Ανάλυση δεδομένων της πανδημίας COVID-19 στην Ελλάδα χρησιμοποιώντας τον νόμο του Benford ως επιδημιολογικό εργαλείο

Καγκαράς Οδυσσέας*, **Καριοφύλλη Δανάη****, **Καρπούζι Ελισάβετ*****, **Κουφοπούλου Ειρήνη******, **Λημναίος Δημήτρης*******, **Μανέτα Χρύσα********

Πρότυπο ΓΕΛ Βαρβακείου Σχολής, Αθήνα, Ελλάδα

*odysseaskagaras@gmail.com, **danaikariofylli@gmail.com, ***elisavetkarpouzi@gmail.com, ****renakouf6@gmail.com, *****dimL05@outlook.com.gr, ******christosmanetas@hotmail.com

Επιβλέπων Καθηγητής: Λυγάτσικας Ζήνων

zligatsikas@gmail.com

Περίληψη

Έχετε ποτέ φανταστεί ότι η φύση μπορεί να μην είναι όσο ομοιόμορφη όσο φαίνεται; Ή ότι μπορεί να έχει μια τάση προς τους μικρότερους αριθμούς; Στα τέλη του 19ου αιώνα, με έναυσμα μια παρατήρηση ότι στους λογαριθμικούς πίνακες οι αρχικές σελίδες (που ξεκινούσαν με 1) ήταν αρκετά πιο φθαρμένες από τις υπόλοιπες, η ιδέα του νόμου του Benford γεννήθηκε. Σύμφωνα με τον νόμο, σε αριθμητικά σύνολα από εμπειρικά δεδομένα δημιουργημένα με μια φυσική διεργασία, το αρχικό ψηφίο είναι πιο πιθανό να είναι μικρό. Σε σύνολα που υπακούνε τον νόμο του Benford, για παράδειγμα, ο αριθμός 1 εμφανίζεται ως πρωταρχικό ψηφίο στο 30% των περιπτώσεων, ενώ το 9 σε λιγότερο από 5%. Ως εκ τούτου, με αυτή την εργασία αποσκοπούμε στο να εξετάσουμε την ιδιομορφία του νόμου του Benford και να την εφαρμόσουμε στα πλαίσια της κρίσιμης υγειονομικής κατάστασης που βιώνουμε σήμερα, την πανδημία COVID-19, ως ένα εργαλείο για την επαλήθευση της εγκυρότητας των στατιστικών δεδομένων κρουσμάτων και απωλειών της Ελλάδας με το κριτήριο αν ακολουθούν αυτή την φυσική ανωμαλία. Αξίζει να σημειωθεί το γεγονός ότι παίρνοντας τα κρούσματα σε χρονικά διαστήματα ενός μηνός δεν ακολουθείται ιδιαίτερα ο νόμος, ενώ συνολικά -όταν το δείγμα δεδομένων είναι μεγάλο- σε ικανοποιητικό βαθμό. Επιπλέον, ως μέρος της έρευνας, ήρθαμε αντιμέτωποι με την ψευδαίσθηση γνωστή στους ψυχολόγους ως «μεροληψία των ισοπίθανων ενδεχομένων» (Flehinger, 1966); με λίγα λόγια, είναι μια ανθρώπινη τάση να σκεφτόμαστε ότι η «πραγματική» πιθανότητα συνεπάγεται ομοιομορφία. Αλλά ισχύει όντως κάτι τέτοιο;

Λέξεις-κλειδιά: νόμος του Benford, COVID-19, Ελλάδα, λογάριθμοι, επιδημιολογία

Εισαγωγή

Σήμερα, την περίοδο 2020 – 2022, η ανθρωπότητα ήρθε αντιμέτωπη με μια πρωτοφανή υγειονομική κατάσταση, την πανδημία του COVID-19. Συνολικά, εκατοντάδες εκατομμύρια κρούσματα έχουν επιβεβαιωθεί και εκατομμύρια συνάνθρωποί μας έχουν αποβιώσει. Παράλληλα, αλλεπάλληλα lockdown έχουν κλονίσει την ψυχολογία των πολιτών και έχουν αλλάξει ριζικά την καθημερινότητά μας. Σε έναν κόσμο λοιπόν όπου όλοι είναι εκτεθειμένοι στην παραπληροφόρηση, πόσο εύκολο είναι να αποκτήσουν έγκυρη ενημέρωση για την ακριβή υγειονομική κατάσταση (κρούσμα, θάνατοι) της κοινωνίας τους. Για άλλη μια φορά, τα μαθηματικά θα δώσουν την λύση…

