

ANOVA

PROBLEMA (comparación de dos promedios)

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

El ingeniero sabe por experiencia que la resistencia es influida por el porcentaje de algodón presente en la fibra. Además, él sospecha que elevar el contenido de algodón incrementará la resistencia, al menos inicialmente.

También sabe que el contenido de 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).

El ingeniero decide probar muestras a dos niveles de porcentaje de algodón: 20% y 30 % y quiere saber si el diferente porcentaje de algodón “aplicado” produce diferencias significativas en la resistencia a la tensión de la nueva fibra sintética. Asimismo decide ensayar cinco trozos de materiales en cada nivel de contenido de algodón.

Este es el caso en que se tiene una población objetivo de material (fibra sintética) de la que se tomarán 10 “unidades experimentales” o “probetas” o “trozos de fibra”, y en forma aleatoria se asignará a 5 de ellas un “tratamiento”: 20 % de algodón, y a otros 5 el “tratamiento”: 30 % de algodón.

Este es así un experimento en el cual se tienen

- 10 unidades experimentales : 10 trozos de fibra sintética
- 1 tratamiento: % de algodón presente en la fibra
- 2 niveles del tratamiento : 20 % y 30 %.
- 5 repeticiones por nivel de tratamiento.
- La variable en estudio: resistencia a la tensión, medida en lb/plg².

Recordemos que en muchos casos no se habla de factores y niveles de factores, sino simplemente de “tratamientos”.- Y en ese caso, cada uno de los % de algodón que “se aplican”, en general se llaman tratamientos.

Entendemos así ahora al experimento en que se considera que cada una de las muestras de 5 repeticiones o trozos de fibra a los que se aplicó un tratamiento son una muestra aleatoria de una población “teórica” que tienen su promedio y su varianza.

Para probar si las resistencias promedio de las fibras tratadas con los dos “tratamientos” o porcentajes de algodón” difieren significativamente, previo análisis de las variabilidades, para “no tener evidencias suficientes para decir que no son iguales” y poder entonces encontrar la variabilidad amalgamada requerida (y verificar el cumplimiento de los supuestos), se comparan los promedios poblacionales a partir de una prueba “t” de Student.

Las hipótesis de interés son:

$$H_0 : \mu_{20\%} = \mu_{30\%}$$

$$H_1 : \mu_{20\%} \neq \mu_{30\%}$$

Recordemos que hemos reiterado en el análisis rápido (repaso) que hemos hecho que:

ESTADÍSTICA ANOVA

- debe existir “**aleatorización**” (para garantizar la no correlación entre el efecto de tratamiento y el error)
- deben existir muestras de tamaño mayor que 1 (**repetición**) para poder calcular las variabilidades de las muestras.
- Las **variabilidades** correspondientes a los niveles del tratamiento deben ser **iguales**, ya que σ^2_e debe ser **único**, común.

Tendríamos

$$Y_{ij} = \mu_i + e_{ij} \quad \text{para } i = 1 \text{ y } 2 \\ j = 1, \dots, 5$$

