f) Calcule el intervalo de confianza del 95% para $\sigma_r^2 / (\sigma^2 + \sigma_r^2)$
6. Suponga un diseño completamente aleatorizado donde el factor en estudio es aleatorio y hay tres observaciones en cada nivel del factor

Nivel 1	Nivel 2	Nivel 3	Nivel 4
15	16	16	20
12	18	15	21
13	19	14	22

- a) Indique:
 - i) Objetivo del estudio
 - ii) Factor en Estudio
 - iii) Unidad experimental
 - iv) Variable respuesta
 - b) Realice un estudio descriptivo de los datos.
 - c) Plantee un modelo que crea adecuado para estos datos
 - d) Existe variación significativa en los valores observados de nivel a nivel?
 - e) Estime los componentes de variancia
 - f) Calcule el intervalo de confianza del 95% para $\sigma_r^2 / (\sigma^2 + \sigma_r^2)$
7. Siguiendo con la investigación vista en el punto 6 de la práctica 2, le llegan al biólogo Toledo información de una nuevo alimento colombiano que tuvo buenos resultados en investigaciones aplicadas en el sur colombiano sobre los *Prionotus maximus*. Por tal motivo, lleva a cabo una investigación similar al anterior, seleccionando al azar 21 animales de iguales características, y en forma aleatoria se le asignaron cada uno de los 3 tratamientos al azar. Se les tomó el peso inicial y el final al cabo de un año, siendo los aumentos en los pesos los siguientes (en Kg)

Alimento tradicional: 8.01 – 8.00 – 8.02 – 8.01 – 8.01 – 8.02 – 8.00
 Alimento nuevo: 8.03 – 8.02 – 8.01 – 8.02 – 8.02 – 8.03 – 8.01
 Alimento colombiano: 8.04 – 8.02 – 8.03 – 8.03 – 8.04 – 8.02 – 8.04

- a) Indique:
 - i) Objetivo del estudio
 - ii) Factor en Estudio
 - iii) Unidad experimental
 - iv) Variable respuesta
- b) Realice un estudio descriptivo de los datos.
- c) Plantee un modelo que crea adecuado para estos datos
- d) Existe diferencia en el aumento de peso entre alimentos a un nivel de significación de 0.05?
- e) Estime la media general y los efectos de los tratamientos.
- f) Si observa los valores observados en este caso y las vistas en el punto 6 de la práctica 2, va a notar ciertas diferencias. A qué podría atribuir las?

PRÁCTICA N° 4 – COMPARACIÓN DE MÁS DE DOS POBLACIONES – COMPARACIONES MÚLTIPLES

3. Una cierta planta ha sido cultivada con cinco fertilizantes distintos. Se desea estudiar el efecto del tipo de fertilizante en la longitud de la planta, para lo cual se midieron cinco series de 10 plantas, obteniéndose:

Fertilizante					
1	2	3	4	5	
10	11	7	12	7	
12	18	14	9	6	
8	12	10	11	10	
10	15	11	10	7	
6	13	9	7	7	
13	8	10	8	5	
9	15	9	13	6	
10	16	11	14	7	
8	9	7	10	9	
9	13	9	11	6	

- i) Indique si es significativo el efecto del fertilizante. Ayuda:

$$\sum_{i=1}^5 \sum_{j=1}^{10} y_{ij}^2 = 5351 \quad \bar{y}_{1\cdot} = 9.5 \quad \bar{y}_{2\cdot} = 13 \quad \bar{y}_{3\cdot} = 9.7 \quad \bar{y}_{4\cdot} = 10.5 \quad \bar{y}_{5\cdot} = 7$$

