



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
IZTACALA

DISEÑO DE UN PROGRAMA DE COMPUTO PARA LA ENSEÑANZA DE ANOVA EN BIOLOGIA

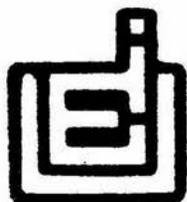
T E S I S

QUE PARA OBTENER EL TITULO DE

B I O L O G O

P R E S E N T A :

MOISES SILVA CERVANTES



Los Reyes, Iztacala

Febrero de 1989.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A JAIME

- POR SU DIRECCION EN ESTA TESIS -,

CORRIGIO, SUGIRIO Y AGREGO

DONDE FUE NECASARIO.

DOCTOR, MAESTRO, BIOLOGO

Y SOBRETUDO : EXCELENTE AMIGO.

MI AMPLIO AGRADECIMIENTO, DR. JAIME CURTS.

A LOS BIOLOGOS ANGEL DURAN Y AGUSTIN GRANADOS ...

- PROFESORES DEL DEPTO. DE MATEMATICAS

DE LA ENEPI -,

DE QUIENES TOME IDEAS MUY VALIOSAS.

AL M. EN C. ENRIQUE KATO

Y AL BIOLOGO ANTONIO CISNEROS

LOS CINCO ANTERIORES

SINODALES Y JURADO CALIFICADOR

EN MI EXAMEN PROFESIONAL.

MIL GRACIAS.

A MIS PADRES

QUE, AUN SIN UNA FORMACION PROFESIONAL,
ME ENCAUSARON
HASTA LOGRAR EN MI UN ANHELO.

A TODOS MIS FAMILIARES Y AMIGOS

CONSECUENTARON MI FORMA DE SER
Y ME ESTIMULARON HACIA EL LOGRO.

A MIS TIOS FILOMENA Y MAYOLO

- HOMENAJE POSTUMO -

QUE EN VIDA PARTICIPARON DE MIS REALIZACIONES
Y DESEABAN COMO YO ESTE MOMENTO.

LO HEMOS LOGRADO.

A QUIEN CONSULTE ESTE TRABAJO

EN FORMA MUY ESPECIAL,

- YA QUE

ES LA FINALIDAD DEL MISMO -.

QUE LE SEA DE PROVECHO;

QUE LO CRITIQUE Y LO ENRIQUEZCA.

Y, FINALMENTE, A DIOS GRACIAS

EN MI CONCEPTO QUE TENGO DE EL :

IDENTIDAD UNIVOCA Y EQUIDISTANTE;

NATURALEZA INOBJETABLE E INEXPLICABLE;

PRESENCIA LATENTE E INVISIBLE;

CULMINACION Y PRINCIPIO;

.....

SUBUNIVERSO INMERSO Y UNIVERSO.

I N D I C E

	pag
INTRODUCCION	
PARTE I GENERALIDADES	7
PARTE II LA COMPUTADORA EN LA ESTADISTICA	10
PARTE III ANALISIS DE VARIANZA	13
DISEÑO DE EXPERIMENTOS	17
PRUEBAS DE HIPOTESIS	21
PRUEBA DE "t"	23
ANALISIS DE VARIANZA	26
DISEÑO ALEATORIO SIMPLE	27
DISEÑO EN BLOQUES AL AZAR	33
DISEÑO EN CUADRO LATINO	38
DISEÑO PARA DOS FACTORES	45
DESCRIPCION DEL SISTEMA ANOVA-MSC	55
CONCLUSIONES	63
APENDICE I. DIAGRAMAS DE FLUJO	65
BIBLIOGRAFIA	68

DESDE SIEMPRE, EN SU AFAN POR CONOCER MAS. EL HOMBRE HA EXPERIMENTADO, A VECES CON UN ENFOQUE NETAMENTE CIENTIFICO, OTRAS BASADO EN EL MAS PURO EMPIRISMO.

LO CIERTO ES QUE AL EXPERIMENTAR, SEA DE UNA O DE OTRA FORMA, SE PUEDEN ENCONTRAR FACTORES HASTA ENTONCES IGNORADOS O BIEN DESCARTAR EL O LOS QUE HASTA ESE MOMENTO ERAN IMPORTANTES.

LA ESTADISTICA JUEGA UN PAPEL MUY IMPORTANTE EN EL CURSO DE LA EXPERIMENTACION CIENTIFICA, CUANDO SE TIENE QUE DECIDIR SOBRE CIERTOS FACTORES.

EN UN DISEÑO EXPERIMENTAL SE PRUEBAN ESTADISTICAMENTE LOS DATOS. ES UNA MANERA CIENTIFICA DE ASUMIR ALGUNA DECISION.

POR EJEMPLO : SI ES FUNCIONAL O NO UN NUEVO PLAN DE ESTUDIOS, SI LA APLICACION DE UN FERTILIZANTE TIENE EFECTO SOBRE LOS CULTIVOS, SI LA EDAD INFLUYE EN LA ASIMILACION DE UN CURSO DETERMINADO O COMPROBAR EL EFECTO DE CIERTA DROGA.

ASUMIR ESTADISTICAMENTE UN PROBLEMA, IMPLICA QUIZA TENER MAYOR SEGURIDAD A LA HORA DE CONCLUIR.

SIN EMBARGO, TRAE CONSIGO ALGUNOS PROBLEMAS, COMO LO SON EL HECHO MISMO DE TENER QUE APRENDER ESTADISTICA, DE FAMILIARIZARSE CON CIERTOS PRINCIPIOS Y PROCEDIMIENTOS MATEMATICOS.

POR LO ANTERIOR, CONTINUAMENTE SE BUSCA DESPERTAR EL INTERES DEL ALUMNO POR APRENDER ESTADISTICA, PERO HACER QUE LOS TEMARIOS LE SEAN ACCESIBLES Y LOS SEPA UTILIZAR PARA ENFRENTAR OBJETIVAMENTE ALGUN PROBLEMA DADO.

DEBE SER CAPAZ DE MANIPULAR FORMULAS A LA VEZ QUE RELACIONARLAS CON UN PROBLEMA REAL O PROPUESTO.

OSCAR KENPTHORNE (1980), SUGIERE QUE LA ESTADISTICA NO ES LA MERA MANIPULACION DE DATOS EN FORMULAS, SINO QUE SIEMPRE SURGEN PREGUNTAS DERIVADAS DEL PROBLEMA ABORDADO :

- ¿CUANTAS MEDIDAS DEBEN HACERSE?
- ¿QUE TAMAÑO DE MUESTRA?
- ¿COMO EVITAR LA INFLUENCIA DE ALGUNA O ALGUNAS VARIABLES?
- ¿QUE ANALISIS ESTADISTICO EMPLEAR PARA EL TRATAMIENTO DE LOS DATOS?

SON ALGUNAS DE LAS PREGUNTAS QUE DEBEN HACERSE ANTES DE INICIAR EL EXPERIMENTO.

UN EXPERIMENTO BIEN PLANEADO NO REQUIERE DE GRANDES ARTIFICIOS MATEMATICOS.

EL ANALISIS ESTADISTICO QUE SE APLIQUE A LOS DATOS, UNA VEZ QUE SE HA CONCLUIDO EL EXPERIMENTO, SERA MUCHO MAS FACIL Y CONFIABLE SI, PREVIO AL EXPERIMENTO, SE OPTO POR APLICAR LAS TECNICAS DEL DISEÑO EXPERIMENTAL.

EL DISEÑO DEL EXPERIMENTO DEBE HACERSE, ANTES DE INICIAR ESTE ULTIMO, PRECISAMENTE PARA DEFINIR EL TIPO DE DATOS A TOMAR.

ES DECIR, QUE EL EXPERIMENTADOR TRAZA SU PLAN DE TRABAJO CON LA FINALIDAD DE AHORRAR TIEMPO Y DINERO.

PARA DAVIES (1954), UN BUEN DISEÑO EXPERIMENTAL ES AQUEL QUE PROPORCIONA LA INFORMACION REQUERIDA CON EL MINIMO ESFUERZO EXPERIMENTAL.

USANDO LA TEORIA MATEMATICA ES POSIBLE OBTENER MEDIDAS DE LA CANTIDAD DE INFORMACION REQUERIDA PARA EL ARREGLO EXPERIMENTAL.

SIN EMBARGO, SIEMPRE HAY QUE HACER BUEN USO DEL CONOCIMIENTO Y EXPERIENCIA PARA PLANEAR ADECUADAMENTE EL TRABAJO.

FINNEY (1960), SUGIERE QUE MEDIANTE EL DISEÑO DE EXPERIMENTOS :

- A) SE SELECCIONAN LOS TRATAMIENTOS A COMPARAR
- B) SE ESPECIFICAN LAS UNIDADES EXPERIMENTALES, A LAS QUE SE RAN APLICADOS LOS TRATAMIENTOS
- C) SE DEFINEN LAS REGLAS POR LAS QUE LOS TRATAMIENTOS SERAN ASIGNADOS A LAS UNIDADES EXPERIMENTALES
- D) SE ESPECIFICA QUE MEDIDAS DEBE HACERSE A CADA UNIDAD EXPERIMENTAL

JORDAN (1984), PROPONE UN MODELO CONCEPTUAL PARA PRESENTAR LAS TECNICAS ESTADISTICAS, DENTRO DE UN CONTEXTO ORGANIZACIONAL, DE ACUERDO A LO QUE SE CONOCE COMO PERSPECTIVA DE SISTEMAS.

LA FIGURA 1, TOMADA DE SU ARTICULO "A SYSTEMS PERSPECTIVE OF STATISTICS", NOS MUESTRA SU IDEA DE ABORDAR LA RESOLUCION DE UN PROBLEMA ESTADISTICO COMO UN PROCESO DE CUATRO FACES.

PARA ESTE AUTOR, EL METODO DE MUESTREO JUEGA UN PAPEL MUY IMPORTANTE Y DESTACA TAMBIEN LA UTILIDAD DE LAS COMPUTADORAS, PARA LA CAPTURA Y PROCESO DE LOS DATOS.

EN LA FIGURA 1 PUEDE OBSERVARSE QUE, POR UNA PARTE HAY ACCIONES PROPIAS DE UN DIRECTIVO Y OTRAS PROPIAS DEL RESPONSABLE EN ESTADISTICA.

ESTE ULTIMO ES ENCAUSADO POR EL PRIMERO, DURANTE LA PRIMERA FASE (DISEÑO DEL ESTUDIO).

LAS SIGUIENTES DOS FASES (COLECCION Y ANALISIS DE LOS DATOS), SON RESPONSABILIDAD EXCLUSIVA DEL ESTADISTICO.

DURANTE LA ULTIMA FASE (ESTUDIO DE IMPLEMENTACION), EL MISMO ESTADISTICO TIENE QUE PRESENTAR LOS RESULTADOS DE UNA MANERA CLARA Y LEGIBLE A LOS DIRECTIVOS, PARA QUE ESTOS TOMEN UNA DECISION Y ACTUEN CONVENIENTEMENTE.

RESPONSABILIDADES
DE LOS DIRECTORES

FASES/
PASOS

RESPONSABILIDADES
DEL ESTADÍSTICO

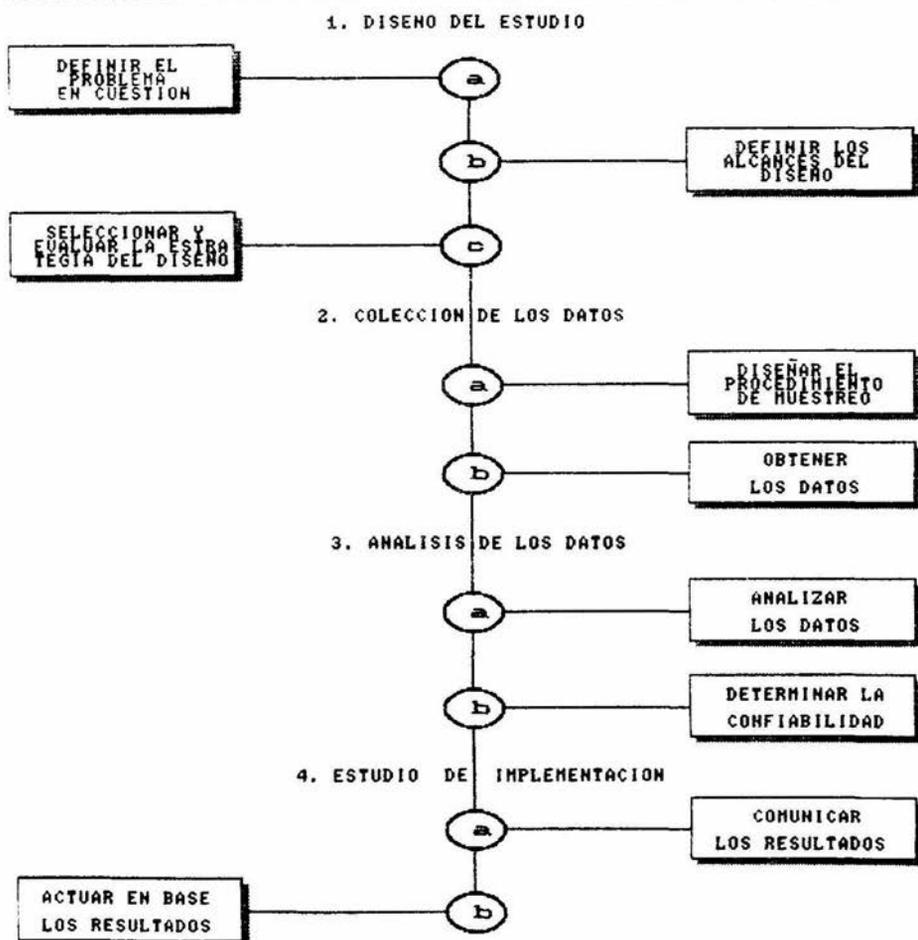


FIGURA 1. DIVISION DE RESPONSABILIDADES PARA EL PROCESO ESTADISTICO

ES OPORTUNO MENCIONAR QUE, PARA MUCHOS AUTORES, LA COMPUTADORA ES UNA HERRAMIENTA INDISPENSABLE PARA EL BUEN DESEMPEÑO ESTADISTICO.

DESDE SU APARICION HA TENIDO CRECIENTE Y HOY EN DIA CASI TOTAL ACEPTACION, EN CAMPOS TAN VARIADOS DE LA ADMINISTRACION, LA ENSEÑANZA Y LA INVESTIGACION.

ADEMAS DE SU EFICIENCIA Y VELOCIDAD PARA EL CALCULO ESTA EL PODER ALMACENAR Y MANIPULAR GRANDES CANTIDADES DE INFORMACION.

LA COMPUTADORA EN SUS POTENCIALIDADES ACTUALES PERMITE ANALIZAR FENOMENOS, QUE DE OTRA MANERA SERIA IMPOSIBLE ABORDAR.

ES UN IMPORTANTE AUXILIAR EN EL AULA, AL GRADO DE QUE SE PODRIA INDIVIDUALIZAR LA ENSEÑANZA (LEHMAR, 1985).

"DISEÑO ASISTIDO POR COMPUTADORA", "INTELIGENCIA ARTIFICIAL", "PROGRAMAS TUTORIALES", "PAQUETES DE APLICACION ADMINISTRATIVA", SON TERMINOS MUY REFERIDOS EN LA ACTUALIDAD.

LA SIMULACION EN PARTICULAR OFRECE MUCHO AL PROFESOR.

TOPICOS TAN VARIADOS COMO GENETICA, CONTAMINACION, PESQUERIAS, CONTROL DE PLAGAS Y TANTOS OTROS, NO SOLO DE LA BIOLOGIA, SE PUEDEN ABORDAR CONVENIENTEMENTE EN LA CLASE A TRAVEZ DE LOS MODELOS DE SIMULACION.

YA QUE EN LA COMPUTADORA SE PUEDE COMPACTAR O EXPANDER EL TIEMPO, ES FACTIBLE INCORPORAR AL ALUMNO EN LA TOMA DE DECISIONES.

AL PODER, ESTE, DE UNA MANERA ALMENOS TEORICA, VARIAR LIBREMENTE LOS PARAMETROS DE UNA POBLACION Y OBSERVAR DE INMEDIATO LOS EFECTOS QUE ESTO OCASIONA, SE ESTARA REFORZANDO ASI MISMO EN LA COMPRENSION Y SOLUCION DE PROBLEMAS REALES.

