UNIVERSIDAD DE LOS ANDES
VENEZUELA

Diseño de Experimentos

Introducción

Rossana Timaure de Valery
rttg27@gmail.com

Instituto de Estadística Aplicada y Computación

Agenda

Introducción

Diagnóstico ?

Análisis de influencia

Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 5 (2017) 1116–1122

Structural Integrity Procedia

www.elsevier.com/locate/procedia

2nd International Conference on Structural Integrity, ICSI 2017, 4-7 September 2017, Funchal, Madeira, Portugal

Statistical Analysis of the Influence of Several Factors on Compressive Strength of Alkali Activated Fly Ash

Adelaide Cerveira[a,*] , Elisete Correia[b], Nuno Cristelo[c], Tiago Miranda[d], Fernando Castro[e], Ana Fernández-Jiménez[f]

[a] INESC–TEC and Department of Mathematics, UTAD, 5001-801 Vila Real, Portugal, cerveira@utad.pt
[b] CM-UTAD and Department of Mathematics, UTAD
[c] CQ-VR and Department of Engineering, UTAD
[d] ISISE and Department of Civil Engineering, UMinho
[e] W2V and Department of Mechanical Engineering, UMinho
[f] Eduardo Torroja Institute for Construction Science, Madrid

Abstract

The use of industrial by-products to produce new types of cement-substitute binders is gaining significant momentum, especially through the alkaline activation technique. However, the exact curing conditions that should be considered with each binder variation have not yet been fully understood. The aim of the present work is thus the statistical analysis of the effects of several factors, namely filler/precursor ratio and curing humidity, on the compressive strength of different mixtures prepared with mine tailings (filler), fly ash (precursor) and an alkali activator based on sodium hydroxide. Five different types of mixture were prepared, with filler/precursor ratios of 80/20, 60/40, 40/60, 20/80 and 0/100. All the specimens were cured at 80ºC for 7 days, after which they were submitted to a uniaxial compression strength (UCS) test. Three different values of curing humidity were considered, namely 25%, 50% and 75%. Each UCS value was the average of 9 different specimens tested.

The aim of the present research is to establish how much these two factors (inert/precursor ratio and curing humidity) influence the UCS. For that purpose, a two-way Analysis of Variance (ANOVA), with interaction, was performed; followed by a Tuckey's Post hoc test. The results showed statistically significant differences for at least one humidity value $F(2,127) = 31.647$ ($p<0.001$) – as well as one inert/precursor ratio – $F(4,127) = 371.64$; ($p<0.001$) and for interaction $F(8,127) = 9.33$; ($p<0.001$). To evaluate which level or levels are different a Tuckey's Post hoc test was performed. This test revealed that the humidity value of 50% presented

Formulación del modelo

$$y = X\beta + \epsilon$$

$$y = X\beta + Zb + \epsilon$$

Donde: y es un vector $(n \times 1)$ de datos observados.
X es una matriz diseño $(n \times p)$ de de efectos fijos de rango k.
Z es una matriz de diseño $(n \times g)$ de efectos aleatorios.
β es un vector $(p \times 1)$ de efectos fijos.
b es un vector $(g \times 1)$, de efectos aleatorios $b \sim N(0, \sigma^2 G)$.
ϵ es un vector de error $\epsilon \sim N(0, \sigma^2 R)$.
b y ϵ son independiente y los elementos de varianza desconocidos (parámetros de covarianza) se encuentran en un vector θ, en el modelo lineal general esta matriz es la identidad

Laird y Ware, 1982; Finucane, Samet y Horton, 2007; Van Rie y otros, 2009; Lachos, Bandyopadhyay y Dey; 2011. Estrella y otros, 2015; Mukerji y otros 2016. 1130-1139. Stirrup, Babiker y Copas, 2016.

Validación del modelo

Análisis de lo supuestos **(residuos)**

- Residuos marginales $\hat{\varepsilon} = y - X\hat{\beta} \Rightarrow \varepsilon = y - \mathbb{E}(y) = y - X\beta$.
- Residuos condicionales $\hat{\epsilon} = y - X\hat{\beta} - Z\hat{b} \Rightarrow \epsilon = y - \mathbb{E}(y|b) = y - X\beta - Zb$.
- Residuos de los efectos aleatorios: $Zb = \mathbb{E}(y|b) - \mathbb{E}(y)$.