donde

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

siendo τ_i el efecto del i – ésimo tratamiento

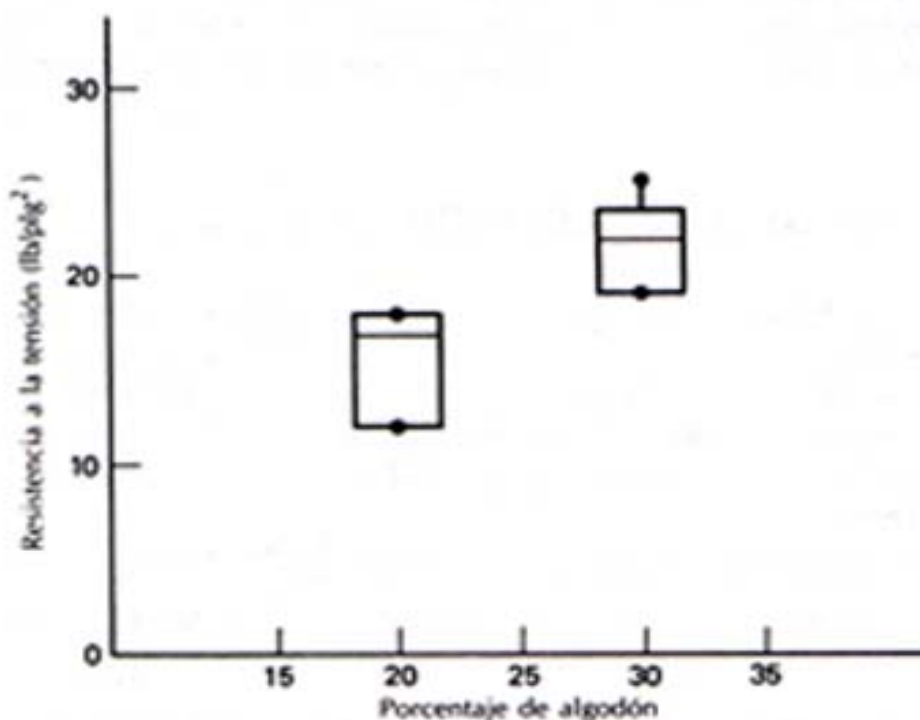
y entonces es

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

Los datos son(en este momento y para repasar, no interesa tener los valores sino analizar el caso)

Datos del experimento de resistencia a la tensión (en lb/plg ²)							
Porcentaje de algodón	Observaciones					Total	Promedio
	1	2	3	4	5		
20	12	17	12	18	18	77	15.4
30	19	25	22	19	23	108	21.6

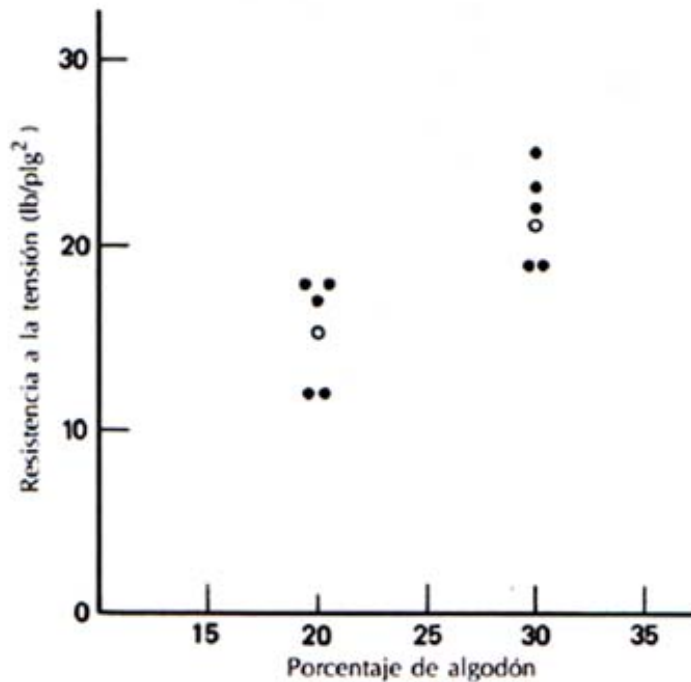
Los diagramas de caja son :



Es

Los diagramas de dispersión son:

.....



Realizados los cálculos se tiene que...

$t_c = 2,19$, para un nivel de significación del 5% es $t_t = \pm 2,306$

y entonces se concluye que no existe diferencia significativa entre los promedios correspondientes a 20 % y 30 % de algodón, lo cual servirá al ingeniero para tomar una decisión.

Obsérvese que si bien parecía, con la información proporcionada por las representaciones gráficas, que “podría” haber diferencias, ***no existe diferencia significativa***)

Ahora sí, repasados conceptos apprehendidos, relacionados temas conocidos, vista una forma mas apropiada para expresarse formalmente (modelos) generalicemos este análisis para el caso en que se tienen más de dos promedios a comparar.

