

Evaluación de la Aplicabilidad del Meta-Análisis de Efectos Aleatorios en Ingeniería del Software

Amatriain H., Dieste O., Fernández O., García-Martínez R.

Grupo de Investigación en Sistemas de Información. Departamento de Desarrollo Productivo y Tecnológico. Universidad Nacional de Lanús. Argentina.

Grupo de Ingeniería de Software Experimental. Facultad de Informática. Universidad Politécnica de Madrid. España.

Escuela de Posgrado. Facultad de Informática. Universidad Nacional de La Plata. Argentina.
hamatriain@frba.utn.edu.ar, odieste@fi.upm.es, enriquefernandez@educ.ar,
rgarcia@unla.edu.ar

Resumen. La síntesis cuantitativa consiste en combinar los resultados de varios estudios experimentales con el objeto de generar nuevas piezas de conocimiento. Estas nuevas piezas de conocimientos serán más generales y fiables que los resultados obtenidos por los estudios individuales, ya que dichas piezas de conocimiento están sustentadas por una mayor cantidad de evidencia empírica. El objetivo del presente trabajo es determinar cuáles de los modelos de Meta-Análisis existentes conviene aplicar en el contexto experimental que hoy día presenta la Ingeniería de Software Experimental.

Palabras clave. Meta-Análisis. Modelo de efecto fijo. Modelo de efectos aleatorios. Ingeniería de Software Experimental.

1. Introducción

La Ingeniería en Software tiene el objetivo de guiar el proceso de construcción de productos software mediante la definición de métodos, técnicas y herramientas. Toda ingeniería debe estar basada en métodos científicos para ser considerada como tal. En la Ingeniería en Software generalmente no se cuenta con técnicas ni métodos que cuenten con una justificación científica ni un “estudio objetivo de su efectividad” [1]. Es necesario un marco que permita a los ingenieros poder conocer cuales son los mejores métodos y herramientas que se deben aplicar a través de un método científico y por lo tanto objetivo. Este marco en la Experimentación, como sucede en otras disciplinas, permite brindar información objetiva y no sólo opiniones sobre lo mejor a aplicar en cada etapa de un proyecto software según las circunstancias. Ahora bien, la experimentación dentro del campo de la Ingeniería de Software es costosa, por lo que se torna muy difícil, sino imposible, realizar experimentos a gran escala (experimentos con gran cantidad de sujetos experimentales) [2]. Esto hace que la única alternativa sea la de realizar experimentos reducidos, lo que acarrea el problema de la poca representatividad de los resultados a los que se arriba. Este problema podría paliarse mediante la agregación de resultados (también conocida

como síntesis cuantitativa) para poder obtener conclusiones avaladas con la mayor evidencia empírica posibles. La síntesis cuantitativa de experimentos [3], más conocida como Meta-Análisis [4], consiste en combinar los resultados de varios estudios experimentales con el objeto de generar nuevas piezas de conocimiento. Los Meta-Análisis son habituales en disciplinas con una fuerte componente empírica, tales como la medicina, la psicología o la física y fue introducido en Ingeniería del Software por [5]. Este trabajo despertó el interés de muchos investigadores por realizar este tipo de trabajos, por ejemplo podemos citar a: [6]-[8] entre otros. El hecho de poder combinar los experimentos permite mejorar en gran medida la calidad del conocimiento empírico que se genera en Ingeniería del Software, donde analizar los resultados de los experimentos de manera conjunta mejora la calidad de los conocimientos respecto de si se los analiza individualmente como se muestran en [2]. Existen dos estrategias bien diferenciadas para combinar los resultados de los experimentos: el modelo de efecto fijo (supone que la existencia de un único resultado poblacional, el cual se irá estabilizando a medida que se incorporen experimentos al meta-análisis) y el modelo de efectos aleatorios (supone que existe un conjunto de variables no controladas que influyen en los resultados de los experimentos provocando que los resultados cambien a medida que se incorporan experimentos de distintas vertientes). Dado que en Ingeniería del Software los estudios experimentales que se incorporan a un Meta-Análisis provienen de distintos momentos (año de desarrollo), de diferentes lugares (donde existen distintas componentes culturales) y muchas veces son desarrollados por personas de diferentes perfiles (por ejemplo profesionales o estudiantes), a priori el modelo de efectos aleatorios sería la mejor opción para combinar los resultados [9]. Ahora bien algunos autores [10] indican que este modelo posee un alto nivel de error cuando se agregan menos de diez experimentos, hecho habitual en Ingeniería del Software donde a excepción del trabajo de [7] que logró combinar 15 experimentos, el resto de los trabajos no llegó a combinar más de nueve estudios [6]. Este hecho convierte al modelo de efecto fijo como la única alternativa para agregar experimentos en Ingeniería del Software. Este hecho motiva la necesidad de desarrollar un trabajo que permita determinar, para el contexto de Ingeniería del Software, cual es el modelo de Meta-Análisis más apropiado. El presente trabajo pretende establecer los niveles de fiabilidad y potencia estadística de cada modelo, mediante una simulación de Monte-Carlo (como se hizo en [8]) en el contexto experimental que hoy día presenta la Ingeniería de Software, para determinar cuál es el modelo más conveniente. La estructura del artículo es la siguiente: en la sección 2 se describe cómo funcionan los modelos de agregación de Efecto Fijo y Efectos aleatorios; en la sección 3 se especifican los objetivos del trabajo; en la sección 4 se describe la metodología de investigación a aplicar; en la sección 5 se presentan los resultados de la simulación de Monte Carlo realizada; en la sección 6 se discuten los resultados obtenidos y, finalmente, la sección 7 avanza algunas conclusiones provisionales.