Validación del modelo

Análisis de lo supuestos **(residuos)**

Planteamiento de la hipótesis
Tipo de error

Tipo de gráfica

Marginal
($\varepsilon = y - E(y) = y - X\beta$)
Relación con la distribución de la variable dependiente
Relación de linealidad de los efectos fijos

⇒ Histogramas, gráficos de la función de densidad
ε vs variables independientes

Condicional
($e = y - E(y|b) = y - X\beta - Zb$)

Los errores dentro del grupo son independientes e idénticamente distribuidos normalmente, con media cero y varianza σ^2
(Homocedásticos e insesgados)

⇒ ε vs valores ajustados

⇒ Inspección visual
⇓
Docimar la hipótesis
⇓
Conclusiones

Aleatorios
($Zb = E(y|b) - E(y)$)

Normalidad de los efectos aleatorios

⇒ Gráficos Q-Q

Demidenko y Stukel 2005; Buja, *et al*, 2009.

Validación del modelo

Datos atipicos

- Asimetría y curtosis por el método de lo momentos.

$$C = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} w_i^2 \left(\frac{x_i - \bar{x}_w}{s_w} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

$$S = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} w_i^{\frac{3}{2}} \left(\frac{x_i - \bar{x}_w}{s_w} \right)^3$$

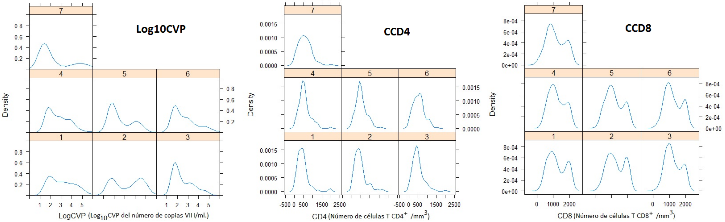
Donde:
x_i, es el valor de la i-ésima observación.
n es el número de observaciones.
s_w, es la media de las n observaciones.
s_w, es la desviación estándar de las n observaciones.
w_i, es el peso de la i-ésima observación.

- Datos atípicos por mayor diferencia, paquete *outliers* del *software R*.

Wendell y Refior, 1990; Komsta, 2015

Asimetría y curtosis por el método de lo momentos

Datos atipicos (ejemplo)



- Asimetría (S) y curtosis (C) diferentes a 0, con asimetría positiva en el Log_{10}CVP y el CCD4; además de presentar una distribución de valores con respecto a los centrales, platicúrticos para el Log_{10}CVP y leptocúrticos para el CCD4.
- Datos atípicos o extremos (8 % Log_{10}CVP, $<$ 1 % CCD4 y CCD8.)
- Alta heterogeneidad, dado elevados valores de σ^2 .
- Correlación significativa entre las observaciones en el tiempo del Log_{10}CVP y el CCD4 que aumenta con el tiempo.

Mellors *et al.*, 1996; Mellors, Muñoz y Giorgi, 1997; Lachos, Bandyopadhyay, Dey, 2011; De Andrade, De Oliveira y Geraldo,

Diagnóstico

Manejo de los datos faltantes

Siendo Y el vector de observaciones, se establece la partición $Y_i = (Y_i^{(o)}; Y_i^{(m)})$, donde $Y_i^{(o)}$ representa los datos observados y $Y_i^{(m)}$ representa los datos no observados. Se defne $R_i = (R_{i1}, \ldots, R_{in_i})$ como el indicador de los índices de los datos faltantes.

$$R_{ij} = \begin{cases} 1 \text{ si } Y_{ij} \text{ (es dato observado)} \\ \\ 0 \text{ si } Y_{ij} \text{ (es dato no observado)} \end{cases}$$

Diagnóstico

Manejo de los datos faltantes

Los mecanismos de datos faltantes son los siguientes

- Los datos fatantes son completamente aletorios (MCAR)

$$P(R_i|Y_i^o, Y_i^m, X_i) = P(R_i)$$

- Los datos faltantes son aleatorios (MAR).

$$P(R_i|Y_i^o, Y_i^m, X_i) = P(R_i|Y_i^o, X_i)$$

- Los datos faltantes no son aleatorios (NMAR)

$$P(R_i|Y_i^o, Y_i^m, X_i) = P(R_i|Y_i^m, X_i)$$

Rubin DB. Inference and missing data. Biometrika. 63(3). 1976. 581-592.

