

**ANÁLISIS DE VALORES EXTREMOS PARA CONTROLES
DE VALIDACIÓN DE LA ENCUESTA COYUNTURAL
DE LA CONSTRUCCIÓN**

Cristina Prado



**EUSKAL ESTATISTIKA ERAKUNDEA
INSTITUTO VASCO DE ESTADISTICA**

Donostia-San Sebastián, 1
01010 VITORIA-GASTEIZ
Tel.: 945 01 75 00
Fax.: 945 01 75 01
E-mail: eustat@eustat.es
www.eustat.es

Índice

INDICE	3
FINALIDAD DEL ESTUDIO	4
MÉTODO APLICADO AL ESTUDIO.....	5
Cálculo de estadísticos y distribución de frecuencias de cada uno de los ratios por grupos de actividad.....	5
CÁLCULO DE LOS LÍMITES INFERIOR Y SUPERIOR DE UNA DISTRIBUCIÓN ASIMÉTRICA.	5
DETERMINACIÓN DE LOS VALORES EXTREMOS INFERIOR Y SUPERIOR PARA CADA UNO DE LOS RATIOS, POR ACTIVIDADES Y DISTINGUIENDO EDIFICACIÓN Y INGENIERÍA CIVIL.....	6
CÁLCULO DE LA MEDIA DE LOS VALORES EXTREMOS POR RATIOS DISTINGUIENDO ÚNICAMENTE EDIFICACIÓN Y INGENIERÍA CIVIL.....	6
CONCLUSIONES	7
Antiguo 501-502-504 Edificación	3

Finalidad del estudio

La finalidad de este estudio es la determinación de valores extremos de los ratios económicos que sirvan para los controles de validación utilizados en la Encuesta Coyuntural de la Construcción (ECC), ya que los existentes actualmente han quedado desfasados y no sirven para la función que le está asignada como es la de servir de filtro para la detección de errores, fundamentalmente de coherencia económica del cuestionario.

La definición de nuevos valores extremos se plantea desde el análisis de la serie de ratios existentes (aproximadamente 20 trimestres) y esto a través de la distribución de frecuencias de cada uno de ellos.

De igual forma este análisis de detección de valores extremos es de aplicación para controles de validación o para cualquier otra finalidad extensiva a otras operaciones de tipo económico, bien referente a los datos de un año como a varios, tanto a valores absolutos de ratios, como a crecimientos de variables (de gran interés en los controles de validación de las operaciones coyunturales).

Método aplicado al estudio

Calculo de estadísticos y distribución de frecuencias de cada uno de los ratios por grupos de actividad

Los ratios calculados son los utilizados en la validación de la ECC: Sueldos y Salarios sobre el Valor de los Trabajos Realizados, Sueldos y Salarios sobre el Personal Remunerado, Valor de los Trabajos Realizados sobre Personal Ocupado, Horas Trabajadas sobre Obreros Asalariados, Valor de los Trabajos Realizados sobre Horas Trabajadas.

El estudio de cada una de estos ratios se ha realizado por grupos de actividad (CNAEs 501, 502, 503, 504) distinguiendo en cada una de ellas la Edificación y la Ingeniería Civil.

El cálculo de los estadísticos así como las representaciones gráficas de la distribución de frecuencias se ha realizado en SAS a través de los procedimientos PROC UNIVARIANTE (Estadísticos) y el PROC CHART (Distribución de frecuencias).

Cálculo de los límites inferior y superior de una distribución asimétrica.

Una gran parte de las distribuciones de frecuencias de las series económicas siguen unas distribuciones que son diferentes a la distribución normal y por tanto son asimétricas, para estas últimas la fórmula de valores extremos es la siguiente:

$$LI = Q_1 - 1.5(Q_3 - Q_1) \quad LS = Q_3 + 1.5(Q_3 - Q_1)$$

Donde:

LI : Límite Inferior

LS: Límite Superior

Q1 : 1er. cuartil (que deja a la izquierda el 25% de las observaciones)

Q3 : 3er. cuartil (que deja a la izquierda el 75% de las observaciones)

Este cálculo se realizó en el mismo programa de sas en que se introducían los procedimientos anteriores.

Determinación de los valores extremos inferior y superior para cada uno de los ratios, por actividades y distinguiendo Edificación y Ingeniería Civil

La fórmula comentada fue de gran utilidad para la determinación del límite superior ya que deja fuera un número de observaciones aceptable en torno al 4 o 5%, y de hecho se ha seleccionado para la práctica totalidad de los límites superiores, en cambio los valores que da la fórmula para los extremos inferiores no son válidos para la mayoría de los casos, ya que dan valores negativos que no tienen sentido para trabajos de esta naturaleza. Por tanto el método seguido ha sido otro diferente, como es el de analizar la distribución para los valores inferiores y seleccionar aquel valor a partir del cual hubiera un salto importante en la porcentaje de distribución de frecuencias acumuladas.

