



Visual Factory Calibre – MSA4

Acerca del RR por Atributos Kappa y Efectividad

Rev. 01 – 14/10/2022

INFORMACIÓN SOBRE EL DOCUMENTO

Editado	Enric Martorell
---------	-----------------

Revisado	Jordi Marín
----------	-------------

LISTA DE REVISIONES

REV.	FECHA	DESCRIPCIÓN
------	-------	-------------

00	29/09/2022	
----	------------	--

01	14/10/2022	Error tipográfico
----	------------	-------------------



Estudio R&R por Atributos Kappa y Efectividad

Introducción:

El estudio R&R por Atributos según MSA4 mediante el índice Kappa comprobando también la efectividad es una prueba que se realiza en Visual Factory Calibre con 3 operarios. El número de ensayos lo puede elegir el usuario hasta un máximo de 50. Consiste en realizar primero un ensayo de Referencia con un determinado número de pruebas. Dichas pruebas pueden ser, por ejemplo, medidas continuas. Seguidamente establecemos unos criterios de aceptación y de rechazo, lo cual convertirá dichas medidas en “discretas”, es decir: Solo podrán adoptar dos valores: Aceptables o rechazables.

Lo recomendable es coger cada una de las muestras del ensayo previo y aleatorizar el orden en que van a ser presentadas a cada uno de los operarios.

A continuación, cada uno de los operarios, de forma independiente, realiza la evaluación de cada una de las muestras, y realiza 3 mediciones de cada una de ellas, determinando si el resultado es “Apto” o “No apto”.

Una vez recogidos los datos, se traspasan a VFCalibre para que realice todos los análisis que definen la Norma.

Las pruebas de hipótesis que se realizan en esta Norma son:

1. H_0 : los evaluadores concuerdan consigo mismos
2. H_0 : Los evaluadores concuerdan consigo mismos y con el valor de referencia de la pieza.
3. H_0 . Los evaluadores concuerdan consigo mismos y con los otros evaluadores.
4. H_0 : Los evaluadores concuerdan consigo mismos, con los otros evaluadores y con la referencia.

Por otro lado, la evaluación de riesgos del método consiste en determinar criterios de aceptación para:

1. % de acuerdos de cada evaluador
2. % de error de aceptación de piezas malas
3. % de error de rechazo de piezas buenas.

Para hacer más comprensible esta Normativa, partiremos de un ejemplo sencillo con 10 datos. El hecho de trabajar con una muestra tan pequeña facilitará la comprensión de los conceptos que se detallan y la revisión de todos los cálculos del procedimiento. Un ejemplo de datos mayor se calcularía de la misma forma.



Datos de partida

Partamos de un ejemplo como el siguiente:

		Usuarios									
		A			B			C			
Obj. N°	Referencia	1	2	3	1	2	3	1	2	3	Cód.
1	+	+	+	+	+	+	+	+	+	+	+
2	+	+	+	+	+	+	+	+	+	+	+
3	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	+	+	+	-	-	-	X
5	-	-	-	-	-	-	-	-	-	-	-
6	+	+	+	-	+	+	-	+	-	-	X
7	+	-	-	-	+	+	+	+	-	+	X
8	+	+	+	+	+	+	+	+	-	+	X
9	-	-	-	-	-	-	-	-	-	-	-
10	+	+	+	+	+	+	+	+	+	+	+

Antes de realizar cálculos, observemos un poco los resultados obtenidos:

- En los datos 1, 2, 3, 5, 9 y 10 los 3 operarios concuerdan con la Referencia. En los demás hay discrepancias. Es decir: En solo 6 de los 10 casos hay concordancia total.
- En el dato 4, B dictamina homogéneamente un “-” que la referencia da como resultado “+”. Es decir, B ha sesgado hacia “-” (Tiende a dar por malo algo que es bueno, error de 1ª especie).
- En el dato 7, A dictamina homogéneamente un “+” que la referencia da como resultado “-”. Es decir, A ha sesgado hacia “+” (Tiende a dar por bueno algo que es malo, error de 2ª especie).
- Las observaciones 6 y 7 presentan “mezclas”, o falta de concordancia de un operario consigo mismo: En la 6ª A y B concuerdan entre ellos, pero no con C, y todos no concuerdan entre sí mismos. En la 7ª observación, es solo C quien no concuerda consigo mismo.

