

Maestría en Enseñanza de la Matemática

Tercera Promoción

Autor: Bossano Cueva José Roberto.
Director: M.Sc. Flores Sánchez Miguel Alfonso

Universidad de las Fuerzas Armadas ESPE

2019

APLICACIÓN DE UNA NUEVA METODOLOGÍA BASADA EN EL MÉTODO BOOTSTRAP PARA LA DETECCIÓN DE RESULTADOS ATÍPICOS EN ESTUDIOS INTERLABORATORIO: UN CASO PRÁCTICO EN LABORATORIOS CLÍNICOS.

El Problema

El procedimiento para realizar Estudios Interlaboratorio se encuentra detallado en las normas:

- ISO 5725 (*Accuracy - Trueness and precision - of measurement methods and results*).
- ASTM E-691 (*Standard practice for conducting an interlaboratory study to determinate the precision of a test method*).

En las mismas se explica que para realizar un Estudio Interlaboratorio, los datos deben seguir una distribución normal, y las medias de los laboratorios deben ser aproximadamente iguales.

Los Estudios Interlaboratorio (ILS)

Un Estudio Interlaboratorio tiene las siguientes características:

- Son estudios colaborativos donde un grupo de 8 a 15 laboratorio se juntan para analizar el mismo material.
- El objetivo es obtener una estimación de la desviación σ de los laboratorios que aplican un método de medición en función de la repetibilidad y reproducibilidad.
- En un ILS, un nivel de significancia de 0.1 % significa que existe un 0.1 % de probabilidad de decidir erróneamente que dos grupos de mediciones son diferentes cuando provienen de la misma población.

El estadístico h de Mandel

Este estadístico detecta variaciones en las medias de los laboratorios, cuando son comparados con otros.

$$h_l = \frac{d_l}{S_{\bar{x}}} = \frac{\bar{x}_l - m}{\sqrt{\sum_{l=1}^L \frac{(\bar{x}_l - m)^2}{L-1}}}; l = 1, \dots, L$$

Donde m es la media de las medias de todos los laboratorios, es decir:

$$m = \frac{\sum_{l=1}^L \bar{X}_l}{L}$$

El estadístico k de Mandel

Este estadístico detecta variaciones en las desviaciones estándar de los laboratorios, cuando son comparados con otros.

El estadístico viene dado por:

$$k_l = \frac{S_i}{S_r} = \frac{\sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}}}{\sqrt{\frac{\sum_{i=1}^L S_i^2}{L}}}$$

Los estadísticos h y k de Mandel

Nuestro objetivo es calcular los estadísticos h y k por métodos de remuestreo, en particular, por el método bootstrap.

El método Bootstrap

De manera general, el método bootstrap consta de los siguientes pasos:

1. Se toma una muestra aleatoria y se generan grandes cantidades de muestras con reemplazamiento a partir de la muestra original.
2. Se calcula el valor del estadístico en cada muestra Bootstrap, de esta forma se obtiene la distribución bootstrap del estadístico.
3. Se emplea la distribución bootstrap que ha sido estimada para obtener información sobre la forma, la localización, y la dispersión de la distribución muestral del estadístico.

El método Bootstrap

Las distribuciones bootstrap tienen la misma forma y dispersión que la distribución muestral del estadístico, pero centradas en el valor del estadístico en la muestra original, por lo que es muy importante que la muestra original sea representativa de la población.

El método Bootstrap

El bootstrap puede servir para realizar pruebas para determinar si existe diferencias entre grupos. Esta prueba se realiza en base a la hipótesis nula de que no existen diferencias entre tales grupos (todos los laboratorios provienen de la misma distribución).

el método Bootstrap

1. Se mezclan las observaciones de los k laboratorios en un solo conjunto, conocido como el conjunto *pool* de los datos.
2. Se aplica el muestreo con reemplazamiento para generar una remuestra del mismo tamaño.
3. Se vuelven a formar k grupos del mismo tamaño que los originales, asignando de forma aleatoria las observaciones obtenidas en el paso anterior.
4. Se calcula el valor del estadístico de interés en cada nuevo grupo.
5. Se repite este proceso una gran cantidad de veces (por ejemplo 500, 1000, 10000) para poder formar la distribución Bootstrap.

Hipótesis de reproducibilidad

La hipótesis nula de reproducibilidad que proponemos establece que no hay diferencia entre las medias de los laboratorios, es decir:

$$H_o = \mu_1 = \mu_2 = \dots = \mu_l$$

Hipótesis de repetibilidad

La hipótesis nula de reproducibilidad que proponemos establece que no hay diferencia entre las desviaciones estándar de los laboratorios, es decir:

$$H_o = \sigma_1 = \sigma_2 = \dots = \sigma_l$$

Metodología

El algoritmo bootstrap propuesto consta de los siguientes pasos:

1. Se agrupan los datos provenientes de los diferentes laboratorios.
2. Se elabora un diagrama de caja de los datos, luego se separan los datos que están fuera del rango intercuantíl.
3. Se realiza muestreo con reemplazamiento con los datos que no fueron separados para generar una remuestra del mismo tamaño del conjunto en el que se mezclaron los datos originalmente.
4. Se asignan de manera aleatoria las observaciones de la remuestra en grupos de igual tamaño a los laboratorios originales.
5. Dentro de la remuestra se calcula el valor del estadístico de interés (h o k) para cada laboratorio.
6. Se repiten los pasos 3, 4, 5, 6 un gran número de veces, por ejemplo, $B=500$. Con esto se obtiene la distribución Bootstrap del estadístico.

Metodología

1. **Para el estadístico h:** Se determinan los valores críticos del estadístico h, calculando los cuantiles de la distribución bootstrap del estadístico para los niveles $\frac{\alpha}{2}$ y $1 - \frac{\alpha}{2}$.
2. Se calculan los estadísticos h en la muestra original, si un valor de h es mayor que el valor absoluto del valor crítico del estadístico h, el laboratorio correspondiente se considera como atípico.
3. **Para el estadístico k:** Se determina el valor crítico del estadístico k, calculando el cuantil de la distribución bootstrap del estadístico k para el nivel $1 - \alpha$.
4. Se calculan los estadísticos k en la muestra original, si un valor de k es mayor que valor crítico del estadístico, el laboratorio correspondiente se considera como atípico.

Estudio de Simulación

Fueron considerados los siguientes valores durante el estudio de simulación:

- Tamaños de las muestras: $\{3, 6\}$.
- Número de laboratorios: $\{5, 10\}$.
- Nivel de significancia: $\alpha = 0.010$.
- Variaciones de la media: $\Delta_{\mu} \in \{-3, -2, -1, 0, 1, 2, 3\}$.
- Variaciones de la desviación estándar: $\Delta_{\sigma} \in \{1, 1.5, 2, 2.5, 3, 3.5, 4\}$

Incumplimiento de la hipótesis de reproducibilidad para la distribución normal

L	n	Estadístico h Bootstrap							Estadístico h de Mandel						
		(-3)	(-2)	(-1)	(0)	(1)	(2)	(3)	(-3)	(-2)	(-1)	(0)	(1)	(2)	(3)
5	3	0.367	0.170	0.045	0.009	0.044	0.167	0.353	0.416	0.180	0.042	0.009	0.043	0.188	0.408
	6	0.634	0.360	0.090	0.009	0.077	0.343	0.640	0.673	0.350	0.092	0.010	0.076	0.358	0.676
10	3	0.869	0.495	0.108	0.012	0.105	0.496	0.889	0.889	0.508	0.107	0.008	0.094	0.499	0.919
	6	0.996	0.856	0.258	0.008	0.241	0.841	0.996	0.997	0.860	0.259	0.009	0.242	0.849	0.999

- Los resultados obtenidos por ambos métodos son similares.
- Cuando aumenta el tamaño n de la muestra, aumenta la potencia de la prueba.

Incumplimiento de la hipótesis de reproducibilidad para la distribución de Laplace

L	n	Estadístico h Bootstrap							Estadístico h de Mandel						
		(-3)	(-2)	(-1)	(0)	(1)	(2)	(3)	(-3)	(-2)	(-1)	(0)	(1)	(2)	(3)
5	3	0.233	0.106	0.029	0.012	0.028	0.108	0.219	0.247	0.111	0.029	0.012	0.029	0.109	0.240
	6	0.386	0.178	0.067	0.010	0.043	0.189	0.442	0.412	0.177	0.066	0.010	0.041	0.188	0.470
10	3	0.560	0.265	0.061	0.019	0.067	0.244	0.534	0.605	0.268	0.053	0.015	0.067	0.246	0.573
	6	0.877	0.527	0.120	0.011	0.119	0.503	0.873	0.899	0.529	0.117	0.011	0.113	0.522	0.893

- Los resultados obtenidos por ambos métodos son similares.
- Cuando aumenta el tamaño n de la muestra, aumenta la potencia de la prueba.

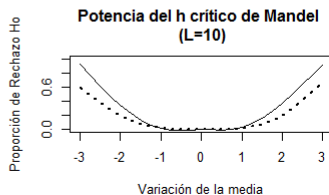
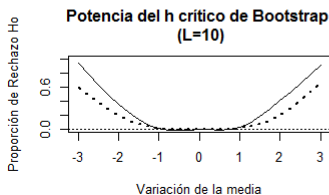
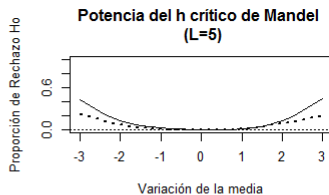
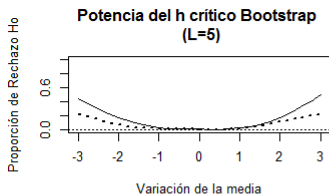
Incumplimiento de la hipótesis de reproducibilidad para la distribución sesgada a la derecha

L	n	Estadístico h Bootstrap							Estadístico h de Mandel						
		(-3)	(-2)	(-1)	(0)	(1)	(2)	(3)	(-3)	(-2)	(-1)	(0)	(1)	(2)	(3)
5	3	0.232	0.066	0.018	0.004	0.020	0.116	0.216	0.236	0.058	0.010	0.002	0.010	0.092	0.202
	6	0.440	0.170	0.028	0.006	0.022	0.174	0.490	0.424	0.124	0.026	0.002	0.012	0.126	0.440
10	3	0.608	0.202	0.018	0.002	0.018	0.210	0.668	0.596	0.182	0.010	0.002	0.012	0.190	0.656
	6	0.940	0.364	0.016	0.004	0.036	0.414	0.918	0.934	0.344	0.014	0.002	0.030	0.382	0.916

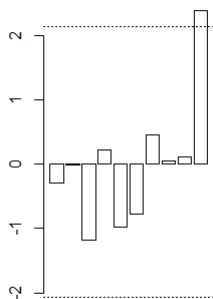
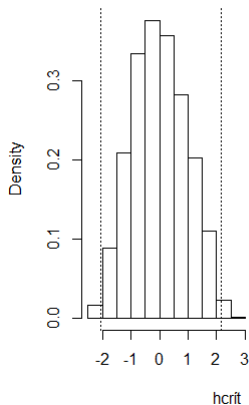
- Los resultados obtenidos por ambos métodos son similares.
- Cuando aumenta el tamaño n de la muestra, aumenta la potencia de la prueba.

Curvas de potencia del estadístico h para la distribución sesgada.

Para $n = 3$ línea segmentada, para $n = 6$ línea continua.



Distribución muestral bootstrap para la distribución sesgada.



Incumplimiento de la hipótesis de repetibilidad para la distribución normal

L	n	Estadístico k Bootstrap							Estadístico k de Mandel)						
		(0)	(0.5)	(1)	(1.5)	(2)	(2.5)	(3)	(0)	(0.5)	(1)	(1.5)	(2)	(2.5)	(3)
5	3	0.013	0.082	0.206	0.320	0.420	0.529	0.609	0.006	0.070	0.187	0.298	0.410	0.518	0.595
	6	0.016	0.183	0.483	0.683	0.791	0.902	0.932	0.010	0.144	0.443	0.650	0.778	0.898	0.928
10	3	0.021	0.097	0.263	0.408	0.572	0.651	0.705	0.013	0.083	0.238	0.387	0.537	0.633	0.684
	6	0.020	0.234	0.564	0.756	0.880	0.923	0.954	0.009	0.190	0.513	0.717	0.861	0.913	0.946

- El método bootstrap detecta mejor los valores atípicos que el método tradicional.
- Cuando aumenta el tamaño n de la muestra, aumenta la potencia de la prueba.

Incumplimiento de la hipótesis de repetibilidad para la distribución de Laplace

L	n	Estadístico k Bootstrap							Estadístico k de Mandel)						
		(0)	(0.5)	(1)	(1.5)	(2)	(2.5)	(3)	(0)	(0.5)	(1)	(1.5)	(2)	(2.5)	(3)
5	3	0.039	0.124	0.199	0.303	0.383	0.436	0.544	0.034	0.176	0.387	0.590	0.686	0.793	0.852
	6	0.034	0.176	0.387	0.590	0.686	0.793	0.852	0.028	0.165	0.390	0.592	0.688	0.795	0.856
10	3	0.023	0.121	0.250	0.361	0.435	0.525	0.596	0.021	0.122	0.250	0.364	0.438	0.525	0.593
	6	0.064	0.249	0.436	0.615	0.773	0.833	0.901	0.059	0.238	0.427	0.605	0.762	0.829	0.898

- Los resultados obtenidos son similares .
- Cuando aumenta el tamaño n de la muestra, aumenta la potencia de la prueba.

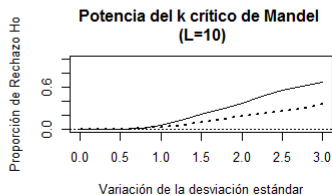
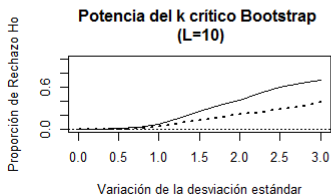
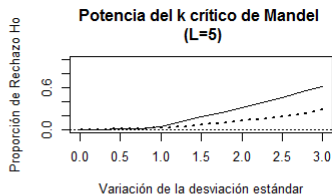
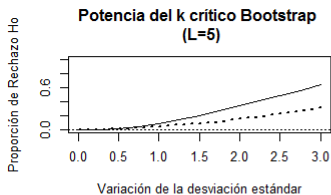
Incumplimiento de la hipótesis de repetibilidad para la distribución sesgada a la derecha

L	n	Estadístico k Bootstrap							Estadístico k de Mandel)						
		(0)	(0.5)	(1)	(1.5)	(2)	(2.5)	(3)	(0)	(0.5)	(1)	(1.5)	(2)	(2.5)	(3)
5	3	0.005	0.014	0.046	0.088	0.150	0.221	0.317	0.001	0.008	0.033	0.065	0.124	0.184	0.288
	6	0.003	0.012	0.061	0.216	0.334	0.492	0.631	0.001	0.007	0.047	0.177	0.308	0.458	0.612
10	3	0.002	0.012	0.051	0.135	0.215	0.288	0.384	0.002	0.008	0.033	0.103	0.192	0.266	0.356
	6	0.002	0.018	0.083	0.258	0.425	0.601	0.705	0.001	0.013	0.060	0.214	0.382	0.560	0.670

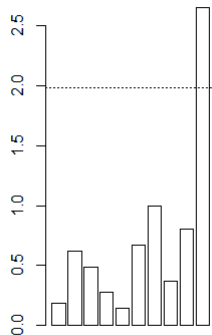
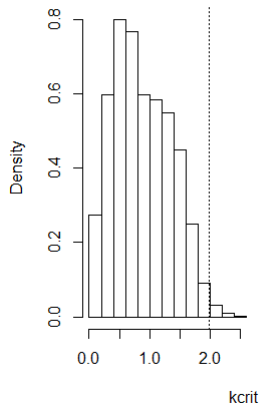
- El método bootstrap detecta mejor los valores atípicos que el método tradicional.
- Cuando aumenta el tamaño n de la muestra, aumenta la potencia de la prueba.

Curvas de potencia del estadístico k para la distribución sesgada.

Para $n = 3$ línea segmentada, para $n = 6$ línea continua.



Distribución muestral bootstrap para la distribución sesgada.



Estudio Interlaboratorio para evaluar la precisión del método empleado para realizar exámenes sanguíneos de Biometría Hemática

- El Estudio Interlaboratorio que se llevó a cabo tuvo como objetivo determinar la precisión del método de medición utilizado para realizar exámenes sanguíneos mediante técnica computarizada, siguiendo las políticas y procedimientos establecidos en las normas ASTM E-691 e ISO 5725.
- Para este estudio hemos seleccionado las características comunes que fueron consideradas por los nueve laboratorios que participaron del Estudio Interlaboratorio, con lo cual se conformaron 11 niveles de medición, o materiales, por valorar dentro de la misma muestra.

Aplicación de la metodología propuesta

Los 11 niveles evaluados durante el Estudio Interlaboratorio fueron:

- (A) nivel de Hemoglobina.
- (B) Nivel de Hematocrito.
- (C) Cantidad de Glóbulos Rojos.
- (D) Cantidad de Glóbulos Blancos.
- (E) Porcentaje de Neutrófilos.
- (F) Porcentaje de Linfocitos.
- (G) Porcentaje de Células Intermedias.
- (H) Hemoglobina globular media.
- (I) Concentración media de hemoglobina modular.
- (J) Volumen globular medio.
- (K) Cantidad de Plaquetas.

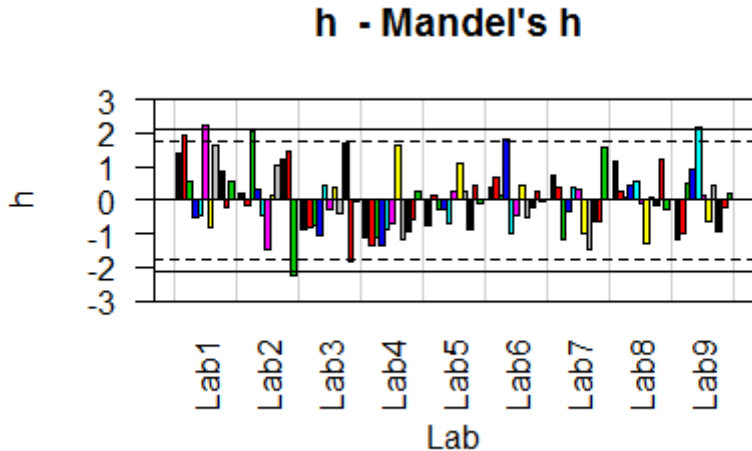
Estadísticos h de Mandel calculados mediante la metodología tradicional.

Los cálculos se realizaron a través del paquete metRology.

Laboratorio	Material										
	A	B	C	D	E	F	G	H	I	J	K
Lab1	1.38	1.94	0.57	-0.50	-0.48	2.23	-0.80	1.64	0.85	-0.20	0.54
Lab2	0.23	-0.18	2.04	0.31	-0.46	-1.44	0.16	1.04	1.20	1.47	-2.22
Lab3	-0.88	-0.82	-0.77	-1.07	0.41	-0.29	0.38	-0.37	1.70	-1.81	-0.04
Lab4	-1.10	-1.37	-1.10	-1.35	-0.90	-0.68	1.63	-1.15	-0.91	-0.58	0.27
Lab5	-0.76	0.16	-0.30	-0.29	-0.67	0.24	1.11	0.28	-0.88	0.42	-0.09
Lab6	0.38	0.67	0.16	1.83	-1.01	-0.44	0.47	-0.54	-0.23	0.26	-0.03
Lab7	0.76	0.37	-1.16	-0.36	0.40	0.34	-0.98	-1.48	-0.65	-0.61	1.60
Lab8	1.17	0.24	0.06	0.47	0.54	-0.09	-1.31	0.11	-0.15	1.24	-0.27
Lab9	-1.18	-0.99	0.49	0.95	2.16	0.13	-0.65	0.46	-0.93	-0.19	0.23

Diagrama de barras de los estadísticos h de Mandel

El diagrama fue realizado por el paquete metRology.



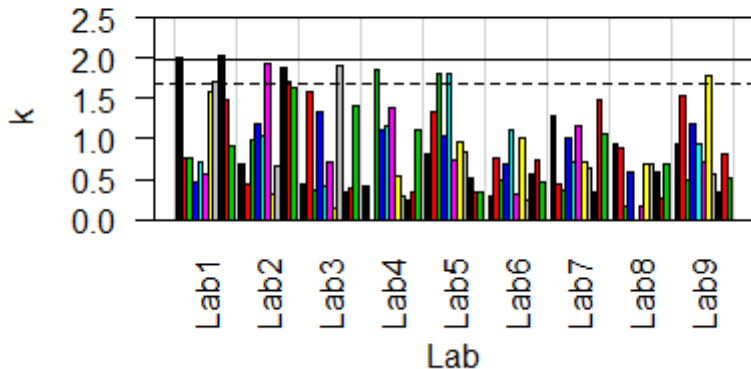
Estadísticos k de Mandel calculados mediante la metodología tradicional.

Los cálculos se realizaron a través del paquete metRology.

Laboratorio	Material										
	A	B	C	D	E	F	G	H	I	J	K
Lab1	2.01	0.77	0.76	0.45	0.70	0.56	1.59	1.70	2.04	1.49	0.91
Lab2	0.68	0.44	0.98	1.19	1.05	1.94	0.31	0.67	1.87	1.71	1.64
Lab3	0.43	1.59	0.37	1.33	0.41	0.70	0.14	1.90	0.34	0.39	1.42
Lab4	0.41	0.00	1.85	1.12	1.16	1.38	0.54	0.28	0.24	0.35	1.20
Lab5	0.80	1.32	1.81	1.04	1.81	0.74	0.95	0.85	0.52	0.35	0.35
Lab6	0.28	0.77	0.49	0.68	1.12	0.32	1.01	0.24	0.56	0.73	0.47
Lab7	1.27	0.44	0.37	1.02	0.70	1.16	0.70	0.65	0.34	1.48	1.07
Lab8	0.94	0.88	0.17	0.60	0.00	0.16	0.68	0.69	0.58	0.26	0.69
Lab9	0.94	1.53	0.49	1.19	0.93	0.70	1.79	0.56	0.34	0.80	0.51

Diagrama de barras de los estadísticos k de Mandel

k - Mandel's k



El

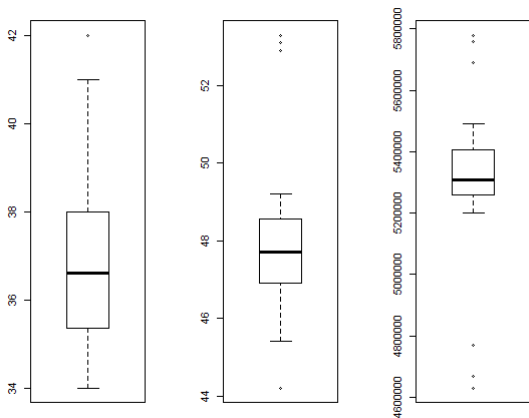
diagrama fue realizado por el paquete metRology.

Resumen de resultados

- Para el estadístico h vemos que los laboratorios 1, 2 y 9 presentan valores atípicos en los niveles F, K y E respectivamente. Mientras que el laboratorio cuatro muestra un sesgo negativo en los resultados, lo que debería ser investigado y reportado al laboratorio respectivo.
- Para el estadístico k Podemos ver que el laboratorio 1 presenta datos atípicos en los niveles A e I, por lo que este laboratorio debería ser investigado.

Aplicación de la metodología bootstrap

Para el material E se encontró un dato atípico, en el material F se encontraron cuatro datos atípicos, y en los resultados del material K se encontraron seis datos atípicos

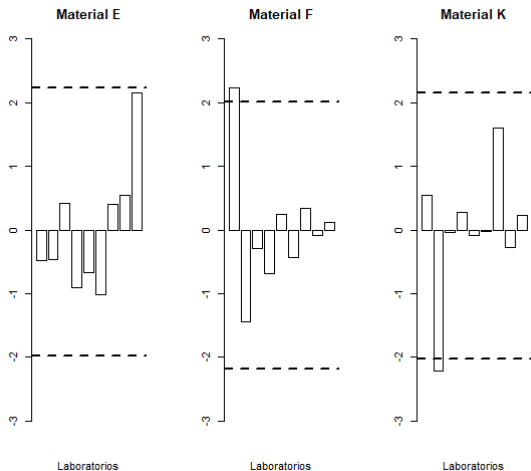


Valores críticos h bootstrap

Se han realizado $B=1000$ remuestras sobre los datos del ILS.

Nivel	Límite inferior	Límite superior
A	-2.02	2.15
B	-2.04	2.11
C	-1.96	2.20
D	-2.06	2.14
E	-1.97	2.24
F	-2.19	2.04
G	-2.06	2.14
H	-2.12	2.10
I	-1.94	2.16
J	-2.17	2.11
K	-2.02	2.15

El primer laboratorio para el nivel F, y el segundo laboratorio para el nivel K, fueron marcados como laboratorios con datos atípicos

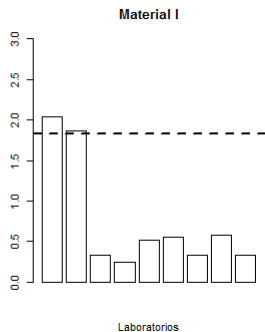
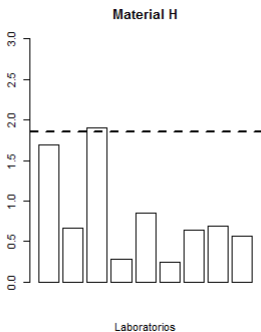
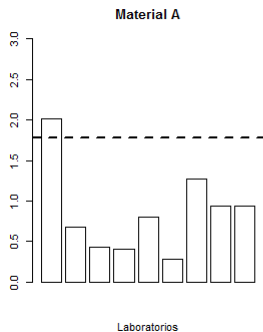


Valores críticos k bootstrap

Se han realizado $B=1000$ remuestras sobre los datos del ILS.

Nivel	Límite superior
A	1.78
B	1.94
C	1.97
D	1.90
E	2.01
F	1.96
G	1.80
H	1.86
I	1.83
J	1.92
K	1.87

Se detectarán como atípicos el primer laboratorio para el nivel K, el tercer laboratorio para el nivel H, y el primer y segundo laboratorio para el nivel I



Conclusiones

Las distribuciones de probabilidad de los estadísticos h y k se estimaron mediante el procedimiento bootstrap propuesto, con el cual fue posible obtener los valores h críticos y k críticos como los cuantiles de las correspondientes distribuciones bootstrap de los estadísticos para un nivel de significación de $\alpha = 0.01$.

Conclusiones

Para obtener la distribución bootstrap de los estadísticos h y k fue necesario realizar remuestreo a partir de la hipótesis nula de que todos los laboratorios provienen de la misma población.

Conclusiones

Durante el estudio de simulación, se comprobó que los laboratorios que fueron simulados como atípicos fueron detectados exitosamente mediante la metodología propuesta.

GRACIAS



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA