PRUEBAS DE ANDERSON-DARLING MODIFICADAS CON MEJORA SELECTIVA DE LA POTENCIA.

A. Cabaña *y E. M. Cabaña †

Resumen

Se construye una prueba de ajuste inspirada en la prueba clásica de Anderson y Darling, en la que se reemplaza el proceso empírico por un *Proceso Empírico Transformado* (PET), y posteriormente se integra para distribuir el efecto de la singularidad que aparece en el extremo izquierdo del recorrido de la variable.

La *función de pesos* del PET se elige de manera tal que la potencia frente una sucesión dada de alternativas contiguas se magnifique.

La distribución asintótica del estadístico de prueba es la misma en todos los casos, bajo la hipótesis nula, y también bajo la sucesión normalizada de alternativas para la cual ha sido diseñado el estadístico.

Se proporcionan tablas de los niveles críticos y potencias asintóticas.

1 Introducción. Procesos empíricos transformados y su distribución asintótica.

Denotemos por $\{X_1, X_2, \ldots, X_n\}$ a una muestra de variables aleatorias reales, independientes con función de distribución F, y consideremos una familia de distribuciones de probabilidad $F^{(\tau)}$ (con τ en un entorno de 0^+) *contigua* a $F^{(0)} = F_0$ (ver [5],[6]), con densidad $f^{(\tau)}$ respecto de F_0, y tal que existe una función k en $L^2(\mathrm{R}, dF_0)$ de norma $\left(\int k^2 dF_0\right)^{1/2} = 1$ que satisface

$$\left\| \frac{1}{\tau}\left(\sqrt{f^{(\tau)}} - 1\right) - \frac{k}{2} \right\|_{L^2} \to 0 \text{ as } \tau \to 0^+, \tau \neq 0, \tag{1}$$

de donde se deduce que $\int k(x) dF_0(x) = 0$.

En [2], uno de los autores introduce ciertas familias de procesos empíricos transformados dependientes de un parámetro funcional (*función de pesos*) con el propósito de diseñar pruebas de bondad de ajuste de tipo Kolmogorov-Smirnov para la hipótesis nula $\mathcal{H}_0$: "$F = F_0$". Tales pruebas resultan ser consistentes frente a cualquier alternativa fija "$F \neq F_0$", y son especialmente sensibles a la sucesión de alternativas "$\mathcal{H}_n$: "$F = F^{(\delta/\sqrt{n})}$"

Siguiendo [3], definimos el Proceso Empírico Transformado (PET) de la muestra $\{X_1, X_2, \ldots, X_n\}$, asociado a la función de distribución F_0, la isometría T de $L_2 = L_2(\mathrm{R}, dF_0)$ con rango $1^\perp$ (el complemento ortogonal de la función constante 1), y la *función de pesos* a con $\|a\|^2 = \int a^2(x) dF_0(x) = 1$ como

$$w_n^{(a,T)}(A) = \int T(a\mathbf{1}_A) db_n, \tag{2}$$

*Departamento de Matemáticas, Instituto Venezolano de Investigaciones Científicas, Apartado 21827, Caracas 1020-A, Venezuela.

†Centro de Matemática, Facultad de Ciencias, Eduardo Acevedo 1139, 11200 Montevideo, Uruguay.

donde 1_A es la función indicatriz A, $b_n(x) = \sqrt{n}(F_n(x) - F_0(x))$ es el proceso empírico, y $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \le x\}}$ es la función de distribución de la muestra.

Sea $\mathcal{J}$ la familia de todos los intervalos en $\mathbb{R}$, y denotemos por $w^{(V)}$ un proceso de Wiener en $\mathbb{R}$ con medida de covariancia

$$V(A) = \mathrm{E}\left(w^{(V)}\right)^2 = \int_A a^2(t)dF_0(t), \; \mathrm{E}w^{(V)}(A)w^{(V)}(B) = V(A \cap B). \tag{3}$$

En [3] se muestra que, bajo condiciones apropiadas, el PET $w_n^{(a,T)}(A), A \in \mathcal{J}$ converge en ley a

$$w^{(V)}(A) + \delta \int kT(a1_A)dF_0, A \in \mathcal{J} \tag{4}$$

cuando n tiende a infinito. En particular, $w_n^{(a,T)}$ converge en ley a $w^{(V)}$ bajo $\mathcal{H}_0$.

