

Οδηγός GWASMetAnalysis

1. Περιγραφή Κλάσεων

- **Research.java**

Κλάση που διαμορφώνει τον τρόπο με τον οποίο χειριζόμαστε τα δεδομένα εισόδου όταν εκείνα είναι διακριτά και αρχικοποιεί τις get() συναρτήσεις για να μπορέσουμε να εκμαιεύσουμε δεδομένα εντός του λογισμικού.

- **ContResearch.java**

Κλάση που διαμορφώνει τον τρόπο με τον οποίο χειριζόμαστε τα δεδομένα εισόδου όταν εκείνα είναι συνεχή και αρχικοποιεί τις get() συναρτήσεις για να μπορέσουμε να εκμαιεύσουμε δεδομένα εντός του λογισμικού.

- **MetaStats.java**

Στην MetaStats.java υλοποιούνται όλες οι χρήσιμες στατιστικές συναρτήσεις για την διεξαγωγή της μετα-ανάλυσης κυρίως για τα διακριτά δεδομένα. Πολλές συναρτήσεις όμως είναι κοινές και στους δύο τύπους δεδομένων όπως γίνεται αντιληπτό και στο άνωθεν flowchart. Οι Codom-, Dom- και Rec-OddsRatio() χρησιμοποιούνται για την εύρεση του odds ratio με βάση το συνεπικρατές, το επικρατές και το υπολειπόμενο μοντέλο αντίστοιχα. Οι Codom-, Dom- και Rec-FixedInvVariance() χρησιμοποιούνται για τον υπολογισμό της αντίστροφης διακύμανσης στο μοντέλο σταθερών επιδράσεων (Fixed Effects Μοντέλο) για κάθε επιμέρους μοντέλο (όπως αναλύθηκαν παραπάνω) ξεχωριστά. Η WeightedMean() υπολογίζει το σταθμισμένο μέσο των ερευνών, ενώ η StdErrorCombined() το συνολικό τυπικό σφάλμα Standard Error που προκύπτει από τις έρευνες. Έπειτα μέσω της Limits εντοπίζονται τα όρια για διάστημα εμπιστοσύνης 95% και μέσω της Zeta υπολογίζουμε τον αριθμό που εισέρχεται στη συνάρτηση $\Phi(Z)$. Τέλος, μέσω της Pvalue() υπολογίζεται το τελικό p-value που θα δοθεί στο χρήστη. Οι συναρτήσεις CochransQ(), {TauSquared()} και RandomInvVariance() χρησιμοποιούνται συνδυαστικά με τις προηγούμενες συναρτήσεις, όταν ο χρήστης επιλέγει να εκτελέσει τους υπολογισμούς του με το Random Effects Model. Οι τρεις αυτές συναρτήσεις υπολογίζουν το Q του Cochran, το τ^2 και τέλος χρησιμοποιώντας αυτά με την RandomInvVariance() προσθέτουν την επιπλέον διακύμανση στη FixedInvVariance().

- **Gaussian.java**

Open source κλάση την οποία χρησιμοποιούμε για να βρούμε την τιμή $\Phi(Z)$.

- **ContMetaStats.java**

Η μοναδική διαφορά που υπάρχει εν συγκρίσει με την απλή MetaStats είναι πως εδώ χειριζόμαστε συνεχή δεδομένα. Μέσω των Codom-, Dom- και Rec-

Οδηγός GWASMetAnalysis

EffSize() υπολογίζουμε το effect size για κάθε ένα εκ των μοντέλων (χρησιμοποιείται αντι για την odds ratio) και εν συνεχεία υπολογίζουμε το την αντίστροφη διακύμανση με τις υπόλοιπες τρεις συναρτήσεις. Τα αποτελέσματα που προκύπτουν χρησιμοποιούνται από τις συναρτήσεις της MetaStats, δηλαδή για τις υπόλοιπες μετρικές δεν υπάρχει καμία περαιτέρω αλλαγή.

- **thread_research.java**

Κλάση για τον χειρισμό των threads όσον αφορά το τι να εκτελέσουν για διακριτά δεδομένα.

- **thread_cont_research.java**

Κλάση για τον χειρισμό των threads όσον αφορά το τι να εκτελέσουν για συνεχή δεδομένα.

- **PreferenceHandling.java**

Αρχικά μέσω της FileReading() διαβάζουμε το αρχείο με τις οδηγίες που μας δίνει ο χρήστης. Αφού αποθηκεύουμε τις επιλογές περνάμε στην Choice() όπου ανοίγουμε ένα αρχείο output και ανάλογα με τις επιλογές που έχουμε διαβάσει από την FileReading() εκτελούμε το αντίστοιχο σετ εντολών και γράφουμε τα αποτελέσματα στο output αρχείο. Όλοι οι πιθανοί συνδυασμοί που μπορούν να προκύψουν αναλύονται στο παραπάνω διάγραμμα ροής.

