**PREGUNTAS:**

**UNIDAD 6 – Prueba de Hipótesis**

**Definición de hipótesis estadística**   
**¿Qué es una hipótesis estadística?**

Una hipótesis estadística es una **aseveración o conjetura** respecto a una o más poblaciones.  
Conjetura: Juicio u opinión formada a partir de indicios o datos incompletos o supuestos.

La **verdad o falsedad** de una hipótesis estadística **nunca se sabe con absoluta certeza**, a menos que se examine toda la población, lo cual, por supuesto, sería poco práctico en la mayoría de las situaciones.

En vez de eso se toma una muestra aleatoria de la población de interés y se utilizan los datos contenidos en ella para proporcionar evidencia que respalde o no la hipótesis.   
Si la evidencia de la muestra es **inconsistente** con la hipótesis planteada conduce al rechazo de la misma.

La idea de que el rechazo de una hipótesis implica que fue **refutada por la evidencia** de la muestra.   
Esto quiere decir que **el rechazo significa que existe una muy pequeña probabilidad** de obtener la información muestral observada cuando, de hecho, la hipótesis es verdadera.  
Es muy importante enfatizar que **la aceptación es en realidad la falta de evidencia para realizar un rechazo**, esto quiere decir que no descarta otras posibilidades.

Lo anterior implica que cuando el analista de datos formaliza la evidencia experimental con base en la prueba de hipótesis, es **muy importante el planteamiento formal de la hipótesis.**¿A qué se llama hipótesis nula y a qué hipótesis alternativa?

**La hipótesis nula y la hipótesis alternativa**

La estructura de la prueba de hipótesis se establece usando el término hipótesis nula, el cual se refiere a cualquier hipótesis que se desea probar y se denota con H0.

El rechazo de H0 conduce a la aceptación de una hipótesis alternativa, que se denota con H1 o HA.

La hipótesis alternativa H1 por lo general representa la pregunta que se responderá o la teoría que se probará, por lo que su especificación es muy importante. La hipótesis nula H0 anula o se opone a H1 y a menudo es el complemento lógico de H1. Se llegará a una de las siguientes dos conclusiones:  
 **- Rechazar H0** a favor de H1 debido a **evidencia suficiente** en los datos  
 - Aceptar H0 o mejor dicho **No rechazar H0 debido a evidencia insuficiente** en los datos.

Las conclusiones no implican una “aceptación de H0” formal y literal.

Buen ejemplo: Este es un dilema que enfrenta el jurado en un juicio:

H0: el acusado es inocente.   
H1: el acusado es culpable.

La acusación proviene de una sospecha de culpabilidad. La hipótesis H0 (el status quo) se establece en oposición a H1 y **se mantiene a menos que se respalde H1** **con evidencia** “más allá de una duda razonable”. Sin embargo, en este caso “**no rechazar H0**” no implica inocencia, sino sólo que **la evidencia fue insuficiente** para lograr una condena. Por lo tanto, el jurado **no necesariamente acepta H0, sino que no rechaza H0**.  
El rechazo significa que existe una pequeña probabilidad de obtener la información muestral observada cuando, de hecho, la hipótesis es verdadera.

Como resultado, el analista de datos establece una conclusión firme cuando se rechaza una hipótesis.

**Estadístico de prueba**

El **estadístico de prueba** se basa en el estadístico que utilizamos para realizar conclusiones con respecto a la hipótesis nula. Nuestra decisión es X. Consideramos ciertos valores críticos todos los posibles valores entre dichos valores conforman la región de aceptación, todos los valores los posibles valores que sean excedan dichos valores (Ya sea por derecha o por izquierda) constituyen la **región crítica**.  
El último número que observamos al pasar a la región de aceptación a la region crítica se llama **valor crítico**.

**¿Cómo se eligen las hipótesis nula y alternativa?**Con frecuencia **la hipótesis nula H0 se plantea usando el signo de igualdad**.   
Con este método se observa claramente cómo se controla la probabilidad de cometer un error tipo I.   
Sin embargo, hay situaciones en que “no rechazar H0” implica que el parámetro θ podría ser cualquier valor definido por el complemento natural de la hipótesis alternativa. Sin embargo, es evidente que en el caso de las pruebas de una cola la consideración más importante es el planteamiento de la alternativa.

