مجله علوم آماری، پاییز و زمستان ۱۳۸۸
جلد ۳، شماره ۲، ص ۱۴۱-۱۵۹

روش‌های رگرسیونی هیسمن-الستون در تحلیل پیوستگی ژنتیکی

مهدی اکبرزاده[۱]، حمید علوی مجد[۲]، یدالله محرابی[۳]، مریم السادات دانشپور[۴]، انور محمدی[۵]

[۱]گروه آمار زیستی و اپیدمیولوژی، دانشگاه علوم پزشکی همدان،
[۲]گروه آمار زیستی، دانشگاه علوم پزشکی شهید بهشتی،
[۳]گروه اپیدمیولوژی، دانشگاه علوم پزشکی شهید بهشتی،
[۴]پژوهشکده علوم غدد درون‌ریز و متابولیسم، دانشگاه علوم پزشکی شهید بهشتی،
[۵]گروه آمار، دانشگاه تربیت مدرس

تاریخ دریافت: ۱۳۸۸/۶/۱۰ تاریخ آخرین بازنگری: ۱۳۸۹/۱/۲۲

چکیده : یکی از مسائل مهم در علم ژنتیک مکان‌یابی یک ژن بخصوص به‌منظور رسم نقشه ژنتیکی و در نهایت تولید داروهای مؤثرتر برای درمان است. مکان‌یابی ژن با تحلیل پیوستگی ژنتیکی انجام می‌شود. یکی از روش‌های آماری که در تحلیل پیوستگی ژنتیکی مورد استفاده قرار می‌گیرد، روش رگرسیونی هیسمن-الستون است که در سال ۱۹۷۲ مطرح شد، و از آن پس ایرادات و پیشنهادات بسیاری در جهت تکمیل، به آن وارد شد. در مقاله حاضر این روش رگرسیونی و به کارگیری در تحلیل پیوستگی ژنتیکی معرفی و سیر تکاملی آن در طی سال‌های ۱۹۷۲ تا ۲۰۰۹ ارائه می‌گردد. در نهایت نحوه کاربست این روش در یک مثال کاربردی نشان داده خواهد شد.

° آدرس الکترونیک مسئول مقاله: حمید علوی مجد، alavimajd@gmail.com
° کد موضوع‌بندی ریاضی (۲۰۰۰): 62J99 و 62P10

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

واژه‌های کلیدی : تحلیل پیوستگی ژنتیکی، روش رگرسیونی هیسمن-الستون، تسهیم آلل‌ها به شیوه IBD.

۱ مقدمه

یکی از مهم‌ترین اهداف محققان در زمینهٔ بالینی تلاش در جهت یافتن راه‌کارهایی برای تشخیص و درمان‌های جدید و مؤثرتر بیماری‌ها است. این پیشرفت‌های بالینی می‌توانند عامل دیدگاه‌های جدید در علوم دیگر باشند و ترکیب آن‌ها با علوم پایه، مانند آمار، می‌تواند باعث پربار شدن این تحقیقات شود. در علم ژنتیک، مکان‌یابی یک ژن بخصوص، با استفاده از تحلیل پیوستگی ژنتیکی به منظور رسم نقشه ژنی[۱] یا تولید داروهای موثرتر برای درمان صورت می‌پذیرد.

در ژنتیک انسانی، به صفتی که حالت توارث[۲] آن از قوانین مندل پیروی نکند، صفت مرکب[۳] گویند (اوت، ۱۹۹۹). لذا در این نوع بیماری‌ها قانون معینی برای تعیین مکان[۴] دقیق ژن بیماری بر روی یک کروموزوم خاص وجود ندارد. در این حالت ناحیهٔ ژنی مربوط به بیماری موردنظر، که در مطالعات قبلی معین شده است، را درنظر گرفته و از تحلیل پیوستگی آن‌ها (نواحی ژنی انتخابی) با مارکرهای انتخابی، برای تعیین هرچه دقیق‌تر این مکان استفاده می‌شود. روش رگرسیونی هیسمن-الستون[۵] (HE) براساس میزان تشابه بین نتاج[۶] و والدین آنهاست، که احتمال رخداد این پیشامد متغیر مستقل در این مدل است و به آن تسهیم آلل‌ها به شیوه IBD[۷] می‌گویند. همچنین متغیر وابسته این مدل، براساس اختلاف در فنوتیپ کمی مورد نظر در بین نتاج است.

---

۱ Gene mapping
۲ Inheritance mode
۳ Complex trait
۴ Locus
۵ Haseman-Elston
۶ Offspring
۷ Allele Sharing Identical By Descend

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

در این مقاله ابتدا به معرفی روش رگرسیونی هیسمن-الستون در تحلیل پیوستگی ژنتیکی و سیر تکاملی آن در طی سال‌های ۱۹۷۲ تا ۲۰۰۹ می‌پردازیم. سپس این روش برای تحلیل داده‌های مربوط به صفت دور کمر افراد مبتلا به سندرم متابولیک مطالعه قند و لیپید تهران به کار گرفته خواهد شد.

۲ مدل ژنتیکی فالکونر

برای n زوج فرزندی[۸] مستقل، مدل ژنتیکی فالکونر[۹] به صورت

$$x_{tj} = \mu + g_{tj} + e_{tj}\ ,\quad t = ۱, ۲,\quad j = ۱, ..., n$$

است، که در آن x_{tj} مقدار فنوتیپ زوج فرزندیِ jام در فرزند tام[۱۰]، g_{tj} اثر ژنتیکی ناشی از مکان صفت موردنظر و e_{tj} جمله خطای مدل هستند، به طوری که

$$Corr(x_{۱j}, x_{۲j}) = Corr(e_{۱j}, e_{۲j}) = \rho_s,\ E(e_{tj}) = ۰$$

و برای $e_j = e_{۱j} - e_{۲j}$ مقدار $E(e_j^۲)$ با $\sigma_e^۲$ نشان داده می‌شود. در این مدل فرض بر این است که اثرات محیطی[۱۱] و مولفه پلی ژنتیک[۱۲] وجود ندارد و با یک مکان ژنی دو آللی با آلل‌های B و b، به ترتیب با فراوانی‌های p و q سروکار داریم که از قانون آمیزش تصادفی در حالت توازن هاردلی-وینبرگ[۱۳] (WHE) و بدون اثر متقابل ژنی[۱۴] و اپیستاسی[۱۵]، پیروی می‌کنند. در مدل ژنتیکی فالکونر، اثر ژنتیکی به صورت

$$g_{tj} = \begin{cases} a & \text{درصورتی که ژنوتیپ فرد } BB \text{ باشد} \\ d & \text{درصورتی که ژنوتیپ فرد } Bb \text{ باشد} \\ d & \text{درصورتی که ژنوتیپ فرد } bb \text{ باشد} \end{cases}$$

[۸] Sib-pair
[۹] Falconer genetic model
[۱۰] Sib
[۱۱] Environmental effects
[۱۲] Polygenic component
[۱۳] Hardly-Weinberg equilibrium
[۱۴] Genetic interaction effect
[۱۵] Epistasis effect

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

است و داریم

$$E(x_{ij}) = \mu + a(p-q) + ۲pqd$$

در این مدل واریانس ژنتیکی، $\sigma_g^۲$، به دو مولفه واریانس افزایشی، $\sigma_a^۲$، و واریانس غالبیت، $\sigma_d^۲$،به صورت $\sigma_g^۲ = \sigma_a^۲ + \sigma_d^۲$ تجزیه می‌شود، که در آن $\sigma_a^۲ = ۲pq[a-d(p-q)]^۲$ و $\sigma_d^۲ = ۴p^۲q^۲d^۲$ هستند.

۳ روش رگرسیونی هیسمن-الستون

موضوع بررسی پیوستگی و کشف آن بین مکان یک نشانگر و صفت کمی با مطالعه بر زوج فرزندی‌ها را اولین‌بار، پنروس در سال ۱۹۳۸ مطرح کرد (اوت، ۱۹۹۹). روشی را که هیسمن-الستون (۱۹۷۲) برای بررسی پیوستگی صفت‌ها معرفی کردند، مربوط به صفت کمی بوده و محاسبات و روش‌های ارائه شده را به داده‌های جمع‌آوری شده از زوج فرزندی‌ها محدود کردند. به علاوه آن‌ها توانستند با استفاده از این روش کسر نوترکیبی[۱۶] را بین مکان ژنی صفت مورد نظر و مکان نشانگر نیز برآورد کنند و تنها صفت کمی دو آللی را مورد بررسی قرار دادند. روش ارائه شده در این مقاله توسط بسیاری از محققین بررسی و شبیه‌سازی شد و مرجع بسیاری از محققین قرار گرفت، تا جایی که مجلهٔ Human Heredity در سال ۲۰۰۳ سالگرد سی‌امین سال استفاده از این مقاله را گرامی داشتند (زیگر و اینکه، ۲۰۰۶) و در مقالات بعدی از این روش به نام OHE[۱۷] یاد کردند.

۱.۳ مقادیر مورد انتظار شرطیِ توان دوم اختلاف زوج فرزندی‌ها

فرض کنید $y_j = (x_{۱j} - x_{۲j})^۲$ توان دوم اختلاف مقدار صفت مورد نظر برای jامین زوج فرزندی باشد، در این صورت برای مقادیر ثابت e_j، y_j برابر با یکی از هفت

[۱۶] Recombination fraction

[۱۷] Original HE method

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

مقدار ممکن، با توجه به ژنوتیپ فرزند اول و دوم، خواهد شد. این مقادیر در ستون دوم جدول ۱ آورده شده است.

جدول ۱: توزیع شرطی Y_j، به شرط π_j

π_j				
۱	$\frac{1}{2}$	۰	y_j	زوج فرزندی
p^2	p^3	p^4	e_j^2	BB-BB
q^2	q^3	q^4	e_j^2	bb-bb
$2pq$	pq	$4p^2q^2$	e_j^2	Bb-Bb
۰	p^2q	$2p^3q$	$(a-d+e_j)^2$	BB-Bb
۰	p^2q	$2p^3q$	$(-a+d+e_j)^2$	Bb-BB
۰	pq^2	$2pq^3$	$(a+d+e_j)^2$	Bb-bb
۰	pq^2	$2pq^3$	$(-a-d+e_j)^2$	bb-Bb
۰	۰	p^2q^2	$(2a+e_j)^2$	BB-bb
۰	۰	p^2q^2	$(-2a+e_j)^2$	bb-BB

هر زوج فرزندی می‌تواند صفر، یک یا دو ژن را به صورت IBD در مکان صفت داشته باشد. بنابراین نسبت ژن‌هایی که به‌صورت IBD هستند، باید یکی از مقادیر ۰، $\frac{1}{2}$ یا ۱ را داشته باشد. اگر این نسبت را برای jامین زوج فرزندی با π_j نشان داده شود، توزیع شرطی زوج فرزندی‌ها به شرط π_j در جدول ۱ نشان داده شده است.

وقتی $\pi_j = 0$، در مکان مورد نظر، فرزندها به هم وابسته نبوده و توزیع آمیزش ژن‌ها مانند جامعه‌ای با آمیزش تصادفی است و مقادیر احتمال شرطی مورد نظر با توجه به قانون HW محاسبه شده است (ستون سوم جدول ۱). برای بررسی حالت $\pi_j = \frac{1}{2}$ و $\pi_j = 1$، زوج فرزندی‌های با ژنوتیپ $BB - BB$ را درنظر بگیرید.

برای $\pi_j = \frac{1}{2}$، دوتا از آلل‌های B از یک والد و دوتای دیگر هر یک از والدی جدا به ارث رسیده‌اند، لذا مقدار این احتمال شرطی برابر با p^3 خواهد بود. برای $\pi_j = 1$، دوتا از آلل‌های B از یک والد و دوتای دیگر از والد دیگر آمده‌اند، لذا مقدار احتمال شرطی برابر p^2 است. سایر احتمال‌ها هم به همین صورت قابل

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

محاسبه هستند. حال با استفاده از جدول ۱ داریم

$$E(Y_j|\pi_j) = \begin{cases} \sigma_e^2 & \pi_j = 1 \\ \sigma_e^2 + \sigma_a^2 + 2\sigma_d^2 & \pi_j = \frac{1}{2} \\ \sigma_e^2 + 2\sigma_a^2 + 2\sigma_d^2 & \pi_j = 0 \end{cases} \qquad (1)$$

با توجه به (۱) واضح است که اگر غالبیت برابر صفر باشد، $d = 0$ یا به طور معادل $\sigma_d^2 = 0$، داریم

$$\begin{aligned} E(Y_j|\pi_j) &= (\sigma_e^2 + 2\sigma_g^2) - 2\sigma_g^2\pi_j \\ &= \alpha + \beta\pi_j\ , \qquad \pi_j = 0, \frac{1}{2}, 1 \end{aligned} \qquad (2)$$

که در آن $\alpha = (\sigma_e^2 + 2\sigma_g^2)$ و $\beta = -2\sigma_g^2$. اگر π_jها معلوم باشند و مدل رگرسیونی خطی ساده (۲) را به داده‌ها برازش داده شود، $-\frac{1}{4}\widehat{\beta}$ یک برآورد نااریب برای σ_g^2 خواهد بود که در آن $\widehat{\beta}$ برآوردگر کمترین توان‌های دوم پارامتر β است. این نتایج حتی زمانی که بین آلل‌ها غالبیت وجود داشته باشد نیز صحیح هستند (هیسمن و الستون، ۱۹۷۲).

۲.۳ برآورد π_j برای یک مکان نشانگر

اگر f_{ji} احتمال آن که jامین زوج فرزندی، دارای i ژن به صورت IBD در مکان نشانگر باشد، آن‌گاه برآوردگر π_j، به صورت

$$\widehat{\pi}_j = f_{j2} + \frac{1}{2}f_{j1} \qquad (3)$$

خواهد شد. وقتی ژنوتیپ‌های زوج فرزندی‌ها و والدین معلوم باشند، محاسبهٔ $\widehat{\pi}_j$ آسان است. برای حالت کلی چند آللی، هفت نوع آمیزش و مشابهاً هفت نوع زوج فرزندی وجود دارد. برای مثال $A_1A_1 \times A_1A_1$ و $A_2A_2 \times A_2A_2$ از لحاظ ژنوتیپی متفاوت است ولی از لحاظ نوع آمیزش یکسان است. همهٔ هفت حالت ممکن به همراه احتمال‌های آنها و مقادیر $\widehat{\pi}_j$ در جدول ۲ نشان داده شده‌اند. در این جدول عدد داخل هر یک از پرانتزها تعداد زوج فرزندی‌های متفاوت از لحاظ ژنوتیپی را

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

نشان می‌دهد. برای مثال در آمیزش نوع IV، ممکن است حاصل آمیزش زوج فرزندی‌های از نوع V به صورت $A_iA_j - A_iA_j$ یا $A_iA_j - A_iA_k$ باشد که هردو حالت با احتمال یکسان f_{ji} رخ خواهند داد.

جدول ۲: مقادیر $\widehat{\pi}$ برای ژنوتیپ والدین و فرزندان معلوم

$\widehat{\pi}_j$	f_{j2}	f_{j1}	f_{j0}	احتمال	نوع زوج فرزندی	نوع آمیزش
۱/۲	۱/۴	۱/۲	۱/۴	p_i^4	$I : A_iA_i - A_iA_i$	$I : A_iA_i \times A_iA_i$
۱/۲	۱/۴	۱/۲	۱/۴	$2p_i^2p_j^2$	$V : A_iA_j - A_iA_j$	$II : A_iA_i \times A_jA_j$
۱/۴	۱/۲	۱/۲	۰	$p_i^3p_j$	$I : A_iA_i - A_iA_i$	$III : A_iA_i \times A_iA_j$
۱/۴	۰	۱/۲	۱/۲	$2p_i^2p_j$	$III : A_iA_i - A_iA_j$	
۱/۴	۱/۲	۱/۴	۰	$p_i^3p_j$	$V : A_iA_j - A_iA_j$	
۱/۴	۱/۲	۱/۴	۰	$p_i^2p_jp_k$	$V : (2)$	$IV : A_iA_i \times A_jA_k$
۱/۴	۰	۱/۲	۱/۲	$2p_i^2p_jp_k$	$VI : A_iA_j - A_iA_k$	
۱	۱	۰	۰	$p_i^2p_j^2/4$	$I : (2)$	$V : A_iA_j \times A_iA_j$
۰	۰	۰	۱	$p_i^2p_j^2/2$	$II : A_iA_i - A_jA_j$	
۱/۲	۰	۱	۰	$p_i^2p_j^2$	$III : (2)$	
۱/۲	۱/۲	۰	۱/۲	$p_i^2p_j^2$	$VI : A_iA_j - A_iA_j$	
۱	۱	۰	۰	$p_i^2p_jp_k/2$	$I : A_iA_i - A_iA_i$	$VI : A_iA_j \times A_iA_k$
۱/۲	۰	۱	۰	$p_i^2p_jp_k$	$III : (2)$	
۰	۰	۰	۱	$p_i^2p_jp_k$	$IV : A_iA_i - A_jA_k$	
۱	۱	۰	۰	$p_i^2p_jp_k/2$	$V : (3)$	
۰	۰	۰	۱	$p_i^2p_jp_k$	$VI : A_iA_j - A_iA_k$	
۱/۲	۰	۱	۰	$p_i^2p_jp_k$	$VI : A_iA_j - A_iA_k$	
					$VI : A_iA_j - A_iA_k$	
۱	۱	۰	۰	$p_ip_jp_kp_i/2$	(۴)	$VII : A_iA_j \times A_kA_i$
۱/۲	۰	۱	۰	$p_ip_jp_kp_i$	(۴)	
۰	۰	۰	۱	$p_ip_jp_kp_i$	(۲)	

وقتی برخی از ژنوتیپ‌ها در والدین معلوم نباشند، محاسبهٔ f_{ji}ها مشکل خواهد شد و می‌توان آنها را با استفاده از الگوریتم ارائه شده توسط کاترمن (۱۹۶۹) محاسبه نمود (هیسمن و الستون، ۱۹۷۲).

۳.۳ مقدار مورد انتظار ضریب رگرسیونی

در یک حالت خاص فرض کنید مکان نشانگر دو آللی باشد، غالبیت نداشته و اطلاعات والدین معلوم باشند. در این صورت نشان می‌دهیم

$$E(Y_j|\widehat{\pi}_j) = \alpha + \beta\widehat{\pi}_j \qquad (۴)$$

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

به‌طوری که

$$\beta = -۲(۱ - ۲c)^۲\sigma_g^۲ \qquad (۵)$$

که در آن c کسر نو ترکیبی بین مکان نشانگر و صفت مورد نظر است. حال اگر نسبت π_j برای صفت موردنظر با π_{jt} و برای نشانگر با π_{jm} نشان دهیم، هر دو با استفاده از رابطه (۳) قابل محاسبه می‌باشند. با فرض آنکه بین مکان صفت موردنظر و نشانگر تعادل پیوستگی[۱۸] برقرار است، یعنی احتمال پیوستگی بین این دو مکان در همه نقاط این فاصله برابر باشند، برای π_{jt} ثابت، y_j و $\widehat{\pi}_{jm}$ مستقل‌اند و برای π_{jm} ثابت، π_{jt} و $\widehat{\pi}_{jm}$ از هم مستقل‌اند. بنابراین داریم

$$\begin{aligned} E(Y_j|\widehat{\pi}_{jm}) &= \sum E(Y_i|\pi_{jt})P(\pi_{jt}|\widehat{\pi}_{jm}) \\ &= \sum\sum E(Y_i|\pi_{jt})P(\pi_{jt}|\pi_{jm})P(\pi_{jm}|\widehat{\pi}_{jm}) \end{aligned} \qquad (۶)$$

که مجموع‌یابی بر روی هر سه مقدار π_{jm} و π_{jt} صورت می‌گیرد.

جدول ۳: توزیع توام π_{jm} و π_{jt}

	π_{jm}			
π_{jt}	۰	$\frac{۱}{۲}$	۱	کل
۰	$\frac{\psi^۲}{۴}$	$\frac{\psi(۱-\psi)}{۲}$	$\frac{(۱-\psi)^۲}{۴}$	$\frac{۱}{۴}$
$\frac{۱}{۲}$	$\frac{\psi(\psi-۱)}{۲}$	$\frac{(۱-۲\psi+۲\psi^۲)}{۲}$	$\frac{\psi(۱-\psi)}{۲}$	$\frac{۱}{۲}$
۱	$\frac{(۱-\psi)^۲}{۴}$	$\frac{\psi(۱-\psi)}{۲}$	$\frac{\psi^۲}{۴}$	$\frac{۱}{۴}$
کل	$\frac{۱}{۴}$	$\frac{۱}{۲}$	$\frac{۱}{۴}$	۱

توزیع توأم π_{jm} و π_{jt} در جدول ۳ ارائه شده است، که در آن $\psi = c^۲ - (۱-c)^۲$ است. توزیع توأم π_{jm} و $\widehat{\pi}_{jm}$ در یک حالت خاص، از ژن نشانگر دو آللی، بدون غالبیت و اطلاعات والدین کامل در جدول ۴ آمده است. برای مثال $\widehat{\pi}_{jm} = ۰$ است، اگر و فقط اگر آمیزش از نوع $Aa \times Aa$ و نوع زوج فرزندی، $AA \times aa$ باشد و احتمال رخداد این پیشامد $\frac{p^۲q^۲}{۲}$ است. طبعاً برای این زوج فرزندی داریم:

[۱۸] Linkage equilibrium

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

$\pi_{jm} = ۰$. اگر فرزندها به‌صورت $AA - AA$ یا $aa - aa$، هر یک با احتمال $\frac{p^۲q^۲}{۴}$، باشند آن‌گاه $\widehat{\pi}_{jm} = \pi_{jm} = ۱$.

جدول ۴: توزیع توام π_{jm} و $\widehat{\pi}_{jm}$

	π_{jm}			
کل	۱	$\frac{۱}{۲}$	۰	$\widehat{\pi}_{jm}$
$\frac{p^۲q^۲}{۲}$	۰	۰	$\frac{p^۲q^۲}{۲}$	۰
$p^۳q + pq^۳$	۰	۰	$p^۳q + pq^۳$	$\frac{۱}{۴}$
$p^۴ + ۵p^۲q^۲ + q^۴$	$\frac{p^۴+۴p^۲q^۲+q^۴}{۴}$	$\frac{p^۴+۶p^۲q^۲+q^۴}{۲}$	$\frac{p^۴+۴p^۲q^۲+q^۴}{۴}$	$\frac{۱}{۲}$
$۲(p^۳q + pq^۳)$	$p^۳q + pq^۳$	$p^۳q + pq^۳$	۰	$\frac{۳}{۴}$
$\frac{p^۲q^۲}{۲}$	$\frac{p^۲q^۲}{۲}$	۰	۰	۱
۱	$\frac{۱}{۴}$	$\frac{۱}{۲}$	$\frac{۱}{۴}$	کل

از جدول‌های ۳ و ۴ و رابطه‌های (۱) و (۶) داریم:

$$E(y_j|\widehat{\pi}_{jm}) = [\sigma_e^۲ + ۲(۱ - ۲c + ۲c^۲)\sigma_g^۲] - ۲(۱ - ۲c)^۲\sigma_g^۲\widehat{\pi}_{jm}$$

این نتیجه برای مکان نشانگرهای چند آللی، اگر در مکان صفت غالبیت نداشته باشیم، قابل تعمیم است. همچنین زمانی که غالبیت وجود دارد نیز، به‌طور مجانبی درست است و برای نمونه‌های به حجم بزرگ اریبی کوچک است (هیسمن و الستون، ۱۹۷۲). بنابراین در رابطه رگرسیونی (۴) می‌توان با جایگذاری $\widehat{\pi}_{jm}$ به‌جای π_{jm}، فرض $\beta = ۰$ را بااستفاده از آزمون t تک دامنه‌ای آزمون کرد. زیرا اگر فرض $\beta < ۰$ معنی‌دار شود، پیوستگی بین مکان صفت موردنظر و نشانگر پذیرفته می‌شود، زیرا اگر $\beta < ۰$، آن‌گاه $c < \frac{۱}{۲}$ و پیوستگی برقرار خواهد بود.

حال فرض کنید k مکان از صفت موردنظر موجود است، که هر یک با مکان مربوط به نشانگر پیوسته‌اند. آن‌گاه رابطهٔ (۵) را می‌توان برای هر مکان صفت به صورت جداگانه درنظر گرفت و اگر مکان‌های صفت مورد نظر دو به دو ناپیوسته باشند و اپیستاسی وجود نداشته باشد، داریم

$$E(\widehat{\beta}) = -۲\sum_{i=۱}^{k}(۱ - ۲c_i)^۲\sigma_i^۲$$

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

که در آن σ_i^2 سهم واریانس ژنتیکی مکان iام صفت مورد نظر و c_i کسر نو ترکیبی بین آن مکان و مکان نشانگر هستند. بنابراین مقادیر معنی‌دار $\beta > 0$ نشان‌دهندهٔ وجود پیوستگی بین مکان نشانگر با یک یا چند مکان صفت موردنظر است.

۴ تعمیم‌های روش HE

یکی از ایرادهای روش oHE این است که از توان‌های دوم اختلاف‌های مربوط به زوج فرزندی‌ها در هر خانواده، به‌عنوان معیاری برای شباهت فنوتیپی افراد استفاده می‌کند، که ممکن است مقداری از اطلاعات در مورد هر زوج فرزندی را از دست بدهد. به عنوان مثال اگر x_1 و x_2 دارای توزیع نرمال دو متغیره با میانگین صفر و ماتریس کوواریانس غیرهمانی باشند، متغیر $(x_1 - x_2)$ همه اطلاعات موجود در داده‌های خام را نخواهند داشت و تنها توزیع دو متغیره تمام اطلاعات موجود در داده‌های خام را خواهد داشت. برای مثال، زوج متغیرهای $(x_1 - x_2, x_1 + x_2)$ همه اطلاعات را دارند. زیرا می‌دانیم که این دو متغیر ناهمبسته‌اند، از طرفی در صورت وجود توزیع نرمال دو متغیره، این دو متغیر مستقل هم خواهند بود. لذا اگر از وجود توأم این دو متغیر در تحلیل پیوستگی استفاده شود، به مقدار ناچیزی به توان آزمون اضافه خواهد شد (آلمسی و بلانگر، ۱۹۹۷).

۱.۴ روش HE بازنگری شده

رایت (۱۹۷۷) نشان داد اگر روش oHE را یک‌بار برای متغیر وابستهٔ توان‌های دوم اختلاف‌ها و بار دیگر برای متغیر وابستهٔ توان‌های دوم مجموع‌های مربوط به مقادیر صفت مورد نظر در هر زوج فرزندی به‌کار برده شود، دو خط رگرسیونی موازی به دست خواهد آمد. لذا می‌توان میانگین دو شیب خطوط رگرسیونی را به‌عنوان شیب خط رگرسیونی نهایی برآورد کرد و اگر واریانس دو خط رگرسیونی برابر باشد، این برآوردگر بهترین برآوردگر خواهد بود، که در واقع در این حالت خط رگرسیونی نهایی معادل با خط رگرسیونی است که متغیر وابستهٔ آن حاصل‌ضرب مقادیر صفت زوج فرزندی‌ها باشد (درینگالکو، ۱۹۹۸). براین اساس، روش HE

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

بازنگری شده[۱۹](rHE) توسط آلستون و همکاران (۲۰۰۰) ارائه شد. در این روش از حاصل‌ضرب متقاطع مرکزی شده با میانگین مقادیر صفت[۲۰] مربوط به زوج فرزندی‌های تنی[۲۱]، $(x_{۱j}-\mu)(x_{۲j}-\mu)$، استفاده شد. استفاده از این روش باعث افزایش توان آزمون می‌شود و همچنین این روش برای صفات کیفی نیز مورد استفاده است. همچنین در سال ۲۰۰۰ تعدادی از محققین استواری خطای نوع I را در صورت وجود انواع انحرافات ممکن مورد بررسی قرار دادند و نشان دادند که این روش بهتر از روش‌های قبلی، oHE، است (آلیسون و همکاران، ۲۰۰۰). همچنین در یک مطالعه شبیه‌سازی نشان داده شد که اگر از میانگین‌های هر خانواده (نه یک میانگین برای همه خانواده‌ها) برای مرکزی کردن در روش rHE استفاده شود، این روش دارای توان آزمون بالاتری خواهد بود و در صورت وجود همبستگی میان عوامل ژنتیکی و محیطی در خانواده‌ها، توان تجربی روش rHE از روش oHE کمتر می‌باشد (پالمر و همکاران، ۲۰۰۰).

۲.۴ روش HE وزنی

به دنبال ایرادهای روش‌های rHE روش‌های HE وزنی[۲۲] (wHE)، توسط زیگلر و اینکه (۲۰۰۶) پیشنهاد شد. فرض کنید برآوردگر شیب خط رگرسیونی مربوط به توان‌های دوم اختلاف‌ها برابر $\widehat{\beta}_{۱}$ و شیب خط رگرسیونی مربوط به توان‌های دوم مجموع‌ها برابر با $\widehat{\beta}_{۲}$ باشد. در این‌صورت

$$\widehat{\beta}=\frac{\widehat{\sigma}_{۲}^{۲}}{\widehat{\sigma}_{۱}^{۲}+\widehat{\sigma}_{۲}^{۲}}\widehat{\beta}_{۱}+\frac{\widehat{\sigma}_{۱}^{۲}}{\widehat{\sigma}_{۱}^{۲}+\widehat{\sigma}_{۲}^{۲}}\widehat{\beta}_{۲}$$

یک ترکیب بهینه از ضرایب رگرسیونی است و برخی از محققین پیشنهاد کردند به جای واریانس ضرایب از واریانس تجربی آن استفاده شود (ویسکر و هاپر، ۲۰۰۲). به‌طور کلی اگر w وزن مربوطه باشد، کلاس برآوردگرهای وزنی β به صورت

[۱۹] Revisited HE method

[۲۰] Mean-corrected cross-product trait

[۲۱] Full-sib-pair

[۲۲] Weighted HE

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

$w\widehat{\beta}_1 + (1-w)\widehat{\beta}_2$ خواهد بود. برآوردگر وزنی دیگری توسط خو و لی (۲۰۰۰) با وزن‌های

$$w = \frac{\widehat{\sigma}_2^2 - \widehat{\sigma}_{12}^2}{\widehat{\sigma}_1^2 + \widehat{\sigma}_2^2 - 2\widehat{\sigma}_{12}^2}$$

معرفی شد، که در آن $\widehat{\sigma}_{12}^2$ کوواریانس بین دو ضریب است. این برآوردگر وقتی که واریانس‌ها و کوواریانس‌ها کاملاً معلوم باشند دارای کمترین واریانس در بین تمام برآوردگرهای خطی دو ضریب مربوط به دو خط رگرسیونی است. به علاوه استفاده از برآوردهای وزنی برای زوج فرزندهای بزرگ‌تر با استفاده از همبستگی‌های زوجی بین فرزندها پیشنهاد شد (شته و همکاران، ۲۰۰۴). در این بین، برآوردگر وزنی‌ای که توسط شم و همکاران (۲۰۰۱) پیشنهاد شد، از همه مشهورتر است و وزن‌های آن به صورت

$$w = \frac{(X_1 + X_2)^2}{(1+r)^2} - \frac{(X_1 - X_2)^2}{(1-r)^2} \qquad (۷)$$

است که در آن r ضریب همبستگی بین فرزندها است. این روش به این دلیل مشهور شده است که به همراه روش‌هایی مانند oHE، rHE در مدل یکپارچه و کلی‌تر GEE[۲۳]، قرار می‌گیرند و از این روش به نام HE-COM یاد کرده‌اند (چن و همکاران، ۲۰۰۴).

۳.۴ تعمیم روش oHE به حالت‌های گسترده‌تر از زوج فرزندی‌های تنی

یکی از پیش‌فرض‌های روش oHE این است که همهٔ فرزندهای مربوط به زوج فرزندی‌های تنی مستقل هستند. البته این روش برای نمونه‌های بزرگ، حتی زوج فرزندی‌های تنی مربوط به یک خانواده هم اعتبار خود را از دست نمی‌دهد، (آموس و همکاران، ۱۹۸۹، کولیس و مورتن، ۱۹۹۵). هم‌چنین با استفاده از خانواده‌های به حجم ۳، این روش توان خود را به مقدار بسیار ناچیزی از دست می‌دهد (بکلودر و السنون، ۱۹۸۲). در تحلیل پیوستگی خانواده‌های بزرگ‌تر از زوج فرزندی‌ها بیشتر

[۲۳] Generalize estimation equation

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

مدنظر هستند، زیرا اطلاعات بیشتری را در مورد پیوستگی به ما ارائه می‌کنند (ونگ، ۲۰۰۶)، ولی تنها همبستگی بین زوج توان‌های دوم اختلاف مربوط به زوج فرزندی‌های تنی‌ای برابر با صفر است که هیچ فرزند مشترکی باهم نداشته باشند، در صورتی‌که یک فرزند مشترک بین آن‌ها باشد، این همبستگی عددی بین $\frac{1}{4}$ و $\frac{1}{3}$ خواهد بود. با درنظر گرفتن توزیع نرمال چند متغیره، برای مقادیر صفت مربوط به هر یک از فرزندها، این همبستگی برابر با $\frac{1}{2}$ خواهد شد (بوستین و همکاران، ۱۹۸۰). اگر در مطالعه پیوستگی در خانواده‌های بزرگ، از رگرسیون GLS استفاده شود، توان آزمون بهبود خواهد یافت. ماتریس همبستگی بین متغیرهای وابسته، به‌صورت همانی نخواهد بود و به صورت قطری بلوکی[۲۴] است که هر بلوک روی قطر را با $w(\rho)$ نشان می‌دهند، که در آن ρ همبستگی بین دو زوج فرزندی‌های تنی دارای یک فرزند مشترک است. ونگ (۲۰۰۶) نشان داد برای همبستگی ρ و خانواده‌ای به اندازهٔ s، معکوس ماتریس همبستگی را می‌توان به‌دست آورد و با استفاده از روش GLS عمل برآورد را به سرعت انجام داد.

۴.۴ تعمیم روش HE به شجره‌نامه‌ها

واضح است که اگر بتوان روش HE را به شجره‌نامه‌ها تعمیم داد، توان آن به‌طور چشم‌گیری بهبود خواهد یافت. آموس و همکاران (۱۹۸۹) روش oHE را به زوج خویشاوندانی که از یک نژاد نیستند[۲۵]، تعمیم دادند. اما این روش یک ایراد مهم دارد، و آن این است که برای استفاده از آمارهٔ آزمون موردنظر، فرض استقلال برای زوج خویشاوندان برقرار نشده است. توزیع آمارهٔ مورد استفاده‌برای نمونه‌های کوچک توسط ویلسون و الستون (۱۹۹۳) تقریب زده شد. شید و همکاران (۲۰۰۰) نیز نظریهٔ ترکیب اطلاعات حاصل از زوج فرزندی‌های تنی و زوج فرزندی‌های ناتنی را برای تحلیل پیوستگی ارائه دادند. در تحلیل پیوستگی، صفت‌های همبسته از زوج‌های خویشاوندی در یک شجره‌نامه، خوشه‌ای از مشاهدات همبسته را

[۲۴] Blocked diagonal

[۲۵] Noninbred relative pairs

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

تشکیل می‌دهند.

روش GEE، از ابتدای معرفی توسط لیانگ و زِیگر (۱۹۸۶) محبوبیت بسیاری کسب کرده است. برای داده‌های شجره‌نامه نیز که مقادیر متغیر وابسته (مربوط به زوج‌های خویشاوندی) نسبت به هم وابسته‌اند، از این روش استفاده می‌شود. البته کارایی این برآوردگر در حالتی که ساختار کوواریانس بهتر تعیین شود، بیشتر خواهد بود. همچنین روش GEE یک حالت کلی از روش‌های oHE، rHE و wHE است. در واقع با تعویض ماتریس کوواریانس در روش GEE، متناسب با هر روش، به هر یک از روش‌های مذکور رسیده خواهد شد (چن و همکاران، ۲۰۰۴).

۵.۴ روش HE دو سطحی

در تمام تعمیم‌های روش HE به شجره‌نامه‌ها، که تا این‌جا از آن‌ها یاد شد، این اشکال وجود دارد که نتوانسته‌اند بررسی تحلیل پیوستگی را به شجره‌نامه‌های کلی[۲۶] بسط دهند. به عبارتی در هیچ‌یک از آن‌ها نمی‌توان، متغیرهای تبیینی سطح فرد[۲۷] و سطح شجره‌نامه[۲۸] را لحاظ کرد. این کار با ارائه روش رگرسیون HE دو سطحی[۲۹] (tHE) توسط ونگ و الستون (۲۰۰۵) انجام شد. ونگ و الستون (۲۰۰۷) این روش را در بررسی پیوستگی صفات فشارخون و شاخص تودهٔ بدنی یا BMI به کار گرفتند و الستون (۲۰۰۹) برنامهٔ مربوط به استفاده از این روش را در نرم‌افزار S.A.G.E[۳۰] ارائه نمود. در این روش با استفاده از متغیر تبیینی سطح فرد در مدل رگرسیونی خطای مانده کاهش داده می‌شود و با در نظرگرفتن سطح شجره‌نامه نیز ناهمگنی داده‌ها کنترل می‌گردد. به عبارت دیگر اثر متقابل محیط و ژنتیک در مدل منظور می‌گردد. روش HE دو سطحی تحت چارچوب کلی رگرسیون چندسطحی و با استفاده از الگوریتم IGLS[۳۱] انجام می‌شود (ونگ و الستون، ۲۰۰۵).

---

۲۶ General pedigree

۲۷ Individual-level

۲۸ Pedigree-level

۲۹ Two level HE or tHE

۳۰ Statistical Analysis for Genetic Epidemiology

۳۱ Iterative generalized least squares

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

۵ مثال کاربردی

افراد مورد بررسی در این مطالعه از بین شرکت کنندگان در مطالعه قند و لیپید تهران انتخاب شده‌اند که شامل ۹۱ خانواده ایرانی (۴۹۳ نفر) بوده و انتخاب آنها به این صورت بوده است که حداقل یک نفر از اعضای این خانوارها مبتلا به سندرم متابولیک (طبق معیار ATP *III*) و حداقل دو نفر از اعضای آنها دچار کاهش HDL-C بودند. به منظور تحلیل پیوستگی ژنتیکی صفت دور کمر افراد مبتلا به سندرم متابولیک، ۱۲ مارکر مختلف در ۴ ناحیه کروموزومی انتخاب شدند و تکثیر این مارکرها با استفاده از تکنیک Fragment Analysis انجام گردید و تحلیل آماری آن با روش‌های رگرسیونی oHE، HE بازنگری شده، HE وزنی و HE دوسطحی انجام گردید. همچنین در این روش‌ها مدل‌های رگرسیونی با متغیرهای سن و جنسیت تعدیل شده‌اند. در این تحلیل از نرم‌افزارهای power.HE، SPSS، Excel، PowerMarker و S.A.G.E استفاده شد. نتایج حاصل در جدول ۵ آمده است. هانطور که دیده می‌شود ژن مربوط به صفت دور کمر با هیچ یک از مارکرهای مورد نظر پیوستگی معنی‌داری نداشته است.

جدول ۵: p-مقدار حاصل از روش‌های رگرسیونی HE

	روش رگرسیونی			
مارکر	oHE	HE بازنگری شده	HE وزنی	HE دوسطحی
D8S1132	۰/۲۶	۰/۲۶	۰/۱۲	۰/۱۳
D8S1779	۰/۴۵	۰/۴۵	۰/۳۱	۰/۱۷
D8S514	۰/۶۸	۰/۶۸	۰/۷۲	۰/۶۱
D8S1743	۰/۹۷	۰/۹۷	۰/۹۲	۱
D11S1998	۰/۴۸	۰/۴۸	۰/۳۹	۰/۴۴
D11S934	۰/۱۵	۰/۱۵	۰/۲۴	۰/۴۰
D11S1304	۰/۱۷	۰/۱۷	۰/۲۴	۰/۱۸
D12S1632	۰/۱۸	۰/۱۸	۰/۱۴	۰/۴۲
D12S96	۰/۸۲	۰/۸۲	۰/۶۴	۰/۲۹
D12S329	۰/۶۶	۰/۶۶	۰/۷۱	۱
D16S2624	۰/۱۴	۰/۱۴	۰/۲۱	۱
D16S3096	۰/۳۶	۰/۳۶	۰/۲۴	۰/۰۸

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

بحث و نتیجه‌گیری

امروزه یکی از مسائل مهم در علم آمار ژنتیک، بحث مکان‌یابی ژنی صفت کمی می‌باشد و یکی از روش‌های آماری که به طور عمده در حل این مساله بکار گرفته می‌شود، روش رگرسیونی هیسمن-الستون است. در این مقاله حالت اصلی این روش توضیح داده شد و در مورد سیر تاریخی آن تا سال ۲۰۰۹ بحث شد. و در نهایت با مثالی کاربردی نحوه استفاده از این روش‌ها در یافتن ژن مربوط به صفت کمی دور کمر پرداخته شد، که البته ژن مربوطه با هیچ یک از مارکرهای موردنظر پیوستگی معنی‌داری نداشته و شاید علت کافی نبودن حجم نمونه برای تحلیل‌های موردنظر باشد. با توجه به سیر بهبود این روش در سال‌های اخیر، می‌توان گفت دقت در روش HE دو سطحی از سایر روش‌های رگرسیونی HE بیشتر است و این روش در تعیین مکان ژنی صفت کمی نسبت به سایر روش‌های HE قابل اعتمادتر می‌باشد. لذا توصیه می‌شود در استفاده از این روش از متغیرهای مربوط به هر یک از سطوح فرد و خانواده استفاده شود تا بتوان در مدل رگرسیونی مورد نظر واریانس بیشتری را کنترل کرد.

مراجع

Ansley, C. F. and Kohn, R. (1983), Exact Likelihood of Vector Autoregressive-Moving Average Process with Missing or Aggregated Data.,*Biometrika* **70**, 275-278.

Allison, D. B., Fernández, J. R., Heo M. and Beasley, T. M. (2000), Testing the Robustness of the new Haseman-Elston Quantitative-Trait Loci-Cmapping Procedure, *The American Journal of Human Genetics*, **67**, 249-252.

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

Almasy, L. and Blangero, J. (1997), Multipoint Oligogenic Linkage Analysis of Quantitative Traits. *Genetic Epidemiology.* **14**, 959-964.

Amos, C. I., Elston, R. C., Wilson, A. F., Bailey-Wilson, J. E. and Rao, D. C. (1989) A more Powerful Robust Sib-pair Test of Linkage for Quantitative Traits. *Genetic Epidemiology.* **6**: 435-449.

Blackwelder, W. C. and Elston, R. C. (1982), Power and Robustness of Sib-Pair Linkage Tests and Extension to Larger Sibships. *Communications in Statistics-Theory and Methods.* **11**, 449-484.

Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. (1980), Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *American Journal of Human Genetics.* **32**, 314-331.

Chen, W. M., Broman, K. W. and Liang, K. Y. (2004), Quantitative Trait Linkage Analysis by Generalized Estimating Equations: Unification of Variance Components and Haseman-Elston Regression. *Genetic Epidemiology.*, **26**, 265-272.

Collins, A. and Morton, N. E. (1995), Nonparametric Tests for Linkage with Dependent Sib Pairs. *Hum. Hered.*, **45**, 311-318.

Drigalenko, E. (1998), How Sib Pairs Reveal Linkage. *The American Journal of Human Genetics.* **63**, 1243-1245.

Elston, R. C. (2009) Statistical Analysis for Genetic Epidemiology (S.A.G.E). 6.0.1 ed. Cleveland, Ohio: Department of Epidemiology and Biostatistics, Case Western Reserve University; April, 2009.

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

Elston, R. C, Buxbaum, S., Jacobs, K. B. and Olson, J. M. (2000), Haseman and Elston Revisited. *Genetic Epidemiology* **19**, 1-17.

Haseman, J. K. and Elston, R. C. (1972), The Investigation of Linkage Between a Quantitative Trait and a Marker Locus. *Behavior Genetics.* **2**, 3-19.

Ott, J. (1999), *Analysis of Human Genetic Linkage. 3*, Editors: Baltimore, Maryland: Johns Hopkins University Press.

Palmer, L. J., Jacobs, K. B. and Elston, R. C. (2000), Haseman and Elston Revisited: The Effects of Ascertainment and Residual Familial Correlations on Power to Detect Linkage. *Genetic Epidemiology.* **19**, 456-460

Schaid, D. J., Elston, R. C., Tran, L. and Wilson, A. F. (2000), Model-Free Sib-Pair Linkage Analysis: Combining Fullsib and Half-Sib Pairs. *Genetic Epidemiology.* **19**, 30-51.

Sham, P. C. and Purcell, S. (2001), Equivalence Between Haseman-Elston and Variance-Components Linkage Analyses for Sib Pairs. *The American Journal of Human Genetics.* **68**, 1527-1532.

Shete, S., Jacobs, K. B. and Elston, R. C. (2003), Adding Further Power to the Haseman and Elston Method for Detecting Linkage in Larger Sibships: Weighting Sums and Differences. *Hum. Hered.* **55**, 79-85.

Visscher, P. M. and Hopper, J. L. (2002), Power of Regression and Maximum Likelihood Methods to Map QTL From Sib-Pair and DZ Twin Data. *Annals of Human Genetics.* **65**, 583-601.

[DOR: 20.1001.1.17358183.1388.3.2.11.4]

Wang, T. (2006) *Extensions of Haseman-Elston Regression for Linkage Analysis.* Cleveland, Ohio: CASE WESTERN UNIVERSITY.

Wang, T. and Elston, R. C. (2005), Two-Level Haseman-Elston Regression for General Pedigree Data Analysis. *Genetic Epidemiology.* **29** 12-22.

Wang, T. and Elston, R. C. (2007), Regression-Based Multivariate Linkage Analysis with an Application to Blood Pressure and Body Mass Index. *Annals of Human Genetics.* **71**, 96-106.

Wei-Min Chen and KWBK-YL. (2004), Quantitative Trait Linkage Analysis by Generalized Estimating Equations: Unification of Variance Components and Haseman-Elston Regression. *Genetic Epidemiology.* **26**, 265-272.

Wilson, A. F. and Elston, R. C. (1993), Statistical Validity of the Haseman-Elston Sib-Pair Test in Small Samples. *Genetic Epidemiology.* **10**, 593-598.

Xu, X., Weiss, S. and Wei. L. J. (2000), A Unified Haseman-Elston Method for Testing Linkage with Quantitative Traits. *The American Journal of Human Genetics.*, **67**, 1025-1028.

Zeigler, A. and Inke, R. K. (2006), *A Statistical Approach to Genetic Epidemiology, Concepts and Applications.* Germany, Lubeck: WILEY-VCH Verlag GmbH&Co.KGaA WeinHeim.

Ziegler A. (2001), The new Haseman-Elston Method and the Weighted Pairwise Correlation Statistic are Variations on the Same Theme. *Biometrical Journal*, **43**, 697-702.

Journal of Statistical Sciences, Autumn & Winter 2009 **228**

Vol. 3, No. 2, pp 141-159.

Haseman-Elston Regression Methods in Genetic Linkage Analysis

Akbarzadeh M., Alavi Majd H., Mehrabi Y., Daneshpour M. S. and Mohammadi A.

Dept. of Biostat. and Epidemiology, Hamedan University, Hamedan, Iran.

Dept. of Epidemiology, Shahid Beheshti University, M. C., Tehran, Iran.

Obesity Research Center, Shahid Beheshti University, M. C., Tehran, Iran.

Dept. of Statistics, Tarbiat Moddares University, Tehran, Iran.

Abstract: One of the important problems that bring up in genetic fields is determining of loci of special gene in order to gene mapping and generating more effective drugs in medicine. Genetic linkage analysis is one important stage in this way. Haseman-Elston method is a quantitative statistical method that is used by biostatisticians and geneticists for genetic linkage analysis. The original Haseman-Elston method is presented in the year 1972 and ever after many investigators recommended some suggestions to make better it. In this article, we introduce the Haseman-Elston regression method and its extensions through 1972 to 2009. and finally we show performance of these methods in a practical example.

Keywords: Genetic linkage analysis, Haseman-Elston regression method, Allele Sharing IBD

Mathematics Subject Classification (2000): 62J99, 62P10