Contrastes de Hipótesis

Método Científico y contrastación de hipótesis

De forma resumida y algo simplista, podemos decir que el método científico propone soluciones tentativas a los problemas, en forma de hipótesis, deduce de esas hipótesis consecuencias verificables que somete a comprobación y mantiene la hipótesis o la rechaza, de acuerdo con el resultado de la comprobación.

En general, se consideran estrategias de investigación más potentes aquellas que buscan obtener evidencia contradictoria con la hipótesis establecida, ya que basta un resultado de este tipo para poder afirmar la falsedad de la hipótesis, mientras que los resultados confirmatorios sólo mantienen provisionalmente la hipótesis, quedando siempre abierta la posibilidad de que comprobaciones posteriores la contradigan. Es decir, si un enunciado A implica otro enunciado B, la negación de B implica la negación de A:

$$Si\ \ A \Rightarrow B \quad Entonces\ \overline{B} \Rightarrow \overline{A}$$

Si expresamos esta regla en términos de ocurrencia de sucesos, puede que quede más claro su significado. Si siempre que ocurre A ocurre B, entonces el que yo observe que ha ocurrido B no me dice nada sobre la ocurrencia de A. Por el contrario, si en esas condiciones, observo que no ha ocurrido B podré derivar lógicamente que no ha ocurrido A. Ejemplo, siempre que llueve se moja el suelo, si observamos que el suelo está mojado eso no quiere decir que necesariamente halla llovido, pueden haber regado, puede haberse roto una tubería, etc. Sin embargo, si observo que el suelo está seco, eso indica, sin ninguna duda, que no ha llovido.

Dentro de la Naturaleza y especialmente en la investigación del campo psicológico las leyes no pueden establecerse de forma tan categórica, debido a la variabilidad inherente al proceso que se estudia o debido a la falta de control sobre algunos factores que lo determinan. En general, no podremos hacer afirmaciones del tipo *"siempre que ocurre A ocurre B"*, sino que nuestras afirmaciones serán de la forma *"siempre que ocurre A es probable que ocurra B y poco probable que no ocurra B"*. Los contrastes de hipótesis estadísticos estudian bajo que condiciones y en que forma una afirmación del tipo anterior nos puede conducir a que la comprobación de la no ocurrencia de B haga poco probable la ocurrencia de A.

Contrastes de Hipótesis Estadísticas

Si analizamos los datos de los estudios psicológicos desde un punto de vista probabilístico y con las herramientas de la Estadística Matemática, es porque numerosas características de los individuos no son determinísticas y en las mismas condiciones pueden manifestarse con distintos valores. Esto hace que tales características se nos presenten como variables aleatorias que toman distintos valores con diferentes probabilidades. En estas condiciones nuestras hipótesis tendrán que ser afirmaciones acerca de la distribución de esa variable en la población y la verificación de tales hipótesis se hará mediante la evidencia empírica que nos proporcionen los datos de una muestra. Definimos:

Hipótesis Estadística

Una hipótesis estadística es una afirmación sobre la distribución de un atributo o variable en la población.

Ejemplos de hipótesis estadísticas son afirmar que el C.I. medio de los alumnos de primero de Psicología es mayor que 100, o decir que el número de aprobados en el primer parcial de Análisis de Datos sigue una distribución Binomial. En el primer caso se trata de una afirmación relativa al valor de un parámetro de la distribución, mientras que la segunda hace referencia a la forma de la distribución. Nosotros nos centraremos en el estudio de las hipótesis estadísticas relativas a un parámetro de la distribución de la población. Desde esta perspectiva podemos diferenciar dos contextos o modelos de actuación.

Contrastes Paramétricos

Son procedimientos que contrastan una afirmación acerca del valor de un parámetro, pero suponiendo conocida la forma de la distribución. Por ejemplo, contrastar la anterior hipótesis de que la media del C.I. de los alumnos es mayor que 100, suponiendo que la distribución de las puntuaciones es una distribución Normal.

Contrastes No Paramétricos

En estos métodos la hipótesis que se somete a prueba puede ser la misma que en el caso anterior, pero no se especifica la forma de la distribución de la población, a lo más, se exige como condición que la distribución sea continua.