**La decisión de plantear una prueba de una cola o una de dos colas depende de la conclusión que se obtenga si se rechaza H0.**   
La ubicación de la región crítica sólo se puede determinar después de que se plantea H1.

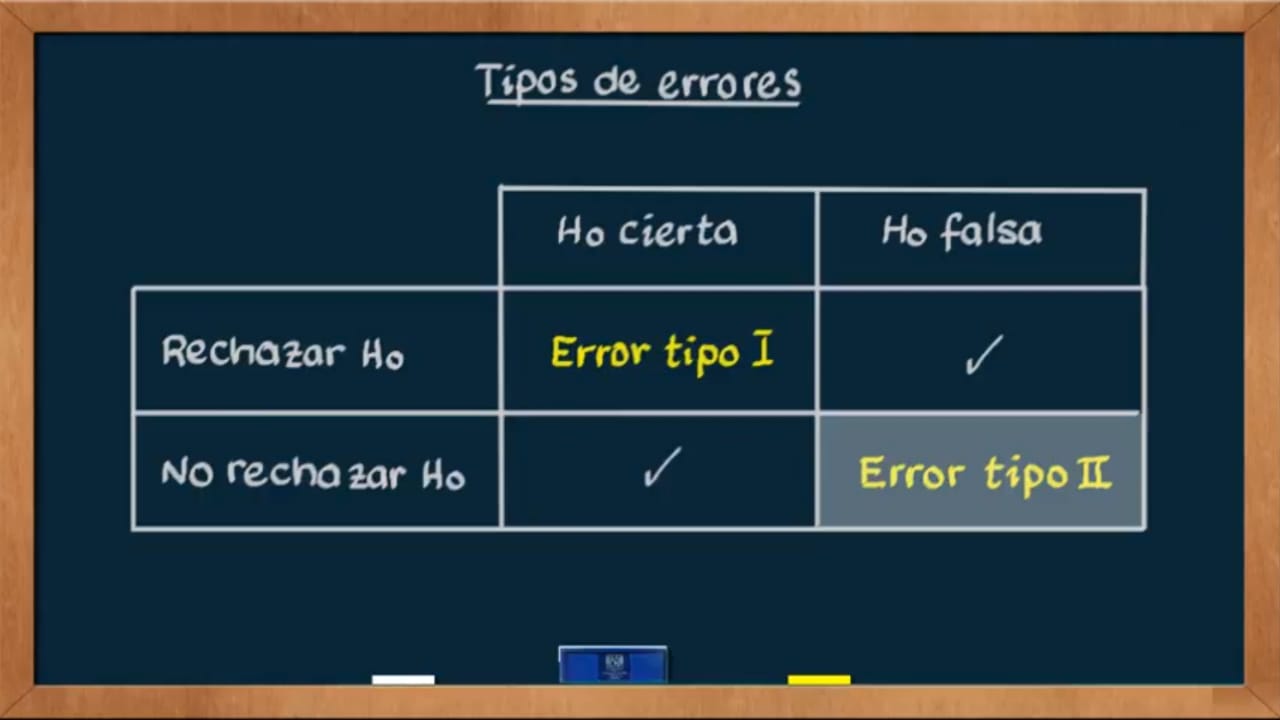
**Prueba de una o dos colas**

Una prueba de cualquier hipótesis estadística donde la alternativa es unilateral, como   
H0: θ = θ0, H0: θ = θ0,  
H1: θ > θ0, o quizás H1: θ < θ0, se denomina prueba de una sola cola.   
Anteriormente en esta sección se hizo referencia al estadístico de prueba para una hipótesis. Por lo general la región crítica para la hipótesis alternativa θ > θ0 yace en la cola derecha de la distribución del estadístico de prueba, en tanto que la región crítica para la hipótesis alternativa θ < θ0 yace por completo en la cola izquierda. (En cierto sentido el símbolo de desigualdad señala la dirección en donde se encuentra la región crítica).   
La prueba de cualquier hipótesis alternativa donde la alternativa es bilateral, como H0: θ = θ0, H1: θ ≠ θ0, se denomina prueba de dos colas, ya que la región crítica se divide en dos partes, a menudo con probabilidades iguales en cada cola de la distribución del estadístico de prueba. La hipótesis alternativa θ ≠ θ0 establece que θ < θ0 o que θ > θ0.

