

ΔΠΜΣ Βιοστατιστική

και Επιστήμη Δεδομένων Υγείας

Πέτρος Τζαβέλλας

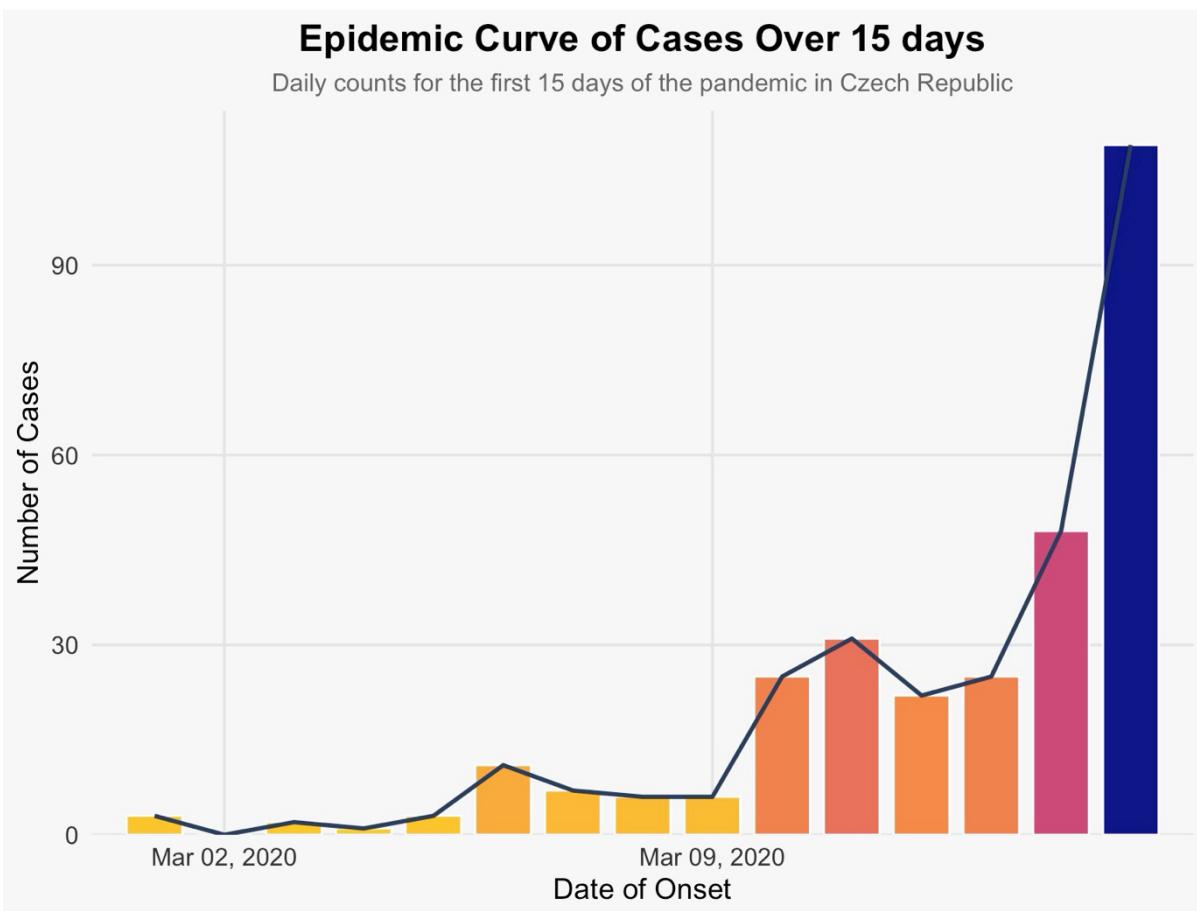
ΑΜ: 7450022400026

Εργασία στα Μαθηματικά Μοντέλα Λοιμωδών Νοσημάτων

Σημείωση: Παρότι θα ανεβάσω μαζί με το παρόν αρχείο και τον κώδικα, το excel με τα δεδομένα, επέλεξα να τα διαχειριστώ απευθείας μέσα από το RStudio.

Ερώτημα 1ο:

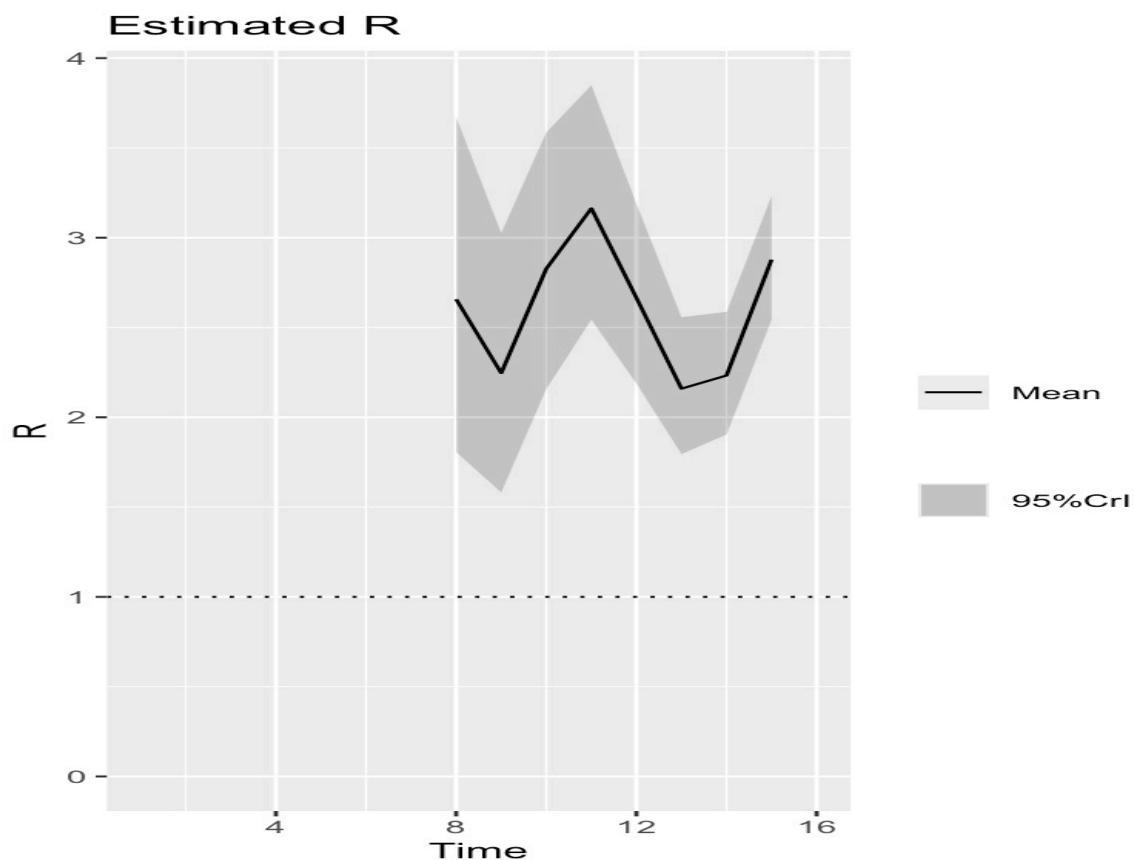
Στην επιδημική καμπύλη της Τσεχίας, από την πρώτο κρούσμα, το οποίο συνέβη την 01/03/2020, μέχρι και της 15/03/2020 παρατηρούμε μια σταδιακή αύξηση των κρουσμάτων, μέχρι την απότομη αύξηση την 15^η μέρα της πανδημίας.