Το 1881, ο Καναδός αστρονόμος Simon Newcomb παρατήρησε ότι, στους πίνακες λογαρίθμων, οι αρχικές σελίδες (που ξεκινούσαν με 1) ήταν πολύ πιο φθαρμένες από τις υπόλοιπες. Το δημοσίευμα του Newcomb είναι το πρώτο γνωστό παράδειγμα της παρατήρησης ενός νόμου γνωστού ως νόμου του Benford. Ο Newcomb, ακόμη, πρότεινε ότι η πιθανότητα ενός μονού αριθμού N να είναι το πρώτο ψηφίο ενός αριθμού είναι ίση με $\log(N + 1) - \log(N)$. (Newcomb, 1881)

Αργότερα, το 1938, ο Frank Benford ξαναφέρνει στην επιφάνεια την παρατήρηση του Newcomb, αλλά κάνει κάτι καίριο: αναγνωρίζοντας τις στατιστικές προεκτάσεις του φαινομένου, αποφασίζει να το δοκιμάσει σε ένα τεράστιο πλήθος αριθμητικών δεδομένων από 20 διαφορετικούς τομείς. Το σύνολο δεδομένων του περιλάμβανε τα μήκη 335 ποταμών, τα μεγέθη 3259 πληθυσμών των ΗΠΑ, 104 φυσικές σταθερές, 1800 μοριακά βάρη, 5000 καταχωρήσεις από ένα μαθηματικό εγχειρίδιο, 308 αριθμούς που περιλαμβάνονται σε ένα τεύχος του Reader's Digest, τις διευθύνσεις των πρώτων 342 ατόμων που αναφέρονται στο American Men of Science και 418 ποσοστά θανάτου. Ο συνολικός αριθμός των παρατηρήσεων που χρησιμοποιήθηκαν στο άρθρο ήταν 20229. (Benford, 1938)

Σήμερα, ο νόμος του Benford προβλέπει την κατανομή συχνότητας των αρχικών ψηφίων σε σύνολα εμπειρικών αριθμητικών δεδομένων. Με άλλα λόγια, σύνολα που δεν είναι τυχαία, αλλά έχουν δημιουργηθεί σύμφωνα με μια φυσική διεργασία. Εν συντομία, ο νόμος αναφέρει ότι το αρχικό ψηφίο είναι πιθανότερο να είναι μικρό. Έτσι, λοιπόν, με την παρούσα εργασία θα αναλύσουμε τα δεδομένα κρουσμάτων και θανάτων που ανακοινώνει η ελληνική κυβέρνηση για να δούμε αν υπακούνε αυτή την φυσική ανωμαλία.

Μέθοδος

Η έρευνα υλοποιήθηκε τους μήνες Απρίλιο 2021 - Φεβρουάριο 2022 από μαθητές λυκείου του ομίλου Μαθηματικών του Προτύπου ΓΕΛ Βαρβακείου Σχολής. Εξετάστηκαν ενδελεχώς αριθμητικά δεδομένα κρουσμάτων και θανάτων που ανακοίνωνε η επίσημη ιστοσελίδα του Εθνικού Οργανισμού Δημόσιας Υγείας (Ε.Ο.Δ.Υ. - https://eody.gov.gr/) χρονικού διαστήματος 2 ετών (Φεβρουάριος 2020 – Φεβρουάριος 2022). Το πρόγραμμα που χρησιμοποιήθηκε για την επεξεργασία των δεδομένων ήταν το Microsoft Excel και οι αντίστοιχες συναρτήσεις: LEFT για την εύρεση του αρχικού ψηφίου, COUNTIF για την μέτρηση του πλήθους των αριθμών που έχουν ένα συγκεκριμένο αρχικό ψηφίο και SUM για την εύρεση του συνολικού πλήθους δεδομένων. Τέλος, δημιουργήθηκαν κατάλληλα διαγράμματα συχνότητας έναντι αρχικό ψηφίο για την μετέπειτα σύγκριση με τις προβλέψεις που κάνει ο νόμος του Benford.

Εμβαθύνοντας στον νόμο του Benford

Έχουμε ήδη αναφέρει ότι κατά τον νόμο το αρχικό ψηφίο είναι πιθανότερο να είναι μικρό, αλλά η ακριβής συχνότητα δίνεται από μια μαθηματική σχέση βασισμένη σε λογαριθμικές συναρτήσεις:

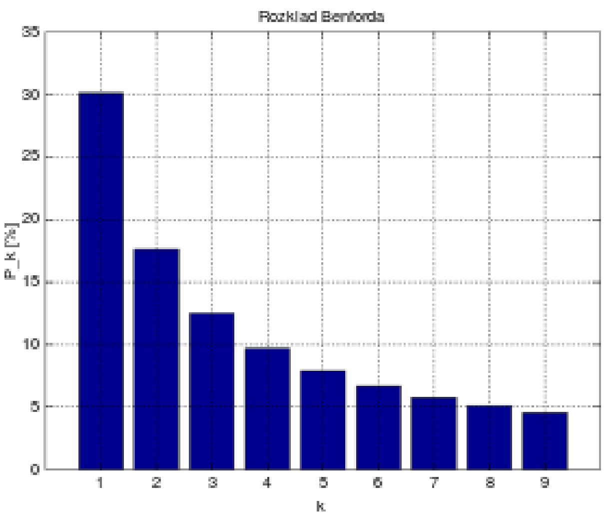
$$P(d) = \log_{10}(d+1) - \log_{10} d = \log_{10}\left(\frac{d+1}{d}\right) = \log_{10}(1+\frac{1}{d}) \text{ ,}$$

όπου d ∈ {1, ..., 9} το αρχικό ψηφίο.

