

UNIVERSIDAD DE SEVILLA
NEGOCIADO DE TESIS

Queda registrado como Tesis de Doctor al
folio 82 número 110 del libro
correspondiente.

Sevilla, ENE. 2000

El Jefe del Negociado

Pere Raffel

UNIVERSIDAD DE SEVILLA

Depositado en el D^{to} de Estadística e I.O.
de la Univ. de Sevilla
de esta Universidad desde el día 24-1-00
hasta el día 14 de febrero de 2000

Sevilla 19 de febrero de 2000
EL DIRECTOR DEL Departamento

M. J. J.

DESARROLLOS ORTOGONALES PARA ESTIMADORES EN FAMILIAS PARAMÉTRICAS

TESIS DOCTORAL

M. TERESA GÓMEZ GÓMEZ

SEVILLA, 2000

UNIVERSIDAD DE SEVILLA
FACULTAD DE MATEMÁTICAS
BIBLIOTECA

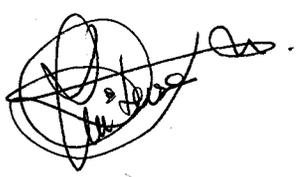
UNIVERSIDAD DE SEVILLA

Facultad de Matemáticas

Departamento de Estadística e Investigación Operativa

DESARROLLOS ORTOGONALES PARA ESTIMADORES EN FAMILIAS PARAMÉTRICAS

Memoria presentada por
M. Teresa Gómez Gómez
para optar al
Grado de Doctor.



M. Teresa Gómez Gómez.

Vº Bº de los Directores



Fdo. Fernando López Blázquez
Fdo. Juan Luis Moreno Rebollo

Sevilla, Enero 2000

A Manolo, Teresa y Carmen

Quiero expresar mi más sincero agradecimiento a todos mis compañeros de Departamento y a mis compañeras de la Biblioteca de la Facultad de Matemáticas que me han animado en todo momento y que en parte, han hecho posible la realización de esta memoria.

En especial, quiero dar las gracias al Prof. Dr. Don Rafael Infante Macías por sus consejos y por la confianza que ha depositado en mí, al Prof. Dr. Don Fernando López Blázquez por su dedicación y por todas las aportaciones científicas que han permitido la realización de este trabajo y al Prof. Dr. Don Juan Luis Moreno Rebollo por su constante apoyo moral y científico.

Sevilla Enero de 2000.

Introducción

Un tema clásico de la Estadística Matemática es el estudio de estimadores insesgados de mínima varianza en familias paramétricas. Trabajos pioneros en este área podemos citar, entre otros, los de Barankin (1949) [8], Stein (1950) [47], Washio et al. (1956) [52], Bahadur (1957) [5], Tate (1959) [49]. La Teoría básica de la estimación insesgada puede encontrarse en libros de texto tales como *The theory of statistical inference* (cap. 3) Zacks(1971) [55], *Theory of point estimation* (cap. 2) Lehman (1983) [28], o *Mathematical statistics* (cap. 3) Shao (1999) [43]. Dicha teoría básica se fundamenta en los bien conocidos teoremas de Blackwell-Rao y Lehman-Scheffé.

Tras la teoría básica son varias las cuestiones que quedan abiertas. Basaremos esta introducción sobre dichas preguntas y las soluciones a las mismas que hemos propuesto en la presente memoria.

En primer lugar, debemos proporcionar un marco adecuado para el estudio de los estimadores insesgados. También debemos delimitar el conjunto de funciones que admiten estimadores insesgados. Dado que el requisito de que un estimador tenga varianza finita no es restrictivo para la mayor parte de los problemas que vamos a tratar, resulta natural considerar espacios de funciones de cuadrado integrable (espacios L^2). De esta forma, un estimador con

varianza finita puede ser visto como un elemento o una función de un espacio L^2 . Estos espacios presentan la ventaja matemática de que son muy ricos en propiedades. Nuestro objetivo será explotar dichas propiedades matemáticas para el estudio de problemas relacionados con la estimación insesgada. Por otro lado, el conjunto de funciones del parámetro que admiten un estimador insesgado (con varianza finita) es la imagen del espacio L^2 considerado mediante el operador esperanza.

Una de las propiedades más interesantes de los espacios L^2 que consideraremos es la existencia de bases ortogonales y la posibilidad de encontrar desarrollos para funciones de cuadrado integrable en términos de dicha base. Desde el punto de vista estadístico, esto se traduce en la posibilidad de asociar a cada estimador un desarrollo ortogonal. Por motivos obvios, resulta conveniente que las bases ortogonales sean simples. Por dicha razón consideraremos desarrollos en términos de polinomios ortogonales.

Una vez delimitado el conjunto de posibles estimadores y funciones estimables surge una pregunta obvia: ¿Cómo se obtiene un estimador insesgado con mínima varianza para una cierta función paramétrica?. Una respuesta a esta pregunta la proporciona el Teorema de Lehman-Scheffé, según el cuál el procedimiento a seguir sería:

- (i) Obtener un estimador insesgado.
- (ii) Condicionar a un estadístico suficiente y completo.

El procedimiento anterior es insatisfactorio en multitud de casos, pues resulta paradójico que para encontrar un estimador insesgado sea necesario encontrar previamente otro estimador insesgado. Este inconveniente del pro-

cedimiento de Lehmann-Scheffé, ha motivado que en la literatura aparezcan métodos alternativos para la obtención de estimadores insesgados de mínima varianza. Entre los métodos alternativos pueden citarse: el método de igualación de desarrollos en series de potencias de Roy y Mitra (1958) [40] para familias de series de potencias; el uso de transformadas funcionales, por ejemplo la transformada de Laplace en familias exponenciales o la transformada de Mellin, ver Washio et al. (1956) [52].

El procedimiento que presentamos para la obtención de estimadores insesgados consiste en obtener directamente su desarrollo ortogonal. Un hecho notable es que tal desarrollo puede obtenerse en la mayor parte de los casos sin conocer el propio estimador, tan solo es necesario conocer la función paramétrica que deseamos estimar. La obtención de desarrollos ortogonales para los estimadores insesgados en familias exponenciales con función de varianza cuadrática tienen su origen en los trabajos de Abbey y David (1970) [1] y Morris (1982, 1983) [35] [36]. En la presente memoria mostramos como estos desarrollos pueden obtenerse para otras muchas familias, no necesariamente exponenciales.

Un paso previo a la obtención de los desarrollos de los estimadores consiste en obtener una familia de funciones ortogonales. Como ya se expresó anteriormente trataremos fundamentalmente con familias de polinomios ortogonales. Estos presentan la ventaja de que son objetos matemáticos simples, muy ricos en propiedades. Estudiaremos procedimientos para la obtención de dichos polinomios, destacando los métodos matriciales para la obtención de los mismos, pues tales métodos pueden implementarse de manera sencilla en un ordenador.

Una vez que han sido obtenidos los desarrollos para los estimadores insesga-

dos, investigamos propiedades que se derivan de tales desarrollos. Sobre este punto destacamos el hecho de que numerosas propiedades tanto ya conocidas como nuevas pueden ser deducidas a partir de los desarrollos. De forma general, podemos apuntar que a partir de los desarrollos podemos obtener, entre otros, resultados acerca de

- *Cotas inferiores para la varianza*, tales como extensiones de la cota de Fréchet-Cramer-Rao o de Bhattacharyya en familias regulares o cotas más finas que las de Chapman-Robins-Kieffer en familias no regulares.
- *Propiedades asintóticas*, en este sentido podemos destacar que a partir de los desarrollos es posible obtener todas las posibles distribuciones límites asociadas a estimadores insesgados.
- *Comparaciones con otros estimadores*, veremos que medidas tales como la eficiencia o la deficiencia asintótica pueden ser obtenidas estudiando los coeficientes de Fourier asociados a los desarrollos.
- *Obtención de estimadores aproximadamente insesgados*, pues por ejemplo existen funciones paramétricas que no admiten estimadores insesgados para los cuales es posible obtener estimadores con menor error cuadrático que el estimador de máxima verosimilitud.

La memoria que presentamos, está organizada tal como sigue:

En el primer capítulo se presenta la teoría general sobre la que se fundamenta el resto de la memoria, dedicando una sección al estudio de procedimientos para la obtención de sistemas de polinomios ortogonales.

El capítulo segundo está dedicado a la aplicación de la teoría general al

caso de las familias exponenciales. Precisamente, los primeros trabajos sobre desarrollos ortogonales para estimadores fueron estudiados para ciertas subfamilias de la familia exponencial. En este capítulo presentamos una extensión de dichos resultados a cualquier familia exponencial. Describimos, en primer lugar, procedimientos específicos para la obtención de polinomios ortogonales en familias exponenciales. Utilizando la teoría descrita en el capítulo primero, observamos como los desarrollos de los estimadores pueden obtenerse conociendo tan solo las derivadas de la función paramétrica a estimar. El procedimiento para la obtención de dichos estimadores es totalmente algorítmico y ha sido implementado en el lenguaje de programación MAPLE V R-4.

También se obtienen ciertas propiedades de las familias de polinomios ortogonales que serán útiles para obtener, posteriormente, propiedades de los estimadores.

Estudiamos el comportamiento asintótico de los estimadores obtenidos, destacando el teorema 2.7.2 que caracteriza los límites en distribución de los estimadores insesgados. Este resultado generaliza al ya obtenido por López y Castaño (1999) [31] que obtuvieron un resultado similar para el caso particular de la familia exponencial con varianza cuadrática. Además, hacemos notar la diferencia de comportamiento asintótico entre el estimador insesgado y el de máxima verosimilitud en determinadas circunstancias.

A continuación, mostramos que los desarrollos ortogonales permiten obtener fácilmente cotas inferiores para la varianza. Observamos como en cualquier familia exponencial uniparamétrica el límite de las cotas de Battacharyya es la varianza del estimador $umvue$. Este resultado extiende el de Khan (1984) [25] que probó esta misma propiedad para el caso de la familia exponencial

con función de varianza cuadrática.

Los desarrollos son también útiles para la comparación entre estimadores. En particular, obtenemos expresiones para la deficiencia asintótica esperada con respecto al estimador de máxima verosimilitud, que coincide con la obtenida por otros métodos por Hwang y Hu (1990) [23]. También obtenemos el límite de la eficiencia del riesgo, ver teorema 2.9.3.

Para finalizar este capítulo, mostramos como los resultados expuestos pueden extenderse al caso biparamétrico, estudiando desarrollos para la distribución $N(\mu, \sigma^2)$ en el caso en que ambos parámetros sean desconocidos.

El tercer capítulo muestra como la teoría desarrollada puede aplicarse a familias que no verifican las condiciones de regularidad usuales pues el soporte de dichas distribuciones depende del parámetro. Comenzamos haciendo un estudio en la familia $U(0, \theta)$. Al final del mismo mostramos como los resultados obtenidos pueden extenderse a otras familias truncadas.

En primer lugar hacemos un estudio de las propiedades de los polinomios ortogonales asociados a la distribución $U(0, \theta)$.

Contrariamente a lo que ocurre en la familia exponencial, para la distribución $U(0, \theta)$ es fácil obtener una expresión explícita para los estimadores insesgados. Sin embargo, daremos sus desarrollos ortogonales pues a partir de ellos veremos como es posible obtener numerosas propiedades de los mismos, tales como una familia de cotas inferiores para la varianza, semejante a las de Bhattacharyya para la familia exponencial, y las propiedades asintóticas reflejadas en el teorema 3.6.3.

Por último, en el cuarto capítulo estudiamos situaciones que se apartan de la teoría clásica. Así, consideraremos situaciones en que la muestra es censurada

(censuramiento tipo II) y otras en las que no disponemos de todas las observaciones, tal como es el caso en que se observan tan solo los valores records.

En el caso de observaciones censuradas es posible considerar casos en el que la distribución de los datos no sea continua. En particular, hacemos un estudio de los estimadores basados en la media winsorizada de datos procedentes de una distribución geométrica. Para ello previamente hemos obtenido en el teorema 4.4.1 la distribución de la media winsorizada en la familia de distribuciones geométrica.

Con la presente memoria no se agota en modo alguno el estudio y las aplicaciones de los desarrollos asociados a los estimadores. Por el contrario se abren ciertas líneas de investigación futura que parecen prometedoras. Destacamos entre ellas:

- La extensión de los resultados obtenidos a familias multiparamétricas.
- El uso de los desarrollos para la obtención de aproximaciones a distribuciones de probabilidad.
- La posibilidad de reducción del sesgo y el error cuadrático en procedimientos bootstrap paramétricos.
- Obtención de transformadas inversas de Laplace a partir de desarrollos en la familia gamma.

Índice

1	Generalidades	11
1.1	Introducción	11
1.2	Estimación p -insesgada	11
1.3	Desarrollos Ortogonales	14
1.4	Cotas inferiores para la varianza	19
1.5	Polinomios ortogonales: Métodos de construcción	20
1.5.1	Método I: Ortogonalización de Gram-Schmidt	20
1.5.2	Método II	21
1.5.3	Método III	23
1.5.4	Método IV	25
2	Estimación en la familia NEF	27
2.1	Introducción	27
2.2	Familia NEF: Definición y propiedades	28
2.3	Polinomios asociados a la familia NEF	31
2.4	Polinomios ortogonales en la familia NEF	36
2.5	Estimación en la familia NEF	38
2.5.1	Estimación de la varianza del UMVUE	50

2.6	Implementación Numérica	51
2.7	Propiedades asintóticas de los estimadores insesgados	54
2.8	Cotas inferiores para la varianza de estimadores insesgados	66
2.9	Comparaciones con el estimador de máxima verosimilitud	67
2.10	Aplicación a la distribución Normal	76
3	Estimación en familias no regulares	81
3.1	Introducción	81
3.2	Estimación en distribuciones truncadas	82
3.3	Polinomios Ortogonales	82
3.4	Propiedades límites de los polinomios ortogonales	84
3.5	Estimación insesgada	88
3.6	Propiedades límites de los estimadores insesgados	90
3.7	Cotas para la varianza de los estimadores insesgados	95
3.8	Comparación con el estimador de máxima verosimilitud	96
3.9	Generalización	97
4	Otras aplicaciones	102
4.1	Introducción	102
4.2	Muestras censuradas	103
4.3	Observaciones récords	109
4.4	Distribución Geométrica	112
4.5	Distribución χ^2 no centrada	120
4.6	Estimación insesgada	121
4.7	Propiedades Asintóticas	131

ÍNDICE 10

Bibliografía 134

Capítulo 1

Generalidades

1.1 Introducción

En este capítulo presentamos la teoría básica sobre la cual se fundamenta el resto de la memoria. Esencialmente, estudiamos espacios de estimadores de funciones paramétricas. En dichos espacios, obtenemos desarrollos ortogonales para los estimadores y mostramos como ciertas propiedades básicas de estos pueden obtenerse a partir de dichos desarrollos.

1.2 Estimación p -insesgada

Sea X_1, X_2, \dots, X_n una muestra aleatoria simple procedente de una familia paramétrica con función de densidad $f(x, \theta)$ respecto a cierta medida σ -finita ν y $\theta \in \Theta$ un parámetro desconocido. Por simplicidad supondremos que θ es un parámetro real y que las observaciones son unidimensionales, aun-

que hemos de tener en cuenta que muchos conceptos aquí expuestos pueden extenderse al caso multiparamétrico tal como se verá en el capítulo 2. Sea $S_n = S_n(X_1, X_2, \dots, X_n)$ un estadístico cuya función de densidad la denotaremos por $f_n(s, \theta)$.

Para cada $\theta \in \Theta$ y $p \geq 1$, definimos los siguientes espacios de funciones

$$L_{\theta,n}^p = \{T : E_{\theta}[|T(S_n)|^p] < \infty\}, \quad (1.2.1)$$

es decir, el espacio $L_{\theta,n}^p$ está formado por todas las funciones medibles del estadístico S_n , para las que $|T(S_n)|^p$ es integrable con respecto a la densidad $f_n(s, \theta)$. En $L_{\theta,n}^p$ dos funciones son iguales si son iguales salvo conjunto de medida ν -nula. Para cada $p > 1$, $L_{\theta,n}^p$ es un espacio de Banach considerando la norma

$$\|T\|_{p,\theta} = E_{\theta}[|T(S_n)|^p]^{\frac{1}{p}}. \quad (1.2.2)$$

Además para $p = 2$ se tiene que dicha norma procede de un producto escalar, definido por

$$\langle T_1, T_2 \rangle_{\theta} = E_{\theta}[T_1(S_n)T_2(S_n)],$$

con lo cual, para cada $\theta \in \Theta$, $L_{\theta,n}^2$ es un espacio de Hilbert.

Para cada Θ_0 subconjunto de Θ , con interior no vacío, se define

$$\mathcal{L}_n^p(\Theta_0) = \bigcap_{\theta \in \Theta_0} L_{\theta,n}^p,$$

que, por ser intersección de espacios vectoriales, es un espacio vectorial.

Vamos a considerar funciones paramétricas a las que se les puede asociar un elemento de $\mathcal{L}_n^p(\Theta_0)$, surgiendo así los siguientes conceptos:

Definición 1.2.1 Una función paramétrica $h(\theta)$ se dirá p -estimable en Θ_0 , para un tamaño muestral n , si existe una función $T_n \in \mathcal{L}_n^p(\Theta_0)$ verificando

$$E_\theta[T_n(S_n)] = h(\theta), \quad \forall \theta \in \Theta_0.$$

Así, diremos que $T_n(S_n)$ es un estimador p -insesgado de $h(\theta)$ en Θ_0 basado en S_n .

El conjunto de todas las funciones p -estimables en Θ_0 con estimadores basados en S_n lo notaremos por $\mathcal{U}_n^p(\Theta_0)$, es decir

$$\mathcal{U}_n^p(\Theta_0) = \{h : \Theta_0 \rightarrow \mathbb{R} \mid \exists T_n \in \mathcal{L}_n^p(\Theta_0), E_\theta[T_n(S_n)] = h(\theta), \forall \theta \in \Theta_0\}.$$

En particular si $p = 1$ y $\Theta_0 = \Theta$, diremos que $h(\theta)$ es una función estimable para el tamaño de muestra considerado y T_n será pues, un estimador insesgado de $h(\theta)$. En el caso que S_n sea suficiente y completo para la familia de distribuciones considerada, $p \geq 2$ y $\Theta_0 = \Theta$, se dirá que $h(\theta)$ es una función umvu-estimable y $T_n(S_n)$ será el umvue de $h(\theta)$, notándose por $T_n(S_n) = \text{UMVUE}_n(h(\theta))$. Si Θ_0 es estrictamente un subconjunto de Θ , $T_n(S_n)$ se denominará estimador umvue en Θ_0 y la correspondiente función $h(\theta)$ se dirá umvu-estimable en Θ_0 .

La motivación para la introducción del concepto de función estimable en un subconjunto Θ_0 del espacio paramétrico se debe al hecho de que en determinadas circunstancias de interés no es posible encontrar un estimador insesgado en todo Θ , pero si en determinados subconjuntos del espacio paramétrico. Por ejemplo, es bien sabido que no existe un estimador insesgado (en el sentido clásico) basado en la media muestral \bar{X}_n para el parámetro $\frac{1}{\lambda}$ de la distribución de Poisson. Sin embargo, dado $\lambda_0 > 0$ arbitrario, el estimador

$$T_n(\bar{X}_n) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \lambda_0 n^k} C_k^{(n\lambda_0)}(\bar{X}_n), \quad (1.2.3)$$

donde $C_k^{(\alpha)}$ el k -ésimo polinomio de Charlier

$$C_k^{(\alpha)}(s) = \sum_{j=0}^k \binom{k}{j} \binom{s}{j} j! (-\alpha)^{k-j}, \quad (1.2.4)$$

es un estimador insesgado para $\frac{1}{\lambda}$ en $\Theta_0 = (0, 2\lambda_0)$, ver López et al. (1995) [33]. El estimador dado en (1.2.3) puede ser utilizado en situaciones en las que interese obtener una estimación insesgada de $\frac{1}{\lambda}$, cuando se dispone de la información adicional de que el verdadero valor del parámetro pertenece a un cierto intervalo acotado $(0, 2\lambda_0)$, λ_0 conocido.

1.3 Desarrollos Ortogonales

En la sección 1.2 habíamos mencionado que para cada $\theta \in \Theta$, $L_{\theta,n}^2$ tiene estructura de espacio de Hilbert, siendo además separable, Kingman y Taylor (1966) [26], por tanto es posible extraer un conjunto ortonormal completo de funciones.

Sea $\{\phi_{k,n}(s, \theta)\}_{k \geq 0}$ dicho sistema ortonormal, sin pérdida de generalidad se puede considerar $\phi_{0,n}(s, \theta) = 1$. Cualquier función $T_n \in \mathcal{L}_{\theta,n}^2$ admite un desarrollo en serie en términos de dicho sistema ortonormal de la forma:

$$T_n \sim \sum_{k=0}^{\infty} a_{k,n}(\theta) \phi_{k,n}(\cdot, \theta), \quad (1.3.1)$$

convergente en media cuadrática, es decir,

$$\lim_{m \rightarrow \infty} \|T_n - \sum_{k=0}^m a_{k,n}(\theta) \phi_{k,n}(\cdot, \theta)\|_{2,\theta} = 0,$$

con

$$a_{k,n}(\theta) = \langle T_n, \phi_{k,n}(\cdot, \theta) \rangle_\theta,$$

los coeficientes de Fourier correspondientes.

La serie de Fourier de $T_n \in L^2_{\theta,n}$, respecto de cualquier conjunto ortonormal, es convergente pues $L^2_{\theta,n}$ es completo. Se trata de una convergencia en el sentido de la norma definida en (1.2.2), no de convergencia puntual. Debido a las propiedades anteriores, se tiene el siguiente resultado sobre la expresión de la varianza del estimador basado en S_n de una función paramétrica.

Teorema 1.3.1 *Dada una función $h(\theta)$ p -estimable en Θ_0 con $p \geq 2$ y $T_n(S_n)$ su correspondiente estimador p -insesgado en Θ_0 , entonces T_n admite un desarrollo de la forma (1.3.1), válido para todo $\theta \in \Theta_0$, verificando*

$$(i) \ a_{0,n}(\theta) = h(\theta), \quad \forall \theta \in \Theta_0.$$

$$(ii) \ Var_\theta(T_n(S_n)) = \sum_{k=1}^{\infty} a_{k,n}^2(\theta), \quad \forall \theta \in \Theta_0.$$

Demostración.

Si $\{\phi_{k,n}(s, \theta)\}_{k \geq 0}$ es un sistema ortonormal de funciones en $L^2_{\theta,n}$, dado que $T_n \in L^2_{\theta,n}$, $\forall \theta \in \Theta_0$, es obvio que T_n admite un desarrollo tipo (1.3.1).

En el caso $p > 2$, al trabajar en espacios de probabilidad, se verifica la relación $L^p_{\theta,n} \subset L^2_{\theta,n}$ para cada $\theta \in \Theta_0$, en particular $T_n \in \mathcal{L}^2_{\theta,n}$ y por tanto admite dicho desarrollo.

Además:

$$a_{0,n}(\theta) = \langle T_n, \phi_{0,n}(\cdot, \theta) \rangle_\theta = \langle T_n, 1 \rangle_\theta = E_\theta[T_n(S_n)] = h(\theta) \quad y$$

$$\text{Var}_\theta(T_n(S_n)) = E_\theta[T_n^2(S_n)] - h^2(\theta) = \|T_n\|_{2,\theta}^2 - a_{0,n}^2(\theta), \quad \forall \theta \in \Theta_0.$$

Como $\|T_n\|_{2,\theta}^2 = \sum_{k=0}^{\infty} a_{k,n}^2(\theta)$, para cada $\theta \in \Theta_0$, se tiene

$$\text{Var}_\theta(T_n(S_n)) = \sum_{k=1}^{\infty} a_{k,n}^2(\theta), \quad \forall \theta \in \Theta_0. \quad (1.3.2)$$

□

Dada una función medible cualquiera $T_n(S_n)$, de una manera formal se le puede asociar una sucesión de funciones $\{a_{k,n}(\theta)\}_{k \geq 0}$ que son los coeficientes de Fourier

$$a_{k,n}(\theta) = \int_{\mathbb{R}} T_n(s) \phi_{k,n}(s; \theta) f_n(s, \theta) d\nu_n(s), \quad \forall \theta \in \Theta_0, \quad (1.3.3)$$

con $f_n(s, \theta)$ la densidad de S_n respecto de la medida ν_n .

Si las funciones $a_{k,n}(\theta)$ para cada $k \geq 0$, están bien definidas en Θ_0 y verifican que

$$\sum_{k=1}^{\infty} a_{k,n}^2(\theta) < \infty, \quad \forall \theta \in \Theta_0,$$

entonces $T_n \in \mathcal{L}_n^2(\Theta_0)$ y $h(\theta) = E_\theta[T_n(S_n)] \in \mathcal{U}_n^2(\Theta_0)$.

Si $\{\phi_{k,n}(s, \theta)\}_{k \geq 0}$ es un sistema ortogonal de funciones se tiene que :

- Los coeficientes del desarrollo de Fourier de $T_n(S_n)$, respecto de dicho sistema son

$$a_{k,n}(\theta) = \frac{\langle T_n(s), \phi_{k,n}(s, \theta) \rangle_\theta}{\|\phi_{k,n}\|_{2,\theta}^2}. \quad (1.3.4)$$

- La expresión de la varianza de T_n es

$$\text{Var}_\theta(T_n(S_n)) = \sum_{k=1}^{\infty} a_{k,n}^2(\theta) \|\phi_{k,n}\|_{2,\theta}^2, \quad \forall \theta \in \Theta_0,$$

y una cota inferior de orden k para la varianza de T_n es

$$b_{k,n}(T_n, \theta) = \sum_{j=1}^k a_{j,n}^2(\theta) \|\phi_{j,n}\|_{2,\theta}^2. \quad (1.3.5)$$

En general, dada una función medible cualquiera T_n , si los coeficientes de Fourier asociados $\{a_{k,n}(\theta)\}_{k \geq 0} \quad \forall \theta \in \Theta_0$, dados por (1.3.3) verifican que

$$\sum_{k=1}^{\infty} a_{k,n}^2(\theta) \|\phi_{k,n}\|_{2,\theta}^2 < \infty, \quad \forall \theta \in \Theta_0,$$

entonces $T_n \in \mathcal{L}_n^2(\Theta_0)$ y $h(\theta) = E_\theta[T_n(S_n)] \in \mathcal{U}_n^2(\Theta_0)$.

Teorema 1.3.2 *Si S_n es suficiente y completo, los espacios vectoriales $\mathcal{U}_n^p(\Theta_0)$ y $\mathcal{L}_n^p(\Theta_0)$ para cada $p \geq 1$ son isomorfos.*

Demostración.

La aplicación que a cada $T_n \in \mathcal{L}_n^p(\Theta_0)$ le asigna la función $h \in \mathcal{U}_n^p(\Theta_0)$, definida como $h(\theta) = E_\theta[T_n(S_n)]$ es lineal y, debido a la completitud del estadístico S_n , es biyectiva.

□

Teorema 1.3.3 *Sea $\{h_m(\theta)\}_{m \geq 1}$ una sucesión de funciones p -estimables en Θ_0 con $p \geq 1$ y $T_m(S_n)$ sus correspondientes estimadores p -insesgados en Θ_0 , tal que para cierta función $T \in \mathcal{L}_n^p(\Theta_0)$ se verifica*

$$\lim_{m \rightarrow \infty} \|T_m - T\|_{p,\theta} = 0, \quad \forall \theta \in \Theta_0,$$

entonces la función $h(\theta)$, definida como

$$h(\theta) = \lim_{m \rightarrow \infty} h_m(\theta)$$

pertenece a $\mathcal{U}_n^p(\Theta_0)$, con T su correspondiente estimador p -insesgado en Θ_0 .

Demostración.

Puesto que la convergencia en media de orden p , implica la convergencia en media de orden 1,

$$\lim_{m \rightarrow \infty} \int_{\mathbb{R}} T_m(s) f_n(s, \theta) d\nu_n(s) = \int_{\mathbb{R}} T(s) f_n(s, \theta) d\nu_n(s), \quad \forall \theta \in \Theta_0,$$

o sea

$$\lim_{m \rightarrow \infty} h_m(\theta) = h(\theta) = E_\theta[T(S_n)], \quad \forall \theta \in \Theta_0$$

y por hipótesis $T \in \mathcal{L}_n^p(\Theta_0)$, por tanto $h \in \mathcal{U}_n^p(\Theta_0)$ lo que demuestra el resultado. □

Siempre que sea posible, consideraremos el caso particular en que el estadístico $S_n = S_n(X_1, \dots, X_n)$ sea suficiente y completo para θ .

Si $h(\theta)$ es una función 2-estimable en Θ_0 , y $T_n(S_n)$ un estimador 2-insesgado de $h(\theta)$ en Θ_0 , entonces $T_n(S_n)$ es una función del suficiente y completo que posee varianza finita para cada $\theta \in \Theta_0$, ligeras modificaciones en los teoremas de Blackwell-Rao y Lehmann-Scheffé, prueban que $T_n(S_n)$ es el único estimador umvue de $h(\theta)$ en Θ_0 .

En el caso particular en que la función característica del suficiente y completo sea analítica en $t = 0$, o bien si $E_\theta(S_n^k) < \infty$ para todo $k \geq 0$, y $\theta \in \Theta_0$, se tiene que el conjunto de las funciones polinómicas es denso en $\mathcal{L}_{\theta,n}^2$ para cada $\theta \in \Theta_0$, Abbey y David (1970) [1], con lo cual a partir del conjunto denso $\{1, s, s^2, \dots\}$ podemos construir por el método de Gram-Schmidt un sistema

completo de polinomios ortogonales. Variantes del método de Gram-Schmidt y otros procedimientos para la obtención del sistema de polinomios ortogonales se verán en el capítulo siguiente. Como consecuencia del Teorema 1.3.3 el espacio $\mathcal{U}_n^2(\Theta_0)$ está formado por las funciones del parámetro que son combinaciones lineales finitas de los momentos del estadístico suficiente y completo y aquellas funciones que son límites de estas combinaciones lineales en la topología inducida en $\mathcal{U}_n^2(\Theta_0)$ por el isomorfismo obtenido en el Teorema 1.3.2.

1.4 Cotas inferiores para la varianza

Una de las aplicaciones inmediatas de los desarrollos de los estimadores en términos de polinomios ortonormales es la obtención de cotas inferiores para la varianza de estimadores insesgados, tal como se refleja en el siguiente resultado:

Teorema 1.4.1 *Sea $h \in \mathcal{U}_n^2(\Theta_0)$ y $V_n = V(X_1, \dots, X_n)$ un estimador insesgado de $h(\theta)$ en Θ_0 , entonces para cualquier $m \geq 1$*

$$\text{Var}_\theta(V_n) \geq B_{m,n}(h, \theta) = \sum_{k=1}^m a_{k,n}^2(\theta), \quad \forall \theta \in \Theta_0, \quad (1.4.1)$$

donde $a_{k,n}(\theta)$ son los coeficientes de Fourier correspondientes al estimador umvue de $h(\theta)$ en Θ_0 .

Demostración.

Sea $T_n(S_n)$ el estimador umvue de $h(\theta)$ en Θ_0 , por tanto se tiene que, para cualquier $V_n = V(X_1, \dots, X_n)$ estimador insesgado de $h(\theta)$ en Θ_0 ,

$$\text{Var}_\theta(V_n) \geq \text{Var}_\theta[T_n(S_n)], \quad \forall \theta \in \Theta_0,$$

y de aquí, aplicando el Teorema 1.3.1, obtenemos

$$\text{Var}_\theta(V_n) \geq \sum_{k=1}^{\infty} a_{k,n}^2(\theta) \geq \sum_{k=1}^m a_{k,n}^2(\theta), \quad \forall \theta \in \Theta_0, m \geq 1.$$

Notese que las cotas obtenidas en (1.4.1), coinciden con las definidas en (1.3.5).

□

Observación 1.4.1 *El teorema anterior proporciona un resultado bastante general, válido tanto para distribuciones regulares como para distribuciones no regulares o que no pertenezcan a la familia exponencial.*

1.5 Polinomios ortogonales: Métodos de construcción

Con objeto de aplicar la teoría desarrollada en las secciones anteriores de este capítulo daremos diferentes métodos para la construcción de polinomios ortogonales asociados a una función de densidad. Algunos de los métodos que a continuación se describen tienen validez general para cualquier familia de distribuciones.

1.5.1 Método I: Ortogonalización de Gram-Schmidt

Si $f(x, \mu)$ es una familia tal que $E_\mu[X^k] < \infty \quad \forall k \geq 0, \quad \forall \mu \in \Omega$, se tiene que el conjunto $\{1, x, x^2, \dots\}$ es denso en L_μ^2 , por tanto, aplicando el método

de ortogonalización de Gram-Schmidt, podemos construir un sistema completo de polinomios ortogonales

$$P_0(x, \mu) = 1,$$

$$P_{k+1}(x, \mu) = x^{k+1} - \sum_{j=0}^k d_{k,j}(\mu) P_j(x, \mu),$$

siendo

$$d_{k,j}(\mu) = \frac{\langle x^{k+1}, P_j(x, \mu) \rangle_\mu}{\|P_j\|_{2,\mu}^2}.$$

Este método puede aplicarse a cualquier familia paramétrica.

