

Coveriance, Correlation and Regression :

در اکثر مطالعات ژنتیکی فهمیدن این موضوع که آیا ارتباطی بین 2 خصوصیت داده شد. در یک

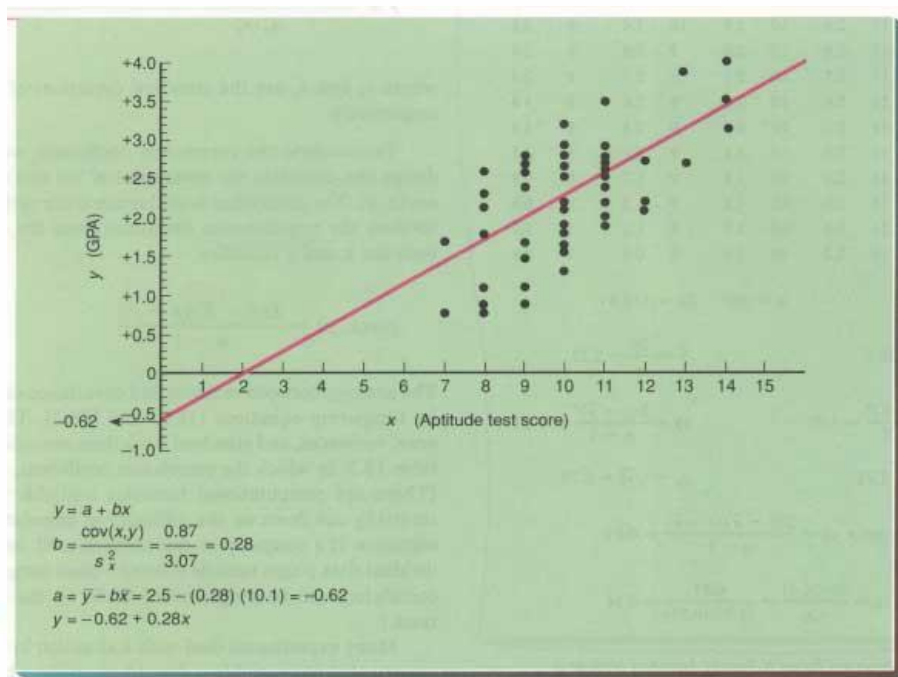
سری از افراد وجود دارد یا نه، مورد توجه است.

برای مثال، آیا ارتباط بین ارتفاع یک درخت و وزن آن وجود دارد؟ آیا ارتباطی بین استعداد تحصیلی

فرد و نمرات وی وجود دارد؟ اگر یکی از این دو خصوصیت افزایش یابد دیگری نیز افزایش خواهد یافت؟

در جدول 1 یک مثال آورده شده است و داده ها به صورت گرافیکی در شکل 1 به

صورت *scatter plot* نمایش داده شده اند.



جدول 1

TABLE 18.3 The Relationship between Two Variables, x (Aptitude Test Score) and y (Grade-Point Average)

x	y	x	y	x	y	x	y	x	y
14	4.0	11	2.9	10	2.9	10	1.4	8	2.6
14	3.4	11	2.8	10	2.8	9	2.8	8	2.4
14	3.2	11	2.7	10	2.7	9	2.7	8	2.3
13	3.7	11	2.6	10	2.6	9	2.6	8	1.8
13	2.7	11	2.5	10	2.2	9	2.4	8	1.4
12	2.7	11	2.4	10	2.1	9	2.1	8	1.1
12	2.4	11	2.2	10	1.9	9	1.7	8	0.9
12	2.2	11	2.0	10	1.8	9	1.5	8	0.8
12	2.1	11	1.9	10	1.7	9	1.1	7	1.7
11	3.5	10	3.2	10	1.6	9	0.9	7	0.8

$$\Sigma x = 505 \qquad n = 50 \qquad \Sigma y = 112.4$$

$$\bar{x} = \frac{\Sigma x}{n} = 10.1 \qquad \bar{y} = \frac{\Sigma y}{n} = 2.25$$

$$s_x^2 = \frac{\Sigma(x - \bar{x})^2}{n - 1} = 3.07 \qquad s_y^2 = \frac{\Sigma(y - \bar{y})^2}{n - 1} = 0.57$$