**Defina error de tipo I y error de tipo II.**

**Error de tipo I y II**

Al probar cualquier hipótesis estadística, hay cuatro situaciones posibles que determinan si nuestra decisión es correcta o errónea. Estas cuatro situaciones se re sumen en



El procedimiento de toma de decisiones recién descrito podría conducir a cualquiera de dos conclusiones erróneas.

El rechazo de la hipótesis nula cuando es verdadera se denomina **error tipo I.**

No rechazar la hipótesis nula cuando es falsa se denomina **error tipo II.**

Probabilidad del **error tipo I (α)**

La probabilidad de cometer un error tipo I, también llamada nivel de significancia, se denota con la letra griega α.

Con un nivel de significancia pequeño tenemos una región crítica de tamaño pequeña y, por lo tanto, es poco probable que se cometa un error de tipo I.

Esto indica que sería poco probable cometer el error de tipo I.

Probabilidad de **error tipo II (β)**   
La probabilidad de cometer un error tipo II, que se denota con β, es imposible de calcular a menos que tengamos **una hipótesis alternativa específica.**

Por ejemplo: Si X≤8 No rechazamos cuando n=20

Si probamos la hipótesis nula H0: p = 1/4 contra la hipótesis alternativa H1: p = 1/2, entonces podremos calcular la probabilidad de no rechazar H0 cuando es falsa.

Para calcular el error tipo 1 de, Calculo la probabilidad de P(X>8) con p=1/4

Luego, para calcular el error tipo II calculo la probabilidad de que P(X≤8) con p=1/2

Cabe destacar que la probabilidad de cometer un error tipo II aumenta rápidamente cuando el valor verdadero de μ se aproxima al valor hipotético

**Mencione propiedades de α y de β.**

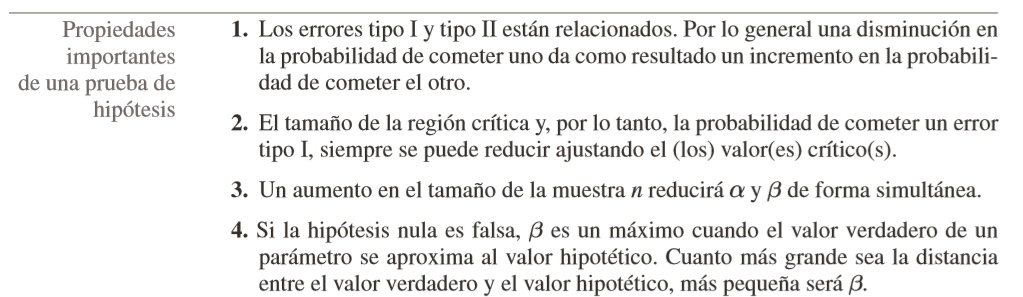
Para un **tamaño muestral fijo,** una disminución en la probabilidad de un error por lo general tendrá como resultado un incremento en la probabilidad del otro error.

En algunos casos la reducción de α no es suficiente por sí misma para garantizar un buen procedimiento de prueba. También debemos evaluar β para varias hipótesis alternativas.

De manera ideal, es preferible utilizar un procedimiento de prueba con el cual haya pocas probabilidades de cometer el error tipo I y el error tipo II.  
Por fortuna, la probabilidad de cometer **ambos tipos de errores se puede reducir aumentando el tamaño de la muestra.**

Evidentemente, los errores tipo I y tipo II rara vez ocurren si el experimento consta de 100 individuos.

También debemos tener en cuenta que la probabilidad de cometer un **error tipo II aumenta rápidamente cuando el valor verdadero de μ se aproxima al valor hipotético**, pero no es igual a éste. Desde luego, **ésta suele ser la situación en la que no nos importa cometer un error tipo II.  
¿A qué se llama potencia de una prueba?**La potencia de una prueba es la **probabilidad de rechazar H0** **dado que una alternativa específica es verdadera**.



La potencia de una prueba es una medida muy descriptiva y concisa de la sensibilidad de una prueba estadística, donde se entiende por sensibilidad a la capacidad de una prueba para “detectar diferencias”

La potencia de una prueba se puede calcular como 1 – β.

