COMENTARIOS SOBRE ARTICULOS PUBLICADOS EN MESES ANTERIORES

Comentarios al artículo: "El método de simulación y la resistencia característica del hormigón", de Valentín Martín Jadraque (publicado en la Revista de Obras Públicas de mayo de 1971).

Por JORGE DOU MAS DE XEXAS Arquitecto

He leído con interés el artículo "El método de simulación y la resistencia característica del hormigón" del número 3073 de la Revista, y su lectura me ha movido a un comentario.

Si bien el comentario es sobre el estimador R^* que se propone en el artículo, son inevitables referencias de orden más general.

En el artículo parece justificarse el estimador propuesto en consideraciones sobre su mejor centrado y su mejor comportamiento respecto al α^* (de la instrucción) evitando ciertas inconveniencias.

Dejando de lado la cuestión de si la resistencia característica se define simplemente por el estimador α^*, prescindiendo de su significado para el caso de hipótesis de población normal o de otro tipo, y admitiendo, de momento, que es deseable que un estimador cumpla la condición de que por debajo de su valor medio se encuentre el 5 por 100 de la población supuesta de distribución normal, vamos a referirnos al centrado de α^* y R^*.

Tanto el estimador α^* como el R^* son casos particulares del:

$$\rho^* = \sum_1^{2N} K_i \, x_i \quad \text{con} \quad \sum_1^{2N} K_i = 1$$

donde x_p es la observación de lugar p en orden creciente de las $n = 2N$ observaciones.

El estimador α^* se obtiene haciendo:

$$K_i = \frac{3}{2N} \text{ para } i \leq N \text{ y } K_i = -\frac{1}{2N} \text{ para } i > N$$

El R^* resulta haciendo:

$$K_i = \frac{2}{N-1} \text{ para } i < N; K_N = 1 \text{ y } K_i = 0 \text{ para } i > N$$

El valor medio de ρ^* será:

$$E(\rho^*) = \sum_1^{2N} K_i \, E(X_i)$$

donde X_i es la v. a. constituida por el valor que ocupa el lugar i en orden creciente de una serie de $2N$ observaciones al azar.

Ahora bien, si consideramos la v. a. Z_p valor pésimo de una serie de n observaciones al azar extraídas de una v. a. normal [0; 1], su función de densidad será:

$$f_p(z) = p \binom{n}{p} \varphi(z) \cdot \left[\varnothing(z)\right]^{p-1} \cdot \left[1 - \varnothing(z)\right]^{n-p}$$

donde:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \quad \text{y} \quad \varnothing(x) = \int_{-\infty}^{x} \varphi(t)\, dt$$

El valor medio de Z_p será:

$$E(Z_p) = p \binom{n}{p} \cdot \int_{-\infty}^{+\infty} x \cdot \varphi(x) \cdot \left[\varnothing(x)\right]^{p-1} \cdot \left[1 - \varnothing(x)\right]^{n-p} \cdot dx$$

Como las funciones φ y $\varnothing$ están tabuladas es fácil el cálculo de las integrales.

Para el caso $2N = 6$ se ha procedido mediante un corto cálculo; tomando tan sólo una docena de intervalos de 0,4 de amplitud se ha obtenido:

$$E(Z_1) = -1{,}268; \quad E(Z_2) = -0{,}642; \quad E(Z_3) = -0{,}198$$

y por consiguiente:

$$E(\alpha^*) = \frac{2}{3}(-1{,}268 - 0{,}642 - 0{,}198) = -1{,}405$$

$$E(R^*) = -1{,}268 - 0{,}642 + 0{,}198 = -1{,}712$$

Claro está que con los valores $E(Z_i)$ se puede controlar el centrado de cualquier estimador del tipo ρ^*.

En particular, el estimador α^* podría retocarse conservando su idea y dejarlo más centrado. Bastaría tomar α_1^*:

$$\alpha_1^* = \left(2+\frac{1}{N}\right)\frac{x_1+x_2+\ldots+x_N}{N} - \left(1+\frac{1}{N}\right)\frac{x_1+x_2+\ldots+x_{2N}}{2N}$$

Veamos ahora un criterio de bondad fácilmente aplicable como el llamado de "eficiencia relativa" que consiste simplemente en la comparación de los momentos de segundo grado de α^* y R^* respecto del valor estimado.

Utilizando los resultados que figuran en el artículo se tendrá una eficiencia relativa de α^* respecto R^* de:

$$\varepsilon = \frac{E[(R^* - 1{,}64)^2]}{E[(\alpha^* - 1{,}64)^2]}$$

Para $2N = 6$ $\varepsilon = 153\ \%$ y para $2N = 30$ $\varepsilon = 112\ \%$.

O sea, que incluso pretendiendo que el valor estimado sea 1,64, el estimador α^* resulta más eficiente que R^*.

Tratemos ahora la cuestión de la pretendida contradicción.

El hecho de que agregando a una muestra observaciones de valores más altos se pueda dar lugar a una muestra cuya estimación resulte más baja, es considerado por algunos como una contradicción.

Sea el ejemplo del artículo:

$$M_1 = \{165, 172, 175, 180, 182, 188\};$$
$$M_2 = \{225, 232, 238, 242\};$$
$$M_3 = M_1 \cup M_2$$

Se tiene $\alpha^*[M_1] = 164{,}3$; $\alpha^*[M_3] = 149{,}7$, siendo los valores de M_2 superiores a los de M_1.

Se dice que con el estimador R^* se intenta corregir esta pretendida contradicción. Para el citado ejemplo se tiene:

$$R^*[M_1] = 162\ ; \quad R^*[M_3] = 164$$

de modo que en lugar de disminuir aumenta.

Pero el caso es que el mejor ajuste de una curva normal a la muestra realizado por el método de máxima verosimilitud, que en este caso coincide con el de igualación de momentos, da lugar a una estimación del valor del 5 por 100 muy parecida a la que proporciona α^* y sigue la misma pauta.

Por otra parte parece claro que si un estimador como el R^* de valor medio algo inferior al de α^* para unos tipos de muestras proporciona estimaciones más altas que las de α^* deberá compensar con estimaciones más bajas.

Una sencilla extensión del ejemplo anterior aclarará mejor:

Sea:

$$M'_2 = \{225, 232, 238, 242, 245, 250, 252, 254\}$$

cuyos valores son manifiestamente más altos que los de M_1. Consideremos la muestra $M'_3 = M_1 \cup M'_2$ de 14 observaciones.

Se tendrá $\alpha^*[M'_3] = 154$, mientras $R^*[M'_3] = 129$, con lo que resulta que la mal llamada penalización que se intentaba corregir resulta en R^* de una intensidad de más de tres veces mayor que en α^*.

Ciertamente que no es correcto poner a prueba un estimador mediante muestras inventadas, cuya densidad de ocurrencia para su entorno no se tiene en cuenta; de ello, no obstante, se presentan cuatro muestras esquemáticas con las correspondientes estimaciones según α^*, R^* y según método de igualación de momentos θ^*.

MUESTRAS	α^*	R^*	θ^*
$\mathcal{M}_1 = \{175, 175, 175, 205, 205, 205\}$	160	175	165
$\mathcal{M}_2 = \{170, 170, 170, 180, 180, 180\}$	165	170	167
$\mathcal{M}_3 = \{170, 170, 170, 170, 180, 180\}$	167	170	165
$\mathcal{M}_4 = \{170, 170, 180, 180, 180, 180\}$	170	160	169

Para todas estas muestras, y en general, el estimador, α^* se comporta más parecidamente al θ^* que el R^*. Ello parece explicarse en gran parte por cargar el estimador un peso muy grande sobre la observación x_N y acarreando toda su dispersión.

Si tratamos de conjeturar sobre la forma de la distribución de las resistencias del hormigón en una cierta obra, es de presumir que sea acampanada y no simétrica, puesto que en la ejecución preside la idea, no simétrica, de obtener resistencias superiores a un nivel, a veces convencional, sin preocupación por un máximo.

En general, ni siquiera se podrá prejuzgar el signo de la asimetría.

Teniendo que actuar la instrucción sobre variadas formas posibles de distribución, parece razonable tomar como referencia la distribución normal como más próxima a todas ellas.

Considerando que el valor a estimar es un valor situado muy a la izquierda, parece apropiado intentar el ajuste del ala izquierda de la curva normal con la mitad izquierda de la muestra.

Por un lado se pierde información, pero por otro el ajuste a la parte de información más próxima al objetivo resultará más perfecta.

Con esta idea, de dudosa eficacia, se pueden elaborar numerosas estimaciones del tipo ρ^*:

$$\rho^* = \sum_{1}^{2N} K_i x_i \; ; \quad \sum_{1}^{2N} K_i = 1 \quad K_i = 0 \text{ para } i > l$$

Para el caso $2N = 6$ se han redactado unos cuantos que, admitiendo la normalidad, se centran sensiblemente en el 5 por 100:

$\rho_1^* = 1{,}6\, x_1 - 0{,}6\, x_2$ (con sólo dos observaciones)

$\rho_2^* = 1{,}44\, x_1 - 0{,}22\,(x_2 + x_3)$

$\rho_3^* = 0{,}95\,(x_1 + x_2) - 0{,}9\, x_3$ (casi igual al R^*)

$\rho_4^* = 0{,}86\,(x_1 + x_2) - 0{,}36\,(x_3 + x_4)$

$\rho_5^* = 0{,}68\,(x_1 + x_2 + x_3) - 1{,}04\, x_4$ ... etc...

y análogamente para otros valores de $2N$.

El estimador más simple de este tipo se presenta para $2N = 12$:

$$\rho_0^* = x_1 \quad \text{con} \quad P\{X < E(\rho_0^*)\} = 0{,}052$$

Una propiedad curiosa del estimador R^* es que "cualquiera que sea $2N$ si se admite que la muestra es extraída de una población uniforme en el intervalo $[a, b]$, el valor medio de R^* es exactamente el extremo inferior a". El estimador α_1^* (retocado del α^*) también tiene esta propiedad.

Es de observar que la condición de centrado del valor medio de un estimador sólo tiene sentido para una determinada hipótesis de distribución y, por tanto, la probabilidad $P\{X < E(\alpha)\}$ depende de dicha hipótesis. Tiene tan sólo valor orientativo el decir que supuesto X normal dicha probabilidad se aproxima más o menos a un cierto porcentaje que, por otra parte, no es necesario que sea el 5 por 100.

La precisión del centrado no suele considerarse cualidad importante para un estimador, pero el que tenga la menor dispersión posible sí que lo es.

Para cerrar el comentario, nos parece que un estimador del tipo ρ^* tendrá menos dispersión si utilizando todas las observaciones tiene los coeficientes K_i de valores absolutos pequeños. El α^* cumple bastante con estas condiciones, y tiene la elegancia de una simplicidad insuperable.