$$s_x = \sqrt{s_x^2} = 1.75 \qquad s_y = \sqrt{s_y^2} = 0.75$$

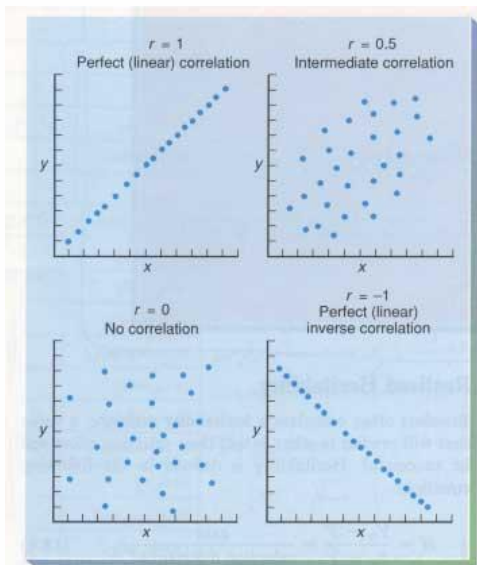
$$\text{cov}(x, y) = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{n - 1} = 0.87$$

$$r = \frac{\text{cov}(x, y)}{s_x s_y} = \frac{0.87}{(1.75)(0.75)} = 0.66$$

به نظر می رسد که بین 2 متغیر ارتباطی وجود دارد. زمانی که استعداد (محور x) افزایش می یابد می توان افزایش را در نمرات (محور y) مشاهده کرد. می توان تعیین نمود که 2 متغیر چه میزان به یکدیگر وابسته هستند و برای این کار باید ضریب همبستگی را محاسبه نموده - شاخصی که بین -1 تا +1 تغییر می کند و مقدار آن وابسته به میزان ارتباط بین 2 متغیر است.

اگر هیچ ارتباطی بین 2 متغیر موجود نباشد (2 متغیر مستقل از هم باشند) ضریب همبستگی برابر صفر خواهد بود. اگر یک همبستگی کامل بین 2 خصوصیت وجود داشته باشد به طوری که افزایش یک خصوصیت همراه با افزایش متناسب خصوصیت دیگر باشد ضریب معادل +1 خواهد بود. اگر افزایش در یک

خصوصیت همواره با کاهش متناسب در خصوصیت دیگر باشد ضریب معادل با 1- خواهد بود. (شکل 2)



فرمول ضریب همبستگی (r) :

$$4. r = \frac{\text{Covariance al Xandy}}{S_x \cdot S_y}$$

S_x معادل با انحراف معیار برای X , S_y معادل با انحراف معیار برای y است.

برای محاسبه ضریب همبستگی در درجه اول نیاز به محاسبه و تعیین کوواریانس 2 متغیر داریم که به

صورت $COV(X,Y)$ نمایش داده می شود. کوواریانس مشابه واریانس است ولی در کوواریانس انحراف از

میانگین را برای y, x را به طور همزمان محاسبه می کنیم.

$$5. Cov(x, y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{n - 1}$$

تشابه بین کوواریانس و واریانس را می توان از طریق مقایسه کردن معادلات آن ها مشاهده کرد.)

فرمول های 4 و 5)

در جدول 1 کوواریانس، واریانس و انحراف معیار محاسبه شده اند و ضریب همبستگی، r ، معادل

0/66 است.

(برخی فرمول های محاسباتی وجود دارد که محاسبه این اعداد را تسهیل می کند. اگر یک کامپیوتر

یا ماشین حساب مورد استفاده قرار گیرد تنها لازم است که داده ها را وارد کنیم - بسیاری از کامپیوترها و

ماشین حساب ها قابلیت برنامه ریزی شدن برای انجام این محاسبات را دارند)

در بسیاری از آزمایش ها با شرایطی روبرو می شویم که فرض می کنیم یک متغیر به متغیرهای دیگر

وابسته است.

برای مثال، ممکن است سوال کنیم، چه ارتباطی بین مقاومت به DDT در دروزوفیلا و افزایش تعداد

الل های مقاوم به DDT وجود دارد؟ هر چه تعداد این الل ها افزایش یابد (شکل زیر)





مقاومت به *DDT* در مگس ها الزاماً افزایش می یابد. تعداد الل های ایجاد کننده مقاومت به *DDT*

متغیر مستقل و مقاومت مگس ها یک متغیر وابسته است. در نتیجه، مقاومت یک مگس در برابر *DDT*

وابسته به تعداد الل های ایجاد کننده مقاومتی است که آن مگس دارد.

به شکل «1» باز می گردیم، می توانیم فرض کنیم که میانگین نمرات وابسته به استعداد تحصیلی

است. اگر اینگونه باشد می توان از یک سیستم آنالیز به نام آنالیز رسیدن استفاده کرد. این آنالیز به ما

اجازه می دهد که میانگین نمره (متغیر y) را بر اساس تعداد استعداد تحصیلی داده شد. (متغیر x) پیش

بینی کنیم (قابل به تذکر است که وقتی یک رابطه علت و معلولی مورد فرض قرار می گیرد آنالیزهای

همبستگی دیگر معتبر نیستند. همبستگی یا رگوسیون می تواند روی یک سری داده محاسبه شود نه هر

