

## Descripción y aplicación de la versión corregida del Test de Langbein para verificar homogeneidad regional

### *Description and Application of the Corrected Langbein Test for Regional Homogeneity Verification*

Campos-Aranda D.F.

*Profesor jubilado de la Universidad Autónoma de San Luis Potosí*

*Correo: campos\_aranda@hotmail.com*

Información del artículo: recibido: junio de 2009, reevaluado: mayo de 2011, aceptado: septiembre de 2011

#### Resumen

Inicialmente se destaca la importancia del análisis de frecuencia para la estimación de las crecientes de diseño, las cuales permiten el dimensionamiento hidrológico de todo tipo de obras hidráulicas. También se citan las ventajas de los análisis regionales de frecuencia para obtener más información y por consecuencia mayor confiabilidad de resultados, cuando la región es homogénea. En seguida se citan los errores encontrados en la teoría y aplicación del Test de Langbein original. Posteriormente se describe con detalle la versión corregida de esta prueba estadística, la cual utiliza como estimador la variable reducida de la distribución Gumbel. Se exponen de manera concisa cuatro aplicaciones numéricas y se formula una conclusión que recomienda la aplicación de esta prueba en la verificación de la homogeneidad regional, dada su consistencia y sencillez.

#### Descriptor:

- test de Dalrymple
- distribución Gumbel
- coeficiente de variación
- sesgo del Cv
- variancia del Cv

#### Abstract

*Frequency analysis allows the estimation of design floods, which are used in the hydrological dimensionality of hydraulic structures. In particular, regional frequency analyses are convenient to obtain a higher amount of information, and thus more reliability on the calculated results (in the case of homogeneous regions). Hitherto some errors have been found in the theory and application of the original Langbein Test. Here, a corrected version of this statistic test is described in detail, which uses the reduced variate of Gumbel distribution as estimator. Four numerical applications are exposed briefly and their results are discussed. Based on this discussion, the systematic application of this corrected test is recommended when verifying regional homogeneity due to its simplicity and consistency.*

#### Keywords:

- Dalrymple test
- Gumbel distribution
- coefficient of variation
- Cv bias
- Cv variance

## Introducción

Las crecientes o avenidas máximas son riesgos naturales que pueden causar pérdida de vidas humanas y daños graves a todo tipo de propiedades. La estimación hidrológica de las crecientes es necesaria en la planeación, diseño, construcción y operación de las obras hidráulicas, así como en el manejo de las planicies de inundación. El procedimiento más confiable para realizar la estimación hidrológica citada es el análisis de frecuencia, el cual consiste en ajustar una función de distribución de probabilidades (FDP) a los datos históricos disponibles de gastos máximos anuales ( $Q$ ) medidos en una estación hidrométrica, y con base en tal modelo probabilístico, realizar las predicciones buscadas.

Casi siempre el análisis de frecuencia de crecientes implica un problema de escasez de datos, ya que habitualmente no existen registros de los sitios de interés y cuando se dispone de uno cercano su amplitud no conduce a predicciones confiables en los periodos de retorno de diseño. Una manera de obtener más información consiste en utilizar todos los registros disponibles en una región, cuyo comportamiento en relación con las crecientes es similar. Este es el enfoque básico del análisis regional de frecuencia de crecientes, el cual se basa en el concepto de región homogénea: grupo de sitios, con cuencas colindantes o no, cuyos datos de crecientes cuando se escalan con su media ( $Q/Q$ ) presentan la misma FDP (Fill y Stedinger, 1995).

En la práctica, la homogeneidad regional es verificada a través de la variabilidad que existe entre sitios o localidades, en su coeficiente de variación, de asimetría, de sus equivalentes en momentos  $L$  (Hosking y Wallis, 1997), o bien, de ciertas propiedades estadísticas adimensionales. Con este último planteamiento trabaja el test de Langbein, el cual fue difundido por Tate Dalrymple en 1960.