- j) Realice tests de comparaciones múltiples a fin de comparar todos los pares de medias de tratamientos.
- k) Suponga que quien realiza la experiencia quiere plantear contrastes ortogonales considerando que el fertilizante 5 es el que se utiliza actualmente, mientras que los fertilizantes 1 y 3 son dos nuevas marcas de la fábrica A y los fertilizantes 2 y 4 son dos nuevas marcas de la fábrica B.
4. Realizar para el ejercicio 1 de la práctica 3 los tests de comparaciones múltiples a fin de comparar todos los pares de medias de tratamientos.
- Se quieren plantear contrastes ortogonales considerando que los semáforos programados son los utilizados actualmente y que los semiactivados y activados pueden ser dos soluciones al problema de las demoras en el tránsito.
5. Realizar para el ejercicio 2 de la práctica 3 los tests de comparaciones múltiples a fin de comparar todos los pares de medias de tratamientos.
- Se quieren plantear contrastes ortogonales considerando que las máquinas A y B fueron son de una marca nacional, mientras que las máquinas C y D resultan ser de marca extranjeras.
6. Realizar para el ejercicio 3 de la práctica 3 los tests de comparaciones múltiples a fin de comparar todos los pares de medias de tratamientos.
- Se quieren plantear contrastes ortogonales considerando que el recubrimiento 1 es el más económico pero poseen una demora en la entrega de 1 mes, el 2 y el 3 son de un precio medio y la entrega es de 15 días, y el 4 es el más caro de todos, pero la entrega es en 24 horas.
7. Realizar para el ejercicio 7 de la práctica 3 los tests de comparaciones múltiples a fin de comparar todos los pares de medias de tratamientos.

PRÁCTICA Nº 5 – COMPROBACIÓN DE SUPUESTOS

- Un biólogo marino estaba interesado en la relación entre distintos ambientes costeros y las poblaciones de cangrejos ermitaños que habitan el lugar. El biólogo contó los cangrejos ermitaños en 25 secciones, ubicadas al azar en seis sitios diferentes de la costa. La siguiente proporciona el número de cangrejos contados en cada sección:

Fertilizante						
1	2	3	4	5	6	
0	415	0	0	0	0	0
0	463	0	0	1	0	0
22	6	4	0	1	0	0
3	14	13	4	2	2	2
17	12	5	2	2	3	3
0	0	1	2	1	0	0
0	3	1	5	2	0	0
7	1	4	4	29	4	4
11	16	4	2	2	0	0
11	55	36	1	2	5	5
73	142	407	0	0	4	4
33	10	0	12	13	22	22
0	2	0	1	0	0	0
65	145	18	30	19	64	64
13	6	4	0	1	4	4
44	4	14	3	3	4	4
20	5	0	28	26	43	43
27	124	24	2	30	3	3
48	24	52	21	5	16	16
104	204	314	8	4	19	19
233	0	245	82	94	95	95
81	0	107	12	1	6	6
22	56	5	10	9	22	22
9	0	6	2	3	0	0
2	8	2	0	0	0	0

Plantear los supuestos y comprobar los mismos. Plantear las soluciones para los casos en que no se cumplen.

- Realizar las mismas comprobaciones y plantear soluciones para los datos de los ejercicios 1 al 7 de la práctica 3.

