



ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

Τμήμα Γεωγραφίας

Πρόγραμμα Μεταπτυχιακών Σπουδών "Εφαρμοσμένη

Γεωγραφία και Διαχείριση του Χώρου"

Μάθημα "Εφαρμογές Γεωπληροφορικής στη Διαχείριση

Καταστροφών"

ΕΚΤΙΜΗΣΗ ΕΠΙΔΕΚΤΙΚΟΤΗΤΑΣ ΚΑΤΟΛΙΣΘΗΣΕΩΝ ΜΕ ΤΗ ΧΡΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ "ΛΟΓΙΣΤΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ "

Χρήστος Πολυκρέτης, Υποψ. Διδάκτορας

e-mail: chpolycretis@hua.gr

Αθήνα, Μάιος 2017

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

- Η λογιστική παλινδρόμηση (logistic regression – LR) περιλαμβάνει το σχηματισμό μιας πολυμεταβλητής συσχέτισης μεταξύ μιας εξαρτημένης μεταβλητής και πολλών ανεξάρτητων μεταβλητών (Atkinson and Massari, 1998).
- Βασική αρχή της είναι ότι η εξαρτημένη μεταβλητή (dependent variable) είναι διχοτομική, δηλαδή μπορεί να πάρει μόνο τιμές 1 και 0, ή «αληθής» και «ψευδής».
- Οι ανεξάρτητες μεταβλητές (independent variables) προβλέπουν την εξαρτημένη και μπορεί να είναι είτε συνεχείς, είτε διακριτές, είτε οποιοσδήποτε συνδυασμός των δύο αυτών τύπων.

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

- Στην εκτίμηση της επιδεκτικότητας κατολισθήσεων, στόχος της LR είναι να βρει το «βέλτιστο» μοντέλο για να περιγράψει τη σχέση μεταξύ της απουσίας ή παρουσίας (τιμή 0 ή 1) κατολισθήσεων (εξαρτημένη μεταβλητή) και ενός συνόλου παραγόντων που επιδρούν στην εκδήλωση του φαινομένου (ανεξάρτητες μεταβλητές).
- Το μοντέλο μπορεί να εκφραστεί στην απλούστερη μορφή του ως εξής:

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

$$P = \frac{1}{1 + e^{-z}}$$

όπου P είναι η πιθανότητα εκδήλωσης ενός κατολισθητικού γεγονότος η οποία κυμαίνεται από 0 έως 1.

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

n = ο αριθμός των ανεξάρτητων μεταβλητών

x_i ($i = 1, 2, \dots, n$) = οι ανεξάρτητες μεταβλητές

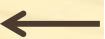
b_0 = η σταθερά του μοντέλου

b_i ($i = 1, 2, \dots, n$) = οι συντελεστές βαρύτητας (coefficients)

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

- Ένας άλλος τρόπος γραφής είναι ο λογαριθμικός μετασχηματισμός (logit transformation):

$$\text{Logit}(p) = \ln \left(\frac{p(Y=1)}{1-p(Y=1)} \right) = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$



Λόγος πιθανοφάνειας (likelihood ratio) όπου αντιπροσωπεύει την πιθανότητα ένα γεγονός να συμβεί διαιρούμενο με την πιθανότητα να μη συμβεί.

- Συνεπώς ένας παράγοντας επηρεάζει σημαντικά την εκδήλωση του φαινομένου, εάν ο λόγος πιθανοφάνειας του είναι >1 και η στατιστική σπουδαιότητα (significance) του είναι μέσα στα όρια ενός δεδομένου διαστήματος εμπιστοσύνης.

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

- Δύο μορφές εφαρμογής του μοντέλου:

1) Τυπική (enter):

Μέσω της μεθόδου μέγιστης πιθανοφάνειας (*maximum likelihood* ή $-2LL$), το μοντέλο υπολογίζει τους συντελεστές και τα στατιστικά μέτρα απόδοσης για το σύνολο των ανεξάρτητων μεταβλητών (Kundu et al., 2013).