2. Estado de la cuestión

En este apartado se introducen los conceptos de meta-análisis (sección 2.1) y errores estadísticos (sección 2.2).

2.1 Meta-Análisis

El Meta-Análisis consiste en la integración de los resultados de un conjunto de experimentos, previamente identificados, que analizan el desempeño de un par de tratamientos predefinidos con el fin de dar una estimación cuantitativa sintética de todos los estudios disponibles [11]. Si todos los estudios incluidos en el proceso de Meta-Análisis fueran igualmente precisos y utilizaran exactamente las mismas variables respuesta, bastaría con promediar los resultados de cada uno de ellos para obtener así una conclusión final [10]. Sin embargo, en la práctica no todos los estudios tienen la misma precisión, por ello cuando se los combine se debe asignar un mayor peso a los estudios que permiten obtener información más fiable. Esto se logra combinando los resultados mediante un promedio ponderado. Por otra parte, para poder solucionar los problemas vinculados a la no uniformidad de las variables respuesta, los métodos Meta-Analíticos expresan sus resultados mediante un índice de "Tamaño de Efecto", el cual es un estimador no escalar de la relación entre una exposición y un efecto [11] y es aplicable a cualquier medida de diferencia de los resultados de dos grupos. El método de síntesis cuantitativa para variables continuas (las más utilizadas en Ingeniería del Software) más utilizado es diferencias medias ponderadas [12]. Este método es conceptualmente sencillo: el estimador de tamaño de efecto (representa la tasa de mejora de un tratamiento respecto del otro en cada experimentos) se estima como el cociente de las diferencias entre las medias y el desvío estándar conjunto:

$$g = \frac{Y^E - Y^C}{S_p} \quad \begin{array}{l} g: \text{tamaño de efecto} \\ Y^s: \text{medias de los tratamientos} \\ S_p: \text{desvío estándar conjunto} \end{array} \quad (1)$$

$$S_p = \sqrt{\frac{(n^E - 1)(s^E)^2 + (n^C - 1)(s^C)^2}{n^E + n^C - 2}} \quad \begin{array}{l} S^s: \text{desvío estándar del grupo experimental (E) y} \\ \text{de control (C)} \\ n^s: \text{numero de sujetos del grupo experimental (E)} \\ \text{y de control (C)} \end{array} \quad (2)$$

Esta función (1), desarrollada por [4], fue optimizada por [12] quienes incorporaron un factor de corrección que aumenta la fiabilidad cuando se trabaja con pocos estudios (3). Convirtiendo a la nueva función en el método de Meta-Análisis más difundido en la actualidad y el recomendado para ser utilizado en Ingeniería del Software [5].