PROBLEMA (comparación de más de dos promedios)

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

El ingeniero sabe por experiencia que la resistencia es influida por el porcentaje de algodón presente en la fibra. Además, él sospecha que elevar el contenido de algodón incrementará la resistencia, al menos inicialmente.

También sabe que el contenido de 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).

El ingeniero decide probar en trozos de fibra sintética a cinco niveles de porcentaje de algodón: 15, 20, 25, 30 y 35%.

Asimismo, decide ensayar cinco muestras en cada nivel de contenido de algodón.

Éste es un ejemplo de experimento unifactorial con $t = 5$ niveles del factor y $n = 5$ repeticiones, en el que se tienen entonces $5 \times 5 = 25$ unidades experimentales a las que se asignaran aleatoriamente los “tratamientos”.

En este experimento se tienen:

- Una variable en estudio = la resistencia a la tensión de la fibra sintética que recibe cierto % de algodón, medida en lb/plg²
- 1 tratamiento o factor = % de algodón que se “aplica” en la fibra sintética
- 5 niveles de tratamiento o de factor = 15-20-25-30 y 35% de algodón
- 5 repeticiones por nivel de tratamiento = 5 trozos de fibra sintética por % de algodón que se aplica
- 25 unidades experimentales = 25 trozos de fibra sintética

RECORDEMOS LOS CONCEPTOS ANALIZADOS

TRATAMIENTO O FACTOR:

NIVELES DE FACTOR: O TAMBIEN TRATAMIENTO?

UNIDADES EXPERIMENTALES:

REPETICIONES:

Analícemos. . .

Supongamos que tomamos las 25 unidades experimentales “tal como aparecen”, esto es, “sin aleatorizar”.

Por ejemplo, supongamos que podemos “aplicar” los tratamientos “ % de algodón ” y tomemos los trozos de tela “uno después de otro...” y adjudiquemos a los cinco primeros trozos el 15% de algodón, a los cinco siguientes el 20%, ... y así siguiendo.

De esta manera se tendría:

Porcentaje de algodón	Número de Corrida experimental				
15	1	2	3	4	5
20	6	7	8	9	10
25	11	12	13	14	15
30	16	17	18	19	20
35	21	22	23	24	25

¿Qué podría pasar?

¿Es este un procedimiento correcto?

Evidentemente NO

¿Por qué? ¿Como aleatorizamos?

Una forma sería. . .

Secuencia de prueba	Número de corrida	Porcentaje de algodón
1	8	20
2	18	30
3	10	20
4	23	35
5	17	30
6	5	15
7	14	25
8	6	20
9	15	25
10	20	30
11	9	20
12	4	15
13	12	25
14	7	20
15	1	15
16	24	35
17	21	35
18	11	25
19	2	15
20	13	25
21	22	35
22	16	30
23	25	35
24	19	30
25	3	15

Aleatorizando en los dos “lados”.

¿Puede haber otra forma? Piénselo...

Ahora:

Cómo expresamos el problema?

Podríamos escribir cómo se componen las observaciones, en general?

Describamos los elementos del modelo

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \begin{cases} i = 1, 2, \dots, t \\ j = 1, 2, \dots, n_i \end{cases}$$

Recordemos los modelos que hemos usado al probar hipótesis, también en los problemas de regresión . . .

Recordemos y analicemos si se trata de parámetros o estimadores . . .

Y a una observación, cómo la expresamos?

$$Y_{11} = \mu + \tau_1 + \epsilon_{11}$$

$$Y_{35} = \mu + \tau_3 + \epsilon_{35}$$

etc.

el modelo es . . .

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad \begin{matrix} \text{para } i = 1, 2, \dots, t \\ j = 1, \dots, n_i \end{matrix}$$

donde

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

siendo τ_i el efecto del i – ésimo tratamiento

y entonces es

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

recordando que ϵ_{ij} = **var. aleatoria del error**,
con distribución normal y Esperanza cero y
varianza “común” e
igual para todos los tratamientos

Falta

RAZONEMOS . . .

