

TEMA 2: Diseño de experimentos

2.1.– Introducción

2.2.– Análisis de la varianza

2.2.1.– Introducción al ADEVA y al diseño de experimentos

2.2.2.– Pruebas de igualdad de varianzas en leyes normales

2.2.3.– Diseño completamente aleatorizado

- ADEVA simple: aplicaciones
- Modelos de diseño: efectos fijos y efectos aleatorios
- Conclusiones de un ADEVA significativo en cada modelo

2.2.4.– Contrastes múltiples simultáneos a posteriori

2.3.– Otros modelos de diseño de experimentos

2. 1. INTRODUCCIÓN

Los modelos de **Diseño de Experimentos** son modelos estadísticos clásicos cuyos objetivos son:

- (i) Averiguar si unos determinados factores influyen en una variable de interés y,
- (ii) En el caso de que haya influencia de algún factor, cuantificar dicha influencia.

Ejemplos donde podríamos utilizar estos modelos son:

- En el rendimiento de un determinado tipo de máquina (unidades producidas por día): se desea estudiar la influencia del trabajador que la maneja y la marca de la máquina.
- Una compañía de software está interesada en estudiar la variable *porcentaje en que se comprime un fichero, al utilizar un programa de compresión* teniendo en cuenta el tipo de programa utilizado y el tipo de fichero que se comprime
- Se quiere estudiar el rendimiento de los alumnos en una asignatura y, para ello, se desean controlar diferentes factores: profesor que imparte la asignatura; método de enseñanza; sexo del alumno.

El **objetivo del diseño de experimentos** es estudiar si cuando se utiliza un determinado tratamiento se produce una mejora en el proceso o no.

Para ello se debe experimentar aplicando el tratamiento y no aplicándolo. Si la variabilidad experimental es grande, sólo se detectará la influencia del uso del tratamiento cuando éste produzca grandes cambios en relación con el error de observación.

La **metodología del diseño de experimentos** se basa en la experimentación. Es sabido que si se repite un experimento, en condiciones indistinguibles, los resultados presentan una cierta variabilidad.

- * Si la experimentación se realiza en un laboratorio donde la mayoría de las causas de variabilidad están muy controladas, el error experimental será pequeño y habrá poca variación en los resultados del experimento.
- * Si se experimenta en procesos industriales o administrativos la variabilidad será mayor en la mayoría de los casos.

La **metodología del diseño de experimentos** estudia cómo variar las condiciones habituales de realización de un proceso empírico para aumentar la probabilidad de detectar cambios significativos en la respuesta; de esta forma se obtiene un mayor conocimiento del comportamiento del proceso de interés. Para que la metodología de diseño de experimentos sea eficaz es fundamental que el experimento esté bien diseñado.

TIPOS DE VARIABILIDAD

Uno de los principales objetivos de los modelos estadísticos y, en particular, de los modelos de diseño de experimentos, es controlar la variabilidad de un proceso aleatorio que puede tener diferente origen. De hecho, los resultados de cualquier experimento están sometidos a tres tipos de variabilidad:

- **Variabilidad sistemática y planificada:** Viene originada por la posible dispersión de los resultados debida a diferencias sistemáticas entre las distintas condiciones experimentales impuestas en el diseño por expreso deseo del experimentador. Es el tipo de variabilidad que se intenta identificar con el diseño estadístico.

Cuando este tipo de variabilidad está presente y tiene un tamaño importante, se espera que las respuestas tiendan a agruparse formando grupos (clusters). Es deseable que exista esta variabilidad y que sea identificada y cuantificada por el modelo.

- **Variabilidad típica de la naturaleza del problema y del experimento:** Es la variabilidad debida al ruido aleatorio. Este término incluye, entre otros, a la componente de variabilidad no planificada denominada error de medida. Es una variabilidad impredecible e inevitable. Esta variabilidad es la causante de que si en un laboratorio se toman medidas repetidas de un mismo objeto ocurra que, en muchos casos, la segunda medida no sea igual a la primera y, más aún, no se puede predecir sin error el valor de la tercera. Sin embargo, bajo el aparente caos, existe un patrón regular de comportamiento en esas medidas: todas ellas tenderán a fluctuar en torno a un valor central y siguiendo un modelo de probabilidad que será importante estimar. Esta variabilidad es inevitable pero, si el experimento ha sido bien planificado, es posible estimar (medir) su valor, lo que es de gran importancia para obtener conclusiones y poder hacer predicciones. Es una variabilidad que va a estar siempre presente pero que es tolerable.

