

CAPÍTULO 2

MARCO TEORICO

2. MARCO TEÓRICO

En la presente sección se expondrán los conceptos estadísticos, analíticos y diferentes pruebas relacionadas con los mismos. El respaldo teórico de esta tesis es diverso y complejo por lo que algunos conceptos necesarios serán tratados en el presente capítulo. Sería imposible abarcar todos aquellos que el tema requiere, recalcando que ese no es el objetivo de esta investigación, por lo que se hablará fundamentalmente de aquellos que son los más relevantes.

Los conceptos a estudiar en el presente capítulo son:

- Concepto de Familia
- Comparaciones dependientes e independientes
- Diseño Completamente Aleatorizado.
- Alfa Nominal (α) VS Alfa Real ($\hat{\alpha}$)
- Pruebas de Rango Múltiple
- Problemática de las Pruebas de Rango Múltiple
- Una nueva prueba de rango múltiple a través de simulación. (Prueba D.E)
- Pruebas Protegidas y No Protegidas
- Tipos de errores
- Comportamiento de los errores en comparaciones independientes.
- Control de los diferentes tipos de error

2.1 Concepto de Familia

Cualquier colección de inferencias para las cuales sea significativo tomar en cuenta algunas medidas combinadas de errores se le llama familia. Ver Hochberg y Tamhane (1987). Por conveniencia también Hochberg y Tamhane (1987) utilizan el término de familia como una colección de problemas o parámetros sobre los cuales las inferencias son hechas como una familia.

La forma en la cual se construya una familia depende mucho del tipo de investigación que se quiera realizar. Si es puramente exploratoria, las preguntas de interés o líneas de inferencia son generadas por la exploración o curioso en los datos. Si es puramente confirmatoria las preguntas surgen conforme la investigación avanza. Algunos estudios empíricos combinan aspectos de ambos tipos de investigación.

Existen diferentes tipos de familias:

- Familias Finitas.- La familia consiste en una serie de comparaciones limitada a un valor real contable.
- Familias Infinitas.- Cuando se desea examinar cualquier contraste entre cualquier grupo de medias, la familia consiste en todos los contrastes posibles, cuya cardinalidad es infinita no contable.

2.2 Comparaciones Dependientes e Independientes.

En muchos de los métodos de comparaciones múltiples se utiliza el concepto de contraste, el cual es definido por Montgomery (2002) como aquel que utiliza comparaciones convexas de tratamientos. Un contraste estándar es una combinación lineal de parámetros de la forma:

$$\Gamma = \sum_{i=1}^a c_i \mu_i$$

donde las constantes de los contrastes c_1, c_2, \dots, c_a suman cero, es decir:

$$\sum_{i=1}^a c_i = 0$$

Para más detalles Ver Montgomery (2002).

Los contrastes o comparaciones de tratamientos pueden ser planteados antes de llevar a cabo el experimento o una vez que los datos han sido evaluados por el experimentador. En este caso las comparaciones generadas por los contrastes estándar se definen como comparaciones dependientes. Caso aparte se encuentran los contrastes ortogonales, dos contrastes con coeficientes $\{c_i\}$ y $\{d_i\}$ son ortogonales si:

$$\sum_{i=1}^a c_i d_i = 0$$

Para a tratamientos, en el conjunto de $a-1$ contrastes ortogonales se hace la partición de la suma de cuadrados debida a los tratamientos en $a-1$ componentes independientes con un solo grado de libertad, por lo tanto las pruebas realizadas con contrastes ortogonales son comparaciones independientes. Ver Montgomery (2002)

En los contrastes ortogonales por lo general debe existir un elemento que sugiera las comparaciones de interés, el planteamiento de esta comparación por lo general se lleva a cabo antes de realizar el experimento.

2.3 Diseño Completamente Aleatorizado

A manera de introducción a este punto es necesario definir el concepto de aleatorización. Para Montgomery (2002) “*La aleatorización es el hecho de que tanto la*

asignación del material experimental como el orden en que se realizan las pruebas individuales o ensayos se determinan aleatoriamente”

Una de las ventajas que acarrea trabajar con un diseño que se encuentre completamente aleatorizado es que logra que todas las observaciones involucradas en este se conviertan en variables aleatorias independientes, punto que resulta necesario en el uso de modelos estadísticos; otra ventaja alcanzada con esto es el de anular los efectos de factores no previstos que pudieran presentarse a lo largo de la investigación.