دوی آن ها. آنالیز رگوسیون وجود یک ارتباط علت و معلولی را مورد فرض قرار میدهد ولی آنالیز همبستگی

اینگونه عمل نمی کند)

فرمول مربوط به وجود یک ارتباط خطی بین 2 متغیر (خط رگوسیون) به صورت $y = a + bx$ است.

b به عنوان شیب خط است (تغییرات y را به تغییرات x تقسیم می کنیم) $(\frac{\Delta y}{\Delta x})$ و a عرض از مبدا خط

است (شکل 1 را ببینید) برای مشخص کردن هر خطی تنها لازم است که شیب و عرض از مبدا آن را

محاسبه کنیم.

$$6. b = \frac{Cov(x, y)}{S^2_X}$$

$$7. a = \bar{y} - b\bar{x}$$

در نتیجه اگر ارتباط علت و معلولی بین 2 متغیر وجود داشته باشد. می توانیم مقدار y را برای هر

مقدار x داده شده پیش بینی کنیم می توانیم از فرمول $y = a + bx$ استفاده کنیم یا برای حساب کردن y

برای هر x مورد نظر از رسم خط رگوسیون استفاده کنیم. حال به بررسی خود در مورد خصوصیات کمی

ژنتیکی ادامه می دهیم.

POLYGENIC INHERITANCE IN BEANS

در سال 1909، W. Johnnsen، کسی که در مورد وزن دانه در گیاه لوبیای کوتاه

phaneolus Vugaris تحقیق می کرد بیان نمودن که خصوصیات چند ژنی توسط ژن های زیادی کنترل

می شوند. جمعیت والدی از دانه های لوبیا با توزیع پیوسته ای از وزن تشکیل شد.

Johansen این گروه جمعیتی را براساس وزن آن ها دسته بندی کرده و آن ها را کشت نموده و دانه های حاصل از خود باروری آن ها را به عنوان F_1 وزن نمود. وی پی برد که والدین دارای سنگین ترین دانه ها فرزندان با سنگین ترین دانه ها را تولید کرده اند و والدینی که سبکترین دانه ها را داشتند فرزندان با سبکترین دانه های را تولید کرده اند. (جدول)

TABLE 18.4 Johansen's Findings of Relationship between Bean Weights of Parents and Their Progeny

Weight of Parent Beans	Weight of Progeny Beans (centigrams)																n	Mean \pm SE
	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90		
65-75				2	3	16	37	71	104	105	75	45	19	12	3	2	494	58.47 \pm 0.43
55-65			1	9	14	51	79	103	127	102	66	34	12	6	5		609	54.37 \pm 0.41
45-55			4	20	37	101	204	281	234	120	76	34	17	3	1		1,138	51.45 \pm 0.27
35-45	5	6	11	36	139	278	498	584	372	213	69	20	4	3			2,238	48.62 \pm 0.18
25-35		2	13	37	58	133	189	195	115	71	20	2					835	46.83 \pm 0.30
15-25			1	3	12	29	61	38	25	11							180	46.53 \pm 0.52
Totals	5	8	30	107	263	608	1,068	1,278	977	622	306	135	52	24	9	2	5,491	50.39 \pm 0.13

در این آزمایش یک همبستگی قابل تشخیص بین وزن دانه های والدین و فرزندان وجود داشت ($r = 0/34 \pm 0/01$) وی کار خود را با شروع 19 خط لوبیا (جمعیت لوبیا) ادامه داد که والدین از جاهای متفاوتی از توزیع انتخاب شده بودند سپس وی نسل های متوالی را از طریق خودباروری طی چندین سال به دست آورد. بعد از چند نسل میانگین و واریانس در هر خط بدون تغییر و ثابت شد. در نتیجه وقتی *Johansen* از هر خط فردی را به عنوان والد انتخاب می کرد، والدین سنگین تر از میانگین یا والدین سبکتر از میانگین، فرزندان حاصل از خود باروری این والد میانگین و واریانس برابر میانگین و واریانس دیده شد. در نسل قبل را نشان میدادند.

برای مثال در یک نسل گیاهان دارای سبکترین میانگین وزنی دانه (سانتی گرم 24) و گیاهان دارای

سنگین ترین میانگین وزنی دانه (47 cg) با هم آمیزش داده شدند و فرزندان با وزن متوسط دانه (cg

37) تولید کردند.

Johnnsen بوسیله استفاده از خود باروری در هر نسل میزان هموزیگوت بودن فرزندان را افزایش

داد به همین دلیل تعداد *segregfing polyqenes* را کاهش داد و به همین دلیل هر یک از خطوط برای

چند ژنی های خاصی هموزیگوت شدند که این هموزیگوت شدن در خطوط مختلف روی ژن های مختلفی رخ

میداد و در این زمان هر گونه تنوع در وزن دانه ها به دلیل عوامل محیطی بود. در نتیجه *Johnnsen* نشان

داد که ویژگی های کمی توسط چندین لوکوس مجزا کنترل می شوند.

SELECTION EXPERIMENTS :

بررسی های انتخابی به اهداف متفاوتی انجام می شوند. پرورش دهندگان حیوانات و گیاهان مطلوب

ترین محصولات خود را به عنوان والد انتخاب می کنند تا نسل آن ها را بهتر کنند. ژنتیک جمعیت

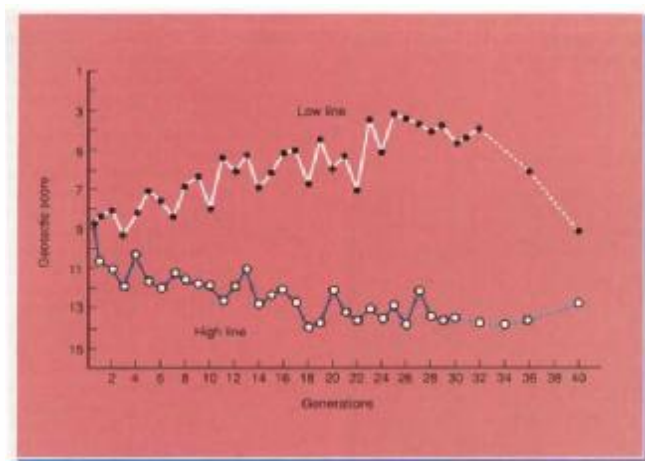
خصوصیات ویژه ای را انتخاب کرده و آن ها را مورد بررسی قرار می دهد تا ماهیت کنترل ژنتیک کمی را

درک کند.

برای مثال، دروزفیلا در یک ماز که دارای 15 انتخاب بود برای بررسی واکنش ژنوتاکتیک مورد

آزمایش قرار گرفت (شکل 3).





ماز به پهلو قرار داده شده بود به همین دلیل در هر تقاطع مگس مجبور به انتخاب بین بالا رفتن یا پایین آمدن بود. مگس هایی که دارای بیشترین امتیاز بودند به عنوان والدین برای خط بالا انتخاب شدند (دارای *geotaxies* مثبت - جهت حرکت به سمت پایین) و مگس های با پایین ترین امتیاز به عنوان والدین برای خط پایین انتخاب شدند (دارای *geotaxies* منفی - جهت حرکت به سمت بالا). در هر نسل همین انتخاب انجام شد. با گذشت زمان 2 خط به طور قابل توجهی از هم فاصله گرفته و دور شدند. این آزمایش نشان داد که یک مولفه ژنتیکی در این پاسخ نقش دارد. آزمایشات بطور موفقیت آمیزی موجب گردآمدن تعداد بیشتری از ال های "به سمت پایین" در خط بالا و ال های "به سمت بالا" در خط پایین شدند. نکات دیگری نیز از این گراف قابل دریافت است. در درجه اول، پاسخ بالا و پایین به مقدار بسیار کمی متفاوت هستند به بیان دیگر این 2 پاسخ متقارن نیستند. خط بالا با سرعت بیشتری واکنش نشان میدهد و نمودار آن سریعتر هموار و صاف می شود و این نمودار با سرعت کمتری به حالت پایه باز می گردد وقتی انتخاب به صورت آزادانه انجام می شود (انتخاب آزادانه به گونه ای است که والدین به طور تصادفی از بین افراد

انتخاب می شوند و در این انتخاب توجهی به امتیاز ژنوتایپیک نداریم)

خط پایین با سرعت کمتری پاسخ داد. و تغییرات آن نامنظم است و وقتی انتخاب به صورت آزادانه

انجام می شود بازگشت به حالت اولیه با سرعت بیشتری انجام می شود.

ماهیت این پاسخ ها (شکل 3) مشخص می کند که خط بالا به میزان نسبت به خط پائین هموزیگوت

شده است. این امر به وسیله پاسخ ناچیز به انتخاب آزادانه نشان داده می شود.

It has exhuoted a good deal of itn variability for the polyqenes

responoible for geotaxies.

به نظر می رسد که خط پایین درصد بیشتری از تنوع ژنتیکی اولیه خود را حفظ کرده باشد چرا که

میانگین امتیاز این خط با سرعت بیشتری در زمان انتخاب آزادانه افزایش می یابد و تنوع در بین افراد این

خط به حدی است که به میانگین جمعیت اولیه برسد.

واکنش و پاسخ نسبت به آزمایشات انتخابی راهی است که پرورش دهندگان حیوانات می توانند از

طریق آن پاسخ های بعدی موجود را پیش بینی کنند.