Supongamos ahora que queremos comparar los cinco promedios. ¿Cómo procedemos?
¿podremos usar la variable pivotal “t” que hemos empleado para comparar de a dos?
¿podemos ahora con una “t” comparar más de dos promedios? Evidentemente NO.

¿Podremos comparar de a pares, en cada caso con una “t”? Puede parecer lógico. ¿No?
Pero analicemos qué pasaría en ese caso.

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

si consideramos $\alpha = 0.05$

entonces es $P(\text{decisión correcta}) = 0.95$,

y entonces para las diez comparaciones sería:

$P(\text{correcta 1ra. vez y correcta 2da. vez y } \dots) = (0,95) (0,95) (0,95) \dots (0,95) = (0,95)^{10} = 0,599$

$P(\text{al menos una errónea}) = 1 - P(\text{todas correctas}) = 1 - 0,599 = 0,401 = \alpha ? ? ?$

Esto significa que mas del 40% de las veces se estaría cometiendo un error de tipo I
esto no es correcto . . .

se necesitará encontrar otro procedimiento . . .

PRESENTAMOS A RONALD A. FISHER

R. A. Fisher (1890-1962)

Sus importantes contribuciones a la Estadística matemática, entre las que destacan el análisis de la varianza y el diseño de experimentos, permitieron hablar de la “revolución fisheriana”, cuyo origen se sitúa en Inglaterra por los años treinta.



Trabajemos entonces para aplicar la técnica del análisis de la varianza para comparar más de dos poblaciones.

Tendremos como hipótesis iniciales . . .

Est. Ana María ROSSI de SALVAGNI.-.2000.-

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

$H_1 : \text{Al menos una } \mu_i \text{ es distinta}$

O bien

$$H_0 : \tau_i = 0, \text{ para todos los } i$$

$$H_0 : \tau_i \neq 0, \text{ para algún } i$$

Comencemos por considerar que se tienen los datos observados de las repeticiones de varios ("t") tratamientos

armemos la tabla para colocar luego los valores observados en el experimento

los n_i pueden ser iguales o distintos para cada tratamiento

consideremos en este caso los n_i iguales

observaciones	tratamientos (i)						totales
	1	2	...	i	...	t	
obs.(j)	y_{11} y_{12} y_{13}	y_{24}					
				y_{ij}			
						$y_{t ni}$	
totales							
$\sum_j y_{ij} = y_{i.}$	$Y_{1.}$	$Y_{2.}$		$Y_{i.}$	$Y_{t.}$	$\sum_i y_{1.} = y_{..} = T$	$\sum_i \sum_j y_{ij} = T$ total general
promedios							
$\sum_j y_{ij} / n_i$	$\tilde{y}_{1.}$	$\tilde{y}_{2.}$		$\tilde{y}_{i.}$	$\tilde{y}_{t.}$	$\sum_i \tilde{y}_{i.} = \tilde{y}_{..}$	promedio general

Total general

$$\sum_i \sum_j y_{ij} = \sum_i y_{i.} = y_{..} = T$$

promedio general

$$\tilde{y}_{..} = \sum_i \sum_j y_{ij} / N = \sum_i y_{i.} / t = y_{..} / N = T / N$$

donde N es el total de U E (unidades experimentales) $\sum_i n_i = N$
 $t \times n$

si el numero de rep. por trat. es el mismo promedio de tratamiento

$$\bar{y}_i = \sum_j y_{ij} / n_i$$

REVISEMOS QUÉ PODEMOS HACER CON CADA TRATAMIENTO. . .

Podríamos calcular \bar{y}_i y S^2_i

Tendríamos $\bar{y}_i = \sum_j y_{ij} / n_i$

$$Y \quad S^2_i = \sum_{(j=1 \text{ a } n_i)} (y_{ij} - \bar{y}_i)^2 / (n_i - 1)$$

Por ejemplo, para el primer tratamiento tendríamos

$$\bar{y}_1 = \sum_j y_{1j} / n_1$$

$$Y \quad S^2_1 = \sum_{(j=1 \text{ a } n_1)} (y_{1j} - \bar{y}_1)^2 / (n_1 - 1)$$

(S^2_e para el tratamiento “1”)

¿Y para todos los tratamientos?