2.4 Alfa Nominal (α) VS Alfa Real ($\hat{\alpha}$)

Primeramente es necesario definir el concepto de alfa nominal (α) como aquel nivel de significancia que el investigador coloca al inicio de la prueba, el alfa real ($\hat{\alpha}$) es el verdadero tamaño estimado al cual se está realizando la prueba; lo ideal para cualquier experimento es que tanto el alfa nominal como el alfa real correspondan al mismo valor.

El Valor P se define como el mínimo nivel de significancia que llevará al rechazo de la hipótesis nula H_0 . Según Montgomery (2002) el Valor P, es la probabilidad de que la prueba estadística tomará en el valor que es al menos tan extremo como el valor observado del estadístico cuando H_0 es verdadera, por lo cual es posible asegurar que el Valor P cubre mucha información acerca del peso de la evidencia en contra de H_0 lo que permite elaborar una conclusión con respecto a cualquier nivel de significancia. En otras palabras este concepto se traduce como la probabilidad de que el valor de F de tablas sea mayor que el valor de F calculada.

A manera de resumen el Valor P es aquel valor que indica el mínimo nivel α para el cual los datos son significativos, liberando a cualquier investigador que retome el análisis sujetarse a un nivel de significancia preestablecido. El conocer este concepto ayuda a que el investigador que elaboró el análisis o investigadores que pudieran tomar

el caso en un futuro puedan reconocer que tan lejos o cerca se encuentra el resultado de la región de rechazo; por lo general en estas investigaciones no se reporta tanto nivel de detalle, de hecho en algunos casos solo se muestran conclusiones, lo que impone y predispone el nivel de α nominal elegido originalmente a otros experimentadores para el mismo análisis.

Obsérvese las figuras 2.1 y 2.2. En la figura 2.1, se ilustra el ejemplo de cuando la F calculada cae en la zona de aceptación.

Si $\hat{\alpha}$ real $>$ α nominal entonces la F Calculada se encuentra en la zona de Aceptación.

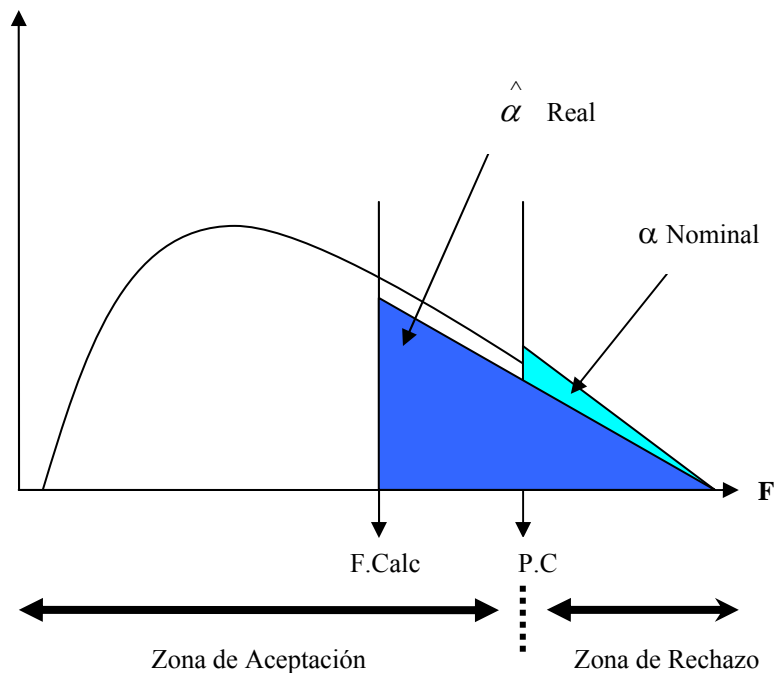


FIGURA 2.1 Caso en el que la F calculada cae en la zona de aceptación, donde el punto crítico (P.C) esta definido por α nominal.

