Intervalos de Confianza Jackknife para Cuantiles en Muestreo con Probabilidades Desiguales

Mario J. Pacheco[1]
Hugo A. Brango[2]

Resumen

Se considera la estimación de cuantiles en poblaciones finitas mediante la técnica *jackknife*. Se emplea un estimador *jackknife* de varianza para muestreo con probabilidades desiguales que funciona mejor que el estimador *jackknife* clásico. La calidad del intervalo de confianza hallado se demuestra vía simulación. El intervalo de confianza propuesto mostró probabilidades de cobertura cercanas al nivel de confianza nominal, así como longitudes promedio y varianzas mucho menores que las longitudes promedio y las varianzas de los intervalos empleando la metodología *jackknife* tradicional.

Palabras clave

Jackknife, intervalos de confianza, muestreo con probabilidades desiguales.

1 Departamento de Matemáticas y Estadística, Universidad de Córdoba, mariopachecolopez@gmail.com
2 Departamento de Matemáticas y Estadística, Universidad de Córdoba, habrango@hotmail.com

Fecha de recepción: 30 de Julio de 2010
Fecha de aceptación: 21 de Noviembre de 2010

Abstract

We consider the estimation of quantiles in finite populations using the jackknife technique. We use a jackknife variance estimator for unequal probability sampling that works better than the classical jackknife estimator. The quality of the confidence interval found is demonstrated via simulation. The proposed confidence interval showed coverage probabilities that were close to the nominal confidence level, and mean lengths and variances much smaller than the mean lengths and variances of the intervals using the traditional jackknife methodology.

Keywords

Confidence intervals, jackknife, sampling with unequal probabilities.

1. INTRODUCCIÓN

Es común en la práctica estadística y en especial en la aplicación de diseños de muestreo probabilístico que se plantee como parámetro de interés funciones no lineales de otros parámetros, Román-Montoya, et al. (2008). Un parámetro que resulta ser de mucho interés en la práctica son los cuantiles poblacionales y en particular la mediana de la variable en estudio.

En la literatura de muestreo encontramos diferentes formas de estimar cuantiles poblacionales. Entre ellas se destacan algunos estimadores directos e indirectos. Entre los estimadores que se consideran directos se destaca la función inversa de la función de distribución acumulada. Entre los estimadores indirectos se tienen aquellos que emplean información auxiliar en el cálculo del estimador a través de estimadores de razón, de diferencia y de regresión, como los propuestos por Kuk & Mak (1989), Rao et al. (1990) y Rueda et al. (2003). Además de estos estimadores se tienen los estimadores *jackknife* directos e indirectos de los cuantiles poblacionales propuestos por Román-Montoya et al. (2008) los cuales demuestran tener un buen comportamiento.

En relación a la construcción de intervalos de confianza para cuantiles poblacionales se tienen los intervalos de confianza que podemos llamar asintóticos que funcionan bajo muestreo aleatorio simple y un tamaño muestral grande. Usando la técnica *jackknife* se encuentra el estimador propuesto por Román-Montoya et al. (2008). En su propuesta se remplaza la varianza asintótica con la varianza *jackknife* tradicional.

Existen numerosos estudios del método *jackknife* en muestreo aleatorio simple y muestreo aleatorio simple estratificado. Cochran (2000), Wolter (1985) y Särndal et al. (1992) aplican el método *jackknife* en la estimación de la varianza de estimadores basados en muestras aleatorias sin remplazo. Jones (1974) aplica el *jackknife* en muestreo estratificado de poblaciones multivariadas de tamaño finito. Shao & Tu (1995) derivan estimadores de varianza de una estadística dada como parte crucial de muestreo por encuestas; además introducen las ideas básicas, fórmulas, implementaciones, propiedades y aplicaciones de este método para datos muestrales.

En relación a muestreo con probabilidades desiguales, Berger y Skinner (2005) presentan un estimador *jackknife* de varianza en muestreo con probabilidades desiguales análogo al estimador de varianza con la técnica de linealización de primer orden de Taylor expuesto en Särndal et al. (1992). Pacheco & Martínez (2007) emplean este estimador para obtener un estimador de la varianza en muestreo en dos fases con probabilidades desiguales. Este estimador resulta ser más consistente que el estimador *jackknife* de varianza tradicional.

