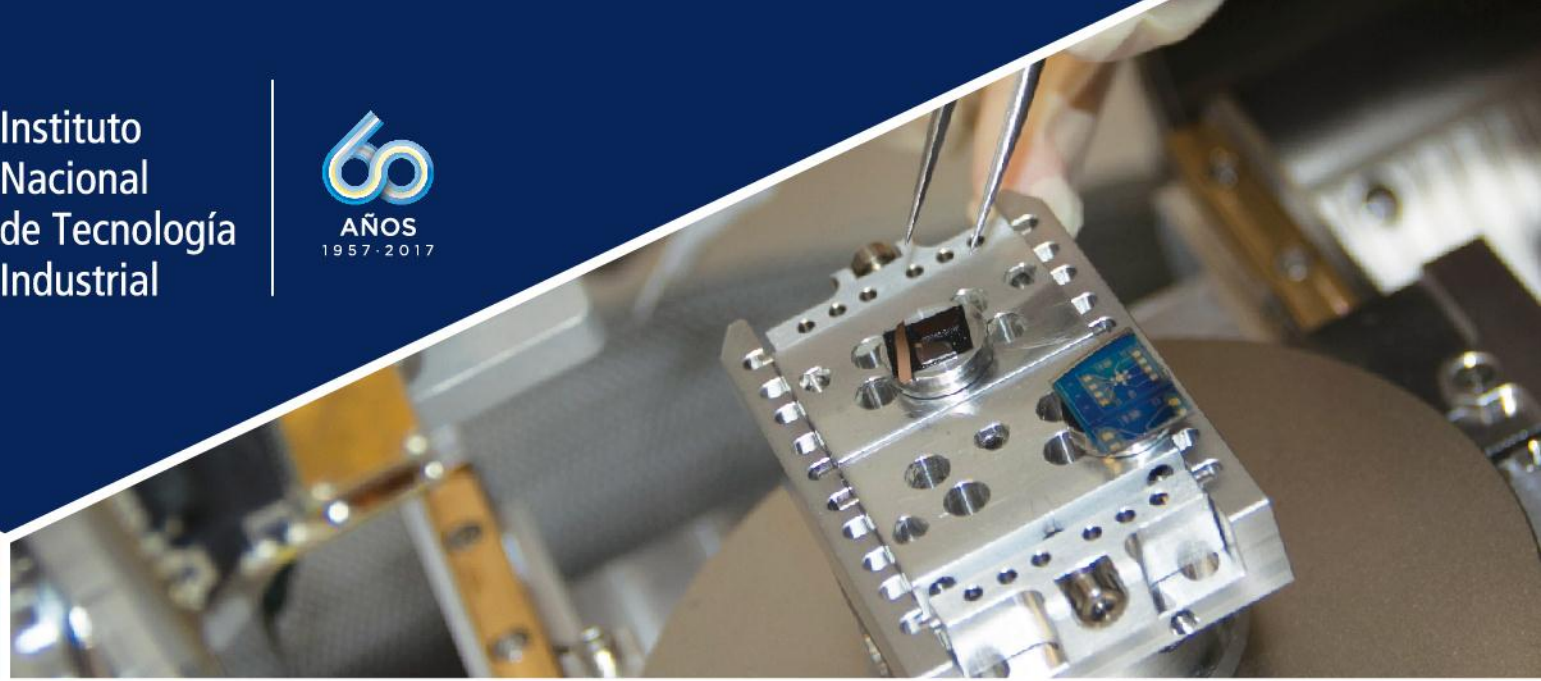




Instituto
Nacional
de Tecnología
Industrial



**MÉTODOS ESTADÍSTICOS
APLICADOS A LA EVALUACIÓN DE
RESULTADOS
INTERLABORATORIALES**



MÉTODOS ESTADÍSTICOS APLICADOS EN LA EVALUACIÓN DE RESULTADOS DE INTERLABORATORIOS

- ***ISO 5725 Accuracy (Trueness and Precision) of measurement methods and results***
- ***ISO 13528, Statistical methods for use in proficiency testing by interlaboratory comparisons***

- ***Control de la homogeneidad de las muestras***
- ***Control de la estabilidad de las muestras***
- ***Análisis de los resultados enviados por los participantes***
- ***Determinación del valor medio***
- ***Determinación de la desviación estándar interlaboratorio σ_{IL}***
- ***Determinación de indicadores de desempeño de los participantes***

- **Control de la homogeneidad de las muestras**
- **Control de la estabilidad de las muestras**
- *Análisis de los resultados enviados por los participantes*
- *Determinación del valor medio*
- *Determinación de la desviación estándar interlaboratorio σ_{IL}*
- *Determinación de indicadores de desempeño de los participantes*

Control de la homogeneidad

Cuando un lote de material es fraccionado para distribuir a varios laboratorios, las unidades pueden presentar diferencias entre ellas en relación a los parámetros a medir, Se busca demostrar que esas diferencias son no significativas en relación a la variación introducida por las mediciones hechas por los participantes,

- Se realiza antes de enviar las muestras
- Al menos para los parámetros más representativos
- Se analizan por duplicado una cantidad de muestras (al menos 10)
- Se aplica ANOVA

- O bien, debe cumplirse $s_m \leq 0.3 \cdot \hat{\sigma}_{IL}$

Control de la estabilidad

Asegurar que los ítems de ensayo permanezcan estables a lo largo de todo el periodo de medición,

De esta forma, las posibles diferencias entre resultados del EA pueden explicarse por variaciones de reproducibilidad inter-laboratorio o por efectos propios de los participantes, más que a efectos causados por la inestabilidad de las muestras,

Las muestras deben ser estables para los parámetros incluidos en el EA, durante el período que va desde su preparación hasta culminar la fecha propuesta para las mediciones

$$|\bar{y}_I - \bar{y}_F| \leq 0.3 \cdot \hat{\sigma}_{IL}$$

- *Control de la homogeneidad de las muestras*
- *Control de la estabilidad de las muestras*
- ***Análisis de los resultados enviados por los participantes***
- *Determinación del valor medio*
- *Determinación de la desviación estándar interlaboratorio σ_{IL}*
- *Determinación de indicadores de desempeño de los participantes*

- Los participantes tratan y miden la muestra del mismo modo que las muestras de rutina
- Se piden resultados por triplicado. Se trabaja con el promedio.
- El interlaboratorio evalúa el desempeño global del laboratorio
- Esto incluye la transcripción de resultados

Qué pasa cuando se reciben datos con unidades erróneas?

Los datos son analizados tal como han sido enviados, Es por esto que los participantes deben tener especial cuidado con los errores de tipeo o transcripción. Por ejemplo, si el resultado analítico

$$1,33 \text{ g} / 100 \text{ g}$$

es informado, por error, como

$$13,3 \text{ g} / 100 \text{ g} \quad \text{o como} \quad 133 \text{ g} / 100 \text{ g}$$

el participante obtendrá un mal resultado,

- *Control de la homogeneidad de las muestras*
- *Control de la estabilidad de las muestras*
- *Análisis de los resultados enviados por los participantes*
- ***Determinación del valor medio***
- *Determinación de la desviación estándar interlaboratorio σ_{IL}*
- *Determinación de indicadores de desempeño de los participantes*

¿Cuál es el verdadero valor de cada analito?

No conocemos el “valor verdadero” de cada muestra, La complejidad intrínseca de la matriz hace imposible para la mayoría de los parámetros obtener valores metrológicamente trazables

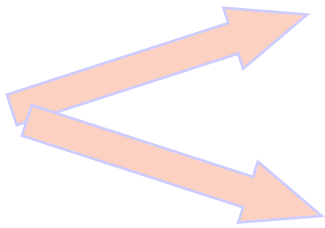
El valor de referencia de cada parámetro es calculado “por consenso” a partir de los resultados aportados por los propios participantes, usando métodos estadísticos adecuados,

Un “buen” resultado en el interlaboratorio, por tanto, debe ser interpretado como un resultado cercano a la mayoría de los participantes

Consenso entre todos los participantes

El valor asignado (VMIL) se calculan a partir de los valores obtenidos por los laboratorios participantes, utilizando procedimientos estadísticos robustos

Estimaciones de valor medio



Promedio (medida eficiente)

Mediana (medida robusta)

Buscamos medidas robustas y eficientes al mismo tiempo:

ALGORITMO A

- *Control de la homogeneidad de las muestras*
- *Control de la estabilidad de las muestras*
- *Análisis de los resultados enviados por los participantes*
- *Determinación del valor medio*
- ***Determinación de la desviación estándar interlaboratorio σ_{IL}***
- *Determinación de indicadores de desempeño de los participantes*

Desviación estándar del interlaboratorio

- Por consenso
- Valor determinado a priori (por ejemplo, normalizado)

Riesgos y desventajas de la estimación por consenso:

Puede ser el modo más natural, Sin embargo...

El valor puede variar sustancialmente entre ronda y ronda, dificultando a los participantes interpretar la evolución de sus propios desempeños,

Por ejemplo, si el desempeño global del grupo de participantes va continuamente mejorando, las sucesivas estimaciones de producirán valores decrecientes, y la mejora no será apreciada

Lo mismo en el caso inverso, cuando el desempeño global tienda a ir empeorando,

- *Control de la homogeneidad de las muestras*
- *Control de la estabilidad de las muestras*
- *Análisis de los resultados enviados por los participantes*
- *Determinación del valor medio*
- *Determinación de la desviación estándar interlaboratorio σ_{IL}*
- ***Determinación de indicadores de desempeño de los participantes***

➤ Indicadores de desempeño

Se define el indicador “z” (z-score) de la siguiente manera:

$$z = \frac{x - VMIL}{\sigma_{IL}}$$

Diferentes modos de evaluar el desempeño

$$\sigma_{\text{consenso}} \rightarrow z_{\text{consenso}} = \frac{x - x_{\text{asignado}}}{\sigma_{\text{consenso}}}$$

$$\sigma_{\text{norma}} \rightarrow z = \frac{x - x_{\text{asignado}}}{\sigma_{\text{norma}}}$$

¿Más exigente?

Ejemplo: C oxidable

(Norma IRAM-SAGyP 29571-2)

Tabla B – Valores de repetibilidad y reproducibilidad

Identificación de la muestra	Muestras									
	A	B	C	D	E	F	G	H	I	J
Número de laboratorios participantes	17	17	17	17	17	17	17	17	17	17
Número de resultados de ensayo aceptados	51	51	51	51	51	51	51	51	51	51
Contenido promedio de carbono orgánico oxidable (mg/g), base seca	12,8	16,5	7,6	7,8	21,6	16,1	21,6	11,1	20,6	25,9
Desviación estándar de la repetibilidad, S_r	0,42	0,43	0,32	0,34	0,64	0,46	0,55	0,45	0,62	0,54
Varianza de repetibilidad, S_r^2	0,18	0,19	0,11	0,11	0,41	0,21	0,30	0,20	0,39	0,30
Límite de repetibilidad, $r(2,8 \cdot S_r)$	1,17	1,21	0,91	0,94	1,80	1,28	1,54	1,25	1,75	1,52
Coefficiente de variación (S_r relativo)	3,26	2,62	4,26	4,32	2,98	2,84	2,55	4,01	3,03	2,10
Desviación estándar de la reproducibilidad, S_R	0,96	1,03	0,55	0,64	1,48	1,15	1,58	0,91	1,47	1,81
Varianza de reproducibilidad, S_R^2	0,93	1,06	0,30	0,41	2,18	1,31	2,43	0,83	2,16	3,27
Límite de reproducibilidad, $R(2,8 \cdot S_R)$	2,69	2,88	1,54	1,78	4,14	3,21	4,38	2,54	4,11	5,06
Coefficiente de variación (S_r relativo)	7,49	6,25	7,22	8,17	6,85	7,14	7,22	8,16	7,13	6,99

CV promedio: 7,3%

Resumen de las últimas tres rondas PROINSA– C oxidable

Ronda		2014		2015		2016	
Muestra		A	B	A	B	A	B
Algoritmo A	VMIL / (g/100 g)	1,24	2,00	1,09	2,85	1,96	1,25
	$s_{consenso}$ / (g/100 g)	0,10	0,13	0,13	0,27	0,15	0,10
	CV	8,1%	6,5%	11,9%	9,5%	7,7%	8,0%
IRAM-SAGyP 29571-2	CV	7,3%	7,3%	7,3%	7,3%	7,3%	7,3%
	σ_{norma} / (g/100 g)	0,09	0,15	0,08	0,21	0,14	0,09

**Sólo para los participantes que
declararon haber aplicado la
norma**

Ronda		2014		2015		2016	
Muestra		A	B	A	B	A	B
Algoritmo A	VMIL / (g/100 g)	1,33	2,11	1,19	3,00	2,10	1,33
	s_{IL} / (g/100 g)	0,20	0,29	0,20	0,51	0,31	0,19
	CV	15%	14%	17%	17%	15%	14%
IRAM-SAGyP 29571-2	CV	7,3%	7,3%	7,3%	7,3%	7,3%	7,3%
	σ_{IL} / (g/100 g)	0,10	0,15	0,09	0,22	0,15	0,10

**Todos los
participantes**

Diferentes modos de evaluar el desempeño

$$\sigma_{consenso} \rightarrow z_{consenso} = \frac{x - x_{asignado}}{\sigma_{consenso}}$$

$$\sigma_{norma} \rightarrow z = \frac{x - x_{asignado}}{\sigma_{norma}}$$

Más exigente para los participantes que aplican otros métodos

Por qué dos muestras (A y B)?

- Para evaluar el desempeño de los laboratorios en diferentes rangos de concentración
- La duplicación de muestras permite evaluar tipos de errores (sistemáticos o aleatorios)
- De esta manera, el participante tiene más información para corregirlos

Muestra **A**: $z = 1.9$

Muestra **B**: $z = -0.6$



Indicio de errores aleatorios

En cambio

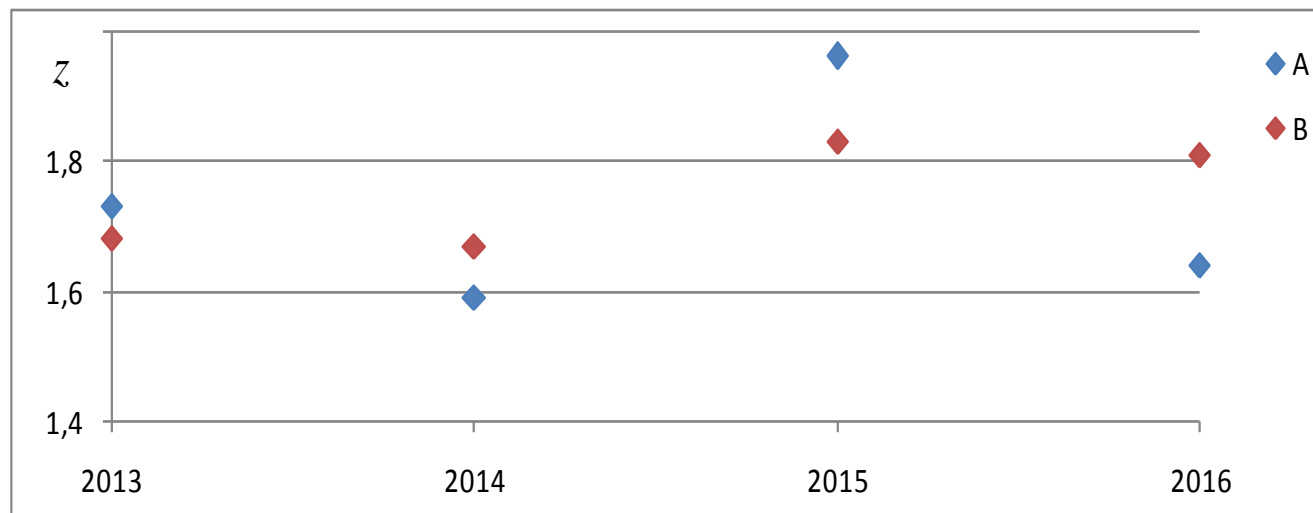
Muestra **A**: $z = 1.9$

Muestra **B**: $z = 1.7$



Indicio de error sistemático

La participación
sistemática
también ofrece
información
muy valiosa



OBJETIVOS PROINSA:

“,,,Propender a mejorar la calidad de los resultados analíticos de los ensayos de los laboratorios de suelos del país, públicos y privados”



INTI



AÑOS
1957 - 2017

Preguntas?.....