Variabilidad sistemática y no planificada: Esta variabilidad produce una variación sistemática en los resultados y es debida a causas desconocidas y no planificadas. En otras palabras, los resultados están siendo sesgados sistemáticamente por causas desconocidas. La presencia de esta variabilidad supone la principal causa de conclusiones erróneas y estudios incorrectos al ajustar un modelo estadístico. Existen dos estrategias básicas para tratar de evitar la presencia de este tipo de variabilidad: la aleatorización y la técnica de bloques. Este tipo de variabilidad debe de intentar evitarse y su presencia lleva a conclusiones erróneas.

PLANIFICACIÓN DEL EXPERIMENTO

La experimentación forma parte natural de la mayoría de las investigaciones científicas e industriales, en muchas de las cuales, los resultados del proceso de interés se ven afectados por la presencia de distintos factores, cuya influencia puede estar oculta por la variabilidad de los resultados muestrales. Es preciso localizar los factores que influyen realmente y estimar esta influencia.

Etapas en el desarrollo de un diseño de experimento

1. Definir los objetivos del experimento.

2. Identificar todas las posibles fuentes de variación. Incluyendo:

(i) Factores tratamiento y sus niveles: Los factores tratamiento pueden ser cualitativos o cuantitativos.

Ejemplos de factores cualitativos: Fertilizantes (distintos tipos de fertilizantes), Sexo (hombre o mujer), tipo de procesador (procesadores con los que se quiere comparar su velocidad de ejecución)

Ejemplo de factores cuantitativos: Droga (distintas cantidades de droga); tamaño de memoria (diferentes tamaños). En el tratamiento matemático de los factores cuantitativos estos son tratados como cualitativos y sus niveles son codificados.

(ii) Unidades experimentales

(iii) Factores nuisance (molestos): factores bloque, factores ruido y covariables.

3. Elegir una regla de asignación de las unidades experimentales a las condiciones de estudio (tratamientos)

4. Especificar las medidas con que se trabajará (la respuesta), el procedimiento experimental y anticiparse a las posibles dificultades

5. Ejecutar un experimento piloto. Un experimento piloto es un experimento que utiliza un número pequeño de observaciones

6. Especificar el modelo

7. Esquematizar los pasos del análisis

8. Determinar el tamaño muestral

9. Revisar las decisiones anteriores. Modificarlas si se considera necesario

PLANIFICACIÓN DEL EXPERIMENTO

1. Definir los objetivos del experimento

Hay que hacer una lista completa de las preguntas concretas a las que debe dar respuesta el experimento. Deben ser cuestiones fundamentales. Esta lista se irá refinando a medida que se van ejecutando las etapas del procedimiento.

2. Identificar todas las posibles fuentes de variación.

(i) Factores tratamiento y sus niveles: Se denomina factor tratamiento a cualquier variable de interés para el experimentador cuyo posible efecto sobre la respuesta se quiere estudiar. Los niveles de un factor tratamiento son los tipos o grados específicos del factor que se tendrán en cuenta en la realización del experimento. Los factores tratamiento pueden ser cualitativos o cuantitativos.

(ii) Unidades experimentales: Son el material donde evaluar la variable respuesta y al que se le aplican los distintos niveles de los factores tratamiento.

(iii) Factores nuisance (molestos): factores bloque, factores ruido y covariables. Los factores nuisance son factores de escaso interés en el estudio, pero cuya influencia sobre la respuesta puede aumentar significativamente la variabilidad no planificada. Dependiendo de su naturaleza pueden ser de distintos tipos:

Factores bloques: Un factor tratamiento es un factor en el que se está interesado en conocer su influencia en la variable respuesta y un factor bloque es un factor en el que no se está interesado en conocer su influencia pero se incorpora al diseño del experimento para disminuir la variabilidad residual del modelo.

Covariables: Si el factor nuisance es una propiedad cuantitativa de las unidades experimentales que puede ser medida antes de realizar el experimento (el tamaño de un fichero informático, la presión sanguínea de un paciente en un experimento médico o la acidez de una parcela de tierra en un experimento agrícola). El factor se denomina covariable y juega un papel importante en el análisis estadístico.