Οπότε, η συχνότητα εμφάνισης του κάθε συγκεκριμένου ψηφίου δίνεται στον Πίνακας 1 και το Γράφημα 1 παρουσιάζει μια άλλη απεικόνιση των υπολογισμών.

d	P(d)	≅	%
One(1)	$\log_{10}(1+\frac{1}{1})$	0,301	30,1
Two(2)	$\log_{10}(1+\frac{1}{2})$	0,176	17,6
Three(3)	$\log_{10}(1+\frac{1}{3})$	0,125	12,5
Four(4)	$\log_{10}(1+\frac{1}{4})$	0,097	9,7
Five(5)	$\log_{10}(1+\frac{1}{5})$	0,079	7,9
Six(6)	$\log_{10}(1+\frac{1}{6})$	0,067	6,7
Seven(7)	$\log_{10}(1+\frac{1}{7})$	0,058	5,8
Eight(8)	$\log_{10}(1+\frac{1}{8})$	0,051	5,1
Nine(9)	$\log_{10}(1+\frac{1}{9})$	0,046	4,6

Πίνακας 1 Η συγκεκριμένη συχνότητα εμφάνισης του κάθε ψηφίου ως αρχικό σε ένα πλήθος δεδομένων που υπακούνε τον νόμο του Benford.



Γράφημα 1 Η κατανομή των πρώτων ψηφίων, σύμφωνα με τον νόμο του Benford. Κάθε στήλη αντιπροσωπεύει ένα ψηφίο και το ύψος της είναι το ποσοστό των αριθμών που ξεκινούν με αυτό το ψηφίο.

Έτσι, σε σύνολα που συμμορφώνονται με το νόμο του Benford, ο αριθμός 1 εμφανίζεται ως πρώτο ψηφίο στο 30% περίπου των περιπτώσεων, ενώ το 9 σε λιγότερο από το 5%. Αντιθέτως, εάν τα ψηφία κατανέμονταν ομοιόμορφα, το καθένα θα εμφανίζονταν στο 11,1% των περιπτώσεων. Ο νόμος του Benford κάνει επίσης προβλέψεις για την κατανομή δευτέρων και τρίτων ψηφίων, αλλά και πιθανούς συνδυασμούς αυτών. (Berger & Hill, 2020)

Σήμερα, ο νόμος του Benford χρησιμοποιείται στην ανίχνευση οικονομικών απατών, σε νομικές υποθέσεις, σε εκλογικά δεδομένα, σε μακροοικονομικά δεδομένα, σε δεδομένα γονιδιώματως (DNA sequencing) και στην επαλήθευση επιστημονικών ερευνών.

Υποθέσεις για τον μηχανισμό που προκαλεί την κατανομή Benford

Πολλοί υποστηρίζουν ότι ο νόμος του Benford έχει άμεση σχέση με το γεγονός ότι το σύμπαν –και συνεπώς το περιεχόμενό του– είναι πεπερασμένο. Έτσι, καθώς οι πόροι είναι περιορισμένοι, τα πάντα τείνουν προς την εξάλειψη και συνεπώς προς τους μικρότερους αριθμούς. (Sanchez, 2020) Μια άλλη άποψη, αναφέρει πως ένας από τους λόγους για την «προκατάληψη» της φύσης προς τους πιο μικρούς αριθμούς είναι ότι αυτοί αποτελούν την βάση για να φτιαχτούν οι μεγάλοι και για αυτό οι μικροί θα πρέπει να είναι πολυπληθέστεροι. Έτσι, στα πλαίσια μιας πανδημίας, πρώτα θα εμφανιστούν τα μικρά κρούσματα και μετά τα μεγάλα ως συνέπεια των πρώτων, άρα θα προηγηθούν και πιθανόν να είναι πολυπληθέστερα.

Benford: Νόμος ή Φαινόμενο – Πιθανά τυχαία σφάλματα στην έρευνα

Στα πλαίσια της μαθηματικής επιστήμης, υπάρχει μια διάκριση μεταξύ των όρων «νόμος» και «φαινόμενο». Από την μία, ο νόμος είναι κάτι ισχυρό, που μπορεί να αποδειχτεί μέσω των μαθηματικών και να εφαρμόζεται με ακρίβεια σε κάθε περίπτωση που ορίζει ο νόμος. Από την άλλη, το φαινόμενο είναι κάτι πιο αδύναμο, συνήθως πειραματικό, χωρίς κάποια ορθή απόδειξη, το οποίο εφαρμόζεται κατά προσέγγιση.