1.5.2 Método II

Este método es válido para cualquier familia de distribuciones y se basa en la siguiente proposición:

Proposición 1.5.1 *Sea m_k el momento de orden k respecto al origen de la distribución en estudio. Denotamos por $\mathcal{M}_k(x)$ a la matriz:*

$$\mathcal{M}_k(x) = \begin{pmatrix} m_0 & m_1 & \dots & m_k \\ m_1 & m_2 & \dots & m_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_{k-1} & m_k & \dots & m_{2k-1} \\ 1 & x & \dots & x^k \end{pmatrix} \quad (1.5.1)$$

y por Δ_j al adjunto de x^j . Si para todo j , $\Delta_j \neq 0$, los polinomios

$$P_k(x, \mu) = \frac{\det \mathcal{M}_k(x)}{\Delta_k},$$

forman un sistema de polinomios mónicos ortogonales respecto a la densidad considerada.

Demostración.

Por definición, $P_k(x)$, $k \geq 0$ son mónicos. Para demostrar que son ortogonales, hemos de demostrar que

$$\langle P_j(x), P_k(x) \rangle = 0, \quad \forall j < k,$$

lo cual es equivalente a demostrar que

$$\langle x^j, P_k(x) \rangle = 0, \quad \forall j < k,$$

ver Chihara (1978) [14]. Desarrollando a partir de la última fila del determinante de la matriz $\mathcal{M}_k(x)$ dado en (1.5.1), podemos expresar $P_k(x)$ como

$$P_k(x) = \sum_{l=0}^k (-1)^{k-l} x^l \frac{\Delta_l}{\Delta_k}.$$

Por tanto, para todo $j < k$

$$\langle x^j, P_k(x) \rangle = \frac{1}{\Delta_k} E \left[\sum_{l=0}^k (-1)^{k-l} x^{j+l} \Delta_l \right] = \frac{1}{\Delta_k} \sum_{l=0}^k (-1)^{k-l} m_{j+l} \Delta_l.$$

Esta última expresión es el desarrollo de un determinante con dos filas iguales, por tanto es 0.

□

Para cualquier familia de distribuciones, los primeros polinomios calculados a partir de este método tienen la expresión:

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x - \frac{m_1}{m_0}, \\ P_2(x) &= x^2 + \frac{m_1 m_2 - m_3}{m_2 - m_1^2} x + \frac{m_1 m_3 - m_2^2}{m_2 - m_1^2}. \end{aligned} \quad (1.5.2)$$

1.5.3 Método III

Este método también es válido para cualquier familia de distribuciones.

Denotamos por $P_j(x) = \sum_{l=0}^j c_{jl}x^l$, $j \geq 0$. Podemos suponer, sin pérdida de generalidad, que dichos polinomios son mónicos, es decir, $c_{jj} = 1$, $\forall j \geq 0$.

Explícitamente, los $k + 1$ primeros polinomios podemos escribirlos

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= c_{10} + x, \\ P_2(x) &= c_{20} + c_{21}x + x^2, \\ &\dots = \dots\dots\dots \\ P_k(x) &= c_{k0} + c_{k1}x + c_{k2}x^2 + \dots + x^k. \end{aligned}$$

Estas igualdades pueden ser expresadas matricialmente como

$$\mathcal{P}_k(x) = C_k \mathcal{X}_k(x),$$

siendo C_k la matriz de coeficientes

$$C_k = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ c_{10} & 1 & 0 & \dots & 0 \\ c_{20} & c_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k0} & c_{k1} & c_{k2} & \dots & 1 \end{pmatrix}, \quad (1.5.3)$$

$\mathcal{P}_k(x)$ el vector de polinomios

$$\mathcal{P}_k^t(x) = \left(P_0(x), P_1(x), P_2(x), \dots, P_k(x) \right),$$

y $\mathcal{X}_k(x)$ el vector de potencias de x

$$\mathcal{X}_k(x) = \left(1, x, x^2, \dots, x^k \right).$$

Multiplicando por las traspuestas tenemos

$$\mathcal{P}_k(x) \cdot \mathcal{P}_k^t(x) = C_k \mathcal{X}_k(x) \mathcal{X}_k^t(x) C_k^t \quad (1.5.4)$$

y tomando esperanzas, en ambos miembros de (1.5.4) e imponiendo que $\{P_k(x)\}_{k \geq 0}$ es un sistema de polinomios ortogonales, es decir

$$E[P_j(x)P_l(x)] = 0 \quad \text{si } j \neq l$$

y

$$E[P_j^2(x)] = \|P_j\|_2^2,$$

obtenemos

$$\text{diag} \left(1, \|P_1\|_2^2, \|P_2\|_2^2, \dots, \|P_k\|_2^2 \right) = C_k \mathcal{M}_k C_k^t,$$

donde

$$\mathcal{M}_k = (m_{i,j})_{i,j=0,1,\dots,k}$$

con $m_{i,j} = E[X^{i+j}]$.

A partir de aquí, un procedimiento para obtener los polinomios ortogonales hasta el grado k consiste en construir la matriz auxiliar

$$\mathcal{A} = (\mathcal{M}_k \mid \mathcal{I}_{k+1}),$$

de dimensión $(k+1) \times 2(k+1)$, obtenida al solapar la matriz identidad de orden $(k+1)$ a la derecha de \mathcal{M}_k . Aplicando el método de eliminación de Gauss por filas a la matriz \mathcal{A} puede transformarse en una matriz de la forma

$$(\mathcal{D} \mid \mathcal{T})$$

con \mathcal{D} una matriz diagonal y \mathcal{T} triangular inferior, cuyos elementos diagonales son la unidad. Se tiene entonces, que los elementos de la fila i -ésima de \mathcal{T} son (en orden ascendente) los coeficientes del i -ésimo polinomio mónico ortogonal y que $(\mathcal{D})_{ii}$ es la norma de dicho polinomio.

1.5.4 Método IV

Es en realidad una variante del método anterior.

Suponemos que el j -ésimo polinomio ortonormal lo expresamos de la forma

$$O_j(x) = \sum_{r=0}^j \rho_{jr} (s - m_1)^r, \quad j \geq 0$$

Usando notación matricial tenemos

$$\begin{bmatrix} O_0 \\ O_1 \\ \vdots \\ O_k \end{bmatrix} = \begin{bmatrix} \rho_{00} & 0 & 0 & \cdots & 0 \\ \rho_{10} & \rho_{11} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k0} & \rho_{k1} & \rho_{k2} & \cdots & \rho_{kk} \end{bmatrix} \begin{bmatrix} 1 \\ x - m_1 \\ \vdots \\ (x - m_1)^k \end{bmatrix},$$

o bien

$$O_k(x) = \mathcal{R}_k \mathcal{X}_k(x). \tag{1.5.5}$$

Notemos que \mathcal{R}_k es una matriz triangular inferior formada por los coeficientes de los polinomios ortonormales de grado menor o igual que k . De (1.5.5), tenemos

$$O_k(x) O_k^t(x) = \mathcal{R}_k \mathcal{X}_k(x) \mathcal{X}_k^t(x) \mathcal{R}_k^t \tag{1.5.6}$$

Tomando valor esperado, como los polinomios son ortonormales,

$$Id = \mathcal{R}_k \mathcal{M}_k \mathcal{R}_k^t$$

de donde

$$\mathcal{M}_k = (\mathcal{R}_k)^{-1} (\mathcal{R}_k^t)^{-1},$$

siendo $\mathcal{M}_k = (E[X - m_1]^{i+j})_{i,j=0,\dots,k}$.

Entonces, un método para obtener la matriz \mathcal{M}_k es el siguiente:

1. Calcular la matriz de momentos centrales \mathcal{M}_k .
2. Encontrar la descomposición de Cholesky \mathcal{M}_k , es decir, encontrar \mathcal{E}_k tal que $\mathcal{M}_k = \mathcal{E}_k \mathcal{E}_k^t$.
3. $\mathcal{R}_k = \mathcal{E}_k^{-1}$.

Capítulo 2

Estimación en la familia NEF

2.1 Introducción

En este capítulo vamos a aplicar la teoría general, vista en el anterior, al caso de familias exponenciales. Desarrollamos, en primer lugar, procedimientos específicos para la obtención de polinomios ortogonales en familias exponenciales. Construimos estimadores insesgados mediante desarrollos ortogonales, a través de un procedimiento totalmente algorítmico que ha sido implementado en un lenguaje de programación. También se obtienen ciertas propiedades de las familias de polinomios ortogonales que nos permiten estudiar, fácilmente, propiedades de los estimadores.

2.2 Familia NEF: Definición y propiedades

Sea F_0 una función de distribución que posee función generatriz de momentos en un entorno de cero.

Consideramos la función

$$\psi(\theta) = \log \int \exp(\theta x) dF_0(x), \quad \theta \in \Theta,$$

con $\Theta \neq \emptyset$ el mayor intervalo para el cual $\psi(\theta)$ está definida. Entonces la familia paramétrica de distribuciones $\{F_\theta, \theta \in \Theta\}$ definida por

$$dF_\theta(x) = \exp\{\theta x - \psi(\theta)\} dF_0(x) \quad (2.2.1)$$

es una familia exponencial natural de distribuciones (NEF).

La función $\psi(\theta)$ se denomina función de cumulantes de la familia. El cumulante j -ésimo es

$$\kappa_j = \psi^{(j)}(0).$$

La función generatriz de momentos viene dada por:

$$M_\theta(t) = \exp\{\psi(t + \theta) - \psi(\theta)\}, \quad (2.2.2)$$

de donde, la media y varianza de la familia son, respectivamente,

$$\mu = E_\theta(X) = \psi'(\theta),$$

$$V(\mu) = \text{Var}_\theta(X) = \psi''(\theta).$$

Puesto que $V(\mu) = \psi''(\theta) > 0$, se tendrá que la media, $\mu = \psi'(\theta)$, es una función 1 - 1 del parámetro θ , por lo que en algunas ocasiones se utilizará

μ como el parámetro de la familia, cuyo dominio viene dado por $\Omega = \psi'(\Theta)$ denominado espacio de medias.

Al par $(\Omega, V(\mu))$ se le denomina función de varianza de la NEF, la cuál es de gran importancia pues caracteriza la NEF dentro de todas las familias de distribuciones NEF, (aunque no en una clase más amplia de distribuciones).

A continuación consideraremos una serie de propiedades que verifica la familia de distribuciones exponencial natural, útiles a lo largo del desarrollo de la presente memoria. Para más detalles sobre familias exponenciales, consúltense los libros de Brown (1968) [13] y Barndorf-Nielsen (1978) [9].

Proposición 2.2.1 *Sea X una variable aleatoria cuya distribución viene dada por (2.2.1), entonces:*

(i) *Existen y son finitos todos los momentos no centrados, es decir*

$$m_k(\theta) = E_\theta[X^k] < \infty \quad k = 0, 1, 2, \dots$$

(ii) *Se verifica la siguiente relación de recurrencia*

$$m_{k+1}(\theta) = D_\theta m_k(\theta) + m_1(\theta)m_k(\theta), \quad (2.2.3)$$

siendo $m_0(\theta) = 1$, $m_1(\theta) = \psi'(\theta) = \mu$.

Proposición 2.2.2 *Sea X una variable con distribución dada en (2.2.1). Entonces*

(i) *La función generatriz de momentos centrales puede expresarse*

$$M_c(t, \theta) = \exp\{-\mu(\theta)t\} M_\theta(t). \quad (2.2.4)$$

(ii) Se verifica la siguiente relación de recurrencia

$$\mu_{k+1}(\theta) = \mu'_k(\theta) + \mu_1(\theta)\mu_k(\theta), \quad (2.2.5)$$

siendo $\mu_1(\theta) = 0$, $\mu_k(\theta) = E_\theta[(X - \mu)^k]$

Demostración.

(i) El momento central de orden k puede expresarse:

$$\begin{aligned} E_\theta[(X - \mu)^k] &= \sum_{j=0}^k \binom{k}{j} (-\mu)^{k-j} E_\theta[X^j] = \\ &= \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-\mu)^{k-j} m_j(\theta), \end{aligned}$$

por tanto

$$\begin{aligned} M_c(t, \theta) &= \sum_{k=0}^{\infty} \frac{1}{k!} E_\theta[(X - \mu)^k] t^k \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k \mu^k t^k}{k!} \sum_{k=0}^{\infty} \frac{m_k(\theta) t^k}{k!} \\ &= \exp\{-\mu(\theta)t\} M_\theta(t). \end{aligned}$$

obteniéndose así (2.2.4).

(ii) Teniendo en cuenta la expresión (2.2.2) de la función generatriz de momentos podemos expresar

$$M_c(t, \theta) = \exp\{\psi(\theta + t) - \psi(\theta) - \mu(\theta)t\}. \quad (2.2.6)$$

Como (2.2.6) es una función de tipo exponencial, sus derivadas respecto a t son de la forma:

$$D_t^k M_c(t, \theta) = \Upsilon_k(t, \theta) M_c(t, \theta)$$

donde las funciones Υ_k pueden obtenerse recurrentemente mediante la relación

$$\Upsilon_{k+1}(t, \theta) = D_t \Upsilon_k(t, \theta) + \Upsilon_1(t, \theta) \Upsilon_k(t, \theta) \quad (2.2.7)$$

siendo $\Upsilon_1(t, \theta) = D_t \psi(t + \theta) - \mu(\theta)$.

Debido a que el momento central de orden k se calcula a partir de

$$\mu_k(\theta) = \Upsilon_k(0, \theta),$$

sustituyendo $t = 0$ en (2.2.7) obtenemos

$$\mu_{k+1}(\theta) = \mu'_k(\theta) + \mu_1(\theta) \mu_k(\theta),$$

con $\mu'_k(\theta) = D_t \Upsilon_k(t, \theta)|_{t=0}$.

□

2.3 Polinomios asociados a la familia NEF

En esta sección estudiaremos una familia de polinomios que, de forma natural, tiene asociada la familia de distribuciones en estudio. Sea $f(x, \theta)$ la función de densidad asociada a la familia de distribuciones (2.2.1). Definimos las funciones

$$Q_k(x, \mu) = V^k(\mu) \frac{D_\mu^k f(x, \theta)}{f(x, \theta)}. \quad (2.3.1)$$

Cabe destacar, que en la expresión anterior, se está derivando con respecto a la media, μ , y no con respecto al parámetro natural θ .

Se verifica el siguiente teorema, cuya demostración puede verse en Abbey y David (1970) [1].

Teorema 2.3.1 *Para cada $k \geq 0$, $Q_k(x, \mu)$ es un polinomio de grado k en $(x - \mu)$. El conjunto $\{Q_k(x, \mu)\}_{k \geq 0}$ es denso en L_μ^2 .*

Veremos a continuación que estos polinomios pueden ser generados a través de una relación de recurrencia o de la función generatriz de polinomios.

Lema 2.3.1 *Los polinomios definidos en (2.3.1) verifican la siguiente relación de recurrencia*

$$Q_{k+1}(x, \mu) = \{Q_1(x, \mu) - kV'(\mu)\} Q_k(x, \mu) + V(\mu)D_\mu Q_k(x, \mu), \quad (2.3.2)$$

siendo $Q_0(x, \mu) = 1$ y $Q_1(x, \mu) = x - \mu$.

Demostración.

Teniendo en cuenta la expresión (2.3.1) y que $D_\mu f(x, \theta) = f(x, \theta)Q_1(x, \mu)V^{-1}(\mu)$, obtenemos

$$\begin{aligned} D_\mu^{k+1} f(x, \theta) &= Q_{k+1}(x, \mu)V^{-(k+1)}(\mu)f(x, \theta) = \\ &= \{[-kV'(\mu) + Q_1(x, \mu)]Q_k(x, \mu) + D_\mu Q_k(x, \mu)V(\mu)\} V^{-(k+1)}(\mu)f(x, \theta), \end{aligned}$$

por tanto

$$Q_{k+1}(x, \mu) = \{Q_1(x, \mu) - kV'(\mu)\} Q_k(x, \mu) + V(\mu)D_\mu Q_k(x, \mu),$$

con $Q_0(x, \mu) = 1$ y $Q_1(x, \mu) = x - \mu$.

□

Para calcular la expresión explícita de los polinomios $Q_k(x, \mu)$, necesitaremos el siguiente lema:

Lema 2.3.2 Sea g una función indefinidamente derivable y

$$G_k(\mu) = \frac{D_\mu^k \exp\{-g(\mu)\}}{\exp\{-g(\mu)\}}, \quad k \geq 0, \quad (2.3.3)$$

entonces

$$\sum_{k=0}^{\infty} \frac{G_k(\mu) z^k}{k!} = \exp\{-g(\mu + z) + g(\mu)\}. \quad (2.3.4)$$

Demostración.

Demostrar que

$$\sum_{k=0}^{\infty} \frac{G_k(\mu) z^k}{k!} = \exp\{-g(\mu + z) + g(\mu)\}$$

es equivalente a demostrar que

$$\sum_{k=0}^{\infty} \frac{G_k(\mu) \exp\{-g(\mu)\}}{k!} z^k = \exp\{-g(\mu + z)\}.$$

Sustituyendo (2.3.3), esta expresión puede escribirse

$$\sum_{k=0}^{\infty} \frac{D_\mu^k \exp\{-g(\mu)\}}{k!} z^k = \exp\{-g(\mu + z)\},$$

por tanto, bastará demostrar que

$$D_\mu^k \exp\{-g(\mu)\} = D_z^k \exp\{-g(\mu + z)\} \Big|_{z=0}.$$

Por tratarse de funciones de tipo exponencial, podemos escribir

$$D_z^k \exp\{-g(\mu + z)\} = J_k(z, \mu) \exp\{-g(\mu + z)\} \quad (2.3.5)$$

y que

$$D_\mu^k \exp\{-g(\mu + z)\} = J_k(z, \mu) \exp\{-g(\mu + z)\}. \quad (2.3.6)$$

Sustituyendo $z = 0$ en (2.3.6), obtenemos

$$D_\mu^k \exp\{-g(\mu)\} = J_k(0, \mu) \exp\{-g(\mu)\} = D_z^k \exp\{-g(\mu + z)\} \Big|_{z=0},$$

donde la última igualdad se obtiene haciendo $z = 0$ en (2.3.5).

□

En general, la función generatriz de una familia de polinomios $\{R_k(x)\}_{k \geq 0}$ es

$$\Pi_R(z, x) = \sum_{k=0}^{\infty} \frac{R_k(x)}{k!} z^k.$$

Teorema 2.3.2 *La función generatriz de los polinomios $Q_k^*(x, \mu) = \frac{Q_k(x, \mu)}{V^k(\mu)}$ viene dada por*

$$\Pi_{Q^*}(z, x, \mu) = \exp \{x(\theta(\mu + z) - \theta(\mu)) - (\psi(\theta(\mu + z)) - \psi(\theta(\mu)))\}.$$

Demostración.

Para simplificar denotaremos por $g(\mu) = \psi(\theta(\mu))$ en la expresión de $f(x, \theta)$, dada en (2.2.1).

Utilizando la fórmula de Leibnitz para la derivada de un producto se obtiene:

$$D_\mu^k f(x, \theta) = \sum_{r=0}^k \binom{k}{r} D_\mu^r \exp\{x\theta(\mu)\} D_\mu^{k-r} \exp\{-g(\mu)\}. \quad (2.3.7)$$

Ahora bien, la derivada j -ésima de $\exp\{x\theta(\mu)\}$ es de la forma:

$$D_\mu^j \exp\{x\theta(\mu)\} = \exp\{x\theta(\mu)\} A_j(x, \mu), \quad (2.3.8)$$

donde

$$A_{j+1}(x, \mu) = x\theta'(\mu)A_j(x, \mu) + D_\mu A_j(x, \mu),$$

con $A_0(x, \mu) = 1$.

Análogamente,

$$D_\mu^j \exp\{-g(\mu)\} = \exp\{-g(\mu)\} G_j(\mu), \quad (2.3.9)$$

siendo

$$G_{j+1}(x, \mu) = -g'(\mu)G_j(\mu) + G_j'(\mu),$$

con $G_0(\mu) = 1$.

Sustituyendo estas expresiones en (2.3.7), obtenemos

$$D_\mu^k f(x, \theta) = f(x, \theta) \sum_{r=0}^k \binom{k}{r} A_r(x, \mu) G_{k-r}(x, \mu),$$

de donde

$$Q_k(x, \mu) = V^k(\mu) \frac{D_\mu^k f(x, \theta)}{f(x, \theta)} = V^k(\mu) \sum_{r=0}^k \binom{k}{r} A_r(x, \mu) G_{k-r}(x, \mu),$$

y de aquí

$$\frac{Q_k(x, \mu)}{k!V^k(\mu)} = \sum_{r=0}^k \frac{A_r(x, \mu)}{r!} \frac{G_{k-r}(x, \mu)}{(k-r)!}$$

con lo cual

$$\begin{aligned} \Pi_{Q^*}(z, x, \mu) &= \sum_{k=0}^{\infty} \frac{Q_k(x, \mu)}{k!V^k(\mu)} z^k = \\ &= \sum_{k=0}^{\infty} \frac{A_k(x, \mu)}{k!} z^k \sum_{k=0}^{\infty} \frac{G_k(x, \mu)}{k!} z^k. \end{aligned} \quad (2.3.10)$$

Teniendo en cuenta las expresiones (2.3.8), (2.3.9) y el lema anterior, podemos expresar

$$\sum_{k=0}^{\infty} \frac{A_k(x, \mu)}{k!} z^k = \exp\{x\theta(\mu + z) - x\theta(\mu)\} \quad (2.3.11)$$

y

$$\sum_{k=0}^{\infty} \frac{G_k(x, \mu)}{k!} z^k = \exp\{-g(\mu + z) + g(\mu)\}. \quad (2.3.12)$$

Al sustituir (2.3.11) y (2.3.12) en la expresión (2.3.10) obtenemos el resultado esperado.

□

Corolario 2.3.1 *Los polinomios $Q_k(x, \mu)$, pueden obtenerse como*

$$Q_k(x, \mu) = V^k(\mu) \cdot D_z^k \Pi_{Q^*}(z, x, \mu) \Big|_{z=0}.$$

Ejemplo 2.3.1 *En la siguiente tabla se muestran las funciones generatrices de los polinomios $\{Q_k^*\}_{k \geq 0}$ de la familia exponencial con función de varianza cuadrática (NEF-QVF)*

$f(x, \theta)$	μ	$\Pi_{Q^*}(z, x, \mu)$
$N(\mu, \sigma^2)$	θ	$\exp\left\{\frac{-z^2}{2\sigma^2} + \frac{(x-\mu)z}{\sigma^2}\right\}$
$Ga(\alpha, 1/\theta)$	$\alpha\theta$	$\left(\frac{\mu}{\mu+z}\right)^\alpha \exp\left\{\frac{zx\alpha}{\mu(\mu+z)}\right\}$
$P(\theta)$	θ	$\exp\{-z\} \left(1 + \frac{z}{\mu}\right)^x$
$Bi(N, \theta)$	$N\theta$	$\left(1 + \frac{z}{\mu}\right)^x \left(1 - \frac{z}{N-\mu}\right)^{N-x}$
$BN(r, \theta)$	$\frac{r(1-\theta)}{\theta}$	$\left(1 + \frac{z}{\mu+r}\right)^{-r} \left(1 + \frac{zr}{\mu(\mu+z+r)}\right)^x$
$GHS(\theta)$	$r \cdot \text{tag}(\theta)$	$\left(\frac{r^2+(\mu+z)^2}{r^2+\mu^2}\right)^{-r/2} \exp\left\{\frac{x}{\text{tag}^{-1}\left(\frac{\mu+z}{r}\right)} - \frac{x}{\text{tag}^{-1}\left(\frac{\mu}{r}\right)}\right\}$

2.4 Polinomios ortogonales en la familia NEF

Los diferentes métodos, desarrollados en la sección 1.5, para la construcción de polinomios ortogonales asociados a una función de densidad son aplicables a la familia exponencial. Además, podemos incluir una variante del método I:

- Ortogonalización de la familia $Q_k(x, \mu), k \geq 0$.

Teniendo en cuenta que el sistema de polinomios $Q_k(x, \mu)$ definidos en

(2.3.1) es denso en L^2_μ , ver teorema 2.3.1, podemos construir, a partir de estos y aplicando el método de ortogonalización de Gram-Schmidt, un sistema completo de polinomios ortogonales $P_k(x, \mu)$, $k \geq 0$:

$$P_0(x, \mu) = Q_0(x, \mu) = 1,$$

$$P_{k+1}(x, \mu) = Q_{k+1}(x, \mu) - \sum_{j=0}^k q_{k,j}(\mu) P_j(x, \mu),$$

siendo

$$q_{k,j}(\mu) = \frac{\langle Q_{k+1}(x, \mu), P_j(x, \mu) \rangle_\mu}{\|P_j\|_{2,\mu}^2}.$$

En el capítulo anterior, utilizando el método II, fueron calculados los primeros polinomios para cualquier familia de distribuciones, ver (1.5.2). En particular, para la familia NEF, calculando los momentos a partir de la ecuación (2.2.3) y sustituyendo en (1.5.2), obtenemos

$$P_0(x, \mu) = 1,$$

$$P_1(x, \mu) = x - \mu,$$

$$P_2(x, \mu) = x^2 - (2\mu + V'(\mu))x + (\mu^2 + \mu V'(\mu) - V(\mu)),$$

siendo sus normas

$$\|P_0\|_{2,\mu}^2 = 1,$$

$$\|P_1\|_{2,\mu}^2 = V(\mu),$$

$$\|P_2\|_{2,\mu}^2 = V^2(\mu) (2 + V''(\mu)).$$

2.5 Estimación en la familia NEF

Sea X_1, X_2, \dots, X_n una muestra aleatoria simple procedente de una familia NEF con función de varianza $V(\mu)$, $\mu \in \Omega$. El estadístico $\bar{X}_n = \sum_{i=1}^n \frac{X_i}{n}$ es suficiente y completo para dicha familia y, debido a que la familia NEF es cerrada por convoluciones y combinaciones lineales, su distribución también pertenece a la NEF, siendo su función de varianza $\frac{V(\mu)}{n}$, $\mu \in \Omega$. La densidad de \bar{X}_n , $f_n(s, \theta)$, será de la forma

$$f_n(s, \theta) = \exp\{ns\theta - n\psi(\theta)\},$$

por lo tanto todos los resultados obtenidos en la sección 2.2 pueden ser extendidos para \bar{X}_n . Así,

- La función generatriz de momentos de \bar{X}_n es

$$M_{\bar{X}_n}(t, \theta) = \left\{ \exp\left\{\psi\left(\theta + \frac{t}{n}\right) - \psi(\theta)\right\} \right\}^n,$$

ver (2.2.2).

- La relación de recurrencia de los momentos de \bar{X}_n será

$$m_{j+1}(\bar{X}_n, \theta) = \frac{D_\theta m_j(\bar{X}_n, \theta)}{n} + m_1(\bar{X}_n, \theta) m_j(\bar{X}_n, \theta), \quad (2.5.1)$$

ver (3.4.1).

- La función generatriz de momentos centrales vendrá dada por

$$M_c(\bar{X}_n, t, \theta) = \exp(-\mu t) \exp\left\{n\psi\left(\theta + \frac{t}{n}\right) - n\psi(\theta)\right\}, \quad (2.5.2)$$

ver (2.2.4).

- La relación de recurrencia para los momentos centrales será

$$\mu_{k+1}(\bar{X}_n, \theta) = \mu'_k(\bar{X}_n, \theta) + \mu_1(\bar{X}_n, \theta)\mu_k(\bar{X}_n, \theta),$$

siendo

$$\mu_k(\theta) = E[(\bar{X}_n - \mu)^k] = \Upsilon_k(0, \theta),$$

$$\Upsilon_{k+1}(t, \theta) = D_t \Upsilon_k(t, \theta) + \Upsilon_1(t, \theta)\Upsilon_k(t, \theta),$$

$$\Upsilon_1(t, \theta) = D_t \psi\left(\theta + \frac{t}{n}\right) - \mu(\theta),$$

$$\mu'_k(\theta) = D_t \Upsilon_k(t, \theta)|_{t=0},$$

ver proposición 2.2.2.

- La familia de polinomios asociadas vendrá dada

$$Q_{k,n}(s, \mu) = \frac{V^k(\mu) D_\mu^k f_n(s, \theta)}{n^k f_n(s, \theta)}, \quad (2.5.3)$$

verificando la siguiente relación de recurrencia

$$Q_{k+1,n}(s, \mu) = \left\{ Q_{1,n}(s, \mu) - k \frac{V'(\mu)}{n} \right\} Q_{k,n}(s, \mu) + \frac{V(\mu)}{n} D_\mu Q_{k,n}(s, \mu),$$

siendo $Q_{0,n}(s, \mu) = 1$ y $Q_{1,n}(s, \mu) = \bar{X}_n - \mu$, ver (2.3.2).

La función generatriz de los polinomios $Q_{k,n}^*(s, \mu) = \frac{Q_{k,n}(s, \mu)n^k}{V^k(\mu)}$ tendrá la expresión

$$\Pi_{Q_n^*(z,s,\mu)} = \exp \{ ns [\theta(\mu + z) - \theta(\mu)] - n [\psi(\theta(\mu + z)) - \psi(\theta(\mu))] \}, \quad (2.5.4)$$

ver el teorema 2.3.2.

- La familia de polinomios mónicos ortogonales asociados a $f_n(s, \theta)$ puede ser calculada a través de cualquiera de los métodos mencionados en la sección 2.4. Dicha familia de polinomios será denotada por $\{P_{k,n}(s, \mu)\}_{k \geq 0}$.

Para posteriores desarrollos, es interesante la siguiente propiedad de la norma de los polinomios $P_{k,n}(s, \mu)$.

Lema 2.5.1 *Sea X_1, \dots, X_n una muestra aleatoria procedente de una familia exponencial natural con función de densidad $f(x, \theta)$. Sea $\{P_{k,n}(s, \mu)\}_{k \geq 0}$ el sistema de polinomios ortogonales asociados a la densidad de \bar{X}_n . Se verifica que*

$$\|P_{k,n}\|_{2,\mu}^2 = \frac{V^k(\mu)k!}{n^k} \left[1 + \frac{\gamma k(k-1)}{n} + O(n^{-2}) \right] \quad (2.5.5)$$

siendo $\gamma = \frac{\kappa_2\kappa_4 - \kappa_3^2}{\kappa_2^3}$ con κ_i el cumulante i -ésimo de la distribución $f(x, \theta)$.

Demostración.

Sea X_1, \dots, X_n una muestra aleatoria perteneciente a la NEF, denotamos por $Y_n = \sqrt{n}(\bar{X}_n - \mu)$ y por λ_p a la razón

$$\lambda_p = \frac{\Lambda_p \Lambda_{p-2}}{\Lambda_{p-1}^2},$$

siendo Λ_k el determinante de la matriz de momentos $\mathcal{M}_k(\mu) = (m_{i,j}(\mu))_{i,j=0,\dots,k}$ con $m_{i,j}(\mu) = E_\mu[Y_n^{i+j}]$. Se verifica que,

$$\lambda_p = \frac{\Lambda_p \Lambda_{p-2}}{\Lambda_{p-1}^2} = pV(\mu) \left[1 + (p-1) \frac{\gamma}{n} + O(n^{-2}) \right],$$

ver Lindsay (1989) [29].

Sea $\{B_{k,n}(Y_n, \mu)\}_{k \geq 0}$ el sistema de polinomios ortogonales respecto a la densidad de Y_n . Dichos polinomios podemos obtenerlos, mediante un cambio de variable, a partir de $P_{k,n}(\bar{X}_n, \mu)$

$$B_{k,n}(y, \mu) = P_{k,n} \left(\frac{y}{\sqrt{n}} + \mu \right) = \sum_{j=0}^k c_{kj}(\mu) \left(\frac{y}{\sqrt{n}} \right)^j = n^{-k/2} B_{k,n}^*(y, \mu).$$

Por ser $\{B_{k,n}^*(y, \mu)\}_{k \geq 0}$ un sistema de polinomios ortogonales mónicos, se verifica que

$$\|B_{k,n}^*\|_{2,\mu}^2 = \lambda_1 \cdot \lambda_2 \dots \lambda_k,$$

ver Chihara (1978) [14]. Por tanto

$$\|B_{k,n}\|_{2,\mu}^2 = n^{-k} \|B_{k,n}^*\|_{2,\mu}^2 = n^{-k} \lambda_1 \cdot \lambda_2 \dots \lambda_k.$$

Calculamos

$$\begin{aligned} \lambda_1 \cdot \dots \cdot \lambda_k &= \prod_{j=1}^k jV(\mu) \left\{ 1 + (j-1) \frac{\gamma}{n} + O(n^{-2}) \right\} \\ &= k! V^k(\mu) + \sum_{j_0=1}^k \left(\prod_{j \neq j_0} jV(\mu) \right) \frac{j_0(j_0-1)\gamma V(\mu)}{n} + O(n^{-2}) \\ &= k! V^k(\mu) + k! V^k(\mu) \frac{\gamma}{n} \sum_{j_0=1}^k (j_0-1) + O(n^{-2}) \\ &= V^k(\mu) k! \left\{ 1 + \frac{\gamma k(k-1)}{n \cdot 2} + O(n^{-2}) \right\}, \end{aligned}$$

y de aquí:

$$\|B_{k,n}\|_{2,\mu}^2 = n^{-k} V^k(\mu) k! \left\{ 1 + \frac{\gamma k(k-1)}{n \cdot 2} + O(n^{-2}) \right\}$$

Por otro lado, debido a que

$$\|B_{k,n}\|_{2,\mu}^2 = \|P_{k,n}\|_{2,\mu}^2$$

se tiene el resultado enunciado. □

En la familia de distribuciones NEF, si $h \in \mathcal{U}_n^p(\Omega)$ $p \geq 1$, se verifica que h es analítica en Ω , Abbey y David (1970) [1]. Hay que notar que el hecho de que

$h(\mu)$ sea analítica no garantiza que sea p -estimable, pues podemos encontrar funciones analíticas en las familias consideradas que no son estimables. Por ejemplo, si la muestra procede de una distribución Poisson de parámetro μ , donde $\Omega = (0, +\infty)$, se puede demostrar que la función $h(\mu) = \frac{1}{\mu}$ es analítica en $(0, +\infty)$, y no es estimable en $(0, +\infty)$ ver Lehmann (1983) [28]. En este caso, las únicas funciones paramétricas estimables son aquellas que se pueden desarrollar en serie de potencias no negativas de μ y analíticas en todo $(-\infty, +\infty)$, Glasser (1962) [20].