En la figura 2.2 se muestra el caso contrario donde F calculada cae en zona de rechazo; en base al criterio de que si $\text{Valor } P < \alpha \text{ nominal}$ entonces la F Calculada se encuentra en la zona de Rechazo.

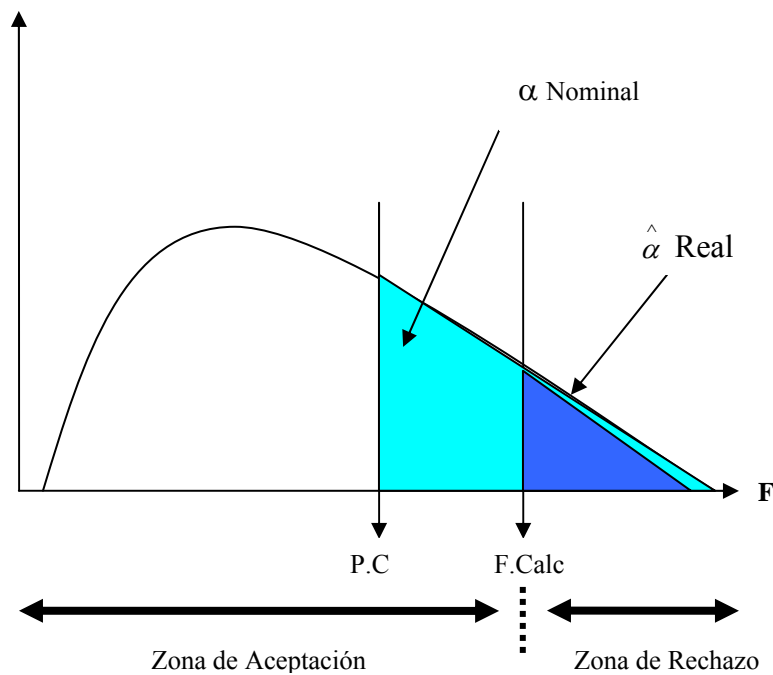


FIGURA 2.2. Caso en el que la F calculada cae en la zona de rechazo, donde el punto crítico (P.C) esta definido por α nominal.

2.5 Pruebas de Rango Múltiple.

El estudio de las comparaciones dependientes se realiza a través de las pruebas de rango múltiple, donde se desea probar una hipótesis nula:

$$H_0 : \mu_i = \mu_j, i \neq j$$

Lo anterior analizado a través del criterio de rechazo:

$$|\bar{Y}_i - \bar{Y}_j| > k(i, j, a, gl.error)S_{\bar{y}}$$

$i = i'$ smo tratamiento

$j = j'$ smo tratamiento

$a =$ tratamientos

$gl.error =$ grados libres error ($gl.total - gl.tratamientos$)

Donde:

$$S_{\bar{y}} = \sqrt{\frac{CM_E}{n}}$$

y la $k(i, j, a, gl.error)$ varía dependiendo la prueba que desee utilizarse, así como los parámetros i , j , a ó grados libres del error que se utilizan acorde a la prueba seleccionada. En general todas las pruebas de rango múltiple se describen de esta manera. A continuación se muestran las pruebas de rango múltiple más utilizadas y sus respectivos estadísticos:

2.5.1 Comparaciones entre Pares

- **Prueba de Tukey**

La Prueba de Tukey declara que dos medias son significativamente diferentes si el valor absoluto de sus diferencias muestrales excede:

$$T_\alpha = q_\alpha(a, f) \sqrt{\frac{CM_E}{n}}$$

Donde $q = \frac{\bar{y}_{\max} - \bar{y}_{\min}}{\sqrt{CM_E/n}}$, que es lo que se conoce como estadístico del rango

studentizado. Para más detalles sobre los valores de q consultar tablas en el Apéndice de Montgomery (2002).

- **Prueba DMS**

Este procedimiento utiliza el estadístico F para probar la H_0 de igualdad entre pares de medias.

$$DMS = t_{\alpha/2, N-a} \sqrt{CM_E \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

donde $t_{\frac{\alpha}{2}, N-a}$ es el $\left(1 - \frac{\alpha}{2}\right)100\%$ cuantil de la distribución t_{N-a} grados de libertad.