La idea es ver cómo se reflejarán en el informe todas estas impresiones, y las que nos propone la Normativa.

Grados de Concordancia 2 a 2

El primer paso es evaluar a los diferentes protagonistas 2 a 2. Es decir: Se trata de comparar:

A vs B, A vs C, B vs C, A vs REF, B vs REF y C vs REF

La forma de evaluar en todos los casos dicho grado de concordancia es la misma. El procedimiento es el siguiente:

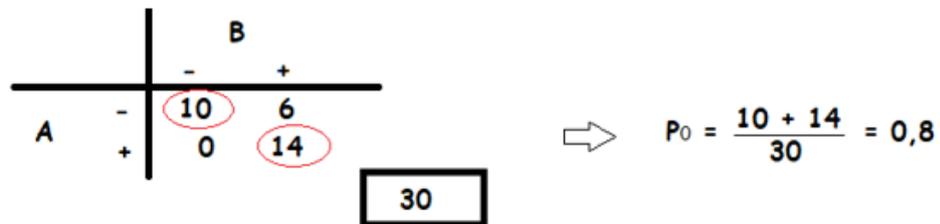


A vs B

Tabla cruzada A con B				
A		B		Total
		VNC(-)	VC(+)	
VNC(-)	Contados	10	6	16
	Esperados	5,3	10,7	16,0
VC(+)	Contados	0	14	14
	Esperados	4,7	9,3	14,0
Total Contados		10	20	30
Total Esperados		10,0	20,0	30,0

Usuarios		A			B		
Obj. N°	Referencia	1	2	3	1	2	3
1	+	+	+	+	+	+	+
2	+	+	+	+	+	+	+
3	-	-	-	-	-	-	-
4	-	-	-	-	+	+	+
5	-	-	-	-	-	-	-
6	+	+	+	-	+	+	-
7	+	-	-	-	+	+	+
8	+	+	+	+	+	+	+
9	-	-	-	-	-	-	-
10	+	+	+	+	+	+	+

Valores contados: Se cuentan todas las coincidencias entre A y B, tanto en positivo como en negativo, respecto el total:



Valores esperados:



$$P_e = (P_+) + (P_-) = 0,4888..8$$

Kappa de Cohen

En nuestro caso, nos queda:

$$K = \frac{P_0 - P_e}{1 - P_e} = 0,61$$

Del mismo modo se procede con las demás comparativas.



Obtenemos así los siguientes Índices Kappa:

Concordancia	A-B	B-C	A-C
Índice Kappa	0,61	0,61	0,73

Concordancia	A-Ref	B-Ref	C-Ref
Índice Kappa	0,74	0,71	0,74

Si no hay discrepancias (valores en la diagonal / iguales a cero) -> Kappa = 1

Criterios de aceptación:

- Si todos los índices Kappa > 0,75 Apto
- Si todos los índices Kappa >= 0,4 Apto con Reservas
- Si uno o más índices Kappa < 0,4 No Apto

El resultado del Sistema de Medida del equipo es:

Apto con reservas

Y por tanto tenemos un primer veredicto en el grado de Concordancia.

Pruebas de Estabilidad

Además, el estudio presenta los siguientes resultados:

CÁLCULO DE LA EFECTIVIDAD

EFECTIVIDAD = Número de Aciertos / Número total de ensayos

Notas: (1) El usuario concuerda en todas sus elecciones. (2) El usuario concuerda en todas sus elecciones y con la referencia. (3) Todos los usuarios concuerdan entre sí mismos y entre ellos. (4) Todos los usuarios concuerdan entre sí mismos, entre ellos y con la referencia.

