

La estadística de los ensayos clínicos

Russell Bowater

Los ensayos clínicos se llevan a cabo para permitir la recolección de los datos acerca de la seguridad y la eficacia de nuevos tipos de intervenciones médicas tales como nuevos medicamentos, dispositivos médicos y técnicas quirúrgicas. Este taller se centrará en los métodos estadísticos que son relevantes para el análisis de este tipo de datos. Los temas que se tratan son: los tipos de ensayos clínicos, la calidad de la evidencia, las medidas estadísticas de eficacia, las propiedades estadísticas de estas medidas y las pruebas de hipótesis. La segunda parte del taller examinará cómo la precisión de una medida de eficacia se puede mejorar mediante la combinación de datos de varios ensayos clínicos. Los temas que se tratan son: la ciencia de meta-análisis y los modelos de efectos fijos y aleatorios para la formación de los estimadores globales.

Programa del taller:

Sesión 1: Los tipos de ensayos clínicos, la aleatorización, el enmascaramiento, los puntos finales binarios y continuos, el riesgo relativo y el odds ratio.

Sesión 2: La razón de hazards, el modelo de riesgos proporcionales de Cox, los intervalos de confianza y las pruebas de hipótesis tal como la prueba de log-rank.

Sesión 3: El peligro de la agregación simple, la selección de una medida robusta de eficacia, el método de la varianza inversa y el método de Mantel-Haenszel.

Sesión 4: El método de Peto, los modelos de efectos aleatorios y el software para el meta-análisis.



*5° Coloquio del
Departamento de
Matemáticas*

La estadística de los ensayos clínicos

Russell Bowater

Enero del 2012, Metepec, Atlixco, Puebla

**5^{to} Coloquio del Departamento
de Matemáticas**

La Estadística de los Ensayos Clínicos

Dr. Russell J. Bowater



Comité Organizador

Dr. Mario Pineda Ruelas

Dra. Blanca Rosa Pérez Salvador

Dr. Joaquín Delgado Fernández

Dr. Constancio Hernández García

Mat. Daniel Espinosa

Beatriz Arce Vargas (Apoyo logístico)

La Estadística de los Ensayos Clínicos

Dr. Russell J. Bowater

Departamento de Matemáticas, UAM-I



Universidad Autónoma Metropolitana

Contenido

Capítulo 1. Los ensayos clínicos	1
1.1. Las fases de los ensayos clínicos	1
1.2. Control y asignación al azar	1
1.3. Los ensayos controlados aleatorios	2
1.4. El enmascaramiento	2
1.5. Los puntos finales	2
1.6. El riesgo relativo	3
1.7. Un estimador de riesgo relativo	3
1.8. Un ejemplo simple de un riesgo relativo	3
1.9. Comentarios acerca de los riesgos relativos	3
1.10. El odds ratio (la razón de odds)	4
1.11. Comentarios acerca de la “razón de odds”	4
1.12. La Reducción del Riesgo Absoluto (RRA)	4
1.13. El número necesario para tratar (NNT)	4
1.14. Hazard Ratio (La razón de hazards)	5
1.15. La relación entre el hazard ratio y el riesgo relativo	5
1.16. El modelo de riesgos proporcionales de Cox	5
1.17. Las distribuciones muestrales de RR y OR	5
1.18. El uso de $\log(\text{RR})$ y $\log(\text{OR})$	6
1.19. Error estándar de $\log(\text{RR})$	6
1.20. Error estándar de $\log(\text{OR})$	6
1.21. Prueba de la hipótesis de que $\text{RR}=1$ o $\text{OR}=1$	6
1.22. Prueba de diferencias en los tiempos reales de supervivencia	7
1.23. La prueba de Log Rank	7
Capítulo 2. Meta-análisis	9
2.1. Introducción	9
2.2. El meta-análisis y el promediar	9
2.3. Las revisiones sistemáticas	9
2.4. Meta-análisis en el contexto de las revisiones sistemáticas	10
2.5. Promediando sobre una variable robusta	10
2.6. El método de la varianza inversa (MVI)	10
2.7. La aplicación de MVI a los datos binarios	11