- **Prueba de Duncan**

El criterio de Duncan se muestra a continuación:

$$R_p = r_\alpha(p, f) \sqrt{\frac{CM_E}{n}}$$

Donde los valores de $r_\alpha(p, f)$ para $p = 2, \dots, a$ se obtienen de tablas; α es el nivel de significación y f el número de grados de libertad del error. Para más detalles sobre los valores de R consultar tablas en el Apéndice de Montgomery (2002).

- **Prueba SNK**

Procedimiento similar a la prueba de Duncan, salvo que las diferencias críticas entre las medias se calculan en una forma un tanto diferente, su criterio es el siguiente:

$$K_p = q_\alpha(p, f) \sqrt{\frac{CM_E}{n}}$$

- **Prueba de Bonferroni**

La desigualdad de Bonferroni, especifica que la probabilidad de unión de eventos es menor que la suma de las probabilidades de los eventos individuales, al especificar Bonferroni únicamente una cota la convierte en un procedimiento de un solo paso. Su criterio se muestra a continuación:

$$t_{\alpha, f}^* \sqrt{\frac{CM_E}{n}}$$

Donde:

$$t_{\alpha, f}^* = t_{\frac{\alpha}{2b}, f}$$

b = Número de comparaciones

f = Grados libres del error

t = Distribución de student con N-a grados de libertad.

2.5.2 Comparaciones con Contrastes

En la mayoría de los experimentos el investigador no sabe a ciencia cierta las comparaciones que desea realizar solo hasta que se ha llevado a cabo un examen preliminar de los datos, o en algunas ocasiones está interesado en realizar más de a-1 posibles comparaciones, es por ello que Scheffé (1959) propuso un método que entre uno de sus propósitos comparara cualquier tipo contraste.

- **Prueba Scheffé**

El método de Scheffé se utiliza para la comparación ya no solo entre pares de medias sino de contrastes completos, es decir combinaciones de medias. El criterio utilizado es:

$$S_{\alpha, u} = S_{C_u} \sqrt{(a-1)F_{\alpha, a-1, N-a}}$$

Donde:

$$S_{C_u} = \sqrt{CM_E \sum_{i=1}^a \frac{C_{iu}^2}{n_i}}$$

Para mayores detalles de las pruebas anteriores Ver Montgomery 2002.

2.6 Problemática de las Pruebas de Rango Múltiple

La gran mayoría de Pruebas de Rango Múltiple que actualmente se presentan en la literatura conllevan una serie de deficiencias debido a la dependencia existente entre los datos a comparar, a continuación se muestran los comentarios hechos por Burguete, Tamborero y García Pajares (1999)

➤ **DSM**

Para medidas alejadas $\hat{\alpha}$ real es mayor al α nominal y para medias cercanas el valor $\hat{\alpha}$ real es menor que el α nominal.

➤ **DUNCAN**

Esta prueba resulta muy parecida a la prueba DMS, pero su demostración matemática es incorrecta, Scheffé (1959).

➤ **SNK**

Es una prueba adecuada pero únicamente para comparar la media mínima contra la máxima, para el resto de las medias se vuelve restrictiva.

➤ **TUKEY**

Caso parecido a SNK donde resulta una prueba exacta, pero únicamente comparando la media mínima contra la máxima, esto resulta a que utiliza una misma métrica para comparar todas las medias, lo cual hace que la prueba se vuelva restrictiva.

➤ **BONFERRONI**

Resulta una prueba restrictiva para cualquier comparación, esto debido a que la prueba arroja el resultado de una cota.

➤ **SCHEFFE**

Restrictiva para cualquier comparación ya que usa un intervalo de confianza para el contraste que concentra toda la suma de cuadrados de los tratamientos

2.7 Una nueva prueba de rango múltiple a través de simulación. (Prueba D.E)

Con la problemática planteada anteriormente se muestra la complicación de las pruebas de rango múltiple tanto en su planteamiento como en su demostración, es por ello que Burguete, Tamborero y Morales (2003) presentan una prueba de rango múltiple basada en la simulación, logrando superar la complicación matemática del problema; aunque la prueba no genera resultados exactos si son aproximados y correctos, manteniendo el mismo nivel de confianza tanto para comparaciones externas como internas, esto debido al adecuado planteamiento de la prueba.