Es obvio que las hipótesis relativas a un parámetro de la población podrían someterse a verificación midiendo la característica en cuestión, en todos los individuos de la población, y calculando el valor del parámetro, pero por las razones que se expusieron en el tema de inferencia, imposibilidad de acceder a la totalidad de la población, economía de los estudios y mayor calidad de los datos, en general para someter a prueba una hipótesis, observaremos los valores de una muestra, convenientemente seleccionada, y determinaremos si la hipótesis establecida es compatible o no con los datos de la muestra. En este sentido definimos:

Contraste de Hipótesis

Es una regla de decisión que nos indica si una hipótesis relativa a la población es compatible (en términos de probabilidad) con los datos de la muestra, o si es incompatible con los mismos y en consecuencia debe ser rechazada.

Señalaremos que una regla de decisión es una función que asigna a cada conjunto de datos la acción a emprender. En general, las reglas de decisión son construidas de tal forma que la acción que seleccionan es óptima, de acuerdo con algún criterio prefijado.

Si recordamos que las hipótesis estadísticas que vamos a considerar son afirmaciones acerca de los parámetros de la distribución de la población, podremos ver que toda hipótesis se identifica, o viene determinada por un subconjunto del espacio paramétrico.

Espacio Paramétrico

El espacio paramétrico, que designaremos por Ω, es el espacio o conjunto donde toma valores el parámetro.

Por ejemplo, si el parámetro que estamos considerando, es la media de una distribución Normal, el espacio paramétrico será la recta real. Si es el número de pruebas de una distribución Binomial, será el conjunto de los números naturales, etc.

Al subconjunto del espacio paramétrico que identifica una hipótesis, se le designa por ω.

Una vez establecida una hipótesis y su correspondiente subconjunto del espacio paramétrico, surge automáticamente su hipótesis alternativa que estará definida por el resto del espacio paramétrico Ω-ω.

Ejemplo: Cuando asegurábamos que el C.I. medio de los alumnos de primero de Psicología estaba por encima de 100, establecíamos una hipótesis que se correspondía con el intervalo $(100, \infty)$ que es un subconjunto de la recta real $\Re$. La hipótesis alternativa sería la afirmación contraria de que el C.I. medio, de estos mismos alumnos, es menor o igual que 100, a la cual correspondería el intervalo $(-\infty, 100]$ que es precisamente igual a $\Re$ - $(100, \infty)$.

Hipótesis simple

Una hipótesis se dice que es simple cuando consta de un solo punto del espacio paramétrico. Por ejemplo, decir que el C.I. medio de los alumnos de primero es igual 112 es enunciar una hipótesis simple.

Hipótesis compuesta

Una hipótesis es compuesta cuando consta de más de un punto del espacio paramétrico. Por ejemplo, la hipótesis ya vista de que la media de los C.I. de los alumnos es mayor que 100, sería una hipótesis compuesta.

Hipótesis Nula

La hipótesis que se somete a contraste se denomina Hipótesis Nula y se designa por H_0.

Por razones del método que se sigue para contrastar una hipótesis, la hipótesis nula debe ser una hipótesis de igualdad, de no diferencia entre las poblaciones, de falta de efectos de los tratamientos, es por esta razón que se denomina nula. la Hipótesis Nula H_0 es la que se somete a contraste y por tanto será la que aceptaremos o rechazaremos.

Hipótesis Alternativa

Es la contraria de la Hipótesis Nula. Esta hipótesis afirma la existencia de una diferencia entre las poblaciones, o la presencia de un efecto no nulo de los tratamientos y suele ser la hipótesis que desea probar el experimentador. Se le designa por H_1.