PRÁCTICA N° 6 – DETERMINACIÓN DEL TAMAÑO DE MUESTRA

8. En el ejercicio 1 de la práctica 3 se planteaba un estudio de ingeniería de tránsito sobre los retrasos en las intersecciones con semáforos en las calles de la ciudad de Rosario. Se usaron tres tipos de semáforo: 1) programado, 2) semiactivado y 3) activado, considerando cinco intersecciones para cada tipo de semáforo.
Utilizar esos datos para determinar cuál es la potencia del test si se utilizan 3, 6 y 11 intersecciones para cada tipo de semáforo para rechazar la hipótesis nula a un nivel de significación de 0.05, si los retrasos medios respectivos en los tres tipos de señal fueron 22, 18 y 14 segundos.
9. Una fábrica produce tuercas con máquinas de 20 marcas distintas de iguales características. Desea saber si puede existir un efecto máquina que produzca un cambio en el diámetro de las mismas. Para ello selecciona al azar 5 marcas y se registran los diámetros promedios durante un número de días. Se desea realizar un análisis de variancia para testar la hipótesis planteada con un nivel de significación de 0.05. De un estudio anterior se obtuvieron estimaciones de las componentes de variancia, las cuales fueron $\hat{\sigma}_e^2 = 5.82$ y $\hat{\sigma}_r^2 = 6.81$.
 - a. Qué potencia tendría el test si se utilizaran 3, 5, y 8 días?
 - b. Cuántos días se deberían utilizar si se deseara una potencia de 0.95?
10. Suponga que $\mu_1 = 50$, $\mu_2 = 60$, $\mu_3 = 50$, y $\mu_4 = 60$ son las medias de cuatro poblaciones normales. Cuántas observaciones deben tomarse de cada población para que la probabilidad de rechazar la hipótesis nula de igualdad de medias sea al menos de 0.90?. Supóngase que $\alpha = 0.05$ y que una estimación razonable de la variancia del error es $\sigma^2 = 25$.
11. Se desean utilizar los datos del ejercicio 2 de la práctica 3 para calcular un tamaño de muestra para un experimento similar, considerando un nivel de significación de 0.05. Considerar distintos tamaños, indicando la potencia que obtendría con cada uno de ellos.
12. Cuántas muestras tendrá que tomar un patólogo de plantas de cada uno de cinco lotes de semilla de algodón para detectar una razón $\sigma_r^2 / \sigma_e^2 = 2$, con un nivel de significación de 0.01 y potencias de 0.80, 0.90 y 0.95.
13. Supongamos que por razones de tiempo los tamaños calculados en el ejercicio 5 deben ser reducidos en un 50%. Calcular las nuevas potencias.
14. Una empresa de transporte por camiones desea saber si 4 marcas de neumáticos poseen la misma vida útil (medida en miles de km). De acuerdo a resultados brindados por el estadístico de la empresa, las medias medidas hace 3 años para cada una de las marcas fueron de 56, 76, 58, 67, mientras que en aquella oportunidad el CME fue de 75. Cuántas observaciones se deben tomar de cada población para tener una potencia de al menos 0.95? Considerar $\alpha = 0.05$.

PRÁCTICA N° 7 – DATOS CLASIFICADOS SEGÚN DOS O MÁS CRITERIOS

- Un gran número de plantas se asignó al azar entre macetas, obteniéndose cuatro plantas por macetas. Seis tratamientos se asignaron al azar entre las macetas, siendo tres el número de macetas en cada uno de ellos. Los tratamientos responden a un diseño factorial 2x3. El segundo factor era horas de exposición a la luz diurna (8, 12, 16) y el primero constaba de dos niveles, temperatura alta o baja del invernadero durante el tiempo que permanecía en la oscuridad. Los datos se presentan en el siguiente cuadro:

PLANTA	BAJAS TEMPERATURAS								
	HORAS DE EXPOSICION								
	8			12			16		
	MACETA			MACETA			MACETA		
	1	2	3	1	2	3	1	2	3
1	3.5	2.5	3.0	5.0	3.5	4.5	5.0	5.5	5.5
2	4.0	4.5	3.0	5.5	3.5	4.0	4.5	6.0	4.5
3	3.0	5.5	2.5	4.0	3.0	4.0	5.0	5.0	6.5
4	4.5	5.0	3.0	3.5	4.0	5.0	4.5	5.0	5.5
\bar{Y}_{ijk}	15.0	17.5	11.5	18.0	14.0	17.5	19.0	21.5	22.0

PLANTA	BAJAS TEMPERATURAS								
	HORAS DE EXPOSICION								
	8			12			16		
	MACETA			MACETA			MACETA		
	1	2	3	1	2	3	1	2	3
1	8.5	6.5	7.0	6.0	6.0	6.5	7.0	6.0	11.0
2	6.0	7.0	7.0	5.5	8.5	6.5	9.0	7.0	7.0
3	9.0	8.0	7.0	3.5	4.5	8.5	8.5	7.0	9.0
4	8.5	6.5	7.0	7.0	7.5	7.5	8.5	7.0	8.0
\bar{Y}_{ijk}	32.0	28.0	28.0	22.0	26.5	29.0	33.0	27.0	35.0