Ερώτημα 2ο:

Αρχικά για χρόνο γενιάς (serial interval) πήρα τις 4 μέρες κατά μέσο όρο, μετά από σχετική αναζήτηση σε αρκετές πηγές που αναφέρονται παρακάτω, το οποίο συμφωνεί με το αντίστοιχο που δόθηκε στο μάθημα, ωστόσο για τυπική απόκλιση, θεώρησα το 2.8 μέρες που χρησιμοποιήθηκε και στο ανάλογο εργαστήριο του μαθήματος.

[<https://www.ecdc.europa.eu/sites/default/files/documents/RRA-seventh-update-Outbreak-of-coronavirus-disease-COVID-19.pdf>]



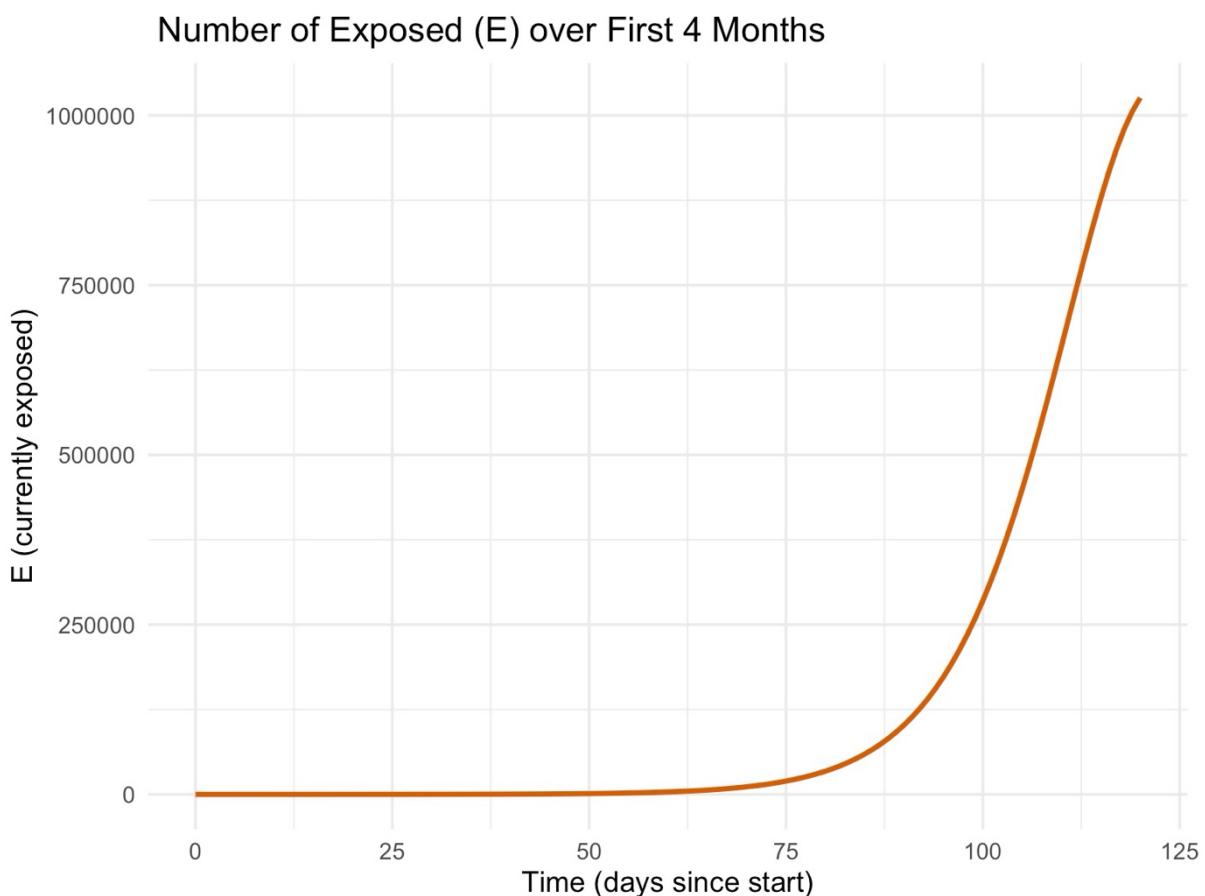
Το εκτιμώμενο R_t παραμένει σαφώς πάνω από το 1 σε όλη την περίοδο, με μεταβολές από περίπου 2,3 έως 3,2. Αυτό δηλώνει ότι τα κρούσματα είναι σε αυξητική πορεία και ότι η πανδημία αυξάνεται και δεν ενδέχεται να σταματήσει.

Ερώτημα 3ο:

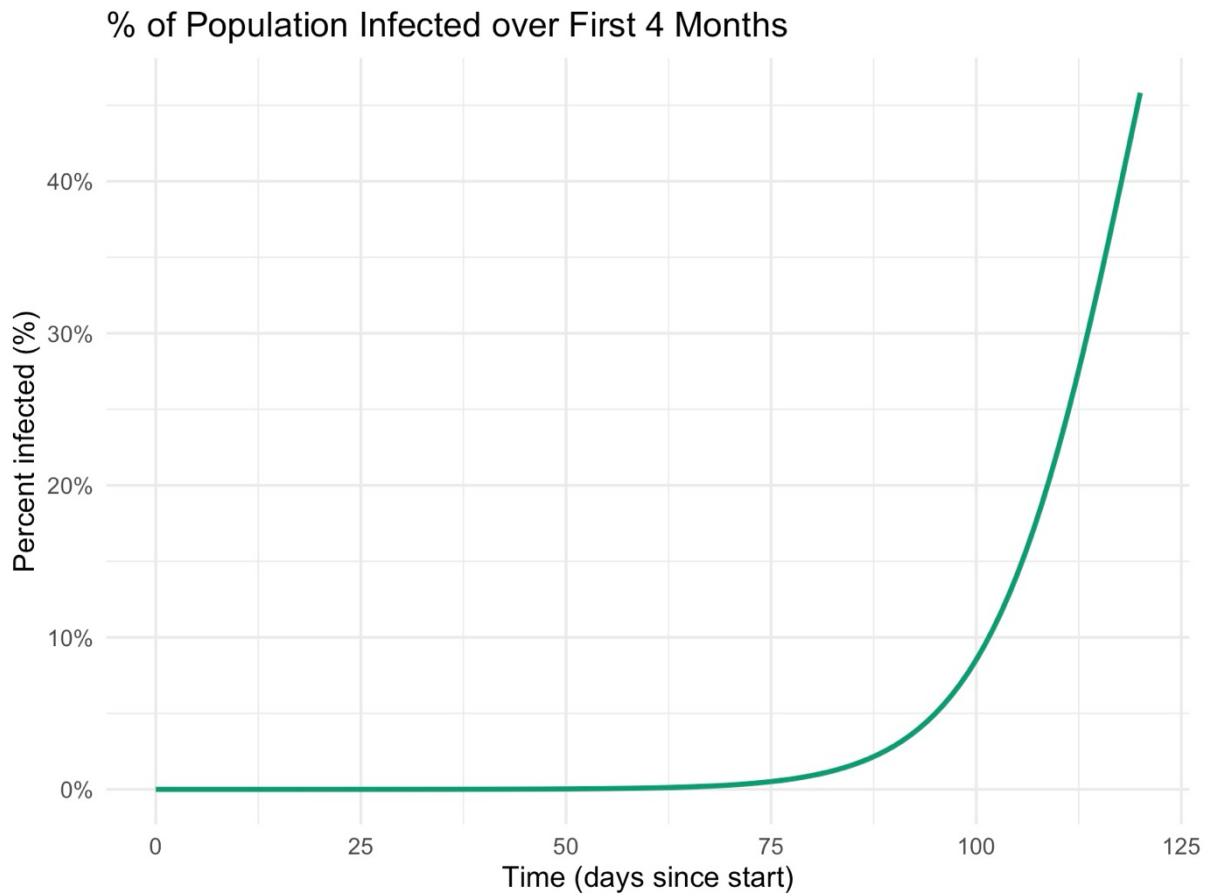
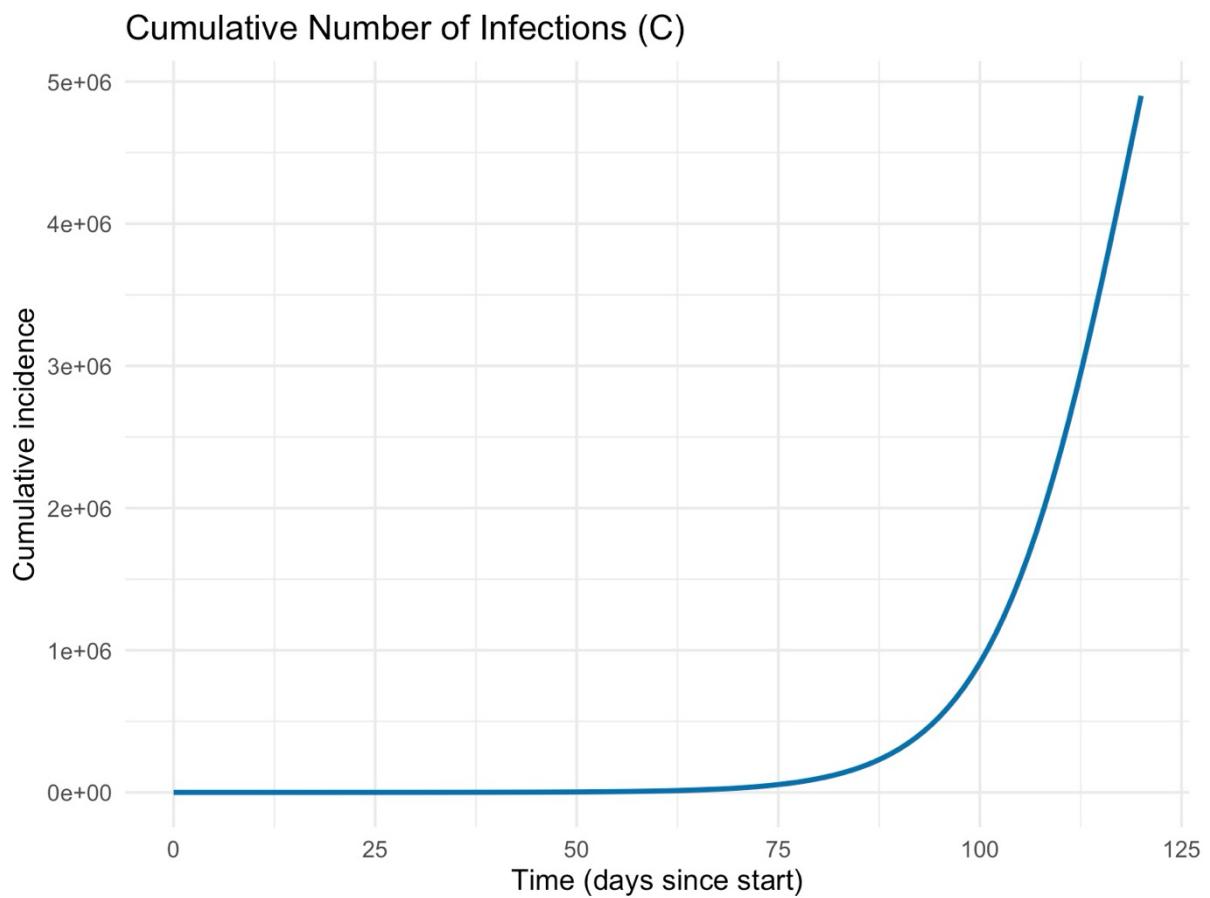
A)

R_0	2.6	S	10693929
β	0.371	E	10
γ	0.14	I	0
σ	0.25	R	0

B) Ο συνολικός αριθμός μολυσμένων (E)στην διάρκεια των 4 πρώτων μηνών αυξάνεται συνεχώς και 120^η μέρα της πανδημίας είναι 1,087,511 άτομα που είναι περίπου το 10% του συνολικού πληθυσμού.



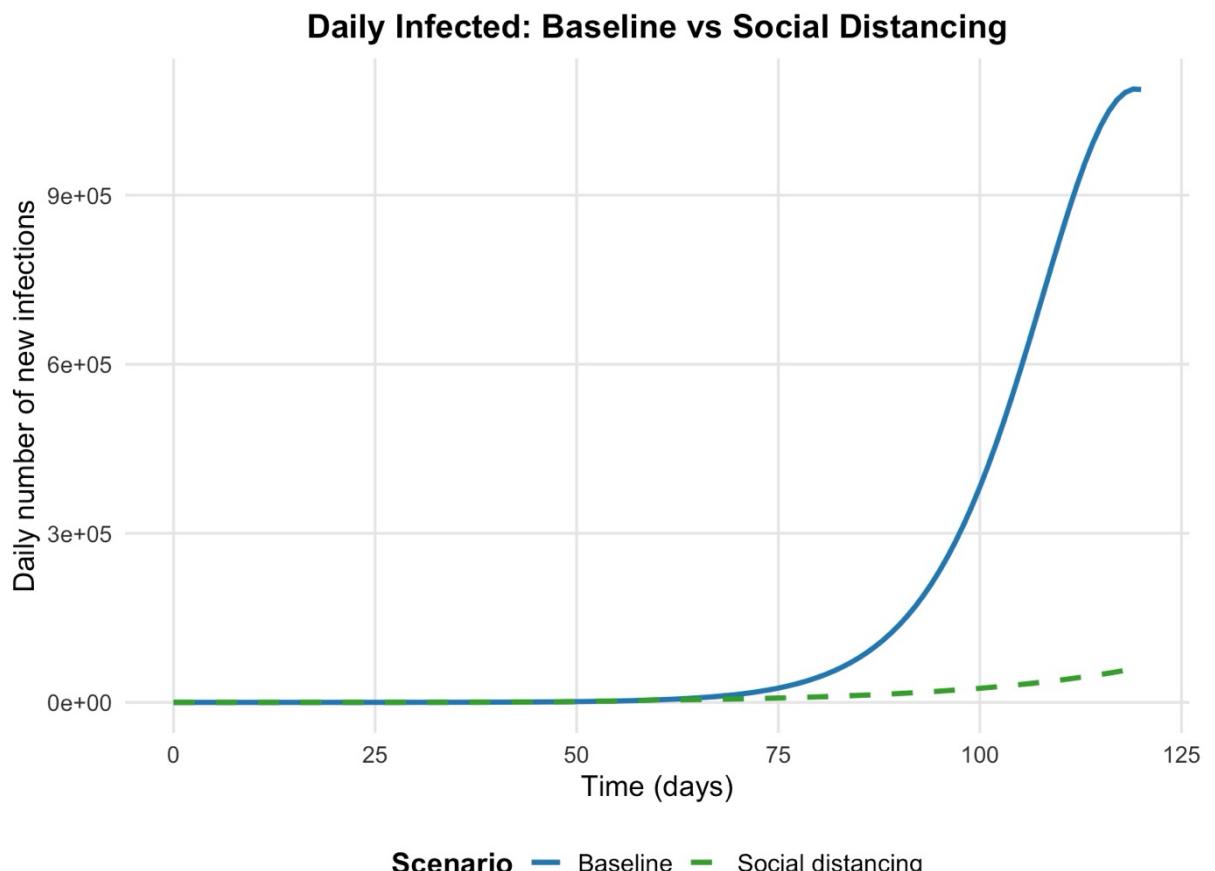
Γ) Ο αθροιστικός αριθμός μολυσμένων ατόμων στο διάστημα των 4 μηνών της πανδημίας είναι 4,899,911, δηλαδή έχει μολυνθεί το 45,81% του συνολικού πληθυσμού της Τσεχίας. Δηλαδή, σε μόλις 4 μήνες, άμα δεν παρθούν μέτρα κοινωνικής αποστασιοποίησης και άλλου τύπου μέτρα, περίπου το 45% του συνολικού πληθυσμού θα έχει μολυνθεί με Covid-19.



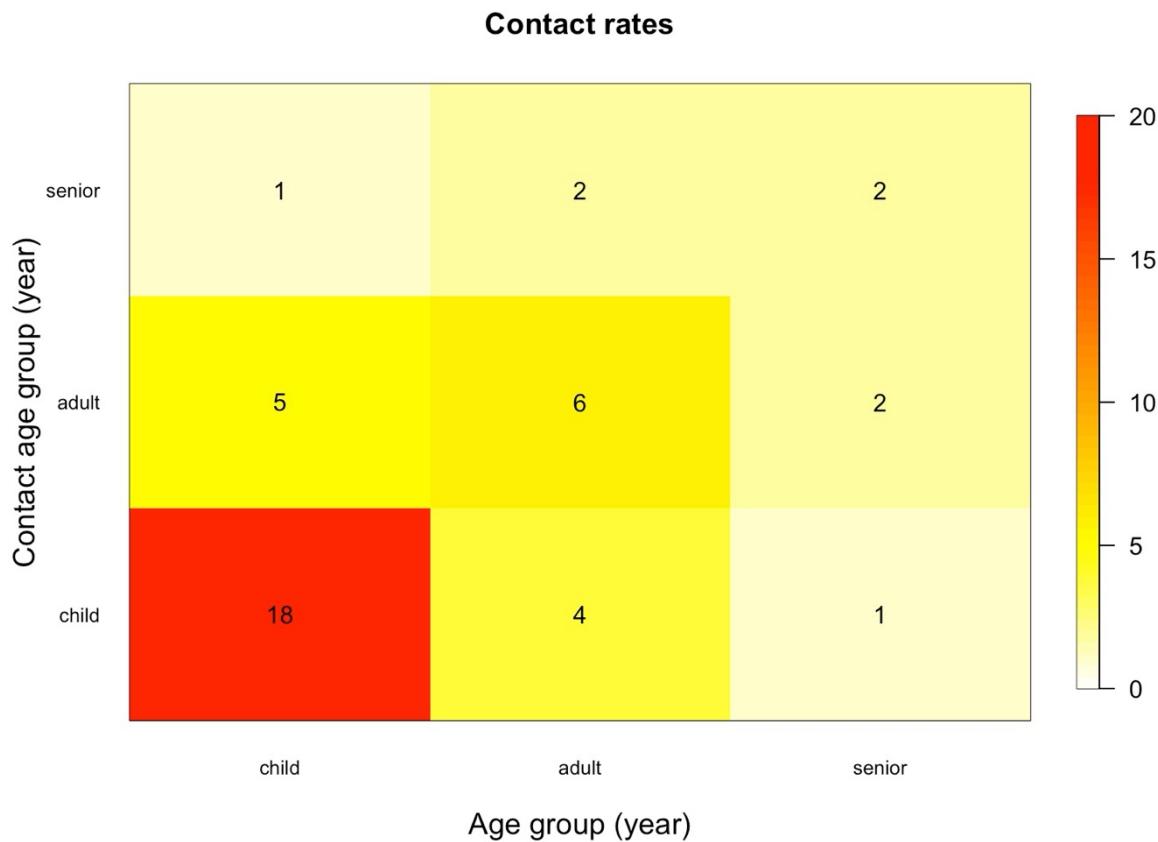
Δ) Έτσι, σε μία ανεξέλεγκτη πανδημία του Covid-19, το τελικό μέγεθος της πανδημίας (δηλαδή ο συνολικός αριθμός των ατόμων που θα έχουν προσβληθεί από τον ιό, μέχρις ότου η πανδημία να εκλείψει) θα είναι 9,747,326 , που αντιστοιχεί στο 0.91% του συνολικού πληθυσμού.

Ερώτημα 4º:

Παρατηρούμε ότι μετά την θέσπιση μέτρων κοινωνικής αποστασιοποίησης ο αριθμός των μολυσμένων (E) ανά μέρα, στο ίδιο διάστημα των 2 επόμενων μηνών είναι αρκετά μικρότερος σε σχέση με τον αντίστοιχο αριθμό μολυσμένων, χωρίς να έχουν ληφθεί μέτρα, κάτι που είναι αναμενόμενο, αφού μειώθηκε ο ρυθμός πραγματοποίησης των επαφών κατά 40% μετά τους 2 πρώτους μήνες. Επίσης υπάρχει μικρή αύξηση μεταξύ δύο συνεχόμενων ημερών στο δεύτερο σενάριο, σε σχέση με το βασικό σενάριο, όπου ο αριθμός των μολυσμένων μέρα με την μέρα έχει ραγδαία αύξηση.



Ερώτημα 5ο:

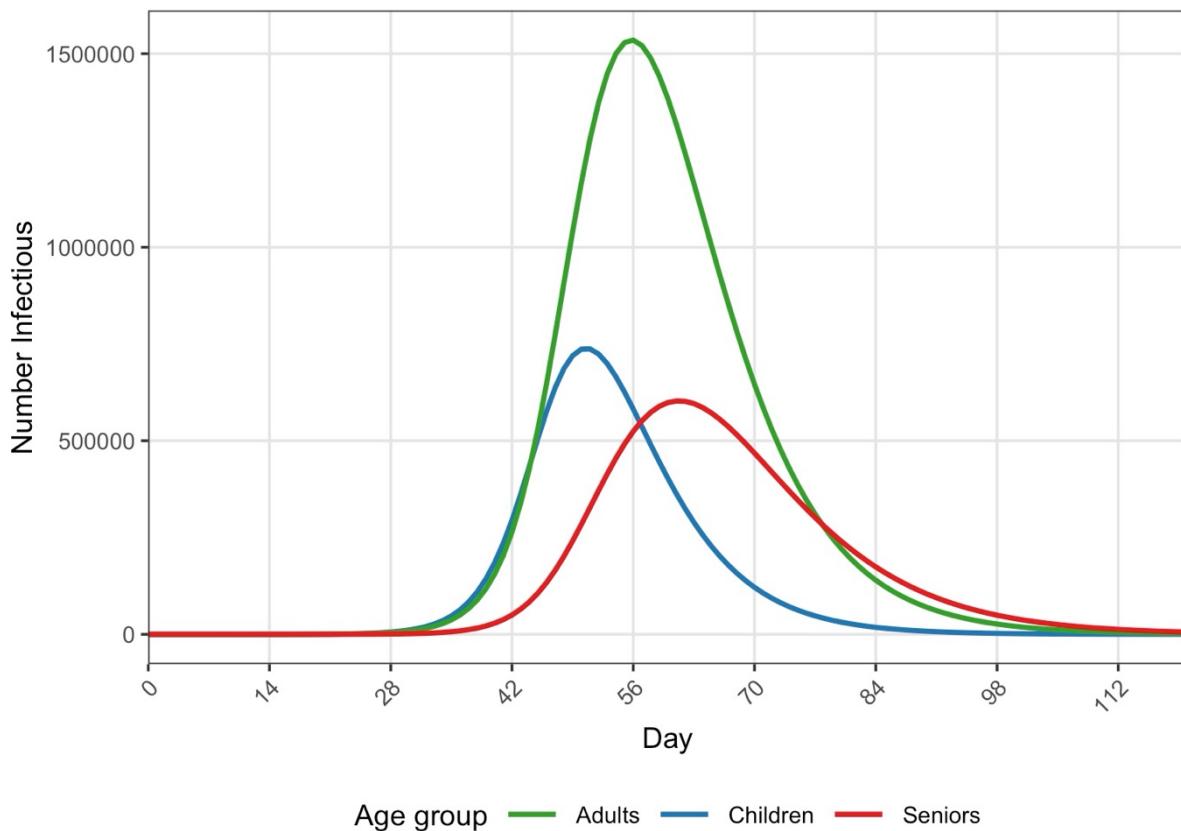


Τα παιδιά παρότι αποτελούν το μικρότερο ποσοστό του πληθυσμού, μόλις 20%, επειδή κάνουν καθημερινά 24 επαφές, εκ των οποίων οι 18 μεταξύ τους, μολύνονται γρήγορα και φτάνουν το μέγιστο αριθμό μολυσματικών ανά μέρα περίπου την μέρα 50 και μετά σιγά σιγά μειώνεται ο αριθμός αυτός.

Οι ενήλικες, παρότι έχουν μειωμένες επαφές σε σχέση με τα παιδιά (12/μέρα), επειδή αποτελούν το 50% του πληθυσμού, έχουν το κύριο μέρος της πανδημίας, έτσι υπάρχει μεγάλη αύξηση στον καθημερινό αριθμό μολυσματικών που φτάνει το μέγιστο περίπου την μέρα 56, με περίπου 1500,000 μολυσματικούς.

Τέλος, οι ηλικιωμένοι, εξαιτίας του μειωμένου αριθμού επαφών, καθώς και του μειωμένου πληθυσμού αυτών σε σχέση με τον συνολικό πληθυσμό, έχουν πιο «πλατιά καμπάνα», δηλαδή καθυστερεί η διάδοση του ιού και χρειάζεται περισσότερο χρόνο να μολύνει το σύνολο του πληθυσμού.

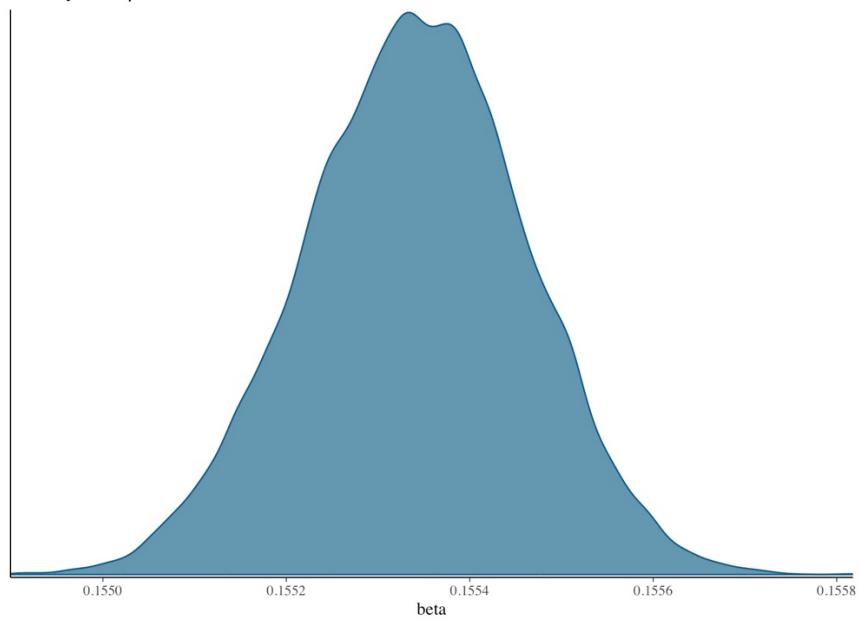
SEIR Model: Infectious over Time by Age Group



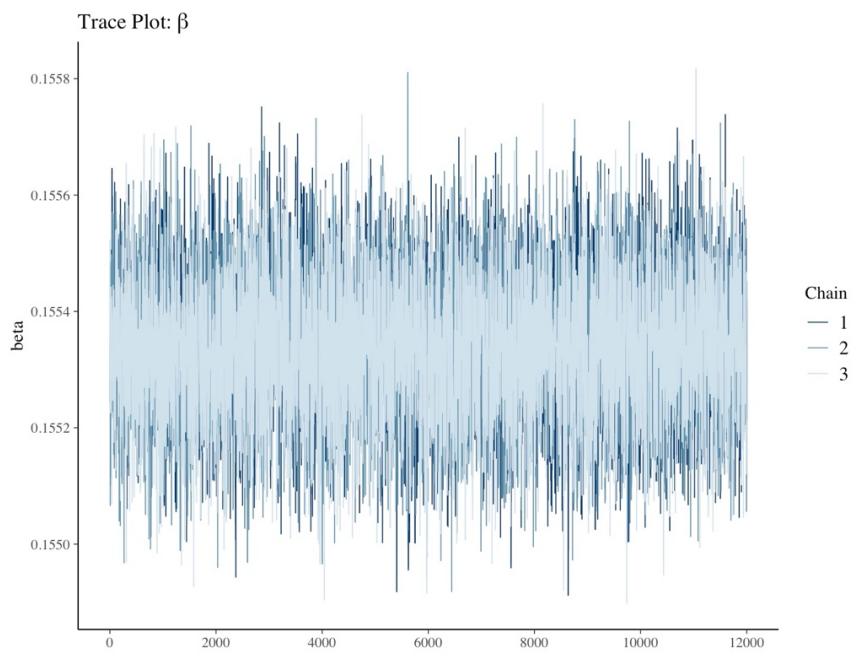
Ερώτημα 6ο:

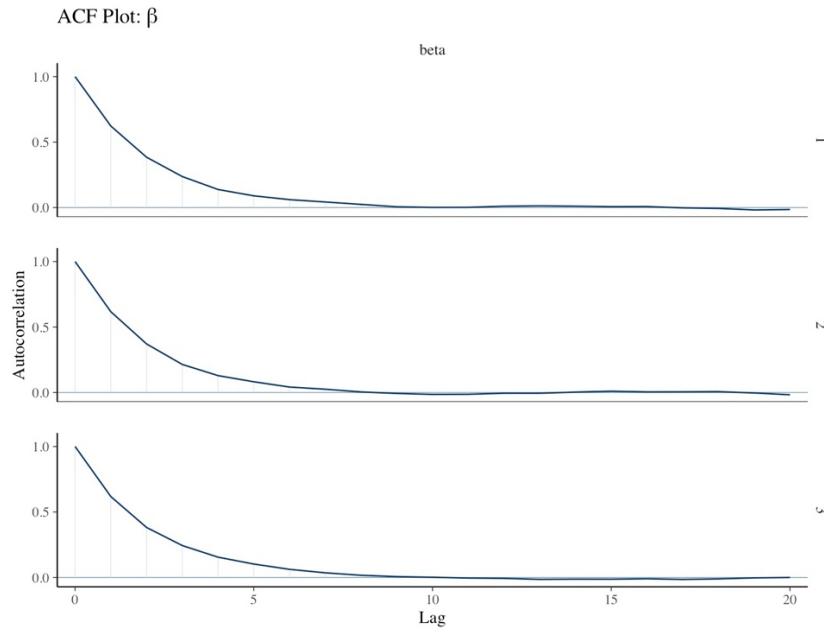
Για να τρέξω το Bayesian chain binomial χρησιμοποίησα το nimble. Τα removals τα όρισα ως εξής, θεώρησα ότι η μέση διάρκεια της μόλυνση είναι 7 ημέρες και άρα μετά από αυτήν την περίοδο, κάποιο κρούσμα θα είναι recovered (είτε θα έχει πεθάνει, είτε θα έχει αναρρώσει), άρα τα κρούσματα της i μέρας θα είναι τα removals της i+7 μέρας, για τις πρώτες 7 μέρες θεώρησα μόνο τους θανάτους που μας δίνει από τα δεδομένα. Επίσης, το β είναι 0,155. Τέλος, ο συνολικός πληθυσμός της Τσεχίας είναι 10,693,939.

Density Plot: β



Trace Plot: β

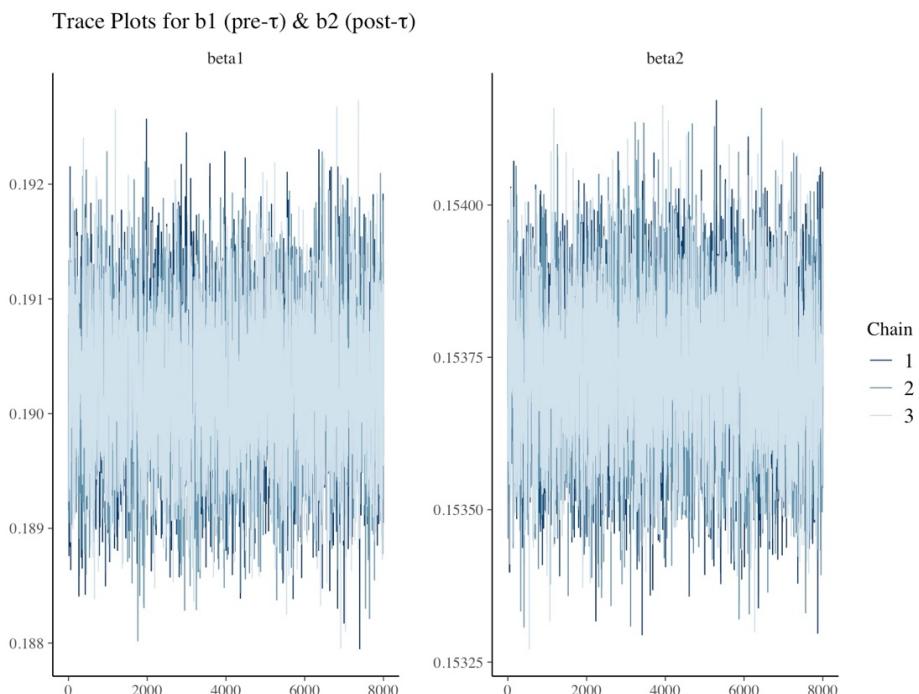




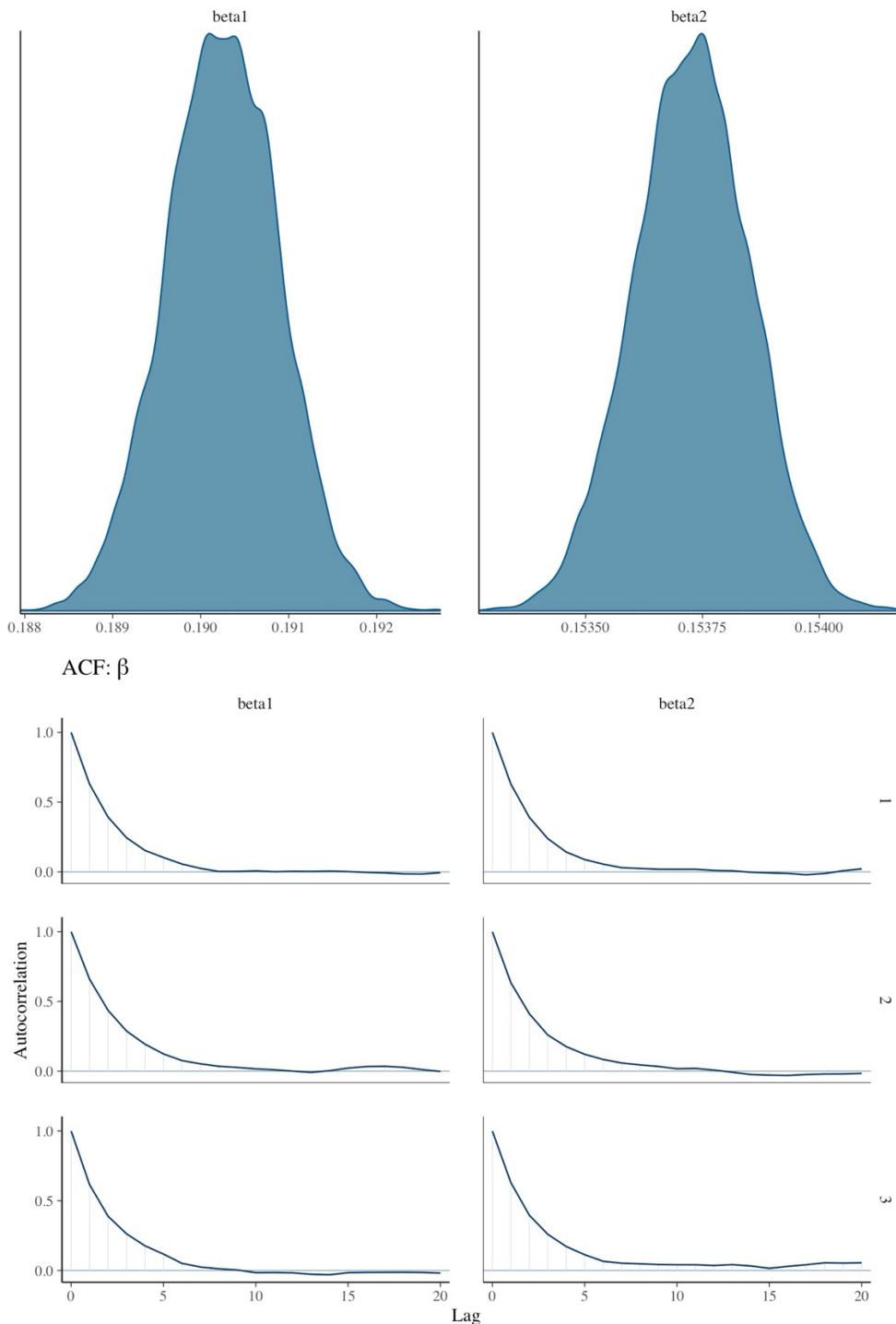
Ερώτημα 7ο:

Στην αρχή του συγκριμένου ερωτήματος έτρεξα ένα διάγραμμα των καθημερινών κρουσμάτων (cases) ανά μέρα για να δω που περίπου που παρουσιάζεται αλλαγή της μονοτονίας, για να επιλέξω το changing point.