EN OTRO ANGULO LOS PROGRAMAS TUTORIALES AYUDAN AL ESTUDIANTE A DECIDIR QUE PRUEBA APLICAR Y LOS DE APLICACION PERMITEN EVADIR LOS TEDIOSOS CALCULOS, QUE MUCHAS VECES OBSTACULIZAN QUE EL ALUMNO CONCLUYA Y TOME UNA DECISION.

PODER ALMACENAR GRUPOS DE DATOS Y MANIPULARLOS CON RELATIVA FACILIDAD, APLICARLES DETERMINADA PRUEBA ESTADISTICA Y OBSERVAR DE INMEDIATO LOS RESULTADOS, SIN DUDA ALGUNA QUE REAFIRMARA CONCEPTOS, AL CENTRAR LA ATENCION EN EL ANALISIS DE RESULTADOS.

DE HECHO LA COMPUTADORA ES UNA HERRAMIENTA INDISPENSABLE PARA LA ESTADISTICA, QUIZA TANTO COMO LO ES EL MICROSCOPIO PARA LA MICROBIOLOGIA (BUTLER Y DWASS, 1976).

COLLIS (1983), CONCUERDA CON LO HASTA AHORA DICHO, MENCIONANDO QUE LA COMPUTADORA PERMITE :

- A) GENERAR FACILMENTE GRAFICOS CLAROS Y ATRACTIVOS
- B) ILUSTRAR CONCEPTOS
- C) EVADIR CALCULOS TEDIOSOS

WEGMAN Y GERE (1972), SUGIEREN QUE LA COMPUTADORA ES UNA INVITACION PARA EL ESTUDIANTE, RESULTANDO ESTE GRANDEMENTE MOTIVADO.

SIN EMBARGO, NO SIEMPRE EL USO POTENCIAL DE LA COMPUTADORA CON FINES EDUCATIVOS RESULTA APROPIADO.

HAY ALGUNAS DESVENTAJAS, DERIVADAS GENERALMENTE DE UNA MALA PLANEACION :

- A) LA SIMULACION, POR UNA PARTE, DEBE SER APROPIADA A LOS EDUCANDOS Y MANEJADA SIEMPRE COMO UN REFUERZO DE LABORATORIO, SIN QUE SU USO IMPLIQUE ABANDONAR LOS TRADICIONALES EJERCICIOS DE CLASE.

B) POR OTRO LADO, TANTO LOS PROGRAMAS DE SIMULACION COMO LOS DE APLICACION ESTADISTICA, QUE COMUNMENTE SE COMERCIALIZAN, NO SIEMPRE SON ACORDES A LAS NECESIDADES CURRICULARES.

C) TAMBIEN HABRA QUE ADVERTIR QUE EL PROGRAMA NO DEBE HACER TODO POR EL ALUMNO.

ESTE DEBE SER CAPAZ DE EJECUTAR LOS CALCULOS MANUALMENTE.

LEHMAN (1983), ACONSEJA QUE PARA QUE LA SIMULACION EN CLASE SEA FRUCTIFERA, SE DEBEN SELECCIONAR TOPICOS ADECUADOS, ADEMAS DE OBTENER LA MAYOR INFORMACION POSIBLE.

POR OTRA PARTE SE PUEDE INTRODUCIR AL ALUMNO EN ALGUN LENGUAJE DE PROGRAMACION.

ASI, ENTRE ALUMNOS Y PROFESORES GENERAR SUS PROPIOS PROGRAMAS.

PARA DEAN (1982), LOS PROFESORES DEBIERAN TENER CONOCIMIENTOS DE UN LENGUAJE DE PROGRAMACION.

EL BASIC ES UN LENGUAJE MUY SENCILLO, PARECIDO EN SU SINTAXIS A LA GRAMATICA DEL IDIOMA INGLES Y PRESENTE EN CASI TODOS LOS EQUIPOS DE COMPUTO, POR LO QUE PUEDE SER LA VIA DE ACCESO A LA PROGRAMACION.

ELABORAR PROGRAMAS, COMO UN REFUERZO PARA LA CLASE, OBTIENE IMPLICAR GASTOS EN TIEMPO Y DINERO PERO REPORTA GRANDES BENEFICIOS.

CREAR UN PROGRAMA DE COMPUTADORA, IMPLICA PONER MUCHO CUIDADO A LA HORA DE DEFINIR LOS PASOS Y EL ORDEN DE LAS OPERACIONES, Y ESTO, DE ALGUN MODO, SE TRADUCE EN CONOCIMIENTO BIEN LOGRADO .

EN RELACION AL PAPEL QUE LAS COMPUTADORAS JUEGAN EN LA ENSE-

ÑANZA DEL ANALISIS ESTADISTICO DE DATOS, CURTS Y RAZCON (1988), OPINAN QUE LOS PROGRAMAS DE COMPUTO ESTADISTICO HAN ADQUIRIDO UNA ENORME IMPORTANCIA, DEBIDO A LA IMPLEMENTACION DE SUBROUTINAS PARA MICROCOMPUTADORAS, CON AMPLIAS POSIBILIDADES DE CALCULO Y GRAFICACION.

LO ANTERIOR HA PERMITIDO QUE EL ANALISIS ESTADISTICO DE LOS DATOS PUEDA HACERSE DE MANERA FACIL Y A BAJO COSTO.

ESTOS DOS AUTORES HAN REFLEXIONADO EN LOS ASPECTOS DEL ANALISIS ESTADISTICO QUE SE HAN VISTO INFLUENCIADOS FUERTEMENTE POR EL USO DE LAS COMPUTADORAS Y QUE A SU JUICIO SON :

1) COMPRENSION DEL CONTEXTO DE LOS DATOS.

EL USO DE LA COMPUTADORA EN LOS CURSOS DE ESTADISTICA PERMITE CONSIDERAR UN MAYOR NUMERO DE PROBLEMAS REALES.

LOS DATOS NO TIENEN QUE SER FICTICIOS, PRESELECCIONADOS O MODIFICADOS, PARA FACILITAR LA REALIZACION DE LOS CALCULOS, NI PARA SATISFACER LOS SUPUESTOS DE MUESTRAS ALEATORIAS PROVENIENTES DE DISTRIBUCIONES GAUSSIANAS.

DE ESTA MANERA SE PUEDE UTILIZAR UNA GRAN VARIEDAD DE LOTES DE DATOS REALES RELACIONADOS CON LAS AREAS DE INTERES A LOS ALUMNOS.

LO ANTERIOR HA SIDO, EN EXPERIENCIA DE LOS AUTORES, UNA GRAN VENTAJA PEDAGOGICA, YA QUE PERMITE ENSEÑARLES UNA DISCIPLINA ENFOCADA, DESDE UN PUNTO DE VISTA PRACTICO Y REALISTA, SATISFACIENDO ASI SUS INTERESES PERSONALES.

2) VERIFICACION DE LOS DATOS.

EL USO DE LAS COMPUTADORAS FACILITA EN GRAN MEDIDA LA VERIFICACION DE LOS DATOS, GRACIAS A LA VELOCIDAD CON QUE SE PUEDEN IMPRIMIR HISTOGRAMAS, DIAGRAMAS DE TALLO Y HOJA, GRAFICOS DE RESIDUALES, Y OTROS.

EJEMPLOS DE DIVERSOS DIAGRAMAS UNIVARIADOS, ELABORADOS POR COMPUTADORA, PUEDEN SER CONSULTADOS EN CURTS (1987).

3) VISUALIZACION DE LOS DATOS.

ES EN ESTE PUNTO DONDE LAS COMPUTADORAS, DE ACUERDO CON ESTOS AUTORES, MUESTRAN SU GRAN EXELEN-CIA.

ELLOS MISMOS HAN AFIRMADO QUE: "NO EXISTE ESTADISTICO MAS PODEROSO QUE UNA GRAFICA BIEN SELECCIONADA".

4) TRANSFERENCIA DE LOS DATOS.

LAS COMPUTADORAS FACILITAN LA REEXPRESION DE LOS DATOS, PARA SER POSTERIORMENTE GRAFICADOS Y RECALCULADOS.

LA REEXPRESION DE LOS DATOS, UTILIZANDO PROGRAMAS DE COMPUTO, HACEN QUE ESTE PROCESO SEA SUMAMENTE SENCILLO Y NATURAL.

CONVIENE AGREGAR QUE ESTOS AUTORES COMENTAN QUE LA REEXPRESION DE LOS DATOS SE DEBE A QUE ESTOS ULTIMOS NO SIEMPRE ESTAN DISPONIBLES EN UNA FORMA FACIL DE RESUMIR, VISUALIZAR Y RAZONAR, Y EN ALGUNAS OCASIONES LA FORMA NATURAL DE LOS DATOS NO ES LA MAS CONVENIENTE PARA EFECTUAR PRUEBAS DE HIPOTESIS O PARA REALIZAR EL PROCESO DE INFERENCIA ESTADISTICA.

LA SOLUCION DE TALES PROBLEMAS CONSISTE EN EXPRESAR MATEMATICAMENTE LA ESCALA ORIGINAL DE LOS DATOS EN OTRA QUE FACILITE SU INTERPRETACION.

5) REALIZACION DE LOS CALCULOS.

LAS COMPUTADORAS FACILITAN ENORMEMENTE LAS OPERACIONES NUMERICAS Y PERMITEN ALMACENAR Y EDITAR DATOS CRUDOS, RESULTADOS Y GRAFICAS.

EL ANALISIS DE VARIANZA, SIN DUDA ALGUNA, ES UNA HERRAMIENTA MUY UTILIZADA POR EL BIÓLOGO, PERO NO SIEMPRE ES BIEN ASIMILADA POR EL ESTUDIANTE DE LICENCIATURA.

SOBRETODO EN LOS DISEÑOS MAS COMPLICADOS, COMO LO ES EL DE 3 FACTORES, EL ALUMNO SE PIERDE EN LOS CALCULOS Y LLEGA A RESULTARLE TEDIOSO, POR LO QUE SU ENSEÑANZA, DE UNA MANERA TRADICIONAL Y AUN MAS SU APLICACION, SE DIFICULTA GRANDEMENTE.

EN RELACION A LOS PROBLEMAS DE ENSEÑANZA DE LOS METODOS CUANTITATIVOS EN BIOLOGIA, CURTS (1988) HA SEÑALADO LO SIGUIENTE :

1) LA IMPORTANCIA DE ESTOS METODOS RESIDE EN EL VALOR DE LOS MODELOS COMO GENERADORES DE IDEAS.

ES NECESARIO VINCULAR LOS ASPECTOS CUANTITATIVOS A LA REALIDAD BIOLOGICA A TRAVES DEL CONCEPTO DE MODELO.

POR EJEMPLO, EL MODELO :

$Y_i = \mu + at + e_i$ DEBE ESTAR SIEMPRE LIGADO AL PROCESO REAL QUE SE DESEA ANALIZAR.

2) EL CARACTER DE LA BIOLOGIA ES ESENCIALMENTE INTERDISCIPLINARIO, YA QUE LOS MEDIOS UTILIZADOS PARA COMPRENDER LOS PROCESOS BIOLÓGICOS, TIENDEN A VOLVERSE IDENTICOS EN LAS DIVERSAS RAMAS DE LA CIENCIA.

POR EJEMPLO, LAS DIVERSAS APLICACIONES DEL ANALISIS DE VARIANZA EN BIOLOGIA, PSICOLOGIA, SOCIOLOGIA Y OTRAS CIENCIAS.

3) LA DIFICULTAD DE HABER IMPLEMENTADO, CON ÉXITO, PROGRAMAS DE MATEMÁTICAS EN BIOLOGIA, SE DEBE A QUE UNA GRAN MAYORÍA DE LOS ESTUDIANTES DESCONOCE LOS ASPECTOS BÁSICOS DEL MODELADO MATEMÁTICO Y SU RELACION CON EL PROCESO DE APRENDIZAJE DEL CONOCIMIENTO CIENTÍFICO.

POR CONSIDERAR QUE ESTE ÚLTIMO PUNTO ES FUNDAMENTAL Y COMO BIEN AFIRMA ESTE AUTOR "QUE ES UNA FALACIA PENSAR" QUE LOS CONCEPTOS DEL DISEÑO EXPERIMENTAL SE APRENDERÁN SI PRIMERO SE LE EXIJE AL ALUMNO CALCULAR SUMA DE CUADRADOS, SE ILUSTRÓ A TRAVÉS DE UN DIAGRAMA, FIGURA # 2, LA RELACION DEL PROCESO DE MODELADO MATEMÁTICO CON EL DEL APRENDIZAJE DEL CONOCIMIENTO CIENTÍFICO.

EN DICHA FIGURA SE ILUSTRÓ QUE LA MODELACION MATEMATICA SIGUE DOS CAMINOS :

- A) EL DE LA EXPERIMENTACION U OBSERVACION DE CAMPO
- B) EL DE LA METODOLOGIA CONCEPTUAL

EL PRIMERO SE RELACIONA EVIDENTEMENTE CON LA MODELACION ESTADISTICA Y EL SEGUNDO CON LA MODELACION TEORICA.

ES ASI QUE, EN ESTE DIAGRAMA, SE REFUERZA LA IDEA QUE PARA LA OBTENCION DE UN MODELO MATEMATICO, COMO POR EJEMPLO LA FUNCION POTENCIAL ($Y = (bx)^k$), SE PUEDEN SEGUIR DOS CAMINOS :

A) LA OBTENCION DE DATOS QUE SE AJUSTAN, POR LA TECNICA DE MINIMOS CUADRADOS, A LA FUNCION POTENCIAL

B) SUPONIENDO EL PRINCIPIO DE EQUILIBRIO DE MASAS, DERIVAR LA ECUACION DIFERENCIAL :

$$dy / dx = k(y/x)$$

EN SINTESIS, COMO SEÑALA ESTE AUTOR, ES IMPORTANTE DEMOSTRAR A LOS ALUMNOS QUE LAS MATEMATICAS SON UNA FORMA DE ORDENAR Y EXPRESAR LAS OPINIONES ACERCA DE COMO ESTA ORGANIZADO EL MUNDO.

APARTIR DE LAS IDEAS ANTERIORES, Y DESPUES DE HACER UNA REVISION MINUCIOSA DEL TEMA, SE REA-

LIZO EL PRESENTE TRABAJO, COMO PROYECTO DE TESIS PARA TITULACION EN LA CARRERA DE BIÓLOGO.

EL PROPOSITO ES PLASMAR, DE UNA MANERA CLARA, LOS PRINCIPIOS DEL ANALISIS DE VARIANZA, VALIENDOSE DE LA COMPUTADORA, COMO IMPORTANTISIMO RECURSO DE APOYO.

EL PRESENTE DOCUMENTO PUEDE SERVIR COMO MANUAL DE CLASE, CON LA VENTAJA DE QUE, PARALELAMENTE, SE ELABORO UN SISTEMA EN COMPUTADORA, PARA LA APLICACION DEL ANALISIS DE VARIANZA.

TANTO EL PROGRAMA, COMO LAS NOTAS PUEDEN SER VALIOSOS RECURSOS PARA "LA ENSEANZA DEL ANOVA EN BIOLOGIA".

POR TANTO LOS OBJETIVOS SON :

- A) QUE EL ALUMNO COMPRENDA EL DISEÑO DE EXPERIMENTOS
- B) DOTARLO DE UNA HERRAMIENTA PRACTICA, PARA EL TRATAMIENTO DE DATOS, POR MEDIO DEL ANALISIS DE VARIANZA
- C) INTRODUCIRLO AL AREA COMPUTO-ESTADISTICA

NO OBSTANTE QUE HAYA ABUNDANTE PAGUETERIA ESTADISTICA EN EL MERCADO, ES PRECISO ACLARAR QUE LOS DISEÑOS QUE SE EXPONEN, EN ESTE TRABAJO, SE ADECUARON A LO EXPUESTO EN EL CURSO "DISEÑO DE EXPERIMENTOS", QUE SE IMPARTE EN LA ENEP IZTACALA.

POR LO CUAL AL MENOS RESULTARA OBJETIVO A DICHO CURSO.