Συνεπώς, ο συλλογισμός του Benford το 1938, λόγω της αφηρημένης φύσης του, δεν έχει οδηγήσει σε μια ορθή απόδειξη με κοινή αποδοχή και έτσι φαίνεται να συγκεντρώνει πιο πολλά από τα χαρακτηριστικά του φαινομένου, παρά του παραπλανητικού ονόματός του ως νόμο.

Έτσι, για την μετέπειτα στατιστική ανάλυση των δεδομένων πρέπει να λάβουμε υπόψη την φύση φαινομένου και όχι νόμου που έχει ο νόμος του Benford. Άλλα τυχαία σφάλματα (random errors) που σίγουρα θα επηρεάσουν τα αποτελέσματα της εργασίας μας αποτελούν τα δείγματα που εξετάζονται τα οποία δεν είναι ιδανικά: ανά περιόδους της πανδημίας ο αριθμός των συνολικών διαγνωστικών τεστ άλλαξε, τα εμβόλια ομοίως επηρέασαν τον αριθμό των κρουσμάτων, όπως και τα διαφορετικά μέτρα προστασίας που ελήφθησαν την συγκεκριμένη περίοδο (lockdown).

Συνεπώς, η πανδημία κατά το διάστημα 2 ετών που εξετάζεται δεν είχε μια συνεχή πορεία, αλλά υπέστη εξωτερικές παρεμβάσεις. Ιδιαίτερης σημασίας είναι και ότι τους αρχικούς μήνες του 2020, λόγω της πιο παρορμητικής διαχείρισης της κατάστασης, αναμένεται να υπάρχει μεγαλύτερη επιρροή από τυχαία σφάλματα. Ωστόσο, θεωρούμε ότι καθώς το πλήθος δεδομένων που εξετάζουμε ανέρχεται στα 2 έτη (724 μέρες δεδομένων) και η μελέτη είναι κατά βάση συνολική, τα τυχαία σφάλματα θα μειωθούν, αυξάνοντας την ακρίβεια (precision) των αποτελεσμάτων.

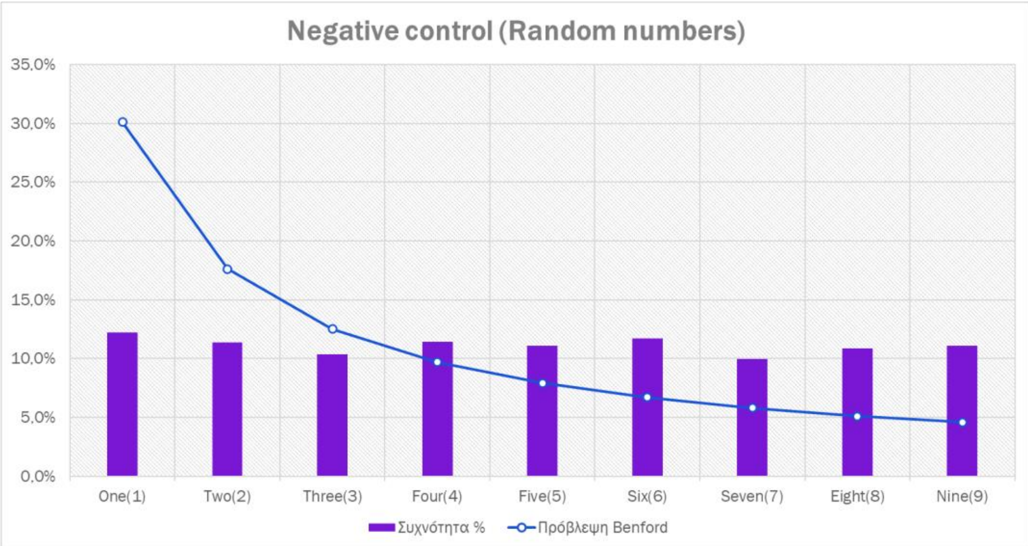
Το παράδοξο του νόμου

Αν παρατηρήσει κάποιος το σύμπαν σε μια αστοφωτογραφία, χαρακτηριστική θα είναι η ομοιομορφία που το κυριεύει μακροσκοπικά. Συνεπώς, κάποιος θα περίμενε ότι η ισοκατανομή είναι θεμελιώδες χαρακτηριστικό της φύσης. Ωστόσο, μια πρώτη παρατήρηση του νόμου του Benford, δημιουργεί την εντύπωση ότι η φύση φαίνεται να ευνοεί τους μικρούς αριθμούς. Κάποιος εύκολα θα σκεφτόταν ότι η "πραγματική" πιθανότητα συνεπάγεται ομοιομορφία, αλλά για μια ακόμη φορά η επιστήμη μας υπενθυμίζει ότι η φύση δεν έχει την υποχρέωση να είναι εύκολα κατανοήσιμη και προβλέψιμη από το ανθρώπινο είδος… (DK, 2021)