Ejemplos: Supongamos que queremos verificar si existen diferencias, por término medio, entre el C.I. de las alumnas y de los alumnos. Independientemente de cual sea la hipótesis que nos interese probar, que las alumnas tienen mayor C.I. que los alumnos, o viceversa que es mayor en los alumnos, la hipótesis que debemos someter a contraste, es decir la Hipótesis Nula, deberá ser que ambos C.I. medios son iguales:

$$H_0: \ \overline{CI_M} = \overline{CI_H}$$

Nuestras intenciones podrán reflejarse en la Hipótesis Alternativa que en este caso podría adoptar cualquiera de las siguientes formas:

$$H_1: \ \overline{CI_M} \neq \overline{CI_H}$$

$$H_1: \ \overline{CI_M} > \overline{CI_H}$$

$$H_1: \ \overline{CI_M} < \overline{CI_H}$$

La aparente contradicción de que puedan existir distintas Hipótesis Alternativas a una misma Hipótesis Nula, será explicada posteriormente cuando hallamos visto los conceptos de potencia y nivel de confianza de un test.

De igual forma, si quisiéramos contrastar que el C.I. medio de una población es distinto de 100 tendríamos que establecer como Hipótesis Nula:

$$H_0: \ \mu_{CI} = 100$$

Mientras que la Hipótesis Alternativa podría ser:

$$H_1: \ \mu_{CI} \neq 100$$

$$H_1: \ \mu_{CI} > 100$$

$$H_1: \ \mu_{CI} < 100$$

Estadístico de Contraste

La aceptación o rechazo de la Hipótesis Nula se hace considerando su compatibilidad o incompatibilidad con los datos muestrales, pero esta decisión no se establece mediante la inspección directa de los datos muestrales, sino a partir de los valores de un estadístico, recuérdese que un estadístico es una función de los valores de la muestra. Si el estadístico es menor que un cierto valor crítico **C**, se aceptará la Hipótesis Nula H_0. Por el contrario, si el estadístico es mayor que el valor crítico, se rechazará la Hipótesis Nula. De esta forma el problema de tomar la decisión a partir de la información multidimensional que proporciona la muestra, que si se abordase directamente obligaría a definir una región del espacio muestral que es n-dimensional, se transforma en un problema unidimensional y para establecer la regla de decisión basta con establecer un punto de corte en la recta real, que si es superado indica que debe rechazarse la hipótesis. El valor critico **C** se determinará atendiendo a minimizar, en su conjunto, las probabilidades de los distintos tipos de error que pueden producirse.

Requisitos del Estadístico de Contraste

Para poder construir adecuadamente la regla de decisión y mantener controladas las probabilidades de error, antes mencionadas, el estadístico de contraste debe de cumplir las siguientes condiciones:

Deberá de ser elegido de manera que pueda ser conocida su distribución, cuando sea cierta la Hipótesis Nula.

Cuando la Hipótesis Nula sea falsa su distribución debe ser diferente. El estadístico será más útil cuanto más fácil sea discernir esta diferencia, por lo que es preferible que ambas distribuciones se hallen diferenciadas por sus promedios y no por su forma.

Ejemplo: Supongamos que tenemos una población, en la cual el C.I. sigue una distribución Normal de media μ y desviación típica σ, y queremos contrastar si la media del C.I. vale 100. Por tanto la Hipótesis Nula será:

$$H_0 : \ \mu_{CI} \ = \ 100$$

El estadístico de contraste para esta hipótesis será:

$$g(x_1, x_2, \ldots x_n) \ = \ \frac{\bar{x} - 100}{\frac{S}{\sqrt{n-1}}}$$

En virtud del teorema de Fisher sabemos que el estadístico:

$$t \ = \ \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n-1}}}$$

sigue una distribución t de Student con n-1 grados de libertad. Por consiguiente, cuando la Hipótesis Nula sea cierta y la media de la población μ sea 100, tendremos que ambos estadísticos coincidirán y por tanto el estadístico de contraste seguirá una t de Student. Por el contrario, cuando la Hipótesis Nula sea falsa, estaremos restando una cantidad, que será distinta del valor de la media de la población, y el estadístico de contraste seguirá una distribución que ya no tendrá de media cero, y por ello estará desplazada, a la derecha o a la izquierda, de la t de Student.

