**2η Εργασία Γραμμικά Γενικευμένα Γραμμικά Μοντέλα 2019-2020**

1. Χρησιμοποιώντας την R, προσομοιώστε δεδομένα από ένα απλό μοντέλο τυχαίων συντελεστών (τυχαίων ευθειών) για 50 άτομα (subjects) και δέκα συνεχείς ημέρες. Το μέσο (πληθυσμιακό) intercept να είναι ίσο με 50 και η (πληθυσμιακή) μέση κλίση να είναι ίση με 1. Οι διακυμάνσεις να είναι: για τα (subject-specific) intercepts,  1, για τις (subject-specific) κλίσεις, 1 , και για το σφάλμα παρατήρησης , 16. Η συν-διακύμανση μεταξύ των intercept και των κλίσεων να είναι ίση με, = -0.05. Χρησιμοποιώντας τις κατάλληλες εντολές στην R, προσαρμόσετε το μοντέλο και εκτιμήσετε τις παραμέτρους. Χρησιμοποιώντας το output κάνετε ελέγχο της υπόθεση έλλειψης χρονικής επίδρασης στην μέση απόκριση. Δώστε 95% δ.ε. για: το μέσο (πληθυσμιακό) intercept, τη (πληθυσμιακή) μέση κλίση και την μέση (πληθυσμιακή) τιμή της απόκρισης την 5η ημέρα.
2. (Συνέχεια της προηγούμενης) Προσομοιώστε δεδομένα από ένα μοντέλο τυχαίων συντελεστών για 100 άτομα και δέκα ημέρες. Τα 50 να είναι στο ‘placebo’ group και τα υπόλοιπα στο ‘drug group’. Το μέσο intercept για ‘placebo’ group να είναι 50 και η μέση κλίση για ‘placebo’ group να είναι 1. Το μέσο intercept για ‘drug’ group να είναι 50 και η μέση κλίση για ‘drug’ group να είναι 1.25. Οι διακυμάνσεις να είναι: για τα intercept 1, για τις κλίσεις 1 , για το σφάλμα παρατήρησης 16 και η συνδιακύμανση μεταξύ intercept και κλίσης να είναι -0.05. Χρησιμοποιώντας τις κατάλληλες εντολές στην R, προσαρμόσετε το μοντέλο και εκτιμήσετε τις παραμέτρους. Επίσης, προσαρμόσετε ένα μοντέλο που να μη λαμβάνει καθόλου υπόψη το φάρμακο (θεραπεία). Χρησιμοποιώντας τα output

Α) Ελέγξτε την υπόθεση που αναφέρει πως το φάρμακο δεν επηρεάζει την μέση απόκριση.

Β) Ελέγξτε την υπόθεση που αναφέρει πως το φάρμακο δεν επηρεάζει την χρονική πορεία της μέσης απόκρισης.

Γ) Δώστε 95% δ.ε. για την διαφορά των κλίσεων στις 2 ομάδες.

Δ) Δώστε 95% δ.ε. για την διαφορά της μέσης απόκρισης μεταξύ των 2 ομάδων κατά την 5η ημέρα.

1. Θεωρήστε πως η μέση τιμή ημερήσιων θανάτων σε μια χώρα από μία πανδημία ακολουθεί το μοντέλο:  *, ,* όπου αναφέρεται στη ημέρα, και είναι η πυκνότητα πιθανότητας μιας τυπικής κανονικής κατανομής. Για τους παρατηρούμενους ημερήσιους αριθμούς θανάτων, , υποθέσετε πως είναι ανεξάρτητοι μεταξύ τους και ακολουθούν την , κατανομή.

Α) Δείξτε πως το παραπάνω μοντέλο είναι ένα γενικευμένο γραμμικό μοντέλο.

Β) Σε ποια ημέρα αναμένουμε τους περισσότερους θανάτους? Ποια η αναμενόμενη τιμή των θανάτων σε αυτή την μέρα?

Γ) Χρησιμοποιώντας την R, προσομοιώσετε τιμές ημερησίων θανάτων, για

Δ) Προσαρμόστε το κατάλληλο γενικευμένο γραμμικό μοντέλο στα δεδομένα. Εκτιμήσετε την αναμενόμενη τιμή ημερησίων θανάτων για την 121η ημέρα και δώστε ένα 95% διάστημα εμπιστοσύνης για αυτήν. Σχολιάστε!

1. Θεωρήστε πως η τιμή μιας αιματολογικής εξέτασης, , είναι ενδεικτική μιας ασθένειας (μεγαλύτερη τιμή σημαίνει ‘περισσότερη’ ένδειξη ασθένειας).

Έστω , , η τιμή της αιματολογικής εξέτασης για το i-οστό άτομο και η δείκτρια ασθένειας (=1 αν το i-οστό άτομο βρέθηκε, μετά από βιοψία, ότι είχε τελικά ασθένεια ). Εδώ είναι η απόκριση, ενώ είναι η συμμεταβλητή.

Α) Γράψτε ένα απλό γενικευμένο γραμμικό μοντέλο για τα δεδομένα.

Β) Ποιες είναι οι τιμές των παραμέτρων του μοντέλου σας αν δίνεται πως a) τα οdds ασθένειας στην τιμή είναι ίση με , b) τα οdds ασθένειας στην τιμή είναι ίση με ?

Γ) Χρησιμοποιώντας την R, και χρησιμοποιώντας τιμές

προσομοιώσετε 200 τιμές των , με τις παραμέτρους που βρήκατε.

Χρησιμοποιώντας τα δεδομένα που γεννήσατε, εκτιμήσετε, με χρήση R, τις παραμέτρους κι ελέγξετε τη στατιστική σημαντικότητα της παλινδρόμησης.

**Σημείωση:** Στις προσομοιώσεις σας ξεκινάτε με seed= ημερομηνία γέννησής σας και προσθέτετε +1 για την επόμενη.

**BONUS**

**Στις Ασκήσεις 1 και 2 βρείτε τις εκτιμήσεις των παραμέτρων μεγιστοποιώντας απευθείας τις πιθανοφάνειες (αφού πρώτα τις ορίσετε στη R, σύμφωνα με τις εξισώσεις 1.23-1.32 στις σημειώσεις.**