- Escriba el modelo adecuado para estos datos considerando que las macetas son una muestra aleatoria de macetas y ambos factores son fijos.
 - Diseñe el cuadro ANOVA con fuentes de variación, grados de libertad y esperanzas de los cuadrados medios.
 - Indique cómo realizaría el test de que no existe interacción entre los dos factores. Calcule el valor de F para esa sola hipótesis con los datos del cuadro. Escriba sus conclusiones explicando claramente el significado de la interacción
- Se realizaron experimentos en bloques completos en tres localidades diferentes. En cada experimento el número de tratamientos (variedades) era igual a 6 y los bloques eran cuatro. Los factores localidad, tratamientos y bloques, se encuentran anidados o cruzados? Cuál es el modelo que corresponde a la experiencia? Si las localidades y bloques son aleatorios, cómo realizaría el test respecto de la igualdad de los tratamientos? Es posible estimar las variancias entre bloques y localidades?
 - Se desea probar la toxicidad de una droga sobre ratas. Se utilizaron en la experiencia 4 camadas de ratas, de 6 animales cada uno. Los animales de una misma camada se dividieron en dos grupos de igual tamaño. A uno se le aplicó la droga, al otro no. Al final de la experiencia se realizaron análisis de los hígados de las ratas experimentales y las determinaciones se hicieron por duplicado.
Considere los factores: camadas, ratas, drogas, determinaciones. Cuáles están cruzados y cuáles anidados?
Escriba el modelo y el ANOVA correspondiente. Cómo realizaría el test de que la droga no es tóxica? Cómo estimaría la variancia entre camadas y entre animales?
 - Se condujo un experimento para estimar la variancia genética y ambiental del maíz utilizando las 48 progenies obtenidas de cruzar cada planta masculina (en número de 12) con 4 diferentes plantas femeninas. El campo se subdividió en 3 bloques. En cada bloque de 16 parcelas se sembraron las progenies de 4 plantas masculinas cruzadas con 4 diferentes plantas femeninas. De cada parcela se tomaron 4 muestras donde se pesó la producción de grano. El plano de campo es el siguiente (antes de aleatorizar)

	BI				BII				BIII			
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
F1	-	-	-	-	-	-	-	-	-	-	-	-
F2	-	-	-	-	-	-	-	-	-	-	-	-
F3	-	-	-	-	-	-	-	-	-	-	-	-
F4	-	-	-	-	-	-	-	-	-	-	-	-

Los factores a considerar son bloques, plantas masculinas, plantas femeninas, y muestras. Escriba el modelo, el ANOVA con fuentes de variación, grados de libertad y esperanzas de los cuadrados medios. Cómo estimarían la variación entre plantas femeninas?

- En una fábrica de fideos se recibieron 6 cargamentos de harina de trigo provenientes de diferentes molinos que eran los proveedores habituales de la empresa y de cada carga se extrajeron 2 bolsas con el objeto de determinar el contenido total de humedad de la harina, realizándose en cada bolsa dos determinaciones. Escribir el modelo correspondiente. Plantear las sumas de cuadrados, grados de libertad, cuadrados medios y E(CM) de los efectos correspondientes al modelo postulado. Indicar la expresión de los estimadores de las componentes de variancia.
- En una industria carbonífera se deseaba comparar la materia prima de varios proveedores y varios métodos para determinar el porcentaje de ácido H en las mismas. Se tomaron 3 proveedores de carbón. Del material de cada proveedor se tomaron tres muestras. Cada muestra se dividió en dos partes. En una parte se realizó la determinación por el método A y en la otra se usó el método B. Considere los factores: proveedores, muestras y métodos, ellos están cruzados o anidados? Escriba el ANOVA (fuentes de variación, grados de libertad y E(CM)). Cómo realizaría el test de que los métodos no difieren entre sí, considerando que las muestras son aleatorias pero proveedores y métodos son fijos?
- Considere el ejemplo siguiente, sobre el número de huevos puestos por cada una de 12 hembras de 25 razas en los cuatro primeros días de postura. El experimento se llevó a cabo cuatro veces (repeticiones). Los factores razas, hembras y repeticiones se consideran aleatorios. Los días de postura son fijos. Son estos efectos (dos a dos) cruzados o anidados? Escriba la tabla de ANOVA con fuentes de variación, grados de libertad y esperanza de los cuadrados medios. La componente más importante es la variación entre experimentos (o repeticiones), cómo haría para estimarla? Y el test de que esa variancia es nula? Indique cómo realizaría el test de que los cuatro días tienen efectos iguales sobre la producción de huevos.
- En una experiencia sobre los efectos de dos fungicidas sobre una variedad de caléndulas se escogieron 5 plantas al azar, de cada planta se tomaron 4 hojas. Cada hoja se dividió en dos mitades. En una mitad se aplicó el fungicida F1 y en la otra el F2. Considere los factores plantas hojas y fungicidas, cuáles están cruzados y cuáles anidados?. Considere que las plantas representan un factor aleatorio, al igual que las hojas. Escriba el modelo, el cuadro ANOVA con fuentes de variación, grados de libertad y esperanza de los cuadrados medios. Cómo estima la variancia entre plantas y entre hojas? Indique cómo realizar el test de que los fungicidas no difieren entre sí.
- Johnson y Tsao condujeron un experimento psicológico para determinar la diferencia en los umbrales de conciencia respecto a los aumentos de peso, cuando estos varían a tasas constantes. Dos tipos de personas fueron seleccionados, videntes y no videntes. Dos hombres y dos mujeres fueron elegidos para representar cada clase, dando un total de ocho personas. Fueron sometidos a una prueba que consistía en adivinar el peso de objetos que tenían pesos diferentes y se calculó un índice de conciencia en base al número de aciertos. Dos meses después se tomó a esas mismas personas una nueva prueba. En total 16 observaciones. Los factores a considerar son: videncia (2 niveles), sexo (2 niveles), personas y pruebas. Las personas dentro de cada sexo y condición se consideran una muestra aleatoria de personas.
Cómo sería el ANOVA? (FV, GL, cuadrados medios y esperanza de los cuadrados medios)
Cómo habría sido el ANOVA si la repetición se hubiera realizado con otras personas?
- Una industria posee 3 institutos de capacitación para sus empleados, una para cada una de las regiones donde opera. Cada instituto posee 2 docentes que dan clases a aproximadamente 15 empleados 3 veces en una semana. Los directivos están preocupados por los efectos que puedan tener tanto los institutos como los docentes en la capacitación

dada. Para investigar esto, se forman cursos en cada uno de los institutos de la manera usual y se los asigna en forma aleatoria a los docentes de cada instituto. Esto se desarrolla en 2 oportunidades (curso), y al final de cada una de las mismas se mide a través de una serie de medidas resúmenes los conocimientos aprendidos.

Los factores instituto, docentes y cursos se encuentran anidados o cruzados? Cuál es el modelo que corresponde a la experiencia? Plantear las sumas de cuadrados, grados de libertad, cuadrados medios y E(CM) de los efectos correspondientes al modelo postulado. Como realizaría el test para ver si hay efecto instituto? Y docente?

PRÁCTICA N° 8 – ANÁLISIS DE COVARIANCIA

1. Se realizó un experimento sobre la resistencia de puntos de soldadura para tres tipos de aleación de acero. Se hicieron seis soldaduras en cada aleación y se midió la fuerza necesaria para separarlas. Se midió el diámetro de la soldadura porque se pensó que afectaba su resistencia. Los datos se muestran en la siguiente tabla, donde y es la resistencia de soldadura y x representa el diámetro de la soldadura