Έτσι, επέλεξα την ημερομηνία 10/10/2020, ως σημείο αλλαγής του ρυθμού μετάδοσης της νόσου, έτσι το β_1 από 0.19 που ήταν έπεισε στο β_2 σε 0.154. Η μείωση του σημαίνει ότι μειώνεται η εξάπλωση του ιού.



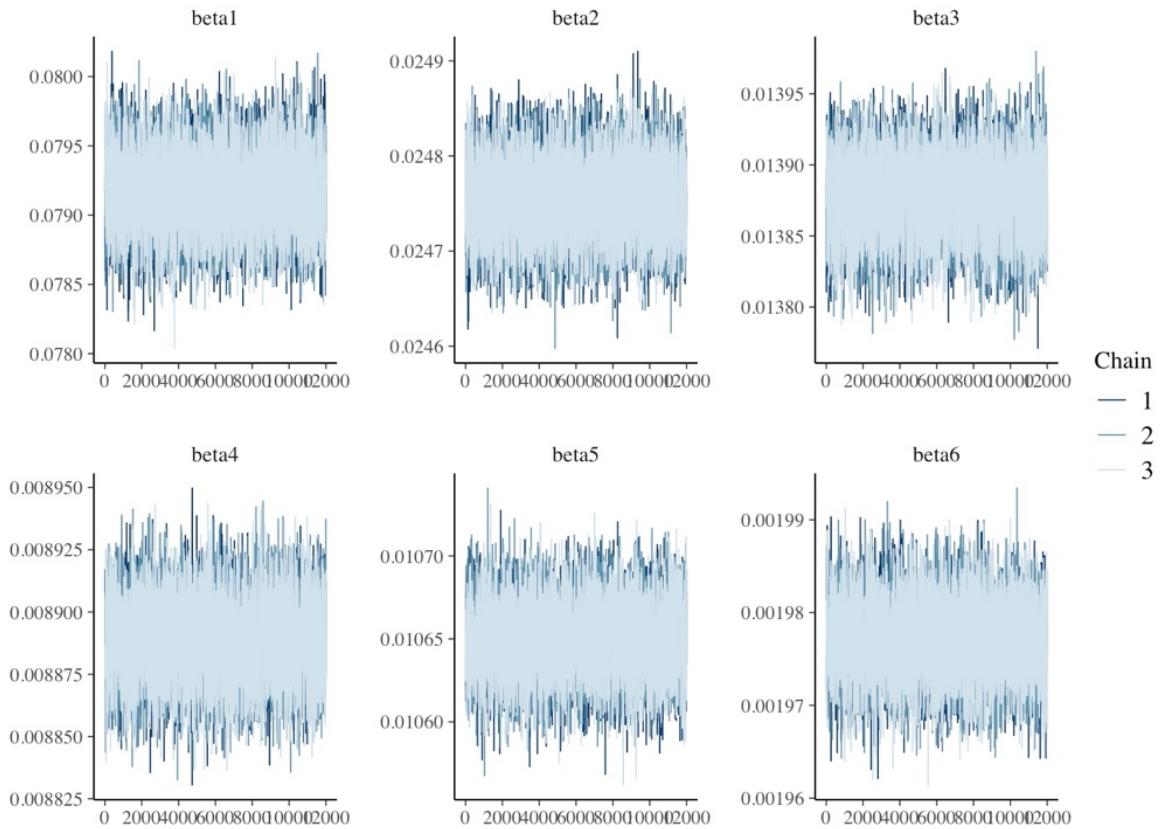
Posterior Densities for $b1$ (pre- τ) & $b2$ (post- τ)



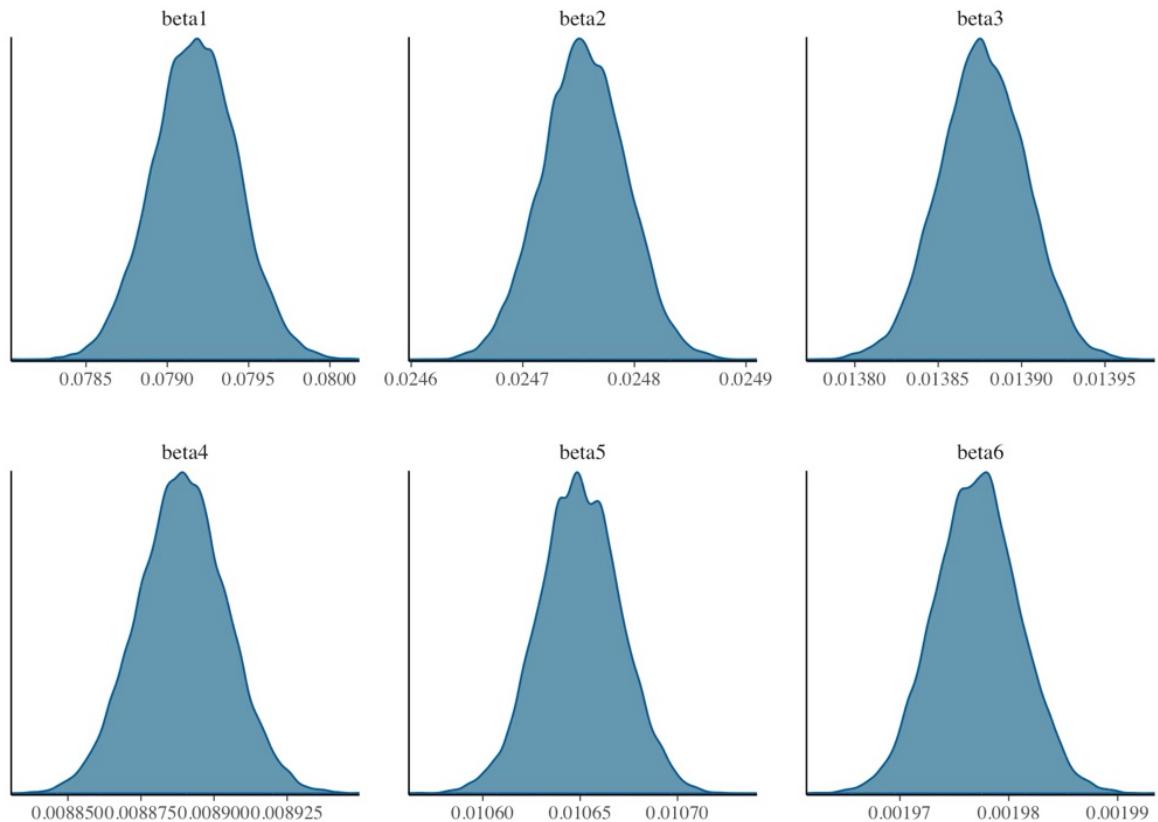
Ερώτημα 8ο:

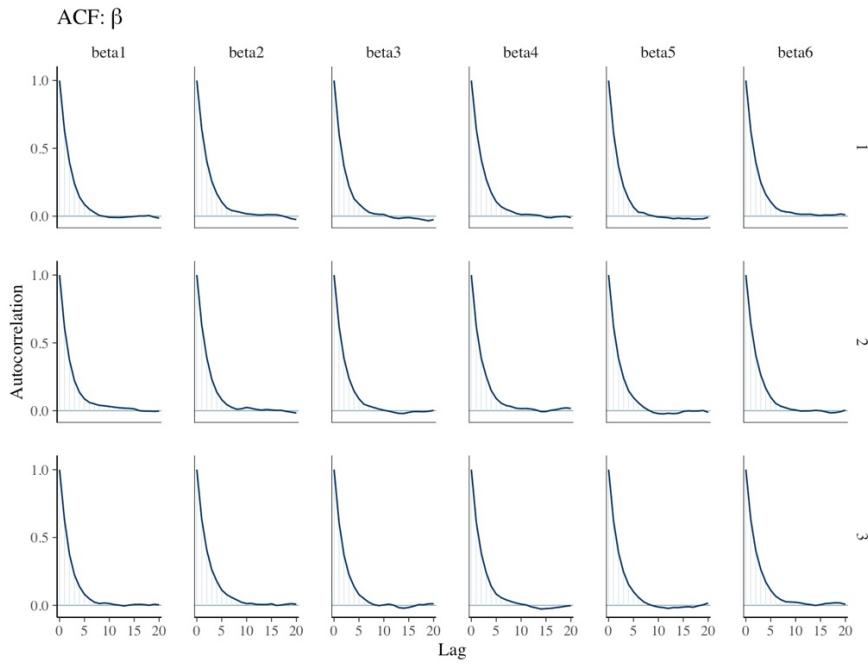
Με την ίδια λογική με το προηγούμενο ερώτημα, αυτήν την φορά για κάθε changing point έβαλα την ημερομηνία που περίπου βρίσκεται σε κάποιο ακρότατο, δηλαδή 10-10-2020, 05-12-2020, 08-01-2021, 21-02-2021, 15-03-2021.

Trace plots of $\beta_1 \dots \beta_6$ for 5 fixed changepoints



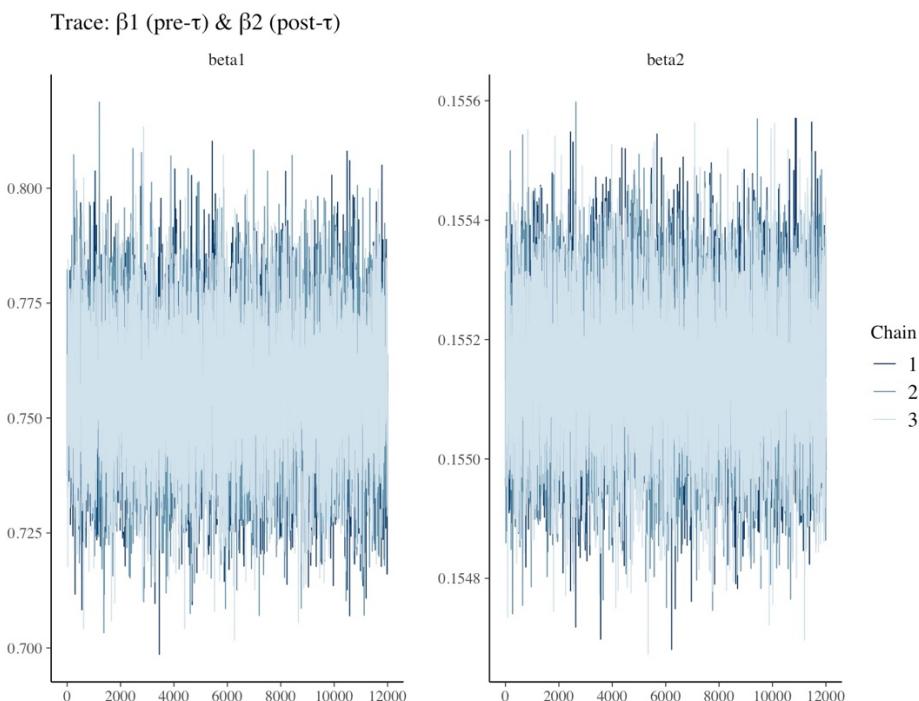
Posterior densities of $\beta_1 \dots \beta_6$ for 5 fixed changepoints



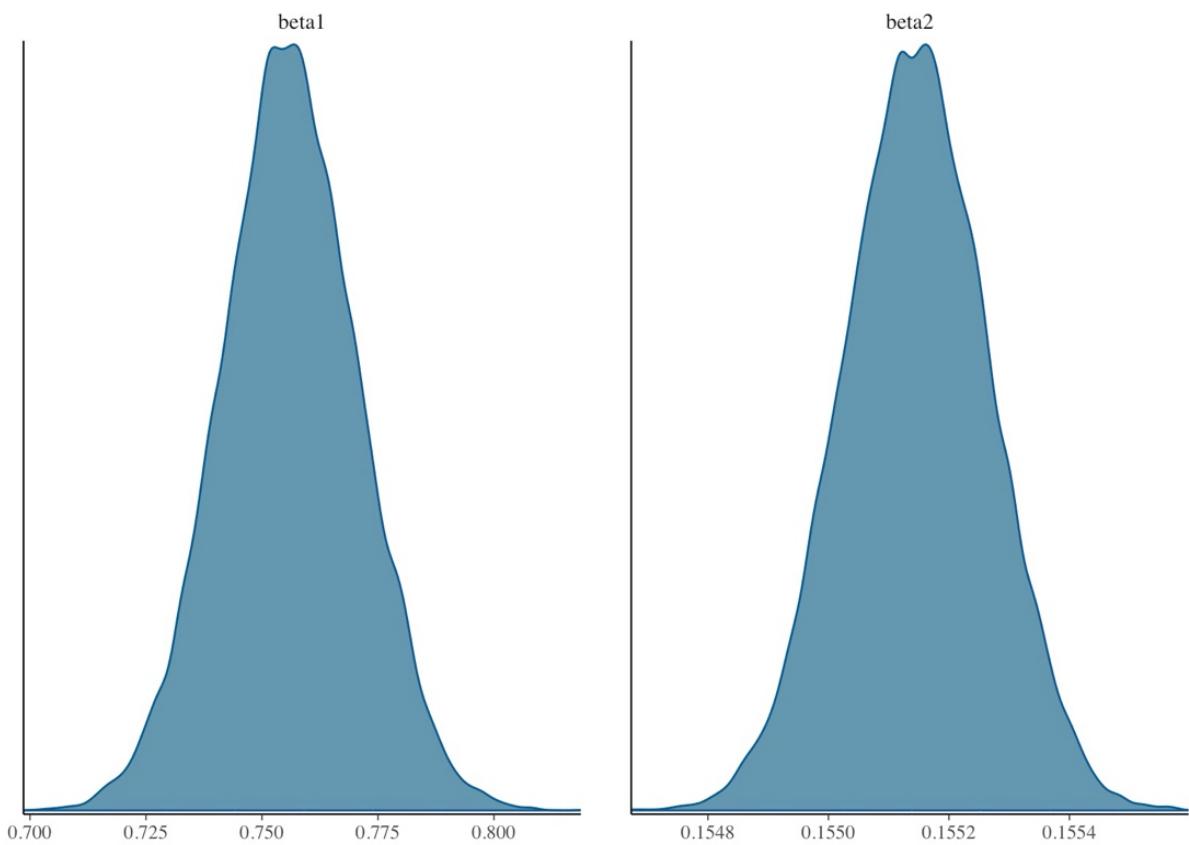


Ερώτημα 9ο:

Στο συγκεκριμένο ερώτημα για να εκτιμήσουμε το changing point από τα δεδομένα θεωρούμε ότι το changing point, έχει prior μια discrete ομοιόμορφη κατανομή και ελέγχει κάθε μια μέρα με ίδια πιθανότητα και κοιτάει που αλλάζει το β . (παράδειγμα Stagnant). Τρέχοντας τον αλγόριθμο, changing point που καταλήγει είναι η ημέρα 04-09-2020 και με $\beta 1$ να είναι 0.756 και $\beta 2$ είναι 0.155.



Density: β_1 (pre- τ) & β_2 (post- τ)



ACF: β

