ANOVA

Análisis de la Varianza en diseño de experimentos

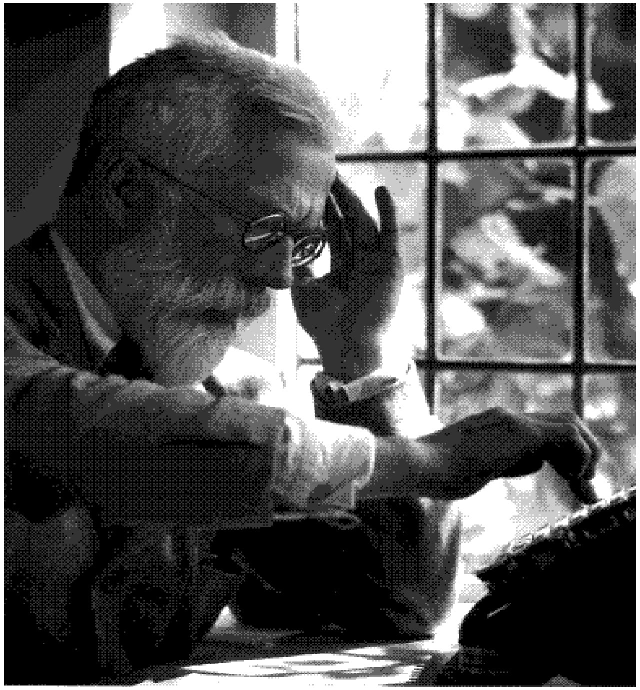
NATURALEZA DEL DISEÑO EXPERIMENTAL

El diseño experimental tiene sus orígenes en los trabajos de Ronald Aylmer Fisher (1890 – 1962) desarrollados en la Estación Agrícola Experimental de Rothamsted, en el Reino Unido, donde introduce los conceptos de aleatorización y el análisis de varianza.

Los cimientos de la materia han llegado a conocerse por el título de su libro The Design of Experiments (1935).

NATURALEZA DEL DISEÑO EXPERIMENTAL

Desde entonces la teoría del diseño experimental ha sido desarrollada y ampliada.
Aplicaciones de esta teoría se encuentran hoy en las investigaciones en ciencias naturales, ingeniería y casi todas las ramas de las ciencias sociales.



DISEÑOS EXPERIMENTALES

Un diseño estadístico de experimentos (DEE) es el proceso de *planificación* del experimento mediante el cual se *recolectan* datos apropiados que al ser *analizados* por métodos estadísticos estos nos conducen a *conclusiones* metodológicamente válidas.

***Es la secuencia completa de pasos a realizar para asegurar que los datos conduzcan a conclusiones válidas con respecto al problema establecido*.**

DISEÑOS EXPERIMENTALES

SURGEN PREGUNTAS COMO:

¿Cómo se va a medir el efecto?,
¿Cuáles son las variables a analizar?
¿Qué otras variables afectan a las variables bajo análisis?
¿Cuántas veces deberá ejecutarse el experimento?
¿Cuál será la forma de análisis?

PROPÓSITOS DE UN DEE

- Proporcionar métodos que permitan obtener *la mayor cantidad* de información en forma objetiva, confiable y *al mínimo costo* (dinero, personal, tiempo, etc.).

- *Controlar la variabilidad* que está presente en todo experimento y que afecta a sus resultados.

DEFINICIONES

UNIDAD EXPERIMENTAL

Es la parte más pequeña de material experimental a la que se le aplica el tratamiento y en donde se mide o registra la variable que se investiga.

La unidad experimental por lo general esta conformada por:

-en los experimentos pecuarios *un animal* (vaca, cerdo, pato, etc.)

-en los experimentos forestales, la mayoría de los casos esta conformada por *un árbol.*

-en experimentos agrícolas suele ser *una parcela* de tierra en lugar de una planta individual.

-en geología muestra de *roca* o *suelo*, un *fósil*, un *corte delgado*, etc.

TRATAMIENTOS, FACTORES, NIVELES

- Tratamiento: Conjunto de condiciones experimentales o procedimientos que se van a aplicar a una unidad experimental en un diseño elegido y cuyos efectos van a ser registrados (respuesta).

- Observación: valor que asume la variable respuesta en una determinada realización.

TRATAMIENTOS, FACTORES, NIVELES

- Factor: Variable controlada por el experimentador o variable independiente. Se estudia su efecto sobre la variable dependiente o respuesta.

- Nivel del Factor: es cada una de las categorías, valores o formas específicas del Factor.

EJEMPLOS

Factor: Tipos de Riego, Dosis del fertilizante, variedades de cultivo, manejo de crianzas, ración alimentaria, profundidad del sembrado, distanciamiento entre plantas, etc.
Se acostumbra simbolizar con A, B, C.

Niveles: Se acostumbra simbolizar por la letras minúscula del factor y con un subíndice que representan al nivel.

Ejemplo:

A: Tipos de Riego: Goteo Aspersión Secano

Niveles : a_0 a_1 a_2

TIPOS DE FACTORES

1.-Factores Cuantitativos. Son aquellos factores cuyos niveles son cantidades numéricas.

Ejemplo:

Factor A: Dosis de Fertilización.

Niveles : 10 kg/Ha (a_0), 20kg/Ha(a_1), 30kg/Ha(a_2)

2.- Factores Cualitativos. Son aquellos factores cuyos niveles son procedimientos o cualidades.

Ejemplo:

Factor A: Variedad de Cultivo

Niveles: Variedad 1(a_0), Variedad 2(a_1).

EXPERIMENTOS UNIFACTORIALES Y MULTIFACTORIALES

- Unifactorial: un solo factor con sus respectivos niveles.

  Cada nivel del factor es un tratamiento.

- Multifactorial: dos o más factores.

  Un tratamiento es la combinación de un nivel de cada factor.

	RIEGO Nivel 1 (500mm)	RIEGO Nivel 2 (700mm)	RIEGO Nivel 3 (900mm)
FERTILIZACIÓN Nivel 1 (100kg UREA)	10 11 13 14	11 12 14 15	13 14 16 17
FERTILIZACIÓN Nivel 2 (200kg UREA)	14 15 17 18	15 16 18 19	16 17 19 20
FERTILIZACIÓN Nivel3 (300kg UREA)	17 18 20 21	17 18 20 21	18 19 21 22

Tn grano/ha

PRINCIPIOS BÁSICOS DEL DISEÑO EXPERIMENTAL

- Repetición
- Aleatorización
- Control Local

REPETICIÓN

Es la *reproducción* o *réplica* del experimento básico (asignación de un tratamiento a una unidad experimental). Son observaciones de un mismo tratamiento en diferentes unidades experimentales.