¿Cómo haríamos?

Deberíamos calcular una varianza promedio o amalgamada!!!!

Como el S^2_a ! ! ! !

$$\begin{aligned} S^2_a &= (n_1 - 1) S^2_1 + (n_2 - 1) S^2_2 + \dots + (n_t - 1) S^2_t = \\ &= \sum_{(i=1 \text{ a } t)} \left[\sum_{(j=1 \text{ a } n_i)} (y_{ij} - \bar{y}_i)^2 / \sum_{(i=1 \text{ a } t)} (n_i - 1) \right] \\ &= S_{Ce} / gl = S_{Ce} / (N - t) \\ &= C_{me} \\ &\text{varianza del error} \\ &\text{(estimacion de } \sigma^2_e \text{)} \end{aligned}$$

Qué es entonces el CMe?

$$\hat{C}_{Me} = \sigma^2_e \text{ estimada! ! !}$$

Estimación de la varianza común de todos los t tratamientos

Promedio “ponderado”, “amalgamado” de las estimaciones de las varianzas de error de cada tratamiento

Y hemos trabajado con una variabilidad “debida al error, residual, dentro de los tratamientos”

Recordemos que

$$\sigma_e^2 = CMe = SCe / (N - t)$$

donde N es el total de unidades experimentales

por otra parte . . .

¿cómo calcularíamos una **variabilidad “debida a los tratamientos”**?

Seguramente, por intuición, diríamos que debería hacerse una suma de cuadrados de diferencias en las cuales se compararía la media de cada tratamiento con la media general y luego se promediarían esas diferencias al cuadrado, hasta obtenerse una “varianza” debida a los tratamientos

Y ESO HAREMOS . . .

Tomemos . . .

$$\sum_{(i=1 \text{ a } t)} (\tilde{y}_i - \tilde{y}.)^2$$

¿ SC tratamientos ? $\sum_{(i=1 \text{ a } t)} (\tilde{y}_i - \tilde{y}.)^2$?

recordemos: $V(Y) = \sigma^2$

$$V(\tilde{y}) = \sigma^2 / n$$

$$n V(\tilde{y}) = \sigma^2$$

y con este razonamiento vemos que . . .

$$SC_{\text{tratamientos}} = n \sum_{(i=1 \text{ a } t)} (\tilde{y}_i - \tilde{y}.)^2$$

$$CM_{\text{tratamientos}} = SC_{\text{tratamientos}} / \text{g.l.}$$

?
 ←
 g.l. = t - 1

$$CM_{\text{tratamientos}} = n \sum_{(i=1 \text{ a } t)} (\tilde{y}_i - \tilde{y}.)^2 / (t - 1)$$

Y entonces tendremos que . . .

$$SC_{\text{tratamientos}} + SC_{\text{error}} = ? = SC_{\text{total}}$$

y como encontramos la SC error?

. . . con el mismo razonamiento . . .

y entonces llegamos a que . . .

$$. . . SC_{\text{total}} = \sum_{(i=1 \text{ a } t)} \sum_{(j=1 \text{ a } n_i)} (y_{ij} - \tilde{y}.)^2$$

y ahora nos preguntamos. . .

¿existe el CM total ?

Recordemos

el teorema de la partición del chi cuadrado. . .

Se había llegado a la conclusión de que. . .

Las SUMAS DE CUADRADOS son ADITIVAS
Los GRADOS DE LIBERTAD son ADITIVOS

PERO. . . ¿ los COCIENTES . . . ?

NO !!! LOS COCIENTES NO SON ADITIVOS !!!

Habíamos visto y recordamos

que : $X^2 = \sum N(0,1)^2$

combinación lineal de var. al. $N(0,1)$ al cuadrado...