El objetivo de este trabajo consiste en exponer la versión corregida del Test de Langbein según Fill y Stedinger (1995), también en describir dos aplicaciones numéricas en regiones hidrológicas que previamente habían sido declaradas homogéneas con el test original y otras dos que posteriormente han sido verificadas homogéneas con otras pruebas estadísticas.

## Desarrollo

### Versión original del Test de Langbein

Desde que Dalrymple (1960) expuso el test desarrollado por Langbein W.B., éste se difundió y se convirtió en

la técnica clásica para verificar la homogeneidad de una zona o región, por ello fue expuesto en los textos básicos de hidrología, por ejemplo en Chow (1964), Kite (1977) y Singh (1992). En México ha sido expuesto y aplicado por Campos (1994, 2006a, 2006b); Gutiérrez y Ramírez (2005).

### Errores encontrados al test original

Fill y Stedinger (1995) indican que no está claro cómo se dedujo la expresión para la variancia del estimador utilizado en la prueba y deducen teóricamente su expresión correcta, encontrando un polinomio cuadrático del coeficiente de variación ( $Cv$ ) de los datos en lugar de la expresión sencilla original. El uso de la nueva expresión de la variancia del estimador define unas curvas de control más estrechas, función del  $Cv$  y del número total de datos. Estos autores también citan que Lu (1991) contrastó mediante simulación Monte Carlo el desempeño de la ecuación original y encontró grandes diferencias.

Wiltshire (1986) ha señalado que tal vez el Test de Langbein no es particularmente eficiente o potente, ya que en todas las regiones en las que ha sido aplicado, éstas resultaron homogéneas.

El otro aspecto erróneo del Test de Langbein original radica en no permitir que existan algunos sitios o puntos fuera de sus curvas de control. Cuando ello ocurría, tales estaciones hidrométricas se eliminaban del grupo. Lo anterior significa que el test no contempla la homogeneidad de la región como un todo (Fill y Stedinger, 1995). Esto también se corrige en la nueva versión del Test de Langbein; además ya no se aplica la prueba con un periodo común de datos, lo cual lo hace más versátil y cómodo al no requerir deducir o estimar valores faltantes de gastos máximos anuales.

### Versión corregida del Test de Langbein

Su aplicación se lleva a cabo a través de los cuatro pasos siguientes (Fill y Stedinger, 1995):

*Paso 1.* Se calcula el coeficiente de variación de cada sitio  $i$  ( $Cv^i$ , ecuación 1) y se obtiene su promedio ponderado por las longitudes de registro  $n_i$  con la expresión 4.

$$Cv^i = \frac{S^i}{Q} \quad (1)$$

donde

$$\bar{Q}^i = \frac{\sum_{j=1}^{n_i} Q_j^i}{n_i} \quad (2)$$

y

$$S^i = \sqrt{\frac{\sum_{j=1}^{n_i} (Q_j^i - \bar{Q}^i)^2}{n_i - 1}} \quad (3)$$

$$Cv^R = \frac{\sum_{i=1}^m n_i \cdot Cv^i}{\sum_{i=1}^m n_i} \quad (4)$$

$m$  es el número de estaciones hidrométricas de la región que se analiza.  $Cv^R$  es el coeficiente de variación regional.

*Paso 2.* El valor anterior se corrige por sesgo, sumándole el resultado de la ecuación siguiente:

$$\text{sesgo } (Cv^R) = \frac{Cv^R \cdot \psi(Cv^R)}{\frac{1}{m} \sum_{i=1}^m n_i} \quad (5)$$

donde

$$\psi(Cv) = Cv^2 - 0.57 \cdot Cv - 0.55 \quad (6)$$

*Paso 3.* Se determinan, con base en los  $Cv^i$ , los intervalos de confianza del periodo de retorno ( $T_S, T_I$ ) necesarios para las longitudes de registro  $n_i$ , con base en las ecuaciones siguientes:

$$T_{S/I} = 1 / \left\{ 1 - \exp \left[ - \exp \left( - \left( 2.2504 + \text{sesgo } y_{10}^i \pm 1.96 \cdot \sqrt{\text{var } y_{10}^i} \right) \right) \right] \right\} \quad (7)$$

donde

$$\text{sesgo } y_{10}^i = 1.6732 \left[ \frac{\varphi(Cv) - \psi(Cv)}{n_i} - \frac{\varphi(Cv) - m \cdot \psi(Cv)}{\sum_{i=1}^m n_i} \right] \quad (8)$$

y

$$\text{var } y_{10}^i = 2.80 \cdot \varphi(Cv) \cdot \left( \frac{1}{n_i} - \frac{1}{\sum_{i=1}^m n_i} \right) \quad (9)$$

donde

$$\varphi(Cv) = Cv^2 - 1.14 \cdot Cv + 1.10 \quad (10)$$

En las fórmulas anteriores,  $y_{10}^i$  es el estimador de la variable reducida de Gumbel de periodo de retorno 10 años en el sitio  $i$ .

*Paso 4.* Para cada sitio se calcula la estimación de la variable reducida con la expresión siguiente:

$$y_{10}^i = 1.6732 \frac{Cv^R}{Cv^i} + 0.5772 \quad (11)$$

y se obtiene su correspondiente periodo de retorno con la ecuación:

$$T(Q_{10}^i) = 1 / \left\{ 1 - \exp \left[ - \exp \left( - y_{10}^i \right) \right] \right\} \quad (12)$$

Se contabiliza el número de veces ( $k$ ) que el valor anterior está fuera de los intervalos de confianza respectivos, calculados en el paso anterior.

El número máximo de valores de  $k$  que son admisibles en una región homogénea se obtienen mediante la distribución de Bernoulli, considerando cada sitio como un experimento con probabilidad de éxito 0.05 y de fracaso 0.95, entonces (Fill y Stedinger, 1995):

$$P(k \geq k_c) = \sum_{j=k_c}^m \frac{m!}{j! (m-j)!} (0.05)^j \cdot (0.95)^{m-j} = \alpha \quad (13)$$

en donde  $\alpha$  es la probabilidad de cometer error tipo I en el test, es decir, 5%. Las mejores aproximaciones son:  $\alpha = 4.44\%$  y  $\alpha = 5.03\%$ , se obtuvieron con  $m = 7$  y  $k_c = 2$  y con  $m = 17$  y  $k_c = 3$ , respectivamente. Para fines prácticos se puede aceptar  $k_c = 2$  para  $5 \leq m \leq 10$  y  $k_c = 3$  cuando  $11 \leq m < 19$ . Entonces, cuando  $k < k_c$  la región es homogénea, en caso contrario es no homogénea. Conviene recordar, que no es recomendable con fines prácticos, formar regiones con más de 20 sitios.

En caso de querer aplicar la prueba con  $\alpha = 1.0\%$ , la única aproximación se obtiene con  $m = 15$  y  $k_c = 2$ , lo cual implica que con confiabilidad de 99% sólo en regiones con 15 o más sitios es aceptable tener  $k_c = 2$ . Por el contrario, cuando en la prueba se quiera utilizar  $\alpha = 10\%$ , sus tres mejores aproximaciones se obtienen con  $m = 6$ ,  $k_c = 2$ ,  $m = 11$ ,  $k_c = 3$  y  $m = 18$  y  $k_c = 4$ ; por lo cual, con confiabilidad de 90% se puede aceptar, de manera aproximada, que  $k_c = 2$  en regiones con menos de 7 sitios,  $k_c = 3$  cuando  $8 \leq m < 16$  y  $k_c = 4$  regiones con más de 17 sitios.

### Primera aplicación numérica

Campos (1994) presentó la información de gastos máximos anuales ( $m^3/s$ ) de 13 estaciones hidrométricas de la

Región Hidrológica Núm. 10 (Sinaloa), cuyos lapsos de registro variaron de 1939 a 1981, con un máximo de 42 datos y un mínimo de 19. Estos gastos están asociados a tormentas ciclónicas, ya que la mayoría de tales estaciones tienen áreas de cuenca de varios miles de km<sup>2</sup>. Utilizando un periodo común de 24 años (1957-1980) se probó la homogeneidad con base en el Test de Langbein original, encontrando que únicamente la estación Tamazula debía ser excluida (Campos, 1994, 2006b).

Los resultados de la aplicación de la versión corregida del Test de Langbein (ecuaciones 1 a 13) se muestran en la tabla 1, observándose que ahora son dos las estaciones hidrométricas que caen fuera de las curvas de control: Tamazula y Toahayana. Como  $k = 2$  es menor que  $k_c$  la región se considera homogénea, con  $\alpha = 5.0\%$ .

Núm.	Estación hidrométrica	$n_i$	$\bar{Q}^i$	$Cv^i$	$T_i$	$T_s$	$T(Q_{10}^i)$	$k$
1	Naranjo	42	567.452	1.1157	6.5	16.7	7.5	no
2	Huites	41	3,173.171	1.0425	6.6	16.5	8.2	no
3	Jaina	40	1,125.975	0.9929	6.6	16.4	8.7	no
4	San Francisco	40	1,787.800	0.7975	6.7	15.8	12.5	no
5	Santa Cruz	40	1,434.250	0.8655	6.7	16.0	10.8	no
6	Zopilote	42	342.191	0.7831	6.8	15.5	12.9	no
7	Acatitán	26	1,006.654	0.8764	6.1	18.5	10.6	no
8	Badiraguato	25	574.720	1.3669	5.3	22.4	5.9	no
9	Choix	26	321.500	0.6716	6.3	17.9	17.6	no
10	Ixpalino	29	1,283.448	0.9764	6.2	18.2	9.0	no
11	Palo Dulce	25	1,049.200	1.2349	5.6	21.1	6.6	no
12	Tamazula	24	615.792	0.4556	6.2	18.5	50.5	si
13	Toahayana	24	953.208	0.6069	6.2	18.4	22.3	si

Tabla 1. Resultados de la aplicación de versión corregida del Test de Langbein, en la Región Hidrológica Núm. 10 (Sinaloa).

$$Cv^R = 0.9147$$

$$\text{sesgo } (Cv^R) = -0.0066$$

Núm.	Estación hidrométrica	$n_i$	$\bar{Q}^i$	$Cv^i$	$T_i$	$T_s$	$T(Q_{10}^i)$	$k$
1	Mante	53	65.419	0.8871	7.1	14.8	11.6	no
2	Sabinas	42	356.929	0.5818	7.0	15.3	29.6	si
3	La Servilleta	43	446.372	0.6663	7.0	15.3	20.9	si
4	La Encantada	53	361.009	1.6130	6.2	18.1	5.4	si
5	San Gabriel II	56	401.684	1.4492	6.5	16.8	6.0	si
6	Magiscatzin II	49	1,521.592	0.7519	7.1	14.8	16.0	si
7	Tamesí	29	1,279.703	0.4598	6.5	17.3	61.7	si

Tabla 2. Resultados de la aplicación de versión corregida del Test de Langbein, en la Cuenca del Río Guayalejo (Región Hidrológica Núm. 26).

$$Cv^R = 0.9751$$

$$\text{sesgo } (Cv^R) = -0.0033$$

Núm.	Estación hidrométrica	$n_i$	$\bar{Q}^i$	$Cv^i$	$T_i$	$T_s$	$T(Q_{10}^i)$	$k$
1	San Nicolás	27	112.370	0.5310	6.4	17.2	13.1	no
2	Tonalá	25	158.532	0.8987	6.1	18.4	6.2	no
3	Cacaluta	33	186.555	0.5374	6.7	16.0	12.8	no
4	Pijijiapán	31	279.223	0.6945	6.6	16.5	8.5	no
5	Cahuacán	40	181.615	0.4762	6.9	15.1	16.3	si
6	Despoblado	31	276.265	0.5604	6.6	16.4	11.9	no
7	Malpaso	23	340.539	0.8735	6.0	18.9	6.4	no

Tabla 3. Resultados de la aplicación de versión corregida del Test de Langbein, en la Región Hidrológica Núm. 23 (Costa de Chiapas).

$$Cv^R = 0.6313$$

$$\text{sesgo } (Cv^R) = -0.0108$$

### Tercera aplicación numérica

Campos (2008) ha expuesto los datos disponibles en el periodo de 1950 a 1997 de gastos máximos anuales ( $\text{m}^3/\text{s}$ ) en las siete estaciones hidrométricas de la Región Hidrológica Núm. 23 (Costa de Chiapas), cuyos registros varían de 23 a 40 valores. En la tabla 3 se indican los resultados de la aplicación de la versión corregida del Test de Langbein, resultando que tal región es homogénea, pues sólo la estación Cahuacán queda fuera de las curvas de control, entonces:  $k = 1 < k_c$ , con  $\alpha = 5.0\%$ .

### Cuarta aplicación numérica

Campos (2011) presenta los datos generales de siete estaciones hidrométricas localizadas dentro de la cuenca

del Río Verde, el cual desemboca en el Río Santiago al noreste de la ciudad de Guadalajara.

De esas estaciones (tabla 4), únicamente las dos primeras tienen régimen virgen, el resto afuera cuencas altamente aprovechadas como Ajojucar y Paso del Sabino sobre el río Aguascalientes, o bien, Lagos y San Gaspar en cuya cuenca se ubica la Presa El Cuarenta en el río Lagos.

La nueva versión del test de Langbein indica que dos estaciones quedan fuera de sus curvas de control, Agostadero y Paso del Sabino, como se muestra en la tabla 4. Al suprimir la segunda, cuyo coeficiente de variación ( $C_v$ ) difiere notablemente del resto (tabla 4), la prueba deja fuera a la primera ( $k = 1$ ) pero en el límite, por ello se acepta que esas seis estaciones forman una región homogénea.

Núm.	Estación hidrométrica	$n_i$	$\bar{Q}^i$	$C_v^i$	$T_i$	$T_s$	$T(Q_{10}^i)$	k
1	Agostadero	39	67.269	0.6179	6.9	15.5	27.6	si
2	Valle de Guadalupe	47	83.338	0.8842	7.0	15.1	12.4	no
3	Lagos	31	178.861	0.8919	6.4	17.2	12.2	no
4	San Gaspar	61	178.661	0.8735	7.3	14.1	12.7	no
5	Ajojucar	23	66.683	0.9979	5.9	19.8	10.1	no
6	Paso del Sabino	31	209.610	2.3006	4.2	32.6	4.2	si
7	La Cuña	52	525.239	0.8730	7.1	14.7	12.7	no

Tabla 4. Resultados de la aplicación de versión corregida del Test de Langbein, en la cuenca del Río Verde de la Región Hidrológica Núm. 12 Parcial (Río Santiago).

$$C_v^R = 1.0079$$

$$\text{sesgo } (C_v^R) = -0.0027$$

### Conclusiones

Los resultados de los contrastes expuestos para la versión corregida del Test de Langbein, son totalmente coincidentes con las conclusiones que se han obtenido en esas regiones con otras pruebas estadísticas. Además, la prueba es muy fácil de realizar y no requiere definir un periodo común. Por lo anterior, se sugiere su aplicación sistemática en la verificación de la homogeneidad de la zona, previa al análisis regional de frecuencia de datos hidrológicos, no sólo de crecientes, como las aplicaciones aquí expuestas.

### Referencias

Campos-Aranda D.F. Aplicación del método de índice de crecientes en la Región Hidrológica Número 10, Sinaloa. *Ingeniería Hidráulica en México*, volumen IX (número 3), septiembre-diciembre 1994: 41-55.

Campos-Aranda D.F. Contraste de métodos regionales de estimación de crecientes en la cuenca del Río Guayalejo, en Tamaulipas. *Pláloc*, volumen 37, mayo-agosto 2006a: 14-24.

Campos-Aranda D.F. *Análisis Probabilístico Univariado de Datos Hidrológicos*, Avances en Hidráulica 13, IMTA-AMH, Jiutepec, Morelos, 2006b, pp. 133-161.

Campos-Aranda D.F. Aplicación de la prueba de discordancias a las crecientes de la costa de Chiapas, en: Memorias del XX Congreso Nacional de Hidráulica, tema: Hidrología superficial y subterránea, ponencia 5, Toluca, Estado de México del 15 al 18 de octubre de 2008.

Campos-Aranda D.F. Estimación de envolventes de diseño por subregiones hidrológicas. *Tecnología y Ciencias del Agua*, volumen II (número 4), octubre-diciembre 2011: 175-194.

Chow V.T. Frequency Analysis, en: *Handbook of Applied Hydrology*, Te-Chow V. (editor), Nueva York, USA, McGraw-Hill Book Co, 1964, pp. 8.1-8.42.

- Dalrymple T. Flood-Frequency Analyses, en: *Manual of Hydrology*, part 3: Flood-Flow Techniques, US Geological Survey, Water-Supply Paper 1543-A, Washington, USA, 1960.
- Fill H.D., Stedinger J.R. Homogeneity Test Based Upon Gumbel Distribution and a Critical Appraisal of Dalrymple's Test. *Journal of Hydrology*, volumen 166, 1995: 81-105.
- Gutiérrez-López A., Ramírez-Orozco A.I. Predicción hidrológica mediante el método de la avenida índice para dos poblaciones. *Ingeniería Hidráulica en México*, volumen XX (número 2), abril-junio 2005: 37-47.
- Hosking J.R.M., Wallis J.R. *Regional Frequency Analysis. An Approach Based on L-Moments*, Cambridge, United Kingdom, Cambridge University Press, 1997, 224 p.
- Kite G.W. *Frequency and Risk Analyses in Hydrology*, Colorado, USA, Water Resources Publications, Fort Collins, 1977, pp. 169-199.
- Lu L.H. *Statistical Methods for Regional Flood Frequency Investigations*, (Ph. D. Dissertation), Cornell University, Ithaca, Nueva York, 1991, 236 p.
- Singh V.P. *Elementary Hydrology*, Englewood, Nueva Jersey, Prentice Hall, 1992, pp. 824-839.
- Wiltshire S.E. Regional Flood Frequency Analysis I: Homogeneity Statistics. *Hydrological Sciences Journal*, volumen 31, 1986: 321-333.

#### Este artículo se cita:

##### Citación Chicago

Campos-Aranda, Daniel Francisco. Descripción y aplicación de la versión corregida del Test de Langbein para verificar homogeneidad regional. *Ingeniería Investigación y Tecnología*, XIII, 04 (2012): 411-416.

##### Citación ISO 690

Campos-Aranda D.F. Descripción y aplicación de la versión corregida del Test de Langbein para verificar homogeneidad regional. *Ingeniería Investigación y Tecnología*, volumen XIII (número 4), octubre-diciembre 2012: 411-416.

#### Semblanza del autor

*Daniel Francisco Campos-Aranda.* Obtuvo el título de ingeniero civil en diciembre de 1972 en la entonces Escuela de Ingeniería de la Universidad Autónoma de San Luis Potosí. Durante el primer semestre de 1977, realizó en Madrid, España un diplomado en hidrología general y aplicada. Posteriormente, durante 1980-1981, llevó a cabo estudios de maestría en ingeniería en la especialidad de hidráulica en la División de Estudios de Posgrado de la Facultad de Ingeniería de la UNAM. En esta misma institución, inició (1984) y concluyó (1987) el doctorado en ingeniería con especialidad en aprovechamientos hidráulicos. Ha publicado artículos principalmente en revistas mexicanas de excelencia: 39 en *Tecnología y Ciencias del Agua*, 13 en *Agrociencia* y 9 en *Ingeniería. Investigación y Tecnología*. En congresos internacionales y nacionales ha presentado 24 y 73 ponencias, respectivamente. Fue investigador nacional (nivel I: expediente 7273) desde el 1º de julio de 1991 hasta el 31 de diciembre del 2007. Actualmente es profesor jubilado de la UASLP, desde el 1º de febrero del 2003. En 2008 la AMH le otorgó el premio nacional "Francisco Torres H." a la práctica profesional de la hidráulica.