Βέβαια, υπάρχουν και περιπτώσεις που όντως παρατηρείται ισοκατανομή, όπως τα τυχαία σύνολα πραγματικών αριθμών που υπακούνε τον νόμο των Μεγάλων Αριθμών. Ας δημιουργήσουμε, για παράδειγμα, στο Excel μια βάση δεδομένων με 1500 τυχαίους αριθμούς από το 1 έως το 100000, χρησιμοποιώντας την συνάρτηση RANDBETWEEN. Υπολογίζοντας την εκατοστιαία συχνότητα (πλήθος μέρους /

πλήθος συνόλου) του κάθε ψηφίου ως αρχικό ψηφίο αυτών των αριθμών, παρατηρούμε ότι είναι ομοιόμορφη για κάθε ψηφίο (≅11%).

Επιπλέον, η συγκεκριμένη βάση δεδομένων τυχαίων αριθμών και οι συχνότητες που ανακτήθηκαν θα χρησιμεύσουν ως negative control, ένα σύνολο δηλαδή που δεν έχει τα χαρακτηριστικά των συνόλων που υπακούνε τον νόμο του Benford (δεν είναι δηλαδή φτιαγμένο σύμφωνα με μια φυσική διεργασία) και δεν αναμένουμε να φανεί η συσχέτιση πρώτου ψηφίου και συχνότητας που προβλέπει ο νόμος. Αυτό θα συνεισφέρει στην εγκυρότητα της παρούσας έρευνας δείχνοντας ότι δεν είναι η μέθοδος ανάλυσης δεδομένων που ακολουθήθηκε ή κάποιος άλλο παράγοντας που έδειξε συσχέτιση πρώτου ψηφίου και συχνότητας ανάλογης αυτής που περιγράφει ο Benford, αλλά τα χαρακτηριστικά του ίδιου του συνόλου που εξετάζεται. Το Γράφημα 2 δείχνει τις συχνότητες ενός συνόλου τυχαίων αριθμών σε σχέση με αυτές που προβλέπει ο νόμος του Benford.



Γράφημα 2 Με μωβ χρώμα οι συχνότητες αρχικού ψηφίου σε σύνολο τυχαίων αριθμών και με μπλε η πρόβλεψη Benford (Negative control). Εύκολα παρατηρείται η ίση κατανομή και η έλλειψη συσχέτισης των μεταβλητών.

Ψυχολογικές προεκτάσεις

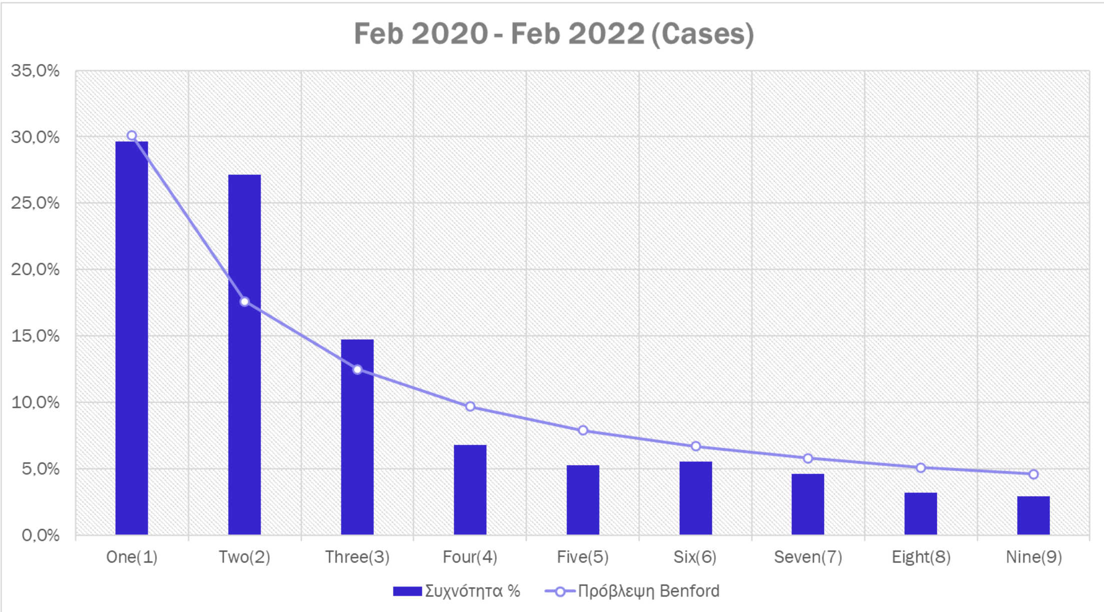
Είδαμε ότι σε ένα σύνολο τυχαίων αριθμών ακολουθείται ο νόμος των Μεγάλων Αριθμών, δηλαδή υπάρχει ίση πιθανότητα το αρχικά ψηφίο να είναι οποιοσδήποτε αριθμός. Τι γίνεται, όμως, αν αυτά τα σύνολα δεν είναι απλά τυχαίοι αριθμοί, αλλά αριθμοί με μια επιπλέον διάσταση, ορισμένοι από τα έμφυτα χαρακτηριστικά της φύσης, που δημιουργήθηκαν με έναν συγκεκριμένο μηχανισμό όπως αυτόν της πανδημίας; Σε ένα τέτοιο ερώτημα πολλοί θα ήταν αυτοί που θα συνέχιζαν να πιστεύουν πάλι σε μια ομοιόμορφη κατανομή κατά την διαίσθηση τους...