El producto y el cociente no son aditivos...

(salvo si los denominadores son iguales)...

La suma sí es distributiva...etc...

AHORA SI. . .

Armemos la tabla de ANOVA o ANDEVA

Similar a la empleada para el análisis de la regresión...

Cuáles son las partes?

Recordemos. . .

Fuentes de variación (F V)

Sumas de cuadrados (S C)

Grados de libertad de las S C (g l) o (d l)

Cuadrados Medios (C M)

Cocientes SC / gl

(Medias de Cuadrados = M C)

estimadores de varianzas. . .

Esperanzas de C M (E(CM))

Varianzas poblacionales

F

(Calculada, operación...? F_c)

(Probabilidad encontrada para valor critico...? $p < o p > \dots$)

(tabulada para cierto nivel de significación? F_t)

Armemos la tabla y veamos. . .

ESTADÍSTICA ANOVA

F.V.	S.C.	g.l.	C.M.	F _c	C.M.Esperad
Tratamientos (entre)	$(\tilde{y}_i - \tilde{y}.)$	$(t - 1)$			
Error (dentro)	$(y_{ij} - \tilde{y}_i.)$	$(N - t)$			
<u>Total</u>	$(y_{ij} - \tilde{y}.)$	$(N-1)$			

donde N es el numero de U.E.

INCOMPLETO !!!!!

CUIDADO !!!!!

esta tabla es sólo para razonar cómo se forman las variabilidades . . . y para asegurarnos que entendemos que significa cada fuente de variación.

OBSERVEMOS COMO LA COMPLETAMOS . . .

La tabla completa y correctamente escrita será . . .

.V.	S C	g.l.	C.M.	F _c	C.M. Espera dos
Tratamientos (entre)	$n\sum(\tilde{y}_i - \tilde{y}.)^2$	$(t - 1)$	$SCt / (t-1)$ C.M.t	?	$\sigma^2 + t \sum \tau_j^2 / (t-1)$
Error (dentro)	$\sum \sum (y_i - \tilde{y}_i.)^2$	$(N - t)$	$SCe / (N-t)$ C.M.e		
<u>Total</u>	$\sum \sum (y_{ij} - \tilde{y}.)^2$	$(N-1)$?		

donde N es el numero de U.E.

Recordemos las hipótesis iniciales . . .

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

H1 : Al menos una μ_i es distinta

O bien

$$H_0 : \tau_i = 0, \text{ para todos los } i$$

$$H_0 : \tau_i \neq 0, \text{ para algún } i$$

Est. Ana María ROSSI de SALVAGNI.-.2000.-

Sabemos, por la forma en que hemos construido el C.M.error,
(como promedio ponderado de las estimaciones no viciadas de las varianzas comunes) y ya que sabemos que e_{ij} tiene una distribución normal con media 0 y varianza σ_e^2 que su esperanza es

$$E (CMe) = \sigma_e^2$$

Por otra parte, al calcular el valor esperado del $CM_t = CM$ entre, se tendrá

$$E (CM_t) = \sigma_e^2 + n \sum \tau_j^2 / (t-1) = \sigma_e^2 + n \sum (\mu_i - \mu)^2 / (t-1)$$

Cómo interpretamos esta varianza?

Supongamos que no existan evidencias para rechazar la hipótesis propuesta.

Esto significaría que no existen evidencias suficientes para decir que los promedios poblacionales son distintos . . .

lo que nos permitiría suponer que

“pueden ser iguales”,

y esto qué indicaría?

Si suponemos que

todas las medias son iguales, e iguales a la media general μ , lo que también señalaría que los efectos de los tratamientos son todos ceros . . .

entonces . . .

$$E (CM_t) = \sigma_e^2$$

Y por ello,



$$E (CM_t) = E (CMe)$$

tanto el CM_t como el CMe serían estimadores de la misma varianza σ_e^2 ,

porque

$$n \sum \tau_j^2 / (t-1) = 0$$

pero en la medida en que haya

al menos una media de tratamiento

distinta a μ

será

$$n \sum (\mu_i - \mu)^2 / (t-1) > 0$$

siempre positiva y

$$E (CM_t) > E (CMe)$$

Este razonamiento nos debe llevar a pensar que si comparamos los CM , podremos decidir y concluir en la prueba de hipótesis que nos hemos planteado

Esto es:

Estaremos comparando varianzas y decidiendo sobre la comparación de promedios.

(Usaremos una F y no una t)

Pero también nos debe llevar a pensar que

la región crítica de la variable pivotal F que usaremos, deberá **indefectiblemente** ser unilateral derecha . . .

dibujamos la distribución . . .

ubicamos la región crítica . . .

Si el nivel de significación fuera del 5% entonces el $F_{\text{Crítico}} = F_{(t-1);(N-t); 0,95}$

y concluiríamos como en cualquier otra prueba de hipótesis, rechazando cuando corresponda, y decidiendo luego en términos del problema.

Revisaremos el significado de los dos cuadrados medios que aparecen en la tabla de análisis de la varianza.

El estadístico **cuadrado medio dentro de grupos** (CM_{dentro}) o ***intravarianza*** mide la dispersión promedio de las observaciones individuales respecto a la media de su grupo. Luego, si las unidades experimentales son asignadas aleatoriamente a los t grupos o si las muestras son muestras aleatorias de una población homogénea, el cuadrado medio *intragrupo* sirve para estimar σ^2 .

Los cuadrados medios intragrupos permanecen siempre como un residuo representando la variabilidad inherente a las unidades experimentales. Aunque el investigador controle todas las fuentes de variabilidad en su experiencia, siempre quedará cierta variabilidad (residuo) en las respuestas de unidad en unidad.

Por esta razón, el cuadrado medio dentro de grupos se denomina también **cuadrado medio de error o cuadrado medio residual**.

El *cuadrado medio de error*, no es un indicador de un error en el sentido de cometer una falta sino en el sentido de proporcionar al investigador *una medida de la variación* con la que hay que contar cuando se desean calcular diferencias significativas entre los grupos.

El *cuadrado medio de error* está compuesto por *desviaciones individuales para cada individuo u observación*.

Ellas se simbolizan con

e_{ij} , la componente aleatoria del valor individual j -ésimo en el i -ésimo grupo.

El **cuadrado medio entre grupos** (CM_{entre}) o ***intervarianza*** mide las desviaciones entre las medias de cada grupo. Es decir, describe la dispersión de las t medias de grupo con respecto a la media total.

Ya hemos visto que si no existe efecto de tratamiento, el cuadrado medio entre grupos es un estimador de σ^2 .

Por el contrario, cuando el tratamiento modifica las medias de grupos, el cuadrado medio entre grupos es un estimador de la varianza poblacional más la componente añadida por efecto del tratamiento.

Ahora bien, al calcular el valor esperado de estos cuadrados medios entre y dentro de grupos, se tiene que:

$$E(CM_{dentro}) = \sigma^2$$

$$E(CM_{entre}) = \sigma^2 + n * (\sum t_i^2 / (t - 1)) =$$

$$E(CM_{entre}) = \sigma^2 + n * (\sum (\mu_i - \mu)^2 / (t - 1))$$

Entonces, cuando las medias son todas iguales entre sí, la expresión:

$$n * (\sum (\mu_i - \mu)^2 / (t - 1)) = 0$$

y, en consecuencia $E(CM_{entre}) = E(CM_{dentro})$,

pero, en la medida que haya al menos una μ_i distinta de μ ,

$$n * \sum (\mu_i - \mu)^2 / (t - 1) > 0$$

será siempre positiva y luego, $E(CM_{entre}) > E(CM_{dentro})$

Si en el conjunto de datos observados encontramos que CM_{entre} es muy superior al CM_{dentro} , existirán entonces sospechas fundadas como para creer que $n * (\sum (\mu_i - \mu)^2 / (t - 1))$ no es nulo y debemos admitir que no todas las medias poblacionales son iguales entre si.

El problema se circunscribe entonces en *buscar un criterio estadístico* para comparar CM_{entre} y CM_{dentro} , de manera tal que *si el valor calculado para el cociente $CM_{\text{entre}} / CM_{\text{dentro}}$ excede un cierto límite, rechazaremos la hipótesis* de que:

$$E(CM_{\text{entre}}) = E(CM_{\text{dentro}})$$

y, *en caso contrario, la aceptaremos*.

Si optamos por el primer curso de, acción, estaremos al mismo tiempo admitiendo que $n * \sum (\mu_i - \mu)^2 / (t - 1) \pm 0$

Luego debemos concluir que no todas las medias poblacionales son iguales; una, algunas o todas son diferentes.

En cambio, si optamos por lo segundo, la conclusión final deberá ser que, en efecto, todas las medias son iguales.

El test F de Fisher es el test apropiado para comparar el valor calculado para $E(CM_{\text{entre}}) = E(CM_{\text{dentro}})$ con el valor teórico de la tabla.

LA CARACTERÍSTICA, TAL VEZ MÁS IMPORTANTE DEL ANÁLISIS DE LA VARIANZA ES QUE MEDIANTE ESTE PROCEDIMIENTO "INFERIMOS" LA IGUALDAD DE MEDIAS A TRAVÉS DE LA IGUALDAD DE VARIANZAS.

Luego, el test F, de igualdad de varianzas, se transforma indirectamente en un test de igualdad de medias.

TEOREMA DE COCHRAN

Sean Z_i variables aleatorias $NID(0,1)$, para $i = 1, 2, \dots, v$
y

$$\sum Z_i^2 = Q_1 + Q_2 + \dots + Q_s$$

en donde $s < v$, y Q_i tiene v grados de libertad ($i = 1, 2, \dots, s$).

Entonces Q_1, Q_2, \dots, Q_s son variables aleatorias independientes con distribución ji- cuadrada y v_1, \dots, v_s grados de libertad, si y sólo si

$$v = v_1 + v_2 + \dots + v_s$$

Como la suma de los grados de libertad de $SC_{\text{Tratamientos}}$ y de SC_E es igual a $N - 1$, es decir, el total de los grados de libertad, el teorema de Cochran implica que

$$SC_{\text{Tratamientos}} / \sigma^2 \text{ y } SC_E / \sigma^2$$

son variables aleatorias independientes con distribución
ji cuadrada.

Por tanto, si la hipótesis nula de igualdad de medias de los tratamientos es verdadera,
la razón

$$F_0 = \left(SC_{\text{Tratamientos}} / (t - 1) \right) / \left(SC_E / (N - t) \right)$$

$$= CM_{\text{Tratamientos}} / CM_E$$

tiene una distribución F con t - 1 y N - t grados de libertad.

La $F = CM_{\text{Tratamientos}} / CM_E$

es la estadística o variable pivotal para probar la hipótesis de igualdad de medias de los tratamientos.

Del valor esperado de la media de cuadrados se observa que, en general, CM_E es un estimador insesgado de σ^2 . Por otra parte, si la hipótesis nula es verdadera, $CM_{\text{Tratamientos}}$ resulta ser un estimador insesgado de σ^2 . Sin embargo, si la hipótesis nula es falsa, el valor esperado de $CM_{\text{Tratamientos}}$ es mayor que σ^2 .

Por tanto, el valor esperado del numerador en la estadística de prueba, es mayor que el valor esperado del denominador si la hipótesis alternativa es verdadera y, en consecuencia, debe rechazarse H_0 si el valor de tal estadística es demasiado grande. Esto implica una región crítica unilateral superior. En otras palabras, se rechaza H_0 si

$$F_c > F_{p,(t-1, N-t)} \text{ donde } F_{c_0} \text{ se calcula usando la Ecuación dada.}$$

Es posible obtener formulas de cálculo para las sumas de cuadrados al escribir y simplificar las definiciones