Las principales razones por las cuales es deseable la repetición son:

- Permite estimar el ***error experimental***

- Aumenta el alcance de la inferencia del experimento por selección y uso apropiado de las u.e.

ALEATORIZACIÓN

Consiste en la *asignación al azar* de los tratamientos a las unidades experimentales con el propósito de asegurar que un determinado tratamiento no presente sesgo.

La aleatorización *hace válidos* los procesos de inferencia estadísticas y los supuestos asegurando conclusiones confiables.

Una de las suposiciones más frecuentes es que las observaciones, o los *errores en ellas están distribuidos independientemente*.

CONTROL LOCAL

(Control del error experimental) consiste en tomar medidas (acciones del experimentador) para hacer el diseño experimental más eficiente, de tal manera que pueda *permitir la reducción del error experimental* y así hacerla más sensible a cualquier prueba de significación.

El error experimental puede reflejar variación del material experimental, errores de medición u observación, errores de experimentación, etc.

CONTROL LOCAL

El error experimental se puede reducir mediante:

♦ El uso de un diseño experimental apropiado.

♦ La selección minuciosa del material a usar (lo más homogéneo posible).

♦ El incremento del número de repeticiones en el experimento.

♦ El perfeccionamiento de la técnica experimental y el mayor cuidado al dirigir el experimento.

♦ La utilización de la información proporcionada por variables relacionadas a la variable en estudio (uso de la técnica de covarianza).

TIPOS DE DISEÑOS

- Diseño Completamente Aleatorizado **(DCA)**
- Diseño en Bloques completos aleatorizados **(DBCA)**
- Diseño en Cuadrado Latino **(DCL)**
- Diseño en Parcelas Divididas **(DPD)**
- Diseño en Cuadrado Grecolatino
- Diseño factorial
- Diseño factorial fraccionario
- Otros

DISEÑO COMPLETAMENTE ALEATORIZADO (DCA)

En este diseño, los tratamientos en estudio se distribuyen al azar en forma *irrestricta* sobre todas las unidades experimentales; siendo el número de repeticiones por tratamiento **igual** ó diferente.

Este diseño se emplea cuando la variabilidad en todo el material experimental es relativamente pequeño y uniformemente distribuido.

VENTAJAS Y DESVENTAJAS DEL DCA

- *Ventajas:*

Fácil de planear y analizar; además es flexible en el empleo del número de tratamientos y repeticiones. Finalmente, permite tener dentro del análisis de varianza el máximo número de grados de libertad para la suma de cuadrados del error.

- *Desventaja:*

La principal desventaja que presenta este diseño está relacionado a la homogeneidad del material experimental; el cual es difícil de encontrar en experimentos de campo, por lo que su uso se restringe con mucha frecuencia a experimentos de laboratorio, o donde se pueda tener control de los efectos no considerados en el estudio (ambiente, temperatura, luz, etc.).

Datos

Grupo 1	Grupo 2	Grupo 3	…	Grupo t
Y_{11}	Y_{21}	Y_{31}	…	Y_{t1}
Y_{12}	Y_{22}	Y_{32}	…	Y_{t2}
Y_{13}	Y_{23}	Y_{33}	…	Y_{t3}
:	:	:		:
Y_{1n_1}	Y_{2n_2}	Y_{3n_3}	…	Y_{tn_k}

El Grupo 1 tiene n_1 observaciones, el Grupo 2 tiene n_2 observaciones, y así sucesivamente.

Comenzaremos con el caso en que todos los tratamientos tienen igual cantidad de repeticiones. (Diseño balanceado) $n_1 = n_2 = ... = n_t = r$

DCA Fijo o Modelo I

$$Y_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1, \ldots, t; \quad j = 1, \ldots, r$$

Donde:

Y_{ij}: Es el valor observado de la *j-ésima* observación del *i-ésimo* tratamiento.

μ_i : Es el valor medio poblacional del *i-ésimo* tratamiento.

ε_{ij}: Es el error aleatorio de la *j-ésima* observación correspondiente al i-ésimo tratamiento.

O si en el modelo anterior denotamos a $\mu_i = \mu + \tau_i \quad i = 1, \ldots, t$

se obtiene la siguiente forma alternativa del modelo

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad i = 1, \ldots, t; \quad j = 1, \ldots, r$$

Donde:

Y_{ij}: Es el valor observado de la *j-ésima* observación que recibió el *i-ésimo* tratamiento.

μ : Es el valor medio poblacional general.

τ_i: Es el efecto medio poblacional del *i-ésimo* tratamiento.

ε_{ij}: Es el error aleatorio de la *j-ésima* observación correspondiente al i-ésimo tratamiento.

SUPUESTOS PARA EL MODELO $Y_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1, \ldots, t; \quad j = 1, \ldots, r$

1. Para cada tratamiento la variable aleatoria respuesta Y_{ij} se distribuye normalmente . Para cada T_i $Y_{ij} \sim N(\mu_i; \sigma_i^2)$

2. Las Y_{ij} son variables aleatorias independiente entre si

$$\mathrm{cov}(Y_{ij} Y_{kl}) = 0 \quad \forall i \neq k; \quad \forall j \neq l$$