A esta prueba se le llamará prueba D.E (Prueba de la Diferencia Estudentizada), debido a que generaliza el rango estudentizado a cualquier comparación de dos medias. Burguete, Tamborero y Morales (2003), elaboraron un programa computacional en lenguaje Fortran para demostrar el funcionamiento de la prueba propuesta, dicho programa se basa en el algoritmo que se muestra a continuación:

- 1.- Solicitud de datos de entrada.- número de tratamientos (a), número de repeticiones (n), datos reales de un experimento que correspondan al número de tratamientos y repeticiones dados y número de simulaciones (B).
- 2.- Inicializar CONT = 0
- 3.- Generar las diferencias estudentizadas de los datos reales (DE)
- 4.- Generar a*n observaciones de una $\sim N(0,1)$, los autores manejan que es posible manejar una $\sim N(\mu,\sigma)$.
- 5.- Aplicar Estadísticos de Orden
- 6.- Calcular las Diferencias Estudentizadas de los datos simulados (DEs):

$$DES_{ij} = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{\frac{CME_s}{n}}}$$

7.- Comparar:

$$DE_s > DE \rightarrow \text{CONT} = \text{CONT} + 1$$

8.- Lo anterior se repite un número determinado de repeticiones = B

9.- Se calcula:

$$P.\text{Value} = \frac{\text{CONT}}{B}$$

La principal ventaja que muestra esta prueba, es que el P.Value o α estimada se mantiene constante a lo largo de todas las comparaciones, sin importan la distancia que exista entre las medias analizadas. Para reforzar lo anterior se menciona el concepto de Estadísticos de Orden que se muestra a continuación:

2.7.1 Estadísticos de Orden

Los estadísticos de orden, también conocidos como ordinales, son indispensables para el planteamiento y desarrollo dentro de las pruebas de rango múltiple ya que son importantes para ordenar las muestras aleatorias de las diversas poblaciones utilizadas.

Como su nombre lo indica su principal fin es “ordenar los datos” que serán utilizados dentro de los experimentos, esto en base a determinados criterios que pueden ser de orden creciente o decreciente. Para el caso de la presente investigación se utiliza el concepto de Mood y Graybill (1972):

“Sea X_1, X_2, \dots, X_n una muestra aleatoria procedente de la densidad $f(x)$ y suponga que se ordenan las Y_1, Y_2, \dots, Y_n en donde las Y_i son las mismas que las X_i dispuestas en orden de magnitud creciente $Y_1 < Y_2 < \dots < Y_n$. Las Y_1, Y_2, \dots, Y_n son conocidas como los estadísticos de orden u ordinales de la muestra aleatoria X_1, X_2, \dots, X_n “

Parzen (1960) comparte el criterio de orden creciente de Mood y Graybill (1972)

2.7.2 Prueba para contrastes y su equivalencia con las pruebas de rango múltiple.

Dentro de las pruebas de rango múltiple, a excepción de la prueba de Scheffé, el problema se centra fundamentalmente en la comparación entre pares de medias; pero se sabe que al momento que la prueba F resulta significativa es posible que el rechazo de la hipótesis nula se deba a una diferencia entre pares de medias o quizá a la presencia de algún contraste que resulte significativo. Las pruebas como Tukey, Duncan, DMS, SNK y Bonferroni no pueden utilizarse para el cálculo y análisis de contrastes y la única prueba que actualmente se conoce en la literatura capaz de realizar dicho análisis es la de Scheffé, ya que esta puede ser utilizada para detectar diferencia entre pares de medias o contrastes.

La Prueba de Scheffé, logra que el error Tipo I sea menor o igual a α en cualquiera de las posibles comparaciones entre las medias, resultando en una prueba conservadora, significando que rechaza poco y arroja en algunos casos resultados erróneos que llevarán al experimentador a tomar decisiones incorrectas. En la investigación elaborada en Ruelas (2000), utilizando las bases de la prueba D.E ya comentada, se plantea una prueba más exacta a la de Scheffé debido a su correcto planteamiento; aunque su cálculo sea aproximado si se realiza un número de simulaciones suficientes resulta adecuada.