A menudo diferentes tipos de pruebas se comparan **contrastando propiedades de potencia**.

La potencia es una medida más sucinta de cuán sensible es la prueba para detectar diferencias

Sucinta: Que está expresado de manera breve, concisa y precisa.

Se interpreta como que la prueba como se describe rechazará de forma adecuada H0 sólo 1 – β % de las veces.

**¿Qué es el valor P?**

**Un valor P es el nivel de significancia más bajo en el que el valor observado de la estadística de prueba es significativo.  
Un valor P es el nivel (de significancia) más bajo en el que el valor observado del estadístico de prueba es significativo.  
Uso de valores P para la toma de decisiones en pruebas de hipótesis**El método está diseñado para dar al usuario una alternativa (en términos de una probabilidad) a la mera conclusión de “rechazo” o “no rechazo”.   
El cálculo del valor P también proporciona al usuario información importante cuando el valor z cae dentro de la **región crítica ordinaria.**  
Por generaciones enteras de análisis estadístico se ha vuelto costumbre elegir una **α de 0.05 o 0.01** y seleccionar la región crítica de acuerdo con esto.   
Entonces, desde luego, **el rechazo o no rechazo estrictos de H0 dependerá de esa región crítica.  
En la estadística aplicada los usuarios han adoptado de forma extensa el método del valor P.  
Por ejemplo: En el caso de H0: μ = 10, contra H1: μ ≠ 10, se observa un valor z = 1.87.** En términos estrictos, con α = 0.05 el valor no es significativo; pero el riesgo de cometer un error tipo I si se rechaza H0 en este caso difícilmente se podría considerar grave. De hecho, en una situación de dos colas, el riesgo se cuantifica como

P = 2P(Z > 1.87 cuando μ = 10) = 2(0.0307) = 0.0614.  
Como resultado, 0.0614 es la probabilidad de obtener un valor de z tan grande o mayor (en magnitud) que 1.87 cuando, de hecho, μ = 10.

Aunque esta evidencia en contra de H0 no es tan firme como la que resultaría de un rechazo a un nivel α = 0.05, se trata de información importante para el usuario.

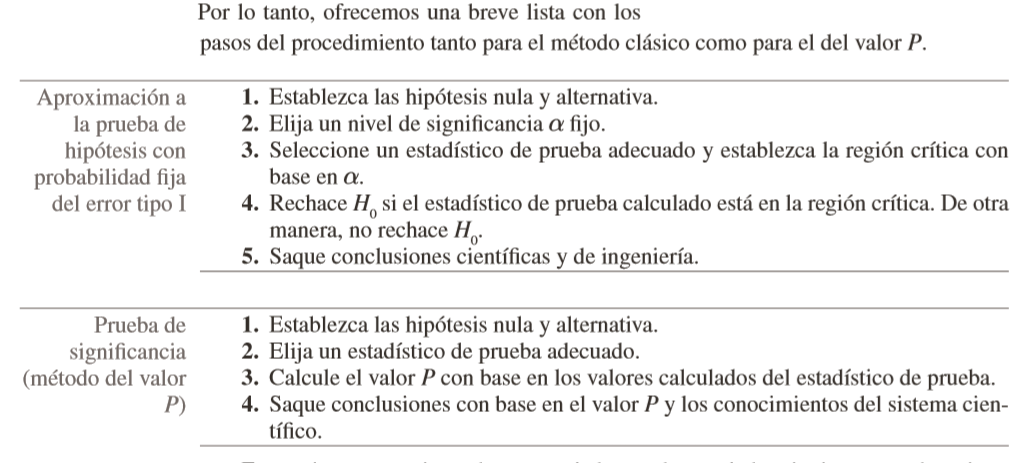
**Otro ejemplo:** El cálculo del valor P también proporciona al usuario información importante cuando el valor z cae dentro de la región crítica ordinaria. Por ejemplo, si z es 2.73, resulta informativo para el usuario observar que P = 2(0.0032) = 0.0064. Y por consiguiente, el valor z es **significativo a un nivel considerablemente menor que 0.05.**

**¿En qué difiere el uso de los valores P de la prueba de hipótesis clásica?**

En el método del valor P no se determina una α fija y las conclusiones se obtienen con base en el tamaño del valor P, **según la apreciación subjetiva** del ingeniero o del científico.  
  
**¿Cómo se usa el valor P para la toma de decisiones?**

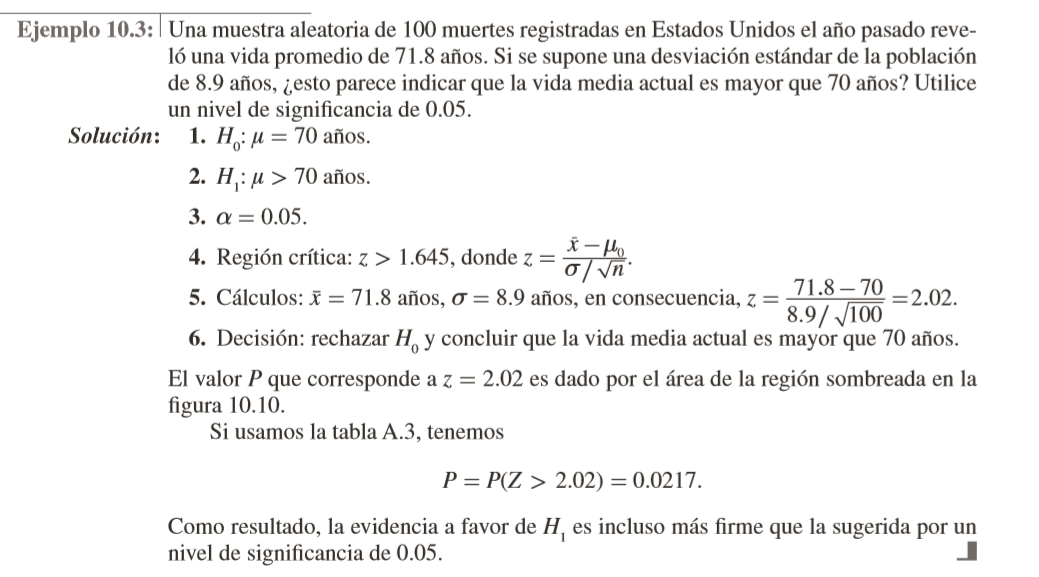
La utilidad del valor P yace en analizar la probabilidad dado un valor aleatorio observado de la muestra. Es el nivel (de significancia) más bajo en el que el valor observado del estadístico de prueba es significativo.

**Recuerde los pasos para plantear y resolver un problema de pruebas de hipótesis.**



**Ejemplo**

**Identifique las condiciones que se deben dar para el uso de cada estadístico de prueba.**



**Relación de pruebas de hipótesis con la estimación del intervalo de confianza**

La estimación del intervalo de confianza incluye el cálculo de límites dentro de los cuales es “razonable” que resida el parámetro en cuestión.

Para el caso de **una sola media** de la población μ con σˆ2 conocida, la estructura tanto de la prueba de hipótesis como de la estimación del intervalo de confianza se basa en la **variable aleatoria**

**La prueba de H0: μ = μ0** contra **H1: μ ≠ μ0** a un **nivel de significancia α es equivalente** **a calcular un intervalo de confianza del 100(1 – α)% sobre μ** y **rechazar H0, si μ0 está fuera del intervalo de confianza.**

* Si μ0 está dentro del intervalo de confianza, no se rechaza la hipótesis.

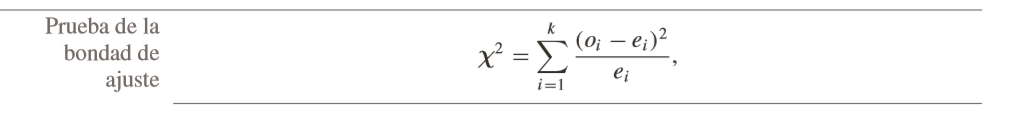
La equivalencia de la estimación del intervalo de confianza con la prueba de hipótesis se extiende a las diferencias entre dos medias, varianzas, cocientes de varianzas, etcétera.

El estudiante de estadística **no debería considerar la estimación del intervalo de confianza y la prueba de hipótesis como formas separadas de inferencia estadística**.

**Prueba de la bondad de ajuste**

A lo largo de este capítulo nos ocupamos de la prueba de hipótesis estadística acerca de parámetros de una sola población, como μ, σ2 y p.   
Ahora **consideraremos una prueba para determinar si una población tiene una distribución teórica específica**.

Básicamente esta prueba nos afirma con cierto grado de confianza si esta podemos describir cierta probabilidad con cierta distribución.  
La prueba se basa en el **nivel de ajuste** que existe **entre la frecuencia de ocurrencia** de las observaciones en una muestra observada **y las frecuencias esperadas** que se obtienen a partir de la distribución hipotética.   
Una **prueba de la bondad** de ajuste entre las frecuencias observadas y esperadas se basa en la cantidad.



donde χ 2 es un valor de una variable aleatoria cuya distribución muestral se **aproxima** muy de cerca a la distribución **chi cuadrada con v = k – 1-r grados de libertad.**

Los símbolos oi y ei representan las frecuencias observada y esperada, respectivamente, para la i-ésima celda.

Si las frecuencias observadas se acercan a las frecuencias esperadas correspondientes, el valor χ 2 **será pequeño**, lo cual indica un **buen ajuste.**

Si las frecuencias observadas difieren de manera considerable de las frecuencias esperadas, el valor χ 2 **será grande** y el **ajuste deficiente**.

Un **buen ajuste** conduce a la **aceptación de H0**, mientras que un ajuste deficiente conduce a su rechazo. Por lo tanto, la región crítica caerá en la cola derecha de la distribución chi cuadrada.

El criterio de decisión que aquí se describe **no se debería utilizar** **a menos que cada una de las frecuencias esperadas sea por lo menos igual a 5**

**Esta restricción podría requerir la combinación de celdas adyacentes, lo que dará como resultado una reducción en el número de grados de libertad.**

La **prueba de bondad** de ajuste chi cuadrada es un recurso **importante**, en particular debido a que **muchos procedimientos estadísticos en la práctica dependen, en un sentido teórico, de la suposición de que los datos reunidos provienen de un tipo de distribución específico.**

**Prueba de independencia**

El procedimiento de prueba de chi cuadrada que se presentó en la sección anterior también se puede usar para probar la **hipótesis de independencia** de dos variables de clasificación.

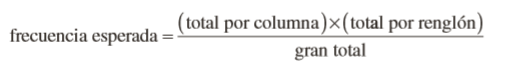
Una tabla de contingencia con r renglones y c columnas se denomina tabla r × c (“r × c” se lee “r por c”).

Los totales de renglones y columnas en la tabla 10.6 se denominan **frecuencias marginales**.

La **prueba de independencia** se basa en tomar una decisión de aceptar o rechazar la hipótesis nula, H0, teniendo en cuenta qué tan bien se **ajustan** las frecuencias observadas a las frecuencias esperadas en el caso de ser independientes.

Las **frecuencias esperadas** se obtienen multiplicando la **probabilidad de cada celda** por el número total de observaciones o también llamado **Gran total**.

Aunque, por regla general se utiliza:

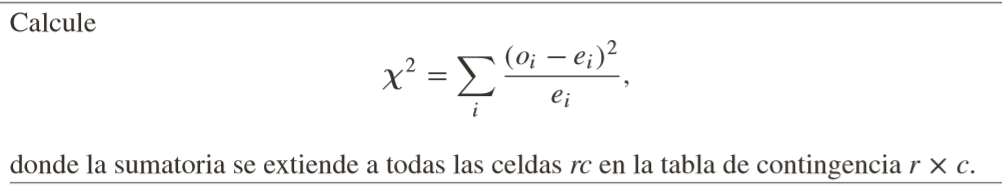


Para registrar dichas frecuencias se puede realizar un nuevo cuadro (Más prolijo) o escribirlas entre paréntesis al lado de la frecuencia observada

Para esta prueba se utiliza el estadístico Ji-Cuadrado.   
El número de grados de libertad asociados con la prueba chi cuadrada que aquí se usa es igual al número de frecuencias de celdas que se pueden llenar libremente cuando se nos proporcionan los totales marginales y el gran total: Como formula práctica se utiliza



Luego, relacionamos el valor observado con el estadístico de prueba:

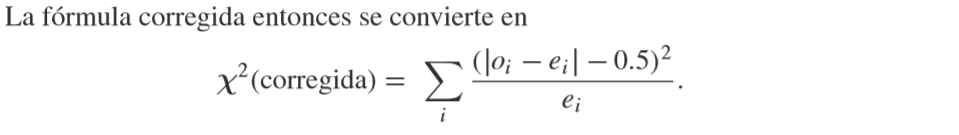


Si > con v = (r – 1)(c – 1) grados de libertad, **rechace la hipótesis nula** de independencia al nivel de significancia α; en otro caso no la rechace.