En el presente artículo se emplea el estimador *jackknife* para muestreo con probabilidades desiguales propuesto por Berger y Skinner (2005) para la construcción de intervalos de confianza del estimador de un cuantil poblacional. Se compara el estimador *jackknife* clásico en la construcción del intervalo con el estimador *jackknife* para muestreo con probabilidades desiguales. Adicionalmente se comparan los intervalos de confianza empleando diferentes estimadores de los cuantiles poblacionales.

2. ESTIMACIÓN DE CUANTILES

Sea $y_1, y_2, \dots, y_N$ los valores de los elementos de una población finita U, para una variable de estudio y. Para cualquier número dado t $(-\infty < t < \infty)$ la función de distribución acumulada poblacional $F_y(t)$ se define como la proporción de elementos en la población para los que $y \leq t$. Dicha función de distribución la podemos calcular como:

$$F_y(t) = \frac{1}{N}\sum_{k\in U} z_k \tag{1}$$

De esta forma, a partir de (1), el cuantil poblacional podrá ser calculado mediante la expresión

$$Q_y(\beta) = \text{ínf}\{t: F_y(t) \geq \beta\} \tag{2}$$

con $z_k = 1$, si $y_k \leq t$ y $z_k = 0$ en otro caso. El interés entonces es estimar el cuantil poblacional (2) mediante una muestra s obtenida de un diseño de muestreo con probabilidades desiguales.

Primero, para estimar $F_y(t)$, para un valor t dado, con una muestra s se emplea la expresión

$$\hat{F}_y(t) = \frac{\sum_{k \in s} z_k / \pi_k}{\hat{N}} \tag{3}$$

con $\hat{N} = \sum_{k \in s} \pi_k^{-1}$ y π_k la probabilidad de inclusión del individuo k-ésimo en la muestra s.

Luego un estimador natural para (2) consiste en estimar la función de distribución acumulada $F_y(t)$ mediante (3) y así estimar el cuantil poblacional con la función inversa de la función de distribución acumulada estimada. Esto es,

$$\hat{Q}_y(t) = \hat{F}_y^{-1}(t) = inf\{t: \hat{F}_y(t) \geq \beta\} \tag{4}$$

2.1 Estimadores que Emplean Información Auxiliar

El uso de información auxiliar para mejorar la precisión de las estimaciones es característico de la teoría de muestreo (Särndal et al., 1992). El uso de esta información auxiliar altamente correlacionada con nuestra variable de estudio y es empleada para construir estimadores más eficientes mediante los estimadores de razón, diferencia o de regresión, teniendo en cuenta que esta información auxiliar es usualmente conocida para todos los elementos de la población.

Kuk y Mak (1989) proponen un estimador de razón de $Q_y(\beta)$ definido por la expresión

$$\hat{Q}_r(\beta) = \hat{Q}_y(\beta) \frac{Q_x(\beta)}{\hat{Q}_x(\beta)} \tag{5}$$

con $Q_x(\beta)$ el cuantil poblacional de la variable x y $\hat{Q}_x(\beta) = inf\{t: \hat{F}_x(t) \geq \beta\}$ su estimador y $\hat{F}_x(t)$ la función de distribución acumulada estimada de la variable x.

Los estimadores de diferencia por su parte son de la forma

$$\hat{Q}_d(\beta) = \hat{Q}_y(\beta) + b\left(Q_x(\beta) - \hat{Q}_x(\beta)\right) \tag{6}$$

con algunas selecciones para b que optimizan el estimador. Entre los valores de b encontramos, uno que optimiza el estimador en términos de la minimización de la varianza es

$$b = \frac{cov\left(\hat{Q}_y(\beta), \hat{Q}_x(\beta)\right)}{v\left(\hat{Q}_x(\beta)\right)} \tag{7}$$

una estimación para b se encuentra en Rueda et al. (2003). Otra elección para b es dada por Rao et al. (1990), tomando a b como el estimador de la razón entre los totales de las variables x y y,

$$b = \hat{R} = \frac{\sum_{i\in s}\frac{y_i}{\pi_i}}{\sum_{i\in s}\frac{x_i}{\pi_i}} \tag{8}$$

Una última elección para b da como resultado el estimador de regresión dado en Rueda et al. (2003),

$$b = \frac{\sum_{i\in s}\frac{y_i x_i}{\pi_i}}{\sum_{i\in s}\frac{x^2{}_i}{\pi_i}} \tag{9}$$

Notar además que si seleccionamos $b = \hat{Q}_y(\beta)/\hat{Q}_x(\beta)$ en el estimador de diferencia y si $\hat{Q}_y(\beta) \neq 0$ y $Q_x(\beta) - \hat{Q}_x(\beta) \neq 0$ entonces $\hat{Q}_r(\beta) = \hat{Q}_d(\beta)$.