Diagnóstico del modelo

Análisis de influencia

- Distancia de verosimilitud; $LD_{(U)}$, $RLD_{(U)}$

$$LD_{(U)} = 2\left\{ l(\hat{\psi}) - l(\hat{\psi}_{(U)}) \right\}; RLD_{(U)} = 2\left\{ l_R(\hat{\psi}) - l_R(\hat{\psi}_{(U)}) \right\}$$

  Donde (U) son las estimaciones sin un conjunto de observaciones y ψ es el cálculo de las estimaciones.

- DFBETA

$$DFBETA = \frac{\hat{\gamma}_i - \hat{\gamma}_{i(-j)}}{se(\hat{\gamma}_{i(-j)})}$$

  Donde $\hat{\gamma}_i$ es la estimación de parámetro con datos completos i y $\hat{\gamma}_{i(-j)}$ la estimación del parámetro i sin la observación o grupo j.

Diagnóstico del modelo

Análisis de influencia

- La distancia de Cook's y MDFFITS

$$D(\beta) = \frac{1}{Rango[X]}\left((\hat{\beta} - \hat{\beta}_{(U)})' var(\hat{\beta})^{-1}(\hat{\beta} - \hat{\beta}_{(U)})\right)$$

$$MDFFITS(\beta) = \frac{1}{Rango[X]}\left((\hat{\beta} - \hat{\beta}_{(U)})' var(\widehat{\beta_{(U)}})^{-1}(\hat{\beta} - \hat{\beta}_{(U)})\right)$$

- Radio de varianza de información, mide la variación en el determinante de la matriz de información en la estimación.

$$IR(\gamma) = \frac{det(Inf(\hat{\psi}_{(i)}))}{det(Inf(\hat{\psi}))}$$

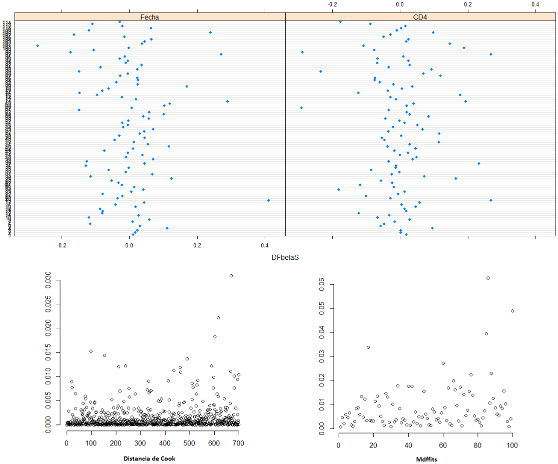
$$IR(\beta) = \frac{det(var(\hat{\psi}_{(i)}))}{det(var(\hat{\psi}))}$$

Diagnóstico

Influencia - Datos completos. (Ejemplo)

Evaluación de la influencia, medidas locales y globales: Distancia de Cook's, DFBeta

Análisis de influencia, para el modelo ajustado para datos con variable dependiente con distribución normal.



Bibliografía recomendada.

- Douglas C. Montgomery. Design and analysis of experiments. 3 Edition. 1991.
- Cocharan W.G. and Cox G.M. Experimental desing. 2 Edition. 1957 (May 1992).
- Hicks, Charles R., Turner Jr., Kenneth V. Fundamental Concepts in the Design of Experiments. 5 Edition. 1973 (1983)
- Hinkelmann K. Kempthorne D. Design and analysis of experiments . 1988. (2007)