El algoritmo de la prueba D.E para contrastes se explica a continuación:

- 1.- Solicitud de datos de entrada.- número de tratamientos (a), número de repeticiones (n), datos reales de un experimento que correspondan al número de tratamientos y repeticiones dados y número de simulaciones (B), colocar el contraste que se desea analizar.
- 2.- Cálculo de la ANOVA y obtención del CM_E
- 3.- Cálculo de las medias del experimento

- 4.- Ordenar las medias en base a estadísticos de orden.
- 5.- Leer el contraste.
- 6.- Para cada contraste obtener una F calculada :

$$F_{Calc} = \frac{SC_{Contraste}}{CM_E}$$

- 7.- Incrementar el número de simulaciones en uno $NB = NB + 1$.
Inicializar el número de repeticiones para la simulación $B = 0$
- 8.- $B = B + 1$, incrementar el número de repeticiones por simulación.
- 9.- Generar an números aleatorios $N(0,1)$ y asignar n de esos aleatorios a cada uno de los tratamientos, calcular las medias y aplicar estadísticos de orden a las mismas.
- 10.- Calcular el CM_E de los datos simulados.
- 11.- Calcular la SC del contraste (SC_{cont}^B) de los datos simulados para posteriormente calcular la F calculada para el contraste de los datos simulados (F_{calc}^B).
- 12.- Comparar:

$$F_{Calc}^B > F_{Calc} \Rightarrow Cont = Cont + 1$$

$$F_{Calc}^B < F_{Calc} \Rightarrow Continua$$

Lo anterior se lleva a cabo hasta que el número de repeticiones por simulación sea concluido.

- 13.- Calcular el Valor P:

$$ValorP = \frac{Cont}{B}$$

Esto se lleva a cabo hasta que el número total de simulaciones haya sido cubierto.

- 14.- Calcular los intervalos de confianza para los valores P.

2.8 Pruebas Protegidas y No Protegidas

Existen dos maneras diferentes de aplicar las pruebas de rango múltiple, a continuación se muestra el planteamiento de cada una de estas:

➤ Pruebas Protegidas

Se conocen como pruebas protegidas a aquellas pruebas de rango múltiple que se aplican después de llevarse a cabo un análisis de varianza y resultar la F significativa, a dichas pruebas también se les llama de dos pasos o post-ANOVA.

La mayoría de los autores, entre ellos Montgomery (2002) y Guitiérrez y De la Vara (2003), recomiendan utilizar las pruebas de rango múltiple como protegidas es decir una vez que la hipótesis nula de igualdad entre medias de todos los tratamientos ha sido rechazada.

➤ Pruebas No Protegidas

A estas pruebas se les conoce como de un solo paso, esto quiere decir que las pruebas de rango múltiple se aplican directamente sin realizar con antelación un análisis de varianza. Es posible calificar a estas pruebas como un análisis puro o menos manipulado, por lo cual los resultados se esperan diferentes a los arrojados por las pruebas de dos pasos.

2.9 Tipos de Errores

En esta sección se usará la siguiente notación:

\mathcal{F} = Familia de inferencias

\mathcal{P} =Procedimiento de comparación múltiple para esta familia.

$M(\mathcal{F}, \mathcal{P})$ = número aleatorio de inferencias equivocadas.

$N(\mathcal{F})$ = cardinalidad de la familia F (la cual solo es cierta en familias finitas).

Para más detalles ver Hochberg y Tamhane (1987).

Familiwise Error Rate (FWE)

Este error se define como la probabilidad de que por lo menos una de las comparaciones hechas dentro de la familia esté equivocada, es decir:

$$\text{FWE} (\mathcal{F}, \mathcal{P}) = \Pr \{ M (\mathcal{F}, \mathcal{P}) > 0 \}$$

Per-Family Error Rate

Cantidad que representa el valor esperado de errores dentro de una familia de inferencias, su representación acorde a la notación presentada es:

$$\text{PFE} (\mathcal{F}, \mathcal{P}) = E \{ M (\mathcal{F}, \mathcal{P}) \}$$

Per-Comparison Error Rate

Valor que representa la probabilidad de falla para cada una de las comparaciones en forma individual sin considerar la familia completa, esto es:

$$\text{PCE} = \frac{E \{ M (\mathcal{F}, \mathcal{P}) \}}{N (\mathcal{F})}$$

La relación de los errores anteriores puede ser vista como:

$$\text{PCE} \leq \text{FWE} \leq \text{PFE}$$

Lo anterior indica que el Per-Family resulta una cota superior para el Familiwise y el Per-Comparison representa una cota inferior para el mismo.