Región Crítica o de rechazo

Es el conjunto de muestras, o puntos del espacio muestral, que determinan un valor del estadístico de contraste superior al valor crítico y por tanto conllevan la decisión de rechazar la Hipótesis Nula. Se designa por W_n y será:

$$W_n \quad = \quad \{(x_1, x_2, \ldots x_n) / g(x_1, x_2, \ldots x_n) > C\}$$

Región de Aceptación

Es el conjunto de puntos del espacio muestral es decir, las diferentes muestras, que hacen que el estadístico tome un valor inferior o igual al valor crítico y que por consiguiente conducen a que se acepte la Hipótesis Nula. Si designamos por R_n al espacio muestral, tendremos:

$$R_n - W_n \quad = \quad \{(x_1, x_2, \ldots x_n) / g(x_1, x_2, \ldots x_n) \leq C\}$$

Tipos de Error

Como hemos visto, en un contraste de hipótesis podemos tomar dos decisiones, rechazar la Hipótesis Nula o aceptarla. A su vez la Hipótesis Nula puede ser verdadera o puede ser falsa, lo cual conduce a que se puedan producir cuatro resultados en un test de hipótesis, de los cuales dos son correctos y dos son erróneos. De forma esquemática

	H_0 Verdadera	H_0 Falsa
Rechazar H_0	Error Tipo I	Rechazo correcto
Aceptar H_0	Aceptación correcta	Error Tipo II

Error Tipo I

Es el error que se comete cuando se rechaza una Hipótesis Nula que es cierta.

Error Tipo II

Es el error que se comete al aceptar una Hipótesis Nula que es falsa.

Es obvio, que lo ideal sería contar con un método de contrastación en el que no se cometieran errores, pero la naturaleza del procedimiento, que basa su decisión en los datos de una muestra, hace que siempre exista la posibilidad de tomar la opción equivocada debido a las fluctuaciones muestrales. Lo relevante es conseguir tener controlado este riesgo y en este sentido definimos:

Nivel de Significación

Es la probabilidad de rechazar una Hipótesis Nula que es cierta o lo que es lo mismo, es la probabilidad de cometer Error de Tipo I. Se designa habitualmente por α.

Nivel de Confianza

Es la probabilidad de aceptar la Hipótesis nula cuando es verdadera. Su valor es, uno menos el nivel de significación, 1-α.

Potencia de un contraste

Es la probabilidad que tiene el contraste de rechazar la Hipótesis Nula cuando es falsa. Será igual a uno menos la probabilidad de cometer Error de Tipo II, como esta última se suele designar por β, la potencia será 1- β.

Criterio de Selección

A la hora de seleccionar un test estadístico, para contrastar una hipótesis, el óptimo sería aquel que hiciese mínimas la s probabilidades de Error de Tipo I y Tipo II. Sin embargo, esto no es posible ya que al disminuir el nivel de significación se reduce la región crítica y por consiguiente se pierde potencia. En sentido contrario, para disminuir la probabilidad de Error de Tipo II, debe aumentar el tamaño de la región crítica lo que conduce a que aumente el nivel de significación. Se impone por tanto una solución de compromiso entre ambos tipo de error. El criterio establecido parte del hecho de que la distribución del estadístico es conocida, cuando la Hipótesis Nula es cierta, por tanto podemos fijar el nivel de significación en el valor que deseemos. Una vez establecido el nivel de significación que queremos, de entre todos los tests existentes, con ese nivel de significación, seleccionaremos el que tenga máxima potencia.

Para ilustrar el criterio anterior volvamos al ejemplo en que queríamos contrastar si la media del C.I. era igual a 100. Sabíamos que cumpliéndose la Hipótesis Nula :

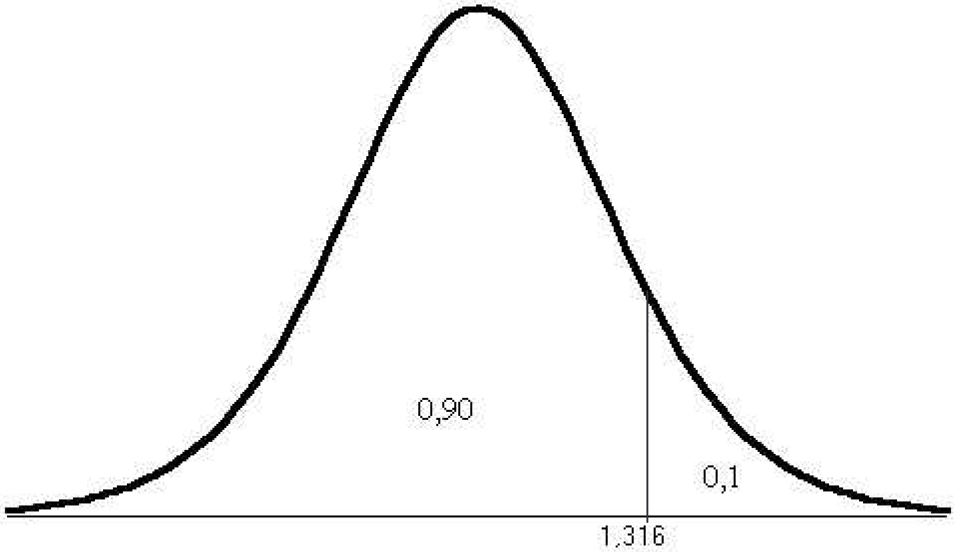
$$H_0: \ \mu_{CI} = 100$$

el estadístico de contraste:

$$t = \frac{\bar{x} - 100}{\frac{S}{\sqrt{n-1}}}$$

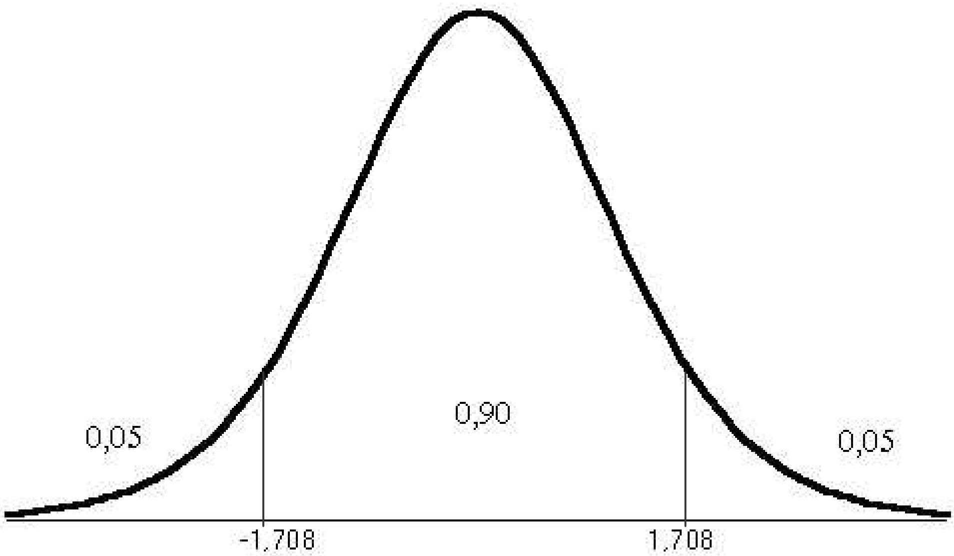
seguía una t de Student con n-1 grados de libertad.

Partiendo de estas premisas puedo fijar el nivel de significación que desee y obtener el valor crítico correspondiente. En efecto, supongamos que la muestra es de tamaño 26 y que fijamos el nivel de confianza en el 90%, que es igual que decir que el nivel de significación será del 10%. Buscando en las tablas de la t de Student, con 25 grados de libertad, obtengo que 1,316 es el valor de la variable que deja por debajo de si una probabilidad de 0,9 y por consiguiente, la probabilidad de obtener un valor mayor que 1,316 es 0,1.



Queda así definido el test, cuya región de aceptación será el conjunto de muestras que hagan que el estadístico tome un valor inferior o igual a 1,316. La región crítica será el conjunto de puntos del espacio muestral para los cuales el valor del estadístico sea mayor que 1,316, caso de que la muestra obtenida al realizar el test pertenezca a este conjunto, rechazaremos la Hipótesis Nula.