Κάτι τέτοιο, το εξέτασε το 1966 ο B. J. Flehinger, ο οποίος το χαρακτήρισε ως μια ψευδαίσθηση του ανθρώπινου νου –γνωστή στους ψυχολόγους και ως «μεροληψία των ισοπίθανων ενδεχομένων» (Flehinger, 1966). Πρόκειται, λοιπόν, για μια ανθρώπινη τάση να θεωρούμε ότι η “πραγματική” πιθανότητα συνεπάγεται ομοιομορφία. Ωστόσο, η φύση μπορεί στην πραγματικότητα να μην είναι τόσο τέλεια και ομοιόμορφη όσο φαίνεται…

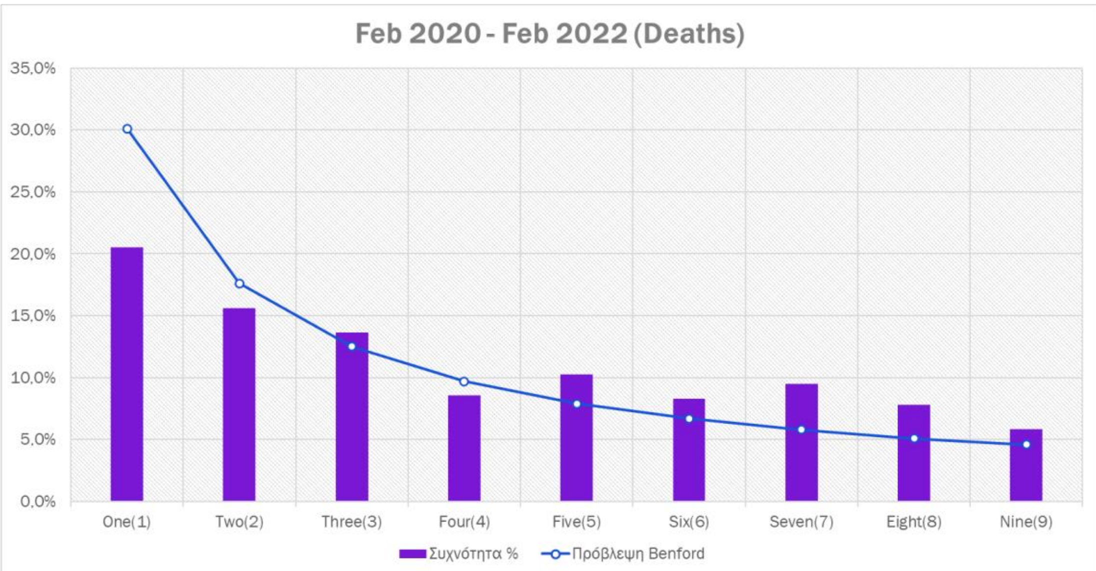
Αποτελέσματα – Στατιστική Ανάλυση

Παρακάτω παρουσιάζονται τα αποτελέσματα της έρευνάς μας. Το Γράφημα 3 δείχνει την κατανομή των πρώτων ψηφίων των κρουσμάτων σε σύγκριση με αυτή που προβλέπει ο νόμος του Benford για όλο το σύνολο των δεδομένων (Φεβρουάριος 2020 – Φεβρουάριος 2022), ενώ αντίστοιχα το Γράφημα 4 για τους θανάτους. Το Γράφημα 5 και 6 δείχνει κατανομές για μεμονωμένους μήνες της περιόδου που εξετάζεται.

Ακόμη, εξετάστηκε κατά πόσο η τάξη μεγέθους επηρεάζει τα αποτελέσματα, γι' αυτό το Γράφημα 9 δείχνει κατανομή για δεδομένα θανάτων μεταξύ του 1 και του 10, η οποία είναι πιο συνετή με τον νόμο του Benford, από ότι με την κατανομή του που εξετάζει όλα τα δεδομένα ανεξάρτητα της τάξης μεγέθους που ανήκουν. Ομοίως, το Γράφημα 7 και 8 εξετάζει κρούσματα της συγκεκριμένης τάξης μεγέθους που αναγράφεται.

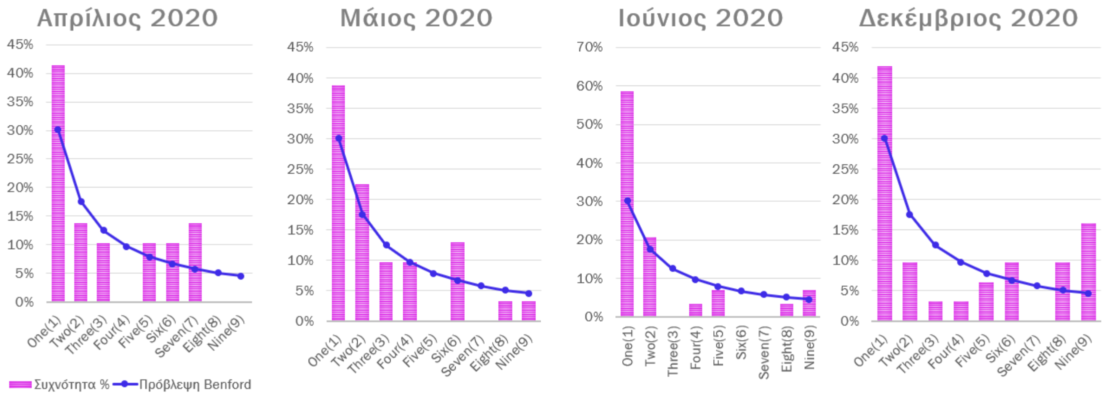


Γράφημα 3

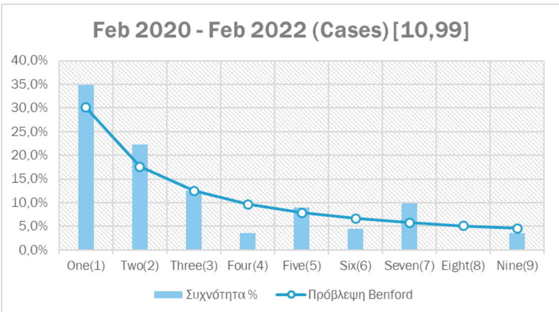


Γράφημα 4

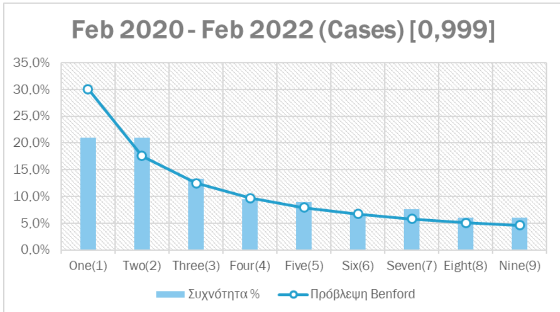
Γράφημα 5



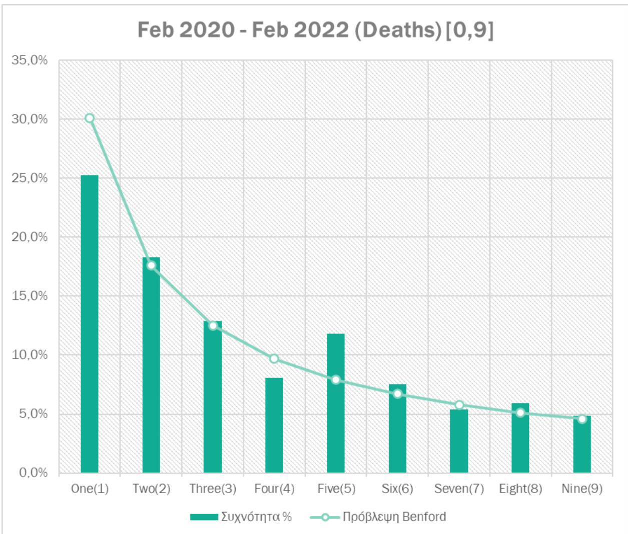
Γράφημα 6



Γράφημα 7



Γράφημα 8



Γράφημα 9

Ερμηνεία διαγραμμάτων

Μετά την εξέταση των διαγραμμάτων Benford, μπορούμε να συμπεράνουμε ότι στην περίπτωση των κρουσμάτων όλης της περιόδου ο νόμος ακολουθείται σε μεγάλο βαθμό και φαίνεται η αρχική μας υπόθεση να επαληθεύεται. Αυτό μπορούμε να το αποδώσουμε στο μεγάλο δείγμα που χρησιμοποιήθηκε και στην ποικιλία των κρουσμάτων (μεγάλο εύρος τιμών). Αρά, υπάρχει ισχυρή συσχέτιση αρχικού ψηφίου και συχνότητας εμφάνισης (strong correlation). Ομοίως, όσον αφορά τους θανάτους πάλι παρατηρείται ένα μοτίβο όμοιο με του Benford, αλλά αυτή τη φορά υπάρχει αδύναμη συσχέτιση (weak correlation), μιας και δεν είναι τόσο ξεκάθαρο όσο στην προηγούμενη περίπτωση. Αυτό μπορούμε να το αποδώσουμε στην μικρότερη ποικιλία των θανάτων μιας και το εύρος των δεδομένων ήταν από 0 έως 134, ενώ των κρουσμάτων ήταν από 0 έως 50126, ενώ το πλήθος δεδομένων ήταν το ίδιο και στις δύο περιπτώσεις. Συνεπώς, ένα πρώτο συμπέρασμα που μπορούμε να κάνουμε είναι ότι, λόγω και της λογαριθμικής φύσης της συνάρτησης του Benford, καθώς αυξάνονται οι τάξεις μεγέθους στο εύρος τιμών των αριθμητικών συνόλων που εξετάζονται, τόσο πιο εμφανείς είναι οι επιπτώσεις στην άνιση κατανομή που ορίζει ο νόμος.