Aleación	y	x	Aleación	y	x	Aleación	y	x
1	37.5	12.5	2	57.5	16.5	3	38.0	15.5
1	40.5	14.0	2	69.5	17.5	3	44.5	16.0
1	49.0	16.0	2	87.0	19.0	3	53.0	19.0
1	51.0	15.0	2	92.0	19.5	3	55.0	18.0
1	61.5	18.0	2	107.0	24.0	3	58.5	19.0
1	63.0	19.5	2	119.5	22.5	3	60.0	20.5

- Considerando a x como covariable, escriba un modelo lineal para el experimento, identificando cada uno de los términos.
- Plantee los supuestos necesarios para realizar un ANCOVA.
- Realizar el test para establecer si existen diferencias entre las aleaciones.
- Realizar el test para establecer si el efecto de la covariable es significativo.
- Verificar el supuesto de que los coeficientes de regresión son iguales en todos los grupos.
- Establecer la estimación de la línea de regresión para cada aleación y graficarlas.

En todos los ejercicios resolver cada uno de los puntos planteados en el ejercicio 1.

2. Un nutriólogo realizó un experimento para evaluar los efectos de cuatro complementos vitamínicos en el aumento de peso de animales de laboratorio. El experimento con un diseño totalmente aleatorizado, con cinco animales, en jaulas separadas, para cada tratamiento. Como el consumo de calorías difiere entre los animales e influye en el aumento de peso, el investigador lo midió en cada animal. Los datos del aumento de peso en gramos (y) y el consumo de calorías (x) son

Dieta	y	x	Dieta	y	x	Dieta	y	x	Dieta	y	x
1	48	35	2	65	40	3	79	51	4	59	53
1	37	44	2	49	45	3	52	41	4	50	52
1	78	44	2	37	37	3	63	47	4	59	52
1	69	51	2	73	53	3	65	47	4	42	51
1	53	47	2	63	42	3	67	48	4	34	43

3. Un ingeniero está estudiando el efecto que tiene la velocidad de corte en una máquina-herramienta sobre la eliminación de material en una operación de maquinado. Sin embargo, la eliminación de material está relacionada con la dureza de la probeta. Se recopilan cinco observaciones para cada rapidez de corte. La cantidad de material eliminado (y) y la dureza de la probeta (x) aparece en la siguiente tabla.

Velocidad	y	x	Velocidad	y	x	Velocidad	y	x
1000	68	120	1200	112	165	1400	118	175
1000	90	140	1200	94	140	1400	82	132
1000	98	150	1200	65	120	1400	73	124
1000	77	125	1200	74	125	1400	92	141
1000	88	136	1200	85	133	1400	80	130

PRÁCTICA N° 9 – PRINCIPIOS DEL DISEÑO DE EXPERIMENTOS

4. Se realizó un experimento en un invernadero para comparar el efecto de 7 fungicidas y un control sobre el crecimiento de brotes de maíz. Se usaron 4 bloques y en cada uno se adjudicaron los tratamientos al azar, disponiéndose de 25 semillas por unidad experimental. La variable medida fue la altura de los brotes medidos en cm. Los tratamientos fueron:

A: control

B y C: fungicidas con mercurio

D y H: fungicidas sin mercurio, compañía I

E, F y G: fungicidas sin mercurio, compañía II; donde F y G son nuevas formulaciones de E

Tratamientos	Bloques			
	1	2	3	4
A	34.45	34.45	36.87	31.95
B	53.13	60.67	78.46	69.73
C	48.45	53.13	48.45	46.15
D	39.23	41.55	43.85	34.45
E	34.45	31.95	11.54	11.54
F	34.45	34.45	20.27	20.27
G	31.95	29.33	29.33	29.33
H	43.85	60.67	36.87	41.55

- Qué hipótesis se desea ensayar?
 - Realice el análisis de la variancia completo
 - Calcule contrastes ortogonales de interés y pruebe su significación estadística
 - Use el test de Tukey para comparar medias de tratamientos
 - Calcule la eficiencia relativa del diseño en bloques completos con respecto al completamente aleatorizado
 - Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.
 - Suponga que se perdió la observación del tratamiento F, bloque 2. Cómo realizaría el análisis?
5. 16 sujetos excedidos de peso participaron en un estudio para comparar cuatro regímenes para reducir de peso. Los sujetos se agruparon de acuerdo con el peso inicial y cada uno de cuatro sujetos de cada grupo de peso inicial fue asignado aleatoriamente a uno de los cuatro regímenes reductores. Al final del período experimental, se registraron las siguientes pérdidas de peso (en kg.)