Es importante tener en cuenta para la selección de b que la propuesta en (7) puede resultar, en la práctica, de muy difícil cálculo, mientras que la (8) y (9) requieren de cálculos mucho más

sencillos, aunque la expresión dada en (8) es mucho más comúnmente usada.

3. INTERVALOS DE CONFIANZA PARA CUANTILES USANDO VARIANZA JACKKNIFE

Para la estimación *jackknife* primero se muestra el uso del estimador tradicional para la estimación del cuantil $Q_y(\beta)$ de una población U de N individuos a partir de la cual seleccionamos una muestra s de tamaño n con un diseño muestral cualquiera con probabilidades desiguales.

Sea la muestra s_i la muestra luego de eliminar el elemento i-ésimo de s y $\hat{Q}_y^{(i)}(\beta)$ un estimador de $Q_y(\beta)$ de la misma forma funcional que $\hat{Q}_y^*(\beta)$ (un estimador cualquiera de $Q_y(\beta)$ calculado con la muestra s_i. Román-Montoya *et al*. (2008) proponen un estimador *jackknife* para $Q_y(\beta)$ de la forma

$$\hat{Q}_y^{JK}(\beta) = \frac{1}{n}\sum_{i=1}^{n} \hat{Q}_y^i(\beta) \tag{10}$$

donde $\hat{Q}_y^i(\beta) = n\hat{Q}_y^*(\beta) - (n-1)\hat{Q}_y^{(i)}(\beta)$ son los denominados pseudo-valores *jackknife*.

Una expresión para la estimación de la varianza del estimador de $Q_y(\beta)$ empleando el estimador (10) se consigue como

$$\hat{V}_{JK} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{Q}_y^i(\beta) - \hat{Q}_y^{JK}(\beta)\right)^2 \tag{11}$$

Román-Montoya et al. (2008) muestran algunas propiedades del estimador *jackknife* para muestreo aleatorio simple.

Para la estimación de intervalos de confianza *jackknife* Román-Montoya et al. (2008) proponen un estimador basado en el supuesto de normalidad del estimador *jackknife* y en el estimador de varianza *jackknife* (11) de la forma

$$\left[\hat{Q}_y^*(\beta) - z_{\frac{\alpha}{2}}\hat{V}_{JK}, \hat{Q}_y^*(\beta) + z_{\frac{\alpha}{2}}\hat{V}_{JK}\right] \tag{12}$$

dado que los estimadores asintóticos para los estimadores de razón, de diferencia y de regresión presentan dificultades para hallar las varianzas asintóticas de los estimadores. Además de esto, Román-Montoya et al. (2008) muestra vía simulación que los intervalos *jackknife* presentan mejores resultados que los estimadores asintóticos bajo muestreo aleatorio simple.

Berger y Skinner (2005) presentan un estimador *jackknife* de varianza en muestreo con probabilidades desiguales. El estimador propuesto modifica los pseudo-valores *jackknife* para lograr un estimador de varianza análogo al estimador de varianza con la técnica de linealización de primer orden de Taylor expuesto en Särndal et al. (1992). Este estimador resulta ser más consistente que los estimadores *jackknife* alternativos en la estimación de la varianza del estimador del parámetro de interés.

Para construir un estimador *jackknife* en muestreo con probabilidades desiguales empleando el estimador de Berger y Skinner (2005), definimos nuevamente s_i como la muestra luego de eliminar el elemento i-ésimo de s. Luego la expresión para la varianza estimada es

$$\hat{V}_{BS} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}\varepsilon_{(i)}\varepsilon_{(j)} \tag{13}$$

con π_{ij} la probabilidad de inclusión de los individuos i y j en la muestra s,

$$\begin{aligned}\varepsilon_{(i)} &= (1 - w_i^{-1})\left(\hat{Q}_y^*(\beta) - \hat{Q}_y^{(i)}(\beta)\right)\\ &= w_i^{-1}\left(\hat{Q}_y^*(\beta) + \hat{Q}_y^i(\beta)\right)\end{aligned} \tag{14}$$

donde $w_i = \hat{N}\pi_i$ y $\hat{Q}_y^i(\beta) = w_i\hat{Q}_y^*(\beta) - (w_i - 1)\hat{Q}_y^{(i)}(\beta)$. Los valores $\varepsilon_{(i)}$ en (14) son equivalentes a los pseudo-valores *jackknife* empleados en el estimador de varianza *jackknife* clásico en (10).