- **DataRead.java**

Εντός της παρούσας κλάσης γίνεται ορισμός λιστών για την αποθήκευση των δεδομένων που εισάγονται από τους χρήστες. Η κλάση αυτή περιέχει δύο συναρτήσεις, την DiscreteDataRead() και την ContinuousDataRead(), οι οποίες χρησιμοποιούνται για διάβασμα σταθερών ή συνεχών δεδομένων αντίστοιχα και για την αποθήκευση των δεδομένων αυτών για μελλοντική χρήση από τις υπόλοιπες κλάσεις του λογισμικού. Η επιλογή μεταξύ των δύο συναρτήσεων, γίνεται αυτόματα από το λογισμικό ανάλογα με την επιλογή που έχει διαβαστεί από το αρχείο επιλογών του χρήστη.

2. Δεδομένα και αρχεία εισόδου

Το παρόν λογισμικό είναι ένα command line app σε πρώτη φάση. Για τον λόγο αυτό κατά την εκτέλεση του δέχεται ως πρώτο όρισμα ένα .txt αρχείο με συγκεκριμένη δομή που θα αναλυθεί εν συνεχεία, το οποίο θα περιέχει τις ρυθμίσεις που θέλει να δώσει ο χρήστης έτσι ώστε το πρόγραμμα να εκτελεί τις ανάλογες μεθόδους και ως δεύτερο όρισμα το .txt αρχείο με τα δεδομένα εισόδου τα οποία είναι space-delimited.

Οδηγός GWASMetAnalysis

Για την ορθή λειτουργία του προγράμματος τα αρχεία δεδομένων πρέπει να ακολουθούν την κάτωθι δομή:

- **Διακριτά Δεδομένα**

Αρχείο		Επεξεργασία			Μορφή		Προβολή		Βοήθεια	
snp	id	aa1	ab1	bb1	aa0	ab0	bb0	Chrom	Pos	
rs1	1	40	587	1325	72	684	2180	1	345	
rs1	2	97	620	1100	88	579	1958	1	345	
rs1	3	37	343	1256	64	589	2001	1	345	
rs1	4	63	445	1182	94	644	1991	1	345	
rs1	5	72	680	1005	78	499	2067	1	345	
rs2	1	40	587	1325	72	684	2180	5	987	
rs2	2	97	620	1100	88	579	1958	5	987	
rs2	3	37	343	1256	64	589	2001	5	987	
rs3	1	63	445	1182	94	644	1991	7	243	
rs3	2	72	680	1005	78	499	2067	7	243	
rs1	1	40	587	1325	72	684	2180	1	345	
rs1	2	97	620	1100	88	579	1958	1	345	
rs1	3	37	343	1256	64	589	2001	1	345	
rs1	4	63	445	1182	94	644	1991	1	345	
rs1	5	72	680	1005	78	499	2067	1	345	
rs2	1	40	587	1325	72	684	2180	5	987	
rs2	2	97	620	1100	88	579	1958	5	987	
rs2	3	37	343	1256	64	589	2001	5	987	
rs3	1	63	445	1182	94	644	1991	7	243	
rs3	2	72	680	1005	78	499	2067	7	243	
rs1	1	40	587	1325	72	684	2180	1	345	
rs1	2	97	620	1100	88	579	1958	1	345	
rs1	3	37	343	1256	64	589	2001	1	345	

- **Συνεχή Δεδομένα**



ContinuousTestDataSet.txt - Σημειωματάριο

Αρχείο	Επεξεργασία	Μορφή	Προβολή	Βοήθεια								
snp	id	PopAA	PopAB	PopBB	MeanAA	MeanAB	MeanBB	StdDevAA	StdDevAB	StdDevBB	Chrom	Pos
rs1	1	181	296	88	1.18	1.24	1.33	0.32	0.33	0.36	2	334
rs1	2	500	896	317	0.79	0.84	0.9	0.25	0.25	0.27	2	334
rs1	3	328	596	210	1.09	1.12	1.25	0.31	0.29	0.4	2	334
rs1	4	845	1236	409	1.19	1.22	1.3	0.37	0.37	0.41	2	334
rs1	5	168	256	100	1.36	1.38	1.48	0.36	0.38	0.42	2	334
rs2	1	181	296	88	1.18	1.24	1.33	0.32	0.33	0.36	12	238
rs2	2	500	896	317	0.79	0.84	0.9	0.25	0.25	0.27	12	238
rs2	3	328	596	210	1.09	1.12	1.25	0.31	0.29	0.4	12	238
rs4	1	845	1236	409	1.19	1.22	1.3	0.37	0.37	0.41	7	121
rs4	2	168	256	100	1.36	1.38	1.48	0.36	0.38	0.42	7	121

Οδηγός GWASMetAnalysis

Αντίστοιχα, το αρχείο με τις προτιμήσεις του χρήστη έχει την παρακάτω δομή και αναλόγως τι επιθυμεί να εκτελέσει πληκτρολογεί και το ανάλογο γράμμα.



TestPreferences.txt - Σημειωματάριο

Αρχείο Επεξεργασία Μορφή Προβολή Βοήθεια

Continuous(c)/Discrete(d)-Data: d

Random(r)/Fixed(f)-Effects-Model: r

Dominant(d)/Recessive(r)/Codominant(c)-Model: c