2.10 Comportamiento de los errores en comparaciones independientes.

Los diferentes tipos de error varían según la familia a la cual el investigador haga referencia y la relación entre las inferencias que integren a dicha familia (independientes o dependientes).

Para familias finitas y con inferencias mutuamente independientes, el porcentaje de error para una inferencia individual es α por lo tanto:

$$\text{PCE} = \alpha \quad (1)$$

$$\text{PFE} = \alpha N (F) \quad (2)$$

$$\text{FWE} = 1 - (1 - \alpha)^{N (F)} \quad (3)$$

Para mayor detalle ver Hochberg y Tamhane (1987).

Si α cercana a cero y $N (F)$ es pequeña, entonces $\text{FWE} \approx \alpha N (F)$, lo que corresponde al PFE. En la siguiente tabla, con un $\alpha = 0.05$, se muestran diferencias entre el FWE y el PFE conforme se incrementa $N (F)$

$N (F)$	PCE	FWE	PFE
1	0.05	0.05	0.05
5	0.05	0.23	0.25
10	0.05	0.40	0.50
15	0.05	0.54	0.75
20	0.05	0.64	1.00
30	0.05	0.78	1.50
40	0.05	0.87	2.00
50	0.05	0.92	2.50

Tabla 2.1. Comportamiento de los errores en comparaciones independientes.

Para el caso de familias infinitas el cálculo del PFE se convierte en un aspecto complicado, en el caso del PCE se considera el mismo que para cada subfamilia finita,

por lo tanto es posible considerar el PCE como un valor común para familias finitas e infinitas.

Cómo una observación importante que hacen Hochberg y Tamhane (1987) podría recalcarse que se asume de manera general como porcentaje de error, al error Tipo I (Probabilidad de rechazar H_0 dado que H_0 es cierta), a menos que un error diferente se especifique con antelación.

Con lo presentado anteriormente se puede conocer el nivel de α requerido para lograr un FWE deseado, simplemente despejando la ecuación (3), como se muestra en el siguiente ejemplo:

$$\begin{aligned} FWE &= 1 - (1 - \alpha)^{N(F)} \\ 1 - FWE &= (1 - \alpha)^{N(F)} \\ \sqrt[N(F)]{1 - FWE} &= 1 - \alpha \\ \alpha &= 1 - \sqrt[N(F)]{1 - FWE} \\ \alpha &= 1 - (1 - FWE)^{\frac{1}{N(F)}} \end{aligned}$$

Suponiendo que se desea un $FWE = 0.05$, en una familia de 63 comparaciones el α requerido se obtendría de la siguiente manera:

$$\alpha = 1 - (1 - 0.05)^{\frac{1}{63}} = .0008$$

2.11 Control de los diferentes tipos de error.

Al tratar el tema de los diferentes tipos de error, una de las preguntas más comunes a responder es ¿Cuál error es el que se debe de controlar en un proceso de comparación múltiple?

La respuesta a dicha pregunta varía demasiado según el punto de vista de los diferentes autores, por ejemplo Tukey (1953), después de examinar los diferentes tipos de errores llega a la conclusión de que el Familiwise (FWE) es el que debe ser controlado, argumentando: *“El FWE debe ser un error estándar, por lo tanto rara vez cualquier otro tipo de error será apropiado controlar”*, Tukey rechaza la importancia del PCE basándose en la idea de que los errores deben permitirse incrementar en proporción al número de comparaciones hechas, recordando que el PCE es una medida promedio del error encontrado en cada comparación.