EL TRABAJO ESCRITO COMPRENDE LAS SIGUIENTES PARTES :

- 1) INTRODUCCION AL DISEÑO DE EXPERIMENTOS.

- 2) COMPARACION DE SOLO DOS MUESTRAS (PRUEBA DE "t").

- 3) COMPARACION DE MAS DE DOS MUESTRAS (ANALISIS DE VARIANZA).

- 4) GUIA PARA LA UTILIZACION DE LOS PROGRAMAS.

EN SU TOTALIDAD SE CAPTO EN MICROCOMPUTADORA IBM-COMPATIBLE, O PC-COMPATIBLE COMO MAS COMUNEMENTE SE LES LLAMA, UTILIZANDO EL PROCESADOR DE PALABRAS "WORD-PERFECT", VERSION 4.1.

LOS PROGRAMAS POR SU PARTE FUERON DESARROLLADOS EN EL MISMO EQUIPO, UTILIZANDO EL LENGUAJE BASIC ("GWBASIC", VERSION 3.2).

EL INDICE, CARATULA Y FIGURAS CON EL EDITOR DE GRAFICOS "FLOW" VERSION II+.

TANTO LA CAPTURA DEL TEXTO COMO DE LOS PROGRAMAS, SE INICIO EN EQUIPO CROMENDO C-10, TENIENDO A LA VEZ LA OPCION DE HACERLO EN COMMODORE 64 O EN APPLE IIe.

SIN EMBARGO, DADA LA GRAN POPULARIDAD QUE HAN LOGRADO LAS PC-COMPATIBLES, SE OPTO POR HACERLO EN ESTA MAQUINA, PRETENDIENDO ASI LLEGAR A UN MAYOR NUMERO DE USUARIOS.

LA ESTRUCTURA DEL TRABAJO, TANTO ESCRITO COMO DE LOS PROGRAMAS ("SOFTWARE"), SE PROCURO FUERA LO MAS MODULAR POSIBLE.

ESTO CON LA FINALIDAD DE AMPLIARLO A FUTURO, ENTENDIENDO OTROS DISEÑOS DEL ANOVA, NO CONSIDERADOS AQUI.

O INCLUSO OTROS RECURSOS ESTADISTICOS IGUALMENTE IMPORTANTES COMO LO SON "EL ANALISIS DE REGRESION" Y "EL ANALISIS DE CUMULOS".

LOS PROGRAMAS SE HICIERON
"PENSANDO EN EL USUARIO".

ES DECIR, QUE SU MANEJO ES
SENCILLO Y CLARA SU PRESENTACION,
PROCURANDO EVITAR, EN LO MAS PO-
SIBLE, CIERTAS EVENTUALIDADES
DE OTROS PROGRAMAS.

SE CONFIA ENTONCES QUE EL
TRABAJO CUMPLIRA ADECUADAMENTE
SUS PROPOSITOS DE ENSEÑANZA,
ADEMAS DE QUE LOS PROGRAMAS ESTAN
LISTOS PARA APLICACIONES REALES.

EL DISEÑO DE EXPERIMENTOS ES UNA TÉCNICA MUY APLICADA EN ESTUDIOS BIOLÓGICOS, MÉDICOS Y AGROPECUARIOS.

SE DESPRENDE DE LAS IDEAS INICIALES DE R. A. FISHER, AMPLIADAS POSTERIORMENTE POR D. FINNEY, G.W. SNEDECOR Y W.S. COCHRAN, ENTRE OTROS (MENDEZ, 1981) .

EN EL DISEÑO DE EXPERIMENTOS SE INVOLUCRAN LOS PRINCIPIOS DE CONTROL, ALEATORIZACIÓN Y COMPARACIÓN (SPECTOR, 1981) .

SU IMPORTANCIA ESTIBA EN QUE AYUDA A REDUCIR EL ERROR EXPERIMENTAL, DURANTE LA COLECCIÓN DE LOS DATOS (HICKS, 1973) .

EL PROCESO DE ALEATORIZACIÓN TIENDE A PROMEDIAR EL EFECTO DE AQUELLAS VARIABLES QUE NO SE PUEDEN CONTROLAR DURANTE EL EXPERIMENTO.

EL MODELO MATEMÁTICO QUE DESCRIBE UN EXPERIMENTO, EN QUE SE APLICA EL PROCESO DE ALEATORIZACIÓN, IMPLICA, POR LO TANTO, QUE LA VARIABLE DE RESPUESTA ESTA EN FUNCIÓN DE LOS FACTORES ESTUDIADOS, PERO ADEMÁS DE CIERTAS RESTRICCIONES IMPUESTAS AL EXPERIMENTO , DURANTE DICHO PROCESO DE ALEATORIZACIÓN.

EN BASE A CIERTOS FACTORES, CONSIDERADOS COMO IMPORTANTES, SE ELIGE UN DISEÑO ADECUADO AL EXPERIMENTO.

POSTERIORMENTE AL CHECAR LOS RESULTADOS, EL MODELO MISMO Y, POR ENDE EL DISEÑO, SON EVALUADOS.

DE ACUERDO CON ANDERSON Y Mc. LEAN (1974), EN UN EXPERIMENTO CIENTÍFICO SE DEBEN CONSIDERAR LOS SIGUIENTES PASOS, QUE CONFORMAN EL REFERIDO "DISEÑO EXPERIMENTAL" :

- A) IDENTIFICAR, FORMULAR Y DELIMITAR EL PROBLEMA
- B) DESGLOZARLO EN FACTORES Y NIVELES
- C) ESPECIFICAR LAS VARIABLES A UTILIZAR
- D) DEFINIR EL ESPACIO INFERENCIAL
- E) SELECCIONAR ALEATORIAMENTE LAS UNIDADES DE MUESTREO
- F) ASIGNAR LOS TRATAMIENTOS A LAS UNIDADES DE MUESTREO
- G) DEFINIR EL TIPO DE ANÁLISIS A SEGUIR
- H) COLECTAR LOS DATOS
- I) APLICAR LOS CÁLCULOS
- J) ANALIZAR LOS RESULTADOS
- K) CONCLUIR
- M) IMPLEMENTAR

PRIMERAMENTE DEFINIR ADECUADAMENTE EL PROBLEMA Y LOS LÍMITES A QUE SE PRETENDE LLEVAR LA INFERENCIA.

ES DECIR, INICIALMENTE PLANTEAR LOS ALCANCES DEL EXPERIMENTO, YA QUE DE ELLO SE DESPRENDE EL NÚMERO DE MUESTRAS Y FACTORES A CONSIDERAR.

LA UNIDAD EXPERIMENTAL (MATERIAL EXPERIMENTAL QUE RECIBE LA APLICACIÓN DE UN TRATAMIENTO ESPECÍFICO) DEBE ENTONCES SER REPRESENTATIVA DEL ESPACIO INFERENCIAL DESEADO.

POR EJEMPLO, AL INVESTIGAR LOS EFECTOS DE VARIOS TIPOS DE FERTILIZANTES SOBRE PLANTAS CULTIVADAS EN MACETAS, LA UNIDAD EXPERIMENTAL SERÁ PRECISAMENTE UNA MACETA.

LAS INFERENCIAS PODRAN HACERSE EXTENSIVAS A LAS CONDICIONES DE CAMPO DE LA POBLACION ESTUDIADA (ESPACIO INFERENCIAL), TODA VEZ QUE EL NUMERO DE ESTAS UNIDADES SEA EL ADECUADO Y ALEATORIA SU ELECCION, PARA APLICARLES ALGUN TIPO PARTICULAR DE FERTILIZANTE.

ESTO ULTIMO ES MUY IMPORTANTE, SOBRETODO QUE NO SIEMPRE SE PUEDEN LOGRAR CONDICIONES HOMOGENEAS PARA LAS UNIDADES EXPERIMENTALES Y ALGUN RESULTADO PUEDE SER INFLUENCIADO POR ALGUN FACTOR DESCONOCIDO (EN ESTE CASO, ESE FACTOR NO CONSIDERADO PODRIA SER EL SUELO).

LA ALEATORIZACION NUNCA ELIMINA LA VARIACION CAUSADA POR VARIABLES EXTRAÑAS.

SINO, MAS BIEN, CONCEDE IGUAL MARGEN DE ERROR AL APLICAR LOS TRATAMIENTOS.

A PROPOSITO DE ESTE ULTIMO CONCEPTO, CONVIENE ACLARAR CIERTA CONFUSION DESPRENDIDA DURANTE LA ENSEÑANZA DEL DISEÑO DE EXPERIMENTOS.

SE CREE, QUE EL SOLO HECHO DE ASIGNAR ALEATORIAMENTE LAS UNIDADES EXPERIMENTALES A LOS TRATAMIENTOS SATISFACE EL PROCESO DE ALEATORIZACION.

ESTO ULTIMO NO ES ASI, Y HAY QUE ENFATIZAR EN CLASE, COMO LO SEÑALA CURTS (1986), LA NECESIDAD DE CRUZAR EL PROCESO DE ALEATORIZACION TANTO DE LOS SUJETOS COMO DE LOS NIVELES DE TRATAMIENTOS.

EN LA FIGURA # 3 SE ILUSTRAS, CON UN EJEMPLO, UNA FORMA MUY SENCILLA DE ALEATORIZAR.

HASTA ESTE MOMENTO, CUANDO SE HAN CONSIDERADO LOS FACTORES IMPORTANTES, SE ELIGE EL MODELO ADECUADO AL EXPERIMENTO, MISMO QUE POSTERIORMENTE SERA EVALUADO.

AUNQUE LA COLECCION DE DATOS PAREZCA DE LO MAS SENCILLO, DEBE TOMARSE EN CUENTA, PROCURANDO QUE ESTOS SEAN CLAROS, MANEJABLES Y QUE NO SE PRESTEN A FALSAS INTERPRETACIONES.

ADEMAS TAMBIEN DEBEN GUARDARSE EN UN MEDIO ADECUADO, QUE FACILITE SU POSTERIOR MANIPULEO.

NATURALMENTE EL ANALISIS DE LOS DATOS DEPENDE DEL DISEÑO ELEGIDO PARA EL EXPERIMENTO.

SE PUEDE OPTAR POR EMPLEAR ALGUN PROGRAMA DE COMPUTADORA ACORDE CON ESE DISEÑO.

EN BASE A LOS CALCULOS, RESUMIDOS EN LA TABLA DE ANOVA, SE PROCEDE AL ANALISIS DE RESULTADOS.

SE RECOMIENDA ACOMPAÑAR ESTO CON LA REPRESENTACION GRAFICA DE LOS DATOS, E INCLUSO HACER COMPARACIONES ENTRE LAS DIFERENTES MUESTRAS (COMPARACIONES MULTIPLES).

TODOS LOS PASOS SON IMPORTANTES Y DE CUIDAR QUE SE CUMPLAN ADECUADAMENTE, PODEMOS LOGRAR CONCLUSIONES MAS CONVINCENTES, Y ASI DESPRENDER DEL EXPERIMENTO RECOMENDACIONES PRACTICAS Y CERTERAS.

PARA CONCLUIR ESTA BREVE REVISION, DEL DISEÑO EXPERIMENTAL, EN LA FIGURA # 4 SE MUESTRA UN DIAGRAMA DE FLUJO, QUE PERMITE SELECCIONAR UNO DE CUATRO POSIBLES MODELOS DEL ANALISIS DE VARIANZA.

NATURALMENTE ESTOS MODELOS CONFORMAN EL SISTEMA DESARROLLADO Y SON LOS QUE SE DISCUTEN EN ESTE TRABAJO.

SE RECOMIENDA REVISAR ANTES EL APENDICE I, DONDE SE HACE UNA INTRODUCCION AL MANEJO DE LOS DIAGRAMAS DE FLUJO, ASI COMO UNA BREVE EXPOSICION DE TERMINOS ASOCIADOS AL DISEÑO DE EXPERIMENTOS.

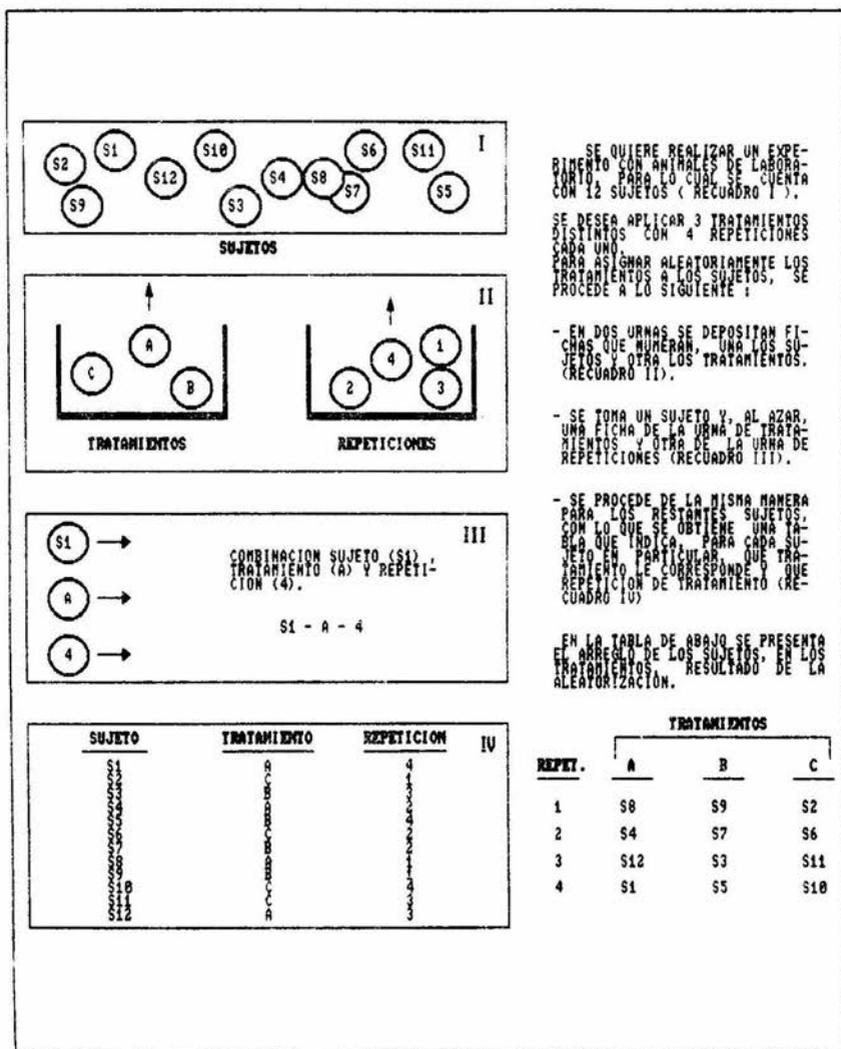


FIGURA 3. UN METODO SENCILLO DE ALEATORIZACION

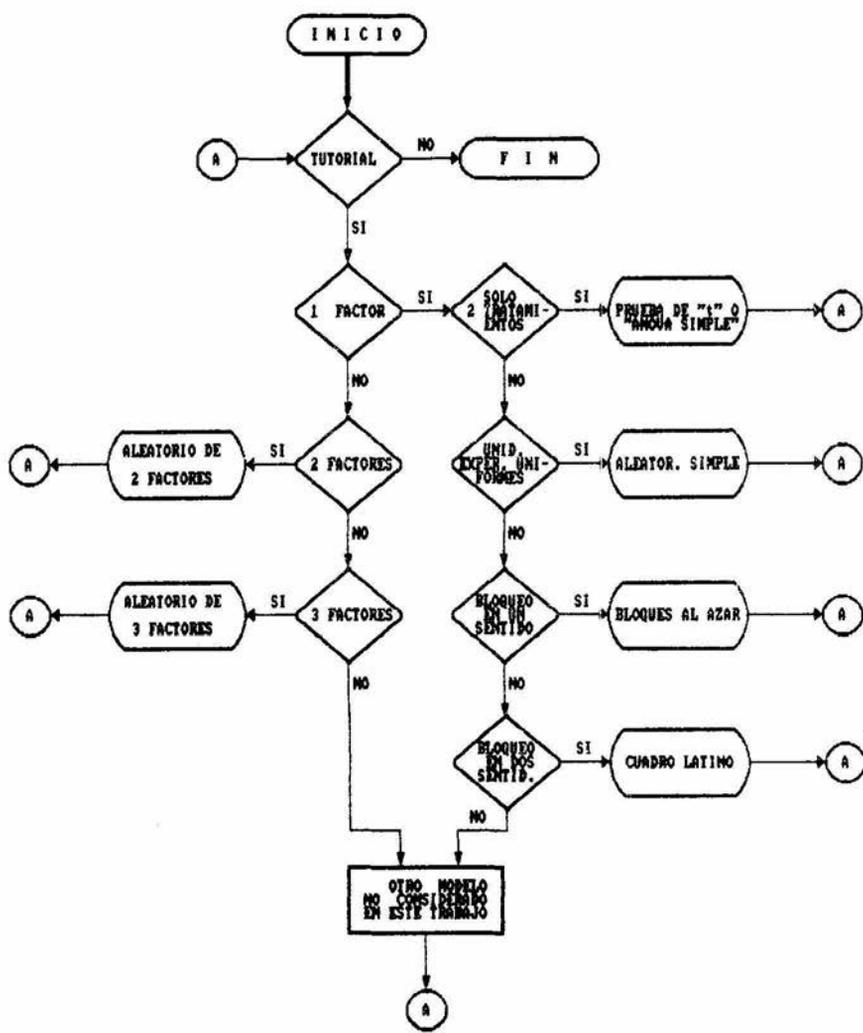


FIGURA 4. DIAGRAMA DE FLUJO QUE PERMITE SELECCIONAR -
ALGUNO DE LOS DISEÑOS MAS COMUNES DE ANOVA .

PRUEBAS DE HIPOTESIS

UNA HIPOTESIS PUEDE DEFINIRSE COMO UNA AFIRMACION ACERCA DE UNA O MAS POBLACIONES.

EL PROPOSITO DE LAS PRUEBAS DE HIPOTESIS ES AYUDAR AL INVESTIGADOR A TOMAR UNA DECISION REFERENTE A UNA POBLACION DETERMINADA, PARA LO CUAL SE EXAMINA UNA MUESTRA REFERENTE A ESA POBLACION.

SE CONTRASTAN LOS RESULTADOS EXPERIMENTALES O VALORES OBSERVADOS, CON CIERTOS VALORES TEORICOS O VALORES ESPERADOS.

SE PLANTEAN LAS HIPOTESIS DE TRABAJO, PARTIENDO DE LA NULIDAD DE DIFERENCIAS ENTRE VALORES OBSERVADOS Y VALORES ESPERADOS.

ESTA ULTIMA ACEVERACION SE CONOCE COMO HIPOTESIS NULA Y SE SIMBOLIZA POR "H₀".

COMO CONTRAPARTE SE PLANTEA LA HIPOTESIS ALTERNATIVA O DE NO NULIDAD DE DIFERENCIAS, QUE SE SIMBOLIZA POR "H_a". EN "H_a." SE ASUME QUE EL EXPERIMENTO TUBO EFECTO.

EL PLANTEAMINETO DE LAS HIPOTESIS SERIA COMO SIGUE :

H ₀ : $\mu_1 = \mu_2 = \dots \mu_n$	TODAS LAS MEDIAS DE LAS MUES TRAS SON IGUALES.
H _a : $\mu_1 \diamond \mu_2 \diamond \dots \mu_n$	NO TODAS LAS MEDIAS DE LAS MUESTRAS SON IGUALES.

LA ESTADISTICA PONE A NUESTRA DISPOSICION UNA SERIE DE HERRAMIENTAS QUE NOS PERMITEN EVALUAR HIPOTESIS COMO LAS ANTERIORMENTE PLANTEADAS.

ES DECIR, DAR UN CIERTO PESO MATEMATICO A NUESTRA DECISION.

EN LA REGLA DE DECISION SE RECHAZA LA HIPOTESIS NULA SI EL VALOR DEL ESTADISTICO DE PRUEBA, CALCULADO A PARTIR DE LA MUESTRA, CAE EN LA REGION DE RECHAZO DE LA GRAFICA DE DISTRIBUCION DEL ESTADISTICO.

POR EL CONTRARIO, SE ACEPTA LA HIPOTESIS NULA SI EL VALOR DEL ESTADISTICO DE PRUEBA CAE EN LA REGION DE ACEPTACION.

LAS REGIONES DE ACEPTACION Y RECHAZO DE LA GRAFICA DE DISTRIBUCION DEL ESTADISTICO, SE DEFINEN EN BASE AL NIVEL DE SIGNIFICANCIA ELEGIDO (VALOR " α ").

UN VALOR DEL ESTADISTICO DE PRUEBA QUE CAE EN LA REGION DE RECHAZO, SE DICE QUE ES SIGNIFICATIVO Y SIGNIFICATIVAS LAS DIFERENCIAS ENTRE LAS POBLACIONES IMPLICADAS.

A UN " α " GRANDE MAYOR SERA LA REGION DE RECHAZO DE LA CURVA DE DISTRIBUCION DEL ESTADISTICO, Y ENTONCES SE TENDRA CIERTA PROBABILIDAD DE RECHAZAR UNA HIPOTESIS NULA VERDADERA (ERROR TIPO I).

A LA INVERSA, UN " α " PEQUEÑO IMPLICA UNA REGION DE RECHAZO TAMBIEN PEQUEÑA, Y ENTONCES SE PODRIA ACEPTAR UNA HIPOTESIS NULA FALSA (ERROR TIPO II).

CON LA FINALIDAD DE HACER PEQUEÑA LA PROBABILIDAD DE RECHAZAR UNA HIPOTESIS NULA VERDADERA, SE HA CONVENIDO EMPLEAR LOS VALORES " α " DE 0.01 , 0.05 Y 0.1 .

PRUEBA "t" DE STUDENT

EL DISEÑO EXPERIMENTAL MAS SENCILLO, CORRESPONDIENTE A LA COMPARACION ESTADISTICA DE UN LOTE EXPERIMENTAL Y UN LOTE CONTROL (O DICHO DE OTRA MANERA, LA COMPARACION DE SOLO DOS MUESTRAS), ES EL REFERIDO COMO PRUEBA "t" DE STUDENT.

SE PUEDE ESTAR INTEREZADO EN COMPARAR MUESTRAS DE DIFERENTE INDOLE.

POR EJEMPLO, SABER SI HAY DIFERENCIA EN EL PESO DE CABRITOS DE DOS RAZAS DIFERENTES.

IMAGINEMOS QUE LOS PESOS REGISTRADOS AL MES DE NACIDOS FUERAN LOS SIGUIENTES :

RAZA	OBSERVACIONES									
1	23	43	24	19	29	47	6	14	29	26
2	17	14	10	9	3	18	3	19	11	

PLANTEAMOS NUESTRAS HIPOTESIS :

$H_0 : \mu_1 = \mu_2$ NO HAY DIFERENCIA EN EL PESO DE CABRITOS DE AMBAS RAZAS.

$H_a : \mu_1 \neq \mu_2$ SI HAY DIFERENCIA.

EN EL EJEMPLO PLANTEADO Y EN GENERAL CUANDO SE COMPARAN SOLO 2 MUESTRAS , SE UTILIZA COMUNMENTE EL ESTADISTICO t-STUDENT

$$t_{\text{de Student}} = \frac{P1 - P2}{\sqrt{\frac{S1^2}{n1} + \frac{S2^2}{n2}}} \quad \text{donde}$$

P1,P2 son los promedios de las muestras 1 y 2, respectivamente.

S1²,S2² son respectivamente la varianza de la muestra 1 y 2.

n1,n2 son el numero de observaciones de las muestras 1 y 2.

Una varianza se estima sumalizando las diferencias al cuadrado de cada dato respecto a su media y dividiendo la suma total entre "n-1" observaciones :

$$S1^2 = \frac{\Sigma(Xi-P1)^2}{n1 - 1} \quad S2^2 = \frac{\Sigma(Xi-P2)^2}{n2 - 1}$$

Xi es una observacion particular

PARA NUESTRO EJEMPLO :

$$P1 = (23 + 43 + \dots 26)/10 = 26$$

$$P2 = (17 + 4 + \dots 11)/9 = 11.56$$

$$\Sigma(Xi-P1)^2 = (23-26)^2 + (43-26)^2 + \dots (26-26)^2 = 1354$$

$$S1^2 = 1354 / 9 = 150.444$$

$$\Sigma(Xi-P2)^2 = (17-11.56)^2 + (14-11.56)^2 + \dots (11-11.56)^2 = 228.2233$$

$$S2^2 = 228.2233 / 8 = 36.0278$$

$$t = \frac{26 - 11.56}{\sqrt{\frac{150.4444}{10} + \frac{36.0278}{9}}} = 3.3086$$

EL VALOR "t" CALCULADO, O VALOR OBSERVADO, SE COMPARA CON UN VALOR TEORICO DE TABLAS, O VALOR ESPERADO, A UN NIVEL DE SIGNIFICANCIA DADO Y PARA LOS GRADOS DE LIBERTAD RESPECTIVOS.

TODOS LOS LIBROS DE ESTADISTICA BASICA INCLUYEN TABLAS DE VALORES "t" TEORICOS.

SIN AHONDAR MUCHO EN DEFINICIONES BASTE DECIR QUE EN BIOLOGIA LOS NIVELES DE SIGNIFICANCIA, - VALORES " α "-, MAS UTILIZADOS SON EL 0.01 % Y EL 0.05 % .

ESTO DE ALGUNA MANERA INDICA LA CONFIABILIDAD QUE SE ESPERA TENER EN LAS DEDUCCIONES.

LOS GRADOS DE LIBERTAD, - "G.L."-, PARA ESTE ESTADISTICO SE ESTIMAN POR LA FORMULA $n_1 + n_2 - 2$

EL PLANTEAMIENTO DE HIPOTESIS PARA LA TOMA DE DECISION SE ESTABLECE ENTONCES :

$H_0 : \mu_1 = \mu_2$
$H_a : \mu_1 <> \mu_2$
Si t CALCULADO > $t(\alpha=0.01), (G.L.=n_1+n_2-2)$ Se rechaza H_0

EN NUESTRO EJEMPLO EL VALOR "t" TEORICO ES :

$$t(\alpha=0.01), (G.L.=10+9-2=17) = 2.898$$

COMO t CALCULADO = 3.3086 > t DE TABLAS RECHAZAMOS H_0

CONCLUIMOS QUE HAY SUFICIENTE EVIDENCIA EXPERIMENTAL ($P \leq 0.01$) DE QUE LOS CABRITOS DE UNA RAZA DIFIEREN, EN SU PESO, DE LOS CABRITOS DE LA OTRA RAZA.

LA COMPRESION DE LAS PRUEBAS DE HIPOTESIS Y SU ADECUADO PLANTEAMIENTO, ES FUNDAMENTAL APOYO PARA LA INVESTIGACION CIENTIFICA.

EL ANALISIS DE VARIANZA (ANOVA) ES UNA TECNICA ESTADISTICA DESARROLLADA POR FISHER, PARA LA PRUEBA DE HIPOTESIS DE VARIAS MUESTRAS.

ES MUY APLICADA EN LA INVESTIGACION BIOLOGICA, AGROPECUARIA, MEDICA, PSICOLOGICA, Y EN GENERAL EN TODA AQUELLA DISCIPLINA QUE SE APLIQUE LA OBSERVACION Y LA EXPERIMENTACION (DANIEL, 1980).

PERTENECE A LA LLAMADA ESTADISTICA PARAMETRICA; GIRANDO EN TORNO A UNA SERIE DE SUPUESTOS, MISMOS QUE SE DEBE VIGILAR SE CUMPLAN, PARA QUE LAS CONCLUSIONES SEAN DEL TODO CIERTAS.

LA TECNICA DEL ANOVA, DE MANERA GENERAL, CONSISTE EN PARTICIONAR LA VARIACION TOTAL OBSERVADA EN SUS CAUSAS PARCIALES.

POSTERIORMENTE, PARA CADA UNA DE ESTAS CAUSAS CALCULAR :

- LOS GRADOS DE LIBERTAD
- LA SUMA DE CUADRADOS DE LAS DESVIACIONES DE CADA OBSERVACION CON RESPECTO A LA MEDIA
- LA VARIANZA O CUADRADO MEDIO

FINALMENTE SE APLICA LA PRUEBA DE HIPOTESIS, EN ESTE CASO LA PRUEBA DE FISHER O DE RAZON DE VARIANZAS.

EN ESTE TRABAJO SE EXPONEN 4 DIFERENTES DISEÑOS, SIGUIENDO UN PATRON DE PROCEDIMIENTO MUY SEMEJANTE EN TODOS ELLOS.

ADEMAS SE EMPLEA UNA NOMENCLATURA MUY GENERAL, CON LO QUE SEGURAMENTE SE FACILITARA LA COMPRENSION DE LOS DISEÑOS MAS COMPLICADOS.

EN TODOS LOS CASOS SE PRESENTA LO SIGUIENTE :

- A TABLA TIPICA DE DATOS Y BREVE EXPLICACION DE LA MISMA
 - B MODELO
 - C SUPOSICIONES DEL MODELO
 - D PLANTEAMIENTO DE LA HIPOTESIS
 - E ESTADISTICO DE PRUEBA
 - F EJEMPLO RESUELTO EN DETALLE :
- 1) PLANTEAMIENTO DEL PROBLEMA.
 - 2) TABLA DE DATOS.
 - 3) PLANTEAMIENTO DE LA HIPOTESIS.
 - 4) DETALLE DEL PROCEDIMIENTO DE CALCULO.
 - 5) REGLA DE DECISION.
 - 6) CONCLUSIONES.

Tabla típica para un diseño Completamente Aleatorizado

	Tratamiento				
	1	2	3	4 k
	Y11	Y12	Y13	Y14	Y1k
	Y21	Y22	Y23	Y24	Y2k
	Y31	Y32	Y33	Y34	Y3k
	Y41	Y42	Y43	Y44	Y4k

	Yn1	Yn2	Yn3	Yn4	Ynk
tot. por tratam.	T.1	T.2	T.3	T.4	T.k
					tot. global T..

Explicación:

Sean "k" tratamientos o muestras con "n" repeticiones cada uno. Se estiman los "T.k" totales por muestra y el "T.." total global.

Tratamientos $i = 1 \dots\dots\dots k$
 repeticiones $j = 1 \dots\dots\dots n$

Modelo

$Y_{ij} = \mu + Z_i + e_{ij}$

Y_{ij} es la "j-ésima" repetición del "i-ésimo" tratamiento.

μ es la media poblacional.

Z_i es el efecto del "i-ésimo tratamiento".

e_{ij} es el error aleatorio de la observación.

Nota : puede ocurrir que el número de repeticiones sea diferente en cada tratamiento; sin embargo, tanto el modelo, como el procedimiento para los cálculos serán los mismos (se estiman promedios "Ti." por tratamiento o muestra).

Suposiciones

- a) Los "k" conjuntos de datos, constituyen "k" muestras aleatorias de las poblaciones respectivas.
- b) Todas las poblaciones tienen la misma varianza.
- c) Los "e_{ij}" están normal e independientemente distribuidos, con media = 0 y varianza sigma cuadrada .

Hipótesis

H₀ : No hay efecto de ningún tratamiento.

$$\mu_1 = \mu_2 = \dots = \mu_j$$

H_a : Hay diferente efecto en al menos un par de tratamientos.

$$\mu_1 \neq \mu_2 \neq \dots = \mu_j$$

Estadístico de Prueba

F₀ ~ F

con (k - 1) G.L. en el numerador

(N - k) G.L. en el denominador

y " α " nivel de significancia.

Ejemplo:

Se estudio el efecto de la glucosa sobre la liberación de insulina. Para ello se trataron especimenes de tejido pancreático de animales experimentales, con 5 concentraciones diferentes de glucosa.

Posteriormente se determinó la cantidad de insulina liberada, resultando los datos que muestra la tabla .

Por medio del Análisis de Varianza, el investigador puede conocer si existe o no diferencia significativa, entre las 5 poblaciones implicadas, respecto a la insulina liberada.

Insulina liberada a diferentes concentraciones de Glucosa

	Concentración de Glucosa				
	1	2	3	4	5
	1.53	3.15	3.89	8.18	5.86
	1.61	3.96	4.80	5.64	5.46
	3.75	3.59	3.69	7.36	5.69
	2.89	1.89	5.70	5.33	6.49
	3.26	1.45	5.62	8.82	7.81
	2.83	3.49	5.79	5.62	9.03
	2.86	1.56	4.75	8.75	7.49
	2.59	2.44	5.33	7.10	8.98
totales	21.32	21.53	39.56	56.71	56.81
				tot.	
				glogal	193.53

Se trata de un diseño con "k" = 5 tratamientos o muestras, de "n" = 8 repeticiones en todos los casos.

Y_{3,7} = 4.75 es la repetición # 7 del tratamiento # 3.
T_{1.} = 21.32 es la suma de todas las repeticiones del tratamiento 1.

Considerando que se cumplen los supuestos, planteamos las siguientes Hipótesis:

H_0 : La concentración de glucosa no tiene efecto sobre la liberación de insulina.

$$\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

H_a : Al menos en dos concentraciones ocurre diferente efecto.

$$\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5$$

En la decisión se empleará el nivel de significancia del 0.05 .

Procedimiento para el Análisis de Varianza

a) Factor de Corrección.

$$\begin{aligned} \text{F.C.} &= \frac{T..^2}{N} & N &= \text{número total de datos} \\ &= \frac{195.93^2}{40} \\ &= 9.59.7141 \end{aligned}$$

b) Suma de Cuadrados Total.

$$\begin{aligned} \text{S.C.Tot} &= \sum_{j=1}^k \sum_{i=1}^n Y_{ij}^2 - \text{F.C.} \\ &= 1.53^2 + 1.61^2 + \dots + 8.98^2 - 959.7141 \\ &= 201.8814 \end{aligned}$$

c) Suma de Cuadrados de los Tratamientos.

$$\begin{aligned} \text{S.C.Trat.} &= \sum_{j=1}^k T.j^2 / n - \text{F.C.} \\ &= (21.32^2 + 21.53^2 + \dots + 56.81^2) / 8 - 959.7141 \\ &= 156.0955 \end{aligned}$$

d) Suma de Cuadrados del Error.

$$\begin{aligned} \text{S.C.Error} &= \text{S.C.Tot} - \text{S.C.Trat.} \\ &= 201.8814 - 156.0955 \\ &= 45.7859 \end{aligned}$$

e) Cuadrados Medios de los Tratamientos.

$$\begin{aligned} \text{C.M.Trat} &= \text{S.C.Trat.} / \text{G.L.Trat.} \\ \text{G.L.Trat} &= k - 1 = 5 - 1 = 4 \\ &= 156.0955 / 4 \\ &= 39.0239 \end{aligned}$$

f) Cuadrados Medidos del Error.

$$\begin{aligned} \text{C.M.Error} &= \text{S.C.Error} / \text{G.L.Error} \\ \text{G.L.Error} &= (k) (n - 1) = 5 * 7 = 35 \\ &= 45.7859 / 35 \\ &= 1.3082 \end{aligned}$$

g) Razón de Varianza.

$$\begin{aligned} F_0 &== \text{C.M.trat.} / \text{C.M.Error} \\ &= 39.0239 / 1.3082 \\ &= 29.83 \end{aligned}$$

Regla de Decisión

$$F_t \left(\alpha = 0.05 \right) \begin{matrix} (k-1 = 4) \\ (N - k = 35) \end{matrix} = 2.69$$

Como $F_0 = 29.83 > F_t = 2.69 \implies$ Se rechaza H_0 .

Se concluye que hay suficiente evidencia experimental ($P \leq 0.05$), de que al menos un par de concentraciones de glucosa tienen efecto sobre la liberación de insulina.

Tabla típica para un diseño en Bloques completamente Aleatorizados

Bloques	Tratamientos						tot. por bloque
	1	2	3	4 k		
1	Y11	Y12	Y13	Y14		Y1k	T1.
2	Y21	Y22	Y23	Y24		Y2k	T2.
3	Y31	Y32	Y33	Y34		Y3k	T3.
.
.
.
n	Yn1	Yn2	Yn3	Yn4		Ynk	Tn.
tot. por tratam.	T.1	T.2	T.3	T.4		T.k	T.. tot. gral.

Explicación:

Sea un experimento con "k" número de tratamientos y "n" número de bloques. Es decir, existe algún gradiente de variación, que se considera factor de blocaje.

Se calculan los totales por bloque, totales por tratamiento y el total general.

Bloques i=1 n
 Tratamientos j=1 k

Modelo

$$Y_{ij} = \mu + \beta_i + \gamma_j + e_{ij}$$

Y_{ij} es a la observación del "i-ésimo" bloque, "j-ésimo" tratamiento.

μ es la media poblacional.

β_i es el efecto del "i-ésimo" bloque.

e_{ij} es el error aleatorio de la observación.

Nota : En éste diseño una observación es a la vez una muestra.

Suposiciones

- a) Cada Y_{ij} es una muestra aleatoria de tamaño 1.
- b) Los e_{ij} están normal e independientemente distribuidos con media = 0 y varianza σ^2 .
- c) No existe interacción entre los tratamientos y los bloques.

Hipótesis

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_j$$

No hay efecto de ningún tratamiento.

$$H_a : \mu_1 \neq \mu_2 \neq \dots \neq \mu_j$$

Al menos en un par de tratamientos hay diferencia.

Estadístico de Prueba y Regla de Decisión

$$F_0 \sim F$$

con $(k - 1)$ G.L. en el numerador

$(n - 1) - (k - 1)$ G.L. en el denominador

y " α " nivel de significancia.

Ejemplo :

En un instituto, para comparar tres métodos de enseñanza del idioma inglés, se cuantificó el número de meses necesarios para dominar el idioma.

Los alumnos, seleccionados al azar, tenían diferentes edades. Pensando que esto influiría de manera decisiva en los resultados, se optó por crear 5 grupos de edad, tomando tres alumnos por cada uno de los grupos; esto es, un alumno por cada método de enseñanza.

Al final del experimento se obtuvieron los siguientes resultados :

Grupo de Edad	Método de enseñanza			tot. por bloque
	1	2	3	
"-" de 15 años	7	9	10	26
15 - 20 "	8	9	10	27
20 - 30 "	9	9	12	30
30 - 50 "	10	9	12	31
"+" de 50 "	11	12	14	37
tot. por Tratam.	45	48	58	151

Se trata de un diseño con "n" = 5 bloques y "k" = 3 tratamientos. O dicho de otra manera, hay 5 bloques, con 3 repeticiones cada uno.

$Y_{2,3} = 10$ es el dato correspondiente a el bloque 2, tratamiento 3.

$T_{.2} = 48$ es la suma de los datos correspondientes al tratamiento 2.

Asumiendo que fueran ciertos los supuestos, planteamos las siguientes hipótesis:

H_0 : No hay diferencia entre los métodos de enseñanza ($\mu_1 = \mu_2 = \mu_3$).

H_a : Al menos en dos métodos ocurre un diferente efecto ($\mu_1 \neq \mu_2 \neq \mu_3$).

Se empleará el nivel de significancia del 0.05

Procedimiento para el Análisis de Varianza.

a) Factor de Corrección.

$$\begin{aligned} F.C. &= T_{..}^2 / nk = (151)^2 / (3)(5) = 22801 / 15 \\ &= 1520.0667 \end{aligned}$$

b) Suma de Cuadrados Total.

$$\begin{aligned} S.C.Tot. &= \sum_{i=1}^n \sum_{j=1}^k Y_{ij}^2 - F.C. \\ &= 7^2 + 9^2 + \dots + 14^2 - 1520.0667 \\ &= 18.5333 \end{aligned}$$

c) Suma de Cuadrados de los Bloques.

$$\begin{aligned} S.C.Block &= \sum_{i=1}^n T_{i.}^2 / n - F.C. \\ &= (26^2 + \dots + 37^2) / 5 - 1520.0667 \\ &= 18.5333 \end{aligned}$$

d) Suma de Cuadrados de los tratamientos.

$$\begin{aligned} S.C.Trat. &= \sum_{j=1}^k T_{.j}^2 / k - F.C. \\ &= (45^2 + 48^2 + 58^2) / 3 - 1520.0667 \\ &= 24.9333 \end{aligned}$$

e) Suma de Cuadrados del Error.

$$\begin{aligned} S.C.Error &= S.C.tot. - (S.C.block + S.C.trat.) \\ &= 49.9333 - (18.5333 + 24.9333) \\ &= 3.4667 \end{aligned}$$

f) Cuadrados medios de Tratamientos.

$$\begin{aligned} \text{C.M.trat.} &= \text{S.C.trat.} / (k - 1) \\ &= 18.5333 / 2 \\ &= 9.2667 \end{aligned}$$

g) Cuadrados Medios del Error.

$$\begin{aligned} \text{C.M.Error} &= \text{S.C.Error} / ((n-1)(k-1)) \\ &= 3.4667 / 8 \\ &= 0.4333 \end{aligned}$$

h) Razón de Varianza.

$$\begin{aligned} F_0 &= \text{C.M.trat.} / \text{C.M.Error} \\ &= 9.2667 / 0.4333 \\ &= 21.39 \end{aligned}$$

Regla de Decisión

$$F_t (\alpha = 0.05) \begin{matrix} (k-1 = 2) \\ [(n-1)(k-1) = 8] \end{matrix} = 4.46$$

Como $F_0 = 21.36 > F_t = 4.46 \implies$ se rechaza H_0 .

En éste caso se concluye que existe suficiente evidencia experimental ($P \leq 0.05$), para decir que hay diferencia en los métodos de enseñanza.

Tabla típica de un Diseño en Cuadro Latino.

Hilera	Columna					tot. por hilera
	1	2	3	4 a	
1	Distribución de tratamientos en el campo.					T1.
2						T2.
3						T3.
4						T4.
.						.
.	.					
a						Ta.
tot. por columna	T.1	T.2	T.3	T.4 T.a	T..
	A	B	C	D	E...a	
tot. por tratam.	TA	TB	TC	TD	TE ...Ta	

Explicación:

Este modelo es de uso amplio en experimentos agrícolas y en general en aquellos casos en que se quiere controlar la variación debida a dos fuentes distintas (bloqueo en dos direcciones).

El número total de tratamientos es igual al número total de repeticiones por tratamiento. Por lo que cada tratamiento aparece una vez en cada hilera y una vez en cada columna; siendo ésta la característica básica del Cuadro Latino.

Se calculan, totales por hilera (T_i), totales por columna (T_j), totales por tratamiento (T_k) y el total general.

Tratamientos	k=1 a
Block's tipo 1	i=1 a
Block's tipo 2	j=1 a

Modelo

$$Y_{ijk} = \mu + Z_i + B_j + \alpha_k + e_{ijk}$$

Y_{ijk} observación del "i-ésimo" tratamiento; que en el campo corresponde al "j-ésimo" bloque tipo 1 (hileras), "k-ésimo" bloque tipo 2 (columnas).

μ es la media poblacional.

Z_i es el efecto del "i-ésimo" tratamiento.

B_j es el efecto del "j-ésimo" bloque tipo 1.

α_k es el efecto del "k-ésimo" bloque tipo 2.

e_{ijk} es el error aleatorio de la observación.

Suposiciones

- a) En cada bloque deben estar representados todos los tratamientos. Es decir, que el número de bloques (sean éstos tipo 1 o tipo 2), corresponde con el número de repeticiones.

Nota : En la siguiente página se indica el procedimiento para conformar los tratamientos, en un arreglo en Cuadro Latino.

- b) Los " e_{ijk} " están normal e independientemente distribuidos con media = 0 y varianza sigma cuadrada.
- c) No hay interacción entre bloques, ni entre bloques con tratamientos.

Estadístico de Prueba

$$F_0 \sim F$$

con $(a - 1)$ G.L. en el numerador

$(a - 1)(a - 2)$ G.L. en el denominador

y " α " nivel de significancia.

Procedimiento para elaborar un Cuadro Latino.

Los tratamientos se asignan aleatoriamente dentro de cada block, de la siguiente manera:

- 1) Se parte de un arreglo básico, en que se ordenan los tratamientos en diagonal.

A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A

- 2) Se aleatorizan renglones.

C	D	A	B
B	C	D	A
A	B	C	D
D	A	B	C

- 3) Finalmente se aleatorizan columnas.

D	B	C	A
C	A	B	D
B	D	A	C
A	C	D	B

Con ello se obtiene la distribución de muestras en el campo. lo que debe hacerse antes de iniciar el experimento. Es decir, que a partir de ella se aplicarán los diferentes tratamientos.

Ejemplo:

En un experimento se estudiaron 7 variedades de maíz:

- A Híbrido H-42.
- B Híbrido Favorita.
- C Sintético Precoz, Blanco.
- D Sintético Plus-1, Blanco.
- E Sintético Carmen, Blanco.
- F Sintético Carmen, Amarillo.
- G Sintético Carmen, Amarillo P. L.

Se empleo el diseño en Cuadro Latino, procurando eliminar diferencias de fertilidad entre hileras, así como diferencias de fertilidad entre columnas.

La distribución de las variedades en el campo, con su respectiva producción de grano seco, en Kg/parcela, fue la siguiente:

		C o l u m n a							
		1	2	3	4	5	6	7	tot. por hileras
Hilera									
		F	B	D	C	E	A	G	
1		12.0	10.6	10.6	9.3	11.0	9.0	11.2	73.7
		B	E	G	F	A	D	C	
2		8.7	8.6	8.3	9.6	7.5	9.0	8.2	59.9
		G	C	E	D	F	B	A	
3		9.3	7.9	7.9	8.2	8.6	9.9	10.1	61.9
		C	F	A	G	B	E	D	
4		7.0	7.9	8.0	8.1	9.1	10.8	12.7	63.6
		D	G	B	A	C	F	E	
5		8.6	9.9	8.8	8.1	8.9	10.7	10.0	65.0
		A	D	F	E	G	C	B	
6		7.7	7.1	6.6	6.6	7.3	7.5	8.7	51.5
		E	A	C	B	D	G	F	
7		6.3	6.6	7.8	8.2	7.1	8.6	7.3	51.9
tot. por columna		59.9	58.6	58.0	58.1	59.5	65.5	68.2	427.5

La producción en grano seco por variedad y repetición fué la siguiente:

A	B	C	D	E	F	G	
9.0	10.6	9.3	10.6	11.0	12.0	11.2	
7.5	8.7	8.2	9.0	8.6	9.6	8.3	
10.1	9.9	7.9	8.2	7.9	8.6	9.3	
8.0	9.1	7.0	12.7	10.8	7.9	8.1	
8.1	8.8	8.9	8.6	10.0	10.7	9.9	
7.7	8.7	7.5	7.1	6.6	6.6	7.3	
6.6	8.2	7.8	7.1	6.3	7.3	8.6	
57.0	64.0	56.6	63.3	61.2	62.7	62.7	
						tot.	
						global =	427.5

Es un diseño en Cuadro Latino con "a" = 7 tratamientos, = 7 bloques tipo 1, = 7 bloques tipo 2, = 7 repeticiones por tratamiento.

Y_i = 8.6 es una determinada repetición del tratamiento "F", que en el campo quedó distribuida como "j"=3 bloque tipo 1, "k"=5 bloque tipo 2.

T₁ = 59.9 es el total para el bloque 1 de columnas.

T₄ = 63.6 es el total para el bloque 4 de hileras.

Considerándose ciertos los supuestos, planteamos las siguientes hipótesis:

H₀ : Las variedades de maíz estudiadas, no reportan diferencias, en cuanto a su producción en grano seco .
 $\mu_1 = \mu_2 = \dots \mu_7$

H_a : Al menos en dos variedades se observa diferente rendimiento.
 $\mu_1 \langle \rangle \mu_2 \langle \rangle \dots \mu_7$

Se empleará el nivel de significancia del 0.05 .

Procedimiento para el Análisis de Varianza.

a) Factor de Corrección.

$$\begin{aligned} \text{F.C.} &= T_{..}^2 / a^2 = (427.5)^2 / (7)^2 \\ &= 3729.72 \end{aligned}$$

b) Suma de Cuadrados Total

$$\begin{aligned} \text{S.C.tot.} &= \sum_{i=1}^a \sum_{j=1}^a X_{ij}^2 - \text{F.C.} \\ &= 9^2 + 7.5^2 + \dots + 8.6^2 - 3729.72 \\ &= 102.17 \end{aligned}$$

c) Suma de Cuadrados por Hileras.

$$\begin{aligned} \text{S.C.Hil} &= \sum_{i=1}^a T_{i.}^2 / a - \text{F.C.} \\ &= 73.7^2 + 59.9^2 + \dots + 51.9^2 - 3729.72 \\ &= 51.3 \end{aligned}$$

d) Suma de Cuadrados por Columnas.

$$\begin{aligned} \text{S.C.Col} &= \sum_{j=1}^a T_{.j}^2 / a - \text{F.C.} \\ &= 59.6^2 + 58.6^2 + \dots + 68.2^2 - 3729.72 \\ &= 14.2 \end{aligned}$$

e) Suma de Cuadrados de los Tratamientos.

$$\begin{aligned} \text{S.C.Trat.} &= \sum_{k=1}^a T_{k.}^2 / a - \text{F.C.} \\ &= 57^2 + 64^2 + \dots + 62.7^2 - 3729.72 \\ &= 7.92 \end{aligned}$$

f) Suma de Cuadrados del Error.

$$\begin{aligned} \text{S.C.Error} &= \text{S.C.Tot} - (\text{S.C.Hil} + \text{S.C.Col} + \text{S.C.Trat}) \\ &= 102.17 - (51.3 + 14.2 + 7.92) \\ &= 28.75 \end{aligned}$$

g) Cuadrados Medios de los Tratamientos.

$$\begin{aligned} \text{C.M.Trat.} &= \text{S.C.Trat.} / \text{G.L.Trat.} \\ \text{G.L.Trat.} &= a - 1 = 7 - 1 = 6 \\ &= 7.92 / 6 \\ &= 1.32 \end{aligned}$$

h) Cuadrados Medios del Error.

$$\begin{aligned} \text{C.M.Error} &= \text{S.C.Error} / \text{G.L.Error} \\ \text{G.L.Error} &= (a-1)(a-2) = (6)(5) = 30 \\ &= 28.75 / 30 \\ &= 0.958 \end{aligned}$$

i) Razón de Varianza.

$$\begin{aligned} F_0 &= \text{C.M.Trat.} / \text{C.M.Error} \\ &= 1.32 / 0.958 \\ &= 1.38 \end{aligned}$$

Regla de Decisión

$$F_t (\alpha = 0.05) \quad (a-1 = 6) \quad = \quad 3.0 \\ (a-1)(a-2) = 12$$

Como $F_0 = 1.38 < F_t = 3.0 \implies$ no se rechaza H_0 .

La F_0 es muy pequeña (al menos para el nivel de significancia adoptado), por lo que se concluye, que no hay suficiente evidencia experimental ($P \leq 0.05$), para afirmar que hay diferencia en las variedades de maíz estudiadas, en cuanto a sus rendimientos en grano seco.

Tabla típica para un diseño completamente aleatorizado de dos Factores.

Factor A niveles	Factor B niveles				tot. para factor A
	1	2	3 b	
<u>1</u>	Y11'1	Y12'1	Y13'1	Y1b'1	T1..
	Y11'n	Y12'n	Y13'n	Y1b'n	
<u>2</u>	Y21'1	Y22'1	Y23'1	Y2b'1	T2..
	Y21'n	Y22'n	Y23'n	Y2b'n	
<u>3</u>	Y31'1	Y32'1	Y33'1	Y3b'1	T3..
	Y31'n	Y32'n	Y33'n	Y3b'n	
<u>a</u>	Ya1'1	Ya2'1	Ya3'1	Yab'1	Ta..
	Ya1'n	Ya2'n	Ya3'n	Yab'n	
tot. para factor B	T.1.	T.2.	T.3.	T.b.	T...

Explicación:

Sea un diseño con dos factores : A y B. Con "a" número de niveles para el factor A, y "b" número de niveles para el factor B. Teniendo además, "n" número de repeticiones en cada conjunción de un determinado "i" nivel A con otro "j" nivel B.

Se calculan los totales para cada nivel de cada factor (los T_i y los T_j) y el total global ($T...$).

Factor	Niveles	Repeticiones
A	$i = 1 \dots a$	$r = 1 \dots n$
B	$j = 1 \dots b$	

Modelo

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij}$$

Y_{ij} es el valor que suma los datos de la conjunción "i" nivel del factor A con "j" nivel del factor B.

μ es la media poblacional.

α_i es el efecto del "i-esimo" nivel del factor A

β_j es el efecto del "j-ésimo" nivel del factor B

$(\alpha\beta)_{ij}$ es efecto de la interacción A B.

e_{ij} es el error aleatorio de la observación.

Suposiciones

- Cada una de las "ij" combinaciones, constituye una muestra aleatoria de "n" observaciones.
- Cada una de las "ij" combinaciones, provienen de poblaciones normalmente distribuidas.
- Todas las poblaciones tienen la misma varianza.

Hipótesis

- Ho : No hay efecto del Factor A, en ninguno de sus niveles.
 $\alpha_1 = \alpha_2 = \dots \alpha_i$

Ha : Hay efecto en al menos algún(os) nivel(es) del factor A.
 $\alpha_1 \langle \rangle \alpha_2 \langle \rangle \dots \alpha_i$
- Ho : No hay efecto del Factor B, en ninguno de sus niveles.
 $\beta_1 = \beta_2 = \dots \beta_j$

Ha : Hay efecto en al menos algún(os) nivel(es) del Factor B.
 $\beta_1 \langle \rangle \beta_2 \langle \rangle \dots \beta_j$
- Ho : No tiene ningún efecto la interacción de los Factores AB.
 $(\alpha\beta)_{1,1} = (\alpha\beta)_{1,2} = \dots (\alpha\beta)_{i,j}$

Ha : Existe efecto de interacción, en al menos alguna(s) combinación(es) AB.
 $(\alpha\beta)_{1,1} \langle \rangle (\alpha\beta)_{1,2} \langle \rangle \dots (\alpha\beta)_{i,j}$

Estadístico de Prueba

$F_0 \sim F$

con $(a - 1)$ G.L. en el numerador

$(n - 1)(ab - 1)$ G.L. en el denominador

y " α " nivel de significancia.

Ejemplo:

En un campo agrícola se cuantificaron los rendimientos, en Ton./Ha. de grano de maíz, combinando cuatro variedades de éste (factor A) con tres dosis de Nitrógeno (factor B). Los resultados aparecen a continuación:

Factor A (variedad)	Factor B (dosis de Nitrógeno)			tot. varied.
	50 Kg/Ha	100 Kg/Ha	150 Kg/Ha	
H-1	3	4	6	65
	2	4	5	
	3	4	6	
	3	5	7	
	2	5	6	
H-2	5	6	7	98
	5	5	8	
	6	6	8	
	6	7	9	
	5	7	8	
V-1	2	2	5	45
	1	2	4	
	1	3	5	
	2	3	6	
	1	2	6	
V-2	4	6	7	82
	5	6	8	
	4	5	7	
	3	5	8	
	2	4	8	
tot. dosis	65	91	134	290

Se trata de un experimento bifactorial con "a" = 4 niveles del factor A, "b" = 3 niveles del factor B y "n" = 5 repeticiones por tratamiento.

Sumando las repeticiones de cada tratamiento (una combinación "ai , bj" particular), la tabla se simplifica:

Variedad	Dosis			tot. varied.
	50	100	150	
H-1	13	22	30	65
H-2	27	31	40	98
V-1	7	12	26	45
V-2	18	26	38	82
tot. dosis	65	91	134	290

Y _{4,2}	=	26	es el valor de la interacción : nivel 4 de A, nivel 2 de B.
T2..	=	98	es la suma de todos los valores - correspondientes al nivel 2 de A.
T.3.	=	134	es la suma de todos los valores - correspondientes al nivel 3 de B.

De cumplirse los supuestos, planteamos las siguientes Hipótesis.

- a) Ho : Las variedades de maíz estudiadas, tienen en promedio el mismo efecto.

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$$

- Ha : Al menos un par de variedades tienen diferente rendimiento.

$$\alpha_1 \langle \rangle \alpha_2 \langle \rangle \alpha_3 \langle \rangle \alpha_4$$

- b) Ho : Las dosis de Nitrógeno tienen en promedio el mismo efecto.

$$\beta_1 = \beta_2 = \beta_3$$

- Ha : Al menos un par de dosis de Nitrógeno, tienen efecto diferente.

$$\beta_1 \langle \rangle \beta_2 \langle \rangle \beta_3$$

- c) Ho : Las diferentes combinaciones de niveles de factores AB (variedad/dosis de Nitrógeno), tienen en promedio el mismo efecto.

$$(\alpha\beta)_{1,1} = (\alpha\beta)_{1,2} = \dots (\alpha\beta)_{4,3}$$

- Ha : Resultan diferentes efectos en al menos un par de combinaciones de niveles de factores AB.

$$(\alpha\beta)_{1,1} \langle \rangle (\alpha\beta)_{1,2} \langle \rangle \dots (\alpha\beta)_{4,3}$$

Para evaluar las Hipótesis, emplearemos el nivel de significancia del 0.05, en todos los casos.

Procedimiento para el Análisis de Varianza

a) Factor de Corrección.

$$\begin{aligned} \text{F.C.} &= \frac{T \dots^2}{a b n} \\ &= \frac{290^2}{(4)(3)(5)} \\ &= 1401.67 \end{aligned}$$

b) Suma de Cuadrados Total.

$$\begin{aligned} \text{S.C.Tot} &= \sum_{i=1}^a \sum_{j=1}^b \sum_{r=1}^n Y_{ijk}^2 - \text{F.C.} \\ &= 3^2 + 4^2 + \dots + 8^2 - 1401.67 \\ &= 254.33 \end{aligned}$$

c) Suma de Cuadrados de los Tratamientos.

$$\begin{aligned} \text{S.C.Trat} &= \sum_{i=1}^a \sum_{j=1}^b T_{ij}^2 / n - \text{F.C.} \\ &= (13^2 + 27^2 + \dots + 38^2) / 5 - 1401.67 \\ &= 229.53 \end{aligned}$$

d) Suma de Cuadrados del Error.

$$\begin{aligned} \text{S.C.Error} &= \text{S.C.Tot.} - \text{S.C.Trat.} \\ &= 254.33 - 229.53 \\ &= 24.8 \end{aligned}$$

e) Suma de Cuadrados del Factor A.

$$\begin{aligned} \text{S.C.A} &= \sum_{i=1}^a T_{i..}^2 / b n - \text{F.C.} \\ &= (65^2 + 98^2 + 45^2 + 82^2) / (3)(5) - 1401.67 \\ &= 103.53 \end{aligned}$$

f) Suma de Cuadrados del Factor B.

$$\begin{aligned} \text{S.C.B} &= \sum_{j=1}^b T.j.^2 / a n - F.C. \\ &= (65^2 + 91^2 + 134^2) / (4)(5) - 1401.67 \\ &= 121.43 \end{aligned}$$

g) Suma de Cuadrados de la interacción AB.

$$\begin{aligned} \text{S.C.AB} &= \text{S.C.Trat.} - \text{S.C.A} - \text{S.C.B} \\ &= 229.53 - 103.53 - 121.43 \\ &= 4.57 \end{aligned}$$

h) Cuadrados Medios del Factor A.

$$\begin{aligned} \text{C.M.A} &= \text{S.C.A} / \text{G.L.A} \\ \text{G.L.A} &= a - 1 = 4 - 1 = 3 \\ &= 103.53 / 3 \\ &= 34.51 \end{aligned}$$

i) Cuadrados Medios del Factor B.

$$\begin{aligned} \text{C.M.B} &= \text{S.C.B} / \text{G.L.B} \\ \text{G.L.B} &= b - 1 = 3 - 1 = 2 \\ &= 121.43 / 2 \\ &= 60.72 \end{aligned}$$

j) Cuadrados Medios de la Interacción AB.

$$\begin{aligned} \text{C.M.AB} &= \text{S.C.AB} / \text{G.L.AB} \\ \text{G.L.AB} &= (a-1)(b-1) = (3)(2) = 6 \\ &= 4.57 / 6 \\ &= 0.76 \end{aligned}$$

k) Cuadrados Medios del Error.

$$\begin{aligned} \text{C.M.Error} &= \text{S.C.Error} / \text{G.L.Error} \\ \text{G.L.Error} &= (n-1) (ab-1) = (5-1) (4*3-1) \\ &= (4) (11) = 44 \\ &= 24.8 / 44 \\ &= 0.56 \end{aligned}$$

l) Razón de Varianza del Factor A.

$$\begin{aligned} F_{0A} &= \text{C.M.A} / \text{C.M.Error} \\ &= 34.51 / 0.56 \\ &= 61.62 \end{aligned}$$

m) Razón de Varianza del Factor B.

$$\begin{aligned} F_{0B} &= \text{C.M.B.} / \text{C.M.Error} \\ &= 60.72 / 0.56 \\ &= 108.42 \end{aligned}$$

n) Razón de Varianza de la interacción AB.

$$\begin{aligned} F_{0AB} &= \text{C.M.AB} / \text{C.M.Error} \\ &= 0.76 / 0.56 \\ &= 1.3571 \end{aligned}$$

Regla de Decisión

i) Factor A (dosis de Nitrógeno).

$$F_t (\alpha = 0.05) \begin{matrix} (a-1 = 3) \\ [(n-1) (ab-1) = 44] \end{matrix} = 2.84$$

Como $F_0 = 61.62 > F_t = 2.84$ se rechaza H_0 .

Se concluye que existe suficiente evidencia experimental ($P \leq 0.05$) para afirmar que, en al menos 2 niveles del factor A, las diferentes dosis de Nitrógeno, tienen efecto diferente sobre el rendimiento en grano seco de maíz.

ii) Factor B (Variedad).

$$F_t \quad (\alpha = 0.05) \quad \begin{array}{l} (b-1 = 2) \\ [(n-1)(ab-1) = 44] \end{array} = 3.23$$

Como $F_0 = 108.42 > F_t = 3.23 \implies$ se rechaza H_0 .

Se concluye que existe suficiente evidencia experimental ($P \leq 0.05$), para afirmar que, en al menos 2 niveles del factor B, las variedades estudiadas reportan diferentes rendimientos en grano seco de maiz.

iii) Interacción AB.

$$F_t \quad (\alpha = 0.05) \quad \begin{array}{l} [(a-1)(b-1) = 6] \\ [(n-1)(ab-1) = 44] \end{array} = 2.34$$

Como $F_0 = 1.35 < F_t = 2.34 \implies$ No se rechaza H_0 .

La interacción AB, resulto no ser significativa para el "a" seleccionado. Por tanto, se concluye que no existe evidencia experimental suficiente ($P \leq 0.05$), para asegurar que las diferentes combinaciones variedad-dosis de Nitrógeno, tengan efectos diferentes sobre los rendimientos en grano seco de maiz.

*** ANOVA - M. S. C. ***

PAQUETE DE PROGRAMAS PARA LA EN-
SEÑANZA Y APLICACION DEL ANALISIS
DE VARIANZA.

EL SISTEMA COMPRENDE :

- A) DISEÑO ALEATORIO SIMPLE
- B) DISEÑO EN BLOQUES AL AZAR
- C) DISEÑO EN CUADRO LATINO
- D) DISEÑO DE DOS FACTORES

AUTOR : MOISES SILVA CERUANTES
SEPTIEMBRE 1988

DESCRIPCION DEL SISTEMA ANOVA-MSC

EL PROGRAMA CUMPLE DOS OBJETIVOS PRIMORDIALES :

- A) PROPORCIONAR ASISTENCIA A QUIENES ESTAN INICIANDOSE EN EL MANEJO DEL ANALISIS DE VARIANZA.
- B) SER UNA HERRAMIENTA PARA EL TRATAMIENTO DE DATOS.

POR TANTO LOS POSIBLES USUARIOS DEL MISMO SERAN :

- A) PERSONAS QUE DESEEN APRENDER LOS PRINCIPIOS BASICOS DEL ANALISIS DE VARIANZA.
- B) PERSONAS QUE QUIERAN APLICAR A UN GRUPO DE DATOS ALGUNO DE LOS DISEÑOS DISPONIBLES.