Es importante recordar que el estadístico sobre el cual basamos nuestra decisión tiene una distribución que sólo se **aproxima** por la distribución chi cuadrada.

Hay que tener en cuenta que **la distribución chi** cuadrada continua **parece aproximarse muy bien a la distribución de muestreo** discreta de χ 2, siempre y **cuando el número de grados de libertad sea mayor que 1.**

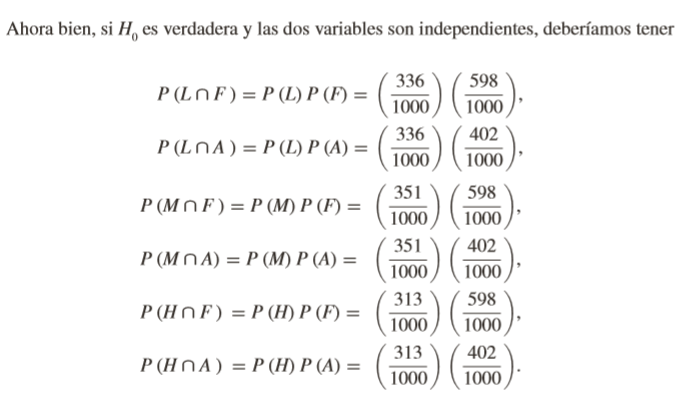
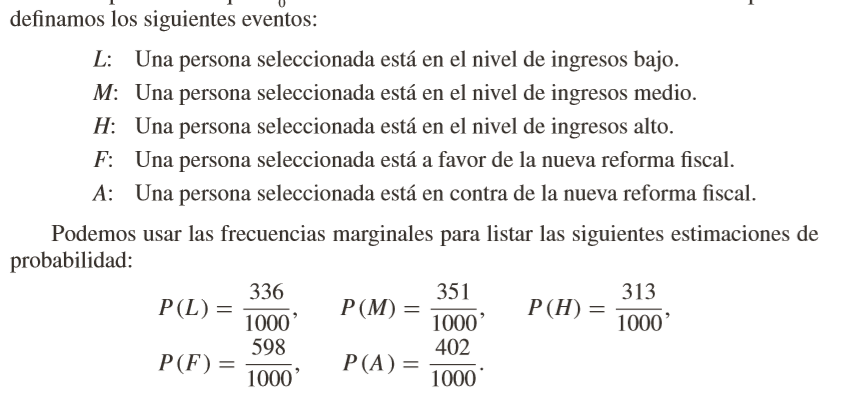
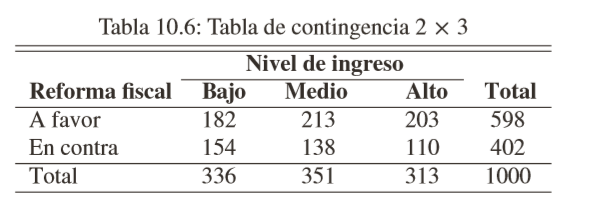
En una tabla de contingencia de 2 × 2, **donde sólo tenemos 1 grado de libertad**, se aplica una corrección llamada **corrección de Yates** para continuidad**.**



Si las frecuencias de las celdas esperadas son grandes, vamos a observar que los resultados corregidos y sin corrección son casi iguales. Por lo tanto, cuando las frecuencias esperadas están **entre 5 y 10**, se debe aplicar la **corrección de Yates**.   
Para **frecuencias esperadas menores que 5** se debería utilizar la **prueba exacta de Fisher-Irwin.**

Todos estos inconvenientes se pueden evitar tomando muestras más grandes, pero no siempre es posible por distintas circunstancias.

Ejemplo:



Las frecuencias esperadas se obtienen multiplicando la probabilidad (x= n p) de cada celda por el número total de observaciones.