Όσον αφορά τα διαγράμματα περιόδου ενός μήνα, τα μοτίβα είναι άλλοτε πιο δυσδιάκριτα άλλοτε πιο εμφανή. Το Γράφημα 5 εμφανίζει κακά παραδείγματα για την επαλήθευση αν ισχύει ο νόμος, λόγω ύπαρξης περιορισμένου πλήθους και εύρους αριθμητικών δεδομένων. Έτσι, δεν υπάρχει πλεόνασμα στα μικρότερα ψηφία ή/και υπάρχει παντελής έλλειψη δεδομένων με αρχικό ψηφίο κάποιο συγκεκριμένο νούμερο. Αντιθέτως, το Γράφημα 6 εμφανίζει καλά παραδείγματα για την επαλήθευση αν ισχύει ο νόμος, καθώς τους συγκεκριμένους μήνες καθημερινά υπήρχε ικανός αριθμός κρουσμάτων, αλλά και μεγάλο εύρος τιμών. Άρα, ένα δεύτερο συμπέρασμα είναι ότι όσο μεγαλύτερο είναι το πλήθος των δεδομένων, τόσο πιο πιστά ακολουθείται ο νόμος.

Τέλος, όσον αφορά τα διαγράμματα που εξετάζουν όλη την περίοδο άλλα εστιάζουν σε συγκεκριμένες τάξεις μεγέθους εύρους τιμών, είτε εμφανίζουν καλύτερη συνοχή με τον νόμο είτε όχι. Για παράδειγμα, το Γράφημα 7 και Γράφημα 8 εστιάζει σε ημερήσιο αριθμό κρουσμάτων από 10 έως 99 και από 0 έως 999 αντίστοιχα, δείχνοντας ευδιάκριτες κατανομές Benford. Ομοίως, το Γράφημα 9 εστιάζει σε ημερήσιο αριθμό θανάτων από 0 έως 9, που περιλαμβάνει 257 ημέρες από το σύνολο των 2 ετών, και παρουσιάζει καλύτερη κατανομή από το διάγραμμα που εξετάζει το σύνολο των θανάτων ανεξάρτητα τάξης μεγέθους.

Συμπεράσματα

Τα αποτελέσματα της έρευνας επιβεβαιώνουν ότι μια πανδημία είναι ένα δυναμικό φαινόμενο που εξελίσσεται και χαρακτηρίζεται από συνέχεια, δηλαδή κάθε φάση της πανδημίας δεν δρα ανεξάρτητα από την άλλη. Ακόμη, φαίνεται και η εγκυρότητα των δεδομένων που ανακοινώνει η κυβέρνηση μιας και, αν λάβουμε υπόψη τα τυχαία σφάλματα, η υπόθεση ότι ακολουθούν τον νόμο του Benford επιβεβαιώνεται. Άρα, φαίνεται και η αξία του νόμου του Benford ως εργαλείο στατιστικής ανάλυσης για την επιστήμη της επιδημιολογίας, που εκτός από το να επιβεβαιώνει ή να διαψεύδει δεδομένα, θα μπορούσε ακόμη μελλοντικά με τη χρήση αυτού του εργαλείου να κάνει και προβλέψεις για το πώς θα εξελιχτεί μια πανδημία.

Αναφορές

Benford, F. (1938). The law of anomalous numbers. *American Philosophical Society*, Vol. 78, No. 4, pp.551-572.

Berger, A., & Hill, T. P. (2020). The mathematics of Benford's law: a primer.

DK. (2021, February 11). *A Treatise On Benford's Law*. Ανάκτηση από https://aenigmaenterprises.medium.com/newcomb-benford-effect-part-1-39555b4080d5

Flehinger, B. (1966). On the probability that a random integer has initial digit A. *The American Mathematical Monthly*, 73:1056—1061.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*.

Sanchez, E. (2020, August 9). *Behind Benford's Law is Inherent Scarcity*. Ανάκτηση από https://ericlsanchez.medium.com/behind-benfords-law-is-inherent-scarcity-91ca9badb72e