El siguiente teorema nos proporciona la expresión ortogonal del estimador insesgado de una función paramétrica $h(\mu)$, así como la de su varianza.

Teorema 2.5.1 *Sea $\{P_{k,n}(s, \mu)\}_{k \geq 0}$ el sistema de polinomios ortogonales asociados a \bar{X}_n ,*

$$P_{k,n}(s, \mu) = \sum_{j=0}^k c_{kj}(\mu) s^j. \quad (2.5.6)$$

Entonces, si $h(\mu) \in \mathcal{U}_n^2(\Omega_0)$ y $T_n(\bar{X}_n)$ es su correspondiente estimador insesgado, se verifica:

1. *Los coeficientes del desarrollo de Fourier de $T_n(\bar{X}_n)$ respecto del sistema de polinomios $\{P_{k,n}(s, \mu)\}_{k \geq 0}$ son:*

$$a_{k,n}(\mu) = \frac{1}{\|P_{k,n}\|_{2,\mu}^2} \sum_{j=0}^k c_{kj} \frac{1}{n^j} D_{\theta}^j (h(\mu) \exp \{n\psi(\theta)\}) \exp \{-n\psi(\theta)\}, \quad (2.5.7)$$

para cualquier $\mu \in \Omega_0$.

2. *La varianza del $T_n(\bar{X}_n)$ es*

$$Var_{\mu}(T_n(\bar{X}_n)) = \sum_{k=1}^{\infty} a_{k,n}^2 \|P_{k,n}\|_{2,\mu}^2, \quad \mu \in \Omega_0. \quad (2.5.8)$$

Demostración.

Si $h(\mu) \in \mathcal{U}_n^2(\Omega_0)$ y $T_n(\bar{X}_n)$ es su correspondiente estimador insesgado, este puede expresarse como

$$T_n(s) = \sum_{k=0}^{\infty} a_{k,n}(\mu) P_{k,n}(s, \mu), \quad \forall \mu \in \Omega_0, \quad (2.5.9)$$

donde, según (1.3.4), los coeficientes de Fourier son:

$$a_{k,n}(\mu) = \frac{\langle T_n(s), P_{k,n}(s, \mu) \rangle_{\mu}}{\|P_{k,n}\|_{2,\mu}^2}.$$

Teniendo en cuenta la expresión (2.5.6), tendremos que los coeficientes $a_{k,n}(\mu)$ son

$$a_{k,n}(\mu) = \frac{1}{\|P_{k,n}\|_{2,\mu}^2} \sum_{j=0}^k c_{kj}(\mu) \langle T_n(s), s^j \rangle_{\mu}.$$

Por tanto para determinar la expresión del estimador insesgado de $h(\mu)$ bastará calcular los productos

$$\langle T_n(s), s^j \rangle_{\mu} = \int T_n(s) s^j f_n(s, \theta) d\nu_n(s).$$

Para ello teniendo en cuenta que $h(\mu)$ es insesgado en Ω_0

$$\int T_n(s) \exp\{ns\theta - n\psi(\theta)\} d\nu_n(s) = h(\mu),$$

igualdad que podemos expresar como

$$\int T_n(s) \exp(ns\theta) d\nu_n(s) = h(\mu) \exp\{n\psi(\theta)\}.$$

Al ser $h(\mu)$ una función analítica en Ω_0 , derivando ambos miembros de la igualdad anterior respecto de θ , k veces, se tiene

$$\int n^k s^k T_n(s) \exp(ns\theta) d\nu_n(s) = D_{\theta}^k (h(\mu) \exp\{n\psi(\theta)\}),$$

y, multiplicando ambos miembros por $\exp\{-n\psi(\theta)\}$

$$\begin{aligned}\langle T_n(s), s^k \rangle_\mu &= \int n^k s^k T_n(s) \exp\{ns\theta - n\psi(\theta)\} d\nu_n(s) \\ &= \frac{1}{n^k} D_\theta^k (h(\mu) \exp\{n\psi(\theta)\}) \exp\{-n\psi(\theta)\}.\end{aligned}\quad (2.5.10)$$

A partir de aquí, dividiendo por la norma $\|P_{k,n}\|_{2,\mu}^2$ se obtiene la expresión (2.5.7).

Finalmente, la expresión de la varianza de $T_n(\bar{X}_n)$ se obtiene directamente de la aplicación del teorema 1.3.1.

□

Nótese que los $a_{k,n}(\mu)$ son combinaciones lineales de las sucesivas derivadas $D_\theta^j h(\mu)$, $j = 0, \dots, k$. En el caso particular de la familia NEF-QVF, los coeficientes $a_{k,n}$ coinciden con $\frac{D_\theta^k h(\mu)}{k!}$ ver Morris (1983) [36]. La importancia del teorema 2.5.1 es que podemos obtener el desarrollo ortogonal para el estimador insesgado, $T_n(\bar{X}_n)$, sin un conocimiento explícito del mismo.

Observación 2.5.1 *En apariencia el segundo término de la expresión (2.5.9) depende de μ , en realidad tal dependencia no existe puesto que, para cualquier valor de $\mu \in \Omega_0$, la serie de Fourier converge (por lo menos en media cuadrática) hacia la misma función y esta función es independiente de μ . De todas formas, el valor de μ sí que es importante en las aplicaciones prácticas pues determina la velocidad de convergencia de la serie de Fourier. Así, podemos obtener, en general, expresiones numéricamente eficientes con una elección adecuada de $\mu \in \Omega_0$. En particular, dada una realización mues-*

tral x_1, x_2, \dots, x_n procedente de la NEF, sugerimos tomar $\mu = \bar{x}_n$, obteniendo

$$T_n(\bar{x}_n) = \sum_{k=0}^{\infty} a_{k,n}(\bar{x}_n) \varphi_{k,n}(\bar{x}_n) \quad (2.5.11)$$

con $\varphi_{k,n}(\bar{x}_n) = P_{k,n}(\bar{x}_n, \bar{x}_n)$.

Ejemplo 2.5.1 Familia chi-cuadrada transformada.

Según Rahman y Gupta (1993) [39], una familia uniparamétrica exponencial con función de densidad

$$f(x, \theta) = \exp\{a(x)b(\theta) + c(\theta) + q(x)\} I_D(x) \quad (2.5.12)$$

pertenece a la familia chi-cuadrado transformada si

(i) θ es un parámetro desconocido, $\theta \in \Theta$, siendo Θ un abierto de \mathbb{R} .

(ii) D es un conjunto de Borel y no depende de θ .

(iii) $b(\theta)$ y $c(\theta)$ son funciones reales, derivables, definidas en Θ .

(iv) $b(\theta) < 0$ y $b'(\theta) \neq 0$, para cualquier $\theta \in \Theta$.

(v) Para cualquier $\theta \in \Theta$, $\frac{2c'(\theta)b(\theta)}{b'(\theta)} = v$, siendo $v > 0$ independiente de θ .

A continuación se muestra una tabla con algunos miembros de esta familia:

Función de densidad	a(x)	b(θ)	c(θ)	v
<p><i>Gamma</i> $x > 0, \theta > 0,$ $p > 0$ (conocido) $f(x, \theta) = \frac{1}{\Gamma(p)\theta^p} x^{p-1} \exp\{-\frac{x}{\theta}\}$</p>	x	$-\frac{1}{\theta}$	$-p \log \theta$	$2p$
<p><i>Normal</i> $x \in \mathbb{R}, \theta > 0$ $f(x, \theta) = \frac{1}{\sqrt{2\pi\theta}} \exp(-\frac{x^2}{2\theta^2})$</p>	x^2	$-\frac{1}{2\theta^2}$	$-\log \theta$	1
<p><i>Lognormal</i> $x > 0, \theta > 0$ $f(x, \theta) = \frac{1}{\sqrt{2\pi\theta x}} \exp(-\frac{\log^2 x}{2\theta^2})$</p>	$\log^2 x$	$-\frac{1}{2\theta^2}$	$-\log \theta$	1
<p><i>Rayleigh</i> $x > 0, \theta > 0$ $f(x, \theta) = \frac{x}{\theta^2} \exp(-\frac{x^2}{2\theta^2})$</p>	x^2	$-\frac{1}{2\theta^2}$	$-2 \log \theta$	2
<p><i>Pareto</i> $x > 1, \theta > 0$ $f(x, \theta) = \theta x^{-(\theta+1)}$</p>	$\log x$	$-\theta$	$\log \theta$	2
<p><i>Weibull</i> $x > 0, \theta > 0,$ $p > 0$ (conocido) $f(x, \theta) = \theta p x^{p-1} \exp(-\theta x^p)$</p>	x^p	$-\theta$	$\log \theta$	2
<p><i>Maxwell</i> $x \in \mathbb{R}, \theta > 0$ $f(x, \theta) = \sqrt{2/\pi} \theta^{3/2} x^2 \exp(-\frac{\theta x^2}{2})$</p>	x^2	$-\frac{\theta}{2}$	$\frac{3}{2} \log \theta$	3
<p><i>Inversa Gaussiana</i> $x > 0, \theta > 0,$ $\mu > 0$ (conocido) $f(x, \theta) = (\frac{\theta}{2\pi x^3})^{1/2} \exp(-\frac{\theta(x-\mu)^2}{2\mu^2 x})$</p>	$\frac{(x-\mu)^2}{2\mu^2 x}$	$-\theta$	$\frac{\log \theta}{2}$	1

Si X_1, \dots, X_n es una muestra aleatoria de una distribución chi-cuadrado transformada tal que $E[|a(X)|^r] < \infty$, para cualquier $r > 0$, entonces el estadístico $S_n = \sum_{i=1}^n \frac{a(X_i)}{n}$ es suficiente y completo y sigue una distribución Gamma con parámetros $\frac{nv}{2}$ y $\frac{nv}{2\lambda} = -nb(\theta)$, ver López-Blázquez (1997) [30].

Como $b(\theta)$ es una función uno a uno, es conveniente considerar λ , la media de S_n , como el parámetro desconocido de la familia. La densidad de S_n la podemos escribir:

$$f_n(s, \lambda) = \frac{1}{\Gamma\left(\frac{nv}{2}\right)} \left(\frac{nv}{2\lambda}\right)^{\frac{nv}{2}} \exp\left\{-\frac{nv s}{2\lambda}\right\} s^{\frac{nv}{2}-1},$$

para $s > 0$ y $\lambda > 0$.

Asociados a esta densidad existe un sistema de polinomios mónicos ortogonales que vienen dados por

$$P_{j,n}^{(\frac{nv}{2}-1)}(s, \lambda) = (-1)^j j! \left(\frac{2\lambda}{nv}\right)^j L_j^{(\frac{nv}{2}-1)}\left(\frac{nv s}{2\lambda}\right), \quad j = 0, 1, \dots$$

donde $L_j^{(\alpha)}$, $j \geq 0$, son los polinomios de Laguerre generalizados, ortogonales respecto de la función peso $w(t) = e^{-t} t^\alpha$, $t > 0$, cuya expresión explícita viene dada por

$$L_j^{(\alpha)}(x) = \sum_{m=0}^j \binom{j+\alpha}{j-m} \frac{(-x)^m}{m!}, \quad \alpha > 0. \quad (2.5.13)$$

Dichos polinomios verifican la relación de recurrencia

$$P_{j,n}^{(\frac{nv}{2}-1)} = \left(s - \lambda - \frac{4\lambda(j-1)}{nk}\right) P_{j-1,n}^{(\frac{nv}{2}-1)} - 4\lambda^2 \left(j - 2 + \frac{nk}{2}\right) \frac{(j-1)}{(nk)^2} P_{j-2,n}^{(\frac{nv}{2}-1)},$$

para $j \geq 1$, con $P_{-1,n}^{(\frac{nv}{2}-1)}(s, \lambda) = 0$ y $P_{0,n}^{(\frac{nv}{2}-1)}(s, \lambda) = 1$.

Para cualquier $\lambda > 0$, los coeficientes, $a_{j,n}(\lambda)$, del desarrollo en serie del estimador insesgado de $h(\lambda)$, $T_n(S_n)$, pueden ser calculados en términos de las

derivadas de h , siendo su expresión:

$$a_{j,n}(\lambda) = \frac{D_\lambda^j h(\lambda) \left(\frac{nv}{2}\right)^j \Gamma\left(\frac{nv}{2}\right)}{j! \Gamma\left(\frac{nv}{2} + j\right)},$$

por tanto, el UMVUE de $h(\lambda)$ es:

$$T_n(S_n) = \sum_{j=0}^{\infty} \frac{D_\lambda^j h(\lambda) \Gamma\left(\frac{nv}{2}\right)}{j! \Gamma\left(\frac{nv}{2} + j\right)} \left(\frac{nv}{2}\right)^j P_{j,n}^{(\frac{nv}{2}-1)}(S_n, \lambda), \quad \lambda > 0. \quad (2.5.14)$$

siendo su varianza

$$Var_\lambda(T_n(S_n)) = \sum_{j=1}^{\infty} \frac{\lambda^{2j} (D_\lambda^j h(\lambda))^2 \Gamma\left(\frac{nv}{2}\right)}{j! \Gamma\left(\frac{nv}{2} + j\right)} \quad (2.5.15)$$

Debido a la unicidad del UMVUE, la elección de λ en el término de la derecha de la expresión (2.5.14) es irrelevante. En particular, si tomamos $\lambda = S_n$ obtenemos:

$$T_n(S_n) = \sum_{j=0}^{+\infty} \frac{(-1)^j D_\lambda^j h(S_n) S_n^j \Gamma\left(\frac{nv}{2}\right)}{\Gamma\left(\frac{nv}{2} + j\right)} L_j^{(\frac{nv}{2}-1)}\left(\frac{nv}{2}\right) \quad (2.5.16)$$

En el caso de la distribución Gamma, Woodward y Gray (1975) [53] derivan una expresión que coincide con la expresión (2.5.16) aplicando un método que consiste en aplicar jackknife generalizado para eliminar el sesgo del estimador de máxima verosimilitud y posteriormente aplicar el teorema de Blackwell-Rao al resultado.

El siguiente corolario nos da una caracterización de las funciones paramétricas UMVU-estimables en Ω_0 .

Corolario 2.5.1 $h \in \mathcal{U}_n^2(\Omega_0)$ sii h es analítica en Ω_0 y

$$\sum_{k=1}^{\infty} a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2 < \infty, \quad \forall \mu \in \Omega_0, \quad (2.5.17)$$

con $a_{k,n}(\mu)$ dados en el teorema 2.5.1.

Demostración.

Si $h \in \mathcal{U}_n^2(\Omega_0)$ es analítica y por el Teorema 2.5.1, se tiene (2.5.17).

Por otra parte, dada una función paramétrica h , analítica en Ω_0 tal que se verifica (2.5.17), por el Teorema de Riesz-Fisher y la completitud del estadístico suficiente se tiene que $T_n(\bar{X}_n)$ definido como en (2.5.9) pertenece al espacio $\mathcal{L}_n^2(\Omega_0)$, verificando

$$E_\mu[T_n(\bar{X}_n)] = h(\mu), \quad \forall \mu \in \Omega_0,$$

lo que demuestra el resultado. □

Según el teorema anterior, el conjunto de funciones UMVU-estimables es un subconjunto del conjunto de funciones analíticas que verifican la restricción (2.5.17), o de forma equivalente

$$\mathcal{U}_n^2(\Omega_0) = \left\{ h \in \mathcal{A}(\Omega_0) : \sum_{k=1}^{\infty} a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2 < \infty \right\}, \quad (2.5.18)$$

donde por $\mathcal{A}(\Omega_0)$ denotamos al conjunto de funciones analíticas en Ω_0 .

Corolario 2.5.2 *La sucesión $\{\mathcal{U}_n^2(\Omega_0)\}_{n \geq 1}$, es creciente en el sentido de la inclusión.*

Demostración.

Sea $h \in \mathcal{U}_{n_0}^2(\Omega_0)$ y $T_{n_0}(\bar{X}_{n_0})$ su correspondiente UMVUE.

Para muestras de tamaño $n \geq n_0$, el estadístico \bar{X}_n es suficiente y completo.

Si consideramos

$$T_n(\bar{X}_n) = E_\mu [T_{n_0}(\bar{X}_{n_0}) | \bar{X}_n],$$

se verifica que es insesgado y su varianza

$$\text{Var}_\mu (T_n(\bar{X}_n)) = \text{Var}_\mu (E_\mu [T_{n_0}(\bar{X}_{n_0}) | \bar{X}_n]) \leq \text{Var}_\mu (T_{n_0}(\bar{X}_{n_0})) < \infty$$

por tanto, $h \in \mathcal{U}_n^2(\Omega)$.

□

Como consecuencia del Corolario anterior, si una función paramétrica h , admite un estimador UMVUE en Ω_0 para un tamaño muestral n_0 , también lo admite para cualquier tamaño de muestra mayor que n_0 . Esta propiedad posibilita el estudio de las propiedades asintóticas del estimador UMVUE en Ω_0 .

2.5.1 Estimación de la varianza del UMVUE

Dada $h \in \mathcal{U}_n^2(\Omega_0)$, sea $T_n(\bar{X}_n)$ su estimador UMVUE en Ω_0 , por el Teorema 2.5.1 una expresión para la varianza viene dada por:

$$g(\mu) = \text{Var}_\mu (T_n(\bar{X}_n)) = \sum_{j=1}^{\infty} a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2, \quad \mu \in \Omega_0. \quad (2.5.19)$$

donde los $a_{k,n}(\mu)$ vienen dados por la expresión (2.5.7).

Si suponemos que $h^2 \in \mathcal{U}_n^2(\Omega_0)$ y $M_n(\bar{X}_n)$ es su estimador insesgado, entonces el estimador insesgado en Ω_0 de (2.5.19) basado en el suficiente y completo viene dado por:

$$V_n(\bar{X}_n) = T_n^2(\bar{X}_n) - M_n(\bar{X}_n).$$

La expresión de $M_n(\bar{X}_n)$ puede ser calculada en términos de los polinomios ortogonales y junto con la expresión de $T_n^2(\bar{X}_n)$ podremos obtener la de $V_n(\bar{X}_n)$.

El estimador $V_n(\bar{X}_n)$ es insesgado.

Si se desea una estimación de máxima verosimilitud de (2.5.19) basta aplicar la propiedad de invarianza obteniéndose

$$g(\bar{X}_n) = \sum_{j=1}^{\infty} a_{k,n}^2(\bar{X}_n) \|P_{k,n}\|_{2,\bar{X}_n}^2.$$

2.6 Implementación Numérica

Según vimos en la observación 2.5.1 dada una realización x_1, \dots, x_n el estimador insesgado puede calcularse como:

$$T_n(\bar{x}_n) = \sum_{j=0}^{\infty} a_{j,n}(\bar{x}_n) \varphi_{j,n}(\bar{x}_n) \quad (2.6.1)$$

con $\varphi_{j,n}(\bar{x}_n) = P_{j,n}(\bar{x}_n, \bar{x}_n)$.

Dada la imposibilidad, desde el punto de vista práctico, de manipular series infinitas, con objeto de implementar en un ordenador el procedimiento de estimación visto en las secciones anteriores, los desarrollos (2.6.1) deben ser truncados. Así pues, obtendremos aproximaciones al verdadero valor del estimador insesgado que vienen dadas por

$$T_{n,k}(\bar{x}_n) = \sum_{j=0}^k a_{j,n}(\bar{x}_n) P_{j,n}(\bar{x}_n, \bar{x}_n). \quad (2.6.2)$$

Esta aproximación al estimador insesgado ha sido implementada mediante un programa escrito en MAPLE V R.4.

Las variables del programa se describen a continuación:

- k: Número de polinomios ortonormales.
- C: Función cumulante.

- m: Momentos no centrados.
- M: Matriz de los coeficientes de los polinomios
- nor: Norma de los polinomios ortogonales.
- pp: Evaluación de los polinomios ortogonales.
- s: valor del estadístico suficiente y completo.
- h: función a estimar.
- pe: producto escalar $\langle T_n(s), s^j \rangle$.
- a: coeficientes de Fourier del desarrollo de $T_n(\bar{X}_n)$.
- T: Aproximación dada en (2.6.2).

- Inicialización.

```

> theta:='theta'; k:='k'; s:='s'; C:=C(theta); mu:='mu'; n:='n';
> with(linalg);
> m:=array(0..2*k); M:=matrix(k,2*k);
> nor:=array(0..k-1);
> pp:=array(0..k-1); pp[0]:=1;
> pe:=array(0..k-1); a:=array(0..k-1);

```

- Cálculo de los momentos mediante la fórmula de recurrencia (2.5.1).

```

> m[0]:=1; m[1]:=diff(C,theta);
> for j from 1 to 2*k-3 do
> m[j+1]:=expand(diff(m[j],theta)/n + m[1]*m[j]); od;

```

- Cálculo de la matriz de coeficientes de los polinomios, ver (1.5.3).

Para ello:

1.- Determinamos la matriz de momentos.

```
> for i from 1 to k do
> for j from i to k do
> M[i,j]:= m[i+j-2]; od;od;
> for i from 2 to k do
> for j from 1 to i-1 do
> M[i,j]:= M[j,i]; od; od;
```

2.- Creamos la matriz auxiliar $M = (M|I)$.

```
> for j from 1 to k do for d from k+1 to 2*k do
> if d-j=k then M[j,d]:=1 else M[j,d]:=0 fi; od; od;
```

3.- Aplicamos el método de eliminación de Gauss.

Así obtenemos una nueva matriz $M=(\text{normas}|\text{coeficientes})$ que nos calcula las normas de los polinomios y los coeficientes de los mismos.

```
> for j from 1 to k-1 do
> M:=pivot(M,j,j,j+1..k); od;
```

4.- Creamos el vector norma de los polinomios y la matriz de coeficientes.

```
> for j from 1 to k do nor[j-1]:=M[j,j]; od;
> M:=delcols(M,1..k);
```

- Evaluación de los polinomios ortogonales.

```
> r:='r'; for j from 1 to k-1 do
> pp[j]:=sum(M[j+1,r+1]*s^r, r=0..j); od;
```

- Cálculo de la aproximación (2.6.2).

Para ello:

1.- Cálculo del producto escalar $\langle T_n(s), s^j \rangle$.

```
> mu:=m[1]; pe[0]:=h;
> for r from 1 to k-1 do
> pe[r]:= simplify(diff((1/n^r) * h*exp(n*C),theta$r)*exp(-1*n*C)); od;
```

2.- Cálculo de los coeficientes de Fourier.

```
> for m from 0 to k-1 do a[m]:=0;
> for j from 0 to m do a[m]:=a[m]+M[m+1,j+1]*(pe[j]/nor[j]); od; od;
```

3.- Cálculo de la aproximación (2.6.2).

```
> T:=0 ;for j from 0 to k-1 do T:=expand(T+a[j]*pp[j]); od;
```

2.7 Propiedades asintóticas de los estimadores insesgados

En esta sección estableceremos propiedades asintóticas de los estimadores obtenidos mediante desarrollos ortogonales. Para ello, será necesario estudiar, en primer lugar, algunos resultados sobre el comportamiento asintótico de los

polinomios y de los coeficientes de Fourier del desarrollo ortogonal de los estimadores insesgados.

Para posteriores desarrollos necesitaremos utilizar los polinomios de Hermite, los cuales serán denotados por $H_{e_k}(z)$, $k \geq 0$ y vienen definidos de la forma

$$H_{e_k}(z) = k! \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} \frac{(-1)^j}{2^j j! (k-2j)!} z^{k-2j}. \quad (2.7.1)$$

Dichos polinomios verifican la siguiente relación de recurrencia

$$H_{e_k}(z) = zH_{e_{k-1}}(z) - (k-1)H_{e_{k-2}}(z), \quad k \geq 1, \quad (2.7.2)$$

siendo $H_{e_{-1}}(z) = 0$, $H_{e_0}(z) = 1$,

por tanto, la expresión explícita de los primeros polinomios será

$$H_{e_1}(z) = z, \quad H_{e_2}(z) = z^2 - 1, \quad (2.7.3)$$

Los polinomios de Hermite también pueden ser obtenidos a partir de su función generatriz, cuya expresión es

$$\Pi_H(w, z) = \sum_{k \geq 0} \frac{H_{e_k}(z)}{k!} w^k = \exp \left\{ -\frac{w^2}{2} + zw \right\}.$$

Lema 2.7.1 *Consideremos la familia de polinomios $\{Q_{k,n}(s, \mu)\}_{k \geq 0}$ definidos en (2.5.3), asociados a la densidad $f_n(s, \theta)$. Se verifica que:*

1. *La función generatriz de los polinomios $Q_{k,n}^*(s, \mu) = \frac{Q_{k,n}(s, \mu)n^k}{V(\mu)^k}$ puede ser expresada como*

$$\Pi_{Q_n^*}(w, s, \mu) = \frac{f_n(s, w + \mu)}{f_n(s, \mu)}. \quad (2.7.4)$$

2.

$$\lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{k}{2}} Q_{k,n}(\mu + zn^{-1/2}V(\mu)^{1/2}, \mu) = H_{e_k}(z) \quad k \geq 0. \quad (2.7.5)$$

Demostración.

1. Teniendo en cuenta la expresión (2.5.4) de la función generatriz de los polinomios $\{Q_{k,n}^*(s, \mu)\}_{k \geq 0}$ y la densidad de la NEF, $f_n(s, \mu)$, se obtiene la expresión (2.7.4).

2. Denotamos por

$$G_{k,n}(z, \mu) = \left(\frac{n}{V(\mu)} \right)^{\frac{k}{2}} Q_{k,n}(\mu + zn^{-1/2}V(\mu)^{1/2}, \mu), \quad k \geq 0, \quad (2.7.6)$$

y sea

$$\Pi_{G_n}(w, z, \mu) = \sum_{k \geq 0} G_{k,n}(z, \mu) \frac{w^k}{k!},$$

su función generatriz. Bastará demostrar que:

$$\lim_n \Pi_{G_n}(w, z, \mu) = \exp \left\{ \frac{-w^2}{2} + zw \right\}, \quad n \rightarrow \infty. \quad (2.7.7)$$

Teniendo en cuenta que

$$\Pi_{G_n}(w, z, \mu) = \Pi_{Q_n^*}(wn^{-1/2}V^{1/2}(\mu), \mu + zn^{-1/2}V^{1/2}(\mu), \mu),$$

y aplicando el apartado anterior

$$\Pi_{G_n}(w, z, \mu) = \frac{f_n(\mu + zn^{-1/2}V^{1/2}(\mu), \mu + wn^{-1/2}V^{1/2}(\mu))}{f_n(\mu + zn^{-1/2}V^{1/2}(\mu), \mu)},$$

de donde

$$\begin{aligned} \log \Pi_{G_n}(w, z, \mu) &= n \left\{ (\mu + zn^{-1/2}V^{1/2}(\mu)) \left[\theta(\mu + wn^{-1/2}V^{1/2}(\mu)) - \theta(\mu) \right] \right\} \\ &\quad - n \left\{ \psi \left(\theta(\mu + wn^{-1/2}V^{1/2}(\mu)) \right) - \psi \left(\theta(\mu) \right) \right\} \\ &= n \left(\mu + zn^{-1/2}V^{1/2}(\mu) \right) \Pi_1 - n \Pi_2. \end{aligned} \quad (2.7.8)$$

Desarrollando, por una parte

$$\begin{aligned}\Pi_1 &= \theta(\mu + wn^{-1/2}V^{1/2}(\mu)) - \theta(\mu) \\ &= \theta'(\mu)wn^{-1/2}V^{1/2}(\mu) + \frac{1}{2}\theta''(\mu)w^2n^{-1}V(\mu) + o(n^{-1}),\end{aligned}\quad (2.7.9)$$

y por otra

$$\begin{aligned}\Pi_2 &= \psi(\theta(\mu + wn^{-1/2}V^{1/2}(\mu))) - \psi(\theta(\mu)) \\ &= \mu\theta'(\mu)wn^{-1/2}V^{1/2}(\mu) + \frac{1}{2}V(\mu)(\theta'(\mu))^2w^2n^{-1}V(\mu) \\ &\quad + \frac{1}{2}\mu\theta''(\mu)w^2n^{-1}V(\mu) + o(n^{-1}),\end{aligned}\quad (2.7.10)$$

y sustituyendo los resultados (2.7.9) y (2.7.10) en (2.7.8), obtenemos:

$$\log \Pi_{G_n}(w, z, \mu) = V(\mu)\theta'(\mu)zw - V(\mu)(\theta'(\mu))^2\frac{w^2}{2} + O(n^{-1/2})$$

teniendo en cuenta que $V^{-1}(\mu) = \theta'(\mu)$ se obtiene, fácilmente, (2.7.7) y de aquí el resultado (2.7.5).

□

Lema 2.7.2 *Se verifica que*

$$\lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{k}{2}} \left\{ P_{k,n}(\mu + zn^{-1/2}V^{1/2}(\mu), \mu) - Q_{k,n}(\mu + zn^{-1/2}V^{1/2}(\mu), \mu) \right\} = 0. \quad (2.7.11)$$

Demostración.

Demostraremos el resultado enunciado mediante inducción.

Debido a que $P_{1,n}(s, \mu)$ es igual a $Q_{1,n}(s, \mu)$ es evidente que

$$\lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{1}{2}} \left\{ P_{1,n}(\mu + zn^{-1/2}V^{1/2}(\mu), \mu) - Q_{1,n}(\mu + zn^{-1/2}V^{1/2}(\mu), \mu) \right\} = 0.$$

Suponemos que se verifica (2.7.11) para cualquier $j < k$.

Teniendo en cuenta que los polinomios $\{P_{k,n}(s, \mu)\}_{k \geq 0}$ pueden ser obtenidos, según se vió en la sección 1.5, ortogonalizando por el método de Gram-Schmidt la familia $\{Q_{k,n}(s, \mu)\}_{k \geq 0}$, mediante la expresión

$$P_{k,n}(s, \mu) = Q_{k,n}(s, \mu) - \sum_{j=0}^{k-1} q_{k,j} P_{j,n}(s, \mu),$$

siendo $q_{k,j} = \frac{\langle Q_{k,n}, P_{j,n} \rangle_{\mu}}{\|P_{j,n}\|_{2,\mu}^2}$, para probar (2.7.11), bastará con demostrar que

$$\lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{k}{2}} \sum_{j=0}^{k-1} q_{k,j} P_{j,n}(s_n, \mu) = 0, \quad (2.7.12)$$

denotando por $s_n = \mu + zn^{-1/2}V^{1/2}(\mu)$.

Para ello, podemos expresar

$$\begin{aligned} \left(\frac{n}{V(\mu)} \right)^{\frac{k}{2}} \sum_{j=0}^{k-1} q_{k,j} P_{j,n}(s_n, \mu) &= \sum_{j=0}^{k-1} \left(\frac{n}{V(\mu)} \right)^{\frac{k-j}{2}} q_{k,j} \left(\frac{n}{V(\mu)} \right)^{\frac{j}{2}} P_{j,n}(s, \mu) \\ &= \sum_{j=0}^{k-1} A_{k,j,n} B_{k,j,n}, \end{aligned}$$

siendo

$$A_{k,j,n} = \left(\frac{n}{V(\mu)} \right)^{\frac{k-j}{2}} q_{k,j} \quad \text{y} \quad B_{k,j,n} = \left(\frac{n}{V(\mu)} \right)^{\frac{j}{2}} P_{j,n}(s, \mu),$$

y estudiar el comportamiento límite de $A_{k,j,n}$ y de $B_{k,j,n}$.