Causa	% Usuario (1)			% = Puntuación vs. Atributos (2)		
	Usuario A	Usuario B	Usuario C	Usuario A	Usuario B	Usuario C
Total ensayos	10	10	10	10	10	10
# Aciertos	9	9	7	8	8	7
Falsos negativos (operario sesga hacia rechazo)				0	1	0
Falsos positivos (operario sesga hacia aceptación)				1	0	0
Mezclas				2	2	3
95% UCI	100%	100%	92%	97%	97%	92%
Puntuación	90%	90%	70%	80%	80%	70%
95% LCI	55%	55%	35%	44%	44%	35%

	% Puntuación Efectiva del Sistema (3)		% Puntuación Efectiva del Sistema (4) vs. Referencia	
	Total ensayos	# Aciertos	Total ensayos	# Aciertos
Total ensayos	10		10	
# Aciertos	6		6	
95% UCI		100%		100%
Puntuación		60%		60%
95% LCI		25%		25%



Analicemos este segundo cuadro de resultados paso por paso:

CÁLCULO DE LA EFECTIVIDAD

EFECTIVIDAD = Número de Aciertos / Número total de ensayos

Notas: (1) El usuario concuerda en todas sus elecciones. (2) El usuario concuerda en todas sus elecciones y con la referencia. (3) Todos los usuarios concuerdan entre sí mismos y entre ellos. (4) Todos los usuarios concuerdan entre sí mismos, entre ellos y con la referencia.

Usuarios		A			B			C			Cód.
Obj. Nº	Referencia	1	2	3	1	2	3	1	2	3	
1	+	+	+	+	+	+	+	+	+	+	+
2	+	+	+	+	+	+	+	+	+	+	+
3	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	+	+	+	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-
6	+	+	+	+	+	+	+	+	+	+	X
7	+	+	+	+	+	+	+	+	+	+	X
8	+	+	+	+	+	+	+	+	+	+	X
9	-	-	-	-	-	-	-	-	-	-	-
10	+	+	+	+	+	+	+	+	+	+	+

Causa	% Usuario (1)			% = Puntuación vs. Atributos (2)		
	Usuario A	Usuario B	Usuario C	Usuario A	Usuario B	Usuario C
Total ensayos	10	10	10	10	10	10
# Aciertos	9	9	7	8	8	7
Falsos negativos (operario sesga hacia rechazo)				0	1	0
Falsos positivos (operario sesga hacia aceptación)				1	0	0
Mezclas				2	2	3
95% UCI	100%	100%	92%	97%	97%	93%
Puntuación	90%	90%	70%	80%	80%	70%
95% LCI	55%	55%	35%	44%	44%	35%

Causa	% Puntuación Efectiva del Sistema (3)		% Puntuación Efectiva del Sistema (4) vs. Referencia	
	Usuario A	Usuario B	Usuario A	Usuario B
Total ensayos	10	10	10	10
# Aciertos	6	6	6	6
95% UCI	100%	100%	100%	100%
Puntuación	60%	60%	60%	60%
95% LCI	25%	25%	25%	25%

Coincidencias de un Usuario consigo mismo (1)

Si nos fijamos en la columna de **#Aciertos**, vemos que A no concuerda consigo mismo en el caso 6, B no concuerda consigo mismo en el caso 6, y C no concuerda consigo mismo en los casos 6, 7 y 8.

La columna **Puntuación** refleja dichos resultados en porcentaje, y coincide con los valores mostrados en **#Aciertos**.

Coincidencias de un Usuario consigo mismo y con la Referencia (2)

Si nos fijamos en la columna de **#Aciertos**, vemos que A no concuerda consigo mismo y con la Referencia en los casos 6 y 7, B no concuerda consigo mismo en los casos 4 y 6, y C no concuerda consigo mismo en los casos 6, 7 y 8.

La columna **Puntuación** refleja dichos resultados en porcentaje, y coincide con los valores mostrados en **#Aciertos**.

Falsos negativos (Operario sesga hacia rechazo)

El usuario A es consecuente consigo mismo pero falla en su diagnóstico respecto a la Referencia en la Observación 7, evaluando como “-“ un “+“. Es un caso de falso negativo o tendencia a rechazar algo que es bueno (Error de 1ª especie). No se ve nada de este tipo en los demás usuarios.

Falsos positivos (Operario sesga hacia acierto)

El usuario B es consecuente consigo mismo pero falla en su diagnóstico respecto a la Referencia en la Observación 4, evaluando como “+“ un “-“. Es un caso de falso positivo o tendencia a aceptar algo que es malo (Error de 2ª especie). No se ve nada de este tipo en los demás usuarios.