2) Βηματική (stepwise):

α) "Προς τα εμπρός (forward)" :

- Στην αρχή **καμία** μεταβλητή δεν περιέχεται στο μοντέλο.
- Εάν η τιμή *significance* μιας μεταβλητής είναι **μικρότερη** από την τιμή *πιθανότητας (probability)* για παραμονή στο μοντέλο, η μεταβλητή αυτή εισέρχεται στο μοντέλο και τα στατιστικά υπολογίζονται εκ νέου έτσι ώστε να εξεταστεί η είσοδος και των υπολοίπων μεταβλητών (Dai and Lee, 2003; Mousavi et al., 2011).

β) "Προς τα πίσω (backward)" :

- Στην αρχή **όλες** οι μεταβλητές περιέχονται στο μοντέλο.
- Η ίδια διαδικασία με την "forward" για κάθε έξοδο μιας μεταβλητής.

ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

- Πριν την εφαρμογή του μοντέλου, είναι απαραίτητη η διάγνωση της πολυσυγγραμμικότητας (multicollinearity) για να ελεγχθεί η συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών.
- Ο πιο διαδεδομένος τρόπος διάγνωσης είναι ο υπολογισμός δύο δεικτών:

Ανοχή (tolerance – TOL): μια τιμή μικρότερη από 0,2 υποδεικνύει πολυσυγγραμμικότητα (Menard, 2001).

Συντελεστής πληθωριστικής διακύμανσης (variance inflation factor – VIF): μια τιμή μεγαλύτερη από 10 υποδηλώνει την ύπαρξη ενός σοβαρού προβλήματος πολυσυγγραμμικότητας.

ΠΡΑΚΤΙΚΟ ΠΛΑΙΣΙΟ

- Βήματα εφαρμογής του μοντέλου:

- 1) Δειγματοληψία κατολισθητικών δεδομένων (εξαρτημένη μεταβλητή)
- 2) Επεξεργασία παραγόντων (ανεξάρτητες μεταβλητές)
- 3) Αντιστοίχιση training dataset με τις κατηγορίες των "επεξεργασμένων" παραγόντων.
- 4) Εισαγωγή συγκεκρινητικού πίνακα σε λογισμικό στατιστικό αναλύσεων.
- 5) Υπολογισμός συντελεστών βαρύτητας για τους παράγοντες (μέσω εξίσωσης Z) και στατιστικών μέτρων απόδοσης του μοντέλου.
- 6) Υπολογισμός της πιθανότητας εκδήλωσης του φαινομένου (μέσω εξίσωσης P)
 - δημιουργία τελικού επιπέδου επιδεκτικότητας κατολισθήσεων.

ΕΝΔΕΙΞΗ ΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- ❖ Atkinson PM, Massari R (1998) Generalized linear modeling of susceptibility to landsliding in the central Apennines, Italy. Computers and Geosciences, 24, 373–385
- ❖ Dai FC, Lee CF (2003) A spatiotemporal probabilistic modelling of storm-induced shallow landsliding using aerial photographs and logistic regression. Earth Surface Processes and Landforms, 28, 527–545
- ❖ Kundu S, Saha AK, Sharma DC, Pant CC (2013) Remote Sensing and GIS Based Landslide Susceptibility Assessment using Binary Logistic Regression Model: A Case Study in the Ganeshganga Watershed, Himalayas. Journal of the Indian Society of Remote Sensing, 41(3), 697–709
- ❖ Menard SW (1995) Applied logistic regression analysis. SAGE, Thousand Oaks
- ❖ Mousavi SZ, Kaviani A, Soleimani K, Mousavi SR, Shirzadi A (2011) GIS-based spatial prediction of landslide susceptibility using logistic regression model. Geomatics, Natural Hazards and Risk, 2(1), 33–50