Factores Ruido: Si el experimentador está interesado en la variabilidad de la respuesta cuando se modifican las condiciones experimentales, entonces los factores nuisance son incluidos deliberadamente en el experimento y no se aísla su efecto por medio de bloques. Se habla entonces de factores ruido.

3. Elegir una regla de asignación de las unidades experimentales a las condiciones de estudio (tratamientos)

La regla de asignación o diseño experimental especifica que unidades experimentales se observarán bajo cada tratamiento. Hay diferentes posibilidades:

- diseño factorial o no,

- anidamiento,

- asignación al azar en determinados niveles de observación, el orden de asignación, etc.

En la práctica, existen una serie de diseños estándar que se utilizan en la mayoría de los casos.

4. Especificar las medidas con que se trabajará (la respuesta), el procedimiento experimental y anticiparse a las posibles dificultades

5. Ejecutar un experimento piloto

Un experimento piloto es un experimento que utiliza un número pequeño de observaciones. El objetivo de su ejecución es ayudar a completar y chequear la lista de acciones a realizar. Las ventajas que proporciona la realización de un pequeño experimento piloto son las siguientes:

- (a) permite practicar la técnica experimental elegida e identificar problemas no esperados en el proceso de recogida de datos,
- (b) si el experimento piloto tiene un tamaño suficientemente grande puede ayudar a seleccionar un modelo adecuado al experimento principal,
- (c) los errores experimentales observados en el experimento piloto pueden ayudar a calcular el número de observaciones que se precisan en el experimento principal.

6. Especificar el modelo

El modelo matemático especificado debe indicar la relación que se supone que existe entre la variable respuesta y las principales fuentes de variación identificadas en el paso 2. Es fundamental que el modelo elegido se ajuste a la realidad con la mayor precisión posible.

El modelo más habitual es el modelo lineal:

$$Y = \sum_{i=1}^k \alpha_i + \epsilon$$

PLANIFICACIÓN DEL EXPERIMENTO

En este modelo la respuesta viene dada por una combinación lineal de términos que representan las principales fuentes de variación planificada más un término residual debido a las fuentes de variación no planificada. Los modelos que se estudian en este texto se ajustan a esta forma general. El experimento piloto puede ayudar a comprobar si el modelo se ajusta razonablemente bien a la realidad. Los modelos de diseño de experimentos, según sean los factores incluidos en el mismo, se pueden clasificar en:

1. Modelo de efectos fijos: Los factores del modelo son de efectos fijos.

En un factor de efectos fijos los niveles han sido seleccionados por el experimentador. Es apropiado cuando el interés se centra en comparar el efecto sobre la respuesta de esos niveles específicos. Ejemplo: un empresario está interesado en comparar el rendimiento de tres máquinas del mismo tipo que tiene en su empresa.

2. Modelo de efectos aleatorios: Los factores del modelo son de efectos aleatorios.

Un factor de efectos aleatorios, es un factor del que sólo se incluyen en el experimento una muestra aleatoria simple de todos los posibles niveles del mismo. Evidentemente se utilizan estos factores cuando tienen un número muy grande de niveles y no es razonable o posible trabajar con todos ellos. En este caso se está interesado en examinar la variabilidad de la respuesta debida a la población entera de niveles del factor. Ejemplo: una cadena de hipermercados que tiene en plantilla 300 trabajadores de caja está interesada en estudiar la influencia del factor trabajador en la variable tiempo en el cobro a un cliente.

3. Modelos mixtos: modelo en el que hay factores de efectos fijos y factores de efectos aleatorios.

7. Esquematizar los pasos del análisis.

Se deben esquematizar los pasos del análisis a realizar que deben incluir: estimaciones que hay que calcular, contrastes a realizar, intervalos de confianza que se calcularán, diagnóstico y crítica del grado de ajuste del modelo a la realidad.

8. Determinar el tamaño muestral.

9. Revisar las decisiones anteriores.

Modificarlas si se considera necesario. Los pasos del listado anterior no son independientes y en un determinado momento puede ser necesario volver atrás y modificar decisiones tomadas en algún paso previo.

RESUMEN DE LOS PRINCIPALES CONCEPTOS

Unidad muestral o experimental: son los objetos, individuos, intervalos de espacio o tiempo sobre los que se experimenta.