De esta manera para muestreo con probabilidades desiguales se propone remplazar la varianza *jackknife* tradicional en (11) basada en la teoría clásica del *jackknife* por el estimador *jackknife* para muestreo con probabilidades desiguales en (13) propuesto por Berger y Skinner (2005). Esto es,

$$\left[\hat{Q}_y(\beta) - z_{\frac{\alpha}{2}}\hat{V}_{BS}, \hat{Q}_y(\beta) - z_{\frac{\alpha}{2}}\hat{V}_{BS}\right] \tag{15}$$

4. ESTUDIO POR SIMULACIÓN

En esta sección se lleva a cabo un estudio de simulación para comparar los estimadores de los intervalos de confianza para cuantiles mediante la metodología *jackknife*.

Para hacer esto se considera un muestreo aleatorio con probabilidades proporcionales al tamaño, pps, sin remplazo, mediante la metodología de Sunter como se describe en Särndal et al. (1992).

El cuantil a estimar fue la mediana ($\beta = 0.5$) dada su importancia práctica y los intervalos considerados fueron con un nivel de confianza nominal del 95%.

Para las simulaciones se considera la población MU284 de Särndal et al. (1992) eliminando 3 observaciones atípicas. Un primer estudio se realiza sobre las variables RMT85 ($= y$) y CS82 ($= x$), las cuales tienen un coeficiente de correlación de 0,6575. Un segundo estudio se realiza sobre las variables P85 ($= y$) y P75 ($= x$), las cuales tienen un coeficiente de correlación de 0,9950. También se considera la población Lucy descrita en el paquete TeachingSampling de R. Las variables empleadas con esta población fueron Income ($= y$) y Taxes ($= x$), las cuales tienen un coeficiente de correlación de 0,9169.

Para cada estudio se seleccionaron 1000 muestras de tamaño n =25 y 50 para la población MU284 y n =25, 50 y 100 para la población Lucy, y para cada una de ellas se calculó el intervalo *jackknife* clásico (12) y el intervalo *jackknife* para muestreo con probabilidades desiguales (15).

Para la comparación de los métodos usados para la estimación de los intervalos, se estimó su probabilidad de cobertura como el porcentaje de veces que el intervalo contiene el verdadero valor del parámetro, la longitud promedio y la varianza de la longitud del intervalo. Un buen método debe proponer intervalos con probabilidades de cobertura muy cercanas a los niveles de confianza nominal y con valores pequeños del promedio y de la varianza de su longitud.

Antes de analizar los resultados del estudio por simulación, se muestran en la Tabla 1, para un mejor entendimiento, los resultados del cálculo de un intervalo de confianza para el cuantil de la variable P85 de la población MU284 antes descrita. En esta se encuentra el valor real del cuantil Q_y y su estimación por intervalo mediante el *jackknife* clásico y propuesto mediante una muestra pps de tamaño 50 y el estimador de razón. Se observa que para esta muestra el estimador *jackknife* propuesto tiene una longitud menor a la del estimador *jackknife* clásico aunque esta evidencia no es suficiente para mostrar las ventajas de la nueva estrategia y por tanto se realiza el estudio por simulación.

Tabla 1. Estimación por intervalo para una muestra de tamaño 50

Población	y	x	Q_y	$\hat{Q}_R$	Intervalo estimado	
					$Linf$	$Lsup$
MU284	P85	P75	16	$\hat{Q}_{JK}$	-10,61	40,61
				$\hat{Q}_{BS}$	5,81	24,19

Para el estudio de simulación, los resultados se resumen en las Tablas 2 a 5. En las Tablas 2 y 3 se encuentran las probabilidades de cobertura cuando el nivel de confianza nominal es del 95%. En la Tablas 4 y 5 se encuentran el promedio de las longitudes y la varianza de las longitudes. Las Tablas 2 y 3 muestran que las probabilidades de cobertura para los intervalos empleando la varianza *jackknife* clásica supera el nivel nominal del 95% para los tres estimadores (4), (5) y (6) empleados cuando se toma la población MU284 y las variables RMT85 y CS82 las cuales tienen grado de correlación medio. Para la población MU284 y las

variables P85 y P75 que tienen un coeficiente de correlación cercano a uno se tienen probabilidades de cobertura inferiores al nivel de confianza nominal cuando se emplea el *jackknife* clásico, excepto cuando se tiene el estimador de regresión y un tamaño de muestra de 50. Para la población Lucy las probabilidades de cobertura empleando el *jackknife* clásico superan el nivel de confianza nominal. Para el caso del *jackknife* para muestreo con probabilidades desiguales, los intervalos de confianza construidos con la varianza de esta metodología no exceden el nivel de significancia nominal para un tamaño de muestra de 25 y una correlación media. También presenta niveles de significancia inferiores al nivel de confianza nominal cuando se emplea el estimador $\hat{Q}_a$. Para el resto de casos se tienen probabilidades de cobertura superiores al nivel nominal.

Tabla 2. Niveles de confianza real cuando el nivel nominal es del 95%, Población MU284.

Población	n	Estimador	JK	BS
MU284 RMT85, CS82 ($\rho = 0{,}675$)	25	$\hat{Q}_a$	0,963	0,869
		$\hat{Q}_R$	0,979	0,942
		$\hat{Q}_r$	0,985	0,946
	50	$\hat{Q}_a$	0,975	0,916
		$\hat{Q}_R$	0,986	0,964
		$\hat{Q}_r$	0,992	0,948
MU284 P85, P75 ($\rho = 0{,}995$)	25	$\hat{Q}_a$	0,903	0,928
		$\hat{Q}_R$	0,921	1,000
		$\hat{Q}_r$	0,941	1,000
	50	$\hat{Q}_a$	0,867	0,919
		$\hat{Q}_R$	0,931	0,996
		$\hat{Q}_r$	0,964	0,996

Tabla 3. Niveles de confianza real cuando el nivel nominal es del 95%, Población Lucy.

Población	n	Estimador	JK	BS
Lucy Income, Taxes ($\rho = 0{,}9169$)	25	$\hat{Q}_a$	0,962	0,896
		$\hat{Q}_R$	0,960	0,975
		$\hat{Q}_r$	0,997	0,992
	50	$\hat{Q}_a$	0,984	0,917
		$\hat{Q}_R$	0,982	0,994
		$\hat{Q}_r$	1,000	1,000
	100	$\hat{Q}_a$	0,992	0,944
		$\hat{Q}_R$	0,993	0,992
		$\hat{Q}_r$	1,000	0,999
	200	$\hat{Q}_a$	0,999	0,952
		$\hat{Q}_R$	1,000	0,993
		$\hat{Q}_r$	1,000	1,000

Tabla 4. Promedio y varianza de las longitudes de los intervalos cuando el nivel nominal es del 95%, Población MU284.

Población	n	Estimador	JK		BS	
			Prom.	Var.	Prom.	Var.
MU284 RMT85, CS82 ($\rho = 0{,}675$)	25	$\hat{Q}_a$	770,87	450585,4	148,27	11296,6
		$\hat{Q}_R$	858,90	367727,5	192,94	12946,6
		$\hat{Q}_r$	1009,18	417108,5	227,72	16392,8
	50	$\hat{Q}_a$	779,02	408978,3	103,64	4742,7
		$\hat{Q}_R$	916,47	366347,1	137,40	6010,6
		$\hat{Q}_r$	1121,61	548390,6	166,91	10240,9
MU284 P85, P75 ($\rho = 0{,}995$)	25	$\hat{Q}_a$	110,68	8685,45	20,79	133,97
		$\hat{Q}_R$	57,78	1721,30	23,18	89,02
		$\hat{Q}_r$	63,28	2130,86	23,82	94,47
	50	$\hat{Q}_a$	116,66	8835,17	15,16	73,01
		$\hat{Q}_R$	94,15	4049,04	18,84	79,57
		$\hat{Q}_r$	101,29	4821,21	19,55	88,02

Aunque los niveles de confianza simulados de los intervalos de confianza empleando la metodología clásica del *jackknife* supera en algunos casos a los niveles de confianza de los intervalos que emplean el estimador de varianza para muestreo con probabilidades desiguales, se observa en las Tablas 4 y 5 que las longitudes promedio de los intervalos con el *jackknife* tradicional son mucho mayores, y en algunos casos es dos o hasta 8 veces la longitud empleando el *jackknife* para muestreo con probabilidades desiguales. Igual situación sucede con la varianza de las longitudes, siendo mucho mayores en el *jackknife* tradicional.