Peso Inicial (kg.)	Régimen			
	A	B	C	D
(75-87]	6.0	13.0	12.0	11.5
(87-100]	7.5	14.5	11.5	12.5
(100-112]	7.5	13.5	12.5	12.0
Más de 112	9.0	19.0	16.5	15.5

- Especifique la variable respuesta y la ecuación del modelo de sus observaciones
 - Qué hipótesis se desea ensayar?
 - Realice el ANOVA completo y extraiga conclusiones
 - Use el test de Duncan para comparar medias de tratamientos
 - Calcule la ER respecto a un diseño completamente aleatorizado
 - Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.
 - Explique cómo realizaría el análisis si se hubieran perdido 2 observaciones en el experimento
6. En una plantación de cítricos de la localidad de San Pedro, se llevó a cabo un experimento de irrigación con un diseño en bloques completamente aleatorizado en el que se usaron seis tratamientos de irrigación en ocho bloques de árboles. Los siguientes datos se refieren al peso en libras de la fruta cosechada en cada parcela.

Método	Bloque							
	1	2	3	4	5	6	7	8
Goteo	450	469	249	125	280	352	221	251
En canal	358	512	281	58	352	293	283	186
Rocío	331	402	183	70	258	281	219	46
Aspersor	317	423	379	63	289	239	269	357
Aspersor/rocío	479	341	404	115	182	349	276	182
Anegado	245	380	263	62	336	282	171	98

- Escriba un modelo lineal para el experimento
 - Qué hipótesis se desea ensayar?
 - Realice el análisis de la variancia completo
 - Calcule el error estándar estimado para una media de tratamiento de irrigación
 - Considere que el método anegado es la práctica normal. Use el método de Dunnett para probar la diferencia entre el anegado y cada uno de los otros métodos.
 - Calcule la eficiencia relativa de este diseño respecto a un diseño totalmente aleatorizado. Cuáles son sus conclusiones?
 - Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.
7. Se midió la autoinductancia de bobinas con núcleos de óxido de hierro con diferentes condiciones del puente de medición. La temperatura de la bobina se mantuvo constante, se usaron cinco bobinas en el experimento y se midió la autoinductancia para cuatro temperaturas (22°, 23°, 24° y 25°) del puente de medición; las temperaturas se utilizaron en orden aleatorio para cada bobina. Los siguientes datos son las desviaciones porcentuales de una medida estándar.

Temperatura	Bobina				
	1	2	3	4	5
22	1.400	0.264	0.478	1.010	0.629
23	1.400	0.235	0.467	0.990	0.620
24	1.375	0.212	0.444	0.968	0.495
25	1.370	0.208	0.440	0.967	0.495

- Escriba un modelo lineal para el experimento
 - Qué hipótesis se desea ensayar?
 - Realice el análisis de la variancia completo
 - Use el método de Tukey para probar la diferencia entre el anegado y cada uno de los otros métodos.
 - Calcule la eficiencia relativa de este diseño respecto a un diseño totalmente aleatorizado. Cuáles son sus conclusiones?
 - Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.
8. Se encuentra bajo estudio el efecto que tienen cinco reactivos distintos (A, B, C, D y E) sobre el tiempo de reacción de un proceso químico. Cada lote de material nuevo es lo suficientemente grande para permitir que sólo se realicen 5 ensayos. Cada ensayo tarda, aproximadamente una hora y media, por lo que sólo pueden realizarse cinco ensayos por día. Los resultados se presentan en la siguiente tabla:

Lote	Día				
	1	2	3	4	5
1	8 (A)	7 (B)	1 (D)	7 (C)	3 (E)
2	11 (C)	2 (E)	7 (A)	3 (D)	8 (B)
3	4 (B)	9 (A)	10 (C)	1 (E)	5 (D)
4	6 (D)	8 (C)	6 (E)	6 (B)	10 (A)
5	4 (E)	2 (D)	3 (B)	8 (A)	8 (C)

- Escriba un modelo lineal para el experimento
- Qué hipótesis se desea ensayar?