3. Las varianzas de los t tratamientos son iguales (homoscedasticidad)

$$\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_t^2 = \sigma^2$$

En consecuencia

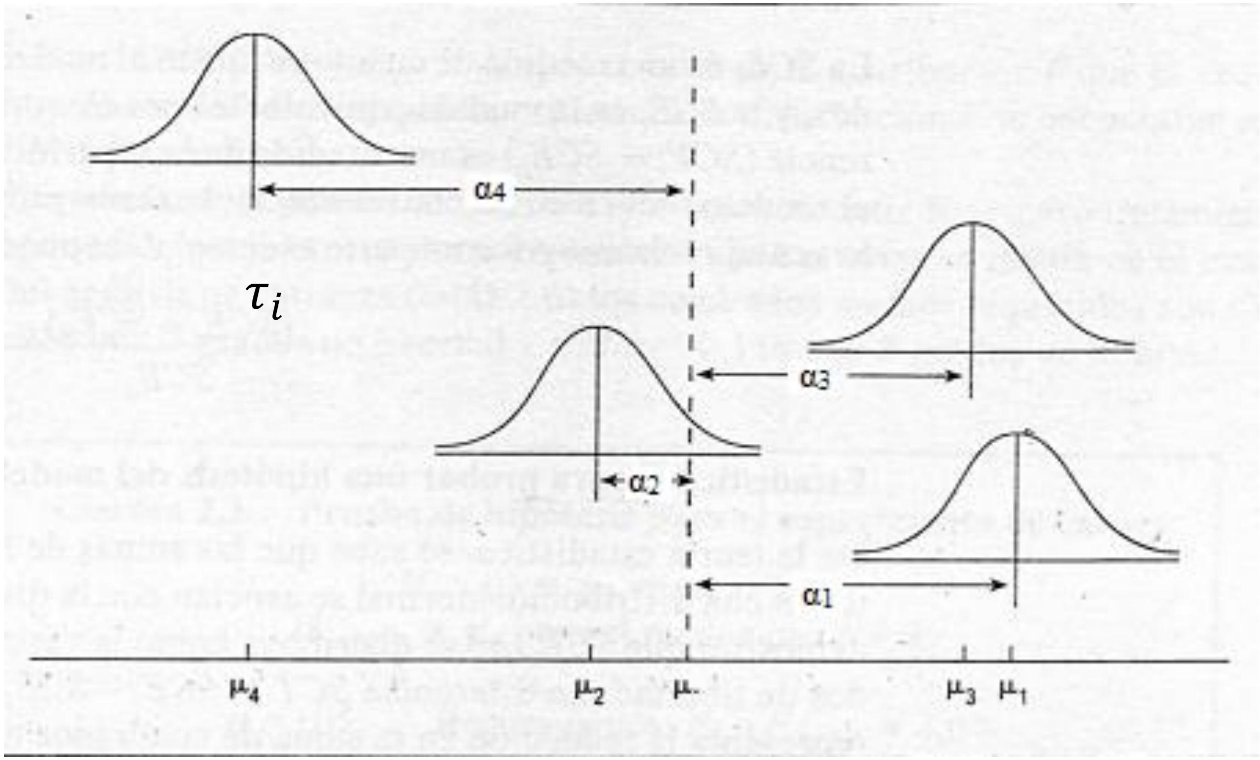
$$\varepsilon_{ij} \sim N(0; \sigma^2)$$

$$E(y_{ij}) = \mu_i$$

$$Var(Y_{ij}) = \sigma^2$$

DISEÑO COMPLETAMENTE ALEATORIZADO (DCA)

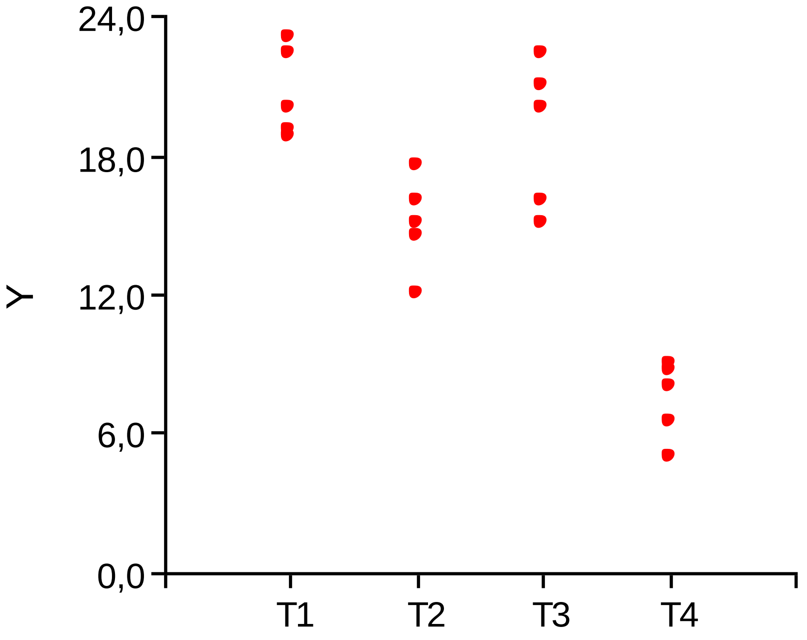
En teoría tenemos las distribuciones correspondientes a los tratamientos para una determinada variable respuesta