2.8.	El método de Mantel-Haenszel (MH)	11
2.9.	El método de Peto para el odds ratio (1985)	11
2.10.	El método de Peto para hazard ratios	12
2.11.	Modelos de efectos fijos y aleatorios	13
2.12.	Un modelos de efectos aleatorios (DerSimonian and Laird 1986)	13
2.13.	La estimación de la varianza entre los estudios	13
2.14.	El estimador DSL del efecto de tratamiento	14
2.15.	El método DSL en comparación con el método MVI	14
2.16.	Efectos fijos versus efectos aleatorios en meta-análisis	14

Los ensayos clínicos

1.1. Las fases de los ensayos clínicos

Los ensayos clínicos se llevan a cabo para permitir la recolección de datos sobre la seguridad y la eficacia de nuevos medicamentos o dispositivos médicos. Un nuevo fármaco tiene que pasar por las distintas etapas (o fases) de las pruebas. El comportamiento del medicamento debe ser satisfactorio en cada fase de los ensayos para poder llegar al final del proceso de prueba.

Pruebas de fase 1: Participan un número pequeño (20-80) de voluntarios sanos. El propósito de la prueba consiste en comprobar la seguridad del medicamento.

Pruebas de fase 2 y 3: El propósito de estos ensayos es comprobar la eficacia del medicamento, es decir, si funciona o no. El número de pacientes es más alto en la fase 3 (300-3000 pacientes) que la fase 2 (20-300 pacientes).

1.2. Control y asignación al azar

Ensayos clínicos no controlados: Todos los pacientes incluidos en el ensayo reciben el tratamiento en estudio. Es difícil evaluar los beneficios del tratamiento en relación a las alternativas.

Ensayos controlados no aleatorios: Los pacientes se dividen en aquellos que reciben el tratamiento en estudio (el grupo de tratamiento) y los que reciben un tratamiento de control (el grupo de control). Sin embargo, esta división no es al azar.

Problema con los ensayos controlados no aleatorios: Los médicos pueden recomendar que sólo el peor de los pacientes reciba el tratamiento en estudio en vez del tratamiento de control.

Los ensayos controlados aleatorios: Estos son el estándar de oro para la determinación objetiva de la eficacia del tratamiento. La división de los pacientes entre los grupos de tratamiento y control se realiza de forma aleatoria.

1.3. Los ensayos controlados aleatorios

- (1) Inscribir a los pacientes en la prueba que satisfagan los criterios previamente establecidos, por ejemplo, pacientes que tienen una enfermedad determinada.
- (2) Aleatorizar los pacientes a 2 grupos.
- (3) Grupo 1 recibe el tratamiento en estudio. Grupo 2 recibe el tratamiento de control.
- (4) Seguir, observar y evaluar los pacientes en ambos grupos en la misma manera.

Muy fácil! Soluciona el problema de la causalidad y los factores de confusión.

1.4. El enmascaramiento

Ensayos abiertos: Los pacientes y los médicos saben cual es el tratamiento que ha sido administrado. Muchas veces esto es inevitable. Por ejemplo, es difícil de evitar esta situación en los ensayos quirúrgicos.

Un ensayo ciego simple: Los médicos saben cual es el tratamiento que ha sido administrado pero los pacientes no lo saben. Muchas veces el enmascaramiento de pacientes es más fácil que el enmascaramiento de médicos.

El problema con el ensayo ciego simple: Los médicos que no están enmascarados pueden comportarse de forma diferente según el tipo de tratamiento que los pacientes reciben.

Un ensayo doble ciego: Tanto los pacientes como los médicos no saben cual es el tratamiento que ha sido administrado. Siempre que sea posible este diseño debería ser siempre aplicado. Puede exigir el uso de un placebo.

1.5. Los puntos finales

Los puntos finales binarios (basados en eventos): Muy a menudo los puntos finales de los ensayos clínicos se basan en eventos tales como la supervivencia global, la supervivencia libre de enfermedad y la enfermedad que requiere hospitalización. Los eventos de estos tres ejemplos son la muerte, el regreso de la enfermedad y la hospitalización.

Los puntos finales continuos (no basados en eventos): Por ejemplo, el nivel de colesterol en la sangre. A menudo este tipo de puntos finales son los puntos finales sustitutos del punto final de interés principal, por ejemplo, el colesterol puede causar los problemas del corazón que conducen a la muerte. Medidas de calidad de la vida son medidas

continuas que se hacen cada vez más populares. Sin embargo, es difícil medir la calidad de la vida (es un punto final “suave”).

1.6. El riesgo relativo

Para un punto final basado en eventos, a menudo estamos interesados en el riesgo (o la probabilidad) de que el evento ocurra en un periodo de tiempo determinado, por ejemplo el riesgo de muerte dentro de un período de 2 años. El riesgo relativo de que un evento ocurra en un período de tiempo x es el siguiente:

$$RR = \frac{p_t}{p_c} \quad (1.1)$$

donde,

p_t = La probabilidad en el grupo de tratamiento que el evento ocurra

p_c = La probabilidad en el grupo de control que el evento ocurra

1.7. Un estimador de riesgo relativo

Generalmente, el riesgo relativo se estima como:

$$\widehat{RR} = \frac{n_t/N_t}{n_c/N_c} \quad (1.2)$$

donde,

n_t = Número de eventos en el grupo de tratamiento en tiempo x

n_c = Número de eventos en el grupo de control en tiempo x

N_t = Número de pacientes en el grupo de tratamiento

N_c = Número de pacientes en el grupo de control

1.8. Un ejemplo simple de un riesgo relativo

De los 100 pacientes en el grupo de tratamiento, 6 mueren dentro de 2 años y de los 80 pacientes en el grupo de control, también 6 mueren dentro de 2 años.

$$\widehat{RR} = \frac{6/100}{6/80} = 0.8 \quad (1.3)$$

1.9. Comentarios acerca de los riesgos relativos

Un RR que es menor de 1 obviamente favorece el tratamiento, mientras un RR que es más de 1 favorece el control. Los problemas se producen si no hay ningún eventos. Los riesgos relativos son populares entre los médicos porque son fáciles de entender. Por otra parte, los matemáticos generalmente no les gustan los riesgos relativos porque son problemáticos en la construcción de modelos estadísticos.

1.10. El odds ratio (la razón de odds)

Dado que el desarrollo de la teoría matemática sobre la base del riesgo relativo puede ser problemático, el odds ratio fue inventado. El odds ratio de que un evento ocurra en un período de tiempo x es:

$$OR = \frac{p_t(1 - p_t)}{p_c(1 - p_c)} \quad (1.4)$$

Si $p_t = 0.06$ y $p_c = 0.075$ (como en el ejemplo anterior) entonces $OR = 0.813$. En general RR y OR son muy similares cuando p_t y p_c son pequeños.

1.11. Comentarios acerca de la “razón de odds”

Ventaja: El logaritmo de la razón de odds forma la base de regresión logística y es generalmente útil en la construcción de modelos estadísticos.

Desventaja: La mayoría de los clínicos interpretan los odds ratios como si fueran los riesgos relativos!

1.12. La Reducción del Riesgo Absoluto (RRA)

La reducción del riesgo absoluto (RRA) es simplemente la diferencia de riesgos entre los grupos de tratamiento y control.

$$RRA = p_c - p_t \quad (1.5)$$

RRA es cada vez más popular como una medida de la eficacia debido a que está estrechamente relacionada con la medición del costo-efectividad de un tratamiento. En otras palabras es lo que hay que pagar para obtener un determinado incremento de beneficio para la salud.

1.13. El número necesario para tratar (NNT)

La inversa multiplicativa de la reducción del riesgo absoluto se conoce como el número necesario para tratar (NNT).

$$NNT = 1/RRA \quad (1.6)$$

Por ejemplo, si el RRA de la muerte en tiempo x es 0.02, entonces se podría esperar que 50 pacientes deben ser tratados para evitar una muerte más en el periodo de tiempo x .

1.14. Hazard Ratio (La razón de hazards)

El hazard rate (la fuerza de mortalidad) es

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{eventos observados}[t, t + \Delta t] / N(t)}{\Delta t} \quad (1.7)$$

donde $N(t)$ es el número (de personas) en riesgo al tiempo t .

Bajo un modelo de riesgos proporcionales (por ejemplo el modelo de Cox) la razón de hazard rates entre los grupos de tratamiento y control es una constante que no depende del tiempo t . Esta razón se conoce como el hazard ratio.

1.15. La relación entre el hazard ratio y el riesgo relativo

Bajo un modelo de riesgos proporcionales, se supone que el riesgo relativo es el mismo sin importar el período de tiempo que se está estudiando y este riesgo relativo es igual al hazard ratio. Por este motivo, el hazard ratio puede ser considerado como un riesgo relativo que es independiente de tiempo.

1.16. El modelo de riesgos proporcionales de Cox

En la investigación médica, este modelo es el modelo más utilizado para obtener un estimador del hazard ratio (y el intervalo de confianza). Cuando hay sólo una covariable dicotómica (que es el caso estándar con respecto a ensayos clínicos), el modelo de riesgos proporcionales de Cox se puede expresar como el siguiente:

$$h(t) = h_0(t)e^{bX} \quad (1.8)$$

donde $h(t)$ y $h_0(t)$ son las fuerzas de mortalidades (hazard functions) para los grupos de tratamiento y control respectivamente. Aquí $X = 1$ si el paciente está en el grupo de tratamiento y $X = 0$ si el paciente está en el grupo de control. El parámetro b , que es de hecho el logaritmo del odds ratio, puede ser estimado por el método de verosimilitud parcial. El modelo de Cox es popular porque no hay necesidad de hacer ningún supuesto sobre la forma de la fuerza de mortalidad de base $h_0(t)$.

1.17. Las distribuciones muestrales de RR y OR

Ambos $\log(\text{RR})$ y $\log(\text{OR})$ tienen distribuciones muestrales que son aproximadamente gaussiana (normal). De hecho, estas dos distribuciones son asintóticamente normal. La aproximación normal por lo general es bueno con la condición de que el número de los eventos en ambos grupos de tratamiento y de control no son demasiado pequeñas.

1.18. El uso de $\log(\text{RR})$ y $\log(\text{OR})$

La transformación de RR y OR a la escala logarítmica también tiene sentido desde un punto de vista intuitivo. Por ejemplo, un RR de 0,5 es como el reflejo en un espejo de un RR de 2 pero estos dos valores no son equidistantes de 1. Sin embargo, $\log(0.5) = -0.693$ y $\log(2) = 0.693$ que son obviamente, equidistante de $\log(1) = 0$. Los estimadores de RR y OR por lo general se trazan sobre la escala logarítmica.

1.19. Error estándar de $\log(\text{RR})$

Utilizando el método delta, se puede demostrar que la varianza asintótica de la distribución muestral de $\log(\text{RR})$ se puede expresar como:

$$\text{Var}(\log \widehat{\text{RR}}) = \frac{1 - \hat{p}_t}{N_t \hat{p}_t} + \frac{1 - \hat{p}_c}{N_c \hat{p}_c} \quad (1.9)$$

Habida cuenta de la aproximación normal, un intervalo de confianza del 95% para el verdadero $\log(\text{RR})$ está dada por:

$$\log(\widehat{\text{RR}}) \pm 1,96 \times \sqrt{\text{Var}(\log \widehat{\text{RR}})} \quad (1.10)$$

Está claro que un intervalo de confianza del 95% para el RR puede obtenerse tomando el anti-logaritmo de estos dos límites.