Variable de respuesta: es la variable que se desea estudiar y controlar su variabilidad.

Factor: son las variables independientes que pueden influir en la variabilidad de la variable de interés. Es el conjunto de condiciones a los que se someten las unidades

Factor tratamiento: es un factor del que interesa conocer su influencia en la respuesta.

Factor bloque: es un factor en el que no se está interesado en conocer su influencia en la respuesta pero se supone que ésta existe y se quiere controlar para disminuir la variabilidad residual.

Niveles: cada una de las condiciones de las posibles que tiene un factor. Según sean elegidos por el experimentador o elegidos al azar de una amplia población se denominan factores de efectos fijos o factores de efectos aleatorios.

Tratamiento: es una combinación específica de los niveles de los factores en estudio. Son, por tanto, las condiciones experimentales que se desean comparar en el experimento. En un diseño con un único factor son los distintos niveles del factor y en un diseño con varios factores son las distintas combinaciones de niveles de los factores.

Observación experimental: es cada medición de la variable respuesta.

Réplica de un nivel: cada una de las unidades muestrales que reciben ese nivel.

Tamaño del Experimento: es el número total de observaciones recogidas en el diseño.

Interacción de factores: existe interacción entre dos factores F_I y F_J si el efecto de algún nivel de F_I cambia al cambiar de nivel en F_J . Esta definición puede hacerse de forma simétrica y se puede generalizar a interacciones de orden tres o superior.

2. 2. ANÁLISIS DE LA VARIANZA

2.2.1. Introducción al ADEVA y al Diseño de Experimentos

Definición del ANOVA según Scheffé (1959)

“El análisis de la varianza es una técnica estadística que sirve para analizar medidas que dependen de varias clases de efectos que actúan simultáneamente, para decidir qué clases de efectos son importantes y estimar esos efectos. Las medidas pueden proceder de experimentos o de observaciones”.

El ANOVA es un contraste paramétrico global de varias hipótesis sobre igualdad entre las medias de varias distribuciones normales o entre los efectos producidos por las distintas condiciones de un diseño de experimentos

Este método estadístico fue introducido en 1925, por el estadístico-genetista Sir Ronald A. Fisher, pionero de la aplicación de los métodos estadísticos en el diseño de experimentos, que introdujo también el concepto de aleatorización.

El ANOVA simple es el tipo de ANOVA más sencillo, que se aplica al análisis de las posibles diferencias entre las medias de varias distribuciones normales y de los efectos que producen los distintos niveles del factor de un diseño de experimentos completamente aleatorizado.

COMPARACIÓN DE VARIAS MEDIAS

Se comparan I distribuciones normales con medias desconocidas $\mu_1, \mu_2, \dots, \mu_I$, con la misma varianza común desconocida, σ_e^2 , llamada varianza de los errores, tomando I muestras aleatorias independientes de tamaños n_1, n_2, \dots, n_I , para contrastar si existen diferencias significativas entre algunas de las I medias que se comparan.

DISEÑO DE EXPERIMENTOS

Un grupo de elementos de una población se divide al azar en varios subgrupos y cada subgrupo se somete a una combinación de condiciones elegidas y controladas por un investigador con el objeto de contrastar si las distintas combinaciones dan lugar a respuestas significativamente distintas respecto de un carácter medible en esos elementos.

El carácter medible se llama **variable de respuesta**. Los distintos tipos de condiciones que se analizan se llaman **variables explicativas o factores**. Cada factor se puede presentar en diferentes formas, grados o valores, que se llaman **niveles del factor**.

La técnica de este contraste consiste en descomponer una SUMA DE CUADRADOS TOTAL, que representa la variabilidad global de los datos, en suma de varias SUMAS DE CUADRADOS PARCIALES, que representan la variabilidad que se debe a las diferencias entre las medias o a los distintos efectos de las condiciones impuestas en un diseño.