Mezclas

Reflejan los casos donde un Operario no coincide consigo mismo en una evaluación determinada o con la Referencia. Esto ocurre con el operario A en las observaciones 6 y 7, con el B en las observaciones 4 y 6, y con el C en las observaciones 6, 7 y 8.

Puntuación Efectiva del Sistema (3)

En el ejemplo se aprecia que las filas 4, 6, 7 y 8 marcan diferencias de criterios entre los Usuarios A, B y C. Por tanto son 6 las Observaciones plenamente coincidentes entre los Usuarios entre sí.

La columna **Puntuación** refleja dichos resultados en porcentaje, y coincide con los valores mostrados en **#Aciertos**.

Puntuación Efectiva del Sistema versus la Referencia (4)

En el ejemplo se aprecian 4 filas marcadas como discrepantes. Por tanto, son 6 las Observaciones plenamente coincidentes entre los Usuarios entre sí y con la Referencia.

Determinación de los Intervalos de Confianza

Todos los Intervalos de Confianza expresados en este estudio se calculan de la misma manera. Se parte de cálculos a partir de la prueba de Fischer.

Estudio de Efectividad

Es la última parte del estudio:

Tabla de Estudio de Efectividad

	Efectividad	Proporción perdida	Proporción de falsa alarma
Usuario A	80%	0,0%	22,2%
Usuario B	80%	25,0%	5,6%
Usuario C	70%	0,0%	22,2%

Aplica"
Aplica"

Criterios de aceptación

Efectividad	perdida	alarma
>=90%	<=2%	<=5%
>=80%	<=5%	<=10%
<80%	>5%	>10%

Observaciones:

Usuario A parcialmente aceptable para el Sistema de medida - puede necesitar mejora.

Usuario B parcialmente aceptable para el Sistema de medida - puede necesitar mejora.

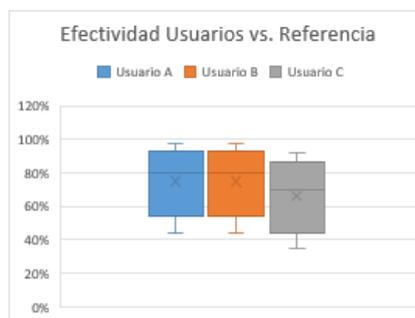
inaceptable para Usuario C, para el Sistema de medida - ha de mejorar.

Tendencia a aprobar defectos por parte de Usuario B.

Tendencia a rechazar conformes por parte de Usuario A, Usuario C

Resultado del estudio de Efectividad:

Efectividad	Proporción perdida	Proporción de falsa alarma
A. Res.	APTO	NO APTO
A. Res.	NO APTO	A. Res.
NO APTO	APTO	NO APTO





Efectividad

Recoge los datos siguientes calculados anteriormente:

Causa	% Usuario (1)			% = Puntuación vs. Atributos (2)		
	Usuario A	Usuario B	Usuario C	Usuario A	Usuario B	Usuario C
Total ensayos	10	10	10	10	10	10
# Aciertos	9	9	7	8	8	7
Falsos negativos (operario sesga hacia rechazo)				0	1	0
Falsos positivos (operario sesga hacia aceptación)				1	0	0
Mezclas				2	2	3
95% UCI	100%	100%	92%	97%	97%	93%
Puntuación	90%	90%	70%	80%	80%	70%
95% LCI	55%	55%	35%	44%	44%	35%

Proporción perdida

Se obtiene de las comparaciones de cada uno de los usuarios con la Referencia de la siguiente forma:

	Efectividad	Proporción perdida
Usuario A	80%	0,0%
Usuario B	80%	25,0%
Usuario C	70%	0,0%

$0 / 12 = 0$

C	Ref		Total
	VNC(-)	VC(+)	
VNC(-)	Contados: 12	4	16
	Esperados: 6,4	3,6	10,0
VC(+)	Contados: 0	14	14
	Esperados: 5,6	8,4	14,0
Total Contados	12	18	30
Total Esperados	12,0	18,0	30,0

$0 / 12 = 0$

A	Ref		Total
	VNC(-)	VC(+)	
VNC(-)	Contados: 12	4	16
	Esperados: 6,4	3,6	10,0
VC(+)	Contados: 0	14	14
	Esperados: 5,6	8,4	14,0
Total Contados	12	18	30
Total Esperados	12,0	18,0	30,0

$3 / 12 = 0,25$

B	Ref		Total
	VNC(-)	VC(+)	
VNC(-)	Contados: 3	1	10
	Esperados: 4,0	6,0	10,0
VC(+)	Contados: 3	17	20
	Esperados: 8,0	12,0	20,0
Total Contados	12	18	30
Total Esperados	12,0	18,0	30,0

VNC(-) = Valores No Conformes VC(+) = Valores Conformes

Valores por debajo del 5% dan resultado de "APTO", valores entre 5 y 10% dan como resultado "APTO CON RESERVAS". Valores mayores dan resultado de "NO APTO".

En este apartado se evalúa la tendencia del usuario a aceptar productos no conformes (es decir, a cometer lo que en Calidad se conoce como "Error de 2ª especie").



Proporción de falsa alarma

Se obtiene de las comparaciones de cada uno de los usuarios con la Referencia de la siguiente forma:

Tabla de Estudio de Efectividad

	Efectividad	Proporción de falsa alarma
Usuario A	80%	22,2%
Usuario B	80%	5,6%
Usuario C	70%	22,2%

Tabla cruzada A con Ref

		Ref		Total
		VNC(-)	VC(+)	
VNC(-)	Contados	12	4	16
	Esperados	6,4	3,6	10,0
VC(+)	Contados	0	14	14
	Esperados	5,6	8,4	14,0
Total Contados		12	18	30
Total Esperados		12,0	18,0	30,0

Tabla cruzada B con Ref

		Ref		Total
		VNC(-)	VC(+)	
VNC(-)	Contados	9	1	10
	Esperados	4,0	6,0	10,0
VC(+)	Contados	3	17	20
	Esperados	8,0	12,0	20,0
Total Contados		12	18	30
Total Esperados		12,0	18,0	30,0

Tabla cruzada C con Ref

		Ref		Total
		VNC(-)	VC(+)	
VNC(-)	Contados	12	4	16
	Esperados	6,4	3,6	10,0
VC(+)	Contados	0	14	14
	Esperados	5,6	8,4	14,0
Total Contados		12	18	30
Total Esperados		12,0	18,0	30,0

Valores por debajo del 2% dan resultado de "APTO", valores entre 2 y 5% dan como resultado "APTO CON RESERVAS". Valores mayores dan resultado de "NO APTO".

En este apartado se evalúa la tendencia del usuario a rechazar productos conformes (es decir, a cometer lo que en Calidad se conoce como "Error de 1ª especie").

Evaluación final:

Y en función de los valores obtenidos se determinan los resultados:

Tabla de Estudio de Efectividad

	Efectividad	Proporción perdida	Proporción de falsa alarma
Usuario A	80%	0,0%	22,2%
Usuario B	80%	25,0%	5,6%
Usuario C	70%	0,0%	22,2%

Resultado del estudio de Efectividad:

Efectividad	Proporción perdida	Proporción de falsa alarma
A. Res.	APTO	NO APTO
A. Res.	NO APTO	A. Res.
NO APTO	APTO	NO APTO

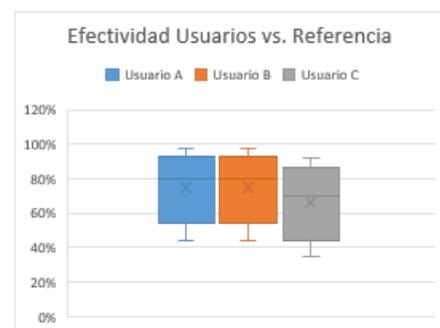
Aplica"
Aplica"

Criterios de aceptación

Efectividad	perdida	alarma
>=90%	<=2%	<=5%
>=80%	<=5%	<=10%
<80%	>5%	>10%

Observaciones:

- Usuario A parcialmente aceptable para el Sistema de medida - puede necesitar mejora.
- Usuario B parcialmente aceptable para el Sistema de medida - puede necesitar mejora.
- inaceptable para Usuario C para el Sistema de medida - ha de mejorar.
- Tendencia a aprobar defectos por parte de Usuario B.
- Tendencia a rechazar conformes por parte de Usuario A, Usuario C



En **Observaciones** se detallan las Conclusiones obtenidas en este apartado del Estudio.