Los argumentos que este autor utiliza para sustentar la defensa del FWE son:

- El controlar el FWE para la familia entera de inferencias potenciales asegura que la probabilidad de cualquier error en un conjunto de inferencias determinado, pueda estar controlado.
- Para una familia infinita el FWE puede ser controlado, caso contrario al PFE.
- Cuando los requerimientos de correcciones simultaneas para todas las inferencias sean satisfechas, el FWE es la única opción de control.

Otro autor que se promueve a favor de la idea de Tukey es Mollet (1981) que recomienda el FWE porque... *“La idea de que todas las comparaciones están correctas con una alta probabilidad parece dar serenidad al investigador, caso contrario lo que le provoca observar un número esperado de errores”*

Spjotvoll (1972), recomienda el uso del PFE para el caso de familias finitas, sustentando sus razones en los siguientes puntos:

- El PFE es técnicamente más fácil de trabajar que el FWE.
- El PFE impone una pena directa en base a la proporción de errores.
- El PFE es un límite superior del FWE por lo cual el control de este último va implícito en el control del PFE.

Por otra parte Duncan (1955) nombra aproximaciones para controlar los errores dentro de las familias finitas de inferencias; en esta aproximación el PCE controla los niveles de α tradicionales, pero el FWE al incrementar solamente como $1 - (1 - \alpha)^{N(F)}$ muestra el número de comparaciones estadísticamente independientes en F , reflejado en $N(F)$.

Como observación, es posible decir que el control de los diferentes tipos de error debe considerar la naturaleza y el propósito de la investigación.

Para los casos relacionados con investigaciones exploratorias, la idea es centrarse fundamentalmente en una validez estadística como argumentan Hochberg y Tamhane (1987), que será estipulada por los resultados de la exploración y el tipo de familia. Por validez estadística debe entenderse a los niveles de probabilidad del error Tipo I.

Cuando se incurre en el error Tipo I se comete el error de rechazar la hipótesis nula cuando esta debe ser aprobada; como caso contrario a esto existe el error Tipo II, sin embargo se emplea con mayor frecuencia el Tipo I debido a que es más fácil de analizar y controlar. Una consideración importante es el recalcar que la familia deberá ser especificada con antelación para poder controlar cualquier tipo de error Hochberg y Tamhane (1987).

Estos mismos autores argumentan que el FWE presenta una ventaja en el control por encima del PCE y del PFE, esto para el caso de familias infinitas, ya que controlar el FWE a un nivel α provee un límite superior de α en el conjunto de inferencias seleccionado, por el contrario el PCE no garantiza que la proporción de errores sea controlada en todo el conjunto de inferencias seleccionadas a ese nivel α y el PFE no es muy dócil de controlar en familias infinitas.

Para el caso de familias finitas es más sencillo controlar cualquiera de los 3 diferentes tipos de error, sin embargo Hochberg y Tamhane (1987) opinan que el FWE parece ser

el mejor candidato ya que provee el límite superior sobre cualquier nivel de una hipótesis seleccionada.

Es importante recalcar que sin importar el tipo de error que se decida controlar el experimentador tiene la facultad de elegir el nivel de α que desea emplear. Si este requiere una alta validez estadística para controlar sus errores los niveles tradicionales de α para las inferencias simples más comunes se encuentra en los rangos de $0.01 \leq \alpha \leq 0.10$, cuando el experimentador tolera una baja validez estadística puede controlar α a niveles más liberales como $0.10 \leq \alpha \leq 0.25$. (Rangos dados por Hochberg y Tamhane (1987) para controlar los niveles del FWE)

Para el caso de investigaciones confirmatorias por lo general se presentan con familias finitas, los tres diferentes tipos de error son buenos candidatos para controlarse, sin embargo para llegar a una conclusión correcta es necesario conocer las inferencias individualmente, esto lo mencionan Hochberg y Tamhane (1987) “ *Para que una decisión final sea la correcta, es necesario que todas las inferencias estén simultáneamente correctas entonces claramente el FWE puede ser controlado*”

A conclusión del análisis anterior, es posible decir que es necesario determinar con claridad la naturaleza del experimento así como lo que se desea alcanzar con la investigación, para que de esta manera se conozca acorde a las necesidades y propiedades del estudio el error que se requiere controlar.