$$d = J(N - 2) \frac{Y^E - Y^C}{S_p} \quad \begin{array}{l} d: \text{tamaño de efecto} \\ J(N - 2): \text{factor de corrección} \\ Y^s: \text{media de los grupos experimental (E) y de} \\ \text{control (C)} \\ S_p: \text{desvío estándar conjunto} \\ N: \text{número de sujetos conjunto } (n_E + n_C) \\ J: \text{factor de ajuste} \\ N: \text{número de sujetos conjunto } (n_E + n_C) \end{array} \quad (3)$$

$$J = 1 - \frac{3}{4N - 9} \quad (4)$$

Una vez estimado el tamaño de efecto para cada estudio, puede estimarse el efecto global (representa la tasa de mejora de un tratamiento respecto del otro a nivel general) aquí es donde se debe determinar qué modelo de agregación aplicar:

El modelo de efectos fijo, que asume que la variación entre los resultados de los experimentos se debe únicamente al error experimental, pero existe un valor de efecto único para toda la población [12], cuya función de estimación es:

$$d^* = \frac{\sum d_i / \sigma^2_i(d)}{\sum 1 / \sigma^2_i(d)}$$

d^* : tamaño de efecto global
 $\sum d_i / \sigma^2_i(d)$: suma de los efectos individuales dividido por la varianza
 $\sum 1 / \sigma^2_i(d)$: suma de la inversa varianza
 $\sigma^2_i(d)$: varianza total

El modelo de efectos aleatorios, que se base en la premisa de que la variación los resultados se debe no solo al error experimental, sino también a ciertas variables no controladas que los afecta y, por ende, no existe un único efecto global, sino que este depende de los estudios que hacen parte del proceso de síntesis [12]. La función de estimación es:

$$\Delta = \frac{\sum d_i / \gamma^2_i}{\sum 1 / \gamma^2_i}$$

Δ : efecto global
 $\sum d_i / \gamma^2_i$: sumatoria de los efectos individuales
 $\sum 1 / \gamma^2_i$: sumatoria de la inversa de las varianzas entre-estudios e intra-estudios

Como puede verse en las funciones 5 y 6 la diferencia entre ambos modelos radica en que para el modelo de efecto fijo no se requiere estimar la varianza entre estudios porque supone que esta debería anularse entre los distintos resultados, mientras que para el modelo de efectos aleatorios es fundamental determinar cuál es el nivel de variación entre los resultados (la varianza entre estudios) ya que este parámetro permitirá mitigar el ruido generado por las variaciones. Bajo estas condiciones, se dice, que si un meta-análisis incluyera experimentos con valores similares (es decir, no existe ruido producido por variables externas) ambos métodos deben dar igual resultado [10].

2.2 Errores estadísticos

Existen básicamente dos tipos de errores: El error tipo I, que se define como el rechazo de la hipótesis nula H_0 cuando ésta es verdadera, es equivalente a encontrar un resultado falso positivo, porque el investigador llega a la conclusión de que existe una diferencia entre las hipótesis cuando en realidad no existe; y el error tipo II, que se define como la aceptación de la hipótesis nula cuando ésta es falsa, es equivalente a la probabilidad de un resultado falso negativo, ya que el investigador llega a la conclusión de que ha sido incapaz de encontrar una diferencia que existe en la realidad [13]. En la práctica el error de tipo I es fijado de antemano (habitualmente como $\alpha=0,05$) y el error de tipo II queda supeditado, en gran medida, al tamaño del experimento (la cantidad de sujetos que el mismo incluya) que en IS muchas veces es menor a lo que la teoría estadística sugiere. Este hecho, en IS, muchas veces se debe a la falta de presupuesto económico, tiempo, recursos humanos calificados y, porque no

decirlo, falta de cultura experimental, y provoca que se asuma que dos tratamientos (técnicas o métodos) se comportan de igual forma cuando en realidad no lo hacen.

3. Metodología de Investigación

Nuestro trabajo apunta a establecer en qué condiciones es seguro aplicar cada uno de los modelos de efectos de Meta-Análisis en las actuales condiciones experimentales de la IS. Para ello, se analizará de forma conjunta como influyen la en la fiabilidad y potencia estadística de los modelos: la cantidad de experimentos, la cantidad de sujetos que los mismos poseen y la diferencia en el tamaño de efecto, mediante una simulación de Monte Carlo [14][15]. Tomando como base los parámetros definidos en [8]:

La cantidad de sujetos por experimentos oscilará entre 4 y 20 sujetos por experimento, tomando como valores de referencia: 4, 8, 10, 14 y 20 sujetos.