Una variedad de trigo ha sido cultivada con 5 fertilizantes distintos, para estudiar el efecto de los mismos en la longitud de la planta. Se miden cinco series de 10 plantas. Los datos son los siguientes

F1	10	12	8	10	6	13	9	10	9	9
F2	11	18	12	15	13	8	15	16	9	13
F3	7	14	10	11	9	10	9	11	7	9
F4	12	9	11	10	7	8	13	14	10	11
F5	7	6	10	7	7	5	6	7	9	6

ELEMENTOS DE UN DISEÑO COMPLETAMENTE ALAETORIZADO CON UN SOLO FACTOR

Unidades muestrales: elementos que se han elegido en el diseño (plantas en este ejemplo)

Variable respuesta: la magnitud que se mide en cada unidad (la longitud de la planta)

Factor: Conjunto de condiciones de las posibles que tiene un factor (uso de fertilizantes)

Nivel: cada una de las condiciones de las posibles que tiene un factor (uno de los fertilizantes)

Réplica de un nivel: cada una de las unidades muestrales que reciben ese nivel

PLANTEAMIENTO DEL MODELO:

El planteamiento del ADEVA simple consiste en descomponer cada variable aleatoria X_{ik} en la suma

$$X_{ik} = \mu + \alpha_i + \epsilon_{ik}$$

donde

- X_{ik} representa el dato que encontraremos en la unidad k del nivel i
- μ es la media general de la población, un parámetro desconocido que representa un efecto común a todas las unidades
- α_i es el efecto producido por el nivel A_i
- ϵ_{ik} es el efecto del azar, que representa la variabilidad incontrolada de la unidad

La media de la i -ésima muestra o del nivel i es $\mu_i = \mu + \alpha_i$

HIPÓTESIS DEL ADEVA

H₀: todas las medias μ_i son iguales entre sí, o que todos los efectos α_i son nulos

$$\mu_1 = \mu_2 = \mu_3 = \cdots = \mu_I = \mu \quad \text{ó} \quad \alpha_i = 0 \quad \forall i$$

H₁: alguna de las medias μ_i es distinta de las otras, o que alguno de los efectos α_i son nulos

$$\text{existe } \mu_i \neq \mu_j \quad \text{ó existe } \alpha_i \neq 0$$

CONDICIONES DE VALIDEZ DEL ADEVA

Para que las conclusiones de un ADEVA sean válidas, los efectos debidos al azar, llamados residuales, deben ser variables aleatorias, denominadas variables de error, ϵ_{ik} , que verifiquen las tres condiciones siguientes:

Independencia: Todas estas variables deben ser independientes entre sí y respecto de los diferentes factores

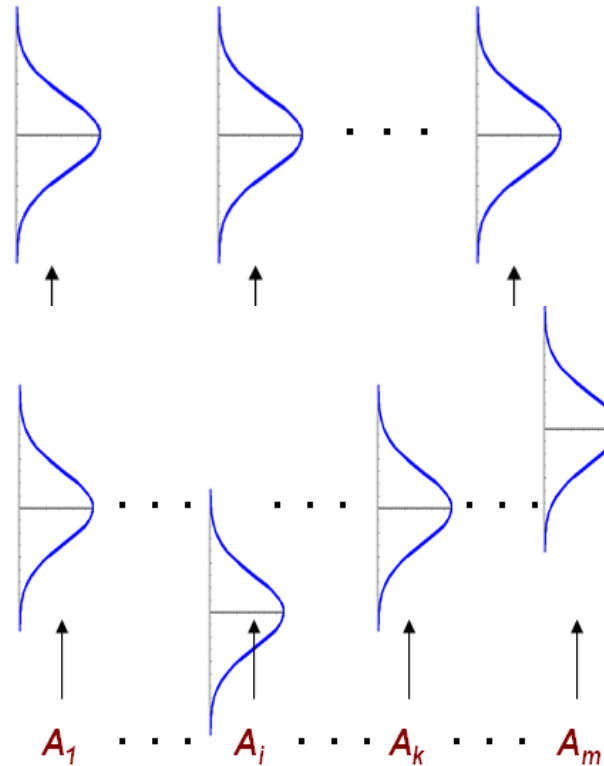
Normalidad: Cada una de estas variables debe seguir una ley normal de media nula

Homogeneidad o homocedasticidad: Las varianzas de las distribuciones de todos los errores son idénticas a un único valor, llamado varianza residual o varianza de los errores, σ_e^2

SUPUESTOS DE VALIDEZ DEL ANOVA

1 – **INDEPENDENCIA:** *Las variables aleatorias que representan la variabilidad propia de las unidades muestrales son mutuamente independientes.*

2 – **NORMALIDAD:**
Todas estas variables aleatorias siguen leyes Normales.



