Estimación de parámetros para indicador TBF
Software RMES

Centro de Desarrollo de Gestión Empresarial.
1 Oriente 1097, Viña del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079
empresa@mes.cl

Diciembre, 2010

Resumen

El siguiente artículo explicará los algoritmos referentes a la estimación de parámetros para la distribución TBF utilizados en el software RMES. Tiene como objetivo documentar exhaustivamente el proceso probabilístico de cálculo del indicador TBF para equipos y no poseerá lenguage técnico informático asociado a la implementación de los algoritmos. Con ello se pretende dejar en claro exactamente lo que ocurre dentro del software, de forma que sea posible evaluar en un futuro su correctitud. Se asume por parte del lector un conocimiento básico del software RMES junto con la jerga asociada y tecnicismos relativos a Ingeniería en Mantención.

Índice

1. Introducción **3**

2. Ajuste con Censura a la Derecha **4**

3. Estimación de Parámetros **7**
3.1. Estimación Exponencial [1] 8
3.2. Estimación Weibull [2] 9

4. Test de Hipótesis **10**
4.1. ChiCuadrado [3] 12
4.2. Kolmogorov-Smirnov [4] 15
4.3. Método Adaptativo 18
4.4. Valor P para parámetros estimados [6] 19

5. Conclusiones **20**

1. Introducción

El objeto más básico en el software RMES es el equipo, que en sí, representa la unidad más básica existente en una planta. Por supuesto, esta unidad no necesariamente debe ser un equipo, puede ser algún componente de mucho menos nivel, dependiendo de la especifidad que se necesite o, simplemente, de la información que se disponga.

La Ingenieria en Mantención, en muy simples palabras y de forma muy burda, tratará de estudiar el comportamiento de este componente y los conjuntos al que pertenece mediante el análisis de las detenciones que este objeto contenga, entendiéndose como detención un no funcionamiento del equipo durante cierto periodo de tiempo. Las detenciones se grafican en un cronograma como lo muestra la figura 1.

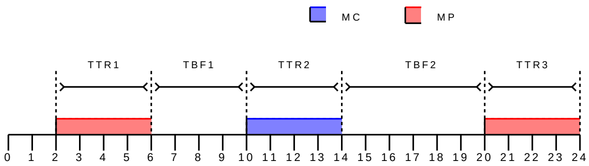

Figura 1: Cronograma

La figura 1 muestra 2 indicadores muy importantes:

- TTR (Time To Repair): Tiempo para Reparar, que es el tiempo que el componente estuvo detenido.
- TBF (Time Between Failures): Tiempo entre Fallas, que es el tiempo existente entre el fin de una falla y el inicio de la siguiente.

Como en muchos procesos, es posible estimar el comportamiento de cualquier indicador utilizando fórmulas estadísticas. Particularmente, este documento se enfocará a explicar el cómo RMES estima el comportamiento del indicador TBF.

2. Ajuste con Censura a la Derecha

Las detenciones pueden tener distintos tipos:

- *Mantención*: parada del equipo.
  - *Correctivas* (MC): parada por falla inadvertida.
  - *Preventivas* (MP): parada preventiva para mantención.
- *Detención*: parada de planta, ajena al equipo.
  - *Operativas* (DO): parada por falla inadvertida.
  - *No Operativas* (DONP): parada preventiva para mantención.

Dado que las detenciones de tipo DO o DONP no reflejan un comportamiento propio del equipo, no se tomarán en cuenta. Entonces, sólo se trabajará con detenciones MC y MP.

En sí, sólo las detenciones MC proporcionan información directa del comportamiento del equipo, dado que las detenciones MP son paradas provocadas intencionalmente. El algoritmo denominado **Ajuste de Curva con información censurada a la derecha** consiste en incorporar la información de las detenciones MP al ajuste. Este es el algoritmo utilizado por el software RMES.

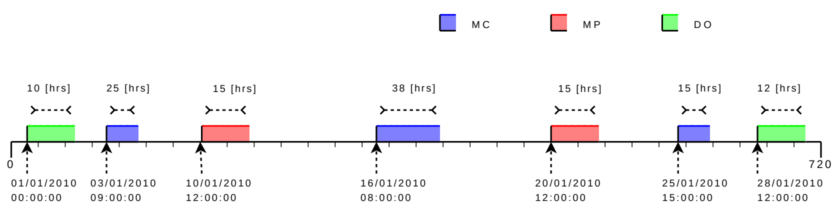

Figura 2: Diagrama para ejemplo de ajuste

La figura 2 representa un ejemplo a desarrollar a lo largo del documento para explicar los algoritmos. Como se puede observar, el ejemplo contiene detenciones MC, MP y DO. En RMES no existe un registro del inicio de funcionamiento del equipo (información que permitiría obtener el primer TBF antes de la primera falla de tipo MC/MP). Intuitivamente, lo mejor que podemos hacer es que nuestro algoritmo comience a obtener los TBF desde el fin de la primera falla de tipo MC/MP. La figura 3 muestra el primer paso a realizar, el cual es quitar todas las detenciones que no sean MC/MP.