Las condiciones generales que garantizan la convergencia a (4) se pueden ver en [3]. Cuando T es la L-isometría (ver [4], o §6.1 en [3])

$$(Tg)(x) = (T_{L,F_0}g)(x) = g(x) - \int_{-\infty}^{x} \frac{g(t)dF_0(t)}{1 - F_0(t)}, \tag{5}$$

la afirmación precedente sobre la distribución límite del PET se cumple si

$$\frac{|a|}{(1 - F_0)^{\alpha}} \in L_2(\mathbb{R}, dF_0), \text{ para algún } \alpha > 0. \tag{6}$$

2 El estadístico de Anderson-Darling modificado.

La prueba clásica de Anderson-Darling se basa en el estadístico cuadrático $S_n = \int (b_n(x))^2 \frac{dF_0(x)}{F_0(x)(1-F_0(x))}$ que se distribuye asintóticamente como $\int (b(s))^2 \frac{ds}{s(1-s)}$, donde b es un puente browniano estándar.

Cuando el proceso empírico b_n y su variancia $F_0(x)(1 - F_0(x))$ se reemplazan por el PET $w_n^{(a,T)}$ y $\mathbf{Var}w_n^{(a,T)}(x) = V(x)$ respectivamente, se obtiene el nuevo estadístico

$$T_n^0 = \int (w_n^{(a,T)}(x))^2 \frac{dV(x)}{V(x)} \tag{7}$$

que tiene una singularidad en el extremo izquierdo del soporte de V, por la misma razón por la cual el estadístico de Anderson-Darling tiene singularidades en ambos extremos del soporte de F_0. Vamos a evitar este comportamiento distribuyendo la singularidad sobre todo el recorrido de V de la siguiente manera:

Introducimos en primer lugar la familia de estadísticos cuadráticos

$$T_{n,x}^{(a,T)}(x) = \int \left(\int c_{(x,y)}(z)dw_n^{(a,T)}(z)\right)^2 \frac{dV(y)}{\int c_{(x,y)}(z)dV(z)} \tag{8}$$

con

$$c_{(x,y)}(s) = \begin{cases} 1 & \text{if } x < s < y, y < x < s \text{ or } s < y < x, \\ 0 & \text{si no.} \end{cases} \tag{9}$$

Definimos el nuevo estadístico $T_n^{(a,T)}$ integrando $T_{n,x}^{(a,T)}(x)$ sobre el recorrido de V:

$$T_n^{(a,T)} = \iint \left(\int c_{(x,y)}(z)dw_n^{(a,T)}(z)\right)^2 \frac{dV(x)dV(y)}{\int c_{(x,y)}(z)dV(z)}. \tag{10}$$

El proceso (4) que tiene la distribución límite de $\{w_n^{(a,\mathcal{T})}(A) : A \in \mathcal{J}\}$ bajo la sucesión de alternativas $\mathcal{H}_n$: "$F = F^{(\delta/\sqrt{n})}$", también puede escribirse como $\{w^{(V)}(A) + \delta \int_A a(\mathcal{T}^{-1}k)dF_0\}$, dado que $\int k\mathcal{T}(1_A a)dF_0 = \int_A a\mathcal{T}^{-1}kdF_0$ ya que $\mathcal{T}$ es una isometría.

Entonces, $T_n^{(a,\mathcal{T})}$ se distribuye asintóticamente bajo $\mathcal{H}_n$ como

$$\iint \left(\int c_{(x,y)}(z)(dw^{(V)}(z) + \delta a(\mathcal{T}^{-1}k)dF_0)\right)^2 \frac{dV(x)dV(y)}{\int c_{(x,y)}(z)dV(z)}$$

$$= \iint \left(\int c_{(x,y)}(z)(dw(V(z)) + \delta h(V(z))dV(z))\right)^2 \frac{dV(x)dV(y)}{\int c_{(x,y)}(z)dV(z)}, \tag{11}$$

donde $h(V(z))a(z) = (\mathcal{T}^{-1}k)(z)$. Como $c_{(x,y)}(z) = c_{(V(x),V(y))}(V(z))$, entonces, con las nuevas variables $r = V(x)$, $s = V(y)$, $t = V(z)$, (11) se reduce a

$$\iint \left(\int c_{(r,s)}(t)(dw(t) + \delta h(t)dt)\right)^2 \frac{dr\,ds}{\int c_{(r,s)}(t)dt}$$

$$= \int_0^1 \int_0^1 C(t,u)(dw(t) + \delta h(t)dt)(dw(u) + \delta h(u)du), \tag{12}$$

con

$$C(t,u) = \int_0^1 \int_0^1 c_{(r,s)}(t)c_{(r,s)}(u)\frac{dr\,ds}{\lambda(r,s)}, \quad \lambda(r,s) = \int c_{(r,s)}(t)dt. \tag{13}$$

La distribución de (12) depende sólamente de la función de pesos elegida a través de la función h. En particular, el sesgo asintótico bajo la alternativa es

$$b(a) = \delta^2 \int_0^1 \int_0^1 C(t,u)h(t)h(u)dt\,du. \tag{14}$$

El comportamiento límite descrito por (12) sugiere rechazar la hipótesis nula $\mathcal{H}_0$ cuando $T_n^{(a,\mathcal{T})}$ es mayor que una constante adecuada, y, para aumentar la sensibilidad de la prueba respecto de la sucesión de alternativas contiguas dada, proponemos elegir la función de pesos a que maximice el sesgo asintótico $b(a)$.

Proposición 1 *El sesgo asintótico $b(a)$ dado por (14) es máximo cuando la función de pesos a se elige igual a $\hat{a} = \mathcal{T}^{-1}k$, y su valor máximo es $\delta^2/2$.*

Nota. El peso óptimo $\hat{a} = \mathcal{T}^{-1}k$ es el mismo que optimiza la potencia para las pruebas de tipo Kolmogorov-Smirnov (ver [2, 3]).

Demostración. Calculemos, para $t < u$,

$$C(t,u) = \int_0^t dr \int_u^1 \frac{ds}{s-r} + \int_0^t dr \int_0^r \frac{ds}{1+s-r}$$

$$+ \int_t^u dr \int_t^r \frac{ds}{1+s-r} + \int_u^1 dr \int_u^r \frac{ds}{1+s-r} = \gamma(|u-t|), \tag{15}$$

con $\gamma(y) = 1 + |y|\log(|y|) + (1-|y|)\log(1-|y|)$. La expresión (15) también es válida para $u < t$, porque depende simétricamente de t y u.

La función γ es simétrica respecto del 0 y de 1/2, y esto implica que $\int_0^1 C(t,u)du$ no depende de t y es igual a $\int_0^1 \gamma(y)dy = 1/2$ así que, cuando h es la constante 1, y por lo tanto $a = \hat{a}$, tenemos $b(\hat{a}) = 1/2$. Por otra parte, por la desigualdad de Cauchy-Schwarz ,

$$b(a) \leq \int_0^1 \int_0^1 h^2(t)C(t,u)dt\,du = \frac{1}{2}\int_0^1 h^2(t)dt = \frac{1}{2}$$

porque la restricción $||k|| = 1$ implica que h debe satisfacer

$$\int_0^1 h^2(s)ds = \int (h(V(x)))^2 dV(x) = \int (T^{-1}k(x))^2 dF_0 = ||T^{-1}k||^2 = 1.$$

Esto concluye la demostración de la proposición.

3 Implementación de la prueba.

3.1 Cálculo del estadístico de prueba.

Abreviemos $T_n = T_n^{(\hat{a},T)}$. Los mismos cambios de variables hechos en §2 nos conducen a escribir la variable de prueba óptima como

$$\begin{aligned} T_n &= \int_0^1 \int_0^1 C(t,u) dw_n^{(\hat{a},T)}(V^{-1}(t)) dw_n^{(\hat{a},T)}(V^{-1}(u)), \\ &= \frac{1}{n} \sum_{i,j=1}^n \int_0^1 \int_0^1 C(t,u) dw_{X_i}^{(\hat{a},T)}(V^{-1}(t)) dw_{X_j}^{(\hat{a},T)}(V^{-1}(u)). \end{aligned}$$

Es fácil de verificar que para cualquier g medible, $\int g(x) dw_X^{(a,T)}(x) = T(ag)(X)$, y por lo tanto, con las notaciones

$$T_x g(x,y)|_{x=X} = T(g(\bullet,y))(X), \ /; T_y g(x,y)|_{y=Y} = T(g(x,\bullet))(Y), \tag{16}$$

llegamos a la expresión $T_n = \frac{1}{n}\sum_{i,j=1}^n S(X_i, X_j)$ con

$$S(X,Y) = T_x T_y \hat{a}(x)\hat{a}(y) C(V(x), V(y))|_{x=X,y=Y} \tag{17}$$

que muestra que T_n es un U-estadístico de segundo orden.

3.2 Regiones críticas y potencia.

La región crítica $T_n > \kappa(\alpha)$ con κ definido por

$$\mathbf{P}\left\{\int_0^1 \int_0^1 C(t,u) dw(t) dw(u) > \kappa(\alpha)\right\} = \alpha$$

proporciona una prueba para $\mathcal{H}_0$ consistente frente a cualquier alternativa fija, con nivel asintótico α. Su potencia asintótica es

$$\pi(\delta) = \mathbf{P}\left\{\int_0^1 \int_0^1 C(t,u)(dw(t) + \delta dt)(dw(u) + \delta du) > \kappa(\alpha)\right\}. \tag{18}$$

Los valores de $\kappa(\alpha)$ y $\pi(\delta)$ que se indican en las tablas 3.2 y 3.2, fueron obtenidos por medio de simulaciones basadas en 8000 réplicas. También se indica la potencia asintótica π^* de la prueba de Kolmogorov-Smirnov modificada introducida en [2] para la función de pesos óptima, con el propósito de hacer comparaciones. Se notará que ambas pruebas presentan un comportamiento muy similar.

3.3 Ejemplo: Ajuste a una normal estándar.

Finalmente, calculemos el estadístico $S(X,Y)$ definido en (17) para la isometría particular

$$Tg = g - \int_{-\infty}^{\bullet} \frac{g(t) d\Phi(t)}{1 - \Phi(t)}, \ T^{-1}h = h + \frac{1}{1-\Phi(\bullet)} \int_{-\infty}^{\bullet} h(t) d\Phi(t)$$

(los detalles acerca de ésta y otras isometrías se pueden ver en [3]) en dos casos sencillos:

α	1%	2.5%	5%	10%
$\kappa(\alpha)$	3.78	2.97	2.40	1.84

Tabla 1: Aproximación numérica de los valores críticos $\kappa(\alpha)$ para niveles α = 1, 2.5, 5 y 10%.

δ	0	0.2	0.4	0.6	0.8	1.0	1.2	1.4
$\pi(\delta)(\%)$	5.0	5.3	6.6	8.9	12.3	16.8	22.3	28.7
$\pi^*(\delta)(\%)$	5.0	5.4	6.7	8.9	12.0	16.2	21.3	27.3
δ	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
$\pi(\delta)(\%)$	35.7	43.2	52.0	59.9	67.2	74.3	80.2	85.0
$\pi^*(\delta)(\%)$	34.1	41.4	49.1	56.9	64.4	71.4	77.7	83.1

Tabla 2: Potencia asintótica $\pi(\delta)$ de la prueba propuesta ,y $\pi^*(\delta)$ de la prueba de K-S modificada con pesos óptimos,ambas con nivel 5%, como función de δ.

3.3.1 Caso 1: prueba sensible a cambios de localización.

Supongamos que $F_0(x) = \Phi(x) = \int_{-\infty}^{x} \varphi(t)dt$, $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $f^{(\tau)}(x) = \frac{\varphi(x-\tau)}{\varphi(x)} = e^{-\frac{\tau^2-2x\tau}{2}}$, por lo tanto $k(x) = \lim_{\tau\to 0^+} \frac{2}{\tau}(e^{-\frac{\tau^2-2x\tau}{4}} - 1) = x$, $||k||^2 = \int_{-\infty}^{\infty} x^2\varphi(x)dx = 1$, $\hat{a}(x) = \mathcal{T}^{-1}(x) = x + \frac{1}{1-\Phi(x)}\int_{-\infty}^{x} t\varphi(t)dt = x - \frac{\varphi(x)}{1-\Phi(x)}$, y entonces $V(x) = \int_{-\infty}^{x}\left(t - \frac{\varphi(t)}{1-\Phi(t)}\right)^2\varphi(t)dt = \int_{-\infty}^{x} t^2\varphi(t)dt + \frac{\varphi^2(x)}{1-\Phi(x)}$. Una vez obtenidas las expresiones analíticas de C, $\hat{a}$ y V, S puede ser calculada mediante de un algoritmo sencillo que involucra integración numérica.

3.3.2 Caso 2: prueba sensible a cambios de dispersión.

Sea $F^{(\tau)}(x) = \Phi((1-\frac{\tau}{\sqrt{2}})x)$,de manera que $f^{(\tau)}(x) = (1-\frac{\tau}{\sqrt{2}})\varphi((1-\frac{\tau}{\sqrt{2}})x)/\varphi(x) = (1-\frac{\tau}{\sqrt{2}})e^{\frac{1}{2}x^2(1-(1-\frac{\tau}{\sqrt{2}})^2)}$, $k(x) = \lim_{\tau\to 0^+}\frac{2}{\tau}\left(\sqrt{1-\frac{\tau}{\sqrt{2}}}e^{\frac{1}{4}x^2(\sqrt{2}\tau-\tau^2/2)} - 1\right) = \frac{1}{\sqrt{2}}(x^2-1)$, and $||k||^2 = \frac{1}{2}\int_{-\infty}^{\infty}(x^2-1)^2\varphi(x)dx = 1$.

La función de pesos es $\hat{a}(x) = \frac{\mathcal{T}^{-1}(x^2-1)}{\sqrt{2}} = \frac{x^2-1}{\sqrt{2}} + \frac{1}{1-\Phi(x)}\int_{-\infty}^{x}\frac{t^2-1}{\sqrt{2}}\varphi(t)dt = \frac{1}{\sqrt{2}}\left[x^2 - 1 - \frac{x\varphi(x)}{1-\Phi(x)}\right]$, y por lo tanto, $V(x) = \frac{1}{2}\int_{-\infty}^{x}\left[t^2 - 1 - \frac{t\varphi(t)}{1-\Phi(t)}\right]^2 dt = \int_{-\infty}^{x}(t^2-1)^2\varphi(t)dt - \frac{x^2\varphi^2(x)}{1-\Phi(x)}$. Como en el caso anterior, se puede utilizar integración numérica para hacer cada evaluación de S.

Referencias

[1] Anderson, T. W. and Darling, D. A., *A test of goodness of fit*. J. Amer. Statist. Assoc., 49 (1954), 765-769.

[2] Cabaña, A., *Transformations of the empirical process and Kolmogorov-Smirnov tests*. Ann. Statist.,Vol.24, No. 5 (1996).

[3] Cabaña, A. and Cabaña, E.M., Transformed Empirical Processes and Modified Kolmogorov-Smirnov Tests for multivariate distributions, Mathematics Preprint Series **208**, June 1996, Universitat de Barcelona, sometido a publicación.

[4] Groeneboom, P. and Wellner, J.A., *Information Bounds and Nonparametric Maximum Likelihood Estimation.*, DMV Seminars, Band 19, Birkhauser, New York, 1992.

[5] Le Cam, L. and Yang, G.L. *Asymptotics in Statistics. Some basic concepts.* Springer-Verlag, New York, 1990.

[6] Oosterhoff, J; van Zwet, W. R. *A note on contiguity and Hellinger distance.* Contributions to Statistics, 157-166, Reidel, Dordrecht, 1979.