3 – **HOMOGENEIDAD DE LAS VARIANZAS:**
Todas estas leyes normales tienen la misma varianza, que se llama “varianza de los errores”

Equivale a decir que todas estas leyes normales tienen la misma forma.

Existen pruebas que contrastan la igualdad de las varianzas de varias leyes normales, como son:

La prueba de Cochran, que sólo sirve cuando los tamaños muestrales son iguales.

La prueba de Bartlett, que es útil cuando los tamaños muestrales no son iguales.

La prueba de Levene, que es válida en cualquier caso, incluso si no se cumple la normalidad

Estas tres condiciones son esenciales en la realización del ADEVA por las siguientes razones:

Normalidad: es esencial porque el ADEVA es un contraste paramétrico y, por lo tanto, la distribución del estadístico de contraste depende de la normalidad de las variables aleatorias ϵ_{ik} que representan la variabilidad de los individuos

Homogeneidad o homocedasticidad: Las varianzas de estas leyes normales deben ser iguales porque el contraste de igualdad de medias o efectos consiste en realizar dos estimaciones de esa varianza común llamada varianza residual

- Una de esas estimaciones sólo depende de la variabilidad individual de las unidades muestrales
- La otra estimación depende conjuntamente de la variabilidad individual y de los efectos de los distintos niveles

Si estas dos estimaciones son aproximadamente iguales, será más creíble la hipótesis nula de ausencia de efectos o de igualdad de medias.

Si la segunda estimación es mucho mayor que la primera, será más creíble la hipótesis alternativa de presencia de efectos o medias diferentes

Independencias entre las variables: es esencial para la construcción del estadístico de contraste.

2. 2. 3 Diseño completamente aleatorizado

El ADEVA simple se fundamenta en la separación de la variabilidad total de los datos en dos sumandos

$$SCT = SCE + SCD$$

donde

SCT es la **Suma de Cuadrados Total**, que mide la variabilidad de los posibles resultados respecto de la media global

SCE es la **Suma de Cuadrados Entre los niveles del factor**, que mide la variabilidad debida a los efectos de los niveles del factor

SCD es la **Suma de Cuadrados Dentro de los niveles del factor**, que mide la variabilidad debida al azar.

Para contrastar la hipótesis nula de igualdad entre los efectos de los niveles del diseño, en el ADEVA se calculan dos estimaciones de la varianza residual, y se comparan entre sí. Los estimadores de la varianza residual son:

MCD: media cuadrática intra-grupos $MCD = \frac{SCD}{N - I}$

Siempre es centrado porque sólo depende de los individuos

MCE: media cuadrática entre grupos $MCE = \frac{SCE}{I - 1}$

Es centrado si todos los niveles producen el mismo efecto.

Es sesgado si algunos niveles dan los efectos distintos.

Las dos estimaciones resultantes se comparan entre sí dividiendo la media cuadrática inter-grupos entre la media cuadrática intra-grupos

Por tanto, el estadístico de contraste del ANOVA es el cociente

$$F_{exp} = \frac{MCE}{MCD}$$

que sigue una ley F -Snedecor con $I - 1$ y $N - I$ grados de libertad, cuando H_0 es cierta.

F_{exp} sólo toma valores positivos. Cuanto más diferencia exista entre las medias o los efectos, mayor será este valor experimental por lo que este contraste es unilateral por la derecha.

La hipótesis nula se rechaza cuando este cociente es mayor que uno y supera el valor crítico asociado al nivel de significación fijado para el contraste ANOVA

CONSTRUCCIÓN DE LA TABLA DEL ADEVA SIMPLE:

Las distintas sumas de cuadrados, con sus grados de libertad, las medias de cuadrados y el valor observado del estadístico de contraste del ADEVA se resumen en la siguiente tabla, que se conoce como “**TABLA DEL ADEVA SIMPLE**”

Fuente de variación	Sumas de cuadrados	Grados de libertad	Medias de cuadrados	Valor de F_e
Diferencias entre grupos	SCE	$I - 1$	$MCE = \frac{SCE}{I - 1}$	$\frac{MCE}{MCD}$
Variación dentro de grupos	SCD	$n - I$	$MCD = \frac{SCD}{n - I}$	—
Variación total	SCT	$n - 1$	—	—

Si H_0 es cierta, F_e sigue una ley F de Snedecor con $I - 1$ y $n - I$ grados de libertad. El valor crítico del contraste, al nivel de significación α , es el percentil $F_{I-1, n-I; (1-\alpha)}$. Se rechaza la hipótesis nula, al nivel α , si $F_e \geq F_{I-1, n-I; (1-\alpha)}$.

Si se acepta H_0 , se admitirá que no existen diferencias entre los efectos de los distintos niveles. Si se rechaza H_0 , se admitirá que algunos niveles producen efectos distintos a los de otros niveles, pero no podemos determinar cuáles son los niveles que difieren entre sí.

La varianza de los errores, σ_e^2 , se estima con la media de cuadrados: $\hat{\sigma}_e^2 = MCD$.

Modelo I o de efectos FIJOS:

En este modelo los niveles del factor se eligen a voluntad para analizar los efectos que producen los niveles elegidos

Los efectos α_i se consideran parámetros fijos , aunque desconocidos, que verifican

$$\frac{1}{n} \sum_{i=1}^I n_i \alpha_i = 0$$

Si se quiere repetir un experimento en un modelo de este tipo, se volverán a elegir los mismos niveles del factor

Se contrasta la igualdad de los efectos α_i y de las medias $\mu_i = \mu + \alpha_i$ de los niveles elegidos

Hipótesis Nula:

Todos los efectos de los niveles A_i son nulos: $\alpha_1 = \alpha_2 = \dots = \alpha_I = 0$

Todas las medias de los I niveles son iguales: $\mu_1 = \mu_2 = \dots = \mu_I = \mu$

Hipótesis Alternativa:

Existe un nivel A_i con efecto $\alpha_i \neq 0$

Al menos existen dos niveles A_i, A_j con : $\mu_i \neq \mu_j$

El contraste sobre la igualdad de las medias de varias leyes normales independientes con la misma varianza se realiza de la misma forma que el ANOVA de un diseño de efectos fijos.

Un campus universitario tiene cuatro facultades. Se quiere estudiar la variable tiempo que tarda un alumno en hacer una consulta en la base de datos de la biblioteca de su facultad. Para ello se ha recogido una muestra aleatoria cuyos resultados son los de la tabla adjunta. Analizar estos datos y estudiar la influencia del factor facultad en la variable de interés.

Arq.	48	31	31	36	39	37	29	24	38	41					
Inf.	24	16	22	10	25	11	18	6	24	30	24	15			
Der	37	40	51	49	36	24	35	26	43	40	35	33	39	55	40
I.Cam.	19	13	26	21	31	26	13	24	12	12	16	21	30		

Modelo II o de efectos ALEATORIOS

En este modelo los niveles del factor se toman al azar, para extraer conclusiones sobre la variabilidad que existe entre los efectos producidos por todos los niveles de ese conjunto.

Los niveles de cada actor se toman al azar, entre un amplio conjunto de posibles niveles.

Los efectos α_i son variables aleatorias, independientes entre sí y de los errores, y siguen una misma ley normal, $N(0, \sigma_A^2)$, donde σ_A^2 es la varianza de los efectos de los niveles A_i , del factor A .

Si se quiere repetir un experimento en el modelo II, sus niveles se volverán a elegir al azar entre todos los del conjunto inicial.

Se contrasta la variabilidad de los efectos α_i de todos los niveles del conjunto inicial.

Hipótesis Nula: la varianza de los efectos debido a los niveles A_i es cero $\sigma_A^2 = 0$

Hipótesis Alternativa: la varianza de los efectos debido a los niveles A_i es positiva $\sigma_A^2 > 0$

En una empresa de montaje trabajan 135 operarios que realizan un determinado trabajo (T). La dirección de la empresa está interesada en conocer si influye el factor operario en la variable “tiempo de realización del trabajo T”. Para ello se eligen cinco operarios al azar y se les controla el tiempo en minutos que tardan en realizar el trabajo T en diez ocasiones. Los resultados del experimento son los de la tabla adjunta. ¿Qué conclusiones se deducen de este experimento?

OP1	72	75	71	69	67	71	75	73	69	65
OP2	75	70	77	73	79	77	72	78	73	69
OP3	78	79	84	72	83	77	80	83	71	85
OP4	69	65	61	75	70	68	67	63	76	72
OP5	65	60	63	68	70	64	62	64	69	62