Tabla 5. Promedio y varianza de las longitudes de los intervalos cuando el nivel nominal es del 95%, Población Lucy.

Población	n	Estimador	JK		BS	
			Prom.	Var.	Prom.	Var.
Lucy Income, Taxes (ρ=0,9169)	25	$\hat{Q}_a$	2661.65	5107873.69	608.42	93499.72
		$\hat{Q}_R$	5707.45	71072045.27	2177.57	7284657.90
		$\hat{Q}_r$	1190.52	548073.02	1001.88	218938.24
	50	$\hat{Q}_a$	3020.38	4451129.12	514.96	48461.38
		$\hat{Q}_R$	6627.93	89166796.98	1712.26	3427554.95
		$\hat{Q}_r$	1567.65	787851.82	879.16	129858.32
	100	$\hat{Q}_a$	3272.71	4332472.27	399.86	26291.37
		$\hat{Q}_R$	7225.58	11728769.59	1250.28	1807800.61
		$\hat{Q}_r$	2168.78	1563346.41	702.50	58249.58
	200	$\hat{Q}_a$	3392.23	4073209.00	294.30	11133.87
		$\hat{Q}_R$	8164.87	106844700.0	911.82	602959.10
		$\hat{Q}_r$	3401.00	5224345.00	557.27	36236.00

5. CONCLUSIONES

Este estudio propone la construcción de intervalos de confianza en muestreo con probabilidades desiguales mediante la técnica *jackknife*. Se muestra como la escogencia de la metodología clásica del *jackknife* no es adecuada al entregar intervalos de confianza excesivamente amplios. Los intervalos propuestos se basan en la

estimación de la varianza de los cuantiles estimados mediante la metodología de estimación *jackknife* de varianza para muestreo con probabilidades desiguales de Berger y Skinner (2005). Los intervalos propuestos presentan probabilidades de cobertura cercanas al nivel de confianza nominal la cual mejora con el empleo de estimadores que usen información auxiliar, como es el caso de los estimadores de razón y de regresión. Las longitudes de los intervalos propuestos resultaron además mucho menores que las longitudes de los intervalos empleando el *jackknife* clásico.

6. REFERENCIAS

Berger, Y., Skinner, C., (2005); A jackknife variance estimator for unequal probability sampling, Journal of the Royal Statistical Society B, 67, 79–89.

Cochran, W.G., (2000); Técnicas de muestreo, Compañía Editorial Continental, México.

Jones, H.L., (1974); Jackknife estimation of functions of stratum means, Biometrika 61(2), 343–348.

Kuk, A., Mak, T.K., (1989); Median estimation in the presence of auxiliary information. J R Stat Soc B 51(2), 261–269.

Pacheco, M., Martínez, G., (2007); Un estimador jackknife de varianza en muestreo en dos fases con probabilidades desiguales. Revista Colombiana de Estadística, 30, 203–212.

Rao, J.N.K., Kovar, J.G., Mantel, H.J., (1990); On estimating distribution functions and quantiles from survey data using auxiliary information. Biometrika 77, 365–375.

Román-Montoya, Y, Rueda, R., Arcos, A., (2008); Confidence intervals for quantile estimation using Jackknife techniques, Comput. Stat., 23, 573–585.

Rueda, M.M., Arcos, A., Martínez, M.D., (2003); Difference estimators of quantiles in finite populations. Test 12(2), 481–496.

Särndal, C.E., Swensson, B., Wretman, J.H., (1992); Model Assisted Survey Sampling, Springer-Verlag, New York.

Shao, J., Tu, D., (1995); The Jackknife and Bootstrap, Springer-Verlag, New York.

Wolter, K.M., (1985); Introduction to Variance Estimation, Springer-Verlag, Berlín.