Para este último punto era necesario por lo general volver a calcular la distribución de frecuencias para estos valores inferiores con el fin de hacer los intervalos de la distribución más pequeños y detectar mejor dicho salto. Este nuevo cálculo no presenta problemas ya que el PROC CHART es muy versátil a la hora de cambiar los intervalos de la distribución.

Cálculo de la media de los valores extremos por ratios distinguiendo únicamente Edificación y Ingeniería Civil

Una vez calculados los valores extremos en el apartado anterior se ha tratado de llegar para cada ratio a dos límites uno por Edificación y otro por Ingeniería Civil, tal y como está diseñado el programa de validación actual. La aplicación de valores extremos podría ser por actividades y Edificación y Ingeniería Civil pero no se ve grandes razones para cambiar a esta segunda opción. Por lo tanto únicamente se ha calculado la media de los límites dentro de Edificación y Ingeniería Civil tal y como aparece en el CUADRO I.

**ANÁLISIS DE VALORES EXTREMOS PARA CONTROLES
DE VALIDACIÓN DE LA ECC**

Cuadro 1. Análisis de valores extremos para controles de validación a partir de la distribución de frecuencias

	Sueldos /Vtr				Sueldos/Pr				VTR/Pr				Horas/Obreros				Vtr/Horas			
	LI	%<=LI	LS	<=LS	LI	%<=LI	LS	%<=LS	LI	%<=LI	LS	%<=LS	LI	%<=LI	LS	%<=LS	LI	<=LI	LS	%<=LS
Antiguo 501-502-504 Edificación	0,12	-	0,30	-	150	-	500	-	800	-	2400	-	430	-	480	-	3,000	-	8,000	-
Antiguo 501-503-504 Ing.Civil	0,12	-	0,30	-	150	-	500	-	800	-	2400	-	430	-	480	-	3,000	-	8,000	-
501 Edificación	0,07	11,220	0,70	-	304	2,000	998	1,800	357	2,000	5954	2,400	414	4,500	584	1,000	1,000	5,770	15,900	5,000
501 Ing.Civil	0,08	5,820	0,59	4,510	300	1,100	1024	1,000	553	4,200	5743	6,000	415	-	577	-	1,100	3,050	16,000	-
502 Edificación	0,08	3,500	0,74	2,300	300	3,760	894	2,700	496	5,400	4521	5,500	412	6,050	580	1,100	1,040	4,730	12,400	7,000
503 Ing.Civil	0,07	12,000	0,62	3,300	308	4,800	1038	1,000	467	4,000	6524	8,260	415	0,720	577	6,000	1,070	4,000	17,700	5,440
504 Edificación	0,10	4,900	0,79	1,500	288	3,750	894	3,600	493	2,100	3963	5,000	417	6,000	574	1,200	1,030	2,000	11,700	4,600
504 Ing.Civil	0,12	4,400	0,78	-	317	5,140	991	3,000	577	5,000	4242	2,600	393	6,400	588	1,000	1,400	1,800	12,300	4,300
MEDIA	0,09	-	0,70	-	303	-	973,17	-	490,50	-	5157,83	-	411,000	-	580,00	-	1,11	-	14,33	-
DESVIACIÓN TÍPICA	0,02	-	0,08	-	8,80	-	58,10	-	70,47	-	958,70	-	8,185	-	4,73	-	0,13	-	2,29	-
COEFICIENTE VARIACIÓN	0,20	-	0,11	-	0,03	-	0,06	-	0,14	-	0,19	-	0,020	-	0,01	-	0,12	-	0,16	-
EDIFICACIÓN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEDIA	0,08	-	0,75	-	297,33	-	928,67	-	448,67	-	4812,67	-	414,333	-	579,33	-	1,02	-	13,33	-
DESVIACIÓN TÍPICA	0,02	-	0,04	-	8,33	-	60,04	-	79,40	-	1027,05	-	2,517	-	5,03	-	0,02	-	2,25	-
COEFICIENTE VARIACIÓN	0,18	-	0,06	-	0,03	-	0,06	-	0,18	-	0,21	-	0,006	-	0,01	-	0,02	-	0,17	-
INGENIERIA CIVIL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MEDIA	0,09	-	0,66	-	308,33	-	1017,67	-	532,33	-	5503,00	-	407,667	-	580,67	-	1,19	-	15,33	-
DESVIACIÓN TÍPICA	0,03	-	0,10	-	8,50	-	24,13	-	57,84	-	1159,78	-	12,702	-	6,35	-	0,18	-	2,76	-
COEFICIENTE VARIACIÓN	0,29	-	0,15	-	0,03	-	0,02	-	0,11	-	0,21	-	0,0312	-	0,01	-	0,15	-	0,18	-