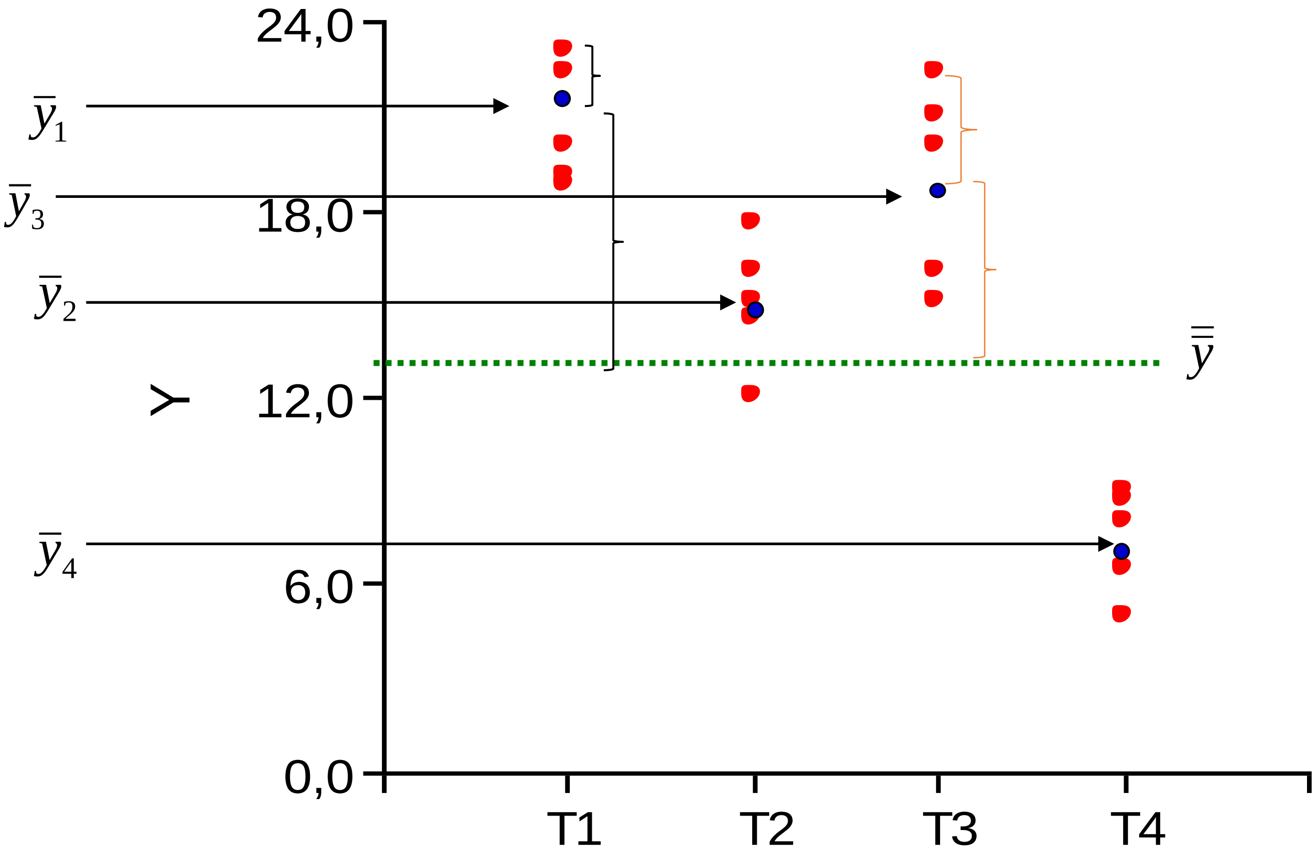


$$\tau_i = \alpha_i = \mu_i - \mu$$

DISEÑO COMPLETAMENTE ALEATORIZADO (DCA)

En la práctica sólo tenemos las muestras correspondientes a los tratamientos y sus correspondientes valores respuesta





24,0
18,0
12,0
6,0
0,0
Y
T1
T2
T3
T4
$\bar{y}_1$
$\bar{y}_3$
$\bar{y}_2$
$\bar{y}_4$
$\bar{\bar{y}}$

PARTICIÓN DE LA SUMA DE CUADRADOS

$$\sum_i \sum_j (y_{ij} - \bar{\bar{y}})^2 = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 + \sum_i \sum_j (\bar{y}_i - \bar{\bar{y}})^2$$

Estas sumas de cuadrados se llaman respectivamente

$$SC_{tot} = SC_{ee} + SC_{Trat} \qquad \text{o} \qquad SC_{tot} = SC_{dentro} + SC_{entre}$$

$SC_{tot} \to r.t - 1$ $SC_{ee} \to t.(r-1)$ $SC_{Trat} \to (t-1)$

rt: cant. de u.e.

y tienen distribución Chi cuadrado y son independientes por lo tanto los grados de libertad son:

Dividiendo la suma de cuadrados por sus respectivos grados de libertad se obtienen los *cuadrados medios* que son varianzas muestrales.

$$CM_{error} = \frac{SC_{error}}{t(r-1)} = \hat{\sigma}^2 \qquad CM_{trat} = \frac{SC_{trat}}{t-1}$$

Estas varianzas muestrales son estimadores de varianzas poblacionales (parámetros) que están relacionadas con las fuentes de variación correspondientes

$$E(CM_{ee}) = \sigma^2 \qquad E(CM_{trat.}) = \sigma^2 + r\frac{\sum_{i=1}^{t} {\tau_i}^2}{t-1}$$

El cociente entre éstos tiene distribución F

$$F = \frac{CM_{trat}}{CM_{ee}} \approx F_{(t-1;t(r-1))}$$

TABLA ANOVA

Fuentes Variación	Sumas Cuadrados	g.l	Cuadrados Medios	*F*
Entre/Trat.	SCTrat.	t-1	SCTrat/g.lTr (I)	*Fcalc=* I/II
Dentro/Trat. (EE)	SCError	t(r-1)	SCError/g.lee (II)	
Total	SCTotal	tr - 1		

Hipótesis en términos de los efectos

$H_0 : \tau_i = 0$ i = 1,2, ..., t (Todos los tratamientos tienen el mismo efecto sobre la variable rta.)
H_1: algún $\tau_i \neq 0$ (No todos los tratamientos tienen el mismo efecto sobre la variable en estudio).

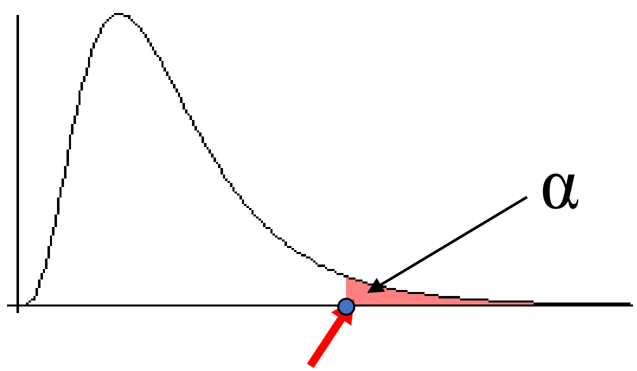
Dócimas en términos de los promedios

H_0: $\mu_1 = \mu_2 = ...= \mu_t = \mu$
H_1: algún $\mu_i \neq \mu$.

PASOS si se realiza a mano la dócima en el DCA

- **Paso Nº 1 Planteo de las hipótesis:**
- **Paso Nº 2 Elección del nivel de significación α**
- **Paso Nº 3 Variable pivotal :** $F = \dfrac{CM_{trat}}{CM_{ee}} \approx F_{((t-1);t(r-1))}$
- **Paso Nº 4 Región Crítica y regla de decisión:**

La región crítica $F_{((t-1);\, t(r-1))} \geq F_{crit}$



$Fcrit = F\ ((t-1);\ t(r-1));\ \alpha$

- **Regla de decisión**

Rechazo H_0 si $Fcalc \geq F_{crit;}$ no Rechazo H_0 si $Fcalc < F_{crit}$

EJERCICIO

Un ingeniero está interesado en maximizar la resistencia a la tensión de una nueva fibra sintética que se empleará en la manufactura de tela para camisas.

El ingeniero sabe por experiencia que la resistencia es influida por el porcentaje de algodón presente en la fibra. Además, sospecha que elevar el contenido de algodón incrementará la resistencia, al menos inicialmente. También sabe que el contenido del algodón debe variar aproximadamente entre 20 y 40 % para que la tela resultante tenga otras características de calidad que se desean (como capacidad de recibir un tratamiento de planchado permanente).

EJERCICIO: PRIMERA PARTE

El ingeniero decide probar muestras con **5 niveles de porcentaje** de algodón: 15, 20, 25, 30 y 35% y quiere saber si el diferente porcentaje de **algodón "aplicado"** produce **diferencias significativas** en la **resistencia a la tensión** de la nueva fibra sintética medida en lb/pul^2.
A si mismo decide ensayar **cinco trozos** de materiales en cada nivel de contenido de algodón.

Éste es un ejemplo de experimento unifactorial con t=5 niveles de factor y n=5 repeticiones, en el que se tienen entonces 5*5= 25 unidades experimentales a las que se asignan aleatoriamente los "tratamientos".

EJERCICIO: PRIMERA PARTE

Es así un experimento en el cual se tienen:
25 unidades experimentales: 25 trozos de fibra sintética

1 Factor: % de algodón presente en la fibra

5 niveles del Factor: 15, 20, 25, 30 y 35% de algodón

5 repeticiones por nivel del Factor: 5 trozos del material por cada tratamiento

La variable en estudio: Resistencia a la tensión, medida en lb/plg^2

Las hipótesis de interés son:

$$\begin{cases} H_0: \mu_{15\%} = \mu_{20\%} = \mu_{25\%} = \mu_{30\%} = \mu_{35\%} = \mu \\ H_1: \text{Alguna } \mu_i \text{ es} \neq \text{de } \mu \end{cases}$$

¿Cómo deben ser las varianzas? Iguales

Los principios que se deben cumplir son: Aleatorización, repetición y control local.

EJERCICIO: PRIMERA PARTE

El modelo que representa el problema es:

$$\boxed{Y_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1, 2,3,4,5; \quad j = 1, \dots, 5}$$

$$\mu_i = \mu + \tau_i$$

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad i = 1,2,3,4,5; \quad j = 1, \dots, 5$$

EJERCICIO: PENSEMOS…

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad i = 1,2 \ldots.; \quad j = 1,2\ldots, n$$

$$\mu_i = \mu + \tau_i$$

Para plantear una dócima, ….razonemos… ¿cuál o cuáles serían las hipótesis y la forma de trabajar?

Supongamos 5 muestras, 10 comparaciones de a dos,….

Si α=0,05, entonces P(de tomar la decisión correcta)=0,95

Entonces para las 10 comparaciones sería:

P(correcta 1ra. Vez y correcta 2da vez y…)=

= (0,95) (0,95) (0,95) …. (0,95) =$(0,95)^{10}$=0,599

Por lo que

P(al menos una errónea)=1-(todas correctas)=1-0,599=0,401=α ??

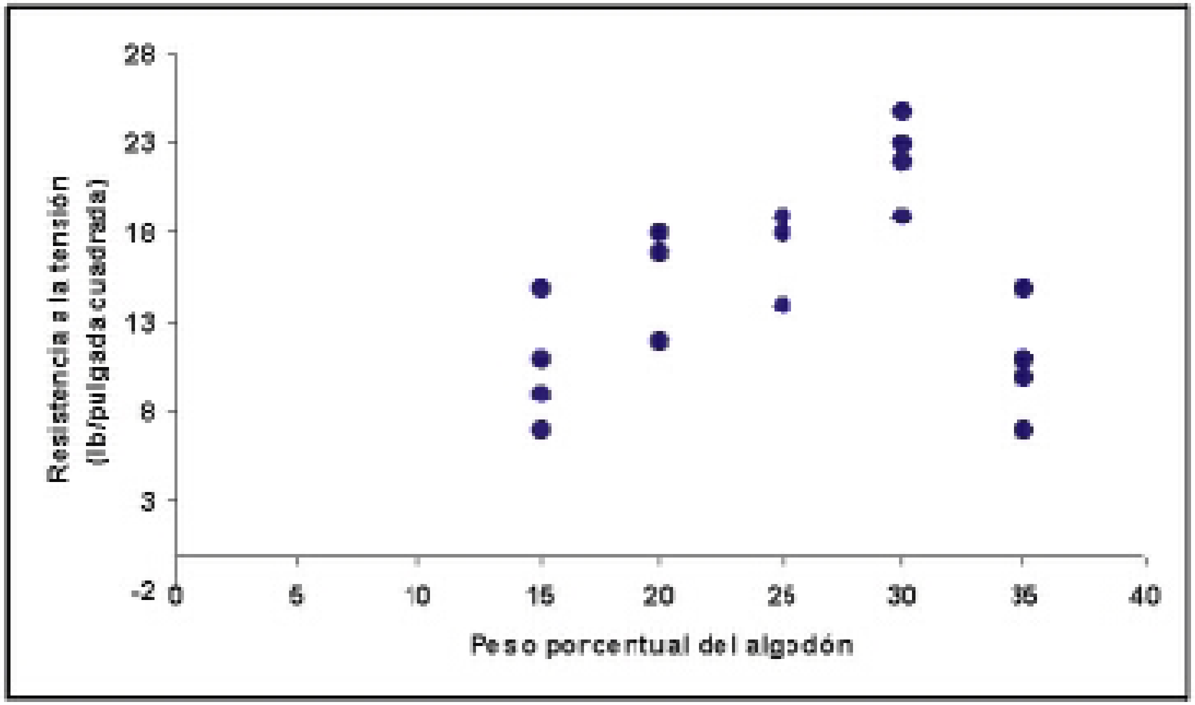
¡¡Esto significa que: más del 40% de las veces se estaría cometiendo un error de tipo I !! (Rechazar H_o siendo verdadera)

…esto no es correcto… se necesitará encontrar otro procedimiento…

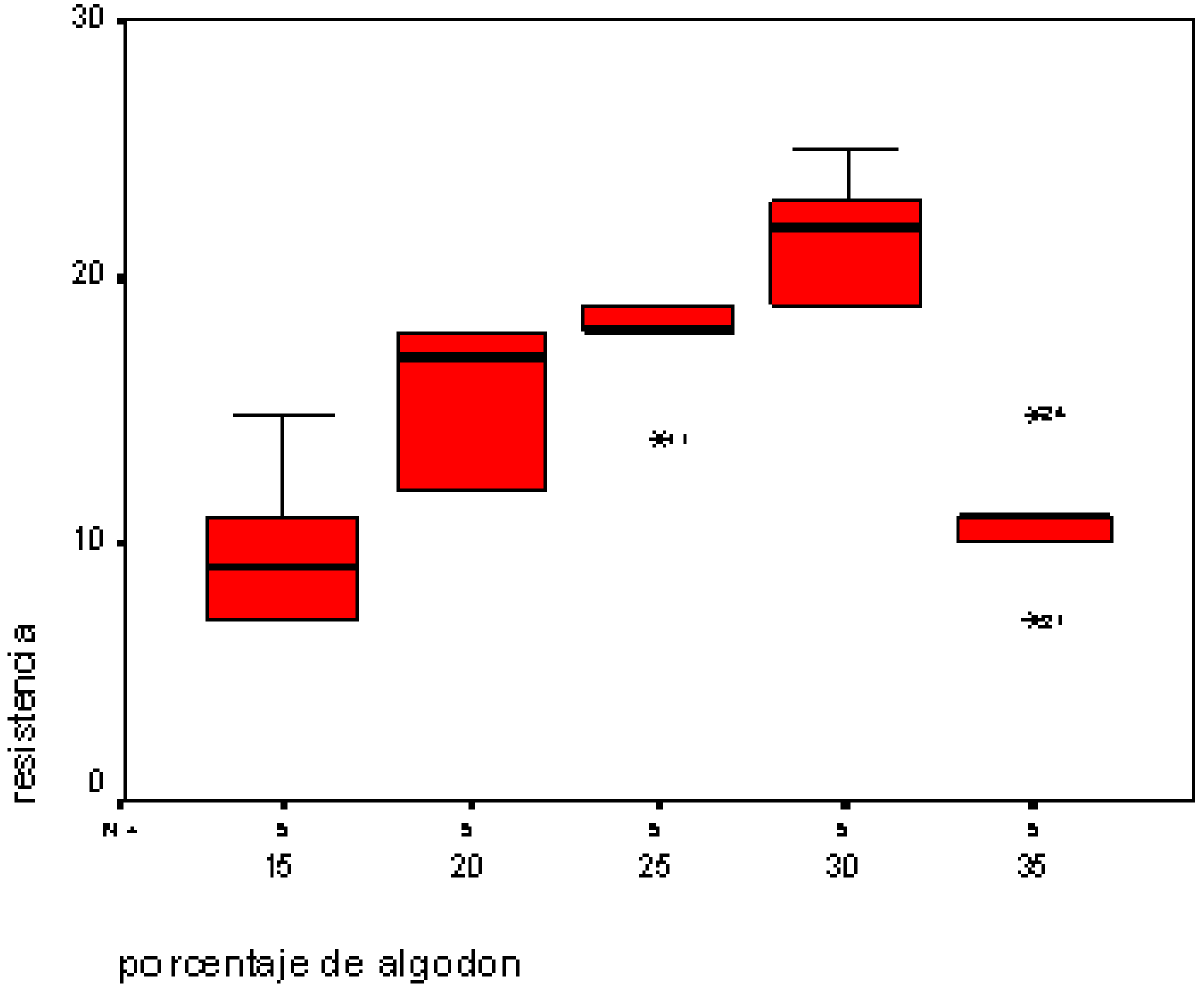
EJERCICIO: SEGUNDA PARTE

El ingeniero decide probar muestras a cinco niveles de porcentaje de algodón: 15, 20, 25, 30 y 35 %. Así mismo, decide ensayar cinco muestras a cada nivel del contenido de algodón.

Porcentaje de algodón	1	2	3	4	5
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11



EJERCICIO: SEGUNDA PARTE



DCA Fijo o Modelo I

$$Y_{ij} = \mu_i + \varepsilon_{ij} \quad i = 1,\dots,t; \quad j = 1,\dots,r$$

Donde:

Y_{ij}: Es el valor observado de la *j-ésima* observación del *i-ésimo* tratamiento.

μ_i : Es el valor medio poblacional del *i-ésimo* tratamiento.

ε_{ij}: Es el error aleatorio de la *j-ésima* observación correspondiente al i-ésimo tratamiento.

O si en el modelo anterior denotamos a $\mu_i = \mu + \tau_i \quad i = 1,\dots,t$ se obtiene la siguiente forma alternativa del modelo

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad i = 1,\dots,t; \quad j = 1,\dots,r$$

Donde:

Y_{ij}: Es el valor observado de la *j-ésima* observación que recibió el *i-ésimo* tratamiento.

μ : Es el valor medio poblacional general.

τ_i: Es el efecto medio poblacional del *i-ésimo* tratamiento.

ε_{ij}: Es el error aleatorio de la *j-ésima* observación correspondiente al i-ésimo tratamiento.

EJERCICIO: SEGUNDA PARTE

PARA TENER EN CUENTA Y COMPLETAR…

-¿cuáles y cuántos son los tratamientos en este caso?
-¿Cuáles y cuántas son las unidades experimentales en este caso?
Si el modelo seleccionado es

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$
$$i = 1\ a\ \ldots\ldots..;$$
$$j = 1\ a\ \ldots\ldots.$$

Diga cuáles son cada uno de sus elementos en términos del problema.

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

Donde:

Y_{ij}: .....

μ : .....

τ_i: ....

ε_{ij}: ....

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

$$i = 15\ a\ 35\ \%\ de\ algodón; \quad j = 1\ a\ 5\ trozo\ de\ tela$$

Donde:

Y_{ij}: Es el valor observado de la resistencia a la tensión del *j-ésimo trozo de tela* que recibió el *i-ésimo* % de algodón.

μ : Es la resistencia media poblacional a la tensión.

τ_i: Es el efecto medio poblacional del *i-ésimo* % de algodón.

ε_{ij}: Es el error aleatorio correspondiente al *j-ésimo* trozo de tela que recibió el i-ésimo % de algodón.

Complete la siguiente Tabla ANOVA

Fuentes Variación (F.V.)	Sumas Cuadrados (S.C.)	g.l	Cuadrados Medios (C.M.)	*F*
....	....	t-1	118,94	
....	....	t(r-1)	....	....
Total	636,96	tr - 1		

Escriba las Hipótesis en términos de los efectos y de los promedios.

(Use α=0,05) dibuje la región crítica y concluya en términos del problema.

Fuentes Variación (F.V.)	Sumas Cuadrados (S.C.)	g.l	Cuadrados Medios (C.M.)	F
Entre Tratamientos	475,76	4	118,94	14,757
Dentro Trat. (EE)	161,2	20	8,06	
Total	636,96	24		

Hipótesis en términos de los efectos

$$\begin{cases} H_0 : \tau_i = 0 \quad i = 1,2, \ldots, 5 \\ H_1 : \text{algún } \tau_i \neq 0 \end{cases}$$

H_0: (Todos los tratamientos tienen el mismo efecto sobre la variable rta.)

H_1: (No todos los tratamientos tienen el mismo efecto sobre la variable en estudio).

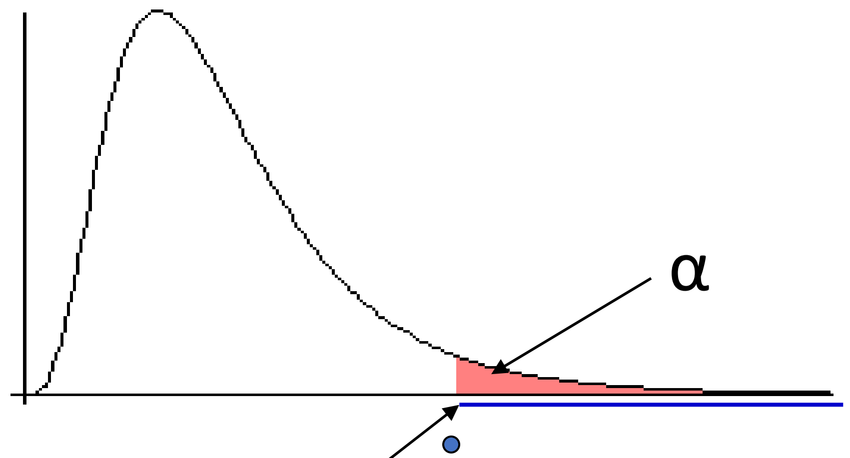
Dócimas en términos de los promedios

$$\begin{cases} H_0: \mu_1 = \mu_2 = \ldots = \mu_5 = \mu \\ H_1: \text{algún } \mu_i \neq \mu. \end{cases}$$

2-Nivel de significación α= 0,05

3 .Variable pivotal o estadístico de la prueba $F = \frac{CM_{trat}}{CM_{ee}} \approx F_{(t-1;t(r-1))}$

4- Región Crítica y regla de decisión: La región crítica $F_{(t-1;\ t(r-1))} \geq F_{crit}$



$F_{crit} = F_{(4;\ 20);\ \alpha=0,05} = 2,87$

Regla de decisión

Rechazo H_0 si Fcalc ≥ 2,87; no Rechazo H_0 si Fcalc < 2,87

Como $14{,}757 \geq 2{,}87$; Rechazo H_0

<u>Conclusión</u>: Con un nivel de significación de 0,05 o del 5%, existe suficiente evidencia para rechazar la hipótesis nula y decir que el porcentaje de algodón influye en la resistencia a la tensión de la fibra sintética.

DATOS DE ALGODÓN

15%	**20%**	**25%**	**30%**	**35%**		
7	12	14	19	7		
7	17	18	25	10		
15	12	18	22	11		
11	18	19	19	15		
9	18	19	23	11		
49	**77**	**88**	**108**	**54**	**376**	**Σy**
2401	**5929**	**7744**	**11664**	**2916**	**30654**	**$(\Sigma y)^2$**
525	**1225**	**1566**	**2360**	**616**	**6292**	**Σy^2**

OTRO EJEMPLO

En un experimento se comparan 3 métodos para enseñar un idioma extranjero (auditivo, traducción y combinado). Para evaluar la instrucción, se administró una prueba de vocabulario de 50 preguntas a los 24 estudiantes repartidos al azar de a 8 por grupo y se midió el puntaje obtenido por cada estudiante.

A-¿Cuál es la variable respuesta?
B- ¿cuáles y cuántos son los tratamientos?
C-¿Se cumplen los principios?
D- ¿cuáles son los supuestos que se deben verificar?
E-Asuma que se cumplen los supuestos y plantee las hipótesis en términos de los promedios y de los efectos de los tratamientos.

Complete la siguiente Tabla ANOVA

Fuentes Variación (F.V.)	Sumas Cuadrados (S.C.)	g.l	Cuadrados Medios (C.M.)	*F*
Tratamientos				
Error			...	.....
Total				

Hipótesis en términos de los efectos

H_0 :
H_1:.

Dócimas en términos de los promedios

H_0:
H_1:

Use α=0,05 decida y concluya en términos del problema.

A-¿Cuál es la variable respuesta? El puntaje obtenido en la evaluación

B- ¿cuáles y cuántos son los tratamientos? El Factor es el método de enseñanza, tiene 3 niveles: auditivo, traducción y combinado.

C-¿Se cumplen los principios? Aleatorización, repetición y control local

D- ¿cuáles son los supuestos que se deben verificar?

1. Para cada tratamiento la variable aleatoria respuesta Y_{ij} se distribuye normalmente . Para cada T_i
2. Las Y_{ij} son variables aleatorias independiente entre si
3. Las varianzas de los t tratamientos son iguales (homoscedasticidad)

E-Asuma que se cumplen los supuestos y plantee las hipótesis en términos de los promedios y de los efectos de los tratamientos.

Hipótesis en términos de los efectos de los tratamientos

$H_0 : \tau_i = 0$ $i = 1, 2$ y 3(Todos los tratamientos tienen el mismo efecto sobre la variable rta.)

H_1: algún $\tau_i \neq 0$ (No todos los tratamientos tienen el mismo efecto sobre la variable en estudio).

Hipótesis en términos de los promedios

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu$

H_1: algún $\mu_i \neq \mu$.