- c. Realice el análisis de la variancia completo
 - d. Use el método de Duncan para probar la diferencia entre los reactivos.
 - e. Calcule la eficiencia relativa del bloqueo por día y la ER del bloqueo por lote. Cuáles son sus conclusiones?
 - f. Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.
9. Los siguientes datos corresponden a un experimento llevado a cabo sobre un "factor de difusión", una sustancia que cuando está presente en la inoculación hecha en laboratorio a conejos, hace que se extienda la ampolla causada por esa inoculación. Los investigadores inoculan a 6 animales en 6 lugares distintos del lomo. El diseño permitió estudiar el efecto del "orden de administración".

Posición	Animales					
	1	2	3	4	5	6
1	7.9 (3)	8.7 (5)	7.4 (4)	7.4 (1)	7.1 (6)	8.2 (2)
2	6.1 (4)	8.2 (2)	7.7 (6)	7.1 (5)	8.1 (3)	5.9 (1)
3	7.5 (1)	8.1 (3)	6.0 (5)	6.4 (6)	6.2 (2)	7.5 (4)
4	6.9 (6)	8.5 (1)	6.8 (3)	7.7 (2)	8.5 (4)	8.5 (5)
5	6.7 (2)	9.9 (4)	7.3 (1)	6.4 (3)	6.4 (5)	7.3 (6)
6	7.3 (5)	8.3 (6)	7.3 (2)	5.8 (4)	6.4 (1)	7.7 (3)

Los números entre paréntesis indican el orden de inoculación y los otros corresponden al área de la ampolla luego de la inoculación.

- a. Escriba un modelo lineal para el experimento
 - b. Qué hipótesis se desea ensayar?
 - c. Realice el análisis de la variancia completo
 - d. Use el método de Tukey para probar la diferencia entre los reactivos.
 - e. Calcule la eficiencia relativa del bloqueo por animal y la ER del bloqueo por posición. Cuáles son sus conclusiones?
 - f. Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.
10. Un ingeniero de tránsito realizó un estudio para comparar el tiempo sin uso de la luz roja para cinco secuencias distintas de semáforo. El experimento se llevó a cabo con un diseño de cuadrado latino en el que los dos factores de bloque eran: 1) cinco cruces de calles elegidos al azar y 2) cinco períodos. En la tabla de datos, las cinco secuencias de tratamiento se muestran entre paréntesis como A, B, C, D, E y los valores numéricos son el tiempo de luz roja sin uso expresado en minutos.

Cruce	Período				
	1	2	3	4	5
1	15.2 (A)	33.8 (B)	13.5 (C)	27.4 (D)	29.1 (E)
2	16.5 (B)	26.5 (C)	19.2 (D)	25.8 (E)	22.7 (A)
3	12.1 (C)	31.4 (D)	17.0 (E)	31.5 (A)	30.2 (B)
4	10.7 (D)	34.2 (E)	19.5 (A)	27.2 (B)	21.6 (C)
5	14.6 (E)	31.7 (A)	16.7 (B)	26.3 (C)	23.8 (D)

- a. Escriba un modelo lineal para el experimento
- b. Qué hipótesis se desea ensayar?
- c. Realice el análisis de la variancia completo
- d. Use el método de Tukey para probar la diferencia entre los reactivos.
- e. Calcule la eficiencia relativa del bloqueo por animal y la ER del bloqueo por posición. Cuáles son sus conclusiones?
- f. Plantee un estimador para los residuos con los cuales realizaría la comprobación de los supuestos.