La cantidad de experimentos a agregar en cada meta-análisis oscilará entre 2 y 10. Incrementándose de dos en dos. Los tamaños de efecto poblacional (δ) a analizar son: bajo (0,2), medio (0,5), alto (0,8) y muy alto (1,2). En cuanto al proceso de simulación: la media poblacional del tratamiento de control (μ^c) es fijada en 100 a efectos de cálculo y el desvío estándar en el 10%, 40% y 70% de la misma y la media poblacional del tratamiento experimental se estimará de la siguiente forma $\mu^E = 100 + \delta * \sigma$. Se realizarán 1000 simulaciones con cada una de todas las combinaciones posibles de sujetos y experimentos. Los resultados vinculados a la fiabilidad indican el porcentaje de veces que el intervalo de confianza estimado (a un nivel de $\alpha = 0.05$) contuvo el valor del tamaño de efecto poblacional, mientras que los resultados vinculados a la potencia estadística indican el porcentaje de veces que dicho intervalo de confianza no contuvo el valor 0. Para facilitar la comprensión de los mismos, en las tablas finales, se resaltan las celdas en las cuales los porcentajes estimados superaban al valor mínimo fijado, $1 - \alpha = 95\%$ para la fiabilidad y $1 - \beta = 80\%$ para la potencia estadística.

4. Resultados

A continuación, en la tabla 1, se presenta un cuadro comparativo general del desempeño de los métodos analizados basado en los resultados indicados en las tabla 3, 4 y 5 del Anexo A. Para facilitar la comprensión hemos realizado dos simplificaciones a los resultados detallados: En primer lugar, hemos agrupado los valores de los parámetros (varianza, tamaño de efecto, número de experimentos y número de sujetos por experimentos) conforme a si superaban o no los criterios de fiabilidad fijados de antemano (fiabilidad $\geq 95\%$). Para ello hemos eliminado los porcentajes sustituyéndolos por los símbolos “+” y “-”, queriendo indicar que el método se comporto de modo fiable y no fiable respectivamente. En segundo lugar, hemos observado que el número de experimentos no afecta a la fiabilidad de los métodos, sino que el factor relevante es el número de sujetos totales. Por este motivo

hemos agrupado los parámetros número de experimentos y número de sujetos en uno solo (número de sujetos/experimento).

Tabla 1. Comparación de la fiabilidad de los métodos de agregación

Varianza	Efecto	Experimentos	Sujetos / experimento	Efecto fijo	Efectos aleatorios
Alta, media y baja	Bajo	< 8	Irrelevante	+	-
		>=8	Irrelevante	+	+
	Medio	< 8	Irrelevante	+	-
		>=8	< 80	+	+
		>=8	>= 80	+	+
	Alto y Muy Alto	< 8	Irrelevante	-	-
>=8		Irrelevante	-	+	

Análogamente, y usando el criterio empleado en las tablas de fiabilidad, en la tabla 2, se presenta un cuadro comparativo general del desempeño de los métodos analizados basado en los resultados indicados en las tabla 6, 7 y 8 del Anexo A.

Tabla 2. Comparación de la potencia de los métodos de agregación

Varianza	Efecto	Experimentos	Sujetos / experimento	Efecto fijo	Efectos aleatorios
Alta, media y baja	Bajo	Irrelevante	Irrelevante	-	-
	Medio	< 8	Irrelevante	-	-
		>= 8	< 80	-	-
		>= 8	>= 80	+	-
	Alto	Irrelevante	< 40	-	-
		Irrelevante	>= 40	+	-
	Muy Alto	< 8	< 32	-	-
>= 8		>= 32	+	-	