COMO YA SE INDICO EN LA INTRODUCCION, LOS PROGRAMAS FUERON CREADOS EN EQUIPO PC-COMPATIBLE. POR TANTO, LOS REQUERIMIENTOS MINIMOS PARA MANEJARLOS SON :

- A) 256 K DE MEMORIA.
- B) SISTEMA OPERATIVO MS-DOS .

EL DISKETTE "ANOVA-MSC" CONTIENE, ADEMÁS DE LOS PROGRAMAS, EL INTERPRETE DE BASIC.

LO UNICO QUE HAY QUE HACER PARA UNA SESION, ES LO QUE SIGUE :

- 1) ENCENDER EL EQUIPO
- 2) CARGAR EL SISTEMA OPERATIVO
- 3) INSERTAR EL DISKETTE "ANOVA-MSC"
- 4) TECLEAR "ANOVA"

CON ESTA PALABRA SE INVOCA EL INICIO DEL SISTEMA.

EL MANEJO EN SU TOTALIDAD ES MUY SENCILLO.

SELECCIONANDO A TRAVES DE MENUS LAS OPCIONES ACORDES A LOS DATOS QUE SE ESTAN OPERANDO.

EN EL MODULO DE ENTRADA SE PRESENTA EL PROGRAMA (FIGURA # 5) Y SE CUESTIONA AL USUARIO SI DESEA ACCESAR EL PROGRAMA "ENSEÑANZA DEL ANOVA" (FIGURA # 6) .

EN EL SE VA PREGUNTANDO HASTA SELECCIONAR EL DISEÑO DE INTERES.

POR SUPUESTO QUE EL USUARIO CON CONOCIMIENTOS DEL ANOVA PUEDE LIBRAR LA PARTE TUTORIAL Y ACCESAR, SIN MAYORES PERDIDAS DE TIEMPO, LA PARTE DE CALCULOS.

LA FIGURA # 7 MUESTRA EL MENU GENERAL, TAL COMO SE PRESENTA EN LA PANTALLA.

EL USUARIO SELECCIONA, PULSANDO EL NUMERO RESPECTIVO, EL DISEÑO QUE DESEE TRABAJAR.

CON LA OPCION 4 SE POR CONCLUIDA LA SECCION, REGRESANDO AL SISTEMA OPERATIVO.

SUPONIENDO QUE SE QUIERAN TRABAJAR DATOS CON EL DISEÑO EN CUADRO LATINO, ENTONCES SE TECLEA 3 Y DE INMEDIATO SE PRESENTA EN LA PANTALLA OTRO MENU, INDICANDO LAS OPCIONES DEL ANOVA EN CUADRO LATINO, FIGURA # 8 .

EL PROGRAMA DE CAPTURA (OPCION 1 DEL SUBMENU) SE EMPLEA TODA VEZ QUE SE GENERA UN ARCHIVO NUEVO. QUEDANDO GUARDADO EN MEMORIA DE DISCO Y PODRA EMPLEARSE DESPUES CUANTAS VECES SE QUIERA CON EL PROGRAMA DE CALCULO RESPECTIVO.

SI EL USUARIO INSERTA ALGUN DATO QUE LA COMPUTADORA RECONOCE COMO NO VALIDO (POR EJEMPLO PULSAR UN VALOR FUERA DE RANGO O UNA

LETRA EN VEZ DE UN NUMERO), SE ESCUCHARA UN "BEEP" Y DE NUEVO SE SOLICITA EL DATO.

SE ACEPTAN MAXIMO 20 DATOS POR TRATAMIENTO, LO CUAL PODRIA PARECER UNA LIMITANTE.

SIN EMBARGO, EN LA PRACTICA ES POCO PROBABLE REBASAR ESTA CANTIDAD.

EL PROGRAMA DE CALCULO (OPCION 2) REQUIERE SIEMPRE DE UN ARCHIVO DE TRABAJO.

DE AHI QUE LO PRIMERO QUE SOLICITA ES UN NOMBRE DE ARCHIVO.

DE FIGURAR UN ARCHIVO CON EL NOMBRE QUE EL USUARIO HA INDICADO, EL PROGRAMA VERIFICA QUE ESTE SEA ADECUADO EN SU ESTRUCTURA AL DISEÑO EN CUESTION.

ES DECIR, QUE SOLO PODRA TRABAJAR ARCHIVOS QUE HAN SIDO CAPTURADOS CON EL RESPECTIVO PROGRAMA DE CAPTURA.

ESTE PROGRAMA REPORTA LA TABLA DE ANOVA, FIGURA # 9.

AL TERMINAR LA CAPTURA O TERMINAR LOS CALCULOS DE ALGUN ARCHIVO, AUTOMATICAMENTE SE REINTEGRA EL MANDO AL SUBMENU DEL DISEÑO.

SI SE DESEA TERMINAR LA SESION PULSAR LA OPCION 3 QUE LLEVA AL MENU PRINCIPAL Y DE ESTE SALIRSE COMO YA SE INDICO.

CON ESTE PROCEDIMIENTO TAN SIMPLE SE PUEDE TRABAJAR CUALQUIERA DE LOS DISEÑOS QUE INCLUYE EL SISTEMA.

LOS PROGRAMAS " LLEVAN DE LA MANO " AL USUARIO Y HABRA CONFUSION SOLO QUE NO SE TENGAN NOCIONES DEL DISEÑO QUE SE ESTA TRABAJANDO.

EN TAL CASO CONSULTE E LA GUIA DE REFERENCIA ESCRITA, EL LA QUE SE EXPONEN LOS FUNDAMENTOS, ASI COMO UN EJEMPLO DETALLADO DE CADA UNO DE LOS DISEÑOS.

PERO... !! MEJOR AUN !! , OBSERVE LA CORRIDA DE LOS PROGRAMAS EJEMPLO.

LOS PROGRAMAS EJEMPLO HAN SIDO PREPARADOS PARA SER VERDADEROS TUTORES EN LA ENSEÑANZA DEL ANOVA.

USTED SELECCIONA ALGUNO Y DE INMEDIATO TENDRA FRENTE A SI :

- PIZARRA (PANTALLA O SCREEN)
- GIS (CURSOR)
- PROFESOR (PROGRAMAS)

EN LA PANTALLA SE EXPLICA. A DETALLE, ALGUN ASPECTO ESPECIFICO O "CLASE".

DESDE PRESENTAR EL PROBLEMA, HASTA EMITIR UNA CONCLUSION.

EN LA PARTE INFERIOR DE LA PANTALLA Y DESPUES DE CADA "CLASE", APARECE PARPADEANTE :

< ENTER ó 1 >

PARA AVANZAR A OTRA CLASE PRESIONE "ENTER"; PARA RETROCEDER Y REPETIR CLASES PRESIONE "1".

A LA VEZ NOTARA QUE EN LA ULTIMA LINEA APARECE FIJO UN LETRERO EN QUE SE LEE :

2 RAPIDO

3 LENTO

4 NORMAL

5 FIN

LOS NUMEROS SE REFIEREN A LAS TECLAS PROGRAMABLES F2, F3, F4 Y F5 DE SU COMPUTADORA PERSONAL.

CON ELLAS USTED "LE MARCA LA PAUTA AL PROFESOR", PARA QUE SE VAYA MAS RAPIDO, MAS LENTO, EN TERMINO MEDIO (NORMAL), O BIEN DAR POR FINALIZADAS LAS CLASES.

EN RESUMEN USTED PUEDE AVANZAR O RETROCEDER PANTALLAS, O BIEN, REGULAR LA VELOCIDAD O TERMINAR LA PRESENTACION EN EL MOMENTO QUE LO DESEE.

*** A N O V A - M . S . C . ***

PAQUETE DE PROGRAMAS PARA LA ENSEÑANZA Y APLICACION DEL ANALISIS DE VARIANZA.

EL SISTEMA COMPRENDE :

- A) DISEÑO ALEATORIO SIMPLE
- B) DISEÑO EN BLOQUES AL AZAR
- C) DISEÑO EN CUADRO LATINO
- D) DISEÑO DE DOS FACTORES

AUTOR : MOISES SILVA CERVANTES
SEPTIEMBRE 1988

EL PROGRAMA 'ENSEÑANZA DEL ANOVA' LE ORIENTA SOBRE QUE ANALISIS DE VARIANZA APLICAR A SUS DATOS.

POSTERIORMENTE, PARA CADA UNO DE LOS DISEÑOS EXPERIMENTALES CONSIDERADOS, USTED PUEDE ACCESAR LOS PROGRAMAS EJEMPLO, QUE RESUELVEN, CON LUJO DE DETALLE, Y CON LAS ''PECULIARIDADES DIDACTICAS'' DE LA COMPUTADORA, UN EJEMPLO DE CADA UNO DE LOS CUATRO DISEÑOS EXPERIMENTALES MAS COMUNMENTE UTILIZADOS EN BIOLOGIA .

SEA USTED BIENVENIDO A UNA CLASE DE ESTADISTICA POR COMPUTADORA, TODO LO QUE TIENE QUE HACER ES ESTAR ATENTO A LAS INSTRUCCIONES QUE OPORTUNAMENTE APARECEN EN ESTA PANTALLA.

FIGURA 5. PANTALLAS DE PRESENTACION DEL SISTEMA "ANOVA-MSC"

VAMOS A SELECCIONAR UNO DE CUATRO POSIBLES DISEÑOS DEL ANALISIS DE VARIANZA.

USTED UNICAMENTE RESPONDA A LAS PREGUNTAS QUE SE LE HACEN :

SU EXPERIMENTO COMPRENDE UN SOLO FACTOR ? S ó N

SOLO CUENTA CON DOS TRATAMIENTOS ? S ó N

LE CONVIENE APLICAR MEJOR LA PRUEBA 't' DE STUDENT (AUNQUE TAMBIEN PUEDE EMPLEAR EL ANOVA SIMPLE).

VAMOS A SELECCIONAR UNO DE CUATRO POSIBLES DISEÑOS DEL ANALISIS DE VARIANZA.

USTED UNICAMENTE RESPONDA A LAS PREGUNTAS QUE SE LE HACEN :

SU EXPERIMENTO COMPRENDE UN SOLO FACTOR ? S ó N

SOLO CUENTA CON DOS TRATAMIENTOS ? S ó N

SON UNIFORMES LAS UNID. EXPERIMENTALES ? S ó N

SE BLOQUEA EN UNA SOLA DIRECCION ? S ó N

SE BLOQUEA EN DOS DIRECCIONES ? S ó N

AL HABER BLOQUES DE DOS TIPOS. SE APLICA EL DISEÑO EN CUADRO LATINO.

FIGURA 6. PANTALLAS DEL PROGRAMA "ENSEÑANZA DEL ANOVA"

DISEÑOS DEL SISTEMA

- 1.- ALEATORIO SIMPLE
- 2.- BLOQUES AL AZAR
- 3.- CUADRO LATINO
- 4.- DOS FACTORES
- 5.- PROG. DE ENSEÑANZA
- 6.- FIN < SALIR DE SISTEMA >

INDIQUE EL # DE LA OPCION DESEADA

FIGURA 7 MENU GENERAL DEL SISTEMA
"ANOVA-MSC"

ANOVA EN CUADRO LATINO

- 1.- CAPTURA DE DATOS
- 2.- TABLA DE ANOVA
- 3.- PROGRAMA EJEMPLO

- 4.- FIN < MENU PRINCIPAL >

INDIQUE EL # DE LA OPCION DESEADA

FIGURA B SUBMENU DEL SISTEMA
"ANOVA-MS"

CALCULOS PARA UN DISEÑO EN CUADRO LATINO
Análisis de varianza de los datos del archivo : TRIGO

FTE DE VARIACION	gl	SUM. CUAD.	CUAD. MED.	F
TRATAMIENTOS	3	78.92468	26.30823	58.02046
REGLONES	3	1.954834	.6516113	
COLUMNAS	3	6.799805	2.266602	
ERROR	6	2.720581	.4534302	
TOTAL	15	90.3999		

PRESIONE < RETURN >

FIGURA 9. TABLA DE ANOVA REPORTADA POR EL SISTEMA "ANOVA-MSC"

CONCLUSIONES

EN ESTE TRABAJO SE ABORDO EL PROBLEMA DE COMO ENSEÑAR EL ANOVA, A TRAVES DE UN PROGRAMA DE COMPUTO.

PARA TAL FIN, SE DISCUTIO EN LA INTRODUCCION LA PROBLEMÁTICA DE LA ENSEÑANZA DE LOS METODOS CUANTITATIVOS EN BIOLOGIA.

SE HIZO INCAPIE EN COMO ABORDAR ESTADISTICAMENTE UN PROBLEMA, VALIENDOSE DE LA COMPUTADORA COMO IMPORTANTE AUXILIAR, Y EL VALOR DE LOS MODELOS COMO GENERADORES DE IDEAS.

SE ABORDO EL PROBLEMA DEL DISEÑO DE EXPERIMENTOS, PRESENTANDO LOS PASOS FUNDAMENTALES PARA SU IMPLEMENTACION.

ADEMAS SE DISCUTIO EL PROBLEMA DE LA ENSEÑANZA DEL CONCEPTO DE ALEATORIZACION, SUGIRIENDO QUE EN CLASE SE PRACTIQUE ESTE PROCESO ANTES DE INICIAR LOS DISEÑOS EXPERIMENTALES.

POR OTRA PARTE SE PRESENTO UN DIAGRAMA DE FLUJO, COMO LA GUIA DE TRABAJO QUE SE SIGUIO PARA LA IMPLEMENTACION DE LOS PROGRAMAS DE COMPUTO.

LA PRESENTACION DE LOS DIFERENTES DISEÑOS SIGUIO EL SIGUIENTE FORMATO :

- 1) PLANTEAMIENTO DEL PROBLEMA.
- 2) TABLA DE DATOS.
- 3) PLANTEAMIENTO DE LA HIPOTESIS.
- 4) DETALLE DEL PROCEDIMIENTO DE CALCULO.
- 5) REGLA DE DECISION.
- 6) CONCLUSIONES.

LA DESCRIPCION DEL SISTEMA "ANOVA-MSC", ABORDO LAS SIGUIENTES CARACTERISTICAS :

- 1) INICIAR SECCION EN EL SISTEMA.
- 2) MANEJO DE LOS PROGRAMAS TUTORIALES.
- 3) MANEJO DE LOS PROGRAMAS DE CALCULO.

FINALMENTE, LAS EXPERIENCIAS DE TRABAJO QUE SE PUEDEN ABORDAR EN EL FUTURO SON :

A) PODER VISUALIZAR LAS SUPOSICIONES DEL ANALISIS DE VARIANZA (NORMALIDAD, HOMOCEDASTICIDAD, INDEPENDENCIA Y ADITIVIDAD), POR METODOS GRAFICOS Y A TRAVES DE ALGORITMOS QUE PUEDAN COMPLEMENTAR EL TRABAJO AQUI DESARROLLADO.

LO ANTERIOR IMPLICA ABORDAR LOS CLASICOS MODELOS FISHERIANOS, COMO MODELOS DE SUPERFICIE DE RESPUESTA A TRAVES DE LA METODOLOGIA DEL MODELO DE REGRESION LINEAL SEMI-CUANTITATIVO (KIRK, 1982).

LA IMPORTANCIA DE EVALUAR LOS SUPUESTOS DEL MODELO, HAN SIDO DISCUTIDOS POR CURTS (1984 Y 1987), YA QUE COMO DISCUTE ESTE AUTOR: EL SESGO QUE PUEDEN TENER LAS PRUEBAS ESTADISTICAS, COMO LA DE "t" Y LA DE "F", PUEDEN DEBERSE FUNDAMENTALMENTE A PROBLEMAS DE DISTRIBUCION NORMAL O A LA PRESENCIA DE CASOS ABERRANTES E INFLUENCIALES.

B) LA PROBABILIDAD DE EXPLORAR, PREVIO A LOS SUPUESTOS DEL MODELO DEL ANALISIS DE VARIANZA, EL CONJUNTO DE DATOS, A TRAVES DE LAS PRUEBAS INFORMALES O DE BAJA SOFISTICACION, COMO LAS PROPUESTAS POR TUKEY (1977), Y CUYOS PRINCIPIOS FUNDAMENTALES HAN SIDO RECIENTEMENTE PUBLICADOS EN ESPAÑOL POR CURTS Y SILVA (1988).