Las dificultades surgen porque el test, anteriormente descrito, no es el único con un nivel de significación de 0,1. Por ejemplo, si yo defino como región de aceptación el conjunto de muestras que hacen que el estadístico esté comprendido entre -1,708 y 1,708 y como región de rechazo aquellas muestras que hacen que el estadístico sea mayor que 1,708 o menor que -1,708, tendré, como puede verse en la figura, un contraste distinto del anterior, pero que también posee el nivel de significación del 10%.

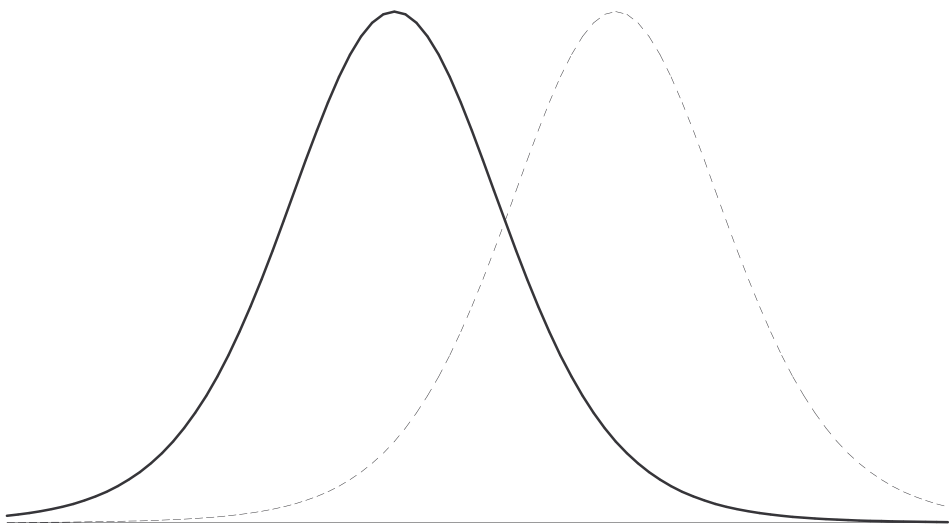


Se comprueba fácilmente que el número de tests con un nivel de significación del 10% es infinito, hay tantos tests como formas de distribuir la probabilidad 0,1 entre las dos colas de la distribución, 0,07 y 0,03, 0,06 y 0,04, etc.

La solución a esta dificultad es seleccionar de entre todos estos tests, con el mismo nivel de significación, el más potente y para ello entra en juego la hipótesis alternativa que se considere. Supongamos que se verificase la Hipótesis Alternativa de que la media es mayor que 100.

$$H_1: \quad \mu_{CI} \quad > \quad 100$$

entonces, en el numerador del estadístico, estaríamos restando una cantidad inferior a la media y por consiguiente la distribución del estadístico ya no tendría de media cero, sino un número positivo y estaría desplazada a la derecha de la t de Student, como vemos en la figura:

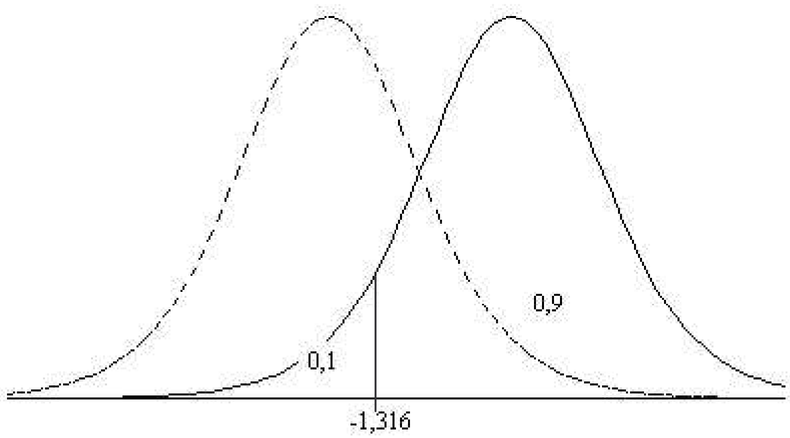


Por ello, el primer contraste que vimos, que tenía toda la región crítica a la derecha, será el más potente, ya que sitúa la región crítica en los valores más probables bajo la Hipótesis Alternativa.