5. Conclusión

La presente simulación ha mostrado la baja potencia del modelo de efectos aleatorio dentro del contexto de simulación desarrollado, esto provoca en la práctica que el resultado final del meta-análisis hecho con este tipo de modelo tienda a dar diferencias no significativas en todo momento, no permitiendo de esta forma poder afirmar que un trata miento es mejor que otro cuando en realidad los es. Igualmente no pudo corroborarse lo indicado en [10] respecto a la baja fiabilidad de los resultados indicado por este método, ya que la fiabilidad del mismo, en general, fue superior a la del modelo de efecto fijo. Esto surge en principio como consecuencia directa de los amplios tamaños de los intervalos de confianza que el método arroja. A partir del presente trabajo, podemos observar que el modelo de efecto fijo se comporta mejor que el modelo de efectos aleatorios, presentando potencia con mas de 80 sujetos/experimentos cuando el modelo de efecto aleatorio no posee potencia en ninguno de los casos analizados; y fiabilidad para todos los casos en que la varianza es baja o media. Si bien el actual trabajo arroja luz sobre las hipótesis de trabajo originales respecto de cual de los modelos es aplicable al actual contexto experimental de la Ingeniería de Software, es difícil que en la práctica todos los experimentos que se incorporan a un Meta Análisis posean la misma cantidad de

sujetos experimentales, por ello se hace necesario desarrollar un trabajo de simulación complementario que permita determinar si agregar experimentos de diferentes tamaños afecta a la fiabilidad y potencia de los métodos. Es decir, se analizó que es lo que sucede al aplicar los dos modelos a una población con un único tamaño de efecto, queda pendiente como trabajo futuro realizar un análisis similar sobre una población que no posea un único efecto.

6. Financiamiento

Las investigaciones que se reportan en este artículo han sido financiadas parcialmente por los proyectos TIN2008-00555 y HD2008-00046 del Ministerio Español de Ciencia e Innovación y por el Proyecto de Investigación 33A105 del Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús.

7. Bibliografía

1. Juristo, N., Moreno, A. 2001. *Basics of Software Engineering Experimentation*. Kluwer 2001
2. Dieste, O.; Fernández, E.; García, R.; Juristo, N. 2010. *Hidden Evidence Behind Useless Replications*, 1st International Workshop on Replication in Empirical Software Engineering Research (RESER) en ACM/IEEE 32nd International Conference on Software Engineering (ICSE 2010).
3. Goodman C. 1996. *Literature Searching and Evidence Interpretation for Assessing Health Care Practices*. SBU. Estocolmo.
4. Glass, G. 1976 *Primary, secondary, and meta-analysis of research*. Educational Researcher 5: 3-8
5. Basili, V., Green, S., Laitenberger, O., Lanubile, F., Shull, F., Sörumgård, S., Zelkowitz, M. 1996. *The empirical investigation of perspective-based reading*. International Journal on Empirical Software Engineering, 1(2);m133-164.
6. Miller, J. 1999. *Can Results from Software Engineering Experiments be Safely Combined?*. IEEE Software Metrics Symposium, Pág. 152-158
7. Dyba, T., Aricholm, E.; Sjöberg, D.; Hannay J.; Shull, F. 2007. *Are two heads better than one? On the effectiveness of pair programming*. IEEE Software; 24(6): 12-15
8. Dieste, O., Fernández, E., García, R., Juristo, N. 2011. *Comparative analysis of meta-analysis methods: when to use which?*. 6th EASE Durham (UK)
9. Schmidt, F. and Hunter, J. 2003. *Handbook of Psychology, Research Methods in Psychology*, Chapter 21, "Meta-Analysis", Schinka, J., Velicer, W., Weiner, I. Editors, Volume 2
10. Borenstein, M.; Hedges, L.; Rothstein, H. 2007. *Meta-Analysis Fixed Effect vs. random effect*; www.Meta-Analysis.com.
11. Cochrane; 2008. *Curso Avanzado de Revisiones Sistemáticas*. www.cochrane.es/?q=es/node/198
12. Hedges, L., Olkin, I. 1985. *Statistical methods for meta-analysis*. Academic Press
13. Noortgate, W., Onghena, P. 2003. *Estimating the mean effect size in a meta-analysis: Bias, precision, and mean squared error of different weighting methods*. Behavioral research methods, instruments and computers; 35:504-511
14. Metropolis, N., Ulam, S. 1949. *The Monte Carlo Method*. Journal of the American Statistical Association, 44(247): 335-341
15. Rogers, D., 2006. *Fifty years of Monte Carlo simulations for medical physics*. Physics in Medicine and Biology, 51: R287-R301
16. Kitchenham, B. 2004. *Procedures for performing systematic reviews*. Keele University; TR/SE-0401. Keele University Technical Report.