C) LA IDEA DE ELABORAR PROGRAMAS TUTORIALES COMO ESTE, EXTENDIENDOLOS A OTROS PROCESOS MAS COMPLEJOS, COMO EL DE LA REGRESION MULTIPLE Y EL ANALISIS MULTIVARIADO.

ES ASI, QUE CON EL APOYO DE SISTEMAS TUTORIALES COMPUTARIZADOS SE PUEDE, COMO BIEN AFIRMA CURTS (1988), QUE EL ALUMNO ABORDE LA

PROBLEMATICA BIOLÓGICA Y LOS
INSTRUMENTOS TEÓRICO-METODOLÓGICOS
QUE LE SEAN NECESARIOS PARA
COMPRENDER :

- QUE ES LA BIOLOGÍA
- COMO SE ESTUDIA
- QUE DISCIPLINAS APOYAN ESTE CO-
NOCIMIENTO Y LOS PROCEDIMIENTOS
QUE ESTOS APORTAN.

PARA RESOLVER CUALQUIER PROBLEMA PRIMERO HAY QUE DETERMINAR LOS OBJETIVOS DEL MISMO.

ENSEGUIDA ELEGIR Y APLICAR ALGUN METODO QUE PERMITA LA RESOLUCION.

EL SIGUIENTE ESQUEMA, ILUSTRALOS PASOS QUE HABRIA QUE SEGUIR PARA RESOLVER UN PROBLEMA DADO :

- 1) DEFINIR EL PROBLEMA (OBJETIVOS QUE SE DESEAN).
- 2) DETERMINAR LO QUE PAREZCA SER EL MEJOR METODO A UTILIZAR (DE ACUERDO A LA EXPERIENCIA, INVESTIGACION, AYUDA).
- 3) PROPORCIONAR LAS ENTRADAS NECESARIAS PARA LA RESOLUCION DE EL PROBLEMA.
- 4) UTILIZAR EL METODO SELECCIONADO (OPERACIONES).
- 5) COMPROBAR LOS RESULTADOS (SALIDA), PARA VER SI EL METODO SELECCIONADO FUNCIONA EFICIENTEMENTE.
- 6) SI LOS RESULTADOS NO SON ADECUADOS, VOLVER AL PASO 2.

EN TERMINOS DE INFORMACION, SE DICE QUE LAS OPERACIONES TRANSFORMAN LOS DATOS DE ENTRADA EN INFORMACION UTIL (RESULTADOS).

AUNQUE NO SE UTILIZA LA SIMBOLOGIA CONVENCIONAL, A LA DERECHA APARECE UN DIAGRAMA DE FLUJO QUE RESUME LO EXPUESTO EN ESTA COLUMNA.

UN DIAGRAMA DE FLUJO SE LEE SIEMPRE DE ARRIBA HACIA ABAJO. SIGUIENDO EL FLUJO MARCADO POR LAS LINEAS.

SOLO EN EL CASO DE UNA DESVIACION CONDICIONAL, SE ALTERA EL FLUJO NORMAL .



EN NUESTRO DIAGRAMA OCURRE UN SALTO CONDICIONAL.

CUANDO NO SE CUMPLE QUE LOS RESULTADOS SEAN CORRECTOS, ENTONCES, SE VUELVE A REVISAR EL METODO DE RESOLUCION.

EL MANEJO DE LOS DIAGRAMAS DE FLUJO ES TODA UNA DISCIPLINA. POR LA QUE SE ESQUEMATIZA UN DETERMINADO PROCEDIMIENTO .

EL ESFUERZO INVERTIDO EN PLASMAR EL DIAGRAMA, SE COMENZA AL LOGRAR TAL CLARIDAD DE LOS PASOS A SEGUIR EN LA RESOLUCION DEL PROBLEMA.

ES BASTA LA INFORMACION AL RESPECTO, Y BIEN VALDRIA LA PENA ENCAUSAR AL ESTUDIANTE A LA COMPRENCION DE LOS MISMOS.

EN SEGUIDA SE HACE UNA BREVE REVISION DE LA SIMBOLOGIA COMUNMENTE UTILIZADA.

SE EXPLICAN SOLO UNOS CUANTOS SIMBOLOS, SIN EMBARGO, SON SUFICIENTES PARA DESGLOZAR GLOBALMENTE CUALQUIER PROBLEMA.

YA QUE LO IMPORTANTE ES LA VISION LOGICO-ANALITICA DE QUIEN ELABORA EL DIAGRAMA, NO ES NECESARIO APRENDER TODO ACERCA DE LA DIAGRAMACION.

MAS BIEN, LO ACONSEJABLE ES EJERCITARSE EN PRESENTAR DE ESTA FORMA UN ALGORITMO (LA SERIE DE PASOS QUE RESUELVEN UN PROBLEMA DADO).

SI BIEN LA MERA DIAGRAMACION ES YA UN GRAN SOPORTE, EN EL DIARIO DESEMPEÑO DEL PROGRAMADOR DE COMPUTADORAS, SE SIGUEN LOS SIGUIENTES PASOS :

- 1) DEFINICION DEL PROBLEMA
- 2) ALGORITMO
- 3) DIAGRAMA DE FLUJO
- 4) PROGRAMA DE COMPUTADORA

EN LA FIGURA # 10 SE MUESTRAN EL ALGORITMO (PASOS NUMERADOS) Y EL DIAGRAMA DE FLUJO, QUE RESUELVEN UN PROBLEMA MUY SIMPLE :

— CALCULAR EL VALOR DE "C", COMO LA SUMA DE "A" + "B" .

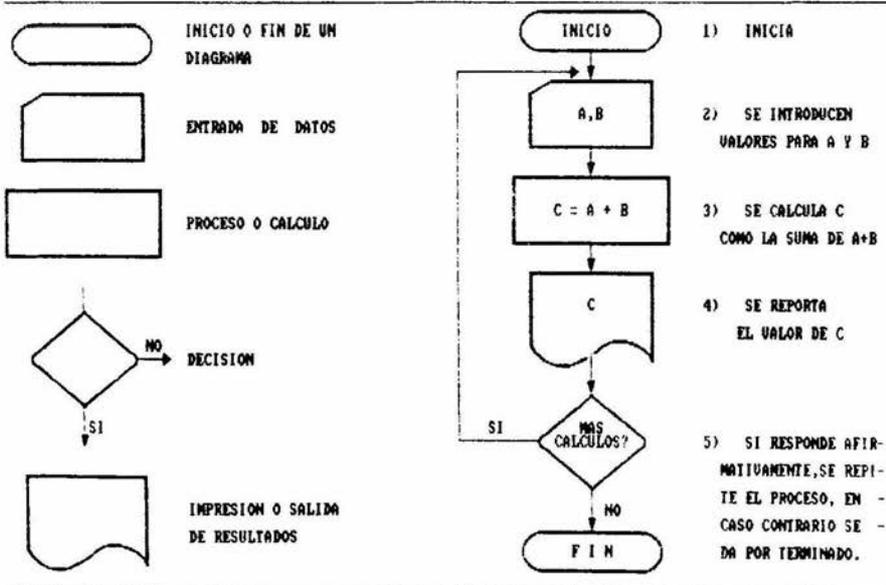


FIGURA 10. ALGUNOS DE LOS PRINCIPALES SIMBOLOS DE DIAGRAMACION

EN LA FIGURA # 3 SE EXPONE UN DIAGRAMA DE FLUJO, CON UN PROPOSITO MAS COMPLICADO QUE LA SUMA DE "A" + "B".

SIN EMBARGO, EN PRINCIPIO ES EL MISMO CRITERIO :

SEGUIR LAS FLECHAS DESDE DONDE SE MARCA EL INICIO, Y DE ACUERDO A LAS RESPUESTAS EMITIDAS, EN LAS DIFERENTES PREGUNTAS(DECISIONES) .

POR EJEMPLO, SUPONGA QUE DESEA INVESTIGAR EL EFECTO DE CIERTA DROGA, SOBRE ANIMALES DE LABORATORIO.

OBSERVE LA FIGURA # 3 Y TRATE DE ADAPTARLA A LOS PLANTEAMIENTOS SIGUIENTES, PARA TOMAR UNA DECISION DETERMINADA.

- 1) USTED CONSIDERA QUE LA EDAD, EL PESO Y EL SEXO INFLUYEN EN LOS RESULTADOS EXPERIMENTALES, TANTO EN FORMA AISLADA COMO DE CONJUNTO (INTERACCION) .
ES DECIR, QUE SEPARARA SUS LOTES EXPERIMENTALES, POR GRUPOS DE CIERTA EDAD, CIERTO PESO Y UN SEXO DETERMINADO.
EN TAL CASO ESTA CONTEMPLANDO 3 FACTORES EN SU EXPERIMENTO :
"A" CON "a" NIVELES O AGRUPAMIENTOS DIFERENTES, "B" CON "b" NIVELES Y "C" CON "c" NIVELES.
UN TRATAMIENTO EN PARTICULAR, SERA LA COMBINACION DE "i", "j" Y "k" NIVELES DE LOS FACTORES "A", "B" Y "C" RESPECTIVAMENTE.
EN RESUMIDAS CUENTAS USTED TENDRA QUE APLICAR EL DISEÑO DE TRES FACTORES.
- 2) CONSIDERA QUE SOLO INFLUYEN EL PESO Y LA EDAD, TANTO EN FORMA AISLADA COMO DE INTERACCION.
SU PLANTEAMIENTO ES MUY SEMEJANTE AL ANTERIOR, SOLO QUE, AHORA SOLO CONTEMPLA 2 FACTORES CON SUS RESPECTIVOS NIVELES.
UN TRATAMIENTO SERA LA COMBINACION DE "i" NIVEL DE "A", CON "j" NIVEL DE "B".

USTED APLICARA EL DISEÑO DE 2 FACTORES.

- 3) SI CONSIDERA QUE SOLO INFLUYE LA EDAD, SOLO 1 FACTOR, ENTONCES, DE ACUERDO A OTROS CRITERIOS. TENDRA ELEGIR ENTRE APLICAR LA PRUEBA DE "t", EL ANOVA SIMPLE, EL ANOVA EN BLOQUES O EL ANOVA EN CUADRO LATINO.
ESTOS CRITERIOS SON LOS SIGUIENTES:
 - LA PRUEBA DE "t" O, INCLUSO, EL ANOVA SIMPLE SI SOLO CUENTA CON 2 TRATAMIENTOS DIFERENTES.
POR EJEMPLO, APLICAR A UN LOTE UNA DOSIS CONCENTRADA Y A OTRO UNA DOSIS LIGERA DE LA DROGA) .
 - EL ANOVA SIMPLE SI APLICARA MAS DE 2 DOSIS, A LA VEZ QUE SUS LOTES EXPERIMENTALES, O UNIDADES EXPERIMENTALES, SEAN UNIFORMES : MISMA EDAD, MISMO PESO, MISMO SEXO, ETC. DICHO DE OTRA MANERA LOS GRUPOS SON INDEPENDIENTES Y HOMOGENEOS.
 - EL ANOVA EN BLOQUES SI, POR EJEMPLO, SEPARARA SUS LOTES EXPERIMENTALES EN GRUPOS HOMOGENEOS DE PESO (BLOQUES). USTED CONSIDERA COMO UNICO FACTOR LA CONCENTRACION DE DROGA Y NO LE INTERESA AVERIGUAR EL EFECTO DEL PESO, POR LO QUE PROCEDE A " BLOQUEAR EN UN SENTIDO " .
 - EL ANOVA EN CUADRO LATINO SI LE INTERESA HACER UN " BLOQUEO EN DOS DIRECCIONES " .
POR EJEMPLO, QUE ADEMAS DE EVITAR EL EFECTO DEL PESO, DESEA ELIMINAR EL DE LA EDAD. DEBERA TENER TANTOS BLOQUES DE PESO Y TANTOS DE EDAD COMO CONCENTRACIONES DIFERENTES DE DROGA QUIERA APLICAR. A CAUSA DE ESTO ULTIMO SE GENERA UNA CONFORMACION EN CUADRO, DE DONDE DERIVA EL NOMBRE DEL DISEÑO .

BIBLIOGRAFIA

- Anderson, V. & Mc. Lean, R. (1974), "Design of Experiments" Marcel Dekker Inc. , New York.
- Butler, M. J. , Dwas, M. , Joiner, B. , Swanson, J. (1976). "Report on the role of computers", in J. R. O'fallon Service (eds), Modular Instruction in Statistics (Report of the American Statistical Association, Study of Modular Instruction), Washington, DC, American Statistical Association.
- Ceibson, R. J. (1968), "Experimental design or happiness is planning the experiment", Bioscience volume 18 , number 3 .
- Collis, Betty (1983), "Teaching descriptive and inferential statistics, using a classroom microcomputer", Mathematics teacher, volume 5, number 76 .
- Curts, J. (1984), "Introducción al análisis de residuos en biología", Biótica, Vol. 9-3, 271-278 .
- Curts, J. (1986), "Teaching college biology students the simple linear regression model using an interactive microcomputer graphics software package", Dissertation Abstracts International, volume 46, number 7 .
- Curts, J. (1988), "Biología y modelos matemáticos : aspectos sobre su enseñanza", Biología, Vol. 17, No. 1-4 .
- Curts, Alcantara y Chiappa (1987), "Introducción al análisis de datos multidimensionales", Ciencias, Revista de Difusión, No. 1, U.N.A.M.
- Curts, J. y Rascón, G. (1988), "El papel de las computadoras en la enseñanza del análisis estadístico de datos", Memorias de la Cuarta Conferencia Internacional : Las Computadoras en Instituciones de Educación Superior, U.N.A.M. , México, D.F.
- Curts y Silva (1988) , "Introducción a los métodos cuantitativos en psicología", Cap. 10, Ed. Trillas, México, D.F.
- Daniel, W. (1980) , "Bioestadística : base para el análisis de las ciencias de la salud". C.E.C.S.A. , México D.F.
- Davies, O. L. (Ed.) , "design and analysis of industrial experiments" , Hafner Publishing Company, New York.
- Finney, D. J. (1960) , "An introduction to the theory of experimental design", The University of Chicago Press, Chicago.
- Hicks, C. R. (1973) , "Fundamental concepts in the design of experiments", (2nd ed.), Holt Rineart, Winston, New York.

-
- IBM (1984) , "BASIC , manual de consulta", IBM, Barcelona.
 - IBM (1985), "Disk operating system version 3.1", IBM, Florida.
 - Jordan, E. (1982), "A systems perspective of statistics. proceedings of the section on statistical education", Paper Presented at the Meeting of the American Statistical Association, Cincinnati, Washington.
 - Kempthorne , O. (1980) , "The teaching of statistics : content versus form" , The American Statistical Association , volume 76, number 1-8.
 - Kirk, R. E. (1982) , "Experimental design : procedures for the behavioral sciences", (2nd edition) , Brooks Cole Publishing Co. , Monterrey , CA.
 - Krumm, R. (1988) , "La potencia de word-perfect 4.2", Mc. Graw Hill, Madrid.
 - Lehman, J. D. (1983) , "Microcomputer simulations for the biology classroom" , The Journal of computers in mathematics and science teaching" , vol. 2 , no. 4 , 10-13.
 - Mendez, I. (1981), "La estadística y la biología", Monografías del Instituto de Investigaciones en Matemáticas aplicadas y Sistemas, U.N.A.M. , México, D.F.
 - Reyes, C. (1985), "Diseño de experimentos aplicados", Trillas, México D. F.
 - Schafer, W. , Dayton, M. (1979) , N E C C , 336-338.
 - Snedecor, G.W. (1981) , "Métodos estadísticos", C.E.C.S.A. , México, D. F.
 - Spector, P.E. (1981), "Research Designs", Spage Publications, Beverly Hills.
 - Steal, R. & Torrie, J. (1980) , "Principies and procedures of statistics (a biometrical approach)", Mc. Graw Hill, U.S.A.
 - Tukey, J. (1977) , "Exploratory data analisis", Addison-Wesley Publishing Company, Reading, MA.
 - Yamane, T. (1979) , "Estadística" , Harla , México D. F.