De forma análoga, cuando la Hipótesis Alternativa sea que la media es menor que 100:

$$H_1: \quad \mu_{CI} \quad < \quad 100$$

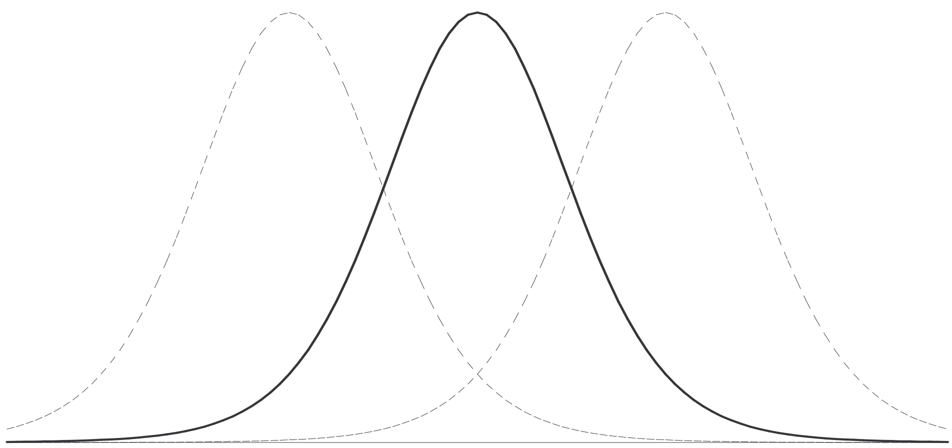
tendremos que si la Hipótesis Alternativa es cierta la distribución del estadístico estará desplazada a la izquierda de la t de Student y por consiguiente el test más potente sería aquel que tiene como región crítica los puntos que hacen que el estadístico tome valores inferiores a -1,316.



Por el contrario, si la Hipótesis Alternativa, a considerar, es que la media es distinta de 100:

$$H_1: \ \mu_{CI} \neq 100$$

si se verifica la Hipótesis Alternativa, la distribución del estadístico puede situarse tanto a la derecha como a la izquierda de la t de Student.



Por ello el contraste de hipótesis que tiene como región de aceptación el intervalo [-1,708 , 1,708] es, de entre todos los tests con un nivel de confianza del 90%, el más potente. En este caso la región crítica es el conjunto de muestras que hacen que el estadístico sea mayor que 1,708 o menor que -1,708, por ello este tipo de tests se denominan contrastes bilaterales o de dos colas, mientras que los anteriores son conocidos como contrastes unilaterales o de una sola cola.

Contrastes Bilaterales y Unilaterales

Como hemos visto, el elemento determinante para que el contraste más potente sea un contraste bilateral o unilateral, es la forma que adopte la Hipótesis Alternativa.

Nos que da por dilucidar como es que una misma Hipótesis Nula puede aparentemente tener distintas Hipótesis Alternativas.

En realidad a la Hipótesis Nula:

$$H_0: \ \mu_{CI} \ = \ 100$$

le corresponde únicamente la Hipótesis Alternativa:

$$H_1: \ \mu_{CI} \ \neq \ 100$$

y las otras Hipótesis Alternativas corresponden a contrastes en que se someten a prueba Hipótesis Nulas diferentes, en concreto los contrastes serían:

$$H_0: \ \mu_{CI} \ \leq \ 100 \quad \text{frente a} \quad H_1: \ \mu_{CI} \ > \ 100$$

$$H_0: \ \mu_{CI} \ \geq \ 100 \quad \text{frente a} \quad H_1: \ \mu_{CI} \ < \ 100$$

Lo que sucede es que en estos contraste la Hipótesis Nula consta de más de un punto del espacio paramétrico y no tendríamos un solo nivel de significación, sino tantos como valores de la media considerásemos. Por ello el cálculo de nivel de significación se hace para la situación más desfavorable que correspondería al caso en que $\mu = 100$, y si se rechaza la Hipótesis Nula de que la media es 100, con mayor razón se haría para los valores inferiores, en el caso del primer test. Un argumento semejante puede realizarse para el segundo.