1.20. Error estándar de $\log(\text{OR})$

No existe un método estándar para estimar el error estándar de $\log(\text{OR})$. El más popular método da como resultado el logit based fórmula de Woolf (que se basa en el método delta aunque no es una aplicación directa de este método). Esta fórmula es:

$$\text{Var}(\log \widehat{\text{OR}}) = \frac{1}{N_t \hat{p}_t} + \frac{1}{N_t(1 - \hat{p}_t)} + \frac{1}{N_c \hat{p}_c} + \frac{1}{N_c(1 - \hat{p}_c)} \quad (1.11)$$

Dado que se puede expresar $\text{Var}(\log \widehat{\text{RR}})$ como:

$$\text{Var}(\log \widehat{\text{RR}}) = \frac{1}{N_t \hat{p}_t} + \frac{1}{N_t} + \frac{1}{N_c \hat{p}_c} + \frac{1}{N_c \hat{p}_c} \quad (1.12)$$

es claro que $\text{Var}(\log \widehat{\text{OR}}) > \text{Var}(\log \widehat{\text{RR}})$ con estos métodos.

1.21. Prueba de la hipótesis de que $\text{RR}=1$ o $\text{OR}=1$

Para probar la hipótesis nula de que $\log(\text{RR}) = 0$ o $\log(\text{OR}) = 0$, los errores estándares que se acaban de describir pueden ser usados bajo el supuesto de normalidad. La prueba resultante no es equivalente a la prueba exacta de Fisher para una tabla de 2×2 (que se considera la prueba más favorable que se puede usar). Sin embargo, a menos de que las muestras sean muy pequeñas, el valor de p producido por

esta prueba, por lo general, se aproxima al valor de p de la prueba exacta (de Fisher) lo suficientemente bien para determinar que alguna diferencia no es importante desde un punto de vista práctico. Nótese que la prueba habitual de la ji-cuadrado para una tabla de 2×2 también intenta aproximar la prueba exacta de Fisher.

1.22. Prueba de diferencias en los tiempos reales de supervivencia

Si los tiempos reales de supervivencia están disponibles, entonces no es usual que una prueba de la diferencia entre los grupos de tratamiento y control tenga una base en la hipótesis nula de que el Hazard Ratio=1. Esto se debe a que esta prueba se basaría en el supuesto de riesgos proporcionales (que se describió anteriormente), el cual no es necesario de hacer si el objetivo es simplemente probar las diferencias globales en los tiempos de supervivencia entre los dos grupos.

En consecuencia los errores estándares para el hazard ratio no suelen ser convertidos en una prueba de las diferencias en la supervivencia global. La prueba estándar a utilizar es de hecho la prueba de Log Rank. Esta prueba compara los hazard rates de los grupos de tratamiento y control en pequeños intervalos a lo largo de todo el eje del tiempo.

1.23. La prueba de Log Rank

Sean $k = 1, \dots, m$ los tiempos observados de muerte en un grupo u otro. Sean N_{tk} y N_{ck} los pacientes que siguen vivos al tiempo k para los grupos de tratamiento y control respectivamente. Sean O_{tk} and O_{ck} el número de muertes que ocurren al tiempo k para los grupos de tratamiento y control respectivamente.

Bajo H_0 : ninguna diferencia entre los grupos, el valor esperado de O_{tk} es

$$E_k = O_k \frac{N_{tk}}{N_k} \quad (1.13)$$

donde $O_k = O_{tk} + O_{ck}$ y $N_k = N_{tk} + N_{ck}$.

El estadístico de log rank (que compara cada O_{tk} con su esperanza E_k) es:

$$Z = \frac{\sum_k (O_{tk} - E_k)}{\sqrt{\sum_k V_k}} \quad (1.14)$$

donde V_k es la varianza de O_{Tk} .

El estadístico de log-rank Z es aproximadamente normal estándar.

Meta-análisis

2.1. Introducción

El meta-análisis es el método para formar estimadores combinando resultados de múltiples estudios científicos que versan sobre asuntos similares.

“Muchos de los grupos . . . son demasiado pequeños para permitir que una opinión definitiva se forme dada la magnitud del error probable que está involucrada” Karl Pearson 1904.

2.2. El meta-análisis y el promediar

Se podría decir que el meta-análisis es sólo un método para obtener una media. Sin embargo, es una forma inteligente de obtener una media porque hay que hacerlo de una manera que tenga en cuenta la variabilidad en el diseño y el contexto de los estudios. Esto se hace:

- (1) Promediando sobre la variable más adecuada.
- (2) Haciendo una buena decisión sobre la forma de los pesos de los estudios.

2.3. Las revisiones sistemáticas

Una revisión sistemática es una revisión de la literatura que se centra en una sola pregunta y se trata de identificar, evaluar, seleccionar y sintetizar todos los datos de investigación de alta calidad que son pertinentes a esa pregunta. Las revisiones sistemáticas se consideran generalmente como el más alto nivel de evidencia médica por profesionales de la salud. Todos los días, pacientes en todo el mundo reciben tratamientos sobre la base de la evidencia presentada en revisiones sistemáticas.

A pesar de esto, el ascenso de los revisiones sistemáticas a una posición de dominio es relativamente reciente. La Colaboración Cochrane (que establece las directrices mundiales para la realización de revisiones sistemáticas) apenas fue creada en 1993.

2.4. Meta-análisis en el contexto de las revisiones sistemáticas

Una buena revisión sistemática debe contener al menos un solo meta-análisis (aunque esto puede significar que sólo se evalúa el “potencial” para combinar las estimaciones de los diversos estudios). Mientras que la salida de una revisión sistemática es tanto descriptiva (una discusión de los estudios involucrados) como cuantitativa (el meta-análisis), nos limitaremos a concentrarnos en la parte cuantitativa.

En términos de un meta-análisis, una revisión sistemática puede ser considerado como un gran pedazo de equipo científico que imparcialmente recoge todos los estudios de buena calidad que sean de interés para una determinada cuestión científica. El propósito de un meta-análisis es establecer cómo se sintetiza (se combina) esta información de buena calidad (que también se espera sea imparcial).

2.5. Promediando sobre una variable robusta

Se puede observar que la agregación simple da más ponderación a los estudios con tasas altas de eventos. Por ejemplo, estudios basados en pacientes que están en un alto riesgo de morir (ya sea en el grupo de control o tratamiento). Lo que se requiere es una variable que sea razonablemente robusta a la heterogeneidad entre los estudios; por ejemplo, robusta a las variaciones en la tasa de mortalidad de base.

Para los datos de la tasa de eventos, una solución simple a este problema es que se elige esta variable como el riesgo relativo o el odds ratio. En algunos casos, la reducción absoluta del riesgo puede ser apropiada. Para los puntos finales continuos (no basados en eventos) la solución de este problema puede no ser tan fácil.

2.6. El método de la varianza inversa (MVI)

La manera matemática que es más natural para combinar las estimaciones que tienen precisiones diferentes es utilizar el método de la varianza inversa. En general, para los datos binarios y continuos, el estimador combinado que es producido por este método está dado por:

$$\hat{\theta} = \frac{\sum_i w_i \hat{\theta}_i}{\sum_i w_i} \quad (2.1)$$

donde, $\hat{\theta}_i$ es el estimador de θ para estudio i y,

$$w_i = \frac{1}{\text{Var}(\hat{\theta}_i)} \quad (2.2)$$

Es claro que la varianza de $\hat{\theta}$ es $1 / \sum_i w_i$

2.7. La aplicación de MVI a los datos binarios

Al combinar los riesgos relativos y los razones de odds (odds ratios), es natural que se toma los logaritmos para llegar a la variable θ , es decir, $\theta = \log(RR)$ o $\log(OR)$. En este escenario, los pesos w_i son simplemente la inversa de $Var(\log \widehat{RR})$ o $Var(\log \widehat{OR})$, que fueron calculados anteriormente.

El método de la varianza inversa (MVI) muchas veces funciona bien para los datos continuos, pero para combinar los RR y OR no es un método popular. Una razón de esto es que cuando los estudios son pequeños y/o las tasas de eventos son bajas, los estimadores de $Var(\log \widehat{RR})$ y $Var(\log \widehat{OR})$ pueden ser muy poco fiables.

2.8. El método de Mantel-Haenszel (MH)

Este método sólo puede utilizarse para los datos binarios. Sin embargo, es el método estándar para la combinación de los RR y OR en un modelo de efectos fijos.

La fórmula para $\hat{\theta}$ es el mismo que para el método de MVI, salvo que:

- (1) Las estimaciones se combinan en la escala natural (no en la escala logarítmica) es decir, $\theta = RR$ o $\theta = OR$ (que es una razón para la mayor estabilidad del método).
- (2) Los pesos w_i para el caso $\theta = RR$ están dados por

$$w_i = \frac{\hat{p}_c N_c N_t}{N_c + N_t} \quad (2.3)$$

Los pesos w_i para el caso $\theta = OR$ están dados por

$$w_i = \frac{\hat{p}_c(1 - \hat{p}_c)N_c N_t}{N_c + N_t} \quad (2.4)$$

2.9. El método de Peto para el odds ratio (1985)

Para los odds ratios, una alternativa al método de Mantel-Haenszal (MH) es el método de Peto.

Tabla I. Observado

	Muertos	Vivos
Control	a	$N_c - a$
Tratamiento	$O = b$	$N_t - b$

Tabla 2. Esperado (con un efecto nulo)

	Muertos	Vivos
Control	$(a + b)N_c/N$...
Tratamiento	$E = (a + b)N_t/N$...

(donde $N = N_c + N_t$)

El estimador de Peto para el odds ratio es:

$$\widehat{OR} = \exp\left(\frac{O - E}{V}\right) \quad (2.5)$$

donde V es la varianza hipergeométrica de b

Para un meta-análisis, el estimador combinado de Peto para el odds ratio es:

$$\widehat{OR} = \exp\left(\frac{\sum_i (O_i - E_i)}{\sum_i V_i}\right) \quad (2.6)$$

Esta es una extensión muy natural del escenario en el que hay sólo un estudio.

Peto describe su método como libre de supuestos, pero probablemente se puede clasificar mejor como un método de efectos fijos (como es el caso para el método de Mantel-Haenszal). La aproximación de Peto ha demostrado ser poco confiable cuando OR es muy grande o pequeño, y cuando N_c difiere mucho de N_t . Sin embargo, el método funciona bien cuando las tasas de eventos (es decir, n_t y n_c o a y b) son muy bajas y por esta razón el uso de este método es popular en meta-análisis.

2.10. El método de Peto para hazard ratios

Siguiendo el mismo formato como el método de Peto para el odds ratio, un estimador del logaritmo del Hazard Ratio (HR) para un ensayo dado es:

$$\log(\widehat{HR}) = \frac{\sum_k (O_k - E_k)}{\sum_k V_k} \quad (2.7)$$

donde O_k es el número de las muertes que ocurren al tiempo k en el grupo de tratamiento. Nótese que $k = 1, \dots, m$ son los tiempos observados de muerte en ambos grupos combinados. Además, E_k y V_k son la esperanza y varianza de O_k bajo H_0 : no hay ninguna diferencia entre los grupos, las cuales fueron definidas anteriormente para la prueba de log-rank.

Para un meta-análisis, el estimador combinado de Peto para el hazard ratio es:

$$\log(\widehat{HR}) = \frac{\sum_i \sum_k (O_{ik} - E_{ik})}{\sum_i \sum_k V_{ik}} \quad (2.8)$$

Nota importante: También los hazard ratios pueden ser calculados mediante un modelo de riesgos proporcionales y luego pueden ser combinados en un meta-análisis utilizando el método de la varianza inversa que fue descrito anteriormente (en la escala logarítmica).

2.11. Modelos de efectos fijos y aleatorios

Todos los métodos descritos hasta ahora han sido de efectos fijos, entre ellos el método de la inversa varianza, el método de Mantel-Haenszel y el método de Peto.

Un modelo de efectos fijos para una medida dada de eficacia supone que cada estudio estima el mismo valor de esta medida. Sin embargo, este modelo no supone que otras variables se mantienen constantes en todos los estudios tales como las tasas de mortalidad de base.

2.12. Un modelos de efectos aleatorios (DerSimonian and Laird 1986)

Un modelo de efectos aleatorios supone que el efecto del tratamiento para un estudio determinado es elegido al azar de una población de los efectos del tratamiento.

El modelo de DerSimonian y Laird (DSL) está dado por:

$$\mu_i = \mu_* + \varepsilon_i \quad (2.9)$$

donde,

μ_i es el verdadero efecto (no el efecto estimado) para el estudio i ,

μ_* es el verdadero efecto global

y, $\varepsilon_i \sim N(0, \tau^2)$

La varianza τ^2 se conoce como la varianza entre los estudios.

Por lo tanto, para cualquier estudio i , el estimador de $\log(RR)$ está dado por

$$\log(\widehat{RR}) = \mu_* + \varepsilon_i + \lambda \quad (2.10)$$

donde λ se distribuye normalmente (Gaussian) con media 0 y varianza igual a la varianza muestral de $\log(\widehat{RR})$ definida anteriormente.

2.13. La estimación de la varianza entre los estudios

No es estándar que el método de máxima verosimilitud es aplicado a la estimación de τ^2 . Esto tiene un impacto tanto en la estimación puntual de τ^2 como en la manera en que el error estándar del estimador τ^2 es tratado. En el enfoque de DerSimonian y Laird (DSL), τ^2 se estima utilizando el método de momentos mientras que el error estándar del estimador para τ^2 no es considerado tan importante en el contexto de todos los demás errores muestrales que están involucrados.

2.14. El estimador DSL del efecto de tratamiento

El estimador DSL del efecto global θ_* es

$$\hat{\theta}_{DSL} = \frac{\sum_i w_i \hat{\theta}_i}{\sum_i w_i} \quad (2.11)$$

Donde los pesos w_i están dados por

$$w_i = \frac{1}{\text{Var}(\hat{\theta}_i) + \tau^2} \quad (2.12)$$

Se nota que $\text{Var}(\hat{\theta}_i)$ es la varianza de $\hat{\theta}_i$ dentro de los estudios que habría sido llamada $\text{Var}(\lambda)$ anteriormente. Para los casos $\theta = \log(RR)$ o $\theta = \log(OR)$, los estimadores de $\text{Var}(\hat{\theta}_i)$ son los que fueron definidos anteriormente. La varianza de $\hat{\theta}_{DSL}$ es simplemente $1 / \sum_i w_i$

2.15. El método DSL en comparación con el método MVI

Si $\tau^2 = 0$ (es decir, la varianza entre los estudios es cero), entonces el estimador DSL es equivalente al estimador MVI que fue definido anteriormente. Si $\tau^2 > 0$ entonces los pesos del estimador DSL serán más pequeños y más parecidos entre sí que los pesos del estimador MVI. Esto significa que:

- (1) Un meta-análisis DSL (de efectos aleatorios) será más conservador que un meta-análisis MVI (de efectos fijos), es decir, los intervalos de confianza serán más amplios ya que $\text{Var}(\hat{\theta}_{DSL})$ y $\text{Var}(\hat{\theta}_{MVI}) = 1 / \sum w_i$.
- (2) Un meta-análisis DSL dará más ponderación a los estudios más pequeños que un meta-análisis MVI (que algunos ven como una desventaja).

2.16. Efectos fijos versus efectos aleatorios en meta-análisis

En las principales revistas de medicina se encuentran los meta-análisis de efectos tanto fijos como aleatorios. Sin embargo, es muy raro que se encuentren ambos tipos de métodos en el mismo artículo. La elección entre los métodos se puede decidir tanto por la filosofía de los investigadores como por una razón relacionada con los datos que se analizan.

Un meta-análisis de efectos fijos no tiene en cuenta la heterogeneidad entre los estudios de la variable que se estima y por esta razón produce las estimaciones que son demasiado precisas. Por otra parte, se puede argumentar que la variable estimada en un meta-análisis de efectos aleatorios no es muy intuitiva ya que es la media de una distribución normal hipotética de los efectos en los estudios. La elección entre los métodos es una batalla entre estas desventajas.