Tabla ANOVA

Fuentes Variación (F.V.)	Sumas Cuadrados (S.C.)	g.l	Cuadrados Medios (C.M.)	*F*
Tratamientos	647,584	2	323,792	8,360
Error	813,374	21	38,732	
Total	1460,958	23		

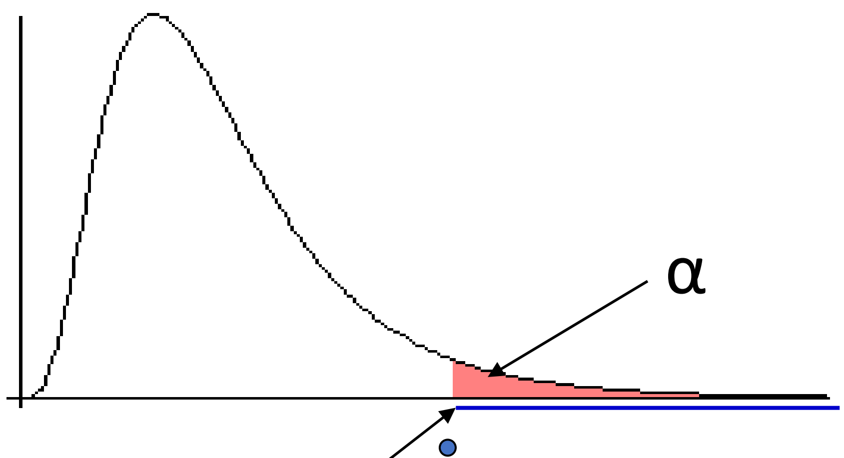
Use α=0,05 decida y concluya en términos del problema.

2-Nivel de significación α= 0,05

3 .Variable pivotal o estadístico de la prueba

$$F = \frac{CM_{trat}}{CM_{ee}} \approx F_{(t-1;t(r-1))}$$

4- Región Crítica y regla de decisión: La región crítica $F_{(t-1;\ t(r-1))} \geq F_{crit}$



Fcrit = F (2; 21); α=0,05=3,467

Regla de decisión

Rechazo H_0 si Fcalc ≥ 3,467; no Rechazo H_0 si Fcalc < 3,467

Como Fcalc (8,369) ≥ 3,467; Rechazo H_0

<u>Conclusión</u>: Con un nivel de significación de 0,05 o del 5%, existe suficiente evidencia para rechazar la hipótesis nula y decir que los métodos de enseñanza influyen en el puntaje obtenido en las pruebas de vocabulario.

ÚLTIMO EJEMPLO

Para estudiar los efectos del etanol en el tiempo de sueño, se seleccionó una muestra al azar de 20 ratas de edad semejante, a cada rata se le administró una dosis oral con una concentración en particular de etanol por peso corporal (0g/kg;1g/kg; 2g/kg, y 4g/kg). Se registró el movimiento ocular rápido (REM) en el tiempo de sueño para cada rata.

A-¿Cuál es la variable respuesta?
B- ¿cuáles y cuántos son los tratamientos?
C-¿cuáles y cuántas son las unidades experimentales en este caso?
D- Asuma que se cumplen los supuestos y plantee las hipótesis en términos de los promedios y de los efectos de los tratamientos.

Complete la siguiente Tabla ANOVA

Fuentes Variación (F.V.)	Sumas Cuadrados (S.C.)	g.l	Cuadrados Medios (C.M.)	*F*
Tratamientos				…..
Error			…	
Total				

Hipótesis en términos de los efectos

H_0 :
H_1:.

Dócimas en términos de los promedios

H_0:
H_1:

Use α=0,05 decida y concluya en términos del problema.

Tabla ANOVA

Fuentes Variación (F.V.)	Sumas Cuadrados (S.C.)	g.l	Cuadrados Medios (C.M.)	*F*
Tratamientos	5882,4	3	1960,785	**21,09**
Error	1487,4	16	92,9625	
Total	7369,8	19		

A-¿Cuál es la variable respuesta? El movimiento ocular rápido REM en el tiempo de sueño.

B- ¿cuáles y cuántos son los tratamientos? El Factor es la dosis oral de etanol, tiene 4 niveles (0g/kg; 1g/kg; 2g/kg, y 4g/kg).

C-¿cuáles y cuántas son las unidades experimentales en este caso? Las unidades experimentales son las ratas. Son 20 unidades experimentales, 20 ratas

D- Asuma que se cumplen los supuestos y plantee las hipótesis en términos de los promedios y de los efectos de los tratamientos.

E-Asuma que se cumplen los supuestos y plantee las hipótesis en términos de los promedios y de los efectos de los tratamientos.

Hipótesis en términos de los efectos de los tratamientos

H_0 : $\tau_i = 0$ $\quad$ i = 1,2 y 4 (Todos los tratamientos tienen el mismo efecto sobre la variable rta.)

H_1: algún $\tau_i \neq 0$ (No todos los tratamientos tienen el mismo efecto sobre la variable en estudio).

Hipótesis en términos de los promedios

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu$

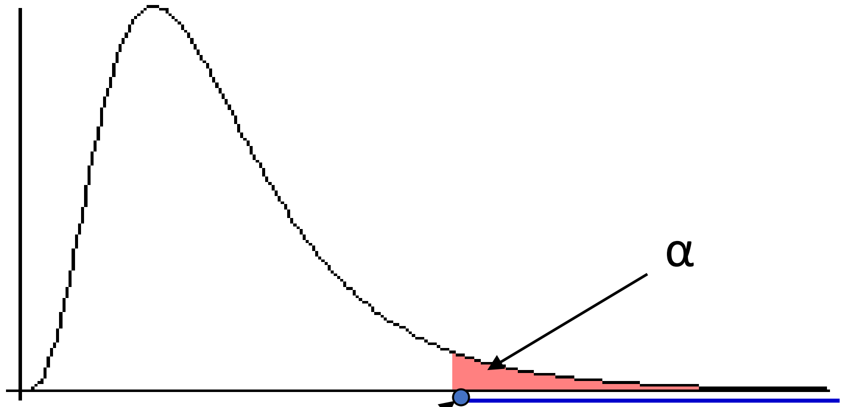
H_1: algún $\mu_i \neq \mu$.

2-Nivel de significación α= 0,05

3 .Variable pivotal o estadístico de la prueba

$$F = \frac{CM_{trat}}{CM_{ee}} \approx F_{((t-1);t(r-1))}$$

4- Región Crítica y regla de decisión: La región crítica $F_{((t-1);\, t(r-1))} \geq F_{crit}$



Fcrit = F (3; 16);
α=0,05=3,239

Regla de decisión

Rechazo H_0 si Fcalc ≥ 3,239; no Rechazo H_0 si Fcalc < 3,239

Como Fcalc (21,09) $\geq$ 3,239; Rechazo H_0

Conclusión: Con un nivel de significación de 0,05 o del 5%, existe suficiente evidencia para rechazar la hipótesis nula y considerar que el movimiento ocular rápido REM medio no es igual en las diferentes concentraciones de etanol.

¡Muchas Gracias por su atención!