Debido a la hipótesis de inducción, $\forall j < k$ tenemos

$$\begin{aligned} \lim_{n \rightarrow \infty} B_{k,j,n} &= \lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{j}{2}} P_{j,n}(s, \mu) \\ &= \lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{j}{2}} Q_{j,n}(s, \mu) = H_{e_j}(z), \end{aligned} \quad (2.7.13)$$

y además,

$$\begin{aligned}
 \lim_{n \rightarrow \infty} A_{k,j,n} &= \lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{k-j}{2}} \frac{\langle Q_{k,n}, P_{j,n} \rangle_{\mu}}{\|P_{j,n}\|_{2,\mu}^2} = \\
 &= \lim_{n \rightarrow \infty} \frac{\langle \left(\frac{n}{V(\mu)} \right)^{\frac{k}{2}} Q_{k,n}, \left(\frac{n}{V(\mu)} \right)^{\frac{j}{2}} P_{j,n} \rangle_{\mu}}{\left(\frac{n}{V(\mu)} \right)^j \|P_{j,n}\|_{2,\mu}^2} = \\
 &= \frac{\langle H_{e_k}(z), H_{e_j}(z) \rangle}{j!} = 0,
 \end{aligned} \tag{2.7.14}$$

debido a la ortogonalidad de los polinomios de Hermite.

A partir de (2.7.13) y (2.7.14), se obtiene el resultado (2.7.12). \square

El polinomio de grado k , ortogonal respecto a la densidad de \bar{X}_n , puede ser expresado de la forma

$$P_{k,n}(s, \mu) = \sum_{j=0}^k \pi_{k,j,n}(\mu) (s - \mu)^j. \tag{2.7.15}$$

Como consecuencia inmediata del lema anterior se obtiene el siguiente corolario:

Corolario 2.7.1 *Si denotamos por $\text{coef}(j, H_{e_k})$ al coeficiente j del k -ésimo polinomio de Hermite, $\forall \mu \in \Omega$, se verifica que*

$$\lim_{n \rightarrow \infty} \left(\frac{n}{V(\mu)} \right)^{\frac{k-j}{2}} \pi_{k,j,n}(\mu) = \text{coef}(j, H_{e_k}), \quad 0 \leq j \leq k.$$

Estudiaremos, a continuación el comportamiento asintótico de los coeficientes de Fourier de la expresión ortogonal de los estimadores insesgados.

Teorema 2.7.1 *Sea $h \in \mathcal{U}_{n_0}^2(\Omega_0)$, y $T_n(\bar{X}_n) = \sum_{k=0}^{\infty} a_{k,n}(\mu) P_{k,n}(\bar{X}_n, \mu)$, el estimador insesgado de $h(\mu)$ en Ω_0 para cada $n \geq n_0$. Se verifica*

$$\lim_n a_{k,n}(\mu) = \frac{D_{\mu}^k h(\mu)}{k!} \quad \forall k \geq 0, \mu \in \Omega_0. \tag{2.7.16}$$

Demostración.

Los coeficientes de Fourier del desarrollo de $T_n(\bar{X}_n)$ vienen dados por

$$a_{k,n}(\mu) = \frac{\langle T_n(\bar{X}_n), P_{k,n}(\bar{X}_n, \mu) \rangle_\mu}{\|P_{k,n}\|_{2,\mu}^2},$$

ver (1.3.4). Teniendo en cuenta (2.7.11) y tomando límite en la expresión anterior obtenemos

$$\begin{aligned} \lim_{n \rightarrow \infty} a_{k,n}(\mu) &= \lim_{n \rightarrow \infty} \left(\frac{V(\mu)}{n} \right)^{\frac{k}{2}} \frac{\langle T_n(\bar{X}_n), n^{k/2} V(\mu)^{-k/2} P_{k,n}(\bar{X}_n, \mu) \rangle_\mu}{\|P_{k,n}\|_{2,\mu}^2} \\ &= \lim_{n \rightarrow \infty} \left(\frac{V(\mu)}{n} \right)^{\frac{k}{2}} \frac{\langle T_n(\bar{X}_n), n^{k/2} V(\mu)^{-k/2} Q_{k,n}(\bar{X}_n, \mu) \rangle_\mu}{\|P_{k,n}\|_{2,\mu}^2}. \end{aligned} \quad (2.7.17)$$

Por otra parte, si derivamos k veces respecto de μ la igualdad

$$E_\mu[T_n(\bar{X}_n)] = \int T_n(s) f_n(s, \theta) ds = h(\mu),$$

debido a la definición de los polinomios $Q_{k,n}(s, \mu)$, ver (2.5.3), obtenemos

$$V^k(\mu) \frac{D_\mu^k h(\mu)}{n^k} = \int T_n(s) Q_{k,n}(s, \theta) f_n(s, \theta) ds = \langle T_n(\bar{X}_n), Q_{k,n}(\bar{X}_n, \mu) \rangle_\mu.$$

Sustituyendo este resultado en (2.7.17), y teniendo en cuenta la expresión (2.5.5) de la norma de los polinomios, se obtiene de forma inmediata que,

$$\lim_n a_{k,n}(\mu) = \frac{D_\mu^k h(\mu)}{k!}, \quad \mu \in \Omega_0 \quad \forall k \geq 0.$$

□

Sea $h \in \mathcal{U}_n^2(\Omega_0)$, $n \geq n_0$ y $k_0 \geq 1$. Denotaremos por $R_{k_0}(\bar{X}_n, \mu)$ al resto de

la serie de Fourier

$$R_{k_0}(\bar{X}_n, \mu) = \sum_{k=k_0}^{\infty} a_{k,n}(\mu) P_{k,n}(\bar{X}_n, \mu).$$

El siguiente lema muestra el comportamiento asintótico de dicho resto.

Lema 2.7.3 Consideremos $h \in \mathcal{U}_n^2(\Omega_0)$ $n \geq n_0$ y sea $k_0 \geq 1$. Se verifica que $\forall \mu \in \Omega_0, \forall \alpha < k_0/2$

$$n^\alpha R_{k_0}(\bar{X}_n, \mu) \xrightarrow{P_\mu} 0, \quad n \rightarrow \infty.$$

Demostración.

Aplicando la desigualdad de Tchebychev se tiene que

$$P_\mu[n^\alpha |R_{k_0}(\bar{X}_n, \mu)| > \epsilon] \leq \frac{n^{2\alpha}}{\epsilon^2} \text{Var}_\mu(R_{k_0}(\bar{X}_n, \mu)),$$

por lo que bastará demostrar que

$$\lim_n n^{2\alpha} \text{Var}_\mu(R_{k_0}(\bar{X}_n, \mu)) = 0, \quad \forall \alpha < k_0/2 \quad \forall \mu \in \Omega.$$

Para ello,

$$\begin{aligned} \lim_n n^{2\alpha} \text{Var}_\mu(R_{k_0}(\bar{X}_n, \mu)) &= \lim_n n^{2\alpha} \sum_{k=k_0}^{\infty} a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2 = \\ &= \lim_n n^{2\alpha} \sum_{k=k_0}^{\infty} \frac{a_k^2(\mu) V^k(\mu) k!}{n^k} \leq \lim_n n^{2\alpha - k_0} n_0^{k_0} \sum_{k=k_0}^{\infty} \frac{a_k^2(\mu) V^k(\mu) k!}{n_0^k} = 0, \end{aligned}$$

donde $a_k(\mu) = \lim_n a_{k,n}(\mu) = \frac{D_\mu^k h(\mu)}{k!}$.

□

Lema 2.7.4 *Se verifican*

(i)

$$\left(\frac{n}{V(\mu)}\right)^{\frac{k}{2}} P_{k,n}(\bar{X}_n, \mu) \xrightarrow{d} H_{e_k}(Z),$$

con $Z \sim N(0, 1)$.

(ii)

$$\overline{\lim}_n \left(\frac{n}{2V(\mu) \log \log n}\right)^{\frac{k}{2}} P_{k,n}(\bar{X}_n, \mu) = 1, \quad (\text{c.s.}) \quad (2.7.18)$$

Demostración.

(i) Definiendo $Z_n = \sqrt{n} \frac{\bar{X}_n - \mu}{V(\mu)^{\frac{1}{2}}}$ y debido al lema 2.7.2 podemos escribir

$$\begin{aligned} p \lim_n \frac{n^{\frac{k}{2}}}{V(\mu)^{\frac{k}{2}}} P_{k,n}(\bar{X}_n, \mu) &= p \lim_n \frac{n^{\frac{k}{2}}}{V(\mu)^{\frac{k}{2}}} P_{k,n}\left(\mu + \frac{Z_n}{\sqrt{n}} V(\mu)^{\frac{1}{2}}, \mu\right) = \\ &= p \lim_n \frac{n^{\frac{k}{2}}}{V(\mu)^{\frac{k}{2}}} Q_{k,n}\left(\mu + \frac{Z_n}{\sqrt{n}} V(\mu)^{\frac{1}{2}}, \mu\right), \end{aligned}$$

a partir de lo cual, teniendo en cuenta el lema 2.7.1 y el teorema central del límite, el resultado es inmediato.

(ii) Teniendo en cuenta la expresión (2.7.15) de $P_{k,n}(\bar{X}_n, \mu)$, podemos expresar

$$\begin{aligned} \overline{\lim}_n \left(\frac{n}{2V(\mu) \log \log n}\right)^{\frac{k}{2}} P_{k,n}(\bar{X}_n, \mu) &= \\ &= \overline{\lim}_n \sum_{j=1}^k \left(\frac{n}{2V(\mu) \log \log n}\right)^{\frac{k}{2}} \pi_{k,j,n} (\bar{X}_n - \mu)^j \\ &= \overline{\lim}_n \sum_{j=1}^k \left(\frac{n}{2V(\mu) \log \log n}\right)^{\frac{k-j}{2}} \pi_{k,j,n} \left(\frac{n}{2V(\mu) \log \log n}\right)^{\frac{j}{2}} (\bar{X}_n - \mu)^j \\ &= 1 \quad (\text{c.s.}), \end{aligned}$$

resultado que se obtiene a partir del corolario 2.7.1 y de la ley del logaritmo iterado.

□

Teorema 2.7.2 *Sea $h \in \mathcal{U}_n^2(\Omega_0)$, $n \geq n_0$, $k_0 = \min\{j > 0 : D_\mu^j h(\mu_0) \neq 0\}$ y $\mu_0 \in \Omega_0$. Sea $T_n(\bar{X}_n)$ el estimador insesgado de $h(\mu_0)$ en Ω_0 , para cada $n \geq n_0$. Entonces,*

$$(i) \quad \forall \alpha < k_0/2, \quad n^\alpha \{T_n(\bar{X}_n) - h(\mu_0)\} \xrightarrow{P_{\mu_0}} 0, \quad n \rightarrow \infty.$$

(ii) *La sucesión $\{T_n(\bar{X}_n)\}$ tiene un límite asintótico*

$$n^{\frac{k_0}{2}} \{T_n(\bar{X}_n) - h(\mu_0)\} \xrightarrow{d} \frac{D_\mu^{k_0} h(\mu_0)}{k_0!} V(\mu_0)^{\frac{k_0}{2}} H_{e_{k_0}}(Z), \quad n \rightarrow \infty. \quad (2.7.19)$$

(iii)

$$\overline{\lim}_n \left(\frac{n}{2V(\mu) \log \log n} \right)^{\frac{k_0}{2}} \{T_n(\bar{X}_n) - h(\mu_0)\} = \frac{D_\mu^{k_0} h(\mu_0)}{k_0!} V(\mu_0)^{\frac{k_0}{2}}, \quad (\text{c.s.}). \quad (2.7.20)$$

Demostración.

(i) Teniendo en cuenta la igualdad

$$n^\alpha \{T_n(\bar{X}_n) - h(\mu_0)\} = n^\alpha R_{k_0}(\bar{X}_n, \mu_0),$$

y aplicando el lema 2.7.3 obtenemos el resultado esperado.

(ii) De igual forma,

$$\begin{aligned} n^{\frac{k_0}{2}} [T_n(\bar{X}_n) - h(\mu_0)] &= n^{\frac{k_0}{2}} a_{k_0, n} P_{k_0}(\bar{X}_n, \mu_0) + n^{\frac{k_0}{2}} R_{k_0+1}(\bar{X}_n, \mu_0) \\ &= T_{1, n} + T_{2, n}. \end{aligned} \quad (2.7.21)$$

Aplicando el lema 2.7.3

$$T_{2,n} = n^{\frac{k_0}{2}} R_{k_0+1}(\bar{X}_n, \mu_0) \xrightarrow{P_{\mu_0}} 0, \quad n \rightarrow \infty.$$

Por otra parte, teniendo en cuenta los resultados del lema 2.7.4 y del teorema 2.7.1

$$T_{1,n} = n^{\frac{k_0}{2}} a_{k_0,n} P_{k_0}(\bar{X}_n, \mu_0) \xrightarrow{d} \frac{D_{\mu}^{k_0} h(\mu_0)}{k_0!} V(\mu_0)^{\frac{k_0}{2}} H_{e_{k_0}}(Z).$$

La conclusión del teorema se sigue de la aplicación del lema de Slutsky.

(iii) La demostración de (2.7.20) es similar a partir de (2.7.18).

□

Resulta interesante comparar las distribuciones límites del umvue en Ω_0 y del estimador de máxima verosimilitud. Por eso incluimos este lema cuya demostración puede verse en Serfling (1980) [42].

Lema 2.7.5 *Sea $\{Y_n\}$ una sucesión de variables aleatorias con distribuciones $AN(\mu, \sigma_n^2)$ con $\sigma_n \rightarrow 0$. Sea g una función real, k_0 veces diferenciable ($k_0 \geq 1$) respecto de μ con $D_{\mu}^{k_0} g(\mu) \neq 0$ y $D_{\mu}^j g(\mu) = 0 \forall j < k_0$. Entonces*

$$\frac{g(Y_n) - g(\mu)}{\frac{1}{k_0!} D_{\mu}^{k_0} g(\mu) \sigma_n^{k_0}} \xrightarrow{d} Z^{k_0}, \quad Z \sim N(0, 1).$$

Por el teorema de Zehna es obvio que el estimador de máxima verosimilitud

de $h(\mu)$ es $h(\bar{X}_n)$. A partir aquí, tomando $Y_n = \sqrt{n}(\bar{X}_n - \mu_0)$, es fácil demostrar que

$$n^{k_0/2} \{h(\bar{X}_n) - h(\mu_0)\} \xrightarrow{d} \frac{D_\mu^{k_0} h(\mu_0)}{k_0!} V^{\frac{k_0}{2}}(\mu_0) Z^{k_0}, \quad n \rightarrow \infty. \quad (2.7.22)$$

Teniendo en cuenta la expresión (2.7.3) de los primeros polinomios de Hermite se tiene que :

- Si $h'(\mu_0) \neq 0$, entonces:

$$\sqrt{n}\{T_n(\bar{X}_n) - h(\mu_0)\} \xrightarrow{d} N(0, D_\mu h(\mu_0)^2 V(\mu_0)). \quad (2.7.23)$$

- Si $h'(\mu_0) = 0$ pero $h''(\mu_0) \neq 0$, entonces:

$$\frac{2n}{D_\mu^2 h(\mu_0) V(\mu_0)} \{T_n(\bar{X}_n) - h(\mu_0)\} + 1 \xrightarrow{d} \chi_1^2.$$

Nótese que si $k_0 = 1$ la distribución límite del umvue y del estimador de máxima verosimilitud coinciden, es decir el umvue y el estimador de máxima verosimilitud son asintóticamente equivalentes. Sin embargo, para $k_0 > 1$, este resultado no es cierto tal y como se pone de manifiesto comparando (2.7.22) y (2.7.19).

2.8 Cotas inferiores para la varianza de estimadores insesgados

Según vimos en el primer capítulo, teorema 1.4.1, para cualquier $V_n = V(X_1, \dots, X_n)$ estimador insesgado de $h(\mu)$,

$$\text{Var}_\mu(V_n) \geq B_{m,n}(h, \mu) = \sum_{k=1}^m a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2 \quad \forall \mu \in \Omega_0.$$

La sucesión de cotas $\{B_{m,n}(h, \mu)\}_{m \geq 1}$ constituye una sucesión no decreciente de cotas inferiores para la varianza de los estimadores insesgados de $h(\mu)$ para un tamaño muestral n . De hecho, la cota de orden 1 es

$$B_{1,n}(h, \mu) = \frac{(D_\mu h(\mu))^2 V(\mu)}{n},$$

que es la cota de Fréchet-Cramer-Rao. En general, $B_{m,n}(h, \mu)$ es la m -ésima cota de Bhattacharyya. Dicha cota se alcanza si y solo si $T_n(\bar{X}_n)$ es un polinomio de grado m en \bar{X}_n , este resultado es debido a Fend (1959) [18].

Es interesante estudiar si la varianza de los estimadores insesgados es el límite de las cotas de Bhattacharyya. Por ejemplo, Khan (1980) [25] investiga este problema en el caso particular de la familia exponencial natural con varianza cuadrática. Si $h \in \mathcal{U}_n^2(\Omega_0)$ la serie $\sum_{k=1}^m a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2$ es finita para cualquier $\mu \in \Omega_0$, entonces es evidente que

$$\lim_m B_{m,n}(h, \mu) = \sum_{k=1}^{\infty} a_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2 = \text{Var}_\mu(T_n(\bar{X}_n)),$$

y como consecuencia el límite de las cotas de Bhattacharyya es la varianza del estimador insesgado en cualquier familia exponencial. Por lo que contrariamente a lo que conjetura Khan, el límite de las cotas de Bhattacharyya

coincide con la varianza del umvue en cualquier familia paramétrica.

Teniendo en cuenta el comportamiento límite de los polinomios $P_{k,n}(\bar{X}_n, \mu)$, ver el lema 2.7.4, y el de los coeficientes de Fourier $a_{k,n}(\mu)$, ver el teorema 2.7.1, podemos deducir algunas propiedades límites de las cotas, $B_{k,n}(h, \mu)$, cuando el tamaño muestral tiende a infinito:

$$\lim_{n \rightarrow \infty} n^{(m+1)} (B_{m+1,n}(h, \mu) - B_{m,n}(h, \mu)) = \frac{(D_\mu^{m+1} h(\mu))^2 V^{m+1}(\mu)}{(m+1)!}.$$

2.9 Comparaciones con el estimador de máxima verosimilitud

En este apartado, basándonos en los desarrollos obtenidos en términos de los polinomios ortogonales, vamos a realizar comparaciones asintóticas entre el estimador de máxima verosimilitud y el estimador umvue de una función $h(\mu)$.

Consideremos $h \in \mathcal{U}_{n_0}^2(\Omega_0)$ y sea $T_n(\bar{X}_n)$, $n \geq n_0$, su correspondiente estimador insesgado en Ω_0 .

Proposición 2.9.1 *Para cada $\mu \in \Omega_0$ se verifica*

$$n\{T_n(\bar{X}_n) - h(\bar{X}_n)\} \xrightarrow{P_\mu} -\frac{D_\mu^2 h(\mu)V(\mu)}{2}. \quad (2.9.1)$$

Demostración.

Como $h \in \mathcal{U}_{n_0}^2(\Omega_0)$, h es analítica en Ω_0 , por lo tanto

$$h(\bar{X}_n) = \sum_{j=0}^{\infty} \frac{D_\mu^j h(\mu)}{j!} (\bar{X}_n - \mu)^j, \quad n \geq n_0. \quad (2.9.2)$$

Teniendo en cuenta la expresión del estimador insesgado

$$T_n(\bar{X}_n) = \sum_{j=0}^{\infty} a_{j,n}(\mu) P_{j,n}(\bar{X}_n, \mu), \quad (2.9.3)$$

podemos escribir

$$\begin{aligned} T_n(\bar{X}_n) &= h(\mu) + D_\mu h(\mu)(\bar{X}_n - \mu) + \\ &+ \frac{D_\mu^2 h(\mu)}{2} \left[(\bar{X}_n - \mu)^2 - \frac{V'(\mu)}{n}(\bar{X}_n - \mu) - \frac{V(\mu)}{n} \right] \\ &+ \sum_{j=3}^{\infty} a_{j,n}(\mu) P_{j,n}(\bar{X}_n, \mu). \end{aligned} \quad (2.9.4)$$

Restando de (2.9.4) la expresión (2.9.2), se obtiene

$$\begin{aligned} n \{T_n(\bar{X}_n) - h(\bar{X}_n)\} &= -\frac{D_\mu^2 h(\mu)}{2} V(\mu) - \frac{D_\mu^2 h(\mu)}{2} V'(\mu)(\bar{X}_n - \mu) + \\ &+ nR_3(\bar{X}_n, \mu) + n \sum_{j=3}^{\infty} \frac{D_\mu^j h(\mu)}{j!} (\bar{X}_n - \mu)^j \\ &= T_1 + T_2 + T_3 + T_4. \end{aligned}$$

Veamos la convergencia en probabilidad de los tres últimos sumandos:

- $T_2 = -\frac{D_\mu^2 h(\mu)}{2} V'(\mu)(\bar{X}_n - \mu)$, converge a 0 en probabilidad, por la consistencia de \bar{X}_n .
- Aplicando el lema 2.7.3, se tiene que $T_3 = nR_3(\bar{X}_n, \mu)$, converge a 0 en probabilidad.
- El último sumando, $T_4 = n \sum_{j=3}^{\infty} \frac{D_\mu^j h(\mu)}{j!} (\bar{X}_n - \mu)^j$, utilizando la expresión del resto de Lagrange, verifica que

$$n \sum_{j=3}^{\infty} \frac{D_\mu^j h(\mu)}{j!} (\bar{X}_n - \mu)^j = n \frac{D_\mu^3 h(\xi_n)}{3!} (\bar{X}_n - \mu)^3$$

con $\xi_n = \mu + \alpha_n(\bar{X}_n - \mu)$, para algún $\alpha_n \in (0, 1)$, una variable aleatoria, con la propiedad

$$\xi_n \xrightarrow{p} \mu \quad n \rightarrow \infty.$$

De esta forma se tiene la convergencia a 0 en probabilidad de este último sumando.

Por lo tanto, teniendo en cuenta la convergencia en probabilidad de cada sumando, se obtiene el resultado esperado.

□

Según (2.7.23), la eficiencia asintótica relativa del estimador de máxima verosimilitud con respecto al umvue en Ω_0 , $ARE[h(\bar{X}_n), T_n(\bar{X}_n)]$, es igual a 1. Se tiene que el estimador de máxima verosimilitud y el umvue de $h(\mu)$ en Ω_0 son equivalentes en términos de la eficiencia asintótica relativa. Cuando esto ocurre, se pueden establecer comparaciones de segundo orden entre ambos estimadores, calculando para ello la deficiencia asintótica esperada, ver Hodges y Lehmann [22].

Definición 2.9.1 Sean $T_{1,n}$ y $T_{2,n}$, dos sucesiones de estimadores de $h(\mu)$, con errores cuadráticos medios dados por

$$ECM(T_{i,n}) = \frac{a}{n^r} + \frac{b_i}{n^{r+s}} + o\left(\frac{1}{n^{r+s}}\right), \quad i = 1, 2.$$

Se define la deficiencia asintótica esperada de $T_{2,n}$ respecto a $T_{1,n}$ como

$$AED[T_{2,n}, T_{1,n}] = \begin{cases} \frac{b_2 - b_1}{ar} & \text{si } s = 1 \\ \infty & \text{si } 0 < s < 1 \\ 0 & \text{si } s > 1. \end{cases} \quad (2.9.5)$$

El siguiente teorema nos proporciona la expresión de la deficiencia asintótica esperada del estimador de máxima verosimilitud, $h(\bar{X}_n)$ con respecto a $T_n(\bar{X}_n)$, lo cual nos permite establecer comparaciones de orden superior.

Teorema 2.9.1 *Sea $h \in \mathcal{U}_{n_0}^2(\Omega_0)$ para cierto natural $n_0 \geq 1$. Entonces*

(i) *El sesgo del estimador de máxima verosimilitud de $h(\mu)$ viene dado por*

$$\begin{aligned} \text{sesgo}_\mu(h(\bar{X}_n)) &= \\ &= \frac{D_\mu^2 h(\mu)V(\mu)}{2n} + \frac{V(\mu)}{n^2} \left[\frac{D_\mu^3 h(\mu)V'(\mu)}{6} + \frac{D_\mu^4 h(\mu)V(\mu)}{8} \right] + o(n^{-2}). \end{aligned} \quad (2.9.6)$$

(ii) *Para el error cuadrático medio del estimador de máxima verosimilitud de $h(\mu)$ se obtiene la siguiente expresión*

$$\begin{aligned} ECM_\mu(h(\bar{X}_n)) &= \\ &= \frac{(D_\mu h(\mu))^2 V(\mu)}{n} + \frac{D_\mu h(\mu)V(\mu)}{n^2} [D_\mu^3 h(\mu)V(\mu) + D_\mu^2 h(\mu)V'(\mu)] \\ &+ \frac{3}{4n^2} (D_\mu^2 h(\mu))^2 V^2(\mu) + o(n^{-2}). \end{aligned} \quad (2.9.7)$$

(iii) *Si $D_\mu h(\mu) \neq 0$ la deficiencia asintótica esperada del estimador de máxima verosimilitud, $h(\bar{X}_n)$, con respecto al umvue en Ω_0 , $T_n(\bar{X}_n)$, viene dada por*

$$\begin{aligned} AED_\mu[h(\bar{X}_n), T_n(\bar{X}_n)] &= \\ &= \frac{D_\mu^3 h(\mu)V(\mu)}{D_\mu h(\mu)} + \frac{D_\mu^2 h(\mu)V'(\mu)}{D_\mu h(\mu)} + \frac{1}{4} \left(\frac{D_\mu^2 h(\mu)}{D_\mu h(\mu)} \right)^2 V(\mu). \end{aligned} \quad (2.9.8)$$

Demostración.

- (i) Sea $g_n(\mu) = E_\mu[h(\bar{X}_n)]$ está bien definida en Ω_0 . Teniendo en cuenta la expresión (2.9.2) de $h(\bar{X}_n)$ e intercambiando la serie con la integral, obtenemos

$$g_n(\mu) = \sum_{j=0}^{\infty} \frac{D_\mu^j h(\mu)}{j!} E[(\bar{X}_n - \mu)^j] \quad \mu \in \Omega_0 \quad (2.9.9)$$

A partir de (2.5.2), función generatriz de momentos centrales de \bar{X}_n , podemos obtener

$$\mu_1 = 0$$

$$\mu_2 = \frac{V(\mu)}{n}$$

$$\mu_3 = \frac{V(\mu)V'(\mu)}{n^2}$$

$$\mu_4 = \frac{V^2(\mu)V^2(\mu) + (V'(\mu))^2V(\mu)}{n^3} + \frac{3V^2(\mu)}{n^2}$$

siendo $\mu_j = E[(\bar{X}_n - \mu)^j]$

Sustituyendo en (2.9.9),

$$g_n(\mu) = h(\mu) + \frac{D_\mu^2 h(\mu)}{2n} V(\mu) + \frac{D_\mu^3 h(\mu)}{6n^2} V(\mu)V'(\mu) + \frac{D_\mu^4 h(\mu)}{8n^2} V^2(\mu) + o(n^{-2}) \quad (2.9.10)$$

y de forma inmediata se obtiene (2.9.6).

- (ii) Para determinar la expresión del error cuadrático medio de $h(\bar{X}_n)$, tendremos en cuenta que

$$ECM_\mu(h(\bar{X}_n)) = Var_\mu(h(\bar{X}_n)) + (sesgo_\mu(h(\bar{X}_n)))^2.$$

Para n suficientemente grande, $Var_\mu(h(\bar{X}_n)) < \infty$ (ya que $h(\bar{X}_n)$ es asintóticamente eficiente) para cada $\mu \in \Omega_0$. Entonces $h(\bar{X}_n)$ será el estimador insesgado de $g_n(\mu)$ en Ω_0 , cuyo desarrollo en términos de po-

linomios ortogonales es

$$h(\bar{X}_n) = \sum_{k=0}^{\infty} d_{k,n}(\mu) P_{k,n}(\bar{X}_n, \mu), \quad \forall \mu \in \Omega_0$$

con varianza

$$Var_{\mu} (h(\bar{X}_n)) = \sum_{k=1}^{\infty} d_{k,n}^2(\mu) \|P_{k,n}\|_{2,\mu}^2, \quad (2.9.11)$$

siendo

$$d_{k,n}(\mu) = \frac{\langle h(\bar{X}_n), P_{k,n}(\bar{X}_n, \mu) \rangle_{\mu}}{\|P_{k,n}\|_{2,\mu}^2}.$$

En particular

$$d_{0,n} = g_n(\mu) \quad d_{1,n} = D_{\mu} g_n(\mu) \quad d_{2,n} = \frac{D_{\mu}^2 g_n(\mu)}{2}.$$

Derivando k veces la expresión (2.9.10) obtenemos

$$D_{\mu}^k g_n(\mu) = D_{\mu}^k h(\mu) + \frac{1}{2n} D_{\mu}^k (D_{\mu}^2 h(\mu) V(\mu)) + o(n^{-1})$$

y aplicando la fórmula de Leibnitz, se obtiene

$$D_{\mu}^k g_n(\mu) = D_{\mu}^k h(\mu) + \frac{1}{2n} \sum_{j=0}^k \binom{k}{j} D_{\mu}^{j+2} h(\mu) D_{\mu}^{k-j} V(\mu)$$

En particular, para $k = 1$

$$(D_{\mu} g_n(\mu))^2 = (D_{\mu} h(\mu))^2 + \frac{1}{n} D_{\mu} h(\mu) [D_{\mu}^3 h(\mu) V(\mu) + D_{\mu}^2 h(\mu) V'(\mu)] + o(n^{-1}), \quad (2.9.12)$$

y para $k = 2$,

$$\begin{aligned} (D_{\mu}^2 g_n(\mu))^2 &= \\ &= (D_{\mu}^2 h(\mu))^2 + \frac{D_{\mu}^2 h(\mu)}{n} [2D_{\mu}^3 h(\mu) V'(\mu) + D_{\mu}^4 h(\mu) V(\mu) + D_{\mu}^2 h(\mu) V''(\mu)] \\ &+ o(n^{-1}). \end{aligned} \quad (2.9.13)$$

Sustituyendo (2.9.12) y (2.9.13) en la expresión (2.9.11) de la varianza de $h(\bar{X}_n)$,

$$\begin{aligned} \text{Var}_\mu (h(\bar{X}_n)) &= \\ &= \frac{(D_\mu h(\mu))^2 V(\mu)}{n} + \frac{D_\mu h(\mu) V(\mu)}{n^2} (D_\mu^3 h(\mu) V(\mu) + D_\mu^2 h(\mu) V'(\mu)) + \\ &+ \frac{(D_\mu^2 h(\mu))^2 V^2(\mu)}{2n^2} + o(n^{-2}). \end{aligned} \quad (2.9.14)$$

Además, teniendo en cuenta (2.9.6),

$$(\text{sesgo}_\mu(h(\bar{X}_n)))^2 = \frac{(D_\mu^2 h(\mu))^2 V^2(\mu)}{4n^2} + o(n^{-2}). \quad (2.9.15)$$

Por tanto, sumando (2.9.14) y (2.9.15) obtenemos la expresión para el error cuadrático medio de $h(\bar{X}_n)$.

(iii) Para obtener la deficiencia asintótica esperada de $h(\bar{X}_n)$ con respecto a $T_n(\bar{X}_n)$, observamos que

$$\begin{aligned} \text{ECM}_\mu (T_n(\bar{X}_n)) &= \text{Var}_\mu (T_n(\bar{X}_n)) = \\ &= \frac{(D_\mu h(\mu))^2 V(\mu)}{n} + \frac{(D_\mu^2 h(\mu))^2 V^2(\mu)}{2n^2} + o(n^{-2}), \end{aligned} \quad (2.9.16)$$

y teniendo en cuenta la expresión (2.9.7) del error cuadrático medio del estimador de máxima verosimilitud de $h(\mu)$, y la definición de la deficiencia asintótica (2.9.5) obtenemos,

$$\begin{aligned} \text{AED} [h(\bar{X}_n), T_n(\bar{X}_n)] &= \frac{D_\mu h(\mu) V(\mu)}{D_\mu h(\mu)^2 V(\mu)} [D_\mu^3 h(\mu) V(\mu) + D_\mu^2 h(\mu) V'(\mu)] \\ &+ \frac{1}{V(\mu) (D_\mu h(\mu))^2} \left[\frac{3}{4} (D_\mu^2 h(\mu))^2 V^2(\mu) - \frac{1}{2} (D_\mu^2 h(\mu))^2 V^2(\mu) \right] \\ &= V(\mu) \left[\frac{D_\mu^3 h(\mu)}{D_\mu h(\mu)} + \frac{1}{4} \left(\frac{D_\mu^2 h(\mu)}{D_\mu h(\mu)} \right)^2 \right] + V'(\mu) \frac{D_\mu^2 h(\mu)}{D_\mu h(\mu)} \end{aligned}$$

y por tanto se obtiene la expresión (2.9.8) para la deficiencia asintótica esperada.

□

Para establecer los resultados anteriores, hemos comparado la expresión del estimador insesgado utilizando polinomios ortogonales y la del error cuadrático medio del estimador de máxima verosimilitud utilizando el desarrollo en serie de Taylor. Se obtiene el mismo resultado si se aproxima la varianza del estimador insesgado utilizando el desarrollo de Taylor, Hwang y Hu (1990) [23].

Si para cierto $\mu_0 \in \Omega_0$, $D_\mu h(\mu_0) = 0$, no es posible establecer comparaciones asintóticas entre $h(\bar{X}_n)$ y $T_n(\bar{X}_n)$ mediante la eficiencia asintótica relativa (ARE) ó la deficiencia asintótica esperada (AED). Una alternativa a esta situación sería estudiar el comportamiento límite de la eficiencia del riesgo, LRE.

Definición 2.9.2 (ver Lehman (1983) [28]) *Supongamos que T_{1,n_1} y T_{2,n_2} son dos series de estimadores de $h(\mu)$, siendo $n_2 = n_2(n_1)$ tal que sus funciones riesgo $R(\mu, T_{i,n_i})$ $i = 1, 2$ verifican que*

$$n_i^\tau R(\mu, T_{i,n_i}) \quad i = 1, 2$$

tienden a un límite común τ^2 . Diremos que el límite de la eficiencia del riesgo de T_{1,n_1} respecto a T_{2,n_2} es

$$LRE(T_{1,n_1}, T_{2,n_2}) = \lim \frac{n_2(n_1)}{n_1}$$

supuesto que este límite existe cuando $n_1 \rightarrow \infty$ y es independiente de la secuencia particular $n_2(n_1)$ elegida.

Se verifica el siguiente teorema, cuya demostración puede verse en Lehman (1983) [28].

Teorema 2.9.2 Sean $T_{1,n}$ y $T_{2,n}$ dos series de estimadores de $h(\mu)$, tales que sus funciones riesgo verifican que $\lim_{n \rightarrow \infty} n^r R(\mu, T_{i,n}) = \tau_i^2 > 0$, $i = 1, 2$ entonces, el límite de la eficiencia del riesgo de $T_{2,n}$ con respecto a $T_{1,n}$ es

$$LRE_{\mu}(T_{2,n}, T_{1,n}) = \left(\frac{\tau_1^2}{\tau_2^2} \right)^{1/r}.$$

Consideraremos como función de riesgo el error cuadrático medio.

Teorema 2.9.3 Si para cierto $\mu_0 \in \Omega_0$, $D_{\mu}h(\mu_0) = 0$ entonces, el LRE de la sucesión de estimadores $h(\bar{X}_n)$ respecto a la sucesión $T_n(\bar{X}_n)$ es

$$LRE_{\mu_0}(h(\bar{X}_n), T_n(\bar{X}_n)) = \sqrt{\frac{2}{3}}.$$

Demostración.

Si $D_{\mu}h(\mu_0) = 0$, según (2.9.16) y (2.9.7), se tiene, respectivamente

$$Var_{\mu_0}(T_n(\bar{X}_n)) = \frac{(D_{\mu}^2 h(\mu_0))^2 V^2(\mu_0)}{2n^2} + o(n^{-2}),$$

$$ECM_{\mu_0}(h(\bar{X}_n)) = \frac{3}{4n^2} (D_{\mu}^2 h(\mu_0))^2 V^2(\mu_0) + o(n^{-2}).$$

Aplicando el teorema anterior, con $r = 2$

$$LRE_{\mu_0}(h(\bar{X}_n), T_n(\bar{X}_n)) = \left[\frac{\frac{1}{2}(D_{\mu}^2 h(\mu_0))^2 V^2(\mu_0)}{\frac{3}{4}(D_{\mu}^2 h(\mu_0))^2 V^2(\mu_0)} \right]^{1/2} = \sqrt{\frac{2}{3}}.$$

□

Podemos concluir que el estimador umvue en Ω_0 de $h(\mu)$ es más eficiente que el estimador de máxima verosimilitud, en el sentido de la definición 2.9.2.

2.10 Aplicación a la distribución Normal

La teoría desarrollada para la obtención de estimadores insesgados mediante desarrollos ortogonales puede ser extendida a distribuciones que dependan de más de un parámetro. Este es el caso de la distribución Normal.

Sea X_1, X_2, \dots, X_n una m.a.s. procedente de una población $N(\mu, \sigma^2)$. Nuestro objetivo es determinar el estimador insesgado de mínima varianza de una función $h(\mu, \sigma^2)$, basada en el estadístico conjuntamente suficiente y completo $(\bar{X}_n, S_{c_n}^2)$.

En primer lugar estudiamos el caso en que solo μ es desconocido. Un estadístico suficiente y completo es $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$ cuya distribución es $N\left(\mu, \frac{\sigma^2}{n}\right)$. Basándonos en la teoría general, el sistema de polinomios mónicos ortogonales $\{P_{k,n}(s, \mu)\}$ asociados a esta densidad viene dado por

$$P_{k,n}(s, \mu) = \left(\frac{\sigma}{\sqrt{n}}\right)^k H_{e_k}\left(\frac{(s - \mu)\sqrt{n}}{\sigma}\right),$$

con $H_{e_k}(t)$ los polinomios de Hermite, definidos en (2.7.1), siendo su norma

$$\|P_{k,\mu}\|_{2,\mu}^2 = k! \left(\frac{\sigma^2}{n}\right)^k.$$

Se tiene el siguiente teorema de caracterización de funciones del parámetro μ que admiten umvue, ver López Blázquez et al. (1997) [32]:

Teorema 2.10.1 *Sea $\mathcal{U}_n^2(\mathbb{R})$ el conjunto de funciones reales del parámetro μ que admiten un estimador insesgado para una muestra de tamaño n . Entonces, si $h \in \mathcal{U}_n^2(\mathbb{R})$ y $T_n(\bar{X}_n)$ es el estimador insesgado de $h(\mu)$, se verifica que*

(i) La expresión del desarrollo ortogonal de $T_n(\bar{X}_n)$ viene dada por

$$T_n(\bar{X}_n) = \sum_{k=0}^{\infty} \frac{D^k h(\mu) \sigma^k}{k! n^{k/2}} H_{e_k} \left(\frac{(\bar{X}_n - \mu) \sqrt{n}}{\sigma} \right) \quad (2.10.1)$$

(ii)

$$Var_{\mu}(T_n(\bar{X}_n)) = \sum_{k=1}^{\infty} \frac{(D^k h(\mu))^2}{k!} \left(\frac{\sigma^2}{n} \right)^k \quad \forall \mu \in \mathbb{R}.$$

$$(iii) \mathcal{U}_n^2(\mathbb{R}) = \left\{ h \in \mathcal{A}(\mathbb{R}) : \sum_{k=1}^{\infty} \frac{(D^k h(\mu))^2 \sigma^{2k}}{n^k k!} < \infty \quad \mu \in \mathbb{R} \right\},$$

siendo $\mathcal{A}(\mathbb{R})$ el conjunto de funciones analíticas en \mathbb{R} .

Según se vió en (2.5.11), escogeremos $\mu = \bar{X}_n$, obteniéndose la siguiente expresión para el umvue y la varianza del mismo:

$$T_n(\bar{X}_n) = \sum_{j=0}^{\infty} \frac{D^j h(\bar{X}_n) \sigma^j}{j! n^{j/2}} H_j(0) = \sum_{j=0}^{\infty} \frac{(-1)^j D^{2j} h(\bar{X}_n) \sigma^{2j}}{j! (2n)^j},$$

$$Var_{\mu}(T_n(\bar{X}_n)) = \sum_{j=1}^{\infty} \frac{(D^j h(\bar{X}_n))^2 \sigma^{2j}}{j! n^j}.$$

Obsérvese que si en la expresión (2.10.1) hacemos $\mu = 0$, el estimador obtenido coincide con el dado por Gray et al. (1973) [21]. Ellos obtuvieron su estimador por un procedimiento completamente distinto: aplicando jackknife al estimador de máxima verosimilitud y después el procedimiento de Blackwell-Rao.

Si μ es conocido y σ^2 es el parámetro desconocido, un estadístico suficiente y completo es $S_n^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{n}$, siendo su distribución $Ga\left(\frac{n}{2}, \frac{n}{2\sigma^2}\right)$.

El sistema de polinomios ortogonales $\{Q_{k,n}(s, \sigma^2)\}$ asociados a esta densidad viene dado por

$$Q_{k,n}(s, \sigma^2) = \frac{(-2)^k k! \sigma^{2k}}{n^k} L_k^{(\frac{n}{2}-1)} \left(\frac{ns}{2\sigma^2} \right)$$

con $L_k^{(\frac{n}{2}-1)}(t)$ los polinomios de Laguerre generalizados, ortogonales respecto a la función peso $w(t) = e^{-t} t^{\frac{n}{2}-1}$, definidos en (2.5.13). Por tanto

$$\|Q_{k,n}\|_{2,\sigma^2}^2 = \frac{(2\sigma^2)^{2k} k! \Gamma(k + \frac{n}{2})}{n^{2k} \Gamma(\frac{n}{2})}$$

Se tiene el siguiente teorema de caracterización de funciones del parámetro σ^2 que admiten umvue.

Teorema 2.10.2 *Sea $\mathcal{U}_n^2(\mathbb{R}^+)$ el conjunto de funciones reales del parámetro σ^2 que admiten umvue para una muestra de tamaño n . Entonces, si $h \in \mathcal{U}_n^2(\mathbb{R}^+)$ y $T_n(S_n^2)$ su correspondiente umvue*

(i) *La expresión del desarrollo orthogonal de $T_n(S_n^2)$ viene dada por*

$$T_n(S_n^2) = \sum_{k=0}^{\infty} \frac{(-1)^k D^k h(\sigma^2) \sigma^{2k} \Gamma(\frac{n}{2})}{\Gamma(k + \frac{n}{2})} L_k^{(\frac{n}{2}-1)} \left(\frac{nS_n^2}{2\sigma^2} \right) \quad (2.10.2)$$

(ii) $Var_{\sigma^2}(T_n(S_n^2)) = \sum_{k=1}^{\infty} \frac{(D^k h(\sigma^2))^2 (\sigma^{4k}) \Gamma(\frac{n}{2})}{k! \Gamma(k + \frac{n}{2})} \quad \forall \sigma^2 > 0$

(c) $\mathcal{U}_n^2(\mathbb{R}^+) = \left\{ h \in \mathcal{A}(0, +\infty) : \sum_{k=1}^{\infty} \frac{(D^k h(\sigma^2))^2 (\sigma^{4k}) \Gamma(\frac{n}{2})}{k! \Gamma(k + \frac{n}{2})} < \infty \quad \sigma^2 > 0 \right\}$,

siendo $\mathcal{A}(0, +\infty)$ el conjunto de funciones analíticas en $(0, +\infty)$.

Escogiendo $\sigma^2 = S_n^2$ se obtiene la siguiente expresión para el umvue y la varianza del mismo:

$$T_n(S_n^2) = \sum_{k=0}^{+\infty} \frac{(-1)^k D^k h(S_n^2) S_n^{2k} \Gamma\left(\frac{n}{2}\right)}{\Gamma\left(k + \frac{n}{2}\right)} L_k^{(\frac{n}{2}-1)}\left(\frac{n}{2}\right),$$

$$Var(T(S_n^2)) = \sum_{k=1}^{+\infty} \frac{(D^k h(S_n^2))^2 (S_n^2)^{2k} \Gamma\left(\frac{n}{2}\right)}{\Gamma\left(k + \frac{n}{2}\right) k!}.$$

Si tanto μ como σ^2 son desconocidos y nuestro objetivo es determinar el umvue de una función $h(\sigma^2)$, consideraremos como estadístico suficiente y completo $S_{c_n}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{n-1}$, cuya distribución es $Ga\left(\frac{n-1}{2}, \frac{n-1}{2\sigma^2}\right)$. El desarrollo es análogo al caso anterior sin más que sustituir n por $n-1$ y S_n^2 por $S_{c_n}^2$. Por tanto, el desarrollo ortogonal del umvue es

$$T_n(S_{c_n}^2) = \sum_{k=0}^{\infty} \frac{(-1)^k D^k h(\sigma^2) S_{c_n}^{2k} \Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(k + \frac{n-1}{2}\right)} L_k^{(\frac{n-3}{2})}\left(\frac{(n-1)S_{c_n}^2}{2\sigma^2}\right). \quad (2.10.3)$$

Si tanto μ como σ son desconocidos y queremos determinar el umvue de una función $h(\mu, \sigma^2)$, aplicaremos consecutivamente los resultados (2.10.1) y (2.10.3). Suponiendo σ^2 fijo, podemos considerar $h(\mu, \sigma^2)$ función de μ y aplicando (2.10.1) tendremos que el umvue, para esta nueva función, viene dado por

$$T_{1,n}(\bar{X}_n, \sigma^2) = \sum_{k=0}^{+\infty} \frac{D_1^k h(\bar{X}_n, \sigma^2) \sigma^k}{k! \sqrt{n^k}} H_k(0). \quad (2.10.4)$$

La función así obtenida depende solo de σ^2 , por tanto aplicando (2.10.3) a $T_{1,n}(\bar{X}_n, \sigma^2)$ obtendremos el umvue de $h(\mu, \sigma^2)$ cuya expresión es

$$T_n(\bar{X}_n, S_{c_n}^2) = \sum_{j=0}^{+\infty} D_2^j T_1(\bar{X}_n, S_{c_n}^2) \frac{(-1)^j S_{c_n}^{2j} \Gamma((n-1)/2)}{\Gamma(j + (n-1)/2)} L_j^{(\frac{n-3}{2})}\left(\frac{n-1}{2}\right). \quad (2.10.5)$$

Desarrollos similares al obtenido en (2.10.5), han sido dados por Woodward y Gary (1977) [54], pero utilizando un método completamente distinto.

Ejemplo 2.10.1 *Estimación insesgada de mínima varianza de $P[Y > X]$, bajo la hipótesis de que X e Y son variables aleatorias independientes, siendo la distribución de Y conocida. Sin pérdida de generalidad podemos suponer que $Y \sim \mathcal{N}(0, 1)$ y $X \sim \mathcal{N}(\mu, \sigma^2)$, μ y σ^2 desconocidos.*

Podemos expresar la función a estimar como

$$P[Y < X] = \phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right) = h(\mu, \sigma^2), \quad (2.10.6)$$

donde por ϕ representamos a la función de distribución de la $\mathcal{N}(0, 1)$.

Teniendo en cuenta que

$$D_{\mu}^k \phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right) = (-1)^{k-1} \left(\frac{1}{\sqrt{1 + \sigma^2}}\right)^k \phi\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right) H_{e_k-1}\left(\frac{\mu}{\sqrt{1 + \sigma^2}}\right),$$

y sustituyendo en (2.10.4) obtenemos

$$T_{1,n}(\bar{X}_n, \sigma^2) = \sum_{k=0}^{+\infty} \frac{(-1)^{k-1} \sigma^k H_{e_k}(0)}{k! \sqrt{n^k} \sqrt{1 + \sigma^2}^k} \phi\left(\frac{\bar{X}_n}{\sqrt{1 + \sigma^2}}\right) H_{e_k-1}\left(\frac{\bar{X}_n}{\sqrt{1 + \sigma^2}}\right).$$

A partir de esta expresión y de (2.10.5) obtenemos el estimador insesgado de mínima varianza de (2.10.6):

$$T_n(\bar{X}_n, S_{c_n}^2) =$$

$$= \sum_{j,k=0}^{\infty} D_{\sigma^2}^j \left(S_{c_n}^k D_{\mu}^k h(\bar{X}_n, \sigma^2) \right) \frac{(-1)^{k+j-1} S_{c_n}^2 j \Gamma\left(\frac{n-1}{2}\right)}{k! n^{\frac{k}{2}} \Gamma\left(j + \frac{n-1}{2}\right)} H_{e_k}(0) L_j^{\left(\frac{n-3}{2}\right)}\left(\frac{n-1}{2}\right),$$

siendo,

$$D_{\sigma^2}^j \left(S_{c_n}^k D_{\mu}^k h(\bar{X}_n, \sigma^2) \right) = \sum_{m=0}^j A_{m,j}(S_{c_n}^2) \phi\left(\frac{\bar{X}_n}{\sqrt{1 + S_{c_n}^2}}\right) H_{e_{k+2m-1}}\left(\frac{\bar{X}_n}{\sqrt{1 + S_{c_n}^2}}\right).$$

y

$$A_{m,j}(S_{c_n}^2) = \binom{j}{m} \left(\frac{1}{\sqrt{1 + S_{c_n}^2}}\right)^{k+2m} \frac{S_{c_n}^{k - \left(\frac{j-m}{2}\right)} \Gamma\left(\frac{k}{2} + 1\right)}{2^m \Gamma\left(\frac{k}{2} - (j-m) - 1\right)}.$$

Capítulo 3

Estimación en familias no regulares

3.1 Introducción

La teoría desarrollada en el capítulo 1, tiene un ámbito de aplicabilidad bastante general. En este capítulo mostramos como dicha teoría puede aplicarse a familias de distribuciones que no son regulares en el sentido clásico. Así pues, consideraremos el caso de una familia de distribuciones cuyo recorrido depende del parámetro.

A la vista de los resultados de este capítulo y del anterior podemos afirmar que el estudio de los desarrollos ortogonales de los estimadores proporciona una herramienta bastante general para el estudio de problemas relacionados con la estimación insesgada.

3.2 Estimación en distribuciones truncadas

Consideramos la familia de distribuciones

$$f(x, \theta) = \frac{1}{a(\theta) - a(\theta_0)}, \quad a(\theta_0) < x < a(\theta), \quad \theta \in \Theta. \quad (3.2.1)$$

Sin pérdida de generalidad podemos suponer $a(\theta_0) = 0$, quedando la densidad de la forma

$$f(x, \theta) = \frac{1}{a(\theta)}, \quad 0 < x < a(\theta). \quad (3.2.2)$$

La función de distribución asociada es

$$F(x, \theta) = \frac{x}{a(\theta)}, \quad 0 < x < a(\theta).$$

Consideraremos X_1, \dots, X_n una m.a.s. de dicha distribución. El estadístico $S_n = \max_{1 \leq i \leq n} \{X_i\}$ es suficiente y completo para θ siempre que $a(\cdot)$ sea estrictamente creciente, ver Stuart y Ord (1991) [48].

La función de densidad de S_n viene dada por

$$f_n(s, \theta) = \frac{ns^{n-1}}{a(\theta)^n}, \quad 0 < s < a(\theta).$$

Esta distribución queda determinada unívocamente por la secuencia de momentos $\{m_{k,n}(\theta) = E_\theta[S_n^k]\}_{k \geq 0}$. Esto es debido a que siempre que una distribución de probabilidad esté concentrada en un intervalo finito, los momentos la determinan completamente, ver Shirayev (1984) [45].

3.3 Polinomios Ortogonales

Para cada $\theta \in \Theta$ consideramos el espacio de funciones $L^2_{\theta,n}$ definido en (1.2.1).

Todos los momentos de S_n existen y son

$$E[S_n^k] = \frac{n}{n+k} a(\theta)^k, \quad \forall \theta, \quad k = 1, 2, \dots$$

Entonces podemos encontrar, en $L_{\theta,n}^2$ un sistema de polinomios mónicos ortogonales respecto a la función de densidad $f_n(s, \theta)$.

Vamos a obtener explícitamente dicho sistema, que lo denotaremos por $\{P_{k,n}(s, \theta)\}_{k \geq 0}$. Para ello consideraremos los polinomios de Jacobi, cuya expresión viene dada por

$$G_{k,p,q}(x) = \frac{\Gamma(p+k)}{\Gamma(p+2k)} \sum_{i=0}^k (-1)^i \binom{k}{i} \frac{\Gamma(p+2k-i)}{\Gamma(q+k-i)} x^{k-i}, \quad x \in (0, 1),$$

siendo $p - q > -1$, $q > 0$.

El sistema $\{G_{k,p,q}\}_{k \geq 0}$ es ortogonal respecto a la función peso

$$w(x) = (1-x)^{p-q} x^{q-1}, \quad p - q > -1, \quad q > 0, \quad x \in (0, 1),$$

siendo su norma

$$\|G_{k,p,q}\|^2 = \frac{k! \Gamma(p+k) \Gamma(q+k) \Gamma(k+p-q+1)}{(2k+p) \Gamma^2(2k+p)}.$$

Mediante un sencillo cambio de variables puede comprobarse que

$$P_{k,n}(s, \theta) = a(\theta)^k G_{k,n,n} \left(\frac{s}{a(\theta)} \right)$$

forman un sistema de polinomios mónicos ortogonales respecto a la función de densidad $f_n(s, \theta)$. Su expresión explícita es

$$P_{k,n}(s, \theta) = a(\theta)^k \frac{\Gamma(n+k)}{\Gamma(n+2k)} \sum_{i=0}^k (-1)^i \binom{k}{i} \frac{\Gamma(n+2k-i)}{\Gamma(n+k-i)} \left(\frac{s}{a(\theta)} \right)^{k-i}, \quad (3.3.1)$$

cuya norma es

$$\|P_{k,n}\|_{2,\theta}^2 = na(\theta)^{2k} \|G_{k,n,n}\|^2.$$

En particular,

$$\begin{aligned} P_{0,n}(s) &= 1, \\ P_{1,n}(s) &= s - \frac{n}{n+1}a(\theta), \\ P_{2,n}(s) &= s^2 - a(\theta)\frac{2(n+1)}{n+3}s + \frac{a^2(\theta)n(n+1)}{(n+2)(n+3)}. \end{aligned}$$

con normas respectivas

$$\begin{aligned} \|P_{0,n}\|_{2,\theta}^2 &= 1, \\ \|P_{1,n}\|_{2,\theta}^2 &= a(\theta)^2 \frac{n}{(n+2)(n+1)^2}, \\ \|P_{2,n}\|_{2,\theta}^2 &= a(\theta)^4 \frac{2n}{(n+4)(n+3)^2(n+2)^2}. \end{aligned}$$

3.4 Propiedades límites de los polinomios ortogonales

Para estudiar las propiedades límites de los polinomios ortogonales necesitamos algunos resultados previos.

Usando la caracterización de la convergencia en ley mediante la función característica, se tiene el siguiente resultado

Lema 3.4.1 Sean $\{Y_n\}_{n \geq 1}$ e Y variables aleatorias con momentos finitos de todos los ordenes, $\forall n, k \quad m_{k,n} = E[|Y_n|^k] < \infty \quad m_k = E[|Y|^k] < \infty$, tal que las secuencias $\{m_{k,n}\}$ y $\{m_k\}$ caracterizan a las distribuciones de Y_n y de Y respectivamente. Son equivalentes:

$$(i) Y_n \xrightarrow{d} Y \quad n \rightarrow \infty,$$

$$(ii) m_{k,n} = E[Y_n^k] \rightarrow m_k = E[Y^k], \quad n \rightarrow \infty, \quad \forall k \geq 0.$$

Corolario 3.4.1 *Sea g una función continua definida en \mathbb{R}^r (o en subconjunto de \mathbb{R}^r). Sean Y_n e Y en las condiciones del lema anterior. Entonces*

$$\lim_n g(m_{1,n}, \dots, m_{r,n}) = g(m_1, \dots, m_r).$$

De forma más general, si $\{g_n\}_{n \geq 1}$ es una sucesión de funciones continuas tal que g_n converge puntualmente a una función continua g , entonces

$$\lim_n g_n(m_{1,n}, \dots, m_{r,n}) = g(m_1, \dots, m_r).$$

Una aplicación importante de este corolario es la convergencia de los determinantes de las matrices de momentos de Y_n a la correspondiente matriz de momentos de Y .

Teorema 3.4.1 *Sean Y_n e Y variables aleatorias en las condiciones del lema anterior. Sean $\mathcal{M}_{k,n} = (m_{i+j,n})_{i,j=1,\dots,k}$ y $\mathcal{M}_k = (m_{i+j})_{i,j=1,\dots,k}$ las matrices de momentos asociadas a Y_n e Y respectivamente. Si $\det \mathcal{M}_k \neq 0$ entonces*

$$\lim_n P_{k,n}(s) = P_k(s) \quad y \quad \lim_n \|P_{k,n}\|^2 = \|P_k\|^2$$

donde $P_{k,n}(s)$ y $P_k(s)$ son los polinomios mónicos ortogonales de grado k asociados a las densidades de Y_n e Y respectivamente.

Demostración.

La demostración de este teorema es inmediata si tenemos en cuenta el método

de construcción polinomios ortogonales mediante determinantes de las matrices de momentos visto en la sección 1.5 de la presente memoria.

□

Sea X_1, \dots, X_n una muestra aleatoria simple procedente de la familia (3.2.2).

Se verifica que

$$W_n = n \left(1 - \frac{S_n}{a(\theta)} \right) \xrightarrow{d} W \sim \text{Exp}(1), \quad (3.4.1)$$

cuando $n \rightarrow \infty$, para cualquier $\theta \in \Theta$.

Sean $\{R_{k,n}(w, \theta)\}_{k \geq 0}$ y $\{P_{k,n}(s, \theta)\}_{k \geq 0}$ los sistemas de polinomios mónicos ortogonales asociados a las densidades de W_n y S_n respectivamente. Se puede comprobar fácilmente que

$$R_{k,n}(w, \theta) = P_{k,n} \left(\left(1 - \frac{w}{n} \right) a(\theta), \theta \right) \frac{(-1)^k n^k}{a(\theta)^k},$$

con norma, en sus respectivos espacios $L_{\theta,n}^2$:

$$\|R_{k,n}\|_{\theta}^2 = \frac{n^{2k+1} k!^2 \Gamma(n+k)^2}{(n+2k) \Gamma(n+2k)^2}.$$

Consideremos los polinomios de Laguerre generalizados definidos de forma recurrente por

$$kL_k^{(\alpha)}(w) = (2k + \alpha - 1 - w)L_{k-1}^{(\alpha)}(w) - (k + \alpha - 1)L_{k-2}^{(\alpha)}(w), \quad (3.4.2)$$

$$L_0^{(\alpha)} = 1, \quad L_1^{(\alpha)}(w) = (\alpha + 1) - w,$$

cuya expresión explícita fue dada en (2.5.13).

Los polinomios $\{(-1)^k k! L_k^{(0)}(w)\}_{k \geq 0}$ son mónicos y ortogonales con respecto a la densidad de una $f(w) = \exp(-w) w > 0$. Como la distribución de

W_n está determinada unívocamente por sus momentos, por el teorema 3.4.1 podemos decir que

$$\lim_n R_{k,n}(w, \theta) = (-1)^k k! L_k^{(0)}(w), \quad \forall \theta \in \Theta, k \geq 0, w \in \mathbb{R}^+,$$

y además

$$\lim_n \|R_{k,n}\|_{2,\theta}^2 = \|(-1)^k k! L_k^{(0)}\|^2 = k!^2, \quad \forall \theta.$$

Entonces, se tiene el siguiente resultado:

Teorema 3.4.2 Sean X_1, \dots, X_n variables aleatorias independientes e idénticamente distribuidas según (3.2.2) y sea $\{P_{k,n}(s, \theta)\}_{k \geq 0}$ el sistema de polinomios mónicos ortogonales asociados a la densidad de $S_n = \max_{1 \leq i \leq n} \{X_i\}$.
Entonces

$$(i) \lim_n \frac{n^k}{k! a(\theta)^k} P_{k,n} \left(\left(1 - \frac{w}{n}\right) a(\theta) \right) = L_k^{(0)}(w), \quad \forall \theta \in \Theta, k \geq 0, w \in \mathbb{R}^+.$$

$$(ii) \lim_n \left(\frac{n}{a(\theta)} \right)^{2k} \|P_{k,n}\|_{2,\theta}^2 = k!^2, \quad \forall \theta \in \Theta, k \geq 0.$$

Supongamos que el polinomio ortogonal de grado k con respecto a la densidad de S_n es expresado como

$$P_{k,n}(s, \theta) = \sum_{j=0}^k \pi_{k,j,n}(\theta) (a(\theta) - s)^j. \quad (3.4.3)$$

A partir del teorema anterior obtenemos

Corolario 3.4.2

$$\lim_n \frac{n^{k-j}}{k! a(\theta)^{k-j}} \pi_{k,j,n}(\theta) = \text{coef}(j, L_k^{(0)}), \quad \forall \theta \in \Theta, \quad 0 \leq j \leq k. \quad (3.4.4)$$

siendo $\text{coef}(j, L_k^{(0)})$ el coeficiente j -ésimo del k -ésimo polinomio de Laguerre.

Demostración.

Sea $w = n \left(1 - \frac{s}{a(\theta)} \right)$. Entonces

$$\begin{aligned} \lim_n \frac{n^k}{a(\theta)^k} P_{k,n}(s, \theta) &= \lim_n \frac{n^k}{a(\theta)^k} \sum_{j=0}^k \pi_{k,j,n}(\theta) (a(\theta) - s)^j \\ &= \lim_n \frac{n^k}{a(\theta)^k} \sum_{j=0}^k \pi_{k,j,n}(\theta) \left(a(\theta) - \left(1 - \frac{w}{n} \right) a(\theta) \right)^j \\ &= \lim_n \frac{n^k}{a(\theta)^k} \sum_{j=0}^k \pi_{k,j,n}(\theta) \left(\frac{w}{n} \right)^j a(\theta)^j = k! L_k^{(0)}(w). \end{aligned}$$

Por tanto, igualando término a término

$$\frac{n^{k-j}}{k! a(\theta)^{k-j}} \pi_{k,j,n}(\theta) \longrightarrow \text{coef}(j, L_k^{(0)}), \quad n \rightarrow \infty.$$

□

Como consecuencia de este corolario y debido a que $\pi_{k,k,n}(\theta_0) = 1$ y a que $\text{coef}(j, L_k^{(0)}) = \frac{(-1)^k}{k!}$ tenemos

$$\begin{aligned} \lim_n P_{k,n}(s, \theta_0) &= \sum_{j=0}^k \lim_n \frac{n^{k-j}}{a(\theta_0)^{k-j}} \pi_{k,j,n}(\theta_0) \lim_n \frac{a(\theta_0)^{k-j}}{n^{k-j}} (a(\theta_0) - s)^j = \\ &= k! \text{coef}(k, L_k^{(0)}) (a(\theta_0) - s)^k = (s - a(\theta_0))^k. \end{aligned}$$

En general, si $\{s_n\}_{n \geq 1}$ es una sucesión de números reales tales que $\lim_n s_n = a(\theta)$ entonces

$$\lim_n P_{k,n}(s_n, \theta_0) = (a(\theta) - a(\theta_0))^k. \quad (3.4.5)$$

3.5 Estimación insesgada

Sea X_1, \dots, X_n una muestra aleatoria simple procedente de la familia de distribuciones (3.2.2). La caracterización de las funciones estimables para esta

familia, ver Tate (1959) [49], es:

Una función $h : (0, +\infty) \rightarrow \mathbb{R}$ es una función estimable si y solo si

1. $h(\theta)\theta^n$ es absolutamente continua.

2. $\lim_{\theta \rightarrow 0^+} h(\theta)\theta^n = 0$.

Además, un estimador insesgado de $h(\theta)$ es

$$T_n(S_n) = h(a^{-1}(S_n)) + \frac{S_n}{n} Dh(a^{-1}(S_n)). \quad (3.5.1)$$

Sea $h(\theta)$ una función estimable y $T_n(S_n)$ el estimador insesgado dado en (3.5.1). De acuerdo con la teoría general si $T_n \in L^2_{\theta,n}$ puede expresarse

$$T_n(S_n) = \sum_{k \geq 0} a_{k,n}(\theta) P_{k,n}(S_n, \theta). \quad (3.5.2)$$

Con objeto de calcular los coeficientes de Fourier en (3.5.2), calcularemos $E_{\theta}[T_n(S_n)P_{k,n}(S_n, \theta)]$. De (3.5.1), obtenemos

$$E_{\theta}[T_n(S_n)S_n^j] = \frac{a(\theta)^j}{n+j} \left(nh(\theta) + j(n+j)! \sum_{i \geq 1} (-1)^{i+1} \frac{D_{\theta}^i h(\theta) a(\theta)^i}{(n+j-i)!} \right), \quad (3.5.3)$$

y a partir de este resultado, de la expresión (3.3.1) de los polinomios ortogonales y teniendo en cuenta que

$$\sum_{i=0}^k (-1)^i \binom{k}{i} \frac{\Gamma(n+2k-i)}{\Gamma(n+k-i+1)} = 0,$$

ver Feller (1968) [17], podemos concluir que, para $k \geq 1$,

$$\begin{aligned} E_{\theta}[T_n(S_n)P_{k,n}(S_n, \theta)] &= \\ &= \frac{k\Gamma(n+k)a(\theta)^k}{\Gamma(n+2k)} \sum_{i=0}^{k-1} \left(\binom{k-1}{i} \Gamma(n+2k-i) \sum_{j \geq 1} \frac{D^j h(\theta) a(\theta)^j (-1)^{i+j+1}}{\Gamma(n+k-i+j+1)} \right) \end{aligned} \quad (3.5.4)$$

La expresión $E_\theta[T_n(S_n)P_{k,n}(S_n, \theta)]$ es una combinación lineal de los coeficientes del k -ésimo polinomio ortogonal y de las derivadas de $h(\theta)$. Nótese que en esta combinación lineal no aparece $h(\theta)$, tan solo sus derivadas, de donde se deduce que los coeficientes de Fourier, $a_{k,n}(\theta)$, del desarrollo ortogonal del estimador $T_n(S_n)$ pueden ser obtenidos sin conocer explícitamente dicho estimador.

3.6 Propiedades límites de los estimadores insesgados

A partir de los desarrollos ortogonales estudiamos propiedades asintóticas de los estimadores insesgados .

En primer lugar estudiaremos el comportamiento asintótico de los coeficientes de Fourier de los desarrollos ortogonales.

Teorema 3.6.1 *Sea $h \in \mathcal{U}_{n_0}^2(\Theta_0)$, y $T_n(S_n) = \sum_{k \geq 0} a_{k,n}(\theta)P_{k,n}(S_n, \theta)$, estimador insesgado para una muestra de tamaño $n \geq n_0$. Denotamos por $g(\beta) = h \circ a^{-1}(\beta)$. Entonces*

$$p \lim_n a_{k,n}(\theta) = \frac{D_\beta^k g(\beta)}{k!}, \quad \forall k \geq 0, \forall \theta \in \Theta_0.$$

Demostración.

Para cualquier $\theta_0 \in \Theta$, el estimador insesgado puede escribirse

$$T_n(S_n) = \sum_{k \geq 0} a_{k,n}(\theta_0)P_{k,n}(S_n, \theta_0).$$

Debido a que $\lim_n \text{Var}_\theta(T_n(S_n)) = 0$, la sucesión de estimadores $\{T_n(S_n)\}$ es consistente, ver la demostración del corolario 2.5.2. Esto implica que tomando

límite en probabilidad (bajo $\theta_1 \in \Theta_0$), $p \lim_n T_n(S_n) = h(\theta_1)$, por tanto

$$g(\beta_1) = h(\theta_1) = p \lim_n \sum_{k \geq 0} a_{k,n}(\theta_0) P_{k,n}(S_n, \theta_0).$$

La serie $\sum_{k \geq 0} a_{k,n}(\theta) P_{k,n}(S_n, \theta)$ es absolutamente convergente en sentido cuadrático, entonces es absolutamente convergente en probabilidad (para cualquier $\theta \in \Theta_0$), por tanto podemos intercambiar el sumatorio con el límite y usando (3.4.5)

$$g(\beta_1) = \sum_{k \geq 0} \lim_n a_{k,n}(\theta_0) \lim_n P_{k,n}(S_n, \theta_0) = \sum_{k \geq 0} \lim_n a_{k,n}(\theta_0) (\beta_1 - \beta_0)^k, \quad (3.6.1)$$

donde $\beta_1 = a(\theta_1)$ y $\beta_0 = a(\theta_0)$.

Por otro lado, el desarrollo de Taylor de la función la función $g(\beta)$ en un entorno de β_0 es:

$$g(\beta_1) = \sum_{k \geq 0} \frac{D_{\beta}^k g(\beta_0)}{k!} (\beta_1 - \beta_0)^k, \quad (3.6.2)$$

para β_1 perteneciente a un entorno de β_0 . Igualando los coeficientes en (3.6.1) y (3.6.2) se sigue el resultado.

□

Estudiamos ahora el comportamiento límite de $P_{k,n}(S_n, \theta)$. Se verifica el siguiente teorema:

Teorema 3.6.2 Para todo $\theta \in \Theta_0$,

(i)

$$\frac{n^k}{k! a(\theta)^k} P_{k,n}(S_n, \theta) \xrightarrow{d} L_k^{(0)}(W), \quad W \sim Exp(1). \quad (3.6.3)$$

(ii)

$$\overline{\lim}_n \frac{(-1)^k n^k}{(a(\theta) \log \log n)^k} P_{k,n}(S_n, \theta) = 1, \quad (\text{c.s.}). \quad (3.6.4)$$

Demostración.

(i) A partir de (3.4.1)

$$\left(\frac{n}{a(\theta)}\right)^j (a(\theta) - S_n)^j \xrightarrow{d} W^j, \quad W \sim \text{Exp}(1),$$

y de (3.4.4)

$$\lim_n \frac{n^{k-j}}{k! a(\theta)^{k-j}} \pi_{k,j,n}(\theta) = \text{coef}(j, L_k^{(0)}),$$

se obtiene el resultado esperado.

(ii) La demostración es análoga a la del resultado (2.7.18) del lema 2.7.4 y es consecuencia inmediata de (3.4.4) y de la siguiente igualdad debida a Barndorff-Nielsen (1961) [10]:

$$\overline{\lim}_n \frac{n}{a(\theta) \log \log n} (a(\theta) - S_n) = 1, \quad (\text{c.s.}).$$

□

A partir de aquí podemos establecer la distribución límite del estimador insesgado. Supongamos que X_1, \dots, X_n son v.a. independientes e idénticamente distribuidas con función de densidad $f(x, \theta_0)$ de la forma (3.2.2), donde θ_0 puede ser interpretado como el verdadero, pero desconocido, valor del parámetro. Sea $\beta_0 = a(\theta_0)$, $h(\theta) \in \mathcal{U}_{\theta_0}^2(\Theta_0)$ y $g(\beta) = h(a^{-1}(\beta))$ (la función $g(\beta)$ es la función $h(\theta)$ reparametrizada en términos de β). Los operadores D_θ y D_β están intrínsecamente relacionados mediante

$$D_\theta h(\theta) = D_\beta(g(\beta)) D_\theta(\beta). \quad (3.6.5)$$

Teorema 3.6.3 Sea $h(\theta) \in \mathcal{U}_n^2(\Theta_0)$, $n \geq n_0$, $k_0 = \min\{j > 0 : D_\theta^j g(\beta_0) \neq 0\}$, es decir, k_0 es el orden de la primera derivada de $g(\beta) = h(a^{-1}(\beta))$ no nula en $\beta_0 = a(\theta_0)$. Sea $T_n(S_n)$ el estimador insesgado de $h(\theta_0)$ en Θ_0 , para cada $n \geq n_0$. Entonces,

(i) La sucesión $\{T_n(S_n)\}$ tiene un límite asintótico

$$n^{k_0} \{T_n(S_n) - h(\theta_0)\} \xrightarrow{d} D_\theta^{k_0} h(\theta_0) a^{k_0}(\theta_0) \sum_{j=1}^{k_0} (-1)^{k_0+j} \binom{k_0-1}{j-1} L_j^{(0)}(W), \quad (3.6.6)$$

$n \rightarrow \infty$, con $W \sim \text{Exp}(1)$.

(ii)

$$\overline{\lim}_n \left(\frac{n}{\log \log n} \right)^{k_0} (T_n(S_n) - h(\theta_0)) = \frac{(-1)^{k_0} D_\theta^{k_0} h(\theta_0)}{k_0!} a^{k_0}(\theta_0), \quad (\text{c.s.}). \quad (3.6.7)$$

Demostración.

(i) Consideremos la expresión de $T_n(S_n)$, dada por Tate (1959) [49]

$$\begin{aligned} T_n(S_n) &= h(a^{-1}(S_n)) + \frac{S_n}{n} Dh(a^{-1}(S_n)) \\ &= g(S_n) + \frac{S_n}{n} Dg(S_n). \end{aligned} \quad (3.6.8)$$

Teniendo en cuenta que k_0 es el orden de la primera derivada no nula de $g(\beta)$ en β_0 , en un entorno de β_0 podemos expresar

$$\begin{aligned} g(S_n) &= g(\beta_0) + \frac{D_\theta^{k_0} g(\zeta)}{k_0!} (S_n - \beta_0)^{k_0} \\ &= g(\beta_0) + \frac{D_\theta^{k_0} g(\zeta)}{k_0!} \frac{(-1)^{k_0} \beta_0^{k_0} W_n^{k_0}}{n^{k_0}}, \end{aligned} \quad (3.6.9)$$

$$\begin{aligned}
 Dg(S_n) &= D_\theta g(\beta_0) + \frac{D_\beta^{k_0} g(\zeta)}{(k_0 - 1)!} (S_n - \beta_0)^{k_0 - 1} \\
 &= \frac{D_\theta^{k_0} g(\zeta)}{(k_0 - 1)!} \frac{(-1)^{(k_0 - 1)} \beta_0^{(k_0 - 1)} W_n^{(k_0 - 1)}}{n^{(k_0 - 1)}}, \quad (3.6.10)
 \end{aligned}$$

siendo $W_n = n \left(1 - \frac{S_n}{a(\theta_0)} \right)$.

Sustituyendo (3.6.9) y (3.6.10) en la expresión (3.6.8) obtenemos

$$\begin{aligned}
 n^{k_0} [T_n(S_n) - g(\beta_0)] &= \frac{(-1)^{k_0} D_\theta^{k_0} g(\zeta) \beta_0^{k_0}}{k_0!} W_n^{k_0} + \\
 &+ \frac{(-1)^{(k_0 - 1)} D_\theta^{k_0} g(\zeta) \beta_0^{(k_0 - 1)}}{(k_0 - 1)!} S_n W_n^{(k_0 - 1)},
 \end{aligned}$$

Tomando límite en la expresión anterior, teniendo en cuenta que

$W_n \xrightarrow{d} W \sim \text{Exp}(1)$, $h(\theta_0) = g(\beta_0)$ y $a(\theta_0) = \beta_0$, obtenemos

$$\begin{aligned}
 \lim_n n^{k_0} \{T_n(S_n) - h(\theta_0)\} &= \frac{(-1)^{k_0} D_\theta^{k_0} h(\theta_0) a^{k_0}(\theta_0)}{(k_0 - 1)!} W^{k_0 - 1} \left(\frac{W}{k_0} - 1 \right) \\
 &= D_\theta^{k_0} h(\theta_0) a^{k_0}(\theta_0) \sum_{j=1}^{k_0} (-1)^{k_0 + j} \binom{k_0 - 1}{j - 1} L_j^{(0)}(W).
 \end{aligned}$$

(ii) La demostración de (3.6.7) es similar, a partir del resultado (3.6.4) debido a Barndorff-Nielsen (1961) [10].

□

En particular, si $k_0 = 1$ del teorema anterior obtenemos

$$\frac{n}{D_\theta h(\theta_0) a(\theta_0)} (h(\theta_0) - T_n(S_n)) \xrightarrow{d} W - 1, \quad W \sim \text{Exp}(1). \quad (3.6.11)$$

El estimador de máxima verosimilitud de $g(\beta)$ es $g(S_n)$. Se verifica que

$$n\{g(S_n) - g(\beta_0)\} \xrightarrow{d} \beta_0 g'(\beta_0)W, \quad W \sim Exp(1) \quad (3.6.12)$$

ver Bickel-Doksum (1977) [12]. Comparando (3.6.12) y el resultado (3.6.11), tenemos que la distribución límite del estimador de máxima verosimilitud está desplazada con respecto a la distribución límite del estimador insesgado. Este comportamiento difiere del obtenido en la familia exponencial, pues en esta, las distribuciones límite del estimador insesgado y de el de máxima verosimilitud, coinciden para $k_0 = 1$. En general, si $k_0 \geq 2$ las distribuciones del EMV y el insesgado no coinciden.

3.7 Cotas para la varianza de los estimadores insesgados

Según la teoría general, teorema 1.4.1, para cualquier $V_n = V(X_1, \dots, X_n)$ estimador insesgado de $h(\theta)$,

$$Var_{\theta}(V_n) \geq \sum_{k=1}^m a_{k,n}^2(\theta) \|P_{k,n}\|_{2,\theta}^2 = B_{m,n}(h, \theta) \quad \forall \theta \in \Theta_0.$$

La sucesión de cotas $\{B_{m,n}(h, \theta)\}_{m \geq 1}$ constituye una sucesión no decreciente de cotas inferiores para la varianza de los estimadores insesgados de $h(\theta)$ para un tamaño muestral n . Así, la cota de orden 1 es

$$B_{1,n}(h, \theta) = \frac{n+2}{n} \left(\sum_{j \geq 1} \frac{(-1)^{j+1} a(\theta)^j D^j h(\theta) (n+1)!}{(n+j+1)!} \right)^2.$$

Esta cota es mas fina que la Chapman-Robins-Kieffer, ver Barranco et al. (1999) [11]. La m -ésima cota $B_{m,n}(h, \theta)$ se alcanza si y solo si $T_n(S_n)$

es un polinomio de grado m en S_n y equivalentemente $h(\theta)$ ha de ser un polinomio de grado m en θ . Si $h \in \mathcal{U}_n^2(\Theta_0)$ la serie $\sum_{k=1}^{\infty} a_{k,n}^2(\theta) \|P_{k,n}\|_{2,\theta}^2$ es finita para cualquier $\theta \in \Theta_0$, entonces es evidente que la sucesión de cotas $\{B_{m,n}(h, \theta)\}$ converge a $Var_{\theta} T_n(S_n)$.

Podemos deducir algunas propiedades límites de las cotas $B_{k,n}(h, \theta)$ cuando el tamaño muestral tiende a infinito, como

$$\lim_n n^{2(m+1)} (B_{m+1,n}(h, \theta) - B_{m,n}(h, \theta)) = (D^{m+1}h(\theta))^2 a(\theta)^{2(m+1)}. \quad (3.7.1)$$

3.8 Comparación con el estimador de máxima verosimilitud

Dada una m.a.s. X_1, \dots, X_n de una distribución perteneciente a la familia de distribuciones (3.2.2), sabemos que el estimador de máxima verosimilitud de $a(\theta)$ es $S_n = \max_{1 \leq i \leq n} X_i$. Dada una función $h : \Theta \rightarrow \mathbb{R}$, el estimador de máxima verosimilitud de $h(\theta)$ será $h(a^{-1}(S_n))$ y el estimador insesgado de mínima varianza es, según vimos en (3.5.2),

$$T_n(S_n) = \sum_{k \geq 0} a_{k,n}(\theta) P_{k,n}(S_n, \theta), \quad \forall \theta \in \Theta.$$

La elección de θ en el término de la derecha de la expresión anterior es arbitraria. Si elegimos $\theta = \hat{\theta}_n = a^{-1}(S_n)$, estimador de máxima verosimilitud de θ , como $a_{0,n}(\theta) = h(\theta)$ tenemos

$$T_n(S_n) = h(a^{-1}(S_n)) + \sum_{k \geq 1} a_{k,n}(\hat{\theta}_n) P_{k,n}(S_n, \hat{\theta}_n). \quad (3.8.1)$$

De (3.4.3) tenemos que $P_{k,n}(S_n, \hat{\theta}_n) = \pi_{k,0,n}(\hat{\theta}_n)$ y a partir (3.4.4) y del teorema 3.6.1 tenemos

$$\lim_n \frac{n}{a(\theta)} \pi_{1,0,n}(\theta) = \text{coef}(0, L_k^{(0)}) = 1, \quad \forall \theta \in \Theta,$$

$$\lim_n a_{1,n}(\theta) = D_\beta^g(\beta), \quad \forall \theta \in \Theta,$$

podemos concluir que

$$n\{T_n(S_n) - h(a^{-1}(S_n))\} \xrightarrow{p} \beta D_\beta g(\beta).$$

Un resultado más general se obtiene considerando los estimadores que se obtienen truncando el desarrollo (3.8.1) hasta el término j -ésimo:

$$C_{j,n}(S_n) = \sum_{k=0}^j a_{k,n}(\hat{\theta}_n) P_{k,n}(S_n, \hat{\theta}_n),$$

que pueden ser considerados como correcciones del estimador de máxima verosimilitud. Siguiendo argumentos similares a los anteriores, podemos concluir que

$$n^j \{T_n(S_n) - C_{j-1,n}(S_n)\} \xrightarrow{p} \beta^j D_\beta^j g(\beta).$$

3.9 Generalización

Vamos a estudiar la estimación de funciones paramétricas en familias, $\mathcal{F}(\xi)$, con densidad de la forma

$$f(y, \xi) = \frac{g(y)}{G(\xi)}, \quad a_1 < y < a_2(\xi), \quad \forall \xi \in \Theta. \quad (3.9.1)$$

siendo

$$G(\xi) = \int_{a_1}^{a_2(\xi)} g(y) dy.$$

Este tipo de distribuciones se les denomina truncadas del tipo II, ver Tate (1959) [49]. La distribución (3.2.2), definida en la sección primera de este capítulo, pertenece a esta familia $\mathcal{F}(\xi)$.

Sin pérdida de generalidad, podemos considerar $a_1 = 0$ y $a_2(\xi) = \xi$.

Consideremos la transformación definida por $X = G(Y)$. La variable transformada pertenece a la familia $U(0, \theta)$ con $\theta = G(\xi)$. Si Y_1, \dots, Y_n es una muestra aleatoria de la familia $\mathcal{F}(\xi)$, por ser G no decreciente, obtenemos

$$\begin{aligned} Y_1, \dots, Y_n &\xrightarrow{G} X_1, \dots, X_n. \\ V_n = \max_{1 \leq i \leq n} Y_i &\xrightarrow{G} G(V_n) = S_n = \max_{1 \leq i \leq n} X_i. \end{aligned}$$

El estadístico $V_n = \max_{1 \leq i \leq n} Y_i$ es un estadístico suficiente y completo para ξ , ver Stuart-Ord (1991)[48], siendo su función de densidad

$$f_{V_n}(v) = \frac{ng(v)G^{n-1}(v)}{G^n(\xi)}, \quad 0 < v < \xi. \quad (3.9.2)$$

Si consideramos una función $\tau(\xi)$ estimable existirá $H_n(V_n)$ tal que

$$\int_0^\xi H_n(v) \frac{ng(v)G^{n-1}(v)}{G^n(\xi)} dv = \tau(\xi),$$

haciendo el cambio de variable $G(v) = s$ obtenemos

$$\int_0^\theta H_n(G^{-1}(s)) \frac{ns^{n-1}}{\theta^n} ds = \tau(G^{-1}(\theta)),$$

por lo tanto $\tau \circ G^{-1}$ es una función estimable en la familia $U(0, \theta)$, siendo su estimador $H_n(G^{-1}(s))$. Podemos así identificar los conjuntos de funciones estimables en la distribución $\mathcal{F}(\xi)$ y en la $U(0, \theta)$, trasladando el problema de la estimación insesgada en la familia $\mathcal{F}(\xi)$ a la estimación insesgada en la $U(0, \theta)$ que es un caso particular del tratado en las secciones anteriores de este capítulo.

Si consideramos el espacio de funciones

$$L_{n,\xi}^2 = \left\{ H_n : \int_0^\xi H_n^2(v) f_{V_n}(v, \xi) dv < \infty \right\},$$

mediante un cambio de variables puede probarse que un sistema ortogonal de funciones en $L_{n,\xi}^2$ viene dado por

$$Q_{k,n}(v, \xi) = P_{k,n}(G(v), G(\xi)). \quad (3.9.3)$$

siendo $\{P_{k,n}(s, \theta)\}_{k \geq 0}$ el sistema de polinomios ortogonales en el espacio

$$L_{n,\theta}^2 = \left\{ T_n : \int_0^\theta T_n^2(s) f_n(s, \theta) ds < \infty \right\}.$$

En particular,

$$Q_{0,n}(v) = 1.$$

$$Q_{1,n}(v) = G(v) - \frac{n}{n+1} G(\xi).$$

$$Q_{2,n}(v) = G^2(v) - \frac{2(n+1)}{n+3} G(\xi)G(v) + \frac{n(n+1)}{(n+2)(n+3)} G^2(\xi).$$

La norma de dichas funciones vendrá dada por

$$\|Q_{k,n}\|_{2,\xi}^2 = G^{2k}(\xi) \frac{nk!^2 \Gamma(n+k)^2}{(n+2k)\Gamma^2(n+2k)}.$$

Si $\tau(\xi)$ es una función estimable, siendo $H_n(V_n)$ el estimador insesgado, podemos expresarlo respecto al sistema $\{Q_{k,n}(v, \xi)\}$ como

$$H_n(V_n) = \sum_{k=0}^{\infty} d_{k,n}(\xi) Q_{k,n}(V_n, \xi).$$

Para calcular los coeficientes $d_{k,n}(\xi)$, obsérvese que:

$$H_n(G^{-1}(S_n)) = \sum_{k=0}^{\infty} a_{k,n}(\theta) P_{k,n}(S_n, \theta) = \sum_{k=0}^{\infty} a_{k,n}(G(\xi)) P_{k,n}(G(V_n), G(\xi))$$

y, teniendo en cuenta la unicidad de los desarrollos y (3.9.3) podemos concluir que

$$d_{k,n}(\xi) = a_{k,n}(G(\xi)),$$

donde $a_{k,n}(G(\xi))$ son los coeficientes de Fourier correspondientes al estimador insesgado de $\tau(G^{-1}(\theta))$ en una $U(0, \theta)$.

Dado que hemos reducido el problema de la estimación en una familia del tipo $\mathcal{F}(\xi)$ al caso de la $U(0, \theta)$ cabe esperar que pueden obtenerse resultados similares a los obtenidos en la familia (3.2.2). Así por ejemplo, podemos:

- Dar cotas inferiores para los estimadores de $\tau(\xi)$ con varianza finita.

$$B_{m,n}(\tau, \xi) = \sum_{k=1}^m d_{k,n}^2(\xi) \|Q_{k,n}\|_{2,\xi}^2, \quad \forall \xi.$$

La cota de orden 1 sería

$$B_{1,n}(\tau, \xi) = \frac{n+2}{n} \left(\sum_{j \geq 1} \frac{(-1)^{j+1} G(\xi)^j D^j \tau(\xi) (n+1)!}{(n+j+1)!} \right)^2.$$

La m -ésima cota $B_{m,n}(\tau, \xi)$ se alcanza si y solo si $H_n(V_n)$ es una combinación lineal de las m primeras funciones ortogonales del sistema $\{Q_{k,n}(v, \xi)\}_{k \geq 0}$.

De igual forma, según lo obtenido en (3.7.1), se deducen algunas propiedades límites de las cotas $B_{k,n}(\tau, \xi)$ cuando el tamaño muestral tiende a infinito:

$$\lim_n n^{2(m+1)} (B_{m+1,n}(\tau, \xi) - B_{m,n}(\tau, \xi)) = (D^{m+1} \tau(\xi))^2 G(\xi)^{2(m+1)}$$

- Obtener propiedades asintóticas.

Si $\tau(\xi)$ es una función estimable para muestras de tamaño $n \geq n_0$, y $k_0 = \min\{j > 0 : D_\xi^j \tau(\xi_0) \neq 0\}$, es decir k_0 es el orden de la primera derivada de τ no nula en ξ_0 . Entonces, aplicando el teorema 3.6.3, obtenemos:

(i)

$$n^{k_0} \{H_n(V_n) - \tau(\xi_0)\} \xrightarrow{d} D_\theta^{k_0} \tau(\xi_0) G^{k_0}(\xi_0) \sum_{j=1}^{k_0} (-1)^{k_0+j} \binom{k_0-1}{j-1} L_j^{(0)}(W),$$

$n \rightarrow \infty$, con $W \sim \text{Exp}(1)$, $\theta_0 = G(\xi_0)$.

(ii)

$$\overline{\lim}_n \left(\frac{n}{\log \log n} \right)^{k_0} \{H_n(V_n) - \tau(\xi_0)\} = \frac{(-1)^{k_0} D_\theta^{k_0} \tau(\xi_0)}{k_0!} G^{k_0}(\xi_0). \text{ (c.s.)}$$

UNIVERSIDAD DE SEVILLA
 FACULTAD DE MATEMÁTICAS
 BIBLIOTECA

Capítulo 4

Otras aplicaciones

4.1 Introducción

La teoría general vista en el capítulo primero puede ser aplicada a situaciones distintas de las estudiadas en los capítulos 2 y 3. Por ejemplo:

- Situaciones en que la muestra es censurada por la derecha (censuramiento tipo II).
- Otras situaciones en las que no disponemos de la muestra completa, como el caso en que tan solo se observan los récords superiores.
- Otras familias de distribuciones, que no posean estadístico suficiente y completo como la familia χ^2 no centrada.

En el caso de censuramiento por la derecha y récords estudiamos el problema de la estimación en una subfamilia de la familia χ^2 -transformada.

Cuando el censuramiento se realiza por la derecha la distribución que genera los datos no necesariamente tiene por qué ser continua, así pues veremos

el problema de estimación en la distribución geométrica censurada por la derecha.

Notaremos que para la aplicación de la teoría general no se requiere ni la suficiencia ni la completitud de los estadísticos sobre los que se construyen los estimadores. Ejemplo de ello es el caso de la estimación en la χ^2 no centrada, en el que, aun disponiendo de una muestra no censurada, es imposible obtener un estadístico minimal suficiente, si $n > 1$, sin embargo si completo.

4.2 Muestras censuradas

El término censura es bien conocido en la literatura estadística. Por datos censurados entendemos el caso en que en una muestra de tamaño n un número conocido o no de observaciones (antes del fin o a ambos lados) son no consideradas.

Un ejemplo importante ocurre en los test de vida cuando se decide parar la experimentación tan pronto como $k < n$ items bajo test fallen. Aquí la censura es por la derecha y podemos obtener razonablemente buenos estimadores de parámetros mejor que esperar a que todos los items fallen (la pérdida en eficiencia de los estimadores será menor comparada con la ganancia en ahorro de tiempo). El tipo de censura descrita, en la cual el experimentador decide parar el experimento después de un número preasignado de observaciones fallidas, es llamada tipo II para distinguirla de las situaciones en las que la muestra es cortada antes o después de unos límites predeterminados (tipo I), aquí el número de observaciones censuradas es una v.a.. Notemos que para el

tipo II el número de observaciones censuradas es fijo, mientras que los límites de censura pueden ser considerados aleatorios

Ambas formas de censura son diferentes del truncamiento donde la población antes que la muestra es cortada y el número de observaciones perdidas es desconocido.

Veremos como es posible aplicar los resultados generales del capítulo 1 al caso de muestras censuradas por la derecha, tipo II, procedentes de una familia con función de densidad

$$f(x, \theta) = \exp\{-a(x)b(\theta)\}a'(x)b(\theta)I_D(x), \theta \in \Theta, \quad (4.2.1)$$

que pertenece a la familia χ^2 -transformada, definida en (2.5.12), con $\nu = 2$. Según vimos en el capítulo 2, $a(X) \sim G\left(1, \frac{1}{\lambda}\right)$, siendo $\lambda = b(\theta)^{-1} > 0$.

Supongamos que las d observaciones mayores son censuradas, un procedimiento usual es realizar nuestras inferencias a partir los estadísticos ordenados

$$X_{(1:n)}, \dots, X_{(n-d:n)}.$$

La densidad conjunta de la muestra ordenada es

$$\begin{aligned} f(x_1, \dots, x_{n-d}) &= \frac{n!}{d!} (1 - F(x_{n-d}))^d \prod_{i=1}^{n-d} f(x_i) \\ &= \frac{n!}{d!} \exp\left\{\frac{-n}{\lambda} w_{n,d}\right\} \prod_{i=1}^{n-d} \frac{a'(x_i)}{\lambda}, \quad x_1 \leq \dots \leq x_{n-d}, \end{aligned}$$

siendo

$$W_{n,d} = W(x_{(1)}, \dots, x_{(n-d)}) = \frac{da(X_{(n-d)}) + \sum_{i=1}^{n-d} a(X_{(i)})}{n},$$

la media winsorizada. Dado que $a(x)$ es una función uno a uno creciente, el estadístico $W_{n,d}$ es suficiente y completo para la distribución conjunta de la muestra ordenada

$$X_{(1:n)} \leq \dots \leq X_{(n-d:n)},$$

ver Bain (1978) [6], y utilizando las propiedades usuales de la distribución gamma, se sigue que $W_{n,d} \sim G\left(n-d, \frac{n}{\lambda}\right)$. Teniendo esto en cuenta, si una función paramétrica $h(\lambda)$ admite un estimador insesgado con varianza finita de la forma $T_{n,d}(W_{n,d})$, entonces $T_{n,d}(W_{n,d})$ tiene mínima varianza entre la clase de todos los estimadores insesgados basados en los estadísticos ordenados $X_{(1:n)}, \dots, X_{(n-d:n)}$.

Usando el hecho de que $W_{n,d} \sim G\left(n-d, \frac{n}{\lambda}\right)$ podemos determinar un sistema de polinomios mónicos ortogonales respecto a dicha densidad. Estos son

$$P_{j,n}^{(n-d-1)}(w, \lambda) = \frac{(-1)^j \lambda^j j!}{n^j} L_j^{(n-d-1)}\left(\frac{nw}{\lambda}\right),$$

donde $L_j^{(n-d-1)}$ denota al j -ésimo polinomio de Laguerre generalizado de orden $n-d-1$, ver (2.5.13).

Teniendo en cuenta la teoría desarrollada el estimador insesgado de $h(\lambda)$ basado en $W_{n,d}$ viene dado por la siguiente expresión

$$T_{n,d}(W_{n,d}) = \sum_{j=0}^{\infty} \frac{D_{\lambda}^j h(\lambda) \Gamma(n-d)}{\Gamma(n-d+j) j!} n^j P_{j,n}^{(n-d-1)}(W_{n,d}, \lambda), \quad \forall \lambda > 0, \quad (4.2.2)$$

siendo su varianza

$$Var_{\lambda}(T_{n,d}(W_{n,d})) = \sum_{j=1}^{\infty} \frac{D_{\lambda}^j h(\lambda)^2 \Gamma(n-d)}{\Gamma(n-d+j) j!} \lambda^{2j}. \quad (4.2.3)$$

La expresión (4.2.2) nos permite obtener propiedades de $T_{n,d}(W_{n,d})$:

- Cotas para la varianza de los estimadores insesgados.

Sea $V_{n,d} = V(X_{(1:n)}, \dots, X_{(n-d:n)})$ un estimador insesgado de $h(\lambda)$, obtenido a partir de la muestra censurada. Teniendo en cuenta la expresión (4.2.3) y las propiedades de mínima varianza de $T_{n,d}(W_{n,d})$ obtenemos

$$\text{Var}_\lambda(V_{n,d}) \geq \text{Var}_\lambda(T_{n,d}(W_{n,d})) \geq B_{m,n-d}(h, \lambda), \quad m \geq 1,$$

con

$$B_{m,n-d}(h, \lambda) = \sum_{j=1}^m \frac{D_\lambda^j h(\lambda)^2 \Gamma(n-d)}{\Gamma(n-d+j)j!} \lambda^{2j},$$

por tanto $\{B_{m,n-d}(h, \lambda)\}_{m \geq 1}$ es una sucesión de cotas inferiores.

En particular, para $m = 1$ se obtiene una cota inferior análoga a la de Fréchet-Cramer-Rao:

$$B_{1,n-d}(h, \lambda) = \frac{(D_\lambda h(\lambda))^2 \lambda^2}{n-d}.$$

- Eficiencia relativa.

Estudiamos la eficiencia relativa de los estimadores basados en la media winsorizada con respecto a los estimadores obtenidos usando toda la muestra, es decir basados en $S_n = \sum_{i=1}^n \frac{a(X_i)}{n}$.

El estimador de mínima varianza de $h(\lambda)$ basado en S_n viene dado por

$$T_n(S_n) = \sum_{j=0}^{\infty} \frac{D_\lambda^j h(\lambda) \Gamma(n)}{\Gamma(n+j)j!} n^j P_j^{(n-1)}(S_n, \lambda),$$

ver (2.5.14), con varianza

$$\text{Var}_\lambda(T_n(S_n)) = \sum_{j=1}^{\infty} \frac{D_\lambda^j h(\lambda)^2 \Gamma(n)}{\Gamma(n+j)j!} \lambda^{2j}. \quad (4.2.4)$$

De (4.2.3) y, teniendo en cuenta que $\frac{\Gamma(n-d)\Gamma(n+j)}{\Gamma(n-d+j)\Gamma(n)}$ es creciente para $n > n-d$, tenemos

$$\begin{aligned} \text{Var}_\lambda(T_{n,d}(W_{n,d})) &= \sum_{j=1}^{\infty} \frac{D_\lambda^j h(\lambda)^2 \Gamma(n-d) \lambda^{2j}}{\Gamma(n-d+j) j!} \frac{\Gamma(n) \Gamma(n+j)}{\Gamma(n+j) \Gamma(n)} \\ &\geq \frac{n}{n-d} \text{Var}_\lambda(T_n(S_n)), \end{aligned}$$

por lo que la eficiencia relativa del estimador basado en la media winsorizada respecto al estimador usual (muestra completa) es

$$ef_\lambda(T_{n,d}(W_{n,d}) | T_n(S_n)) = \frac{\text{Var}_\lambda(T_n(S_n))}{\text{Var}_\lambda(T_{n,d}(W_{n,d}))} \leq 1 - d/n,$$

La igualdad se alcanzará si y solo si $D_\lambda^k h(\lambda) = 0$, para $k > 1$, es decir si $h(\lambda)$ es una función lineal de λ .

Si $\frac{d}{n}$, tiende a $\alpha \in (0, 1)$, cuando $n \rightarrow \infty$, de (4.2.3) y (4.2.4), obtenemos

$$\lim_n n \text{Var}_\lambda(T_{n,d}(W_{n,d})) = \frac{(D_\lambda h(\lambda))^2 \lambda^2}{(1-\alpha)},$$

$$\lim_n n \text{Var}_\lambda(T_n(S_n)) = (D_\lambda h(\lambda))^2 \lambda^2,$$

de aquí que la eficiencia relativa asintótica de $T_{n,d}(W_{n,d})$ respecto a $T_n(S_n)$ viene dada por

$$ARE(T_{n,d}(W_{n,d}), T_n(S_n)) = \lim_n \frac{\text{Var}_\lambda(T_n(S_n))}{\text{Var}_\lambda(T_{n,d}(W_{n,d}))} = 1 - \alpha.$$

- Propiedades asintóticas.

Denotaremos por $\mathcal{U}_{n-d}^2(0, \infty)$ a la clase de funciones infinitamente dife-

renciables que satisfacen la condición

$$\sum_{k=1}^{\infty} \frac{D_{\lambda}^k h(\lambda)^2 \Gamma(n-d)}{\Gamma(n-d+k) k!} \lambda^{2k} < \infty.$$

Sea $h \in \mathcal{U}_{n-d}^2(0, \infty)$, $\lambda_0 > 0$ y $k_0 = \min\{j \geq 1 : D_{\lambda}^j h(\lambda_0) \neq 0\}$. Si $\frac{d}{n} \rightarrow \alpha \in (0, 1)$ cuando $n \rightarrow \infty$, entonces

(i) Para cualquier $\gamma < k_0/2$,

$$n^{\gamma} \{T_{n,d}(W_{n,d}) - h(\lambda_0)\} \xrightarrow{P_{\lambda_0}} 0, \quad n \rightarrow \infty.$$

(ii)

$$n^{k_0/2} \frac{k_0! (1-\alpha)^{k_0/2}}{\lambda_0^{k_0} D_{\lambda}^{k_0} h(\lambda_0)} \{T_{n,d}(W_{n,d}) - h(\lambda_0)\} \xrightarrow{d} H_{e_{k_0}}(Z),$$

siendo $Z \sim N(0, 1)$, $n \rightarrow \infty$.

En particular si $k_0 = 1$ obtenemos la normalidad asintótica de $T_{n,d}(W_{n,d})$, explícitamente:

$$n^{1/2} \{T_{n,d}(W_{n,d}) - h(\lambda_0)\} \xrightarrow{d} N\left(0, \frac{\lambda_0^2 (D_{\lambda} h(\lambda_0))^2}{1-\alpha}\right).$$

Si $k_0 = 2$,

$$\frac{2n(1-\alpha)}{\lambda^2 D_{\lambda}^2 h(\lambda_0)} \{T_{n,d}(W_{n,d}) - h(\lambda_0)\} + 1 \xrightarrow{d} \chi_1^2.$$

4.3 Observaciones récords

Aunque pueda pensarse que es un caso particular de los estadísticos de orden, el estudio de récords es independiente del estudio de los extremos muestrales. El interés por los récords va aumentando, así como el número de aplicaciones a diversos campos como son la meteorología, deportes, hidrología, análisis de stock, etc.. Por este motivo incluimos este apartado en nuestro estudio.

Sea $\{X_n\}$ una sucesión de v.a. con función de distribución $F(x)$. Consideramos la siguiente sucesión de v.a. definidas sucesivamente por

$$L(1) = 1,$$

$$L(r) = \min\{j : X_j > X_{L(r-1)}\},$$

a la variable aleatoria $X_{L(r)}$ se le denomina el r -ésimo récord superior de la sucesión $\{X_n\}_{n>0}$.

Si $F(x)$ es absolutamente continua y $f(x)$ es la función de densidad asociada, la densidad conjunta de los r primeros récords viene dada por, ver Shorrock (1972) [46].

$$f(y_1, \dots, y_r) = (1 - F(y_r)) \prod_{j=1}^r \frac{f(y_j)}{1 - F(y_j)}, \quad y_1 < y_2 < \dots < y_r,$$

y la función de densidad del r -ésimo récord por

$$f_{L(r)}(y) = \frac{1}{(r-1)!} \left(\log \frac{1}{1 - F(y)} \right)^{r-1} f(y).$$

Sea $\{X_k\}_{k>0}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con densidad (4.2.1). La densidad conjunta de los r primeros récords viene dada

$$f(y_1, \dots, y_r) = \exp \left\{ \frac{-a(y_r)}{\lambda} + \log \sum_{i=1}^r a(y_i) - r \log \lambda \right\},$$

de donde $a(X_{L(r)})$ es suficiente y completo. Por tanto, si $T_{n,r} \left(\frac{a(X_{L(r)})}{r} \right)$ es una función tal que $E_\lambda \left(T_{n,r}^2 \left(\frac{a(X_{L(r)})}{r} \right) \right) < \infty$, entonces $T_{n,r} \left(\frac{a(X_{L(r)})}{r} \right)$ es el estimador insesgado de menor varianza, basado en la muestra de récords de la función, $h(\lambda) = E_\lambda \left(T_{n,r} \left(\frac{a(X_{L(r)})}{r} \right) \right)$.

La distribución de $a(X_{L(r)})$ la calculamos a partir de:

$$f_{L(r)}(x) = \frac{1}{(r-1)!} \lambda^{-r} \exp \left\{ \frac{-a(x)}{\lambda} \right\} a(x)^{r-1} a'(x),$$

haciendo el cambio de variable $t = a(x)$, así

$$f_{a(X_{L(r)})}(t) = \frac{1}{(r-1)!} \lambda^{-r} \exp \left\{ \frac{-t}{\lambda} \right\} t^{r-1}.$$

Por lo que $\frac{a(X_{L(r)})}{r} \sim G \left(r, \frac{r}{\lambda} \right)$. Podemos determinar un sistema de polinomios mónicos ortogonales respecto a dicha densidad. Dichos polinomios son

$$P_{j,r}^{(r-1)}(w, \lambda) = \frac{(-1)^j \lambda^j j!}{r^j} L_j^{(r-1)} \left(\frac{rw}{\lambda} \right),$$

siendo $L_j^{(r-1)}$ el j -ésimo polinomio de Laguerre generalizado de orden $r-1$, ver (2.5.13).

Siguiendo la teoría general, la expresión ortogonal del estimador insesgado de una función $h(\lambda)$ basado en $\frac{a(X_{L(r)})}{r}$ es

$$T_r \left(\frac{a(X_{L(r)})}{r} \right) = \sum_{j=0}^{\infty} \frac{D_\lambda^j h(\lambda) \Gamma(r)}{j! \Gamma(r+j)} r^j P_j^{(r-1)} \left(\frac{a(X_{L(r)})}{r}, \lambda \right), \quad (4.3.1)$$

siendo su varianza

$$Var_\lambda \left(T_r \left(\frac{a(X_{L(r)})}{r} \right) \right) = \sum_{j=1}^{\infty} \frac{D_\lambda^j h(\lambda)^2 \Gamma(r) \lambda^{2j}}{j! \Gamma(r+j)}. \quad (4.3.2)$$

La expresión (4.3.1) nos permite obtener propiedades de estos estimadores:

- Cotas para la varianza de los estimadores insesgados.

Si $V_r = V(X_{L(1)}, \dots, X_{L(r)})$ un estimador insesgado de $h(\lambda)$ basado en la muestra de observaciones r cords, teniendo en cuenta la expresi n (4.3.2) y las propiedades de m nima varianza de $T_r \left(\frac{a(X_{L(r)})}{r} \right)$ obtenemos que para $m \geq 1$,

$$Var_\lambda(V_r) \geq Var_\lambda \left(T_r \left(\frac{a(X_{L(r)})}{r} \right) \right) \geq \sum_{j=1}^m \frac{D_\lambda^j h(\lambda)^2 \Gamma(r) \lambda^{2j}}{j! \Gamma(r+j)} = B_{m,r}(h, \lambda),$$

por tanto $\{B_{m,r}(h, \lambda)\}_{m \geq 1}$ es una sucesi n de cotas inferiores.

En particular, para $m = 1$ se obtiene una cota inferior an loga a la de Fr chet-Cramer-Rao:

$$B_{1,r}(h, \lambda) = \frac{h'(\lambda)^2 \lambda^2}{r}.$$

- Propiedades asint ticas.

Denotaremos por $\mathcal{U}_r^2(0, \infty)$ a la clase de funciones infinitamente diferenciables que satisfacen la condici n

$$\sum_{k=1}^{\infty} \frac{D_\lambda^k h(\lambda)^2 \Gamma(r)}{\Gamma(r+k) k!} \lambda^{2k} < \infty.$$

Sea $h \in \mathcal{U}_r^2(0, \infty)$, $\lambda_0 > 0$ y $k_0 = \min\{j \geq 1 : D_\lambda^j h(\lambda_0) \neq 0\}$. Se verifica

- (i) Para cualquier $\gamma < k_0/2$,

$$r^\gamma \left\{ T_r \left(\frac{a(X_{L(r)})}{r} \right) - h(\lambda_0) \right\} \xrightarrow{P_{\lambda_0}} 0 \quad r \rightarrow \infty.$$

- (ii)

$$r^{k_0/2} \frac{k_0!}{\lambda_0^{k_0} D_\lambda^{k_0} h(\lambda_0)} \left\{ T_r \left(\frac{a(X_{L(r)})}{r} \right) - h(\lambda_0) \right\} \xrightarrow{d} H_{e_{k_0}}(Z),$$

siendo $Z \sim N(0, 1)$, $r \rightarrow \infty$.

En particular si $k_0 = 1$ obtenemos la normalidad asintótica de $T_{n,r} \left(\frac{a(X_{L(r)})}{r} \right)$, explícitamente:

$$r^{1/2} \left\{ T_r \left(\frac{a(X_{L(r)})}{r} \right) - h(\lambda_0) \right\} \xrightarrow{d} N \left(0, \lambda_0^2 D_\lambda h(\lambda_0)^2 \right).$$

4.4 Distribución Geométrica

En esta sección vamos a tratar el problema de estimación de funciones paramétricas en la distribución geométrica cuando la información nos la proporciona una muestra censurada tipo II.

En primer lugar estudiaremos algunas propiedades de los polinomios de Meixner de primera clase que serán de gran utilidad para la obtención de posteriores resultados:

Sea Y una v.a. con distribución binomial negativa de parámetros $r > 0$ conocido y $p \in (0, 1)$ desconocido, $\mathcal{BN}(r, p)$. La función de probabilidad es de la forma:

$$P_p [Y = y] = \binom{r + y - 1}{y} p^r (1 - p)^y \quad y = 0, 1, \dots$$

Cuando resulte conveniente reparametrizaremos usando $\xi = \frac{q}{p}$, siendo $q = 1 - p$, así la función de probabilidad de dicha variable podemos escribirla

$$P_\xi[Y = y] = \binom{r + y - 1}{y} \xi^y (1 + \xi)^{-(r+y)} \quad y = 0, 1, \dots \quad (4.4.1)$$

Consideremos los polinomios de Meixner de primera clase, $m_k(y, r, q)$, cuya expresión explícita es, ver Chihara (1978) [14],

$$m_k(y, r, q) = (-1)^k k! \sum_{j=0}^k \binom{y}{j} \binom{-y-r}{k-j} q^{-j}.$$

Estos polinomios verifican las siguientes propiedades:

$$(1) \frac{d^k}{d\xi^k} P_\xi[Y = y] = (-1)^k (1 + \xi)^{-k} m_k(y, r, q) P_\xi[Y = y], \quad k = 0, 1, \dots$$

Basta derivar k veces con respecto a ξ , usando la fórmula de Leibnitz para la derivada de un producto, en la expresión (4.4.1).

$$(2) E_\xi [m_k(Y, r, q) m_j(Y, r, q)] = \delta_{kj} q^{-k} k! (r)_k, \text{ con } \delta_{kj} \text{ la delta de Kronecker y } (r)_k = r(r+1) \cdots (r+k-1), \text{ ver Chihara (1978) [14].}$$

(3) La función generatriz de los polinomios de Meixner de primera clase es

$$\Pi_m(w, y, q) = (1 - w/q)^y (1 - w)^{-y-r},$$

ver Chihara (1978) [14].

(4) Si d_n es tal que $d_n/n \rightarrow \alpha \in (0, 1)$, $n \rightarrow \infty$, entonces $\forall k \geq 0$, se verifica

$$\lim_n Q_{k,n}(z, \xi) = He_k(z),$$

siendo $Q_{k,n}(z, \xi) =$

$$= (-1)^k \left(\frac{q}{n - d_n} \right)^{\frac{k}{2}} m_k \left((n - d_n)\xi + z \left((n - d_n)\xi(1 + \xi) \right)^{\frac{1}{2}}, n - d_n, q \right),$$

y $He_k(z)$ el k -ésimo polinomio de Hermite, ver (2.7.1). Este resultado es consecuencia directa del lema 2.7.1.

Sea X_1, \dots, X_n una muestra aleatoria simple procedente de una distribución Geométrica de parámetro $p \in (0, 1)$. El estadístico $n\bar{X}_n$ es suficiente y completo para esta familia y sigue una distribución $\mathcal{BN}(n, p)$. Sea $h(\xi)$ una función analítica que admite un estimador insesgado, de la forma $T(\bar{X}_n)$. Siguiendo la teoría general, y teniendo en cuenta que los polinomios de Meixner son ortogonales con respecto a la densidad (4.4.1), dicho estimador puede ser expresado mediante un desarrollo ortogonal de la forma

$$T(\bar{X}_n) = \sum_{k=0}^{\infty} a_{k,n}(\xi) m_k(n\bar{X}_n, n, q), \quad (4.4.2)$$

siendo los coeficientes

$$a_{k,n}(\xi) = \frac{(-\xi)^k D_{\xi}^k h(\xi)}{k!(n)_k}, \quad k \geq 0. \quad (4.4.3)$$

Sustituyendo (4.4.3) en (4.4.2) obtenemos la siguiente expresión para el estimador insesgado

$$T(\bar{X}_n) = \sum_{k=0}^{\infty} \frac{(-\xi) D_{\xi}^k h(\xi)}{k!(n)_k} m_k(n\bar{X}_n, n, q), \quad (4.4.4)$$

y la de su varianza

$$\text{Var}_{\xi}(T(\bar{X}_n)) = \sum_{k=1}^{\infty} \frac{\xi^k (D_{\xi}^k h(\xi))^2}{k!(n)_k (1 + \xi)^k}. \quad (4.4.5)$$

Debido a que la distribución geométrica pertenece a la familia exponencial, a partir de la expresión (4.4.4) podemos obtener todas las propiedades de los

estimadores insesgados estudiadas en el capítulo segundo.

Para la distribución en estudio tratamos a continuación el problema de la estimación insesgada en el caso de muestras censuradas, tipo II, por la derecha. Análogamente al caso de la familia (4.2.1) supondremos que las d observaciones mayores son censuradas y que trabajamos con los estadísticos ordenados

$$X_{(1:n)}, \dots, X_{(n-d:n)}.$$

Vamos a determinar la distribución de la media winsorizada

$$W_{n,d} = \frac{S_{n,d}}{n} = \frac{\sum_{i=1}^{n-d} X_{(i:n)} + dX_{(n-d:n)}}{n}.$$

Teorema 4.4.1 *Se verifica que $nW_{n,d} \sim \mathcal{BN}(n-d, p)$ $d > 1$.*

Demostración.

Primeramente vamos a determinar la distribución de los $n-d$ primeros estadísticos ordenados. Para $x_1 \leq \dots \leq x_{n-d}$, enteros no negativos, consideramos el suceso

$$\begin{aligned} A_\sigma &= A_\sigma(x_1, \dots, x_{n-d}) \\ &= \{X_{\sigma(1)} = x_1, \dots, X_{\sigma(n-d)} = x_{n-d}, X_{\sigma(j)} \geq x_{n-d}, j = n-d+1, \dots, n\}, \end{aligned}$$

donde σ es un elemento de \mathcal{P}_n , conjunto de las permutaciones de $\{1, \dots, n\}$.

Definimos en \mathcal{P}_n la relación de equivalencia $\sigma_1 \mathcal{R} \sigma_2$ si y solo si $A_{\sigma_1} = A_{\sigma_2}$. Sea

$\mathcal{Q} = \mathcal{Q}(x_1, \dots, x_{n-d})$ un subconjunto de \mathcal{P}_n formado por un único elemento de cada clase de equivalencia y $C(x_1, \dots, x_{n-d})$ el cardinal de \mathcal{Q} . Entonces

$$\{X_{\sigma(1)} = x_1, \dots, X_{\sigma(n-d)} = x_{n-d}\} = \bigcup_{\sigma \in \mathcal{Q}} A_\sigma, \quad (4.4.6)$$

estando la unión formado por conjuntos disjuntos. Teniendo en cuenta que

$$P_p(A_\sigma) = (1-p)^s p^{n-d}, \quad \text{con } s = \sum_{i=1}^{n-d} x_i + dx_{n-d},$$

a partir de (4.4.6) obtenemos

$$P_p [X_{\sigma(1)} = x_1, \dots, X_{\sigma(n-d)} = x_{n-d}] = C(x_1, \dots, x_{n-d})(1-p)^s p^{n-d}. \quad (4.4.7)$$

De donde se deduce que la función de probabilidad del estadístico

$$S_{n,d} = \sum_{i=1}^{n-d} X_{(i:n)} + dX_{(n-d:n)},$$

es de la forma

$$P_p [S_{n,d} = s] = C(s)(1-p)^s p^{n-d}, \quad s = 0, 1, \dots$$

e imponiendo la condición de ser función de probabilidad, obtenemos que

$$C(s) = \binom{n-d+s-1}{s}, \quad s = 0, 1, \dots,$$

por tanto, $S_{n,d} = nW_{n,d}$ sigue una distribución binomial negativa de parámetros $n-d$ y p .

□

En particular, si $d = 0$ obtenemos $\sum_{i=1}^n X_i \sim \mathcal{BN}(n, p)$.

A partir de la expresión (4.4.7) observamos que $S_{n,d}$ es un estadístico suficiente y completo para la muestra censurada.

Teniendo en cuenta estos resultados, si una función paramétrica $h(\xi)$ admite un estimador insesgado con varianza finita de la forma $T_{n,d}(W_{n,d})$, entonces $T_{n,d}(W_{n,d})$ tiene mínima varianza entre los estimadores insesgados basados en los estadísticos ordenados $X_{(1:n)}, \dots, X_{(n-d:n)}$. Usando el hecho de

que $nW_{n,d} \sim \mathcal{BN}(n-d, p)$, $d > 1$, obtenemos la expresión mediante desarrollo ortogonal de $T_{n,d}(W_{n,d})$, estimador insesgado de $h(\xi)$,

$$T_{n,d}(W_{n,d}) = \sum_{k=0}^{\infty} \frac{(-\xi)^k D_{\xi}^k h(\xi)}{k!(n-d)_k} m_k(nW_{n,d}, n-d, q), \quad \forall \xi > 0, \quad (4.4.8)$$

y la de su varianza,

$$\text{Var}_{\xi}(T_{n,d}(W_{n,d})) = \sum_{k=1}^{\infty} \frac{\xi^k (D_{\xi}^k h(\xi))^2}{k!(n-d)_k (1+\xi)^k}. \quad (4.4.9)$$

La expresión (4.4.8) nos permite establecer propiedades de $T_{n,d}(W_{n,d})$:

- Cotas para la varianza de estimadores insesgados.

Sea $V_{n,d} = V(X_{(1:n)}, \dots, X_{(n-d:n)})$ un estimador insesgado de $h(\xi)$. Por las propiedades de mínima varianza de $T_{n,d}(W_{n,d})$ y la expresión (4.4.9) obtenemos

$$\text{Var}_{\xi}(V) \geq \text{Var}_{\xi}(T_{n,d}(W_{n,d})) \geq B_{m,n-d}(h, \xi), \quad m \geq 1,$$

con

$$B_{m,n-d}(h, \xi) = \sum_{k=1}^m \frac{\xi^k (D_{\xi}^k h(\xi))^2}{k!(n-d)_k (1+\xi)^k},$$

siendo $\{B_{m,n-d}(h, \xi)\}_{m \geq 1}$ una sucesión de cotas inferiores similares a las de Bhattacharyya, ver Khan (1984)[25]. Notemos que la cota $B_{m,n-d}(h, \xi)$ es alcanzada por el estimador insesgado de $h(\xi)$ si y solo si $D_{\xi}^k h(\xi) = 0$, para todo $k > m$, es decir, si solo si $h(\xi)$ es un polinomio de grado k .

En particular, para $m = 1$ se obtiene una cota inferior análoga a la de Fréchet-Cramer-Rao

$$B_{1,n-d}(h, \xi) = \frac{(1+\xi)\xi(D_{\xi}h(\xi))^2}{(n-d)}.$$

- Eficiencia relativa del estimador basado en la media winsorizada respecto al estimador basado en la media muestral.

De las expresiones (4.4.5) y (4.4.9), y teniendo en cuenta que $\left\{ \frac{(n)_k}{(n-d)_k} \right\}_k$ es creciente para $d > 1$, podemos escribir

$$Var_{\xi}(T_{n,d}(W_{n,d})) = \sum_{k=1}^{\infty} \frac{\xi^k (D_{\xi}^k h(\xi))^2}{k!(n)_k(1+\xi)^k} \frac{(n)_k}{(n-d)_k} \geq \frac{n}{n-d} Var_{\xi}(T_n(\bar{X}_n)),$$

de donde

$$ef_{\xi}(T_{n,d}(W_{n,d})|T_n(\bar{X}_n)) = \frac{Var_{\xi}(T_n(\bar{X}_n))}{Var_{\xi}(T_{n,d}(W_{n,d}))} \leq 1 - d/n. \quad (4.4.10)$$

La igualdad (4.4.10) se obtendrá si y solo si $D_{\xi}^k h(\xi) = 0$, para todo $k > 1$, es decir si $h(\xi)$ es una función lineal de ξ .

Si $\frac{d}{n}$, tiende a $\alpha \in (0, 1)$, cuando $n \rightarrow \infty$, de (4.4.9) y (4.4.5), obtenemos

$$\lim_n n Var_{\xi}(T_{n,d}(W_{n,d})) = \frac{\xi(1+\xi)(h'(\xi))^2}{(1-\alpha)},$$

$$\lim_n n Var_{\xi}(T_n(\bar{X}_n)) = \xi(1+\xi)(h'(\xi))^2$$

de aquí que la eficiencia relativa asintótica de $T_{n,d}(W_{n,d})$ respecto a $T_n(\bar{X}_n)$ viene dada por

$$ARE(T_{n,d}(W_{n,d}), T_n(\bar{X}_n)) = \lim_n \frac{Var_{\xi}(T_n(\bar{X}_n))}{Var_{\xi}(T_{n,d}(W_{n,d}))} = 1 - \alpha.$$

- Propiedades asintóticas.

Denotamos por $\mathcal{U}_{n-d}^2(0, \infty)$ a la clase de funciones infinitamente diferenciables en $(0, \infty)$ que satisfacen la condición

$$\sum_{k=1}^{\infty} \frac{\xi^{2k} (D_{\xi}^k h(\xi))^2}{k!(n-d)_k \eta^k} < \infty, \quad \xi > 0.$$

Sea $h \in \mathcal{U}_{n-d}^2(0, \infty)$. Denotamos por

$$R_j(W_{n,d}, \xi) = \sum_{k=j}^{\infty} \frac{(-\xi)^k D_{\xi}^k h(\xi)}{k!(n-d)_k} m_k(nW_{n,d}, n-d, q) \quad j \geq 1, \text{ fijo.}$$

Se verifica que para cualquier $\xi > 0$ y $\gamma < \frac{j}{2}$, si $\frac{d}{n}$ tiende a $\alpha \in (0, 1)$ cuando $n \rightarrow \infty$ entonces

$$n^{\gamma} R_j(W_{n,d}; \xi) \xrightarrow{P_{\xi}} 0, \quad n \rightarrow \infty. \quad (4.4.11)$$

Este resultado nos permite estudiar el comportamiento asintótico de los estimadores insesgados basados en la media winsorizada:

Sea $h \in \mathcal{U}_{n-d}^2(0, \infty)$, y $k_0 = \min\{j > 0 : D_{\xi}^j h(\xi_0) \neq 0\}$. Entonces, si $\frac{d}{n} \rightarrow \alpha \in (0, 1)$, cuando $n \rightarrow \infty$, se verifica

(i) Para cualquier $\gamma < \frac{k_0}{2}$,

$$n^{\gamma} \{T_{n,d}(W_{n,d}) - h(\xi_0)\} \xrightarrow{P_{\xi_0}} 0, \quad n \rightarrow \infty.$$

(ii)

$$\frac{n^{k_0/2} k_0!}{D_{\xi}^{k_0} h(\xi_0)} \left(\frac{1-\alpha}{\xi(1+\xi)} \right)^{k_0/2} \{T_{n,d}(W_{n,d}) - h(\xi_0)\} \xrightarrow{d} He_{k_0}(Z),$$

$n \rightarrow \infty$, con $Z \sim \mathcal{N}(0, 1)$.

En particular, si $k_0 = 1$, $He_1(z) = z$, obtenemos la normalidad asintótica del estimador $T_{n,d}(W_{n,d})$, explícitamente,

$$n^{1/2} \{T_{n,d}(W_{n,d}) - h(\xi_0)\} \xrightarrow{d} \mathcal{N} \left(0, \frac{\xi_0(1+\xi_0)}{1-\alpha} (D_{\xi} h(\xi_0))^2 \right).$$

Si $k_0 = 2$, entonces $He_2(z) = z^2 - 1$, con lo cual

$$\frac{2n(1-\alpha)}{D_{\xi}^2 h(\xi_0) \xi_0(1+\xi_0)} \{T_{n,d}(W_{n,d}) - h(\xi_0)\} + 1 \xrightarrow{d} \chi_1^2.$$

4.5 Distribución χ^2 no centrada

Vamos a aplicar la teoría general, vista en el capítulo primero, al estudio de estimadores insesgados de una función analítica del parámetro de descentralización de una distribución chi-cuadrado no centrada.

Sea $U = (U_1, \dots, U_\nu)$ una variable aleatoria ν dimensional que sigue una distribución normal multivariante con vector de medias $(\delta_1, \dots, \delta_\nu)$ y matriz de covarianzas la identidad I_ν . Entonces la variable $\sum_{j=1}^{\nu} U_j^2$ se distribuye según una chi-cuadrado no centrada con ν grados de libertad y parámetro de descentralización $\lambda = \sum_{j=1}^{\nu} \delta_j^2$.

Existen diferentes formas de presentar la función de densidad, $f(x, \lambda, \nu)$, de la distribución chi-cuadrado no centrada, ver Johnson y Kotz (1970) [24]. Por ejemplo, puede ser expresada como mixtura de distribuciones chi-cuadrado centradas

$$f(x, \lambda, \nu) = \exp\left\{-\frac{\lambda}{2}\right\} \sum_{j=0}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^j}{j!} f(x, 0, \nu + 2j), \quad x \geq 0, \quad (4.5.1)$$

donde $f(x, 0, \nu + 2j)$ denota la función de densidad de una distribución chi-cuadrado con $\nu + 2j$ grados de libertad. Tiku (1965) [50] obtuvo la función de densidad en términos de los polinomios de Laguerre generalizados

$$f(x, \lambda, \nu) = f(x, 0, \nu) \sum_{j=0}^{\infty} \frac{\left(-\frac{\lambda}{2}\right)^j \Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2} + j\right)} L_j^{(\frac{\nu}{2}-1)}\left(\frac{x}{2}\right), \quad x \geq 0, \quad (4.5.2)$$

donde $L_j^{(\alpha)}(y)$ es el j -ésimo polinomio de Laguerre generalizado, definido en (2.5.13).

La distribución chi-cuadrado no centrada es reproductiva bajo convolución. Los grados de libertad y los parámetros de descentralización son aditivos bajo

convolución.

La estimación del parámetro de descentralización, λ , nos proporciona información útil sobre la potencia de diferentes test, por lo que hay diversos estudios sobre este tema. Por ejemplo, Meyer (1967) [34], obtuvo el estimador de máxima verosimilitud para λ con $\nu = 2$. Un estudio más general sobre el estimador de máxima verosimilitud y sus propiedades ha sido hecho por Saxena y Alam (1982) [41].

Por otra parte, el estudio de estimadores insesgados no es siempre satisfactorio, por ejemplo, para muestras de tamaño uno, el estimador insesgado de mínima varianza del parámetro de descentralización es inadmisibles, puesto que es posible encontrar otro estimador cuyo error cuadrático medio es menor, ver Chow (1987) [15]. En este sentido, justificamos el estudio de la estimación insesgada porque puede ser vista como punto de partida para la obtención de mejores estimadores, ver Perlman y Rasmussen (1975) [38], Neff y Strawderman (1976) [37] y Shao y Strawderman (1995) [44].

4.6 Estimación insesgada

Sea X_1, \dots, X_n una m.a.s. de una población con función de densidad (4.5.1), $\nu > 0$, conocido y $\lambda > 0$, desconocido. Basaremos nuestras inferencias en el estadístico $\bar{X}_n = \sum_{i=1}^n \frac{X_i}{n}$, cuya función de densidad viene dada por

$$g_n(y, \lambda, \nu) = w_n(y, \nu) \sum_{j=0}^{\infty} \frac{\left(-\frac{n\lambda}{2}\right)^j \Gamma\left(\frac{n\nu}{2}\right)}{\Gamma\left(\frac{n\nu}{2} + j\right)} L_j^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right), \quad y \geq 0, \quad (4.6.1)$$

donde

$$w_n(y, \nu) = \frac{\left(\frac{n}{2}\right)^{\frac{n\nu}{2}} \exp\left\{-\frac{ny}{2}\right\} y^{\frac{n\nu}{2}-1}}{\Gamma\left(\frac{n\nu}{2}\right)}, \quad y \geq 0,$$

es la función de densidad de una distribución gamma de parámetros $\frac{n\nu}{2}$ y $\frac{n}{2}$.

Consideremos los polinomios,

$$Q_{j,n}(y, \lambda, \nu) = \frac{j!(-2)^j}{n^j} \sum_{k=0}^j \frac{\left(\frac{n\lambda}{2}\right)^{j-k}}{(j-k)!} L_k^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right), \quad j = 0, 1, \dots \quad (4.6.2)$$

La expresión explícita de los cuatro primeros polinomios es

$$Q_{0,n}(y, \lambda, \nu) = 1,$$

$$Q_{1,n}(y, \lambda, \nu) = \left(\frac{V}{n}\right)^{\frac{1}{2}} z,$$

$$Q_{2,n}(y, \lambda, \nu) = \frac{V}{n} \left\{ z^2 - \frac{4}{(nV)^{\frac{1}{2}}} z - 1 \right\},$$

$$Q_{3,n}(y, \lambda, \nu) = \left(\frac{V}{n}\right)^{\frac{3}{2}} \left\{ z^3 - \frac{12}{(nV)^{\frac{1}{2}}} z^2 - 3\left(1 - \frac{8}{nV}\right) z + \frac{8(2\nu\lambda + 3\lambda)}{(nV^3)^{\frac{1}{2}}} \right\},$$

donde

$$z = \left(\frac{n}{V}\right)^{\frac{1}{2}} (y - (\lambda + \nu)) \quad y \quad V = Var_{\lambda}(X_1) = 2(2\lambda + \nu).$$

Los polinomios de grados superiores pueden ser obtenidos usando (4.6.2) y la relación de recurrencia de los polinomios de Laguerre generalizados :

$$\begin{aligned} 2kL_k^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right) &= (4(k-1) - n(y-\nu))L_{k-1}^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right) \\ &\quad - (2(k-2) + n\nu)L_{k-2}^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right), \end{aligned}$$

para $k \geq 1$, con $L_{-1}^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right) = 0$, $L_0^{(\frac{n\nu}{2}-1)}\left(\frac{ny}{2}\right) = 1$, ver (3.4.2) .

Los polinomios $\{Q_{j,n}(y, \lambda, \nu)\}_{j \geq 0}$, verifican las siguientes propiedades:

(1) $E_\lambda [Q_{j,n}(\bar{X}_n, \lambda_0, \nu)] = (\lambda - \lambda_0)^j, j \geq 0.$

(2) Para cualquier w e y ,

$$\sum_{j=0}^{\infty} \frac{(-1)^j Q_{j,n}(y, \lambda, \nu)}{j! 2^j} w^j = \left(1 - \frac{w}{n}\right)^{-\frac{n\nu}{2}} \exp\left\{-\left(\frac{ny}{n-w} - \lambda\right) \frac{w}{2}\right\}. \tag{4.6.3}$$

(3) $\lim_{n \rightarrow \infty} \left(\frac{n}{V}\right)^{\frac{j}{2}} Q_{j,n}(y_n, \lambda, \nu) = He_j(z)$, con $y_n = \lambda + \nu + z \left(\frac{V}{n}\right)^{\frac{1}{2}}$, z arbitrario, y $He_j(z)$ el j -ésimo polinomio de Hermite, definido en (2.7.1).

(4) $\lim_{n \rightarrow \infty} Q_{j,n}(y_n, \lambda, \nu) = (y - (\lambda + \nu))^j, j \geq 0.$

(5) Si denotamos por $y_n = \lambda + \nu + o(n^{-\alpha})$, para cierto $\alpha > 0$. Entonces ,

$$\sum_{j=k+1}^{\infty} \frac{(-1)^j Q_{j,n}(y_n, \lambda, \nu)}{j! 2^j} w^j = o(n^{-(k+1)\alpha}).$$

Demostración.

(1) Teniendo en cuenta (4.6.1) y dado que los polinomios de Laguerre generalizados verifican la relación de ortogonalidad

$$\int_0^\infty L_k^{(\frac{n\nu}{2}-1)} \left(\frac{ny}{2}\right) L_j^{(\frac{n\nu}{2}-1)} \left(\frac{ny}{2}\right) w_n(y, \nu) dy = \delta_{kj} \frac{\Gamma(k + \frac{n\nu}{2})}{\Gamma(\frac{n\nu}{2}) k!}. \tag{4.6.4}$$

donde, $w_n(y, \nu)$ es la función de densidad de una distribución gamma con parámetros $\frac{n\nu}{2}, \frac{n}{2}$, y δ_{kj} la delta de Kronecker, se demuestra que

$$E_\lambda \left[L_k^{(\frac{n\nu}{2}-1)} \left(\frac{n\bar{X}_n}{2}\right) \right] = \frac{\left(-\frac{n\lambda}{2}\right)^k}{k!}.$$

A partir de aquí y de la expresión de los polinomios (4.6.2) se tiene el resultado.

(2) De (4.6.2), tenemos

$$\frac{(-1)^j n^j}{j! 2^j} Q_{j,n}(y, \lambda, \nu) = \sum_{k=0}^j \frac{\left(\frac{n\lambda}{2}\right)^{j-k}}{(j-k)!} L_k^{\left(\frac{n\nu}{2}-1\right)}\left(\frac{ny}{2}\right), \quad j = 0, 1, \dots \quad (4.6.5)$$

La función generatriz de los polinomios de Laguerre generalizados, para cualquier x y t , es

$$\sum_{k=0}^{\infty} L_k^{(\alpha)}(x) t^k = (1-t)^{-(\alpha+1)} \exp\left(\frac{-xt}{1-t}\right), \quad (4.6.6)$$

A partir de (4.6.5) y de (4.6.6) obtenemos

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{(-1)^j n^j Q_{j,n}(y, \lambda, \nu)}{j! 2^j} t^j &= \sum_{j=0}^{\infty} \frac{\left(\frac{n\lambda}{2}\right)^j}{j!} t^j \sum_{j=0}^{\infty} L_j^{\left(\frac{n\nu}{2}-1\right)}\left(\frac{ny}{2}\right) t^j \\ &= \exp\left\{\frac{n\lambda}{2}\right\} (1-t)^{-\frac{n\nu}{2}} \exp\left\{-\frac{nyt}{2(1-t)}\right\}, \end{aligned}$$

de donde (2) resulta inmediato eligiendo $t = \frac{w}{n}$.

(3) Denotamos por $\Pi_{Q_n^*}(t, y, \lambda, \nu)$ a la función generatriz de los polinomios

$Q_{j,n}^* = \frac{(-1)^j Q_{j,n}(y, \lambda, \nu)}{j! 2^j}$, con $Q_{j,n}(y, \lambda, \nu)$ definidos en (4.6.2). Para

$z > 0$, sea $y_n = \lambda + \nu + z \left(\frac{V(\lambda, \nu)}{n}\right)^{\frac{1}{2}}$ y $t_n = -2t \left(\frac{n}{V(\lambda, \nu)}\right)^{\frac{1}{2}}$.

Entonces,

$$\lim_{n \rightarrow \infty} \Pi_{Q_n^*}(t_n, y_n, \lambda, \nu) = \exp\left\{-\frac{t^2}{2} + tz\right\} = \sum_{k=0}^{\infty} He_k(z) \frac{t^k}{k!}, \quad (4.6.7)$$

teniendo en cuenta (4.6.2) se obtiene el resultado.

(4) Se sigue fácilmente del hecho

$$\lim_{n \rightarrow \infty} \Pi_{Q_n}(t, y, \lambda, \nu) = \exp\left\{-\frac{(y - (\lambda + \nu))t}{2}\right\}.$$

(5) Puede demostrarse que

$$\lim_{n \rightarrow \infty} \sum_{j=k+1}^{\infty} \frac{(-1)^j Q_{j,n}(y, \lambda, \nu)}{j! 2^j} t^j = \sum_{j=k+1}^{\infty} \frac{(-1)^j (y - (\lambda + \nu))^j}{j! 2^j} t^j. \quad (4.6.8)$$

Notemos que la parte derecha de (4.6.8) es un resto del desarrollo de Taylor de $\exp\left\{-\frac{(y - (\lambda + \nu))t}{2}\right\}$, entonces el resultado se tiene usando la expresión de Lagrange para el resto de una serie de Taylor.

□

Si $n = 1$, es evidente que \bar{X}_n es un estadístico suficiente para λ . Para $n > 1$, teniendo en cuenta (4.5.1), no es posible factorizar la función de densidad de las n observaciones como nos dice el criterio de factorización de Fisher-Neyman, y consecuentemente \bar{X}_n no es suficiente para λ .

Por otra parte, se verifica:

Teorema 4.6.1 *El estadístico \bar{X}_n es completo para todo $n \geq 1$.*

Demostración.

Para probar que \bar{X}_n es completo, tenemos que demostrar que si T es una función integrable con respecto a la densidad $g_n(y, \lambda, \nu)$, $\forall \lambda > 0$ tal que

$$\int_0^{\infty} T(y) g_n(y, \lambda, \nu) dy = 0, \quad \forall \lambda > 0, \quad \text{entonces } T(y) = 0 \text{ (c.s.)}. \quad (4.6.9)$$

Como los polinomios son densos en el espacio de funciones integrables, es suficiente demostrar (4.6.9) cuando T sea un polinomio.

Sea $\lambda_0 > 0$, un número real arbitrario. Si T es un polinomio de grado k , podemos escribirlo de la forma

$$T(y) = \sum_{j=0}^k a_j(\lambda_0) Q_{j,n}(y_n, \lambda, \nu).$$

Usando la propiedad (1), tenemos

$$\int_0^{\infty} T(y)g_n(y, \lambda, \nu)dy = \sum_{j=0}^k a_j(\lambda_0)(\lambda - \lambda_0)^j = 0 \quad \forall \lambda > 0,$$

esto implica necesariamente que $a_j(\lambda_0) = 0$, para $j = 0, \dots, k$ y de aquí que $T(y) = 0$.

□

La completitud del estadístico \bar{X}_n implica la unicidad de los estimadores insesgados basados en la media aritmética.

Supongamos que $h(\lambda)$ es una función analítica, tal que para todo λ_0 el desarrollo de Taylor

$$h(\lambda) = \sum_{j=0}^{\infty} \frac{D_{\lambda}^j h(\lambda_0)}{j!} (\lambda - \lambda_0)^j,$$

es convergente para cualquier λ . Por la propiedad (1), es inmediato que el estimador insesgado de $h(\lambda)$ es

$$T_n(\bar{X}_n) = \sum_{j=0}^{\infty} \frac{D_{\lambda}^j h(\lambda_0)}{j!} Q_{j,n}(\bar{X}_n, \lambda_0, \nu). \quad (4.6.10)$$

Nótese que diferentes elecciones de λ_0 en el término derecho de (4.6.10) nos darán diferentes expresiones, pero la unicidad del estimador insesgado significa que todas estas expresiones, aparentemente, tienen el mismo límite (c.s.). Esto quiere decir que hay un único estimador insesgado basado en la media aritmética con un número infinito de diferentes expresiones, una para cada elección de λ_0 . En particular, eligiendo $\lambda_0 = 0$, se obtiene

$$T_n(\bar{X}_n) = \sum_{j=0}^{\infty} \frac{(-2)^j D_{\lambda}^j h(0)}{n^j} L_j^{(\frac{n\nu}{2}-1)}(n\bar{X}_n/2). \quad (4.6.11)$$

En la práctica recomendamos la elección de $\lambda_0 = \hat{\lambda} = \bar{X}_n - \nu$, para la cual se tiene

$$T_n(\bar{X}_n) = \sum_{j=0}^{\infty} D_{\hat{\lambda}}^j h(\hat{\lambda}) \gamma_j(\hat{\lambda}, \nu), \quad (4.6.12)$$

donde $\gamma_j(\hat{\lambda}, \nu) = \frac{Q_{j,n}(\bar{X}_n, \hat{\lambda}, \nu)}{j!}$. La expresión explícita de γ_j , $j = 0 \dots 6$ es

$$\gamma_0(\hat{\lambda}, \nu) = 1, \quad \gamma_1(\hat{\lambda}, \nu) = 0,$$

$$\gamma_2(\hat{\lambda}, \nu) = -\frac{2}{n} \{ \hat{\lambda} + \nu \},$$

$$\gamma_3(\hat{\lambda}, \nu) = \frac{4}{n^2} \left\{ \hat{\lambda} + \frac{2\nu}{3} \right\},$$

$$\gamma_4(\hat{\lambda}, \nu) = \frac{2}{n^2} \left\{ \hat{\lambda}^2 + \left(\nu - \frac{4}{n} \right) \hat{\lambda} + \nu \left(\frac{\nu}{4} - \frac{3}{n} \right) \right\},$$

$$\gamma_5(\hat{\lambda}, \nu) = -\frac{8}{n^3} \left\{ \hat{\lambda}^2 - \left(\frac{2}{n} - \frac{7\nu}{6} \right) \hat{\lambda} + \left(\frac{8}{5n} - \frac{\nu}{3} \right) \nu \right\},$$

$$\begin{aligned} \gamma_6(\hat{\lambda}, \nu) = & -\frac{4}{n^4} \left\{ \hat{\lambda}^3 + \left(\frac{3\nu}{2} - \frac{18}{n} \right) \hat{\lambda}^2 + \left(\frac{3\nu^2}{4} - \frac{23\nu}{n} + \frac{24}{n^2} \right) \hat{\lambda} \right\} \\ & - \frac{4}{n^4} \left\{ \left(\frac{20}{n^2} - \frac{43\nu}{6n} + \frac{\nu^2}{8} \right) \nu \right\}. \end{aligned}$$

Aproximaciones para el estimador insesgado pueden obtenerse truncando (4.6.12). En general, no es fácil calcular la varianza de $T_n(\bar{X}_n)$, pero un estimador insesgado de dicha varianza es $V_n(\bar{X}_n) = T_n^2(\bar{X}_n) - M_n(\bar{X}_n)$, donde $M_n(\bar{X}_n)$ es el estimador insesgado de $h^2(\lambda)$. Para obtener $M_n(\bar{X}_n)$, necesita-

mos las derivadas de $h^2(\lambda)$, que podemos calcularlas usando las relaciones

$$D_\lambda^k h^2(\lambda) = 2 \sum_{m=0}^{r-1} \binom{k}{m} D_\lambda^m h(\lambda) D_\lambda^{k-m} h(\lambda) + \binom{k}{r} (D_\lambda^r h(\lambda))^2, \quad \text{si } k = 2r,$$

$$D_\lambda^k h^2(\lambda) = 2 \sum_{m=0}^r \binom{k}{m} D_\lambda^m h(\lambda) D_\lambda^{k-m} h(\lambda) \quad \text{si } k = 2r + 1.$$

Una expresión aproximada para $V_n(\bar{X}_n)$, es

$$\begin{aligned} V_n &= \frac{2}{n} (D_\lambda h(\lambda))^2 (2\hat{\lambda} + \nu) - \frac{2\hat{\lambda}^2}{n^2} (8D_\lambda h(\lambda) D_\lambda^3 h(\lambda) + 4(D_\lambda^2 h(\lambda))^2) \\ &\quad - \frac{2\hat{\lambda}}{n^2} (12D_\lambda h(\lambda) D_\lambda^2 h(\lambda) + 8\nu D_\lambda h(\lambda) D_\lambda^3 h(\lambda) + 4\nu (D_\lambda^2 h(\lambda))^2) \\ &\quad - \frac{2\nu}{n^2} (8D_\lambda h(\lambda) D_\lambda^2 h(\lambda) + 2D_\lambda h(\lambda) D_\lambda^3 h(\lambda) + \nu (D_\lambda^2 h(\lambda))^2) + o(n^{-2}). \end{aligned}$$

Aunque la expresión explícita para el estimador insesgado (4.6.10), (4.6.11) y (4.6.12) y la del estimador de la varianza del mismo parecen complicadas, pueden ser fácilmente implementadas en un ordenador.

Ejemplo 4.6.1 Sea $h(\lambda) = a_2 \lambda^2 + a_1 \lambda + a_0$. Teniendo en cuenta (4.6.12),

$$T_n(\bar{X}_n) = a_2 \hat{\lambda}^2 + \left(a_1 - 2\frac{a_2}{n}\right) \hat{\lambda} + \left(a_0 - \frac{2\nu a_2}{n} + \frac{4a_2}{n^2}\right),$$

con $\hat{\lambda} = (\bar{X}_n - \nu)$. En particular, si $h(\lambda) = \lambda$, es decir, $a_0 = a_2 = 0, a_1 = 1$, entonces $T_n(\bar{X}_n) = \hat{\lambda}$. El estimador insesgado de la varianza del estimador insesgado, $\hat{\lambda}$, es $V_n(\bar{X}_n) = \frac{2}{n} \left(2\hat{\lambda} + \nu - \frac{4}{n}\right)$.

Ejemplo 4.6.2 Para obtener un estimador insesgado de la función de densidad (4.5.1), necesitamos la expresión de las derivadas de $f(x, \lambda, \nu)$. Si

$$p(j, \lambda) = \frac{1}{j!} \exp \left\{ -\frac{\lambda}{2} \right\} \left(\frac{\lambda}{2} \right)^j,$$

entonces

$$D_{\lambda}^k p(j, \lambda) = \frac{1}{\lambda^k} C_k^{(\frac{\lambda}{2})}(j) p(j, \lambda), \quad k \geq 0, \quad (4.6.13)$$

donde $C_k^{(\frac{\lambda}{2})}$ es el k -ésimo polinomio de Charlier definido en (1.2.4), ver Chihara (1978) [14]. A partir de (4.6.13), tenemos

$$D_{\lambda}^k f(x, \lambda, \nu) = \frac{1}{\lambda^k} \sum_{j=0}^{\infty} C_k^{(\frac{\lambda}{2})}(j) p(j, \lambda) f(x, 0, \nu + 2j). \quad (4.6.14)$$

Con el fin de dar un ejemplo numérico, hemos generado una muestra de tamaño $n = 20$ de una distribución chi-cuadrado no centrada con parámetros $\lambda = 5$ y $\nu = 3$. En este caso, la función de densidad es

$$f(x, 5, 3) = \exp \left(-\frac{x}{2} \right) x^{\frac{1}{2}} P(x), \quad (4.6.15)$$

con

$$P(x) = 1.637 \cdot 10^{-2} + 1.364 \cdot 10^{-2} x + 3.411 \cdot 10^{-3} x^2 + \dots$$

Para esta muestra, hemos obtenido $\bar{X}_n = 7.1522$, entonces $\hat{\lambda} = 4.1552$. Usando (4.6.14) y siete términos de la serie (4.6.12), la aproximación del estimador insesgado de la función de densidad (4.6.15) es

$$\hat{f}(x, 5, 3) = \exp \left(-\frac{x}{2} \right) x^{\frac{1}{2}} \hat{P}(x), \quad (4.6.16)$$

con

$$\hat{P}(x) = 2.332 \cdot 10^{-2} + 1.726 \cdot 10^{-2} x + 3.712 \cdot 10^{-3} x^2 + \dots$$

En la figura 1, la línea de puntos representa la verdadera función de densidad dada en (4.6.15) y la línea continua es la densidad estimada dada en (4.6.16).

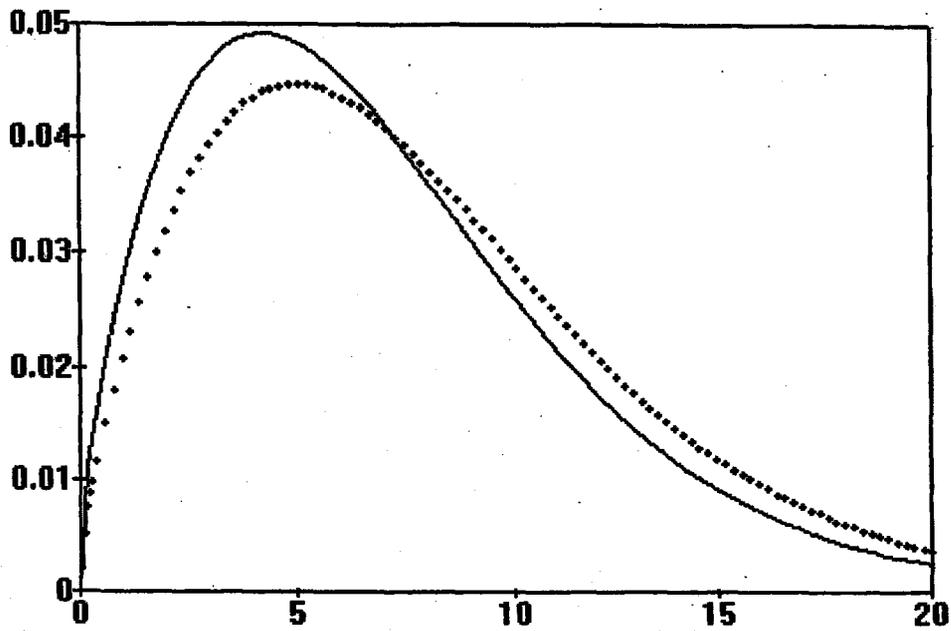


Figura 1

4.7 Propiedades Asintóticas

Sea $h(\lambda)$ una función que admite un estimador insesgado $T_n(\bar{X}_n)$. Las propiedades asintóticas de $T_n(\bar{X}_n)$ se derivan de las propiedades de los polinomios $Q_{j,n}$ anteriormente definidos. Sea λ_0 el verdadero, pero desconocido, valor del parámetro y $k_0 = \min\{j \geq 1 : D_\lambda^j h(\lambda_0) \neq 0\}$.

Teorema 4.7.1 *Si existe $R_0 > 0$ tal que*

$$\overline{\lim}_{j \rightarrow \infty} \left(\frac{2}{R_0} \right)^j |D_\lambda^j h(\lambda_0)| < \infty, \quad (4.7.1)$$

entonces,

(i)

$$k_0! \left(\frac{n}{V_0} \right)^{\frac{k_0}{2}} (D_\lambda^{k_0} h(\lambda_0))^{-1} \{T_n(\bar{X}_n) - h(\lambda_0)\} \xrightarrow{d} H_{e_{k_0}}(Z), \quad (4.7.2)$$

con $Z \sim \mathcal{N}(0, 1)$ y H_{e_k} el k -ésimo polinomio de Hermite.

(ii)

$$\overline{\lim}_{n \rightarrow \infty} \left(\frac{n}{\log(\log n)} \right)^{\frac{k_0}{2}} |T_n(\bar{X}_n) - h(\lambda_0)| = \frac{2^{\frac{k_0}{2}} V_0^{\frac{k_0}{2}}}{k_0!} |D_\lambda^{k_0} h(\lambda_0)|, \quad P_{\lambda_0} \text{ c.s.}, \quad (4.7.3)$$

donde $V_0 = 2(2\lambda_0 + \nu)$.

Demostración.

(i) A partir de (4.6.10),

$$\begin{aligned} n^{\frac{k_0}{2}} (T_n(\bar{X}_n) - h(\lambda_0)) &= \frac{D_\lambda^{k_0} h(\lambda_0) n^{\frac{k_0}{2}}}{k_0!} Q_{k_0,n}(\bar{X}_n, \lambda_0, \nu) \\ &+ n^{\frac{k_0}{2}} \sum_{j=k_0+1}^{\infty} \frac{D_\lambda^j h(\lambda_0)}{j!} Q_{j,n}(\bar{X}_n, \lambda_0, \nu) = \\ &= T_{1,n} + T_{2,n}. \end{aligned}$$

Sea $Z_n = \left(\frac{n}{V_0}\right)^{\frac{1}{2}} (\bar{X}_n - (\lambda_0 + \nu))$. Para n tendiendo a infinito, el teorema central del límite nos dice que Z_n converge en distribución (bajo P_{λ_0}) a una v.a. Z que sigue una distribución $\mathcal{N}(0, 1)$. Entonces, usando la propiedad (3), tenemos,

$$T_{1,n} = \frac{D_{\lambda}^{k_0}(\lambda_0)}{k_0!} \left\{ V_0^{\frac{k_0}{2}} H_{e_{k_0}}(Z_n) + o_{P_{\lambda_0}}(n^{-\frac{1}{2}}) \right\} \xrightarrow{d} \frac{V_0^{\frac{k_0}{2}} D_{\lambda}^{k_0} h(\lambda_0)}{k_0!} H_{e_{k_0}}(Z),$$

con $Z \sim \mathcal{N}(0, 1)$. Para probar el resultado, hemos de demostrar que

$$T_{2,n} \xrightarrow{P_{\lambda_0}} 0, \text{ cuando } n \rightarrow \infty. \quad (4.7.4)$$

Consideremos la serie

$$T_{2,n} = n^{\frac{k_0}{2}} \sum_{j=k_0+1}^{\infty} \frac{D_{\lambda}^j h(\lambda_0)}{j!} Q_{j,n}(\bar{X}_n, \lambda_0, \nu), \quad (4.7.5)$$

$$T_{2,n}^* = n^{\frac{k_0}{2}} \sum_{j=k_0+1}^{\infty} (-1)^j Q_{j,n}(\bar{X}_n, \lambda_0, \nu) \frac{R_0^j}{(2^j j!)}. \quad (4.7.6)$$

La hipótesis (4.7.1) implica que para demostrar (4.7.4), es suficiente demostrar que $T_{2,n}^*$ converge en P_{λ_0} -probabilidad a 0. Como para cualquier ε tal que $0 < \varepsilon < 1/(2(k+1))$, $\bar{X}_n = \lambda_0 + \nu + o_{P_{\lambda_0}}(n^{-1/2+\varepsilon})$, entonces de la propiedad (5) tenemos

$$T_{2,n}^* = n^{k/2} o_{P_{\lambda_0}}(n^{-(k+1)(1/2-\varepsilon)}),$$

y de aquí la convergencia de $T_{2,n}^*$ en P_{λ_0} -probabilidad hacia 0.

- (ii) La ley del logaritmo iterado, Shirayev (1984) [45], aplicada a la sucesión de variables aleatorias i.i.d. $\{X_i - (\lambda_0 + \nu)\}_{i \geq 1}$ nos da,

$$\overline{\lim}_{n \rightarrow \infty} \frac{1}{(2 \log \log n)^{\frac{1}{2}}} Z_n = 1, \quad P_{\lambda_0} \text{ c.s.} \quad (4.7.7)$$

Entonces , (ii) es inmediato a partir de (4.7.7), del hecho de que $H_{e_{k_0}}$ es un polinomio mónico de grado k_0 y de

$$k_0! \left(\frac{n}{V_0} \right)^{\frac{k_0}{2}} \left(D_\lambda^{k_0} h(\lambda_0) \right)^{-1} \{ T_n(\bar{X}_n) - h(\lambda_0) \} = H_{e_{k_0}}(Z_n) + o_{P_{\lambda_0}}(n^{-\frac{k_0}{2}}).$$

□

La condición (4.7.1) es satisfecha por una gran clase de funciones que incluye los polinomios y las exponenciales. En particular, si $k_0 = 1$, (4.7.3) y (4.7.2) implica la consistencia y normalidad asintótica del estimador insesgado. En el caso en que $k_0 = 2$, la distribución límite, bajo P_{λ_0} , de la variable aleatoria $\frac{2n}{D_\lambda^2 h(\lambda_0) V_0} \{ T_n(\bar{X}_n) - h(\lambda_0) \} + 1$, es una distribución chi-cuadrado con un grado de libertad.

Bibliografía

- [1] ABBEY, J.L., DAVID, H. T. (1970). The construction of uniformly minimum variance unbiased estimators for exponential distributions. *Ann. Math. Statist.*, **41** (4), 1217-1222.

- [2] ABRAMOWITZ, M., STEGUN, I.A. (1965). *Handbook of mathematical functions*. Dover, New York.

- [3] AKAHIRA, M. (1982). Asymptotic optimality of estimators in non-regular cases. *Ann. Inst. Math.*, **34** (A), 69-82.

- [4] ARNOLD, C.A., BALAKRISHNAN, N., NAGARAJA, H.N. (1998). *Records*. Wiley & Sons, New York.

- [5] BAHADUR, R.R. (1957). On unbiased estimates of uniformly minimum variance. *Sankhyä*. **18**, 211-224.

- [6] BAIN, L.J. (1978). *Statistical analysis of reliability and life-testing models*. Marcel Dekker, New York.
- [7] BALAKRISHNAN, N., CLIFFORD COHEN, A. (1991). *Order statistics and inference*. Academic Press Inc.
- [8] BARANKIN, E.W. (1949). Locally best unbiased estimates. *Ann. Math. Statist.*, **20**, 477-501.
- [9] BARNDORFF-NIELSEN, O. (1978). *Information and exponential families in statistical theory*. Wiley, New York.
- [10] BARNDORFF-NIELSEN, O. (1961). On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables. *Math. Scand.* **9**, 383-394.
- [11] BARRANCO CHAMORRO, I., LÓPEZ BLÁZQUEZ, F., MORENO REBOLLO, J.L. (1999). Estimation in Uniform distributions using orthogonal polynomials. *Commun. Statist. Theory and Methods*, **28** (5), 1145-1167.

- [12] BICKEL, P.J., DOKSUM, K.A. (1977) *Mathematical statistics: Basic ideas and selected topics*. Holden-Day.
- [13] BROWN, L.D. (1986). *Fundamentals of statistical exponential families with applications in statistical decision theory*. Lecture notes-monograph series, Vol.9, Institute of Matematical Statistics.
- [14] CHIHARA, T. (1978). *An introduction to orthogonal polynomials*. Gordon & Breach, New York.
- [15] CHOW, M.S. (1987). A complete class theorem for estimating a non-centrality parameter. *Ann. Statist.*, 15 (2), 800-804.
- [16] DAVID, H.A. (1970). *Order statistics*. Wiley & Sons, New York.
- [17] FELLER, W. (1968). *An introduction to probability theory and its application*. 3^a ed. Wiley, New York.
- [18] FEND, A.V. (1959). On the attainment of Cramer-Rao and Bhattacharya bounds for the variances of an estimate. *Ann. Math. Statist.*, 30, 381-388.

- [19] GALAMBOS, J. (1987). *The asymptotic theory of extreme order statistics*. Krieger, Malabar.
- [20] GLASSER, G.J. (1962). Minimum variance unbiased estimators for Poisson probabilities. *Technometrics*, 4 (3), 409-418.
- [21] GRAY, H.L., WATKINS, T.A., SCHUCANY, W.R. (1973). On the jackknife statistic and its relation to UMVU estimators in the normal case. *Commun. Statist.*, 2, 285-308.
- [22] HODGES, J.L., LEHMANN, E.L. (1970). Deficiency. *Ann. Math. Statist.*, 41 (3), 783-801.
- [23] HWANG T., HU, CH. (1990). More comparisons of mle with umvue for exponential families. *Ann. Inst. Math.*, 42 (1), 65-75.
- [24] JOHNSON, N.L., KOTZ, S. (1970). *Continuous univariate distributions*, Vol. 2. Wiley, New York.
- [25] KHAN, R.A. (1984). On umvu estimators and Bhattacharya bounds in exponential distributions. *J. Statist. Plann. Infer.*, 9, 199-206.

- [26] KINGMAN, J.F.C., TAYLOR, S.J. (1966). *Introduction to measure and probability*. Cambridge University Press, Cambridge.
- [27] LEHMANN, E.L. (1959). *Testing statistical hypotheses*, Wiley, New York.
- [28] LEHMANN, E.L. (1983). *Theory of point estimation*, Wiley, New York.
- [29] LINDSAY, B.G. (1989). On the determinants of moments matrices. *Ann. Statist.*, **17** (2), 711-721.
- [30] LÓPEZ BLÁZQUEZ, F. (1998). A note on umvu estimation in the transformed Chi-square family. *Commun. Statist. Theory and Methods*, **27** (8), 2065-2073.
- [31] LÓPEZ BLÁZQUEZ, F., CASTAÑO MARTÍNEZ, A. (1999). Asymptotic properties of unbiased estimators in the natural exponential family with quadratic variance function. *Sankhyā* **61** (A), 292-297.
- [32] LÓPEZ BLÁZQUEZ, F., GÓMEZ GÓMEZ, T., ZAFRA GARRIDO, M.J. (1997). Estimación de mínima varianza en distribuciones normales. *XXIII Congreso nacional de Estadística e Investigación Operativa*.

26.9-26.10.

- [33] LÓPEZ BLÁZQUEZ, F., GÓMEZ GÓMEZ, T., ZAFRA GARRIDO, M.J., BARRANCO CHAMORRO, I. (1995). *Estimación insesgada de mínima varianza y de máxima verosimilitud en distribuciones Poisson. XXIII Congreso nacional de Estadística e Investigación Operativa*. 191-192.
- [34] MEYER, P.L. (1967). The maximum likelihood estimate of the noncentrality parameter of a noncentral chi-squared variable. *J. Amer. Statist.*, **62**, 1258-1264.
- [35] MORRIS, C.N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.*, **10**, 65-80.
- [36] MORRIS, C.N. (1983). Natural exponential families with quadratic variance functions: Statistical theory. *Ann. Statist.*, **11**, 515-529.
- [37] NEFF, N., STRAWDERMAN, E. (1976). Further remarks on estimating the parameter of a noncentral chi-square distribution. *Commun. Statist. Theor. Meth., A*, **5** (1), 65-76.

- [38] PERLMAN, M.D., RASMUSSEN, A. (1975). Some remarks on estimating a noncentrality parameter. *Commun. Statist.*, 4 (3), 433-468.
- [39] RAHMANN, H.S., GUPTA, R.P. (1993). Family of transformed chi-square distributions. *Comm. Statist. Theory and Methods*, 22 (1), 135-146.
- [40] ROY, J., MITRA, S.K. (1958). Unbiased minimum variance estimation in a class of discrete distributions. *Sankhyā*, 18, 371-378.
- [41] SAXENA, L., ALAM, K. (1982). Estimation of the noncentrality parameter of a chi squared distribution. *Ann. Statist.*, 10 (3), 1012-1016.
- [42] SERFLING, R.J. (1980). *Aproximations theorems of mathematical statistics*. John Wiley & Sons, New York.
- [43] SHAO, J. (1999). *Mathematical statistics*. Springer-Verlag, New York Inc..
- [44] SHAO, Y.S., STRAWDERMAN, E. (1995). Improving on the positive part of the umvue of a noncentrality parameter of a noncentral chi square distribution. *J. Multivariate Anal.*, 53, 52-56.
- [45] SHIRYAYEV, A.N. (1984). *Probability*. Springer-Verlag, New York Inc..

- [46] SHORROCK, R.W. (1972). A limit theorem for inter-record times. *Journal of Applied Probability*, **9**, 210-223.
- [47] STEIN, C. (1950). *Unbiased estimates with minimum variance*. Ann. Math. Statist., **21**, 406-415.
- [48] STUART, A., ORD, J.K. (1991). *Kendall's advanced theory of statistics. Classical inference and relationship*. (5^a ed., Vol.1 y 2). Edward Arnold, London.
- [49] TATE, R.F. (1959). Unbiased estimation: Functions of location and scale parameters. *Annals of Mathematical Statistics*, **30**, 341-366.
- [50] TIKU, M.L. (1965) Laguerre series form of non-central χ^2 and F distributions. *Biometrika*, **52**, 3 y 4, p.415.
- [51] VASUDEVA, R. (1999). Law of the iterated logarithm for moving maxima. *Sanhkya*, **61** (2), 166-173.
- [52] WASHIO, Y., MORIMOTO, H., IKEDA, N. (1956). Unbiased estimation based on sufficient statistics. *Bull. Math. Statist.*, **6**, 69-94.

- [53] WOODWARD, W.A., GRAY, H.L. (1975). Minimum variance unbiased estimation in the Gamma distribution. *Commun. Statist.*, **4** (10), 907-922.
- [54] WOODWARD, W.A., GARY, D.K. (1977). Minimum variance unbiased of $P[Y < X]$ in the normal case. *Technometrics*, **19** (1), 95-98.
- [55] ZACKS, S. (1971). *The theory of statistical inference*. Wiley, New York.
- [56] ZACKS, S., EVEN, M. (1966). The efficiencies in small samples of the maximum likelihood and best unbiased estimators of reliability functions. *J. Am. Statist. Assoc.*, **61**, 1033-1051.

UNIVERSIDAD DE SEVILLA

Reunido el Tribunal integrado por los abajo firmantes en el día de la fecha, para juzgar la Tesis Doctoral de D. M^o Teresa Gomez Gomez titulada "Desarrollo atípico para estimadores en familias paramétricas"

acuerdo otorgarle la calificación de Sobresaliente Cum Laude por unanimidad

Sevilla, 21 de marzo

2000

El Vocal,

El Vocal,

El Vocal,

Francisco Coma

Rafael Infante

Antonio Saez
El Doctorado

El Secretario

El Secretario

Luis Barras

José Muñoz

M^o Teresa Gomez Gomez