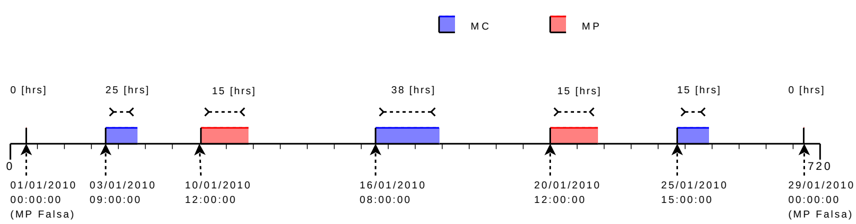

Figura 3: Ejemplo sin DO's y con MP's Falsas

Si comparamos la figura 2 con la figura 3 podemos observar que la primera detención de tipo DO fue eliminada, al igual que la última. Esto implica una perdida de información, debido que al menos sabemos que el equipo funcionó correctamente desde el **inicio** de la primera DO hasta el **inicio** de la primera MC, y también desde el **fin** de la última MC hasta el **fin** de la última DO. Para soslayar este problema, se agregan 2 tuplas falsas de tipo MP con duración cero a la lista de detenciones originales bajo las condiciones:

1. La primera tupla falsa se agregará al inicio de la primera detención original sólo si ésta no es MC/MP.
2. La segunda tupla falsa se agregará al final de la última detención original sólo si ésta no es MC/MP.

Estas condiciones se deben a que si se las tuplas falsas agregan ante un detenciones de tipo MC/MP, el TBF entre la tupla falsa y la tupla (o viceversa) será cero, lo que implica una perturbación negativa sobre el comportamiento de los TBF's. La figura 3 muestra la ubicación de las tuplas falsas luego de la eliminación de las detenciones de tipo DO.

A continuación, se calculan los TBF's entre fallas consecutivas. Se verifica también si de entre las 2 detenciones (de donde se calcula el TBF) la última es o no MC. La tabla 1 muestra el cálculo de los TBF's para las tuplas representados por la figura 3.

Fecha	Hora	Duración [hrs]	Tipo	TBF [hrs]	Es MC
01/01/2010	00:00:00	0	MP	-	-
03/01/2010	09:00:00	25	MC	57	SI
10/01/2010	12:00:00	15	MP	146	NO
16/01/2010	08:00:00	38	MC	125	SI
20/01/2010	12:00:00	15	MP	62	NO
25/01/2010	15:00:00	15	MC	108	SI
29/01/2010	00:00:00	0	MP	66	NO

Cuadro 1: Tabla de Resultados Parciales para los TBF's

Es importante destacar que, luego de calcular el TBF entre 2 fallas, este tiempo se multiplica por un valor llamado "ratio", que representa la relación entre cantidad de Horas de Operación Anual dividido la cantidad de horas en un año (8760).

$$ratio = \frac{\textbf{Horas de operacion anual}}{\textbf{Horas en un año}}$$

Para nuestro ejemplo, el valor de "ratio" es igual a 1.

Luego, los TBF's se ordenan de menor a mayor y se les asigna un índice correlativo, comenzando desde 1. La tabla 2 presenta el resultado parcial hasta ahora.

Indice	TBF [hrs]	Es MC
1	57	SI
2	62	NO
3	66	NO
4	108	SI
5	125	SI
6	146	NO

Cuadro 2: Tabla de Resultados Parciales para los TBF's

1 Oriente 1097, Vina del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079

El siguiente paso es crear un ranking ajustado que se utilizará para realizar la estimación de parámetros. Este ranking se construye con la siguiente fórmula:

$$r_0 = 0$$
$$r_i = r_{i-1} + \frac{(n+1) - r_{i-1}}{(n+1) - (i-1)} \qquad \forall_{i=1,\ldots,n}$$

donde r_i es el valor del ranking, r_{i-1} el valor del ranking anterior, n es la cantidad de TBF's e i es el índice correlativo.

En este ranking sólo se consideran aquellos TBF's asociados a tipos MC, que desde ahora, serán conocidos como los valores x's. Esto quiere decir que el valor r_i no se calculará si el tipo asociado al TBF es MP. De esta forma, las detenciones MP modifican el ranking final que las detenciones MP tendrán.

Para nuestro ejemplo, $n = 6$, entonces, la fórmula quedaría:

$$r_i = r_{i-1} + \frac{7 - r_{i-1}}{7 - (i-1)} \qquad \forall_{i=1,\ldots,6}$$

La tabla 3 presenta el resultado parcial hasta ahora en base a la fórmula personalizada.

Indice (i)	TBF [hrs]	Es MC	x	$i-1$	r_{i-1}	r_i
1	57	SI	57	0	0	1
2	62	NO	-	-	-	-
3	66	NO	-	-	-	-
4	108	SI	108	3	1	2.5
5	125	SI	125	4	2.5	4
6	146	NO	-	-	-	-

Cuadro 3: Tabla de Resultados Parciales para los TBF's y su ranking

La tabla 4 presenta el resultado final del algoritmo.

r	x
1	57
2.5	108
4	125

Cuadro 4: Tabla de Resultado Final para algoritmo con Censura a la Derecha

3. Estimación de Parámetros

Luego de tener los resultados generados por el algoritmo de **Ajuste con Censura a a la Derecha** se comienza con el fitteo de las curvas. RMES considera 3 casos posibles según la cantidad de datos que este algoritmo entrege:

1. Si se generaron entre 0 y 2 datos, no se realiza estimación dado que se posee muy poca información.
2. Si se generaron entre 3 y 20 datos, se estiman los parámetros para una Exponencial y una Weibull y se decide cual de las dos distribuciones se ajusta mejor a los datos utilizando Test de Hipótesis Kolmogorov-Smirnov.
3. Si se generaron más de 20 datos, se estiman los parámetros para una Exponencial y una Weibull y se decide cual de las dos distribuciones se ajusta mejor a los datos utilizando Test de Hipótesis ChiCuadrado.