Anexo

Tabla 3. Fiabilidad de los métodos con $\alpha = 0.05$ y desvío = 10%

Análisis de Fiabilidad												
Varianza baja (10 %)	Columnas: experimentos Filas: sujetos y efecto		Modelo de Efecto Fijo					Modelo de Efecto Aleatorio				
			2	4	6	8	10	2	4	6	8	10
Varianza baja (10 %)	EF 0,2	4	97,9	100	100	99,3	100	92,7	95,8	100	100	100
		8	100	100	100	100	100	84,9	99	100	100	100
		10	98,7	100	100	100	100	84,8	99,3	98,7	100	100
		14	87	97,8	100	100	100	88	94,4	100	99,2	100
		20	95,6	100	100	100	100	99,3	100	100	100	100
	EF 0,5	4	97,7	100	100	98,5	96,8	87,6	92	98,5	100	100
		8	100	100	98,3	98,1	98,9	74,8	90,4	98,9	100	100
		10	96,6	100	98,4	100	93	79,7	88,1	96,4	100	100
		14	85,2	96,4	100	100	97,8	89,9	84,5	92,4	97,4	100
		20	94,6	100	100	100	100	98	78,9	90,1	98,5	100
	EF 0,8	4	96,5	99,1	98,9	95	93,6	82,4	93,1	98,6	99,1	100
		8	100	98,1	96,6	93,1	94,5	71,1	83,1	93,3	97,1	98,8
		10	96,9	99,1	94,5	97,6	83,5	74,7	78,3	86,1	97,5	99
		14	86,1	92,7	91	98,3	88,7	87,5	84,5	91,8	88,7	94
		20	88,8	100	100	81,6	100	98	79	85,8	96,5	96,4
	EF 1,2	4	96,7	95,8	94,3	88,8	80,8	75	91,7	93	98,9	99,3
		8	98,9	93,9	83,4	83,5	79,7	70,1	82,7	92,1	89,3	93,3
		10	95,9	96,6	83,3	83,4	60,3	77,3	79,5	85	91,9	95,1
		14	86,2	80,3	71,5	72	57,3	86,5	83,8	95,7	91,6	98,9
		20	84	90,4	83,1	58,2	51,3	99,2	90,2	95,7	93,5	98,3

Tabla 4. Fiabilidad de los métodos con $\alpha = 0.05$ y desvío = 40%

Análisis de Fiabilidad												
Varianza media (40 %)	Columnas: experimentos Filas: sujetos y efecto		Modelo de Efecto Fijo					Modelo de Efecto Aleatorio				
			2	4	6	8	10	2	4	6	8	10
Varianza media (40 %)	EF 0,2	4	97,4	100	100	98,4	100	93,2	95,3	100	100	100
		8	100	100	100	100	100	85,2	99,1	100	100	100
		10	98,4	100	100	100	100	84,2	98,9	98,9	100	100
		14	86,4	96,9	100	100	100	89	94,1	100	98,7	100
		20	95	100	100	100	100	99,5	100	100	100	100
	EF 0,5	4	97,8	100	100	99,5	95,6	88,1	93,3	98,4	100	100
		8	100	100	99,2	98,2	98,4	77,5	89,9	98,9	100	100
		10	97,7	100	99,1	100	92,1	74,7	89,5	96,6	100	100
		14	86,4	96,8	100	100	98,6	91,3	83,5	92,9	96,9	100
		20	93,2	100	100	100	100	98,5	78,2	88,9	97,7	100
	EF 0,8	4	95,5	98,8	98,8	95,9	94,5	85,2	94,3	99	98,6	100
		8	100	98,3	96,6	94,8	92,3	73,4	86,1	91	97,9	98,6
		10	97,2	98,1	92,9	96,2	84,1	73,6	77,5	89,3	97,8	98,7
		14	85,9	92,7	89	97,7	87,8	91,7	82,8	91,5	89,6	94,7
		20	88,8	100	100	83,2	100	98,1	79,4	89	96,7	96,8
	EF 1,2	4	95,8	95,4	95	87,1	80,7	77,3	91,4	94,4	98,1	99,6
		8	98,5	94,9	83	84,6	82,8	69	84	92	89,1	93,4
		10	96,8	96,3	83	82,2	59	78	79,6	77,8	92,1	95,8
		14	84,1	79,5	70,8	73,2	52,3	85	82,3	94,1	92,5	99,5
		20	84,8	91,7	84,3	59	51,3	99,2	89,4	97	93,6	98,1

Tabla 5. Fiabilidad de los métodos con $\alpha = 0.05$ y desvío = 70%

Análisis de Fiabilidad												
Varianza alta (70%)	Columnas: experimentos Filas: sujetos y efecto		Mdelo de Efecto Fijo					Modelo de Efecto Aleatorio				
			2	4	6	8	10	2	4	6	8	10
			2	4	6	8	10	2	4	6	8	10
	EF 0,2	4	98,4	100	100	99,7	100	91,1	96,5	100	100	100
		8	100	100	100	100	100	87	98,8	100	100	100
		10	99,4	100	100	100	100	84,1	99,4	99	100	100
		14	86,2	97,3	100	100	100	88,7	93,8	100	99,6	100
		20	96,1	100	100	100	100	98,9	100	100	100	100
	EF 0,5	4	98,1	100	100	98,9	96,9	86,5	92,4	98,6	100	100
		8	100	100	99,3	97,7	97,6	76,4	91,5	99,3	100	100
		10	96,5	100	99,3	100	94,3	79,6	89,7	96,6	100	100
		14	83,7	96	100	100	98,7	90,6	82,5	92,3	97	100
		20	94,7	100	100	100	100	98,4	77,7	90,7	97,9	100
	EF 0,8	4	95,8	99,2	99,3	95	93,1	83,9	93,6	99,3	99	100
		8	100	98,3	98,3	93,2	93,7	69,7	84,3	93,6	97,1	99,4
		10	96,5	99,1	94	96	80,7	73,6	79	87	97,8	98,7
		14	84,9	93,3	90,8	98,2	91,2	89,9	81,2	92,8	91,2	95,3
		20	88,2	100	100	84,5	100	97,9	77,3	87,5	96,2	96,3
	EF 1,2	4	96	96,1	93,5	86,8	82,7	76,9	91,5	93,5	97,6	99,5
		8	99,2	93,4	82	81,5	79,8	71,8	81,7	90,2	89,5	94,1
		10	95,8	96,5	85,1	82,3	62,4	77	78,1	80,1	91,5	95,3
		14	83,9	83,9	73,4	75,5	57,4	85,5	83,9	92,5	91,6	99
		20	82,3	92,2	81,3	57,6	49,2	99,1	90,5	95,8	92,4	97,6

Tabla 6. Potencia de los métodos con $\alpha = 0.05$ y desvío = 10%

Análisis de la Potencia Estadística												
Varianza baja (10%)	Columnas: experimentos Filas: sujetos y efecto		Mdelo de Efecto Fijo					Modelo de Efecto Aleatorio				
			2	4	6	8	10	2	4	6	8	10
			2	4	6	8	10	2	4	6	8	10
	EF 0,2	4	1,3	2,6	0,6	0,5	1,2	0	0	0	0	0
		8	1,6	5,4	2	4,3	7,1	0	0	0	0	0
		10	8,3	0	7,9	5,2	13,4	0	0	0	0	0
		14	21,1	17,1	9,5	2,9	17,7	0	0	0	0	0
		20	5,5	2,6	0	14,1	0	0	0	0	0	0
	EF 0,5	4	6,9	9,6	17,3	20,7	24,1	0	0,9	0	0	0
		8	7,9	27	48,5	70,1	80,9	0	1,9	0	0	0
		10	29	35,6	60,7	82,1	85,1	0	1	0	0	0
		14	41,2	57,4	75	100	100	0	0	0	0	0
		20	40,7	98,2	100	100	100	0	0	0	0	0
	EF 0,8	4	13,9	34,6	43,3	63	80,3	0,9	0	0	0	0
		8	40,2	75,1	96,2	100	100	0	0,9	0	1,1	0
		10	57,9	95,3	99,3	100	100	0	0,5	0	0	0
		14	64,2	96,3	100	100	100	0	0,9	0	1,3	0
		20	80,2	100	100	100	100	0	0,9	0	2	0
	EF 1,2	4	32,7	74,7	96	99,2	97,6	0,8	0	2,5	0	0
		8	89,2	100	100	100	100	0	0	1,3	0	1,8
		10	93	100	100	100	100	0,5	0,6	2,9	0	0
		14	92,2	100	100	100	100	2,2	1,6	0	0,6	0
		20	100	100	100	100	100	0	0	1,2	0	0

Tabla 7. Potencia de los métodos con $\alpha = 0.05$ y desvío = 40%

Análisis de la Potencia Estadística												
	Columnas: experimentos sujetos y efecto	Filas:	Mdolo de Efecto Fijo					Modelo de Efecto Aleatorio				
			2	4	6	8	10	2	4	6	8	10
Varianza media (40 %)	EF 0,2	4	1,7	1,8	0,7	1	2,2	0	0	0	0	0
		8	1	5,3	2,5	4,3	4,6	0	0	0	0	0
		10	9,1	0	6,8	5,3	13,4	0	0	0	0	0
		14	21,9	16,6	9,7	4,4	15,1	0	0	0	0	0
		20	5,8	2,3	0	16,8	0	0	0	0	0	0
	EF 0,5	4	6,1	9,3	14,2	18,8	22,6	0	0,6	0	0	0
		8	9,2	22,8	47,4	68,6	83,8	0	1,9	0	0	0
		10	27,2	37,2	64,3	83,1	84,3	0	0,4	0	0	0
		14	40,6	59,3	73,7	100	100	0	0	0	0	0
		20	41,6	98,9	100	100	100	0	0	0	0	0
	EF 0,8	4	13,3	33,7	44,2	66	81,4	1,1	0	0	0	0
		8	41,5	75,3	95,8	100	100	0	1,5	0	0,8	0
		10	56,8	94,4	98,8	100	100	0	0,8	0	0	0
		14	67,6	96,1	100	100	100	0	1,3	0	1,2	0
		20	81,2	100	100	100	100	0	0,4	0	2,2	0
	EF 1,2	4	34,4	70,3	96,3	98,8	97,5	1	0	1,6	0	0
		8	88,3	100	100	100	100	0	0	1	0	1,4
		10	94,1	100	100	100	100	0,4	1,3	3	0	0
		14	90,7	100	100	100	100	2,7	1,2	0	0,6	0
		20	100	100	100	100	100	0	0	1,2	0	0

Tabla 8. Potencia de los métodos con $\alpha = 0.05$ y desvío = 70%

Análisis de la Potencia Estadística												
	Columnas: experimentos sujetos y efecto	Filas:	Mdolo de Efecto Fijo					Modelo de Efecto Aleatorio				
			2	4	6	8	10	2	4	6	8	10
Varianza alta (70 %)	EF 0,2	4	0,9	2,1	0,8	0,6	1,3	0	0	0	0	0
		8	2,6	6,2	1,8	5,4	6	0	0	0	0	0
		10	9,7	0	7,4	4,3	13	0	0	0	0	0
		14	18,2	15,4	10,4	5	16	0	0	0	0	0
		20	5,1	2,3	0	15,5	0	0	0	0	0	0
	EF 0,5	4	6,2	8,2	13,7	17	21,4	0	1,1	0	0	0
		8	9,2	25,4	47,2	68,2	81,1	0	0,9	0	0	0
		10	26,4	35,9	63,7	83,3	86,1	0	0,7	0	0	0
		14	39,2	54,9	76,7	100	100	0	0	0	0	0
		20	42,1	99	100	100	100	0	0	0	0	0
	EF 0,8	4	12,5	31,9	40,2	67,3	79,8	0,8	0	0	0	0
		8	41,5	78	96,9	100	100	0	1,3	0	0,6	0
		10	53,8	95,2	99	100	100	0	0,9	0	0	0
		14	65,3	96,1	100	100	100	0	1,2	0	1,2	0
		20	81,1	100	100	100	100	0	1	0	2,6	0
	EF 1,2	4	32,9	71,9	95,7	98,9	98,1	1,3	0	1,9	0	0
		8	88,6	100	100	100	100	0	0	1,3	0	0,8
		10	93,1	100	100	100	100	0,9	1,2	3,4	0	0
		14	89,9	100	100	100	100	2	1,1	0	1,2	0
		20	100	100	100	100	100	0	0	0,8	0	0