1 Oriente 1097, Vina del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079

3.1. Estimación Exponencial [1]

Partiremos desde la tabla 4 que contiene el ranking y los datos asociados. Según lo implementado en RMES, el algoritmo de estimación de parametros en base al ranking para una distribución exponencial es el siguiente:

$$y_i = 1 - \frac{r_i - 0{,}3}{n + 0{,}4}$$
$$\lambda = -\frac{\sum_i \ln(y_i) \cdot x_i}{\sum_i x_i^2}$$

Donde x_i es el dato i-ésimo de los TBF's asociados a MC y n es la cantidad total de **detenciones** que se consideraron para el ajuste, o sea, los asociados a MC's y MP's, en nuestro caso particular, $n = 7$, resultado que se puede obtener contando las detenciones en la tabla **??** (las detenciones falsas también se consideran). La tabla 5 resume los cálculos para el ajuste exponencial, de los cuales podemos desprender el valor del parámetro λ para la exponencial.

x	r	y_i	$ln(y_i)$	x_i^2	$ln(y_i) \cdot x_i$
57	1.0	0.9054	-0.099372	3249.0	-5.6642
108	2.5	0.7027	-0.352821	11664.0	-38.1047
125	4.0	0.5	-0.693147	15625.0	-86.6434
				30538.0	-130.4123

Cuadro 5: Ajuste Exponencial

$$\lambda = -\frac{-130{,}4123}{30538{,}0} = 0{,}00427$$

1 Oriente 1097, Vina del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079

3.2. Estimación Weibull [2]

La tabla 4 también nos será útil para el desarrollo del ejemplo. En este caso, en necesario filtrar todos los valores menores o iguales a 0. Dado que en nuestro caso, esto no ocurre, este filtro es inncesario. Los parámetros α y β se calculan con las siguientes ecuaciones:

$$y_i = 1 - \frac{r_i - 0{,}3}{n + 0{,}4}$$

$$v_i = \ln(x)$$
$$w_i = \ln(-\ln(y_i))$$

$$o = m \cdot \sum_i v_i \cdot w_i - \sum_i v_i \cdot \sum_i w_i$$
$$p = \sum_i w_i \cdot \sum_i v_i^2 - \sum_i v_i \cdot \sum_i v_i \cdot w_i$$
$$q = m \cdot \sum_i v_i^2 - (\sum_i v_i)^2$$
$$r = \frac{o}{q}$$
$$s = \frac{p}{q}$$

$$\beta = r$$
$$\alpha = e^{-\frac{s}{\beta}}$$

donde m es la cantidad de datos que quedaron después del filtro. En nuestro caso particular, $m = 3$.

x	r	v	w	v_i^2	$v \cdot w$
57	1.0	4.0430	-2.3088	16.3462	-9.3349
108	2.5	4.6821	-1.0417	21.9223	-4.8778
125	4.0	4.8283	-0.3665	23.3126	-1.7696
		13.5534	-3.7172	61.5812	-15.9824

Cuadro 6: Ajuste Weibull

De la tabla 6 se pueden obtener los valores de $\sum_i v_i$, $\sum_i w_i$, $\sum_i v_i \cdot w_i$ y $\sum_i v_i^2$ para el cálculo de las variables o, p y q.

$$o = 3 \cdot -15{,}9824 - 13{,}5534 \cdot -3{,}7172 = 2{,}4337$$
$$p = -3{,}7172 \cdot 61{,}5812 - 13{,}5534 \cdot -15{,}9824 = -12{,}2918$$
$$q = 3 \cdot 61{,}5812 - (13{,}5534)^2 = 1{,}0464$$

Con estos valores, podemos calcular r y s.

$$r = \frac{2{,}4337}{1{,}0464} = 2{,}3257$$

$$s = \frac{-12{,}2918}{1{,}0464} = -11{,}7464$$

Y finalmente, podemos calcular primero β y luego α.

$$\beta = 2{,}3257$$

$$\alpha = e^{-\frac{-11{,}7464}{2{,}3257}} = 156{,}1139$$

1 Oriente 1097, Vina del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079

4. Test de Hipótesis

Muchos métodos estadísticos asumen un distribución subyacente en la derivación de sus resultados. Tales métodos se denominan *paramétricos* (aquellos que no asumen una distribución subyacente en los datos se denominan *no paramétricos*).

Existen procedimientos formales para evaluar la correctitud de la distribución elegida y estos son los Test de Hipótesis, basados en la Teoría Estadística, los cuales generalmente son numéricamente complejos y requieren de algún tipo de soporte computacional.

Un Test de Hipótesis es, en simples palabras, una "regla", una herramienta que permite medir la "distancia" que existe entre un conjunto de datos y una distribución candidata. El concepto de "distancia" lo define propiamente cada Test de Hipótesis, es así como el test ChiCuadrado utiliza su método particular para calcular distancia que difiere de Kolmogorov-Smirnov. Esta "distancia" formalmente se conoce como TSV (Test Statistical Value, o Valor Estadístico del Test).

Para saber cuán bueno o cuán malo es nuestro TSV existen las distribuciones que rigen el comportamiento del estadístico. Por ejemplo, para el test ChiCuadrado, la distribución que rige al estadístico es una ChiCuadrado (de ahí su nombre). En sí, lo único que nos dice el test es si nuestro TSV es un valor cómun (probable) dentro de la distribución del estadístico.

Pensar por ejemplo en el hecho de intentar clasificar a un animal en base a su altura. Respecto a los seres humanos, mientras mayor sea la altura del animal que estemos midiendo (TSV), diremos con mayor certeza que el animal no es un ser humano. Al contrario, si altura es pequeña, podremos decir que el animal tiene una probabilidad de ser un humano, o de ser un mono, o de ser un perro, etc. En este caso, el test simplemente nos dice, respecto a la altura, si se puede o no clasificar al animal dentro de la categoría. Es en este punto es donde tiene sentido el término Valor Crítico (CV, Critical Value), asociado otro término muy importante que es el Nivel de Significancia α.

El Nivel de Significancia α permite obtener un CV para el test en particular que se este realizando. Supongamos que $D(x)$ es la función de distribucion acumulada CDF del estadístico de distancia de nuestro test. La relacion entre CV y α es la siguiente:

$$D(X > CV) = \alpha$$

Es decir, la probabilidad de que un número se encuentre entre $[CV, \infty]$ es α. Si por ejemplo nuestro TSV se encuentra en este intervalo, quiere decir que es poco probable que lo queramos verificar sea verdadero.

El concepto asociado a "lo que se quiere verificar" se conoce como Hipótesis Nula H_0. La negación de la Hipótesis Nula se conoce como Hipótesis Alternativa H_1. Siguiendo con nuestro ejemplo, queremos saber si el animal es humano o no, entonces diremos que:

H_0: el animal es un ser humano.

H_1: el animal NO es un ser humano.

Podemos decir sin ser estrictos que un valor crítico para la altura de un ser humano es de 2.05 [m] con un nivel de significancia de 0.05. Si nuestro animal mide más de tal altura, es poco probable que sea un ser humano. Se cumple que:

1 Oriente 1097, Vina del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079

$$D_{Humano}(X > 2{,}05) = 0{,}05$$

Para la categoría de monos, diremos que:

H_0: el animal es un mono.

H_1: el animal NO es un mono.

Un valor critico podría ser 1.40 [m] con el mismo Nivel de Significancia de 0.05. Si nuestro animal mide más de tal altura, es poco probable que sea un mono. Se cumple que:

$$D_{Mono}(X > 1{,}40) = 0{,}05$$

Supongamos ahora que el animal mide 1.30[m], aplicando Test de Hipótesis para los seres humanos y para los monos ambos nos dirían que las respectivas H_0 son verdaderas, o sea, que el animal puede ser tanto como un ser humano o como un mono. Finalmente, es en este punto es donde nace el último término importante denominado Valor P (P-Value, en inglés).

El Valor P es la probabilidad que existe entre el intervalo $[TSV, \infty]$. Si observamos la figura 4, podemos ver que el área abarcada para la distribución asociada a los monos es mayor que el área asociada a la distribución de los seres humanos. O sea:

$$D_{Mono}(X > 1{,}30) > D_{Humano}(X > 1{,}30)$$
$$Pvalue_{Mono} > Pvalue_{Humano}$$

Diremos finalmente que es más probable que el animal se trate de un mono que de un ser humano. Por supuesto, podemos equivocarse y errar en el supuesto H_0, pero es lo mejor que se puede hacer en base a la información que se tiene.

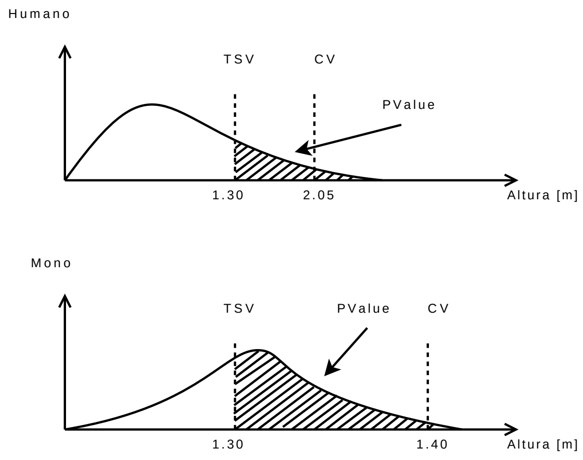

Figura 4: Conceptualización de un Test de Hipótesis

4.1. ChiCuadrado [3]

El test de hipótesis ChiCuadrado cabe dentro de los procedimientos denomidados "tests de área", dado que utiliza la función de densidad (PDF) en el proceso.

Para evaluar los datos, existe un esquema bien definido. Primero, se asume que los datos provienen de cierta distribución (i.e: Normal). Luego, se estiman los parámetros de la distribución candidata (i.e: media y varianza) o simplemente se obtienen apriori de la experiencia. Tal proceso lleva a crear una hipótesis "compuesta" de distribución, dado que varios elementos juntos deben ser verdaderos, la cual se denomina Hipótesis Nula (H_0), la negación de que los datos no se explican con distribución candidata se conoce como Hipótesis Alternativa (H_1). Luego, se testea nuestra distribución usando el conjunto de datos. Finalmente, H_0 es rechazada si alguno (o varios) de los elementos en la hipótesis H_0 no son soportados por los datos.

El test ChiCuadrado esta basado conceptualmente sobre la PDF de la distribución candidata. Si la distribución es correcta, ésta debería abarcar el rango del conjunto de datos de forma muy precisa (pensar en datos que se distribuyen como una normal N(0,1) y nosotros asumimos una normal N(1000,1000), claramente, los rangos no coinciden). De este rango, se eligen convenientemente valores que lo dividan en varios subintervalos. Luego, se calcula el número de puntos en cada intervalo (al igual que en un histograma de frecuencia absoluta). Estos puntos son denominados valores "observados". Luego, se calcula la cantidad de puntos que deberían haber caído en esos mismos subintervalos, de acuerdo a la PDF de la distribución candidata. Éstos son denominados valores "esperados" y el Test ChiCuadrado requiere de al menos 5 de ellos en cada subintervalo. Finalmente, se comparan estos 2 resultados. Si ellos concuerdan (probabilísticamente), entonces los datos soportan la distribución candidata. De lo contrario, se rechaza la suposición. La fórmula (estadístico o indicador que sigue algún comportamiento) utilizada para saber la diferencia entre los valores "esperados" y "observados" sigue una distribución ChiCuadrado. De ahí el nombre del test.

La fórmula para el estadístico ChiCuadrado es la siguiente:

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} \sim \chi^2_{k-1-nep}$$

donde

e_i	número de datos esperados en el subintervalo i ($e_i \geq 5$)
o_i	número de datos observados en el subintervalo i
k	total de subintervalos en el rango
n	total de datos para la implementación del Test ChiCuadrado ($n \geq 5 \cdot k$)
nep	número estimado de parámetros (depende de la distribución)
$k - 1 - nep$	grados de libertad o DF para la distribución ChiCuadrado ($DF > 0$)

De todo esto, se pueden desprender algunas observaciones:

1. Por ejemplo para una exponencial, la cantidad de subintervalos mínima k a escoger debe ser de 3, según la formula de DF donde $DF = k - 1 - neq > 0$.
2. Luego, teniendo el número de intervalos k, se debe cumplir también que $n >= 5 \cdot k$. O sea, que para nuestra misma exponecial deberíamos

tener al menos 15 datos.

3. Para una Weibull, el k mínimo es 4, por ende, la cantidad de datos mínima es de 20.
4. Como conclusión, este test necesita de varios datos para realizar el procedimiento.
5. La elección de los puntos que definen el inicio y el fin de los subintervalos es aparentemente arbitraria. No se exige que los subintervalo sean iguales en tamaño de rango, sólo se pide que en cada uno de ellos al menos la cantidad de valores esperados sea mayor que 5. Además, estos puntos no necesariamente tiene que pertenecer al conjunto de datos. El test necesita de cierto criterio para la elección de subintervalos basado en la PDF/CDF de la distribución candidata.

Para claridad del test, se presentará el siguiente ejemplo. La tabla 7 presenta datos ordenados que deseamos contrastar con una distribución Normal que asumimos como candidata. La tabla 8 resume algunos indicadores para los datos.

6.1448	6.6921	6.7158	7.7342	9.6818	12.3317	12.5535	13.0973	13.6704
14.0077	14.7975	15.3237	15.5832	15.7808	15.7851	16.2981	16.3317	16.8147
16.8860	17.5166	17.5449	17.9186	18.5573	18.8089	19.2541	19.5172	19.7322
21.9602	23.2046	23.2625	23.7064	23.9296	24.8702	25.2669	26.1908	26.9989
27.4122	27.7297	28.0116	28.2206	28.5598	29.5209	30.0080	31.2306	32.5446

Cuadro 7: Datos para Test de Bondad de Ajuste Normal

Variable	N	Media	Mediana	StDev	Min	Max	Q1	Q3
Valor	45	19.50	18.56	7.05	6.14	32.54	15.06	25.73

Cuadro 8: Estadísticas Descriptivas para datos Normales

Según la tabla 8, estimamos una distribución Normal $N(19{,}5; 7{,}05)$. Dado que la distribución normal utiliza 2 parámetros, podemos calcular los grados de libertad de nuestro test ChiCuadrado como $DF = k - 1 - 2$. Podemos elegir sin riesgo un valor para $k = 5$, dado que $DF > 0$ ($DF = 2$) y se cumple que $n \geq 5 \cdot k$ ($45 \geq 15$).

Luego, seleccionaremos arbitrariamente los puntos finales para cada subintervalo: 14, 17, 22 y 26. Ellos definirán 5 subintervalos o celdas, donde cada una contendrá más de 5 valores esperados.

La tabla 9 muestra los resultados intermedios del algoritmo. La primera columna muestra los puntos finales de cada subintervalo. La segunda, muestra el punto final estandarizado. La tercera, muestra sus valores acumulados (obtenidos desde una tabla Normal común y corriente). Por ejemplo, para el primer punto final (14), su valor estandarizado será:

$$StdEnd_1 = \frac{14 - 19{,}5}{7{,}05} = -0{,}78014$$

Y su valor acumulado (la probabilidad entre $[-\infty, -0{,}78014]$) será:

$$CumProb_1 = N_{19{,}5;7{,}05}(-0{,}78014) = 0{,}2176$$

En la cuarta columna se muestra la probabilidad del área de la celda, la cual se calcula restando las probabilidades acumuladas de la columna anterior inmediatamente adyacentes. Por ejemplo, $CellProb_1 = 0{,}217654 - 0 =$

1 Oriente 1097, Vina del Mar, Chile.
Fono:(56) (32)688987, Fax:(56) (32)2684079

End	StdEnd	CumProb	Cellprob	Expected	Observed	$(e-o)^2/e$
14	-0.78014	0.217654	0.217654	9.7944	9	0.06443
17	-0.35461	0.361441	0.143787	6.4704	10	1.92535
22	0.35461	0.638559	0.277118	12.4703	9	0.96574
26	0.92199	0.821732	0.183173	8.2428	6	0.61024
∞	∞	0.999999	0.178300	8.0235	11	1.10420
		Totales	1.000000	45.0010	45	4.67000

Cuadro 9: Estadísticas Descriptivas para datos Normales

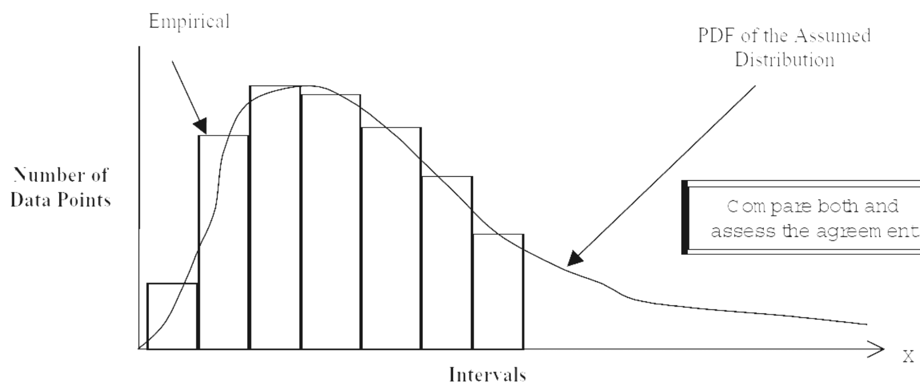

Figura 5: Aproximación Conceptual del Test ChiCuadrado

0,217654, $CellProb_2 = 0{,}361441 - 0{,}217654 = 0{,}143787$, y así con el resto. La quinta columna contiene la multiplicación del área de la celda por la cantidad de datos (45) dando a lugar el valor esperado para cada celda. La sexta columna es el valor observado que se puede obtener directamente de la tabla 7 y la última columna presenta el estadístico parcial para cada celda. La sumatoria de esta columna es el estadístico para el test:

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} = 4{,}67$$

Este resultado se debe comparar con el valor crítico para ChiCuadrado. Este se obtiene directamente desde una tabla ChiCuadrado. Con un nivel de significancia α de 0.05 y con 2 grados de libertad el valor crítico es:

$$\chi^2_2(0{,}05) = 5{,}99$$

En vista de que el valor calculado 4,67 es menor que el valor crítico 5,99 se acepta la hipótesis nula y se dice que con un error del 5 % es probable de que los datos se comporten como una $N(19{,}5; 7{,}05)$.

4.2. Kolmogorov-Smirnov [4]

A diferencia del Test ChiCuadrado, este procedimiento se denomina "test de distancia" dado que trabaja con la función de densidad acumulada (CDF). Al igual que el Test Anderson-Darling, es uno de los mejores test en presencia de pocos datos.

Su implementación sigue una serie bien definida de pasos. Primero, se asume una distribución candidata para los datos y se estima sus parámetros. Tal proceso conduce a la creación de la hipótesis nula (H_0), en donde varias de los supuestos de la hipótesis deben ser verdaderos para que ésta sea aceptada. La negación será la hipótesis alternativa (H_1).

En un test de distancia, cuando la distribución candidata es correcta, la CDF teórica (denotada F_0) sigue muy de cerca a la Función de Densidad Acumulada Empírica, una función escalón de la CDF (denotada F_n). Esto se ilustra conceptualmente en la figura 6. Aqui, los datos se encuentran ordenados, $\{X_1 \leq X_2 \leq ... \leq X_n\}$ y la distribución candidata H_0 tiene una CDF, $F_0(x)$. Luego, se obtienen los correspondientes valores estadísticos para el test de bondad de ajuste (GoF). Si concuerdan, probabilísticamente, los datos soportan a la distribución candidata. De lo contrario, se rechaza la distribución.

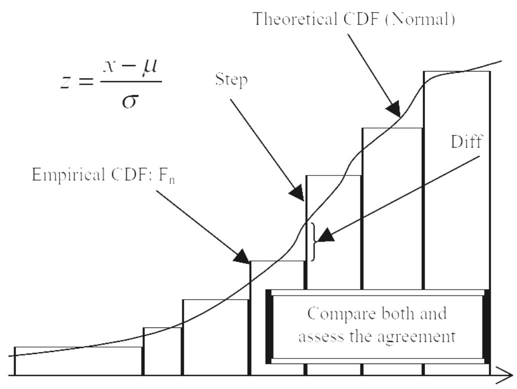

Figura 6: Aproximación Conceptual del Test Kolmogorov-Smirnov

Existen varias formas de realizar el test KS. A continuación se explicará una de ellas. Primero, se ordenan los datos y se obtiene los parámetros de la distribución candidata (H_0). Luego, se obtienen la función CDF teórica (F_0) y la función CDF empírica (F_n).

$F_0(X_i)$ es la función de distribución acumulada (de la distribución candidata) evaluada en el punto X_i y $F_n(X_i)$ es la función de distribuión acumulada empírica obtenida de la proporción de la cantidad de datos menores que X_i en el conjunto de datos de tamaño n. Dado que KS es un test de distancia, necesitamos encontrar la distancia máxima entre las dos funciones $|F_0 - F_n|$. Estas se definen como:

$$F_0(X_i) = P_0(X \leq X_i) = CDF(X_i)$$

$$F_n(X_i) = \frac{\# \ de \ X's \leq X_i}{n} = \frac{i}{n} \qquad i = 1, ..., n$$

Definamos para cada punto X_i:

$$D^+ = F_n - F_0$$

$$D^- = F_0 - F_n$$

.

El estadístico KS se define como:

$$D = Maximo\ de\ todos\ los\ D^+\ y\ D^-\ (\geq 0) \quad i = 1, ..., n$$

La lógica de KS es la siguiente: si la máxima separación entre la CDF candidata y la CDF empírica es pequeña, entonces aceptamos que la distribución candidata explicada al conjunto de datos. De lo contrario, se rechaza.

Para una mejor compresión del test, se utilizará un ejemplo. La tabla 10 contiene seis datos con los que se trabajará durante la explicación del test. La tabla 11 contiene estadísticas de los datos. Se asumirá que los datos siguen una distribución Normal(μ,σ).

294.2	308.5	313.1	317.7	322.7	338.7

Cuadro 10: Datos para Test KS GoF

Variable	N	Media	Mediana	SD
Datos	6	315.82	315.40	14.85

Cuadro 11: Datos para Test KS GoF

Como ejemplo, calcularemos los valores para el primer punto de la tabla 10. Asumiremos que los datos se comportan como una $N(315{,}82, 14{,}85)$. Esta será nuestra distribución candidata.

$$F_0(294{,}2) = N_{315{,}82,14{,}85}\left(\frac{294 - 314{,}82}{14{,}85}\right) = 0{,}0727$$

$$F_n(294{,}2) = \frac{1}{6} = 0{,}167$$

$$F_{n-1}(294{,}2) = \frac{0}{6} = 0$$

Los estadísticos D^+ y D^- serán:

$$D^+(294{,}2) = F_n(294{,}2) - F_0(294{,}2) = 0{,}167 - 0{,}0727 = 0{,}0939$$

$$D^-(294{,}2) = F_0(294{,}2) - F_{n-1}(294{,}2) = 0{,}0727 - 0 = 0{,}0727$$

El resto de los valores se presenta en la tabla 12.

Row	DataSet	F_0	F_n	F_{n-1}	$D^+ = F_n - F_0$	$D^- = F_0 - F_{n-1}$
1	294.1	0.072711	0.16667	0.00000	0.093955	0.072711
2	308.5	0.311031	0.33333	0.16667	0.022302	0.144365
3	313.1	0.427334	0.50000	0.33333	0.007266	0.094001
4	317.7	0.550371	0.66667	0.50000	0.116295	0.050371
5	322.7	0.678425	0.83333	0.66667	0.154908	0.011759
6	338.7	0.938310	1.00000	0.83333	0.061690	0.104977
Max:					0.154908	0.144365

Cuadro 12: Valores intermedios para test de normalidad con KS GoF.

El resultado para el Test estadístico KS GoF para este conjunto de datos y la distribución candidata es:

$$D = Max(D^+, D^-) = 0{,}1549$$

Este resultado se debe comparar con el valor crítico. Este se obtiene directamente desde una tabla Kolmogorov-Smirnov. Con un nivel de significancia α de 0.05 y con una cantidad de datos igual a 6 el valor crítico es:

$$KS_6(0{,}05) = 0{,}521$$

En vista a que el valor calculado 0,1549 es menor que el valor crítico 0,521 se acepta la hipótesis nula y se dice que con un error del 5 % es probable de que los datos se comporten como una $N(315{,}82, 14{,}85)$.

4.3. Método Adaptativo

Las tablas de inferencia de valores críticos (CV) se pueden encontrar en varios textos. Estos CV son válidos cuando los parámetros de la distribución son conocidos. Cuando los parámetros se estiman (como en nuestro caso), estos CV son sólo una aproximación. Un procedimiento adaptativo utiliza un CV para un error $\alpha^{'}$ cuatro veces mayor al que se desea testear.

Veamos la tabla 13, asociada a los valores criticos para un test Kolmogorov-Smirnov. Supongamos que nuestro estadístico KS fue de $D = 0{,}155$ con 6 datos y el nivel de significancia es $\alpha = 0{,}05$. Analizando la tabla, el valor correspondiente al CV será de 0.521 para $n = 6$ con un error $\alpha = 0{,}05$. Si los verdaderos parámetros de la normal hubiesen sido usados, el resultado nos permitiría aceptar la hipótesis nula. Dado que los parámetros son estimados desde los datos, se debe utilizar un método adaptativo de testeo.

n	CV(0.20)	CV(0.05)
4	0.494	0.
5	0.446	0.564
6	0.411	0.521
7	0.381	0.486
8	0.358	0.457
9	0.339	0.432
10	0.322	0.411
Para $n > 10$, usar la aproximación $CV(0{,}05) = 1{,}36/\sqrt{n}$ $CV(0{,}20) = 1{,}07/\sqrt{n}$ Error aproximado < 0.02		

Cuadro 13: KS CV para $\alpha = 0{,}05$, $\alpha^{'} = 0{,}20$ y $n = 4, ..., 10$

Se testea un nivel $\alpha = 0{,}05$ y se compara con un $\alpha = 0{,}20$, que es 4 veces el nivel tolerable de error. Dado que aún con error el estadístico D sigue siendo menor $CV(0{,}20) = 0{,}411$, se acepta la hipótesis nula.

4.4. Valor P para parámetros estimados [6]

Otra forma de validar (o rechazar) el valor del estadístico D es calculando su **valor p**:

$$Pvalue = P(D >= d)$$

Si $Pvalue > \alpha$ se acepta H_0.

Para calcular el **valor p** se puede realizar una simulación en base a la distribución candidata.

Primero, se debe obtener el estadístico TSV para los datos originales.

Luego, se deben realizar m simulaciones, en donde, en cada simulación se deben generar n números aleatorios respecto a la distribución candidata y calcular el estadístico TSV_s asociado a los datos simulados.

Finalmente obtener la ponderación de cuántos TSV'fueron mayores que el TSV original. O sea:

$$Pvalue = \frac{\#TSV_s > TSV}{m}$$

5. Conclusiones

El articulo tuvo como objetivo presentar y explicar los procedimientos implementados en RMES con el fin de que todo actor asociado al software tenga conocimiento de lo que sucede dentro.

Aún queda mucho por investigar. Existen pocos casos de prueba formales y es necesario que un experto verifique todo el proceso probabilistico. Además, en necesario estipular concensos dentro del grupo de trabajo para ciertas situaciones.

Para el algoritmo de ajuste con censura a la derecha **no se tuvo respaldo teórico de ningún tipo**. Por ende, todo lo que en este documento aparece respecto al tema es una traducción directa de lo implementado en código Java a lenguaje común matemático.

Aún queda por comparar el resultado de la estimación de RMES respecto a otros softwares. Lamentablemente, es una tarea bastante compleja dado que cada software tiene algoritmos propios para realizar las estimaciones de parámetros y los test de hipótesis.

En el futuro, este documento se irá perfeccionando y probablemente, dividiendo en varios otros para abarcar de forma particular y más precisa cada tema, en pos de una mejor explicación y entendimiento.

Referencias

[1] Statistical Assumptions of an Exponential Distribution. Volume 9, Number 2. Dr. Jorge Luis Romeu. Reliability Analysis Center. Documento de libre distribución disponible en http://src.alionscience.com/pdf/E_ASSUME.pdf a la fecha 10/12/2010.

[2] Empirical Assessment of Weibull Distribution. Volume 10, Number 3. Dr. Jorge Luis Romeu. Reliability Analysis Center. Documento de libre distribución disponible en http://src.alionscience.com/pdf/WEIBULL.pdf a la fecha 10/12/2010.

[3] The Chi-Square: a Large-Sample Goodness of Fit Test. Volume 10, Number 4. Dr. Jorge Luis Romeu. Reliability Analysis Center. Documento de libre distribución disponible en http://src.alionscience.com/pdf/Chi_Square.pdf a la fecha 10/12/2010.

[4] Kolmogorov-Simirnov: A Goodness of Fit Test for Small Samples. Volume 10, Number 6. Dr. Jorge Luis Romeu. Reliability Analysis Center. Documento de libre distribución disponible en http://src.alionscience.com/pdf/K_STest.pdf a la fecha 10/12/2010.

[5] Técnicas de validación estadística, Bondad de ajuste. Patricia Kisbye. Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba. Documento de libre distribución disponible en http://www.famaf.unc.edu.ar/ kisbye/mys/clase16_pr.pdf a la fecha 13/12/2010.

[6] Técnicas de validación estadística, Bondad de ajuste, Kolmogorov-Smirnov. Patricia Kisbye. Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba.